2011

A ROLE FOR THE ANTERIOR CINGULATE CORTEX IN REINFORCEMENT-GUIDED LEARNING FOR COVERT ATTENTIONAL SELECTION

Michelle Lynne Bale

A ROLE FOR THE ANTERIOR CINGULATE CORTEX IN REINFORCEMENT-GUIDED LEARNING FOR COVERT ATTENTIONAL SELECTION

[Spine title: The ACC in Reinforcement Guided Learning]

[Thesis format: Monograph]

by

Michelle Lynne Bale

Graduate Program in Neuroscience

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

School of Graduate and Postdoctoral Studies,

The University of Western Ontario,

London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisors

Dr. Stefan Everling

Dr. Thilo Womelsdorf

Examiners

Dr. Bruce Morton

Dr. Britt Anderson

Dr. Susanne Schmid

Advisory Committee

Dr. Brian Corneil

Dr. Ravi Menon

The thesis by

Michelle Lynne Bale

entitled:

A ROLE FOR THE ANTERIOR CINGULATE CORTEX IN REINFORCEMENT-GUIDED LEARNING FOR COVERT ATTENTIONAL SELECTION

is accepted in partial fulfillment of the requirements for the degree of Master of Science

Date

Chair of the Thesis Examination Board

Abstract:

The anterior cingulate cortex (ACC) has been associated with a variety of functions including conflict monitoring, error detection and more recently reward based learning. In this study we recorded from the ACC, the ventromedial prefrontal cortex (vmPFC) and the dorsolateral prefrontal cortex (dlPFC) while macaque monkeys performed a variably rewarded spatial attention task. First, we found dynamic encoding of reward outcome and reward expectancy associated with attentional targets within the ACC, mPFC and dlPFC. These results expand the function of the ACC beyond merely action value associations and suggest this area serves a broader role in reinforcement guided learning and decision making. Secondly, analysis of outcome encoding relative to reward reversal revealed two distinct types of neurons: positive/negative prediction error neurons and positive/negative prediction certainty neurons. Prediction error neurons encoded outcome information only when reward associations had recently changed and thus new outcome information was most informative for establishing new reward expectations. Prediction certainty neurons on the other hand signaled the certainty of the reward prediction itself and encoded outcome information only later, when reward expectations had been built up. Prediction error neurons showed a correlation between reward selectivity during outcome periods and reward selectivity preceding subsequent reward predictive events. This finding could serve as a link between prediction error signals and behavioural adjustment. Finally, prediction error neurons predominated in the ventral ACC whereas prediction certainty neurons predominated in dlPFC area 9. Though not definitive this supports proposals that

outcome predictions are developed and adjusted within the ACC and mPFC and these predictions are used by the dlPFC to determine the behavioural response.

Key Words: anterior cingulate cortex, prediction error, attention, reward, prefrontal cortex, reinforcement learning, decision making, prediction certainty, macaque

Co-Authorship:

M. L. Bale [1], I. Janemi [1], D. Kaping [2], S. Everling [1,2,3], T. Womelsdorf [1,2]

1. Department of Neuroscience , University of Western Ontario, London, Ontario N6A 5C1, Canada.
2. Department of Physiology and Pharmacology, University of Western Ontario, London, Ontario N6A 5C1, Canada.
3. Robarts Research Institute, London, Ontario N6A 5K8, Canada

Iman Janemi helped with data collection and recording location reconstruction. Dr. Daniel Kaping also helped with data collection and provided helpful feedback during the writing process. Dr. Stefan Everling provided guidance and mentorship throughout the entirety of this research project. Dr. Thilo Womelsdorf, as my primary supervisor, was present throughout every stage of this project and provided invaluable help in data collection, data analysis and research direction. He also provided guidance throughout the writing process. A version of the study presented here is in preparation for submission.

For my family whose love strengthens me in everything I do.

And for Rory who has helped me so much during this time- I love you and appreciate you more than you know.

Thank you.

Acknowledgements:

This work was supported by the Canadian Institutes of Health Research and I was supported by an Ontario Graduate Scholarship in Science and Technology.

I thank all the members of the Everling Lab for their help and support. I extend a special thank you to Dr. Thilo Womelsdorf whose advice and expertise was invaluable to the completion of this project, and to Iman Janemi for his help and support throughout many hours of data collection and recording location reconstruction.

I thank Drs. Brian Corneil and Ravi Menon who served on my advisory committee for their advice and guidance. I also extend special thanks to my co-supervisor Dr. Stefan Everling for his time and mentorship throughout my time in his laboratory.

Finally, I thank my friends and family who have supported and encouraged me over the past two years. I am extremely thankful for all of you.

Table of Contents:

List of Figures:

List of Appendices:

List of Abbreviations:

ACC	- anterior cingulate cortex
ACCd	- dorsal bank of anterior cingulate cortex
ACCv	- ventral bank of anterior cingulate cortex
BOLD	- blood oxygenation level dependent
dACC	- dorsal anterior cingulate cortex
dlPFC	- dorsolateral prefrontal cortex
FEF	- frontal eye fields
fMRI	- functional magnetic resonance imaging
IPS	- inferior parietal sulcus
lPFC	- lateral prefrontal cortex
m1	- monkey 1
m2	- monkey 2
MCC	- midcingulate cortex
mPFC	- medial prefrontal cortex
MRI	- magnetic resonance imaging
mvPFC	- medial ventral prefrontal cortex
NAcc	- nucleus accumbens
nPC	- negative prediction certainty
nPE	- negative prediction error
OFC	- orbitofrontal cortex
PCC	- posterior cingulate cortex
PFC	- prefrontal cortex
pPC	- positive prediction certainty
pPE	- positive prediction error
RSC	- retrosplenial cortex
SNc	- substantia nigra pars compacta
SPL	- superior parietal lobule
VTA	- ventral tegmental area

List of Symbols:

cm	- centimeter(s)
Hz	- hertz (cycles per second)
kg	- kilogram(s)
kHz	- kilo-hertz
mm	- millimeter(s)
msec	- millisecond(s)
MΩ	- mega ohm(s)
%	- percent
T	- Tesla

Chapter 1 - Introduction and Literature Review:

1.1 - The Anterior Cingulate Cortex an Overview

The anterior cingulate cortex (ACC) is a brain area whose function has long been under debate. Historically there have been two predominant views. The first, based on strong error-related negativity signals generated in the ACC, suggests a specific role in error detection and compensation (Posner & DiGirolamo, 1998). The second posits that the ACC serves to detect the occurrence of conflict in information processing and to adjust levels of cognitive control so as to prevent conflict in the future (Carter et al., 1998). Recently, single unit recording, lesion and functional magnetic resonance imaging (fMRI) studies have expanded the role of the ACC into positive and negative outcome processing, outcome expectancy and behavioural adjustment. It is now generally believed to be a key player in reward-based learning.

The cingulate cortex is situated on the medial wall of the frontal lobes of each hemisphere following roughly the curvature of the corpus callosum. It includes areas 32, 25, 24, 23 and 31 with 32, 25 and 24 forming the ACC and 23 and 31 forming the posterior cingulate cortex (PCC) (Carmichael & Price, 1994; Barbas 1992). There is no strict delineation separating the anterior and posterior components of the cingulate cortex but rather a gradual increase in layer IV cells as one moves posteriorly from an agranular ACC to a dysgranular/granular PCC (Vogt et al., 1987; Morecraft et al., 2004). Because these areas occupy slightly different locations in humans and non-human primates and because strict delineations of the areas is impossible with human imaging,

broader names are commonly used. The ventral medial prefrontal cortex (vmPFC) generally includes anterior cingulate areas 25, 32 as well as orbitofrontal areas (Averbeck & Seo, 2008; Haber & Knutson, 2010). Reference to medial prefrontal cortex (mPFC) tends to exclude orbitofrontal areas. Area 24 is sometimes included in mPFC or independently referred to as dorsal anterior cingulate cortex (dACC) (Averbeck & Seo, 2008).

Diverse connectivity profiles along the ACC make this area an interface of multiple functional systems. Extensive, reciprocal cortico-cortical connections with the lateral and dorsolateral prefrontal cortex (lPFC and dlPFC respectively) support a role in cognition (Barbas & Pandya, 1989). Areas receiving input from primary motor, premotor and supplementary motor cortices and projecting to the spinal cord outline the cingulate motor field (Morecraft & Van Hoesen, 1992; Dum & Strick, 1991). The anterior cingulate is also heavily connected with limbic areas receiving input from the amygdala (Barbas & De Olmos, 1990), hypothalamus (Carmichael & Price, 1998) and ventral striatum (Kunishio & Haber, 1994). It is also a major target of the meso-cortical dopamine system which originates in the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) (Williams & Goldman-Rakic, 1993; Berger, 1992).

Though generally agreed upon, details of the structural and functional organization of the cingulate cortex are still under debate. Brodmann divided the cingulate cortex into two regions: the anterior cingulate cortex and the posterior cingulate cortex (ACC and PCC respectively) (Brodmann, 1909). This two-region division has been challenged, or at least expanded, by neurocytology, imaging, electrical

stimulation and lesion studies (Vogt et al. 1997, 2004) suggesting four rather than two cingulate divisions. These four divisions are the anterior cingulate, midcingulate, posterior cingulate, and retrosplenial cortices (ACC, MCC, PCC and RSC respectively) (Vogt et al. 2005). Of note is the distinction of Brodmann's ACC into perigenual and midcingulate regions with distinct connectivity and functions. The perigenual cingulate cortex (Vogt et al. 2005's true ACC) has reciprocal connections with the amygdala. It also has projections to the nucleus of the solitary tract and to the dorsal motor nucleus of the vagus allowing for regulation of autonomic output (Neafsey et al. 1993, Vogt et al. 1997, 2003). The midcingulate cortex has only modest amygdala input. It projects to the spinal cord allowing for regulation of skeletomotor function (Vogt et al. 1997, 2004). These anatomical distinctions imply a role for the midcingulate in action and motor-related responses and the anterior (perigenual) cingulate cortex in limbic, emotional responses. Within the limbic system, the ACC as a whole is a major component of the brain's larger reward circuitry.

Many areas of the brain show reward modulated activity, however the cortical-basal ganglia system forms the basis of the brain's reward system (Olds & Milner, 1954; Haber & Knutson, 2010). Key components of this system include the dopaminergic neurons of the nucleus accumbens (NAcc) and the ventral tegmental area (VTA) (Hikosaka et al. 2008; Schultz, 2000), the ventral striatum, ventral pallidum, orbitofrontal cortex (OFC) and the ACC (Haber & Knutson, 2010). The ACC has been shown not only to encode the presence or absence of reward (Matsumoto et al. 2003) but also multiple dimensions of the outcome itself including reward magnitude (Amiez et al.

2005), reward probability (Kennerley et al. 2009) and even the cost of obtaining the reward (Preuschoff et al. 2006; Kennerley et al. 2009). Monkey electrophysiology studies of reward encoding in the ACC as well as event-related functional magnetic resonance imaging (fMRI) has improved our understanding of what information the ACC encodes and particularly *when* this information is encoded. The ACC has been shown to encode outcome information at essentially all time points throughout a decision making period from as early as before choice alternatives appear (Niki & Watanabe, 1979), to pre and peri-response selection (Niki & Watanabe, 1979; Nishijo et al. 1997) and also after the outcome is incurred (Amiez et al., 2005; Matsumoto et al., 2003; Quilodran et al., 2008). This extensive outcome encoding expands the role of the ACC beyond simply error detection or conflict monitoring but points rather to a reiterant circular process (Amiez et al. 2006) of behavioural adaptation including outcome processing, outcome-association updating and outcome expectancy adjustment (*Figure 1*).

Extending from its role in reward circuitry, it has also been proposed that the ACC is involved in biasing attention toward emotionally relevant stimuli by influencing the frontoparietal attention network (Mohanty et al. 2009). The frontoparietal attention network (Corbetta & Shulman, 2002) includes the inferior parietal sulcus (IPS), the superior parietal lobule (SPL) and the frontal eye fields (FEF) and serves to guide top-down signals for spatial attention. This network is "short circuited" by particularly salient (high sensorial intensity) stimuli through the bottom-up capture of attention (Corbetta & Shulman, 2002) but should also be influenced by non-salient cues which

develop emotional relevance through experience. This incorporation of cognition and emotion in attentional allocation is purported to be achieved through limbic (amygdala, orbitofrontal cortex, anterior cingulate cortex) influences on the canonical spatial attention network via the cingulate gyrus (ACC and PCC). According to this view, limbic input influences the spatial attention network via the PCC (Mohanty et al. 2009) which is reciprocally connected with the ACC (Pandya et al., 1981; Baleydier & Mauguiere, 1980). The recent expansion of the function of the ACC into outcome processing, reward expectancy and attentional biasing has made it an area particularly implicated in learning and decision making.

1.2 - Reinforcement-Guided Learning

Classroom-style learning has become so prominent in our society that we sometimes overlook the most basic, innate type of learning: the learning that arises with no formal instructor but simply from interaction with our environment. From birth, this type of learning is the first to occur as we learn which sound productions make words or which food choices nourish us and which make us sick. We learn by making choices and seeing what outcomes result. Choices that lead to positive outcomes are reinforced while choices resulting in negative outcomes decrease (Thorndike, 1898). This idea forms the basis of reinforcement-guided learning and is a key component in the study of the larger process of decision making.

This notion first fueled research in areas of artificial intelligence and led to the development of multiple reinforcement learning algorithms (Sutton & Barto, 1998).

These attempt to model the behaviour of an organism motivated by needs and goals and to explain how trial and error (or trial and success) type activity shapes and optimizes behaviour in order to achieve them. One way for this learning to take place is for the organism to have a representation of its 'best guess' of what will happen if it makes a particular choice and then learn based on errors in its predictions. This idea forms the basis of adaptation rules in engineering (Kalman, 1960; Widrow & Stearns, 1985) and learning rules in psychology (Rescorla & Wagner, 1972; Dickinson, 1980). A popular algorithm, and one well-suited to be used in the study of brain and neural control in general is the *Q-learning* algorithm (Watkins, 1989; Watkins & Dayan, 1992). Put simply, this model proposes that an organism repeats three basic steps in the process of reinforcement learning: 1) It predicts expected outcomes of its potential choices, 2) It selects the choice with the greatest expected outcome, 3) It updates its predictions if there are discrepancies in the expected outcome (Doya, 2007) (*Figure 1*). By doing so this organism fine-tunes its predictions when faced with alternative options and thus also fine-tunes and optimizes the choices it makes. This model also permits for adjustments to the organism's choice-outcome predictions should the environment and thus outcome contingencies change.

Though these algorithms were developed purely as computational theories, the mergence of such theories with work in the fields of neuroscience and psychology revealed that reinforcement guided learning models fit surprisingly well with findings from neuronal recordings and brain imaging data (Doya, 2007). Great advances have been made in fitting human architecture to such models (Glimcher et al., 2005, Santesso

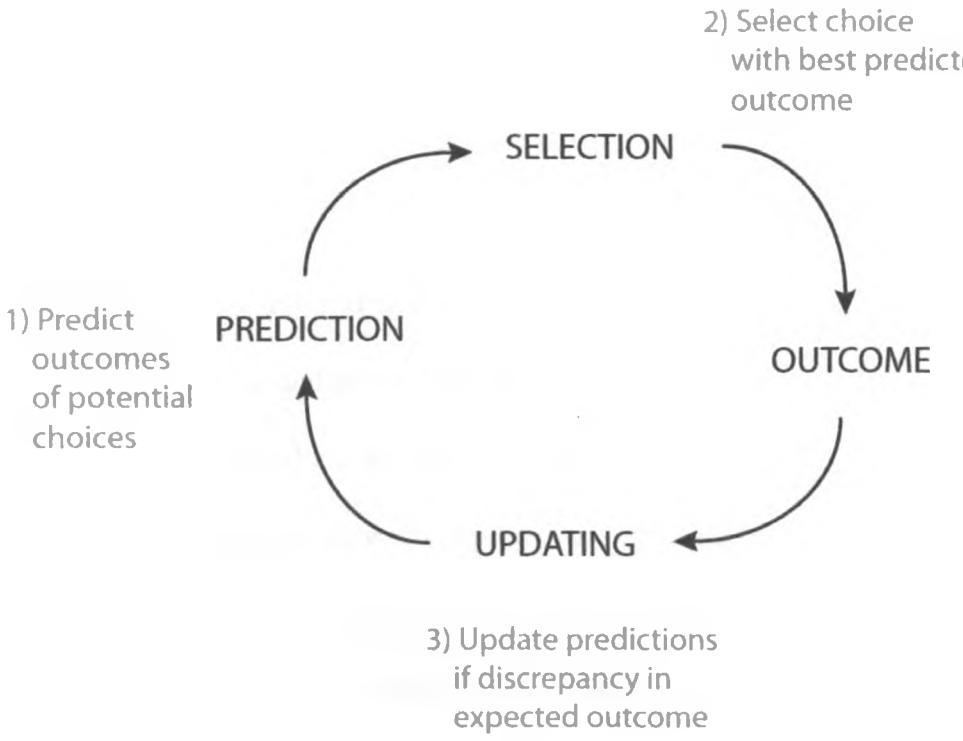


Figure 1. **Diagrammatic representation of the circular process of behavioural adaptation** as suggested by Amiez et al., 2006. Numbered descriptors identify the 3 basic steps of reinforcement guided learning as described in the *Q-learning* algorithm (Watkins, 1989; Watkins & Dayan, 1992).

et al., 2008; Francesco et al, 2007). However, much more work is required to elucidate the brain circuitry involved as well as the specific functional roles each area serves.

