Neural mechanisms underlying the influence of associative learning on valuation and decision-making in humans

Thesis by

Signe Bray

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2008

(Submitted July 5, 2008)

Acknowledgments

I joined John O'Doherty's lab as a rotation student in my first year. At the time it was just him and one postdoc, Hackjin Kim, in the corner of the basement of the Broad building. Having John almost all to myself for a time, I learned to do fMRI studies (for that matter, science) from a true expert, and I am thankful for the time and energy he spent training me. John has been a truly dedicated advisor, always encouraging and supportive, and most importantly a master at keeping projects on track. Thank you.

We were soon joined by my fellow CNS grad student Alan Hampton, and from there the lab blossomed to include Vivian Valentin, Klaus Wunderlich, Jan Glaescher, Eebie Tricomi, Vikram Chib, Saori Tanaka, and Kerstin Preuschoff. These have been a wonderful group of people to work with, and I feel extremely fortunate to have been a part of the O'Doherty lab at Caltech. It is sad to watch it dwindle with John's departure for the greener pastures of Dublin, but I'm sure the lab will be an equally cheerful and productive place to work at its new home.

Twice blessed, I also became a member of the Shimojo lab, which undoubtedly served to broaden my horizons. I am grateful to Shin for taking me on as a co-advisor, giving me great opportunities to discuss and present my work, and for many fruitful conversations on the Broad patio. Many thanks to friends in the Shimojo lab: Neil Halelamien, Daw An Wu, Virginie van Wassenove, Junghyun Paryk, Yasuki Noguchi, Iris and family, Michael Campos, Lindsay Lewis, and Zoltan Nadasdy.

I also had the opportunity to collaborate with Antonio Rangel and Bernard Balleine; they were both amazingly helpful and giving with their time. I would also like to thank my committee members, Richard Andersen and Ralph Adophs.

The CBIC is staffed with many great people, who made it much easier and more pleasant to do my job: Mary Martin, Steve Flaherty, and Ralph Lee. Extra-special thanks to Mike Tyszka, who was instrumental in getting my real-time project off the ground. I also enjoyed commiserating with fellow imagers Asha Iyer, Hilke Plassmann, and Todd Hare.

Many people helped me with the sometimes rough transition to Pasadena, Caltech, and grad school. A big thanks to my CNS classmates Neil Halelamien, Jonathan Young and Joe Levine (who did triple time as classmate, roommate, and friend), and good friends Guido Maretto, Chris Alabi and Alex Holub. Of course in my first year I also met a tall Canadian dreamboat named Paul, without whom I might have split long ago.

I would like to thank the ladies of GOGA for making Tuesday night the highlight of my week, swarm!, Jen and Lars, David and Liz, Sabrina (and Matt?), and of course all the folks I can still call friends despite short biannual visits and spotty email correspondence, you remain dear to my heart: Anouk, Veronique, Amanda, Nevena, Mariusz, Marty, Zainab, Jason, and Richy.

Most importantly I need to acknowledge my mum and dad, who not only raised me, loved and supported me, but also *strongly* encouraged me to go to Caltech despite my misgivings; seriously, they kinda pushed me to come here and without their nudging I might still be living happily in Toronto. The real reason for their urging revealed itself in time: Canadians love an excuse to visit SoCal, and I am so glad that they and my sister Jane were able to share this experience with me.

Abstract

Reward is a powerful modulator of behavior. Animals and humans are endowed with the ability to learn to associate events and actions with reinforcing stimuli, and flexibly adapt their behavior. The experiments described in this thesis use functional magnetic resonance imaging (fMRI) to study the neural mechanisms of reward learning in humans, the neural substrates by which reward associations influence behavior, and the neural plasticity that can be induced by provision of reward.

Attractive faces have been shown to be a form of visual reward, but their influence on behavior has yet to be characterized. We show that reward prediction errors in the nucleus accumbens are engaged when subjects learn associations between neutral cues and attractive faces, as has been shown with other reinforcers such as juice and money. This learning increases the subjective value of cues associated with attractive faces.

Animal studies have shown that Pavlovian cues can influence response vigor and decision-making. We present the first investigation into the neural mechanisms by which Pavlovian cues influence human decision-making. We find that activity in the ventral striatum differentiates between decisions to act in a manner compatible or incompatible with a concurrently presented Pavlovian cue.

In the next section we apply associative learning techniques to directly instrumentally condition neural activity, using reward feedback derived from fMRI images processed and analyzed in real time. This technique presents an alternative to standard bio/neurofeedback approaches and may prove useful in many clinical and research applications. We demonstrate that this method can be used to probe the causal influence of regional brain activity; specifically we test the impact of medial

orbitofrontal cortex (mOFC) activity on affective judgments. Subjects learn to elevate mOFC activity on cue and elevated mOFC activity increases the propensity to make a positive subjective valuation judgment.

Taken together these studies advance our understanding of the functional contributions of ventral striatum and orbitofrontal cortex in influencing decision-making and valuation, and illustrate the utility of applying associative learning techniques in combination with real-time fMRI in order to evaluate the causal contribution of specific brain regions toward particular cognitive functions.

Contents

List of Figures

List of Tables

Nomenclature

AC-PC line anterior commissure-posterior commisure line

BOLD response blood oxygenation level dependent response

EPI echo-planar imaging

FDR correction False discovery rate correction

HRF hemodynamic response function

NAcc nucleus accumbens

OFC orbitofrontal cortex

voxel 3D pixel

Chapter 1

Introduction

Preamble

If I examine the list of things I did this morning before sitting down to write this, I can evaluate each one in terms of the benefits it produced. I brushed my teeth; this made my mouth taste better, will prevent future trips to the dentist, and will improve my social interactions. I showered, dressed, and ate, all of which served to refresh and energize me for the day ahead. In fact, I would be hard pressed to think of any behavior I perform that does not have an expected consequence of delivering me from an unpleasant situation or improving the situation in which I find myself.

The ability to adapt my behavior in order to meet my basic and not-so-basic needs is one that I share with most animals. Natural selection should in fact favor animals who can flexibly adapt their behavior to changes in their environment. Indeed, animals from aplysia, to drosophila, to dogs, cats, and humans are endowed with the neural faculties required to learn.

The scientific study of animal learning took a huge leap forward in the late 19^{th} and early 20^{th} century, as researchers such as Thorndike and Pavlov began rigorous empirical studies of learning [1, 2]. The paradigms that they, along with Watson, Skinner, and others, pioneered laid the groundwork for studying learning behavior [3, 4]. Behavioral neuroscientists studying animals took this work a step further, side-stepping the nebulous issue of the mind by directly tapping into the brain, and quantifying the neural processes involved in learning.

Animal lesion and electrophysiology work has provided invaluable insight into the neural underpinnings of reward representation and learning. While animal learning is fascinating in its own right, many of us are interested in how these findings relate to human behavior. While studying humans is appreciably less amenable to the high degree of experimental control afforded in animals, studying humans directly is the only way of truly addressing this question.

Studies of human reward learning at the behavioral level have provided many insights, but methods for probing the neural bases of learning in humans were for a long time rather limited. The advent of functional magnetic resonance imaging (fMRI) in the late 90s, for whole brain functional imaging, facilitated significant advances towards the goal of understanding how the human brain processes and learns about rewards. This technique has allowed researchers to directly measure neural responses to reward and observe dynamic changes in neural activity due to learning.

fMRI studies of human reward learning have identified several key brain regions implicated in representing and learning about rewards. The experiments described in this thesis build on this work, and are primarily concerned with the neural mechanisms of learning, the neural substrates by which reward associations influence behavior and the neural plasticity that can be induced by provision of reward.

Rewards, reinforcers, and emotions

How do you teach an old dog a new trick? Reward performance with food. How do you train a student to study for a test? Reward correct answers with good grades. Clearly many different types of stimuli have the ability to elicit behavioral changes on the part of an organism. Thus, we define reinforcing stimuli not by their physical characteristics, but rather by the responses they elicit. Reinforcers are things that we seek to obtain or avoid; they are powerful modulators of behavior.

The term 'reinforcer' is sometimes preferred to reward, because reward has an emotional connotation of providing some kind of satisfaction. Many different conse-

quences can cause an animal to repeat a behavior, without necessarily putting the animal in a 'satisfied' state [5]. However, in this thesis we will generally use the terms 'reward' and 'punishment' to refer to positive and negative reinforcers.

While animals have dedicated sensory systems for detecting light, sound, and touch, they do not have dedicated receptors for reward. Rewards exist across a range of modalities, including auditory, visual, gustatory, olfactory, and social. What is rewarding to one animal, may not be to another, and will also vary depending on the animal's internal state. Rewards may cause measurable physiological responses, such as symptoms of arousal (pupil dilation, heart rate, respiration, blood pressure) and approach or avoidance behaviors [6].

Primary reinforcers are those that elicit responses without any prior learning, while secondary reinforcers become rewarding only through an association with a primary reinforcer. Money, for example, becomes a very powerful reinforcer over the course of human cognitive development, as it is associated with food, material goods, and social status.

Emotions, distinct from feelings, can be thought of as reactions to stimuli that move us. Under this definition, physiological responses to reinforcers qualify as emotions; paradigms such as fear learning have been used to study the neural basis of emotions.

Passive learning of reward prediction

While it is intuitive that animals able to learn to predict rewards in their environment will be more likely to survive and thrive, methods for objectively measuring associative learning in animals were not developed until the 1920s. Using a preparation designed to measure dogs' salivary secretions for the purpose of studying digestive processes, Ivan Pavlov first observed dynamic changes in physiological responses as a function of learning. Specifically, he noted that while salivation occurred in response to food delivery, it also occurred at other times, such as when the dog heard the experimenter's footsteps approaching. It became obvious to him that stimuli predictive of food

delivery could come to elicit responses similar to delivery of the food itself. By counting the drops of saliva produced in response to a stimulus, Pavlov's researchers were able to quantify learning. The importance of this discovery cannot be overstated. Since animal learning cannot be measured by subjective report, and even in humans subjective reports are not always reliable, this finding opened the door for empirical study of animal learning [6, 1].

To describe learning associations, Pavlov coined terminology that remains in common use. A primary reinforcer, or unconditioned stimulus (US), elicits an unconditioned response (UR).

US $\rightarrow$ UR

A conditioned stimulus (CS) can be paired with a US, which provokes a UR.

CS + US $\rightarrow$ UR

By association, the CS comes to elicit a conditioned response (CR).

CS $\rightarrow$ CR

Pavlov and the legions he inspired set out to describe the conditions which facilitated successful learning. Temporal contiguity was presumed to be an important factor in the strength of a learned association, and indeed this is the case. The strongest associations are those formed when the CS is present until the time of delivery of the US [1]. However, some types of conditioning, such as taste avoidance, are effective even if the US follows the CS by 24 hours [7], indicating that contiguity is not strictly necessary.

While contiguity can play a role in the strength of an association, it was later shown to be not only unnecessary but also insufficient for conditioning. In 1966, a paper by Robert Rescorla [8] highlighted the importance of contingency between the CS and the US. That is, he demonstrated that it is not simply the number of times that two stimuli are paired that determines how successfully the association will be learned. What is much more important is the relative number of pairings among the total number of times that the stimulus is presented. The predictive power of the CS over the US is diminished when the US is presented at times other than predicted by the CS. Therefore the effectiveness of a number of contiguous pairings in

creating an association can be severely degraded and even abolished by manipulating the contingency of the US.

The importance of contiguity was further reduced when Leo Kamin's influential 1969 paper on the phenomenon of blocking was published [9]. In this study he demonstrated that an outcome that is already fully predicted by a cue will not generate any new learning to a second cue, when the two cues are presented together and paired with the outcome.

CS1 + US $\rightarrow$ UR

CS1 $\rightarrow$ CR

(CS1 + CS2) + US $\rightarrow$ UR

CS2 $\rightarrow$ no CR

That is, learning to the second cue is blocked by the first cue, because the first cue already fully predicts the outcome. Thus, the contiguity between the second cue and the outcome does not seem to matter when the outcome is already fully explained by the first cue.

Computational models of reward learning

The blocking paradigm served as the foundation for one of the most influential theories of learning: learning requires an error in prediction. That is, in order for learning to take place, there must be a discrepancy between the actual outcome and the outcome that was expected. This rule was mathematically formalized by Rescorla and Wagner in 1972 [10].

The Rescorla-Wagner learning rule describes the process of value acquisition, in a trial-by-trial fashion. Over learning, the predictive stimulus acquires value V[i], where 'i' is the ith trial.

$\mathbf{V[i]} = \mathrm{V[i\text{-}1]} + \alpha\delta$

where

α is the learning rate

$$\delta = \mathrm{R[i]} - \mathrm{V[i-1]} = \text{prediction error}$$

In this model the prediction error signal is generated at the time of expected outcome, and influences the value of the cue on the subsequent trial.

The power of this rule lies in its ability to reconcile learning phenomena unexplained by previous theories. The phenomenon of blocking described above is gracefully accommodated: the first cue fully predicts the reward, therefore the prediction error is zero and the value of the second cue does not change.

An obvious shortcoming of this model is how it might account for the not unrealistic situation in which the time of reward delivery relative to the cue is variable. Reinforcement learning theorists [11] adopted this model as the inspiration for a temporally extended version in which trials are subdivided into a number of temporal epochs with value and prediction errors in each epoch. In temporal difference learning, the value signal represents the total expected value for the remainder of the trial; the timing of value onset shifts backwards in time from the time of outcome to the time of the cue. Upon initiation of conditioning trials, the cue is meaningless and does not elicit value or prediction errors, while the presentation of the outcome is unexpected and therefore generates a large error signal. After several similar trials, the value will have shifted temporally closer to the time of the cue. Finally, learning reaches an asymptote when the prediction error is positive at the time of the cue, and neutral at the time of the fully predicted reward, and the value signal is elevated starting at the time of the cue.

Reinforcing behavior: operant conditioning

Pavlovian conditioning describes a process of passive learning of reward associations, but animals can also learn to modify their behavior in response to their environment. Anecdotal reports of animals learning new and intelligent behaviors circulated in the late 19th century, prompting EL Thorndike [2, 6] to begin formally studying processes of animal learning of behavior. Using a device he termed a puzzle box, essentially

a wooden crate with a door that could be opened by a mechanism inside, such as a latch or rope, he studied the effects of reinforcement on cats' behavior. Hungry cats placed inside the box with a bowl of food in view outside would scramble around until they stumbled upon the response that would release them. When placed back in the box on subsequent trials they scrambled in a similar way, but the latency with which they performed the response to escape from the box gradually decreased over repeated trials. He deduced that the animals were not learning the physics of the latch mechanism, but rather that the reward of getting out of the box was 'stamping in' responses that led to this outcome.

Based on his observations, Thorndike formalized the Law of Effect, which basically states that the effect of an action has a strong influence over whether that action will be repeated [2]. He called this learning 'instrumental' because the animals in his studies learned to manipulate an instrument (e.g., a latch).

As with Pavlovian learning, many researchers have advanced our understanding of the principles of instrumental conditioning. Among the topics of interest were developing effective schedules of reinforcement, and understanding what was learned in instrumental conditioning: associations between the response and the outcome, or between the cue to respond and the response, or between the cue and the outcome? Several clever experiments have addressed this question, one which we will examine in more detail below and in Chapter 2.

An important methodological advance was the development of automatic tools for measuring responses. BF Skinner designed an 'operant chamber' also known as a Skinner box, to automatically measure responding and present reinforcers to experimental animals [3]. Skinner coined the term 'operant' to describe the responses made, because responses operate on the environment; the term 'operant level' refers to the baseline level of responding prior to learning. Thus the study of conditioning behavior with reinforcers became known as operant or instrumental conditioning.

Many types of reinforcers can exert control over behavior. Primary reinforcers such as food, water, sex, and sensory stimulation [12] do not require any learning to reinforce behavior. Secondary reinforcers that are learned through experience, such

as money or tokens, can be equally powerful.

Indeed, reinforcing behavior is a powerful tool; animals will respond for many hours for access to visual stimuli [13] or food rewards. But in order to reinforce a behavior, the behavior must first be performed, so how does one train a behavior that does not come naturally? One method for training complex behaviors is shaping, a method of approximations whereby responses are reinforced which take an incremental step towards the goal behavior [4]. Shaping has been applied with some success to smoking cessation [14] and academic task engagement [15], to cite only a few. However, in practice applying shaping techniques is still somewhat subjective. How large a step must the subject take in order to earn reinforcement? While some attempts have been made to formalize shaping procedures [16], finding optimal parameters is still largely a matter of trial and error.

Reward representation and learning in the brain

Rewards are stimuli that are defined by the responses they evoke in organism, rather than by their sensory properties. Rewards exist in gustatory, olfactory, visual, and auditory domains, and thus information about rewards is conveyed to the brain by a range of sensory modalities. It is therefore an important question as to whether these inputs converge in regions of the brain that respond specifically to the rewarding properties of stimuli, independent of modality. We know that rewards from different modalities can have a similar impact on behavior; it would therefore be parsimonious to have a common system for representing reward value which could be used for decision-making and expression of reward-mediated behaviors.

A related issue is how secondary or conditioned reinforcers are learned and represented in the brain. When Pavlov described conditioned reflexes, he hypothesized about the neural basis of this association. In particular he presumed that when an animal is presented with food, there is a food center in the brain which is activated and which in turn stimulates the physiological responses commonly seen to food. There is also a region of the brain which is stimulated by exposure to the conditioned

stimulus, for example areas which respond to the visual and auditory stimuli accompanying the arrival of an experimenter. He hypothesized that when these sensory and reward regions are co-active, the connections between them are strengthened, and this associative strengthening leads to expression of conditioned responses [6, 1].

This idea later gained strength at the synaptic level: the psychologist Donald Hebb proposed that 'cells that fire together, wire together', that is when neurons are coincidently active the strength of the ability of one cell to cause the firing of the other increases [17]. Empirical studies have since shown that there is some truth to this proposal [18].

Brain regions involved in both sensory and reward processing have been identified. However, we do not yet have a full systems-level description of the regional interactions supporting learning. Nonetheless, significant advances in understanding how the brain processes and learns about rewards have been made. Here we shall briefly review some of the known functions of reward-sensitive regions of the human brain.

Orbitofrontal cortex

Lesion studies in animals and humans have identified the orbitofrontal cortex (OFC) as important for representing the current value of stimuli. Lesions to this region cause deficits in tasks that require response behavior to flexibly adjust when the value of a given action changes [19, 20].

In terms of connectivity, the orbitofrontal cortex is well poised to integrate stimulus properties across a range of modalities, as it is highly connected to sensory processing areas, in visual, gustatory, and olfactory modalities [21, 22, 23]. Indeed it has been shown that OFC neurons are sensitive to stimuli from all of these domains. In non-human primates, both unimodal and multimodal food responsive neurons have been found in this region [24]; some cells respond preferentially to specific food objects independent of the modality in which they are presented.

While the sights, tastes, and smells of foods are certainly rewarding, the possibility remained that OFC activity represented the sensory properties of stimuli, rather

than their reward value per se. Sensory-specific satiation procedures have been used to address this question [25]. In these experiments, monkeys were presented with differently flavored food-related stimuli in visual, olfactory [26], or gustatory forms [27], and subsequently fed to satiety on one of the foods. Following satiation, OFC neurons showed decreased responding specifically to stimuli related to the food on which they had been satiated, independent of the modality in which the stimulus was presented. These results suggest that these OFC neurons are sensitive to the current value of the food stimuli.

While single/multi-unit recordings are less feasible in human subjects, whole-brain imaging techniques such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have been useful in delineating human brain regions responsive to rewarding stimuli. A typical methodology for studying reward representation in the brain is to present subjects with stimuli from a common sensory modality but differing in reward value, for example O'Doherty et al. [28] compared neural responses to the taste of glucose (pleasant), neutral, and salty (unpleasant) solutions. These authors found non-overlapping differential activity in the OFC in response to both pleasant and unpleasant flavors, relative to the neutral flavor, suggesting that OFC encodes both positive and negative valence, but potentially in distinct neural populations. Similar studies have been performed in the olfactory [29], visual [30, 31], and auditory [32] domains, demonstrating similar patterns of OFC activity. Human OFC responses also appear to be sensitive to the current value of stimuli: specific satiation on a food item significantly decreases OFC responses to that food [33].

The OFC also responds to secondary reinforcers such as money and social feedback [34, 35, 36], and cues learned as predictors of reward [37]. Sensory-specific satiety has again been used to test whether the OFC is sensitive to changes in the value of a reinforcer with which a cue has been paired [38]. Indeed, OFC activity selectively decreases in response to a Pavlovian cue when the outcome predicted by that cue is no longer valuable.

Several studies have made a distinction between medial OFC and more lateral

parts of OFC as representing pleasant and aversive stimuli respectively [28, 31, 39]. Although this medial/lateral distinction is not universally found in the literature, activity in medial OFC seems to correlate specifically with increasing subjective value [31, 40, 41].

Thus, converging evidence points to the OFC as an important structure for representing current value. A task that has consistently shown medial OFC responses is probabilistic reversal learning, in which subjects are required to choose between stimuli delivering reward probabilistically. However, the probabilities of reward delivery are sometimes reversing, forcing subject to pay close attention to reward rates and update the values of the cue stimuli. In Chapter 4 of this thesis, we use probabilistic reversal learning as a functional localizer for medial OFC, and probe the specific impact of elevated activity in this region on affective judgments.

Amygdala

The amygdala, a pair of almond shaped nuclei located bilaterally deep in the temporal lobes, have also been implicated in expression of emotion and emotional learning. The amygdala are composed of a heterogeneous group of sub-nuclei, differing in both composition and connectivity. These sub-nuclei are reciprocally connected to each other, and other brain regions, in such a way as to be well poised to integrate information about, and associate, unconditioned (US) and conditioned stimuli (CS) (for a review see [42]).

Early reports of temporal lobe resections in animals described a range of aberrant behaviors, including a lack of fear responses. The disrupted fear response was later localized to the amygdala [43]; thus, much of the work on the amygdala has concentrated on fear responses. Pavlovian fear conditioning has been an important model paradigm for understanding acquisition and expression of fear, as well as emotional processing in the brain more generally. In particular it has been shown that lesioning the basolateral complex (BLA) interferes with acquisition [44, 45], while lesioning the central nucleus (CE) interferes with expression of learned fears [46, 47]. Human

patients with amygdala lesions also fail to acquire conditioned physiological fear responses, despite having explicit knowledge of, and normal responses to, the aversive outcome [48].

The amygdala are also involved in representing appetitive stimuli. Single unit recordings in amygdala during a task in which a cue predicts first an appetitive stimulus and later switches to predicting an aversive stimulus, showed that some amygdala cells reflect positive predicted valence, while others reflect a negative predicted outcome; the activity in these cells changed to reflect the reversal in stimulus contingencies [49]. Human fMRI studies have corroborated amygdala involvement in representation of both appetitive and aversive emotional stimuli. Both happy and fearful faces activate the amygdala, relative to neutral faces [50], as do faces which have been associated with either positive or negative emotional characteristics. The amygdala also responds to both pleasant and aversive taste [28].

Evidence points to the involvement of human amygdala in representation of not only primary but also learned reinforcers. Patients with amygdala lesions show impaired acquisition of conditioned preference relative to healthy controls and patients with frontal lesions [51]. fMRI studies of reward learning have implicated amygdala in representing the current value of stimuli [38]. In conditioning, it has frequently been found that responses to the US decrease with repeated presentation, a process known as habituation. Some studies have also found that amygdala response to the cue habituates with repeated presentations [52, 53], which is related to behavioral habituation. We will discuss the amygdala further in terms of Pavlovian learning and expression of learned behaviors in Chapters 1 and 2.

Mesolimbic dopamine system

In 1954, Olds and Milner [54] reported that animals would work (e.g., lever press) for direct electrical stimulation to certain parts of the brain. This discovery led to a method for mapping out which brain regions animals seek to stimulate by means of natural reinforcers. Dopamine is a neurotransmitter produced primarily by neu-

rons in the midbrain. Neurons in the mesolimbic dopamine system project from the ventral tegmental area (VTA) of the brainstem to frontal and temporal cortices and limbic structures of the basal forebrain such as the nucleus accumbens. Olds and Milner found that stimulation to the nucleus accumbens, a known target of dopaminergic neurons, was a very powerful reinforcer. In fact, stimulating all along this dopaminergic pathway elicits strong reward responses. Similarly, direct injections of rewarding drugs provide a neurochemical means for studying reward function. Animals will self-stimulate the nucleus accumbens with amphetamine [55], a dopamine releaser, and nomifensine or cocaine [56], which inhibit dopamine reuptake, indicating that the presence of dopamine in this region is rewarding (for a review see [57]).

Dopamine responses to natural reinforcers

Electrophysiology work in animals has helped to clarify the role of dopamine in reward representation and prediction. Approximately 75% of dopamine neurons increase their firing rates in response to unexpected rewards [58]. Their response is relatively indiscriminate among different types of food or liquid reinforcers, suggesting that they respond to the reward value as opposed to the specific sensory properties of each stimulus. However, these cells do distinguish between food and non-food objects [59]. Aversive events provoke a phasic increase in activity in only about 14% of cells, but cells do show depressions or activations with slower time courses [60], suggesting that aversive events may be coded by a depression rather than an activation.

Reward predicting stimuli, learned through Pavlovian or instrumental conditioning tasks, elicit activation in 55-70 % of dopamine neurons [61, 62]. Conditioned stimuli are somewhat less effective than actual reinforcers, however they are similarly indiscriminate among reinforcers and preferentially responsive to cues predictive of appetitive stimuli [62].

Dopamine signals and learning

Of particular interest is the response of dopaminergic neurons over the course of learning. Before learning has taken place and rewards are unexpected, dopamine

neurons fire in response to primary rewards. However, once a predictive cue for reward has been established, it is the cue which elicits dopaminergic firing, and not the reward itself [63]. There is a transient learning period during which both the cue and reward elicit some amount of activity, but after learning only the predictive cue and unexpected rewards generate a response [63, 61].

This pattern of responses bears a strong resemblance to the prediction error signals postulated by learning theorists [11]. An important similarity between neuronal responses and a prediction error signal, derived from a temporal difference model, is temporal sensitivity. Dopamine neurons respond when rewards arrive earlier or later than predicted, even when it is certain that the rewards are to occur [64]. Dopamine neurons are depressed when a predicted reward is omitted; this depression occurs at the specific time the reward was predicted [64].

Kamin's blocking paradigm [9] was an important source of inspiration for the development of prediction- error-based theories of learning, and makes specific predictions about the role of expectation and surprise in learning. If dopaminergic neurons were in fact coding for something like a prediction error, responses in a blocking test should comply with the observed behavior [9, 65]. That is, learning of an initial cue-outcome association:

CS1 → outcome

CS1 → CR (licking response)

should block learning of the predictive power of a second cue when the two cues are presented together followed by the outcome:

CS1 + CS2 → outcome

In a blocking paradigm tested in monkeys, Waelti et al. [66] found that both licking behavior and dopamine responses complied with a prediction-error based account of blocking. The blocked cue did not elicit any licking behavior or dopaminergic firing:

CS2 → no licking response

A second pair of stimuli were used to control for repeated exposure and to show that learning can take place for one member of a pair of cues. A cue was first presented alone, and then with a second cue predicting reward:

CS3 → no outcome

CS3 + CS4 → outcome

The result was that the added cue became predictive of reward, as shown by both licking behavior and dopaminergic firing:

CS4 → licking response

These studies provide very strong evidence that dopaminergic signals from the midbrain play an important role in learning. However, it is important to note that this is not the only role played by dopamine in the brain. Patients with Parkinson's disease show deficits in movement, cognition, and motivation as a result of degeneration of the nigrostriatal dopamine system. This and evidence from lesion and pharmacological manipulation studies [67] point to a role for tonic and medium scale (on the order of seconds to minutes) dopamine release in wide range of behaviors [68].

Functional MRI of reward prediction error signals

Mapping out the activity of reward responsive regions in the human brain quickly turned from characterizing responses to primary rewards, to understanding how conditioned stimuli acquire value. Animal studies provided very compelling evidence that the mesolimbic dopamine system is involved in signaling reward prediction errors related to learning. The search for prediction error signals in the human brain began with the hypothesis that unexpected rewards should elicit increased dopaminergic activity. This was tested in a block-design fMRI study by [69]; these authors compared neural responses during a block when juice rewards were delivered at unexpected times, to a block during which the same number of juice rewards were given at predictable intervals. They found significantly greater activity in the ventral striatum/nucleus accumbens during the block of temporally unpredictable rewards.

Pagnoni et al. [70] further demonstrated that a temporally delayed reward elicits a positive deflection at the time that the delayed reward is delivered. However, this is but one of the three components of prediction-error-based models. These models predict that 1) unexpected rewards should elicit phasic activity, 2) omitted expected rewards should elicit a depression in activity, and 3) fully predicted rewards should

not elicit any change. O'Doherty et al. [71] demonstrated that responses in the ventral striatum and orbitofrontal cortex correlate with the full range of prediction error activity. Further evidence for temporal prediction errors in the human brain comes from Seymour et al. [72], who demonstrated that the time course of striatal BOLD responses comply with predictions from a temporal difference model for a range of situations in a second-order conditioning procedure.

Prediction-error-related activity in the ventral striatum has subsequently been demonstrated for a wide range of natural and secondary reinforcers such as food [73], money [74], and pain [72]. As discussed above, the blocking paradigm is an important test for determining whether behavior and neural activity comply with formal learning theory. Tobler et al. [73] tested blocking with human fMRI and demonstrated that indeed, blocking behavior was evident in a subset of subjects, and in these subjects activity in the ventral striatum showed phasic activity in response to a non-blocked cue relative to a cue for which learning had been blocked.

Prediction error signals to visual rewards

While our understanding of human neural responses to rewarding stimuli is arguably still preliminary, whole-brain fMRI studies have shown remarkable consistency in correlating activity in the OFC, amygdala, and striatum with specific aspects of reward processing. Meanwhile, the reward value of specific classes of stimuli, for example social stimuli [75], and the associated neural responses, continues to be probed. Attractive faces have recently been shown to be a form of visual reinforcer; male subjects are willing to exert effort to prolong viewing of attractive female faces [30]. This block design fMRI study and another event-related fMRI study of passive viewing of faces [31] both showed increased neural activity in reward structures such as OFC and the ventral striatum in response to attractive, relative to unattractive faces.

In Chapter 1 of this thesis we extend this work, exploring the effects of exposure to attractive faces on behavior, and the neural bases of this effect. We demonstrate that attractive faces can serve as a visual reinforcer in a classical conditioning task, and

that some of the value of attractive faces transfers to a previously neutral cue over the course of learning. We further provide evidence that a reward prediction error signal in the ventral striatum is engaged during classical conditioning with attractive faces.

Pavlovian-instrumental transfer

Operant conditioning, in which an animal learns to perform an action in order to modify their environment, is one example of how reward learning influences behavior. By definition, Pavlovian learning affects passive responding, but has additionally been shown to influence active responding. It is quite natural to think that Pavlovian associations can influence decisions: once a stimulus has acquired value through Pavlovian learning, decisions made concerning that stimulus may be affected. For example, foods are often associated with emotions felt at the time of their consumption, affecting future choices related to those foods; so-called 'comfort foods' are increasingly prevalent in cookbooks and on restaurant menus [76].

These effects have been formally studied with Pavlovian-instrumental transfer paradigms, which test the ability of a Pavlovian cue to influence an instrumental response. This occurs despite no formal learning of the effect of performing the action in the presence of the cue. [77, 78] first described that a cue predictive of an outcome could cause an increase in the rate of performing a response that had been associated with the same outcome.

1. Tone → food delivery
2. Lever press → food delivery
3. Extinction: Lever press → no food delivery
4. Tone: Lever press response increases

The implication was that the rats had learned a stimulus-outcome association for the first association that could then exert control over other behaviors resulting in the same outcome.

This work was important for theories of instrumental learning, which had long

been concerned with what exactly is learned during instrumental conditioning. Do animals learn response-outcome, stimulus-outcome or stimulus-response associations? Arguably, Pavlovian-instrumental transfer paradigms provide evidence that both stimulus-outcome and response-outcome associations are learned [79], since the stimulus could not affect the response unless they had both been associated with the outcome. Extensions of the original transfer paradigm have shown that Pavlovian cues can bias action choice towards responses associated with specific outcomes, in what is known as outcome-specific transfer. An example of this can be seen in the work of Blundell et al. [80], who trained rats to respond on one lever (on the left) for sucrose and another lever (on the right) for pellets.

L → sucrose

R → pellets

They also trained the rats on Pavlovian associations between auditory stimuli and reward: a click train predicted sucrose while a tone predicted pellets.

Click-train → sucrose

Tone → pellets

Following Pavlovian and instrumental training they performed a transfer test, during which the animals were exposed to both levers in extinction (no reinforcement was provided), and the auditory cues were presented in alternation with baseline periods during which no Pavlovian cues were presented. During Pavlovian cue presentation, the rate of responding increased, and was significantly higher on the lever associated with the same outcome as the Pavlovian cue.

