ANTERIOR CINGULATE CORTEX CELLS IDENTIFY ERRORS OF ATTENTIONAL CONTROL PRIOR TO PREFRONTAL DISENGAGEMENT

Chen Shen

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

Thesis Defense Date: January 2014

Abstract

The anterior cingulate cortex (ACC) is implicated in the detection of errors and the allocation of correctional adjustments. However, error detection alone is not sufficient to resolve and prevent future mistakes since errors can occur in various ways, subsequently requiring different adjustments. I therefore investigated whether the ACC tracks specific processing states that give rise to errors in order to identify which specific processing aspects need readjustment. To do this, my lab recorded from cells in the prefrontal cortex (PFC) of macaques while they were performing a selective-attention task that elicited three types of error. My study provides support for the functional role of the ACC in performance monitoring and specifying correctional adjustments through the tracking of specific sources of erroneous task outcomes.

Acknowledgements:

My research was funded by the Canadian Institutes of Health Research, and The Natural Sciences and Engineering Research Council of Canada.

I would like to thank all members of the Womelsdorf Lab and the Hoffman lab for their advice and expertise, as well as their support and encouragement over the past two years. In particular, I would like to extend a special thanks to Dr. Thilo Womelsdorf for his guidance and mentorship, to Dr. Daniel Kaping and Dr. Stefan Everling for providing me with a rich set of monkey physiology data, as well as to Dr. Salva Ardid and Dr. Stephanie Westendorff for providing valuable input to my analysis. I would also like to thank my advisory committee member, Dr. Kari Hoffman, for her advice and guidance.

Table of Contents

List of Figures:

List of Abbreviations:

ACC – Anterior cingulate cortex

dlPFC – dorsolateral prefrontal cortex

ERN – Event-related potential

ERP – Error-related negativity

mPFC – medial prefrontal cortex

PFC – Prefrontal cortex

lPFC – Lateral prefrontal cortex

Chapter 1 – Introduction and Literature Review

1.1 - Introduction

Envision a situation where a western tourist attempts to order iced coffee from a vending machine in Japan. A tentative process of pressing one button that he believes to produce iced coffee instead delivers iced tea. During the next few days, he presses more wrong buttons, but ultimately succeeds in learning the right symbol that servers as a reliable predictor of iced coffee. The proposed hypothetical situation highlights the importance of errors, or failures to meet one's prior expectation, in the process of acquiring reliable predictions of one's actions to achieve the desirable outcome.

Error typically carries a negative connotation due to its potentially severe consequences; yet it plays a vital role in our daily lives by signaling that a given behaviour is no longer appropriate, and that proper behavioural adjustments may be needed. According to the roman philosopher Seneca, "To err is human, to persist is devilish". Indeed, failure to learn from previous harmful behaviours is one of the characteristic causes underlying the symptoms demonstrated by patients suffering from mental illnesses such as schizophrenia (Franken et al. 2007; Becerril et al. 2010). Taken together, it is not surprising that recent literature on learning has emphasized the importance of error monitoring (Shenhav et al. 2013; Khamassi et al. 2013; Alexander & Brown 2011). The aim of my master's thesis is to examine how errors are detected by the brain and how it might be utilized in learning. The

following chapters will first introduce two models of error-facilitated learning and discuss how these models are implemented by the brain. Then I'll present my research examining whether our brain can not only detect errors, but also discriminate different types of errors during an attentional-selection task.

1.2 – Models of error-facilitated learning

Error and learning

According to associative learning theory, learning consist of associating external stimuli or behavioural responses with their subsequent outcomes. For instance, the sound of thunder is an external stimuli associated with the outcome of oncoming rain (stimulus-outcome association). An example of an association between behavioural responses and outcome (response-outcome association) would be learning that pressing the right button on the drink dispenser gives you a can of coke. Based on an array of these learned associations, the subject can then direct his behaviours in order to maximize favourable outcomes while avoiding unfavourable outcomes.

In order to establish either stimulus-outcome or response-outcome associations, the following two conditions must be met. First, the external stimulus or behavioural response must be closely followed in time by their subsequent outcome. Second, the associated outcome must be surprising or unpredictable for

the prior stimulus or response to be established as its predictor (Schultz & Dickinson 2000).

The first condition was proposed based on classical and instrumental conditioning. In classical conditioning, a tone can form association with fear if fear is predictively followed in time. Similarly for instrumental conditioning, the organism can learn to perform a particular action in order to obtain the desirable outcome if the outcome is predictively preceded by the action. However, this simple pairing of stimulus and outcome alone is not sufficient for explaining the blocking effect during learning (**Figure 1**), thus a second condition was added (Martin & Levey 1991; Kamin 1969). To illustrate the blocking effect, imagine two cues A and B that are equally predictive of the outcome X. If A and B are both presented to a subject, then by our first criteria of learning, A and B will both be associated with outcome X. But suppose that the subject learned previously that A by itself is predictive of the outcome, then the ability of B to predict the outcome will not be learned. The blocking effect demonstrated that associative learning is not based solely on the sequential occurrence of cue and outcome, but the outcome must also be surprising or unpredictable for the stimulus to be established as a predictor (condition 2). In this way, associated learning can discriminate against redundant and relatively uninformative stimuli or responses.

Taking both conditions into account, two prominent theories have been proposed to explain learning: the reinforcement learning model which assumes that learning is directly driven by prediction errors, and an attentional model which argues that errors indirectly affect learning by allocating attention to the predictive stimuli or responses. In the following sections, I'll attempt to explain these two theories in more detail.

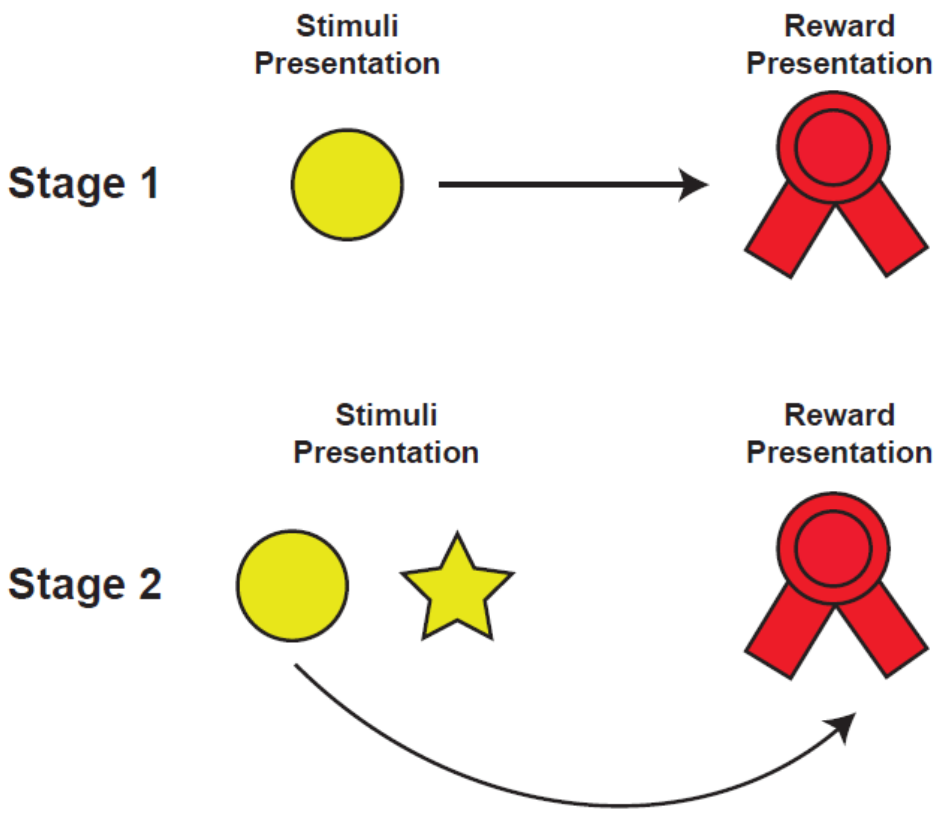

Figure 1. Blocking experiment demonstrated that learning is sensitive to the unexpectedness of the reinforcer. In stage 1, the circle was repeatedly paired with the reinforcer so that the circle became established as a predictor of the reinforcer. In stage 2, the pre-trained circle was presented in compound with the star and paired with the same reinforcer. According to simple associative pairing rule, the circle and the star should be established as equivalent predictors of the reinforcer. However, because the reinforcer was already fully predicted by the circle and the star supplied no additional information, the star was unable to be

established as a predictor of the reinforcer (Schematic adapted from Scheultz & Dickinson 2000).

Direct learning with prediction errors

In the reinforcement learning model, an organism must first learn to correctly predict the reward values (outcome) associated with the occurrence of a particular stimulus or action through trial-and-error, then use the array of learned associations to maximize the values gained and minimize the values lost. The underlying assumption here is that learning is directly driven by prediction errors or the degree of mismatch between the predicted reward values and the actual reward values associated with the stimulus or action (Rescorla & Wagner 1972; Sutton & Barto 1998; Rushworth & Behrens 2008).

In this model, the expected value associated with a stimulus or action is constantly updated by the equation:

$$V_{t+1} = V_t + \sigma \cdot \alpha$$

Where V_t is the expected reward value on the current trial, sigma is the prediction error, or the discrepancy between the expected reward value and the actual reward value, and alpha determines the learning rate, or the extent to which the prediction error influences the expected reward value on the upcoming trial, represented by

the symbol V_{t+1}. Here, the prediction error plays a direct role in minimizing the deviation in actual and expected reward value: the prediction error is positive when the outcome is better than expected, and negative when the outcome is worse than expected, and the magnitude of the prediction error signifies the degree of discrepancy.

Under this model, once the subject accurately predicted the values of reward received using cue A, there would be no prediction error. Thus, no learning would occur during the blocking experiment when cue B was presented.

Indirect learning via attention induced by error

A competitive theory first proposed by Mackintosh (1975) argues that errors do not have direct effect on learning, but rather they indirectly modulate learning by allocating attention. The underlying assumption is that the subject will assign greater attention to the relevant stimulus or action following an error. The more attention that is allocated to a given stimulus or action, the more readily that stimulus or action can be associated with its outcome (Mackintosh 1975; Pearce & Hall 1980; Kaye & Pearce 1984).

In this model, the degree of allocated attention (alpha) is directly proportional to the absolute degree of discrepancy between the predicted (V_t) and

the actual outcome (V_{t+1}). The allocated attention would then serves to facilitate learning:

$$\alpha \propto |V_{t+1} - V_t| \propto Learning$$

In the blocking experiment, the occurrence of cue B allocated little attention since the outcome could be accurately predicted from the occurrence of cue A. Thus, the ability to form association between cue B and the outcome was weak due to the lack of attention allocated to cue B.

Although the two models differ greatly on the role of errors in learning, empirical studies have shown that both appear to be involved in learning processes (Shenhav et al. 2013; Holroyd & Yeung 2012). The following sections will cover the past and recent biological evidence supporting both of these models of learning.

1.3 – Biological signatures of error

Despite their differences, both reinforcement learning and attentional learning take into account the important role of error. In attempt to link error-driven learning with biological systems, Konorski (1948) and Hebb (1949) proposed the idea that associations are encoded by the strength of synaptic connections between neurons encoding predictive stimuli or actions and neurons encoding the expected outcomes. The weight of their synaptic connections are thought to be modified by

prediction errors signaled by dopamine and norepinephrine neurons (Rumelhart et al. 1986).

Dopamine neurons and reinforcement learning

Midbrain dopamine neurons have been shown to encode for prediction errors by phasically increasing or inhibiting their firing rates following unexpected reward delivery (positive prediction error) or omission of rewards (negative prediction error) respectively. As demonstrated by Schultz et al (1998), dopamine neurons showed phasic increase in firing rate following unpredicted delivery of a juice reward. With continuous pairing of juice reward and an audio tone, phasic activation following the delivery of juice reward decreased as the reward delivery became predictable by the audio tone. However, when the predicted reward failed to occur following the tone after training, dopamine neurons were phasically depressed at the time when the reward was expected **(Figure 2**).

The observation that positive-prediction errors induced opposite modulation in firing rates compared to negative-prediction errors suggested that errors are directly implemented in adjusting reward expectations so that future outcomes can be accurately predicted. The empirical study of dopamine neurons provided strong support for reinforcement learning. The biological support for attentional learning, however, came from studies of norepinephrine neurons.

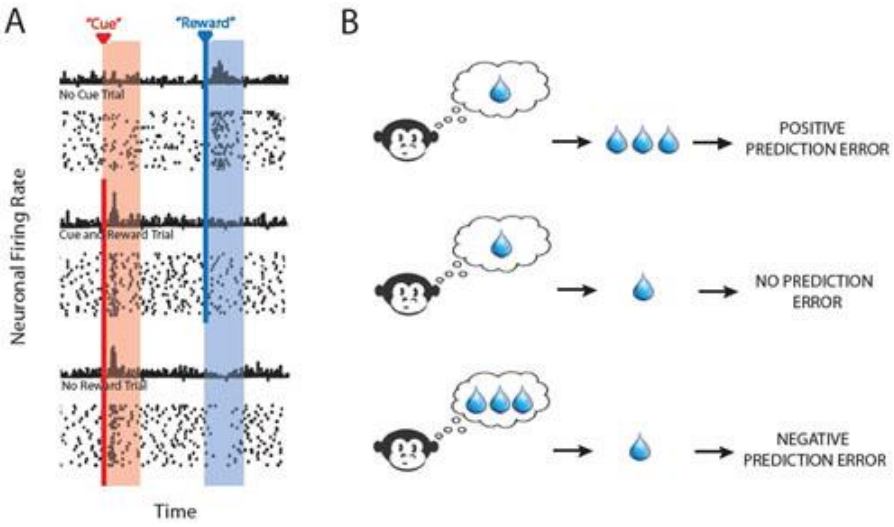

Figure 2. Dopamine neuron and prediction errors. (A) Firing rate of a dopamine neuron signaled prediction errors. In the top panel, an unexpected reward was signaled by increased firing rate following reward delivery (positive prediction error). In the middle panel, the cue was established as a predictor of the reward, and the firing rate was not modulated following reward delivery (no prediction error). In the bottom panel, the expected reward following cue presentation was omitted and triggered a decrease in firing rate following reward omission (negative prediction error). (B) Diagrammatic representation of the corresponding prediction errors shown in A (Scheme adapted from Schultz et al. 1997).