1.3 - Prediction Errors

One of the biggest and first parallels drawn between computational theories of reinforcement learning and actual neuronal recordings was through the firing pattern of dopamine neurons (Schultz et al., 1997). Computational models require the organism to be aware of discrepancies between expected and received outcomes and the dopamine neurons of the VTA and SNc seem to fit this function (Schultz et al., 1997; Schultz 2002). In a landmark study, Schultz et al. (1997) recorded the activity of dopamine neurons in these areas while thirsty monkeys received juice rewards on a regimen of varying stimulus-reward associations (*Figure 2A*). While the monkeys simply rested, the dopaminergic neurons fired at their basal rate of roughly 3 spikes per second (3 Hz). First, the monkeys received a drop of juice sporadically without any cue the reward was coming. This unexpected reward resulted in a sharp, transient increase in dopaminergic neuron firing rate (from 3 to roughly 80Hz for 100msec). Next, the monkeys began to receive a drop of juice preceded by an auditory tone. At first this reward resulted again in a sharp, transient increase in dopaminergic neuron firing rate but as the stimulus-reward combination was repeated the dopaminergic neuronal response dissipated until the tone and reward evoked no change in firing rate. Finally, the auditory tone was presented but no reward followed. This unexpected absence of reward led to a transient decrease in firing rate following the expected time of reward delivery, a phenomenon

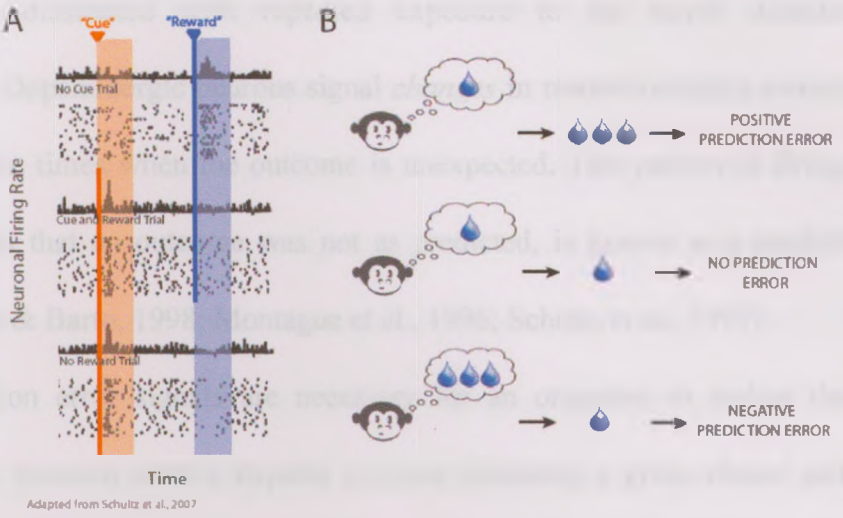


Figure 2. **Neuronal and diagrammatic representation of prediction errors.** (A) Prediction error signals from dopaminergic neurons. Top panel shows an un-cued and therefore unexpected reward delivery: positive prediction error. Middle panel shows a cued and therefore expected reward delivery: no prediction error. Lower panel shows a cued and therefore expected reward delivery but no reward follows: negative prediction error. Adapted from Schultz et al., 2007. (B) Diagrammatic representations of the corresponding prediction errors shown in (A).

which again dissipated with repeated exposure to the novel stimulus-reward combination. Dopaminergic neurons signal *changes* in reward-outcome associations or, in other words, times when the outcome is unexpected. This pattern of firing, when a neuron signals that an outcome was not as predicted, is known as a prediction error signal (Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997).

Prediction error signals are necessary for an organism to realize there is an inconsistency between what it expects to occur following a given choice and what in fact happens. They are necessary both for shaping behaviour, in order to maximize positive or desirable outcomes, and also to adjust behaviour in light of changing environmental conditions and thus changing choice-outcome contingencies. Prediction errors are defined by an organism's outcome expectation while it is making a choice relative to the actual outcome that results. Any discrepancy between the two is a prediction error. (*Figure 2B*) When an organism's choice results in a payoff that is *better* than expected, the outcome is more positive than predicted and this is referred to as a *positive* prediction error. When an organism's choice results in a payoff that is *less* than expected, the outcome is more negative than predicted and this is referred to as a *negative* prediction error (Sutton & Barto, 1998).

The prediction error-like pattern of activity seen in dopamine neurons, though pronounced and well-studied, has its weaknesses. Among them is the fact that both positive and negative prediction errors are conveyed by the same cells, via an increase or decrease in firing rate respectively (Schultz et al., 1997). Though dopaminergic neurons encode the size of the error itself for positive prediction errors (Bayer &

Glimcher, 2005), this encoding of the magnitude of the prediction error is not as reliable for negative prediction errors (Bayer & Glimcher, 2005). With the basal firing rate of dopaminergic neurons in the range of 3 Hz (Glimcher et al., 2005), negative prediction errors, signaled by a decrease in firing rate, do not have nearly the range across which the magnitude of the error can be conveyed as do positive prediction errors (Bayer & Glimcher, 2005). Dopaminergic signaling of increasingly negative prediction errors is also capped when firing rates hit 0 Hz. These limitations have led to the proposal that there is an opponent system to the dopaminergic system, perhaps serotonergic, which specializes in the signaling of negative prediction errors (Daw et al., 2002; Cools et al., 2008). The ACC receives strong dopaminergic (Williams & Goldman-Rakic, 1993; Berger, 1992) as well as serotonergic input (Varnas et al., 2004; Jacobs & Azmitia, 1992) and may provide a more behaviourally useful prediction error signal.

1.4 - The Anterior Cingulate in Reinforcement Learning

Dopaminergic neurons have been shown to encode reward outcome information, expectations of reward magnitude and probability (Tobler et al., 2005; Fiorillo et al., 2003) and errors in these expectations (Schultz et al., 1997; Haruno & Kawato, 2006). The ACC has been shown to encode similar aspects of reward. ACC cells discriminate rewarding from non-rewarding outcomes (Matsumoto et al., 2003) and also distinguish differences in the magnitudes of received rewards (Amiez et al., 2006). In addition, ACC cells have been found to encode predictions of outcomes in terms of both reward magnitude and reward probability (Kennerley et al., 2009). Recently, studies recording

single neuron activity from the dorsal bank and fundus of the anterior cingulate sulcus have found neurons which encode prediction-error like signals (Seo & Lee, 2007; Amiez et al., 2005; Matsumoto et al., 2007). These studies however, have some limitations. Amiez and colleagues (2005) recorded ACC single unit activity during a task which included 3 reward predicting stimuli: high, medium and zero. When the animal broke fixation during a trial and failed to receive the expected reward, they found a small proportion of cells (~ 4.5%) which showed error-related activity that varied according to the size of the expected reward: negative prediction error neurons type neurons. Their task, however did not allow for the occurrence of positive prediction errors. Seo and Lee (2007) fit a linear regression model combining value functions and positive and negative prediction errors to their recorded ACC unit activity. Their task however, a binary choice between two identical stimuli possessing computer assigned values changing on a trial by trial basis, is unlikely to evoke strong prediction errors. Using a continually reversing action-reward association task, Matsumoto and colleagues (2007) recorded quantitative reward prediction error signals from ACC neurons at the time of trial outcome (*Figure 3*). These results clearly show neurons coding the amount and direction (positive or negative) of errors made in estimating the value of executed actions.

Though similar to signals encoded by dopaminergic neurons, ACC cell activity has important differences. First of all, as mentioned above, single dopaminergic neurons encode both positive and negative prediction errors with an increase or decrease in firing rate respectively. ACC prediction error neurons recorded by Matsumoto and colleagues (2007) encode positive and negative prediction errors in different neurons. This allows

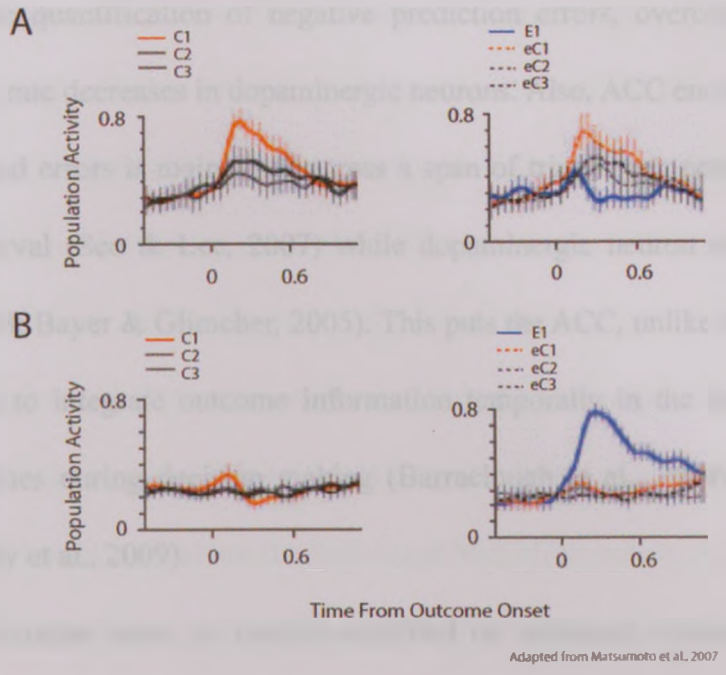


Figure 3. **Prediction error neurons recorded from the anterior cingulate cortex of the macaque.** Activity of a population of ACC cells (*y-axis*) relative to time from stimulus onset (*x-axis*) during the process of learning which actions are rewarded. (A) Positive feedback preferring cells. (B) Negative feedback preferring cells. In the left panels, the red line indicates the monkey guessed correctly and chose the rewarded action on the first try. The black lines indicate the second and third subsequent correct responses. In the right panels, the blue line indicates the monkey guessed incorrectly and chose the unrewarded action on the first try. The red line indicates a correct choice following the first error. The black line indicates the second and third subsequent correct responses following the first correct response. Adapted from Matsumoto et al., 2007.

more accurate quantification of negative prediction errors, overcoming the issue of limited firing rate decreases in dopaminergic neurons. Also, ACC encoding of outcomes, predictions and errors is maintained across a span of trials, even persisting through the intertrial interval (Seo & Lee, 2007) while dopaminergic neuron activity is transient (Schultz, 1998; Bayer & Glimcher, 2005). This puts the ACC, unlike dopamine neurons, in a position to integrate outcome information temporally in the brain and reference recent outcomes during decision making (Barraclough et al., 2004; Kennerley et al., 2006; Buckley et al., 2009).

ACC neurons seem to encode modified or enhanced versions of information encoded in dopaminergic neurons. There are, however, many other aspects of rewards and outcomes that influence optimal decision making. These should be considered in the decision making process of organisms as complex as human and non-human primates and include the cost of making a given choice (Kennerley et al., 2009), how certain one is of current value estimates (Behrens et al., 2007), the volatility of current environmental conditions (Courville et al., 2006), and even the potential value of new information gained from a choice (Behrens et al., 2007). Taking into account these other relevant criteria creates a much richer representation of the decision making environment and allows for greater optimization of choice behaviour both in terms of immediate outcomes and longer term beneficence --a consideration important for goal-driven organisms. The ACC has been shown to encode a variety of these "higher order" aspects of decision making.

Rodent work has proven useful in investigating the role of the ACC in weighing the cost required to obtain a reward versus the magnitude of the reward itself. One study had rats select between two food rewards of differing magnitude, each requiring different amounts of effort to reach them (Walton et al., 2003). Large rewards were blocked by a 30 cm barrier the rats had to scale to reach the food while the small rewards, though only half the size, were not blocked by a barrier. Control rats were able to compare the increased value of the large reward with the extra effort required to achieve it and consistently chose the high value/high effort option. After lesioning of the ACC however, there was nearly a complete reversal with the animals almost never willing to put in extra effort for a greater payout. The involvement of the ACC in effort-based decision making has been further studied with single unit neuronal recordings in the macaque monkey. A recent study required monkeys to select between two juice rewards of equal magnitude but that required different amounts of effort, in the form of lever presses, to obtain (Kennerley et al., 2009). The activity of neurons recorded from the ACC was significantly modulated by the amount of effort required to obtain the reward while the monkeys were making their decision.

ACC neurons have also been implicated in encoding how certain an organism is of its current reward estimates. In a task that contrasted a trial-and-error search period with a reward exploitation period (Procyk et al., 2000), a population of "search-preferring neurons" in the ACC was found to significantly increase activity during the search and return to baseline once the solution could be inferred. The search period can be presumably associated with uncertainty regarding choice-outcome associations while

the reward-exploitation period implies a degree of confidence in expected outcomes. These cells thus seem to be signaling uncertainty in one's reward estimates. In a more definitive study, the likelihood of specific choice-outcome associations was manipulated so as to produce periods of relative certainty in subjects' outcome expectations contrasted with periods of relative uncertainty (Behrens et al., 2007). ACC blood oxygenation level dependent (BOLD) activity during the post-reward period correlated with the degree of certainty or uncertainty inherent in the task. Importantly, the root of the uncertainty in the tasks mentioned above is the result of a changing or volatile environment. Environmental volatility is an aspect of learning tied to uncertainty and one increasingly relevant in statistical learning models (Dayan et al., 2000; Courville et al., 2003).

Reinforcement-guided learning follows a pattern of Bayesian probability where current probabilities or estimates are updated in light of new, relevant data or outcomes (Courville et al., 2006). This type of modeling treats the environment as unchanging or at best changing in a simple steady manor (Tenenbaum et al., 2001; Courville et al., 2006). Higher level cognitive functions, such as learning and decision making in goal-directed organisms fail to be adequately represented by such statistical models and, as such, are often described as stochastic and unpredictable in nature (Kahneman & Tversky, 2000). Taking into account the current state of the environment, whether steady and predictable or changing and volatile, more accurately predicts behaviour. Unlike in classic Bayesian theory, not every new outcome or prediction error has the same influence on the next decision or even necessarily leads to any behavioural

adjustment (Bayer & Glimcher, 2005). The influence of a prediction error on value-association updating depends on an additional factor "α" termed the learning rate (Rescorla & Wagner, 1972). This variable serves as a multiplier of the prediction error (δ) and modulates its effect on the current reward estimate (V_t) when generating the updated reward estimate (V_{t+1}).

$$V_{t+1} = V_t + \delta \alpha$$

The learning rate depends on the current level of certainty in value estimates and the reliability of the reward environment (Behrens et al., 2007). In a volatile environment, where reward contingencies are changing rapidly and the organism is uncertain, new outcome information should be weighted more heavily as it is more informative of novel reward associations than is past "out of date" outcome information. In a stable environment, new outcomes are weighted lightly as they provide little, if any, information beyond what is already known (Courville et al., 2006). This corresponds with findings that surprising outcomes enhance the speed of learning (Pearce & Hall, 1980). As mentioned above, the ACC has been shown to track uncertainty: an indicator of environmental volatility. Projections from the ACC to the ventral striatum (Kunishio & Haber, 1994) could serve as a means for the ACC to modulate prediction error signals directly based on "higher order" factors such as prediction certainty and environmental volatility. In support of this idea are findings from a macaque lesion study (Kennerley et al., 2006) in which subjects executed a lever-movement task where performance was entirely guided by reward feedback. Pre-operatively, behaviour on a given trial was estimated by multiple logistic regression to be dependent on five trials into the past.

After lesioning of the ACC, only the outcome on the most recent trial exerted any influence on subsequent choices. It was as if the prediction error signal of dopaminergic neurons was left as the sole modulator of behaviour. These findings point to a role for the ACC in encoding a more comprehensive, behaviourally relevant representation of the reward environment used to guide the reinforcement learning process.

Many ACC studies, particularly the lesion study mentioned immediately above, conclude the ACC encodes *action*-value associations rather than *stimulus*-value associations. This conclusion may be the result of the recording locations of a majority of ACC studies. Most recordings are made in locations interconnected with adjacent rostral cingulate motor areas and thus in a position to be influenced by action-selection processes (Rushworth & Behrens, 2008). A pair of studies seem to support this notion (Kennerley et al., 2006; 2009). The earlier study (Kennerley et al., 2006) made circumscribed lesions spanning a large area of the ACC including posterior regions impinging on cingulate motor areas and found deficits in action value encoding but not stimulus value encoding. The later study (Kennerley et al., 2009) recorded from a more confined and more anterior area much more distant from cingulate motor areas. This study showed clear stimulus value encoding in terms of reward magnitude, probability and cost. Though the earlier study shows intact stimulus-value representations despite near abolishment of the entire ACC, this maintained ability is likely the result of contributions from other prefrontal areas such as the orbitofrontal cortex discussed below. The presence of clear stimulus-value encoding in the later study however

confirms the contribution of the ACC to the representation of stimulus-reward associations.

1.5 - Frontal Areas in Reinforcement Learning

Many brain areas encode information related to reward value or expectation including the striatum (Kawagoe et al., 1998; Samejima et al., 2005), globus pallidus (Pasquereau et al., 2007), thalamus (Komura et al., 2005), ACC, OFC, dlPFC (Kennerley et al., 2009), parietal areas and medial temporal lobe areas (Liu & Richmond, 2000). Though the ACC is a key area for combining or "multiplexing" various aspects of reward and choice (Kennerley et al., 2009, Hayden & Platt, 2010), these widespread reward signals are not likely redundant. It is therefore necessary and relevant to explore differences or specializations in these areas. This section will focus on frontal areas relevant in reward based learning and decision making. Two areas, apart from the ACC detailed above, have been implicated in the literature based on lesion, brain damage, and single unit recording studies: the OFC (Fellows & Farah, 2007; Bechara et al., 1994) and the lPFC (Miller & Cohen, 2001).

The lPFC occupies the anterior part of the frontal lobes of the brain including Brodmann areas 46 and 9. It is historically associated with working memory tasks such as the Wisconsin Card Sorting Test and "match" / "nonmatch" rule tasks where subjects must flexibly shift between abstract rules, whether learned or cued. In reinforcement-guided learning, the lPFC has been implicated in maintaining a "state" representation (Lee et al., 2007) or in other words the context in which the learning and decision

making is taking place. Associated values can be modulated by a variety of environmental and internal factors. The value of a food decreases as one consumes it to satiety for example, or as mentioned above the state of the environment, whether stable or volatile, can influence values and even the effect of prediction errors. Activity encoding this state representation persists through delays and is often modulated by expected reward outcomes (Watanabe, 1996). The lPFC therefore encodes information about environmental context and the properties of expected rewards concurrently (Watanabe & Sakagami, 2007) making it a site of integration of cognitive and motivational information in the brain (Watanabe & Sakagami, 2007).

