

# **Neural mechanisms for reducing uncertainty in 3D depth perception**

Aidan Peter Murphy



A thesis submitted to the University of Birmingham for the degree of  
DOCTOR OF PHILOSOPHY

October 1, 2014

Binocular Vision Laboratory  
School of Psychology  
College of Life and Environmental Sciences  
University of Birmingham, UK

Section on Cognitive Neurophysiology and Imaging  
Laboratory of Neuropsychology  
National Institute of Mental Health  
NIH, Bethesda MD, USA

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

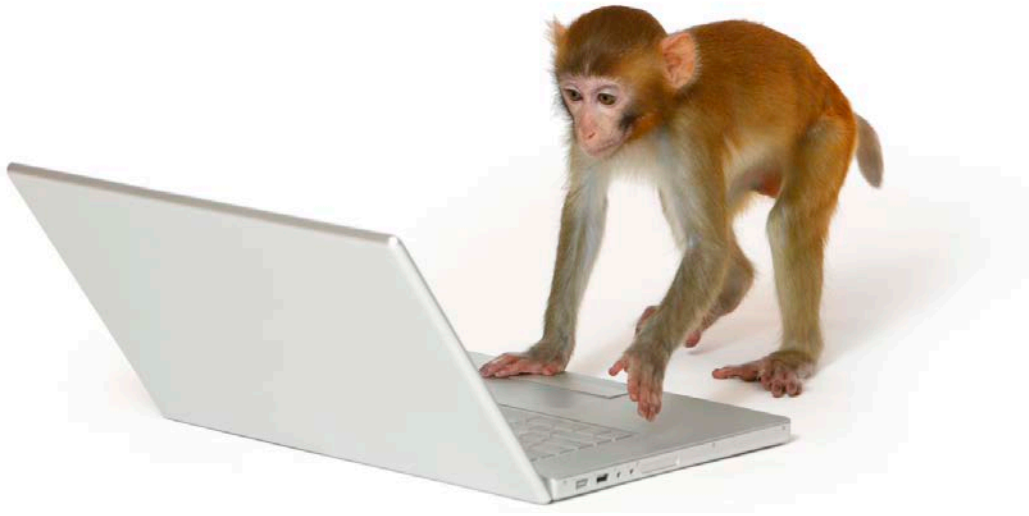
This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

# Abstract

---

In order to navigate and interact within their environment, animals must process and interpret sensory information to generate a representation or ‘percept’ of that environment. However, sensory information is invariably noisy, ambiguous, or incomplete due to the constraints of sensory apparatus, and this leads to uncertainty in perceptual interpretation. To overcome these problems, sensory systems have evolved multiple strategies for reducing perceptual uncertainty in the face of uncertain visual input, thus optimizing goal-oriented behaviours. Two available strategies have been observed even in the simplest of neural systems, and are represented in Bayesian formulations of perceptual inference: sensory integration and prior experience. In this thesis, I present a series of studies that examine these processes and the neural mechanisms underlying them in the primate visual system, by studying depth perception in human observers. In the first section (Chapters 2 & 3), we used functional brain imaging to localize cortical areas involved in integrating multiple visual depth cues, which enhance observers’ ability to judge depth. Specifically, we tested which of two possible computational methods the brain uses to combine depth cues. Based on the results we applied disruption techniques to examine whether these select brain regions are critical for depth cue integration. The second section (Chapters 4 & 5) addresses the question of how memory systems operating over different time scales interact to resolve perceptual ambiguity when the retinal signal is compatible with more than one 3D interpretation of the world. Finally, we examine the role of higher cortical regions (parietal cortex) in depth perception and the resolution of ambiguous visual input by testing patients with brain lesions (Chapter 6).



*Dedicated to Armstrong, Dexter, Layla and Maliha.*

# Acknowledgements

---

One of the many aspects of scientific research that I find so appealing is the opportunity/necessity for collaboration. It is arguably this characteristic of primate species that has led to their enhanced cognitive capacities and ultimately the development of human civilizations and the modern enterprise of science as we know it. Consequently, I am indebted to a great many people, past and present, for the opportunity to conduct this research in the first place, as well as the enormous knowledge base without which it would not have been possible to build my small contributions.

Thanks to the many lab members—past and present—of the Binocular Vision Lab and Cognitive NeuroImaging Lab at the University of Birmingham and the Section on Cognitive Neurophysiology and Imaging, Neuroimaging Facility, and other sections within the Laboratory of Neuropsychology at NIMH. Special mentions must go to Matt Patten, Dicle Dövençioğlu, Arthur Lugtigheid, Alex Murry, Vassilis Pelekanos, Matthew Dexter, Hiroshi Ban, Shuguang Kuai, Satoshi Endo, and Chunshan Deng for their outstanding support. Whether it was participating in my experiments when there were no undergrads around, sharing code, assisting scans, or just keeping me entertained, you all helped in some way. I am particularly grateful to my postdoc collaborators at the two institutions: Hiroshi Ban and Chunshan Deng, for their time and patience in teaching and assisting me, and for their dedication to making our collaborative projects successes. Thanks also to my examiners Andrew Parker and Chris Miall, as well as Zoe Kourtzi, Glyn Humphreys, Carmel Mevorach, Amanda Wood, Max DiLuca, Rob Hardwick, Elise Lesage, Alex Maier, Brian Russ and Dave McMahon for their advice and discussion. I am eternally indebted to my PhD supervisors Andrew Welchman and David Leopold, for their support, wisdom and guidance throughout the last 4 years of my academic journey. Thanks also to my student collaborators Latoya Piper and Xiaoping Li for their hard work and dedication, and the many participants who willingly gave their time to this research whilst generally making the process of data collection such a pleasure. Thanks to the NIH-Wellcome Trust PhD program for their generous support and for providing this amazing opportunity for transatlantic collaboration. Finally, thanks to my amazing family and friends from all around the world for their relentless love and support throughout this wonderful, challenging experience.

Aidan Murphy  
September 30, 2014  
Washington, DC

# Contents

---

<b>Contents.....</b>	<b>5</b>
<b>Acknowledgements by Chapter.....</b>	<b>8</b>
<b>List of Figures.....</b>	<b>10</b>
<b>Glossary.....</b>	<b>12</b>
<b>Chapter 1: Introduction.....</b>	<b>14</b>
1.1. Animals as Empiricists .....	14
1.2. Organization of the primate visual system .....	18
1.3. Perception as statistical inference.....	24
1.4. Reducing sensory uncertainty through cue integration .....	27
1.5. Resolving perceptual ambiguity through prior experience .....	32
1.6. Additional methods for reducing uncertainty .....	37
<b>Chapter 2: Integration of texture and disparity cues to surface slant in dorsal visual cortex.....</b>	<b>40</b>
2.1. Introduction .....	41
2.2. Materials and Methods .....	46
Observers.....	46
Stimuli .....	46
Vernier Task.....	50
Psychophysics.....	50
Imaging.....	53
Defining regions of interest.....	53
fMRI design .....	57
fMRI data analysis.....	57
2.3. Results .....	61
Psychophysics.....	61
fMRI.....	62
2.4. Discussion .....	70
Surface orientation components: slant and tilt.....	71
Surface orientation processing in the dorsal visual pathway .....	71
Relation between psychophysical and fMRI results .....	76
Conclusions.....	79
<b>Chapter 3: The effects of V3B/KO disruption on the integration of texture and disparity cues to surface slant .....</b>	<b>80</b>
3.1. Introduction .....	81
3.2. Methods.....	85
Observers.....	85
Stimuli and presentation.....	85
fMRI acquisition.....	87
fMRI analysis .....	88
Experimental procedure .....	90
Behavioural data analysis.....	91
3.3. Results .....	93
3.3.1. Baseline performance.....	93
3.3.2. Offline TMS effects .....	96

3.3.3. Reaction times .....	98
3.3.4. Eye movements.....	99
<b>3.4. Discussion .....</b>	<b>100</b>
3.4.1. Relation to fMRI studies .....	100
3.4.2 Experimental design limitations.....	102
3.4.3. Limitations of TMS.....	104
3.4.4. Interpretations of a genuine ‘null result’ .....	106
3.4.5. Conclusion.....	108
<b>3.5. Supplementary materials.....</b>	<b>109</b>
<b>Chapter 4: Perceptual memory drives learning of retinotopic biases for bistable stimuli.....</b>	<b>111</b>
<b>4.1. Introduction .....</b>	<b>112</b>
<b>4.2. Results .....</b>	<b>115</b>
Effects of training stimuli on perceptual bias.....	116
Effects of spontaneous alternation on perceptual bias .....	121
<b>4.3. Discussion .....</b>	<b>125</b>
Classical conditioning versus perceptual association of location-contingent biases .....	129
Retinotopic specificity of perceptual memory affects perceptual association learning	130
The nature of perceptual resolution .....	132
Conclusion.....	133
<b>4.4. Methods.....</b>	<b>135</b>
Hardware and software.....	135
Stimuli.....	135
Task.....	136
Participants.....	137
Analysis .....	138
<b>4.5. Supplementary Materials .....</b>	<b>140</b>
<b>4.6. Appendices .....</b>	<b>143</b>
Pilot experiment .....	143
<b>Chapter 5: Long-term biases for bistable stimuli are modified by recent perceptual experience.....</b>	<b>146</b>
<b>5.1. Introduction .....</b>	<b>147</b>
<b>5.2. Methods.....</b>	<b>150</b>
Participants.....	150
Hardware and software .....	150
Stimuli .....	151
Task.....	152
Analysis .....	154
<b>5.3. Results .....</b>	<b>155</b>
5.3.1. Experiment 1: Long-term effects of perceptual bias disruption.....	155
5.3.3. Eye of origin specificity.....	166
<b>5.4. Discussion .....</b>	<b>169</b>
5.4.1. Long-term bias .....	169
5.4.2. Pre-existing biases remain stable without stabilization.....	171
5.4.3. Ocular transfer .....	173
5.4.4. Conclusion.....	174
<b>5.5. Supplementary materials.....</b>	<b>176</b>
<b>Chapter 6: The role of parietal cortex in the perception of depth from motion and stereopsis.....</b>	<b>180</b>
<b>6.1. Introduction .....</b>	<b>181</b>
6.2.1. Stereopsis .....	182
6.2.2. Bistable structure-from-motion.....	184

<b>6.2. Methods and Materials .....</b>	<b>189</b>
Participants.....	189
Stimuli .....	191
Stereopsis tasks .....	192
Perceptual bistability tasks.....	194
Tasks.....	196
Imaging and analysis .....	197
Behavioural Analysis .....	197
<b>6.3. Results .....</b>	<b>200</b>
6.3.1 Stereopsis .....	200
6.3.2. Perceptual Rivalry .....	204
<b>6.4. Discussion .....</b>	<b>213</b>
6.4.1. Stereopsis .....	213
6.4.2. Motion perception and structure-from-motion .....	217
6.4.3 Perceptual alternations during bistable viewing.....	218
6.4.4. Limitations of between-subjects studies of percept duration.....	222
6.4.4. Conclusion.....	224
<b>6.5. Supplementary Materials .....</b>	<b>225</b>
<b>Summary .....</b>	<b>230</b>
<b>References.....</b>	<b>234</b>



# Acknowledgements by Chapter

## **Collaborators:**

Hiroshi Ban<sup>1,3</sup>, Glyn W. Humphreys<sup>1,4</sup>, David A. Leopold<sup>2</sup>, Andrew E. Welchman<sup>1,5</sup>

1. School of Psychology, University of Birmingham, UK
2. Section on Cognitive Neurophysiology & Imaging, Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, MD, USA
3. Center for Information and Neural Networks, National Institute of Information and Communications Technology, Osaka, 565-0871, Japan
4. School of Experimental Psychology, Oxford University, UK
5. Department of Psychology, University of Cambridge, UK

All research was supported by a Wellcome Trust-NIH PhD studentship (WT091467MA) to Aidan P. Murphy and a Wellcome Trust Senior Research Fellowship (095183/Z/10/Z) to Andrew E. Welchman. David A. Leopold was supported by the Intramural Research Program of the National Institute of Mental Health. Hiroshi Ban was supported by the Japan Society for the Promotion of Science (H22,290). Glyn W. Humphreys was supported by grants from the MRC and the Stroke Association.

## **Chapter 1:**

Text and figures by APM.

## **Chapter 2:**

**Reference:** Murphy AP, Ban H & Welchman AE (2013). Integration of texture and disparity cues to surface slant in dorsal visual cortex. *J. Neurophysiol.*, 110: 190-203. doi:10.1152/jn.01055.2012. [PubMed] [J.Neurophysiol.]

**Author contributions:** APM, HB, and AEW conception and design of research; APM, HB, and AEW performed experiments; APM, HB, and AEW analyzed data; APM, HB, and AEW interpreted results of experiments; APM, HB, and AEW prepared figures; APM and AEW drafted manuscript; APM, HB, and AEW edited and revised manuscript; APM, HB, and AEW approved final version of manuscript.

Thanks to M. Nardini for discussion on the project; D. Leopold for comments on the manuscript; M. Dexter for technical assistance; and A. Meeson for writing MVPA analysis tools.

## **Chapter 3:**

**Author contributions:** APM and AEW conceived and designed the study. APM performed experiments. APM and AEW analyzed and interpreted the data. APM wrote the chapter and prepared the figures.

Thanks to C. Mevorach and R. Hardwick for advice on stimulation protocols, C. Miall and G. Humphreys for the use of TMS equipment, and M. Dexter for assistance in calibrating displays. Thanks to M. DiLuca for advice on the psychophysics design. Data from a previous study submitted by APM for the degree of M.Res Cognitive Neuroscience and

Brain Imaging at the University of Birmingham is referenced and included in supplementary materials (Murphy and Welchman, 2009).

#### **Chapter 4:**

**Reference:** Murphy AP, Leopold DA, & Welchman AE (2014). Perceptual memory drives learning of retinotopic biases for bistable stimuli. *Front. Psychol.*, 5: 60. doi: [10.3389/fpsyg.2014.00060](https://doi.org/10.3389/fpsyg.2014.00060). [PubMed] [Frontiers]

**Author contributions:** APM and AEW conceived and designed the study. APM performed experiments. APM and AEW analyzed the data. APM, AEW, and DAL interpreted the data. APM drafted the manuscript and prepared figures. APM, DAL and AEW edited and revised the manuscript.

Thanks to Massimiliano Di Luca for helpful comments on an earlier draft, and Xiaoping Li for running participants.

#### **Chapter 5:**

**Author contributions:** APM and AEW conceived and designed the study. APM performed experiments. APM and AEW analyzed and interpreted the data. APM wrote the chapter and prepared the figures.

Thanks to Latoya Piper for testing participants.

#### **Chapter 6:**

**Author contributions:** APM, DAL, GWH and AEW conceived and designed the study. APM performed experiments. APM and AEW analyzed and interpreted the data. APM wrote the chapter and prepared the figures. We thank all of our participants for their time.

A small proportion of the data presented here was collected as part of an M.Res dissertation project conducted by APM and supervised by AEW and GWH. Specifically, data were previously collected from 2 patients (PM and PF) on the bistable and dynamic stereo tasks. These data are integrated here with new material collected during the PhD: both patients were retested, including on new tasks, and additional data were collected from another 3 patients and 2 age matched controls.

## List of Figures

Figure 1.1. Primate vision aids object manipulation and navigation	18
Figure 1.2. Geometry of binocular vision	21
Figure 1.3. Sensory integration in the Bayesian framework	29
Figure 1.4. Visual illusions of depth perception	33
Figure 1.5. Depth cues and priors for an ambiguous 3D stimulus	36
Figure 2.1. Illustration of stimulus space and predictions for discrimination performance	43
Figure 2.2. Illustration of stimulus and task design	47
Figure 2.3. Cortical flat map of regions of interest and integration index	54
Figure 2.4. Behavioural results for cue integration	60
Figure 2.5. Classifier prediction performance for discriminating slant angle across ROIs	62
Figure 2.6. Distribution of bootstrapped fMRI integration indices for all ROIs	65
Figure 2.7. Subjective and objective eye movement analysis	68
Figure 3.1. Stimuli and task design	84
Figure 3.2. Cortical surface rendering indicating V3B/KO	87
Figure 3.3. Psychometric functions for one observer	93
Figure 3.4. Baseline sensitivities to surface slant for all observers	95
Figure 3.5. Group averaged sensitivities to surface slant after TMS	96
Figure 3.6. Mean slant sensitivities arranged by stimulus condition	97
Figure S3.1. Mean reaction times across sessions	108
Figure S3.2. Vergence eye movements at stimulus onset	108
Figure S3.3. Correlation between fMRI decoding and psychophysical performance	109
Figure 4.1. Cropped screen shots of the stimuli	115
Figure 4.2. Individual probit transformed bias summed across retinal locations	116
Figure 4.3. Time course of median perceptual bias across subjects throughout the block	117
Figure 4.4. Time courses of perceptual bias for individual observers	118
Figure 4.5. Illustration of experiment 2 design	121
Figure 4.6. Perceptual bias in the trained direction for each retinal location	121
Figure 4.7. Perceptual bias at the continuous location during block 2	123
Figure S4.1. Time course of perceptual bias in the pilot experiment	139
Figure S4.2. Time course of perceptual bias over 40 minutes	140
Figure S4.3. Individual observers' perceptual time courses	141
Figure A4.1. Individual observers' perceptual time courses from the pilot experiment	144
Figure 5.1. Correlation between perceptual alternations and perceptual bias	151
Figure 5.2. Task design for experiment 2	154
Figure 5.3. Perceptual biases across the visual field for individual observers	156
Figure 5.4. Scatter plots of perceptual bias for clockwise rotation	158
Figure 5.5. Correlations between change in bias and perceptual alternation rates	159
Figure 5.6. Change in perceptual bias by final percept	160
Figure 5.7. Perceptual bias across blocks for individual observers	161
Figure 5.8. Stimulus geometry for Experiment 2	163

Figure 5.9. Individual observers' perceptual biases across the visual field	164
Figure 5.10. Ocular transfer of perceptual bias in Experiment 3	167
Figure S5.1. All individual data from Experiment 1 (n = 32)	174
Figure S5.2. Distribution of inter-stimulus interval durations	176
Figure S5.3. Analysis of perceptual alternation synchrony across the visual field	177
Figure 6.1. Parietal ROIs on sagittal slices of structural MRIs	187
Figure 6.2. Schematic illustrations of stereoscopic stimuli	191
Figure 6.3. Schematic illustrations of perceptual rivalry stimuli	193
Figure 6.4. Example of perceptual alternation in eye movement data	197
Figure 6.5. Stereo thresholds for dynamic and fine disparity stimuli.	201
Figure 6.6. Psychometric functions for the dynamic, fine, and coarse disparity tasks	202
Figure 6.7. Distribution of individual percept durations for two bistable stimuli	206
Figure 6.8. Mean percept durations for all bistable stimuli	207
Figure 6.9. Distribution of normalized percept durations for the rotating sphere task	208
Figure 6.10. Perceptual alternation during continuous and intermittent viewing	209
Figure 6.11. A coronal slice through patient MH's posterior parietal lobe	213
Figure S6.1. Percept duration distributions for individual patients on the SFM experiment	225
Figure S6.2. Axial slices of spatially normalized MR images	226
Figure S6.3. Scatter plot comparing mean percept durations	227
Figure S6.4. Perceptual stabilization in the presence of disambiguating disparity	227
Figure S6.5. Correlations between distribution fit parameters	227

# Glossary

## *Associative learning:*

The formation or strengthening of a connection between a stimulus and a response. This term is sufficiently broad to include Pavlovian or classical conditioning, auto-shaping and other reward-contingent learning paradigms. Importantly, modern definitions of associative learning emphasize the basis on knowledge (i.e. unambiguous sensory information) rather than perceptual interpretations (Rescorla, 2003). For this reason, we introduce the term ‘perceptual association’ (see below) to describe a new type of associative learning based on perceptual interpretations.

## *Bias:*

In the case of bistable stimuli or any other two alternative forced choice (2IFC) experimental paradigm, the expected probability of either response is 0.5. Over multiple trials, a deviation from the expected response probability represents a bias.

## *Cue recruitment:*

The term cue recruitment was introduced by Haijiang and colleagues (2006) to describe both an experimental paradigm and the type of learning that results from it. During cue recruitment an uninformative signal (such as a specific retinal location) is repeatedly paired with a trusted cue (such as binocular disparity), so that the former signal becomes associated with the percept (e.g. depth configuration) specified by the latter. Cue recruitment elicits long-term perceptual biases lasting from 5-minutes to 1 week, but only affects perception of ambiguous stimuli. However, attempts to train observers to recruit uninformative cues other than retinal location have generally been unsuccessful, suggesting that this type of learning relies heavily on short-term biasing processes.

## *Short-term bias:*

A range of processes capable of influencing perception and acting on the timescale of milliseconds to seconds. Examples include adaptation to a recently experienced percept, which inhibits perception of that same percept during subsequent ambiguous stimulation (i.e. a negative bias). A contrasting example is perceptual stabilization (see below), in which recent experience of a given percept facilitates subsequent perception of the same percept.

## *Long-term bias:*

A range of processes acting on the timescale of minutes to weeks. Examples include pre-existing biases that are esoteric between individuals and across the visual field but consistent over time (Carter and Cavanagh, 2007). The biases induced by the cue recruitment paradigm are another example, arising in minutes but lasting for weeks.

## *Pavlovian conditioning:*

Named after the work of Ivan Pavlov (Pavlov, 1927), in which the pairing of an unrelated conditioning stimulus (e.g. a bell) with a second unconditioned stimulus (e.g. food) that naturally elicits a behavioural response (e.g. salivation), causes the former stimulus to become associated with the behavioural response even in the absence of the unconditioned stimulus.

*Perceptual stabilization:*

The term perceptual stabilization was introduced by Leopold and colleagues (Leopold et al., 2002) to describe the phenomenon whereby intermittent presentation of a bistable stimulus elicits repeated selection of the same percept across consecutive appearances (Orbach et al., 1963). It is hypothesized to rely on an interaction of short-term processes including adaptation and short-term ‘perceptual memory’.

*Perceptual association:*

We introduce this term in Chapter 4 to provide a more accurate description of the processes underlying cue recruitment. Perceptual association differs from Pavlovian conditioning and conventional definitions of associative learning in that it does not depend on reinforcing stimuli or sensory information, but instead perceptual information.

*Quadratic summation:*

A hypothesized method of computing depth cue integration that requires independent representations of depth estimates from each cue. If the two cues are independent then they can be represented orthogonally in stimulus space, and the diagonal distance between two stimuli in this space (i.e. their discriminability) can be derived from Pythagoras’s theorem ( $a^2 + b^2 = c^2$ ). The discriminability based on two cues is thus greater than that based on either cue alone. See **Figure 2.1** for details.

*Fusion:*

A hypothesized method of computing depth cue integration that requires fused representations of depth estimates from each cue. Under a fusion computation, the distance between the means of the combined depth estimates for two stimuli remains the same regardless of the number of cues, but the variance of the combined depth estimates is reduced, thus improving discriminability. See **Figure 2.1** for details.

# Chapter 1: Introduction

---

*“In all demonstrative sciences the rules are certain and infallible; but when we apply them, our fallible and uncertain faculties are very apt to depart from them, and fall into error. ... By this means all knowledge degenerates into probability.”*

— David Hume, *A Treatise of Human Nature* (1739) (2012)

## 1.1. Animals as Empiricists

The primary function of a nervous system is to execute actions that maximize survival, and this requires information about the state of the world<sup>1</sup>. As animals (and as scientists), we gather empirical evidence about our environment by making observations (or measurements), which inform our internal representations (or models) of the world and allow us to make predictions that can be used to guide our actions. In doing so, we must solve the inverse problem: to infer information about the physical environment based on signals received through our sensory apparatus. One important example of this inference problem pertains to a fundamental feature of the physical world that we inhabit, namely the existence of three spatial dimensions. The optics of the eye project light from the 3D world onto the two-dimensional surface of the retina, and from this 2D input our brains must reconstruct the missing depth dimension. Although we perform this task constantly, it is far from trivial because sensory information is frequently ambiguous, incomplete, and prone

---

<sup>1</sup> Specifically, it is survival of an organism’s genes rather than the organism itself that is maximized. In rare cases where these two goals diverge, actions that preserve genes at the cost of the organism’s life are selected. Examples include kin-selection in eusocial insect species (Bourke, A.F.G. (2011). The validity and value of inclusive fitness theory. *Proceedings of the Royal Society B: Biological Sciences*, rspb20111465.), and the post-copulatory suicide behavior of male orb-weaving spiders (Foellmer, M.W., and Fairbairn, D.J. (2003). Spontaneous male death during copulation in an orb-weaving spider. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, S183-S185.)

to noise. This leads to uncertainty when attempting to infer the physical properties of the environment responsible for the observed sensory signals. Such uncertainty presents a fundamental challenge for animals because the successful execution of an appropriate action is dependent on accurate estimation of environmental variables in general, and depth in particular.

Animals have evolved several strategies for dealing with uncertainty and noise in sensory input. For example, multiple independent sources of sensory information can be integrated in order to reduce uncertainty in the perceptual outcome (Trommershäuser et al., 2011), and this can be achieved both within (Landy et al., 1995) and between sensory modalities (Ernst and Banks, 2002;Angelaki et al., 2009). Another strategy is to use memory systems to bias perception in ways that are optimal based on previous experiences, or environmental statistics encountered in the recent past (Leopold et al., 2002;Girshick et al., 2011;Chopin and Mamassian, 2012). At longer time scales, neural systems for sensory processing can become structured in ways that reflect their input during development (Shatz and Stryker, 1978;Stryker et al., 1978), and some circuits may even have evolved to be genetically predisposed for interpreting regularities in an organism's environment that have been present throughout its evolution (Barlow, 1961;Isbell, 2006). A further strategy is to allocate neural resources towards processing specific objects or features of behavioural relevance by directing attention, and thus increasing sensitivity (Carrasco et al., 2000;Reynolds et al., 2000). To continue the analogy with the scientific method, these various strategies are akin to performing repeated data measurements, either simultaneously using different instruments (sensory integration) or across time (memory), or alternatively to fine-tune or focus the measurement devices (attention).

Several of these strategies can be observed in even the simplest of neural systems, such as that of the nematode *Caenorhabditis elegans*, which consists of only 302 neurons



(White et al., 1986). For example, this species is capable of detecting mechanical forces, light, and a variety of volatile and water-soluble chemicals (senses analogous to touch, vision, olfaction and gustation respectively), and these multiple sources of sensory information must be integrated in order to determine an appropriate motor response (Bargmann, 2006). Information from the chemosensory organs on the head and tail are integrated to form a crude spatial map of chemical gradients in the environment (Hilliard et al., 2002). Further, the behaviour of *C. elegans* can be modified by past experience of developmental environment, sensory adaptation, and associative learning (Nuttley et al., 2002; Bono and Villu Maricq, 2005; Kindt et al., 2007). While it is unlikely that these sensory processes contribute to a subjective experience of perception as we experience in such a simple system, these examples demonstrate a remarkable conservation of fundamental principles for sensory processing over the course of animal evolution<sup>2</sup>.

If simple neural systems are capable of performing basic sensory processing functions such as integration and learning, then why did some animals evolve large, costly brains<sup>3</sup>? One short answer is that greater computational power is required to perform more complex and flexible sensory-motor couplings that allowed animals to better exploit their environment and adapt to new environments. Consider the sensory modality of vision. In a simple organism like *C. elegans*, photoreceptor cells are capable of mediating phototaxis: the one-to-one stimulus-response mapping between the detection of light and locomotion (Ward et al., 2008). Through the evolution of the eye, however, animals acquired the capability to discriminate the direction of light (Arendt and Wittbrodt, 2001; Erclik et al., 2008). In order for more complex eyes to be of evolutionary advantage, increased image

---

<sup>2</sup> The “urbilaterian ancestor” is the hypothesized common ancestor of all animals having bilateral symmetry, that lived 500-600 million years ago (Knoll, A.H., and Carroll, S.B. (1999). Early animal evolution: emerging views from comparative biology and geology. *Science* 284, 2129-2137.).

<sup>3</sup> In vertebrate species, the evolutionary trend for nervous systems to become concentrated at one end of the organism (cephalization) permits increased connectivity. However, past a certain point, larger brains do not necessarily confer enhanced computational ability, due to tradeoffs (Kaas, J.H. (2000). Why is brain size so important: design problems and solutions as neocortex gets bigger or smaller. *Brain and Mind* 1, 7-23.).

processing capabilities are necessary. For example, the approach of a predator and that of a potential mate might produce similar patterns of photoreceptor activation ('looming'), but demand very different behavioural responses. While the detection of such a looming pattern can be achieved even by ganglion cells within the mammalian retina (Münch et al., 2009), only once certain distinguishing visual features are extracted from the image can these possibilities be discriminated between and the appropriate motor response executed. This task requires the enhanced processing power that cephalization (development of a brain at one end of the central nervous system) yields, but provides far greater flexibility in the sensory-motor repertoire (i.e. an expanded range of stimulus-response mappings) through more intermediate stages of processing. The functional flexibility permitted by the structural complexity of the brain is most clearly evident in the neocortex of mammalian species (Diamond and Hall, 1969), and reaches its current pinnacle in primate species, who rely on vision as their primary sensory modality (Barton, 1998; Martin and Ross, 2005).

Two notable features of primate vision are high visual acuity and a wide binocular field of view, which are made possible due to large eyes and convergent orbits, both of which assist with the execution of precise visually-guided movements for manipulating



**Figure 1.1.** Primates have a highly developed visual system that allows them to recognize and manipulate objects as well as navigate and locate objects in their environment. This system is highly flexible, allowing primates to adapt to new environments and rapidly acquire visuo-motor skills. A) A wild Japanese macaque explores the interface of a smart phone he stole from a tourist (photograph by Marsel van Oosten, squiver.com). B) A wild rhesus macaque in India swigs cola from a can. C) Wild rhesus macaques in Thailand navigate the city streets (photograph by backpackology.org).

objects and navigating their environment (**Figure 1.1**). The evolution of depth perception from binocular disparity is not unique to primates, but in the case of this order is hypothesized to represent an adaptation for either arboreal locomotion or nocturnal predation, both of which demand accurate depth judgments for hand-eye coordination (Heesy, 2009). Despite these remarkable abilities, the primate visual system experiences the same fundamental problems when attempting to infer the depth structure of the world from visual input as many other animals. In this introductory chapter, I consider the processes used to address this problem and their neural bases, with an emphasis on the cortical processing of visual information and the perception of 3D depth structure.

## 1.2. Organization of the primate visual system

At a macroscopic scale, a fundamental principle of organization of brain areas involved in processing sensory inputs and motor outputs is topography or ‘maps’. This means that sensitivity to a given component of sensory information (e.g. visual space, auditory frequency, or body part) is continuously distributed within regions of the cortical sheet (or

within a thalamic nucleus) in a manner that preserves the organization of the input signal (e.g. retinotopy, tonotopy, somatotopy) or output effectors (e.g. muscle groups controlled by primary motor cortex and the cerebellum). In the case of vision, the organizing principle of retinotopy is inherited from the optics of the eye, which results in the projection of a two-dimensional image of the environment onto the retina. This organization propagates throughout sensory processing, and in primates, similar topographic organization extends to motor and cognitive functions related to visual perception, such as the spatial locus of attention, memory, and oculomotor planning and execution (Konen and Kastner, 2008; Silver and Kastner, 2009). In chapters 4 and 5, we examine the retinotopic organization of visual perceptual memory in human observers –a process for resolving ambiguity in sensory input that inherits organizational features of the underlying neural architecture.

The dominant conceptual framework for cortical visual processing posits two anatomically and functionally distinct pathways that both begin in primary visual cortex (Mishkin et al., 1983; Goodale and Milner, 1992). Specifically, these are a ventral pathway from V1 to inferotemporal cortex that is responsible for object processing, and a dorsal pathway from V1 to posterior parietal cortex that is responsible for action-oriented, visuo-spatial processing. Each pathway consists of functionally and cytoarchitectonically distinct visual areas that are reciprocally interconnected (Felleman and Van Essen, 1991). According to the received view, these areas form a processing hierarchy, so that visual information is processed serially as it is fed forward along each pathway, yielding progressively higher levels of abstraction that rely less on the retinotopic coordinate frame of the sensory input. For example, in the ventral pathway, which receives input primarily from central vision, this abstraction is evident in the increasing receptive field sizes of neurons in progressively later cortical areas (Kravitz et al., 2013; Orban et al., 2014). This

relates directly to the goal-oriented demands of visual perception: animals need to be able to recognize an object's identity irrespective of its position on the retina. Thus, an efficient solution to neural representation of object information is proposed to involve the extraction of visual features from the sensory input, followed by the integration of these features to achieve a 'view-invariant' or concept-level representation (Booth and Rolls, 1998; Rolls, 2000; Quiroga et al., 2005; Freiwald and Tsao, 2010).

Another coarse level feature of cortical organization that has been revealed by brain imaging studies is the specialization of areas for processing visual features such as color (Engel et al., 1997; Hadjikhani et al., 1998; Lafer-Sousa and Conway, 2013), motion (Watson et al., 1993; Morrone et al., 2000; Pitzalis et al., 2010), and shape (Kourtzi and Kanwisher, 2001; Pasupathy and Connor, 2002; Brincat and Connor, 2004), as well as categories of objects including faces (Kanwisher et al., 1997; Tsao et al., 2006; Tsao et al., 2008), bodies (Downing et al., 2001) and scenes (Epstein and Kanwisher, 1998). Several studies have noted the retention of spatial information (most notably, eccentricity) within these object selective areas (Levy et al., 2001; Hasson et al., 2002; Schwarzlose et al., 2008), and more recent evidence suggests that many of these feature and category selective regions retain varying degrees of retinotopic organization (Larsson and Heeger, 2006; Kolster et al., 2010; Janssens et al., 2014; Kolster et al., 2014). Thus, the independence between form and location processing implied by the distinction between 'what' and 'where' (ventral and dorsal) pathways is not entirely accurate.

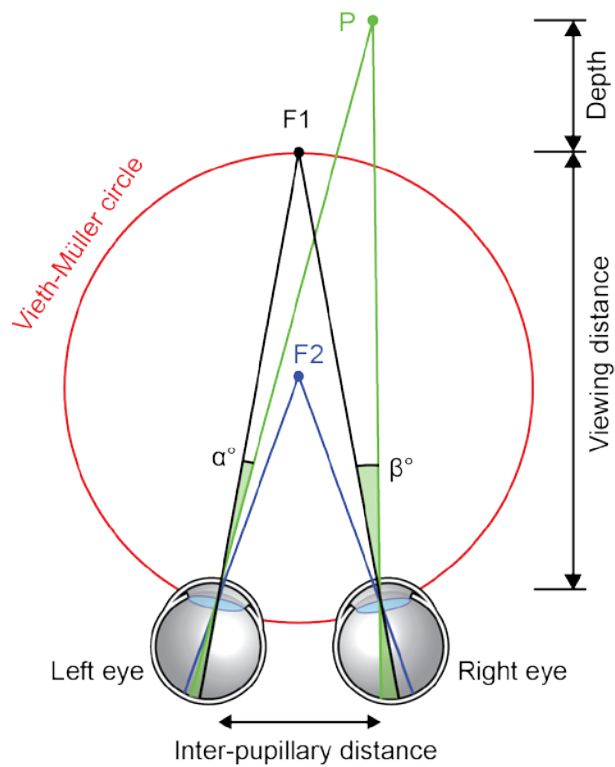


Figure 1.2. The geometry of binocular vision, viewed in the horizontal plane. When an observer fixates a point in space ( $F1$ ), this point is projected onto the fovea of each retina. Light from a point at a different depth to fixation ( $P$ ) projects to positions on the retinae that have different horizontal distances from the fovea in each eye. The angular difference ( $\alpha - \beta$ ) is the absolute disparity of point  $P$ . In contrast, if the observer were to fixate another point in space ( $F2$ ), then the same angular difference would instead reflect the *relative* disparity between points  $F1$  and  $P$ . The Vieth-Müller horopter (shown in red) passes through the point of fixation and all other points in the horizontal plane that project to geometrically corresponding points in the two retinae (i.e. all points that have zero absolute disparity).

While the 2D spatial relation between objects in the retinal image is preserved throughout much of the visual system through retinotopic organization, the recovery of depth information requires more computationally demanding processing. Pictorial depth cues in the retinal image, such as occlusion, perspective, texture and shading, result from the optics of the eye and the physical laws governing how light travels. In contrast, other monocular depth cues such as motion-parallax and structure-from-motion require the integration of information over time. Similar computations are required in order to utilize binocular depth cues, which arise from the spatial separation of the two eyes and result in subtle differences between the two retinal images (**Figure 1.2**). Unlike other visual features such as color and motion that appear to be processed in networks of dedicated cortical areas, the processing of various types of visual depth information appears to be relatively distributed throughout the visual cortex, and occurs in both dorsal and ventral visual pathways (Orban et al., 2006b; Parker, 2007). Intuitively, this makes sense, since

depth information is useful (and sometimes essential) for object recognition, as well as for estimating the 3D position of objects. However, it is clear that the type of depth information required for these two perceptual tasks is quite different.

Binocular disparity provides perhaps the most vivid perception of depth in primates, and is one of the main benefits of having two eyes for animals with forward-facing eyes. It is also a more computationally demanding cue to extract, and thus more neural resources appear to be dedicated to it compared to other depth cues. Experimentally, disparity processing is typically studied using random-dot stereogram stimuli (Julesz, 1971) that isolate binocular cues. Binocular cells in primary visual cortex show tuning to binocular disparities, and specifically to absolute disparity information (Cumming and Parker, 1999), but at this early processing stage, neural responses do not reflect a depth percept (Cumming and Parker, 1997). In contrast, a variety of cortical areas in the dorsal visual pathway show responses that correlate with observers perceptual judgments of depth from binocular disparity information (Tsao et al., 2003;Preston et al., 2008;Georgieva et al., 2009;Minini et al., 2010;Anzai et al., 2011;Krug and Parker, 2011). Notably, dorsal areas appear better at extracting disparity information from noise, which requires processing of absolute disparities rather than relative disparity information (Neri et al., 2004;Uka and DeAngelis, 2004;2006), although this is not strictly the case (Tsao et al., 2003;Krug and Parker, 2011). In contrast, temporal cortex has been shown to processes fine, relative disparity information (Janssen et al., 1999;Tanaka et al., 2001;Uka et al., 2005), and integrates disparity information globally across surfaces to support perception of 3D shapes and objects.

The two-pathway framework has proven conceptually useful and successfully accounts for a wide range of experimental findings, but it is an obvious simplification of a highly complex system, and thus faces numerous challenges (Schenk and McIntosh,

2010;de Haan and Cowey, 2011). Importantly, there is both anatomical and functional evidence for connectivity between the two pathways (Baizer et al., 1991;Rockland and Van Hoesen, 1999;Ding et al., 2000;Zhong and Rockland, 2003;Verhoef et al., 2011), and the hierarchical model based on direct cortico-cortical connections does not take into account cortico-thalamo-cortical connections which may play a major role in visual processing (Kaas and Lyon, 2007;Saalmann et al., 2012). Recent refinements of the two-pathway model have therefore elaborated on subdivisions within each pathway, as well as connections between them (Milner and Goodale, 2008;Kravitz et al., 2011;Kravitz et al., 2013).

At a finer spatial scale, the neocortex has a canonical neural architecture, organized in layers and columns with consistent patterns of cortico-cortical and cortico-thalamic connectivity. This is particularly true of sensory processing areas, such as barrel cortex of rodents (Petersen, 2007;Feldmeyer et al., 2013), and primary visual and auditory cortices of cats and primates (Hubel and Wiesel, 1962;Abeles and Goldstein Jr, 1970;Hubel and Wiesel, 1977). Evidence that attests to the canonical nature of neural architecture within sensory cortices comes from experiments in which the retinal inputs of newborn ferrets were redirected from the LGN to the medial geniculate nucleus (MGN), which is the thalamic nucleus responsible for relaying auditory information to primary auditory cortex (Roe et al., 1990;Sharma et al., 2000). Remarkably, the area of auditory cortex normally containing a tonotopic map (a 1-dimensional representation of auditory frequency) develops a 2D retinotopic representation in these animals, and neurons in this area show orientation tuning, similar to that normally observed in primary visual cortex. This finding has inspired some to suggest that the cortex in general may be structured as a knowledge-seeking, unsupervised learning algorithm (Beatty, 2001). Such a hypothesis, along with the experimental findings and a wealth of other experimental evidence, also supports the



notion that sensory input (and thus afferent activity) during development strongly determines neural connectivity (Hubel and Wiesel, 1970a; Shatz and Stryker, 1978; Stryker et al., 1978; Le Vay et al., 1980). In the case of normal development, the visual system should therefore become optimized for processing based on natural environmental statistics. The concept of efficient coding extends this idea to suggest that the role of early sensory processing stages is to produce an efficient representation of the sensory input based on the statistical properties of the world (Simoncelli and Olshausen, 2001; Simoncelli, 2003).

### 1.3. Perception as statistical inference

The notion that brains represent the structure of the world in a probabilistic manner was proposed by Herman von Helmholtz as ‘unconscious inference’ (Helmholtz, 1925), and has more recently received substantial support from studies of human behaviour (Knill and Richards, 1996; Kersten et al., 2004; Knill and Pouget, 2004; Ma et al., 2006; Yuille and Kersten, 2006; Maloney and Mamassian, 2009). In particular, Bayesian probability has also been used to successfully model the neural activity of both human and non-human observers in perception and decision making (Yang and Shadlen, 2007; Gu et al., 2008) and motor control (Todorov and Jordan, 2002; Körding and Wolpert, 2004; Scott, 2004). Bayesian decision theory postulates that inferences are made using conditional probabilities (Bayes and Price, 1763). Bayes’ theorem is formally expressed as:

$$P(X|O) = \frac{p(X) * p(O|X)}{p(O)}$$

(Equation 1.1)

where  $X$  is some state of the world,  $O$  is the observed sensory information, and  $P$  is probability. The posterior probability distribution  $P(X|O)$  is equal to the product of the prior probability distribution  $p(X)$  and the likelihood function,  $p(O|X)$ , divided by a normalization constant  $p(O)$ . According to a Bayesian model of perceptual inference, there are therefore two primary sources of information that can be used to calculate the probable state of the world given the sensory information, which correspond to the prior and likelihood functions. This thesis examines how each of these theoretical components contributes to perceptual judgments of depth and investigates their neural underpinnings.

One recent functional brain imaging study examined the neural representations of uncertainty for prior and likelihood information during a perceptual decision task (Vilares et al., 2012). Human observers were asked to estimate the horizontal position of an invisible target disk based on the distribution of five visible cue disks. From trial to trial, likelihood uncertainty was manipulated by varying the distribution of cue locations, while prior uncertainty was manipulated independently by varying the location of the actual target (which was revealed at the end of each trial) over the course of the experiment. The behavioural results indicated that, as expected, observers used both types of information to guide their responses in a manner that was dependent on the uncertainty of each source of information. The fMRI results suggested that while early sensorimotor areas (i.e. primary and dorsal visual cortex) were more strongly activated by higher likelihood uncertainty, uncertainty in prior information was associated with limbic and paralimbic decision-related areas (specifically the putamen, amygdala, insula and orbitofrontal cortex). The areas that were found to represent prior uncertainty are known to receive input from sensory areas (Cavada et al., 2000; Amaral et al., 2003; Freese and Amaral, 2005) as well as projecting to motor areas, and have previously been implicated more generally in decision-making

under uncertainty for non-perceptual tasks (Hsu et al., 2005), especially when reward is involved, as was the case in this study.

The concept of statistical inference has important implications for the neural architecture of sensory processing and perception. Specifically, neural structures even at the microscopic scale face the same fundamental problems experienced by behaving observers (Lochmann and Deneve, 2011). Unfortunately, evidence of probabilistic neural coding of sensory information remains sparse at present (Averbeck et al., 2006; Ma et al., 2006). One important aspect of probabilistic coding of perceptual decisions in the brain is that it must occur at the level of neural populations, since individual neurons are generally not very informative about sensory variables (Pouget et al., 2002; Beck et al., 2008; Ma and Pouget, 2008; Angelaki et al., 2009). In fact, the Poisson-like noise observed at the single cell levels is proposed to facilitate the reduction of Bayesian inference operations to simple linear combinations of activity of neural populations (Ma et al., 2006), such as those representing depth configurations from different visual cues (Morgan et al., 2008).

Further evidence that the brain approximates a representation of the posterior distribution comes from the phenomenon of perceptual multistability (Schrater and Sundareswara, 2006; Sundareswara and Schrater, 2008; Gershman et al., 2012). While many attempts to model the dynamics of stochastic perceptual alternation observed during ambiguous visual input have emphasized the role of neural noise (Brascamp et al., 2006; Noest et al., 2007; Deco et al., 2009), others have postulated that such patterns can emerge naturally from rational algorithms for probabilistic inference (Gigante et al., 2009; Gershman et al., 2012). Instead of considering random spike timings primarily as a source of neural noise (e.g. with respect to sensory representation), the latter approach suggests that, to some extent, these fluctuations reflect the visual system's attempt to approximate the posterior distribution through a dynamic sample-generating process

known as Markov chain Monte Carlo. However, many competing models are able to simulate the dynamics of bistable perception to varying degrees, and other accounts of neural noise have been proposed. For example, neural variability may reflect sub-optimal inference (Beck et al., 2012), or fluctuations in excitability (Goris et al., 2014). Interestingly, the onset of a visual stimulus reduces neural response variability even in the absence of changes in mean firing rate (Churchland et al., 2010; Klink et al., 2012), which supports the notion that neurons encode the posterior probability distribution for a variable of interest at all times (Beck et al., 2008).

Finally, the concept of the prior distribution is closely related to the principle of predictive coding. This suggests that far from being a purely passive, ‘bottom-up’ driven process, visual processing and perceptual inference involve active prediction of sensory signals based on ‘top-down’ information such as prior experience, and contextual information. Incoming sensory evidence is evaluated against predictions, and since sensory input that deviates from the expected norm is most informative, prediction error should be signaled via feedback connections to update the prediction (Rao and Ballard, 1999; Huang and Rao, 2011). This principle is manifested at the level of the cortical microcircuit (Bastos et al., 2012) but also extends to the recursive hierarchical structure of cortical areas. The effects of predictive coding can thus be observed as reductions in activity in primary visual cortex in response to stimulus predictability (Murray et al., 2002; Alink et al., 2010; den Ouden et al., 2010; Kok et al., 2012).

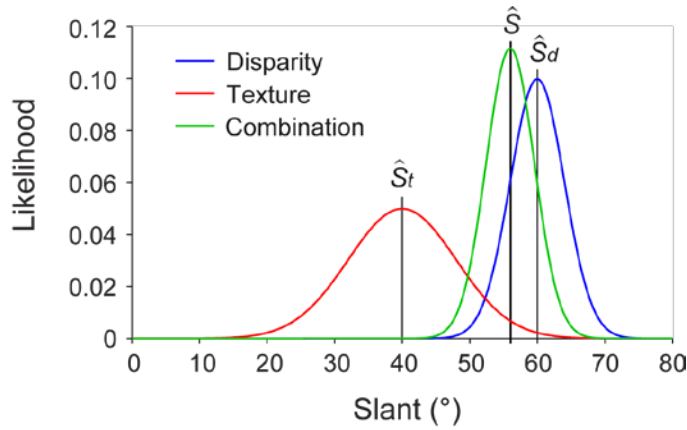
## 1.4. Reducing sensory uncertainty through cue integration

When two or more independent sources of information about the state of the world are available, how does the brain process these signals? This problem clearly applies to multisensory signals (e.g. auditory, haptic, vestibular, or visual), but it can also occur

within a single modality, such as in the case of visual cues to depth. For example, due to the optics of the eye, light reflected from an object that is closer to us is incident on a larger area of the retina than if that object were further away (i.e. nearer objects appear larger in retinal size). At the same time, due to the spatial separation of the eyes, each element of the visual scene projects onto spatially distinct areas of the two retinae. These two visual features (perspective and binocular disparity) are just two of the many types of information in the retinal images that provide cues about the 3D structure of the environment. Behavioural studies have consistently demonstrated that adult human observers not only perceive greater depth when multiple visual cues are available (Bülthoff and Mallot, 1988; Johnston et al., 1993; Johnston et al., 1994), but are objectively better at making perceptual judgments of depth (Ernst and Banks, 2002; Knill and Saunders, 2003; Hillis et al., 2004; Knill, 2007; Lovell et al., 2012). Intriguingly, the advantage gained from the availability of multiple visual depth cues appears to emerge around the age of 12 in humans (Nardini et al., 2010b), and multisensory visual haptic integration begins a few years earlier (Gori et al., 2008). This suggests that the developing visual system may still be learning the relationships between different sources of visual information, permitting recalibration as the body grows (for example, the inter-pupillary distance, which determines the relationship between binocular disparities and depth).

One way in which multiple depth cues might be utilized is for the visual system to compute depth estimates from each cue independently (i.e. modularly) before combining the two estimates linearly to obtain a unified depth percept (Marr, 1982). This model is known as weak fusion (Clark and Yuille, 1990), and stands in contrast to strong fusion models that assume non-linear interactions between depth cues that are not processed modularly (Nakayama and Shimojo, 1992). If the strong observer model were correct, cue interactions are completely unconstrained, and thus untestable. On the other hand, the

weak fusion model faces the problem that different depth cues produce qualitatively different depth information, some of which can be converted to absolute depth units given knowledge of scaling factors (e.g. binocular disparity, given viewing distance), while others produce purely relative depth information (e.g. texture). It is unclear how such dissimilar modular depth outputs could be integrated without first promoting the individual depth maps to common units. To overcome this problem, a modified weak fusion model was proposed by Landy and colleagues (Landy et al., 1995), which suggests that individual depth cue modules interact by sharing depth information, and that each module produces not only an independent depth map but also a corresponding reliability map that is used to weight the depth estimates during fusion.



**Figure 1.3.** A Bayesian perceptual system represents the world in terms of conditional probability distributions, given the sensory data. In this example, estimates about the state of the world (the slant angle of an object) can be made from two independent sources of visual information (texture and disparity). By integrating both sources of information, variance (i.e. uncertainty) in the perceptual estimate is reduced compared to the estimates made from either source alone.

Consider an example in which an observer is required to make a perceptual judgment of the slant angle of a planar surface, for which only texture and disparity cues are available. The optimal strategy when both cues are available is to make an estimate based on a weighted average of the cues:

$$\hat{S} = w_t \hat{S}_t + w_d \hat{S}_d$$

(Equation 1.2)

where each cue is given a weight ( $w$ ) that is proportional to its reliability ( $\sigma$ ):

$$w_t = \frac{1/\sigma_t^2}{1/\sigma_t^2 + 1/\sigma_d^2}, \quad w_d = \frac{1/\sigma_d^2}{1/\sigma_d^2 + 1/\sigma_t^2} \quad (\text{Equation 1.3})$$

The reliability of each cue is dependent on a number of factors that may be specific to the viewing configuration or variables specific to the individual observer. For example, the reliability of perspective information for judging surface slant is dependent on the uniformity of the texture, and the estimation of surface slant from texture relies on assumptions that textures are isotropic and homogenous (Rosenholtz and Malik, 1997). Further, the reliability of texture information for discriminating surface slant is higher at slant angles closer to horizontal than at angles closer to vertical, since equivalent changes in angle produce larger deformations of the 2D image (Knill, 1998a). Similarly, the reliability of disparity information is dependent on viewing distance (Johnston, 1991; Johnston et al., 1993). Thus cue reliability can vary across a scene, and this could be represented in a 2D reliability map corresponding to the depth map of a scene (Landy et al., 1995).

Irrespective of the specific individual cue reliabilities, the final perceptual estimate of slant ( $\hat{s}$ ) based on a weighted average of the two cues will have less variance ( $\sigma$ ) than either cue alone (**Figure 1.3**):

$$\sigma = \frac{\sigma_t^2 \sigma_d^2}{\sigma_t^2 + \sigma_d^2} \quad (\text{Equation 1.4})$$

The ‘optimal observer’ suggested by the modified weak fusion model employs simple weighted averages to accurately model the behavior of human and animal observers during perceptual judgments of depth from multiple visual cues, as well as on a variety of multisensory tasks. However, little is known about the applicability of such models to

neural populations. Existing evidence suggest that responses to multiple individual depth cues occur in spatially overlapping areas of both ventral cortex, such as LO (Moore and Engel, 2001;Kourtzi et al., 2003;Welchman et al., 2005;Georgieva et al., 2008), as well as dorsal and parietal cortex (Taira et al., 2001;Tsutsui et al., 2002;Vanduffel et al., 2002;Welchman et al., 2005;Georgieva et al., 2008;Durand et al., 2009). Further, neurophysiological studies in the macaque suggest that neurons in cortical area MT jointly encode motion and disparity information (Krug et al., 2013;Nadler et al., 2013), while neighboring area MSTd encodes visual-vestibular cue integration that supports depth perception from motion-parallax (Morgan et al., 2008). Importantly, signals in these areas have been shown to weight sensory information to reflect cue reliability, as predicted by weak fusion.

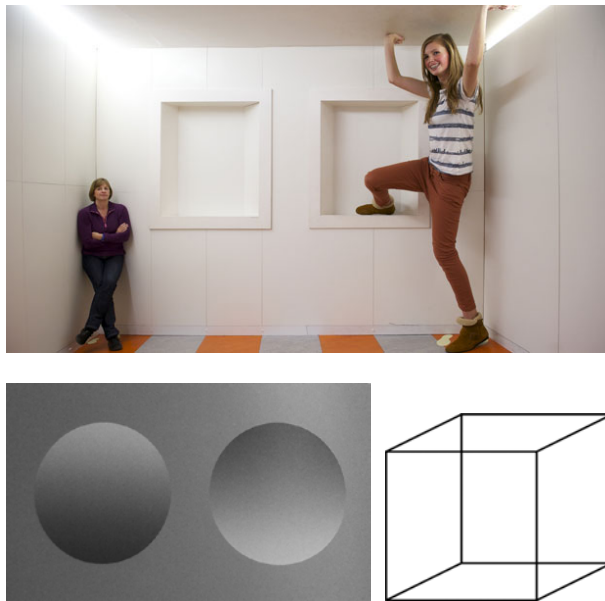
In **Chapter 2**, we use functional brain imaging to answer several questions about the integration of multiple depth cues (texture and disparity). First, we test two competing hypotheses regarding the neural computation of cue integration: independence versus cue fusion. The former predicts that co-located depth representations of two (or more) cues provide independent depth estimates, the combination of which is the quadratic sum of the individual estimates. In contrast, the fusion hypothesis predicts that component cues are fused into a single depth estimate via multiplication of the probability distributions associated with each cue, thus increasing sensitivity beyond that predicted by quadratic summation. To test these hypotheses, we compared the sensitivity of a machine learning algorithm for decoding the depth structure of visual stimuli from patterns of observers' fMRI activity. Specifically, we asked whether sensitivity to congruent cue combinations exceeded the quadratic summation of sensitivities to individual cues, as predicted by the fusion hypothesis, and if so, in which parts of the visual system? The answers to these questions were used to guide a follow-up experiment presented in **Chapter 3**, which was



designed to go beyond the correlative methods used in the previous chapter through the use of perturbation. Using transcranial magnetic stimulation to temporarily suppress cortical activity in target areas guided by the fMRI results, we tested the effects on healthy observers' ability to perform perceptual judgments of 3D stimuli containing individual or multiple cues.

## 1.5. Resolving perceptual ambiguity through prior experience

The projection of the 3D environment onto the 2D surface of the retina results in a specific pattern of retinal stimulation. However, there is often more than one possible state of the world that could have given rise to the observed pattern of retinal stimulation. This fact has long been exploited by artists, who create the impression of depth within visual scenes that are painted on 2D surfaces (Mamassian, 2008). The visual signal is therefore under-constrained or 'ambiguous', and this complicates the process of perceptual inference. One strategy that the brain uses to cope with sensory uncertainty is to constrain interpretation using prior knowledge gained from past experience. In Bayesian statistics, such prior information is represented as a prior probability distribution, which is combined with current sensory information to determine perceptual outcome (the 'posterior distribution'). Previous experience can inform explicit beliefs or 'knowledge' about the world, as well as implicit expectations or 'priors'. For example, when we enter a room we typically obtain sufficient sensory evidence to accurately perceive its geometry. Having experienced the geometry of many rooms, we accumulate statistical evidence that rooms with four walls tend to be cuboid in shape. In contrast, when sensory information is limited, perception can be strongly influenced by prior experience. In the Ames room illusion (**Figure 1.4A**), observers view the unusual case of a trapezoidal room through a carefully positioned peephole, depriving them of sufficient visual information (binocular disparity) to reject



**Figure 1.4.** Visual illusions demonstrate the influence of previous experience on visual perception of depth. **A)** The Ames room illusion. When sensory input is ambiguous (in this case, because the scene is viewed monocularly), observers rely heavily on their prior assumptions about the geometry of building interiors. **B)** Shape-from-shading stimuli create the impression of depth that is strongly dependent on prior assumptions about the direction of illumination. **C)** The Necker cube is an example of a bistable stimulus: the orthographic projection of a wire-frame cube removes perspective and occlusion pictorial cues that would normally be informative of which face of the cube is nearest to the observer. Consequently, the surface perceived to be the front of the cube spontaneously alternates during continuous viewing.

their expectation that the room is a cuboid. This misperception of 3D space results in scaling errors when estimating object sizes, and makes for a compelling demonstration of the role that expectations play in shaping perception.

Another example of ambiguity that arises in the absence of additional disambiguating depth information is the perception of shape-from-shading for curved surfaces. In **Figure 1.4B**, observers perceive a convex bump on the left and a concave dimple on the right. This interpretation of the 3D structure of the scene is based on a strong prior expectation: that scenes are normally illuminated from above, and slightly from the left (Mamassian and Goutcher, 2001). If the example figure were actually lit from below, then our interpretation of the 3D structure would be incorrect, and convex and concave shapes would be reversed (Ramachandran, 1988). The experience of light-from-above is one that we not only tend to experience everyday, but one that all seeing animals throughout evolution have tended to experience. Evidence suggests that illumination priors

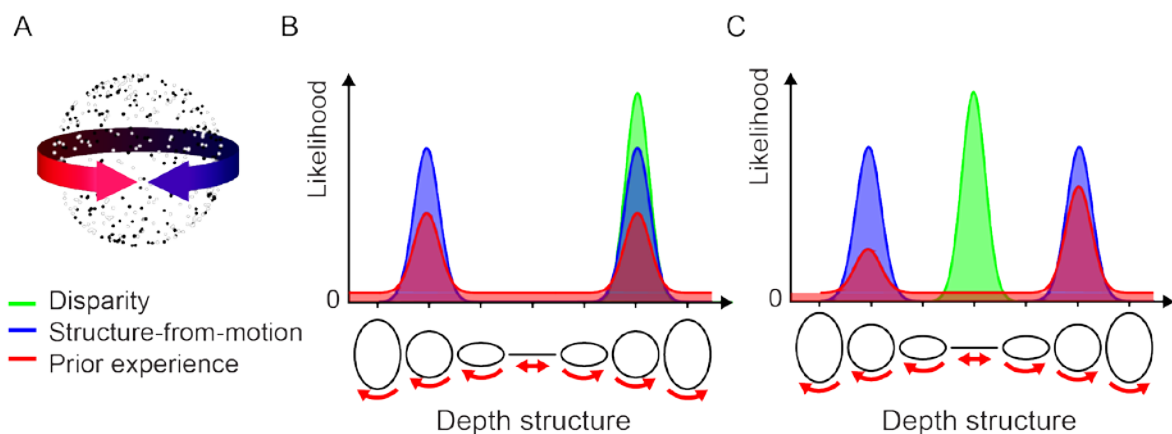
affect neural activity at the early stages of visual processing (Mamassian et al., 2003;Gerardin et al., 2010). It is therefore all the more surprising that such a strong, stable prior can in fact be modified by recent experience (Adams et al., 2004). This is a testament to the plasticity of the brain, which allows animals to adapt to changing environments.

Ambiguous or ‘bistable’ stimuli, such as the Necker Cube (Necker, 1832) shown in **Figure 1.4C**, provide a means of testing the effects of past experience on perception experimentally (Leopold and Logothetis, 1999). Bistable stimuli typically have two possible interpretations, and when viewed continuously, perception alternates spontaneously between the two. In contrast, when bistable stimuli are presented intermittently (for example, with 1 second blank intervals between each 1 second presentation), observers tend to perceive the same interpretation on consecutive presentations, and experience alternations between percepts far less frequently (Orbach et al., 1963;Leopold et al., 2002). This short-term effect of prior perception is hypothesized to reflect an interaction of adaptation and a ‘perceptual memory’ trace within neural populations that represent the two competing percepts (Nawrot and Blake, 1989;Brascamp et al., 2009;Pastukhov and Braun, 2011). For example, in the case of bistable structure-from-motion stimuli such as the rotating sphere (**Figure 1.5A**), the activity of neurons in area MT, which are selective for both motion direction and depth from disparity, have been shown to correlate with macaques’ perceptual reports (Bradley et al., 1998;Dodd et al., 2001). Similar results were obtained using pattern classification of fMRI response in human hMT+ (Brouwer and van Ee, 2007). Intermittent presentation of such stimuli elicits adaptation effects in MT neurons, but also reduces spike rate variability proportionally to the interruption duration (Klink et al., 2012). This ‘perceptual stabilization’ phenomenon may reflect a simple prior that favours constancy of perceptual interpretation across short durations of disappearance, such as those experienced in natural vision when an object

may be briefly occluded by nearer objects in the environment due to either observer or object motion.

Despite the high choice probabilities observed in MT, some evidence suggests that the depth percept and associated perceptual decision emerges later in the processing stream, such as parietal areas that read out population level activity from visual areas (Beck et al., 2008). Indeed, it has been proposed that the particularly high magnitude of choice probabilities observed in MT for this task might reflect feedback from areas that have rendered a perceptual decision rather than the feedforward of noisy sensory evidence (Gold and Shadlen, 2007). For example, perceptual decisions involving binary variables such as opposite motion directions have been shown to be encoded in several ways in lateral intraparietal area (LIP) of macaque monkeys (Platt and Glimcher, 1999; Shadlen and Newsome, 2001). When monkeys are trained to make a saccade to one of two targets, the response of a subset of LIP neurons is proportional to the probability that the saccade will be to the target in their receptive field (Platt and Glimcher, 1999). In contrast, some LIP neurons integrate information over time in a manner that is proposed to reflect computation of a likelihood ratio (Gold and Shadlen, 2001; 2007).

Other types of contextual information have also been shown to bias perception of ambiguous structure-from-motion stimuli. For example, when a pair of bistable spheres rotating about parallel axes are positioned such that their surfaces appear to be in contact, perception becomes biased towards counter-rotation, consistent with the mechanics of friction (Gilroy and Blake, 2004). In contrast, when two bistable spheres rotate around a common axis, they tend to be perceived as rotating in the same direction (Gillam, 1972; Eby et al., 1989; Grossmann and Dobbins, 2003). Similar effects have been demonstrated for other types of bistable stimuli, where completion of temporal sequences bias perception (Maloney et al., 2005; Denison et al., 2011). These phenomena demonstrate



**Figure 1.5.** **A)** A rotating sphere stimulus commonly used in studies of perceptual bistability. When the only depth cue provided is structure-from-motion (SFM), observers experience a vivid 3D percept of a rotating sphere. However, the depth order of the front and back surfaces is ambiguous, and thus the direction of rotation is bistable. When binocular disparities consistent with the sphere's 3D structure are added to the surface dots, the direction of rotation becomes unambiguous. **B-C)** Schematic illustration of probability distributions for stimulus depth structure (illustrated on the x-axis as a view from above the axis of rotation) from different sources of information. **B)** When disparity and SFM cues are available, the depth structure is unambiguous. **C)** When disparity information is not added to the dots, the disparity cue signals that the object is 2D (in the plane of the screen). Despite this, the SFM cue provides a compelling impression of depth that is compatible with either a leftward or rightward rotating sphere. The perceptual system may therefore recruit information from prior experience (e.g. the scenario described in B) in order to disambiguate the direction of sphere rotation. Note that the peaks in the prior distribution at either spherical interpretation reflect the fact that only these two depth structures are consistent with the rotation of a rigid object. In contrast, for the stimulus to be consistent with any other depth structure would entail a deformable sheet of dots flowing over the surface of a static object – a scenario that observers are less likely to have prior experience of.

the integration of contextual information across both time and space in order to resolve perceptual ambiguities in the visual input in a manner that is consistent with prior experience.

