The Curious Brain:

Behavioural and Neural Mechanisms of Spatial Updating and Reward Surprise

Urs Schüffelgen

Department of Experimental Psychology

Linacre College

Michaelmas Term 2016

A dissertation in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Abstract

The Curious Brain: Behavioural and Neural Mechanisms of Spatial Updating and Reward Surprise

In an ever-changing environment, organisms learn about non-random events and use that information to form predictive models to facilitate the processing of expected stimuli or planned actions and their outcomes. Predictions about stimuli enable for an efficient and appropriate response and predictions about action outcomes determine the value and therefore likelihood of repetition of that action. In primate brains, these models have been described in the framework of spatial attention and reward expectation. Unexpected observations in space trigger an attentional reorienting process and surprising action outcomes are defined as reward prediction errors. These error signals can be used by the brain to update and refine its internal model that contains probabilistic information about events in the environment (Chapter 1). In what follows, I will report effects of spatial surprise and unexpected rewards on behaviour and brain activity in humans and macaques. I will show how humans react to visually surprising events and how they use that information to form future predictions (Chapter 2). Brains reacting to surprise have been primarily described in the attention literature but there are dissociable networks for an attentional shift and for an expectation updating process following unexpected but informative observations (Chapter 3). Macaques also form spatial predictions and violations of these predictions result in a behavioural cost. Furthermore, both expectations about the size of a reward and prediction errors influence behaviour (Chapter 4). FMRI in macaques performing a task reveals a network of reward sensitive areas. Nodes of this network however carry different information about the reward history and therefore play different roles in forming a reward expectation. Spatially surprising events activate networks similar to those described in human fMRI studies (Chapter 5). In summary, unexpected observations and action outcomes influence behaviour and activate brain networks that process the surprise as well as form new predictions.

Urs Schüffelgen
Linacre College, Michaelmas Term 2016
Department of Experimental Psychology
Submitted for the degree of Doctor of Philosophy
Approximate word count: 45,000

Acknowledgements

First, I would like to thank my supervisor Matthew who has led me through my DPhil with patience and encouragement. His calm, thoughtful and knowledgeable way of supervising helped me beyond anything I can express here and I could not have wished for a better supervisor.

I would also like to thank my co-supervisors Jill, Jerome and Nils who all helped me selflessly during different stages of my long journey to finish this dissertation and without whom none of the projects could have been done.

A huge thank you goes to the entire Rushworth Lab, especially Bolton, Rogier, Jacquie, Georgios, Franz and Marco for helping me with scanning or commenting on earlier versions of this dissertation.

Also, without the care and effort from Maria, the whole Level 3 team and the vets, I could not have collected the data so many thanks to them too.

FMRI can be a tricky business but here in Oxford we are fortunate enough to have one of the world's leading MRI units just around the corner. Thanks to Karla, Stuart, Wilfred and Sam H. for continuous MRI physics support.

Doing a DPhil is not only about lab work though and I wish to express my thanks to all the people who made my life in Oxford a wonderful one: First of all Philly but also thanks to the (extended) Linacre crew.

Of course, none of this would have been possible without my parents and I wish to dedicate this dissertation to them.

Table of Contents

Chapter 1: General Introduction

1.1 THEORETICAL FRAMEWORK

In this section I will introduce concepts in cognitive neuroscience that are of importance to the empirical studies I will present in chapters two to five. I will summarise theoretical approaches that try to explain reward processing in the brain and review supporting empirical evidence (section 1.1.1). Furthermore, I will outline how humans and animals react to visually surprising events and what the consequences of these events are (section 1.1.2).

1.1.1 Reward prediction errors and learning about rewards

What is reward or rewarding?

Most commonly reward is perceived as something that you receive when you have done something well or accomplished a certain task (instrumental conditioning). This becomes clear after typing the word 'reward' into Google (Figure 1.1). When clicking on 'more definitions' in this Google search however, one finds more conceptualisations of what can be rewarding. One of the other ways of thinking about something rewarding is when an event 'makes you feel good', i.e. the pleasantness of an event (hedonic function of reward).

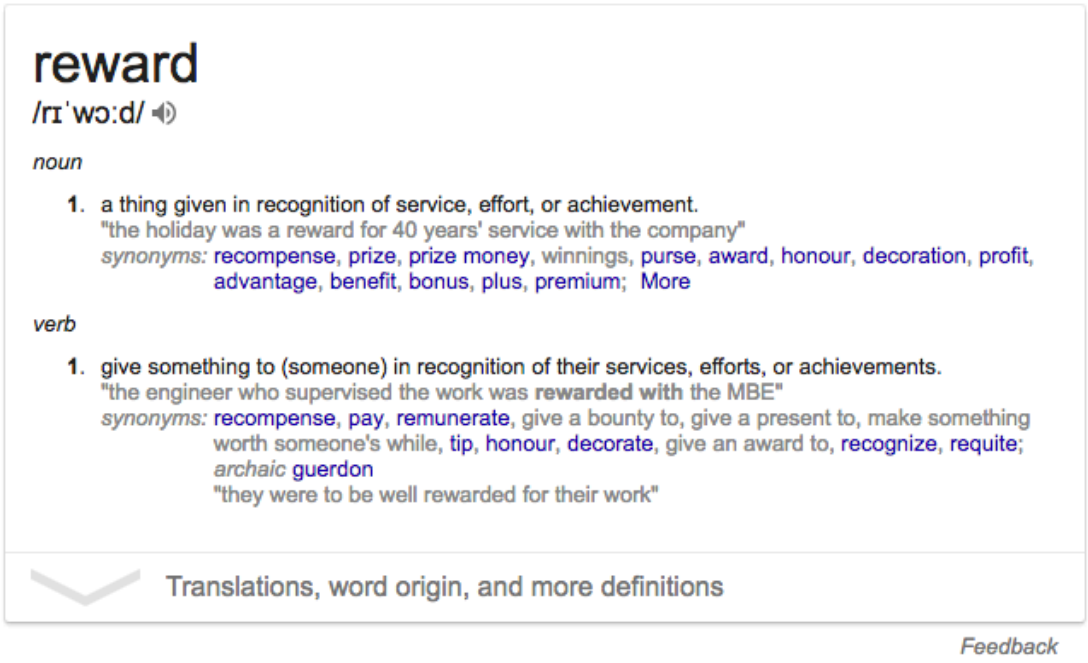

Figure 1.1 Screenshot of Google search for 'reward'

As Google is predominantly designed for humans and - generally speaking - not for looking up scientific definitions but rather the everyday usage of a word, it is not surprising that an important aspect of reward is not mentioned in Google's definition: the reinforcing effect of reward. Items of food or amounts of liquids can act as positive reinforcers for an action, meaning that when delivery of food or liquid follows a particular action then the animal becomes more likely to perform the action again.

Although there is no doubt about that preferred foods or liquids can be rewards, it remains debated which aspect of the items actually triggers a reward response in the brain. Shape or colour of the item might play a role, as might its texture when chewing or swallowing it. It is believed that the taste plays a crucial part in having a rewarding effect as sweeteners with no nutritional value like

saccharin increase behavioural responses (Schultz, 2006). But also the item's nutritious value and the way in which it influences the body's vegetative functions seem to be related to the way in which an item is experienced as rewarding. Animals tend to avoid foods with low nutrient content such as essential amino acids (Delaney & Gelperi, 1986; Hrupka et al. 1997; Rogers & Harper, 1970; Wang et al., 1996).

How predictions about reward are formed

One of the earliest reports of learning about the predictability of rewards are Pavlov's famous series of experiments (Pavlov, 1927). Through what was later termed 'classical conditioning', Pavlov used a neutral acoustic stimulus and paired its presentation with the delivery of food. The subject (in this case a dog) soon started to salivate on hearing the formerly neutral sound before the food was even presented. Intrigued, Pavlov invented the term 'psychic secretion' which referred to the salivation caused by conditioned (or conditional) reflexes (Buser, 2006). At around the same time as Pavlov conducted his experiments on dogs, Thorndike (Thorndike, 1911) tested cats in experiments in which they had to learn to interact with a box in which they were trapped. By pushing a certain lever, a flap opened and the cat could get out. This seems to have been rewarding for the cats as, over time, the animals found and pushed the lever more and more quickly. This observation lead to Thorndike formulating the psychological principle of the Law of Effect which states that "responses that produce a satisfying effect in a particular situation become more likely to occur again in that situation, and responses that produce a discomforting effect become less likely to occur again in that situation." The words 'satisfying' and 'dissatisfying' were later changed to reinforcing and punishing.

Schultz has summarized the Law of Effect as meaning that "rewards make you come back for more" (Schultz, 2006). Unlike Pavlov's dog that was completely inactive during the conditioning process, Thorndike's cat had to do something before getting the reward (i.e. freedom) and thus the term instrumental (or operant) conditioning was formed. Actions are repeated if they yield positive reinforcement or avoided if followed by negative reinforcements or punishments.

Over the last century classical and instrumental conditioning have been researched in great detail. More conditioning phenomena have been described such as 'extinction' where, after establishing a relationship by pairing conditioned stimulus (CS) and reward (unconditioned stimulus (US)), the relationship is subsequently broken by presenting the CS with no US. Thus the conditioning effect of the CS is extinguished (Pavlov, 1927).

Other aspects of conditioning include conditioned inhibition, where presentation of a stimulus in addition to the CS means that the CS-US relationship no longer holds. Conditioned inhibitors diminish animals' reactions to the CS. In another phenomenon, blocking, the fact that one CS-US relation has already been learned blocks the acquisition of a relationship between a second CS and the same US (Kamin, 1969). Based on these findings it has been concluded that in order for conditioning to occur, three factors need to be taken into account. First, a temporal component called contiguity. It is required that CS and reward occur in proximity in time for them to be paired, i.e. the US has to appear a few seconds after the CS and the US must not precede the CS (Figure 1.2a). The second factor is called contingency which states that the US must occur more often after a presentation of the CS than without the presentation of a CS (Figure 1.2b). The third factor is a

prediction error. Prediction errors are the difference between the expected outcome and the outcome that is actually received (Figure 1.2c). For example, when a food pellet is first delivered after the first occasion that a choice is made, then the food pellet is completely unexpected. The animal had no prior expectation of reward (a 'zero' expectation reward) but actually receives one pellet. The discrepancy, or prediction error, between the expectation of zero and the actual reward of one pellet is positive and corresponds to one food pellet. Gradually with more and more experience of receiving the reward after making the choice, the animal will have some degree of reward expectation. It might not fully expect the reward (it might not have the full and certain expectation of one pellet) but it might have some degree of reward expectation. The level of reward expectation will increase after each occasion a reward is delivered (Figure 1.2c) and the expectation will correspond increasingly closely to the reward amount actually delivered. If, however, the reward is not delivered after the predictive stimulus is seen or the predictive response is made then the actual outcome, zero reward, will be less than the animal's prior reward expectation and there will be a negative prediction error. On the trial after a negative as opposed to a positive prediction error, the animal's reward expectation will be reduced rather than increased.

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.2 Basic assumptions of animal learning theory defining the behavioral functions of rewards. (*a*) Contiguity refers to the temporal proximity of a conditioned stimulus (CS), or action, and the reward. (*b*) Contingency refers to the conditional probability of reward occurring in the presence of a conditioned stimulus as opposed to its absence (modified from Dickinson (1980)). (*c*) Prediction error denotes the discrepancy between an actual reward that has been received and its prediction. Learning (ΔV, associative strength) is proportional to the prediction error (λ–V) and reaches its asymptote when the prediction error approaches zero after several learning trials. All three requirements need to be fulfilled for learning to occur. US, unconditioned stimulus. (Schultz, 2006)

Violation of predictions: prediction errors and learning

As briefly mentioned in the section above, prediction errors are a necessary condition for learning to take place. If there is no prediction error then no learning will take place and no stimulus reward association will be formed. Rescorla and Wagner (1972) built a formal model of learning based on this assumption. They refer to learning as a change in associative strength (ΔV) between CS and US over time.

Learning is proportional to the prediction error, which means that it is highest whenever there is a big mismatch between prediction and actual outcome. Next trial's prediction will thus be adjusted which reduces the prediction error for the next trial's outcome. Adjusting the prediction over time diminishes the prediction error and therefore the change in associative strength. During learning the prediction error asymptotically approaches 0. In other words:

$$\Delta V = \alpha\beta\,(\lambda - V),$$

where ΔV is the change in associative strength and α a the learning rate which can be seen as a salience of the stimulus as different stimuli-reward associations are formed at different rates. β is the US association value and λ-V is the prediction error (λ is the prediction and V the reward outcome). If the reinforcement is fully predicted, associative strength remains unchanged. If, however, reward is omitted the prediction error is negative and associative strength is weakened. After each episode of learning (event *n*), the associative strength (V_n) is updated by adding ΔV_n:

$$V_{n+1} = V_n + \Delta V_n$$

Importantly the model explains a number of the learning phenomena mentioned above such as blocking. If an animal already knows that stimulus A predicts a particular reward then this blocks the animal's ability to learn that stimulus B also predicts the same reward when the animal is subsequently exposed to presentations of both A and B followed by the reward. According to the Rescorla-Wagner model this happens because there is no prediction error; the reward is already fully predicted by stimulus A and so there is no positive

reward prediction error if B is presented in conjunction with A. B therefore does not become associated with reward expectation even if A, B, and the reward are repeatedly presented contiguously and contingently.

Neural responses to reward learning and prediction errors

Neural responses to reward delivery can be found in various different brain areas. Studies have shown an increase in firing rate to liquid or food rewards in the cortex (orbitofrontal, prefrontal, premotor cortex) as well as in subcortical structures (striatum, amygdala, dopaminergic midbrain) (Amador et al., 2000; Apicella et al., 1991; Bowman et al., 1996; Hikosaka et al., 1989; Ljungberg et al., 1992; Markowitsch & Pritzel, 1976; Nakamura et al., 1992; Nishijo et al., 1988; Pratt & Mizumori, 1998; Ravel et al., 1999; Schultz & Tremblay, 1999; Shidara et al., 1998; Thorpe et al., 1983). As discussed above, in order for a stimulus-outcome association to be formed, the requirements of Contiguity, Contingency and a prediction error need to be fulfilled. In their seminal paper, Schultz et al. demonstrated that activity in dopaminergic neurons encodes a prediction error signal when an animal is learning a stimulus-outcome association (Figure 1.3, (Schultz et al., 1997)). Based on neurophysiological recordings in monkey midbrain, Schultz and colleagues formulated a model that uses a temporal difference (TD, (Sutton & Barto, 1998)) algorithm that accounts for the functional role of dopamine reward signalling in the brain. Dopamine therefore is no longer considered a 'reward signal' in the sense of a 'pleasure molecule' but instead it is seen as carrying a reinforcement signal. More recently, it has been shown that reward prediction errors can also be measured using fMRI in humans (e.g. Jocham et al., 2011; Kahnt et al., 2009; O'Doherty et al.,

2003; Pessiglione et al., 2006; Tobler et al., 2006). A recent meta-analysis including 35 human fMRI studies found overlapping activity patches predominantly in the basal ganglia (dorsal and ventral putamen but also in the nucleus accumbens). Furthermore, clusters were found in the anterior cingulate cortex as well as prefrontal areas (Garrison et al., 2013).

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.3 RPE signalling in dopaminergic neurons (Schultz 1997). Top, increase in firing rate after reward (R) is received. Middle, phasic discharging of dopamine cells when a predictive stimulus (CS) is presented (in this case a CS+, predicting reward delivery) but not when the actual reward (R) is delivered. Bottom, When CS+ is no longer followed by a reward, dopamine neurons are phasically silent at time when reward is omitted.

1.1.2 Spatial prediction errors and updating information in space

Animals (including humans) are agents in their environment. In order to successfully interact with or survive in their environment, they need spatial awareness and knowledge regarding the non-random processes surrounding them. The brain uses

past experience to form an internal model of the world. Such a model reflects statistics of the environment which are relevant for future behaviour. The model facilitates processing of incoming sensory information and it optimises the selection of appropriate actions. Furthermore, it makes predictions about events and outcomes of actions based on the environment's stochastic elements (Friston et al., 2009).

As these predictions are based on stochastic events, they themselves have certain probabilities which means that some events may be strongly predicted (have a high probability under the model) while others are surprising (have a low probability under the model). This surprise (or prediction error) has been intensely studied in the field of visual perception and attention (Posner et al., 1980). Posner et al. described behavioural effects of situations in which the model predictions were correct (validly cues targets) and reaction times were low and cases in which the model made an inaccurate prediction (invalidly cued targets). The latter was followed by a behavioural cost (i.e. longer reaction times). Posner et al. attributed the reaction time cost to the process of shifting attention in space. The process of shifting or reorienting attention in space can be roughly categorised into two classes: Stimulus-driven or goal-directed (Corbetta & Shulman, 2002). The first process describes the phenomenon of an attention-grabbing event, a sudden appearance of a stimulus that had a low probability under the internal model like a flashing light or the appearance of a target that had been invalidly cued. The second process refers to an attentional shift that is driven by a top-down mechanism. It's the shifting of attention based on model predictions in order to react optimally to an upcoming event. Corbetta and Shulman (2002) derived a model that links separate (but

interconnected) brain networks to the two types of attentional shifts. The two attention networks in the brain have also been thoroughly investigated (Astafiev et al., 2006; Corbetta et al., 2008; Patel et al., 2015; Shulman et al., 2003). In its essence, the model of the two systems describes a dorsal and a ventral frontoparietal network. While the dorsal network (consisting of areas in the intraparietal sulcus (IPs) and superior parietal lobe (SPL) and the frontal eye fields (FEF)) is more active during goal-directed shifts of attention, the ventral network (temporoparietal junction (TPJ) and ventral frontal cortex (VFC)) seems to be more involved when reacting to surprising stimuli and reorienting one's attention in a stimulus-driven manner.

In most of the studies mentioned above however, the focus lies on reorienting processes evoked by surprise including reallocation of resources to a previously deprioritized region of space and/or replanning a motor response to an unexpected stimulus. It is possible however, that the brain carries out another, distinct process between trials that is triggered by surprising events: updating the internal model to predict future observations more accurately. It is worth noting that the concepts of reacting to a surprise and updating the model are behaviourally dissociable. Surprise need not necessarily lead to updating, and updating can be triggered without surprise but in most studies both concepts have been strongly correlated. Updating one's spatial model is not dissimilar to learning a reward expectation as described in Section 1.1.1. A spatial prediction error represents an event with a low probability in the set or distribution of possible observations as predicted by the internal model (prior expectation). Subsequently learning or

updating occurs and probabilities of observations as defined by the prior expectation will be adjusted (for a computational definition, see Section 1.3.3).

1.2 ANATOMY AND FUNCTION

In this section I will discuss anatomy and functions of brain areas that I will focus on in the subsequent chapters. As the studies I will present in chapters two to five investigate the behavioural and neural effects of surprising rewarding and spatial events and the formation of new spatial and reward predictions in both the human and the macaque brain, I will introduce areas that have been described as important for these functions.

1.2.1 VTA and ventral Striatum

For more than half a century, scientists have studied brain structures that affect motivation. Before human neuroimaging was invented many electrophysiological and pharmacological studies with animal models suggested that there was a network of brain areas that underlie the processing of rewards (Olds & Milner, 1954; Phillips & Fibiger, 1978). There has been a particular focus on two interconnected brain regions that have been shown to be responsive to rewarding stimuli such as food treats and where electrical stimulation is itself apparently rewarding: the nucleus accumbens, a part of the ventral striatum (vSTR), and the dopaminergic midbrain, particularly the ventral tegmental area (VTA). Importantly, the VTA contains a great number of dopaminergic neurons which project to the vSTR among other areas (Schultz, 2002). This and the fact that neither of the two areas receive any direct

sensory input has inspired decades of research on dopamine (DA) and anatomical connections to and from the vSTR and VTA.

Dopaminergic neurons in the midbrain have been divided into categories based on their projections. There are three main projection pathways: mesolimbic (connects VTA to vSTR), mesocortical (connects VTA to prefrontal cortex areas), and nigrostriatal (connects the substantia nigra pars compacta (SNc), another midbrain region adjacent to the VTA, to the caudate nucleus and putamen). Thus, most of the midbrain's dopamine projections connect to the striatum and the cortex (Lynd-Balta & Haber, 1994; Williams & Goldman-Rakic, 1998). The organisation of projections from DA neurons in the midbrain to cortical areas is diffuse. Neurons connecting to functionally and anatomically different prefrontal cortical areas don't follow a clear topography but are intermingled. In contrast, mesolimbic and nigrostriatal neurons seem to have a topographic organisation with ventral and lateral cells projecting to the dorsal and lateral striatum whereas dorsal and medial neurons project to vSTR (Figure 1.4, (Haber & Behrens, 2014)). The vSTR occupies over 20% of the striatum in macaques and about the same in humans (Haber et al., 2006; Tziortzi et al., 2014) and has particularly prominent reward responses. It consists of the nucleus accumbens (NAcc) which in turn includes shell (NAccS)) and core subdivisions as well as the medial wall of the rostral caudate nucleus and the ventral and rostral putamen.

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.4 Schematic illustration of the complex connections between the striatum and substantia nigra pars compacta (SNc) in the midbrain. The arrows indicate how the ventral striatum can influence the dorsal striatum through the midbrain dopamine cells. Colors indicate functional regions of the striatum, based on cortical inputs. Midbrain projections from the shell target both the VTA and ventromedial SNc. Projections from the VTA to the shell (S) form a ''closed,'' reciprocal loop but also project more laterally to impact on dopamine cells projecting to the rest of the ventral striatum, forming the first part of a feedforward loop (or spiral). The spiral continues through the striato-nigro-striatal projections through which the ventral striatum impacts on cognitive and motor striatal areas via the midbrain dopamine cells. Red, inputs from the vmPFC; orange, inputs from the OFC and dACC; yellow, inputs from the dPFC; green and blue, inputs from motor control areas (Haber & Behrens, 2014).

In addition to their role in processing rewards, dopaminergic midbrain regions and other parts of the basal ganglia (BG) like putamen and caudate nucleus were historically considered relevant for motor functions due to their prominent role in Parkinson's or Huntington's disease. Today, we know that BG together with connecting cortical areas such as OFC and ACC play a part in a range of complex functions and goal-directed behaviours including emotions, motivation and cognition. A complex interplay between midbrain, ventral and dorsal striatum, and prefrontal cortex areas determines learning and adaptation. It relies on calculation of risk, value, and probability of events or options in the world and develops and executes an appropriate action plan.

1.2.2 vmPFC

The precise boundaries of the ventromedial prefrontal cortex are somewhat disputed but most definitions agree that it comprises medial orbitofrontal cortex (mOFC) areas such as Brodmann areas 14r, 14c, and 14m as well as ventral anterior cingulate areas such as 24, 25 and ventral 32 in humans (Figure 1.5A, (Brodmann, 1909; Rushworth et al., 2012; Rushworth et al., 2011)). Several of these areas are also found in macaque (Figure 1.5B, (Mackey & Petrides, 2010). VmPFC is medial to areas 11 and 13 in the central part of the orbitofrontal cortex (OFC), ventral to the dorsal anterior cingulate cortex (dACC) and caudal to area 10 of the frontal pole.

In its connectivity, there is a rough distinction between caudal vmPFC and rostral vmPFC. Caudal areas (14c, 25 and 32) have strong projections to the hypothalamus, amygdala and NAccS (Freedman et al., 2000; Ongür et al., 1998; Rempel-Clower & Barbas, 1998) and receive inputs from insula, hippocampus and hypothalamus. More rostral vmPFC areas, despite a connection with caudal vmPFC areas, have closer connections to areas 9, 10 and 45 (Freedman et al., 2000; Ongür & Price, 2000). In short, caudal vmPFC seems to be more strongly linked to visceral and emotional processing areas whereas rostral regions have a stronger connection to cognitive-related regions (Figure 1.5C).

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.5 Anatomy and connectivity of vmPFC. (A) Human vmPFC with areas 14r, 14m, 14c, 24, 25 and 32. (B) Macaque vmPFC with areas 14m, 24, 25 and 32 ((A) and (B) from Mackay and Petrides 2010). (C) Schematic illustration of vmPFC connections (from Haber and Behrens (2014)).

Functionally, vmPFC has been associated with different aspects of decision-making using BOLD-fMRI. In particular, it has been shown that the BOLD signal correlates with the reward value of a choice (Hare et al., 2008; Noonan et al., 2011; Plassmann et al., 2007; Plassmann et al., 2010). The vmPFC can encode the sum of all available options (Hare et al., 2011) or is sensitive to the difference between the option chosen and the unchosen option (Boorman et al., 2009). These different types of signals might reflect different stages of a decision-making process as populations of cells representing different values of options seem to compete with one another (Hunt et al., 2012) which is also supported by electrophysiological recordings in macaques (Strait et al., 2014). Nevertheless, recordings in macaques in vmPFC are rare and a substantial part of neuronal computations underlying decision-making processes remains unknown. One study showed that neurons in vmPFC (areas 14m and 25) are more sensitive to the internal values of options (for example hunger or thirst can increase the subjective value of a reward) (Bouret & Richmond, 2010) but less responsive to the external, sensory cues that predict reward. Furthermore, in

addition to value (Strait et al., 2014), vmPFC cells encoded an animal's motivational state and predicted the animal's willingness to work (San-Galli et al., 2016). Lesions of vmPFC neurons result in an impairment in distinguishing the values of options when those values are close together and also when a choice-irrelevant distractor is introduced (Noonan et al., 2010).

1.2.3 OFC

The OFC can be distinguished from vmPFC both in its connectivity and function. Anatomically, it includes Brodmann areas 11l, 12r, 12m, 12s, 13l, 13m, 13b and 13a in humans and macaques (Ongür & Price, 2000). It can be divided into central (cOFC, areas 11 and 13) and lateral (lOFC, area 12) OFC. Its connectivity varies between caudal and rostral subdivisions. Caudal areas have strong inputs from primary sensory areas whereas more rostral areas receive more highly processed sensory input. Furthermore, lOFC regions have tight links to ventrolateral prefrontal areas (vlPFC) and more medial OFC regions connect to vmPFC, area 10, and dorsal anterior cingulate cortex (dACC) (Barbas, 2007; Carmichael & Price, 1995a, 1995b).

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.6 Anatomy of OFC in (A) humans and (B) macaques. Dark blue areas form the lateral orbital network whereas light blue areas are related intermediate areas (from Stalnaker (2015)).

Functionally, OFC is thought to be involved in a variety of different cognitive tasks. Some of the earliest accounts of what OFC neurons do stem from the famous case study of Phineas Gage who suffered substantial PFC (including OFC) damage and as a result became socially and cognitively disinhibited (Harlow, 1868). Harlow's finding, led to interest in the general idea that OFC controls inhibition, in particular response inhibition (Ferrier, 1876; B. Jones & Mishkin, 1972; Kringelbach & Rolls, 2004). Other theories such as Damasio's 'somatic markers' hypothesis emphasized the processing of emotions (Bechara et al., 1997; Damasio, 1994). In parallel, OFC areas have been studied under the framework of value-guided decision-making and neuroeconomic cognition. While initially thought to be a flexible look-up table for stimulus-outcome (S-O) associations (Rolls et al., 1996) and heavily involved in tracking S-O associations

during reversal learning (Thorpe et al., 1983), it later was thought that OFC integrates all information about possible outcomes (e.g. probability, magnitude, costs, etc.) into a common neural currency that thus encodes the economic value of choice options. The existence of a common currency means that a choice can be made even between different types and modalities of goods (Levy & Glimcher, 2012; Montague & Berns, 2002; Padoa-Schioppa, 2011).

In addition to an outcome's value, OFC also represents its specific identity (Gottfried & Zelano, 2011; Klein-Flügge et al., 2013; Passingham & Wise, 2012; Schultz & Tremblay, 1999). The idea that OFC plays a crucial role in learning S-O associations was supported by the finding that OFC lesions disrupt rapid S-O learning in tasks such as stimulus-reward reversal learning (Gallagher et al., 1999; Izquierdo et al., 2004).

More recently, however, Rudebeck and colleagues (2013) have reported that performance of reversal learning tasks by macaques is not impaired when the lesions are made with an excitotoxic injection rather than by aspiration. Confusingly, however, excitotoxic lesions in OFC in other primate species, new world monkeys such as marmosets and rats do appear to disrupt reversal learning performance (Roberts, 2006; Stalnaker et al., 2015).

One possibility is that there is a specific region just outside the central OFC region most frequently investigated in macaques that is critical for discrimination reversal learning. This region may have been deafferented by the aspiration lesions that do impair discrimination reversal learning but it may not have been affected by the excitotoxic injections of the central OFC region that Rudebeck and colleagues (2013) made. A candidate lies in or just lateral to the posterior lateral orbitofrontal

cortex; Chau and colleagues (2015) have reported that this region has activity that is related to key aspects of the reversal task.

In addition Chau and colleagues (2015) point out that activity in this posterior lateral OFC (lOFC) area may be related to assigning credit for a specific outcome to a specific choice – a process necessary for learning which stimulus is currently rewarded and which is not. The idea is that it is not sufficient simply to detect the occurrence of a prediction error but that prediction errors must be accredited to specific past choices. This is particularly important when the recent choice history is variable; when animals have been shifting between making one choice or another over recent trials it is particularly important to be able to credit each specific instance of an outcome to a specific choice (Chau et al., 2015). Monkeys with lOFC lesions have shown impairments in credit assignment and incorrectly attributed rewards received on one trial to choices made in previous or even subsequent trials (Walton et al., 2010). Several human neuroimaging studies have drawn a similar conclusion; they have each found that a corresponding region (Neubert et al., 2015) in human lOFC is concerned with precise reward credit assignment (Akaishi et al., 2016; Boorman et al., 2016; Jocham et al., 2016; Noonan et al., 2011).

Although there may have been uncertainty about the role that OFC plays in S-O reversal learning, it does seem that there is agreement that it is important when stimulus values are inferred from knowledge or a model of task structure. In the reward devaluation task animals learn that some stimuli predict one reward, A, while other stimuli predict another reward, B. Animals are then given choices between stimuli predicting rewards A or B after a procedure has been carried out to devalue one reward or the other. For example, an animal might be left to feed to satiety on

reward A before being given a test in which it chooses between stimuli associated with reward A and stimuli associated with reward B. In such a scenario reward A is devalued. Normal control rats and monkeys integrate knowledge of the new value of reward A into the decisions that they take; they pick stimuli associated with reward B. They can do this even before they have experience of eating reward A and B in association with particular stimulus choices; rats will make more choices of the B-associated stimulus even if the test is conducted in extinction and monkeys will make more choices of the B-associated stimulus even if the test is conducted with many different stimuli each of which is only experienced once. After lesions are made in OFC (in the central OFC region between the lateral and medial orbitofrontal sulci in the case of monkeys), however, it is no longer possible for either rats or monkeys to rapidly integrate new value expectations into their decisions in the devaluation task (Jones et al., 2012; Murray et al., 2015; Rudebeck & Murray, 2014; Stalnaker et al., 2015).

Adaptive decision making in the devaluation task must depend on stored knowledge of specific S-O associations – associations between specific stimuli and specific types of reward. Schoenbaum and colleagues have argued that neural activity in rat OFC is best interpreted as reflecting such associations (Lopatina et al., 2015; McDannald et al., 2014) Multivariate decoding and repetition suppression-based analyses of human neuroimaging data suggest that such associations are also stored in human OFC (Howard et al., 2015; Klein-Flügge et al., 2013).

It has been argued that many theories of OFC function can explain certain aspects of phenomena observed in OFC but fail to do so with other findings (for a review, see Stalnaker (2015)). More recently, there has been the attempt of a unified

theory of OFC function. Wilson et al. hypothesise that OFC uses currently available information about the environment or task to calculate a current task state. Over time, OFC acquires a cognitive map of task space and thus provides state information which can then be used by other areas to adapt and react with an appropriate action plan (Wilson et al., 2014).

1.2.4 ACC

The anterior cingulate cortex (ACC) is located on the medial surface of the brain in the medial longitudinal fissure (also called interhemispheric fissure). It comprises the anterior parts of the cingulate gyrus and the cingulate sulcus and it extends caudally from the premotor cortex and curves rostrally along the shape of the corpus callosum. It mainly consists of dorsal area 32 and ventral areas 24 and 25 in humans and macaques. Both functionally and in terms of connectivity, it is becoming increasingly clear that the ACC can be split into dorsal ACC (dACC) and ventral and caudal ACC (often called perigenual ACC or pgACC) regions with different roles in decision making. In this introduction, I will mainly focus on anatomy and function of the dorsal part of the ACC.

The dACC is part of the agranular cortex which separates it from prefrontal areas like vmPFC (Wallis, 2011). Area 24 can be subdivided based on whether it is located in the cingulate sulcus or gyrus. 24c is mainly located in the sulcus and its caudal part is known as the rostral cingulate motor area (rCMA) whereas areas 24a/b are found in the gyrus. Area 24c has strong links to PFC including medial and lateral orbital cortices (areas 10,11,12,13,14, (Morecraft et al., 2012; Vogt & Pandya, 1987; Vogt et al., 1987)) but is also highly connected to adjacent subdivisions of the

anterior and posterior cingulate cortex. Furthermore, it connects to supplementary motor, premotor and motor cortices and there is an approximately somatotopically organised representation in parts of the cingulate in humans and monkeys (Amiez & Petrides, 2014; Procyk et al., 2014). Area 24a/b has connections to amygdala and vSTR.

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.7 Anatomy and connections of ACC in the (A) macaque and (B) human brain. (Rushworth et al., 2004). Dorsomedial cortex (areas 6 and 9) shown in green, ACC sulcus areas shown in red, ACC gyrus areas in light blue. (C) Connections from and to dACC in the human brain (from Haber and Behrens (2014)).

Dorsal ACC is one of the most debated and studied brain areas in cognitive neuroscience (Behrens et al., 2013). Like OFC, dACC is associated with a variety of cognitive functions. In contrast to OFC however, dACC lacks major sensory input but is rather strongly connected to motor areas; it even has direct connections to both primary motor cortex and to the ventral horn of the spinal cord (Dum & Strick, 1996; He et al., 1995; Picard & Strick, 1991). Thus, rather than detecting changes in the environment and making predictions based on sensory cues, its role in decision making may be more closely related to selection between different actions. For

example, dACC lesions only affect reversal learning if a reward is associated with different actions but not different sensory cues (Cai & Padoa-Schioppa, 2012; Rudebeck et al., 2008).

Traditionally, dACC was studied in the context of the Stroop task and error monitoring and it was hypothesised that it was mainly concerned with motor conflict processing (Cohen et al., 1990). Botvinick et al. went one step further and suggested that the signal in dACC is rather a more general response conflict signal and proposed a response conflict monitoring model with dACC in its core (Botvinick, 2008; Botvinick et al., 1999). This model comprises an explicit conflict monitoring unit which is active whenever competing response representations coexist. However, such a unit seems physiologically redundant and is at odds with other models of decision making (Gold & Shadlen, 2007; Wang, 2008). In these models, decision uncertainty is not represented explicitly but conflict is a mere by-product of competing populations of neurons that encode different action values. Moreover, this decision uncertainty can be associated with differences in duration of decisions (Hunt et al., 2012) or in differences in the subjective values of actions, i.e. decision confidence (De Martino et al., 2012). This interpretation is further supported by the fact that so far no single neuron recordings have provided evidence for a coding of response difficulty *per se* independent of coding a specific action (Cai & Padoa-Schioppa, 2012; Ebitz & Platt, 2015; Hayden et al., 2011; Nakamura et al., 2004).

The idea that dACC carries information about the valence of outcomes and subsequent action re-planning has been studied in depth since. It has been shown that ACC activity is correlated with error-monitoring and post-error slowing on the next trial (Ullsperger & von Cramon, 2001). This function of signalling behavioural

adaptation has in particular been studied in the context of reward which is motivated by the fact that dACC is located at the connectional intersection of reward and action selection networks. Many studies showed that ACC activity can be related to the monitoring of choice outcomes (both positive and negative), in particular when the information is subsequently used to update response strategies (Amiez et al., 2005; Behrens et al., 2007; Kennerley et al., 2006; Kennerley et al., 2011; Matsumoto et al., 2007; Nieuwenhuis et al., 2004; Quilodran et al., 2008; Walton et al., 2004). Interestingly, functional coupling of dACC depends on and changes strength to whichever motor cortex is involved in executing the response (Hare et al., 2011). DACC receives input from both the dopaminergic midbrain and vSTR and its activity encodes the value of feedback in a context dependent manner. For example, dACC responses to reward and error outcomes depend on the significance of those outcomes for changes in behaviour (Hayden et al., 2011) and on the type of choice that was made that led to the rewarding outcome (Kolling et al., 2014). In line with the idea that dACC plays a key role in updating an action plan, it has been shown that it is active in cases where humans and animals show a behaviour that explores possible action spaces (Kolling et al., 2012; Quilodran et al., 2008).

A key role of the dACC is in mediating the balance between exploration and exploitation (Kolling et al., 2016; Kolling et al., 2016). When an animal identifies a valuable choice it is likely to exploit it and repeatedly take it. If, however, the animal only ever focuses on exploiting known choices it may fail to identify even more profitable choices. It is therefore also adaptive for animals to explore the values of alternative choices.

Kolling and colleagues (2016) have argued that some of the ideas used to describe foraging behaviour in animals (Charnov, 1976; Stephens & Krebs, 1987) can be adapted to facilitate understanding of the exploitation/exploration balance. The idea is that an animal exploits, for example, a food source until the cost of obtaining a food item exceeds the reward being obtained. This is because many food sources become depleted with repeated sampling; for example a monkey harvesting fruit from a tree will eventually consume all the fruit in the tree and it will encounter fruit more frequently if it moves to a new tree. The animal should explore new opportunities if the cost of moving to a new food source is low and the environment is sufficiently rich that other opportunities are plentiful.

It has been shown that neurons in dACC ramp up their activity each time a monkey receives a reward while it exploits a depleting food patch. The ramping in activity increases until it reaches a threshold which determines an exploratative behaviour and the monkey moves to a new food patch (Hayden et al., 2011). Activity in human dACC encodes the value of searching amongst alternative options and, when the decision to explore is taken, it encodes the cost of exploring (Kolling et al., 2012). DACC carries different but related signals during decision making and during subsequent outcome monitoring. When making a choice, dACC activity encodes information about the values of available options and weighs them against the associated costs which leads to an assessment of how rich or valuable the environment is. During the outcome phase, activity reflects the value of a change in behaviour from exploitation to exploration.

1.2.5 FEF

The frontal eye field (FEF) has been extensively studied originally in macaques but later also in human PET and fMRI studies. Pioneering work was conducted as early as the 19th century when Ferrier demonstrated that electrical stimulation of certain parts of frontal cortex elicits eye movements (Ferrier, 1876). Since then, many electrophysiological studies have tried to decipher the activity of FEF neurons and its role in occulomotor control and attention. There are several different definitions of FEF in macaques that are based on either neurophysiological or anatomical criteria, but most definitions agree that the anterior bank (area 8Ar) and parts of the arcuate sulcus (area 8Ac and area 45B) constitute the FEF and that neurons with related activity patterns are found more rostrally in areas dorsal and ventral to the most caudal parts of the principle sulcus (area 46d and 46v) (Schall, 1997). Over the last century, studies using stimulation have revealed that FEF is topographically organised. The directions of evoked saccades can be mapped (Crosby et al., 1952; Mott & Schaefer, 1890) and saccade eccentricity seems to increase from ventral parts to dorsal parts of FEF (Bruce & Goldberg, 1985; Gottlieb et al., 1994; Suzuki & Azuma, 1983). Today's localisation of FEF was originally part of both the granular and agranular frontal cortex with rostral parts of A8 showing a granular layer IV which disappears when moving caudally midway into the depth of the arcuate sulcus. The FEF may even including a rostral part of area 6 (Stanton et al., 1989). Furthermore, FEF (area 8 and 45b but not 46v/d) has relatively large pyramidal cells in layers III and V. In contrast to the situation in macaques, FEF in humans is less well described and located. Most studies using PET, fMRI, or stimulation methods however agree on a location on the lateral-frontal convexity in the intersection of the inferior part of the superior precentral sulcus and the superior frontal sulcus (Amiez et al., 2006;

Fox et al., 1985; O'Driscoll et al., 1995; Paus et al., 1993; Paus, 1996). FEF is a highly interconnected area with strong input and output connections with both cortical and subcortical structures. Subcortically, it connects to the superior colliculus, the thalamus (mainly mediodorsal and ventroanterior nucleus), basal ganglia (striatum, subthalamic nucleus and caudate nucleus) and to a small degree to nuclei in the midbrain and pons (Leichnetz et al., 1984). Cortically, FEF, as a key area for sensorimotor integration, is connected to other frontal cortical areas including the supplementary eye field (SEF), areas 12 and 46, dACC area 24, and premotor areas on the postarcuate gyrus (Barbas & Mesulam, 1981; Huerta et al., 1986, 1987; Stanton et al., 1995). Furthermore, FEF is interconnected with almost the entire extrastriate visual cortex (Schall et al., 1995; Stanton et al., 1995) receiving convergent inputs from both the dorsal and the ventral visual stream (Schall, 1997).

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.8 (A) Anatomy of dlPFC of the macaque after Petrides and Pandya (1994). FEF consists of 46v/d, 8Av/d and parts of 45A/B. Abbreviations: AS=arcuate sulcus, CS=central sulcus,

PS=principle sulcus, STS=superior temporal sulcus. Functional connectivity of FEF in (B) human and (C) macaque brain as derived by resting state fMRI. From Hutchison (2012)

Functionally, FEF neurons are involved in nearly all types of eye movements. Neuronal discharge has been found in macaques during saccade planning and programming, saccade execution and saccade target selection. Nearly 50% of FEF neurons show a response at some stage during a saccade when the saccade is required for an operant reward task (e.g. Bruce et al., 1985; Schall et al., 1995). These signals can be phasic and represent the onset or offset of a visual stimulus or they can be longer lasting during a delay period. Furthermore, neurons are active during the production of a saccade (pro- or anti-saccade), pursuit eye movements or maintenance of fixation. Imaging studies in humans confirm these functions even though spatial and temporal resolution of imaging techniques (PET and fMRI) don't allow for such a fine-grained conclusion (Anderson et al., 1994; Paus, 1996). However, it has been shown that human FEF contains a topographic map of visual space (Kastner et al., 2007).