Click train: L (increased)

Tone: R (increased)

A Pavlovian cue can also exert a non-specific influence over responding, that is it can have a more general influence on motivation as measured by response vigor. For example, several studies have shown that a Pavlovian association learned during one drive state (stimulus predicts food when hungry) can influence responding under a different drive state (lever predicts water when thirsty) [81, 82].

A growing body of work is concerned with investigating the neural mechanisms of

transfer. Lesion studies in rats have shown that regions of nucleus accumbens [83], amygdala [84], and dorsal striatum [85] selectively abolish certain transfer behaviors. In humans, Talmi at al. [86] demonstrated that general motivational enhancement was correlated with activity in the nucleus accumbens and amygdala. In Chapter 2 we present the first investigation of the neural correlates of outcome-specific Pavlovian-instrumental transfer in humans. We scanned human subjects with fMRI while they underwent sessions of both Pavlovian and instrumental learning, followed by a transfer test. During the transfer test we were able to discern which regions of the brain were preferentially engaged when subjects acted under the influence of the Pavlovian cue.

Direct instrumental conditioning of neural activity

As we have seen, reward is a powerful modulator of behavior, and necessarily exerts influence over neural activity in the process; environmental stimuli come to evoke behavioral responses as outcome contingencies are learned. Accompanying any behavioral response is a neural response; at the very least we can assume a motor command to execute the action. It would therefore seem possible that if we could record neural activity and make reward feedback contingent on that activity, we could condition a neural response directly, perhaps even in absence of overt behavior.

Fetz developed a technique for recording single cell activity and making reward contingent on neural firing rates, to study the activity of motor cortex neurons in relation to muscle activity [87]. It had previously been shown that human subjects can learn to control the activity of single motor units (consisting of a spinal anterior horn cell, its axon, and all of the muscle fibers on which the terminal branches of the axon extend) [88], and that neural activity can be conditioned in rats [89]. Fetz [87] combined these techniques and described isolating single motor neurons and reinforcing the animal when neuronal firing rates were elevated above operant levels. Reinforcing elevated firing rates proved to be an effective technique for training an animal to learn operant control of newly isolated cells, whose firing rates increased 50-500% in response to reinforcement.

Later work showed that it was also possible to train monkeys to perform this response while suppressing motor activity [90], and to some extent suppress neural activity while performing a muscle contraction. This work demonstrated that neural activity previously associated with a motor response can be dissociated from the response, implying that while neural activity was correlated with the response it might not have been strictly necessary for its performance. Another possibility is that the neural activity was necessary to generate the motor response, but rapid plasticity occurred, recruiting different motor cortex cells for performance of the response.

This work exemplifies how operant training of behavioral and neural responses can be used to investigate the precise causal relationship between local brain activity and behavior. Measuring brain activity during performance of a task provides information about correlations between brain and behavior, but from this evidence alone we cannot infer that they are causally related, i.e. that the behavior could not be performed in absence of the brain activity.

Causality can be investigated using lesion studies: a brain region can be inferred to be necessary for a task if elimination of that region renders task performance impossible. However it can be difficult to find subjects with very specific lesions and this method does not allow testing of the impact of varying levels or patterns of neural activity on behavioral performance. Temporary lesions induced by transcranial magnetic stimulation (TMS) are another method for establishing causal relations between brain activity and behavior [91]. However, these techniques are limited in spatial scope (most stimulating techniques are active only at the cortical surface). Training subjects to voluntarily activate or suppress neural activity in specific brain regions has the potential to complement these techniques. An interesting avenue of investigation, not possible with lesion methods, is that if fine-grained control over regional activity can be achieved, the effects of varying levels of activity on behavior can be investigated.

Training subjects to voluntarily control neural activity that can be measured by an external device also raises the possibility of using these neural recordings as a communication tool. Locked-in syndrome describes a condition in which a patient is

paralyzed and has lost the ability to voluntarily control muscle activity except for eye blinking. A large body of research has been concerned with helping these patients, and also those with less severe paralysis, to communicate. It has been demonstrated that human subjects can learn to modulate recordings of electroencephalographic (EEG) recordings of neural activity at the scalp [92, 93]. Multi-unit recordings in monkeys have also demonstrated that signals from posterior parietal [94] and motor cortex [95] can be interpreted and used for control of a cursor or physical manipulandum.

Recent technical advances have allowed development of neurofeedback techniques using functional MRI data processed in real time (see Appendix A for more details). Despite the disadvantage of poor temporal resolution, fMRI has the very important advantage of allowing whole-brain imaging at much finer spatial resolution than that afforded by scalp EEG, while remaining less invasive than intracranial recordings.

Several groups have provided proof-of-concept that subjects can learn to enhance, and sometimes suppress, brain activity in specific regions including motor cortex [96], amygdala [97], auditory cortex [98], supplementary motor area and parahippocampal place area [99], and rostral anterior cingulate cortex (rACC) [100]. A particularly interesting application of this technique is choosing regions in which it is hypothesized that elevated activity will result in specific changes in behavior. A nice demonstration of this was provided by deCharms et al. [100], in which subjects were trained to elevate and suppress activity in the rACC, a region involved in pain processing. They showed that when subjects were successfully modulating activity in this area, their perception of a painful stimulus was enhanced or suppressed.

Previous studies have typically used ongoing graphical feedback of neural signals to facilitate learned control. In general it is unclear what kind of feedback is necessary or most effective. In typical instrumental conditioning studies, reinforcing behavior with reward is sufficient to modify behavioral responses. We were therefore interested in investigating whether reward would be sufficient to induce instrumental learning of neural responses. In Chapter 3 we present a series of experiments in which we rewarded subjects for elevating brain activity in regions of motor cortex related to hand and foot movements, respectively, on separate trials. We demonstrate that

subjects can learn to modulate neural activity in this region using only monetary reward feedback. We further investigated the impact of training on behavior by measuring reaction times in a cued response task.

Next, we extended this work and attempted to condition neural activity in an emotional brain region: the orbitofrontal cortex; this work is described in Chapter 4. Activity in medial OFC correlates strongly with the subjective value of stimuli; we were interested in investigating the causal influence of OFC activity on subjective judgments. To this end, we asked subjects to judge the attractiveness of rapidly presented faces interspersed with the OFC-activate blocks. We included a control condition in which subjects tried to regulate activity in a hand-motor area, in order to control for the effects of elevating neural activity in order to obtain reward.

Contributions of this thesis

Attractive faces can act as visual reinforcers: male subjects are willing to exert effort in order to gain access to attractive female faces [30], and fMRI studies have shown that attractive faces engage known reward circuitry in the brain, such as the orbitofrontal cortex and nucleus accumbens [30, 31]. Advertising campaigns have long paired products with pictures of attractive models in order to enhance the desirabiliry of their product, a practice supported by marketing research [101, 102]. However it has remained unclear how attractive faces exert an influence over behavior. In Chapter 1 we investigate the possibility that attractive faces can transfer value to cues via classical conditioning, and use fMRI to investigate neural signals related to learning.

In Chapter 2 we take a deeper look at the influence of Pavlovian cues on behavior, specifically related to decision-making. Pavlovian cues predictive of a rewarding outcome can influence response behavior when presented during a period of instrumental responding [77, 78], despite no learning of the consequences of performing the response in the presence of the cue. In general transfer, the Pavlovian cue may be associated with an outcome other than those available on the instrumental responses, and causes a general enhancement in response vigor [82, 84, 86]. In specific transfer,

a Pavlovian cue associated with a particular outcome will bias responding towards a response associated with the same outcome as the cue [85, 84, 83, 80]. Lesion studies in animals have probed the neural circuitry involved in expression of transfer effects; these studies have found that partly dissociable regions of nucleus accumbens [83], amygdala [84] and dorsal striatum [85] are necessary for expression of general and specific transfer effects. However these studies have not addressed the mechanism by which transfer effects are expressed. In Chapter 2 we implement an outcome-specific Pavlovian-instrumental transfer paradigm in humans, and use fMRI to investigate the neural mechanisms underlying the influence of Pavlovian cues over decision-making.

Recent advances in fMRI technology have made it possible to perform fMRI analyses in real-time, and present subjects with information about their neural responses. This technique has been used to train human subjects to learn control over neural activity in circumscribed brain regions [96, 100, 103, 98, 104, 105, 106, 99, 107, 108]. In Chapters 3 and 4 we investigate the use of this technique for instrumentally conditioning neural activity directly. We apply associative learning techniques, providing subjects with reward for making specific neural responses upon presentation of discriminative cues. This technique may prove important for a host of research and clinical applications. In Chapter 3 we apply this technique to conditioning regions of motor cortex related to hand and foot movements. In Chapter 4 we extend this work to condition increased activity in medial orbitofrontal cortex, a brain region in which has been frequently correlated with subjective value [30, 109, 40, 31, 41], and probe the behavioral effects of increased activity in this area on an affective judgment task.

Chapter 2

Prediction error signals to attractive faces*

Attractive faces can be considered to be a form of visual reward. Previous imaging studies have reported activity in reward structures including orbitofrontal cortex and nucleus accumbens during presentation of attractive faces. Given that these stimuli appear to act as rewards, we set out to explore whether it was possible to establish conditioning in human subjects by pairing presentation of arbitrary affectively neutral stimuli with subsequent presentation of attractive and unattractive faces. Furthermore, we scanned human subjects with fMRI while they underwent this conditioning procedure in order to determine whether a reward prediction error signal is engaged during learning with attractive faces, as is known to be the case for learning with other types of reward such as juice and money. Subjects showed changes in behavioral ratings to the CS stimuli when comparing post- to pre- conditioning evaluations, notably for those CSs paired with attractive female faces. We used a simple Rescorla-Wagner learning model to generate a reward prediction error signal and entered this into a regression analysis with the fMRI data. We found significant prediction-error-related activity in the ventral striatum during conditioning with attractive compared to unattractive faces. These findings suggest that an arbitrary stimulus can acquire conditioned value by being paired with pleasant visual stimuli just as with other types of reward such as money or juice. This learning process elicits a reward prediction error signal in a main target structure of dopamine neurons: the ventral striatum. The findings we describe here may provide insights into the neural mechanisms tapped into by advertisers seeking to influence behavioral preferences by repeatedly exposing consumers to simple associations between products and rewarding visual stimuli such as pretty faces.

*Adapted with permission from: Bray S, O'Doherty J (2007) Neural coding of reward-prediction error signals during classical conditioning with attractive faces. Journal of Neurophysiology 97:3036-3045. Copyright 2007 *Journal of Neurophysiology*

Introduction

Faces convey a wealth of information, and are perhaps the most important visual stimuli for humans in social environments [110]. The attractiveness of a face is a feature which we can perceive quite automatically [111], can subsequently motivate our behavior in terms of mate choice [112], and bias our beliefs about others' personality and expected success in life [113]. The effect of attractiveness on human behavior has been documented in the workplace, where it has been shown that attractive individuals enjoy higher salaries [114] and better employment prospects [115]. These observations have led to the suggestion that preference for facial attractiveness may have evolved to enhance reproductive success [116, 117].

Recent evidence indicates that attractive faces may act as a form of visual reinforcer, as human subjects are prepared to work in order to gain access to them [30]. Although much is now known about the neural circuitry involved in processing the perceptual [118, 119, 120] and affective aspects of facial stimuli [110, 121, 122], the neural substrates of facial attractiveness are much less well understood. Nevertheless, some preliminary studies investigating processing of facial attractiveness have implicated brain regions known to be involved in reward processing, such as the orbitofrontal cortex and ventral striatum [30, 123, 124, 125, 31, 109].

Here we aim to address the manner in which facial attractiveness can influence one important aspect of human behavior: behavioral preference. Attractive faces have long been used in advertising as a means of modulating behavioral preferences for specific products. Indeed, marketing research has shown that people will evaluate products more favorably when they are presented alongside physically attractive models [102, 101]. One possible mechanism for this preference modulation effect is through classical conditioning, whereby an arbitrary neutral stimulus acquires affective value through repeated pairing with a stimulus that has pre-established value such as an attractive face.

In this study we set out to elucidate the neural mechanisms of this phenomenon, by scanning human subjects with functional magnetic resonance imaging (fMRI) while

they learned an association between arbitrary affectively neutral visual stimuli (fractals) and attractive and unattractive male and female faces. Before and after the conditioning procedure, we took ratings of pleasantness and preference for the arbitrary fractal stimuli, in order to establish whether behavioral preferences for these stimuli could be modulated as a function of conditioning with attractive faces.

We aimed to characterize the neural processes underlying learning of these preference associations. Modern learning theories propose that such reward-dependent learning is driven by the degree of surprise or unpredictability of a rewarding outcome, or more specifically, errors in predictions of reward [10]. Electrophysiological studies in non-human primates implicate the phasic firing of midbrain dopaminergic neurons in encoding reward prediction errors [126]. fMRI studies of human learning have found evidence of reward-prediction-error-related activity in known projection sites of dopaminergic cells, especially the ventral striatum, during learning with other forms of natural and abstract rewards such as juice or money [124, 127, 73]. Given that attractive faces can also be considered as a form of reward, we hypothesized that learning with attractive faces would also engage brain structures known to be involved in reward prediction error coding such as the ventral striatum.

Materials and Methods

Subjects

28 subjects participated in this study (15 females and 13 males), ranging in age from 18 to 27 (mean 20.8 $\pm$ 2.24 S.D). All subjects gave informed consent, which was approved by the local research ethics committee. Due to technical difficulties (for one subject the experiment stopped during the study due to a software problem, and for two other subjects part of the data set was lost in transfer), three subjects were excluded from the imaging analysis (N = 25), one subject was excluded from response time analysis (N=27) and one subject's preference data were lost (N=27).

Stimuli

The visual conditioned stimuli (CS) were complex abstract fractal images, and the unconditioned stimuli (UCS) were photographs of human male and female faces, attractive and unattractive. A set of 148 faces were previously rated by a separate group for attractiveness on a scale from 1 to 7 [31]. Based on these ratings, 8 faces were chosen to make up each of four conditions: female-attractive, male-attractive, female-unattractive, and male-unattractive. The faces had forward head position, and gazed forward with neutral to mildly happy expressions. The face images were masked to remove hair, were adjusted to be of approximately equal size and luminance, and centered in a 450 x 450 pixel grey background. We also used six abstract fractals, each centered in a 170 x 170 pixel grey background. Stimuli were presented at a screen resolution of 800x600. Example stimuli are shown in the time course of a conditioning trial in Figure 2.1a, and additional example face stimuli are shown in Figure 2.1b. Stimuli were presented using Cogent 2000, developed by the Cogent 2000 team at the FIL and the ICN, and Cogent Graphics, developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience.

Behavioral measures

Sexual orientation

Subjects first completed a questionnaire, in which they were asked to describe their sexual orientation by choosing from a set of labels (heterosexual, homosexual, bisexual, transgender, polyamorous, none). They were also asked to rate on a 10-point scale how interested they are in having sex with men and women and how sexually attractive they find men and women.

Behavioral measures of learning and preference modulation

During conditioning, attractive and unattractive faces were paired with affectively neutral fractal pictures. Subjects were first exposed to the fractal stimuli before conditioning, in order to obtain pleasantness ratings and preference rankings. Pleas-

antness ratings were performed once subjects were installed in the scanner, before fMRI data collection began. Subjects were first shown a screen with six fractal images; they were then shown each of the fractals once in random order and asked to verbally report a pleasantness rating between -10 and +10 (where -10 = very unpleasant, 0 = neutral, 10 = very pleasant). Next, subjects were presented with pairs of fractals and asked to indicate which of the two they preferred by pressing the left or right button on a two-button pad. Pairs were presented in random order, with each combination presented three times, and fractals randomly assigned to the left or right side of the screen. Subjects responded to a total of 45 pairs, with each fractal appearing a total of 15 times.

During the conditioning procedure, subjects were asked to respond with a button press to indicate which side of the screen the fractal stimulus was presented on each trial. During conditioning, subjects were presented with each fractal a total of 48 times. These reaction times provided an additional on-line measure of conditioning [128].

After conditioning, the preference and ratings tasks were repeated in that order; subjects were given the additional instruction that they should not try to match their previous answers, but rather respond according to their present evaluation.

In order to assess explicit awareness of the contingencies, subjects were shown each of the six fractals in random order and asked how likely they thought it was that the fractal had been paired with an attractive face, using a scale from 0 to 10 (where 0 = not at all likely, and 10 = very likely). Subjects were also asked how unlikely it was that a stimulus was paired with an attractive face. We then asked subjects how likely and unlikely it was that each fractal had been paired with an unattractive face.

Evaluation of attractiveness of face stimuli

The final task in the experiment was to evaluate the attractiveness of the faces. Subjects were presented with each of the 32 faces in random order and asked to verbally report a subjective rating of facial attractiveness on a scale from -10 to +10.

Behavioral Data Analysis

We used differential ratings and preferences as an index of conditioning. We hypothesized that fractals paired with attractive faces would increase in pleasantness and become more preferred over fractals paired with unattractive faces, especially for fractals paired with faces opposite in gender to the subject. A preliminary inspection of these data indicated that they were not normally distributed. Consequently, we used non-parametric statistics for all behavioral analyses in this study.

The pair-wise preference results were ranked based on the number of times a fractal was chosen as preferred, and category differences in ranking changes before and after conditioning were compared (e.g., change in rankings for fractals paired with attractive female faces compared to unattractive female faces).

Neuroimaging

Conditioning Procedure

Four of the fractals were randomly assigned to be paired with faces from one of the four face gender/attractiveness categories, and two were assigned to never be paired with any faces. The fractal/face categories were: attractive female, attractive male, unattractive female, unattractive male, and unpaired. Each trial began with the presentation of a fractal image, randomly displayed either to the left or right of a central fixation cross. This fractal remained on the screen for 1.5 seconds. On reinforced CS+ trials, after 1 second a picture of a face appeared in the middle of the screen, next to the fractal. The two appeared together for 500 milliseconds, the fractal then disappeared while the face remained on the screen for another full second, followed by a fixation cross for 500 milliseconds. The duration of each trial was 3 seconds, with the face and fractal each presented for 1.5 seconds, with 0.5 seconds of overlap. We chose to use a delay conditioning paradigm with a short inter-stimulus interval in order to maximize conditioning efficacy. The time course of a CS+ trial is shown in Figure 2.1. In order to enhance conditioning, the first three trials of each condition were reinforced CS+ trials in which the face followed

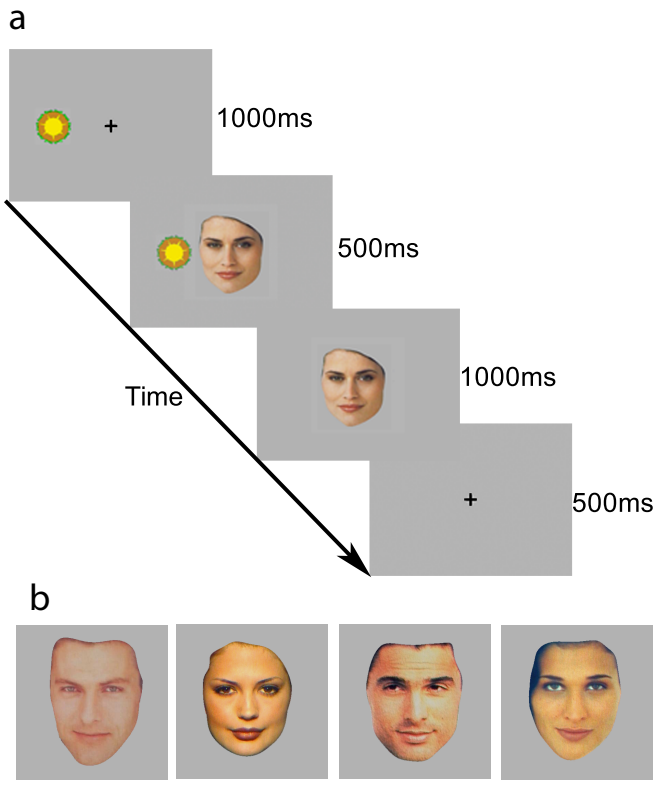

Figure 2.1: a) Sample time course of a reinforced CS+ conditioning trial with a female attractive face. b) Example face stimuli

the fractal, while for the remainder of the experiment 50% of trials were reinforced. In order to obtain a trial-based behavioral measure of learning during the study and also to ensure that subjects attended to the task, subjects were instructed to press the left or right button on a two button pad, to indicate which side of the screen the fractal appeared. They were also instructed to keep their attention directed toward the center of the screen throughout the experiment. There were 48 trials of each type, 50% of which were reinforced, and each of the 8 faces in a category was presented up to three times. Along with the 288 conditioning trials we included a set of 96 null events, during which the fixation cross was presented for 3 seconds, in order to mimic the effect of a jittered inter-trial interval and facilitate separation of neural responses from consecutive trials. Trials were randomly ordered, and the total duration of the conditioning session was approximately 20 minutes.

Prediction error signals

We used a simple trial-based Rescorla-Wagner rule to model trial-by-trial prediction errors in learning [10]. This model uses a prediction error signal δ which reflects the difference between the value of the outcome received on a given trial (R) and the value of the expected outcome on that trial (V): $\delta = R - V$. The expected value V is then updated by adding delta weighted by a learning rate α : V=V+$\alpha\delta$.

In a follow up region of interest analysis, we employed a real-time extension of the Rescorla-Wagner learning rule, a temporal difference model [11, 129], in which the prediction error shifts backwards in time from the face presentation to the cue presentation. We divided trials of each type in to early, middle, and late epochs, and modeled the prediction error signal at the time of the face, 0.5s before the face, and time of cue (1s before the face), respectively.

The specific values used in our implementations of the model were the following: we modeled the presentation of a face with R = 1, the omission of a face with R = 0 (for both attractive and unattractive faces), and derived the learning rate (α) from subjects' behavioral responses. We used reaction times (responses to the conditioned stimuli) as a trial-by-trial measure of learning, to derive model parameters from subjects' behavior. Reaction times have previously been shown to be modulated as a function of conditioning, and changes in reaction times over time have previously been found to correlate with reinforcement learning models [130, 131, 38]. We derived learning signals for each subject based on their individual conditioning histories for a range of learning rates α (ranging from 0.01–0.5). For each type of trial we averaged log-adjusted trial-by-trial response times across subjects and fit these to a regression model which included the averaged learning signal curves. In order to account for general changes in reaction time that would occur over the experiment we included an additional regressor as a covariate of no interest that reflected the change in reaction times across the experiment in the neutral trials (specifically a spline-smoothed fit of the averaged reaction times from the unpaired trials). This method allowed us to determine the learning rates that gave the best fit to subjects' behavior (on average

across subjects). We used the learning rates resulting from this procedure to model the fMRI data for all subjects, by regressing these signals against the brain imaging data as described below (the specific learning rates are given in the Results section).

fMRI scanning procedure

fMRI data were acquired on a Siemens AG (Erlangen, Germany) 3T TRIO MRI scanner; Blood Oxygenation Level Dependent (BOLD) contrast was measured with gradient echo T2* weighted echo-planar images (EPI). Imaging parameters were optimized to minimize signal dropout in medial ventral prefrontal and anterior ventral striatum: we used a tilted acquisition sequence at 30° to the AC-PC line [132], and an 8 channel phased array coil which yields a ~40% signal increase in this area over a standard coil. The first 5 volumes of 620 were discarded to permit T1 equilibration. Other parameters were as follows: in-plane resolution, 3 x 3 mm; slice thickness, 3 mm; repetition time, 2s; echo time, 30 ms; field of view, 192 x 192 mm. A T1 weighted structural image was also acquired for each subject.

Imaging data analysis

fMRI data were preprocessed in SPM2 (`http://www.fil.ion.ucl.ac.uk/spm/software/spm2/`). Images were corrected for slice acquisition time within each volume, motion corrected by aligning to the first volume [133], and unwarped to correct for estimated movement-related deformations in the EPI field [134]. They were normalized to a standard EPI template in Montreal Neurological Institute space, and spatial smoothing was applied using a Gaussian kernel with full width at half maximum of 8 mm. The normalization parameters estimated for each subject were also applied to their T1-weighted structural scans.

Statistical analysis was carried out using the general linear model, with the canonical hemodynamic response function (HRF) as a basis set. We describe results from two main analyses, designed to examine stimulus-driven effects and learning-related effects, respectively. In the first analysis fractal and face presentation events were

modeled as delta functions. In the second analysis, prediction error signals were entered as parametric regressors for each trial-type, at the time a face would be presented, independently of whether a face had actually been shown. For all models, the six ongoing motion parameters estimated during realignment were included as regressors of no interest. The results from each subject were taken to the random effects level by applying t-tests between contrast images to produce group statistical parametric maps. We focused our analyses on brain regions of interest, specifically the striatum, orbitofrontal cortex, and amygdala.

Results: Behavioral Measures

Face attractiveness ratings

Consistent with previous studies using the same set of faces [31], subjects rated the faces in the attractive category as significantly more attractive than those in the unattractive category, for both female (Wilcoxon signed ranks test $|Z| = 4.264$, $N = 28$, $p<.001$) and male faces ($|Z| = 4.623$, $N = 28$, $p<.001$). Gender differences were observed in evaluations of male faces, as female subjects rated them as significantly more attractive than did male subjects (Mann-Whitney $|Z| = 2.374$, $N = 15, 13$, $p<.05$ and $|Z| = 2.097$, $N = 15, 13$, $p<.05$, attractive and unattractive respectively). There were no significant gender differences in evaluations of female faces. Attractiveness evaluations are shown in Figure 2.2a.

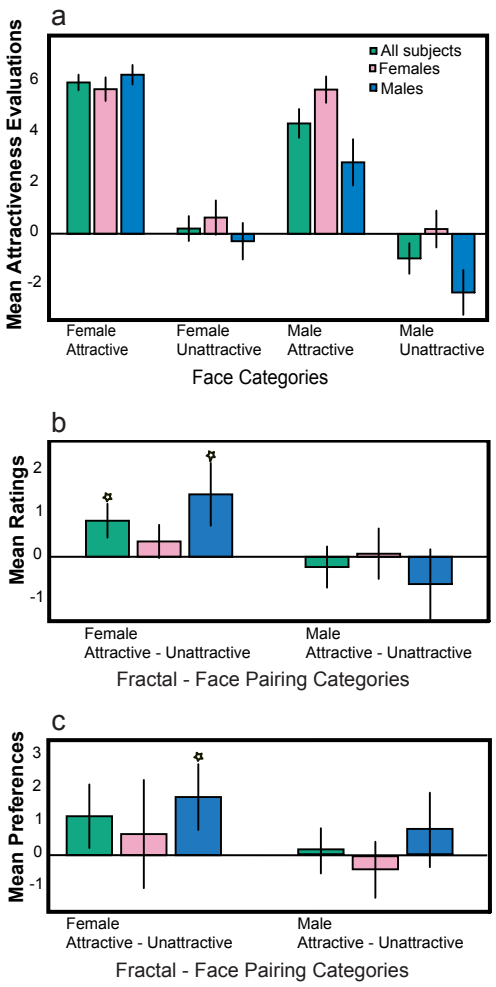

Figure 2.2: a) Evaluations of attractiveness of faces in each category (attractive and unattractive, male and female), averaged across the 8 faces in each category. For both face genders, the unattractive mean was subtracted from the attractive mean and the differences averaged across subjects in three groups: all subjects, males and females. Bars indicate standard error. b) Difference in pleasantness ratings for fractals pre- and post- conditioning, unattractive difference subtracted from attractive difference, and this difference averaged across subjects in three groups: all subjects, males and females. Bars indicate standard error. Stars indicate differences that are significant (Wilcoxon signed ranks test $p<0.05$). c) Difference in number of times fractal was chosen as preferred pre- and post-conditioning, unattractive difference subtracted from attractive difference, and this difference averaged across subjects in three groups: all subjects, males and females. Bars indicate standard error. Star indicates significant difference (Wilcoxon signed ranks test $p<0.05$).

Sexual orientation questionnaire

Based on self-reports of sexual orientation, our subject group consisted of 14 heterosexual females and 1 bisexual female, 11 heterosexual males and 2 bisexual males. Labels of sexual orientation were corroborated by ratings of attraction and sexual interest to the opposite sex: female heterosexual subjects rated their attraction to males to be on average 8±0.6, whereas male heterosexual subjects rated their attraction to females as 9±0.28. Similar scores were obtained on ratings of sexual interest in the opposite sex: 7.47±0.41 in female subjects for males, 8.5±1.5 in male subjects for females. The bisexual subjects rated their level of attraction and sexual interest for the opposite sex within the same range as the heterosexual subjects, and were therefore included in all analyses described here, unless explicitly stated otherwise.

Changes in pleasantness ratings of stimuli as a function of conditioning

Significant differences in pleasantness ratings for fractal stimuli were found from before to after conditioning for the stimuli paired with attractive female faces (Wilcoxon signed ranks test, $|Z| = 2.169$, $N = 28$, $p<0.05$) across all subjects (both male and female). This effect was also significant in the sub-group of male subjects (Wilcoxon signed ranks test, $|Z| = 1.992$, $N = 13$, $p<0.05$), but not female subjects. We did not find a similar effect for the fractals paired with male faces, in any of the subject groups. Male and female subjects showed no significant differences in pleasantness ratings. Figure 2.2b shows differences in pleasantness ratings from before to after conditioning for stimuli paired with attractive and unattractive faces, plotted for all subjects and males and females separately.

Changes in behavioral preference for fractal stimuli as a function of conditioning

In male subjects, the increase in preference for fractals paired with highly attractive female faces was significantly greater than for those paired with unattractive female faces (Wilcoxon signed ranks test, $|Z| = 2.428$, $N = 13$, $p<0.05$), although this effect was not significant across all subjects. On the other hand, we did not find a similar effect in female subjects for those fractals paired with either male or female faces. No significant effects were found in either male or female subjects for stimuli paired with same-sex attractive faces, and no significant gender differences in preference ratings were observed. Figure 2.2c shows differences in preference rankings for the fractal stimuli as a function of conditioning, plotted separately for stimuli paired with male, and female faces, and groups of all, male and female subjects.

Correlations between reaction times and learning model

Our regression analysis showed that the Rescorla-Wagner model with the best fitting learning rate was significantly correlated with changes in subjects' reaction time data over the experiment for all four trial types in which subjects learned the predictive value of the fractal cues. The learning rates obtained for each trial type were [Attractive female: 0.026 ($R^2 = 0.59$, $p <0.05$) , Attractive male: 0.04 ($R^2 = 0.43$, $p <0.05$), Unattractive female: 0.038 ($R^2 = 0.48$, $p <0.05$), Unattractive male: 0.04 ($R^2 = 0.57$, $p <0.05$)]. Subject averaged reaction times are shown separately for low (0.0-0.2) and high (0.2-0.5) value trials for each condition in Figure 2.3. This figure shows that for all four face-paired conditions, in both genders, there is a slowing in reaction times as model-predicted reward value increases. The mean reaction times for each condition are (mean±se, in ms): 469.19±88.66 (attractive female), 462.42±87.38 (attractive male), 465.16±87.90 (unattractive female), 460.33±86.99 (unattractive male).

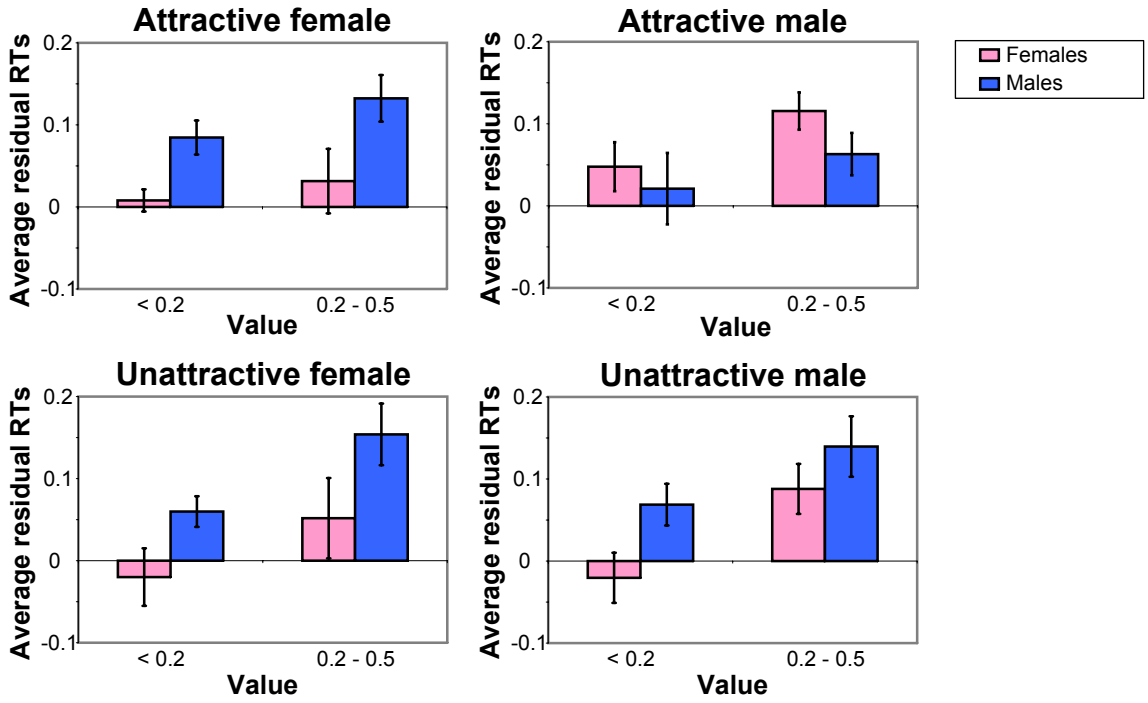

Figure 2.3: Relationship between reaction times and predicted value from the Rescorla-Wagner learning model. Individual subjects reaction times (RTs) were corrected for drift by taking the residuals from a regression onto the averaged reaction times for the neutral (never paired) conditions. Corrected RTs were then binned according to the predicted value derived from the Rescorla-Wagner learning model, using the derived learning rates for each trial type. The RTs were binned into low (0-0.2) and high (0.2-0.5) value trials. The plot shows that trials high in value show increased RTs compared to trials low in value.