Norepinephrine neurons and attention

Norepinephrine (NE) neurons are located within the brainstem and have wide-spread projections to cortical regions, including the hippocampus and the neocortex. Like dopaminergic neurons, NE neurons exhibit phasic increase in activity following unexpected outcomes, and habituate quickly upon repeated exposures (Aston-Jones et al 1994). However, unlike dopaminergic neurons, NE neurons phasically increase their activities following both positive and negative prediction errors (Sara & Segal 1991; Foote et al. 1980).

The activation of NE neurons following prediction errors has been demonstrated to play a critical role in learning (Devauges & Sara, 1990; Harley 2004). Kety (1970) was the first to propose that phasic NE activation might serve as a learning signal by facilitating synaptic strength between two neurons that fire in conjunction with the release of NE neurotransmitters. Therefore, the phasic release of NE neurotransmitters following prediction errors may enhance ongoing learning by facilitating associations. This theory is supported by the finding that formation of memory traces in the hippocampus were highly sensitive to NE release (Berridge & Waterhouse 2003). Studies have also demonstrated that learning can be facilitated by pharmacologically boosting NE release, and impaired when NE receptors were inactivated (Devauges & Sara 1990; Collier et al. 2004).

Altogether, empirical studies of NE neurons provided biological support for the attentional model of learning. Phasic release of NE neurotransmitters following errors increased the responsiveness of the learning system by facilitated synaptic plasticity, and supported the indirect role of prediction error in facilitating learning.

1.4 – Neuronal substrates of learning and action selection

Anatomy of the Prefrontal cortex

The prefrontal cortex is a density heterogonous region located at the frontal lobe of the brain, and can be further sub-divided into several regions. In particular, the brain regions implicated in learning and implanting behavioural responses have been broadly credited to the medial and lateral sub-regions of the prefrontal cortex (Miller & Cohen 2001).

The lateral prefrontal cortex (lPFC) is composed of areas with distinct cytoarchitectonic properties (area 8, area 9, and area 46) and is similarly organized in both humans and monkeys. However, in humans, the lPFC is folded with three sulci (superior frontal sulcus, frontal sulcus, inferior frontal sulcus), whereas only one sulcus (principle sulcus) is found in monkeys (Petrides & Pandya 2004). The medial prefrontal cortex can be separated into several regions. Barba and Zikopoulos (2007) classified area 24 as the anterior cingulate cortex (ACC) while designating the adjust area 32 as a part of the medial frontal cortex. Similar to the

lPFC, the medical prefrontal cortex is also similarly organized in both humans and monkeys, with an additional sulcus sometimes found within the human ACC. However this sulcus, termed paracingulate sulcus, is presented in only 30 to 60% of the human population (Petrides & Pandya 2004; **Figure 3**).

Complex cognitive learning studies have segregated the function of the lPFC and ACC. The functional segregation of these two regions will be discussed in more detail in the sections below. Briefly, the ACC has been shown to play a critical role in tracking performance failures and signaling a need to make adjustments. When an error is detected, the ACC signals to the lPFC which exerts control over the cognitive processes in order to optimize task performance (Miller & Cohen 2001).

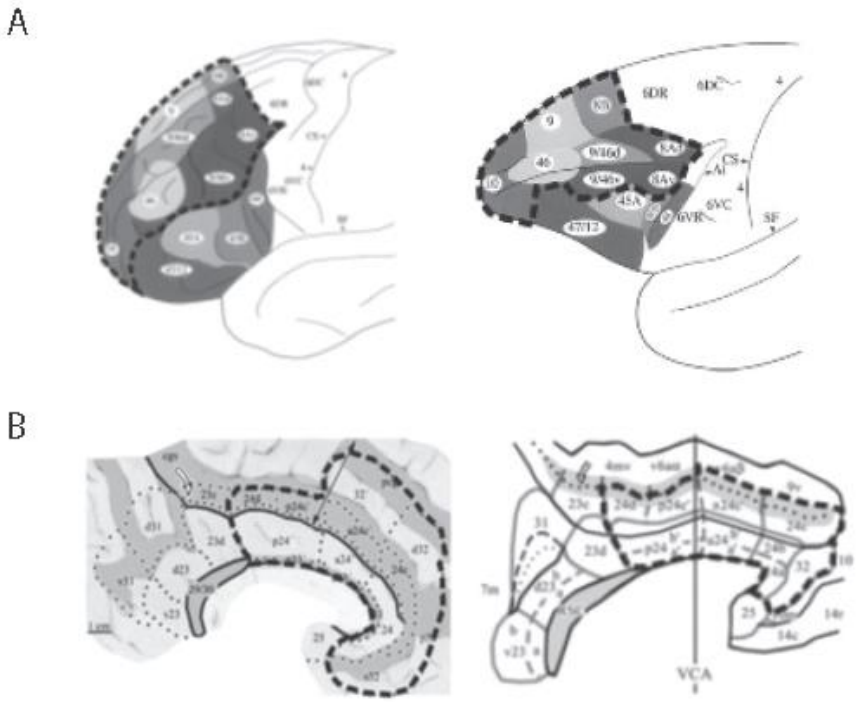

Figure 3. Anatomy of the medial and lateral prefrontal cortex. A, lateral view of the prefrontal cortex in humans (left) and in monkeys (right). B, medial view of the prefrontal cortex in humans (left) and monkeys (right). The lateral prefrontal cortex and the anterior cingulate cortex has long been suggested to play a criticle role in learning. The lateral prefrontal cortex usually refers to boardman area 8, 9 and 46. The anterior cingulate cortex generally refers to boardman area 24. (Diagrams obtained from Petrides & Pandya 2004)

ACC - Performance monitoring and allocating adjustments

The ACC is the most prominent brain region associated with error processing. The earliest description of error signals came from electroencephalograph studies that showed a negative deflection in event-related potential (ERP) peaking around 100ms after the onset of an error. This ERP component, referred to as error-related negativity (ERN), was later localized to the ACC (Falkenstein et al., 1991, 2000; Miltner et al., 1997; Gehring et al. 1993; **Figure 4A**). ERN has been shown to reflect a genuine error-locked response, since it is only observed following sub-optimal performances and can be elicited by a wide variety of sensory modalities (Holroyd & Coles, 2002; Miltner et al., 1997). More recently, single neuronal recordings have also confirmed the existence of error-encoding neurons within the ACC. Neurons within the ACC have been reported to encode for prediction errors, with positive and negative prediction errors encoded by distinct populations (Matsumoto et al. 2007; **Figure 4B**). Other recent electrophysiological studies have reported neurons within the ACC that encoded for inappropriate motor responses such as break fixations, and incorrect response mappings such as incorrect choice errors (Quilodran et al., 2008; Ito et al., 2003).

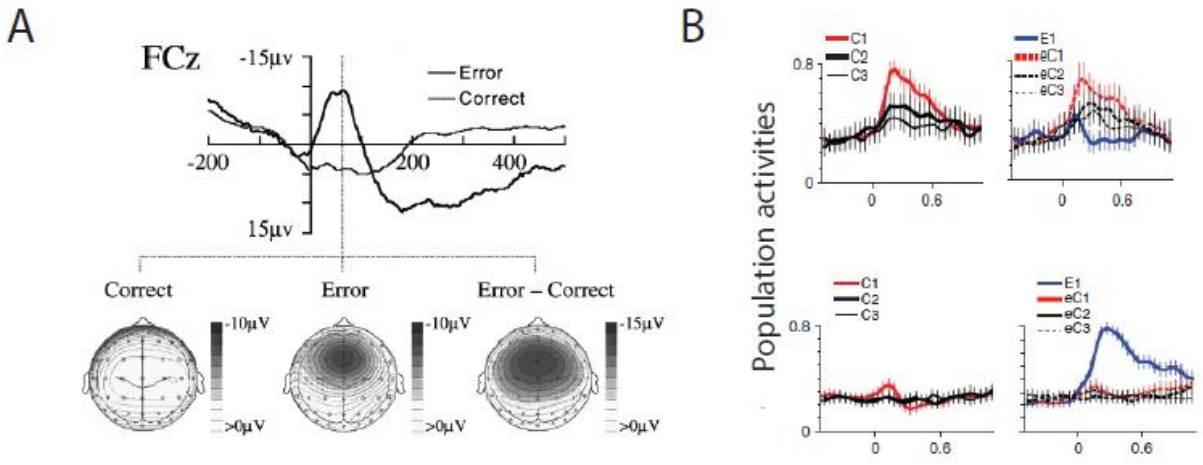

Figure 4. Evidence of error processing within the ACC. (A) Error-related negativity. Top panel shows response-locked waveform for correct and error trials at mid-frontal electros. Bottom panel shows the scalp voltage maps at the time of peak error-related negativity (56ms after the response). (B) Activity of anterior cingulate prediction-error encoding cells aligned to the time of response. Top panel shows the population activity of positive feedback cells, where C1, C2 and C3 respectively indicate first correct responses, and second and third subsequent correct responses. Bottom panel shows the population activity of negative feedback cells, where E1, eC1, eC2, and eC3 respectively indicate error responses, and first second third subsequent correct responses following E1 (Diagrams adapted from Yeung et al. 2004 and Matsumoto et al. 2007).

Errors are salient indicators of the need for adjustments, and the error monitoring role of the ACC has been viewed as a small part of its overall function in optimizing performance by overcoming a bad streak of failures (Shenhav. et al. 2013, Khamassi et al. 2013). Under the reinforcement learning framework, the ACC plays a direct role in updating expectations through prediction errors and allocating correctional adjustments. Evidence in support of this theory came from studies demonstrating ACC's sensitivity to not only prediction errors, but also to the degree of expected rewards and to internally selected actions (Khamassi et al. 2013; Lapish et al. 2008; Hyman et al. 2010, 2011). On the other hand, attentional theory proposes that the ACC indirectly optimizes task performance by adjusting motivational control or the overall responsiveness of the system. In support of this, ACC lesion studies showed changes in behavioral responses that were consistent with the idea of motivational deficits. Lesions of the ACC in rats shifted their selection criteria such that the lesioned rats would often choose the less effortful actions that yield less reward over the more effortful action that yield larger rewards (Walton et al. 2003). In humans, patients with ACC lesions showed signs of a deficit in self-initiated, or willed, movements (Nemeth et al. 1988).

lPFC - Implementing adjustments

In contrast to the monitoring and regulatory role of the ACC, the lPFC is thought to be involved in implementing and executing purposeful behaviors. In particular, the ventrolateral prefrontal cortex is thought to be involved in first-order executive processes, such as the active representation of stimulus-response associations and the selection of task-relevant versus the task-irrelevant information. Whereas the dorsolateral prefrontal cortex is thought to be involved in higher-order executive processes, such as strategic planning, which involves the integration of multiple pieces of information (Petrides 1996).

Given the close functional relationship between the ACC and lPFC, performance errors often elicit the co-activation of both regions. A study by Cavanagh et al (2009) demonstrated that theta band phase coherence reflected communication between the lateral and the medial prefrontal cortex, and that resolutions following errors occur when these two regions interact with each other. In their study, the medial prefrontal cortex responded to error events with increased inter-areal theta power and phase coherence, consistent with the view that the ACC monitors performance. Although the lPFC did not show an increase in inter-areal theta power, theta phase synchrony between the medial and the lateral prefrontal cortex was increased on error trials and this synchrony is predictive of post-error behavioral adjustment (**Figure 5**).

These highlighted studies demonstrated the division of labor between ACC and lPFC: the ACC is involved in performance monitoring, while the lPFC is involved in implementing correctional adjustments to ensure future successes. In order to optimize performance during cognitively demanding tasks, the ACC monitors task performance and signals to the lPFC a need for adjustments.

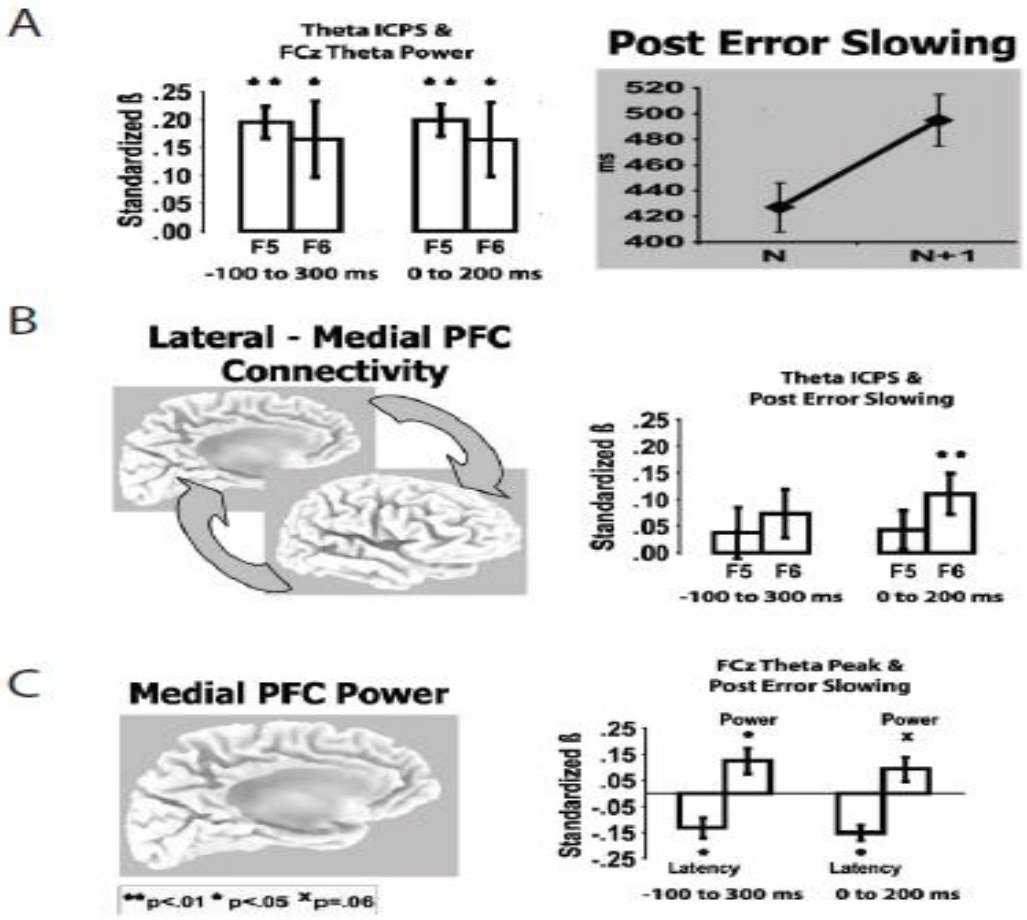

Figure 5. Intra-areal communications between the lateral and medal prefrontal cortex facilitated adjustments following error. Left panel A, medial prefrontal theta power is positively correlated with the power of intra-areal theta synchronization between lateral and medial prefrontal cortex. F5 and F6 represents electros placed on the lateral prefrontal cortex of the left and right hemisphere. Right panel A, post-error response slowing is present in trials following error, where N is error trials and N+1 is the trials following error. Panel B, intra-areal

theta phase synchronization between lateral prefrontal cortex in the right hemisphere and medial prefrontal cortex is correlated with post-error slowing. Similar in C, the inter-areal medial prefrontal cortex theta power after error response is also shown to be correlated with post-error slowing (Diagrams adapted from Cavanagh et al. 2009).

