Brain mechanisms determining the dynamics of bistable perception

Megumi Fukuda

Institute of Cognitive Neuroscience, Division of Psychology & Language Sciences

Primary Supervisor: Prof. Geraint Rees
Secondary supervisor: Prof. Nikolaus Weiskopf

Dissertation submitted for the degree of Doctor of Philosophy, University College London

DECLARATION OF ORIGINALITY

I, Megumi Fukuda, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

December 2015

Megumi Fukuda

ABSTRACT

The neural correlate of visual perception has been one of the central issues in neuroscience. Bistable perception, two distinct percepts spontaneously alternate in every few seconds while the physical input remains the same, has been used to investigate how our visual system resolves ambiguity in visual information. In this thesis, I interrogate the brain mechanisms of visual bistability in a series of experiments.

In chapter 3, I recorded brain activity while participants engaged in a LGN localizer task. Activation in LGN could not be identified successfully, suggesting that we may need to establish a better approach to localize LGN.

The experiment in chapter 4 answers how multiple brain regions, two parietal areas and the visual cortex, interact with each other in perceptual switches. Previous TMS studies suggested that the right anterior superior parietal lobule (r-aSPL) and the right posterior superior parietal lobule (r-pSPL) have opposite roles in triggering perceptual reversals. Using dynamic causal modeling (DCM), I found that resolving such perceptual ambiguity was specifically associated with reciprocal interactions between these parietal regions and V5/MT, and the strength of bottom-up coupling between V5/MT to r-pSPL and from r-pSPL to r-aSPL predicted individual mean dominance duration.

The third study addressed which functional networks and brain regions would contribute to successful prediction of individual switch frequencies. Applying graph theoretical analysis to resting state data, I found centrality measures, which are used as proxy of hubness of the region in the entire network, predict individual switch frequencies and attack (removal of edges) to fronto-parietal network and visual network decreased prediction accuracy.

Finally, in chapter 6 I investigated what determines trial-by-trial dynamics of bistable perception. I have developed a new experimental paradigm to test

how an observer forms expectation from statistical information of the stimulus sequence. The subjective percept of ambiguous bistable perception was strongly biased towards expected stimuli, and such expectation quantified by a hierarchical Bayesian model was represented in multiple brain regions, including the fronto-parietal areas as well as the visual cortex.

The series of experiments showed that the fronto-parietal network and visual network involve in forming conscious visual percept. My results favor predictive coding theory in bistable perception, which explains formation of consciousness as continuous input from the environment and update of the internal belief.

ACKNOWLEDGEMENTS

First of all, I would like to express my eternal gratitude to my supervisor Prof. Geraint Rees. Geraint has been very supportive and enthusiastic, and I cannot thank enough for his support.

I am very grateful to Geraint lab members – Adeel, Bahador, Becky, Ben, Chen, Elaine, Javier, Jin, Joel, Klaartje, Kristian, Ryota, Maren, Micah, Sam, Sarah, Spas, and Takamitsu (and David – he left the group before I join, but he taught me how to run bistable perception experiments through his online video). I really liked our lab members as we had nice diversity in the research field as well as personal backgrounds, and everyone had special talent. Bahador, Joel, and Ryota were my office mates in G09 at ICN. I will miss dim lighting in G09, which has been boosting our productivity (Steidlea and Werth, 2013). Kristian, Spas, and Takamitsu helped me a lot as collaborators and I always appreciate their help and kindness.

I was very lucky to meet great people at ICN and FIL. Nik is my secondary supervisor and helped me to run fMRI experiments. Gareth is the very best MEG evangelist ever and I really enjoyed warm and welcoming atmosphere in MEG meeting. Christoph helped me for the Bayesian model analysis for MEG/behavioural data. I would like to thank the FIL methods group members for the advice with data analysis in Monday method seminar and SPM support.

Also I would like to acknowledge other staff at FIL and ICN – imaging support team members, IT support team members, Peter, David, Tom, Alison, Marcia, and above all, Rosalyn.

I was supported by Japan Student Service Organization (JASSO) during my PhD and I am very grateful to their support.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PUBLICATIONS DURING MY PHD

Papers were published under 'Fukuda Megumi' (not 'Megumi Fukuda' – there are too many 'Fukuda M' in the world).

Publication

Megumi F, Bahrami B, Kanai R, and Rees G (2015). Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. NeuroImage 107, 190-197.

- Experiment in Chapter 4

Other manuscripts

Megumi F, Kanai R, Getov S, Urner M, Watanabe T, and Rees G. (under revision). Resting-state network architecture predicts individual switch frequencies in bistable perception.

- Experiment in Chapter 5

Megumi F, Barnes GR, Bahrami B, Mathys C, Sandberg K, Kanai R, and Rees G. (in prep). Prior expectation shapes the dynamics of visual perception.

- Experiment in Chapter 6

Other Publications published during PhD

Watanabe T, Masuda N, Megumi F, Kanai R, Rees G (2014) Energy landscape and dynamics of brain activity during human bistable perception. Nature communications, 5:4765.

Chapter 1 GENERAL INTRODUCTION

1.1 Preface

The discovery of the columnar structure in the cat sensory cortex (Mountcastle, 1957) and the primate visual cortex (Hubel & Wiesel, 1968) suggests that sensory information is mapped in the sensory cortices as a regular pattern. Such vertical column structure is thought to be a minimum unit of neural representation of sensory stimuli.

We can even describe our subjective experience, consciousness, as patterns of neural activity. Psychologists, engineers, and more recently, neuroscientists joined forces and started to explore the "Neural correlates of consciousness (NCC)" (Crick & Koch, 2003), meaning the temporal or spatial patterns of neural activity that are sufficient to cause a specific conscious percept. If we wish to seek the NCC, we need to find out how neurons represent conscious percepts as spatial/temporal activity patterns.

Specifically, the visual system has been used to explore NCC. This is partly because the visual system is relatively simple and well understood. To investigate the visual system, what we need to do is to associate visual inputs (visual stimuli) with the perceptual experience of the participants (based on their reports). In addition, the anatomy of the visual system has been investigated extensively, from retina to higher brain regions, in humans as well as other animals (e.g. (Felleman & Van Essen, 1991)).

The development of non-invasive brain imaging techniques has enabled us to measure human brain activity *in vivo*. One of the most effective techniques for investigating human brain functions is functional magnetic resonance imaging (fMRI). fMRI measures changes in the ratio of deoxy- to oxy-haemoglobin associated with blood flow changes caused by neuronal activity (the so-called Blood Oxygen Level Dependent (BOLD) signals) (Logothetis *et al.*, 2001) and its good spatial resolution allows us to examine the functional roles of specific brain regions.

1.2 Neural correlates of visual bistable perception

1.2.1 Measuring NCC with bistable perception stimuli

One approach to investigate NCC is to use bistable stimuli. Bistable stimuli, as the name indicates, induce two distinct (and alternating) conscious percepts while the physical input to retina remains the same.

The advantage of using bistable stimuli is we can observe fluctuations of conscious perception without changing external inputs, and therefore any changes in brain signal are a reflection of changes in conscious percept[1]. To understand the biological and neural mechanism of perceptual switches, 'replay' condition, which is often used to control for the effects of button press or stimulus presentation (rather than subjective percept) is often included in the experiment and compared with the pure rivalry condition.

There is no guarantee that subjective report is the 'true' percept, not imagery or guess, as there is no solid way to define the 'correct' response during bistable perception view. However, sophisticated experimental paradigms and analysis methods (e.g. using fMRI multivariate analysis, optokinetic nystagmus data; see the later section) enable us to access contents of the (expected) reports.

1.2.2 Bistable perception tasks

There are many types of bistable stimuli, and they have been used to investigate our brain function and the basis of visual awareness.

Here I briefly introduce different types of bistable stimuli. 'Bistable' means there are two possible states (i.e. perception, or interpretation of the stimuli), but some stimuli may elicit more than two distinct percepts. For instance, during binocular rivalry view, there might be three different types

[1] Although recent studies suggest that making report of one's percept alters brain activity pattern. Details will be discussed in Chapter 1.2.3.

of reports; perceiving the image shown to the right eye, perceiving the image show on the left eye, and the mixture of the two images. These stimuli can be referred as 'multistable', but here for simplicity I just call all stimuli as 'bistable, as most of the stimuli are bistable, and the stimulus (spherical-shape structure-from-motion) used in the experiments is indeed bistable.

Typical behavioral measures to describe the dynamics of bistable perception are mean dominance duration and switch frequency. Mean dominance duration is how long each percept remains as dominant percept. Switch rate quantifies how often the subjective percept changes between the exclusive percepts. Basically, they both characterize stabilization of the subjective percept, and they can be converted each other quite easily (i.e. take the inverse).

1.2.2.1 Binocular rivalry

Normally, we see the world with two eyes and the inputs to the retina are quite similar between the two eyes. However, when an observer sees different two images with the eyes – one image exposed to the right eye, the other image to the left eye - they initially often perceive one of the monocular images, rather than mixture of the two images. Subsequently the two monocular images alternate in perception. This is called 'binocular rivalry', as the two images compete and suppress each other to enter perception, and as a consequence, only one of the two is perceived. Such binocular rivalry typically requires dichoptic stimulus presentation with anaglyph glasses, prism glasses, or mirror stereoscope (Carmel *et al.*, 2010a).

Continuous flash suppression (CFS) paradigm (Tsuchiya & Koch, 2005) is similar to binocular rivalry, but with a monocular flashing image presented to one eye. Typically, the monocular flashing images are dominant and subjective percepts do not alternate frequently; when one eye is exposed to flushing images such as Mondrian patch, the other image presented to the other eye is suppressed and typically takes a long time to

emerge into conscious percept. This paradigm is often used to explore unconscious information processing or effect of suppression (e.g. (Seitz *et al.*, 2009)). Although the CFS paradigm is similar to binocular rivalry, the biological mechanism of CFS may not be binocular rivalry but flash suppression (Tsuchiya *et al.*, 2006). Hence, the results and findings of CFS experiments may not be generalized to binocular rivalry or bistable perception.

1.2.2.2 Structure-from-motion

Ambiguous structure-from-motion (SFM) describes a bistable perception paradigm that utilizes the kinetic depth effect (KDE). KDE was first reported in (Wallach & O'Connell, 1953); when a rotating 3D object is projected on a 2D screen, its 2D shadow is perceived as a rotating object with depth, especially if the stimulus is smoothly curved. What is even more interesting is the direction of rotation remains ambiguous, as the depth cue is lost in the silhouette and there is no 'correct' interpretation any more.

Many objects have been used for SFM bistable perception stimuli; a cylinder (Ullman, 1979; Treue *et al.*, 1991), the Necker cube (Dosher *et al.*, 1986), a spinning dancer (Troje & McAdam, 2010), and a sphere (Jiang *et al.*, 1998; Maier *et al.*, 2003). (Ullman, 1979) pointed out that objects viewed in orthographic projection elicit spontaneous perceptual switches.

SFM stimuli can be presented to the participants on a screen without need for glasses or stereoscope for stimulus presentation.

1.2.2.3 Apparent motion

Apparent motion is a motion-induced illusion; in this paradigm, observers perceive continuous movement from a series of static images when they are presented in rapid succession. For example, von Schiller (1933) used bistable apparent motion paradigm ("apparent motion quartet"), where two pairs of dots induce parallel or vertical movements of the pairs of dots.

This stimulus is easy to implement and there is no need to prepare

prism glasses or stereoscope to present the stimuli; however, a difficulty of using this paradigm is that it is difficult to prepare a replay condition, which mimics the dynamics of reported percept with the disambiguated stimuli. Replay condition is essential in imaging experiments because researchers wish to find out the brain activity related to resolving the ambiguity of the stimuli by comparing rivalry and replay conditions.

1.2.2.4 Motion-induced blindness

Motion-induced blindness (Bonneh *et al.*, 2001) is a relatively new bistable perception paradigm. When a global moving pattern superimposed on high-contrast objects is presented to an observer with a central fixation point, the observer experiences the objects disappearing and reappearing spontaneously every few seconds.

1.2.2.5 Figure-based bistable perception stimuli

Figures (or paintings) that elicit two different types of percepts can be used to elicit bistable perception (sometimes they are mentioned as 'optical illusions'). The Necker cube is a famous optical illusion (Necker, 1832). A line drawing of a skeleton of a cube is ambiguous; observers are not sure which face is the front or back of the cube. 'My wife and mother-in-law' is a picture with two different interpretations – a young woman or an old woman, and this was first used as stimuli for bistable perception task in (Boring, 1930). Rubin's vase is a monochrome painting that can be seen as two faces looking each other, or a vase on the center; the subjective percept changes depending on the figure-ground interpretation.

These stimuli are very easy to prepare, but some participants may have difficulty to switch their perception and dominance duration may become longer than other types of bistable perception stimuli.

1.2.3 Shared or different mechanisms for perceptual switches?

As reviewed, there are many different stimuli that may be used in bistable perception tasks, which induce spontaneous perceptual switches between two (or more than three) percepts. Is there commonality in the mechanisms behind the perceptual switches across different bistable perception stimuli?

The dynamics of bistable perception are similar across different types of stimuli. Switch rates (per seconds) follow the beta distribution and the gamma rate distribution, and this confirmed with binocular rivalry, slant rivalry, and the Necker cube task (Brascamp *et al.*, 2005). The frequency of perceptual switches, or mean dominance duration, correlates with each other across different tasks; MIB and binocular rivalry, (Carter and Pettigrew (2003); but Gallagher and Arnold (2014) did not find correlation in switch frequencies between MIB and binocular rivalry), and the Necker cube and binocular rivalry (Shannon *et al.*, 2011), suggesting there is a common mechanism across different bistable perception tasks.

Kondo *et al.* (2012) compared individual switch frequencies across the Necker cube, Rubin's vase, and plaid motion stimuli and found significant correlations in switch frequency across different types of bistable perception tasks. Importantly, the correlation of number of switches between the Necker cube and the Rubin's vase were higher than those between the Necker cube and plaid motion, and the Rubin's base and plaid motion. In addition, Kondo *et al.* (2012) reported that performance in auditory bistable perception tasks correlated with those in the Necker cube, the Rubin's vase, and plaid motion tasks (c.f. Pressnitzer and Hupe (2006)). The Necker cube and the Rubin's vase are figure-based bistable perception stimuli, whereas plaid motion is a motion-related, and this should lead the difference in strength of correlations. Collectively, there should some commonality across different types of bistable perception tasks and it would be related to no-modality-specific factors.

1.3 What would affect dynamics of perception in bistable perception task?

In this section, I overview the cognitive processes that influence the

dynamics of bistable perception. There are some overlaps between factors – for example, attention may influence the pattern of eye movements (and vice versa).

1.3.1 Attention

Visual attention is our ability to process information selectively over the visual field (Palmer, 1999). This term has been widely used in cognitive neuroscience, and, indeed, 'attention' has many aspects. Here I categoize attention into two types - endogenous attention and exogenous attention – and discuss the roles of the two attentions in bistable perception tasks. Endogenous attention is controlled by an observer's voluntary effort, whereas exogenous attention is controlled by external triggers, such as appearance of a stimulus and cannot be controlled by the observer's will.

1.3.1.1 Endogenous attention

During bistable perception view, possible interpretations of the stimuli compete each other, and therefore attending one of these interpretations may help to resolve ambiguity (people may refer this as 'volitional control', e.g. (Lack, 1978)). Series of experiments with four different types of bistable perception tasks (slant rivalry, orthogonal grating binocular rivalry, house-face binocular rivalry, and the Necker cube rivalry) suggested that such control affects the speed of perceptual switches (i.e. slower switches or longer dominance duration) (van Ee *et al.*, 2005).

The role of attention may be to increase the gain of task-relevant signals. (van Ee *et al.*, 2009) used auditory and visual bistable perception paradigms to explore the role of selective attention. The ability to control perception is enhanced by multimodal congruency, and such enhancement requires participants' active attention to the stimuli.

On the other hand, similar experiments concluded that the voluntary control is not the only force to govern the dynamics of bistability (Meng & Tong, 2004). Participants attempted to control the content of the

subjective perception by selectively attending one of the stimuli, but the corresponding influences on perceptual dynamics were weaker in a binocular rivalry task than in the Necker cube paradigm. Selective attention, the effort to control the percept, is expected to enhance the attended stimulus and to suppress the unattended stimulus. Hence, the failure of successful control of percept suggested that interocular competition should involve in the formation of perception under binocular rivalry task. Importantly, (van Ee *et al.*, 2005) also pointed out that volitional control provokes greater influence on the dynamics of subjective percept in figure-based rivalry (i,e, slant rivalry and the Necker cube rivalry) than in binocular rivalry.

1.3.1.2 Exogenous attention

Exogenous attention is controlled by external factors, such as saliency of the stimuli, and is an automatic process and cannot be controlled voluntarily. Studies suggested that object-based exogenous attention, cued by stimuli, helps to resolve ambiguity in binocular rivalry task (Ooi & He, 1999; Mitchell *et al.*, 2004). This effect is stronger at the onset of rivalry than during maintenance of the percept (Chong & Blake, 2006).

The question is whether exogenous attention is required to induce perceptual switches, or it is not always necessary to switch between exclusive percepts in bistable perception. (Zhang *et al.*, 2011) employed a dual task paradigm and showed that attention is indispensable for the resolution of binocular competition. They hypothesized that rivalries would occur with unattended stimuli if binocular rivalry happens as a fully automatic process that does not require any attention. The results suggested that binocular rivalry stops when the participants diverted their attention from the stimuli.

Similar results were reported in a different binocular rivalry task. (Brascamp & Blake, 2012) developed a binocular rivalry paradigm consisting of there periods; monocular stimulation periods, dual-task intervening period, and rivalry period. They speculated that the image an

observer is perceiving allows one to predict which stimulus is more likely to be the dominant image a given time interval later, as the dominance duration that is close to the mean individual dominance duration is much more likely than shorter or longer ones (Levelt, 1967). Thus, the relation between perception before and after an intervening period would tell whether binocular rivalry occurred during the intervening period, where attention was reverted from the rivaling stimuli even if subjective percept is not reported during the intervening period. The dynamics of the percept indicated that binocular rivalry does not occur during the dual-task intervening period, suggesting attention is essential to cause perceptual rivalry.

Taken together, these results support the idea that both attention processes, exogenous and endogenous attention, contribute to achieve unitary percept by suppressing irrelevant information (i.e. non-dominant percept or stimulus).

1.3.2 Expectation

Top-down (or feedback) information, or the signal coming from higher parts of the brain architecture hierarchy, influences visual information processing by facilitating relevant information processing or (possibly) suppressing irrelevant information (cf. (Lavie *et al.*, 2004)..

It has been reported that knowledge and previous experience may alter visual perception (e.g. Mooney image perception, (Gorlin *et al.*, 2012)). This also happens during bistable perception. For example, (Sterzer *et al.*, 2008) demonstrated that expectation induced by experimental manipulation alters the subjective percept in a bistable perception task. In their experiment, participants were asked to look at SFM stimuli with red-green anaglyph glasses and report their percept. They were also given a false instruction that the anaglyph glasses are special glasses, and they would bias their percept toward right-rotation or left-rotation. During the learning phase, the stimuli contained depth cue and the participants response was biased according to the physical cue. Such bias continued in the test session,

where the stimuli did not contain depth so that the stimuli were physically ambiguous, suggesting that the persisting response bias was purely due to the participants' enduring false belief about the anaglyph glasses. The bias went away in the baseline session (SFM viewing without the anaglyph glasses), further confirming the hypothesis that expectation alters subjective perception.

1.3.3 Duration of stimulus presentation

There are two presentation paradigms for bistable perception experiments – continuous presentation and intermittent presentation.

In the continuous presentation paradigm, typically the stimulus is shown for a few tens of seconds or more. Participants are asked to hold buttons to indicate their percept or press buttons when a perceptual switch occurs. On the other hand, in an intermittent presentation paradigm, bistable stimuli are presented briefly and repeatedly (for a few hundred milliseconds to a few seconds).

The dynamics of subjective perception are very different between the two paradigms. Intermittent presentation causes many fewer perceptual switches than continuous presentation (Orbach *et al.*, 1963; Leopold *et al.*, 2002). (Carter & Cavanagh, 2007) pointed out that the mechanisms behind rivalry are different comparing the two paradigms, and subjective reports in the intermittent presentation paradigm are more likely to be a reflection of participant's preference towards stimuli. (Brascamp *et al.*, 2009) showed that the intermittent presentation paradigm induces periodic and slow perceptual switches compared to the continuous presentation paradigm, suggesting that perceptual switches do occur in the intermittent presentation paradigm.

The important point here is these studies have changed the view of bistable perception. Traditionally, it was thought that there were no memory effects in bistable perception tasks, as the successive dominance duration is random in a continuous presentation paradigm (e.g. (Blake *et al.*, 1971)). However, as explained in 1.3.1.2, the intermittent presentation

paradigm induces strong stabilization effects on the sequence of perceptual reports (Orbach *et al.*, 1963; Leopold *et al.*, 2002). This indicates that sensory memory influences interpretation of bistable perception stimuli (Leopold & Logothetis, 1999; Pearson & Brascamp, 2008), and it may be important to consider such effect in the experiment (Brascamp *et al.*, 2008).

1.3.4 Eye movements

Eye position and eye movements change the sampling of visual information, and the relationship between eye movements and perception under rivalry has been investigated (e.g. the Necker cube, (Hubel & Wiesel, 1968)).

The relationship between gaze position and switch frequency can be partially explained by influence of instructions to the participants. (Einhauser *et al.*, 2004) showed a close link between dominant percept and eye gaze position, and that moving eye positions are necessary for perceptual switches. Their data implied that shift of eye position might provide a negative feedback signal, that suppresses the current percept. Similarly, there is a relationship between the patterns of saccades and fixation positions in bistable perception paradigms (van Dam & van Ee, 2005; 2006). Indeed (van Dam & van Ee, 2006) showed that it is retinal image shifts that causes perceptual switches (i.e. not eye movements).

Pupil dilation also shows predictability of perceptual switches (Einhauser *et al.*, 2008), but it might be explained by the timing of decisions for button response rather than specific to changes in subjective percept (Hupe *et al.*, 2009). Hence it is less likely that these changes in pupil size or subtle eye movements are the "driving forces" of perceptual switches in rivalry tasks. Still, the optokinetic reflex correlates with perceptual switches, and can be used as index for perceptual switches (Logothetis & Schall, 1990; Naber *et al.*, 2011).

1.3.5 Individual difference – Genes, psychiatric condition, and the brain

It is known that the dynamics of bistable perception show individual

differences, and it can be an interesting avenue for investigating the brain mechanisms serving visual perception.

(Miller *et al.*, 2010) showed that dynamics in binocular rivalry are much more similar between monozygotic twin pairs (i.e. identical twins) comparing to dizygotic twin pairs or unmatched twin pairs, suggesting that genes influence subjective perception in bistable perception tasks. (Shannon *et al.*, 2011) replicated the results with a binocular rivalry task and the Necker cube task.

Factor analysis reveals a relationship between bistable perception and neurotransmitters. Kondo *et al.* (2012) (See 1.2.3) compared individual switch frequency across genotype groups to ask if the genes contributing to regulation of neurotransmitters involve in bistability. Their analysis revealed significant differences in switch frequency in auditory bistable perception tasks across COMT genotype groups, and significant differences in switch frequency in shape bistable perception tasks (the Necker cube and the Rubin's vase) across HTR2A genotype groups, suggesting that resolving ambiguity in auditory stimuli involves dopaminergic systems, whereas resolving shape ambiguity requires serotoninergic systems.

Psychiatric disorders can influence subjective perception of bistable stimuli. Bipolar patients tend to experience fewer perceptual switches than healthy controls (Pettigrew & Miller, 1998), but such a tendency was not found in major depression patients or schizophrenia patients (Miller *et al.*, 2003).

Focal volumes of the parietal cortex also correlates with individual differences in bistable perception tasks. (Kanai *et al.*, 2010) analyzed structural brain images with voxel-based morphometry (VBM), which quantifies volume of brain tissues voxel-by-voxel (Ashburner & Friston, 2000). The volume of the superior parietal lobule correlated with individual mean dominance duration in SFM, suggesting that differences in the brain structure link to individual tendency in visual perception performance (also see 1.4.1.4).

1.4 Neural correlate of perceptual switches

In the previous section, I have reviewed how cognitive factors impact on bistable perception performance. How are such processes are implemented in the brain? There are many studies addressing this question, and indeed the neural correlates of perceptual switches or subjective perception have been reported in many areas in the brain (see Table 1-1).

Table 1-1 Summary of brain activity related to bistable perception tasks.

Study	Stimuli	Comparison	Regions
Kleinschmidt et al. (1998)	Rubin's vase, and 'my wife and mother-in-law'	Perceptual switches vs stable percept	occipital lobe, frontal lobe parietal lobe, thalamus, and cingulate
Tong et al. (1998)	Face-house binocular rivalry	House vs face	FFA/PPA
Lumer et al. (1998)	Face-grating binocular rivalry	Grating vs face	Visual cortex / fronto-parietal areas
Haynes et al. (2005)	Binocular rivalry with gratings	Right-eye image vs left-eye image	LGN
Wunderlich et al. (2005)	Binocular rivalry with gratings	Right-eye image vs left-eye image	LGN
Sterzer and Kleinschmidt (2007)	Apparent motion	Spontaneous perceptual switch vs stimulus-driven switches	inferior frontal area
Wilcke et al. (2009)	Grating binocular rivalry	Rivalry vs fusion	Superior parietal lobule / precentral gyrus etc
Knapen et al. (2011)	Binocular rivalry	Spontaneous perceptual switch vs stimulus-driven switches	Visual cortex; no fronto-parietal activity for Rivalry > matched-replay
Weilnhammer et al. (2013)	Lissajous stimuli	Spontaneous perceptual switch vs stimulus-driven switches	hMT+ and IFG

1.4.1.1 Thalamus

1.4.1.1.1 LGN

Perceptual switches in the binocular rivalry paradigm induce changes in activity in lateral geniculate nucleus (LGN). Wunderlich *et al.* (2005) and Haynes *et al.* (2005) showed that activity in LGN is modulated during binocular rivalry view. (Wunderlich *et al.*, 2005) used high- and low-contrast sinusoidal gratings as stimuli for their binocular rivalry task, and found that activity in LGN is modulated depending on subjective percept; activity in LGN increase when a high-contrast grating is perceived, and decreased when a low-contrast grating is perceived. (Haynes *et al.*, 2005) showed that eye-specific activity in LGN is modulated during binocular rivalry paradigm; when an observer perceives the image exposed to their right eye, the eye-specific voxels in LGN, which increase activity during right-eye stimulation, increases its activity. Thus these studies suggest that selection of competing stimuli may happen as early as in LGN.

However, the role of LGN is still controversial. As explained above, human fMRI studies indicated involvement of LGN in perceptual switches. On the other hand, primate electrophysiology reveals no change in LGN activity during binocular rivalry (Lehky & Maunsell, 1996). This may be due to difference in experimental paradigms. Lehky and Maunsell (1996) did not record subjective reports (e.g. button press) from the animals (primate). On the other hand Wunderlich *et al.* (2005) and Haynes *et al.* (2005) asked participants to give explicit report of their percept by pressing buttons. The difference in the report conditions causes notable differences in brain activity in human (e.g. (Frässle *et al.*, 2014)), and therefore it is difficult to conclude whether LGN actually is involved in perceptual switches in binocular rivalry or not.

1.4.1.1.2 Pulvinar

The pulvinar is a part of the thalamus and a part of the visual pathways in the primate brain (Van Essen, 2005). Lesion studies revealed that the pulvinar contributes to visual perception, especially attention and oculomotor function (Arend *et al.*, 2008).

Perceptual switches of ambiguous figures (Rubin's vase and 'my wife and mother-in-law') transiently activate the pulvinar as well as the visual cortex (Kleinschmidt *et al.*, 1998), suggesting these regions may play functional roles in the alternation of subjective percepts. Furthermore, primate electrophysiological studies show that neural activity in pulvinar is related to states of visual awareness in generalized flash suppression task (Wilke *et al.*, 2009).

1.4.1.2 Visual cortex

It has been shown that the contents of subjective perception correlate with activity in the visual cortex (Lumer *et al.*, 1998; Polonsky *et al.*, 2000; Tong & Engel, 2001). Especially, (Lee *et al.*, 2005) demonstrated the 'travelling wave' effect. In this experiment, high- and low-contrast orthogonal annulus-shape monocular gratings were presented to the participants. When the contrast of the low-contrast image was increased in a small region for a short period, the low-contrast image becomes the dominant image, and the observer felt as if the high-contrast image spreads over the annuals (Figure 1-1).

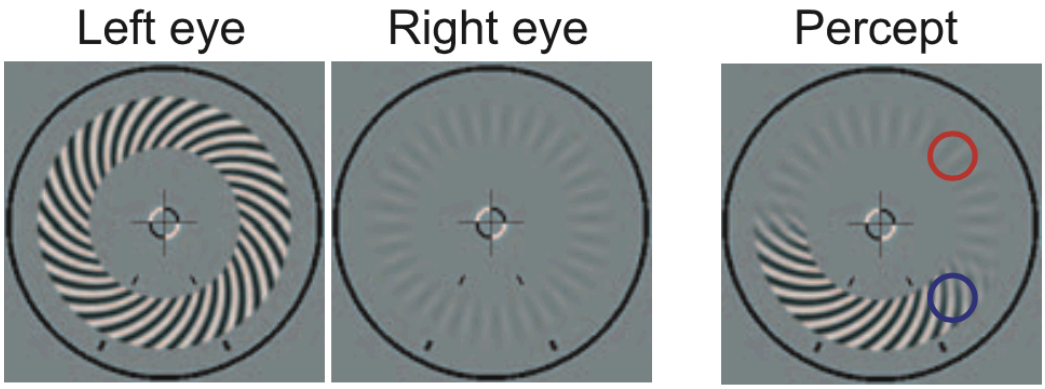

Figure 1-1Travelling wave stimuli.

Brief increase in the contrast of he low-contrast image (right eye) induces perceptual switch and the participant perceive as if the low-contrast image (red circle) spreads over the annual. Image adopted from (Lee *et al.*, 2005).

In addition to focal activation, spatial patterns of fMRI signal in visual cortex represent the contents of conscious percept. Multivariate pattern analysis (MVPA) is an fMRI data analysis methodology that characterizes the predictive role of patterns across more than one voxel at once (Norman *et al.*, 2006). MVPA was first applied to human fMRI data by Haxby *et al.* (2001). They reported that the pattern of activity in ventral temporal cortex shows stimulus-category specificity. This report has opened up the possibility of decoding analysis with human fMRI data.

Kamitani and Tong (2005) and Haynes and Rees (2005) reported that activity in visual cortex predicts subjective perception of visual stimuli. Kamitani and Tong (2005) analyzed fMRI data with support vector machine, a supervised learning algorithm, and found that activity in V1 are predictive for attended motion direction of the plaid stimuli. Haynes and Rees (2005) employed binocular rivalry paradigm (blue and red gratings) and found that prediction accuracy for subjective percept is better in early visual cortex (V1-V3) comparing to V5/MT. These studies suggest that subjective percept is represented at the earliest stages of visual cortex.

1.4.1.3 Temporal regions

The dominant (conscious) percept during bistable visual perception elicits activity in temporal areas when the stimuli have corresponding category specificity. Fusiform face area (FFA) and parahippocampal place area (PPA) increase their activity when face or house stimuli become dominant (Tong *et al.*, 1998). FFA shows face-selectivity and is thought to be a module representing faces (Sergent *et al.*, 1992; Kanwisher *et al.*, 1997), and PPA is thought to encode place or scene (Epstein & Kanwisher, 1998). During binocular rivalry view, retinal inputs remain constant, and thus the coupling of conscious percept and activity in PPA/FFA indicates that activity in these areas reflects the subjective percept and suppression of the unperceived stimulus. Single-neuron recording from patients further validated this finding; neurons showing category-selectivity for visual stimuli increased firing rate when the favored stimuli were perceived (Kreiman *et al.*, 2002).

1.4.1.4 Frontal and parietal regions

Leopold and Logothetis (1999) and Sterzer *et al.* (2009) proposed that fronto-parietal regions have a causal role in generating perceptual switches. Lumer *et al.* (1998) performed binocular rivalry experiments with a face image and a grating image. They compared brain activity during the rivalry period and the replay period, where stimuli alternate to mimic the dynamics of subjective reports during the rivalry period. They reported that frontal and parietal regions increased their activity in the rivalry period comparing to the replay period, indicating that these regions play a role in resolving ambiguity in stimuli. Similarly, activity in fronto-parietal regions during perceptual switches was reported with apparent motion task (Sterzer & Kleinschmidt, 2007) and Lissajous stimuli (Weilnhammer *et al.*, 2013). Cross-correlation of activity in visual cortex, and prefrontal and parietal cortex were stronger in rivalry conditions than in replay condition (Lumer & Rees, 1999). EEG experiment also supported the importance of

fronto-parietal activation - pre-stimulus activity in parietal area is related to perceptual switches (Britz *et al.*, 2009). The findings in the previous paragraph are all correlational relationship between focal brain activation and perceptual switches in bistable perception.

Transcranial magnetic stimulation (TMS) experiments showed the causal involvement of the parietal areas in perceptual switches (Carmel *et al.*, 2010b; Kanai *et al.*, 2010; Zaretskaya *et al.*, 2010; Kanai *et al.*, 2011). Importantly, impairing two subregions in superior parietal lobule (SPL), right anterior SPL (r-aSPL) and right posterior SPL (r-pSPL), caused opposite effects in mean dominance duration (Kanai *et al.*, 2010; Kanai *et al.*, 2011). While repetitive TMS stimulation on r-aSPL prolongs dominance duration (Kanai *et al.*, 2011), stimulation on r-pSPL shortens mean dominance duration (Kanai *et al.*, 2010). Similarly, TMS on r-aSPL shortens dominance duration (Carmel *et al.*, 2010b) and repetitive TMS on the right intraparietal sulcus prolongs dominance duration (Zaretskaya *et al.*, 2010). These all suggests functional involvements of fronto-parietal areas in forming conscious perception during bistable perception.

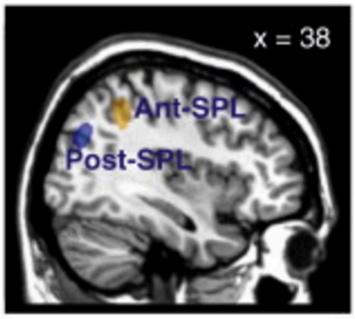

Figure 1-2 r-aSPL and r-pSPL.

Kanai et al. (2011) applied repetitive TMS stimulation to disrupt activity in r-aSPL (yellow, (*x*, *y*, *z*) = 36, -45, 51) and r-pSPL (blue, (38, -64, 32)), and they found the stimulation on these two regions showed opposite effects on mean dominance duration. Picture is adopted from (Kanai *et al.*, 2011).

However, recently, the causal involvement of fronto-parietal areas' in bistable perception has been questioned; (Knapen *et al.*, 2011) proposed that activity in fronto-parietal cortices may be associated with the start of

perceptual transitions, and not directly related to perceptual switches. Binocular rivalry stimuli often induces a "mixed percept", where perception is mixed and neither of the two monocular images becomes dominant. (Knapen *et al.*, 2011) modified the binocular rivalry paradigm to consider the duration of mixed percept; they mimicked gradual transitions of dominant percepts during a replay condition and found activation in fronto-parietal cortex decreased when better matched transitions were included in the analysis. Frässle *et al.* (2014) showed that fMRI activation in fronto-parietal areas may be associated with active reports of the contents of perception (i.e. button press). These findings suggested that activity in fronto-parietal regions might not be necessary or sufficient to cause perceptual transition; it is not driving-force of the switches, but is instead related to some top-down or higher-order visual information processing.

1.5 Computational model for visual perception – suppression and predictive coding

I have reviewed the empirical studies and discussed factors that influence the dynamics of conscious percept during bistable perception. This section now explores how such factors could be accounted for from a computational perspective.

1.5.1 Inhibition and suppression

1.5.1.1 Interocular suppression

Bistable perception stimuli elicit two distinctive percepts. To form a stable percept, the observer may need to suppress sensory signal for unperceived stimuli as they distract the perceived image. This idea led to computational models with inhibitory connections between layers or populations of neurons. Suppression models are popular to explain the dynamics of binocular rivalry (Tong *et al.*, 2006).

(Tong & Engel, 2001) and colleagues provided some empirical

evidence to support the intraocular suppression model for binocular rivalry. They presented superimposed rivalry stimuli in the part of the visual field corresponding to one of the blind spots (note that blind spot location differs between the eyes, see Figure 1-3). They found that BOLD signals in the V1 blind spots increased when the grating on the ipsilateral side became the dominant percept, whereas the signal was suppressed when the blind-spot grating became dominant. This study provides evidence that binocular rivalry emerges from intraocular inhibition, rather than inhibition for the stimulus representation.

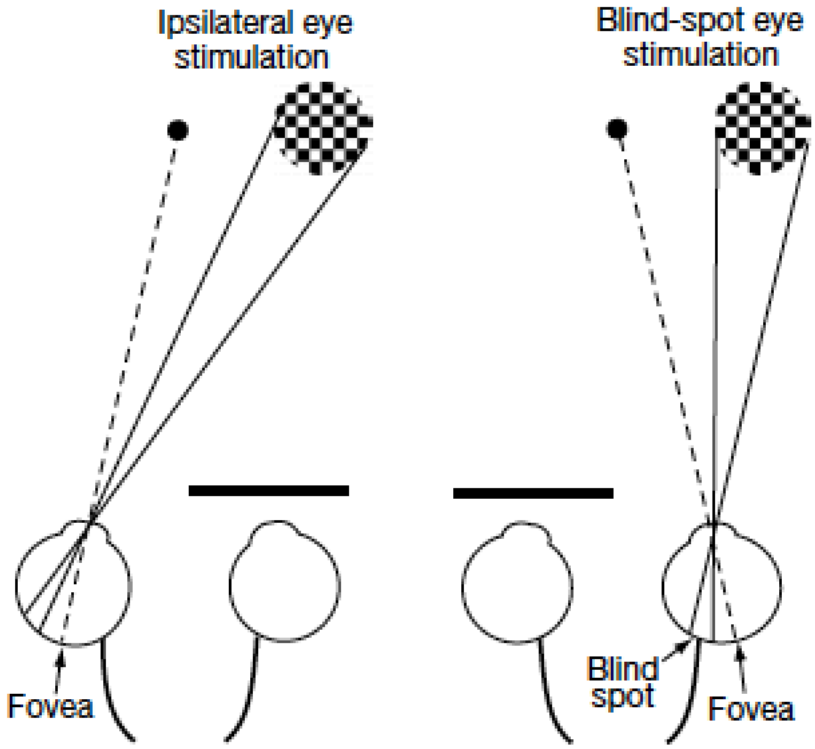

Figure 1-3 Blind spots

Participants performed binocular rivalry task. As the picture indicates, location of the blind spot is different between the right and the left eye. Therefore, only one eye receives visual input on the blind spot when superimposed rivalry gratings are presented. The image adopted from (Tong & Engel, 2001).

The results from (Tong & Engel, 2001) suggest that rivalry can be mediated by intraocular competition. This may not be generalized to other bistable perception task or other types of binocular rivalry, but biased competition theory suggests a role for inhibitions at different levels (see 1.5.2.2) and therefore inhibition should be one important factor to explain

perceptual switches and their dynamics.

(van Loon *et al.*, 2013) reported that concentration of inhibitory neurotransmitter, gamma-aminobutyric acid (GABA), may account for individual variability in the dynamics of bistable perception. In their study, the participants underwent three different bistable perception task – binocular rivalry, MIB, and SFM and magnetic resonance spectroscopy[2] session. Participants with higher concentration of GABA in visual cortex show slower perceptual switches (i.e. more stabilized percept, longer mean dominance duration). Furthermore, pharmacological intervention with lorazepam, a GABA agonist, prolongs individual dominance duration for MIB and SFM[3]. These results further support the role of inhibition in the visual cortex, not only for binocular rivalry, but also possibly for other bistable perception tasks.

1.5.1.2 Biased competition

Biased competition theory is one framework to understand the mechanism of visual perception. The amount of information in visual inputs is quite rich, and we obviously cannot perceive everything. For example, it is often the case that we do not consciously perceive objects even when they are physically present in the field of view (e.g. attentional blink (Raymond *et al.*, 1992)). To maximize the efficiency of information processing, our brain needs to put weight on useful information and suppress trivial or unimportant visual information.

Biased competition theory proposes that the objects in the visual field compete each other to emerge into conscious percept. Attention is often interpreted as enhancement of attended areas or objects, but biased competition theory focuses on the complementary 'suppression' of unattended stimuli (Desimone, 1998).

---

[2] Magnetic resonance spectroscopy (MRS) is a non-invasive in vivo imaging method to investigate the distribution of molecules.

[3] Interestingly, the authors reported that participants had difficulty to perform binocular rivalry task during lorazepam administration.

As discussed in the previous section, attention plays important roles in forming the subjective percept in bistable perception. Indeed, binocular rivalry can be interpreted as a form of biased competition (Dieter & Tadin, 2011). This view can be also applied to other forms of bistable perception stimuli, as the bistable perception stimuli are ambiguous and the observer needs to suppress the unperceived percept to form a unitary percept.

1.5.1.3 Central oscillator hypothesis and interhemispheric switches

(Miller *et al.*, 2000) proposed the interhemispheric switch hypothesis, which explains bistable perception as a competition between each hemisphere's higher visual regions. They found that unilateral caloric vestibular stimulation, as well as TMS stimulation of the temporo-parietal cortex, induces changes in the dynamics of bistable perception, and they claimed that activation or disruption of a brain hemisphere may thus affect the perceptual switches. Although this hypothesis was not supported by the results from tristable perception experiments (Wallis & Ringelhan, 2013), the results in (Miller *et al.*, 2000) possibly suggested that bistability may arise from oscillatory activity from the brain stem, and this may explain why individuals show similar dynamics under different bistable perception tasks (Carter & Pettigrew, 2003).

1.5.2 Predictive coding theory

1.5.2.1 Brain as inference machine

Predictive coding theory has its origins in the ideas originally proposed by Helmholtz (e.g. (Helmholtz, 1910)). He proposed that the perception is a form of inference based on sensory inputs, and this idea has been further elaborated by cognitive neuroscientists (e.g. (Shannon *et al.*, 1956; Neisser, 2014)) and more recently formulated as 'predictive coding' with structure of the retina (Srinivasan *et al.*, 1982) and the visual cortex (Rao & Ballard,

1999).

Bayesian approaches provide a sophisticated implementation for predictive coding theory in the visual cortex (Rao & Ballard, 1999). Under a Bayesian scheme, prediction is represented as a Bayesian prior, and the prediction error, which explains away the difference between sensory inputs and the prior, is implemented as the posterior (i.e. update of the prior). Such model can account for attentional modulation effect in non-human primate V4 (Rao, 2005).

Evidence in favor of predictive coding theory has been reported in several fields, such as reward processing (Iglesias *et al.*, 2013) and visual decision making (Summerfield *et al.*, 2006). Predictive coding theories have been employed to explain many aspects of human cognition, but some of them have not been discussed under the context of prediction or anticipation traditionally (Bubic *et al.*, 2010). We need more empirical evidence to test feasibility of the predictive coding theory and to generalize this idea as a principle to explain human perception (Friston, 2010).

1.5.2.2 Predictive coding in bistable perception

Hohwy *et al.* (2008) proposed a scheme to understand binocular rivalry as representing the outcome of predictive coding between the visual cortex and other brain areas higher in the brain hierarchy. During binocular rivalry, representations of the two different visual stimuli (sensory inputs) compete with each other. To resolve the competition and form a unitary percept, our brain – specifically brain areas at a hierarchically higher level – should generate hypotheses about the causes of the perception, and send such a prediction to sensory areas to suppress irrelevant information and stabilize a unitary percept.

Figure 1-4 illustrates how predictive coding theory accounts for perceptual switches in binocular rivalry. Initially, one of the images becomes the dominant percept, and this yields diminishing returns for the (ongoing) prediction. Still, the brain needs to seek the best explanation of the unexplained error signal coming from the suppressed stimulus, causing

decrease in the prior and exploration of the free energy landscape (which describes the dynamics of the state of the brain).

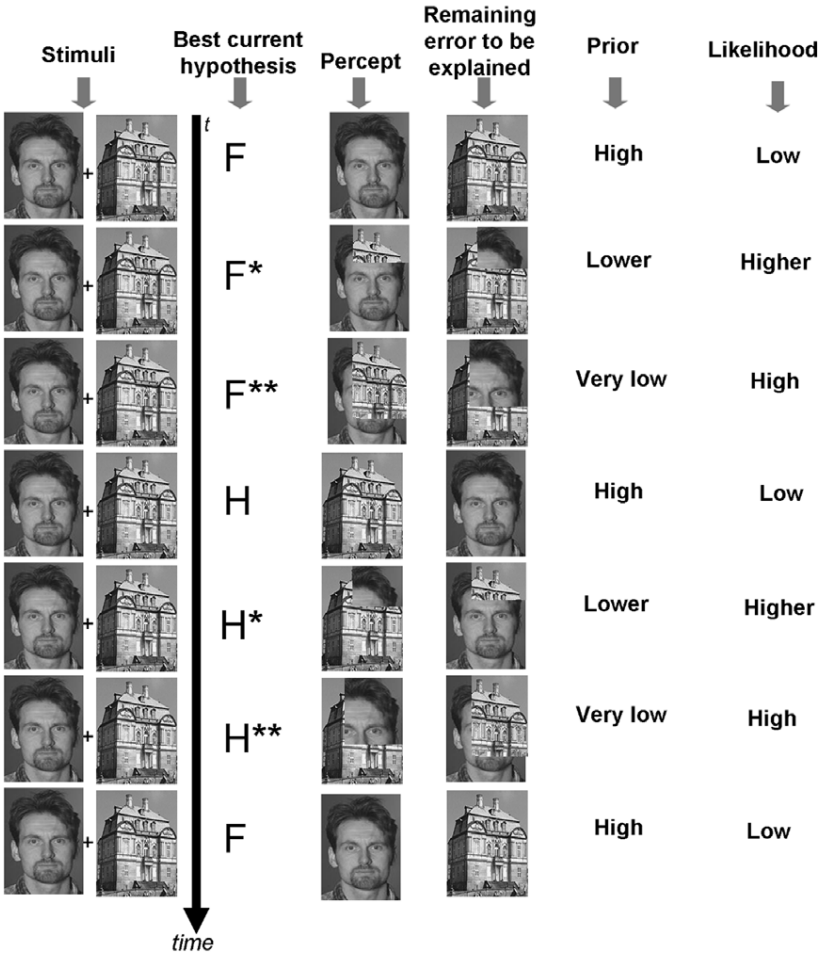

Figure 1-4 Predictive coding theory explains mechanisms behind perceptual switches in binocular rivalry task.

Binocular rivalry with face-house stimuli. Starred hypotheses signify explorations of the free-energy landscape (i.e. the states of entity). The brain choose dominant percept based on the best current hypothesis. The picture adopted from (Hohwy *et al.*, 2008).

The advantage of predictive coding is Predictive coding models are compatible with biased competition theory that predicts attention facilities information processing by suppressing irrelevant information (see 1.5.2.1). (Spratling, 2008a) and (Spratling, 2008b) developed a predictive coding model which accounts for different aspects of attention by defining error-detecting nodes and prediction nodes.

