Regretful Choices: Neural Representations of Choice in the Orbitofrontal Cortex

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Adam Patrick Steiner

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Aaron David Redish

October 2014

Acknowledgements

I am fortunate to have had extensive guidance from my advisor David Redish. His guidance and critical thinking have drastically shaped my ability to design, think and answer questions as a scientist. Our conversations have nurtured my endless curiosity for science and have significantly increased my understanding of neuroscience. I have been extremely lucky to be a member of the Redish lab and surrounded by the many brilliant and insightful members. I am extremely grateful for entertaining and enlightening discussions that not only guided my work but also clarified my thought processes: in particular Nate Powell, Andrew Wikenheiser, Andy Papale, Jeff Stott, Paul Reiger, Chris Boldt, Kelsey Seeland and Ayaka Sheehan.

I am extremely grateful for all the guidance and assistance previous members of the lab, Jadin Jackson, Matt van der Meer and John Ferguson, provided as well as the guidance and thought provoking discussion from my committee, Matt Chafee, James Ashe, and Mark Thomas. In addition, the support provided by the Graduate Program Neuroscience at the University of Minnesota has been exceptional.

Last but not least, I am very grateful for the loving support of my family, mom, dad, mom2, dad2, Angela, Laura, Lynn, my beautiful wife, Ellen, and my friends who helped maintain my sanity in this process.

Dedication

This thesis is dedicated to my wife, Ellen Schofield. Without your encouragement, love and constant patience this work would not have been possible. Thank you for your continued support and endless love. There are not enough words (~48,000) in this thesis to describe how much I love and appreciate you.

Abstract

In order to successfully acquire reward under many different circumstances, a decision maker must learn to expect a particular outcome in a specific situation. The ability to predict an expected outcome simplifies the decision process and enables a decision maker to accurately choose between multiple options without having to experience each option independently. The orbitofrontal cortex represents specific outcomes and aids in the selection of actions to acquire specific outcomes without experiencing the direct action-outcome sequence. Without the orbitofrontal cortex, humans, non-human primates, and rats all exhibit an inability to modify their actions to changing reward conditions. In addition, an intact orbitofrontal cortex is required to correctly identify cues paired with and predictive of reward. Without the orbitofrontal cortex the ability to distinguish between differing expectations collapses.

The ability to create counterfactuals, a representation of the alternative would-have-been received outcome, is thought to aid in the process of simulating the expected outcomes of a situation. The representation of counterfactuals has been found in the orbitofrontal cortex in humans and non-human primates. While there is some evidence that rats can represent the counterfactual, the only direct neural evidence is presented in this thesis. Without the representation of a counterfactual, it is impossible to experience

iv

regret. If the orbitofrontal cortex is homologous across species and if rats can represent the counterfactual, can rats represent the counterfactual during regret inducing situations?

Chapters 1 and 2 of this thesis examine the homology of the orbitofrontal cortex comparing the structure of OFC in humans, rats and non-human primates. These chapters also summarize the current hypotheses regarding Orbitofrontal cortex function. Orbitofrontal cortex has been shown to be largely homologous among humans, non-human primates and rats.

The third chapter of this thesis introduces the concept of regret and counterfactuals and explains the etymology, psychology and economics that describe how regret and counterfactuals can be studied.

Chapter 4 of this thesis shows that rats are capable of representing the counterfactual. One key aspect of orbitofrontal function is in the representation of counterfactuals and regret. Counterfactuals are defined as the alternative, would-have-been option. Rats performing on a multiple T maze, stopped at a decision point and looked both directions before continuing to reward. During this pause, orbitofrontal neurons represented reward after ventral striatal neurons represented reward. The reward representation was general. Once rats arrived at the reward sites, orbitofrontal neurons reliably represented the reward and continued to do so for every reward encounter. However, when rats arrived at reward locations that were not active, orbitofrontal neurons represented the other reward site, representing the counterfactual.

Chapters 5 and 6 of this thesis expand on the necessity of counterfactuals to decision making and shows that rats are capable of experiencing regret on an economic foraging task. Rats showed distinct economic preferences on a circular foraging task. Rats waited longer for flavors of reward that they preferred and spent less time waiting for rewards they did not prefer. Neural responses reliably differentiated between reward flavors and the zones associated with the reward flavors.

When a rat left a preferred reward early, without receiving reward, then encountered a non-preferred reward with a longer wait, the rat's behavior matched the economic definition of regret. Regret occurs when a decision maker selects between two options and the option finally received is less valuable than what the alternative option would have produced. Importantly this can be differentiated from disappointment, where an option received is less than expected even though it was not the fault of the decision maker. Counterfactuals are necessary for the experience of regret. Humans without an orbitofrontal cortex do not experience regret. However, regret has traditionally been known as a human experience. Economic definitions of regret make it possible to measure regret in rodents. During these regret instances neural ensembles in the orbitofrontal cortex represented the missed, previous action. These representations agree with psychological accounts of regret that state decision makers regret the action that led to outcome more than outcome.

In the final chapter, chapter 7, the role of orbitofrontal cortex is re-examined in the context of regret and counterfactuals. Specifically, the previously introduced hypotheses from Chapter 2 are reviewed under the framework introduced by counterfactuals and regret.

Table of Contents

List of Figures

xiv

xv

List of Tables

1. Neuroanatomy of the orbitofrontal cortex

The orbitofrontal cortex (OFC) is located on the ventral medial surface of the forebrain and receives a large number of inputs. The OFC receives sensory information from many sources, such as the gustatory and olfactory cortices. Other inputs from structures such as the ventral striatum and prefrontal cortex support integration of sensory information, reward information and higher level, cognitive processes potentially providing the information necessary to instruct future choices. The OFC has been implicated in a variety of decision making tasks and is essential in reversal tasks as well as in the recognition of regret.

1.1 Basic connectivity of OFC

The OFC receives a wide variety of inputs from many different regions throughout the brain, both as direct connections and as thalamic connections from the medial dorsal nucleus of the thalamus[45,46,86,147-149,151-153,221,248]. The OFC has strong

reciprocal connections with the medial prefrontal cortex and the limbic system, two areas associated with decision making[160-162,252,295,297,360,361].

The internal connectivity of OFC reveals two separate connectivity patterns. There is little interconnectivity between the two relatively distinct networks within the OFC[221]. Even though previous results from humans and non-human primates are reported to be from different regions within OFC, the connectivity patterns between OFC and other areas of the cortex are largely consistent. The connectivity similarities between human and non-human primates has been supported by diffusion tensor imaging of the fiber tracts in human and non-human primate OFC[61]. In both humans and non-human primates prefrontal connectivity with subcortical regions was homologous.

The more medial aspects of OFC are reciprocally connected to the ventral medial prefrontal cortex. Central and lateral OFC show a much weaker reciprocal connectivity with ventral medial prefrontal cortex and receive more visceral afferents[43,221].

In humans and non-human primates OFC receives and projects to the amygdalar complex (specifically the basal lateral nucleus of the amygdala), anterior hippocampus, hypothalamus, nucleus accumbens, and cingulate cortex[45,46,86,104,221]. Regions of the cortex that process somatosensory, olfactory and viscera inputs project strongly to the OFC[45,47,221,275]

1.2 Comparing OFC in rats with humans and non-human primates

While human and non-human primate OFC are generally agreed to be homologous, there is still debate as to whether rat OFC is homologous. The neuronal profile of OFC in humans and non-human primates consists of pyramidal cells as well as small granule cell layers[44,181,248]. However rat OFC consists entirely of large bodied pyramidal cells[248]. Rats lack any evidence of smaller cell body, granule layers in OFC[221] and until recently little evidence supported distinct differences in projection patterns from the different regions of OFC to other areas of the brain[183]. However neuroanatomical evidence now indicates that rat OFC may be homologous to non-human primate and human OFC.

If the orbitofrontal cortex is classified by thalamic projections[277], the connectivity pattern of rat OFC is homologous to non-human primate and human[101]. By using thalamic inputs and outputs, the medial dorsal nucleus of the thalamus projects to the same orbitalfrontal areas in rats, non-human primates and humans[45,101,152,153,164]. Based on these projections and their relatively similarity across species, the orbital and agranular insular areas of rat orbitofrontal cortex can be considered to be homologous to non-human primate and human OFC[101,164,221,246,306]. The OFC in the rat includes ventral orbital, lateral orbital and to some extent the agranular insular cortex. However OFC classification in the rat does not include the region along the medial wall of the ventral

3

forebrain which is likely equivalent to ventral medial prefrontal cortex (vmPFC) in humans and non-human primates[183,212,221,304].

Like humans and non-human primates, the OFC in rats receives inputs from and projects to the amygdalar complex (the majority of reciprocal connections terminate in the basal lateral amygdala), anterior hippocampus[250], hypothalamus and nucleus accumbens[183,221]. In addition, the strong connectivity of OFC with the basal lateral amygdala in both non-human primates and rats is thought to contribute to the motivational and emotional constituents of learning[9,68,80,118,145,161,295]. Lesions to amygdala and OFC in both non-human primates and rats produce similar deficits[85,90,127,186,233,295,296,301,306,363]. In addition in both rats and non-human primates the OFC projects to the nucleus accumbens, overlapping with inputs from the amygdala[100,104,194,220,348]. Most primate neurophysiology has focused on the anterior, central OFC[225-228] while most rat physiology has focused on the more posterior aspects of the orbitofrontal-medial/lateral cortex[142,184,296,298,299,327,328,356-358]. In contrast, most prior research in humans is reported from the ventral medial prefrontal cortex[107,130,137]. The discrepancy in recording locations, lesion boundaries and the experimental results in rats, non-human primates and humans raises a very important question: are all regions of OFC functionally equivalent?

4
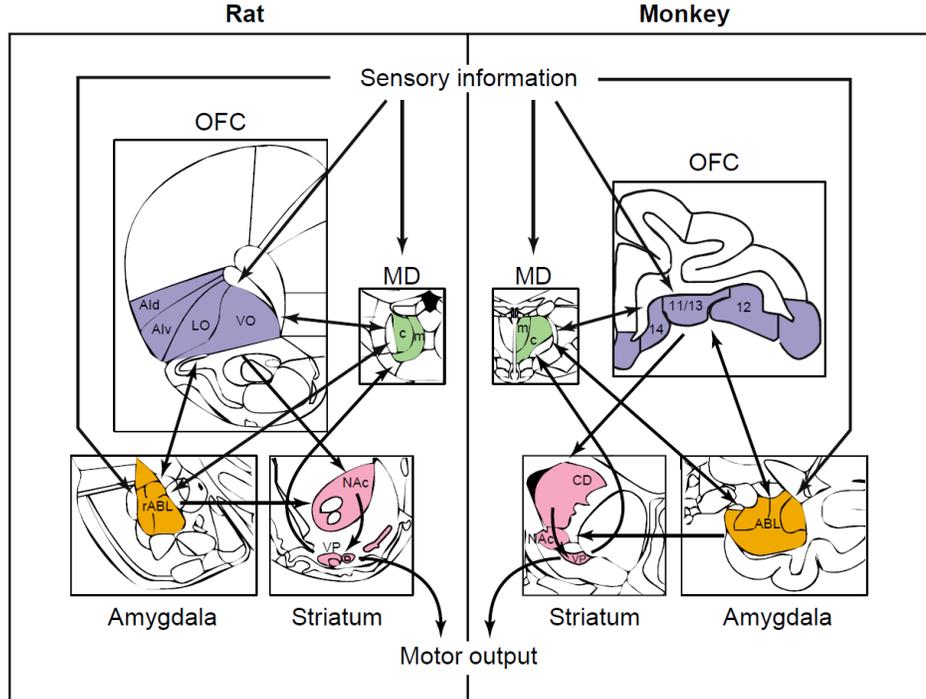
Figure 1-1 **Connectivity patterns of OFC in rats and non-human primates.**
OFC (blue), striatum (pink), amygdala (orange), medial dorsal thalamus (green).The connectivity patterns of medial dorsal thalamus and orbitofrontal cortex are largely consistent between rats and non-human primates. Inputs and outputs to and from OFC are conserved across species. (used with permission[303])

1.3 Homogeneity of the orbitofrontal cortex

Other areas of prefrontal cortex are functionally distinct. Similarly, there may be functional distinctions in OFC regions. Anatomical evidence suggests that there are regional OFC differences in connectivity and function in non-human primates[45], humans[137] and finally in rats[152,153,247]. These regional distinctions may contribute to the different results reported in a variety of past experiments regarding OFC function in humans, non-human primates and rats.

5

A growing body of literature suggests that there is indeed a behavioral and neurophysiological distinction between the different regions of OFC in humans (ventral-medial versus lateral)[144,196,234,329,373], non-human primates (central versus lateral)[141,213,215,228,265,342] and rats (ventral-lateral versus dorsal-lateral)[81,142,283,296,298,299,327,328,356-358]. In human, the ventral medial OFC may be responsible for the representation of how valuable a reward is compared to other rewards (economic value)[21,74,97,108-110,130,131,214,229,241,349]. Lateral aspects of OFC may be involved in assigning reward to the actions required to receive reward (credit assignment)[196,214,282]. In non-human primates, there is little evidence to suggest the coding of economic value in the vmPFC, rather all current evidence suggests that economic value is being represented in central OFC[228] and that the more lateral aspects of OFC process credit assignment [365]. In rats, most of the neurophysiological and behavioral evidence comes from the ventral-lateral and dorsal-lateral OFC.

If the rat OFC is homologous to human and non-human primate OFC, then the ventral-lateral aspects of OFC in rats should play a role in credit assignment and exhibit reward related activity for any action or cue leading to reward. Several previous experiments have shown reward/cue related activity in ventral-lateral rat OFC as well distinct representations of credit assignment[91,142,184,296,298,299,308,327,328]. These experimental findings in rat ventral-medial OFC agree with experimental data from humans[99,349] and non-human primate lateral OFC[213,215,279,280,364,365]. In addition, the specific coding of

economic value as seen in non-human primates[228] may exist in the ventral medial aspects of rat OFC close to the ventral wall (medial orbital cortex and ventral pre limbic) where little neurophysiological evidence has yet to be published. What little evidence that exists regarding the role of ventral-medial orbital cortex in rats indicates that the medial orbital cortex is indeed involved in the calculation of value. Previous studies have found that the medial orbitofrontal cortex of rats is involved in the probability of receiving reward[320], selecting a learned behavior for a desired value[92] and choosing between delayed options effectively weighing the value of the available options[185,202,283,324]. Although sparse, these findings are consistent with human[13,25,196] and non-human primate data[117,215,261-263,265].

Other experiments in humans and non-human primates have shown that central OFC (in non-human primates)[1] and aspects of medial and lateral OFC (in humans) [41,56,57] are involved in the representation of hypothetical outcomes. Specifically, in humans the medial OFC represents missed past outcomes[56,57] and imagined future outcomes[33]. In non-human primates single neurons in the central OFC increased their activity for the cued, hypothetical outcome[1]. If rat ventral-lateral OFC is equivalent to these regions in human and non-human primates, then it would be expected to see similar responses when encountering situations that induce the representation of hypothetical outcomes. Indeed initial evidence suggests that rat ventral-lateral OFC does contain representations of imagined outcomes[333]. These experiments as well as their consistent findings across

animal models support the functional segregation of orbitofrontal regions in rats as well as the homology between rat, human and non-human primate OFC.

2. Hypotheses of orbitofrontal function

2.1 OFC function: evidence from behavioral experiments

Early experiments had indicated that the OFC was involved in multiple roles: i) decision making and planning[12-19,62-67,91,182,207,208,237,241,271,272,274,290,296-299,306,338], ii) judging whether decisions made had failed and had long term severe consequences[13,19,62,66,290,360], iii) and in the ability to identify and recall the implied meaning or importance of events/situations[13,62-67,78,91,272,274,290,297-299,311,338,342,343]. These early hypotheses regarding the function of OFC were broad and ill defined. Later research in humans, non-human primates and rats would more precisely describe the role of the OFC in decision making.

Early evidence indicated that OFC lesions resulted in highly emotional behavior[13,64,201]. Human subjects often behaved as if they had no recognition of future consequences and were only responding to simple Pavlovian cues[13,62,290]. Damasio postulated that the OFC played a primary role in the recognition of physiological changes that induced emotional changes such as anger or sadness[62,64,66]. He theorized that the role of the OFC would be to store different association patterns of external inputs with the

internal somatic states that may be produced by the external stimuli. A decision would

then activate the somatic state most closely related to the paired external stimuli, biasing

the decision. Damage to the OFC would make it difficult for human subjects to activate a

particular somatic state which meant all choices would be emotionally equivalent.

In reference to his patient EVR who suffered from bilateral OFC lesions,

Damasio stated:

> *"What we are proposing is that normal individuals can be assisted in this complex decision-making process by the appearance of a somatic signal that marks the ultimate consequences of the response option with a negative or positive somatic state. In other words, response option "A," regardless of its predictable immediate reward, can evoke a future scenario that is potentially threatening to the individual, and is marked by a negative somatic state. The perceiver would then experience the reenactment of punishment."*

Although EVR still possessed the ability to categorize social situations and emotions he

lacked the ability to use the information to successfully modify his behavior and select

the appropriate responses[290]. EVR was not perserverative or impulsive and had no

deficits in working memory or changes in intelligence[13,290].

Early evidence suggested that the OFC damage led to disruptions in behavior;

human subjects were oblivious to the future consequences of their actions and continually

selected large gains even when this action led to increasing large losses[13,15,62]. The Iowa

Gambling Task was designed to test subjects with ventral medial prefrontal damage.

They hypothesized, based on the deficits exhibited by EVR, that ventral medial prefrontal

damaged patients would be oblivious to the future consequences of their actions and

would be guided only by the immediate rewards. In a simplified version of the Iowa Gambling Task, human patients selected cards from one of two decks (the original version of the task had four decks). Each deck had a different probability of reward. Deck A was high loss/high gain. Continued selections from Deck A would result in substantial losses. Deck B was low loss/low gain. Continued selections from Deck B would result in a small net gain. Damasio's theory suggested that patients with OFC lesions would continue to select Deck A even though they were incurring substantial losses. Control subjects should switch to the low loss deck, Deck B. In order to make sure each deck was sampled, the first few cards in both decks were guaranteed wins.

Subjects were allowed to freely choose cards from either deck. The first few cards in Deck A led to much higher gains than the first few cards in Deck B (loss cards were positioned in each deck such that a subject would have to draw multiple times from a deck to encounter a loss card). A typical control subject sampled each of the decks and after approximately 10 cards and encountering at least 1 loss card in each deck began to select almost exclusively from the low loss/low gain deck. A typical OFC lesioned subject, like control subjects, sampled each of the decks. However in contrast to controls, a typical OFC lesioned subject after the first approximately 5 cards began to select almost exclusively from Deck A. A typical OFC lesioned subject generally did not encounter any loss cards in the first 5 cards from Deck A. High loss cards were not encountered until after approximately 10 cards were drawn from Deck A.

Initially the most advantageous action would have been to continually select cards from Deck A until learning that Deck A resulted in large losses, then switch to Deck B. Control subjects readily changed their selections from Deck A to Deck B after encountering one loss card from each deck. However, OFC lesioned subjects continued to select cards from Deck A, preferring the high gain rewards even after encountering the high loss cards.

The behavioral evidence that OFC lesioned subjects would continue to select the high risk deck leading to much greater losses supported the hypothesis that OFC lesioned patients disregarded future outcomes and instead were guided by immediate rewards. Patients without an intact OFC would disregard the high loss cards and continue to select from the high loss deck seemingly driven to select from Deck A to obtain the high gain cards[13,16,19].

2.1.1 Reversal learning

Unpublished data from Bechara *et. al.* suggested that if the punishment was immediate in the Iowa Gambling task, OFC lesioned subjects would perform similarly to controls and avoid the large loss deck[13]. This evidence suggested that their previous hypothesis regarding the insensitivity to future losses was incomplete. Evidence from non-human primate data and rats expanded upon the role OFC played in behavior. OFC was hypothesized to be necessary for behavioral modification following a change in a

reward-cue association; OFC was necessary for the modification of choices, reversing

learning, following a change in reward/cue association[19,27,39,53,85,119,122,127,189,199,274,307].

In a typical experiment designed to test reversal learning in rats, one type of

reward, a banana flavored reward, would be paired with one cue, a yellow light. Every

presentation of the yellow light would lead to a small banana reward. Another reward, a

larger cherry flavored reward, would be paired with a red light. Every presentation of the

red light would lead to the larger cherry reward. After a rat had learned the pairing, the

rewards would be reversed and paired with the other cues; the red light would lead to

banana reward and the yellow light would lead to the larger cherry reward. All rats with

an intact OFC would originally respond to the red light as it would lead to the larger

cherry reward; however, once the cues and rewards were reversed (yellow = cherry, red =

banana), the rats with the intact OFC would learn that the cues had reversed and begin

responding for the yellow light to receive the larger, cherry reward. In contrast, a rat

without a functioning OFC would persevere and continually select the red light even

though it now led to the smaller banana reward.

Rats[53,189], humans [19,85,119] and non-human primates[39,122,127,199,201,270] with lesions to

the OFC cannot perform the reversals. However, OFC lesioned rats[53,301], humans[85,119],

and non-human primates[13,19,199] have no issues learning the initial cue and reward

associations. Animals with bilateral OFC lesions are still capable of learning that a

stimulus predicts food and still make correct choices to acquire

13

food[91,122,123,179,180,295,301,306,307]. These results suggest that OFC is only responsible for learning when cue-reward pairs have changed and OFC is not necessary during initial learning.

A later experiment using a modified Iowa Gambling Task in humans by Fellows and Farah showed that if the decks were shuffled and the cards intermixed, effectively eliminating specific rules from each deck, OFC lesioned subjects and control subjects were no different in the cards they selected[83]. In addition in a variation of the Iowa Gambling Task designed for rats, the Rat Gambling Task, Zeeb *et. al.* showed that OFC lesioned rats had no impairments in correctly selecting the most advantageous option; without a rule reversal OFC damage did not affect the performance of the lesioned rats[139,140]. These experimental results in human and rats were consistent with the hypothesis that the OFC was necessary for reversal learning and that the Iowa Gambling Task deficits reviewed above were dependent on the planned reversals within the card decks rather than the inability of the human subjects to reject losses[75,84,138,189,249,273,279,280,295,301,305,306,338,374].

One study casts doubt on the role of OFC in reversal learning at least in non-human primates. Rudebeck et al demonstrated that prior lesion data in the non-human primate OFC may have been damaging fiber tracts[281,282]. This damage to the connecting fiber tracts may have resulted in the reversal learning deficits. This result conflicts with other results in non-human primates showing that OFC lesions result in reversal learning

14

deficits[122,127,199]. They indicated that if excitotoxic methods are used to accurately lesion

OFC, rather than aspiration lesions (removal of tissue), non-human primates did not lose

behavior flexibility in a reversal learning task. This result may not be generalizable to all

non-human primates; the authors indicate that their results may apply only to humans and

Old World monkeys (those primate species found in Africa and Asia which are more

closely related to apes).