Results: fMRI

Prediction error contrasts

We performed a linear contrast between prediction error signals for attractive and unattractive faces, and found significant activation in one of our a priori regions of interest: nucleus accumbens (NAcc) as shown in Figure 2.4a ([-9 15 -3], z = 3.38 and [9 15 -9], z = 3.12; significant at $p<0.001$, uncorrected). These areas survive small volume correction using a sphere of 8mm radius defined around co-ordinates derived from a previous demonstration of reward prediction error activity in the NAcc ([-11 11 -2] and [11 11 -2] [135]). The peak in the left NAcc is also significant in the contrast of learning with opposite-sex attractive compared to opposite-sex unattractive faces in all subjects ([-9 15 -6], z = 3.52; $p<0.001$ uncorrected), and the subset of heterosexual subjects ([-9 15 -6], z = 3.79; $p<0.001$ uncorrected). The contrast between learning with same-sex attractive compared to unattractive faces did not show any significant activations. Activations for prediction error contrasts are shown in Table 2.1.

The prediction error contrast for learning with opposite sex attractive compared to unattractive faces also showed activity in some of our other a priori regions of interest, namely bilateral medial orbitofrontal cortex ([-6, 33, -9], z = 3.63; $p<0.001$ uncorrected) and ([9, 33, -12], z = 3.22; $p<0.001$ uncorrected), and caudate nucleus ([9, 15, 6], z = 3.37; $p<0.001$ uncorrected).

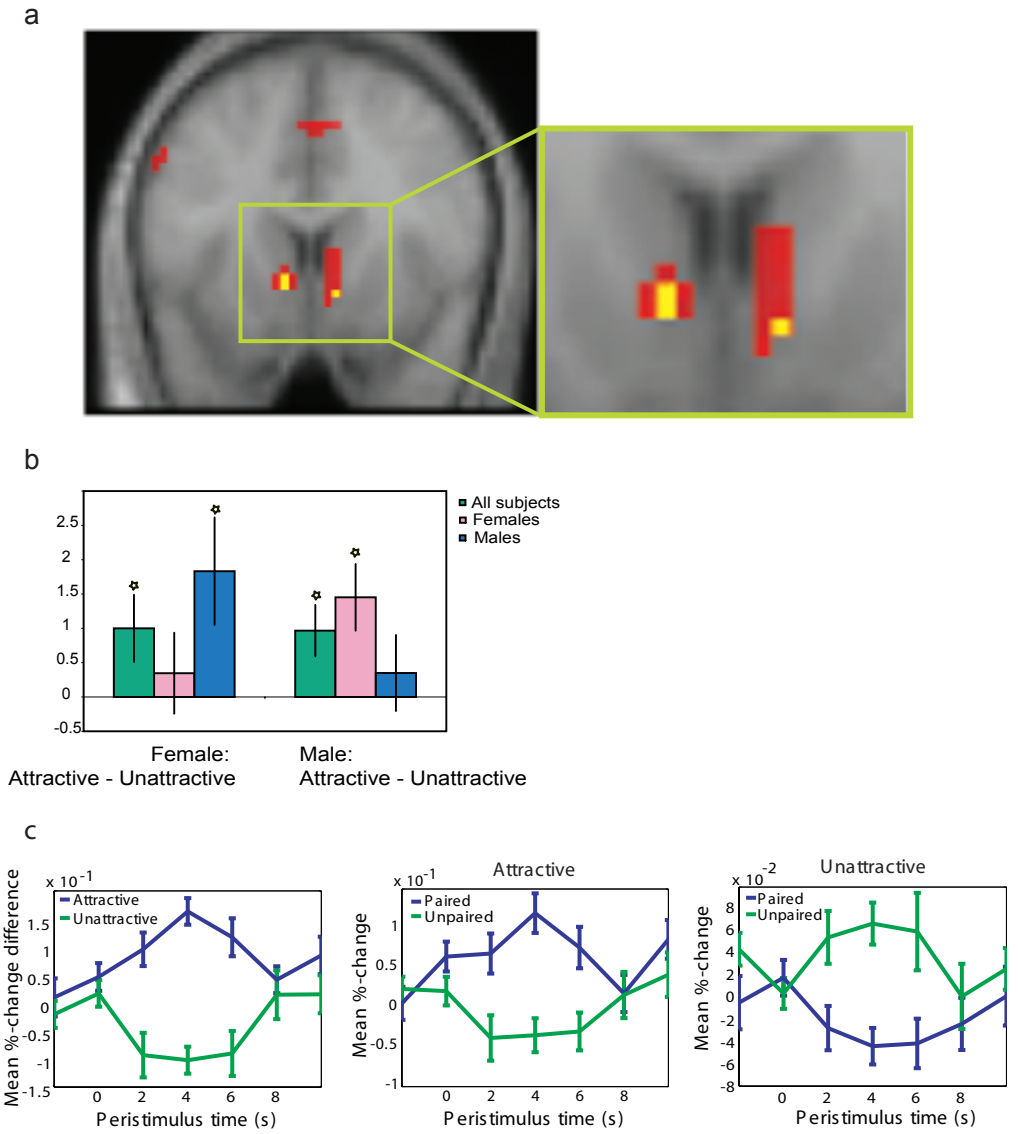

Figure 2.4: a) Voxels in the nucleus accumbens were significantly activated in a contrast of prediction error signals for attractive faces vs. unattractive faces, voxels in yellow are significant at $p<0.001$, voxels in red are significant at $p<0.01$. b) Parameter estimates for prediction error at the peak NAcc voxel from the attractive-unattractive contrast [-9 15 -3], averaged across subjects in three groups: all subjects, males, and females. Bars indicate standard error in the mean. Stars indicate differences that are significant (one-tailed t-test, $p<0.05$). c) Subject averaged time courses, aligned to the beginning of a trial, i.e., onset of the fractal cue; faces were presented at 1 second. Bars indicate standard errors. Time courses extracted from each subjects peak voxel in the left NAcc region. The leftmost plot shows the averaged over attractive and unattractive trials, unpaired trials subtracted from paired. The middle and rightmost plots show paired and unpaired trials separately for attractive and unattractive faces, respectively.

Table 2.1: Prediction error contrasts

Z scores and MNI coordinates of peak activation foci. Minimum cluster 5 contiguous voxels, thresholded at p<0.001 uncorrected.

	Prediction error contrast			
	Attractive-Unattractive		Opposite sex Attractive - Unattractive	
Region	# voxels	Z	#Voxels	Z
Right inferior frontal gyrus	23	3.74(39, 30, 18)	42	3.79(36, 30, 18)
Left inferior frontal gyrus	32	3.7(-36, 33, 15)	12	3.38(-39, 30, 18)
Left nucleus accumbens	5	3.38(-9, 15, -3)	7	3.52(-9, 15, -6)
Left medial OFC			38	3.63(-6, 33, -9)
Right medial OFC			5	3.22(9, 33, -12)
Right caudate			5	3.37(9, 15, 6)

In the contrast of prediction error for learning with attractive compared to unattractive faces, we also found significant effects in the right and left inferior frontal gyrus (see Table 2.1). These areas remain significant when we restrict this contrast to opposite sex faces.

Prediction error responses to learning with same and opposite sex faces

We explored prediction error responses to fractals paired with opposite and same sex faces by conducting a post-hoc statistical analysis on the contrast estimates derived from the left ventral striatum (plotted in Figure 2.4b) in heterosexual subjects. In male subjects, we found a significant difference in responses for attractive compared to unattractive female faces ($|t| = 3.01$, dof $= 8$, $p<0.05$), but no difference for male faces, whereas in female subjects, we found a significant difference in contrast estimates for attractive compared to unattractive male faces ($|t| = 3.16$, dof $= 12$, $p<0.05$), but not female faces. Pooling male and female subjects we found a significant effect of attractiveness when subjects were presented with opposite ($|t| = 4.31$, dof $= 21$, $p<.001$) but not same sex faces.

Time-course plots

We also extracted time courses for peak voxels in the region of individual subjects' left NAcc, and averaged over attractive and unattractive trials, subtracting the averaged trials for which no face was presented from the averaged trials for which a face was presented. The resulting time courses are shown in Figure 2.4c (left panel) and indicate a positive increase in BOLD signal for positive prediction errors and a decrease in signal for negative prediction errors. Figure 2.4c (middle panel) shows that the NAcc responds positively to the presentation of an attractive face, and negatively to the omission of an attractive face, while Figure 2.4c (rightmost panel) shows that for unattractive faces this relationship is inverted, with increased activation seen to the omission of a face.

Test for learning related changes over time

In order to establish whether activity in NAcc is associated with a temporally evolving learning signal as opposed to non-learning related effects induced by the presence or absence of a face, we performed an additional analysis on the time-series data extracted from the peak voxel in NAcc (at [9 15 -9]). For this, we included in the analysis a prediction error regressor that temporally shifted from the time of face presentation to the time of cue presentation, using a real-time extension of the Rescorla Wagner learning rule: temporal difference learning [129]. We included in the same analysis a regressor at the time of face presentation, only when faces were actually presented. The temporal difference prediction error signal was a significantly better fit to activity in the NAcc than the face regressor at $p<0.05$, suggesting that activity in this structure reflects dynamic learning related changes and not merely effects relating to the presence of absence of a face.

Separate prediction errors during learning with attractive and unattractive faces

Comparing prediction error responses during learning with attractive and unattractive faces produced robust differences, due to the opposing direction of the prediction error signal. We also examined learning signals in response to both attractive and unattractive faces independently. A simple contrast for areas showing a positive correlation with the prediction error signal for attractive faces produced a peak in the right NAcc (at [6 15 -12]) which survived correction for small volume at $p<0.05$ FDR corrected in a 5 mm sphere centered on the peak identified above ([9 15 -9]). A simple contrast to detect areas showing a negative correlation with the prediction error signal for unattractive faces, also produced a peak in the left NAcc (at [-6 12 -3]) which survived correction for small volume at $p<0.05$ FDR corrected in a 5 mm sphere centered on the NAcc peak identified above ([-9 15 -3]).

We found evidence for a positive correlation with prediction error signals during learning with both attractive and unattractive faces in the amygdala, another of our a priori regions of interest: for attractive faces in right amygdala ([24 0 -25], $z = 3.31$ $p<0.001$ uncorrected) and for unattractive faces in right amygdala ([18 -6 -21], $z = 4.02$ $p<0.001$ uncorrected) and left amygdala ([-18 -6 -18], $z = 3.26$ $p<0.001$ uncorrected).

Main effect of attractiveness

We also tested for regions responding to receipt of the attractive faces themselves. For this we performed a linear contrast of attractive–unattractive faces at the random effects level (Figure 2.5a) and found significant effects in medial OFC ([12 39 -9] $z = 2.93$) extending into medial prefrontal cortex, a region previously shown to be responsive to the receipt of attractive faces [30, 31]. The OFC area survived correction for small volume at $p<0.05$ FDR corrected in an 8 mm sphere centered around coordinates from a previous study of facial attractiveness (at [16 45 -11] from [30].

A number of other regions show responses to facial attractiveness (clusters larger

Table 2.2: Main effect of attractiveness

Z scores and MNI coordinates of peak activation foci. Minimum cluster 5 contiguous voxels, thresholded at p<0.001 uncorrected

	Prediction error contrast	
	Attractive-Unattractive	
Region	# voxels	Z
Right inferior frontal gyrus	7	3.70(39, 24, 18)
Left inferior frontal gyrus	57	4.7(-39, 36, 15)
Left nucleus accumbens	8	4.28(-9, 15, -6)
Medial anterior cingulate	32	3.76(0, 36, 12)
Medial posterior cingulate	49	4.51(-3, -30, 30)

than 5 voxels significant at p<0.001 uncorrected are tabulated in Table 2.2), including the left NAcc at [-9 15 -6] (z = 4.28; p<0.001 uncorrected). The NAcc activity is in the same region we found to be responsive to prediction error. A post-hoc inspection of the time course plots from these two regions shows that the NAcc demonstrates a response profile consistent with a reward prediction error and not face presentation per se, as this region not only increases following presentation of an attractive face, but also increases following the omission of an unattractive face (Figure 2.4c). On the other hand, the OFC area only showed increased activity to the presentation of an attractive face, and showed no change in activity to any other condition, suggesting that this area is responding to the receipt of an attractive face and not a prediction error (Figure 2.5b).

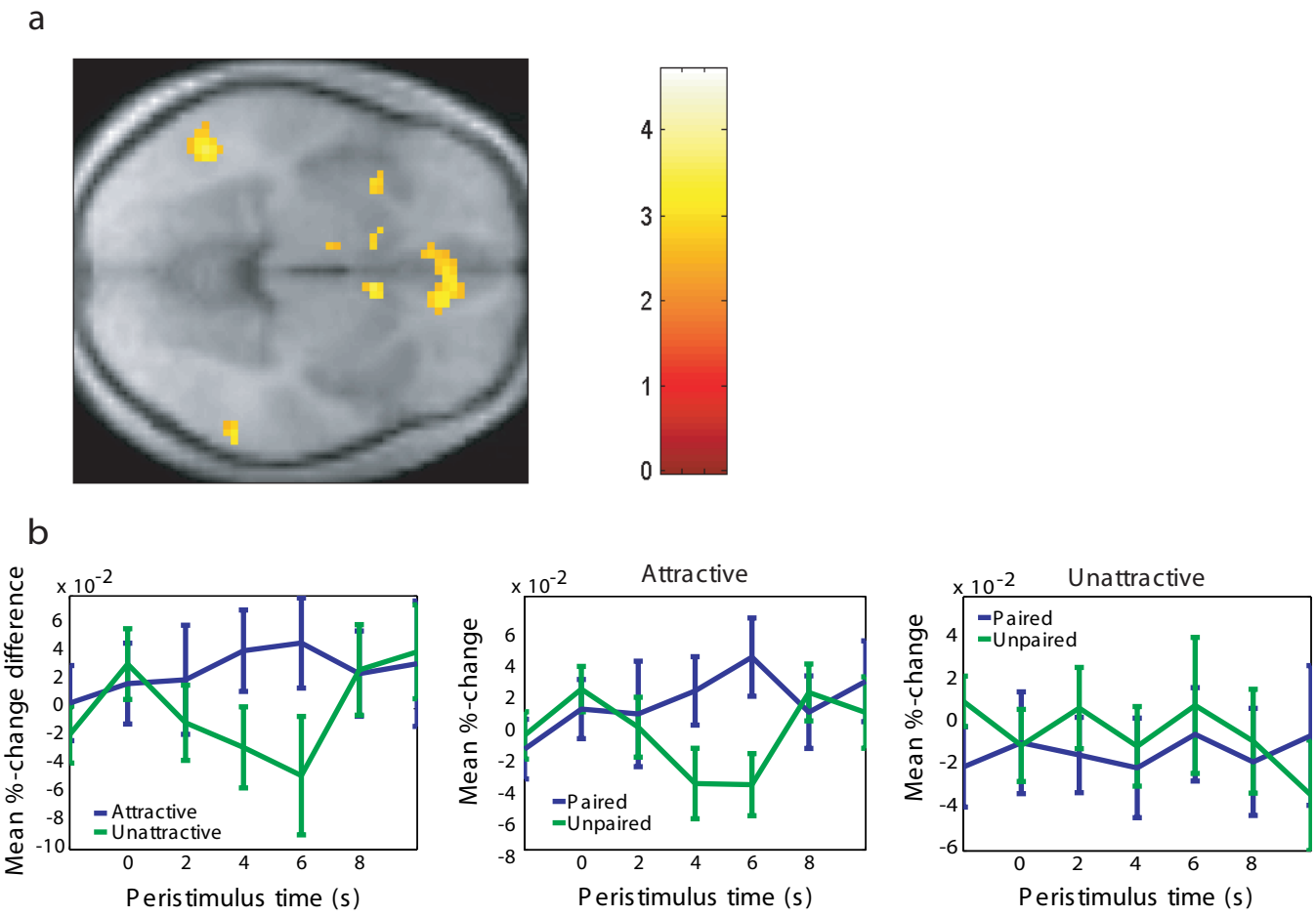

Figure 2.5: a) Voxels in the orbitofrontal cortex extending into medial prefrontal cortex were significantly activated in a contrast of attractive faces vs. unattractive faces. The peak in medial OFC ([12 39 -9] z = 2.93), survived correction for small volume at $p<0.05$ FDR corrected in an 8 mm sphere centered around co-ordinates from a previous study (see Results). For visualization, the threshold is set at $p<0.01$ uncorrected. b) Subject averaged time courses, aligned to the beginning of a trial, i.e. onset of the fractal cue; faces were presented at 1 second. Bars indicate standard errors. Time courses extracted from the medial OFC peak in response to the main effect of attractiveness. The leftmost plot shows the averaged over attractive and unattractive trials, unpaired trials subtracted from paired. The middle and rightmost plots show paired and unpaired trials separately for attractive and unattractive faces, respectively.

Discussion

The impact an attractive face can have on human behavior, from product choice [102] to hiring preference [115], has been well documented. However, to our knowledge, this study marks the first demonstration of modulation of behavioral preference for a neutral visual stimulus by conditioned association with an attractive face. Subjects in our study rated neutral fractal images as significantly more pleasant after they had been repeatedly paired with attractive female faces. Our finding of a modulation of behavioral preference to previously neutral stimuli as a function of conditioning with attractive faces resonates with other studies that have found similar effects through acquisition of conditioned associations with other types of reinforcers such as food and money [136, 51, 128].

By measuring neural activity with fMRI while subjects acquired this association, we were able to observe learning-related activity in the brain as the association was formed. We found that reward prediction errors were engaged in the ventral striatum, differentially for stimuli paired with attractive compared to unattractive faces. Prediction errors have been observed during learning with other types of reward, such as juice and money [74, 73]. The observation that attractive faces also engage these signals further reinforces the proposal that attractive faces can be considered to be a form of visual reward [30, 31]. The present result also provides insight into the mechanism by which attractive faces transfer their rewarding properties to other stimuli.

It is notable that increases in activity occurred in the striatum in response to positive prediction error signals following the unexpected presentation of an attractive face, but the opposite effect was found in response to the unexpected presentation of an unattractive face, in which case a decrease in signal was observed. These findings suggest that ventral striatum shows a very different response profile to prediction error signals during learning with attractive as opposed to unattractive faces. These results are similar to effects found for prediction error signals generated during learning with monetary reward and punishment [137]. These results are also compatible

with response profiles reported in striatum in fMRI studies involving delivery of monetary reward and punishment, whereby ventral striatum has been shown to increase in response to receipt of monetary reward, and decrease in response to receipt of monetary loss [74]. By contrast, a different pattern of responses has been observed in ventral striatum in response to prediction errors produced during learning with other types of reinforcers such as somatosensory pain and even non-preferred flavors [128, 72]. In these cases, an increase in signal has been reported in ventral striatum following the unexpected delivery of a cue signaling subsequent pain or unpleasant flavor. Thus, ventral striatum appears to show very different neural responses as a function of learning with different types of reinforcers. This raises the question as to the nature of the difference between reinforcers that leads to such divergent response profiles. One possibility is that ventral striatum responds differently to learning with primary as opposed to secondary reinforcers. Money can be considered to be a secondary or learned reinforcer, whereas pain and food can be argued to be primary reinforcers [109]. However, facial attractiveness is often suggested to be a primary reinforcer, as judgments of facial attractiveness are suggested to be culturally invariant [138], and attractiveness has been argued to signal reproductive fitness [112]. As a consequence, the fact that attractive faces and money are similar in the way they activate the striatum would appear to argue against a primary vs. secondary reinforcer account of differential striatal function. An alternative possibility is that ventral striatum is involved not in learning about the sensory properties or abstract value of unconditioned stimuli, but instead learns associations between arbitrary stimuli and the unconditioned responses produced by an unconditioned stimulus. Differences in the nature of the unconditioned responses produced by different reinforcers could potentially account for differential activity in the striatum. Future studies will be needed to investigate this possibility further. Although we found an overall effect of attractiveness on prediction error activity in the ventral striatum, we also found that in this area the effect was significant when heterosexual subjects were presented with opposite sex faces, but not same sex faces. That is, prediction error responses were enhanced when learning about attractive faces of the opposite sex in both genders.

This suggests that ventral striatum may be involved in mediating learning about attributes linked to sexual preference, as opposed to learning about more general aspects of visual aesthetics [123].

In contrast to the nucleus accumbens, the amygdala showed positive correlations with prediction error signals during learning with both attractive and unattractive faces, consistent with previous findings of a role for amygdala in conditioning involving both appetitive and aversive stimuli [52, 38, 139, 49]. More generally, these results add to an extensive prior literature implicating the amygdala in processing stimuli of both positive and negative valence [140, 141, 142].

The results of this study also have important implications for understanding the underlying mechanisms by which product advertising can influence behavioral preference in the marketplace. Marketers have long attempted to bias consumer preference by pairing a particular product with another stimulus that is already highly valued, such as an attractive face. Indeed, changes in product evaluations and preference have been observed in behavioral experiments as a function of such pairing procedures [143, 101]. However, the precise mechanism by which preference modulation takes place has remained an open question. One possibility is that changes in preference evaluations occur through cognitive appraisal or top down modulation of affect (as in cognitive appraisal cf Folkman [144]). Another possibility is that preference evaluations occur as a function of classical conditioning [145]. We directly tested this hypothesis using classical conditioning. Our results provide evidence that the change in preference likely occurs as a function of classical conditioning, by showing that similar neural mechanisms are engaged during evaluative preference modulation as are engaged during other types of classical conditioning. Moreover, the fact that evaluative preference modulation specifically engages prediction error signals in the ventral striatum, suggests that this procedure may recruit dopamine neurons in the midbrain, as is known to be the case during learning with other kinds of reward in non-human primates [64]. Consistent with the above suggestion, a recent fMRI study has shown that prediction error signals expressed in the ventral striatum during reward-learning can be modulated through pharmacological manipulation of dopamine levels in hu-

mans [146], indicating that the source of such signals in human fMRI studies may in part be attributable to the afferent input from dopamine neurons.

To conclude, in the present study we have found significant prediction error-related activity in the ventral striatum during conditioning with attractive compared to unattractive faces. These findings suggest that an arbitrary stimulus can acquire conditioned value by being paired with pleasant visual stimuli just as with other types of reward, like money or juice. Such a learning process elicits a reward prediction error signal in a main target structure of dopamine neurons: the ventral striatum. The learning process we describe here may provide insights into the neural mechanisms used in advertising to influence behavioral preferences, whereby consumers are exposed repeatedly to simple associations between products and rewarding visual stimuli such as pretty faces.

Chapter 3

Pavlovian cues influence decision-making*

In outcome-specific transfer, Pavlovian cues that are predictive of specific outcomes bias action choice towards actions associated with those outcomes. This transfer occurs despite no explicit training of the instrumental actions in the presence of Pavlovian cues. The neural substrates of this effect in humans are unknown. To address this we scanned 23 human subjects with fMRI while they made choices between different liquid food rewards in the presence of Pavlovian cues previously associated with one these outcomes. We found behavioral evidence of outcome-specific transfer effects in our subjects, as well as differential BOLD activity in a region of ventrolateral putamen when subjects chose, respectively, actions consistent and inconsistent with the Pavlovian-predicted outcome. Our results suggest that choosing an action incompatible with a Pavlovian-predicted outcome might require the inhibition of feasible but non-selected action-outcome associations. The results of this study are relevant for understanding how marketing actions can affect consumer choice behavior as well as for how environmental cues can influence drug seeking behavior in addiction.

*Adapted with permission from: Bray S, Rangel A, Shimojo S, Balleine B, O'Doherty JP (2008) The neural mechanisms underlying the influence of pavlovian cues on human decision making. J Neurosci 28:5861-5866. Copyright 2008 *Journal of Neuroscience*

Introduction

It is well known that Pavlovian cues associated with rewarding outcomes can exert biasing effects on action selection [147, 82]. A form of this effect relevant for decision-making is outcome-specific transfer [79, 83, 84, 85, 148]. In outcome specific transfer, an animal's choice between multiple simultaneously available instrumental responses leading to different outcomes can be biased by the presentation of a Pavlovian cue previously associated with one of those outcomes, such that the animal will tend to favor the instrumental action corresponding to the particular outcome with which that cue has been associated. Outcome-specific transfer effects are evident, for example, in the impact that in-store advertisements and other marketing strategies have on consumer behavior [149], as well as in addictive behavior [150].

Lesion studies in rodents indicate that the following structures are necessary for outcome-specific transfer to occur: the striatum, including the nucleus accumbens shell [83] and the dorsolateral striatum [85], and structures afferent to these regions including the medio-lateral orbitofrontal cortex [151] and basolateral amygdala [84].

Outcome-specific transfer can be differentiated from another form of Pavlovian-instrumental interaction called general-transfer in which a Pavlovian cue exerts a non-specific energizing effect on instrumental behavior by increasing the vigor of instrumental responses [148, 84]. General transfer seems to depend on circuitry involving the ventral striatum and amygdala that is clearly dissociable from that involved in the outcome-specific transfer effect: lesions of the nucleus accumbens core and amygdala central nucleus affect general transfer but leave specific transfer intact, whereas lesions in the nucleus accumbens shell and basolateral amygdala have the converse effect [83, 84]. In humans, a recent fMRI study has implicated human nucleus accumbens and amygdala in general transfer [86], but the brain systems underlying outcome-specific transfer in the human or primate brain more generally have yet to be identified. Furthermore, whereas rodent lesion studies have identified regions that appear to be necessary for specific transfer [83, 84, 85], the precise functional contribution of each of these regions to this process has yet to be characterized.

The aim of the present study was twofold: Firstly to determine the neural substrates of the outcome-specific transfer effect in the human brain, and, secondly, to gain insight into the neural computations within these regions that might underlie this function. To address these aims we used event-related fMRI to measure BOLD responses in human subjects while they made instrumental choices in the presence of Pavlovian cues that were either associated with the liquid food reward outcomes generated by some of the actions, or associated with an affectively neutral (control) outcome. On the basis of the animal studies, we focused our analysis on the striatum, particularly its ventral aspect, including the nucleus accumbens and adjacent ventral putamen.We also tested for specific-transfer effects in the amygdala.

Methods

Subjects

Twenty-three healthy, right-handed subjects participated in this study (6 females), ranging in age from 18 to 40 (mean 24 $\pm$ 5.3 S.D). One additional subject did not complete the study and was not included in the analysis. All subjects gave informed consent and the study was approved by the Caltech Institutional Review Board.

Stimuli

Visual stimuli were presented via a projector positioned at the back of the room. Subjects viewed a reflection of the projected image (800 x 600 pixels) in a mirror attached to the scanner head coil. The food rewards were delivered by means of four separate electronic syringe pumps (one for each liquid) positioned in the scanner control room. For each rewarded trial, these pumps pushed 0.6 ml of liquid to the subject's mouth via ~10 m long polyethylene plastic tubes, the other end of which were held between the subject's lips like a straw while they lay supine in the scanner. Stimulus presentation and response recording were controlled with the Cogent 2000 Matlab (Mathworks, Natick, MA) toolbox.

Behavioral procedures

During both the Pavlovian and instrumental training subjects were explicitly asked to learn the cue-outcome and action-outcome relationships. All 4 training and test sessions described below were performed in the scanner.

Pavlovian training

Pavlovian training consisted of the presentation of associations between simple geometrical visual stimuli (see Figure 3.1a for an example) and one of four liquid outcomes, three of which were rewarding: chocolate milk (Hershey's, dist by Dean National Brand Group, Dallas, TX), cola (Coca-Cola, Atlanta, GA), and orange juice (Trader Joe's, Monrovia, CA) and an affectively neutral tasteless control solution, which consisted of the main ionic components of human saliva (25mM KCl and 2.5mM NaHCO3) (Figure 3.1a). Cues were presented at the center of the screen for 1.75 s, then 3 s after cue offset rewards were delivered with a probability of 50%. The intertrial interval varied uniformly between 1 and 5 s.

Instrumental training

During instrumental training trials subjects were asked to choose between two button-press actions. Four grey squares at the bottom of the screen corresponded to the four buttons on a response box (Current Designs, Philadelphia, PA) that the subjects held in their right hand. Specific actions were made available for selection when the corresponding squares changed color from grey to brown, two at a time. As in the Pavlovian trials, the response cues appeared for 1.75 s. Subjects were asked to make a choice during this time. The choice was followed by a 3 s delay before the outcome associated with the chosen action was delivered on 50% of trials (Figure 3.1b). The intertrial interval varied uniformly between 1 s and 5 s. Responses on each button earned distinct outcomes: two of the buttons led to rewarding outcomes (for example, orange juice and chocolate milk) and two led to the neutral outcome. Therefore, during Pavlovian training subjects experienced 4 different outcomes, while

in the instrumental trials they experienced only 3.

Training schedule

The first training session consisted entirely of Pavlovian trials, 10 of each type for a total of 40 trials and a duration of approximately 6 minutes (see Table 3.1). The second training session consisted entirely of instrumental trials, 6 of each type for a total of 36 trials and a duration of approximately 5 minutes. In the first two sessions, Pavlovian and instrumental trials were presented separately to enhance learning of the respective associations. In the third session, Pavlovian and instrumental trials were randomly intermixed, 60 (15 x 4) Pavlovian trials and 60 (10 x 6) instrumental trials, for a duration of approximately 18 minutes. Before training and after each session, subjects rated the pleasantness of the stimuli as described below.

Outcome-specific transfer

Following the three training sessions subjects performed a series of transfer trials. During transfer trials one of the Pavlovian cues was presented simultaneously with the instrumental cues (Figure 3.1c), and as in instrumental training, subjects were asked to choose between two available options. This phase was conducted in extinction, meaning that no outcomes were delivered. The reason for performing this phase in extinction was to allow assessment of the influence of the Pavlovian cues on instrumental responding without the confounding effects of the outcomes themselves. Testing for outcome-specific effects in extinction is standard in animal learning studies of this phenomenon [79, 80, 83].

There were five different types of trials. Two of the trial types were designed to test for outcome-specific transfer effects. On these trials subjects chose between actions associated with two particular reward outcomes: O1 and O2 (for example, orange juice and chocolate milk), while the concurrently presented Pavlovian cue was associated with one of these specific outcomes. One of the specific trial types involved the Pavlovian cue paired with outcome O1, and the other specific trial type involved the Pavlovian cue paired with O2. Evidence for an outcome-specific transfer effect

would be seen if the presence of the Pavlovian cue biased choice towards the action associated with the same outcome as the Pavlovian cue. In the subsequent analysis we pooled over both of the specific trial types, but differentiated between trials in which subjects made choices compatible with the Pavlovian outcome from those trials in which subjects made choices that were not compatible.

In a third 'Pavlovian reward control' trial type, subjects were again presented with the choice between two reward outcomes (O1 and O2), but instead the Pavlovian cue was previously associated with a different outcome (for example, cola), that was not compatible with either response option.

In the fourth 'Pavlovian neutral control' trial type, subjects were again presented with the choice between two reward outcomes (O1 and O2), but the Pavlovian cue presented this time was that associated with the affectively neutral outcome.

In the final 'neutral choice control' trial type subjects made choices between actions associated with the affectively neutral outcome, in the presence of a Pavlovian cue also associated with a neutral outcome. This last trial type was intended to be a baseline condition for choosing between two options in the presence of a visual cue but in the absence of predicted rewards. Each type of trial was presented 25 times, for a total of 125 trials and a duration of approximately 20 minutes.

Behavioral measures

Reaction times

Reaction times to choices were recorded both during the learning trials and the transfer test trials; these can be used as an online measure of learning [128].

Pupillary dilation

Pupil diameter was continuously measured during scanning using an Applied Science Laboratories (Bedford, MA, USA) MRI compatible eyetracking system. Pupil reflex

Table 3.1: Trial composition for training and transfer sessions

S1-4: visual cues. R1-4: 4 button response actions, O1-3: liquid rewards, ON: affectively neutral tasteless control solution.