The predictive coding model can explain subjective percept in SFM

bistable perception task (Figure 1-5). The hypothesis (or prediction), prediction error, and sensory input are computed in aSPL, pSPL, and V5 respectively (see Kanai et al., 2011). Hypothesis, prediction error, and sensory input of light-spin and right-spin should be represented by different population of neurons, as subjective percept of SFM can be read out from BOLD signal in these regions ((Brouwer & van Ee, 2007)). The competing hypotheses (right spin or left spin) should suppress each other and this also contributes to 'explain away' the suppressed percept.

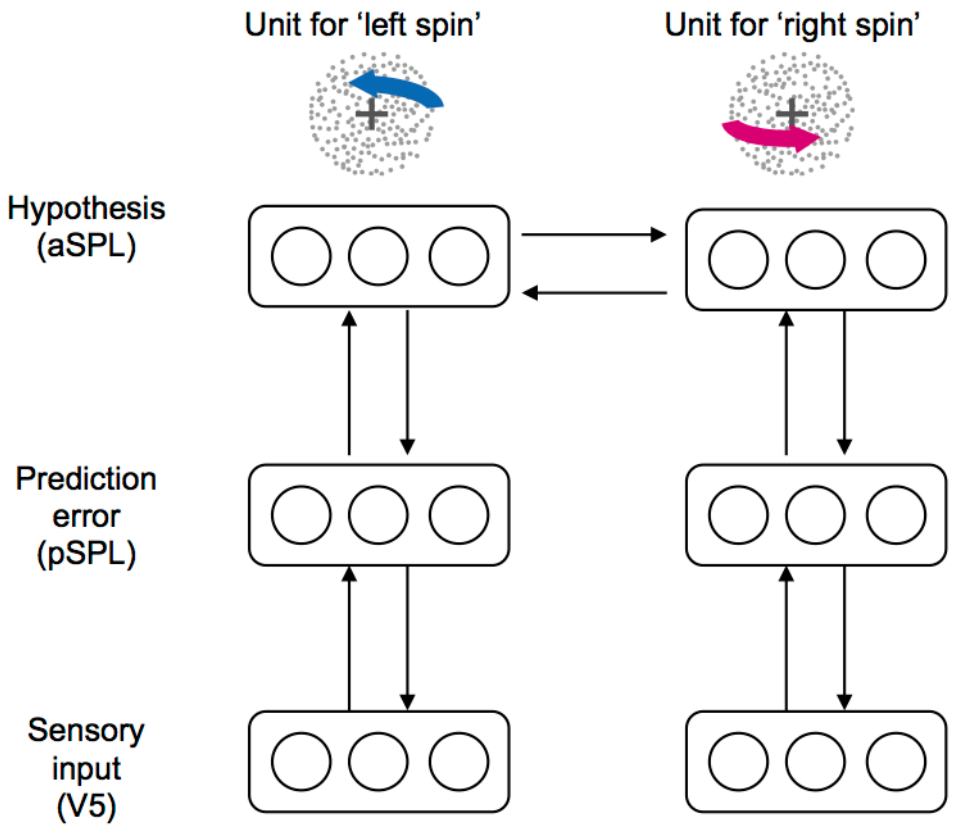

Figure 1-5 Schematic illustration explaining perception in SFM bistable perception task with predictive coding model.

Basic concept is explained in Howehy et al. (2008) and Kanai et al. (2011).

1.5.2.3 What is top-down and button-up?

Although predictive coding theories have attracted much popularity in the

field of cognitive neuroscience, the theories have two broad problems; how to define top-down and bottom-up, and how to experimentally test feasibility of the theory?

People would agree that the words top-down and bottom-up refer to information flows between different layers in the hierarchical structure of the brain. The problem is how to define the 'top' and the 'bottom' in the brain hierarchy; many studies consider various psychological processes as top-down or bottom-up processing without a clear definition of top and bottom in the brain or neuroanatomical evidence (Bohland *et al.*, 2009). (Felleman & Van Essen, 1991) proposed a hierarchical model of the visual system based on the connectivity map, and it would be sensible to define the hierarchical structure of the brain based on (anatomical) connectivity. However, this strategy may difficult to apply to the human brain, as exploring connectivity rests on speculations – for example diffusion tensor imaging can estimate anatomical connectivity, but it is a probabilistic value and may not be the 'real' connectivity. The terms 'top-down' and 'bottom-up' need more scientific rationale based on anatomy and structure of the brain, and the network models of the brain should consider the validity of the proposed network.

(Rauss & Pourtois, 2013) discussed how to define top-down and bottom-up in the brain. They summarized claims in previous psychophysics studies, and pointed out that psychologists have assumed that (1) the brain has hierarchical organization of information processing with bidirectional information exchange between layers, and (2) Lower levels of the hierarchy serve to represent detailed stimulus information, while higher levels represent more integrated information. However, these are based on psychology definition, and (Rauss & Pourtois, 2013) claimed that top-down and bottom-up needs to be defined based on predictive coding theory – the bottom-up and top-down may not be totally independent each other or opposite, and both *ascending* and *descending* connections contribute to form bottom-up and top-down signal.

1.5.2.4 Implement of prediction and prediction-error as neural activity

The predictive coding theory does not give clear prediction regarding the dynamics of neural activity. Prediction may increase brain activity by facilitating information favoring the likely hypothesis (prior), but other studies suggest that predictability of the stimuli decreases activity by suppressing the irrelevant information.

(Hesselmann *et al.*, 2010) performed an auditory and a visual experiments and compared BOLD signals across hit, miss, false alarm, and correct rejection conditions. They tested the role of pre-stimulus activity by comparing the four conditions, and found that the activity in sensory areas is related to perceptual decisions towards correct inference rather than the stimulus itself. They concluded that the sensory signal encodes the precision of prediction errors (i.e. inverse of variance in the prediction error), not the sensory evidence or prediction errors. Their experiments implied that the sensory evidence or prediction errors per se might be represented in some other areas.

Furthermore, expectation may facilitate perception by sharpening neural representation of the stimuli (Kok *et al.*, 2012). In their experiment, the participants performed visual discrimination task with auditory cues so that they developed associations between the cues and the grating stimuli. While activation in the visual cortex reduced when presented gratings with an expected orientation, comparing to an unexpected orientation.

1.6 Conclusion and research questions

The human brain continuously resolves ambiguity in visual information and forms a unitary conscious percept. Such brain function has been studied using bistable perception stimuli. In this chapter, I have discussed the factors affecting the dynamics of conscious percept in bistable perception tasks and the neural representation of such tasks in the human brain. In summary, previous studies have tried to address two important questions – what causes perceptual switches in bistable perception, and which brain

regions play roles in perceptual switches. These questions are difficult to address, as many cognitive factors involve in stabilizing and switching subjective percepts, so many regions in the brain are implicated.

One way to understand this complicated process is to view the brain as a hierarchical system of brain networks. As discussed, many brain regions – LGN, visual cortex, fronto-parietal areas – have been reported to involve in bistable perception. Hence it is less likely that there is a 'responsible' region for the emergence of conscious percept, and we need to understand how the multiple brain regions work together and form the complex dynamics of visual bistable perception. Predictive coding theory is suitable to explain the dynamics of complicated interactions by assuming reciprocal connections between different layers in the brain hierarchy.

In this thesis, I will focus on these questions:

- Which brain regions play roles in causing perceptual switches?
- What causes perceptual switches or stabilize percept in bistable perception? Especially, how does predictive coding theory account for perceptual switches and stabilization?

1.7 Overview of this thesis

The primary motivation of the work presented in this thesis is to understand how the brain generates complex dynamics of visual perception. Using bistable perception stimuli, I interrogate how the brain processes visual stimuli and forms conscious visual experience.

In chapter 3, I investigate the possibility of LGN modulation. The role of LGN in binocular rivalry is still controversial. One possible way to test the hypothesis is to modulate the activity in LGN through fMRI-based neurofeedback training and test the effect of neurofeedback training on binocular rivalry task performance. For successful neurofeedback training, it is necessary to establish a methodology to localize LGN at individual level. I will explore stability of functional LGN localizer with high-resolution

imaging sequence.

In chapter 4, I seek the role of interactions between visual cortex and frontoparietal regions in perceptual switches. VBM analysis revealed that volume of the sub-regions in parietal cortex is correlated with individual switch frequency, and TMS stimulation of these regions causes different effects in behavioral performance. I will further explore the role of those two regions in bistable perception with dynamic causal modeling.

Chapter 5 focuses on the relationship between resting-state and visual bistable perception performance. It has been suggested that innate factors (e.g. genes) are related to the dynamics of visual perception (including bistable perception). Recent studies indicated that our brain organizes functional networks, and the structure of these networks is predictive for demographic data, personality, and psychiatric conditions (e.g. (Cole *et al.*, 2012)). Using graph theoretical analysis and multivariate analysis technique, I test if the structure of functional networks can predict individual performance in bistable perception task.

In Chapter 6, I inspect what determines the trial-by-trial dynamics of visual perception. Especially, I focus on the role of expectation formed by previous precepts. As reviewed in this chapter, predictive coding theory predicts that expectation influences interpretation of the visual inputs and such expectation is updated by new experiences. To test the feasibility of predictive coding theory to explain the dynamics of bistable perception and find neural representation of such expectations, I perform bistable perception experiment and measure brain activity using MEG. This study directly addresses the role of expectation in resolving bistability.

Finally, in chapter 7, I summarize findings in the experiments and discuss possible directions of future studies.

Chapter 2 GENERAL METHOD

2.1 Introduction

Human neuroscience studies use various techniques to measure brain activity and analyze the data to expand our understanding for human brain functions. For the work presented in this thesis, I used MRI and MEG to measure human brain activity and associate it with visual perception. In this chapter, I overview the basics of MRI and MEG, and the methods used to analyze such data.

2.2 Human brain imaging and its neural basis

2.2.1 Measuring signals from a human brain

The advent of functional magnetic resonance imaging (fMRI) has enabled us to explore human brain function by observing brain activity *in vivo,* and we have benefitted from its high spatial resolution, while it is non-invasive.

2.2.2 Neural mechanisms behind human in-vivo recordings

Neurons mainly convey information by spikes, or action potentials. These are electrical signals, and therefore we can measure them with electrodes. Intracellular recording, inserting electrodes inside neurons or axons to measure the voltage difference inside and outside the cell, can record the action potentials directly. On the other hand, extracellular recording, placing electrodes near the membrane to measure the transmembrane current induced by spikes, describes the ion current induced by spikes.

The signal from non-invasive recording, such as fMRI, EEG, and MEG, is thought to be related to local field potentials (LFP). LFP is extracellular signal that is a summation of the electric current induced by firing of a small population of neurons. When the action potentials reach the

terminal of the presynaptic neuron (synapse), the synapse releases glutamate. This opens cation channels and the flow of the positive ions makes the extracellular fluid negative (post synaptic potential). This current eventually flows out from the dendrites, making the extracellular fluid positive. As explained, the main source of LFP is synaptic activity, but other factors, for example fast action potentials, calcium spikes, and intrinsic current, also influence the signal (Buzsaki *et al.*, 2012).

I further discuss the relationships between LFPs and fMRI signal, and LFPs and MEG signal in the latter sections.

2.3 Functional magnetic resonance imaging

2.3.1 Magnetic resonance

Magnetic resonance imaging (MRI) was first developed in the 1970s. MRI utilizes magnetic resonance to visualize brain activity. Magnetic resonance is a phenomenon whereby atoms in static magnetic field show synchronous oscillation with specific frequency band of electromagnetic wave.

Protons spin around its axis (Larmor precession), and such spins generate magnetic field (Figure 2-1A). When the directions of the protons' axis are not aligned, the protons interfere each other, and therefore there should not be magnetic field (Figure 2-1B). When strong magnetic field (B_0, usually above 1 Tesla for MRI scanning) is generated, the atoms are aligned to the direction of the magnetic field and the magnetic field emerges as macroscopic magnetic field (Figure 2-1C, D).

Most importantly, the presence of protons in a static magnetic field changes the direction of precession when a radiofrequency pulse excites the protons (magnetic resonance). This radiofrequency pulse changes the direction of the protons' spin, and also the relaxation time to the equilibrium state. The relaxation time differs depending on tissues. Therefore, the differences in tissues can be visualized by varying the magnetic field.

Figure 2-1 Schematic explanation of proton's spin.

(A) Each proton's spin (Larmor precession) generates magnetic field around the proton. (B) There should not be magnetic field when the directions of protons' spin are not the same. (C, D) The protons are aligned to the same direction when magnetic field (B_0) is generated in the MRI scanner. Images are adapted from (Carter & Shieh, 2010).

2.3.2 MR Imaging

MRI observes resonance of protons using a specific radiofrequency pulse (magnetic resonance). Thanks to this, we can visualize tissues by controlling frequency of signal with different magnetic gradient pulses (spin echo imaging).

MRI uses phase-encoding gradient, frequency-encoding gradient, and slice-selective gradient; phase-encoding and frequency-encoding to create 2D images (x-axis for phase-information and y-axis for frequency-information), and slice-selective gradient to reconstruct z-axis information. The MRI scanner releases gradient pulse many times (repeat

as many times as number of pixels in each slice) with different strength of gradient, and we obtain a phase-encoded image. The acquired MRI data is represented as 2D data (i.e. phase-encoding and frequency-encoding information), and this time-space data is called K-space. The combination of two different gradients yield spatial specificity, and therefore, by applying Fourier transformation to the K-space, we obtain an MRI image. In other words, depending on the position of the slice, the difference of the gradient change differs so that information of the position (x, y coordinate in the image) can be reconstructed through Fourier transformation.

T1 relaxation (or spin-lattice relaxation) describes a process whereby protons recover the lower energy state (i.e. restoring longitudinal magnetization) after removing the radio frequency pulse, and T2 relaxation (or spin-spin interaction) means a process whereby transverse magnetization decreases and disappears. T1 and T2 are different across tissues, and hence we can visualize the difference between tissues by spin echo imaging. For functional imaging, echo planer imaging (EPI) was used for the experiments described in this thesis. EPI generates multiple gradient echoes while T2* relaxation (T2 relaxation considering B_0) to fill the K-space.

2.3.3 BOLD effect

The Blood-Oxygen-Level Dependent (BOLD) effect was first reported with mice (Ogawa *et al.*, 1990) and confirmed with cats (Turner *et al.*, 1991). When group of neurons increase their firing rate, the ratio of hemoglobin and deoxyhemoglobin changes; although the amount of oxygen consumed by the neurons increase, more blood flows into the area and therefore ratio of deoxyhemogrobin decreases in the activated area. Thanks to this effect, now we can observe human brain activity *in vivo* ((Kwong *et al.*, 1992; Ogawa *et al.*, 1992)) and relate the signal to cognitive process or behavioral performance.

The amount of BOLD signal change between resting state (i.e. states without external inputs or task conditions) and task state is not very

big, perhaps 6% at maximum (Ogawa *et al.*, 1990).

2.3.4 What we are measuring with fMRI?

fMRI signal rests on changes in local blood flow. It has been known that neural activity increases blood flow (Roy & Sherrington, 1890), and the precise relationship between neural activity and BOLD signal has been investigated using combination of MRI imaging and electrophysiology techniques. The current consensus is that the BOLD signal best correlates with LFPs. Simultaneous intracortical recordings of fMRI responses and neural activity (LFPs, single-unit spiking activity, and multi-unit spiking activity (MUA)) from the visual cortex of anesthetized monkeys revealed that BOLD signal correlates with LFPs, suggesting that that the source of the BOLD should be synaptic input and intra-cortical information flow (Logothetis *et al.*, 2001)

2.4 Basics of fMRI data analysis

2.4.1 Preprocessing of fMRI data

The main purpose of an fMRI experiment is to associate behavior or personal traits to temporal or spatial brain activity patterns. Appropriate preprocessing for the fMRI data is required before applying statistical analysis and make inference regarding the roles of specific brain regions or network. In this section, I will overview the preprocessing steps used for this thesis (experiments in chapter 3, 4, 5). Statistical Parametric Mapping software (SPM: http://www.fil.ion.ucl.ac.uk/spm/) was used for brain imaging analysis.

fMRI images have four dimensions (three spatial dimension plus time), and the level of brain activation (i.e. BOLD signal) is represented as brightness of volumetric pixel (voxel). Most preprocessing steps work on spatial information rather than temporal information.

2.4.1.1 Correcting image distortion using fieldmap images

EPI images are often distorted by magnetic field inhomogeneity. To remove such artefact, mapping information of the spatial distribution of the field inhomogeneities can be obtained by fieldmap imaging (Hutton *et al.*, 2002). Dual-echo EPI sequence (shorter and longer time echo) produces fieldmap images to compute distortions of phase difference and this information was applied to EPI data.

2.4.1.2 *Realignment*

During realignment, EPI images were aligned to the first EPI volume of the scanning run. The resolution of EPI images in this thesis was 3×3×3mm, and therefore even subtle head movements may affect the results. Such movement was modeled with affine transformation (i.e. linear transformation), as there should not be changes in the shape of the individual brain. Six parameters (movement towards x, y, and z axis, rotation towards pitch, roll, and yaw) are computed with the least-square method, which minimizes difference in the six parameters across images (Friston *et al.*, 2003a).

2.4.1.3 Normalize

To make statistical analysis with fMRI data, individual brains should be normalized to standard brain space such as Montreal Neurological Institute (MNI) template. There are several ways to normalize brain images to template images, and most popular approach should be to obtain normalization parameters from individual structural images. First, individual anatomical brain images are aligned to fMRI images (again this transformation should be represented as linear transformation) and a conversion matrix is computed between standard MNI space and the individual brain. Then, the conversion matrix is applied to individual EPI images.

2.4.1.4 Spatial smoothing

The last step of preprocessing is spatial smoothing[4]. The purpose of spatial smoothing is to enhance the power for statistical analysis by reducing noise in the image. The signal of interest (i.e. BOLD signal modulated by experimental manipulation) extends over several voxels, while a proportion of the noise in fMRI images should be independent and therefore it should not show any spatial correlation. According to random field theory, applying spatial smoothing should suppress the effect of noise but does not hurt the signal of interest and thus improves statistical power. In addition, smoothing helps to cancel out subtle difference in locus of brain activation over individuals (Frackowiak *et al.*, 2004).

Spatial smoothing involves averaging over voxels and increase spatial correlation of the signals. Specifically, SPM achieves spatial smoothing by convolving bell-shape Gaussian kernel. Optimal value for full width at half maximum of the kernel may depend on the image resolutions. In this thesis, I used 8mm for experiments in Chapter 4 and 5, and 3mm for the experiment in Chapter 3.

2.4.2 Analysis of BOLD activation with general linear model

When performing fMRI measurement, researchers may be interested in which brain regions or populations of neurons carry specific information about events or perception. This means one needs to compare brain activity associated to specific events (e.g. experimental conditions, perceptual

---

[4] There is debate about utility of spatial smoothing for multi-voxel pattern analysis – some researchers insist smoothing hurt information representation Kamitani, Y. & Sawahata, Y. (2010) Spatial smoothing hurts localization but not information: pitfalls for brain mappers. *Neuroimage*, **49**, 1949-1952.
, Op de Beeck, H.P. Ibid.Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? , 1943-1948.. In this thesis, I have applied spatial smoothing to the EPI images as there is no concerns for applying spatial smoothing for GLM/DCM/connectivity analysis.

states).

One of the ways to do it is to model BOLD responses using the general linear model (Kiebel & Holmes):

$$\boldsymbol{y} = \boldsymbol{X\beta} + \varepsilon.$$

where y is a response variable, X is a matrix which corresponds to explanatory variables. ß is parameters we wish to estimate, and ε is the errors. The model assumes linearity between neural activity (i.e. the sources of the BOLD signal) and hemodynamic response, and that the parameters are normally distributed.

Parameter ß can be obtained as follows:

$$\hat{\beta} = (X^T X)^{-1} X^T \underline{Y}.$$

This is derived from the normal equation.

Statistical tests are performed against combination of experimental condition and the estimated parameter ß. Statistical tests interrogate if the contrast, linear combination of the parameters (across conditions), of estimated ß shows significant difference (i.e. not zero). T-test or F-test was used depending on the experiment's interest.

2.5 Connectivity analysis

2.5.1 Beyond localizing function: local activity to network analysis

Functional neuroimaging studies have focused on investigating regional brain activity and finding out which brain area is 'responsible' for specific cognitive functions. This approach has been successful and revealed specific modules in the human brain.

Another line of studies has shown that the brain forms intrinsic networks, which is often manifested as synchronous fluctuation between regions. One intriguing aspect of the intrinsic brain networks is these networks can be found even without any external input or cognitive performance. Biswal *et al.* (1995) first discovered that low frequency BOLD

signal (below 0.1 Hz) in motor cortex tends to show strong coupling in resting state. Discovery of default-mode network (Raichle *et al.*, 2001) has further validated the utility of resting state functional connectivity analysis. These studies found that some brain regions, for instance the post cingulate cortex (PCC), tend to show greater activation in the resting state[5] than in goal-directed task condition.

2.5.2 Functional connectivity

If brain areas which have similar function for a certain cognitive function show synchronous activity during task engagement, the time course of ROI activity in each of the brain areas should highly correlate each other. Such relationship is called 'functional connectivity'.

The first report of "functional connectivity" was Biswal *et al.* (1995), which presented correlated low-frequency activity inside motor cortex during the resting state as well as during motor task. Although such correlations in BOLD signal were first thought to represent physiological noise and not so much attention was paid to their importance, the discovery of task-positive and task-negative networks (Raichle *et al.*, 2001) (nowadays 'task-positive' and 'task-negative' are not favored, as (1) magnitude of activation are different across tasks and (2) some tasks actually activate 'task-negative' default mode regions (e.g. cite mind-wondering papers, etc). More detailed and systematic exploration of the intrinsic network have been performed by various groups (e.g. Power *et al.* (2011)) showed that 'intrinsic' property exists in the brain even without external input and functional connectivity can characterize architecture of the intrinsic network.

There are still ongoing debates - what determines functional connectivity and what is the basis of the functional network? Functional connectivity is mainly supported by anatomical connections (Greicius *et al.*,

[5] There are various definitions of 'resting state', and indeed the conditions of resting state measurements (i.e. darkness of the room, eyes open or closed) influence functional connectivity per se. Participants for the experiment in chapter 5 were instructed to close their eyes, relax, and to not fall asleep (see chapter 5).

2009), and this was confirmed by an electrophysiological study (Vincent *et al.*, 2007).

2.5.3 Preprocessing for functional connectivety analysis

The BOLD signal itself is very noisy, and functional connectivity is very vulnerable to noise. It is therefore essential to perform preprocessing to remove such signals from the images to detect networks in the brain.

The standard preprocessing protocol for functional connectivity analysis (Van Dijk *et al.*, 2010) requires (1) applying high-pass filtering (2) regression using average signal from whole brain, white matter, gray matter, and cerebrospinal fluid and six motion realignment parameters (x, y, z, pitch, roll, and yaw) estimated during preprocessing stage. Applying high-pass filter is justified as the communication between distant brain regions should be represented as a low frequency signal (Gusnard *et al.*, 2001; Fox & Raichle, 2007), while physiological noise (heart beat, breezing) and movement artefact often appear in high-frequency domain.

2.5.4 Computation of functional connectivity

Functional connectivity is quantified with Person's correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where x_i, y_i (i = 1, 2, 3, …, n) represents time series of BOLD signal in ROI 1 and ROI 2 (averaged across all voxels within ROIs) and $\hat{x}$ and $\hat{y}$ is mean of the activity.

Correlation coefficient is often z-transformed:

$$z = \frac{1}{2}\ln(\frac{1+r}{1-r})$$

Where r represents Pearson's correlation coefficient. This treatment increases flexibility of the analysis, as z-transformed correlation follows normal distribution.(Power *et al.*, 2011)

There are several different strategies for functional connectivity

analysis. The most common approach is the ROI-based approach; compute functional connectivity between ROIs (anatomically or functionally defined) or seed ROI and all voxels. Although the ROI approach is easy to implement and fast to compute, this approach is hypothesis-driven (constraint from ROI definition) and the results should be interpreted with caution. For hypothesis-free approach, independent component analysis (ICA) has been used for resting state and task fMRI data to detect intrinsic brain network (Smith *et al.*, 2009).

In this thesis, I have used an ROI-based approach; ROIs were defined from meta-analysis in previous study (Power *et al.*, 2011), and the ROI set covers the whole brain. Functional connectivity were computed for all the possible ROI pairs, and this should complement any weakness of such an ROI-based approach.

2.5.5 Effective connectivity and dynamic causal modeling

In general, effective connectivity analysis deals with predictability of brain activity between/across regions (Friston, 1994). Popular effective connectivity measures used in neuroimaging studies are Granger causality (Granger, 1969) and Dynamic causal modelling (DCM) (Friston *et al.*, 2003b).

Granger causality was first proposed in the context of economic research (Granger, 1969). Granger causality is computed from predictability between two variables; if time-series of variable A could predict future of variable B, A is the Granger-cause of B. Note that these relationship doesn't have to be mutual; in this case, B may not be Granger-cause of A. Granger causality is widely used for brain imaging data analysis. The advantage of this method is low computational cost and simplicity. The MATLAB toolbox customised for neuroimaging data analysis is also available (Seth, 2010).

DCM computes effective-connectivity analysis by modeling LFPs. DCM analysis for fMRI data aims to infer influence of neural activity by describing changes of BOLD signal as a function of experimental condition. DCM models describe changes in connectivity as follows:

$$\frac{dx(t)}{dt} = \left(A + \sum_{j=1}^{m} u^j B^{(j)} \right) x(t) + C_u$$

Here, $x(t)$ represents brain activations in ROIs and t expresses time, so the equation describes time-dependent changes in the dynamics of brain activation. The right side of the equation shows that the rate of change in brain activity in an ROI can be represented by the combination of brain dynamics of other regions and experimental conditions: A represents the endogenous connectivity matrix (A-matrix) and thus represents context-independent connectivity between regions. Therefore in the present experiment, the values in the A-matrix were the same across conditions. Contextual variables (e.g. experimental conditions) are denoted by the vector u and the matrix B represents the modulations on endogenous connectivity (B-matrix). C represents the driving input (C-matrix) and models the effect of experimental condition (u) on the brain dynamics in the ROI. The values of the B-matrix and C-matrix vary depending on the experimental conditions. A recent study employing electrophysiological recording confirmed that DCM analysis can locate the source of neural activations better than similar analytic tools such functional connectivity and Granger causality, implying DCM parameters can characterize neural dynamics in a biologically and functionally meaningful way (David *et al.*, 2008).

2.5.6 Graph analysis

To understand the topological structure of brain network, graph theoretical analysis has been applied to neuroimaging data. Graph theory deals with complexity of a given network and characterizes the structure of a network by describing connection (edges) between nodes of the graph. Graph is defined as a set of nodes and edges, which define connections between nodes. The graph can be directed or undirected, weighted or unweighted (i.e. binary), and these are determined by how the edges are defined. Graph

theoretical measures can be computed from functional and structural MRI data (Diffusion-tensor imaging).

2.5.6.1 Degree

In chapter 5, resting state data is analyzed based on graph theory. Specially, several centrality measures were computed based on functional connectivity. The centrality measures used in this thesis are computed based on degree and shortest path.

Degree (or node degree) is a quite simple measure. Degree describes how well nodes are connected with each other. It is defined as a number of edges connected with the node.

Node degree for node i in a binary undirected graph [6] can be obtained as follows:

$$k_i = \sum_{j \in N} a_{ij}$$

i is a set of all nodes which belongs to the binary, undirected graph and a_{ij} is a connection status between node i and j. This equation simply means that node degree, a_i, is obtained by summing up the number of connections.

The idea of node degree has been used for brain imaging analysis. Voxel-by-voxel degree based on functional connectivity (degree of connectivity) has been used to characterize patient's brain (Alzheimer's disease (Buckner *et al.*, 2009) and obsessive-compulsive disorder (Beucke *et al.*, 2013)), and to investigate structure of brain and its relationship to demographic data (e.g. (Hampson *et al.*, 2012)).

2.5.6.2 Shortest paths

[6] I computed graph theoretical measures from an undirected binary graph using resting state fMRI data (Chapter 5). Therefore, here I focus on undirected binary graph. See Rubinov, M. & Sporns, O. (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, **52**, 1059-1069. for weighted or directed graph.

The shortest path quantifies the efficiency of a connection and is defined as follows:

$$d_{ij} = \sum_{a_{uv} \in g_{i \leftrightarrow j}} a_{uv}$$

where $g_{i \leftrightarrow j}$ indicates the shortest path between i and j.

2.6 Magnetoencephalography

2.6.1 Measuring brain activity with an MEG system

Neural activity is electrical signal. This means the magnetic field when a neuron fires and magnetoencephalography (MEG) measures such subtle changes in magnetic field[7].

Electrical signal flow induces magnetic field in orthogonal direction and MEG measures the changes in magnetic fields with superconducting coils. The first MEG recording was done with room temperature coil (Cohen, 1968); he measured alpha-band signal from normal and epilepsy patients. Modern MEG uses a large number of superconducting quantum interference devices (SQUIDs), which are magnetometers that can detect subtle magnetic field and its changes.

2.6.2 Structure of the SQUID coils and measurements

2.6.2.1 SQUID coils

The SQUID sensors (Figure 2-2) can detect the magnetic field and its

[7] Electroencephalogram (EEG) also measures electrical signal. EEG is measured by electrode placed on the scalp or the surface of the brain (intracranial EEG) etc. and thus EEG measures electrical potential caused by neural activity, not magnetic field.

changes accompanied with firing of a population of neurons. The current produced by neural activity generates magnetic field. Cooled down the SQUID sensors with helium, the sensors become superconductors, which means they lose electrical resistance (approximately 4 Kelvin, i,e, -269C°) (Hansen *et al.*, 2010). The electrical current on the material generates a perpetual magnetic field. SQIOD sensors have Josephson junctions. Electrical current on Josephson junctions interfere the superconducting loop, generating the voltage. The MEG system detects the magnetic field by measuring this voltage. DC SQUID, which has two Josephon junctions, is preferred for better sensitivity to the signal.

The average voltage across the SQUID depends on the bias current and shows a periodic function of the magnetic flux. The system sends feedback signal to suppress such the periodic oscillation and obtain linear relationship between the input and the signal.

2.6.2.2 Flux transformer

MEG signals are detected using flux transformers coupled with the SQUID sensors. Most MEG systems use gradiometers (axial gradiometer or planar gradiometer) for their robustness against environmental noise. Gradiometers consist of two coils; one coil for picking up signals from neural activity, and the other for detecting environmental noises. Changes in magnetic field should show spatial selectivity, whereas the environmental noises should be relatively uniform across locations. To cancel out the environmental noises, the difference between the signals from the sensors is sent to the SQUID.

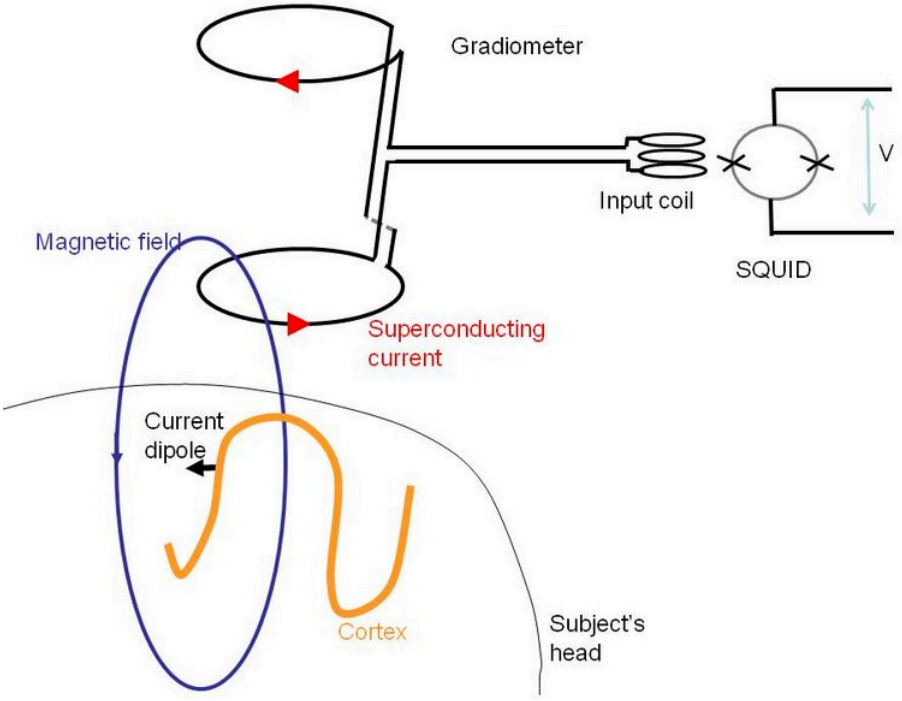

Figure 2-2 Schematic illustration of a dc SQUID magnetometer with a first order gradiometer.

This picture illustrates how the MEG system measures the changes in the magnetic fields. Neural activity generates the magnetic field and the flux transformer pickes up such activity as changes in the magnetic flux. Picture adapted from(Barnes, 2010).

2.6.2.3 Environmental noises

Changes in magnetic fields induced by neural activity are much smaller than the environmental artefacts (Figure 2-3). Therefore MEG systems should be located in a magnetically shielded room. The shielded room is composed of layers of materials with different magnetic properties, such as aluminum, coppers. This is because materials show different resistance to specific frequency magnetic interferences.

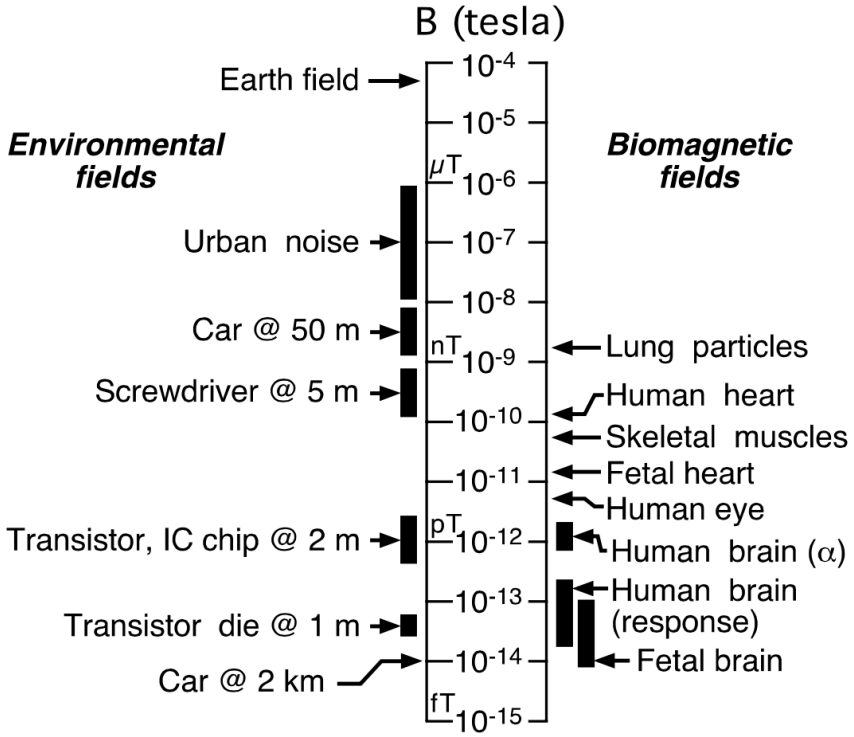

Figure 2-3 Strength of magnetic field

We are surrounded by sources of magnetic field and the strength of the fields varies. Biomagnetic fields are weaker than environmental fields. Note that B (tesla) is log-scaled. Picture from (Vrba, 2002).

2.6.3 Source of MEG signals

As explained in 2-2, MEG should be predominantly measuring LFPs. Especially, excitatory post synaptic potentials (EPSPs), cause relatively big and synchronous current across neurons, particularly from superficial layers, and thus LFPs should be what we measure with MEG. Pyramidal cells will be the main source of such currents.

Considering the sensitivity of the sensors, approximately a million synapses must be simultaneously active during a typical evoked response observed with MEG. Since there are approximately 10^5 pyramidal cells per mm^2 of cortex and thousands of synapses per neuron, the simultaneous activation of as few as one synapse in a thousand over an area of one square millimeter would suffice to produce a detectable signal (Hämäläinen *et al.*,

1993). Current may be cancelled out by neighboring neurons, so more neurons would be required to produce a dipole moment for detectable MEG signals.

2.7 Basics of MEG data analysis

2.7.1 Preprocessing steps

As explained in the previous section, MEG needs to measure magnetic fields produced by neuronal populations and they are relatively smaller than those caused by environmental noise. Thus preprocessing and attenuating the effects of artefacts in the signal is very important. Here I briefly review the preprocessing steps.

2.7.1.1 Data preparation

MEG data has high time-resolution and researchers may want to associate specific events with the MEG signal. This can be done using trigger signal, which enables synchronizing MEG scanner and external devices (e.g. the stimulus presentation screen). The MEG data is divided and marked using the trigger signal. Unnecessary channels can be removed at this step. Also the data can be down-sampled to minimize the file size[8].

2.7.1.2 Filtering

MEG data should be filtered with high-pass filter (at 0.1Hz) and stop-band filter (attenuate 48-50Hz) to remove artefacts. The design of the filters can vary depending on the purpose of the experiments or signal-of-interests. MEG tends to be good at capturing high-frequency signals compared to EEG,

[8] Downsampling should be done carefully, especially if you are interested in high-frequency signal.

so low-pass filtering is not always necessary.

2.7.1.3 Artefact detection

Similar to MRI, heartbeat, breathing, eyeblink, and oculomotor signals can alter the MEG signal dramatically and thus it is necessary to remove such artefact from the MEG signal.

Especially, eyeblink and oculomotor action causes quite big artefacts in frontal sensors and therefore these artefacts should be removed. It is recommended to record Electrooculography (EOG) signal or eye-movements to clean the MEG signal (Gross *et al.*, 2013). In addition, if the experimental design may be sensitive to such artefacts, independent component analysis can be applied to find pattern of activity, which is caused by eye-movement, or eye-blink related noise.

In addition to use the difference in amplitudes, there are several ways to detect artefact signals and they are implemented in standard analysis software such as SPM, EEGLAB (http://sccn.ucsd.edu/eeglab/), and Fieldtrip (http://www.fieldtriptoolbox.org/).

2.7.2 Source estimation

Inferring the source of MEG signal is known as the 'inverse problem'. When we wish to investigate the source(s) of MEG signal(s), the source of neural activity needs to be estimated from the signal obtained with MEG sensors. The problem is there is no single, correct solution; while the number of sensors is limited, there may be infinite number of sources of the current. To solve this inverse problem, localization methods need to incorporate assumptions regarding the source of the MEG signal. For example, dipole fitting assumes there is one brain region that shows time-locked neural activity to stimulus or the event of interest. Such constraints make it easier to obtain a solution. Nowadays, it is typical to employ a distributed solution, which uses many fixed dipoles to estimate the sources of MEG signals (Lopez *et al.*, 2014).

We can describe the relationship between MEG signals and their sources as follows:

$$Y = LJ + \varepsilon$$

Where $Y \in RN_c \times N_n$ is the MEG dataset, N_c are sensors and N_t are time samples. $J \in RN_d \times N_n$ represent sources of neural activity, and N_d are current dipoles distributed over the cortical surface. L, the lead field matrix, determines the dynamics of the magnetic field that influence the signal. To solve this, the source should be estimated as the expectation of the posterior distribution given the data Y:

$\hat{J} = E[p(J|Y)]$.

$p(J|Y)$ be described using Bayesian theorem:

$$p(J|Y) = \frac{p(Y|J)p(J)}{p(Y)}$$

The evidence can be removed, as it is a constant for a given data. Therefore we obtain the following formula:

$$p(J|Y) \propto p(Y|J)p(J).$$

The right hand side of the equation can be expressed as multivariate normal distribution. MEG source inversion is finding the parameters which minimizes the right hand side of the formula with the given data Y and L, which is defined by the head model.

By default SPM solves this by multiple sparse priors (MSP) analysis (Friston *et al.*, 2008). As the name indicates, the algorithm assumes multiple spatial priors, which are enough to cover the whole brain but fewer than dipole, and selects cortical sources.

2.8 Conclusion

I have reviewed human brain imaging methods (fMRI and MEG) and analysis methods for the imaging data - generalized linear model (used in chapter 3 and 4), functional and effective connectivity analysis (chapter 4 and 5), graph theoretical analysis (chapter 5), and MEG source reconstruction (chapter 6).

Chapter 3 LOCALIZING THE LGN AT AN INDIVIDUAL LEVEL

3.1 Summary of this chapter

fMRI studies show that the lateral geniculate nucleus (LGN) is involved in processing visual information during bistable perception (Haynes *et al.*, 2005). The LGN could be an interesting target for neuromodulation (e.g. real-time fMRI neurofeedback) as whether it plays a causal role in bistable perception remains unclear. For successful neurofeedback training, it is important to localize the target region across different training days. Therefore, I sought to establish a reliable way to localize LGN using a functional localiser approach. However, the results showed that localizing bilateral LGN across different scanning days was somewhat difficult even though I obtained comparable amounts of EPI data as for previous LGN fMRI studies. The failure of LGN localization suggested that a more reliable imaging method is needed for future LGN imaging studies.

3.2 Introduction

3.2.1 LGN and binocular rivalry

The detailed anatomy of the LGN has been investigated with primate. Light is first projected on the retina and the neural signals elicited from the retina are sent to the LGN. The LGN receives sensory inputs from retinal ganglion cells and sends output to primary visual cortex (Livingstone & Hubel, 1988). The LGN has a distinctive six-layered structure under the microscope with each layer containing specific types of cells (Figure 3-1); magnocellular cells and parvocellular cells. Magnocellular cells (M-cells) are found in layer 1 and 4, and they have good time-resolution and are highly sensitive to stimulus contrast. Parvocellular cells (P-cells) are found in layers 3-6, and they have high spatial resolution and selectivity to colours (Howard & Rogers, 1995).

As explained in Chapter 1, binocular rivalry paradigms, where each

eye views a different image, such as a face image presented to the left eye, and a house image to the right eye, have been used to investigate the neural correlates of consciousness. Many regions have been reported as being involved in such binocular rivalry paradigms. In the human brain, signals in the lateral geniculate nucleus (LGN) (Haynes *et al.*, 2005; Wunderlich *et al.*, 2005), early visual cortex (Polonsky *et al.*, 2000), the parahippocampal place area (PPA), and the fusiform face area (FFA) (Tong *et al.*, 1998) correlate with the ongoing percept reported by the individual viewing the rivalry stimuli. In particular, LGN neurons are driven strongly by monocular input. In humans, some voxels in the LGN show eye-specific response patterns during binocular rivalry (Kastner *et al.*, 2004; Haynes *et al.*, 2005; Wunderlich *et al.*, 2005), suggesting that the NCC appears at an early stage of the visual information stream.

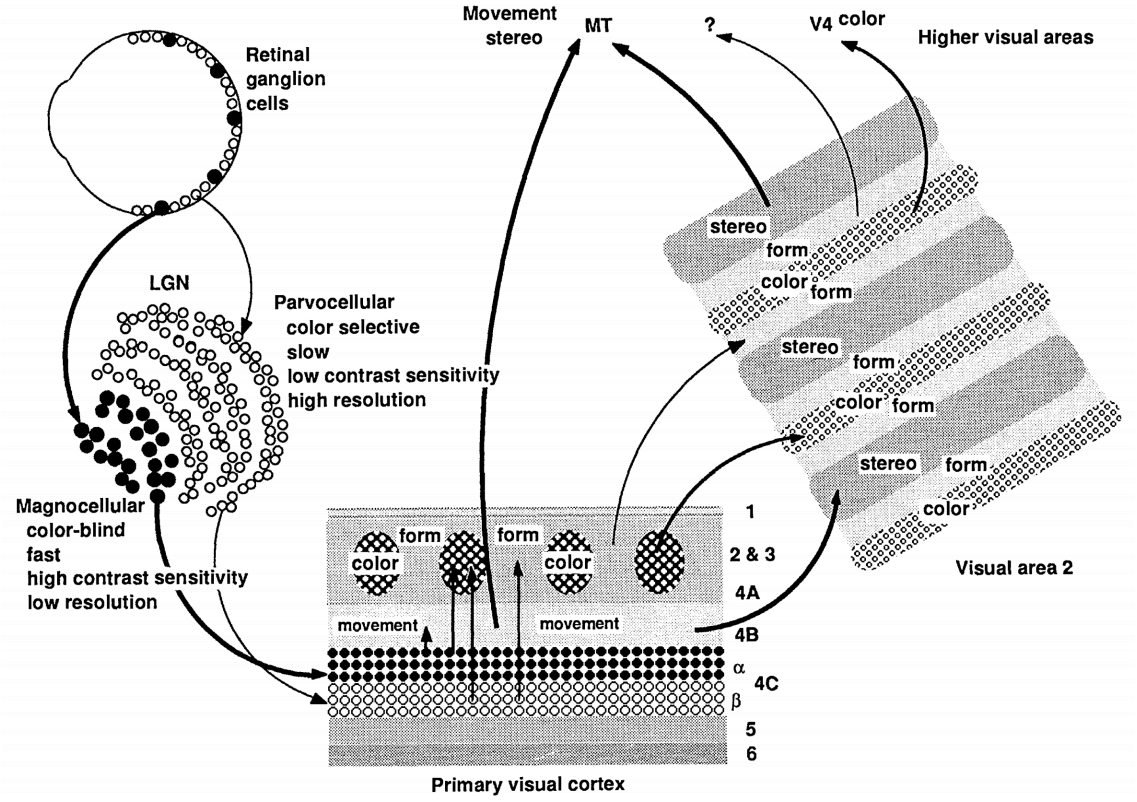

Figure 3-1 Retina, LGN, V1, and V2.

Neurons in the primate LGN receive input from retinal ganglion cells, and

they project to neurons in V1[9]. LGN has laminar structure and two different types of cell are found in different layers. Illustration from (Livingstone & Hubel, 1988).

3.2.2 What is the functional role of LGN in binocular rivalry?

Although human fMRI studies show that activity in subcortical structures such as the LGN correlates with the conscious percept in binocular rivalry task (Haynes *et al.*, 2005; Wunderlich *et al.*, 2005), the functional role of the LGN is controversial (Tong *et al.*, 2006). As the LGN has bidirectional anatomical connections with early visual cortex, it might be the case that the LGN activation found in binocular rivalry tasks simply reflects feedback signals associated with binocular competition in early visual cortex. One recent fMRI study suggested that attention is required to resolve perceptual competition during binocular rivalry task (Zhang *et al.*, 2011), and indeed it has been suggested that attention modulates activation patterns in the LGN (Ling *et al.*, 2015). These studies suggest that fMRI activation in LGN may reflect feedback modulation from visual cortex or is caused by attention, and LGN and its activation may not play a causal role to build up conscious percepts[10].

3.2.3 Real-time fMRI neurofeedback training approach for neuroscience

---

[9] In addition to magnocellular cells and parvocellular cells, there are "koniocellular cells" between the LGN layers.