This study may indicate most previous data regarding reversal learning in non-

human primates (and to a certain extent in humans) was an unintended result of extensive

damage beyond the OFC. In humans, most subjects with OFC lesions had quite extensive

damage that extended beyond the ventral medial OFC into more posterior and lateral

regions of the frontal cortex[15,16]. Damage to these more posterior medial areas would

have also damaged the fiber tracts consistent with the data shown in Rudebeck et al,

implying that reversal learning impairments in humans and rats may also be the result of

extensive fiber tract damage, rather than a specific function of OFC. Other data has

suggested the damage to limbic structures in the temporal lobe resulted in impaired

reversal learning[52,123,211] and that these structures in non-human primates and rats heavily

project to the posterior OFC (especially to the ventral and medial orbital areas in rats), via

the uncinate fascicle[45,46,149,292,339]. Consequently this may mean that previous reversal

learning deficits were a result of damage to the fiber tract connecting OFC to the limbic

15

structures in the temporal lobe. However, no experiments in rats have been able to dissociate OFC lesions and impairments in reversal learning.

2.1.2 Response inhibition and other behavioral tests of OFC

The previous research showing that rats, humans, and non-human primates could not perform reward reversals led to the hypothesis that OFC was responsible for response inhibition, withholding a response when conditions leading to a reward had changed. This description, while accurately describing the inability of subjects to correctly select a reward following a reversal, could not explain the fact that subjects had no issues withholding responses prior to the reversal of the cues/rewards.

More recently, devaluation experiments have proven that the OFC was not necessary for response inhibition[239,301,307]. In a devaluation experiment, one of the rewards is devalued by pairing that reward with illness after training. As before, the large cherry reward might be paired with the red light and the small banana with the yellow light. In a different environment, the large cherry reward would then be paired with lithium chloride which induced illness. OFC lesioned animals continue to respond to the red light and approach the large cherry reward, however, they do not consume the large cherry reward (as it had made them ill in a different location) (Figure 2-1). Rats[36,91,238,239] and non-human primates[10,122,123] had no issue inhibiting their response to consume the now devalued food. One shortcoming of reversal learning experiments was that they were confounding the learning impairments (such as the inability of animals to

learn or acquire a new behavior or to extinguish an old behavior) with the performance of

the animal using new information on a task. Animals rarely performed well following a

reversal. Was this result a case of impaired performance or impaired learning?
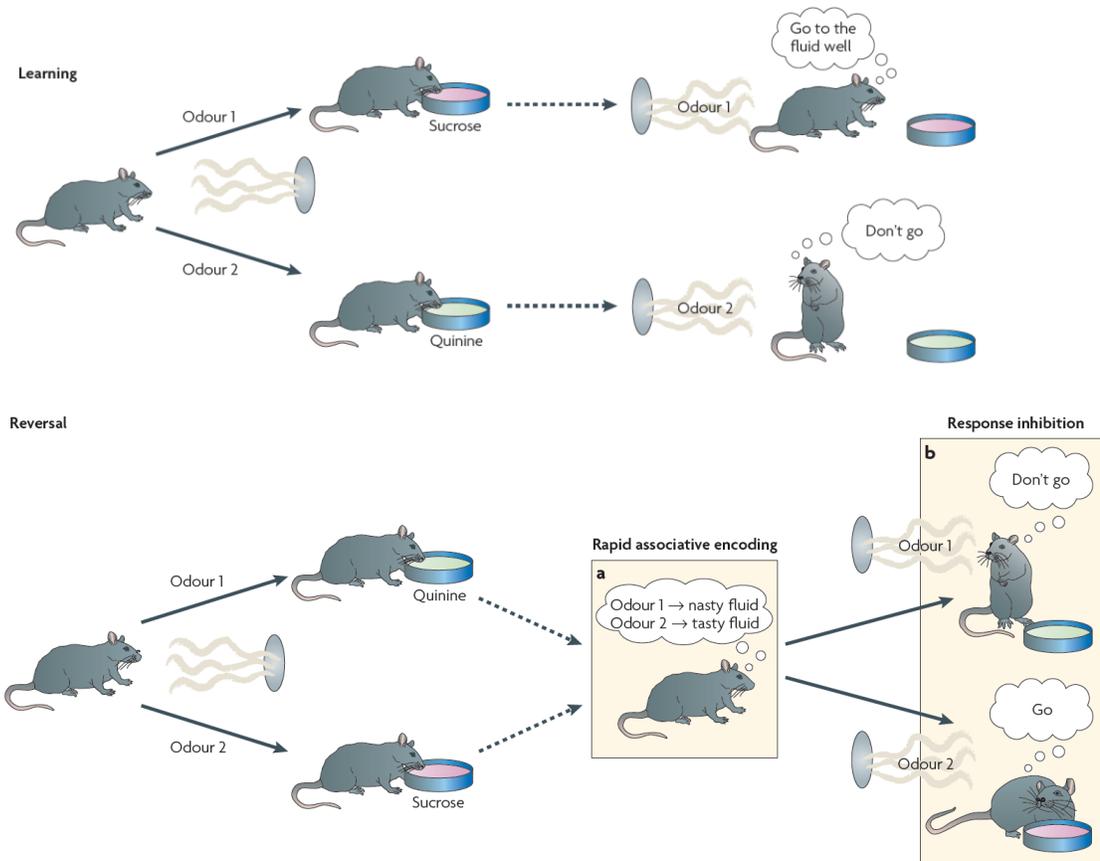
Figure 2-1 **Rats with OFC lesions fail to modify their behavior during odor discrimination reversal tasks.**
Top Panel. During training a rat learns to pair a specific odor with reward. Odor 1 is paired with sucrose. Odor 2 is paired with a quinine solution. When consumed the quinine solution makes the rat ill. When the correct odor is presented (odor 1) the rat will travel to the fluid well to receive the sucrose reward. If the other odor is encountered the rat will not travel to the reward well. The rat has learned that odor 2 predicts illness. Bottom Panel. Odor 1 is now paired with the quinine solution and odor 2 is paired with the sucrose solution. After several trials rats with an intact OFC will switch their responses; on odor 2 presentation the rat will travel to the well to receive the sucrose solution. On odor 1 presentation the rat will not travel to the

17

fluid well. Rats with OFC lesions will continue to respond to odor 1 even though it now predicts the quinine solution. Without a functioning OFC rats fail to modify their behavior (inhibit their responses) following cue or reward reversals. (used with permission[304])

The question then becomes how the OFC contributes to learning if it is not necessary for behavioral acquisition or for inhibiting reward consumption responses. This led to the hypothesis that OFC was representing some aspect of the relationship between the cue and the reward such as acquisitions of associative information[256]. Several experiments in rats have shown that the OFC is necessary for a cue to evoke a representation of the outcome, such as in a Pavlovian task where a cue is followed by a reward.

In a Pavlovian over-expectation task, a rat is trained that several different cues predict different amounts of reward, then the cues are combined to measure if there was increased response for the reward. Following the previous example, a red light would produce a large cherry reward and a yellow light a small banana reward. Presenting both stimuli can be hypothesized to lead to overexpectation; both the red and the yellow light would be presented simultaneously. If the rat had learned that each cue predicted a different reward, the presentation of both lights would imply that both the cherry and the banana reward were available. According to the theory that OFC integrated cue and reward information, if cues signaled that more reward was available (by combining different reward cues), a rat should respond faster and more vigorously for that reward. However, when both cues were presented the rat found that only the large cherry reward

was available. No banana reward was present even though the cues had signaled banana reward was present. On subsequent presentations of the red light, the rat responded more slowly. The spontaneous reduction in the response to red light indicated that the previous expectation of more reward from both lights (and the violation of that expectation) had subsequently lowered the rat's expectation of reward upon presentation of the red light. If the OFC was inactivated during the combination of the red and yellow lights, rats did not exhibit reduced response on subsequent presentations of the red light. This result suggested that the OFC was responsible for the integration of information from different cues predicting a new situation in which reward should have been increased[222,333,334]. The OFC has been found to be necessary to associate a cue that provided information regarding the expected reward with the actual reward, including delays[283], spatial categories[81,261], reward size[228,342] and the specific flavors of the reward[128,192].

McDannald et al[128,192,193] found that while OFC was necessary for representing the specifics of the outcome, such as identity (flavor), the ventral striatum was much more necessary for the formulation of general value of the expected reward. The OFC was necessary when learning an inferred value, which required that the information provided be integrated to form an estimate about a reward, but not when a stored, previously learned value was deemed sufficient. This indicated the OFC was important in the representation of specific expectations based on flavor, and differentiated itself from ventral striatum and other areas that may represent value. This research suggested that the

OFC had a much broader role in the representation and prediction of an outcome, instead of only being necessary for reversal learning and response inhibition. The OFC was hypothesized to have a different role in decision making, one that involved the formation of outcome expectations[302,305,306]; if I am in condition *X*, I can expect outcome *Y*.

An important recent result has shown that the OFC is necessary to learn inferred value[128]. Rats were first trained to associate cue *B* with a reward and cue *D* with no reward. Once rats were able to successfully respond to the cue-reward pair to receive reward, another cue, cue *A*, was paired with the reward cue; cue *A* preceded and predicted cue *B* which led to reward: *A→B→Reward*. A separate set of cues did not lead to reward, cue *C* preceded and predicted cue *D* which led to no reward; *C→D→No Reward*. The first cue preconditioned the rat to respond to the subsequent cue that was directly associated with reward. Rats reliably responded more for the cue that predicted the cue-reward pair, showing increased responses to *A* compared to *C*. When OFC was inactivated, rats no longer responded differently to *A* compared to *C*. This result indicated that without a functioning OFC rats could no longer make the association between the preconditioned cue *A* and the cue associated with reward, cue *B*. However inactivation of the OFC did not affect the rats' ability to respond for reward; rats still reliably responded to cue *B* which led to reward. Control rats continued to respond to cue *A* and approached the reward cup.

To test the effect of OFC inactivation on learning, a group of rats from the preconditioning experiment underwent an additional experiment[128]. Rats were tested on an inferred value blocking task. Blocking occurred when one of the preconditioned cues (*A* or *C*) was paired with a new cue (*X* or *Y*); if *X* was paired with *A* and because *A* already predicted reward, *X* would provide no additional information (*X+A→Reward* compared to the preconditioning *A→B→Reward*). Because *X* would not provide additional information, learning about *X* would be blocked. However, if *Y* was paired with *C* and reward was present, because *C* did not initially lead to reward, *Y* would supply information that reward was available (*Y+C→Reward* compared to the preconditioning *C→D→No reward*). As a result, there should be no behavioral responses when *X* was presented by itself and increased responding when *Y* was presented by itself. When OFC was inactivated, rats showed similar responses for *X* and *Y*. This result indicated that a functioning OFC was necessary to infer that the previous cue *A* completely predicted the reward. In contrast, control rats with a functioning OFC did not respond for *X* and instead only showed increased response for *Y*. These results showed that the OFC was necessary to learn in complex inferences when a new cue either predicted a future reward or provided no additional information about reward.

The preconditioning results as well as other results showing OFC is necessary for devaluation in non-human primates and rats[59,91,122,238,241,366] support the conclusion that OFC is responsible for more than reversal learning and is necessary for integrating

information about reward, cues and potential situations. Additionally, these results imply that OFC is not necessary for the representation of value; rats with an inactivated OFC still respond to the first order cue predicting reward, *B→Reward*.

If OFC were necessary for value guided decisions in this preconditioning task, then rats with an inactivated OFC would not have been able to reliably respond to cue *B* for reward. They would not have been able to associate reward value with cue *B* and respond to cue *B*. The necessity of an intact and functioning OFC in preconditioning, blocking and learning an inferred value is also supported by other data which suggests that OFC is necessary for responding to changes in reward identity[192,193,222,309]. Although, without cellular responses (single neuron, fMRI, and neuronal population activity) any conclusion regarding the exact function of OFC and how it does or does not represent value or identity is difficult to support.

2.2 OFC function: neural representations of value and outcome expectations

Orbitofrontal cortex may have a much more specialized function in decision making, one that primarily involves the generation of expected outcomes and the value associated with those expected outcomes. Early neural recording experiments showed that neurons in OFC robustly responded to reward delivery as well as the cue that came to

predict reward delivery. This view was consistent with the idea that OFC was responsible for tracking reward receipt and modifying behavior if outcomes were no longer matching what was previously learned. What role does OFC play in decision making and can the neural activity in OFC be entirely explained by the representations of outcomes or specific value?

2.2.1 Tracking outcome and reward

Recording from single neurons in non-human primates and rats offers insight into the neural activity at the time of a particular behavior. If the OFC's role in decision making is more complicated than simple reversal learning, then it should be expected that neuronal activity would not reliably follow reward reversal. Instead, the neuronal activity should reflect a combination of the cues necessary for reward receipt. Initial reports of single neuron activity seemed to agree with a simple role in reversal learning. A neuron would show preference for a specific reward by responding with increased activity to a particular reward or cue that predicted that reward. Single neuron studies reported that some neurons reversed their representation of reward following the reversal[59,91,272,338,362]. However, in one of the initial experiments in non-human primates that showed OFC neurons reversed their reward representations following reward reversal, only 22% of the neurons (70 of 317) varied their activity following the reversal[338].

In rats, 65% of neurons in the amygdala modify their activity during reward reversals compared to only 25% in the OFC[296,324]. Neurons in non-human primate

23

amygdala reverse their activity to match a change in outcome much more readily than OFC neurons[232]. If the majority of neurons in OFC are not tracking reward reversals; OFC must have a larger role than modulating reversals. Instead OFC must represent some other aspects related to reward and how those different variables contribute to or result in a decision.

The abundant evidence regarding the responses of OFC neurons to a variety of different aspects related to reward led to a different hypothesis, one that agreed with aforementioned findings in the lesion and inactivation studies; namely that OFC was responsible for encoding the relevant associations required to receive reward. This hypothesis again was incomplete since it meant that OFC neurons would rapidly reverse their activity following reward reversals. Single neuron recordings have shown that even if OFC neurons rapidly reversed their activity for reward, behavior in rats was much slower to reverse[323-325].

Although the previous idea that OFC was required for combining reward information and forming associations to facilitate rapid, flexible reversals was incomplete, it still captured a new idea in terms of OFC function; OFC neurons (and OFC as a whole) were encoding some aspect of the reward that allowed subjects to change their behavior to match the new cue and its associated outcome. Single neuron recordings indicated that many neurons in the OFC developed selective firing for reward. These neurons also started to become active during the cue that predicted the reward. Following

prolonged training, OFC neurons would fire for the cue that predicted the reward and again for the reward, often differentiating between several reward types. This activity occurred on every presentation of the reward and the cue[297]. This anticipatory activity occurred first in OFC before appearing elsewhere in the brain[297]. Single neuron recordings in non-human primates also showed that OFC neurons responded to the reward and the cue that predicted the reward[116,117,228,342,343,363]. While there are few direct single neuron recordings from human OFC, functional magnetic resonance imaging studies have shown increased activation of the OFC during cues and reward presentations. However, this may not have reflected individual neuron activity and instead only reports increased blood flow to that area of cortex[99,219].

The neurophysiological evidence suggests that the OFC is representing the expected outcome. In rat OFC many neurons fire differently in anticipation of different rewards and the parameters of those rewards. These parameters include the delay until the reward is received or the odor predicting reward[261] and different populations of neurons represent magnitude differences in reward[260,261]. Neurons representing magnitude for a reward do not fire for the other reward option[261]. Delay and reward identity are also represented but in different populations of neurons[260,261,327]. These different neuronal populations can be used to predict the expected reward fairly reliably[356,357].

2.2.2 Reward prediction error

While some data suggests the presence of reward-prediction-error information in OFC [218,331,340], other experiments suggest that OFC more closely tracks outcomes and value rather than prediction error[69,110,228]. If OFC were tracking reward prediction error, then OFC neurons would demonstrate similar activity to dopamine neurons found in the ventral tegmental area. Neurons in the VTA respond to unpredicted changes in reward by increasing their firing rate for unpredicted received rewards and decreasing their firing rate for unpredicted omitted rewards. If reward receipt does not change and can entirely be predicted, then neurons show no change in their responses[11,310]. OFC neurons continue to respond even when the reward was expected. Compared to ventral tegmental area (VTA) dopamine neurons[11,230,311,359], OFC neurons do not modify their activity when reward is received or withheld unexpectedly[262-264,297,308,326-328] (Figure 2-2). In addition, in the absence of OFC input, VTA neurons fail to signal prediction error, indicating that OFC is necessary for reward prediction error to be calculated but does not directly compute reward prediction error[335]. These results indicate that while OFC is not directly calculating reward prediction error, an intact OFC is necessary for prediction error to occur downstream in the VTA[334,335].

26
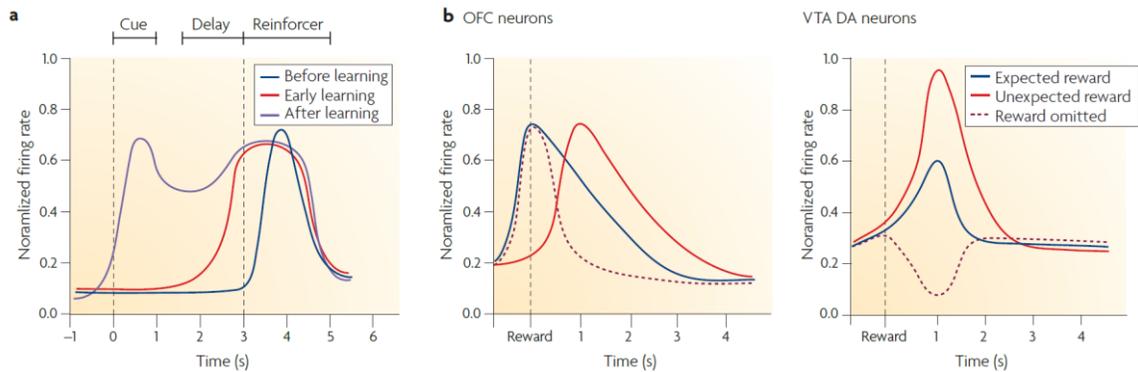
Figure 2-2 **Differences between OFC and VTA during reward receipt.**
a During early learning, OFC reward responses increase immediately following reward receipt (dark blue trace). As the animal learns the task, reward related activity begins to increase immediately prior to receiving the reward (red trace). Once the animal has learned the task, reward activity increases following the cue that predicts the reward as well as increasing after the reward has been received. **b** During unexpected reward receipt (red trace), OFC neurons increase their activity shortly after receiving reward, potentially reflecting the somatosensory response to the reward receipt. Once the animal begins to expect reward, OFC neurons increase their response immediately upon reward receipt followed by a gradual decrease in activity (blue trace). If reward is omitted unexpectedly without any prior training, OFC neurons still respond to the moment in which reward would-have-been received then immediately decrease their firing rate (dotted red trace). This indicates that OFC neurons are representing an expectation of reward. **c** In contrast to OFC neurons, VTA neurons strongly increase following reward receipt (red trace). With each subsequent reward receipt, VTA neurons decrease their firing rate (blue trace). Eventually VTA neurons respond very little to expected reward (blue trace). If reward is omitted unexpectedly, VTA neurons decrease their activity (dotted red trace). (used with permission[304])

2.2.3 Specific economic value

If OFC neurons are tracking different parameters of reward, it is possible that OFC neurons are representing some aspect of economic value, which is defined as how useful a reward is to the subject at that given instance. Past functional magnetic resonance imaging studies in humans alluded to the representation of value in OFC, specifically that cortical responses in OFC increased during the selection of potentially more valuable options[236]. Individual neuron responses in non-human primates have been found to track value as well[228].

27

Padoa-Schioppa and Assad found that neurons in the OFC tracked the economic value of different juice/reward offers[228]. Non-human primates were trained to select between different rewards, water and unsweetened kool-aid. The quantity and type of reward was represented by color and number of indicators seen on a screen. Offers were paired. The offer ratios varied. For example, they could select either 1 unit of water or 3 units of kool-aid. In another trial the offer might have been 3 units of water to 1 unit of kool-aid. If subjects behaved rationally and economically, comparing the different offers to each other, each subject should exhibit preferences for the different rewards in each offer pair. A measured indifference point would then indicate under which offer quantities each juice was equally preferred.

Consistent with this hypothesis, each subject showed distinct preferences between the offer pairs and had an indifference point where the two offers were equally preferred. Non-human primates behaved economically and generally equated 4(water):1(kool-aid)[228]. For example, if the subject chose 1 kool-aid when offered with 1, 2, or 3 water, but not when 5 water were available, then 1 kool-aid would be equivalent to 4 water. If 4 water and 1 kool-aid were offered both would be selected with equal frequency. Based on these preferences, non-human primates would select the offer, either kool-aid or water, that provided the best economic value to that subject.

Neurons recorded from the medial OFC increased their firing rates for a preferred reward matching the increased selection preferences for that same reward. For example, a

neuron that coded the economic value of water fired maximally when a higher value water offer was paired with a lower value kool-aid offer. In this example, if the economic value of water and kool-aid were equivalent at 1(water):2 (kool-aid), then the neuron's firing rate would not increase. As the amount of water offered increased when compared to the amount of kool-aid offered, the values offered were no longer equivalent. If 1 water and kool-aid were offered (conceptually 1 water was valued approximately twice as much as 1 kool-aid), the neuron's firing rate increased, potentially representing the increased economic value of the current offer (the water was more valuable economically to the subject than the kool-aid). In this example, the neuron would reach a maximum firing rate when the offer was 6 water:1 kool-aid. The maximal firing rate of that neuron would have indicated that the offer of 6 water to 1 kool-aid was the most economically valuable offer.

Approximately 18% of neurons recorded in OFC represented some aspect of chosen economic value. However, this report was at odds with representation of expected outcome mentioned previously. If the OFC was representing the economic value of the outcome then the OFC could not be representing information regarding the outcome expectation, such as the attributes that predicted a specific outcome.

The representation of economic value is not equivalent to the representation of expected outcome. Economic value requires that different outcomes be compared and evaluated against each other. The representation of expected outcome requires that the

components necessary to predict the reward are included in the neural representation; it does not assign a value to each of the outcomes. If OFC neurons predict the outcome based on the current cues, situation, and other available indicators rather than formulate a specific economic value between the available rewards, then OFC neurons should show a change in activity to any cue or object that predicts a certain reward rather than only responding for reward.

A key tenet of economic and rational decision making is that the choices between options be range adapting and menu invariant[97]. Range adaptation allows a representation of reward to dynamically scale to the magnitude of the reward options available. This change would allow a single neuron to represent economic value in multiple situations when comparing multiple different types of rewards. Menu invariance specifies that all rewards offered are reduced to a single value unit, a common currency that can be maintained and directly compared across all pairs and combinations of reward. The responses for a given reward remain the same even if that reward is paired with a different reward. Menu invariance may lead to stable economic preferences over time[223,227].

