Title
Encoding externally and internally accredited value in prefrontal cortex

Author
Luk, Chung-Hay

Publication Date
2011

Peer reviewed|Thesis/dissertation

Encoding externally and internally accredited value in prefrontal cortex

by

Chung-Hay Luk

A dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jonathan D. Wallis, Chair
Professor Silvia A. Bunge
Professor Richard B. Ivry
Professor Stuart J. Russell

Fall 2011

Encoding externally and internally accredited value in prefrontal cortex

by

Chung-Hay Luk

Abstract

Encoding externally and internally accredited value in prefrontal cortex

by

Chung-Hay Luk

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Jonathan D. Wallis, Chair

Given the prefrontal cortex (PFC) is involved in tracking consequences from both one's environment and actions, this work attempts to characterize the involved neuronal processes. In the first portion of this thesis, we focus on internally accredited value, namely the value of actions. We recorded the spiking activity from single neurons of two subjects while they performed a task that required them to monitor relationships between their actions and resulting outcomes on a trial-by-trial basis. We contrasted the activity from two brain areas known to have projections to motor areas but different limbic and attentional roles: dorsolateral prefrontal cortex (PFdl) and medial prefrontal cortex, dorsal to the cingulate sulcus (PFcs). During the epoch after the first action-outcome (AO) association was revealed to the subject, nearly twice as many PFdl neurons encoded the action compared to the outcome. Conversely, more PFcs neurons encoded the outcome as compared to the action. In both brain areas, information about action and outcome were separated in two populations: few neurons encoded action and outcome. When the subjects learned a subsequent AO association, that outcome was encoded relative to the first, i.e., better or worse, rather than as the identity of the outcome, i.e., apple juice or quinine. Again, PFcs had more reward encoding than PFdl at this time. These data support past studies that implicate PFcs in monitoring the value of actions for value-based decision making.

The second portion of this work contrasts neuronal representations of AO associations with that of stimulus-outcome (SO) associations. Lesion and imaging studies have suggested that AO encoding relies on PFcs and SO encoding on orbitofrontal cortex (PFo). We hypothesized similar dissociation at the neuronal level. To test this idea, we trained the same two subjects on two tasks, one that relied on AO associations to solve and another that relied on SO associations. The SO task was analogous to the aforementioned AO task, except that pictures, rather than actions, were associated with the delivery of juice rewards. While the subjects performed these tasks, we recorded the neuronal activity from PFcs and PFo. Both areas had neurons encoding stimuli, actions, and outcomes across the tasks. There

was only a subtle bias for PFo neurons preferring the SO task. Most notable was the prominent action selectivity in PFo for the action the subject would perform to select his more preferred reward in the SO task. No such bias was seen in the AO task. Collectively, this work implicates PFcs and PFo as key regions in goal-directed behavior given their ability to flexibly represent value as it is attributed to either external cues or internally generated motor actions.

Acknowledgements

Jon Wallis

Jon runs the lab as a CEO would at an efficient business. He has cultivated a great team with his meticulous training style and on-task, but relaxed culture. I admire his ability to lead by example and keep a level head. He is always dedicated to our work, which has included (to the shock of his fellow Britons) knowingly scheduling and performing an experiment-related surgery during a World Cup game with England playing. I learned a lot while in Jon's lab from the expected neuroscience, statistics, and technical writing to the more surprising NATO alphabet, air-filtering properties of spider plants, and Justin Bieber's latest hairdo. Thank you for all the insight and making this mental marathon a rewarding one!

My labmates and thesis commitee

They have provided me with much help (code, loosening screws that I can't physically undo, paper recommendations...), good questions (lots of them!), and solid advice (grant writing tips, analyses for my work, the 2π rule - the multiplier for the amount of time some task takes in actuality). Special thanks to Steve Kennerley, my go-to person in lab. He is the one I cite most often when I discuss my work. To this day, I aspire to have his energy while discussing science.

Shaowen Bao

My introduction to animal work was in Shaowen's lab. He was a good mentor and made rat electrophysiology easy enough to learn that I could build up to primate work. As I finish my doctoral studies, it has been his work in conjunction with my doctoral work that has guided me in my future path. Thank you for always supporting me.

Foxtrot, Hotel, and Juliet

They are the brains beyond all my work, and I thank them for their dedication in learning such complex tasks that no normal animal could.

Eric Boyd, Lynne Bruning, Yu-Chun Huang

You literally brighten up my days. Thank you for the collaborations and support!

My passport

How I cherish you, golden ticket to fun.

Amy Orsborn

If anyone could build the cyborg equivalent of Thomas Keller, Amy would be that person. Thanks for the delicious food and witty conversations.

Erica Warp

Fellow neuroscientist-by-day, artist-by-night, I love seeing what Erica creates and cheering her on as she does for me.

Jascha Sohl-Dickstein

Playing with fire, stargazing, Maker Faire outings, and adventures to come. Jascha keeps life captivating for all of us around him.

Ipsheeta Furtado

Her personable demeanor makes her my favorite wingman at networking events. Thanks for joining me at my first hackathon and being my first customer in my wearable electronics endeavor!

Jen Sloan

I cannot express how lucky I am to have such a great friend working in the same lab as I. I can always rely on her discerning eye, leadership, and sensibility.

Alejandra Dubcovsky Joseph

Former roommate, best friend, and superhuman in all the activities she juggles and excels at. Her energy rubs off on me. It was always a joy starting the day with our breakfast meet-ups.

Matt Sparks

He keeps me sane by being my voice of reason, chocolate provider, and personalized tech support with tutorials and filtered TechCrunch newsfeeds. Thank you for getting me started with git and fostering better code and back-up habits. I cherish our think sessions.

My family

My parents and brother have been infinitely supportive. I owe much of my fascination with the natural world to my dad, who introduced my brother and me to lasers, soldering, and national parks early on. Many thanks to my mom for my pragmatic sense of design. I am so fortunate to have my family close-by should I ever need a home-cooked meal, some life advice, or to check-in before spending vacation time traveling.

Dedicated to my mom and dad

Contents

Chapter 1

Introduction

We live in a dynamic world, and things are constantly changing. For us to successfully navigate such an environment, we must keep track of the numerous meanings of environmental cues, other people's actions as well as our own actions. We quickly learn that an order of chips in England yields you thick, fried wedges of potato (French fries) but thin, fried slices of potato in America. Kissing on the cheek is a welcomed gesture in France, but frowned upon in Japan. For more complex scenarios, we may learn connections between a cue, an action, and a subsequent outcome. Suppose you move to Hong Kong for work and see the sign " 拉 " on your office door for the first time. You deduce it to mean either push or pull. But until you have executed one of those actions, you do not know the association between that particular stimulus, action, and outcome. Hence, you decide to apply force on the door and realize that it does not budge. " 拉 " must mean pull then, you conclude as you test that out. After months of associating " 拉 " with the action pull, you have mastered this Chinese character.

How exactly does your mind learn to associate items in the environment with actions and consequences? One of the areas in the brain that may be critical in this learning is the prefrontal cortex (PFC). PFC is important in novel situations [Schoenbaum *et al.*, 2002; Wallis and Miller, 2003b; Antzoulatos and Miller, 2011]. It can encode diverse aspects of the environment, such as objects [Wilson *et al.*, 1993; Rao *et al.*, 1997; Wallis and Miller, 2003b], faces [Tsao *et al.*, 2008], spatial location of stimuli [Funahashi *et al.*, 1993; Rao *et al.*, 1997; Kennerley and Wallis, 2009], sounds [Fuster, 1973; Fuster *et al.*, 2000] and odors [Eichenbaum *et al.*, 1983; Schoenbaum *et al.*, 1998]. PFC also carries action information [Matsumoto *et al.*, 2003; Matsumoto *et al.*, 2007; Furuyashiki *et al.*, 2008; Tsujimoto *et al.*, 2009]. Lastly, it is well known for having a rich representation of outcomes, capturing physical components of rewards [Thorpe *et al.*, 1983; Tremblay and Schultz, 1999], effects of satiety [Gottfried *et al.*, 2003; Bouret and Richmond, 2010], how reliable an outcome is [Behrens *et al.*, 2007; Kennerley *et al.*, 2009], economic value [Breiter *et al.*, 2001; Glimcher, 2003], as well as emotional and social components [Damasio, 1994; Rudebeck *et al.*, 2006]. In order to form

associations between stimuli, actions, and outcomes, these pieces of information need to be store across periods of time [Fuster, 1973] and integrated together in a goal-relevant manner [Miller and Cohen, 2001], both of which are characteristic functions of PFC.

1.1 Outline of thesis

The thesis consists of four chapters beyond this current one. The following Chapter 2 describes the procedures and equipment used across experiments. Afterward, each of the two experiments is described in full. The first experiment is in Chapter 3 and addresses the neuronal signatures of AO associations in medial and lateral PFC. The second experiment is in Chapter 4 and contrasts the encoding of AO associations with SO associations in medial and orbital PFC. Lastly, Chapter 5 closes with conclusions for each experiment and future directions.

In the remainder of this chapter, we introduce concepts for understanding PFC and its role in learning associations between stimuli, actions, and outcomes. We begin with background on the anatomy of the three PFC regions, starting with medial PFC, then lateral PFC, and lastly orbital PFC. We also introduce specific subsections of these regions, which are the subsequent foci of our experiments. Next we provide a theoretical framework for studying stimulus-action-outcome (SAO) associations. Specifically, we explain two functions of PFC necessary for learning SAO association. Then we describe the functionality of the brain areas in relation to SAO encoding. Lastly, we show parallels between SAO learning and the clinical problem of addiction.

1.2 Anatomy of the prefrontal cortex

The prefrontal cortex (PFC) occupies the front portion of the frontal lobe just anterior to the motor regions of the brain. The border between PFC and the adjacent premotor cortex is approximately at the arcuate sulcus in the monkey [Barbas and Pandya, 1989] but lacks a dividing landmark in the human [Semendeferi *et al.*, 2002]. Rats also lack a clear landmark, because they do not have sulci. Consequently, the generally accepted method for defining the prefrontal region across species is to locate the projection zone from the mediodorsal nucleus of the thalamus [Rose and Woolsey, 1948]. From so doing, PFC occupies nearly a third of the human neocortex [Fuster, 1997]. It is the slowest region to develop and reaches full maturation only after a person enters their twenties [Giedd *et al.*, 1999]. Likewise, in the monkey, prefrontal development is slow relative to the other cortices [Fuster, 2001]. PFC has connections with subcortical structures, motor areas and all higher-level sensory areas. Thus, it is central for integrating information [Modha and Singh, 2010]. Further, the connections are comparable across monkey and human [Croxson *et al.*, 2005; Klein *et al.*, 2010]. As illustrated in fig. 1.1, PFC can be separated into three regions – medial, lateral, and orbital

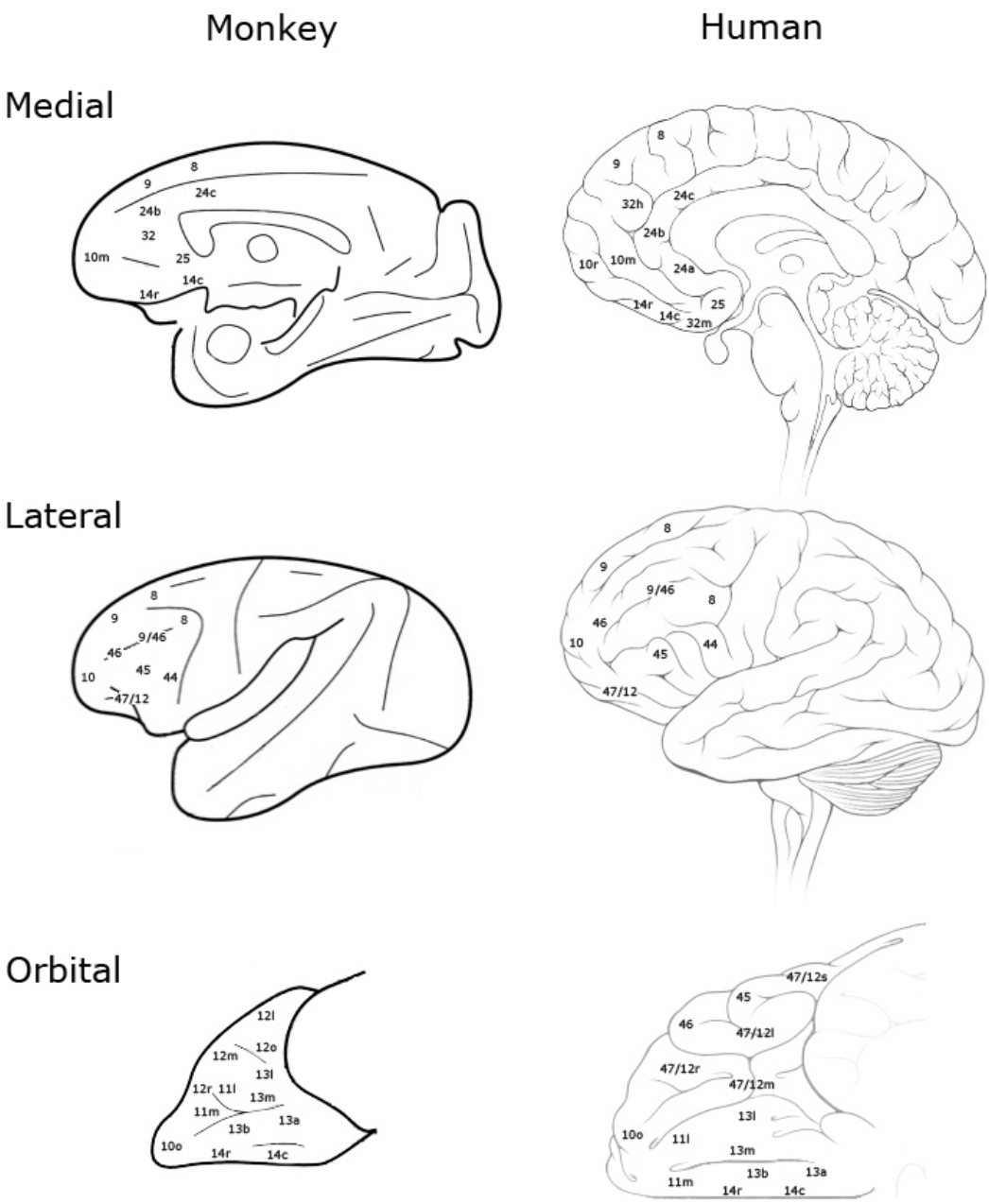

Figure 1.1: Medial, lateral, and orbital views of the prefrontal cortex in monkey and human. Brain outlines taken from [Carmichael and Price, 1996; Petrides and Pandya, 1999] for monkey and [Öngür and Price, 2000] for human.

– all of which are further subdivided into areas based on cytoarchitecture [Brodmann, 1909; Petrides and Pandya, 1994; Carmichael and Price, 1996; Öngür and Price, 2000]. Each area is addressed individually in the following sections with descriptions of its composition and morphology across species, as well as patterns of connectivity with other brain areas. In discussing connectivity, it is important to note that, unless otherwise stated, connections are always reciprocal between areas.

1.2.1 Anatomy of medial prefrontal cortex

The medial prefrontal region occupies the medial wall of the prefrontal cortex and overlaps with the anterior end of the anterior cingulate cortex (ACC). In monkeys, it consists of the medial portions of areas 9 and 8 at the dorsal extent. Ventral to those are areas 32 and 24. The most ventral part of the medial wall consists of area 25 and medial portions of areas 14 and 10 (Figure 1.1). Moving rostrocaudally along the cingulate sulcus, the cytoarchitecture transitions from agranular lamination (tissue lacks layer IV) to dysgranular (tissue with a poorly defined layer IV) lamination. Medial PFC in humans is comparable with subtle differences. It contains similar divisions with the notable difference that area 10 is greatly expanded [Öngür *et al.*, 2003]. Cytoarchitecture follows the monkey for the most part, though area 24 in the cingulate sulcus has noticeably larger pyramidal neurons in layer V, which is not shared by monkeys or rats [Nimchinsky *et al.*, 1996]. Large pyramidal neurons in layer V may provide strong efferents from the area, similar to Betz cells in the primary motor cortex that project to the spinal cord. Compared to monkeys and humans, medial PFC in rats is more rudimentary. It contains fewer divisions and simpler cell morphology. Medial PFC subregions are infralimbic and prelimbic cortices, dorsal and ventral ACC, and medial precentral cortex [Öngür and Price, 2000]. Dorsal and ventral ACC can also be referred to as areas Cg1 and Cg2, while prelimbic area can be referred to as Cg3 [Zilles and Wree, 1995]. Furthermore, the cytoarchitecture in rodent medial PFC is completely agranular. In fact, the entire rat PFC consists of agranular tissue.

To date, most primate electrophysiological studies have focused on a single area of medial PFC: the region surrounding the portion of the cingulate sulcus anterior to the genu of the corpus callosum [Matsumoto *et al.*, 2003]. This is, in part, because this region is most accessible for electrophysiological studies. It lies close to the surface of the brain and is sufficiently lateral to avoid any accidental contact with the central sinus. We refer to this prefrontal subregion as PFcs due to its location within the cingulate sulcus. Although PFcs can be easily defined by landmarks, it lacks clear classification based on cytoarchitecture. In such studies, monkey PFcs is classified as either area 9 [Vogt *et al.*, 2005], 32 [Petrides and Pandya, 1994], 9/32 [Paxinos *et al.*, 2000], or 24b [Carmichael and Price, 1994]. Consequently, we will place more weight on functional differences based on connectivity for our study of PFcs.

Medial PFC has strong connections with limbic areas, motor-related areas, and other

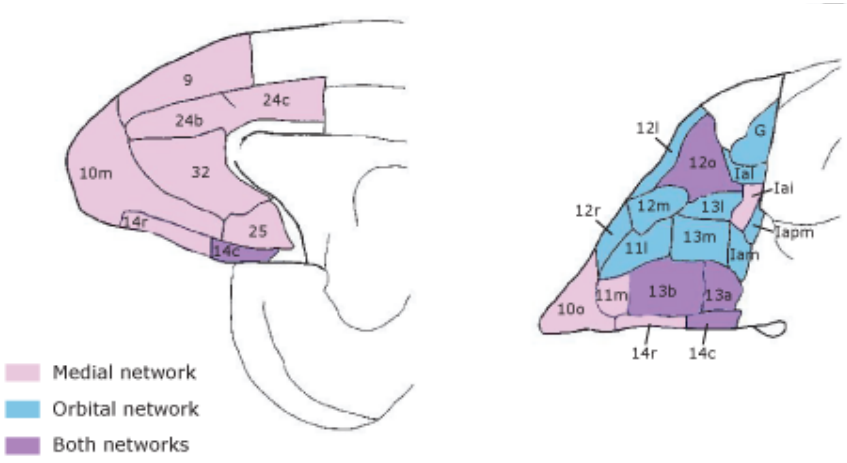

Figure 1.2: Dual networks in the monkey prefrontal cortex - "medial prefrontal network" and "orbital prefrontal network" from [Carmichael and Price, 1996].

parts of PFC. Of the limbic areas, the amygdala, a key structure for processing affective value and emotion [Adolphs *et al.*, 1994], is the prominent structure that connects with the medial most portions of areas 24a, 24b (PFcs) and 32 [Carmichael and Price, 1995a]. As for motor connectivity, PFcs strongly connects to the area immediately posterior to it - cingulate motor area (CMA) - which has direct projections to the spinal cord for arm and leg movements [Dum and Strick, 1991; Dum and Strick, 1996]. CMA is also connected with the supplementary motor area (SMA), and together their spinal projections make up 40% of all corticospinal projections in the frontal lobe [Dum and Strick, 1991]. Aside from controlling muscle movements, medial PFC processes motor-related information. Area 24 shares similar connectivity patterns as premotor cortex (area 6). Both have inputs from sensory association areas, but not from primary auditory or visual cortices. They are interconnected: area 24c connects with the ventral portion of the premotor cortex (areas 6va and 6vb) [Carmichael and Price, 1995b]. Lastly, both are interconnected with lateral PFC. Area 24 does differ from the premotor cortex in how balanced its interconnectivity with lateral PFC is though. Area 24 sends stronger inputs to lateral PFC than lateral PFC sends back [Passingham, 1995]. Aside from lateral PFC, medial PFC forms a "medial prefrontal network" within itself and a few subregions of orbital PFC (Figure 1.2). Its complement, the "orbital prefrontal network" mostly interconnects orbital regions and has few connections with medial areas. The two networks converge at areas 12o, 13a, 13b and 14c [Carmichael and Price, 1996]. Given the numerous connections between areas in medial PFC, limbic areas, and motor areas, the "medial prefrontal network" is also referred to as the "visceromotor network".

1.2.2 Anatomy of lateral prefrontal cortex

Lateral PFC covers the lateral surface of the anterior side of the brain. In monkeys and humans, lateral PFC has several subregions: areas 9, 9/46, 46 at the dorsal end and areas 44, 45, 47/12 at the ventral end. These areas are flanked by the frontopolar cortex (area 10) at the anterior extent and the frontal eye fields (area 8) at the posterior extent [Petrides and Pandya, 1999]. The tissue is largely granular, and more subtle details in tissue lamination have been used to isolate out the specific areas. In rodents, it is difficult to locate a homologous lateral region, since their PFC is strictly agranular. However, based on structure and function, Uylings and colleagues claim portions of rat PFC, particularly the dorsomedial shoulder region, parallel the dorsolateral portion of primate lateral PFC [Uylings *et al.*, 2003]. This thesis focuses on that dorsolateral analog in monkey: area 9/46 surrounding the principal sulcus. We abbreviate it as PFdl. The equivalent region in humans is the middle frontal gyrus [Petrides and Pandya, 1999].

In terms of connectivity, PFdl receives inputs from sensory and motor areas. Polysensory information enters from the superior temporal sulcus [Passingham, 1995]. Visual information from inferotemporal cortex (IT) and V5 (also known as MT) provide it high-level visual signals related to object recognition and location, respectively. PFdl also gets additional projections carrying abstract visual information from posterior parietal cortex, particularly area 7a, 7b and the medial portion of area 7. Projections from IT enter ventrolateral PFC, while parietal projections enter PFdl [Petrides and Pandya, 1984]. Based on diffusion-weighted imaging tractography, a sizeable portion of the human parietal cortex also connects with PFdl [Croxson *et al.*, 2005]. This implies that lateral PFC can represent complex and integrated stimuli information. As for motor connectivity, PFdl is interconnected with premotor cortex and frontal and supplementary eye fields. These motor areas contain high-level motor representations rather than low-level signals controlling muscle movement. The aforementioned connections with areas in the parietal cortex further include regions that integrate body position with visual space [Mountcastle *et al.*, 1975]. Together with the sensory inputs, PFdl has rich connections that can integrate sensory and motor information.

In contrast, PFdl has little limbic connectivity. It has some reciprocal connections with the hippocampus [Goldman-Rakic *et al.*, 1984], which gives it the ability to process memory-related information. Much of the limbic connectivity it gets instead derives from other regions in PFC. It interconnects with medial prefrontal area 32 [Passingham, 1995] and non-motor regions in area 24 [Bates and Goldman-Rakic, 1993]. Both of these areas connect directly to limbic areas. Moreover, there is directionality in the connectivity between area 24 and PFdl: more projections go from area 24 to PFdl than the other way around. This asymmetry suggests an order in information processing that starts in medial PFC (i.e., area 24) and then goes to PFdl. PFdl also gets limbic information secondhand from its interconnections with orbital prefrontal areas 13, 14, and particularly area 11 [Petrides and Pandya, 1999]. Collectively, these anatomical connections suggest that PFdl gets highly

processed information related to memory and motivation rather than direct information from the subcortical structures.

1.2.3 Anatomy of orbital prefrontal cortex

Orbital PFC occupies the ventral surface of PFC, right above the orbits, the eye cavities in the skull. It includes Brodmann areas 10 through 14 in primates. The cytoarchitecture varies along the anterior-posterior axis. The anterior end consists of granular tissue, the middle dysgranular tissue and the posterior agranular tissue [Morecraft *et al.*, 1992]. Human orbital PFC has similar cellular pattern [Mackey and Petrides, 2010]. In contrast, rat orbital PFC lacks clear homologous Brodmann areas. Its subregions are ventrolateral PFo, lateral PFo, and agranular insular cortex [Öngür and Price, 2000]. We refer to orbital PFC as PFo. PFo has few connections with motor regions. Ventral premotor cortex connects to areas 12m and 13l, as well as to the posterior end of medial PFC, area 24c. Area 24c in turn connects to ventromedial PFC at area Ial. Eye-related information from the supplementary eye field enters area 12, though at different subregions: 12l, 12o, and 12r [Carmichael and Price, 1995b]. Finally, PFo can also indirectly affect motor behavior through its interconnections with lateral PFC. These connections suggest a limited role of PFo in motor processing and execution.

PFo has an array of connections with sensory areas. All senses converge in PFo. Inferotemporal cortex (IT) provides visual information mostly to area 12l. Regions in the superior temporal cortex (i.e., TPg, STGa and STGp) send polysensory information to PFo. Secondary somatosensory cortex and anterior infraparietal area (AIP) send tactile information about the face, hand, and forelimb mostly to areas 12m, 13m and 13l. Meanwhile, insula and opercular cortex provide taste information, while pyriform cortex provides olfactory information. Most of these inputs enter in lateral PFo. More medial regions have fewer sensory afferents [Passingham, 1995; Carmichael and Price, 1995b]. Subtle regional differences aside, PFo in general has the ability to process a rich array of sensory signals.Adding to the sensory afferents, PFo has a plethora of connections carrying reward-related and emotional signals. It has far more limbic connectivity compared to either medial or lateral PFC. All limbic structures (amygdala; hippocampus; temporal pole; entorhinal, perirhinal, parahippocampal, and cingulate cortices) connect to central orbital areas 11m, 13a, 13b, and medial orbital areas 14c and 14r [Carmichael and Price, 1995a]. The amygdala, a key structure for processing value information [Murray and Izquierdo, 2007], has most abundant connections with medial PFo across both monkey and human [Croxson *et al.*, 2005]. The lateral portion of PFo lacks such diverse limbic inputs and is limited to having mostly projections from amygdala, temporal pole, and cingulate cortex [Carmichael and Price, 1995a]. Despite the differences within PFo, the overall region has diverse limbic information.

1.2.4 Closing remarks on anatomy

PFC can be divided into three general regions: medial, lateral and orbital [Fuster, 1997], each of which has its own specialization. The medial portion may be particularly important for encoding action-outcome (AO) information, as it receives robust limbic inputs and connects with motor areas. Lateral PFC also connects strongly with motor areas, but has fewer limbic connections compared to medial PFC. The information it receives is abstract, which suggests lateral PFC has a higher-order cognitive role than medial PFC. The first experiment in this thesis will assess whether medial PFC is preferentially involved in encoding AO associations relative to lateral PFC. The second experiment will contrast medial PFC with orbital PFC. Both areas have plenty of limbic connections. What differentiates the areas is their connectivity with sensory and motor areas. Medial PFC connects to various motor areas and the spinal cord, whereas orbital PFC has limited motor connections. Rather orbital PFC receives information from all sensory modalities. Given these striking anatomical differences, the second experiment will investigate the extent to which orbital PFC encodes SO associations and medial PFC encodes AO associations. We now turn to some of the theory that underpins how such associations may be learned.

1.3 Theoretical considerations underlying credit assignment

Taking from the field of artificial intelligence, the credit assignment problem addresses how specific outcomes get correctly paired with their causes [Sutton and Barto, 1998]. The credit assignment problem is an important problem to resolve, because it provides an agent with the understanding of how the world operates so as to better behave in it. Suppose you teach your robot or dog a trick, it must first realize that it is to learn something new. Then it must sort through all the gestures you make and all the actions it tries to determine which components are important, i.e., may be potential causes of the desired outcome. From the important components, it has to deduce the underlying temporal structure so as to complete its learning. This problem becomes particularly difficult if the agent is in a dynamic environment. The amount of information to track can grow dramatically, and the agent becomes more prone to falsely evaluating a random component as important. Additionally, the amount of information may exhaust memory capacity. The credit assignment problem is a general framework. From the example, you may have noticed that causes can be separated into those external to the agent (stimuli) and those internal to the agent (actions). Also, the agent could easily be a machine (robot) or animal (dog). Given such flexibility, the credit assignment problem can be applied to biological contexts. In our case, the credit assignment problem becomes a study of how the agent – the brain area, PFC – combines stimuli, actions, and outcomes into meaningful relationships via neural mechanisms. In this form, the credit assignment problem becomes similar to a division of psychology called animal learning theory [Pearce, 1997].

In animal learning theory, learning that one event predicts another is termed associative learning. There are two methods of evoking this type of learning: pavlovian conditioning and instrumental conditioning. Pavlovian conditioning is used to associate a neutral stimulus such as the Starbucks logo with a motivationally relevant stimulus such as a cup of freshly ground coffee. The neutral stimulus is the one being conditioned; hence it is referred to as the conditioned stimulus (CS), whereas the motivating stimulus is the unconditioned stimulus (US). The US follows with a reflexive action, such as approaching the stimulus, Starbucks in the example. After much iteration over time, this action becomes conditioned and linked with the CS. At that point, merely presenting the CS automatically prompts execution of the action, formally known as the conditioned response (CR). In the example, approaching the stimulus Starbucks is also referred to as a pavlovian approach response. The approach response is an automatic event following the CS. Studies have found that subjects with PFo lesions exhibit this approach response despite motivational changes to the US. These impaired subjects will continue to touch pictures on a touch screen previously associated with reward even after the reward has been removed and dissociated from the picture [Rolls *et al.*, 1994; Pears *et al.*, 2003]. These results suggest that PFo may be necessary for monitoring these pavlovian associations. In contrast, instrumental conditioning handles voluntary actions. It is used to associate an action an animal willfully makes such as rolling over with its outcome such as a dog treat. An SO association can be thought of as the learnt product of pavlovian conditioning, whereas an AO association is the learnt product of instrumental conditioning.

SO and AO associations can be combined. Often a stimulus precedes an AO association, and the fully associated sequence becomes SAO. Returning to the example, a dog may learn that rolling over only upon seeing a hand gesture yields a treat. This process of learning the SAO association involves a pavlovian-instrumental interaction that can be broken down into a couple components. First, the pavlovian CS (hand gesture predictive of reward) motivates the animal to respond. This response (approaching) can be slowly shaped into an instrumentally conditioned action (rolling over). Second, the pavlovian CS serves as a response-cueing signal. Only when the dog sees that particular hand gesture does he act in hopes of reward. After much rehearsal of the SAO association, the dog automatically rolls over upon seeing the hand gesture without a need for the treat. At this point, his behavior has become habitual and governed by the SA association (more commonly referred to as a stimulus-response [SR] association). Neural signatures of this transition have been found: lateral PFC neurons encode SR associations more strongly as the associations become more learned [Pasupathy and Miller, 2005]. The transition has also been modeled as an outcome-sensitive system (that uses value iteration in a Bayesian tree-search) switching behavioral control over to a habit-like system (that uses Bayesian Q-learning) [Daw *et al.*, 2005]. As a generalization, SA associations can be thought of as rules, how one behaves within a specific context. Rule learning, in addition to SO and AO associations, plays an important role in prefrontal research. They are incorporated into many experimental paradigms used to

investigate prefrontal function in rodents [Balleine and Dickinson, 1998; Pickens *et al.*, 2003; Ostlund and Balleine, 2007], monkeys [Pears *et al.*, 2003; Wallis and Miller, 2003b; Rudebeck *et al.*, 2008], and humans [O'Doherty *et al.*, 2003; Valentin *et al.*, 2007].

Neural representations of AO and SO contingencies could follow the anatomical limitations of the prefrontal brain areas. Since medial PFC connects with motor and limbic areas, it can potentially encode the conjunction between action and outcome. Although lateral PFC has fewer limbic connections and equally strong motor connections, it receives strong projections from medial PFC. Therefore, we might still see it encode AO associations. Differentiating the two brain areas' ability to encode AO associations will be the objective of the first experiment in this thesis. Orbital PFC does not have many connections with motor areas. Hence, it may not encode AO associations. Instead, given its numerable sensory and limbic connections, it may encode associations between stimuli and outcomes. The second experiment will address this hypothesis.In order to learn associations, the agent must track all potentially useful pieces of information. A common method of retaining those pieces of information would be to place them in memory. However, most of the information becomes useless after the agent discovers the exact pieces necessary for learning an association. Hence, it is energy-efficient to keep all of those pieces of information in a short-term memory buffer rather than storing them in more permanent ways. In the brain, this memory buffer is called working memory.

This ability to hold and manipulate information in mind across short periods of time is not well understood, despite its relatively long history. During the 1930s, Carlyle Jacobsen observed that a lesion in PFC limited a subject's memory. Jacobsen had chimpanzees perform a delayed response learning task, where the subject saw two bowls. One bowl contained food, the other nothing. After seeing which bowl had food, lids were placed on top of both bowls. Then a delay was enforced, after which the subject could select the bowl it preferred. Animals with lesions could not remember the bowls even when delays were one second long [Jacobsen and Nissen, 1936]. The neuronal signature for this process was not observed until forty years thereafter when Fuster observed sustained firing in PFC neurons after removing a stimulus from the environment [Fuster, 1973]. We now know that many brain areas display working memory capacity: PFC, parietal cortex [Jonides *et al.*, 1993], anterior cingulate cortex [Jonides *et al.*, 1993], and basal ganglia [McNab and Klingberg, 2008]. Within PFC, the details on how different regions maintain sustained activity instantiating working memory remains unclear. One theory says that different areas encode different types of information in working memory based on their anatomical connections [Goldman-Rakic, 1987]. Another says that prefrontal neurons adapt to task demands [Miller and Cohen, 2001] and hence may exhibit working memory if needed in the task. Regardless, of all prefrontal areas, PFdl is most frequently studied for its robust working memory capabilities [Funahashi *et al.*, 1993; Miller *et al.*, 1996; Romo *et al.*, 1999; Rowe, 2000]. In short, both PFdl and PFcs, which is part of the anterior cingulate cortex, have been shown to encode information in working memory. However, PFdl may have a stronger capacity to do so, given the numerous studies

regarding its role in working memory studies. Hence, to measure the extent to which PFdl can encode AO associations, we should characterize its ability to encode AO associations in working memory. This is what we do in the first experiment (Chapter 3).