		Cues			
Phase	Presentations	Pavlovian	Instrumental		Outcome	
1	10	S1			O1	
		S2			O2	
		S3			O3	
		S4			ON	
2	6		R1	R2	O1	O2
			R1	R3	O1	ON
			R1	R4	O1	ON
			R2	R3	O2	ON
			R2	R4	O2	ON
			R3	R4	ON	ON
3	10	S1			O1	
		S2			O2	
		S3			O3	
		S4			ON	
	15		R1	R2	O1	O2
			R1	R3	O1	ON
			R1	R4	O1	ON
			R2	R3	O2	ON
			R2	R4	O2	ON
			R3	R4	ON	ON
4		S1	R1	R2		
		S2	R1	R2		
		S3	R1	R2		
		S4	R1	R2		
		S4	R3	R4		

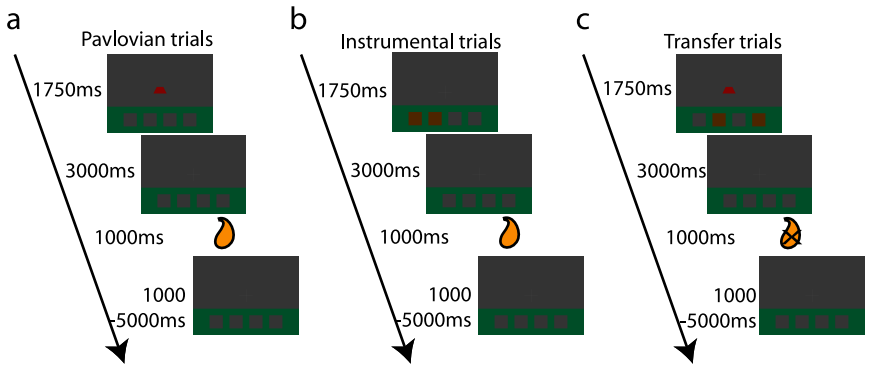

Figure 3.1: Illustration of trial types a) Pavlovian trial: A visual shape stimulus was presented at the center of the screen for 1.75 s followed by a fixation cross for 3 s. The liquid outcome corresponding to the stimulus was then delivered with a probability of 50 percent. 1 s was allotted for consumption and the interval between trials varied uniformly between 1 s and 5 s. b) Instrumental trial: Two of the four squares at the bottom of the screen changed color from grey to brown for 1.75 s during which time subjects were instructed to push one of the buttons. The liquid outcome corresponding to their response was delivered after 3s, with a probability of 50 percent. 1 s was allotted for consumption and the interval between trials varied uniformly between 1 s and 5 s. c) Transfer trial: A visual shape stimulus was presented simultaneously with two squares changing color. Subjects were instructed to press one of the corresponding buttons. Timing was similar to the Pavlovian and instrumental trials; however no outcomes were delivered during these trials.

amplitude has been shown to be modulated by arousal level and can thus be used as an index of conditioning [152].

Affective evaluations of stimuli

Before the start of the training procedure, and after each scanning session, we asked subjects to rate the pleasantness of the shape images and the liquid outcomes. Within each category, stimuli were presented in random order and subjects reported their evaluation by moving a cursor along a scale from -5 to +5.

Swallowing motion

A motion sensitive inductive coil was positioned on top of the subjects' throat using a Velcro strap around the neck. This measured the motion of the subjects' throat during swallowing. The time course derived from this measure was used as a regressor

of no interest in the fMRI data analysis. We do not have recordings for one subject who found the coil uncomfortable.

fMRI scanning procedure

fMRI data were acquired on a Siemens AG (Erlangen, Germany) 3T TRIO MRI scanner; Blood Oxygenation Level Dependent (BOLD) contrast was measured with gradient echo T2* weighted echo-planar images (EPI). Imaging parameters were optimized to minimize signal dropout in medial ventral prefrontal and anterior ventral striatum: we used a tilted acquisition sequence at 30° to the AC-PC line [132], and an 8 channel phased array coil which yields a ~40% signal increase in this area over a standard coil. The first 5 volumes of each session were discarded to permit T1 equilibration. Other parameters were as follows: 36 slices, in-plane resolution, 3 x 3 mm; slice thickness, 3 mm; repetition time, 2.25 s; echo time, 30 ms; field of view, 192 x 192 mm. A T1 weighted structural image was also acquired for each subject, as well as a 49 slice EPI to improve coregistration.

Imaging data processing and analysis

Data were pre-processed using the SPM5 software package (`http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`). Images were corrected for slice timing and spatially realigned to the first image from the first session. One of the 49 slice EPI images collected at the end of the experiment was used to improve coregistration and spatial normalization. The 36 slice EPI images were coregistered to a 49 slice EPI, which was in turn coregistered to the T1-weighted anatomical scan. The T1 image was segmented into white and grey matter, and the grey matter was coregistered and normalized to the template grey matter image distributed with SPM5 (in Montreal Neurological Institute space). These parameters were subsequently applied to the T1 image itself as well as the set of 36 slice EPI images. Spatial smoothing was then applied to the 36 slice EPI images using a Gaussian kernel with full width at half maximum of 8 mm.

Statistical analysis was carried out using a general linear model (GLM). The transfer session was modeled separately from the three training sessions, and here we report results only from the transfer phase of the experiment. The GLM included regressors at the time of cue onset, for 5 conditions: specific transfer when the option compatible with the Pavlovian cue was chosen, specific transfer when the incompatible option was chosen, 'Pavlovian reward control', 'Pavlovian neutral control', and 'neutral choice control'. We also included regressors at the time of expected outcome. Each regressor was modeled as an impulse function (0s), and convolved with the canonical hemodynamic response function. Regressors of no interest included missed trials when no option was chosen, the six ongoing motion parameters estimated during realignment, and motion due to swallowing. The results from each subject were taken to the random effects level by applying t-tests between contrast images to produce group statistical parametric maps.

Results

Behavioral results

Results of Pavlovian training

Behavioral results indicate that the Pavlovian stimulus-outcome associations were successfully learned. Following each training session, subjects were asked to rate on a scale from -5 to 5 how pleasant they found each shape stimulus and each liquid. After training, subjects rated the stimuli associated with rewarding outcomes as significantly more pleasant than the stimulus associated with the neutral outcome (paired t-test; $t(22) = -3.0840$, $p<0.01$)(Figure 3.2a). Pupil reflex amplitude also discriminated between reward and neutral conditions (Figure 3.2b). In the 16 subjects who showed reliable amplitude changes in pupil diameter following cue presentation the peak amplitude is significantly smaller for rewarded outcome trials which indicates a higher degree of arousal when subjects saw reward predictive cues (paired t-test; $t(15) = 2.4173$, $p <0.05$) [152, 72].

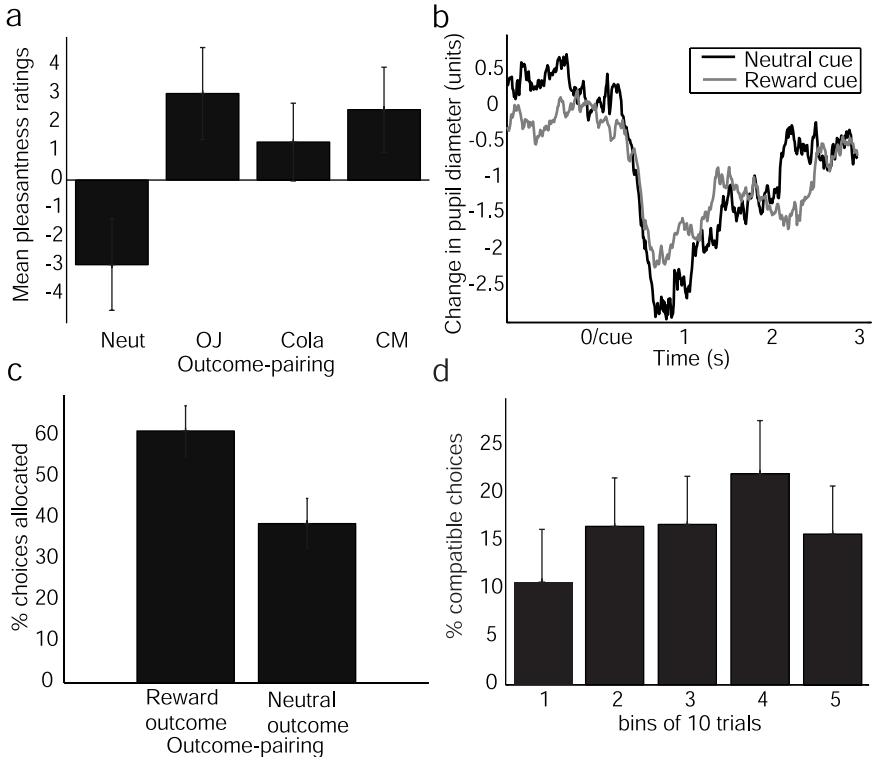

Figure 3.2: Behavior during training and test sessions a) Mean pleasantness ratings for visual cue stimuli following the training sessions, plotted by outcome pairing. The cues paired with the neutral outcome were rated as significantly less pleasant than the cues paired with reward outcomes (paired t-test; t(22) = -3.0840, $p<0.01$) b) Pupil diameter in response to visual cues. The peak amplitude is significantly smaller for the cues paired with reward outcomes, for the 16 subjects who showed reliable amplitude changes following cue presentation (paired t-test; t(15) = 2.4173, $p<0.05$) c) Choice behavior during the second session of instrumental trials, above cue invariant responding (50 percent). Plotted are responses during trials in which subjects chose between a reward outcome and the neutral outcome. Subjects were significantly more likely to choose the action leading to the reward outcome (1-sided paired t-test; t(22) = 1.8399, $p<0.05$) d) Choice data binned into 5, 10-trial bins. There is no significant linear trend across the session (linear regression of percent compatible choice allocation onto bin number, $p=0.239$).

Initial learning of instrumental associations

Subjects' choice behavior in the instrumental trials indicated that the instrumental associations were acquired. During the final training session subjects were significantly more likely to choose the action delivering a reward outcome when the alternative action delivered the neutral solution (Figure 3.2c) (1-sided paired t-test; $t(22)=1.8399$, $p<0.05$).

Outcome-selective transfer effects during test phase

We found evidence for an outcome-specific transfer effect in subjects' choice behavior during the transfer test phase. During the transfer phase, subjects chose the compatible option on average 66% of the time; this is significantly higher than cue invariant responding which averages to 50% over the two outcome-specific conditions (paired t-test; $t(22) = 3.6348$, $p<0.005$). There were a total of 50 specific transfer trials for each subject and separating these into 5, 10-trial bins, we found that there was neither a significant increase or decrease in choice allocation across time (Figure 3.2d), indicating that the biasing effect of the Pavlovian cues on choice persisted for the duration of the extinction test and did not attenuate.

fMRI results

In order to gain insight into the mechanisms underlying outcome-specific transfer in humans we performed two analyses. First, we compared brain activity during trials assessing outcome-specific transfer when subjects chose the option compatible with the Pavlovian cue to trials when they chose the incompatible option (one subject who never chose the incompatible cue was excluded from this analysis) (Figure 3.3a). We found significant activation in right ventrolateral putamen ($t(21) = 3.79$, $p<0.001$ uncorrected; $x = 27$, $y = -3$, $z = -3$) extending posteriorly towards the pallidum ($t(21) = 3.81$, $p<0.001$ uncorrected; $x = 24$, $y = -18$, $z = 0$). The left pallidum also showed a peak at this threshold ($t(21) = 3.82$, $p<0.001$ uncorrected; $x = -27$, $y = -15$, $z = -3$). These were the only regions to meet our significance criterion in this contrast.

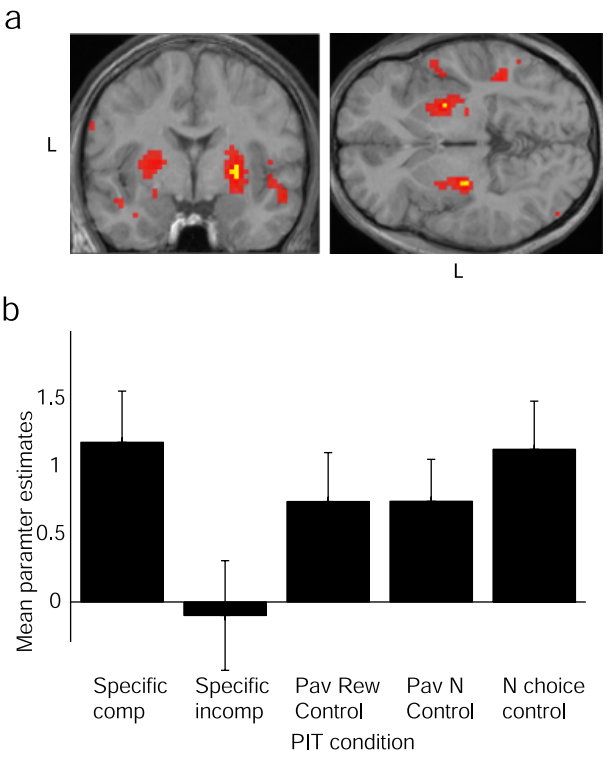

Figure 3.3: Imaging results from the Pavlovian-instrumental transfer phase a) fMRI results from the contrast comparing the outcome-specific transfer trials in which the action compatible with the Pavlovian cue is selected to those in which the incompatible action is selected (red = $p<0.01$, yellow = $p<0.001$). At a threshold of $p<0.001$ uncorrected we find significant activation in the ventrolateral putamen ($t(21) = 3.79$; $p<0.001$ uncorrected; x = 27, y = -3, z = -3), and bilateral pallidum ($t(21) = 3.81$; $p<0.001$ uncorrected; x = 24, y = -18, z = 0) and ($t(21) = 3.82$; $p<0.001$ uncorrected; x = -27, y = -15, z = -3) b) Parameter estimates from the peak putamen voxel for each subject, for each of the 5 experimental conditions during the transfer phase (specific compatible, specific incompatible, Pavlovian reward control, Pavlovian neutral control, neutral choice control). Parameter estimates in the specific compatible condition do not differ significantly from any condition other than specific incompatible (paired t-tests, $p>0.05$).

Second, we plotted the average parameter estimates taken from the general linear model estimates at the peak putamen voxel for each subject (Figure 3.3b). We found that the difference between conditions was due to a significant decrease in signal during the outcome-specific trials where the incompatible response was chosen, relative to the outcome-specific trials when the compatible response was chosen and to the other control conditions. In fact, activity in the compatible condition did not differ significantly from activity during any of the other control conditions (paired t-tests; $p>0.05$), and more generally, activity in the outcome-specific trials did not differ from the control conditions (paired t-tests; $p>0.05$).

Discussion

Our results provide insights into the neural mechanisms by which Pavlovian cues can modulate choice between different instrumental courses of action in humans. In outcome-specific transfer, subjects are more likely to choose an action that is associated with a particular outcome in the presence of a Pavlovian cue that was previously associated with the presence of that outcome. We found neural correlates of outcome-specific transfer in a very circumscribed region of extended ventral striatum in the ventral caudolateral putamen. This region and an adjacent region of ventral pallidum were the only areas to meet our statistical criterion for significance.

These findings add to an accumulating body of evidence from human fMRI studies of a role for an extended region of ventral parts of putamen alongside nucleus accumbens in functions related to reward-learning and prediction errors [37, 124, 71] and now in interactions between Pavlovian and instrumental conditioning. Such findings resonate with anatomical and histochemical studies in primates which indicate that ventral parts of putamen share many of the cytoarchitectonic characteristics of nucleus accumbens, as well as sharing similar inputs [153, 154, 155, 156].

The present findings do suggest however, that different parts of the ventral striatum may contribute differentially to distinct forms of Pavlovian-instrumental transfer in humans. This suggestion is based on a comparison of our finding that ventrolat-

eral putamen is involved in outcome-specific transfer in humans with the results of a previous study implicating nucleus accumbens in the general excitatory effects of Pavlovian cues on instrumental performance [86]. It is well established that Pavlovian cues can exert a general, non-specific excitatory effect on the performance of instrumental actions [77, 78, 157, 147, 79, 148] an effect that Talmi et al. [86] demonstrated is mediated by activation of nucleus accumbens and of amygdala. In the context of the present study, these findings suggest that outcome-specific and general transfer may depend on quite distinct neural substrates in humans, mirroring clear double dissociations between the neural circuits known to be involved in implementing these effects in rodents [83, 84]. Although the present study was not designed to assess the effects of general transfer, in future it will be important to compare and contrast outcome-specific and general-transfer effects within the same fMRI study in order to provide a more direct test of the hypothesis that, as in rodents, outcome-specific and general transfer in humans depends on distinct components of ventral striatum.

Note that although we found a remarkably good correspondence between our findings and those from the rodent lesion studies at the level of the ventral striatum, other regions besides ventral striatum have been implicated in specific PIT in rodents including basolateral amygdala [84] and dorsolateral striatum [85]. We did not find any evidence for a differential contribution of these regions in the present study. One possibility is that these areas do play a role in specific transfer effects in humans, but this does not result in a global increase in activity between conditions, and thus does not become manifest with BOLD fMRI. The present results go beyond merely pointing to homologies between outcome–specific transfer effects in rodents and humans. Previous animal studies on this topic have all involved lesion manipulations, which, though important for identifying whether a given region is necessary for implementing specific transfer effects, cannot provide insight into the neural computations underlying such an effect. Here we measured dynamic changes in BOLD responses as subjects made choices that were either consistent with the specific-transfer effect or inconsistent. Responses consistent with the specific transfer effect occurred when subjects chose the outcome compatible with the Pavlovian cue, and inconsis-

tent responses occurred when subjects chose the incompatible option. Even though subjects showed a significant bias toward the compatible action overall, sometimes they chose the incompatible action; this allowed us to compare activity when transfer guided behavior, with activity under identical stimulus conditions when subjects chose independently of the cue. Activity in ventrolateral putamen was not significantly elevated on trials when an outcome-specific cue was presented compared to control trials where cues for other, unavailable, outcomes were presented, suggesting that outcome-specific transfer effects are not mediated by an overall increase in activity in this area. Furthermore, even on outcome-specific trials where subjects chose the action compatible with the Pavlovian cue, there was no increase in activity compared to non-outcome-specific control trials. Instead, we found a significant decrease in signal on those outcome-specific trials where subjects chose the action incompatible with the outcome, compared to compatible choice outcome-specific trials.

This finding provides insight into the computations that might be taking place in the ventral striatum during outcome-specific transfer effects. Outcome-specific transfer effects are thought to be mediated by outcome-response (O-R) associations that are activated by the Pavlovian cues [79, 158]. A natural hypothesis is that when the action plan activated by the O-R association is feasible (because such an action is available), it must be inhibited before another action can be taken. Note that under this hypothesis, the O-R association needs to be inhibited during the outcome-specific transfer trials when the incompatible response is chosen, but not when the compatible response is selected, or in any of the other control trials. This provides a computational explanation for why suppression of activity in the ventrolateral putamen is observed only in the incompatible outcome-specific transfer trials.

Specific transfer effects from Pavlovian cues have been argued to play a role in addictive behaviors [159]. For example, Hogarth et al. [150] demonstrated specific transfer of a tobacco-seeking response in the presence of a tobacco predicting cue, relative to a money predicting cue. Here we demonstrate similar behavioral results, using non-addictive outcomes, indicating that the observed transfer effects reflect a general property of reward-associated cues that are not specifically related to addictive

stimuli. Nonetheless, there are clear parallels between our experimental design and the potential influence of environmental cues on drug-seeking behavior. Our fMRI results suggest the hypothesis that suppression of an outcome-response association might contribute toward biasing behavior away from cue-compatible responding. This raises the possibility of a future therapeutic intervention in addiction, in which ventrolateral putamen circuitry could potentially be targeted (for instance via a neurofeedback procedure; [100, 160]) to suppress effects of environmental drug cues on drug-seeking behavior.

In this study we have demonstrated an outcome-specific Pavlovian-instrumental transfer effect in humans, which serves to bias action choice towards actions associated with an outcome consistent with a concurrently presented cue. BOLD fMRI measured while subjects performed this task demonstrated a signal decrease in ventrolateral putamen when subjects chose the action incompatible with the cue. This finding points to a computational role for this region in suppressing outcome-response associations, necessary in order to perform an action incompatible with the Pavlovian cue only when a compatible action is feasible. This work adds to our understanding of the neural mechanisms of stimulus-outcome guided decision-making in both animals and humans, which is fundamental for understanding maladaptive choice behaviors such as addiction.

Acknowledgements

We are grateful for assistance from Vivian Valentin, and Hackjin Kim, Sam Huang, and Shawn Wagner for developing the coil we used to detect swallowing movement. This work was supported by a Searle scholarship and a grant from the National Science Foundation (NSF 061714) to JOD, and by grants from the Gordon and Betty Moore Foundation to JOD, AR, and the Caltech Brain Imaging Center, and the Japan Science and Technology Agency (ERATO) to SS.

Chapter 4

Direct instrumental conditioning of neural activity in motor cortex*

Successful learning is often contingent on feedback. In instrumental conditioning, an animal or human learns to perform specific responses in order to obtain reward. Instrumental conditioning is often used by behavioral psychologists in order to train an animal (or human) to produce a desired behavior. Shaping involves reinforcing those behaviors which in a step-wise fashion are successively closer to the desired behavior until the desired behavior is reached. Here, we aimed to extend this traditional approach in order to directly shape neural activity instead of overt behavior. To achieve this we scanned 22 human subjects with fMRI and performed image processing in parallel with acquisition. We delineated regions of interest (ROIs) in finger and toe motor/somatosensory regions, and used an instrumental shaping procedure to induce a regionally specific increase in activity by providing an explicit monetary reward to reinforce neural activity in the target areas. After training, we found a significant and regionally specific increase in activity in the ROI being rewarded (finger or toe) and a decrease in activity in the non-rewarded region. This demonstrates that instrumental conditioning procedures can be used to directly shape neural activity, even without the production of an overt behavioral response. This procedure offers an important alternative to traditional biofeedback-based approaches, and may be useful in the development of future therapies for stroke and other brain disorders.

*Adapted with permission from: Bray S, Shimojo S, O'Doherty JP (2007) Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. Journal of Neuroscience 27:7498-7507. Copyright 2007 *Journal of Neuroscience*

Introduction

In instrumental conditioning, an animal learns to increase the probability of making a particular response in order to obtain reward or avoid punishments. Traditionally, the response consists of overt behavioral actions, such as pulling a lever, traversing a maze, or pressing a button [161, 2, 3, 162, 163]. As the ability to measure neural responses has improved, it has become possible to perform experiments in which an animal is rewarded merely for generating neural activity instead of actually performing an overt motor response [87]. Musallam et al. [94] demonstrated that by recording from neurons in parietal cortex, monkeys could be trained to generate neural responses in order to obtain juice rewards, without emitting any behavior.

Parallel advances in human neuroimaging techniques have enabled neural activity measured by fMRI to be processed and analyzed in parallel with image acquisition (real-time fMRI), making it possible to provide rapid feedback of activity in specific brain regions to the subject during an on-going experiment [164, 165, 105]. This technique has previously been used to assess human subjects' ability to modulate their own brain activity, by providing an on-line graphical representation of activity in a specific brain region [96, 100, 107]. This approach has much in common with traditional biofeedback techniques that have provided on-line feedback of physiological responses such as heart rate or scalp EEG [166, 92].

In the present study we explore an alternative approach for modulating neural activity to the standard biofeedback paradigm. Here, instead of providing an on-line representation of neural activity and requiring subjects to actively modulate that activity in order to reach a specified goal, we used procedures derived from instrumental conditioning, whereby an actual reward (monetary gain) is the only feedback subjects receive contingent on their performance. This instrumental training procedure allows one to employ 'shaping' [4], in which the threshold for reward is gradually increased in order to induce incremental improvements in performance.

The aim of the present study was to determine whether it is possible to use instrumental conditioning techniques to modulate neural activity in the human brain.

For this we delineated two regions of left sensorimotor cortex (activated by imagined flexion and extension of fingers and toes), and attempted to train subjects to activate one region in response to a visual cue, while suppressing the second region. A further aim was to determine the extent to which learned modulation of motor cortex in the absence of movement might subsequently influence overt motor behavior as assessed by a speeded reaction time task. Subjects performed a task in which the cues used during conditioning were alternately displayed on a screen and intermittent cues instructed them to respond as quickly as possible with fingers or toes. Modulation of reaction times by exposure to instrumental cues offers a measure of how learning of cue contingencies affects concurrent processing of motor responses.

Materials and Methods

Experiment 1

Subjects

A total of 26 right-handed healthy normal human subjects participated in the experiment, 14 males and 12 females, aged 18 to 39 years with a mean age of 25.4 years. All subjects gave informed consent, which was approved by the local research ethics committee. The first 7 subjects performed only the pre-training and conditioning components of the study. The remaining 19 subjects also performed a reaction time task before and after conditioning.

Four subjects were removed from the imaging analysis, three of which were also removed from the reaction time analysis. One subject was eliminated from the imaging analysis due to excessive head movements during the final run. Two other subjects were eliminated from all analyses due to inability to learn the task. An additional subject was removed from all analyses for failing to comply with task instructions. For one subject the experiment terminated on the 9th trial of the last block due to equipment failure. This left a total of 16 subjects in the reaction time analysis and 22 in the imaging analysis.

Stimuli

During the conditioning task, subjects were presented with one of three brightly colored abstract fractal images (100 x 100 pixels) centered on a grey background (800 x 600 pixels). One of the fractals had the word 'Rest' written across it in white letters, to clarify that this cue meant that the subject should be resting. A different set of stimuli were used during the pre-training and functional localizer, where each task (real or imagined / hand or foot movements) was associated with a centrally presented colored circle with a radius of 100 pixels, and lettering coding for each task as described below. During the reaction time task, the fractal images were presented at an offset of 125 pixels above center and responses were prompted using the brightly colored circles used in the localizer task. All stimuli were presented using the Cogent 2000 Matlab toolbox.

Pre-Training and Functional Localizer Tasks

The functional localizer task consisted of blocks of real and imagined movement, alternating with periods of rest. Subjects ran through this task once outside the scanner as pre-training, so that they could familiarize themselves with the task. Movement tasks consisted of: 1) bending fingers II-V at the metacarpophalangeal joint and 2) flexing and extending all five toes through their full range of movement. During imagination blocks, subjects were instructed to imagine what it would feel like to produce these movements without actually moving. The functional localizer sequence of [resting, finger tapping, resting, imagined finger tapping, resting, toe tapping, resting, imagined toe tapping] blocks was repeated five times. During pre-training, blocks were 10 seconds in duration and during the functional localizer performed in the scanner, blocks lasted 15 seconds. Subjects were cued as to which task to perform by brightly colored visual stimuli with letters coding for the task: red circle with an 'R' for rest, green circle with a 'HaT' for hand/finger tapping, a blue circle with 'HaI' for imagined hand/finger tapping, an orange circle with 'FoT' for foot/toe tapping and a yellow circle with 'FoI' for foot/toe imagined tapping. During both the pre-training

and scanner sessions, subject motion was recorded, as described below. Additionally, during pre-training, we were able to observe subjects at close range to confirm that they were not moving during the imagined movement periods.

ROI selection

After completion of the functional localizer task in the scanner, the resulting images were sorted into resting and task periods, and t-tests applied to generate probabilistic activation maps. Two regions of interest were selected for each subject: one for hand-motor areas activated by imagined finger tapping and one for foot-motor areas activated by imagined toe tapping. In both cases a mask was generated from the contrast of actual movement vs. resting periods. The mask was used to spatially constrain the results of a second contrast comparing imagined movement of fingers to imagined movement of toes. This contrast was chosen to identify regions associated with imagining moving each body part specifically, rather than areas activated by motor imagery in general. From this second map, an ROI center was chosen among the most significant regions, using prior anatomical knowledge of where finger/toe motor cortical areas should be located. A rectangular area of 6x6 voxels in the x-y plane and 3 voxels in the z-direction was generated around the chosen center. The ROI for each subject comprised a maximum of 108 voxels; in some subjects this number was smaller if the volume defined by the rectangle stretched beyond the spatial extent of the brain.

Neuroconditioning Procedure Task and Instructions

Subjects were instructed that during this part of the experiment, they should never perform any real movements, but must only use their imagination or state of mind to increase activity in the specific brain regions defined during the localizer task, corresponding to imagined finger and toe tapping respectively. A reinforced conditioning trial is illustrated in Figure 4.1a. Each trial began with a resting cue for a variable duration between 15 and 20 seconds. Next, the subjects saw one of two fractal cues for 15 seconds. Each 'active' cue meant that if the subject sufficiently activated one

of the regions of interest, they could earn a reward. Data were analyzed online after 14 seconds, and after the 15th second subjects received visual feedback indicating whether they had successfully earned a reward. Positive reward feedback consisted of a picture of a dollar and the phrase 'You have won ONE dollar', while negative feedback was represented by a picture of a scrambled dollar, along with the phrase 'You have not won ONE dollar'. Dollars earned during the task corresponded to real money paid to the subject at the end of the experiment. At the start of the experiment, subjects did not know which cue corresponded to which brain region. They were told that they would have to proceed by trial and error in order to discover the meaning of each cue, and that once they learned the meaning it would stay the same for the duration of the experiment.

Subjects were told that the 'resting' period preceding each active period would serve as a baseline against which the activity during the 'active' periods would be compared. Therefore, they should try to relax as much as possible during 'rest' periods and not practice mental imagery similar to during the 'active' periods. They were also told that in order to earn a reward they would have to activate one region specifically and not both regions. Subjects were told that any kind of mental imagery could be appropriate as long as it specifically activated brain regions delineated by the imagined finger and toe tapping tasks, but that strategies involving motor imagery might be more likely to succeed, given the known functional responses of these regions. Subjects were told that the threshold defining the minimum activity required to get rewarded would be slowly increasing, therefore they would have to improve on their strategy in order to continue earning rewards.

The total duration of the experiment was approximately 1.5 h in a single session. In this time subjects performed reaction time tasks, pre-training, a functional localizer, and 4 conditioning blocks consecutively with 14 trials in each; trials were ordered pseudorandomly so that each trial type appeared 7 times within a block without 3 consecutive trials being of the same type. Each block was approximately 8 minutes long for a total of 32 minutes of training.

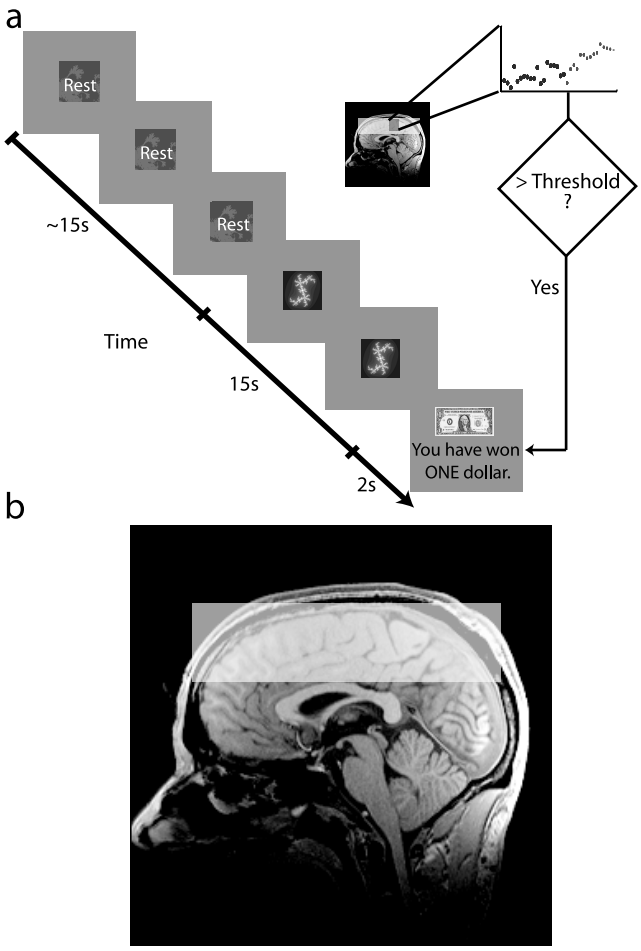

Figure 4.1: a) Example time course for a conditioning trial. Subjects were presented with a resting cue for a variable interval between 10 and 20 s, followed by a cue to activate a specific brain region for 15 s. A percent-change value from resting to active was computed online and compared to the current threshold. If the threshold was exceeded, subjects were shown a picture of a dollar bill, indicating that they had won one dollar, otherwise a scrambled picture of a dollar was shown, for 2 s. b) Diagram showing typical fMRI slice coverage, overlaid on a sagittal slice from a single subjects anatomical scan. We imaged 16 3 mm slices, straight across the top of cortex

Post-experimental debriefing

After the experiment was completed, subjects were asked to complete a short questionnaire. This form asked them to briefly describe what they were thinking about when they saw each of the three cues (one rest, two active) on the display, and to indicate how their strategy might have changed across runs.