On a macroscopic level, FEF is a key node in the dorsal frontoparietal attention network (Corbetta & Shulman, 2002) and is thus thought to play a major role in top-down control of (visual) attention. So, in addition to planning and executing saccades, FEF is active when shifting one's attention in space. This can either be an overt (by moving one's eyes) or a covert (without eye movement) attentional shift. Corbetta and Shulman proposed a model in which there are two separate but interconnected attention networks in the brain, each involved in different aspects of attentional control. The ventral frontoparietal attentional network, consisting of parietal areas including the temporoparietal junction (TPJ) and ventrolateral PFC areas including the inferior frontal gyrus (IFg) and medial

frontal gyrus (MFg), is active when an attentional shift is stimulus driven, i.e. when attention is reoriented in space due to the infrequent or unexpected appearance of a stimulus outside the attentional focus. Note that the ventral attentional network is highly lateralised and comprises mainly the right TPJ and right ventral frontal cortex. This has also important clinical implications as lesions to parts of the right ventral attention system (but rarely the left) can cause the neurological syndrome of neglect in which patients show an impairment in detecting or responding to stimuli or information in the contralesional visual hemifield (Corbetta et al., 2002; Posner et al., 1984; Stone et al., 1992). The dorsal frontoparietal attention network includes areas in the intraparietal sulcus (IPs) and superior parietal lobule (SPL) and, as mentioned above, FEF in both hemispheres of the brain. Both IPs and FEF contain retinotopically organised maps of the contralateral visual space which underlines their role in maintaining spatial priority maps for covert spatial attention, saccade planning, and visual working memory (Jerde et al., 2012). Generally, the dorsal attention network is thought to control attention in a top-down or goal-oriented manner. As both systems are crucial for effectively controlling attention in space, the dorsal and ventral networks need to interact but until today, it is still debated how such interaction might be implemented in the brain (Vossel et al., 2014). The original idea that TPJ acts as a 'circuit-breaker' for the dorsal network (Corbetta & Shulman, 2002), i.e. redirecting attention to unexpected but behaviourally relevant stimuli, needs to be revised as it has been shown that latencies for visual responses are shorter in the dorsal network than in TPJ (Corbetta et al., 2008) which suggests that TPJ is involved in later evaluation of sensory events rather than sending an early reorienting signal to the dorsal network. Another candidate for orchestrating

interactions between both networks is the medial frontal gyrus (MFG) as it has been shown that resting state activity in MFG correlates with activity in both networks suggesting a functional coupling. The interaction patterns of both networks are highly flexible and depend on task demands. For example, activation differences between validly and invalidly cued targets in a Posner task can be seen in both frontoparietal networks (Corbetta & Shulman, 2011) whereas during visual search, only the dorsal network is involved and the ventral network seems to be actively suppressed (Shulman et al., 2003).

1.2.6 Parietal Cortex

The human parietal cortex extends caudally from the central sulcus to the occipitoparietal sulcus. Ventrally, its boundary is the Sylvian fissure. It can be (depending on nomenclature) separated into several subareas: the postcentral gyrus, superior parietal lobule (SPL), inferior parietal lobule (IPL) and precuneus. The postcentral gyrus and SPL are separated by the postcentral sulcus and the intraparietal sulcus (IPS) separates SPL and IPL. IPL can further be divided into supramarginal gyrus rostrally and angular gyrus caudally. The macaque parietal cortex consists of the somatosensory cortex in the posterior bank just caudal to the central sulcus (Brodman areas 1, 2 and 3), SPL (consisting of area PE/5) and IPL (consisting of areas PF/7b and PG/7a). Again an IPS separates the SPL from IPL and the ventral border of the parietal lobe is the Sylvian fissure while caudally, the parietal lobe is separated from the occipital lobe by the lunate sulcus.

Although the gross anatomy of human and macaque parietal cortex is similar, there are important differences in overall size and in the sizes of component areas;

the ratio of cortex occupied by the IPS is twice as large in humans (Orban et al., 2004). Moreover there are differences in morphology; the IPS has a complex folding and branching pattern in the human brain but a simpler form in macaques making it a challenge to identify human homologues of some macaque areas. There has therefore been considerable interest in understanding how areas in the macaque parietal cortex map to areas in the human brain. Here, I will mainly focus on areas in IPS and IPL, particularly those that play a role in attentional control and spatial representation, in particular the lateral intraparietal sulcus area (LIP) and IPL area 7a.

Macaque area LIP has been particularly well studied with both tracer injections and electrophysiological recordings. It has strong connections to the visual (V2, V3, V4, MT) and oculomotor system (FEF) and projects to the superior colliculus (Blatt et al., 1990; Lewis & Van Essen, 2000). The connectivity profile of area 7a is similar to that of LIP. However, 7a shows fewer and weaker connections to early visual areas and has stronger connections with prefrontal areas like area 45, cingulate and parahippocampal cortex (Carmen Cavada & Goldman-Rakic, 1989a, 1989b). Parcellation studies using a combination of diffusion tractography imaging (DTI) and resting state fMRI have tried to identify correspondences between monkey and human parietal areas (e.g. Mars et al., 2011; Neubert et al., 2015). Mars et al. found two areas that match the connectivity profile of LIP and area 7a and called these human areas IPS3 and PGp, respectively (Mars et al., 2011).

Both LIP and 7a encode spatial information and many neurons have visual receptive fields (RF). Furthermore, neurons in LIP and 7a form a retinotopic map of the contralateral visual hemifield. In contrast to many visual areas in the occipital lobe however, LIP and 7a neurons respond selectively to task-relevant or salient

stimuli that enter their RFs rather than firing as soon as stimuli enter the RF. This is implemented by a modulation of visually evoked activity through non-spatial (predominantly top-down) signals. For example, Oristaglio et al. (2006) recorded from LIP neurons while monkeys performed a covert visual search task that required them to release a bar with either the right or the left hand, depending on the cue. Most LIP neurons were active when the cue fell in their RF, however, in roughly 50% of the cells, the activity was modulated by the action the animal was asked to perform. Neurons encoded either a right hand response or a left hand response (Figure 1.9A). Importantly, this modulation disappeared when the animals' attention was directed outside the RF of the LIP neuron (Oristaglio et al., 2006). In another study from the same year, Freedman and Assad showed that not only limb movement but also stimulus categorisation modulates activity of LIP neurons (Freedman & Assad, 2006). But LIP activity can be modulated by other non-spatial signals too. Reward expectation and valence has proven to be an effective modulator of neuronal activity in LIP (Platt & Glimcher, 1999; Sugrue et al., 2004). Peck et al. showed that reward (CS+) increased the response to visual cues and that this increase was followed by a sustained activity in the delay period following the reward (Figure 1.9B). Non-reward (CS-) on the other hand elicited a lower initial response to the visual cue which was then followed by a sustained suppression of neural activity during the delay (Peck et al., 2009).

The image originally presented here cannot be made freely available via ORA because copyright. The image source can be found in the figure caption

Figure 1.9 (A) LIP neuronal activity is modulated by type of response. Left, increased firing rate for left hand response (blue) compared to right hand response (red). Right, no modulation of neural activity when distractor occurred in the RF and not the behaviourally relevant cue. Adapted from Oristaglio et al. (2006) (B) Reward valence modulates LIP activity. Left, CS+ triggers stronger initial response to visual stimulus (blue) compared to CS- (red) and evokes sustained activity during delay phase compared to suppression during delay for CS-. Right, LIP activity when CS is inside (black) or outside RF. Sustained signal increase after CS+ (top) and suppression after SC- (bottom). From Gottlieb (2010)

1.3 METHODOLOGY

In this section I will give a brief overview of scientific methods that I used for my studies. I used fMRI in macaques and humans in combination with analysing behaviour and computational modelling to investigate effects of spatial and reward surprise.

1.3.1 Brief History of fMRI in Humans and Macaques

FMRI is a popular method to study the *in vivo* brain in a non-invasive manner. It is sensitive to slight changes in the magnetic properties of haemoglobin that depend on its degree of oxygenation. Neuronal activation consumes energy and therefore requires oxygen. As a consequence, neural activation leads to localized changes in the balance of oxygenated and deoxygenated hemoglobin. FMRI can be used to detect changes in this balance. Hence fMRI is an indirect measure of brain activity; rather than measuring electrical activity in neurons, it measures the haemodynamic changes with neural activity (as explained in greater detail below). Within a few years of its inception, fMRI became one of the most frequently used techniques to study human brain functions as apparent from more than 19,000 publications using this technique up to 2008 (Logothetis, 2008).

While investigating the molecular structure of haemoglobin in 1936, Pauling and Coryell discovered that the molecule has different magnetic properties depending on whether or not it is bound to oxygen (Pauling & Coryell, 1936). When it is oxygenated (oHb), haemoglobin has no unpaired electrons and no magnetic moment which makes it diamagnetic, unlike deoxygenated haemoglobin (dHb) which is paramagnetic due to the presence of unpaired electrons and its magnetic moment. Because it is paramagnetic, dHb will distort the local magnetic field. As a consequence, adjacent protons will precess at different frequencies as the precessing frequency (Lamor frequency) depends on the magnetic field strength. This results in a more rapid decay of transverse magnetization, which is today commonly referred to as T_2^* decay. The implications of this discovery for MRI, that magnetic properties of haemoglobin change dependent on its oxygenation, was first

shown by Thulborn et al. (1982). They found that the decay of transverse magnetization is dependent on the relative amount of oHb within a sample of blood. This proved that MRI is sensitive to slight changes in blood oxygenation. Further, they showed that the relaxation difference in transverse magnetization between oxygenated and deoxygenated blood depends on the magnetic field strength (Thulborn et al., 1982).

Based on the findings made by Thulborn and colleagues, Ogawa devised a way of examining brain physiology using MRI. It was realized that for MRI to provide a useful physiological measure, it would need to be sensitive to some indirect measure of metabolism. One possibility was measuring blood flow, since metabolic processes require oxygen that is supplied through haemoglobin within red blood cells. In their seminal work, Ogawa et al. varied the amount of oxygen that anesthetized rodents breathed in a high-field (7.0 T) MRI. They found differences in the images between conditions in which rodents breathed pure oxygen, normal air (21% oxygen), and 0% oxygen. The less oxygen was provided to the animals, the more distortions were observable in the images. They attributed these distortions to the magnetic susceptibility effects of paramagnetic dHb in blood vessels (Ogawa et al., 1990). The first study using BOLD fMRI in the human brain was published two years later in 1992 by Kwong et al. They used periodic visual stimulation with flickering light and showed a signal intensity increase in occipital regions when visual stimulation was compared to non-stimulation (Kwong et al., 1992). Only months later Ogawa et al. replicated the findings by Kwong et al. They also investigated the impact of echo time (TE) on the measured signal intensity increase at 1.5 T and found a reduced stimulus-induced signal change at TE = 8ms compared to TE = 40ms

(Ogawa et al., 1992). Both studies described here made use of a blocked experimental design with long alternating periods of stimulation and darkness (70 - 150s). Blamire et al. were the first to show that a stimulation period of only two seconds is sufficient to observe a significant signal increase (Blamire et al., 1992).

Since then, the popularity of fMRI increased rapidly and it became the method of choice to investigate human brain functions. The advantage of this technique is that it is non-invasive and that it is associated with no known long-term tissue damage. In addition, compared to other non-invasive neuroimaging techniques, it has good spatial resolution. Unlike PET, fMRI does not require an injection of exogenous (radioactive) substances. These properties make it very suitable for the investigation of the human brain. But BOLD fMRI also has its limitations. The haemodynamic response is slow compared to neuronal activity which means that fMRI alone cannot be used for investigating the timescale of neuronal processes. However, it can be particularly powerful when it is complemented by other techniques such as electrophysiological recordings, where the focus is narrowed down to local brain processes. Since recordings with electrodes are problematic to perform on humans and only possible in some exceptional cases (a specific group of human patients), research on the brains of non-human primates remains essential for the investigation of many cognitive processes.

In 1998, Stefanacci et al. and Dubowitz et al. were the first to conduct fMRI studies with awake monkeys. Both studies found significant BOLD responses to visual stimulation in the occipital cortex while the monkey performed a passive viewing task (Dubowitz et al., 1998; Stefanacci et al., 1998). Several following studies

were concerned with the improvement of data quality such as increasing signal-to-noise ratio (SNR) by using exogenous contrast agents (Brewer et al., 2002; Leite et al., 2002; Vanduffel et al., 2001) or reducing motion induced artifacts (Keliris et al., 2007; Pfeuffer et al., 2007). Over the last decade, simultaneous use of fMRI and intracranial recordings (Logothetis et al., 2001) or stimulations (Moeller et al., 2008) became more and more established. In 2001 Logothetis et al. simultaneously recorded fMRI and electrophysiological data from the primary visual cortex of monkeys. Their aim was to elucidate the relation between BOLD signal and different types of electrophysiological signals (single unit activity (SUA), multi unit activity (MUA) and local field potential (LFP)). SUA and MUA occurred transiently at the onset of the stimulus and did not persist over time (returned to baseline several seconds after stimulus onset), while the LFP activity - which includes both post-synaptic potentials and integrative activity occurring at the soma - better predicted the BOLD signal change compared to SUA and MUA (Logothetis et al., 2001).

However, fMRI in awake macaques is still rare compared to fMRI in humans or electrophysiology in macaques. A simple search on pubmed (https://www.ncbi.nlm.nih.gov/pubmed) revealed fewer than 100 studies when querying for 'awake macaque fMRI'.

1.3.2 Why do fMRI in Macaques?

The reasons for the relative rarity of fMRI studies using awake macaques are manifold. One of the obvious reasons is a logistic consideration. An MRI scanner needs to be easily accessible from and close to a macaque colony for researchers to

use it efficiently and this infrastructure is still rare. Another contributing factor might be the costs associated with MRI scanning. MRI scanners are expensive and so is peripheral equipment (coils, primate chairs, etc.), that is needed for conducting the research. Furthermore, training of an animal to reliably perform a task while being scanned can take years (depending on the task). So it seems natural that the main use of fMRI in cognitive neuroscience is for scanning human subjects. Especially when considering the relative abundance of MRI scanners suitable for research at university research institutes or hospitals. But why use fMRI in macaques at all when one can do the same experiments in humans?

An argument for conducting fMRI studies in awake macaques is that fMRI in macaques can serve as a bridge between the two big data pools of human fMRI and macaque electrophysiology. By inter-species comparison of fMRI signal (e.g. Mars et al., 2011; Neubert et al., 2015) and relating it to recordings obtained in macaques, one might draw conclusions about cell behaviour in the human brain (Paradiso, 1999). Furthermore, fMRI enables collection of whole brain data images that reveal distributed activity patterns and functional networks in cognitive tasks, thus it can complement and guide electrophysiological recordings or lesion studies by providing maps of task active brain areas. As mentioned in Section 1.3.1, fMRI in macaques can also be used in combination with invasive methods. Logothetis et al. (2001) used concurrent fMRI and electrophysiology to investigate which aspect of neuronal activity is driving the BOLD signal and found that BOLD fluctuations correlate more with LFP than with single neurons firing (Logothetis et al., 2001). Furthermore, fMRI can be combined with lesion studies which enables pre- and post-lesion scanning, something that is difficult if not impossible in human subjects. Even if a human

subject's lesion were confined to a single functional brain area, it would be rarely the case that pre-lesion data was obtained from that subject. Thus, fMRI and lesion studies in macaques can provide crucial evidence for causal relations within a functional network. It can also show potential compensation mechanisms in the brain if nodes of a network are damaged.

Recently, advances in low-intensity focused ultrasound (FUS) technology based on animal research have provided evidence for its use as a non-invasive brain stimulation method in rabbits (Bystritsky et al., 2011; Yoo et al., 2011), rodents (King et al., 2013; Tufail et al., 2011; Younan et al., 2013) and macaques (Deffieux et al., 2013) that can be used concurrently with fMRI. FUS has several advantages over other stimulation techniques. It is non-invasive and thus ultimately applicable in human neuroscience. Furthermore, it can provide superior spatial resolution and focus compared to other non-invasive stimulation techniques like transcranial magnetic stimulation (TMS) (Bystritsky et al., 2011). Recently, FUS has been used in human studies (Lee et al., 2016a) even concurrently with fMRI (Lee et al., 2016b). This underlines the importance of pre-clinical studies in animal models using the help of fMRI for the development of new techniques in cognitive neuroscience.

1.3.3 Brief Introduction to Information Theory and the Brain

Information theory is often considered the corner stone for modern computer technology (it is no coincidence that computer technology is usually referred to as information technology or IT). One of the most influential approaches was proposed by Shannon (Shannon, 1948). In his seminal work 'A Mathematical Theory of

Communication' he provided the foundation for modern communication and information technology by defining a quantifiable unit of information: Bits. In communication theory, a transmitter sends a signal obtained from an information source through a channel to a receiver. He uses the concept of Entropy to describe the predictability of the signal. The more predictable the signal is overall, the lower the (Shannon) Entropy H_S. Each item (e.g. alphabetical letter in a text) of the signal carries a distinct amount of information (I_S). Shannon defined the amount of bits per item α as

$$\mathrm{I_S}(\alpha) = -\log \mathrm{p}(\alpha), \tag{1}$$

where p(α) is the probability that α is observed. As H_S is the sum of all items, it follows

$$\mathrm{H}_S(\alpha) = \sum \mathrm{p}(\alpha)\mathrm{I_S}(\alpha) = \sum -\mathrm{p}(\alpha)\log \mathrm{p}(\alpha), \tag{2}$$

From eq. 1 it becomes clear that the lower the probability of event or item α is (i.e. the more surprising it is), the higher is its I_S. If p(α) = 1, its I_S therefore is zero, i.e. 0 bits.

Biological systems (including brains) are designed to maintain their state and form in a constantly changing environment (e.g. Ashby, 1947; Haken, 1978). So biologically speaking a system regulates its internal environment to maintain its state despite influences from the external milieu - a process that is referred to as homeostasis (Bernard, 1974). Mathematically, it means for the brain that the probability for internal states as well as for external sensory input needs to have low entropy. So there is a high probability that the brain will be in a small number of

states and a low probability that it will be in the remaining states (Friston, 2010). As mentioned above, the brain has to achieve this in a constantly changing environment in which there is permanent input of sensory information. It has been proposed that the brain uses stochastic information (or Bayesian probability information) about the environment to create a model of the world (Knill & Pouget, 2004). In order to keep entropy low, the model should be designed so that the input's I_S is low. It does so by having a set of expectations or, in Bayesian terms, a prior: a distribution of probabilities that encodes beliefs about probabilities of possible future sensory inputs or observations. If subsequently an observation is made that does not match the probabilities as defined in the prior (hence the observation has a high I_S), probabilities in the prior are updated using Bayes' rule:

$$\mathrm{p}(\mathrm{post}|\mathrm{prior}) = \frac{\mathrm{p}(\mathrm{prior}|\mathrm{post})}{\mathrm{p}(\mathrm{prior})}\,\mathrm{p}(\mathrm{post}). \tag{3}$$

Since posterior and prior are both distributions of probabilities, we can calculate the difference between the two using the Kullback-Leibler divergence $\mathrm{D_{KL}}$ (Kullback & Leibler, 1951):

$$\mathrm{D_{KL}}(\mathrm{post}||\mathrm{prior}) = \sum_{\alpha} \mathrm{p}(\alpha|\mathrm{prior})\,[\log \mathrm{p}(\alpha|\mathrm{prior}) - \log \mathrm{p}(\alpha|\mathrm{post})] \tag{4}$$

where $\mathrm{p}(\alpha|\mathrm{prior})$ is the probability that the observation α would be made, given the model just before α was observed, and $\mathrm{p}(\alpha|post)$ is the same quantity, given the updated model just after α was observed.

Animals (including humans) are agents in their environment. In order to successfully interact with or survive in their environment, they need spatial awareness and knowledge regarding non-random processes and action outcomes.

The brain uses past experience about these processes and outcomes to form an internal model of the world. Such a model reflects statistics of the environment which are relevant for future behaviour. In what follows, I will describe two studies that investigate how humans and non-human primates form expectations about spatial events. Furthermore, I will show behavioural and neural effects using fMRI in non-human primates while the animals are receiving a certain amount of reward and how that might affect forming a subsequent reward expectation.

In chapters 2 and 3, I will show behavioural and neural results of a study using human volunteers. The aim of the study is to disentangle the operations a spatially surprising event might trigger: an attentional reorienting process and a subsequent learning or updating of the internal model that takes the new and surprising event into account. In past studies, these two operations have often been strongly correlated and it was therefore not possible to computationally disentangle them. In our study, we explore the possibility that there are indeed two separate and distinguishable operations that have differentiable effects on both behaviour and brain activity. With the use of computational models based on Bayesian statistics, we are able to model surprise independently from an updating of the internal model. This will enable us to reveal distinct brain areas and brain networks that are involved in processing either surprise or update and where these brain networks interact and overlap.

In chapters 4 and 5, I will present data that has been collected with fMRI in non-human primates while the animals were performing a cognitive task. The studies aim is to show different brain networks that are involved in processing

different types of prediction errors. These prediction errors can either be based on spatial expectations and reflect an event or stimulus appearing at an unexpected location in space or they can be based on reward expectations which in turn are influenced by past reward events. By using a carefully designed task that manipulates spatial and reward events in a decorrelated manner, we are able to look at prediction errors in space and the reward domain independently. Furthermore, by using fMRI, we will be able to reveal whole brain networks that might be involved in processing unexpected events and forming new predictions in non-human primates – something that previously has only been studied in a very narrow and focused manner using electrophysiological methods.

Chapter 2: Behavioural Effects of Spatial Reorienting vs. Updating

In the following two chapters, I will present and discuss data that concern the manner in which humans change and adapt their behaviour in response to surprising events (O'Reilly, Schüffelgen et al., 2013). In chapter 2 I will present behavioural effects that relate to spatial surprise and the updating of spatial expectation. Furthermore, I will describe changes in pupil size that we measured with pupillometry when participants experienced spatially surprising events or updated their future spatial expectations. I discuss these changes in the framework of the brain's noradrenaline system. Chapter 3 will report the neural findings made in the same study.

2.1 INTRODUCTION

In a nonrandom environment, brains can and should make use of past experience to facilitate the processing of incoming sensory information and the selection of actions, through prediction (Friston, 2010; Friston et al., 2009). An important aspect of brain function is therefore the construction and tuning of internal models to represent statistics of the environment that are relevant for future behavior.

The use of predictive internal models implies that not only are some events well predicted (high probability under the model) but, conversely, some events (which have a low probability under the model) are surprising (Posner et al., 1980). Surprising events may be associated with behavioral costs; for example, although valid attentional cues speed reaction times (RTs), invalid cues lengthen them (Posner

et al., 1980; Posner et al., 1984). However, surprising events can have a further significance for the observer in that they sometimes provide evidence for a change in the environment, which would imply a need to update the brain's internal models to predict future events accurately.

Here, we explore the possibility that the brain carries out at least two distinct operations when a surprising event occurs: (i) within trial reorienting processes evoked by surprise, including reallocation of resources to a previously deprioritised region of space and/or replanning a motor response to an unexpected stimulus, and (ii) between-trial processes, particularly the possible need to update the internal model to predict future observations accurately in a changeable environment.

The within- and between-trial processes could be broadly characterised in terms of surprise (or rather the reprogramming/reorienting response caused by the surprising stimulus) and updating (of the entire model), respectively. Consider, for example, the classic Posner orienting task (Posner et al., 1980), in which the locations of visual events are predicted either explicitly by symbolic cues or implicitly by the fact that targets appear more frequently in certain locations. Invalidly cued (surprising) targets evoke behavioral reorienting (redirecting of attention or gaze to the surprising target location), but they may also cause the participant to update his/her beliefs about the probable locations of future targets.

Information theory gives distinct definitions of surprise and updating that formalise the distinction between the surprise (and consequent behavioral reorienting) evoked by a particular stimulus and updating of the overall model of the environment. In information theory, the surprise associated with a particular stimulus value, α, is characterised by its Shannon information [$\mathrm{I_S}(\alpha)$]:

$$\mathrm{I_S}(\alpha) = -\log \mathrm{p}(\alpha|\mathrm{prior}), \tag{5}$$

where $\mathrm{p}(\alpha|\mathrm{prior})$ is the prior probability that the observation α would be made, given the brain's internal model just before the data point was observed. Therefore, the $\mathrm{I_S}$ captures how unexpected or unlikely a particular observation is, given the internal model. In contrast, updating of the internal model is captured by the Kullback–Leibler divergence ($\mathrm{D_{KL}}$) between the posterior and the prior:

$$\mathrm{D_{KL}}(\mathrm{post}||\mathrm{prior}) = \sum_{\alpha} \mathrm{p}(\alpha|\mathrm{prior})\,[\log \mathrm{p}(\alpha|\mathrm{prior}) - \log \mathrm{p}(\alpha|\mathrm{post})] \tag{6}$$

where $\mathrm{p}(\alpha|\mathrm{prior})$ is the probability that the observation α would be made, given the model just before α was observed, and $\mathrm{p}(\alpha|post)$ is the same quantity, given the updated model just after α was observed.

As evident in Eq. 6, the $\mathrm{D_{KL}}$ is the probability-weighted average change in the $\mathrm{I_S}$ across all possible stimuli as a consequence of updating the model. Hence, although the $\mathrm{I_S}$ describes the degree of surprise evoked by observing a particular data point α, the $\mathrm{D_{KL}}$ describes how the model, as a whole, is updated as a consequence of observing α.

Although surprise and updating have distinct computational definitions, they are usually strongly correlated in experimental paradigms. As already noted, invalid targets in a Posner paradigm could evoke just surprise or both surprise and updating, depending on the task design. More theoretically, in temporal difference learning (Sutton & Barto, 1998), the updating of action or stimulus values is driven by prediction error (surprise). Furthermore, the terminology used is confusing because the $\mathrm{D_{KL}}$ (updating of the model) has been described as "Bayesian" surprise (Baldi & Itti, 2010). Note, however, that the constructs are behaviorally dissociable: Surprise

need not necessarily lead to updating, and updating can be triggered without surprise. The relationship between surprise and updating depends, among other things, on the learning rate (Rescorla & Wagner, 1972), the degree of expected stochasticity in the environment (Courville et al. 2006; Knight, 1921), and the expected frequency or rate of change in the underlying environment (Behrens et al. 2007). Furthermore, there exist scenarios in which we might hypothesise that updating should be triggered in the absence of surprising observations; for example, if the observer moves into a new context in which he knows a priori that his old internal models are unlikely to be valid (Dayan, 2012).

In the present study, we wished to investigate the physiological and neural mechanisms by which expectations (internal models) are updated, as distinct from the mechanisms by which the immediate behavioral response is reprogrammed to a surprising stimulus. We used saccadic planning as a context in which to investigate this problem. The choice of a saccadic planning task was motivated as follows. First, spatial attention and saccadic planning are clear examples of predictive models, in which the violation of predictions (as in invalid trials on a Posner task) evokes a process of behavioral reorienting that is measurable in terms of RT costs (Posner et al., 1980). Second, the neural nature of predictive models over saccadic target locations (the representations to be updated) is relatively well understood. Networks of spatially tuned cells in macaque parietal cortex represent maps over space of variables relevant to saccadic planning, such as the probability of targets appearing or the reward value associated with making a saccade to different points in space (Colby et al., 1996; Gold & Shadlen, 2000; Platt & Glimcher, 1999; Yang & Shadlen, 2007). Third, the contrast between saccadic reprogramming (as evoked by surprise)

and updating of internal models has intriguing parallels in the neurological literature. Although it is generally recognised that the posterior parietal cortex plays a key role in control of eye movements and spatial attention, it has long been noted that there is a fundamental difference between the neurological syndromes that follow damage to the intraparietal sulcus (IPS), including the lateral intraparietal area (LIP), and to the inferior parietal lobule (IPL) (Milner, 1996; Milner & Goodale, 1995). On the one hand, impairments in the ability to reorient or redirect the eyes (Pierrot-Deseilligny & Muri, 1996) in Balint syndrome or the covert focus of attention (Posner et al., 1984) in the extinction syndrome, for example, are associated with damage to the IPS that includes the LIP. By contrast, the quite distinct syndrome of neglect is associated with damage to the IPL (Milner, 1996; Mort et al., 2003). In the neglect syndrome, patients' problems cannot be described as simply the inability to reorient their attention to the neglected visual field when two stimuli are present as is the case in extinction. Instead, patients with the neglect syndrome are unable to acquire, despite experience and instruction, an expectation of any event of significance in the neglected field.

We developed a paradigm in which participants made speeded saccades to visual targets, the locations of which could be predicted by estimating their underlying spatial distribution. To dissociate the neural processes associated with surprise and updating in the experimental paradigm, we used a simple manipulation of the relevance of surprising targets in terms of the extent to which they predicted the locations of future targets, so that both surprise (probability of observations under the internal model) and updating (change in the internal model) varied independently across trials. Using functional MRI (fMRI) and behavioral measures,

including pupillometry, we observed both neural and behavioral dissociations between surprise and the updating of an internal model.

2.2 METHODS

Participants were 17 healthy volunteers (age range 21–32 years, 7 females). All participants gave written informed consent in accordance with the National Health Service Office for Research Ethics Committees (ref 07/ Q1603/11). The participants completed two sessions of the same task: a behavioral session in which high-quality eye tracking data were acquired and a session with fMRI and concurrent eye tracking (See Chapter 3). Before each session, participants were given explicit instructions about the meaning of different target colours and the distributions from which target locations were drawn. We sought to distinguish the neural response to surprise per se from neural activity occurring when an internal model was to be updated. To do this, we designed a saccade task in which the surprise associated with target locations (I_S) and the updating of the internal model (D_{KL}) were dissociated by a relevance manipulation.

The task was a simple saccadic eye movement response task in which participants could use prior knowledge about the spatial distribution of saccadic targets to facilitate speeded eye movement responses to them. On each trial, participants began by fixating a central cross. First, a warning signal appeared (the cross brightened for 400ms); immediately afterward, a coloured dot, the saccadic target, appeared for 350ms. Participants were instructed to move their eyes as quickly as possible to look at the dot and then return fixation to the central cross.

We measured saccadic RT using an infrared eye tracker [EyeLink 2000 (SR Research, Ottawa, Ontario, Ca) sampling at 500 Hz].

Participants could anticipate the location of the target dot because the dots appeared in similar locations over a block of several trials (mean block length of 15 trials, range of 10–20 trials). The distance of the dot from the central fixation point was fixed (so all targets appeared on a circular perimeter); hence, the position of the dot was governed by a single parameter, angle from vertical. This angle, α, followed a circular Gaussian distribution with a mean and variance that remained fixed during each block but moved abruptly to new values between blocks. Blocks were not temporally separated (i.e., the first trial of block *n* followed directly from the last trial of block *n-1*), but the start of each new block was explicitly signaled by a change in dot colour; thus, all the targets in block *n* may have been red and all the targets in block *n+1* may have been blue, etc.

Although most trials (75%) were drawn from Gaussian distributions as described above, we included another set of trials, called “one-offs” (25% of trials), that had locations randomly selected from a uniform distribution over the whole circle and were randomly interspersed with the other trial types. Hence, the generative probability density function over dot locations was the sum of a normalised Gaussian scaled by 75% and a normalised uniform over the whole circle scaled by 25%:

$$\mathrm{p}(\alpha) = 0.75\ \mathrm{p}\left(\alpha \middle| \alpha \sim N(\mu, \sigma^2)\right) + 0.25\ \mathrm{p}\left(\alpha | \alpha \sim U(0°, 360°)\right) \qquad (7)$$

Importantly, participants received explicit signals informing them when the current trial was the start of a new block or a one-off trial. Most target dots were

chromatically coloured and the beginning of a new block was signalled explicitly by a change in the colour of the target dots (so all the dots in one block would be, say, red and all the dots in the next block would be, say, blue). In contrast, the dots on one-off trials were always coloured grey. Therefore, although both one-off and update trials shared the feature of the target dot appearing in an unexpected location, the two types could be easily distinguished; the target dot was grey on update trials but chromatically coloured on trials drawn from the Gaussian distribution and took a new chromatic colour on update trials. The task is illustrated in Figure 2.1.

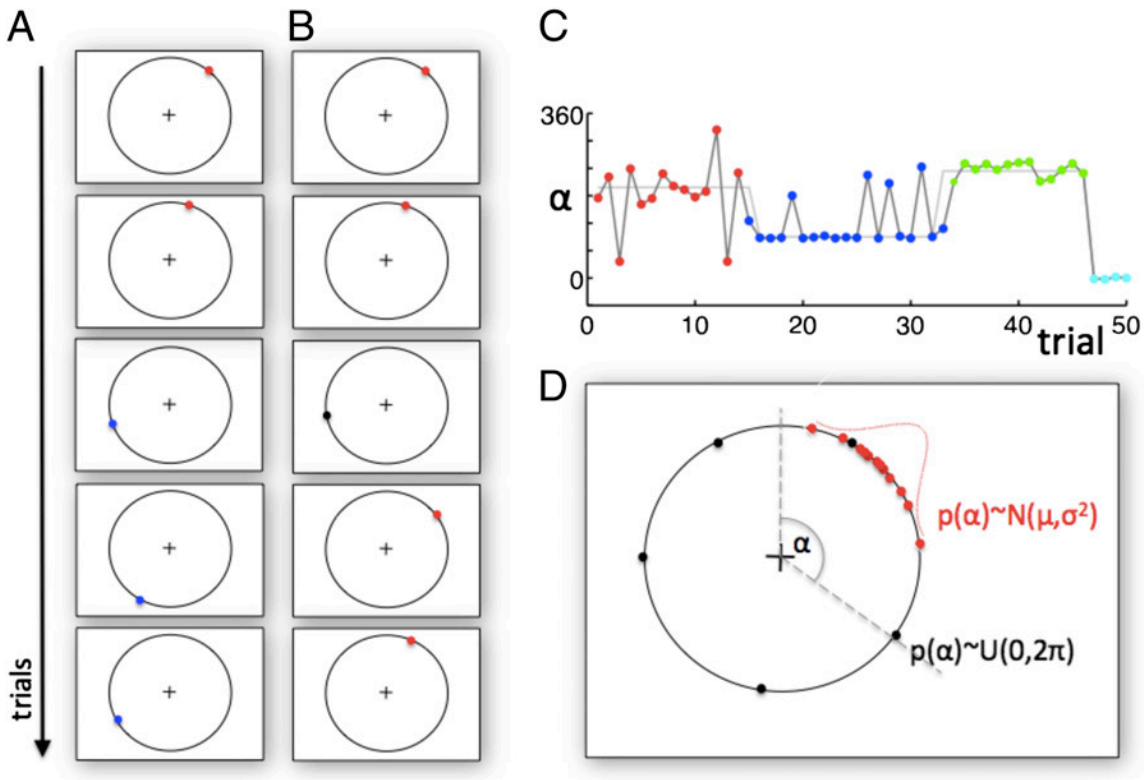

Figure 2.1 On each trial of the task, participants began by fixating a central cross. A target (coloured dot) appeared on a circular perimeter. Its location was predictable because target locations were similar over block s of 10–20 trials (mean of 15 trials). Two types of unexpected target locations could be observed, as illustrated in A and B. (A) Transition between block s

(update trial). Initially, red dots are observed in the upper right; subsequently, blue dots are observed in the lower left. (B) One-off trial. A target appears in an unexpected location (targets are expected in the upper right, but the one-off target appears in the lower left). However, this is a one-off trial, and future targets revert to the original distribution in the upper right. Targets on one-off trials are coloured grey. (C) Plot of target locations (angle α from vertical) over 150 trials. Different coloured targets are from different block s. One-off targets are shown in grey. (D) Distribution of target locations within a block is a combination of a circular Gaussian, shown in red, and a uniform distribution, shown in black, from which one-off trials are drawn.

2.2.1 Saccadic Reaction Times

We used the rate of gaze displacement to parse eye data into fixations and saccades. We then defined our response measures as follows. Onset time was defined as the end time of the fixation immediately preceding the saccade in which gaze first moved outside a 50-pixel window around the fixation cross (target locations were 300 pixels from the fixation cross). Arrival time was defined as the time of onset of the first fixation within a 50-pixel radius of the target location. Dwell time was defined as the duration of the first fixation within a 50-pixel radius of the target. To remove outliers, onset and arrival times less than 0ms from the target onset and more than 800ms (double the average arrival time) were disregarded; values more than 3 SDs from the mean for each measure were also discarded. We use the statistical measures of t-tests despite the potential issue of multiple comparison problems. We have a strong hypothesis about effects on specific trials and can therefore make a planned comparison rather than having to perform a variance analysis that would check for all interactions.

2.2.2 Pupillometry

For the general linear model (GLM) analyses of pupil size, pupil data were divided into epochs beginning 600ms before the onset of the warning cue and ending

2,000ms after the target onset. Each epoch was baseline-corrected by subtracting the mean pupil area in the first 200ms of the epoch. We used a multiple regression on the pupil size at each time point to determine the effects of four regressors: the main effect (which captured target-related changes in pupil diameter over time across all trial types), the entropy of the model, surprise [Shannon information (I_S)], and updating [Kullback–Leibler divergence (D_{KL})]. We plotted the mean effect size (beta value) across subjects ± SEM. Effects were considered significant for time points when the Z-score of the mean (mean/SEM) was greater than ±2.3, corresponding to a P value of 0.01 uncorrected. For the raw pupil analysis presented in Figure 2.7, each epoch was baseline-corrected by dividing by the mean pupil area in the first 200ms of the epoch.

2.2.3 Additional Pupillometry Experiment.

To replicate the unexpected finding that updating negatively modulates pupil diameter, we conducted an additional behavioral experiment with 18 participants to check potential confounding factors that could have driven the pupillometric effect. The additional experiment was similar in design to the main experiment, except that instead of making a saccade to a coloured dot on a circular perimeter as in the main experiment, participants maintained central fixation while a Gabor grating stimulus was presented on the circular perimeter. The task was to discriminate the grating's orientation. Trial types (expected, one-off, and update) were indicated by the colour of a warning stimulus presented 500ms before the Gabor patches.

2.3 RESULTS

2.3.1 Behavioural Confirmation of Task Strategy

The logic of including one-off trials was to dissociate the effects of surprise and behavioural reprogramming from the process of updating an internal model. On both update and one-off trials, we expected participants to be surprised by the target location and to engage motor reprogramming to cancel the planned saccade and plan a saccade to the new target location. However, updating should only occur on the update trials, because participants were explicitly informed that one-off trials had no predictive value relating to future trials.

Participants learn on update trials but not on one-off trials.

RT data confirmed that participants did learn on update trials (Figure 2.2A). Although participants' gaze took longer to reach the target on the first trial of a new block (the update trial) compared with "expected" (non-one-off, non-update) trials (t_{16} = 4.7, P = 0.0001, paired samples t test), there was no further behavioural cost on subsequent trials; on the second and subsequent trials of a new block, there was no significant difference in time for the gaze to reach the target compared with the average of expected trials (t_{16} = 0.48, P = 0.32). However, RTs were significantly (t_{16} = 3.0, P = 0.0045) faster than average on the last trial of the block (trial -1 in Figure 2.2A), perhaps because of a gradual increase in response speed across the block. Note that the notion 'behavioural cost' is used throughout this thesis to describe a mere reaction time cost, i.e. a relative slowing of a response.

As well as confirming that participants learned on update trials, it was important to confirm that they did not learn on one-off trials. One possible problem with our design would be if participants, even involuntarily, used the location of the

target on one-off trials to update their internal model of the underlying distribution for non–one-off trials (because we would then expect update-related brain activity on one-off trials, potentially leading to a false-negative result). If this were the case, we would expect to see slower RTs on trials that immediately followed one-off trials. This is because if the internal model had been updated on a one-off trial, participants should now expect subsequent trials to appear near the one-off target. However, we did not observe slowing on responses following a one-off trial; although participants' gaze took longer to reach the target on one- off trials than on expected trials (t_{16} = 2.0, P = 0.03, paired samples t test), RTs on trials following a one-off were very similar to the average expected trial (mean time to reach target for a trial immediately following a one-off was 408ms, and mean time for all expected trials was 402ms; t_{16} = 0.48, P = 0.32 for a paired samples t test between these values across participants). This suggests that participants successfully refrained from updating their internal model for chromatic dots based on the one-off trials. Equivalently, one could say that there was no evidence that the learning rate on one-off trials was above zero.

RTs following update- and one-off trials are illustrated in Figure 2.2A, which shows RT (time to fixate the target) by trial number for one-off and update trials, where trials are sorted so that trial 0 is the one-off or update trial and trial +1 is the subsequent trial, followed by trials +2, +3, etc. Trial -1 is the trial before the update or one-off trial. To reiterate, although the plots for one-off and update trials appear similar, they logically indicate opposite effects. To predict the position of the dot on trial +1 correctly, participants would need to update their internal model on update trials but not on one-off trials. Analysis of RTs on the update and one-off trials

themselves indicated no overall difference in time to reach the target between update and one-off trials (t_{16} = 1.22, P = 0.12, paired samples t test).