If the OFC was calculating the economic value, then the representations in OFC must be range adapting and menu invariant (neuronal activity would not depend on the rewards being paired). Range adapting neurons should scale their activity to the type and value of the rewards available. A range adapting neuron's responses will adapt to the

minimum and maximum scale of reward available. If the options range from one drop of juice to three drops of juice, a juice preferring OFC neuron would respond maximally for three drops of juice. If in a different instance, the offers are now three drops of juice and five drops of juice, the neuron would now respond maximally for five drops of juice and minimally for the three drops of juice (even though three drops was the previous maximum reward available). The neuron has adapted to the range of the available rewards. Menu invariance refers to the ability of the neurons to respond to rewards independent of the reward types being paired; grape juice responsive neurons will respond the same for grape juice regardless if the grape juice is paired with orange juice or apple juice.

Later work by Padoa-Schioppa in non-human primates indicated that neurons in non-human primate OFC did in fact adapt to the range of the values represented [226]. Neurons in OFC encoded values in a linear function. In addition, OFC neuron responses were menu invariant, neuronal activity was consistent according to the good being represented and did not differ based on the goods available [227].

It should be noted subjects in the non-human primate experiments had *substantial* experience with each of the presented offer pairs. They had been trained on the task for 6-8 months prior to the recording sessions[225,228]. A single recording session consisted of up to 2400 trials with the same set of reward pairs (on average 1500 per day). Under these conditions, it may be expected that the animal would become quite adept at selecting and

ranking the available rewards. Due to their extensive training, it is quite possible that the primates became experts at determining the tradeoffs between each reward offered. Additional data from rat neural recordings indicates that rats performing in a similar task with much less training showed stronger neural activity for outcome representations independent of value[260,261].

Results from human imaging experiments have been less clear about economic value representations in OFC. General value signals have been found in OFC[4,42,99]. However, these studies have not specifically correlated OFC activity with any measure of economic value. The results instead implied that OFC was representing rewards in a common currency, reducing each reward offer into a value unit that can be directly compared. Increased activity as measured by fMRI showed that higher valued items, regardless of the item types being compared, elicited stronger responses[50,130,177].

If value is being represented in OFC as a common currency, a neuron's firing rate should increase proportionally to the summed value of the combined rewards; if two reward predicting cues are combined indicating twice as much reward is available, a neuron should double its firing rate to reflect that the current reward is twice as valuable. Neural data from rats suggests that OFC cannot be encoding economic value because the firing rate of OFC neurons does not sum to reflect that the amount of reward may have doubled[333].

In an over-expectation task, Takahashi et. al. trained rats to expect different rewards from four different cues, $A1 \rightarrow 3$ Banana pellets, $A2 \rightarrow 0$ pellets, $A3 \rightarrow 3$ Grape pellets, $V \rightarrow 3$ Banana pellets[333]. When rats encountered combined cues for the first time, $A1+V$, OFC cue-reward responsive neurons spontaneously increased their activity. Neurons that had responded to each cue independently did not double their firing rate. Instead the new firing rate in response to the combined cues was much stronger than a doubling of the firing rate ($A1 \ \& \ V > A1 + V$)[333].

These results are inconsistent with the common currency and value explanation of OFC function. The summation of firing rates did not reflect the summation of the rewards. Instead the firing rate was much higher than twice the firing rate as would have been predicted by twice as much reward. In addition, because the firing rate for the combined cues exceeded the neurons previous firing rates for a single cue, the neural activity was not range adapting, in contrast to the data reported from non-human primates by Padoa-Schioppa[226]. This raises an important question; does OFC activity reflect economic value, expected outcome or something else entirely?

2.2.4 Representation of task parameters

The conflict between the specific economic value hypothesis and the outcome expectation hypothesis can be addressed by looking at neural representations of all the different aspects leading up to and predicting the reward including the reward representations themselves. If multiple, different variables that predict reward are being

represented in OFC neural activity then OFC is representing more than just specific economic value.

Single neuron activity in rats (as well as a recent study in non-human primates[178]) strongly suggests that OFC is representing the different aspects of outcome expectation rather than just economic value[81,142,261-263,296-299,327,328]. The neurons in rat OFC respond to many attributes and cues predictive of the reward such as the reward direction left or right of the rat[81,261], the odor that predicts the current reward[91,296,298,299,321,356,357], the specific flavor of the reward[298,327], the delay until reward[261], imagined outcomes if reward cues are combined signaling a potential for increased reward[333] and even the reward that could have been received, in rats[327,328] and non-human primates[1].

If OFC is not representing the specific value and is instead representing general value, including the salient aspects of reward, then the previous evidence of economic value representations in OFC should be sensitive to task settings (framing effects). Newer evidence from Padoa-Schioppa in non-human primates indicates that the OFC represents a more general value/outcome and is subject to framing effects, such as when the situation changes or the general state changes, the values modify[223,228,229]. In addition, recent evidence in human fMRI supports the hypothesis that the OFC is involved in the specific representation of outcome expectations[50,79,99,110,146,154,318,347], both real and imagined[33], that the OFC is sensitive to the task context[242] and that the OFC is more likely to represent the specific categories of the reward rather than just the specific

economic value[196]. All of this evidence from fMRI supports the hypothesis that OFC is responsible for the generation of outcome expectations.

The increasing evidence, both neural and behavioral, suggests that the OFC is representing specific, expected outcomes. This hypothesis predicts that the OFC is involved in reward expectations under specific conditions identified by any variables relating to the reward. These representations include the formulation of the counterfactual, an alternative reward not received, and the representation of regret[300]. If this is true, OFC neurons are capable of integrating the current information about expected rewards and representing previously missed rewards; the representation of the reward that could have been received. The additional knowledge provided by these representations in OFC will expand the current understanding of how OFC works to include the representation of imaginary past outcomes (a corollary to the neural representations of imagined future outcomes in rats[333] and the neural representation of cued, hypothetical outcomes in non-human primates[1]), linking the representations of outcomes to episodic future thinking and how errors in decision making modify future actions.

3. Defining regret

What differentiates regret and disappointment? Both require the knowledge that the current offer is less than desirable. In the psychology literature, disappointment is defined as the realization that an expected outcome is less than the previously conceived outcome[95,156-158,197,198,258]; more specifically, disappointment is 'sorrow caused by circumstances beyond one's control' or as the outcome decided by nature[15,24-26,31-33]. Economics defines disappointment as the value of an action being less than expected in the presence of a singular choice[22-24,173,174].

Regret is defined as the value of an action being less than expected in the presence of multiple choices. Regret is the realization that the action selected, of multiple available actions, has led to a value less than the alternative actions. In the simplest sense regret can be conceptualized as an emotion of the counterfactual, where the emotional response generated by regret results from the realization that the alternative choice would have been a better choice[22-24,41,56-58,173,174,286]. Regret has been described as the 'mental undoing of events' in the presence of an action/outcome that is less than the previously

available action would have produced[133-135,156,158,346]. Regret cannot occur without the prior counterfactual[22-24,105,106,173,174,287-289].

3.1 Counterfactuals

Counterfactual thinking was defined by its linguistic meaning, as 'contrary to the fact' [156,370]. Most definitions of counterfactuals were relegated to uses in defining linguistic arguments. Gradually, counterfactual thinking began to influence the formation of laws as well, with the principle *sine qua non*, Latin for "could the result have happened with the act"[157]. This form of counterfactual reasoning required that the alternative outcome be considered when determining the accountability of person's action.

Counterfactuals have been described as a set of 'conditionals that were based on the different similarity of the worlds, or states'[167]. In situations that could lead to different outcomes, there were a finite number of states that could be potentially described or imagined to which the current situation could hypothetically transition. The American philosopher David Lewis postulated that for counterfactuals to exist and be based on reality they must have been consistent with the laws of nature (of a given state, where the state included all the necessary information and possible actions a person could use to make a decision), that the potential states could in fact occur, and that the counterfactual in no way could violate the laws of nature[167]. The imagination of future states and consequences may be an inevitable consequence of conflicting choices and situations in

37

nature[105]. In relationship to animal behavior, counterfactuals could be associated with approach and avoidance behavior. If an encounter was unpleasant, a counterfactual representation of the alternative would allow an animal to switch its behavior to avoid unpleasantness[34,267,268]. This would functionally allow an animal to switch its behavior to a more appropriate behavior without having to experience the unpleasantness again. Deviations from the expected behaviors or expected results would cause the formation of the counterfactual especially if the previous outcome were controllable by the animal[200]. The counterfactual would result in the representation of the alternate expectation[35,132,135].

In order to represent counterfactuals, the alternative outcomes must be (or could have been) possible to achieve from the current situation, the decision maker must be capable of selecting the options available to achieve the alternative outcome and the decision maker must be able to identify the causal nature of the actions preceding the outcome[23,51,105,106,156,157,171,173,174,267,268,289]. By definition the counterfactual is a causal statement acknowledging that an action will lead to a alternative, specific outcome[136]. The counterfactual simulates the outcomes of potentially different results had a different action been chosen[155]. For example if two dice were thrown, there would be 35 different 'worlds' possible, only one of those worlds could be physically represented, all other representations would be the counterfactuals to the current world[155]. The construction of the counterfactual occurs by converting potential outcomes into possible precursors thus regenerating the expectancies[268].

38

The representation of the counterfactual must be present to compare alternative outcomes to each other[98]. In decision making, counterfactuals are necessary for the experience of regret. Without the representation of the counterfactual, there are no hypothetical outcomes to compare to and therefore regret cannot exist.

3.2 Psychology of regret

In the simplest sense, regret is the recognition that the current situation is worse than an alternative situation. This definition of regret first arose from the conceptualization of regret as a pure emotion, the recognition of the emotional response following a bad series of choices[20,278]. Only later was regret defined in terms of actions and outcomes both real and imagined. The definition of regret in terms of actions and outcomes would lead to the economic definition of regret.

The word regret is derived from Scandinavian and Norse words, *regrete* and *grata* which is defined as 'to weep'[156]. Early mentions of regret in classic literature and psychology describe regret purely as an affective emotion[38,95,156,158]. The transition of regret from an emotional descriptor to a cognitive state began in the 1940's. Burks postulated that regret is the recognition that the actions within a given state of nature did not satisfy the expectation and that alternative actions would have produced a more desirable outcome, thus satisfying the expectation[38].

If regret is assumed to occur in a specific instance of nature, then the experience of regret would require that the alternative outcome is known as well as the proprieties of that outcome (antecedents of) and how they relate to the current situation. Any decision between multiple options would therefore result in outcomes with 'real potencies', or real-world implications. The state in which the regret experience occurred was described as an emotional recognition[23,102,174,294,312].

Landman proposed that regret was generated more readily in instances that produced less than ideal events[156]. Regret was hypothesized to be a larger function of inductive thought involving facts and counter-facts, from a given set of cases to a wider set of cases. Though she reasoned that regret, while depending on the counterfactuals 'was a distinctly psychological phenomenon', regret could also be conceptualized as an emotional phenomenon. Others had specified that regret depended upon the representation of alternative selves, that contained the complete cognitive and emotional information in a hypothetical state[58,62,66,95,197,258,285].

Regret has been suggested to be a consequence of rationality, a measure of how realistic a subject's goals are in relation to the current reality[129]. Thus, if regret is a consequence of rationality, then a subject behaving rationally would avoid the experience regret [105,156]. However, if a subject experiences regret, then a subject has made an irrational decision; therefore regret is the failure to rationalize one's behavior. If economics can be used to operationalize what is rational and what is not rational behavior, then it is possible to describe regret using economics. An economic definition of regret provides an explicit set of circumstances under which regret can occur. With an explicit set of conditions that should induce regret, testing for the existence of regret becomes feasible.

3.3 Economics of regret

The economists Loomes, Sugden and Bell stated that regret played an important role in rational, economic decision making because many subjects behaved in ways that violated rational decision making. Regret theory could potentially explain violations of rational behavior such as why a subject would exhibit preference reversals; initially preferring a chocolate to a grape, but later preferring a grape to a chocolate. Subjects could reverse their preferences if they regretted choosing a chocolate to a grape, subsequently selecting a grape when offered a grape and a chocolate again. Introducing regret into the decision making process allowed regret to modify future choices by comparing the difference between the possible rewards and the currently received rewards; a suboptimal (and seemingly irrational) choice may be selected to avoid regret.

Regret is defined in economics as the difference in value between two known actions/outcomes in a specific situation, where the outcomes and results of the actions are assumed to be known or capable of being conceptualized[23,174]. Regret Theory assumes that an agent is rational and maintains the following:

1) There exists a usefulness function that is capable of being represented without an agent choosing and experiencing the outcome (choiceless), where the value of the options is assigned independent of the choices and prior to obtaining the outcome of the options (the choice has yet to be made); the value of the goods is anticipated prior to the actual decision.

2) The usefulness function is defined and independent to (and prior to) the choice of the agent.

3) The agent has knowledge of what all action/outcome pairs will be and their relative value in the given situation.

4) The measure of regret is dependent only on the choiceless usefulness of the two possible consequences (the counterfactual, 'what might-have-been', and the reality, 'what is').

5) The regret function assigns a value to every possible action/outcome pair and the value assignment is explicit to the state the agent is in; any future states cannot be assumed to have the same regret function unless their conditions exactly match the initial state.

For a specific situation, the utility in regret theory is defined as *Utility(choice₁,choice₂)* where utility is the usefulness of the outcome in the current situation and *choice₁* is selected simultaneously rejecting *choice₂*. The best possible action minimizes the regret function defined as the *value(choice₁)-value(choice₂)* where *R(.)* is the regret function [23].

$$Utility(choice_1, choice_2) = v(choice_1) + R[v(choice_1) - v(choice_2)]$$

If an agent were to choose *choice ₁* over *choice ₂* in an explicit state of the world the resultant value *v(choice ₁)* would occur and the counterfactual value/action would be

represented as *v(choice $_2$).*This formulation lets the regret function represent a real value for each possible increase or decrease in the value of an option. The regret function can increase or decrease the relative value of the selected action.

When translating Regret Theory to the economics of choice, the experience of regret relies on the probabilities of receiving the specific outcomes in that situation being known. Regret is defined as the difference between the payoff on a given trial and the maximum (expected) payoff. In this context regret again modifies the future actions by either increasing or decreasing the value of a choice, where the value of a choice is highly correlated with the prior experience of regret, higher potential regret over the loss of an outcome associated with a choice increases the willingness to select that choice and its resultant outcome (the regret function increases the value of the action that leads to the outcome)[171,174,175,258,375,379].

Economic definitions of regret rely on simulation heuristics, formulating the potentially experienced regret prior to a decision and minimizing future regret by either prospectively avoiding situations that would induce regret, or by modifying actions following a regret experience to avoid future regret[23,174,375,379]. Prior to the explicit definition of regret, regret was identified as a special form of decision-making; regret was the result of mental simulation and required the comparison of states and expected outcomes[133]. If the outcome violated the expected outcome in the subsequent state, regret was experienced. This early economic definition of regret could not be differentiated

from disappointment. Therefore regret as defined later by economists Bell, Loomes and Sugden differentiated between regret and disappointment, regret being a comparison between two outcomes where an error was the result of an agent's action and disappointment was a singular outcome that was not the fault of the agent[23,174].

The economic definition of regret agrees with psychological definitions of regret, both definitions state that regret is unique to each experience[23,156,157,172,198,258]. Even if a subject is in a similar situation again, the regret experienced will not be the same; 'If I was in that situation again, I would probably make a different choice'[23,174]. The regret function cannot be expected to be the same across multiple situations; each experience leads to a different experience of regret and potentially different choices[23,174].

By using these definitions of regret, the effects of regret on decision making in economics can be identified and separated from disappointment. The modifications introduced by regret theory to rational decision making address how a subject's preferences may shift after experiencing regret and can explain irrational choices. Once the counterfactual exists, the experience of regret is possible. Disappointment and regret are differentiated through the representation of the counterfactual. Regret is the representation (the counterfactual) of the more valuable alternative action and the associated outcome had the agent selected the more valuable alternative option.

If disappointment is strictly defined as the outcome being worse than expected determined by conditions beyond the agents' control then disappointment can be differentiated from regret through the representation of the counterfactual. These definitions of regret and disappointment make it possible to design and conduct experiments that identify:

- the counterfactual
- regret
- disappointment

This raises some important questions; can the counterfactual and regret be identified in rats by using the economic framework outlined above and does OFC play a role in the formulation of the counterfactual and regret agreeing with counterfactual representations in non-human primate OFC neurons[1] and that the activation of OFC during the experience of regret in humans[56]?

4. Identifying reward representations on Multiple-T Left-Right-Alternate: evidence of the counterfactual in rats

4.1 Introduction

When faced with a difficult decision, rats and humans sometimes pause and orient back-and-forth towards options or paths, a behavior termed *vicarious trial and error* (VTE) [125,150,210,231,341]. During VTE and similar behaviors, neural signals in hippocampus and ventral striatum show evidence of covert decision-making processes [125,126,351,353]. Here we report that neural representations of reward in the OFC of behaving rats increased following VTE events at a decision-point, implying representation of the expected rewards during an internal, self-initiated decision.

Neural representations of reward in rodent OFC increased immediately following each re-orientation, implying a transient representation of the expected outcome following self-initiated decisions. Upon reaching reward locations and finding no reward (having made an error), OFC representations of reward decreased locally indicating a disappointment signal that then switched to represent the unrewarded, non-local, would-have-been rewarded site. These results illustrate that following a decision to act, neural ensembles in OFC represent reward, and upon the realization of an error, represent the

reward that could have been. We report that, at reward locations on error trials, when no reward was present, neural representations of reward in OFC of behaving rats decreased, indicating a neural correlate of disappointment.

Here, we report that, when faced with a lack of delivered reward (disappointment) after making a decision (implying the potential for regret), neural representations in rat OFC switched from encoding the local, unrewarded site, to encoding the non-local would-have-been-rewarded site, representing a neural signal of the counterfactual necessary for regret. In summary, following a decision to act, neural ensembles in OFC represent the expectation of reward, potentially guiding future evaluative processes, and upon the realization of an error, represent the reward that could-have-been.

4.2 Methods

4.2.1 Animals

Four Fisher Brown Norway rats aged 10-12 months at the start of behavior were used in this experiment. Prior to training, rats were handled for two weeks. On the last six days of the two week period, normal Teklad pellets were replaced with flavored pellets within the rats' home cage. Rats had access to 15 grams of white (unflavored), fruit-flavored, or banana-flavored, presented in random order during handling. Each flavor was presented once per day no more than twice during the six day sequence. Rats were housed on a 12 h light/dark cycle and training/probe trials occurred during the same time each day. During testing, rats were maintained at roughly 85% of their free feed weight. Rats had access to water at all times. All training procedures were in accordance with the National Institutes of Health guidelines and approved by the Institutional Animal Care and Use Committee at the University of Minnesota.

4.2.2 Behavior: the Multiple-T-LRA task

We trained 4 rats on a continuous loop, multiple choice, maze task (Figure 4-1)**.** Reward was delivered under Left (L), Right (R), or Alternation (A) schedules (MT-LRA) [30,103]. The Multiple T maze consisted of a figure-8 topology, with a central navigation sequence leading to a high-cost choice point. The choice-point led to two, different return rails. Each had two feeders (Med-Associates, St. Albans VT) and potentially provided 2x

49

45mg food pellets (Research Diets, New Brunswick, NJ, USA) each. The navigation sequence consisted of three low-cost choice points, at which the rat could turn around if he made a wrong choice. After a choice at the high-cost choice point at the end of the navigation sequence, the rat had to continue down the return rails before coming around for another lap. The left return rail provided banana-flavored pellets at the first feeder-site, and unflavored (white) pellets at the second feeder-site; the right return rail provided fruit-flavored pellets at the first feeder and unflavored (white) pellets at the second feeder-site. During training, if a rat tried to run backwards on the navigation sequence or backwards from the second feeder to the first feeder on one of the return rails or from one feeder side to the other across the top rail, the rat's path was blocked by the experimenter with a PVC pipe. However, by the recording sessions, rats never turned around and did not need to be blocked.

The navigation sequence remained constant within a day, but changed from day to day. Whether reward was provided on a return rail or not depended on the choices made by the rat. Three reward contingencies were used: (L) turn left for reward, in which the left return rail always provided reward and the right did not, (R) turn right for reward, in which the right return rail always provided reward and the left did not, and (A) alternate for reward, in which the return rail not previously visited was rewarded. In the alternation (A) contingency, the first return rail visited was always rewarded on a given day. All

reward site locations and flavors at each reward site were constant across all sessions. On

a correct lap, reward was always presented. On error laps, reward was never presented.
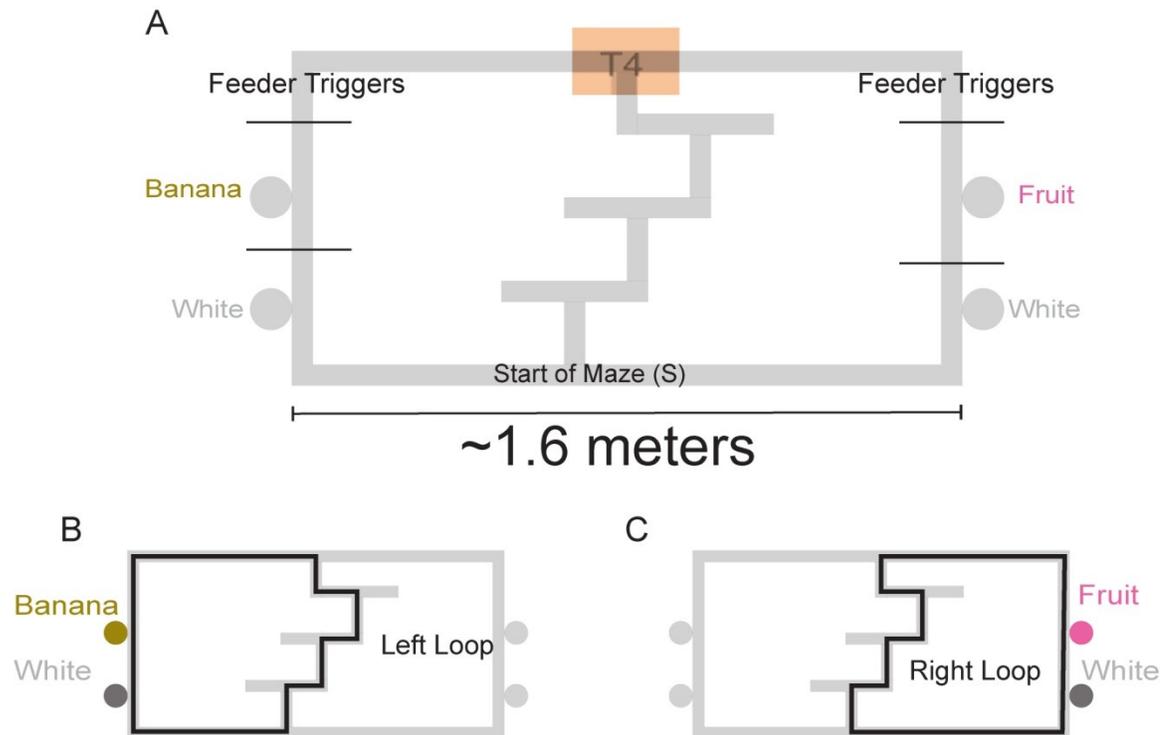
Figure 4-1 **Multiple T–LRA Task Behavior.**
Rats were placed at the start of the maze (S). Reward was delivered as animals crossed the reward-trigger
lines. Reward trigger zones were spatially defined and did not vary from day to day. On any given the day,
if the rat chose correctly, once he passed into the zone, pellets were delivered. Each side had two feeders,
one flavored, one unflavored. Flavors at reward sites remained constant across all sessions[328].