Motion Recordings

To control for subject motion during periods of imagined movement, we recorded EMG from the forearm (flexor digitorum superficialis muscle) to measure muscle activity related to finger flexion and extension. We also used a finger twitch sensor (Biopac Systems Inc., Goleta, CA, USA), placed lengthwise along the bottom of the foot and attached by Velcro around the big toe and below the ball of the foot. The sensor is essentially a variable resistor sensitive to bending and compression, and therefore generated a potential difference when subjects bent their toes downwards. Both movement recording devices that we used are MRI compatible, but fMRI scanning introduced noise into the recordings. These data were analyzed by comparing the RMS signal value during resting, active, and imagined periods. Recordings obtained while the scanner was running were smoothed using a 15 point (twitch sensor) or 25 point (EMG) median filter to reduce the impact of scanner noise on signal detection.

fMRI scanning procedure

fMRI data were acquired on a Siemens AG (Erlangen, Germany) 3T TRIO MRI scanner; Blood Oxygenation Level Dependent (BOLD) contrast was measured with gradient echo T2* weighted echo-planar images (EPI). We used an 8 channel phased array coil. The first 5 volumes were discarded to permit T1 equilibration. In order to keep the repetition time (TR) at 1 s, we imaged only 16 3 mm slices across the top of cortex. Typical slice coverage is illustrated on a single subjects' anatomical scan in Figure 4.1b. Scan coverage was therefore limited to superior and middle frontal gyri, pre- and post-central gyri, and superior parietal lobule. Other scan parameters were

the following: in-plane resolution, 3 x 3 mm; echo time, 30 ms; field of view, 192 x 192 mm. After the conditioning procedure a T1 weighted structural image was acquired for each subject, as well as a set of ~6 32-slice EPI images (to improve coregistration and normalization of images to a template).

Concurrent fMRI analysis and processing

As soon as images were reconstructed, they were transferred in real-time via TCP/IP socket to an external Intel Xeon workstation (3.8 MHz 64-bit processor running Redhat Linux); data processing was performed using MATLAB 7.0 (The Mathworks Inc., Natick, MA).

Pre-processing Image pre-processing consisted of motion correction using AFNI [167], and linear detrending to correct for low-frequency scanner drift. During functional localizer scans spatial smoothing using a two-dimensional Gaussian of 5 mm width was performed prior to performing statistical tests. During the conditioning task no temporal or spatial smoothing was performed.

Reward Criterion Two thresholds shared equal priority in the decision rule for determining whether a subject had earned a reward on a particular trial: one threshold on the minimum %-change within a region and a second threshold on the difference between the %-change in the rewarded region and the non-rewarded region. Both thresholds had to be exceeded for a subject to earn reward on a given trial, and both were adapted according to a modified percentile reinforcement schedule [16]. They both started at 0, and increased only after the current threshold had been exceeded 4 times. At this time, both thresholds were set to be the lowest of the four values that had beaten the previous value. If a reward was not obtained on one of the next 4 trials, one or both thresholds was reset to its' previous value, depending if one or both conditions was not met. In this way, the thresholds for the signal level and the difference increased together, but were reset separately.

As images arrived on the external workstation, they were pre-processed and the

signal was averaged over all voxels in the previously defined ROIs. After one variable length baseline period (10 - 20 seconds) and one 14 second active period had elapsed, the ROI signal was averaged over each time period (excluding the first two seconds to allow for some lag in the hemodynamic response), and a percent change from baseline to active was computed for each ROI. For the ROI being rewarded, the %-change was compared to the current threshold, and the difference between the %-change in the two ROIs was compared to the difference threshold. If both conditions were met the current trial was 'rewarded' after the 15th second, when the subject would see the 'reward' feedback for 2 seconds. If the reward conditions were not met, they would see the 'no reward' feedback for 2 seconds.

Performance-based grouping of subjects

For some analyses, we divided subjects into groups depending on their performance during the last experimental run: subjects who earned fewer than 5 rewards on the final run were classified as poor learners, relative to those who earned more than 5 rewards. Performance during the last conditioning run was especially relevant to analysis of the reaction time measures taken immediately afterwards, since that should give the most current estimate of the subjects' level of learning. Some subjects reported tiring towards the end of the experiment, which could corrupt learning related effects in the reaction time analysis. This criterion put 17 subjects in the 'good-learner' category and 6 in the 'poor-learner' category.

Group fMRI %-change analysis

We performed a group analysis on the trial-by-trial percent-change values measured during conditioning. Trials in which twitching movements were visible in the EMG traces were eliminated, as were trials in which large head movements caused sharp deflections in the BOLD signal time course. We performed a repeated measures ANOVA on the averaged %-change values during each run, with within-subject factors of ROI (3 levels: hand ROI, foot ROI, and whole-brain background ROI), rewarded ROI (2 levels: hand rewarded and foot rewarded) and run (4 levels), and a single

between subjects factor, a binary value indicating whether or not the subject was a 'good-learner'.

To look for trial-by-trial increases in signal difference between the two ROIs, we averaged the %-change value across subjects on each trial and performed a linear regression on the difference between the signals in each ROI.

SPM analysis

Data were pre-processed using the SPM5 software package (`http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`). Images were corrected for slice timing and spatially realigned to the first image from the functional localizer. One of the 32 slice EPI images collected at the end of the experiment was used to improve coregistration and spatial normalization. The 16 slice EPI images were coregistered to a 32 slice EPI, which was in turn coregistered to the T1-weighted anatomical scan. The T1 image was segmented into white and grey matter, and the grey matter was coregistered and normalized to the template grey matter image distributed with SPM5 (in Montreal Neurological Institute space). These parameters were subsequently applied to the T1 image itself as well as the set of 16 slice EPI images. Spatial smoothing was then applied to the 16 slice EPI images using a Gaussian kernel with full width at half maximum of 8 mm.

The four conditioning sessions for each subject were modeled in SPM using a finite impulse response model, with separate regressors for hand and foot rewarded trials, for each run. The six ongoing motion parameters estimated during realignment were included as regressors of no interest.

Parameter estimates were modeled with a full factorial model with 2 factors: rewarded region (2 levels) and session (4 levels). This created an 8 column design matrix for each subject, each column corresponding to a session x rewarded-region interaction term. Linear contrast images from these design matrices were taken to the random effects level by applying t-tests between them to produce group statistical parametric maps.

Reaction Time Task

Subjects performed a simple reaction time task before and after conditioning. They were randomly presented with one of the three fractal cues used in the conditioning task (referred to here as the background cue), slightly above center on the screen for a time uniformly distributed between 1 and 2 seconds. After this time the background cue remained on the screen and a second cue appeared in the center of the screen (referred to here as the response cue), either a green circle containing the letters 'HaT' or an orange circle containing the letters 'FoT'. This second cue instructed subjects to respond by pressing a button on the keypad in their hand (HaT) or strapped to the bottom of their foot (FoT). Both cues remained on the screen for 1 second. Subjects responded 30 times to each of the six possible combinations of background cue and response cue.

During the conditioning task, subjects learned to associate the fractal cues with either a hand-imagine or foot-imagine response, so that after the experiment the background cues can be considered either compatible (e.g., hand imagine cue and hand response cue) or incompatible (e.g., hand imagine cue and foot response cue). Trial-by-trial reaction times measured before and after the conditioning task were divided into three blocks, early, middle, and late, and averaged within each trial type. The block-averaged reaction times were analyzed with a repeated measures ANOVA, with within-subject factors of block (3 levels), time (2 levels: before and after conditioning), cue-response relation (3 levels: compatible cue, incompatible cue, rest cue), response type (2 levels: hand and foot), and a single between-subjects factor, a binary value indicating whether or not a subject performed well on the task.

Experiment 2

In Experiment 1 the behavioral reaction time measure was taken outside the scanner, before and after the experiment. This meant that subjects were exposed to the cues in a different context and that any response evoked by the fractal cues could diminish since the test was performed in extinction (responses were not rewarded).

We conducted a follow-up experiment to test the effect of performing the reaction time measure in a similar context as the conditioning and interleaved reaction time measurements with conditioning trials in order to minimize the effects of extinction. In Experiment 2 we also included a condition to control for the effects of repeated practice alone, without contingent feedback.

Subjects

We scanned an additional 9 subjects in an alternate version of the conditioning task that included a reaction time measure taken during the conditioning trials (aged 21-38 mean 24.9, 3 males). One subject performed only 3 out of 4 sessions due to discomfort, and the imaging data from another were not analyzed due to excessive head movements. We also scanned 9 subjects (aged 21-34 mean 23.3, 4 males) in a control condition.

Conditioning with interleaved reaction time task

A separate group of 9 subjects underwent a conditioning procedure nearly identical to experiment 1, but with additional sets of reaction time trials randomly inserted among the regular conditioning trials in each block. For this task subjects held a button pad in their right hand and a second button pad was held against the bottom of their foot in a sandal so that they could push a button with their toe. The reaction time trials began with a central fixation cross presented for 250 ms, followed by one of the two fractal cues from the conditioning trials for 1.75 s. Either 'HaT' or 'FoT' then appeared on the fractal for 250 ms, instructing subjects to respond by pressing the hand button or foot button, respectively. The fixation cross appeared for 1.75 s, during which time subjects made their response. Each session consisted of 14 conditioning trials with two sets of 30 consecutive reaction time trials inserted at pseudorandom intervals. In the first session they always appeared after the 12th and 14th trial, to give subjects the opportunity to learn the response associated with each fractal. In subsequent sessions, the blocks of reaction time trials appeared at random intervals, with the condition that two blocks could not be presented consecutively.

In each session, 15 of each type of trial: hand-cue/hand-response, hand-cue/foot response, foot-cue/hand-response, foot-cue/foot-response, were presented in random order in two sets of 30 trials. Across the four conditioning blocks, a total of 60 reaction times from each type of trial were collected.

Control Task

We ran a second version of the task designed to control for the effects of repeated practice of motor imagery. During the conditioning trials, these subjects were instructed to imagine either hand or foot movements when they saw the corresponding cue, and to ignore the reward feedback. Unlike in the feedback task the rewards delivered to subjects were not linked to neural activity but instead each subject in the control group experienced the rewards obtained by a randomly assigned 'yoked' subject from the feedback group. The control task also included reaction time trials identical to those in the feedback task.

Reaction time analysis

During the conditioning task, subjects learned to associate the fractal cues with either a hand-imagine or foot-imagine response, so that the background cues can be considered either compatible (e.g., hand-imagine background cue and hand response cue) or incompatible (e.g., hand-imagine background cue and foot response cue) with the response. We hypothesized that there would be a facilitation for compatible stimuli relative to incompatible, i.e., faster reaction times. We log transformed these data and entered them into a 3-way repeated measures ANOVA, with within-subject factors of cue (hand/foot), response (hand/foot), and session (1-4), separately for the feedback and control groups.

Results

Experiment 1

Behavioral Results

Post-experimental debriefing According to the questionnaire responses, all subjects excepting two correctly discriminated between the two cues, and were aware which cue instructed them to activate hand or foot areas. The two subjects who did not learn correctly performed very poorly at the task and were eliminated from both the behavioral and imaging analyses. The self-reported strategies for activating these areas all involved motor imagery of some kind. During the resting period subjects either relaxed and let their mind wander or distracted themselves by repeating a song or numbers in their head. Some subjects reported making eye movements during these periods.

Subject performance on conditioning task Most subjects were able to successfully obtain rewards in the task. The mean number of rewards obtained per run were [3.27 $\pm$ 0.40, 3.41 $\pm$ 0.31, 2.82 $\pm$ 0.30, 3.27 $\pm$ 0.35] for hand rewarded trials and [3.05 $\pm$ 0.32, 2.82 $\pm$ 0.37, 3.09 $\pm$ 0.40, 2.82 $\pm$ 0.39] for foot rewarded trials. The number of rewards remained relatively constant across runs, a repeated measures ANOVA with between-subjects factors of rewarded region (2 levels) and run (4 levels) yielded no significant main effects or interactions. However the threshold for the activation level which subjects had to achieve in order to obtain reward increased across trials. A linear regression on the trial-by-trial mean threshold across subjects, shows a significant increase, both for the hand rewarded threshold (β = 9.05x10-5, R^2 = 0.8717, p$<$0.001) and foot rewarded threshold (β = 1.51x10-4, R^2 = 0.9634, p $<$0.001). Since subjects were able to maintain a constant rate of reward despite the increasing difficulty of the task, we consider this a measure of overall success of the conditioning procedure.

Table 4.1: Movement recording comparisons

	Condition (vs. rest)	Subjects showing sig difference
Pre-training (Wilcoxon ranksum, n1 = 5, n2 = 5; $p<0.05$)	Finger Tapping	16/22
	Imagined Finger Tapping	0/22
	Toe Tapping	21/22
	Imagined Toe Tapping	0/22
Functional Localizer (Wilcoxon ranksum, n1 = 5, n2 = 5; $p<0.05$)	Finger Tapping	19/22
	Imagined Finger Tapping	0/22
	Toe Tapping	20/22
	Imagined Toe Tapping	0/22
Conditioning Task (Wilcoxon ranksum, n1 = 7, n2 = 7; $p<0.05$)	Imagined Hand Movement	0/22
	Imagined Foot Movement	0/22

Movement Recordings Movement recordings during the localizer task, both during pre-training and in the scanner, confirmed that subjects were able to perform the imagination task without actually moving. We compared RMS values during resting periods to real and imagined movement periods, during the pre-training, functional localizer and conditioning task. The results are summarized in Table 4.1, and example recordings for real and imagined movements with fingers and toes are shown in Figure 4.2. For some subjects, the difference between rest and movement did not reach significance; inspection of the movement time courses showed that these subjects had probably adjusted their position during the resting period of one of the blocks and due to the small number of blocks (n = 5) the comparison did not reach significance.

Although subjects were instructed that they should keep their hands and feet still, some subjects showed evidence of hand twitches during certain trials. Trials in which sharp spikes in the EMG indicated a small twitch in the hand or arm, either during rest, hand-imagined, or foot-imagined periods, were removed from further analysis. The mean number of trials eliminated per subject was 5/56, with a standard deviation of 4.33.

Reaction times The repeated measures ANOVA performed on the reaction times yielded a significant main effect of response type ($p<0.001$; $F(1,14) = 48.374$), and

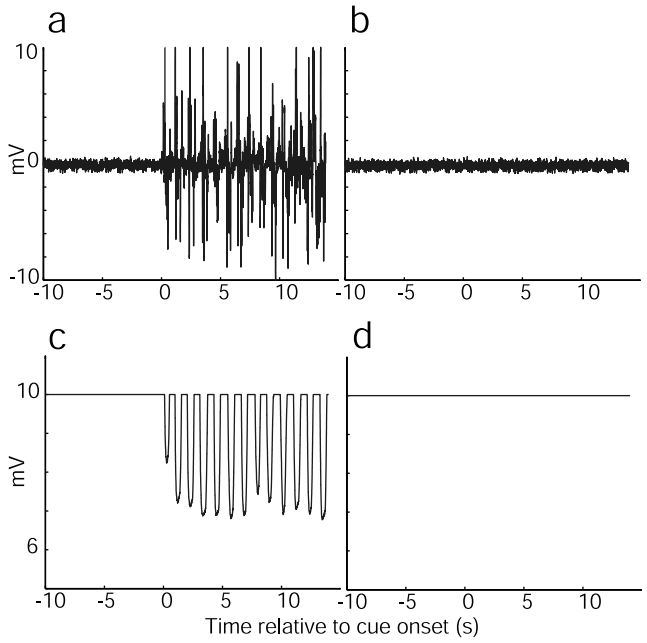

Figure 4.2: Sample movement recordings from a single subject during Experiment 1. a) EMG during real hand movement. b) EMG during imagined hand movement. c) goniometer recording during real foot movement. d) goniometer recording during imagined foot movement

significant interactions of time x cue-response relation x learner type ($p<0.01$; $F(2, 28) = 6.011$), and time x response type ($p<0.01$; $F(1,14) = 13.292$). Planned t-contrasts showed that in good learners there was a significant difference between both compatible (paired t-test $p<0.05$; $N = 11$ $|t| = 2.1871$) and incompatible (paired t-test $p<0.05$; $N = 11$; $|t| = 1.8013$) cue types after conditioning compared to before, with both becoming slower after conditioning. However, contrary to our hypothesis, we did not observe a significant difference between responses that were compatible or incompatible with the background cue. To address the possibility that the absence of this effect was a consequence of the reaction time measure being performed outside the scanner and therefore in extinction, in Experiment 2, reaction times were tested in the scanner interleaved with conditioning trials to reduce extinction of the response (see Methods: Experiment 2 for more details).

fMRI Results

ROI location The mean ROI center for the hand region in Montreal Neurological Institute (MNI) space was [-35 ± 1.24, -26 ± 1.9, 65 ± 0.941] located on the left precentral gyrus (Broadmann area 4a, 6, 1) [168]; individual subject ROI centers were located near the hand knob [169] on the pre- and post-central gyri. The mean ROI center for the foot region was [-6 ± 0.729, -25 ± 1.5, 69 ± 1.1] located on the left paracentral lobule (Broadmann area 4a, 6); individual subject ROI centers were distributed from the posterior part of the superior frontal gyrus along the length of the paracentral lobule. These areas are highly consistent with finger- and toe-imagery-specific locations found in Ehrsson et al. [170].

Trial-by-trial %-change in regions of interest Averaging the trial-by-trial %-change data across trials within each session, and over subjects, we see a general increase in signal in the rewarded ROI, and a decrease in the non-rewarded ROI, corresponding to an overall increase in the signal difference between the rewarded and non-rewarded regions; these data are plotted in Figure 4.3. In addition to the two pre-defined ROIs, we also looked at the signal in a large background ROI which included all brain voxels outside of the two task-related ROIs. The background ROI did not show the same increase as the rewarded ROI, confirming that the activation in response to the cue was specific to the rewarded ROI rather than reflecting a nonspecific increase in brain activity.

To test for a learning effect we performed a repeated-measures ANOVA on the trial-averaged %-change measures within each session from each ROI. Across all 22 subjects, we found a significant main effect of ROI ($p<0.005$; $F(6,120) = 6.246$), and significant interactions of ROI x rewarded ROI ($p<0.001$; $F(2, 40) = 14.308$), as well as an interaction between ROI x rewarded ROI x session that approached significance ($p = .064$), suggesting a learning effect. Restricting our analysis to a subgroup who successfully met a learning criterion of 5 or more rewards during the last session ($N = 17$), this interaction became significant ($p<0.05$; $F(6,96) = 3.907$). Taking the trial-by-trial average across all subjects and regressing the mean difference between ROIs

onto trial number, we found a significant positive increase, both for [hand ROI-foot ROI] in trials when an increase in the hand ROI was rewarded (β = 0.0001, R^2 = 0.3840, $p<0.05$) and [foot ROI-hand ROI] in trials when an increase in the foot ROI was rewarded (β = 0.0001, R^2 = 0.2557, $p<0.05$).

Random effects analysis with SPM We generated a contrast to detect regions in which signal increased during hand rewarded trials and decreased during foot rewarded trials, and likewise a second contrast to detect regions with signal increase during foot rewarded trials and decreases during hand rewarded trials. Taken to the random effects level, the contrast to detect activity during foot rewarded trials showed a significant cluster with peaks surviving small volume correction around the mean foot ROI center [-6 -24 69] at [-3 -24 75] (t = 5.06; $p<0.01$ FDR-corrected), [0 -21 72] (t = 4.95; $p<0.01$ FDR-corrected), and [-3 -18 69] (t = 4.71; $p<0.01$ FDR-corrected). The results of this contrast are shown in Figure 4.3c.

The contrast to detect activity during hand rewarded trials shows a large cluster with a peak at [-39 -33 66] which survives small volume correction in an 8 mm sphere around the mean of subjects' hand ROI centers [-36 -27 66] (k = 38; t = 3.13; $p<0.05$ FDR-corrected). The results of this contrast are shown in Figure 4.3f.

The hand-region and foot-region activation tasks engaged a network of brain regions in addition to the regions of interest, though activations in these regions remained relatively constant across the study (see Table 4.2). As would be expected, there was substantial overlap between regions activated by imagined hand and foot movements, in dorsal pre-motor (PMd) extending into supplementary motor (SMA) and pre-SMA as well as bilateral regions of the parietal cortex and precentral gyri. In Figure 4.4 we have plotted the parameter estimates for the hand and foot rewarded trials in each of the 4 sessions: in 4.4a the significant regions in the foot-region activation task, and in 4.4b in the hand-region activation task. Despite the fact that these regions were generally activated by subjects performing the task, our protocol caused selective enhancement and depression of activity only in the delineated regions of interest. This can be seen from the slopes and divergence of the curves in the topmost

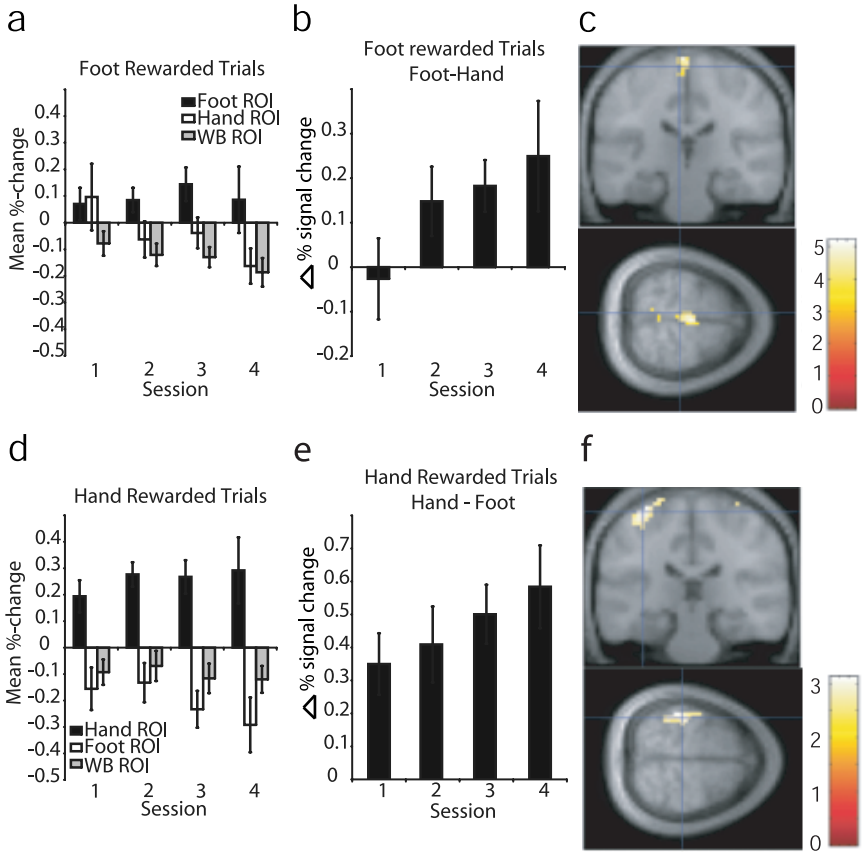

Figure 4.3: fMRI results from Experiment 1. a) Mean percent-change data averaged over subjects within runs in each ROI during trials in which the foot ROI was rewarded, b) Difference in mean percent-change data, averaged over subjects within runs, between foot and hand ROIs during foot rewarded trials. c) Results of random effects analysis in SPM from Experiment 1; t-test on contrast increasing during foot rewarded trials and decreasing during hand rewarded trials, thresholded at $p<0.01$, crosshairs indicate mean of subjects ROI centers for the foot ROI [-6, -25, 69]. d) Averaged responses in each ROI during trials in which the hand ROI was rewarded e) Difference between hand and foot ROIs during hand rewarded trials. f) Results of random effects analysis in SPM from Experiment 1; t-test on contrast increasing during hand rewarded trials and decreasing during foot rewarded trials, thresholded at $p<0.001$, crosshairs indicate mean of subjects ROI centers for the hand ROI [-35, -26, 65]

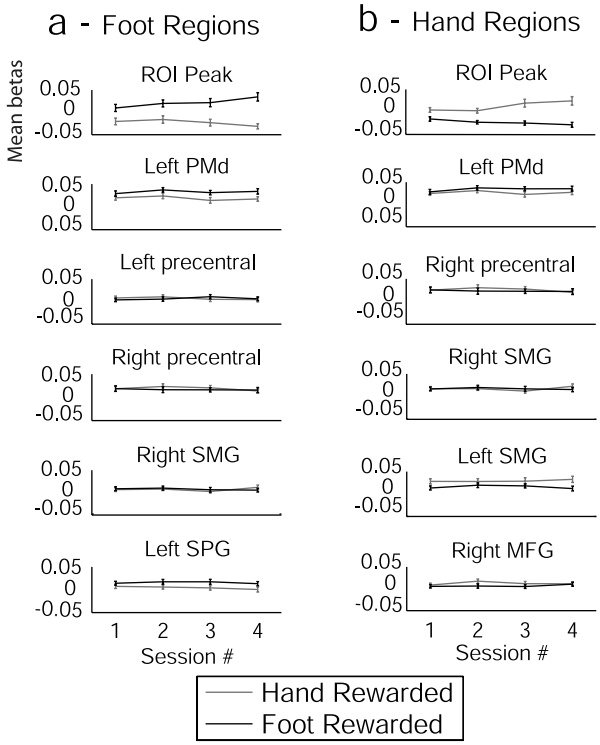

Figure 4.4: Subject averaged parameter estimates across sessions from Experiment 1. Hand and foot rewarded trials are plotted separately. Bars indicate standard errors. Regional coordinates in Table 2. a) Regions identified as significant during trials when subjests were rewarded for activating the foot region. b) Regions identified as significant during trials when subjects were rewarded for activating the hand region. (PMd = dorsal pre-motor, MFG = middle frontal gyrus, SMG = supramarginal gyrus, SPG = superior parietal gyrus)

plots (from the peak voxel near the ROI centers described above), compared to the other regions significantly activated by a general task–baseline contrast.

Experiment 2

Behavioral Results

Reaction times The results of the ANOVA on the feedback group showed that subjects were significantly faster to make a response when the background cue was compatible with the type of response, as demonstrated by a significant interaction between cue and response ($p<0.05$; $F(1,7) = 7.23$). We also found significant main effects of session ($p<0.05$; $F(3,21) = 4.134$) and response ($p<0.01$; $F(1,7) = 17.7$) –

Table 4.2: Motor imagery tasks

Z scores and MNI coordinates of peak activation foci, $p<0.001$, minimum cluster size 5 voxels

	Contrast			
	Hand Rewarded		Foot Rewarded	
Region	voxels	Z	# voxels	Z
Dorsal pre-motor (PMd)	1001	5.79(-12 -3 69)	997	5.94(-12 -6 72)
Left precentral gyrus			6	3.29(-48 0 54)
Right precentral gyrus	68	4.79(57 0 48)	44	4.22(57 0 48)
Right parietal (supramarginal gyrus)	64	5.40(60 -27 51)	47	5.47(60 -27 54)
Left parietal (supramarginal gyrus)	79	4.49(-36 -48 60)		
Left superior parietal gyrus			71	4.35(-18 -60 69)
Right middle frontal gyrus	5	3.51(30 -3 72)		

subjects responded more quickly with fingers than toes. In the control group only the main effect of response was significant ($p<0.05$; $F(1,7) = 11.29$), in that subjects were faster responding during hand than foot movements, but no significant cue or cue x response effects were found in this group.

fMRI Results

ROI location The ROIs identified in Experiment 2 were similar to Experiment 1. For the feedback group, the mean ROI center for the hand region in MNI space was [-39 ± 2.2, -25 ± 2.1, 58 ± 1.8], and for the foot region [-6 ± 0.7, -25 ± 1.5, 69 ± 1.1]. The ROI centers for the control group were statistically indistinguishable from the feedback group, with the mean hand ROI center at [-37 ± 2.2, -23 ± 1.1, 56 ± 1.3] and the mean foot ROI center at [-7.6 ± 0.6, -29 ± 2.5, 69 ± 1.0].

Trial-by-trial %-change in regions of interest We averaged the trial-by-trial %-change data across trials within each session, and over subjects. In the feedback group we see a general increase in signal in the rewarded ROI, and a decrease in the non-rewarded ROI, corresponding to an overall increase in the signal difference between the rewarded and non-rewarded regions. In the control group the difference between the two regions is stable or decreasing. These data are plotted in Figure 4.5:

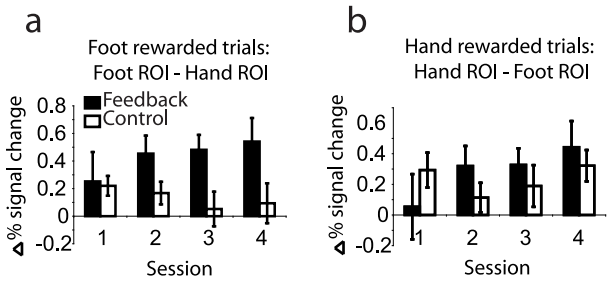

Figure 4.5: Percent-signal change plots across sessions for feedback and control groups from Experiment 2 a) difference in percent-change between foot and hand ROIs when foot responses were rewarded. b) Difference in percent-change between hand and foot ROIs when hand responses were rewarded

4.5a shows the %-change difference for the foot-region activation task and 4.5b for the hand-region activation task. Taking the trial-by-trial average across all subjects and regressing the mean difference between ROIs onto trial number in the feedback group, we found a significant positive increase both for [hand ROI - foot ROI] in trials when an increase in the hand ROI was rewarded ($R^2 = 0.23$, $p<0.05$) and [foot ROI -hand ROI] in trials when an increase in the foot ROI was rewarded ($R^2 = 0.22$, $p<0.05$). By contrast, no significant linear increase was seen in the control group, either in the hand-imagine or foot-imagine conditions, suggesting that repeated practice of motor imagery is not sufficient to explain the shaping of neural responses demonstrated here and in Experiment 1.

Random effects analysis with SPM We generated contrasts comparing activity during hand-imagine periods and foot-imagine periods, and took them to the random effects level. Consistent with the results from Experiment 1, significant activity was found in the foot region in the contrast of foot-cue trials > hand-cue trials (Figure 4.6a), within an 8 mm sphere corrected for small volume around the mean center of the foot ROIs for the feedback group at [-6 -27 69] ($t = 4.04$; $p<0.05$ FDR-corrected). Significant activity was also found in the hand region in the contrast of hand-cue trials > foot-cue trials (Figure 4.6b), which survived correction for small volume within an 8mm sphere centered around the mean of the hand ROIs for the feedback group at [-42 -33 54] ($t = 5.66$; $p<0.05$ FDR-corrected).

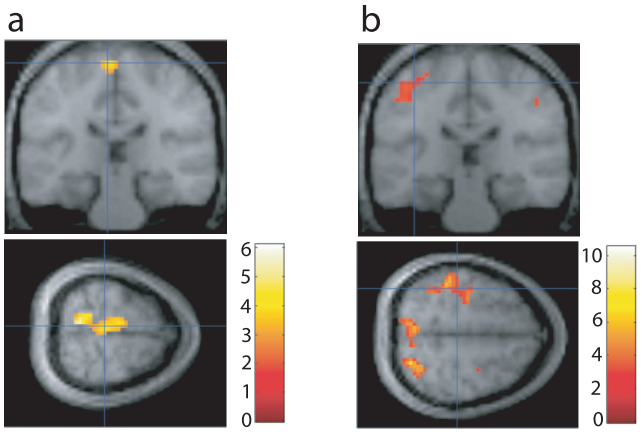

Figure 4.6: Random effects and ROI analyses from Experiment 2. a) Results of a contrast of foot-cue vs. hand-cue conditions across all four sessions. Crosshairs are centered on mean of subjects ROI centers for the foot ROI [-6 -30 69]. Results are shown at $p<0.01$ for visualization, but survive correction for small volume at $p<0.05$. b) Results of a contrast of hand-cue rewarded vs. foot-cue rewarded trials across all four sessions from the feedback group. Crosshairs are centered on mean of subjects ROI centers for the hand ROI [-39 -27 57]. Results are shown at $p<0.01$ for visualization, but survive correction for small volume at $p<0.05$.

ROI based comparison of effects in feedback and control groups We next compared the mean parameter estimates from each ROI between the feedback and control groups. During the hand-cue condition, neural activity in the hand ROI was significantly greater in the feedback than the control group during the last 2 sessions once learning was consolidated in the feedback group ($t(15) = 1.9$, $p<0.05$ one-tailed). During the last two sessions of the foot-cue condition, neural activity in the foot ROI was also significantly greater in the feedback group than in the control group ($t(15) = 3.2$; $p<0.005$).

Discussion

In this study we have shown that it is possible to directly condition neural activity using reward feedback derived from fMRI. Subjects were able to discriminate between two cues and respond to each by activating the appropriate region of their left sensorimotor cortex, while suppressing activity in a second region. Post-hoc analysis

showed that the brain regions significantly increasing in response to rewarded cues and decreasing in response to non-rewarded cues were spatially limited to the specific brain regions where activity was reinforced in our procedure. We also demonstrated in a control group that repeated practice of motor imagery alone is not sufficient to account for this effect. A behavioral reaction time measure showed that in the context in which the association was learned, a neural response to a cue can have a facilitatory effect on reaction times, when the physical response engages regions similar to those activated by the learned neural response. Taken together these findings could lead to development of therapies for patients who have suffered stroke damage to the motor system.

Behavioral shaping has long been known to be a powerful method for behavioral modification in both humans and animals [2, 4]. Here we have used the methods derived from behavioral shaping to directly shape neural activity. Our goal in this study was to show that by using a reward schedule based on behavioral shaping we could train subjects to increase the level of their neural responses in a specific brain region over time. Shaping schedules constantly adjust the threshold required to earn reward, based on subjects' prior performance, thus ensuring that subjects are in a state of constant learning [171]. Our procedure succeeded not only in increasing activity over time, but also in selectively increasing and decreasing activities in the specific regions of interest, while activities in other regions recruited by this task remained stable.