1.5 – Research Questions & Hypothesis

In order to achieve the desired outcome, one must monitor his performance and strategize his behaviors based on learned associations between the external stimuli or behavioral responses and their subsequent outcomes. In the brain, the ACC has been suggested to play a role in performance monitoring and regulating the implementation of executive functions by the lPFC. Metaphorically, the ACC can be viewed as the coach while the lPFC can be viewed as the players. As a coach, the ACC is on the field with the players, experiencing the joy and pain of wins and losses. It monitors the current environment as well as keeping track of the recent performance, and guides the players to overcome a bad streak in performance or to maintain a string of success.

In order for the ACC to function effectively as a coach, it must be able to discriminate performance failures based on their underlying causations in order to make the necessary corrections through the lPFC. Past studies have shown that

ACC neurons encode for different types of errors related to reward expectation, such as positive and negative prediction errors, uncertainty errors, and incorrect choice errors (Matsumoto et al. 2007; Quilodran et al. 2008; Ito et al. 2003; Amiez et al. 2005). However, few studies to date have examined whether the ACC can discriminate errors caused by the failure of different processes of attentional controls. In support of the current theories on the role of ACC and lPFC in decision making, I hypothesize that the ACC can distinguish between different errors types caused by failures of different processes of attentional controls. The rest of the chapters will outline the experimental approaches used in our lab to segregate different processes of attentional control, and highlight my findings as to whether the ACC can distinguish different attentional control failures.

Chapter 2 – Methods

2.1 – Experimental Procedures and Paradigm.

In my analysis, I used a set of previously collected and processed data from a recent publication by our lab (Kaping et al. 2011). The data were collected from two male macaque monkeys following the guidelines of the Canadian Council of Animal Care on the use of laboratory animals and of Western University's Council on Animal Care. The following experimental procedures and data acquisition protocols have been described in detail in Kaping et al. (2011).

Extracellular recordings.

Extra-cellular recordings commenced with 1-6 tungsten microelectrodes (impedance 1.2-2.2 MΩ, *FHC, Bowdoinham, ME*) through standard recording chambers (19mm inner diameter) implanted over the right hemisphere in both monkeys. The recording chambers allowed access to the anterior aspects of the prefrontal cortex and cingulate sulcus, and to align recordings to the same anterior-to-posterior extent in both animals (*see* Kaping et al. 2011; and below: *Reconstruction of Recording Sites*). Electrodes were lowered in guide tubes with software controlled precision microdrives (*NAN Instruments Ltd., Israel*) on a daily basis, through a recording grid with 1mm inter-hole spacing. Before recordings

began, anatomical 7T MRIs were obtained from both monkey's to allow reconstruction of electrode trajectories and recording sites.

Data acquisition.

Data amplification, filtering, and acquisition was done with a multi-channel processor (Map System, Plexon, Inc.) using unity gain headstages. Spiking activity was obtained following a 100-8000 Hz passband filter, further amplification and digitization at 40kHz sampling rate. During recording, the threshold was adjusted to have always a low proportion of multiunit activity visible against which we could separate single neuron action potentials in a 0.85 to 1.1 ms time window. Sorting and isolation of single unit activity was performed offline with Plexon Offline Sorter (*Plexon Inc., Dallas, TX*), based on principal component analysis of the spike waveforms. The monkeys eye position was tracked continuously with an infra-red system (*ISCAN*, Woburn, US) running on a DOS platform, with eye fixation controlled within a 1.4-2.0 degree radius window.

Visual stimulation.

Stimuli were presented on a 19'' CRT monitor placed 57cm from the monkeys' eyes, running at 1024x768 pixel resolution and 85 Hz refresh rate. Behavioural control and visual stimulation was accomplished with Pentium III PCs

running the open-source software Monkeylogic (http://www.monkeylogic.net/), which has been benchmarked and validated in two previous publications (Asaad and Eskandar 2008a, 2008b). Grating stimuli was used with 'rounded off' edges and were presented at 4.2° eccentricity to the left and right of fixation. Monkeys had to detect a transient and smooth clockwise/counterclockwise rotation of the grating movement (see below). The rotation was adjusted to ensure ≥ 85% of overall correct responses to the grating and ranged between ±13° and ± 19°. The rotation proceeded smoothly from standard direction of motion towards maximum tilt within 60 ms, staying at maximum tilt for 235 ms, and rotating back to the standard direction within 60 ms.

Experimental paradigm

Monkeys performed a selective attention task requiring a two-alternative forced-choice discrimination on the attended stimulus (**Figure 6**). Monkeys initiated a trial by directing their gaze to a centrally presented, gray fixation point. Following a fixed 0.4 sec period, the two black/white grating stimuli were colored red/green ('*Color Cue' onset*). Within 0.05 to 0.75 sec after color onset, the central fixation point changed to red or green cueing the monkeys to covertly shift attention towards the location with the color matching stimulus. I label the period of sustained spatial attention the '*Attention Epoch*' (**Figure 6**, time epoch with red

colored solid line). Errors during the Attention Epoch were fixation breaks and triggered abortion of the trial. At random times (drawn from a flat random distribution) within 0.05-4 sec after cue onset the cued target grating transiently rotated clockwise or counterclockwise. In half of the trials the un-cued distractor grating transiently rotated before the target. The stimulus rotation of the distractor stimulus had to be ignored, or filtered, which lead me to label this trial period the '*Filter Epoch*' that required control of bottom-up stimulus interference (**Figure 6**, time epoch with red colored dashed line). Monkeys had to discriminate the rotation of the target stimulus by making a saccadic eye movement up- or downwards to one of two response targets within 70-550 ms following rotation onset. I label the time after target stimulus rotation that required a choice from the monkeys the '*Choice Epoch*' (**Figure 6**, blue colored time epoch). The monkeys received fluid reward after a further delay of 0.4 sec after correct saccadic responses. Errors during the Filter and Choice Epoch were incorrect saccadic responses to the response target, or saccadic responses elsewhere Stimulus colors were differentially associated with high and low reward in alternating blocks (for more details *see* Kaping et al. 2011).

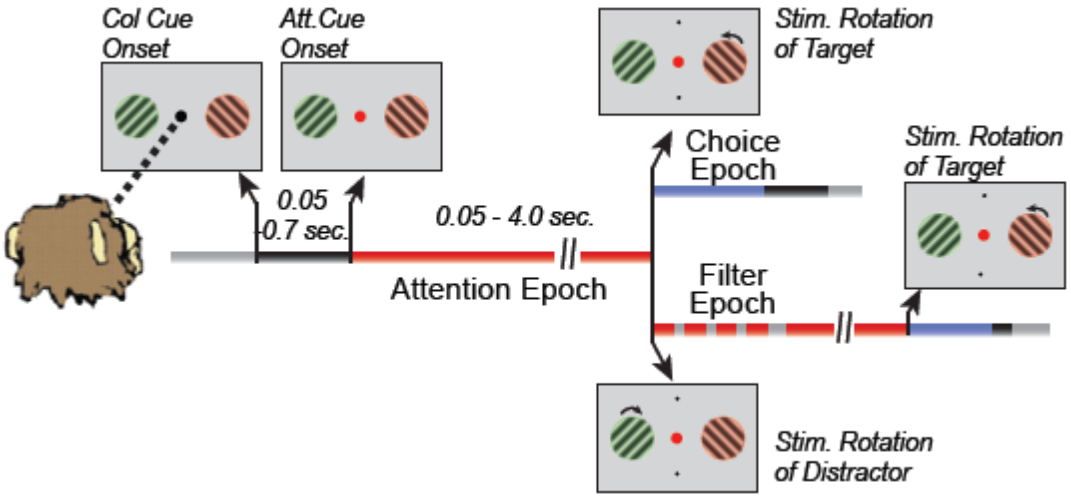

Figure 6. Selective attentional task paradigm. Two macaque monkeys performed a selective attention task (red / green) requiring a two-alternative forced-choice discrimination (clockwise vs. counterclockwise rotation) on the attended stimulus. Trials begin with monkeys fixating a central fixation point bounded by green and red colored grating stimuli. During the attentional period, the color of the central fixation point changed to either red or green cuing the monkey to covertly shift its attention towards the target grating stimulus with color matching that of the fixation cue. Within 0.05-4s following cue onset, either the target or the distractor transiently changed. In 50% of the trials, the distractor change occurred before the target change (distractor period) and had to be ignored. To receive a liquid reward, the monkeys had to discriminate the rotation of the cued stimulus by making an

upward/downward saccade to clockwise/counter clockwise rotations (response period).

2.2 – Reconstruction of Recording Sites.

I utilized a set of data that had been previously processed and prepared for a recent publication from my lab (Kaping et al. 2011). Briefly, the anatomical sites of each recorded neuron were reconstructed and projected onto a flat map representation of the macaque brain (*see* Kaping et al. 2011) (**Figure 7**). My lab used the area subdivision scheme outlined by Barbas and Zikopoulus (2007) and refer to prefrontal areas 46, 8, and 9 as 'lateral prefrontal cortex' (lPFC), area 24 as the anterior cingulate cortex (ACC), and area 32 as the (ventro-) medial prefrontal cortex (mPFC) (see also Passingham and Wise, 2012). Very similar area assignments would follow when considering two other major anatomical subdivision schemes (Saleem et al. 2008, Petrides and Pandya 2007, *see* Suppl. Fig. 2 in Kaping et al. 2011). To briefly summarize the reconstruction steps, my lab began by projecting each electrodes trajectory onto the two dimensional brain slice obtained from 7T anatomical MRI images, using the open-source OsiriX Imaging software and custom-written Matlab programs (*Mathworks Inc.*), and utilizing the iodine visualized electrode trajectory within the electrode grid placed within the recording chamber during MR scanning. They drew the coronal outline of the cortical folding of the MR gray scale image to ease the comparison of the

individuals monkey brain slices to standard anatomical atlases, and to ease using major landmarks (Van Essen et al. 2001). They then projected, manually and under visual guidance, the electrode position to the matched location in the standard F99 brain in Caret. Our lab estimated that the complete procedure from documenting precisely the recording depth, identification of the recording location in the monkeys MR slice, and up to the placement of the electrode position in the standard F99 brain introduces a potential maximal error of 3mm.

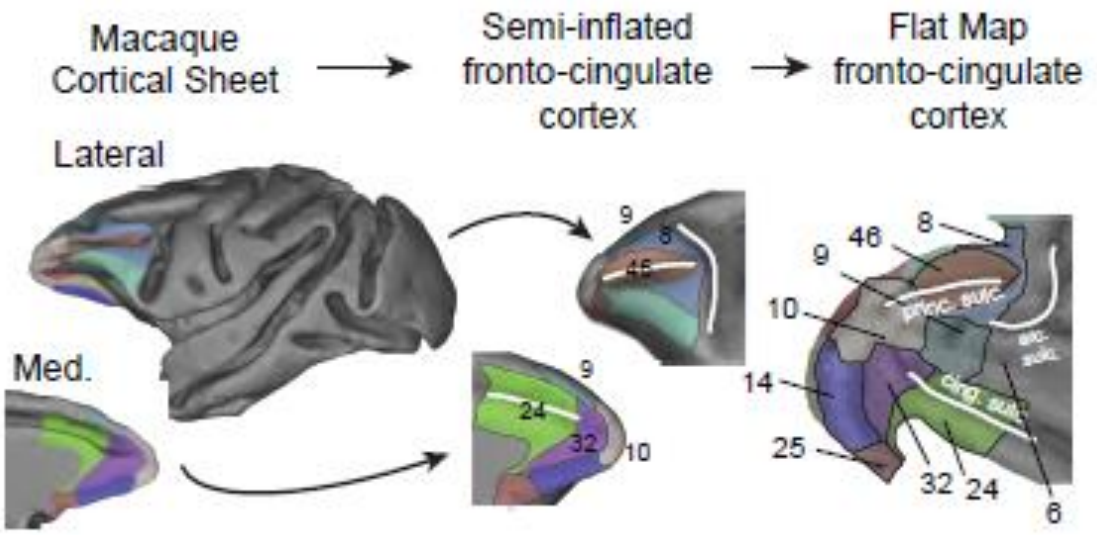

Figure 7. Anatomical subdivisions of the prefrontal cortex. A, lateral and medial view of the macaque brain. Left panel shows a typical 3D model of the macaque frontal cortex, with subdivisions colored according to Barbas and Zikopoulus (2007). The middle panel shows a partially inflated brain, and the right panel shows the flattened representation of the brain.