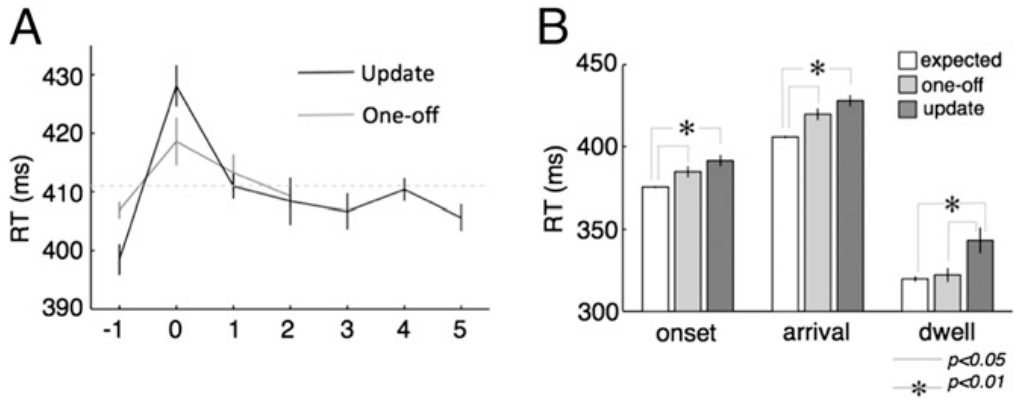

Figure 2.2 RTs indicate that participants performed the task as instructed. (A) RTs (time for gaze to arrive at the target) on update and one-off trials, as well as on surrounding trials. The dashed line is the mean RT for all trials. (B) Breakdown of saccadic RTs into onset time (time to leave central fixation), arrival at target (onset of fixation at the target), and dwell time (duration of fixation at the target).

Saccadic RTs reflect surprising target locations, but dwell time reflects updating

We conducted a more detailed analysis of saccadic responses in terms of RTs for saccadic onset; arrival at the target (as reported above); and dwell time, the time spent looking at the target. These comparisons are illustrated in Figure 2.1B. In line with the hypothesis that within-trial response production is slowed by unexpected target locations, saccade onset and arrival times were slower on both of the trial types containing surprising target locations (update and one-off trials), with both update and one-off trials having slower RTs than expected trials (paired sample t test for saccade onset time: t_{16} = 4.3, P = 0.0003 and t_{16} = 2.24, P = 0.02 for update and one-off trials vs. expected trials, respectively; no difference between up- date and expected trials: t_{16} = 0.86, P = 0.20).

In contrast, dwell time was specifically affected by updating, with update trials differing significantly from the other trial types: Dwell time was longer on update trials than on one-off trials (t_{16} = 2.02, P = 0.03). Dwell time was also longer on update trials compared with expected trials (t_{16} = 2.6, P = 0.01), and there was no significant difference in dwell time between one-off and expected trials (t_{16} = 0.26, P = 0.6).

RTs were driven by spatial, not feature-based, surprise

It could be argued that RT costs were partly driven by surprise associated with nonspatial, or nonoculomotor, aspects of the one-off and update stimuli, notably their colour. However, behavioural data suggest this was not the case. Because one-off trial locations were uniformly distributed around the target circle, some of the one-off targets fell within the expected spatial range. We compared RTs on one-off trials that fell within the expected spatial range (hence evoking feature-based but not spatial surprise) with RTs on one- off trials that fell outside the expected spatial range (hence evoking spatial surprise as well as feature-based surprise). Only on one-off trials that were spatially surprising were RTs slower than on expected trials, suggesting that feature-based surprise did not affect behaviour (Figure 2.3).

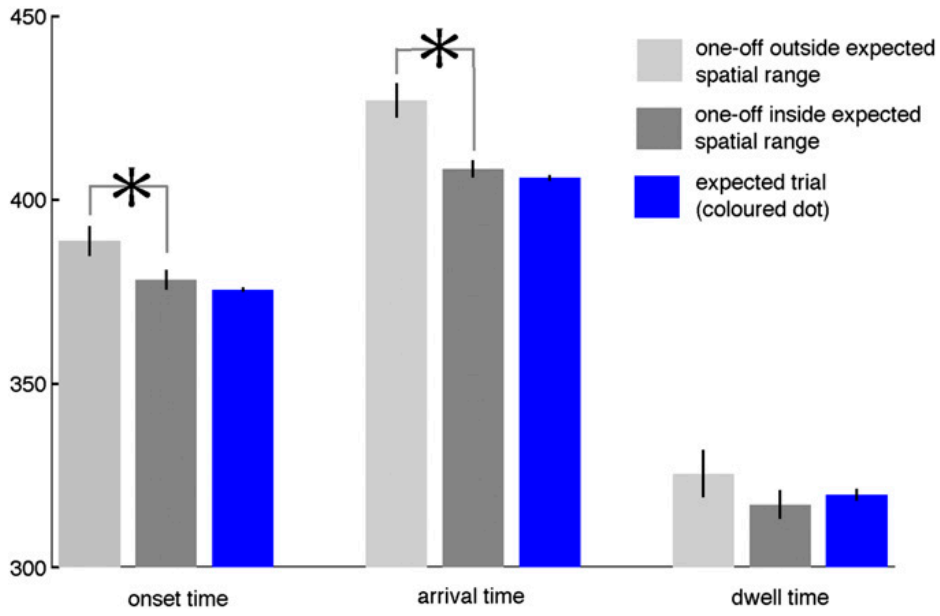

Figure 2.3 Behaviour on one-off trials is determined by spatial surprise, not feature-based surprise. One-off trial targets could fall anywhere on the circle; hence, they sometimes fell within the expected spatial range by chance. We divided the one-off trials into those with locations within 2 SDs of the mean of the expected distribution and those outside that range. Onset times were significantly slower on one-off trials outside the expected spatial range than inside it (t_{16} = 2.53, P = 0.011, paired samples t test for the group of 18 participants comparing mean onset time in each condition within subjects), but there was no evidence of a difference between one-off trials within the expected spatial range and expected trials (t_{16} = 0.11, P = 0.46). Similarly, arrival times were significantly slower on one-off trials outside the expected spatial range than on one-off trials within the expected spatial range (t_{16} = 3.63, P = 0.0010), but there was no evidence of a difference between one-off trials within the expected spatial range and expected trials (t_{16} = 0.15, P = 0.44). There was no evidence for a difference between conditions in dwell time (t_{16}= 0.37, P = 0.36; t16 = 0.016, P = 0.51); note that in the main behavioural analysis, dwell times were not reported to be longer on one-off trials than on expected trials. In summary, behaviour on the one-off trials that fell within the expected spatial range was similar to that on expected trials, whereas one-off trials outside the expected spatial range showed slower responses. This suggests that behavioural costs were determined by spatial reorienting/surprise rather than the nonspatial features (e.g., colour) of the targets. *P < 0.05.

2.3.2 Model-Based Measures of Surprise and Updating

The RT data suggest that participants performed the task as instructed [i.e., they used prior knowledge of the distribution of target locations to facilitate speeded eye movements (because time to arrive at the target was longer when the target location was unexpected)] and, furthermore, that participants updated their internal model, as instructed, on update trials but not on one-off trials, as indicated by the

data in Figure 2.2A. They also indicated that whereas the production of the saccadic response within trials was slowed for all surprising target locations (linking surprise and behavioural reorienting), the processing of the target location (indexed by dwell time) was specifically affected by updating.

To give a formal account of the two processes (surprise/reorienting and updating), we wished to analyze the behavioural and neural correlates of two information theoretical measures, I_S (surprise) and D_{KL} (updating), as described in the introduction. These measures must be calculated with reference to a model of the state of the environment; for example, the I_S is the log probability of an observed data point, given some model. Specifically, we wished to calculate the I_S and D_{KL} on each trial with reference to the participants' internal model of the target distribution.

We ran the Bayesian learning model with the same sets of target locations experienced by our human subjects to obtain an estimate of what their internal models of the environment would be on each trial. This allowed us to determine the surprise (I_S) associated with each stimulus and the degree of updating (D_{KL}) on each trial. In addition, we modeled the strength of participants' prior expectations on each trial as the Shannon entropy of the prior distribution (H_S), because behavioural work indicates that saccadic RTs may depend on this factor (Vossel et al., 2014). The H_S depended on both the variance of the generative targets' distribution, which was experimentally manipulated, and on learning. The formal definitions of these quantities in terms of the model are given in Methods in Chapter 3.

RT data indicate that both surprise and updating influence RT

As before, we broke down RTs into onset, arrival, and dwell times, and we modeled these values using a general linear model (GLM) analysis in which RT was modeled as a linear combination of effects of surprise/I_S, updating/D_{KL}, and H_S. The results are shown in Figure 2.4; there was an effect of both surprise and updating on all three measures: onset, arrival, and dwell times (statistics for effect of surprise are $t_{16} > 5.1$, $P < 0.00005$; $t_{16} > 5.1$, $P < 0.00005$; and $t_{16} = 1.9$, $P = 0.037$ and statistics for effect of updating are $t_{16} = 2.8$, $P = 0.0059$; $t_{16} = 4.3$, $P = 0.0003$; and $t_{16} = 2.9$, $P = 0.0049$ for saccadic onset, arrival, and dwell times, respectively, in one-sample t tests against zero). The entropy of the prior had no significant effect on any measure ($t_{16} = 1.3$, $P = 0.10$; $t_{16} = 1.3$, $P = 0.10$; and $t_{16} = 0.76$, $P = 0.23$ for one-sample t tests against zero as above).

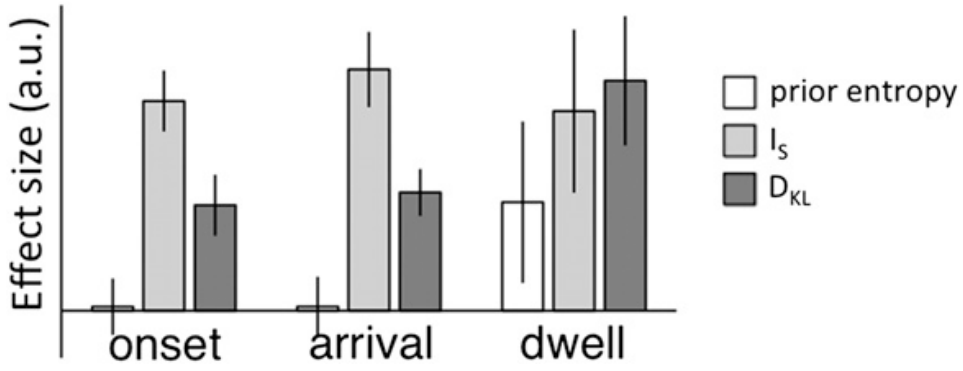

Figure 2.4 Effect of prior entropy, surprise, and updating on saccadic responses. Results of a GLM analysis on saccadic onset, arrival, and dwell times in which the regressors were entropy of the prior (the strength by which the stimulus location was predicted), I_S, and D_{KL}. The bar height is the mean effect size (beta value) for each parameter across participants, and error bars are the group SEM. a.u., arbitrary units.

2.3.3 Pupil dilation shows opposite responses to surprise and updating

Recent work has linked pupil dilation responses to psychological constructs similar to those investigated in this study. For example, Nassar et al. (2012) observed that pupil dilation increased on trials when participants detected change points in the environment (which should trigger both surprise and updating), whereas Preuschoff et al. (2011) have linked pupil dilation to uncertainty about the state of the environment and to changes in that uncertainty (and hence to learning).

To investigate whether pupillometric effects were correlated with surprise, updating, or both, we analyzed pupil dilation data in our task. We again used a GLM approach in which the effects of surprise, updating, prior entropy, and the main effect of task were modeled as a function of time in each trial. These were the same regressors used in the analysis of RT data presented in Figure 2.4. Pupillometric effects are shown in Figure 2.5, whereas the raw pupil dilation data (raw pupil area on each trial type) are shown in Figure 2.7. There was a positive effect of surprise on pupil diameter. On trials with higher values of I_S, there was a greater dilation of the pupil after viewing the target. This effect is in line with previous work that reported increases in pupil dilation when participants make observations that have a low probability, given their current model of the environment (Nassar et al., 2012) However, we also observed that updating was associated with a negative effect on pupil diameter [i.e., there was a relative decrease in pupil size on trials with a high D_{KL} (update trials)]. This effect peaked at a slightly later time than the surprise effect (Figure 2.5). Inspection of the raw data (Figure 2.7) indicates that pupil diameter was actually decreased in this period on update trials compared with expected trials (and

one-off trials).

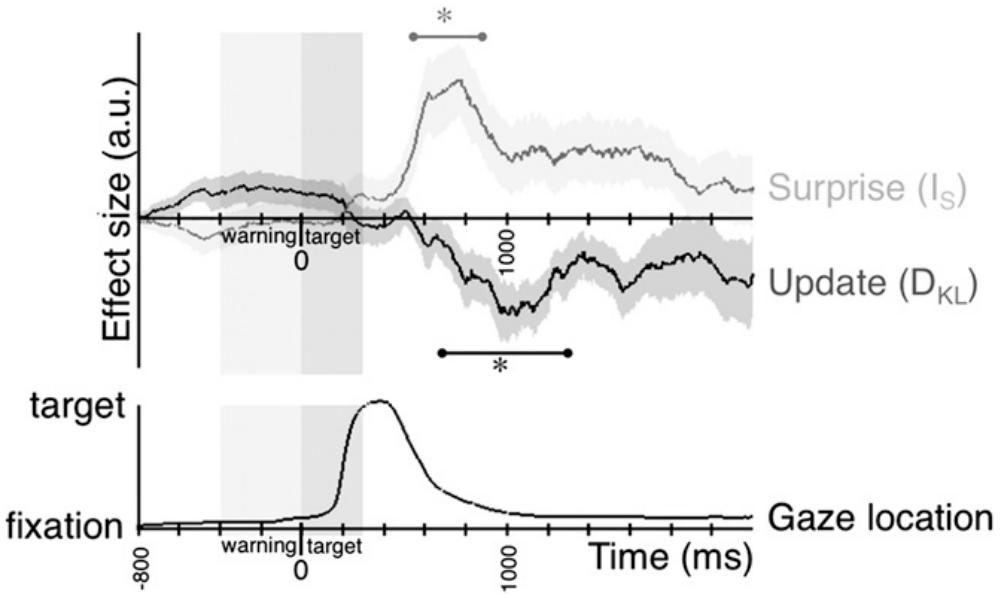

Figure 2.5 Pupil dilation is affected by both surprise and updating. The data are plotted as a function of time in the trial (shown on the x axis; labels in the lower panel apply also to the upper panel). The shaded regions represent the period in which the warning stimulus (a brightening of the central fixation cross) and the target were present on the screen. (Lower) Average eye displacement as a function of time, which gives an indication of when participants were looking at the target. (Upper) Results of a GLM in which pupil dilation at each time point was modeled with the same regressors used in the analysis of RTs (Figure 2.4): entropy of the prior, surprise (I_S), and updating (D_{KL}) on the trial. The main effect (mean pupil dilation at that time point across all trials) was also modeled. The results show that although surprise (I_S) was correlated with increased pupil diameter, updating (D_{KL}) was associated with a relative decrease in pupil diameter. The time periods in which these effects were significant [Z-score is >2.3 (i.e., P < 0.01 uncorrected, two-tailed)] are indicated by the starred bars above and below the plot. The effect of D_{KL} is significant in an interval 730–1,240ms after target onset; for I_S, the interval is 540–908ms.

To replicate the unexpected finding that updating negatively modulates pupil diameter, we conducted an additional behavioural experiment with 18 participants to check potential confounding factors that could have driven the pupillometric effect. In the replication experiment, participants could predict target positions as in the main task, but the task was to make orientation judgments about isoluminant Gabor patches while fixating centrally, eliminating potential confounds of target colour and eye movement. This additional experiment addressed two potential concerns about the pupillometric results: first, that the effect could be driven by the

different colours of the dots on update and one-off trials (because all targets were isoluminant Gabor patches in the additional experiment) and, second, that the effect could be contaminated by eye movements (because there were no eye movements in the additional experiment). We observed a very similar pattern of opposite effects of surprise and updating as in the main experiment (Figure 2.6).

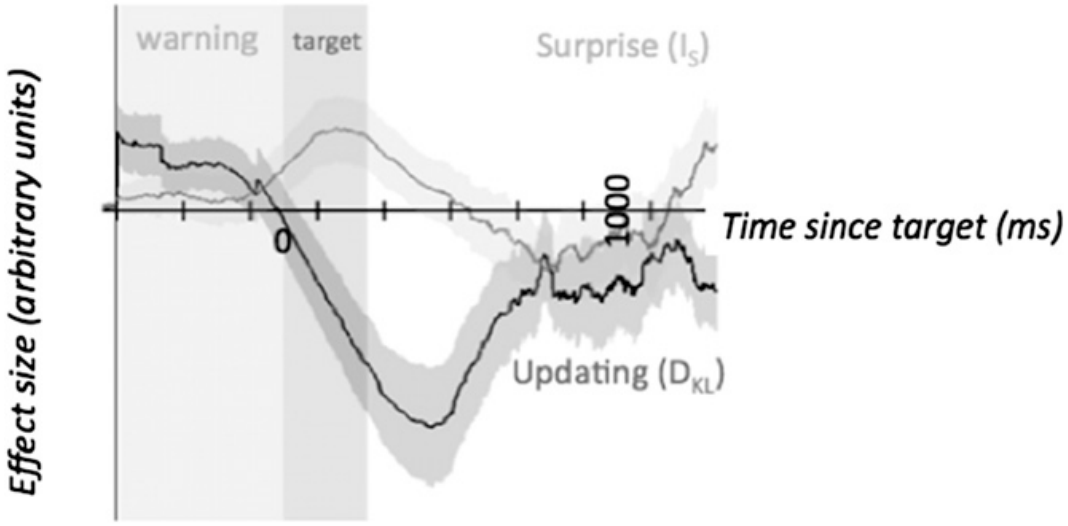

Figure 2.6 Pupil area effects from a separate behavioural experiment. Pupil effects for updating and surprise in an additional behavioural experiment, as reported in the main text, are shown. Note that the pattern of effects is similar to those seen in the main experiment, with opposite effect directions for surprise and updating and an earlier time of effect for surprise compared with updating.

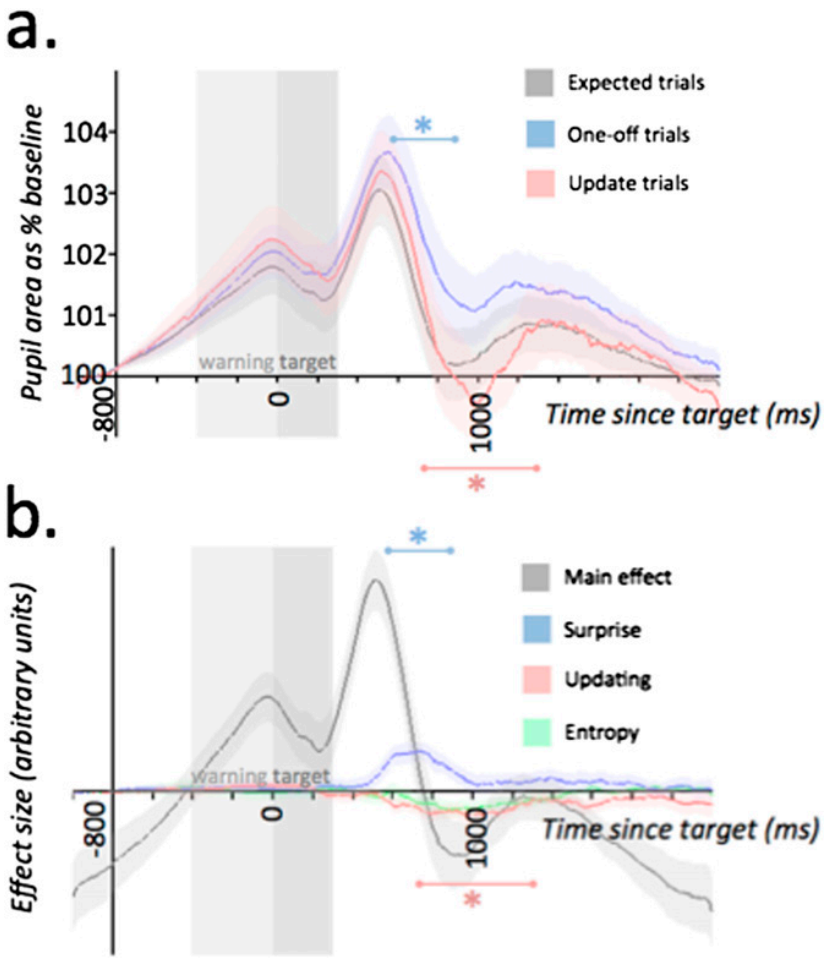

Figure 2.7 Raw pupil response on update, one-off, and regular trials. (A) Data on which the GLM analysis in Figure 2.5 is based: the pupil area on each of three trial types (update, one-off, and expected). Pupil area is expressed as a percentage (%) of the area on a trial-by-trial baseline, made up of the first 200ms of the trial (this was the same baseline correction approach used in the GLM analysis). The data lines are the mean pupil area across participants for each trial type, and the shaded surrounds are the SEM. The grey-shaded boxes indicate the period in which the warning signal and the target dot were on the screen. The blue and red lines with an asterisk indicate the period in which the GLM effects of surprise and updating were significant, and they are shown in this figure to aid comparison with Figure 2.5. There is a large pupil response to the presentation of the warning signal and the target that ramps up before the warning signal (as in the behavioural session, the timing of stimuli was predictable). Following the target, there is a differential response to update and one-off trials. (B) GLM results show all four regressors (i.e. the red and blue lines for effect of surprise and updating are the same as the light- and dark-grey lines indicating the effects of surprise and updating in Figure 2.5). Note that the "main effect," target-evoked effects that do not vary across trial types or with the computational regressors, dwarf the computationally specific effects. The fact that these nonspecific evoked effects are modeled out makes the GLM approach particularly sensitive.

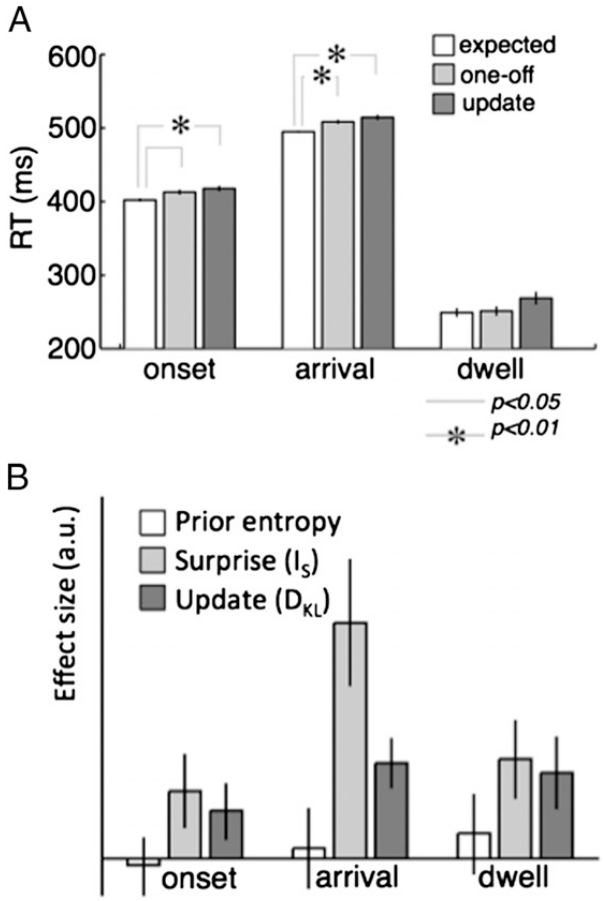

Figure 2.8 Behavioural eye data from the fMRI session. (A) Breakdown of saccadic reaction times (RTs) into saccade onset, arrival, and dwell times for data recorded during the fMRI session. These are defined in Methods. The plotting conventions are exactly as in Figure 2.2; this figure shows that the saccadic behaviour in the fMRI session was similar to that recorded in the separate eye-tracking session. Onset time was significantly longer on both update trials (t_{16} = 5.8, P = 1.1 × 10^{-5}, paired samples t test) and one-off trials (t_{16} = 5.3, P = 2.8 × 10^{-5}) compared with expected trials; there was no significant difference in onset time between update and one-off trials (t_{16} = 0.94, P = 0.18). Arrival time was significantly longer on both update trials (t_{16} = 3.7, P = 8.2 × 10^{-4}) and one-off trials (t_{16} = 2.6, P = 0.010) compared with expected trials; there was no significant difference in

arrival time between update and one-off trials (t_{16} = 0.46, P = 0.32). Dwell time was significantly longer in the update trials compared with both the one-off trials (t_{16} = 2.7, P = 0.0073, paired samples t test) and expected trials (t_{16} = 3.5, P = 0.0013). (B) Results of a GLM on saccadic onset, arrival, and dwell times for data recorded during the fMRI session. The analysis and plotting conventions are as in Figure 2.4.

2.4 DISCUSSION

2.4.1 Behavioural correlates of information theoretical measures

Participants' internal models of the targets' distribution could be expected to differ from the true (generative) distribution of target locations because they observed a limited number of data points and data were observed sequentially; hence, for example, in the first few trials after a change of distribution, the internal model would be based on relatively few data points. To obtain an estimate of a participant's beliefs (i.e., the state of the participant's internal model of the environment) on a trial-to-trial basis, we constructed a normative Bayesian learner (see details in Methods of Chapter 3). Its key features were that it updated its estimate of the underlying distribution using data only from non–one-off trials and the estimated distribution on each trial was updated with the new data point using Bayes' rule. Two features of the model merit emphasis here. First, no updating occurred on one-off trials (equivalently, one might say the learning rate was set to zero on one-off trials). This was in accordance with the instructions given to participants (that the locations of one-off targets did not predict the location of future targets). Behavioural evidence suggests participants complied with these instructions, because RTs on the trials following one-off trials were not slowed, as they should have been if participants had erroneously updated their model on the

one-off trial. However, it is worth noting that if participants did update on one-off trials, this could only lead to a false-negative result in terms of detecting update-specific brain activation because it would reduce the contrast between update and nonupdate trials. Second, on the first trial of a new block, the prior in force before the update was first "blanked" (replaced with a uniform distribution); hence, the new posterior model for the update trial (on which the prior for trial +1 was based) was obtained by updating from a uniform distribution using Bayes' rule. This "blanking" step was introduced to reflect a feature of the task design, of which participants were explicitly informed: that the location of one "block" of target dots and the next block (after an update trial) were totally independent, such that expectations about target locations based on previous targets should be disregarded.

The model was Bayes' optimal, implying that it returned the best possible estimate of the underlying distribution based on the data points it was given (the same data points that human participants observed) (Cox, 1946). Furthermore, the model was optimal in that it was supplied with correct assumptions about the structure of the world. For example, the model correctly assumed that when an update occurred, the new distribution of target locations did not depend on the previous distribution; this reflected the true form of the generative process by which target locations were chosen in the experiment. Therefore, the internal model generated by our Bayesian learning algorithm represents an upper bound on the accuracy with which any participant could have estimated the underlying distribution.

The one-off trials are evenly spread so they appeared at both the beginning

of a block (when the model is relatively uncertain) and the end of a block (when the model is relatively certain). While we cannot exclude an effect of 'normal' trials throughout a block, we are confident that it did not matter whether the one-off trials appeared at the first or the second half of a block. The maximum width or a prior is comparatively small with regard to the probability distribution of one-off trials and participants' reaction time is fast for trials t+1 after an update. Given the sparseness of one-off trials, we cannot test statistically whether these trials are affected by a certainty about a prior.

2.4.2 Is the updating effect simply due to a quantitative enhancement of surprise?

Given that RTs showed a parametric effect of surprise with an additional effect of updating, it might be argued that the occurrence of an update stimulus simply caused a quantitative enhancement of the surprise effect. However, several results argue that, in fact, the additional RT cost on update trials reflects a qualitatively different neural process to the effect of surprise. First, participants' behaviour on the trial following an update or one-off trial (Figure 2.2A) indicates that participants adjusted their internal model of targets' distribution on update trials but not on one-off trials. Second, dwell time was enhanced on update trials compared with one-off trials, a possible behavioural correlate of updating. Finally, there were dissociable effects of surprise and updating on both pupil diameter and behaviour.

It is worth noting that although the additional experiment demonstrates that the pupillometric effects reported in the main experiment were not an artifact of

target colour or eye movements, even within the main experiment, these confounds were not likely to have driven the observed effects because:

i) In the original experiment, although the target dots were different colours on one-off and update trials, the targets on update trials were the same colours as the targets on expected trials (i.e., if an update trial had a red target, all the subsequent expected trials in that run also had red targets). Therefore, any small luminance differences could not account for the difference between update and expected trials shown in Figure 2.7B.

ii) Although participants did make eye movements during the task, these eye movements were complete by the time point at which the effect of updating became significant (Figure 2.5).

Dissociable effects of surprise and updating on pupil diameter have not previously been reported, probably because these parameters are generally strongly correlated in most experimental designs. However, the negative effect of D_{KL}/updating on pupil diameter was somewhat unexpected because of theoretical work linking pupil dilation and noradrenaline/norepinephrine to learning and uncertainty. Pupil dilation in the absence of luminance changes may correlate with the firing of cells in the locus coeruleus, and hence with the release of noradrenaline, although the strength of the correlation is debated (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005; Preuschoff et al., 2011). Theoretically, it has been argued that noradrenaline signals uncertainty about the state of the environment (i.e., estimation uncertainty) (Dayan, 2012; Yu & Dayan, 2005) or that it acts as a kind of "reset" signal for internal models (Bouret & Sara, 2005; Sara, 2009).

It is unclear how the update-specific decrease in pupil area fits with

computational theories of noradrenaline (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005), which have linked learning, uncertainty, noradrenaline release, and pupil dilation and would tend to predict an increase in pupil dilation on update trials. Our results are rather in line with a 'network reset' function of NA than a mere increase of learning rate ('gain' model of NA) given that an old prior is 'blanked' by the occurrence of a new prior. However, our task design does not allow for a direct comparison of the two concepts, as the learning rate for update trials is almost 1 and 0 for the remaining trials. A key difference between the present paradigm and probabilistic learning tasks used in previous studies (Nassar et al., 2012) is that in the present task, unlike in probabilistic learning tasks, updating is not driven by increased uncertainty; in fact, at the same moment that the observer detects a change in the environment (an update target), he also gains information about the new target distribution, reducing uncertainty. This interpretation would suggest that pupil increases in learning tasks are driven by uncertainty, or the influence of uncertainty on learning, rather than by learning or change per se.

The observation of a decrease in pupil diameter on update trials, at a different time point from previously reported uncertainty-related pupil dilations, is unique (previous experiments have not reported results in this time period; indeed, previous studies have not dissociated surprise and updating). The timing of the effect raises the intriguing possibility that the late dip in pupil diameter on update trials reflects neural processing of the stimulus itself (modulated by its relevance) rather than a modulation of the prior preparation for learning (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005).

In relation to these results, it should be noted that multiple psychological

factors, including arousal and attention, can affect pupil dilation and that the effect should not be overinterpreted in the absence of pharmacological work. Nonetheless, the effect is replicable and provides robust evidence for differential processing of update and non-update stimuli.

Chapter 3: Dissociable Neural Correlates of Surprise and Model Update

3.1 INTRODUCTION

In this chapter, I will present and discuss the neural activity patterns associated with surprise and behavioural updating (O'Reilly, Schüffelgen et al., 2013). In the previous chapter, I demonstrated that surprising visual stimuli can trigger both, a behavioural response to the surprise (i.e. attentional reorienting and replanning of a saccade to the unexpected location) and a process of updating one's prior expectation about the locations in which future stimuli will fall. Furthermore, differences in pupil size measured when participants were either merely reorienting or updating suggest the possibility of identifying separate underlying neural correlates of these two processes. Here, I will show effects measured with fMRI that correlate with information theoretical concepts, in particular with surprise [Shannon information (I_S)] and updating [Kullback–Leibler divergence (D_{KL})] as introduced in the previous chapter.

3.2 METHODS

Participants were asked to perform the same task as described in Chapter 2. While in the MRI scanner, we recorded eye movements using an MRI compatible eye tracker [EyeLink 2000 (SR Research, Ottawa, Ontario, Ca) sampling at 500 Hz].

3.2.1 Scanning Parameters and Analysis

A 3T scanner (Magnetom Verio, Siemens Healthcare, Erlangen, Germany) was used to acquire the T1-weighted high-resolution structural and the functional MRI data. For the functional MRI we used an echo planar imaging (EPI) sequence with the following parameters: Repetition time (TR) = 2.41s, voxel-resolution 3x3x3 mm^3, echo time (TE) = 30ms, Flip angle = 90°. T1-weighted structural images were acquired for subject alignment using a magnetization-prepared radio-frequency pulses and rapid gradient-echo (MPRAGE) sequence with the following parameters: Voxel resolution 1x1x1 mm^3 on a 174x192x192 grid, TE = 4.70ms, inversion time (TI) = 900ms, TR = 2.04s.

FMRI data processing and preprocessing were carried out using Centre for Functional MRI of the Brain (FMRIB) Expert Analysis Tool (FEAT), version 6.0, which is part of the FMRIB's Software Library (FSL; www.fmrib.ox.ac.uk/fsl, (Smith et al., 2004)). The following prestatistics processing was applied: motion correction using MCFLIRT (Motion Correction with FMRIB Linear Image Registration Tool), slice timing correction using Fourier-space time-series phase-shifting, non-brain removal using BET (FMRIB Brain Extraction Tool), spatial smoothing using an 8.0-mm FWHM Gaussian kernel, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted, least-squares, straight-line fitting; σ = 50 s).

At the first level (individual subject analysis), time-series statistical analysis was carried out using FILM with local autocorrelation correction. A GLM analysis was carried out with four task regressors (main effect of task, entropy of the model, $\mathrm{I_S}$, and $\mathrm{D_{KL}}$) as described in Chapter 2, plus six regressors of no interest capturing the

effects of head motion. Z-statistics (Gaussianised t-statistics) were passed to the second (group)-level analysis. At the group level, individual participants' Z-statistic images were normalised to the Montreal Neurological Institute brain using nonlinear registration [FNIRT (FMRIB Non-linear Image Registration Tool)] and entered into a random effects analysis in which the average effect of each regressor across the group was tested against zero. Higher-level analysis was carried out using the FMRIB's Local Analysis of Mixed Effects (FLAME) stage 1 with automatic outlier detections and deweighting. Group statistics were corrected for multiple comparisons using a cluster size-based approach (as implemented in FSL), with a cluster-forming threshold of $Z > 2.3$ and a cluster significance threshold of $P < 0.05$ (corrected).

The model was a normative Bayesian learner implemented numerically in MATLAB (MathWorks). It estimated the probability density distribution from which each target location, α_t, was drawn, based on the current and previous observations (i.e., $\alpha_{1:t}$), by assigning probabilities to putative parameters of the underlying Gaussian distribution $\mathrm{p}(\mu_t, \sigma_t|\alpha_{1:t})$. The posterior probabilities assigned to each parameter pair $\mathrm{p}(\mu_t, \sigma_t|\alpha_{1:t})$were calculated as follows. If the trial type is one-off, no updating occurs; that is,

$$\mathrm{p}(\mu_t, \sigma_t\,|\alpha_{1:t}, \mathrm{one-off}_t = \mathrm{p}(\mu_t, \sigma_t\,|\alpha_{1:t-1}), \tag{8}$$

Otherwise, probabilities are updated using Bayes' rule:

$$\mathrm{p}(\mu_t, \sigma_t|\alpha_{1:t}) \propto \mathrm{p}(\alpha_t|\mu_t, \sigma_t)\mathrm{p}(\mu_t, \sigma_t|\alpha_{1:t-1}, \varphi), \tag{9}$$

where the variable φ denotes dependence of the prior on trial type as follows. If the trial type is update, a uniform prior over (μ_t, σ_t) is used:

$$\mathrm{p}(\mu_t, \sigma_t \,|\alpha_{1:t-1}, \mathrm{update}_t) = \frac{1}{360 \times 50} \tag{10}$$

where the denominator 360 × 50 represents the number of possible values for (μ_t, σ_t) probed in the numerical implementation of the model. If the trial type is expected (i.e., any trial that is not one-off or update), then the prior on trial *t* is obtained without modification from the posterior on trial *t* − 1:

$$\mathrm{p}\big(\mu_t, \sigma_{t,} \big|\alpha_{1:t-1}, \mathrm{expected}_t\big) = \mathrm{p}(\mu_{t-1}, \sigma_{t-1} \,|\alpha_{1:t-1}) \tag{11}$$

A prior probability density function for the subsequent trial, *t*+1, was obtained from the posterior on trial *t*, and expressed as a function of target location (rather than parameters $\mu_t, \sigma_{t,}$) as follows:

$$\begin{aligned} &\mathrm{p}(\alpha_{t+1}|\alpha_{1:t}) = \\ &\mathrm{p}(\mathrm{expected}_{t+1}) \sum_{\mu_{t+1}} \sum_{\sigma_{t+1}} \mathrm{p}\big(\alpha_{t+1}\big|\alpha_{t+1} \sim N(\mu_{t+1}, \sigma_{t+1})\big) \\ &\quad \times \mathrm{p}(\,\mu_{t+1}, \sigma_{t+1}|\alpha_{1:t}) \\ &\quad + \mathrm{p}(\mathrm{one-off}_{t+1} \cup \mathrm{update}_{t+1})\, \mathrm{p}(\alpha_{t+1}|\mathrm{uniform}(0°, 360°)) \end{aligned} \tag{12}$$

The probabilities of each trial type occurring on trial *t* + 1, p(one-off$_{t+1}$), p(update$_{t+1}$), and p(expected$_{t+1}$) were simply set to the true probability of each trial type in the generative model [i.e., p(one-off$_{t+1}$) = 1/4, p(update$_{t+1}$)= 1/15, and p(expected$_{t+1}$) = 1 − (p(update$_{t+1}$) + p(one-off$_{t+1}$)].

3.2.2 Information Theoretic Regressors

For the analyses presented in Figure 2.4 (analysis of saccadic RTs), Figure 2.5 (analysis of pupil area), and the fMRI whole-brain analysis and ROI analysis presented in Figure 3.3 and Figure 3.5, respectively, we used a GLM analysis with information theoretic regressors based on the learning model. The four task

regressors were main effect of task, entropy of the model, I_S, and D_{KL}. Each regressor used in fMRI analysis was defined as a series of 300 delta functions (events of 0.1 s in duration) at the times of target appearance (for 300 trials), weighted by task parameters and convolved with the haemodynamic response function. The weighting of each regressor on each trial was defined as follows. All trials took value 1 as the main effect of task. Entropy of the model on trial *t* was expressed as

$$H(t) = \sum_{\alpha=1}^{360} \mathrm{p}(\alpha_t|\alpha_{1:t-1}) \log\left(\frac{1}{\mathrm{p}(\alpha_t|\alpha_{1:t-1}}\right), \tag{13}$$

where $\mathrm{p}(\alpha_t|\alpha_{1:t})$ is defined analogously as in Eq. 12. The I_S of trial *t* is defined as

$$\mathrm{I_S}(\alpha) = -\log \mathrm{p}(\alpha_t|\alpha_{1:t-1}), \tag{14}$$

where $\mathrm{p}(\alpha_t|\alpha_{1:t-1})$ defined analogously as in Eq. 12. The D_{KL} on trial t is defined as

$$\mathrm{D_{KL}} = \sum_{\alpha=1}^{360} \mathrm{p}(\alpha_t|\alpha_{1:t-1}) \log\frac{\mathrm{p}(\alpha_t|\alpha_{1:t-1})}{\mathrm{p}(\alpha_t|\alpha_{1:t})}, \tag{15}$$

where $\mathrm{p}(\alpha_t|\alpha_{1:t-1})$ is defined analogously as in Eq. 12 and $\mathrm{p}(\alpha_t|\alpha_{1:t})$ is defined analogously to the definition of $\mathrm{p}(\alpha_{t+1}|\alpha_{1:t})$in Eq. 12:

$$\begin{aligned}\mathrm{p}(\alpha_t|\alpha_{1:t}) = \mathrm{p}(\mathrm{expected}_t) \sum_{\mu_{t+1}} \sum_{\sigma_{t+1}} \mathrm{p}\big(\alpha_t\big|\alpha_t{\sim}N(\mu_t,\sigma_t)\big) \times \mathrm{p}(\mu_t,\sigma_t|\alpha_{1:t}) \\ + \mathrm{p}(\mathrm{one-off}_t \cup \mathrm{update}_t)\,\mathrm{p}(\alpha_t|\mathrm{uniform}(0°,360°))\end{aligned} \tag{16}$$

Output of the model is illustrated in Figure 3.1 and Figure 3.2. The shared variance between each pair of regressors was relatively low; the R^2 value for each pair was as follows: prior entropy and $\mathrm{I_S}$/ R^2 = 0.14; prior entropy and $\mathrm{D_{KL}}$/ R^2 = 0.14; and $\mathrm{I_S}$ and $\mathrm{D_{KL}}$/ R^2 = 0.15.