1.4 Functional properties of prefrontal regions

PFC is regarded as the center for executive control. It has functions related to organizing and planning, motivation and attention, rationalizing and processing emotions, as well as personality. In the realm of planning, PFC tracks goals and helps determine what potentially appropriate actions there are for achieving each goal. It promotes the appropriate one, while inhibiting the less appropriate options. It also partakes in rule acquisition, which can be thought of as the acquisition of SA associations [White and Wise, 1999], as well as the more abstract forms: category learning [Freedman *et al.*, 2001; Antzoulatos and Miller, 2011] and language acquisition [Opitz and Friederici, 2003]. Beyond affecting actions, PFC can control voluntary attention in scenarios, where you need to seek out a particular stimulus (like your run away puppy in a crowd) or ignore distracting elements (like a fender-bender while driving in a new city). PFC also keeps emotions in check and affects moral judgment [Koenigs *et al.*, 2007]. It is further implicated for fluid reasoning, the ability to think logically in novel situations. Additionally, information can be stored in PFC across short periods of time in working memory for manipulation [Funahashi *et al.*, 1993; Rainer *et al.*, 1998; Constantinidis *et al.*, 2001]. Subregions specialize in different aspects of the aforementioned diversity of functions. In the following sections, these specializations within PFcs, PFdl, and PFo will be described.

1.4.1 Properties of medial prefrontal cortex (PFcs)

Consistent with its anatomical connections to motor and limbic areas, PFcs carries both action and outcome information. Actions and their predicted outcomes appear to be encoded by separate but intermingled populations of neurons [Hayden and Platt, 2010]. PFcs not only integrates those pieces of information, but monitors changes to them. Neurons have higher activity monitoring outcomes of executed actions when monkeys are learning by trial and error which action yields reward [Matsumoto *et al.*, 2003]. Other studies have found that PFcs signals erroneous actions [Carter *et al.*, 1998], probabilistic or unexpected changes to the reward [Ito *et al.*, 2003], and uncertainty in the environment affecting choice [Behrens *et al.*, 2007]. Lesions to PFcs render rats [Balleine and Dickinson, 1998; Dias and Aggleton, 2000] and monkeys [Hadland *et al.*, 2003] incapable of adjusting their actions based on changes to the actions' associated outcomes. Building upon this role in monitoring and adjusting behaviors based on AO associations, PFcs is also known to encode prediction errors of action values (e.g., signals capturing the difference in the reward the subject actually received and

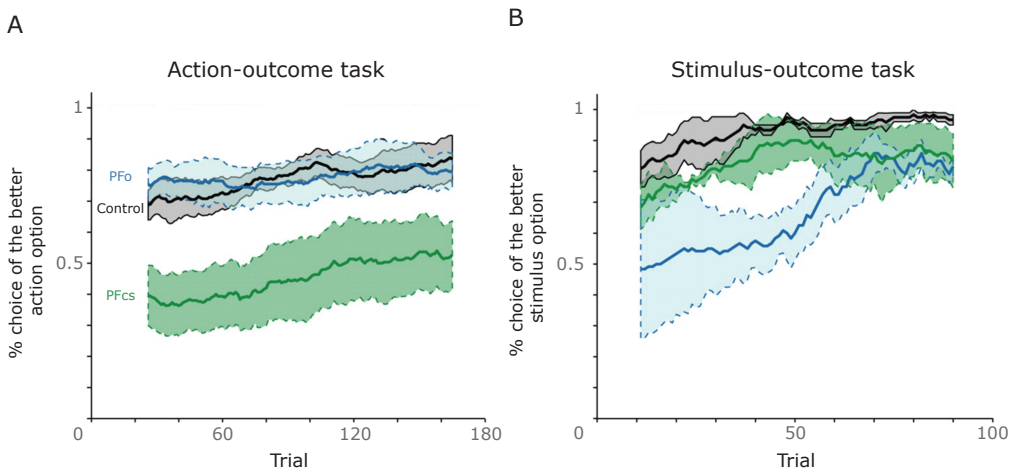

Figure 1.3: Accuracy in selecting the better action or stimulus that yielded more reward. **A**, Average number of trials in which the monkeys selected the action associated with more reward. Monkeys with PFo lesions (blue line) performed as well as controls (black line), whereas the monkeys with PFcs lesions (green line) were impaired. **B**, In the stimulus-outcome task, monkeys with PFo lesions were impaired, but those with PFcs lesions were not. From [Rudebeck *et al.*, 2008].

the reward the subject thought the action would yield him) in monkeys [Seo and Lee, 2007; Matsumoto *et al.*, 2007] and humans [Ribas-Fernandes *et al.*, 2011].

In regards to SO associations, PFcs does have the ability to process information related to stimuli, given its inputs from sensory association areas. It can encode the value of visual cues that predict different magnitudes of reward [Kennerley *et al.*, 2009; Kennerley and Wallis, 2009], the likelihood of receiving a fixed amount of reward [Knutson *et al.*, 2005; Kennerley *et al.*, 2009], or the physical effort needed to obtain the reward [Walton *et al.*, 2002; Kennerley *et al.*, 2009; Croxson *et al.*, 2009]. But compared to PFo, PFcs is slower to encode rewards that are instructed by a visual cue [Kennerley and Wallis, 2009]. As for monitoring changes to SO contingencies, PFcs may not be necessary. As shown in fig. 1.3, monkeys with PFcs lesions can still monitor changes in SO associations. Specifically, if they are presented with different pictures, each of which has a different likelihood of receiving reward, they can track changing SO associations and select the picture associated with more reward [Rudebeck *et al.*, 2008]. To date, it remains unknown whether PFcs neurons are involved in choice tasks that rely solely on SO associations.These studies collectively implicate PFcs in having a strong role in assigning value to actions and tracking changes in the value. Its role in representing the value of stimuli may be minimal, based on lesion studies, but remains largely inconclusive, based on electrophysiological studies. Hence, Chapter 4 of this thesis will characterize the neuronal mechanisms related to PFcs' role in encoding SO associations

and contrast that with its encoding of AO associations. We expect stronger encoding of AO associations.

1.4.2 Properties of dorsolateral prefrontal cortex (PFdl)

PFdl processes high-level cognitive information given its connections with associative sensory and motor areas and general lack of limbic afferents. Studies suggest that it keeps track of rules in the form of stimulus action (SA) associations. PFdl neurons show one activity pattern to a stimulus when it instructs a particular action and a different activity pattern to the same stimulus when it instructs a different action [Asaad *et al.*, 2000; Wallis *et al.*, 2001; Wallis and Miller, 2003a]. When a monkey is performing a task that requires both SA and AO associations, PFcs neurons tend to encode the AO association while PFdl encodes the SA association [Matsumoto *et al.*, 2003]. Human studies suggest less clear distinctions between the two brain areas. Both areas have been shown to activate when subjects needed to suppress prepotent SA associations in a Stroop task, i.e., suppress reading aloud the word that was the name of a color and rather say aloud the font color of the word [Duncan and Owen, 2000]. Yet others find that only PFdl is involved in this suppression during a Stroop task [MacDonald *et al.*, 2000].Studies, however, tend to agree that one of PFdl's distinguishing feature is its ability to store and manipulate these rules in working memory. Neurons sustain activity following a stimulus even after it has been extinguished [Miller *et al.*, 1996; Asaad *et al.*, 2000; Constantinidis *et al.*, 2001]. PFdl neurons also maintain delay activity with integrated information about the reward value of the stimulus and its corresponding action [Leon and Shadlen, 1999; Wallis and Miller, 2003b]. Aside from just maintaining information over a short period of time, information in working memory can be computed within a population of neurons. For example, in a task, where monkeys see a visual cue which instructs a delayed response of either a saccade to the target or an anti-saccade to a target 180 degrees away, many PFdl neurons encode the location of the visual cue across the delay. On anti-saccade trials, those neurons subsequently suppress their activity for the cue location during the delay, while other neurons encoding the correct action increase their activity [Funahashi *et al.*, 1989].With PFdl's preferential role maintaining representations in working memory, the contribution of PFdl in decision-making may be better characterized when information related to the valuation of choice options needs to be held in working memory. Most working memory studies probing PFdl, however, lack a decision-making component. Tasks do not require determining the motivational value of either stimuli or actions. Take Jacobsen's delayed response task where the monkey needed to remember over a short period of time which covered bowl contained a food reward. This task does not require remembering either a specific SO or AO association, because the qualities of the food outcome are unnecessary for selecting the better bowl. The animal merely needs to remember the rule of the task and keep track of the bowl with more stuff, which is a simple perceptual discrimination. Other studies lacking associative learning include the n-back

memory task. In this task, subjects must recall the piece of information they had already experienced that was the nth piece counting backwards in time. The motivational value of the pieces of information is irrelevant. Simpler memory recall tasks, i.e., remembering stimuli that had been shown briefly, also require no value-based decision-making. All these tasks activate PFdl, yet do not provide us with insight into PFdl's role in evaluating outcomes of stimuli or actions.

Similarly, value-based decision-making tasks often do not require working memory components. Most studies keep contingencies between choice options and reward outcomes the same across long periods of time. Matsumoto and colleagues had monkeys perform lever movements guided by AO contingencies that were held constant across blocks of many trials. They found that PFcs and not PFdl neurons tended to carry information about AO associations [Matsumoto *et al.*, 2003]. By limiting the variability, subjects could rely on long-term memory processes rather than working memory processes that strongly engage PFdl [Funahashi *et al.*, 1989; Miller *et al.*, 1996; Fuster, 1997], more so than PFcs [MacDonald *et al.*, 2000]. Consequently, PFdl may not appear to encode AO contingencies. To ensure memory demands rely on working memory, our first experiment in Chapter 3 will require AO associations change trial to trial. Changing AO contingencies often should require that contingencies to be stored in working memory, thereby involving PFdl. We hypothesize that PFdl will store strong AO representations in working memory. Consequently, its AO representations, in contrast to past studies, may be more similar to that of PFcs.

1.4.3 Properties of orbitofrontal cortex (PFo)

Value processing is a central function of PFo, owning to its rich projections from the limbic system. Value appears to reflect computations regarding specific rewards rather than a generalized level of motivation [Baxter *et al.*, 2000; Roesch and Olson, 2004]. Specifically, neurons can encode the value of choice options, as well as the value of the chosen option [Padoa-Schioppa and Assad, 2006]. PFo carries context information, which gives rise to its ability to represent relative values of outcomes [Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2008], as well as hypothetical outcomes [Kang *et al.*, 2011; Abe and Lee, 2011]. Similar conclusions hold in humans and often with an additional emotional component. For instance, control subjects feel regret to negative hypothetical outcomes. In other words, controls become upset over their decision even when they win money if they discover that they would have won even more money had they chosen the alternative option. Patients with PFo damage, on the other hand, do not feel the emotional effect of the hypothetical option and only feel the immediate valence of their choice (whether or not they won or lost money) [Camille *et al.*, 2004]. Given that PFo also has sensory inputs from all modalities, in addition to limbic inputs, it has the infrastructure for processing SO relationships. It is not necessary for learning the simple association between a stimulus and an outcome like a novel light predicting food [Pickens *et al.*, 2003]. Rather it

is important for updating the value of stimuli. Monkeys with PFo lesions can still discriminate between pictures, as well as select between familiar foods presented in front of them. They are instead impaired at selecting among two or more options the stimulus yielding the most reward when the rewarded option changes between the options [Izquierdo *et al.*, 2004; Rudebeck *et al.*, 2008]. PFo lesions produce similar value updating impairments in rodents [Pickens *et al.*, 2003] and humans [Rolls *et al.*, 1994]. PFo can also monitor value across time [Sul *et al.*, 2010], thereby enabling it to associate a stimulus with the expected value of a probabilistic reward [Kennerley *et al.*, 2009]. Collectively, these studies emphasize PFo's key role in gauging changes to various types of SO associations.With its scant anatomical connections to motor regions, PFo has a minor role in tracking AO associations. To begin, many primate studies find almost zero action encoding [Padoa-Schioppa and Assad, 2006] or only a moderate amount compared to other types of encoding or to other brain areas [Wallis and Miller, 2003b]. Lesions in rodent and primate PFo do not produce any impairment in instrumental conditioning. Subjects with such lesions will stop executing motor actions if the associated rewards are made less desirable, e.g., food reward is spoiled [Ostlund and Balleine, 2007]. Furthermore, monkeys with PFo lesions can track how valuable several actions are even when the values change across time. They are quick to find the action that yields the most reward and switch to a different action when the values have changed to favor the other action. See figure 1.3 for illustration [Rudebeck *et al.*, 2008]. Human studies also report PFo encoding when subjects choose between different stimuli of different values regardless of what actions were used to indicate choice [Arana *et al.*, 2003; Valentin *et al.*, 2007]. Thus there is consistency across species that PFo selects for SO pairings and not AO pairings.

In short, PFo plays a major role in monitoring outcomes as they relate to stimuli. It has far less importance encoding the value of actions. Hence, we hypothesize that PFo will be engaged while subjects rely on SO associations to guide their decisions. We will test this speculation and describe the results in Chapter 4.

1.4.4 Closing remarks on functionality

The introduction has set up two contrasts between pairs of prefrontal areas. Firstly, PFcs and PFdl both have rich connections with motor regions in the brain and therefore may readily encode values of actions. Indeed PFcs studies have found that area to represent AO associations. In contrast, many PFdl studies implicate PFdl in representing information in working memory. To more accurately measure PFdl's involvement in encoding AO associations, we need to involve working memory. Hence, our first experiment will compare the encoding of AO associations in working memory across PFcs and PFdl. We hypothesize that PFcs will be strongly selective of AO associations and consistent with past studies. Furthermore, we anticipate that PFdl will be selective for AO associations particularly during portions of the task requiring memory storage.Our second contrast will be between PFcs and PFo.

Both areas have strong connections with the limbic system that provides reward-related information. Hence, these areas can encode value information. Studies have found that PFcs selects for AO associations, whereas PFo selects for SO associations. However, it remains unknown what neuronal mechanisms underlie this dissociation. Our second experiment will characterize the activity from PFcs and PFo neurons while subjects perform an AO task and an SO task in alternation. We expect a majority of PFcs neurons to represent actions and outcomes and a majority of PFo neurons to represent stimuli and outcomes.

1.5 Clinical implications in addiction

Several studies have found evidence of PFC dysfunction in addicts [Volkow *et al.*, 1992; Biggins *et al.*, 1997; Bolla *et al.*, 1998; Catafau *et al.*, 1999]. There is a decrease in PFC volume in human polysubstance abusers [Liu *et al.*, 1998], while in rats, structural changes in the dendrites of PFC pyramidal neurons have been observed after extended cocaine self-administration [Robinson *et al.*, 2001]. PFo activity is associated with craving for a variety of drugs, including nicotine [Brody *et al.*, 2002], cocaine [Grant *et al.*, 1996], and heroin [Sell *et al.*, 2000]. Finally, addicts show a variety of neuropsychological disorders that they share with patients with PFC damage such as difficulty making decisions, myopia, apathy, and altered personality [Bolla *et al.*, 1998; Bolla *et al.*, 2003]. Together these studies illustrate the importance of PFC dysfunction in addictive behavior. Furthermore, addiction is often characterized as a transition from outcome-directed behavior to stimulus-driven habitual behavior, as addicts develop cravings upon seeing stimuli associated with drug-taking [Everitt and Robbins, 2005]. This dissertation aims to understand the neuronal mechanisms that might underlie the instantiation of these processes within PFC. We will focus on the outcome-directed behavior and compare specific mechanisms in PFdl, PFcs, and PFo that are used to track contingencies between actions and outcomes, and stimuli and outcomes. By so doing, we contribute to the scientific foundation that may lead to the development of novel therapeutic strategies for breaking addictions and habits.

Chapter 2

General Methods

2.1 Overview

This chapter will describe the methods used in performing all experiments in this thesis. Specific details related to each experiment will be included in the chapter describing that experiment.

2.2 Behavioral training materials and methods

2.2.1 Subjects

We used the same two male rhesus monkeys (*Macaca mulatta*) for both experiments. During neuronal recording for experiment 1: AO task contrasting PFcs and PFdl, subject H was 5 years old and weighed 7.5 kg, while subject J was 5 years old and weighed 9.8 kg. For experiment 2: AO and SO encoding across PFcs and PFo, subject H was 6 years old and 7.5 kg, while subject J was 7 years old and 12.5 kg. Subjects were housed in pairs as part of a 14-member colony living within a large room. They were fed twice a day and obtained daily enrichment. They experienced a 13-hour long light cycle that begins each day at 7 am. The subject's fluid intake was regulated so as to maintain motivation on the tasks.

2.2.2 Behavioral training

We trained the subjects to perform the behavioral tasks using positive reinforcement. They sat in front of a video monitor and chose pictures to get a juice reward. The subject continued to do this until he received as much reward as he wanted. Once the subject had learned the task, we started the recording sessions. Training was typically carried out for 5 or 6 days a week.

2.2.3 Materials and methods for training

Subjects performed tasks seated in a primate chair facing a 19-inch LCD computer screen placed 50.5 cm from the chair. A system of computers controlled the display of behavioral events (Figure 2.1). They ran the NIMH Cortex program for data acquisition and experimental control (http://www.cortex.salk.edu/) over the timing and presentation of stimuli. The central 'Cortex' computer sent commands via a COM port to a receiving computer, which then presented the visual stimuli on the LCD screen. The stimuli were mirrored via a video splitter onto a third screen in the sound-attenuating boxed room, in which the monkey sat. By mirroring the display, we could monitor exactly what the subject saw without disturbing him as he worked. The 'Cortex' computer handled compiling and running timing routines, as well as interfacing with various external devices via a PCI-DAS1602/12 data acquisition (DAQ) card (Measurement Computing, Norton, MA). Each behavioral event in the trial was marked with a code that was sent as an 8-bit number from the DAQ to the MAP system (see section 2.3). The MAP system read in this number and recorded a time stamp of when it occurred along with its value. Its time stamp was stored along with neurophysiological data in a single '.plx' data file. The 'Cortex' computer ran with a single interrupt routine that triggered every 1 millisecond and updated both a software clock and sampled all data lines. Thus, the control of the behavioral contingencies, the presentation of stimuli and the monitoring of behavioral events all took place with a 1 ms resolution. Visual stimuli in the SO task were isoluminant as measured by the Spyder luminance-meter (Datacolor, Lawrenceville, NJ).

Actions of the subjects were registered using a 4-TPS-E1 Touch Sensor Module (Crist Instrument, Damascus, MD) connected to the digital input port of the DAQ card. The touch sensor was a contact sensitive device designed to send a 5-volt TTL pulse when a grounded subject touched it. Actions were executed by the subject via custom-made joysticks that connected to the Touch Sensor Module. The joystick was constrained, such that the subject could only make left and right movements.

Eye position was recorded using an infrared eye monitoring system (ISCAN, Burlington, MA). An infrared camera focused on the subject's eye and visualized the results using proprietary image tracking software. The software tracked the center of the subject's pupil as x and y coordinates, in addition to the pupil diameter. These three measures were fed separately to three DAQ analog input channels and recorded for the duration of the session.

Juice rewards were delivered by commands from the DAQ digital output ports to the juice pump and a particular juice's corresponding solenoid valve. The ISMATEC-IPC8 peristaltic juice pump (ISMATEC SA, Glattbrug, Switzerland) took a 5-V TTL pulse that delivered fluid at 0.62 mL/s during the duration of the TTL pulse. To deliver one juice at a time, we built a custom circuit that opened a single specified solenoid valve in response to a TTL pulse from the DAQ card. The valve circuit could deliver four lines of juice, of which we used three. Those three lines ended at a custom-made mouthpiece with separate cannulae. When

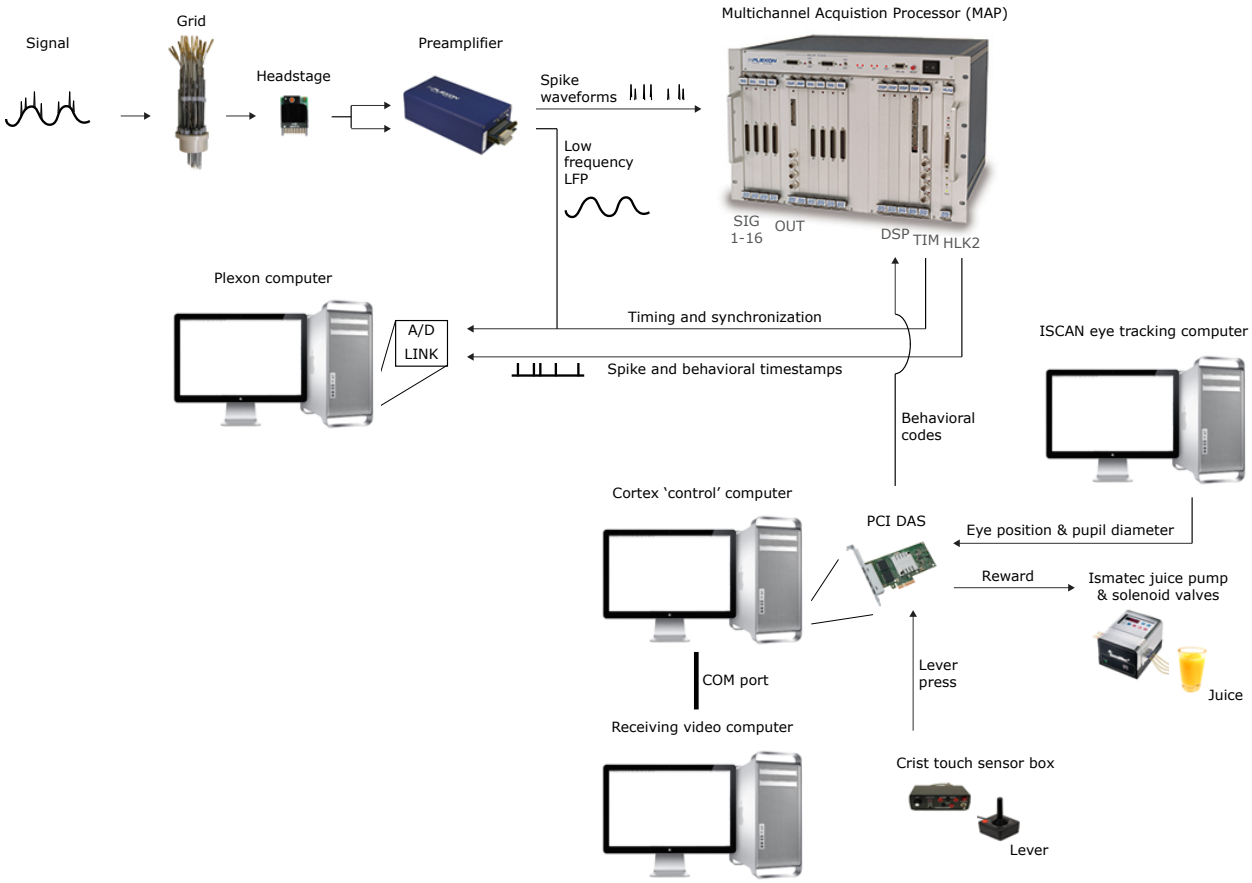

Figure 2.1: Experimental setup used to control behavioral events and record neurophysiological data

one juice line was delivering juice to the subject, the other two lines routed their respective juices back into their juice containers.

2.3 Neurophysiological techniques

2.3.1 Isolation of recording sites

To determine the exact location for performing neuronal recordings, we began by placing recording chambers over the brain areas of interest. Magnetic resonance images (MRIs) of the subjects brain were taken at the U.C. Davis Center for Imaging Sciences with a 1.5 T scanner prior to the animal's arrival at Berkeley. Those digital images were then imported into commercial graphics software (Adobe Illustrator CS5, San Jose, CA), where we calculated stereotactic coordinates for the chamber. We verified the correspondence between the MRI scans and electrode placement by mapping the position of sulci and the boundaries of gray and white matter during recording sessions. Upon completing data collection, we plotted the electrode positions and sulci and created an unfolded cortical map of our recording locations and positioning of task-selective neurons.

In experiment 1: AO task contrasting PFcs and PFdl, both subjects were each implanted with one recording chamber through which we could reach PFdl and PFcs. The chamber in subject H was over the left hemisphere centered at 29 mm anterior of the interaural line (i.e., AP 29) and 1.3 mm lateral of the mid-sagittal plane (i.e., LM 1.3) on the skull and angled at 22 degrees outward from the vertical (Figure 2.2). Subject J also had one chamber over the left hemisphere. It was centered at AP 24, LM 1.3 with a 25-degree angle tilt outward from the vertical.

Experiment 2: AO and SO encoding across PFcs and PFo required access to more brain areas. Both subjects were each implanted with two recording chambers through which we could reach PFdl and PFcs in one and PFo in the other. The PFdl-PFcs chamber in both subjects over the left hemisphere remained the same from experiment 1. The additional PFo chamber in subject H was over the right hemisphere centered at AP 31 and LM 1.9 with an angle of 23.1 degrees. In subject J the equivalent chamber was over the right hemisphere located at AP 27, LM 1.6 with a 23-degree angle tilt from the vertical. Subject J fell ill due to a possible infection during recording, and we took off his chambers to allow him to properly heal. After about half a year, we re-implanted one chamber over the right hemisphere to finish collecting neuronal data across all brain areas. The chamber was centered at AP 29, LM 1.2 with a tilt of 25 degrees.

2.3.2 Surgery

After determining the locations of chambers, we performed two operations, one to implant the chambers, then another to create the craniotomy, through which recording electrodes

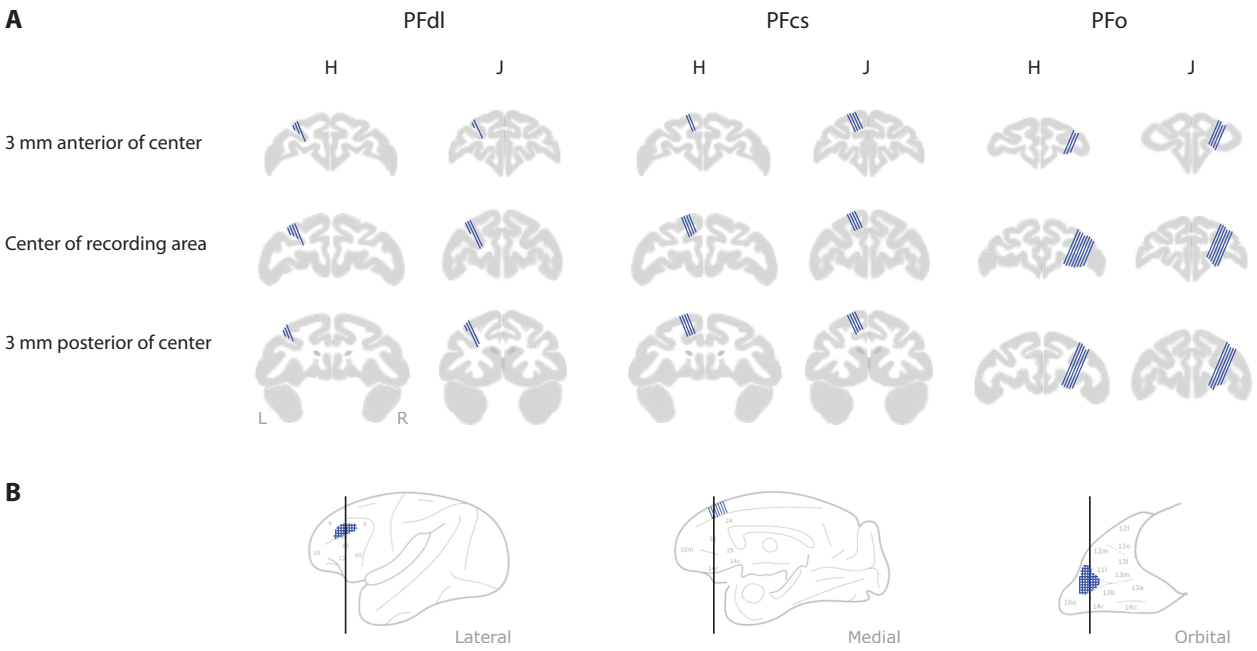

Figure 2.2: Examples of the recording sites **A**, Coronal brain slices are overlaid with electrode tracks (blue lines) from where we recorded within each brain area. The six columns are separated into three pairs of columns, one for each brain area. Within the pair of columns, MRI scans from subject H and J are on the left and right, respectively. The rows refer to three different AP locations of the brain slices with respect to the center of the recording region within each brain area. **B**, A representative view of the span of recording sites across the subjects and experiments. Black vertical lines roughly correspond to the AP position of the middle row of MRI scans in **A**. Electrodes enter at the intersection points on the blue grids for the left and right images pertaining to PFdl and PFo sites, respectively. The middle image shows electrode tracks in the same format as the above MRI scans.

enter the brain. Prior to those surgeries, each subject underwent a surgery to implant a custom-made titanium head-positioning post secured with titanium orthopedic screws. This post kept the subject's head immobile such that we could track the subject's eye position and have electrodes in the brain during neuronal recording. The following surgery was to implant cylindrical titanium recording chambers on to the skull above the brain area of interest. The chambers were secured via bone cement and titanium screws. The final surgery was to create craniotomies within the chambers, thus enabling access to the brain. Chambers were covered with a cap (made of either polypropylene CILUX from Crist Instrument, Hagerstown, MD or polyetherimide Ultem®, custom-built that tightens via screws. Before surgery, we anesthetized the animals with ketamine (10mg/kg IM). Xylocaine or lidocaine spray (14%) or ointment was used as a local anesthetic to facilitate intubation. During surgery, the surgical team monitored the animal closely, to ensure appropriate level of analgesia. Anesthesia was maintained with isoflurane (2-4%). The level was monitored by heartbeat (90-150 beats/min), respiration rate (17-23 breaths/min), temperature ($36-39°C$) and absence of responses to stimuli such as toe pinch. Blood oxygen saturation was also monitored with a pulse-oxymeter. A lactated ringer IV drip (2-4 ml/kg/hr) was used during surgery to ensure that the animal received sufficient hydration, and a heating pad and towels were used to maintain body temperature. Following surgery, gas anesthesia was discontinued, and once the animal showed signs of coming out of anesthesia, the intubation tube was removed and buprenorphine (0.01-0.03 mg/kg SC or IM) or morphine (1-2 mg/kg SC) were injected for relief of post-operative pain. Immediately after surgery the animal was checked at least every half hour by members of the lab and veterinary staff, and the time between checks was gradually lengthened as the animal recovered from anesthesia. After initial recovery, the animal was checked several times per day at which time appropriate analgesics and antibiotics were administered. Typically, buprenorphine (0.01-0.03 mg/kg SC or IM) or morphine (1-2 mg/kg SC) was administered 2-3 times per day for 3-5 days. In general, all appropriate measures were taken to minimize postoperative pain. To ensure adequate control of post-operative pain, choice of analgesia and frequency of administration were made in consultation with veterinary staff. All procedures were in accord with the National Institute of Health guidelines and the recommendations of the U.C. Berkeley Animal Care and Use Committee.

2.3.3 Recordings

We recorded neuronal activity from three brain areas: PFcs, the dorsal bank of the cingulate sulcus anterior of the corpus callosum; PFdl, the dorsal bank of the principal sulcus; and PFo, areas 11-13, mostly between the medial and lateral orbital sulci. Activity from these areas were simultaneously recorded by lowering tungsten electrodes (FHC, Bowdoin, ME) attached to custom-designed screw microdrives that independently moved pairs of electrodes. The microdrives were mounted to a custom-made plastic grid containing an array of 24-gauge

holes spaced 1 mm apart. Stainless steel 24-gauge thin-wall hypodermic needles (Terumo, Somerset, NJ) were glued to the bottom of the grid, such that the beveled tip of the needles pointed out of the grid. These tips served to puncture the dura and guide the electrodes to the desired recording location. Electrodes were lowered manually by handheld screwdriver to depths predetermined from the MRI scans. When an electrode approached a cell layer, we slowed the lowering and stopped when we isolated single neurons. After neurons were located on most or all of the channels, we waited 1-2 hours for the brain to settle, thereby ensuring stability during the recording session, which typically lasted 1.5 hours. Neurons tended to stay the full duration of the recording session without much drift. We randomly sampled neurons to ensure fairer comparison of neuronal properties between the different brain regions. During the recording session, no changes to the channels were made, as that would distract the subject from performing the task.

We cleaned the chambers both before and after the recording session to minimize infection. The cleaning began by removing the cap or grid and sterilizing the outside of the chamber with alcohol. Next under sterile conditions, the inside of the chamber was flushed with sterile saline, then Nolvasan or povidone, an iodine antiseptic, followed again with sterile saline. The tissue was then dabbed dry with a sterile cotton swab. If cleaning was prior to recording, the plastic grid with microdrives and electrodes was fit on top of the chamber. The grid was sterilized overnight in Cidex. If cleaning was after recording, a new sterilized cap was placed on top of the chamber.

2.3.4 Materials and methods for neurophysiology

Voltage signals were taken from the tip of the electrodes with respect to the reference, the head-positioning post affixed to the subject's skull. Those neuronal signals were recorded and amplified via hardware and software from Plexon Inc. (Dallas, TX) as shown in figure 2.1. The signal was first amplified 20 times from an op-amp based circuit in the headstage, which connected directly to the electrodes. The signal was then further amplified 1000 times through a preamplifier and filtered for spikes in the 100 Hz - 8 kHz frequency band and local field potentials (LFPs) in the 1 Hz - 300 Hz frequency band. Signals next fed into a Multichannel Acquisition Processor (MAP system) for further amplification and filtering.