At longer-time scales, perceptual biases for ambiguous stimuli can be instantiated through the induction of explicit beliefs (Sterzer et al., 2008;Schmack et al., 2013), or implicitly through training (Haijiang et al., 2006;Harrison and Backus, 2010a;b;van Dam and Ernst, 2010;Harrison et al., 2011). Such training uses a combination of unambiguous priming stimuli (**Figure 1.5B**) and ambiguous versions of the same stimuli (**Figure 1.5C**) in order to form associations between a given percept and a previously uninformative cue, such as retinal location. Interestingly, short-term biasing factors such as priming and perceptual memory also appear to be specific to retinal location (Chen and He, 2004;Knapen et al., 2009a). In **Chapter 4** we therefore examined how short-term and long-term perceptual experience interact when the visual system must select between competing interpretations of ambiguous sensory information. In **Chapter 5** we extend this investigation to longer time scales, to reveal that the perceptual statistics of the most recent past are most predictive of future perception.

## 1.6. Additional methods for reducing uncertainty

The strategies for dealing with the problem of perceptual inference discussed thus far are fundamental: in some basic format, they are implemented in all living neural systems, and they are represented as separate components in the Bayesian formula. Using the analogy of data collection, I previously suggested that these strategies are comparable to making measurements with multiple different devices simultaneously (sensory integration) or making repeated measurements over time (prior experience). However, there are other obvious ways in which one might reduce uncertainty when making measurements, starting

with pointing the measurement device at the object of interest. This is particularly important for vision, since the distribution of photoreceptors in the mammalian retina results in a pattern of visual acuity that is highest at the fovea and falls off in the peripheral visual field. A simple solution to sensory uncertainty concerning objects in the periphery is therefore to move the eyes so that objects of interest are projected onto the fovea. Oculomotor control is of course more complicated than this, in particular due to the 3D nature of the world. Shifting gaze between two objects at different distances from the observer therefore requires coordination of conjugate and vergence eye movements as well as accommodation. Accordingly, sensitivity to visual depth information has been found in various areas involved in oculomotor control (Ferraina et al., 2000; Fukushima et al., 2002). However, the relationship between sensory input and motor output is not unidirectional. Translational motion of the head provides a strong visual cue to depth in the form of motion parallax (Rogers and Graham, 1979; Kral, 2003; Wexler and Van Boxtel, 2005), and the vestibular signals provided by head motion provide important sensory feedback for this purpose.

Another function that is closely related to the oculomotor system is the cognitive capacity of spatial attention (Corbetta et al., 1998). Endogenous attention can be allocated based on prior experience, and enhances the sensitivity of neural populations in visual areas in a manner that may be specific to features (Corbetta et al., 1990; Treue and Trujillo, 1999; Maunsell and Treue, 2006) or spatial location (Somers et al., 1999; O'Connor et al., 2002), even prior to the onset of a visual stimulus (Kastner et al., 1999; Doherty et al., 2005). By attending to an object or a feature, sensory uncertainty is therefore diminished compared to when the same stimulus is not the focus of attention. Attention is known to influence perception of ambiguous 3D stimuli (Meng and Tong, 2004) in addition to various other ambiguous stimuli (Mitchell et al., 2004; Chong et al., 2005; Kanai et al.,

2006;Paffen et al., 2006;Zhang et al., 2011). Empirical effects of attention on neural populations in visual areas can also be accounted for using Bayesian models (Rao, 2005;Chikkerur et al., 2010). However, despite the important role of attention on the visual system, the precise nature of the relationship between conscious perceptual awareness and attention remains contentious (Van Boxtel et al., 2010;Cohen and Dennett, 2011;Cohen et al., 2012;Koch and Tsuchiya, 2012). In **Chapter 6**, we examine the perception of bistable 3D stimuli in patients with lesions to parietal cortex – an area of the brain that is strongly implicated in the process of spatial attention (Silver et al., 2005;Buschman and Miller, 2007) and perceptual decisions in general.



## Chapter 2: Integration of texture and disparity cues to surface slant in dorsal visual cortex.

---

Reliable estimation of 3D surface orientation is critical for recognizing and interacting with complex 3D objects in our environment. Human observers maximize the reliability of their estimates of surface slant by integrating multiple depth cues. Texture and binocular disparity are two such cues, but they are qualitatively very different. Existing evidence suggests representations of surface tilt from each of these cues coincide at the single neuron level in higher cortical areas. However, the cortical circuits responsible for 1) integration of such qualitatively distinct cues, and 2) encoding the slant component of surface orientation, have not been assessed. We tested for cortical responses related to slanted plane stimuli that were defined independently by texture, disparity, and combinations of these two cues. We analysed the discriminability of functional MRI responses to two slant angles using multivariate pattern classification. Responses in visual area V3B/KO to stimuli containing congruent cues were more discriminable than those elicited by single cues, in line with predictions based on the fusion of slant estimates from component cues. This improvement was specific to congruent combinations of cues: incongruent cues yielded lower decoding accuracies, which suggests the robust use of individual cues in cases of large cue conflicts. These data suggest that area V3B/KO is intricately involved in the integration of qualitatively dissimilar depth cues.

## 2.1. Introduction

---

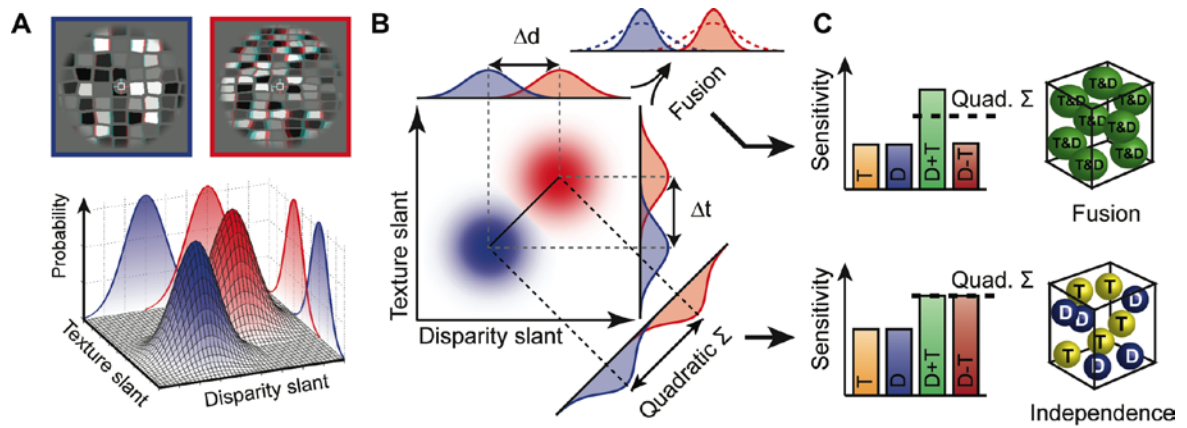
Humans experience the solid objects that populate their surroundings from multiple viewpoints. To recognise and physically interact with such objects, the visual system requires an estimate of the orientation of the object's surfaces with respect to the line of sight. This orientation is composed of tilt (rotation in the image plane, such as the changing orientation of the hands of a clock over time) and slant (rotation away from frontoparallel, such as changes in the steepness of a hillside road while driving). In estimating slant, the visual system is faced with the difficult challenge of inferring the 3D surface structure using the information contained within a pair of two-dimensional (2D) retinal images. Perception of surface slant represents a critical step in recognizing and interacting with 3D objects and has been studied extensively using psychophysical methods (Gillam and Ryan, 1992; van Ee and Erkelens, 1998; Backus et al., 1999; Knill, 2003; Hillis et al., 2004; Norman et al., 2006), yet its neural basis remains poorly understood.

To achieve slant estimation, the human visual system exploits various sources of information that may be present within the retinal inputs, such as binocular disparity, motion parallax, texture, perspective, and shadow (Gibson, 1950; Braunstein, 1968; Marr, 1982; Howard and Rogers, 1995). Given this range of depth cues, the visual system is believed to integrate signals to reduce noise and enhance perceptual judgements (Landy et al., 1995). For instance, adults have greater sensitivity to slant when two cues are available compared to the individual component cues (Knill, 2003; Hillis et al., 2004). However, relatively little is known about the neural implementation of such cue integration computations.

Here we aimed to test the neural basis of the integration of disparity and texture slant cues using human fMRI. To this end, we exploited multi-voxel pattern analysis (MVPA) to discriminate patterns of fMRI activity that were evoked by viewing surfaces with different slants. Thereby, we sought to quantify the information about differences between the viewed stimuli in different portions of the visual cortex. Logically, we might expect a number of different outcomes based on the sensitivity of a given area to different depth cues. First, consider an area that was exclusively sensitive to a given cue, for example, binocular disparity. The presence or absence of other cues should have little effect of the information about disparity contained in this area. Moreover, if different stimuli were presented that contained the same disparity, the area would be insensitive to stimulus differences. Such an area would constitute a pure disparity-processing module, and while potentially elegant from an engineering perspective, it would be biological unwieldy and is empirically unlikely (Cumming and DeAngelis, 2001; Parker, 2007; Orban, 2011; Welchman, 2011). Second, an area might be sensitive to different aspects of the cues that compose a slanted surface, but represent this information using different subpopulations. In this case information about depth from different cues would be collocated, but not integrated. The responses of this area would be expected to change in line with manipulations of different cues, and stimulus discrimination (based on the area's activity) would improve when two (or more) cues specified differences in slant between the stimuli. Finally, an area might integrate information from the different depth signals, thereby enhancing discrimination performance when two or more cues specified the same information about slant. Given that improved discrimination performance is expected under scenarios two and three, how can we differentiate them?

We can conceptualize the information about different depth configurations in terms of bivariate probability density distributions for different slanted stimuli (**Fig. 2.1a, b**).

One way for an ideal observer to discriminate between two such distributions would be to use the optimal decision boundary. Under this solution, slant estimates from the two cues remain independent. The improved discrimination performance that this computation would yield can be calculated from the quadratic sum of the discriminabilities of the component cues (**Fig. 2.1b**). Alternatively, the visual system might fuse component cue dimensions into a single depth estimate. Under this solution, improved discrimination



**Figure 2.1.** Illustration of the stimulus space and predictions for the discrimination performance of mechanisms based on the independent or fused use of depth cues. **A)** An observer is presented with two different slanted stimuli (anaglyph stereograms at the top) which we denote red and blue. We can conceptualise detectors for these stimuli based on independent depth estimates from disparity and texture. These detectors yield an estimate for each stimulus with a certain probability density function. The outputs of these detectors are conceptualised as bivariate Gaussian probability distributions (3D plot below) – with one distribution for the blue stimulus, and one for the red stimulus. The marginal projections (i.e. performance of the single cue detectors) are illustrated on the walls of the 3D plot. **B)** A planar projection of the bivariate Gaussians (colour saturation indicates probability density) to illustrate two possible computations when using the information from two depth cues. The red and blue stimuli could potentially be discriminated using either cue alone, but with some uncertainty (overlapping probability distributions). Two computations for reducing this uncertainty are illustrated: Independent – separation orthogonal to the optimal discriminating boundary (negative diagonal) and Fusion – multiplication of the probability densities associated with each cue. **C)** The idealized performance of a mechanism based on Fusion (top) or Independence (bottom). Illustrations on the right depict neural implementations within single voxels for the extreme cases of these two alternative computations. Each sphere represents a neuronal population that encodes depth from either component cue (yellow and blue) or from a combination of cues (green). An intermediate implementation (not depicted here) would contain all of these population types within a voxel.

occurs through a reduction in variance of the depth estimate.

To test which of these computations (independence or fusion) the visual system uses to achieve improved discrimination performance for combined cues, we assessed changes in fMRI responses related to independent manipulations of each cue. Our logic was that fMRI responses to different depth configurations should be more discriminable when texture and disparity cues were congruent. Further, the extent of this improved discriminability provides a means of distinguishing between the two alternative computations that might underlie cue integration (Ban et al., 2012). In particular, we expect that sensitivity to ‘single cue’ stimuli (**Fig. 2.2**: Texture; Disparity stimuli) should be attenuated by the presence of the conflicting cue only if depth estimates from the two cues are fused into a common representation. In that case, performance for congruent cue stimuli will exceed that predicted by quadratic summation of performance for ‘single cue’ stimuli (**Fig. 2.1c**). In contrast, if depth representations are co-located but remain independent, then performance for combined cue stimuli should be predicted by quadratic summation. Using this logic, we used MVPA to decode information about surface slants defined by disparity, texture and their congruent and incongruent combinations.

While the experimental logic and methods employed here have previously been used to test for cue fusion (Ban et al., 2012), there are two fundamental differences. First, we test two qualitatively different types of cue: disparity and texture. Whereas motion and disparity are computationally similar cues that can both be used to compute absolute depth (given knowledge of ego-motion, inter-pupillary distance, and gaze angle), texture provides only relative depth information (Landy et al., 1995). Outputs from independent texture and disparity processing circuits cannot therefore simply be averaged in order to estimate depth. It remains an open question as to whether the integration of qualitatively similar (motion and disparity) and dissimilar (texture and disparity) cue pairings share a common

neural substrate. Second, the previous study assessed cue combination for simple depth order relationships between frontoparallel surfaces. Here we test neural responses to slanted surfaces - a far more behaviourally relevant depth structure that is important for the estimation of extended surfaces and a critical step in the processing of complex/ curved surfaces and solid 3D objects.

## 2.2. Materials and Methods

### *Observers*

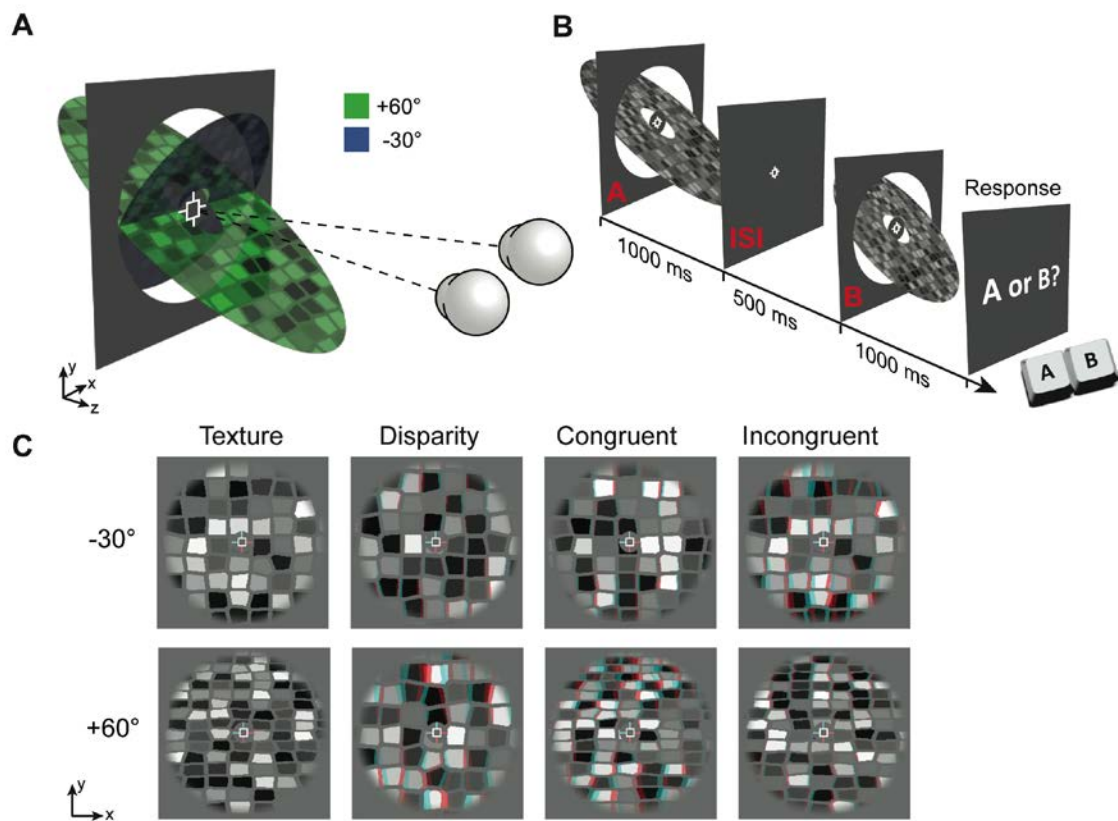
Fifteen healthy observers from the University of Birmingham with normal or corrected-to-normal vision participated in the fMRI experiment. The mean age was 25.8 years (range = 21-37, 6 female). Three participants' data were excluded from analysis due to excessive head movements during scanning, which prevented voxel correspondence necessary for multi-voxel pattern analysis (MVPA). Excessive head movement was defined as more than 10 sharp head movements ( $\geq 1$  mm displacement between consecutive volumes), or  $\geq 4$  mm total displacement over the course of a 7 minute run. After excluding runs that exceeded these thresholds, three participants had fewer than 5 remaining runs, and so were excluded entirely as there was insufficient data for MVPA. In general, participants were able to remain still throughout scanning (the mean maximum head displacement per run for included participants was 1.2 mm vs. 2.8 mm for excluded participants). Participants were screened for stereoacuity and contraindications to MRI prior to the experiment. All experiments were conducted in accordance with the ethical guidelines of the Declaration of Helsinki and were approved by the University of Birmingham STEM ethics committee.

### *Stimuli*

Stimuli were grayscale stereograms of textured planar surfaces, with slant (rotation about a central horizontal axis) defined independently by texture and disparity (**Figs. 2.2a, c**). Surface textures were generated by voronoi tessellation of a regular grid of points ( $1^\circ \pm 0.1^\circ$  point spacing) randomly jittered in two dimensions by up to  $0.3^\circ$  (Knill, 2003; Nardini et al., 2010a). Each texture element (texel) was randomly assigned a grey level and shrunk

about its centroid by 20%. This created the appearance of ‘cracks’ between the texels, the width of which also varied with surface slant and thus provided additional texture information. Textures were mapped onto a vertical virtual surface and rotated about the horizontal axis by the specified texture-defined slant angle, before a perspective projection consistent with the physical viewing geometry was applied. From this cyclopean view, binocular disparity was then calculated and applied to each vertex based on the specified disparity-defined slant angle.

Surfaces were presented inside a circular aperture with a radius of  $3.5^\circ$ , and a cosine edge profile in order to blur the appearance of depth edges. Stimuli were presented



**Figure 2.2.** A) A cartoon illustrating the opposing slanted surfaces tested in the study. The two eyes are illustrated fixating the marker at the center of the display. B) The sequence of events for a single trial in the behavioural experiment. C) Examples of the two slant levels for the 4 conditions tested in the experiment. Stimuli are rendered as red-cyan anaglyphs for stereoscopic viewing (red filter over left eye). Note that for the Incongruent condition, the stimuli are labeled according to the slant specified by texture – i.e. in the top ( $-30^\circ$ ) stimulus, texture signals  $-30^\circ$  and disparity signals  $+60^\circ$  slant. In the bottom stimulus, texture signals  $+60^\circ$  and disparity  $-30^\circ$ .



on a mid-grey background, surrounded by a grid of black and white squares (75% density) designed to provide an unambiguous background reference. Texture-defined position in depth—which corresponds to mean texel size—was randomized for each stimulus presentation by increasing point spacing in the initial grid of points by  $\pm 10\%$ . This reduced the reliability of mean texel size in a given portion of the stimulus as a cue to surface slant. Disparity-defined position in depth was kept constant across all stimulus presentations.

Stimuli in the ‘single cue’ conditions (Texture, and Disparity) were intentionally designed to contain a cue conflict for two reasons. First, the premise of our test for cue fusion was that the presence of cue conflict would reduce sensitivity to ‘single cue’ conditions only if cues are fused, but not if they remain independent. Second, the presence of both component cues in all stimulus conditions minimized low-level differences that would otherwise complicate comparison of responses to different conditions. For example, monocular presentation (which is frequently used to isolate texture cues in behavioural studies) is known to significantly affect both univariate and multivariate fMRI responses (Büchert et al., 2002; Schwarzkopf et al., 2010). We presented all stimuli binocularly, and stimuli in the Texture condition were given a disparity-defined slant of  $0^\circ$ . To attenuate the reliability of the disparity cue for this condition, the disparity-defined position in depth of each texel was randomly jittered between  $\pm 2$  arcmin. Similarly, stimuli in the Disparity condition had a texture-defined slant of  $0^\circ$ . Stimuli in the Incongruent condition consisted of each cue signaling one of the two base slant angles (i.e. texture signaled  $-30^\circ$  and disparity signaled  $+60^\circ$  or vice versa). For the behavioral experiment, the increments applied to create test stimuli shifted texture and disparity defined slants in opposite directions to each other (i.e. the absolute slant of both cues either increased or decreased relative to the base slants).

Stimuli were programmed and presented in MATLAB (The MathWorks, Inc., Matick, MA) using Psychophysics Toolbox extensions (Brainard, 1997;Pelli, 1997). Stereoscopic presentation in the scanner was achieved using two projectors (JVC, D-ILA SX21) containing separate spectral interference filters (INFITEC). The two projections were optically combined using a beam-splitter and entered the scanner room through a wave guide. This method produced negligible cross talk between the two images, since the filtered emission spectra for the two projectors contained little overlap. Stimuli were back-projected (1280 x 1024, 52 x 41.6cm) onto a translucent screen inside the bore of the magnet and viewed via a front-surfaced mirror attached to the head coil and angled at 45° above the observers' heads. This resulted in a viewing distance of 65cm, from which the entire slanted plane stimulus was visible within the binocular field of view. In a separate session, a subset of participants (n = 3) repeated the experiment in the scanner while eye movement was recorded monocularly using a CRS limbus eye tracker (Cambridge Research Systems Inc.).

For the initial psychophysical experiment, binocular presentation was achieved using a stereoscope setup consisting of a pair of ViewSonic FB2100x CRT monitors (1600 x 1200, 100Hz) viewed through front-surfaced mirrors at a viewing distance of 50cm. A subset of participants (n = 5) repeated the experiment on a second stereoscope, which consisted of a pair of Samsung 2233RZ LCD monitors (1280 × 1024, 120Hz). These monitors were viewed through front surfaced mirrors, at a viewing distance of 50cm and participants' head position was stabilized by an eye mask, head rest and chin rest. For all display devices in all experiments, linearized gamma tables were calculated based on photometric measurements (Admesy, Ittervoort, The Netherlands), and paired display devices were calibrated to produce matched luminance outputs.

### *Vernier Task*

During experimental scans, participants performed an attentionally demanding Vernier task at fixation and were not required to discriminate slant. This served two purposes. First, it ensured consistent attentional allocation between conditions, making it unlikely that systematic differences in attention could modulate decoding performance. Second, it provided a subjective measure of eye position, allowing us to assess whether there were any systematic differences in eye vergence between conditions. Participants were instructed to fixate a central crosshair fixation marker that was present throughout experimental scans. The fixation marker consisted of a white square outline (side length,  $0.5^\circ$ ) and horizontal and vertical nonius lines (length, 22 arcmin) presented inside a mid-grey disk (diameter  $1.25^\circ$ ). One horizontal and one vertical line were presented to each eye in order to promote stable vergence and to provide a reference for a Vernier task (Poppo et al., 1998). The Vernier target line subtended 6.4 arcmin in height by 2.1 arcmin in width, and was presented at 7 evenly spaced horizontal offsets of between  $\pm 6.4$  arcmin for 500ms (with randomized onset relative to stimulus) on 50% of TRs. Participants were instructed to indicate, by button press, which side of the central upper vertical nonius line the target appeared, and the target was presented monocularly to the contralateral eye. We obtained a subjective measure of observers' vergence by fitting a cumulative Gaussian psychometric function to the proportion of 'target on the right' responses against horizontal target offset. Bias (deviation from the desired vergence position) in participants' judgments was close to zero.

### *Psychophysics*

In the initial psychophysics experiment conducted prior to the fMRI scans, participants performed a two interval forced choice discrimination task in which they were sequentially presented a reference stimulus (one of the 8 conditions: 2 slants x 4 cue

configurations, **Fig. 2.2c**) and a test stimulus in a randomized order (**Fig. 2.2b**). Each stimulus was presented for 1000ms (as in the scanner) with an inter-stimulus interval of 500ms between the two. Participants were instructed to indicate which interval contained the more slanted surface (i.e. which was further from frontoparallel). The difference in slant angle ( $\Delta d$  and/or  $\Delta t$ ) between reference and test stimuli was controlled via an adaptive staircase method (Watson and Pelli, 1983), with a maximum limit of  $\pm 20^\circ$ . This limit was chosen in order to avoid presentation of test slants of greater than  $80^\circ$  (at which point the surface required to fill the stimulus aperture becomes prohibitively large, and disparities cannot be fused (Burt and Julesz, 1980) or smaller than  $10^\circ$  (at which point textures cannot be reliably discriminated from frontoparallel). Further, previous studies reported discrimination thresholds of less than  $20^\circ$  for this type of stimulus (Knill, 2003).

In the main behavioural experiment, a staircase consisting of 20 trials was run for each base slant in each condition, and all conditions were interleaved in a pseudo-randomized order. Data were then combined to calculate the just noticeable difference (j.n.d.) for the four different cue configuration conditions (Disparity, Texture, Congruent, Incongruent). Threshold estimates greater than 125% of the maximum slant difference ( $20^\circ$ ) were substituted with a value of  $25^\circ$ , since these estimates were based on insufficient sampling of the psychometric function (Wichmann and Hill, 2001a;b). This typically occurred for the Incongruent cue condition, as observers required more than  $20^\circ$  difference between stimulus pairs to reliably judge slant differences. Thus, the mean sensitivity for Incongruent stimuli is likely an over-estimate of the true sensitivity, and the relatively low S.E.M. in this condition is due to capping upper values of the threshold.

To obtain more detailed threshold estimates, a subset of participants ( $n = 5$ ) were retested over 7 separate sessions, in which interleaved staircases were programmed to converged at six different points distributed broadly along the psychometric function. For

each condition tested (Texture, Disparity, and Congruent), data from 820 trials were fitted with a cumulative Gaussian psychometric function using a bootstrapping method (Psignifit toolbox; Wichmann and Hill, 2001a;b). Parameters for lapse rate ( $\lambda$ ) and guess rate ( $\gamma$ ) were constrained to the range 0-0.1, while the point of subjective equality ( $\beta$ ) was constrained to equal the base slant. Threshold estimates from these data were consistent within participants across sessions, and were used in place of the original threshold estimates for the observers and conditions that were retested. This method generally produced slightly lower threshold estimates compared to the single staircase method, since incorrect responses early in a session had less influence on the subsequent sampling range. This was most noticeable for the Texture condition, where all threshold estimates measured using multiple staircases were less than the maximum slant difference tested ( $20^\circ$ ).

To test for behavioural performance compatible with a fusion computation, we calculated predictions for the minimum bound based on quadratic summation (**Fig. 2.1b**). Quadratic summation was calculated using the formula:

$$\text{Quadratic } \Sigma = \sqrt{S_D^2 + S_T^2} \quad (1)$$

where  $S_T$  and  $S_D$  are the observer's sensitivities (1/j.n.d.) in each of the single cue conditions. To further contrast observers' performance on combined cue conditions against the performance predicted by quadratic summation, we calculated a psychophysical integration index ( $\psi$ ):

$$\psi = \frac{S_{DT}}{\sqrt{S_D^2 + S_T^2}} - 1. \quad (2)$$

A value of zero for this index would indicate that sensitivity to combined cues was equal to that predicted by quadratic summation of individual cues (the minimum bound for

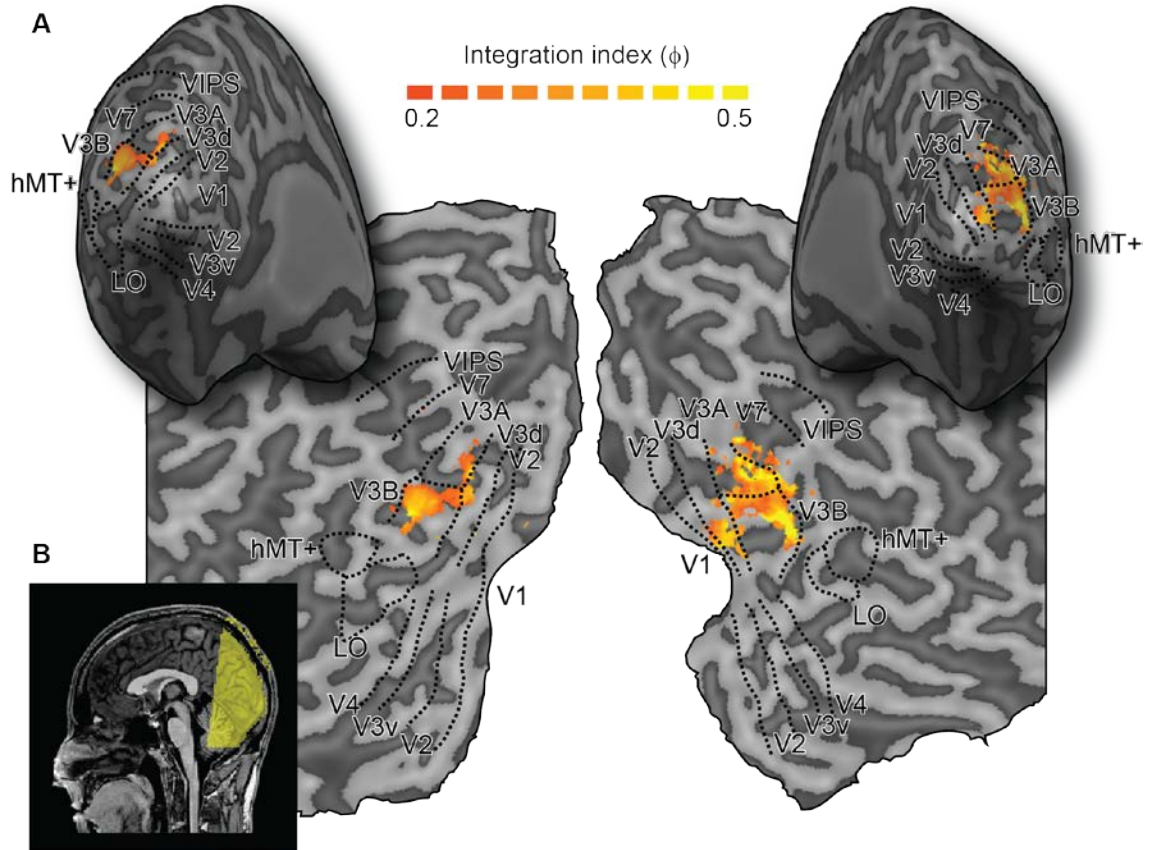
fusion). Integration indices were calculated for each observer and then bootstrapped using 10,000 iterations.

### *Imaging*

Data were collected at the Birmingham University Imaging Centre using a 3 tesla Philips Acheiva MRI scanner with an 8-channel head coil. Blood oxygen level-dependent (BOLD) fMRI data were acquired using a gradient echo echo-planar imaging (EPI) sequence [echo time (TE), 34ms; repetition time (TR), 2000ms; voxel size, 1.5 x 1.5 x 2 mm, 28 slices covering occipital cortex] for experimental and localizer scans. Slice orientation was close to coronal section but rotated about the mediolateral axis with the ventral edge of the volume more anterior in order to ensure coverage of lateral occipital cortex (**Fig. 2.3b**). A high resolution T1-weighted anatomical scan (voxel size, 1mm isotropic) was additionally acquired for each participant.

### *Defining regions of interest*

We localized regions of interest (ROIs) for each participant using standard retinotopic mapping procedures (**Fig. 2.3a**). Retinotopically organized visual areas V1, V2, V3v, V4, V3d, V3A, V3B, and V7 were defined using polar and eccentricity maps, which were generated by fMRI responses to rotating wedge and expanding concentric ring stimulus presentations respectively (DeYoe et al., 1996; Sereno et al., 1995). V3 was



**Figure 2.3.** **A)** Representative inflated cortical surface reconstructions and flat maps showing the visual regions of interest from one participant. These include retinotopic areas, V3B/KO, human middle temporal complex (hMT+) and the lateral occipital (LO) area. Sulci are rendered a darker grey than the gyri. Superimposed on the maps are thresholded fMRI integration indices (see Methods, Eq. 3). These were calculated from the results of group searchlight classifier analyses that moved iteratively throughout the measured volume of cortex, discriminating between stimulus slant angles ( $-30^\circ$  or  $+60^\circ$ ) for the Congruent condition and each of the component cue conditions.. This procedure confirmed that we had not missed any important areas outside those localized independently. **B)** Sagittal view of a structural MRI scan of one participant, with coloured overlay illustrating slice orientation for functional scans.

divided into dorsal and ventral quadrants in each hemisphere (V3d and V3v) in line with previous delineation based on both functional and cytoarchitectonic distinctions (Wilms et al., 2010). Area V4 was defined as the region of ventral visual cortex adjacent to V3v containing a representation of the upper quadrant of the contralateral visual field (Tootell and Hadjikhani, 2001; Tyler et al., 2005). Area V7 was defined as the region of retinotopic



activity anterior and dorsal to V3A (Tootell et al., 1998;Tsao et al., 2003;Tyler et al., 2005).

Area V3B/KO (Zeki et al., 2003;Tyler et al., 2006) was defined as the area comprising the union of retinotopically mapped V3B and an area ‘KO’ that was functionally localized by its preference for motion-defined boundaries compared to transparent motion of black and white dots ( $p < 10^{-4}$ ) (Dupont et al., 1997;Tootell and Hadjikhani, 2001;Zeki et al., 2003;Tyler et al., 2006). This area contained a full hemifield representation, inferior to V7 and anterior to, and sharing a foveal confluence with, V3A (Tyler et al., 2005). A previous study revealed no clustering of voxels selective for cue integration within V3B/KO that might provide grounds for separating the two areas (Ban et al., 2012; their Supplementary Fig. S7).

Human middle temporal complex (hMT+/V5) was defined as the set of voxels in lateral temporal cortex that responded significantly more strongly ( $p < 10^{-4}$ ) to transparent dot motion compared to a static array of dots (Zeki et al., 1991). Lateral occipital complex (LOC) was defined as the set of voxels in lateral occipito-temporal cortex that responded significantly more strongly ( $p < 10^{-4}$ ) to intact compared to scrambled images of objects (Kourtzi et al., 2005). LOC subregion LO, which extended into the posterior inferotemporal sulcus, was defined based on the overlap of functional activations and anatomical structures, consistent with previous studies (Grill-Spector et al., 2000). For a subset of participants ( $n = 8$ ) regions of the intra-parietal sulcus (VIPS, POIPS, DIPSM) were functionally localized based on significantly stronger responses to 3D structure-from-motion than for 2D transparent motion (Orban et al., 1999;Orban et al., 2006a), and defined using anatomical landmarks.

### *fMRI design*

In the fMRI experiment, each run consisted of 24 blocks and began and ended with a 16s fixation period during which only the background and fixation marker were presented. Each block was 16s duration and consisted of a stimulus from one of the eight conditions (a new stimulus was created for each trial) being presented for 1s followed by a 1s fixation period. Block order was randomized and each of the 8 conditions was presented for 3 blocks per run. Eleven observers completed 9 runs and one completed 8.

### *fMRI data analysis*

Anatomical scans of each participant were transformed into Talairach space (Talairach and Tournaux, 1988) and inflated and flattened surfaces were rendered using BrainVoyager QX™ (BrainInnovation, Maastricht, The Netherlands). Functional data were preprocessed using slice timing correction, head motion correction and high-pass filtering before being aligned to the participant's anatomical scan and transformed into Talairach space. We ran univariate analyses using random effects (RFX) group GLM analysis to contrast BOLD responses to the two slant angles, for each of the four conditions. Additionally we contrasted the combination of responses to the two component cue conditions (Texture + Disparity) vs. the Congruent condition, and Incongruent vs. Congruent.

For multivariate analyses, we tested the discriminability of BOLD responses to the two different slant angles ( $60^\circ$  vs.  $-30^\circ$ ). Since visual areas differ considerably in size, and ROIs that contain more voxels could yield higher classification accuracies, we selected the same number of voxels across ROIs and observers (Haynes and Rees, 2005; Serences and Boynton, 2007; Preston et al., 2008). For each ROI we selected gray matter voxels from both hemispheres and ranked them based on their response (as indicated by  $t$ -values calculated using the standard general linear model) to all stimulus conditions compared to

fixation baseline across experimental runs. Pattern size was restricted to voxels showing a  $t$  value greater than 0 for the contrast of all stimuli versus fixation. Assessment of the point at which classification accuracies tended to saturate as a function of pattern size resulted in the selection of 300 voxels per ROI. In 11 ROIs (out of 120 across all participants) where fewer than 300 voxels reached threshold, all voxels above threshold were included in the ROI (mean =  $199 \pm 16$  SEM). This voxel selection process has been validated by previous studies (Kamitani and Tong, 2005; Preston et al., 2008).

We normalized each voxel time course separately for each experimental run (by calculating z-scores) to minimize baseline differences across runs. The multivariate analysis data vectors were produced by applying a 4s temporal shift to the fMRI time series to compensate for the lag time of the hemodynamic response function, and then averaging all time series data points of an experimental condition. Univariate signals were removed before conducting SVM analysis, by subtracting the mean response of all voxels in a given volume, from each voxel in that volume. All volumes therefore had the same mean value across voxels, but different patterns of activity. We used a support vector machine (SVM<sup>light</sup> toolbox, <http://svmlight.joachims.org>) for classification and performed a ninefold leave-one-out cross-validation, in which data from eight runs were used as training patterns (24 patterns, 3 per run) and data from the remaining run were used as test patterns (3 patterns). For each participant, the mean accuracy across cross-validations was taken.

To quantify differences in SVM prediction accuracies between combined cue conditions and the minimum bound prediction, we calculated an fMRI integration index:

$$\phi = \frac{d'D + \tau}{\sqrt{d'D + d'T}} - 1 \quad (3)$$

where  $d'_{D+T}$ ,  $d'_D$ , and  $d'_T$  are SVM discriminability in d-prime units, calculated using the formula:

$$d' = 2 \times \text{erfinv}(2p-1) \quad (4)$$

where  $\text{erfinv}$  is the inverse error function and  $p$  is the proportion of correct classifications.

To test for transfer of classifier learning between individual cue conditions, we used recursive feature elimination (RFE) to select the voxels in each ROI that made the greatest contribution to classifier performance (De Martino et al., 2008). Voxels were iteratively discarded independently for each cross-validation, for each condition, and the final voxels used for SVM analysis were selected from the intersection of corresponding validations.

In addition to performing our standard region of interest analysis, we used a ‘searchlight’ MVPA approach (Kriegeskorte et al., 2006), which sequentially sampled small volumes of cortical voxels within a 9mm radius, throughout the entire cortical volume. Prior to analysis, a mask was applied to exclude voxels located outside of cortex, and voxels with a  $t$  value less than 0 for the contrast of all stimuli versus fixation were excluded. For each sample, we ran three SVM classification analyses discriminating the activity evoked by different slants in the Texture, Disparity and Congruent cue conditions. From these sensitivity values we then calculated the fMRI integration index ( $\phi$ ) for each sample (**Fig. 2.3a**). Using this approach we sought to determine whether we had missed any important areas outside of the independently localized regions of interest.

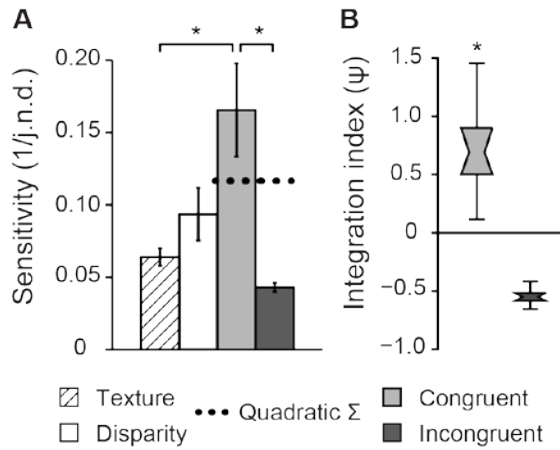
Analysis of variance was performed in SPSS (IBM Corporation), and Greenhouse-Geisser correction was applied where the assumption of sphericity was not met. We report the results of two different statistical tests that both assess cue fusion: repeated measures ANOVA and bootstrapped fMRI integration indices. The former test estimates variances

based on the sample of SVM sensitivities (12 participants), assuming a Normal distribution. The latter estimates variances in a manner that is model free, since the population (10,000 data points) is calculated from permutations of the sample. While ANOVA is a standard parametric test, we include the bootstrap analysis for a more sensitive and direct assessment of our key statistic of interest: sensitivity to Congruent cues relative to quadratic summation of component cue sensitivities.

## 2.3. Results

### Psychophysics

We measured participants' sensitivity to small differences in surface slant defined in the Disparity, Texture, Congruent and Incongruent conditions. Specifically, we presented two stimuli and asked observers to decide which had the greater slant, using an adaptive staircase procedure to control the slant difference between the two stimuli. We found significant differences in the ability of observers to judge slant under the different conditions ( $F(1.5,16.9)=7.191$ ,  $p<0.01$ ), with best discrimination performance observed when slant differences were congruently indicated by both texture and disparity (**Fig. 2.4a**). This performance was significantly better than that in the two 'single' cue conditions



**Figure 2.4.** Behavioural tests of cue integration. **A)** Bars represent the mean sensitivity across observers ( $n = 12$ ) and base slants for the four experimental conditions. Sensitivity is calculated as the reciprocal of the just noticeable difference (j.n.d.), which is in degrees. Error bars show between-observers SEM. The dashed horizontal line indicates predicted sensitivity for combined cue conditions based on quadratic summation of the observed sensitivities to individual cue conditions.  $*p<0.05$ . **B)** Distribution plots of bootstrapped psychophysical integration indices for combined cue conditions. A value of zero indicates that combined cue performance was equal to that predicted by quadratic summation. Notches represent median, boxes represent 68% confidence intervals, and the error bars represent 95% confidence intervals.  $*p<0.001$ .

(Texture condition:  $F(1,11)=8.96$ ,  $p<0.01$ ; Disparity condition:  $F(1,11)=3.60$ ,  $p=0.042$ , one-tailed) confirming the expectation that slant sensitivity improves for congruent combinations of cues. As discussed in the Introduction, an improvement in performance *per se* would be expected on the basis of both a fusion mechanism and a discrimination based on the outputs of independent detectors. We were therefore interested in the extent of the improved performance in the Congruent cue condition, relative to performance for the ‘single cue’ component conditions.

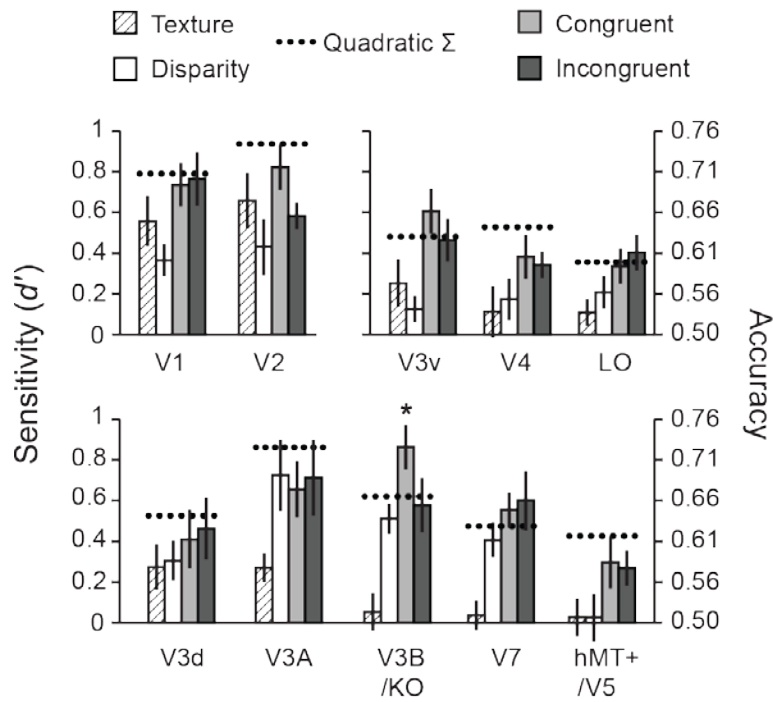
To assess whether participants integrated information from disparity and texture in line with cue fusion, we used a psychophysical integration index ( $\psi$ ) that compared performance in the Congruent cue case to the minimum bound expected for fusion using the quadratic summation of component cue accuracies (**Fig. 2.4b**). Using this index, a value of zero would suggest performance improvements in line with independent access to texture and disparity signals, while a value in excess of zero suggests that signals from the two cues are fused to support perceptual judgments. We found performance in the Congruent cue condition exceeded the minimum bound for fusion ( $p<0.001$ ). However, when the cues were Incongruent,  $\psi$  did not exceed zero ( $p>0.999$ ). Indeed, performance in the Incongruent condition was, on average, lower than the single cue component conditions, although the low variance observed for this condition was a result of capping maximum threshold estimates (see *Materials and methods*). Together, these results suggest that participants fused the information from disparity and texture cues to estimate surface slant.

### *fMRI*

We measured fMRI responses in regions of interest within the visual cortex (**Fig. 2.3**) while participants were presented with slanted stimuli ( $-30^\circ$  and  $+60^\circ$ ) defined by

Disparity, Texture, Congruent or Incongruent cues (**Fig. 2.2**). Univariate analyses revealed no significant differences in BOLD response to different slants in any condition, except for in the Texture condition. For this condition, small clusters of voxels in foveal regions of early visual areas (V1 and V2) showed significant changes in mean response to different texture-defined slants. This result likely reflects differences in spatial frequency between these stimuli. However, we did not find this for any other conditions, suggesting that differences in BOLD responses related to spatial frequency were only detectable in the absence of concomitant differences in disparity gradient. Importantly, none of the cortical areas for which subsequent multivariate analysis suggest cue fusion (see below) showed any reliable differences in univariate response between slants or conditions.

We next quantified the information about the presented stimuli in each region of



**Figure 2.5.** Classifier prediction performance for discriminating +60° versus -30° slants for all regions of interest. Dotted lines indicate the performance predicted for combined cue conditions based on quadratic summation of performance observed in the Texture and Disparity (single cue) conditions. Error bars indicate SEM. Asterisks indicate areas where sensitivity for congruent cues (green bars) significantly exceeded performance predicted by quadratic summation (red horizontal lines) ( $p < 0.05$ ).



interest by training a machine learning classifier to discriminate the patterns of fMRI activity evoked by the different slants. We were able to reliably decode the presented stimuli in all regions of interest, and decoding accuracy varied by condition (**Fig. 2.5**). Notably, classifier sensitivity to Texture was close to zero in higher dorsal and ventral areas, yet these same regions showed improved sensitivity for Congruent compared to Disparity conditions. This suggests that these areas do process texture in the absence of cue conflict, and the low classification accuracies observed for the Texture condition may have been related to the dominance of disparity signals in that condition. In contrast, decoding accuracies for Disparity were relatively high in dorsal areas, in line with previous findings for decoding of fMRI responses to disparity (Preston et al., 2008; Ban et al., 2012). Our primary interest, however, was not to compare classification accuracies between ROIs (since these are influenced by various factors besides neural activity), but to assess the relative performance between conditions within each area.

In early visual cortex (V1 and V2) classifier accuracy was above chance for all conditions. However, it is unlikely that these areas explicitly encode representations of surface slant from texture and/or disparity. Instead, the representations of local disparities and image features represented in these areas are likely to provide important inputs to higher areas that encode depth gradients. This low-level image feature information differs between the different stimulus classes, and as such supports decoding of one slant from the other. Reassuringly, classifier performance for the Congruent condition did not exceed quadratic summation in early visual areas, suggesting that signals related to texture and disparity signals remain independent at this stage of processing.

In contrast, prediction accuracies for the Congruent condition were significantly higher than for either component cue in V3v ( $F(3,33) = 5.41$ ,  $p < 0.01$ ) and V3B/KO ( $F(3,33) = 13.16$ ,  $p < 0.001$ ). To assess integration, we calculated the minimum bound prediction based on quadratic summation of classifier sensitivities to individual cues (**Fig. 2.5**, red lines). Classifier performance for Congruent stimuli was higher than predicted by quadratic summation in areas V3v, V3B/KO, and V7. This difference was only statistically significant in area

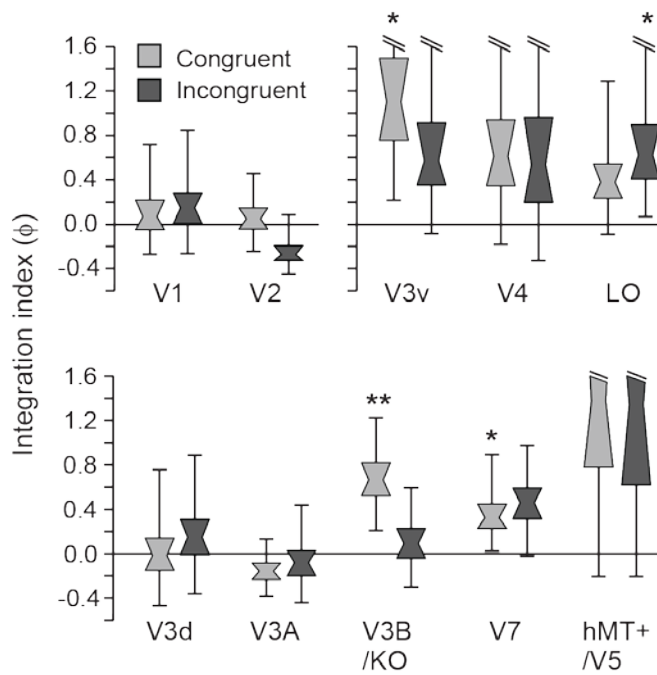
V3B/KO, and—importantly—only when cues were congruent ( $F(1,11) = 4.16$ ,  $p = 0.033$ ), but not when they were incongruent ( $F(1,11) < 1$ ,  $p = 0.39$ ; **Fig. 2.5**). The pattern of classifier prediction accuracies across conditions in V3B/KO is therefore reminiscent of the pattern of observers' behavioral accuracy for discriminating between slants (**Fig. 2.4**).

To quantify the differences between classifier accuracies for combined cue conditions and the minimum bound prediction, we calculated fMRI integration indices (*Materials and Methods*, equation 3). Based on previous findings, we anticipated that the cue integration index would exceed zero in V3B/KO for the Congruent condition (Ban et al., 2012). This expectation was confirmed as statistically reliable (after Bonferroni correction for multiple comparisons) in V3B/KO ( $p = 0.005$ ). In addition we found integration indices for the Congruent condition above zero (uncorrected significance

**Table 2.1.** Significance tests for the integration index ( $\phi$ )

Cortical region	P-value	
	Congruent $\phi > 0$	Incongruent $\phi \leq 0$
V1	0.373	0.314
V2	0.43	0.909
V3v	0.01	0.081
V4	0.125	0.212
LO	0.099	0.023
V3d	0.033	0.300
IPPA	0.81	0.637
V3B/K	<b>0.00</b>	0.358
V7	0.037	0.051
hMT+/V5	0.077	0.102

Bold  $p$ -values indicate significance after Bonferroni correction for multiple comparisons; italicized  $p$ -values indicate significance at the uncorrected threshold ( $p < 0.05$ ). LO, lateral occipital cortex; hMT+, human middle temporal complex.



**Figure 2.6.** Notched distribution plots of bootstrapped fMRI integration indices for all regions of interest. Plots are shown for the results obtained with Congruent and Incongruent stimuli. A value of zero indicates that combined cue performance was equal to that predicted by quadratic summation. Notches represent the median, boxes represent 68% confidence intervals, and the error bars represent 95% confidence intervals. \*  $p < 0.05$ ; \*\*  $p < 0.001$ .

threshold) in areas V3v and V7 (**Fig. 2.6; Table 2.1**). These results suggest cortical representations that are compatible with the fusion of slant information from the disparity and texture signals.

The inclusion of the Incongruent condition afforded an additional test of the idea that representations were fused, rather than collocated but independent. Specifically, if the representations of the two stimulus dimensions (disparity and texture) remain independent, then the performance predicted for the Incongruent condition will be equal to the quadratic sum of the component cues. This can be understood by considering that the separation between two stimulus distributions (as conceptualized in Fig. 1) remains the same, irrespective of the sign of the differences between component cues (Fig. 1b). In contrast, fusion predicts one of two possible outcomes, based on previous theoretical work (Clark and Yuille, 1990; Landy et al., 1995). Under ‘strict’ fusion, the variance of the combined cue estimate for an Incongruent stimulus would increase, and performance would deteriorate below that of either component cue. Alternatively, under ‘robust’ fusion, performance would revert to the level of one of the two component cues. We found that

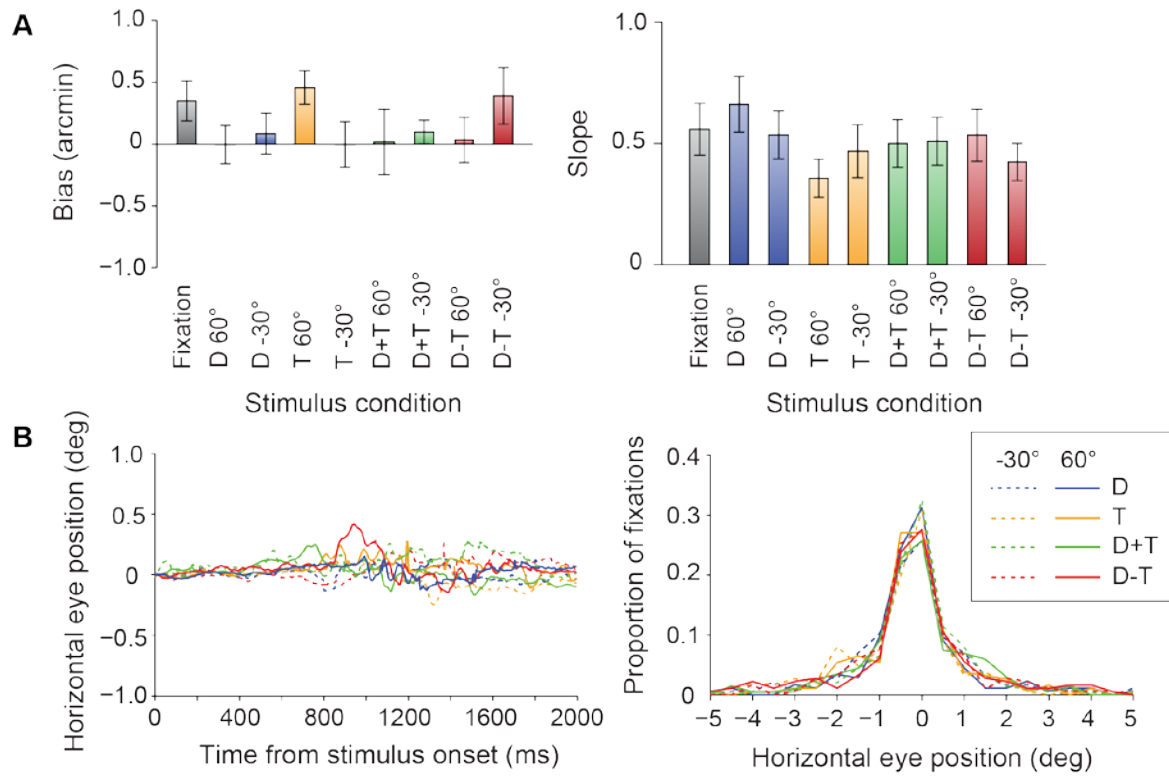
cue integration indices exceeding the minimum bound were not observed when cues were highly inconsistent in areas V3B/KO and V3v (Fig. 5) while results in area V7 were similar for Congruent and Incongruent cue configurations. In particular, accuracy for Incongruent stimuli in V3B/KO was slightly below that predicted by quadratic summation, and not separable from performance in the single cue disparity cue condition. This result is compatible with the robust fusion of cues in V3B/KO, and consistent with the results of Ban et al. (2012).

To assess the extent to which fMRI responses were related to depth structure in general, we tested whether training the classifier on data from one single cue condition (e.g. Texture) would allow predictions to be made for data obtained in the other single cue condition (e.g. Disparity). Such transfer between depth signals would suggest generalized depth representations. However, we failed to find any evidence of transfer between cues – specifically, accuracies for transfer between cues were not significantly different from baseline accuracies obtained by performing within-cue tests when the labels given to the classifier were randomly permuted. This result suggests that identical slants in the two single cue conditions elicited very different patterns of activity. This might reflect insufficient generalized representation of depth structure within individual voxels in the cortical regions tested. However, it might equally reflect the cue conflicts present in the single cue stimuli. For example, training the classifier to use disparity-related activity to discriminate between slants would not be useful for decoding slant in the Texture condition, where disparity always signaled zero slant.

We also considered the involvement of regions of the intraparietal sulcus (IPS). However, searchlight analyses (Kriegeskorte et al., 2006; Preston et al., 2008; Ban et al., 2012) revealed that sensitivity to Congruent stimuli did not significantly exceed quadratic summation of sensitivities to component cue stimuli in regions beyond extrastriate cortex

in many of our participants (**Fig. 2.3**). Correspondingly, we did not find reliable decoding accuracies in any of the IPS regions tested (VIPS, POIPS, DIPSM) for the subset of participants for whom we had independently localized parietal regions ( $n = 8$ ). This might indicate that slant representation is sparsely encoded in these regions (as appears to be the case in inferior temporal cortex; Liu et al., 2004), and therefore unable to significantly bias the activity of individual voxels. Further, these parietal regions might be more sensitive to tilt than slant (Shikata et al., 2001; Shikata et al., 2008), or to 3D boundaries than to large surfaces (Theys et al., 2012), since in-plane orientation and object segmentation are more critical for guiding eye-movements and grasping responses.

A possible concern is that the decoding accuracies reported here might reflect fMRI activity not related to sensory processes. Our inclusion of control measures was therefore important. In particular, participants performed an attentionally demanding Vernier task at fixation during scans, which ensured consistent attentional allocation between conditions, and provided a subjective measure of eye position. Using this method, we found no significant differences in estimated vergence position between conditions (**Fig. 2.7**). Although it was not possible to measure eye vergence directly during the scan, we measured eye position (monocularly) while participants performed the task in the scanner during separate test sessions. These measures suggested that observers were able to reliably maintain fixation and revealed no systematic differences in eye position between conditions (**Fig. 2.7**). Taken together, these control measures suggest that differential attentional allocation or differences in eye movements are unlikely to account for our findings.



**Figure 2.7.** Subjective and objective eye movement analysis. **A)** Participants performed a Vernier task during scanning, in which they reported the perceived direction of horizontal offset of a small target relative to the central nonius line of the fixation marker. For each participant we fitted a cumulative Gaussian to the proportion of ‘right’ responses as a function of the horizontal offsets of the targets, to obtain bias and slope measurements. Bias (deviation from the desired vergence position) in participants’ judgments was close to zero. Repeated measures ANOVA revealed no significant effect of condition ( $F(3,33)=1.15$ ,  $p=0.34$ ), slant angle ( $F(1,11)<1$ ,  $p=0.95$ ), or interaction ( $F(3,33)=2.32$ ,  $p=0.09$ ). Similarly we found no differences in the slope of the psychometric functions by condition ( $F(3,33)<1$ ,  $p=0.41$ ), slant ( $F(1,11)<1$ ,  $p=0.87$ ), or interaction ( $F(3,33)=2.81$ ,  $p=0.07$ ). Performance on the Vernier task therefore suggests that participants were able to maintain stable eye vergence (Poppo et al., 1998). Error bars show SEM between subjects ( $n = 12$ ). **B)** We recorded horizontal eye movements for a subset of participants ( $n=3$ ) while they lay in the scanner bore, using a monocular eye tracker (CRS, Ltd, Rochester, UK) with a stated accuracy of  $< 0.25$  degrees visual angle. No significant differences were observed across conditions for eye position ( $F(3,6)=1.39$ ,  $p=0.34$ ), number of saccades ( $F(3,6)=2.26$ ,  $p=0.18$ ), or saccade amplitude ( $F(3,6)=1.38$ ,  $p=0.34$ ). Traces of mean eye position aligned to the start of each trial showed only small deviations from fixation, with no systematic differences between conditions.

## 2.4. Discussion

---

Reliable estimation of 3D surface orientation is a critical step in recognizing and interacting with complex 3D objects, and can be enhanced by integrating information from multiple depth cues. Binocular disparity and texture are two important—but qualitatively very different—cues to surface orientation, yet little is known about the neural processes involved in integrating these sources of information. Here we assessed fMRI responses in visual cortex to slanted surfaces defined by texture and disparity cues using two criteria. First, we used a quadratic summation prediction to test for regions where performance for congruent cue stimuli exceeded the performance expected if depth representations from texture and disparity are co-located but independent. Second, we tested whether this result was specific to conditions where the two cues congruently signaled the same depth configuration. We found that the first criterion was met in visual area V3B/KO and to a lesser extent in V3v and V7, while the second criterion was again met most reliably in V3B/KO, but to some extent in V3v. These results suggest that area V3B/KO in particular integrates texture and disparity cues by fusing depth estimates, while other extrastriate visual areas (V3v) might also support fusion computations in addition to independent representations of these cues. Further, the pattern of slant discriminability across stimulus conditions in V3B/KO was consistent with the pattern of behavioural sensitivity to surface slant when observers were tested on a discrimination task. These findings suggest that area V3B/KO plays an important role in the integration of qualitatively and computationally different cues to surface slant.

### *Surface orientation components: slant and tilt*

Previous neurophysiological studies have reported neuronal responses to surface tilt defined by either texture or disparity in both parietal and temporal cortices (Tsutsui et al., 2001; Liu et al., 2004). In contrast to our manipulation of surface slant, these studies manipulated surface tilt by rotating surfaces with a fixed slant within the plane of the screen. The neural selectivity measured in these studies was therefore for the orientation of the depth gradient, rather than the magnitude of that gradient. Behavioural evidence suggests that sensitivities to tilt and slant components of surface orientation are independent of one another (Stevens, 1983; Norman et al., 2006), and neurophysiological evidence also suggests that neural selectivity for these two components may be mostly independent (Liu et al., 2004; Sanada et al., 2012). It is not clear whether the presence of multiple depth cues produces comparable enhancements in perceptual judgment of tilt as for slant. Specifically, tilt selectivity might be less critically dependent on cue integration. Further work is required to examine how cue integration benefits perception of surface orientation involving both tilt and slant components.

### *Surface orientation processing in the dorsal visual pathway*

The current findings complement recent evidence that disparity and relative motion depth cues are integrated in area V3B/KO (Ban et al., 2012). However, the computations involved in extracting depth from motion parallax and disparity cues are qualitatively similar, requiring integration of retinal information across time and space respectively (Rogers and Graham, 1982; Anzai et al., 2001). Taken together, these findings suggest that V3B/KO plays a central role in the integration of depth cues in general, even for qualitatively or computationally quite different sources of information. More importantly, the present findings demonstrate that the same cortical circuits involved in processing



simple depth order relations (Ban et al., 2012) also support estimation of more complex and behaviourally relevant depth structures, such as surface slant.

Our finding that V3B/KO (and to a lesser extent V7) plays an important role in the integration of depth cues to surface slant is consistent with existing evidence that emphasizes dorsal pathway involvement in depth processing. Other dorsal areas in extrastriate cortex that have been implicated in the processing of surface orientation include motion selective areas MT/V5 and MSTd (Xiao et al., 1997; Sugihara et al., 2002; Nguyenkim and DeAngelis, 2003; Sanada et al., 2012), which encode information about surface orientation from both motion and disparity cues. Recent evidence suggests that single MT neurons integrate motion and disparity gradient information, but display comparatively less selectivity for texture (Sanada et al., 2012). Encoding of depth gradients in these areas could support higher level representations of 3D surface orientation that are found further along the dorsal stream in posterior parietal cortex (Shikata et al., 2001; Tsutsui et al., 2002).