[10] Another study to be mentioned here is Lehky, S.R. & Maunsell, J.H. (1996) No binocular rivalry in the LGN of alert macaque monkeys. *Vision research*, **36**, 1225-1234.. This study reported that there was no change in spiking rate depending on task conditions (rivalry/spontaneous perceptual switches or replay/stimulus-driven perceptual switches). Their data suggested that changes in BOLD signal during rivalry task are not necessary to resolve ambiguity in sensory input. One thing to be noted is monkeys did not actively report their perceptual contents; recent fMRI study revealed that active report alters pattern or intensity of BOLD signal in binocular rivalry task Frässle, S., Sommer, J., Jansen, A., Naber, M. & Einhäuser, W. (2014) Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J Neurosci*, **34**, 1738-1747.. Therefore, it is not clear if the lack of changes in LGN activity in Lehky, S.R. & Maunsell, J.H. (1996) No binocular rivalry in the LGN of alert macaque monkeys. *Vision research*, **36**, 1225-1234. supports the idea that LGN does not have functional importance for resolving rivalry, or it was simply due to task condition (i.e. no-report paradigm). The influence of active reports on fMRI signal is discussed in other chapters.

studies

A key problem in resolving this question is there are no good tools in humans to investigate LGN function in a causal manner. Typically studies have observed neural activation that correlates with specific conscious experiences and describe such structures as being neural elements that assist in forming conscious experience. However, this approach cannot conclude what element of neural activation is necessary to generate conscious experience. Invasive approaches to disrupting brain function such as transcranial magnetic stimulation (TMS) allow us to partially overcome this issue, but there are still three difficulties in studying LGN. First of all, TMS stimulation may not reach the LGN. The second is that the effect of TMS stimulation may spread though anatomical or functional connectivity (Ruff *et al.*, 2009). If the effect of stimulation spreads to other brain regions, we cannot draw a clear conclusion regarding the relationship between regional stimulation and behavioral change. The third issue is that TMS does not manipulate brain activity on a voxel-by-voxel basis. There are a number of TMS protocols, but all of them induce or suppress activation in a region or brain network, but not voxels. However, recent advances in fMRI data analysis technique suggest that some aspects of conscious visual perception are represented by the fine scale patterns of activation across multiple voxels, rather than the average or aggregate activity across a region or brain network (Haynes & Rees, 2005; Kamitani & Tong, 2005). Consequently, for a better understanding of the potential causal role of brain activity in conscious visual perception, it is necessary to examine and manipulate brain activation in a voxel-by-voxel fashion.

One promising way to achieve this goal is real-time fMRI neurofeedback training. The advent of real-time fMRI techniques enables the analysis of signals recorded using fMRI on a near-real time basis (Cox *et al.*, 1995). With this approach, a participant in the scanner can learn to modulate their own brain activation by receiving feedback computed from their own recent history of brain activation (Weiskopf *et al.*, 2004) (Figure 3-2). Neurofeedback can improve chronic pain perception when participants

learn to modulate brain activation in areas associated with the processing of pain (deCharms, 2005). In addition, with real-time fMRI neurofeedback training, it is possible to change brain activation (and the sensitivity of retinotopically specific perception) in visual region (Scharnowski *et al.*, 2012) and the voxel-level representation of visual perception (Shibata *et al.*, 2011).

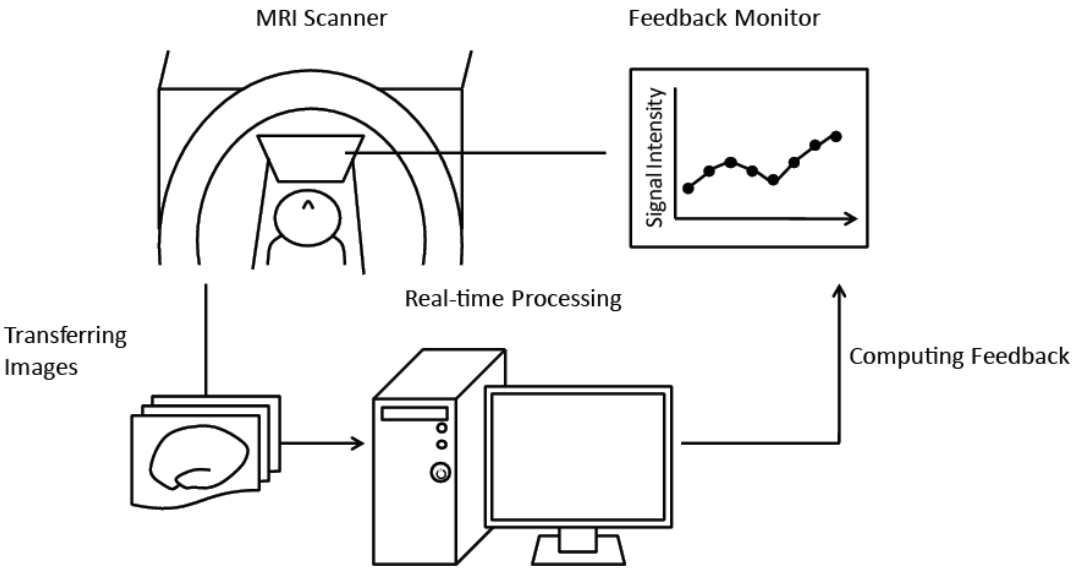

Figure 3-2 Schematic description of real-time fMRI neurofeedback training.

While in an fMRI scanner, a participant is asked to execute an imagery task, such as motor imagery or altering attention. At the same time, with real-time fMRI, Blood-Oxygen Level Dependent signals (BOLD signals) are acquired and analysed very rapidly, and then the participant receives graphical feedback about their neural activity that they can view in the scanner. By viewing the feedback and understanding how it varies with mental effort, the participants can learn to increase or decrease their brain activity in a specific region of interest (ROI) or associated with a particular brain activity pattern.

3.2.4 Hypothesis and technical challenges

To interrogate the role of LGN in binocular rivalry experiments, we could utilize real-time fMRI neurofeedback approach to train participants to modulate their own LGN activity. I planned to perform such a real-time fMRI experiment to investigate whether modulating activation in

eye-specific voxels in LGN would lead to plastic changes in those regions that in turn modulated the dynamics of binocular rivalry. I hypothesized that the training effect (increment of signal in eye-specific voxels in LGN) should prolong dominant the percept duration of image exposed to the trained eye.

An important technical challenge here is how to identify target regions and voxels reliably. Typically, target regions for neurofeedback training are specified at an individual level and participants for neurofeedback training experiments are asked to come to the MRI scanner for training on multiple days (e.g. (deCharms *et al.*, 2005; Scharnowski *et al.*, 2012; Megumi *et al.*, 2015b)). To find a link between brain regions and their possible causal role for cognitive functions, it is crucial to specify the region of interest at an individual level across experimental sessions or days. For the first step of the project, I conducted a pilot study to establish a reliable method to localize bilateral LGN. I followed protocols reported in previous human fMRI studies (Haynes *et al.*, 2005; Wunderlich *et al.*, 2005) and tested if I could replicate the results at an individual level.

3.3 Method

3.3.1 Participant

Ten healthy young adults (8 females, 2 males; right-handed, ages 20 to 25, mean age ± standard deviation, SD: 22.9 ± 1.8 years) participated in this study. All participants had normal or corrected-to-normal vision with contact lenses. Written informed consent was obtained from all participants. The local ethics committee approved procedure of the experiment.

3.3.2 Experimental Design

3.3.2.1 General procedure

Ten participants underwent an LGN localizer task. Participants were

instructed to gaze at the fixation point on the screen. To ensure that participants focused on the task, they were also asked to press a button when the fixation point changed its colour during the localizer tasks.

3.3.2.2 Apparatus

Stimuli were presented on the screen mounted on the MRI head coil using a JVC DLA-SX21 projector. Participants viewed the screen (the screen size was 27cm x 21cm; spatial resolution was 1024 x 768) through a mirror attached to the MRI coil. A viewing distance was approximately 72cm.

3.3.2.3 Stimuli

I used a visual hemifield stimulation protocol (Kastner *et al.*, 2004; Schneider *et al.*, 2004; Haynes *et al.*, 2005; Anderson *et al.*, 2009) to localize right and left LGN. Black and wide full contrast-reversing checkerboard patterns (reversing rate: 0.5 Hz) for left or right visual field were created with Psychtoolbox3 (Brainard, 1997) and presented on gray background. In each visual hemifield, the wedges of the checkerboard extended between 0.6° and 8.0° from a hawk-eye fixation point centred on the screen.

3.3.2.4 Experimental procedure

Each participant underwent 3 runs of the LGN localizer task (Figure 3-3). Each run started with a brief fixation period followed by right visual hemifield stimulation (right-stimulation, RStim, 10 volumes i.e. 21.28 seconds) then left visual hemifield stimulation (left-stimulation, LStim, 10 volumes), Subsequent blocks followed with alternate hemifield stimulation. Each run contained five right/left-stimulation blocks. Between right-stimulation and left-stimulation blocks, a blank screen with a fixation point (fixation period; 5 volumes) was presented.

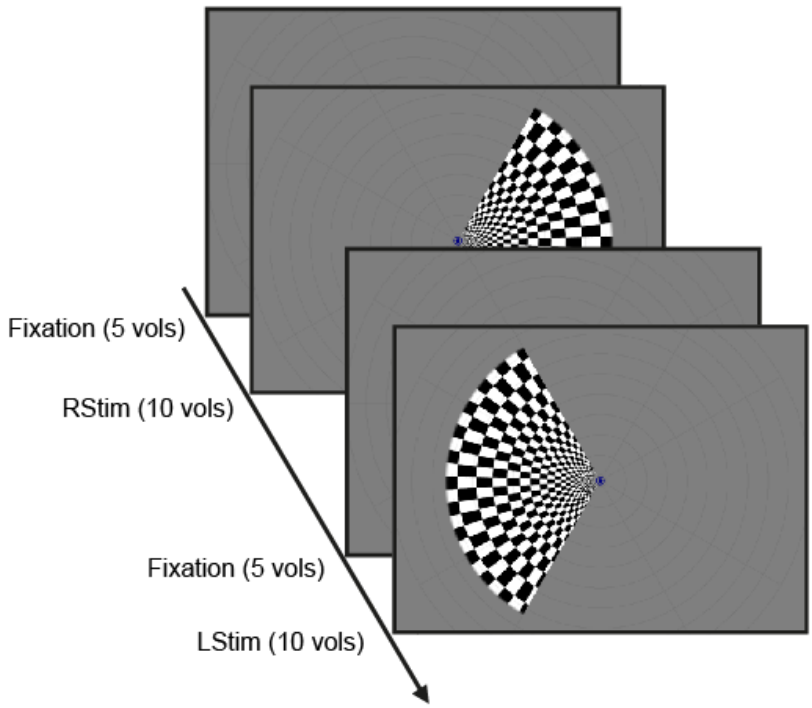

Figure 3-3 Checkerboard stimuli and experimental procedure.

I followed the LGN localization protocol reported in previous studies (e.g. Kastner *et al.* (2004)). A full-contrast checkerboard pattern was presented in either the right or the left visual hemifield for 10 EPI volumes (21.28 seconds). The stimulated visual field alternated every stimulation block.

3.3.3 MRI Data Acquisition

Images were obtained using a 3-T Siemens Magnetom Trio MRI at the Wellcome Trust Centre for Neuroimaging, University College London. MRI data was collected with the fitted 32-channel head coil. Blood-oxygen-level-dependent (BOLD) signals were measured using three-dimensional (3D) high-resolution echo planar imaging (EPI) sequence (slice/volume repetition time, 69 ms / 2.128 seconds; echo time, 34.1 ms; Flip angle, 20°). EPI image contains 24 axial slices (1.5 mm thickness), voxel size was 1.5 mm × 1.5 mm × 1.5 mm, and field of view was 192 mm × 192 mm. T1-weighted structural images were acquired with 1 × 1 × 1 mm. Phase image and magnitude images were also obtained to compute the fieldmap (Hutton *et al.*, 2002).

3.3.4 MRI Data Processing

3.3.4.1 Preprocessing

Statistical Parametric Mapping software (SPM8; Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm) was used for processing MRI data. The EPI images were realigned and unwrapped based on fieldmap images using the FieldMap toolbox in SPM8 (Hutton *et al.*, 2002). Images were normalized with standard SPM normalization module. Then Data were smoothed spatially with a Gaussian kernel of 3 mm full-width at half-maximum and used for statistical analysis.

3.3.4.2 General linear model analysis

Right and left LGN were identified using the general linear model (GLM) approach. First GLM parameters were computed at the individual-level (fixed effects). The design matrix of the model contained three regressors: right-stimulation, left-stimulation, and fixation. Regressors were modelled using a box-car function, which represented the onset and duration of stimulus presentation and fixation period. Six head-motion parameters were included in the GLM model as regressors of no interest to eliminate noise on EPI images due to motion during the scanning. I computed four T-maps representing the statistical contrast of: Right-stimulation > left-stimulation, left-stimulation>right-stimulation, right-stimulation > fixation, and left-stimulation > fixation. These T-maps were used to localize LGN in an individual brain space. Also the contrasts files were used for second level GLM analysis (group-level statistics) and tested by one-sample t-test.

To compare the current results with previous results, I used small volume correction analysis and investigated if the activation in the current study is close to the previous results.

3.4 Results

3.4.1 First level GLM analysis results

I investigated whether LGN could be reliably identified with the real-time optimized EPI sequence. Firstly, I explored if bilateral LGN could be identified at individual GLM activation map.

Figure 3-4, Figure 3-5, and Table 3-1 present right and left LGN identified by LStim > RStim (right LGN) and RStim > LStim (left LGN) contrasts ($p < 0.001$, uncorrected; number of voxels, $k > 10$). Table 3epicts the results of small volume correction (SVC) for GLM. Five out of ten participants showed some amount of activation in right LGN and four out of ten showed activation in left LGN. Three participants showed significant activation in bilateral activation in LGN localizer task.

I tried different threshold methodology to explore better way to localize LGN at individual level. Figure 3-6, Figure 3-7, and Table 3-2 show right and left LGN with different contrasts. In order to identify the right and left LGN, voxels within an anatomically defined region of the posterior thalamus showing greater activation for contralateral visual field stimulation compared to the ipsilateral visual field stimulation (p=0.05, cluster threshold 20 voxels) were masked inclusively with those voxels showing greater activation for contralateral visual field stimulation compared to fixation period (p=0.05, cluster threshold > 20 voxels).

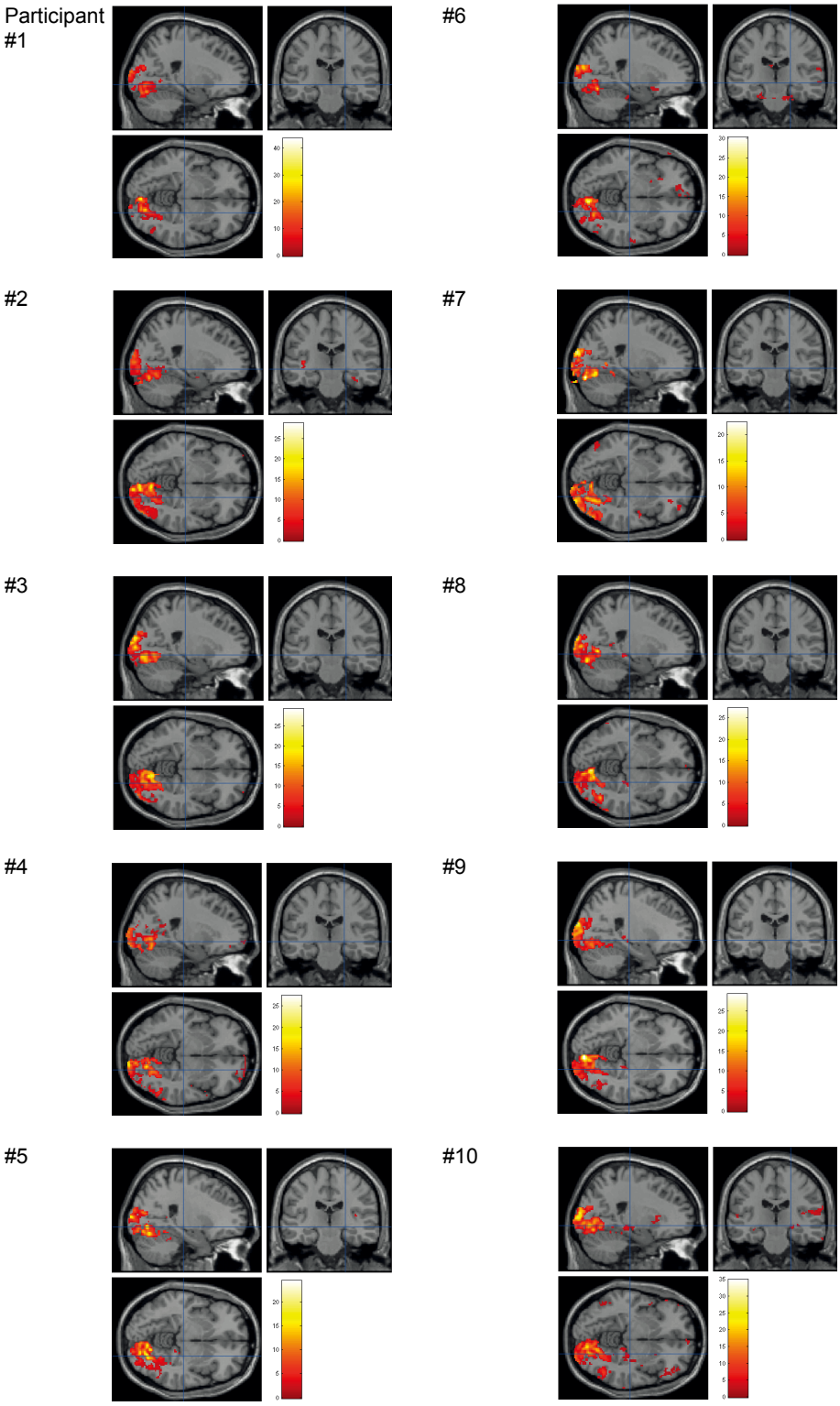

Figure 3-4 Activation in the right LGN (all participants)

5 out of 10 participants showed activation in right LGN (RStim > L Stim contrast, overlaid on the standard MNI template). The color scale represents t-value. Thresholded at $p < 0.001$, $k > 20$. The crosshair indicates centre of right LGN (anatomical) reported in O'Connor *et al.* (2002).

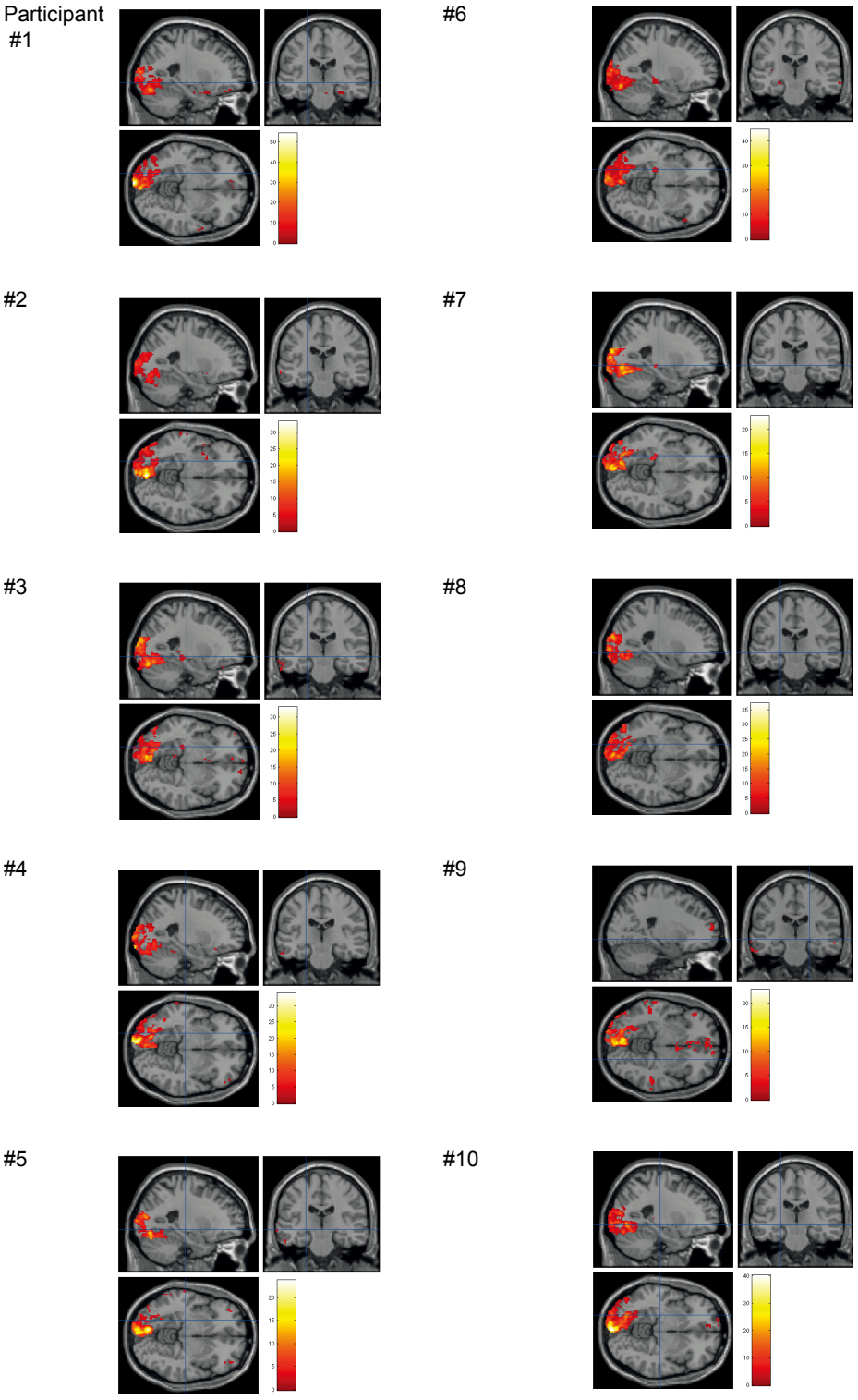

Figure 3-5 Activation in the left LGN (all participants).

4 out of 10 participants showed activation in right LGN (RStim > L Stim contrast, overlaid on the standard MNI template). The color scale represents t-value. Thresholded at $p < 0.001$, $k > 20$. The crosshair indicates centre of right LGN (anatomical) reported in O'Connor *et al.* (2002).

Table 3-1 Small volume correction results for right and left LGN (LStim > RStim, RStim > LStim, $p < 0.001$, $k > 20$; reporting the peak of the cluster)

Right LGN								
Participant #	*p (FWE-corr)*	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
#1	< 0.001	5.94	5.82	< 0.001	22	-28	-6	37
#3	< 0.002	5.28	5.20	< 0.001	20	-30	-6	28
#7	< 0.003	5.51	5.41	< 0.001	20	-28	-8	42
#8	< 0.004	5.01	4.93	< 0.001	22	-30	-8	14
#10	0.001	4.59	4.53	< 0.001	22	-28	-8	33
Left LGN								
	p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
#1	< 0.001	5.15	5.07	< 0.001	-20	-30	-6	17
#3	< 0.001	4.83	4.76	< 0.001	-22	-28	-8	29
#6	< 0.001	6.02	5.89	< 0.001	-20	-28	-8	73
#7	< 0.001	7.32	7.09	< 0.001	-20	-28	-8	38

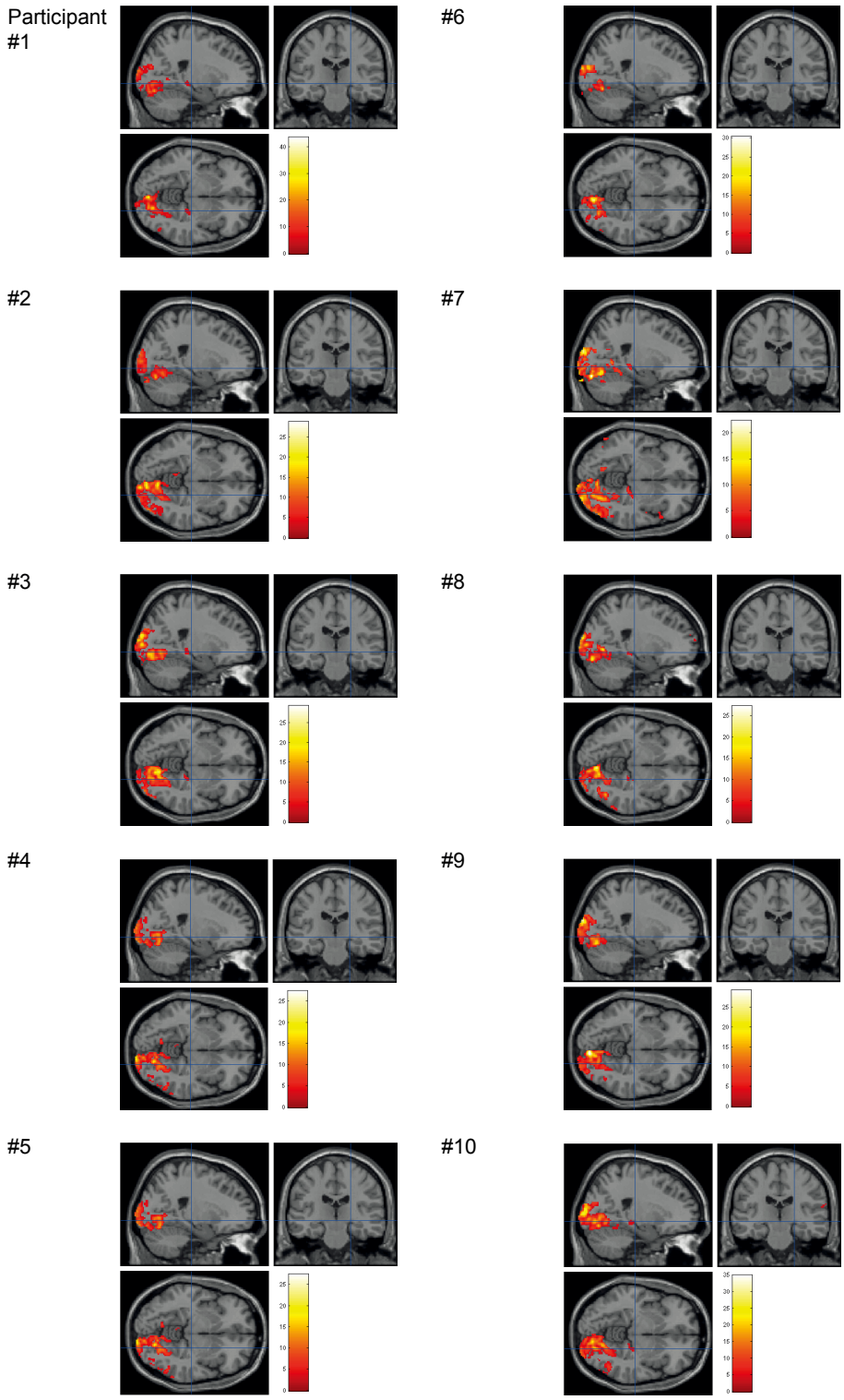

Figure 3-6 Activation in right LGN (all participants)

4 out of 10 participants showed activation in right LGN in RStim > LStim ($p < 0.05$) masked with RStim > Fixation ($p < 0.05$) contrast ($k > 20$)., The individual-level contrast map was overlaid on the standard MNI template. The color scale represents t-value. The crosshair indicates centre of right LGN (anatomical) reported in O'Connor *et al.* (2002).

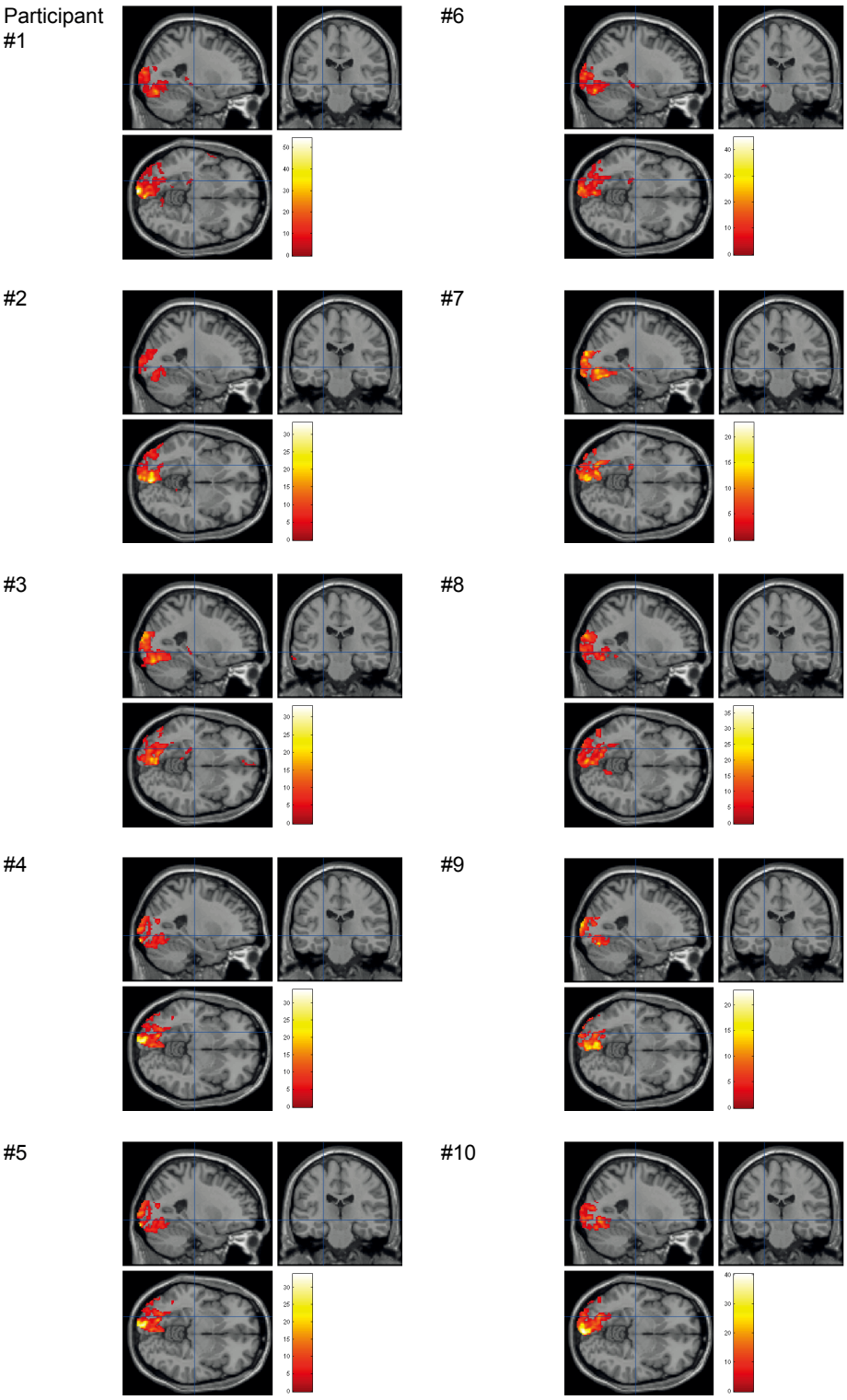

Figure 3-7 Activation in left LGN (all participants)

3 out of 10 participants showed activation in left LGN in RStim > LStim ($p < 0.05$) masked with RStim > Fixation ($p < 0.05$) contrast ($k > 20$). The individual-level contrast map was overlaid on the standard MNI template. The color scale represents t-value. The crosshair indicates centre of right LGN (anatomical) reported in O'Connor *et al.* (2002).

Table 3-2 Small volume correction results for right and left LGN (LStim > RStim masked with LStim > Fixation, RStim > LStim masked with RStim > Fixation, $p < 0.05$, $k > 20$; reporting the peak of the cluster)

Right LGN								
	p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
#5	< 0.001	4.89	4.82	< 0.001	20	-28	-8	12
#8	< 0.001	5.01	4.93	< 0.001	22	-30	-8	22
#9	0.057	3.63	3.60	< 0.001	24	-28	0	22
#10	0.873	2.04	2.04	0.021	-30	-20	-10	5
Right LGN								
	p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
#3	< 0.001	4.83	4.76	< 0.001	-22	-28	-8	48
#6	< 0.001	6.02	5.89	< 0.001	-20	-28	-8	72
#7	< 0.001	7.32	7.09	< 0.001	-20	-28	-8	33

3.4.2 Second level GLM analysis results

Finally, second-level GLM activation maps were computed based on the individual –level activation maps shown in Figure 3-8 and Figure 3-9 (peak coordinate and more details are shown in Table 3-3). Although I could not find consistent activation patterns in the LGN at an individual GLM level, bilateral activation in LGN was found in the second-level analysis, as reported in previous studies (O'Connor *et al.*, 2002; Kastner *et al.*, 2004; Haynes *et al.*, 2005). To compare the current second level GLM results with those previous reports (Kastner *et al.*, 2004), small volume correction was performed (Table 3-4). The analysis showed the significant (corrected for multiple comparisons) activation in both right and left LGN were close to the results reported in previous studies (coordinate from (Kastner *et al.*, 2004); right LGN, 10mm sphere centred at (21, -19, -5), $t = 11.95$, $p < 0.001$; left LGN, 10mm sphere centred at (-22, -21, -4), $t = 7.31$, $p < 0.001$).

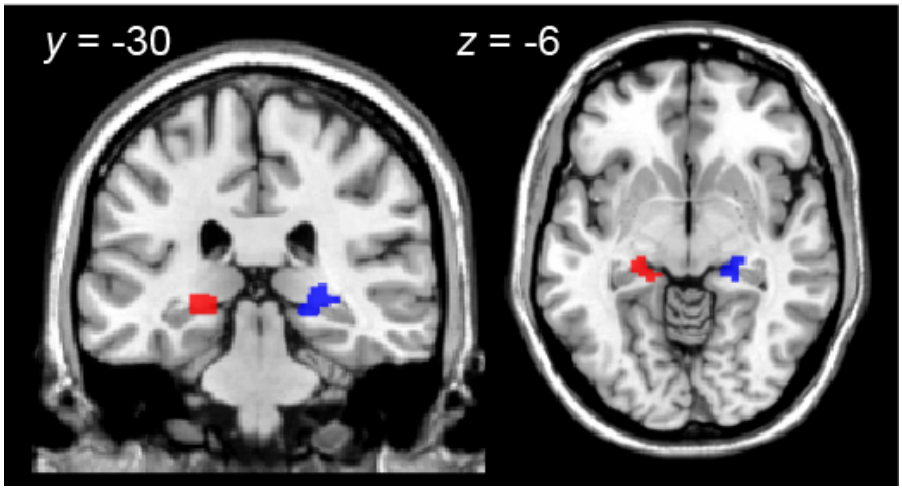

Figure 3-8 Bilateral LGN.

The right LGN (red, LStim > RStim contrast) and the left LGN (blue, RStim > LStim, contrast) was identified in second level GLM statistics (contrast images were thresholded at $p < 0.001$, uncorrected; $k > 20$). The picture was made with MRICron (http://www.mccauslandcenter.sc.edu/mricro/mricron/).

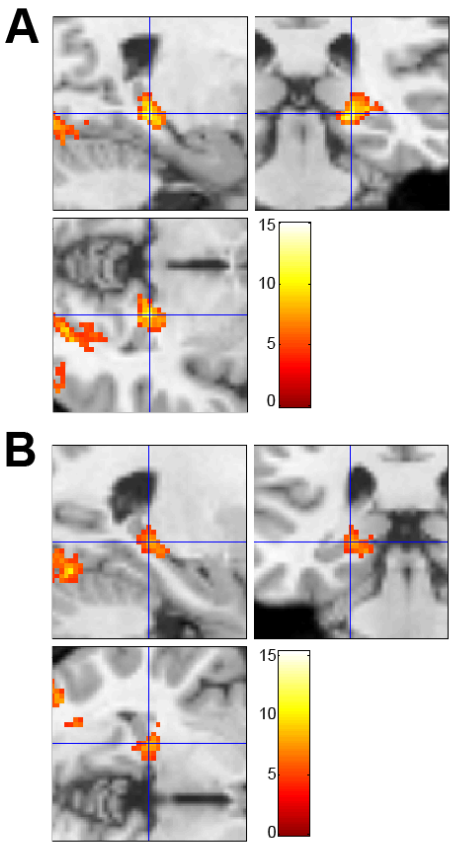

Figure 3-9 The second level analysis: Activation in (A) right and (B) left LGN.

Activation in the bilateral LGN identified by second level GLM analysis. The crosshair indicates peak voxels in the LGN clusters (A: (22, -30, -4), B: (-22,

-32, -2)). One side of the square image is equivalent to 80mm. The color map indicates t-value of the contrast.

Table 3-3 Peak of the right and left LGN cluster at second level

Right LGN							
p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
0.027	11.95	4.94	< 0.001	22	-30	-4	160
Left LGN							
p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
0.673	8.13	4.27	< 0.001	-22	-32	-2	106

Table 3-4 Small volume correction results for right and left LGN at second level

Right LGN							
p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
< 0.001	11.95	4.94	< 0.001	22	-30	-4	83
Left LGN							
p (FWE-corr)	*T*	*Z*	*p (unc)*	x	y	z	Number of voxels
0.006	7.9	4.22	< 0.001	-22	-30	-2	49

3.5 Discussion

3.5.1 Overview of the findings

To achieve successful neurofeedback training on the human LGN, it is necessary to establish a protocol to identify bilateral LGN and eye-specific voxels across training days and sessions in an individual. As a first step, I aimed to establish a method to localize LGN functionally. Following the methods used in previous studies (O'Connor *et al.*, 2002; Haynes *et al.*, 2005), however, I could only find significant bilateral LGN activation in 4 out of 12 participants (33 %) at the first level. However, right and left LGN were successfully identified in a second-level analysis and the MNI coordinates were similar to those reported in previous studies.

3.5.2 Comparison of current results with previous reports

Although right and left LGN were localized in the second level analysis with the location of peak activation being similar to those reported in previous fMRI studies (O'Connor *et al.*, 2002; Haynes *et al.*, 2005), localizing LGN at the first level of analysis wasn't successful.

One possible reason for the failure of LGN localization at the first level is the lack of data. LGN is a small region and located in a subcortical area and these make it difficult to localize LGN. (O'Connor *et al.*, 2002) and (Haynes *et al.*, 2005) used a 3T head scanner (Allegra, Siemens) and the localizer tasks take approximately 10-20 minutes (5 minutes runs, 2-4 runs for each participants). Although I followed this earlier procedure, the data were not sufficient to yield reliable LGN activation at the individual level.

A notable difference between the current experiment and previous experiment is the particular nature of the EPI sequence. Haynes et al. (2005) and O'Connor et al. (2005) used a single-shot gradient echo. On the other hand, I used multi-shot EPI sequence and thus benefited from its high spatial resolution. However, it is known that images acquired with a multi-shot EPI sequence tend to show higher signal-to-noise ratio and physiological noise can cause serious artefact in such data (Lutti *et al.*, 2013). This could therefore be the cause of the difference in the results. The artefact problem can be avoided by using pulse and respiration data in the preprocessing stage (see (Lutti *et al.*, 2013)).

3.5.3 Alternative ways to identify LGN at individual level

If the functional localizer paradigm does not work well at the individual level, what could be an alternative way to localize LGN at the individual level?

One solution would be using structural/anatomical data or combine them with functional data. The structure of LGN can be distinguished from neighboring regions. (Ling *et al.*, 2015) using proton-density weighted images to anatomically localize LGN.

Another possible solution is using diffusion tractography-based segmentation. Diffusion tensor imaging (DTI) enables us to explore structural connectivity. The anatomy of LGN and its connections are well-known - most of LGN should have anatomical connections to early visual cortex and the connectivity pattern should be different from neighbouring regions such as pulvinar. This possibly means that structural connectivity could dissociate LGN from other regions using DTI data. Detecting anatomical regions using DTI is proven to be successful with subcortical regions (thalamus, (Johansen-Berg *et al.*, 2005)), and similar analysis should be useful for LGN localization as well.

3.6 Conclusion

As a precursor to real-time fMRI neurofeedback training for eye-specific voxels in LGN, I sought to develop a method to locate right and left LGN cluster at individual level. Although 20 minutes LGN localizer scanning could detect bilateral LGN activation at second level, only a few participants showed reliable right and left LGN activation at an individual level. Possible solutions include doubling the duration of the localizer task or using anatomical structure to locate LGN.

Chapter 4 BRAIN ACTIVITY DYNAMICS IN PARIETAL REGIONS DURING SPONTANEOUS SWITCHES IN BISTABLE PERCEPTION

4.1 Summary of this chapter

The neural mechanisms underlying conscious visual perception have been extensively investigated using bistable perception paradigms. Previous functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) studies suggest that the right anterior superior parietal (r-aSPL) and the right posterior superior parietal lobule (r-pSPL) have opposite roles in triggering perceptual reversals. It has been proposed that these two areas are part of a hierarchical network whose dynamics determine perceptual switches. However, how these two parietal regions interact with each other and with the rest of the brain during bistable perception is not known. In this chapter, I investigated such a model by recording brain activity using fMRI while participants viewed a bistable structure-from-motion stimulus. Using dynamic causal modeling (DCM), I found that resolving such perceptual ambiguity was specifically associated with reciprocal interactions between these parietal regions and V5/MT. Strikingly, the strength of bottom-up coupling between V5/MT to r-pSPL and from r-pSPL to r-aSPL predicted individual mean dominance duration. my findings are consistent with a hierarchical predictive coding model of parietal involvement in bistable perception and suggest that visual information processing underlying spontaneous perceptual switches can be described as changes in connectivity strength between parietal and visual cortical regions.

4.2 Introduction

4.2.1 Function of right superior parietal lobule in bistable perception

Visual bistable perception stimuli induce different and spontaneously varying percepts while visual information projected on the retina remains unchanged. Functional magnetic resonance imaging (fMRI) studies suggest that human fronto-parietal brain regions may play a critical role in resolving such ambiguity in visual information and forming a unitary conscious percept (Kleinschmidt *et al.*, 1998; Lumer *et al.*, 1998; Sterzer & Kleinschmidt, 2007). In addition, as explained in the previous chapter, transcranial magnetic stimulation (TMS) of the human parietal cortex demonstrates the causal involvement of distinct parietal regions in perceptual changes during bistable perception (Carmel *et al.*, 2010b; Kanai *et al.*, 2010; Zaretskaya *et al.*, 2010; Kanai *et al.*, 2011). Specifically, stimulation of the right anterior superior parietal (r-aSPL) and the right posterior superior parietal lobule (r-pSPL) have led to opposite effects on perceptual reversals (Kanai *et al.*, 2011) leading to the suggestion that these two areas may be parts of a hierarchical network whose dynamics play a causal role in perceptual switches in bistable perception.

4.2.2 Hypothesis, experimental paradigm, and analysis strategy

To test this hypothesis, I used functional magnetic resonance imaging (fMRI) to record brain activity, while participants viewed a structure-from-motion stimulus (see Figure 4-1), which leads to spontaneous alternations between two exclusive perceptual states (sphere rotating toward or away from the viewer). I applied dynamic causal modeling (DCM) analyses to test a specific model of effective connectivity proposed previously (Kanai et al 2011). The advantage of using DCM is that we can express changes in brain dynamics associated with an experimental condition and directly compare the quantitative agreement between

competing models and empirically observed Blood Oxygenation Level Dependent (BOLD) dynamics.

Based on a previous study (Kanai et al 2011), I hypothesized that the anterior and posterior subregions of the right superior parietal lobule (r-aSPL and r-pSPL, respectively) and the motion sensitive visual area V5/MT form a hierarchical network structure with area V5/MT at the bottom and aSPL at the top level of the hierarchy. I predicted that reciprocal connections between them should mediate their dynamical interactions during perceptual rivalry and that the strength of the dynamical modulations of these connections should correlate with individual differences observed in participants' behavior in bistable perception. I functionally identified the three regions of interest (r-aSPL, r-pSPL, and, and right V5/MT, r-V5) using standard fMRI analysis approaches and then asked which DCM model structure and dynamics best explained information flow among these three regions and whether brain dynamics represented as parameters in the DCM model predicted the inter-individual variance in percept dominance duration.

Regions from the right hemisphere, not the left hemisphere, were included to the DCM models. This is because the DCM model space was motivated by previous VBM and TMS studies, and they focused on the right hemisphere as well.

4.3 Method

4.3.1 Participants

Eighteen healthy participants (10 females, right-handed, ages 18 to 39, mean age ± standard deviation, SD: 26.0 ± 6.2 years) participated in this study. All participants had normal or corrected-to-normal vision with contact lenses. Written informed consent was obtained from all participants. The local ethics committee approved the experiments.

4.3.2 Experimental Design

4.3.2.1 General procedure

I used structure-from-motion (SFM) stimuli (see Figure 4-1) and recorded participants' behavioural reports of spontaneous fluctuations in bistable perception (rivalry condition) plus stimulus-driven changes (replay condition). Participants were instructed to look at the screen through prism glasses (Schurger, 2009) and report their subjective percept (the direction of rotation of the sphere) by holding one of three buttons; one for each of the two rotation directions, and one for mixture of two percepts or when the direction of rotation was unclear).

Prior to the fMRI session, participants underwent a short behavioral testing period outside the scanner to ensure that they could achieve stereopsis with my experimental setup and that their percept durations were in a suitable range (3 to 10 seconds) for the fMRI experiment.

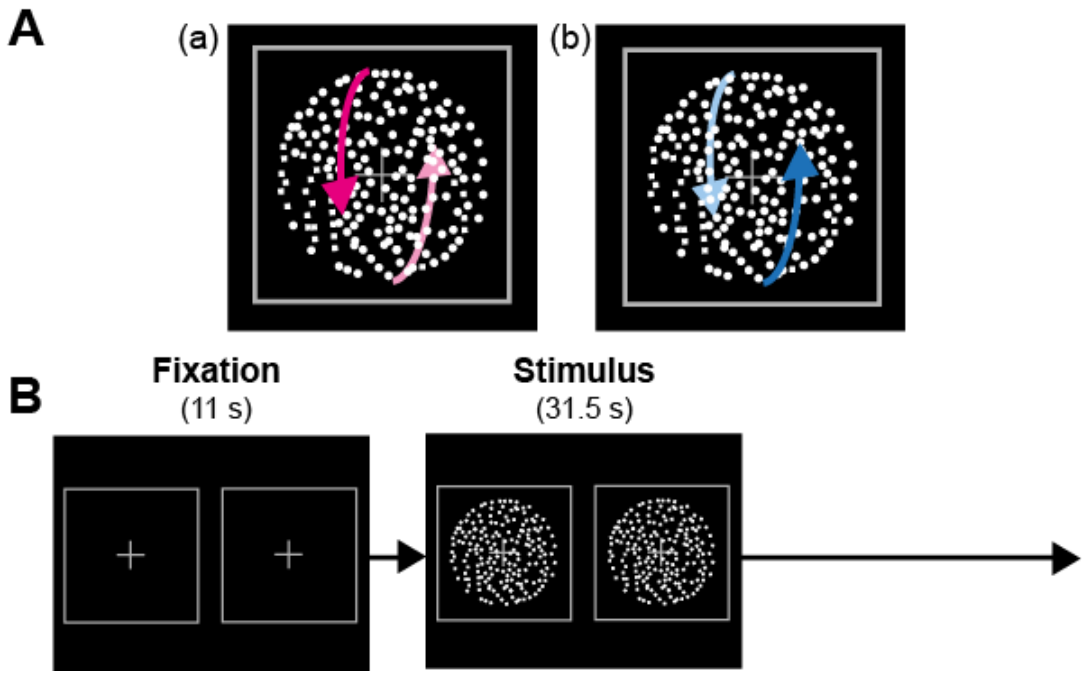

Figure 4-1 Illustration of the stimuli and the experimental procedure.