Spikes and LFPs were digitalized at 40 kHz with 12-bit resolution per channel. For the spikes, thresholds were set to ensure that neuronal signals were a minimum of 4 standard deviations above background noise (calculated over a 10s period immediately prior to recording). When the spike waveform crossed a manually set threshold, the program recorded the time stamp of the threshold crossing, along with the actual waveform in a 1400 μs window around the time of crossing. Waveforms that did not cross the threshold were discarded. The digitized waveforms were then sorted offline, using Offline Sorter software (Plexon Inc., Dallas, TX). This constructed 2D or 3D plots of a subset of 12 waveform features including the first three projections from principal components analysis, peak-valley difference

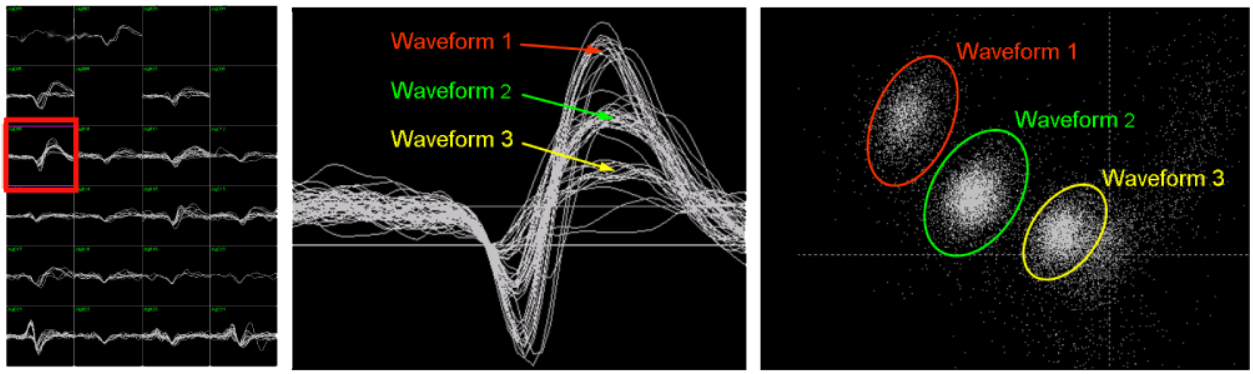

Figure 2.3: Cluster isolation of neuronal waveforms. On the left are 24 channels, of which many have neurons on them, i.e., detected waveforms. The channel highlighted in red is zoomed in in the middle panel to reveal three distinct waveforms. Those waveforms are decomposed into components in Plexon's offline sorter, and each waveform is plotted as a single dot in the right panel. Clusters of waveforms are then isolated manually and entered into neuronal analyses.

and widths, and waveform energy. From those 2D or 3D plots, clusters of waveforms were grouped together manually and classified as single units (Figure 2.3). We ensured the separation of neuronal waveforms by rejecting channels where more than 0.1% of the waveforms were separated by intervals of less than 1.5 ms or where neuronal 'drift' occurred. Typically, approximately 15% of the channels were discarded.

2.4 Statistical analysis

We used MATLAB (MathWorks, Natick, MA) to perform all analyses. Analyses were restricted to correct trials, as there were typically too few error trials to permit a meaningful analysis of activity on such trials. To characterize the selectivity of a neuron, we first calculated its mean firing rate in each trial during a defined time epoch. We compared differences in firing rate between experimental conditions with the null hypothesis that the neuron did not encode a given type of information. We specified the independent variables and statistical test in our description of the individual experiments. Once we had classified the neurons according to the type of information they encoded, we assessed differences between brain areas in the prevalence of these different types of neurons using chi-squared tests. We also characterized strength and earliest latency of encoding in selective neurons. Specific statistical analyses are described in detail in the experimental chapters.

Chapter 3

Action and outcome encoding in working memory

3.1 Introduction

Two regions of frontal cortex are particularly implicated in goal-directed behavior: dorsolateral prefrontal cortex (PFdl) and medial prefrontal cortex by the cingulate sulcus (PFcs). Both regions have strong connections with motor areas [Dum and Strick, 1993; Carmichael and Price, 1995b; Petrides and Pandya, 1999] and are activated by tasks requiring high-level cognition [Duncan and Owen, 2000]. Several studies implicate PFcs in the control of behavior via action-outcome (AO) associations. Lesions of PFcs impair the ability to use outcome information to guide behavior [Balleine and Dickinson, 1998; Hadland *et al.*, 2003; Rudebeck *et al.*, 2008] and reduce the influence of past reward history on motor selection [Kennerley *et al.*, 2006]. PFcs neurons tend to encode both rewards and the actions that led to the rewards [Matsumoto *et al.*, 2003; Williams *et al.*, 2004]. The involvement of PFdl in AO learning is less clear since studies of PFdl lesions in the primate typically focus on cognitive tasks, such as those underpinning working memory [Funahashi *et al.*, 1993; Petrides, 1995] or rule learning [Halsband and Passingham, 1982; Petrides, 1982; Gaffan *et al.*, 2002]. From an anatomical perspective, whereas PFcs strongly connects with areas processing reward, such as the amygdala, orbitofrontal cortex, and insular cortex, PFdl only weakly connects with these areas [Carmichael and Price, 1995a; Carmichael and Price, 1996; Cipolloni *et al.*, 1999].

Studies directly contrasting the neuronal properties of PFdl and PFcs showed that the latter is more important for encoding AO associations [Matsumoto *et al.*, 2003; Matsumoto *et al.*, 2007]. However, in these tasks the outcome is the presence or absence of reward. Thus, it is difficult to determine whether the neurons are encoding an association between a specific action and a specific outcome, or whether the neuronal activity is reflective of a more general cognitive process such as assessing the success or failure of an action [Brown and Braver, 2005;

Modirrousta and Fellows, 2008]. To distinguish between these possibilities, we used a task that required monitoring which of multiple possible outcomes occurred following a specific action.

In addition, tasks in which reward-related contingencies are continually changing have found strong encoding of rewards and actions in PFdl [Barraclough *et al.*, 2004; Seo *et al.*, 2007]. This is consistent with the notion that PFdl is most implicated in cognitive control when contingencies change on a trial-by-trial basis necessitating flexible, online control in the form of working memory [Miller and Cohen, 2001; Rossi *et al.*, 2007]. Thus, we trained two subjects to perform a task that required them to associate a particular outcome with a particular action on a trial-by-trial basis, while we recorded the activity of single neurons from PFcs and PFdl simultaneously.

In summary, by having multiple potential outcomes associated with an action, we aimed to test whether PFcs neurons encoded specific AO associations, and by constantly changing these contingencies on a trial-by-trial basis, we aimed to tax working memory, thereby providing a stronger test of whether PFdl neurons encode AO information. We predicted that even under these conditions, PFcs neurons would show a stronger encoding of AO information than PFdl neurons.

3.2 Materials and methods

3.2.1 Behavioral task

Each trial consisted of a sampling phase and a choice phase. During the sampling phase, the subject made two sample motor actions, each of which resulted in the delivery of a small drop of one of three juices. During the choice phase, the subject then chose to repeat one of the actions, and received a larger amount of the juice that was associated with that action earlier in the trial. Thus, to receive juices that are more preferable at the choice phase of the task, the subject had to remember which action produced which outcome during the sample phases of the task. The AO contingencies changed on a trial-by-trial basis. All possible pairings were equally likely to occur, and all juices were equally likely to be paired with either movement.

A trial began with the subject fixating a central fixation spot for 1 s (Fig. 3.1). The subject then made two sample responses by moving a lever, once to the left and once to the right, with a delay separating the two movements. The lever consisted of a 4.5-cm-long sprung joystick handle, which the subject needed to displace laterally by 1.5 cm to register a correct movement. We organized trials into blocks of 30 trials each. Depending on the block, the subject's sample actions had to consist of either a leftward movement followed by a rightward movement, or a rightward movement followed by a leftward movement. A cue presented over the fixation spot indicated to the subject the current block of trials and the

time at which to make the movement. We instructed the movements in this way to ensure that during each task epoch we had sufficient conditions of each possible AO association. A 3 s intertrial interval separated each trial.

Over the course of a trial, the subjects made one of two different types of error. First, their eye position had to remain within $\pm 2.5^\circ$ of the fixation spot until the delivery of the final reward. Otherwise they experienced a 5 s timeout (the screen turned red and the subject had to wait) before the trial resumed from the point at which they had broken fixation. Second, at the sample stages of the task, the subjects could make the wrong movement, i.e., making a rightward lever movement when they should have made a leftward movement and vice versa. If this occurred, they experienced a 5 s timeout (the screen turned yellow) and the trial resumed from the point at which the wrong action occurred. We excluded both types of trial from our statistical analysis of the neurophysiological data.

We tailored the specific juices to each subject to ensure robust, stable preferences. For subject H, we used orange juice (Minute Maid), apple juice (50% dilution, Safeway), and quinine (1.1 mM, Sigma-Aldrich). For subject J, we used orange juice (50% dilution), apple juice, and quinine (3.3 mM). We note that the quinine solution was sufficiently dilute that it was not aversive: both subjects were willing to drink large quantities (>300 ml) of a more concentrated solution of quinine (4.2 mM) from a water bottle. We tailored the quantities of juice to ensure that each subject received their daily aliquot of fluid over the course of a single session. For subject H, the sample juices lasted 0.4 s (0.25 ml) and the final reward amount was 1.1 s (0.68 ml) on 75% of the trials (small reward trials) and 2.1 s (1.3 ml) on 25% of the trials (large reward trials). For subject J, the sample juices lasted 0.3 s (0.19 ml) and the final reward amount was 1 s (0.62 ml) on 75% of the trials and 1.9 s (1.2 ml) on 25% of the trials. The subjects had no way of knowing the volume of the final reward amount until its delivery. Nevertheless, this variation in its size motivated the subject to work for more trials.

3.2.2 Statistical methods

We excluded trials in which a break fixation occurred and trials in which the subject moved the lever in the wrong direction during one of the sample responses. We constructed spike density histograms by averaging activity across the appropriate conditions using a sliding window of 100 ms. We quantified neuronal selectivity during the performance of the task using several defined time epochs. We determined these time epochs based on where the most prominent encoding of the action and the outcome was evident in the spike density histograms. The presample epoch consisted of the 800 ms immediately preceding the subject's first sample action. The first-outcome epoch consisted of an 800 ms period beginning 200 ms after the delivery of the first juice. The 200 ms offset allowed for the latency to encode juice information in frontal cortex [Lara *et al.*, 2009]. The second-outcome epoch consisted of the equivalent period after the delivery of the second juice. The prechoice epoch

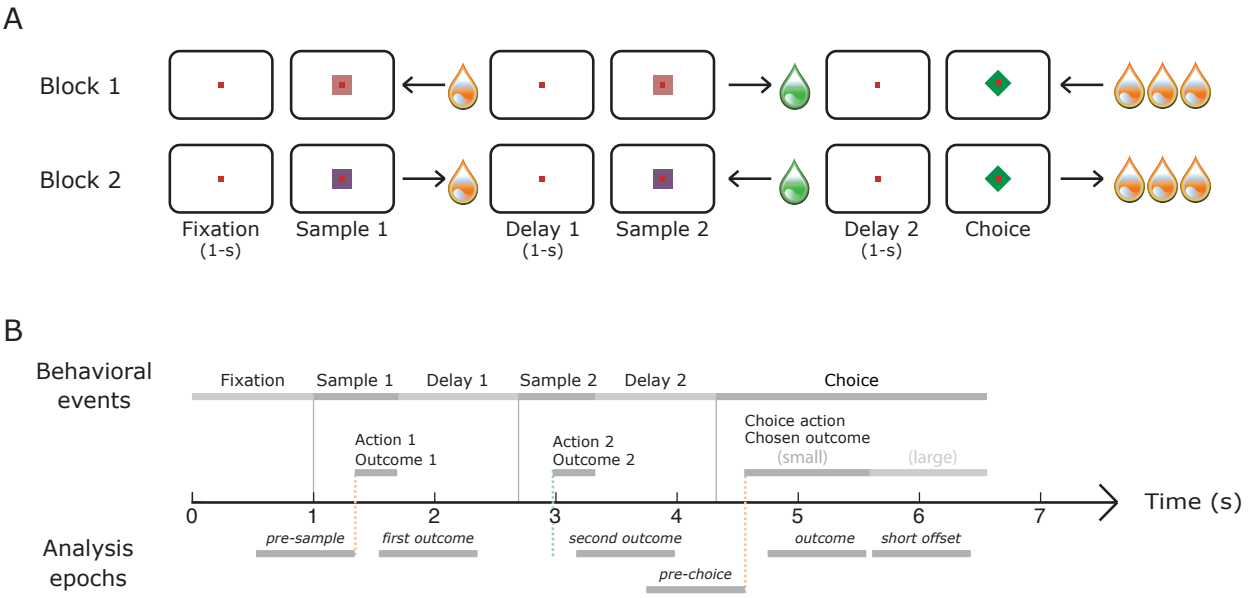

Figure 3.1: **A**, Illustration of the sequence of events in the behavioral task. The subject made two lever movements sequentially, indicated by the horizontal arrows, and received one of three juices for each movement. At the choice phase, the subject repeated one of the movements to receive the juice associated with that movement earlier in the trial. The two different blocks consist of trials in which the subject either made a leftward followed by a rightward lever movement or a rightward followed by a leftward lever movement. The block is signaled by a colored square, but this provided no information about the specific AO contingencies. The figure illustrates trials for subject H, who preferred orange juice to apple juice. In the two example trials, subject H receives orange juice for the first movement and apple juice for the second movement and elects to repeat the movement that led to orange juice at the choice phase of the task. **B**, Timeline indicating the timing of behavioral events during the course of a single trial, as well as the corresponding epochs that we used for statistical analysis. The vertical gray lines indicate the onsets of the visual cues, and the vertical colored lines indicate the time of the subjects' actions and the delivery of juice. The precise timing of these events depends on how long it takes for the subject to initiate and perform his action. For the purposes of illustration, we used the subjects' median reaction times.

consisted of the 800 ms immediately preceding the subject's final choice action. We defined two additional epochs at the choice phase of the task. The chosen epoch consisted of an 800 ms period beginning 200 ms after the delivery of the chosen juice, and the small-offset epoch consisted of an 800 ms period beginning 200 ms after the offset of the juice on the small reward trials. For each neuron, we then calculated its mean firing rate during each epoch on every trial. We used statistical tests, assessed using an α level of 0.01, to determine whether the neuron's mean firing rate depended on the various experimental factors.

During the presample epoch, we wished to determine whether neurons encoded the direction of the first sample movement. For each neuron, we performed a t test with the neuron's firing rate during the presample epoch as the dependent variable and whether the first sample action was a leftward or rightward lever movement as the independent variable. We used χ^2 tests to assess differences between PFcs and PFdl in the proportion of neurons encoding the upcoming sample movement. In addition, we calculated a direction index of the strength of encoding by subtracting the neuron's mean firing rate on trials in which the first sample action was a rightward lever movement from its firing rate on those trials in which the first sample action was a leftward lever movement. We then normalized the resulting value by dividing by the standard deviation of the neuron's mean firing rate across all trials. For example, a value of 1 on this index would indicate that the firing rate of the neuron was 1 standard deviation higher when the subject intended to make a leftward lever movement than when the subject intended to make a rightward lever movement.

To quantify selectivity during the first-outcome epoch, for each neuron we performed a two-way ANOVA with the experimental factors of action (whether the subject had just made a leftward or rightward sample movement) and outcome (which of the three juices was delivered). We then classified neurons according to whether they encoded just the action, just the outcome, or a combination of the two. We used χ^2 tests to assess differences in the proportion of selective neurons between the two brain areas. We also used the two-way ANOVA to calculate the strength of selectivity. We did this by calculating the magnitude of our statistical effects using η^2. This is equivalent to the percentage of explained variance attributable to a specific experimental factor. We calculate its value by dividing the sum of squares associated with the experimental factor by the total sum of squares in the analysis and multiplying the result by 100%. In addition, for each neuron we calculated the direction index in the same manner as for the presample epoch.

During the first-outcome epoch, we also examined the latency at which neurons first began to encode the identity of the juice, by performing a "sliding" ANOVA analysis. For each neuron, we took a 200 ms window of time, beginning 1000 ms before the delivery of the first outcome, and performed a two-way ANOVA on the neuron's mean firing rate during that window, using the factors of action and outcome. We then calculated the percentage of variance in the neuron's firing rate that was attributable to the outcome factor. We then moved the window forward by 10 ms, and repeated the analysis. We continued in this manner until we had analyzed the entire first-outcome epoch. We defined the latency of

selectivity as the time when the p value fell below 0.005 for three consecutive time bins. We then compared neuronal latencies between brain areas using Wilcoxon's rank-sum test. We chose our criterion by comparing the results from the sliding ANOVA analysis for each neuron with the selectivity evident in their spike density histograms. To verify that this criterion resulted in a reasonable rate of type I errors, we examined how many neurons would have reached the criterion in the 500 ms before the delivery of the juice (i.e., when it would have been impossible for the neurons to encode the juice's identity). Just 5/284 (1.8%) neurons reached criterion in this time, indicating that our choice of criterion indeed yielded a reasonable rate of type I errors.

To quantify selectivity during the second outcome, we first separated the trials into three sets depending on the identity of the first outcome. Then, for each neuron and each set of trials in turn, we performed a t test using the neuron's mean firing rate on each trial as the dependent variable and the identity of the second outcome as the independent variable. To examine the time course of this selectivity at the population level, we performed a sliding analysis analogous to our analysis for the first out- come, to quantify the percentage of variance in the neuron's firing rate that was attributable to the identity of the second outcome. With regard to the choice phase of the task, we analyzed neuronal activity during three epochs. First, we analyzed data during the prechoice epoch in an analogous manner to our analysis of neuronal activity in the presample epoch, although the independent variable on which we focused was the direction of the subject's choice action. Next, we analyzed data during the chosen epoch. For each neuron, we calculated its mean firing rate during both the chosen and small-offset epochs. For both epochs in turn, we then performed a three-way ANOVA, with the neuron's mean firing rate as the dependent variable, and juice (whether the chosen juice was the most preferred or intermediately preferred), magnitude (small or large reward trials) and action (whether the subject made a leftward or rightward lever movement to choose the juice) as the independent variables.

3.3 Results

3.3.1 Behavior

Both our subjects exhibited clear preferences between the three available juices throughout recording (Fig. 3.2A). Subject H preferred orange juice to apple juice, and preferred both juices to quinine. Across 17 sessions, he performed a mean of 429 $\pm$ 13 trials, and picked orange juice over apple juice 98 $\pm$ 0.6% of the time, apple juice over quinine 98 $\pm$ 0.6% of the time, and orange juice over quinine 99 $\pm$ 0.2% of the time. Subject J preferred apple juice to orange juice, and preferred both juices to quinine. Across 25 sessions, he performed a mean of 434 $\pm$ 10 trials, and picked apple juice over orange juice 96 $\pm$ 1.3% of the time, apple juice over quinine 96 $\pm$ 1.7% of the time, and orange juice over quinine 85 $\pm$ 3.2% of the

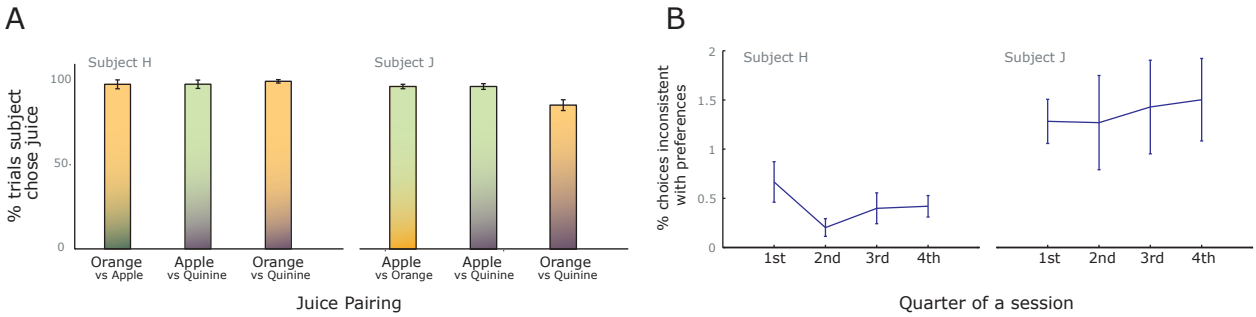

Figure 3.2: **A**, Mean percentage of trials within a session in which the subject chose in a manner consistent with his preferences for each of the three potential juice pairings. Subjects rarely made choices that were not consistent with their preferences. **B**, Mean percentage of trials during each quarter of a session in which the subject chose in a manner inconsistent with his preferences. There was no evidence that the likelihood of an inconsistent choice varied across the course of the recording session.

time. Moreover, both subjects maintained clear preferences across the course of individual recording sessions. We divided each session into quarters and, for each quarter, calculated the proportion of trials in which the subject chose the less preferred reward (Fig. 3.2B). The percentage of choices that were inconsistent with the subjects' preferences did not differ between the session quarters for either subject (Kruskal-Wallis one-way ANOVA, $p < 0.1$ for both subjects). Neither subject had a movement bias during the choice phase: both subjects made leftward movements on 51% of the trials. Thus, both subjects appeared to perform the task as we anticipated: they showed clear preferences between the outcomes, and monitored which outcome occurred following each sample movement to ensure that they consistently chose their preferred juice at the choice phase of the task.

We excluded two types of errors from our analysis of the neuronal data. The first was when the subject failed to maintain fixation for the duration of the trial. For subject H, 5.0 $\pm$ 0.6% of the trials included break fixation errors, while for subject J, 12.5 $\pm$ 1.2% of the trials included break fixation errors. The identity of the outcome did not affect the likelihood of a break fixation for either subject in either delay (one-way ANOVA, $F < 1, p > 0.1$ in all cases). The second type of error was when subjects made the wrong movement given the current block of trials. Subject H made wrong movements on 5.3 $\pm$ 0.5% of trials, while subject J made wrong movements on 8.4 $\pm$ 0.6%. Most of the wrong movements tended to occur on the first trial of a block. For subject H, 47% of his wrong movements occurred on the first trial of the block, while 38% did so for subject J. The remaining wrong movements occurred throughout the remainder of the block. The identity of the juice delivered as the first outcome did not influence the likelihood of making a wrong movement for sampling the

second outcome. For subject H, the proportion of incorrect movements made in sample 2 was 31 ± 7% following orange juice, 22 ± 7% following apple juice, and 23 ± 7% following quinine. For subject J, the proportion of incorrect movements was 38 ± 5% following apple juice, 23 ± 3% following orange juice, and 35 ± 5% following quinine. These proportions did not significantly differ from one another for either subject (Kruskal-Wallis one-way ANOVA, $p > 0.05$ for both subjects). In summary, both subjects performed the task at a high level of accuracy making errors on only a small proportion of trials.

For both subjects, the first sample action (H: median 402 ms, J: median 285 ms) was significantly slower than the second sample action (H: median 319 ms, J: median 247 ms, Wilcoxon's rank-sum test, $p < 1 \times 10^{-15}$ for both subjects). In addition, the second sample action was significantly slower than the choice action (H: median 248 ms, J: median 231 ms, Wilcoxon's rank-sum test, $p < 1 \times 10^{-15}$ for both subjects). The identity of the first outcome did not affect subsequent reaction times for the second sample action (Kruskal-Wallis one-way ANOVA, $p > 0.1$ for both subjects), nor did the identity of the second outcome affect reaction times for the choice action (Kruskal-Wallis one-way ANOVA, $p > 0.1$ for both subjects). Thus, the subjects' responses became quicker as the trial progressed, but the juices that they received did not affect their reaction times.

3.3.2 Neurophysiology

We recorded neuronal activity simultaneously from two brain areas, PFdl and PFcs, for a total of 284 neurons. We recorded from 172 PFdl neurons (H: 77, J: 95) and 112 PFcs neurons (H: 60, J: 52).

3.3.3 Comparison of action and outcome encoding across brain areas

Our initial characterization of the neuronal response properties focused on the first-outcome epoch. It is in this epoch that the subject learns which outcome is associated with the first sample action. We found that the majority of the neurons encoded either action-related information or information about the reward the subject received for making that action. To quantify this selectivity, for each neuron we performed a two-way ANOVA on the neuron's mean firing rate on each trial, using the factors of which action the subject made (the leftward or rightward lever movement) and which outcome he received (orange, apple, or quinine). Figure 3.3A illustrates a PFdl neuron that showed a significant main effect of action evident as a higher firing rate when the subject had made a rightward as opposed to the leftward action ($F_{(1,427)} > 562, p < 1 \times 10^{-15}$), but no main effect or interaction with outcome. Figure 3.3B illustrates a PFcs neuron that showed a significant main effect of outcome ($F_{(2,392)} > 61, p < 1 \times 10^{-15}$), but no main effect or interaction with action. The neuron showed its highest firing rate when the subject had just received quinine, a lower firing rate when the subject had received apple juice, and its lowest firing rate when the subject

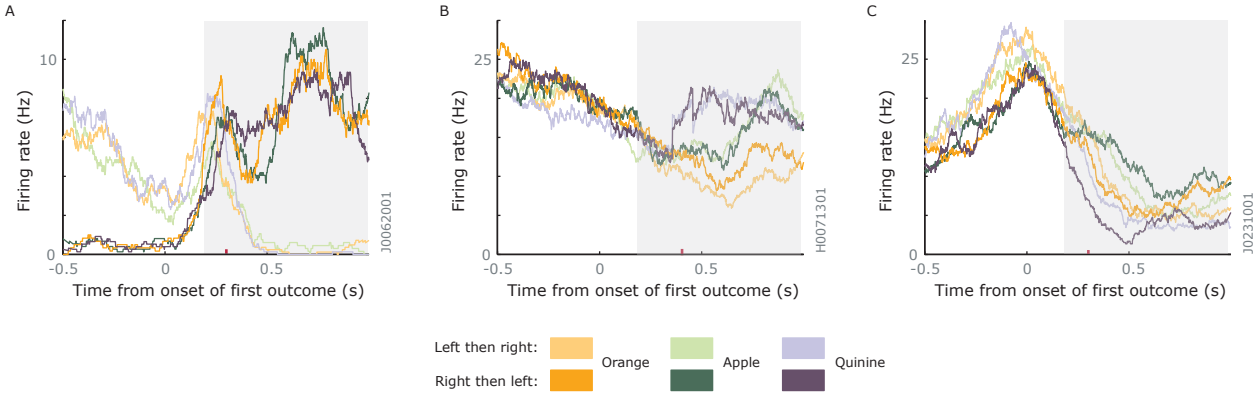

Figure 3.3: **A**, Spike density histograms illustrating a PFdl action-selective neuron. The plots are color-coded according to which movement the subject made for the first sample action and which juice he received. Time along the x-axis is plotted from the onset of the juice delivery, and the red tick mark indicates the offset of the juice. The gray shading illustrates the extent of the first-outcome epoch that we used for statistical analysis. This neuron had a significantly higher firing during the first-outcome epoch when the subject had made a leftward lever movement, but did not discriminate between the delivery of the three different juices. **B**, A PFcs neuron that had a higher firing rate to the delivery of quinine than to apple juice and a higher firing rate to apple juice than to orange juice. This neuron was from subject H, and so the neuronal firing rate inversely correlated with the subject's preferences. The neuron did not discriminate between the different lever movements. **C**, A PFcs neuron that showed a higher firing rate for apple juice than for orange juice and a higher firing rate for orange juice than for quinine. This neuron was from subject J, and so the neuronal firing rate directly correlated with the subject's preferences. The encoding of outcome information was stronger when the subject had made a rightward movement than when the subject had made a leftward movement.

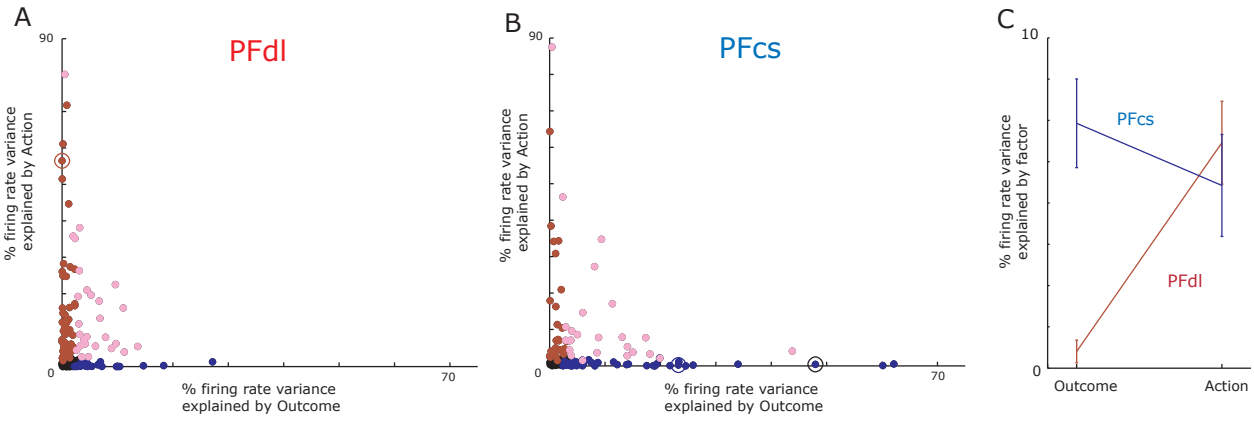

Figure 3.4: For each neuron from PFdl (**A**) and PFcs (**B**), we plotted the percentage of variance in the neuron's firing rate during the first-outcome epoch that could be attributable to either the action or outcome factor as determined from a two-way ANOVA. Those neurons whose firing rates showed a significant main effect of one or both of the experimental factors are color-coded (red, action only; blue, outcome only; pink, action and outcome). The data points circled in red, blue, and black indicate the data from the neurons illustrated in Figure 3.3A-C, respectively. **C**, Plot of the mean percentage of variance in the PFdl or PFcs neurons' firing rates that was explainable by either the action or the outcome factor.

had received orange juice. This neuron was from subject H, and so the neuronal firing rate inversely correlated with the subject's preferences. Figure 3.3C illustrates a PFcs neuron that showed a significant interaction between action and outcome ($F_{(2,371)} > 5.0, p < 0.01$). The neuron showed a higher firing rate for apple juice than for orange juice, and a higher firing rate for orange juice than for quinine. This neuron was from subject J, so the neuronal firing rate directly correlated with the subject's preferences. However, the encoding of the outcome information was stronger when the subject had made a rightward movement than for a leftward movement.

We contrasted the prevalence of these different types of neuronal encoding in PFdl and PFcs. In PFdl, nearly twice as many neurons showed a main effect of action (83/172 or 48%) as showed a main effect of outcome (49/172 or 28%). In contrast, in PFcs, more neurons showed a main effect of outcome (67/112 or 60%) than showed a main effect of action (52/112 or 46%). Few neurons in either area showed an action x outcome interaction (PFdl: 20/172 or 12%, PFcs: 10/112 or 9%). A comparison of the proportion of selective neurons between the areas revealed that outcome-selective neurons were significantly more prevalent in PFcs than in PFdl ($\chi^2 = 26.3, p < 5 \times 10^{-6}$), while the proportion of action-selective neurons did not significantly differ between the areas ($\chi^2 = 0, p > 0.1$).

We next examined the magnitude of the experimental effects using η^2 for each of the experimental factors (see 3.2.2 Statistical methods). For each neuron, this provided us with two selectivity measures: the degree to which the neuron encoded the action and the degree to which the neuron encoded the outcome. We then plotted these two measures against one another (Fig. 3.4A, B). This confirmed our findings from the prevalence analysis: the main difference between the two areas was that PFcs encoded both the action and the outcome, whereas PFdl encoded the action alone. We performed a two-way ANOVA using these selectivity measures as the dependent variable and independent variables of encoding (outcome or action) and area (PFdl or PFcs). There was a significant interaction between the factors (Fig. 3.4C) ($F_{(1,564)} > 12.7, p < 0.0005$). A post hoc simple effects analysis revealed that there was significantly higher action encoding in PFdl than outcome encoding ($F_{(1,564)} > 19, p < 0.00005$), whereas there was no significant difference in the strength of action and outcome encoding in PFcs ($F_{(1,564)} < 1, p > 0.1$). Furthermore, outcome encoding was significantly higher in PFcs than in PFdl ($F_{(1,564)} > 18, p < 0.00005$), but there was no significant difference in the strength of action encoding between the two areas ($F_{(1,564)} < 1, p > 0.1$).

In summary, while PFdl encoded information about the action, only PFcs encoded both pieces of information that were necessary to solve the task, namely the action and the outcome. In addition, the low incidence of neurons showing an interaction between the action and outcome factors, as well as the L-shaped pattern of the scatter in Figure 3.4B, suggests that the two experimental factors were encoded by separate populations of neurons in PFcs. In other words, neurons either strongly encoded the outcome or strongly encoded the action but were unlikely to encode both factors. Our subsequent analyses focus on determining more precisely what aspect of the outcome and action neurons are encoding.