Along with the expansion of the prefrontal cortex, humans and non-human primates have developed the ability to exercise self control: to impose cognitive context information on motivational value information. The lPFC has been shown to modulate value signals encoded in the ACC when self control is required (Hare et al., 2009). Subjects were required to select between two foods: a tasty but unhealthy option or a healthy but less tasty alternative. In the context of dieting, selecting the less appetizing, healthier option is the best decision but making this choice requires modulation of taste value signals. When subjects made exactly this decision, Hare and colleagues found increased dlPFC activity and correspondingly suppressed ACC activity. The lPFC has also been implicated in delay discounting of rewards when environmental conditions are stable. Activity in the lPFC was shown to increase when subjects select larger, delayed monetary payouts over smaller but immediate ones (McClure et al., 2004). This is

suggestive of a role for the dlPFC in modulating value signals encoded in the ACC in order to overcome the desire for immediate rewards and exercise self control.

The OFC is located on the medial ventral surface of the frontal lobes, above the orbits of the eye. It is a key area in reinforcement-guided learning, strongly connected in function with the ACC. Deficits from lesions of the OFC were at first difficult to detect. One human case study for example, where the patient suffered bilateral OFC damage, failed to reveal any deficits in a battery of cognitive tests including tests of general intelligence, memory, visuospatial ability and working memory (Damasio 1994; Eslinger & Damasio 1985). This patient however, struggled to make everyday life decisions that required him choose between a variety of options. The OFC is now known to be essential for integrating various sources of information, from different sensory modalities, in order to calculate the overall value of choices (Padoa-Schioppa & Assad, 2006) and for adjusting these values in light of novel information (Mishkin 1964). One of the most prominent deficits seen with OFC lesions is difficulty in updating value associations when choice-value contingencies have changed, for example in stimulus-reward reversal tasks (Mishkin 1964) or in the Iowa gambling task (Bechara et al., 1994).

Like the ACC, the OFC encodes predictions of value based on reward magnitude, probability and cost (Kennerley et al., 2009) however there are some key differences. The connectivity of the OFC, with its strong sensory input from all modalities, weak motor connections and extensive limbic connections (Carmichael & Price 1995a,b) make it ideal for integrating sensory and reward information to determine the value of

stimuli. The ACC on the other hand has extensive motor connections as well as strong limbic connectivity (Chiba et al., 2001) making it ideal for combining motor and reward information to determine the value of actions. Though the OFC and ACC appear to specialize in stimulus-reward and action-reward associations respectively, single unit recordings and lesion studies show they are not mutually exclusive structures in this regard. As mentioned above, the ACC encodes clear stimulus value signals relating to reward size, likelihood and the cost of obtaining the reward (Kennerley et al., 2009). Additionally, extensive OFC lesions do not destroy ones ability to make stimulus-reward associations but rather slows their updating when contingencies change and makes it more difficult to perform value judgements when stimuli vary across multiple dimensions (Fellows & Farah, 2005).

Another difference is in what aspect of "cost" each area encodes. As previously discussed, the ACC encodes cost particularly in terms of the *effort* to obtain a reward. Alternatively, the OFC has been implicated in encoding cost particularly in terms of the *delay* in receiving reward. In a double dissociation between the two areas and the two types of costs, Rudebeck et al., 2007 trained rats on an effort-manipulated reward task. In this study, control animals would generally opt to either wait for a timed delay or climb a large barrier in order to receive a larger reward than they would had they chosen not to invest this time or effort. ACC lesioning biased animals toward the less effortful but less rewarding option yet did not affect choices concerning time delays. OFC lesioning on the other hand biased animals toward the more immediate but less rewarding option yet had no effect on choices involving effort manipulations. The OFC

is therefore important when considering cost in terms of delay while the ACC is necessary in considering costs in terms of effort.

A final difference between the ACC and OFC is the extent and complexity of reward encoding. Though they both encode a variety of aspects of reward, the ACC contains more prevalent, stronger and multiplexed outcome information (Kennerley et al., 2009). Over half of ACC neurons were found to encode reward information along at least one dimension for example, compared with less than 10% of OFC neurons. ACC neurons, unlike OFC neurons, also multiplex multiple dimensions of reward information. These findings are consistent with a role of the OFC in updating stored value associations and the ACC in integrating abstract value information and deriving overall behavioural values to guide choices (Wallis 2007).

In line with this hypothesis are the results from a recent study investigating the effect of circumscribed frontal lesions on a rule reversal task (Buckley et al., 2009). OFC lesioned animals made the majority of their errors in proximity to other errors, when rule-value associations were poorly defined. Following a string of correct responses however, and thus the strengthening of these associations, OFC lesioned animals performed at a similar level to controls. ACC lesioned animals performed consistently worse than controls with no tendency for errors to cluster suggesting a consistent impairment in making use of rule-value associations. More interesting than the performance accuracy data was the reaction time data. OFC lesioned animals responded slower than control animals, consistent with the presence of weak value associations making decisions harder to reach. ACC lesioned animals on the other hand

responded faster than controls suggestive of a failure to properly integrate reward information when reaching a decision, leading to hastier yet less optimal choices.

Overall, the OFC may excel in stimulus valuation while the ACC favours action valuation but they are not exclusive in this regard. The ACC and OFC seem to work together to develop value expectations adjusted for delay, effort and uncertainty which can be compared to actual outcomes and adjusted to optimize behaviour. The ACC, with its stronger, richer reward related activity is a prime candidate for integrating all of this information in order to generate functionally useful signals to guide changes in behaviour. The lPFC seems to govern this system by putting reward information into context and by modulating the strength of ACC activity when more highly rewarded options must be sacrificed in favour of long-term, higher order goals.

1.6 - Research Question

Further investigation into the nature and function of prediction error-like signals found in the ACC is critical for elucidating this structure's role in learning and decision making. The prediction error signals recorded from Matsumoto et al. (2007) are particular for errors in expectation of *action* value. Evidence from single unit recording and lesion studies points to a broader function of the ACC beyond merely action related valuation. Exploring these signals in terms of other types of value expectation will help delineate the role of the ACC in reinforcement-guided learning.

ACC prediction error signals carry richer representations of the reward environment than do dopaminergic prediction error signals and, very likely, serve a

behavioural function. However, studies have failed to find a link between these error signals and behavioural adjustment (Amiez et al., 2005). Exploring how these signals relate to and affect neuronal activity while choices are being made will help explain how the ACC contributes to the decision making process.

1.7 - Hypothesis

The purpose of this study is to further our understanding of how we use past experience to guide current decisions. Without explicit instructions, we learn through both the positive and negative outcomes of our choices how to adjust and improve our behaviour. Such reinforcement learning is a critical piece in the broader study of the neurobiology of decision making which has become a prominent area of research in the fields of neurophysiology, economics, evolutionary biology and computer science over the past decade (Kable & Glimcher, 2009), yet the neurological mechanisms behind it are still weakly understood.

The hypothesis was that the ACC encodes discrepancies between experienced outcomes and predicted outcomes and uses this information to influence decisions by influencing the allocation of attention. From this follow two predictions: (1) recordings from ACC neurons should show cells that follow a "prediction error-like" pattern of activity during the reward period. (2) the prediction error related signal of these cells should correlate with activity during subsequent cue periods, when the animal is directing his attention.

Chapter 2 - Methods:

2.1 - Animal Preparation

All experimental recordings were done according to the guidelines set forth by the Canadian Council of Animal Care on the use of laboratory animals and the University of Western Ontario's Council on Animal Care (see Appendix A). Two male rhesus monkeys (*Macaca mulatta*) served as research subjects. They were between 6 and 10 years old and weighed between 5-12 kg at the time of data collection. The monkeys will be referred to as subjects m1 and m2 respectively. Each animal was surgically implanted with a head post prior to the start of training to stabilize head and eye position during recording sessions. The animals were trained to sit in a primate chair with adjustable base height and neck plates to allow for comfortable fixation of the head throughout the entirety of the recording session.

To allow for extracellular recordings, each animal was fitted with two recording chambers each measuring 19mm in diameter. These chambers were implanted with the use of a stereotaxic frame and provided access to the ACC, mPFC and dlPFC. Each chamber was fitted with a recording grid containing tracts with 1mm inter-hole spacing which served to guide electrode insertion. Before recording, each animal's brain was imaged with a 7T magnetic resonance imaging (MRI) scanner in 1mm slices to ensure correct positioning of each of the recording chambers. During the scan, the animal's ear canals were marked with vitamin E capsules for later horizontal alignment and grid holes were filled with iodine to visualize electrode trajectories for later reconstruction of

recording sites. Both vitamin E and iodine are clearly visible in magnetic resonance images.

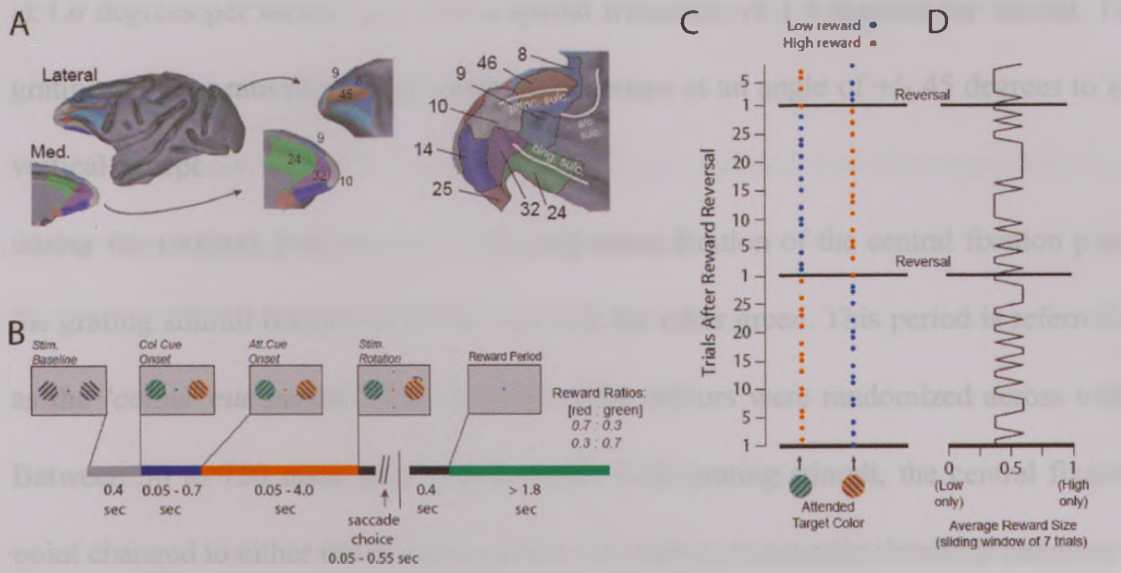
Subject m2's anterior chamber was surgically repositioned once during the recording period in order to access more anterior sites within the prefrontal and cingulate cortices and to align the recording locations with the same anterior-posterior axis covered with recordings obtained from subject m1. A second MRI scan was then performed to precisely determine the new location of the chamber.

2.2 - Experimental Paradigm

The animals were trained on a selective attention task with varying target-reward associations (*Figure 4B*). In each trial, subjects were required to covertly direct their attention to one of two lateral stimuli based on information from a central colour cue. They then had to discriminate the direction of rotation of their chosen attentional target and respond with the appropriate saccadic eye movement in order to receive reward.

Trials were initiated when the subject directed their gaze toward a central grey fixation point. Following 300 msec fixation of the central point, two black and white moving circular grating stimuli with radius of 1.5 to 2.2 degrees appeared at 4.2 degrees eccentricity to the left and right. This period is referred to as the '*stimulus baseline period*'. The gratings were an intermediate between a square wave and gabor type grating and moved within a circular aperture

Figure 4. **Fronto-Cingulate Anatomy and Task Design.** (A) Lateral and medial view of the macaque brain with anatomical subdivisions (colored) according to Barbas and Zikopoulus (2007). Middle panels highlight fronto-cingulate subdivisions in a partially inflated brain. The right panel shows the flattened representation of the fronto-cingulate cortex, covering anterior cingulate cortex (areas 24 and 32) and lateral prefrontal cortex (areas 10, 9, 46, and 8). (B) Basic task design: Monkeys initiated a trial by directing and keeping their gaze on a centrally presented fixation point. Following 0.3 sec two moving grating stimuli appeared (*Stim. Baseline*), which were colored red/green after 0.4 sec (*Color Cue onset*). Within 0.05 to 0.7 sec after color onset the central fix. point changed to red or green cueing the monkeys to covertly shift attention towards the location with the color matching stimulus (*Att. Cue onset*). At random times within 0.05-4 sec the cued target grating smoothly rotated clockwise or counterclockwise. In half of the trials the uncued distractor changed before the target. Monkeys discriminated the rotation of the target stimuli by saccading up- or downwards to one of two response targets. Reward (drops of water) was delivered 0.4 sec. after the saccadic choice. Stim. color was associated with high/low liquid reward with reward ratios (0.7 : 0.3) reversed every 30 correct trials. (C) Illustration of the trial progression (*y-axis*) with the attended target color (*x-axis*) changing randomly from trial-to-trial. Reward-color association reversed every 30 correct trials with an equal number of high (*red dots*) and low rewarded (*blue dots*) trials within each block. (D) Average reward size calculated for a sliding average encompassing 7 trials illustrate that the task allowed only minimal variation of reward size across trials.


A
Lateral
Med.
9
46
8
10
14
25
32
24
principal sulc.
arc. sulc.
cing. sulc.
B
Stim. Baseline
Col Cue Onset
Att.Cue Onset
Stim. Rotation
Reward Period
Reward Ratios: [red : green]
0.7 : 0.3
0.3 : 0.7
0.4 sec
0.05 - 0.7 sec
0.05 - 4.0 sec
saccade choice
0.05 - 0.55 sec
0.4 sec
> 1.8 sec
C
Low reward
High reward
Trials After Reward Reversal
Reversal
Reversal
Attended Target Color
D
0 (Low only)
0.5
1 (High only)
Average Reward Size (sliding window of 7 trials)

at 1.0 degrees per second and with a spatial frequency of 1.4 degrees per second. The gratings always moved upward within the aperture at an angle of +/- 45 degrees to the vertical except

during the rotation. Following a further 400 msec fixation of the central fixation point, the grating stimuli became coloured: one red, the other green. This period is referred to as the *'colour cue period'*. The location of the colours were randomized across trials. Between 50 to 750 msec after colour onset of the grating stimuli, the central fixation point changed to either red or green cueing the monkey to covertly direct his attention to the corresponding coloured stimulus. This period is referred to as the *'attention cue period'*. Within 50 msec to 4 seconds after cue onset, the cued grating transiently rotated either clockwise or counterclockwise. Rotation was smooth and proceeded from the standard direction of motion to maximum tilt within 60 msec, remained at maximum tilt for 235 msec, and rotated back to the standard direction of motion again within 60 msec. The times of target rotation were drawn from a flat random distribution to prevent subjects from anticipating the time of the event and in half of the trials the un-cued, distracting stimulus transiently rotated *before* the target stimulus. The direction of rotation indicated to the monkey which response, either an upward or downward saccade, would result in reward. Eye movement responses had to be made within 70 to 550 msec following rotation onset and subjects were required to maintain fixation of the selected response target for a minimum of 50 msec. The saccade direction associated with each stimulus rotation direction was reversed for the two subjects. Following a

delay of 400 msec after a correct response, the monkey received a fluid reward through a sipper tube placed in his mouth.

Importantly, the cued stimulus' colour was associated with the magnitude of the reward the subject would receive upon successful completion of the trial. Reward magnitudes for the red and green targets were set to a ratio of 0.7 : 0.3 and reversed after a block of 30 correctly performed trials. Colour-reward associations reversed multiple times within a recording day, the precise number depending on the amount of time the animal was willing to work. The animal received no indication of the reward block change apart from the receipt of an outcome that was better or worse than expected.

All task stimuli were displayed on a 19 inch CRT monitor running at 1024x768 pixel resolution and 85 Hz refresh rate. The monkeys were both seated 57 cm from the screen. Control of eye movements within the task and the generation of task stimuli were done with the open-source software Monkeylogic (*http://www.monkeylogic.net*) running on a Pentium III PC. Reward delivery was also controlled by Monkeylogic software and was accomplished by the opening of the mechanical valve of a custom-made air-compression controlled reward system. The system was placed a sufficient distance from the subject and recording equipment so as to minimize any auditory cues associated with the reward as well as any signal interference in the recordings.

Trial conditions were drawn from a custom condition-selection function which equalized the count of correct trial types with a variable trial lag of 2 to 5 trials to avoid immediate repetition of trial features such as target location (left/right), reward magnitude (high/low) or rotation direction (clockwise/counterclockwise). If at any point

during a trial the monkey broke fixation outside of the response window, responded to the wrong target or failed to respond to the correct target, the trial was aborted and counted as incorrect for the purposes of condition selection. For data analysis, only trials in which the monkey broke fixation after a stimulus change (either of the target *or* the distractor) were counted as true errors.

2.3 - Animal Training

The animals were first trained to detect the rotation direction of a single grating stimulus while permitted to fixate it directly. They would indicate their choice with either a right or leftward saccade. Again, rotation direction and saccade response direction were reversed for the two subjects. To determine the degree of transient stimulus rotation for each subject, complete psychometric curves were created prior to recording. The rotation was adjusted to ensure 85% or more overall correct responses to rotation discrimination and resulted in grating rotations between +/- 13 and +/- 19 degrees. The locations of the saccadic response targets were then shifted gradually until they occupied positions above and below the grating stimulus thus prompting upward or downward saccadic responses to the stimulus rotation direction. Gradually, the grating target was moved out of the animal's fixation window and replaced with a fixation point. The animals then had to detect the stimulus' rotation direction by covertly directing their attention to the spatial location of the target in their peripheral vision while maintaining central fixation. Finally, a "distractor" grating stimulus was added to the mirror image location of the target stimulus and red and green colouring was added

to indicate which grating stimulus should be the target of the monkey's attentional allocation.

Throughout training, sets of 5 "neutral" fixation trials were included at the end of blocks in which the animals were simply required to maintain fixation on a central yellow fixation point for 5 seconds. Successful completion of a fixation trial resulted in a small reward (60 to 70% of the low reward for attention trials) but proved invaluable for maintaining subject motivation on a difficult task throughout the training period. These fixation trials were maintained after training as the animals proved quite sensitive to the varying reward magnitudes throughout the task and successful completion of the less demanding fixation trials boosted motivation.