However, we felt that despite this potential distortion, which we cannot rule out despite our confidence that the typical (unsystematic) error is more in the 1mm range, the assignment of recording locations to standard brains is highly beneficial. Note in particular, that anatomical reconstruction was conducted entirely independent of the analysis of neuronal data and of projecting functional results onto the anatomical 2D map. After identifying all recording sites within the standard F99 brain, my lab used the Caret software package to render the standard brain into a three dimensional volume, which was then spherically inflated and cut in order to unfold the brain into two dimensional space, i.e. the flat map.

2.3 – Data Analysis.

Analysis was performed with custom Matlab code (Mathworks, Natick, MA), utilizing functionality from the open-source fieldtrip toolbox (Oostenveld et

al. 2011). Analysis of spiking activity was based on convolving spiketrains of individual trials with a Gaussian (SD 30ms) sampled every 5 ms.

Classification of errors & saccade directions

Analysis of errors was time aligned to the onset of the erroneous trial outcome. For the analysis, I considered errors that were committed during either of three time epochs. First, I considered errors committed during the sustained Attention Epoch within 0.3-3 sec. following attention cue onset. These errors occurred during a state of selective attention and may best reflect rather unspecific motivational lapses of attentional control, which entailed lost inhibitory control of eye fixation. The second error type occurred during the Filter Epoch. I restricted analysis of these errors to the typical response time window that animals were granted for the true attentional target stimulus. Errors in this Filter Epoch reflected the erroneous bottom-up capture of attention by the salient distractor, or erroneous top-down attention to the wrong stimulus that was of a color different to the attention cue color. The third error type was seen in the Choice Epoch, where the animals either responded to the wrong response target (i.e. made a wrong choice), or incorrectly directed attention to the wrong stimulus. Errors during choice period reflected erroneous motor response due to erroneous stimulus-response mapping. I

restricted analysis of errors in the choice period to the time window that would have been allowed for the correct response to occur.

I categorized the saccade direction of errors in each of the three time epochs during the trial relative to the four major quadrants (up: 0° ± 45°, left 90° ± 45°, down: 180° ± 45°, and right 270° ± 45°). For each of the three errors, I statistically tested whether erroneous saccades were biased in any particular direction (Chi square test). First, I compared whether saccadic responses up versus down (and right vs left) were statistically equally likely. Since the response targets were along the vertical axis and the peripheral stimuli in the horizontal dimension, I then tested the null hypothesis that saccade directions were equally distributed in the horizontal vs. vertical axis.

Identifying error specific firing rate modulation.

I imposed two main criteria to identify whether cells encoded error specific information in their firing. First, firing in response to the error had to be significantly different from pre-error baseline firing. Baseline was defined as the average firing within 300 ms before error onset. Error activity was averaged in sliding windows of ±150 ms in steps of 50 ms. Secondly, the activity modulation following errors had to be significantly different from the firing in correct trials. For this comparison, I randomly realigned correct trials so that their activity at time

"zero" corresponded to the time when errors were committed (i.e. drawing the alignment times for the times of error commission).

For both criteria, I required a cell to show significant modulation for at least five consecutive sliding windows in the time windows between 0 and 0.7 sec. after error onset ($p < 0.05$, Mann-Whitney U test). To focus on the transient detection processes I additionally required that there was no significant modulation prior to error-onset, and after 0.7 sec. following error onset. For some cells with low and no firing in a large subset of trials, I used a non-parametric two-part model for statistical comparison that has been shown to have more power and be more accurate for data with large proportions of discrete (0's) over continuous values (Lachenbruch 2002).

For all neuronal analysis I used only errors committed ≥0.2 sec. following the onset of cue (Attention Epoch errors), the onset of distractor rotation (Filter Epoch errors) and the onset of target rotation (Choice Epoch errors). This selection prevented interference from transient onset related effects. I further required at least ten error trials per isolated cell to test for error-selectivity and proceed with subsequent analysis. Error modulated firing was either characterized as enhanced or suppressed firing, based on the sign of the firing rate difference between baseline corrected correct and error trials.

Anatomical distribution of cells with error-selective firing.

To test for fine anatomical clustering of error-detecting cells, I calculated the proportion of error-selective cells, relative to all recorded cells, at intersections of a virtual grid within a 4mm radius of the flat map. I only considered pixels in the map that included at least ten recorded cells. This anatomical mapping of error activity allowed testing for spatial clustering of error detection using randomization statistics that controlled for uneven sampling of cells across the grid (*see* Kaping et al. 2011). To test for a higher or lower proportion of error-detecting cells than expected by chance, I calculated n=2000 random distributions of proportions of cells with a significant effect, after randomly shuffling the location label while maintaining the number of recorded cells per pixel. A cluster was significant if the observed proportion of error-selective cells exceeded the 95^{th} percentile of the random distributions, corresponding to one-tailed $p < 0.05$ threshold.

Independent of the fine anatomical mapping, I also compared the proportions of error detecting cells between three major subdivision of the PFC using multiple comparison corrected (Bonferroni threshold: $p < 0.0165$) chi-square tests of independence, comparing ACC vs. lPFC, lPFC vs. mPFC and mPFC vs. ACC.

Latency analysis of error-selective firing.

For each cell with significant error-selective firing, I calculated the latency as the time of maximal difference in post-error to pre-error firing baseline. I aimed then to test for significant discrepancies in the latency associated with different response types (enhanced vs. suppressed firing), different area subdivisions (lateral PFC, ACC, mPFC), and different error epochs (Attention-, Filter-, and Choice-Epoch). However, the small sample size of some cell subsets limited the reliability of the latency distribution (i.e. proportion of error-selective cells with a given latency). Therefore, instead of simply comparing medians from unreliably characterized distributions, I computed the cumulative latency distribution for each cell subset, fitted them with sigmoidal functions, and utilized the C_{50} parameter of the fits – latency at which the cumulative distribution reaches 50% – as an improved estimate of the overall cell subset latency. Prior to fitting, response latencies were tested for unimodality using Hartigan's Dip Test (Hartigan and Hartigan 1985). Those that were confirmed to be multimodal were fitted with two separate sigmoidal functions. Confidence intervals surrounding the fit were computed from the Jacobian matrix and the residuals using Matlab toolbox.

Analysis of error-selective firing: enhancement versus inhibition.

For ACC and lPFC, I directly tested whether the proportion of cells with error-locked firing enhancement exceeded, or fell behind, the proportion of cells with error-locked firing inhibition across time relative to error onset. To do so, I first estimated the distribution of cells showing error-specific enhancement/suppression by differentiating the fitted cumulative distribution of error latencies (*see* above). The subtraction of both derivatives indexed the relative balance of enhanced and inhibited firing. I evaluated statistical significance using randomization statistics by shuffling the assignment of cells to the enhanced/suppressed cell subset and repeating the procedure 2000 times.

Classification of putative cell types.

Putative cell types were classified using a method developed by one of our post-doctoral colleagues. Following his guidance, I aligned, normalized and averaged all action potentials for the set of highly isolated neurons (n = 404) of the sample. Each neuronal waveform was then fitted with cubic interpolation from an original precision of 25 μs to 2.5 μs. On the resultant waveform, I analyzed two measures (**Figure 8**): the peak-to-trough duration and the time for repolarization. The time for repolarization was defined as the time at which the waveform amplitude decayed 25% from its peak value. These two measures were highly

correlated ($r = 0.68$, $p < 0.001$, Pearson correlation). I computed the Principal Component Analysis and used the first component (84.5 % of the total variance), as it allowed for better discrimination between narrow and broad spiking neurons, compared to any of the two measures alone. I used the calibrated version of the Hartigan Dip Test (Hartigan and Hartigan 1985) that increases the sensitivity of the test for unimodality (Cheng and Hall 1998; Henderson et al. 2008). Results from the calibrated Dip Test discarded unimodality for the first PCA component ($p < 0.01$) and for the peak to trough duration ($p < 0.05$) but not for the duration of 25% repolarization ($p > 0.05$). In addition, I applied Akaike's and Bayesian information criteria for the two- vs one- Gaussian model to determine whether using extra parameters in the two-Gaussian model is justified. In both cases, the information criteria decreased (from -669.6 to -808.9 and from -661.7 to -788.9, respectively), confirming that the two-Gaussian model is better. I then used the two-Gaussian model and defined two cutoffs that divided neurons into three groups (**Figure 8**). The first cutoff was defined as the point at which the likelihood to be a narrow spiking cell was 10 times larger than a broad spiking cell. Similarly, the second cutoff was defined as the point at which the likelihood to be a broad spiking cell was 10 times larger than a narrow spiking cell. Thus, 95% of neurons ($n = 384$) were reliably classified: neurons at the left side of the first cutoff were reliably classified as narrow spiking neurons (20%, $n = 79$), neurons at the right side of the second cutoff were reliably classified as broad spiking neurons (75%, $n = 305$). The

remaining neurons were labeled as ‘fuzzy’ neurons as they fell in between the two cutoffs and were not reliably classified (5%, n = 20).

Analysis of cell type distributions

To test whether the ratio of narrow and broad cells were different between the error encoding neurons within the ACC and lPFC, I applied randomization statistics. From the recorded neurons within each region, I randomly sampled a set of neurons equal in size to error population. I repeat this process 2000 times for both lPFC and ACC to obtain the expected narrow and board cell proportions found within the two regions. I then compared the observed difference between narrow and board cells to the expected different between narrow and board cells obtained from randomization. Statistical significances were identified if the observed values fall outside the 95% confidence interval. This method corresponds to a two-tailed test with $p < 0.05$.

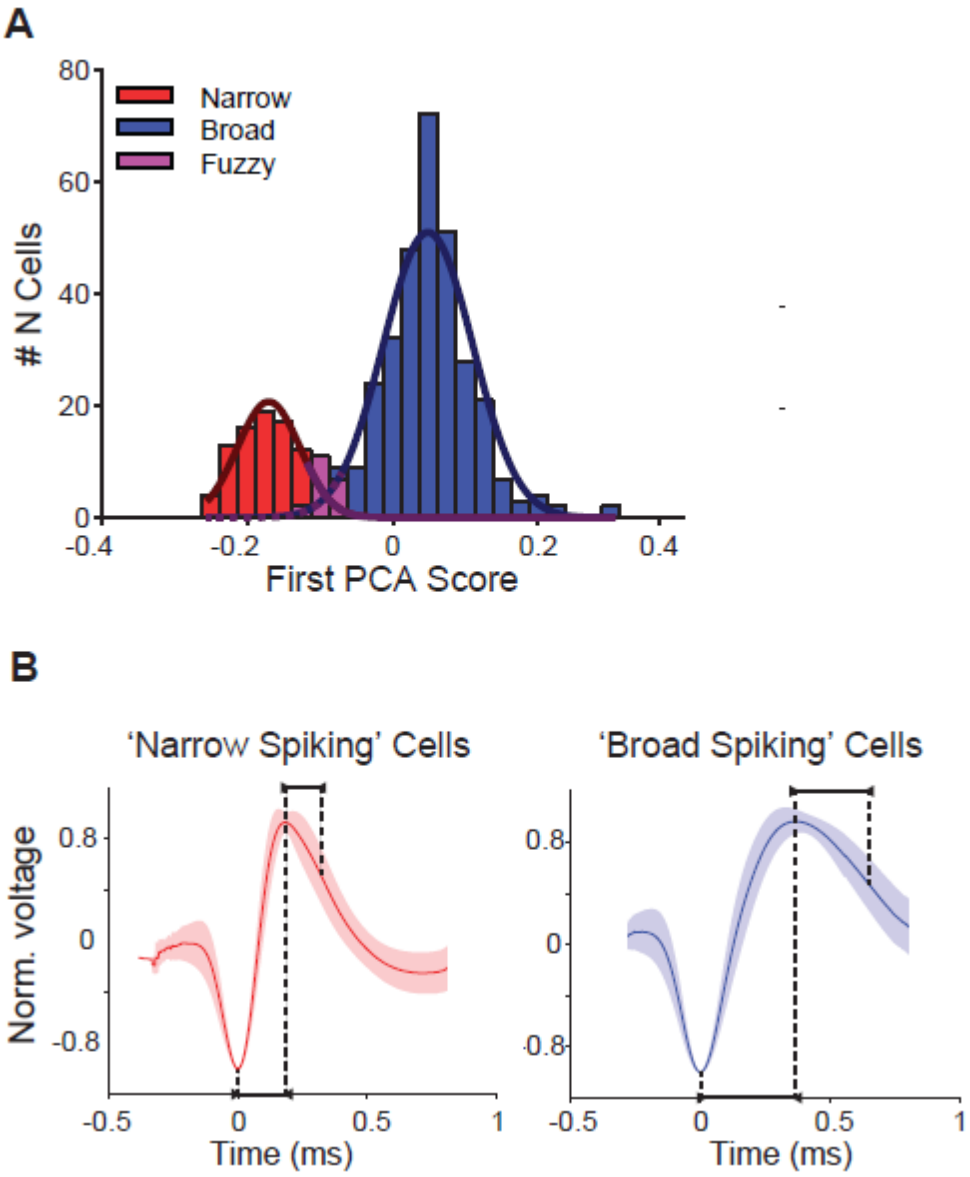

Figure 8. Putative cell types classification. A, cells towards the left and right of the bimodal distribution were classified as 'narrow spiking' and 'board spiking' cells respectively, with those falling in the middle labeled as 'fuzzy' cells. The bimodal histogram distribution was plotted using PCA score calculated from peak-

to-trough time and repolarization time. B, averaged population wave-form for narrow (red) and board spiking cells (blue). The black line represent the mean normalized voltage, and the shaded region represents stand error around the mean.

Analysis of oculomotor activities

To examine whether error-related activities are confounded by other factors, such as motor related modulations, spike activities of error-feedback neurons were re-examined after aligning to saccade onset. For a particular error-feedback neuron, saccadic aligned correct and error trials were tested for significance against their respective baselines as described above. For error neurons that displayed significantly modulated activity from baseline, during both correct and error trial, their respective magnitude of maximum modulation from baseline were compared pair-wise (paired-ranksum test, $p < 0.05$).

Chapter 3 – Results

3.1 – Behavioral Error Classification.