(A) my structure-from-motion (SFM) stimuli, which consisting of moving white dots, typically cause two exclusive alternating percepts: a sphere rotating either toward (a) or away (b) from the viewer. Note that the size of the white dots is magnified in this figure for visualization. (B) In fMRI

session, SFM stimuli were presented on the screen for 30.5 seconds (15 EPI volumes). Participants were asked to report their percept by pressing or holding one of three buttons (toward, away, or not sure/mixture) during stimulus presentation. Note that stimuli were presented dichoptically in order to add disparity information for the replay condition: the participants used prism glasses and the screen was split by a black cardboard divider to aid fusion and ensures monocular presentation of each image.

4.3.2.2 Apparatus

Stimuli were presented on the screen mounted on the MRI head coil using a JVC DLA-SX21 projector. Participants viewed the screen (the screen size was 27cm x 21cm; spatial resolution was 1024 x 768) through a mirror attached to the MRI coil. A viewing distance was approximately 72cm. For dichoptic stimulus presentation, participants used prism glasses (lenses with 4 prism dioptres base out) and a black cardboard partition was attached to the head coil to divide the screen and the mirror into two areas for separate presentations to the left and right eye.

4.3.2.3 Stimuli

For the rivalry condition, a vertically spinning sphere (3.1° diameter) comprising 200 full-contrast white dots was presented to each eye for a structure from motion task (Kanai *et al.*, 2010). Spheres were created using PsychToolbox 3 under MATLAB (The Mathworks, Inc.) and they were presented against black background. The white dots moved sinusoidally upwards and downwards at an angular velocity of 120 degree/s. A fixation cross (0.1° in height and width) was superimposed at the center of each sphere. The spheres were surrounded by a square frame to help participants to maintain stable vergence and were presented at the same position relative to the fixation points to ensure that direction of spin was ambiguous in the rivalry condition.

For the replay condition, binocular disparity was computed for each dot so that stimuli were embedded with unambiguous disparity cues and

participants could perceive stereoscopic depth without difficulty. Fixation points, the spheres, and squares were aligned to the center of the illusory 3D spheres. Unlike binocular rivalry (Knapen *et al.*, 2011) or the Lissajous figure (Weilnhammer *et al.*, 2013), SFM typically does not induce a high proportion of mixed percepts. I confirmed for my stimulus configuration that the total duration of mixture of two alternative percepts was indeed very short (2.23% of the total duration of stimulus presentation per MRI run). I therefore focused on perceptual switches between two alternative percepts for the analysis.

4.3.2.4 Experimental procedure

On each trial, the ambiguous rotating sphere was presented continuously for 31.5 seconds (15 EPI volumes) followed by a fixation period (11 seconds, 5 EPI volumes).

Each MRI run consisted of 10 trials (five trials for rivalry condition and replay condition respectively), started with rivalry trial, and the order of rivalry and replay trials was pseudo-randomised. In a subset of the subsequent trials of the same run, the percept reported during the rivalry condition was replayed. The order of rivalry and replay trials was randomized across runs and participants. Participants performed the task for 4 to 7 runs in the MRI scanner (Mean ± SD: 6.4 ± 0.9). Mean dominance duration during rivalry condition and replay rate (percentage of correct button response to disambiguated sphere's spin on the screen, judged at each screen frame) was computed from MRI-compatible button response.

4.3.3 MRI Data Acquisition

Images were obtained using a 3T Siemens Magnetom Trio MRI at the Wellcome Trust Centre for Neuroimaging at University College London. MRI data were collected with the fitted 32-channel head coil. Blood Oxygen Level Dependent (BOLD) signals were measured using an echo planar imaging (EPI) sequence (volume repetition time, 2.1 s; echo time, 30 ms; Flip

angle, 90°). EPI image contained 30 axial slices (3 mm thickness, ascending slice order), voxel size was 3 mm × 3 mm × 3 mm, and the field of view was 192 mm × 192 mm. T1-weighted structural images were acquired with 1 × 1 × 1 mm. Phase image and magnitude images were also obtained to compute a fieldmap (Hutton *et al.*, 2002).

4.3.4 MRI Data Processing

4.3.4.1 Preprocessing

Statistical Parametric Mapping software (SPM8 and SPM12: Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm) was used to process MRI data. The first five EPI volumes were discarded to allow for T1 equilibration. For preprocessing the EPI images, first, the EPI images were then realigned and unwrapped based on fieldmap images using the FieldMap toolbox in SPM8 (Hutton *et al.*, 2002). EPI images were spatially normalized to the Montreal Neurological Institute (MNI) stereotactic template. Data were smoothed spatially with a Gaussian kernel of 8 mm full-width at half-maximum.

4.3.4.2 General linear model analysis

Statistical parametric mapping analysis was performed using the general linear model (GLM) approach. As a first step, GLM parameters were computed at the individual-level (fixed effects). The design matrix of the model contained four regressors: fixation, visual stimulation, spontaneous perceptual switch (rivalry-switch), and stimulus-driven perceptual switch (replay-switch). Visual stimulation and fixation periods were modeled using a box-car function, which represented the onset and duration of stimulus presentation and fixation period. Rivalry-switch and replay-switch were modeled with an impulse function. All regressors were then convolved with a canonical hemodynamic response function implemented in SPM8. In order to estimate actual timing of switch events, reaction time to press a button

(mean reaction time across participants ± SD: 0.90 ±0.43 seconds) was computed as the interval between replay stimulus change and participant's button press. The estimated average reaction time was subtracted from the time of button presses to model the actual timing of rivalry-switch and replay-switch events. Six head-motion parameters were also included in the GLM model as regressors of no interest to model and eliminate any noise on EPI images due to motion during the scanning.

4.3.4.3 Additional GLM analysis

In addition, I further analyzed my fMRI data using an additional GLM (alternative GLM) to confirm that activation in r-aSPL and r-pSPL associated with the comparison of rivalry-switch and replay-switch did not merely reflect minor differences in stimuli between the conditions (i.e. presence of binocular disparity). I employed the same preprocessing procedure as the original GLM but the design matrix of the new GLM model now comprised five main regressors: fixation, spontaneous perceptual switch (rivalry-switch), stimulus-driven perceptual switch (replay-switch), stable percept in rivalry block (rivalry-stable, i.e. stimulus presentation without perceptual switch in rivalry block), and stable percept in replay block (replay-stable, stimulus presentation without perceptual switch in replay block). Fixation periods, rivalry-stable, replay-stable were modeled using a box-car function, which represented the onset and duration of fixation period and stable percept. Rivalry-switch and replay-switch were modeled with an impulse function; the estimated average reaction time was subtracted from the time of button presses to model the actual timing of rivalry-switch and replay-switch. Six head motion parameters derived from the preprocessing were also included in this GLM to regress out nuisance signal due to head movement.

4.3.5 Dynamic Causal Modeling Analysis

Dynamic causal modeling (Friston *et al.*, 2003b) was performed using DCM12 in SPM12 (Wellcome Trust Centre for Neuroimaging). The DCM analysis estimates neural dynamics with A-matrix (endogenous connectivity, fixed across experimental conditions), B-matrix (modulatory parameters), and C-matrix (driving force).

In my study, I was particularly interested in investigating B-matrix parameters during rivalry-switch and replay-switch events and relate them to individual differences in behavioral data. I hypothesized that r-aSPL, r-pSPL, and r-V5 constitute a three-layer hierarchical model with reciprocal interactions between areas during spontaneous perceptual transitions (Kanai *et al.*, 2011) and tested this hypothesis by estimating these coupling parameters (B-matrix) in DCM. The three ROIs were selected based on univariate fMRI results: r-aSPL, r-pSPL, and r-V5. ROIs for DCM analysis were defined by the following procedure. First, peak voxel coordinates were found for each ROI based on anatomically defined ROIs: 10mm radius sphere centered (x, y, z) = (36, -45, 51) for r-aSPL (Carmel *et al.*, 2010b), 10mm radius sphere centered (38, -64, 32) for r-pSPL (Kanai *et al.*, 2011), 10mm radius sphere centered (44, -67, 0) for V5/MT (Dumoulin *et al.*, 2000; Mars *et al.*, 2011). Then the 10mm-sphere masks centered on the peak voxels were created with PickAtlas (Maldjian *et al.*, 2003) and applied to group level fMRI results (Rivalry-switch > Replay-switch contrast, thresholded at $p < 0.001$, uncorrected) to create DCM ROIs (See Figure 3). Averaged BOLD signals in each region were extracted using the *Volume of Interest* module in SPM8 and used for DCM analysis.

DCM analysis was performed in two steps. First, I explored the optimal model structure that best described neuronal responses using Bayesian model selection (DCM model selection). Four conditions (fixation, visual stimulation, rivalry-switch, and replay-switch) in the GLM model were included in DCM models. A previous functional connectivity study has shown that the posterior part of parietal lobule is specifically coupled to V5/MT in the resting state (Mars *et al.*, 2011) and therefore my DCM models specifically posited endogenous connectivity between r-pSPL and r-V5 but not r-aSPL and r-V5. It is recommended to utilize such prior knowledge

about brain connectivity when defining DCM model space (Stephan *et al.*, 2010) and both monkey electrophysiology (Vincent *et al.*, 2007) and imaging (Greicius *et al.*, 2009) studies suggest that functional connectivity reflects anatomical connectivity. Direct input to r-V5 (C-matrix) was explicitly modeled during visual stimulation, rivalry-switch, and replay-switch events but not during the fixation period. Human area V5/MT is involved in visual motion processing (Tootell *et al.*, 1995) and a previous structure-from-motion study also detected activation in V5/MT associated with subjective perceptual switches (Freeman *et al.*, 2012), suggesting V5/MT is involved in perceptual switches even when the stimulus remains unchanged (albeit with a repetitively fluctuating retinal input due to the sinusoidal motion of individual dots). In addition, theoretical work proposes that adaptation of neurons in visual cortex may be the driving force for perceptual switches in bistable perception (Dayan, 1998). These studies indicate that direct input to r-V5 should be included in my modeling of perceptual switches. I therefore modeled modulatory effects on four connections between ROIs (from r-aSPL to r-pSPL, from r-pSPL to r-aSPL, from r-pSPL to r-V5, and from r-V5 to r-pSPL – see Figure 3 for ROI positions) and two driving inputs (r-aSPL, r-pSPL) in rivalry-switch and replay-switch events and therefore the total number of tested models was 64 (6 dimensions, 2^4 for B-matrix × 2^2 for C-matrix). I then divided all the models into four groups based on modulatory parameters: no modulation, top-down, bottom-up, and bidirectional (See figure 4). The exceedance probability of each family and model was computed with random-effect assumptions using a Bayesian approach and the best-fit model in the winning group was selected as the wining model (Penny *et al.*, 2004; Penny *et al.*, 2010). The exceedance probability represents the probability that a model or model family is more likely than the other models or families (Stephan *et al.*, 2009; Penny *et al.*, 2010)[11]. Note that it was assumed that

[11] It has been pointed out that exceedance probability may not be as good as Bayesian omnibus risk, which is a normalized version of Bayes factor (Rigoux et al., 2014). The exceedance probability rests on given dataset. Bayes factor is a comparison between two

the optimal model structure was common between rivalry-switch and replay-switch condition and this assumption enabled us to handle DCM parameters quantitatively across the two conditions and relate them to behavioral performance.

To validate the DCM result, two additional DCM analyses were performed. First, I tested a different DCM ROI selection approach. Previous network connectivity studies show that inappropriate signal in ROIs may lead to failure to detect network connectivity (Smith *et al.*, 2011) and therefore I accounted for individual differences in the ROI peak location across participants by instead selecting voxels for ROIs using the individual first level GLM activation map. DCM ROIs were created as follows: 10mm spherical ROI masks centered on the original second level GLM were created and applied to each individual's GLM activation map for the contrast Rivalry > Replay. Voxels that exceeded a liberal threshold of $P < 0.3$ (uncorrected) within that sphere were used collectively as ROIs for DCM analysis. Note that two participants out of eighteen were excluded from this analysis due to the absence of any activated voxels at this threshold inside the 10mm ROI spheres.

I also performed family-level DCM model comparison with the alternative GLM to explore whether differences in regressors would affect DCM family level inference. I defined DCM ROIs with a similar procedure to that used for the main DCM analysis: 10mm-sphere masks centered on the group level peak voxel coordinates ((34, -52, 50) for r-aSPL, (30, -70, 32) for r-pSPL, and (44, -74, 6) for r-V5) were applied to group level fMRI results (Rivalry-switch > Replay-switch contrast, thresholded at $p < 0.008$, uncorrected) and used as DCM ROIs.

4.3.6 Individual Difference Analysis

---

models and there is no straight-foward way to find out 'the best' model among more than three models. Rigoux et al., (2014) propose 'protected exceedance probability', which considers the variability of model evidences when computing the goodness of the model.

To investigate whether the DCM parameters could explain the variations in individual behavioral differences between participants, a multiple regression analysis was performed using SPSS software (International Business Machines Corporation, New York). I tested the hypothesis that the variability in network dynamics is related to inter-individual variability in behavior. I asked whether the difference in the strength of modulatory effect (B-parameters) for rivalry-switch and replay-switch conditions were predictive of the behavioral variability across participants. Thus, I chose the difference between B-parameters of the winning model (i.e. rivalry – reply) as regressors and tested if those regressors could predict an individual's mean dominance duration.

To ensure a good fit to regression model and remove outlier effects, three participants whose Cook's distance was larger than one were eliminated from the analysis. Cook's distance is a measure to detect influential data points in regression analysis and used for detecting outliers (Cook, 1977). I report R^2 and corrected R^2 value (adjusted for degree of freedom to account for number of repressors).

4.4 Results

4.4.1 Behavioral data analysis

The average perceptual dominance duration across the two fluctuating bistable percepts was 4.50 seconds (SD: 0.99). In the replay condition, mean dominance duration was 4.48 seconds (SD: 1.00) and there was no significant difference in mean dominance duration of the two conditions ($t(17) = 0.69$, $p = 0.50$, n.s.), suggesting that participants successfully reported replay based on the depth information added to the stimuli.

4.4.2 GLM analysis

To investigate which brain regions showed activation associated with perceptual transitions, I constructed a general linear model (GLM) that

included stimulus presentation, fixation, spontaneous perceptual change (rivalry-switch), and stimulus-driven perceptual change (replay-switch) as regressors. Figure 4-2A shows the brain activations correlated with rivalry-switch compared to replay-switch (see Table 4-1). Given my prior hypothesis (Lumer *et al.*, 1998; Carmel *et al.*, 2010b; Kanai *et al.*, 2010) I used small volume correction (SVC) and validated that the activation evoked during rivalry-switch in r-aSPL (sphere radius = 10mm, sphere center, (36, -45, 51); peak voxel, (32, -48. 48), $t(17) = 5.38$, $p = 0.002$, $p < 0.01$, corrected for small-volume) and r-pSPL (sphere radius = 10mm, sphere center, (38, -64. 32); peak voxel, (30, -70. 32), $t(17) = 4.47$, $p = 0.01$, $p < 0.01$, corrected) (see Figure 4-2B). Moreover, motion-sensitive visual area V5/MT in the right hemisphere also showed greater activation associated with rivalry versus replay switches (sphere radius = 10mm, sphere center, (44, -67, 0); peak voxel, (48, -62, -8), $t(17) = 4.82$, $p = 0.005$, $p < 0.01$, corrected). In addition, I also observed activation evoked by spontaneous perceptual switches in frontal cortex, visual cortex, insula, and middle frontal gyrus as reported in previous studies of bistable perception (Kleinschmidt *et al.*, 1998; Sterzer & Kleinschmidt, 2007; Zaretskaya *et al.*, 2010; Knapen *et al.*, 2011).

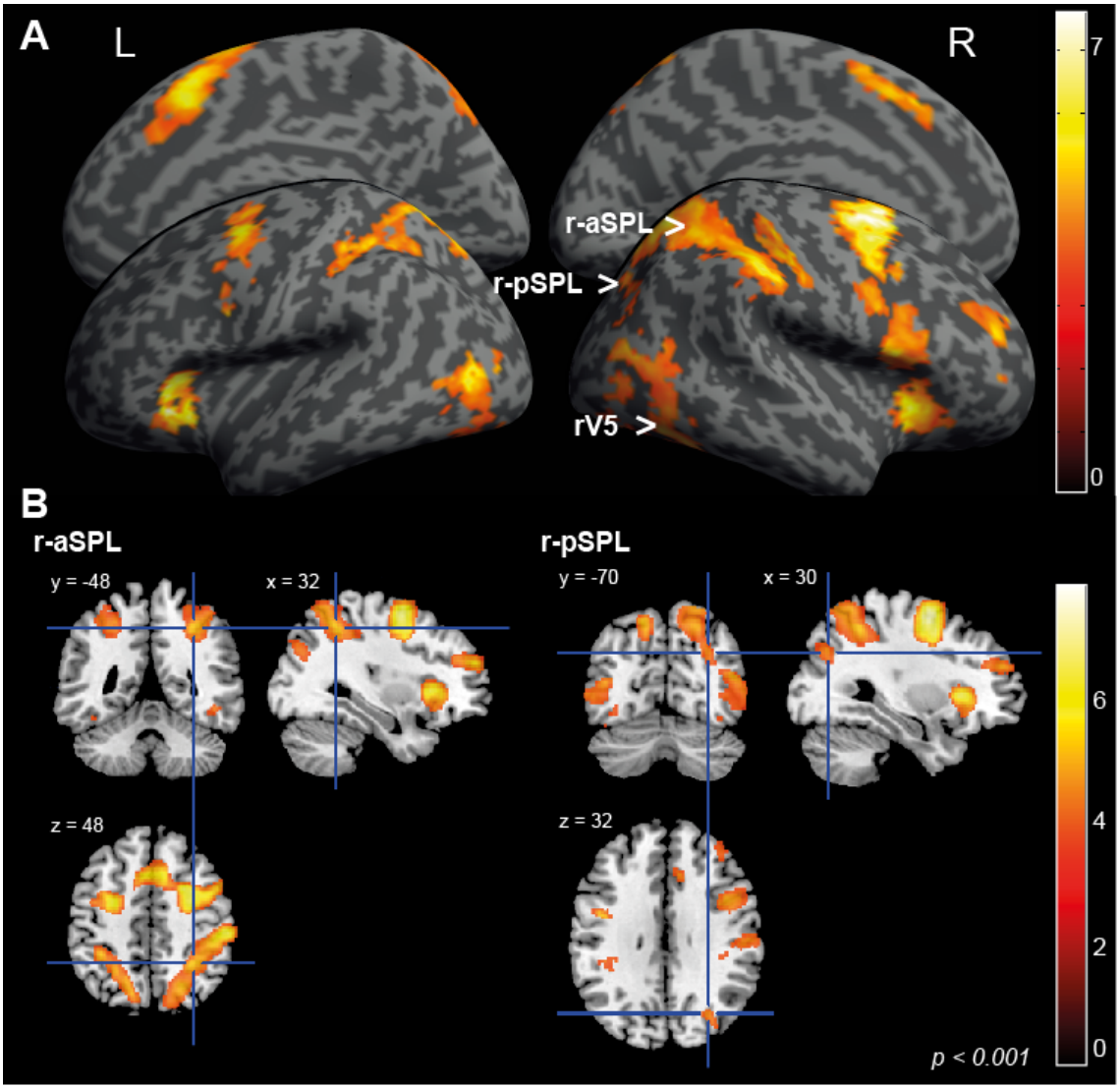

Figure 4-2 Brain activation evoked by perceptual switch.

(A) Activation associated with perceptual transitions in the rivalry condition (rivalry-switch) comparing to the replay condition (replay-switch) is shown in this figure. The color bar indicates the T-value of the GLM activation map which is overlaid on a MNI template artificially 'inflated' using SPM8. (B) The figure shows the peak voxel coordinates of r-aSPL (the left panel, (32, -48, 48), $p < 0.01$, corrected for small-volume) and r-pSPL (the right panel, (30, -70, 32), $p < 0.01$, corrected for small-volume). The color bar indicates T-value of the GLM activation map overlaid on an MNI anatomical template brain using MRICron (http://www.mccauslandcenter.sc.edu/mricro/mricron/).

Table 4-1 Activation in Rivalry-switch > Replay-switch contrast

AAL Label	*t*(17)	*p*	Peak coordinate	Number

		(uncorrected)	*x*	*y*	*z*	of Voxels
Frontal_Sup_R	8.65	< 0.001	18	8	66	5187
Insula_L	8.24	< 0.001	-34	22	2	718
Occipital_Mid_L	6.71	< 0.001	-44	-78	10	775
Postcentral_R	6.54	< 0.001	56	-26	48	4888
Frontal_Mid_L	6.31	< 0.001	-26	-4	50	820
Parietal_Sup_L	6.18	< 0.001	-18	-62	52	1559
Frontal_Mid_R	5.60	< 0.001	32	52	22	544
Occipital_Inf_R	4.01	< 0.001	38	-78	-14	62
Temporal_Inf_L	4.01	< 0.001	-42	-48	-16	8
Thalamus_R	3.71	0.001	8	-14	0	3

To investigate if the existence of binocular disparity in stimuli affected the brain activation in rivalry > replay contrast, I also performed additional GLM analyses with different regressors (Figure 4-3). This alternative GLM showed results similar to the original GLM; r-aSPL and r-pSPL showed greater activation in rivalry switch comparing to replay-switch (r-aSPL peak voxel, (34, -52, 50), $t(17) = 3.75$, $p = 0.001$, uncorrected; r-pSPL peak voxel, (30, -70. 32), $t(17) = 3.11$, $p = 0.003$, uncorrected).

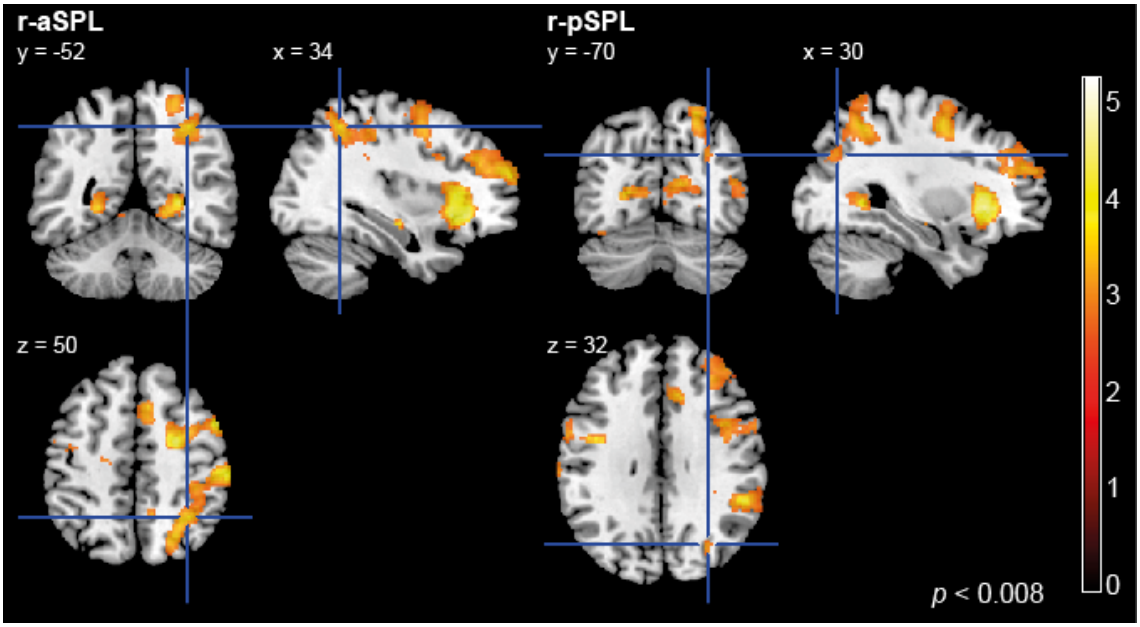

Figure 4-3 GLM validation analysis (second level analysis; *p* < 0.008, uncorrected).

I found activation in r-aSPL and r-pSPL (r-aSPL peak voxel, (34, -52, 50), t(17) = 3.75 , p = 0.001, uncorrected; r-pSPL peak voxel, (30, -70. 32), t(17) = 3.11, p = 0.003, uncorrected).

4.4.3 Dynamic causal modeling analysis

Having established that activity in r-aSPL and r-pSPL was associated with perceptual switches, DCM was performed to characterise the dynamic coupling between three ROIs (Figure 3). r-aSPL (435 voxels), r-pSPL (152 voxels), and r-V5 (356 voxels) were selected based on the GLM results (rivalry-switch > replay-switch) as described in the previous section (Figure 4-4).

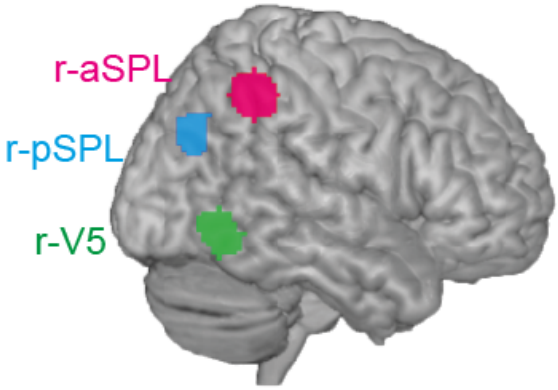

Figure 4-4 Regions of interest for DCM analysis.

r-aSPL (red), r-pSPL (blue), and r-V5 (green) were identified based on anatomical coordinate and univariate analysis (rivalry-switch > replay-switch; see Method and Result for details). I confirmed that r-aSPL and r-pSPL ROIs were consistent in location with previous reports (Kanai *et al.*, 2011)(shown as magenta and cyan in the figure respectively; see main text for details).

To find the optimal DCM model structure that described the interaction between these regions associated with each experimental condition, I used family-level Bayesian model selection. 64 models (combination of all possible models) were divided into four families based on their underlying B-matrix: no modulation, bottom-up, top-down, and bidirectional (Figure 4-5). I found that the exceedance probability was largest for the bidirectional family of the models (Figure 5A; the exceedence probability of the winning model family was 0.83).

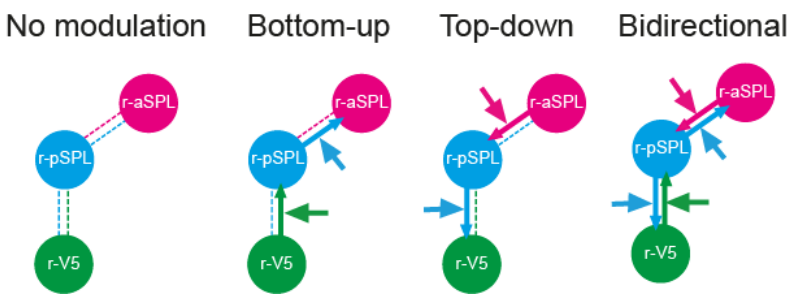

Figure 4-5 DCM model families for model comparison are illustrated.

64 models were divided into four model families (No modulation, 4 models; Bottom-up, 12 models; Top-down, 12 models; Bidirectional, 36 models)

according to modulatory effect. Figure describes modulatory effect (B-matrix) and models with different driving inputs (four patterns) were included in the family.

The winning model in the bidirectional family was the model described in Figure 4-6B (the exceedance probability of the wining model was 0.44 among all 64 models). There were four modulatory effects in the winning model: (r-aSPL to r-pSPL), (r-pSPL to r-aSPL), (r-pSPL to r-V5), and (r-V5 to r-pSPL). In addition, r-V5 received driving input during perceptual transitions. In DCM, the strength of parameters characterises how the rate of activation changes in a region is affected by activation in a given connected region. Here, positive values indicate that increasing activation in a region facilitates the rate of change in the connected region whereas negative values mean increasing activation in a region suppresses the rate of change in the connected region (Seghier *et al.*, 2010).

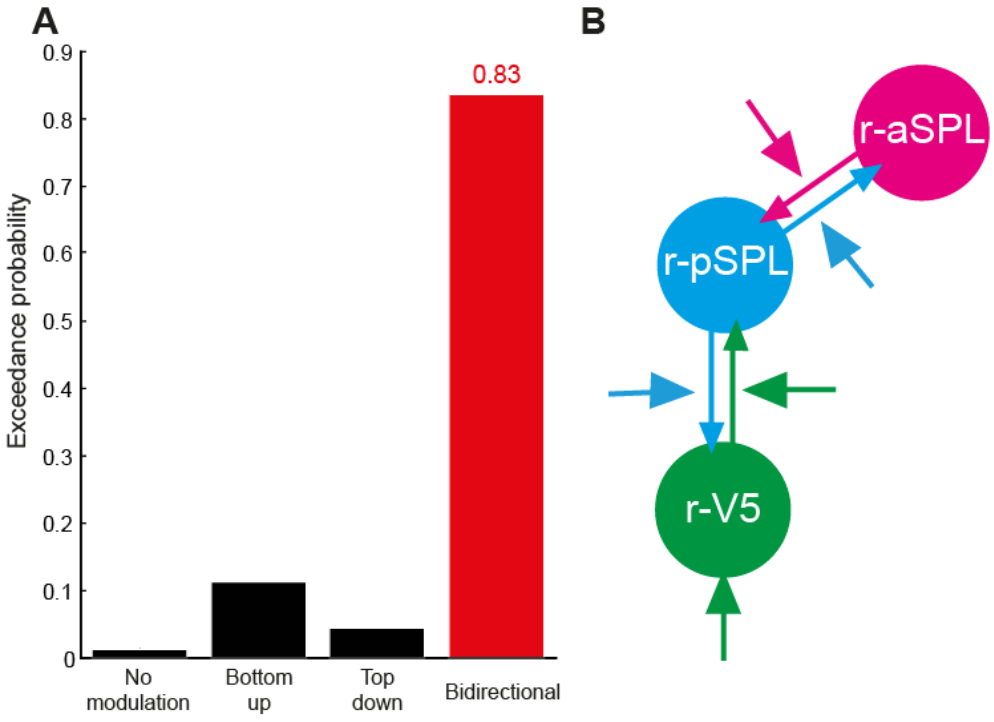

Figure 4-6 Results of DCM analysis.

(A) DCM family-level model comparison result is shown in this figure. Random-effect Bayesian comparison indicates that bidirectional model family (i.e. models containing bottom-up and top-down modulatory effects) was the best among the four families. (B)Winning model contains four

modulatory inputs to all connections and driving input to r-V5 (exceedance probability for the winning model was 0.44 among 64 models).

I additionally performed DCM family-level model selection with two different DCM model construction approaches. First, I looked at if different ROI definition approach would change model selection results. Here, DCM ROI were selected based on individual GLM results. I identified individual ROI coordinates for the three DCM ROIs (Table 4-2). Based on these coordinates, I created individual ROI masks for DCM and performed DCM model selection. Bayesian family-level comparison revealed that the bidirectional model family was the best among the four families (Figure 4-7A; exceedence probability 0.92).

Also I performed family-level DCM model comparison with the alternative GLM to explore if difference in regressors would affect DCM family level inference (Figure 4-7B). I defined DCM ROIs with similar procedure used for the main DCM analysis: 10mm-sphere masks centered on the group level peak voxel coordinates ((34, -52, 50) for r-aSPL, (30, -70, 32) for r-pSPL, and (44, -74, 6) for r-V5) were applied to group level fMRI results (Rivalry-switch > Replay-switch contrast, thresholded at $p < 0.008$, uncorrected) and used as DCM ROIs. Again, I found that the bidirectional model family was the most likely DCM model family among the four families (exceedance probability was 0.98).

These two additional DCM analysis results confirmed the validity of the bidirectional model to account for neural dynamics during spontaneous perceptual switches.

Table 4-2 Individual peak coordinate of r-aSPL, r-pSPL, and r-V5 within DCM ROIs.

	r-aSPL			r-pSPL			r-V5		
	X	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
Subject 1	26	-52	52	30	-62	38	42	-58	-14
Subject 2	38	-48	54	30	-74	30	48	-58	-14

Subject 3	34	-52	48	32	-64	34	52	-66	0
Subject 4	36	-42	54	30	-78	38	52	-62	0
Subject 5	32	-40	42	38	-70	38	54	-58	-6
Subject 6	36	-48	42	28	-66	30	42	-58	-10
Subject 7	36	-50	54	24	-66	38	48	-66	0
Subject 8	38	-42	50	32	-76	28	48	-66	-6
Subject 9	28	-54	46	34	-68	36	54	-58	-14
Subject 10	38	-54	52	34	-64	26	52	-66	-10
Subject 11	36	-42	50	32	-64	36	50	-58	0
Subject 12	36	-40	52	28	-72	24	44	-64	-16
Subject 13	28	-40	44	26	-72	36	48	-70	-2
Subject 14	32	-42	40	26	-76	36	48	-68	0
Subject 15	36	-46	44	30	-62	38	42	-58	-6
Subject 16	40	-44	46	26	-62	36	54	-56	-6

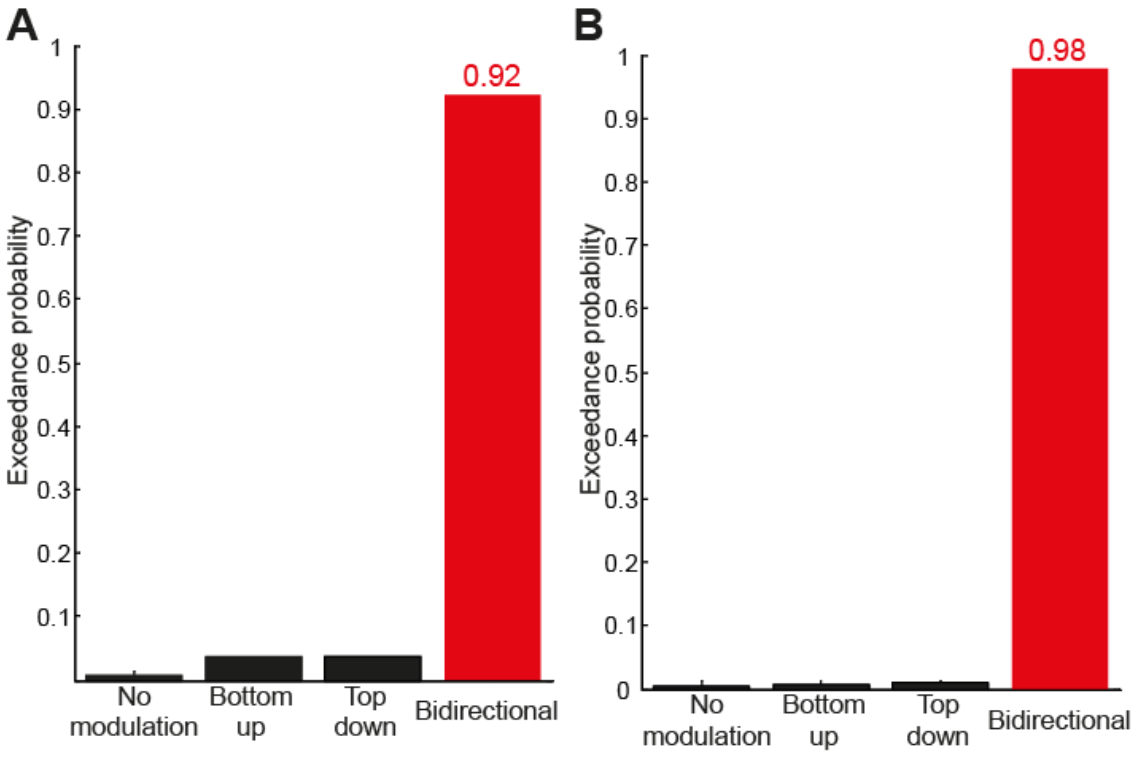

Figure 4-7 DCM validation analyses.

(A) I also performed another DCM analysis with individual-based ROI selection approach. Based on the original GLM (See Figure 4-2 for activation map), three ROI masks were created based on individual GLM activation map (See Table 4-2 for peak coordinates) and DCM family-model comparison revealed bidirectional model family was the winning model family. (B) To address if difference in GLM would affect DCM result, I performed DCM analysis with the alternative GLM (Figure 4-3). Again that bidirectional

model family was the best among four families with the alternative GLM.

Finally, multiple regression analysis was performed to explore if the parameters in the winning DCM model could predict variation in the individual behavioural data (mean percept dominance duration). Differences in the four DCM B-parameter values between the two switch conditions were entered into a multiple linear regression model as predictors. The model successfully predicted individual mean dominance duration (R^2 = 0.77 (adjusted R^2 = 0.67), $F(4, 10) = 8.18$, $p = 0.003$; see Figure 4-8A). Specifically, two bottom-up B-parameters were significantly correlated with mean dominance duration in the full model (ß = -1.896, $t(10) = -3.919$, $p = 0.003$, $p < 0.01$ for r-pSPL to r-aSPL; ß = 2.18, $t(10) = 4.30$, $p = 0.002$, $p < 0.01$ for r-V5 to r-pSPL; see Figure 6B): suppressive modulation from r-pSPL to r-aSPL and facilitative modulation from r-V5 to r-pSPL were associated with a longer dominance duration. I did not observe such trends in the two top-down B-parameters ((ß = -0.25, $t(10) = -1.28$, $p = 0.23$, *n.s.* for r-aSPL to r-pSPL; ß = 0.65, $t(10) = 2.05$, $p = 0.07$, *n.s.* for to r-pSPL to r-V5).

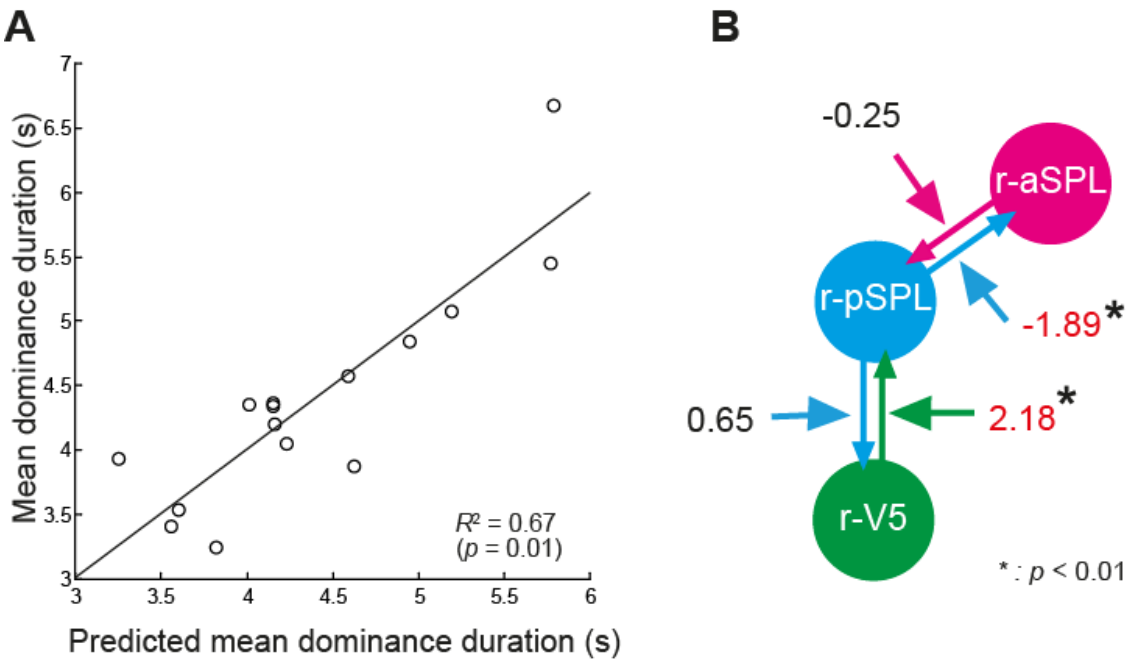

Figure 4-8 Result of multiple regression analysis.

(A) Multiple regression analysis showed that a combination of four B-parameters could predict an individual's mean dominance duration. The R^2 value given in the figure is adjusted R^2. (B) Two bottom-up modulatory parameters (r-pSPL to r-aSPL and r-V5 to r-pSPL) were the significant

predictors for individual mean dominance duration (r-pSPL to r-aSPL, p = 0.003, p < 0.01; r-V5 to r-pSPL, p = 0.002, p < 0.01). Values besides the arrows indicate β (standardized coefficients) of each predictor in the full-model.

4.5 Discussion

4.5.1 Overview of the findings

Based on evidence from previous TMS and VBM studies, I investigated how two focal areas of parietal cortex and the motion-sensitive area V5/MT of the human brain interacted with each other during visual perceptual switches in bistable perception. Using DCM analyses, I formally characterised reciprocal modulatory interactions between these brain areas which were designated by my prior hypothesis (Kanai *et al.*, 2011). Furthermore, I found that the strength of bottom-up modulations accounted for inter-individual variability in percept dominance duration.

4.5.2 Role of fronto-parietal regions in perceptual switches

I first replicated the previously described functional association between activity in human parietal regions and perceptual switches. Lumer *et al.* (1998) showed that higher BOLD signals in the superior parietal lobule (SPL) are observed time-locked to perceptual switches in binocular rivalry. Kanai and colleagues (Kanai *et al.*, 2010; Kanai *et al.*, 2011) showed that cortical grey matter volume and thickness of r-aSPL and r-pSPL correlate with perceptual switch rate for structure from motion (SFM). In addition, modulation of subjective perception by application of TMS to these areas confirmed a causal role for these regions in bistable perception (Carmel *et al.*, 2010b; Kanai *et al.*, 2010; Kanai *et al.*, 2011). Despite this compelling collection of evidence for the role of right human SPL in fluctuations of subjective awareness, the functional interplay between these parietal subregions and lower visual areas has not previously been shown.

It has been suggested that perceptual switches are caused by continuous cortical interactions between fronto-parietal regions and sensory regions rather than just "bottom-up (feed-forward)" or "top-down (feedback)" neural communication (Sterzer *et al.*, 2009). Previous TMS and fMRI studies (Zaretskaya *et al.*, 2010; Kanai *et al.*, 2011) pointed to a role for connectivity between a number of parietal and visual brain areas in the human brain in bistable perception. Multiple brain regions, including visual cortex and fronto-parietal regions, show activation when perceptual switches occur and this has been replicated several times; see Rees (2007). In addition, a recent fMRI study (Wang *et al.*, 2013) suggested that changes in functional connectivity (Friston *et al.*, 2013) between multiple brain regions is enhanced during a bistable perception task further supporting the role of connectivity changes in bistable perception. Despite this wide range of previous findings, whether fronto-parietal activation associated with perceptual switches directly contributes to conscious perception is contested: for example, activations of fronto-parietal regions could reflect top-down information processes such as selective attention (Sterzer *et al.*, 2009). Alternatively, a recent study proposed that activations observed in the fronto-parietal regions are due to ambiguity in visual information rather than a driving force of perceptual switches (Knapen *et al.*, 2011). Yet another more recent study has attributed this brain activity to introspection and report of perceptual states (Frässle *et al.*, 2014) rather than a change in the subjective content of consciousness. These results cast doubt on the involvement of fronto-parietal areas in perceptual alternation.

To address this issue directly, I used DCM analysis to identify the dynamics of network level interactions between parietal and motion sensitive visual areas during perception of bistable structure from motion. The winning model comprised four bidirectional connections in which r-V5 is both a driving force as well as modulated by perceptual switches. This structure indicates that sensory input to r-V5 propagates to higher brain areas (r-pSPL and r-pSPL); and r-V5 and r-pSPL both receive feedback modulation, suggesting that perceptual switches are induced as a result of bidirectional modulation between fronto-parietal and sensory areas.

Furthermore, I found that the variation in bottom-up modulatory parameters (B-parameters) between the rivalry and replay conditions could predict the individual participant's mean dominance duration. Although further studies will be required to understand the precise nature of the biological mechanisms that account for the difference between the two bottom-up modulations, the correlation between DCM parameters and mean dominance duration is evidence for the involvement of these two parietal regions in perceptual switches.

4.5.3 Bistable perception and predictive coding

How does the bidirectional interaction described here give rise to changes in perceptual states? The predictive coding theory of brain function (Helmholtz, 1910; Rao & Ballard, 1999; Clark, 2013) offers a framework to answer this question. This theory proposes that the brain seeks to infer the causes in the external world that give rise to the signals gathered through sensory organs. Based on these inferences, the brain constructs expectations or predictions about subsequently forthcoming sensory inputs which are then iteratively updated by comparing the expectations with the observation and computing the "prediction error" (Hohwy *et al.*, 2008). The neuronal correlates of such iterative prediction and comparison processes have been documented in several brain regions when participants engage in visual tasks (Murray *et al.*, 2002; Muckli *et al.*, 2005; Summerfield *et al.*, 2006). Recent theoretical (Hohwy *et al.*, 2008) and empirical work (Denison *et al.*, 2011) have also suggested that predictive coding theory could account for perceptual alternation in bistable perception.

Kanai *et al.* (2011) employed this framework to propose a connectivity hypothesis consisting of r-aSPL, r-pSPL and visual cortex that might account for bistable perception. Based on the observation that impairing r-pSPL and r-aSPL function by TMS prolongs and shortens, respectively, the mean dominance duration in bistable structure from motion perception, they proposed that r-aSPL generates a prediction about the causes of sensory evidence (i.e. structure of the environment) and

r-pSPL computes the prediction error between that expectation and the sensory evidence it receives from the visual cortex. The current results showed that the connectivity structure of the winning DCM model is consistent with the connectivity hypothesis proposed by Kanai et al (2011).

Taken together, I speculate that the bottom-up modulation from r-pSPL to r-aSPL (and from V5/MT to r-pSPL) corresponds to a hierarchical process of "explaining away" which may serve to balance out the difference between prediction (represented in r-aSPL) and sensory information (represented in r-V5). In this view, smaller prediction errors (i.e. less bottom-up modulation from r-pSPL to r-aSPL and from V5/MT to r-pSPL; it means the sensory input matches the current percept) would lead to stabilized perception (longer mean dominance duration) as demonstrated in my findings. Indeed, the coefficient of the regression model indicated that participants with 'suppressive' bottom-up modulation had longer mean-dominance duration.

It is still hard to interpret the biological meaning of fMRI-DCM parameters and further electrophysiology validation may be needed to understand more detailed neural mechanism behind the perceptual switches and neural modulations.

4.5.4 Comparison with other studies

Another recent fMRI study drew a rather different conclusion regarding regional interactions in perceptual switches during multistable perception. Weilnhammer *et al.* (2013) explored perceptual alternations associated with a rotating Lissajous figure and demonstrated that a DCM model with top-down modulation (but no bottom-up modulation) from the right inferior frontal gyrus (rIFG) to right V5/MT could account for the neural dynamics of spontaneous perceptual switches. The difference between the present and previous studies may be associated with differences in paradigm, but may also come from the ROI selection process. In this study, I included both anterior SPL and posterior SPL in DCM models separately based on the hypothesis from previous study (Kanai *et al.*, 2011) and did not include rIFG. DCM analysis should be performed using

anatomical or functionally connected regions (Stephan *et al.*, 2010) and parietal regions and V5/MT are indeed anatomically and functionally connected (Mars *et al.*, 2011). Most importantly, I found the strength of two bottom-up connections predicted individual mean dominance duration and this implies involvement of bottom-up connectivity in defining the timing of perceptual alternation, at least for my structure-from-motion stimulus.