3.3.4 Specific information encoded by outcome-selective neurons

To characterize more specifically how the neurons were encoding outcome information, we began by determining, for each neuron, its mean firing rate when each of the different outcomes occurred. We then determined the rank order of those means. There were three different outcomes and consequently six potential rank orderings of those three outcomes. Two of the six possible rank orderings were consistent with the subject's preferences: either a monotonically increasing or decreasing relationship between firing rate and preference. The other four possible rank orderings were not consistent with the subject's preferences. In PFcs, 47/67 (70%) of the outcome-selective neurons encoded the outcomes in a manner consistent with the subject's preferences, which was significantly greater than the one-third expected by chance (binomial test, $p < 1 \times 10^{-9}$). In PFdl, a similar proportion of outcome-selective neurons encoded the outcome in a manner consistent with the subject's preferences (28/49 or 57%, binomial test, $p < 0.0005$). In other words, despite consistent orderings making up only one-third of the possible orderings, approximately two thirds of the neurons encoded

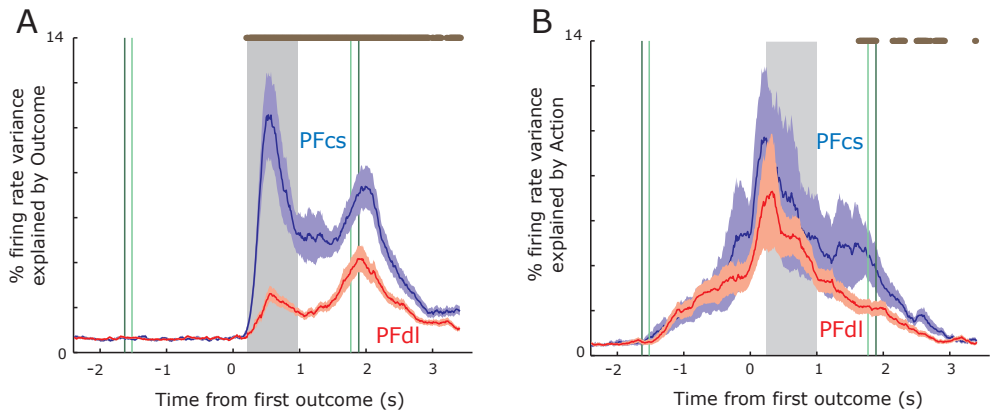

Figure 3.5: **A**, Time course of selectivity related to encoding the first outcome across the two neuronal populations as determined by the sliding ANOVA analysis. The bold line indicates the mean of the population, while the colored shading indicates the SEM. The gray shaded bar indicates the extent of the first-outcome epoch that we used for statistical analysis. The left pair of green vertical lines indicate the mean time across all trials at which the subject acquired fixation (subject H, dark green; subject J, light green). The right pair of green vertical lines indicate the mean time at which the subjects made their movements to sample the second outcome. The horizontal brown bar indicates time points at which the outcome selectivity was significantly stronger in PFcs relative to PFdl as determined by a t test. **B**, Time course of selectivity related to encoding the first action as determined by the sliding ANOVA analysis.

the outcomes in a manner consistent with the subject's preferences. Of these neurons, there was a tendency to show a negative relationship between the subject's preferences and their firing rate (PFcs: 32/47 or 68%, binomial test, $p < 0.005$, PFdl: 23/28 or 82%, binomial test, $p < 0.0001$). In other words, these neurons tended to show their highest firing rate to the least preferred juice and their lowest firing rate to the most preferred juice.

We confirmed that the neuronal selectivity was not a simple sensory response to the juice, by comparing each neuron's firing rate during the first- and second-outcome epochs. If the neurons were encoding the sensory properties of the juice, then their firing rate should show the same rank ordering of the juices in both epochs. To determine this, we performed the same analysis of neuronal activity during the second-outcome epoch as we had performed for the first-outcome epoch. Of the 116 neurons (PFcs: 67, PFdl: 49) that were outcome selective in the first- outcome epoch, 76 (PFcs: 43, PFdl: 33) were also selective during the second- outcome epoch. However, only 8 of these neurons (PFcs: 6, PFdl: 2) showed the same rank ordering of the outcomes in both epochs. Thus, the neuronal response to the juice did not depend solely on the physical properties of the juice, but also depended on the context within the task in which the juice occurred.

We also examined the time course over which neurons encoded outcome information.

We performed a sliding two-way ANOVA and calculated for each time point the percentage of variance in the neuron's firing rate that could be attributable to the outcome factor (see 3.2.2 Statistical methods). Figure 3.5A shows that the encoding of the outcome was stronger in PFcs than PFdl. An analysis of the latency at which neurons reached our criterion for encoding of outcome information revealed that this occurred earlier in PFcs (median 360 ms) than in PFdl (median 520 ms, Wilcoxon's rank-sum test $> 8855, p < 0.00005$), although our interpretation of this must be tempered by the very weak encoding of outcome information that was present in PFdl.

Previous studies have shown that as a subject becomes satiated with a specific outcome, the value of that outcome diminishes, causing changes in the subject's choice behavior (Balleine and Dickinson, 1998). Thus, we examined whether there was any evidence that the neuronal responses changed across the course of the session. We divided each session into quarters and repeated our previous analyses across these quarters. We did not find any evidence that encoding of the outcome differed across the course of the session. Across the four quarters, there was no difference in the proportion of neurons that encoded the outcome, the action, or the interaction between the two factors (Fig. 3.6A) ($\chi^2, p > 0.1$ for both areas and all comparisons). Nor was there a difference between the quarters in the proportion of outcome-selective neurons that encoded the outcome in a manner that was consistent with the subjects' preferences (Fig. 3.6B) ($\chi^2, p > 0.1$ for all comparisons). Finally, there was no difference in the mean strength encoding of the outcome across the session as determined by calculating the percentage of variance in neuronal firing rates attributable to the identity of the outcome (Fig. 3.6C). Thus, there was no evidence for any change in the neuronal encoding of the outcome across the course of the session. This was consistent with the lack of any behavioral change (Fig. 3.2C), which might suggest that our subjects were not fully sated on any of the specific juices by the end of the recording session.

In summary, PFcs neurons showed stronger encoding of the three different juice outcomes. The nature of this encoding was consistent with the neurons signaling the value of the juices in terms of the subject's preferences. Furthermore, the encoding was dynamic, as the neurons did not encode the second juice outcome in the same way as the first. We examined the encoding of the second juice outcome in more detail in subsequent analyses (see below).

3.3.5 Specific information encoded by action-selective neurons

To determine more precisely what information the action-selective neurons were encoding, we compared the encoding of action information during the first-outcome epoch with the encoding during the presample epoch. Over a third of the neurons showed significant action encoding during the presample epoch, and there was no significant difference between the two areas in the proportion of such neurons (PFdl: 60/172 or 35%, PFcs: 43/112 or 38%, $\chi^2 < 1, p > 0.1$). A comparison of the neuronal selectivity during the presample and first-outcome epochs revealed that neurons appeared to be encoding upcoming motor responses.

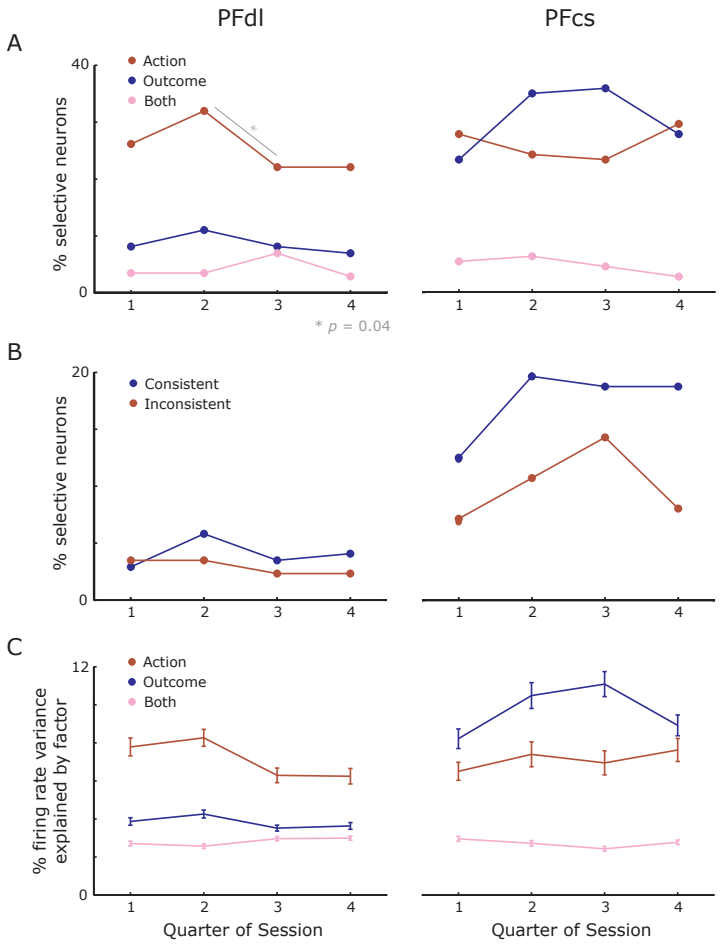

Figure 3.6: **A**, There was no difference across the quarters of the session ($\chi^2, p > 0.1$ for both areas and all comparisons) in terms of the prevalence of neurons encoding the action, outcome, or the interaction of the two factors. (Although there were significantly fewer action-selective neurons between the second and third quarters in PFdl, this effect did not survive the correction for multiple comparisons.) **B**, There was no difference across the session in the prevalence of neurons encoding outcome information in a manner that was either consistent or inconsistent with the subjects' preferences across the course of the session ($\chi^2, p > 0.1$ for all comparisons). **C**, There was no difference across the session in terms of the strength of outcome selectivity. For each neuron and each quarter in turn, we calculated the percentage of variance in the neuron's firing rate that could be attributed to either the outcome or the action. We performed a three-way ANOVA using the strength of selectivity as the dependent variable, and quarter, area, and encoding (outcome or action) as independent variables. None of the main effects or interactions with the quarter factor were significant ($p > 0.1$ in all cases), although there was a significant area $\times$ encoding interaction ($F_{(1,2254)} = 43, p < 1 \times 10^{-10}$), consistent with the interaction shown in Figure 3.4C.

For example, the neuron in Figure 3.3A showed a higher firing rate during the presample epoch on left-then-right trials, and a higher firing rate in the first-outcome epoch on right-then-left trials. Thus, the neuron showed a higher firing rate whenever the subject intended to make a leftward movement.

To investigate these effects at the population level, for each neuron and each epoch in turn we calculated a direction index (see 3.2.2 Statistical methods). This index was positive when the neuron showed a higher firing rate for upcoming leftward lever movements and negative when the neuron showed a higher firing rate for upcoming rightward movements. There was a significant correlation between the value of this measure in the two epochs in both PFdl and PFcs (Fig. 3.7A, B). This pattern of results is consistent with neurons encoding the upcoming motor action, with activity during the presample epoch encoding the upcoming first sample action and activity during the first-outcome epoch encoding the upcoming second sample action.

We also examined the value of the direction index during the prechoice epoch. If a neuron was encoding the upcoming action during this epoch, then its activity should reflect the direction of the lever movement that the subject intended at the choice phase of the task, in the same way that it encoded the intended first sample action during the presample epoch. Indeed, there was a positive correlation between the direction index in the prechoice and presample epochs in PFdl (Pearson product-moment correlation coefficient, $r = 0.34, p < 0.00001$), although not in PFcs ($r = 0.1, p > 0.1$). However, in both areas action encoding during the prechoice epoch was noticeably weaker than in the presample epoch, with 44/172 (26%) of PFdl neurons and 20/112 (18%) of PFcs neurons encoding the upcoming action. We confirmed this by contrasting the absolute magnitude of the direction index during the presample epoch with that in the prechoice epoch (Fig. 3.7C). We performed a two-way ANOVA on these index values with the factors of epoch (presample or prechoice) and area (PFdl or PFcs). There was a significant main effect of epoch ($F_{(1,564)} = 15, p < 0.0005$) with no other significant main effects or interactions ($F < 1, p > 0.1$). Thus, the direction index was significantly smaller during the prechoice epoch, and this effect was consistent across both areas.

We also examined whether the strength of action encoding at the choice phase depended on the identity of the first outcome. For example, there might be weaker encoding at the choice phase when the first outcome is the most preferred, since the subject could potentially decide their final choice when they receive the first outcome. To examine whether this was the case, we grouped the trials according to the identity of the first outcome and calculated the direction index for each epoch and for each group of trials (Fig. 3.7D). We then performed a three-way ANOVA using the direction index values as the dependent variable and epoch (presample or prechoice), area (PFcs or PFdl), and juice (the subject's preference for the first outcome) as independent variables. This revealed a significant epoch $\times$ juice interaction ($F_{(1,1692)} = 19, p < 0.00005$). A post hoc analysis of the simple effects revealed that direction selectivity was significantly weaker in the prechoice epoch when the first outcome was either

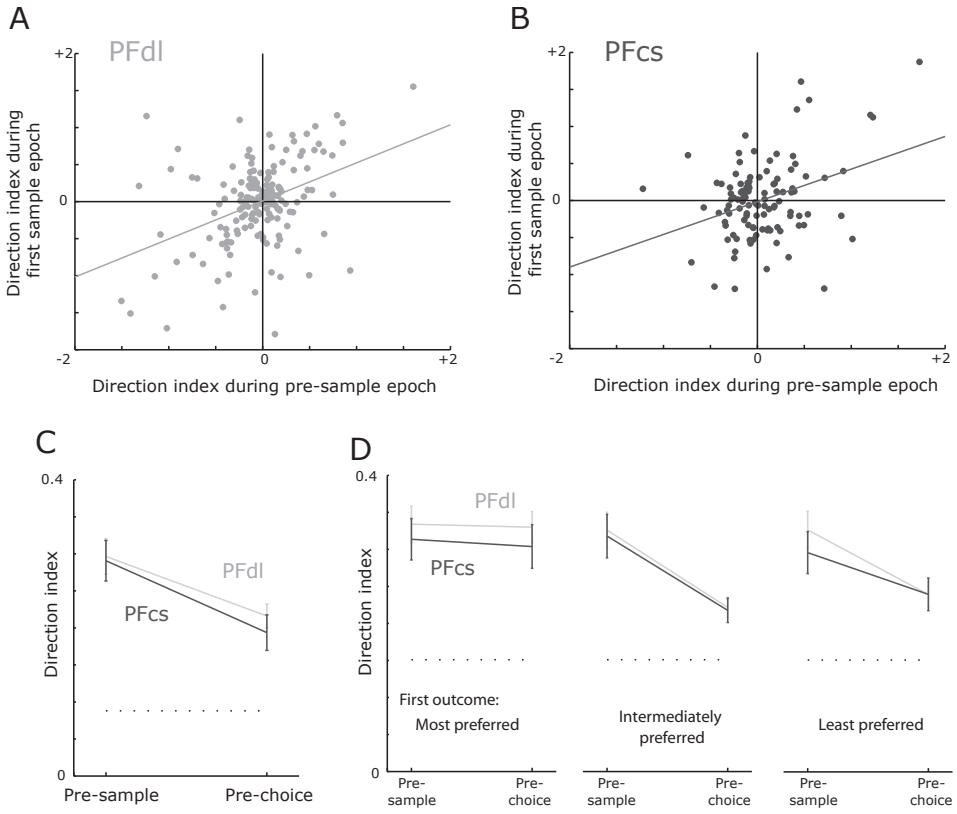

Figure 3.7: For each neuron from PFdl (**A**) and PFcs (**B**), we plotted the direction index during the presample and first-outcome epochs. The direction index consisted of the normalized difference in the neuron's firing rate on those trials in which the subject made a leftward lever movement to sample the first outcome and those in which he made a rightward lever movement. There was a significant correlation between the value of this measure in the two epochs in both PFdl (Pearson product-moment correlation coefficient, $r = 0.4, p < 5 \times 10^{-7}$) and PFcs ($r = 0.35, p < 0.0005$). **C**, Comparison of the mean action selectivity, defined as the absolute magnitude of the direction index, in PFdl and PFcs during the presample and prechoice epochs. Action selectivity was significantly weaker in the prechoice epoch, and the effect was consistent across both areas. The horizontal dotted line indicates the chance level of selectivity as calculated via a Monte Carlo analysis. **D**, Comparison of the mean action selectivity in PFdl and PFcs during the presample and prechoice epochs, separated according to the identity of the first outcome. Action selectivity was significantly weaker in the prechoice when the first outcome was either the intermediately preferred juice or the least preferred juice, but not when it was the most preferred juice. The horizontal dotted line indicates the chance level of selectivity as calculated via a Monte Carlo analysis (the chance level of selectivity was higher than in Fig. 3.6C, as the analyses were based on fewer trials, and so the index showed more variability).

the least preferred ($F_{(1,1692)} = 11, p < 0.005$) or the intermediately preferred ($F_{(1,1692)} = 19, p < 0.00005$) outcome, but not when the first outcome was the most preferred ($F_{(1,1692)} < 1, p > 0.1$). Thus, although the subject could potentially make their choice earlier in the trial when the first outcome was the most preferred juice, this did not appear to be what produced weak encoding of the final choice action.

We examined the time course of action encoding by performing a sliding two-way ANOVA and calculating for each time point the percentage of variance in the neuron's firing rate that could be attributable to the action factor (see 3.2.2 Statistical methods). Encoding of the upcoming sample actions began shortly after the presentation of the fixation cue and peaked shortly after the performance of the first sample action (Fig. 3.5B). It dropped rapidly once the subject had made the second sample action. The time course of action selectivity in PFcs and PFdl was very similar, although it was slightly stronger in PFcs around the time of the second sample action.

In summary, when the subject needed to monitor the outcome associated with his action there was strong encoding of the action in both PFcs and PFdl. In contrast, once the subject had received both the first and second outcome, such that he could make his choice between them, there was very little encoding of the movement that would lead to the final delivery of the reward. This result begs the question as to where in the brain the final choice action is encoded. Although PFcs appears to be important for encoding action and outcome information, the final action selection may occur in an area downstream of PFcs.

3.3.6 Encoding of outcome information during the second delay

Our analysis of the first sample action and its associated outcome revealed that neither PFcs nor PFdl encoded specific AO associations. Furthermore, while both areas encoded the action, a critical difference was that PFcs encoded the value of the outcome associated with that action.

We next examined neuronal selectivity during the second sample action. To quantify neuronal selectivity, we first separated the trials into three sets depending on the identity of the first outcome. Then, for each set of trials, we determined whether the neuron's firing rate was affected by the identity of the two possible second outcomes (see 3.2.2 Statistical methods). Thus, we performed three t tests comparing firing rates for (1) apple juice versus quinine (when the first outcome was orange juice), (2) orange juice versus quinine (first outcome was apple juice), and (3) orange juice versus apple juice (first outcome was quinine). Significantly more neurons in PFcs than in PFdl showed a significant difference between the two outcomes for at least one of these t tests (PFcs: 71/112 or 63%, PFdl: 53/172 or 31%, $\chi^2 = 28, p < 5 \times 10^{-6}$).

In both areas, the majority of the selective responses (PFcs: 47/71 or 66%, PFdl: 44/53 or 83%, binomial test, $p < 0.005$ in both cases) consisted of a higher firing rate when the second outcome was the less preferable of the two potential outcomes (as was the case for all

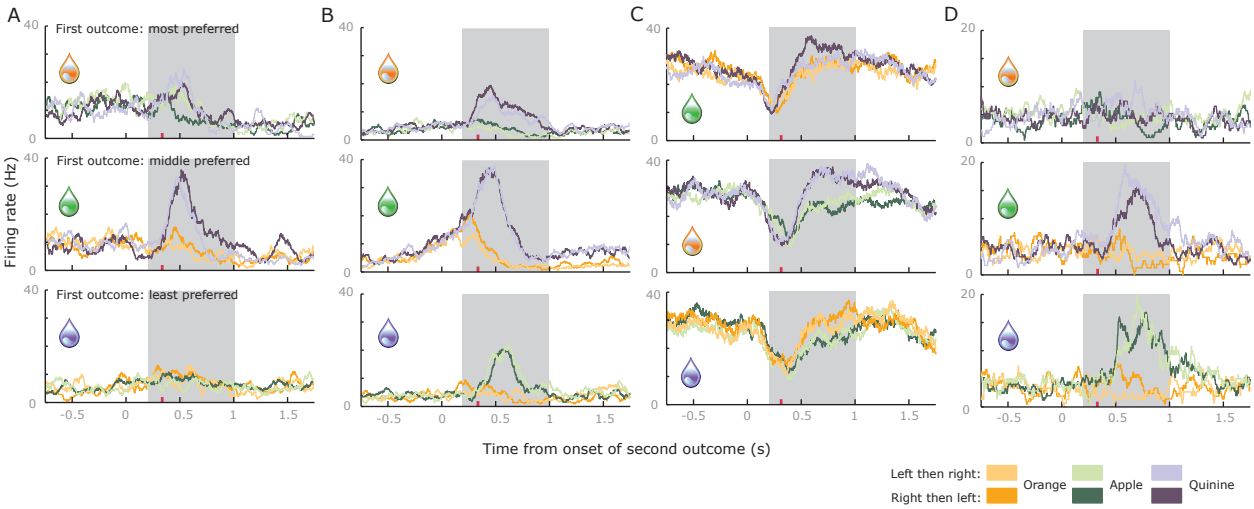

Figure 3.8: Spike density histograms illustrating the firing rate of neurons to the delivery of the second outcome. In each plot, the top panel indicates trials in which the first outcome was the most preferred juice, the middle panel indicates trials in which the first outcome was the intermediately preferred juice, and the bottom panel indicates trials in which the first outcome was the least preferred juice. The colored drop indicates the identity of the first outcome (orange, orange juice; green, apple juice; blue, quinine). The red tick mark indicates the offset of the juice delivery. The gray shading illustrates the second-outcome epoch that we used for statistical analysis. **A**, A PFcs neuron that had a significantly higher firing rate when the second outcome was quinine, but only when the intermediately preferred reward (apple juice for subject H) was the first outcome. **B**, A PFcs neuron, recorded from subject H, which showed a higher firing rate to the less preferable of the two potential outcomes that might follow each of the first outcomes. **C**, Another PFcs neuron that showed a higher firing rate to the less preferable of the two potential outcomes, this time recorded from subject J. **D**, A PFcs neuron that had a significantly higher firing rate to quinine following apple juice and to apple juice following quinine. In both cases, the subject would receive apple juice for his final choice, and consequently the most likely explanation for this pattern of selectivity is that it reflected the subject's expectancy of the reward that he would receive for his choice.

of the neurons in Fig. 3.8). These actions were not simply to the sensory properties of the juice, since the action to the second outcome depended on the identity of the first outcome. For example, the neuron in Figure 3.8A showed a strong response to quinine when it followed the delivery of apple, but not when it followed the delivery of orange. This type of selectivity accounted for 35/71 (49%) of the PFcs neurons that encoded the second outcome and 35/53 (66%) of the selective PFdl neurons. In PFdl, the vast majority of these selective responses (34/35 or 97%) occurred when the intermediately preferred juice was the first outcome. In PFcs, encoding of the second outcome was also most prevalent when the intermediately preferred juice was the first outcome (15/35 or 43%), but selective responses also occurred when the first outcome was the most preferred juice (5/35 or 15%) or the first outcome was the least preferred juice (13/35 or 37%).

Some neurons appeared to encode a more abstract version of this information. For example, the neuron in Figure 3.8B consistently showed a higher firing rate whenever the less preferred of the two possible second outcomes occurred, which for subject H was quinine following orange, quinine following apple, and apple following quinine. Figure 3.8C shows an analogous neuron in subject J. Thus, these neurons appeared to encode the relative value of the two outcomes that could potentially follow the first outcome, regardless of the specific identities of those outcomes. Such responses accounted for 10/71 (14%) of selective PFcs neurons and 2/53 (4%) of PFdl neurons. Of these neurons, the majority showed a higher firing rate to the less preferable outcome, with only one neuron (in PFcs) showing a higher firing rate to the more preferable outcome.

Finally, some neurons appeared to encode the juice that the subject expected in the choice phase of the task. For example, the neuron in Figure 3.8D showed a strong response to quinine following apple juice and apple juice following quinine. This neuron was from subject H, and these are the two juice combinations where he will receive apple juice as his final reward rather than orange juice. Such neurons accounted for 17/71 (24%) of the selective PFcs neurons and 8/53 (15%) of the PFdl neurons. This left just 12/71 (17%) of the selective PFcs neurons and 8/53 (15%) of PFdl neurons whose responses could not be accounted for by one of the above patterns of selectivity.

We also examined the time course over which neurons encoded the identity of the second outcome. We performed a sliding one-way ANOVA and calculated for each time point the percentage of variance in the neuron's firing rate that could be attributable to the identity of the second outcome (Fig. 3.9A, B). The results supported our conclusions from the analysis of the single neuron activity. Encoding of the second outcome was stronger in PFcs than in PFdl, and it was most evident when the first outcome was the intermediately preferred outcome. In PFcs, encoding of the second outcome was also present when the first outcome was either the most or least preferred, but this was not the case in PFdl.

In summary, encoding of the second outcome was most prevalent in PFcs. Most neurons appeared to encode the value of the second outcome relative to the first. This signal could take the form of a response to a specific second outcome in the context of a specific first

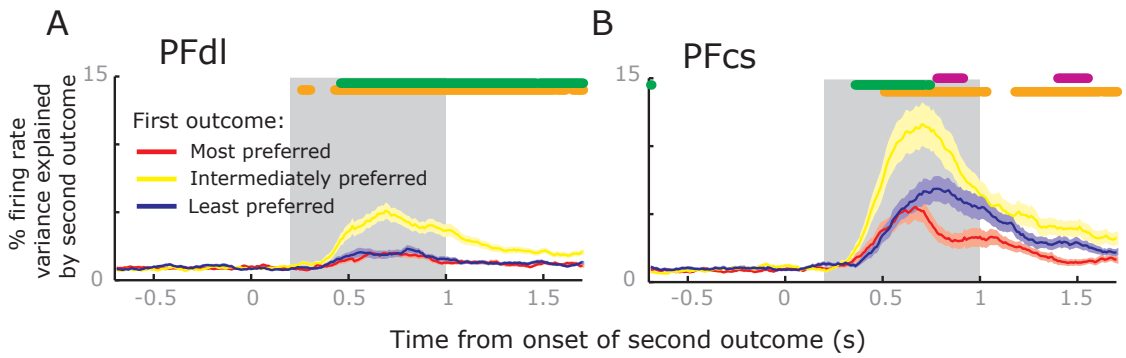

Figure 3.9: Time course of selectivity related to encoding the second outcome in PFdl (**A**) and PFcs (**B**), separated according to the identity of the first outcome. The bold line indicates the mean of the population, while the colored shading indicates the SEM. The horizontal colored lines indicate significant differences between the plots as determined by a t test. Orange points indicate a significant difference between the yellow and red plot, green points indicate a significant difference between the yellow and blue plot, and pink points indicate a significant difference between the blue and red plot. For example, the orange points indicate those times when the selectivity for the second outcome was stronger when the first outcome was intermediately preferred (yellow plot) than when the first outcome was the most preferred (red plot).

outcome, or it could take a more abstract form in which the neuron responded whenever the less preferable of the two potential outcomes (as indicated by the first outcome) occurred. This was particularly evident when the first outcome was the intermediately preferred juice. It is these trials in which the identity of the second juice is most critical. This is because on half the trials following the intermediately preferred outcome, the second outcome will be less preferable than the first, while on the other half of the trials the second outcome will be more preferable than the first. In addition, they are also the trials in which the difference in value between the two potential outcomes (the most and least preferred outcomes) is greatest. Either or both of these reasons may explain the bias toward encoding the second outcome when it follows the intermediately preferred juice.

3.3.7 Encoding of the chosen outcome

Our final analysis of neuronal selectivity focused on the actions that occurred when the subject received their chosen outcome. Many neurons encoded the identity of the chosen outcome. For example, the neuron in Figure 3.10A showed a higher firing rate when the subject received apple juice than when the subject received orange juice. In addition, on 25% of the trials the subject received a reward that was approximately twice as large as the standard reward size. Many neurons responded differentially depending on the magnitude

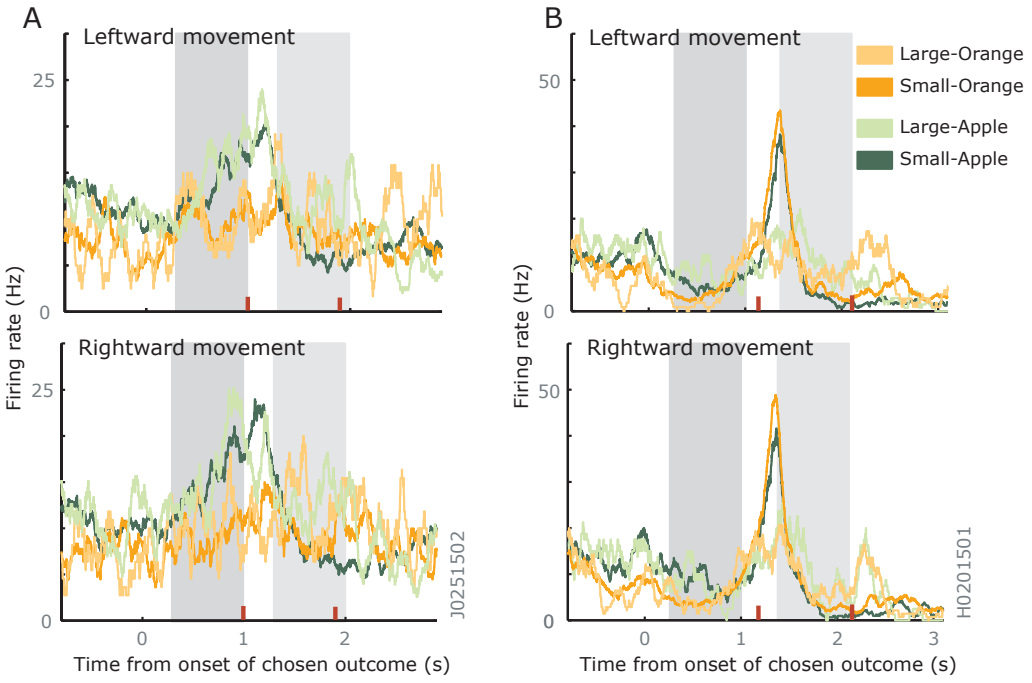

Figure 3.10: **A**, Spike density histograms illustrating a PFcs neuron that encodes the identity of the chosen juice. The plots are color-coded according to the identity and magnitude of the juice reward. The red tick marks illustrate the offset of the small and large juice reward. The dark gray shading illustrates the extent of the chosen epoch, and the light gray shading indicates the extent of the small-offset epoch. This neuron had a significantly higher firing rate when the subject received his preferred juice (apple for subject J) than when he received his intermediately preferred juice. **B**, A PFcs neuron that fired in response to the offset of the smaller reward. Note that this was not a generic reward offset response, since a similar strong response did not occur to the offset of the large reward.

of the received reward. For example, the neuron in Figure 3.10B showed a phasic response at the time of the offset of the small reward, which did not occur when the subject received a large reward.

To quantify the prevalence of these different types of neuronal selectivity in our neuronal populations, we performed a three-way ANOVA for each neuron and both the chosen and small-offset epochs in turn. We used the neuron's mean firing rate on each trial as the dependent variable and independent variables of juice (whether the subject had chosen the most preferred or intermediately preferred juice), action (whether the subject had made a leftward or rightward lever movement to indicate their choice), and magnitude (small or large reward trials). During the chosen epoch, a small number of neurons in each area showed a main effect of juice or a main effect of action (Fig. 3.11A). No other main effects or interactions accounted for $> 3\%$ of the neurons. During the small-offset epoch, the majority of selective neurons encoded the magnitude of the received juice while a smaller population encoded the juice's identity. The prevalence of neurons encoding the magnitude of the juice was greater in PFcs than PFdl, and PFcs tended to respond more strongly to the most preferred juice. In summary, the identity of the juice and its magnitude both drove neuronal activity, and the encoding of reward-related parameters was biased toward PFcs.

We also examined the relationship between encoding juice information during the sample phase of the task (specifically the first-outcome epoch) and the choice phase of the task (the chosen and small-offset epochs). There was no evidence that neurons that encoded juice information during the first-outcome epoch were any more or less likely to encode juice information during the chosen and small-offset epochs than were neurons that did not encode juice information during the first-outcome epoch (χ^2 test, $p > 0.1$ in all cases). For those neurons that did encode juice information during both phases of the task, we examined how their selectivity compared. (Too few PFdl neurons encoded juice information at both phases of the task to permit a meaningful statistical test of their response properties, so for this and subsequent analyses, we focused solely on PFcs neurons.) PFcs neurons tended to encode juice information during the choice phase of the task in the opposite direction to how they encoded it during the sampling phase of the task (Fig. 3.11B). To ensure sufficient neurons to permit us to examine this statistically, we combined the two groups, i.e., those that encoded juice information during the first-outcome epoch and the chosen epoch, and those that encoded juice information during the first-outcome epoch and the small-offset epoch. (Only one neuron was common to both groups, and our analyses were unaffected by whether we included or excluded this neuron as a member of both groups.) This analysis confirmed that neurons were more likely to switch their juice preference between the sample and choice phases of the task than they were to maintain the same preference (binomial test, $p < 0.05$). We also compared the firing rates that were elicited by the juices in the sample and choice phases of the task (Fig. 3.11C). For each neuron and each epoch, we calculated the neuron's preferred juice, defined as the juice that elicited the highest firing rate, and determined how this firing rate compared between those epochs. For those neurons encoding juice in

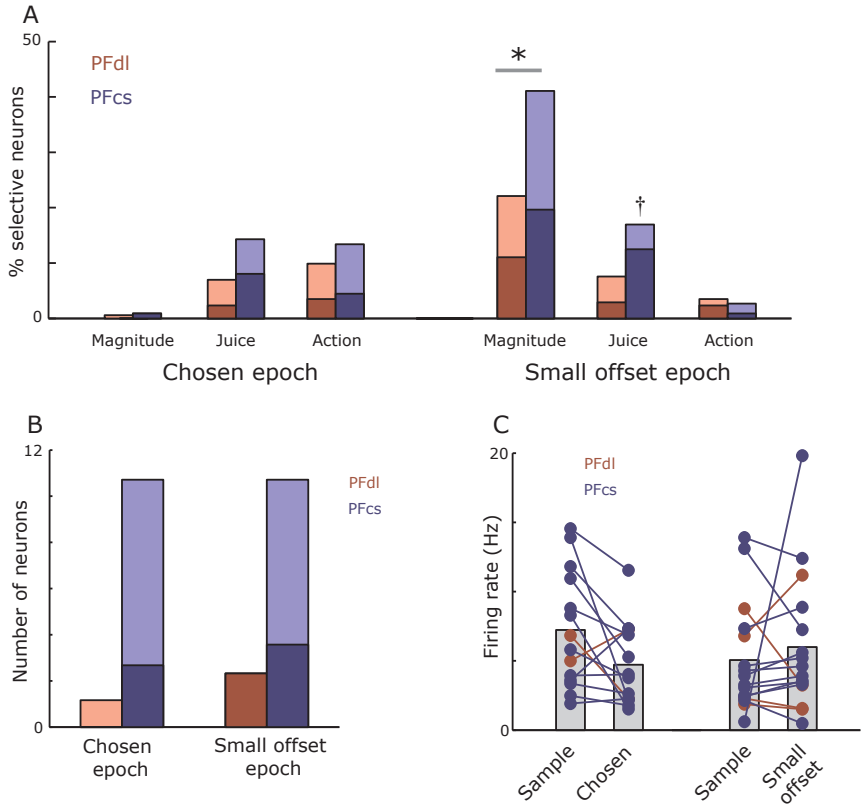

Figure 3.11: **A**, Percentage of neurons that encoded either the magnitude of the chosen reward (magnitude), the identity of the chosen reward (juice), or the direction of the behavioral response to select the chosen reward (action). For those neurons encoding the magnitude of the reward, the darker shading indicates the proportion that showed a stronger response to the larger reward. For those encoding the juice, the darker shading indicates the proportion that showed a stronger response to the more preferred juice. For those neurons encoding the action, the darker shading indicates the proportion that showed a stronger response for a leftward movement. The asterisk indicates that the prevalence of encoding the magnitude of the reward was significantly greater in PFcs than in PFdl ($\chi^2, p < 0.05$). The dagger indicates that there were significantly more PFcs neurons that showed a stronger response to the more preferred juice than the less preferred juice (binomial test, $p < 0.05$). **B**, Proportion of neurons that encoded outcome information during both the first-outcome epoch and the choice phase of the task that showed either the same (dark shading) or opposite (light shading) preference among the juices. **C**, Comparison of the firing rate elicited by the neuron's preferred juice for those neurons that encoded outcome information during both the sample and choice phases of the task.

both the first-outcome and chosen epochs, the firing rate to the neuron's preferred juice was significantly less in the chosen epoch (Wilcoxon signed-rank test, $p < 0.05$). In contrast, for those neurons encoding juice in the first-outcome and small-offset epochs, there was no significant difference between the mean firing rates (Wilcoxon signed-rank test, $p > 0.1$).