Our lab have recorded single neuronal activities across a large extent of the lateral and medial prefrontal cortex in two macaque monkeys while they were performing a selective attention task (*see* Kaping et al. 2011). The task required the monkeys to use an instructional cue stimulus to covertly shift their attention to one of two peripheral stimuli, and maintain spatial attention on that stimulus until it transiently rotated. The monkeys were then required to make the correct saccadic movements towards one of the two targets based on the direction of the rotation (**Figure 6**). In half of the trials, the non-attended distracting stimulus transiently rotated before the attended stimulus, which had to be filtered (ignored) by the animals. Overall, the task was performed at 78% accuracy levels (monkeys M and R: 76.6 / 83.9%; SD: 10%), with errors falling into three major task epoch: First, errors committed during sustained spatial attention, but prior to any stimulus rotations (*Attention epoch*), were breaks of fixations without a particular strong directional bias to the response targets (46% horizontally) or peripheral stimuli (54% vertically) (**Figure 9**). These errors mostly reflected lapses in inhibitory control of eye fixation and were indicative of lapses in the top-down maintenance of attention. Monkeys committed errors in the Attention Epoch for an average of 8.5% of all trials (monkey M: 6% and R: 11%). The second epoch with large numbers of errors

followed the distractor rotation that had to be filtered from influencing the behavior of the monkeys. Saccadic errors in this *Filter Epoch* partly reflect erroneous top-down attention to the distractor (as inferred from the 78% vertical preference in the saccade directions), but also partly a bottom-up capture of attention by the peripheral stimuli (22% horizontal saccades, **Figure 9**). Monkeys made on average 8% of the trial errors in the Filter Epoch (monkey M: 6% and R: 10%). The third type of error was committed within the allowed response time window to the rotation of the attended target stimulus (**Figure 9**). Errors in this *Choice Epoch* were predominantly saccades to the wrong response target (91% vertically) located vertically up and down from the fixation point (**Figure 9**). Choice Epoch errors were thus mostly reflecting incorrect sensory-response mappings from rotation directions to saccade directions. Monkeys made on average 11.5% of trial errors in the Choice Epoch (monkey M: 18% and R: 5%).

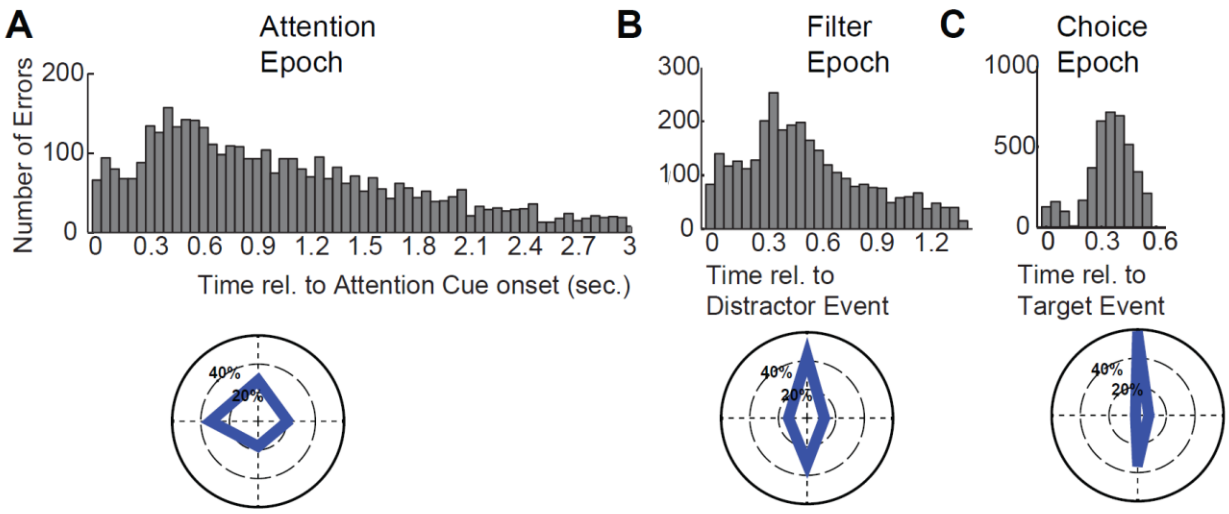

Figure 9. Behavioral characteristics of erroneous saccades within attentional, filter, and choice epochs. A, distribution of error commission times during attentional, filter, and choice epoch. B, proportion of erroneous saccades that were made toward one of the four quadrants (up, down, left, right) during attentional, filter, and choice epochs.

3.2 – Selective error-detection signals in neuronal firing.

Next, I attempted to isolate prefrontal cells encoding for error outcomes in the three task epochs described above. **Figure 10** shows the proportion of cells that transiently modified their firing upon error commission in each of the task epochs: for the Attention Epoch, the recordings of 867 cells included at least 10 error trials, and 128 (15%) showed error selectivity; for the Filter Epoch, the recordings of 544 cells included at least 10 error trials, and 72 (13%) showed error selectivity; and for the Choice Epoch, the recordings of 728 cells included at least 10 error trials, with 84 (12%) showing error selectivity. The error-locked firing modulations were evident not only in increased firing, but also through transient post-error response inhibition for a substantial number of neurons (44% of error-selective cells).

Cells may show error-locked responses only for errors in one task epoch, or they could generalize and signal errors across epochs. I quantified this 'tuning' to specific errors in different epochs by calculating the proportion of cells that showed joint error-locked response modulation (**Figure 11**). The joint selectivity of error outcomes across epochs was evident in less than one third of the cells (28% attentional epoch, 33% filter epoch, 24% choice epoch) that were error-selective through increased activations (**Figure 11 A**). Joint-error coding through response inhibition was less equally distributed compared to response enhancement (22% attentional epoch, 28% filter epoch, 9% choice epoch), with proportion of joint-error encoding being significantly less in choice epoch than in attentional epoch

(**Figure 11 A**, Chi squared test, $p < 0.05$). **Figure 11 B,C** illustrates the specific combinations of task epochs to which single cells showed error locked firing increases (**Figure 11 B**) and decreases (**Fig. 11 C**). Joint-error selectivity through response enhancement was similarly likely for all combinations of error types. In contrast, cells showing response inhibition following errors in the Choice Epochs were largely unaffected by errors committed during the Attention Epoch and Filter Epoch (**Figure 11C**).

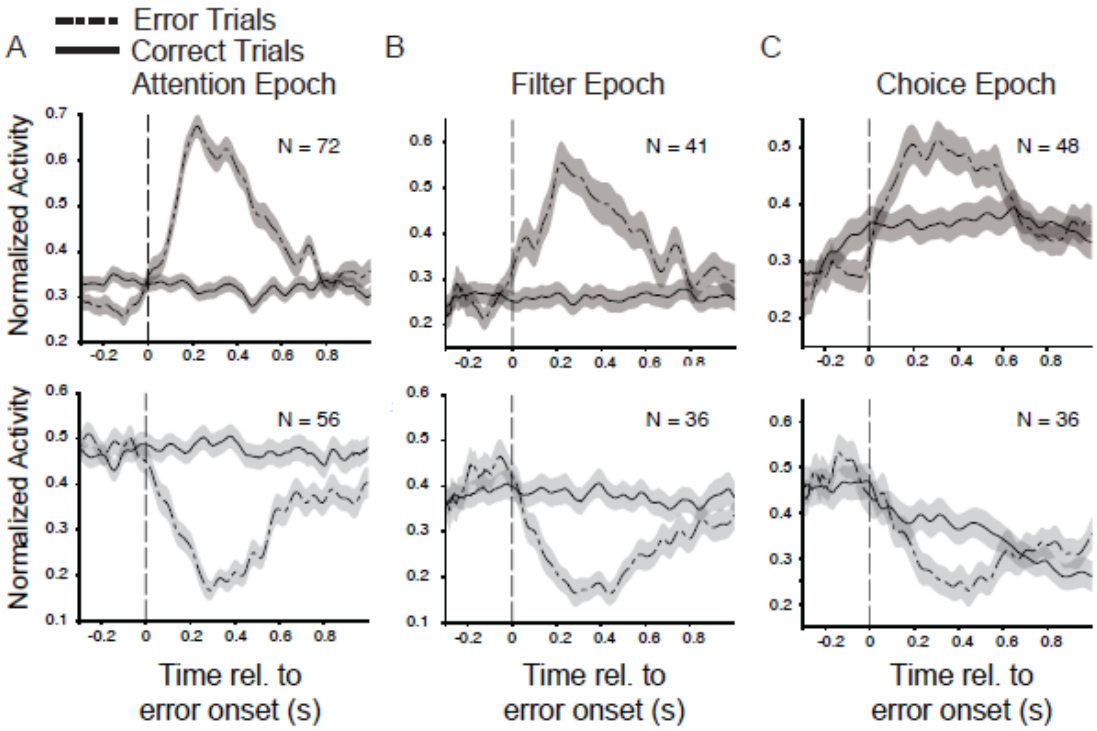

Figure 10. Error-related firing rate modulations for errors in three different task epochs. (A) Normalized firing of cells that increase (upper panels) and decrease (lower panel) their firing rate transiently following error commission in the Attention Epoch (see Fig. 1A). Dashed and Solid lines correspond to error and correct trials. (B, C) Same format as in A, but for the set of cells with significant activity modulation in response to error commissions in the Filter Epoch (B), and the Choice Epoch (C). Normalization of single cell firing rates used the formula (Rate - min(Rate))/(max(Rate)-min(Rate)). Gray shading shows SE.

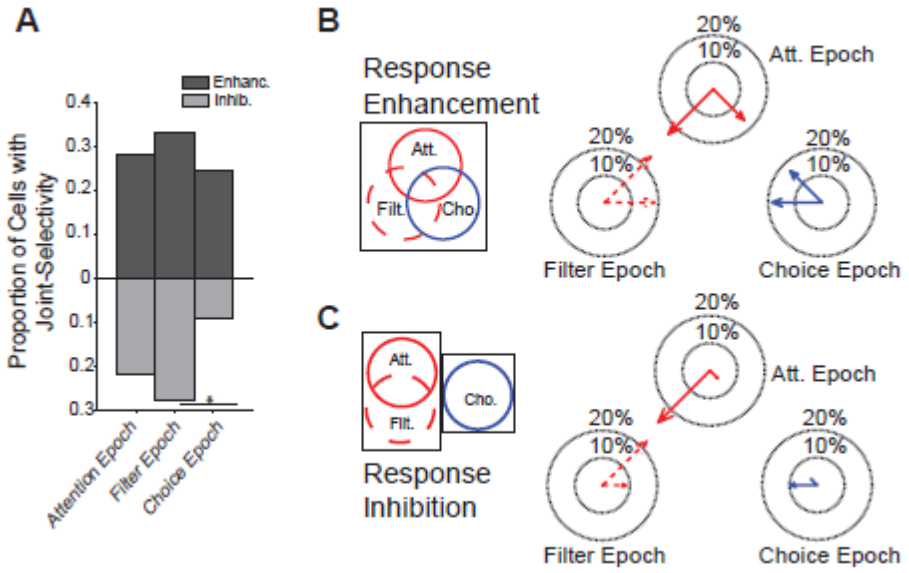

Figure 11. Proportion of cells encoding error outcomes in more than one task epoch. (A) Proportion of cells (y-axis) that encode errors in one task epoch (as indicated on the x-axis) and in at least one other epoch ('joint encoding' cells). The dark gray bars reflect join encoding for cells that increased their firing upon errors. The light gray bars denote cells showing transient firing suppression following errors. (B) The distribution of joint encoding of errors for cells that increase firing upon error commission. The length of arrows in each plot denotes the proportion of cells that jointly encode errors in the epoch where the arrows originate and where they point to. The sketched circles on the left illustrate the main result of partly

overlapping error types. (C) Same format as B but for cells with firing rate suppression upon error commission.

3.3 – Functional topography of error detection.

Next, I tested whether neurons encoding errors in a particular task epoch were anatomically located in specific subareas within the PFC of the macaque monkeys. For this purpose, I used the previously reconstructed recording locations of neurons that were projected onto a two dimensional flat map of the PFC (*see* Materials and Methods, and Kaping et al. 2011). Across the three major subdivision of the PFC that the neurons were recorded from, the ACC (area 24) hosted the largest proportion of error cells (**Figure 12A, 12B**). The salient cluster in area 24 (**Figure 12B**, randomization test, $p < 0.05$) is highly consistent with existing evidence about the functional role of the ACC in error detection (Alexander and Brown, 2011; Shenhav et al. 2013). Beyond the ACC I also found small satellite spots with significantly higher proportions of error-detecting neurons than expected by chance in dorsolateral prefrontal cortex (dlPFC) area 9 and at the border of medial frontal (mPFC) area 32 (**Figure 12B**, randomization test, $p < 0.05$). Taken together, these findings showed first that error detection was particularly prominent in the ACC with more than 25% (and locally up to 40%) of cells signaling error

outcomes across task epochs in an attention task. Secondly, the results also showed that error signals were not strictly confined to the ACC.