Rats ran one 40 minute session per day. Contingencies were presented in a

pseudorandom order across days. The rat did not receive any cues informing it of the

rewarded contingency. On each day, the rat was placed at the start of the maze and

allowed to run through the navigation sequence and turn left or right at the final choice

51

point for reward, but it did not know which contingency it faced. Rats were trained for an average 24 days on this task before surgery, until they were performing all three contingencies (L, R, A) reliably.

Following surgery, rats were allowed to recover for 2-4 days, during which they had free access to food and water. After 2-4 days, rats were returned to the Multiple-T-LRA task. Recordings commenced when the rats returned to running a number of laps comparable to pre-surgery. To acclimate to the additional weight of the tether and hyperdrive implant before the probe sequence began, the rats were trained for several more days while tetrodes were advanced to target sites.

Following adaption to the increased running weight and achievement of large ensemble sizes, rats began the 6 day probe sequence. A probe day entailed a change in contingency after 18 to 22 minutes. Thus, the rat faced one of the three contingencies (left, right, or alternation) for approximately 20 minutes, and then faced a new contingency for the second 20 minutes. Rats were not removed from the maze at the switch, nor were they signaled as to the switch. During probe sessions, the fourth T was always aligned to the middle of the top rail. This ensured that the path length from the high-cost choice point did not change to either the left return-rail or the right return-rail. We ran six probe days: left/right, right/left, left/alternation, right/alternation, alternation/left, alternation/right. Each rat saw all six probe days, but the order of the six probe days was randomized between rats.

4.2.3 Surgery

After an initial phase of pre-training and after the rats had reached behavioral criterion, rats were chronically implanted with 14 tetrode-hyperdrives (Kopf). Targets were the ventral OFC, AP +3.5, ML +2.5. Implants were alternated on each rat, such that two implants were right centered and two were left centered. Surgical procedures were performed as described previously [125]. All tetrode locations were histologically verified to lie in the ventral orbitofrontal cortex (Figure 4-2).
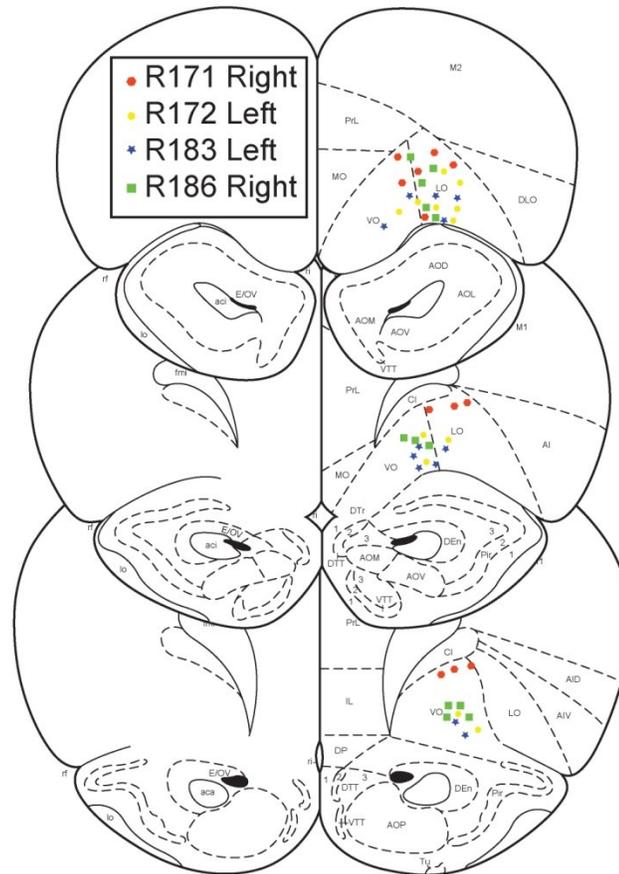
Figure 4-2 **Histology.**

All recordings were confirmed to be in orbitofrontal cortex. The implantation sites were alternated. All recordings were in the lateral orbital frontal cortex, which corresponds to more lateral aspects in human and primate neuroanatomy [152,328].

4.2.4 Data acquisition

Rats were tracked by an overhead camera system via Neuralynx (camera 1). A second camera (camera 2) was centered on T4 to increase positional recording accuracy and to serve as a set spatial window for high-cost choice point (T4) passes. Data for the calculation of orient-reorient behavior, defined as VTE, was taken exclusively from the spatial window defined by the second camera at T4. Before surgery, rats were tracked from an LED attached to a in-house-built backpack; after surgery, rats were tracked from LEDs built into the headstage attached to the implanted hyperdrive.

Unit recording

Unit and local field potential activity was monitored as the tetrodes were advanced. Once the tetrodes began to approach approximately 4.2mm in depth, tetrodes were advanced no more than 80 microns per day to allow the tissue to stabilize. Once LFP and units were stable, tetrodes were moved to find the largest possible ensemble.

We recorded neural activity on a 64 channel Cheetah recording system (Neuralynx, Bozeman MT). Session data were recorded to disk and units were identified offline using MClust 3.5. Pre-clusters were formed automatically using Klustakwik. During recordings the position of the rat was tracked using colored LEDs on the headstage. The position was time stamped and recorded in Cheetah by the overhead

camera and a second camera centered on T4. A total of 712 cells were recorded. Cell yields were distributed across four rats; R171: 173 cells; R172: 252 cells; R183: 137 cells; R186: 150 cells. Because the recordings were conducted over multiple days it is difficult to rule out that some cells may have been recorded multiple times. Because results were consistent across multiple rats, we remain confident that our results are not due to re-sampling. Analyses that are over-conservative for re-sampling also produce similar results.

4.2.5 Data analysis

Behavior: Path Linearization

In order to compare multiple sessions of differing paths, the 2D tracking data was mapped to the closest point in a 1D path [293,352]. Each path had 7 landmarks (Start of Maze, T1-T4, both feeders) with a set number of points between landmarks. The data between each landmark was assigned to a fixed number of spatial bins. Because T4 was centered along the top rail on probe sessions, the path length from T4 to the first feeder on either side was equidistant on all probe sessions.

Behavior: Laps

A lap was defined as a complete loop from the start of the maze to the middle of the bottom rail prior to the start of maze zone. Lap times were defined as the time elapsed from when the rat crossed into the navigation sequence, passed through the feeder zones

and finally crossed back into the start of the maze zone. Laps that did not include feeder passes, either correct or incorrect, were excluded. In practice this only occurred when the 40 minute session ended with the rat between the start of the maze and T4. On correct laps, the rat was rewarded by 2x pellets at each feeder. On error laps no reward was presented.

Behavior: Vicarious-trial and error behaviors (VTE, zIdPhi)

In order to quantify VTE behaviors, we measured the integrated angular velocity (*zIdPhi*) through the choice-point pass [231]. A choice-point (T4) pass was defined by entry and exit times through the field of view of camera 2. First, the velocity of the animal was calculated using a modified, discrete-time adaptive window for velocity estimation [124]. We used the change in the velocity vectors, *dx* and *dy*, to calculate an angle of motion, and then used the velocity estimation algorithm to calculate the momentary change in angle, *dPhi*. Integrating *dPhi* over the duration of the choice point pass, defined by the box in Figure 4-1, resulted in a measure of *IdPhi* which we used to quantify the behavior on a single lap. The *IdPhi* scores were normalized by z-scoring across laps for each session for each rat. The z-scored measure, *zIdPhi*, was compared across all animals and sessions. This measure proved to be a reliable assessment of the rat's behavior see Figure 4-3 [231]. The behavior we observed, previously classified as VTE, was quantitatively

56

defined as *zIdPhi*>0.5, during which rats reliably demonstrated visible orienting-reorienting behavior [210,341].

<u>Behavior: Reorientation events</u>

Reorientations were identified at times when the rat performed an abrupt change in direction at T4 (Figure 4-3, ***black arrows***). These events were clearly visible in the tracking data.
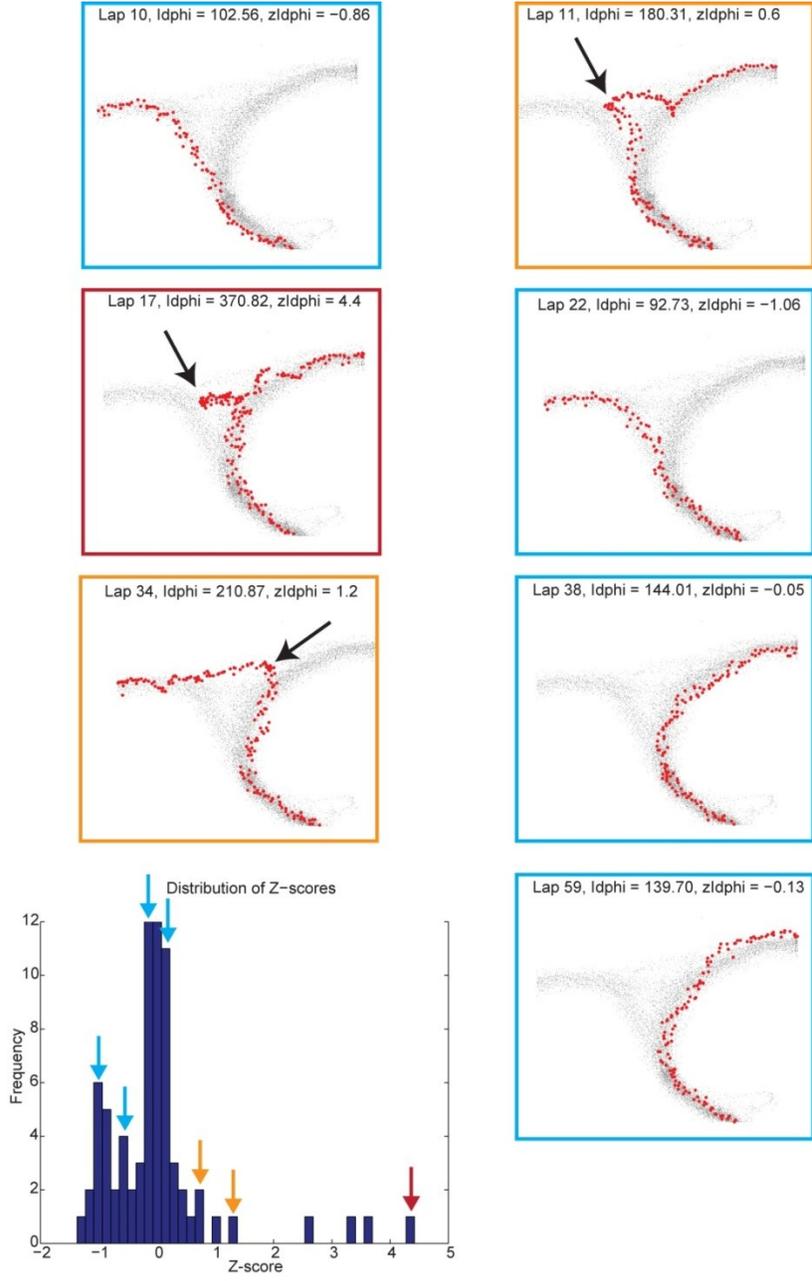
Figure 4-3 **Identifying vicarious trial and error (VTE) events.**
The colored boxes refer to the VTE distribution in the bottom left corner as scored by zIDphi. Each pass
through camera 2's field of view is shown in light grey, individual passes (once per lap) through the field of
view are highlighted in red. Low zIdPhi passes were the most common and demonstrated little behavior
indicative of zIdPhi (Blue squares). High VTE passes were less common however they demonstrated large,

head swings from one direction to the other (Yellow and Red boxes/ arrows on the histogram on bottom left, black arrows on individual behavioral traces from camera 2). [328]

4.2.6 Single-unit analysis

<u>*Reward sensitivity*</u>

To determine the reward responsivity of a unit, we first calculated a peri-event-time-histogram (PETH) from -1.5 seconds to 3.5 seconds following feeder triggering using a time step of 100ms. In order for a neuron to be classified as reward responsive, we compared the number of spikes fired during the 2.5 seconds following triggering of the feeder (reward delivery, 0 seconds to 2.5 seconds) to 500 bootstrapped samples of the same duration aligned to random times throughout the session. If the activity during the reward epoch was significantly different than the bootstrapped samples the cell was classified as <u>reward responsive</u>.

<u>*Decoding*</u>

All decoding was performed using a one-step Bayesian decoding method with a time-step of 250ms [380], measuring the probability that the neural ensemble decoded to a given spatial location on the maze. Only cells with >100 spikes and data sets with > 14 cells were included in the analyses. We first calculated the linearized tuning curves for each cell during each session. Training sets were extracted from steady state

performance. To control for tautology, any test sets used were excluded from the tuning curves in the training sets.

Shuffled control data.

In order to ensure that the non-local decoding seen in the results does not arise from random firing, we tested our decoding algorithm using tuning curves derived from actual firing patterns and shuffled spike trains. Shuffled spike trains preserves cell identity and the first-order firing statistics of each cell. This allowed us to test whether increased random activity during pauses, at the choice point or at the feeders, could account for the increased decoding to the reward locations.

Decoding p(Reward)

To construct *p(Reward)*, each side of the maze was linearized to control for differing lengths in the central portion of the maze (T1 through T4) on different days. This produced two separate loops, left and right (Figure 4-4**a,b**). Once the maze was linearized, we calculated the spatial tuning curves for all cells on left and right portions of the maze. Because rewards are only delivered at specific locations on the maze, cells which fire primarily in response to reward will drive the spatial location on each loop towards the reward locations. On this task, reward reliably occurs at specific locations on the maze. For example, a cell that fired for banana flavored reward would fire most on

the left loop (Figure 4-4**b**). It is important to note that successful decoding to reward

locations does not imply that spatial information is encoded in OFC ensembles. Rather

we are using a spatial algorithm to provide information regarding the presence or absence

of reward. During VTE, decoding was calculated using all cells.

p(Reward) = Probability of being at a feeder given the spikes observed
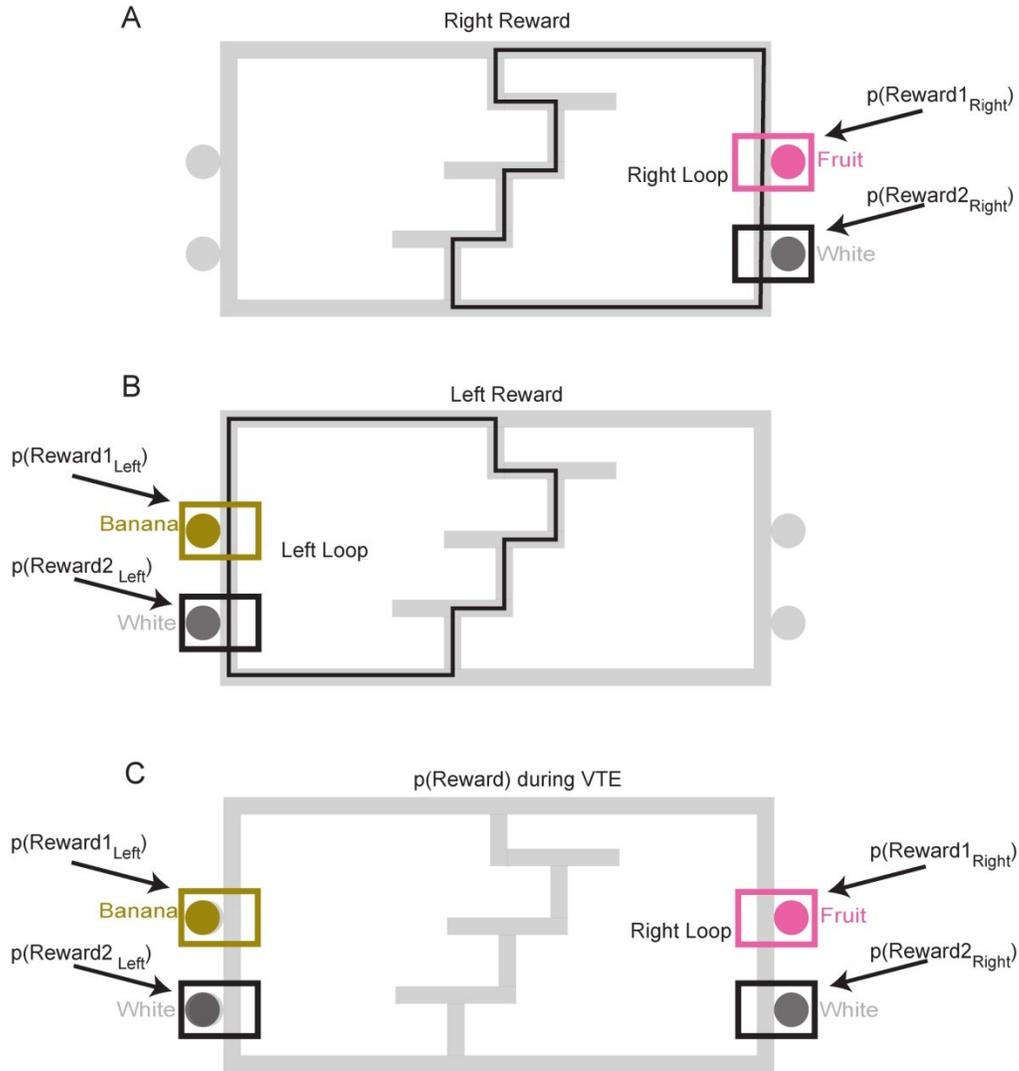
Figure 4-4 **Decoding Locations.**
Probability of reward refers to the probability of the rat being at the rewarded site given the spikes that are observed. Because reward sites are fixed in space, a decoding algorithm initially designed to determine predictions in space also reflects the probability of the rat being at a reward site. Since neurons are quantified as being reward responsive based on the presence or absence of reward, the spatial decoding algorithm allows us to the measure the likelihood of receiving reward when the animal is at the reward site, *p(Reward)*. It should be noted that we make no claims that OFC represents space. When calculating the probability of reward (*p(Reward)*) for the rightward loop, only the two right feeder locations (**a**) are considered in the calculation of *p(Reward)*. Conversely, when considering the leftward loop (**b**), only left

62

side feeders are considered in the calculation of *p(Reward)*. During VTE events, *p(Reward)* is averaged across all four feeder locations (**c**). [328]

Decoding at VTE events

Entry and exit times through the T4 choice point were recorded for each pass using the field of view of camera 2. Orient-reorientations were noted. On instances where multiple orient-reorient behaviors were observed, we calculated *p(Reward)* for each event. All cells were used during decoding at VTE events.

Counterfactual Representations

Because different sets of OFC cells responded to each of the four reward-sites, it was possible to measure *p(Reward)* for a given site. As above, spatial tuning curves for the entire maze were defined for each cell for each loop, right and left, and then separate *p(Reward)* measures were taken from the decoded posteriors at each feeder site (Figure 4-5). During counterfactual calculations only reward responsive cells were used for decoding *p(Reward)* (see Figure 4-7).
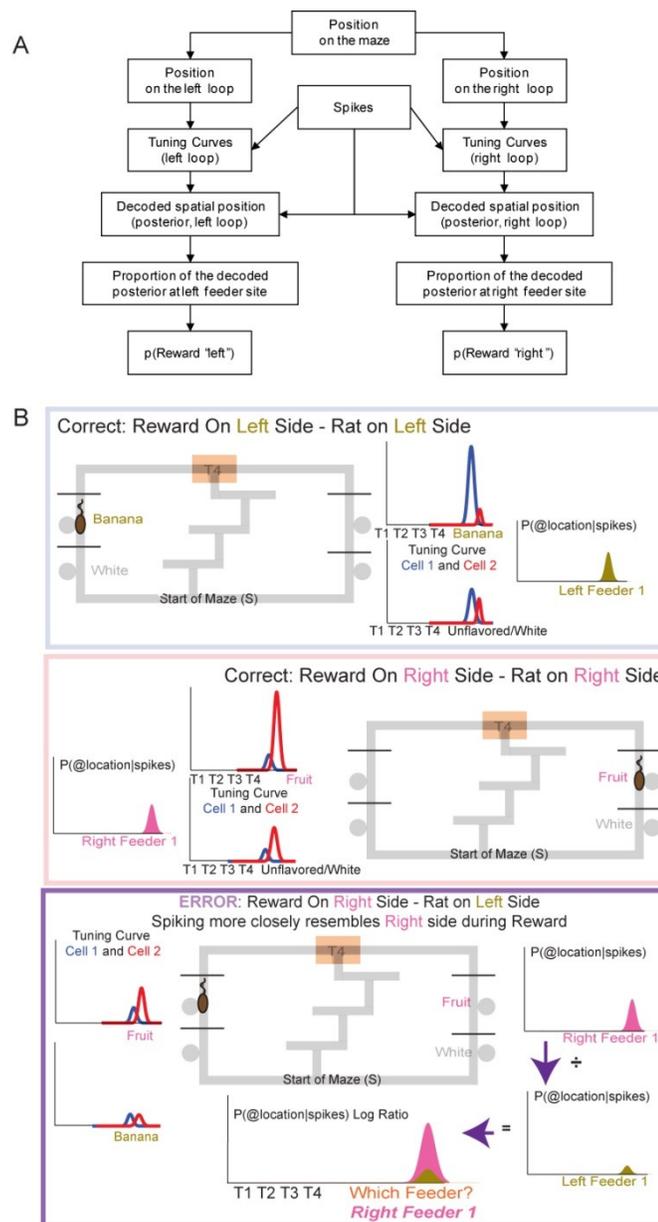
Figure 4-5 **The Decoding process.**
a Any decoding algorithm consists of three parts: (1) a set of tuning curves which defines the expected firing as a function of the variable in question, (2) a set of spikes or firing rates (In order to prevent a tautology, the spikes used in (2) should not be taken from the same set used to generate tuning curves in (1). We accomplish this by a leave-one-out approach in which the tuning curve definition does not include the lap in question), and (3) the posterior probability calculated from (1) and (2). We use two independent decoding processes – one in which the tuning curves are defined from spatial position on the leftward loop

only, and the other in which the tuning curves are defined from spatial position on the rightward loop only. Each of these decoding processes provides us with a posterior probability of spatial position around the maze. It is important to note that we do not require that OFC cells be spatial in order to derive spatial decoding posteriors. Because rewards are only delivered at specific locations on our maze, cells which fire primarily in response to reward will drive our spatial decoding to those reward locations (the feeder sites). A cell that fires in response to any reward will drive decoding to all of the reward sites on the maze; a cell that only fires in response to banana-flavored pellets will drive decoding to the first-left-feeder, etc. As shown in Figure 4-8 we have a diversity of cells which respond to subsets of feeders. We define "reward decoding, *p(Reward)*" as the amount of posterior probability that has been spatially located to those feeder sites.

b Calculating the counterfactual**.** For two cells with differentiable tuning curves, cell 1 and cell 2, we can use the activity of both cells to determine where on the maze the firing rate best represents the location of the rat. In the top panel, cell 1 prefers banana, and is more active at left feeder 1 when the animal receives banana flavored pellets. Cell 2 does not respond to banana flavored pellets. To calculate the decoding we combine information from cell 1 and cell 2 and ask where on the maze does this activity represent? If cell 1 is very active and cell 2 inactive, combing that information increases the probability of decoding to left feeder 1, where the rat has received banana flavored pellets. This represents an increased probability of decoding to the local reward, *p(Reward$_{same}$)*. If cell 2 is now active and cell 1 inactive, we would expect that the probability of decoding to reward would now shift to right feeder 1 (fruit) where the rat just received fruit flavored pellets (middle panel). Again this would increase the probability of decoding to the local reward site, *p(Reward$_{same}$)*. Because both cells differentiate between feeders, we can determine the probability of decoding to a non-local reward as well as a local reward.