In summary, PFcs neurons tended to encode juice preference at the choice phase in the opposite direction to how they encoded this information during the sample phase. In addition, the neurons showed lower firing rates to chosen rewards than to rewards received during the sample phase. Thus, PFcs neurons showed their strongest activity when the subject had to monitor the outcome of an action, as in the sampling phase, or when it was not possible to predict the outcome of the action, as in the small-offset epoch, which is when the subject learned whether he would receive a small or large reward.

3.3.8 Functional anatomy within areas

Figure 3.12 illustrates the location of neurons that were action selective and outcome selective during the first-outcome epoch. To determine whether there was a relationship between the incidence of selective neurons and their anterior-posterior position, we performed a logistic regression using whether or not a neuron was selective as the dependent variable and the anterior-posterior position of the recording location as the predictor variable. We did this for each area and each variable (action or outcome) in turn. Action-selective neurons were significantly more prevalent in more posterior recording locations in PFdl ($p < 0.00005$). However, there was no relationship between location and action selectivity in PFcs, and no relationship between location and outcome selectivity in either area.

In addition, we have previously reported that neurons encoding the value of multiple decision-related parameters are restricted to the anteriormost portion of PFcs, with a marked drop-off in the prevalence of such neurons as one moves posterior to the genu of the corpus callosum [Kennerley *et al.*, 2009]. To examine whether such a functional dissociation existed in the current task, we recorded an additional dataset that focused on the dorsal bank of the cingulate sulcus posterior to the genu of the corpus callosum (antero-posterior +20 to +23) in subject J. We recorded 34 neurons from this area ($\text{PFcs}_{postgenual}$). We treated this as a separate dataset: it was not included in our previous analyses.

Neuronal selectivity in $\text{PFcs}_{postgenual}$ was markedly different from the anterior part of PFcs. During the first-outcome epoch, 28/34 (82%) of the neurons encoded the action compared with 5/34 (15%) that encoded the outcome. When we plotted the magnitude of the experimental effects of outcome and action, it was clear that the majority of neurons encoded the action with little effect of the outcome (Fig. 3.13A). To confirm that the neuronal encoding in $\text{PFcs}_{postgenual}$ was indeed different from our original PFcs dataset, we performed a two-way ANOVA using the magnitude of the experimental effects as the dependent variable and independent variables of encoding (outcome or action) and area ($\text{PFcs}_{postgenual}$ or PFcs). There was a significant interaction between the factors (Fig.

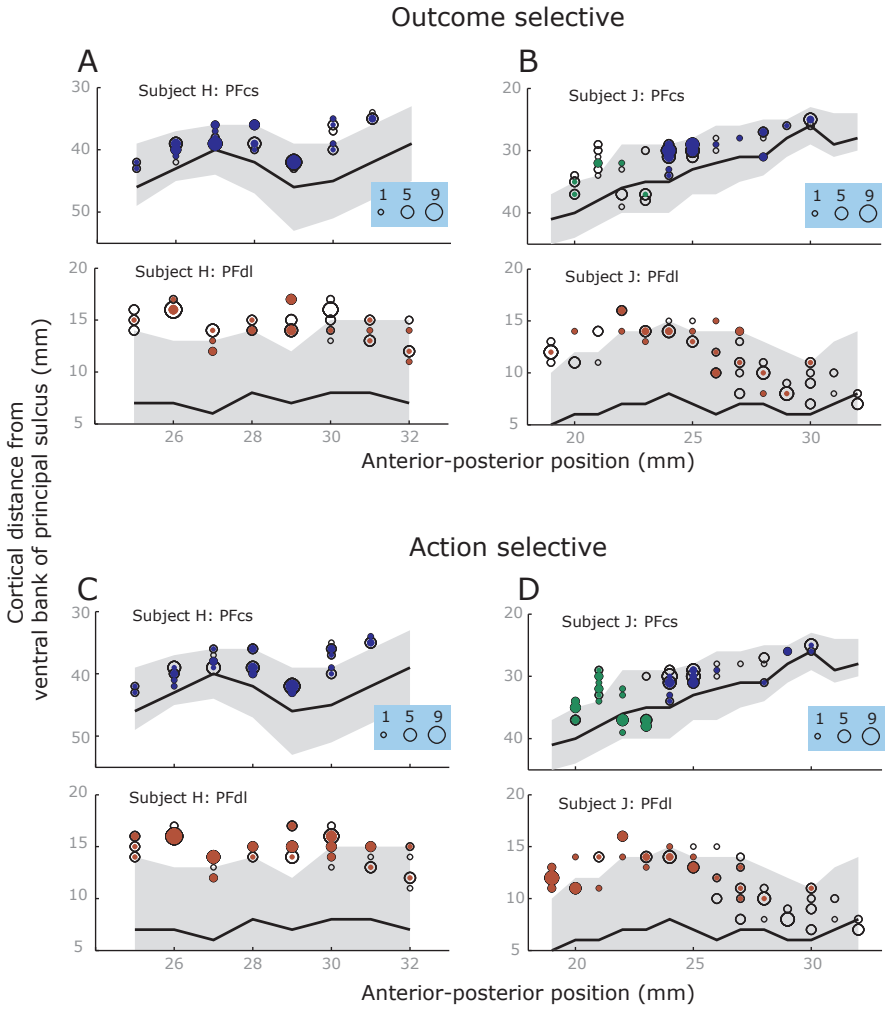

Figure 3.12: Flattened reconstructions of the cortex indicating the location of all recorded neurons (open circles) and outcome-selective neurons (filled circles) in subject H (**A**) and subject J (**B**). The size of the circles indicates the number of neurons recorded at that location. Red circles indicate neurons recorded from PFdl, blue circles indicate those recorded from PFcs, and green circles indicate those recorded from PFcs$_{postgenual}$. Gray shading indicates the unfolded cingulate sulcus in PFcs and the unfolded principal sulcus in PFdl. We measured the anterior'posterior position from the interaural line (x-axis) and the dorsoventral position relative to the lip of the ventral bank of the principal sulcus (0 point on y-axis). For ease of viewing, each plot is presented so that dorsal regions are at the top and ventral regions are at the bottom. C and D indicate the location of action-selective neurons recorded from subject H and subject J, respectively.

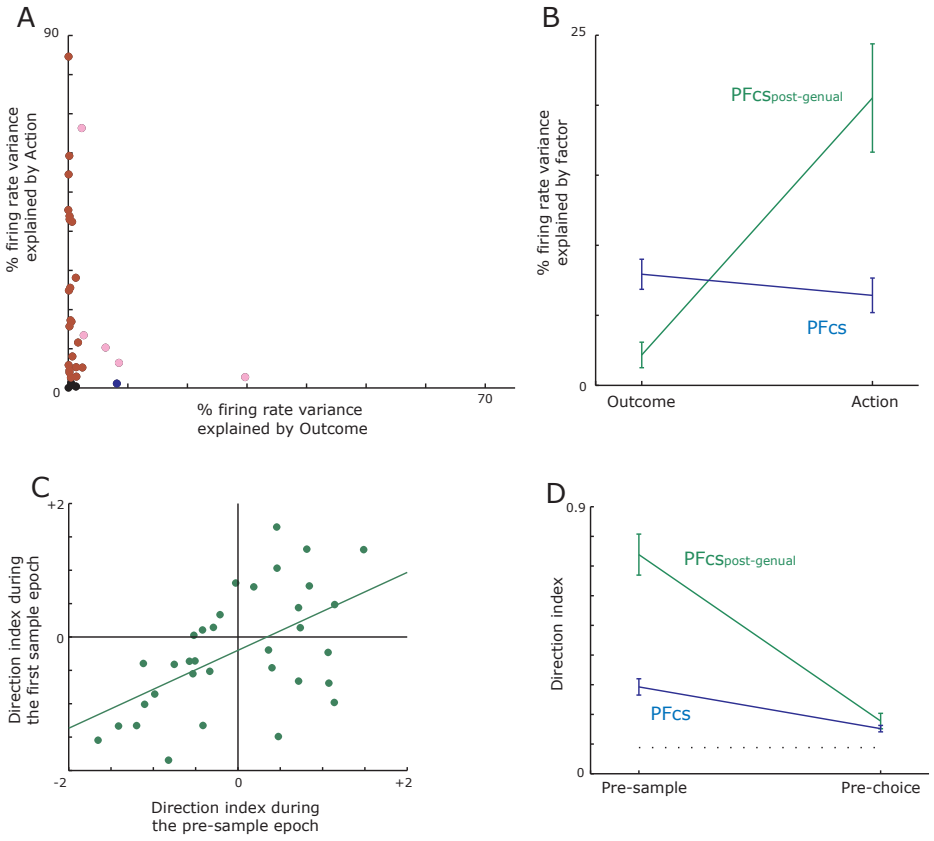

Figure 3.13: **A**, Percentage of variance in PFcs$_{postgenual}$ neuronal firing rates during the first-outcome epoch that could be attributable to either the action or outcome factor as determined from a two-way ANOVA. Conventions are as in Figure 3.4. **B**, Plot of the mean percentage of variance in PFcs$_{postgenual}$ or PFcs neuronal firing rates that was explainable by either the action or the outcome factor. **C**, Scatter plot of the direction index values during the presample and first-outcome epochs for every neuron in PFcs$_{postgenual}$. There was a significant positive correlation (Pearson product-moment correlation coefficient, $r = 0.55, p < 0.001$) indicative of the neurons encoding the upcoming action. **D**, Comparison of the mean action selectivity, defined as the absolute magnitude of the direction index, in PFcs$_{postgenual}$ and PFcs during the presample and prechoice epochs. The horizontal dotted line indicates the chance level of selectivity as calculated via a Monte Carlo analysis.

3.13B) ($F_{(1,288)} = 29, p < 5 \times 10^{-6}$). A post hoc simple effects analysis revealed that there was significantly greater action encoding in PFcs$_{postgenual}$ than outcome encoding ($F_{(1,288)} = 32, p < 5 \times 10^{-7}$). In addition, outcome encoding was significantly weaker in PFcs$_{postgenual}$ than in the original PFcs dataset ($F_{(1,288)} = 5, p < 0.05$), and action encoding was significantly greater in PFcs$_{postgenual}$ ($F_{(1,288)} = 29, p < 5 \times 10^{-6}$). These results supported our previous findings [Kennerley *et al.*, 2009]: in PFcs$_{postgenual}$ there is a clear decrease in the encoding of outcome-related information and a clear increase in the encoding of action-related information relative to more anterior regions of PFcs.

To examine the nature of the action selectivity evident in PFcs$_{postgenual}$, we compared our direction index in the presample and first-outcome epochs. There was a significant positive correlation indicative of the neurons encoding the upcoming action (Fig. 3.13C). The predominance of action selectivity raised the possibility that PFcs$_{postgenual}$ might be an area downstream of the anterior regions of PFcs that could be responsible for encoding the final choice action. In fact, there was no evidence that this was the case: selectivity relating to the encoding of the final action did not differ between PFcs$_{postgenual}$ and the original PFcs dataset during the prechoice epoch (Fig. 3.13D). We confirmed this by performing a two-way ANOVA on the absolute magnitude of the direction index values with the factors of epoch (presample or prechoice) and area (PFcs$_{postgenual}$ or PFcs). There was a significant interaction between the two factors ($F_{(1,288)} = 39, p < 5 \times 10^{-8}$). A post hoc analysis of the simple effects revealed that the direction index was significantly smaller in the prechoice epoch than in the presample epoch in both areas (PFcs$_{postgenual}$: $F_{(1,288)} = 91, p < 1 \times 10^{-15}$, PFcs: $F_{(1,288)} = 18, p < 0.00005$). Furthermore, whereas the direction index was significantly stronger in PFcs$_{postgenual}$ than our original PFcs dataset during the presample epoch ($F_{(1,288)} = 88, p < 1 \times 10^{-15}$), there was no difference between the areas in its strength during the prechoice epoch ($F < 1, p > 0.1$). Thus, although PFcs$_{postgenual}$ encoded the upcoming action strongly during the presample epoch, there was little evidence that it was responsible for implementing the final choice action.

3.4 Discussion

Using a task that required subjects to maintain information about which of three potential outcomes was associated with one of two possible actions, we observed marked differences in the properties of PFcs and PFdl neurons. Whereas both areas encoded the action, PFcs neurons also encoded the value of the outcomes associated with those actions. PFcs neurons encoded the value of the first outcome with respect to the subjects' preferences. They subsequently encoded the value of the second outcome relative to the first outcome. This information would be sufficient for the subject to determine their choice, but neurons showed much weaker encoding of both actions and outcomes during the choice phase of the task, compatible with the notion that PFC is more important for monitoring behaviors than

implementing them.

3.4.1 Differential control of action selection by PFdl and PFcs

Our results suggest that PFcs is important for monitoring the value of an outcome that results from a specific behavioral action. This conclusion is compatible with recent theories of PFcs function, which suggest this region may be important for valuing actions [Rushworth *et al.*, 2007; Quilodran *et al.*, 2008; Rushworth and Behrens, 2008]. For example, PFcs lesions render monkeys unable to sustain rewarded action in an outcome-guided choice task [Kennerley *et al.*, 2006], while PFcs neurons tend to encode both rewards as well as the action that led to the reward [Matsumoto *et al.*, 2003; Williams *et al.*, 2004]. The results of the current study help to define what information PFcs encodes. Neurons encoded outcome information in a manner that was consistent with the subjects' preferences. This supports the idea that PFcs neurons encode outcome information via an abstract value signal [Amiez *et al.*, 2006; Wallis, 2007; Kennerley *et al.*, 2009].

In contrast, despite a good deal of focus on PFdl in the control of goal-directed behavior [Shallice and Burgess, 1991; Duncan *et al.*, 1996; Watanabe, 1996; Miller and Cohen, 2001], it is becoming increasingly apparent that its role in this process is relatively constrained. Neuropsychological studies show that PFdl damage does not impair decision making [Bechara *et al.*, 1998; Fellows, 2006; Fellows and Farah, 2007; Baxter *et al.*, 2008]. Neurophysiological studies show that PFcs neurons encode multiple decision parameters, such as the amount of available juice, the amount of work necessary to earn the juice or the probability of juice delivery, while PFdl neurons show much weaker encoding of such information [Kennerley *et al.*, 2009]. Results from [Matsumoto *et al.*, 2003] have also supported this conclusion. In a task that required subjects to learn associations between stimuli (two different pictures), actions (holding or releasing a lever), and outcomes (receiving or not receiving a juice reward), PFcs neurons encoded the AO association, while PFdl neurons encoded the stimulus-action association.

In addition, we saw a marked decline in the prevalence of encoding outcomes in the more posterior parts of PFcs. In a previous study, we saw a marked decrease in the complexity of the encoding of decision parameters in the more posterior PFcs [Kennerley *et al.*, 2009]. This raises the possibility that the dorsal bank of the cingulate sulcus anterior to the genu of the corpus callosum is functionally distinct from the more posterior regions of the dorsal bank. We note that there is considerable disagreement regarding the cytoarchitectonic designation of the dorsal bank of the anterior PFcs, with it variously labeled as area 9 [Vogt *et al.*, 2005], 32 [Petrides and Pandya, 1994], 9/32 [Paxinos *et al.*, 2000], or 24b [Carmichael and Price, 1994]. Regardless of the cytoarchitectonic labels, our results highlight the functional diversity of the cingulate sulcus. In addition to the anterior/posterior distinction, the dorsal and ventral banks of the cingulate sulcus also appear to be functionally distinct [Hoshi *et al.*, 2005]. These differences in functional anatomy must be considered if we are to understand

PFcs function [Rushworth *et al.*, 2004].

3.4.2 The specific nature of action encoding in PFcs

Some neurophysiological studies of PFcs find strong encoding of behavioral responses [Matsumoto *et al.*, 2003], while others find little evidence for such encoding [Ito *et al.*, 2003; Matsumoto *et al.*, 2007]. Our own data seem to suggest that PFcs action encoding is most evident at the point in the trial when the subject must monitor the outcome associated with an action. Those studies that did not observe strong action encoding, may not have required such monitoring, either because a cue instructed which movement to make [Ito *et al.*, 2003] or because the subject could adopt a simple 'win-stay, lose-shift' strategy [Matsumoto *et al.*, 2007]. A further issue relates to why our study found segregation of neuronal response properties: neurons tended to encode either the outcome or the action but not both. Although similar segregation occurs in the striatum [Lau and Glimcher, 2007], previous studies of PFcs found that neurons integrated action and outcome information [Matsumoto *et al.*, 2003]. The segregation that we observed might be a consequence of our task design, since information in our task was only relevant for the current trial. Integration of action and outcome information may only occur when it is necessary to keep track and update an action's value across multiple trials.

Both PFcs and PFdl strongly encoded the movements to sample the outcomes but only weakly encoded the final choice movement. These findings are compatible with accounts of prefrontal function that emphasize its role in monitoring behavior, rather than planning and executing actions [Petrides, 1996]. They are also consistent with studies that show stronger encoding in both PFcs [Matsumoto *et al.*, 2007; Quilodran *et al.*, 2008] and PFdl [Procyk and Goldman-Rakic, 2006] when a subject is determining which action is rewarded than when they are repeating a known rewarded action. More generally, we can interpret such findings within the framework of exploration (discovering what outcomes are associated with specific actions) versus exploitation (repeating actions that lead to known outcomes), with prefrontal neurons more involved in exploration than exploitation. This raises the question as to whether there is a downstream motor area responsible for implementing the final choice, although our data seem to exclude PFcs$_{postgenual}$ from such a role. For example, reward-dependent modulation of motor responses increases in progressively downstream motor structures [Roesch and Olson, 2003]. Alternatively, exploitation may be a default strategy that exploration must override, with the consequence that neuronal responses to exploration will always be stronger than exploitative responses [Daw *et al.*, 2006].

3.4.3 Biased encoding of valence in PFcs

Both EEG [Gehring *et al.*, 1993; Miltner *et al.*, 1997] and fMRI [Carter *et al.*, 1998; Ullsperger and von Cramon, 2003; Holroyd *et al.*, 2004] studies have found that PFcs activity is stronger

to failures than successes. Our finding that neurons responded more strongly to the less preferable outcomes during both the first- and second-outcome epochs supports this view. This valence bias is the opposite of that reported in other areas involved in value-based decision making, such as orbitofrontal cortex, where neuronal responses were stronger for positive events [Roesch and Olson, 2004]. This has prompted some theories to argue that choice behavior involves the integration of benefits from orbitofrontal cortex and costs from PFcs [Cohen *et al.*, 2007]. However, such a conceptualization may be overly simplistic.

Several studies have found comparable responses in PFcs to successes and failures [Knutson *et al.*, 2000; Holroyd and Coles, 2002; Walton *et al.*, 2004]. We previously found an even split between neurons that increased their firing rate as the value of a choice increased compared with those that increased their firing rate as the value of a choice decreased in both PFcs and orbitofrontal cortex [Kennerley *et al.*, 2009]. PFcs neurons are equally likely to respond to gains (when a subject was cued that they would receive more juice than expected) and losses (when a subject learned they would receive less juice than expected) [Sallet *et al.*, 2007]. Furthermore, PFcs neurons respond to rewards in a manner that is highly sensitive to the reward context in which they occur [Sallet *et al.*, 2007]. For example, the response of a PFcs neuron to a large reward will be larger if it occurs in a block of relatively small rewards versus a block of relatively large rewards. This suggests that PFcs neurons may be relatively susceptible to framing [Tversky and Kahneman, 1981]: their response may depend on the context in which the choice is presented. A possible explanation for why we observed larger signals for negative outcomes may be because subjects tend to be overly optimistic about the likelihood of a positive outcome [Miller and Ross, 1975]. Thus, our subjects may have hoped for their most preferred juice to occur following a sample, and the PFcs signaled when this did not occur.

3.5 Conclusion

In summary, our results support the notion that PFcs is important for monitoring the value of an outcome produced by a specific behavioral response and highlight the relative lack of outcome encoding in PFdl. These neuronal populations in PFcs could contribute to many of the functions in which PFcs has been implicated, including action valuation, error detection, and decision making.

Work in this chapter has been published in the *Journal of Neuroscience* [Luk and Wallis, 2009].

Chapter 4

Neuronal representations of action-outcome and stimulus-outcome associations

4.1 Introduction

In a dynamic world, values of stimuli in the environment and of one's actions change. To obtain desired outcomes efficiently, it becomes necessary to track the value of stimuli and actions as stimulus-outcome (SO) and action-outcome (AO) associations, respectively. Prefrontal cortex may be critical in tracking these associations, particularly the prefrontal region of the cingulate sulcus (PFcs) and orbitofrontal cortex (PFo). Both areas are strongly connected to limbic areas [Carmichael and Price, 1995a], thus providing them with reward outcome information. PFcs has robust connections to motor regions in the cingulate motor area and ventral premotor cortex [Carmichael and Price, 1995b]. It is also part of a medial prefrontal network referred to as the "visceromotor network" [Carmichael and Price, 1996]. PFo is in a complementary network focused on processing sensory information from inferotemporal cortex, insula and pyriform cortex, among other sensory areas [Passingham, 1995; Carmichael and Price, 1995b].

Consistent with the anatomy, studies suggest that there is a division of labor across PFcs and PFo, with PFcs involved in encoding AO associations and PFo involved in encoding SO associations. PFcs neurons encode values of a spectrum of actions like instructed movements [Matsumoto *et al.*, 2003; Luk and Wallis, 2009], chosen actions [Hayden and Platt, 2010], effort requirements to obtain reward [Walton *et al.*, 2002; Croxson *et al.*, 2009] among other decision parameters. On the other hand, PFo neurons capture the value of outcomes predicted by specific stimuli and contexts [Thorpe *et al.*, 1983; Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2008; Bouret and Richmond, 2010]. BOLD activity in PFcs correlates with the subject's actions, while BOLD activity in PFo increases with changing SO

associations [Noonan *et al.*, 2011]. Lesions of PFcs render rats [Balleine and Dickinson, 1998; Dias and Aggleton, 2000] and monkeys [Hadland *et al.*, 2003] incapable of adjusting their actions based on changes to outcomes based on responses, but they can link changing outcomes to associated stimuli [Rudebeck *et al.*, 2008]. In contrast, PFo lesions impair rats [Pickens *et al.*, 2003] and monkeys [Izquierdo *et al.*, 2004] on SO associations, but do not affect action encoding [Ostlund and Balleine, 2007; Rudebeck *et al.*, 2008]. Collectively, these studies provide strong evidence for a double dissociation between the medial and orbital prefrontal areas in how they represent associated value, suggesting that PFcs is required for processing AO associations, and PFo for SO associations. The neuronal mechanisms instantiating this differential encoding across areas have yet to be determined. Thus, we aimed to determine the precise neuronal mechanisms that take place in PFcs and PFo that enable them to process information on AO and SO associations.

Past experiments often required subjects to focus on either AO and SO associations throughout an entire experiment or used a task that required both types of associations. By merging AO and SO associations into single task, it becomes difficult to know how much attention is placed on individual components of the task. Subjects may have utilized SO, AO, or both associations to choose optimally [Valentin *et al.*, 2007]. Thus, we had subjects alternate between performing either AO and SO tasks within single recording sessions, and we contrasted the neuronal representation of AO and SO associations in PFcs and PFo. The specific associations changed trial to trial, so the subject needed to remember both the predictor (A or S) and its paired outcome to perform optimally. We then compared PFcs and PFo in their encoding of AO and SO associations.

4.2 Materials and methods

4.2.1 Behavioral task

The experiment comprised of two tasks, an AO and an SO task (figure 4.1A). Both tasks followed the same structure as was described in section 3.2.1 for the AO task. A trial began with a sampling period and followed with a choice period. In an SO trial, the subject viewed two different pictures that yielded two different rewards during the sampling period. Then in the choice period, the subject saw both pictures side-by-side and moved the lever in the direction of the picture associated with the more preferred reward. In order for the subject to obtain more of his preferred reward, he needed to remember which picture was linked with which outcome. The AO trial was similar, but instead of viewing pictures, the subject executed two different lever movements and then repeated one of them in the choice period. For more details on the AO task, see 3.2.1 Behavioral task. We made one change to the AO task described in the previous chapter: the colors of the block cues were identical in this experiment. This lack of visual information instructing actions encouraged the subject

to ignore stimuli in the AO task. The SO and AO contingencies changed trial to trial. All cue-outcome pairings were equally likely to occur. The spatial order of the pictures during the choice period of the SO task were counterbalanced and randomized.

Cues were actions in the AO task (left and right lever movements) and pictures in the SO task (natural scene, portrait images of isoluminance and square dimensions with sides of 5 degrees of visual angle). Pictures were presented for 0.6 s. Reward was presented such that its offset coincided with the offset of the picture. The rewards we used were unique to each subject. We gave subject H orange juice (Minute Maid), apple juice (50% dilution, Safeway), and quinine (1.1 mM, Sigma-Aldrich). Subject J received orange juice (75% dilution), apple juice, and quinine (4.4 mM). Reward amounts were also tailored to each subject, so as to ensure that they showed consistent preferences between the juices and so that they received their daily fluid aliquot within a single recording session. Sample juices lasted approximately 0.4 s (0.25 mL). Choice rewards varied between two sizes, one that was small and occurred 75% of the time, and the other that was large and occurred 25% of the time. The small reward was typically 3-5 times larger than the sample juice, and the large reward was typically 5-8 times larger.

The two tasks were organized into blocks: 35 trial-long blocks of AO trials, 45 trial-long blocks of SO trials. Because behavioral performance was lower on the SO task, this ensured that we had a similar number of correct trials for each task. Within a block, the order of the cues did not change. Hence, we had two different blocks of the AO task and two other blocks of the SO task. During a recording session, blocks alternated from being one of the AO blocks to being one of the SO blocks.

4.2.2 Statistical methods

We excluded error trials from our statistical analyses. Subjects made two types of errors, which have been described in section 3.2.1. Briefly, subjects needed to maintain their eye position within $\pm 2.5^\circ$ of the fixation spot throughout the trial prior to the choice period. If the subject's eye went beyond the $\pm 2.5^\circ$ boundary, the subject experiences a 5 s timeout (red screen) before the trial returned to where the subject left off. The second error only occurred in the AO task. During the sampling period, the subject could make the wrong movement, i.e., a leftward movement instead of the correct rightward movement. If the subject made a wrong movement, he experienced a 5 s timeout (yellow screen). Afterward, the trial resumed from the point where the subject was prior to the error. If the error occurred prior to the delivery of the first sample reward, then the trial was included in our statistical analyses, but all other error trials were excluded.

We also excluded recording sessions in which the subjects exhibited significant side bias in their choices. Specifically, we divided a session into blocks of 30 trials. Within each block, we performed a binomial test on the direction of the choice response. If there was a significant difference in the number of leftward and rightward movements, we removed the

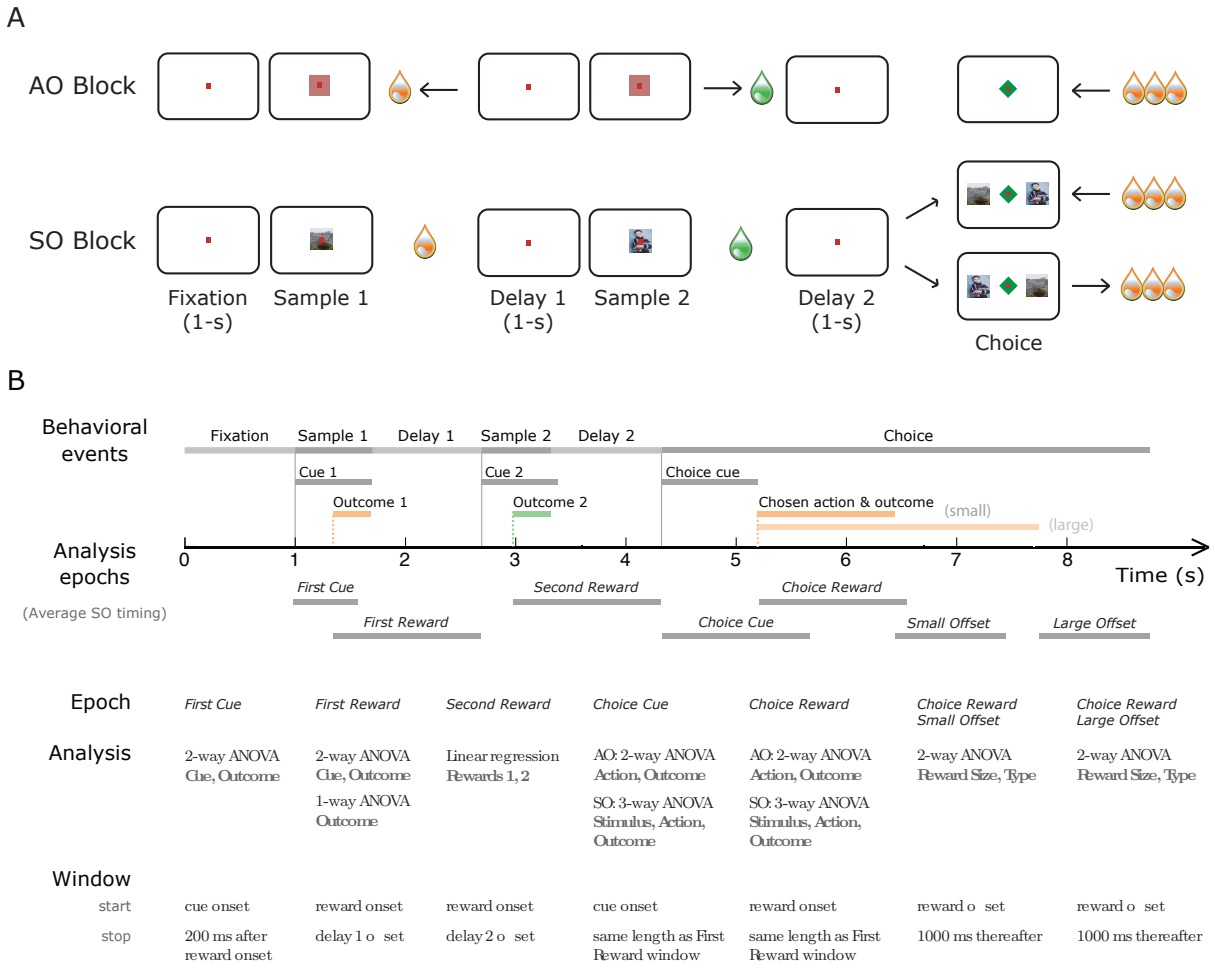

Figure 4.1: **A**, The experiment consisted of two tasks presented in alternating blocks. In the AO block, the subject moved the lever in two directions to receive two different liquid rewards. Then he repeats one of the movements to receive more of that corresponding reward. For instance, the subject could move left to receive orange juice, then right to receive apple juice. Supposing the subject prefers orange juice, he will repeat the leftward movement in the choice period to gain more orange juice. In the SO block, the subject viewed two pictures that were paired with different rewards. Then he selected the picture associated with his preferred reward via a lever movement. Pictures were equally likely to appear on either side at the choice phase. **B**, Timeline of the behavioral events on a typical SO trial. Vertical lines denote onsets of cues (duration in gray bars) and rewards (duration in colored bars). The two possible durations of the chosen reward are shown in shades of orange. The shorter one occurs 75% of the time. Time from choice cue onset to chosen reward onset was the median reaction time for subject H. Below the timeline are the analysis epochs we used time locked to the timeline. Below those are descriptions of the epochs regarding what statistical model was used and bounds for the sliding window of analysis.

block. Then we totaled the number of remaining trials. If that number was greater than 240 trials, we kept the session in our analyses.

We visualized rasters of the spikes for each neuron and removed spikes that were likely artifacts from the solenoid valve delivering rewards. For each neuron, we compared the number of spikes within 15 ms from reward onset to the number of spikes between 15 and 30 ms from reward onset. If there were more than two times as many spikes within 15 ms from reward onset, then those spikes were removed. We repeated this for reward offset. For each task, we visualized spike density histograms by averaging activity across the appropriate conditions using a sliding window of 150 ms. We quantified neuronal selectivity in each task during specified windows following several defined events of a trial. These specified windows and events are listed in Fig. 4.1B. For a given event and corresponding window, we performed a "sliding" analysis on each neuron using either an ANOVA or linear regression separated by task.