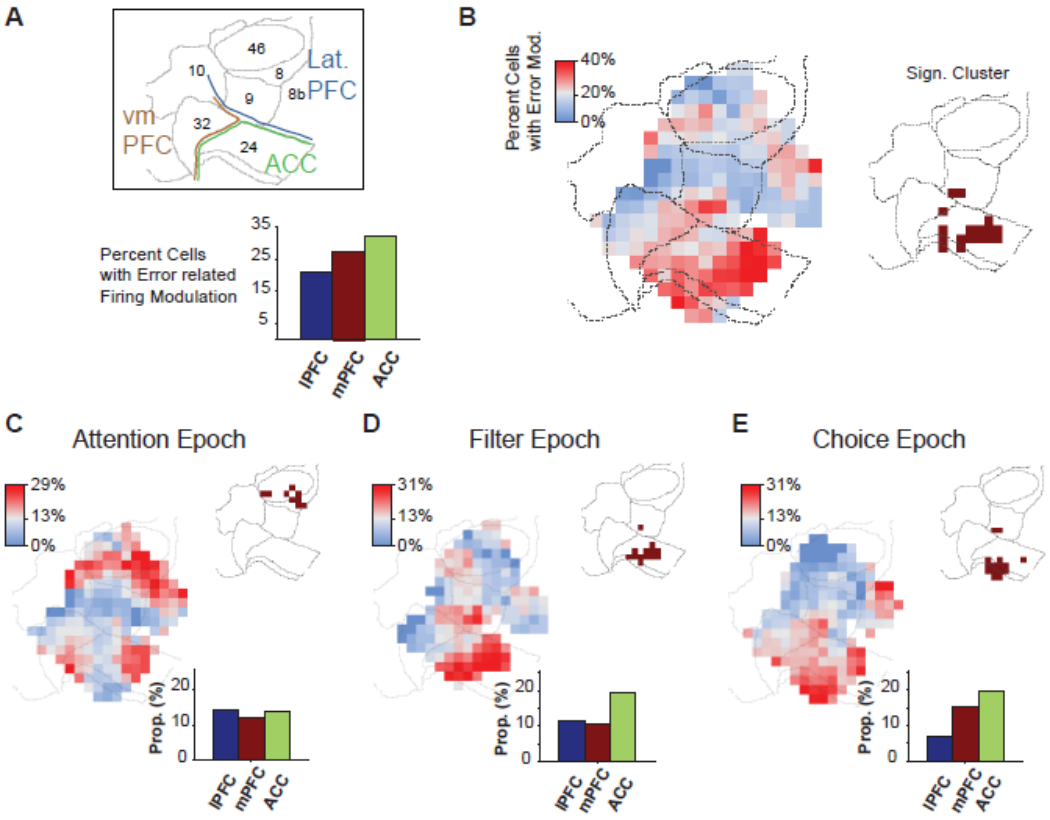

Figure 12. Anatomical distribution of error-encoding cells in the prefrontal and anterior cingulate cortex. (A) The overall proportion of cells with significant error-related firing rate modulation in the lPFC (areas 46, 8 and 9), the ACC (area 24), and the medial prefrontal cortex (mPFC, area 32). The contour map shows these areas as patches on a 2D flat map (see also Figure 1E). (B) Fine grained

functional topography of the proportion of error-encoding cells (color scale) in prefrontal and cingulate cortex. Significant spatial clustering, particularly within ACC, is shown in the contour map to the right. (C-E) Same format as in B but only for cells that signal error outcomes in the Attention Epoch (C), in the Filter Epoch (D), and in the Choice Epoch (E). The bar histograms show the average proportions of error-encoding cells per sub-area. The small contour maps show significant spatial clustering of error encoding cells for each error type. The colorbars are scaled with the distribution medians' as the intermediate (white) color.

Topographic clustering of cells encoding errors in different task epochs.

One reason for the anatomical clustering of error signals in frontal areas beyond the ACC could be an anatomically specific tuning to errors committed in selective task epochs. As shown above (**Figure 11**) more than two thirds of error-encoding neurons were exclusively selective to one error type. I therefore analyzed the topographic distribution of error-encoding neurons for each task epoch separately (**Figure 12C-E**). For the Attention Epoch the overall proportion of error-encoding cells did not vary between the dlPFC, mPFC and ACC, but fine grained functional mapping revealed a significant cluster of error-encoding cells in the dlPFC areas 46 and 8 (**Figure 12C**, randomization test, $p < 0.05$). Error encoding during the Filter Epoch was evident in up to 31% of cells in local clusters within

ACC (area 24) (**Figure 12D**, randomization test, $p < 0.05$). Errors committed during the Choice Epoch were likewise mostly encoded in local clusters of cells within the ACC (**Figure 12E**, randomization test, $p < 0.05$). The anatomical clusters of error-detecting cells in the Filter Epoch and the Choice Epoch overlapped only partly, suggesting that largely independent populations of ACC cells encoded errors in each epoch.

Anatomical dissociation of error-locked enhancement and inhibition.

In principle, the functional topographies described above could reflect a local circuit of primitive cells and interneurons tuned to error outcomes by locally coordinating their firing rates. However, errors could also trigger the activation of intra-areal brain circuits, where the activation of interneurons within one cluster triggers the inhibition of clusters within different sub-regions of the PFC. I tested whether these possibilities may be realized in the PFC by investigating the anatomical distribution of neurons that showed error-locked inhibition independent from neurons that signaled errors through response enhancement. **Figure 13A** shows that, across all PFC subregions, the proportion of neurons that transiently increased their firing upon error commission was statistically indistinguishable across different task epochs. In contrast, error-locked inhibition varied across PFC sub-regions as a function of the task epoch: within the ACC error-locked inhibition

was more than twice as likely to occur in the Choice Epoch than in the Attention Epoch ($p < 0.016$, multiple comparison corrected); the opposite pattern was found within the lPFC ($p < 0.016$, multiple comparison corrected). As illustrated in **Figure 13B**, in lPFC a large proportion of cells (up to 18% of cells in local clusters) showed error-locked inhibition in the Attention Epoch, while there were virtually no error-signaling cells with response inhibition in the Choice Epoch. Analogously, large proportion of cells in the ACC (up to 21% of cells) showed error-locked inhibition in the Choice Epoch, whereas virtually no cells encoded errors through response inhibition in the Attention Epoch.

To validate this double dissociation (ACC-lPFC and Attention-Choice Epoch), I evaluated significant spatial clustering of cells with a permutation test that corrected for uneven sampling of cells across the flat map (see material and methods). This analysis (**Figure 13C**) confirmed that a larger proportion of cells than expected by chance ($p < 0.05$) showed error-locked inhibition during the Choice Epoch within the ACC, and during the Attention Epoch within the lPFC (area 8 and 46).

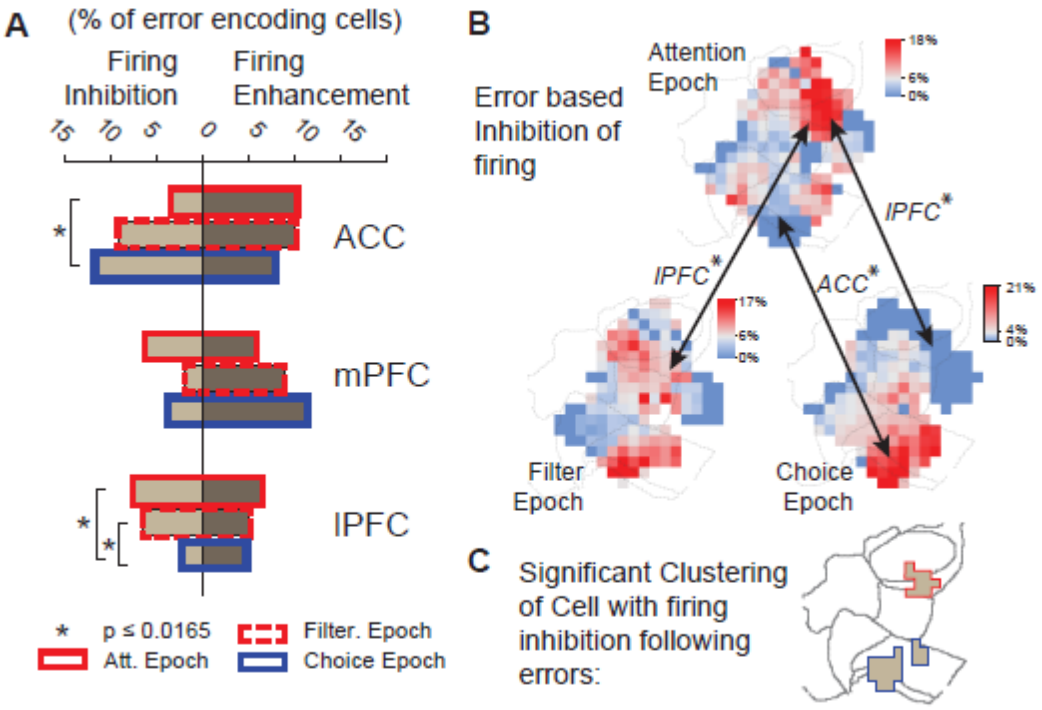

Figure 13. Anatomical distribution of cells with enhanced and suppressed firing upon errors across anterior cingulate and prefrontal cortices. (A) Proportion of cells encoding error outcomes with enhanced firing (dark gray bars) and suppressed firing (light gray bars). Significantly different proportions are indicated with a star. (B,C) Finer grained functional topography of cells with suppressed firing following errors in the Attention Epoch (top), the Filter Epoch (left bottom), and the Choice Epoch (right bottom). The double pointing arrows connect the spatial clusters where there were significantly larger proportions of error encoding cells than expected by chance as shown in the clustering map in

panel C. Red and blue bounded gray patches denote cell clusters with error-locked inhibition in the Attention Epoch (red) and Choice Epoch (blue). The colorbars in B are scaled with the distribution medians as the intermediate (white) color.

3.4 – Relative spike timing of error-locked responses.

Next, I tested the relative timing of error-locked enhancement and inhibition. For each cell with significant error-selective firing, I first calculated the latency of maximal error-locked response modulation, and then compared the cumulative distribution of cells with error-locked firing increases versus the cumulative distribution of cells with error-locked firing inhibition. I restricted this analysis to cells in ACC and lPFC, as my aim was to explain the dissociation between both areas described above (**Figure 13B**).

Figure 14A-C illustrates the cumulative distributions of error-specific response enhancement and inhibition cells. The cumulative distributions were well fitted by a sigmoidal function (or, in one case, with the combination of two sigmoids, *see* Materials and Methods), which allowed using the C_{50} parameter of the fits – latency at which the cumulative distribution reached 0.5 – as a good estimate of the latency for error encoding per sub-area (see Materials and Methods). **Figure 14D** summarizes the latencies of error-signaling through response enhancement and response inhibition, illustrating two main findings: first, error-locked firing

enhancement preceded in time to error-locked inhibition in all task epochs and brain areas with only one exception - the lPFC error signaling in the Filter Epoch that showed similar strengths of firing enhancement and inhibitions. Secondly, error-locked firing enhancement in the ACC preceded the error-locked signaling in the lPFC. This finding of the earliest latency in the ACC corroborates the special role of the ACC to detect errors in the attention task. In summary, error-detections were the fastest in the population of ACC cells that increased firing upon error commission, followed by transient response inhibition in the ACC and the lPFC.

I further validated these conclusions with an additional analysis that, unlike the cumulative distributions, took into account the unequal proportion of cells signaling distinct error types in different prefrontal regions through either enhancement or inhibition (**Figure 14E-G**). For this analysis I subtracted the estimated distribution of the proportions of cells showing error-locked firing enhancement with respect to the cells showing firing inhibition (Materials and Methods). As shown in **Figure 14E-G**, error locked enhancement activity clearly emerged earliest and strongest in the ACC compared to the lPFC in the Attention Epoch, the Filter Epoch, and the Choice Epoch.

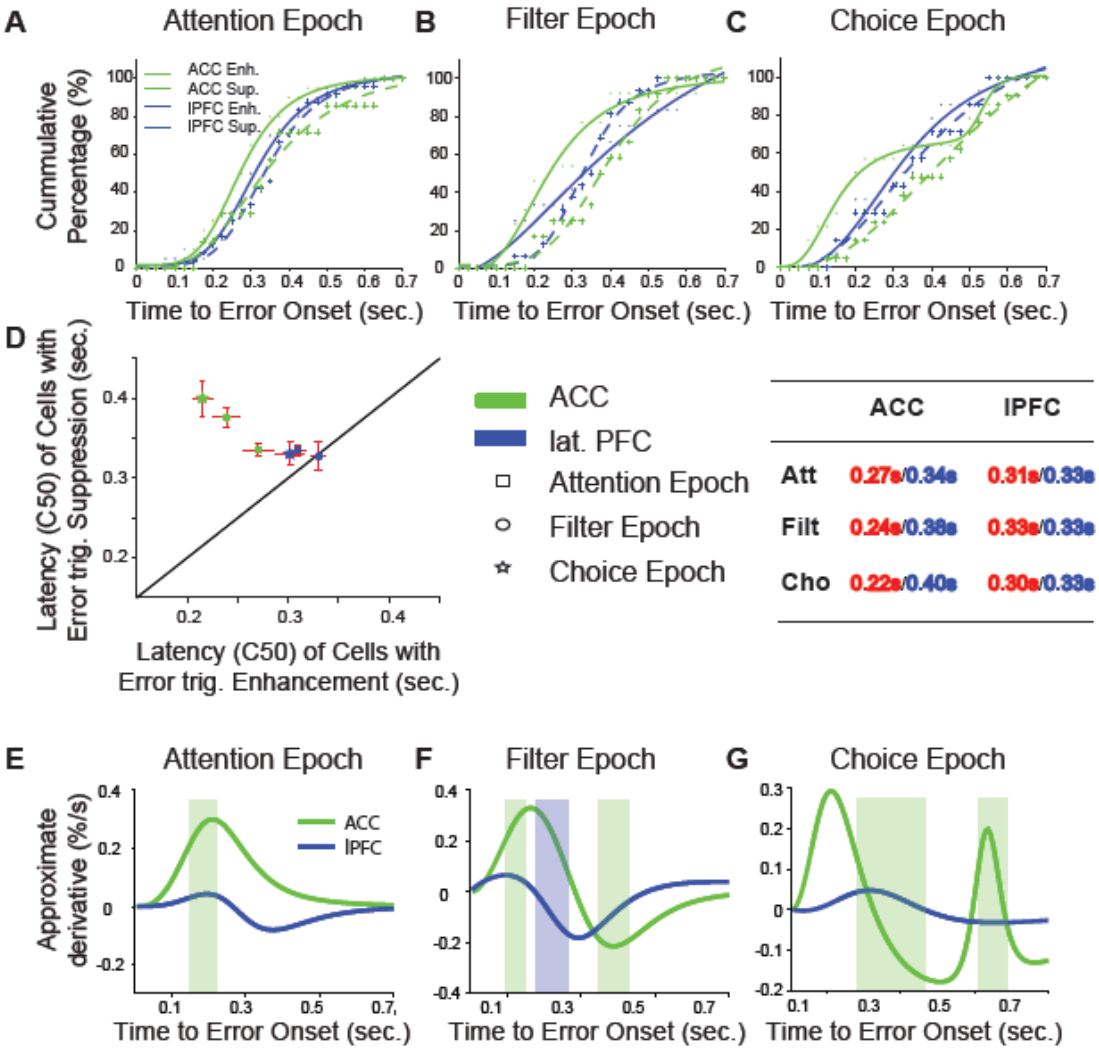
A
Attention Epoch
B
Filter Epoch
C
Choice Epoch
Cummulative Percentage (%)
ACC Enh.
ACC Sup.
IPFC Enh.
IPFC Sup.
Time to Error Onset (sec.)
D
Latency (C50) of Cells with Error trig. Suppression (sec.)
Latency (C50) of Cells with Error trig. Enhancement (sec.)
ACC
lat. PFC
Attention Epoch
Filter Epoch
Choice Epoch
ACC IPFC
Att 0.27s/0.34s 0.31s/0.33s
Filt 0.24s/0.38s 0.33s/0.33s
Cho 0.22s/0.40s 0.30s/0.33s
E
Attention Epoch
F
Filter Epoch
G
Choice Epoch
Approximate derivative (%/s)
ACC
IPFC
Time to Error Onset (sec.)