2.4 - Data Collection

Throughout the recording sessions, subjects were seated in the primate chair within a wooden box enclosure which blocked ambient light and dampened ambient noise. Their eye movements were tracked with an infrared eye tracking system (*ISCAN, Woburn, US*) running on a DOS platform. The system permitted only deviations from the task fixation point that did not exceed a 1.4 to 2.0 degree radius. Subject fluid intake was controlled to maintain motivation on the task and recording sessions lasted between 1 to 3 hours dependent entirely on the animal's continued performance of the task.

Extracellular recordings of single neuron action potentials were obtained using between 1 to 6 single tungsten electrodes per recording session. The electrodes had an impedance of 1.2 - 2.2 MΩ (*FHC, Bowdoinham, ME*). Electrodes were inserted using

software-controlled precision microdrives (*NAN Instruments Ltd., Israel*) and guided into the cortex with stainless steel guide tubes nested within the recording grid tracts. The coordinate location of the tract within the grid and the depth of electrode insertion determined the recording location and each was selected beforehand using the 7T MR images and major cortical landmarks such as the anterior cingulate sulcus and the principle sulcus. Both the grid coordinates and depth as well as the activity profile observed while lowering the electrodes were all used for later reconstruction of recording locations.

Data acquisition, filtering and amplification were done with a multi-channel processor (*Map System, Plexon Inc.*) using head stages with unit gain. Spiking activity was isolated using a 100-8000 Hz bandpass filter, further amplified and digitized at a 40 kHz sampling rate. Action potentials were recorded in a 0.85 to 1.1 msec time window and single unit, mostly single unit and multi unit activity was isolated. Waveforms were preliminarily isolated online then final isolation and sorting were done offline with Plexon Offline Sorter (*Plexon Inc., Dallas, TX*) based on principal component analysis of the waveforms. To visualize the spike rate data, spike density functions were generated by convolution with a Gaussian kernel function.

2.5 - Reconstruction of Recording Sites

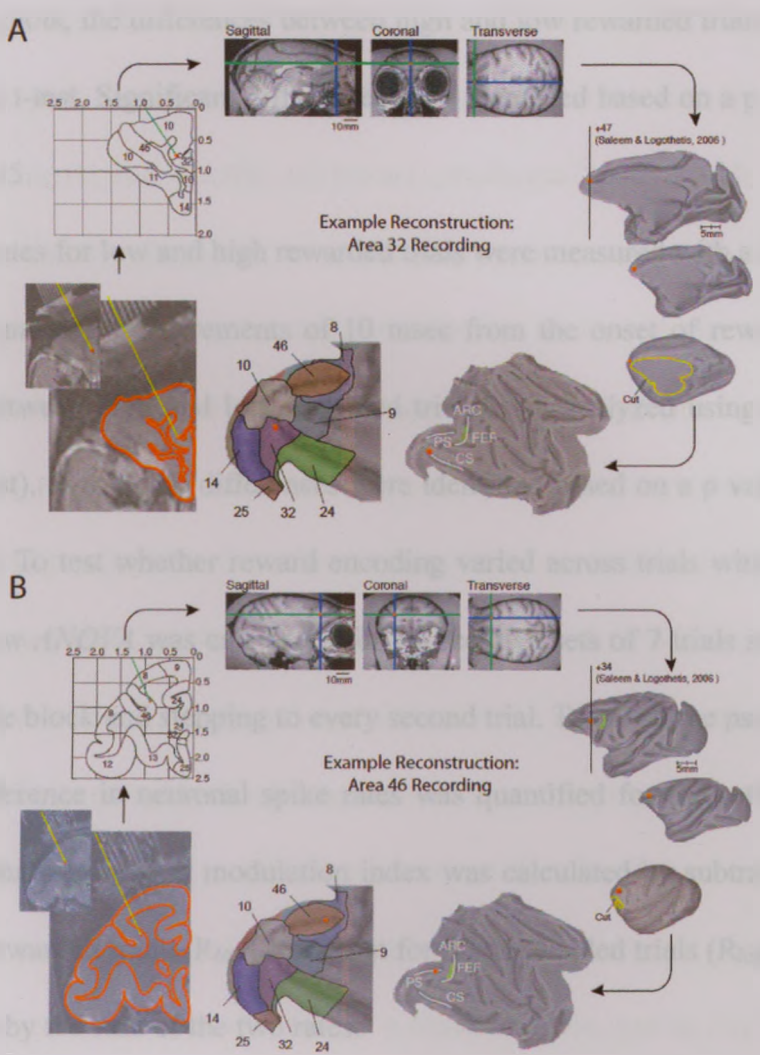
The images obtained from the MRI scan were viewed using open-source OsiriX Imaging software and traced with commercial graphical software (*Adobe Illustrator, San Jose, CA*) to produce 2-dimensional bitmap brain slice images. These images were

outlines of cortical folding within the coronal plane and assisted in comparing individual brain images with standard anatomical atlases. Using the coordinate location of each electrode within the recording grid and the noted depth of each recording site, the trajectory of each electrode was projected onto the bitmap brain images using custom-written MATLAB programs *(Mathworks, Natick, MA*). Individual monkey brains were reproduced using the open-source software CARET (Computerized Anatomical Reconstruction and Editing Toolkit) to validate use of the standard F99 brain and to derive scaling factors to match the individual images to the standard brain. Recording positions were then projected manually from the MATLAB-reconstruced images of the individual brains onto the standard F99 macaque brain. After identifying all recording sites on the standard brain, CARET software was used to generate a 3-dimensional reconstruction of the recordings. The 3-dimensional brain was then inflated within CARET and cut to yield an unfolded flat-map upon which to better visualize the data (*Figure 5*). The estimated error from the entire reconstruction procedure can reach a maximum of 3mm, yet the typical, unsystematic error is much more likely within the range of 1mm. Importantly, the entire anatomical reconstruction was done *independently* of any analysis of neuronal data.

2.6 - Data Analysis

All statistical analyses were done using MATLAB (*Mathworks, Natick, MA*). For behavioural analysis, only trials in which the monkey broke fixation after a stimulus change (either target or distractor) were counted as true errors. For accuracy, error and

Figure 5. **Example reconstructions of two recording sites in area 32 (A) and in area 46 (B).** Reconstruction began from 7T anatomical MRs of each individual monkeys brain. The MR scan was obtained with (iodine based) visualization of electrode trajectories within the electrode grid placed inside the recording chamber. The outline of the cortical folding was sketched on the two-dimensional MR to ease identification of areas and landmarks according to standard brain atlases, and to align reconstruction of the electrode tip with custom MATLAB software. The electrode position was then placed into the standardized F99 macaque brain available in the CARET software package. CARET allowed for rendering of the MR slice into a three dimensional volume and to inflate the volume in order to finally cut (indicated as yellow line) the spherically inflated brain for representing it in a two - dimensional flat map. White lines on the flat map indicate the principal sulcus (PS), the arcuate sulcus (ARC), and the cingulate sulcus (CS). The location of the frontal eye field within the arcuate sulcus is indicated by the yellow shading. As a last step, the anatomical subdivision of areas in the fronto-cingulate cortex were visualized (here: area subdivision derived from Barbas and Zikopoulus, 2007). Note that the recorded site is visualized throughout the panels as red dot.


A
Sagittal
Coronal
Transverse
10mm
+47
(Saleem & Logothetis, 2006)
5mm
Example Reconstruction:
Area 32 Recording
8
46
10
9
14
25
32
24
ARC
FEF
PS
CS
Cut
B
Sagittal
Coronal
Transverse
10mm
+34
(Saleem & Logothetis, 2006)
5mm
Example Reconstruction:
Area 46 Recording
8
46
10
9
14
25
32
24
ARC
FEF
PS
CS
Cut

reaction time plots, the differences between high and low rewarded trials were analyzed using a paired t-test. Significant differences were identified based on a p value less than or equal to 0.05.

Spike rates for low and high rewarded trials were measured with a sliding window of 300 msec moving in increments of 10 msec from the onset of reward. Firing rate differences between high and low rewarded trials were analyzed using a one-factorial *ANOVA* (F-test). Significant differences were identified based on a p value less than or equal to 0.05. To test whether reward encoding varied across trials within a block, the sliding window *ANOVA* was calculated for consecutive sets of 7 trials starting with the first trial in the block and stepping to every second trial. To create the pseudocolour heat plots, the difference in neuronal spike rates was quantified for the sets of 7 trials, as above. A normalized reward modulation index was calculated by subtracting the firing rate for low rewarded trials (R_{low}) from that for high rewarded trials (R_{high}) and dividing the difference by the sum of the two rates:

$$(R_{high} - R_{low}) / (R_{high} + R_{low})$$

This resulted in values above zero if there was stronger activity for high rewarded outcomes and values below zero if there was stronger activity for low rewarded outcomes.

To characterize neuronal patterns of reward modulation across trials, the time after reward onset with the maximal spike rate difference between high and low outcome trials was identified for each neuron. The reward modulation index at this time point was plotted against trial number for the entire block. A linear regression slope was

then fitted to the distribution of reward modulation indices across trials. A significantly positive or negative slope was identified as having a p value less than or equal to 0.05.

To investigate the transfer of reward prediction error signals in the reward outcome period to the attentional cue period two analysis steps were performed. First, the reward modulation index was calculated for the attentional cue period across all trials in the block in 300 msec analysis windows as described above. Next, a correlation analysis of the reward outcome modulation to the attentional modulation in the cue period was performed. To do this, the average reward modulation between 300 to 600 msec after reward onset in trials 1 through 7 was calculated for each neuron. This value was then correlated with the reward modulation in the attentional cue period observed across all trials in the block and across all time windows following cue onset. Significant correlations were identified as having p values less than or equal to 0.05.

To visualize the anatomical locations of different functional classes of neurons, recording locations were reconstructed on a flattened brain map as described above. A regular grid with 2mm inter-node spacing was overlain on the map for visualization. The number of neurons at each intersection of the grid falling within a radius of 4mm around the intersection were counted. The proportion of neurons fitting the functional classification was then calculated relative to the total number of neurons recorded at that location.

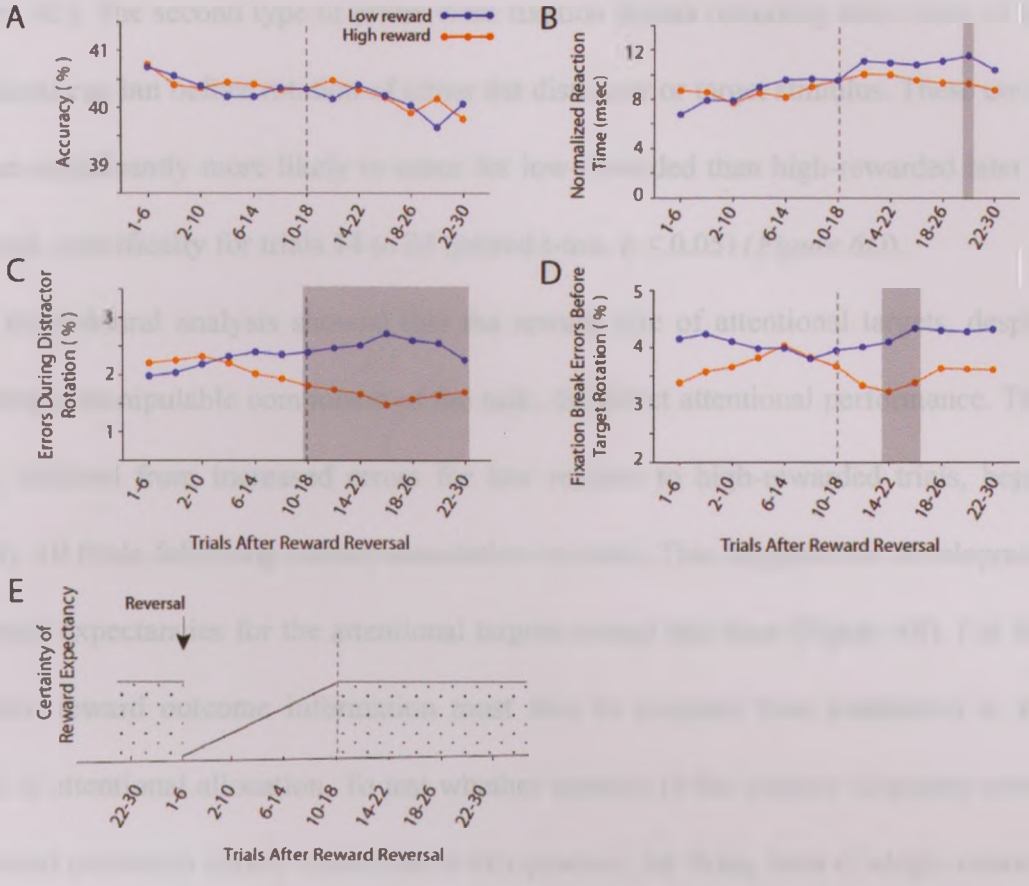
Chapter 3 - Results:

3.1 - Behavioural Sensitivity to Attentional Target Value

Behavioural performance was recorded on the selective spatial attention task outlined in (*Figure 4B*) across 133 experimental recording sessions in two monkeys (n=63 for m1, n=70 for m2). The monkeys were required to perform equal numbers of correct trials for high and low rewarded targets with the use of a custom condition-selection function (*see Methods*). This ensured the average reward rate was kept nearly constant. If low and high reward magnitudes were considered 0 and 1 respectively, the average reward rate ranged from 0.375 to 0.625 (*Figure 4C,D*). Reward itself could not be manipulated by the monkeys yet they were consistently directing their attention to both high and low rewarded targets.

Average task performance was roughly 80% correct. To detect any difference in performance between high-rewarded and low-rewarded trials, accuracy and reaction times were plotted separately relative to trial from reward-association reversal. There was no significant difference in performance accuracy between high and low rewarded trials (*Figure 6A*). Reaction times became slower for low-rewarded trials later on in the block. Trials 20-28 showed significantly slower reaction times for low relative to high-rewarded outcomes (paired t-test, $p \leq 0.05$) (*Figure 6B*). Errors fit into two general categories. The first occurred within 500 to 600msec following a rotation of the distracting stimulus. These errors became significantly more likely to occur for low-rewarded than high-rewarded trials during the last half of a block (paired t-test, $p \leq 0.05$)

Figure 6. **Behavioural performance across trials relative to reward-association reversal.** Areas of significant differences between high (red) and low (blue)-rewarded trials (paired t-test, $p < 0.05$) are encased in grey shading. (a) Overall performance accuracy in percentage correct (*y-axis*) plotted relative to trial from reward-association reversal. (b) Normalized reaction times plotted relative to trial from reward-association reversal. (c) Proportion of errors made shortly after rotation of the distracting stimulus plotted relative to trial from reward-association reversal. (d) Proportion of breaks in fixation made between attentional cue onset and the first stimulus change plotted relative to trial from reward-association reversal. (e) Diagrammatic representation of the animal's certainty in its estimate of target value (*y-axis*) across trials as inferred from the development of increased errors for low-value trials later in the block.


A
Low reward
High reward
Accuracy (%)
41
40
39
1-6
2-10
6-14
10-18
14-22
18-26
22-30
B
Normalized Reaction Time (msec)
12
8
4
0
C
Errors During Distractor Rotation (%)
3
2
1
Trials After Reward Reversal
D
Fixation Break Errors Before Target Rotation (%)
5
4
3
2
Trials After Reward Reversal
E
Reversal
Certainty of Reward Expectancy
Trials After Reward Reversal

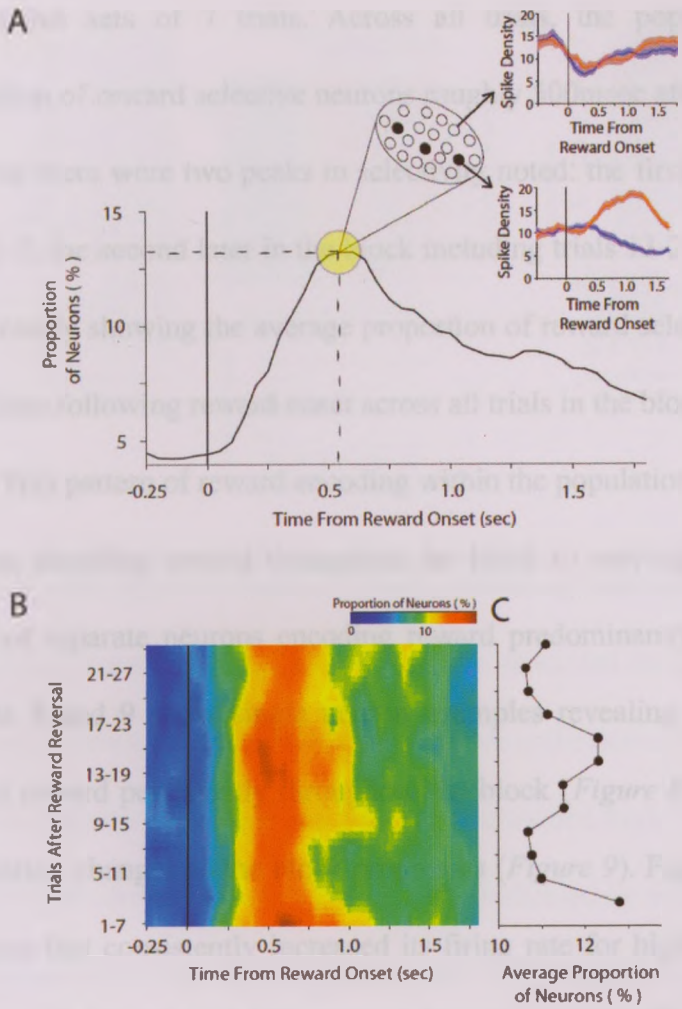
(*Figure 6C*). The second type of errors were fixation breaks occurring after onset of the attentional cue but before rotation of either the distractor or target stimulus. These errors became significantly more likely to occur for low-rewarded than high-rewarded later in the block, specifically for trials 14 to 24 (paired t-test, $p \leq 0.05$) (*Figure 6D*).

Behavioural analysis showed that the reward size of attentional targets, despite not being a manipulable component of the task, did affect attentional performance. This effect, inferred from increased errors for low relative to high-rewarded trials, began roughly 10 trials following reward-association reversal. This suggests the development of reward expectancies for the attentional targets around this time (Figure *6E*). For this to occur, reward outcome information must first be encoded then transferred to the period of attentional allocation. To test whether neurons in the anterior cingulate cortex and lateral prefrontal cortex contribute to this process, the firing rates of single neurons from these areas were analyzed during the outcome evaluation period, shortly following reward delivery, to see if they showed reward modulation.