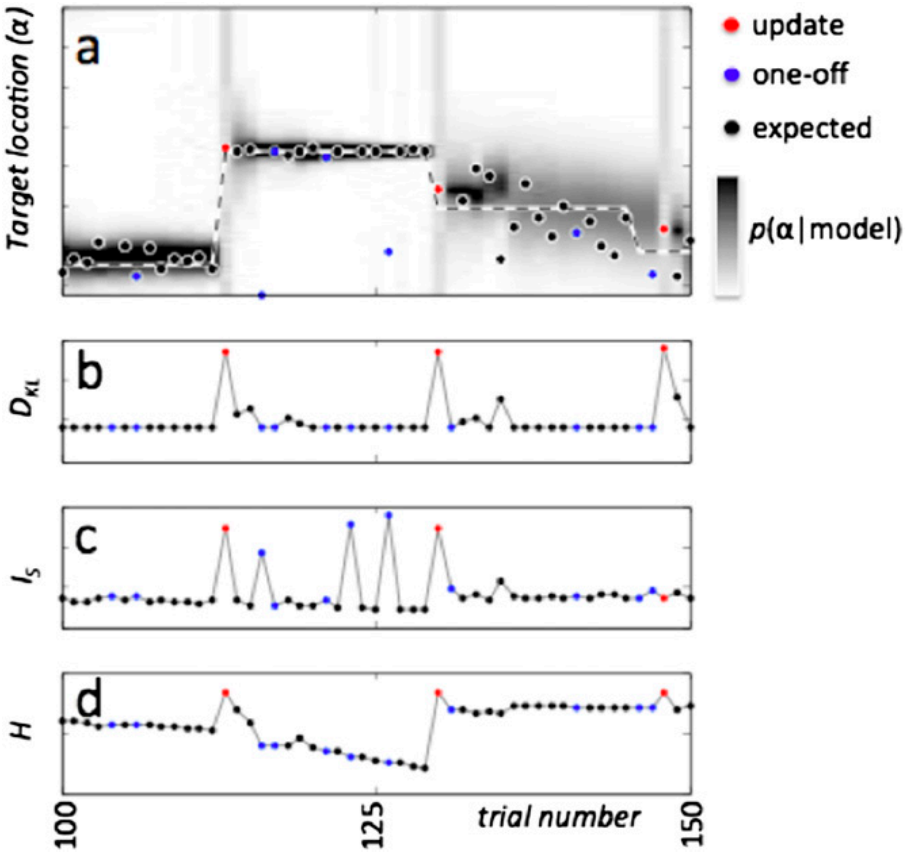

Figure 3.1 Model and regressors. All panels show the data from 50 trials, with trials on the x-axis. Trial types are indicated by dot colour as in the key. (A) Plot of the state of the learning model over 50 trials. On the y-axis is target location (α = 0–360°). The dashed line indicates the mean of the actual (generative) Gaussian from which expected and update trials were drawn. Dots indicate the actual values of target location on each trial. Shading indicates the probability that a target will appear at the location, given the state of the model on that trial. Note that the distribution of probability is broadest at the start of each run and that it is broader when the generative distribution has a higher variance [increased spread of data points about the mean (e.g., toward the right of the plot)]. (B–D) Computational regressors used in data analysis and based on the model in A. The equations by which these regressors were derived are given in Methods. Again, the coloured dots are the actual values (of the regressors) on each trial. (B) Update regressor: D_{KL} between the state of the model before and after observing the target location on each trial. (C) Surprise regressor: I_S of each target location, given the model. (D) Entropy (H) of the model is roughly equivalent to the spread of probability density over possible values of α, as shown by the spread of the shaded area in A.

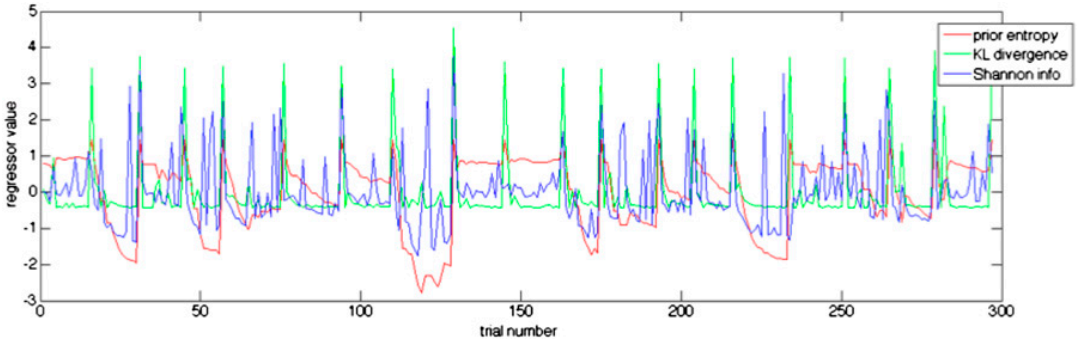

Figure 3.2 Regressors over all 300 trials. This plot illustrates the values of the three information theoretic regressors (prior entropy, I_S, and D_{KL}) over all 300 trials of the task (this was one of three different schedules used). Each regressor was normalised such that its mean value across all 300 trials was 0 and its variance across all 300 trials was 1. Note that although there is some correlation between regressors (e.g., on update trials, I_S, D_{KL}, and entropy are all high), the regressors were not strongly correlated overall. The shared variance (R^2) between prior entropy and I_S was 0.14, that between prior entropy and D_{KL} was 0.14, and that between D_{KL} divergence and I_S was 0.15.

3.3 RESULTS

3.3.1 Distinct Brain Networks for Surprise and Updating

The behavioural results described in the previous chapter suggested that although the within-trial behaviour (saccadic RT) was affected by both surprise and updating, stimuli that elicited updating (high D_{KL}) were processed differently because: (i) behaviour on the trials subsequent to the one-off surprise trials and the update trials supported the hypothesis that participants updated on update trial but not one-off trials; (ii) dwell times were greater on update trials than on one-off surprise trials; and (iii) measures of updating, D_{KL},were associated with an opposite effect on pupil dilation to measures of surprise such as I_S.

To determine the neural correlates of surprise and updating, we collected fMRI data from the same group of 17 participants for whom behaviour was presented in the

previous chapter while performing the task. (replications of Figure 2.2 and Figure 2.4, using the eye tracking data from the fMRI session, are provided in Figure 2.8); however, pupillometry was not possible with the in-scanner setup.

We began by analyzing the pattern of activity associated with surprise and updating across the whole brain, by means of a GLM analysis implemented using the Centre for Functional MRI of the Brain (FMRIB) software library (Smith et al., 2004). The task was modeled in terms of four model-derived regressors: main effect of task (i.e., all trials, modeled as a delta function at the time of the target appearance) and three information theoretical parameters, the prior entropy (the predictability of the environment), the I_S of the observed target location (surprise), and the D_{KL} of the posterior from the prior (updating).

We observed a spatial dissociation between activity correlated with surprise (I_S) and activity correlated with updating (D_{KL}); no significant activity associated with prior entropy was observed (in line with the lack of behavioural effects relating to this parameter, see Chapter 2). Surprise (I_S) was strongly correlated with activity in the posterior parietal cortex, with the peak activation in the superior parietal lobule (SPL) and extending along the IPS. In contrast, updating (D_{KL}) was strongly correlated with activity in the anterior cingulate cortex (ACC), particularly the rostral cingulate motor area (rCMA) and extending into the ventral part of the adjacent presupplementary motor area (pre-SMA). These spatially distinct effects, which were the only effects to survive multiple comparisons correction from the whole-brain analysis, are shown in Figure 3.3A and Figure 3.5A. For illustrative purposes, the uncorrected statistical maps from which the corrected statistics were drawn are shown in Figure 3.4.

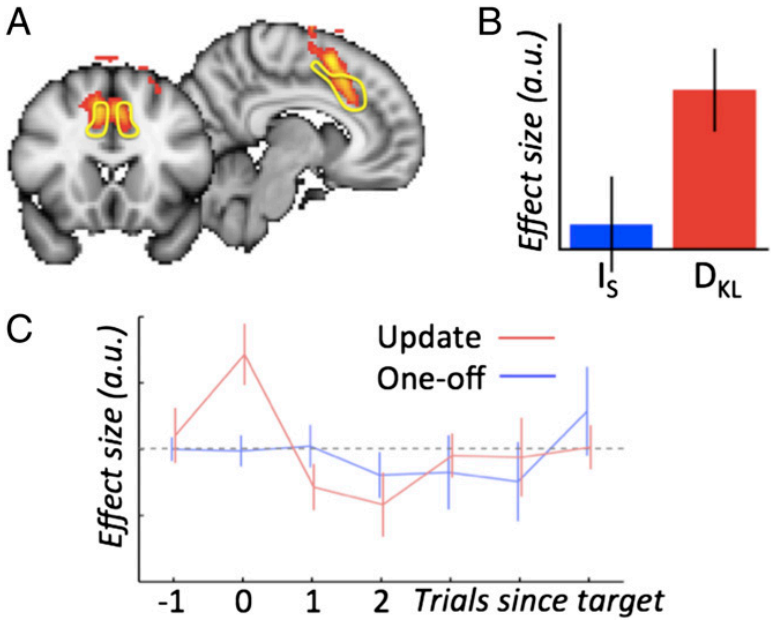

Figure 3.3 Effect of updating in the ACC. (A) Results of whole-brain fMRI analysis. This region in the ACC and pre-SMA was the only area in which there was a significant effect of updating (contrast shows all voxels with a parametric effect of D_{KL} as defined in Methods in Chapter 2), using cluster size-based multiple comparisons correction. The colour scale is 2.3 < Z < 3; the peak Z-score is 3.1 at MNI coordinates (6, 10, 54). The ROI denoted by the yellow line is the ACC ROI, a region described as the rCMA zone in a diffusion- weighted parcellation of the cingulate cortex (32). This was the ROI used in the analyses shown in B and C. (B) Effect size for I_S and D_{KL} in the ACC ROI, where bar height is the mean effect across the group of participants and error bars are the SEM. (C) Raw activity in the ACC ROI plotted as a function of trial-in-run; plotting conventions are as in Fig. 2.

Although we had specific parametric predictions about brain activity based on the I_S and D_{KL}, as a confirmatory analysis, we also analyzed the effects of different trial types: update trials + one-off trials (both types evoke surprise/behavioural reorienting) and update trials - one-off trials (to isolate behavioural reorienting). The maps obtained are very similar to those obtained using our parametric regressors (Figure 3.8). Region of interest (ROI) effects using the trial type and parametric regressors are compared in Figure 3.9. To characterise these effects further, we extracted and analyzed the fMRI signal from anatomically defined ROIs as follows:

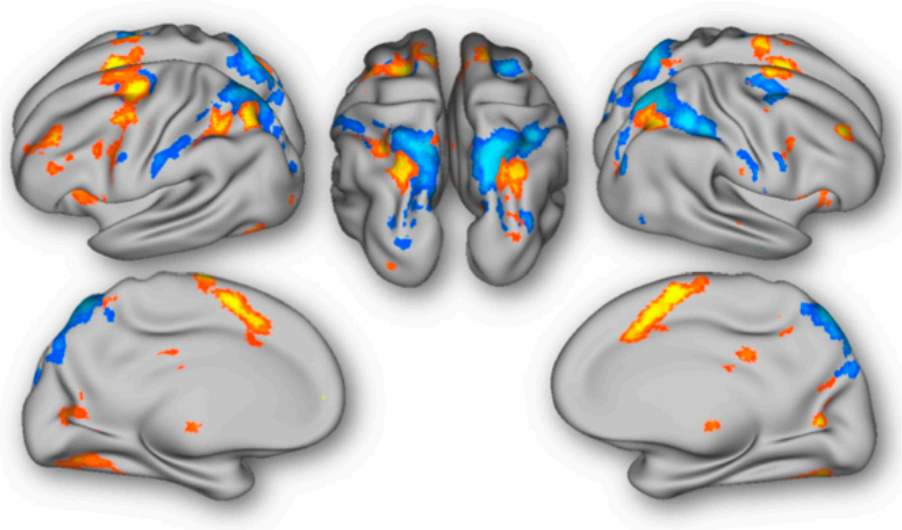

Figure 3.4 Uncorrected whole-brain activity patterns for surprise (I_S) and updating (D_{KL}). These maps are not corrected for multiple comparisons, and hence are shown only to give an impression of the overall pattern of activity that underlies the multiple comparisons-corrected effects reported in the main text. The maps are thresholded at Z > 2.0 uncorrected. Blue colours indicate surprise, and red/yellow colours indicate updating.

3.3.2 ACC: Updating

The presence of updating-specific activity in the ACC was of particular interest because the ACC has previously been implicated in the updating of expectations about the reward environment (Behrens et al. 2007; Jocham et al. 2009; Kennerley et al. 2006) but not in spatial updating. At the same time, the ACC has been implicated in studies of spatial reorienting (Mesulam et al. 2001), but the nature of its specific contribution has remained unclear so far.

To characterise the role of the ACC in our task, we extracted the fMRI signal from an anatomical ROI, the rCMA, using a mask derived from a diffusion-imaging parcellation of the human cingulate cortex (Beckmann, Johansen-Berg, & Rushworth, 2009). This ROI corresponded to the more ventral portion of the updating activity observed in the whole-brain analysis (Figure 3.3). Figure 3.3B shows the effect size in this ROI for updating (D_{KL}, red) and surprise (I_S, blue); as expected, in this ROI, there

was a significant effect of updating (P = 0.015, two- tailed t test against zero), but note also the lack of effect for surprise (P = 0.61). To characterise the pattern of ACC activity across a block of trials, we plotted raw activity (averaged across a time bin 6-7s after the target, timed to coincide with the peak of the hemodynamic response function) in the rCMA ROI by trial number within a block for block of trials around both update and one-off trials (using the same plotting conventions as in Figure 2.2). The rCMA showed a pattern of activity that suggested it was uniquely activated on the first trial of a block: The rCMA was activated on update trials (t_{16} = 2.7, P = 0.0158, two-tailed t test of raw activity in the rCMA across a time bin 6-7s after the target on update trials vs. mean activity at the same time point across all other trials) but not on subsequent trials or on one-off trials.

3.3.3 Posterior parietal cortex: Action reprogramming evoked by surprise

Surprise (I_S) was correlated with activity in the posterior parietal cortex. The activity extended along the SPL and into the IPS. Its peak fell in a region on the posterior medial bank of the IPS and adjacent SPL [Montreal Neurological Institute (MNI) coordinates -18, -60, 58]. This region, which we will call IPS3 in agreement with the terminology used by Mars et al. (2011), has been postulated to be the human homolog of the monkey LIP because, like macaque LIP, it has strong connections with the frontal eye fields (Mars et al., 2011) and superior colliculus (Rushworth et al., 2006), and is retinotopically organised (Sereno et al., 2001; Swisher et al., 2007). Given this motoric connectivity profile, the presence of the surprise (I_S)-evoked activity in the IPS may well reflect reprogramming of the current saccade to account

for the new (surprising) target location.

In macaque parietal cortex, LIP is one of two regions that are especially linked to covert attention and overt eye movements; the other is area 7a (Gottlieb & Snyder, 2010). We were intrigued by the possibility that the two areas could play different roles in behavioural reorienting and updating in our task, because in contrast to LIP, area 7a is less strongly connected with oculomotor regions but more strongly connected with the ACC, in which we observed update- specific activity (Cavada & Goldman-Rakic, 1989a, 1989b). Additionally, there is a suggestion from the patient literature that lesions of the IPS (including human LIP homolog IPS3) and IPL (including human 7a homolog PGp) differently affect motoric and nonmotoric aspects of orienting (Milner, 1996; Pierrot-Deseilligny & Muri, 1996).

To address this hypothesis, we defined anatomical ROIs in the two functional areas. The ROIs were based on the diffusion-imaging parcellation of Mars et al. (2011) and were defined based on the two connectivity-defined regions labeled IPS3 (LIP homolog) and PGp [a region on the posterior lateral bank of the IPS and adjacent posterior IPL that has been identified as a putative homolog of monkey area 7a (Caspers et al., 2008; Mars et al., 2011)]; the peak voxels were at MNI coordinates (±26, -54, 60) and (±32, -64, 44), respectively. There was indeed a significant difference between ROIs in their responsiveness to updating (D_{KL}), as opposed to surprise (I_S). We extracted the effect size from each of these ROIs in each participant. A within-subjects ANOVA with the factors ROI (IPS3 vs. PGp) and effect (I_S vs. D_{KL}) indicated a significant interaction between ROI and condition ($F(63,3) = 12.5$, $p < 0.05$); in Figure 3.5B, it can be seen that the IPS3 region was more active for surprise than updating, whereas the PGp region was more active for updating than surprise.

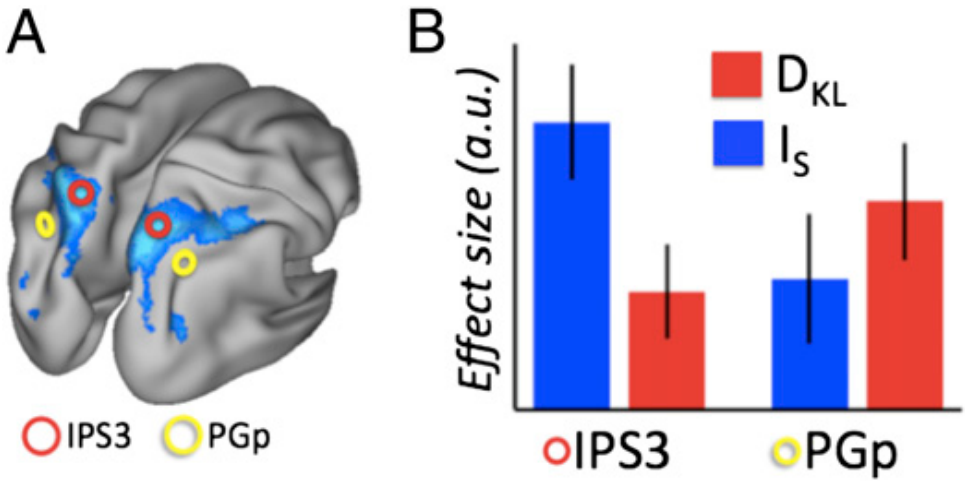

Figure 3.5 Effect of surprise in the posterior parietal cortex. (A) Results of whole-brain fMRI analysis. This region in the posterior parietal cortex was the only area in which there was a significant effect of surprise (contrast shows all voxels with a parametric effect of I_S as defined in Methods in Chapter 2), using cluster size-based multiple comparisons correction. The colour scale is 2.3 < Z < 3, and the peak Z-score is 4.8 at MNI coordinates (-18, -60, 58). The red circle is the IPS3 ROI (with coordinates ±26, -54, 60), and the yellow circle is the PGp ROI (with coordinates ±32, -64, 44) (33). (B) Effect size for the I_S and D_{KL} in each ROI; the bar height is the mean effect across the group of participants, and the error bars are the SEM. There is a significant ROI × condition interaction.

3.3.4 Parietal activity correlates with within-trial reprogramming costs

Throughout this analysis we have interpreted the neural activity associated with I_S in terms of the behavioural reprogramming of saccades, necessitated when a low-probability stimulus location occurs. In support of this hypothesis, activity in both the IPS and PGp was significantly related to RTs (higher activity on longer RT trials), whereas there was no significant relationship between RTs and ACC activity. We conducted a multiple regression in which RTs (taken separately for one-off trials and update trials) were modeled in terms of activity in the IPS3, PGp, and ACC. For both one-off and update trials, there was a significant effect of activity in the PGp and

IPS3, but not in the ACC, on RT. Results are shown in Figure 3.6. The p values for a group t-test of effect size (the effect of ROI activity on RT in the multiple regression) were as follows: for one-off trials, $t_{16} = -1.6$, $p = 0.13$; $t_{16} > 4.0$, $p = 0.001$; and $t_{16} > 4.0$, $p = 0.001$ for the ACC, PGp, and IPS3, respectively, and for up- date trials, $t_{16} = 1.4$, $p = 0.181$; $t_{16} = 3.8$, $p = 0.0016$; and $t_{16} = 4.3$, $p = 0.0005$ for the ACC, PGp, and IPS3, respectively (two-tailed t test against null hypothesis of zero effect).

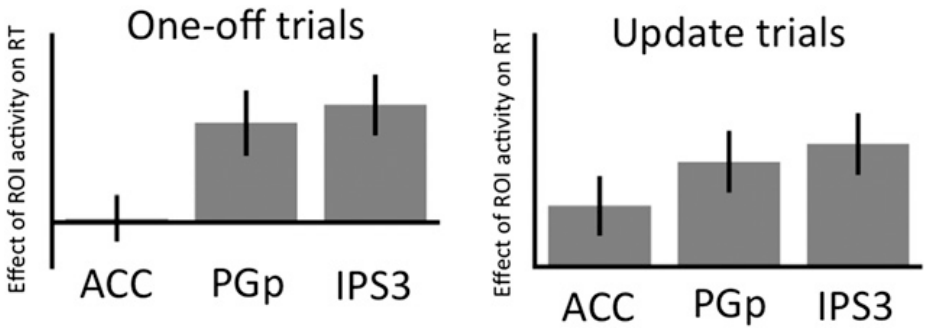

Figure 3.6 Parietal activity, but not ACC activity, predicts RT on both one-off and update trials. Bars show the effect size (beta value) ± SEM from a multiple regression in which activity in the three ROIs (ACC, PGp, and IPS3) was used to predict RT. The effects of PGp and IPS3 activity on RT are significant for both one-off trials and update trials, and the effect of ACC activity is not significant in either case. A breakdown of this effect into onset, arrival, and dwell times is shown in Figure 3.7.

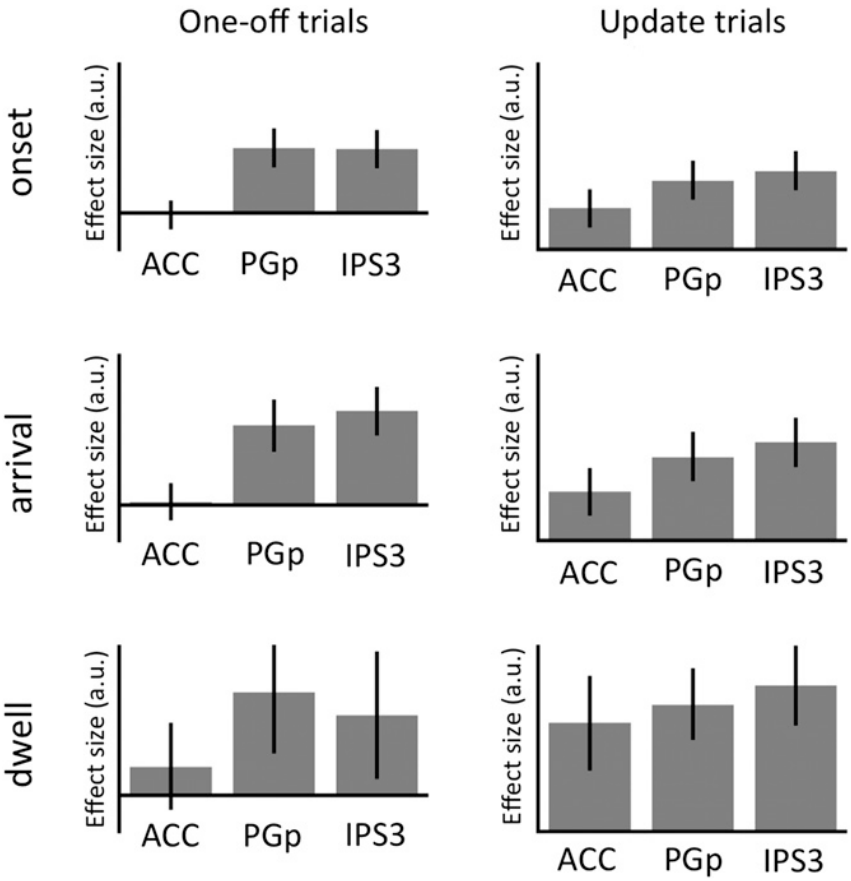

Figure 3.7 Regression of onset, arrival, and dwell times on activity in the ACC, IPS3, and PGp. Bars show the effect size (beta value) ± SEM from a multiple regression in which activity in the three ROIs (ACC, PGp, and IPS3) was used to predict each behavioural measure, as in Figure 3.6. Note that the parietal ROIs are more strongly associated with all behavioural measures; the relationship between ACC activity and behaviour does not reach statistical significance in any case.

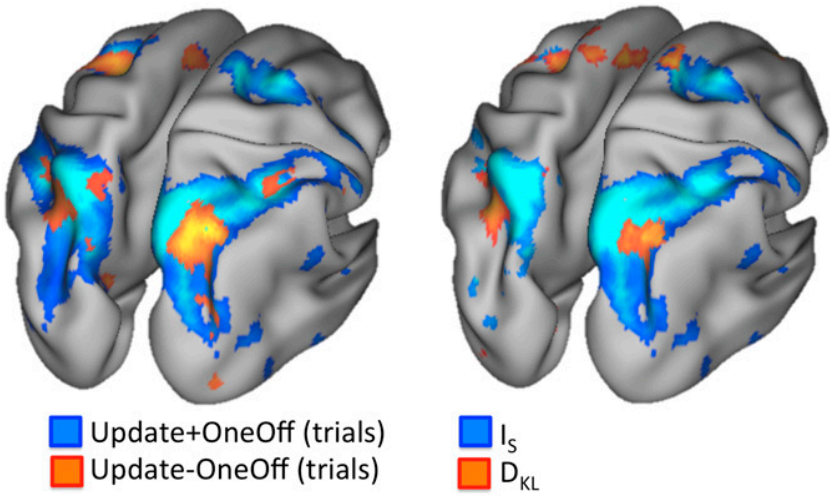

Figure 3.8 fMRI analysis on trial type (update ± one-off trials). Although the parametric regressors I_S and D_{KL} provide a specific model to compare against brain activity, it is useful to check that the activations evoked by different trial types (update and one-off trials) are consistent with these predictions. Because surprise/behavioural reorienting is present on both trial types but updating is present only on update trials, the relevant contrasts are (update trials + one-off trials) and (update trials - one-off trials). Maps are thresholded at p < 0.01 uncorrected to illustrate the similarities between maps for the different contrasts

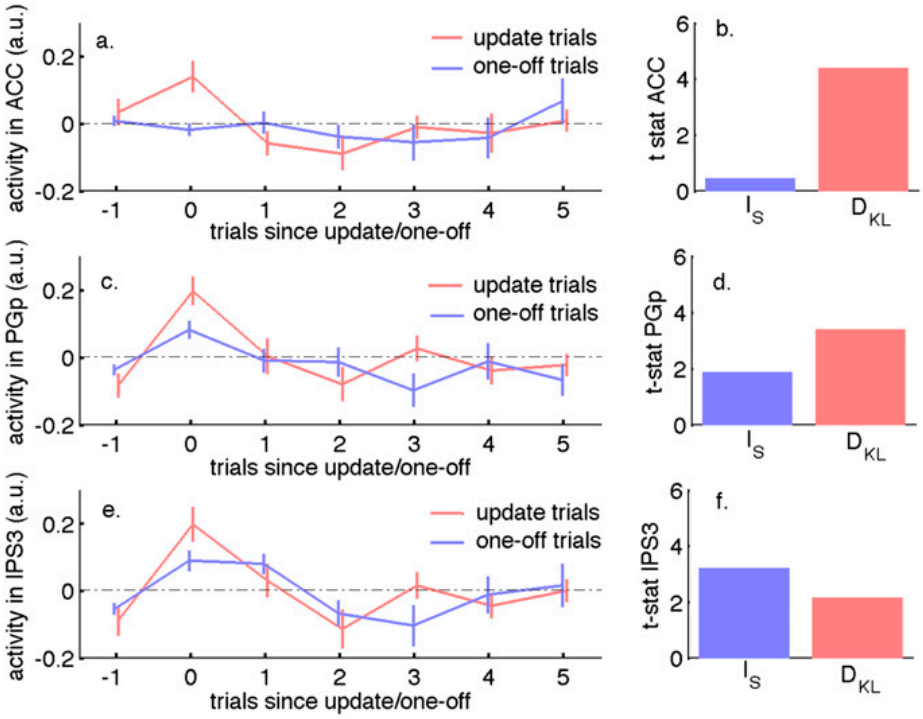

Figure 3.9 Activity in cingulate and parietal regions of interest (ROIs). (A, C, and E) Raw activity level on update and one-off trials and the trials around update and one-off trials. Trials are numbered so that the update or one-off trial is 0, the preceding trial is -1, and trials after the

update/one-off trial are numbered + 1, +2, etc. Note that in both parietal ROIs [parietal region PGp and interparietal sulcus 3 (IPS3)], activity on both update trials and one-off trials is above baseline (baseline is the average activity across expected trials). In the anterior cingulate cortex (ACC), activity is above baseline on update trials but not on one- off trials. (B, D, and F) Group t-statistics for I_S and D_{KL} in each ROI, from a GLM analysis of ROI activity. Note that there is a parametric effect of surprise/ behavioural reorienting (I_S) in both parietal ROIs but not in the ACC. In all three ROIs, there is an effect of updating (D_{KL}). Within the parietal cortex, IPS3 is more activated in relation to the parametric effect of surprise/behavioural reorienting, whereas PGp is more activated in relation to the parametric effect of updating. a.u., arbitrary units.

3.4 DISCUSSION

We observed specific neural signals associated with the updating of an internal model (D_{KL}) and the parametric effect of surprise (I_S). Updating was associated with activity in the rCMA zone of the ACC and ventral pre-SMA. On the other hand, activity in the posterior parietal region IPS3, a putative homolog of monkey saccade planning area LIP (Mars et al., 2011), was correlated with the parametric effect of surprise (I_S).

Activity has been reported in both the posterior parietal cortex and the ACC when stimuli are presented that would be surprising given participants' expectations (in other words, their internal models) of the environment (Molenberghs et al., 2007; R. Vandenberghe et al., 2001). Lesions in both regions have been linked to failures of reorienting (Molenberghs et al., 2008; Vandenberghe et al., 2012). However, the specific roles of the parietal and cingulate components of the "reorienting network" have so far been unclear (Mesulam et al., 2001). Furthermore, there has been little attempt to relate the role of the two regions in responding to surprising stimuli with their other reported functions, such as the role of the ACC in reward prediction (Jocham et al., 2009; Kennerley et al., 2006) and error monitoring (Debener et al., 2005; Holroyd et al., 2004). By identifying computational functions for the two regions, we hope the present results cast some light on this debate.

We observed activity in the IPS (area IPS3) that was parametrically related to the degree of surprise elicited by a stimulus (I_S); hence, stimuli that occurred in low-probability locations (given an optimal model of the environment) were associated with high levels of activity in the SPL. This activity can be understood in terms of the need to re-program planned eye movements in response to surprising target locations. Both surprise and posterior parietal activity were correlated with longer RTs, suggesting a link between behavioural reprogramming or spatial remapping (Molenberghs et al., 2007; R. Vandenberghe et al., 2001) and parietal activity. This remapping may be described as a within-trial effect because it is concerned with producing or reprogramming a behavioural response to the unexpected stimulus itself, whether or not predictions are updated for future trials.

3.4.1 Saccadic Reprogramming and Updating Expectations in Parietal Cortex

Although it is widely agreed that the posterior parietal cortex is active when either movements or covert attention are redirected or reoriented from one location to another (Hartwigsen et al., 2012; Mars et al., 2007; Mort et al., 2003; Rushworth et al., 2001a; Rushworth et al., 2001b; Thiel et al., 2004), there is disagreement about which regions within the parietal cortex are most important for redirecting attention and eye movements (Vandenberghe et al., 2012). One possibility is that different parietal regions are concerned with different aspects of redirection (Gillebert et al., 2013).

A closer analysis of anatomically defined ROIs in the parietal cortex suggests that the computational functions defined in this study may differentiate two regions

of the parietal cortex, both of which have been linked to saccadic planning. Region IPS3, a putative homolog of macaque region LIP, which has a generally motoric connection profile, was activated when surprise necessitated saccadic reprogramming. In contrast, posterior IPS/IPL region PGp, a putative homolog of macaque region 7a, was more strongly activated during updating. Human PGp and macaque area 7a are connected to the ACC, the main region activated during updating in the present study (Cavada & Goldman-Rakic, 1989a, 1989b).

This functional specialization within the parietal saccadic system is reminiscent of a long-standing observation from the neurological literature (Milner, 1996; Milner & Goodale, 1995) that lesions of the IPS and IPL produce different deficits in patients. Lesions of the IPS are associated with impairments in the ability to redirect the eyes, as in Balint syndrome (Pierrot-Deseilligny & Muri, 1996), or the covert focus of attention, as in extinction (Posner et al., 1984). In contrast, spatial neglect is associated with damage to the IPL (Milner, 1996; Mort et al., 2003). Neglect, unlike Balint syndrome and extinction, cannot be characterised as an inability to reorient to surprising events; rather, it seems to reflect a persistent inability to orient to the neglected field in the first place, which could be characterised as a distorted spatial prior.

3.4.2 ACC Is Specifically Involved in Updating

We observed activity in the ACC that was specific to trials on which the internal model was updated. ACC activity was not modulated by surprising stimuli that did not cause updating.

Although the ACC has previously been associated with the reorienting of attention and eye movements, and deficits therein, its specific role has remained unclear (Mesulam et al., 2001; Rushworth et al., 2009). At the same time, little in the way of theory has been offered to relate the role of the ACC in reorienting to the role it clearly has in error monitoring and reward prediction.

The current findings suggest a specific computational function for the ACC: It is involved in updating internal models to facilitate future information processing. This mechanistic theory of ACC function suggests a possible framework in which observations that the ACC is involved in reorienting and attention may be reconciled with the abundant evidence for it having a role in reward processing, learning, exploration, and foraging.

3.4.3 Errors and Updating in the ACC

A major line of research on the ACC has focused on observations that the ACC is active in situations when participants make errors. It has long been known that a negative-going potential, the error- or feedback-related negativity, is evoked from the region of the ACC (Debener et al., 2005) when participants make errors in psychological tasks or are presented with feedback on their performance (Holroyd et al., 2004). However, these findings are not incompatible with a role for the ACC in updating of internal models; indeed, it could be argued that the functional role of error- or feedback-related signals is to facilitate the updating of internal models from

which future action is generated (Rushworth & Behrens, 2008). This interpretation is supported by the findings that activity in ACC cells is particularly strong in instrumental tasks (Matsumoto et al., 2007) and, furthermore, that the magnitude of the error related negativity is proportional to the degree to which participants modify their behaviour on future trials (Cohen & Ranganath, 2007; Debener et al., 2005).

3.4.4 Exploration, Updating, and Estimation Uncertainty

A separate line of research has linked ACC activity to exploratory behaviour. The ACC is more active on experimental trials when participants explore alternative options rather than exploiting a known source of reward (Karlsson et al., 2012; Quilodran et al., 2008). Furthermore, the ACC is active during foraging, when participants decide to "forage" (seek alternative options) as opposed to choosing from among the options immediately presented to them (Kolling et al., 2012). Note that this role of the ACC in exploring and engaging in alternative actions is quite distinct from that of another frontal area, the orbitofrontal cortex, in updating the value or significance of specific stimuli (Nobre et al., 1999; Noonan et al., 2011).

Although exploration and foraging may appear to be only loosely related to learning and updating, there is a computational concept that relates the two functions: control of estimation uncertainty. Estimation uncertainty (Knight, 1921) is uncertainty about the parameters of the environment. It is distinct from expected uncertainty (or risk), which characterises the uncertainty or stochasticity that is inherent within a state of the environment and would persist even if the observer were certain about the parameters of the environment. In contrast, estimation

uncertainty is not inherent in the environment but reflects the agent's knowledge of the environment and can be reduced if the agent has the opportunity to make further observations of the environment. Computationally, the level of estimation uncertainty affects the ability of an internal model to learn (or be updated), because a model that is totally certain about the state of the world should not learn anything new (Courville et al., 2006; Yu & Dayan, 2005); this idea is embodied in learning theory by the concept of "associability" (Dayan et al., 2000).

Although estimation uncertainty is driven up by unexpected observations as in the present study (Courville et al., 2006), it can also be argued that estimation uncertainty should be elevated during foraging and exploration, because in exploring, the observer deliberately seeks new information; therefore, his neural networks should be in a state to accept new learning (Dayan, 2012).

The pattern of activity of the ACC in learning tasks could support an interpretation of its role in terms of controlling estimation uncertainty. In a one-armed bandit task (Behrens et al., 2007), it was observed that the ACC was more active in response to new observations when the environment was volatile (and estimation uncertainty was hence high). Furthermore, recent recordings from rat ACC (Karlsson et al., 2012) support the hypothesis that the ACC is active when estimation uncertainty should be elevated. In a two-armed bandit task, after the reward probability associated with different actions changed, there was a radical shift in the pattern of activity across ACC neurons; this shift occurred at the start of a period of exploratory behaviour rather than during the acquisition of new information per se, suggesting that the ACC was active at the point at which estimation uncertainty increased (when a model of the environment was abandoned

in favor of “knowing nothing”). This result is particularly closely related to the present study, in which participants were instructed to disregard their previous experience as soon as the onset of a new experimental run was signaled; this was represented in our model by the application of a uniform probability distribution over all possible target locations at the start of each run.

Chapter 4: Behavioural Effects of Reward and Spatial Surprise

In this chapter, I will show effects of spatial surprise and reward magnitude on the animals' performance. The following chapter 5 will then discuss the neural findings related to different aspects of reward and spatial surprise.

4.1 INTRODUCTION

The ability to predict events in space and time is of great evolutionary value. Therefore, all animals learn from past events and use learned information to make predictions about future events, as this helps organisms to adjust to their environment and choose differently and thus is key to survival. Accurate predictions of events also allow more efficient behavioural responses because the animal can prepare a future action in advance. If a currently held prediction or expectation turns out to be inaccurate or outdated, e.g. an event takes place in an unexpected location, the prepared action will be inappropriate and the action plan will need to be adjusted. The necessary adjustment normally comes with a reaction time cost of slowed responses, or if adjustment fails, incorrect responses that may remain in line with previous expectations rather than the new current situation.

In fact, it has been well established that responses to spatial behavioural cues are slowed down when those cues appear in an unexpected location (Corbetta & Shulman, 2002; Posner et al., 1980). It as been argued that this slowing might be due to the prepared (motor) response being discarded or refined which takes time (Posner et al., 1984). After reprogramming an action and responding to the cue in

the appropriate manner, the new spatial information can be used to form and refine future predictions. This constitutes an updating of spatial expectation (O'Reilly et al., 2013). Generally, a mismatch between expectation and actual cue position represents a spatial prediction error that triggers a learning process.

In addition to spatial information, other factors can influence performance and speed of a planned action as well. Animals also form expectations about the reward outcomes following actions. Prediction errors for outcomes (or rewards) affect how likely an animal is to repeat the same action and the speed with which it is executed and degree to which it is invigorated. Better than expected outcomes can increase motivation and worse ones can demotivate. In both cases there may be an impact on performance (Pessoa, 2009). Generally, It is assumed that reward facilitates and motivates, i.e. invigorates behaviour, especially in a learning context (Wise, 2004). This is supported by studies that found a performance improvement associated with reward in different behavioural domains like response speed accuracy (Bijleveld et al., 2010; Knutson et al., 2001) or visual search (Engelmann & Pessoa, 2007).

Here, we designed a study that looks at both, spatial updating and effects of different reward sizes on behaviour. While rewards were not contingent on option selection, i.e. only one target was presented on either the right or left side of the screen, rewards were omitted if the animal did not successfully touch the target stimulus. Furthermore, we manipulated the frequency of different reward magnitudes and the fluctuation of average reward rate over the course of a session. With those manipulations, we were able to look at the animals' behaviour when reacting to visually surprising stimuli and at effects on reaction times after the

animals experienced different reward magnitudes or find themselves in a 'rich' environment (currently high average reward rate) or not (currently low average reward rate). We expected the animals to be sensitive to the different reward magnitudes and also to form a spatial expectation about where a target was likely to appear in space.

We are using non-human primates for this study in order to be able to combine functional MRI with a direct brain manipulation method at later stage. The aim is to intervene in the normal information processing throughout different networks using locally precise lesions. This manipulation will allow for a causal investigation of brain area and cognitive function and will enable us to track changes in behaviour in a pre-post lesion comparison. Furthermore, through post-lesion fMRI, we would be able to identify changes in a network due to inactivation of a certain network node. In the absence of any behavioural effect, we might be able to see how the network compensates for inactivated nodes.

4.2 METHODS

4.2.1 Subjects

Six rhesus monkeys (Macaca mulatta, one female) participated in the experiment. The animals weighed between 8.6 and 13.5 kg and were 7-8 years of age. They were kept on a 12-h light dark cycle, with ad-lib access to water for 12-16h after testing and throughout the day on non-testing days. All procedures were conducted under licences from the United Kingdom (UK) Home Office in accordance with the UK Animals (Scientific Procedures) Act 1986.

4.2.2 Head post surgery

Prior to MRI data acquisition all animals underwent aseptic surgery to implant an MRI compatible head post. Before any surgeries, monkeys were treated with a steroidal anti-inflammatory [20mg/kg methylprednisolone injected intramuscularly (i.m.)] and an antibiotic (8.75 mg/kg amoxicillin, i.m.) a minimum of 12 hours prior to surgery so as to reduce the risk of postoperative infection, inflammation or edema. During surgery, extra steroidal supplements were provided at 4 to 6 hour intervals. On the morning of the surgery monkeys were sedated with ketamine (10 mg/kg, i.m.) and xylazine (0.5 mg/kg, i.m.) and were injected with non-steroidal anti-inflammatory agents (0.2 mg/kg meloxicam), atropine (0.05 mg/kg), and opioid (0.01 mg/kg buprenorphine). These were provided to reduce secretions and provide analgesia. They were further treated with a histamine H2 receptor antagonist (1 mg/kg ranitidine) for gastric ulceration protection due to the administration of both steroidal and non-steroidal anti-inflammatory treatments. Animals were then moved to the operating theatre where they were intubated, switched to sevoflurane inhalational anaesthesia and placed in a stereotactic frame. Their head was shaved and cleaned using alcohol and antimicrobial scrub. Throughout surgeries, respiration rate, heart rate, body temperature, blood pressure, and expired CO_2 and sevoflurane were continuously monitored.

To implant MRI compatible Polyether Ether Ketone (PEEK) head posts (Rogue Research, Montreal, Canada) a midline incision was made, the tissue was blunt-dissected in anatomical layers and retracted. The exposed bone was then prepared for the implant. 12-14 holes (3mm in diameter) were drilled in an ellipsoidal shape around the future position of the head post using a mechanical hand drill (Thomas

Recordings GmbH, Giessen, Germany). Ceramic screws (Thomas Recordings GmbH, Giessen, Germany) were then carefully screwed into the drill holes. A layer of dental cement (Lang Dental, Wheeling, IL, USA) was then applied to cover the screws and the head post was embedded in it. Finally dental cement was applied around the head post to create a smooth surface. The skin was then stitched in anatomical layers.