The approach used here offers an important alternative to that employed in previous fMRI neurofeedback training studies [100, 96, 107, 99]. In these previous studies, explicit visual feedback was provided to subjects, signaling the level of activity in a particular area. Subjects were then instructed to modulate their activity in order to attain a specific target level of activation. However, in the present study no visual feedback was presented. Subjects were instructed to activate a specific brain region and received an actual tangible reward (here winning one US dollar) if they succeeded in reaching a criterion on a given trial. One potential advantage of the present technique over the classical biofeedback approach is that provision of tangible rewards

may be much more motivating for subjects than the instruction to reach a target activation level in the absence of extrinsic reward. Another possible advantage of the present technique is that the use of instrumental conditioning instead of a visual biofeedback procedure may render the task much less 'cognitive' and thus less likely to require high level or effortful cognitive processing. Thus, the present technique may be efficacious even under situations when subjects are either incapable or unwilling to engage in effortful cognitive processing, or when a cognitively demanding task is concurrently imposed. Furthermore, the present technique may not even require subjective conscious awareness of task progress to be effective, given that instrumental conditioning procedures are known to work in a wide variety of animal species including rats, pigeons, and even aplysia [172, 173, 174], which one might speculate are unlikely to have developed conscious subjective awareness to the same degree as in humans. This raises the intriguing possibility that human brain regions may differ in the degree to which successful neural conditioning is associated with a subjective conscious correlate. Here, subjects reported using, and were instructed to use, a conscious strategy of imagining movement during task performance. Future studies could probe the subjective correlates of conditioning in different brain regions to examine whether, for example, subjective correlates of neural conditioning in higher cortical areas are qualitatively different than those associated with sub-cortical structures. Finally, the use of an approach based on instrumental conditioning means that we can benefit from the extensive work done in this area to inform our understanding of the neural and behavioral processes mediating this learning [163, 131, 125, 175, 176].

The task of differentially activating two motor cortical regions seemed to engage parallel learning processes: as the signal in the ROI being rewarded increased over time, we saw a corresponding decrease in the ROI not being rewarded. Subjects reported activating the rewarded ROI using kinesthetic motor imagery; however the signal decrease observed in the non-rewarded ROI may not be attributable to the same deliberate control. In order to continue earning rewards throughout the task, subjects had to increase the difference in signal between the two ROIs. Such differential neural sensitivity to the reward conditions may tap into covert associative learning mecha-

nisms over and above the explicit imagery strategy the subjects reported employing, as demonstrated in previous instrumental conditioning experiments [177, 178].

While not all functional imaging studies of motor imagery have reported activations in primary motor cortex (M1) [179, 180], several fMRI studies have shown evidence for somatotopically organized activations in primary motor cortex during motor imagery [170, 181]. We report here that activation in somatotopically specific regions of primary motor and sensory cortices increased over the course of conditioning. This enhancement could arguably be a side-effect of repeated practice of mental imagery, and not dependent on the reward feedback. However, it is difficult to explain the suppression in the non-rewarded ROI without the requirement that we imposed for differential activity in order to earn reward, suggesting that in our study provision of reward based on neural activity led to specific shaping of the neural response.

Nyberg et al. [182] compared the effects of mental practice to physical practice in a recent fMRI study. They found that practice in general led to a more regionally specific activation in motor cortex. They also found a differential increase in visual cortex activity in the mental practice group. Studies comparing kinesthetic and visual imagery have found that they evoke different patterns of neural activity [183, 184]. Since we found an increase in activity specific to sensorimotor cortex, perhaps the feedback from this area caused subjects to refine their imagery strategy to favor kinesthetic rather than visual. A similar effect was found in Yoo et al. [98], in which verbal feedback of auditory cortex activation was found to influence subjects' strategies during selective attention to auditory stimuli. Similarly, Posse et al. [97] gave subjects feedback of amygdala activation during sad mood induction, resulting in amygdala activations that correlated with sad mood. Generally speaking, training subjects to activate a particular part of their brain while performing a task could be a way of enhancing task performance or correcting deficits. Training subjects to make more efficient use of neural resources could potentially lead to long-term alterations in neural plasticity related to performance of specific tasks.

In summary, we have presented an instrumental conditioning technique which succeeds in shaping an increase in sensorimotor cortical responses over time, as measured

with fMRI. We have also used a behavioral measure to explore the effects of training on behavior. The method presented here extends previous work [97, 107, 96, 100, 98] by incorporating a well-studied operant conditioning paradigm with fMRI derived neurofeedback training. This method was successful in conditioning a differential response between two regions with a very high neuroanatomical precision– a finding that could have clear benefit in future clinical applications.

Acknowledgments

We would like to thank Mike Tzyska for invaluable assistance with technical setup, as well as Steven Flaherty and Ralph Lee. We also thank Masamichi Sakagami and Shinsuke Kobayashi for their input. This work was supported by a DANA foundation grand to J. O'Doherty and the JST ERATO Shimojo Implicit brain function project, as well as a Gordon and Betty Moore grant to the Caltech Brain Imaging Center.

Chapter 5

Direct instrumental conditioning of neural activity in orbitofrontal cortex

Functional neuroimaging studies have found correlations between activity in human orbitofrontal cortex (OFC; especially its medial aspect) and subjective ratings of reward value for a diverse range of stimuli including attractive faces. However, it is unclear from these studies whether orbitofrontal cortex activity has a causal influence on subjective evaluations of reward value, or is merely an epiphenomenon. To address this question, we used a real-time fMRI procedure involving instrumental conditioning with monetary reward in 13 male subjects who were trained to selectively increase activity in medial OFC, interleaved with binary attractiveness judgments on a set of female faces. Over several sessions subjects showed increased differential activity in OFC relative to a control condition during which they were instead conditioned to activate the hand area of motor cortex. After controlling for other factors such as whether a trial was rewarded or not, the differential OFC signal was found to be a significant predictor of increased attractiveness ratings, while differential activity in the hand-motor area was not. These results demonstrate that a neurofeedback procedure can be used to condition increased activity in OFC and that by selectively modulating activity levels in medial OFC it is possible to influence subjective judgments of attractiveness.

Introduction

Activity in medial OFC (mOFC) correlates with subjective ratings of reward value for a diverse range of stimuli, including liquid rewards [28], willingness to pay, [40] and attractive faces [30, 31]. Recent evidence also suggests that mOFC represents the experienced pleasantness of a stimulus, which can be influenced by factors other than the sensory properties of a stimulus, and the internal state of the subject [41]. While these studies have demonstrated correlations between subjective value and mOFC activity, they have not established whether mOFC has a causal influence on subjective judgments. Patients with OFC lesions are able to express a range of emotions [185], however they show impairment at relative valuation of sets of options [186]. OFC-lesioned patients are able to express preferences, but their preferences are internally inconsistent, pointing to a causal role for OFC in accessing stimulus value. However, it remains unclear whether elevated mOFC activity causes positive evaluations, or is merely epi-phenomenal. In this study we sought to test whether elevated mOFC activity can positively bias affective judgment of a concurrently presented stimulus.

To elicit reliable increases in mOFC activity on a trial-by-trial basis, we implemented a neural conditioning procedure [96, 100], in which subjects were rewarded for elevating mOFC activity upon being presented with a discriminative cue. Similar procedures have been employed in emotional brain regions such as rostral anterior cingulate cortex [100], and insula [108], but as of yet neural conditioning of mOFC has not been reported.

It is unclear what mental strategy is the most effective for elevating activity in the OFC. Based on prior reports of emotional recall activating OFC [187, 188], and the known responses of this region to both primary and abstract reinforcers [33, 189, 109], we instructed subjects that imagining things that they find personally rewarding would be a good initial strategy.

To test for effects of mOFC activity on affective judgments, subjects were asked to evaluate the attractiveness of a face at the end of each trial. Prior reports have shown that mOFC activity correlates with ratings of facial attractiveness [30, 31].

To control for general effects of elevating regional neural activity for reward, we included a second condition in which subjects were trained to elevate activity in a region of motor cortex related to hand movements, and received rewards contingent on successfully activating that region past a given threshold [160, 96].

Because the mOFC is important for subjective evaluations and decision making, the ability to train reliable increases in mOFC activity has important clinical implications for treating patients who show impaired decision-making abilities, such as depressed or addicted individuals. The results of this study also have important implications for understanding the precise influence of mOFC activity on expressed preferences.

Materials and Methods

Subjects

A total of 18 healthy right-handed male subjects aged 19 to 29 years (mean age 21 $\pm$ 2.3 years) participated in the experiment. All subjects gave informed consent, which was approved by the local research ethics committee. One additional subject did not complete the entire session.

Scan-to-scan motion can have a detrimental impact on learning, as subjects cannot tell whether changes in BOLD response are due to head motion or neural activity. We therefore eliminated subjects who showed large amounts of head motion. We summed the total scan-to-scan motion, estimated during SPM pre-processing (as described below) over all three directions in each session, and eliminated subjects who showed $>$ 30 mm of movement in more than one session. According to this criterion, five subjects were removed from the imaging analysis. This study included a behavioral response on each trial, as described below; one subject was excluded from the behavioral analysis for making the same response on every trial. This left 13 subjects in the imaging analysis and 12 in the behavioral analysis.

Stimuli and tasks

Visual stimuli were presented via a projector positioned at the back of the room. Subjects viewed a reflection of the projected image (800 x 600 pixels) in a mirror attached to the scanner head coil. Stimulus presentation and response recording were controlled with the Cogent 2000 toolbox in Matlab (Mathworks, Natick, MA).

Functional localizer tasks and ROI selection

Functional localizers for the three ROIs were run in two sessions, one for mOFC and one for hand-motor (HaM) and V5/MT, as described below. The order of the two localizer sessions was counterbalanced across subjects. Following the two functional localizer scans, subjects waited in the scanner for several minutes as the ROIs were selected.

OFC functional localizer: Probabilistic reversal learning

Evidence from both human and animal lesion and functional imaging work has shown that the mOFC is engaged during tasks that require subjects to keep track of varying stimulus values. One such task that has consistently been shown to recruit the mOFC is probabilistic reversal learning [189, 125, 20, 190]; thus we used this task as a functional localizer for parts of medial OFC sensitive to reward value. The implementation of this task was similar to Hampton et al. [190]. On each trial subjects were presented with the same two abstract fractal images, randomly assigned to the left or right side of a central fixation cross. These stimuli were presented for 2.9 s, during which time the subject was asked to choose between the two images, and press the left or right button on a button box held in their right hand (Current Designs, Philadelphia, PA), to choose the image on the left or right side of the screen. The chosen image then became brighter for 2.9 s, followed by feedback indicating whether the subject had won a quarter or lost a quarter for 2.9 s. The next trial immediately followed. Rewarding feedback was indicated with a picture of a US quarter in the center of the screen, while punishing feedback was indicated by a picture of a quarter

with a red X across the image. A running total of subjects' earnings during this task was presented above the quarter. Missed trials were indicated with a red X in the center of the screen and no change in the running total.

The images were randomly assigned to be the 'correct' or 'incorrect' choice. Choosing the 'correct' option was associated with the subsequent delivery of a monetary reward (gaining 0.25 USD) on 80% of trials, and a monetary punishment (losing 0.25 USD) on 20% of trials. Subjects were instructed to sample both choices in order to ascertain which was more rewarding (they are not told the exact probabilities, but merely that one image delivered rewards more often). The subjects were also instructed that sometimes the contingencies associated with the images would reverse, that is the image that delivered reward more often would begin to deliver less often and vice versa. Subjects were not informed of the specific details of the reversal probabilities, but contingency reversals would only occur after they demonstrated learning which was the 'correct' image, by choosing this image on 3 consecutive trials. Once this association had been acquired, the contingencies had a 1 in 4 probability of reversing on each subsequent trial. Subjects practiced this task for several minutes outside the scanner during the pre-training session. In the scanner, subjects performed a session that included 40 task trials with 20 null events (during which the fixation cross was presented for the duration of a normal trial) randomly interspersed, for a duration of ~8.5 min.

Within the same scan, subjects then saw a fixation cross at the center of the screen for 8.5 s, followed by the letters 'ImR' presented in the middle of the screen for 17 s. This was a cue for subjects to begin a period of reward imagery. Subjects were not given specific instructions about the contents of the imagery that they should use, but were simply asked to conjure imagery that they found personally rewarding. The imagery condition, alternating with the fixation cross, was presented 6 times.

mOFC ROI selection

A t-test was performed comparing reward to punishment scans, and the resulting statistical map was thresholded at $p<0.001$ and overlaid on the subjects' anatomical scan.

The p-value was iteratively increased or decreased until the lowest p-value was found that included some activity in the region of ventromedial prefrontal/orbitofrontal cortex (typically between $p<0.05$ and $p<0.001$). This image was used as a mask for a second contrast, comparing imagine-reward to rest scans. With this method we aimed to identify regions that were activated by both real and imagined rewards. An ROI center was chosen among voxels active in this masked contrast, near medial orbitofrontal cortex; the ROI extended 3 slices in the vertical direction and 6 x 6 voxels in plane.

Hand-motor and V5/MT localizer

The hand-motor and V5/MT localizer tasks were run in a single session. The stimuli presented to the subject consisted of a central fixation cross for 8.5 s, which then alternated with blocks of task specific stimuli, each presented for 14.5 s. All stimuli were low contrast, light gray presented on a darker gray background. The first three task-specific stimuli consisted of an array of 10,000 dots arranged in a circle at the center of the screen with a 100 pixel radius. In the first two blocks the dots moved outward from or inward to the center of the circle at a rate of 66 pixels/sec. The third stimulus was a similar array of dots, but not moving. The fourth stimulus was the letters 'ImM', which indicated to the subjects to imagine visual motion. That is, any kind of visual imagery of motion in the visual field, similar to the moving dot patterns. The fifth stimulus was the letters 'HaT', for 'hand tap'; here subjects were instructed to bend fingers II-V at the metacarpophalangeal joint at a rate of approximately 1 Hz. The sixth and final stimulus was the letters 'ImHaT', in response to which subjects were instructed to imagine the sensation of tapping their fingers as in the previous task, without actually moving. This series of stimuli cycled through five times. This task was practiced with a slightly shorter duration (3 cycles) outside the scanner during pre-training, where subjects could be observed at close range to ensure that they were not making real movements during the imagine-moving blocks.

Hand-motor ROI selection

A t-test was performed comparing reward to punishment scans, and the resulting statistical map was thresholded at $p<0.001$ and overlaid on the subjects' anatomical scan. The p-value was iteratively increased or decreased until the lowest p-value was found that included some activity in sensorimotor areas related to hand movements [169] (typically between $p<0.05$ and $p<0.001$). This image was then used as a mask for the contrast of imagine-hand motion vs. imagine-visual motion. With this method we aimed to identify brain regions responsive to both real and imagined hand movements. An ROI center was chosen near the hand region of sensorimotor cortex [169].

V5/MT ROI selection

A t-test was performed comparing reward to punishment scans, and the resulting statistical map was thresholded at $p<0.001$ and overlaid on the subjects' anatomical scan. The p-value was iteratively increased or decreased until the lowest p-value was found that included some activity near the ascending limb of the inferotemporal sulcus/lateral occipital sulcus [191] (typically between $p<0.05$ and $p<0.001$). This image was then used as a mask for the contrast of imagine-visual motion vs. imagine-hand motion. With this method we aimed to identify brain regions activated by both real and imagined visual motion. A region of interest was then chosen near the inferotemporal sulcus/lateral occipital sulcus [191].

Neural conditioning Task

Subjects were instructed that during this part of the experiment, they would be asked to activate specific brain regions on cue, using only their imagination. They were specifically instructed not to make any real motions while the scan was running. Two gray-colored shape cues, one triangle and one parallelogram, were assigned to either the mOFC-activate or HaM-activate condition, counterbalanced across subjects. A third hexagon shape with the word 'Rest' written in the center in white lettering, was assigned to the resting/baseline condition.

A reinforced conditioning trial is illustrated in Figure 5.1a. Each trial began with presentation of the rest-cue for a variable interval between 11.5 and 23 s, followed by either the mOFC-activate cue or the HaM-activate cue. During this time subjects were instructed to try to elevate activity in the target region, using only their imagination. Before scanning, subjects were told which target brain region was assigned to each cue.

Subjects were given ongoing feedback of neural activity with a thermometer-style graphic [108]. At the start of each block, below the shape cues, a white bar (50 x 150 pixels) was displayed with two black lines 25 pixels from the bottom and 25 pixels from the top. After two scans were acquired, the bar filled with a grey rectangle representing neural activity in the target area. The signal depicted the %-change from baseline to active in the background ROI (V5/MT) subtracted from the %-change from baseline to active in the target ROI. During rest blocks the bar appeared similarly, but the height of the bar fluctuated randomly. The signal was calibrated so that 100 pixels corresponded to the difference between zero and the current threshold. The bottom (zero) line indicated the level at which there was no difference between the target and background regions. Below this level, the signal appeared as a darkly colored bar. If the signal was greater than zero, it appeared as a medium grey bar between the bottom and top lines, and if the signal difference exceeded the current threshold, the medium grey bar extended up above the top line and the part above the line was colored in lighter grey. After a minimum of three scans (~9 s), if the signal exceeded the current threshold, the trial ended immediately and following the face attractiveness judgment, described below, the subject was presented with reward feedback. Reward feedback consisted of a picture of a dollar bill with the words 'You have earned ONE dollar'. If 23 s elapsed and the subject was not able to bring the activity level above threshold, the trial ended and the no-reward feedback was displayed. No-reward feedback consisted of a picture of a scrambled dollar and the words 'You have NOT earned one dollar'. Dollars earned during the task corresponded to real money paid to the subject at the end of the experiment.

Subjects were told that the 'rest' period preceding each 'active' period would

serve as a baseline against which the activity during the 'active' periods would be compared. Therefore, they should not practice mental imagery similar to during the 'active' periods. They were also told that in order to earn rewards they would have to activate the specific brain region being targeted in each condition. Subjects were told that any kind of mental imagery could be appropriate as long as it specifically activated brain regions delineated by finger tapping or reward, but that strategies involving motor imagery and reward imagery might be more likely to succeed, given the known functional responses of these regions. Subjects were told that the threshold defining the minimum activity required to get rewarded would be slowly increasing, therefore they would have to improve on their strategy in order to continue earning rewards.

The total duration of the experiment was approximately 2 h in a single session. In this time subjects performed pre-training, two functional localizer sessions, and 4 consecutive conditioning blocks with 12 trials in each; trials were ordered pseudorandomly so that each trial type appeared 6 times within a block without 3 consecutive trials being of the same type. Blocks were on average 8 minutes long, for a total of 32 minutes of training.

Attractiveness ratings

Upon completion of each activate block, but before the reward feedback was shown, subjects were asked to make a rapid subjective judgment. A single female face was presented at the center of the screen for 250ms, followed by the instruction: 'Please press the left button for below average attractive and the right button for above average attractive'. Subjects responded using a button pad in their right hand (Current Designs, Philadelphia PA). Faces were generated using computer software (FaceGen; Singular Inversions), and were all forward gazing with a neutral expression; 48 faces were used during the conditioning task. In order to get accustomed to responding to rapidly presented faces, subjects practiced responding to a set of 20 faces outside the scanner, during the pre-training session (this set of faces was different than the faces

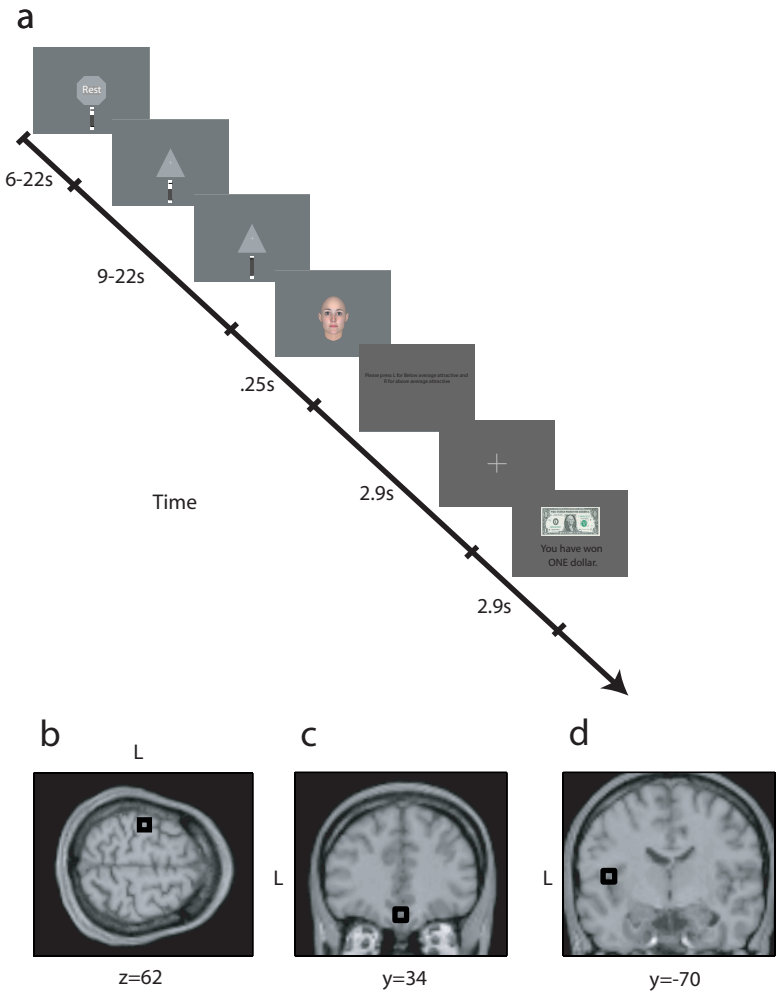

Figure 5.1: a) Illustration of a conditioning trial. Subjects are first presented with a resting cue for 6-22s, followed by a cue to activate one of the two target regions (subjects are informed prior to scanning of the cue-ROI assignment). The trial ends when the current threshold is exceeded or 22 s has elapsed, at which point subjects are presented with a face for 250 ms. They have 2.9 s to respond to the face, and finally receive reward or no-reward feedback. b) Square around mean center of subject ROIs for OFC region. c) Square around mean center of subject ROIs for HaM region. d) Square around mean center of subject ROIs for V5/MT region

used during the conditioning task).

We hypothesized that faces were more likely to be considered above average attractive during the OFC-activate trials, and that the signal level in OFC might impact on the attractiveness judgment. To test this we used a mixed effects logistic regression to model the effects of local signal levels on the decision to rate a face as above or below average attractive (function lmer, the R Foundation for Statistical Computing). Because OFC responses were indistinguishable between conditions in the fourth session, we restricted our analysis to the first three sessions. At the time that subjects were asked to evaluate a face, they were aware of whether or not the trial has been rewarded, thus reward can be a potential confounding influence; we therefore restricted this analysis to the trials which had been rewarded which included a total of 100 OFC-activate trials and 106 HAM-activate trials.

Post-experimental questionnaire

After leaving the scanner, subjects were asked to complete a short questionnaire. They were asked to briefly describe what they were thinking about during the imagine blocks of the functional localizer and the conditioning tasks, and to indicate if and how their strategy changed across sessions.

Motion Recordings

To control for subject motion during periods of imagined movement, we recorded EMG (BIOPAC, Goleta, CA) from the forearm (flexor digitorum superficialis muscle) to measure muscle activity related to finger flexion and extension. These data were recorded at 200 Hz. This recording device is MRI compatible, but fMRI scanning introduced noise into the recordings. Due to technical difficulties, recordings from 6 of the 13 subjects used for imaging are not usable, leaving 7 subjects. Single trials were individually inspected for signs of motion artifact and those trials removed from further analysis (0-6 per subject, mean 2).

fMRI scanning procedure

fMRI data were acquired on a Siemens AG (Erlangen, Germany) 3T TRIO MRI scanner; Blood Oxygenation Level Dependent (BOLD) contrast was measured with gradient echo T2* weighted echo-planar images (EPI). Imaging parameters were optimized to minimize signal dropout in medial ventral prefrontal and anterior ventral striatum: we used a tilted acquisition sequence at 30° to the AC-PC line [132], and an 8 channel phased array coil which yields a ~40% signal increase in this area over a standard coil. The first 3 volumes of each session were discarded to permit T1 equilibration. Other parameters were as follows: 49 slices, in-plane resolution, 3 x 3 mm; slice thickness, 3 mm; repetition time, 2.88 s; echo time, 30 ms; field of view, 192 x 192 mm. T1 and T2 weighted structural images were also acquired for each subject

Concurrent fMRI analysis and processing

As soon as images were reconstructed, they were transferred in real-time via TCP/IP socket to an external Intel Xeon workstation (3.8 MHz 64-bit processor running Redhat Linux); data processing was performed using MATLAB 7.0 (The Mathworks Inc., Natick, MA).

Pre-processing

Image pre-processing consisted of motion correction using AFNI [167], and linear detrending to correct for low-frequency scanner drift. During functional localizer scans spatial smoothing using a two-dimensional Gaussian of 5 mm width was performed prior to performing statistical tests. During the conditioning task no temporal or spatial smoothing was performed.

Online analysis

As images arrived on the external workstation, they were pre-processed and the signal was averaged over all voxels in the previously defined ROIs. Each trial began with a

variable length baseline period (6 - 23 s), followed by an 'activate' period. Starting from the third scan into each 'activate' block, the %-change from baseline to active was computed in both the target ROI (either HaM or mOFC) and the background region (V5/MT), and activity in the background subtracted from the target. This quantity was compared to the current threshold, and after a minimum of three scans the trial ended if the threshold was exceeded.

The threshold was updated according to a modified percentile reinforcement schedule [16]. The threshold started at 10^{-4}, for the first three trials in each condition. Following that, the threshold on each trial was set to the smallest value from the last three trials, so that the condition for reward on each trial was to improve on 1/3 of the previous responses. This rule was augmented with a shifting baseline parameter, so that if the lowest of the three most recent trials was below the baseline, the threshold value was instead set to the baseline value. When a trial was rewarded, the baseline increased to the smallest value in the entire trial history greater than the previous baseline. This ensured that the threshold did not significantly decrease across the experiment.

Group fMRI %-change analysis

In order to test whether regional signal levels increased over sessions as a function of condition, we performed a group analysis on the trial-by-trial %-change values measured during conditioning. We modeled the differential signal level (mOFC-V5 or HaM-V5) in a linear mixed effects model, with fixed effects of experimental session and condition, a session x condition interaction, and a random subject intercept (function lme, the R Foundation for Statistical Computing).

Post hoc SPM analysis

Data were pre-processed using the SPM5 software package (SPM5 `http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`). Images were corrected for slice timing and spatially realigned to the first image from the functional localizer. The EPI images

were coregistered to the T1-weighted anatomical scan. The T1 image was segmented into white and grey matter, and the grey matter was coregistered and normalized to the template grey matter image distributed with SPM5 (in Montreal Neurological Institute space). These parameters were subsequently applied to the T1 image itself as well as the set of EPI images. Spatial smoothing was then applied to the EPI images using a Gaussian kernel with full width at half maximum of 8 mm.

For each subject, we constructed a general linear model in SPM with all four conditioning sessions. We modeled trials from the mOFC-activate and HaM-activate conditions separately, face presentation for faces during rewarded trials and non-rewarded trials separately, and onset of reward feedback and no-reward feedback separately. The six ongoing motion parameters estimated during realignment were included as regressors of no interest. Linear contrast images from the single subject analyses were taken to the random effects level by applying t-tests between them to produce group statistical parametric maps.

Results

Post-experiment questionnaire

After the experiment, we asked subjects to complete a short questionnaire about their experience during the experiment. Specifically, we asked subjects what they were thinking about during the imagery portion of the functional localizer: 1) imagine visual motion, 2) imagine hand tapping, 3) imagine reward. During the imagine visual motion task, subjects reported imagining the moving dots or other moving patterns, moving in a vehicle or other things moving around them like baseballs and joggers. During the imagine hand tapping task, subjects reported using motor imagery including tapping, contracting muscles and squeezing motions. During the imagine reward task, subjects reported imagining monetary rewards, praise, compliments and erotic imagery.

We also asked them what they were imagining during the conditioning task in each

type of trial, and whether their strategies changed across the study. The strategies reported were generally similar to those used during the localizer tasks. During the HaM-activate condition, 5/12 subjects reported that they used the same imagery throughout, while 4/12 subjects reported that they had to imagine more intensely or different kinds of movements during later trials. During the mOFC-activate condition, 7/12 subjects reported having to imagine different scenarios in order to continue earning rewards while 2/12 just modulated the intensity of a particular scenario (e.g., increasing amounts of money), and 2/12 settled on one strategy after trying different things.

ROI locations

The mean ROI center (± SE) for the hand-motor region in Montreal Neurological Institute (MNI) space was [-34 ± 0.5, -17 ± 0.5, 62 ± 0.4]. This is similar to our previous study [-35 ± 1.24, -26 ± 1.9, 65 ± 0.941] [160], and consistent with reports of localization for hand-motor activity [169] and imagery [170]. The mean ROI center for the OFC was [0 ± 0.3, 34 ± 0.5, -19 ± 0.3]; previous probabilistic reversal learning studies have reported similar regions in reward-punishment contrasts, for example [12 36 -18] in [125] and [6 24 -24] in [189]. The mean ROI center for V5/MT was [-45 ± 0.4, -69 ± 0.5, 8.5 ± 0.7], consistent with [-38 -74 8] from [192] and [-47 -76 2] in [191]. ROIs are overlaid on a single subject's anatomical scan in Figure 5.1bcd.

Subject performance on conditioning task

All subjects were able to earn rewards during both tasks. The mean number of rewards (± SE) per session (out of 6 possible) for the mOFC-activate condition was [3.2 ± 0.4, 3.5 ± 0.3, 2.9 ± 0.3, 2.1 ± 0.3] and for the HaM-activate condition [4 ± 0.5, 4 ± 0.3, 2.7 ± 0.4, 3 ± 0.4]. A repeated measures ANOVA with between-subject factors of condition (2 levels) and session (4 levels) and their interaction, showed a significant main effect of session ($F(1,96) = 5.993$, $p<0.05$), reflecting a slight decrease in reward count across sessions, that did not depend on condition. We also constructed a linear

mixed effects model on the trial-by-trial reward threshold, with trial and condition as factors, and a random subject intercept. This analysis showed a linear increase in threshold across sessions, independent of condition (β = 1.496 x 10−4, t(608) = 6.39, p<0.0001). Thus, as the threshold for reward increased, the rate of reward slightly decreased.

Trial-by-trial %-change in regions of interest

We were interested in whether activity in the regions of interest showed a condition specific increase across sessions (Figure 5.2ab). Post-hoc plots of the differential mOFC-V5 signal, showed a sharp decline in the 4th session. We therefore constructed a linear model to test for an effect of learning across the first three sessions. We found a significant session x condition interaction (β = 0.174; t(442) = 2.9093315, p<0.05), indicating that the mOFC-V5 signal difference increased in a condition-specific manner across the first 3 sessions.

We also tested for a condition-specific increase in the HaM region (Figure 5.2cd). Using a similar model, the session x condition interaction does not reach significance across 3 (p = 0.4) or 4 (p = 0.19) sessions. However if we compare signal levels in the first session with their peak value in the third or fourth session, we find that signal in later sessions is significantly higher (paired t-test, t(12) = -1.8994, p<0.05 one-tailed). Restricting this analysis to those for which we have good EMG recordings, we see the same trend (paired t-test, t(6) = -1.798, p = 0.06 one-tailed). Despite some differences in training procedure, this result is highly consistent with the findings from our earlier study [160].

Post-hoc SPM analysis

We performed a post-hoc SPM analysis in order to validate the findings from our ROI-based analysis, and also to test for involvement of brain regions outside the target ROIs. We tested for regions with relatively stronger activity in each of the two activate conditions, across all four training sessions.