Figure 14. Latency of error encoding in the ACC and lateral PFC. (A-C) Cumulative latency distributions (y-axis) in the Attention Epoch (A), the Filter Epoch (B), and the Choice Epoch (C), with respect to time (x-axis). Raw data (small point- or cross- symbols) were fitted with a sigmoidal function (see text and Materials and Methods for details). Green / Blue lines show results from ACC and lateral PFC, respectively. Solid and dashed lines show cumulative latency distributions for cells with enhanced (solid) and suppressed (dashed) firing to errors. (D) The latency of cells with error related response inhibition (y-axis) against the latency of cells with error related response enhancement (x-axis). Latency is estimated by the C50 parameter of the sigmoidal fits shown in (A-C). Vertical and horizontal error bars denote the 95% confidence intervals around the C50 estimates. (E) The table lists the latencies (in seconds) from panel D for cells with error-locked enhancement (1st number) and inhibition (2nd number). (F-H) Difference in the proportion of error-selective cells that showed firing enhancement vs inhibition within ACC and lPFC in the Attention Epoch (E), the Filter Epoch (F), and the Choice Epoch (G) (see and Material and Methods). The green (blue) shading denotes the time intervals for which the difference in proportion between enhancement and inhibition neurons was significantly ($p<0.05$) different from zero (null hypothesis) in the ACC (green) and the lPFC (blue).

3.5 – Putative neuron types underlying error detection and error-locked inhibition.

The previous results illustrated that transient post-error inhibition followed the error-locked firing enhancement. This finding suggested that post-error enhanced firing was instrumental in triggering the post-error inhibition, and thus one can hypothesize that error signaling relies on inhibitory interneurons whose error-locked activation will impose error specific inhibition on the connected pyramidal cell populations (Medalla & Barbas 2012).

I tested this prediction by classifying the set of n=404 maximally isolated neurons into putative inhibitory interneurons and putative pyramidal cells according to their action potential waveform parameters (**Figure 8**, see Materials and Methods). The trough-to-peak duration and the time of 25%-repolarization of the cells' action potentials provided a bimodal distribution of waveform parameters with a clean separation of narrow spiking ('NS', 20%, N=79) and broad spiking ('BS' 75%, n=305) neurons, and 5% (N=20) of cells without unequivocal classification (which I labeled fuzzy cells). Of the classified cells, I found that n=70 cells showed significant error-locked firing modulation (38/32 with error-locked enhancement and suppression, respectively), comprising n=11 NS cells (16%) and n=55 BS cells (79%), (and n=4 fuzzy cells) (**Figure 15A**). The overall proportion of NS to BS cells did not differ between lateral PFC (NS: N=4, 14%; BS: N=22, 79%), medial PFC (NS: N=2, 10%; BS: N=18, 90%) and ACC (NS: N=5, 23%; BS: N=15, 68%). **Figure 15B** documents that on average, the population of cells with

error-locked response enhancement contained 74% BS cells (n=28) and 18% (N=7) NS cells. Similar overall proportions of 84% BS cells (n=27) and 13% NS cells (N=4) showed error-locked response inhibition. However, considering the subarea's of recorded cells revealed that a significant larger proportion of narrow spiking cells encoded errors in the ACC than in the lPFC through response enhancement (**Figure 15B,D**, randomization statistics, $p < 0.05$). This finding is consistent with the outlined scenario that error signaling in the ACC involves a large fraction of putative inhibitory interneurons which use their error information to inhibit connected cells within ACC and possibly within the connected lPFC (Medalla & Barbas 2009, 2012; Morecraft et al. 2012).

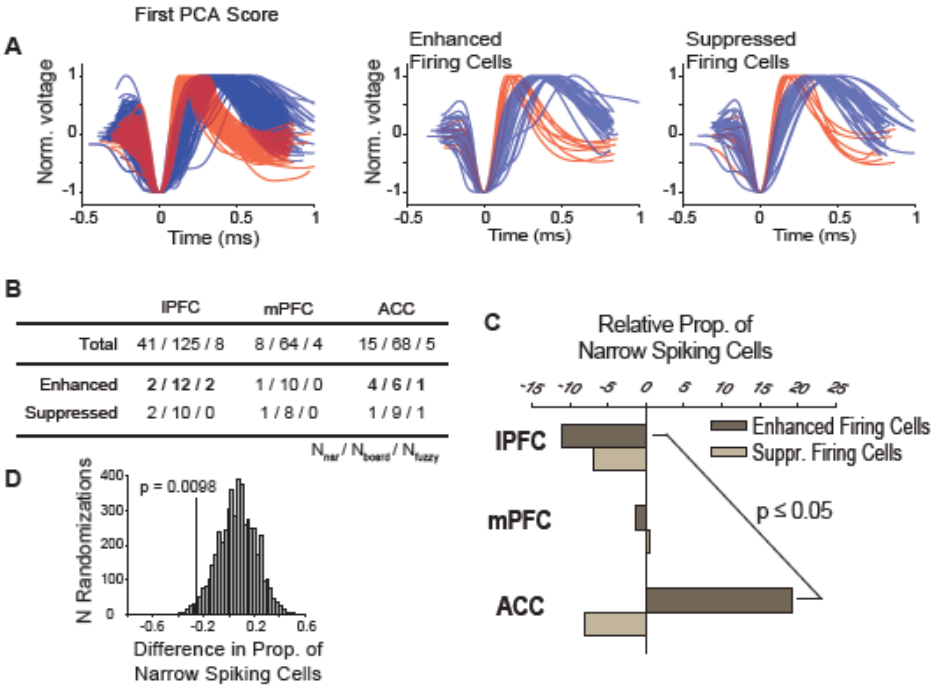

Figure 15. Putative neuron types underlying error detection and error-locked inhibition. (A) Normalized waveforms of all recorded single cells (left) and of those single cells that encoded error outcomes by enhanced and suppressed firing (middle and right panels, respectively). Red/Blue colored waveforms reflect grouping into NS/BS cells. (B) Overview of the number of NS and BS cells encoding error in lateral PFC, mPFC, and ACC. (C) Proportion of NS cells within the classes of error encoding cells in lateral PFC, mPFC and ACC. The difference between proportions of error-selective NS cells in lateral PFC and ACC was statistically significant for cells with enhanced firing following an error. (D) Shuffled random distribution of the relative difference in the proportion of NS cells

increasing firing upon error commission between lateral PFC and ACC. The observed difference is statistically lower than expected by chance, showing that there are more putative inhibitory interneurons in ACC that encode errors through enhanced firing (the comparison is shown in bold font in the table in B).

3.6 – Error-locked firing and oculomotor activity.

Previous studies have shown that error-related activity in the ACC and adjacent prefrontal cortex is largely separate, and goes beyond, motor-related activation (Ito et al. 2003; Hyman et al. 2013). To relate to these findings I tested whether the error-locked modulations that I have described were separated from saccade-related activity modulations. For this analysis I aligned the data to saccade onset in correct trials and compared to the saccade onset aligned activity modulation on error trials. My analysis revealed that cells that increased their firing upon error commission showed on average a stronger activity increase on error aligned trials compared to post-saccadic modulation on correct trials ($p < 0.05$ for all error types, Wilcoxon rank test). Similarly, cells with error-locked inhibition showed stronger firing inhibition on error trials than on saccade aligned correct trials ($p < 0.05$ for all error types, **Figure 16**).

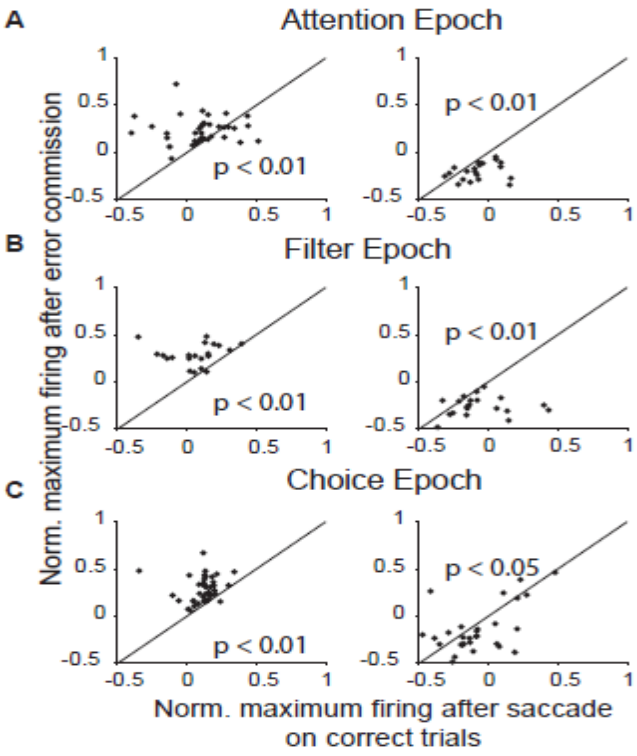

Figure 16. Error-related modulation exceeds oculomotor aligned activity modulation. (A-C) Comparison of maximal firing after errors (y-axis) and after saccadic responses on correct trials (x-axis). Panels in A-C show results for the Attention Epoch, the Filter Epoch and the Choice Epoch. Panels on the left show results for the population of cells with enhanced firing after errors and panels on the right shows results for cells with suppressed firing following errors. Statistical p values are based on paired Wilcoxon signed-rank test.

Chapter 4 – Discussion

4.1 – Behaviorally separable error types in selective attention task

Our task required the monkeys to covertly attend to the peripheral stimulus while fixating at the centre fixation cue, and discriminate changes in the rotation of the peripheral stimulus by saccading to the correct response target. In our task, the peripheral stimuli were located to the left and right of the centre fixation cue while the response targets were located above and below the fixation cue (**Figure 1**). Thus, to examine behavioral differences between error types, I categorized erroneous saccades into up, down, left, and right quadrants centred on the fixation points (see material and methods).

The overall direction of erroneous saccades were distinctively different between error types, and these behavioral differences likely reflected differences in the underlying cognitive deficits leading to the commission of each error type. Amongst the three error types, errors committed during attentional epoch were the most evenly distributed within the four quadrants. While errors committed during choice epoch were the most biased toward the target stimuli in the vertical quadrants. Errors committed during filtering epoch, however, fell in between the two extremes. Filter epoch errors were disturbed mostly toward the target stimuli in the vertical quadrants but were also distributed towards the peripheral stimuli in the horizontal quadrants.

These results confirmed that errors committed during different epochs were likely elicited by different underlying cognitive deficiencies. Break fixation errors during attentional epoch likely reflected the lack of top-down attentional control, or the lack of inhibition of impulsive saccades. Vertically orientated response errors during choice period were likely caused by incorrect sensory-response mapping between the direction of stimulus rotation and the correct saccadic response. Whereas filter epoch errors were likely due to erroneous top-down facilitation of attention to the distractor (vertical response errors), or the failure to prevent bottom up capture of attention by the salient distractor stimuli (horizontal break fixation errors).

I hypothesize that these different error types must be discriminated by the monkeys in order for them to implement the correct adjustment measures. Thus, my behavioral results warrant the investigation for whether different neurons within the prefrontal cortex discriminately code for these three error types.

4.2 - Performance monitoring and selective error encoding by the ACC activation

Previous electrophysiological studies reported up to 14% of the ACC neurons were activated following error events (Matsumoto et al. 2007, Quilodran et al. 2008, Ito et al. 2003, Amiez et al. 2005). Albeit, these studies examined different performance failures in isolation: negative prediction errors were studied

by Matsumoto (2007), break fixation errors were studied by Quildran (2008) and Amiez (2005), while Ito (2003) studied incorrect choice errors. Consistent with the previous findings, 12% of the our recorded units within the ACC were found to be encoding errors during attentional epoch, 17% encoded for errors during filter epoch, and 18% encoded for errors during choice epoch **(Figure 12 C-E**). Surprisingly, our findings showed that more than two thirds of our error-encoding neurons were exclusively selective to one error type, and up to 29% of the recorded neurons within the ACC encoded for any one given error type (**Figure 12A**). This result is a strong indicator that neurons within the ACC monitor and categorize error events, and that previous electrophysiological studies have largely under-estimated the ACC's involvement in performance monitoring.

The timing difference in error-selective encoding within the ACC and lPFC has been reported in an earlier study. Rothe and colleagues (2011) observed that an early error response within the ACC was followed by shortly by a sustained response within the lPFC. Our findings were consistent with this earlier report, and demonstrated that error-driven activation of the ACC is followed by error-driven suppression of other PFC regions, suggesting that error detection is likely reflected by the former. A recent model proposed by Alexander and Brown (2011) predicted that error or unexpected events are reflected by the increased firing-rate of neurons within the ACC. Our finding that the activation of ACC is reflective of error monitoring resonates well with this proposed model.