4.2.3 Behavioural training

After a post surgical recovery period of at least two months, the animals were trained to perform the task in an MRI compatible chair in the so-called sphinx position (See Figure 4.1A). The chair was placed inside a custom mock scanner that simulated the MRI scanning environment under head fixation. While exposed to recorded fMRI noise, the animals were trained to use a pair of infra-red touch sensors to respond to stimuli presented on a screen, approximately 30cm away from the eyes. The mock scanner training was considered complete once the animals performed the task with greater than 90% accuracy (See Experimental Task for details) for at least three consecutive training sessions. The animals were then moved to the actual MRI scanner for data acquisition.

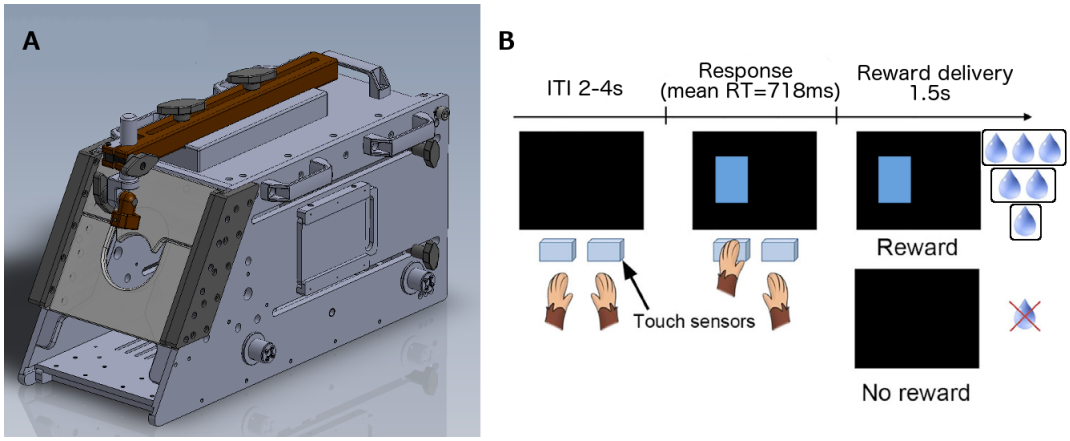

Figure 4.1 A) MRI chair. The animals were sitting in the so-called sphinx position in the MRI compatible chair. With the mounting bar and the head holder (brown) the animal's head post could be attached to the chair, thus helping the animal to keep the head immobile. Custom-made MRI coils were then placed around the animal's head and the reward tube was positioned into the animal's mouth. A box containing the response sensors was placed in front of the chair so that animals could easily reach both sensors. B) Task design. A trial began with a blank screen (ITI) followed by the presentation of the cue. After a correct response, the cue reappeared and the animals received one, two or three drops of juice. If responded incorrectly, the stimulus disappeared and no reward was given.

4.2.4 Experimental Task

We wanted to investigate the behavioural and neural effect of spatial and reward surprise. We therefore designed an experiment that enabled us to look at both surprising effects separately. On each trial, animals were presented with a single blue rectangle appearing on either the left or the right side of the screen. They received a reward after touching the response sensor on the corresponding side of the stimulus (See Figure 4.1B). Each trial began with a blank screen (2-4 seconds, mean 3 seconds) followed by a presentation of the rectangle. Each response was classified in either correct (touched the response sensor on the side of the stimulus appearance) or incorrect (touched the other sensor). Each correct response yielded a juice reward of one, two or three drops (~1.5ml each drop) after a delay of 200ms. The juice delivery took 1.5 seconds and after an inter-trial interval of 2-4 seconds (mean 3 seconds) the next trial began. Importantly, the spatial cue position (left or

right) varied independently of reward magnitude (one, two, or three drops) that was given for each correct trial. Reward magnitude was thus decorrelated from spatial position of the stimulus.

The side on which the stimulus was shown reversed after a sequence of 11-19 trials on the same side (mean 15 trials). Each sequence could be interrupted by one stimulus presentation on the opposing side (outlier) followed by a stimulus on the original side, thus the sequence would continue after such an outlier. 10% of trials were outlier trials with a maximum of one outlier per sequence.

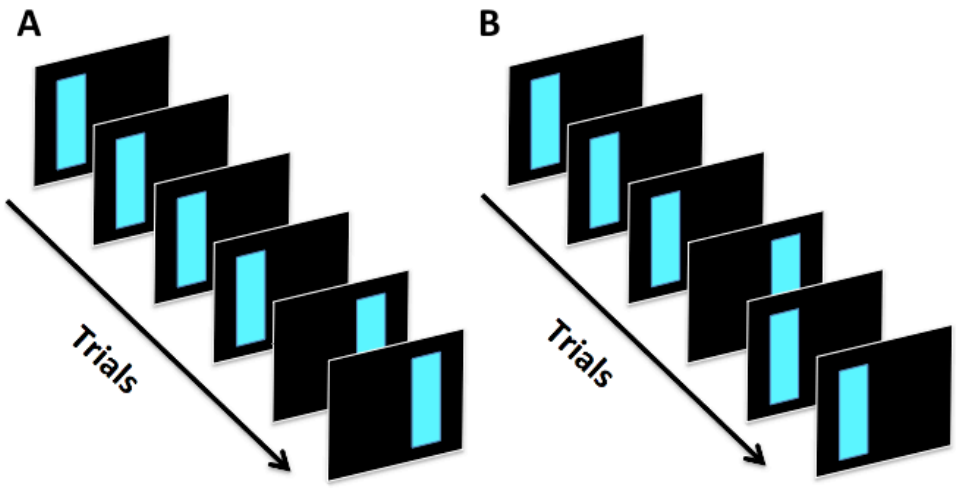

Figure 4.2 Trial sequence. A) Example of a reversal trial. Each screen represents one trial. After four trials, the side of stimulus presentation changes from left to right. B) Example of an outlier trial. In each trial sequence could be interrupted by an outlier trial which was followed by a stimulus on the original side.

12 sessions of 150 rewarded (i.e. correct) trials were collected for each animal. In six sessions, one and three drops were delivered randomly in 90% of the cases (45% for each reward size) and two drops in 10% of correct trials. Mean reward rate over a

session was kept at two drops (Figure 4.3). We will refer to these sessions as 'simple' sessions henceforth.

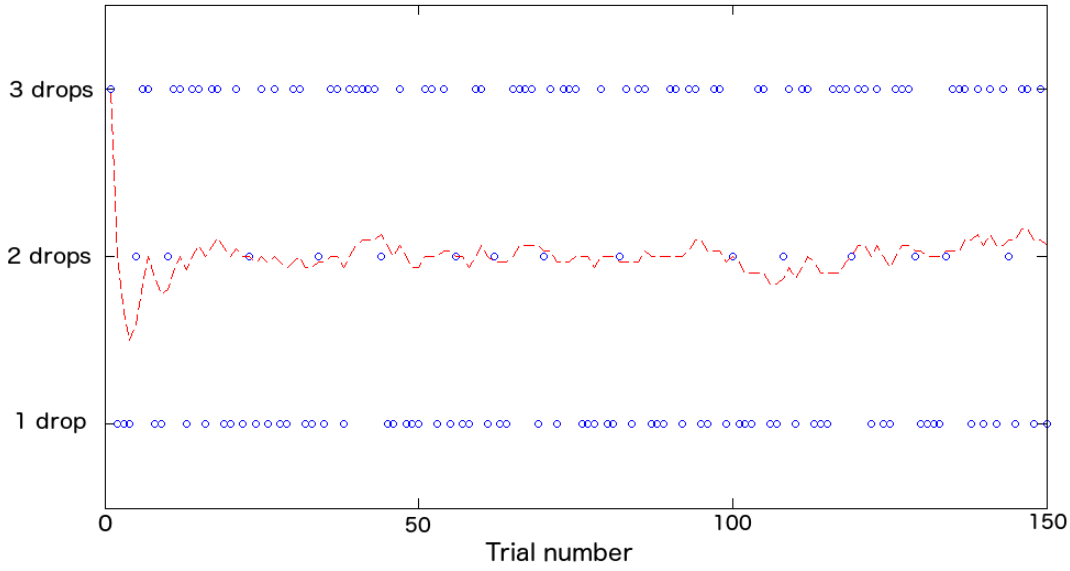

Figure 4.3 Reward schedule for one example session ('simple' session). Blue dots represent each reward for that session. The red dashed line represents the average reward rate as calculated with a sliding window (window size = 15 trials)

In addition to the six 'simple' sessions and in order to further investigate the effects of average reward rate we designed control reward schedules in which the reward rate varied over time. In two different reward schedules, the mean reward rate drifted either from 1.5 drops to 2.5 drops ('drift up' session) or in the opposite direction, i.e. from 2.5 drops down to 1.5 drops ('drift down' session). In these drifting sessions, the occurrence of one and three drops was kept at 45%. In a third control condition, we kept the average reward rate stable but each reward magnitude had a probability of 1/3 (Figure 4.4), therefore eradicating any reward frequency or identity effects ('equiprob' sessions). With the exception of the drift

sessions, the animals were not able to predict the next trials reward magnitude. We designed the reward schedules in such a way that reward magnitude occurred pseudorandomly based on its relative frequency, i.e. 45% for one and three drops, 10% for two drops in the simple sessions; 33% for all different magnitudes for the equal probability sessions. The drift sessions have a reward probability distribution of 75% three drops, 15% one drop, 10% two drops (drift down) and 75% one drop, 15% three drops, 10% two drops (drift up) respectively.

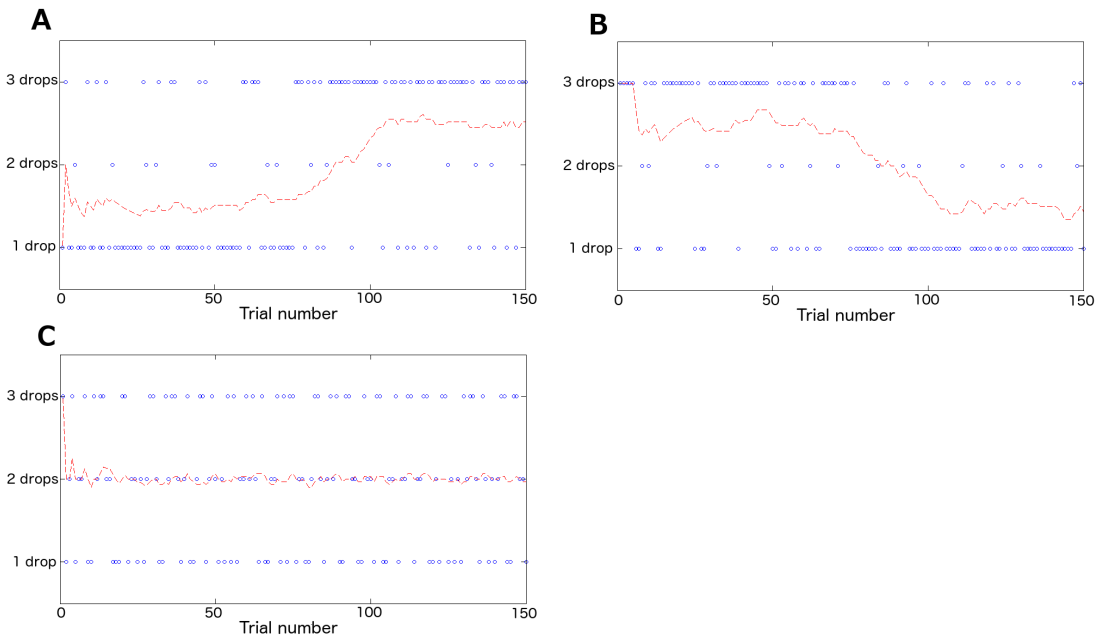

Figure 4.4 Example reward schedules (like Figure 4.3). A) + B) Reward schedules in which the average reward rate goes either up ('drift up' session, A) or down ('drift down' session, B). C) Reward schedule in which the average reward rate was kept at two drops but all three different reward magnitudes had a probability of 1/3 ('equiprob' session). The red dashed line represents the average reward rate as calculated with a sliding window (window size = 15 trials)

4.2.5 Behavioural analysis

In order to evaluate the effects of spatial and reward surprise on behaviour, we took reaction times (RTs) as a measure of the animals' performance. A hard cut-off of 4 seconds was applied to the raw RTs before removing all responses that took longer

than 2.5 standard deviations from the RT distribution of an individual session. The hard cut-off discards all responses that are slower than a 'reasonable' amount of time, e.g. when the animal got distracted or even fell asleep. I acknowledge that the amount of four seconds is somewhat arbitrary but it reflects the subjective impression of the maximum time an animal took to respond while focused on the task. Then, medians of RTs of each session were taken separately and averaged on a higher level. We expected a spatially surprising event to be associated with a longer RT on that trial. We therefore averaged all trials immediately preceding spatially surprising events (t-1), spatially surprising trials themselves (t) and the trials following the spatially surprising event (t+1, t+2, t+3) separately. Note that all behavioural effects of reward are referring to the subsequent trial, i.e. after an animal has experienced a certain reward. For example, the regressor 'last reward' refers to effects that the reward size from one trial ago has on the reaction time of the current trial. On the other hand, spatial effects are referring to the current trial, i.e. reaction time effects are analysed based on the level of spatial surprise of the trial the animals are about to perform.

We also looked to see if surprising reward events caused RT differences. Since the reward delivery occurred after the stimulus presentation and animals' responses were made, we looked at effects of reward surprise on the trial following the reward delivery. We binned the data into trials that were preceded by a single occurrence of a certain reward magnitude (one, two or three drops of juice) and trials that were preceded by at least one repetition of a certain reward magnitude.

To get a better understanding of which components of the task had an effect on behaviour, we fitted two regression models to the RTs of each session (for correct

trials only). GLM1 contained the following regressors: Spatial position (coded as 1 when stimulus was on the right and -1 when on the left), spatial surprise (coded as 1 for all surprising stimuli, i.e. reversals and outliers and 0 otherwise), reward size (coded as 3 for three drops, 2 for two drops and 1 for one drop), two drop reward (coded as 1 for every instance of a two-drop-reward; these trials are potentially of particular interest because they are unsurprising in the sense that the reward level corresponds to the average reward rate but, at least in the simple and drift sessions, but they are surprising in the sense that a reward of that precise magnitude rarely occurs), last reward (coded as 3, 2 and 1 according to the reward magnitude of the previous reward), trial number (coded as 1 to n where n is number of trials in a session), and a constant.

We designed a second GLM in order to evaluate the effect of repetition of a certain reward magnitude. GLM2 contained spatial position, spatial surprise, two drop reward, trial number, a constant (all as in GLM1), and in addition reward repetition (coded as 1 whenever there was a repetition of a certain reward magnitude) and reward size (coded as 3,2 and 1 for each reward but only if the previous reward was different from the current one, see Table 1).

All regressors (for GLM1 and GLM2) were z-score normalised. A regression analysis was done using Matlab's glmfit function on the distribution of RTs. In order to test whether on average the coefficient estimates are different from 0, we used t-tests on each of the beta weight distributions.

Table 1: List of regressors for behavioural GLMs 1 and 2

GLM1 regressors	GLM2 regressors
Constant	Constant
Spatial position (left or right)	Spatial position (left or right)
Spatial surprise	Spatial surprise
Reward Size	Reward size without repetition
Two drop reward	Reward size repetition
Last reward	Two drop reward
Trial number	Trial number

We checked the degree of correlation of the different regressors in both GLMs to make sure we could be able to identify separable neural correlates for our variables of interest. Figure 5 and Figure 7 shows the mean value of |r| over sessions in every animal. Figure 6 and Figure 8 show the mean level of correlation between regressors for each animal and the highest level of correlation between regressors for each animal in a given session. For GLM1, the maximal absolute mean correlation is $|r|<0.23$ in any animal (Fig.5, highest:$|r| = 0.221$ for reward size and trial number in Pat). The highest correlation in a single session is $|r|<0.52$ (Fig. 6, highest: $|r| = 0.514$ for reward size and trial number in Pat).

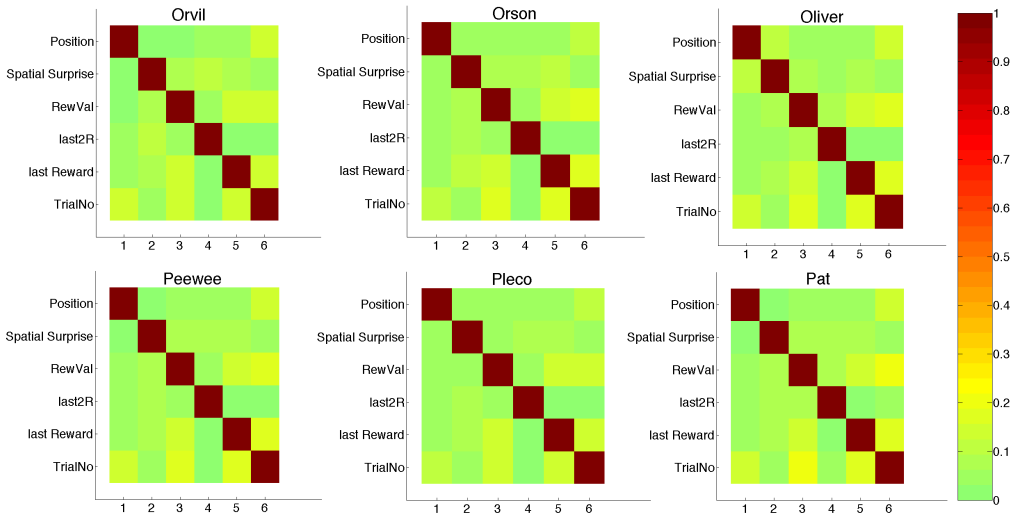

Figure 4.5 Mean level of correlation between regressors used in GLM1, plotted for each animal separately (mean absolute correlation)

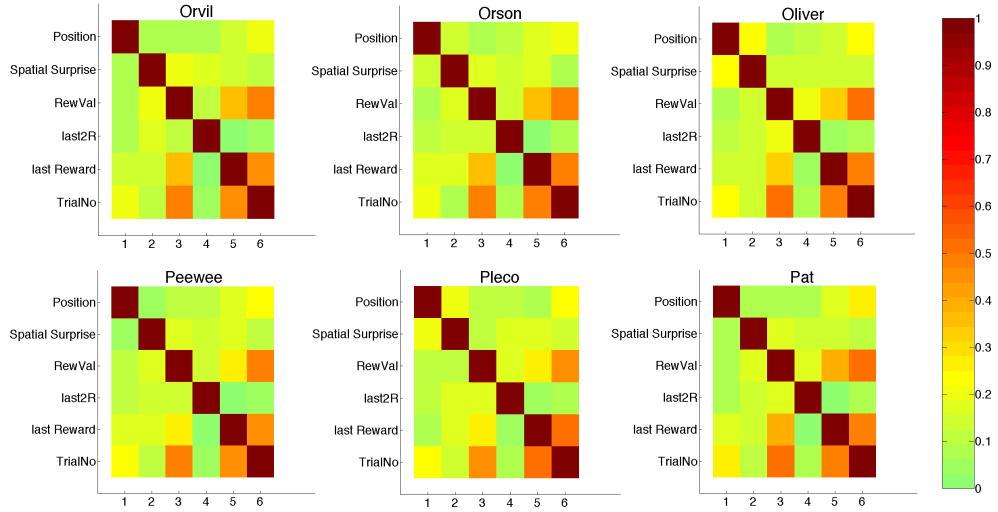

Figure 4.6 Highest level of correlation between regressors for a given session in each animals for GLM1

For GLM2, the maximal absolute mean correlation is |r|<0.23 for all animals (Fig. 7, highest:|r| = 0.228 for reward repetition and trial number in Pat). The highest correlation in a single session is |r|<0.59 (Fig. 8, highest: |r| = 0.589 for reward repetition and trial number in Pat).

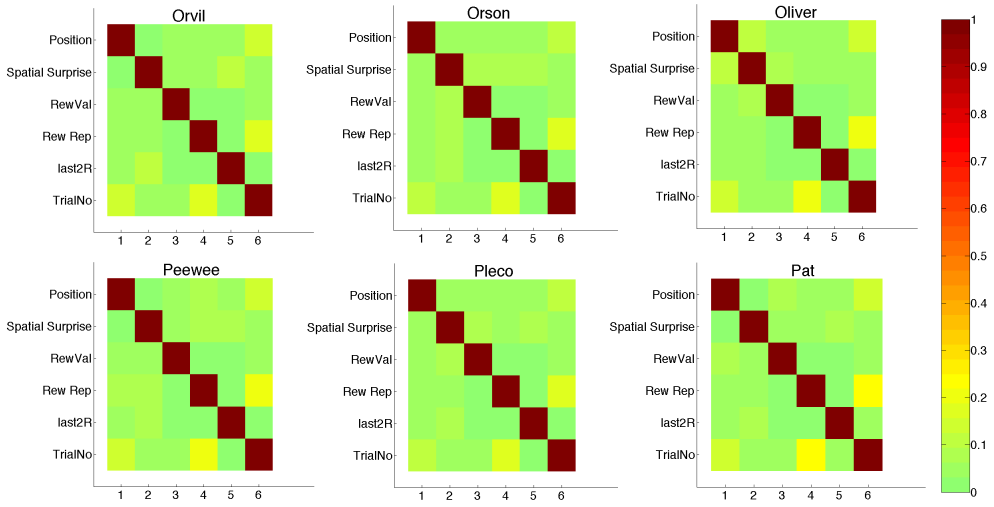

Figure 4.7 Mean level of correlation between regressors used in GLM2 (mean absolute correlation) for each animal

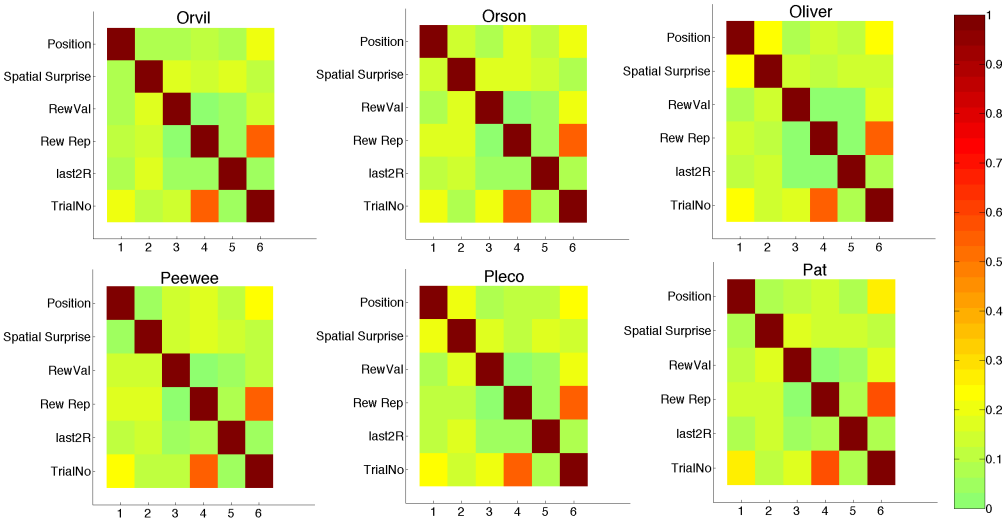

Figure 4.8 Highest level of correlation between regressors for a given session in each animals for GLM2.

4.3 RESULTS

4.3.1 Binned reaction time analysis of reward effect

The animals' accuracy was high (95% correct on average; mean error rate: 5%, SEM=0.85%) and there was no significant difference in errors over the course of a session (errors in the first half compared to second half errors of a session: $t_{(70)} = 1.2$, p=0.23, paired t-test). Mean raw response time (RT) over all animals and sessions was 3.46s (SEM: 417ms). After removing RTs from very long outlier responses, we analysed an RT distribution with a mean of 718ms (SEM: 19.6ms).

We analysed the effects of reward magnitude and spatial location on response times. Reward related behavioural effects were observed following single occurrences of one, two, and three drops of reward respectively. Animals were faster after they received only one drop and slowest after a large reward (Figure 4.9). This difference, however, disappeared when a reward type was repeated. There is no significant difference between RTs on the trial following a repetition of either one or three drops (all effects are statistically tested in the GLM2 later).

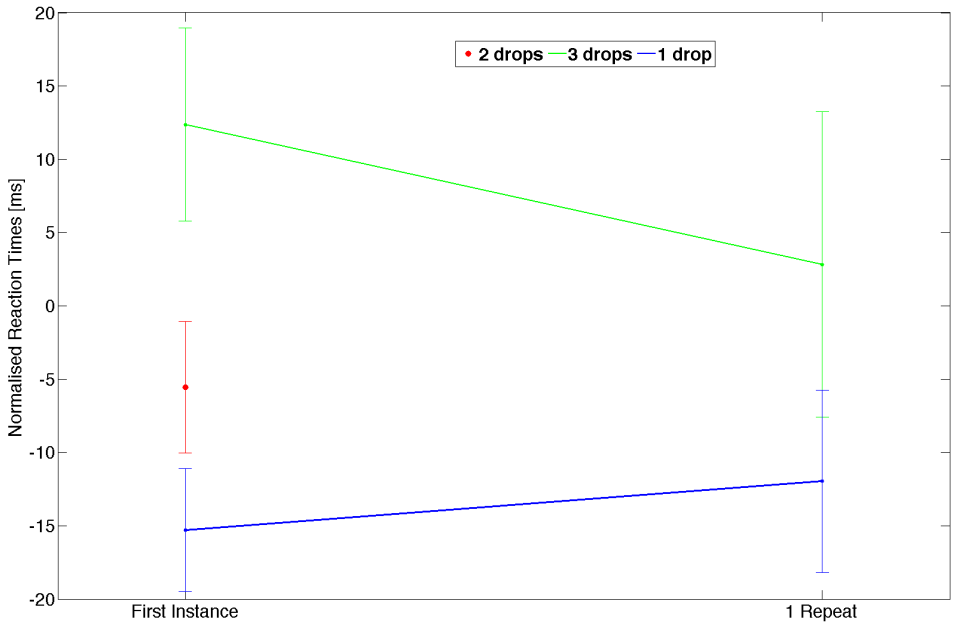

Figure 4.9 Mean effect of reward on RTs. If animals experience a certain reward magnitude only once, they behave differently on the following trial. After receiving only one drop, animals tend to speed up their response while a large reward of three drops slows them down. If the same reward magnitude is given on the following trial, this difference in RT seems to disappear. Error bars: SEM.

4.3.2 Binned reaction time analysis of spatial effect

Spatially surprising trials were divided into outlier and reversal trials. Reversal trials started a sequence of stimuli that appeared on the same side. That sequence could be interrupted by outlier trials. Both, reversal and outlier trials (trial *t*) are associated with a significant increase in RT in comparison with the preceding trial (paired sample t-test, p=0.0021 and p=0.0051, respectively, Figure 4.10). However, there was no difference in RT between the two types of spatially surprising stimuli: reversal and outlier trials (paired sample t-test, p=0.11). While RTs following a reversal trial returned to baseline levels (no difference between t-1 and t+1 for reversal trials (paired sample t-test, p=0.77)), RTs following an outlier remained slow; although they were not slower than on the outlier trial itself (paired sample t-

test, p=0.2) they were significantly slower than trials following a reversal trial (paired sample t-test, p=0.026) and slower than trials preceding an outlier trial (paired sample t-test, p=0.042). On trial *t+2* reaction times in both cases were close to baseline again.

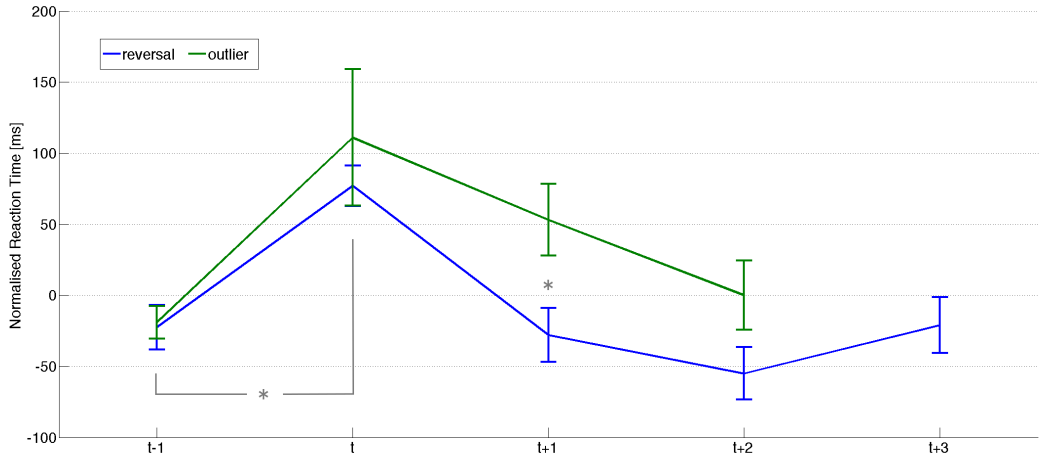

Figure 4.10 Effect of reversal trials and outlier trials on RTs. RTs are plotted against surprising trials *t* (reversal in blue and outlier trials in green) and the surrounding trials. Asterisks indicate significant differences, error bars: SEM.

4.3.3 Multiple Regression based reaction time analysis I

Two GLMs were used to further analyse behaviour (GLM1 and GLM2 described in Methods). The first included the following regressors: Constant (not shown), spatial position, spatial surprise, reward size, two drop reward, last reward and trial number (Figure 4.11). Spatial surprise has a beta weight significantly larger from zero (one-sample t-test, p<0.001), i.e. spatially surprising stimuli were accompanied by an RT cost. Furthermore, the positive beta weights for trial number indicates that the animals are slowing down during the course of the experiment (one-sample t-test,

p<0.001). As shown in Figure 4.11, there is a positive beta weight for reward size. This shows that animals are significantly slower after larger reward sizes (one-sample t-test, p<0.005). The reward from two trials ago, however, seems to be associated with a slight speeding up of animals' responses (one-sample t-test, marginally significant: p=0.0564).

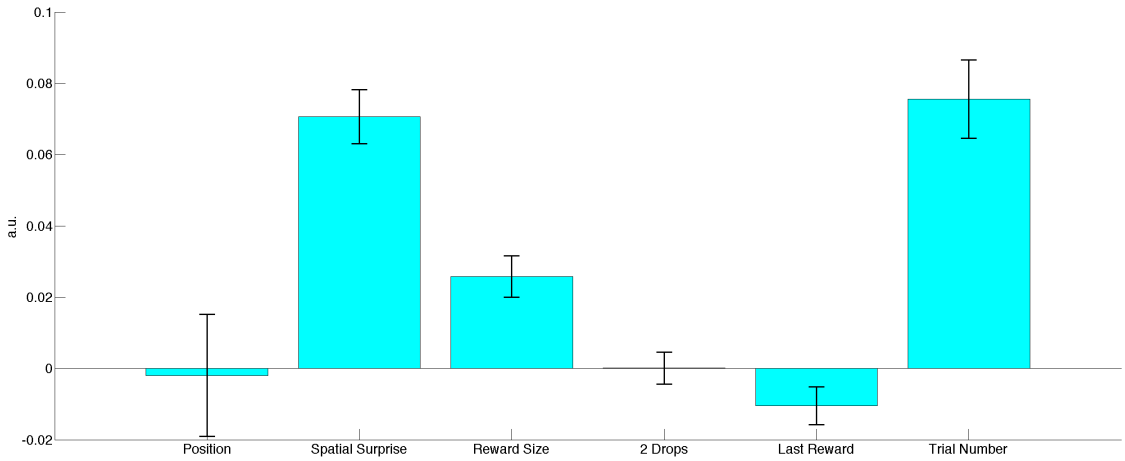

Figure 4.11 Behavioural GLM1. Bars indicate mean beta weights for each regressor across all animals and sessions. Error bars: SEM. Note that 'Spatial Surprise' and 'Position' refer to the current trial whereas reward related regressors refer to the reward experiences in the previous trial.

4.3.4 Multiple Regression based reaction time analysis II

The second GLM resembled GLM1 but instead of including a reward size regressor it included a regressor coding for whenever a certain reward magnitude was repeated and another regressor coding for whenever a certain reward magnitude appeared for the first time. The beta weights for the reward repetition regressor were not

different from 0 (one-sample t-test, p=0.72) but there was a significant difference in effect size between the reward size repetition and the reward size first instance regressor (paired sample t-test, p=0.01, Figure 4.12) as illustrated in Figure 4.9.

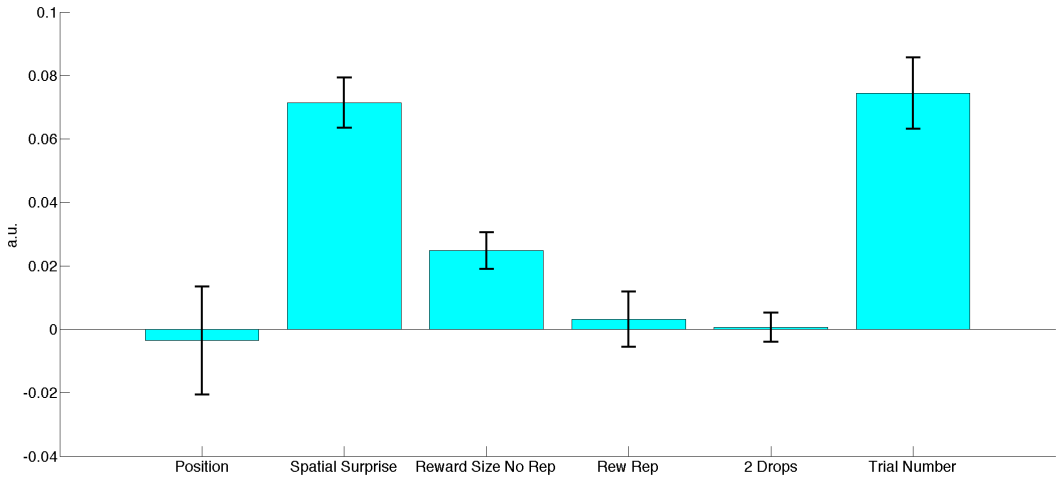

Figure 4.12 Behavioural GLM2. Bars indicate mean beta weights for each regressor across all animals and sessions. Error bars SEM

4.4 DISCUSSION

Our results show that the animals noticed spatially surprising events and used them to inform and update new spatial expectations after both reversal and outlier trials. Spatial surprise leads to at least two distinguishable processes. First, it triggers an attentional orienting process in which attention is reallocated in space towards the unexpected event or stimulus (Posner et al., 1980). Second, it can trigger a learning process that updates one's prior expectation about where in space events are likely to happen (See chapter 2 and 3 and O'Reilly et al., 2013).

In Chapter 2 and 3 I used a saccade paradigm where humans were asked to make an eye movement towards a dot on the screen. The location of the target

could be predicted to a certain degree and thus participants could prepare for a saccade. Every so often, however, following a cue, participants had to learn a new potential target location. We could replicate some of the results obtained in humans in chapters 2 and 3 with the monkey studied here in chapter 4. We found no behavioural difference in RTs for reversal (update) versus outlier trials. Note, however, that we explicitly cued outlier trials for the humans, but not for the animals. This simple task is likely to explain the difference in updating behavior on outlier trials; it was simply not obvious that outlier trials were not reversal trials in the current paradigm. The monkeys in the current study appeared to update their expectations about future stimulus positions whenever they saw any surprising visual stimulus. This means that on the trial after an update trial their RTs returned to low levels similar to those prior to the update cue. By contrast, there was a further RT cost on trials following outlier trials (t+1). This suggests that animals had indeed updated their visuospatial expectation after every visually surprising event and so had to update their visuospatial expectations once again after t+1 stimulus following the outlier trial (Figure 4.10).

Analysis of RTs after different reward magnitudes showed that the animals were able to distinguish between low (one drop), middle (two drops), and high rewards (three drops). Reward anticipation, which is dependent on past rewards, is known to have a facilitating effect on behaviour (e.g. Bijleveld et al., 2010; Engelmann & Pessoa, 2007; Knutson et al., 2001) as rewarding events are motivating and thus can produce an increase in mental or physical effort (Schmidt 2012). No reward, relatively low reward, or punishment can trigger the opposite behaviour; lower motivation impairs performance (e.g. Engelmann & Pessoa, 2007). However,

by contrast, in our study, we found an immediate decrease in performance (longer RTs) following a single instance of high reward (three drops) as opposed to a relatively low reward (one drop). This difference diminishes when high or low rewards are repeated (Figure 4.9 and Figure 4.12).

One interpretation is that surprising 3 drop events might capture an animals attention more and lead to longer consumatory behavior, slowing down the response to the next stimulus. Furthermore, it could reflect a process where the animals are trying to figure out the 'richness' of their current environment. After the first instance of a high reward, animals might pay more attention to the subsequent trial which can lead to a responsive slowing. Experiencing a low reward might increase the animals' impatience and which could make them respond more quickly on the subsequent trial in order to get more reward more promptly. This difference disappears if small and large rewards are repeated. This could reflect a reversal of the attention/motivation effect that we observed after the first instance. If animals receive multiple low rewards in succession, they might consider themselves being in a 'low reward' environment which decreases their motivation. The opposite might happen after a succession of high rewards and animals' performance might increase.

Trial number also had a positive correlation with RT which means that animals responded slower later in sessions. The fact that session length influences an animal's performance has been previously reported (San-Galli et al., 2016; Stoll et al., 2016). Recently, it has been shown that beta oscillation power in frontal areas correlates with the likelihood of the animal having a spontaneous break but beta power 'resets' after the break. Such a pattern suggests that beta power is a marker for attentional effort rather than fatigue or saturation (Stoll et al., 2016), however

their data does not allow for such a clear distinction as these concepts were correlated in their study. The firing rate of vmPFC neurons also encodes the willingness to engage and predicts when animals are about to take a break (San-Galli et al., 2016). In comparison to experimental sessions in the studies reported by Stoll and colleagues and San-Galli and colleagues, our sessions were relatively short (150 rewarded trials, 28 minutes on average) and the animals took breaks only very infrequently. Furthermore, given that the error rate is so low, there was no difference between error rates in the first half of a session compared to the second half. Therefore it is possible that the increase in RT over the course of a session is due to saturation or fatique. However, in the next chapter, I will discuss neural correlates of trial number which we found using fMRI. In short, I found evidence of changing levels of activity in a circumscribed neural network as a function of session length.

Chapter 5: Neural Correlates of Reward and Spatial Surprise in the Macaque Brain

5.1 INTRODUCTION

In the last chapter, I demonstrated the effect of both reward size as well as spatial surprise on behaviour. In short, the animals showed a significant reaction time (RT) cost when stimuli occurred at unexpected locations, while the size and frequency of different reward sizes affected the animals' RTs in several ways. Overall, this showed that the animals were clearly noticing and tracking different reward sizes as well as trying to predict the location of the next visual stimulus.

However, while the behavioural effects of reward and surprise have been very well described in the previous chapter here we describe the neural networks that mediate the impact of these factors on behaviour. Even though there have been a number of electrophysiological studies examining reward related signals at the neuronal level in NHPs conducted in a number of areas including orbitofrontal, premotor and prefrontal cortex, striatum, amygdala, and dopamine neurons (Amador et al., 2000; Apicella et al., 1991; Bowman et al., 1996; Hikosaka et al., 1989; Ljungberg et al., 1992; Markowitsch & Pritzel, 1976; Nakamura et al., 1992; Nishijo et al., 1988; Pratt & Mizumori, 1998; Ravel et al., 1999; Schultz & Tremblay, 1999; Shidara et al., 1998; Thorpe et al., 1983) those that have focused more specifically on reward prediction errors and have been conducted in a limited

number of anatomical regions most notably the dopaminergic midbrain and striatum (Schultz, 2006) and areas interconnected with these regions such as the border region of the globus pallidus and the lateral habenula (Hong & Hikosaka, 2008, 2013, Matsumoto & Hikosaka, 2007, 2009); It is not clear if similar signals exist elsewhere. A number of experiments have been conducted on reward networks in the human brain (O'Doherty, 2016a; O'Doherty et al., 2016; Rangel et al., 2008; Rushworth et al., 2011; Schmidt et al., 2012) and although these have often had the virtue of examining activity across the whole brain, very few of them have examined neural responses to primary rewards for which human subjects are strongly motivated to work. Moreover, few of these studies directly test whether activity is better described as reflecting a response to reward delivery per se, reward expectation, reward prediction error (RPE) or some other function of current and past rewards. This means we do not have a complete picture which areas exhibit which type of reward signals.

Interestingly, there might exist signals that go beyond the canonical RPE. The canonical RPE is the breach of an expectation about the size or amount of reward. It is sometimes described as a scalar prediction error (Niv & Schoenbaum, 2008). However, the reward expectations that humans and other animals hold are not just about the amount of reward that they might obtain but concern the identity of the reward that they might obtain (Tinklepaugh, 1928; Watanabe, 1996). Contrary to other studies investigating reward identity by using different types of rewards, we made use of the relative frequency (or infrequency) of the two-drop reward. Given that high and low rewards are equally frequent and that two drops are comparatively rare but yet match the average reward expectation, we look at the

two-drop rewards as unsigned RPEs or reward identity surprise. We have known for some time that dopamine very likely plays a critical role in prediction error based learning (Schultz, 2006). We designed our task so that we could map different kinds of reward change, such as reward identity change in addition to change in average expected value, and compare their impact on putatively dopaminergic and dopaminoceptive regions such as the ventral tegmental area (VTA) and ventral striatum (vSTR).

A reward prediction error response is composed of three components. First, there is a response to reward delivery but in addition there are also responses that are functions of the reward expectation. These occur both prior to reward delivery and at the time of reward delivery. The reward expectation-related response at the time of reward delivery is opposite in sign to the reward delivery-related response. For example the prediction error response in the dopaminergic midbrain, particularly in VTA, comprises a positive response to the reward delivery itself and a negative response to the prior reward expectation at the time of reward delivery (Schultz review). Human neuroimaging studies have suggested a similar prediction error signal may also exist in a brain area to which the dopaminergic midbrain strongly projects: the striatum (O'Doherty, 2016; O'Doherty et al., 2016; Rangel et al., 2008). In the lateral habenula, a region from which the dopaminergic midbrain receives anatomical projections, there is a prediction error response but both of the component parts of the response at the time of reward delivery are flipped in sign (Matsumoto & Hikosaka, 2007, 2009). The reward delivery related response is negative and the response to the prior reward expectation that is recorded at the time of reward delivery is positive. In our experiment it was possible to dissociate

the reward delivery per se and the reward prediction error (reward delivery given the context of prior reward expectation).