3.2 - Neuronal Encoding of Reward Outcome for Attentional Targets

A total of 742 neurons were recorded from the fronto-cingulate cortex, including the ACC, vmPFC and dlPFC. Figure 7*A* shows the proportion of neurons with significant (*ANOVA* F-test, $p \leq 0.05$) reward modulation. The proportion of neurons rises sharply to a peak at roughly 500msec after reward onset then declines slowly throughout the remainder of the post-reward period. To see whether there were any differences in reward modulation throughout the block, a sliding window *ANOVA* was calculated for

Figure 7. **Reward-selective outcome processing.** (a) Proportion of neurons with a significant main effect (*ANOVA*, *F*-test, $p \leq 0.05$) of reward size (*y-axis*) plotted relative to reward onset (*x-axis*). Expanded from the peak of the plot, highlighted in yellow, is a diagrammatic representation of the proportion of neurons showing significant reward modulation (black circles) within the pool of non-modulated neurons (unfilled circles). Inlayed spike density plots show representative single neuron examples for high (red) and low (blue) rewarded trials. (b) Data from (a) with the proportion of reward modulated neurons (*colour-axis*) expanded across trials within a reward-association block (*y-axis*). (c) Average proportion of reward modulated neurons within 0.25 to 1.25 sec after reward onset (*x-axis*) plotted relative to trial from reward-association reversal.


A
Spike Density
Time From
Reward Onset
Proportion
of Neurons (%)
Time From Reward Onset (sec)
B
Proportion of Neurons (%)
Trials After Reward Reversal
21-27
17-23
13-19
9-15
5-11
1-7
Time From Reward Onset (sec)
C
Average Proportion
of Neurons (%)

consecutive sets of 7 trials. Across all trials, the population showed a maximal proportion of reward selective neurons roughly 500msec after reward onset (*Figure 7B*), however there were two peaks in selectivity noted: the first early in the block including trials 1-7, the second later in the block including trials 13-21. Figure 7*C* illustrates these peaks nicely showing the average proportion of reward selective neurons from 250msec to 1.25sec following reward onset across all trials in the block.

This pattern of reward encoding within the population could be the result of single neurons encoding reward throughout the block to varying degrees, or it could be the result of separate neurons encoding reward predominantly early or late in the block. Figures 8 and 9 show single neuron examples revealing there are both neurons that encode reward persistently throughout the block (*Figure 8*) and neurons whose reward modulation changes as the block progresses (*Figure 9*). Figure 8*A* shows an example of a neuron that consistently increased its firing rate for high rewarded outcomes, Figure 8*C* shows a neuron with the opposite pattern: consistently increasing its firing rate for low rewarded outcomes. Figure 9 shows two example neurons whose reward selectivity varies relative to trial after reward reversal. These neurons responded stronger to high-rewarded versus low-rewarded (*Figure 9A*) or stronger to low-rewarded versus high-rewarded (*Figure 9C*) only during early trials in the block. The latencies of these signals varied across neurons and ranged from early after reward onset (300msec) to late after reward onset (700 to 1000msec). The examples in Figure *9* show there is a subset of neurons in the fronto-cingulate cortex that convey reward information only when

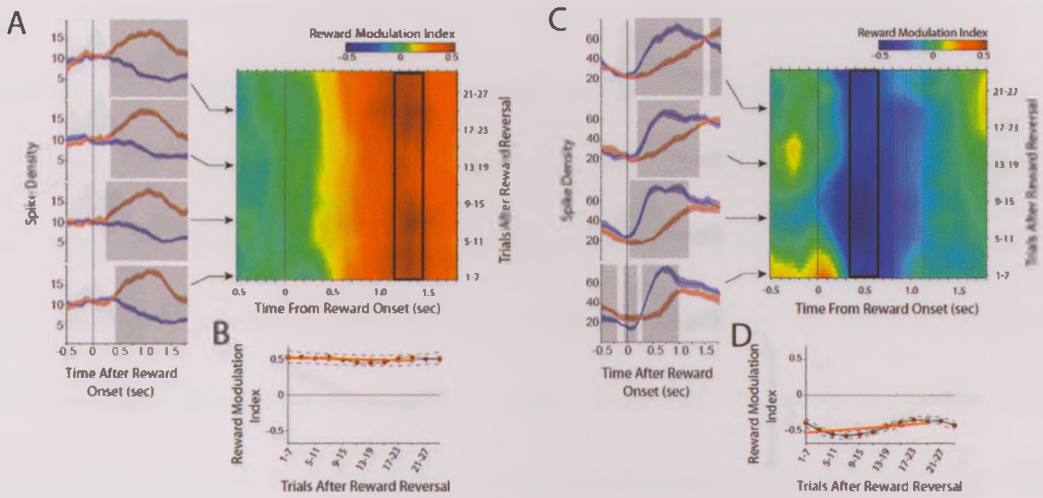


Figure 8. **Examples of constant reward encoding neurons.** (a) Example neuron showing a greater firing rate for high vs. low rewarded outcomes throughout all trials in the block. Shown to the left are representative spike density plots over time from reward onset (*x-axis*) for high (red) and low (blue) rewarded trials. Each panel represents the average spike density for a set of 7 trials as indicated by the arrows. (b) Reward modulation index (*y-axis*) of the firing rate difference between high and low rewarded trials for the same neuron as in (a). The panel shows data from a vertical cross-section of (a) at time 1.3 sec after reward onset, outline by a black rectangle, when reward modulation was at maximal statistical significance (*Anova*, F-test, $p \leq 0.05$). Red circles are actual data points. Unfilled circles show a 2nd order polynomial fit with its confidence range in dashed lines. The red line is a linear regression fit. (c, d) Same format as in (a, b) but showing an example neuron with a greater firing rate for low vs. high rewarded outcomes and a maximal reward modulation effect at 0.5 sec after reward onset.

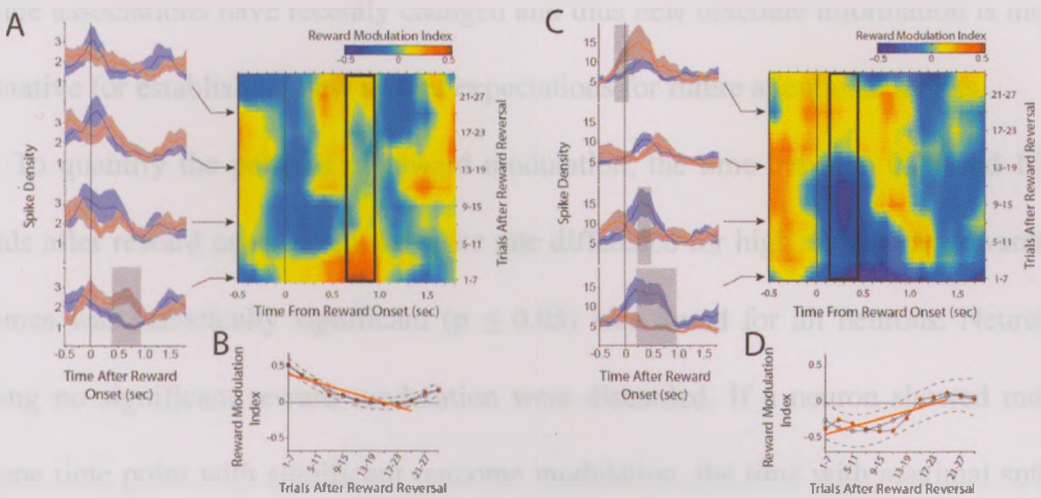


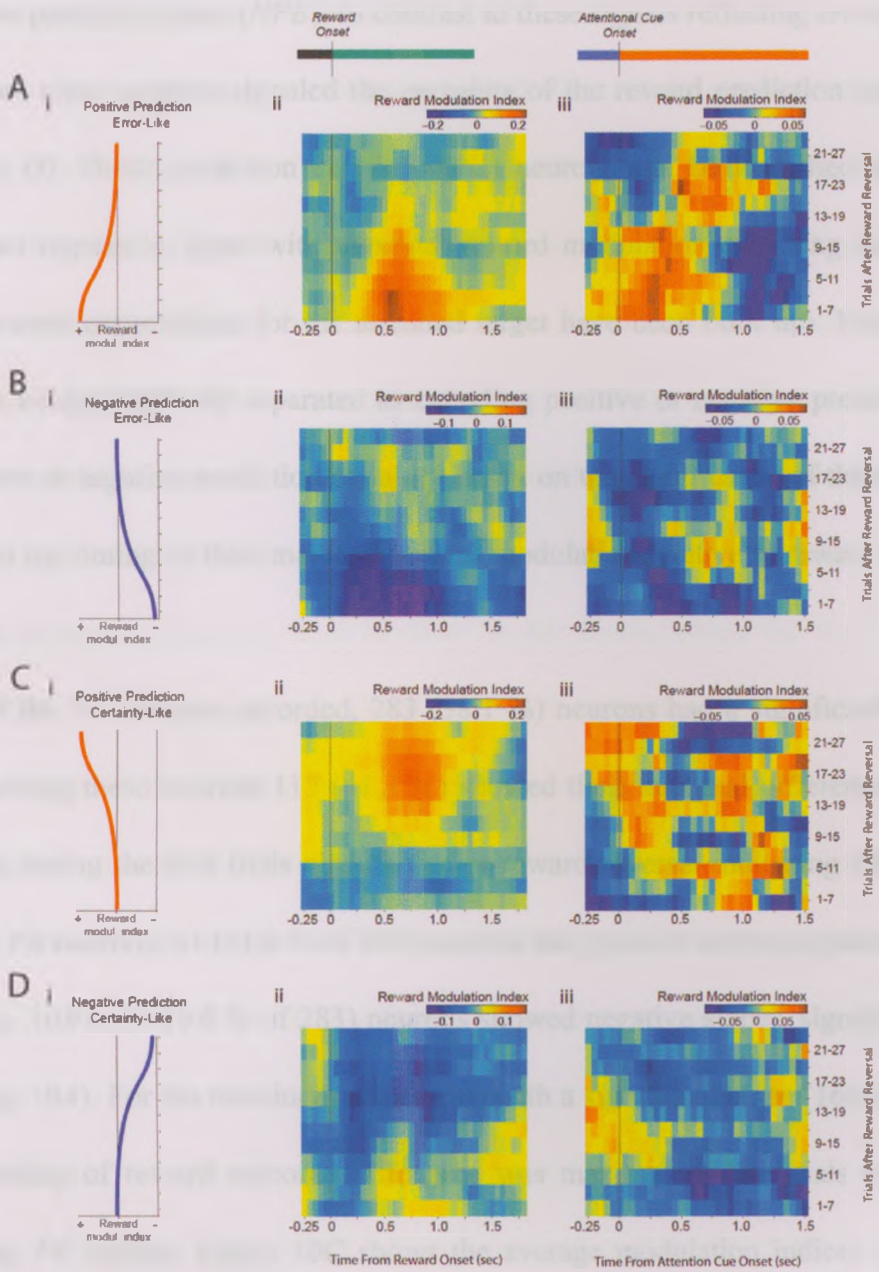
Figure 9. **Examples of changing reward encoding neurons.** (a) Example neuron showing a greater firing rate for high vs. low rewarded outcomes only during the first trials after reward-association reversal. Shown to the left are representative spike density plots over time from reward onset (*x-axis*) for high (red) and low (blue) rewarded trials. Each panel represents the average spike density for a set of 7 trials as indicated by the arrows. (b) Reward modulation index (*y-axis*) of the firing rate difference between high and low rewarded trials for the same neuron as in (a). The panel shows data from a vertical cross-section of (a) at time 0.8 sec after reward onset, outlined by a black rectangle, when reward modulation was at maximal statistical significance (*Anova*, F-test, $p \leq 0.001$). Red circles are actual data points. Unfilled circles show a 2nd order polynomial fit with its confidence range in dashed lines. The red line is a linear regression fit. (c, d) Same format as in (a, b) but showing an example neuron with a greater firing rate for low vs. high rewarded outcomes only during the first half of trials after reward-association reversals. Maximal reward modulation effect for (c,d) at 0.3 sec after reward onset and regression slope fit ($p \leq 0.01$).

outcome associations have recently changed and thus new outcome information is most informative for establishing new reward expectations for future attentional targets.

To quantify the patterns of reward modulation, the time between 0.25 and 1.25 seconds after reward onset when the spike rate difference for high versus low rewarded outcomes was statistically significant ($p \leq 0.05$) was found for all neurons. Neurons showing no significant reward modulation were discarded. If a neuron showed more than one time point with significant outcome modulation, the time with maximal spike rate difference was selected. A linear regression line was then fit to the distribution of reward modulations across trials at this time point. Figures 8*B and D* show the reward modulation and regression fit for the example neurons in Figure 8*A and C* respectively. Figures 9*B and D* show the reward modulation and regression fit for the example neurons in Figure 9*A and C* respectively. A flat regression slope characterizes neurons with a constant reward modulation. A significantly positive or negative slope ($p \leq 0.05$) characterizes neurons with varying reward modulation across trials in a block.

Neurons having a significant slope with the maximal difference in firing rate occurring in early trials (*Figure 9*) are conceptually akin to prediction error signals (Schultz et al., 1997). Neurons responding more strongly to the outcome of a high-rewarded target color, when the same color was recently low rewarded (in the last trials of the previous block), are signaling a positive prediction error (*PPE*): a positive difference between the experienced and expected outcome. Conversely, neurons responding more strongly to a low-reward outcome immediately after the color-reward reversal are signaling a negative difference between experienced and expected outcome:

Figure 10. **Population results for different neuron types signaling reward outcome.** (a) Reward modulation for neurons (n=56) showing increased firing rate for high vs. low reward outcomes in early trials only. (a, i) Illustration of the reward modulation characteristics of this neuronal classification (corresponding to a negative regression line slope in *Figure 4*). (a, ii) Population average of the reward modulation index (*coloured axis*) for trials in a block (*y-axis*)calculated in successive 300msec time windows relative to reward onset (*x-axis*). (a, iii) Reward modulation index for trials in a block (*y-axis*) aligned to attentional cue onset (*x-axis*) for the same population as in (a, i and ii). (b-d) Same format as in (a) but for neurons (n=61) showing increased firing rate for low vs. high reward outcomes early in the block corresponding to a positive regression line slope (b), neurons (n=77) showing increased firing rate for high vs. low reward outcomes late in the block corresponding to a positive regression line slope (c), neurons (n=89) showing increased firing rate for low vs. high reward outcomes late in the block corresponding to a negative regression line slope (d).


Reward Onset
Attentional Cue Onset
A
i
Positive Prediction Error-Like
ii
Reward Modulation Index
-0.2 0 0.2
iii
Reward Modulation Index
-0.05 0 0.05
21-27
17-23
13-19
9-15
5-11
1-7
Trials After Reward Reversal
-0.25 0 0.5 1.0 1.5
+ Reward modul index -
B
i
Negative Prediction Error-Like
ii
Reward Modulation Index
-0.1 0 0.1
iii
Reward Modulation Index
-0.05 0 0.05
C
i
Positive Prediction Certainty-Like
ii
Reward Modulation Index
-0.2 0 0.2
iii
Reward Modulation Index
-0.05 0 0.05
D
i
Negative Prediction Certainty-Like
ii
Reward Modulation Index
-0.1 0 0.1
iii
Reward Modulation Index
-0.05 0 0.05
Time From Reward Onset (sec)
Time From Attention Cue Onset (sec)

a negative prediction error (*NPE*). In contrast to these signals reflecting *errors* in reward prediction, some neurons signaled the *certainty* of the reward prediction itself (*Figure 10C and D*). These prediction certainty (PC) neurons can be identified as having a significant regression slope with maximal reward modulation occurring in later trials (when reward expectations for the attended target have been built up). Neurons could therefore be quantitatively separated as encoding positive or negative prediction errors or positive or negative prediction certainty based on the significance of their regression slope and the timing of their maximal reward modulation relative to reward association reversal.

Of the 742 neurons recorded, 283 (38.1 %) neurons had a significant regression slope. Among these neurons 117 (41.3 %) showed their maximal difference of reward encoding during the first trials after the color-reward reversal indicating *PE* signaling. Of these *PE* neurons, 61 (21.6 % of 283) neurons had positive slopes, signaling negative *PE's* (Fig. 10*B*). 56 (19.8 % of 283) neurons showed negative slopes, signaling positive *PE's* (Fig. 10*A*). For the remainder of neurons with a significant slope (166 of 283, 58.7 %), encoding of reward outcome difference was maximal in late trials in the block indicating *PC* signals. Figure 10*C* shows the average modulation indices for neurons having a positive slope, n = 77 (27.2 % of 283), signifying stronger activity after high reward outcomes late in the trial (*positive prediction certainty* neurons). 89 (31.5 % of 283) neurons had a negative a slope (Fig. 10*D)*, signifying stronger activity after low reward outcomes in late trials of a block (when the certainty of reward outcomes was measurable in the behavioral error analysis) (*see* Fig. 6*C,D)*.

3.3 - Relationship Between Reward Modulation in the Outcome and Cue Period

Neurons conveying positive reward predictions in the reward outcome period were on average likewise modulated during the attentional cue period. Figure 10*A,iii* shows these positive PE neurons were more active following the attentional cue, when the cue directed attention to a high rewarded target stimulus. During the first 1-17 trials, the reward modulation emerged early in response to the cue onset and became evident within the very first analysis windows (0.3 sec width) around the onset of the attention cue period: when the monkey had to select either the high or low target stimulus for covert attentional processing. In later trials in the block, where the reward outcome signals were not evident anymore, these positive PE type neurons continued to show a positive attentional modulation, but with a greater latency relative to cue onset.

Reward modulation during the attentional cue period was likewise observed for the subpopulation of neurons conveying a negative prediction error response in the reward outcome period (Fig. 10*B,ii*). These neurons showed a lower spike rate following the attention cue onset, when the cue directed attention to the low rewarded stimulus, again mimicking the reward response as obtained in the reward outcome period (*Figure 10B,iii*). This reward modulation of the cue induced response remained evident throughout the block of trials (until trial sets encompassing trial 25), and appeared to gradually decrease in latency relative to the onset of the attentional cue.