If the rat is at left feeder and does not receive reward, cell 1 no longer increases its activity. Instead cell 2 increases its activity. We again calculate the probability of decoding. Because cell 1 is inactive, we see very small probability of decoding to the local reward site. However, cell 2 increases its activity. As a result the probability of decoding to the-would-have-been rewarded site, increases (i.e. *right* feeder 1/fruit). To compare both these values, we compare the log ratio of the all the activity and the decoded probability on the local side versus all the activity and the decoded probability on the opposite, non-local side and ask, which side has a higher probability? [328]

p(Reward$_{same}$) was defined as the *p(Reward)* for the side on which the rat

currently was located, while *p(Reward$_{opposite}$)* was defined as the *p(Reward)* for the other

side. When constructing *p(Reward$_{same}$)* and *p(Reward$_{opposite}$)* we created two training sets,

same and opposite. Same side training sets included correct laps from the same loop that

the rat was currently on. Opposite side training sets included correct laps from the

opposite loop than the rat was currently on. If, for example, a rat was at right feeder 1 and

received reward, $p(Reward_{same})$ would be defined from correct rightward tuning curves, while $p(Reward_{opposite})$ would be defined from correct leftward tuning curves. On error passes we compared *error, non-rewarded* passes to the two test sets, $p(Reward_{same})$ and $p(Reward_{opposite})$. On correct passes we compared *correct, rewarded* passes to the two test sets, $p(Reward_{same})$ and $p(Reward_{opposite})$.

As above, these training sets were selected from correct laps, either all left correct or all right correct and excluded the lap that contained the feeder-pass of interest. Correct passes were those at which the rat arrived at the correct feeder and received reward; error passes were those at which the rat arrived at the wrong feeder and no reward was present. Decoding was stopped once the rat left the feeder site. Correct laps were matched to error laps by randomly selecting either the correct lap that immediately preceded the error lap or the correct lap that immediately followed it. Correct laps immediately following the start of the session and the switch were excluded from the test sets.

66

4.3 Results

4.3.1 Behavior

Rats effectively learned the task. During the six probe days with contingency switches, rats started at chance and quickly learned to choose correctly (Figure 4-6**a**). Rats maintained a high percentage of correct laps until the change in reward contingency. Following the contingency switch, the percentage of correct laps dropped below chance and gradually returned to a high percentage of correct laps (Figure 4-6**b**).

Initially, as the rats learned the task, VTE was high while the rats determined the correct contingency. As rats learned the task and the percentage of correct laps increased, the amount of VTE (*zIdPhi*) demonstrated at the choice point decreased **(**Figure 4-6**c)**. When rats encountered the change in contingency, VTE increased drastically and then decreased back to levels seen during stable, correct performance (Figure 4-6**d**).
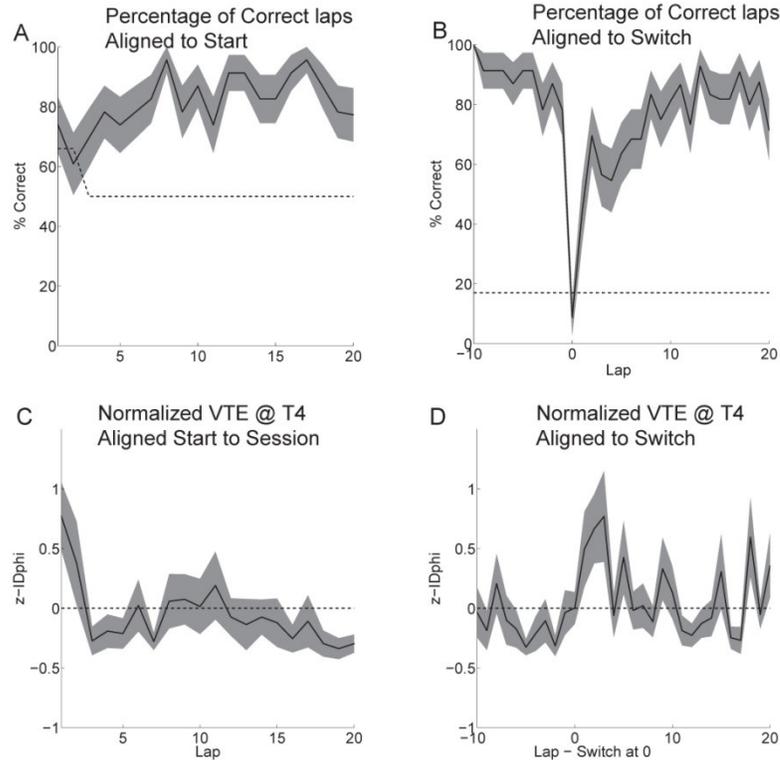
Figure 4-6 **Behavior on MT-LRA.**
a Rats performed at chance (dashed line) during the first few laps on the task. The rat had a 66% chance to receive food on his first lap because the first lap of an alternation session was always rewarded. As rats discovered the correctly rewarded contingency, their behavior stabilized and the majority of laps were correct (reward received). **b** Following the contingency switch, correct accuracy fell below chance (to the expected level the rat would show if it perseverated on the previous first-half-session strategy, dashed line). Rats were not aware of the time of the contingency switch or of the new, correct contingency. **c** VTE (*zIdPhi*) from the start of each session by lap. There was a significant effect of early laps on VTE (ANOVA laps 1-5, F = 4, P < 0.01). **d** VTE following the switch was significantly higher than scores prior (Wilcoxon, P < .01). Comparing the 5 laps pre switch to the 5 laps post switch demonstrated a significant interaction of VTE and lap (ANOVA F = 3, P < 0.01). [328]

4.3.2 Reward responsivity at feeders

Past neural recordings in OFC have demonstrated robust reward responses with a variety of specific responses [30,299,342,356,358]. Some cells responded to different rewards

68

(Figure 4-7).Many cells demonstrated preferential activity for a reward site, some responded more for banana flavored pellets at Left feeder 1 (Figure 4-8**a**) or for fruit flavored pellets at Right feeder 1 (Figure 4-8**b**). Of the 712 cells, 506 (71%) were classified as reward responsive. Cellular reward response dynamics we observed are consistent with prior recordings in this region [91,299,356].

Other data have suggested the OFC encodes value during decision processes [227,228]. Given that an individual rat would be expected to have a preference for one food over the other, if the OFC cells were encoding value, we would expect all of that animal's cells to prefer one food over the other. As shown in the Figure 4-7, cellular firing preferences within animal were equally divided between sides. This suggests that the reward-responses included sensory information. This interpretation is consistent with recent evidence that OFC represents the sensory aspects of rewards in a current state rather than value and is necessary during model-based decision making [128,192,193,262,263].
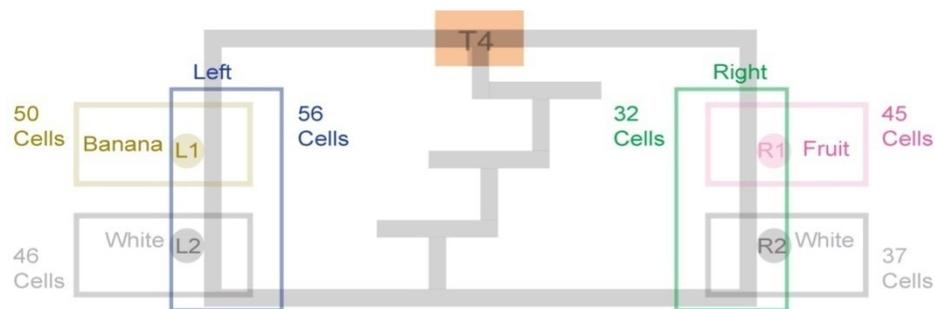
Figure 4-7 **Number of cells that responded to each different reward type/location.**
Many cells responded preferentially for a given reward type. Others responded for a given rewarded side, while some responded for certain combinations of reward sites. [328]

Table 1

	L1 Reward	L2 Reward	R1 Reward	R2 Reward	Left Side Reward
L1 Reward	50				
L2 Reward	56	46			
R1 Reward	10	10	45		
R2 Reward	13	9	32	37	
Left Side Reward			34	15	
Right Side Reward	23	19			108

Table 4-1 **Reward combination responsivity.** Each row and column represents a different feeder location. At each intersection point, numbers indicate how many cells responded for the different locations. If the row and column match, then the number reflects the number of cells for one feeder location (L1Reward x L1Reward, 50 cells). If the column and row intersect and are different, then the number reflects the number of cells that significantly responded at only those two sites (R1Reward x L1Reward, 10 cells). Certain cells were active for combinations: 10 cells responded to both L1 Reward and R1 Reward, while 9 cells responded to both R2 Reward and L2 Reward. 108 Cells responded at all reward locations. [328]

Table 2

Rat	Feeder L1	Feeder R1
R171	9	8
R172	14	19
R183	6	6
R186	21	12

Table 4-2 **Explicit reward responses to individual feeders**. Table 2 shows the number of cells, within rat, that responded explicitly to either Left feeder 1 reward (banana) or Right feeder 1 reward (fruit). [328]

A R172-2009-07-07-TT05_5

B R172-2009-07-08-TT10_7

C R171-2009-05-10-TT08_1

D R186-2009-09-14-TT06_1

71
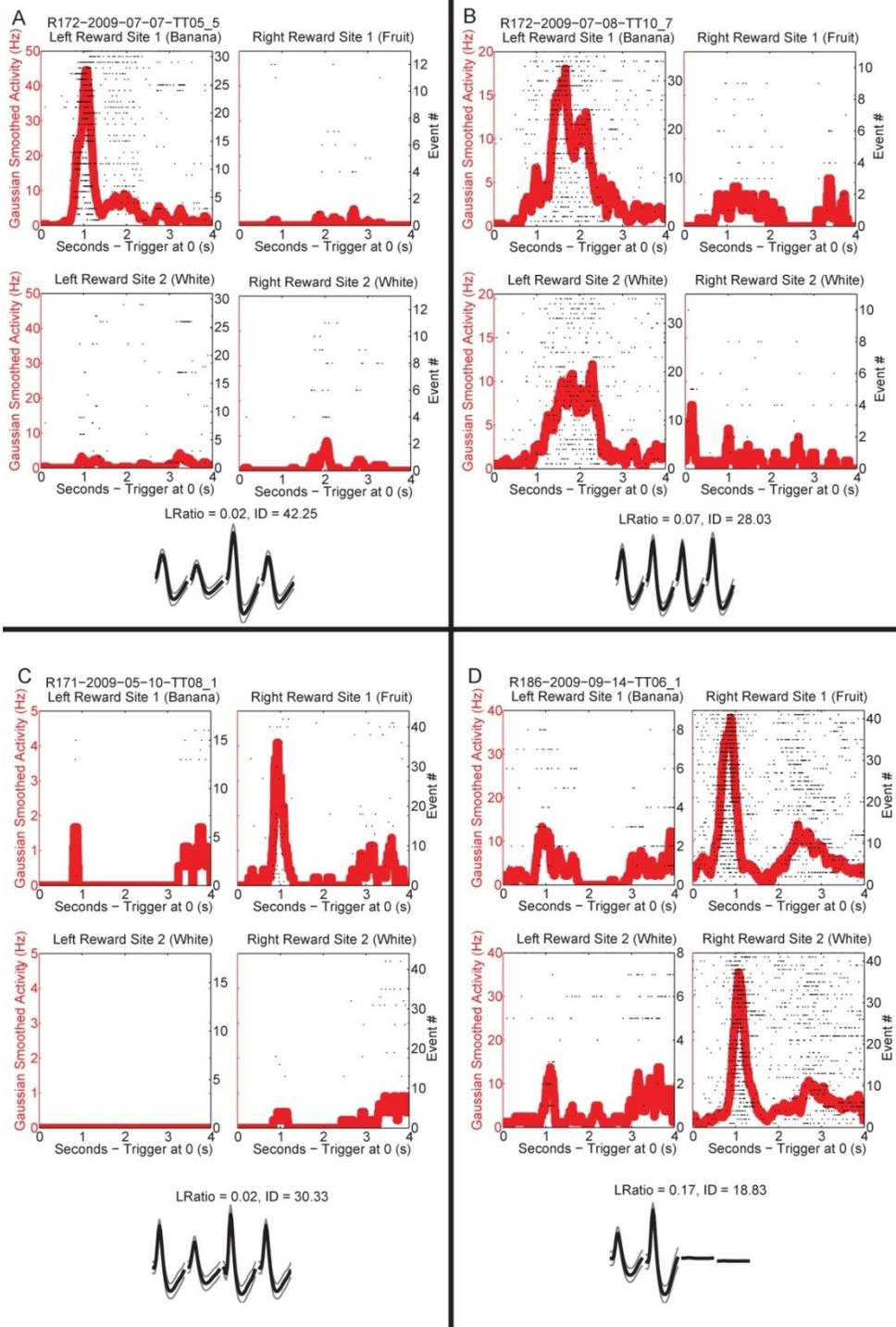
Figure 4-8 **Individual Neurons, different Rewards. Individual Neurons prefer left rewards.**
a This cell, R172-2009-07-07-TT05_5, preferred Left feeder 1 reward (banana). Very little activity was present at all other reward sites. Panels show rasters at each of the four feeder reward sites for each rewarded feeder pass (event #) with the population density overlaid in the red trace. As indicated by the rasters, this neuron responded significantly more to left feeder 1 (banana). The waveform shows the average waveform of the example neuron, because we are recording from tetrodes, waveforms have four components. The LRatio and Isolation Distance indicate that this neuron was well isolated from the other spikes in the session. **b** The bottom panels show a different neuron, R172-2009-07-08-TT10_7.t, which responded to both left feeder rewards (banana left feeder 1 and white, left feeder 2) much more strongly than the right feeder rewards. Panels, as in (b), show individual rasters with overlaid population spike densities convolved with a Gaussian at each of the four feeder sites during each rewarded feeder pass (event #). The waveform shows the waveform for this neuron on each channel of the tetrode. (c) A cell, R171-2009-05-10-TT08_1, that prefers Right feeder 1 reward. (d) R186-2009-09-14-TT06_1, which responded more to the right feeder rewards (fruit right feeder 1 and white, right feeder 2). [328]

4.3.3 Reward representations during VTE

Previous evidence has suggested that evaluative decision making occurs during orienting/re-orienting behaviors, quantified as VTE [125,150,352]. If evaluative processes necessary for deliberation are occurring during VTE, and if reward expectations are being modified or generated during this deliberative process, we should expect to see these expectations reflected in OFC activity.

Over the first 20 laps, 158 of 506 reward responsive cells demonstrated a significant individual firing rate correlation with VTE (Figure 4-9**a,c**). Of those 158 cells, 36 were also correlated with speed (Figure 4-9**e,g**). Even excluding the cells correlated with speed, many cells continued to show a relationship between *zIdPhi* (VTE) and reward firing while the rat was at T4, indicating that speed could not explain the excess neural activity in reward cells during VTE.

As previously stated, VTE reappeared after the contingency switch. To test if VTE and firing rate were still correlated after the switch, we again calculated the individual regressions for the z-scored firing rate of each reward-responsive cell against *zIdPhi* by lap. Again, there was a strong relationship between VTE and firing rate, driven in large part by the first few laps post-switch, when VTE was high. Following the switch in reward contingency, 214 reward responsive cells displayed a significant increase in normalized firing rate with VTE. Of those 214 cells, only 57 were correlated with speed (Figure 4-9**d,f,h**). These correlations imply that OFC reward responsive cells increased their firing rates during VTE. This effect could not be explained as simple correlations with speed.
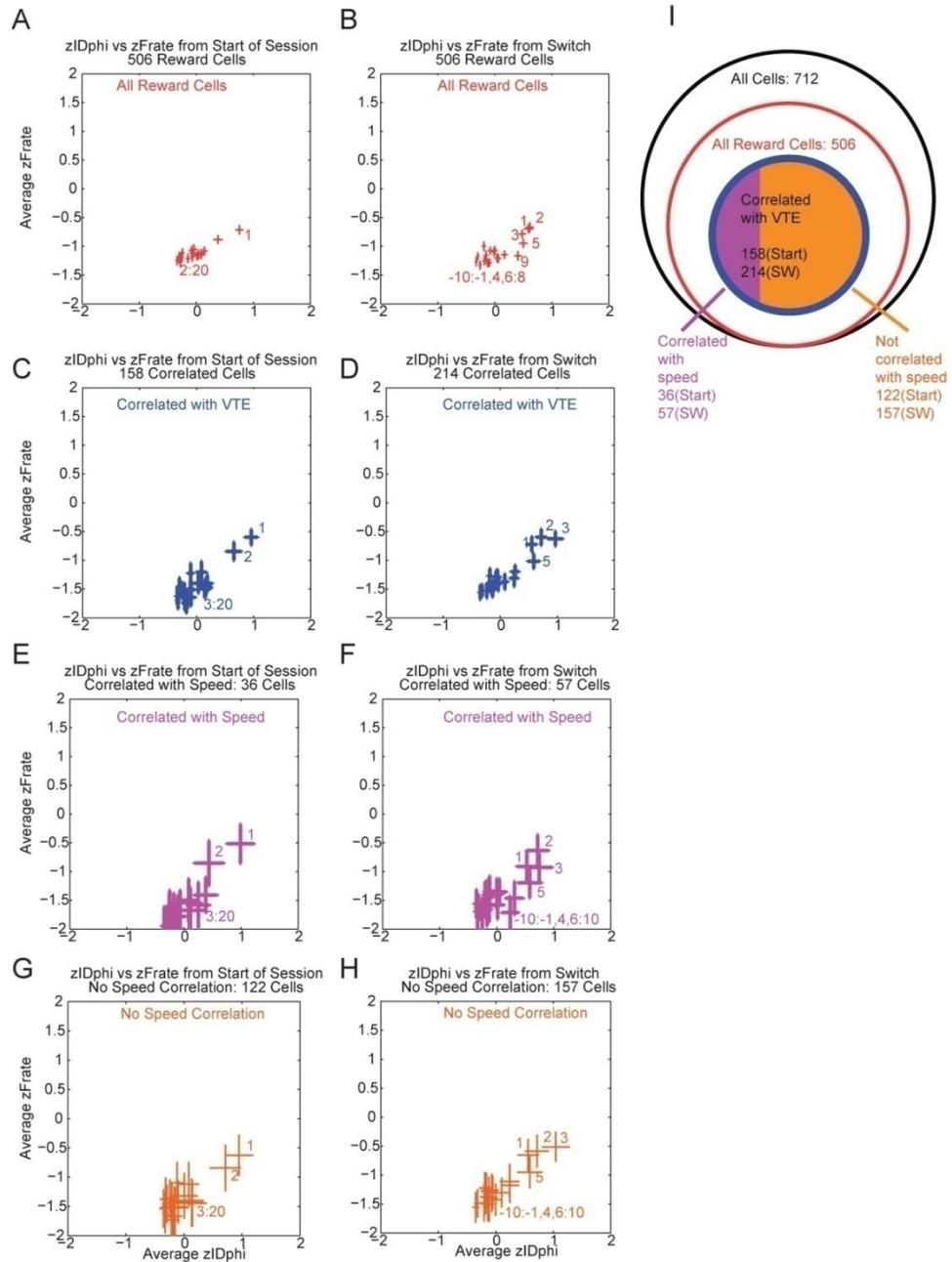
Figure 4-9 **Firing rate and *zIdphi* relationships.**
For each reward responsive cell, the average firing rate through a pass across T4 was calculated. A total of 506 cells responded significantly to reward. We then z-scored that firing rate distribution for each cell, producing a *zFRate* measure for each cell for each pass. Each panel shows the average *zFRate* across cells for each lap as a function of *zIdPhi* for that lap. Left panels used lap numbers aligned to the start of the

session. Right panels used lap numbers aligned to the switch in reward contingency. For many of these cells, so much activity was present at the reward site relative to the rest of the maze, that the zFiring rate was negative at all locations, even at the choice point, even when the cell fired extra spikes at the choice point and not elsewhere. The shift in z-scores during VTE identified that reward responsive cells increased their firing rate, but not to the same level as seen at the reward locations.

The top panels (**a,b**) shows the average *zFRate* versus *zIdPhi* for all reward responsive cells (506 cells). The next row of panels (**c,d**) shows the average *zFRate* versus *zIdPhi* for all cells that had a significant correlation between individual firing rate and *zIdPhi* (158 cells aligned to start of the session, 214 cells aligned to contingency switch). Of the reward responsive cells correlated with *zIdPhi*, some cells were also correlated with the speed of the animal during the choice point pass. The third row of panels (**e,f**) shows the average *zFRate* versus *zIdPhi* for all cells that were also correlated with speed (36/158 cells aligned to start of the session, 57/214 cells aligned to contingency switch). The bottom row of panels (**g,h**) shows the average *zFRate* versus *zIdPhi* for those cells not correlated with speed (122/158 cells aligned to start of session, 157/214 cells aligned to contingency switch).