However, other brain regions might also have activity that not only reflects reward delivery in a positive way, but also activity which reflects prior reward expectations in a positive way. Such a brain region might be tracking average reward rates, in order to control the animal' s vigour, a feature that has already been related to tonic dopamine in the striatum (Niv et al., 2007). Note that the relationship between dopamine release (tonic or phasic) and the BOLD signal is not clear and parallels should therefore be drawn with caution. Behaviour may be faster and more vigorous when the average reward rate in the environment is higher because slower and less effective behaviour is associated with more opportunity costs when there are more opportunities available in the environment. In addition, recently accumulated reward resources may also determine the degree of risk related behaviour; risk-taking can be adaptive when reward resources are low (Kolling et al., 2014).

Beyond reward, other events can be surprising and therefore relevant for any system which tracks expectation violation i.e. a more generalized prediction error. Therefore, it would be useful to be able to compare reward prediction errors with, for example, responses to spatially surprising stimuli. Typically such stimuli are thought to trigger a response in attentional reorienting networks but because direct comparisons have not been made between surprising visual events and surprising reward events it is not clear if both types of expectation violation trigger similar responses. In the current experiment we also test whether there are parts of such a system that might also be engaged when surprisingly rewarding outcomes appear.

In our analyses we break down the reward history into both reward events that are surprising in the short term and changes in reward value that occur over the longer term as animals learn the long term average reward rate in an environment. Furthermore, we can also investigate effects related to animals becoming satiated and less inclined to engage in the task over time.

To look more specifically at learning relevant history effects we generated sessions with schedules in which we kept the reward rate stable or let the animals experience a drift in average reward rate either up or downward, by changing the relative frequency of large vs. small reward sizes. We reasoned that when rates drifted, the current reward rate had to be tracked over time, potentially making animals pay more attention to the reward history. Additionally, as mentioned in the last chapter, we changed the frequency of rewards of different sizes. In most sessions rewards matching the overall average reward rate would be presented very rarely, enabling us to look at signals relating to rare reward events beyond a classic RPE. Specifically, in all but the control sessions in which all reward events were equally likely, two drops would only occur very rarely, while most times animals were given either one or three drops. Furthermore, we manipulated spatial surprise by simply varying the location on the screen in which the target stimulus appeared. Stimuli would appear mostly either on the left or right side for a whole block and the side of stimulus presentation would change between blocks. Since rewards and locations of stimuli were decorrelated in each session, we were able to look at reward and spatial surprise signals independently from each other.

5.2 METHODS

5.2.1 Experimental Task

For this study, we used the same task design as in Chapter 3. All six animals performed 12 sessions each while we recorded their brain activity using fMRI. One session had to be discarded because of poor image quality, yielding a total of 71 scanning sessions.

5.2.2 Imaging Data Acquisition

Imaging data were collected using a 3T clinical MRI scanner and a four-channel phased-array receive coil in conjunction with a radial transmission coil (Windmiller Kolster Scientific, Fresno, CA). FMRI images and reference images for artefact corrections were collected while awake animals were head-fixed in a sphinx position in an MRI compatible chair. FMRI data were acquired using a gradient-echo T2* echo planar imaging (EPI) sequence with 38 horizontal slices, 1.5×1.5×1.5mm^3 voxel resolution, TR=2.28s, TE=30ms, flip angle=90°. Data were taken as accelerated images in GRAPPA mode with foot-head phase-encoding direction and an acceleration factor (IPAT) of 2. In order to avoid motion artifacts that may result from body motion of the subjects during scanning, the resulting EPI raw data were exported from the scanner and reconstructed using an offline SENSE reconstruction (sensitivity encoding; (Pruessmann et al., 1999)). Proton-density-weighted images using a gradient refocused echo (GRE) sequence (TR=10ms, TE=2.52ms, flip angle=25°) were acquired as reference for body motion artefact correction.

T1-weighted MP-RAGE images (no slice gap; 0.5 x 0.5 x 0.5 mm; TR = 2500 ms; TE = 4.01 ms; 128 slices) were acquired in separate anaesthetised scanning sessions. In those sessions anaesthesia was induced using intramuscular injection

of ketamine (10 mg/kg) combined with either xylazine (0.125–0.25 mg/kg) or midazolam (0.1 mg/kg) and buprenorphine (0.01 mg/kg). Macaques also received injections of atropine (0.05 mg/kg, i.m.), meloxicam (0.2 mg/kg, i.v.), and ranitidine (0.05 mg/kg, i.v.). Local anesthetic (5% lidocaine/prilocaine cream and 2.5% bupivacaine injected subcutaneously around the ears to block peripheral nerve stimulation) was also used at least 15 min before placing the macaque in the stereotaxic frame. Anaesthesia was maintained with isoflurane. The anaesthetised animals were placed in an MRI-compatible stereotactic frame (Crist Instrument, Hagerstown, MD) in a sphinx position within a horizontal 3T MRI scanner with a full-size bore (see also Sallet et al., 2011). A total 5 five MPRAGE sequences was acquired and then averaged using FMRIB's Software Library (FSL; (Smith et al., 2004).

5.2.3 FMRI data preprocessing

FMRI data were corrected for body motion artefact by an offline-SENSE reconstruction method (Kolster et al., 2009) (Offline_SENSE GUI, Windmiller Kolster Scientific, Fresno, CA). Image alignment was carried out on two levels. First, a local correction was applied to each EPI image, which accounts for the dynamic differences between images within a session, mostly due to body motion. This is followed by a second, global correction, which accounts for static differences between the sessions due to differences in B0 shimming and positioning inside the scanner. For this purpose, a custom motion-correction was employed (Kolster et al., 2014) to correct the raw EPI images with respect to an EPI reference image chosen from the same session. The correction was applied as a 2D distortion field of the grid

points in the horizontal EPI images and the corrected image values were calculated by interpolation based on the initial image values and the distortion field. Each slice was corrected independently based on the target EPI volume, which was selected from a run within the same session after visual inspection to avoid slanted positions of the odd and even slices. Next, a global correction was applied, in which the GRE reference volumes of each session were aligned to the T1 volumes using a rigid-body, six-parameter alignment tool (www.nitrc.org/projects/jip; (Mandeville et al., 2011). The resulting alignment was applied to the EPI volumes, as these are intrinsically aligned to the GRE volumes. Session EPI volumes were then registered to the T1 volumes using 2D distortion fields in the horizontal imaging plane (www.nitrc.org/projects/jip). The data processing codes for the local and global corrections (Align_EPI GUI and Align_Anatomy GUI, Windmiller Kolster Scientific, Fresno, CA) were written in MATLAB (The Mathworks, Inc.). The aligned data were processed with high-pass temporal filtering (3-dB cut-off of 100s) and Gaussian spatial smoothing (full-width half maximum of 3mm). 4D volumes were intensity normalised and volumes with a higher signal variance than 2.5 std off the mean variance were modelled as confound volumes. The data that were already registered to subject's structural space was registered to the CARET macaque F99 template (Van Essen, 2004) using affine transformation.

5.2.4 FMRI data analysis

Whole-brain analysis was conducted using a univariate General Linear Model (GLM) approach. 17 Regressors were designed (Table 2). All regressors were z-score normalised and the data was analysed using FSL's FEAT (FMRI Expert Analysis Tool).

To model the haemodynamic response function, each regressor was convolved with a single Gamma function (mean lag = 3s, standard deviation = 1.5s).

Analyses were first conducted at the individual session level. Average effects of the GLM across all 71 sessions were calculated using FMRIB's local analysis of mixed effects stage 1 and 2 (FLAME1+2) (Beckmann et al., 2003; Woolrich et al., 2004) and using standard cluster-based thresholding criteria of $z>2.3$ and $p<0.05$ cluster-corrected (Worsley et al., 1992). The approach that we used treats sessions as an independent event. This strategy is appropriate if the variance across sessions performed by a given animal is no different to the variance across sessions performed by different animals.

Table 2 List of regressors used in whole brain GLM

Regressor	Definition
Constant	n/a
Spatial repeat	Regressor value = 1 whenever a spatial position is repeated
Spatial repeat number	Regressor value contains the amount of spatial repeats. Resets when spatially surprising event happens
Spatial surprise	Spatially surprising trials as identified using RL model (α=0.9)
Spatial surprise plus one	Trial following spatially surprising trial
Spatial surprise minus one	Trial preceding spatially surprising trial
Spatial surprise plus one parametric	Regressor value = 1 for trial following a reversal; Regressor value = -1 for trial following an outlier
Reward size	Normalised reward size value
Last reward	Normalised reward size value of trial t-1
Average reward rate	Average reward rate as identified using a RL model on all trials minus the last two (α=0.05)
Two drop reward	Regressor value = 1 for every time two drops were given
Reward repeats	Regressor value = 1 for whenever a

	certain reward magnitude was repeated
Right hand response	Regressor value = 1 for every right hand response
Left hand response	Regressor value = 1 for every left hand response
Right hand response (unconvolved)	Regressor of no interest. Included to account for motion induced artefacts. No HRF convolution, event duration = 1TR (2.28 s)
Left hand response (unconvolved)	Regressor of no interest. Included to account for motion induced artefacts. No HRF convolution, event duration = 1TR (2.28 s)
Main effect reward	Regressor value = 1 for every time a reward is given

5.2.5 ROI analysis

We looked at the whole brain effects of the main effects of two regressors: Reward Size and Spatial Surprise. To illustrate reward related activity in vSTR, vmPFC, VTA and OFC, we placed a region of interest (ROI) over the peak of the vSTR, vmPFC and OFC signal at the time that the visual stimuli were presented on the screen and extracted the timecourse of the BOLD signal from two-voxel (1mm) radius spherical masks (one voxel (0.5mm) for VTA) using the FSLMEANTS function in FSL. In some cases, the time courses are shown for illustration only. For spatial surprise related activity, we looked at two main peaks of activity, namely a cluster in the frontal eye fields covering Area 8 and Are 45b and another cluster in the parietal cortex, at the dorsal end of the STS.

In order to illustrate and identify factors significantly influencing the timecourse of the BOLD signal in these regions we ran multiple regressions using a GLM of the BOLD timecourse (Table 3), after temporal upsampling, aligning it with the event onset of interest and cutting it with a fixed time window (epoching). Note

that the upsampling is necessary to better align an event of interest with the timecourse, because we sampled a volume at jittered timings relative to event onsets between trials. The upsampling method used was spline interpolation. In addition to the reward magnitude regressor (three drops versus one drop), the GLM included regressors representing the animals' reward history because this is a likely determinant of an animal's reward expectation.

To establish the statistical significance of the influence of specific variables on the BOLD signal we defined a window (1-4 seconds after reward delivery or cue onset) within the peak effects of an event should lie. We then used a leave one-out-procedure to find a peak or trough on the group level taking the remaining sample not used for peak estimation for testing. We ran this procedure for all sessions and could therefore calculate the mean peak/trough value across all sessions with a peak/trough position defined independently from the beta-weights being tested. All timeseries analysis was done in MATLAB (The Mathworks, Inc.).

Table 3 List of regressors used in ROI analysis

Regressor	Definition
Constant	n/a
Reward Size	Normalised reward size value
Last reward	Normalised reward size value of trial t-1
Reward t-2	Normalised reward size value of trial t-2
Reward t-3	Normalised reward size value of trial t-3
Long term reward expectation	Average reward rate as identified using a RL model on all trials minus the last three (α=0.05)
Two drop reward	Regressor value = 1 for every time two drops were given
Spatial surprise	Spatially surprising trials as identified using RL model (α=0.9)
Spatial updating	Regressor coding positive for t+1 after

	outlier and negative for t+1 after reversal
Trial number	Ascending regressor value based on trial number
logRT	Log transformed reaction time for current trial

5.3 RESULTS

5.3.1 Reward related effects

Whole brain analysis of reward size

First, we wanted to look for evidence of reward signals, using data from all the sessions we had acquired. In order to do this we ran a multiple regression analysis, i.e. a general linear model (GLM), using FLAME 1+2 from the FMRIB Software Library (FSL). The analysis was carried out across the whole brain using a cluster-based correction procedure for multiple comparisons. The whole brain analysis, revealed three main clusters in which activity was significantly related to the reward size regressor ($p<0.05$ at cluster forming threshold of $z > 2.3$, Table 4). The first and biggest cluster was a diffuse and widespread cluster covering big parts of the occipital and parietal cortex. It extended across brain anatomical and functional areas. The second cluster included large parts of frontal areas and spread into parts of the basal ganglia. The third cluster was localized in the midbrain and parts of the brain stem. Within those clusters, sub-clusters (local maxima) could be identified as anatomical regions ventral striatum (vSTR), a mesencephalic region including ventral tegmental area (VTA), orbitofrontal cortex (OFC), and ventromedial prefrontal cortex (vmPFC) which all showed significant activity increases when the animals received a higher rather than lower reward (areas in which activity was

significantly greater when animals received three drops – as opposed to one drop are shown in Figure 5.1 (z>2.3, cluster corrected, manually thresholded at z>3.1)). Based on these local maxima, we defined four spherical regions of interest (ROIs; illustrated in green in Figure 5.1) over the local activation peaks each with a radius of 1mm (0.5mm for VTA). These ROIs were used to further interrogate activity in these reward sensitive areas. The other regressors in our GLM were either related to the decision phase (e.g. spatial surprise, response side) or factors potentially important in the outcome phase (e.g. long-term reward expectation, trial number), controlling for other aspects that could have changed the BOLD signal.

Table 4 List of clusters and local maxima for whole brain analysis with reward size regressor

Cluster index	Local maxima	Max z-score	Coordinates (mm)			Coordinates (voxels)			Cluster size (# voxels)
			X	Y	Z	X	Y	Z	
1	1	5.1	16.5	-39.5	14.5	41	31	89	39265
2	1	7.26	-9.5	18	7	93	146	74	27092
	2	5.06	1.5	15	-2.5	71	140	55	
	3	4.79	-2	4.5	-3	78	119	54	
3	1	5.01	-3	-29.5	-23.5	80	51	13	9563
	2	4.63	-0.5	-12	-16	80	86	28	

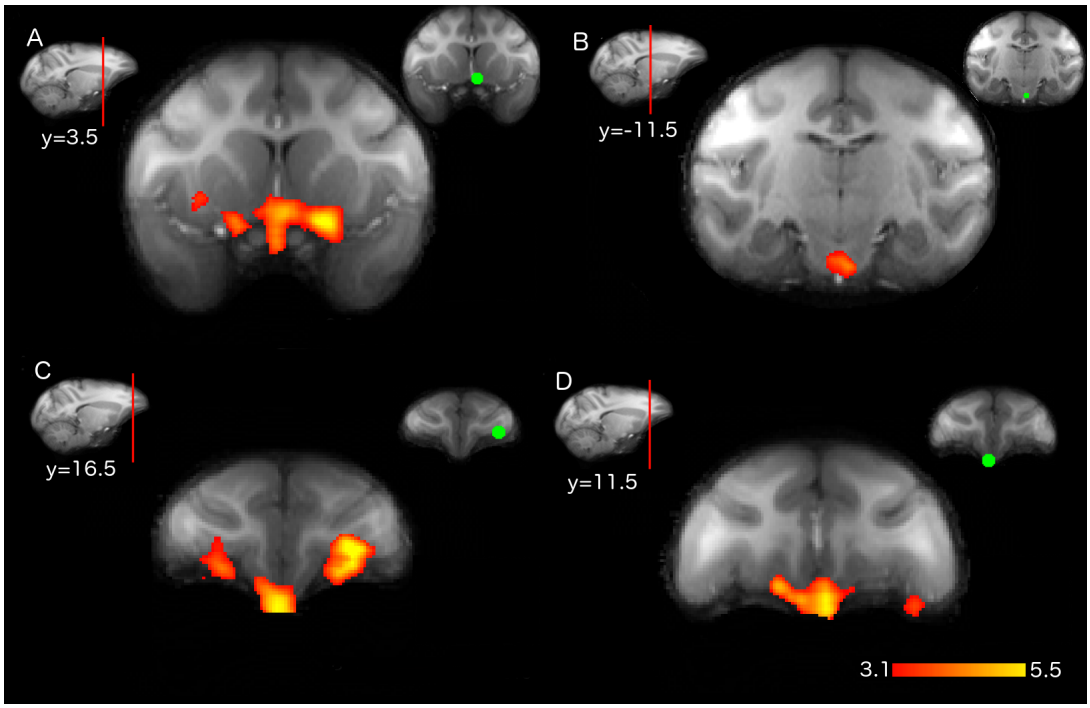

Figure 5.1 Whole brain analysis of reward size regressor. A shows vSTR, B shows VTA, C shows OFC and C and D show vmPFC. Small brains on the top left indicate slice position on the y-axis. Green dots on small brains on the top right represent ROI location.

ROI based effects of learning of reward expectation

We extracted the mean time course from our four ROIs of interest (vSTR, VTA, OFC, and vmPFC). In order to illustrate and identify factors significantly influencing the timecourse of the BOLD signal in these regions we ran multiple regressions using a GLM of the BOLD timecourse. In order to better understand signals relating to the short and long-term reward history, we first looked at the drift sessions. Here the average reward rate was drifting either up or down. Because they entailed a need to track fluctuating reward rates, these sessions are likely to carry the strongest or clearest reward rate-related signals.

First, we examined whether activity reflected not just the reward received on the current trial but whether it also simultaneously reflected

expectations that the animals might have had about reward size given the recent history of reward they had experienced on earlier trials. For example, if an area is coding not just reward receipt but reward prediction errors (the difference between reward received and the animal's prior expectation about what reward would be received) then its activity should be positively influenced by the magnitude of reward received on the current trial but it should also be negatively influenced by the reward received on the previous trial. Unsurprisingly, the reward magnitude regressor had a significantly positive effect in all four tested brain areas (Figure 5.2 A-D). We couldn't find any activity relating to past rewards or long-term reward expectation (LTRE) in vmPFC and OFC (Figure 5.2 A and B). However, we found that in the drift sessions, beyond the effect of the current reward, past reward events also had an impact on the BOLD signal in other areas. While the last (immediately previous) reward regressor had a significant positive effect in vSTR ($p<0.05$, unpaired t-test), this impact decayed with time; even earlier rewards from two trials ago failed to elicit a significant response. However, when going even further into the past, there was a significant negative effect of rewards that had been received three trials ago. Thus reward received recently, just like the reward received on the current trial, have a positive impact on activity (Figure 5.2 C). By contrast, rewards received earlier in time have a negative impact on activity that is reminiscent of the canonical reward expectation suppression necessary for prediction error coding. Interestingly, while the reward receipt on the current trial exerted a positive influence on activity in the VTA, one of the regions most prominently associated with prediction errors, expectation suppression only emerged for reward received

three or more trials ago, without an increase in BOLD signal for more recent past outcomes (Figure 5.2 D).

The same GLM also included a regressor capturing rewards that were even further in the past which we refer to as LTRE. In order to determine LTRE we used a simple reinforcement learning model which included rewards previous to the four reward events we had entered individually (reward on the current trial (t), reward on the previous trials t-1, t-2, and t-3). By also using a relatively low learning rate ($\alpha = 0.05$) we ensured that this regressor captured more long term reward signal. The low learning rate means that reward expectation is based on the distant reward history. We thought this regressor might be particularly interesting for vSTR and VTA as both have previously been associated with coding (negatively) for reward expectation (Figure 5.2 C and D). In other words, we wanted to test whether such a negative reward expectation signal, based on long term reward history, might be present in our task. Surprisingly, in vSTR, activity positively correlates with the LTRE, whereas VTA, as expected, long-term reward history had a negative effect on activity (Both effects are $p<0.05$, unpaired t-test, leave-one-out procedure, Figure 5.2 C and D, right panel). In other words, in VTA activity is positively influenced by the most recent rewards but suppressed by the prior reward history.

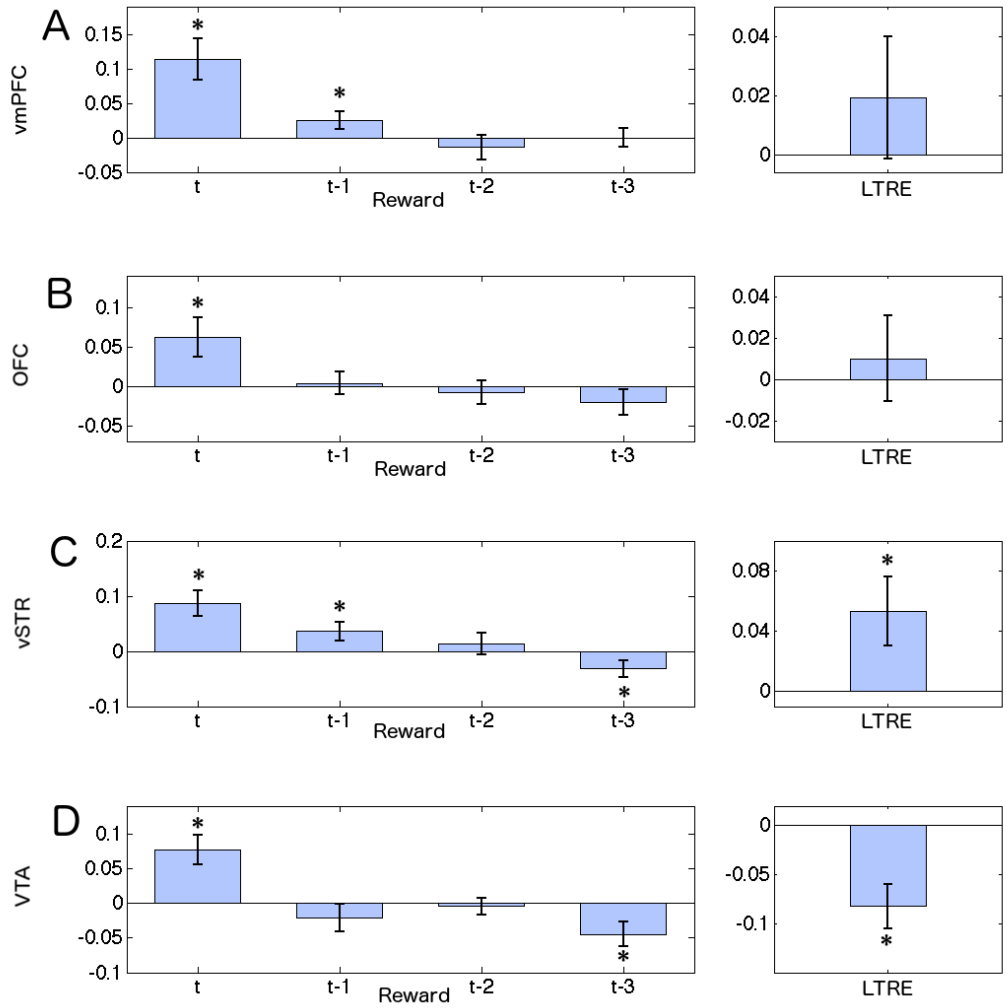

Figure 5.2 Reward related signals in four ROIs: vmPFC (A), OFC (B), vSTR (C) and VTA (D). All areas show a significant reward size effect for trial t. Immediate (t-1 to t-3) and more distant (LTRE) reward history has different effects on different brain areas. Errorbars: SEM

To visualise the results from Figure 5.2, Figure 5.3 shows the average beta weights of reward size and LTRE over the time of six seconds from reward onset. All four areas show a clear BOLD response to when the animals receive larger rather than smaller rewards (Figure 5.3 A-D). But while in vmPFC and OFC no clear LTRE signal can be observed (Figure 5.3 A and B), vSTR shows a late signal increase for the LTRE regressor and VTA a clear suppression (Figure 5.3 C and D). The difference in signal for reward size and LTRE in VTA looks like a classical canonical reward prediction

error. While the signal from vSTR showed positive beta weights for both reward delivery and expectation, signals in VTA showed a clear positive beta weight for reward delivery but a suppression (negative beta weight) for LTRE.

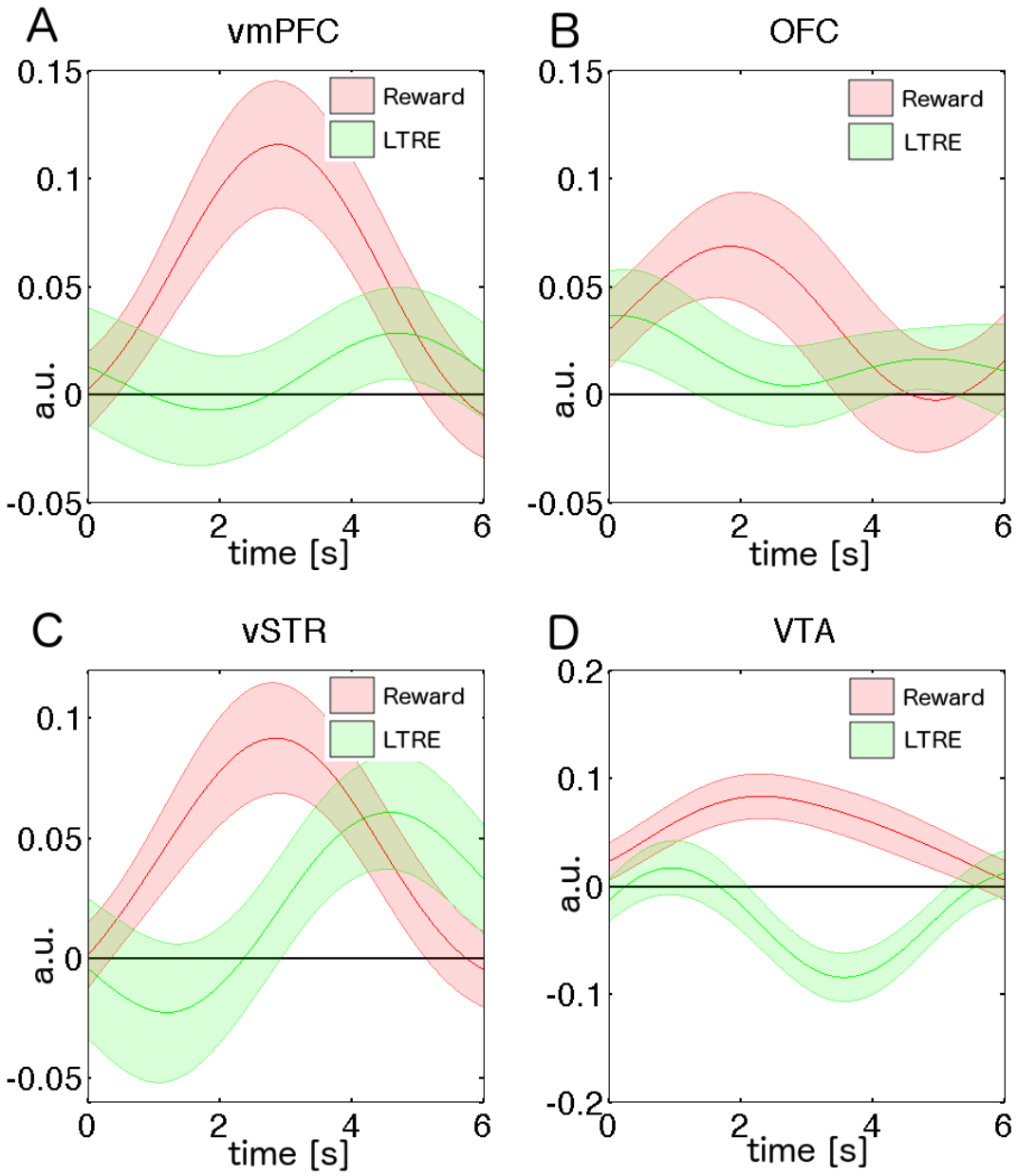

Figure 5.3 Regression coefficients for reward size and LTRE in four ROIs: A: vmPFC, B: OFC, C: vSTR and D: VTA. The difference between the red line (reward size) and green line (LTRE) can be seen as a reward prediction error. Errorbars: SEM

Comparing learning with the static reward sessions

After looking at the past reward signals during learning, we wanted to see whether they generalized to a non-learning context. Importantly, in those, "simple" sessions, reward size fluctuated randomly between 3 and 1 drops, but there was no structure or trend underlying those changes. Very surprisingly, we did not find a significantly positive beta or negative effect of past rewards in the VTA. VSTR on the other hand now seemed to signal rewards from trial t-3 positively. In other words, its activity was influenced by rewards in the intermediate past, on trial t-3 in a positive way just as it was it was influenced by rewards in the more distant past. Reward history-based activity suppression may be tied to prediction errors and learning and may not be possible in environments in which reward rates remain relatively constant.

When looking at vmPFC signals however, past reward events changed the BOLD signal in both learning and non-learning/simple sessions. We ran an area*trial*session_type ANOVA and could identify a triple interaction ($F(6,336)=2.16$, $p=0.046$). When running a area*trial ANOVA for each session type separately, we see a significant main effect for trial in the drift sessions ($F(2,44)=4.508$, $p=0.017$) but not for the simple sessions ($F(2,68)=0.307$, $p=0.737$).

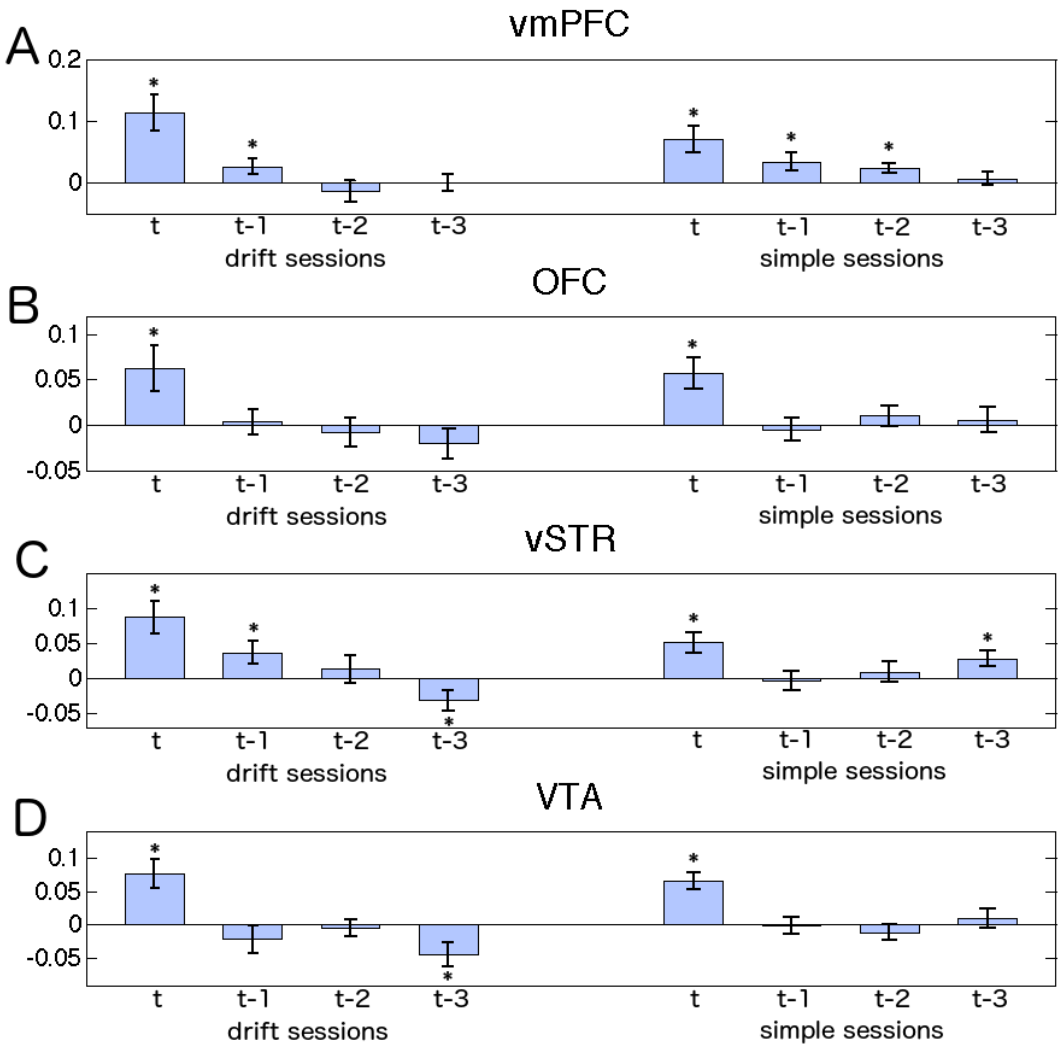

Figure 5.4 Reward in drift (left) and simple sessions (right). All sessions elicit a significant reward size effect across all RIOs for trial t. Immediate reward history (t-1 to t-3) are represented differently across brain areas as well as session types. Errorbars: SEM

Infrequent rewards (two drops) in different reward rate conditions

In addition to manipulating the extent of learning between sessions we also included a variable number of two drop rewards. Those rewards were surprising in the sense that they were encountered rarely, but very unsurprising in the sense that their reward magnitude corresponded to the overall average reward rate in each session. Interestingly, while not all reward sensitive brain regions responded to two drops any differently than expected from its intermediate amount of reward, OFC and vSTR seemed to be more active for two drops, but only in the context of

the learning sessions (Figure 5.5 B and C). In the simple sessions, despite two drops being rare, no increase in activity was seen when they were delivered. Not surprisingly, no activity specifically related to the two drop rewards was recorded in sessions in which two drops were not rarer than one or three drops.

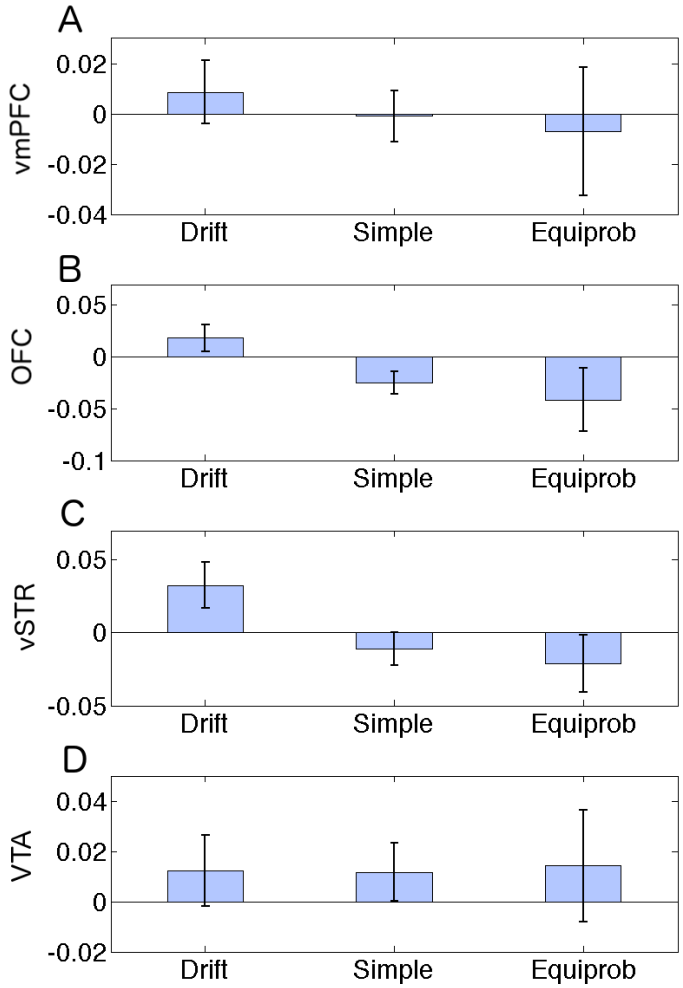

Figure 5.5 Two drop effect (reward identity surprise)) over three session types in all four ROIs. VmPFC and VTA don't show a significant effect whereas OFC and vSTR show a positive reward identity effect in drift sessions and a negative effect in non-drift sessions. Errorbars: SEM

We cannot completely rule out the possibility that the signal in the two drop condition in the learning session is based on the fact that two drops was not exactly the current expectation of reward held by the animal at any given time. Nevertheless, it is important to note that an alternative explanation of the activity

pattern in terms of a small unsigned surprise effect is inconsistent with a lack of an effect for the far larger manipulation of surprise between one and three drops and two drops in the simple sessions. Additionally, while two drops was slightly more surprising in the learning session than in the simple session, even within the learning session itself, one and three drops were by far more surprising when compared to the average reward and effects should have therefore been negative, rather than positive.

5.3.2 Spatial surprise related activity

Whole brain analysis of spatial surprise

On the whole brain level, we created a contrast between spatially surprising and unsurprising trials. Spatially surprising trials include both spatial reversal trials as well as spatial outlier trials. On the whole brain level, we found several clusters that were significantly more active in surprising trials compared to the trial immediately prior to a surprising trial. We found two big clusters ($z>2.3$, cluster corrected): One frontal cluster and one covering parts of the temporal lobe and parietal areas. Within these clusters we found several sub-clusters (see Table 5) and extracted the mean timecourses for two ROIs, one comprising area 8 and adjacent area 45b and one at the dorsal extent of the STS in the inferior parietal lobule in a region that has been described as area 7a or PG depending on nomenclature (Figure 5.6, (Gottlieb & Snyder, 2010)).

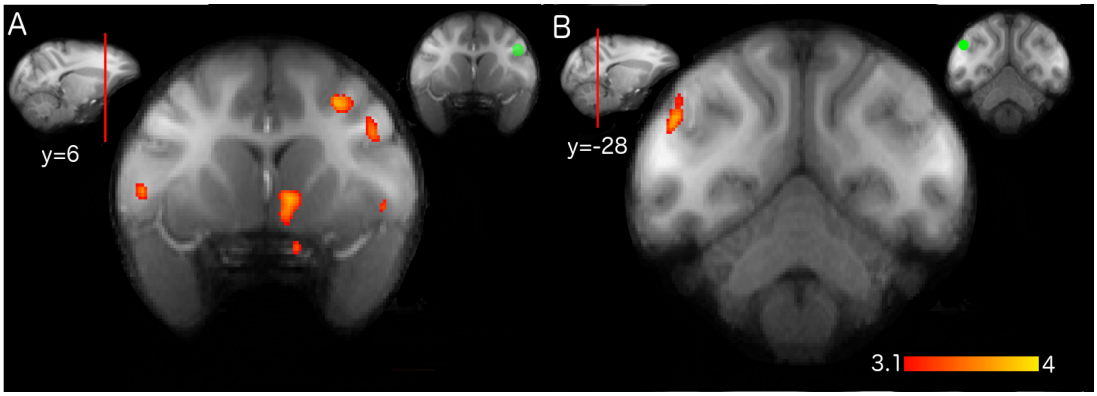

Figure 5.6 Whole brain effects of spatial surprise regressor. (A) Area 8 + 45B, (B) area 7a. Cluster corrected, z>2.3, p<0.05, manually thresholded at z>3.1. Small brains on the top left indicate slice position on the y-axis. Green dots on small brains on the top right represent ROI location.

Table 5 List of clusters for whole brain analysis of spatial surprise regressor, cluster corrected z>2.3, thresholded z>3.1

Cluster index	Max z-score	Coordinates (mm)			Coordinates (voxels)			Cluster size (# voxels)
		X	Y	Z	X	Y	Z	
1	3.85	-5	0.5	0	84	111	60	820
2	3.9	25	-2.5	5.5	24	105	71	733
3	3.94	-18	-9	-8.5	110	92	43	728
4	3.8	-6	-22.5	21.5	86	65	103	454
5	3.82	-11.5	5.5	16	97	121	92	376
6	4.16	-19.5	8	-0.5	113	126	59	350
7	3.83	22.5	-28.5	11.5	29	53	83	227
8	3.85	18.5	3	2.5	37	116	65	200
9	3.61	-21.5	2	3	117	114	66	159

Time courses in ROIs

We looked at the timecourse of activity change related to surprising spatial events in these two ROIs and also in the ROIs identified by the reward regressor. We ran the same multiple regression analysis as before but time locked it to the onset of stimulus appearance. First, we looked at the timecourse in areas that showed a

whole brain effect for spatial surprise (area 8 and area 7a) and could confirm the significant signal (peak $p<0.05$, tested in leave one out procedure). In area 7a (Figure 5.7E) we see a peak of activation at around four seconds after stimulus onset whereas in Area 8 (Figure 5.7F) the peak seems to appear earlier (about two seconds). We then looked in the reward sensitive areas to assess whether we could identify spatial surprise signals there too. VmPFC and OFC did not show a significant timecourse fluctuation (Figure 5.7A + B), but vSTR showed a very early activation and VTA a very early suppression for visual surprise (Figure 5.7C + D).