4.6 Conclusion

In this chapter, I performed an fMRI experiment and DCM analyses to elucidate functional role of the two fronto-parietal subregions for bistable perception. I found that activity in two focal regions of parietal cortex plus motion-sensitive visual cortex influenced each other during bistable perceptual switches; and the strength and direction of modulation of connectivity between regions predicts individual mean percept dominance duration. The results are consistent with a predictive-coding theory of bistable perception and contribute to clarifying the dynamics of a functional network in the brain that contributes emergence of conscious perception.

Chapter 5 RESTING STATE NETWORK ARCHITECTURE PREDICTS INDIVIDUAL SWITCH FREQUENCIES IN BISTABLE PERCEPTION

5.1 Summary of this chapter

In the previous chapter, I demonstrated that bistable perception induces complex brain dynamics and multi-regional interaction. However, it remains unclear whether some aspects of such intrinsic network organization are present even without external input in the resting state and predictive for bistable perception performance. To address this question, I applied graph theoretical analysis to resting state fMRI data recorded from human participants to determine which brain regions and functional networks could predict individual switch frequency in bistable perception. Regions of interest were defined anatomically from a previous meta-analysis and a binary undirected graph was created based on resting-state functional connectivity between these regions. I performed a support vector regression analysis with centrality measures computed from the binary graph, and I found that hubness (PageRank centrality) of right postcentral gyrus, left insula, and left lateral occipitotemporal gyrus all significantly predicted individual differences in mean dominance duration (or its inverse, switch frequency). Furthermore, such prediction performance was significantly deteriorated when the connections of the fronto-parietal and visual sub-network, which is thought to play a significant role in resolving ambiguity in visual information, were removed from the network analysis. These findings show that systematic features of the resting state connectivity can predict task-related behavioral dynamics, suggesting that intrinsic network properties of human brain underlie individual differences in the dynamics of visual awareness.

5.2 Introduction

5.2.1 Background

The experiment and analysis in the previous chapter showed that multiple brain regions, including visual cortex and fronto-parietal areas, form brain network and they communicate with each other when resolving ambiguity in visual input. The strength of communication explained individual difference in the behaviour (mean dominance duration in bistable perception task), suggesting that information flow between regions have functional roles for building up visual perception.

Other empirical studies also reported multiple brain regions showing time-locked activation for perceptual switches, including the lateral geniculate nucleus (Haynes *et al.*, 2005; Wunderlich *et al.*, 2005), visual cortex (Kleinschmidt *et al.*, 1998; Polonsky *et al.*, 2000; Tong & Engel, 2001; Knapen *et al.*, 2011) and fronto-parietal network (Lumer *et al.*, 1998; Sterzer & Kleinschmidt, 2007; Zaretskaya *et al.*, 2010; Weilnhammer *et al.*, 2013). This suggests that these regions interact with each other while observers are resolving ambiguity in stimuli. Moreover, individual characteristics seen in the complex energy landscape associated with neural dynamics during bistable perception can account for individual differences in perceptual switch frequencies (Watanabe *et al.*, 2014).

5.2.2 Hypothesis of the study: Does structure of task-independent intrinsic networks in the brain predict complex individual behavior?

An intriguing question arises concerning whether such mechanisms might also have a signature in the resting state, independently of visual perception per se. Individual differences in bistable switch frequency show trait-like properties, are correlated with gray matter volume in focal cortical regions (Kanai *et al.*, 2010; Kanai *et al.*, 2011), and have high heritability suggesting a genetic component (Miller *et al.*, 2010; Shannon *et al.*, 2011). I therefore hypothesized that the dynamics of resting state brain networks measured by

fMRI would predict individual differences in switch frequency during bistable perception.

To test this hypothesis, I analyzed functional connectivity architecture of the human brain in the resting state using a graph theoretical approach. Graph theoretical approach characterizes detailed features of complex networks and systems (Sporns *et al.*, 2000; Bullmore & Sporns, 2009) and it describes dynamic properties of brain functional and structural networks (Power *et al.*, 2013). In particular, centrality measures quantify how nodes (brain regions) are connected with each other (hubness) and can relate brain network organization to individual traits (Cole *et al.*, 2012; Lord *et al.*, 2012; Warren *et al.*, 2014).

5.2.3 Analysis strategy

I hypothesized that such graph-theoretical measures of functional network organization would predict individual behavioural performance in a subsequent and separate bistable perception task. I therefore obtained resting state fMRI data, independently measured behavioural performance in a separate bistable structure-from-motion task outside the scanner, and finally related the two measures. First, I constructed an undirected binary graph from a resting state functional connectivity map and examined whether centrality measures of each node predicted individual perceptual switch frequencies. Then I determined which brain regions were most predictive of switch frequency (cf. my behavioural measure was mean dominance duration, which is the inverse of switch frequency). I then determined which intrinsic brain network contributed to such a prediction using a 'targeted network attack' approach. If an intrinsic sub-network contributes to prediction accuracy, removing the edges (connection between nodes) of the ROIs in the network will result in increase in errors (i.e., decrease of prediction accuracy).

In this study, I also computed four centrality measures – PageRank centrality, degree centrality, closeness centrality, and betweenness centrality – and asked if which centrality I should use for prediction.

Basically, centrality measures describe how a node (ROI) influences on the network and several measures have been proposed and applied to fMRI data analysis. Figure 5-1 is an example graph and here nodes are coloured with its centrality values (degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality. As shown in the figure, node's centrality measures vary across different centralities as they capture different characteristics.

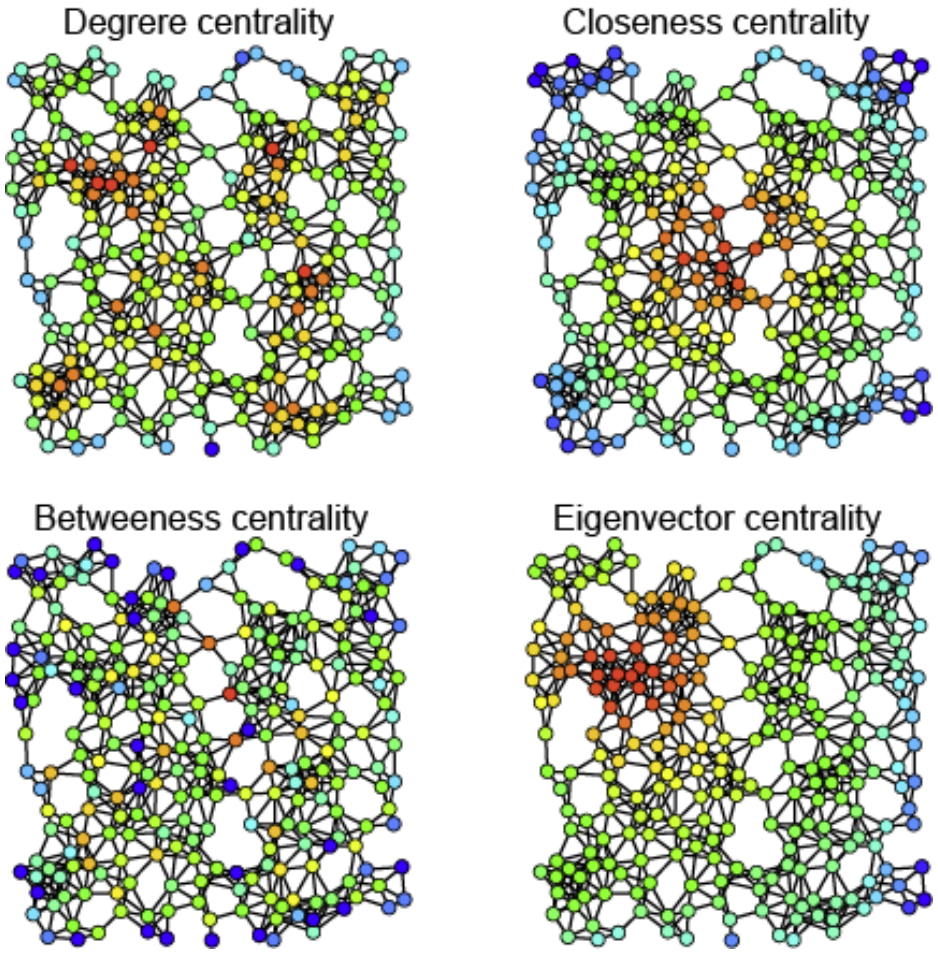

Figure 5-1 Centrality measures.

This figure depicts four centrality measures of the same graph. Color of the node indicates centrality of each node – reddish node means higher centrality whereas blueish node means lower centrality (Jet colormap). Figures adapted from Claudio Rocchini's homepage (http://www.rockini.name/math/index.html CC-BY).

5.3 Methods

5.3.1 Participants

Twenty-seven healthy participants (18 female and 9 male, aged 18-36, mean age ± standard deviation, 24.7 ± 5.2; all right handed and with normal or corrected-to-normal vision) performed the behavioural task and underwent resting state fMRI measurements. All participants provided written informed consent and the experimental procedure was approved by the UCL research ethics committee.

5.3.2 Bistable perception behavioral experiment

We used structure-from-motion stimuli (Figure 5-2) in the bistable perception task (Kanai *et al.*, 2010) and recorded behavioural reports of spontaneous fluctuations in bistable perception.

Figure 5-2 Spherical shape structure-from-motion.

Typically this stimulus elicits one of two exclusive percepts (right-rotating or left-rotating sphere). Participants were asked to report which direction they thought the sphere was rotating using one of two button presses. Perceptual switches are typically crisp so that there shouldn't be long mixture percepts in perceptual transitions.

The stimuli were presented on a 22-inch LCD monitor (Samsung SyncMaster 2233RZ; spatial resolution 1680 × 1050; monitor refresh rate 60Hz) at a viewing distance of 80cm fixed using a chin rest. Structure-from-motion stimuli were created using PsychToolbox 3 (Brainard, 1997) under MATLAB2007b (The Mathworks, Inc.) and were presented against a black background. A configuration of 200 full-contrast white dots was displayed on the center of the computer screen. These dots moved sinusoidally with an angular velocity of 120 degrees/s so that they were perceived as a sphere (diameter 3.1° visual angle) spinning around the

vertical axis with ambiguity in the direction of rotation. A red fixation point was superimposed at the center of the stimulus.

Each trial comprised a brief fixation period followed by stimulus presentation for 48s. Participants were instructed to fixate their eye gaze on the red fixation point and provide continuous report of the perceived direction of rotation by holding one of two buttons. They performed the task for 10 trials (eight minutes in total) and I computed mean and median dominance duration for each participant. Participants were told not to press any buttons if the percept was not clear and the period without button response was not included to the mean dominance duration.

In addition to dominance duration, I also computed no-report duration for each trial, which is sum of the time without observer's button press due to mixture percept or delay in button response. Perceptual switches in SFM are very crisp and times of no button response is very short and therefore I considered no-report duration as 'behaviour-of non-interest' and tested if centrality measures computed from resting state would be predictive for trivial variance in individual behaviour.

5.3.3 MRI data collection

MRI data were acquired at the Wellcome Trust Centre for Neuroimaging at University College London. Structural and resting state functional images were collected using a 3T Siemens Magnetom Trio with a 32-channel head coil. For resting state functional imaging, Blood Oxygen Level Dependent (BOLD) signals were measured using an echo planar imaging (EPI) sequence (volume repetition time, 2.176 s; echo time, 30 ms; flip angle, 90°). EPI images contained 32 axial slices (3 mm thickness, ascending slice order), voxel size was 3 mm × 3 mm × 3 mm, and the field of view was 192 mm × 192 mm. Prior to resting state measurement, all participants were given identical instructions to close their eyes, relax, and to not fall asleep. T1-weighted structural images (3D Modified Driven Equilibrium Fourier Transform Sequence (Deichmann *et al.*, 2004)) were also acquired with 1 × 1 × 1 mm resolution, with 176 sagittal slices covering the whole brain (slice

repetition time, 7.92 ms; echo time, 2.48 ms; flip angle, 16°). Phase image and magnitude images were also obtained to compute a fieldmap (Hutton *et al.*, 2002).

5.3.4 MRI data processing

Brain imaging data were processed using Statistical Parametric Mapping software (SPM12: Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm). Resting state data comprised 276 volumes; the first six EPI volumes were discarded to allow for T1 equilibration. For preprocessing the EPI images, first, distortions of EPI images were corrected. The EPI images were then realigned and unwrapped based on fieldmap images using the FieldMap toolbox in SPM8. EPI images were spatially normalized to the Montreal Neurological Institute (MNI) stereotactic template using DARTEL module in SPM12. Data were smoothed spatially with a Gaussian kernel of 8 mm full-width at half-maximum.

5.3.5 ROI definition for network analysis

We employed a collection of 254 ROIs defined in a previous meta-analysis (Power *et al.*, 2011) that were also within the EPI coverage as network nodes (Figure 5-3; see Appendix for the full ROI list). These ROIs were defined based on task fMRI activation and are putative functional areas. Full coordinates and detailed description of ROI definition can be found in the original paper (Power *et al.*, 2011) and the senior author's website (http://www.nil.wustl.edu/labs/petersen) and Appendix.

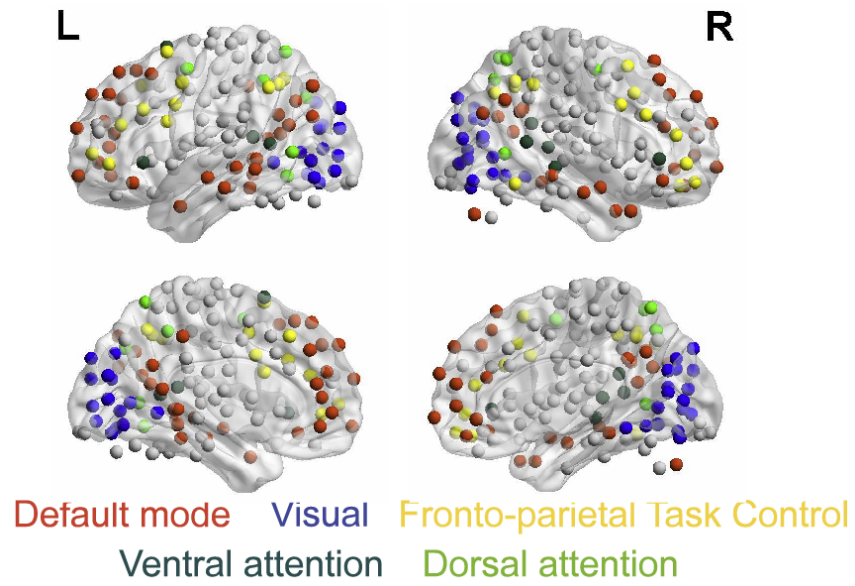

Figure 5-3 ROIs used for the analysis.

259 ROIs covering the whole brain (Power *et al.*, 2011) were used for the analysis. These ROIs were classified into five networks: default mode network (DMN), visual network (VN), fronto-parietal task control network (FPN), ventral attention network (VAN), and dorsal attention network (DAN). Gray spheres are unlabeled ROIs. The figure was created by BrainNet Viewer (http://www.nitrc.org/projects/bnv/).

5.3.6 Resting state data preprocessing

To remove nuisance signals which are unlikely to reflect spatially-specific functional connectivity, resting state fMRI data was processed in accordance with a standard procedure reported in previous studies (Van Dijk *et al.*, 2010; Power *et al.*, 2011). First, using the filtering module in the REST toolbox (Song *et al.*, 2011), I applied a temporal band-pass filter (0.009 Hz < f < 0.08 Hz). Then regression analysis with nine explanatory variables (six realignment parameters, averaged signal over gray matter, white matter, and cerebrospinal fluid) was applied to remove the effect of body movement and scanner-dependent noise. The resultant residual time courses were averaged within each ROI (5mm radius spheres), and were then used to compute functional connectivity between all ROI pairs. The correlation value (r) was transformed to Fisher's z-transformed correlation.

5.3.7 Analysis strategy and pre-process for graph theoretical analysis

A binary graph was created based on the functional connectivity correlation matrix and subjected to graph network analysis (see Figure 5-4 for analysis pipeline). First, the correlation matrix was transformed by Fisher's z-transformation and then converted to a binary adjacency matrix to retain only the strongest functional connectivity (> 80th percentile; computed from group-level z-transformed correlation matrix; thresholded at z = 0.13, equivalent to r = 0.13; i.e. assigned 1 to an edge where functional connectivity was stronger than 0.13, otherwise 0 to the edge). This threshold level was chosen based on a previous fMRI-based graph theoretical study (Ekman *et al.*, 2012).

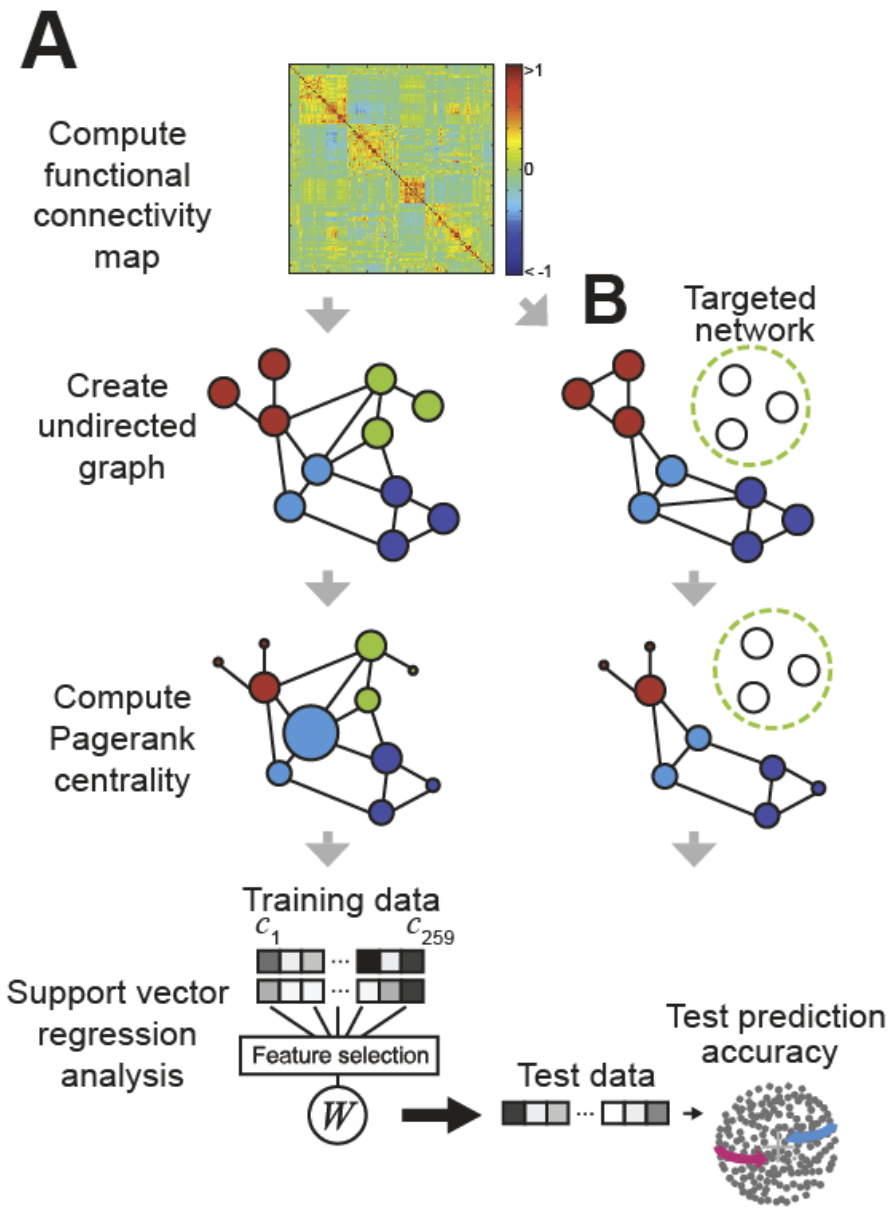

Figure 5-4 Analysis strategy.

I used a graph theoretical approach to address the research question. First functional connectivity matrix was converted to binary undirected graph and then centrality measures were computed from the resultant graphs. (B) To evaluate the contribution of intrinsic brain networks to SVR prediction accuracy, I performed 'tagreted network attack' analysis, where nodes belonging to certain functional networks were systematically removed from the network and centrality measures were computed based on the adjacency matrix.

5.3.8 Centrality measures

In this study, I computed PageRank centrality, degree centrality, closeness

centrality, and betweenness centrality. Centrality measures quantify the importance of nodes (ROI) within the graph, and these four centrality measures characterize different aspects of node's influence over the network. I aimed to interrogate which centrality measure would be appropriate to describe the link between intrinsic network and behaviour during bistable perception.

5.3.8.1 PageRank centrality

The centrality measures (PageRank centrality, betweenness centrality, closeness centrality) were computed from the individual adjacency matrix with NetworkX (Hagberg *et al.*, 2008).

PageRank centrality characterizes the strength of each node's connection with other nodes within the network and it also considers the entire organization of the network by preferring nodes that have connection to central nodes within the network (Lohmann *et al.*, 2010) and has been previously applied to neuroimaging data (Ekman *et al.*, 2012). Higher PageRank centrality means that the node is well-connected with other nodes (i.e. higher 'hubness').

PageRank centrality for a node i is computed as follows:

$$PR_i = \frac{1-\lambda}{n} + \lambda \sum_{\mathrm{j}\epsilon M(\mathrm{i})} \frac{PR(\mathrm{j})}{O(\mathrm{j})},$$

where $M(i)$ is the set of the nodes connected to a node i and $PR(i)$ represents PageRank centrality of node j, and n is the number of nodes in the graph. O(i) denotes the out-degree of the node i. In the present study, I created an undirected binary graph and thus did not consider directionality of the graph. Therefore, I used the same value for out-degree and in-degree parameters and λ was set to 0.85 (the same as Ekman *et al.* (2012)).

5.3.8.2 Degree centrality

Computation of the degree centrality is very simple; degree centrality considers number of connections a node has and degree centrality for a node $i(D_i)$ is computed as follows:

$$D_i = \sum_{j \in N} a_{ij}$$

where a_{ij} is the status of connection between i and j, which is number of edges connected to the node.

5.3.8.3 Closeness centrality

Closeness centrality is a placement on many of the network's shortest paths (Sporns, 2014).

Closeness centrality for a node $i(L_i^{-1})$ is computed as follows:

$$L_i^{-1} = \frac{n-1}{\sum_{j \in N, j \neq i} dij}$$

d_{ij} is a shortest path length between i and j,

$$d_{ij} = \sum_{a_{uv} \in g_{i \leftrightarrow j}} a_{uv}$$

where $g_{i \leftrightarrow j}$ indicates the shortest path between i and j. This equation means that closeness centrality of a node is obtained by counting total number of steps to reach the node from other nodes in the graph.

5.3.8.4 Betweenness centrality

Betweenness centrality (Freeman, 1977) represents how often the node disturbs the shortest paths of other nodes in the graph, and it is obtained as follows:

$$B_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N \\ h \neq i, h \neq j, i \neq j}} \frac{\rho_{hj}(i)}{\rho_{hj}}$$

where ρ_{hj} is the number of shortest paths between node h and j (d_{hj}).

5.3.9 Multivariate support vector regression analysis

Support-vector regression analysis was performed using the LIBSVM toolbox (http://www.csie.ntu.edu.tw/~cjlin/libsvm). Support vector machines are a class of widely-used machine learning algorithms with high generalization ability (see Haynes and Rees (2006)) and have been extensively applied to brain imaging data analysis (Cox & Savoy, 2003; Kamitani & Tong, 2005). Support vector regression (SVR) is an extension of the support vector machine algorithms that enables prediction of continuous values rather than category or class (Smola & Scholkopf, 2004), such as age or maturity of the brain (Dosenbach *et al.*, 2010).

The SVR analysis comprised feature selection, SVR training, and validation using test data, in which I took a leave-one-out cross validation approach to test generalization of the prediction accuracy. The aim of feature selection was to reduce the number of features (here, ROIs). Neuroimaging data tend to be high-dimensional and it is often beneficial to reduce the number of features to achieve high classification performance (Norman *et al.*, 2006; Pereira *et al.*, 2009) in neuroimaging data analyses (Chadwick *et al.*, 2010; Wager *et al.*, 2011). In this study, I used a standard correlation analysis to find the most informative regions for predicting individual mean dominance duration in bistable perception. First, using the training dataset, the Pearson correlation (r) between network measures of each ROI and individual mean dominance duration was computed and then N ROIs showing the strongest correlation (both positive and negative) were used as features (N is a number of features used for SVR). Zero was assigned as the SVR weight for features that were not selected as informative features at the feature selection stage. Note that test data were not used in this feature selection process. Individual mean dominance duration and centrality measures were rescaled between 0 and 1 ($X_{normalized}$) according to the equation:

$$X_{normalized} = \frac{(X - \min X)}{\max X - \min X},$$

where X denotes the feature matrix and contains both training data set and

test data set. I then used epsilon-SVR with a linear kernel on the small set of features (ROIs) identified in the first feature selection phase. The details of the algorithm can be found elsewhere (Chang & Lin, 2011).

We evaluated the accuracy of the SVR prediction by computing correlation between prediction and measured behavioural data. For statistical testing, I used non-parametric approach to avoid assumptions regarding the distribution of the prediction and the error. The Spearman correlation (*rho*) between the normalized mean dominance duration and predicted mean dominance duration was computed and used as a measure of the fit of prediction. In addition, using a MATLAB toolbox (Pernet *et al.*, 2012), Spearman's skipped correlation (Rousseeuw, 1984; Rousseeuw & Driessen, 1999; Verboven & Hubert, 2005) was used to validate the robustness of the fit against outliers in the variables.

5.3.9.1 Targeted network attack analysis

To investigate which functional brain networks contributed to prediction of mean dominance duration, I removed edges (connection between nodes) of targeted network ROIs with other areas, and then performed SVR. This hub attack approach is commonly used in network analysis for neuroimaging data (Achard *et al.*, 2006) to investigate the significance of each node in the context of the entire network. I assumed that each intrinsic network, rather than each ROI, would have a functional role for behaviour, and therefore I removed edges at the network level rather than at ROI level and explored the effect of edge removal on prediction. In sum, I removed all edges connected to the targeted network nodes and computed MSE after performing SVR.

From this analysis I defined (using independently defined ROIs – see above) five well-established intrinsic functional networks that are very likely involved in (visual) bistable perception (attention, self-monitoring, visual introspection) (Knapen *et al.*, 2011; Watanabe *et al.*, 2011; Zhang *et al.*, 2011; Freeman *et al.*, 2012; Karten *et al.*, 2013; Frässle *et al.*, 2014): the default mode network (DMN, 57 regions), the visual network (VN, 31

regions), the fronto-parietal task control network (FPN, 25 regions), the ventral attention network (VAN, 9 regions), and the dorsal attention network (DAN, 11 regions).

I analyzed mean squared error (MSE) of the prediction and used it to compare goodness of the prediction under targeted network attack conditions. For statistical testing, I again employed non-parametric approach (Wilcoxon signed rank test and Friedman's test).

5.4 Results

5.4.1 Behavioral results

I characterized switch frequency of each participant during spherical-shape structure-from-motion bistable perception by calculating dominance duration (the inverse measure, which is the average duration that each percept remains unchanged). Twenty-seven participants performed a bistable perception task with structure-from-motion (SFM) stimuli and continuously reported their perceptual experience using key presses. The mean dominance duration of the two exclusive percepts (clockwise-rotating sphere and anticlockwise-rotating sphere) was 9.30 ± 4.33 seconds, ranged 4.31 to 19.63 seconds (Figure 5-5A). I also computed the median of the duration (Figure 5-5BC). The mean and SD of the individual median dominance duration was 6.69 ± 3.15, raged 2.60 to 14.80. Mean and median dominance duration showed strong correlation (Figure 5-5C, $rho = 0.84$, $p = 0.0000016$) and therefore I used mean dominance duration as behavioral measure for stability of percept in bistable perception task.

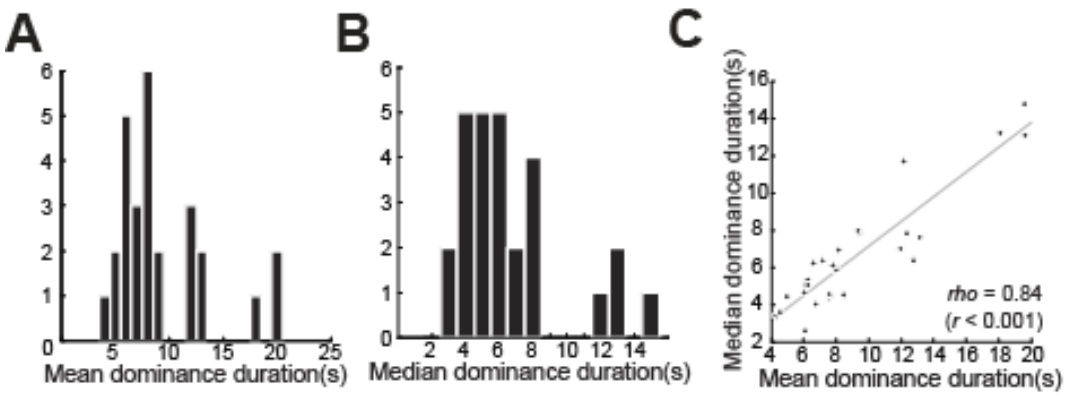

Figure 5-5 Behavioral data analysis.

(A) The histogram represents mean dominance duration of the participants (average and SD, 9.30 ± 4.33s). (B) I also measured median dominance duration, which is the median of dominance duration for the two exclusive percepts. (C) I confirmed that mean and median of individual dominance duration showed strong correlation (*rho* = 0.84, $p = 0.0000016$).

5.4.2 Testing four centrality measures and prediction accuracy

I explored whether centrality measures could predict mean dominance duration in bistable perception (SFM) and which areas contributed most to the prediction. I performed SVR analysis to test the hypothesis: PageRank centrality and eigenvector centrality were calculated from individual binarized functional connectivity matrices and a support vector machine was trained with training data sets before testing the generalization with independent test data sets.

First, I investigated SVR prediction results using PageRank centrality with different number of features (in this case, brain ROIs). The number of features affects the prediction accuracy; in general, too complex a model (i.e. a model with too many features) results in poor prediction performance due to over-fitting of the test data set. Figure 5-6A shows prediction accuracy with different numbers of features. SVR with PageRank centrality achieved the highest prediction accuracy when six features were selected for SVR inputs (*rho* = 0.59, $p = 0.001$). Skipped-correlation analysis, which offers robust computation of correlation between variables (Pernet *et al.*, 2012), validated the results that ROI centrality predicted individual

switch frequency at significant level (Table 5-1) and this indicates SVR achieved robust prediction. I also asked if the no-report duration can be predicted from the ROI PageRank centrality. I found that prediction accuracy was not significant across different numbers of features used for SVR (Figure 5-6B, Table 5-2).

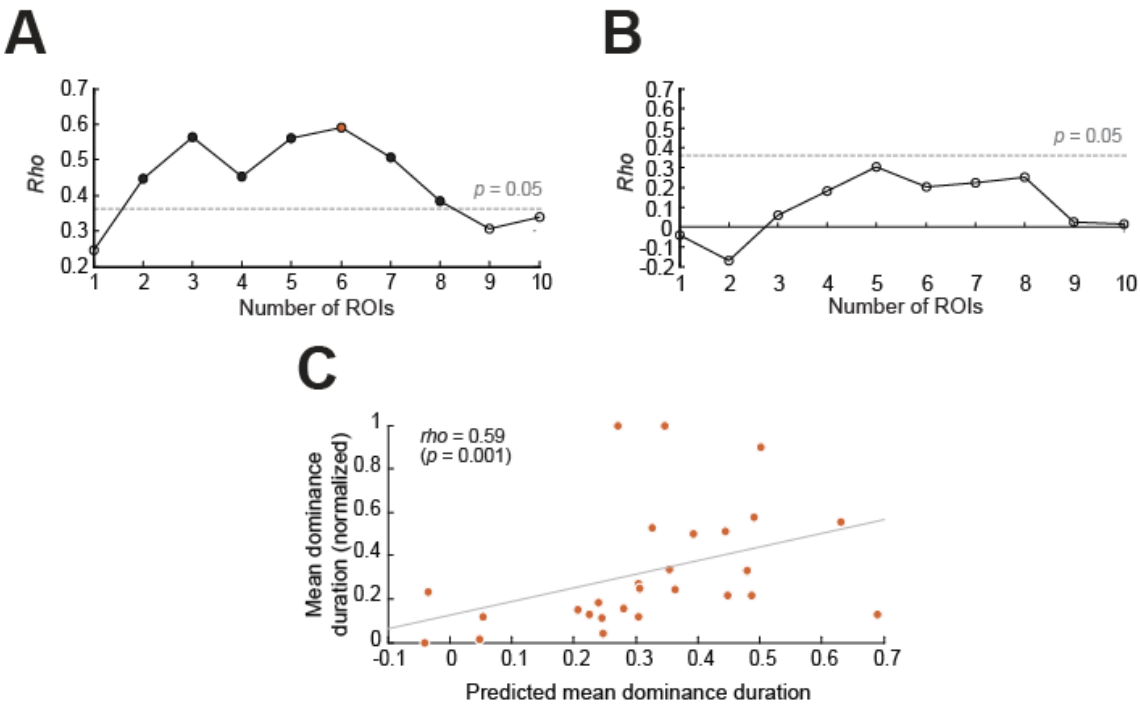

Figure 5-6 SVR prediction results.

(A) The graphs show correlation (*rho*) between prediction and mean dominance duration. The filled circles indicate that p-value for Spearman correlation *rho* were smaller than 0.05. Prediction accuracy was highest when 6 ROIs were used for prediction (orange circle). (B) I also tested whether length of mixed percept periods (duration without any button response) can be predicted from PageRank centrality but and did not find significant ($P<0.05$) correlation between prediction and behavioral data. (C) The figure shows the fit between SVR prediction and behaviorally measured mean dominance duration. (number of features: 6). Note that individual mean dominance duration is rescaled between 0 to 1 (see Methods section for details). SVR revealed that centrality measures successfully predict individual mean dominance duration ($rho = 0.59$, $p = 0.001$).

Table 5-1 SVR prediction accuracy for mean dominance duration

ROI	Spearman's correlation		Spearman's skipped correlation		
	Rho	*p*	*Rho*	CI	
1	0.25	0.22	0.30	-0.13	0.64
2	0.45	0.02	0.53	0.16	0.79
3	0.56	0.00	0.64	0.34	0.85
4	0.45	0.02	0.50	0.11	0.79
5	0.56	0.00	0.65	0.29	0.87
6	0.59	0.00	0.68	0.33	0.89
7	0.51	0.01	0.59	0.22	0.82
8	0.38	0.05	0.41	0.02	0.72
9	0.31	0.12	0.26	-0.17	0.66
10	0.34	0.08	0.42	0.01	0.75

Table 5-2 SVR prediction accuracy for no-report duration.

ROI	Spearman's correlation		Spearman's skipped correlation		
	Rho	*p*	*Rho*	CI	
1	-0.04	0.84	-0.07	-0.57	0.45
2	-0.17	0.39	-0.23	-0.65	0.21
3	0.06	0.76	0.15	-0.34	0.64
4	0.18	0.36	0.04	-0.47	0.49
5	0.30	0.12	0.17	-0.32	0.56
6	0.20	0.31	-0.01	-0.46	0.48
7	0.22	0.26	0.02	-0.42	0.44
8	0.25	0.20	0.10	-0.31	0.51
9	0.03	0.90	-0.15	-0.59	0.28
10	0.01	0.95	-0.21	-0.66	0.31

We also tested if other centrality measures were predictive for bistable perception performance (Figure 5-7A). Degree centrality showed high prediction accuracy but closeness centrality and betweenness centrality were not predictive for mean dominance duration. I also computed skipped Spearman's correlation (Figure 5-7B) and the results were almost the same. One notable difference is the number of features used for the best prediction performance with degree centrality; skipped *rho* indicated that degree centrality achieved highest prediction accuracy when two features were used instead of one feature. This possibly suggests that degree centrality, which is relatively simple and treat all edges the same (cd. PageRank

centrality puts weight on the nodes connected to hub nodes), might be vulnerable to the noise or outliers when used for SVR predictors.

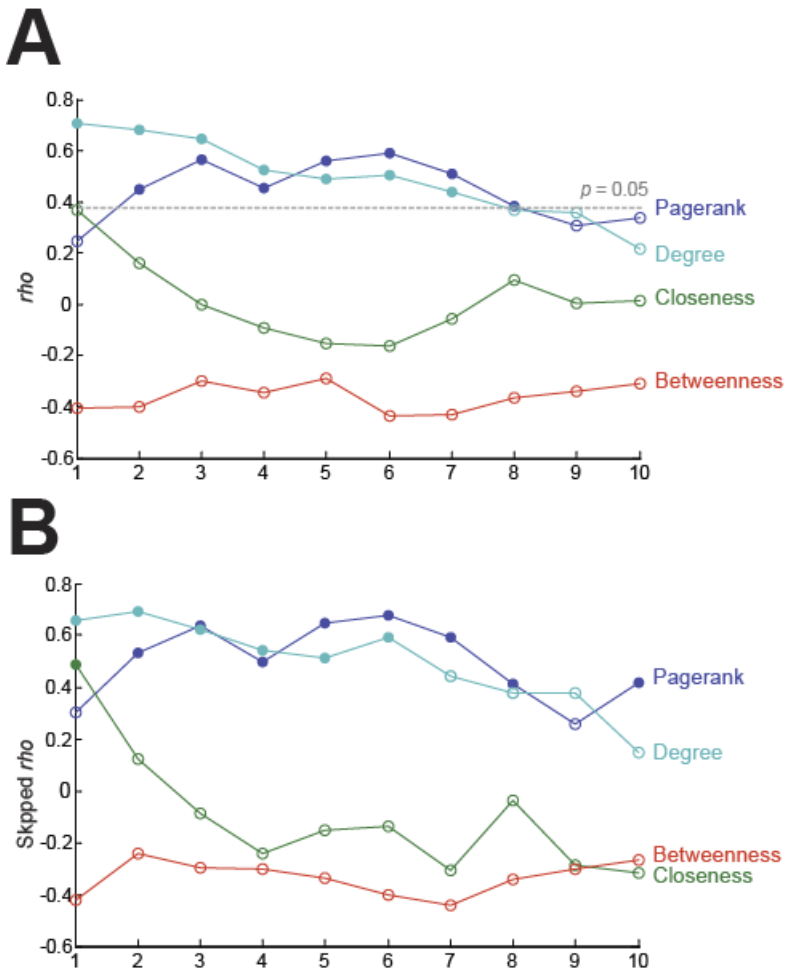

Figure 5-7 SVR prediction accuracy with different centrality measures.

(A) Prediction accuracy was evaluated with Spearman's *rho.* (B) In addition to *rho*, I also computed skipped Spearman's correlation, which allows us to eliminate outliers and obtain robust correlation.

Next, I asked if the degree centrality, which was predictive for mean dominance duration, also predicted no-report duration during the task. Figure 5-8 shows the correlation between predicted and measured individual non-report duration. As shown in Figure 5-6B, PageRank centrality failed to predict this behaviour. On the other hand, combination of ROI degree centrality showed marginally significant prediction performance in 4 ROI condition (Figure 5-8A, $rho = 0.38$, $p = 0.05$).

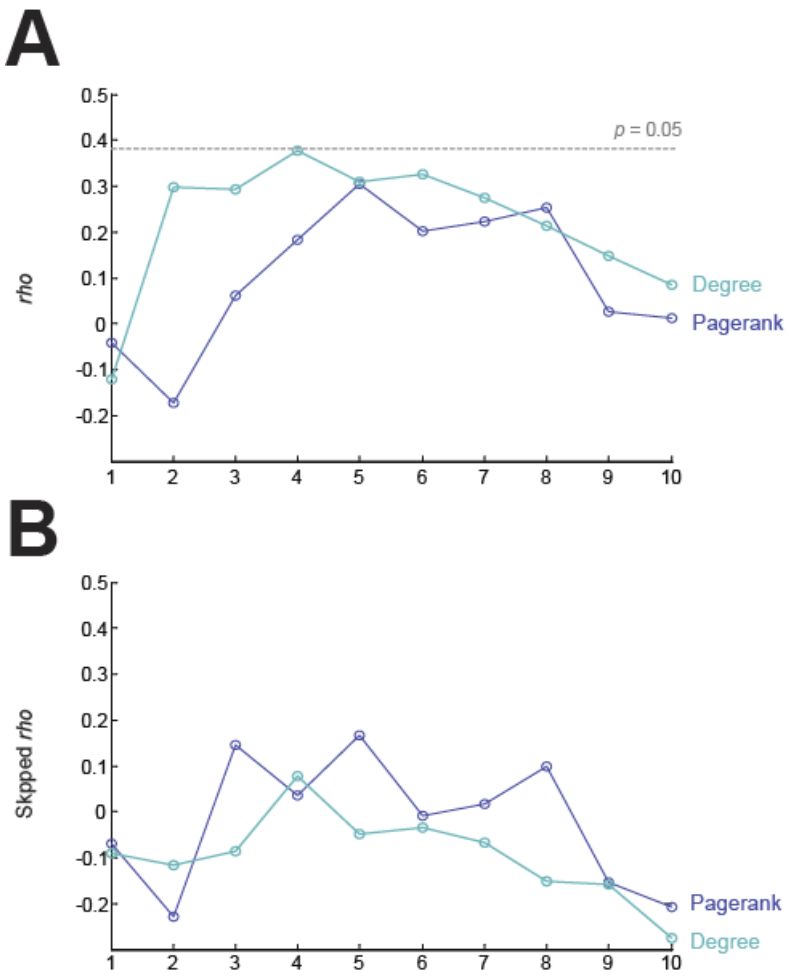

Figure 5-8 SVR with degree centrality for no-report duration.

(A) Prediction accuracy was evaluated with Spearman's *rho*. Rho for degree centrality was almost significant when four ROIs were used as SVR input. (B) Prediction accuracy decreased when skipped-rho was used to remove outlier effect in prediction.

5.4.3 Brain regions predictive for behavior

Table 5-3 and Figure 5-9 depict ROIs and averaged SVR weight for successful SVR prediction with PageRank centrality. Here I listed all features used for the SVR prediction at least one time; feature selection was performed for each training/test dataset (27 times in total) and zero was assigned as ROI weight when the feature was not selected for SVR input. ROI weight shows contribution of the ROI for the prediction. Positive ROI weight means that higher PageRank centrality of that ROI lead to more frequent switches, whereas negative weight indicates that higher PageRank

centrality lead to less frequent switches. ROI labels were derived from ICBM template with PickAtlas based on (*x*, *y*, *z*) coordinate of each ROI (Maldjian *et al.*, 2003; Maldjian *et al.*, 2004). ROIs from occipital areas, insula, and parietal areas were consistently used for the SVR prediction regardless of different number of features. Especially, right postcentral gyrus (ROI 25), left insula (ROI 73), and left lateral occipitotemporal gyrus (ROI 248) were consistently selected as informative ROIs in the feature selection stage across different test datasets for SVR, suggesting that hubness of these regions would account for individual difference in mean dominance duration.

Table 5-3 SVR ROI list for PageRank centrality

2 ROIs						
ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
25	29	-39	59	postcentral gyrus right	-0.250	Sensory/somatomotor Hand
70	-55	-9	12	postcentral gyrus left	-0.011	Auditory
73	-30	-27	12	insula left	-0.028	Auditory
164	26	-79	-16	lateral occipitotemporal gyrus right	0.015	Visual
248	-47	-51	-21	lateral occipitotemporal gyrus left	0.439	Uncertain
3 ROIs						
ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
25	29	-39	59	postcentral gyrus right	-0.241	Sensory/somatomotor Hand
70	-55	-9	12	postcentral gyrus left	-0.012	Auditory
73	-30	-27	12	insula left	-0.261	Auditory
164	26	-79	-16	lateral occipitotemporal gyrus right	0.018	Visual
200	-3	26	44	medial frontal gyrus left	0.012	Fronto-parietal Task Control
248	-47	-51	-21	lateral occipitotemporal gyrus left	0.405	Uncertain
4 ROIs						
ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
25	29	-39	59	postcentral gyrus right	-0.298	Sensory/somatomotor Hand
32	22	-42	69	postcentral gyrus right	-0.007	Sensory/somatomotor Hand
43	36	-9	14	insula right	-0.013	Sensory/somatomotor Mouth
70	-55	-9	12	postcentral gyrus left	-0.049	Auditory
73	-30	-27	12	insula left	-0.184	Auditory
164	26	-79	-16	lateral occipitotemporal gyrus right	0.159	Visual
200	-3	26	44	medial frontal gyrus left	0.014	Fronto-parietal Task Control
219	2	-24	30	cingulate region right	-0.003	Memory retrieval
248	-47	-51	-21	lateral occipitotemporal gyrus left	0.355	Uncertain
5 ROIs						
ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
25	29	-39	59	postcentral gyrus right	-0.248	Sensory/somatomotor Hand

32	22	-42	69	postcentral gyrus right	-0.008	Sensory/somatomotor Hand
43	36	-9	14	insula right	-0.009	Sensory/somatomotor Mouth
64	-38	-33	17	supramarginal gyrus left	-0.003	Auditory
67	43	-23	20	supramarginal gyrus right	0.002	Auditory
70	-55	-9	12	postcentral gyrus left	-0.188	Auditory
73	-30	-27	12	insula left	-0.097	Auditory
146	-28	-79	19	superior occipital gyrus left	0.009	Visual
152	43	-78	-12	lateral occipitotemporal gyrus right	0.008	Visual
164	26	-79	-16	lateral occipitotemporal gyrus right	0.221	Visual
200	-3	26	44	medial frontal gyrus left	0.014	Fronto-parietal Task Control
217	26	50	27	middle frontal gyrus right	0.007	Salience
219	2	-24	30	cingulate region right	-0.001	Memory retrieval
248	-47	-51	-21	lateral occipitotemporal gyrus left	0.316	Uncertain

7 ROIs

ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
2	27	-97	-13	occipital pole right	0.012	Uncertain
25	29	-39	59	postcentral gyrus right	-0.235	Sensory/somatomotor Hand
32	22	-42	69	postcentral gyrus right	-0.007	Sensory/somatomotor Hand
38	-16	-46	73	postcentral gyrus left	-0.008	Sensory/somatomotor Hand
43	36	-9	14	insula right	-0.010	Sensory/somatomotor Mouth
64	-38	-33	17	supramarginal gyrus left	-0.007	Auditory
67	43	-23	20	supramarginal gyrus right	-0.004	Auditory
69	-53	-22	23	N/A	-0.029	Auditory
70	-55	-9	12	postcentral gyrus left	-0.223	Auditory
73	-30	-27	12	insula left	-0.066	Auditory
135	4	-48	51	superior parietal lobule right	0.021	Memory retrieval
137	-10	11	67	superior frontal gyrus left	0.004	Ventral attention
146	-28	-79	19	superior occipital gyrus left	0.057	Visual
152	43	-78	-12	lateral occipitotemporal gyrus right	0.005	Visual
153	-47	-76	-10	lateral occipitotemporal gyrus left	0.006	Visual

164	26	-79	-16	lateral occipitotemporal gyrus right	0.218	Visual
200	-3	26	44	medial frontal gyrus left	0.044	Fronto-parietal Task Control
204	31	33	26	NA	0.001	Salience
217	26	50	27	middle frontal gyrus right	0.008	Salience
219	2	-24	30	cingulate region right	-0.005	Memory retrieval
248	-47	-51	-21	lateral occipitotemporal gyrus left	0.271	Uncertain

8 ROIs

ROI	coordinate			ICBM label	Weight	Network label
	x	y	z			
2	27	-97	-13	occipital pole right	0.013	Uncertain
17	-7	-21	65	medial frontal gyrus left	-0.004	Sensory/somatomotor Hand
25	29	-39	59	postcentral gyrus right	-0.232	Sensory/somatomotor Hand
32	22	-42	69	postcentral gyrus right	-0.010	Sensory/somatomotor Hand
38	-16	-46	73	postcentral gyrus left	-0.008	Sensory/somatomotor Hand
43	36	-9	14	insula right	-0.010	Sensory/somatomotor Mouth
64	-38	-33	17	supramarginal gyrus left	-0.017	Auditory
67	43	-23	20	supramarginal gyrus right	0.005	Auditory
69	-53	-22	23	N/A	-0.033	Auditory
70	-55	-9	12	postcentral gyrus left	-0.226	Auditory
73	-30	-27	12	insula left	-0.069	Auditory
135	4	-48	51	superior parietal lobule right	0.031	Memory retrieval
137	-10	11	67	superior frontal gyrus left	0.005	Ventral attention
146	-28	-79	19	superior occipital gyrus left	0.065	Visual
152	43	-78	-12	lateral occipitotemporal gyrus right	0.004	Visual
153	-47	-76	-10	lateral occipitotemporal gyrus left	0.007	Visual
163	-42	-74	0	inferior occipital gyrus left	-0.001	Visual
164	26	-79	-16	lateral occipitotemporal gyrus right	0.219	Visual
200	-3	26	44	medial frontal gyrus left	0.044	Fronto-parietal Task Control
204	31	33	26	NA	0.001	Salience
217	26	50	27	middle frontal gyrus right	0.017	Salience
219	2	-24	30	cingulate region right	-0.0002	Memory retrieval

248	-47	-51	-21	lateral occipitotemporal gyrus left	0.2541	Uncertain

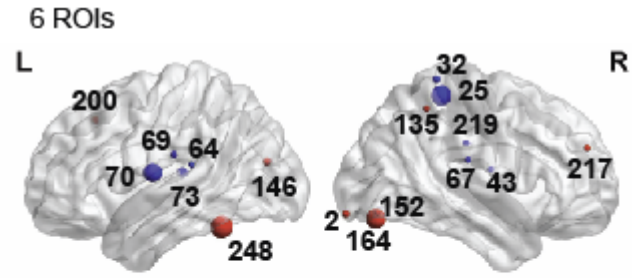

Figure 5-9 ROIs used for SVR prediction (PageRank centrality, 6 ROIs were used as features).