The diagram on the right (**i**) depicts the total number of cells (black), the number of reward responsive cells (red), the number of reward responsive cells correlated with *zIdPhi* (blue), and finally the number of reward cells correlated with *zIdPhi* and speed in orange and the number of reward cells correlated with *zIdPhi* but *not* speed in purple. [328]

Reward cell activity at T4 was significantly different from non-reward cells (Figure 4-10). Previous research has indicated that ventral striatal reward cells show a similar phenomenon [351]. van der Meer and Redish found that this increased activity translated into an increased decoding to reward locations under a Bayesian decoding analysis [351]. In order to determine whether the increased firing in OFC in Figure 4-10 entailed a representation of the reward, we determined the extent to which OFC ensembles decoded to reward locations during VTE.
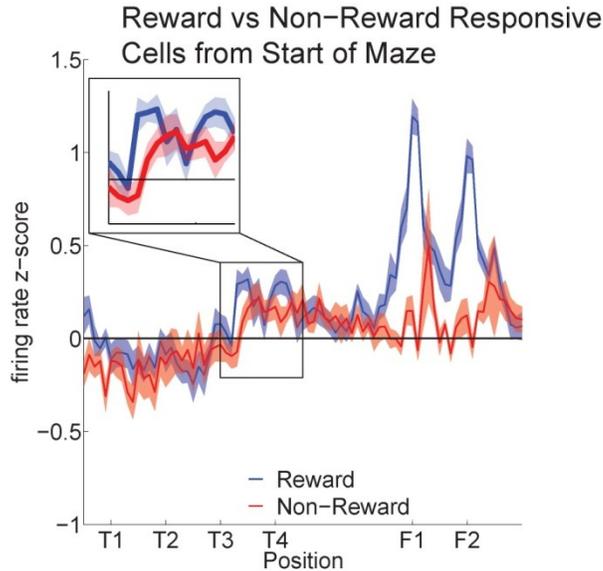
Figure 4-10 **Increased activity of reward cells at the choice point.**
Reward responsive cells showed significantly higher firing rates than non-reward responsive cells at T4 as well as at F1,F2 (ANOVA F = 5.3, P = 0.03; ANOVA F = 23, P < 0.01). We compared the normalized activity of reward responsive cells at salient maze locations by linearizing the tracking data from each session and assigning a set number of points between landmarks (Schmitzer-Torbert and Redish, 2004;van der Meer and Redish, 2009). A two-way ANOVA with lap (laps 1-10) and location on the maze (T1 to halfway to F1) as factors demonstrated a significant variation in firing rate across the navigation sequence (F = 5.6, p << 0.01). A Tukey post hoc comparison indicated that the activity at T4 was higher compared to other portions of the navigation sequence (T4: Mean = 0.44, Std = 0.07; Avg T1-T3: Mean = -0.08, Std = 0.09). [328]

As can be seen in Figure 4-10, *p(Reward)* increased following reorientation

events, then decreased as the rat left the choice point and progressed towards the feeders.

Individual reward responsive cells showed increased activity during VTE events, as

illustrated by the sample cell in Figure 4-11**a**. Interestingly, *p(Reward)* decoding to a

specific reward site was not seen. Instead reward representations remained general. The

non-specific increase in *p(Reward)* seen following reorientation suggests that once the rat

has made a decision, his reward expectation reflects the potential for reward rather than an explicit reward such as banana or fruit. An alternative explanation for increased representation of reward during VTE behaviors could be activity related to the previous lap (which was often an error). Previous reports have indicated that signals in OFC relate to the reward on the previous trial [331]. On our task VTE does increase during similar laps that errors do (early laps and again after the switch) [30], however there was no direct relationship between VTE and error trials (comparing VTE after errors to VTE after

Figure 4-11 **Example Activity and decoded *p(Reward)* during VTE events.**
a This particular reward responsive cell spiked multiple times during VTE. The inset shows the average waveforms of this cell for each channel of the tetrode. The behavior demonstrates the typical 'head sweep' seen during a VTE event. The red dots indicate neural spikes which are overlaid on top of the behavioral tracking data. The current lap is denoted by black dots. Reorientation is identified by the peak deviation in the pass, indicated by the blue arrow. On many sessions, multiple reorientation events occurred at the choice point. For this figure, *p(Reward)* was calculated using decoding based on all cells. *p(Reward)* was defined as the average posterior probability at each of the reward locations on the maze. Both left feeder and right feeder locations were included as part of the average to calculate *p(Reward)*. Average *p(Reward)* peaked immediately after the turnaround (green) on the (**b**) first, (**c**) last and (**d**) all reorientations at the choice point. We performed a control by shuffling (*grey trace in **b, c, d***) (nBoot = 500) the interspike intervals for each cell and recalculating the decoding using the un-shuffled tuning curves. This control determined that random firing would not reliably decode to reward locations. (Steiner and Redish 2012)

matched/following correct laps, Wilcoxon, p=0.1329).

Our data suggests that following VTE, representations of reward remain general and nonspecific (Figure 4-12), failing to represent a given side/specific flavor. During VTE the rat is contemplating the possibility of reward. However, this more than likely reflects the possibility of reward and we cannot distinguish between the sensory aspects of reward or reward value during VTE. During VTE it is possible that the rat has not transitioned to a rewarded state and is still selecting between the two available representations.

Hippocampal representations of the alternate choices do not reliably represent the direction the animal is facing – an animal can face to the left, but show a sweep of hippocampal representation to the right [125]. OFC activity could be confused for value if it represented the simplest sensory aspect of the reward, food or not. Given that the decision on MT-LRA is to remember which side is going to be reward on a given trial (and thus that, although the choice is free, the reward is defined), the specific sensory aspects of reward may be irrelevant at the time of decision. The activity might then seem to reflect the value of the reward during VTE. However, once the animal has reached the reward and consumed it, the sensory aspects of the reward would be fully represented.

The representations at the feeders more than likely represent the sensory aspects of reward rather than only the value of it. Given that an individual rat would expect to have a preference for one food, if the OFC cells were encoding value, we would expect

all of that animal's cells to prefer one food over the other. As shown in the following figure, cellular firing preferences within animal were equally divided between sides. This suggests that the reward-responses included sensory information. This interpretation is consistent with recent evidence that OFC represents the sensory aspects of rewards in a current state rather than value and is necessary during model-based decision making [192,193,263,335].
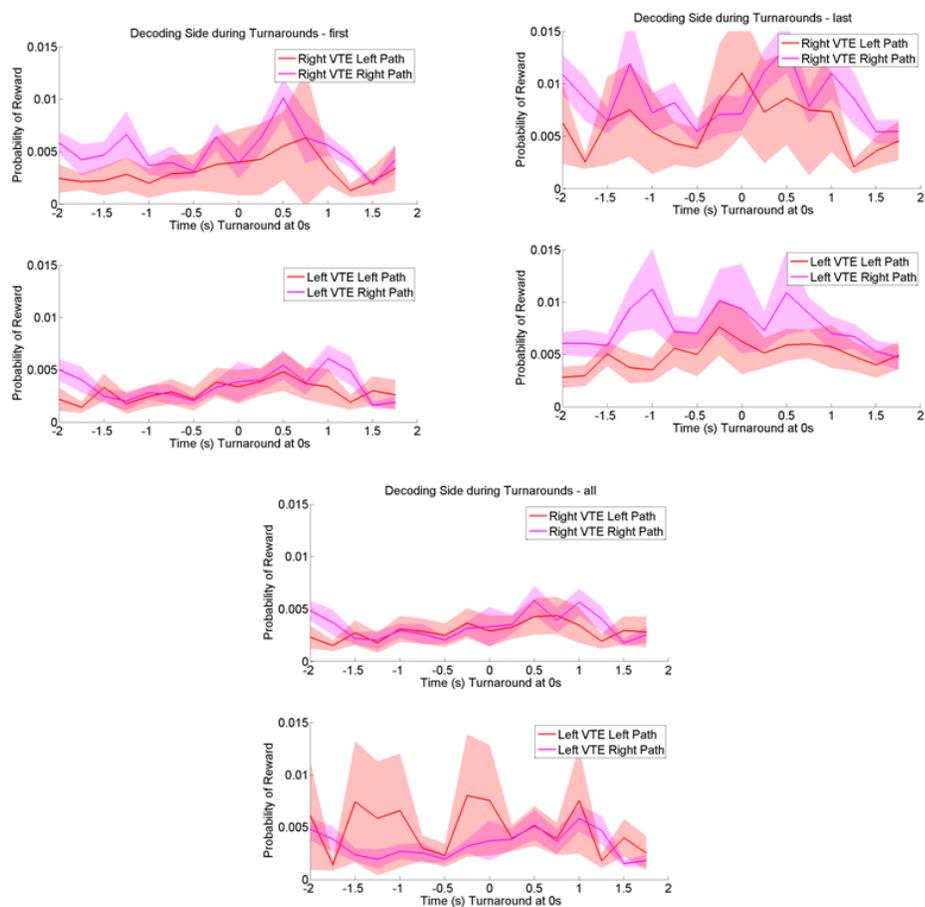
Figure 4-12 **Decoding to specific sides during VTE.**

The direction of the VTE did not predict which side the rat was traveling towards. Decoding to reward during VTE was general.

4.3.4 Disappointment and counterfactuals in OFC

Population responses of reward cells differentiated between laps in which rats received reward (correct laps) and laps in which they did not (error laps). Interestingly, on error laps population responses were more similar to that usually seen at the opposite feeder. For example, cells that responded to fruit reward receipt (Right Feeder 1), often responded at Left Feeder 1 during errors, when no reward was present (Figure 4-13**,** Figure 4-14**, single cell example)**. These results led us to investigate how OFC representations changed during the violation of reward expectation. We determined the difference in *p(Reward)* between correct and error laps and compared decoding on the same side loop to the opposite side loop, that is, *p(Reward$_{opposite}$)* was compared to *p(Reward$_{same}$)*.
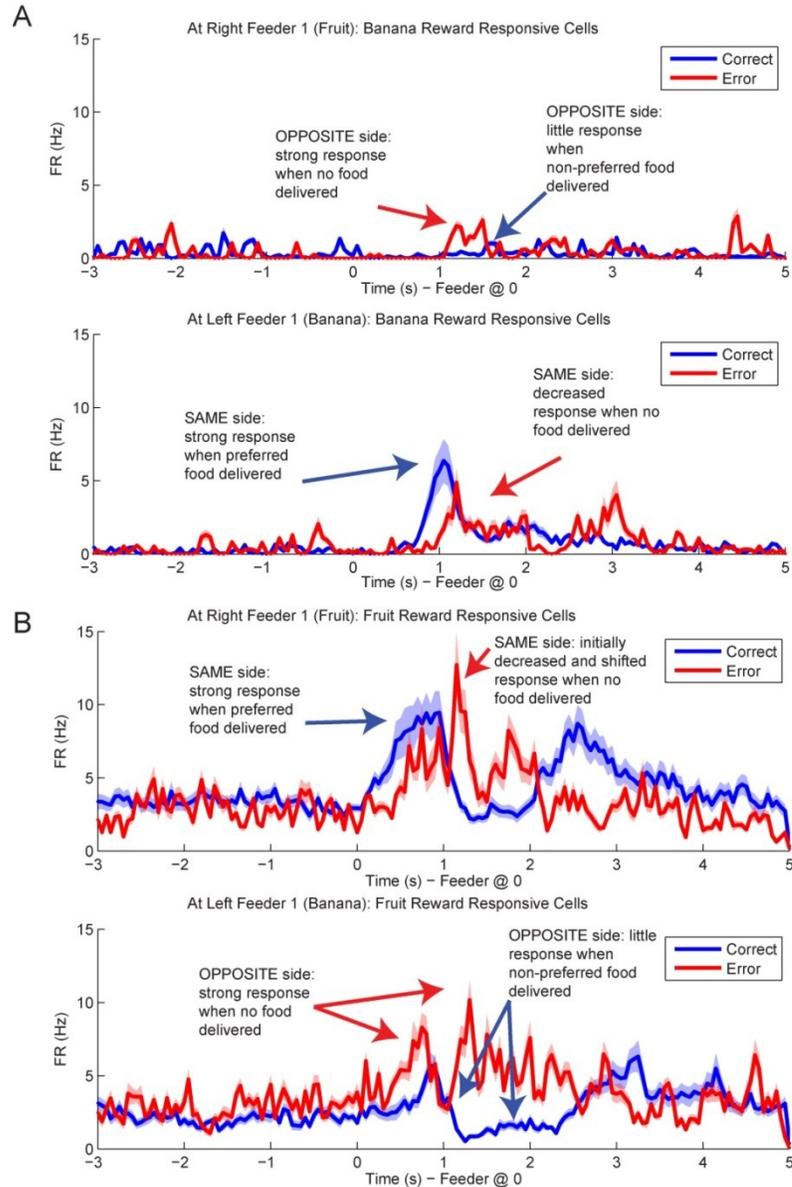
Figure 4-13 **Population Responses during errors.**
The average spike population density was calculated for all cells that preferred Left Feeder 1 (Banana) or Right Feeder 1 (Fruit) rewards. (**a**) Left-Reward/Banana preferring cells. The blue line in the lower plot (correct at left feeder) is larger than the blue line in the upper plot (correct at right feeder). However, note that the activity of left-preferring cells is larger on error laps at the right feeder than on correct laps at the right feeder. (**b**) Right-Reward/Fruit preferring cells. Cells classified as right-reward preferring demonstrated increased firing at the right reward sites. The blue line in the upper plot (correct at right feeder) is larger than the blue line in the lower plot (correct at left feeder). However, the activity of right-

reward preferring cells is larger on error laps at the left feeder than on correct laps at the left feeder (red trace, lower plot). [328]
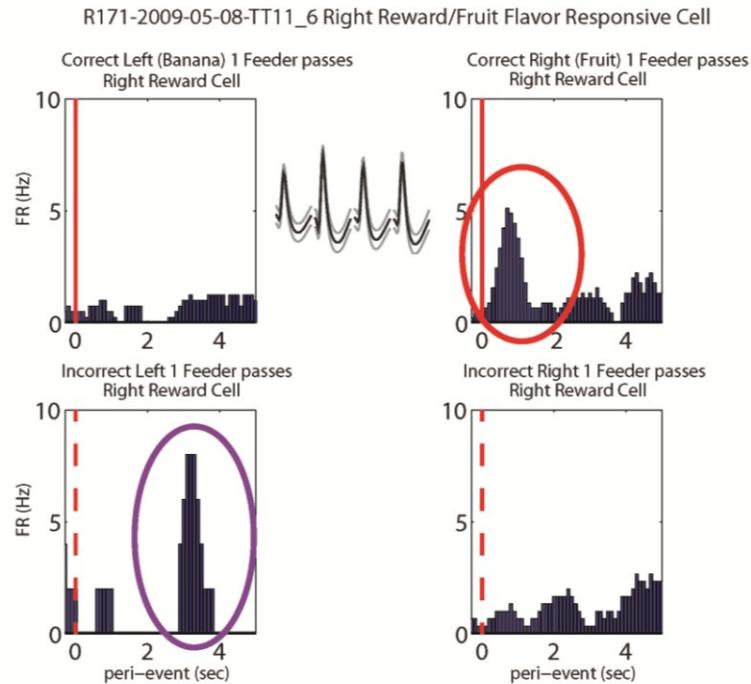
Figure 4-14 **Single cell responses during errors.**
During errors (no reward was present) this fruit preferring cell's average activity was more similar to the reward responses at feeder 1 when reward was present.

By utilizing the two different decoding sets, one for each loop, we were able to compare responses while rats were at the feeders during correct (reward) and error laps (no reward). *p(Reward_same)* was stronger on correct laps than during errors (Figure 4-15**a**). The shift in reward representations on error laps from 1 to 3.5 seconds seen in *p(Reward_same)* represented the neural correlate of disappointment; local representation of the reward decreased when the rat finally realized that he was not going to receive

reward. This observation of disappointment agrees with the economic definition;

disappointment is classified as the realization that available outcome does not match the

expected outcome [22,173].

In contrast, $p(Reward_{opposite})$ was larger on error laps, when no reward was

present, than $p(Reward_{opposite})$ on correct laps. This increase indicates that the spiking

activity seen during errors better matched the *other, would-have-been* rewarded side

during errors (Figure 4-15**b**). The increase in $p(Reward_{opposite})$ during errors

(approximately 1.25 seconds, following the shift in $p(Reward_{same})$ during errors) matches

the definition of the counterfactual and is consistent with observations of neural

representations of counterfactuals in humans; where an alternative, known outcome was

better than the received outcome [23,56,57,172,174].

The shift in $p(Reward)$ is best seen by comparing the ratio between

$p(Reward_{opposite})$ and $p(Reward_{same})$, which is best measured as the difference of the logs:

$log(p(Reward_{opposite})) – log(p(Reward_{same}))$. On correct laps, the difference remains on the

same side (i.e. $p(Reward_{same}) > p(Reward_{opposite})$), which indicates a better representation

of the side the animal is on. However, on error laps, the difference is initially located on

the same side (i.e. $p(Reward_{same}) > p(Reward_{opposite})$), but transitions to the opposite side

as the rat realizes no food is forthcoming (i.e. $p(Reward_{opposite}) > p(Reward_{same})$).

Comparing these two changes during *errors*, demonstrates a sustained shift to the would-

have-been rewarded side, the counterfactual (Figure 4-15**c**). Similar, sustained effects can

be seen at the second feeder-site (Figure 4-16).
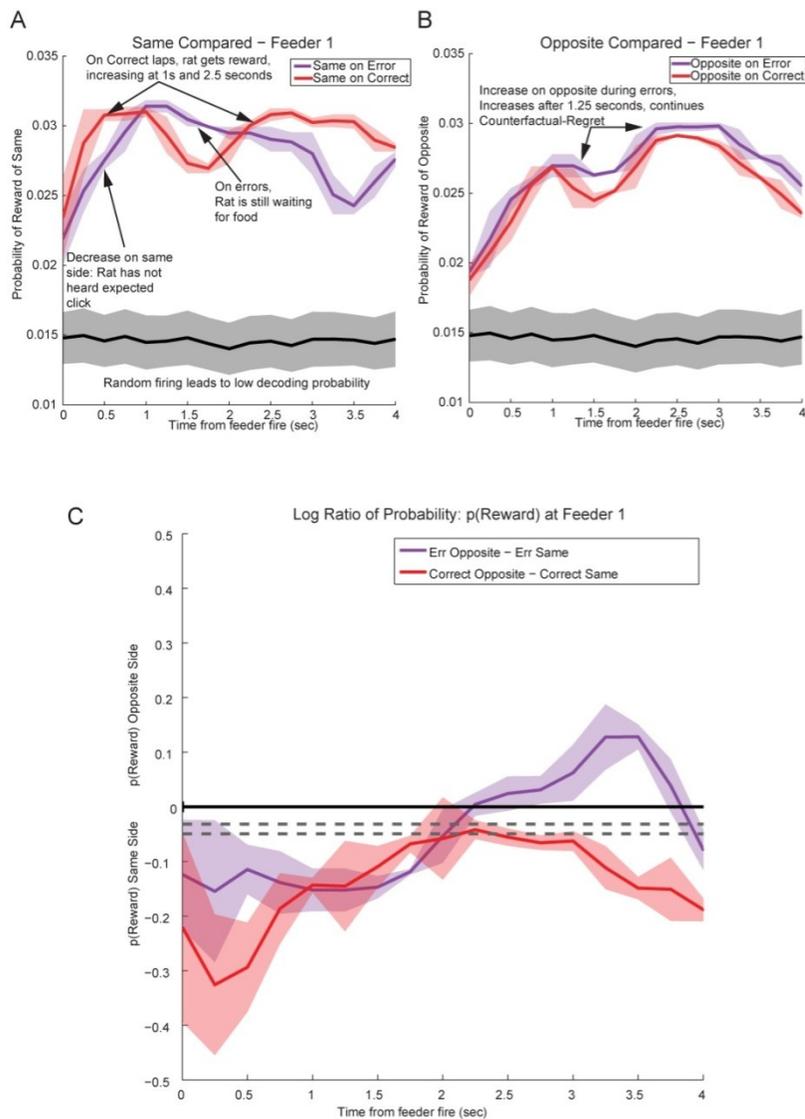
Figure 4-15 **Decoded *p(Reward)* at Feeder 1 Switches Sides during Errors – Disappointment and Counterfactuals.**
On correct feeder passes, reward followed approximately 1.8 seconds after an audible click (see Figure 1A for zone entry locations). On error passes, no reward is present and the animal does not hear the audible solenoid click. (**a**) In order to determine the presence of disappointment, we examined the difference in *p(Reward)* for the same side loop as a function of whether the choice was correct or an error. As can be

seen in the figure, there is a clear shift in the *p(Reward)* on errors (purple trace at approximately 1 to 3 seconds), when the animal realizes that he has not heard, nor will he hear the solenoid click that he expects and consequently he will not be receiving food. The decrease in local reward representation at 3 seconds occurs while the rats are pausing at the reward site, several seconds before rats begin to leave the reward site. Additionally, this difference is not a result of random firing in the absence of reward; shuffling the interspike intervals produces a much smaller *p(Reward)* [shown in the gray traces]. (**b**) Initially *p(Reward_opposite)* on error and correct are similar, however, once the rat realizes his error and begins to experience disappointment, neural representations increased to the opposite would-have-been rewarded site. This increase in *p(Reward)* to the opposite would-have-been rewarded side represents the counterfactual signal. (**c**) The log ratio between the local and non-local representations of reward, *p(Reward_same):p(Reward_opposite)*, for correct feeder passes and error passes. Data was smoothed using a 500 ms moving average. Grey lines represent the upper and lower quartiles for shuffled control, based on shuffling interspike intervals and re-calculating the decoding using unshuffled tuning curves. On errors, the log ratio of *p(Reward)* at feeder 1 remained local, (*p(Reward_same)* > *p(Reward_opposite)*), following arrival at the feeders from 0 to ~2 seconds, then switched to a non-local representation, (*p(Reward_same)* < *p(Reward_opposite)*). In contrast, on correct laps, the log ratio of *p(Reward)* at feeder 1 remained local for the duration of the animal's pause at the reward site. [328]
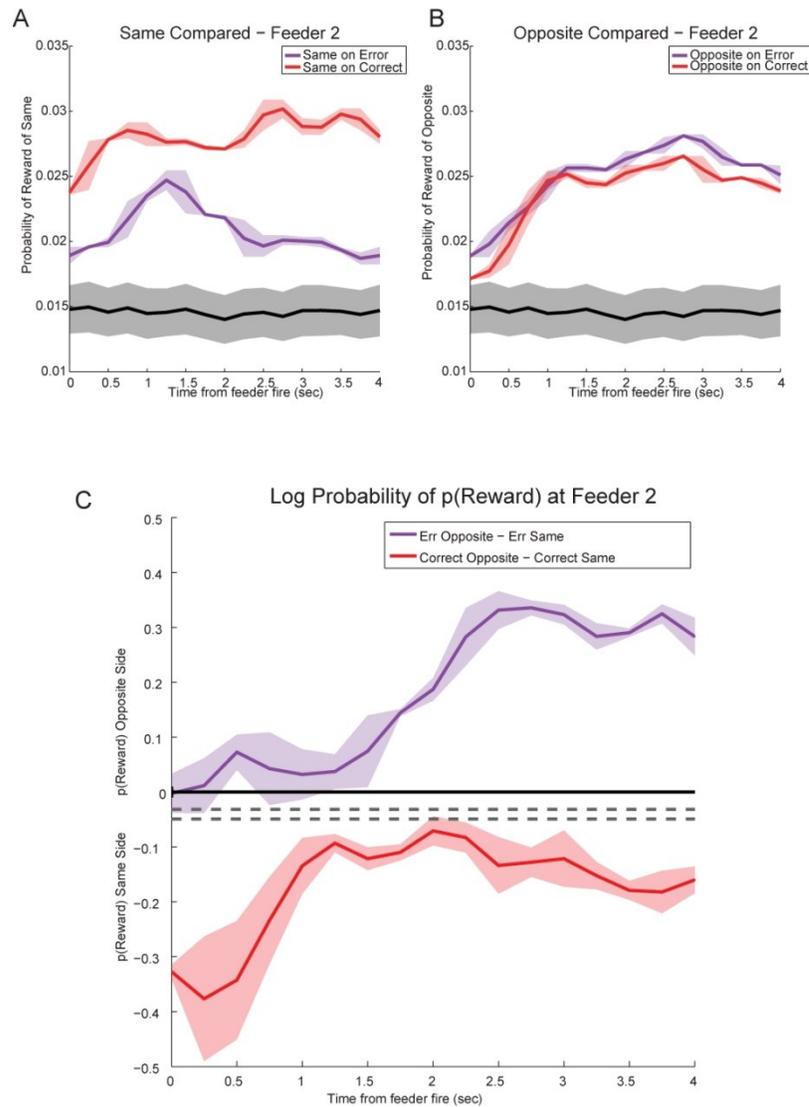
Figure 4-16 **Decoded p(Reward) at Feeder 2.**

a $p(Reward_{same})$ on correct was much higher reflecting the reward received. However, during errors $p(Reward_{same})$ decreased drastically. **b** Initially $p(Reward_{opposite})$ on error and correct remained the same. However, once the rat realized he would not be receiving reward, $p(Reward_{opposite})$ on errors increased and remained stronger than $p(Reward_{opposite})$ on correct. The sustained increase in $p(Reward_{opposite})$ during errors, the would-have-been rewarded side, may reflect a planning signal. **c** The log ratio between the local and non-local representations of reward, $p(Reward_{same}):p(Reward_{opposite})$, for correct feeder passes and error passes. Data was smoothed using a 500 ms moving average. Grey lines represent the upper and lower quartiles for shuffled control, based on shuffling interspike intervals and re-calculating the decoding using unshuffled tuning curves. On errors, the log ratio of $p(Reward)$ at feeder 2 demonstrated a sustained non-local response, ($p(Reward_{same}) < p(Reward_{opposite})$). This sustained response to the opposite, non-local side

during errors may reflect a planning signal. On correct laps, as seen in **Figure 5.14c**, the log ratio of *p(Reward)* at feeder 2 remained local for the duration of the animal's pause at the reward site. [328]

Controls

One potential confound is that in the absence of reward, decoding may become random or drop off. The increased noise (random-firing) hypothesis would predict that decoding would shift away from the representation of the local reward-site to become generally uniform across the entire maze. Similarly, the reward prediction error hypothesis would predict that the decoded probability would merely decrease and not increase on the other side. Neither of these hypotheses predicts self-consistent representations decoding to the opposite side reward feeder location (Figure 4-15).