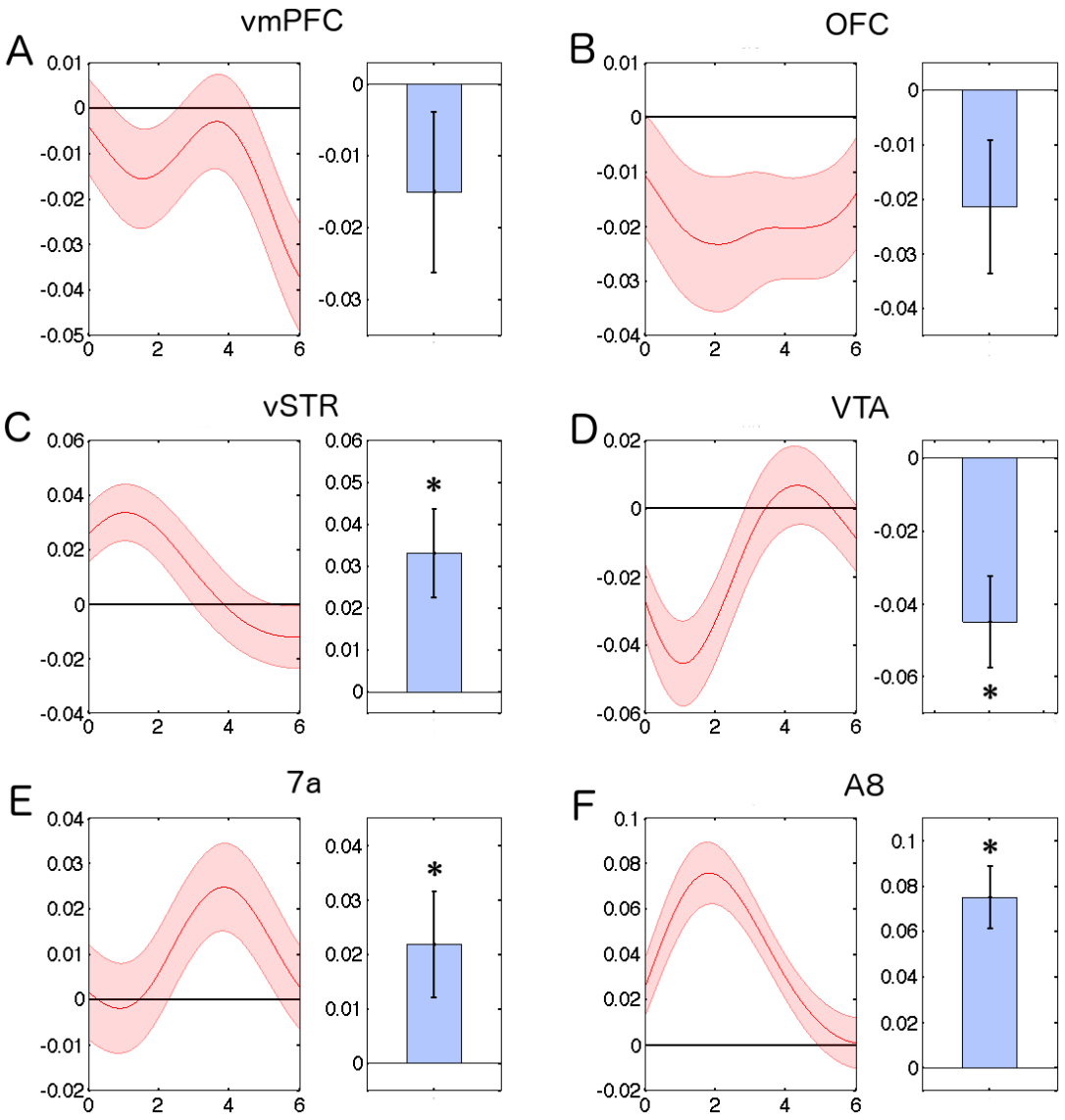

Figure 5.7 Spatial surprise in reward sensitive areas (A-D, left plot) and ROIs as found in the whole brain analysis for spatial surprise (E and F, left plot). Right plots show average peak value of correlation coefficient as determined by leave-one-out procedure. Errorbars: SEM

We also looked at neural effects on trials after outlier and reversal trials to see whether we could find any evidence for animals not only reorienting towards but also updating their expectation of target position in the following trials. In order to do that, we looked at BOLD correlations in area 7a and FEF with the spatial update regressor coding positively for t+1 after an outlier trial and negatively for t+1 after a reversal trial. While we couldn't find any signal that correlates with the spatial

updating regressor in the parietal cortex, FEF (in particular area 8) showed a significantly negative correlation with the spatial updating regressor (Figure 5.8). Note, that the regressor was designed in a somewhat counterintuitive way by coding positively for trials after which less updating should happen, and negatively for trials after which animals show updating behaviour.

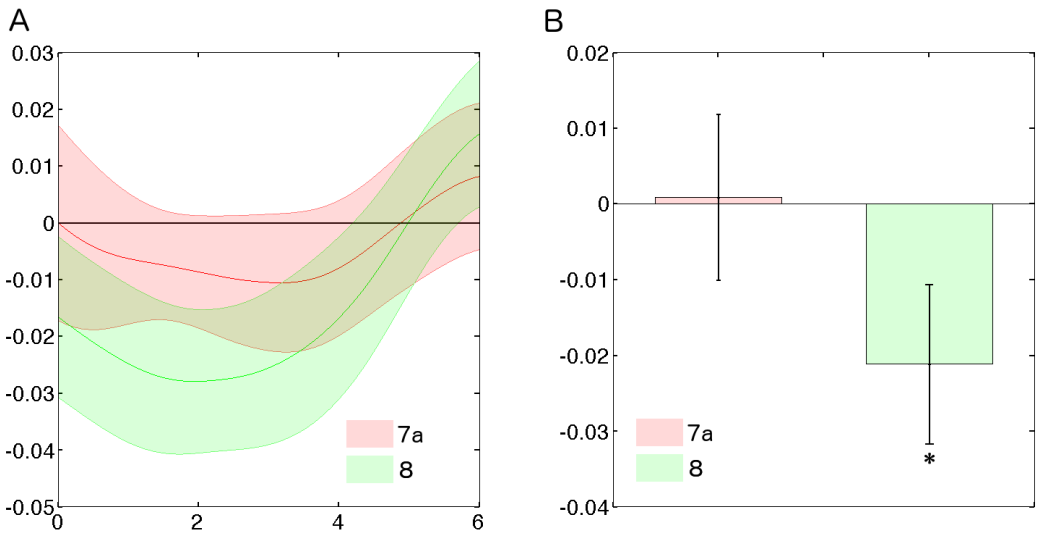

Figure 5.8 Trial n+1 after a reversal versus outlier trial in area 7a (red) and area 8 (green). (A) BOLD timecourse of parametric surprise n+1 regressor which codes positively for trials after an outlier and negatively for trials after a reversal. (B) Mean peak in BOLD timecourse as determined with leave-one-out procedure. Statistically significant result in area 8 (one-sided t-test, p<0.05). Y-axis are arbitrary units, x-axis is time in seconds locked to the onset of the stimulus. Errorbars: SEM

5.3.3 Trial number effect

Intrigued by a consistent pattern of trial number effects across different ROI timecourses we ran the whole brain analysis again but added one regressor, coding for trial number, into the GLM. We found activity that correlated with increasing trial number that is spatially close to the vmPFC ROI we used for extracting reward-related signals (Figure 5.9A). When looking at the beta weights of the multiple regression analysis, we found a fluctuating pattern of brain activity that correlated with trial number (Figure 5.9B). It appears that there is a trial locked pattern of signal

increase and decrease that gets stronger the longer the experiment runs. In fact, this fluctuation could be observed in a variety of brain areas including OFC and parts of the cingulate cortex (see Table 6).

Table 6 List of clusters for whole brain analysis of trial number regressor, cluster corrected z>2.3,

Cluster index	Local maxima	Max z-score	Coordinates (mm)			Coordinates (voxels)			Cluster size (# voxels)
			X	Y	Z	X	Y	Z	
1	1	6.23	-14	2	-10	101	114	40	93201
	2	5.81	15	4.5	1.5	44	119	63	
	3	5.6	-10.5	-2	-4.5	95	106	51	
	4	5.58	17.5	0	-9	39	110	42	
	5	5.5	-9.5	-5	-5	93	100	50	
	6	5.31	13	3	-11	48	116	38	
2	1	4.6	15	-10.5	21.5	44	89	103	8240
	2	4.48	13	-9	23	48	92	106	
	3	4.43	10.5	-3	23.5	53	104	107	
	4	3.7	3	-16.5	27	68	77	114	
	5	3.57	1	-16	17	72	-16	17	
	6	3.52	1.5	-15.5	15	71	79	90	

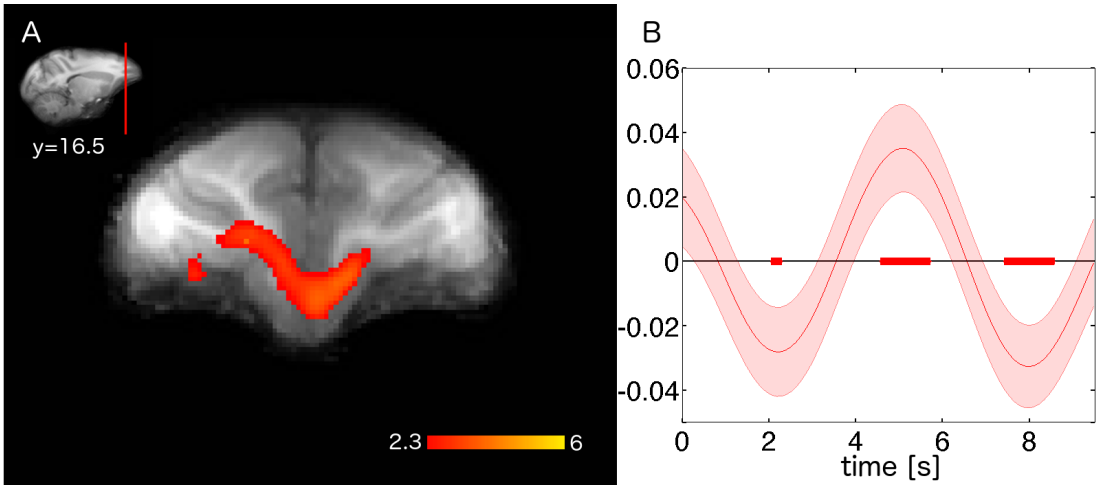

Figure 5.9 Trial number effects in (A) whole brain contrast (cluster corrected, p<0.05) and (B) regression coefficient in the vmPFC ROI (time locked to stimulus onset). Red sections indicate statistically significant difference from 0 (one sided t-test, p<0.05). Errorbars: SEM

5.3.4 Parietal cluster

The reward size regressor also correlated significantly with a cluster over visual areas in occipital and parietal cortex. While these areas are not known to carry reward size signals per se, in our experiment however, we presented a visual feedback (reappearance of the chosen stimulus) concurrent with reward delivery. Thus the visual activity could have been triggered by the visual feedback and was particularly strong when reward size was large. This is in line with research that showed that BOLD response to visual stimuli in early visual areas (even as early as V1) is modulated by its associated reward value (Serences, 2008; Serences & Saproo, 2010). Furthermore, electrophysiological recordings showed a reward modulated increase in activity in neurons in the parietal cortex in areas such as LIP (Bendiksby & Platt, 2006; Peck et al., 2009; Platt & Glimcher, 1999; Sugrue et al., 2004).

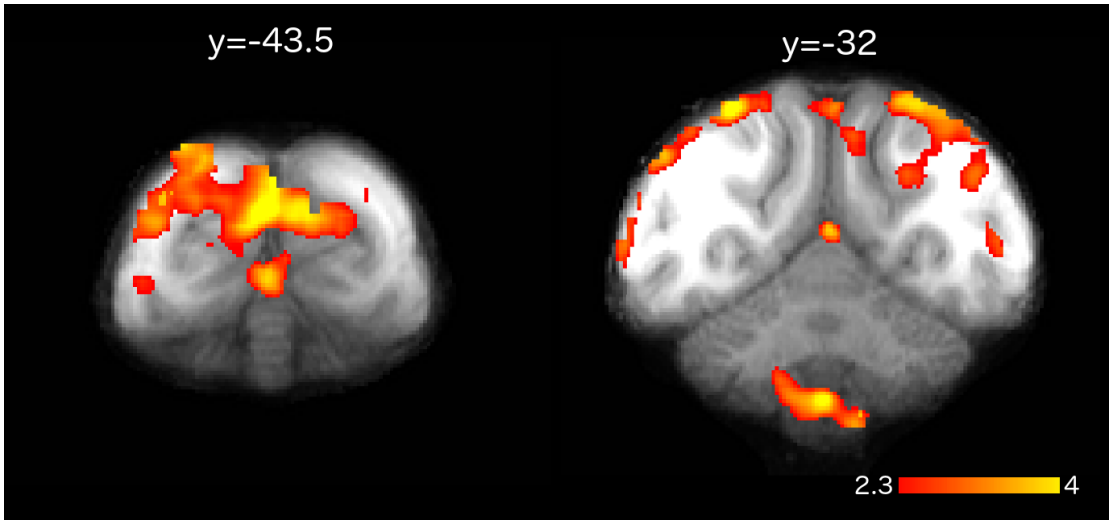

Figure 5.10 Parietal cluster of whole brain reward size regressor (cluster corrected, z>2.3, p<0.05)

5.4 DISCUSSION

When rewards were delivered there were significant changes in activation that were related to reward size in several brain areas, including several that have already been described as reward sensitive areas: dopaminergic midbrain (VTA, (Ljungberg et al., 1992; Romo & Schultz, 1990; Schultz & Romo, 1990; Schultz, 2006)), ventral striatum (Apicella et al., 1991; Schultz et al., 1992), as well as frontal areas such as vmPFC (Bouret & Richmond, 2010) and OFC (Padoa-Schioppa & Assad, 2006; Rolls et al., 1996; Schultz & Tremblay, 1999). All of these areas showed a significant signal increase when the animal received a higher reward compared to a lower reward. When looking at the individual BOLD timecourses in each area however, we found that different brain areas seem to carry information about different aspects of reward history and long-term reward expectation.

5.4.1 VTA

We measured BOLD activity in the vicinity of the dopaminergic nuclei in the midbrain, similarly to previous studies in humans (D'Ardenne et al., 2008; Klein-Flügge et al., 2013). When recording activity with fMRI it is not possible to establish whether the neurons generating the signal are dopaminergic neurons as opposed to adjacent neurons in the same nucleus, as the signal is an indirect measure of neural activity via oxygen consumption with no exclusive relationship to one neurotransmitter. While the activity was centred on the VTA it may have extended to some degree into the adjacent substantia nigra pars compacta (SNpc). In many ways the pattern of activity recorded in VTA is approximately that expected from a brain region encoding reward prediction errors. There was a positive response to the receipt of reward (Figure 5.1B and Figure 5.2D) and, at the time of reward delivery, a negative response as a function of the reward expectation given the previous reward history (Figure 5.2D). Typically, however, activity reflecting a reward prediction error would be most negatively influenced by the most recent rewards (these are particularly important for determining the current reward expectation) followed by an exponential decay in the influence so that rewards more distant in time have less negative influence (Bayer & Glimcher, 2005). The rate of decay will reflect the learning rate (faster decay with faster learning rates) (Behrens et al., 2007; Rescorla & Wagner, 1972; Watkins & Dayan, 1992). However, by contrast, in the current experiment recent rewards had very little impact on VTA activity and it was only rewards received on trials t-3 and earlier that had a negative impact on VTA activity on the current trial. We, however, add a note of caution as this may simply reflect the fact that rewards in our experiment exerted a positive

impact for several seconds after their receipt and the onset of consumption and that this impact continued into subsequent trials. It is possible that some of the larger three drop rewards, approximately 4.5ml in size, were still physically present in the animals' mouths on subsequent trials. Alternatively, short term positive reward effects might also be neutral in nature. Either way, if short term positive and negative effects intersected they would cancel out, leaving only the longer term expectation with an impact on the aggregate signal.

5.4.2 vSTR

By contrast, the vSTR exhibited a very different pattern of activity. Although it, like the VTA, responded positively to receipt of reward on the current trial, both the most recent past reward events and the long-term reward rate average were associated with an activity increase at the time of reward delivery. The signal correlating with LTRE that we found in vSTR not only differs from the signal we found in VTA but also stands in contrast to reward expectation signals that have previously been described in fMRI studies in humans (e.g. D'Ardenne et al., 2008; Klein-Flügge et al., 2013). In fact, reward expectation-related suppression, consistent with prediction error coding, was confined to only the relatively recent past reward history (i.e. t-3). While the overall response pattern is inconsistent with the simplest prediction error based theories, it is consistent with the idea that striatum might additionally carry other information about reward. For example, longer term reward rates reflect the quality of the environment that an animal is in and the opportunities that are available in the environment. Behaviour should be

invigorated when the environment is richer so that as many reward opportunities as possible can be harvested (Niv et al., 2007).

5.4.3 vmPFC

The vmPFC is one of the cortical areas that is has been suggested to hold a reward prediction error coding in humans (Rutledge et al., 2010). However, in the present experiment involving delivery of primary reinforcers it was unclear whether vmPFC activity reflected reward prediction error coding. We found that vmPFC carried a reward magnitude signal which quickly decays over time irrespective of whether animals had to learn about drifting reward rates. (Figure 5.2A and Figure 5.4A). While very recent rewards (t-1) and therefore short term reward expectation were positively represented in the timecourse, we could not find an longer term reward expectation suppression effect (and thus no RPE) in vmPFC. The long term reward expectation effect was instead positive but non-significant.

This result is in line with single cell recordings in vmPFC in monkeys that have shown activity increases as a function of expected reward size, both at the cue and at reward delivery (Bouret & Richmond, 2010) and also that neurons in vmPFC are significantly modulated by immediate reward expectation (San-Galli et al., 2016). Strait et al. showed that some vmPFC neurons are tuned both positively and negatively to reward but their value signal depended on both reward size and probability and whether the reward was related to a choice that might be taken or rejected (Strait et al., 2014). Other studies confirm reward outcome signals in vmPFC and have subdivided the area in two or more subregions, reporting that a ventral

subregion is sensitive to 'appetitive' rewards whereas a dorsal subregion fires more when 'aversive' punishments are received (Monosov & Hikosaka, 2013).

5.4.4 OFC

So far I have concentrated on activity that might be related to scalar reward prediction errors. In addition, however, the experimental design made it possible to look at whether neural activity also reflected expectations about precise reward sizes. In drift blocks, animals mainly received rewards of either one or three drops. The average scalar reward prediction should therefore have been approximately two drops. The expectation would slightly increase or decrease during the session depending on whether the average reward rate was drifting upwards or downwards. However, while the average reward expectation was close to two drops, monkeys rarely actually received two drops; two drop rewards were only received on 10% of trials. Thus two drop reward events were expected in the sense that they were always closer to the average reward expectation of approximately two drops than either three or one drop outcomes regardless of the direction of drift. However, if animals hold expectations about particular reward events then two-drop rewards will be the least expected rewards because they were the rarest events.

We found evidence of activity reflecting expectations about precise types of reward and not just the average amount of reward in vStr and lOFC. In both these regions there was a larger response to the two drop rewards in drift sessions than to either one or three drop rewards (Figure 5.5). There was no evidence that such

reward prediction errors were signalled by activity in the VTA. The pattern of results suggest that multiple types of reward expectation are carried by different brain areas and some reward expectations about specific reward levels may not depend on the dopaminergic midbrain in the same way as other scalar reward prediction errors. Instead neural mechanisms associated with the representation of particular reward types and identities in the orbitofrontal cortex may be important (Rudebeck et al., 2013; Rudebeck & Murray, 2014). Rudebeck and Murray (2014) have emphasized the importance of a central region of orbitofrontal cortex between the lateral and medial orbitofrontal cortex. The current results highlighted a region immediately laterally adjacent in lOFC.

5.4.5 Spatial surprise

The way in which surprising events were processed depended on their nature; surprising visual stimuli were associated with activity patterns that only partially overlapped with those generated in response to surprising reward events. We found that surprising visual activity was associated with activity in a frontoparietal circuit centred on the IPL and FEF. Activity in both these regions increased when visual stimuli appeared in an unexpected location requiring re-orienting of attention.

There has been some debate about the degree to which visuospatial circuits in the macaque and human correspond. For example it is often argued that Brodmann (1909) failed to find an area corresponding to the human IPL region in the macaque brain. Brodmann labelled the human IPL areas 39 and 40 but he did not assign a similar number to any parietal area in the monkey. In humans the IPL region and adjacent temporoparietal junction are associated with the redirecting of

attention to surprising events (Corbetta & Shulman, 2002). But in the macaque it has been suggested that the IPL does not have a similar function (Patel et al., 2015). However, closer reading of Brodmann's report suggests that he found considerable similarities between the IPL region in both humans and macaques and other researchers have emphasized similarities in the patterns of resting state functional connectivity in humans and macaques (Mars et al., 2011). The current data are consistent with a related role for IPL in re-orienting and/or updating the direction of visual attention when surprising visual events occur.

Parts of the parietal cortex are associated with two separate but interconnected networks of attentional reorienting. One network comprises parts of the IPS and FEF (dorsal attention network) and the other includes TPJ and lateral frontal areas (Corbetta & Shulman, 2002). The dorsal network is thought to be involved in goal-directed top down shifts of attention whereas the ventral network is activated by salient but surprising events that trigger a bottom-up or stimulus-driven shift of attention. Attentional reorienting can also trigger a learning or updating process that adjusts the prior expectation about events or observations in the future. Recently, there has been evidence that the parietal cortex houses two functionally dissociable areas that are active in either the mere reorienting of attention (IPS3, human homologue of macaque LIP) or the updating of people's spatial expectation (PGp, human homologue of area 7a) (O'Reilly et al., 2013). In our experiment, we find activation in area 7a for visual surprise but did not find evidence for a signal that corresponds to an updating of expectation. However, signals in FEF correlate with spatial updating (Figure 5.8).

While some areas might be especially concerned with surprising visual events as opposed to reward events, a small number of areas appeared to have a more general role in responding to surprising events. Specifically, in my experiment, vSTR and VTA showed a very early activation and deactivation, respectively, when a spatially surprising stimulus was shown.

5.4.6 Trial number effect

When looking at the trial number regressor in the whole brain analysis, we found activity in frontal areas, adjacent to OFC and vmPFC ROsI as defined by the reward size effect. When looking at the BOLD timecourse in these two areas, we found a fluctuation pattern that was time locked to stimulus onset and positively (and negatively) correlated with trial number. Recent studies have found neural correlates of fatigue, saturation or attentional effort (San-Galli et al., 2016; Stoll et al., 2016). In the study by San-Galli et al. vmPFC neurons encoded an animal's willingness to engage in the task; the neurons' activity (or lack thereof) predicted when animals took a voluntary break during the experiment. Stoll and colleagues found a correlation between frontal beta power in the LFP and the length of experiment. The beta power increased until the animal took a break and then reset to baseline levels. They concluded that because of this reset, beta power represents the attentional effort rather than more general fatigue or satiation. Our results support the idea that the length of the experiment or attentional effort might be encoded in neurons in the ventral PFC. Especially LFP oscillations in the beta range in frontal areas have been linked with cognitive control (Bastos et al., 2015; Buschman & Miller, 2007; Siegel et al., 2012). Furthermore, beta oscillations have been thought

to have a substantial influence on the BOLD signal (Jabbi et al., 2015; Magri et al., 2012; Stevenson et al., 2011).

Chapter 6: General Discussion

6.1 Spatial Updating and Surprise in Humans and Macaques

Although attentional reorienting has been described and studied in great detail (e.g. Corbetta & Shulman, 2002; Posner et al., 1980), previous studies did not dissociate between effects due to shifting attention in space and effects that relate to learning and updating expectations triggered by a surprising observation. In chapters 2 and 3, however, I showed that there are indeed behaviourally and neurally dissociable effects that correspond to either a mere reorienting of the attentional focus or to updating one's internal model about where future observations are likely to appear. RTs confirmed that participants followed the task instructions and did not update on trials on which they were not meant to, but updated on trials on which the spatial distribution of targets changed. Dwell time, i.e. the time participants spent fixating on a target, differed significantly between the two surprising trial types. Participants fixated the target longer on an update trial which could reflect an underlying updating process. Furthermore, effects measured by pupillometry provided more evidence for distinct cognitive processes for reacting to surprise and updating. Using a GLM with regressors derived from information theory, surprise [Shannon information (I_S)], and updating [Kullback–Leibler divergence (D_{KL})] evoked opposite responses in pupil size. While on trials with high I_S we found an increased pupil size, there was a negative correlation of pupil size with D_{KL}. Surprise related increase in pupil size has been previously reported (Nassar et al., 2012) and changes in pupil size in the absence of a change in luminance have been attributed to firing of

noradrenergic neurons in the locus coeruleus (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005; Preuschoff et al., 2011). However, it remains unclear how the update-specific decrease in pupil size fits with current computational theories of noradrenaline (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005) which have linked learning, uncertainty, noradrenaline release, and pupil dilation and would tend to predict an increase in pupil dilation on update trials. It should be noted that multiple psychological factors, including arousal and attention, can also affect pupil dilation and that the effect we found should not be overinterpreted in the absence of pharmacological work. Nonetheless, the effect is replicable and provides robust evidence for differential processing of update and non-update stimuli.

We were also able to dissociate distinct neural effects that reflect surprise and updating. Surprise (I_S) was strongly correlated with activity in the posterior parietal cortex, with the peak activation in the superior parietal lobule (SPL) and extending along the IPS. In contrast, updating (D_{KL}) was strongly correlated with activity in the anterior cingulate cortex (ACC), particularly the rostral cingulate motor area (rCMA) and extending into the ventral part of the adjacent pre-supplementary motor area (pre-SMA) but also with activity in the IPL, specifically in area PGp (Mars 2011). Although previous experiments have not dissociated surprise from updating, parts of the PPC have been known to be active when surprise motivates attentional shifts (Hartwigsen et al., 2012; Mars et al., 2007; Mort et al., 2003; Rushworth et al., 2001a; Rushworth et al., 2001b; Thiel et al., 2004) but there has been disagreement about which regions within the parietal cortex are most important for redirecting attention and eye movements (Vandenberghe et al., 2012). Our finding that activity in a part of the posterior IPS/IPL (PGp) is more correlated with updating than with

surprise supports the idea of different PPC areas being involved in different aspects of redirecting attention. Furthermore, the neurological literature describes different clinical conditions associated with damage to different areas in the parietal cortex. Lesions of the IPS and IPL produce different deficits in patients. Lesions of the IPS are associated with impairments in the ability to redirect the eyes, as in Balint syndrome (Pierrot-Deseilligny & Muri, 1996), or the covert focus of attention, as in extinction (Posner et al., 1984). In contrast, spatial neglect is associated with damage to the IPL (Milner, 1996; Mort et al., 2003).

In the experiment with macaques described in chapters 4 and 5, we replicated some of the effects of spatial surprise reported in chapter 2. Animals were slower when trials were spatially less predictable (in both outlier and reversal trial conditions) but they learned to adapt to that surprise quickly (within one trial); they responded fast on the trial after a reversal trial. Unlike the human study, in the macaque study we did not cue monkeys as to whether a trial was an outlier or reversal trial and thus animals showed some updating behaviour even in trials that were mere outliers. It could also be that the animals updated on every surprising trial, that is on both reversal and on outlier trials. This would mean that they may have "believed" a reversal had happened even on outlier trials so that they would have to update yet again on outlier+t trials as their spatial expectation would be violated again. Neurally, we found areas in the parietal cortex and in the dorsolateral frontal cortex to be active in spatially surprising task conditions. This is in line with the model of frontoparietal attention networks being involved in redirecting attention towards surprising events in humans (Corbetta & Shulman, 2002). However, it has been argued that that parts of the dorsal attention network, in

particular IPL, have a different function in the macaque brain. But other researchers have emphasized similarities in the patterns of resting state functional connectivity in humans and macaques (Mars et al., 2011). The current data are consistent with a related role for IPL in the macaque brain in re-orienting and/or updating the direction of visual attention when surprising visual events occur.

6.2 Reward Signals in the Macaque Brain

In chapter 5 I presented and discussed data from an fMRI study in macaques. We designed an experiment in which the animals received different amounts of reward (one, two or three drops) for responding to a visual stimulus. We kept the average reward rate fixed across all session types but changed the relative frequency of reward magnitudes within a session ('simple' versus 'equiprob' condition) or manipulated the reward rate within a session ('drift' condition) so that it either drifted upwards (from roughly 1.5 drops to 2.5 drops on average) or downwards (from roughly 2.5 to 1.5 drops on average).

Behaviourally, animals showed RT differences on the trial after different reward magnitudes, suggesting that they were able to distinguish between the three different reward sizes. Neurally, we found significant changes in activation that were related to reward size in several brain areas including several that have already been described as reward sensitive areas: in the vicinity of the dopaminergic midbrain (VTA) (Ljungberg et al., 1992; Romo & Schultz, 1990; Schultz & Romo, 1990; Schultz, 2006), ventral striatum (VSTR) (Apicella et al., 1991; Schultz et al., 1992), as well as frontal areas such as ventromedial prefrontal cortex (vmPFC) (Bouret & Richmond, 2010) and orbitofrontal cortex (OFC) (Padoa-Schioppa & Assad, 2006; Rolls et al.,

1996; Schultz & Tremblay, 1999). All of these areas showed a significant signal increase when the animal received a higher reward compared to a lower reward. Furthermore, we looked at the BOLD timecourse extracted from VTA, vSTR, vmPFC and OFC and primarily focused on signal recorded during drift sessions. In drift sessions, animals had to keep track of a varying average reward rate so learning about outcomes and reward history was more important compared to non-drift sessions.

Unsurprisingly, all four ROIs showed a strong effect for reward size (Figure 5.2 and Figure 5.4). Beyond the reward size, different areas seemed to encode different aspects of reward history and therefore reward learning and expectation. VTA exhibited a classic reward prediction error signals; its activity was positively correlated with reward size but, at the time of reward delivery, it was negatively correlated with the level of prior reward expectation; at the time of reward delivery there was a suppression of activity in proportion to reward expectation. This expectation, however, was based on more distant rewards in the past so that the strongest correlation could be found with the long-term reward expectation. VSTR on the other hand showed a complex pattern of reward activation and expectation suppression. The immediate reward history was represented by an activation in relation to reward on trial t-1 but the signal flipped and we found a suppression for trial t-3. When looking further into the past, however, we found that vSTR showed a late activation for long-term reward expectation (in other words reward expectations based on the even earlier history of reward events). This provides evidence that vSTR tracks reward history in several ways. It keeps track of the immediate reward history but it also tracks the reward history over the longer-term.

OFC did not exhibit either clear reward expectation related activation or suppression but there was evidence that coded for reward identity in a more sophisticated way. When comparing drift versus non-drift (simple) sessions, we found activation and suppression, respectively when animals received a (rare) two drop reward or an expected one or three drop reward. VmPFC showed a short lasting reward expectation signal that was present irrespective of session type. Furthermore, we found a fluctuating pattern in the BOLD signal that correlated with trial number. This could represent the attentional effort or willingness to engage in the task (San-Galli et al., 2016; Stoll et al., 2016).

Bibliography

Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue Environment. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *36*(4), 1096–112. https://doi.org/10.1523/JNEUROSCI.3159-15.2016

Amador, N., Schlag-Rey, M., & Schlag, J. (2000). Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. *Journal of Neurophysiology*, *84*(4), 2166–2170.

Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience*, *21*(12), 3447–3452. https://doi.org/10.1111/j.1460-9568.2005.04170.x

Amiez, C., Kostopoulos, P., Champod, A.-S., & Petrides, M. (2006). Local Morphology Predicts Functional Organization of the Dorsal Premotor Region in the Human Brain. *Journal of Neuroscience*, *26*(10).

Amiez, C., & Petrides, M. (2014). Neuroimaging Evidence of the Anatomo-Functional Organization of the Human Cingulate Motor Areas. *Cerebral Cortex*, *24*(3), 563–578. https://doi.org/10.1093/cercor/bhs329

Anderson, T. J., Jenkins, I. H., Brooks, D. J., Hawken, M. B., Frackowiak, R. S. J., & Kennard, C. (1994). Cortical control of saccades and fixation in man A PET study. *Brain*, *117*(5).

Apicella, P., Ljungberg, T., Scarnati, E., & Schultz, W. (1991). Responses to reward in monkey dorsal and ventral striatum. *Experimental Brain Research*, *85*(3), 491–500. https://doi.org/10.1007/BF00231732

Ashby, W. R. (1947). Principles of the Self-Organizing Dynamic System. *The Journal of General Psychology*, *37*(2), 125–128. https://doi.org/10.1080/00221309.1947.9918144

Astafiev, S. V, Shulman, G. L., & Corbetta, M. (2006). Visuospatial reorienting signals in the human temporo-parietal junction are independent of response selection. *The European Journal of Neuroscience*, *23*(2), 591–6. https://doi.org/10.1111/j.1460-9568.2005.04573.x

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–50. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649–666. https://doi.org/10.1016/j.neunet.2009.12.007

Barbas, H. (2007). Specialized elements of orbitofrontal cortex in primates. *Annals of the New York Academy of Sciences*, *1121*, 10–32. https://doi.org/10.1196/annals.1401.015

Barbas, H., & Mesulam, M.-M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *The Journal of Comparative Neurology*, *200*(3),

407–431. https://doi.org/10.1002/cne.902000309
Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., … Fries, P. (2015). Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*, *85*(2), 390–401. https://doi.org/10.1016/j.neuron.2014.12.018
Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020
Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science (New York, N.Y.)*, *275*(5304), 1293–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9036851
Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, *20*(2), 1052–1063. https://doi.org/10.1016/S1053-8119(03)00435-X
Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci*, *29*(4), 1175–1190. https://doi.org/10.1523/jneurosci.3328-08.2009
Behrens, T. E. J., Fox, P., Laird, A., & Smith, S. M. (2013). What is the most interesting part of the brain? *Trends in Cognitive Sciences*, *17*(1), 2–4. https://doi.org/http://dx.doi.org/10.1016/j.tics.2012.10.010
Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. https://doi.org/10.1038/nn1954
Bendiksby, M. S., & Platt, M. L. (2006). Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia*, *44*(12), 2411–2420. https://doi.org/http://dx.doi.org/10.1016/j.neuropsychologia.2006.04.011
Bernard, C. (1974). *Lectures on the phenomena of life common to animals and plants*. Thomas.
Bijleveld, E., Custers, R., & Aarts, H. (2010). Unconscious reward cues increase invested effort, but do not change speed–accuracy tradeoffs. *Cognition*, *115*(2), 330–335. https://doi.org/10.1016/j.cognition.2009.12.012
Blamire, A. M., Ogawa, S., Uğurbil, K., Rothman, D. L., McCarthy, G., Ellermannc, J. M., … Shulman, R. G. (1992). Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging. *PNAS*, *89*, 11069–11073.
Blatt, G. J., Andersen, R. A., & Stoner, G. R. (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *The Journal of Comparative Neurology*, *299*(4), 421–445. https://doi.org/10.1002/cne.902990404
Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, *62*(5), 733–743. https://doi.org/10.1016/j.neuron.2009.05.014
Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., & Behrens, T. E. (2016). Two Anatomically and Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome Associations in Hippocampus. *Neuron*, *89*(6), 1343–54.

https://doi.org/10.1016/j.neuron.2016.02.014
Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(5), 201–208. https://doi.org/10.1016/j.tics.2008.02.009
Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179–181. https://doi.org/10.1038/46035
Bouret, S., & Richmond, B. J. (2010). Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. *The Journal of Neuroscience*, *30*(25), 8591 LP-8601. Retrieved from http://www.jneurosci.org/content/30/25/8591.abstract
Bouret, S., & Sara, S. J. (2005). Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2005.09.002
Bowman, E. M., Aigner, T. G., & Richmond, B. J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *Journal of Neurophysiology*, *75*(3), 1061 LP-1073. Retrieved from http://jn.physiology.org/content/75/3/1061.abstract
Brewer, A. A., Press, W. A., Logothetis, N. K., & Wandell, B. A. (2002). Visual Areas in Macaque Cortex Measured Using Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *22*(23), 10416–10426.
Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth.
Bruce, C. J., & Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology*, *53*(3), 603–35. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3981231
Bruce, C. J., Goldberg, M. E., Bushnell, M. C., & Stanton, G. B. (1985). Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. *Journal of Neurophysiology*, *54*(3), 714–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4045546
Buschman, T. J., & Miller, E. K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science*, *315*(5820).
Buser, P. (2006). Slowly forgetting the Pavlovian adventure? *Comptes Rendus - Biologies*, *329*(5–6), 398–405. https://doi.org/10.1016/j.crvi.2006.03.010
Bystritsky, A., Korb, A. S., Douglas, P. K., Cohen, M. S., Melega, W. P., Mulgaonkar, A. P., … Yoo, S.-S. (2011). A review of low-intensity focused ultrasound pulsation. *Brain Stimulation*, *4*(3), 125–136. https://doi.org/10.1016/j.brs.2011.03.007
Cai, X., & Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(11), 3791–808. https://doi.org/10.1523/JNEUROSCI.3864-11.2012
Carmichael, S. T., & Price, J. L. (1995a). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *The Journal of Comparative Neurology*, *363*(4), 615–641. https://doi.org/10.1002/cne.903630408
Carmichael, S. T., & Price, J. L. (1995b). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *The Journal of Comparative Neurology*, *363*(4), 642–664.

https://doi.org/10.1002/cne.903630409
Caspers, S., Eickhoff, S. B., Geyer, S., Scheperjans, F., Mohlberg, H., Zilles, K., & Amunts, K. (2008). The human inferior parietal lobule in stereotaxic space. *Brain Structure and Function*, *212*(6), 481–495. https://doi.org/10.1007/s00429-008-0195-z
Cavada, C., & Goldman-Rakic, P. S. (1989a). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *The Journal of Comparative Neurology*, *287*(4), 393–421. https://doi.org/10.1002/cne.902870402
Cavada, C., & Goldman-Rakic, P. S. (1989a). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *The Journal of Comparative Neurology*, *287*(4), 393–421. https://doi.org/10.1002/cne.902870402
Cavada, C., & Goldman-Rakic, P. S. (1989b). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *The Journal of Comparative Neurology*, *287*(4), 422–445. https://doi.org/10.1002/cne.902870403
Cavada, C., & Goldman-Rakic, P. S. (1989b). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *The Journal of Comparative Neurology*, *287*(4), 422–45. https://doi.org/10.1002/cne.902870403
Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/http://dx.doi.org/10.1016/0040-5809(76)90040-X
Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. P., Bell, A. H., Walton, M. E., … Nakanishi, S. (2015). Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, *87*(5), 1106–1118. https://doi.org/10.1016/j.neuron.2015.08.018
Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2200075
Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *The Journal of Neuroscience*, *27*(2), 371 LP-378. Retrieved from http://www.jneurosci.org/content/27/2/371.abstract
Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology*, *76*(5), 2841–2852.
Corbetta, M., Kincade, M. J., & Shulman, G. L. (2002). Two neural systems for visual orienting, and the pathophysiology of unilateral spatial neglect. In H.-O. Karnath, A. D. Milner, & G. Vallar (Eds.), *The Cognitive and Neural Bases of Spatial Neglect*. Oxford: Oxford University Press.
Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(March).