Parietal cortex has previously been implicated in processing depth from disparity as well as various monocular cues (Shikata et al., 2001; Sereno et al., 2002; Tsutsui et al., 2002; Georgieva et al., 2008; Durand et al., 2009; Srivastava et al., 2009; Orban, 2011). In the present study, we observed higher classification accuracy for congruent cues compared to quadratic summation in area V7, which may correspond to CIP in the macaque (Tsutsui et al., 2002; although, for more detailed discussions of human-macaque parietal homology, see: Orban et al., 2006a; Shikata et al., 2008). It should be noted that ventral intraparietal sulcus (VIPS) and V7 have been considered to overlap by other groups (Swisher et al., 2007; Georgieva et al., 2009), and they were differentiated here based on different localizer tasks (Orban et al., 1999; Press et al., 2001; Chandrasekaran et al., 2007; Preston et al., 2008). In our study, however, we did not find reliable decoding of activity in intraparietal

areas. One reason for this may have been because participants were not required to attend explicitly to the surface slant, but instead performed an unrelated attentionally demanding task at fixation. Parietal involvement in the processing of 3D surfaces has previously been shown to be strongly modulated by task (Shikata et al., 2001; Chandrasekaran et al., 2007). A further possibility is that surface slant may be sparsely encoded in parietal cortex, thus precluding the possibility of sufficient biases within voxels to support accurate fMRI decoding (Kriegeskorte et al., 2010). Such a finding would be comparable to the low proportion of single cells that were reported to show selectivity for slant from texture and disparity in inferior temporal cortex (Liu et al., 2004).

Ultimately, 3D surface information in the dorsal stream is likely to facilitate visually guided actions (Jeannerod et al., 1995; Sakata et al., 1998; Cohen and Andersen, 2002). Accordingly, various regions of posterior parietal cortex involved in motor planning respond to depth signals, including anterior intraparietal area (AIP) (Taira et al., 2000; Shikata et al., 2001; Durand et al., 2007; Srivastava et al., 2009; Theys et al., 2012) for grasping; the parietal reach region (PRR) (Bhattacharyya et al., 2009); and the lateral intraparietal area (LIP) (Genovesio and Ferraina, 2004; Durand et al., 2007), which is involved in saccadic eye movements. However, responses to visual depth signals in these areas appear to be more concerned with 2D retinal shape than 3D shape. For example, AIP neurons predominantly show selectivity for 3D boundaries rather than 3D surface shape (Theys et al., 2012), suggesting a somewhat more rudimentary representation of 3D surfaces compared to that found in CIP (Sakata et al., 1999; Taira et al., 2001). Such specialization of might enhance efficiency for computing grasping movements.

Similar distinctions have been reported for areas in inferior temporal (IT) cortex in the ventral pathway. Specifically, the lower bank of the superior temporal sulcus (TEs), contains neurons selective for both planar and curved 3D surfaces, whereas neurons in

lateral TE are selective only for 2D shape (Janssen et al., 1999;Janssen et al., 2000;Liu et al., 2004). Recent evidence suggests communication between cortical areas in higher portions of the dorsal and ventral streams, in the form of synchronized activity between anterior parietal and inferior temporal cortical areas (Verhoef et al., 2011), and anatomical connectivity between CIP and TEs has also previously been shown (Baizer et al., 1991). Indeed, given the utility of depth information common to the goals of action and recognition, it is likely that communication between these two processing pathways occurs at earlier stages.

Previous fMRI studies have identified the involvement of visual areas from both the dorsal and ventral pathways in the processing of 3D surfaces from disparity (Chandrasekaran et al., 2007;Durand et al., 2007;Durand et al., 2009), monocular cues (Orban et al., 1999;Shikata et al., 2001;Taira et al., 2001;Serenio et al., 2002;Georgieva et al., 2008), and combinations of the two (Welchman et al., 2005). We noted that sensitivity to congruent cues was statistically higher than for the component cues in area V3v as well as V3B/KO (**Fig. 2.4a**). Although subsequent analysis indicated that V3v does not meet other criteria for cue integration, it is possible that the increased decoding performance (around the level predicted by quadratic summation) is the result of co-located representations of disparity and texture information, as has previously been reported (Janssen et al., 1999;Serenio et al., 2002;Welchman et al., 2005;Georgieva et al., 2008). In contrast to our result, a previous study that measured fMRI adaptation to perspective and disparity cues suggested that both dorsal and ventral extrastriate cortex (specifically, areas hMT+ and LO) represent depth configurations from a combination of these cues (Welchman et al., 2005). This difference might be explained by differences in stimulus configuration: we used randomized Voronoi textures on a single slanted planar surface presented within an aperture, whereas the previous study used regular grids forming a

dihedral angle without the use of an aperture. The stimuli in the previous study therefore had a more object-like structure and appearance (resembling an upright open book), which might increase ventral pathway involvement (Kourtzi & Kanwisher, 2001). Another difference is that the size and shape of individual texture elements were randomized in our stimuli, which reduced the reliability of perspective foreshortening and scaling (Knill, 1998b) compared to those in the previous study. It has previously been reported that both behavioural sensitivity (Rosas et al., 2004) and neural selectivity (Liu et al., 2004) can be dependent on the type of texture pattern presented (e.g. uniform vs. randomized), and might therefore account for differences in results. Finally, although Welchman *et al.* (2005) did not sample from V3B/KO, it is possible that this area provides important inputs to LO that might convey depth information from multiple cues, given the proximity of the two areas.

Previous studies have suggested an important role for higher ventral areas (LO and posterior inferior temporal gyrus) in the processing of 3D shape from disparity (Chandrasekaran et al., 2007; Georgieva et al., 2009). However, there are several critical differences that might explain the relatively low classification accuracy we obtained for the Disparity condition in area LO. First, our stimuli were slanted planar surfaces, while the other studies used more complex 3D shapes that appeared more object-like. Second, our disparity stimuli contained texture cue conflicts, while previous studies used random dot stereograms to minimize the reliability of the texture cue. Finally, previous studies reported that univariate BOLD response in ventral areas corresponded to either shape coherence or depth structure, whereas surface coherence was consistent between different slants in our study.

### *Relation between psychophysical and fMRI results*

Recently, Ban and colleagues (2012) found that training a classifier on fMRI data acquired during the presentation of frontoparallel planes defined by one depth cue (e.g. disparity) produced comparable accuracies whether tested on data from the same cue (disparity) or another cue (motion) in V3B/KO. However, using a similar test, we did not find such cross-cue transfer for texture and disparity cues to surface slant. One possible explanation for the lack of transfer between texture and disparity cue depth representations is that the cue conflicts in our ‘single cue’ Texture and Disparity stimuli were more perceptually salient than those in the stimuli of previous studies. Although the results of our behavioural experiment suggested that observers were sensitive to changes in slant for these stimuli, the Texture condition in particular (which contained a conflicting disparity cue that signaled a frontoparallel surface orientation) may represent a special case for human observers. We frequently witness 2D representations of 3D depth in daily life, such as those projected on a television screen. In such instances we perceive the flat surface of the screen, but recognize that the normally informative disparity signal provides no useful information about the intended depth representation in the image. We are then able to “suspend our disbelief” regarding the depiction of depth for the purpose of interpreting the image (Gibson, 1978; Goldstein, 1987; Koenderink et al., 1992; Vishwanath et al., 2005). Thus while observers showed sensitivity to changes in Texture stimuli when required to make perceptual decisions regarding slant during the behavioural task, changes in slant may have been less perceptually salient for Texture stimuli during the fMRI experiment, in which participants passively viewed stimuli whilst performing an unrelated attention task. Indeed, a previous fMRI study of texture-defined surface orientation reported significantly stronger activation of parietal regions when subjects performed an orientation discrimination task compared to when they performed a comparably attention-demanding

colour discrimination task whilst viewing the same stimuli (Shikata et al., 2001). Top-down processes have been shown to influence choice-related neural activity in sensory cortex (Nienborg and Cumming, 2009), and the absence of such top-down processes in our fMRI experiment might account for the relative lack of slant-related variation in fMRI activity for Texture in the present study.

The pattern of results across conditions observed in V3B/KO (**Fig. 2.5**) is mainly consistent with performance measured at the behavioral level (**Fig. 2.4**). One notable difference is that observers exhibited lower perceptual sensitivity to Incongruent stimuli than for either component cue in the behavioral experiment, whereas classifier performance for the Incongruent condition was either comparable to, or exceeded, performance for the component cues in all cortical regions. This behavioral result is consistent with previous reports of mandatory cue fusion in adult observers (Hillis et al., 2002; Nardini et al., 2010a), and provides evidence against an independence mechanism at the behavioral level, which would predict performance in the Incongruent condition that matches quadratic summation.

A further distinction can be made between two possible forms of fusion based on performance in the Incongruent condition (Clark and Yuille, 1990; Landy et al., 1995). Under strict fusion, depth estimates are combined via weighted averaging using fixed weights, resulting in relative insensitivity to highly incongruent cues. In contrast, robust fusion predicts that one cue (e.g. the less reliable) should be reweighted, and can effectively be ‘vetoed’ (Bülthoff and Mallot, 1988). In this case performance would therefore revert to the same level as the more reliable component cue (Ban et al., 2012). While behavioural performance suggests the former, classifier sensitivity in V3B/KO suggests the latter. One possible explanation for this divergence between psychophysical sensitivity and fMRI discriminability in the Incongruent condition is that observers’

behavioral performance was adversely affected by the use of an excessively large cue conflict ( $90^\circ$ ). This has the potential to produce perceptual bistability (van Ee et al., 2002; Girshick and Banks, 2009), whereby perception alternates between the depth configurations specified by either one of the two cues. For example, when the disparity signal is perceptually dominant, the texture on the surface will appear anisotropic and heterogeneous (Rosenholtz and Malik, 1997; Knill, 2003). Failure to consistently select a single cue (e.g. the more reliable) and veto the other, could lead to the observed decrement in behavioral performance (Girshick and Banks, 2009). As Landy et al. (1995) point out, there is no reason to expect visual performance to be optimized outside of the visual system's normal operating region, such as in the case of our Incongruent stimuli.

Classifier performance for Incongruent stimuli in V3B/KO was slightly above performance for the more reliable of the two single cues (disparity) and slightly below the prediction of quadratic summation (**Fig. 2.5**). This might suggest robust fusion, as was previously reported in area V3B/KO for incongruent motion and disparity stimuli (Ban et al., 2012). However, simulations performed by Ban et al (2012) suggest that it is likely that V3B/KO contains a mixed population of units tuned to both independent and fused cues, whose outputs may be selected based on their relative reliabilities. This idea is consistent with evidence that even in areas further along the visual processing pathways, approximately half of the recorded neurons displayed selectivity for only one cue (Tsutsui et al., 2001; Liu et al., 2004). By averaging over trials, the block design of our fMRI experiment may have aggregated responses to perceptual interpretations of Incongruent stimuli based alternately on independent and fused estimates. Consistent with this idea, previous work suggest that activity in area V3B/KO changes in line with alternating interpretations of ambiguous 3D stimuli (Brouwer et al., 2005; Preston et al., 2009).

## *Conclusions*

Estimation of surface slant is essential for interaction with 3D objects, and may be a critical intermediate stage for processing more complex 3D shapes. By combining information from multiple depth cues, observers can increase the reliability of their slant estimate. Previous studies have suggested a variety of locations within visual cortex where representations of surface orientation from different cues overlap. Here we provide evidence that area V3B/KO not only contains co-located depth representations from texture and disparity cues, but integrates these cues by fusing depth estimates. Importantly, this finding suggests V3B/KO is capable of integrating qualitatively different depth cues, which may be important for the enhancement of depth perception observed for combined cues.



## Chapter 3: The effects of V3B/KO disruption on the integration of texture and disparity cues to surface slant

---

When judging the three-dimensional structure of the environment, observers optimally integrate information from multiple independent visual cues to maximize their sensitivity. Evidence from functional imaging studies in humans suggests that patterns of activity in extrastriate cortical area V3B/KO correlate with the integration of multiple visual cues to 3D depth. However, a fundamental limitation of brain imaging methods is that they can only identify correlations between behaviour and neural activity, and in the case of the blood-oxygenation level dependent (BOLD) signal, such correlations are indirect since they rely on assumptions about neurovascular coupling. Here, we tested for the causal involvement of dorsal (V3B/KO) and ventral (LO) regions of interest in integrating texture and binocular disparity cues to surface slant by temporarily disrupting cortical processing in these regions in healthy observers. Observers discriminated the orientation of slanted surfaces that were defined by texture, disparity, and combinations of these two cues, before and after receiving repetitive transcranial magnetic stimulation (TMS). Specifically we applied continuous theta burst stimulation (TBS) over cortical areas LO and V3B/KO in each hemisphere, plus a vertex control site. The results suggest that unilateral disruption to processing in either visual area is insufficient for eliciting detectable impairment of cue integration. This could reflect parallel/ redundant processing capabilities for cue integration. The methodological limitations of TMS are discussed in relation to interpreting null results.

### 3.1. Introduction

The pair of two-dimensional (2D) retinal images that constitute the visual input to the brain contain multiple sources of information about the three-dimensional structure of the world. For example, when we view a uniformly textured object, deformations of the texture provide perspective information about surface orientation in depth, while differences in the relative position of texture elements in the two retinal images provide binocular disparity information. Once the visual system has extracted these independent depth cues from the sensory input, it must then integrate them in order to form a coherent percept of depth. How does the human visual system accomplish this?

Psychophysical studies demonstrate that human adults have greater sensitivity to depth when multiple cues (e.g. texture and binocular disparity) are available, compared to either cue alone. Further, these two cues are optimally combined by means of weighted averaging based on cue reliability (Knill and Saunders, 2003; Hillis et al., 2004). Recent brain imaging studies have identified area V3B/KO of visual cortex as playing an important role in cue integration for a variety of depth cues (Ban et al., 2012; Dövençioğlu et al., 2013; Murphy et al., 2013). Specifically, these studies used multivoxel pattern analysis (MVPA) of fMRI activity in visual cortex to decode the 3D structure of the visual stimuli presented. All three studies found that the sensitivity of the classifier increased when multiple congruent cues were presented compared to individual cues or multiple incongruent cues, specifically in area V3B/KO. These results suggest that cue combination is performed via a fusion computation in this region.

One fundamental limitation of functional MRI and other neuroimaging methods is that they can only be used to measure correlations between brain activity (or the associated vascular response) and behaviour (Logothetis, 2008). To investigate the causal relationship

between neural activity in V3B/KO and the multiple cue advantage when observers make depth judgements, we therefore sought to disrupt normal cortical function in this area. To achieve this we used fMRI-guided stereotaxic neuronavigation to apply repetitive transcranial magnetic stimulation (rTMS) to regions of interest in visual cortex of healthy human observers. Specifically, we employed a continuous theta burst stimulation (cTBS) protocol for TMS delivery. Under this protocol, repetitive high frequency pulse trains are administered offline (i.e. prior to psychophysical measurement). The effect of cTBS is a localized temporary depression of neural activity in the underlying region of cortex, lasting between 15 to 20 minutes (Huang et al., 2005).

Based on previous fMRI data, we selected two target regions of interest: V3B/KO in dorsal extrastriate visual cortex and lateral occipital complex (LO) in the ventral visual pathway. V3B/KO was selected based on recent evidence from human fMRI data, including data collected using identical stimuli to the ones used in the current experiment (Murphy et al., 2013). Although we did not find a significant improvement in fMRI decoding performance for congruent cue stimuli in LO, several previous imaging studies did find enhanced BOLD activation to multiple depth cues in this area (Kourtzi et al., 2003; Welchman et al., 2005; Georgieva et al., 2008). Further, neurophysiology studies have reported neurons sensitive to surface orientation defined by disparity and texture in both parietal and temporal regions of macaque cortex (Shikata et al., 2001; Tsutsui et al., 2001; Tsutsui et al., 2002; Liu et al., 2004). We were therefore interested to see whether the human ventral visual processing pathway might also be involved in the integration of texture and disparity depth cues.

*Visual areas V3B/KO (AKA V4d-topo) and LO*

Area V3B is a functionally defined region of human visual cortex, located ventro-laterally to V3A, and containing a retinotopic representation of the entire contralateral hemifield (both upper and lower quadrants) (Press et al., 2001). Previous identification of a ‘kinetic occipital’ area overlapping V3B and defined as an area specialized in the processing of kinetic contours (Van Oostende et al., 1997) has been shown to be inaccurate, since this region also responds to shapes defined by color or luminance (Zeki et al., 2003); although see (Gulyas et al., 1994)). More relevantly, KO has been shown to respond to stereoscopic contours (Tyler et al., 2006), while V3B appears to be more selective for stereoscopic contours than kinetic contours. Here we use the term V3B/KO to refer to an area that we define based on both retinotopy (V3B) and responsiveness to motion defined borders (KO). Human V3B/KO does not appear to have a direct homologue in non-human primates. Although human V3B does appear to be the topologue of macaque V4d, it differs in retinotopic organization (containing a full hemifield representation) and does not appear to mirror human V4v (Tootell and Hadjikhani, 2001). This evidence suggests that in humans, V4v and V3B are indeed distinct visual areas rather than retinotopic subdivisions of a single V4.

The lateral occipital complex (LOC) is located ventral and anterior to V3B/KO, and is functionally defined by its stronger response to images of intact objects than scrambled versions of those objects (Kourtzi and Kanwisher, 2001). The LOC is subdivided into the posterior fusiform gyrus (pFs) and LO, which is located at the posterior part of the inferior-temporal sulcus. LO contains two retinotopic representations of the contralateral hemifield, LO1 and LO2 (Larsson and Heeger, 2006), that show sensitivity for orientation and shape respectively (Silson et al., 2013). Although previous fMRI studies of cue integration did not find enhanced sensitivity for combined cues in LOC, several previous

studies have implicated this area in the processing of 3D depth from multiple cues (Kourtzi et al., 2003; Welchman et al., 2005; Georgieva et al., 2008).

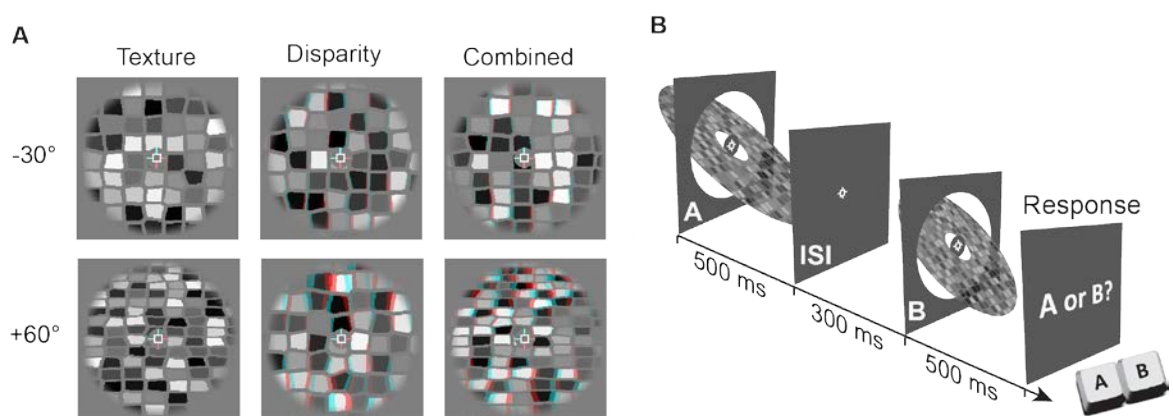
## 3.2. Methods

### *Observers*

Five healthy observers with normal or corrected-to-normal vision participated in the offline TMS experiment. All participants had previously participated in the fMRI experiment (Chapter 2), including functional localizer scans. The mean age was 24.4 years (range = 21-29, 3 female). Participants were screened for stereoacuity on a Wheatstone stereoscope setup in the lab prior to participation, and completed screening forms to check for contraindications for MRI or TMS. All experiments were conducted in accordance with the ethical guidelines of the Declaration of Helsinki and were approved by the University of Birmingham's STEM ethics committee.

### *Stimuli and presentation*

Stimuli were identical to those described in Chapter 2, except that incongruent cue combinations were not presented (**Figure 3.1**). Briefly, stimuli were grayscale stereograms



**Figure 3.1.** Stimulus and task design. A) Examples of the two base slants (-30 and 60) for each of the three cue conditions (texture, disparity, and combined). Stimuli are rendered here as red-cyan anaglyphs for illustration. B) Timeline for a single 2IFC trial. One of the six base slant stimuli shown in A was always presented in one of the two intervals, while the stimulus in the other interval contained the same depth cues but had a different slant angle.

of textured planar surfaces, with slant (rotation about a central horizontal axis) defined independently by texture and disparity (Knill and Saunders, 2003; Nardini et al., 2010b). Surfaces were presented inside a circular aperture with a radius of  $3.5^\circ$ , and a cosine edge profile in order to blur the appearance of depth edges. Stimuli were presented on a mid-grey background, surrounded by a grid of black and white squares (75% density) designed to provide an unambiguous background reference. The fixation marker consisted of a square (length  $0.5^\circ$ ) and nonius lines (length 22 arcmin) bisecting each edge of the square, half of which were presented to each eye in order to promote stable vergence.

In order to prevent observers from relying on absolute disparity information to compare surface slants, the disparity-defined position in depth was randomized for each stimulus presentation (except for stimuli in the texture condition, which contained no disparity-defined slant). This was achieved by randomly shifting the surface relative to the fixation plane ( $0^\circ$  disparity) to between  $\pm 50\%$  of the total surface depth. Similarly, texture-defined position in depth—which corresponds to mean texel size—was randomized for each stimulus presentation by increasing point spacing in the initial grid of points by  $\pm 10\%$ . This reduced the reliability of mean texel size in a given portion of the stimulus as a cue to surface slant.

Stimuli in the single cue conditions ('texture', and 'disparity') were intentionally designed to contain a cue conflict. Specifically, stimuli in the 'texture' condition had a disparity-defined slant of  $0^\circ$ , since all stimuli were presented binocularly. To attenuate the reliability of the disparity cue for this condition, the disparity-defined position in depth of each texel was randomly jittered between  $\pm 4$  arcmin. Similarly, stimuli in the 'disparity' condition had a texture-defined slant of  $0^\circ$ .

Stimuli were programmed and presented in MATLAB (The MathWorks, Inc., Natick, MA) using Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

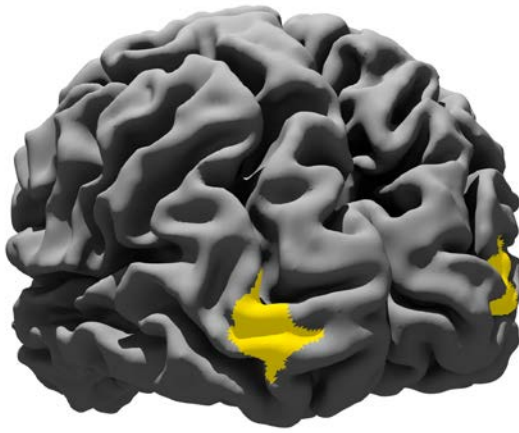
Binocular presentation was achieved using a pair of Samsung 2233RZ LCD monitors (120Hz,  $1280 \times 1024$ ), the technical characteristics of which have been shown to be suitable for vision research (Wang and Nikolić, 2011). The monitors were viewed through mirrors in a Wheatstone stereoscope configuration, at a viewing distance of 51cm and participants' head position was stabilized by an eye mask, head rest and chin rest. Eye movement was recorded binocularly at 1kHz using an EyeLink 1000 (SR Research Ltd., Ontario, Canada) infrared video eye-tracking system. This allowed us to monitor any potential changes in fixation or vergence eye movements during stimulus presentation, including changes that might have occurred as a result of TMS (Kapoula et al., 2001).

#### *fMRI acquisition*

Data were collected at the Birmingham University Imaging Centre using a 3 Tesla Philips Acheiva (Best, The Netherlands) MRI scanner with an 8-channel head coil. Blood oxygen level dependent (BOLD) fMRI data were recorded using a gradient echo EPI sequence [TR=2000ms, TE=34ms, voxel size =  $1.5 \times 1.5 \times 2$  mm, 28 slices covering occipital cortex] for the localizer scans, and a high resolution T1 anatomical scan (voxel size = 1mm isotropic) was also acquired.

Area V3B/KO (Dupont et al., 1997; Zeki et al., 2003) was defined retinotopically based on polar and eccentricity mapping, which was achieved using rotating wedge and expanding concentric ring stimuli respectively (DeYoe et al., 1996; Sereno et al., 1995). V3B was defined as the region of cortex with a full hemifield representation located inferior to, and sharing a foveal representation with V3A (Press et al., 2001; Tyler et al., 2005). This retinotopically defined area overlapped with the set of contiguous voxels that responded significantly more ( $p < 10^{-4}$ ) to kinetic boundaries than transparent motion of a field of black and white dots (Tyler et al., 2006).





**Figure 3.2.** 3D rendering of one subject's cortical surface (left posterior view) reconstructed from MRI with functionally localized area V3B/KO highlighted. Note that the centre of the ROI is located in a sulcus on the lateral surface of the occipital lobe.

Lateral occipital complex (LOC) was identified as the contiguous patch of voxels in lateral occipito-temporal cortex that showed a significantly ( $p < 10^{-4}$ ) greater response to intact versus scrambled images of objects (Kourtzi et al., 2005). LO (a subregion of LOC that extends into the posterior inferotemporal sulcus), was defined based on the overlap of anatomical structures and functional activations, consistent with previous studies (Grill-Spector et al., 2000).

### *fMRI analysis*

Processing of fMRI data from the participants in this study was described in the previous chapter. Briefly, each observer's anatomical volume was transformed into Talairach space (Talairach and Tournaux, 1988) and inflated and flattened surfaces were rendered using BrainVoyager QX™ (BrainInnovation, Maastricht, The Netherlands). Functional data were preprocessed to correct for slice timing and head motion, and high-pass filtered before being aligned with the participant's anatomical scan and transformed into Talairach space. The 1.5x1.5x2 mm voxels of the experimental scans were down-sampled to 2mm isotropic using nearest-neighbour interpolation to minimize spatial smoothing and preserve the fine-scale voxel activity pattern. For TMS sessions, functionally defined ROIs were

exported from BrainVoyager as .vmr files, and imported into BrainSight neuro-navigation software.

## TMS

Biphasic TMS was delivered using a 70mm figure-of-eight coil connected to a Magstim Super Rapid<sup>2</sup> stimulator (Magstim, Withland, UK). Coil position over fMRI-localized regions of interest was guided by stereotaxic neuronavigation using BrainSight software (Rogue Research, Canada). Specifically, participants wore infrared reflective markers on their head, which were tracked by a Polaris Vicra (NDI Medical, Ontario, Canada) optical tracking system with an accuracy of within 0.25mm. Correspondence was established between external anatomical landmarks on the participant's head (inion, nasion, bilateral lower tragus and lateral canthus) and those on a surface mesh reconstruction of observers' heads from their structural T1 MRI. Coregistered ROIs derived from the thresholded functional MRI data were then overlaid. Use of functional imaging data to guide TMS localization has previously been shown to provide greater experimental power, and can therefore permit identification of behavioural effects with smaller sample sizes (Sack et al., 2009). The approximate centroid of each ROI was marked as the target for stimulation, and coil placement was tangential to the scalp, with the coil handle oriented upward. The coil was supported by an adjustable arm, and coil position relative to the cortical target site was monitored and corrected in real-time in order to maintain an accuracy of < 3mm throughout all sessions.

Offline repetitive TMS was delivered as a continuous theta burst stimulation (cTBS) protocol (Huang et al., 2005) consisting of 600 pulses at an intensity of 45% of maximum stimulator output (MSO) for all participants. A fixed output intensity was selected based on evidence that motor and phosphene thresholds (measured at primary

motor, and primary visual cortex, respectively) are not informative about local cortical excitability at other locations, even when scalp-cortex distance is taken into consideration (Stewart et al., 2001; Stokes et al., 2005). The output intensity of 45% MSO was selected based on safety guidelines (Rossi et al., 2009) and previous studies that applied repetitive TMS to adjacent target sites (McKeefry et al., 2008).

### *Experimental procedure*

Participants performed a two interval forced choice (2IFC) discrimination task in which they were sequentially presented a reference stimulus (depicting either +60° or -30° slant) and a test stimulus in a randomized order (**Figure 3.1B**). Each stimulus was presented for 500ms with an inter-stimulus interval of 300ms between the two, and participants were only permitted to give their response after the second stimulus disappeared. Participants were instructed to respond by keypress to indicate the stimulus that contained the least slant. The difference in slant angle between reference and test stimuli was controlled via an adaptive staircase method (Watson and Pelli, 1983), with a maximum limit of  $\pm 20^\circ$ . For each of the two base slants (+60°, -30°) in each of the three conditions (texture, disparity, combined), six interleaved staircases were programmed to converge on the 5<sup>th</sup>, 16<sup>th</sup>, 42<sup>nd</sup>, 68<sup>th</sup>, 84<sup>th</sup> and 95<sup>th</sup> percentiles of the distribution of ‘Test stimulus’ responses. For each observer response, three staircases were updated (either those corresponding to the 5<sup>th</sup>, 16<sup>th</sup> and 46<sup>th</sup> percentiles, or the 64<sup>th</sup>, 84<sup>th</sup> and 95<sup>th</sup>), thus accelerating convergence. Each of the 36 staircases contained 10 trials per block and staircase order was randomized on each of the 10 steps. Each block took approximately 16 minutes to complete.

All participants underwent 6-7 sessions, each on different days. Calibration of the eye tracker was performed immediately prior to the onset of each block in all sessions. The first session was used to train observers on the task and to screen for deficits in perception

of either cue. Observers performed 3 blocks of 396 trials with visual and auditory feedback. In the second session, participants again performed 3 blocks, this time without feedback, and these data served as a measure of baseline performance. Each subsequent session began with observers completing a single block. MRI co-registration was then performed in order to locate the target site (left V3B/KO, right V3B/KO, left LO, right LO or vertex) before cTBS was administered. Following cTBS, participants performed a second block of the behavioural task. The order of stimulation of the five cortical target sites across sessions was randomized between participants. Stimulation of right LO had to be aborted for one participant due to excessive discomfort caused by stimulation of underlying muscle tissue. All other participants reported no physical discomfort.

#### *Behavioural data analysis*

Data for each base slant and each condition were fitted with cumulative Gaussian psychometric functions using a bootstrapping method (Psignifit toolbox; Wichmann and Hill, 2001a;b). Parameters for lapse rate ( $\lambda$ ) and guess rate ( $\gamma$ ) were constrained to the range 0-0.1, while the point of subjective equality ( $\beta$ ) was constrained to equal the base slant.

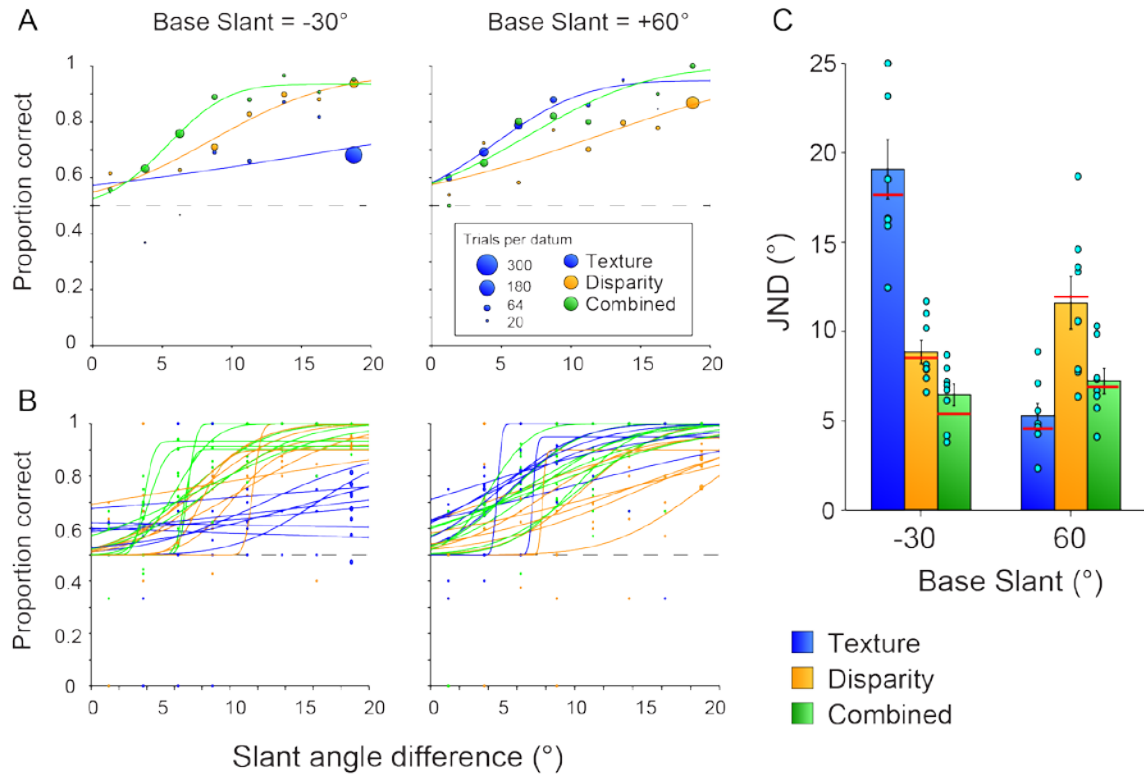
Although observers were not instructed to make speeded responses, we analysed observers' reaction times (RTs) as a potential secondary measure of performance (Luce, 1986). RT was defined as the time between the disappearance of the second of the paired stimuli and the keyboard key press, indicating which of the two stimuli the observer perceived as having less slant. RTs from both correct and incorrect trials were used in the analysis. Since observers were instructed to take their time in responding, a small number of trials showed very long RTs. No lower threshold was applied, but RTs greater than 1s were excluded from analysis since these tended to indicate that observers were withholding

their response in order to take a short break. RTs in the millisecond range indicated that observers had frequently already made a perceptual judgment during the second stimulus interval, and pre-empted the onset of the response period.

## 3.3. Results

### *3.3.1. Baseline performance*

To test observers' baseline sensitivity to surface slant, we pooled data from all blocks in the baseline session and the baseline block (i.e. pre-TMS) from each subsequent session, to obtain a minimum of 2772 data points per observer (462 trials per condition). We then estimated observers' discrimination thresholds (the 'just noticeable difference' or JND) for each condition by fitting cumulative Gaussian psychometric functions to their response data. **Figure 3.3A** illustrates this method with data from one observer, and gives an impression of the reliability of threshold estimates obtained from single blocks (**Figure 3.3B**). Variability in JND estimates based on single blocks was highest for the conditions where thresholds tended to be highest: texture at base slant  $-30^\circ$  and disparity and base slant  $+60^\circ$ . **Figure 3.4** shows baseline thresholds estimated from pooled data and plotted as sensitivities (the reciprocal of the JND) for all five observers.



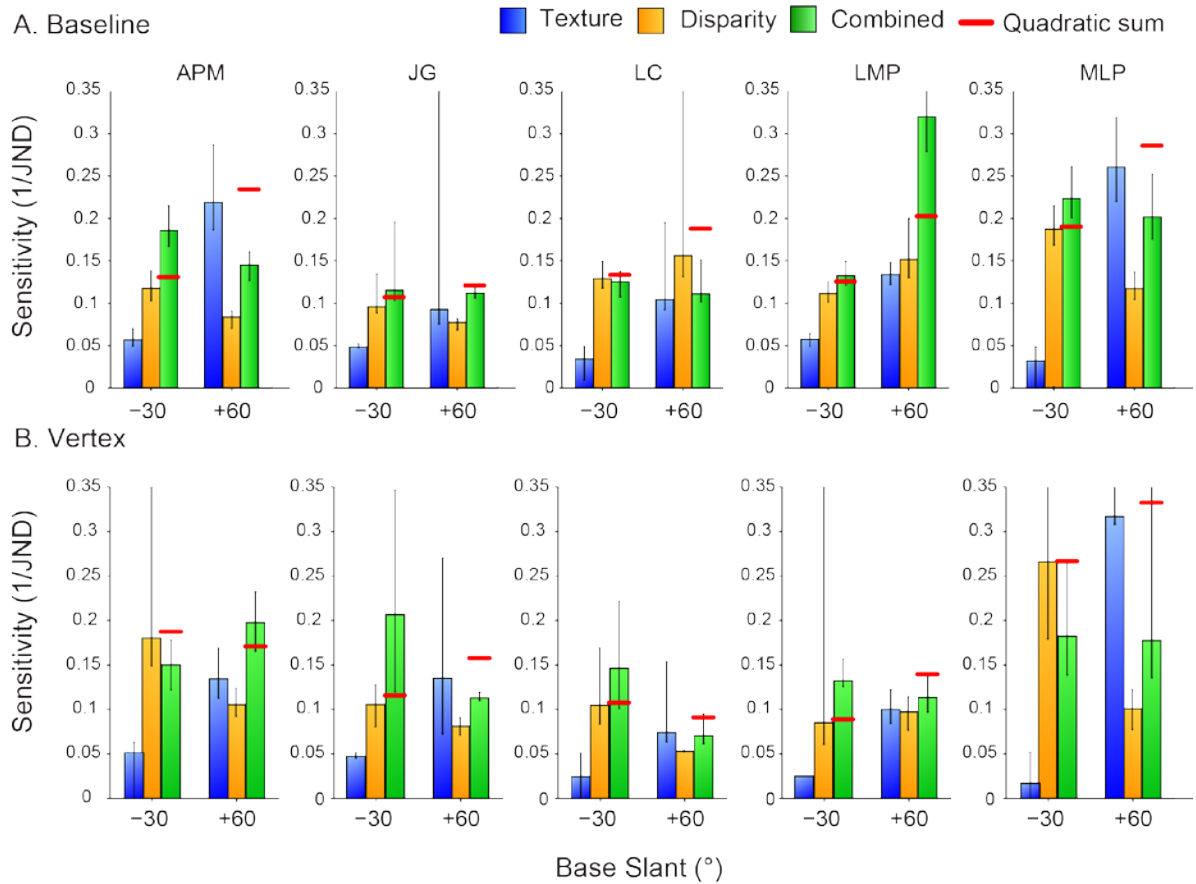
**Figure 3.3. A)** Psychometric functions for one observer and all condition (3 cue types x 2 base slants) at baseline (no TMS). Data from all baseline blocks was first pooled (474 trials per function) and binned (bin width = 2.5°) before bootstrapping to fit a cumulative Gaussian that was used to estimate threshold JND. The diameter of individual data points is proportional to the number of trials in that bin. **B)** The same data with psychometric fitting and threshold estimate performed on each block (54 – 66 trials per function). **C)** Baseline JNDs for all 6 conditions in one observer. Individual data points show the distribution of JND estimates across blocks (B), bars show the means, and error bars show the standard error. Red lines indicate the JNDs derived from pooled data (A).

Sensitivity to the individual depth cues varied across base slants. Specifically, observers showed poor sensitivity to texture information at low (-30°) base slant (mean JND = 46°±17°), resulting in a floor effect since threshold estimates for nearly all observers were outside the tested range (±20°). For the purpose of presentation, we substituted thresholds of 125% of this limit (25°, or  $d' = 0.04$ ), and the data presented for sensitivity to texture at -30° base slant is thus an over-estimate. Sensitivity to texture was significantly higher ( $t(8)=3.5$ ,  $p<0.01$ ) at the larger (+60°) base slant (mean = 13°±4°). In contrast, sensitivity to slant from disparity cues showed the reverse pattern across base

slants, although the reduction in sensitivity at the larger base slant was small ( $-30^\circ$  base slant mean =  $12 \pm 2^\circ$ ;  $60^\circ$  base slant mean =  $15^\circ \pm 1.6^\circ$ ;  $t(8) < 1$ ,  $p > 0.1$ ).

Due to the floor effect for sensitivities to texture at the smaller base slant, the quadratic summation of sensitivities to texture and disparity at this slant (**Figure 3.4**, dotted red line) were not much greater than the sensitivity to disparity alone. Despite this, sensitivities to combined cue stimuli were not significantly greater than those predicted by quadratic summation of individual cue sensitivities across observers at either base slant. The pattern of results was similar for threshold estimates calculated from the block following cTBS to the vertex control site (**Figure 3.4B**).

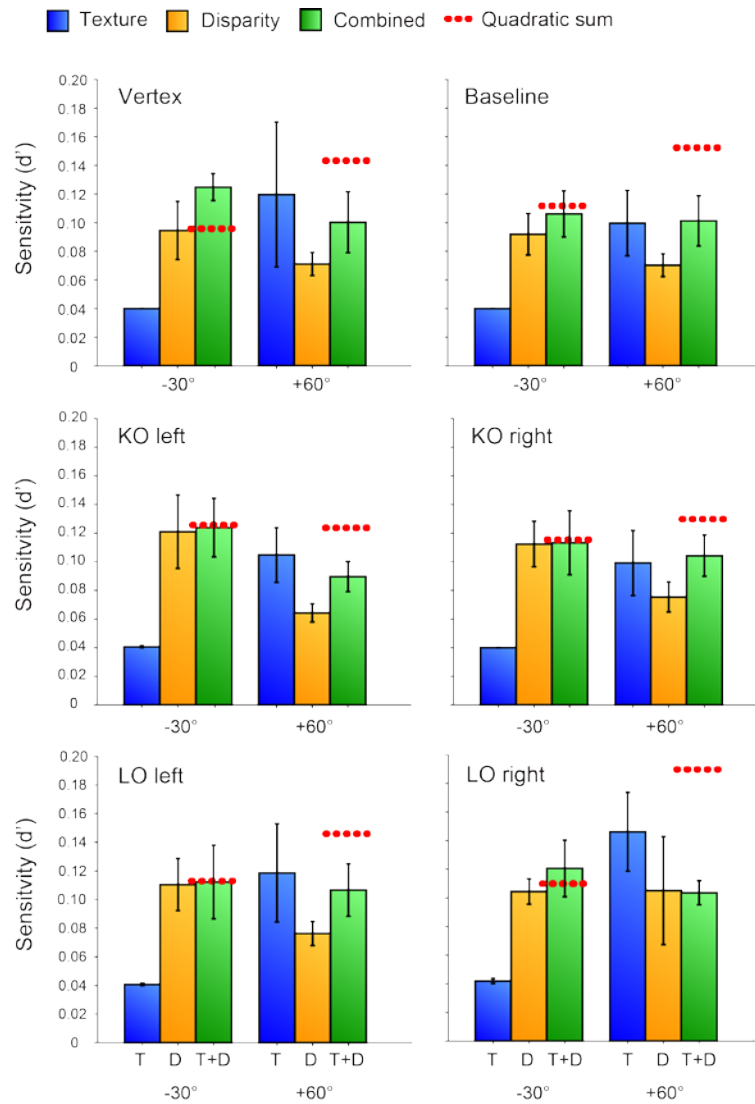




**Figure 3.4.** **A)** Baseline sensitivities to surface slant for all observers. Data was pooled across all baseline (pre-TMS) blocks to obtain a minimum of 262 trials per condition, for each observer. JNDs were calculated for each condition by fitting a psychometric function to slant difference (base slant - test slant) against proportion correct, and sensitivity was calculated as the reciprocal of the JND. **B)** Sensitivities following TMS to vertex control site. Although behavior is not expected to differ from baseline in these data, the estimates are derived from a single block of data (66 trials per condition), leading to increased variability. Error bars show 16% and 84% confidence intervals based on 10,000 bootstrap iterations.

### 3.3.2. Offline TMS effects

Group mean sensitivities at baseline and following TMS to each target site are shown in **Figure 3.5**. The same data are presented again in **Figure 3.6**, but grouped by stimulus condition, with bar colors representing TMS site. Averaging sensitivities across observers produced similar patterns of results across stimulus conditions for the baseline and post-



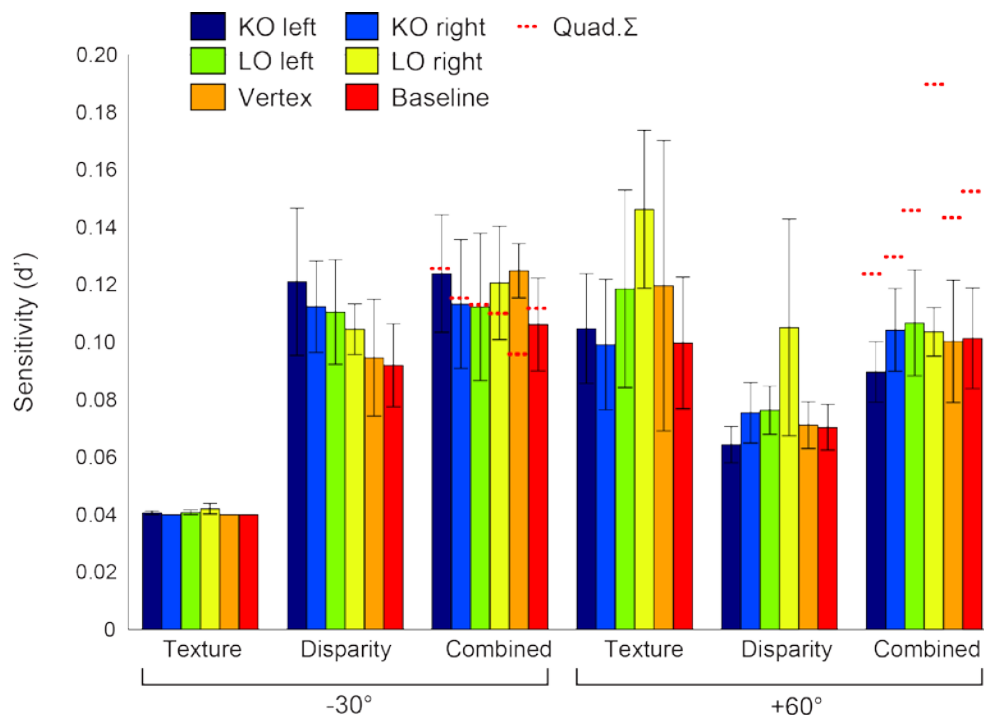
**Figure 3.5.** Group averaged sensitivities to surface slant across experimental conditions. Each plot shows data for the two base slants ( $-30^\circ$  and  $+60^\circ$ ) separately. Dotted red lines show the average predicted sensitivity to combined cues based on quadratic summation of sensitivity to the individual cues. Error bars show standard error from averaging across observers ( $n=5$ ).

vertex TMS conditions (top row). Since vertex TMS is not expected to have a behavioural effect on thresholds for visual judgment of surface slant, any difference between these two conditions provides an indicator of variability in our psychophysical measurements when restricted to a single block of data. The pattern of results for the other 4 post-TMS conditions is generally quite similar to baseline. One difference is that sensitivity to disparity is increased following TMS at all 4 sites, relative to baseline and vertex

stimulation, especially at the lower base slant. A second noticeable difference is that sensitivities to both single cues were higher in the high base slant conditions following TMS of right LO. However, this latter effect in the group averaged data may have resulted from the exclusion of one participant from this condition (due to discomfort during TMS), since her sensitivity for disparity was relatively poor in all other conditions.

### 3.3.3. Reaction times

Although observers were not explicitly instructed to give speeded responses, we considered whether TMS might have affected reaction times differently. Comparison of reaction times (RTs) for all trials in the pre-TMS blocks versus all trials in the post-TMS blocks showed a consistent decrease in RTs, irrespective of condition. Further, RTs in



**Figure 3.6.** Mean slant sensitivities across observers for all stimulus conditions (x-axis) and TMS conditions (bar colors). Error bars show standard error and red dotted lines show predicted combined cue sensitivities based on quadratic summation of sensitivities to individual cues.

baseline blocks for all conditions decreased both within and across sessions, indicating a strong practice effect on RT (Supplementary **Figure S3.1**). Consequently, changes in RT could not clearly be related to effects of TMS.

#### *3.3.4. Eye movements*

Since TMS was applied offline, we did not expect any direct effects on eye movement. However, presentation of binocular disparities can induce reflexive, short-latency changes in vergence eye position (Busetini et al., 1996). We tested version and vergence eye movements averaged across trials and found no clear effect of TMS conditions on eye movement (e.g. Supplementary **Figure S3.2**).

## 3.4. Discussion

We applied an established inhibitory TMS protocol (cTBS) unilaterally to four target sites in visual cortex that have previously been implicated in integrating multiple depth cues: V3B/KO and LO of each hemisphere. We localized the target sites in each individual observer using fMRI data (presented in the previous chapter) and measured sensitivity on a task involving perceptual judgments of surface slant from multiple depth cues. Surprisingly, however, we observed no clear behavioural effects of TMS on observer's thresholds for the conditions tested. Here we consider interpretations of this null result, including limitations of the methods employed.

### *3.4.1. Relation to fMRI studies*

The present study was motivated by previous fMRI evidence that V3B/KO plays an important role in cue integration. Specifically, three studies have shown that the sensitivity of a machine classifier for decoding the 3D structure of congruent cue stimuli from the pattern of BOLD response in this area is higher than would be expected based on quadratic summation of sensitivities to individual cues (Ban et al., 2012; Dövcencioglu et al., 2013; Murphy et al., 2013). While all three studies emphasized the significant result in area V3B/KO, several other ROIs also showed non-significant trends for cue combination, including V3v (for both texture-disparity and motion-disparity combinations), V7 (texture-disparity), and V3A (motion-disparity) and V2 (shading-disparity). In contrast, the results in LO in these studies did not indicate cue integration, although classifier accuracy for individual cues was above chance and close to that predicted by quadratic summation. This suggests representations of depth information from both cues are collocated in LO - a

conclusion supported by evidence from a previous study of univariate responses to perspective and disparity cues (Welchman et al., 2005).

We found a weak effect of TMS to right LO, which *increased* sensitivity to surface slant at high slant angles (60°) when specified by either texture or disparity cues. Inhibition of right LO using repetitive TMS has previously been shown to improve performance on a perceptual task involving the detection of motion-defined contours (Wokke et al., 2014). Further, the absence of an effect for combined cue stimuli following right LO disruption fits with the fMRI evidence suggesting that texture and disparity cues are collocated but not integrated in this region. However, the absence of a similar effect for single cue stimuli at low slant angles is difficult to interpret. In the texture condition at low slant, sensitivity was greater following right LO stimulation than baseline or any other stimulation site, but sensitivity generally remained relatively poor in all conditions. The only single cue condition where right LO stimulation did not enhance sensitivity was disparity at low slant. At low slant angles, observers were more reliant on disparity information for making judgments, and thus there might have been attentional differences between the disparity conditions for high and low slant which could potentially affect LO involvement (Murray and Wojciulik, 2003). Other TMS studies have also reported lateralization of LO disruption effects on perceptual tasks (Mullin and Steeves, 2013; Bona et al., 2014).

A potentially critical difference between the present study and the results of previous fMRI studies that inspired it is that our observers were actively attending to the depth configuration of the stimuli in order to make perceptual judgments, whereas observers in the previous studies performed an orthogonal fixation task during fMRI data collection and were not required to respond to depth information (Ban et al., 2012; Dövcencioğlu et al., 2013; Murphy et al., 2013). In contrast, a previous fMRI study reported enhanced activation in LO when observers were required to make perceptual

judgments of surface slant based on disparity and perspective information (Welchman et al., 2005). Thus, the involvement of LO in depth perception from these cues may be enhanced by actively attending to the stimuli (Murray and Wojciulik, 2003).

One previous study reported disruption of depth perception from binocular disparities following TMS of the superior occipital lobes, 3-4cm above the inion on the midline, which may correspond to V3A (Takayama and Sugishita, 1994). The authors further reported that a stimulus duration of 200ms and stimulation frequency > 10Hz was required to observe any effect. However, in previous unpublished work, we found no effect of offline repetitive 1Hz TMS to KO or LO on the ratio of signal-to-noise required for observers to make perceptual judgments of depth from binocular disparities in random dot stereogram stimuli (Murphy and Welchman, 2009). It is possible that differences in the effect of online TMS (pulses delivered in synchrony with the visual stimulus) compared to offline rTMS (pulses delivered repeatedly, prior to the behavioural task) could account for this. However, there has, to the best of our knowledge, been no further evidence of TMS disruption to visual depth processing in the last 2 decades since that study was published.

#### *3.4.2 Experimental design limitations*

One difficulty for assessing TMS effects on cue integration in the present data is that we did not observe a reliable advantage for combined cue stimuli over individual cues, even in the baseline and control site stimulation conditions, where this might be expected (Knill and Saunders, 2003; Hillis et al., 2004; Nardini et al., 2010b; Murphy et al., 2013). However, in agreement with previous psychophysical findings, our results for the baseline condition showed that sensitivity to combined cues was more likely to exceed the model prediction (quadratic summation of single cue sensitivities) at the  $-30^\circ$  base slant than at  $60^\circ$  (Knill and Saunders, 2003). It is also possible that the short stimulus durations used here (500ms

as compared to 1000ms used in other studies) in order to maximize repetitions, might have reduced observers' accuracies at baseline relative to other studies.

To maintain consistency with the fMRI study reported in chapter 2, we used the same base slant angles:  $-30^\circ$  and  $60^\circ$ . This was desirable because of the known changes in cue weighting across slant angles. However, one trade-off was that we halved the number of trials that we could obtain for each stimulus condition. Further, the presence of a floor effect for the  $-30^\circ$  texture condition (nearly all observers showed thresholds that were outside of the  $\pm 20^\circ$  range tested), meant that 17% of trials were not informative. We attempted to optimize our sampling of observers' psychometric functions by using interleaved adaptive staircases designed to converge at different points along the psychometric function. However, comparing the results obtained at baseline (where data from many blocks could be pooled and the number of trials per condition was close to 500 for each observer), with those collected following TMS to vertex (one block) there are clear changes in patterns of threshold estimates within observers (**Figure 3.4**).

A major constraint of any experiment that is particularly relevant for TMS experiments is the amount of time required to collect sufficient data to reliably estimate a behavioural variable of interest. In the case of offline TMS, as used here, this constraint is specified by the duration of cortical depression following stimulation. Even keeping trial duration to a minimum, we found that observers could only complete approximately 396 trials (66 per condition) within the 20 minute window during which cTBS is reported to reliably maintain cortical inhibition (Huang et al., 2005). This is approximately 20% of the amount of data typically collected per condition for this type of psychophysical experiment (Knill and Saunders, 2003). One solution to this issue is to stimulate the same cortical target across multiple sessions for each participant, conducted on different days (Wokke et al., 2014).



Interestingly, Dövençioğlu and colleagues (2013) previously reported that classifier performance varied not only by ROI but by the psychophysically measured ability of the observer to integrate the shading and disparity cues tested. In contrast, we found no such correlation between psychophysical integration indices and fMRI integration indices in V3B/KO for texture-disparity based on group data from Chapter 2 (Supplementary **Figure S3.3**). Further, we found no clustering of integration indices across observers, and therefore no grounds for categorizing observers based on cue integration performance. Nevertheless, our results clearly show variation between individuals in sensitivity to different depth cues, as well as the ability to integrate depth cues.

#### *3.4.3. Limitations of TMS*

There are many potential explanations for failure to reject the null hypothesis in TMS studies. The most meaningful explanation would be that the areas stimulated are not necessary for the task that was tested. However, other explanations would suggest that the result cannot be meaningfully interpreted, and these arguments fall under three main categories: localization, statistical power, and neural efficacy (De Graaf and Sack, 2011b). With regards to localization, we used the most precise method available by using individual functional localizer data for each observer, overlaid on high-resolution structural MRIs that were stereotactically co-registered to observers. We targeted the center of each ROI and used real-time feedback from the neuronavigation software to minimize the distance between the stimulation coil hotspot and the target during the brief (60 second) stimulation periods. While the number of participants tested in this study was relatively small ( $n = 5$ ), we did not observe any clear trends in the behavioural data, suggesting that this might not be an issue of statistical power. Nevertheless, a larger sample size is –as

always- one potential solution for reducing the problem of measurement noise, mentioned previously, at the group level.

However, one possibility that is difficult to rule out is that our stimulation protocol may have lacked neural efficacy, and there are several possible reasons for this. For example, for safety reasons we selected a conservative stimulation intensity (45% MSO), which therefore would not have penetrated as deeply into cortex as more recent studies that have tended to use higher stimulation intensities (e.g. closer to 70% MSO, (Bolognini et al., 2009)). Since the strength of the magnetic field diminishes as a function of distance from the coil (Ruohonen and Ilmoniemi, 2002), less current is induced in deeper tissue (such as sulci) compared to more superficial areas (e.g. gyri). Further, the TMS coil (which is normally placed tangential to the scalp) generates a magnetic field with lines of flux passing orthogonal to the coil, and this induces electrical current perpendicular to the magnetic field (Barker, 1991). Therefore, current is optimally induced in axons that run parallel to the coil surface, such as those in sulcal walls, rather than gyral crowns or sulcal fundi. In at least one of our subjects (**Figure 3.2**), the center of the V3B/KO ROI was indeed located in a sulcus. Potential solutions to the problem of unknown neural efficacy include measuring neural response after TMS (e.g. using fMRI), or measuring behavioural effects on a second task that the stimulated ROI is known to be necessary for.

While offline repetitive TMS methods have been characterized as inducing cortical inhibition (Huang et al., 2005;Nyffeler et al., 2006;Zafar et al., 2008), the mechanisms of action of TMS remain poorly understood. It should be noted that even positive results obtained using TMS are not without interpretative issues. The effects of stimulation are not locally restricted, and activation (and thus disruption of normal processing) can spread to interconnected regions (Ruff et al., 2009). Thus TMS applied to an area that is not involved in a task can nevertheless disrupt performance on the task if it projects to another

site that is critically involved. This underscores the value of combining TMS with neuroimaging.

#### *3.4.4. Interpretations of a genuine ‘null result’*

Assuming that our TMS protocol did indeed inhibit neural activity in the target ROIs as intended, what does the lack of observable behavioural effect mean? These results suggest that, for each stimulated area, normal processing is not necessary for making perceptual depth judgments from texture or disparity cues. There are several possible reasons for this. First, our ROIs were stimulated unilaterally and stimuli were presented centrally. One possibility is that TMS only affected cue integration in the contralateral visual hemifield, and thus observers were still able to perform the task using visual information from the unaffected hemifield. This is less likely than one might expect since population receptive fields in both LO and V3B (considered part of V3A in some studies) partially cover the ipsilateral visual field (Tootell et al., 1998; Dumoulin and Wandell, 2008), and show biases towards representation of the fovea (Sayres and Grill-Spector, 2008). However, it is possible that a small behavioural effect of TMS to these target areas might be more easily detected by presenting visual stimuli peripherally in one visual field after stimulating the contralateral hemisphere, as used by some previous studies involving TMS of visual cortex (McKeefry et al., 2008; Harvey et al., 2010). Second, both texture and disparity cues are processed in other areas of visual cortex (Tsutsui et al., 2002; Liu et al., 2004; Uka and DeAngelis, 2004; Welchman et al., 2005; Georgieva et al., 2008), including the equivalent areas (V3B/KO and LO) in the contralateral hemisphere. Each stimulated area may therefore reflect redundant processing capability. The pattern of connectivity between areas in the visual hierarchy (based on anatomical evidence) suggests that V3B/KO is not necessarily a critical node for the propagation of visual information along the dorsal visual

pathway, since there are other routes for visual information to pass from primary visual cortex to higher-tier areas (Ungerleider and Desimone, 1986; Felleman and Van Essen, 1991).

It is possible for TMS to affect reaction times on a behavioural task whilst producing no observable effect on error rates (Ellison and Cowey, 2009). Given this possibility, it is unfortunate that our 2IFC task was implemented in such a way that observers were not permitted to give their response immediately after the stimulus appeared in the second interval, but instead required to wait until the second stimulus disappeared (500ms later) before responding. Consequently, the reaction time data may have reflected observers' ability to predict the stimulus disappearance, rather than the difficulty of perceptual judgments, as was evident from the practice effects on reaction times (**Figure S3.2**). However, the effects of offline TMS on RTs can be on the order of milliseconds for some tasks (e.g. saccades (Gerits et al., 2011)).

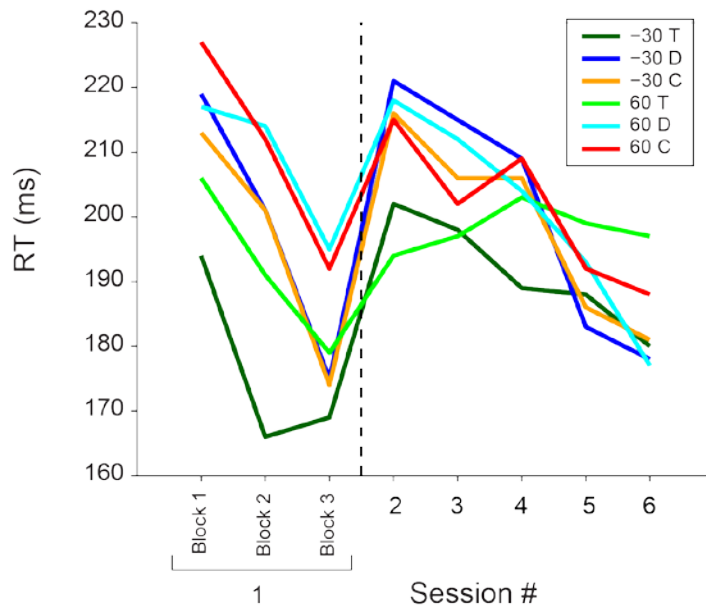
Behavioural evidence indicates similarities between the integration of multiple visual depth cues and integration of multisensory information: the availability of independent sources of information about the environment enhances sensitivity (Ernst and Banks, 2002). TMS of posterior parietal cortex (PPC) has previously been shown to abolish the reaction time advantage normally gained from the availability of information from two senses (vision and touch), without affecting performance for unimodal stimulation (Pasalar et al., 2010). In contrast, rTMS to right inferior parietal lobule (IPL) was shown to increase reaction times for both audio and visual unimodal stimuli, but not crossmodal (Bolognini et al., 2009). While these results indicate that the spatial resolution of TMS is sufficient for establishing such dissociations between multimodal cue integration in association cortex, it remains less clear whether TMS is an appropriate tool for studying visual depth cue integration, or even individual depth cues. The majority of

published TMS studies targeting extrastriate visual cortex have focused on the processing of visual features associated with coarse-scale cortical organization, such as motion, colour, shape, and faces. In contrast, visual depth cues appear to be represented throughout visual cortex in a more diffuse distribution (Orban, 2011), which may complicate their study via TMS. For example, microstimulation of specific functionally-defined subpopulations of MT neurons in macaques has been shown to bias the perception of depth from motion (Krug et al., 2013), yet TMS of the human homolog hMT+ produced only very subtle effects (Brascamp et al., 2010).

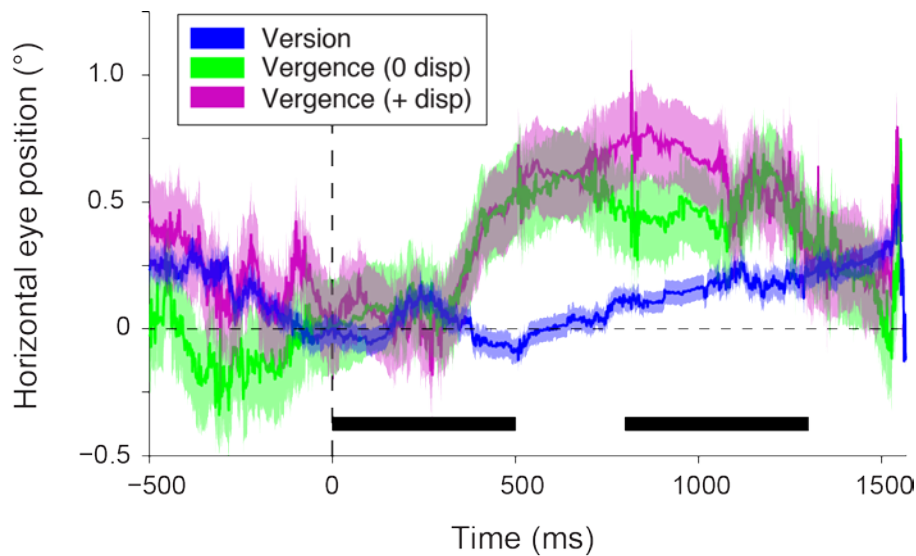
#### *3.4.5. Conclusion*

Precise application of an established inhibitory TMS protocol (cTBS) to functionally localized cortical target sites implicated in integrating multiple depth cues (V3B/KO and LO) produced no reliable changes in observers' sensitivity to surface slant at the fovea, irrespective of cue type (texture, disparity, or a congruent combination). The absence of an observed effect is unlikely to have been due to imprecise localization of cortical targets, although, as in all cases of null results, it is possible that an effect might be observed with greater statistical power (e.g. more subjects). However, an alternative explanation that seems entirely plausible given the lack of reported TMS effects on depth processing to date is that the cortical organization of visual depth information is not amenable to investigation via TMS due to this method's coarse spatial resolution.