Spheres represent ROIs used for SVR prediction, color of the spheres indicates sign of weights (red means positive weight while blue means negative weight). Weight of the ROIs is represented by size of the sphere. Numbers indicate ROI index (See Table 5-3 for details).

5.4.4 Targeted network attack analysis

Having confirmed that ROI centrality could predict individual mean dominance duration, I next explored to what degree each network specifically contributed to prediction accuracy. I employed a 'targeted network attack' approach, removing all nodes and edges of a targeted node and then investigating how much MSE changed compared to a no-attack condition. This approach enabled us to elucidate which intrinsic networks (along with its connections to other network regions) are important for achieving high prediction performance.

Figure 5-10 shows MSE changes resulting from targeted attack of each network. I tested whether targeted attack of these intrinsic networks increased MSE compared to a no-attack condition (i.e. when all ROIs' connections were included). MSE was computed when different numbers of features was used; for statistical tests, MSE was averaged over conditions where SVR prediction accuracy was above chance level (r = 0.38). Nonparametric tests for results of the SVR with PageRank centrality revealed that attack of FPN caused significant or marginally-significant

increment in MSE (Wilcoxon signed rank test, VN, $Z = 2.48$, $p = 0.013$, $p < 0.10$ after Bonferroni correction for multiple comparison; FPN, $Z = 2.76$, $p = 0.006$, $p < 0.05$ after Bonferroni correction) In contrast, attack to DMN, VAN, and DAN did not affect MSE (DMN, $Z = -0.22$, $p = 0.83$, n.s.; VAN, $Z = -1.51$, $p = 0.13$, n.s.; DAN, $z = 0.22$, $p = 0.83$, n.s.). To validate the results, I also performed Friedman test for MSE changes compared to no-attack condition and asked if increment or decrement in MSE were different across five targeted network attack conditions. I found significant effect in attacked network (Friedman's test, $\chi^2_3 = 22,4$, $p = 0.00017$) and post-hoc comparison revealed significant difference between DMN and VN, DMN and FPN, and VN and VAN (detailed results are available in Table 5-4), suggesting removing VN and FPN from the intrinsic network deteriorated prediction accuracy compared to the attack to other networks.

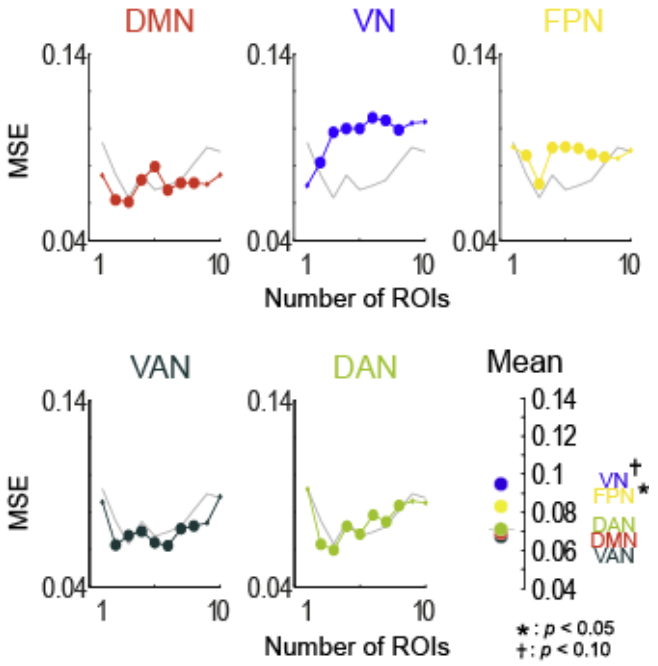

Figure 5-10 Network attack and mean squared error (MSE).

We interrogated changes in MSE due to targeted network attack with different centrality measures ((A) PageRank centrality (B) eigenvector centrality). Bigger markers indicate that prediction accuracy was better than chance at that number of ROIs. Gray lines show MSE in no-attack condition. 'Mean' indicates MSE averaged over conditions where SVR prediction accuracy was above chance in no-attack baseline condition (2, 3, 4, 5, 6, 7, and 8 ROIs were elected as input features,).

Table 5-4 Post-hoc multiple comparison for Friedman test

Comparison		95% CI		*p*
DMN	VN	-4.729	-0.871	0.001
DMN	FPN	-4.129	-0.271	0.019
DMN	VAN	-2.329	1.529	1.000
DMN	DAN	-3.029	0.829	1.000
VN	FPN	-1.329	2.529	1.000
VN	VAN	0.471	4.329	0.007
VN	DAN	-0.229	3.629	0.162
FPN	VAN	-0.129	3.729	0.109
FPN	DAN	-0.829	3.029	1.000
VAN	DAN	-2.629	1.229	1.000

5.5 Discussion

5.5.1 Overview of the findings

I investigated whether resting state brain network properties could predict behavioral performance in a separate bistable perception task, and then asked which of these networks had the highest predictive value. In particular, I tested which of five intrinsic networks including DMN, VN, FPN, VAN, and DAN were the most predictive of perceptual switch frequency. my results indicated that centrality measures successfully predicted individual switch frequency (measured as individual mean dominance duration) in bistable perception and that removing FPN and VN edges from the network decreased prediction accuracy (i.e. increased MSE), suggesting that task-independent and intrinsic functional connectivity supports the task-evoked dynamics of bistable perception.

5.5.2 Predict individual behavioral performance using ROI centrality measures in task-independent resting state activity

I demonstrated that ROI centrality measures computed from task-independent resting state activity could predict an individual

behavioural trait in a visual perception task. The organization and structure of brain networks have been investigated using anatomical connectivity, functional connectivity, and effective connectivity (Friston *et al.*, 2013) and graph theoretical approaches have been used for brain imaging data to characterize more detailed properties of brain intrinsic networks by quantifying the collective influence of each node on the entire network. Typically, functional connectivity and effective connectivity approaches consider interactions between two regions, or between a seed ROI and each voxel. In contrast, graph measures consider the influence of each node and thus enable us to understand the architecture of a complex system, such as the brain, in a different way. This approach has been successful in predicting personal traits (van den Heuvel *et al.*, 2009; Cole *et al.*, 2012; Lord *et al.*, 2012; Warren *et al.*, 2014) and describing differences between patient and healthy-control brain networks (Chennu *et al.*, 2014).

In bistable perception, inter-individual differences in perceptual switch frequency vary substantially across participants and an individual difference approach (Kanai & Rees, 2011) shows that anatomical structure (Kanai *et al.*, 2010; Kanai *et al.*, 2011; Shimono & Niki, 2013), local GABA-concentration level (van Loon *et al.*, 2013), and effective connectivity (Megumi *et al.*, 2015a) all predict individual variability in behavioral switch frequencies (or their reciprocal, mean dominance). In this study, I adopted a novel approach by applying graph theory to resting state (rather than task-related) data and found that centrality measures, which describe the efficiency of information flow within the network, predicted individual switch frequency. This further validates the utility of graph theory for brain imaging data analyses, as well as showing that measures of intrinsic connectivity can predict task-related performance for bistable perception.

5.5.3 Comparison of centrality measures

In this study, I found PageRank centrality of ROIs was informative features for SVR prediction and skipped correlation analysis revealed that the results were tolerant of outliers. Although Degree centrality also showed

high prediction accuracy, it also showed high prediction accuracy for no-report duration. Previous study showed that degree centrality is less consistent across different scanning runs compared to eigenvector centrality, which is a similar measure to PageRank centrality, and it might be due to physiological noise (Lohmann *et al.*, 2010). Degree centrality is quite simple measure; in this study, existence of edges between regions were simply determined with the strength of functional connectivity between the regions, and global physiological noise (e.g. breathing, heartbeat, eye blinks etc) affect connectivity pattern and that would be a possible cause for the spurious prediction accuracy for no-report duration.

Closeness centrality and betweenness centrality failed to predict individual bistable perception performance. Degree centrality and PageRank centrality is somewhat similar; both accounts for the degree of the node, whereas closeness centrality and betweenness centrality use shortest path information. Such information would be useful to predict variance in human behavior during task execution (Ekman *et al.*, 2012), where dynamic reconfiguration of intrinsic brain networks occurs.

5.5.4 Informative regions for SVR prediction and possible functional role

Current data showed that the centrality of occipital areas, parietal areas, and insula (specifically, right postcentral gyrus, left insula, and left lateral occipitotemporal gyrus), predicted individual SFM switch frequency and but not trivial individual variance in task behaviour (no-report duration). The right postcentral gyrus (ROI 25) was close to anterior superior parietal lobule, a region causally mediating stability of perceptual experience during bistable perception (Kanai *et al.*, 2011) and more recent work also suggests this region is functionally involved in bistable perception (Megumi *et al.*, 2015a). Activation in anterior insula during perceptual switches has been reported repeatedly (e.g. Lumer *et al.* (1998); Sterzer and Kleinschmidt (2010)). Left posterior insula (ROI 73) is thought to relate to sensorimotor information processing (Menon & Uddin, 2010) and integration of crossmodal sensory information (Calvert, 2001). Left lateral

occipitotemporal gyrus (ROI 248) is a region involved in visual hallucinations (Ffytche *et al.*, 1998; Meppelink *et al.*, 2009) and synaesthesia (Rich *et al.*, 2006). Such work would predict that the left lateral occipitotemporal gyrus may contribute to conscious visual experience.

5.5.5 Visual and fronto-parietal network in bistable perception

The 'targeted network attack' analysis complements the claim that fronto-parietal areas, together with visual network, are involved in resolving ambiguity in bistable perception tasks. Although many previous studies have suggested that fronto-parietal cortices show greater activation time-locked to perceptual switches (see Rees (2007)) and effective connectivity studies revealed interaction between visual cortex and fronto-parietal area in bistable perception (Weilnhammer *et al.*, 2013; Megumi *et al.*, 2015a) , some recent studies suggest that activation in fronto-parietal regions might reflect a response to ambiguity in visual stimuli, rather than a mechanism for the resolution of this ambiguity (Knapen *et al.*, 2011). Here I investigated the importance of FPN with resting state data and multivariate analysis (i.e. SVR). The advantage of using resting state data for SVR prediction is its task-independence. Active report of perceptual content and the need to direct attention to a bistable task influence percepts per se (Meng & Tong, 2004; Zhang *et al.*, 2011), and hence removing confounds or eliminating the 'report effect' in task fMRI data is very challenging (but see Frässle *et al.* (2014)). In contrast, resting state measurements are not associated with any active task. Functional connectivity manifest through co-variation between activity in distant regions is thought to reflect a fundamental property of brain networks (Fox & Raichle, 2007), mainly supported by anatomical connections (Vincent *et al.*, 2007; Honey *et al.*, 2009). This approach enables us to associate behavior or cognition with the intrinsic network structure of the brain. I interrogated the importance of such an intrinsic network for prediction of perceptual switch frequency and demonstrated that removing edges connected to FPN

and VN increased MSE, which suggests that connectivity in this part of the network plays a significant role in determining perceptual switch frequency. Based on the current results and previous findings, I speculate that individual differences in switch frequency can be explained by differences in more general and inherent visual introspection ability. Frässle *et al.* (2014) claimed that fronto-parietal activation in bistable perception is induced by observers' introspection about visual input, indicating a possible role for FPN in visual introspection. Indeed, VBM analysis shows that gray matter volume of an FPN region (anterior prefrontal cortex, BA10) correlates with visual metacognitive ability (Fleming *et al.*, 2010), providing further evidence for the hypothesis that FPN mediates visual introspection, self-monitoring of one's perceptual experience, and attention, which are indispensable to resolve ambiguity in sensory information (Zhang *et al.*, 2011).

5.5.6 Default mode network in bistable perception

I could not find any influence of resting state activity in the DMN on behavior during bistable perception (as indicated by the lack of any significant change in MSE after DMN attack). DMN is a set of brain regions showing more activation at rest than during goal-directed tasks (Raichle *et al.*, 2001). The BOLD signal in DMN regions is anti-correlated with that in regions that participate in task-related networks, such as the dorsal-attention network (DAN) and FPN (Fox *et al.*, 2005; Hellyer *et al.*, 2014). Freeman *et al.* (2012) showed that posterior cingulate, which is known as a hub region of the default-mode network, was deactivated prior to perceptual switches. This likely indicates DMN's involvement in monitoring or allocation of visual attention, which is thought to be necessary for resolving ambiguity in visual information. Furthermore, the strength of coupling between occipital areas and the two networks, DMN and FPN, can account for perceptual content during bistable perception (Karten *et al.*, 2013). Although these results point to involvement of the DMN in bistable perception, my findings suggest that connection between three informative

ROIs (post central gyrus right, posterior insula left, and lateral occipitotemporal gyrus left) and DMN did not directly contribute to predict individual differences in switch frequency during bistable perception.

The results do not exclude the possibility of DMN's functional involvement in visual information processing during bistable perception. The neuronal dynamics underlying bistable perception could be understood as interplay between different regions, and DMN and FPN have reciprocal influences, to toggle the mode of the brain (Fransson, 2005; Chadick & Gazzaley, 2011). Taken together, during bistable perception, DMN interacts with FPN cooperatively to help form a unitary and clear percept, and the activation of DMN may be a reflection of signal suppression for unnecessary information or noise in perception. This hypothesis should be tested by analyzing connectivity architecture between FPN and DMN during bistable perception task.

5.5.7 Limitation of this study

A potential limitation of my study is the size of the dataset. The number of participants (N=27) was relatively small compared to the number of features I employed for model fitting (254 ROIs) and therefore my findings might change if a larger number of participants were studied. To demonstrate the robustness of my data, I verified the results using different numbers of features and employed robust correlation analysis (skipped-correlation). These additional results all supported the main conclusion that the ROI centrality measure in task-independent resting state activity were predictive for switch frequency in bistable perception. Nevertheless, a study with a larger dataset and different bistable perception paradigms (e.g. binocular rivalry and apparent motion) would be potentially interesting in further elucidating the relationship between intrinsic connectivity and perceptual switching during bistability.

5.6 Conclusion

I showed that resting state intrinsic functional connectivity successfully predicted individual switch frequency in bistable perception. Graph theoretical analyses revealed that centrality of the right precentral gyrus, left insula, and left lateral occipitotemporal gyrus were the most informative, and the coupling between these ROIs and FPN/VN were important for successful prediction. These results give further support for the functional importance of fronto-parietal cortex and visual cortex for mediating visual perceptual switches.

Chapter 6 PERCEPTUAL PRIORS DETERMINE THE NEURAL AND BEHAVIORAL DYNAMICS OF VISUAL BISTABLE PERCEPTION

6.1 Summary of this chapter

The experiments I described in chapter 4 and 5 revealed that effective connectivity and functional connectivity between specific human brain areas are related to individual difference in bistable perception performance. However, it is unclear what determines the contents of visual perception on a trial-by-trial basis. Predictive coding theories explored in the previous chapters propose that the brain uses prior expectations when resolving such ambiguity, and that these prior expectations can be derived from sensory experience. Here, by recording magnetoencephalographic (MEG) signals from participants viewing ambiguous structure-from-motion (SFM) stimuli, I explored how experimental manipulation of such prior expectations affected both subjective perception and corresponding neural representations. Behavioural analysis showed that prior expectations strongly influenced subsequent ambiguous perception. I used a hierarchical Bayesian model to estimate how such prior expectation might be represented in the brain, and found that key parameters of the model were correlated with signals from fronto-parietal cortex, medial parietal and visual cortex. Taken together, the current results reveal how prior sensory expectations are encoded and updated in the human brain and shape the contents of conscious experience.

6.2 Introduction

6.2.1 How do we resolve ambiguity in visual bistable perception stimuli?

Experiments in chapter 4 and 5 revealed that effective connectivity and functional connectivity are related to individual difference in bistable

perception performance (mean dominance duration, switch frequency). A remaining question is what determines the subjective percept in a bistable perception task on a trial by trial basis, rather than the determinants of individual differences. The experiments in chapter 4 and 5 took individual difference approach, and therefore it is still not clear what determines trial-by-trial dynamics of visual perception.

6.2.2 Hypothesis, experimental paradigm, and analysis strategy

Our perception often reflects the consequences of resolving ambiguity in visual input. Visual signals from the retina are inherently ambiguous with respect to their sources in the environment, but the brain needs to solve this inverse problem to guide behaviour (Poggio & Koch, 1985). Our perception is also influenced by information derived from its history when making perceptual decisions (Denison *et al.*, 2011; Wyart *et al.*, 2012a; Wyart *et al.*, 2012b; Akaishi *et al.*, 2014; Fischer & Whitney, 2014). Predictive coding theory (Rao & Ballard, 1999; Clark, 2013) proposes that the brain resolves perceptual ambiguity by combining sensory information with expectation (priors) for the world and such prior expectation is subsequently updated upon the observer's perceptual experience (Hohwy *et al.*, 2008; Clark, 2013).

The statistical properties of stimulus sequences may influence our visual system when the brain needs to resolve ambiguity and form visual awareness. Spherical-shaped structure-from-motion (SFM) bistable perception stimuli elicit rightward rotation or leftward rotation of a sphere (Jiang *et al.*, 1998) and have been used to study nature of visual perception. The stability of one of the two percepts shows marked individual difference and the volume of gray matter in parietal cortex underpins such individual differences (Kanai *et al.*, 2010; Kanai *et al.*, 2011). This suggests the possibility that front-parietal cortex contributes to stabilization of a percept by generating expectation for interpretation of visual inputs.

Although it has been suggested that top-down information processing, such as voluntary control, attention, memory, affects interpretation of bistable stimuli (van Ee *et al.*, 2005; Denison *et al.*, 2011;

Gorlin *et al.*, 2012; Panichello *et al.*, 2012)), there is currently a lack of empirical evidence that such expectation indeed affects interpretation of bistable perception stimuli. The subjective percept in a visual task is determined by both top-down and bottom-up signals (Kornmeier *et al.*, 2009), and the time-course of subjective percepts shows time-varying and complex dynamics (Brascamp *et al.*, 2008). This makes it difficult to establish how the prior expectation, which is build up through accumulation of previous percepts, affects interpretation of ambiguous stimuli.

To overcome this difficulty and investigate the effect of prior expectation, I used SFM stimuli to determine whether the prior expectation biased interpretation of ambiguous visual information and neural substrates of such expectation. I developed a new bistable perception paradigm incorporating implicit statistical learning. In the task, participants were asked to report their percept while a bistable SFM stimulus was intermittently presented. Importantly, trials could either be primes or probes. On prime trials, SFM stimuli were disambiguated by depth information (binocular disparity) intended to influence the subjective percept which would be less ambiguous. On probe trials, stimuli were standard SFM stimuli and were thus completely ambiguous. I used probe trials as a way of establishing participants' belief for incoming sensory stimuli. Responses on probe trials should reflect 'expectation'; the input itself is completely ambiguous without sensory noise so that participants should follow their prior expectation and report the expected direction as their perceived direction. To track and quantify such 'expectation' signals, I employed a hierarchical Gaussian filtering (HGF) model. This model describes participants' beliefs about their environment using a three hierarchy Bayesian model. I also compared the ability of this model to capture behavior with established Rescorla-Wagner and Sutton models. The dynamics of bistable perception change over time (Brascamp *et al.*, 2008), so I compared models with and without volatility, which represents variance of changes.

6.3 Methods

6.3.1 Participant and exclusion criteria

Twenty-four healthy adults (18 female and 6 male; aged 18 – 34, mean ± SD: 24.6 ± 4.8) participated in the MEG study and data from twenty participants (16 female and 4 male; aged 18 – 34, mean ± SD: 24.1 ± 4.8) were included in the analysis. They were all right-handed and had normal (16 participants) or corrected-to-normal vision with contact lenses (4 participants). None had any prior experience with experiments involving structure-from-motion stimuli. All participants provided written informed consent. UCL ethics committee approved the experimental procedures.

Four participants were removed from subsequent analysis; one participant had head motion >2cm during MEG session, two participants whose correct rate on the behavioural task was below chance, and one participant whose behaviour did not fit the Bayesian model.

6.3.2 Apparatus

Visual stimuli were presented on the screen attached to the MEG. They were projected onto a 19" screen at a resolution of 1024 × 768 pixels and at a refresh rate of 60 Hz using a JVC D-ILA, DLA-SX21 projector.

Participants viewed the stimuli through a plastic mirror stereoscope (Geoscope Pro, Stereo Aids). A black cardboard was attached to the stereoscope as a splitter. Viewing distance was approximately 60 cm and the size of the screen was 43cm×32cm.

6.3.3 Stimuli

Spherical shape structure-from-motion (SFM) bistable perception stimuli (Kanai *et al.*, 2010) were created using Psychtoolbox3 (Brainard, 1997). The screen resolution was set to 1024×768.

On probe trials, 180 full-contrast small dots (diameter: 0.12 degree)

showing sinusoidal movement at an angular velocity of 120 degree/s were presented on the black background forming a spherical shape of 3.69 degrees diameter. I added a surrounding line square (4.79 degree) as well as a central fixation cross (0.55 degree in height and width) to each monocular image in order to help participants fuse the spheres.

On prime trials, disambiguated spheres with binocular disparity was presented on the screen. Size and rotation speed of the spheres were the same as those of probe trials. Binocular disparity was computed for each dot and the location of the dots were shifted accordingly (0-0.2 degree, depending on the dots' position in the sphere).

6.3.4 Experiment procedure

Participants were asked to perform a spherical-shape structure-from-motion task. After a fixation period (0.5 seconds), on each trial a spherical shape SFM stimulus appeared on the screen (1.5 seconds). Once the stimulus disappeared from the screen, participants reported the perceived direction of the rotation of the sphere by pressing one of two buttons during a response period of either 3 or 4.5 seconds. Button responses had to be made within the response period. If participants failed to give button response within the response period, the trial was treated as 'non-response' and not included in subsequent analyses.

Two different types of stimuli were presented; prime stimuli, which were disambiguated SFM by binocular disparity (depth information), and probe stimuli, which were ambiguous SFM. The purpose of prime stimuli was to manipulate participants' percept and let them learn environmental dynamics. SFM with binocular disparity (i.e. depth information) was presented for a number of prime trials (30 trials per block); adding binocular disparity to the stimuli decreased perceptual ambiguity so that subjective percept was stimulus-driven. In probe trial (18 trials per block), fully ambiguous structure-from-motion (without binocular disparity) was presented and I tested whether the subjective percept reported on probe trials was biased by prime stimuli.

Prior to MEG recording session, participants underwent 4-8 practice sessions (approximately 3 minutes for each session, 10-30 minutes in total). Each session consisted of 30 prime trials. The purpose of this practice session was to help participants to get used to the prime stimuli. Feedback for their response (correct or incorrect) was given to the participants as the change of fixation point. Participants performed practice up to four sessions. Ambiguous stimuli were never presented to the participant during the practice. Both the practice session and the MEG session were on the same day.

During the MEG recording session, participants viewed stereoscopic bistable stimuli presented intermittently for 1.5 seconds, and were asked to report the perceived direction of rotation (right or left). Participants were shown five right-biased blocks and five left-biased blocks at maximum and the order of the block condition was pseudo-randomized. One participant performed nine blocks (four right-biased blocks and five left-biased blocks) instead of ten due to lack of scanning time caused by MEG technical problems. Participants were instructed to refrain from movement as much as possible (e.g. eye blink, eye movement, and button response) during the stimulus presentation period to reduce any motion-induced artefact in the MEG signal.

After the experiment, participants were interviewed and asked if they noticed the existence of probe trials (ambiguous SFM). This confirmed that no participants realized the existence of probe stimuli explicitly. The main difference between practice session and MEG session was feedback for the prime trials, duration of the experiment, and experimental setup. They were asked if they noticed any difference in stimuli between practice session and MEG session but no one could point out that probe stimuli (ambiguous SFM stimuli) were added to the experiment.

6.3.5 Behavioural data analysis

Data analysis was performed using the MATLAB statistical toolbox and SPSS. Blocks where participant's correct rate was below 30% were excluded

from the analysis to ensure that participants were concentrated on the task (one block from nine participants, two blocks from one participant, and three blocks from one participant). On average, 9.4 (SD: 0.8) blocks per participant were used for the analysis.

To validate any bias effect (Figure 2A), the probability of reporting right or left in right- and left- biased blocks respectively was computed. As some participants showed inherent preferences for right or left choice, the bias effect ($Bias_{RightBlock}$ for right-block and $Bias_{LeftBlock}$ for left-block) was computed as follows:

$$Bias_{RightBlock} = P(Right|Probe_{RightBlock}) \quad - P(Right|Probe_{All})$$

$$Bias_{LeftBlock} = P(Left|Probe_{LeftBlock}) \quad - P(Left|Probe_{All})$$

Where $P(Right|Probe_{RightBlock})$ represents mean probability to report 'right' in right-biased block and $P(Right|Probe_{All})$ denotes mean probability to report 'right' in all probe blocks, and the same applies to $Bias_{LeftBlock}$. These indicate bias effect considers individual tendency (Carter & Cavanagh, 2007) in report in probe trials and corrected the tendency by subtracting such effect as baseline.

6.3.6 Bayesian modelling

To estimate the trajectory of participants' expectation for incoming sensory stimuli (i.e. probability that the sphere spinning towards right) and correlate it with MEG signals, I employed a Hierachical Gaussian Filtering model (HGF) and estimated each observer's belief about environmental causes with HGF toolbox (Mathys *et al.*, 2011; Mathys *et al.*, 2014).

The HGF toolbox enables us to model participants' behaviour with two complementary and related models; the perceptual model, describing the mapping from environmental cause to sensory input, and the response model, representing the mapping from observer's belief to response. This enabled us to explore the representation both of the world (perceptual model) and the choice behaviour (observation model) separately (Daunizeau *et al.*, 2010). The HGF model accounts for individual difference in

behaviour. It is known that behaviour in bistable perception task varies across individuals (e.g. (Kanai *et al.*, 2010)) and therefore it is essential to consider such variance and assign parameters for each participant.

We also tested a Rescorla-Wagner model and a hierarchical Bayesian model for a Gaussian observer model, a softmax function model, and a stick function model for observation model. All blocks were concatenated and model fitting was performed for each participant independently.

6.3.6.1 Perceptual models

The HGF model consists of three layers, representing percept (P), tendency to right-rotation ('Expectation', noted as E), and volatility of tendency to perceive right-rotation ('Volatility', noted as V). Volatility is equivalent to, x_3, expectation is x_2, and percept is x_1 in the original paper (Mathys *et al.*, 2011).

The first layer (P) is given as follows:

$$p(P|E) = s(E)^P\big(1 - s(E)\big)^{1-E} = Bernoulli(P; s(E)),$$

where $s(\cdot)$ represents a sigmoid function ($s(x) = \frac{1}{1+\exp(-x)}$).

The second layer ($x2$) is given as follows:

$$p\left(E^{(k)}\middle|E^{(k-1)}, V^{(k)}\right) = N(E^{(k)};\ E^{(k-1)}, \exp(\kappa V^{(k)} + \omega),$$

where $N(\cdot)$ denotes Gaussian distribution. κ, ω are parameters for Gaussian random walk.

Third layer can be obtained as follows:

$$p\left(V^{(k)}\middle|V^{(k-1)}, \vartheta\right) = N(V^{(k)};\ V^{(k-1)}, \vartheta),$$

where ϑ is a constant parameter for volatility.

Prediction-error is obtained as follows:

$$\mu_i^{k+1} - \mu_i^k \propto \psi_i^k - \delta_{i-1}^k$$
$$= \varepsilon_i^{(k)}$$

Where μ_i indicates mean of the distribution i. ψ_i^k, precision ratio of i at time k, is obtained as follows:

$$\psi_i^k = \frac{\widehat{\pi_{i-1}^{(k)}}}{\pi_i^{(k)}}.$$

Note that $\hat{x}$ is 'prediction' of the variable x. This means the precition ratio ψ_i^k is a precision about the input from the level below.

In this study, the main interests were Volatility and Expectation so these parameters and their precision-weighted prediction-error were used for the MEG analysis.

In addition to HGF, I tested Rescorla-Wagner model:

$$Prediciton^k = Prediciton^{k-1} + \alpha \times \delta^{k-1}$$

Here alpha is a learning rate (fixed parameter) and delta is a prediction error. Indeed this equation is similar to the HGF model, but HGF model offers more flexible representation for learning rate alpha and prediction and error.

6.3.6.2 Response models

In addition to the perceptual model, I compared three different response models. The response model describes how the observer generates decision or response based on stimulus input (Daunizeau *et al.*, 2010). I tested three different conditions; (1) Modelling perceptual trajectory without response model (2) Modelling perceptual trajectory with stick-function (3) Modelling perceptual trajectory with sigmoid function. Combinations of three perceptual models and three observation models were tested and the trajectory of the best model was used for MEG analysis

6.3.6.3 Variational inversion and Model selection

Posterior distributions $p(x(k), x|u(1, \dots k))$ were obtained through variational Bayesian (VB) inversion by maximizing the log-model evidence (LME). Based on log-model evidence, the best-fitting model was selected with Bayesian model selection approach (Stephan *et al.*, 2009) using SPM12 (spm_BMS.m).

6.3.7 MEG recording

MEG data were obtained at the Wellcome Trust Centre for Neuroimaging, University College London. MEG data were recorded in a magnetically shielded room with a 275 channel CTF Omega whole-head gradiometer system (VSM MedTech, Coquitlam, BC, Canada) at a 600 Hz sampling rate.

After participants were comfortably seated, head localizer coils were attached to the nasion and 1 cm anterior of the left and right outer canthus to monitor head movement during the recording sessions. During the MEG session electrooculography (EOG) was recorded bipolarly with two pairs of metal electrodes. Electrodes were placed 2 cm above the lateral canthus of the right eye and 2 cm below to the lateral canthus of the left eye, and on the supraorbital ridge of the left eye and on the left zygomatic bone. Reference signal was taken from the left hand. The impedance was maintained at less than 5KΩ for all participants.

6.3.8 MEG data analysis

6.3.8.1 Preprocessing

All MEG data analysis was carried out with Statistical Parametric Mapping software (SPM12, http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Each trial was epoched from -500ms to 1500ms from the stimulus onset. Each trial was baseline corrected using the average signal of trial (-500ms to 1500ms). 0.1Hz cut-off low-pass temporal filter (5th-order Butterworth filter) and 48-50Hz stop-band filter were applied to remove artefacts.

I used EOG data for artefact detection. The signal was taken from the difference in signal of two paired-electrodes sets (threshold was set to 300 μV). Trials contained eye-movement related artefact was rejected based on the criterion. In average, 17.0 trials (4 % of all trials) were rejected per participants (SD: 15.21, 1-47 out of 480 trials).

6.3.8.2 Source analysis

I used the multiple sparse prior (MSP) function implemented in SPM12. I sought to identify which brain areas represented expectation and its update depending on the history of subjective perception reported by each participate, and therefore I aimed to investigate focal MEG responses and compare them with previous studies rather than MEG sensor-space. The multiple-sparse prior algorithm (Friston *et al.*, 2008) provides a way to solve the M/EEG source-localization inverse problem by considering brain structure or fMRI activation pattern. In this study, I used the canonical brain template (SPM default brain template, contains 8196 vertex) to estimate the source and performed group-inversion, which assumes the same source across participants. The stimulus presentation period was divided into six time windows (100-300ms, 300m-500ms, 550-750ms, 750-950ms, 1000-1200ms, and 1200-1400ms) and the source of the MEG signal was computed for each time window. The source was estimated with greedy-search algorithm and inferred for single-trial basis in surface space.

6.3.8.3 MEG Statistical analysis

I first created regressors representing HGF model parameters (expectation, expectation-PE, volatility, and volatility-PE. Four participants were removed from model-based MEG analysis due to high correlation between parameters.

My hypothesis was that specific brain regions (most likely fronto-parietal regions) represent model parameters such as expectation and evidence accumulation and update of the model parameters would take place throughout stimulus presentation periods, and I sought brain regions whose responses correlated with model parameters.

The stimulus presentation period was divided into six periods to estimate the source of activation (see 6.3.8.2 "Source analysis" section) and performed 1st level analysis for each time window with four regressors (Expectation, Expectatio-PE, Volatility, and Volatility-PE). The contrast

images from six time windows were included for second level analysis, and I characterised which brain regions represented model parameters.

6.4 Results

6.4.1 Behavioural results

To elucidate how the expectation for incoming sensory stimuli evolved over trials and its neural signature, I asked participants to perform structure-from-motion visual perception task in the MEG scanner. The task was to report direction of sphere's rotation (rightward or leftward) (Figure 6-1A and 1B). Each experimental block consisted of 30 prime trials and 18 probe trials (Figure 6-1C). The prime stimuli were disambiguated with binocular disparity (i.e. depth cue) and there were six levels of rotational coherence (proportion of dots with right or left rotation) so that the observers were expected to report the primed direction (Figure 6-1D). On the other hand, probe stimuli were standard spherical SFM (i.e. no depth information was added to the sphere) and the stimuli were completely ambiguous. Inter-trial interval was 3 seconds or 4.5 seconds and the duration was randomized for each trial.

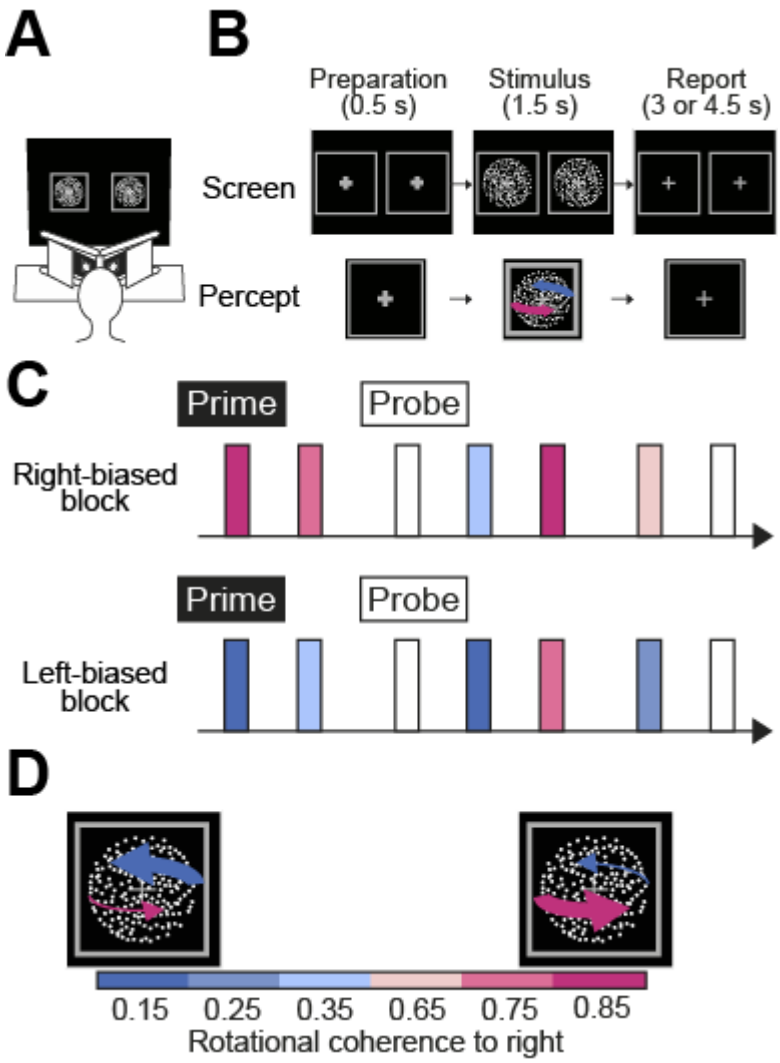

Figure 6-1 Experimental task.

(A) Participants were asked to perform structure-from-motion (SFM) task with mirror-stereoscope to achieve stereopsis. (B) Followed by preparation period (0.5secs), SFM stimuli appeared on the computer screen. Once the stimuli disappeared, participants reported their percept (sphere rotating rightward or leftward). (C) Each experimental block (consisting of 30 prime trials, 18 probe trials) was assigned to either right-biased block or left-biased block. In each block, 70% of prime trials (i.e. 21 trials) showed disambiguated sphere rotating to biased-direction (right for right-biased block, left for left-biased block). (D) Prime stimuli had six different strength of rotational coherence, 85% of dots with right-rotation disparity /15% with left-rotation disparity (magenta) to 15% of dots with right-rotation disparity / 85% with left-rotation disparity (cyan).

The correct rate in prime trials was 0.71 ± 0.11 in right-biased blocks and 0.68 ± 0.11 in left-biased blocks, and there was no significant difference

in the correct rate ($t(19) = 0.91$, $p = 0.37$). Correct rate increased as a function of rotational coherence (Figure 6-2A), and this indicates that the correct rate in prime trials was affected by the strength of sensory evidence. ROC analysis revealed that there was no significant difference in false-positive and false-negative rate between the two block conditions ($t((19) = 0.50$, $p = 0.62$). All these data indicated that experimental manipulation worked well, and that each participant's percept was guided by the depth information added to the sphere in prime trials.

Then I asked if there were any response biases in probe trials. In this study, I hypothesized that the participants formed 'prior expectation' for incoming sensory stimuli from statistical information derived from stimulus sequence and such prior expectation affected perceived direction of the sphere's rotation. I found that the subjective percept reported in right-bias blocks, where depth cue in 70% of prime trials indicated spheres rotated towards right, was significantly biased comparing to left-biased blocks (Figure 6-2B; Wilcoxon signed rank test, $Z = 2.64$, $p = 0.0084$).

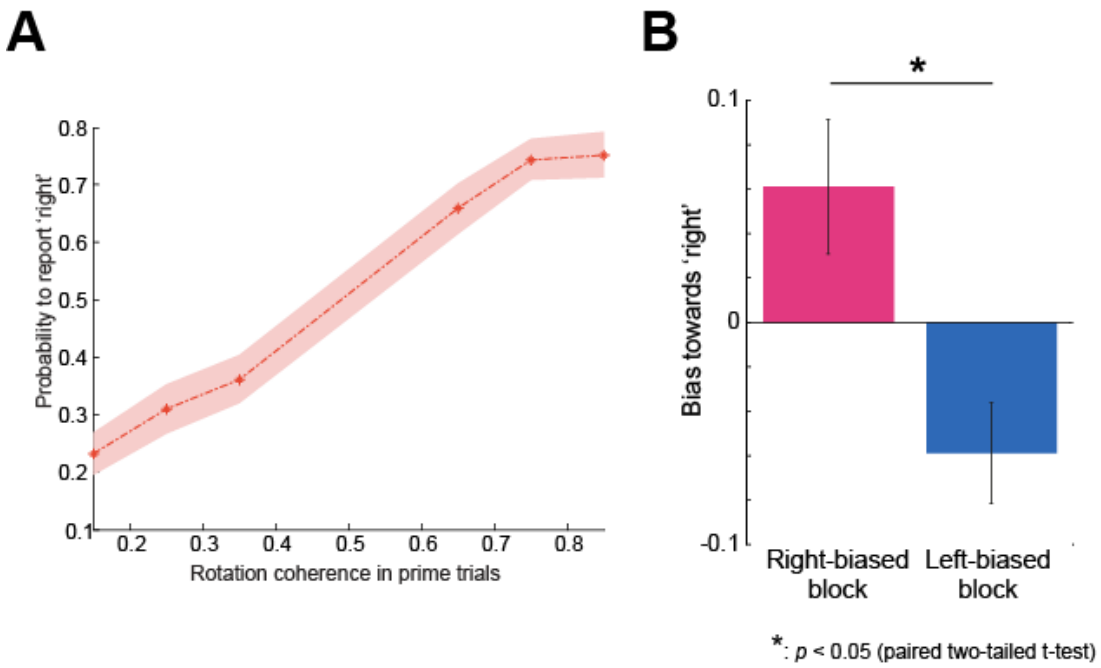

Figure 6-2 Behavioral data analysis.

(A) The plot shows probability of right response as a function of rotational coherence condition. In high-rotational coherence condition, where majority of the dots indicated the sphere was rotating towards right, participants were more likely to report 'right'. (B) I looked at response in probe trials, where SFM had no binocular disparity so that rotation of sphere was completely ambiguous. Wilcoxon signed rank test revealed that participant's percept in probe trial was biased towards biased direction (Z = 2.64, p = 0.0084)

6.4.2 Bayesian model analysis

Having confirmed the presence of bias effects in block conditions, I then fitted a model of participants' responses using the HGF toolbox (Mathys *et al.*, 2011). I reasoned that participants' expectations for incoming sensory stimuli and their interpretation (percept) should be affected by the history of percept and therefore I employed a Bayesian modelling approach to estimate the trajectory of the priors. I tested an HGF model (Mathys *et al.*, 2011; Mathys *et al.*, 2014), the Rescorla-Wagner model (Rescorla & Wagner, 1972), and the Sutton k1 model (Sutton), a Rescorla-Wagner model with variable learning rate. The HGF model contained three layers (Figure 6-3A); percept (the first layer), expectation, which represents tendency to perceive

right-rotation (the second layer), and volatility of the expectation, which express how likely the observer sticks to the current percept (the third layer). Note that I did not explicitly manipulate volatility in the experiment but including such a layer should be justified as the dynamics of bistable perception reports are not stable (Brascamp *et al.*, 2008). In addition, the HGF model computes parameters and trajectories for each individual based on their behavior; bistable perception shows variability across individuals (Kanai *et al.*, 2011) and therefore the HGF model is suitable for this behavioral paradigm.

Bayesian model comparison (Stephan *et al.*, 2009) revealed that the HGF model with optimal Bayesian observer best explained the trajectory of the participants' percepts (Figure 6-3B). The eexpectation computed with the Bayesian model successfully predicted participants' percept on probe trials (Figure 6-3C; data from one representative participant). Expectation (model parameter x2) were higher when participants responded to right comparing to left (Figure 6-3D, $t(19) = 4.78$, $p = 0.00013$). The bar graph represents averaged model parameter in probe trials in right-biased block and left-biased block. The response bias effect (Figure 2B) were also replicated from the model parameter x2 ($t(19) = 4.13$, $p = 0.0006$). Collectively, these results suggest that combination of stimulus and subjective percept forms expectations that subsequently shape interpretation of ambiguous sensory input.

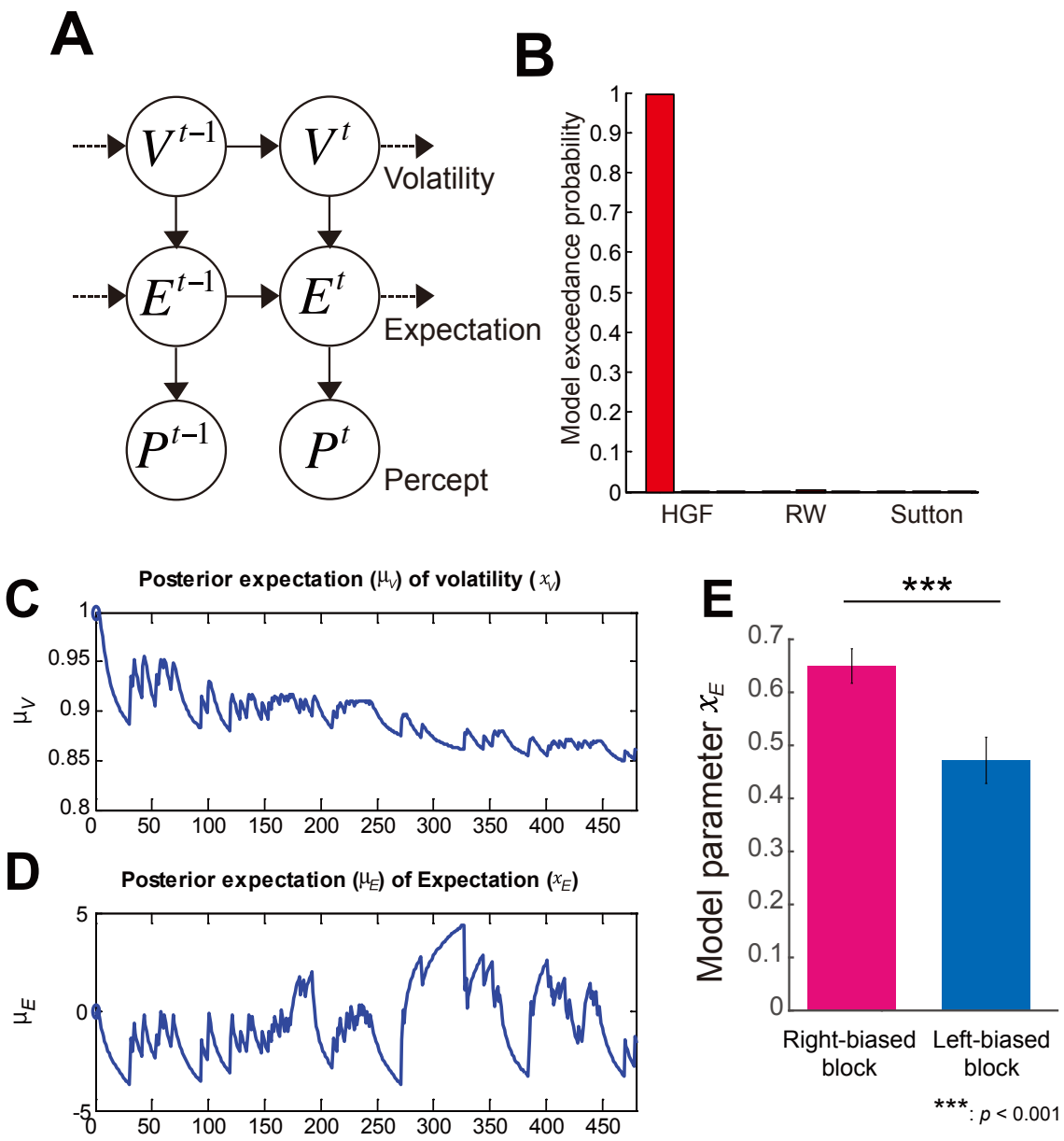

Figure 6-3 Computational model analysis.