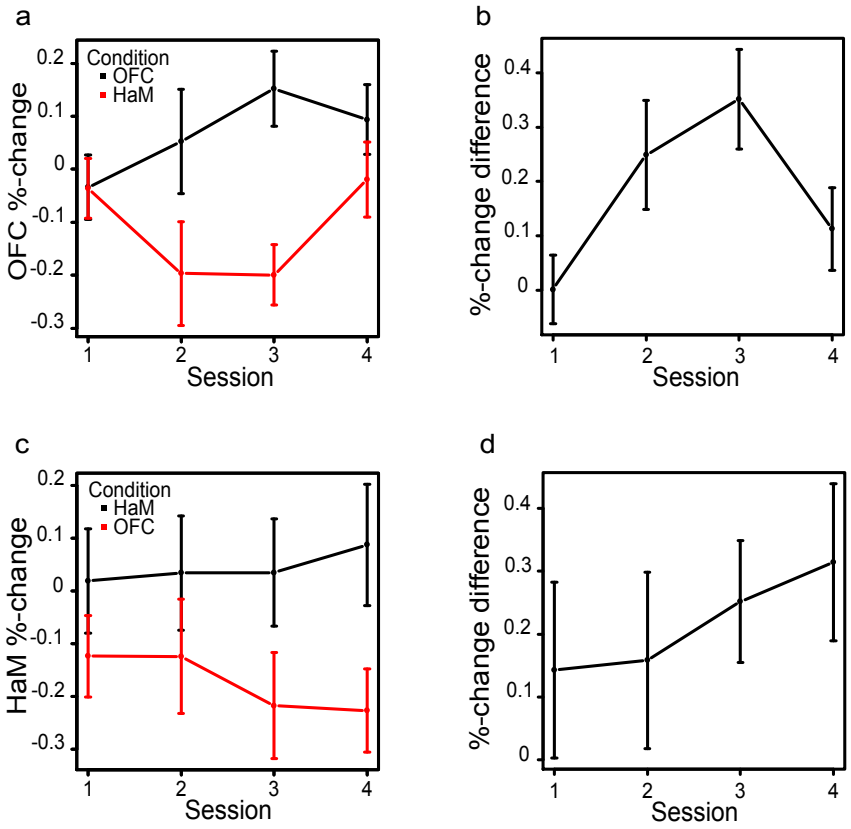

Figure 5.2: ROI percent-change averaged within condition in each session and across subjects. a) Averaged OFC-V5 differential signal in OFC and HaM activate conditions. b) Difference in OFC-V5 signal between OFC and HaM activate conditions, averaged over subjects within each session. c) Averaged HaM-V5 differential signal in HaM and OFC activate conditions. d) Difference in HaM-V5 signal between HaM and OFC activate conditions, averaged over subjects within each session

Table 5.1: Regions activated in OFC > HaM contrast

Regions activated in OFC activate > HaM activate contrast across 4 sessions, p>0.001.

Region	Contrast:OFC-activate>HaM-activate	
	voxels	Z
Superior frontal gyrus	333	4.81(-12 48 12)
Posterior cingulate	68	3.92(-3 -63 27)
Anterior cingulate cortex	14	3.54(9 30 -9)
Left superior frontal gyrus	7	3.45(-9 15 60)

For the contrast showing areas more active during the OFC-activate condition compared to the HAM-activate condition (Figure 5.3a), we find activity in anterior cingulate cortex, orbitofrontal cortex, and extending along the medial frontal gyrus to the superior frontal gyrus (see Table 5.1). Performing a small volume correction in an 8 mm sphere around the mean of subjects' ROI centers, we find peaks at [-3 30 -12], [-6 36 -15], and [0 27 -21] that survive FDR correction at p<0.05.

For the contrast showing areas more active during the HaM-activate condition compared to the OFC-activate condition (Figure 5.3b) across the four sessions, we find activity at the mean HaM ROI center at p<0.005. The peak voxel in this area is at [-27 -12 57]. Small volume correction around the mean ROI center shows activity at [-30 -15 57] surviving FDR correction in an 8 mm sphere around the mean HaM ROI center. This contrast also shows activity in bilateral post-central gyrus extending into inferior parietal lobule at p<0.001 (see Table 5.2).

Face attractiveness

We modeled the effect of the mOFC signal on attractiveness ratings in the subset of rewarded trials (from both conditions) and found that it significantly predicted a positive attractiveness judgment (β = 0.66, z = 2.255, p<0.05). Adding the HaM signal or the V5 signal to this model did not significantly improve the fit (HaM: $\chi^2(1)$ = 4.281, p = -.5 ;V5: $\chi^2(1)$ = 1.85, p=-0.17). Modeling the HaM and V5 signals on their own in separate models also did not result in significant predictors (HaM: z =

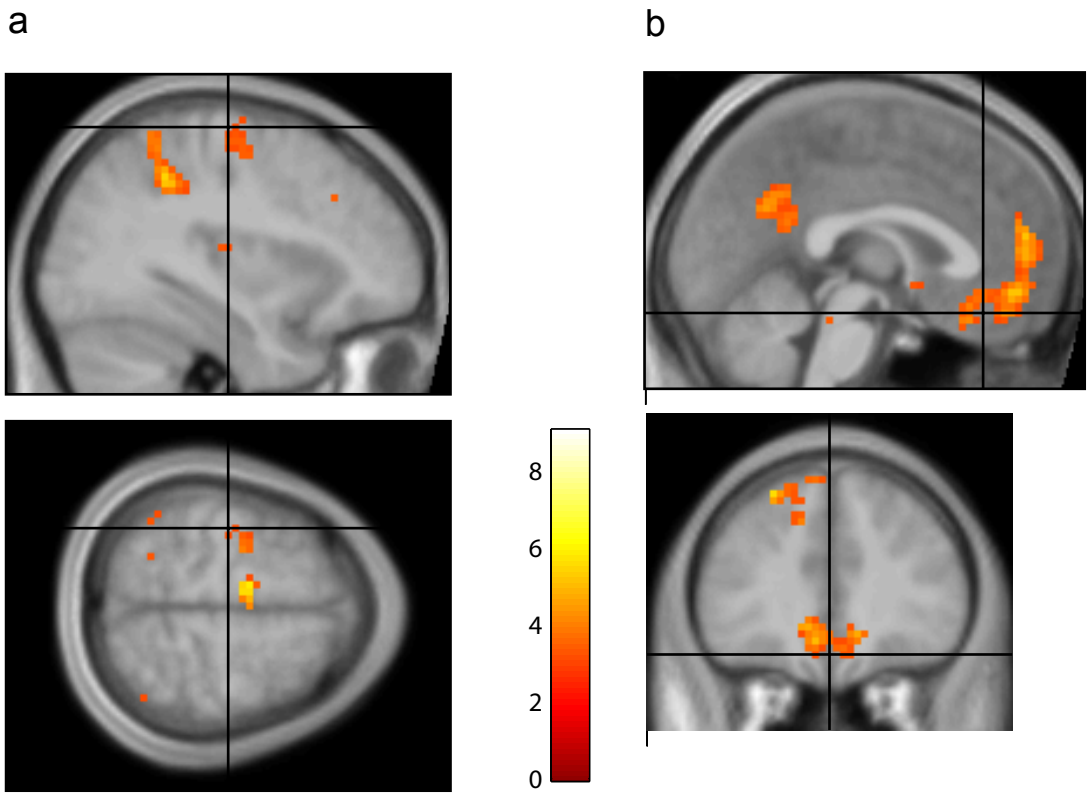

Figure 5.3: Post-hoc SPM analysis. a) Contrast showing regions more active in OFC activate condition over HaM activate condition across all four sessions, shown at $p<0.005$. Crosshairs centered at the mean of subjects OFC ROI centers. b) Contrast showing regions more active in HaM activate condition over OFC activate condition across all four sessions, shown at $p<0.005$. Crosshairs centered at the mean of subjects HaM ROI centers

Table 5.2: Regions activated in HaM activate > OFC contrast

Regions activated in HaM activate > OFC activate contrast across 4 sessions, p>0.001.

Region	Contrast:HaM-activate>OFC-activate	
	voxels	Z
Left inferior frontal gyrus/precentral gyrus	33	4.48(-60 9 33)
Left precentral gyrus	8	3.56(-27 -12 57)
	5	3.46(-60 3 12)
Left post-central gyrus/inferior parietal lobule	285	4.42(-60 -24 33)
Left inferior parietal lobule	11	3.61(-45 -48 57)
Right post-central gyrus/inferior parietal lobule	191	4.31(54 -30 39)
Right superior parietal lobule/precuneus	17	3.78(27 -60 51)
Right precentral gyrus	34	3.73(60 12 9)
Right inferior frontal gyrus	37	3.62(39 33 15)
	7	3.36(63 6 21)
Middle frontal gyrus	32	3.96(-3 -6 57)

1.14, p = 0.25; V5: z = -0.18, p = 0.86). This analysis suggests that mOFC activity is the most significant predictor of the propensity to judge a face as above average attractive for this subset of responses. This means that elevated mOFC activity is more likely to result in a face being rated as above average attractive than relatively lower activity levels, and that this influence is not shared by the other regions of interest that we recorded.

Discussion

In this study we have demonstrated that subjects can learn to voluntarily increase local activity in mOFC. This study represents an important extension of our earlier work using reward feedback to instrumentally condition neural activity. In order to validate this technique it was important to train a region that cannot be activated as easily by endogenously generated overt behaviors. Although we attempted in our previous study to control for real movements [160], motor cortex could potentially be activated by small movements not detectable by our MEG sensor, or by movement of muscles other than the one we monitored. However, the finding that mOFC can also

be conditioned by provision of rewards helps to reinforce the validity of this technique.

We have also demonstrated that mOFC activity affected subjective ratings of facial attractiveness. This finding dovetails nicely with studies showing that mOFC activity represents an integrated value measure, affected by endogenous and exogenous influences over and above sensory input [41].

We note that due to the design of this experiment, reward delivery may confound our estimates of the impact of the regional BOLD signals on attractiveness ratings. At the time that face judgments were made, subjects were aware whether the trial was going to be rewarded, and that could cause a more immediate effect on neural activity that we are unable to account for. Also, since the mOFC and HaM signals are stronger in rewarded trials, this could artificially inflate their impact on the attractiveness response. That said, since the influence of mOFC activity on attractiveness judgments was more consistent across subjects than the influence of HaM, we may infer that activity in this area plays a stronger role in affecting judgments.

It will also be important to test whether this influence is related to the role of mOFC in representing reward value specifically, as opposed to any kind of subjective judgment task. Although it would be impossible to test all possible judgment tasks, we can compare these results with those from a control group who perform a nearly identical paradigm, but who are asked to make a subjective evaluation of the faces that does not depend on differential reward value, for example we could ask subjects whether the faces are above or below average roundness. We predict that this judgment would not be affected by trial-by-trial variation in mOFC signal.

It is impressive that we observed this mOFC-behavior correlation considering the relatively long acquisition time (TR) for each image (2.88 s) and the inherent lag in the BOLD response (~2-6 s behind peak of neural activity). In this study, subjects began to attempt to elevate regional activity upon presentation of the discriminative cue. The trial ended when recordings of the BOLD signal showed that the threshold for differential signal level had been reached. Unfortunately, due to the lag in the hemodynamic response, the time at which the face stimulus was presented probably does not represent the peak in underlying mOFC activity. Although the lag in hemo-

dynamic response is an unavoidable issue in fMRI studies, it might be interesting to perform a similar study with a reduced number of slices centered on OFC, effectively sacrificing brain coverage to gain some temporal precision.

It is notable that the strategies used to continue regulating mOFC and HaM activity over the course of 4 conditioning sessions differed between brain regions. All subjects reported using strategies related to motor activity in order to activate the hand region of motor-cortex. In this condition, variation in strategy was largely a matter of making imagery more intense or vivid, or thinking about different kinds of motor tasks. Conversely, subjects used a wider range of strategies for the OFC-activate task, which included imagining financial rewards, praise, erotic imagery, and food. While this is undoubtedly related to the relative vagueness of the initial instruction, it also reflects the heterogeneous nature of rewards and reward processing in the mOFC. Several subjects also reported that they frequently had to change strategy in order to continue earning reward. This finding has important implications for future implementations of mOFC conditioning: if subjects must continuously change strategies in order to continue earning reward, can this activity be sustainable in the longer term? In this study there was of course the confounding effect that strategies that produced reward would also lead to an increase in the reward threshold. In order to develop a robust procedure for conditioning mOFC activity it will be important to test whether repeatedly imagining the same rewarding scenario, with and without reward feedback, generates diminishing levels of mOFC activity.

Potentially related to the issue of repeatedly using the same strategy, in the mOFC-activate condition we observed an initial period of learning over the first 3 sessions, followed by a sharp decrease in differential activity in the fourth session. We did not observe a similar pattern in the HaM-activate condition. There are several reasons why this might have been the case. The first is that subjects are simply tired by the last session, and that mOFC is more sensitive than HaM to fatigue. Another possibility is that rewards become less valuable and mOFC habituates more rapidly to receipt of reward. We did not explicitly ask subjects whether they found one condition more difficult than the other, however we did observe a trend towards earning

more rewards in the HaM condition, indicating that subjects likely found the mOFC condition more difficult.

In our previous neural conditioning study [160], we used reward feedback to reinforce performance of the neural response. However, following some initial piloting, we decided to supplement this with ongoing feedback of differential activity. We cannot quantify the impact that the addition of ongoing feedback (updated at ~3 s lags) had on subjects' performance. While graphical methods tend to be favored, there is no published evidence that this is the most effective; some groups are attempting to quantify the utility of different feedback methods [193]. Methods for delivering feedback have included graphical visual [107, 99, 106, 108], auditory [103], and visual reward [160]. In general, feedback modality can also be largely dependent on the goal of training: graphical feedback may be inappropriate if subjects are concurrently engaged in attention-demanding tasks. It will be important to test that the feedback we used contributed significantly to learned improvement, and rule out the possibility that the increase in mOFC activity occurred as an effect of repeated practice. This can be tested, as in our previous study [160], by scanning a group of subjects whose feedback and rewards are yoked to those from a previous subject.

An intriguing question for future study is whether it is possible to down-regulate activity in mOFC in order to earn reward. This region is intimately involved in reward representation and expectation. A functional dissociation between mOFC activity and reward would have implications for causally linking mOFC activity with positive reward value. It would also be interesting to test the behavioral impact of down-regulating OFC. Extrapolating on the present findings, we would expect that down-regulating OFC should make subjects less likely to rate a face as attractive.

In general the technique of using feedback training to regulate local brain activity combined with testing of behavioral responding has the potential to complement existing techniques for establishing the causal influence of regional brain activity, such as TMS [91] and lesion studies [194]. Studies similar to the one presented here could be used to probe the precise functional impact of varying levels of regional activity. In this study we used a binary response so that subjects could respond

quickly without excessive cognitive deliberation; however, allowing more variability in response might show an even stronger correlation.

In this study we have shown that with feedback of regional BOLD activity subjects can learn to voluntarily increase activity in mOFC, and that elevated mOFC activity influences a subjective judgment. This work has important implications both for clinical applications of regulating mOFC activity and for our understanding of how mOFC activity can influence subjective judgments.

Chapter 6

Summary

In recent years functional MRI studies of human reward learning have significantly advanced our understanding of how the brain represents rewards and learns reward associations. The studies presented in this thesis build on this work to further characterize the functional contributions of regions such as the orbitofrontal cortex and ventral striatum, with a focus on understanding how neural activity relates to behavior, specifically in terms of valuation and decision-making.

Attractive faces have been shown to be a form of visual reward, suggesting that they should affect behavior and neural activity in a manner similar to other types of reinforcers. In Chapter 1 we tested this hypothesis and demonstrated that attractive faces can act as reinforcers in a classical conditioning paradigm. The affective pleasantness of a set of neutral visual cues increased as a result of repeated pairings with attractive, compared to unattractive, female faces. We found that reward prediction errors in the ventral striatum were engaged during learning, as has been found for other types of reinforcers such as food, pain, and money.

The change in valuation for cues paired with attractive female faces was especially pronounced in male subjects, while female subjects did not show a similar effect in response to male faces. Interestingly, in male subjects prediction error responses were strongest for female faces, and prediction error responses in female subjects were strongest for male faces. This suggests that learning takes place similarly in the brains of male and female subjects, but is expressed differently at the behavioral level. An avenue for future study would be to employ different behavioral probes to

investigate if and how female subjects express this learning.

More generally the results of this study are relevant for marketing studies, which have shown that the presence of an attractive female model in an advertisement can influence customer perception of a product [101, 102]. Our findings suggest that classical conditioning mechanisms may contribute to this effect.

Pavlovian cues elicit passive responses but can also exert control over instrumental responding. In Chapter 2 we presented the first investigation into the neural mechanisms by which Pavlovian cues exert control over human decision-making. We showed that a Pavlovian cue predictive of a specific liquid reward can bias action choice towards responses associated with the same liquid reward. We found that a region of ventrolateral putamen was relatively suppressed when subjects made choices incompatible with the Pavlovian cue. While lesion studies in animals have shown that regions of ventral striatum are necessary for the expression of Pavlovian-instrumental transfer effects [83], this study is the first to show the dynamics of neural activity involved in outcome-specific transfer.

Current theories propose that transfer mechanisms are mediated by stimulus-outcome and outcome-response associations [79]. Our results fit nicely with this theory: we interpret our finding of a relative suppression when an incompatible cue is chosen as related to the suppression of an outcome-response association stimulated by the Pavlovian cue.

We note that the regions we found to be involved in outcome-specific transfer are distinct from those found in a recent fMRI study on general transfer effects, in which a Pavlovian cue enhances response vigor rather than influencing decision-making per se [86]; this mirrors the dissociation in neural circuitry found in animal studies of general and specific transfer effects [83, 84, 85]. However, it will be important in future studies to demonstrate both general and specific transfer effects in the same paradigm.

Several interesting features of Pavlovian-instrumental paradigms have been identified in the animal literature, of particular interest is the effect of reinforcer devaluation. It has been shown that in certain situations devaluing the reinforcer associated

with the Pavlovian cue does not suppress the expression of transfer effects in animals [148, 79]. This finding has clear parallels with addictive behaviors, in which environmental cues trigger drug-seeking, even when the outcome has known aversive effects. Few studies have directly probed the links between transfer effects and addictive behaviors [150], but this line of research could prove important in understanding the neurophysiological underpinnings of addiction. One potential avenue for treatment could involve training subjects to suppress regional activity in order to successfully avoid making choices associated with environmental cues [100, 160].

A further extension of the work presented in chapters 1 and 2, related to the impact of attractive faces used in advertising, would be to test whether cues associated with attractive faces can exert control over instrumental behaviors, as has been shown with other types of reinforcers [150, 86].

In Chapters 3 and 4 we investigate how provision of reward can influence neural plasticity: we trained human subjects to activate specific brain regions in order to earn reward. We demonstrated that a shaping procedure in which subjects were given monetary rewards for making improvements on their past performance was successful in training an increase in differential activity across sessions. This technique presents an alternative to standard bio/neurofeedback approaches and may prove useful in many clinical and research applications.

In the study described in Chapter 3 we successfully trained subjects to differentially activate regions of motor cortex related to hand and foot movements, in absence of overt movements. We investigated behavioral effects of this learning, and showed that reaction times in a cued response task were differentially affected by presentation of the learned cues.

A primary motivation for developing this technique was to condition neural activity in emotional brain regions, in order to study the causal effects of elevated activity on behavior. In the study described in Chapter 4 we trained subjects to activate medial orbitofrontal cortex (mOFC) activity and probed the impact of this training on an affective judgment task. We demonstrated that subjects can improve at elevating mOFC activity on cue, and that elevated activity was associated with a positive bias

in affective evaluations. This study represents a significant advance in our understanding of how mOFC activity affects our perception of value, as previous imaging studies have been unable to establish this causal link.

Taken together, these studies advance our understanding of the functional contributions of ventral striatum and orbitofrontal cortex in influencing decision-making and valuation, and suggest that applying associative learning techniques to real-time fMRI training can be a powerful method for characterizing the causal influence of regional neural activity.

Bibliography

[1] I. Pavlov, "Conditioned refelxes: An investigation of the physiological activity of the cerebral cortex," 1927.

[2] E. Thorndike, *Animal Intelligence*. New York: Macmillan, 1911.

[3] B. F. Skinner, *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.

[4] B. F. Skinner, *Science and human behavior*. New York: MacMillan, 1953.

[5] F. D. Sheffield, J. J. Wulff, and R. Backer, "Reward value of copulation without sex drive reduction," *Journal of comparative and physiological psychology*, vol. 44, pp. 3–8, 1951.

[6] D. A. Lieberman, *Learning: Behavior and Cognition*. Stamford CT: Wadsworth, 3 ed., 2000.

[7] F. Etscorn and R. Stephens, "Establishment of conditioned taste aversions with a 24-hour cs-us interval," *Physiological Psychology*, vol. 1, no. 3, pp. 251–253, 1973.

[8] R. A. Rescorla, "Predictability and number of pairings in pavlovian fear conditioning," *Psychonomic Science*, vol. 4, no. 11, pp. 383–384, 1966.

[9] L. Kamin, "Predictability, surprise, attention, and conditioning," in *Punishment and aversive behavior* (C. R. Campbell, BA, ed.), New York: Appleton-Century-Crofts, 1969.

[10] R. A. Rescorla and A. R. Wagner, "A theory of pavlovian conditioning: Variations in effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning II: Current research and theory* (A. H. Black and W. F. Prokasy, eds.), New York: Appleton Century-Crofts, 1972.

[11] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning*, vol. 3, pp. 9–44, 1988.

[12] E. R. Siquelan and C. A. Delucia, "Visual reinforcement of nonnutritive sucking in human infants," *Science*, vol. 165, no. 3898, p. 1144, 1969. Times Cited: 87.

[13] R. A. Butler, "Incentive conditions which influence visual exploration," *Journal of Experimental Psychology*, vol. 48, no. 1, pp. 19–23, 1954.

[14] R. J. Lamb, A. R. Morral, G. Galbicka, K. C. Kirby, and M. Y. Iguchi, "Shaping reduced smoking in smokers without cessation plans," *Experimental and Clinical Psychopharmacology*, vol. 13, no. 2, pp. 83–92, 2005.

[15] E. S. Athens, T. R. Vollmer, and C. Pipkin, "Shaping academic task engagement with percentile schedules," *Journal of Applied Behavior Analysis*, vol. 40, no. 3, pp. 475–488, 2007.

[16] G. Galbicka, "Shaping in the 21st century: moving percentile schedules into applied settings," *Journal of Applied Behavior Analysis*, vol. 27, no. 4, pp. 739–760, 1994.

[17] D. Hebb, *The organization of behavior*. New York: Wiley, 1949.

[18] G. Q. Bi and M. M. Poo, "Synaptic modification by correlated activity: Hebb's postulate revisited," *Annual Review of Neuroscience*, vol. 24, pp. 139–166, 2001.

[19] R. Dias, T. W. Robbins, and A. C. Roberts, "Dissociation in prefrontal cortex of affective and attentional shifts," *Nature*, vol. 380, no. 6569, pp. 69–72, 1996.

[20] J. Hornak, J. O'Doherty, J. Bramham, E. T. Rolls, R. G. Morris, P. R. Bullock, and C. E. Polkey, "Reward-related reversal learning after surgical excisions in

orbito-frontal or dorsolateral prefrontal cortex in humans," *Journal of Cognitive Neuroscience*, vol. 16, no. 3, pp. 463–478, 2004.

[21] J. L. Price, "Olfactory system," in *The Human Nervous System* (G. Paxinos, ed.), San Diego: Academic, 1990.

[22] J. L. Price, "Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions," 2007.

[23] L. L. Baylis, E. T. Rolls, and G. C. Baylis, "Afferent connections of the caudolateral orbitofrontal cortex taste area of the primate," *Neuroscience*, vol. 64, no. 3, pp. 801–812, 1995.

[24] E. T. Rolls and L. L. Baylis, "Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex," *Journal of Neuroscience*, vol. 14, no. 9, pp. 5437–5452, 1994.

[25] B. J. Rolls, E. T. Rolls, E. A. Rowe, and K. Sweeney, "Sensory specific satiety in man," *Physiology and Behavior*, vol. 27, no. 1, pp. 137–142, 1981.

[26] H. D. Critchley and E. T. Rolls, "Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex," *Journal of Neurophysiology*, vol. 75, no. 4, pp. 1673–1686, 1996.

[27] E. T. Rolls, Z. J. Sienkiewicz, and S. Yaxley, "Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey," *European Journal of Neuroscience*, vol. 1, no. 1, pp. 53–60, 1989.

[28] J. O'Doherty, E. T. Rolls, S. Francis, R. Bowtell, and F. McGlone, "Representation of pleasant and aversive taste in the human brain," *Journal of Neurophysiology*, vol. 85, no. 3, pp. 1315–1321, 2001.

[29] E. T. Rolls, M. L. Kringelbach, and I. E. T. de Araujo, "Different representations of pleasant and unpleasant odours in the human brain," *European Journal of Neuroscience*, vol. 18, no. 3, pp. 695–703, 2003. Times Cited: 83.

[30] I. Aharon, N. Etcoff, D. Ariely, C. F. Chabris, E. O'Connor, and H. C. Breiter, "Beautiful faces have variable reward value: fmri and behavioral evidence," *Neuron*, vol. 32, pp. 537–551, 2001.

[31] J. O'Doherty, J. Winston, H. Critchley, D. Perrett, D. M. Burt, and R. J. Dolan, "Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness," *Neuropsychologia*, vol. 41, pp. 147–155, 2003.

[32] A. J. Blood and R. J. Zatorre, "Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 20, pp. 11818–11823, 2001.

[33] J. O'Doherty, E. T. Rolls, S. Francis, R. Bowtell, F. McGlone, G. Kobal, B. Renner, and G. Ahne, "Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex,"

[34] R. Elliott, C. D. Frith, and R. J. Dolan, "Differential neural response to positive and negative feedback in planning and guessing tasks," *Neuropsychologia*, vol. 35, no. 10, pp. 1395–1404, 1997.

[35] R. Elliott, J. L. Newman, O. A. Longe, and J. F. W. Deakin, "Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: A parametric functional magnetic resonance imaging study," *Journal of Neuroscience*, vol. 23, no. 1, pp. 303–307, 2003.

[36] 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*, vol. 30, no. 2, pp. 619–639, 2001.

[37] J. P. O'Doherty, R. Deichmann, H. D. Critchley, and R. J. Dolan, "Neural responses during anticipation of a primary taste reward," *Neuron*, vol. 33, no. 5, pp. 815–826, 2002.

[38] J. A. Gottfried, J. O'Doherty, and R. J. Dolan, "Encoding predictive reward value in human amygdala and orbitofrontal cortex," *Science*, vol. 301, no. 5636, pp. 1104–1107, 2003.

[39] D. M. Small, R. J. Zatorre, A. Dagher, A. C. Evans, and M. Jones-Gotman, "Changes in brain activity related to eating chocolate - from pleasure to aversion," *Brain*, vol. 124, pp. 1720–1733, 2001. Part 9.

[40] H. Plassmann, J. O'Doherty, and A. Rangel, "Orbitofrontal cortex encodes willingness to pay in everyday economic transactions," *Journal of Neuroscience*, vol. 27, no. 37, pp. 9984–9988, 2007. Times Cited: 1.

[41] H. Plassmann, J. O'Doherty, B. Shiv, and A. Rangel, "Marketing actions can modulate neural representations of experienced pleasantness," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, pp. 1050–1054, 2008. Times Cited: 1.

[42] J. E. LeDoux, "Emotion circuits in the brain," *Annual Review of Neuroscience*, vol. 23, pp. 155–184, 2000.

[43] S. Maren and M. S. Fanselow, "The amygdala and fear conditioning: Has the nut been cracked?," *Neuron*, vol. 16, no. 2, pp. 237–240, 1996.

[44] L. M. Romanski and J. E. Ledoux, "Equipotentiality of thalamoamygdala and thalamocorticoamygdala circuits in auditory fear conditioning," *Journal of Neuroscience*, vol. 12, no. 11, pp. 4501–4509, 1992.

[45] S. Maren and M. S. Fanselow, "Synaptic plasticity in the basolateral amygdala induced by hippocampal-formation stimulation in-vivo," *Journal of Neuroscience*, vol. 15, no. 11, pp. 7548–7564, 1995.

[46] J. Hitchcock and M. Davis, "Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm," *Behavioral Neuroscience*, vol. 100, no. 1, p. 11, 1986.

[47] L. D. Vandekar, R. A. Piechowski, P. A. Rittenhouse, and T. S. Gray, "Amygdaloid-lesions - differential effect on conditioned stress and immobilization-induced increases in corticosterone and renin secretion," *Neuroendocrinology*, vol. 54, no. 2, pp. 89–95, 1991.

[48] A. Bechara, D. Tranel, H. Damasio, R. Adolphs, C. Rockland, and A. R. Damasio, "Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans," *Science*, vol. 269, no. 5227, pp. 1115–1118, 1995.

[49] J. J. Paton, M. A. Belova, S. E. Morrison, and C. D. Salzman, "The primate amygdala represents the positive and negative value of visual stimuli during learning," *Nature*, vol. 439, no. 7078, pp. 865–870, 2006.

[50] H. C. Breiter, N. L. Etcoff, P. J. Whalen, W. A. Kennedy, S. L. Rauch, R. L. Buckner, M. M. Strauss, S. E. Hyman, and B. R. Rosen, "Response and habituation of the human amygdala during visual processing of facial expression," *Neuron*, vol. 17, no. 5, pp. 875–887, 1996.

[51] I. S. Johnsrude, A. M. Owen, N. M. White, W. V. Zhao, and V. Bohbot, "Impaired preference conditioning after anterior temporal lobe resection in humans," *Journal of Neuroscience*, vol. 20, no. 7, pp. 2649–2656, 2000.

[52] C. Buchel, J. Morris, R. J. Dolan, and K. J. Friston, "Brain systems mediating aversive conditioning: an event-related fmri study," *Neuron*, vol. 20, no. 5, pp. 947–957, 1998.

[53] W. Li, J. D. Howard, T. B. Parrish, and J. A. Gottfried, "Aversive learning enhances perceptual and cortical discrimination of indiscriminable odor cues," *Science*, vol. 319, no. 5871, pp. 1842–1845, 2008.

[54] J. Olds and P. Milner, "Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain," *Journal of Comparative and Physiological Psychology*, vol. 47, no. 6, pp. 419–427, 1954.

[55] B. G. Hoebel, A. P. Monaco, L. Hernandez, E. F. Aulisi, B. G. Stanley, and L. Lenard, "Self-injection of amphetamine directly into the brain," *Psychopharmacology*, vol. 81, no. 2, pp. 158–163, 1983.

[56] W. A. Carlezon, D. P. Devine, and R. A. Wise, "Habit-forming actions of nomifensine in nucleus-accumbens," *Psychopharmacology*, vol. 122, no. 2, pp. 194–197, 1995.

[57] R. A. Wise, "Brain reward circuitry: Insights from unsensed incentives," *Neuron*, vol. 36, no. 2, pp. 229–240, 2002.

[58] W. Schultz, "Getting formal with dopamine and reward," *Neuron*, vol. 36, no. 2, pp. 241–263, 2002.

[59] R. Romo and W. Schultz, "Dopamine neurons of the monkey midbrain - contingencies of responses to active touch during self-initiated arm movements," *Journal of Neurophysiology*, vol. 63, no. 3, pp. 592–606, 1990.

[60] W. Schultz and R. Romo, "Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey," *Journal of Neurophysiology*, vol. 57, no. 1, pp. 201–217, 1987.

[61] W. Schultz and R. Romo, "Dopamine neurons of the monkey midbrain - contingencies of responses to stimuli eliciting immediate behavioral reactions," *Journal of Neurophysiology*, vol. 63, no. 3, pp. 607–624, 1990.

[62] J. Mirenowicz and W. Schultz, "Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli," *Nature*, vol. 379, no. 6564, pp. 449–451, 1996.

[63] J. Mirenowicz and W. Schultz, "Importance of unpredictability for reward responses in primate dopamine neurons," *Journal of Neurophysiology*, vol. 72, no. 2, pp. 1024–1027, 1994.

[64] J. R. Hollerman and W. Schultz, "Dopamine neurons report an error in the temporal prediction of reward during learning," *Nature Neuroscience*, vol. 1, pp. 304–309, 1998.

[65] I. Martin and A. B. Levey, "Blocking observed in human eyelid conditioning," *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, vol. 43, no. 3, pp. 233–256, 1991.

[66] P. Waelti, A. Dickinson, and W. Schultz, "Dopamine responses comply with basic assumptions of formal learning theory," *Nature*, vol. 412, no. 6842, pp. 43–48, 2001.

[67] R. S. Burns, C. C. Chiueh, S. P. Markey, M. H. Ebert, D. M. Jacobowitz, and I. J. Kopin, "A primate model of parkinsonism - selective destruction of dopaminergic-neurons in the pars compacta of the substantia nigra by n-methyl-4-phenyl-1,2,3,6-tetrahydropyridine," *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, vol. 80, no. 14, pp. 4546–4550, 1983.

[68] A. M. J. Young, M. H. Joseph, and J. A. Gray, "Increased dopamine release invivo in nucleus-accumbens and caudate-nucleus of the rat during drinking - a microdialysis study," *Neuroscience*, vol. 48, no. 4, pp. 871–876, 1992.

[69] G. S. Berns, S. M. McClure, G. Pagnoni, and P. R. Montague, "Predictability modulates human brain response to reward," *Journal of Neuroscience*, vol. 21, no. 8, pp. 2793–2798, 2001.

[70] G. Pagnoni, C. F. Zink, P. R. Montague, and G. S. Berns, "Activity in human ventral striatum locked to errors of reward prediction," *Nature Neuroscience*, vol. 5, no. 2, pp. 97–98, 2002.

[71] J. P. O'Doherty, P. Dayan, K. Friston, H. Critchley, and R. J. Dolan, "Temporal difference models and reward-related learning in the human brain," *Neuron*, vol. 38, pp. 329–337, 2003.