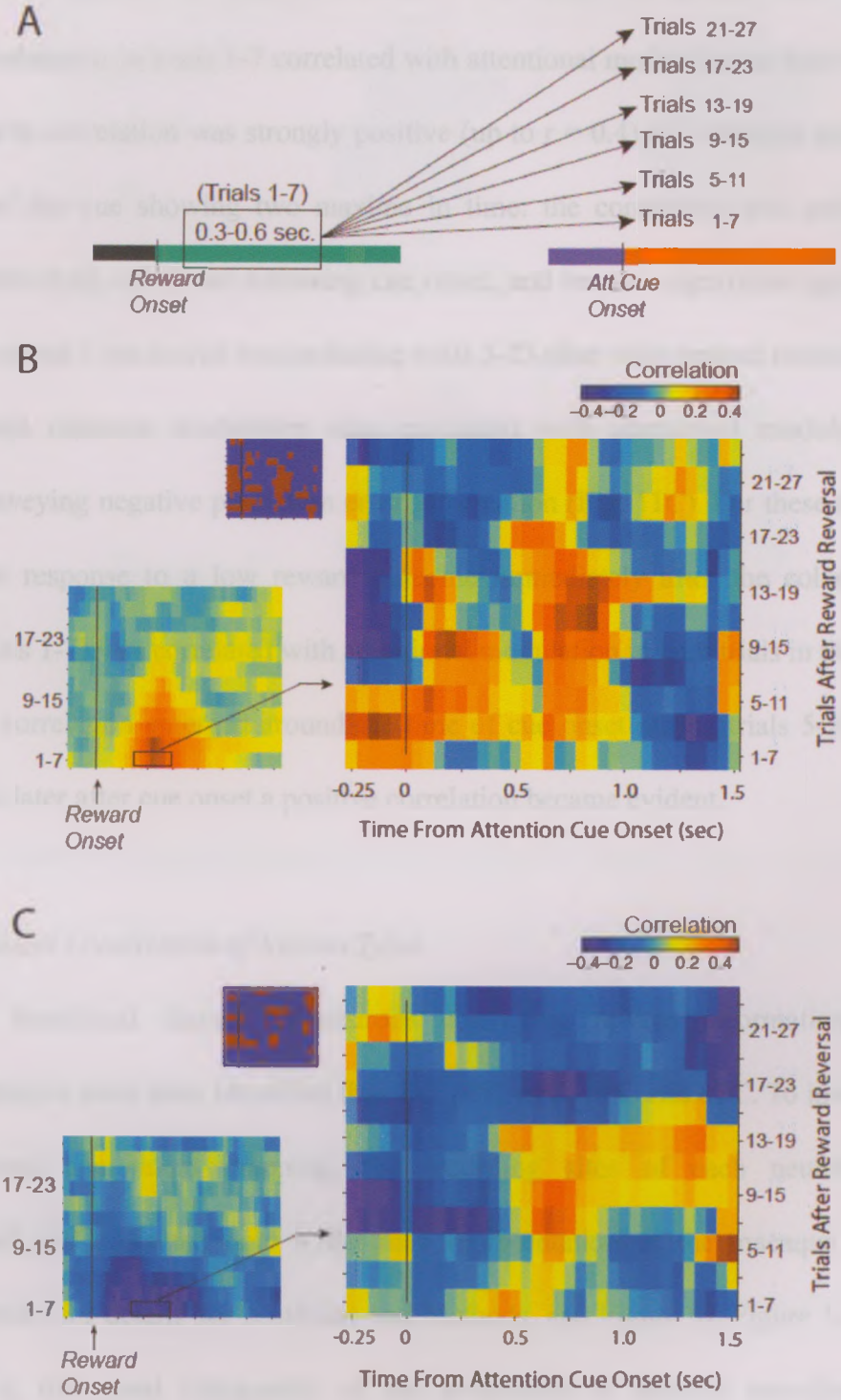
A more complex pattern of attention cue induced modulation was observed for neurons conveying the certainty of reward associated with the attentional target (Fig. 10*C,D*). Neurons responding stronger to high versus low reward outcomes (conveying

positive prediction certainty), showed an overall more positive attentional modulation for the high versus low rewarded target during the attentional cue period at varying times relative to cue onset (Fig. 10*C,iii*). Attentional modulation during the cue period was likewise evident for neurons conveying negative prediction certainty during the reward outcome period (Fig. 10*D,ii*). These neurons reduced their firing rate for low rewarded attentional targets particularly during the early trials of the block. This attentional effect emerged around 0.5 sec. after attention cue onset, and became less evident in later trials of the block (Fig. 10*D,iii*).

This attentional modulation emerging during the attention cue period could be independent of the modulation in the reward period. Alternatively, the outcome signals during the reward period could directly relate to the modulation during attentional selection. This interrelation would be expected for reward prediction error neurons, if their signals are instrumental for future processing (Schoenbaum et al., 2009; Matsumoto et al., 2007). To test this, a correlation analysis of the reward outcome modulation with the attentional modulation in the cue period was performed. Figure 11*A* illustrates the procedure. The average reward modulation for each neuron in the reward epoch (0.3 - 0.6 sec. following reward onset) was calculated during trials 1-7 after the color-reward reversal. This reward outcome modulation was then correlated with the attentional modulation observed across all trials in the block and across the whole time window following the cue onset.

Figure 11*B* shows the resulting correlation matrix for neurons signaling positive *PEs*. For these neurons, the reward outcome modulation showed significant correlations

Figure 11. **Correlation of early reward outcome modulation and attentional selection.** (a) Illustration of the time window for the correlation analysis. Correlation was performed between the average reward outcome modulation between 300 to 600 msec after reward onset in trials 1 to 7 with the reward modulation in all time windows after attentional cue onset in all trial sets. (b) Correlation matrix of early-trial reward modulation (*outlined in left panel*) with reward modulation during the attentional cue period for the same population of neurons with a positive prediction error response (*see Figure 10a, i*). (c) Correlation matrix of early-trial reward modulation (*outlined in left panel*) with reward modulation during the attentional cue period for the same population of neurons with a negative prediction error response (*see Figure 10b, i*).


A
Trials 21-27
Trials 17-23
Trials 13-19
Trials 9-15
Trials 5-11
Trials 1-7
(Trials 1-7)
0.3-0.6 sec.
Reward Onset
Att Cue Onset
B
Correlation
−0.4 −0.2 0 0.2 0.4
21-27
17-23
13-19
9-15
5-11
1-7
Trials After Reward Reversal
17-23
9-15
1-7
Reward Onset
-0.25 0 0.5 1.0 1.5
Time From Attention Cue Onset (sec)
C
Correlation
−0.4 −0.2 0 0.2 0.4
21-27
17-23
13-19
9-15
5-11
1-7
Trials After Reward Reversal
17-23
9-15
1-7
Reward Onset
-0.25 0 0.5 1.0 1.5
Time From Attention Cue Onset (sec)

($p \leq 0.05$) with attention cue induced modulation at various trial epochs. Most notably, outcome modulation in trials 1-7 correlated with attentional modulation in later trials in the block. The correlation was strongly positive (up to $r = 0.4$) and emerged early after the onset of the cue showing two maxima in time: the correlation was statistically evident within 0.05 - 0.45 sec following cue onset, and became significant again from 0.55 sec. to about 1 sec in trial sets including trials 3-23 after color-reward reversal.

Reward outcome modulation also correlated with attentional modulation in neurons conveying negative prediction error information (Fig. 11*C*). For these neurons, the stronger response to a low reward outcome immediately after the color-reward reversal (trials 1-7) was correlated with attentional modulation in later trials in the block. A negative correlation emerged around the time of cue onset during trials 5-23 in the block, while later after cue onset a positive correlation became evident.

3.4 - Anatomical Localization of Neuron Types

Four functional classes of neurons conveying reward information about attentional targets have been identified thus far: pPE, nPE, pPC and nPC. To investigate the functional anatomical mapping, the recording sites of each neuron were reconstructed and projected onto a flat map representation of the macaque fronto-cingulate brain (for details *see* Materials and Methods, and Figure 5). Figure 12 shows the resulting functional topography of the proportion of neurons encoding each functional class. These functional maps reveal remarkable spatial structure, despite the very small number of neurons in each functional class (n = 56-89, *see* above), and the

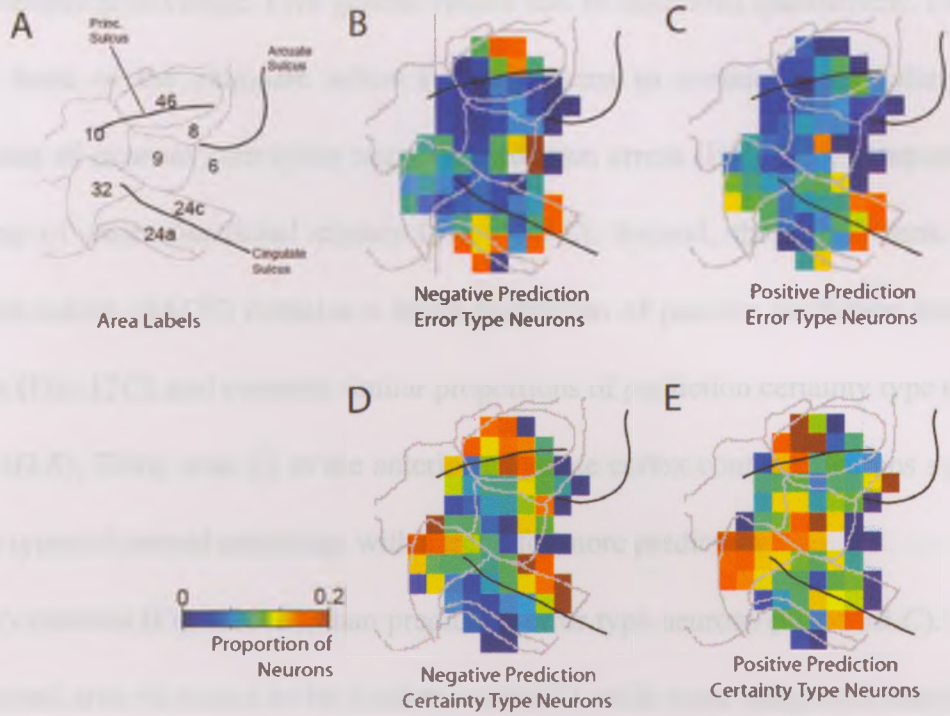


Figure 12. **Anatomical localization of reward prediction signals within the fronto-cingulate cortex.** (a) Flat map outline of fronto-cigulate cortex with area outlines in grey and sulci traced in black (*see* Figure 4a). (b) Flat map outline of fronto-cingulate cortex showing the proportion of neurons (*coloured axis*) with a significant negative prediction error response (*see text for details*). (c-e) Same format as in (b) but showing the proportion of neurons with a significant positive prediction error response (c), negative prediction certainty-type response (d), and positive prediction certainty-type response (e).

large anatomical coverage. Five general results can be described qualitatively. First, the ventral bank of the cingulate sulcus (vACC) seems to contain a particularly large proportion of neurons conveying negative prediction errors (Fig. 12*B*), compared with encoding of other functional classes (Fig. 12*C-E*). Second, the dorsal bank of the cingulate sulcus (dACC) contains a larger proportion of positive prediction error type neurons (Fig. 12*C*), and contains similar proportions of prediction certainty type neurons (Fig. 12*D,E*). Third, area 32 in the anterior cingulate cortex contains neurons signaling all four types of reward encodings with a trend for more prediction certainty neurons (Fig. 12*D,E*), than prediction error type neurons (Fig. 12*B-C*). Fourth, dorsolateral area 46 seems to be similar to area 32, with more ubiquitous encoding of prediction certainty, although the absolute number of reward modulated neurons in area 46 was considerably lower compared to area 32. Finally, lPFC area 9 contains neurons which most likely encode positive prediction certainty (Fig. 12*E*), with few neurons encoding any of the other functional classes.

Chapter 4 - Discussion:

4.1 - Behavioural Patterns resulting from Changing Attentional Target Value

The expectation of reward affects how we behave. The prospect of reward influences decision making, attentional allocation and the configuration of cognitive rules or 'state' representations that put our responses into context (Rowe et al., 2008). When a subject expects reward, accuracy, attention and perceptual acuity have been

shown to increase, and reaction times shown to decrease. The behavioural results from this study showed a pattern of decreasing errors and reaction times for highly rewarded trials as the block progressed (*Figure 6*). The animals began to expect different amounts of reward for correct completion of the trial, and their performance changed in accordance with these differing expected values.

As mentioned, errors fell into two general categories: those occurring 500 to 600msec following rotation of the distracting stimulus and fixation breaks occurring before rotation of either stimulus. For the first type of error, the monkey simply broke fixation without actually discriminating rotation direction, yet this occurred within the response time window for the distracting stimulus. Such errors have occurred in other studies (Taylor & Fragopanagos, 2005; Fries et al., 2008) and are indicative of a capture of attention by the distracting stimulus and a failure to sustain top-down attention on the target stimulus. The second type of error is likely indicative of a lack of motivation to sustain attention as they tended to occur more frequently when the monkey was cued to the lower-rewarded target during later trials in the block.

The effects of reward expectancy on performance developed consistently after roughly 10 trials into the block. Before this, the animal's behaviour showed no significant indication of reward modulation however there was a trend for errors made during rotation of the distractor and the reaction times to reverse their reward modulation from the beginning to the middle of the block. This behavioural pattern sets up a timeline which allows one to infer how certain or uncertain the monkey is about target value associations at any point in the block and thus also where the animal is in

the reinforcement learning process (*see Figure 13*). In the early trials, before reward effects tend to reverse, the monkey feels certain of his reward predictions but because of the preceding reversal these are incorrect and outcomes are thus unexpected. Throughout the middle of the block, when reward effects are showing a reversing trend, the monkey is learning novel reward associations, he is less certain in his reward predictions and thus outcomes are less unexpected. Later in the block, when behavioural data develop significant reward effects, the monkey is again certain in his reward predictions which are proving to be correct, and outcomes are thus entirely expected. This cycle of behaviour when learning in a changing environment is intuitive and well understood, the neuronal mechanisms which underlie this pattern of behaviour, however, are less clear.

4.2 - Patterns of Reward Outcome Encoding for Attentional Targets

Neurons recorded in this study showed a similar proportion and latency of outcome encoding as seen in other studies recording from frontal and cingulate areas (Luk & Wallis, 2009; Matsumoto et al., 2007). The proportion of neurons encoding reward outcome peaked roughly 500msec after reward onset and declined steadily throughout the remaining 1.75sec reward period (*Figure 7a*). When analyzing outcome-encoding neurons across trials in a block, some different patterns of outcome encoding emerged.

The majority of neurons showing significant reward modulation maintained their outcome encoding steadily across all trials within a given block (*Figure 8*), some

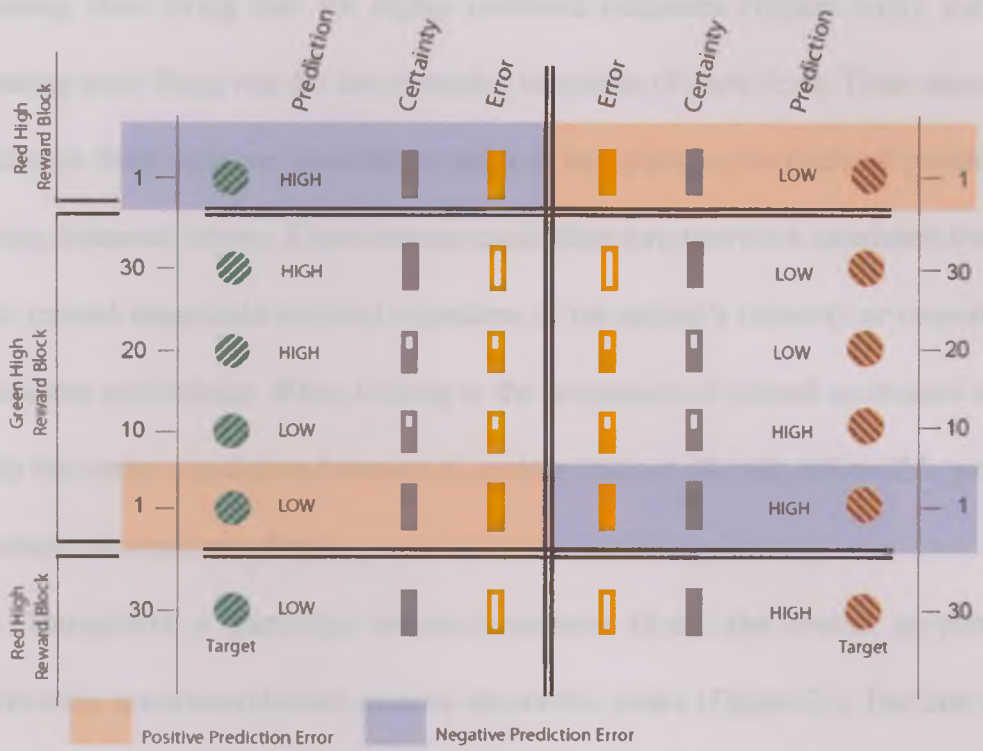


Figure 13. **Diagrammatic representation of inferred predictions, certainty and prediction errors within task blocks.** Going from a "red high rewarded block" to a "green high rewarded block", the monkey maintains predictions of low reward when cued to the green target and high reward when cued to red. Recent outcome history makes him fairly certain in these predictions yet due to the preceding reversal they are erroneous. Inferred predictions, level of certainty and prediction errors are tracked relative to trial from reward reversal (*x-axis*). As the monkey adjusts his outcome expectations and prediction errors no longer occur. Following the next reward reversal, the cycle begins anew. Times of inferred positive prediction errors are encased in pink and times of inferred negative prediction errors are encased in blue.

increasing their firing rate for highly rewarded outcomes (*Figure 8a,b*), the others increasing their firing rate for low rewarded outcomes (*Figure 8c,d*). These neurons did not change their outcome modulation index at any point in the cycle of reward based learning discussed above. These neurons could therefore provide a consistent indication of the reward magnitude received regardless of the animal's certainty or uncertainty in its outcome predictions. When looking at the proportion of reward modulated neurons within the entire population however, it is clear that not all cells follow this pattern of consistent outcome encoding.