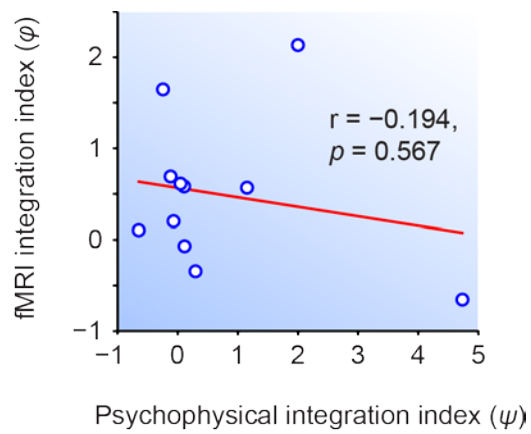
### 3.5. Supplementary materials



**Figure S3.1.** Mean reaction times for each condition in each baseline block in chronological order. RT was measured from the end of the second stimulus interval, which was 500ms duration. The trend for decreasing RTs with increasing session number suggests a practice effect. Differences between stimulus conditions reflect speed vs. accuracy trade-offs. For example, the fastest RTs were for the texture only condition at low base slant (dark green), which all observers were consistently poor at discriminating.



**Figure S3.2.** Example of horizontal eye movements averaged over trials. Green and purple traces indicate vergence eye movement in response to stimuli containing zero disparity-defined slant, and  $+60^\circ$  ( $\pm 20^\circ$ ) disparity-defined slant respectively. There was no clear difference in eye movement responses to these stimuli across TMS conditions.



**Figure S3.3.** Despite variations in integration index across observers in Chapter 2, this measure was not correlated with an equivalent index based on decoding accuracy for patterns of fMRI activity in V3B/KO.

## Chapter 4: Perceptual memory drives learning of retinotopic biases for bistable stimuli.

---

The visual system exploits past experience at multiple timescales to resolve perceptual ambiguity in the retinal image. For example, perception of a bistable stimulus can be biased towards one interpretation over another when preceded by a brief presentation of a disambiguated version of the stimulus (positive priming) or through intermittent presentations of the ambiguous stimulus (stabilization). Similarly, prior presentations of unambiguous stimuli can be used to explicitly “train” a long-lasting association between a percept and a retinal location (perceptual association). These phenomena have typically been regarded as independent processes, with short-term biases attributed to perceptual memory and longer-term biases described as associative learning. Here we tested for interactions between these two forms of experience-dependent perceptual bias and demonstrate that short-term processes strongly influence long-term outcomes. We first demonstrate that the establishment of long-term perceptual contingencies does not require explicit training by unambiguous stimuli, but can arise spontaneously during the periodic presentation of brief, ambiguous stimuli. Using rotating Necker cube stimuli, we observed enduring, retinotopically specific perceptual biases that were expressed from the outset and remained stable for up to forty minutes, consistent with the known phenomenon of perceptual stabilization. Further, bias was undiminished after a break period of five minutes, but was readily reset by interposed periods of continuous, as opposed to periodic, ambiguous presentation. Taken together, the results demonstrate that perceptual biases can arise naturally and may principally reflect the brain’s tendency to favor recent perceptual interpretation at a given retinal location. Further, they suggest that an association between retinal location and perceptual state, rather than a physical stimulus, is sufficient to generate long-term biases in perceptual organization.



## 4.1. Introduction

Our perception can be shaped by past sensory experiences, recent or removed in time. In vision, phenomena such as adaptation and perceptual learning illustrate that sensory experience can affect perception over different timescales (Seitz and Watanabe, 2005; Kohn, 2007). However, sensory signals are often ambiguous, and the brain must constructively process them to achieve a coherent perceptual interpretation of the environment. Bistable stimuli provide a means of dissociating sensory stimulation from perceptual experience, as there are (at least) two valid interpretations of the same sensory input (Blake and Logothetis, 2002; Sterzer et al., 2009). Using such stimuli it has been demonstrated that past perceptual experience over a range of timescales can strongly influence subsequent perception.

At short timescales, perception of bistable stimuli can be biased either towards or away from that of a recently presented unambiguous version of the stimulus, depending on the duration of the inter-stimulus interval (Nawrot and Blake, 1993; Kanai and Verstraten, 2005; Long and Moran, 2007). Similarly, brief intermittent presentations of ambiguous stimuli cause observers to repeatedly experience the same percept on consecutive presentations (Orbach et al., 1963; Leopold et al., 2002; Maier et al., 2003; Brascamp et al., 2009). This perceptual stabilization phenomenon is attributed to a putative short-term perceptual memory trace that accumulates over seconds and can last for tens of minutes (Brascamp et al., 2008; Pastukhov and Braun, 2008; Pearson and Brascamp, 2008; de Jong et al., 2012b). Over extended periods of intermittent presentation, perception alternates between phases of stability for each percept, at a rate that is inversely proportional to the interval between consecutive presentations (Brascamp et al., 2009).

At longer timescales, the resolution of bistable stimuli can be biased in a context-contingent manner through training. In a recently developed paradigm referred to as “cue recruitment”, context-dependent biases that last 24 hours and are resistant to counter-training can be elicited by initial exposure to a mixture of ambiguous and unambiguous versions of the stimulus (Haijiang et al., 2006; Harrison and Backus, 2010a;b; van Dam and Ernst, 2010; Harrison et al., 2011). Such learning has previously been described as a form of Pavlovian conditioning (Haijiang et al., 2006). In contrast to conventional notions of associative learning however, the resolution of perceptual ambiguity appears to be more important than unambiguous sensory stimulation for learning to occur (Harrison and Backus, 2010; van Dam & Ernst, 2010). Learning therefore appears to be driven by past perceptual experiences—internally generated interpretations of ambiguous sensory input. Here we refer to this type of long-term bias as perceptual association.

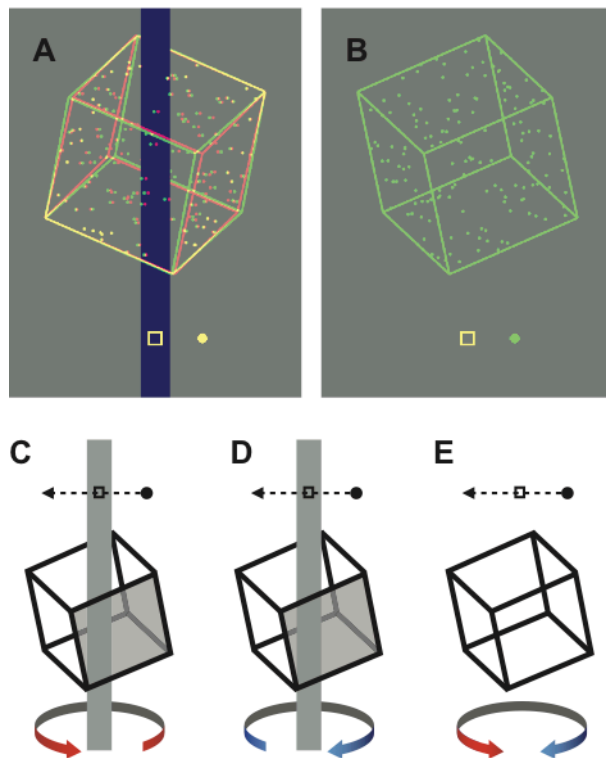
Experience-dependent changes in perception occurring at different timescales have typically been studied separately and treated as independent phenomena. Consequently the relationship between short term perceptual memory and long term perceptual associations remains poorly understood. For example, intermittent presentation of ambiguous stimuli during the training phase of perceptual association learning has been assumed to induce stabilization, but unambiguous stimuli are also typically interleaved. It therefore remains unclear to what extent the learned association between percept and retinal-location is driven by traditional associative learning mechanisms (reliant on unambiguous information) compared to learning guided by perceptual memory. Conversely, stabilization studies to date have not investigated the long-term effects of stabilization on perceptual bias.

Here we tested the hypothesis that long-term associative learning can emerge naturally from repeated instances of short-term perceptual memory, in the absence of

explicit training. First we assessed the relative contribution of these two distinct processes during an extended initial ‘training’ period of intermittent exposure to ambiguous stimuli, that is known to induce long-term perceptual associations. In the second experiment, we examined the ability of perceptual memory to reconfigure a recently learned association. The results demonstrate an important role for perceptual memory in both initiating and updating retinal location-contingent learning of perceptual biases, thus drawing a strong link between these two apparently different forms of visual plasticity. Together, these findings suggest that certain forms of long-term perceptual learning might derive from short-term mnemonic, rather than strictly associative, mechanisms. Specifically, repeated experience of a subjective perceptual state, as opposed to a physical stimulus, can give rise to an enduring bias in the subsequent interpretation of ambiguous patterns.

## 4.2. Results

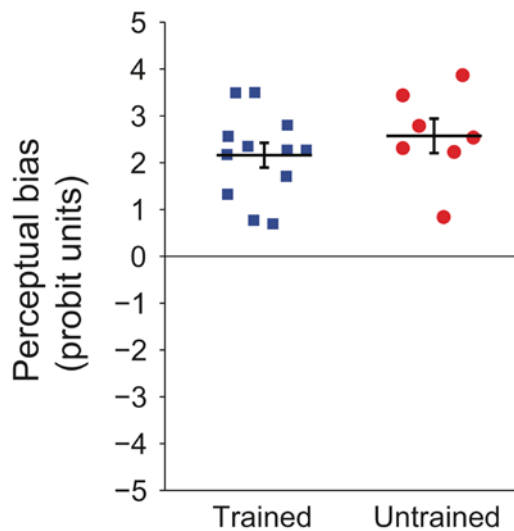
Our experiments investigated the influence of brief visual stimulation with ambiguous and unambiguous stimuli on the establishment and maintenance of perceptual biases. In the first experiment, we investigated the influence of unambiguous stimuli in biasing subsequent perception of rotating Necker cube stimuli. The unambiguous stimuli were similar to the ambiguous versions but incorporated disparity and occlusion cues, which prompted the perception of rotation in one or the other direction (Figure 4.1A). We refer to the disambiguated Necker cube stimuli as *training* stimuli, as they have previously been used to train long-term biases in perception (Harrison and Backus, 2010b; van Dam and Ernst, 2010). In an initial pilot experiment, we first replicated this long-term learning effect (Supplementary Figure S4.1). The results confirmed that the training regimen used in the subsequent experiments was capable of inducing strong long-term learning that persisted over 24 hours and was resistant to counter conditioning.



**Figure 4.1.** Cropped screen shots of the stimuli, rendered here for red-green anaglyph presentation. (A) Unambiguous stimulus disambiguated by occlusion and disparity cues. (B) Ambiguous stimulus presented monocularly. (C,D,E) Schematic representations of the task. Participants reported whether the front face of the cube (shaded in the unambiguous examples) moved in (D) the same direction as the probe dot or (C) the opposite direction to the probe dot. (E) On ambiguous trials the task was the same, but there was no correct answer.

### *Effects of training stimuli on perceptual bias*

In the first experiment, participants viewed a succession of briefly presented ambiguously rotating Necker cubes over a block lasting approximately 20 minutes. Each block consisted of 400 trials, with each trial initiated by the participant with a keypress. This block length was nearly four times longer than those typically used in “cue recruitment” paradigms (Supplementary Figure 1; Harrison and Backus, 2010b; van Dam and Ernst, 2010) and was selected to allow us to look for periodic perceptual dominance phases associated with perceptual stabilization (Leopold et al., 2002; Brascamp et al., 2009). On each trial the stimulus was presented either above or below the fixation point for 1.5 seconds. Following each presentation, the participants indicated the perceived rotation direction relative to a probe dot with a button press (see Figure 4.1 and Methods).



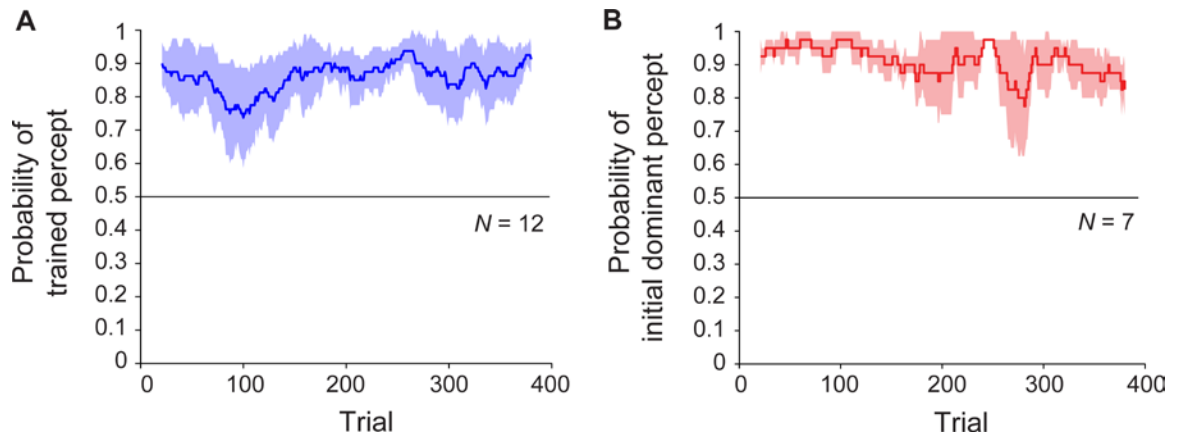
**Figure 4.2.** Individual probit transformed bias summed across retinal locations for participants in the first experiment. Black lines indicate group means and error bars show standard error.

Participants were assigned to one of two conditions. In a *trained* condition ( $n = 12$ ), the first two trials at each retinal location were unambiguous (training) stimuli. The training stimuli always rotated in opposite directions at the two locations, and were counter balanced across participants. These training trials were then followed by 396 ambiguous trials. Previous studies have suggested that initial association of unambiguous rotation

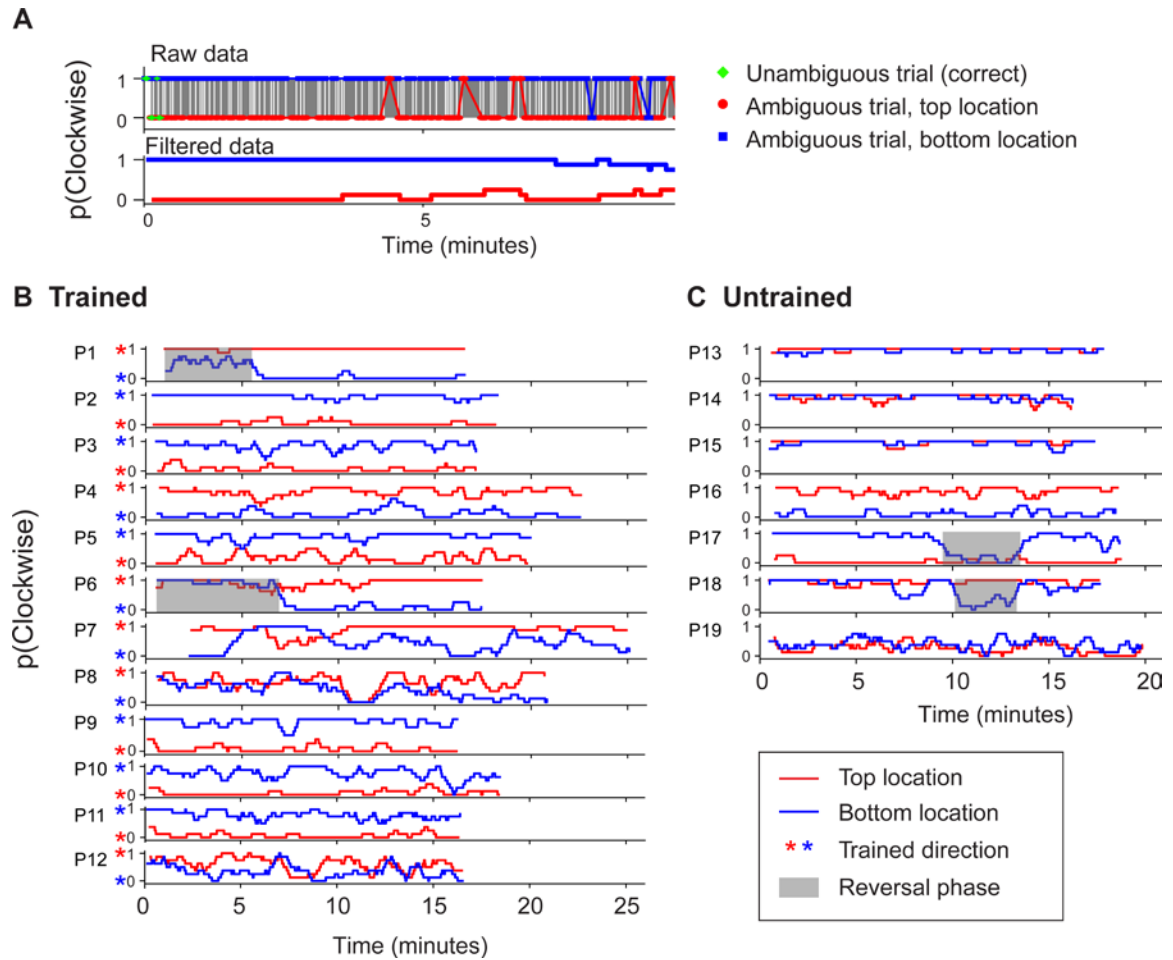
direction with a retinal position leads to long-lasting, position-specific biases in perception (Haijiang et al., 2006; Harrison and Backus, 2010b; van Dam and Ernst, 2010). To test the importance of the initial training stimuli, we also tested participants in an *untrained* condition ( $n = 7$ ), in which the structure of the block was identical except for the absence of any training stimuli.

Prominent perceptual biases emerged in both trained and untrained groups. Figure 4.2 shows the magnitude of this bias for each of the participants. The perceptual bias is expressed as the sum of probit transformed probabilities at each retinal location (probit units, see Methods), per earlier convention (Harrison and Backus, 2010b). For the trained group, the positive values indicate that the majority of trials were seen as moving in the trained direction at each location. For the untrained group, positive values indicate that the initial direction of bias was maintained throughout the session. The results show that the explicit training trials had no significant effect on the magnitude of the bias [ $t(17) < 1$ ,  $p = 0.37$ ].

We next considered whether, given the initial training trials, there might be a difference in the rate of buildup or decay of the perceptual bias, as an observable effect of associative learning. To address this, we examined the time courses of sessions with and without training. Specifically, we asked whether the perceptual bias emerged gradually over the course of the session, and whether or not this change was shared across the trained and untrained conditions. The results are shown in Figure 4.3, which plots the probability of the trained (or initially dominant) percept as a function of trial number over the 20-minute block. For the trained group, the bias in the trained direction was present from the first ambiguous trial and persisted until the end of the session (Figure 4.3A). Analysis using robust regression demonstrated no significant increase or decrease in this variable over time (trained group:  $t(11) = 1.51$ ,  $p = 0.16$ ; untrained group:  $t(6) < 1$ ,  $p = 0.40$ ). Despite the absence of any disambiguated trials, the results for the untrained sessions were similar (Figure 4.3B). Note that in this case, the probability is computed relative to the initially



**Figure 4.3.** Time course of median perceptual bias across subjects throughout the block for observers in (A) trained ( $N = 12$ ) and (B) untrained ( $N = 7$ ) groups. Coloured lines represent group median proportion of trials perceived in the ‘trained’ direction for a sliding window of 40 trials, and shaded regions indicate median absolute deviation.



**Figure 4.** (A) An example to illustrate how raw perceptual report data from individual observers was boxcar filtered (with a window size of 8 trials) to enhance clarity in (B) and (C). (B) Perceptual report data from observers in the trained group in experiment 1 for two consecutive blocks of 400 trials each. Red asterisks indicate the percept specified by initial disambiguated training trials at the top location. (C) Data from observers in the untrained group. Shaded regions indicate phases of stabilized perceptual reversal ( $> 3$  minutes). Such periodic alternations are characteristic of perceptual stabilization (Brascamp et al., 2009).

dominant percept rather the direction of a particular training trial. For a subset of participants ( $n = 8$ ) who were tested on a second block of 400 ambiguous trials immediately after completion of the first block, bias remained similarly stable throughout sessions lasting approximately 40 minutes (Supplementary Figure S4.2).

It has been suggested that group averaged data may be a poor measure of rate of learning (Gallistel et al., 2004). However, in agreement with our group averaged data, stable biases could be observed in the data of individual participants from both groups










(Figure 4.4). For each participant, a boxcar-filtered time course of perceived direction is shown separately for the two retinotopic positions (red and blue traces). Figure 4.4B demonstrates that for the majority of participants in the trained group, training was effective for the upper and lower positions (red and blue stars, respectively). Thus two unambiguous trials at each location were generally sufficient to induce opposite-direction biases above and below the fixation point. For the untrained group (Figure 4.4C), biases were established early in the block and were most often in the *same* direction in the different positions. While established biases remained stable for most participants, we observed occasional phases of reversed dominance (gray highlights), which are a hallmark of perceptual stabilization (Brascamp et al., 2009). The occurrence of such reversals for several individuals contributed to the increased variability seen at certain time points in the group averaged data (Figure 4.3). Note that the observers in the trained group who experienced such reversals began the session reporting the untrained percept, and their subsequent reversal to the trained percept is unlikely to reflect a delayed effect of training.

These results reveal that while explicit training can determine the direction of perceptual biases, similar biases emerge in the absence of training. These spontaneous biases are of similar magnitude and stability to the trained biases and their direction can be the same or opposite at different retinal positions. The results are consistent with the suggestion that learning under the “cue recruitment” paradigm results primarily from associations formed between *percepts* and retinal locations, rather than between the physical stimuli and retinal locations, which might be predicted based on conventional theories of associative or Pavlovian learning (Harrison and Backus, 2010b; van Dam and Ernst, 2010). From this we conclude that the formation of such long-term perceptual associations ultimately depends on perceptual memory holding the percept constant over many presentations of the physically ambiguous stimulus.

### *Effects of spontaneous alternation on perceptual bias*

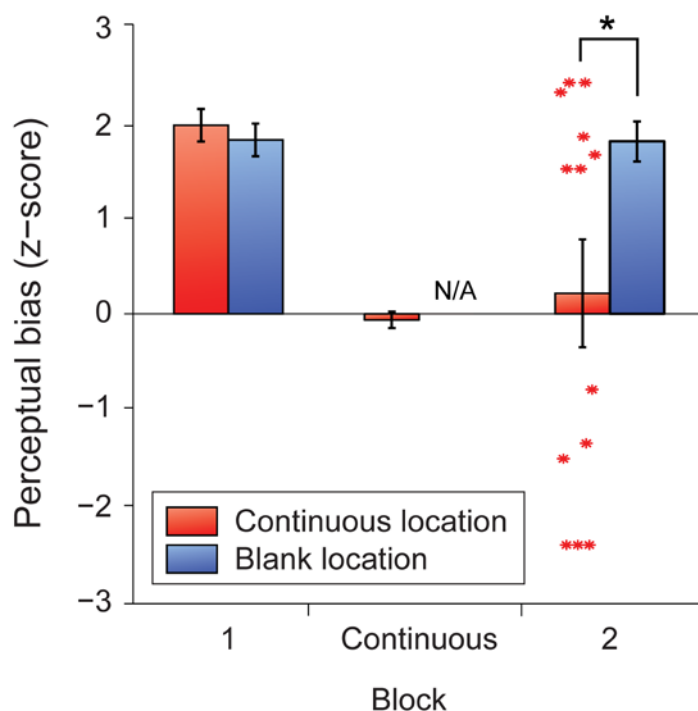
In the experiments described above, we used a presentation paradigm that is known from previous studies as well as our own pilot experiment, to induce long-term retinal location-contingent biases (Harrison and Backus, 2010b; van Dam and Ernst, 2010). The results suggested that perceptual stabilization (achieved through brief intermittent stimulus presentations) plays an important role in establishing such perceptual associations. Given these results, we hypothesized that perceptual alternations, which arise spontaneously during an extended period of continuous stimulus presentation, might therefore have a deleterious effect on learned biases. We thus tested whether periods of spontaneous alternation would affect recently acquired perceptual biases, and further whether they would do so in a retinotopically specific manner.

We modified the original paradigm by adding an extended period of ambiguous stimulation. Specifically we interposed five minutes of continuous rotating Necker cube presentation at one of the two retinal locations into a training block (Figure 4.5). During this period of continuous presentation, the opposite retinal location remained blank. A blank interval of this duration has previously been considered sufficient to demonstrate ‘long term’ learning effects (van Dam & Ernst, 2010). Participants reported their perceptual state for both the continuous and intermittent portions of these extended blocks. In each testing session, participants were initially biased in opposing directions for the two retinal locations, as described above. Initial analysis was restricted to those participants for whom training was effective at both locations during the first block ( $n = 13$ ).

	Block 1	Continuous interval	Block 2
Top location	 x 2  x 58	 x 5 minutes	 x 60
Bottom location	 x 2  x 58	Blank	 x 60

**Figure 4.5.** Illustration of experiment 2 design. All sessions began with a block of 120 trials, of which 4 out of the first 8 (2 at each retinal location) were unambiguous. Observers were then presented with an ambiguous stimulus at either the top or bottom retinal location (counterbalanced across subjects) continuously for 5 minutes, during which they reported perceptual alternations in the perceived direction of stimulus rotation. At the end of the continuous interval, observers completed a final block consisting of 120 ambiguous trials.

We found that the five-minute period of continuous perceptual alternation effectively abolished the trained perceptual biases at the continuously stimulated location, but not at the blank location (Figure 4.6). Subsequently, participants expressed significant biases that were nearly equally divided between the two different directions at the continuously stimulated location (Figure 4.6, red asterisks). By contrast, the unstimulated



**Figure 4.6.** Mean z-scores indicating perceptual bias in the trained direction for each retinal location for the full learners group ( $N = 13$ ). Error bars represent *SEM*. Red asterisks represent z-scores for individual participants in block 2 at the retinal location where the stimulus had previously been presented continuously. The black asterisk indicates a significant difference between means ( $p < 0.05$ ).

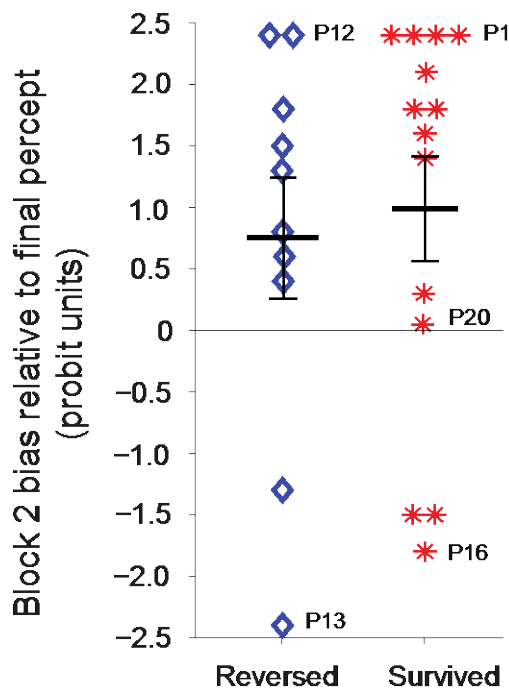
retinotopic location maintained its previous bias and appeared entirely unaffected by the spontaneous alternations at the opposite location.

What factors determine the perceived direction of rotation following periods of continuous ambiguous stimulation? In the previous experiment we observed strong, stable perceptual biases from the start of the session, irrespective of whether explicit training was provided. This suggests that the perceptual interpretation of early ambiguous trials at each location determines which percept an observer becomes biased toward, since the dominant percept at the start of a session tends to become stabilized throughout. We reasoned that perhaps the last direction of rotation perceived during continuous viewing might be important. For example, the final percept experienced by an observer before the stimulus was turned off, might become stabilized during the next intermittent sequence of ambiguous trials. If this manipulation influences subsequent perceptual bias, then this would provide further evidence for the role of short-term perceptual history in long-term bias acquisition, since no disambiguating cues of any sort were applied during this phase of the experiment.

To test the dependence of the bias in the final block on the spontaneous alternation process, we related the bias direction to the ultimate perceptual state reported by all participants during the continuous presentation period. This analysis revealed that the bias in a significant majority of participants (19 of 24) matched the final perceptual state ( $p < 0.05$ ). To test whether more distant aspects of perceptual history influenced subsequent bias, we coded final percept duration as a vector, with both magnitude (seconds) and direction relative to training (same or opposite to trained bias). This measure correlated significantly with block 2 bias ( $R^2 = 0.17$ ,  $p < 0.05$ ). In contrast, we found no correlation between block 2 bias and a variety of other measures of longer-term perceptual history

during the continuous presentation period (including overall bias, total perceptual alternations, frequency of perceptual alternations and bias during the final 30 seconds).

To test whether the tendency for block 2 bias to follow the last percept experienced during continuous viewing was related to the direction of the ‘trained’ perceptual bias prior to continuous stimulation, we divided participants into two groups depending on whether the percept that dominated at the continuous location during block 2 was the same as, or opposite to, the percept that had dominated during block 1, regardless of training (Figure 4.7). We found that for both groups the last continuous percept was the principal predictor of the block 2 bias direction, and there was no significant effect of prior bias direction on the magnitude of subsequent bias.



**Figure 4.7.** Perceptual bias at the continuous location during block 2, for all participants tested ( $N = 24$ ). Participants were divided into two groups depending on whether the percept that dominated at the continuous location during block 2 was the same as (red asterisks), or opposite to (blue diamonds), the percept that had dominated at that location during block 1. Labeled data points indicate the corresponding individual observers’ data. Time courses for all individual observers are shown in **Figure S3**. For the majority of observers the last percept experienced during continuous viewing correctly predicted subsequent perceptual bias at the continuous location.

### 4.3. Discussion

Perceptual resolution of ambiguous sensory input can be influenced by both recent and distant past experiences. Perceptual stabilization (Orbach et al., 1963; Leopold et al., 2002; Brascamp et al., 2008; Brascamp et al., 2009; Knapen et al., 2009a) and “cue recruitment” (Haijiang et al., 2006; Harrison and Backus, 2010a;b; van Dam and Ernst, 2010; Harrison et al., 2011) are two such examples, putatively driven by short-term perceptual memory and associative learning processes respectively. Here we tested the hypothesis that perceptual memory, rather than strictly stimulus-based associative learning, is primarily responsible for driving the learning of retinal location-dependent perceptual biases.

We assessed the stability of perceptual bias over sessions lasting up to 40 minutes and observed no reliable increase in bias that might reflect a gradual learning component. Instead, biases were expressed from the outset and remained relatively stable throughout, with occasional phase reversals of perceptual dominance in individual participants. This phenomenon is characteristic of perceptual stabilization (Brascamp et al., 2009), and cannot be explained by associative learning. Further, we compared perceptual bias between a group who received unambiguous training trials, and a group who received no training. We found no effect of training on the strength or stability of perceptual bias, demonstrating that such biases can emerge exclusively as a result of perceptual stabilization.

The absence of gradual changes in bias might reflect a ceiling effect, since observers in both trained and untrained groups began each session with strong, immediate biases. This result does not therefore rule out the involvement of gradual learning during training. Our results and those of previous studies demonstrate that the training paradigm used here reliably results in the learning of long-term perceptual associations, regardless of

whether participants are tested 5 minutes (Figure 4.6; van Dam and Ernst, 2010) or 24 hours (Supplementary Figure S4.1; Harrison and Backus, 2010b) after training. However, our results suggest that perception is dominated by short-term processes during training, rendering any gradual learning effects unobservable.

For observers in the untrained group, perception of the first stimulus at each location necessarily reflected ‘pre-existing bias’. In contrast, for the trained group, the trained percept was likely to have conflicted with observers’ pre-existing biases in approximately half of all instances. (This conflict perhaps explains the occurrence of ‘partial learners’, for whom training failed at one of the two locations). The finding that trained and untrained observers displayed similar levels of perceptual bias therefore suggests that the relative contribution of pre-existing bias throughout the session was negligible compared to the strong effect of stabilization.

Trained biases were shown to persist following a five minute blank interval, which has previously been considered sufficient for demonstration of ‘long term’ learning, as opposed to stabilization (van Dam and Ernst, 2010). Other studies have considered even shorter inter-stimulus intervals to be sufficient to diminish the contribution of stabilization (Carter and Cavanagh, 2007). Although the exact rate of decay of the short-term perceptual memory trace, which promotes stabilization, is not known, existing models predict decay to baseline within this time frame (Noest et al., 2007; Brascamp et al., 2009), and are supported by physiological evidence for multi-timescale neural adaptation (Fairhall et al., 2001). Thus the processes of short-term perceptual memory and long-term perceptual association are behaviourally distinguishable by their temporal properties. As such, the persistence of trained biases following an extended blank interval must primarily reflect long-term learning, since any contribution of residual perceptual memory will be significantly diminished over this time frame.

Following a period of continuous viewing, during which perception spontaneously alternated, observers were equally likely to exhibit a perceptual bias in the same or opposite direction to that specified by previous ‘training’. One way in which continuous viewing might produce this result is by causing observers to experience both percepts in approximately equal measure through frequent perceptual alternations. This period of unbiased perceptual history could serve to ‘reset’ long-term perceptual bias, resulting in an equal likelihood of either percept dominating subsequent intermittent viewing. Another factor that could explain this result is that subsequent perception is driven by short-term factors, such as the most recent percept. We tested the relative contributions of these two factors to the observed result by comparing measures of both short term and long term perceptual history during continuous viewing with subsequent perceptual bias. We found that the last dominant percept during continuous presentation predicted subsequent perceptual bias, including its direction and to some extent its duration. This result is consistent with several previous findings regarding the probability of a previously stabilized percept regaining dominance following a short period of continuous presentation. That survival probability appears to tend towards chance as a function of the number of spontaneous alternations, and as a function of the duration of dominance of the opposing percept (Brascamp et al., 2008).

Previous findings further suggest that perceptual history preceding the last dominant percept of a presentation period influences perception at the next presentation (Brascamp et al., 2008; Pastukhov and Braun, 2008). However, we found no correlation between block 2 bias and a variety of measures of longer-term perceptual history during the continuous presentation period (including total alternations, total bias, and bias during the final 30 seconds). This discrepancy is likely related to the longer duration of continuous presentation used here, which resulted in a greater number of perceptual



alternations. Thus, by reducing the survival probability of the previously dominant percept towards chance ('resetting'), continuous presentation increases the influence of the final perceptual state on subsequent perception. Our results therefore suggest that the effect of perceptual alternations on a recently learned perceptual association is similar to its effects on perceptual memory. Ultimately, the brain favours repeated selection of the most recently experienced percept when resolving perceptual ambiguity following a period of unstable perceptual history. This emphasizes the importance of short-term priming-like effects in determining perception, even when these conflict with recent learning.

It is important to note that the transfer of perceptual biases was asymmetrical, with those established during continuous viewing affecting perception during intermittent presentation, but not vice-versa. This result agrees with previous findings, which suggest that distinct neural mechanisms underlie stochastic perceptual alternation (under continuous viewing conditions) and initial percept selection at stimulus onset (under intermittent viewing conditions) (Stanley et al., 2011; de Jong et al., 2012a). For example, individuals' perceptual alternation rate during continuous viewing is not correlated with their alternation rate during intermittent viewing (Brascamp et al., 2009). Further, pre-existing (i.e. untrained) long-term perceptual bias is observed during intermittent viewing but not continuous viewing of binocular rivalry stimuli (Carter and Cavanagh, 2007). Nevertheless, our results demonstrate that these two processes do interact, with the transfer from continuous to intermittent viewing being strong enough to disrupt a recently acquired bias.

### *Classical conditioning versus perceptual association of location-contingent biases*

Experience-dependent changes in perception can occur at multiple timescales, yet they have traditionally been studied independently and considered as separate phenomena. The learning of long-lasting retinal location-contingent biases has previously been referred to as “cue recruitment” and described as a form of Pavlovian conditioning (Haijiang et al., 2006). These statements generate testable hypotheses: first, the term “cue recruitment” implies generalizability, and therefore suggests that a variety of typically uninformative signals (other than retinal location) can also be recruited by the visual system as a cue to visual appearance. Second, characterization of this learning as classical conditioning suggests that the association between percept and retinal location (i.e. the conditioned stimulus), relies on unambiguous training trials that elicit a single percept (i.e. an unconditioned response). However, the available evidence presents problems for both of these hypotheses, and instead suggests that the learning of long-term biases is strongly driven by perceptual memory.

Describing the acquisition of long-term perceptual bias as ‘classical’ or ‘Pavlovian’ conditioning would extend the definition of these terms in a fundamentally new direction. Modern learning theorists have previously made such extensions in order to incorporate principles of classical conditioning within a Cognitivist framework. Contemporary views of Pavlovian conditioning therefore emphasize that animals adjust by developing accurate knowledge of the environment (Rescorla, 2003). However, sensory estimation is always subject to uncertainty, so one might argue that even perception of an ‘unambiguous’ stimulus involves some degree of interpretation, and therefore all types of association occur between ‘percepts’. It worth noting that our ‘ambiguous’ and ‘unambiguous’ stimuli do indeed lie on a continuum of perceptual uncertainty: the latter contain multiple depth cues (structure-from-motion, occlusion and binocular disparity), which are combined to

minimize uncertainty. In contrast, bistable structure-from-motion stimuli contain only a single depth cue that constrains the object structure but leaves the depth order of front and back surfaces underconstrained, resulting in high uncertainty between two possible interpretations. The mapping between sensory stimulus and perceptual interpretation differs based on the degree of uncertainty, yielding qualitatively different behavioural outcomes: for the unambiguous stimuli the mapping is more or less deterministic, while for the bistable stimulus it is stochastic and strongly influenced by past experience. In the present experiments, both the dependent variable and the reinforcing stimuli are percepts derived from an ambiguous stimulus, departing substantially from even the revised classical paradigms. For this reason, we prefer to characterize the retinotopic-specific learning observed in the present study as a form of long-term 'perceptual association' rather than any form of classical conditioning.

#### *Retinotopic specificity of perceptual memory affects perceptual association learning*

One striking similarity between experience-dependent biases at short and long timescales is their retinotopic specificity. Previous studies of short term biases have demonstrated that perceptual stabilization of ambiguous figures is tightly specific to retinal location; in particular, stabilization is typically disrupted if the retinal position of the stimulus is changed by more than one degree of visual angle (Chen and He, 2004; Knapen et al., 2009a). Consequently, bistable stimuli presented intermittently at two sufficiently distant retinal locations will stabilize independently of each other and can therefore become stabilized in opposite percepts, as we observed for some individuals in the untrained group. A key difference between using retinal location and other novel signals to train perceptual biases is that, in the former case, only one percept is trained at a given retinal location,

whereas in the latter case, two opposing percepts must be trained at the same location. Stabilization therefore promotes the learning of long-lasting perceptual associations only when retinal location is used as the novel signal, and instead interferes with learning when retinal location is kept constant. Accordingly, biases have indeed been reliably trained using retinal location as the novel signal, while various other signals appear to be ineffective (Haijiang et al., 2006; Harrison and Backus, 2010a; Jain et al., 2010). In the limited cases where non-retinal location signals have been shown to produce biases, these biases were much weaker (Haijiang et al., 2006; Jain and Backus, 2013) and either decayed rapidly after training ceased (Di Luca et al., 2010) or required special manipulations of the training condition (Harrison and Backus, 2012). This lack of generality beyond retinotopic location cues is incompatible with an account of long-term bias based exclusively on conventional associative learning, but supports our finding that perceptual stabilization plays an important role in driving longer-term bias.

Retinal-location appears to be a unique signal because of its importance to perceptual stabilization, which in turn presumably relates to the retinotopic organization of visual cortex. Nevertheless, previous studies using non-retinal cues suggest that some amount of traditional associative learning (learned associations between a perceptually robust or ‘unambiguous’ stimulus and a novel signal) can occur independently of stabilization in the training of long-term bias (Haijiang et al., 2006; Di Luca et al., 2010; Harrison and Backus, 2012). The fact that observable trained biases were produced in these studies suggests that learned associations between novel signals and perceptually robust signals (unambiguous stimuli) can be significant, since stabilization effects would have interfered with training. One way to disentangle these components might be via manipulations that reduce stabilization, such as presentation timing (Leopold et al., 2002)

or possibly physical stimulus properties (Brouwer and van Ee, 2006). However, given the importance of retinal location to perceptual stabilization, it would be of interest to assess whether any of the non-retinal novel signals used in previous studies might have caused systematic changes in observers' eye movements, shifting the retinal position of stimuli in a cue-contingent manner.

### *The nature of perceptual resolution*

Presented with ambiguous sensory input, the visual system must constructively process signals to interpret the structure and content of the environment. Even in dynamic environments, changes typically occur on timescales comparable to the temporal resolution of visual processing, resulting in temporal autocorrelations in natural vision. One useful strategy would therefore be to re-select the most recent interpretation of the same scene. This can be seen experimentally as perceptual stabilization during intermittent presentation of a stimulus, which is analogous to the occlusion and reappearance of a moving, ambiguous object (e.g. a camouflaged predator) being tracked behind a cluttered foreground (e.g. a forest). A second adaptive strategy would be to select the interpretation of the scene that has proven most frequently to agree with other sources of sensory information in the history of one's experience (for a given context). For example, if you had witnessed similar looking moving objects in this forest before, and also heard a roar, then it might be advantageous to be biased towards perceiving the object as a predator. This corresponds to prior knowledge, which has been shown to be modified by recent sensory experience (Adams et al., 2004). In contrast, the learning of perceptual associations demonstrates that recent perceptual experience alone (i.e. internal resolutions of ambiguous sensory information) is also capable of modifying subsequent perception, despite the absence of any feedback on the reliability of the perceptual interpretation.

A recent model of adaptation postulates that perceptual statistics of the remote past are used to estimate the world's statistics (Chopin and Mamassian, 2012). According to this model, adaptation is predictive: the next dominant percept will be the one that brings the statistics of recent perceptual history closer to that of the remote past (Maloney et al., 2005; Denison et al., 2011; Chopin and Mamassian, 2012). Although the present study was concerned primarily with perceptual memory for ambiguous stimuli rather than adaptation to unambiguous stimuli, our results contrast with this idea. Instead, our results suggest that recent perceptual history can play a more important role than the statistics of the remote past when these two sources of information conflict. This might reflect differences in the way the brain estimates world statistics from past experience based on the reliability of the sensory information, i.e. the difference between the same percept arising from an ambiguous or unambiguous stimulus. Differences in the way the brain processes these stimuli are clearly evident from previous work indicating that long-term biases are more strongly driven by ambiguous than unambiguous stimuli (Harrison and Backus, 2010b).

### *Conclusion*

Here we investigated the contribution of short-term perceptual processes to the learning of retinal location-contingent biases for a bistable stimulus. We found that even in the absence of explicit training, bias emerges naturally as a product of stabilized perception at each location. Perceptual biases were immediate, retinotopically specific, and maintained throughout the testing session, irrespective of whether observers viewed disambiguated training trials or not. Further, we observed that a period of spontaneous perceptual alternations, induced by continuous presentation of the ambiguous stimulus, abolished the previously acquired bias. This manipulation had no effect on the bias at a second, spatially removed location, confirming the retinotopic specificity of the effect. Further analysis

revealed that subsequent biases at the continuous presentation location were strongly determined by the final dominant percept reported during the continuous presentation. These results suggest that the learning of long-term biases in the perception of ambiguous stimuli relies heavily on short-term perceptual processes, which promote repeated selection of the most recent perceptual interpretation at a given retinal location.

## 4.4. Methods

### *Hardware and software*

Stimuli were programmed in OpenGL using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for MATLAB™ (Mathworks, Natick, USA). In both experiments, stimuli were presented on a mirror stereoscope in a Wheatstone configuration using a pair of ViewSonic P225f CRT monitors (1600 x 1200, 100Hz). Observers' head position was stabilized by means of a chin rest at a viewing distance of 50cm. An EyeLink 1000 (SR Research Ltd., Ontario, Canada) infrared video eye-tracking system was used to track the position of both eyes, in order to ensure observers maintained fixation throughout the experiments.

### *Stimuli*

Virtual rotating cube stimuli were replications of those used by Harrison and Backus (2010b), with the exception that they were smaller in size in order to permit presentation on our stereoscope. Stimuli consisted of orthographic projections of a white wire-frame cube (Necker, 1832) on a mid-grey background. The centre of the cubes were positioned either 8.5° above or below a central fixation marker, which was a 1.15°x1.15° square outline. Cube edges subtended 6.5° when oriented in the frontoparallel plane and were 4' wide. Each transparent face of the cube contained 25 randomly located white dots with a diameter of 6'. Cubes rotated about a vertical axis corresponding to the vertical meridian at an angular velocity of 45°s<sup>-1</sup>.

The starting orientation of the cubes was set by roll and pitch angles of  $\pm 25^\circ$ , balanced across trials. For unambiguous trials (**Figure 4.1A**) these angles determined



whether the cube appeared to be viewed from above or below at stimulus onset. On unambiguous trials, a vertical strip ( $0.5^\circ$  wide x full screen height) was drawn through the axis of rotation to provide an occlusion depth cue and the cube was rendered stereoscopically with appropriate binocular disparity. On ambiguous trials (**Figure 4.1B**), stimuli were presented monocularly to the right eye in order to remove the cue conflict that would otherwise arise from a structure-from-motion stimulus being presented binocularly on a flat screen.

### *Task*

In all experiments, participants were instructed to fixate the central marker, which was always present. Participants initiated trials by key press, and following a brief delay ( $<0.5$  s), a single rotating cube appeared, either above or below fixation. In the first experiment a probe dot appeared on each trial, which repeatedly moved through the central fixation marker at a speed of  $9.15^\circ\text{s}^{-1}$ , either from left to right, or from right to left. The direction of the probe dot was randomized across trials in order to decouple perceived direction of cube rotation from the motor response. The probe dot was presented at zero disparity on unambiguous trials and monocularly on ambiguous trials. Participants reported (using the computer's keyboard) whether the probe dot was moving in the same direction or the opposite direction to the perceived front face of the cube (**Figure 4.1C-E**). In the second experiment no probe dot was presented and participants instead reported (via the keyboard) whether the front face of the cube moved to the left or right.

On each trial, the stimulus was presented for 1.5 s. This duration was selected based on the observation that participants in a pilot study—who were allowed to give their response at any time between 1.5 and 6 seconds after stimulus onset (following Harrison and Backus, 2010b)—responded on average 1.78 seconds ( $\pm 0.07$  SEM) after stimulus

onset. This corresponds to just 280 milliseconds after the appearance of a cue indicating the start of the response period (a change of fixation marker colour), which suggests that perceptual decisions were—on average—made within the first 1.5 seconds. Fixing presentation duration served to ensure that all observers experienced the same stimulus on duration, which is known to be an important parameter for perceptual stabilization. The minimum duration between consecutive stimulus presentations was set to 1s, although on average it was approximately 5.5 s for a given location since trial onset was self-paced and stimulus location was randomized. Unambiguous stimuli appearing above fixation always rotated in the opposite direction to those appearing below fixation, and this location-rotation contingency was counterbalanced across participants. In blocks that included unambiguous trials, these trials always occurred within the first 8 trials of the block. The first 8 trials therefore included every possible stimulus configuration: ambiguous and unambiguous, viewed from above and viewed from below, and located above or below fixation.

### *Participants*

In total, 55 naïve observers participated (16 – 42 years old, mean age = 21.4 years, 16 males). They were recruited through the School of Psychology at the University of Birmingham and received either £6/hr or research scheme credits for their participation. Participants had normal or corrected-to-normal vision. They provided written informed consent, in line with the ethics approval granted by the University of Birmingham's STEM ethics committee and the Declaration of Helsinki.

Data from participants in the trained group of experiment 1 who exhibited overall perceptual biases opposite to that expected based on training at either retinal location, were excluded from further analysis ( $N = 4$ ). In such cases, perception may have initially been

biased by a strong pre-existing bias or a negative aftereffect from the preceding unambiguous stimulus, but in either case could not be considered ‘trained’. For experiment 2, participants were divided into two groups based on whether training successfully influenced perceptual bias during block 1. All participants correctly identified the direction of rotation on unambiguous training trials. Eye movement data revealed that all participants tested were largely successful in maintaining fixation during stimulus presentations.

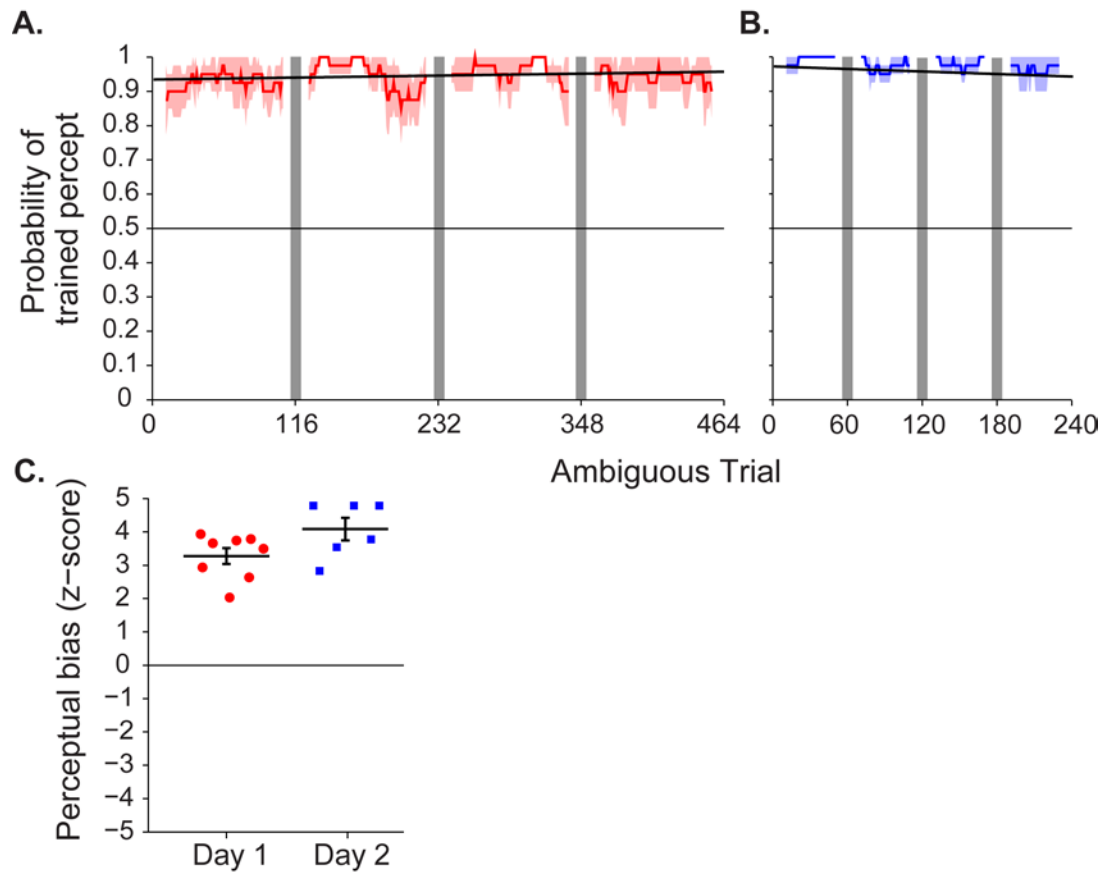
### *Analysis*

To analyse the time course of perceptual bias, data were boxcar filtered using a sliding window of 40 trials (20 trials at each stimulus location). For each window position, the proportion of trials for which the reported direction of rotation matched the direction specified by disambiguated training trials at the given location was calculated. We use the term ‘perceptual bias’ here to refer to the probability of experiencing one percept over the other. We make the assumption that *on average across individuals*, untrained observers begin with an equal probability of perceiving either percept.

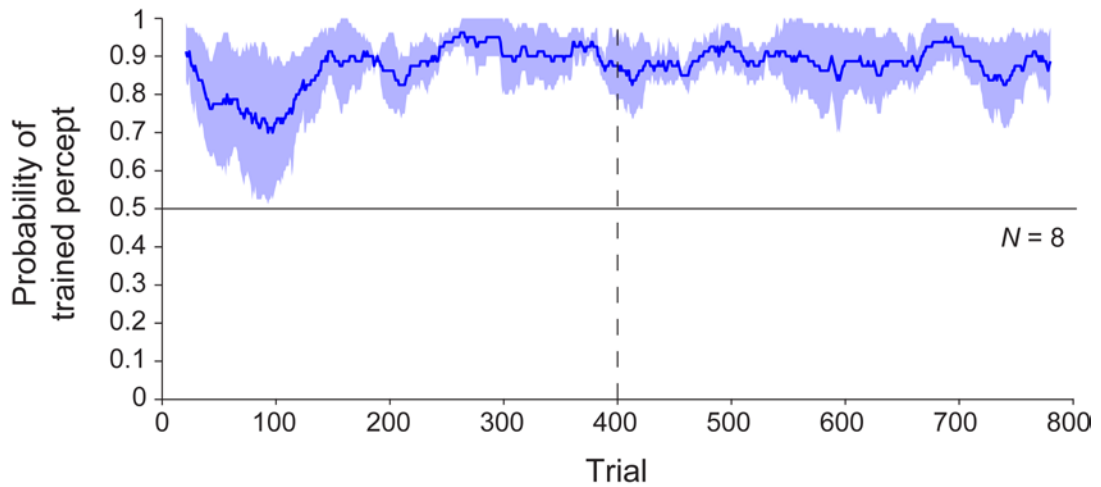
To facilitate comparison of our results with those from previous studies, the proportion of ambiguous stimuli reported as rotating in the direction specified by disambiguated stimuli at each location were converted into z-scores using a probit transformation (inverse cumulative distribution function) (Harrison and Backus, 2010b). Saturated values (probabilities of 0 or 1) were substituted with maximum z-scores equivalent to one response in the opposite direction to training per block ( $\pm 2.394$  in experiment 2). The sum of the z-scores for the two locations was calculated for each subject, to give a measure of training-induced perceptual bias. All analyses were conducted

using MATLAB<sup>TM</sup> (Mathworks, Natick, USA); in particular, robust regressions were implemented using the built-in function *robustfit.m*.

## 4.5. Supplementary Materials



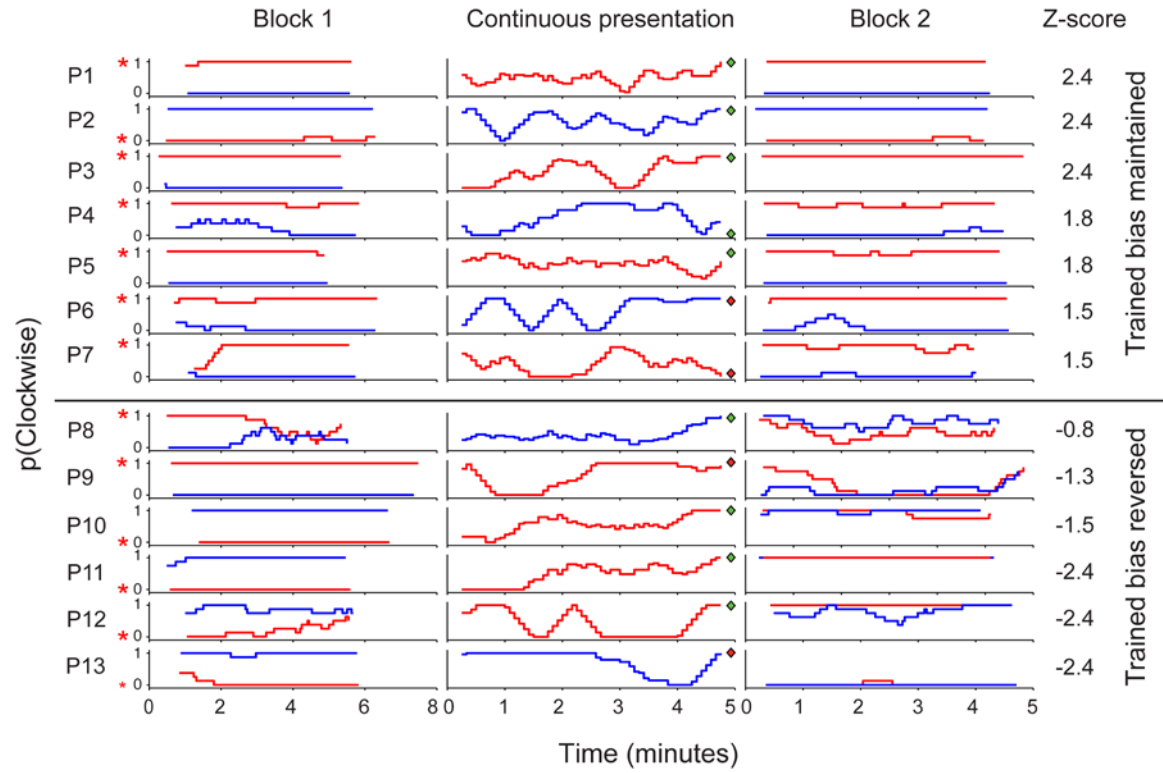
**Figure S4.1.** Time course of perceptual bias in the pilot experiment, which was a replication of Harrison and Backus’s (2010b) ‘uninformative’ training group. (A) On day 1, observers ( $n = 8$ ) completed 4 blocks of 120 trials. Each block began with 2 unambiguous trials (within the first 4 trials) at each retinal location, and observers took 2 minutes break between blocks. The time course reveals that strong biases were expressed immediately and persisted across blocks. However, the short duration of each block and the presence of unambiguous trials at the start of each block made the relative contributions of short-term (perceptual memory) and long-term (associative learning) processes difficult to assess. (B) On day 2, observers ( $n = 6$ ) again completed 4 blocks of 120 trials, but now 50% of trials were unambiguous and specified the opposite direction of rotation to day 1 training at each location. Black lines indicate robust regression fits to data from all 4 blocks in each session. (C) Overall perceptual bias in trained direction for ambiguous trials across days for individual observers. Despite attempted ‘counter-conditioning’ on day 2, the biases specified by day 1 training were, overall, even stronger than on day 1, although this difference did not reach significance. This result confirms that the training regimen employed in Experiments 1 and 2 was capable of eliciting robust, long-term learning effects, as previously described in “cue recruitment”.



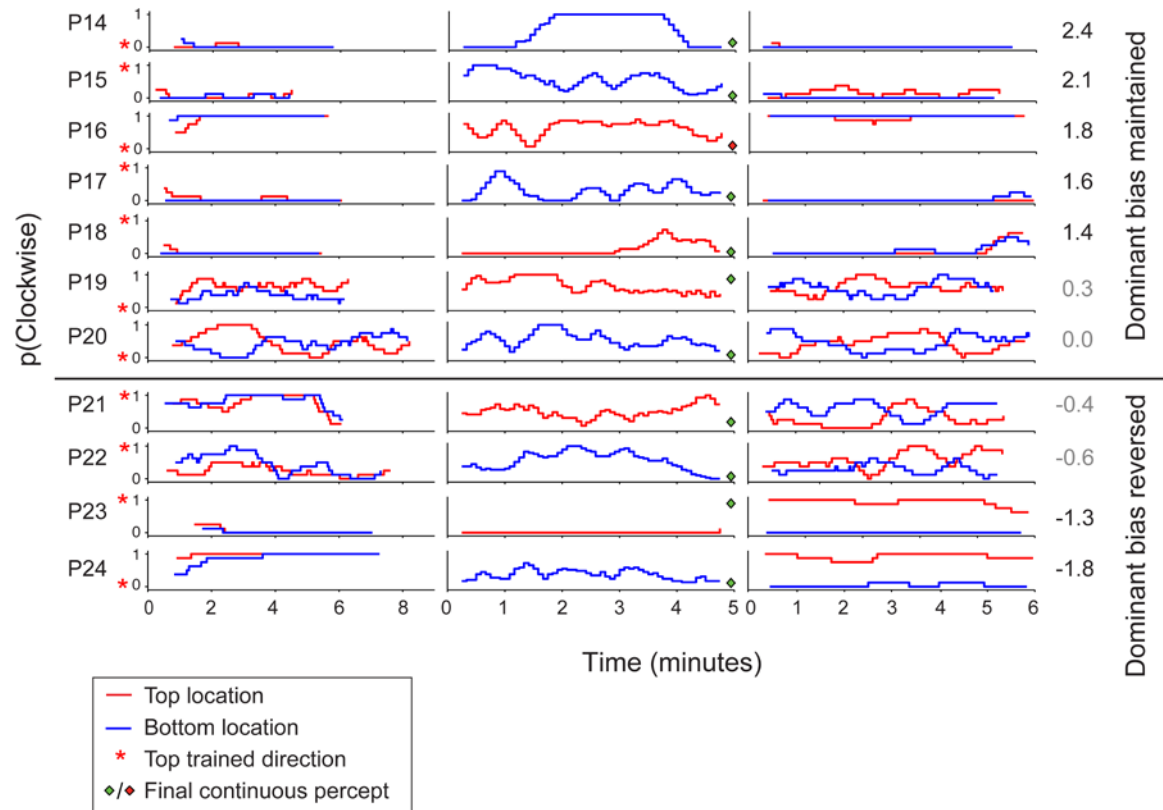
**Figure S4.2.** Time course of perceptual bias for a subset of participants in the ‘trained’ group of experiment 1 ( $N = 8$ ), who performed two ‘superblocks’ of 400 trials each. Biases remained stable throughout the sessions, which took participants approximately 40 minutes

**Figure S4.3** (Next page). Time courses of perceptual bias for individual participants in experiment 2. (A) Full learners - participants who expressed biases in the trained direction at both locations. (B) Partial learners - participants who expressed a bias in the trained direction at only one location. For clarity, raw data was boxcar filtered with a window of 8 trials for intermittent blocks or 28 seconds for continuous blocks. Red asterisks indicate the direction specified by unambiguous ‘training’ trials at the top location in block 1. Diamond symbols indicate the last reported percept during the continuous block, and their colour indicates whether block 2 bias matched that percept (green = yes, red = no). (B) Observers in the partial learners group only acquired the ‘trained’ bias at one retinal location during block 1. Note that all participants correctly identified the direction of rotation for the unambiguous ‘training’ trials, which were presented at the start of block 1. Despite failure to bias perception in the direction specified by training, perception readily stabilized in the opposite direction.

## A. Full learners



## B. Partial learners



## 4.6. Appendices

### *Pilot experiment*

In an initial pilot study we replicated the results of Harrison and Backus (2010b). Stimuli were presented on a Samsung SyncMaster 2493HM in portrait screen orientation (1020 x 1900, 60Hz), while observers sat with their heads supported by a chinrest at a viewing distance of 50cm. This configuration permitted presentation of stimuli equal in visual angle to those used by Harrison and Backus (2010b). Prior to testing a custom gamma table was calculated from photometer readings taken through the red and green filters of the anaglyph glasses worn by observers, and applied during stimulus presentation to minimize cross-talk in the anaglyphs. This step is critical when attempting monocular presentation using anaglyphs, as any weak signals reaching the supposedly ‘covered’ eye will result in the Pulfrich effect (Pulfrich, 1922).