(A) Participants' responses were analyzed using hierarchical Bayesian model (Mathys *et al.*, 2011). Participants' expectation and volatility trajectories were estimated on a trial-by-trial basis . (B) Bayesian model selection (BMS) was performed to determine which computational model best explains participants' choice behavior. six models were tested (two perceptual models: HGF model and Rescorla-Wagner model, three observation model: Gaussian observer, sigmoid function, and stick-function). BMS revealed that HGF model with Gaussian observer model best explains the behavioural data (exceedance probability: 0.82). (C, D) The plots show trajectory of volatility (C) and expectation (D) of a representative participant. The blue line indicates participant's parameter. (E) Expectation (model parameter $x2$) were higher when participants responded to right comparing to left ($t(19) = 4.13$, $p = 0.0006$).

6.4.3 MEG analysis

Finally, using the trajectory of parameters from the winning Bayesian model, I analyzed MEG data to identify brain regions where signals correlated with such a parameter and its temporal evolution on a per-participant basis. Any such regions would thus be identified as candidates for representing the expectation of incoming sensory input. I took a model-based analysis approach, which seeks to identify brain regions whose activity correlates with model parameters. Previous studies suggested that fronto-parietal regions play an important role in resolving conscious percept, especially generating hypotheses about perception and updating them on the basis of expectations (Kanai *et al.*, 2011; Megumi *et al.*, 2015a). In particular, the expression of parameters and their update should be represented in different brain (sub)regions, as theoretical work has suggested (Spratling, 2008a). Therefore I identified sources of activity in the MEG signals and looked for regions where activity represented model parameters. Given the current hypothesis and the model comparison results, I included four key parameters: expectation, expectation-PE, volatility, and volatility-PE for the regression analysis.

Figure 6-4 and Table 6-1 show the results of the MEG analysis. Signals in bilateral inferior frontal cortex were correlated with behaviourally derived parameters encoding expectation, which represents how likely participants were to report right-rotation on each trial. Expectation-PE, the trial-by-trial update of the expectation, was also correlated with signals from left postcentral gyrus and right superior temporal gyrus. Volatility (how quickly an observer switches their percept between the two), was correlated with activity in the bilateral precuneus (right precuneus, 10, -81, 39; left precuneus, -19, -80, 41), which is part of the superior parietal lobule (SPL). Finally volatility-PE, the update of the volatility, was encoded in left supramarginal gyrus, bilateral precuneus, bilateral postcentral gyrus, right superior temporal gyrus, and right inferior parietal lobule.

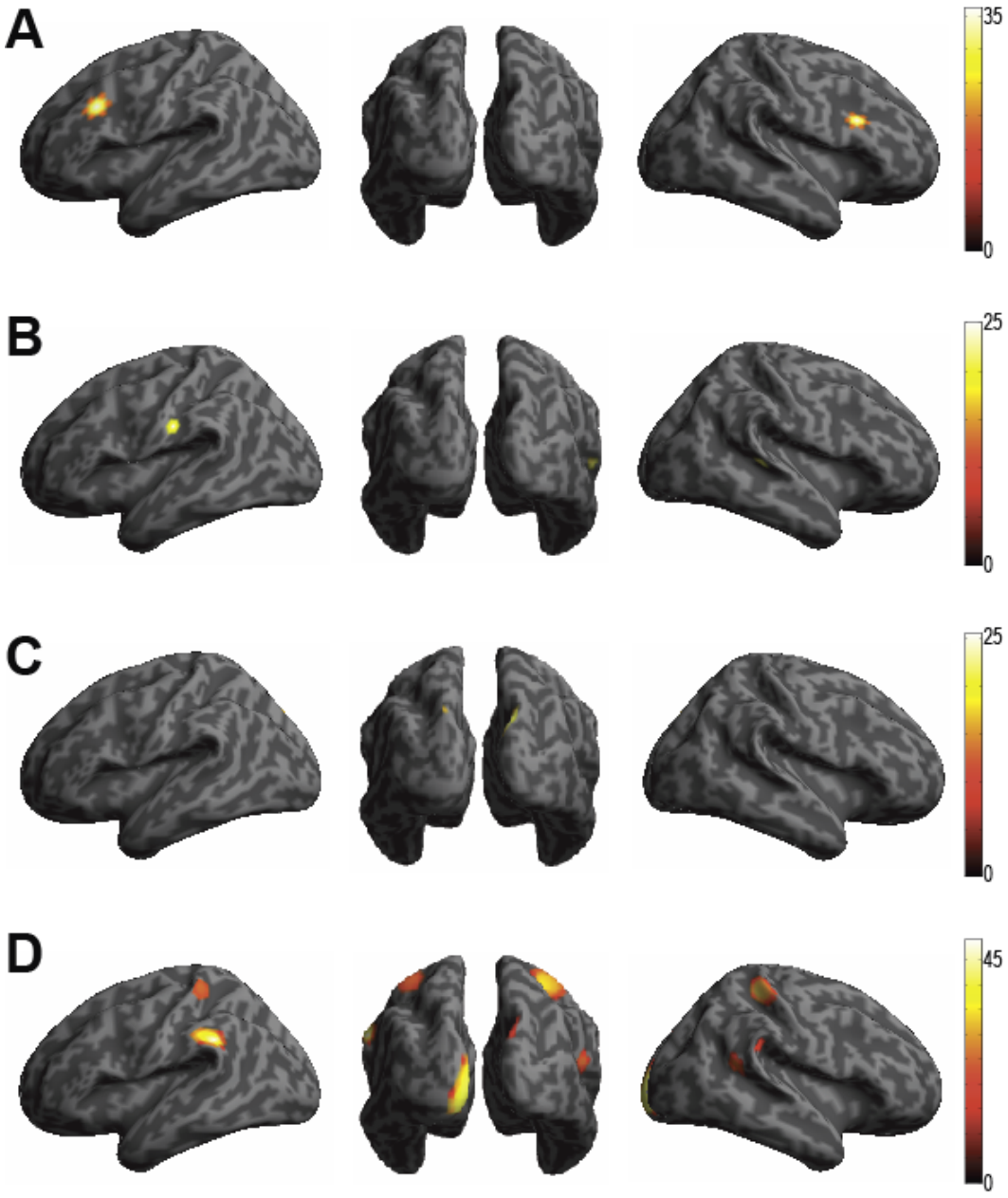

Figure 6-4 MEG analysis.

The figure represents brain areas whose activity correlated with HGF model parameters, Expectation (A), Expectation-PE (B), which is difference in expectation between trial t and t-1, Volatility (C), Volatility-PE (D). F contrast map was thresholded at $p < 0.05$, FEW-corrected at cluster level. The source regions are plotted on the inflated standard MNI template using the SPM visualization tool. The colormap indicates F-values.

Table 6-1 Labels and coordinate of MEG sources

Expectation

Cluster *p* (FWE-corr)	*k*	Peak *p* (FWE-corr)	*Z*	Coordinate *x*	*y*	*z*	Label TD label	IBA SPM116
0.03	13	0.001	5.3	-46	20	29	Middle Frontal Gyrus	Frontal_Inf_Tri_L
0.03	13	0.001	5.3	45	16	27	Middle Frontal Gyrus	Frontal_Inf_Oper_R

Expectation-PE

Cluster *p* (FWE-corr)	*k*	Peak *p* (FWE-corr)	*Z*	Coordinate *x*	*y*	*z*	Label TD label	IBA SPM116
0.033	7	0.002	4.49	-55	-20	27	Postcentral Gyrus	Postcentral_L
0.042	3	0.021	3.95	58	-34	6	Temporal_Sup_R	superior temporal gyrus right

Volatility

Cluster *p* (FWE-corr)	*k*	Peak *p* (FWE-corr)	*Z*	Coordinate *x*	*y*	*z*	Label TD label	IBA SPM116
0.031	11	0.002	4.53	10	-81	39	Precuneus	Cuneus_R
0.048	1	0.037	3.71	-19	-80	41	Precuneus	Occipital_Sup_L

Volatility-PE

Cluster *p* (FWE-corr)	*k*	Peak *p* (FWE-corr)	*Z*	Coordinate *x*	*y*	*z*	Label TD label	IBA SPM116
0.01	28	28	6.02	-59	-45	32	Supramarginal Gyrus	SupraMarginal_L
0.002	57	57	5.73	-8	-91	6	Cuneus	Calcarine_L
0.007	34	34	5.2	28	-36	54	Postcentral Gyrus	Postcentral_R
0.023	14	14	4.66	53	-49	18	Superior Temporal Gyrus	Temporal_Sup_R
0.017	19	19	4.54	-38	-40	60	Postcentral Gyrus	Postcentral_L
0.034	7	7	4.09	52	-36	31	Inferior Parietal Lobule	SupraMarginal_R
0.034	7	7	3.94	13	-84	41	Precuneus	Cuneus_R

6.4.4 Validation analysis

6.4.4.1 Removed trials and model parameters

It is known that eye movements can affect or modify interpretation of bistable stimuli, and one may argue the current results are reflection of MEG artefacts induced by percept-specific eye movements. I therefore tested whether there were any differences in the number of rejected trials across

stimulus conditions and in estimated model parameters. For the comparison of the model parameter, a non-parametric test (Wilcoxon signed rank test) was used to avoid making a strong assumption regarding the distribution of the parameters. There was no significant difference in the ratio of rejected trials across stimulus conditions. I confirmed that there was no significant difference in the ratio of rejected trials across stimulus condition (six prime trial conditions and one probe trial conditions, one-way ANOVA, $F(6, 114) = 0.75$, $p = 0.61$) and no difference in estimated HGF parameters between rejected trials and other trials (Wilcoxon signed rank test; $Z = -1.50$, $p = 0.14$ for expectation, $Z = -0.04$, $p = 0.97$ for expectation-PE, $Z = 0.63$, $p = 0.54$ for volatility, $Z = -0.37$, $p = 0.71$ for volatility-PE, $Z = -1.76$, $p = 0.79$ for surprise; uncorrected p-value).

6.5 Discussion

6.5.1 Overview of the findings

Predictive coding theory proposes that visual percepts are affected by expectation for incoming sensory stimuli, but until now it has remained elusive how such expectation affects interpretation of ambiguous stimuli and is represented in the brain. To address this question, I invented a new intermittent-presentation bistable perception paradigm associated with statistical learning. I found that the history of the percept biased subsequent subjective reports of the percept, and that hierarchical Bayesian modelling with volatility and expectation layers could successfully model participants' perceptual priors for incoming sensory stimuli. Furthermore, MEG data analysis revealed several brain regions, including inferior frontal gyrus, parietal areas, supramarginal gyrus, and cuneus where signals exhibited significant correlations with model parameters.

6.5.2 Top-down signal and bistable perception

Determinants of bistable perception dynamics have been investigated using

different bistable perception tasks (Mamassian & Goutcher, 2005; Brascamp *et al.*, 2008; Denison *et al.*, 2011; Chopin & Mamassian, 2012). These studies show that perceptual switches exhibit complex dynamics and computational models have been proposed to explain the dynamics of bistable perception (Wilson, 2003; Schrater & Sundareswara, 2006; Noest *et al.*, 2007; Shpiro *et al.*, 2009; Moreno-Bote *et al.*, 2011). Above all, predictive coding theory (Dayan, 1998; Hohwy *et al.*, 2008) has attracted much attention for its simplicity as well as biological plausibility. This theory assumes the brain computes an expectation regarding the possible environmental causes of percept, and such expectation is modified by visual inputs and perceptual experience (Clark, 2013). fMRI studies provide empirical evidence that sensory and fronto-parietal cortices show inter-regional coupling during bistable perception (Weilnhammer *et al.*, 2013; Megumi *et al.*, 2015a), consistent with the notion that the brain modifies hypotheses about the environment through connectivity. Furthermore, expectation or experience obtained through learning affects perception of ambiguous stimuli (Sterzer *et al.*, 2008; Murphy *et al.*, 2014). These studies all therefore suggest that expectation plays an essential role in determining interpretation of ambiguous stimuli. However, perceptual switches in ambiguous bistable perception happen spontaneously and it is hard to control them experimentally. This makes it difficult for us to observe or quantify expectation in a standard bistable perception task. Here, by using disambiguated bistable perception stimuli in a novel statistical learning paradigm, participants' expectation was experimentally manipulated. This enables to describe the internal states of participants with a hierarchical Bayesian model and use this to interrogate MEG signals to identify brain regions associated with particular aspects of expectation formation. The current method can potentially be applied other visual paradigms to interrogate how our brain extracts statistical information and forms conscious percept.

We found that the expectation built up through the specific stimulus sequence helped to stabilize the current percept. Intermittent presentation paradigms typically induce slower switch rates than

continuous presentation (Leopold *et al.*, 2002) and therefore I assumed that participants in an intermittent paradigm such as ours would tend to maintain the same percept. On the other hand, computational modelling of bistable perception often speculates that fatigue in sensory neurons caused by sustained percept is the driving force of perceptual switches (Dayan, 1998). Indeed, in bistable perception task, enhancement and suppression of the current percept can happen at the same time (see Pearson and Brascamp (2008)). The two effects yield complex dynamics of bistable perception, making it difficult to interrogate the effect of prior expectation on interpretation of bistable stimuli. In this experiment, I speculated that the speed of changes in expectation should be relatively slow compared to fatigue in sensory neurons and therefore the changes in expectation were concealed due to 'fatigue' effect. The stimuli were presented for a very short period so that fatigue effect in sensory area was thought to small and cancelled out, and this is the reason why dominant percept persisted in the probe trials. A future challenge would be to incorporate variables representing sensory neurons with Bayesian model so that it should explain perceptual switches in continuous presentation paradigms as well (Dayan, 1998).

6.5.3 Role of fronto-parietal regions when building up expectation

Our results suggested that fronto-parietal cortices (inferior frontal area, postcentral gyrus) are involved in the computation of Expectation and Volatility-PE, together with precuneus, superior temporal gyrus, and inferior parietal lobe. While signals from inferior frontal cortex were associated with changes in expectation, which affect the interpretation of the next percept, parietal cortex signals showed significant correlations with the updating of volatility. Volatility represents how often participants experienced perceptual alternation (i.e. how quickly they changed their percept), and this would be conceptually related to the individual tendency to maintain or change current percept (Kanai *et al.*, 2010; Kanai *et al.*, 2011; Watanabe *et al.*, 2014; Megumi *et al.*, 2015a). The current findings may thus

also be consistent with recent suggestions that activation in fronto-parietal cortices mediate top-down information processing rather than being the indispensable cause of perceptual switches (Knapen *et al.*, 2011; Frässle *et al.*, 2014; Brascamp *et al.*, 2015). Signals associated with Expectation-PE were found in superior temporal gyrus. This region is involved in generating prediction of object movements (Schultz *et al.*, 2004) and processing PE signal in social behavior (Zilbovicius *et al.*, 2006); current results added supporting evidence to their arguments.

6.5.4 Relation to other visual tasks

What is special about the current results compared to findings from other visual tasks that require ambiguity to be resolved? SFM stimuli have a prominent feature of inherent ambiguity in the signal. Previous studies of the effect of prior percepts often use visual stimuli that have a clear distinction between 'signal' and 'noise'; in such a situation 'ambiguity' is caused by the noise in the stimulus. In binocular rivalry task, which is also another form of bistable perception, the non-dominant stimulus can be seen as 'noise' and the brain needs to suppress such irrelevant information to form stable percept (Lee & Blake, 2002) so rivalrous stimuli compete with each other to enter consciousness. On the other hand, in our SFM task, the sensory input (motion of the dots) itself is ambiguous and all sensory inputs (i.e. moving dots consisting the sphere) become supporting evidence for either the right or left rotation percept; in this sense there is thus no 'noise' in the visual inputs. This aspect of SFM stimuli makes them particularly suitable to observe 'expectation' effects without strong suppression signal.

6.6 Conclusion

I asked whether people can utilize prior expectations when resolving ambiguity in a visual bistable perception task. Behavioural data suggested that the subjective percept was biased by the statistical properties of the stimulus sequence, and such priors, described with a hierarchical Bayesian

model, were represented in fronto-parietal cortices, superior temporal gyrus, and precuneus. The results suggest that the dynamics of bistable perception task are determined by expectation and its volatility thus representing empirical evidence for predictive coding in visual perception.

Chapter 7 GENERAL DISCUSSION

7.1 Overview of the findings

This thesis aimed to elucidate the determinants of the dynamics of perception for ambiguous visual information (using a common bistable perception task) with an emphasis on (1) the role of multi-regional interactions and brain networks and (2) the role of prediction. A series of experiments have revealed that specific brain regions and functional brain networks support the emergence of visual awareness during bistable perception.

In Chapter 3, I tested whether the functional localizer method used in previous studies can be applied to detect bilateral LGN at individual level. It turned out that this method did not work very well, and it was suggested that it is needed to improve the current methods (e.g. using anatomical masks, improve scanning parameters).

Chapter 4 demonstrates that interactions among the two fronto-parietal areas and the visual cortex support individual difference in bistable perception performance. Previous VBM and TMS studies have suggested that two parietal regions, r-aSPL and r-pSPL differentially involve in spontaneous perceptual switches in bistable perception tasks (Carmel *et al.*, 2010b; Kanai *et al.*, 2010; Kanai *et al.*, 2011). Based on thse findings, I hypothesized that these regions build hierarchical connections together with the sensor cortex. To test this hypothesis, I have conducted an fMRI experiment. Participants performed SFM bistable perception task in the fMRI scanner and reported their percept in spontaneous perceptual switches and stimulus-driven perceptual switches. I found the two parietal regions, r-aSPL and r-pSPL, showed grater activity in spontaneous perceptual switches comparing to stimulus-driven switches. Then, I tested the four hierarchical connectivity models (no modulation, top-down modulation, bottom-up modulation, and bidirectional modulation) with DCM approach, and found both top-down and bottom-up connectivities were

modulated when perceptual switches occurred. Furthermore, the strength of bottom-up modulation explained individual difference in mean dominance duration. These results supported the idea that fronto-parietal regions communicate with sensory cortex, and such interaction affects dynamics of bistable perception.

In the following Chapter 5 investigated influence of intrinsic network architecture for bistable perception performance. The series of analysis in Chapter 4 have shown that bistable perception induces complex brain dynamics and multi-regional interaction, yet it remained unclear whether some aspects of such intrinsic network organization are present even without external input (i.e. in the resting state) and are predictive for bistable perception performance. To address this question, I applied graph theoretical analysis to resting state fMRI data recorded from human participants and sought to determine which brain regions and functional networks could predict individual switch frequency in bistable perception. Regions of interest were selected based on a previous meta-analysis, and a binary undirected graph was created using resting-state functional connectivity between these regions, and graph centrality measures were obtained. A support vector regression analysis with centrality measures revealed that the PageRank centrality of the right postcentral gyrus, the left insula, and the left lateral occipitotemporal gyrus all significantly contributed to predict individual differences in mean dominance duration (or its inverse, switch frequency). Furthermore, such prediction performance was significantly deteriorated when the connections of the fronto-parietal and visual sub-network, which is thought to play a significant role in resolving ambiguity in visual information, were removed from the graph. These findings indicate that systematic features of the resting state connectivity can predict task-related behavioral dynamics, suggesting that intrinsic network properties of human brain underlie individual differences in the dynamics of visual awareness.

In the final experimental chapter, Chapter 6 showed that expectation influences dynamics of bistable perception and revealed brain regions involving in mediating such expectation. Our brain forms conscious

percepts by complementing ambiguous sensory information. Predictive coding theory proposes that the brain derives statistical information and utilizes such information as prior knowledge when resolving ambiguity. Here, using bistable structure-from-motion (SFM) stimuli, I explored how such prior affects subjective percept as well as its neural representation. I measured brain activity with magnetoencephalography (MEG) while participants reported the perceived direction of an intermittently presented bistable stimulus. To manipulate participants' prior expectation, disambiguated SFM were presented. Results proved that subjective percept of ambiguous stimuli were biased towards the primed direction. I estimated each participant's expectation trajectory with hierarchical Bayesian model and found the model parameters were represented in fronto- parietal areas, precuneus, and visual cortex. Together, these results suggest that prior expectations were encoded and updated in the brain and shape the contents of conscious experience.

7.2 Bottom-up and top-down modulation for perceptual switches

7.2.1 Predictive coding account of top-down and bottom-up information processing

As discussed in chapter 1, the roles of top-down and bottom-up signals have been discussed in the psychology literature, and the development of neuroimaging methods enables us to associate top-down and bottom-up information processing to brain activity.

The experiments in chapter 4 (DCM analysis with task fMRI data) and chapter 6 (source analysis with MEG data) aimed to elucidate the driving force (i.e. cause) of perceptual switches in bistable perception task. Results in chapter 4 suggested that both top-down and bottom-up connections are modulated in perceptual switches, but only bottom-up connectivity contributes to explain individual tendency (i.e. mean dominance duration) in bistable perception view. The experiment in chapter 6 provided

empirical evidence that the expectation, which is more likely to be 'top-down' modulation (in psychological definition), helps to resolve ambiguity in visual bistable perception.

These results from the two experiments indeed support the idea that the combination of top-down and bottom-up modulation is essential to cause perceptual switches. Predictive coding theories suggest that the interpretation of the world is made up by the interaction of top-down and bottom-up signal ((Clark, 2013)). DCM analysis in chapter 4 showed that the bidirectional models surpass top-down or bottom-up models, suggesting that the reciprocal interactions between parietal regions (r-aSPL and r-pSPL) and the visual cortex (r-V5) contribute to perceptual switches. The experiment in chapter 6 gave more direct evidence regarding the role and computation of trial-by-trial expectation computed for the subsequent perception, and such expectation is represented in the brain, suggesting the brain does utilize the expectation when resolving sensory ambiguity.

The results in chapter 4 and 6 appear to contradict each other; bottom-up connectivity is more important to explain individual differences, while top-down signals help to resolve ambiguity. To resolve this apparent contradiction, we need to consider two things – the definition of top-down and bottom-up, and the difference in experimental paradigm.

7.2.1.1 Psychological definition and neuronal definition for 'top' and 'bottom'

The experiment in chapter 4 aimed to elucidate the dynamics of brain activity, and the definition of 'top-down' and 'bottom-up' was purely based on the brain hierarchy. On the other hand, the experiment in chapter 6 manipulated expectations by changing the sequence of stimulus presentation. Here, the definition of 'top-down' is based on psychological implications – for example, (Summerfield & de Lange, 2014) uses 'top-down' to indicate observer's knowledge or expectations. The MEG data analysis in chapter 6 focused on finding regions representing the model parameters, and did not consider an explicit hierarchical structure in the brain, or modulations on top-down and bottom-up connectivity.

Of course it is possible to dissociate bottom-up and top-down information processing in psychological definition with clever experimental manipulations (e.g. (Pinto *et al.*, 2013)). However, it may be more difficult to dissociate top-down modulation and bottom-up modulation in the brain. Given the hypothesis that the expectation is built up and updated through the interaction of previous percept and incoming sensory stimuli, the influence of top-down signals (in a psychology definition) may be represented in multiple regions, including fronto-parietal areas and sensory areas (c.f. the results in chapter 6). Indeed, previous studies found the influence of expectation (top-down modulation) in the sensory cortex as well as the fronto-parietal areas, suggesting that top-down and bottom-up influences (in psychological definition) may not be completely separable in the brain (Rauss & Pourtois, 2013).

We are always tempted to expect that top-down processing, such as attention or memory, is associated with signals in 'higher' brain areas or descending connectivity, but the current human neuroimaging studies may not be able to address this issue. Further research is needed to understand how trial-by-trial expectation, which should in essence be treated as 'top-down' influence, is represented in the brain hierarchy and how the connectivity between regions is modulated as a function of expectation, possibly using brain stimulation or animal models.

7.2.1.2 Continuous presentation and intermittent presentation

The mechanism and driving force of perceptual switches might be different in the two different paradigms. The experiment in chapter 4 employed a continuous presentation paradigm, whereas the one in chapter 6 used an intermittent presentation paradigm.

In the continuous presentation paradigm (experiment in chapter 4), the stimuli (sphere-shape SFM) were presented for a few ten seconds and participants were asked to report their percept by holding buttons. While participants perceive one percept, populations of neurons in visual cortex represent for perceived stimuli and therefore it may be possible that

perceptual switches have occurred due to fatigue in sensory neurons. This has been implied by theoretical work; (Dayan, 1998) simulated the neural dynamics of a hierarchical brain model under binocular rivalry task, and it was suggested that oscillators representing fatigue of V1 sensory neurons cause perceptual switches in rivalry. Indeed, longer exposure to the disambiguated SFM stimuli elicits suppressive bias of the perceived interpretation by sensory adaptation (Nawrot & Blake, 1989). This can explain why the results in chapter 3 showed that individual differences in mean dominance duration could be explained by strength of modulation on bottom-up connectivity, but not top-down connectivity.

On the other hand, in the intermittent presentation paradigm, the stimuli are presented for a very brief period (less than a few seconds) and it might be too short to cause the fatigue effect in sensory neurons. In intermittent-presentation paradigm, perceptual switches occur much less often (Leopold *et al.*, 2002), suggesting the sensory memory (also individual tendency or preference (Carter & Cavanagh, 2007)) strongly affects the dynamics of conscious percept. This explains why the expectation biased the interpretation of ambiguous stimuli.

7.3 Functional roles of parietal regions for visual bistable perception – causal signal or epiphenomenal activation?

In the previous section, I have discussed top-down and bottom-up modulation, and its influence on two different paradigms for bistable perception. How should we integrate the implications from the two different paradigms and associate them with particular brain regions or networks, especially fronto-parietal regions?

Previous studies suggested that fronto-parietal regions are involved in bistable perception (Leopold & Logothetis, 1999; Sterzer *et al.*, 2009), and such activation in fronto-parietal regions during bistable perception tasks has been replicated several times. The experiment in chapter 4 also yielded

activation in the frontal areas and the parietal areas, together with the visual cortex. Resting-state analysis in chapter 5 supported that the fronto-parietal network contributes to the prediction of individual performance in a bistable perception task. These results both support the idea that fronto-parietal regions form a macroscopic network and that the organization of such network explains individual tendency in bistable perception performance.

Traditionally, bistability is thought to be a reflection of competition in early stages of cortical processing. fMRI enables us to record large-scale brain activity from humans, rather than focusing on vision-specific areas. Accumulation of empirical evidence suggested that fronto-parietal areas are involved in perceptual switches (Leopold & Logothetis, 1999; Rees, 2007). The idea that fronto-parietal regions are involved in conscious perception is also supported by work in other domains (e.g. conscious- and unconscious-word processing (Gaillard *et al.*, 2009)), and (Dehaene & Changeux, 2011) discussed the possibility that the 'ignition' of the prefrontal and parietal networks play a causal role in accessing the contents of (visual) consciousness and make a report of current perception.

However, recent findings cast doubt on the hypothesis that the fronto-parietal areas are actively involved in perceptual switches. Activation in fronto-parietal cortices is induced by conscious reports (not changes in perception) (Frässle *et al.*, 2014) and ambiguity in stimuli (Knapen *et al.*, 2011; Brascamp *et al.*, 2015). These authors both proposed that fronto-parietal activation is not the driving force of perceptual switches, but reflections of the percept or induced by other cognitive processes (e.g. attention).

How do these earlier findings relate to the experiments presented in this thesis, and what are the implications for understanding the roles of fronto-parietal cortices in bistable perception? I would like to propose that fronto-parietal regions are involved in two different factors controlling the dynamics of visual bistable perception; expectation for incoming sensory stimuli and individual-specific tendency (i.e. switch frequency or volatility).

As explained at the beginning of this section, the experiment in

chapter 5 demonstrated that the architecture of the fronto-parietal network during the task-free resting-state contributes to successful prediction of individual switch frequency, and this suggests that an intrinsic property of the fronto-parietal network may explain the 'baseline' or individual tendency of bistable perception performance. It is known that resting-state dynamics predict task performance for other visual tasks (e.g. resting-state activity and perceptual learning performance (Baldassarre *et al.*, 2012)), suggesting that the intrinsic properties of brain networks mediate performance in a complex task. As reviewed in chapter 1, the dynamics of bistable perception are influenced by inherent factors (e.g. genetic factors, (Miller *et al.*, 2000)), and such factors may be related to individual difference in brain volume in the front-parietal areas. The involvement of fronto-parietal regions in the inherent dynamics of bistable perception performance, rather than directly controlling the trial-by-trial percept, is further supported by previous VBM studies that showed a systematic relationship between the volume of SPLs and individual mean dominance duration (Kanai *et al.*, 2010; Kanai *et al.*, 2011).

Such individual tendency was also found in the Bayesian model used in chapter 6. The model consisted of three layers, representing percept, expectation, and volatility. Volatility quantifies how well the observer sticks to one percept; high-volatility means the observer tends to switch percept frequently (i.e. higher switch frequency), whereas low-volatility indicates the observer tends to keep the current percept (i.e. lower switch frequency). My data showed that these parameters were correlated with activity in the parietal cortex as well as the visual cortex, indicating that fronto-parietal regions play roles in controlling the dynamics of bistable perception (together with the visual cortex).

To better understand the functional roles of fronto-parietal regions, further studies are needed to elucidate the precise relationship between the structures of fronto-parietal regions and volatility; especially how the anatomical structure produces the cortical dynamics and result in individual difference.

7.4 Future research questions

7.4.1 How to define the network?

The experiments in chapter 4 and 5 tested the organization of the brain network architecture. Both experiments utilized findings from previous functional connectivity studies ((Mars *et al.*, 2011; Power *et al.*, 2011)). Although DCM analysis (chapter 4) is an effective connectivity analysis and the combination of different connectivity analysis approaches strengthens the importance of the fronto-parietal network, we still need to consider the difference between functional/effective connectivity and anatomical connectivity.

Functional connectivity is indirect evidence for the existence of connections between regions. Functional connectivity is computed from the time-series of regional brain activity (e.g. BOLD signal averaged within the regions), and anatomical connectivity is found by fiber-tracking methods (e.g. DTI, injecting tracers). It has been suggested that the pattern of functional connectivity can be accounted for by anatomical connectivity (Vincent *et al.*, 2007), but there is still some discrepancy between the maps. Notably, (Honey *et al.*, 2009) compared resting-state functional connectivity and anatomical connectivity (measured with DTI) and found that the resting-state functional connectivity pattern produces robust connections even if the regions are not directly linked.

It is important to consider what we are looking at by investigating functional connectivity, and it would be beneficial for researchers to constrain the connectivity pattern using DTI data first. DCM analysis is hypothesis-based, and choosing the models for the analysis is very important. DCM analysis assumes equal priors across the models, and hence it may be harmful to include 'less-likely' models in the model space. It is strongly recommended to utilize the anatomical-knowledge when choosing the models (Stephan *et al.*, 2010).

The recent development of connectivity database and shared data might be a remedy for this problem. For example, human connectome

project (http://www.humanconnectomeproject.org/) offers human DTI data, which helps researchers to explore the anatomical connectivity between regions.

7.4.2 How to consider weak connections?

The resting-state network analysis in chapter 5 postulates that the brain network can be described as undirected binary graph under resting condition. The comparison of the human brain network revealed that weighted and unweighted (i.e. binary) graphs show similar topological structures; almost all nodes are directly linked to hub nodes (called rich-club nodes) and the rich-club nodes are strongly interconnected (van den Heuvel & Sporns, 2011) (Figure 7-1). Furthermore, attacks to the rich-club nodes (i.e. removing the nodes/edges from the graph) gives more damages to the efficiency of the graph than attacks to the non-rich-club nodes. These findings bring an idea that the few regions in the brain are more 'important' and own dense connections. Rich-club organization can be found in the newborn brain (Ball *et al.*, 2014), and rich-club regions are more likely to show anatomical abnormalities in many brain disorders (Crossley *et al.*, 2014).

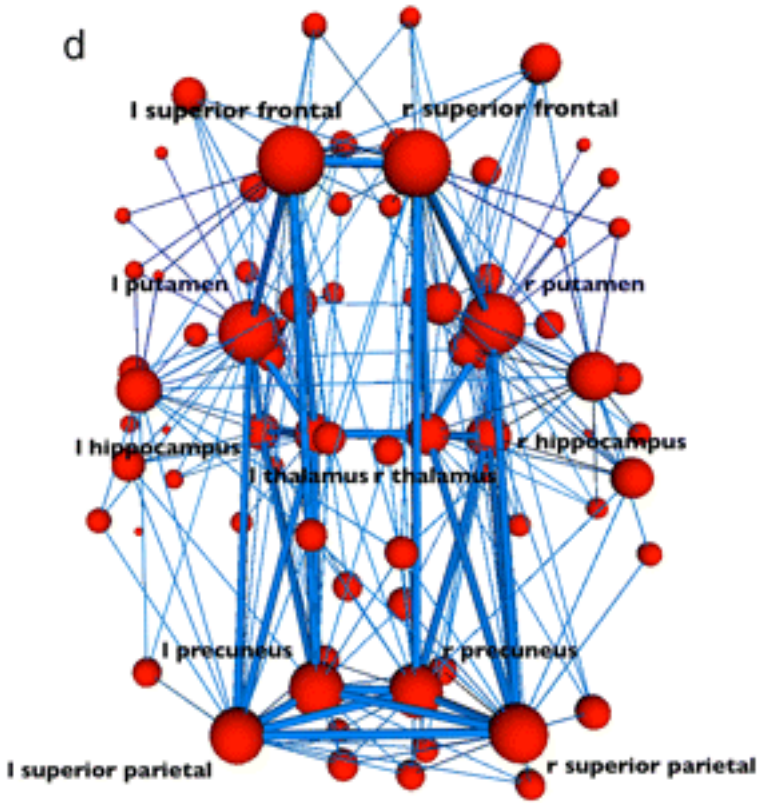

Figure 7-1Rich-club nodes and connections in the human brain.

The image is created from the diffusion tensor imaging data and averaged across participants. The size of the nodes (red cspheres) indicates their degree. The thick edges (blue lines) are rich-club connections, and the thin blue lines are fiber pathway in >75% participants. Adopted from (van den Heuvel & Sporns, 2011).

Although such a threshold approach has been successful in describing human brain dynamics, the ignored 'weak connections' may also contribute to the connectivity profile and emergence of behavior and perception. As explained, connections between nodes are often evaluated using a binarized matrix. When binarizing the connectivity matrix and computing graph theoretical measures, a certain threshold should be set to define existence of edge between nodes. This approach often produces sparse connectivity in the brain and thus many studies support the idea of "small-worldness" in brain connectivity.

(Markov *et al.*, 2014) investigated the network structure of the macaque brain; the weighted and directed connectivity map of the macaque brain was created with retrograde tracer injections, and the analysis revealed that the majority of the inter-regional connections are moderate or

weak in strength. Although the structures of the human brain and the primate brain are different and thus we cannot over-generalize the findings, the discrepancy between the results from human non-invasive studies and (Markov *et al.*, 2014) may be due to thresholding of the connectivity matrix; in resting-state analysis, weak connections are treated as 'no-connection'. Future studies may address the role of the weak connections in the network topology and their contributions for emergence of individual difference in the brain.

7.4.3 Preprocessing steps

When analyzing resting state data and quantifying functional connectivity accurately, it is indispensable to remove nuisance signal and obtain 'true' correlations in BOLD signals reflecting relevant connections. Studies have shown that inappropriate preprocessing procedures result in failure of detecting connectivity and characterize the network organization.

(Power *et al.*, 2012) raised a caution that conventional preprocessing approach – applying band-pass filter to ROI BOLD signal and performing regression analysis using head-motion parameters and averaged signal from gray matter/white matter/cerebrospinal fluid compartments – is not sufficient to eliminate the effect of participant movement during fMRI measurement. Power et al. (2012) reported that spontaneous head movements caused contamination in 'true' signal and spurious correlation in the resting state signal. One possible solution is motion-scrubbing; motion-scrubbing is a preprocessing step that removes EPI volumes where the participant showed significant movements, especially for participants who are likely to produce bigger head motion e.g. young adolescents, elder cohorts, and patients. Of course, there is no straightforward way to determine the 'best' preprocessing steps for functional connectivity analysis as it is not easy to estimate 'true' functional connectivity. It might be interesting to see which preprocessing method for functional connectivity helps the best to identify anatomical connectivity – this prediction accuracy may be an index to determine the goodness of preprocessing method.

7.4.4 The role of specific frequency

The experiment described in chapter 6 focused on exploring neural representation of Bayesian parameters, and here I assumed that such parameters (expectations for incoming sensory stimuli) would be represented in the correlated activity of different brain regions. This assumption is motivated by the predictive coding model ((Rao & Ballard, 1999; Spratling, 2008a)), which consists of the subsets of neural populations representing prediction and prediction-error and send such signals to different areas in the brain hierarchy. Such scheme helps to achieve efficient coding of information. In addition, VBM studies implied that sub-regions in parietal cortex mediate different information (prediction and prediction-error) (Kanai *et al.*, 2010; Kanai *et al.*, 2011), and these findings further support the feasibility of ROI-based representation of the parameters.

On the other hand, specific frequency signals may mediate inter-regional communications (Bastos *et al.*, 2012). It is known that neurons in primate superficial layers and granular layers produces different oscillatory activity patterns; (Buffalo *et al.*, 2011) found that neurons in the superficial layers produces gamma-synchronous activity and are projected to descending areas (i.e. top-down modulation), whereas neurons in the gradual layers show alpha-band synchrony and are projected to ascending areas (i.e. bottom-up modulation). These yield the hypothesis that top-down and bottom-up modulations are carried by different frequency signals.

Although it is not entirely clear how the laminar-level neural activity might be related to macroscopic neural activity such as that measured using fMRI, and how it might relate to perception in bistable perception task, there are some empirical studies supporting the hypothesis that top-down and bottom-up modulations are related to specific frequency of the M/EEG signals. Apha-band signal (8Hz – 16Hz) decrease accompany with perceptual switches (e.g. (Isoglu-Alkac *et al.*, 2000)). (Kloosterman *et*

al., 2014) demonstrated that the power in beta-band frequencies (12-30Hz) can predict perceptual switches in MIB bistable perception task, and concluded that beta-band oscillation corresponds to top-down modulation preceding perception. Non-invasive stimulation with transcranial alternating current stimulation (tACS) revealed that rhythmic gamma stimulation (60Hz) over posterior area of the brain increases the spontaneous switch rate of SFM bistable perception stimuli (Cabral-Calderin *et al.*, 2015). Given the finding that tACS stimulation induces frequency-specific increase of both power and coherence (Helfrich *et al.*, 2014), their data suggests that the involvement of gamma-band in bistable perception [12].

There is no consensus which frequency band represents top-down or bottom-up modulation, but future experiments might address this question by demonstrating more precise relationship between specific frequency band and top-down or bottom-up modulation.

7.5 Conclusion

The series of experiments in this thesis interrogated how the brain resolves ambiguity in visual information, especially how sets of brain regions interact with each other and how such networks helps to stabilize or switch subjective perception. Analysis of fMRI brain activity during structure-from-motion bistable perception task revealed that sub-regions in right parietal cortex and the visual cortex (V5) form a bidirectional network, and the strength of bottom-up connectivity in this network explains inter-individual variance in mean dominance duration, suggesting that the ascending signal plays roles in perceptual switches. Analysis of intrinsic, task-free resting-state activity analysis elucidated that the topological structure of the brain network is predictive for individual mean dominance

[12] There is experimental evidence that tACS stimulation may not be frequency-specific (e.g. Brignani, D., Ruzzoli, M., Mauri, P. & Miniussi, C. (2013) Is transcranial alternating current stimulation effective in modulating brain oscillations? *PloS one*, **8**, e56589.).

duration, and fronto-parietal network, together with the visual network, contributes to successful prediction. Finally, I have established a Bayesian model to explain dynamics of bistable perception, and found that the perceptual priors explain trial-by-trial percept. Such parameters are represented as MEG signals in parietal cortex, frontal cortex, and the visual cortex.

Together, these findings provided further evidence that multiple brain regions, especially fronto-parietal areas, comprise distributed brain networks where communications between regions and networks play important roles in perceptual switches during bistable perception. I propose that the temporal dynamics of visual bistable perception is controlled by expectation about the world and individual tendency (i.e. switch frequency or volatility). The two factors should be represented from the bidirectional interactions of multiple cortical regions, including fronto-parietal regions. This view nicely matches with predictive coding theory.

REFERENCE

Achard, S., Salvador, R., Whitcher, B., Suckling, J. & Bullmore, E. (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci*, **26**, 63-72.

Akaishi, R., Umeda, K., Nagase, A. & Sakai, K. (2014) Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron*, **81**, 195-206.

Anderson, E.J., Dakin, S.C. & Rees, G. (2009) Monocular signals in human lateral geniculate nucleus reflect the Craik-Cornsweet-O'Brien effect. *Journal of vision*, **9**, 14 11-18.

Arend, I., Machado, L., Ward, R., McGrath, M., Ro, T. & Rafal, R.D. (2008) The role of the human pulvinar in visual attention and action: evidence from temporal-order judgment, saccade decision, and antisaccade tasks. *Progress in brain research*, **171**, 475-483.

Ashburner, J. & Friston, K.J. (2000) Voxel-based morphometry--the methods. *Neuroimage*, **11**, 805-821.

Baldassarre, A., Lewis, C.M., Committeri, G., Snyder, A.Z., Romani, G.L. & Corbetta, M. (2012) Individual variability in functional connectivity predicts performance of a perceptual task. *Proc Natl Acad Sci U S A*, **109**, 3516-3521.

Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., Robinson, E.C., Ogundipe, E., Rueckert, D., Edwards, A.D. & Counsell, S.J. (2014) Rich-club organization of the newborn human brain. *Proc Natl Acad Sci U S A*, **111**, 7456-7461.

Barnes, G.a.H., A. and Hirata, M. (2010) Magnetoencephalogram. *Scholarpedia*, **5**, 3172.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P. & Friston, K.J. (2012) Canonical microcircuits for predictive coding. *Neuron*, **76**, 695-711.

Beucke, J.C., Sepulcre, J., Talukdar, T., Linnman, C., Zschenderlein, K., Endrass, T., Kaufmann, C. & Kathmann, N. (2013) Abnormally high degree connectivity of the orbitofrontal cortex in obsessive-compulsive disorder. *JAMA psychiatry*, **70**, 619-629.

Biswal, B., Yetkin, F.Z., Haughton, V.M. & Hyde, J.S. (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, **34**, 537-541.

Blake, R.R., Fox, R. & McIntyre, C. (1971) Stochastic properties of stabilized-image binocular rivalry alternations. *Journal of experimental psychology*, **88**, 327-332.

Bohland, J.W., Wu, C., Barbas, H., Bokil, H., Bota, M., Breiter, H.C., Cline, H.T., Doyle, J.C., Freed, P.J., Greenspan, R.J., Haber, S.N., Hawrylycz, M., Herrera, D.G., Hilgetag, C.C., Huang, Z.J., Jones, A., Jones, E.G., Karten, H.J., Kleinfeld, D., Kotter, R., Lester, H.A., Lin, J.M., Mensh, B.D., Mikula, S., Panksepp, J., Price, J.L., Safdieh, J., Saper, C.B., Schiff, N.D., Schmahmann, J.D., Stillman, B.W., Svoboda, K., Swanson, L.W., Toga, A.W., Van Essen, D.C., Watson, J.D. & Mitra, P.P. (2009) A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS computational biology*, **5**, e1000334.

Bonneh, Y.S., Cooperman, A. & Sagi, D. (2001) Motion-induced blindness in normal observers. *Nature*, **411**, 798-801.

Boring, E.G. (1930) A new ambiguous figure. *The American journal of psychology*.

Brainard, D.H. (1997) The Psychophysics Toolbox. *Spatial vision*, **10**, 433-436.

Brascamp, J., Blake, R. & Knapen, T. (2015) Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nat Neurosci*.

Brascamp, J.W. & Blake, R. (2012) Inattention abolishes binocular rivalry: perceptual evidence. *Psychological science*, **23**, 1159-1167.

Brascamp, J.W., Knapen, T.H., Kanai, R., Noest, A.J., van Ee, R. & van den Berg, A.V. (2008) Multi-timescale perceptual history resolves visual ambiguity. *PloS one*, **3**, e1497.

Brascamp, J.W., Pearson, J., Blake, R. & van den Berg, A.V. (2009) Intermittent ambiguous stimuli: implicit memory causes periodic perceptual alternations. *Journal of vision*, **9**, 3 1-23.

Brascamp, J.W., van Ee, R., Pestman, W.R. & van den Berg, A.V. (2005) Distributions of alternation rates in various forms of bistable perception. *Journal of vision*, **5**, 287-298.

Brignani, D., Ruzzoli, M., Mauri, P. & Miniussi, C. (2013) Is transcranial alternating current stimulation effective in modulating brain oscillations? *PloS one*, **8**, e56589.

Britz, J., Landis, T. & Michel, C.M. (2009) Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb Cortex*, **19**, 55-65.

Brouwer, G.J. & van Ee, R. (2007) Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *J Neurosci*, **27**, 1015-1023.

Bubic, A., von Cramon, D.Y. & Schubotz, R.I. (2010) Prediction, cognition and the brain. *Front Hum Neurosci*, **4**, 25.

Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A. & Johnson, K.A. (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci*, **29**, 1860-1873.

Buffalo, E.A., Fries, P., Landman, R., Buschman, T.J. & Desimone, R. (2011) Laminar differences in gamma and alpha coherence in the ventral stream. *Proc Natl Acad Sci U S A*, **108**, 11262-11267.

Bullmore, E. & Sporns, O. (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, **10**, 186-198.

Buzsaki, G., Anastassiou, C.A. & Koch, C. (2012) The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nat Rev Neurosci*, **13**, 407-420.

Cabral-Calderin, Y., Schmidt-Samoa, C. & Wilke, M. (2015) Rhythmic gamma stimulation affects bistable perception. *Journal of cognitive neuroscience*, **27**, 1298-1307.

Calvert, G.A. (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb Cortex*, **11**, 1110-1123.

Carmel, D., Arcaro, M., Kastner, S. & Hasson, U. (2010a) How to create and use binocular rivalry. *Journal of visualized experiments: JoVE*.

Carmel, D., Walsh, V., Lavie, N. & Rees, G. (2010b) Right parietal TMS shortens dominance durations in binocular rivalry. *Curr Biol*, **20**, R799-800.