To address these potential issues, we compared all decoded locations attained from both correct and error laps. Our decoding algorithm provided posterior probabilities for all possible positions on the linearized maze. By examining the posterior probability at other locations, we can differentiate noise from self-consistent counterfactuals (Figure 4-17). On correct laps, the differences were significantly positive; the neural activity was representative of the local reward location. This analysis replicated the results seen in Figure 4-15**a**, indicating that on rewarded laps, the decoded probability was a better match to the local training sets; *p(Reward)* matched the currently rewarded location of the rat. This analysis also confirmed that on error trials, the decoding better matched the non-local training sets; *p(Reward)* better represented the would-have-been rewarded location. On error laps, the differences were significantly negative. The neural activity

was representative of the alternate reward location, not a general change in representation

of the entire maze.
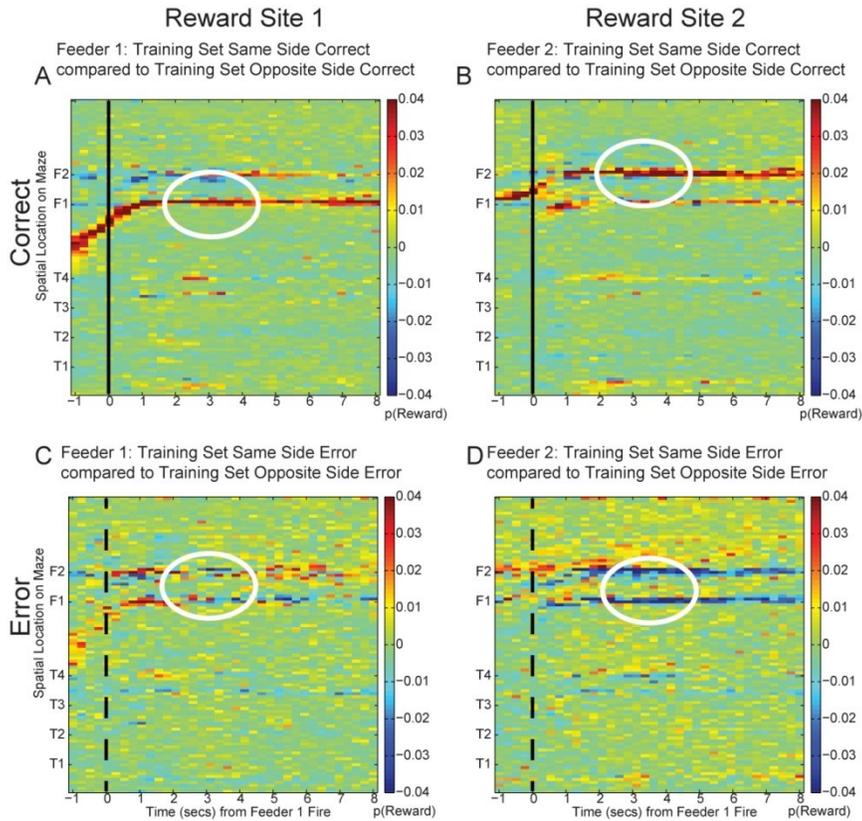
Figure 4-17 **Decoding across space.**
By comparing the log ratio of the decoding generated using all cells and with training sets on same and opposite sides of the rat, we can determine the extent to which neural ensembles reflect the current reward location, the opposite side reward location, or other locations on the maze. Each panel of the figure shows the difference between decoding using tuning curves based on the current side the animal is on and decoding using tuning curves based on the opposite side. As the rat passes the feeder trigger and hears an audible click (solid black line, correct), the rat receives reward and decoding is strongest to the reward site where the rat actually is, as indicated by the strong red color at the correct feeder location (*p(same) > p(opposite)*, **a**: white circle, Feeder 1; **b**: white circle, Feeder 2). This indicates that decoding remains local for the duration of the rats' stay at the feeder and fails to represent other possible locations on the maze. However, when the rat chooses incorrectly, upon crossing into the zone where he would have heard the reward trigger (dashed black line, incorrect), neural ensembles switch and represent the would-have-been rewarded side, as indicated by the strong blue color at the feeder locations (*p(same)<p(opposite)*, **c**: white circle, Feeder 1 and **d**: white circles Feeder 2). This indicates that neural representations during errors more closely resemble the activity at the would-have-been rewarded feeder. The decoding does not become random; instead *p(Reward $_{same}$)* decreases, while *p(Reward $_{opposite}$)* increases. [328]

The representations did not become random during reward, as would be expected from the increased-noise/random-firing hypothesis. Shuffling interspike intervals for the spiking data did not reliably represent reward on the maze (grey traces, Figure 4-15 and Figure 4-16), indicating that the increase we see to the would-have-been rewarded side could not be due to an increase in random firing. Nor did *p(Reward)* remain local, as would be expected by a reward prediction error signal. These decoding results indicated that OFC activity was representing the local, rewarded feeder during correct laps and the unchosen (opposite-side) feeder during errors.

We can differentiate disappointment from the counterfactual necessary for regret because we are separately measuring the amount of posterior probability assigned to each side independently. When the amount of posterior probability assigned to the same-side decreased (i.e. disappointment), it is not necessarily true that the amount of posterior probability assigned to the opposite-side would increase. Thus a local, same side decrease represents disappointment, defined economically as the violation of an expectation in the presence of *one* possible outcome [22,173], while an increase in the opposite-side representation can be interpreted as a true representation of the alternative reward (i.e. the counterfactual necessary for regret which has been defined as the representation of the alternative outcome when the received outcome does not meet expectations [23,174].

89

4.4 Discussion

In his poem "Mountain Interval", the poet Robert Frost postulated that a traveler faced with a decision pauses to consider possible outcomes, and then compares what is to what could have been [88]. Both of these processes require the representations of information about potential and non-local rewards, a process that relies on OFC [13,22,41,56,57,228,243,298,299,342]. This data reports that OFC is involved in both of these processes: (1) During orient-reorient behavior (previously identified as VTE), as rats pause to consider possible outcomes, OFC encodes expectations of reward. (2) During errors, OFC first encodes the disappointment (local) caused by the violated expectation and then transiently encodes the alternative, would-have-been-rewarded option (non-local).

4.4.1 Orientation and Reorientation

During early learning, computationally expensive, action-outcome processes attempt to predict reward through a series of what-if scenarios [70,125,330,350]. These scenarios can be evaluated without the direct execution of the action that leads to the outcome. During this vicarious evaluation, hippocampal ensembles represent prospective paths and ventral striatal ensembles indicate the presence of expected reward [125,354]. Here, we report that OFC ensembles reflect the expectation of general reward after re-orientation. The general representation of reward during VTE agrees with previous data

based on hippocampal representations during VTE. Hippocampal representations of the alternate choices did not reliably represent the direction the animal is facing – an animal could face to the left, but show a sweep of hippocampal representations to the right [125]. In our OFC data, we did not find reliable reward decoding that differentiates outcomes based on the orientation of the animal.

Other data indicates that OFC representations differentiated uncertainty on a trial by trial basis [142]. It is possible that during VTE, as the rat accesses internal representations of a reward expectation, uncertainty about the upcoming reward modulates the representation of reward and contributes to VTE. Additionally, activity in OFC may reflect some degree of decision confidence as the rat executes the turnaround and proceeds to a possibly rewarded site [142,184]. Although a more likely explanation is that the recording locations reported in Kepecs et al are much more lateral than the recording locations reported here. Rat lateral orbital cortex receives different inputs, especially visual input and dorsal lateral prefrontal cortex which has been strongly implicated in working memory, than the ventral orbital cortex, (where these data were obtained from).

4.4.2 OFC and ventral striatum

As shown in Figure 4-11, OFC ensembles decode to represent reward immediately after each reorientation during the VTE process. This suggests that OFC is likely to be involved in expectation following reorientation. In contrast, van der Meer and Redish found that ventral striatal representations of reward generally *preceded*

reorientation [354]. The tasks were identical and these reorientation processes occurred at the same location on similar laps. This suggests a difference between ventral striatal and OFC roles in decision-making processes.

4.4.3 OFC and reward prediction error

Our data suggest that OFC represents potential reward expectations, and our data are not consistent with OFC representations of reward prediction error. With experience, reward prediction error decreases. However, we did not see evidence for this decrease. Because reward-delivery at the two feeder sites on a given return rail were always either both provided (correct lap) or both not provided (error lap), reward prediction error signals would predict no activity at the second feeder site on a given lap. As shown in Figure 4-8, robust reward-related activity was seen at the second feeder-site. In fact, we were able to identify both disappointment signals and counterfactual signals at the second-feeder site.

4.4.4 OFC's contribution to a decision

Previous evidence from Johnson and Redish [125] has identified that hippocampal representations sweep ahead of the animal's location. Additional evidence links normal OFC function to the presence of an intact hippocampus [250]. When an animal approaches a decision point, hippocampal ensembles represent the possible paths. Following the spatial representations, ventral striatal ensembles represent the possible reward that lies at the

end of the represented paths. Our current data suggests that OFC represents the expectation of reward following the representation of reward in ventral striatum and the representation of potential paths in hippocampus. This implies that hippocampus likely contributes information regarding the paths to reward concurrently with reward/value information on the upcoming reward from ventral striatum. This information may be combined in OFC to form a state expectation, which includes the relevant sensory aspects of the reward, the reward type and other unique reward properties.

On MT-LRA, because reward sites are at a fixed location, location and reward are confounded. However, if OFC is representing the state of the expected reward, then all salient features of the reward, including the contingency (side) of the reward may be represented. The representation of state characteristics would allow an animal to make decisions using model-based processes [128,176,193,335] see also [70,350]. The exact timing of the reward representations in OFC and ventral striatum would be of much interest and could potentially provide useful evidence of a functional dissociation between OFC (state expectation) and ventral striatum (value calculation). Our data indicates that reward expectations peak in OFC after the turn-around point. van der Meer and Redish [354] found that the reward expectations in ventral striatum peaked before the turn-around point. However, it is important to note that the recordings came from different tasks and slight variations in the task procedures could have affected these timings. Further experiments recording neural ensembles from both locations simultaneously are likely to be fruitful.

93

4.4.5 Planning and counterfactuals

There is a close relationship between the ability to plan, which must include a representation of the potential outcome, and counterfactuals, a representation of the alternative that might-have-been. Planning often occurs after violations of expectations and the experience of disappointment or regret. Disappointment arises when situational expectations are violated and these violations are beyond the scope of one's control (Bell, 1985).

Disappointment (a lack of delivery of expected reward) is inherently aversive [257], however, disappointment and aversion must be computationally distinct entities because they show different relationships to extinction [254,255] – disappointment has the effect of extinguishing reinforced behaviors while aversion is extinguished by relief. Early economic studies [22,173] defined disappointment explicitly as reward omission. In our task, "disappointment" can be distinguished from the "counterfactual necessary for regret" because disappointment entails the recognition that an expected reward is not going to be delivered, while regret is the recognition that an alternative choice would have produced a better reward [23,174]. This requires the evaluation of the current reward, the expectation, and the possible alternatives.

Our data indicate that when the rat discovers his error at the first feeder, OFC representations of reward decrease at the expected reward location, implying disappointment, in conjunction with a distinct, transient increase in the representation of

94

the other, alternative would-have-been-rewarded option, the representation of the counterfactual. The strong shift during errors in $p(Reward_{same})$ occurred while rats were pausing at the first feeder waiting for reward and then drastically decreased several seconds prior to departure for the second feeder. Therefore we find it unlikely that the decrease in reward representations is related to movement away from the reward site or increased distance from the first reward site. Following the evaluation of the counterfactual, the feedback from the current lap could serve to instruct choice on subsequent laps, which could serve as a planning signal. We find it unlikely that the transient representation at Feeder 1 is a planning signal because the rat still has to go to the second (unrewarded) feeder before proceeding to the next lap. Rats reliably stop and check the second feeder, even when unrewarded. The increased decoding to the would-have-been rewarded side at the second feeder, however, may reflect a planning signal and represent a form of episodic future thinking [125,235,350].

4.4.6 Violation of expectations

Regret entails a comparison between the expected outcome and a better alternative, which implies a comparison between *multiple* choices [23,41]. This means that regret requires a representation of a counterfactual, the better alternative. Although the log ratio analysis in Figure 4-15**c** cannot differentiate between disappointment and regret, the evidence in Figure 4-15**a,b** does differentiate, indicating that the decoding during errors is to the opposite reward location rather than a general diminishment in decoding

quality. This implies that the transient representation includes disappointment (local) and the counterfactual (non-local).

The decrease seen in the representations of the local side reflects the absence of reward, a disappointing outcome. This is followed by a transient representation of the known alternatives, an increase in the decoding to the opposite side, and the comparison of what could have been to what was; the counterfactual and the experiencing of regret. Interestingly, the evidence that OFC is required for the generation of a reward prediction error in the ventral tegmental area, may imply that disappointment and regret drive the formation of the reward prediction error elsewhere [309,335]. In human subjects during fictive learning, OFC activity increased when subjects were considering the possible outcomes of their actions in different conditions [204,319]. Following this hypothesis, reversal learning could be considered to be a form of 'learning from one's regret'. Both humans and animals with OFC lesions are impaired on reversal tasks [13,75,301]. Additionally, human subjects with OFC lesions do not exhibit the negative emotional arousal associated with the experience of regret [41,56]. Regret and disappointment could contribute to the maintenance of reward expectations in OFC for a situation or a given model based representation [128,176,192,193].

In humans, value representations of alternative outcomes (counterfactuals) activate OFC [56,57]. The timing of these counterfactual representations agrees with experiences of regret and is correlated with fMRI BOLD activation in OFC and anterior

96

hippocampus among other structures [56,57,243]. Which leads to the question: are rats

capable of experiencing regret? Regret requires the comparison between an actual

outcome and a counterfactual outcome that would have been the result of an unchosen

action [23,41,56,57]. There may be some evidence of causal reasoning in rats, though this

evidence is contentious [28]. On our task, when a rat chose incorrectly and arrived at non-

rewarded feeder sites, neural ensembles in OFC representing reward switched to

represent reward on the opposite, rejected side, implying that rats can at least represent

the counterfactual necessary for regret.

4.4.7 Conclusion

Are rodents capable of experiencing regret? There is building evidence that they

are capable of causal thought. Causal thought contributes to the counterfactual. If

humans, primates, and now rats are capable of calculating the counterfactual, then it is

possible that they can use the counterfactual signal to calculate regret, as defined in the

economics literature. With the advances in technology and the reliability of ensemble

decoding techniques, it is now possible to measure, on an ensemble level, what the

neurons are representing. By using this technique we can measure what reward the

neurons in OFC are representing and use this information to determine what the rat is

'thinking about'.

These techniques have been successfully used in the past and have demonstrated

remarkable accuracy when considering a rat's future path, the covert expectation of

reward, or the counterfactual. With the added behavior accuracy of the definition of VTE, we can determine if the rat is performing VTE behavior and if that occurrence coincides with the representation of the counterfactual. However, MT-LRA was not designed to test multiple reward representations. Other studies relied on a variety of menu items to determine the reward response dynamics in OFC. As a result in order to successfully determine what is being represented in OFC during counterfactuals and VTE, we need to be able to compare the neural signals to *multiple* different reward signals, essentially asking, 'What flavor am I thinking about? What do I prefer?'.

Human studies have termed this the revealed preference, what is the cost of a menu item and how much does a subject prefer that item over another? These menu items are subject to framing effects and can vary across days. However, during a single session the framing effect is usually assumed to be consistent which allows for a comparison of goods.

5. From counterfactuals to regret

5.1 Necessity of regret to decision making

Economic definitions of regret rely on the value of the outcome (in prospect theory, expected utility and in regret theory [22,23,58,97,171,174,203,206,252,285]. However, when considering the effects of regret, most human subjects report that the action leading up to the negative outcome generated more regret than the actual outcome [95,156,157]. Recent neurophysiology indicates that regret is strongest following an incorrect choice and the reveal of a better alternative [56,57]. However, due to the close proximity of the decision to the reveal, subjects may have been expressing more regret over the actions that lead to the outcomes rather than the less than optimal outcomes. If subjects did in fact experience more regret over their actions, these results would be consistent with psychological measures of regret, bridging the gap between neurophysiology and psychology. However, this does not address the discrepancy seen in economics; previous definitions of regret have placed emphasis on the outcome rather than on the process of the arriving at the outcome.

This difference between the economic definition of regret and psychology definition of regret stems from a reliance on economics to first define the economics of the decision and further fit human behavior to the model derived. Psychological models of regret often fail to describe their findings in terms of replicable decisions models. Instead they allude to equations and principle generated by the economists and offer no translation to what has been postulated instead relying on vague descriptions on human behavior. This massive divide between fields has contributed many inadequate and incorrect models of regret.

Building from the early descriptions of expected utility value, psychologists described regret as a minimax function that followed loosely with formulations of expected utility; an agent selects an option in a given state that if incorrect minimizes the amount of regret. The overall loss of the option is defined as the difference between the best possible payoff and the actual payoff, the smaller this value, the smaller the potential regret[289]. When regret is minimized the income (*I*) is maximized.

$$Regret\ (Loss(a,s)) = I(a_1, s_1)_{max} - I(a_1, s_1)$$

This representation of regret has several issues. This methodology assumes that the actor determines the world to be always in the worst possible state [289]. In addition, the minimax representation ignores any probabilistic representations common to expected utility and prospect theory [190]. Humans generally do not use minimax regret to modify decisions [102].

Instead humans rely on probabilistic outcomes and try to anticipate the perceived cost/benefit of a decision [97,130,131,203,205,252,253,355]. Under this framework, the anticipation of regret can reduce the perceived benefit of a decision, thus increasing the cost of the action that leads to the outcome. Regret modifies the probabilistic representation outcome and its associated value as well as assigning a cost to the action performed to acquire an outcome; the higher the perceived anticipated regret, the lower the value of the outcome. In retrospective decisions, regret serves an update function in the presence of more than one potential action/outcome, a regret function compares the perceived action/value to the alternative action/value [23,174]. When regret is high, the expected/value of the alternative is higher than the performed action/outcome.

5.2 Neural basis for regret

The role of OFC in regret was first identified by Camille et al[41]. They measured the ability of OFC lesioned human subjects to express regret, both the negative emotions associated with regret, as well as the subsequent behavioral changes measured through skin conductivity responses [41]. When subjects with OFC lesions selected a deal that resulted in a less than desired outcome to the alternative deal, they did not self report regret or exhibit any behavioral response to regret. In contrast control subjects under the same circumstances reported regret and exhibited increased skin conductivity indicative of a negative emotional response.

Subjects were asked to choose between two spinners, each of which represented a different distribution of reward probabilities. Once subjects selected a spinner the outcome of the selected spinner was revealed. Two conditions controlled for the effect of feedback in the experience of regret. Under the partial feedback condition, subjects were not shown the result of the not-selected spinner. Under the complete feedback subjects were shown the results of the other spinner. Following the presentation of results, subjects were asked to rate their experience in terms of sadness and happiness. During these reports skin conductivity was recorded as well.