4.3 - Adaptive configuration of cognitive control by the PFC inhibition

Error-driven activation of the ACC was followed immediately by error-driven inhibition within regions of the prefrontal cortex – the ACC and the lPFC. The relative delay between ACC activation and prefrontal inhibition suggested that the error-driven inhibition within the prefrontal cortex was influenced by the activation of ACC inhibitory interneurons following an erroneous outcome. Indeed, the ACC has been shown to contain dense local inter-connection as well as forming long-range connection with the lPFC (Barbas & Zikopoulos 2007). Furthermore, I have shown that a relative large percentage of inhibitory interneurons were activated within the ACC following error onset (**Figure 15**). Taken all together, the activation of ACC interneurons could possibly be the source of local inhibition of the ACC as well as long-range inhibition of the lPFC.

The anatomical distribution of error-driven inhibition differed across error types and likely reflected differences in cognitive demands between error types. During attentional epoch, errors-driven inhibitions were non-homogenously clustered within the dlPFC, area 46 and 9 (**Figure 13**). dlPFC has long been proposed to play a role in directing selective attention and inhibiting inappropriate actions. Lesion studies have suggested that dlPFC damage results in immature or impulsive responses during learning tasks (Aron 2007; Chambers et al. 2006). Studies have also demonstrated that inhibition of the dlPFC via calbindin interneurons increased selective attention by reducing task-irrelevant signals

(Medalla and Barabas 2010). As demonstrated in our behavioral analysis, break fixation errors made during motivational periods were scattered in all directions and were likely due to the lack of top-down maintenance of focus on the task at hand, or the lack of top-down inhibition of inappropriate saccade (**Figure 9A**). Transient inhibition of the dlPFC might be one of the mechanisms by which the dlPFC readjusts top-down control to focus on task-relevant processes and suppress inappropriate responses after error.

During Choice Epoch, error-driven inhibitions were non-homogenously clustered within the ACC (**Figure 13**). Anatomical studies have shown direct innervation of the ACC with premotor cortex, and the ACC has been reported to encode for action-reward association as well as directing reward orientated responses (Hayden & Platt, 2010). As indicated by our behavioral analysis, erroneous saccades during Choice Epoch were almost exclusively towards the response targets and were likely due to inappropriate action-reward mapping (**Figure 9C**). Therefore, inhibitions of the ACC were likely associated with reinforcing the correct action-reward mappings to ensure proper responses in the subsequent trials following error.

Considering that there appeared to be a clear behavioural distinction between saccadic errors committed during Attentional and Choice Epoch, it was not surprising to find distinctly different topographical distribution of inhibition

following error onset during Attentional and Choice Epoch. Errors committed within Filter Epoch, however, triggered wide-spread inhibition within both lPFC and ACC (**Figure 13**). As indicated by our behavioral analysis, saccadic errors during Filter Epoch were oriented mostly towards the target stimuli but a sizeable amount were also distributed towards the peripheral stimuli (**Figure 9B**). Thus, the spread of inhibition following Filter Epoch errors might have reflected the failures of more heterogeneous categories of attentional processes.

This idea that a unique set of error-driven inhibition neurons were elicited by a unique error type was also supported by our joint-signaling analysis. Neurons displaying error-driven inhibition during Attentional and Choice Epoch rarely overlapped with each other. Taken together with our topographical analysis, these findings suggested that errors committed during attentional and choice periods were distinctively separated from each other, and inhibited two very distinct population of neurons. A relatively large percentage of neurons were inhibition following error-onset during both Filter and Attentional epoch, but neurons displaying error-driven inhibition during Filter Epoch were mostly distinct from the neurons inhibited during Choice Epoch. This suggested that errors occurring during Filter Epoch reflected more attentional type of errors than choice type of errors (**Figure 11C**).

4.4 – Ocular motor activity

I have demonstrated that neurons encoding for errors were also modulated by saccade-related activities. However, the error-driven component of the modulation could still be separated by comparing saccade aligned firing rates between correct trials and error trials. My findings were not unusual. Previous studies have demonstrated that prefrontal neurons can track multiple task-relevant factors (Lapish et al. 2008; Hyman et al. 2010, 2011). In particular, error encoding cells within the prefrontal cortex have been reported to also monitor for action-related activities (Hyman et al. 2010; Ito et al. 2003).

Taken together, I have demonstrated that error-encoding neurons also encoded for other task-relevant factors. Our findings supports the idea that the prefrontal neurons encode for multi-dimensional variables. One can hypothesize that these additional task-relevant factors add context to the errors, and can be used to distinguish error types and specify the necessary corrections for subsequent trials.

Chapter 5 – Summary and Conclusions

My results were consistent with the established view that the ACC is involved in error monitoring and error resolution. Importantly, my results suggested that the ACC's role in error monitoring extended beyond mere error detection, but also error identification. Specifically, I have demonstrated that the ACC can distinguish between errors caused by lapse of top-down attention, errors caused by bottom-up attentional capture, and errors due to failures to associate the cued stimulus rotations with the correct responses. This makes intuitive sense, since resolution of performance failures cannot rely solely on error detection, but requires that the type of error to be identified.

From these findings two questions arise: 1) how were error categorizations made immediately following error onset, and 2) how were these error signals being utilized by the prefrontal cortex. My results provided suggestions to both of these questions. To answer the first question, I have demonstrated that error-signaling neurons within the ACC were not only highly tuned to detecting error types, but were also encoding for events beyond mere error detection (e.g. motor-related events). These observations suggested that different subsets of ACC neurons may monitor a particular cognitive state. Errors caused by the failure of a particular cognitive state were then perhaps signaled by the same neurons responsible for monitoring the overall state of that particular process. As for the second question, I

have provided evidence suggesting that the ACC signaled to other regions of the prefrontal cortex for the need of adjustment through inhibitory interneurons. Importantly, we have shown that different error types triggered distinct clusters of inhibition within the prefrontal cortex. Taken all together, this suggested that locally clustered neurons within the ACC can detect and identify error types, and allocate specific correctional adjustments by inhibiting distinct regions of the prefrontal cortex.

Bibliography

Alexander WH, Brown JW (2011) Medial prefrontal cortex as an action-outcome predictor. Nat Neurosci 14:1338–1344.

Amiez C, Joseph J-P, Procyk E (2005) Anterior cingulate error-related activity is modulated by predicted reward. Eur J Neurosci 21:3447–3452.

Aron AR (2007) The neural basis of inhibition in cognitive control. Neuroscientist 13:214–228.

Asaad WF, Eskandar EN (2008a) A flexible software tool for temporally-precise behavioral control in Matlab. J Neurosci Methods 174:245–258.

Asaad WF, Eskandar EN (2008b) Achieving behavioral control with millisecond resolution in a high-level programming environment. J Neurosci Methods 173:235–240.

Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. J Neurosci 14:4467–4480.

Barbas H, Zikopoulos B (2007) The prefrontal cortex and flexible behavior. Neuroscientist 13:532–545.

Becerril KE, Repovs G, Barch DM (2011) Error processing network dynamics in schizophrenia. Neuroimage 54:1495–1505.

Berridge CW, Waterhouse BD (2003) The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. Brain Res Brain Res Rev 42:33–84.

Cavanagh JF, Cohen MX, Allen JJB (2009) Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. J Neurosci 29:98–105.

Chambers CD, Bellgrove M a, Stokes MG, Henderson TR, Garavan H, Robertson IH, Morris AP, Mattingley JB (2006) Executive "brake failure" following deactivation of human frontal lobe. J Cogn Neurosci 18:444–455.

Cheng M-Y, Hall P (1998) Calibrating the excess mass and dip tests of modality. J R Stat Soc Ser B (Statistical Methodol 60:579–589.

Collier TJ, Greene JG, Felten DL, Stevens SY, Collier KS (2004) Reduced cortical noradrenergic neurotransmission is associated with increased neophobia and impaired spatial memory in aged rats. Neurobiol Aging 25:209–221.

Devauges V, Sara SJ (1990) Activation of the noradrenergic system facilitates an attentional shift in the rat. Behav Brain Res 39:19–28.

Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1991) Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. Electroencephalogr Clin Neurophysiol 78:447–455.

Falkenstein M, Hoormann J, Christ S, Hohnsbein J (2000) ERP components on reaction errors and their functional significance: a tutorial. Biol Psychol 51:87–107.

Foote SL, Aston-Jones G, Bloom FE (1980) Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal. Proc Natl Acad Sci U S A 77:3033–3037.

Franken IHA, Rassin E, Muris P (2007) The assessment of anhedonia in clinical and non-clinical populations: further validation of the Snaith-Hamilton Pleasure Scale (SHAPS). J Affect Disord 99:83–89.

Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (1993) A NEURAL SYSTEM FOR ERROR DETECTION AND COMPENSATION. Psychol Sci 4:385–390.

Goldstein DS, Eisenhofer G, McCarty R, Schultz W (1997) The Phasic Reward Signal of Primate Dopamine Neurons. Adv Pharmacol 42:686–690.

Harley CW (2004) Norepinephrine and dopamine as learning signals. Neural Plast 11:191–204.

Hartigan JA, Hartigan PM (1985) The Dip Test of Unimodality. Ann Stat 13:70–84.

Hatzigiakoumis DS, Martinotti G, Giannantonio M Di, Janiri L (2011) Anhedonia and substance dependence: clinical correlates and treatment options. Front psychiatry 2:10.

Hayden BY, Platt ML (2010) Neurons in anterior cingulate cortex multiplex information about reward and action. J Neurosci 30:3339–3346.

Hebb (1949) The Organization of Behavior. New York: Wiley.

Henderson DJ, Parmeter CF, Russell RR (2008) Modes, weighted modes, and calibrated modes: evidence of clustering using modality tests. J Appl Econom 23:607–638.

Holroyd CB, Coles MGH (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109:679–709.

Holroyd CB, Yeung N (2012) Motivation of extended behaviors by anterior cingulate cortex. Trends Cogn Sci 16:122–128.

Hyman JM, Hasselmo ME, Seamans JK (2011) What is the Functional Relevance of Prefrontal Cortex Entrainment to Hippocampal Theta Rhythms? Front Neurosci 5:24.

Hyman JM, Whitman J, Emberly E, Woodward TS, Seamans JK (2013) Action and outcome activity state patterns in the anterior cingulate cortex. Cereb Cortex 23:1257–1268.

Hyman JM, Zilli EA, Paley AM, Hasselmo ME (2010) Working Memory Performance Correlates with Prefrontal-Hippocampal Theta Interactions but not with Prefrontal Neuron Firing Rates. Front Integr Neurosci 4:2.

Ito S, Stuphorn V, Brown JW, Schall JD (2003) Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 302:120–122.

Kaping D, Vinck M, Hutchison RM, Everling S, Womelsdorf T (2011) Specific contributions of ventromedial, anterior cingulate, and lateral prefrontal cortex for attentional selection and stimulus valuation. Behrens T, ed. PLoS Biol 9:e1001224.

Kaye H, Pearce JM (1984) The strength of the orienting response during Pavlovian conditioning. J Exp Psychol Anim Behav Process 10:90–109.

Konorski J (1948) Conditioned reflexes and neuron organization. New York, NY: Cambridge University Press.

Lachenbruch PA (2002) Analysis of data with excess zeros. Stat Methods Med Res 11:297–302.

Lapish CC, Durstewitz D, Chandler LJ, Seamans JK (2008) Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. Proc Natl Acad Sci U S A 105:11963–11968.

M P, DN P (2004) The Human Nervous System, 2end ed. Amsterdam: Elsevier.

Mackintosh NJ (1975) A Theory of Attention: Variations in the Associability of Stimuli with Reinforcement. Psychol Rev 82:276–298.

Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. Nat Neurosci 10:647–656.

Medalla M, Barbas H (2009) Synapses with inhibitory neurons differentiate anterior cingulate from dorsolateral prefrontal pathways associated with cognitive control. Neuron 61:609–620.

Medalla M, Barbas H (2010) Anterior cingulate synapses in prefrontal areas 10 and 46 suggest differential influence in cognitive control. J Neurosci 30:16068–16081.

Medalla M, Barbas H (2012) The anterior cingulate cortex may enhance inhibition of lateral prefrontal cortex via m2 cholinergic receptors at dual synaptic sites. J Neurosci 32:15611–15625.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202.

Miltner WH, Braun CH, Coles MG (1997) Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. J Cogn Neurosci 9:788–798.

Morecraft RJ, Stilwell-Morecraft KS, Cipolloni PB, Ge J, McNeal DW, Pandya DN (2012) Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. Brain Res Bull 87:457–497.

Németh G, Hegedüs K, Molnár L (1988) Akinetic mutism associated with bicingular lesions: clinicopathological and functional anatomical correlates. Eur Arch Psychiatry Neurol Sci 237:218–222.

Nieuwenhuis S, Aston-Jones G, Cohen JD (2005) Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychol Bull 131:510–532.

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869.

Passingham R, Wise S (2012) The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight. Oxford: Oxford University Press.

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

Petrides M (1996) Specialized systems for the processing of mnemonic information within the primate frontal cortex. Philos Trans R Soc Lond B Biol Sci 351:1455–61; discussion 1461–2.

Petrides M, Pandya DN (2007) Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J Neurosci 27:11573–11586.

Quilodran R, Rothé M, Procyk E (2008a) Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. Neuron 57:314–325.

Quilodran R, Rothé M, Procyk E (2008b) Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron 57:314–325.

Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. :64 – 99.

Rothé M, Quilodran R, Sallet J, Procyk E (2011) Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. J Neurosci 31:11110–11117.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. :318–362.

Rushworth MFS, Behrens TEJ (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 11:389–397.

S. K (1970) The biogenic amines in the central nervous system: their possible roles in arousal, emotion and learning. New York: Rockefeller Press.

Saleem KS, Kondo H, Price JL (2008) Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. J Comp Neurol 506:659–693.

Sara SJ, Segal M (1991) Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: implications for cognition. Prog Brain Res 88:571–585.

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. Annu Rev Neurosci 23:473–500.

Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79:217–240.

Sutton R, Barto A (1998) Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning). A Bradford Book.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001) An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc 8:443–459.

Walton ME, Bannerman DM, Alterescu K, Rushworth MFS (2003) Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. J Neurosci 23:6475–6479.

Yeung N, Botvinick MM, Cohen JD (2004) The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol Rev 111:931–959.

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681–692.