Throughout a particular reward-association block, the overall proportion of significantly reward-modulated neurons shows two peaks (*Figure 7c*). The first peak in outcome encoding occurs early in the block, during the first 7 trials. The second peak occurs later, spanning roughly trials 13 to 21. This second peak corresponds with the emergence of significant reward-induced behavioural effects. These two peaks could reflect single neurons encoding reward throughout the block to varying degrees, but likewise it could reflect different populations of neurons encoding reward outcomes at early and later points during the reward based learning process.

4.3 - Patterns of Outcome Encoding: Prediction Error

Further single neurons examples reveal there are neurons whose reward modulation varies across trials and that encode outcomes only at specific times in the block. Figure 9 shows two single neuron examples of cells that encode reward outcome information only early on in the block. These neurons signal reward outcomes only at

the points in the block when the monkey is feeling certain about his reward predictions yet they are proving to be incorrect (*Figure 14B*). These cells are found within the ACC (*discussed in detail in section 4.6*) and are akin to the prediction error neurons recorded by Amiez et. al. (2005) and Matsumoto et al. (2007) but with two important differences. Firstly, unlike the neuronal finding from the Amiez et al. (2005) study these neuronal examples show both positive and negative prediction error encoding. While Amiez et al., 2005 do note the importance of investigating whether their negative prediction error neurons also encode *positive* prediction errors (in a similar style to dopaminergic neurons) the cells recorded here support findings from Matsumoto et al. (2007) showing positive and negative prediction errors to be encoded by different cells. Secondly, unlike the prediction error neurons recorded by Matsumoto et al. (2007) these neurons are encoding errors in the predicted value of attentional stimuli rather than actions. As mentioned above the ACC is persistently linked with action value associations (Hadland et al., 2003; Matsumoto et al., 2007; Kennerley et al., 2006; Gläscher et al., 2009). While some studies have in fact shown stimulus value encoding in the ACC (Kennerley et al., 2009), the results discussed here clearly show encoding of the value of *attentional targets*. This is an entirely different entity from action values or stimulus values and is a novel finding in the ACC. The finding of prediction error signals for attentional targets in the ACC suggests this area follows a pattern of outcome encoding essential for the process of reinforcement guided learning but that it encompasses a much broader range of value associations than is suggested by the current literature.

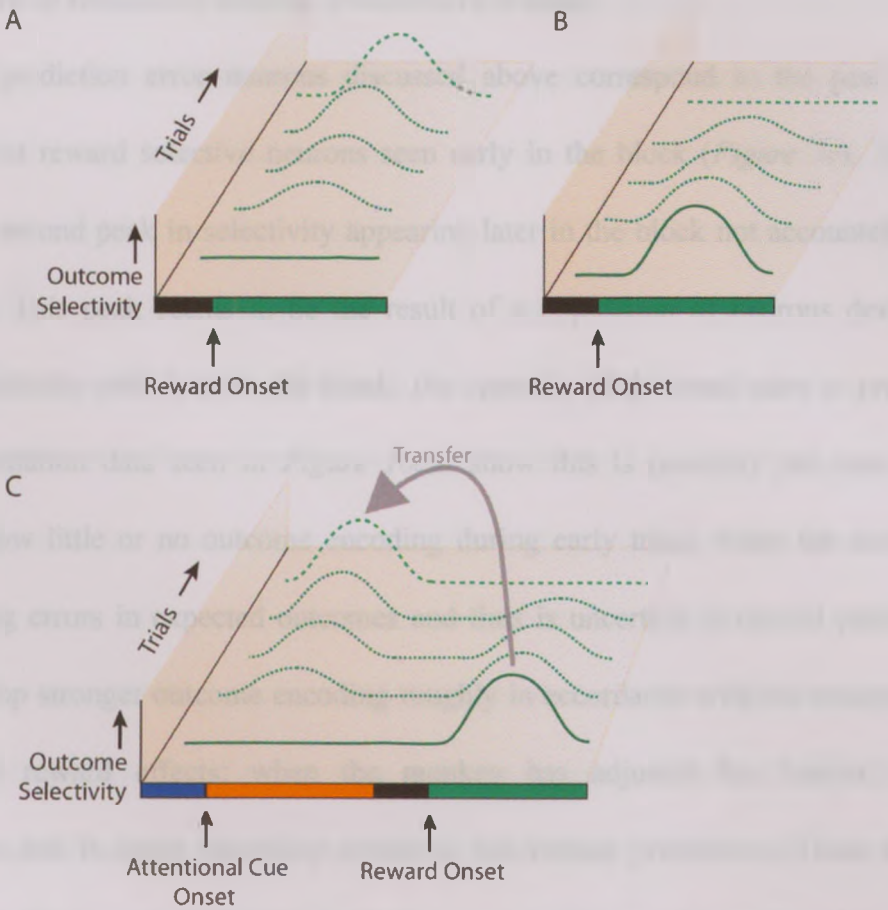


Figure 14. **Diagrammatic representation of prediction certainty neurons, prediction error neurons and the transfer of outcome selectivity.** (A) Illustration of a prediction certainty neuron. Outcome selectivity (*y-axis*) relative to time after reward onset (*x-axis*) increases as trials progress within a block (*z-axis*). (B) Same format as (A) but for a prediction error neuron. Selectivity decreases as trials progress within a block. (C) Illustration of the transfer of outcome selectivity (*y-axis*) by a prediction error neuron from the reward outcome period to the attentional cue period (*x-axis*) as trials progress in a block (*z-axis*).

4.4 - Patterns of Outcome Encoding: Prediction Certainty

The prediction error neurons discussed above correspond to the peak in the proportion of reward selective neurons seen early in the block (*Figure 7c*). There is, however a second peak in selectivity appearing later in the block not accounted for by these cells. This peak seems to be the result of a population of neurons developing reward selectivity only later in the block, the opposite of the trend seen in prediction errors. Population data seen in *Figure 10c,d* show this is precisely the case. These neurons show little or no outcome encoding during early trials, when the monkey is experiencing errors in expected outcomes and thus is uncertain in reward predictions. They develop stronger outcome encoding roughly in accordance with the emergence of behavioural reward effects: when the monkey has adjusted his internal reward associations and is again becoming certain in his reward predictions. These neurons appear to be the inverse of reward prediction error neurons and are referred to here as reward prediction certainty neurons (*Figure 14A*).

As mentioned, the ACC and other frontal areas encode a variety of aspects of the reward environment beyond outcomes and outcome associations. The ACC and lPFC have been shown to encode information such as one's certainty or uncertainty in reward estimates (Behrens et al., 2007), an aspect closely tied to environmental volatility (Dayan et al., 2000; 2006). While neurons have been recorded which encode directionless prediction errors akin to encoding general prediction uncertainty (Procyk et al., 2000) the neurons found here signal prediction *certainty*. As predictable rewards are

also able to influence behaviour (Rowe et al., 2008), these prediction certainty neurons may be useful for maintaining responses in order to exploit known sources of reward.

In this study, four classifications of neurons have been described: positive prediction error (pPE), negative prediction error (nPE), positive prediction certainty (pPC), and negative prediction certainty (nPC). Each class of neuron shows specific patterns of outcome encoding but for these cells to be involved in reinforcement guided learning, their outcome information must somehow transfer from the reward period to the reward predictive event, which is typically a single sensory stimulus conveying information about either the magnitude, probability or delay of reward delivery (Schoenbaum et al., 2009). In the task used here, the reward predicting event is the onset of the attentional cue. This link between outcome signals and behavioural adjustment has remained unclear in studies of reward based learning and decision making (Amiez et al., 2005; Mayr et al., 2003). The following section will focus on the transfer of reward prediction error and prediction certainty signals from the reward period to the attentional cue period.

4.5 - Transfer of Reward Prediction Error and Certainty Signals to the Attentional Cue Period

In reinforcement guided learning, outcome contingencies change and differences in what is expected upon making a given choice and what is obtained must be noted and used to adjust outcome expectations in the future. These continuously updated outcome expectations then serve to bias cognitive control processes toward the most highly

rewarded or most behaviourally beneficial option. For such biasing to occur, neurons noting discrepancies between expected and experienced outcomes -prediction error neurons- must transfer their information from the outcome period to the stimulus or response selection period where it can exert overt behavioural or covert attentional effects. In the context of this study, the prediction error neurons, or the prediction certainty neurons, must transfer their information to the attentional cue period. Figure 9 shows population data for pPE, nPE, pPC and nPC neuron types during the reward period alongside plots of modulation of expected reward during the attentional cue period for these same neuron types. The prediction error neurons tend to show a mimicking of the reward response during the attentional cue period with some variation in onset latency. The prediction certainty neurons showed a more complex pattern of attentional cue-induced reward modulation (*see Results section for a more detailed description*). The pattern of corresponding activity could be the result of independent processes or the two temporally distinct reward modulation signals could be related.

Correlation analysis between early prediction error signals and reward modulation during the corresponding attentional cue period showed significant correlations for both pPE and nPE type neurons (*Figure 11*). This suggests a transfer of information from the outcome period to the behaviourally relevant cue period is occurring within a defined set of frontal neurons (*Figure 14C*). Such an idea is in line with proposals that interractions between reward information and cognitive processes occur within local regions of the prefrontal cortex rather than within a 'global workspace' for decision making (Dehaene & Changeux, 2000). LPFC neurons for example have been shown to encode

combinations of reward magnitude and pending responses (Wallis & Miller, 2003) and ACC neurons have been shown to multiplex information about both reward and selected actions (Hayden & Platt., 2010). Such examples show local regions linking affective information with cognitive processes. The data presented here show classes of neurons linking affective information between cognitively distinct periods.

Importantly, ACC prediction error neurons may transfer their information to a time when they can affect choice behaviour but the ACC must be in a position, anatomically, to influence cognitive processes for this to occur. An fMRI study showed ACC activity predicted both prefrontal cortex (PFC) activity and behavioural adjustment (Kerns et al., 2004). Greater ACC activity was associated with increased PFC activity and with behavioural adjustment on subsequent trials, suggestive of a role of the ACC in influencing cognitive control processes in the PFC. A neuronal tracer study (Medalla & Barbas, 2010) found more definitive evidence linking the ACC with modulation of cognitive processes, particularly with influencing attentional control. Tracer injected into ACC area 32 found this area preferentially innervated presumed inhibitory neurons in dlPFC areas 46 and 9. These inhibitory neurons innervate distal dendrites of pyramidal neurons (DeFelipe et al., 1989), a pattern of connectivity well suited to minimize noise in active neurons (Wang et al., 2004). As attention involves selectively focusing on one aspect of the environment while filtering out irrelevant information (Anderson, 2004), this connectivity profile is suggestive of a role for the ACC in biasing attentional selection processes within the lPFC.

4.6 - Clustering of Prediction Error and Certainty Signals within the Frontal Cortex

As discussed above, the ACC and dlPFC are both important structures in reinforcement guided learning but their precise roles differ. The ACC encodes strong reward outcome and expectancy signals and combines more aspects of reward information than other frontal areas (Kennerley et al., 2009). The ACC also tends to encode rewards along with the choices that led to these rewards (Matsumoto et al., 2003; Luk & Wallis, 2009) forming a type of choice-outcome memory (Matsumoto et al., 2003; Kennerley et al., 2006) which can be updated and referenced during decision making. The dlPFC on the other hand tends to encode reward values along with *upcoming* responses (Wallis 2007). Given the lPFC's role in generating response plans (Kim & Shadlen, 1999; Hoshi & Tanji, 2004) it has been proposed that values associated with particular stimuli or actions are calculated in areas such as the ACC or OFC using outcome histories and transferred to the dlPFC which uses this information to determine future behavioural responses (Wallis 2007).

The four functional classes of neurons described in this study: pPE, nPE, pPC and nPC could be distributed throughout the fronto-cingulate cortex or could be localized to distinct areas. Mapping of each neuronal class onto a fronto-cingulate flat map (*Figure 12*) revealed distinct clustering of functional subtypes into anatomically defined areas. There is a trend for prediction error neurons to predominate in the ventral bank of the ACC (vACC) and prediction certainty neurons to predominate in dorsolateral prefrontal cortex (dlPFC) area 9. Areas 46 and 32 both contain all four neuron types while the dorsal bank of the ACC (dACC) contains predominantly pPE and nPC neurons. This

clustering fits with the proposed functions of ACC and dlPFC areas described above. The ACC is purported to be necessary for using outcome histories to construct and modify choice-outcome associations which then contribute to reward expectations. Prediction error signals convey the need to adjust these associations and the ACC was found to contain a large proportion of nPE neurons. The dlPFC is purported to use reward association information in order to guide behavioural responses. The more likely it is that an association is correct, the more behaviourally useful it will be and dlPFC areas tended to contain a greater proportion of pPC neurons. Interestingly, the proportion of neuron types varied between the dorsal and ventral banks of the ACC with the dorsal bank containing larger proportions of both pPE and nPC neurons. In his expanded definition of cingulate areas, Vogt et. al. (2005) noted cytoarchitectural differences between the dorsal and ventral banks of the cingulate sulcus with the dorsal bank showing more "frontal" like features and the ventral bank being more phylogenetically cingulate in structure. This could explain the increased proportion of reward certainty neurons in the dACC relative to the vACC. While the anatomical localization results shown here are not definitive, they are suggestive of specific contributions of the ACC and dlPFC in reinforcement guided learning and decision making. These results are in line with current understanding of ACC and dlPFC function based on lesion, fMRI, neurocytology and single unit recording studies.

Chapter 5 - Summary and Conclusion:

The anterior cingulate cortex is a functionally heterogeneous brain region whose divisions and function have long been debated. Proposed functions have included conflict monitoring (Carter et al., 1998), error detection (Dehaene et al., 1994) and more recently a role in reinforcement guided learning. In the latter, the ACC is typically associated with integrating action-outcome associations over time in order to develop action value predictions. The results reported here show dynamic encoding of reward outcome and reward expectancy associated with attentional targets within the ACC as well as within the vmPFC and dlPFC. This expands the function of the ACC beyond merely action values and suggests this area serves a broader role in reinforcement guided learning and decision making.

These results also define two distinct patterns of changing reward modulation throughout the reward based learning process, as reward associations and outcome predictions change. Neurons encoding these patterns are classified as negative/positive prediction error neurons and negative/positive prediction certainty neurons. The prediction error neurons corroborate findings by Matsumoto et al., 2007 of separate positive and negative prediction error neurons within the ACC. The prediction certainty neurons show a pattern of activity opposite to that of prediction error neurons and are a novel finding.

Missing from the current understanding of ACC function in reinforcement guided learning is a connection between prediction error signals and behavioural adjustment (Kerns et al., 2004; Amiez et al., 2005). The findings shown here identify a correlation

between the prediction error neurons' reward selectivity during outcome periods with its reward selectivity preceding subsequent reward predictive events (the attentional cue). Given that ACC neurons are likely able to influence attentional processes (Medalla & Barbas), this could serve as a means for prediction error information to lead to adjustments in behaviour and, in particular, in selecting stimuli before a choice is made.

Finally, mapping of the neuron classifications described above onto the fronto-cingulate cortex revealed clustering of neuron types to different anatomical areas. The vACC contains a greater proportion of prediction error neurons while the dlPFC area 9 contains a greater proportion of prediction certainty neurons. Though not definitive, these results seem to support proposals by Wallis (2007) suggesting outcome predictions are developed and adjusted within the ACC and vmPFC and these predictions are used by the dlPFC to determine the behavioural response.

Further research in this area will help determine the precise contributions of the ACC and OFC in reinforcement guided learning and decision making and also increase our knowledge of how these brain areas interact with dlPFC areas. This will hopefully lead to the development of a useful framework for understanding how brain circuits implement reward based learning and flexible attentional control and how we are able to make optimal choices amidst continuously changing environmental circumstances. As schizophrenia, obsessive-compulsive disorder, attention deficit hyperactivity disorder, and depression have been linked to aberrant ACC function (MacDonald et al., 2005; Pittenger et al., 2006; Bush et al., 1999; Mayberg et al., 2005), increasing our

understanding in this area has the potential to increase our understanding and treatment of a variety of psychiatric disorders.

Bibliography:

Amiez, C., Joseph, JP., Procyk, E. Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* 12: 3447-52, 2005.

Amiez, C., Joseph, JP., Procyk, E. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* 7: 1040-55, 2006.

Anderson, VA., Godber, T., Smibert, E., Weiskop, S., Ekert, H. Impairments of attention following treatment with cranial irradiation and chemotherapy in children. *J. Clin Exp Neuropsychol.* 26(5): 684-97, 2004.

Averbeck, BB. and Seo, M. The statistical neuroanatomy of frontal networks in the macaque. *PLoS Comput. Biol.* 4: e1000050, 2008.

Baleydier, C. and Mauguiere, F. The duality of the cingulate gyrus in monkey. Neuroanatomical study and functional hypothesis. *Brain* 103: 525-54, 1980.

Barbas, H. Architecture and cortical connections of the prefrontal cortex in the rhesus monkey *Adv. Neurol.* 57: 91-115, 1992.

Barbas, H. and De Olmos, J. Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 300: 549-71, 1990.

Barbas, H. and Pandya, DN. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J.Comp Neurol.* 286: 353-75, 1989.

Barraclough, DJ., Conroy, ML. and Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7: 404-10, 2004.

Bayer, HM. and Glimcher, PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47: 129-41, 2005.

Bechara, A., Damasia, AR., Damasio, H., Anderson, SW. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition.* 50(1-3): 7-15, 1994.

Behrens, TE., Woolrich, MW., Walton, ME. and Rushworth, MF. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10: 1214-21, 2007.

Berger, B. Dopaminergic innervation of the frontal cerebral cortex. Evolutionary trends and functional implications. *Adv. Neurol.* 57: 525-44, 1992.

Brodmann, K. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. *Leipzig: Barth.*

Buckley, MJ., Mansouri, FA., Hoda, H., Mahboubi, M., Browning, PG., Kwok, SC., Phillips, A. and Tanaka, K. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science*. 325: 52-8, 2009.

Bush, G., Frazier, JA., Rauch, SL., Seidman, LJ., Whalen, PJ., Jenike, MA., Rosen, BR., Biederman, J. Anterior cingulate cortex dysfunction in attention-deficit hyperactivity disorder revealed by fMRI and the Counting Stroop. *Biol. Psychiatry.* 45 (12): 1542-52, 1999.