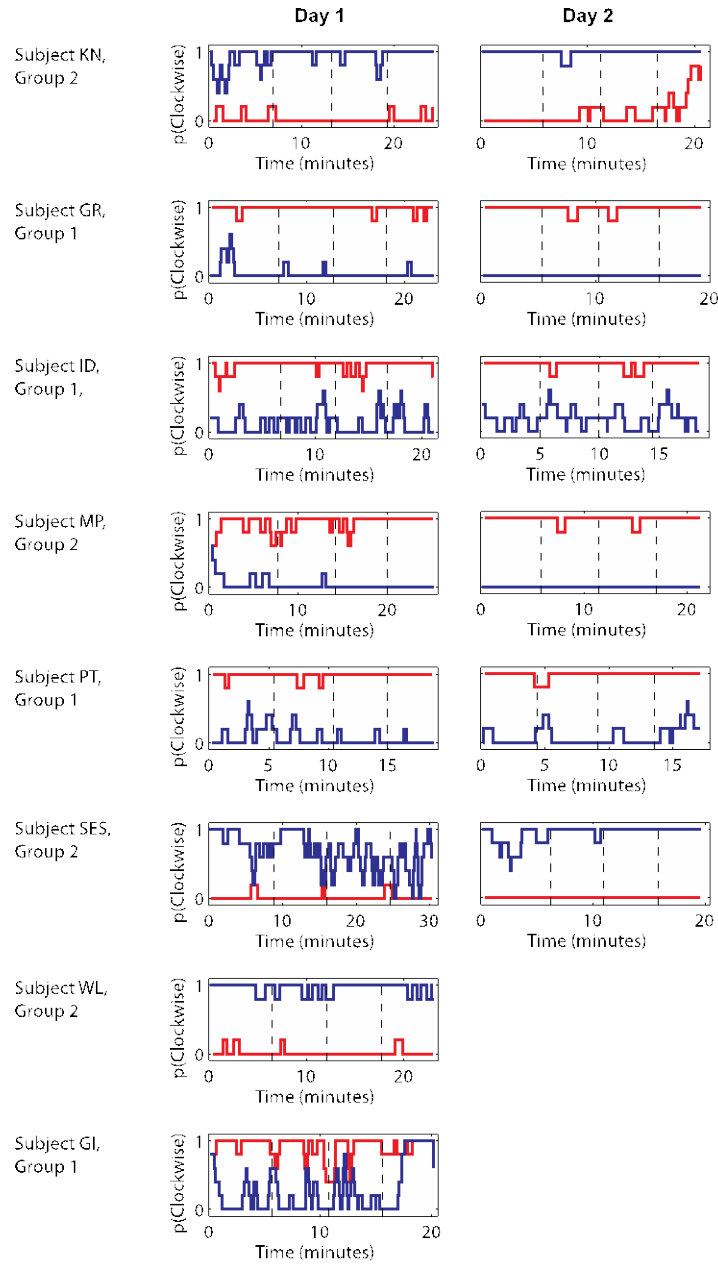
Stimuli were identical to those used by Harrison and Backus (2010b) except in physical size, and identical to those of Experiments 1 and 2, except in both physical and retinal size. The centres of the cubes were positioned either  $12^\circ$  above or below the central fixation marker, which was a  $1.15^\circ \times 1.15^\circ$  square outline. Cube edges subtended  $11.42^\circ$  when oriented in the fronto-parallel plane and were 10.2’ wide.

A total of 15 observers participated in 2 experimental sessions on consecutive days: 8 participants using the anaglyph setup described above and 7 using the Wheatstone stereoscope setup described in **Methods**. Each session consisted of 4 blocks of 120 trials. On day 1, 4 of the first 8 trials in each block ( $<4\%$ ) were unambiguous and the direction of rotation was contingent on retinal location. On day 2, 60 trials in each block (50%) were unambiguous and the location-rotation contingency from day 1 was reversed.



The proportions of ambiguous trials perceived in the same direction as disambiguated training trials at each location were converted into z-scores (as described in General Methods) and added together to produce a ‘z-diff’ value for each observer. The mean z-diff values for day 1 and day 2 were not significantly different (**fig.S1C**; day 1 mean= 3.28, s.e.m.= 0.24; day 2 mean= 3.87, s.e.m.= 0.26;  $t(12) = -1.66$ ,  $p = 0.12$ , 95% CI= [-1.37, 0.18]). This result replicates the findings of Harrison and Backus (2010b), and demonstrates that long-term perceptual bias was induced on day 1 since perception was resistant to priming in the opposite location-rotation contingency on day 2.

As a further analysis, the ambiguous trials from all blocks for each subject were divided into bins of 20 trials and the mean proportion of responses at either location in the trained direction was plotted (**fig.S1A**). Fitting a linear regression to these data revealed that the trained bias took immediate effect from the start of the first block on day 1, ( $\alpha = 0.83$ , 95% CI = 0.80 – 0.85; bin 1 mean = 0.81, s.e.m.= 0.1). Furthermore, the induced bias was maintained at a constant level throughout each session, as revealed by the shallow gradient of the regression line ( $\beta = -4.27 \times 10^{-5}$ , 95% CI=  $-1.45 \times 10^{-4} - 5.91 \times 10^{-5}$ ). A similar pattern of results was observed for ambiguous trials on day 2 (**fig.S1B**;  $\beta = 1.0 \times 10^{-5}$ , 95% CI=  $-1.55 \times 10^{-4} - 1.65 \times 10^{-4}$ ;  $\alpha = 0.96$ , 95% CI = 0.94 – 0.98; bin 1 mean = 0.95, s.e.m.= 0.02).



**Figure A4.1.** Time course of perceptual bias for individual observers in the pilot study.

## Chapter 5: Long-term biases for bistable stimuli are modified by recent perceptual experience.

---

In the previous chapter we observed that recent perceptual experience, driven primarily by either short-term mnemonic processes (perceptual memory) or stochastic processes (perceptual alternation), is capable of influencing perception over longer time scales. Here we extended these time scales to test the limits of recent experience on learning, and potential consolidation during sleep. The first experiment was similar to that in chapter 4, except that the test phase took place 24 hours after the training and disruption phases. The results confirm several findings reported in chapter 4: i) a period of perceptual alternations can disrupt previously learned biases, and ii) the perceptual statistics of the most recent encounter (as opposed to earlier training) influence subsequent perceptual bias. However, in contrast to the results observed at shorter timescales, we found that perceptual biases 24 hours later reflected the integration of net perceptual experience across the entire previous session rather than just the most recent experience. In addition to the long-term learning and short-term stabilization biases tested thus far, naïve observers appear to exhibit esoteric ‘pre-existing’ perceptual biases that vary across the visual field, but remain stable across time. However, previous studies reporting this phenomenon have failed to account for the confounding contribution of short-term memory to the observed result. In a second experiment, we therefore tested the longevity of esoteric perceptual biases across the visual field using manipulations that diminished the biasing contribution of short-term memory. The results suggest that perceptual biases that occur at first exposure to ambiguous visual input are indeed stable across long time periods, even when short-term processes are controlled for. Finally, we report the successful ocular transfer of training, which indicates that the neural mechanism underlying these learned biases does not occur earlier than primary visual cortex.

## 5.1. Introduction

Faced with ambiguous sensory input, the brain must constructively process the available information to arrive at a plausible perceptual interpretation of the environment. In the absence of sufficient sensory evidence to distinguish between competing interpretations, the brain can use the statistics of past perceptual experience to inform percept selection (Leopold et al., 2002;Adams et al., 2004;Maloney et al., 2005;Brascamp et al., 2008;Pastukhov and Braun, 2008;Brascamp et al., 2009;Harrison and Backus, 2010b). However, different mechanisms may mediate the influence of recent and distant past perceptual experience, and the percept promoted by these processes can sometimes conflict - exerting opposing effects on future perception (Chopin and Mamassian, 2012;Murphy et al., 2014).

Bistable stimuli provide a unique means of investigating how past experience biases perception. Short-term influences of past experience on perception are known to occur when bistable stimuli are presented intermittently. Specifically, brief presentations separated by intervals during which the stimulus is removed results in the same percept being repeatedly selected on consecutive presentations (Orbach et al., 1963;Leopold et al., 2002;Maier et al., 2003;Brascamp et al., 2009). While the neural basis of this perceptual stabilization phenomenon remains unknown, it has generally been described in terms of a short-term perceptual memory trace, which accumulates over a timescale of seconds and can last for tens of minutes (Brascamp et al., 2008;Pastukhov and Braun, 2008;Pearson and Brascamp, 2008). Another noticeable feature of stabilization is that it occurs in retinotopic, rather than spatiotopic coordinates (Chen and He, 2004;Knapen et al., 2009a).

Longer lasting perceptual biases for briefly presented ambiguous stimuli can be trained, and these too are specific to retinal location (Haijiang et al., 2006;Harrison and Backus, 2010b;a;van Dam and Ernst, 2010;Jain and Backus, 2011). The acquisition of

such long-term biases appears to be strongly driven by short-term perceptual stabilization, resulting from the brief presentations of stimuli (Harrison and Backus, 2010b; Murphy et al., 2014). However, long-term biases can extend to days (Haijiang et al., 2006; Harrison and Backus, 2010b; Mamassian and Wallace, 2010), or even weeks (Carter and Cavanagh, 2007; Afraz et al., 2010; Harrison and Backus, 2014), and it is possible that sleep may play an important role in the consolidation of learned biases (Stickgold et al., 2000; Stickgold, 2005). Further, some studies have reported ‘pre-existing’ biases for perceptual interpretation of ambiguous stimuli at stimulus onset. These biases are not the result of explicit training, and have been described as idiosyncratic between observers and retinal locations, but stable across time (Carter and Cavanagh, 2007; Stanley et al., 2011).

Several lines of evidence suggest that the process of selecting a perceptual interpretation of a bistable stimulus each time it appears during repeated presentation (‘onset rivalry’) differs from perceptual rivalry during continuous viewing of the same stimulus (Carter and Cavanagh, 2007; Kalisvaart et al., 2011; Stanley et al., 2011). For example, during continuous viewing, the rate of perceptual alternation is reduced when attention is directed away from the stimulus (Paffen et al., 2006), but conversely, increased when the same manipulation is applied during intermittent viewing (Kanai and Verstraten, 2006). Further, prior exposure to bistable stimuli can have opposing influences on subsequent perception depending on whether stimuli are presented continuously or intermittently (de Jong et al., 2012a). However, there seems to be an interaction between onset bias and perceptual alternations that occur spontaneously during continuous viewing. Specifically, recently acquired onset biases can be disrupted by a period of continuous viewing (Brascamp et al., 2008; Murphy et al., 2014).

Here, we examine the long-term effects of trained associations between retinal locations and perceptual biases, and the influence of perceptual alternations on the

longevity of trained biases. In Experiment 1 we tested the long-term effects of perceptual alternations following recent acquisition of learned perceptual bias. In the previous chapter we demonstrated that a 5-minute period of continuous viewing of an ambiguous stimulus reduces recently acquired perceptual bias at that retinal location to chance levels. However, it is not clear how the competing effects of past perceptual experience would affect subsequent perceptual bias on a longer time scale, especially following sleep. To examine this, we tested observers on two consecutive days, as has been used in previous “cue recruitment” studies to measure learned biases (Haijiang et al., 2006; Harrison and Backus, 2010b).

When observers receive no explicit training, their perception of bistable stimuli tends to naturally stabilize on one of the two percepts at a given retinal location. This could reflect strong ‘pre-existing’ biases, short-term perceptual stabilization effects, or some combination of these two biasing factors. While the concept of ‘pre-existing’ bias is plausible (the observer must always make a binary choice on trial 1), studies to date have made no effort to control for the confounding contribution of stabilization in their measurement of onset bias. In Experiment 2 we assessed pre-existing onset biases by increasing the time between consecutive ambiguous trials, so as to minimize the contribution of perceptual stabilization, and allowed for spontaneous perceptual alternations to disrupt perceptual stabilization between trials. Finally, in Experiment 3 we tested whether recently learned associations between retinal location and perceived direction of rotation are specific to the eye in which they were induced. The results suggest that learned bias is represented binocularly, and indicates that the neural locus of such processes is retinotopically organized visual cortex.

## 5.2. Methods

### *Participants*

In total, 38 naïve observers participated (16 – 42 years old, mean age = 21.4 years, 12 males). Participants in **Experiment 1** (n = 26), and **Experiment 3** (n = 10) were recruited through the School of Psychology at the University of Birmingham and received either £6/hr or research scheme credits for their participation. Participants in **Experiment 2** (n = 6) were trained psychophysical observers who were naïve to the aim of the experiment. All participants had normal or corrected-to-normal vision and provided written informed consent, in line with the ethics approval granted by the University of Birmingham's STEM ethics committee and the Declaration of Helsinki. Data from two participants in Experiment 1 were not recorded correctly for last 2 blocks of session 2 due to a programming error. Their available data were included in the analyses where appropriate. Eye movement data from the control participants in **Experiments 2** and **3** revealed that all participants tested were successful in maintaining fixation during stimulus presentations.

### *Hardware and software*

Stimuli were programmed in OpenGL using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for MATLAB™ (Mathworks, Natick, USA). In **Experiment 1**, stimuli were presented on a Samsung SyncMaster 2493HM LCD monitor (1920 x 1200, 60Hz, 52 x 33 cm), with the screen rotated 90° for a portrait orientation. Observers' head position was stabilized by means of a chin rest at a viewing distance of 60cm and binocular presentation was achieved using red-green anaglyph glasses. A custom gamma table was created based on photometric readings taken through the anaglyph filters, in order to minimize cross talk between the red and green channels.

Due to the presentation timing requirements of the experiments, **Experiments 2** was not conducted under laboratory conditions for the majority of participants. Stimuli were presented on a range of standard office LCD monitors (see Appendix Table 1), which had not been calibrated for luminance, and ambient lighting was not controlled. However, for each subject stimulus size was adjusted according to viewing distance (which varied between 48 and 65cm) to maintain constant size in visual angle across the different experimental settings.

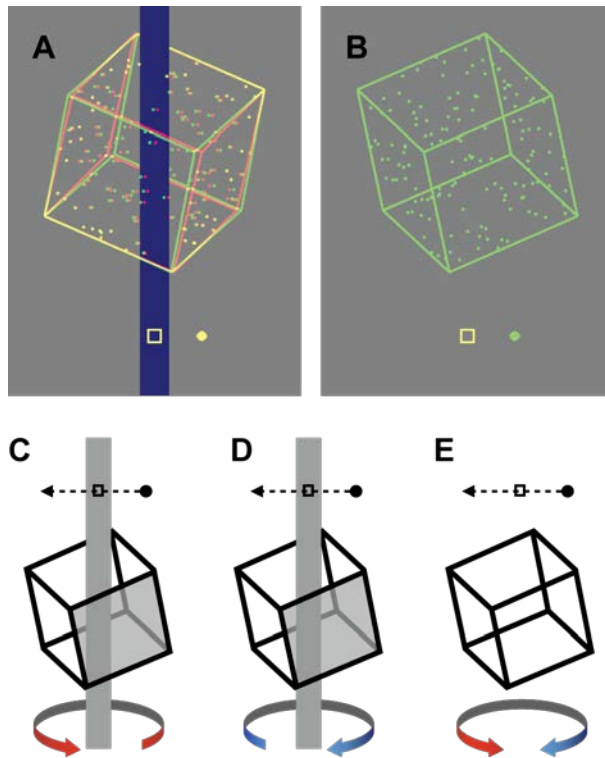
For two control observers in **Experiments 2**, standard laboratory conditions were maintained. Stimuli were presented on a pair of luminance calibrated Samsung RZ2233 LCD monitors (1920 x 1080, 120Hz), which have previously been extensively tested and characterized as suitable for psychophysical experimentation (Wang and Nikolić, 2011). The two monitors were viewed through cold mirrors in a Wheatstone stereoscope configuration. An EyeLink 1000 (SR Research Ltd., Ontario, Canada) infrared video eye-tracking system was used to track the position of both eyes through cold mirrors, in order to ensure observers maintained fixation throughout the experiments.

### *Stimuli*

Virtual rotating cube stimuli were identical to those used in **Chapter 4**, except for being scaled up in size to match those used by Harrison and Backus (2010b). In **Experiment 1** the centers of the cubes were therefore positioned either 12° above or below a central fixation marker, which was a 1.15°x1.15° square outline. Cube edges subtended 11.3° when oriented in the frontoparallel plane and were 10' wide, and each cube surface contained 25 dots (0.28°).

In **Experiment 2**, random dot kinematograms of orthographic projections of rotating cylinders were similar to those used by Knapen et al. (2006) (1° x 1° diameter).





**Figure 5.1.** Cropped screen shots of the stimuli, rendered here for red-green anaglyph presentation. **A)** Unambiguous stimulus disambiguated by occlusion and disparity cues. **B)** Ambiguous stimulus presented monocularly. **C, D, E)** Schematic representations of the task. Participants reported whether the front face of the cube (shaded in the unambiguous examples) moved in **(D)** the same direction as the probe dot or **(C)** the opposite direction to the probe dot. **(E)** On ambiguous trials the task was the same, but there was no correct answer.

Cylinders were composed of 200 white dots ( $1.8'$ ) presented on a black background (to improve on-screen contrast in ambient lit rooms) and rotated at an angular velocity of  $130^\circ\text{s}^{-1}$ . The centre of each cylinder could appear either at fixation or at one of 8 radial locations at  $3^\circ$  eccentricity from fixation (see **Figure 5.8**). Stimulus dimensions were specified in degrees of visual angle and converted to pixels for each observer's viewing conditions based on input of viewing distance and screen dimensions. A circular fixation marker ( $0.1^\circ$ ) was presented at the centre of the screen throughout each block.

### Task

In all experiments, participants were instructed to fixate the central marker, which was always present. Participants initiated trials by key press, and following a brief delay (0.1 s), a single rotating stimulus appeared. In Experiment 1 this was a cube that appeared either above or below fixation, while in Experiment 2 it was a cylinder that could appear either centrally (behind the fixation marker) or at one of 8 polar angles (**Figure 5.8**). In

Experiments 1 and 3 a probe dot appeared on each trial, which repeatedly moved through the central fixation marker at a speed of  $9.15^{\circ}\text{s}^{-1}$ , either from left to right, or from right to left. The direction of the probe dot was randomized across trials in order to decouple perceived direction of cube rotation from the motor response. The probe dot was presented at zero disparity on unambiguous trials and monocularly on ambiguous trials. Participants reported (using the computer's keyboard) whether the probe dot was moving in the same direction or the opposite direction to the perceived front face of the cube (**Figure 5.1C-E**). In Experiment 2 no probe dot was presented and participants instead reported (via the keyboard) whether the front face of the rotating cylinder moved to the left or right.

On each trial, the stimulus was presented for 1.5 s. Fixing presentation duration served to ensure that all observers experienced the same stimulus duration, which is known to be an important parameter influencing perceptual stabilization (Manousakis, 2012). The minimum duration between consecutive stimulus presentations was set to 1s, although on average it was approximately 5.5 s for a given location since trial onset was self-paced and stimulus location was randomized. Unambiguous stimuli appearing above fixation always rotated in the opposite direction to those appearing below fixation, and this location-rotation contingency was counterbalanced across participants. In blocks that included unambiguous trials, these trials always occurred within the first 8 trials of the block. The first 8 trials therefore included every possible stimulus configuration: ambiguous and unambiguous, viewed from above and viewed from below, and located above or below fixation. Participants were required to take a 1-minute break between consecutive blocks within a session. On average, intermittent blocks took observers 5 minutes 42 seconds ( $\pm 12\text{s}$ ) to complete on day 1, and 5 minutes 3 seconds ( $\pm 10\text{s}$ ) on day 2, and were therefore reasonably well matched by the 5-minute block of continuous presentation.






















































### *Analysis*

We calculated perceptual bias as the proportion of ambiguous stimuli reported as rotating in a given direction. The direction could be absolute (e.g. clockwise) or relative (e.g. compared to the direction specified by disambiguated stimuli). Biases were calculated separately for each retinal location, and were converted into z-scores using a probit transformation (inverse cumulative distribution function) (Harrison and Backus, 2010b). Saturated values (probabilities of 0 or 1) were substituted with maximum z-scores equivalent to one response in the opposite direction to training per block ( $\pm 2.394$  in Experiment 1). The sum of the z-scores for the two locations was calculated for each subject, to give a measure of training-induced perceptual bias. All analyses were conducted using MATLAB™ (Mathworks, Natick, USA).

## 5.3. Results

### 5.3.1. Experiment 1: Long-term effects of perceptual bias disruption

Experiment 1 was designed to test the longevity of recently acquired perceptual biases at multiple time scales: from 5-minutes to 24 hours after initial bias acquisition. On the first day, observers completed an initial training block consisting of 120 trials, followed by a ‘manipulation’ block, and a final test block (**Figure 5.2**). For a subset of observers (group 1) the training block contained a small number (4) of unambiguous trials near the start that were intended to prime subsequent perception in opposite directions at the two locations. All other observers received no explicit training, and were presented exclusively ambiguous stimuli on day 1. During the manipulation block, observers were either presented 5-minutes of continuous stimulation, or presented with a block of 120 trials at one of the two locations (randomized across participants). The retinal location where no stimulus was presented during block 2 remained blank and served as a within-subjects

		Day 1			Day 2		
		Block 1	Block 2	Block 3	Block 1	Block 2	Block 3
Group 1 Trained Continuous	Top location	 x 2  x 58	 x 5 minutes	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30
	Bottom location	 x 2  x 58	Blank	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30
Group 2 Untrained Intermittent	Top location	 x 60	 x 120	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30
	Bottom location	 x 60	Blank	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30
Group 3 Untrained Continuous	Top location	 x 60	 x 5 minutes	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30
	Bottom location	 x 60	Blank	 x 60	 x 30  x 30	 x 30  x 30	 x 30  x 30

**Figure 5.2.** Design of Experiment 1. Participants were randomly assigned to one of 3 groups, each of which received different combinations of training and disruption during blocks 1 and 2 of day 1. All blocks on day 2 were the same for all participants: 50% of trials were unambiguous and rotated in the opposite direction to whichever direction had dominated at each location during session 1. Note that the manipulated location in block 2 could be either top or bottom, and this was counterbalanced across observers.

control condition.

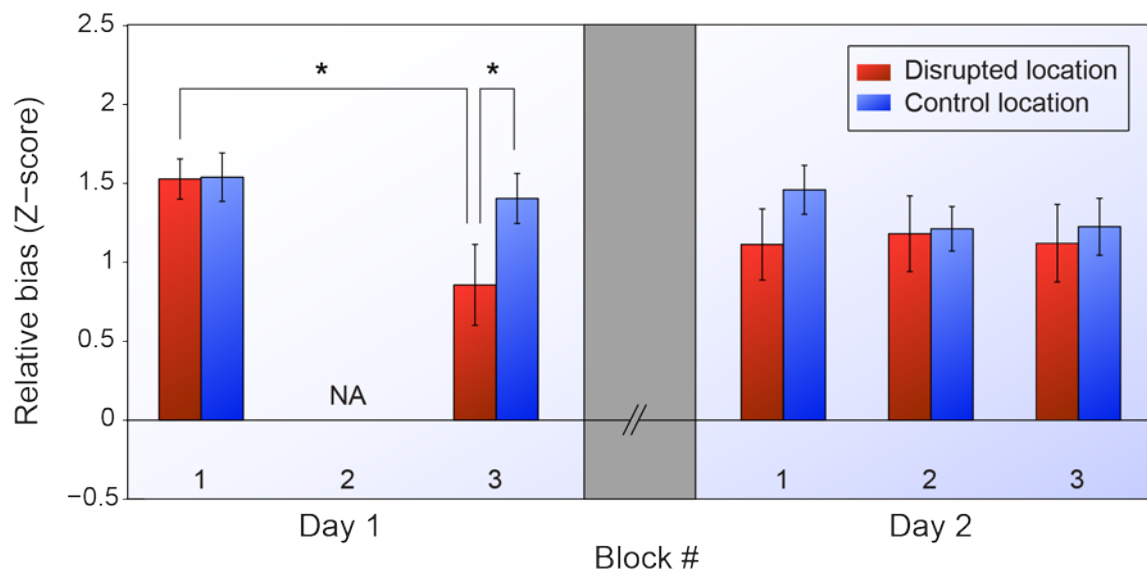
On the second day, observers completed 3 identical test blocks, each of which consisted of 50% unambiguous trials. The direction of rotation on unambiguous trials at each location was opposite to the direction that had dominated during at that location during the previous session. Day 2 blocks were therefore an attempt at ‘counter-conditioning’ of biases that were manifested on day 1, since unambiguous versions of bistable stimuli have been shown to prime subsequent perception (Nawrot and Blake, 1993; Kanai and Verstraten, 2005; Long and Moran, 2007).

We first measured observers’ ability to correctly identify the direction of rotation on unambiguous trials. Observers in group 1 perceived 90% of unambiguous trials correctly, and all observer perceived at least 3 of the 4 unambiguous trials correctly. Similarly, all observers reported 92.8% ( $\pm 1.3\%$ ) of unambiguous trials on day 2 correctly. This suggests that observers were attending to the task of reporting their percepts rather than giving responses randomly. We next examined the efficacy of the unambiguous training trials presented to observers in group 1, in priming perception of subsequent ambiguous trials in block 1. The direction of perceptual bias specified by unambiguous training trials was only successful in biasing perception at both locations during block 1 in 50% of observers in group 1 (6 of 12). For the other observers, perceptual bias in block 1 tended to follow training at only one location, and bias at the other location was therefore in the same direction. Incidentally, training trials were not necessary to elicit opposite bias directions at the two retinal locations: 35% of observers in groups 2 and 3 (7/20) who did not receive any training spontaneously exhibited opposite directional biases at the two locations tested.

The second manipulation we performed was to either continuously or intermittently present an ambiguous stimulus at one of the two retinal locations during block 2. The

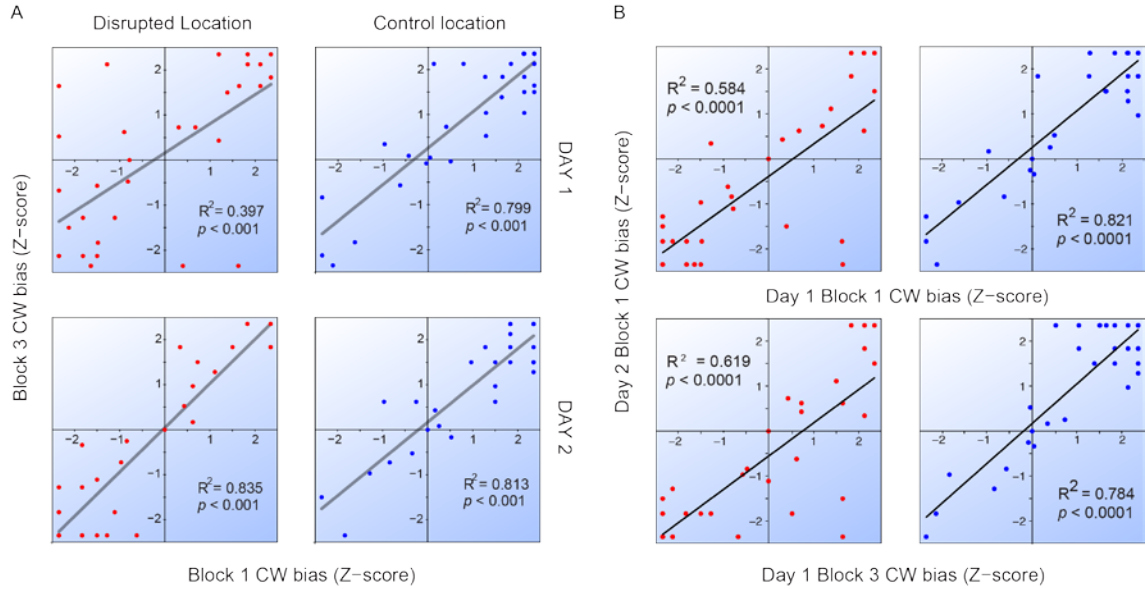
purpose of this manipulation was to induce varying frequencies of perceptual alternation during this block in different observers. Accordingly, our observers exhibited a wide range of perceptual alternation rates during block 2, ranging from zero to 46 alternations over a 5 minute period. In contrast, and as expected based on previous experiments, during intermittent presentation blocks, the majority of observers showed highly stable (and therefore biased) perception of ambiguous stimuli at each retinal location. Filtered versions of the raw perceptual time courses for each observer during each block are shown in Supplementary **Figure S5.1**.

We quantified perceptual bias at each retinal location for each block by calculating the proportion of ambiguous trials at that location that were perceived as rotating clockwise, and converting the data to Z-units using the probit transform (see Methods). A positive Z score therefore reflects a clockwise bias, whereas a negative Z score reflects an anticlockwise bias. For an initial analysis, bias directions were converted so that the sign was relative to block 1 bias rather than a particular direction of rotation, and thus all biases



**Figure 5.3.** Mean perceptual bias for each location during each intermittent block. Bias directions are relative to block 1 bias direction for each location, rather than relative to a particular direction of rotation. Error bars show standard error. \*Significantly different at  $p < 0.05$  – see main text for details.

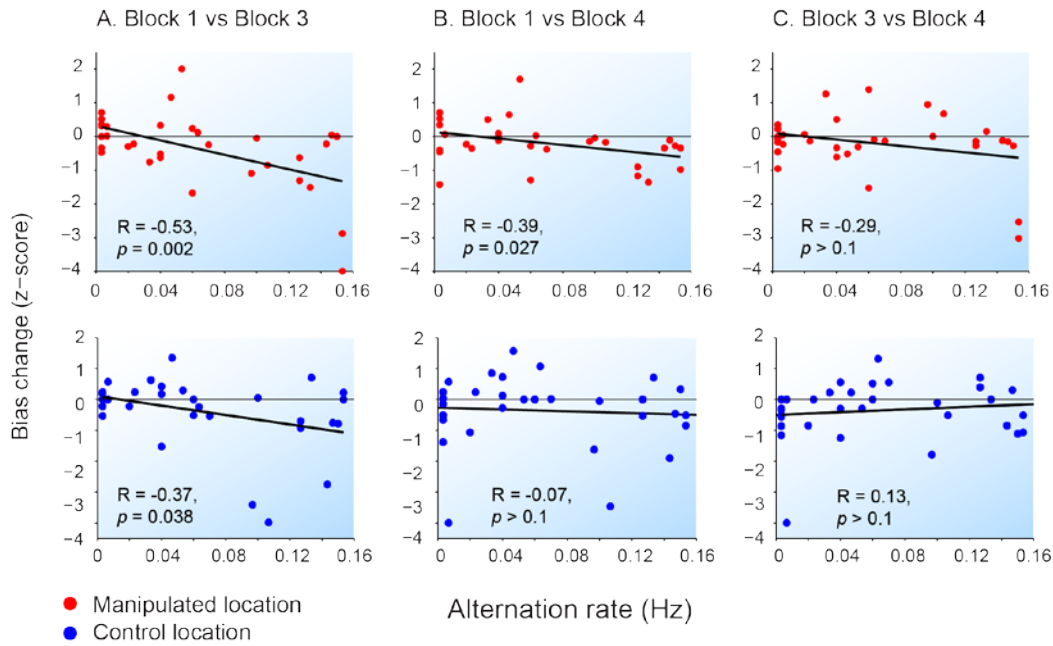
for block 1 (day 1) in this analysis were positive. **Figure 5.3** shows the mean bias across all observers for each location in each block. At the retinal location that was manipulated in block 2, there was a significant decrease in perceptual bias between blocks 1 and 3 (paired T-test,  $t(31) = 2.84$ ,  $p = 0.0079$ ), whereas at the location that remained blank during block 2, there was a far smaller decrease that was not statistically significant ( $t(31)=1.2$ ,  $p = 0.23$ ). As a result, the difference in bias between retinal locations during block 3 also approached significance ( $t(31) = 2$ ,  $p = 0.054$ ). However, the decrease in perceptual bias at the manipulated location (relative to block 1 dominance) partially recovered 24 hours later (block 1 of day 2): although bias remained significantly less than in block 1 ( $t(31) = 2.15$ ,  $p = 0.04$ ), it did not differ significantly from the control location ( $t(31) = 1.35$ ,  $p = 0.19$ ).



**Figure 5.4.** Scatter plots of perceptual bias for clockwise rotation (converted to Z units) during **A**) block 1 (x-axis) versus block 3 (y-axis) of each day. The left column (red markers) shows data from the ‘disrupted location’ – the retinal location that was manipulated during block 2 of day 1. The right column (blue markers) shows data from the control location – the retinal location that remained blank during block 2 of day 1. **B**) Comparisons of bias across days. Biases in block 1 of day 2 (y-axes) are plotted against biases from day 1 block 1 (upper plots) and block 3 (lower plots). Regression lines (grey) indicate high correlation for bias between blocks, with the notable exception of the disrupted location on day 1 (top left), where perceptual experience during block 2 disrupted bias.

We next examined changes in individual observers’ perceptual biases across blocks. Perceptual biases for the first block of each session compared to the last block of each session are plotted for all observers in **Figure 5.4A**, for each retinal location. Data points that lie close to the positive diagonal indicate little change in perceptual bias between blocks. In contrast, points that lie in the upper left or lower right quadrant of each plot indicate a reversal of bias direction between blocks. The results show that reversals in bias were most common at the disrupted location on day 1 (upper left plot), i.e. the manipulation applied in block 2 of day 1 effectively disrupted the trained bias for some observers ( $n = 6$ ). Biases at the manipulated location during day 2 were also affected,





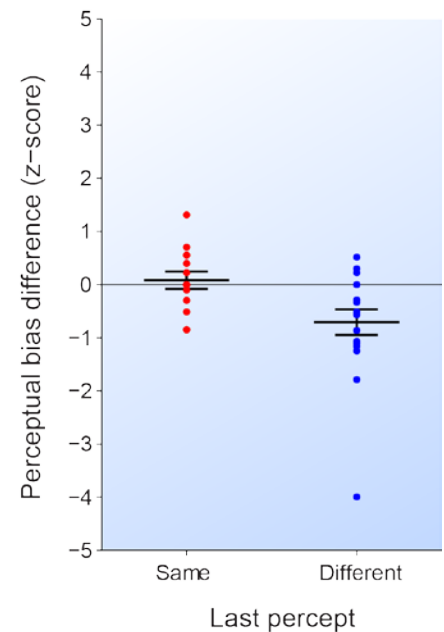
**Figure 5.5.** Scatter plots illustrating the relationship between changes in bias and perceptual alternation rate experienced during block 2 at the manipulated location. The y-axis of each plot shows changes in bias from **A)** block 1 to block 3 of session 1; **B)** block 1 of session 1 to block 1 of session 2; **C)** block 3 of session 1 to block 1 of session 2. The top row of plots shows data for the manipulated location, while the bottom row shows data for the control location. Note that in both cases, alternation rates are based on perception at the manipulated location during block 2, since the control location remained blank during this block. Black lines show regression fits.

compared to biases observed on day 1 both before (block 1) and after (block 3) the manipulation block (**Figure 5.4B**).

These results suggest that the manipulation applied in block 2 had an effect on perceptual bias in some, but not all observers. Following analysis of a similar experiment presented in **Chapter 4**, we investigated the relationship between changes in perceptual bias at the manipulated location and quantitative measures of perceptual experience at that same retinal location during block 2. Linear regression showed a significant relationship between change in bias across blocks 1 and 3 and the rate of perceptual alternations experienced during block 2 at the manipulated location ( $R = -0.53$ ,  $p = 0.002$ ), but also, to a lesser extent, at the control location ( $R = -0.37$ ,  $p = 0.038$ ). In contrast, changes in bias across the first and last blocks on day 2 were not related to perceptual alternation rate observed on day 1 (**Figure 5.5 B & C**). Similarly, the direction of the last percept reported

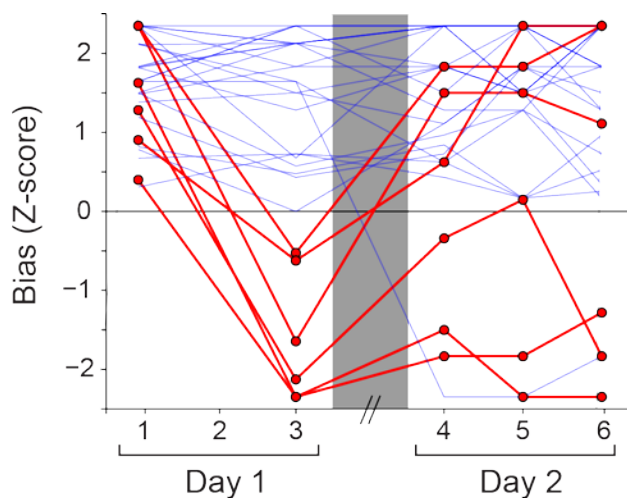
during block 2, relative to block 1 bias, was associated with changes in bias in block 3 (**Figure 5.6**). However, the change in bias for observers who reported the opposite percept (compared to the percept that dominated during block 1) last during block 2 was not significantly greater than for observers who reported the same percept (Same group, mean Z difference =  $0.08 \pm 0.16$  S.E.; Different group, mean Z difference =  $-0.71 \pm 0.25$  S.E.;  $t(30) < 1, p > 0.1$ ).

As expected, the decrease in mean perceptual bias at the manipulated location during block 3 was the result of a reversal in bias direction for some observers ( $n=6$ ) rather than moderate decreases in bias magnitude for all observers. **Figure 5.7** highlights the perceptual bias at the manipulated location for these observers across sessions. Notably, observers who experienced changes in bias direction of smaller magnitude during block 3 were more likely to recover the original (block 1) bias direction on the next day, whereas those with the strongest bias reversal in block 3 retained that bias direction into day 2. Further, those who retained the opposite bias direction during day 2 tended to experience weaker biases during block 1 than they had during block 3. The direction of day 2 bias was therefore predicted by the net perceptual bias across blocks 1 and 3 of day 1 for all but one of the 32 observers.



**Figure 5.6.** Change in perceptual bias between block 1 and block 3, plotted by final percept during the manipulation block relative to block 1 bias.

On day 2, all observers showed a strong resistance to priming from unambiguous trials. One possible explanation for the reduced susceptibility to priming from unambiguous stimuli observed on day 2 compared to day 1 is a change in response times. For a different type of motion stimulus, it has previously been shown that unambiguous priming stimuli can have an inhibitory effect on subsequently presented ambiguous stimuli at short inter-stimulus intervals (ISIs) (<500ms), but become facilitatory at longer ISIs (Kanai & Verstraten, 2005). In contrast, ambiguous stimuli always have a facilitatory effect on perception of subsequent ambiguous stimuli, regardless of the ISI (Brascamp et al., 2009). Since trials are self-paced in the “cue recruitment” paradigm, the average ISI will decrease if observers respond more quickly, and it is plausible that observers might respond more quickly as they become more familiar with the task. Analysis of ISI distributions on day 1 versus day 2 showed a significant difference (day 1 median = 893 ms; day 2 median = 733 ms;  $t(19465) = 17$ ,  $p < 0.001$ ; Supplementary **Figure S5.2**), after exclusion of outliers (RTs > 6s). However, taking into account that stimulus location was randomized across trials, the ISI for a given location was, on average, twice the duration, plus one stimulus on duration (1.5s), giving median ISIs of 2.090s and 1.850s for days 1 and 2 respectively.



**Figure 5.7.** Perceptual bias at the manipulated location for all observers across blocks. The bold red traces are for 6 observers who exhibited the opposite perceptual bias direction in block 3 to the bias they had exhibited in block 1. The original (block 1) bias is seen to recover in half of these observers on day 2. Biases for individuals who did not experience a reversal of bias direction between blocks 1 and 3 are plotted in light blue. Note that only one of these observers experienced a reversal of bias between the last block of day 1 and the day 2.

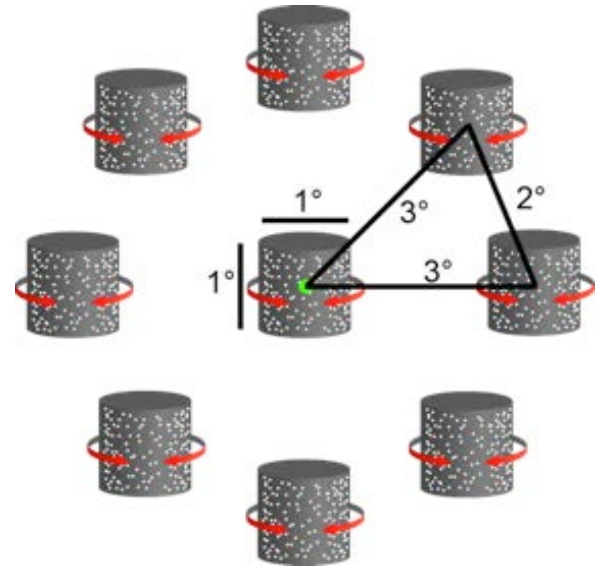
#### *5.4.2. Experiment 2 – Dissociating pre-existing perceptual bias from stabilization*

When a perceptually ambiguous stimulus is presented intermittently for brief periods, perception at each stimulus onset tends to stabilize on a single percept, resulting in a strong perceptual bias (Leopold et al., 2002; Brascamp et al., 2008). This perceptual stabilization is specific to the retinal location of the stimulus, and does not transfer between ambiguous stimuli separated by greater than 2° of visual angle (Chen & He, 2004; Knapen et al., 2009; Carter and Cavanagh, 2007; Stanley et al., 2011).

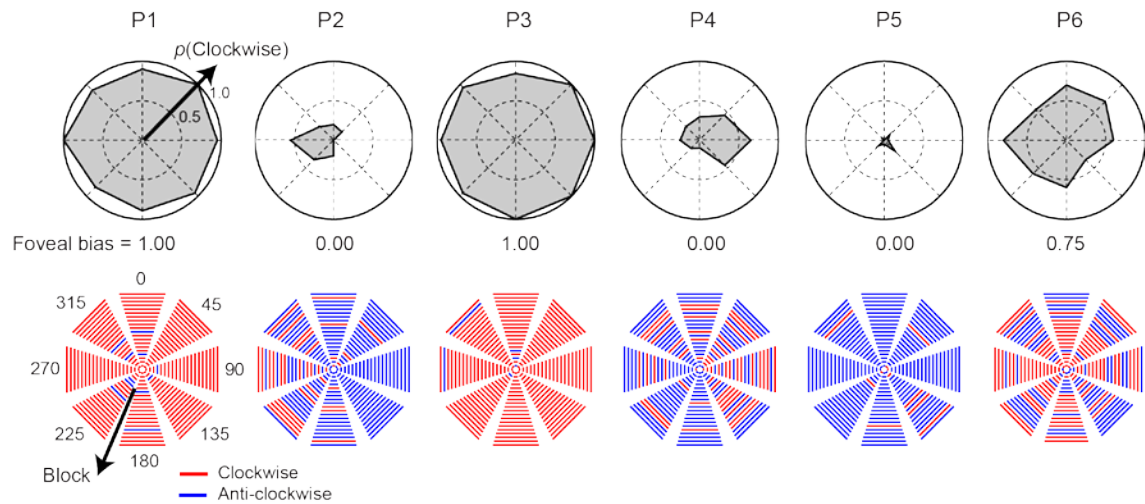
Several studies of the spatial extent of perceptual stabilization have also posed the question of how pre-existing biases (biases that arise prior to and independently of stabilization) vary across the visual field and across individual observers (Knapen et al., 2009; Carter and Cavanagh, 2007). They reported that pre-existing biases are idiosyncratic across the visual field and between observers, but stable across time within individuals. However, the maximum durations between consecutive presentations at any given peripheral retinal location in these studies were 48s (Knapen et al. 2009) and 100s (Carter & Cavanagh, 2007). Inter-stimulus intervals (ISIs) in this range have previously been shown to promote perceptual stabilization (Leopold et al., 2002; Brascamp et al., 2008). Carter and Cavanagh (2007) retested participants 2 weeks after the initial session and reported that perceptual bias remained relatively stable over time for each location for each observer. However, based on the timing of presentations used, only the first trial of each session was truly indicative of pre-existing bias, whereas all subsequent trials were likely to have been influenced by short-term stabilization.

We were interested to investigate the stability of long-term bias over time and retinal location in order to understand its relation to short-term stabilization effects. We therefore replicated the methods of Knapen et al. (2009), with several modifications. First, we reduced the total number of retinal locations to 9, and increased the radial distance between central and peripheral stimuli to  $3^\circ$ . This meant that the distance between adjacent peripheral stimuli was  $2^\circ$ , which is sufficient separation to minimize transfer of stabilization between retinal locations (Knapen et al., 2009; Chen & He, 2004). Random

dot kinematograms of rotating cylinders were presented binocularly (as with the majority of studies of such stimuli, but unlike in Experiment 1 and previous “cue recruitment” studies). This introduces a conflict between disparity and structure-from-motion cues to depth, causing the cylinders to appear flatter, but does not bias perceived direction of rotation.



**Figure 5.8.** Stimulus geometry for Experiment 2. Structure-from-motion cylinders with ambiguous direction of rotation about vertical axes were presented at one of 9 locations on each block, while observers maintained central fixation.



**Figure 5.9.** Individual observers' perceptual biases across retinotopic space and time. A) Polar plots on which the radial axes indicate the probability that an observer perceived the stimulus as rotating clockwise, while the polar angles represent the 8 peripheral stimulus locations. The probability of perceiving clockwise rotation at the central fixation location is given beneath each observer's plot. B) Polar plots on which the radial axes indicate block number and polar angles represent the 8 peripheral stimulus locations. Red bars indicate that the observer perceived the stimulus as rotating clockwise on that trial, while the blue bars indicate anti-clockwise perceived rotation.

Second, we increased the duration between consecutive trials at each retinal location to a minimum of 15 minutes. In order to collect data with such long inter-trial intervals, trials were self-paced and completed by trained psychophysical observers naïve to the goals of the experiment, whilst sat at their office computers (running Psychtoolbox), over the course of 1 week. On each block, one stimulus was presented at each location, in a randomized order and observers reported their first percept when each stimulus appeared. Consecutive blocks were completed anywhere between 15 minutes and 24 hours after one another (mean = 105 minutes).

The results of Experiment 2 are plotted in **Figure 5.9**. The most striking feature is that, although patterns of bias varied considerable across observers, within observers perceptual biases were generally consistent across retinal locations and time (e.g. P1, P3, P5). Several observers (P2 and P4) exhibited perceptual bias that was anisotropic across retinal space. This observation is in line with previous reports of esoteric spatial bias that

was consistent within observers across time (Carter and Cavanagh, 2007;Knapen et al., 2009a). Further, bias at the foveal location tended to correlate with bias at the tested locations across the rest of the visual field.

### 5.3.3. *Eye of origin specificity*

Previous studies of induced perceptual biases for rotating Necker cube stimuli have presented unambiguous stimuli binocularly in order to provide disambiguating disparity depth cues, although it should be noted that monocular disambiguating cues alone are also effective at inducing a perceptual bias (Harrison and Backus, 2010a;Harrison et al., 2011). However, in the context of cue recruitment, ambiguous stimuli are presented monocularly. This removes binocular disparity cues, which would otherwise indicate a flat depth structure and conflict with the depth structure signaled by structure-from-motion cues. The presentation of ambiguous stimuli is important because perceptual biases are longer lasting and more resistant to reverse learning when the proportion of ambiguous trials is greater, which suggests that learning occurs during ambiguous rather than unambiguous trials (Harrison and Backus, 2010b;van Dam and Ernst, 2010). If the neural locus of this learning is low-level, then it is possible that the acquired perceptual bias is specific to the eye to which ambiguous trials are presented (Karni and Sagi, 1991).

Harrison and Backus (2010a) previously tested ocular transfer of perceptual bias across consecutive days of testing by presenting both ambiguous and unambiguous stimuli monocularly, and switching eye of presentation between sessions. They found no significant effect of switching presentation between eyes and concluded that the representation of retinal location as a recruited cue was binocular. In contrast, Maier et al. (Maier et al., 2003) alternated presentation of an intermittent bistable stimulus between

each eye on consecutive presentations, and observed a reduced co-occurrence of perceptual reversals compared to when the same intermittent stimulus was presented binocularly. Together, these findings suggest that long-term perceptual bias, but not short-term perceptual stabilization, fully transfers between eyes. This implies that these two processes are implemented by distinct neural mechanisms, at binocular and monocular levels of encoding respectively.

Eye-of-origin information has also previously been shown to be important for stabilization of binocular rivalry stimuli, where either the percept (high-level binocular representation) or the eye-of-origin (low-level monocular input) can be stabilized (Chen and He, 2004; Pearson and Clifford, 2004; Sandberg et al., 2011). Switching the images presented to the two eyes during a blank interval of intermittent binocular rivalry can result in either stabilization of one eye's input, or stabilization of one image regardless of which eye it is presented to (Logothetis et al., 1996; Sandberg et al., 2011). Given the well-documented similarity in perceptual time courses of binocular rivalry and ambiguous stimuli, it is probable that perception of bistable stimuli recruits neural mechanisms at multiple levels of organization.

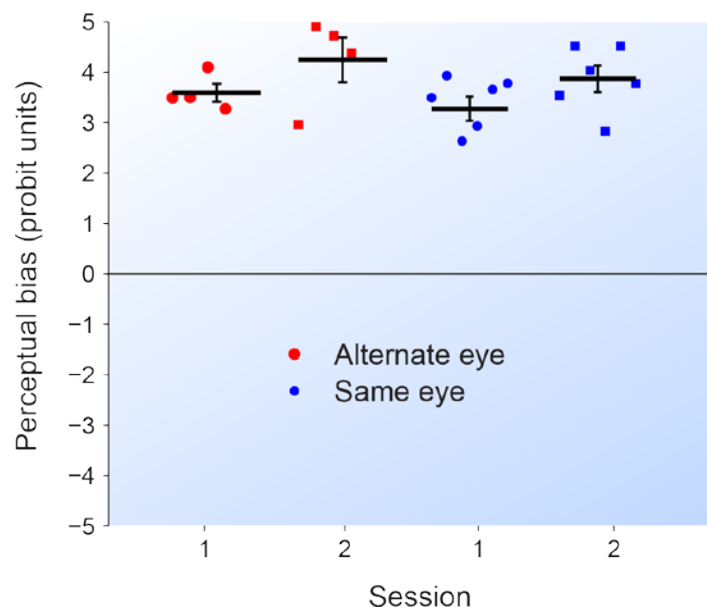
In the present experiment, we therefore examined whether long-term perceptual biases induced for ambiguous stimuli presented to one eye were maintained when presentation was later switched to the other eye. In a single experimental session, 10 participants performed 2 blocks of 200 trials each, separated by a 5-minute break. Four of the first 8 trials in the first block were disambiguated, while the rest were ambiguous and were presented monocularly to either the subject's left or right eye (counterbalanced across participants). In the second block, all trials were ambiguous and were presented monocularly to either the opposite eye ( $n=4$ ) or the same eye ( $n=6$ ) as the previous block.



For each block, saturated values (probabilities of 0 or 1) were substituted with maximum z-scores equivalent to one response in the opposite direction to training per block ( $\pm 2.576$ ).

## Results

The perceptual bias observed in block 1 was not significantly different from the bias observed for a second block, for which monocular presentation was switched to the opposite eye (**Figure 5.10**; Block 1 mean =  $3.59 \pm 0.18$  SE; Block 2 mean =  $4.24 \pm 0.44$ ;  $t(3) = -1.53$ ,  $p = 0.22$ ). The perceptual bias induced during the first block was therefore not specific to the eye of origin, since bias acquired during presentation of ambiguous trials to one eye was maintained when presentation was switched to the other eye. This suggests that changes in neural processing associated with the induced perceptual bias occur at a stage in the visual processing stream after information from the two eyes has been combined, and is in agreement with previous findings (Harrison and Backus, 2010a).



**Figure 5.10.** Perceptual biases for all observers in Experiment 3. For the alternation group (red markers), retinal location-specific perceptual biases were trained in one eye (session 1) and tested in the other eye (session 2). For the same eye group (blue markers), training and testing stimuli were presented to the same eye across sessions.

## 5.4. Discussion

### *5.4.1. Long-term bias*

Experiment 1 was designed to measure how learned retinotopic biases for perceptual interpretation of ambiguous visual information evolve over time. Specifically, we measured perceptual biases (either trained or naturally elicited) at two retinal locations over a period of 24 hours. Perceptual experience was manipulated between training and testing at one of the two locations, through the continuous presentation of the stimulus, which elicited varying rates of perceptual reversal across observers. On day 1, biases were strong during ‘training’ and testing blocks that involved intermittent stimulus presentations, irrespective of whether observers received explicit training in the form of unambiguous priming trials. However, the magnitude of bias was markedly reduced at the manipulated location during subsequent testing, replicating the results of a previous experiment (Chapter 4). Additionally, we found that differences between observers in the rate of perceptual alternations experienced during the manipulation block predicted subsequent disruption of recently acquired biases.

Biases measured 24 hours later tended to follow those experienced during the last block of the previous session. For the manipulated location, day 2 bias was therefore more strongly correlated with the last perceptual experience (24 hours earlier) than with the bias that had initially dominated during training. This maintenance of bias across days occurred despite the presentation of unambiguous trials on day 2 that were intended to prime perception in the opposite direction to day 1 bias. This resistance to counter-conditioning of recently learned biases has been reported previously (Haijiang et al., 2006; Harrison and Backus, 2010b), and was highly robust in our data and could not be accounted for by changes in inter-stimulus interval between days. Although the small number of

unambiguous training trials presented at the start of day 1 (2 per location) were not always effective in priming perception of subsequent ambiguous stimuli, previous studies suggest that such priming can be highly effective (Long et al., 1992; Kanai and Verstraten, 2005; Long and Moran, 2007), making it all the more remarkable that the many unambiguous trials on day 2 (90 per location) had virtually no effect in counteracting previously acquired biases for ambiguous stimuli.

The experience of perceptual alternations resulting from continuous viewing of the stimulus at the manipulated location was sufficiently detrimental to existing bias as to cause reversals in the direction of perceptual bias for a subset (17%) of observers during day 1. Interestingly, the original bias recovered on day 2, only for those observers with the weakest bias magnitude at the end of day 1. In contrast, those with strong reversed biases at the end of day 1 retained this new bias direction on day 2. In other words, net perceptual experience integrated across the session on day 1 best accounted for perceptual bias 24 hours later. This result stands in contrast to perceptual bias on the shorter timescale, where most recent perceptual experience appeared to best predict bias in the next block. This fits with suggestions based on a study of adaptation to unambiguous visual stimulation, which proposed that, in contrast to short-term effects of adaptation on perception, adaptation over longer periods biases perception towards the statistics of the remote past .

The retinal location-specificity of perceptual bias can be understood in terms of local activity in different neural populations within retinotopically organized cortex (Blake et al., 2003; van Boxtel et al., 2008). For example, it has been shown that a rivalrous stimulus moving continuously across the retina will be perceived to reverse less frequently than the same stimulus presented at a constant retinal location, due to local adaptation (Blake et al., 2003). Similarly to the retinotopic biases reported for ambiguous stimuli, a previous study of ambiguous depth-from-motion stimuli reported biases toward perception

of downward or rightward motion as the front surface in transparent opponent motion stimuli that were presented foveally (Mamassian and Wallace, 2010). Notably, retinotopic space, motion direction, and depth order are all visual properties that neurons in motion sensitive cortical area hMT+/V5 are selective for. Intermittent presentation of ambiguous structure-from-motion motion stimuli has been shown to reduce the variability in the firing rates of MT neurons (Klink et al., 2011), and might therefore bias perception through predictive coding at a population level. Neuroimaging studies of perceptual stabilization suggest that unlike the suppressive effects of stimulus repetition on neural activity, perceptual repetition increased activity in early stimulus-specific visual areas (de Jong et al., 2012b; de Jong et al., 2014), as well as involving parietal cortex (Sterzer and Rees, 2008).

#### *5.4.2. Pre-existing biases remain stable without stabilization*

In Experiment 2, we tested natural (i.e. untrained) perceptual biases across the visual field over the course of a one-week period. Importantly, we enforced long intervals between consecutive trials at each retinal location (average > 1 hour) in order to minimize the contribution of short-term perceptual stabilization that is known to result from intermittent presentation of ambiguous stimuli. We observed consistent perceptual biases within observers, despite having no control over the visual experiences of observers during the long ISIs. However, in contrast to previous reports for binocular rivalry stimuli, our observers generally showed the same direction of perceptual bias across visual field locations.

Notably, Maier and colleagues (2003) interleaved presentations of either unrelated stimuli or different bistable stimuli, between intermittent presentations of an ambiguous

rotating sphere, and found no reduction in stabilization. This suggests that perceptually irrelevant visual signals do not impair perceptual memory or disrupt esoteric pre-existing biases, just as they appear not to impair visual short-term memory (Andrade et al., 2002). In contrast, allocation of attentional resources during a brief ( $< 3s$ )—potentially critical—period between consecutive intermittent presentations of rivalrous stimuli have been shown to reduce stabilization (Kanai & Verstraten, 2006). The visual stimulation that participants in Experiment 2 experienced between consecutive trials—generated by getting up from their computer, going for lunch, watching television, and even sleeping—may have had no degrading effect on the perceptual memory trace that supports stabilization of the last percept experienced *for the specific context/ stimulus type* in question.

No previous study has examined the effect of long ISIs on onset bias, as used here (the maximum reported previously was 100s by Carter and Cavanagh, 2007). Consequently, we cannot be certain whether our manipulation of introducing long ISIs effectively reduced the contribution of perceptual stabilization to the observed biases, despite stabilization being generally considered a “short-term” phenomenon due to its rapid onset. Existing models of perceptual stabilization are based on the gradual decay of a perceptual memory trace for the stabilized percept over the course of each ISI (Noest et al., 2007; Brascamp et al., 2009). Although not empirically verified, an implication of this model is that over a sufficiently long ISI, traces for both competing percepts will return to baseline, with the difference between baseline states of the neural populations representing the two competing percepts governing onset bias. Perceptual stabilization is most effective if the duration of stimulus presentation ( $T_{on}$ ) is shorter than the mean dominance duration for continuous viewing. This duration varies between stimulus type, stimulus parameters, and observers, so to ensure stabilization, previous studies have tended to use short stimulus-on durations ( $T_{on} \sim 1s$ ). To determine the potential persistence of stabilization

across long ISIs, stimulus presentation on each trial should continue after observers report their initial percept, until the observer experiences perceptual reversals. Repeated alternations have been shown to diminish the strength of the perceptual memory trace for the percept selected at initial onset (Brascamp et al., 2008), while an odd number of reversals would also cause the dominant percept at stimulus disappearance to be opposite to that perceived at stimulus onset (Pastukhov and Braun, 2008).

Visual adaptations to various visual features, including the gender of a face, and orientation of a line, have previously been shown to operate retinotopically (Afraz and Cavanagh, 2009; Knapen et al., 2009b). Further, when observers are asked to discriminate visual features (e.g. the gender of a face) for stimuli that are objectively ambiguous (e.g. a 50% male-female morph), they exhibit perceptual biases that, like the patterns observed here, are esoteric across observers, heterogeneous across the visual field, but stable across time (Afraz et al., 2010). Afraz and colleagues also found that the magnitude of perceptual biases decreased with increasing stimulus size, leading them to propose that the heterogeneity of perceptual bias across the visual field is the result of undersampling of the visual signal at small spatial scales relative to receptive field sizes in the neural populations that encode the visual attribute. In other words, larger stimuli activate larger pools of individually biased cortical neurons, and the averaging of their responses reduces bias.

#### *5.4.3. Ocular transfer*

Retinotopic location is a low-level property of the visual input signal that is encoded throughout visual the visual system. Given its importance for both short-term stabilization and long-term perceptual biases studied here (in addition to its importance in a variety of other perceptual processes), we were interested to establish just how ‘low-level’ these

processes might be in terms of where in the visual system they are represented. We therefore tested whether perceptual bias for ambiguous 3D motion stimuli is specific to the eye in which it is presented during training. Our results show that trained biases are not eye-specific, and thus confirms that their neural locus must be within or beyond primary visual cortex.

A previous study of ocular transfer for cue recruitment used monocular training trials and reversed location-rotation contingencies, resulting in a significant decrease in perceptual bias between training and testing phases, even when the eye of presentation was not switched (Harrison and Backus, 2010a). However, they noticed a trend towards greater retention of the trained perceptual bias on day 2 when the eye-of-origin was not switched. In contrast, our experimental protocol using binocular training trials and no further unambiguous trials during block 2 resulted in greater retention of the perceptual bias, indicating without doubt that trained perceptual biases were not eye-specific.

#### *5.4.4. Conclusion*

Trained perceptual biases that are specific to retinal location have previously been shown to persist across days and be robust to counter conditioning, although these biases can also emerge naturally without explicit training and can be disrupted through experience of spontaneous perceptual alternations, as we showed in the previous chapter. The current study confirmed that at short time scales, future perception tends to be best predicted by recent perceptual experience. However, at longer time scales, future perception was dependent on the integration of experience across longer periods of perceptual history. In the absence of training or disruption, and with minimal contribution from short-term stabilization, individual observers exhibited pre-existing perceptual biases that were steady

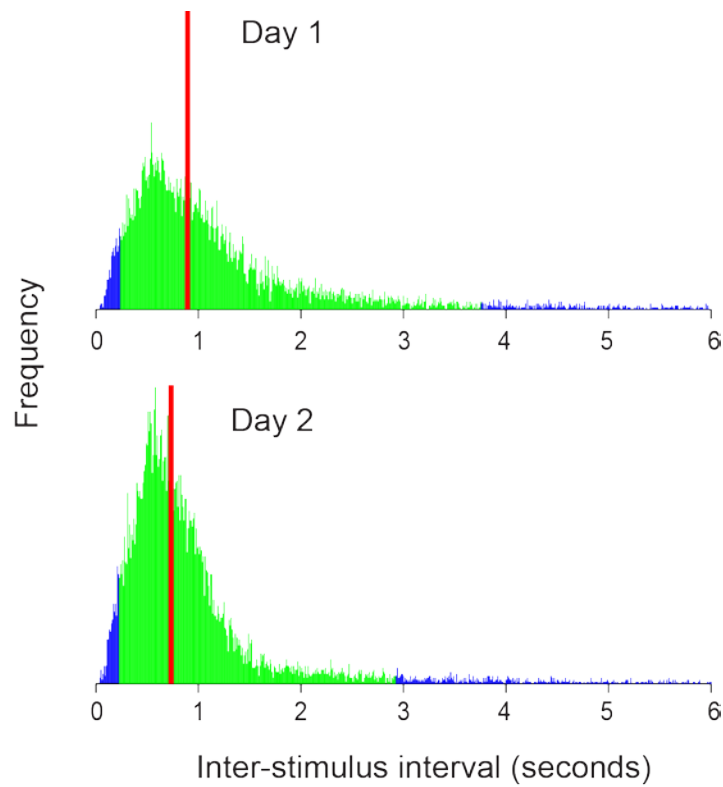
across time and retinotopic space, while differing between individuals.



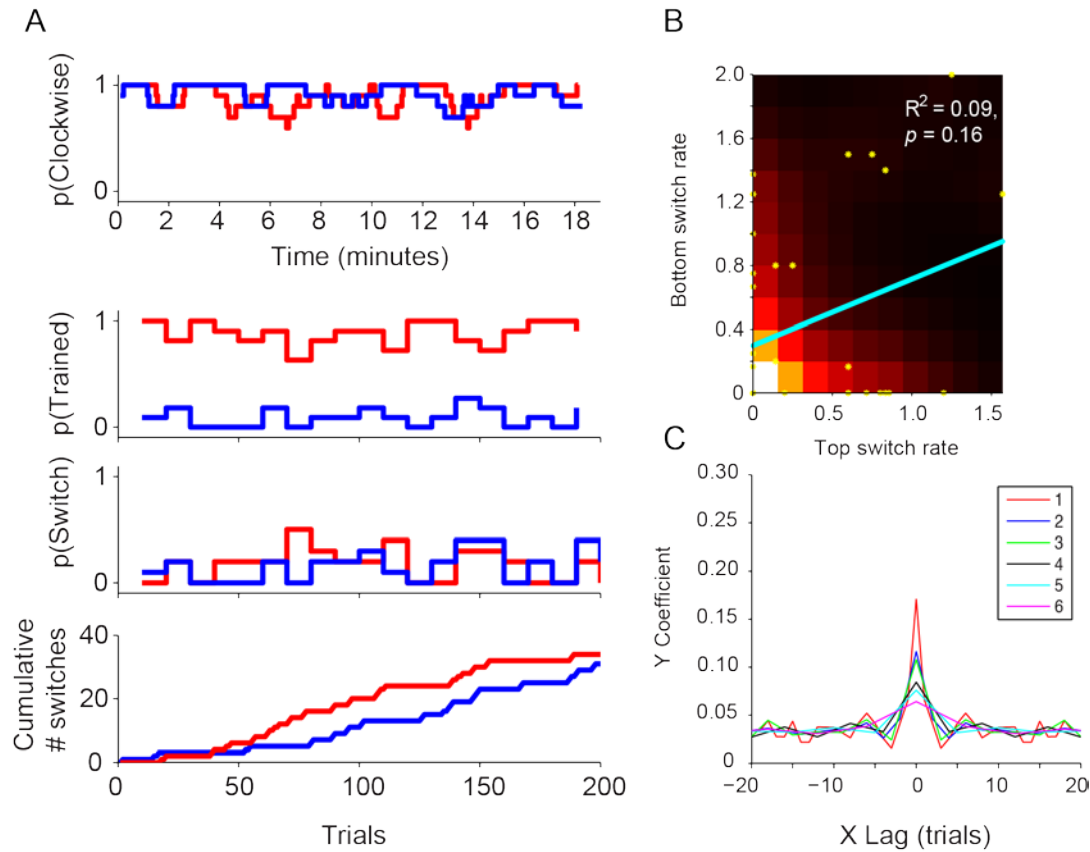
5.5. Supplementary materials



**Figure S5.1.** Individual observers' data from Experiment 1 (n= 32). To enhance clarity, raw data were boxcar filtered using a sliding window of 8 trials. Green dots represent the persistence of block 1 bias at the manipulated location into block 3 (left column) or day 2 (middle column). Red dots indicate a reversal of bias direction at the manipulated location between blocks, while grey dots indicate no clear bias in either direction.