To analyze encoding of the first reward we focused on neuronal selectivity that began at reward onset and ended at the end of the first delay period. This corresponded to the period of time in which the subject needed to encode information about the reward so as to later compare it to the second reward and choose the better option for that given trial. For each neuron, we separated its data by task. In each task, we performed a "sliding" analysis. We took 200-ms bins of time, beginning at reward onset and performed a two-way ANOVA on the neuron's mean firing rate during that bin with factors of cue (which action was made in the AO task or which picture was shown in the SO task) and outcome (which reward was given). In each bin, we used our statistical model to calculate the percentage of variance (PEV) in the neuron's firing rate that was attributed to each factor. Bins were incremented by 10 ms until the end of the window for analysis was reached. We defined each neuron's selectivity based on the time of the maximal PEV obtained from that factor or interaction term. We used this same analysis to examine the time course of neuronal encoding. We determined the time of maximal encoding from the sliding analysis. We further defined the earliest latency of selectivity for that factor as the earliest time bin in which the p value fell below 0.005 for three consecutive time bins. We contrasted the strength and latency of neuronal encoding across brain areas and tasks using ANOVAs.

Following delivery of the second reward, we wanted to measure how much information neurons encoded about both the first and second reward. So for each neuron, we performed a linear regression analysis in a sliding manner starting from the onset of the second reward until the end of the subsequent delay period. The regression model compared how fluctuations in a neuron's firing rate scaled linearly with the value of the two rewards:

$$y = b_o + b_1 \times r_1 + b_2 \times r_2 \,, \tag{4.1}$$

where y was firing rate (Hz) in a given bin of time, r_1 was the normalized value of the

first reward and r_2 was the normalized value of the second reward. All variables were standardized. Normalized value for each reward term equaled one if the reward was the subject's most preferred reward and zero if it was the subject's least preferred reward. The value given for the subject's intermediately preferred reward was a ratio of how often the subject picked the more preferred rewards in all reward pairings. Specifically, the ratio followed this equation:

$$r = \frac{n_{best>intermediate}}{n_{best>intermediate} + n_{intermediate>worst}}, \tag{4.2}$$

where r is either r_1 or r_2 for trials when the intermediately preferred option was presented, $n_{best>intermediate}$ is the number of trials the subject picked the best (most preferred) reward in trials presenting the best and intermediate options, and $n_{intermediate>worst}$ is the number of trials the subject picked the intermediately preferred reward in trials presenting the intermediate and worst options.

We quantified how correlated the independent variables were to check the validity of the model. We used the variance inflation factor as our measure of collinearity. In the 47 recording sessions across the subjects, all variance inflation factors were between 1.2 and 1.6, which is well under the 2.5-valued threshold indicating a problem of collinearity [Allison, 1999].

We performed a different sliding ANOVA analysis following the choice cue onset. Trials in which the subject chose quinine were removed, as they were inconsistent with the subjects' overall preferences and we could not be certain as to the reason for the error. For example, the subject could have forgotten the cue-outcome contingencies or accidentally moved incorrectly. Trials were then separated by task. The window for the analysis was kept as the same duration used for the first reward and ranged from 1300-1500 ms. In the AO task, a sliding two-way ANOVA was performed with factors of chosen action (which direction the subject intended to move) and chosen outcome (which reward the subject had chosen). In the SO task, a sliding three-way ANOVA was performed with the same factors as in the AO task and an additional factor of stimulus (in which spatial order the pictures appeared).

In all sliding analyses, we controlled for multiple comparisons by using the adjusted p value of 0.005. We calculated this value by determining the false alarm rate during the first second of the ITI period. For each neuron, we ran a sliding ANOVA analysis on its mean firing rate with the factors of first cue and first reward shuffled across trials ten times. We then averaged those p values collapsed across task to determine the threshold at which approximately five percent of the neurons reached criterion. Only 13-14 neurons out of 464 (3.6-4%) were falsely classified as selective for each factor. Hence, our choice of threshold yielded a reasonable rate of type I errors.

We defined value neurons as those that scaled their firing rate consistent with the subject's reward preferences. For neurons selective in one task, we focused on the neuron's peak

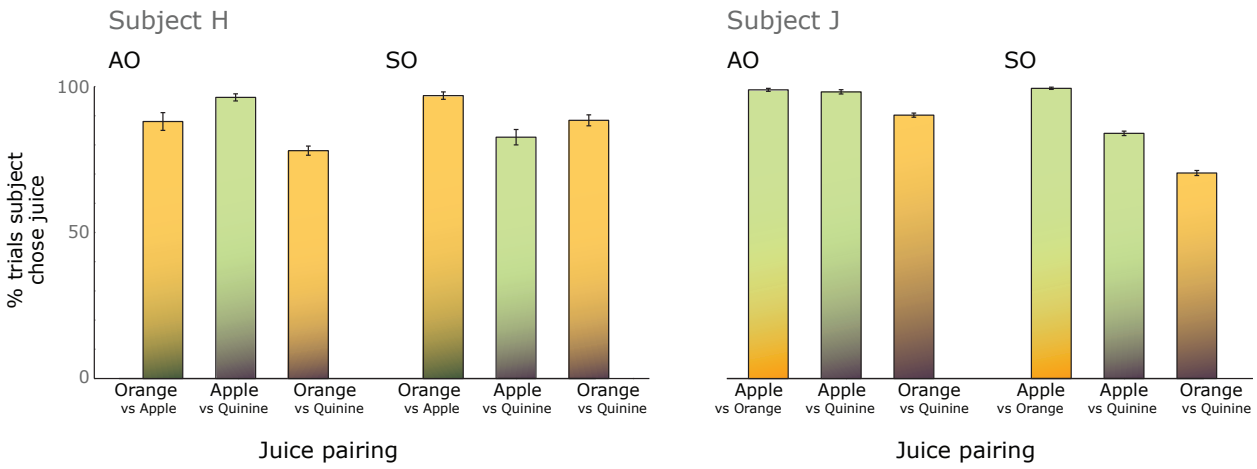

Figure 4.2: Reward preferences across sessions represented as the mean proportion of trials in which the subject selected his more preferred option in each reward pairing.

selectivity, and compared the ordering of firing rates for each reward with the subject's behavioral rank ordering of the rewards. There are six possible orderings of the three rewards, of which two were consistent with the subject's preferences (a positive or negative relationship between firing rate and preference). Those two rankings indicated value encoding. For neurons selective in both tasks, we first collapsed across tasks to determine the team of peak outcome encoding for each neuron. Proportions of value neurons were then compared across brain area and task.

4.3 Results

4.3.1 Behavior

Both subjects showed consistent preferences across both tasks (Figure 4.2). Subject H preferred orange juice to apple juice and quinine, and apple juice to quinine across 27 sessions. Subject J preferred apple juice to both orange juice and quinine, and orange juice to quinine across 20 sessions. To account for possible deterioration in choice behavior towards the end of sessions, we performed a three-way ANOVA on the percentage of choices that were consistent with the subject's preferences. The factors were reward pair, task, and quartile of a session. Subjects' choices depended on the interaction between pairings and task ($F > 8.1, p \leq 0.0003$ for both subjects). Subject H had strongest preferences for orange over apple in the AO task (97 ± 1.2), followed by orange over quinine in the AO task (96 ± 1.1), orange over apple in the SO task (89 ± 2.3), and orange over quinine in the SO task ($82 \pm$

2.2). Weakest preferences were for orange over apple in the SO task (78 ± 1.7). All simple effects were significant (One-way ANOVA, $F > 7.6, p \leq 0.006$; t-test, $t_{107} > 3.9, p < 2 \times 10^{-4}$). Subject J had similar preferences in the AO task (apple over quinine: 99 ± 0.4, apple over orange: 99 ± 0.6, orange over quinine: 98 ± 1.0; simple effects one-way ANOVA, $F_{(2,237)} = 0.75, p > 0.47$). In the SO task, he had strongest preferences for apple over quinine (91 ± 1.0), followed by apple over orange (84 ± 1.1), and weakest preference for orange over quinine (71 ± 1.3, simple effects one-way ANOVA, $F_{(2,237)} > 81, p < 1 \times 10^{-15}$). Further, all simple effects differentiating task levels for each level of reward pairing were significant (t-test, $t_{79} > 9.0, p < 6.6 \times 10^{-14}$). Quartiles had only a limited effect. Specifically, the interaction of reward pairing and quartile affected subject H ($F_{(6,624)} > 2.6, p < 0.02$), but not subject J ($F_{(6,456)} = 0.65, p > 0.69$). It was limited to the apple-quinine pairing (Simple effects one-way ANOVA, $F_{(3,212)} > 5.3, p < 0.002$) in that preferences dropped across the session (quartile 1: 92 ± 1.6%, 2: 92 ± 1.6, 3: 85 ± 2.7, 4: 81 ± 3.1).

Neither subject made many errors. We excluded error trials, which were defined one the basis of two types of errors: 1. Subject broke fixation during a period subsequent to the presentation of the first reward, or 2. Subject made a wrong lever movement during the second sample period in the AO task. The mean number of fixation errors (± S.E.M.) per session was 10 ± 2 for subject H and 68 ± 3 for subject J. Across both tasks, the fixation errors were equally likely to occur in either delay for subject H (t-test, $t_{52} < 0.2, p > 0.7$). Subject J tended to break fixation more often in the first delay during the SO task (t-test, $t_{38} < 2.1, p = 0.046$). In the AO task, however, he tended to break fixation during the second delay (t-test, $t_{38} < -3.1, p < 0.003$). Fixation errors also did not depend on the reward just given (one-way ANOVA; subject H: $F_{(2,78)} < 2.6, p > 0.08$; subject J: $F_{(2,57)} < 2.8, p > 0.07$), except in one case. The exception was delay 2 during SO trials for subject H: he broke fixation most often following his least preferred outcome, quinine, and least often following his most preferred outcome, orange juice (one-way ANOVA, $F_{(2,78)} > 5.6, p < 0.005$). As for moving the wrong way in the sample periods of the AO task, the mean number of errors per session was 17 ± 2 for subject H and 18 ± 1 for subject J. Most errors happened for the first movement of the trial (t-test, $t_{52} > 7, p < 5 \times 10^{-9}$ for both subjects). The likelihood of making wrong movements did not depend on the identity of the reward (one-way ANOVA, $F_{(2,78)} < 1.3, p > 0.2$). After removing all error trials, subject H completed a mean of 138 ± 35 correct trials per session in the AO task and 170 ± 28 correct trials in the SO task. Subject J completed 135 ± 18 correct trials in the AO task and 199 ± 31 correct trials in the SO task.

We removed sessions, in which subjects showed significant side biases (see section 4.2.2 Methods for more details). In the remaining sessions used for analysis, we verified that both subjects showed reward preferences and did not strongly prefer one movement over the other. Subject H selected one lever movement significantly more than the other on only 21% of the trials within a session. For subject J, it was 13%. Of the bias trials, they tended to occur during the SO task (H: 56%, J: 59%, t-test, $p < 2 \times 10^{-6}$ for both subjects).

Subject H gradually responded faster to successive movements in an AO trial, taking 456 ms (median) for the first response, 336 ms for the second response, and 199 ms for the choice response. The choice response was much faster than that in the SO task (median 744 ms, Wilcoxon's rank-sum test, $p < 1 \times 10^{-15}$). Subject J was also slowest at the first response of the AO task (median 333 ms), faster at the second response (median 286 ms) and fastest at the choice response (median 266 ms). AO choice responses were again much faster than the choice response in the SO task (median 709 ms, Wilcoxon's rank-sum test, $p < 1 \times 10^{-15}$). These results suggest that the choice phase of the SO task involves a qualitatively different process than the choice phase of the AO task. This is likely because in the AO task the subject can determine the appropriate response to obtain his preferred juice immediately after the delivery of the second reward. In contrast, in the SO task they must wait until the pictures appear before they can determine the appropriate response. This involves processing the pictures, recalling their association with the rewards and then determining the response.

Subject H had a small bias in the AO task to perform leftward choice movements more quickly (median left: 193 ms, median right: 204 ms, Wilcoxon's rank-sum test, $p = 0.05$). He had the same bias in the SO task (median right: 733 ms, median left: 754 ms, Wilcoxon's rank-sum test, $p < 2 \times 10^{-5}$). He showed no reaction time difference across the two directions during the first sample response of the AO task (median left: 453 ms, median right: 459 ms, Wilcoxon's rank-sum test, $p > 0.65$). There was a difference during the second sample response with him being faster for rightward movements (median left: 353 ms, median right: 311, Wilcoxon's rank-sum test, $p < 4 \times 10^{-25}$). Like subject H, subject J was also faster choosing leftward responses in both tasks (median AO: left 247 ms, right 283 ms; SO: left 579 ms, right 826 ms; Wilcoxon's rank-sum test, $p < 1 \times 10^{-15}$ for both cases). His reaction times on the sample responses of the AO task were faster for rightward movements in the first sample epoch (median left: 355 ms, median right: 314 ms, Wilcoxon's rank-sum test, $p < 2 \times 10^{-7}$) but there was no difference in the second sample epoch (median left: 286 ms, median right: 286 ms; Wilcoxon's rank-sum test, $p > 0.84$).

The AO choice movement reaction time scaled with subject H's reward preferences (Kruskal-Wallis one-way ANOVA, $\chi^2 > 221, p < 1 \times 10^{-15}$). It was quickest following the subject's favorite juice as the preceding second sample juice (median 181 ms), moderately fast following the subject's intermediately preferred juice (median 190 ms), and slowest following the subject's least preferred juice (median 208 ms). The SO choice movements also differed (Kruskal-Wallis one-way ANOVA, $\chi^2 > 51, p < 6 \times 10^{-12}$). From a post-hoc simple effects test, we found that actions were only significantly faster following subject H's favorite reward compared to either of the other two: favorite juice median 727 ms, intermediately preferred juice median 740 ms, least preferred juice median 744 ms (Kruskal-Wallis one-way ANOVA, $\chi^2 > 24, p < 7 \times 10^{-7}$ in both cases). Subject J's choice movements were also affected by previous rewards given in the second sample period. In the AO task, he was fastest to respond following his least preferred quinine option (259 ms) and slowest following his most preferred apple juice option (270 ms) (Kruskal-Wallis one-way ANOVA, $\chi^2 > 7, p < 0.03$). In the SO

task, he was fastest after his intermediately preferred option (692 ms) and slowest after his least preferred option (730 ms) (Kruskal-Wallis one-way ANOVA, $\chi^2 > 13, p < 0.002$).

4.3.2 Neurophysiology

We recorded the activity of 215 neurons (H: 125, J: 90) from the dorsal bank of the cingulate sulcus (PFcs) and 249 neurons (H: 145, J: 104) from areas 11 and 13 in the orbitofrontal cortex (PFo) while our subjects performed the AO and SO choice tasks. In order to choose optimally, the subjects needed to remember which juice outcome was paired with which stimulus or action.

4.3.3 Encoding first outcome following its delivery

In common to both tasks was the requirement to remember the first juice across the first delay and so our analyses begin with this time period. Many of the neurons showed similar encoding of the outcome across the two tasks. Figure 4.3A illustrates an example of a PFo neuron that responded to the delivery of quinine in both the SO and AO task. However, other outcome-selective neurons showed very different patterns of selectivity between the two tasks. Figure 4.3B illustrates a PFcs neuron that encoded the first outcome during the AO task but did not fire at all during the SO task. In contrast, figure 4.3C illustrates a PFo neuron that encoded outcome during the SO task but not during the AO task. We quantified the prevalence of neurons that encoded the first juice outcome using a sliding one-way ANOVA (see section 4.2.2 Methods). A similar proportion of neurons encoded outcome in either one or both of the tasks, and the proportions did not differ between brain areas (Table 4.1).

We also examined the extent to which neurons encoded the value of the outcome, in which case encoding of the outcomes should match the subjects' preferences. For each neuron, we determined the time point of maximum outcome selectivity and then calculated how it rank ordered the juices and compared this to the subject's behavioral rank ordering of the juices. There are six possible ways to rank order three juices, two of which are consistent with the subject's preferences (a positive or a negative relationship between firing rate and preference) and therefore indicative of value encoding. For neurons selective in one task, we focused solely on the task in which they encoded the outcome. For neurons selective in both tasks, we collapsed activity across the tasks and then determined the order in which the neuron ranked the outcomes. Since two of the six possible rankings of three juices were consistent with subjective preference, we would expect by chance a third of the neurons showing these two rankings. Instead we found that a majority of the neurons that encoded outcomes across both tasks or in the SO task reflected subjective preferences (Fig 4.4). The exception was with the neurons in the AO task where there was a variety of different rankings, suggesting

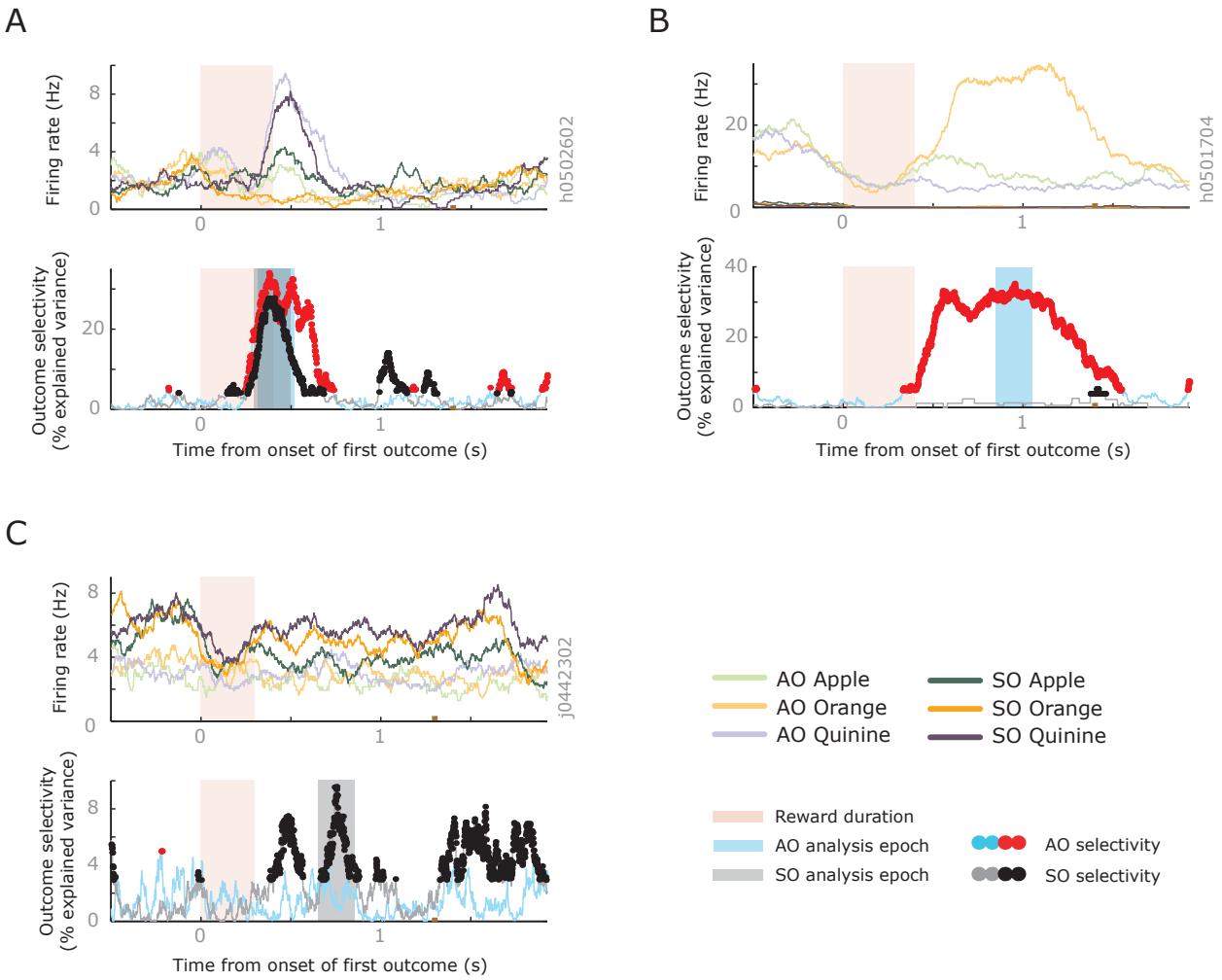

Figure 4.3: **A**, Spike density histogram illustrating a PFo neuron encoding the first reward across both tasks. It had its highest firing rate when the reward was quinine (purple lines) and its lowest firing rate when the reward was orange juice (orange lines). The magnitude of outcome selectivity is shown below the histogram. Significant encoding is denoted with red and black dots in the SO and AO tasks, respectively. Brown tick marks illustrate the end of the delay period. **B**, A PFcs neuron that encoded outcomes only in the AO task. **C**, A PFo neuron that encoded outcomes only in the SO task.

Table 4.1: Neurons with significant encoding for cue or outcome following the first reward. Rows are sorted by which task neurons were selective in and from which brain area neurons were recorded. In the top panel, percentages are out of all recorded neurons. In the bottom panel, percentages are from outcome-selective neurons. The columns of the top panel lists neurons that had significant main effects of cue or outcome or a significant interaction. There were no differences in such selective neurons across brain areas. In the bottom panel, outcome-selective neurons are further divided according to whether they also encoded the cue and according to whether their cue encoding fell in the same or different task as their outcome encoding. Outcome neurons selective in one task carrying cue information tended to do so in the same task as their outcome encoding. However, outcome neurons selective across both tasks rarely had cue encoding in both tasks.

	Cue	**Outcome**	**Cue x Outcome**
AO only: PFo	28 (11.2%)	43 (17.3%)	15 (6.0%)
AO only: PFcs	27 (12.6%)	24 (11.2%)	22 (10.2%)
SO only: PFo	17 (6.8%)	43 (17.3%)	22 (8.8%)
SO only: PFcs	11 (5.1%)	39 (18.1%)	10 (4.7%)
Both: PFo	4 (1.6%)	32 (12.9%)	1 (0.4%)
Both: PFcs	4 (1.9%)	37 (17.2%)	0 (0.0%)

	Outcome + no cue selectivity	**Outcome + same pattern cue selectivity**	**Outcome + different pattern cue selectivity**
AO only: PFo	22 (51.2%)	20 (46.5%)	1 (2.3%)
AO only: PFcs	15 (62.5%)	9 (37.5%)	0 (0.0%)
SO only: PFo	34 (79.1%)	9 (20.9%)	0 (0.0%)
SO only: PFcs	30 (76.9%)	5 (12.8%)	4 (10.3%)
Both: PFo	18 (56.2%)	3 (9.4%)	11 (34.4%)
Both: PFcs	18 (48.6%)	3 (8.1%)	16 (43.2%)

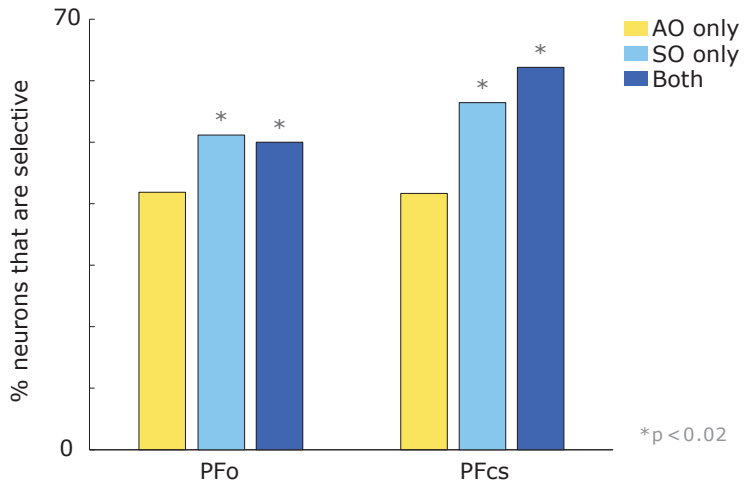

Figure 4.4: Value encoding in outcome neurons at the maximal point of outcome encoding. Outcome neurons in both brain areas that were selective in both tasks or in the SO task only encoded rewards consistent with subjective preference. However, the proportion of AO outcome neurons encoding rewards consistent with preferences did not significantly differ from the proportion expected by chance.

that neurons may have encoded the specific outcome associated with the movement rather than its value.

We next turned our attention to other measures of outcome encoding. We began by looking at the strength of encoding, which was measured as the maximal percent explained variance in mean firing rate attributable to outcome ($PEV_{outcome}$). To compare encoding strength across populations of neurons, we ran a three-way ANOVA on $PEV_{outcome}$ with factors of brain area, task, and whether the cell was selective in one or both tasks. As illustrated in figure 4.5A, outcome encoding of neurons selective in both tasks was significantly stronger than that of neurons selective in one task ($F_{(1,279)} > 38, p < 1 \times 10^{-15}$). In addition, the AO task carried stronger outcome encoding than the SO task ($F_{(1,279)} > 5.2, p < 0.023$). In short, outcome selectivity is more robust in the AO task and strongest when it is present in both tasks.

We next compared the temporal evolution of these signals. We compared the latencies of maximal outcome encoding via a three-way ANOVA analysis again using factors of brain area, outcome and whether neurons were selective in one or both tasks. Maximum outcome encoding occurred slightly earlier in PFo compared to PFcs (PFo: 702 $\pm$ 25 ms, PFcs: 763 $\pm$ 26 ms; $F_{(1,255)} > 3.7, p = 0.053$). We also examined how quickly neurons began encoding outcome information. To measure this, we determined each neuron's earliest latency for outcome selectivity. Earliest latency was defined as the first 200-ms window after reward onset, that the neuron remained selective for three consecutive time bins. We ran a three-

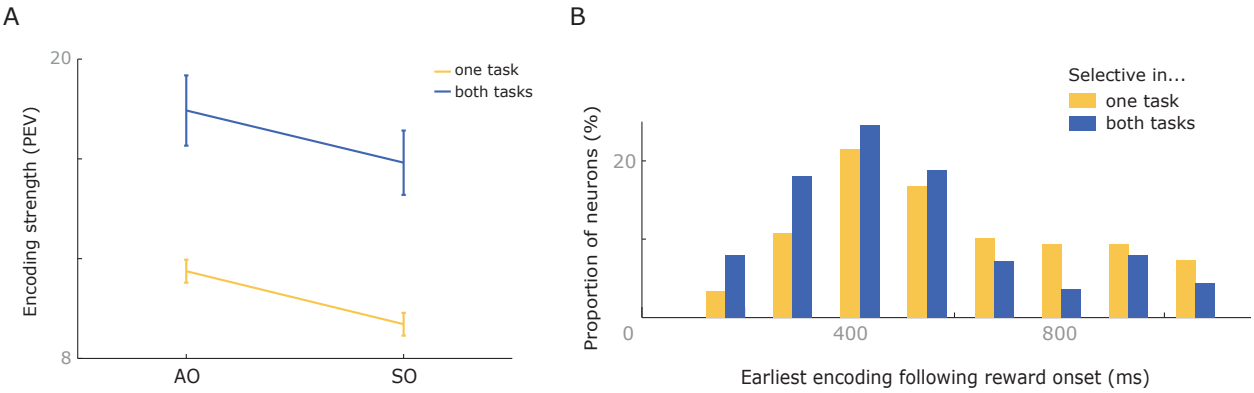

Figure 4.5: **A**, Strength of maximal outcome encoding in neurons that are outcome selective during one or both tasks. Encoding was much stronger in neurons selective in both tasks than one task ($p < 1 \times 10^{-15}$). Encoding in the AO task was also reasonably stronger than the SO task ($p < 0.023$). **B**, Distribution of earliest encoding of outcome in selective neurons. Outcome neurons selective in both tasks (blue) tended to have earlier latencies than those selective only in one task (gold). Significant $p = 0.0008$.

way ANOVA of those latencies with factors of brain area, task, and whether the neuron encoded outcome in one or both tasks. As shown in figure 4.5B, neurons encoding outcome in both tasks were significantly faster than those encoding outcome merely in one task (mean $\pm$ standard error: 557 $\pm$ 25 ms for both task neurons, 672 $\pm$ 25 ms for one task neurons; $F_{(1,276)} > 11, p = 0.0008$). These data suggest that both tasks utilize a task-independent reward signal that is broadcast quickly and simultaneously to PFo and PFcs.

4.3.4 Encoding task cues by outcome neurons

The subject not only had to remember what outcome he received in order to guide his choice, he also had to link it with the associated task cue. One method neurons could do this is by simultaneously encoding outcome and cue. To see if this occurred, we looked for cue encoding during maximal outcome encoding. We began by isolating the outcome neurons that were selective for cue. Nearly 40% (81/218) of outcome neurons also encoded the cue with most of them (49/81 or 60%) encoding the cue simultaneously. See Table 4.1 for more details. These cue-outcome selective neurons were evenly divided across brain area as seen in figure 4.6A.

Because of the nature of the task, it was possible that the cue encoding we observed could reflect neuronal activity encoding the expected second cue, rather than first cue. This would require a very different interpretation of the cue-outcome selective neurons. To distinguish

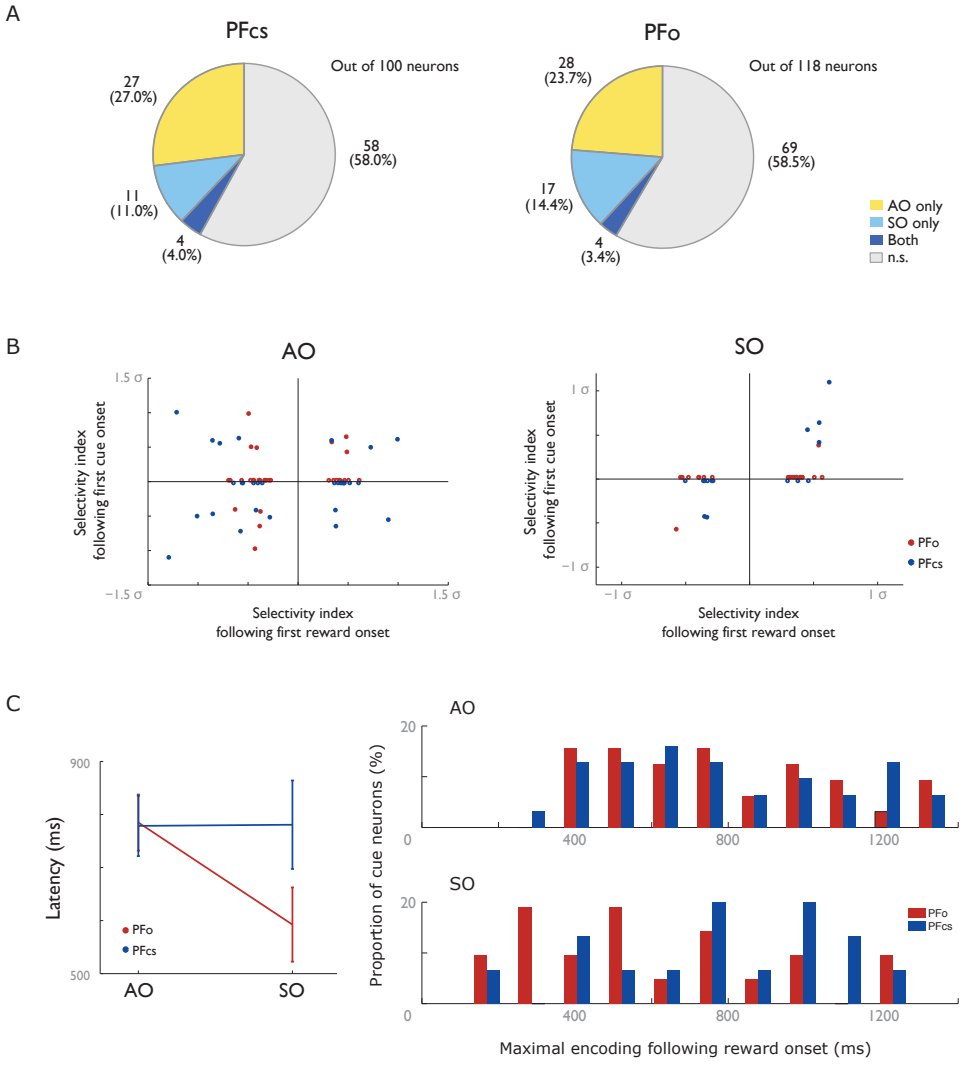

Figure 4.6: **A**, Proportion of outcome neurons that also had cue encoding in the AO task (yellow), SO task (light blue), or both tasks (dark blue). Roughly 40% of outcome neurons in each brain area also carried significant cue encoding. **B**, Direction index prior to and following reward onset of the cue-selective outcome neurons. Neurons are denoted as circles. Those lacking cue selectivity prior to reward onset are unfilled circles. They have been offset from the abscissa to enable visualization of their direction index. There were more neurons with consistent direction indices before and after reward, but that was only significant in the SO task ($p < 1 \times 10^{-15}$). **C**, Latencies for encoding cue and outcome differed across tasks and brain areas. In the AO task, latencies were roughly the same in PFcs (blue) and PFo (red), whereas in the SO task, these latencies were clearly faster in PFo than PFcs ($p < 0.01$).

this possibility, we computed a cue index for neurons at the time of maximal outcome encoding and compared it to a cue index taken from the epoch prior to reward onset. Specifically, we performed a sliding ANOVA to calculate the point of maximal cue encoding from the period of cue presentation until 200-ms after the onset of the reward. (The 200-ms took account of the latency for reward information to reach prefrontal cortex [Lara *et al.*, 2009]). The cue index was the normalized difference in a neuron's firing rate during trials of one cue (leftward movement or picture A) and trials of the other cue (rightward movement or picture B). There were more neurons showing consistent cue encoding before and after the reward onset (figure 4.6B): 16/26 (62%) and 8/8 (100%) in the AO and SO tasks, though it was only significant in the SO task (binomial test, AO: $p = 0.08$, SO: $p < 1 \times 10^{-15}$). Thus, it is possible these neurons could be combining the cue and outcome information necessary for performing the experiment optimally.

Next we compared the encoding strengths and latencies of combined cue and outcome information across neurons. We isolated the outcome selective neurons. Then we used a three-way ANOVA to detect differences due to factors of task, area, and whether or not these neurons encoded cues in one or both tasks. For encoding strength, we found no differences ($p > 0.1$). The ability to encode cue and outcome was similar across tasks, brain areas, and the number of tasks in which cue encoding occurred. The latencies, on the other hand, showed an interaction between task and brain area (figure 4.6C). During the AO task, cue-outcome neurons in both brain regions took about 780 ms following reward onset to reach maximal outcome encoding. In contrast, during the SO task, cue-outcome neurons from PFo were faster. They took 590 ms and cue-outcome neurons from PFcs took 780 ms ($F_{(1,91)} > 7, p < 0.01$). These data suggest PFo and PFcs carry both cue and outcome information on a single neuron level, but that this information for the SO task is first combined in PFo.