https://doi.org/10.1038/nrn755
Corbetta, M., & Shulman, G. L. (2011). Spatial Neglect and Attention Networks. *Annual Review of Neuroscience*, *34*(1), 569–599. https://doi.org/10.1146/annurev-neuro-061010-113731
Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300. https://doi.org/10.1016/j.tics.2006.05.004
Cox, R. T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics*, *14*(1), 1. https://doi.org/10.1119/1.1990764
CROSBY, E. C., YOSS, R. E., & HENDERSON, J. W. (1952). The mammalian midbrain and isthmus regions. Part II. The fiber connections. D. The pattern for eye movements on the frontal eye field and the discharge of specific portions of this field to and through midbrain levels. *The Journal of Comparative Neurology*, *97*(2), 357–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12999993
D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science*, *319*(5867), 1264–1267. https://doi.org/10.1126/science.1150605
Damasio, A. R. (1994). Descartes' error and the future of human life. *Scientific American*, *271*(4), 144. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7939563
Dayan, P. (2012). Twenty-Five Lessons from Computational Neuromodulation. *Neuron*. https://doi.org/10.1016/j.neuron.2012.09.027
Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nat Neurosci*.
De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2012). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279
Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring. *The Journal of Neuroscience*, *25*(50), 11730 LP-11737. Retrieved from http://www.jneurosci.org/content/25/50/11730.abstract
Deffieux, T., Younan, Y., Wattiez, N., Tanter, M., Pouget, P., & Aubry, J.-F. (2013). Low-Intensity Focused Ultrasound Modulates Monkey Visuomotor Behavior. *Current Biology*, *23*(23), 2430–2433. https://doi.org/10.1016/j.cub.2013.10.029
Delaney, K., & Gelperi, A. (1986). Post-ingestive food-aversion learning to amino acid deficient diets by the terrestrial slug Limax maximus. *J Comp Physiol A*, *159*(3), 281–295.
Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.
Dubowitz, D. J., Cheng, D.-Y., Atkinson, D. J., Grieve, K. L., Gilliken, B., Bradley Jr, W. G., & Andersen, R. A. (1998). Functional magnetic resonance imaging in macaque cortex. *NeuroReport*, *9*, 2213–2218.
Dum, R. P., & Strick, P. L. (1996). Spinal cord terminations of the medial wall motor areas in macaque monkeys. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(20), 6513–25. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/8815929
Ebitz, R. B., & Platt, M. L. (2015). Neuronal Activity in Primate Dorsal Anterior Cingulate Cortex Signals Task Conflict and Predicts Adjustments in Pupil-Linked Arousal. *Neuron*, *85*(3), 628–640. https://doi.org/10.1016/j.neuron.2014.12.053
Engelmann, J. B., & Pessoa, L. (2007). Motivation sharpens exogenous spatial attention. *Emotion*, *7*(3), 668–674. https://doi.org/10.1037/1528-3542.7.3.668
Ferrier, D. (1876). *The Functions of the Brain*. New York: GP Putnam's Sons.
Fox, P. T., Fox, J. M., Raichle, M. E., & Burde, R. M. (1985). The role of cerebral cortex in the generation of voluntary saccades: a positron emission tomographic study. *Journal of Neurophysiology*, *54*(2), 348–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3875696
Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature*, *443*(7107), 85–8. https://doi.org/10.1038/nature05078
Freedman, L. J., Insel, T. R., & Smith, Y. (2000). Subcortical projections of area 25 (subgenual cortex) of the macaque monkey. *The Journal of Comparative Neurology*, *421*(2), 172–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10813780
Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787
Friston, K., Kiebel, S., Barlow, H. B., Feynman, R. P., Neal, R. M., Hinton, G. E., & Neisser, U. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–21. https://doi.org/10.1098/rstb.2008.0300
Gallagher, M., McMahan, R. W., & Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(15), 6610–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10414988
Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1297–1310. https://doi.org/10.1016/j.neubiorev.2013.03.023
Gillebert, C. R., Mantini, D., Peeters, R., Dupont, P., & Vandenberghe, R. (2013). Cytoarchitectonic mapping of attentional selection and reorienting in parietal cortex. *NeuroImage*, *67*, 257–272. https://doi.org/http://dx.doi.org/10.1016/j.neuroimage.2012.11.026
Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, *404*(6776), 390–394.
Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*(1), 535–74. https://doi.org/10.1146/annurev.neuro.29.051605.113038
Gottfried, J. A., & Zelano, C. (2011). The value of identity: olfactory notes on orbitofrontal cortex function. *Annals of the New York Academy of Sciences*, *1239*(1), 138–148. https://doi.org/10.1111/j.1749-6632.2011.06268.x
Gottlieb, J. P., MacAvoy, M. G., & Bruce, C. J. (1994). Neural responses related to smooth-pursuit eye movements and their correspondence with electrically elicited smooth eye movements in the primate frontal eye field. *Journal of Neurophysiology*, *72*(4), 1634–53. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/7823092
Gottlieb, J., & Snyder, L. H. (2010). Spatial and non-spatial functions of the parietal cortex. *Current Opinion in Neurobiology*, *20*(6), 731–740. https://doi.org/10.1016/j.conb.2010.09.015
Haber, S. N., & Behrens, T. E. J. (2014). The Neural Network Underlying Incentive-Based Learning: Implications for Interpreting Circuit Disruptions in Psychiatric Disorders. *Neuron*, *83*(5), 1019–1039. https://doi.org/http://dx.doi.org/10.1016/j.neuron.2014.08.031
Haber, S. N., Kim, K.-S., Mailly, P., & Calzavara, R. (2006). Reward-Related Cortical Inputs Define a Large Striatal Region in Primates That Interface with Associative Cortical Connections, Providing a Substrate for Incentive-Based Learning. *Journal of Neuroscience*, *26*(32), 8368–8376. https://doi.org/10.1523/JNEUROSCI.0271-06.2006
Haken, H. (1978). *Synergetics : an Introduction Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology*. Springer Berlin Heidelberg.
Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. *Journal of Neuroscience*, *28*(22), 5623–5630. https://doi.org/10.1523/JNEUROSCI.1309-08.2008
Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P., & Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences*, *108*(44), 18120–18125. https://doi.org/10.1073/pnas.1109322108
Harlow, J. (1868). Recovery after passage of an iron bar through the head. *Publ Massachusetts Med Soc*, (2), 329–346.
Hartwigsen, G., Bestmann, S., Ward, N. S., Woerbel, S., Mastroeni, C., Granert, O., & Siebner, H. R. (2012). Left Dorsal Premotor Cortex and Supramarginal Gyrus Complement Each Other during Rapid Action Reprogramming. *The Journal of Neuroscience*, *32*(46), 16162 LP-16171. Retrieved from http://www.jneurosci.org/content/32/46/16162.abstract
Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise Signals in Anterior Cingulate Cortex: Neuronal Encoding of Unsigned Reward Prediction Errors Driving Adjustment in Behavior. *Journal of Neuroscience*, *31*(11), 4178–4187. https://doi.org/10.1523/JNEUROSCI.4652-10.2011
Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939. https://doi.org/10.1038/nn.2856
He, S. Q., Dum, R. P., & Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *15*(5 Pt 1), 3284–306. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7538558
Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*(4), 814 LP-832. Retrieved from http://jn.physiology.org/content/61/4/814.abstract
Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H., &

Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat Neurosci*, *7*(5), 497–498. Retrieved from http://dx.doi.org/10.1038/nn1238

Hong, S., & Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral habenula. *Neuron*, *60*(4), 720–9. https://doi.org/10.1016/j.neuron.2008.09.035

Hong, S., & Hikosaka, O. (2013). Diverse sources of reward value signals in the basal ganglia nuclei transmitted to the lateral habenula in the monkey. *Frontiers in Human Neuroscience*, *7*, 778. https://doi.org/10.3389/fnhum.2013.00778

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences*, *112*(16), 5195–5200. https://doi.org/10.1073/pnas.1503550112

Hrupka, B. J., Lin, Y. M., Gietzen, D. W., & Rogers, Q. R. (1997). Small Changes in Essential Amino Acid Concentrations Alter Diet Selection in Amino Acid–Deficient Rats. *The Journal of Nutrition* , *127*(5), 777–784. Retrieved from http://jn.nutrition.org/content/127/5/777.abstract

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1986). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys: I. Subcortical connections. *The Journal of Comparative Neurology*, *253*(4), 415–439. https://doi.org/10.1002/cne.902530402

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1987). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys II. cortical connections. *The Journal of Comparative Neurology*, *265*(3), 332–361. https://doi.org/10.1002/cne.902650304

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), 470–476. https://doi.org/10.1038/nn.3017

Hutchison, R. M., Gallivan, J. P., Culham, J. C., Gati, J. S., Menon, R. S., & Everling, S. (2012). Functional connectivity of the frontal eye fields in humans and macaque monkeys investigated with resting-state fMRI. *Journal of Neurophysiology*, *107*(9), 2463 LP-2474. Retrieved from http://jn.physiology.org/content/107/9/2463.abstract

Izquierdo, A., Suda, R. K., & Murray, E. A. (2004). Bilateral Orbital Prefrontal Cortex Lesions in Rhesus Monkeys Disrupt Choices Guided by Both Reward Value and Reward Contingency. *Journal of Neuroscience*, *24*(34), 7540–7548. https://doi.org/10.1523/JNEUROSCI.1921-04.2004

Jabbi, M., Kohn, P. D., Nash, T., Ianni, A., Coutlee, C., Holroyd, T., … Berman, K. F. (2015). Convergent BOLD and Beta-Band Activity in Superior Temporal Sulcus and Frontolimbic Circuitry Underpins Human Emotion Cognition. *Cerebral Cortex (New York, N.Y. : 1991)*, *25*(7), 1878–88. https://doi.org/10.1093/cercor/bht427

Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H., & Curtis, C. E. (2012). Prioritized Maps of Space in Human Frontoparietal Cortex. *Journal of Neuroscience*, *32*(48), 17382–17390. https://doi.org/10.1523/JNEUROSCI.3810-12.2012

Jocham, G., Brodersen, K. H. H., Constantinescu, A. O. O., Kahn, M. C. C., Ianni, A. M.,

Walton, M. E. E., … Behrens, T. E. E. J. (2016). Reward-Guided Learning with and without Causal Attribution. *Neuron*, *90*(1), 177–190. https://doi.org/10.1016/j.neuron.2016.02.018

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-Mediated Reinforcement Learning Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Value-Based Choices. *Journal of Neuroscience*, *31*(5).

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(23), 7489–96. https://doi.org/10.1523/JNEUROSCI.0349-09.2009

Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus--reinforcement associations. *Experimental Neurology*, *36*(2), 362–77. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4626489

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science (New York, N.Y.)*, *338*(6109), 953–6. https://doi.org/10.1126/science.1227489

Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal Striatal–midbrain Connectivity in Humans Predicts How Reinforcements Are Used to Guide Decisions. *Journal of Cognitive Neuroscience*, *21*(7), 1332–1345. https://doi.org/10.1162/jocn.2009.21092

Kamin, L. J. (1969). Selective association and con- ditioning. In N. J. Mackintosh & H. W. K (Eds.), *Fundamental Issues in Instrumental Learning* (pp. 42–64). Halifax, NS: Dalhousie Univ. Press.

Karlsson, M. P., Tervo, D. G. R., & Karpova, A. Y. (2012). Network Resets in Medial Prefrontal Cortex Mark the Onset of Behavioral Uncertainty. *Science*, *338*(6103), 135 LP-139. Retrieved from http://science.sciencemag.org/content/338/6103/135.abstract

Kastner, S., Desimone, K., Konen, C. S., Szczepanski, S. M., Weiner, K. S., Schneider, K. A., … Schneider, K. (2007). Topographic Maps in Human Frontal Cortex Revealed in Memory-Guided Saccade and Spatial Working-Memory Tasks. *J Neurophysiol*, *97*, 3494–3507. https://doi.org/10.1152/jn.00010.2007

Keliris, G. A., Shmuel, A., Ku, S.-P., Pfeuffer, J., Oeltermann, A., Steudel, T., & Logothetis, N. K. (2007). Robust controlled functional MRI in alert monkeys at high magnetic field: Effects of jaw and body movements. *Neuroimage*, *36*, 550–570.

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589. https://doi.org/10.1038/nn.2961

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nat Neurosci*, *9*(7), 940–7. https://doi.org/10.1038/nn1724

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–7. https://doi.org/10.1038/nn1724

King, R. L., Brown, J. R., Newsome, W. T., & Pauly, K. B. (2013). Effective Parameters for Ultrasound-Induced In Vivo Neurostimulation. *Ultrasound in Medicine &*

Biology, *39*(2), 312–331. https://doi.org/10.1016/j.ultrasmedbio.2012.09.009
Klein-Flügge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., & Behrens, T. E. J. (2013). Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(7), 3202–11. https://doi.org/10.1523/JNEUROSCI.2532-12.2013
Knight, F. H. (1921). *Risk, Uncertainty and Profit*. *Hart, Schaffner & Marx, Boston, MA*.
Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007
Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*(17), 3683–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11726774
Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930
Kolling, N., Behrens, T., Wittmann, M., & Rushworth, M. (2016). Multiple signals in anterior cingulate cortex. *Current Opinion in Neurobiology*, *37*, 36–43. https://doi.org/10.1016/j.conb.2015.12.007
Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci*, *19*(10), 1280–1285. Retrieved from http://dx.doi.org/10.1038/nn.4382
Kolling, N., Wittmann, M., & Rushworth, M. F. S. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*, *81*(5), 1190–1202. https://doi.org/10.1016/j.neuron.2014.01.033
Kolster, H., Janssens, T., Orban, G. A., & Vanduffel, W. (2014). The Retinotopic Organization of Macaque Occipitotemporal Cortex Anterior to V4 and Caudoventral to the Middle Temporal (MT) Cluster. *Journal of Neuroscience*, *34*(31), 10168–10191. https://doi.org/10.1523/JNEUROSCI.3288-13.2014
Kolster, H., Mandeville, J. B., Arsenault, J. T., Ekstrom, L. B., Wald, L. L., & Vanduffel, W. (2009). Visual field map clusters in macaque extrastriate visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(21), 7031–9. https://doi.org/10.1523/JNEUROSCI.0518-09.2009
KRINGELBACH, M., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*(5), 341–372. https://doi.org/10.1016/j.pneurobio.2004.03.006
Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86. https://doi.org/10.1214/aoms/1177729694
Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskopf, R. M., Poncelet, B. P., … Rosen, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *PNAS*, *89*, 5675–5679.
Lee, W., Chung, Y. A., Jung, Y., Song, I.-U., & Yoo, S.-S. (2016). Simultaneous acoustic

stimulation of human primary and secondary somatosensory cortices using transcranial focused ultrasound. *BMC Neuroscience*, *17*(1), 68. https://doi.org/10.1186/s12868-016-0303-6

Lee, W., Kim, H.-C., Jung, Y., Chung, Y. A., Song, I.-U., Lee, J.-H., & Yoo, S.-S. (2016). Transcranial focused ultrasound stimulation of human primary visual cortex. *Scientific Reports*, *6*, 34026. https://doi.org/10.1038/srep34026

Leichnetz, G. R., Smith, D. J., & Spencer, R. F. (1984). Cortical projections to the paramedian tegmental and basilar pons in the monkey. *The Journal of Comparative Neurology*, *228*(3), 388–408. https://doi.org/10.1002/cne.902280307

Leite, F. P., Tsao, D. Y., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L. L., … Mandeville, J. B. (2002). Repeated fMRI Using Iron Oxide Contrast Agent in Awake, Behaving Macaques at 3 Tesla. *NeuroImage*, *16*, 283–294.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Lewis, J. W., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *The Journal of Comparative Neurology*, *428*(1), 112–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11058227

Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*, *67*(1), 145 LP-163. Retrieved from http://jn.physiology.org/content/67/1/145.abstract

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*, 869–878. https://doi.org/10.1038/nature06976

Logothetis, N. K., Pauls, J. M., Augath, M. A., Trinath, T., & Oeltermann, a. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–7. https://doi.org/10.1038/35084005

Lopatina, N., McDannald, M. A., Styer, C. V, Sadacca, B. F., Cheer, J. F., & Schoenbaum, G. (2015). Lateral orbitofrontal neurons acquire responses to upshifted, downshifted, or blocked cues during unblocking. *eLife*, *4*, e11299. https://doi.org/10.7554/eLife.11299

Lynd-Balta, E., & Haber, S. N. (1994). The organization of midbrain projections to the striatum in the primate: sensorimotor-related striatum versus ventral striatum. *Neuroscience*, *59*(3), 625–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7516506

Mackey, S., & Petrides, M. (2010). Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *The European Journal of Neuroscience*, *32*(11), 1940–50. https://doi.org/10.1111/j.1460-9568.2010.07465.x

Magri, C., Schridde, U., Murayama, Y., Panzeri, S., & Logothetis, N. K. (2012). The Amplitude and Timing of the BOLD Signal Reflects the Relationship between Local Field Potential Power at Different Frequencies. *Journal of Neuroscience*, *32*(4).

Mandeville, J. B., Choi, J.-K., Jarraya, B., Rosen, B. R., Jenkins, B. G., & Vanduffel, W.

(2011). fMRI of cocaine self-administration in macaques reveals functional inhibition of basal ganglia. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *36*(6), 1187–98. https://doi.org/10.1038/npp.2011.1

Markowitsch, H. J., & Pritzel, M. (1976). Reward related neurons in cat association cortex. *Brain Research*, *111*(1), 185–188. https://doi.org/http://dx.doi.org/10.1016/0006-8993(76)91061-1

Mars, R. B., Jbabdi, S., Sallet, J., O'Reilly, J. X., Croxson, P. L., Olivier, E., … Rushworth, M. F. S. (2011). Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Parietal Cortex and Comparison with Human and Macaque Resting-State Functional Connectivity. *Journal of Neuroscience*, *31*(11), 4087–4100. https://doi.org/10.1523/JNEUROSCI.5102-10.2011

Mars, R. B., Piekema, C., Coles, M. G. H., Hulstijn, W., & Toni, I. (2007). On the Programming and Reprogramming of Actions. *Cerebral Cortex* , *17*(12), 2972–2979. https://doi.org/10.1093/cercor/bhm022

Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, *447*(7148), 1111–1115. https://doi.org/10.1038/nature05860

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–841. https://doi.org/10.1038/nature08028

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007a). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*(5), 647–656. https://doi.org/10.1038/nn1890

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007b). Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci*, *10*(5), 647–656. Retrieved from http://dx.doi.org/10.1038/nn1890

McDannald, M. A., Esber, G. R., Wegener, M. A., Wied, H. M., Liu, T.-L., Stalnaker, T. A., … Schoenbaum, G. (2014). Orbitofrontal neurons acquire responses to "valueless" Pavlovian cues during unblocking. *eLife*, *3*, e02653. https://doi.org/10.7554/eLife.02653

Mesulam, M. M., Nobre, A. C., Kim, Y.-H., Parrish, T. B., & Gitelman, D. R. (2001). Heterogeneity of Cingulate Contributions to Spatial Attention. *NeuroImage*, *13*(6), 1065–1072. https://doi.org/10.1006/nimg.2001.0768

Milner, A. (1996). Neglect, extinction and the cortical streams of visual processing. In P. Thier & H.-O. Karnath (Eds.), *Parietal Lobe Contributions to Orientation in 3D Space* (Experiment). Berlin: Springer.

Milner, D., & Goodale, M. (1995). *The Visual Brain in Action*. *The Visual Brain in Action*. Oxford, United Kingdom: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524724.001.0001

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Supporting Online Material for Patches with Links: A Unified System for Processing Faces in the Macaque Temporal Lobe. *Science*. https://doi.org/10.1126/science.1157436

Molenberghs, P., Gillebert, C. R., Peeters, R., & Vandenberghe, R. (2008). Convergence between Lesion-Symptom Mapping and Functional Magnetic Resonance Imaging of Spatially Selective Attention in the Intact Brain. *The Journal of Neuroscience*, *28*(13), 3359 LP-3373. Retrieved from

http://www.jneurosci.org/content/28/13/3359.abstract
Molenberghs, P., Mesulam, M. M., Peeters, R., & Vandenberghe, R. R. C. (2007). Remapping Attentional Priorities: Differential Contribution of Superior Parietal Lobule and Intraparietal Sulcus. *Cerebral Cortex* , *17*(11), 2703–2712. https://doi.org/10.1093/cercor/bhl179
Monosov, I. E., & Hikosaka, O. (2013). Selective and graded coding of reward uncertainty by neurons in the primate anterodorsal septal region. *Nature Neuroscience*, *16*(6), 756–762. https://doi.org/10.1038/nn.3398
Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, *36*(2), 265–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12383781
Morecraft, R. J., Stilwell-Morecraft, K. S., Cipolloni, P. B., Ge, J., McNeal, D. W., & Pandya, D. N. (2012). Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Research Bulletin*, *87*(4–5), 457–97. https://doi.org/10.1016/j.brainresbull.2011.12.005
Mort, D. J., Malhotra, P., Mannan, S. K., Rorden, C., Pambakian, A., Kennard, C., & Husain, M. (2003). The anatomy of visual neglect. *Brain*, *126*(9), 1986–1997. https://doi.org/10.1093/brain/awg200
Mort, D. J., Perry, R. J., Mannan, S. K., Hodgson, T. L., Anderson, E., Quest, R., … Kennard, C. (2003). Differential cortical activation during voluntary and reflexive saccades in man. *NeuroImage*, *18*(2), 231–246. https://doi.org/10.1016/S1053-8119(02)00028-9
MOTT, F. W., & SCHAEFER, E. A. (1890). ON ASSOCIATED EYE-MOVEMENTS PRODUCED BY CORTICAL FARADIZATION OF THE MONKEY'S BRAIN. *Brain*, *13*(2).
Murray, E. A., Moylan, E. J., Saleem, K. S., Basile, B. M., Turchi, J., Barbas, H., … Andersen, R. (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *eLife*, *4*, 353–375. https://doi.org/10.7554/eLife.11695
Nakamura, K., Mikami, A., & Kubota, K. (1992). Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *Journal of Neurophysiology*, *67*(6), 1447 LP-1463. Retrieved from http://jn.physiology.org/content/67/6/1447.abstract
Nakamura, K., Roesch, M. R., & Olson, C. R. (2004). Neuronal Activity in Macaque SEF and ACC During Performance of Tasks Involving Conflict. *Journal of Neurophysiology*, *93*(2), 884–908. https://doi.org/10.1152/jn.00305.2004
Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046. https://doi.org/10.1038/nn.3130
Neubert, F.-X., Mars, R. B., Sallet, J., & Rushworth, M. F. S. (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proceedings of the National Academy of Sciences*, *112*(20), E2695–E2704. https://doi.org/10.1073/pnas.1410767112
Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004). Sensitivity of Electrophysiological Activity from Medial Frontal Cortex to Utilitarian and Performance Feedback. *Cerebral Cortex*, *14*(7), 741–747.

https://doi.org/10.1093/cercor/bhh034
Nishijo, H., Ono, T., & Nishino, H. (1988). Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *The Journal of Neuroscience*, *8*(10), 3570 LP-3583. Retrieved from http://www.jneurosci.org/content/8/10/3570.abstract
Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4
Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, *12*(7), 265–272. https://doi.org/10.1016/j.tics.2008.03.006
Nobre, A. C., Coull, J. T., Frith, C. D., & Mesulam, M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat Neurosci*, *2*(1), 11–12. Retrieved from http://dx.doi.org/10.1038/4513
Noonan, M. P., Mars, R. B., & Rushworth, M. F. S. (2011). Distinct Roles of Three Frontal Cortical Areas in Reward-Guided Behavior. *The Journal of Neuroscience*, *31*(40), 14399 LP-14412. Retrieved from http://www.jneurosci.org/content/31/40/14399.abstract
Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences* , *107*(47), 20547–20552. https://doi.org/10.1073/pnas.1012246107
O'Doherty, J. P. (2016). Multiple Systems for the Motivational Control of Behavior and Associated Neural Substrates in Humans. In E. H. Simpson & P. D. Balsam (Eds.), *Current Topics in Behavioral Neurosciences: Behavioral Neuroscience of Motivation* (pp. 291–312). Springer International Publishing.
O'Doherty, J. P. (2016). Multiple Systems for the Motivational Control of Behavior and Associated Neural Substrates in Humans. In E. H. Simpson & P. D. Balsam (Eds.), *Behavioral Neuroscience of Motivation* (pp. 291–312). Cham: Springer International Publishing. https://doi.org/10.1007/7854_2015_386
O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2016). Learning, Reward, and Decision Making. *Annual Review of Psychology*. https://doi.org/10.1146/annurev-psych-010416-044216
O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7
O'Driscoll, G. A., Alpert, N. M., Matthysse, S. W., Levy, D. L., Rauch, S. L., & Holzman, P. S. (1995). Functional neuroanatomy of antisaccade eye movements investigated with positron emission tomography. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(3), 925–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7846080
O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), E3660--9. https://doi.org/10.1073/pnas.1305373110
Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance

imaging with contrast dependent on blood oxygenation. *PNAS*, *87*, 9868–9872.
Ogawa, S., Tank, D. W., Menon, R. S., Ellermannc, J. M., Kim, S.-G., Merkle, H., & Uğurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *PNAS*, *89*, 5951–5955.
OLDS, J., & MILNER, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13233369
Ongür, D., An, X., & Price, J. L. (1998). Prefrontal cortical projections to the hypothalamus in macaque monkeys. *The Journal of Comparative Neurology*, *401*(4), 480–505. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9826274
Ongür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex (New York, N.Y. : 1991)*, *10*(3), 206–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10731217
Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, *8*(7), 315–324. https://doi.org/10.1016/j.tics.2004.05.009
Oristaglio, J., Schneider, D. M., Balan, P. F., & Gottlieb, J. (2006). Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(32), 8310–9. https://doi.org/10.1523/JNEUROSCI.1779-06.2006
Padoa-Schioppa, C. (2011). Neurobiology of Economic Choice: A Good-Based Model. *Annual Review of Neuroscience*, *34*(1), 333–359. https://doi.org/10.1146/annurev-neuro-061010-113648
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676
Paradiso, M. A. (1999). Monkey business builds a bridge to the human brain. *Nature Neuroscience*, *2*(6), 491–492.
Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex : anatomy, evolution, and the origin of insight*. Oxford Univ. Press.
Patel, G. H., Yang, D., Jamerson, E. C., Snyder, L. H., Corbetta, M., & Ferrera, V. P. (2015). Functional evolution of new and expanded attention networks in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(30), 9454–9. https://doi.org/10.1073/pnas.1420395112
Pauling, L., & Coryell, C. D. (1936). The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. *PNAS*, *22*, 210–216. https://doi.org/10.1002/jcp.1030080107
Paus, T. (1996). Location and function of the human frontal eye-field: A selective review. *Neuropsychologia*, *34*(6), 475–483. https://doi.org/http://dx.doi.org/10.1016/0028-3932(95)00134-4
Paus, T., Petrides, M., Evans, A. C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a

positron emission tomography study. *Journal of Neurophysiology*, *70*(2), 453–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8410148

Pavlov, P. I. (1927). *Conditioned Reflexes*. Oxford, United Kingdom: Oxford Univ. Press.

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward Modulates Attention Independently of Action Value in Posterior Parietal Cortex. *Journal of Neuroscience*, *29*(36).

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. https://doi.org/http://dx.doi.org/10.1016/j.tics.2009.01.006

Petrides, M., & Pandya, D. N. (1994). Comparative architectonic analysis of the human and the macaque frontal cortex. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology Vol. 9* (pp. 17–58). Amsterdam: Elsevier.

Pfeuffer, J., Shmuel, A., Keliris, G. A., Steudel, T., Merkle, H., & Logothetis, N. K. (2007). Functional MR imaging in the awake monkey: effects of motion on dynamic off-resonance and processing strategies. *Magnetic Resonance Imaging*, *25*, 869–882. https://doi.org/10.1016/j.mri.2007.03.002

Phillips, A. G., & Fibiger, H. C. (1978). The role of dopamine in maintaining intracranial self-stimulation in the ventral tegmentum, nucleus accumbens, and medial prefrontal cortex. *Canadian Journal of Psychology*, *32*(2), 58–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/737577

Picard, N., & Strick, P. L. (1991). Motor areas of the medial wall: a review of their location and functional activation. *Cerebral Cortex*, *6*(3), 342–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8670662

Pierrot-Deseilligny, C., & Muri, R. (1996). Posterior parietal cortex control of saccades in man. In P. Thier & H.-O. Karnath (Eds.), *Parietal Lobe Contributions to Orientation in 3D Space, Experimental Brain Research Series*. Berlin: Springer.

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(37), 9984–8. https://doi.org/10.1523/JNEUROSCI.2131-07.2007

Plassmann, H., O'Doherty, J. P., & Rangel, A. (2010). Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. *The Journal of Neuroscience*, *30*(32), 10799 LP-10808. Retrieved from http://www.jneurosci.org/content/30/32/10799.abstract

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238. https://doi.org/10.1038/22268

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238. https://doi.org/10.1038/22268

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of Pareital Injury on Covert Orienting of Attention 1. *The Journal of Neuroscience*, *4*(7),

1863–1874. https://doi.org/10.1136/jnnp.72.1.73
Pratt, W. E., & Mizumori, S. J. Y. (1998). Characteristics of basolateral amygdala neuronal firing on a spatial memory task involving differential reward. *Behavioral Neuroscience*. US: American Psychological Association. https://doi.org/10.1037/0735-7044.112.3.554
Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, (SEP), 1–12. https://doi.org/10.3389/fnins.2011.00115
Procyk, E., Wilson, C. R. E., Stoll, F. M., Faraut, M. C. M., Petrides, M., & Amiez, C. (2014). Midcingulate Motor Map and Feedback Detection: Converging Data from Humans and Monkeys. *Cerebral Cortex*, *26*(2), bhu213. https://doi.org/10.1093/cercor/bhu213
Pruessmann, K. P., Weiger, M., Scheidegger, M. B., & Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, *42*(5), 952–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10542355
Quilodran, R., Rothé, M., & Procyk, E. (2008a). Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. *Neuron*, *57*(2), 314–325. https://doi.org/10.1016/j.neuron.2007.11.031
Quilodran, R., Rothé, M., & Procyk, E. (2008b). Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. *Neuron*, *57*(2), 314–325. https://doi.org/http://dx.doi.org/10.1016/j.neuron.2007.11.031
Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357
Ravel, S., Legallet, E., & Apicella, P. (1999). Tonically active neurons in the monkey striatum do not preferentially respond to appetitive stimuli. *Experimental Brain Research*, *128*(4), 531–534. https://doi.org/10.1007/s002210050876
Rempel-Clower, N. L., & Barbas, H. (1998). Topographic organization of connections between the hypothalamus and prefrontal cortex in the rhesus monkey. *The Journal of Comparative Neurology*, *398*(3), 393–419. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9714151
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II Current Research and Theory* (Vol. 21, pp. 64–99). https://doi.org/10.1101/gr.110528.110
Roberts, A. C. (2006). Primate orbitofrontal cortex and adaptive behaviour. *Trends in Cognitive Sciences*, *10*(2), 83–90. https://doi.org/10.1016/j.tics.2005.12.002
Rogers, Q. R., & Harper, A. E. (1970). Selection of a solution containing histidine by rats fed a histidine-imbalanced diet. *J Comp Physiol Psychol*, *72*(1), 66–71.
Rolls, E. T., Everitt, B. J., & Roberts, A. (1996). The Orbitofrontal Cortex [and Discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *351*(1346), 1433–1444. https://doi.org/10.1098/rstb.1996.0128
Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, *63*(3), 592–606. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2329363
Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J.,

Walton, M. E., & Rushworth, M. F. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci*, *28*(51), 13775–13785. https://doi.org/10.1523/JNEUROSCI.3541-08.2008

Rudebeck, P. H., Mitz, A. R., Chacko, R. V., & Murray, E. A. (2013). Effects of Amygdala Lesions on Reward-Value Coding in Orbital and Medial Prefrontal Cortex. *Neuron*, *80*(6), 1519–1531. https://doi.org/10.1016/j.neuron.2013.09.036

Rudebeck, P. H., & Murray, E. A. (2014). The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron*, *84*(6), 1143–56. https://doi.org/10.1016/j.neuron.2014.10.049

Rushworth, M. F., Ellison, A., & Walsh, V. (2001). Complementary localization and lateralization of orienting and motor attention. *Nature Neuroscience*, *4*(6), 656–661. https://doi.org/10.1038/88492

Rushworth, M. F., Mars, R. B., & Summerfield, C. (2009). General mechanisms for making decisions? *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2009.02.005

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–97. https://doi.org/10.1038/nn2066

Rushworth, M. F. S., Behrens, T. E. J., & Johansen-Berg, H. (2006). Connection Patterns Distinguish 3 Regions of Human Parietal Cortex. *Cerebral Cortex* , *16*(10), 1418–1430. https://doi.org/10.1093/cercor/bhj079

Rushworth, M. F. S., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: One or many serial or parallel systems? *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2012.04.011

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal Cortex and Reward-Guided Learning and Decision-Making. *Neuron*. https://doi.org/10.1016/j.neuron.2011.05.014

Rushworth, M. F. S., Paus, T., & Sipila, P. K. (2001). Attention Systems and the Organization of the Human Parietal Cortex. *The Journal of Neuroscience*, *21*(14), 5262 LP-5271. Retrieved from http://www.jneurosci.org/content/21/14/5262.abstract

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.07.009

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience*, *30*(40).

Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., … Rushworth, M. F. S. (2011). Social Network Size Affects Neural Circuits in Macaques. *Science*, *334*(6056).

San-Galli, A., Varazzani, C., Abitbol, R., Pessiglione, M., & Bouret, S. (2016). Primate Ventromedial Prefrontal Cortex Neurons Continuously Encode the Willingness to Engage in Reward-Directed Behavior. *Cerebral Cortex* . https://doi.org/10.1093/cercor/bhw351

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews. Neuroscience*, *10*(3), 211–223.

https://doi.org/10.1038/nrn2573
Schall, J. D. (1997). Visuomotor Areas of the Frontal Lobe. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Cerebral Cortex - Volume 12: Extrastriate Cortex in Primates* (pp. 527–638). New York: Springer US. https://doi.org/10.1007/978-1-4757-9625-4
Schall, J. D., Hanes, D. P., Thompson, K. G., & King, D. J. (1995). Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *15*(10), 6905–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7472447
Schall, J., Morel, A., King, D., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *Journal of Neuroscience*, *15*(6).
Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. *PLOS Biology*, *10*(2), e1001266. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pbio.1001266
Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12383780
Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115. https://doi.org/10.1146/annurev.psych.56.091103.070229
Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of Neuroscience*, *12*(12), 4595 LP-4610. Retrieved from http://www.jneurosci.org/content/12/12/4595.abstract
Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306).
Schultz, W., & Romo, R. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *Journal of Neurophysiology*, *63*(3), 607–624. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2329364
Schultz, W., & Tremblay, L. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*(6729), 704–708. https://doi.org/10.1038/19525
Serences, J. T. (2008). Value-Based Modulations in Human Visual Cortex. *Neuron*, *60*(6), 1169–1181. https://doi.org/http://dx.doi.org/10.1016/j.neuron.2008.10.051
Serences, J. T., & Saproo, S. (2010). Population Response Profiles in Early Visual Cortex Are Biased in Favor of More Valuable Stimuli. *Journal of Neurophysiology*, *104*(1), 76 LP-87. Retrieved from http://jn.physiology.org/content/104/1/76.abstract
Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of Contralateral Space in Retinotopic Coordinates by a Parietal Cortical Area in Humans. *Science*, *294*, 1350–1354.
Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System*

Technical Journal, *27*(July 1928), 379–423. https://doi.org/10.1145/584091.584093

Shidara, M., Aigner, T. G., & Richmond, B. J. (1998). Neuronal Signals in the Monkey Ventral Striatum Related to Progress through a Predictable Series of Trials. *The Journal of Neuroscience*, *18*(7), 2613 LP-2625. Retrieved from http://www.jneurosci.org/content/18/7/2613.abstract

Shulman, G. L., McAvoy, M. P., Cowan, M. C., Astafiev, S. V, Tansy, A. P., d'Avossa, G., & Corbetta, M. (2003). Quantitative Analysis of Attention and Detection Signals During Visual Search. *Journal of Neurophysiology*, *90*(5), 3384–3397. https://doi.org/10.1152/jn.00343.2003

Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, *13*(2), 121. https://doi.org/10.1038/nrn3137

Smith, S., Jenkinson, M., Woolrich, M., Beckmann, C., Behrens, T., Johansen-Berg, H., … Matthews, P. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208-19. https://doi.org/10.1016/j.neuroimage.2004.07.051

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208-19. https://doi.org/10.1016/j.neuroimage.2004.07.051

Stalnaker, T. a, Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience*, *18*(5), 620–627. https://doi.org/10.1038/nn.3982

Stanton, G. B., Bruce, C. J., & Goldberg, M. E. (1995). Topography of Projections to Posterior Cortical Areas From the Macaque Frontal Eye Fields. *THE JOURNAL OF COMPARATIVE NEUROLOGY*, 353291–305.

Stanton, G. B., Deng, S.-Y., Goldberg, E. M., & McMullen, N. T. (1989). Cytoarchitectural characteristic of the frontal eye fields in macaque monkeys. *The Journal of Comparative Neurology*, *282*(3), 415–427. https://doi.org/10.1002/cne.902820308

Stefanacci, L., Reber, P., Costanza, J., Wong, E. C., Buxton, R. B., Zola, S., … Albright, T. D. (1998). fMRI of Monkey Visual Cortex. *Neuron*, *20*, 1051–1057.

Stephens, D. W., & Krebs, J. R. (1987). *Foraging Theory*. Princeton: Princeton University Press.

Stevenson, C. M., Brookes, M. J., & Morris, P. G. (2011). β-Band correlates of the fMRI BOLD response. *Human Brain Mapping*, *32*(2), 182–97. https://doi.org/10.1002/hbm.21016

Stoll, F. M., Wilson, C. R. E., Faraut, M. C. M., Vezoli, J., Knoblauch, K., & Procyk, E. (2016). The Effects of Cognitive Control and Time on Frontal Beta Oscillations. *Cerebral Cortex (New York, N.Y. : 1991)*, *26*(4), 1715–32. https://doi.org/10.1093/cercor/bhv006

Stone, S. P., Patel, P., Greenwood, R. J., & Halligan, P. W. (1992). Measuring visual neglect in acute stroke and predicting its recovery: the visual neglect recovery index. *Journal of Neurology, Neurosurgery, and Psychiatry*, *55*(6), 431–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1619406

Strait, C. E., Blanchard, T. C., & Hayden, B. Y. (2014). Reward value comparison via

mutual inhibition in ventromedial prefrontal cortex. *Neuron*, *82*(6), 1357–66. https://doi.org/10.1016/j.neuron.2014.04.032
Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching Behavior and the Representation of Value in the Parietal Cortex. *Science*, *304*(5678).
Sutton, R., & Barto, A. (1998). Reinforcement Learning: An Introduction.
Suzuki, H., & Azuma, M. (1983). Topographic studies on visual neurons in the dorsolateral prefrontal cortex of the monkey. *Experimental Brain Research*, *53*(1), 47–58. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6673997
Swisher, J. D., Halko, M. A., Merabet, L. B., McMains, S. A., & Somers, D. C. (2007). Visual Topography of Human Intraparietal Sulcus. *The Journal of Neuroscience*, *27*(20), 5326 LP-5337. Retrieved from http://www.jneurosci.org/content/27/20/5326.abstract
Thiel, C. M., Zilles, K., & Fink, G. R. (2004). Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: an event-related fMRI study. *NeuroImage*, *21*(1), 318–328. https://doi.org/10.1016/j.neuroimage.2003.08.044
Thorndike, E. L. (1911). *Animal Intelligence: Ex- perimental Studies*. New York: MacMillan.
Thorpe, S. J., Rolls, E. T., & Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res*, *49*(1).
Thulborn, K. R., Waterton, J. C., Matthews, P. M., & Radda, G. K. (1982). Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochim. Biophys. Acta*, *714*(2), 265–270.
Tinklepaugh, O. L. (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology*, *8*(3), 197–236. https://doi.org/10.1037/h0075798
Tobler, P. N., O'doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, *95*(1), 301–10. https://doi.org/10.1152/jn.00762.2005
Tufail, Y., Yoshihiro, A., Pati, S., Li, M. M., & Tyler, W. J. (2011). Ultrasonic neuromodulation by brain stimulation with transcranial ultrasound. *Nature Protocols*, *6*(9), 1453–1470. https://doi.org/10.1038/nprot.2011.371
Tziortzi, A. C., Haber, S. N., Searle, G. E., Tsoumpas, C., Long, C. J., Shotbolt, P., … Gunn, R. N. (2014). Connectivity-Based Functional Analysis of Dopamine Release in the Striatum Using Diffusion-Weighted MRI and Positron Emission Tomography. *Cerebral Cortex*, *24*(5), 1165–1177. https://doi.org/10.1093/cercor/bhs397
Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of Performance Monitoring: A Dissociation of Error Processing and Response Competition Revealed by Event-Related fMRI and ERPs. *NeuroImage*, *14*(6), 1387–1401. https://doi.org/10.1006/nimg.2001.0935
Van Essen, D. C. (2004). Surface-based approaches to spatial localization and registration in primate cerebral cortex. *Neuroimage*, *23*, 97–107.
Vandenberghe, R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2001). Functional Specificity of Superior Parietal Mediation of Spatial Shifting. *NeuroImage*, *14*(3), 661–673. https://doi.org/10.1006/nimg.2001.0860
Vandenberghe, R., Molenberghs, P., & Gillebert, C. R. (2012). Spatial attention

deficits in humans: The critical role of superior compared to inferior parietal lesions. *Neuropsychologia*, *50*(6), 1092–1103. https://doi.org/10.1016/j.neuropsychologia.2011.12.016

Vanduffel, W., Fize, D., Mandeville, J. B., Nelissen, K., Hecke, P. Van, Rosen, B. R., … Orban, G. A. (2001). Visual Motion Processing Neurotechnique Investigated Using Contrast Agent- Enhanced fMRI in Awake Behaving Monkeys. *Neuron*, *32*(November), 565–577.

Vogt, B. A., & Pandya, D. N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. *The Journal of Comparative Neurology*, *262*(2), 271–289. https://doi.org/10.1002/cne.902620208

Vogt, B. A., Pandya, D. N., & Rosene, D. L. (1987). Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *The Journal of Comparative Neurology*, *262*(2), 256–270. https://doi.org/10.1002/cne.902620207

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and Ventral Attention Systems: Distinct Neural Circuits but Collaborative Roles. *The Neuroscientist*, *20*(2), 150–159. https://doi.org/10.1177/1073858413494269

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K. J., & Stephan, K. E. (2014). Spatial attention, precision, and bayesian inference: A study of saccadic response speed. *Cerebral Cortex*, *24*(6), 1436–1450. https://doi.org/10.1093/cercor/bhs418

Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, *15*(1), 13–19. https://doi.org/10.1038/nn.2956

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, *65*(6), 927–39. https://doi.org/10.1016/j.neuron.2010.02.027

Walton, M. E., Devlin, J. T., & Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*(11), 1259–1265. https://doi.org/10.1038/nn1339

Wang, X.-J. (2008). Decision Making in Recurrent Neuronal Circuits. *Neuron*, *60*(2), 215–234. https://doi.org/10.1016/j.neuron.2008.09.034

Wang, Y., Cummings, S. L., & Gietzen, D. W. (1996). Temporal-spatial pattern of c-fos expression in the rat brain in response to indispensable amino acid deficiency I. The initial recognition phase. *Molecular Brain Research*, *40*(1), 27–34. https://doi.org/http://dx.doi.org/10.1016/0169-328X(96)00032-0

Watanabe, M. (1996). Reward expectancy in primate prefrental neurons. *Nature*, *382*(6592), 629–632. Retrieved from http://dx.doi.org/10.1038/382629a0

Watkins, C. J. C. H., & Dayan, P. (1992). Q-Learning. *Machine Learning*, *8*, 279–292.

Williams, S. M., & Goldman-Rakic, P. S. (1998). Widespread origin of the primate mesofrontal dopamine system. *Cerebral Cortex (New York, N.Y. : 1991)*, *8*(4), 321–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9651129

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, *81*(2), 267–279. https://doi.org/http://dx.doi.org/10.1016/j.neuron.2013.11.005

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494. https://doi.org/10.1038/nrn1406

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747. https://doi.org/10.1016/j.neuroimage.2003.12.023

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A Three-Dimensional Statistical Analysis for CBF Activation Studies in Human Brain. *Journal of Cerebral Blood Flow & Metabolism*, *12*(6), 900–918. https://doi.org/10.1038/jcbfm.1992.127

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, *447*(7148), 1075–1080. https://doi.org/10.1038/nature05852

Yoo, S.-S., Bystritsky, A., Lee, J.-H., Zhang, Y., Fischer, K., Min, B.-K., … Jolesz, F. A. (2011). Focused ultrasound modulates region-specific brain activity. *NeuroImage*, *56*(3), 1267–1275. https://doi.org/10.1016/j.neuroimage.2011.02.058

Younan, Y., Deffieux, T., Larrat, B., Fink, M., Tanter, M., & Aubry, J.-F. (2013). Influence of the pressure field distribution in transcranial ultrasonic neurostimulation. *Medical Physics*, *40*(8), 82902. https://doi.org/10.1118/1.4812423

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692. https://doi.org/10.1016/j.neuron.2005.04.026

List of Figures