**Figure S5.2.** Distribution of inter-stimulus interval durations (either location). Since trials were self-paced, ISI was variable. However, perceptual stabilization of ambiguous stimuli is known to be influenced by ISI. The distributions shown here indicate that observers were less likely to allow longer ISIs on day 2 of testing, resulting in a small decrease in median ISI (red lines) across days.



**Figure S5.3.** Additional analyses of example data from one observer. **A)** Binned data (width = 10 trials) showing the time course of perceptual bias at each location (red and blue traces), which can be defined as either the probability of clockwise rotation or trained direction of rotation. The third row shows the probability of a perceptual alternation occurring at a given time, and in some cases the trace for the two locations appeared to be correlated. **B)** Plotting the alternation rate within a time bin (in this case, 30 seconds) for the top retinal location versus the bottom retinal location revealed no significant correlation between alternation rates for any observers. **C)** Further, testing for correlations between switch rate at different lag times and different bin widths did not reveal any significant correlations.

**Table S5.1.** Details of individual displays used for stimulus presentation in Experiment 2 ( $n = 6$ ). Stimulus dimensions in pixels were scaled to maintain constant size in visual angle across different display setups.

Observer	LCD monitor model	Screen resolution	Physical dimensions (cm)	Refresh rate (Hz)	Viewing distance (cm)
P1 DS	Samsung SyncMaster 913B	1280 x 1024	37.6 x 30.4	60 Hz	65
P2 APM	Samsung SyncMaster 2233RZ	1680 x 1050	47.4 x 29.8	120 Hz	51
P3 AM	Samsung SyncMaster 2493HM	1920 x 1200	52.0 x 33.0	60 Hz	60
P4 HB	Sony Vaio	1440 x 900		60 Hz	40
P5 TH	Samsung SyncMaster	1280 x 1024		60 Hz	
P6 VP	Samsung SyncMaster	1280 x 1024		60 Hz	

**Table S5.2.** Summary of visual properties influencing long-term perceptual biases and short-term perceptual stabilization.

Stimulus property		Long-term perceptual bias		Short-term perceptual stabilization	
•	Retinal location	✓	(Haijiang et al., 2006; Harrison and Backus, 2010b)	✓	(Chen and He, 2004; Knapen et al., 2009a)
•	Spatial location	✗	(Backus and Haijiang, 2007; Harrison and Backus, 2010a)	✗	(Knapen et al., 2009a)
•	Eye-of-origin	✗	(Harrison and Backus, 2010a)	✗	(Maier et al., 2003)
•	Object shape	✗?	(van Dam and Ernst, 2010)	✗	(Maier et al., 2003)
			*Measurement did not separate long-term and short-term effects.	✓	(Pastukhov et al., 2013)
		✓?	(Harrison and Backus, 2012) *A small bias was only observed after monocular but not binocular training.		
•	Motion direction	✓?	(Mamassian and Wallace, 2010)	✓	(Maier et al., 2003; Pastukhov and Braun, 2013)
<b>Other features</b>					
•	Immediate onset	✓	(van Dam and Ernst, 2010)	✓	(Orbach et al., 1963; Leopold et al., 2002)
•	Influences percept selection, but not percept maintenance	✓	(Carter and Cavanagh, 2007; Manousakis, 2012)	✓	(Orbach et al., 1963; Leopold et al., 2002)

## Chapter 6: The role of parietal cortex in the perception of depth from motion and stereopsis.

---

In order to execute adaptive behaviours, the brain must construct a spatial representation of the world from the available sensory information. Posterior parietal cortex appears to play an important role in the inference process, including the perception of depth from binocular disparity and motion cues, as well as perceptual interpretation of ambiguous sensory input. However, much of the evidence for parietal involvement in these functions comes from correlational methods such as human brain imaging. To investigate whether parietal regions play a causal role in these visual processes, we tested patients with occipito-parietal lesions on a series of stereoscopic and bistable perception tasks. The bistable tasks measured changes in the perception of 3D structure-from-motion and other rivalry stimuli over time. The stereo tests involved discrimination of depth from random dot stereograms containing coarse and fine disparities, as well as disparity-defined dynamic objects and complex contoured surfaces. Patients with unilateral lesions of left PPC were able to discriminate depth from dynamic and static binocular disparities with sensitivity comparable to that of healthy controls. However, all patients with lesions to right PPC showed poor sensitivity to binocular depth cues, supporting a causal role for the right PPC in stereopsis. In the bistability tasks, the results were similarly heterogeneous: some patients reported perceptual alternations at rates comparable to healthy controls, while several patients with bilateral parietal lesions showed noticeably faster alternations rates than other observers. Despite this, all patients showed reduced rates of perceptual alternation when stimuli were presented intermittently. These results support the idea that the continuous and constructive process of perceptual inference of 3D structure is strongly mediated by posterior parietal cortex, and that this process is lateralized to the right hemisphere in the case of stereopsis.

## 6.1. Introduction

In order to execute appropriate motor responses, the brain must integrate and interpret sensory information to form some internal representation of the environment. Posterior parietal cortex (PPC) is thought to play an essential role in this task, as supported by its involvement in a host of perceptual, cognitive, and motor functions (Cohen and Andersen, 2002; Corbetta and Shulman, 2002; Hubbard et al., 2005; Freedman and Assad, 2006; Cabeza et al., 2008; Sereno and Huang, 2014). As the apex of the dorsal visual pathway, PPC is particularly important for encoding spatial information that is essential for planning visually-guided actions, including saccades, reaching, and grasping (Gottlieb, 2007). The importance of PPC for perception is most strikingly revealed by the neuropsychological condition of spatial neglect, in which damage to PPC causes deficits of both attention and awareness in the contralateral hemifield. While such deficits in the perception of the 2D visual field reflect the topographic functional organization of parietal cortex (Schluppeck et al., 2005; Silver et al., 2005), it is less clear how parietal damage affects perception of 3D space and objects within it.

Here, we examined the causal contribution of PPC to the perception of 3D depth by testing patients with parieto-occipital lesions on two types of depth cue that elicit qualitatively different perceptual experiences. First, we tested the perception of depth from binocular disparity (stereopsis), which gives rise to a robust and stable impression of depth. Second, we tested perception of depth from motion, which is an ambiguous depth cue. In this second case, perception of the stimulus is bistable (there are two equally valid interpretations of depth structure), and healthy subjects report spontaneous alternations between rival percepts during viewing.

### *6.2.1. Stereopsis*

Humans and other primates are able to perceive the three-dimensional structure of the world based on subtle differences between the retinal image in each eye. Neural computation of these binocular disparities begins in primary visual cortex (Hubel and Wiesel, 1962;1970b;Cumming and Parker, 1997), but the perception of depth from this information appears to occur at later stages of visual processing (Cumming and DeAngelis, 2001;Parker, 2007). Depth perception is important for both recognizing objects and guiding actions, and correspondingly disparity signals are processed in both the ventral and dorsal visual streams (Neri et al., 2004;Preston et al., 2008), with more complex coding of 3D shape and position occurring in temporal and parietal cortices (Orban et al., 2006b).

The involvement of PPC in processing binocular disparity was initially established through neuropsychological studies of human patients and animals with parietal lobe lesions (Holmes and Horrax, 1919;Rothstein and Sacks, 1972;Danta et al., 1978;Hamsher, 1978;Cowey and Porter, 1979;Ross, 1983;Miller et al., 1999). Most evidence suggests that PPC in the right hemisphere is necessary for depth processing from disparity in the absence of monocular depth cues (Carmon and Bechtoldt, 1969;Benton and Hécaen, 1970;Durnford and Kimura, 1971;Hamsher, 1978;Ross, 1983;Vaina, 1989). Only one study reported lesions of the left parietal cortex to be associated with greater impairment of stereopsis (Rothstein and Sacks, 1972), while another reported no lateralization effect (Lehmann and Wälchli, 1975). In contrast, intact stereopsis has been reported (for certain types of stimulus configuration) following lesions to the left hemisphere (Vaina, 1989) and posterior ventral areas such as lateral occipital cortex (Riddoch and Humphreys, 1987;Read et al., 2010). More specifically, local but not global stereopsis is retained following more anterior temporal lesions (Ptito et al., 1991), while the same pattern of

dissociation is found following posterior inferotemporal lesions in macaques (Covey and Porter, 1979).

More recently, evidence from neurophysiology and functional imaging in both humans and monkeys has provided insights into how binocular disparity information is processed in spatially localized regions of the dorsal visual stream, and the intraparietal sulcus (IPS) in particular. Human fMRI has confirmed that extrastriate regions of the dorsal visual pathway process disparity information (Backus et al., 2001; Tsao et al., 2003; Neri et al., 2004; Welchman et al., 2005; Chandrasekaran et al., 2007; Preston et al., 2008; Georgieva et al., 2009), and these areas feed forward to posterior IPS (Tsao et al., 2003; Durand et al., 2009; Georgieva et al., 2009). IPS regions appear to support more complex depth representations from disparity information, such as selectivity for slanted and curved surfaces (Shikata et al., 2001; Georgieva et al., 2009), disparity-defined contours (Chandrasekaran et al., 2007) and some degree of invariance to position-in-depth (Durand et al., 2009). Some evidence suggests that several parietal areas of humans and macaque monkeys may be functionally homologous (Tsao et al., 2003; Grefkes and Fink, 2005; Orban et al., 2006a), and neurons in each subdivision of macaque IPS have been shown to respond to stereoscopic visual stimuli (Shikata et al., 1996; Taira et al., 2000; Tsao et al., 2003; Genovesio and Ferraina, 2004; Durand et al., 2007; Bhattacharyya et al., 2009; Srivastava et al., 2009; Yang et al., 2011; Theys et al., 2012).

Despite the wealth of correlational evidence, causal relationships between neural activity in specific regions of parietal cortex and stereopsis are yet to be established. To date, neuropsychological studies suggesting a causal influence of the right parietal lobe in stereopsis have classified lesion locations at a coarse spatial scale, leaving their correspondence with functional cortical areas unknown. In the present study, we therefore



used high resolution structural MRI to examine more carefully which regions of parietal cortex are necessary for the perception of depth from binocular disparities.

### *6.2.2. Bistable structure-from-motion*

When observers view a bistable visual stimulus, their perception of the stimulus alternates spontaneously over time between two possible interpretations, while the physical stimulus remains constant. Structure-from-motion (SFM) stimuli are one example of this, in which motion cues lead to the vivid perception of a three-dimensional object (such as a rotating sphere) but the depth order of the objects' surfaces remains ambiguous (Wallach and O'Connell, 1953). The responses of neurons in cortical visual area V5/MT of macaque monkeys have been shown to correlate with the monkey's reported percept when viewing SFM stimuli (Bradley et al., 1998; Dodd et al., 2001), and perceptual biases can be elicited by micro stimulation of these cells (Krug et al. 2013). MT is known to have strong reciprocal connections with parietal areas, (Ungerleider and Desimone, 1986; Cavada and Goldman-Rakic, 1989; Lewis and Van Essen, 2000; Law and Gold, 2008), which allow feedback signals from parietal cortex to bias activity of neural populations in MT. Several lines of evidence suggest that parietal cortex is important for processing complex motion features. Using a different type of bistable motion stimulus (apparent motion), neurons in LIP—but not MT—respond more strongly on trials when monkeys reported perceiving motion in the neuron's preferred direction (Williams et al., 2003). In humans, the perception of SFM elicits greater activation of MT, LOC and various areas within the IPS compared to 2D motion (Orban et al., 1999; Vanduffel et al., 2002), and the involvement of parietal regions in processing SFM appears to be unique to humans. In one previous neuropsychological study, none of the patients with right parietal lesions were able to perceive structure from motion on a single trial, while all patients with left parietal or right

temporal lesions could (Vaina, 1989). In contrast, perception of SFM was retained in one patient with bilateral occipito-parietal lesions (thought to affect white matter connectivity to V5/hMT) who showed severe impairment in basic motion perception tasks (Vaina et al., 1990).

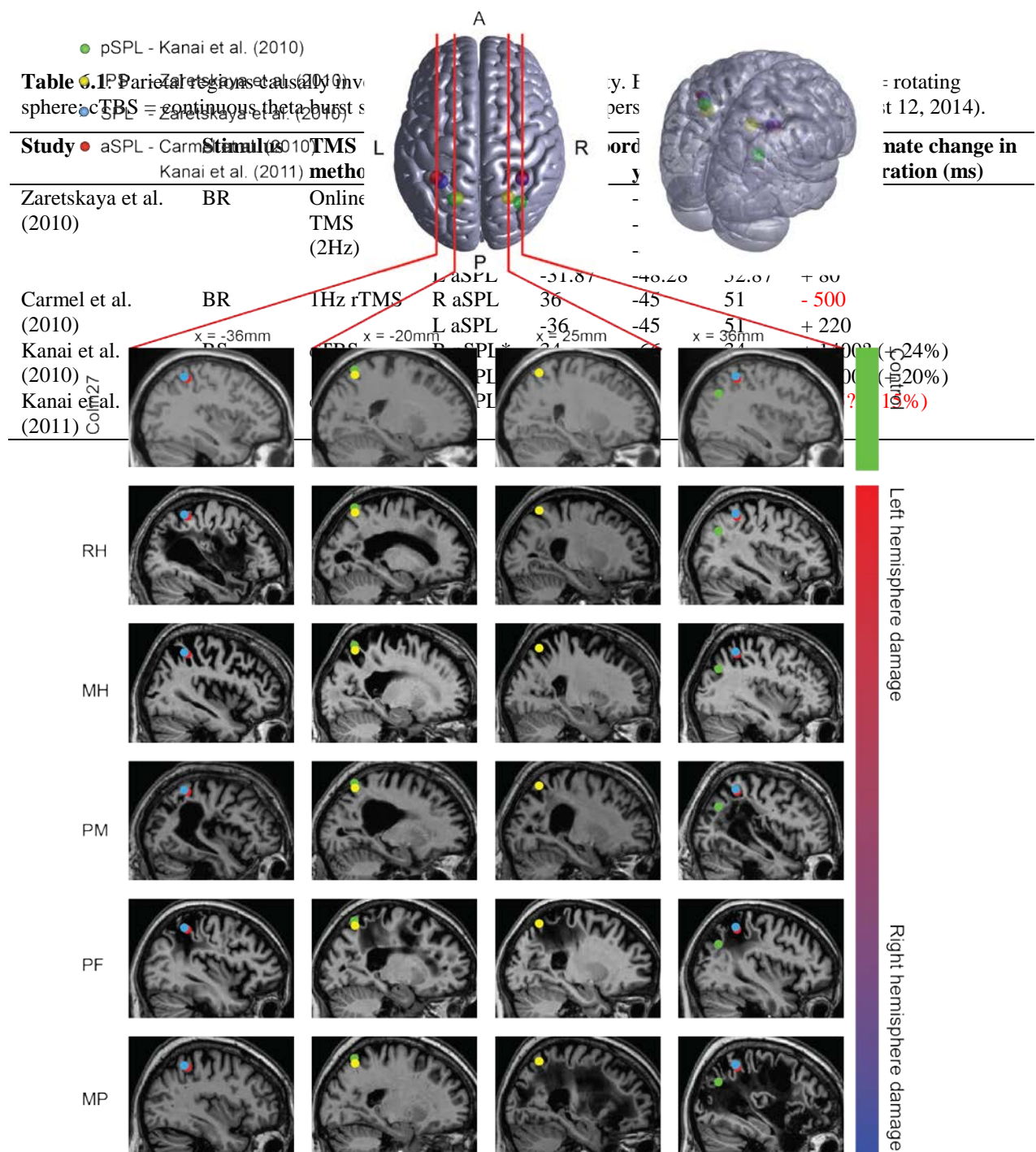
Bistable stimuli have been used in functional brain imaging studies as tools for investigating the neural correlates of conscious perception/ awareness (Lumer et al., 1998; Leopold and Logothetis, 1999; Chalmers, 2000; Blake and Logothetis, 2002; Rees et al., 2002). A wealth of functional neuroimaging data has highlighted the involvement of the fronto-parietal network in perceptual alternations during viewing of bistable figures (Kleinschmidt et al., 1998; Lumer et al., 1998; Blake and Logothetis, 2002; Sterzer and Kleinschmidt, 2007; Britz et al., 2009; Wilcke et al., 2009; Britz et al., 2011). However, this same network is known to be activated during perceptual decision making (Heekeren et al., 2004; Heekeren et al., 2008; Pesaran et al., 2008) and attention demanding tasks (Corbetta and Shulman, 2002), and might therefore reflect introspection or task difficulty as subjects are asked to constantly monitor and report their percept. A recent fMRI study demonstrated that activations of prefrontal cortex previously associated with perceptual alternations, are no longer observed when subjects are not required to actively report their perceptual state, while parietal activations remain (Frässle et al., 2014). Further, patients with lesions of prefrontal cortex show similar rates of spontaneous perceptual alternation to healthy controls, although they are less able to voluntarily facilitate perceptual alternations (Windmann et al., 2006). This evidence suggests that prefrontal cortex does not play a causal role in perceptual alternations during passive viewing, while the necessity of PPC remains an open question.

In addition to the functional evidence implicating PPC in perceptual alternation, structural measures such as grey matter volume, density, and white matter integrity in

certain areas of PPC in healthy observers have been shown to correlate with individuals' perceptual switch rates for bistable stimuli (Kanai et al., 2010). Based on these correlations, several transcranial magnetic stimulation (TMS) experiments have recently demonstrated that the superior parietal lobule (SPL) and intraparietal sulcus (IPS) are causally involved in perceptual alternation (Carmel et al., 2010; Kanai et al., 2010; Zaretskaya et al., 2010). Specifically, anterior and posterior SPL appear to exert opposing effects on perceptual alternation rate, with repetitive TMS of anterior SPL increasing switch rate while stimulation of posterior SPL decreases it (Kanai et al., 2011). However, another study using online TMS (stimulation administered while subjects perform the task) reported conflicting results. They found that stimulating an IPS site located approximately 3mm from Carmel et al.'s 'aSPL' instead reduced switch rate (Zaretskaya et al., 2010), while a second site 6mm from Kanai and colleague's left 'pSPL' produced no significant behavioural effect (**Fig. 6.1** and **Table 6.1**). In all cases the results appear to be partially lateralized to the right hemisphere, although there is considerable individual variability in lateralization (Zaretskaya et al., 2010).

While these conflicting TMS results nevertheless suggest a causal role for PPC in bistable perception, precise interpretation is further complicated by the unknown specificity of the stimulation techniques (Ruohonen and Ilmoniemi, 2002; Ruff et al., 2009). In the present study, we therefore tested patients with circumscribed lesions to PPC that could be objectively measured using high resolution structural MRI. We also tested the same patients' ability to perceive depth from binocular disparity, since the majority of previous neuropsychological studies implicating the right hemisphere in stereopsis did not provide sufficient information to determine the loci of patients' lesions within that hemisphere. Our approach was to focus on a small sample of patients with damage to a range of parietal regions, for whom detailed neurological and behavioural information has

previously been collected, and to test their depth perception using rigorous psychophysical methods.



**Figure 6.1.** A) Parietal target areas reported by previous TMS studies to play a causal role in perceptual alternations. B) Reported MNI coordinates (listed in Table 6.1) are plotted on sagittal slices (one column per medio-lateral position) for each patients' structural MRI (rows), which were spatially normalized to MNI152 space. (aSPL = anterior superior parietal lobule; pSPL = posterior superior parietal lobule; aIPS = anterior intraparietal sulcus). For comparison, the top row shows the T1 of a healthy individual (Colin27) also normalized to MNI152 space (Holmes et al., 1998).

## 6.2. Methods and Materials

### *Participants*

Five patients (4 male and 1 female) were recruited from the pool of neuropsychological volunteers established in the Behavioural Brain Sciences Centre at the School of Psychology, University of Birmingham, and had previously participated in the Birmingham Cognitive Screen battery (Humphreys et al., 2011). All patients had acquired brain lesions to parieto-occipital cortex (see **Figure 6.1** and **Supplementary figure S6.2** for MR images), and had been previously evaluated for clinical deficits of spatial neglect and extinction as part of the Birmingham Cognitive Screen (Humphreys et al., 2011) (described below and summarized in **Table 6.2**). Patients were classed as having a clinical deficit on the basis of whether their test scores were significantly below those of control participants ( $n = 86$ ) with no history of neurological disease (35 males, mean age 67 years, range 47–88 years). Additionally, two healthy age-matched controls (DC and RC, right-handed males, aged 64 and 65) were tested, and twelve younger healthy adults (6 male, ages 18–30). All experiments were conducted in accordance with the ethical guidelines of the Declaration of Helsinki and were approved by the University of Birmingham STEM ethics committee.

Patient MH is a right-handed, 58 year old former garage owner and cook who suffered an anoxic accident (carbon monoxide intoxication) in 1996. His anatomical MRI reveals diffuse and widespread bilateral sulcal widening that is most pronounced in the left parietal lobe, including the occipito-parietal junction, intraparietal sulcus and angular and supramarginal gyri. Subcortically, MH displays ventricular widening and bilateral lesions of the lentiform nuclei and claustrum. Consistent with the left parietal focus of his lesions, MH exhibits unilateral spatial attention bias with extinction in the right hemifield (Kitadono and Humphreys, 2007). MH also exhibits optic ataxia, but only when reaching

to targets in the right visual hemifield with his right arm (Kitadono and Humphreys, 2007; Rice et al., 2008; Cavina-Pratesi et al., 2013). Functional MRI has previously revealed that MH's motion-selective cortical area V5/hMT+ is intact and responsive (Dent et al., 2010).

Patient RH is a left-handed former plumber who was 77 years old at the time of testing. He suffered a left-hemisphere stroke in 1999 that affected the inferior parietal lobe, superior temporal lobe and angular gyrus. Behaviourally, RH exhibits severe extinction of visual events in the right hemifield during bilateral simultaneous stimulation and made left-sided line bisections (Riddoch et al., 2002; Kitadono and Humphreys, 2007; Snow et al., 2013). Consistent with the extensive damage to his left hemisphere, RH also shows deep dysphasia, dyslexia and dysgraphia.

Patient MP is a left-handed, 64 year old former tool worker who suffered an aneurysm of the right middle cerebral artery in 1992, which resulted in occlusion and subsequent infarction. The resulting lesion to the right hemisphere affected the inferior frontal gyrus, the superior temporal gyrus, the supramarginal and angular gyri and the post-central gyrus. The damage caused mild hemiparesis of his previously dominant left arm, although he retained normal sensation. Perceptually, he shows extinction during bilateral simultaneous stimulation (Riddoch et al., 2002; Kitadono and Humphreys, 2007).

Patient PF is a right-handed, 63 year old woman who suffered a stroke in 1999, which resulted in bilateral lesions to the superior and inferior parietal gyri. Specifically, the left sided lesion extends into the left angular gyrus and the right sided lesion extends slightly into the right postcentral gyrus. There are small lesions in both the left and right putamen which extend into the caudate and another small lesion in the left thalamus. PF exhibits right hemifield extinction and dysgraphia. Patient PM is a left-handed, 69 year old male who suffered a stroke that resulted in right superior temporal, left angular and

**Table 6.2.** Summary of patient clinical and behavioural details.

Initials	Sex/age/ hand	Years post injury	Aetiology	Lesion side	Neglect <sup>a</sup> / Extinction <sup>b</sup>	Stereo thresholds				SFM	Mean percept durations (s)		
						D	F	S-n	Contour		RS	AM	BR
RH	M/77/L	12	Stroke	L	R	+	+	NA	NA	+	9.6	NA	7.3
MH	M/58/R	15	Anoxia	L/B	R	+	+	+	95%	+	3.6	7.0	3.1
MP	M/64/L	18	Aneurism	R	L	-	-	-	NA	+	8.8	NA	8.5
PM*	M/69/L	> 10	Stroke	B	L	-	-	NA	77.5%	+	5.1	5.1	NA
PF*	F/63/R	12	Stroke	R	L	-	-	NA	70%	+	8.2	6.5	11.0

<sup>a</sup> Based on Apples Cancellation test scores; <sup>b</sup> Based on Visual Extinction test scores (Humphreys et al., 2011). NA = Not available. \*Some data for patients PM and PF were collected as part of an M.Res dissertation project. See acknowledgements on page 9 for details.

supramarginal lesions. There is also a small lesion in the left postcentral gyrus, bilateral putamen and caudate.

### *Stimuli*

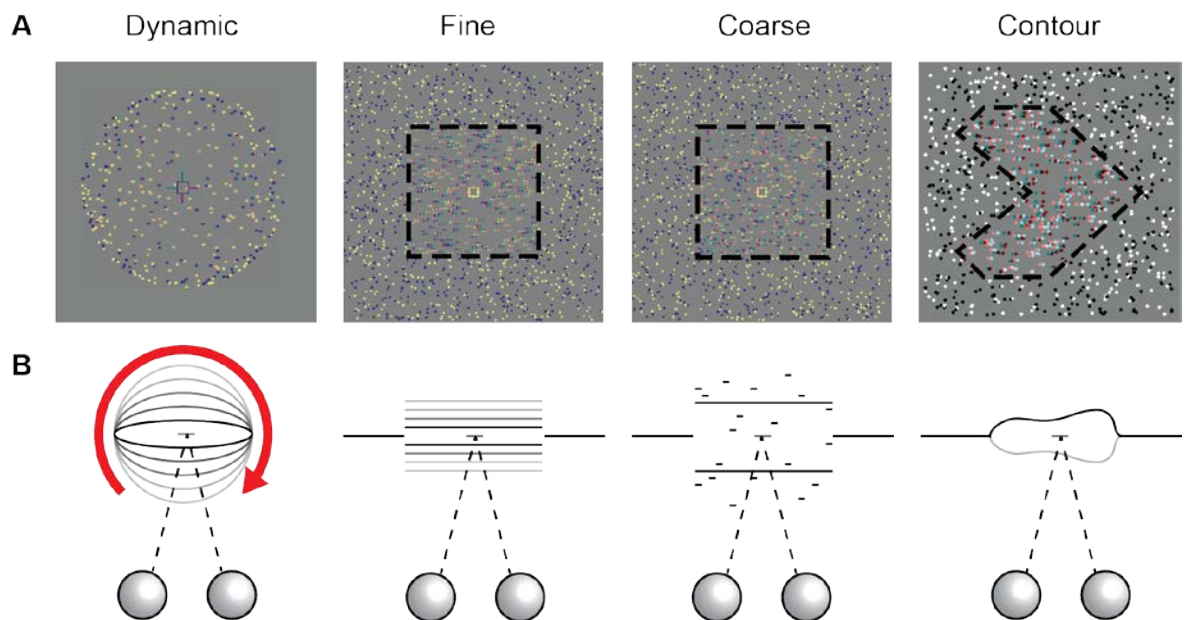
Stimuli were programmed in MATLAB (The MathWorks, Natick, MA) with Psychophysics Toolbox extensions (Brainard 1997; Pelli 1997), and presented binocularly in a Wheatstone stereoscope setup consisting of a pair of ViewSonic P225f CRT monitors (1600 x 1200, 100 Hz) viewed through cold mirrors at a viewing distance of 50 cm. The only exception to this was the disparity-defined contour task, for which stimuli were presented on a single CRT while participants wore red-green anaglyph glasses. Participants' head position was stabilized through use of a chin rest. For the majority of sessions, eye position was recorded binocularly at 1kHz using an EyeLink 1000 video-based eye tracker (SR Research). Photometric measurements were used to calculate linearized gamma tables (Admesy, Ittervoort, The Netherlands) allowing calibration of the two monitors to produce matched luminance outputs. All stimuli were presented centrally on a mid grey-background inside a textured border consisting of black and white squares (75% density, 0.5°x0.5°), which served to promote correct vergence posture. In all tasks, participants gave their responses via a configuration of buttons on a gamepad that was



customized for each patient such that they could respond easily using the middle and index fingers of their preferred hand.

### *Stereopsis tasks*

All random dots stereogram stimuli (Julesz, 1971) consisted of black and white dots ( $0.1^\circ$  radius) on a mid-grey background (**Fig. 6.2**). A fixation marker consisting of nonius lines forming a crosshair was presented on all task except for the disparity-defined contour task, and observers were instructed to maintain fixation during trials.



**Figure 6.2.** Schematic illustrations of stereoscopic stimuli. Four binocular depth perception tasks were presented, using random dot stereograms (A). RDSs are rendered here for red-green anaglyph viewing, and black dotted lines illustrate depth edges but were not present in the stimuli. (B) View-from-above for the depth perception tasks.

- **Dynamic stereo task:** random dot kinematograms (RDKs) depicted a rotating sphere ( $6^\circ$  diameter), defined by structure-from-motion (SFM) and binocular disparity. The sphere consisted of 400 black and white dots (3 arcmin diameter) distributed randomly across the transparent surface, and rotated about a vertical axis at an angular velocity of

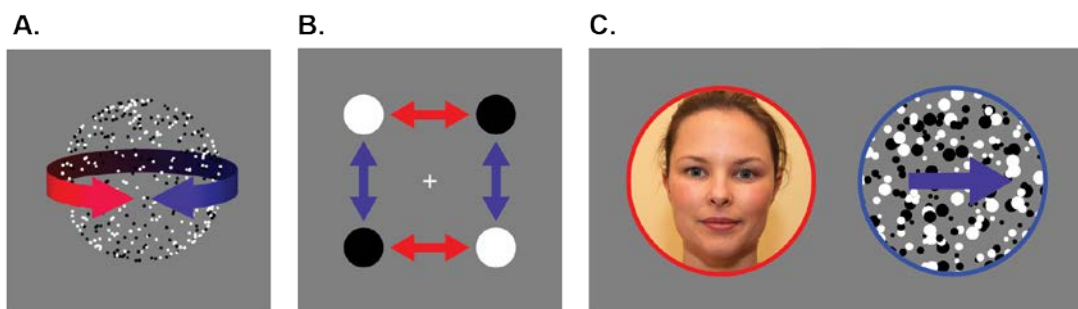
90°s<sup>-1</sup>. Disparity was manipulated parametrically in order to measure psychometric functions, and varied between 0.5 and 14 minutes of arc between the front surface and the fixation plane. Participants were asked to report which direction (left or right) the front face of the sphere was moving in on each trial, and direction as randomized. At smaller binocular disparities, the stimulus becomes bistable, as in the rotating sphere task (see below). On each trial, the stimulus was presented until the observer responded, up to a maximum of 5 seconds.

- **Signal-in-noise disparity discrimination:** random dot stereograms (RDSs) (8 dots/deg<sup>2</sup>) depicted a central square plane (7 x 7°) at either a crossed or uncrossed horizontal disparity ( $\pm 6$  arcmin) relative to the surrounding border dots (14 x 20°), which lay in the fixation plane. The proportion of dots that appeared at the correct depth in the target plane was parametrically varied between 0 and 100% (7 levels), while the remaining dots were assigned a random depth ( $\pm 12$  arcmin). Stimuli were presented for 500ms and participants reported whether the central target appeared near or far relative to the fixation plane.
- **Fine disparity discrimination:** RDSs depicted a central square plane inside a border, similar to that presented in the coarse disparity discrimination task. However, in this case, all dots belonging to the target plane had the same disparity, this was parametrically manipulated across trials ( $\pm 0.3, 0.5, 1, 4, 6, 10$  and  $12$  arcmin). Again, stimuli were presented for 500ms and participants' task was to report whether the central target appeared near or far relative to fixation.
- **Disparity-defined contours:** RDSs depicted convex or concave 3D shapes with symmetrical contours, taken from the 0% noise condition of a previous study by Chandrasekaran et al. (2007). Stimuli were scaled to subtend the same visual angle as in the original study (14.4 x 14.4°), and the peak disparity at the centre of each shape

was  $0.21^\circ$ . Participants viewed the stimuli through red-green anaglyph glasses and were asked to report the orientation of the axis of symmetry for each, which was always either horizontal or vertical. Stimuli remained on screen until the participant had given a response. There were 10 stimulus contour shapes, each of which was presented in both vertical and horizontal orientations as convex and concave surfaces, yielding a total of 40 unique stimuli that were presented once each.

### *Perceptual bistability tasks*

Fixation markers were presented on 50% of trials, except for the apparent motion dot quartet task, where the fixation marker was always present. On trials without fixation, observers were instructed to maintain their gaze on the stimuli.



**Figure 6.3.** Schematic illustrations of perceptually bistable stimuli. Three forms of bistable stimuli were presented: **A)** a rotating sphere defined by structure-from-motion; **B)** an apparent-motion dot quartet; **C)** binocular rivalry. For each stimulus, there are two possible perceptual interpretations (indicated here by red and blue), which alternate over time: **A)** leftward or rightward rotation; **B)** vertical or horizontal motion; **C)** left or right eye image.

- **Rotating sphere (Fig 6.3A):** RDKs depicted the orthographic projection of a rotating sphere ( $6^\circ$  diameter) defined exclusively by structure-from-motion. The sphere consisted of 400 black and white dots (3 arcmin diameter) distributed randomly across the transparent surface, and rotated about a vertical axis at an angular velocity of  $90^\circ\text{s}^{-1}$ . The apparent direction of rotation was bistable, except during catch periods when binocular disparity was added to disambiguate the direction of rotation. The magnitude

of the disparity added during catch periods was set for each observer based on the disparity that enabled a score of 84% correct on the dynamic stereo task (see above). For patients with thresholds outside the tested range of disparities, the maximum disparity was used for catch trials.

- **Apparent motion dot quartet (Fig 6.3B):** The apparent motion dot quartet (v. Schiller, 1933; Anstis and Ramachandran, 1987) was composed of 2 white dots ( $1^\circ$  diameter) located in diagonally opposite corners of a rectangular mid-grey background ( $5^\circ$  wide x  $7\text{--}7.2^\circ$  high), and 2 black dots of equal size in the other two corners. In the ambiguous condition, the dots switched between these two configurations every 300ms, with a 1 frame (10ms) blank interval interleaved, producing a perception of apparent dot motion that was bistable between horizontal and vertical motion. This frame rate has previously been shown to be well below the threshold for apparent motion perception in parietal damaged patients (Battelli et al., 2001). Observers fixated a marker located in the centre of the stimulus. In a square configuration of dots (i.e. when horizontal and vertical distances between dots are equal) observers exhibit a bias toward perceiving vertical motion. We therefore increased the vertical distance between dots until participants showed approximately equal probability of reporting horizontal and vertical motion percepts (Sterzer and Kleinschmidt, 2007).
- **Binocular rivalry (Fig 6.3C):** images subtended a visual angle of  $8^\circ$  and comprised either a face versus moving dots, or oblique orthogonal drifting gratings. Motion direction was randomized between trials. During catch periods, the contrast of one image was gradually reduced to 20% of its original contrast over a period of 2s, while the contrast of the other image remained constant. This reduced the probability of the unchanged image being suppressed and thus increased the probability of it becoming dominant, although it was also possible for observers to perceive the low contrast

image.

- **Stabilization:** additionally, the rotating sphere task was performed as described above, but with intermittent presentation of the stimulus in a 1s on, 1s off cycle. In healthy observers this presentation method is known to increase perception durations, i.e. reduce alternation rate (Orbach et al., 1963; Leopold et al., 2002).

### *Tasks*

Data collection was performed over a period of 18 months. Each session began with a 9-point calibration routine of the eye tracker. For bistability experiments, each block lasted 3-5 minutes, during which observers continuously reported their percept. We randomized the occurrence of ‘catch periods’, during which subtle manipulations were applied to the bistable stimuli that biased observers toward one percept over the other (Leopold et al., 2003; Krug et al., 2008). For the rotating sphere stimulus, this was achieved by adding small binocular disparities to the sphere, thus yielding an objectively correct direction of rotation. For binocular rivalry, the contrast of one image was reduced, and for apparent motion, intermediate frames were added, thus disambiguating the direction of motion. For observers who were capable of discriminating these changes, this ensured that they were attending to the task. Catch periods were triggered by 15% of key presses outside of a catch period, and began at a random interval (1- 5 seconds) following the key press. Catch periods lasted 6 seconds and the disambiguated percept was randomized.

During catch periods in the apparent motion task the direction of movement was disambiguated by briefly presenting an intermediate frame (10ms), in which the dots appeared halfway between their two normal positions – indicating either horizontal or vertical unambiguous motion (Fig.6.3). The contrast of these disambiguating dots was reduced (pixel intensity = 5%) to make their presence less obvious. This dot contrast was

selected as the lowest contrast at which all observers were able to reliably perceive disambiguated dot motion direction.

### *Imaging and analysis*

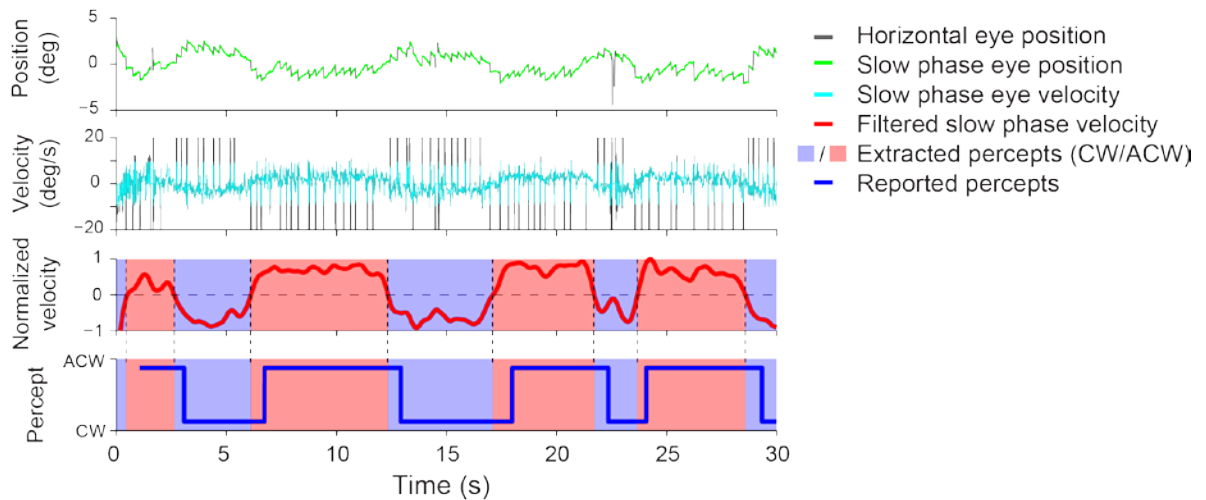
Anatomical images were collected at the Birmingham University Imaging Centre using a 3-T Philips Acheiva MRI scanner with an eight-channel phased array SENSE head coil. T1-weighted MR images (1mm isotropic voxels, TE = 3.8ms, TR = 8.4ms) were acquired and processed using the Statistical Parametric Mapping package SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/spm8.html>) for MATLAB (The MathWorks, Natick, MA). Lesion masks were created for each patient, using ITK-SNAP's active contour segmentation (Yushkevich et al., 2006), and adjusted manually. Patients' structural MR images and lesion masks were then spatially normalized to the MNI152 T1 template using unified segmentation (Ashburner and Friston, 2005;Crinion et al., 2007) and lesion cost function masking (Brett et al., 2001;Andersen et al., 2010). Regions of interest (ROIs) were created based on published MNI coordinates from previous studies of parietal involvement in perceptual bistability (Carmel et al., 2010;Kanai et al., 2010;Zaretskaya et al., 2010;Kanai et al., 2011). For each ROI, the proportion of voxels within a 20mm diameter sphere centred on the MNI coordinate that intersected with the spatially normalized binary lesion mask was calculated. Voxels within the spherical ROI that lay outside of the normalized brain mask were not included.

### *Behavioural Analysis*

For stereoscopic tests, binocular disparity thresholds, sensitivity, and confidence intervals were calculated by fitting a cumulative Gaussian psychometric function using a bootstrapping method (Psignifit toolbox; (Wichmann and Hill, 2001a;b), and lapse rate

was set to 0.01. Binomial maximum likelihood estimates were calculated using MATLAB's `binofit` function, which uses the Clopper-Pearson method to calculate confidence intervals (Clopper and Pearson, 1934). For the disparity-defined contour discrimination task, neither disparity magnitude nor signal intensity were manipulated and therefore the 'proportion correct' was used for analysis.

For perceptual bistability data, percept durations were calculated from observers' active report (via button press). For the rotating sphere and binocular rivalry tasks, percept durations were additionally calculated based on analysis of optokinetic nystagmus (OKN) eye movements (Fox et al., 1975; Leopold et al., 2003; Frässle et al., 2014). To achieve this, eye position data were first low-pass filtered (25Hz cutoff frequency), and saccades and fast phases of OKN eye movements were then removed by applying a velocity threshold of



**Figure 6.4.** Example data from a single trial of continuous presentation for the rotating-sphere task. The upper plot shows the participant's horizontal eye position (grey trace) over time, from which a clear saw-tooth wave characteristic of optokinetic nystagmus (OKN) can be seen. The slow-phase velocities (cyan) are extracted and filtered (red), and the velocity sign (+/-) is used to infer which surface direction was perceived as the front of the rotating sphere (background colour). The results agree closely with participants' subjective perceptual reports (blue), although alternations in OKN pattern can be seen to precede perceptual report.

$\pm 20^\circ\text{s}^{-1}$  in the horizontal direction. The remaining slow phase OKN velocity data were linearly interpolated and low-pass filtered (1Hz), and changes in velocity sign were then

identified as perceptual alternations (**Figure 6.4**). Extracted perceptual phases lasting less than 1 second in duration were discarded. Finally, the extracted perceptual time courses were compared to observers' subjective perceptual reports using cross-correlation. All further analyses were based on observers' subjective report data.

Perceptual dominance periods that were interrupted by catch events or the end of a trial were discarded. Percept durations of less than 300ms were also discarded as they are likely to be due to accidental simultaneous button presses. Within individuals, the distribution of percept durations is positively skewed, so the mean percept duration (or its reciprocal, alternation rate) is not an ideal metric for perceptual stability. We therefore fitted two functions to the data that have previously been shown to closely approximate the empirical distribution: gamma (Levelt, 1967; Borsellino et al., 1972; Logothetis et al., 1996), and lognormal (Zhou et al., 2004; Krug et al., 2008), which are fully described by just two parameters (shape and scale; see Supplementary Materials). Goodness of fit for each theoretical distribution to the empirical data was tested using Chi square. Note that both the distribution of percept durations for a given individual, and the distribution of mean percept durations across individuals are positively skewed (Krug et al., 2008; Kanai et al., 2010), and therefore both require non-parametric statistical tests such as the Mann-Whitney and Kolmogorov-Smirnov tests.



## 6.3. Results

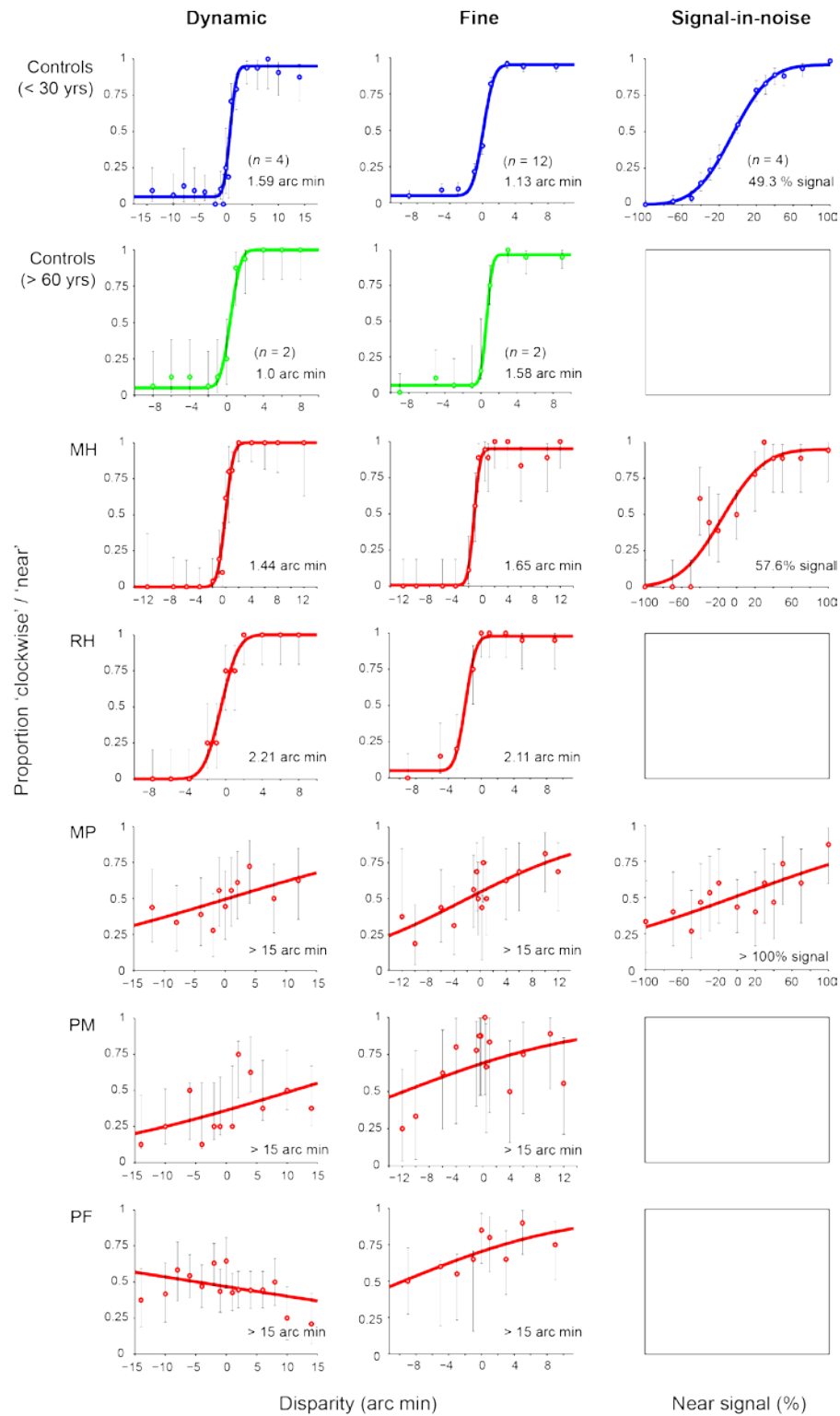
### 6.3.1 Stereopsis

#### *Dynamic Stereopsis*

In the dynamic stereo experiment, observers viewed a transparent virtual sphere covered in dots, which rotated either clockwise or anticlockwise about a vertical axis (**Figure 6.2**). The task was a two-alternative forced choice in which observers reported the direction of motion of the front surface of the sphere (left or right). In the absence of disambiguating depth information, either direction of rotation can be perceived. We parametrically manipulated the amount depth between the front and back surfaces of the sphere by adding binocular disparities, thus disambiguating the direction of rotation. We estimated thresholds (the disparity at which performance reached 84% correct) by fitting cumulative Gaussian psychometric functions to the data for each subject (**Figure 6.5**). Three of the patients performed poorly on this task, and did not achieve accuracy above 84% even at the largest disparities tested (15 arc min). Prior to the dynamic stereo test, these patients reported perceiving a rotating sphere when presented the stimulus, suggesting that their difficulty on the task was specific to processing dynamic disparity rather than object perception. Although the rotating sphere stimuli were presented for up to 5 seconds on each trial and participants were not instructed to give speeded responses, patients took slightly longer to respond on average (across all disparities) compared to other observers (MP =  $2340 \pm 78$ ms; PF =  $1552 \pm 47$ ms; PM =  $1762 \pm 53$ ms; RH =  $3594 \pm 113$ ms; MH =  $1197 \pm 35$ ms; Controls =  $960 \pm 30$ ms).

In contrast, the other two patients—MH and RH—were able to discriminate front and back surfaces using disparities of just 1.44 and 2.21 minutes of arc respectively (**Figure 6.5A**). These results were comparable to the performance of the younger control

group ( $n = 4$ , ages 23-28, mean threshold = 1.59 arc min) and age-matched control subjects DC and RC (0.96 and 1.04 arc min).



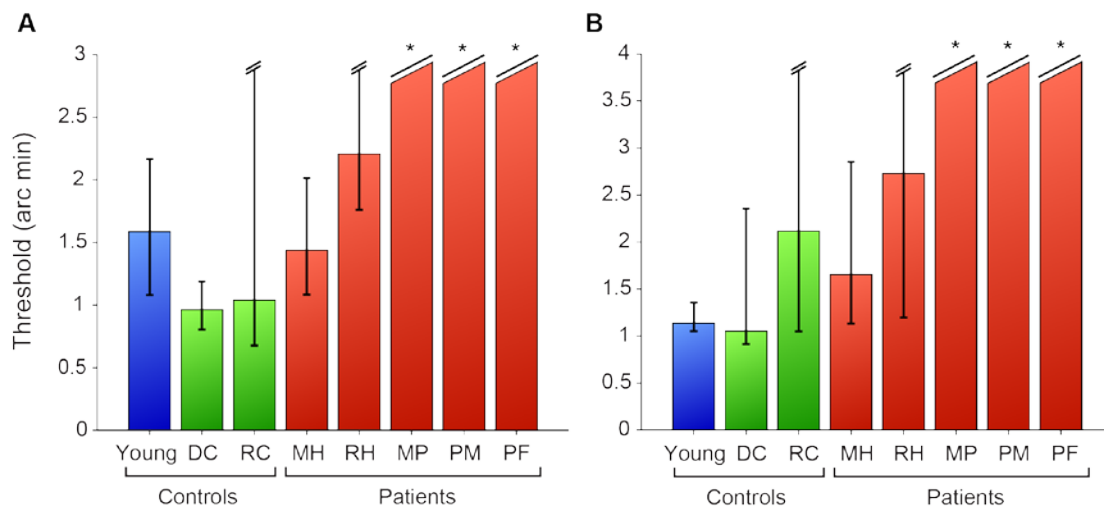
**Figure 6.5.** Psychometric functions for all observers on the dynamic, fine, and signal-in-noise binocular disparity tasks. For dynamic disparity, the proportion of ‘clockwise’ responses are plotted as a function of the relative disparity between the front surface of a clockwise rotating sphere and its axis of rotation. For fine disparity, the proportion of ‘near’ responses are plotted as a function of the disparity of the target plane relative to the border. For signal-in-noise, the proportion of ‘near’ responses are plotted as a function of % signal intensity. Error bars indicate 95% confidence intervals for binomial test. Inset values indicate thresholds at which observers responded correctly to 84% of trials.

### *Relative disparity*

In the fine stereo discrimination task, observers were presented random dot stereograms depicting a central square target plane that was either near or far relative to the surrounding surface (**Figure 6.2**). The pattern of discrimination thresholds across observers on this task was similar to that observed in the dynamic stereo experiment. The same three patients (MP, PM and PF) were unable to reliably discriminate between near and far conditions in the range of disparities tested ( $\pm 15$  arc min). However, the slopes of the psychometric functions were greater than zero, indicating that these observers were able to utilize some disparity information. In contrast, patients MH and RH performed similarly to controls, reliably discriminating surfaces with only 1.65 and 2.11 minutes of arc respectively.

### *Signal-in-noise Stereopsis*

Only MH, MP and young control observers were tested on the signal-in-noise disparity task. MH's performance, shown in Figure 6.5, was again comparable to that of young



**Figure 6.6.** Stereo thresholds for **A**) dynamic stimuli and **B**) fine disparity stimuli. Bars show the disparity at which performance reached 84% correct. Error bars show 95% confidence intervals. \*Thresholds for patients MP, PM and PF could not be estimated accurately since their performance was less than 84% correct at the largest disparities tested ( $\pm 15$  arc min).

healthy control subjects ( $n=4$ ), requiring a signal to noise ratio of 57.6% in order to correctly discriminate depth sign 84% of the time (compared to 49.3% SNR threshold for young controls). In contrast, MP performed poorly even at 100% signal levels, as expected from his performance on the fine task.

### *Disparity-defined contours*

Only patients MH, PM and PF were tested on the disparity contour task. In a previous study, healthy observers were able to correctly discriminate the axis of symmetry (horizontal or vertical) with approximately 98% accuracy when the stimuli contained 10% noise were presented for just 300ms (Chandrasekaran et al., 2007). The patients were presented versions of the same stimuli without noise added and given unlimited time to view and report their percept. PM scored 70% correct for convex (uncrossed disparities) and 85% for concave (crossed disparities) shapes, while PF scored 60% and 80% respectively, both of which were statistically above chance (binomial test,  $p < 0.001$  and  $p < 0.01$ ), but also significantly less than healthy controls ( $p < 0.001$ ). Patient MH scored 95% accuracy on this task (2 incorrect responses).

### *6.3.2. Perceptual Rivalry*

#### *Structure-from-motion*

In order for the ambiguous rotating sphere task to provide meaningful results regarding 3D perception, we first tested whether patients were able to perceive structure-from-motion. This was achieved by monocularly presenting six random dot kinematograms, half of which contained SFM cues that depicted symmetrical rotating 3D structures (a sphere, a cylinder, and a cone). The other half of the stimuli had the same 2D contours as the SFM

stimuli, but dots moved in opposite directions with uniform velocity and thus gave the appearance of two flat surfaces translating behind an aperture. Patients were asked to describe what they saw while viewing each stimulus and verbally reported perception of depth structures for the SFM stimuli (such as “a barrel” for the cylinder, or “a Christmas tree” for the cone).

#### *Catch event accuracy*

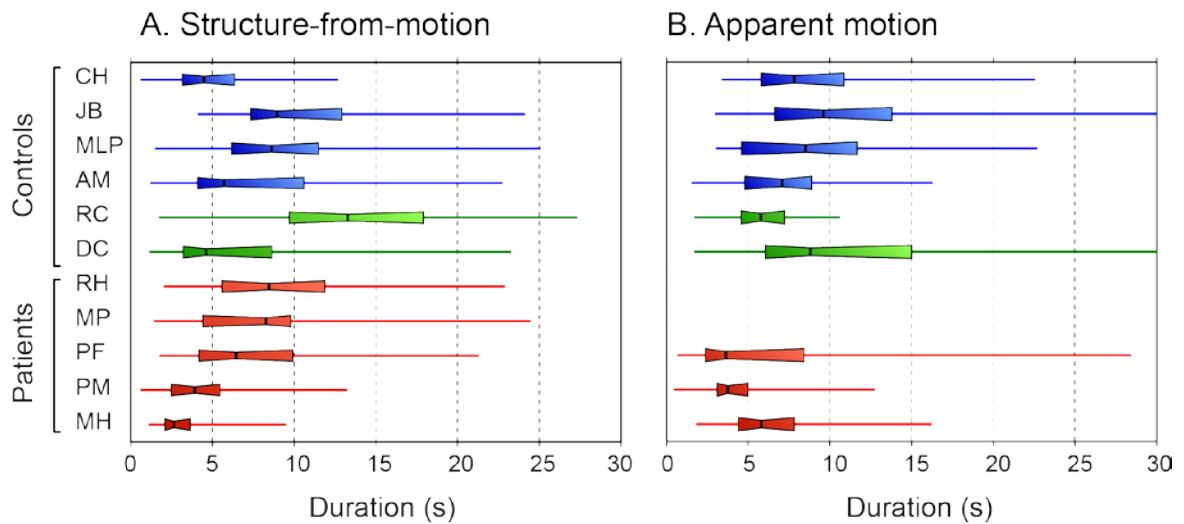
In order to control for attentional effects on perceptual alternation, we introduced randomized catch periods in the perceptual bistability experiments, during which the stimuli were temporarily rendered unambiguous. This resulted in only one objectively correct percept during these periods, which observers were expected to report. In the ambiguous rotating sphere experiment, the stimulus during catch periods was identical to that used in the dynamic disparity task, but with a fixed disparity magnitude for each observer. As expected from performance on the dynamic disparity experiment, patients MH and RH responded correctly to 98% (118/120) and 89% (25/28) of catch events respectively on the rotating sphere task. Similarly, control observers responded correctly to  $93 \pm 7\%$  catch periods on average. For the other patients who had shown large thresholds for the dynamic disparity task, the disparity during catch periods was set at 15 arc min. Surprisingly, two of these patients (MP and PF) were better able to correctly detect the unambiguous direction of rotation from disparity cues in the context of the continuously presented bistability experiment than they had during the individual trials of the dynamic disparity task (MP = 100%, 4/4; PF = 80%, 12/15; PM = 54% 13/24). Similarly, all observers responded correctly on the majority of catch periods for apparent motion and binocular rivalry stimuli (see Supplementary **Table S6.1**). These results suggest that patients were actively engaged in the perceptual tasks and attending to the stimuli.

### *Eye movements*

We recorded eye position binocularly during all bistability tasks, and presented a central fixation marker on only a small proportion of trials for the rotating sphere and binocular rivalry tasks. When observers were not instructed to fixate, these stimuli naturally elicited optokinetic nystagmus (OKN) eye movements, which provide a physiological indicator of perception (see Methods; (Fox et al., 1975; Leopold et al., 2003; Frässle et al., 2014). Comparison of the extracted versus reported perceptual time courses revealed that extracted perceptual transitions tended to precede reported transitions (**Figure 6.4**). Cross-correlations showed that the maximum correlation coefficients (patient group mean  $R = 0.78 \pm 0.04$ ) occurred at a mean lag time of  $1.10 \pm 0.21$ s (see Supplementary **Table S6.2**). Further, visual inspection of the eye movement data revealed no obvious abnormalities in any of the patients' eye movements in relation to the motion stimuli. Thus these data provide further support that observers were actively engaged in the perceptual task and not merely pressing the response keys randomly.

### *Percept durations*

The distributions of percept durations for all observers on the rotating sphere and dot quartet tasks are summarized in **Figure 6.7** (see Supplementary **Figure S6.2** and **Table S6.4** for details). There was considerable overlap in the range of percept durations for patient and control groups, although the shortest median durations were for two patients (MH and PM). Comparing the distribution of percept durations for each patient against the group distribution for all controls, only patients MH and PM showed significantly different percept durations on the rotating sphere task (Mann-Whitney test,  $p < 0.001$ ). For all three patients who were tested on the apparent motion dot quartet task (MH, PM and PF),

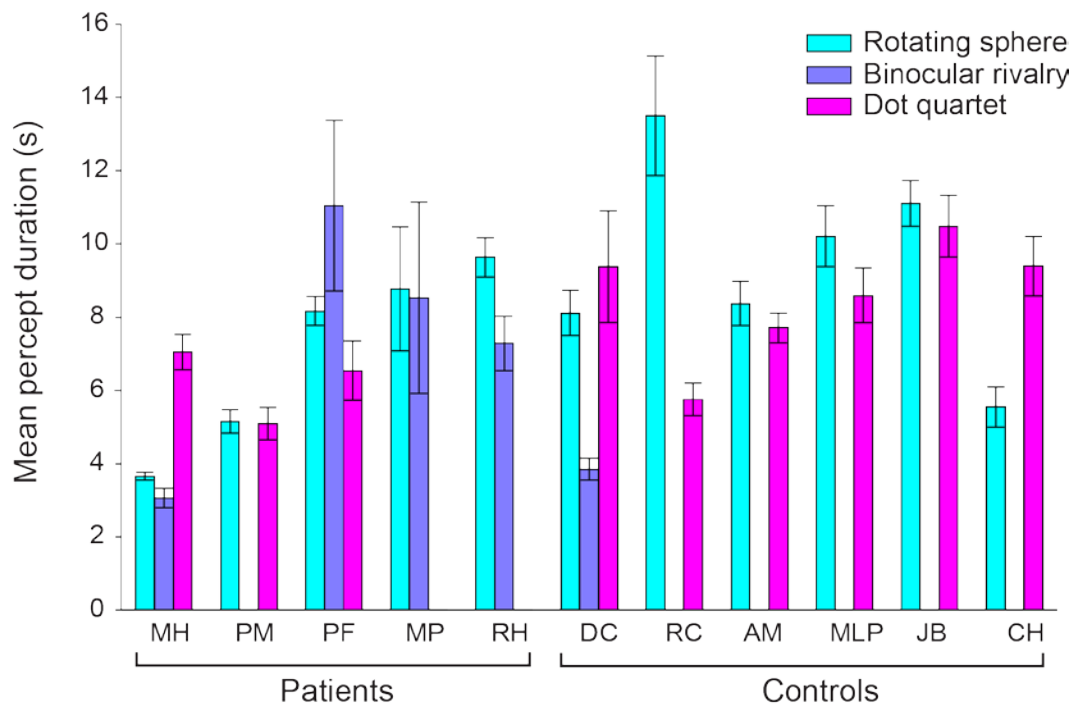


**Figure 6.7.** Percept duration distributions for all observers on A) the bistable rotating sphere and B) the apparent motion dot quartet. Notches represent medians, boxes represent 68% confidence intervals, and error bars represent 95% confidence intervals.

median percept durations tended to be shorter than those of control observers (Mann-Whitney test,  $p = 0.0476$ ).

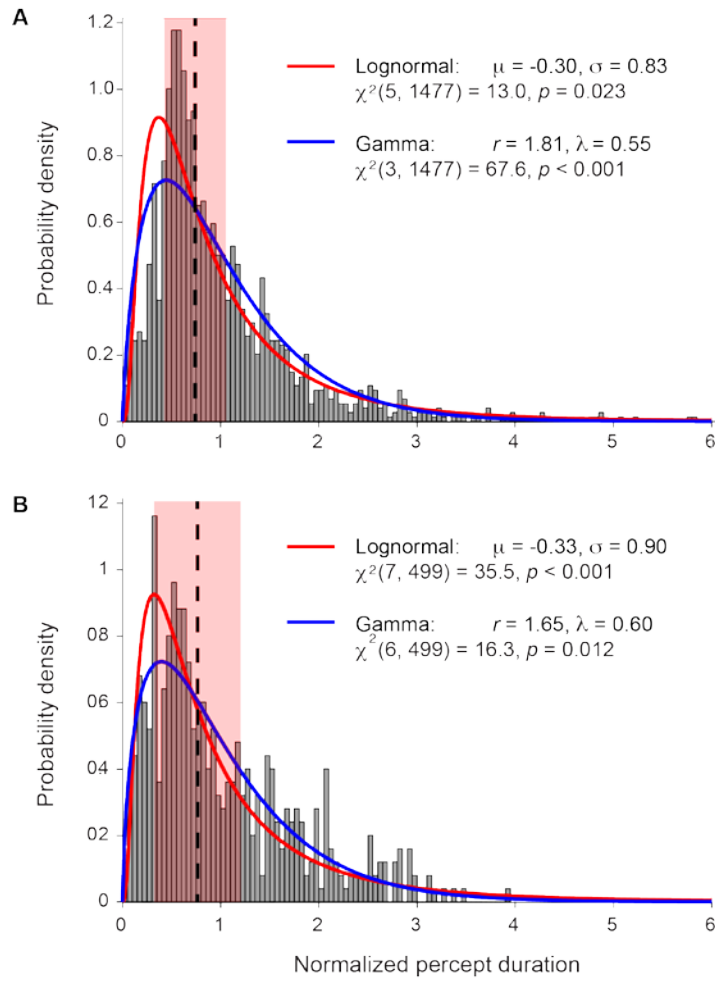
Mean percept durations for rotating sphere, binocular rivalry and apparent motion dot quartet stimuli followed a similar pattern (**Figure 6.8**), and mean percept durations for all three types of stimuli showed a trend towards being correlated with each other, although this did not reach statistical significance (rotating sphere-binocular rivalry,  $R = 0.61$ ; rotating sphere-dot quartet,  $R = 0.13$ ; **Figure S6.3A**). Mean durations for the rotating sphere task were comparable to those described elsewhere for similar stimuli (Brouwer and van Ee, 2006; Kanai et al., 2010). At the group level, the mean percept durations for patients and controls were not significantly different for any of the bistable tasks (Mann-Whitney test,  $p > 0.1$ ).