4.3.5 Encoding outcomes following the second reward

Representation of the first reward was very similar across brain regions. However, there were hints of task-related differences. Outcomes were more strongly represented following lever movements than pictures, and cue-outcome selective neurons from PFo were faster to reach maximal outcome encoding in the SO task than such neurons from PFcs.

We next turned our attention to the second reward epoch. During this epoch, the subject learned the identity of the second reward and could begin comparing both rewards so as to determine which associated cue he would later choose in the trial. We were interested in knowing how much information about each reward neurons encoded. For each neuron, and each task in turn we performed the following linear regression in a sliding analysis starting from the onset of the second reward until the end of the subsequent delay period:

$$y = b_o + b_1 \times r_1 + b_2 \times r_2 \,, \tag{4.3}$$

Table 4.2: Neurons with significant outcome encoding following the second reward. Rows are sorted by which task neurons were selective in and which brain area neurons came from. First column percentages are out of all recorded neurons from the given region. Percentages for the rest of the columns are out of all significant neurons in that given category (first column total). Approximately 15% of the neurons in each brain area showed each type of selectivity (AO only, SO only, Both).

	Significant	**Positive betas**	**Negative betas**	**Anti-correlated betas**
AO only: PFo	43 (17.3%)	14 (32.6%)	21 (48.8%)	8 (18.6%)
AO only: PFcs	36 (16.7%)	10 (27.8%)	12 (33.3%)	14 (38.9%)
SO only: PFo	40 (16.1%)	11 (27.5%)	7 (17.5%)	22 (55.0%)
SO only: PFcs	37 (17.2%)	5 (13.5%)	16 (43.2%)	16 (43.2%)
Both AO: PFo	30 (12.0%)	5 (16.7%)	12 (40.0%)	13 (43.3%)
Both AO: PFcs	39 (18.1%)	7 (17.9%)	11 (28.2%)	21 (53.8%)
Both SO: PFo	30 (12.0%)	5 (16.7%)	13 (43.3%)	12 (40.0%)
Both SO: PFcs	39 (18.1%)	10 (25.6%)	12 (30.8%)	17 (43.6%)

where y was firing rate (Hz) in a given sliding window, r_1 was the normalized value of the first reward and r_2 was the normalized value of the second reward. (See 4.2.2 Statistical methods.) We first checked that our model was a reasonable one without too much correlation in the independent variables. This was a concern, since there is some correlation given that the second reward was never the same as the first. As the measure of collinearity, we used the variance inflation factor. In the 47 recording sessions across the subjects, all variance inflation factors were between 1.2 and 1.6, which is well under the 2.5-valued threshold indicating a problem of collinearity [Allison, 1999]. Hence, we could interpret the results from this model. From the sliding analysis, we isolated the time when each neuron best fit the model and looked to see if the fit was significant.

Roughly 15% of the recorded neurons from each brain area had significant fits to the linear model and therefore carried reward information. Table 4.2 shows the breakdown. We noticed that a number of these neurons modulated their firing rate in opposite directions to the value of the rewards. For example, the neuron in figure 4.7A had highest firing rate when the first reward was orange juice and lowest firing rate when the first reward was quinine ($R^2 > 0.06, p \leq 0.005$ in both tasks). However, when conditions were grouped by the second reward, this neuron had highest firing for quinine and lowest firing for orange. Figure 4.7B shows a similar neuron, except its selectivity was limited to the SO task ($R^2 = 0.08, p < 0.002$). This contrasting behavior to the rewards meant the neuron had one positive and one

Figure 4.7: **A**, Spike density histogram illustrating a PFcs neuron that captured relative reward value as a linear difference of the rewards' values. It had highest firing rate when the first reward was orange juice (orange lines) and lowest firing rate when the first reward was quinine (purple lines). When conditions were grouped by the second reward, this neuron had highest firing for quinine and lowest firing for orange. Brown tick marks illustrate the end of the delay period. **B**, A PFcs neuron that also captured relative reward value, but only in the SO task.

negative beta coefficient. Thus, these neurons could potentially be involved in the decision process, comparing the value of the two rewards by calculating their difference in value.

To determine how many neurons encoded relative rewards as linear differences, we plotted the beta coefficients of all significant neurons and sorted neurons based on the sign of their beta coefficients to see if they were anti-correlated, i.e., had opposite sign. Of the PFo neurons selective in the AO task, 21/73 (29%) had anti-correlated beta coefficients. In contrast, 34/70 (49%) of the PFo neurons selective in the SO task had anti-correlated coefficients ($\chi^2 > 5, p < 0.025$). This significant difference fits with the idea that PFo is sensitive to the context of reward value, in this case whether value related to motor or visual cues. PFcs neurons lacked such task difference. They were evenly split with 35/75 (47%) having opposite signs in the AO task and 33/76 (43%) in the SO task ($\chi^2 < 0.4, p > 0.5$).

Aside from variations in encoding strength, we also examined whether there were differences in the time that encoding emerged. We compared both the time of maximal reward

encoding as well as the earliest reward encoding using a three-way ANOVA with factors of task, brain area, and whether or not the neuron was selective in one or both tasks. The analysis yielded no significant differences. Hence, the temporal dynamics of outcome encoding were similar across brain regions and task following the second reward.

4.3.6 Encoding stimulus, action and outcome at choice cue onset

Our analyses thus far have revealed subtle distinguishing features in the function of PFo and PFcs across the tasks. PFo carried simultaneous stimulus and outcome information faster than PFcs in the SO task. Both areas represented relative value of the two rewards in a linearly manner across both tasks, but only in PFo were there more relative encoding neurons limited to the SO task as compared to those limited to the AO task.

We continued examining regional and task-related variations further in the trial to when the choice cue came on. This event had marked differences across the tasks. In the AO task, no decision necessarily occurred. The subject could have planned his action already. But in the SO task, the subject could not plan ahead. He needed to observe both pictures presented side by side, select the one that was associated with the more preferred option, and then plan his action accordingly. To characterize neurons during this process, we ran two sliding ANOVA analyses per neuron separated by task. For AO trials, we used a two-way ANOVA with the dependent variable of mean firing rate and factors of chosen action and chosen reward. We located the 200-ms bin in which each factor had its maximal encoding. For SO trials, we used a three-way ANOVA with the additional factor of stimulus, the spatial order of the pictures. In both analyses, we removed trials in which quinine was picked since this occurred too infrequently to meaningfully analyze the data. We then compared the number of neurons with main effects for each factor. Roughly 5-10% of the neurons in either brain area encoded stimulus, and 5-20% encoded outcomes in some form. The striking difference arose in action encoding with most neurons, particularly in PFo, selective only in the SO task. As seen in figure 4.7, over 40% (108/249) of all recorded PFo neurons encoded actions only during the SO task. This was more than the proportion of PFcs neurons selective only in the SO task (64/215 or 30%). Both proportions were far greater than either proportion selective in just the AO task or both tasks. These action neurons were not simply motor neurons though. They were not active in the AO task, which also required a motor action to be executed. Further, they were not encoding the spatial location of the better reward either, because the location was equally likely to occur on either side. Instead they could be encoding the output of a decision-making process. In the AO task, the subject could have made his decision earlier in the trial; however, in the SO task, the subject could only plan his response when he saw the two pictures and located the one associated with his preferred reward. Hence, these PFo neurons encoded the means, by which to achieve the subject's goal during the SO task.

It is possible that neurons first locate stimuli within the environment, and then plan the

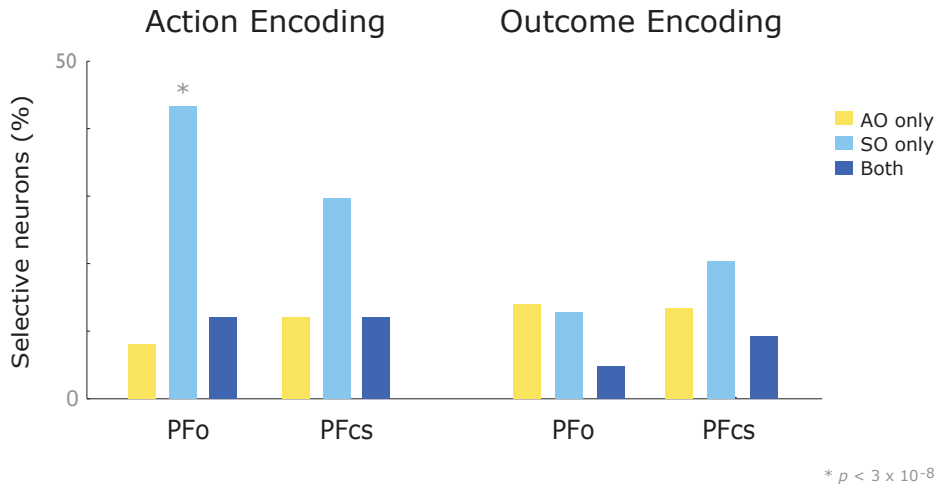

Figure 4.8: Prevalence of action and outcome encoding following presentation of the choice cue in the SO task, i.e., the two pictures presented side-by-side. PFo neurons with action encoding limited to the SO task were more common than any other class of neuron ($p < 3 \times 10^{-8}$). Not shown: less than 10% of the neurons selected for stimulus.

necessary movement towards the better stimulus. If this is the case, we should see stimulus encoding sooner than action encoding. To test this theory we compared the timing of action encoding with the encoding of stimulus and outcome. We performed a two-way ANOVA on the maximal encoding latencies with the factors of which of the three variables neurons encoded and which brain area neurons came from. We found the factor for which variable neurons encoded to be significant (two-way ANOVA, $F_{(2,563)} = 6.5, p < 0.002$). Further analysis of simple effects showed that the latency for actions was similar to that of stimulus but faster than outcome (660 ± 17 [action], 726 ± 56 [stimulus], 763 ± 25 [outcome]; one-way ANOVA, $F_{(2,566)} > 6.2, p < 0.003$). As seen in figure 4.9, both action and stimulus information appeared roughly at the same time. Thus, our data do not support a sequential process in which the picture associated with the preferred juice is first located and then the correct response determined. Rather, these processes appear to occur simultaneously and in parallel, at least within PFo and PFcs.

4.3.7 Encoding action and outcome after choice

Given the subject had to execute a decision in the form of a lever movement following the choice cue in the SO task, our observation of prevalent action encoding right after that cue seemed logical. We would further expect a drop in such encoding after reward was delivered, since the decision had completed. To measure this, we ran the same sliding ANOVA analyses used for each task preceding reward delivery on a window of the exact same size but starting

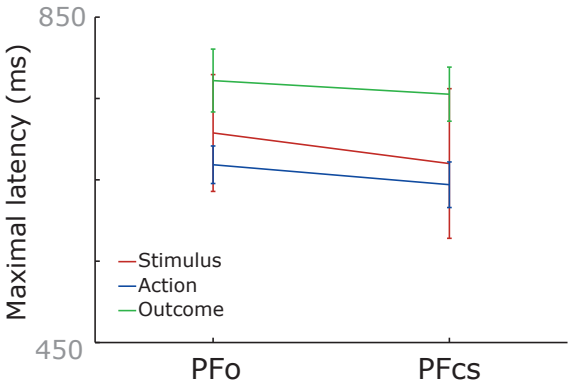

Figure 4.9: The time neurons took to reach their maximal encoding was quickest for actions (blue). Those latencies were significantly faster than those for outcomes (green).

at chosen reward onset. Indeed, after the decision had been executed, significantly fewer neurons selected for actions only in the SO task. Numbers are shown in table 4.3. The drop was significant for both PFo and PFcs populations ($\chi^2 > 9.4, p < 0.003$ in both cases). There was no disparity in the number of neurons encoding action only in the AO task or across both tasks before and after chosen reward onset. Nor was there any difference in the number of neurons encoding stimulus or outcome ($\chi^2 < 1.9, p > 0.1$ in all cases). This data highlights the transience of action encoding limited to the SO task, which coincides with the execution of the decision.

If neurons specific to the SO task were most active during the execution stage of a decision, we would hypothesize fewer to encode subsequent parameters in the experiment after the chosen reward was given. We probed this issue by characterizing neurons to see if they encoded the size of the chosen reward, which on 25% of the time yielded a much bigger reward. We also tracked if neurons encoded the identity of the chosen reward. We ran a sliding two-way ANOVA analysis with dependent variable of mean firing rate and factors of reward identity (most or intermediately preferred juice) and reward size (small or big) during a 1000-ms long window starting at the small offset of the chosen reward. Comparing selectivity only in the SO task, there was a drop in the number of neurons selective for reward identity: from 73/464 (16%) following cue onset to 46/464 (9.9%) during reward delivery ($\chi^2 > 7, p = 0.008$). Roughly equal proportions of neurons encoded reward size as reward identity. The difference became significant when all task selectivity was grouped. 204/464 (44%) neurons encoded size, and 111/464 (24%) neurons encoded identity ($\chi^2 > 41, p < 2 \times 10^{-10}$).

Next we examined encoding strength of reward size and reward identity across brain regions, tasks, and whether neurons were selective in one or both tasks. Strength measures were entered into a three-way ANOVA as the dependent variable, and the factors were brain

Table 4.3: Neurons with significant task-relevant encoding following delivery of the chosen reward. Rows are sorted by which task neurons were selective in and which brain area neurons came from. Columns correspond to which task-relevant event neurons encode: stimulus (spatial location of pictures), action (chosen movement), and outcome (chosen reward). Fewer neurons in both areas encoded actions once the action had been completed and reward delivered.

	Stimulus	**Action**	**Outcome**
AO only: PFo	–	20 (8.0%)	35 (14.1%)
AO only: PFcs	–	26 (12.1%)	29 (13.5%)
SO only: PFo	24 (9.6%)	108 (43.4%)	32 (12.9%)
SO only: PFcs	14 (6.5%)	64 (29.8%)	41 (19.1%)
Both: PFo	–	30 (12.0%)	12 (4.8%)
Both: PFcs	–	26 (12.1%)	20 (9.3%)

area (PFo or PFcs), task (AO or SO), and the neurons' selectivity (one or both tasks). There was a significant three-way interaction with encoding strength for reward size varying with all three factors ($F_{(1,262)} = 4.2, p < 0.042$). We then performed simple effects analyses holding each factor constant at each level and then computing the resulting two-way ANOVA. Significant area differences were limited to the SO task and varied with selectivity (significant interaction, $F_{(1,118)} = 4.6, p = 0.034$): PFo neurons encoding reward size across both tasks had strongest selectivity ($6.4 \pm 0.2\%$), followed by PFcs neurons encoding size in one task ($5.7 \pm 0.2\%$). PFo neurons encoding size in one task had weakest selectivity ($5.5 \pm 0.2\%$). Reward identity was represented more strongly in PFo than PFcs, and this trend was more pronounced for neurons selective in both tasks (Three-way ANOVA, significant area by selectivity interaction, $F_{(1,113)} > 10, p < 0.002$). We continued by analyzing encoding latencies for reward size and identity. For size and identity in turn, we ran the same three-way ANOVA analysis with the corresponding maximum latencies as the dependent variable. We found that encoding of reward size was faster in PFo at 498 ± 17 ms than PFcs at 569 ± 20 ms ($F_{(1,262)} > 7, p < 0.008$). In sum, PFo remains fairly active in monitoring parameters related to reward even after a decision.

4.4 Discussion

Obtaining the optimal goal often requires assessing the value of environmental cues such as visual stimuli, as well as one's own actions. We hypothesized that values associated with stimuli and actions utilize PFo and PFcs, respectively. Contrary to our hypothesis, there was not a clear-cut distinction between PFo and PFcs in terms of encoding these SO and AO associations. Both areas encoded stimulus, action, and outcome information. However, there

were some differences that appeared to support a role for PFo in preferentially encoding SO associations. Neurons selective for the first stimulus and reward were faster to encode these pieces of information within PFo than PFcs. Most surprisingly, PFo neurons showed robust encoding of the chosen actions during the choice period of the SO task.

4.4.1 Overall similarity across brain areas in outcome encoding

We did not support our hypothesis of a double dissociation across brain areas in encoding AO and SO information. Instead we found that both PFcs and PFo represented AO and SO associations. It is particularly surprising that we observed AO encoding in PFo, as it has not been reported before in monkeys. Rather, PFo studies in monkeys have found robust neuronal selectivity for SO associations with little selectivity for AO associations [Tremblay and Schultz, 1999; Wallis and Miller, 2003b; Padoa-Schioppa and Assad, 2006]. In contrast, rat studies have found action encoding in PFo [Feierstein *et al.*, 2006; Furuyashiki *et al.*, 2008; Sul *et al.*, 2010]. However, these studies differ from the monkey studies in behavioral testing procedures, which could explain the different encoding patterns in PFo across species. For instance, rat studies tend to place rewards in different spatial locations, which require different actions to reach them, but also manifest as different stimuli. Our findings instead suggest that PFcs and PFo can flexibly encode action- and stimulus-outcome pairings depending on task demands.

Neuronal representations may reflect the relative attention placed across the task variables, i.e., stimuli, actions, and outcomes. Greater attention to a particular parameter could yield greater neuronal encoding of that parameter within the physiological bounds of a neuron. Another formulation of this idea is that neuronal representations reflect the set of prior probabilities of the task variables being utilized during the experiment. Suppose an experiment required most decisions be computed from how rewarding a set of movements and few decisions be computed from how rewarding a single colored cue is, then there would be more prior probabilities in the set of action-related priors than the set of stimulus-related priors. Having more elements related to a particular task variable should increase neuronal representation for that variable. In perceptual decision-making tasks where subjects must determine which direction semi-randomized dots are moving on a computer screen, subjects will perceive dots moving more often in a certain direction if that direction tends to be the correct answer. On a neuronal level, a majority of parietal neurons correlate with the prior probability of movements [Platt and Glimcher, 1999]. PFC neurons may show similar correlations in a decision-making experiment given the breadth of task details they encode. Hence, experiments that emphasize SO associations should have little action encoding [Wallis and Miller, 2003b; Padoa-Schioppa and Assad, 2006]. These studies place more attention on stimuli and rewards than actions. Specifically, each recording day used novel pictures [Wallis and Miller, 2003b] and different juices [Padoa-Schioppa and Assad, 2006]. Our study does not change the identity of stimuli and outcomes, but rather alters their associations. Further,

if an experiment includes the evaluation of actions, attention or a prior should be allotted to it and manifest as action selectivity. Kennerley and colleagues performed an experiment, in which monkeys selected stimuli that predicted rewards differing either in their size, likelihood of receiving them, and the physical cost to obtain them [Kennerley *et al.*, 2009]. Since cost involved lever movements, the subjects needed to evaluate AO associations, in addition to SO associations. In terms of neuronal encoding, the study found more action encoding compared to the two aforementioned studies. Likewise, our experiment is split across trials where chosen actions require comparisons of either AO associations or on SO associations. In so doing, actions may be more readily represented across both the AO and SO tasks.

4.4.2 Choice representation in PFo

Our study has found a strong representation of the chosen action in PFo. Proportions of action-encoding neurons were higher than those found in other studies [Tremblay and Schultz, 1999; Wallis and Miller, 2003b; Padoa-Schioppa and Assad, 2006; Bouret and Richmond, 2010; Abe and Lee, 2011].

There may be a number of reasons for varying degrees of choice-related action selectivity in PFo across studies. For example, tasks that could utilize Pavlovian approach responses rather than more deliberative movements may have less action selectivity. Governed by Pavlovian influences, the subject reflexively gravitates towards the desirable stimulus. Subsequently, the executed choice action need not be well encoded. In [Padoa-Schioppa and Assad, 2006], monkeys were presented two pictures that predicted two different rewards. The pictures had been well learned and always predicted the same reward across a day of the experiment. Hence, the subjects could select the better-rewarded picture by relying solely on SO associations and being more captivated and physically drawn in by the picture of greater value. Consistent with this idea, the study found a negligible number of neurons encoding actions. [Kennerley *et al.*, 2009] used a similar task set-up, having monkeys choose between two pictures presented side-by-side that denote two different reward outcomes. They also found less action encoding in PFo than our study. In contrast, our SO task could not be solved with Pavlovian approach responses. Both pictures could predict any of the three rewards depending on the trial. In other words, no picture could be consistently associated with a particular outcome, and therefore could not acquire the conditioned value. Thus, chosen action selectivity may scale with whether the action reflects a Pavlovian approach response or a more deliberative action.

The ease to which actions can be encoded as Pavlovian approach responses may be influenced by whether choice movements are indicated by saccade or lever movements. Eye movements could be a reflexive orientation response, much like a Pavlovian approach response. Whereas arm movements needed to control a joystick involve an arbitrary mapping between what is on the screen and the physical movement of the joystick. Such a mapping must be learned and cannot be acquired through a simple approach response. Based on these dis-

tinctions, we would expect less neuronal representation for chosen actions involving saccades as compared to actions involving lever movements. Indeed studies using arbitrary stimulus-motor mappings, including our SO task, have found more prevalent selectivity for the chosen action than studies using saccades [Padoa-Schioppa and Assad, 2006]. Fitting with this idea, Kennerley and colleagues had one subject use saccades and the other use a joystick via an arm movement [Kennerley *et al.*, 2009]. The subject who used saccades had an insignificant number of action encoding neurons in PFo, whereas the subject using the joystick had a sizable number of action encoding neurons in PFo. Other studies in which the subject cannot utilize Pavlovian approach responses include rat work where the subject must move according to odor cues in order to receive different types of rewards [Feierstein *et al.*, 2006; Furuyashiki *et al.*, 2008]. The odors are delivered from a port flanked by the response movement ports. Therefore the location is not in the same spatial position as the response locations and cannot be readily mapped onto a Pavlovian approach action. Rather the subject must translate the odor into the appropriate movement in order to receive a particular type of outcome. Hence these studies require action representation and have findings consistent with ours: there are many PFo neurons encoding deliberative actions.

However, the action encoding we observed also correlated with the direction of the better stimulus option, so that neuron activity could relate to the value of the chosen option. Chosen option encoding has been found in PFo neurons in rodents [Sul *et al.*, 2010], monkeys [Kennerley *et al.*, 2011], and humans [Arana *et al.*, 2003; Valentin *et al.*, 2007]. The choice-related activity in our study was limited to the SO task. This isolation might reflect the fact that only in the SO task could the subjects not anticipate their movements earlier in the trial. Consequently, the subjects would not know the value of their chosen action until the choice phase, when the action could be chosen. Hence, action encoding coincided with the earliest time at which the value of the action could be ascertained. Specifically, we observed the most action encoding prior to the feedback period, i.e., delivery of the chosen reward. Selecting a stimulus in our task reliably produces a reward of known value to the subject. Consequently, it is at this point the subject can transfer values from the two SO associations onto the possible choice actions. The feedback period, when the chosen reward is given, yields no new information about the value of the chosen action, and in fact there are fewer neurons encoding the chosen action at this time. In short, chosen action selectivity peaked at the time when choice options could be first compared.

This temporal selectivity is consistent with other studies that have found the strongest PFo choice-related action selectivity at the earliest time when the assessment of the choice can be made. In tasks where the subject uses the outcome of the current trial's chosen action to plan the next trial's action, action encoding occurs during the feedback period when the subject learns if his action yields reward [Tsujimoto *et al.*, 2009; Sul *et al.*, 2010; Abe and Lee, 2011]. Likewise, action encoding is maximal when the subject learns the value of his chosen action after having selected the action and waited a variable interval of time [Bouret and Richmond, 2010]. The value can be communicated as a binary signal,

i.e., reward for correct choice and no reward for incorrect choice [Tsujimoto *et al.*, 2009; Bouret and Richmond, 2010] or via visual cues that then predict different rewards [Abe and Lee, 2011, our SO task]. In all of these studies, reward delivery changes across time either within the trial or across trials. Thus, action encoding is updated at the end of the trial during the feedback period. Our SO task differs in that reward delivery is fully predicted by the picture values for a given trial. Hence, action encoding is first updated at the onset of the choice phase, when picture values are first mapped onto the potential actions necessary to select them. In light of this comparison, our study validates the idea that PFo can encode the chosen action specifically during the portion of the trial when the action can first be evaluated.Other studies may have less action encoding in PFo owning to the dynamics of encoding. Firstly, some tasks may encourage updating actions in a structured way based on past trials. That is, if actions and their reward contingencies do not change trial to trial, it becomes energy efficient to represent the action in memory and update only when the association changes. Consequently, the trial-averaged encoding of such actions will be less than that in a task where contingencies affecting those actions change trial to trial, such as in our SO task. Our task had the lever movement corresponding to the more preferred picture change randomly, as well as the identity of the corresponding reward. This randomness may promote more updating and subsequent encoding of actions and outcomes. In contrast, studies that rely on updating the value of actions in a structured way find less encoding in PFo. For instance, there is more structure in updating action values when subjects update actions based on actions from past trials. This is the case in a rock-paper-scissors game against a computer opponent that tracks and exploits the subject's choices [Abe and Lee, 2011]. Secondly, aside from structure in trial history accounting for less action encoding, action encoding might happen at different times within single trials. If that is the case, a decision can be made during various times within the trial, then analyses that collapse across all such trials end up diluting neural representations of action selection [Wallis and Miller, 2003b, our AO task]. Consequently, by temporally constraining when choice actions are executed both with respect to past trials and within the current one, we potentially are more likely to see greater action selectivity.

4.4.3 Interpretational issues

We found similar prevalence in action encoding across PFcs and PFo during the sampling period of the trial. Such similarity may stem from anatomical connections between the brain areas via areas 12o, 13a and b, and 14c [Carmichael and Price, 1996]. Consequently, activity across PFcs and PFo could well correlate. Further, we cannot determine the directionality of information flow between the areas when task-related encoding like that of the first reward appears in both areas at roughly the same time. While information may be shared freely through PFcs and PFo, it is also possible that reward information reaches both areas from a common source, accounting for the similar latency to encode value. Cor-

relation of actions and their intended outcomes is a further concern following the onset of the choice cues. During this time a significant number of PFo neurons encoded actions necessary for selecting the preferred picture. The chosen action was always in the direction of the more valuable stimulus. Hence, we cannot discern how much of that activity is purely motor-related.Another point to note is that overall neuronal activity in the AO task was greater than in the SO task. This appears highly correlated with behavioral reward preferences. During the AO task, behaviorally demonstrated preferences were more consistent, and both prevalence and strength of encoding these preferences were greater at the neuronal level. These neurons may have provided a clear signal with which the animal could more easily determine the better reward option. The AO task was also learned to a higher criterion than the SO task, and this difference may contribute to a stronger representation of AO associations at the neuronal level. Well-learned items can be encoded in the prefrontal cortex more strongly than more novel items [Pasupathy and Miller, 2005; Antzoulatos and Miller, 2011].

4.5 Conclusion

Our study has found that both PFcs and PFo are capable of encoding outcomes associated with external environmental stimuli and internally generated actions. The SO task evoked subtly more involvement from PFo as compared to PFcs, but differences were minor with respect to expectations from anatomy and findings from a number of past studies. Given both of these areas encoded multiple value estimates, they may play key roles in enabling the adaptive control of value-based decision making.

Chapter 5

Conclusions

We began this thesis with the question of how your brain learns to associate stimuli in the environment with actions and outcomes. Based on shared anatomical connections, we contrasted three prefrontal areas. We compared two areas with robust motor connections PFcs and PFdl in their roles for storing AO information in working memory. Then we compared two areas with prevalent limbic inputs PFcs and PFo in their ability to differentiate AO and SO information. This chapter will summarize findings from our two experiments in relation to current prefrontal literature. Then it will address remaining questions and future directions related to this work.

5.1 Summary of results

We approached the credit assignment problem by contrasting how different prefrontal regions encoded in working memory the value of either environmental stimuli or the subject's actions. Chapter 3 compared AO encoding across PFcs and PFdl and found that PFcs had prevalent encoding of AO associations compared to PFdl, despite taxing working memory, a process linked with robust PFdl activity. We also observed relative reward encoding of the second outcome with respect to the first, a pattern that was again more common in PFcs. These results strongly implicate PFcs in goal-directed decision-making. Chapter 4 then compared AO and SO encoding across PFcs and PFo. Both areas encoded all task-related information, namely actions, stimuli, and outcomes. There was a subtle bias for PFo to encode information in the SO task, particularly of the action needed to select the more preferred picture. In short, PFcs and PFo both carry all pieces of task information needed for flexible decision-making and credit assignment.

5.2 Prefrontal involvement in the framework for decision-making

A potential theoretical framework for understanding the role of PFC in decision-making is provided by computational accounts of reinforcement learning [Sutton and Barto, 1998]. Specifically, we might expect PFC to have correlates of these four parameters: the set of possible states, the set of possible actions or actions for a given state, the set of transition probabilities for the chance that one state following a given action leads to another particular state and a variable to track reward expectation. This thesis found that both PFcs and PFo can encode information related to states, actions, and rewards. PFdl, on the other hand, had mostly action encoding and was not as involved in decision-making. Our discussion in the remainder of this section will therefore focus on PFcs and PFo. Both of these areas contained neurons that reflected states, actions, and rewards in that they encoded cues (actions in the AO task and pictures in the SO task) and/or outcomes. Some did so only in one task. This diversity of selectivity suggests that the collective activity of neurons may reflect details about what behavioral state the subject is in as well as possible upcoming actions from that state.States, actions and outcomes appear to be encoded dynamically. For instance, attention may alter the span of the parameters to be encoded. From the previous chapter (section 4.4.1), we raised the idea that neuronal representations could scale with the relative attention placed across task variables. Greater attention to a particular parameter would yield greater neuronal encoding of that parameter. Similar responses have been observed in posterior association cortex, such as the parietal cortex [Platt and Glimcher, 1999]. This could explain the diversity of selectivity seen in PFcs and PFo across numerous studies. Further, the attention required to perform different actions can also manifest in the strength of neuronal encoding for that particular action. On one hand, little attention is necessary to execute pavlovian approach responses. Hence experiments that can be solved using such actions may have less action encoding. For example, decision-making tasks that require the animal to indicate its choice by looking at the preferred picture could be solved by pavlovian orienting responses [Padoa-Schioppa and Assad, 2006; Kennerley *et al.*, 2009]. In contrast, more attention is necessary for deliberative actions like when a subject needs to remap a 3D joystick movement onto a 2D movement on a computer screen [Kennerley *et al.*, 2009]; our experiments]. Such an action requires more attention and consequently could evoke more neuronal activity.

The stimuli, actions, and outcomes used in our experiments serve as building blocks for the computations needed in decision-making. Although we do not know the exact computations occurring in either PFo or PFcs, both areas utilize all three parameters. Past literature places a bias for PFo to compute signals related to states and PFcs actions. This thesis suggests that these areas embody computations that do not dissociate stimuli from actions within a local computational unit in either brain area. The computations

related to state may well require information about actions. For instance, it is possible that PFo captures very diverse signals related to state, such as the transition probability of one state switching to a particular one in the future: $P(s_{t+1} = s'|s_t = s, a_t = a)$, where action a in state s at time t leads to state s' at time $t + 1$. If so, it would make sense for PFo to carry signals about pictures and actions in our task. This proposal is consistent with ideas implicating PFo as a goods space [Padoa-Schioppa, 2011] capable of tracking potential choice options and their values [Sul *et al.*, 2010; Abe and Lee, 2011; Kennerley *et al.*, 2011].

As another key area for decision-making, PFcs also needs signals carrying information about states, actions and outcomes. Indeed we find PFcs neurons to encode all those signals. However, the role of this area often centers on action processing. Animal lesion work and human imaging find PFcs to be crucial for action selection in situations where the outcomes of the actions need to be monitored [Dias *et al.*, 1997; Hadland *et al.*, 2003; Rudebeck *et al.*, 2008; Noonan *et al.*, 2011]. But tracking potential actions and their outcomes change depending on the state. Hence, PFcs should encode state-related information as well. The recent predicted response-outcome model includes this fact. This model learns the joint probability of actions and outcomes for each given state, $P(a, o|s)$, and can accurately monitor goal-related actions [Alexander and Brown, 2011]. In sum, PFcs may have a bias for action encoding, but the computations performed within the area needs to account for state. Consequently, it should encode information related to the subject's environment (including the sensory stimuli) as our results suggest.In conclusion, both PFo and PFcs are key areas for processing information involved in goal-directed decision-making. They carry dynamic signals that can identify and evaluate specific items in the state space and action space of the agent's task. This is consistent with our findings: neurons encoded specific cues or outcomes, as well as AO and SO associations, depending on the block of the experiment. PFo and PFcs may be more biased towards processing state-related and action-related computations, respectively, but ultimately both areas need information about stimuli, actions, and outcomes.

5.3 Remaining questions

There are more avenues of research following this thesis. We have only compared PFdl and PFo against PFcs. It remains unclear how divergent are the roles of PFdl compared to PFo and PFcs in encoding either AO or SO associations during a working memory choice task. Another question directly related to the thesis is why neuronal encoding of AO and SO associations does not closely match lesion studies. Many such studies have found PFcs to be critical for updating AO associations based on lesions in rodents [Balleine and Dickinson, 1998; Dias and Aggleton, 2000] and monkeys [Hadland *et al.*, 2003; Kennerley *et al.*, 2006; Rudebeck *et al.*, 2008]. Similarly, studies with PFo lesions in rodents [Pickens *et al.*, 2003]

and monkeys [Dias *et al.*, 1996; Dias *et al.*, 1997; Izquierdo *et al.*, 2004; Rudebeck *et al.*, 2008] found that PFo is necessary for representing SO associations. In contrast, we only see a bias for PFo activity during the SO task and none for PFcs during the AO task. It is likely that neuronal activity in PFcs and PFo are highly correlated, given their strong interconnections. Further, lesions often cover large areas and consequently greatly alter behavior. To bridge the gap between lesion work and our work, future research may consider local and temporary inactivation of each brain area while subjects perform AO and SO tasks. We also cannot determine the source of the choice encoding in PFo neurons during the SO task. Perhaps it reflects the chosen action similar to rat work [Feierstein *et al.*, 2006; Furuyashiki *et al.*, 2008; Sul *et al.*, 2010] or it could merely reflect the selection of the more valuable stimulus.