Carmichael, ST. and Price, JL. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J. Comp. Neurol.* 346: 366-402, 1994.

Carmichael, ST. and Price, JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 363(4): 642-64, 1995a.

Carmichael, ST. and Price, JL. Limbic connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 363(4): 615-41, 1995b.

Carter, CS., Braver, TS., Barch, DM., Botvinick, MM., Noll, D. and Cohen, JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280: 747-9, 1998.

Chiba, T., Kayahara, T., Nakano, K. Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, Macaca fuscata. *Brain Res.* 888(1): 83-101, 2001.

Cools, R., Robinson, OJ. and Sahakian, B. Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology* 33: 2291-9, 2008.

Corbetta, M. and Shulman, GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3: 201-15, 2002.

Courville, AC., Daw, ND. and Touretzky, DS. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* 10: 294-300, 2006.

Damaiso, AR. Descartes' error and the future of human life. *Sci Am.* 271(4): 144, 1994.

Daw, ND., Kakade, S. and Dayan, P. Opponent interactions between serotonin and dopamine. *Neural Netw.* 15: 603-16, 2002.

Dayan, P., Kakade, S. and Montague, PR. Learning and selective attention. *Nat. Neurosci.* 1218-23, 2000.

DeFelipe, J., Hendry, SH., Jones, EG. Visualization of chandelier cell axons by parvalbumin immunoreactivity in monkey cerebral cortex. *Proc Natl Acad Sci USA.* 86(6): 2093-7, 1989.

Dehaene, S. and Changeux, JP. Reward-dependent learning in neuronal networks for planning and decision making. *Prog Brain Res.* 126: 217-29, 2000.

Dickinson, DJ. The direct assessment: an alternative to psychometric testing. *J. Learn Disabil.* 13: 472-6, 1980.

Doya, K. Reinforcement learning: Computational theory and biological mechanisms. *HFSP J.* 1: 30-40, 2007.

Dum, RP. and Strick, PL. The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* 3: 667-89, 1991.

Eslinger, PJ. and Damasio, AR. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology.* 35(12): 1731-41, 1985.

Fellows, LK. and Farah, MJ. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb Cortex.* 15(1): 58-63, 2005.

Fellows, LK. and Farah, MJ. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb Cortex.* 17(11): 2669-74, 2007.

Fiorillo, CD., Tobler, PN. and Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: 1898-1902, 2003.

Fries, P., Womelsdorf, T., Oostenveld, R., Desimone, R. The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci.* 28(18): 4823-35, 2008.

Gläscher, J., Hampton, AN., O'Doherty, JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex.* 19(2): 483-95, 2009.

Glimcher, PW., Dorris, MC. and Bayer, HM. Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* 52: 213-256, 2005.

Haber, SN. and Knutson, B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. 35: 4-26, 2010.

Hadland, KA., Rushworth, MF., Gaffan, D., Passingham, RE. The effect of cingulate lesions on social behaviour and emotion. *Neuropsychologia.* 41(8): 919-31, 2003.

Hare, TA., Camerer, CF., Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science.* 324(5927): 646-8, 2009.

Haruno, M. and Kawato, M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J. Neurophysiol.* 95: 948-59, 2006.

Hayden, BY. and Platt, ML. Neurons in anterior cingulate cortex multiplex information and reward and action. *J Neurosci.* 30(9): 3339-46, 2010.

Hikosaka, O., Bromberg-Martin, E., Hong, S. and Matsumoto, M. New insights on the subcortical representation of reward. *Curr. Opin. Neurobiol.* 2: 203-8, 2008.

Hoshi, E. and Tanji, J. Area-selective neuronal activity in the dorsolateral prefrontal cortex for information retrieval and action planning. *J Neurophysiol.* 91(6): 2707-22, 2004.

Jacobs, BL. and Azmitia, EC. Structure and function of the brain serotonin system. *Physiol Rev.* 72: 165-229, 1992.

Kable, JW., and Glimcher, PW. The neurobiology of decision: consensus and controversy. *Neuron.* 63(6): 733-45, 2009.

Kahneman, D. and Tversky, A. 2000. Choices, values, and frames. In: Kahneman, D. and Tversky, A editors. New York: Cambridge University Press.

Kawagoe, R., Takikawa, Y. and Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci*. 1: 411-416, 1998.

Kennerley, SW., Walton, ME., Behrens, TE., Buckley, MJ. and Rushworth, MF. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 7: 940-7, 2006.

Kennerley, SW., Dahmubed, AF., Lara, AH. and Wallis, JD. Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 6: 1162-78, 2009.

Kerns, JG., Cohen, JD., MacDonald, AW. 3rd, Cho, RY., Stenger, VA., Carter, CS. Anterior cingulate conflict monitoring and adjustments in control. *Science.* 303 (5660): 1023-6, 2004.

Kim, JN. and Shadlen, MN. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci.* 2(2): 176-85, 1999.

Komura, Y., Tamura, R., Uwano, T., Nishijo, H and Ono, T. Auditory thalamus integrates visual inputs into behavioral gains. *Nat. Neurosci.* 8: 1203-1209, 2005.

Knishio, K. and Haber, SN. Primate cingulostriatal projection: limbic striatal versus sensorimotor striatal input. *J. Comp. Neurol.* 350: 337-56, 1994.

Lee, D., Rushworth, MF., Walton, ME., Watanabe, M., Sakagami, M. Functional specialization of the primate frontal cortex during decision making. *J Neurosci.* 27 (31): 8170-3, 2007.

Liu, Z. and Richmond, BJ. Response difference in monkey TE and perirhinal cortex: stimulus association related to reward schedules. *J Neurophysiol.* 83(3): 1677-92, 2000.

Luk, CH. and Wallis, JD. Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J Neurosci.* 29(23): 7526-39, 2009.

MacDonald, AW. 3rd, Carter, CS., Kerns, JG., Ursu, S., Barch, DM., Holmes, AJ., Stenger, VA., Cohen, JD. Specificity of prefrontal dysfunction and context processing deficits to schizophrenia in never-medicated patients with first-episode psychosis. *Am J Psychiatry.* 162(3): 475-84, 2005.

Matsumoto, K., Suzuki, W. and Tanaka, K. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* 5630: 229-32, 2003.

Matsumoto, M., Matsumoto, K., Abe, H. and Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10: 647-56, 2007.

Mayberg, HS., Lozano, AM., Voon, V., McNeely, HE., Seminowicz, D., Hamani, C., Schwalb, JM., Kennedy, SH. Deep brain stimulation for treatment-resistant depression. *Neuron.* 45(5): 651-60, 2005.

Mayr, U., Awh, E., Laurey, P. Conflict adaptation effects in the absence of executive control. *Nat Neurosci.* 6(5): 450-2, 2003.

McClure, SM., Laibson, DI., Loewenstein, G., Cohen, JD. Separate neural systems value immediate and delayed monetary rewards. *Science.* 306(5695): 503-7, 2004.

Medalla, M. and Barbas, H. Anterior cingulate synapses in prefrontal areas 10 and 46 suggest differential influence in cognitive control. *J Neurosci.* 30(48): 16068-81, 2010.

Miller, EK. and Cohen, JD. An integrative theory of prefrontal cortex function. *Ann Rev Neurosci.* 24: 167-202, 2001.

Mishkin, S. The interdependence of clinical neurology and neurophysiology. An historical review of the vestibulo-ocular reflex. *McGill Med J.* 33:80-97, 1964.

Mohanty, A., Gitelman, DR., Small, DM. and Mesulam, MM. The spatial attention network interacts with limbic and monoaminergic systems to modulate motivation-induced attention shifts. *Cereb. Cortex.* 11: 2604-13, 2008.

Mohanty, A., Egner, T., Monti, JM. and Mesulam, MM. Search for a threatening target triggers limbic guidance of spatial attention. *J. Neurosci.* 34: 10563-72, 2009.

Montague, PR., Dayan, P. and Sejnowski, TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16: 1936-47, 1996.

Morecraft, RJ., Cipolloni, PB., Stilwell-Morecraft, KS., Gedney, MT. and Pandya, DN. Cytoarchitecture and cortical connections of the posterior cingulate and adjacent somatosensory fields in the rhesus monkey. *J. Comp. Neurol.* 469: 37-69, 2004.

Morecraft, RJ. and Van Hoesen, GW. Cingulate input to the primary and supplementary motor cortices in the rhesus monkey: evidence for somatotopy in areas 24c and 23c. *J. Comp. Neurol.* 322: 471-89, 2008.

Neafsy, EJ., Terreberry, RR., Hurley, KM., Ruit, KG. and Frysztak, RJ. 1993. Anterior cingulate cortex in rodents: connections, visceral control functions, and implications for emotion. In: Vogt BA, Gabriel M, editors. Neurobiology of cingulate cortex and limbic thalamus. Boston: Birkhauser. p 206–223.

Niki, H. and Watanabe, M. Prefrontal and cingulate unit activity during timing behaviour in the monkey. *Brain Res.* 171: 213-24, 1979.

Nishijo, H., Yamamoto, Y., Ono, T., Uwano, T., Yamashita, J. and Yamashima, T. Single neuron responses in the monkey anterior cingulate cortex during visual discrimination. *Neurosci. Lett.* 226: 57-60, 1997.

Olds, J. and Milner, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.* 47: 419-27, 1954.

Ongür, D., An, X. and Price, JL. Prefrontal cortical projections to the hypothalamus in macaque monkeys. *J. Comp. Neurol.* 401: 480-505, 1998.

Padoa-Schioppa, C. and Assad, JA. Neurons in the orbitofrontal cortex encode economic value. *Nature.* 441(7090):223-6, 2006.

Palomero-Gallagher, N., Vogt, BA., Schleicher, A., Mayberg, HS. and Zilles, K. Receptor architecture of human cingulate cortex: evaluation of the four-region neurobiological model. *Hum. Brain Mapp..* 8: 2336-55, 2009.

Pandya, DN., Van Hoesen, GW. and Mesulam, MM. Efferent connections of the cingulate gyrus in the rhesus monkey. *Exp. Brain Res.* 42: 319-30, 1981.

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., Gross, CE. and Boraud, T. Shaping of motor responses by incentive values through the basal ganglia. *J. Neurosci.* 27: 1176-83, 2007.

Pearce, JM. and Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev.* 87(6): 532-552, 1980.

Pittenger, C., Krystal, JH., Coric, V. Glutamate-modulating drugs as novel pharmacotherapeutic agents in the treatment of obsessive-compulsive disorder. *NeuroRx.* 3 (1): 69-81, 2006.

Posner, MI. and DiGirolamo, GJ. (1998) Executive attention: Conflict, target detection, and cognitive control. In Parasuraman R. *The attentive brain*. Cambridge, Mass: MIT Press.

Preuschoff, K., Bossaerts, P. and Quartz, SR. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 3: 381-90, 2006.

Procyk, E., Tanaka, YL. and Joseph, JP. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci*, 3: 502-8, 2000.

Quilodran, R., Rothé, M. and Procyk, E. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57: 314-25, 2008.

Rescorla, RA. and Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.) *Classical conditioning II.* New York: Appleton-Century-Crofts. 1972.

Rowe, JB., Eckstein, D., Braver, T., Owen, AM. How does reward expectation influence cognition in the human brain? *J Cogn Neurosci.* 20(11): 1980-92, 2008.

Rudebeck, PH., Walton, ME., Millette, BH., Shirley, E., Rushworth, MF., Bannerman, DM. Distinct contributions of frontal areas to emotion and social behaviour in the rat. *Eur J Neurosci.* 26(8): 2315-26, 2007.

Rushworth, MF. and Behrens, TE. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci.* 11(4): 389-97, 2008.

Samejima, K., Ueda, Y., Doya, K. and Kimura, M. Representation of action-specific reward values in the striatum. *Science* 310: 1337-1340, 2005.

Santesso, DL., Dillon, DG., Birk, JL., Holmes, AJ., Goetz, E., Bogdan, R. and Pizzagalli, DA. Individual differences in reinforcement learning: behavioural, electrophysiological, and neuroimaging correlates. *Neuroimage.* 42: 807-16, 2008.

Schoenbaum, G., Roesch, MR., Stalnaker, TA., Takahashi, YK. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci.* 10 (12): 885-92, 2009.

Schultz, W., Dayan, P. and Montague, PR. A neural substrate of prediction and reward. *Science* 275: 1593-9, 1997.

Schultz, W. The phasic reward signal of primate dopamine neurons. *Adv. Pharmacol.* 42: 686-90, 1998.

Schultz, W. and Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23: 473-500, 2000.

Schultz, W. Getting formal with dopamine and reward. *Neuron* 36: 241-63, 2002.

Seo, H. and Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27: 8366-77, 2007.

Sutton, R.S. and Barto, A.G. *Reinforcement Learning* (MIT Press, Cambridge, Massachusetts, USA, 1998).

Taylor, JG. and Fragopanagos, NF. The interaction of attention and emotion. *Neural Netw.* 18(4): 353-69, 2005.

Tenenbaum, JB. and Griffiths, TL. Generalization, similarity, and Bayesian inference. *Behav. Brain Sci.* 24:629-40, 2001.

Thorndike. Some experiments on animal intelligence. *Science* 181: 818-24, 1898.

Tobler, PN., Fiorillo, CD. and Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* 307: 1642-5, 2005.

Varnäs, K., Halldin, C. and Hall, H. Autoradiographic distribution of serotonin transporters and receptor subtypes in human brain. *Hum. Brain Mapp.* 22: 246-60, 2004.

Vogt, BA,. Berger, GR., Derbyshire, SWG.. Structural and functional dichotomy of human midcingulate cortex. *Eur. J. Neurosci.* 18:3134–3144, 2003.

Vogt, BA., Pandya, DN., Rosene, DL. Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *J. Comp. Neurol.* 262: 256-70, 1987.

Vogt, BA., Vogt, L. 2004. Cingulate gyrus. In: Paxinos G, Mai J, editors. The human nervous system. 2nd ed. San Diego: Elsevier. p 915-945

Vogt, BA., Vogt, L., Farber, NB., and Bush, G. Architecture and neurocytology of monkey gyrus. *J. Comp. Neurol.* 485: 218-39, 2005.

Vogt, BA., Vogt, LJ., Nimchinsky, HofPR. 1997. Primate cingulate cortex chemoarchitecture and its disruption in Alzheimer's disease. In: Bloom FE, Bjorklund A, Hokfelt T, editors. Handbook of chemical neuroanatomy, the primate nervous system. Vol. 13. Part I. Amsterdam: Elsevier. p 455-528.

Wallis, JD. Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci.* 30:31-56, 2007.

Wallis, JD. and Miller, EK. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci.* 18 (7): 2069-81, 2003.

Walton, ME., Bannerman, DM., Alterescu, K. and Rushworth, MF. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J. Neurosci.* 23: 6475-9, 2003.

Wang, XJ., Tegnér, J., Constantinidis, C., Goldman-Rakic, PS. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc Natl Acad Sci USA.* 101(5): 1368-73, 2004.

Watanabe, M. Reward expectancy in primate prefrontal neurons. *Nature.* 382(6592): 629-32, 1996.

Watanabe, M. and Sakagami, M. Integration of cognitive and motivational context information in the primate prefrontal cortex. *Cereb Cortex.* Suppl 1: 101-9, 2007.

Williams, SM. and Goldman-Rakic, PS. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb. Cortex.* 3: 199-222, 1993.

Western

Dec. 30, 2008

'This is the Original Approval for this protocol'
'A Full Protocol submission will be required in 2012'

Dear Dr Everling

Your Animal Use Protocol form entitled
Role of Frontal Cortex in cognitive control
Funding Agency CIHR - Grant #R3104A12, NSERC DISCOVERY - Grant - PYFRR

has been approved by the University Council on Animal Care This approval is valid from **Dec. 30, 2008 to Dec. 31, 2009**. The protocol number for this project is **#2008-125 and replaces #2004-099-12.**

1 This number must be indicated when ordering animals for this project
2 Animals for other projects may not be ordered under this number
3 If no number appears please contact this office when grant approval is received
If the application for funding is not successful and you wish to proceed with the project, request that an internal scientific peer review be performed by the Animal Use Subcommittee office
4 Purchases of animals other than through this system must be cleared through the ACVS office Health certificates will be required

ANIMALS APPROVED FOR 4 Years

Species	Strain	Other Detail	Pain Level	Animal # Total for 4 Years
Other, add to detail	NHP - Macaca mullata or Macaca fascicularis	3-16 kg	D	20

REQUIREMENTS/COMMENTS
Please ensure that individual(s) performing procedures on live animals, as described in this protocol, are familiar with the contents of this document

c.c. Approved Protocol - S Everling, T Admans
Approval Letter - S. Everling, T Admans

Western

01.01.11
***This is the 2nd Renewal of this protocol**
*A Full Protocol submission will be required in **2013**

Dear Dr. **Everling**

Your Animal Use Protocol form entitled:

Role of Frontal Cortex in Cognitive Control

has had its yearly renewal approved by the Animal Use Subcommittee.

This approval is valid from **01.01.11 to 01.01.12**

The protocol number for this project remains as **2008-125**

1. This number must be indicated when ordering animals for this project.
2. Animals for other projects may not be ordered under this number.
3. If no number appears please contact this office when grant approval is received.
   If the application for funding is not successful and you wish to proceed with the project, request that an internal scientific peer review be performed by the Animal Use Subcommittee office.
4. Purchases of animals other than through this system must be cleared through the ACVS office. Health certificates will be required.

REQUIREMENTS/COMMENTS
Please ensure that individual(s) performing procedures on live animals, as described in this protocol, are familiar with the contents of this document.
The holder of this *Animal Use Protocol* is responsible to ensure that all associated safety components (biosafety, radiation safety, general laboratory safety) comply with institutional safety standards and have received all necessary approvals. Please consult directly with your institutional safety officers.

c.c. B. Soper

***The* University *of* Western Ontario**
Animal Use Subcommittee / University Council on Animal Care
Health Sciences Centre, ? London, Ontario ? CANADA – N6A 5C1
PH: 519-661-2111 ext. 86770 ? FL 519 -661-2028 ? www.uwo.ca / animal