**Figure 6.8.** Mean percept durations for all three bistable stimuli tested. Error bars show standard error.

Mean percept durations provide a simple measure of perceptual stability for comparing observers and groups, but fail to capture more subtle differences in the percept duration distributions, especially since these distributions are positively skewed. To assess these differences we fitted probability density functions to each observer's distribution of percept durations (**Table S6.2; Figure S6.1**), and compared the parameters that described them. However, comparing the distribution of shape and scale parameters for the best fitting lognormal and gamma functions to each observer's percept duration distribution revealed no clear segregation between groups (**Figure S6.5**), and the parameters for each group were not significantly different from each other (Mann-Whitney test,  $p > 0.1$ ).

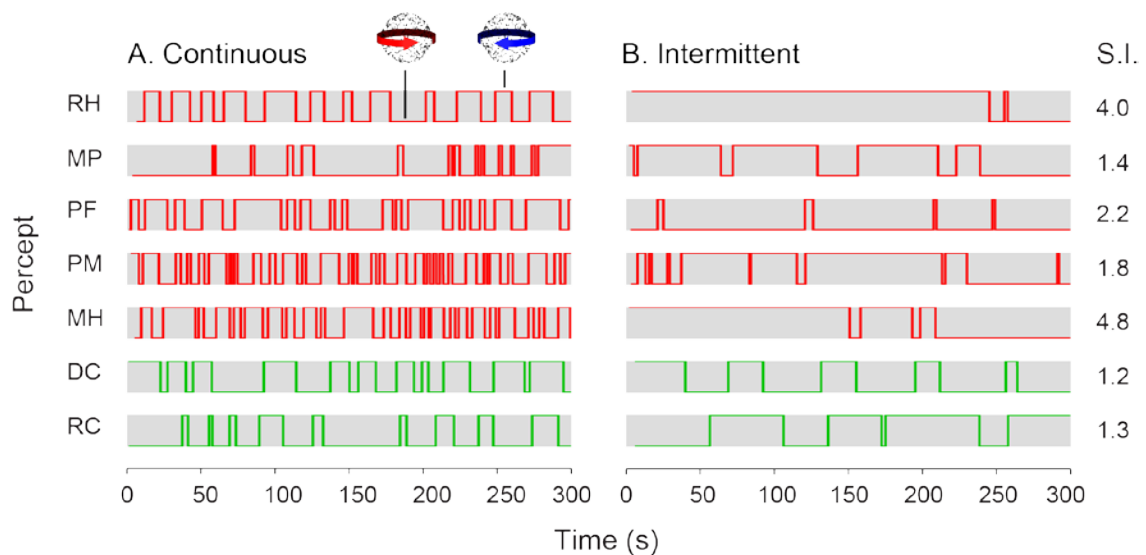


**Figure 6.9.** Distribution of normalized percept durations for **A)** patients and **B)** controls for the bistable rotating sphere task. Vertical dashed lines indicate the distribution medians and the light red bars indicate median absolute deviations.

Finally, we performed a group-level analysis by pooling the normalized percept durations for each group and comparing the resulting distributions (**Figure 6.9**). The gamma distribution provided the closest fit to the patient group's data ( $\chi^2(3, 1477) = 67.6, p < 0.001$ ), while the lognormal distribution provided the best fit to the control group's data ( $\chi^2(7, 499) = 35.5, p < 0.001$ ). This difference was due to a greater proportion of short percept durations (normalized value  $< 1$ ) in the control group compared to patients. Direct statistical comparison of the empirical distributions for the two groups revealed that they were significantly different (Kolmogorov-Smirnov statistic = 0.137,  $p < 0.001$ ).

### Perceptual Stabilization

In addition to measuring alternations in perception during continuous viewing of the bistable stimuli, we tested the effect of intermittent presentation of the same stimuli on observers' perception. Examples of perceptual alternations during 5-minute blocks of the rotating sphere task are presented in **Figure 6.10**, for patients (red) and older control observers (green). For all observers tested, intermittent presentation of the stimulus (right column) reduced the frequency of perceptual alternations compared to continuous viewing (left column). Although the quantity of data available for the intermittent viewing was limited due to the long percept durations, statistical testing of the available data indicated that the distribution of percept durations during continuous viewing were significantly shorter than those during intermittent viewing (Kolmogorov-Smirnov statistic = 0.204,  $p < 0.001$ ). This reduction in percept durations appeared most pronounced in observers with



**Figure 6.10.** Example data illustrating perceptual stabilization. A) During a 5-minute block of continuous presentation of the rotating sphere stimulus, observers reported spontaneous alternations between the two possible percepts (clockwise or anti-clockwise rotation), at a rate that varies between individuals. B) When the same stimulus is presented intermittently (in this case a 1 second on, 1 second off cycle) the frequency of perceptual alternations is reduced. S.I. = stabilization index (see Equation 6.1).

higher frequencies of alternation during continuous viewing. To quantify the stabilizing effect of intermittent presentation beyond that predicted by the net reduction in stimulus presentation duration, we calculated a stabilization index (S.I.):

$$S.I. = \frac{\text{Continuous alternations}}{\text{Intermittent alternations}} \times T_{ON} / (T_{ON} + T_{OFF})$$

(Equation 6.1)

An index value of 1 indicates that the reduction in alternations reported during intermittent viewing is predicted by the reduced duration of the stimulus appearance. Larger index values indicate a larger reduction in the proportion of alternations during intermittent compared to continuous viewing. Index values for each observer are shown in **Figure 6.10**.

### 6.3.3. Relationships between behaviors and lesion locations

Based on previously evidence we hypothesized that the specificity and location of PPC lesions might predict patients' perceptual dynamics. In order to quantify the extent to which different regions of PPC are affected by lesions in each patient, we used spherical regions of interest (ROIs) centered on previously published MNI coordinates for anterior and posterior superior parietal areas (**Table 6.1; Figure S6.2**). We calculated the proportion of voxels in each ROI that overlapped with each patient's lesion mask (**Table 6.3**), in order to provide a coarse metric of local tissue damage. Plotting these values against mean percept durations did not reveal any clear correlation based on the small patient sample tested (**Figure S6.3C**). Further, we found no correlation between observers' age and mean percept durations. Comparing the pattern of behavioural results with the pattern of damage to the parietal ROIs (summarized in **Table 6.3**), one notable finding is that patient RH's lesions did not overlap with any of the ROIs tested and he was the only patient who showed both normal stereo thresholds and perceptual alternation rates.

One hypothesis derived from previous studies is that anterior and posterior SPL exert opposing influences on mean percept duration, with disruption of aSPL reducing it and disruption of pSPL increasing it. For the ROIs based on these studies (Table 6.3, group A), patients MP and PM showed slightly more damage to anterior ROIs than posterior ones. In agreement with the hypothesis, PM showed shorter mean percept durations on both the rotating sphere and apparent motion dot quartet stimuli. Patient MP's mean percept duration was in the normal range, although there were limited data available for this observer (number of percept durations on the rotating sphere = 17), resulting in a larger standard error in the estimate of the mean. In contrast, patients MH and PF's lesions overlapped both anterior and posterior ROIs, for which there is no clear predicted outcome. Indeed, MH showed shorter percept durations than normal, while PF's were variable across stimuli (normal for the rotating sphere but shorter than normal for apparent motion).

**Table 6.3.** Percentage of voxels within each parietal ROI classified as lesioned.

Patient	Stereo	Switch rate	Lesion side	Group A				Group B			
				Left pSPL	Right pSPL	Left aSPL	Right aSPL	Left pSPL	Right pSPL	Left aSPL	Right aSPL
RH	+	+	L	0	0	0	0	0	0	0	0
MH	+	-	L/B	70	63	55	61	59	56	0	66
PM	-	-	B	13	16	40	41	2	5	0	29
PF	-	+	R	86	93	89	88	12	16	7	50
MP	-	+	R	0	0	24	24	0	0	0	30

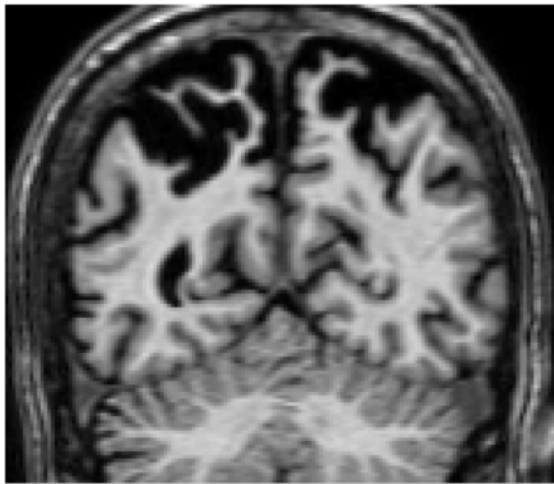
\*A = ROIs based on coordinates from (Carmel et al., 2010; Kanai et al., 2010) and Ryota Kanai, personal communication, August 22<sup>nd</sup> 2014. B = Corresponding but non-overlapping ROIs based on coordinates from (Zaretskaya et al., 2010).

## 6.4. Discussion

We tested the visual perception of depth from binocular disparity and motion in five patients with lesions to the parietal lobe, as well as healthy controls, using psychophysical methods. The behavioural performance of the patient group was heterogeneous: on the stereo tasks, some patients were able to perceive depth with sensitivities comparable to control subjects, while others showed obvious deficits, and these were associated with extensive damage to right PPC. During the presentation of ambiguous structure-from-motion and other bistable stimuli, some patients showed rates of perceptual alternation within the same range as control observers, while others exhibited noticeably faster perceptual alternations. We quantified lesion volumes for specific regions of interest within the PPC that have previously been implicated as playing causal roles in perceptual alternation. Although we found no clear relationship between lesions within these ROIs and the alternation rates for our patients, several findings were in agreement with the hypothesis that anterior SPL may play an inhibitory role in perceptual alternation under ambiguous viewing conditions.

### *6.4.1. Stereopsis*

Two patients (MH and RH) showed sensitivity to binocular disparities in the normal range for behavioural tasks that required the discrimination of relative disparities, comparable to that of healthy control subjects. Given that the damage primarily affects the left hemisphere in these patients, this result is in agreement with previous neuropsychological studies, which suggest that stereopsis is mediated predominantly by the right hemisphere (Carmon and Bechtoldt, 1969; Benton and Hécaen, 1970; Durnford and Kimura,



**Figure 6.11.** A coronal slice through patient MH's posterior parietal lobe ( $y = -60\text{mm}$ ) showing bilateral sulcal enlargement affecting IPS. Despite the damage to these areas, MH retains the ability to discriminate depth from binocular disparity better than many neurologically healthy observers.

1971;Hamsher, 1978;Ross, 1983). More specifically, it has been suggested that the right parietal lobe might be essential for stereopsis (Vaina, 1989). In contrast, human imaging studies have tended not to emphasize lateralization in IPS activation to binocular disparities, although examples of greater right-hemisphere activation are evident (Tsao et al., 2003;Durand et al., 2009;Georgieva et al., 2009). It is therefore of particular interest that patient MH showed both intact local (fine task) and global (signal-in-noise task) stereopsis, despite some bilateral lesions to parietal cortex (**Figure 6.1, 6.11**).

MH's injuries are a result of anoxia, which caused thinning of the cortical sheet most prominently in the left parietal lobe but also in the right IPS (rIPS). However, his excellent performance on all stereo tasks tested in the present study indicates that the areas affected by his injury are not essential for stereopsis. This conclusion conflicts with the hypothesis, supported by some human neuropsychological and imaging evidence, that suggests an important role for the right hemisphere IPS in stereopsis. One possibility for reconciling the current finding with previous work is that rIPS in healthy observers contains redundant disparity processing capacity, and MH may therefore have retained sufficient rIPS grey matter to support stereopsis. Other functions associated with right parietal cortex also remain intact in MH, such as reaching and grasping movements directed to the left hemifield. Indeed, he exhibits an unusual lateralized pattern of optic

ataxia that only affects reaching with his right hand to objects in the right hemifield (Kitadono and Humphreys, 2007; Rice et al., 2008; Cavina-Pratesi et al., 2013). It should be noted that all stereoscopic stimuli were presented centrally in our experiments, and therefore unimpaired stereopsis in just one hemifield would have been sufficient for patients to perform the task.

In contrast, all three patients with lesions to right parietal cortex in the present study (MP, PM, PF) showed disparity discrimination thresholds outside the tested range ( $> 14$  arc min) on both the fine and dynamic disparity tasks. Nevertheless, the psychometric fits for these observers' data mainly had positive gradients, indicating some sensitivity to disparity information (**Fig 6.6**). Similarly, although PM and PF performed worse than healthy controls on the disparity-defined contour task, their performances were significantly above chance. The only case where a psychometric fit did not show the expected positive slope was for PF in the dynamic stereo task. Closer inspection of her responses on this task revealed that she tended to report the same percept across consecutive trials (**Fig S6.4**). This response pattern is consistent with the idea that without the ability to perceive depth from the disparity information, PF perceived only the ambiguous structure-from-motion cue. The temporal structure of the task (short trials and inter-trial intervals) creates an intermittent presentation schedule, which is known to induce strong perceptual stabilization in healthy observers (Leopold et al., 2002), and appears to be preserved in the current patient group (**Fig 6.10**).

Surprisingly, patients MP and PF performed better at discriminating the direction of sphere rotation from binocular disparities during catch periods on the bistability task than for individual trials (of the same maximum duration) on the dynamic disparity task. One possible explanation for this difference in performance on virtually identical tasks might be that the transient of the catch period onset increased the salience of the disparity



signal. However, another possibility is that the strong perceptual biases elicited by intermittent presentation of ambiguous versions of the stimulus interfered with perception of dynamic disparity information in these observers because of their poor stereoacuity. Although the poor sensitivity to disparity cues observed in these patients cannot be attributed specifically to their lesions (Richards, 1970), the performance of the age-matched control subjects within the range of younger controls suggests that these deficits are not necessarily related simply to age (Wright and Wormald, 1992; Brown et al., 1993; Zaroff et al., 2003), in agreement with previous evidence (Krug et al., 2008; Read et al., 2010). Further, lesions to areas outside of PPC could be responsible for the stereo deficits in these patients, such as the left pulvinar (Takayama et al., 1994), which may be affected in patient PF.

One previous study reported that none of the 5 right occipito-parietal lesion patients tested were able to detect global stereopsis (Vaina, 1989). In agreement with that finding, three of the patients in the present study with right PPC damage showed deficits in all forms of stereopsis tested. However, patient MH, who also has some right occipito-parietal damage, showed normal perception on the binocular disparity tasks, including global stereopsis measured by the signal-in-noise task. A previous study of another hypoxia patient, DF (who suffers visual form agnosia resulting from bilateral lesions of lateral occipital cortex), reported that depth perception from absolute and dynamic relative disparities was preserved, but not for static relative disparities (Read et al., 2010). However, that patient also suffered a lesion to left PPC and bilateral atrophy of IPS in addition to other areas (James et al., 2003; Bridge et al., 2013). This evidence suggests that the type of diffuse damage to PPC and IPS that can result from hypoxic/ anoxic injury is therefore not sufficient to induce deficits in all aspects of stereopsis.

#### *6.4.2. Motion perception and structure-from-motion*

Several previous patient studies suggest that both dorsal and ventral visual pathways may be involved in the perception of structure-from-motion (SFM). A previous study of akinetopsic patient LM revealed that her ability to perceive coherent motion, 2D shape from motion and 3D structure-from-motion all broke down with the introduction of moderate levels of noise (Rizzo et al., 1995). Similarly, patients with lesions to ventral visual areas of the right hemisphere were less accurate at discriminating SFM-defined object shapes only when the number of dots defining the object surface was reduced (Gilaie-Dotan et al., 2013). Thus ventral visual areas such as lateral occipital cortex seem to be important not merely for recognition of 3D object shape from motion and stereo cues (Gilaie - Dotan et al., 2002), but for extracting such information from noise.

Parieto-temporal lesions have also been shown to impair perception of motion-defined 2D form (Regan et al., 1992), and none of the 5 right occipito-parietal lesion patients tested in another study were able to perceive SFM based on a single trial (Vaina, 1989). In contrast to those findings, all five patients in the present study (including those with right occipito-parietal lesions) reported that they were able to perceive the 3D structure of the ambiguous rotating sphere stimulus. However, unlike some previous studies, our SFM stimuli were not embedded in noise, and thus the 2D object contour was visible even in a single static frame. Similarly, all of our patients reported being able to perceive apparent motion in a centrally presented dot quartet stimulus. This is in agreement with a previous study demonstrating that patients with lesions to right parietal cortex were able to perceive apparent motion in a dot quartet stimulus, provided that the frame rate was below 4Hz, as was the case for our stimulus (Battelli et al., 2001).

Patient MH's motion-selective area V5/hMT+ has previously been shown to be functionally intact (Dent et al., 2010). This is likely a minimum requirement for perception of depth from motion, given the known involvement of this area in processing SFM cues (Bradley et al., 1998;Dodd et al., 2001) and combinations of motion and disparity (Krug et al., 2013). However, evidence from neurophysiology suggests that parietal areas (such as LIP in the monkey), which receive direct input from MT, are responsible for generating perceptual decisions (Huk and Shadlen, 2005;Hanks et al., 2006) and human neuroimaging also suggests IPS involvement in processing SFM (Orban et al., 1999;Vanduffel et al., 2002). It is therefore somewhat surprising that MH readily perceived structure-from-motion and performed comparably to healthy control observers on the dynamic stereo task. Indeed, MH and RH's performance on the dynamic stereo task was comparable to that previously reported for a patient with lateral occipital lesions and three other age matched controls (aged 53-64) who were tested on a similar task (Read et al., 2010). The reduction in MH's right anterior parietal cortical grey matter therefore appears not to have been sufficient to affect perception of structure-from-motion, despite its known involvement in processing other forms of high-level motion (Battelli et al., 2001).

#### *6.4.3 Perceptual alternations during bistable viewing*

A major concern for any experiment in which participants report their subjective perceptual state is how closely their responses actually reflect their perceptual experience (Leopold et al., 2003). We used three methods to ensure that observers were actively attending to the bistable stimuli and accurately reporting their percepts. First, we included random catch periods during which the stimuli were rendered unambiguous, and for which there was only one objectively correct percept (Leopold et al., 2003;Krug et al., 2008). Second, we employed ambiguous motion stimuli that were capable of eliciting specific

patterns of reflexive optokinetic nystagmus eye movements that are known to correlate strongly with perceptual state (Fox et al., 1975; Leopold et al., 2003; Frässle et al., 2014). Third, we analysed the distributions of reported percept durations and assessed the goodness of fit with theoretical functions known to provide accurate models for these distributions (Levelt, 1967; Borsellino et al., 1972; Zhou et al., 2004). The results of all three tests indicated that all observers were indeed attending to the tasks and reported their perceptual experience as accurately as possible.

In the bistable perception experiments, several patients (MH and PM) showed, on average, significantly shorter percept durations than other observers. While it is difficult to estimate the ‘normal’ range of percept durations given the high variability of this measure within the healthy population (Krug et al., 2008; Kanai et al., 2010), it was notable that these patients showed mean percept durations that were over 3 seconds shorter than any of the other observers tested on the rotating sphere stimuli, and that the same two observers also showed the shortest mean percept durations for binocular rivalry and the apparent motion dot quartet respectively. However, the relationship between the fast perceptual alternation rates and the pattern of parietal damage in these patients is not obvious. Based on the ROI lesion analysis, MH and PM have different patterns of parietal lesions: PM’s SPL lesions affect primarily anterior ROIs whereas MH’s lesions affect both posterior and anterior SPL (**Table 6.3**). Further, for each of the patients who showed fast alternation rates there is another patient with a similar pattern of parietal lesions whose rate of perceptual alternations was not different from controls (MP for PM, and PF for MH).

In contrast to the other patients, RH’s parietal lesions are confined to the left inferior parietal lobe, and showed no overlap with any of the parietal ROIs tested. Despite exhibiting severe extinction in his right hemifield, RH showed normal stereopsis and perceptual alternation rates that were not significantly different from control observers. As

the only patient tested with no damage to the right PPC, these results fit with the suggestion that depth perception from disparity and motion cues are lateralized to the right hemisphere.

Previous studies have hypothesized that an antagonistic relationship between posterior and anterior SPL regulates perceptual alternation rate. Specifically, this was based on the finding that posterior SPL grey matter density correlates negatively with percept duration ( $R = -0.59$ ), while the equivalent correlation for anterior SPL grey matter was positive ( $R = 0.53$ ) (Kanai et al., 2010; Kleinschmidt et al., 2012). Combined with the results of TMS studies that suggest causal relationships between these cortical areas and percept durations (Carmel et al., 2010; Kanai et al., 2010; Kanai et al., 2011), this hypothesis generates testable predictions for the effect of parietal lesions on percept duration. One patient with slightly more damage to anterior ROIs than posterior ones (PM) did indeed exhibit shorter mean percept durations on both the rotating sphere and apparent motion dot quartet stimuli, as predicted by the antagonistic hypothesis. However, other patients either had damage to a mixture of anterior and posterior ROIs that yielded no clear prediction, or else had insufficient data to reliably determine the normality of their percept durations (MP).

Two previous studies have examined perceptual alternations during binocular rivalry in patients with right hemisphere lesions (Bonneh et al., 2004; Daini et al., 2010). In stark contrast to our results (both for binocular rivalry and structure-from-motion), both studies reported that patients with right hemisphere lesions showed a *reduced* frequency of perceptual alternation compared to controls. One of the studies found that this was only the case for patients with unilateral spatial neglect (Bonneh et al., 2004), while the other found this effect for all patients with right hemisphere damage, irrespective of neglect symptoms (Diani et al 2010). Since neglect can also result from lesions to areas other than parietal

cortex, neither study specifically implicates right PPC. However, we found no clear relationship between lesion lateralization and percept durations in our sample, and some evidence to the contrary: one patient with limited right PPC damage (MH) exhibited fast alternation rates, while another (PF) showed alternations rates comparable to controls. Despite the apparent right hemisphere dominance for perceptual alternations emphasized by previous studies (Carmel et al., 2010; Kanai et al., 2011), there appears to be broad individual variation in lateralization (Zaretskaya et al., 2010). Further, a previous study of a split brain patient showed similar distributions of perceptual dominance durations in each visual hemifield for binocular rivalry (O'Shea and Corballis, 2003). However, there is evidence for hemispheric differences in other parietal mediated functions, such as a dissociation in the contributions of left and right PPC to processing salient stimuli (Mevorach et al., 2006).

For a given observer, perceptual alternation rates for different types of bistable stimuli tend to be correlated (Pressnitzer and Hupé, 2006). This trend was apparent in the results for the three bistable stimuli tested here, although the correlations were not significant, possibly due to a lack of statistical power. One previous study reported that increased age was associated with reduction in perceptual alternation rates for binocular rivalry (Ukai et al., 2003), although another study reported that this relationship was not apparent for SFM (Krug et al., 2008). Although the age distribution of our observers was bimodal, we found no correlations between age and perceptual alternation rate for any of the stimuli tested.

In addition to recording the time course of perceptual alternations during extended continuous presentation of ambiguous stimuli, we tested whether intermittent presentation of the same stimuli affected the dynamics of perception. In healthy observers, brief, repeated presentations of ambiguous stimuli separated by blank intervals on the order of

seconds are known to reduce the frequency of spontaneous perceptual alternations (Orbach et al., 1963; Leopold et al., 2002; Pastukhov and Braun, 2008; Brascamp et al., 2009). This phenomenon, often referred to as perceptual stabilization, is hypothesized to result from a form of short-term, implicit perceptual memory that promotes the repeated selection of the same percept across consecutive presentations. However, little is known about the neural mechanisms responsible for implementing this process. Our demonstration of perceptual stabilization through intermittent presentation in patients with parietal lesions suggests that the putative perceptual memory trace that facilitates this process is not mediated exclusively by parietal cortex. This finding is compatible with recent neuroimaging evidence suggesting that implicit memory influences perception of bistable stimuli via modulation of early visual areas (Sandberg et al., 2011; de Jong et al., 2012b; de Jong et al., 2014), including motion sensitive areas (Klink et al., 2012).

#### *6.4.4. Limitations of between-subjects studies of percept duration*

Mean percept durations for a given bistable stimulus vary widely between individuals even within the healthy population. While this variability has been utilized successfully for guiding experiments with a repeated measures design (Kanai et al., 2010), and despite statistically significant differences reported using between-subject designs, it creates difficulty for comparisons between clinical and healthy populations (Krug et al., 2008). Further, perceptual alternation rate is strongly affected by low-level stimulus properties, thus preventing direct comparisons of ‘normal’ alternation rates between studies with differing stimulus parameters (Brouwer and van Ee, 2006). For example, 27% of the healthy participants (14/52) tested by Kanai and colleagues (2010) showed mean percept durations of less than 4 seconds for a similar rotating sphere stimulus, but this trend toward shorter percept durations compared to our data can be accounted for by the increased speed

of stimulus rotation ( $151^{\circ}\text{s}^{-1}$  vs  $90^{\circ}\text{s}^{-1}$ ). Studies involving larger samples, such as those by Pettigrew and Miller (1998) ( $n = 63$ ), Krug et al. (2008) ( $n = 50$ ), and Kanai and colleagues (2010) ( $n = 52$ ), suggest a greater variability in perceptual alternation rates than those reported by previous studies using smaller samples (e.g. Bonnef et al., 2004), although this is could be due to differences in stimulus properties.

Despite the broad variation in individuals' mean perceptual dominance durations, many studies have used this measure to compare differences between groups. For example, several studies have found evidence that bipolar patients tend to experience longer mean percept durations compared to control groups (Miller et al., 2003; Krug et al., 2008; Nagamine et al., 2009; Vierck et al., 2013), although the result may not be sufficiently large or robust to serve as a reliable trait marker (Krug et al., 2008). Individual differences in mean percept duration in healthy observers have been linked to various measures including age (Ukai et al., 2003), mood (Sheppard and Pettigrew, 2006), delusional ideation (Schmack et al., 2013), and phenotype (Miller et al., 2010), while pharmacological effects on mean percept duration include GABA<sub>A</sub> agonist lorazepam (van Loon et al., 2013), 5-HT<sub>1A</sub> agonist tandospirone (Nagamine et al., 2008), and the psychedelics psilocybin (Carter et al., 2005; Carter et al., 2007), LSD (Carter and Pettigrew, 2003), and ayahuasca (Freeska, White & Luna, 2004). This broad range of influences/ correlations suggests that perceptual alternations are governed by a complex interaction of neural systems, for which the front-parietal network remains a primary candidate.

Several recent fMRI studies suggest that fronto-parietal activations previously associated with perceptual alternations actually reflect the gradual transitions between perceptual states (Knapen et al., 2011), or that prefrontal cortex activations reflect active reporting of perceptual state (Frässle et al., 2014). However, neurophysiological evidence suggests that frontal areas show a rapid response that correlates with perceptual transitions



and precedes visual responses (Libedinsky and Livingstone, 2011), or correlates with perceptual state independent of response (Panagiotaropoulos et al., 2012). While patients with PFC lesions show similar rates of spontaneous perceptual alternation to healthy controls, they are less able to voluntarily facilitate perceptual alternations (Windmann et al., 2006), and this finding has been replicated in healthy observers using TMS (De Graaf and Sack, 2011a). Convergent evidence therefore still points towards parietal cortex as being responsible for spontaneous perceptual alternations.

#### *6.4.4. Conclusion*

The results of our study suggest that the ability to perceive depth from binocular disparity in central vision is dependent on the PPC of the right hemisphere. However, the retention of stereopsis in one patient with limited damage to right PPC suggests that there is some level of redundancy in the processing of binocular disparity information in this area. In contrast to the perception of depth from disparity, all patients tested reported the perception of depth structure from motion. During continuous viewing of a perceptually bistable structure-from-motion stimulus, patients with parietal lesions showed variable rates of spontaneous perceptual alternations: either within in the normal range, or faster. This stands in contrast to previous reports for parietal lesion patients viewing binocular rivalry stimuli, which reported reduced alternation rates. Finally, we observed that intermittent presentation of the same ambiguous stimuli induced perceptual stabilization in parietal patients, just as it does in healthy controls. These results suggest a lateralized role of PPC in stereopsis, and suggest that the perceptual memory trace responsible for perceptual stabilization of bistable stimuli is not specific to parietal cortex.

## 6.5. Supplementary Materials

### *Distribution fitting*

Maximum likelihood estimates of parameters for the best lognormal, gamma, and gamma rate fits to the empirical percept duration distributions were calculated using the MATLAB functions `lognfit` and `gamfit`. The lognormal probability density function (PDF) is given by the equation:

$$f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$$

(Equation S6.1)

and the gamma PDF:

$$f(x|k, \lambda) = \frac{1}{\lambda^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\lambda}}$$

(Equation S6.2)

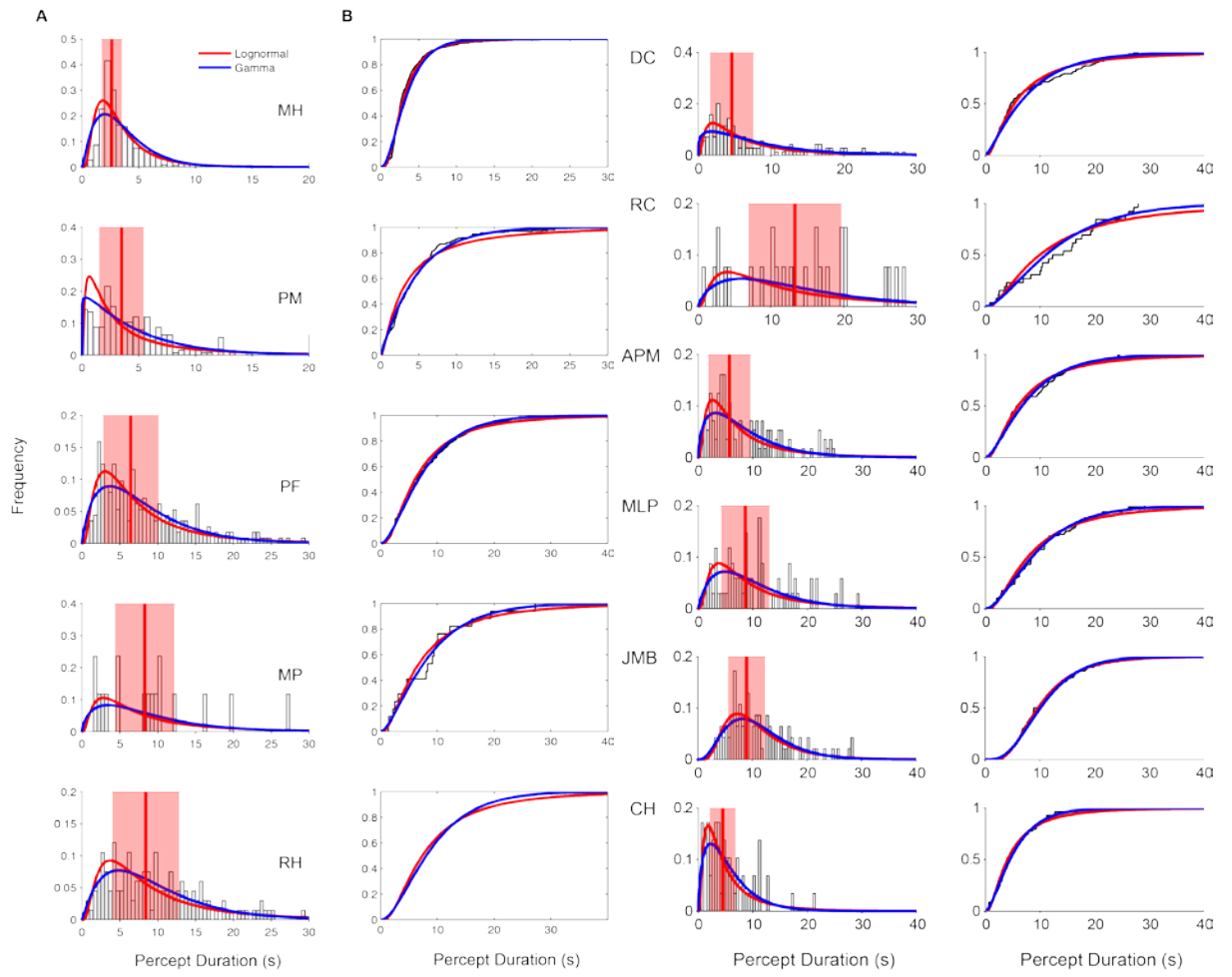
where  $x$  = time and  $\Gamma(k) = (k-1)!$ . For gamma rate, shape ( $k$ ) and scale ( $\lambda$ ) parameters for a gamma distribution were fitted to the distribution of switch rates, i.e. the reciprocal of the duration values ( $1/t$ ) were substituted for  $x$  (Eq.6.2). The result was multiplied by  $1/t$  to obtain the probability, and then divided by the width of a rate bin (Hz) to obtain probability density on a duration axis (seconds).

**Table S6.1.** Percentage correct responses during catch periods for bistable experiments.

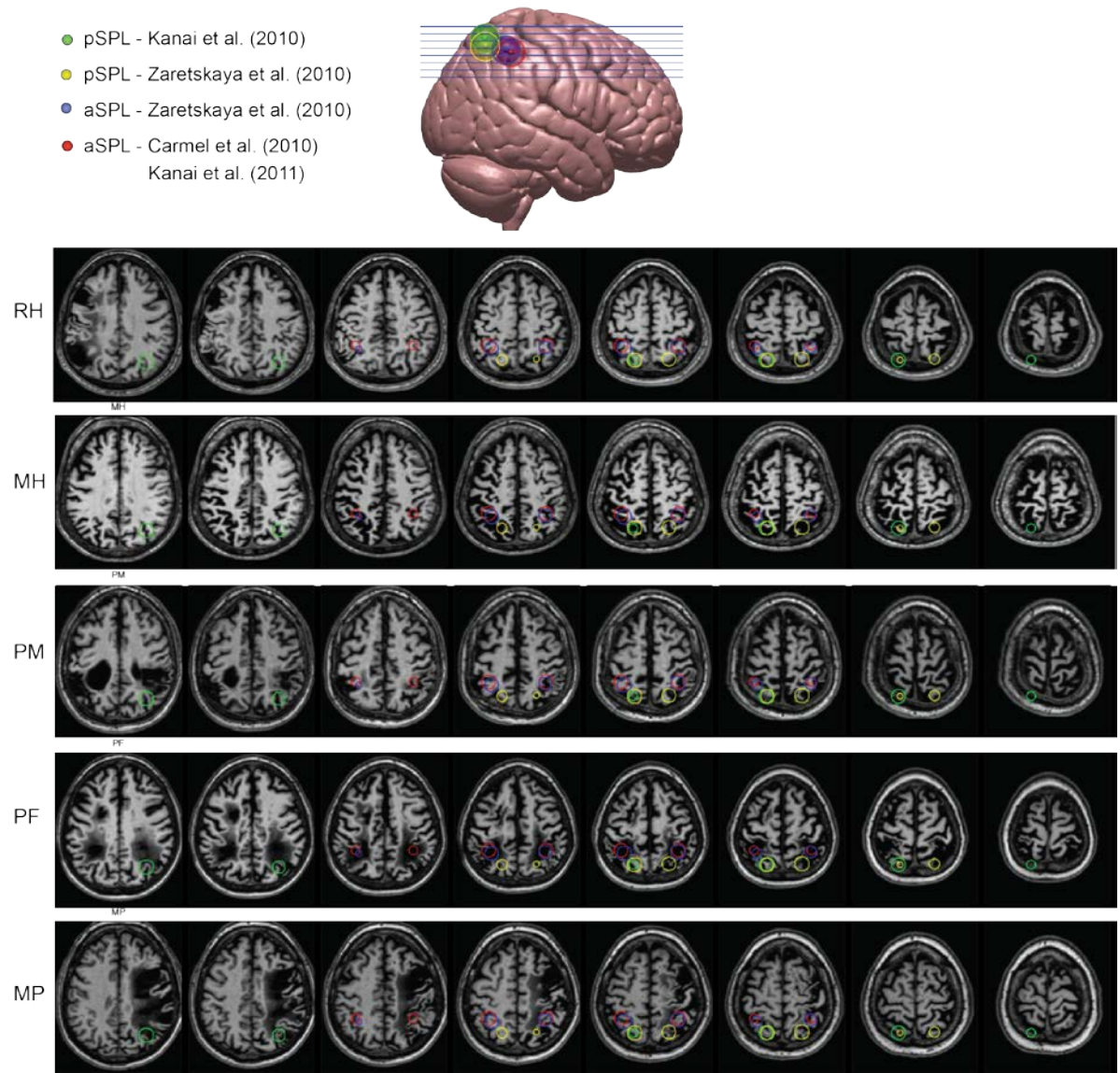
ID	Catch event performance		
	SFM	AM	BR
PF	80% (12/15)	100% (4/4)	NA
MH	98% (118/120)	70% (7/10)	60% (6/10)
PM	54% (13/24)	89% (16/18)	NA
RH	89% (25/28)	NA	75% (3/4)
MP	100% (4/4)	NA	100% (4/4)
DC	86% (12/14)	100% (2/2)	100% (1/1)
RC	100% (4/4)	67% (2/3)	NA

**Table S6.2.** Cross-correlation of perceptual time courses extracted from eye movement data with those reported by observers.

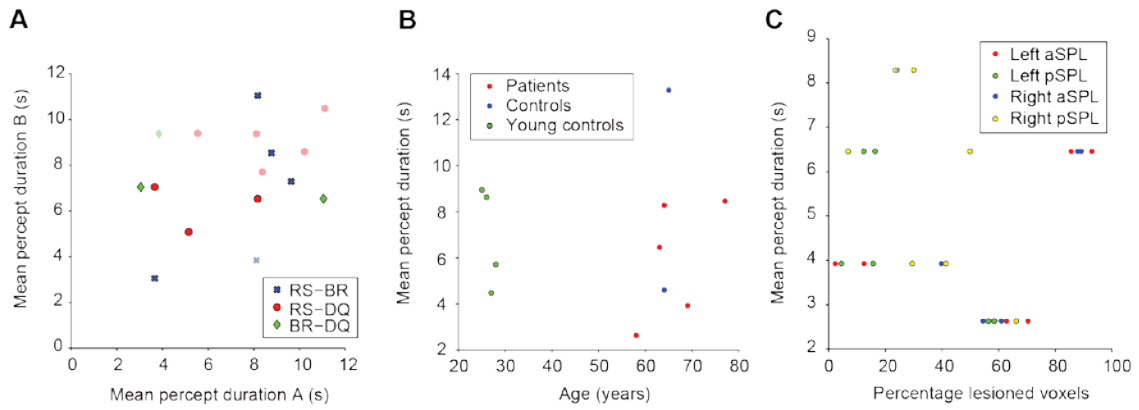
ID	R	Lag(s)
PF	0.771	1.24
MH	0.803	1.178
PM	0.649	1.67
RH	0.839	1.04
MP	0.843	0.37



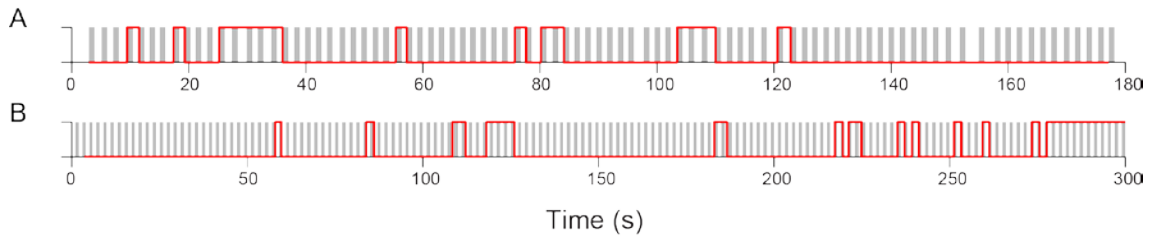
**Figure S.6.1.** Percept duration distributions for individual patients (left 2 columns) and controls (right two columns) on the SFM experiment. **A)** Probability density fitted with lognormal (red) and gamma (blue) distributions. Vertical red lines indicate the distribution medians and the light red bars indicate median absolute deviations. **B)** Cumulative probability distributions.



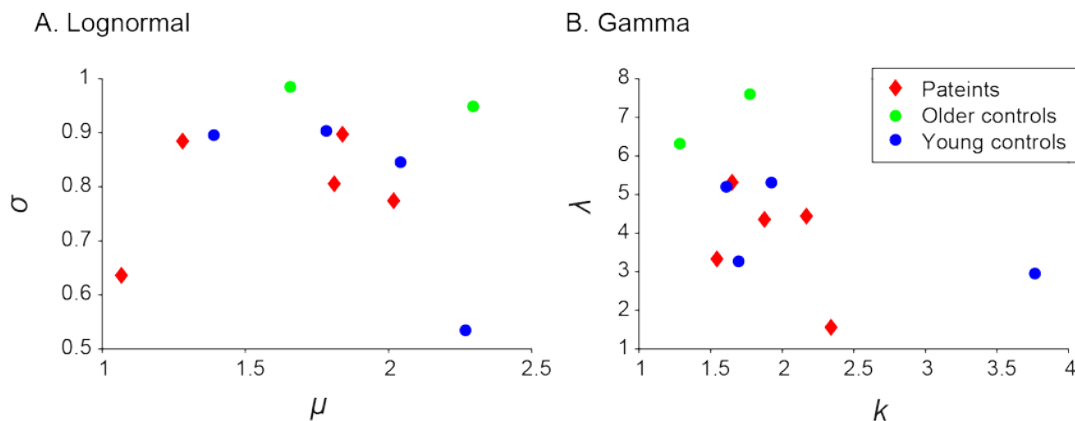
**Figure S6.2.** Axial slices of spatially normalized MR images for all patients with spherical ROI outlines overlaid (left to right shows inferior to superior).



**Figure S6.3.** A) Scatter plot comparing mean percept durations for the three different bistable stimulus types. RS = rotating sphere; BR = binocular rivalry; DQ = dot quartet. Outlined markers indicate patient data. B) Scatter plot of observer age against mean rotating sphere percept duration. C) Scatter plot of the percentage of lesioned voxels against mean rotating sphere percept duration for each ROI for all patients.



**Figure S6.4.** A) Perceptual reports for patient PF during one block of the dynamic disparity task. On this task, inter-trial intervals were always 1 second (grey bars), but stimulus duration was self-paced (i.e. variable) and averaged  $1.03 \pm 0.04$ s. The stimulus rotated clockwise on half of the trials in each block, and this result therefore shows a strong bias. B) Perceptual reports for the same patient during intermittent presentation of the bistable rotating sphere stimulus (i.e. similar to the dynamic disparity task, but without disparity).



**Figure S6.5.** Relationship between shape and scale parameters for best A) lognormal and B) gamma function fits to individual percept duration distributions. The plots reveal no clear segregation of groups based on fit parameters.

**Table S6.3.** Individual percept duration distribution statistics for the rotating sphere stimulus. All summary statistics are given in seconds (3dp).

Summary statistics						
Subject t	Mean	SD	SE	Media n	32% CI	68% CI
MH	3.655	3.246	0.112	2.631	2.099	3.631
PM	5.146	4.818	0.318	3.924	2.474	5.476
PF	8.169	6.038	0.403	6.454	4.171	9.930
MP	8.775	7.002	1.698	8.285	4.422	9.800
RH	9.636	6.249	0.544	8.465	5.577	11.892
DC	8.116	7.360	0.624	4.607	3.205	8.652
RC	13.502	8.354	1.638	13.277	9.698	17.928
APM	8.370	6.444	0.609	5.715	4.101	10.599
MLP	10.214	6.892	0.836	8.627	6.187	11.495
JMB	11.110	5.997	0.622	8.952	7.339	12.930
CH	5.543	4.220	0.554	4.477	3.140	6.367

Lognormal fit data									
$\mu$ (shape)	$\mu$ CI	$\mu$ CI	$\sigma$ (scale)	$\sigma$ CI	$\sigma$ CI	df	N	$\chi^2$	p
1.067	1.025	1.110	0.636	0.607	0.668	3	846	35.721	0.000
1.281	1.166	1.396	0.885	0.810	0.974	3	229	2.935	0.402
1.811	1.705	1.916	0.806	0.737	0.888	5	225	<b>7.702</b>	0.173
1.839	1.378	2.301	0.897	0.668	1.366	0	17	<b>0.777</b>	-
2.018	1.884	2.151	0.774	0.691	0.881	5	132	<b>6.656</b>	0.248
1.657	1.492	1.822	0.985	0.881	1.117	4	139	9.958	0.041
2.296	1.912	2.679	0.949	0.744	1.310	0	26	<b>2.260</b>	-
1.783	1.614	1.952	0.903	0.799	1.040	4	112	<b>10.861</b>	0.028
2.042	1.837	2.247	0.846	0.724	1.018	3	68	<b>5.986</b>	0.112
2.269	2.159	2.379	0.535	0.467	0.625	4	93	1.011	0.908
1.390	1.154	1.625	0.896	0.757	1.097	2	58	<b>1.300</b>	0.522

Gamma fit data									
k (shape)	k CI	k CI	$\lambda$ (scale)	$\lambda$ CI	$\lambda$ CI	df	N	$\chi^2$	p
2.341	2.141	2.560	1.561	1.413	1.725	2	846	<b>99.648</b>	0.000
1.544	1.306	1.825	3.333	2.737	4.058	3	229	<b>3.257</b>	0.354
1.876	1.581	2.225	4.355	3.581	5.297	5	225	4.319	0.504
1.650	0.891	3.056	5.318	2.591	10.912	0	17	0.191	-
2.169	1.732	2.717	4.442	3.448	5.722	5	132	2.256	0.813
1.285	1.040	1.588	6.315	4.882	8.169	5	139	<b>16.904</b>	0.005
1.776	1.076	2.930	7.604	4.266	13.553	0	26	0.412	-
1.609	1.266	2.045	5.202	3.929	6.887	5	112	9.590	0.088
1.924	1.409	2.627	5.308	3.720	7.573	3	68	4.755	0.191
3.765	2.858	4.960	2.951	2.197	3.963	4	93	<b>2.426</b>	0.658
1.696	1.214	2.369	3.268	2.216	4.819	2	58	0.591	0.744

In this thesis, I began by arguing that there are two fundamental strategies for reducing uncertainty when inferring the 3D structure of the world from noisy or ambiguous visual input. The first method is to integrate depth information from multiple independent sources within the retinal image (depth cues). The second strategy used by the primate visual system is to use prior experience to resolve ambiguities. The experimental chapters presented here investigated a specific example of each of these solutions to the inference problem. In chapters 2 and 3, I considered the problem of cue combination for qualitatively different types of depth cue (texture and binocular disparity) and examined the contribution of specific cortical areas to the computation of an integrated depth estimate for surface slant. In chapters 4 and 5, I examined how prior experience, encoded implicitly and within a retinotopic coordinate frame, influences future perception by integrating information over different time scales. Finally, in chapter 6 I presented evidence for a causal role of sub-regions of posterior parietal cortex in perception of depth from binocular disparity signals, but a lack of involvement in short term perceptual memory of the type investigated in the preceding chapters. The findings of these studies advance our understanding of the coarse-scale organization of neural systems for actively interpreting visual depth information, but also highlight some of the challenges of contemporary cognitive neuroscience.

In order to investigate functions of the human brain, it seems intuitive to study human subjects. However, one major limitation of studying human brains is that the non-invasive imaging and perturbation technologies currently available for doing so lack spatial specificity. Thus, our ability to perform informative experiments on human subjects is highly dependent on the scale of neural organization at which a given process of interest is implemented. The success of neuropsychology and neuroimaging in examining the processing of visual features such as motion, color and objects likely reflects the coarse scale organization of neural populations (or “modules”) that process these properties (Sunaert et al., 1999; Lafer-Sousa and Conway, 2013). In contrast, other visual features such as orientation, ocular dominance, and binocular disparity tuning appear to be encoded

at a finer spatial scale (DeAngelis and Newsome, 1999; Kara and Boyd, 2009). Further, due to the coarse spatial scale and uncontrolled nature of brain lesions, some inferences made by early neuropsychology studies have since proven to be inaccurate (Annese et al., 2014).

Several findings reported here point to such limitations (past and present) of neuroscientific methods. For example, studies of stereopsis conducted decades ago using groups of patients with unilateral neglect appeared to support a causal role for parietal cortex in perceiving depth from disparity. However, in chapter 6 we reported two cases of patients with damage to regions within parietal cortex (identified using MRI) who display normal stereopsis. Similarly, in chapter 2 we did not detect differences in the fMRI BOLD responses to different depth cue combinations when we used conventional univariate analysis of regions of interest (mean signal change within  $\sim 1350\text{mm}^3$  of cortex). However, using multivariate analysis methods that preserve information at the single voxel level ( $4.5\text{mm}^3$ ), we were able to detect more information in specific regions when combined depth cues were presented compared to individual cues. While this result indicated the involvement of V3B/KO in the integration of depth cues, in Chapter 3 we failed to find a behavioural effect of disrupting this area using TMS. There are several possible reasons for this result, but the one that is most problematic for contemporary human neuroscience is that our most advanced tools remain too coarse relative to the spatial organization of certain neural functions, such as visual depth perception.

A central tenet of cognitive neuroscience has always been to study function at multiple levels of structural organization (Churchland and Sejnowski, 1988). In order to improve our understanding of functions such as visual depth processing, we must therefore zoom in on brain regions of interest. Several groups are already utilizing developing technologies such as high-field (7T) fMRI to investigate disparity maps within individual visual areas of human cortex (Coullon et al., 2012; Gonçalves et al., 2014). Another solution is to use animal models, where perturbation techniques can be applied with greater spatial and temporal precision and can be more conveniently verified (Krug et al., 2013). Developments in the manufacture of multi-electrode arrays and digital signal processing technologies are allowing researchers to record simultaneously from larger neural ensembles, as well as studying local micro-circuitry at the laminar and columnar levels (Longordo et al., 2013; Maier, 2013). Microscopic imaging methods capable of resolving the structure and activity of individual neurons whilst recording simultaneously from a large volume (or even a whole brain) are becoming available for some animal models



(Dodt et al., 2007; Ahrens et al., 2013; Chung and Deisseroth, 2013), while lower resolution optical imaging methods are already being used to study binocular disparity maps in primate visual cortex (Chen et al., 2008). Integrating the insights into the local processing of visual depth information that these methods can deliver with our current understanding of functional organization at the macroscopic level may guide future systems-level research.

The ability to perceive depth is critical for successfully executing reaching and grasping movements that allow us to interact in the world. However, patients with brain injuries may show reaching and grasping impairments for a variety of reasons, including perceptual and attentional deficits, as well as difficulties executing motor commands. Parietal cortex is strongly implicated in all of these functions, so functional deficits must be assessed behaviourally (Cavina-Pratesi et al., 2010). In Chapter 6 we observed that a subset of patients with lesions to parietal cortex retained normal depth perception from binocular disparity signals, despite deficits in grasping, while others did not. Conversely, lesions to ventral cortex may produce deficits in depth perception without visuomotor deficits (Dijkerman et al., 1996). Behavioural assessment of relative deficits in depth perception and visually-guided actions could therefore play an important role in guiding patient rehabilitation strategies. The plasticity of topographic representations observed in chapters 4 and 5 implies that deficits affecting one part of the visuo-spatial field might be attenuated through re-training. Further, the effect of prior experience at a cognitive level (e.g. expectations or beliefs) may provide patients with alternative strategies for performing reaching and grasping actions (Rossetti et al., 2005).

The experiments presented here were designed to isolate the effects of cue integration and prior experience on the processing of uncertainty in visual depth signals. Of course, in the real world, both strategies interact. For example, we have prior expectations that other road users steer their vehicles to avoid collisions, but we also know that human or mechanical errors can occur. At the same time we can accumulate sensory information (such as perspective and disparity cues) to estimate the trajectory and time-to-impact of an oncoming vehicle, and we integrate this evidence (likelihood distribution) and experience (prior distribution) to reach a decision on whether we need to swerve out of the way. The work presented in this thesis suggests that integration of visual depth cues occurs in early extrastriate cortex, and that the influence of prior experience on depth judgments may be encoded topographically in the visual system. Damage to parietal cortex does not

necessarily interfere with either of these processes, although it can cause deficits in the perception of depth from binocular disparities. A promising future direction will be to investigate the processing of visual depth and sensory uncertainty at finer scales of organization within these cortical areas.

## References

- Abeles, M., and Goldstein Jr, M.H. (1970). Functional architecture in cat primary auditory cortex: columnar organization and organization according to depth. *Journal of Neurophysiology*.
- Adams, W.J., Graf, E.W., and Ernst, M.O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience* 7, 1057-1058.
- Afraz, A., and Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *Journal of Vision* 9.
- Afraz, A., Pashkam, M.V., and Cavanagh, P. (2010). Spatial heterogeneity in the perception of face and form attributes. *Current Biology* 20, 2112-2116.
- Ahrens, M.B., Orger, M.B., Robson, D.N., Li, J.M., and Keller, P.J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nature methods* 10, 413-420.
- Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., and Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience* 30, 2960-2966.
- Amaral, D.G., Behnia, H., and Kelly, J.L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118, 1099-1120.
- Andersen, S.M., Rapsak, S.Z., and Beeson, P.M. (2010). Cost function masking during normalization of brains with focal lesions: still a necessity? *Neuroimage* 53, 78-84.
- Andrade, J., Kemps, E., Werniers, Y., May, J., and Szmalec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology: Section A* 55, 753-774.
- Angelaki, D.E., Gu, Y., and Deangelis, G.C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current opinion in neurobiology* 19, 452-458.
- Annese, J., Schenker-Ahmed, N.M., Bartsch, H., Maechler, P., Sheh, C., Thomas, N., Kayano, J., Ghatan, A., Bresler, N., and Frosch, M.P. (2014). Postmortem examination of patient HM's brain based on histological sectioning and digital 3D reconstruction. *Nature communications* 5.
- Anstis, S., and Ramachandran, V.S. (1987). Visual inertia in apparent motion. *Vision research* 27, 755-764.
- Anzai, A., Chowdhury, S.A., and Deangelis, G.C. (2011). Coding of stereoscopic depth information in visual areas V3 and V3A. *The Journal of Neuroscience* 31, 10270-10282.
- Anzai, A., Ohzawa, I., and Freeman, R.D. (2001). Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect. *Nature neuroscience* 4, 513-518.
- Arendt, D., and Wittbrodt, J. (2001). Reconstructing the eyes of Urbilateria. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356, 1545-1563.
- Ashburner, J., and Friston, K.J. (2005). Unified segmentation. *Neuroimage* 26, 839-851.
- Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience* 7, 358-366.

- Backus, B.T., Banks, M.S., Van Ee, R., and Crowell, J.A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *VISION RESEARCH- OXFORD* 39, 1143-1170.
- Backus, B.T., Fleet, D.J., Parker, A.J., and Heeger, D.J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of neurophysiology* 86, 2054-2068.
- Backus, B.T., and Haijiang, Q. (2007). Competition between newly recruited and pre-existing visual cues during the construction of visual appearance. *Vision Research* 47, 919-924.
- Baizer, J.S., Ungerleider, L.G., and Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *The Journal of neuroscience* 11, 168-190.
- Ban, H., Preston, T.J., Meeson, A., and Welchman, A.E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience*.
- Bargmann, C.I. (2006). Chemosensation in *C. elegans*.
- Barker, A.T. (1991). An introduction to the basic principles of magnetic nerve stimulation. *Journal of Clinical Neurophysiology* 8, 26-37.
- Barlow, H.B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory communication*, 217-234.
- Barton, R.A. (1998). Visual specialization and brain evolution in primates. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265, 1933-1937.
- Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695-711.
- Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M.J., Hénaff, M.-A., Michèl, F., and Barton, J.J.S. (2001). Unilateral right parietal damage leads to bilateral deficit for high-level motion. *Neuron* 32, 985-995.
- Bayes, T., and Price, M. (1763). An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFRS. *Philosophical Transactions (1683-1775)*, 370-418.
- Beatty, J. (2001). The neuronal empiricism hypothesis: Cerebral cortex is an unsupervised knowledge-seeking neural network. *Neurocomputing* 38, 1095-1100.
- Beck, J.M., Ma, W.J., Kiani, R., Hanks, T., Churchland, A.K., Roitman, J., Shadlen, M.N., Latham, P.E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron* 60, 1142-1152.
- Beck, J.M., Ma, W.J., Pitkow, X., Latham, P.E., and Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* 74, 30-39.
- Benton, A.L., and Hécaen, H. (1970). Stereoscopic vision in patients with unilateral cerebral disease. *Neurology* 20, 1084-1084.
- Bhattacharyya, R., Musallam, S., and Andersen, R.A. (2009). Parietal reach region encodes reach depth using retinal disparity and vergence angle signals. *Journal of neurophysiology* 102, 805-816.
- Blake, R., and Logothetis, N.K. (2002). Visual competition. *Nat Rev Neurosci* 3, 13-21.
- Blake, R., Sobel, K.V., and Gilroy, L.A. (2003). Visual motion retards alternations between conflicting perceptual interpretations. *Neuron* 39, 869-878.
- Bolognini, N., Miniussi, C., Savazzi, S., Bricolo, E., and Maravita, A. (2009). TMS modulation of visual and auditory processing in the posterior parietal cortex. *Experimental brain research* 195, 509-517.
- Bona, S., Herbert, A., Toneatto, C., Silvanto, J., and Cattaneo, Z. (2014). The causal role of the lateral occipital complex in visual mirror symmetry detection and grouping: An fMRI-guided TMS study. *Cortex* 51, 46-55.

- Bonneh, Y.S., Pavlovskaya, M., Ring, H., and Soroker, N. (2004). Abnormal binocular rivalry in unilateral neglect: evidence for a non-spatial mechanism of extinction. *Neuroreport* 15, 473-477.
- Bono, M.D., and Villu Maricq, A. (2005). Neuronal substrates of complex behaviors in *C. elegans*. *Annu. Rev. Neurosci.* 28, 451-501.
- Booth, M.C., and Rolls, E.T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex* 8, 510-523.
- Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S., and Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik* 10, 139-144.
- Bourke, A.F.G. (2011). The validity and value of inclusive fitness theory. *Proceedings of the Royal Society B: Biological Sciences*, rspb20111465.
- Bradley, D.C., Chang, G.C., and Andersen, R.A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392, 714-717.
- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433-436.
- Brascamp, J.W., Kanai, R., Walsh, V., and Van Ee, R. (2010). Human middle temporal cortex, perceptual bias, and perceptual memory for ambiguous three-dimensional motion. *The Journal of neuroscience* 30, 760.
- Brascamp, J.W., Knapen, T.H., Kanai, R., Noest, A.J., Van Ee, R., and Van Den Berg, A.V. (2008). Multi-timescale perceptual history resolves visual ambiguity. *PLoS One* 3, e1497.
- Brascamp, J.W., Pearson, J., Blake, R., and Van Den Berg, A.V. (2009). Intermittent ambiguous stimuli: implicit memory causes periodic perceptual alternations. *J Vis* 9, 3 1-23.
- Brascamp, J.W., Van Ee, R., Noest, A.J., Jacobs, R.H.a.H., and Van Den Berg, A.V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of vision* 6, 8.
- Braunstein, M.L. (1968). Motion and texture as sources of slant information. *Journal of Experimental Psychology* 78, 247.
- Brett, M., Leff, A.P., Rorden, C., and Ashburner, J. (2001). Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14, 486-500.
- Bridge, H., Thomas, O.M., Minini, L., Cavina-Pratesi, C., Milner, A.D., and Parker, A.J. (2013). Structural and functional changes across the visual cortex of a patient with visual form agnosia. *The Journal of Neuroscience* 33, 12779-12791.
- Brincat, S.L., and Connor, C.E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature neuroscience* 7, 880-886.
- Britz, J., Landis, T., and Michel, C.M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex* 19, 55-65.
- Britz, J., Pitts, M.A., and Michel, C.M. (2011). Right parietal brain activity precedes perceptual alternation during binocular rivalry. *Human brain mapping* 32, 1432-1442.
- Brouwer, G.J., and Van Ee, R. (2006). Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Res* 46, 3393-3402.
- Brouwer, G.J., and Van Ee, R. (2007). Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *The Journal of neuroscience* 27, 1015.
- Brouwer, G.J., Van Ee, R., and Schwarzbach, J. (2005). Activation in visual cortex correlates with the awareness of stereoscopic depth. *The Journal of neuroscience* 25, 10403-10413.
- Brown, B., Yap, M.K.H., and Fan, W. (1993). Decrease in stereoacuity in the seventh decade of life. *Ophthalmic and Physiological Optics* 13, 138-142.

- Büchert, M., Greenlee, M., Rutschmann, R.M., Kraemer, F., Luo, F., and Hennig, J. (2002). Functional magnetic resonance imaging evidence for binocular interactions in human visual cortex. *Experimental Brain Research* 145, 334-339.
- Bülthoff, H.H., and Mallot, H.A. (1988). Integration of depth modules: stereo and shading. *Josa a* 5, 1749-1758.
- Burt, P., and Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science* 208, 615-617.
- Buschman, T.J., and Miller, E.K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science* 315, 1860-1862.
- Busetini, C., Miles, F.A., and Krauzlis, R.J. (1996). Short-latency disparity vergence responses and their dependence on a prior saccadic eye movement. *Journal of Neurophysiology* 75, 1392-1410.
- Cabeza, R., Ciaramelli, E., Olson, I.R., and Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience* 9, 613-625.
- Carmel, D., Walsh, V., Lavie, N., and Rees, G. (2010). Right parietal TMS shortens dominance durations in binocular rivalry. *Current Biology* 20, R799-R800.
- Carmon, A., and Bechtoldt, H.P. (1969). Dominance of the right cerebral hemisphere for stereopsis. *Neuropsychologia* 7, 29-39.
- Carrasco, M., Penpeci-Talgar, C., and Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vision research* 40, 1203-1215.
- Carter, O., and Cavanagh, P. (2007). Onset rivalry: brief presentation isolates an early independent phase of perceptual competition. *PLoS One* 2, e343.
- Carter, O.L., Hasler, F., Pettigrew, J.D., Wallis, G.M., Liu, G.B., and Vollenweider, F.X. (2007). Psilocybin links binocular rivalry switch rate to attention and subjective arousal levels in humans. *Psychopharmacology* 195, 415-424.
- Carter, O.L., and Pettigrew, J.D. (2003). A common oscillator for perceptual rivalries? *PERCEPTION-LONDON-* 32, 295-306.
- Carter, O.L., Pettigrew, J.D., Hasler, F., Wallis, G.M., Liu, G.B., Hell, D., and Vollenweider, F.X. (2005). Modulating the rate and rhythmicity of perceptual rivalry alternations with the mixed 5-HT<sub>2A</sub> and 5-HT<sub>1A</sub> agonist psilocybin. *Neuropsychopharmacology* 30, 1154-1162.
- Cavada, C., and Goldman - Rakic, P.S. (1989). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *Journal of Comparative Neurology* 287, 393-421.
- Cavada, C., Tejedor, J., Cruz-Rizzolo, R.J., and Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral Cortex* 10, 220-242.
- Cavina-Pratesi, C., Connolly, J.D., and Milner, A.D. (2013). Optic ataxia as a model to investigate the role of the posterior parietal cortex in visually guided action: evidence from studies of patient MH. *Frontiers in human neuroscience* 7.
- Cavina-Pratesi, C., Ietswaart, M., Humphreys, G.W., Lestou, V., and Milner, A.D. (2010). Impaired grasping in a patient with optic ataxia: primary visuomotor deficit or secondary consequence of misreaching? *Neuropsychologia* 48, 226-234.
- Chalmers, D.J. (2000). What is a neural correlate of consciousness. *Neural correlates of consciousness: Empirical and conceptual questions*, 17-39.
- Chandrasekaran, C., Canon, V., Dahmen, J.C., Kourtzi, Z., and Welchman, A.E. (2007). Neural correlates of disparity-defined shape discrimination in the human brain. *Journal of Neurophysiology* 97, 1553-1565.