[72] B. Seymour, N. Daw, P. Dayan, T. Singer, and R. Dolan, "Differential encoding of losses and gains in the human striatum," *Journal of Neuroscience*, vol. 27, no. 18, pp. 4826–4831, 2007.

[73] P. N. Tobler, J. P. O'Doherty, R. J. Dolan, and W. Schultz, "Human neural learning depends on reward prediction errors in the blocking paradigm," *Journal of Neurophysiology*, vol. 95, no. 1, pp. 301–310, 2006.

[74] M. R. Delgado, L. E. Nystrom, C. Fissell, D. C. Noll, and J. A. Fiez, "Tracking the hemodynamic responses to reward and punishment in the striatum," *Journal of Neurophysiology*, vol. 84, no. 6, pp. 3072–3077, 2000.

[75] C. F. Zink, Y. X. Tong, Q. Chen, D. S. Bassett, J. L. Stein, and A. Meyer-Lindenberg, "Know your place: Neural processing of social hierarchy in humans," *Neuron*, vol. 85, no. 2, pp. 273–283, 2008.

[76] B. Wansink, M. M. Cheney, and N. Chan, "Exploring comfort food preferences across age and gender," *Physiology and Behavior*, vol. 79, no. 4-5, pp. 739–747, 2003.

[77] W. K. Estes, "Discriminative conditioning i a discriminative property of conditioned anticipation," *Journal of Experimental Psychology*, vol. 32, no. 2, pp. 150–155, 1943.

[78] W. K. Estes, "Discriminative conditioning 2 effects of a pavlovian conditioned stimulus upon a subsequently established operant response," *Journal of Experimental Psychology*, vol. 38, no. 2, pp. 173–177, 1948.

[79] R. A. Rescorla, "Transfer of instrumental control mediated by a devalued outcome," *Animal Learning and Behavior*, vol. 22, no. 1, pp. 27–33, 1994.

[80] P. Blundell, G. Hall, and S. Killcross, "Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats," *Journal of Neuroscience*, vol. 21, no. 22, pp. 9018–9026, 2001.

[81] G. Bower and R. Kaufman, "Transfer across drives of discriminative effect of a pavlovian conditioned stimulus," *Journal of the Experimental Analysis of Behavior*, vol. 6, no. 3, p. 445, 1963.

[82] B. Balleine, "Asymmetrical interactions between thirst and hunger in pavlovian-instrumental transfer," *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, vol. 47, no. 2, pp. 211–231, 1994.

[83] L. H. Corbit, J. L. Muir, and B. W. Balleine, "The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell," *Journal of Neuroscience*, vol. 21, no. 9, pp. 3251–3260, 2001.

[84] L. H. Corbit and B. W. Balleine, "Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer," *Journal of Neuroscience*, vol. 25, no. 4, pp. 962–970, 2005.

[85] L. H. Corbit and P. H. Janak, "Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of pavlovian stimuli on instrumental responding," *Journal of Neuroscience*, vol. 27, no. 51, pp. 13977–13981, 2007.

[86] D. Talmi, B. Seymour, P. Dayan, and R. J. Dolan, "Human pavlovian-instrumental transfer," *Journal of Neuroscience*, vol. 28, no. 2, pp. 360–368, 2008.

[87] E. E. Fetz, "Operant conditioning of cortical unit activity," *Science*, vol. 163, no. 3870, p. 955, 1969.

[88] J. V. Basmajian, "Control and training of individual motor units," *Science*, vol. 141, no. 357, p. 440, 1963.

[89] J. Olds and M. Olds, "Inference and learning in paleocortical systems," in *Brain mechanisms and Learning* (A. Fessard, R. Gerard, and J. Konorski, eds.), Oxford: Blackwell Scientific Publications, 1961.

[90] E. B. Fetz and FinocchiDv, "Operant conditioning of specific patterns of neural and muscular activity," *Science*, vol. 174, no. 4007, p. 431, 1971. Times Cited: 66.

[91] A. Pascual-Leone, D. Bartres-Faz, and J. P. Keenan, "Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'."

[92] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (ttd) for completely paralyzed patients," *Ieee Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 190–193, 2000.

[93] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An eeg-based brain-computer interface for cursor control," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.

[94] S. Musallam, B. D. C. BD, B. Greger, H. Scherberger, and R. A. Andersen, "Cognitive control signals for neural prosthetics," *Science*, vol. 305, no. 5681, pp. 258–262, 2004.

[95] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz, "Cortical control of a prosthetic arm for self-feeding," *Nature*, 2008. 0028-0836 101038/nature06996 101038/nature06996.

[96] R. C. deCharms, K. Christoff, G. H. Glover, J. M. Pauly, S. Whitfield, and J. D. E. Gabrieli, "Learned regulation of spatially localized brain activation using real-time fmri," *Neuroimage*, vol. 21, pp. 436–443, 2004.

[97] S. Posse, D. Fitzgerald, K. X. Gao, U. Habel, D. Rosenberg, G. J. Moore, and F. Schneider, "Real-time fmri of temporolimbic regions detects amygdala

activation during single-trial self-induced sadness," *Neuroimage*, vol. 18, no. 3, pp. 760–768, 2003.

[98] S. S. Yoo, H. M. O'Leary, T. Fairneny, N. K. Chen, L. P. Panych, H. Park, and F. A. Jolesz, "Increasing cortical activity in auditory areas through neurofeedback functional magnetic resonance imaging," *Neuroreport*, vol. 17, no. 12, pp. 1273–1278, 2006.

[99] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer, "Principles of a brain-computer interface (bci) based on real-time functional magnetic resonance imaging (fmri)," *Ieee Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 966–970, 2004.

[100] R. C. deCharms, F. Maeda, G. H. Glover, D. Ludlow, J. M. Pauly, D. Soneji, J. D. E. Gabrieli, and S. C. Mackey, "Control over brain activation and pain learned by using real-time functional mri," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 51, pp. 18626–18631, 2005.

[101] G. H. Smith and R. Engel, "Influence of a female model on perceived characteristics of an automobile," in *76th Annual Convention American Psychological Association*, 1968.

[102] M. J. Baker and G. A. J. Churchill, "The impact of physically attractive models on advertising evaluations," *Journal of Marketing Research*, vol. XIV, pp. 538–555, 1977.

[103] S. S. Yoo, J. H. Lee, H. O'Leary, V. Lee, S. E. Choo, and F. A. Jolesz, "Functional magnetic resonance imaging-mediated learning of increased activity in auditory areas," *Neuroreport*, vol. 18, no. 18, pp. 1915–1920, 2007.

[104] S.-S. Yoo, T. Faireny, N.-K. Chen, S.-E. Choo, L. P. Panych, H. Park, S.-Y. Lee, and F. A. Jolesz, "Brain-computer interface using fmri: spatial navigation by thoughts," *Neuroreport*, vol. 15, no. 10, pp. 1591–1595, 2004.

[105] S.-S. Yoo and F. A. Jolesz, "Functional mri for neurofeedback: feasibility study on hand motor task," *Neuroreport*, vol. 13, no. 11, pp. 1377–1381, 2002.

[106] N. Weiskopf, F. Scharnowski, R. Veit, R. Goebel, N. Birbaumer, and K. Mathiak, "Self-regulation of local brain activity using real-time functional magnetic resonance imaging," *Journal of Physiology - Paris*, vol. 98, pp. 357–373, 2004.

[107] N. Weiskopf, R. Veit, M. Erb, K. Mathiak, W. Grodd, R. Goebel, and N. Birbaumer, "Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fmri): methodology and exemplary data," *Neuroimage*, vol. 19, pp. 577–586, 2003.

[108] A. Caria, R. Veit, R. Sitaram, M. Lotze, N. Welskopf, W. Grodd, and N. Birbaumer, "Regulation of anterior insular cortex activity using real-time fmri," *Neuroimage*, vol. 35, no. 3, pp. 1238–1246, 2007. Times Cited: 5.

[109] E. T. Rolls, "The orbitofrontal cortex and reward," *Cerebral Cortex*, vol. 10, pp. 284–294, 2000.

[110] R. Adolphs, "The neurobiology of social cognition," *Current Opinion in Neurobiology*, vol. 11, pp. 231–239, 2001.

[111] I. R. Olson and C. Marshuetz, "Facial attractiveness is appraised at a glance," *Emotion*, vol. 5, no. 4, pp. 186–201, 2005.

[112] G. Rhodes, L. W. Simmons, and M. Peters, "Attractiveness and sexual behavior: Does attractiveness enhance mating success?," *Evolution and Human Behavior*, vol. 26, no. 2, pp. 186–201, 2005.

[113] K. Dion, E. Bersheid, and E. Walster, "What is beautiful is good," *Journal of Personality and Social Psychology*, vol. 24, pp. 285–290, 1972.

[114] D. S. Hamermesh and J. E. Biddle, "Beauty and the labor market," *The American Economic Review*, vol. 84, no. 5, pp. 1174–1194, 1994.

[115] R. L. Dipboye, R. D. Arvey, and D. E. Terpstra, "Sex and physical attractiveness of raters and applicants as determinants of resume evaluations," *Journal of Applied Psychology*, vol. 4, pp. 288–294, 1977.

[116] G. Rhodes, "The evolutionary psychology of facial beauty," *Annual Review Psychology*, vol. 57, pp. 199–226, 2006.

[117] H. Thornhill and S. W. Gangestad, "Facial attractiveness," *Trends in Cognitive Sciences*, vol. 3, no. 12, pp. 452–460, 1999.

[118] J. V. Haxby, E. A. Hoffman, and I. M. Gobbini, "The distributed human neural system for face perception," *Trends in Cognitive Sciences*, vol. 4, pp. 223–233, 2000.

[119] A. Ishai, C. F. Schmidt, and P. Boseiger, "Face perception is mediated by a distributed cortical network," *Brain Research Bulletin*, vol. 67, pp. 87–93, 2005.

[120] N. Kanwisher, J. McDermott, and M. Chun, "The fusiform face area: a module in human extrastriate cortex specialized for the perception of faces," *Journal of Neuroscience*, vol. 17, pp. 4302–4311, 1997.

[121] R. J. R. Blair, "Facial expressions, their communicatory functions and neurocognitive substrates," *Philosophical Transactions of the Royal Society London B*, vol. 358, no. 1431, pp. 561–572, 2003.

[122] P. J. Whalen, S. L. Rauch, N. L. Etcoff, S. C. McInerney, M. B. Lee, and M. A. Jenike, "Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge," *Journal of Neuroscience*, vol. 18, no. 1, pp. 411–418, 1998.

[123] F. Kranz and A. Ishai, "Face perception is modulated by sexual preference," *Current Biology*, vol. 16, pp. 63–68, 2006.

[124] S. M. McClure, G. S. Berns, and P. R. Montague, "Temporal prediction errors in a passive learning task activate human striatum," *Neuron*, vol. 38, no. 2, pp. 339–346, 2003.

[125] 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*, vol. 23, no. 21, pp. 7931–7939, 2003.

[126] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, pp. 1593–1599, 1997.

[127] J. O'Doherty, P. Dayan, J. Schultz, R. Deichmann, K. Friston, and R. J. Dolan, "Dissociable roles of ventral and dorsal striatum in instrumental conditioning," *Science*, vol. 304, no. 5669, pp. 452–454, 2004.

[128] J. P. O'Doherty, T. W. Buchanan, B. Seymour, and R. J. Dolan, "Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum," *Neuron*, vol. 49, no. 1, pp. 157–166, 2006.

[129] R. S. Sutton and A. G. Barto, "Time derivative models of pavlovian reinforcement," in *Learning and Computational Neuroscience: Foundations of adaptive networks* (M. Gabriel and J. Moore, eds.), pp. 497–537, Cambridge, MA, USA: MIT Press, 1990.

[130] H. D. Critchley, C. J. Mathias, and R. J. Dolan, "Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy," *Neuron*, vol. 33, no. 4, pp. 653–663, 2002.

[131] J. A. Gottfried, J. O'Doherty, and R. J. Dolan, "Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging," *Journal of Neuroscience*, vol. 22, no. 24, pp. 10829–10837, 2002.

[132] R. Deichmann, J. A. Gottfried, C. Hutton, and R. Turner, "Optimized epi for fmri studies of the orbitofrontal cortex," *Neuroimage*, vol. 19, no. 2, pp. 430–441, 2003.

[133] K. J. Friston, J. Ashburner, J. B. Poline, C. D. Frith, J. D. Heather, and R. S. Frackowiak, "Spatial registration and normalisation of images," *Human Brain Mapping*, vol. 2, pp. 165–189, 1995.

[134] J. L. Andersson, C. Hutton, J. Ashburner, R. Turner, and K. Friston, "Modeling geometric deformations in epi time series," *Neuroimage*, vol. 13, pp. 903–919, 2001.

[135] B. Knutson, J. Taylor, M. Kaufman, R. Peterson, and G. Glover, "Distributed neural representation of expected value," *Journal of Neuroscience*, vol. 25, no. 19, pp. 4806–4812, 2005.

[136] S. M. Cox, A. Andrade, and I. S. Johnsrude, "Learning to like: a role for human orbitofrontal cortex in conditioned reward," *Journal of Neuroscience*, vol. 25, no. 10, pp. 2733–2740, 2005.

[137] H. Kim, S. Shimojo, and J. P. O'doherty, "Is avoiding an aversive outcome rewarding? neural substrates of avoidance learning in the human brain," *PLoS Biology*, vol. 4, no. 8, 2006.

[138] M. R. Cunningham, A. R. Roberts, A. P. Barbee, P. B. Druen, and C.-H. Wu, ""their ideas of beauty are, on the whole, the same as ours": consistency and variability n the cross-cultural perception of female physical attractiveness," *Journal of Personality and Social Psychology*, vol. 68, pp. 261–279, 1995.

[139] K. D. M. LaBar, "Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fmri study," *Neuron*, vol. 20, no. 5, pp. 937–945, 1998. 0896-6273.

[140] H. Garavan, J. C. Pendergrass, T. J. Ross, E. A. Stein, and R. C. Risinger, "Amygdala response to both positively and negatively valenced stimuli," *Neuroreport*, vol. 12, no. 12, pp. 2779–2783, 2001.

[141] S. B. Hamann, T. D. Ely, S. T. Grafton, and C. D. Kilts, "Amygdala activity related to enhanced memory for pleasant and aversive stimuli," *Nature Neuroscience*, vol. 2, no. 3, pp. 289–293, 1999.

[142] J. S. Winston, J. O'Doherty, and R. J. Dolan, "Common and distinct neural responses during direct and incidental processing of multiple facial emotions," *Neuroimage*, vol. 20, no. 1, pp. 84–97, 2003.

[143] G. J. Gorn, M. E. Goldberg, and K. Basu, "Mood, awareness, and product evaluation," *Journal of Consumer Psychology*, vol. 2, no. 3, pp. 237–256, 1993.

[144] S. Folkman, R. S. Lazarus, C. Dunkel-Schetter, A. DeLongis, and R. J. Gruen, "Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes," *Journal of Personality and Social Psychology*, vol. 50, no. 5, pp. 992–1003, 1986.

[145] F. Baeyens, D. Hermans, and P. Eelen, "The role of cs-us contingency in human evaluative conditioning," *Behav Res Ther*, vol. 31, no. 8, pp. 731–737, 1993.

[146] M. Pessiglione, B. Seymour, G. Flandin, R. J. Dolan, and C. D. Frith, "Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans," *Nature*, vol. 442, no. 7106, pp. 1042–1045, 2006.

[147] R. M. Colwill and R. A. Rescorla, "Associations between the discriminative stimulus and the reinforcer in instrumental learning," *Journal of Experimental Psychology-Animal Behavior Processes*, vol. 14, no. 2, pp. 155–164, 1988.

[148] P. C. Holland, "Relations between pavlovian-instrumental transfer and reinforcer devaluation (vol 30, pg 104, 2004)," *Journal of Experimental Psychology-Animal Behavior Processes*, vol. 30, no. 4, pp. 258–258, 2004.

[149] P. M. Smeets and D. Barnes-Holmes, "Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer," *Journal of Economic Psychology*, vol. 24, no. 5, pp. 603–618, 2003.

[150] L. Hogarth, A. Dickinson, A. Wright, M. Kouvaraki, and T. Duka, "The role of drug expectancy in the control of human drug seeking," *Journal of Experimental Psychology-Animal Behavior Processes*, vol. 33, no. 4, pp. 484–496, 2007.

[151] S. B. Ostlund and B. W. Balleine, "Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome," *Learning and Behavior*, vol. 35, no. 1, pp. 43–52, 2007. Times Cited: 0.

[152] P. Bitsios, E. Szabadi, and C. M. Bradshaw, "The fear-inhibited light reflex: importance of the anticipation of an aversive event," *International Journal of Psychophysiology*, vol. 52, no. 1, pp. 87–95, 2004.

[153] F. T. Russchen, I. Bakst, D. G. Amaral, and J. L. Price, "The amygdalostriatal projections in the monkey - an anterograde tracing study," *Brain Research*, vol. 329, no. 1-2, pp. 241–257, 1985.

[154] L. D. Selemon and P. S. Goldmanrakic, "Longitudinal topography and interdigitation of corticostriatal projections in the rhesus-monkey," *Journal of Neuroscience*, vol. 5, no. 3, pp. 776–794, 1985.

[155] J. L. Fudge and S. N. Haber, "Defining the caudal ventral striatum in primates: Cellular and histochemical features," *Journal of Neuroscience*, vol. 22, no. 23, pp. 10078–10082, 2002.

[156] J. L. Fudge, K. Kunishio, P. Walsh, C. Richard, and S. N. Haber, "Amygdaloid projections to ventromedial striatal subterritories in the primate," *Neuroscience*, vol. 110, no. 2, pp. 257–275, 2002.

[157] R. A. Rescorla and R. L. Solomon, "2-process learning theory - relationships between pavlovian conditioning and instrumental learning," *Psychological Review*, vol. 74, no. 3, p. 151, 1967.

[158] B. W. Balleine and S. B. Ostlund, "Still at the choice-point - action selection and initiation in instrumental conditioning," *Reward and Decision Making in Corticobasal Ganglia Networks*, vol. 1104, pp. 147–171, 2007. Times Cited: 0.

[159] A. M. Ludwig, A. Wikler, and L. H. Stark, "First drink - psychobiological aspects of craving," *Archives of General Psychiatry*, vol. 30, no. 4, pp. 539–547, 1974. Times Cited: 223.

[160] S. Bray, S. Shimojo, and J. P. O'Doherty, "Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback," *Journal of Neuroscience*, vol. 27, no. 28, pp. 7498–7507, 2007.

[161] W. Small, "An experimental study of the mental processes of the rat," *American Journal of Psychology*, vol. 12, pp. 206–239, 1901.

[162] T. Ljungberg, P. Apicella, and W. Schultz, "Responses of monkey dopamine neurons during learning of behavioral reactions," *Journal of Neurophysiology*, vol. 67, no. 1, pp. 145–163, 1992.

[163] B. W. Balleine and A. Dickinson, "Goal-directed instrumental action: contingency and incentive learning and their cortical substrates," *Neuropharmacology*, vol. 37, no. 4-5, pp. 407–419, 1998.

[164] R. W. Cox, A. Jesmanowicz, and J. S. Hyde, "Real-time functional magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 33, no. 2, pp. 230–236, 1995.

[165] D. Gembris, J. G. Taylor, S. Schor, W. Frings, D. Suter, and S. Posse, "Functional magnetic resonance imaging in real time (fire): Sliding-window correlation analysis and reference-vector optimization," *Magnetic Resonance in Medicine*, vol. 43, pp. 259–268, 2000.

[166] M. S. Schwartz, *Biofeedback - A Practitioner's Guide*. New York: Guilford Press, 1995.

[167] R. W. Cox and A. Jesmanowicz, "Real-time 3d image registration for functional mri," *Magnetic Resonance in Medicine*, vol. 42, no. 6, pp. 1014–1018, 1999.

[168] S. Eickhoff, K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, and K. Zilles, "A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data," *Neuroimage*, vol. 25, no. 4, pp. 1325–1335, 2005.

[169] T. A. Yousry, U. D. Schmid, H. Alkadhi, D. Schmidt, A. Peraud, A. Buettner, and P. Winkler, "Localization of the motor hand area to a knob on the precentral gyrus: A new landmark," *Brain*, vol. 120, pp. 141–157, 1997.

[170] H. H. Ehrsson, S. Geyer, and E. Naito, "Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations," *Journal of Neurophysiology*, vol. 90, pp. 3304–3316, 2003.

[171] A. Keil, M. M. Muller, T. Gruber, C. Wienbruch, and T. Elbert, "Human large-scale oscillatory brain activity during an operant shaping procedure," *Cognitive Brain Research*, vol. 12, no. 3, pp. 397–407, 2001.

[172] X. Y. Chen and J. R. Wolpaw, "Operant-conditioning of h-reflex in freely moving rats," *Journal of Neurophysiology*, vol. 73, no. 1, pp. 411–415, 1995.

[173] P. R. Green, L. Gentle, T. M. Peake, R. E. Scudamore, P. K. McGregor, F. Gilbert, and W. H. Dittrich, "Conditioning pigeons to discriminate naturally lit insect specimens," *Behavioural Processes*, vol. 46, no. 1, pp. 97–102, 1999.

[174] R. D. Hawkins, G. A. Clark, and E. R. Kandel, "Operant conditioning of gill withdrawal in aplysia," *Journal of Neuroscience*, vol. 26, no. 9, pp. 2443–2448, 2006.

[175] K. Samejima, Y. Ueda, K. Doya, and M. Kimura, "Representation of action-specific reward values in the striatum," *Science*, vol. 310, no. 5752, pp. 1337–1340, 2005.

[176] J. Yacubian, J. Glascher, K. Schroeder, T. Sommer, D. F. Braus, and C. Buchel, "Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain," *Journal of Neuroscience*, vol. 26, no. 37, pp. 9530–9537, 2006.

[177] R. F. Hefferline, B. Keenan, and R. A. Harford, "Escape and avoidance conditioning in human subjects without their observation of the response," *Science*, vol. 130, no. 3385, pp. 1338–1339, 1959.

[178] F. Svartdal, "When feedback contingencies and rules compete - testing a boundary-condition for verbal control of instrumental performance," *Learning and Motivation*, vol. 26, no. 3, pp. 221–238, 1995.

[179] K. M. Stephan, G. R. Fink, R. E. Passingham, D. Silbersweig, A. O. Ceballosbaumann, C. D. Frith, and R. S. J. Frackowiak, "Functional-anatomy of the mental representation of upper extremity movements in healthy-subjects," *Journal of Neurophysiology*, vol. 73, no. 1, pp. 373–386, 1995.

[180] E. Gerardin, A. Sirigu, S. Lehericy, J. B. Poline, B. Gaymard, C. Marsault, Y. Agid, and D. Le Bihan, "Partially overlapping neural networks for real and imagined hand movements," *Cerebral Cortex*, vol. 10, no. 11, pp. 1093–1104, 2000.

[181] M. Leonardo, J. Fieldman, N. Sadato, G. Campbell, V. Ibanez, L. Cohen, M. P. Deiber, P. Jezzard, T. Pons, R. Turner, D. LeBihan, and M. Hallett, "A functional magnetic resonance imaging study of cortical regions associated with motor task execution and motor ideation in humans," *Human Brain Mapping*, vol. 3, no. 2, pp. 83–92, 1995.

[182] L. Nyberg, J. Eriksson, A. Larsson, and P. Marklund, "Learning by doing versus learning by thinking: An fmri study of motor and mental training," *Neuropsychologia*, vol. 44, pp. 711–717, 2005.

[183] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial eeg," *Cognitive Brain Research*, vol. 25, no. 3, pp. 668–677, 2005.

[184] C. M. Stinear, W. D. Byblow, M. Steyvers, O. Levin, and S. P. Swinnen, "Kinesthetic, but not visual, motor imagery modulates corticomotor excitability," *Experimental Brain Research*, vol. 168, no. 1-2, pp. 157–164, 2006.

[185] J. S. Beer, O. P. John, D. Scabini, and R. T. Knight, "Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions," *Journal of Cognitive Neuroscience*, vol. 18, no. 6, pp. 871–879, 2006. Times Cited: 10.

[186] 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*, vol. 17, no. 11, pp. 2669–2674, 2007.

[187] A. P. R. Smith, R. N. A. Henson, R. J. Dolan, and M. D. Rugg, "fmri correlates of the episodic retrieval of emotional contexts," *Neuroimage*, vol. 22, no. 2, pp. 868–878, 2004. Times Cited: 43.

[188] L. Kuchinke, A. M. Jacobs, C. Grubich, M. L. H. Vo, M. Conrad, and M. Herrmann, "Incidental effects of emotional valence in single word processing: An fmri study," *Neuroimage*, vol. 28, no. 4, pp. 1022–1032, 2005.

[189] J. O'Doherty, M. L. Kringelbach, E. T. Rolls, J. Hornak, and C. Andrews, "Abstract reward and punishment representations in the human orbitofrontal cortex," *Nature Neuroscience*, vol. 4, no. 1, pp. 95–102, 2001.

[190] A. N. Hampton and J. P. O'Doherty, "Decoding the neural substrates of reward-related decision making with functional mri," *Proceedings of the National Acad-*

emy of Sciences of the United States of America, vol. 104, no. 4, pp. 1377–1382, 2007. Times Cited: 12.

[191] S. O. Dumoulin, R. G. Bittar, N. J. Kabani, C. L. Baker, G. Le Goualher, G. B. Pike, and A. C. Evans, "A new anatomical landmark for reliable identification of human area v5/mt: a quantitative analysis of sulcal patterning," *Cerebral Cortex*, vol. 10, no. 5, pp. 454–463, 2000.

[192] S. Zeki, J. D. G. Watson, C. J. Lueck, K. J. Friston, C. Kennard, and R. S. J. Frackowiak, "A direct demonstration of functional specialization in human visual-cortex," *Journal of Neuroscience*, vol. 11, no. 3, pp. 641–649, 1991.

[193] G. McCaig and K. Christoff, "Feedback presentation techniques for real-time fmri training of higher cortical regions," 2008.

[194] M. Carlson, "Characteristics of sensory deficits following lesions of brodmann areas-1 and areas-2 in the postcentral gyrus of macaca-mulatta," *Brain Research*, vol. 204, no. 2, pp. 424–430, 1981.

[195] D. T. Gering and D. M. Weber, "Intraoperative, real-time, functional mri," *Jmri-Journal of Magnetic Resonance Imaging*, vol. 8, no. 1, pp. 254–257, 1998.

Appendix A

Real-time fMRI

General principles

In typical fMRI studies, 3D images are collected at a rate of every 1-2 s for the duration of the scan. Each image consists of ~2^{16} voxel elements (equivalent to a 3D pixel), and it is not uncommon to collect hundreds of such images. As the images are acquired on the scanner, they are reconstructed (transformed from k-space, in which they are acquired, into physical space), and stored in a database. At some later point, the images are downloaded, pre-processed, and analyzed by the experimenter. Pre-processing and analysis can take several hours to perform on a large data set.

The earliest fMRI studies were run in blocked designs, due to limitations on scanner technology and to improve signal-to-noise. However, practically since the advent of technologies for rapid event-related imaging, researchers began describing techniques for analyzing fMRI data in real-time [164]. The term 'real-time fMRI' (rt-fMRI) typically describes processing that keeps pace with image acquisition. Lags between image acquisitions are on the order of 1-3 s, therefore with modern computers this is certainly feasible.

Typical processing steps in an rt-fMRI study are listed below, and are in fact quite similar to those employed in ordinary offline fMRI analysis.

1. Image acquisition and reconstruction

2. Pre-processing: may include

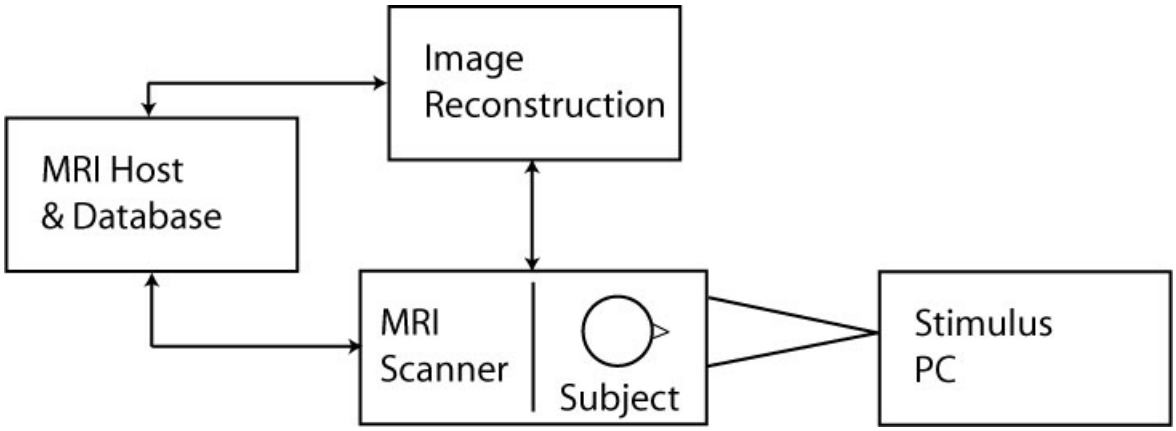

Figure A.1: Conventional fMRI experimental setup
This diagram illustrates the setup of a typical fMRI experiment. Images collected during the scan are stored in a database and downloaded at a later time.

(a) online motion correction

(b) temporal filtering for high frequency noise or slow drift

(c) spatial smoothing

(d) normalization to template

3. Analysis

   Typically signals are extracted from one or several regions of interest and statistical analysis may be performed, e.g., comparing task to rest blocks or comparing target region of interest to background region of interest

Technical implementation

The technical setup employed for conventional fMRI studies at the Caltech Brain Imaging Center is illustrated in Figure A.1; the fMRI scanner is a 3T Siemens TRIO (Siemens, Erlangen, Germany). As subjects lie in the scanner, stimuli are presented to them visually or auditorily, controlled by a dedicated stimulus computer located in the control room. During the experiment, images are acquired, reconstructed, and stored in a database for offline analysis.

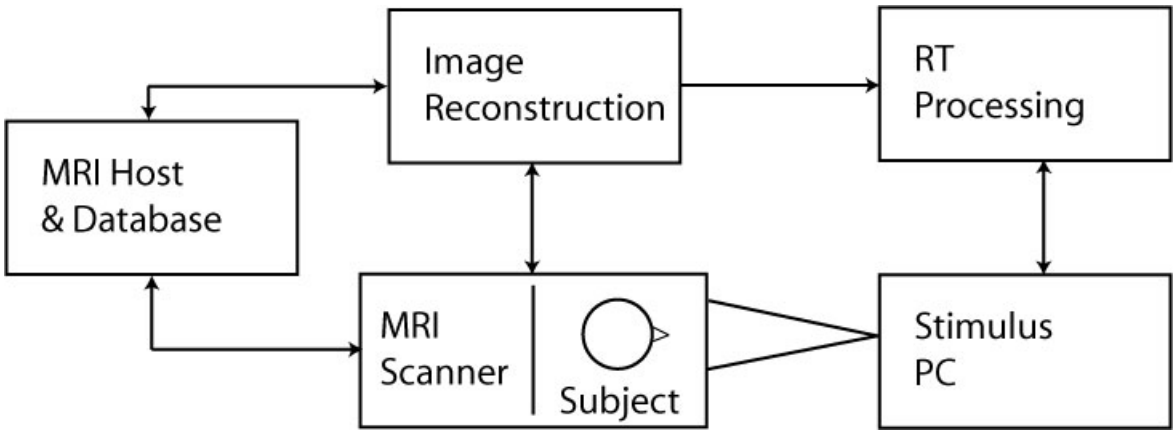

Figure A.2: Real-time fMRI experimental setup
This diagram illustrates the setup of a real-time fMRI experiment. Images are transferred to an external workstation for online analysis. Feedback derived from the analysis can be presented to the subject in the scanner.

In order to implement real-time image processing, we inserted an extra node into the network (Figure A.2): an external Intel Xeon workstation (3.8 MHz 64-bit processor running Redhat Linux). As soon as images are reconstructed, they are transferred in real-time via TCP/IP socket to this dedicated rt-fMRI processor. This is accomplished with a modified pulse sequence program running on the scanner, in which a command to open a TCP/IP socket and transfer the newly reconstructed images was inserted into the regular processing stream.

Online analysis of fMRI images

On the external computer, data processing was performed using MATLAB 7.0 (The Mathworks Inc., Natick, MA). Images were motion corrected [167], and a linear detrend was applied to correct for low-frequency scanner drift. Online analysis consisted of applying a mask over each of the regions of interest and averaging the signal over the region. Temporal averaging was then performed over the baseline and active blocks, and a %-change from baseline to active computed. Information derived from this signal could then be used as feedback to the subjects in the scanner. The rt-fMRI computer communicates with the stimulus PC via Samba (`http://us1.samba.org/`

samba/).

Applications

In the studies described in this thesis, and several others [106, 99, 107, 105, 104, 98, 103, 96, 100, 108], rt-fMRI was applied to deliver feedback of neural responses to subjects as a training signal. Other potential applications of rt-fMRI include surgical applications [195] and online monitoring of experiments for desired effects and quality assurance.