Another constraint to this thesis is its lack of realistic contingency learning. In our tasks, learning does not occur across trials and hence could not utilize reinforcement models. Future work may consider trial-to-trial learning. One possible task could require a subject to select actions in response to various visual cues. Reward delivery then follows and can either be predicted by what action the subject made or by what cue the subject viewed. For the subject to maximize rewards, he would need to learn whether a particular action or cue predicts reward and keep selecting that action or cue. Through this type of learning paradigm, we can observe the neuronal mechanisms involved in a more realistic credit assignment problem. In other words, the animal has to determine whether actions or stimuli are relevant to solving a given problem. This paradigm benefits from the quantitative power reinforcement modeling offers in characterizing both behavioral and neuronal activity.

5.4 Closing thoughts

Credit assignment plays a central role in what the brain is designed to do. It provides biologists with a psychological and computationally rigorous framework for quantifying neuronal mechanisms related to contingency learning. This thesis highlights the differential roles of prefrontal areas in contingency learning over short delays. Future work can expand from this by having subjects learn contingencies across long periods of time, which better reflect natural learning. Ultimately, decision-making studies will benefit greatly from tracking the full learning process in natural settings. This field may achieve this goal from technology and methodologies developed by the engineering and physical science communities. A likely starting point would include adoption of chronic implants that monitor neuronal activity and collaborative code libraries. Such technologies will broaden the scope over which we can monitor neuronal activity and speed up data analysis. After all, a decision is not made by a single neuron. It is the distributed network of neuronal dynamics that we need to observe and understand.

Bibliography

[Abe and Lee, 2011] H. Abe and D. Lee. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*, 70(4):731–41, May 2011.

[Adolphs *et al.*, 1994] R. Adolphs, D. Tranel, H. Damasio, and A. Damasio. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372:669–672, 1994.

[Alexander and Brown, 2011] W. H. Alexander and J. W. Brown. Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 14(10), September 2011.

[Allison, 1999] P. D. Allison. *Multiple Regression: A Primer*. Pine Forge Press, Thousand Oaks, CA, 1999.

[Amiez *et al.*, 2006] C. Amiez, J. P. Joseph, and E. Procyk. Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex*, 16(7):1040–55, July 2006.

[Antzoulatos and Miller, 2011] E. G. Antzoulatos and E. K. Miller. Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron*, 71(2):243–9, July 2011.

[Arana *et al.*, 2003] F. S. Arana, J. A. Parkinson, E. Hinton, A. J. Holland, A. M. Owen, and A. C. Roberts. Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *Journal of Neuroscience*, 23(29):9632–8, October 2003.

[Asaad *et al.*, 2000] W. F. Asaad, G. Rainer, and E. K. Miller. Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, 84(1):451–459, 2000.

[Balleine and Dickinson, 1998] B. W. Balleine and A. Dickinson. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5):407–19, 1998.

[Barbas and Pandya, 1989] H. Barbas and D. N. Pandya. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, 286:353–375, 1989.

[Barraclough *et al.*, 2004] D. J. Barraclough, M. L. Conroy, and D. Lee. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, 7(4):404–10, April 2004.

[Bates and Goldman-Rakic, 1993] J. F. Bates and P. S. Goldman-Rakic. Prefrontal connections of medial motor areas in the rhesus monkey. *Journal of Comparative Neurology*, 336(2):211–28, October 1993.

[Baxter *et al.*, 2000] M. G. Baxter, A. Parker, C. C. Lindner, A. D. Izquierdo, and E. A. Murray. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *Journal of Neuroscience*, 20(11):4311–9, June 2000.

[Baxter *et al.*, 2008] M. G. Baxter, D. Gaffan, D. A. Kyriazis, and A. S. Mitchell. Dorsolateral prefrontal lesions do not impair tests of scene learning and decision-making that require frontal-temporal interaction. *European Journal of Neuroscience*, 28(3):491–9, August 2008.

[Bechara *et al.*, 1998] A. Bechara, H. Damasio, D. Tranel, and S. W. Anderson. Dissociation Of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience*, 18(1):428–37, January 1998.

[Behrens *et al.*, 2007] T. E. J. Behrens, M. W. Woolrich, M. E. Walton, and M. F. S. Rushworth. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9):1214–21, September 2007.

[Biggins *et al.*, 1997] C. A. Biggins, S. MacKay, W. Clark, and G. Fein. Event-related potential evidence for frontal cortex effects of chronic cocaine dependence. *Biological Psychiatry*, 42(6):472–485, 1997.

[Bolla *et al.*, 1998] K. I. Bolla, J. L. Cadet, and E. D. London. The neuropsychiatry of chronic cocaine abuse. *Journal of Neuropsychiatry and Clinical Neurosciences*, 10(3):280–9, January 1998.

[Bolla *et al.*, 2003] K. I. Bolla, D. A. Eldreth, E. D. London, K. A. Kiehl, M. Mouratidis, C. Contoreggi, J. A. Matochik, V. Kurian, J. L. Cadet, A. S. Kimes, F. R. Funderburk, and M. Ernst. Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *NeuroImage*, 19(3):1085–1094, July 2003.

[Bouret and Richmond, 2010] S. Bouret and B. J. Richmond. Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *Journal of Neuroscience*, 30(25):8591–601, June 2010.

[Breiter *et al.*, 2001] H. C. Breiter, I. Aharon, D. Kahneman, A. Dale, and P. Shizgal. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30(2):619–39, May 2001.

[Brodmann, 1909] K. Brodmann. *Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Johann Ambrosius Bart, Leipzig, 1909.

[Brody *et al.*, 2002] A. L. Brody, M. A. Mandelkern, E. D. London, A. R. Childress, G. S. Lee, R. G. Bota, M. L. Ho, S. Saxena, L. R. Baxter, D. Madsen, and E. J. Murray. Brain metabolic changes during cigarette craving. *Arch Gen Psychiatry*, 59:1162–1172, 2002.

[Brown and Braver, 2005] J. W. Brown and T. S. Braver. Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307(5712):1118–21, February 2005.

[Camille *et al.*, 2004] N. Camille, G. Coricelli, J. Sallet, P. Pradat-Diehl, J.-R. Duhamel, and A. Sirigu. The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674):1167–70, May 2004.

[Carmichael and Price, 1994] S. T. Carmichael and J. L. Price. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*, 346(3):366–402, August 1994.

[Carmichael and Price, 1995a] S. T. Carmichael and J. L. Price. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, 363(4):615–641, December 1995.

[Carmichael and Price, 1995b] S. T. Carmichael and J. L. Price. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, 363(4):642–664, December 1995.

[Carmichael and Price, 1996] S. T. Carmichael and J. L. Price. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, 371(2):179–207, July 1996.

[Carter *et al.*, 1998] C. S. Carter, T. S. Braver, D. M. Barch, M. M. Botvinick, D. Noll, and J. D. Cohen. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280:747–749, May 1998.

[Catafau *et al.*, 1999] A. M. Catafau, A. Etcheberrigaray, J. Perez de los Cobos, M. Estorch, J. Guardia, A. Flotats, L. Bernà, C. Marí, M. Casas, and I. Carrió. Regional cerebral blood flow changes in chronic alcoholic patients induced by naltrexone challenge during detoxification. *Journal of Nuclear Medicine*, 40(1):19–24, January 1999.

[Cipolloni *et al.*, 1999] P. B. Cipolloni and D. N. Pandya. Cortical connections of the frontoparietal opercular areas in the rhesus monkey. *Journal of Comparative Neurology*, 458:431–458, 1999.

[Cohen *et al.*, 2007] J. D. Cohen, S. M. McClure, and A. J. Yu. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362(1481):933–42, May 2007.

[Constantinidis *et al.*, 2001] C. Constantinidis, M. N. Franowicz, and P. S. Goldman-Rakic. The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience*, 4(3):311–6, March 2001.

[Croxson *et al.*, 2005] P. L. Croxson, H. Johansen-Berg, T. E. J. Behrens, M. D. Robson, M. A. Pinsk, C. G. Gross, W. Richter, M. C. Richter, S. Kastner, and M. F. S. Rushworth. Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *Journal of Neuroscience*, 25(39):8854–66, September 2005.

[Croxson *et al.*, 2009] P. L. Croxson, M. E. Walton, J. X. O'Reilly, T. E. J. Behrens, and M. F. S. Rushworth. Effort-based cost-benefit valuation and the human brain. *Journal of Neuroscience*, 29(14):4531–41, April 2009.

[Damasio, 1994] A. R. Damasio. *Descartes' Error: Emotion, Reason and the Human Brain*. Grosset/Putnam, New York, 1994.

[Daw *et al.*, 2005] N. D. Daw, Y. Niv, and P. Dayan. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12):1704–11, December 2005.

[Daw *et al.*, 2006] N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, and R. J Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879, 2006.

[Dias and Aggleton, 2000] R. Dias and J. P. Aggleton. Effects of selective excitotoxic prefrontal lesions on acquisition of nonmatching- and matching-to-place in the T-maze in the rat: differential involvement of the prelimbic-infralimbic and anterior cingulate cortices in providing behavioural flexibility. *European Journal of Neuroscience*, 12(12):4457–66, December 2000.

[Dias *et al.*, 1996] R. Dias, T. W. Robbins, and A. C. Roberts. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, 380:69–72, 1996.

[Dias *et al.*, 1997] R. Dias, T. W. Robbins, and A. C. Roberts. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience*, 17(23):9285–97, December 1997.

[Dum and Strick, 1991] R. Dum and P. L. Strick. The Origin of corticospinal projections from the premotor areas in the frontal lobe. *Journal of Neuroscience*, 11(3):667–689, 1991.

[Dum and Strick, 1993] R. P. Dum and P. L. Strick. Cingulate motor areas. In B. A. Vogt and M. Gabriel, editors, *Neurobiology of Cingulate Cortex and Limbic Thalamus*, pages 415–441. Birkhäuser, Boston, 1993.

[Dum and Strick, 1996] R. P. Dum and P. L. Strick. Spinal cord terminations of the medial wall motor areas in macaque monkeys. *Journal of Neuroscience*, 16(20):6513–25, October 1996.

[Duncan and Owen, 2000] J. Duncan and A. M. Owen. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10):475–83, October 2000.

[Duncan *et al.*, 1996] J. Duncan, H. Emslie, P. Williams, R. Johnson, and C. Freer. Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognitive Psychology*, 30(3):257–303, June 1996.

[Eichenbaum *et al.*, 1983] H. Eichenbaum, R. A. Clegg, and A. Feeley. Reexamination of functional subdivisions of the rodent prefrontal cortex. *Experimental Neurology*, 79(2):434–51, February 1983.

[Everitt and Robbins, 2005] B. J. Everitt and T. W. Robbins. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, 8(11):1481–9, November 2005.

[Feierstein *et al.*, 2006] C. E. Feierstein, M. C. Quirk, N. Uchida, D. L. Sosulski, and Z. F. Mainen. Representation of spatial goals in rat orbitofrontal cortex. *Neuron*, 51(4):495–507, August 2006.

[Fellows and Farah, 2007] L. K. Fellows and M. J. Farah. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex*, 17(11):2669–74, November 2007.

[Fellows, 2006] L. K. Fellows. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain*, 129:944–52, April 2006.

[Freedman *et al.*, 2001] D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, 291(5502):312–6, January 2001.

[Funahashi *et al.*, 1989] S. Funahashi, C. J. Bruce, and P. S. Goldman-Rakic. Mnemonic coding of visual space in the monkeys dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61(2):331–349, 1989.

[Funahashi *et al.*, 1993] S. Funahashi, M. V. Chafee, and P. S. Goldman-Rakic. Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature*, 365:753–756, 1993.

[Furuyashiki *et al.*, 2008] T. Furuyashiki, P. C. Holland, and M. Gallagher. Rat orbitofrontal cortex separately encodes response and outcome information during performance of goal-directed behavior. *Journal of Neuroscience*, 28(19):5127–38, May 2008.

[Fuster *et al.*, 2000] J. M. Fuster, M. Bodner, and J. K. Kroger. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, 405(6784):347–350, 2000.

[Fuster, 1973] J. M. Fuster. Unit activity in prefrontal cortex during delayed-resopnse performance: neuronal correlates of transient memory. *Journal of Neurophysiology*, 36:39–60, 1973.

[Fuster, 1997] J. M. Fuster. *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe*. Lippincott, Williams & Wilkins, Philadelphia, PA, 1997.

[Fuster, 2001] J. M. Fuster. The prefrontal cortex–an update: time is of the essence. *Neuron*, 30(2):319–33, May 2001.

[Gaffan *et al.*, 2002] D. Gaffan, A. Easton, and A. Parker. Interaction of inferior temporal cortex with frontal cortex and basal forebrain: double dissociation in strategy implementation and associative learning. *Journal of Neuroscience*, 22(16):7288–96, August 2002.

[Gehring *et al.*, 1993] W. J. Gehring, B. Goss, M. G. H. Coles, D. E. Meyer, and E. Donchin. A neural system for error detection and compensation. *Psychological Science*, 4(6):385–390, 1993.

[Giedd *et al.*, 1999] J. N. Giedd, J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans, and J. L. Rapoport. Brain development during

childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, 2(10):861–3, October 1999.

[Glimcher, 2003] P. W. Glimcher. *Decisions, uncertainty, and the brain: the science of neuroeconomics*. MIT Press, Cambridge, 2003.

[Goldman-Rakic *et al.*, 1984] P. S. Goldman-Rakic, L. D. Selemon, and M. L. Schwartz. Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal corex in te hesus monkey. *Neuroscience*, 12(3):719–743, 1984.

[Goldman-Rakic, 1987] P. S. Goldman-Rakic. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum, editor, *Handbook of Physiology: The Nervous System*, pages 373–417. Am. Physiol. Soc., Bethesda, MD, 1987.

[Gottfried *et al.*, 2003] J. A. Gottfried, J. O'Doherty, and R. J. Dolan. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, 301(5636):1104–7, August 2003.

[Grant *et al.*, 1996] S. Grant, E. D. London, D. B. Newlin, V. L. Villemagne, X. Liu, C. Contoreggi, R. L. Phillips, A. S. Kimes, and A. Margolin. Activation of memory circuits during cue-elicited cocaine craving. *Proceedings of the National Academy of Sciences*, 93(21):12040–5, October 1996.

[Hadland *et al.*, 2003] K. A. Hadland, M. F. S. Rushworth, D. Gaffan, and R. E. Passingham. The anterior cingulate and reward-guided selection of actions. *Journal of Neurophysiology*, 89(2):1161–4, February 2003.

[Halsband and Passingham, 1982] U. Halsband and R. E. Passingham. The role of premotor and parietal cortex in the direction of action. *Brain Research*, 240(2):368–372, 1982.

[Hayden and Platt, 2010] B. Y. Hayden and M. L. Platt. Neurons in anterior cingulate cortex multiplex information about reward and action. *Journal of Neuroscience*, 30(9):3339–46, March 2010.

[Holroyd and Coles, 2002] C. B. Holroyd and M. G.H. Coles. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4):679–709, 2002.

[Holroyd *et al.*, 2004] C. B. Holroyd, S. Nieuwenhuis, N. Yeung, L. Nystrom, R. B. Mars, M. G. H. Coles, and J. D. Cohen. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7(5):497–8, May 2004.

[Hoshi *et al.*, 2005] E. Hoshi, H. Sawamura, and J. Tanji. Neurons in the rostral cingulate motor area monitor multiple phases of visuomotor behavior with modest parametric selectivity. *Journal of Neurophysiology*, pages 640 – 656, 2005.

[Ito *et al.*, 2003] S. Ito, V. Stuphorn, J. W. Brown, and J. D. Schall. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, 302(5642):120–2, October 2003.

[Izquierdo *et al.*, 2004] A. Izquierdo, R. K. Suda, and E. A. Murray. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *Journal of Neuroscience*, 24(34):7540–8, August 2004.

[Jacobsen and Nissen, 1936] C. F. Jacobsen and H. W. Nissen. Studies of cerebral function in primates. IV. The effects of frontal lobe lesions on the delayed alternation habit in monkeys. *Journal of Comparative Psychology*, 23(1):101, 1936.

[Jonides *et al.*, 1993] J. Jonides, E. E. Smith, R. A. Koeppe, and E. Awh. Spatial working memory in humans as revealed by PET. *Nature*, 1993.

[Kang *et al.*, 2011] M. J. Kang, A. Rangel, M. Camus, and C. F. Camerer. Hypothetical and real choice differentially activate common valuation areas. *Journal of Neuroscience*, 31(2):461–8, January 2011.

[Kennerley and Wallis, 2009] S. W. Kennerley and J. D. Wallis. Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. *Journal of Neurophysiology*, 102(6):3352–64, December 2009.

[Kennerley *et al.*, 2006] S. W. Kennerley, M. E. Walton, T. E. J. Behrens, M. J. Buckley, and M. F. S. Rushworth. Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7):940–7, July 2006.

[Kennerley *et al.*, 2009] S. W. Kennerley, A. F. Dahmubed, A. H. Lara, and J. D. Wallis. Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience*, 21(6):1162–78, June 2009.

[Kennerley *et al.*, 2011] S. W. Kennerley, T. E. J. Behrens, and J. D. Wallis. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, 14(12):1581–1591, 2011.

[Klein *et al.*, 2010] J. C. Klein, M. F. S. Rushworth, T. E. J. Behrens, C. E. Mackay, A. J. de Crespigny, H. D'Arceuil, and H. Johansen-Berg. Topography of connections between human prefrontal cortex and mediodorsal thalamus studied with diffusion tractography. *NeuroImage*, 51(2):555–64, June 2010.

[Knutson *et al.*, 2000] B. Knutson, A. Westdorp, E. Kaiser, and D. Hommer. FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*, 12(1):20–7, July 2000.

[Knutson *et al.*, 2005] B. Knutson, J. Taylor, M. Kaufman, R. Peterson, and G. Glover. Distributed neural representation of expected value. *Journal of Neuroscience*, 25(19):4806–12, May 2005.

[Koenigs *et al.*, 2007] M. Koenigs, L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, and A. Damasio. Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446(7138):908–11, April 2007.

[Lara *et al.*, 2009] A. H. Lara, S. W. Kennerley, and J. D. Wallis. Encoding of gustatory working memory by orbitofrontal neurons. *Journal of Neuroscience*, 29(3):765–74, January 2009.

[Lau and Glimcher, 2007] B. Lau and P. W. Glimcher. Action and outcome encoding in the primate caudate nucleus. *Journal of Neuroscience*, 27(52):14502–14, December 2007.

[Leon and Shadlen, 1999] M. I. Leon and M. N. Shadlen. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron*, 24(2):415–25, October 1999.

[Liu *et al.*, 1998] X. Liu, J. A. Matochik, J. L. Cadet, and E. D. London. Smaller volume of prefrontal lobe in polysubstance abusers: a magnetic resonance imaging study. *Neuropsychopharmacology*, 18(4):243–52, April 1998.

[Luk and Wallis, 2009] C.-H. Luk and J. D. Wallis. Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *Journal of Neuroscience*, 29(23):7526–39, June 2009.

[MacDonald *et al.*, 2000] A. W. MacDonald, J. D. Cohen, V. A. Stenger, and C. S. Carter. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472):1835–1838, June 2000.

[Mackey and Petrides, 2010] S. Mackey and M. Petrides. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *European Journal of Neuroscience*, 32(11):1940–50, December 2010.

[Matsumoto *et al.*, 2003] K. Matsumoto, W. Suzuki, and K. Tanaka. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science*, 301(5630):229–32, July 2003.

[Matsumoto *et al.*, 2007] M. Matsumoto, K. Matsumoto, H. Abe, and K. Tanaka. Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 10(5):647–56, May 2007.

[McNab and Klingberg, 2008] F. McNab and T. Klingberg. Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, 11(1):103–7, January 2008.

[Miller and Cohen, 2001] E. K. Miller and J. D. Cohen. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24:167–202, January 2001.

[Miller and Ross, 1975] D. T. Miller and M. Ross. Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 2:213–225, 1975.

[Miller *et al.*, 1996] E. K. Miller, C. A. Erickson, and R. Desimone. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16(16):5154–67, August 1996.

[Miltner *et al.*, 1997] W. H. R. Miltner, C. H. Braun, and M. G. H. Coles. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, 9(6):788–798, 1997.

[Modha and Singh, 2010] D. S. Modha and R. Singh. Network architecture of the long-distance pathways in the macaque brain. *Proceedings of the National Academy of Sciences*, 107(30):13485–90, July 2010.

[Modirrousta and Fellows, 2008] M. Modirrousta and L. K. Fellows. Dorsal medial prefrontal cortex plays a necessary role in rapid error prediction in humans. *Journal of Neuroscience*, 28(51):14000–5, December 2008.

[Morecraft *et al.*, 1992] R. J. Morecraft, C. Geula, and M. M. Mesulam. Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. *Journal of Comparative Neurology*, 323(3):341–58, September 1992.

[Mountcastle *et al.*, 1975] V. B. Mountcastle, J. C. Lynch, A. Georgopoulos, H. Sakata, and C. Acuna. Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *Journal of Neurophysiology*, 38:871–908, 1975.

[Murray and Izquierdo, 2007] E. A. Murray and A. Izquierdo. Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Annals of the New York Academy of Sciences*, 1121:273–96, December 2007.

[Nimchinsky *et al.*, 1996] E. A. Nimchinsky, P. R. Hof, W. G. Young, and J. H. Morrison. Neurochemical, morphologic, and laminar characterization of cortical projection neurons in the cingulate motor areas of the macaque monkey. *Journal of Comparative Neurology*, 374:136–160, 1996.

[Noonan *et al.*, 2011] M. P. Noonan, R. B. Mars, and M. F. S. Rushworth. Distinct roles of three frontal cortical areas in reward-guided behavior. *Journal of Neuroscience*, 31(40):14399–14412, October 2011.

[O'Doherty *et al.*, 2003] J. O'Doherty, H. Critchley, R. Deichmann, and R. J. Dolan. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience*, 23(21):7931–9, August 2003.

[Öngür and Price, 2000] D. Öngür and J. L. Price. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, 10(3):206–19, March 2000.

[Öngür *et al.*, 2003] D. Öngür, A. T. Ferry, and J. L. Price. Architectonic subdivision of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology*, 460(3):425–49, June 2003.

[Opitz and Friederici, 2003] B. Opitz and A. D. Friederici. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, 19(4):1730–1737, August 2003.

[Ostlund and Balleine, 2007] S. B. Ostlund and B. W. Balleine. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *Journal of Neuroscience*, 27(18):4819–25, May 2007.

[Padoa-Schioppa and Assad, 2006] C. Padoa-Schioppa and J. A. Assad. Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090):223–6, May 2006.

[Padoa-Schioppa and Assad, 2008] C. Padoa-Schioppa and J. A. Assad. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, 11(1):95–102, January 2008.

[Padoa-Schioppa, 2011] C. Padoa-Schioppa. Neurobiology of economic choice: a good-based model. *Annual Review of Neuroscience*, 34:333–59, January 2011.

[Passingham, 1995] R. E. Passingham. *The frontal lobes and voluntary action*. Oxford University Press, New York, 1995.

[Pasupathy and Miller, 2005] A. Pasupathy and E. K. Miller. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*, 433(7028):873–6, February 2005.

[Paxinos *et al.*, 2000] G. Paxinos, X. Huang, and A. W. Toga. *The rhesus monkey brain in stereotaxic coordinates*. Academic, San Diego, 2000.

[Pearce, 1997] J. M. Pearce. *Animal learning and cognition: An introduction*. Psychology Press, 2 edition, 1997.

[Pears *et al.*, 2003] A. Pears, J. A. Parkinson, L. Hopewell, B. J. Everitt, and A. C. Roberts. Lesions of the orbitofrontal but not medial prefrontal cortex disrupt conditioned reinforcement in primates. *Journal of Neuroscience*, 23(35):11189–201, December 2003.

[Petrides and Pandya, 1984] M. Petrides and D. N. Pandya. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *Journal of Comparative Neurology*, 228(1):105–16, September 1984.

[Petrides and Pandya, 1994] M. Petrides and D. N. Pandya. *Comparative architectonic analysis of the human and the macaque frontal cortex. In Handbook of Neuropsychology*. Elsevier, Amsterdam, 1994.

[Petrides and Pandya, 1999] M. Petrides and D. N. Pandya. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience*, 11(3):1011–36, March 1999.

[Petrides, 1982] M. Petrides. Motor conditional associative-learning after selective prefrontal lesions in the monkey. *Behavioural Brain Research*, 5(4):407–13, August 1982.

[Petrides, 1995] M. Petrides. Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience*, 15(1 Pt 1):359–75, January 1995.

[Petrides, 1996] M. Petrides. Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 351(1346):1455–61; discussion 1461–2, October 1996.

[Pickens *et al.*, 2003] C. L. Pickens, M. P. Saddoris, B. Setlow, M. Gallagher, P. C. Holland, and G. Schoenbaum. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience*, 23(35):11078–84, December 2003.

[Platt and Glimcher, 1999] M. L. Platt and P. W. Glimcher. Neural correlates of decision variables in parietal cortex. *Nature*, pages 233–238, 1999.

[Procyk and Goldman-Rakic, 2006] E. Procyk and P. S. Goldman-Rakic. Modulation of dorsolateral prefrontal delay activity during self-organized behavior. *Journal of Neuroscience*, 26(44):11313–23, November 2006.

[Quilodran *et al.*, 2008] R. Quilodran, M. Rothé, and E. Procyk. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, 57(2):314–25, January 2008.

[Rainer *et al.*, 1998] G. Rainer, W. F. Asaad, and E. K. Miller. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, 393(6685):577–9, June 1998.

[Rao *et al.*, 1997] S. C. Rao, G. Rainer, and E. K. Miller. Integration of what and where in the primate prefrontal cortex. *Science*, 276(5313):821–824, May 1997.

[Ribas-Fernandes *et al.*, 2011] J. J. F. Ribas-Fernandes, A. Solway, C. Diuk, J. T. McGuire, A. G. Barto, Y. Niv, and M. M. Botvinick. A neural signature of hierarchical reinforcement learning. *Neuron*, 71(2):370–9, July 2011.

[Robinson *et al.*, 2001] T. E. Robinson, G. Gorny, E. Mitton, and B. Kolb. Cocaine self-administration alters the morphology of dendrites and dendritic spines in the nucleus accumbens and neocortex. *Synapse*, 39(3):257–66, March 2001.

[Roesch and Olson, 2003] M. R. Roesch and C. R. Olson. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *Journal of Neurophysiology*, 90(3):1766–89, September 2003.

[Roesch and Olson, 2004] M. R. Roesch and C. R. Olson. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science*, 304(5668):307–10, April 2004.

[Rolls *et al.*, 1994] E. T. Rolls, J. Hornak, D. Wade, and J. McGrath. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57(12):1518–24, December 1994.

[Romo *et al.*, 1999] R. Romo, C. D. Brody, A. Hernández, and L. Lemus. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, 399(6735):470–3, June 1999.

[Rose and Woolsey, 1948] J. E. Rose and C. N. Woolsey. The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep and cat. *Res Publ Assoc Res Nerv Ment Dis*, 27:210–32, January 1948.

[Rossi *et al.*, 2007] A. F. Rossi, N. P. Bichot, R. Desimone, and L. G. Ungerleider. Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. *Journal of Neuroscience*, 27(42):11306–14, October 2007.

[Rowe, 2000] J. B. Rowe. The prefrontal cortex: response selection or maintenance within working memory? *Science*, 288(5471):1656–1660, June 2000.

[Rudebeck *et al.*, 2006] P. H. Rudebeck, M. J. Buckley, M. E. Walton, and M. F. S. Rushworth. A role for the macaque anterior cingulate gyrus in social valuation. *Science*, 313(5791):1310–2, September 2006.

[Rudebeck *et al.*, 2008] P. H. Rudebeck, T. E. Behrens, S. W. Kennerley, M. G. Baxter, M. J. Buckley, M. E. Walton, and M. F. S. Rushworth. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *Journal of Neuroscience*, 28(51):13775–85, December 2008.

[Rushworth and Behrens, 2008] M. F. S. Rushworth and T. E. J. Behrens. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, 11(4):389–97, April 2008.

[Rushworth *et al.*, 2004] M. F. S. Rushworth, M. E. Walton, S. W. Kennerley, and D. M. Bannerman. Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, 8(9):410–7, September 2004.

[Rushworth *et al.*, 2007] M F S Rushworth, T E J Behrens, P H Rudebeck, and M E Walton. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences*, 11(4):168–76, April 2007.

[Sallet *et al.*, 2007] J. Sallet, R. Quilodran, M. Rothé, J. Vezoli, J.-P. Joseph, and E. Procyk. Expectations, gains, and losses in the anterior cingulate cortex. *Cogn Affect Behav Neurosci*, 7(4):327–336, 2007.

[Schoenbaum *et al.*, 1998] G. Schoenbaum, A. A. Chiba, and M. Gallagher. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, 1(2):155–9, June 1998.

[Schoenbaum *et al.*, 2002] G. Schoenbaum, S. L. Nugent, M. P. Saddoris, and B. Setlow. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*, 13(6):885–90, May 2002.

[Sell *et al.*, 2000] L. A. Sell, J. S. Morris, J. Bearn, R. S. Frackowiak, K. J. Friston, and R. J. Dolan. Neural responses associated with cue evoked emotional states and heroin in opiate addicts. *Drug and Alcohol Dependence*, 60(2):207–16, August 2000.

[Semendeferi *et al.*, 2002] K. Semendeferi, A. Lu, N. Schenker, and H. Damasio. Humans and great apes share a large frontal cortex. *Nature Neuroscience*, 5(3):272–6, March 2002.

[Seo and Lee, 2007] H. Seo and D. Lee. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *Journal of Neuroscience*, 27(31):8366–77, August 2007.

[Seo *et al.*, 2007] H. Seo, D. J. Barraclough, and D. Lee. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cerebral Cortex*, 17 Suppl 1:i110–7, October 2007.

[Shallice and Burgess, 1991] T. Shallice and P. W. Burgess. Deficits in strategy application following frontal lobe damage in man. *Brain*, 114:727–41, April 1991.

[Sul *et al.*, 2010] J. H. Sul, H. Kim, N. Huh, D. Lee, and M. W. Jung. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66(3):449–60, May 2010.

[Sutton and Barto, 1998] R. S. Sutton and A. G. Barto. *Reinforcement Learning : An Introduction*. The MIT Press, Cambridge, 1998.

[Thorpe *et al.*, 1983] S. J. Thorpe, E. T. Rolls, and S. Maddison. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Experimental Brain Research*, 49:93–115, 1983.

[Tremblay and Schultz, 1999] L. Tremblay and W. Schultz. Relative reward preference in primate orbitofrontal cortex. *Nature*, 398(6729):704–8, April 1999.

[Tsao *et al.*, 2008] D. Y. Tsao, N. Schweers, S. Moeller, and W. A. Freiwald. Patches of face-selective cortex in the macaque frontal lobe. *Nature Neuroscience*, 11(8):877–9, August 2008.

[Tsujimoto *et al.*, 2009] S. Tsujimoto, A. Genovesio, and S. P. Wise. Monkey orbitofrontal cortex encodes response choices near feedback time. *Journal of Neuroscience*, 29(8):2569–74, February 2009.

[Tversky and Kahneman, 1981] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481):453, 1981.

[Ullsperger and von Cramon, 2003] M. Ullsperger and D. Yves von Cramon. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, 23(10):4308–14, May 2003.

[Uylings *et al.*, 2003] H. B.M. Uylings, H. J. Groenewegen, and B. Kolb. Do rats have a prefrontal cortex? *Behavioural Brain Research*, 146(1-2):3–17, November 2003.

[Valentin *et al.*, 2007] V. V. Valentin, A. Dickinson, and J. P. O'Doherty. Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, 27(15):4019–26, April 2007.

[Vogt *et al.*, 2005] B. A. Vogt, L. Vogt, N. B. Farber, and G. Bush. Architecture and neurocytology of monkey cingulate gyrus. *Journal of Comparative Neurology*, 485(3):218–239, 2005.

[Volkow *et al.*, 1992] N. D. Volkow, R. Hitzemann, G. J. Wang, J. S. Fowler, A. P. Wolf, S. L. Dewey, and L. Handlesman. Long-term frontal brain metabolic changes in cocaine abusers. *Synapse*, 11(3):184–190, 1992.

[Wallis and Miller, 2003a] J. D. Wallis and E. K. Miller. From rule to response: neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, 90(3):1790–806, September 2003.

[Wallis and Miller, 2003b] J. D. Wallis and E. K. Miller. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience*, 18(7):2069–2081, October 2003.

[Wallis *et al.*, 2001] J. D. Wallis, K. C. Anderson, and E. K. Miller. Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840):953–6, June 2001.

[Wallis, 2007] J. D. Wallis. Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience*, 30:31–56, January 2007.

[Walton *et al.*, 2002] M. E. Walton, D. M. Bannerman, and M. F. S. Rushworth. The role of rat medial frontal cortex in effort-based decision making. *Journal of Neuroscience*, 22(24):10996–1003, December 2002.

[Walton *et al.*, 2004] M. E. Walton, J. T. Devlin, and M. F. S. Rushworth. Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, 7(11):1259–65, November 2004.

[Watanabe, 1996] M. Watanabe. Reward expectancy in primate prefrontal neurons. *Nature*, 382(6592):629–632, 1996.

[White and Wise, 1999] I. M. White and S. P. Wise. Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research*, 126(3):315–35, June 1999.

[Williams *et al.*, 2004] Z. M. Williams, G. Bush, S. L. Rauch, G. R. Cosgrove, and E. N. Eskandar. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience*, 7(12):1370–5, December 2004.

[Wilson *et al.*, 1993] F. A. Wilson, S. P. Scalaidhe, and P. S. Goldman-Rakic. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, 260(5116):1955–1958, 1993.

[Zilles and Wree, 1995] K. Zilles and A. Wree. Cortex: areal and laminar structure. In: The rat nervous system. In G. Paxinos, editor, *The rat nervous system*, pages 649–685. Academic Press, San Diego, 1995.