Carter, M. & Shieh, J.C. (2010) *Guide to research techniques in neuroscience*. Elsevier/Academic Press, Amsterdam ; Boston.

Carter, O. & Cavanagh, P. (2007) Onset rivalry: brief presentation isolates an early independent phase of perceptual competition. *PloS one*, **2**, e343.

Carter, O.L. & Pettigrew, J.D. (2003) A common oscillator for perceptual rivalries? *Perception*, **32**, 295-305.

Chadick, J.Z. & Gazzaley, A. (2011) Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nat Neurosci*, **14**, 830-832.

Chadwick, M.J., Hassabis, D., Weiskopf, N. & Maguire, E.A. (2010) Decoding individual episodic memory traces in the human hippocampus. *Curr Biol*, **20**, 544-547.

Chang, C.-C. & Lin, C.-J. (2011) LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, **2**, 27.

Chennu, S., Finoia, P., Kamau, E., Allanson, J., Williams, G.B., Monti, M.M., Noreika, V., Arnatkeviciute, A., Canales-Johnson, A. & Olivares, F. (2014) Spectral signatures of reorganised brain networks in disorders of consciousness.

Chong, S.C. & Blake, R. (2006) Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision research*, **46**, 1794-1803.

Chopin, A. & Mamassian, P. (2012) Predictive Properties of Visual Adaptation. *Curr Biol*.

Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and brain sciences*, **36**, 181-204.

Cohen, D. (1968) Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science*, **161**, 784-786.

Cole, M.W., Yarkoni, T., Repovs, G., Anticevic, A. & Braver, T.S. (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J Neurosci*, **32**, 8988-8999.

Cook, R.D. (1977) Detection of influential observation in linear regression. *Technometrics*, 15-18.

Cox, D.D. & Savoy, R.L. (2003) Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, **19**, 261-270.

Cox, R.W., Jesmanowicz, A. & Hyde, J.S. (1995) Real-time functional magnetic resonance imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, **33**, 230-236.

Crick, F. & Koch, C. (2003) A framework for consciousness. *Nat Neurosci*, **6**, 119-126.

Crossley, N.A., Mechelli, A., Scott, J., Carletti, F., Fox, P.T., McGuire, P. & Bullmore, E.T. (2014) The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain : a journal of neurology*, **137**, 2382-2395.

Daunizeau, J., den Ouden, H.E., Pessiglione, M., Kiebel, S.J., Stephan, K.E. & Friston, K.J. (2010) Observing the observer (I): meta-bayesian models of learning and decision-making. *PloS one*, **5**, e15554.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C. & Depaulis, A. (2008) Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS biology*, **6**, 2683-2697.

Dayan, P. (1998) A hierarchical model of binocular rivalry. *Neural computation*, **10**, 1119-1135.

deCharms, R.C. (2005) Control over brain activation and pain learned by using real-time functional MRI. *Proceedings of the National Academy of Sciences*, **102**, 18626-18631.

deCharms, R.C., Maeda, F., Glover, G.H., Ludlow, D., Pauly, J.M., Soneji, D., Gabrieli, J.D. & Mackey, S.C. (2005) Control over brain activation and pain learned by using real-time functional MRI. *Proc Natl Acad Sci U S A*, **102**, 18626-18631.

Dehaene, S. & Changeux, J.P. (2011) Experimental and theoretical approaches to conscious processing. *Neuron*, **70**, 200-227.

Deichmann, R., Schwarzbauer, C. & Turner, R. (2004) Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. *Neuroimage*, **21**, 757-767.

Denison, R.N., Piazza, E.A. & Silver, M.A. (2011) Predictive Context Influences Perceptual Selection during Binocular Rivalry. *Front Hum Neurosci*, **5**, 166.

Desimone, R. (1998) Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond B Biol Sci*, **353**, 1245-1255.

Dieter, K.C. & Tadin, D. (2011) Understanding attentional modulation of binocular rivalry: a framework based on biased competition. *Front Hum Neurosci*, **5**, 155.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Jr., Barch, D.M., Petersen, S.E. & Schlaggar, B.L. (2010) Prediction of individual brain maturity using fMRI. *Science*, **329**, 1358-1361.

Dosher, B.A., Sperling, G. & Wurst, S.A. (1986) Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision research*, **26**, 973-990.

Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L., Jr., Le Goualher, G., Bruce Pike, G. & Evans, A.C. (2000) A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex*, **10**, 454-463.

Einhauser, W., Martin, K.A. & Konig, P. (2004) Are switches in perception of the Necker cube related to eye position? *The European journal of neuroscience*, **20**, 2811-2818.

Einhauser, W., Stout, J., Koch, C. & Carter, O. (2008) Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc Natl Acad Sci U S A*, **105**, 1704-1709.

Ekman, M., Derrfuss, J., Tittgemeyer, M. & Fiebach, C.J. (2012) Predicting errors from reconfiguration patterns in human brain networks. *Proc Natl Acad Sci U S A*, **109**, 16714-16719.

Epstein, R. & Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature*, **392**, 598-601.

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

Ffytche, D.H., Howard, R.J., Brammer, M.J., David, A., Woodruff, P. & Williams, S. (1998) The anatomy of conscious vision: an fMRI study of visual hallucinations. *Nat Neurosci*, **1**, 738-742.

Fischer, J. & Whitney, D. (2014) Serial dependence in visual perception. *Nat Neurosci*.

Fleming, S.M., Weil, R.S., Nagy, Z., Dolan, R.J. & Rees, G. (2010) Relating introspective accuracy to individual differences in brain structure. *Science*, **329**, 1541-1543.

Fox, M.D. & Raichle, M.E. (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, **8**, 700-711.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C. & Raichle, M.E. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A*, **102**, 9673-9678.

Frackowiak, R.S., Friston, K.J., Frith, C.D., Dolan, R.J. & Mazziotta, J. (2004) Human brain function. *San Diego, CA*.

Fransson, P. (2005) Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum Brain Mapp*, **26**, 15-29.

Frässle, S., Sommer, J., Jansen, A., Naber, M. & Einhäuser, W. (2014) Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J Neurosci*, **34**, 1738-1747.

Freeman, E.D., Sterzer, P. & Driver, J. (2012) fMRI correlates of subjective reversals in ambiguous structure-from-motion. *Journal of vision*, **12**, 35.

Freeman, L.C. (1977) A set of measures of centrality based on betweenness. *Sociometry*, 35-41.

Friston, K. (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci*, **11**, 127-138.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G. & Mattout, J. (2008) Multiple sparse priors for the M/EEG inverse problem. *Neuroimage*, **39**, 1104-1120.

Friston, K., Moran, R. & Seth, A.K. (2013) Analysing connectivity with Granger causality and dynamic causal modelling. *Current opinion in neurobiology*, **23**, 172-178.

Friston, K.J. (1994) Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, **2**, 56-78.

Friston, K.J., Ashburner, J. & Heather, J. (2003a) Statistical parametric mapping. *Neuroscience Databases: A Practical Guide*, 237.

Friston, K.J., Harrison, L. & Penny, W. (2003b) Dynamic causal modelling. *NeuroImage*, **19**, 1273-1302.

Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L. & Naccache, L. (2009) Converging intracranial markers of conscious access. *PLoS biology*, **7**, e61.

Gallagher, R.M. & Arnold, D.H. (2014) Interpreting the temporal dynamics of perceptual rivalries. *Perception*, **43**, 1239-1248.

Gorlin, S., Meng, M., Sharma, J., Sugihara, H., Sur, M. & Sinha, P. (2012) Imaging prior information in the brain. *Proc Natl Acad Sci U S A*, **109**, 7935-7940.

Granger, C.W. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.

Greicius, M.D., Supekar, K., Menon, V. & Dougherty, R.F. (2009) Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex*, **19**, 72-78.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M. & Schoffelen, J.M. (2013) Good practice for conducting and reporting MEG research. *Neuroimage*, **65**, 349-363.

Gusnard, D.A., Raichle, M.E. & Raichle, M.E. (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci*, **2**, 685-694.

Hagberg, A., Swart, P. & S Chult, D. (2008) Exploring network structure, dynamics, and function using NetworkX. Los Alamos National Laboratory (LANL).

Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J. & Lounasmaa, O.V. (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, **65**, 413.

Hampson, M., Tokoglu, F., Shen, X., Scheinost, D., Papademetris, X. & Constable, R.T. (2012) Intrinsic brain connectivity related to age in young and middle aged adults. *PloS one*, **7**, e44067.

Hansen, P., Kringelbach, M. & Salmelin, R. (2010) *MEG: an introduction to methods*. Oxford university press.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L. & Pietrini, P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, **293**, 2425-2430.

Haynes, J.D., Deichmann, R. & Rees, G. (2005) Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, **438**, 496-499.

Haynes, J.D. & Rees, G. (2005) Predicting the stream of consciousness from activity in human visual cortex. *Curr Biol*, **15**, 1301-1307.

Haynes, J.D. & Rees, G. (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci*, **7**, 523-534.

Helfrich, R.F., Schneider, T.R., Rach, S., Trautmann-Lengsfeld, S.A., Engel, A.K. & Herrmann, C.S. (2014) Entrainment of brain oscillations by transcranial alternating current stimulation. *Curr Biol*, **24**, 333-339.

Hellyer, P.J., Shanahan, M., Scott, G., Wise, R.J., Sharp, D.J. & Leech, R. (2014) The control of global brain dynamics: opposing actions of frontoparietal control and default mode networks on attention. *J Neurosci*, **34**, 451-461.

Helmholtz, H.v. (1910) Concerning the Perceptions in General *The Perceptions of Vision*. The Optical Society of America, New York, pp. 1-37.

Hesselmann, G., Sadaghiani, S., Friston, K.J. & Kleinschmidt, A. (2010) Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PloS one*, **5**, e9926.

Hohwy, J., Roepstorff, A. & Friston, K. (2008) Predictive coding explains binocular rivalry: an epistemological review. *Cognition*, **108**, 687-701.

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R. & Hagmann, P. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci U S A*, **106**, 2035-2040.

Howard, I.P. & Rogers, B.J. (1995) *Binocular vision and stereopsis*. Oxford University Press.

Hubel, D.H. & Wiesel, T.N. (1968) Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, **195**, 215-243.

Hupe, J.M., Lamirel, C. & Lorenceau, J. (2009) Pupil dynamics during bistable motion perception. *Journal of vision*, **9**, 10.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J. & Turner, R. (2002) Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage*, **16**, 217-240.

Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., den Ouden, H.E. & Stephan, K.E. (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron*, **80**, 519-530.

Isoglu-Alkac, U., Basar-Eroglu, C., Ademoglu, A., Demiralp, T., Miener, M. & Stadler, M. (2000) Alpha activity decreases during the perception of Necker cube reversals: an application of wavelet transform. *Biological cybernetics*, **82**, 313-320.

Jiang, Y., Pantle, A.J. & Mark, L.S. (1998) Visual inertia of rotating 3-D objects. *Perception & psychophysics*, **60**, 275-286.

Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M. & Matthews, P.M. (2005) Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cereb Cortex*, **15**, 31-39.

Kamitani, Y. & Sawahata, Y. (2010) Spatial smoothing hurts localization but not information: pitfalls for brain mappers. *Neuroimage*, **49**, 1949-1952.

Kamitani, Y. & Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci*, **8**, 679-685.

Kanai, R., Bahrami, B. & Rees, G. (2010) Human parietal cortex structure predicts individual differences in perceptual rivalry. *Curr Biol*, **20**, 1626-1630.

Kanai, R., Carmel, D., Bahrami, B. & Rees, G. (2011) Structural and functional fractionation of right superior parietal cortex in bistable perception. *Curr Biol*, **21**, R106-107.

Kanai, R. & Rees, G. (2011) The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci*, **12**, 231-242.

Kanwisher, N., McDermott, J. & Chun, M.M. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, **17**, 4302-4311.

Karten, A., Pantazatos, S.P., Khalil, D., Zhang, X. & Hirsch, J. (2013) Dynamic coupling between the lateral occipital-cortex, default-mode, and frontoparietal networks during bistable perception. *Brain Connect*, **3**, 286-293.

Kastner, S., O'Connor, D.H., Fukui, M.M., Fehd, H.M., Herwig, U. & Pinsk, M.A. (2004) Functional imaging of the human lateral geniculate nucleus and pulvinar. *Journal of neurophysiology*, **91**, 438-448.

Kiebel, S. & Holmes, A. *The general linear model*. Academic Press.

Kleinschmidt, A., Buchel, C., Zeki, S. & Frackowiak, R.S. (1998) Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings. Biological sciences / The Royal Society*, **265**, 2427-2433.

Kloosterman, N.A., Meindertsma, T., Hillebrand, A., van Dijk, B.W., Lamme, V.A. & Donner, T.H. (2014) Top-Down Modulation in Human Visual Cortex Predicts the Stability of a Perceptual Illusion. *Journal of neurophysiology*, jn 00338 02014.

Knapen, T., Brascamp, J., Pearson, J., van Ee, R. & Blake, R. (2011) The role of frontal and parietal brain areas in bistable perception. *J Neurosci*, **31**, 10293-10301.

Kok, P., Jehee, J.F. & de Lange, F.P. (2012) Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, **75**, 265-270.

Kondo, H.M., Kitagawa, N., Kitamura, M.S., Koizumi, A., Nomura, M. & Kashino, M. (2012) Separability and commonality of auditory and visual bistable perception. *Cereb Cortex*, **22**, 1915-1922.

Kornmeier, J., Hein, C.M. & Bach, M. (2009) Multistable perception: when bottom-up and top-down coincide. *Brain and cognition*, **69**, 138-147.

Kreiman, G., Fried, I. & Koch, C. (2002) Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc Natl Acad Sci U S A*, **99**, 8378-8383.

Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., Turner, R. & et al. (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci U S A*, **89**, 5675-5679.

Lack, L.C. (1978) *Selective attention and the control of binocular rivalry*. Mouton De Gruyter.

Lavie, N., Hirst, A., de Fockert, J.W. & Viding, E. (2004) Load theory of selective attention and cognitive control. *Journal of experimental psychology. General*, **133**, 339-354.

Lee, S.H. & Blake, R. (2002) V1 activity is reduced during binocular rivalry. *Journal of vision*, **2**, 618-626.

Lee, S.H., Blake, R. & Heeger, D.J. (2005) Traveling waves of activity in primary visual cortex during binocular rivalry. *Nat Neurosci*, **8**, 22-23.

Lehky, S.R. & Maunsell, J.H. (1996) No binocular rivalry in the LGN of alert macaque monkeys. *Vision research*, **36**, 1225-1234.

Leopold, D.A. & Logothetis, N.K. (1999) Multistable phenomena: changing views in perception. *Trends Cogn Sci*, **3**, 254-264.

Leopold, D.A., Wilke, M., Maier, A. & 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.

Ling, S., Pratte, M.S. & Tong, F. (2015) Attention alters orientation processing in the human lateral geniculate nucleus. *Nature Neuroscience*, **18**, 496-+.

Livingstone, M. & Hubel, D. (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, **240**, 740-749.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature*, **412**, 150-157.

Logothetis, N.K. & Schall, J.D. (1990) Binocular motion rivalry in macaque monkeys: eye dominance and tracking eye movements. *Vision research*, **30**, 1409-1419.

Lohmann, G., Margulies, D.S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., Schloegl, H., Stumvoll, M., Villringer, A. & Turner, R. (2010) Eigenvector Centrality Mapping for Analyzing Connectivity Patterns in fMRI Data of the Human Brain. *PloS one*, **5**.

Lopez, J.D., Litvak, V., Espinosa, J.J., Friston, K. & Barnes, G.R. (2014) Algorithmic procedures for Bayesian MEG/EEG source reconstruction in SPM. *Neuroimage*, **84**, 476-487.

Lord, L.D., Allen, P., Expert, P., Howes, O., Broome, M., Lambiotte, R., Fusar-Poli, P., Valli, I., McGuire, P. & Turkheimer, F.E. (2012) Functional brain networks before the onset of psychosis: A prospective fMRI study with graph theoretical analysis. *NeuroImage. Clinical*, **1**, 91-98.

Lumer, E.D., Friston, K.J. & Rees, G. (1998) Neural correlates of perceptual rivalry in the human brain. *Science*, **280**, 1930-1934.

Lumer, E.D. & Rees, G. (1999) Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc Natl Acad Sci U S A*, **96**, 1669-1673.

Lutti, A., Thomas, D.L., Hutton, C. & Weiskopf, N. (2013) High-resolution functional MRI at 3 T: 3D/2D echo-planar imaging with optimized physiological noise correction. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, **69**, 1657-1664.

Maier, A., Wilke, M., Logothetis, N.K. & Leopold, D.A. (2003) Perception of temporally interleaved ambiguous patterns. *Curr Biol*, **13**, 1076-1085.

Maldjian, J.A., Laurienti, P.J. & Burdette, J.H. (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage*, **21**, 450-455.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A. & Burdette, J.H. (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, **19**, 1233-1239.

Mamassian, P. & Goutcher, R. (2005) Temporal dynamics in bistable perception. *Journal of vision*, **5**, 361-375.

Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C. & Kennedy, H. (2014) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex*, **24**, 17-36.

Mars, R.B., Jbabdi, S., Sallet, J., O'Reilly, J.X., Croxson, P.L., Olivier, E., Noonan, M.P., Bergmann, C., Mitchell, A.S., Baxter, M.G., Behrens, T.E., Johansen-Berg, H., Tomassini, V., Miller, K.L. & Rushworth, M.F. (2011) Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci*, **31**, 4087-4100.

Mathys, C., Daunizeau, J., Friston, K.J. & Stephan, K.E. (2011) A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci*, **5**, 39.

Mathys, C.D., Lomakina, E.I., Daunizeau, J., Iglesias, S., Brodersen, K.H., Friston, K.J. & Stephan, K.E. (2014) Uncertainty in perception and the Hierarchical Gaussian Filter. *Front Hum Neurosci*, **8**, 825.

Megumi, F., Bahrami, B., Ryota, K. & Rees, G. (2015a) Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. *Neuroimage*, **107**, 190-197.

Megumi, F., Yamashita, A., Kawato, M. & Imamizu, H. (2015b) Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Front Hum Neurosci*, **9**, 160.

Meng, M. & Tong, F. (2004) Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of vision*, **4**, 539-551.

Menon, V. & Uddin, L.Q. (2010) Saliency, switching, attention and control: a network model of insula function. *Brain structure & function*, **214**, 655-667.

Meppelink, A.M., de Jong, B.M., Renken, R., Leenders, K.L., Cornelissen, F.W. & van Laar, T. (2009) Impaired visual processing preceding image recognition in Parkinson's disease patients with visual hallucinations. *Brain : a journal of neurology*, **132**, 2980-2993.

Miller, S.M., Gynther, B.D., Heslop, K.R., Liu, G.B., Mitchell, P.B., Ngo, T.T., Pettigrew, J.D. & Geffen, L.B. (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. & Wright, M.J. (2010) Genetic contribution to individual variation in binocular rivalry rate. *Proc Natl Acad Sci U S A*, **107**, 2664-2668.

Miller, S.M., Liu, G.B., Ngo, T.T., Hooper, G., Riek, S., Carson, R.G. & Pettigrew, J.D. (2000) Interhemispheric switching mediates perceptual rivalry. *Curr Biol*, **10**, 383-392.

Mitchell, J.F., Stoner, G.R. & Reynolds, J.H. (2004) Object-based attention determines dominance in binocular rivalry. *Nature*, **429**, 410-413.

Moreno-Bote, R., Knill, D.C. & Pouget, A. (2011) Bayesian sampling in visual perception. *Proc Natl Acad Sci U S A*, **108**, 12491-12496.

Mountcastle, V.B. (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, **20**, 408-434.

Muckli, L., Kohler, A., Kriegeskorte, N. & Singer, W. (2005) Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS biology*, **3**, e265.

Murphy, A.P., Leopold, D.A. & Welchman, A.E. (2014) Perceptual memory drives learning of retinotopic biases for bistable stimuli. *Frontiers in psychology*, **5**, 60.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P. & Woods, D.L. (2002) Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci U S A*, **99**, 15164-15169.

Naber, M., Frassle, S. & Einhauser, W. (2011) Perceptual rivalry: reflexes reveal the gradual nature of visual awareness. *PloS one*, **6**, e20910.

Nawrot, M. & Blake, R. (1989) Neural integration of information specifying structure from stereopsis and motion. *Science*, **244**, 716-718.

Necker, L.A. (1832) LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid.

Neisser, U. (2014) *Cognitive Psychology: Classic Edition*. Psychology Press.

Noest, A.J., van Ee, R., Nijs, M.M. & van Wezel, R.J. (2007) Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *Journal of vision*, **7**, 10.

Norman, K.A., Polyn, S.M., Detre, G.J. & Haxby, J.V. (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci*, **10**, 424-430.

O'Connor, D.H., Fukui, M.M., Pinsk, M.A. & Kastner, S. (2002) Attention modulates responses in the human lateral geniculate nucleus. *Nat Neurosci*, **5**, 1203-1209.

Ogawa, S., Lee, T.M., Kay, A.R. & Tank, D.W. (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A*, **87**, 9868-9872.

Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H. & Ugurbil, K. (1992) Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci U S A*, **89**, 5951-5955.

Ooi, T.L. & He, Z.J. (1999) Binocular rivalry and visual awareness: the role of attention. *Perception*, **28**, 551-574.

Op de Beeck, H.P. (2010) Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage*, **49**, 1943-1948.

Orbach, J., Ehrlich, D. & Heath, H.A. (1963) REVERSIBILITY OF THE NECKER CUBE: I. AN EXAMINATION OF THE CONCEPT OF" SATIATION OF ORIENTATION". *Perceptual and motor skills*, **17**, 439-458.

Palmer, S.E. (1999) *Vision science : photons to phenomenology*. MIT Press, Cambridge, Mass.

Panichello, M.F., Cheung, O.S. & Bar, M. (2012) Predictive feedback and conscious visual experience. *Frontiers in psychology*, **3**, 620.

Pearson, J. & Brascamp, J. (2008) Sensory memory for ambiguous vision. *Trends in cognitive sciences*, **12**, 334-341.

Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M. & Leff, A.P. (2010) Comparing families of dynamic causal models. *PLoS computational biology*, **6**, e1000709.

Penny, W.D., Stephan, K.E., Mechelli, A. & Friston, K.J. (2004) Comparing dynamic causal models. *Neuroimage*, **22**, 1157-1172.

Pereira, F., Mitchell, T. & Botvinick, M. (2009) Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, **45**, S199-209.

Pernet, C.R., Wilcox, R. & Rousselet, G.A. (2012) Robust correlation analyses: false positive and power validation using a new open source matlab toolbox. *Frontiers in psychology*, **3**, 606.

Pettigrew, J.D. & Miller, S.M. (1998) A 'sticky'interhemispheric switch in bipolar disorder? *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 2141-2148.

Pinto, Y., van der Leij, A.R., Sligte, I.G., Lamme, V.A. & Scholte, H.S. (2013) Bottom-up and top-down attention are independent. *Journal of vision*, **13**, 16.

Poggio, T. & Koch, C. (1985) Ill-Posed Problems in Early Vision - from Computational Theory to Analog Networks. *Proc R Soc Ser B-Bio*, **226**, 303-323.

Polonsky, A., Blake, R., Braun, J. & Heeger, D.J. (2000) Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat Neurosci*, **3**, 1153-1159.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L. & Petersen, S.E. (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, **59**, 2142-2154.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L. & Petersen, S.E. (2011) Functional network organization of the human brain. *Neuron*, **72**, 665-678.

Power, J.D., Schlaggar, B.L., Lessov-Schlaggar, C.N. & Petersen, S.E. (2013) Evidence for hubs in human functional brain networks. *Neuron*, **79**, 798-813.

Pressnitzer, D. & Hupe, J.M. (2006) Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr Biol*, **16**, 1351-1357.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A. & Shulman, G.L. (2001) A default mode of brain function. *Proc Natl Acad Sci U S A*, **98**, 676-682.

Rao, R.P. (2005) Bayesian inference and attentional modulation in the visual cortex. *Neuroreport*, **16**, 1843-1848.

Rao, R.P. & Ballard, D.H. (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*, **2**, 79-87.

Rauss, K. & Pourtois, G. (2013) What is Bottom-Up and What is Top-Down in Predictive Coding? *Frontiers in psychology*, **4**, 276.

Raymond, J.E., Shapiro, K.L. & Arnell, K.M. (1992) Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of experimental psychology. Human perception and performance*, **18**, 849-860.

Rees, G. (2007) Neural correlates of the contents of visual awareness in humans. *Philos Trans R Soc Lond B Biol Sci*, **362**, 877-886.

Rescorla, R.A. & Wagner, A.R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement *Classical conditioning: Current research and theory*.

Rich, A.N., Williams, M.A., Puce, A., Syngeniotis, A., Howard, M.A., McGlone, F. & Mattingley, J.B. (2006) Neural correlates of imagined and synaesthetic colours. *Neuropsychologia*, **44**, 2918-2925.

Rigoux, L., Stephan, K.E., Friston, K.J. & Daunizeau, J. (2014) Bayesian model selection for group studies - revisited. *Neuroimage*, **84**, 971-985.

Rousseeuw, P.J. (1984) Least median of squares regression. *Journal of the American statistical association*, **79**, 871-880.

Rousseeuw, P.J. & Driessen, K.V. (1999) A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, **41**, 212-223.

Roy, C.S. & Sherrington, C.S. (1890) On the Regulation of the Blood-supply of the Brain. *The Journal of physiology*, **11**, 85-158 117.

Rubinov, M. & Sporns, O. (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, **52**, 1059-1069.

Ruff, C.C., Driver, J. & Bestmann, S. (2009) Combining TMS and fMRI: from 'virtual lesions' to functional-network accounts of cognition. *Cortex; a journal devoted to the study of the nervous system and behavior*, **45**, 1043-1049.

Scharnowski, F., Hutton, C., Josephs, O., Weiskopf, N. & Rees, G. (2012) Improving visual perception through neurofeedback. *J Neurosci*, **32**, 17830-17841.

Schneider, K.A., Richter, M.C. & Kastner, S. (2004) Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. *J Neurosci*, **24**, 8975-8985.

Schrater, P.R. & Sundareswara, R. (Year) Theory and dynamics of perceptual bistability. Advances in neural information processing systems. City. p. 1217-1224.

Schultz, J., Imamizu, H., Kawato, M. & Frith, C.D. (2004) Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of cognitive neuroscience*, **16**, 1695-1705.

Schurger, A. (2009) A very inexpensive MRI-compatible method for dichoptic visual stimulation. *Journal of neuroscience methods*, **177**, 199-202.

Seghier, M.L., Zeidman, P., Neufeld, N.H., Leff, A.P. & Price, C.J. (2010) Identifying abnormal connectivity in patients using dynamic causal modeling of FMRI responses. *Front Syst Neurosci*, **4**.

Seitz, A.R., Kim, D. & Watanabe, T. (2009) Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, **61**, 700-707.

Sergent, J., Ohta, S. & MacDonald, B. (1992) Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain : a journal of neurology*, **115 Pt 1**, 15-36.

Seth, A.K. (2010) A MATLAB toolbox for Granger causal connectivity analysis. *Journal of neuroscience methods*, **186**, 262-273.

Shannon, C.E., McCarthy, J. & Ashby, W.R. (1956) *Automata studies*. Princeton University Press Princeton, NJ.

Shannon, R.W., Patrick, C.J., Jiang, Y., Bernat, E. & He, S. (2011) Genes contribute to the switching dynamics of bistable perception. *Journal of vision*, **11**.

Shibata, K., Watanabe, T., Sasaki, Y. & Kawato, M. (2011) Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science*, **334**, 1413-1415.

Shimono, M. & Niki, K. (2013) Global mapping of the whole-brain network underlining binocular rivalry. *Brain Connect*, **3**, 212-221.

Shpiro, A., Moreno-Bote, R., Rubin, N. & Rinzel, J. (2009) Balance between noise and adaptation in competition models of perceptual bistability. *Journal of computational neuroscience*, **27**, 37-54.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R. & Beckmann, C.F. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, **106**, 13040-13045.

Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D. & Woolrich, M.W. (2011) Network modelling methods for FMRI. *Neuroimage*, **54**, 875-891.

Smola, A.J. & Scholkopf, B. (2004) A tutorial on support vector regression. *Stat Comput*, **14**, 199-222.

Song, X.W., Dong, Z.Y., Long, X.Y., Li, S.F., Zuo, X.N., Zhu, C.Z., He, Y., Yan, C.G. & Zang, Y.F. (2011) REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PloS one*, **6**, e25031.

Sporns, O. (2014) Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci*, **17**, 652-660.

Sporns, O., Tononi, G. & Edelman, G.M. (2000) Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb Cortex*, **10**, 127-141.

Spratling, M.W. (2008a) Predictive coding as a model of biased competition in visual attention. *Vision research*, **48**, 1391-1408.

Spratling, M.W. (2008b) Reconciling predictive coding and biased competition models of cortical function. *Frontiers in computational neuroscience*, **2**, 4.

Srinivasan, M.V., Laughlin, S.B. & Dubs, A. (1982) Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci*, **216**, 427-459.

Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J. & Friston, K.J. (2009) Bayesian model selection for group studies. *Neuroimage*, **46**, 1004-1017.

Stephan, K.E., Penny, W.D., Moran, R.J., den Ouden, H.E., Daunizeau, J. & Friston, K.J. (2010) Ten simple rules for dynamic causal modeling. *Neuroimage*, **49**, 3099-3109.

Sterzer, P., Frith, C. & Petrovic, P. (2008) Believing is seeing: expectations alter visual awareness. *Curr Biol*, **18**, R697-698.

Sterzer, P. & Kleinschmidt, A. (2007) A neural basis for inference in perceptual ambiguity. *Proc Natl Acad Sci U S A*, **104**, 323-328.

Sterzer, P. & Kleinschmidt, A. (2010) Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain structure & function*, **214**, 611-622.

Sterzer, P., Kleinschmidt, A. & Rees, G. (2009) The neural bases of multistable perception. *Trends Cogn Sci*, **13**, 310-318.

Summerfield, C. & de Lange, F.P. (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci*, **15**, 745-756.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J. & Hirsch, J. (2006) Predictive codes for forthcoming perception in the frontal cortex. *Science*, **314**, 1311-1314.

Sutton, R.S. (Year) Gain adaptation beats least squares. City.

Tong, F. & Engel, S.A. (2001) Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, **411**, 195-199.

Tong, F., Meng, M. & Blake, R. (2006) Neural bases of binocular rivalry. *Trends Cogn Sci*, **10**, 502-511.

Tong, F., Nakayama, K., Vaughan, J.T. & Kanwisher, N. (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, **21**, 753-759.

Tootell, R.B., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R. & Belliveau, J.W. (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci*, **15**, 3215-3230.

Treue, S., Husain, M. & Andersen, R.A. (1991) Human perception of structure from motion. *Vision research*, **31**, 59-75.

Troje, N.F. & McAdam, M. (2010) The viewing-from-above bias and the silhouette illusion. *Iperception*, **1**, 143-148.

Tsuchiya, N. & Koch, C. (2005) Continuous flash suppression reduces negative afterimages. *Nat Neurosci*, **8**, 1096-1101.

Tsuchiya, N., Koch, C., Gilroy, L.A. & Blake, R. (2006) Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of vision*, **6**, 1068-1078.

Turner, R., Bihan, D.L., Moonen, C.T., Despres, D. & Frank, J. (1991) Echo - planar time course MRI of cat brain oxygenation changes. *Magnetic Resonance in Medicine*, **22**, 159-166.

Ullman, S. (1979) The Interpretation of Structure from Motion. *Proceedings of the Royal Society B: Biological Sciences*, **203**, 405-426.

van Dam, L.C. & van Ee, R. (2005) The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision research*, **45**, 2417-2435.

van Dam, L.C. & van Ee, R. (2006) The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision research*, **46**, 787-799.

van den Heuvel, M.P. & Sporns, O. (2011) Rich-club organization of the human connectome. *J Neurosci*, **31**, 15775-15786.

van den Heuvel, M.P., Stam, C.J., Kahn, R.S. & Hulshoff Pol, H.E. (2009) Efficiency of functional brain networks and intellectual performance. *J Neurosci*, **29**, 7619-7624.

Van Dijk, K.R., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W. & Buckner, R.L. (2010) Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of neurophysiology*, **103**, 297-321.

van Ee, R., van Boxtel, J.J., Parker, A.L. & Alais, D. (2009) Multisensory congruency as a mechanism for attentional control over perceptual selection. *J Neurosci*, **29**, 11641-11649.

van Ee, R., van Dam, L.C. & Brouwer, G.J. (2005) Voluntary control and the dynamics of perceptual bi-stability. *Vision research*, **45**, 41-55.

Van Essen, D.C. (2005) Corticocortical and thalamocortical information flow in the primate visual system. *Progress in brain research*, **149**, 173-185.

van Loon, A.M., Knapen, T., Scholte, H.S., St John-Saaltink, E., Donner, T.H. & Lamme, V.A. (2013) GABA shapes the dynamics of bistable perception. *Curr Biol*, **23**, 823-827.

Verboven, S. & Hubert, M. (2005) LIBRA: a MATLAB library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, **75**, 127-136.

Vincent, J.L., Patel, G.H., Fox, M.D., Snyder, A.Z., Baker, J.T., Van Essen, D.C., Zempel, J.M., Snyder, L.H., Corbetta, M. & Raichle, M.E. (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, **447**, 83-86.

von Schiller, P. (1933) Stroboskopische Alternativbewegungen. *Psychologische Forschung*, **17**, 179-214.

Vrba, J. (2002) Magnetoencephalography: the art of finding a needle in a haystack. *Physica C: Superconductivity*, **368**, 1-9.

Wager, T.D., Atlas, L.Y., Leotti, L.A. & Rilling, J.K. (2011) Predicting individual differences in placebo analgesia: contributions of brain activity during anticipation and pain experience. *J Neurosci*, **31**, 439-452.

Wallach, H. & O'Connell, D.N. (1953) The kinetic depth effect. *Journal of experimental psychology*, **45**, 205-217.

Wallis, G. & Ringelhan, S. (2013) The dynamics of perceptual rivalry in bistable and tristable perception. *Journal of vision*, **13**.

Wang, M., Arteaga, D. & He, B.J. (2013) Brain mechanisms for simple perception and bistable perception. *Proc Natl Acad Sci U S A*, **110**, E3350-3359.

Warren, D.E., Power, J.D., Bruss, J., Denburg, N.L., Waldron, E.J., Sun, H., Petersen, S.E. & Tranel, D. (2014) Network measures predict neuropsychological outcome after brain injury. *Proc Natl Acad Sci U S A*, **111**, 14247-14252.

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K. & Logothetis, N. (2011) Attention but not awareness modulates the

BOLD signal in the human V1 during binocular suppression. *Science*, **334**, 829-831.

Watanabe, T., Masuda, N., Megumi, F., Kanai, R. & Rees, G. (2014) Energy landscape and dynamics of brain activity during human bistable perception. *Nature communications*, **5**, 4765.

Weilnhammer, V.A., Ludwig, K., Hesselmann, G. & Sterzer, P. (2013) Frontoparietal cortex mediates perceptual transitions in bistable perception. *J Neurosci*, **33**, 16009-16015.

Weiskopf, N., Mathiak, K., Bock, S.W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R. & Birbaumer, N. (2004) Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE transactions on bio-medical engineering*, **51**, 966-970.

Wilke, M., Mueller, K.M. & Leopold, D.A. (2009) Neural activity in the visual thalamus reflects perceptual suppression. *Proc Natl Acad Sci U S A*, **106**, 9465-9470.

Wilson, H.R. (2003) Computational evidence for a rivalry hierarchy in vision. *Proc Natl Acad Sci U S A*, **100**, 14499-14503.

Wunderlich, K., Schneider, K.A. & Kastner, S. (2005) Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nat Neurosci*, **8**, 1595-1602.

Wyart, V., de Gardelle, V., Scholl, J. & Summerfield, C. (2012a) Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron*, **76**, 847-858.

Wyart, V., Nobre, A.C. & Summerfield, C. (2012b) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A*, **109**, 3593-3598.

Zaretskaya, N., Thielscher, A., Logothetis, N.K. & Bartels, A. (2010) Disrupting parietal function prolongs dominance durations in binocular rivalry. *Curr Biol*, **20**, 2106-2111.

Zhang, P., Jamison, K., Engel, S., He, B. & He, S. (2011) Binocular rivalry requires visual attention. *Neuron*, **71**, 362-369.

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y. & Boddaert, N. (2006) Autism, the superior temporal sulcus and social perception. *Trends Neurosci*, **29**, 359-366.

APPENDIX

ROI list used in chapter 5

	MNI space			Network label
ROI	**X**	**Y**	**Z**	
1	-25	-98	-12	Uncertain
2	27	-97	-13	Uncertain
3	24	32	-18	Uncertain
4	-56	-45	-24	Uncertain
5	8	41	-24	Uncertain
6	-21	-22	-20	Uncertain
7	17	-28	-17	Uncertain
8	-37	-29	-26	Uncertain
9	65	-24	-19	Uncertain
10	52	-34	-27	Uncertain
11	55	-31	-17	Uncertain
12	34	38	-12	Uncertain
13	-7	-52	61	Sensory/somatomotor Hand
14	-14	-18	40	Sensory/somatomotor Hand
15	0	-15	47	Sensory/somatomotor Hand
16	10	-2	45	Sensory/somatomotor Hand
17	-7	-21	65	Sensory/somatomotor Hand
18	-7	-33	72	Sensory/somatomotor Hand
19	13	-33	75	Sensory/somatomotor Hand
20	-54	-23	43	Sensory/somatomotor Hand
21	29	-17	71	Sensory/somatomotor Hand
22	10	-46	73	Sensory/somatomotor Hand
23	-23	-30	72	Sensory/somatomotor Hand
24	-40	-19	54	Sensory/somatomotor Hand
25	29	-39	59	Sensory/somatomotor Hand
26	50	-20	42	Sensory/somatomotor Hand
27	-38	-27	69	Sensory/somatomotor Hand
28	20	-29	60	Sensory/somatomotor Hand
29	44	-8	57	Sensory/somatomotor Hand
30	-29	-43	61	Sensory/somatomotor Hand
31	10	-17	74	Sensory/somatomotor Hand
32	22	-42	69	Sensory/somatomotor Hand
33	-45	-32	47	Sensory/somatomotor Hand
34	-21	-31	61	Sensory/somatomotor Hand
35	-13	-17	75	Sensory/somatomotor Hand
36	42	-20	55	Sensory/somatomotor Hand
37	-38	-15	69	Sensory/somatomotor Hand
38	-16	-46	73	Sensory/somatomotor Hand
39	2	-28	60	Sensory/somatomotor Hand
40	3	-17	58	Sensory/somatomotor Hand
41	38	-17	45	Sensory/somatomotor Hand
42	-49	-11	35	Sensory/somatomotor Mouth
43	36	-9	14	Sensory/somatomotor Mouth
44	51	-6	32	Sensory/somatomotor Mouth
45	-53	-10	24	Sensory/somatomotor Mouth

46	66	-8	25	Sensory/somatomotor Mouth
47	-3	2	53	Cingulo-opercular Task Control
48	54	-28	34	Cingulo-opercular Task Control
49	19	-8	64	Cingulo-opercular Task Control
50	-16	-5	71	Cingulo-opercular Task Control
51	-10	-2	42	Cingulo-opercular Task Control
52	37	1	-4	Cingulo-opercular Task Control
53	13	-1	70	Cingulo-opercular Task Control
54	7	8	51	Cingulo-opercular Task Control
55	-45	0	9	Cingulo-opercular Task Control
56	49	8	-1	Cingulo-opercular Task Control
57	-34	3	4	Cingulo-opercular Task Control
58	-51	8	-2	Cingulo-opercular Task Control
59	-5	18	34	Cingulo-opercular Task Control
60	36	10	1	Cingulo-opercular Task Control
61	32	-26	13	Auditory
62	65	-33	20	Auditory
63	58	-16	7	Auditory
64	-38	-33	17	Auditory
65	-60	-25	14	Auditory
66	-49	-26	5	Auditory
67	43	-23	20	Auditory
68	-50	-34	26	Auditory
69	-53	-22	23	Auditory
70	-55	-9	12	Auditory
71	56	-5	13	Auditory
72	59	-17	29	Auditory
73	-30	-27	12	Auditory
74	-41	-75	26	Default mode
75	6	67	-4	Default mode
76	8	48	-15	Default mode
77	-13	-40	1	Default mode
78	-18	63	-9	Default mode
79	-46	-61	21	Default mode
80	43	-72	28	Default mode
81	46	16	-30	Default mode
82	-68	-23	-16	Default mode
83	-58	-26	-15	Uncertain
84	27	16	-17	Uncertain
85	-44	-65	35	Default mode
86	-39	-75	44	Default mode
87	-7	-55	27	Default mode
88	6	-59	35	Default mode
89	-11	-56	16	Default mode
90	-3	-49	13	Default mode
91	8	-48	31	Default mode
92	15	-63	26	Default mode
93	-2	-37	44	Default mode
94	11	-54	17	Default mode
95	52	-59	36	Default mode
96	23	33	48	Default mode
97	-10	39	52	Default mode
98	-16	29	53	Default mode
99	-35	20	51	Default mode

100	22	39	39	Default mode
101	13	55	38	Default mode
102	-10	55	39	Default mode
103	-20	45	39	Default mode
104	6	54	16	Default mode
105	6	64	22	Default mode
106	-7	51	-1	Default mode
107	9	54	3	Default mode
108	-3	44	-9	Default mode
109	8	42	-5	Default mode
110	-11	45	8	Default mode
111	-2	38	36	Default mode
112	-3	42	16	Default mode
113	-20	64	19	Default mode
114	-8	48	23	Default mode
115	65	-12	-19	Default mode
116	-56	-13	-10	Default mode
117	-58	-30	-4	Default mode
118	65	-31	-9	Default mode
119	-68	-41	-5	Default mode
120	13	30	59	Default mode
121	12	36	20	Default mode
122	52	-2	-16	Default mode
123	-26	-40	-8	Default mode
124	27	-37	-13	Default mode
125	-34	-38	-16	Default mode
126	28	-77	-32	Default mode
127	52	7	-30	Default mode
128	-53	3	-27	Default mode
129	47	-50	29	Default mode
130	-49	-42	1	Default mode
131	-31	19	-19	Uncertain
132	-2	-35	31	Memory retrieval
133	-7	-71	42	Memory retrieval
134	11	-66	42	Memory retrieval
135	4	-48	51	Memory retrieval
136	-46	31	-13	Default mode
137	-10	11	67	Ventral attention
138	49	35	-12	Default mode
139	8	-91	-7	Uncertain
140	17	-91	-14	Uncertain
141	-12	-95	-13	Uncertain
142	18	-47	-10	Visual
143	40	-72	14	Visual
144	8	-72	11	Visual
145	-8	-81	7	Visual
146	-28	-79	19	Visual
147	20	-66	2	Visual
148	-24	-91	19	Visual
149	27	-59	-9	Visual
150	-15	-72	-8	Visual
151	-18	-68	5	Visual
152	43	-78	-12	Visual
153	-47	-76	-10	Visual

154	-14	-91	31	Visual
155	15	-87	37	Visual
156	29	-77	25	Visual
157	20	-86	-2	Visual
158	15	-77	31	Visual
159	-16	-52	-1	Visual
160	42	-66	-8	Visual
161	24	-87	24	Visual
162	6	-72	24	Visual
163	-42	-74	0	Visual
164	26	-79	-16	Visual
165	-16	-77	34	Visual
166	-3	-81	21	Visual
167	-40	-88	-6	Visual
168	37	-84	13	Visual
169	6	-81	6	Visual
170	-26	-90	3	Visual
171	-33	-79	-13	Visual
172	37	-81	1	Visual
173	-44	2	46	Fronto-parietal Task Control
174	48	25	27	Fronto-parietal Task Control
175	-47	11	23	Fronto-parietal Task Control
176	-53	-49	43	Fronto-parietal Task Control
177	-23	11	64	Fronto-parietal Task Control
178	58	-53	-14	Fronto-parietal Task Control
179	24	45	-15	Fronto-parietal Task Control
180	34	54	-13	Fronto-parietal Task Control
181	-21	41	-20	Uncertain
182	-18	-76	-24	Uncertain
183	35	-67	-34	Uncertain
184	47	10	33	Fronto-parietal Task Control
185	-41	6	33	Fronto-parietal Task Control
186	-42	38	21	Fronto-parietal Task Control
187	38	43	15	Fronto-parietal Task Control
188	49	-42	45	Fronto-parietal Task Control
189	-28	-58	48	Fronto-parietal Task Control
190	44	-53	47	Fronto-parietal Task Control
191	32	14	56	Fronto-parietal Task Control
192	37	-65	40	Fronto-parietal Task Control
193	-42	-55	45	Fronto-parietal Task Control
194	40	18	40	Fronto-parietal Task Control
195	-34	55	4	Fronto-parietal Task Control
196	-42	45	-2	Fronto-parietal Task Control
197	33	-53	44	Fronto-parietal Task Control
198	43	49	-2	Fronto-parietal Task Control
199	-42	25	30	Fronto-parietal Task Control
200	-3	26	44	Fronto-parietal Task Control
201	11	-39	50	Salience
202	55	-45	37	Salience
203	42	0	47	Salience
204	31	33	26	Salience
205	48	22	10	Salience
206	-35	20	0	Salience
207	36	22	3	Salience

208	37	32	-2	Salience
209	34	16	-8	Salience
210	-11	26	25	Salience
211	-1	15	44	Salience
212	-28	52	21	Salience
213	0	30	27	Salience
214	5	23	37	Salience
215	10	22	27	Salience
216	31	56	14	Salience
217	26	50	27	Salience
218	-39	51	17	Salience
219	2	-24	30	Memory retrieval
220	6	-24	0	Subcortical
221	-2	-13	12	Subcortical
222	-10	-18	7	Subcortical
223	12	-17	8	Subcortical
224	-5	-28	-4	Subcortical
225	-22	7	-5	Subcortical
226	-15	4	8	Subcortical
227	31	-14	2	Subcortical
228	23	10	1	Subcortical
229	29	1	4	Subcortical
230	-31	-11	0	Subcortical
231	15	5	7	Subcortical
232	9	-4	6	Subcortical
233	54	-43	22	Ventral attention
234	-56	-50	10	Ventral attention
235	-55	-40	14	Ventral attention
236	52	-33	8	Ventral attention
237	51	-29	-4	Ventral attention
238	56	-46	11	Ventral attention
239	53	33	1	Ventral attention
240	-49	25	-1	Ventral attention
241	-16	-65	-20	Cerebellar
242	-32	-55	-25	Cerebellar
243	22	-58	-23	Cerebellar
244	1	-62	-18	Cerebellar
245	33	-12	-34	Uncertain
246	10	-62	61	Dorsal attention
247	-52	-63	5	Dorsal attention
248	-47	-51	-21	Uncertain
249	46	-47	-17	Uncertain
250	47	-30	49	Sensory/somatomotor Hand
251	22	-65	48	Dorsal attention
252	46	-59	4	Dorsal attention
253	25	-58	60	Dorsal attention
254	-33	-46	47	Dorsal attention
255	-27	-71	37	Dorsal attention
256	-32	-1	54	Dorsal attention
257	-42	-60	-9	Dorsal attention
258	-17	-59	64	Dorsal attention
259	29	-5	54	Dorsal attention