- Chen, G., Lu, H.D., and Roe, A.W. (2008). A map for horizontal disparity in monkey V2. *Neuron* 58, 442-450.
- Chen, X., and He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Current Biology* 14, 1013-1017.
- Chikkerur, S., Serre, T., Tan, C., and Poggio, T. (2010). What and where: A Bayesian inference theory of attention. *Vision research* 50, 2233-2247.
- Chong, S.C., Tadin, D., and Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of vision* 5, 6.
- Chopin, A., and Mamassian, P. (2012). Predictive properties of visual adaptation. *Current biology* 22, 622-626.
- Chung, K., and Deisseroth, K. (2013). CLARITY for mapping the nervous system. *Nature methods* 10, 508-513.
- Churchland, M.M., Byron, M.Y., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., and Scott, B.B. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature neuroscience* 13, 369-378.
- Churchland, P.S., and Sejnowski, T.J. (1988). Perspectives on cognitive neuroscience. *Science* 242, 741-745.
- Clark, J.J., and Yuille, A.L. (1990). *Data fusion for sensory information processing systems*. Kluwer Academic Publishers.
- Clopper, C.J., and Pearson, E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 404-413.
- Cohen, M.A., Cavanagh, P., Chun, M.M., and Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in cognitive sciences* 16, 411-417.
- Cohen, M.A., and Dennett, D.C. (2011). Consciousness cannot be separated from function. *Trends in cognitive sciences* 15, 358-364.
- Cohen, Y.E., and Andersen, R.A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience* 3, 553-562.
- Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., and Van Essen, D.C. (1998). A common network of functional areas for attention and eye movements. *Neuron* 21, 761-773.
- Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., and Petersen, S.E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248, 1556-1559.
- Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* 3, 201-215.
- Coullon, G.S.L., Sanchez-Panchuelo, R.M., Francis, S., Schluppeck, D., Parker, A.J., and Bridge, H. (2012). Investigating disparity organisation in the human early visual cortex with high resolution magnetic resonance imaging (7 Tesla). *Journal of Vision* 12, 41-41.
- Cowey, A., and Porter, J. (1979). Brain damage and global stereopsis. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 204, 399-407.
- Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., and Friston, K. (2007). Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *Neuroimage* 37, 866-875.
- Cumming, B., and Deangelis, G. (2001). The physiology of stereopsis. *Annual review of neuroscience* 24, 203-238.
- Cumming, B.G., and Parker, A.J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389, 280-283.

- Cumming, B.G., and Parker, A.J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *The Journal of neuroscience* 19, 5602-5618.
- Daini, R., Facchin, A., Bignotti, M., Lentini, C., Peverelli, M., O'shea, R.P., and Molteni, F. (2010). Neuropsychological evidence of high-level processing in binocular rivalry. *Behavioural neurology* 23, 233-235.
- Danta, G., Hilton, R.C., and O'boyle, D.J. (1978). Hemisphere function and binocular depth perception. *Brain* 101, 569-589.
- De Graaf, T., and Sack, A. (2011a). Null results in TMS: from absence of evidence to evidence of absence. *Neuroscience & Biobehavioral Reviews* 35, 871-877.
- De Graaf, T.A., and Sack, A.T. (2011b). Null results in TMS: from absence of evidence to evidence of absence. *Neuroscience & Biobehavioral Reviews* 35, 871-877.
- De Haan, E.H.F., and Cowey, A. (2011). On the usefulness of 'what' and 'where' pathways in vision. *Trends in cognitive sciences* 15, 460-466.
- De Jong, M.C., Brascamp, J.W., Kemner, C., Van Ee, R., and Verstraten, F.a.J. (2014). Implicit Perceptual Memory Modulates Early Visual Processing of Ambiguous Images. *The Journal of Neuroscience* 34, 9970-9981.
- De Jong, M.C., Knapen, T., and Van Ee, R. (2012a). Opposite influence of perceptual memory on initial and prolonged perception of sensory ambiguity. *PloS one* 7, e30595.
- De Jong, M.C., Kourtzi, Z., and Ee, R. (2012b). Perceptual experience modulates cortical circuits involved in visual awareness. *European Journal of Neuroscience*.
- De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., and Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage* 43, 44-58.
- Deangelis, G.C., and Newsome, W.T. (1999). Organization of disparity-selective neurons in macaque area MT. *The Journal of neuroscience* 19, 1398-1415.
- Deco, G., Rolls, E.T., and Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Progress in neurobiology* 88, 1-16.
- Den Ouden, H.E.M., Daunizeau, J., Roiser, J., Friston, K.J., and Stephan, K.E. (2010). Striatal prediction error modulates cortical coupling. *The Journal of Neuroscience* 30, 3210-3219.
- Denison, R.N., Piazza, E.A., and Silver, M.A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in Human Neuroscience* 5.
- Dent, K., Lestou, V., and Humphreys, G.W. (2010). Deficits in visual search for conjunctions of motion and form after parietal damage but with spared hMT+/V5. *Cognitive neuropsychology* 27, 72-99.
- Di Luca, M., Ernst, M.O., and Backus, B.T. (2010). Learning to use an invisible visual signal for perception. *Current Biology*.
- Diamond, I.T., and Hall, W.C.T. (1969). Evolution of neocortex. *Science* 164, 251-262.
- Dijkerman, H.C., Milner, A.D., and Carey, D.P. (1996). The perception and prehension of objects oriented in the depth plane. *Experimental Brain Research* 112, 442-451.
- Ding, S.L., Van Hoesen, G., and Rockland, K.S. (2000). Inferior parietal lobule projections to the presubiculum and neighboring ventromedial temporal cortical areas. *Journal of Comparative Neurology* 425, 510-530.
- Dodd, J.V., Krug, K., Cumming, B.G., and Parker, A.J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of neuroscience* 21, 4809.
- Dodt, H.-U., Leischner, U., Schierloh, A., Jährling, N., Mauch, C.P., Deininger, K., Deussing, J.M., Eder, M., Zieglgänsberger, W., and Becker, K. (2007).



- Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature methods* 4, 331-336.
- Doherty, J.R., Rao, A., Mesulam, M.M., and Nobre, A.C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *The Journal of neuroscience* 25, 8259-8266.
- Dövençioğlu, D., Ban, H., Schofield, A.J., and Welchman, A.E. (2013). Perceptual Integration for Qualitatively Different 3-D Cues in the Human Brain. *Journal of Cognitive Neuroscience* 25, 1527-1541.
- Downing, P.E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science* 293, 2470-2473.
- Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647-660.
- Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A.M., Michiels, J., Marchal, G., Mortelmans, L., and Orban, G. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex* 7, 283-292.
- Durand, J.-B., Nelissen, K., Joly, O., Wardak, C., Todd, J.T., Norman, J.F., Janssen, P., Vanduffel, W., and Orban, G.A. (2007). Anterior regions of monkey parietal cortex process visual 3D shape. *Neuron* 55, 493-505.
- Durand, J.-B., Peeters, R., Norman, J.F., Todd, J.T., and Orban, G.A. (2009). Parietal regions processing visual 3D shape extracted from disparity. *Neuroimage* 46, 1114-1126.
- Durnford, M., and Kimura, D. (1971). Right hemisphere specialization for depth perception reflected in visual field differences.
- Eby, D.W., Loomis, J.M., and Solomon, E.M. (1989). Perceptual linkage of multiple objects rotating in depth. *Perception* 18, 427-444.
- Ellison, A., and Cowey, A. (2009). Differential and co-involvement of areas of the temporal and parietal streams in visual tasks. *Neuropsychologia* 47, 1609-1614.
- Engel, S., Zhang, X., and Wandell, B. (1997). Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* 388, 68-71.
- Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature* 392, 598-601.
- Erclik, T., Hartenstein, V., Lipshitz, H.D., and McInnes, R.R. (2008). Conserved Role of the *Vsx* Genes Supports a Monophyletic Origin for Bilaterian Visual Systems. *Current Biology* 18, 1278-1287.
- Ernst, M.O., and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429-433.
- Fairhall, A.L., Lewen, G.D., Bialek, W., and Van Steveninck, R.R.D.R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature* 412, 787-792.
- Feldmeyer, D., Brecht, M., Helmchen, F., Petersen, C.C.H., Poulet, J.F.A., Staiger, J.F., Luhmann, H.J., and Schwarz, C. (2013). Barrel cortex function. *Progress in neurobiology* 103, 3-27.
- Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex* 1, 1-47.
- Ferraina, S., Paré, M., and Wurtz, R.H. (2000). Disparity sensitivity of frontal eye field neurons. *Journal of neurophysiology* 83, 625-629.
- Foellmer, M.W., and Fairbairn, D.J. (2003). Spontaneous male death during copulation in an orb-weaving spider. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, S183-S185.
- Fox, R., Todd, S., and Bettinger, L.A. (1975). Optokinetic nystagmus as an objective indicator of binocular rivalry. *Vision research* 15, 849-853.

- Frässle, S., Sommer, J., Jansen, A., Naber, M., and Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *The Journal of Neuroscience* 34, 1738-1747.
- Freedman, D.J., and Assad, J.A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature* 443, 85-88.
- Freese, J.L., and Amaral, D.G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *Journal of Comparative Neurology* 486, 295-317.
- Freiwald, W.A., and Tsao, D.Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330, 845-851.
- Fukushima, K., Yamanobe, T., Shinmei, Y., Fukushima, J., Kurkin, S., and Peterson, B.W. (2002). Coding of smooth eye movements in three-dimensional space by frontal cortex. *Nature* 419, 157-162.
- Gallistel, C.R., Fairhurst, S., and Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the national academy of Sciences of the united States of america* 101, 13124.
- Genovesio, A., and Ferraina, S. (2004). Integration of retinal disparity and fixation-distance related signals toward an egocentric coding of distance in the posterior parietal cortex of primates. *Journal of neurophysiology* 91, 2670-2684.
- Georgieva, S., Peeters, R., Kolster, H., Todd, J.T., and Orban, G.A. (2009). The processing of three-dimensional shape from disparity in the human brain. *The Journal of Neuroscience* 29, 727-742.
- Georgieva, S.S., Todd, J.T., Peeters, R., and Orban, G.A. (2008). The extraction of 3D shape from texture and shading in the human brain. *Cerebral cortex* 18, 2416-2438.
- Gerardin, P., Kourtzi, Z., and Mamassian, P. (2010). Prior knowledge of illumination for 3D perception in the human brain. *Proceedings of the National Academy of Sciences* 107, 16309-16314.
- Gerits, A., Ruff, C.C., Guipponi, O., Wenderoth, N., Driver, J., and Vanduffel, W. (2011). Transcranial magnetic stimulation of macaque frontal eye fields decreases saccadic reaction time. *Experimental brain research* 212, 143-152.
- Gershman, S.J., Vul, E., and Tenenbaum, J.B. (2012). Multistability and perceptual inference. *Neural computation* 24, 1-24.
- Gibson, J.J. (1950). The perception of visual surfaces. *American journal of psychology* 63, 367-384.
- Gibson, J.J. (1978). The ecological approach to the visual perception of pictures. *Leonardo* 11, 227-235.
- Gigante, G., Mattia, M., Braun, J., and Del Giudice, P. (2009). Bistable perception modeled as competing stochastic integrations at two levels. *PLoS computational biology* 5, e1000430.
- Gilaie-Dotan, S., Saygin, A.P., Lorenzi, L.J., Egan, R., Rees, G., and Behrmann, M. (2013). The role of human ventral visual cortex in motion perception. *Brain* 136, 2784-2798.
- Gilaie - Dotan, S., Ullman, S., Kushnir, T., and Malach, R. (2002). Shape - selective stereo processing in human object - related visual areas. *Human brain mapping* 15, 67-79.
- Gillam, B. (1972). Perceived common rotary motion of ambiguous stimuli as a criterion of perceptual grouping. *Perception & Psychophysics* 11, 99-101.
- Gillam, B., and Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *PERCEPTION-LONDON-* 21, 427-427.

- Gilroy, L.A., and Blake, R. (2004). Physics embedded in visual perception of three-dimensional shape from motion. *Nature neuroscience* 7, 921-922.
- Girshick, A.R., and Banks, M.S. (2009). Probabilistic combination of slant information: Weighted averaging and robustness as optimal percepts. *Journal of Vision* 9.
- Girshick, A.R., Landy, M.S., and Simoncelli, E.P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature neuroscience* 14, 926-932.
- Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in cognitive sciences* 5, 10-16.
- Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535-574.
- Goldstein, E.B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance* 13, 256.
- Goncalves, N., Ban, H., Sanchez-Panchuelo, R., Francis, S., Schluppeck, D., and Welchman, A. (2014). Cortical organization of binocular disparity in human V3A. *Journal of Vision* 14, 970-970.
- Goodale, M.A., and Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences* 15, 20-25.
- Gori, M., Del Viva, M., Sandini, G., and Burr, D.C. (2008). Young children do not integrate visual and haptic form information. *Current Biology* 18, 694-698.
- Goris, R.L.T., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. *Nature neuroscience*.
- Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron* 53, 9-16.
- Grefkes, C., and Fink, G.R. (2005). REVIEW: The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of anatomy* 207, 3-17.
- Grill-Spector, K., Kushnir, T., Hendler, T., and Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature neuroscience* 3, 837-843.
- Grossmann, J.K., and Dobbins, A.C. (2003). Differential ambiguity reduces grouping of metastable objects. *Vision research* 43, 359-369.
- Gu, Y., Angelaki, D.E., and Deangelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature neuroscience* 11, 1201-1210.
- Gulyas, B., Heywood, C., Popplewell, D., Roland, P., and Cowey, A. (1994). Visual form discrimination from color or motion cues: functional anatomy by positron emission tomography. *Proceedings of the National Academy of Sciences* 91, 9965.
- Hadjikhani, N., Liu, A.K., Dale, A.M., Cavanagh, P., and Tootell, R.B.H. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nature neuroscience* 1, 235-241.
- Haijiang, Q., Saunders, J.A., Stone, R.W., and Backus, B.T. (2006). Demonstration of cue recruitment: change in visual appearance by means of Pavlovian conditioning. *Proc Natl Acad Sci U S A* 103, 483-488.
- Hamsher, K. (1978). Stereopsis and unilateral brain disease. *Investigative Ophthalmology & Visual Science* 17, 336-343.
- Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature neuroscience* 9, 682-689.

- Harrison, S.J., and Backus, B.T. (2010a). Disambiguating Necker cube rotation using a location cue: What types of spatial location signal can the visual system learn? *Journal of Vision* 10.
- Harrison, S.J., and Backus, B.T. (2010b). Uninformative visual experience establishes long term perceptual bias. *Vision Res* 50, 1905-1911.
- Harrison, S.J., and Backus, B.T. (2012). Associative learning of shape as a cue to appearance: A new demonstration of cue recruitment. *Journal of Vision* 12.
- Harrison, S.J., and Backus, B.T. (2014). A trained perceptual bias that lasts for weeks. *Vision Research*.
- Harrison, S.J., Backus, B.T., and Jain, A. (2011). Disambiguation of Necker cube rotation by monocular and binocular depth cues: Relative effectiveness for establishing long-term bias. *Vision Research*.
- Harvey, B.M., Braddick, O.J., and Cowey, A. (2010). Similar effects of repetitive transcranial magnetic stimulation of MT+ and a dorsomedial extrastriate site including V3A on pattern detection and position discrimination of rotating and radial motion patterns. *Journal of vision* 10, 21.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., and Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34, 479-490.
- Haynes, J.-D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience* 8, 686-691.
- Heekeren, H., Marrett, S., Bandettini, P., and Ungerleider, L. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* 431, 859-862.
- Heekeren, H.R., Marrett, S., and Ungerleider, L.G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience* 9, 467-479.
- Heesy, C.P. (2009). Seeing in stereo: the ecology and evolution of primate binocular vision and stereopsis. *Evolutionary Anthropology: Issues, News, and Reviews* 18, 21-35.
- Helmholtz, H.V. (1925). Physiological optics. *Optical Society of America* 3, 318.
- Hilliard, M.A., Bargmann, C.I., and Bazzicalupo, P. (2002). *C. elegans* Responds to Chemical Repellents by Integrating Sensory Inputs from the Head and the Tail. *Current Biology* 12, 730-734.
- Hillis, J.M., Ernst, M.O., Banks, M.S., and Landy, M.S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science* 298, 1627-1630.
- Hillis, J.M., Watt, S.J., Landy, M.S., and Banks, M.S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision* 4, 967-992.
- Holmes, C.J., Hoge, R., Collins, L., Woods, R., Toga, A.W., and Evans, A.C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of computer assisted tomography* 22, 324-333.
- Holmes, G., and Horrax, G. (1919). Disturbances of spatial orientation and visual attention, with loss of stereoscopic vision. *Archives of Neurology & Psychiatry* 1, 385-407.
- Howard, I.P., and Rogers, B.J. (1995). *Binocular vision and stereopsis*. Oxford University Press, USA.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680-1683.
- Huang, Y., and Rao, R.P.N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science* 2, 580-593.
- Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P., and Rothwell, J.C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201-206.

- Hubbard, E.M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience* 6, 435-448.
- Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology* 160, 106.
- Hubel, D.H., and Wiesel, T.N. (1970a). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of physiology* 206, 419-436.
- Hubel, D.H., and Wiesel, T.N. (1970b). Stereoscopic vision in macaque monkey: cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225, 41-42.
- Hubel, D.H., and Wiesel, T.N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 1-59.
- Huk, A.C., and Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *The Journal of neuroscience* 25, 10420-10436.
- Hume, D. (2012). *A treatise of human nature*. Courier Dover Publications.
- Humphreys, G., Bickerton, W., Samson, D., and Riddoch, M. (2011). "The Birmingham Cognitive Screen (BCoS)". London: Psychology Press).
- Isbell, L.A. (2006). Snakes as agents of evolutionary change in primate brains. *Journal of Human Evolution* 51, 1-35.
- Jain, A., and Backus, B.T. (2011). Learned bias for 3-D shape perception without object motion. *Journal of Vision* 11, 981.
- Jain, A., and Backus, B.T. (2013). Generalization of Cue Recruitment to Non-moving Stimuli: Location and Surface-texture Contingent Biases for 3-D Shape Perception. *Vision research*.
- Jain, A., Fuller, S., and Backus, B.T. (2010). Absence of cue-recruitment for extrinsic signals: sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. *PLoS One* 5, e13295.
- James, T.W., Culham, J., Humphrey, G.K., Milner, A.D., and Goodale, M.A. (2003). Ventral occipital lesions impair object recognition but not object - directed grasping: an fMRI study. *Brain* 126, 2463-2475.
- Janssen, P., Vogels, R., and Orban, G.A. (1999). Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proceedings of the National Academy of Sciences* 96, 8217-8222.
- Janssen, P., Vogels, R., and Orban, G.A. (2000). Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* 288, 2054-2056.
- Janssens, T., Zhu, Q., Popivanov, I.D., and Vanduffel, W. (2014). Probabilistic and single-subject retinotopic maps reveal the topographic organization of face patches in the macaque cortex. *J Neurosci* 34, 10156-10167.
- Jeannerod, M., Arbib, M.A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in neurosciences* 18, 314-320.
- Johnston, E.B. (1991). Systematic distortions of shape from stereopsis. *Vision research* 31, 1351-1360.
- Johnston, E.B., Cumming, B.G., and Landy, M.S. (1994). Integration of stereopsis and motion shape cues. *Vision research* 34, 2259-2275.
- Johnston, E.B., Cumming, B.G., and Parker, A.J. (1993). Integration of depth modules: Stereopsis and texture. *Vision research* 33, 813-826.
- Julesz, B. (1971). Foundations of cyclopean perception.

- Kaas, J.H. (2000). Why is brain size so important: design problems and solutions as neocortex gets bigger or smaller. *Brain and Mind* 1, 7-23.
- Kaas, J.H., and Lyon, D.C. (2007). Pulvinar contributions to the dorsal and ventral streams of visual processing in primates. *Brain research reviews* 55, 285-296.
- Kalisvaart, J.P., Rampersad, S.M., and Goossens, J. (2011). Binocular onset rivalry at the time of saccades and stimulus jumps. *PloS one* 6, e20017.
- Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience* 8, 679-685.
- Kanai, R., Bahrami, B., and Rees, G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. *Current Biology* 20, 1626-1630.
- Kanai, R., Carmel, D., Bahrami, B., and Rees, G. (2011). Structural and functional fractionation of right superior parietal cortex in bistable perception. *Current Biology* 21, R106-R107.
- Kanai, R., Tsuchiya, N., and Verstraten, F.A.J. (2006). The scope and limits of top-down attention in unconscious visual processing. *Current Biology* 16, 2332-2336.
- Kanai, R., and Verstraten, F.A. (2005). Perceptual manifestations of fast neural plasticity: Motion priming, rapid motion aftereffect and perceptual sensitization. *Vision research* 45, 3109-3116.
- Kanai, R., and Verstraten, F.A.J. (2006). Attentional modulation of perceptual stabilization. *Proceedings of the Royal Society B: Biological Sciences* 273, 1217-1222.
- Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience* 17, 4302-4311.
- Kapoula, Z., Isotalo, E., Müri, R.M., Bucci, M.-P., and Rivaud-Péchoux, S. (2001). Effects of transcranial magnetic stimulation of the posterior parietal cortex on saccades and vergence. *Neuroreport* 12, 4041-4046.
- Kara, P., and Boyd, J.D. (2009). A micro-architecture for binocular disparity and ocular dominance in visual cortex. *Nature* 458, 627-631.
- Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences* 88, 4966.
- Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., and Ungerleider, L.G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751-761.
- Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271-304.
- Kindt, K.S., Quast, K.B., Giles, A.C., De, S., Hendrey, D., Nicastro, I., Rankin, C.H., and Schafer, W.R. (2007). Dopamine Mediates Context-Dependent Modulation of Sensory Plasticity in *C. elegans*. *Neuron* 55, 662-676.
- Kitadono, K., and Humphreys, G.W. (2007). Interactions between perception and action programming: Evidence from visual extinction and optic ataxia. *Cognitive neuropsychology* 24, 731-754.
- Kleinschmidt, A., Büchel, C., Zeki, S., and Frackowiak, R. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265, 2427-2433.
- Kleinschmidt, A., Sterzer, P., and Rees, G. (2012). Variability of perceptual multistability: from brain state to individual trait. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 988-1000.

- Klink, C., Lankheet, M., and Van Wezel, R. (2011). Intermittent motion stimuli stabilize neuronal responses in area MT: Implications for the perceptual stabilization of visual ambiguities. *Journal of Vision* 11, 765.
- Klink, C., Oleksiak, A., Lankheet, M.J.M., and Van Wezel, R.J.A. (2012). Intermittent stimulus presentation stabilizes neuronal responses in macaque area MT. *Journal of Neurophysiology* 108, 2101-2114.
- Knapen, T., Brascamp, J., Pearson, J., Van Ee, R., and Blake, R. (2011). The role of frontal and parietal brain areas in bistable perception. *The Journal of Neuroscience* 31, 10293-10301.
- Knapen, T., Brascamp, J.W., Adams, W.J., and Graf, E.W. (2009a). The spatial scale of perceptual memory in ambiguous figure perception. *J Vis* 9, 16 11-12.
- Knapen, T., Rolfs, M., and Cavanagh, P. (2009b). The reference frame of the motion aftereffect is retinotopic. *Journal of Vision* 9.
- Knill, D.C. (1998a). Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision research* 38, 2635-2656.
- Knill, D.C. (1998b). Surface orientation from texture: ideal observers, generic observers and the information content of texture cues. *Vision Research* 38, 1655-1682.
- Knill, D.C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Research* 43, 831-854.
- Knill, D.C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision* 7, 5.
- Knill, D.C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences* 27, 712-719.
- Knill, D.C., and Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.
- Knill, D.C., and Saunders, J.A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision research* 43, 2539-2558.
- Knoll, A.H., and Carroll, S.B. (1999). Early animal evolution: emerging views from comparative biology and geology. *Science* 284, 2129-2137.
- Koch, C., and Tsuchiya, N. (2012). Attention and consciousness: related yet different. *Trends in cognitive sciences* 16, 103-105.
- Koenderink, J.J., Van Doorn, A.J., and Kappers, A.M.L. (1992). Surface perception in pictures. *Attention, Perception and Psychophysics* 52, 487-496.
- Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of neurophysiology* 97, 3155-3164.
- Kok, P., Jehee, J.F.M., and De Lange, F.P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75, 265-270.
- Kolster, H., Janssens, T., Orban, G.A., and Vanduffel, W. (2014). The retinotopic organization of macaque occipitotemporal cortex anterior to V4 and caudoventral to the middle temporal (MT) cluster. 34, 10168-10191.
- Kolster, H., Peeters, R., and Orban, G.A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *The Journal of Neuroscience* 30, 9801-9820.
- Konen, C.S., and Kastner, S. (2008). Representation of eye movements and stimulus motion in topographically organized areas of human posterior parietal cortex. *The Journal of Neuroscience* 28, 8361-8375.
- Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244-247.
- Kourtzi, Z., Betts, L.R., Sarkheil, P., and Welchman, A.E. (2005). Distributed neural plasticity for shape learning in the human visual cortex. *PLoS biology* 3, e204.

- Kourtzi, Z., Erb, M., Grodd, W., and Bühlhoff, H.H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex* 13, 911-920.
- Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506-1509.
- Kral, K. (2003). Behavioural–analytical studies of the role of head movements in depth perception in insects, birds and mammals. *Behavioural Processes* 64, 1-12.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience* 12, 217-230.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., and Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in cognitive sciences* 17, 26-49.
- Kriegeskorte, N., Cusack, R., and Bandettini, P. (2010). How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? *Neuroimage* 49, 1965-1976.
- Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America* 103, 3863-3868.
- Krug, K., Brunskill, E., Scarna, A., Goodwin, G.M., and 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, 1839-1848.
- Krug, K., Cicmil, N., Parker, A.J., and Cumming, B.G. (2013). A Causal Role for V5/MT Neurons Coding Motion-Disparity Conjunctions in Resolving Perceptual Ambiguity. *Current Biology*.
- Krug, K., and Parker, A.J. (2011). Neurons in dorsal visual area V5/MT signal relative disparity. *The Journal of Neuroscience* 31, 17892-17904.
- Lafer-Sousa, R., and Conway, B.R. (2013). Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nature neuroscience*.
- Landy, M.S., Maloney, L.T., Johnston, E.B., and Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research* 35, 389-412.
- Larsson, J., and Heeger, D.J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience* 26, 13128-13142.
- Law, C.-T., and Gold, J.I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature neuroscience* 11, 505-513.
- Le Vay, S., Wiesel, T.N., and Hubel, D.H. (1980). The development of ocular dominance columns in normal and visually deprived monkeys. *Journal of Comparative Neurology* 191, 1-51.
- Lehmann, D., and Wälchli, P. (1975). Depth perception and location of brain lesions. *Journal of neurology* 209, 157-164.
- Leopold, D., Maier, A., and Logothetis, N.K. (2003). Measuring subjective visual perception in the nonhuman primate. *Journal of Consciousness Studies* 10, 9-10.
- Leopold, D.A., and Logothetis, N.K. (1999). Multistable phenomena: changing views in perception. *Trends in cognitive sciences* 3, 254-264.
- Leopold, D.A., Wilke, M., Maier, A., and Logothetis, N.K. (2002). Stable perception of visually ambiguous patterns. *Nat Neurosci* 5, 605-609.
- Levelt, W.J. (1967). Note on the distribution of dominance times in binocular rivalry. *Br J Psychol* 58, 143-145.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., and Malach, R. (2001). Center–periphery organization of human object areas. *Nature neuroscience* 4, 533-539.



- Lewis, J.W., and Van Essen, D.C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Comparative Neurology* 428, 112-137.
- Libedinsky, C., and Livingstone, M. (2011). Role of prefrontal cortex in conscious visual perception. *The Journal of Neuroscience* 31, 64-69.
- Liu, Y., Vogels, R., and Orban, G.A. (2004). Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. *The Journal of neuroscience* 24, 3795-3800.
- Lochmann, T., and Deneve, S. (2011). Neural processing as causal inference. *Current opinion in neurobiology* 21, 774-781.
- Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869-878.
- Logothetis, N.K., Leopold, D.A., and Sheinberg, D.L. (1996). What is rivaling during binocular rivalry? *Nature* 380, 621-624.
- Long, G.M., and Moran, C.J. (2007). How to keep a reversible figure from reversing: teasing out top-down and bottom-up processes. *PERCEPTION-LONDON-* 36, 431.
- Long, G.M., Toppino, T.C., and Mondin, G.W. (1992). Prime time: fatigue and set effects in the perception of reversible figures. *Percept Psychophys* 52, 609-616.
- Longordo, F., To, M.-S., Ikeda, K., and Stuart, G.J. (2013). Sublinear integration underlies binocular processing in primary visual cortex. *Nature neuroscience* 16, 714-723.
- Lovell, P.G., Bloj, M., and Harris, J.M. (2012). Optimal integration of shading and binocular disparity for depth perception. *Journal of vision* 12, 1.
- Luce, R.D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization* 3. Oxford University Press.
- Lumer, E.D., Friston, K.J., and Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science* 280, 1930-1934.
- Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature neuroscience* 9, 1432-1438.
- Ma, W.J., and Pouget, A. (2008). Linking neurons to behavior in multisensory perception: A computational review. *Brain research* 1242, 4-12.
- Maier, A. (2013). Neuroscience: The Cortical Layering of Visual Processing. *Current Biology* 23, R959-R961.
- Maier, A., Wilke, M., Logothetis, N.K., and Leopold, D.A. (2003). Perception of temporally interleaved ambiguous patterns. *Curr Biol* 13, 1076-1085.
- Maloney, L., Martello, M.F.D., Sahm, C., and Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proceedings of the national academy of Sciences of the united States of america* 102, 3164.
- Maloney, L.T., and Mamassian, P. (2009). Bayesian decision theory as a model of human visual perception: testing Bayesian transfer. *Visual neuroscience* 26, 147-155.
- Mamassian, P. (2008). Ambiguities and conventions in the perception of visual art. *Vision Research* 48, 2143-2153.
- Mamassian, P., and Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition* 81, B1-B9.
- Mamassian, P., Jentzsch, I., Bacon, B.A., and Schweinberger, S.R. (2003). Neural correlates of shape from shading. *NeuroReport* 14, 971-975.
- Mamassian, P., and Wallace, J.M. (2010). Sustained directional biases in motion transparency. *Journal of vision* 10.
- Manousakis, E. (2012). When perceptual time stands still: Long percept-memory in binocular rivalry. *BioSystems* 109, 115-125.

- Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information, Henry Holt and Co. Inc., New York, NY, 2-46.
- Martin, R.D., and Ross, C.F. (2005). The evolutionary and ecological context of primate vision. *The primate visual system: a comparative approach*, 1-36.
- Maunsell, J.H.R., and Treue, S. (2006). Feature-based attention in visual cortex. *Trends in neurosciences* 29, 317-322.
- McKeefry, D.J., Burton, M.P., Vakrou, C., Barrett, B.T., and Morland, A.B. (2008). Induced deficits in speed perception by transcranial magnetic stimulation of human cortical areas V5/MT+ and V3A. *The Journal of Neuroscience* 28, 6848-6857.
- Meng, M., and Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of vision* 4, 2.
- Mevorach, C., Humphreys, G.W., and Shalev, L. (2006). Opposite biases in salience-based selection for the left and right posterior parietal cortex. *Nature neuroscience* 9, 740-742.
- Miller, L.J., Mittenberg, W., Carey, V.M., Mcmorrow, M.A., Kushner, T.E., and Weinstein, J.M. (1999). Astereopsis caused by traumatic brain injury. *Archives of clinical neuropsychology* 14, 537-543.
- Miller, S.M., Gynther, B., Heslop, K., Liu, G.B., Mitchell, P., Ngo, T.T., Pettigrew, J.D., and Geffen, L. (2003). Slow binocular rivalry in bipolar disorder. *Psychological medicine* 33, 683-692.
- Miller, S.M., Hansell, N.K., Ngo, T.T., Liu, G.B., Pettigrew, J.D., Martin, N.G., and Wright, M.J. (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proceedings of the National Academy of Sciences* 107, 2664-2668.
- Milner, A.D., and Goodale, M.A. (2008). Two visual systems re-viewed. *Neuropsychologia* 46, 774-785.
- Minini, L., Parker, A.J., and Bridge, H. (2010). Neural modulation by binocular disparity greatest in human dorsal visual stream. *Journal of neurophysiology* 104, 169-178.
- Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in neurosciences* 6, 414-417.
- Mitchell, J.F., Stoner, G.R., and Reynolds, J.H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature* 429, 410-413.
- Moore, C., and Engel, S.A. (2001). Neural response to perception of volume in the lateral occipital complex. *Neuron* 29, 277-286.
- Morgan, M.L., Deangelis, G.C., and Angelaki, D.E. (2008). Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59, 662-673.
- Morrone, M.C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., and Burr, D.C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature neuroscience* 3, 1322-1328.
- Mullin, C.R., and Steeves, J.K.E. (2013). Consecutive TMS-fMRI Reveals an Inverse Relationship in BOLD Signal between Object and Scene Processing. *The Journal of Neuroscience* 33, 19243-19249.
- Münch, T.A., Da Silveira, R.A., Siegert, S., Viney, T.J., Awatramani, G.B., and Roska, B. (2009). Approach sensitivity in the retina processed by a multifunctional neural circuit. *Nature neuroscience* 12, 1308-1316.
- Murphy, A.P., Ban, H., and Welchman, A.E. (2013). Integration of texture and disparity cues to surface slant in dorsal visual cortex. *Journal of Neurophysiology* 110, 190-203.
- Murphy, A.P., Leopold, D.A., and Welchman, A.E. (2014). Perceptual memory drives learning of retinotopic biases for bistable stimuli. *Name: Frontiers in Psychology* 5, 60.

- Murphy, A.P., and Welchman, A.E. (2009). The effect of repetitive TMS of kinetic occipital and lateral occipital areas in the human visual cortex on stereopsis (Unpublished Masters project). *University of Birmingham, Birmingham UK*.
- Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., and Woods, D.L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences* 99, 15164-15169.
- Murray, S.O., and Wojciulik, E. (2003). Attention increases neural selectivity in the human lateral occipital complex. *Nature neuroscience* 7, 70-74.
- Nadler, J.W., Barbash, D., Kim, H.R., Shimpf, S., Angelaki, D.E., and Deangelis, G.C. (2013). Joint representation of depth from motion parallax and binocular disparity cues in macaque area MT. *The Journal of Neuroscience* 33, 14061-14074.
- Nagamine, M., Yoshino, A., Miyazaki, M., Takahashi, Y., and Nomura, S. (2008). Effects of selective 5-HT<sub>1A</sub> agonist tandospirone on the rate and rhythmicity of binocular rivalry. *Psychopharmacology* 198, 279-286.
- Nagamine, M., Yoshino, A., Miyazaki, M., Takahashi, Y., and Nomura, S. (2009). Difference in binocular rivalry rate between patients with bipolar I and bipolar II disorders. *Bipolar disorders* 11, 539-546.
- Nakayama, K., and Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science* 257, 1357-1363.
- Nardini, M., Bedford, R., Desai, M., and Mareschal, D. (2010a). Fusion of disparity and texture cues to slant is not mandatory in children. *Journal of Vision* 10, 494-494.
- Nardini, M., Bedford, R., and Mareschal, D. (2010b). Fusion of visual cues is not mandatory in children. *Proceedings of the National Academy of Sciences* 107, 17041-17046.
- Nawrot, M., and Blake, R. (1989). Neural integration of information specifying structure from stereopsis and motion. *Science* 244, 716-718.
- Nawrot, M., and Blake, R. (1993). On the perceptual identity of dynamic stereopsis and kinetic depth. *Vision Research* 33, 1561-1571.
- Necker, L.A. (1832). LXI. Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *Philosophical Magazine Series 3* 1, 329-337.
- Neri, P., Bridge, H., and Heeger, D.J. (2004). Stereoscopic processing of absolute and relative disparity in human visual cortex. *Journal of Neurophysiology* 92, 1880-1891.
- Nguyenkim, J.D., and Deangelis, G.C. (2003). Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *The Journal of neuroscience* 23, 7117-7128.
- Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459, 89-92.
- Noest, A.J., Van Ee, R., Nijs, M.M., and Van Wezel, R.J.A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *Journal of vision* 7, 10.
- Norman, J.F., Todd, J.T., Norman, H.F., Clayton, A.M., and McBride, T.R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision research* 46, 1057-1069.
- Nuttley, W.M., Atkinson-Leadbetter, K.P., and Van Der Kooy, D. (2002). Serotonin mediates food-odor associative learning in the nematode *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 99, 12449-12454.
- Nyffeler, T., Wurtz, P., Lüscher, H.-R., Hess, C.W., Senn, W., Pflugshaupt, T., Von Wartburg, R., Lüthi, M., and Müri, R.M. (2006). Repetitive TMS over the human

- oculomotor cortex: comparison of 1-Hz and theta burst stimulation. *Neuroscience letters* 409, 57-60.
- O'Connor, D.H., Fukui, M.M., Pinsk, M.A., and Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature neuroscience* 5, 1203-1209.
- O'shea, R.P., and Corballis, P.M. (2003). Binocular rivalry in split-brain observers. *Journal of Vision* 3, 3.
- Orbach, J., Ehrlich, D., and Heath, H.A. (1963). Reversibility of the Necker Cube. I. An Examination of the Concept of "Satiation of Orientation". *Percept Mot Skills* 17, 439-458.
- Orban, G.A. (2011). The extraction of 3D shape in the visual system of human and nonhuman primates. *Annual Review of Neuroscience* 34, 361-388.
- Orban, G.A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J.T., Wardak, C., Durand, J.-B., and Vanduffel, W. (2006a). Mapping the parietal cortex of human and non-human primates. *Neuropsychologia* 44, 2647-2667.
- Orban, G.A., Janssen, P., and Vogels, R. (2006b). Extracting 3D structure from disparity. *Trends in neurosciences* 29, 466-473.
- Orban, G.A., Sunaert, S., Todd, J.T., Van Hecke, P., and Marchal, G. (1999). Human cortical regions involved in extracting depth from motion. *Neuron* 24, 929-940.
- Orban, G.A., Zhu, Q., and Vanduffel, W. (2014). The transition in the ventral stream from feature to real-world entity representations. *Frontiers in psychology* 5.
- Paffen, C.L.E., Alais, D., and Verstraten, F.a.J. (2006). Attention speeds binocular rivalry. *Psychological Science* 17, 752-756.
- Panagiotaropoulos, T.I., Deco, G., Kapoor, V., and Logothetis, N.K. (2012). Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron* 74, 924-935.
- Parker, A.J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience* 8, 379-391.
- Pasalar, S., Ro, T., and Beauchamp, M.S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience* 31, 1783-1790.
- Pastukhov, A., and Braun, J. (2008). A short-term memory of multi-stable perception. *J Vis* 8, 7, 1-14.
- Pastukhov, A., and Braun, J. (2011). Cumulative history quantifies the role of neural adaptation in multistable perception. *Journal of vision* 11, 12.
- Pastukhov, A., and Braun, J. (2013). Structure-from-motion: dissociating perception, neural persistence, and sensory memory of illusory depth and illusory rotation. *Attention, Perception, & Psychophysics* 75, 322-340.
- Pastukhov, A., Füllekrug, J., and Braun, J. (2013). Sensory memory of structure-from-motion is shape-specific. *Attention, Perception, & Psychophysics* 75, 1215-1229.
- Pasupathy, A., and Connor, C.E. (2002). Population coding of shape in area V4. *Nature neuroscience* 5, 1332-1338.
- Pavlov, I.P. (1927). *Conditioned reflexes*. New York: Dover.
- Pearson, J., and Brascamp, J.W. (2008). Sensory memory for ambiguous vision. *Trends Cogn Sci* 12, 334-341.
- Pearson, J., and Clifford, C.G. (2004). Determinants of visual awareness following interruptions during rivalry. *J Vis* 4, 196-202.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437-442.

- Pesaran, B., Nelson, M.J., and Andersen, R.A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453, 406-409.
- Petersen, C.C.H. (2007). The functional organization of the barrel cortex. *Neuron* 56, 339-355.
- Pettigrew, J.D., and Miller, S.M. (1998). A 'sticky' interhemispheric switch in bipolar disorder? *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265, 2141-2148.
- Pitzalis, S., Sereno, M.I., Committeri, G., Fattori, P., Galati, G., Patria, F., and Galletti, C. (2010). Human V6: the medial motion area. *Cerebral Cortex* 20, 411-424.
- Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233-238.
- Popple, A.V., Smallman, H.S., and Findlay, J.M. (1998). The area of spatial integration for initial horizontal disparity vergence. *Vision Research* 38, 319-326.
- Pouget, A., Deneve, S., and Duhamel, J.-R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience* 3, 741-747.
- Press, W.A., Brewer, A.A., Dougherty, R.F., Wade, A.R., and Wandell, B.A. (2001). Visual areas and spatial summation in human visual cortex. *Vision Research* 41, 1321-1332.
- Pressnitzer, D., and Hupé, J.-M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology* 16, 1351-1357.
- Preston, T.J., Kourtzi, Z., and Welchman, A.E. (2009). Adaptive estimation of three-dimensional structure in the human brain. *J Neurosci* 29, 1688-1698.
- Preston, T.J., Li, S., Kourtzi, Z., and Welchman, A.E. (2008). Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *The Journal of Neuroscience* 28, 11315-11327.
- Ptito, A., Zatorre, R.J., Larson, W.L., and Tosoni, C. (1991). Stereopsis after unilateral anterior temporal lobectomy dissociation between local and global measures. *Brain* 114, 1323-1333.
- Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102-1107.
- Ramachandran, V.S. (1988). Perceiving shape from shading. *Scientific American* 259, 76-83.
- Rao, R.P.N. (2005). Bayesian inference and attentional modulation in the visual cortex. *Neuroreport* 16, 1843-1848.
- Rao, R.P.N., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience* 2, 79-87.
- Read, J.C., Phillipson, G.P., Serrano-Pedraza, I., Milner, A.D., and Parker, A.J. (2010). Stereoscopic vision in the absence of the lateral occipital cortex. *PloS one* 5, e12608.
- Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience* 3, 261-270.
- Regan, D., Giaschi, D., Sharpe, J.A., and Hong, X.H. (1992). Visual processing of motion-defined form: selective failure in patients with parietotemporal lesions. *The Journal of neuroscience* 12, 2198-2210.
- Rescorla, R.A. (2003). Contemporary study of Pavlovian conditioning. *The Spanish journal of psychology* 6, 185-195.

- Reynolds, J.H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron* 26, 703-714.
- Rice, N.J., Edwards, M.G., Schindler, I., Punt, T.D., McIntosh, R.D., Humphreys, G.W., Lestou, V., and Milner, A.D. (2008). Delay abolishes the obstacle avoidance deficit in unilateral optic ataxia. *Neuropsychologia* 46, 1549-1557.
- Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research* 10, 380-388.
- Riddoch, M.J., and Humphreys, G.W. (1987). A case of integrative visual agnosia. *Brain* 110, 1431-1462.
- Riddoch, M.J., Humphreys, G.W., Edwards, S., Baker, T., and Willson, K. (2002). Seeing the action: neuropsychological evidence for action-based effects on object selection. *Nature neuroscience* 6, 82-89.
- Rizzo, M., Nawrot, M., and Zihl, J. (1995). Motion and shape perception in cerebral akinetopsia. *Brain* 118, 1105-1127.
- Rockland, K.S., and Van Hoesen, G.W. (1999). Some temporal and parietal cortical connections converge in CA1 of the primate hippocampus. *Cerebral Cortex* 9, 232-237.
- Roe, A.W., Pallas, S.L., Hahm, J.-O., and Sur, M. (1990). A map of visual space induced in primary auditory cortex. *Science* 250, 818-820.
- Rogers, B., and Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception* 8, 125-134.
- Rogers, B., and Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research* 22, 261-270.
- Rolls, E.T. (2000). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron* 27, 205-218.
- Rosas, P., Wichmann, F.A., and Wagemans, J. (2004). Some observations on the effects of slant and texture type on slant-from-texture. *Vision Research* 44, 1511-1535.
- Rosenholtz, R., and Malik, J. (1997). Surface orientation from texture: isotropy or homogeneity (or both)? *Vision research* 37, 2283-2293.
- Ross, J.E. (1983). Disturbance of stereoscopic vision in patients with unilateral stroke. *Behavioural Brain Research* 7, 99-112.
- Rossetti, Y., Revol, P., McIntosh, R., Pisella, L., Rode, G., Danckert, J., Tilikete, C., Dijkerman, H.C., Boisson, D., and Vighetto, A. (2005). Visually guided reaching: bilateral posterior parietal lesions cause a switch from fast visuomotor to slow cognitive control. *Neuropsychologia* 43, 162-177.
- Rossi, S., Hallett, M., Rossini, P.M., and Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology* 120, 2008-2039.
- Rothstein, T.B., and Sacks, J.G. (1972). Defective stereopsis in lesions of the parietal lobe. *American journal of ophthalmology* 73, 281.
- Ruff, C.C., Driver, J., and Bestmann, S. (2009). Combining TMS and fMRI: from 'virtual lesions' to functional-network accounts of cognition. *Cortex* 45, 1043-1049.
- Ruohonen, J., and Ilmoniemi, R.J. (2002). Physical principles for transcranial magnetic stimulation. *Handbook of transcranial magnetic stimulation*, 18-30.
- Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337, 753-756.

- Sack, A.T., Kadosh, R.C., Schuhmann, T., Moerel, M., Walsh, V., and Goebel, R. (2009). Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. *Journal of Cognitive Neuroscience* 21, 207-221.
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., and Tsutsui, K. (1998). Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 353, 1363-1373.
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tsutsui, K., Tanaka, Y., Shein, W.N., and Miyashita, Y. (1999). Neural representation of three-dimensional features of manipulation objects with stereopsis. *Experimental brain research* 128, 160-169.
- Sanada, T.M., Nguyenkim, J.D., and Deangelis, G.C. (2012). Representation of 3D surface orientation by velocity and disparity gradient cues in area MT. *Journal of neurophysiology*.
- Sandberg, K., Bahrami, B., Lindelov, J.K., Overgaard, M., and Rees, G. (2011). The impact of stimulus complexity and frequency swapping on stabilization of binocular rivalry. *J Vis* 11.
- Sayres, R., and Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology* 100, 249-267.
- Schenk, T., and McIntosh, R.D. (2010). Do we have independent visual streams for perception and action? *Cognitive Neuroscience* 1, 52-62.
- Schluppeck, D., Glimcher, P., and Heeger, D.J. (2005). Topographic organization for delayed saccades in human posterior parietal cortex. *Journal of Neurophysiology* 94, 1372-1384.
- Schmack, K., De Castro, A.G.-C., Rothkirch, M., Sekutowicz, M., Rössler, H., Haynes, J.-D., Heinz, A., Petrovic, P., and Sterzer, P. (2013). Delusions and the Role of Beliefs in Perceptual Inference. *The Journal of Neuroscience* 33, 13701-13712.
- Schrater, P.R., and Sundaeswara, R. (Year). "Theory and dynamics of perceptual bistability", in: *Advances in neural information processing systems*, 1217-1224.
- Schwarzkopf, D.S., Schindler, A., and Rees, G. (2010). Knowing with Which Eye We See: Utrocular Discrimination and Eye-Specific Signals in Human Visual Cortex. *PLoS ONE* 5, e13775.
- Schwarzlose, R.F., Swisher, J.D., Dang, S., and Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences* 105, 4447-4452.
- Scott, S.H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience* 5, 532-546.
- Seitz, A., and Watanabe, T. (2005). A unified model for perceptual learning. *Trends in cognitive sciences* 9, 329-334.
- Serences, J.T., and Boynton, G.M. (2007). The representation of behavioral choice for motion in human visual cortex. *The Journal of Neuroscience* 27, 12893-12899.
- Sereno, M.E., Trinath, T., Augath, M., and Logothetis, N.K. (2002). Three-dimensional shape representation in monkey cortex. *Neuron* 33, 635-652.
- Sereno, M.I., and Huang, R.-S. (2014). Multisensory maps in parietal cortex. *Current opinion in neurobiology* 24, 39-46.
- Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology* 86, 1916.
- Sharma, J., Angelucci, A., and Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature* 404, 841-847.

- Shatz, C.J., and Stryker, M.P. (1978). Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *The Journal of Physiology* 281, 267-283.
- Sheppard, B.M., and Pettigrew, J.D. (2006). Plaid motion rivalry: Correlates with binocular rivalry and positive mood state. *PERCEPTION-LONDON* 35, 157.
- Shikata, E., Hamzei, F., Glauche, V., Knab, R., Dettmers, C., Weiller, C., and Büchel, C. (2001). Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an event-related fMRI study. *Journal of Neurophysiology* 85, 1309-1314.
- Shikata, E., Mcnamara, A., Sprenger, A., Hamzei, F., Glauche, V., Büchel, C., and Binkofski, F. (2008). Localization of human intraparietal areas AIP, CIP, and LIP using surface orientation and saccadic eye movement tasks. *Human brain mapping* 29, 411-421.
- Shikata, E., Tanaka, Y., Nakamura, H., Taira, M., and Sakata, H. (1996). Selectivity of the parietal visual neurones in 3D orientation of surface of stereoscopic stimuli. *Neuroreport* 7, 2389-2394.
- Silson, E.H., McKeefry, D.J., Rodgers, J., Gouws, A.D., Hymers, M., and Morland, A.B. (2013). Specialized and independent processing of orientation and shape in visual field maps LO1 and LO2. *Nature neuroscience* 16, 267-269.
- Silver, M.A., and Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends in cognitive sciences* 13, 488-495.
- Silver, M.A., Ress, D., and Heeger, D.J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *Journal of neurophysiology* 94, 1358-1371.
- Simoncelli, E.P. (2003). Vision and the statistics of the visual environment. *Current opinion in neurobiology* 13, 144-149.
- Simoncelli, E.P., and Olshausen, B.A. (2001). Natural image statistics and neural representation. *Annual review of neuroscience* 24, 1193-1216.
- Snow, J.C., Miranda, R.R., and Humphreys, G.W. (2013). Impaired visual sensitivity within the ipsilesional hemifield following parietal lobe damage. *Cortex* 49, 158-171.
- Somers, D.C., Dale, A.M., Seiffert, A.E., and Tootell, R.B.H. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences* 96, 1663-1668.
- Srivastava, S., Orban, G.A., De Mazière, P.A., and Janssen, P. (2009). A distinct representation of three-dimensional shape in macaque anterior intraparietal area: fast, metric, and coarse. *The Journal of Neuroscience* 29, 10613-10626.
- Stanley, J., Forte, J.D., Cavanagh, P., and Carter, O. (2011). Onset rivalry: the initial dominance phase is independent of ongoing perceptual alternations. *Frontiers in Human Neuroscience* 5.
- Sterzer, P., Frith, C., and Petrovic, P. (2008). Believing is seeing: expectations alter visual awareness. *Current Biology* 18, R697-R698.
- Sterzer, P., and Kleinschmidt, A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences* 104, 323-328.
- Sterzer, P., Kleinschmidt, A., and Rees, G. (2009). The neural bases of multistable perception. *Trends Cogn Sci* 13, 310-318.
- Sterzer, P., and Rees, G. (2008). A neural basis for percept stabilization in binocular rivalry. *J Cogn Neurosci* 20, 389-399.
- Stevens, K.A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological cybernetics* 46, 183-195.



- Stewart, L., Walsh, V., and Rothwell, J. (2001). Motor and phosphene thresholds: a transcranial magnetic stimulation correlation study. *Neuropsychologia* 39, 415-419.
- Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature* 437, 1272-1278.
- Stickgold, R., James, L.T., and Hobson, J.A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience* 3, 1237-1238.
- Stokes, M.G., Chambers, C.D., Gould, I.C., Henderson, T.R., Janko, N.E., Allen, N.B., and Mattingley, J.B. (2005). Simple metric for scaling motor threshold based on scalp-cortex distance: application to studies using transcranial magnetic stimulation. *Journal of neurophysiology* 94, 4520-4527.
- Stryker, M.P., Sherk, H., Leventhal, A.G., and Hirsch, H.V. (1978). Physiological consequences for the cat's visual cortex of effectively restricting early visual experience with oriented contours. *J Neurophysiol* 41, 896-909.
- Sugihara, H., Murakami, I., Shenoy, K.V., Andersen, R.A., and Komatsu, H. (2002). Response of MSTd neurons to simulated 3D orientation of rotating planes. *Journal of neurophysiology* 87, 273-285.
- Sunaert, S., Van Hecke, P., Marchal, G., and Orban, G.A. (1999). Motion-responsive regions of the human brain. *Experimental Brain Research* 127, 355-370.
- Sundareswara, R., and Schrater, P.R. (2008). Perceptual multistability predicted by search model for Bayesian decisions. *Journal of Vision* 8.
- Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A., and Somers, D.C. (2007). Visual topography of human intraparietal sulcus. *The Journal of neuroscience* 27, 5326-5337.
- Taira, M., Nose, I., Inoue, K., and Tsutsui, K. (2001). Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. *Neuroimage* 14, 959-966.
- Taira, M., Tsutsui, K.-I., Jiang, M., Yara, K., and Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *Journal of neurophysiology* 83, 3140-3146.
- Takayama, Y., and Sugishita, M. (1994). Astereopsis induced by repetitive magnetic stimulation of occipital cortex. *Journal of neurology* 241, 522-525.
- Takayama, Y., Sugishita, M., Kido, T., Ogawa, M., Fukuyama, H., and Akiguchi, I. (1994). Impaired stereoacuity due to a lesion in the left pulvinar. *Journal of Neurology, Neurosurgery & Psychiatry* 57, 652-654.
- Tanaka, H., Uka, T., Yoshiyama, K., Kato, M., and Fujita, I. (2001). Processing of shape defined by disparity in monkey inferior temporal cortex. *Journal of Neurophysiology* 85, 735-744.
- Theys, T., Pani, P., Van Loon, J., Goffin, J., and Janssen, P. (2012). Selectivity for Three-Dimensional Shape and Grasping-Related Activity in the Macaque Ventral Premotor Cortex. *The Journal of neuroscience* 32, 12038-12050.
- Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience* 5, 1226-1235.
- Tootell, R.B.H., and Hadjikhani, N. (2001). Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cerebral Cortex* 11, 298-311.
- Tootell, R.B.H., Mendola, J.D., Hadjikhani, N.K., Liu, A.K., and Dale, A.M. (1998). The representation of the ipsilateral visual field in human cerebral cortex. *Proceedings of the National Academy of Sciences* 95, 818-824.
- Treue, S., and Trujillo, J.C.M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399, 575-579.
- Trommershäuser, J., Körding, K.P., and Landy, M.S. (2011). *Sensory cue integration*. Oxford University Press.

- Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., and Livingstone, M.S. (2006). A cortical region consisting entirely of face-selective cells. *Science* 311, 670-674.
- Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences* 105, 19514-19519.
- Tsao, D.Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T.A., Mandeville, J.B., Wald, L.L., Dale, A.M., Rosen, B.R., and Van Essen, D.C. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39, 555-568.
- Tsutsui, K.-I., Sakata, H., Naganuma, T., and Taira, M. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298, 409-412.
- Tsutsui, K.I., Jiang, M., Yara, K., Sakata, H., and Taira, M. (2001). Integration of Perspective and Disparity Cues in Surface-Orientation-Selective Neurons of Area CIP. *Journal of neurophysiology* 86, 2856.
- Tyler, C.W., Likova, L.T., Chen, C.C., Kontsevich, L.L., Schira, M.M., and Wade, A.R. (2005). Extended concepts of occipital retinotopy. *Current Medical Imaging Reviews* 1, 319-329.
- Tyler, C.W., Likova, L.T., Kontsevich, L.L., and Wade, A.R. (2006). The specificity of cortical region KO to depth structure. *Neuroimage* 30, 228-238.
- Uka, T., and Deangelis, G.C. (2004). Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42, 297-310.
- Uka, T., and Deangelis, G.C. (2006). Linking neural representation to function in stereoscopic depth perception: roles of the middle temporal area in coarse versus fine disparity discrimination. *The Journal of neuroscience* 26, 6791-6802.
- Uka, T., Tanabe, S., Watanabe, M., and Fujita, I. (2005). Neural correlates of fine depth discrimination in monkey inferior temporal cortex. *The Journal of neuroscience* 25, 10796-10802.
- Ukai, K., Ando, H., and Kuze, J. (2003). Binocular rivalry alternation rate declines with age. *Perceptual and motor skills* 97, 393-397.
- Ungerleider, L.G., and Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology* 248, 190-222.
- V. Schiller, P. (1933). Stroboskopische alternativversuche. *Psychological Research* 17, 179-214.
- Vaina, L. (1989). Selective impairment of visual motion interpretation following lesions of the right occipito-parietal area in humans. *Biological cybernetics* 61, 347-359.
- Vaina, L.M., Lemay, M., Bienfang, D.C., Choi, A.Y., and Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual neuroscience* 5, 353-369.
- Van Boxtel, J.J.A., Alais, D., and Van Ee, R. (2008). Retinotopic and non-retinotopic stimulus encoding in binocular rivalry and the involvement of feedback. *Journal of Vision* 8, 17.
- Van Boxtel, J.J.A., Tsuchiya, N., and Koch, C. (2010). Consciousness and attention: on sufficiency and necessity. *Consciousness Research* 1, 217.
- Van Dam, L.C., and Ernst, M.O. (2010). Preexposure disrupts learning of location-contingent perceptual biases for ambiguous stimuli. *J Vis* 10, 15.
- Van Ee, R., and Erkelens, C.J. (1998). Temporal aspects of stereoscopic slant estimation: An evaluation and extension of Howard and Kaneko's theory. *Vision research* 38, 3871-3882.

- Van Ee, R., Van Dam, L.C.J., and Erkelens, C.J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision* 2.
- Van Loon, A.M., Knapen, T., Scholte, H.S., St John-Saaltink, E., Donner, T.H., and Lamme, V.A. (2013). GABA Shapes the Dynamics of Bistable Perception. *Current biology*.
- Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., and Orban, G. (1997). The kinetic occipital (KO) region in man: an fMRI study. *Cerebral Cortex* 7, 690-701.
- Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J., and Orban, G. (2002). Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science* 298, 413-415.
- Verhoef, B.-E., Vogels, R., and Janssen, P. (2011). Synchronization between the end stages of the dorsal and the ventral visual stream. *Journal of neurophysiology* 105, 2030-2042.
- Vierck, E., Porter, R.J., Luty, S.E., Moor, S., Crowe, M.T., Carter, J.D., Inder, M.L., and Joyce, P.R. (2013). Further evidence for slow binocular rivalry rate as a trait marker for bipolar disorder. *Australian and New Zealand Journal of Psychiatry* 47, 371-379.
- Vilares, I., Howard, J.D., Fernandes, H.L., Gottfried, J.A., and Kording, K.P. (2012). Differential representations of prior and likelihood uncertainty in the human brain. *Current Biology* 22, 1641-1648.
- Vishwanath, D., Girshick, A.R., and Banks, M.S. (2005). Why pictures look right when viewed from the wrong place. *Nature neuroscience* 8, 1401-1410.
- Wallach, H., and O'connell, D. (1953). The kinetic depth effect. *Journal of experimental psychology* 45, 205.
- Wang, P., and Nikolić, D. (2011). An LCD monitor with sufficiently precise timing for research in vision. *Frontiers in human neuroscience* 5.
- Ward, A., Liu, J., Feng, Z., and Xu, X.Z.S. (2008). Light-sensitive neurons and channels mediate phototaxis in *C. elegans*. *Nature neuroscience* 11, 916-922.
- Watson, A.B., and Pelli, D.G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics* 33, 113-120.
- Watson, J.D.G., Myers, R., Frackowiak, R.S.J., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S., and Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex* 3, 79-94.
- Welchman, A.E. (2011). "Decoding the Cortical Representation of Depth," in *Sensory Cue Integration*, eds. J. Trommershauser, K. Kording & M.S. Landy. . Oxford University Press), 345.
- Welchman, A.E., Deubelius, A., Conrad, V., Bulthoff, H.H., and Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature neuroscience* 8, 820-827.
- Wexler, M., and Van Boxtel, J.J.A. (2005). Depth perception by the active observer. *Trends in cognitive sciences* 9, 431-438.
- White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 314, 1-340.
- Wichmann, F.A., and Hill, N.J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & psychophysics* 63, 1293-1313.
- Wichmann, F.A., and Hill, N.J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & psychophysics* 63, 1314-1329.

- Wilcke, J.C., O'shea, R.P., and Watts, R. (2009). Frontoparietal activity and its structural connectivity in binocular rivalry. *Brain research* 1305, 96-107.
- Williams, Z.M., Elfar, J.C., Eskandar, E.N., Toth, L.J., and Assad, J.A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. *Nature neuroscience* 6, 616-623.
- Wilms, M., Eickhoff, S.B., Hömke, L., Rottschy, C., Kujovic, M., Amunts, K., and Fink, G.R. (2010). Comparison of functional and cytoarchitectonic maps of human visual areas V1, V2, V3d, V3v, and V4 (v). *Neuroimage* 49, 1171-1179.
- Windmann, S., Wehrmann, M., Calabrese, P., and Güntürkün, O. (2006). Role of the prefrontal cortex in attentional control over bistable vision. *Journal of cognitive neuroscience* 18, 456-471.
- Wokke, M.E., Scholte, H.S., and Lamme, V.a.F. (2014). Opposing dorsal/ventral stream dynamics during figure-ground segregation. *Journal of cognitive neuroscience* 26, 365-379.
- Wright, L.A., and Wormald, R.P.L. (1992). Stereopsis and ageing. *Eye* 6, 473-476.
- Xiao, D.K., Marcar, V., Raiguel, S., and Orban, G. (1997). Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *European Journal of Neuroscience* 9, 956-964.
- Yang, T., and Shadlen, M.N. (2007). Probabilistic reasoning by neurons. *Nature* 447, 1075-1080.
- Yang, Y., Liu, S., Chowdhury, S.A., Deangelis, G.C., and Angelaki, D.E. (2011). Binocular disparity tuning and visual-vestibular congruency of multisensory neurons in macaque parietal cortex. *The Journal of Neuroscience* 31, 17905-17916.
- Yuille, A., and Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in cognitive sciences* 10, 301-308.
- Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., and Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31, 1116-1128.
- Zafar, N., Paulus, W., and Sommer, M. (2008). Comparative assessment of best conventional with best theta burst repetitive transcranial magnetic stimulation protocols on human motor cortex excitability. *Clinical Neurophysiology* 119, 1393-1399.
- Zaretskaya, N., Thielscher, A., Logothetis, N.K., and Bartels, A. (2010). Disrupting parietal function prolongs dominance durations in binocular rivalry. *Current biology* 20, 2106-2111.
- Zaroff, C.M., Knutelska, M., and Frumkes, T.E. (2003). Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Investigative ophthalmology & visual science* 44, 891-900.
- Zeki, S., Perry, R., and Bartels, A. (2003). The processing of kinetic contours in the brain. *Cerebral Cortex* 13, 189-202.
- Zhang, P., Jamison, K., Engel, S., He, B., and He, S. (2011). Binocular rivalry requires visual attention. *Neuron* 71, 362-369.
- Zhong, Y.-M., and Rockland, K.S. (2003). Inferior parietal lobule projections to anterior inferotemporal cortex (area TE) in macaque monkey. *Cerebral Cortex* 13, 527-540.
- Zhou, Y., Gao, J., White, K., Merk, I., and Yao, K. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological cybernetics* 90, 256-263.