When considering regret, the two conditions represent disappointment (partial feedback) and potential regret (complete feedback). In the partial feedback condition a loss would be disappointing; however, without an alternative to compare that loss (in value) to, according to the formal economic definition of regret, $\psi[v(x)-v(y)]$, regret could not be represented. In contrast, in the complete feedback condition, with two values presented, if the value of the non-selected option was greater than the value of the selected option, subjects could be expected to experience regret.
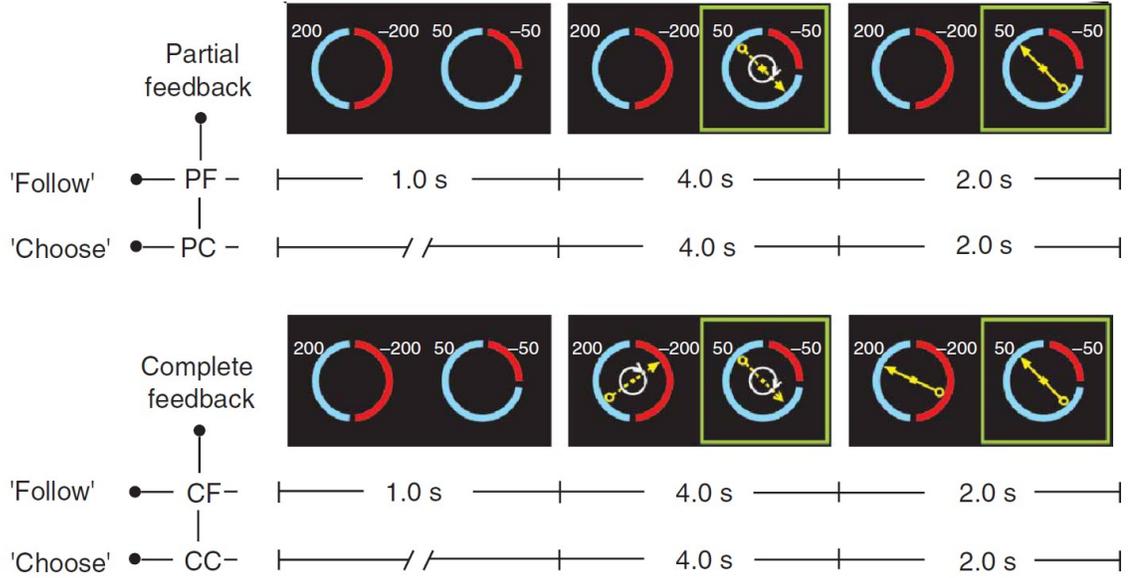
Figure 5-1 **Measuring regret in humans**.
Both OFC lesion and control subjects selected a gamble. Each spinner had a different probability of reward. In the partial feedback condition only the selected gambles arrow and outcome were displayed. In the complete feedback condition, both selected and unselected gambles showed the outcome. In the follow set for partial and complete feedback, subjects were instructed that the computer would choose a gamble. (used with permission [41,56])

Subjects with OFC lesions failed to report any negative experience with the complete feedback condition (Figure 5-1) indicating that they did not experience regret and when faced with future, risky gambles, did not choose the option with the least amount of risk (avoiding potential regret). In contrast, control subjects reported a strong negative emotional response to loss when instead they could have gained (Figure 5-2). On further trials control subjects would select the less risky of the two options to minimize their future potential regret. These results illustrated that not only did OFC play a role in the formulation and expression of regret, lesions to OFC eliminated any form of anticipated regret.

A later study used fMRI and demonstrated that mOFC activation was strongly correlated to the experience of regret [56]. In addition, hippocampus and anterior cingulate cortex activation also showed correlations with the experience of regret. Other structures such as ventral striatum were correlated with value but not the experience or the strength of regret representations. mOFC representations also distinguished between mere disappointment (attenuated responses) and regret representations. Following a disappointing outcome, fMRI increases in OFC were much smaller when compared to an outcome that induced regret. These results are also consistent with the role of the OFC in regret modulation and the formulation of expected outcomes.
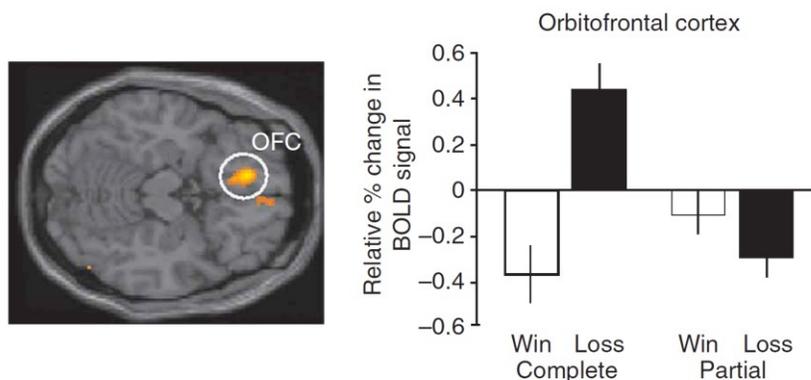
Figure 5-2 **Activity in the OFC increased following regret (loss rather than a gain) in the complete information condition.**
In the partial information condition, activity in OFC did not significantly change between a gain or a loss. (used with permission[41])

When considering OFC's role in the formulation of regret, it is worth noting that the formulation of regret need not be a complex process. The formulation of regret can be achieved by matching the differences in action-outcomes [111,112]. Regret matching is the

process of changing a previous strategy to acquire the previously missed outcome. This methodology seeks to maximize the expected value and minimize regret: "I chose A previously, A was much worse than B, so next time I will choose B". By computing the counterfactual the overall value of the missed outcome is compared and the future action is modified to achieve a greater overall value.

A recent experimental result in primate neurophysiology shows that neurons respond to both the hypothetical and the actual reward. Abe and Lee showed that following a loss during a rock, paper, scissors, game, neurons in the OFC represented the previous, hypothetical option (counterfactual) as well as the actual outcome and the current location of the stimulus on the screen [1]. In addition, recent work identified neural representations of the counterfactual in rodent OFC [328], see chapter 4).

When better alternatives were known to be available, human subjects reported feeling regret [41]. When human subjects reported feeling regret, neural activity increased in OFC [56]. This realization that reward would have been received had an alternative action been taken can be defined as the counterfactual [41,56].

5.3 Exclusivity of regret

New models of OFC function include the actions as well as the values. This indicates that OFC not only represents the value but other aspects as well, such as identity, the possible actions required, and the necessary alternatives: all are parts of the

representation. The implication of these results means that OFC represents the outcome as the combined action/value pairing, indicating that during regret, representations of the alternative outcome would be strongest following initiation of the action that leads to the expected outcome.

By defining regret in terms of the observable action/outcome we introduce an interesting question; is regret exclusive to humans? If regret serves to modify future behavior, then it is advantageous to survival. One can reasonably expect that other animals can experience regret or possess a mental module capable of causal reasoning [276,337]; if two options are available and the selected option is lesser valued, then regret (according to the definition) can and should be experienced. The only limitation in the assignment of regret is the reluctance to classify regret as something more ubiquitous than just a human emotion. Even human causal reasoning can be thought to have existed first in animals [195,244]. There is building evidence that both rats and non-human primates represent the counterfactual [1,28,115,328] and are capable of retrospective and prospective evaluation [8,76,77,87,350,381]. Other animal models such as the New Caledonian crow are capable of causal learning and possess the potential for counterfactual reasoning [336]. Because the representation of the counterfactual is necessary for regret, it is possible that other animals are capable of representing regret according to the economic definition and the effect regret has been shown to have on future behavior.

To measure regret in rats, the experiment performed in Chapter 4 would require *multiple* choices [23,41,174], rather than just one or the other. This would make it possible to differentiate regret and disappointment by comparing the outcomes across multiple conditions, where some errors are the fault of the rat and others are the fault of the world. The analyses in figure 5-15 lacked the ability to measure multiple choices. To measure regret in rats, humans or non-human primates would require that the subject have choices, and that the choices have known outcomes each with a specific preference.

6. Neurophysiological representations of regret in rats

6.1 Summary

Disappointment entails the recognition that one did not get the value one expected. In contrast, regret entails the recognition that an alternate (counterfactual) action would have produced a more valued outcome. Thus, the key to identifying regret is the representation of that counterfactual option in situations in which a mistake has been made.

In humans, the orbitofrontal cortex is active during expressions of regret, and humans with damage to the orbitofrontal cortex do not express regret. In rats and non-human primates, both the orbitofrontal cortex and the ventral striatum have been implicated in decision-making, particularly in representations of expectations of reward. In order to examine representations of regretful situations, we recorded neural ensembles from orbitofrontal cortex and ventral striatum in rats encountering a spatial sequence of wait/skip choices for delayed delivery of different food flavors. This task revealed individual preferences for flavor, allowing the identification of value and cost. Rats occasionally skipped low-cost choices and then encountered a high-cost choice. This

sequence economically defines a potential regret-inducing instance. In these situations, when rats looked backward towards the lost option, the cells within the orbitofrontal cortex and ventral striatum represented that missed action, rats were more likely to wait for the long delay, and rats rushed through eating the food after that delay. That these situations drove rats to modify their behavior suggests that regret-like processes modify decision-making in non-human mammals.

6.2 Introduction

Regret is a universal human paradigm [23,41,57,95,158]. The experience of regret modifies future actions [23,57,174]. However, the experience of regret in other mammals has never been identified; it is not known whether non-human mammals are capable of experiencing regret. Although non-human animals cannot verbally express regret, one can create regret-inducing situations and ask whether those regret-inducing situations influence neurophysiological representations or behavior: Do non-human animals demonstrate the neural correlates of regret in potential regret-inducing situations?

When evaluating the experience of regret, it is important to differentiate regret from disappointment [22,23,173,175]. Disappointment is the realization that a realized outcome is worse than expected [22,173,175]; regret is the realization that the worse than expected outcome is due to one's own mistaken action [41,57,95,156]. Disappointment can be differentiated from regret through differences in the recognition of alternatives [22,41,162,174]. Regret can be defined as

the recognition that the option taken resulted in a worse outcome than an alternative option/action would have. The revaluation of the previous choice in context of the current choice is the economic foundation of regret [23,174].

Humans with damage to the orbitofrontal cortex do not express regret [41], and fMRI experiments reveal activity in the orbitofrontal cortex during regret [56,57]. In rats and non-human primates, the orbitofrontal cortex has been implicated in decision-making, particularly in the role of expectations of future reward and the complex calculations of inferred reward [128,193,228,297,333,342,363,369]. Orbitofrontal cortical neurons represent the chosen value of an expected future reward [1,228,331], and an intact orbitofrontal cortex is critical for reversal learning [85,301] (however recent evidence suggests that OFC may have a more specialized role and is not necessary for reversal learning at least in primates [282]). Orbitofrontal cortex has been hypothesized to be critical for learning and decision-making [128,162,163,192,193,259,302,309], particularly in the evaluation of expected outcomes [228,328].

The ventral striatum has also been implicated in evaluation of outcomes [60,121,216,245,266,354], particularly in evaluation during the process of decision making [159,192,193,309,316,354]. Neural recordings from ventral striatum and orbitofrontal cortex in rats have found representations of reward, value and prediction of expected value in both structures [239,297,328,334,335,354]. In the rat, lesion studies suggest orbitofrontal cortex is necessary for recognition of reward-related changes that require inference, such as flavor

110

and kind, while ventral striatum is necessary for recognition of any changes that affect value [128,192,193,309]. In rats deliberating at choice points, ventral striatal reward representations are transiently active before and during the reorientation process [354], but orbitofrontal cortex reward representations are only active after the reorientation process is complete [328].

We developed a neuroeconomic spatial decision-making task for rats (<u>Restaurant Row</u>) in which the rat encounters a serial sequence of take/skip choices. The Restaurant Row task consisted of a large inner loop, approximately one meter in diameter with four spokes proceeding out from the inner loop (Figure 6-1**a**). The zone's entries were separated by ninety degrees and each one led to a potential reward location approximately half a meter from the entry point on the central, circular track. Each zone supplied a different flavor of food (banana, cherry, chocolate, and unflavored). Flavor locations remained constant throughout the experiment. Rats were trained to run around the loop, making stay/skip decisions as they passed each spoke.
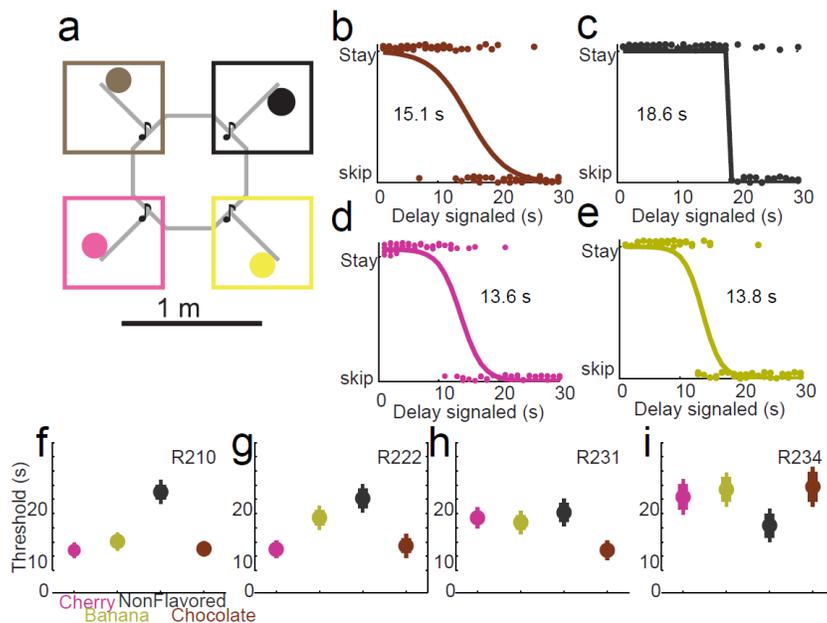
Figure 6-1 **Restaurant Row and revealed preferences in rats.**
a, The Restaurant Row task consisted of a central ring with four connected spokes leading to individual food flavors. Rats ran around the ring, encountering the four invisible zones (square boxes) sequentially. Color reflects flavor: magenta=cherry, yellow=banana, black= nonflavored/plain, brown=chocolate. **b-e,** Rats typically waited through short delays but skipped long delays. Each panel shows the stay/go decisions for all encounters of a single rat running a single session (R210-2011-02-02). A small vertical jitter has been added for display purposes. Thresholds were fit as described in <u>Methods</u>. **f-i,** Each rat demonstrated a different revealed preference that was consistent within rat across all sessions, but differed between rats. Thresholds were fit for each flavor for each session. Each panel shows the mean fit threshold for a given rat, with standard error shown over sessions[327].

As the rats proceeded around the track, the rat's position was tracked from LEDs

on the head via a camera in the ceiling. A spatial zone was defined for each spoke that

included the complete spoke and a portion of the inner loop and aligned with the inner

loop such that a rat could not miss a zone by running past it (boxes in Figure 6-1**a**). Upon

entering each zone, rats encountered different offers of delays. Zone entries were defined

entirely by the detected position of the rat's head and were not explicitly marked on the

track. On entry into a zone, a tone sounded, such that pitch of the tone indicated the delay

112

that the rat had to wait in order to receive reward (higher pitch = longer delay). As long as the rat stayed within the zone, the delay counted down, with each subsequent second indicated by a lower pitch tone. If the rat left the zone, the countdown stopped, no sound was played, and the offer rescinded – the rat's only option was to proceed on to the next spoke and the next zone.

The delays were selected pseudo-randomly from a uniform distribution ranging from 1-30s (for two rats) or 1-45s (for two rats). The delays were independently selected, so the delay offered at each zone encounter was independent of other zones for that lap. When making a decision to stay or skip at a given zone (when offered a given delay), the only information the rat had was the flavor of the food offered (because flavor locations remained constant throughout the experiment), the delay it would have to wait (because delay was signaled by pitch of the auditory cue), and the probability distribution of any future offers (because offers were drawn from a uniform distribution of 1-30s or 1-45s).

Rats were run for one 60 minute session per day. This time-limit meant that rats had a "time-budget" of 60 minutes to spend foraging for food. Because the session was time-limited, the decision to stay or skip a zone was not independent of the other zones. Waiting at one zone was time that could have been spent at another zone. An economically-maximizing rat should distribute its time between the offers, waiting for valuable offers, but skipping expensive offers. Assuming that an animal likes some

113

flavors more than others, the value of an offer should depend on the delay offered and the animal's preferences.

An important consideration is to control for the possibility that rats were waiting for a specific cue before leaving the zone. The fact that rats either stayed through the entire delay or left after a very stable 3 seconds implies that rats were not waiting for a specific delay cue, but were making economic decisions based on the immediate delay offered.

6.3 Methods

6.3.1 Animals

Four Fisher Brown Norway rats aged 10–12 months at the start of behavior were used in this experiment. Rats were maintained at above 80% of their free-feeding weight. All experiments followed approved NIH guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Minnesota.

6.3.2 Experimental Design

The Restaurant Row task consisted of a central ring (approximately three feet in diameter) and four spokes leading off of that ring (Figure 6-1). At the end of each spoke, a feeder (MedAssociates, St. Albans VT) dispensed two 45mg food pellets of a given flavor (*banana, cherry, chocolate, and unflavored[plain],* Research Diets, New Brunswick, NJ). A given flavor remained at a constant spoke through the entire experiment. A trigger zone was defined so as to include a spoke and the portion of the ring nearby. Zones were primed in a sequential manner so that the rat ran in one direction around the loop. When the rat entered a primed zone, a tone sounded indicating the delay the rat would have to wait in that zone to receive food. Offered delays ranged from 1 second (identified by a 750Hz tone) to 45 seconds (12kHz). As long as the rat remained in the active zone, a tone sounded each second, decreasing in pitch (counting down by

250Hz increments). If the rat left the zone, the tones stopped, and the next zone in the sequence was primed.

In practice, rats waiting out a delay would proceed down the spoke and wait near the feeder; rats skipping a zone would proceed directly on to the next trigger zone. Each rat ran one 60 minute session each day. Zones were defined as a box around each reward location and extended onto the circular inner portion of the maze such that a rat was required to pass through a zone. Each reward arm extended approximately 22 inches away from the circular portion of the track. During training, rats were allowed to run the task in any manner they saw fit. However, rewards were only available if they traveled through the zones sequentially, Zone1 to Zone2, to Zone3, to Zone4. If a rat were to travel backwards the rat would have had to complete approximately three laps in order to prime the previous zone. Rats quickly learned that this behavior was not viable. Within seven days, rats learned to travel in only one direction and to pass through each zone sequentially.

Rats were initially handled and accustomed to the different flavors as described previously [328]. Rats were shaped to the task in three stages. In the first stage, all offers were 1 second. Once rats ran 30 laps/session consistently, they progressed to the second stage. In the second stage, each offer was randomly chosen from 1 to 10 seconds (uniform distribution, independent between encounters). Again, once rats ran 30 laps/session consistently, they progressed to the third stage, in which they faced the full

116

Restaurant Row task with offers selected between 1 and 30 seconds (uniform distribution). Two rats often waited out the full 30 seconds at some locations, so delays were increased for those rats to range from 1 to 45 seconds.

Once rats were completing at least 50 laps/session on the full Restaurant Row task, they were implanted with hyperdrives targeting the ventral striatum and orbitofrontal cortex. Rats were then re-introduced to the task until running well. Each day, rats were allowed to run for 60 minutes and often completed upwards of 70 entries per zone. Rats received all of their food on the track each day.

6.3.3 Control task (4x20).

In order to confirm the economic nature of the Restaurant Row paradigm, two rats ran an additional task after completing all recordings. In this modified version, each rat ran one session per day which consisted of four blocks of 20 minutes/block. In each block, one reward site provided three food pellets (of its corresponding flavor), while the other three reward sites provided one food pellet (of their corresponding flavors). Delays ranged from 1 to 45 seconds (uniform distribution). Each of the four sites were the "3-pellet" site for one of the four blocks each day. Which site was improved in which block was pseudo-randomly varied across days. Rats were removed to rest on a nearby flower pot for 60 seconds between each block.

117

6.3.4 Surgery

Rats were implanted with a dual bundle 12 tetrode + 2 reference hyperdrive aimed at ventral striatum (6 tetrodes + 1 reference, M/L +1.8mm, A/P +1.9mm) and orbitofrontal cortex (6 tetrodes + 1 reference, M/L +2.5mm, A/P +3.5mm) (Figure 6-2)[328]. For two rats, the two targets were left vStr and left OFC, while for two rats the two targets were right vStr and right OFC. Following surgery, tetrodes were turned daily until they reached vStr and OFC. Upon acquisition of large neural ensembles and a return to stable behavior on the maze, each rat ran a minimum of 10 recording days. Data reported here came from a total of 47 sessions distributed evenly over the four rats, R210: 12 sessions, R222: 12 sessions, R231: 13 sessions, R234: 10 sessions
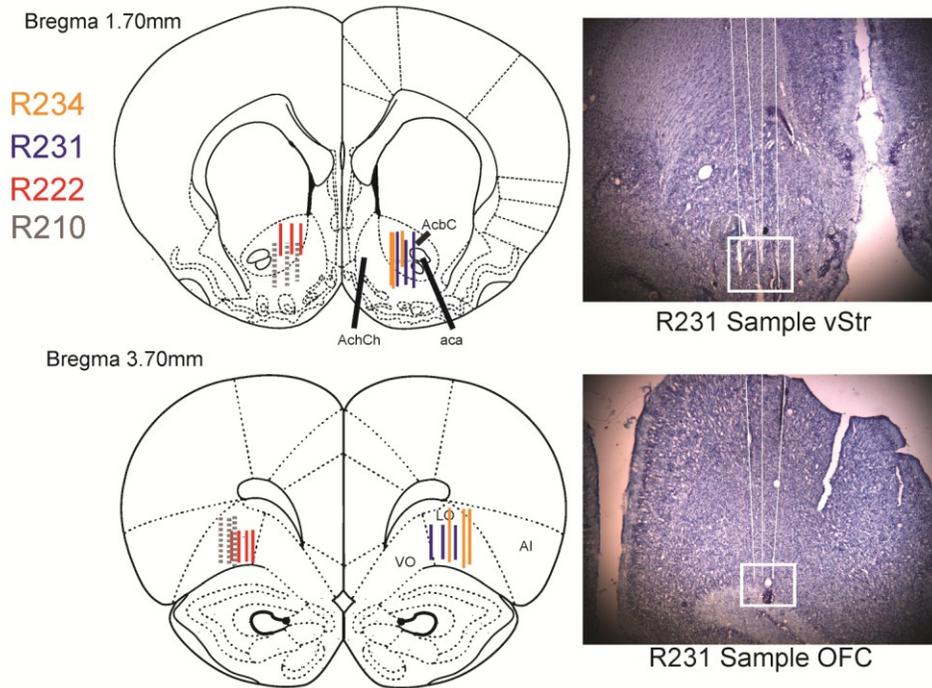
Figure 6-2 **Representative recording locations from OFC and vStr.**

Rat Number	OFC (number of cells)	Vstr (number of cells)
R210	225	243
R222	329	43
R231	336	112
R234	235	61

Table 6-1 **Number of cells recorded from each structure separated by rat**. [327]

119

6.3.5 Data Analysis

<u>Behavior</u>

Threshold calculation. At every encounter with a reward zone, the rat could <u>wait</u> through the delay or <u>skip</u> it and proceed to the next zone. Skips occurred throughout the duration of each session (Figure 6-3). If the rat chose to skip, it tended to do so quickly (Figure 6-4). As can be seen in Figure 6-1, rats tended to wait for short delays and skip long delays, as expected. In order to determine the threshold, we defined stays as 1 and skips as 0 and fit Sigmoid functions of stay/skip as a function of delay using a least squares fit (Matlab, MathWorks, Natick MA). The threshold for "above/below" calculations was defined as the mid-point of the Sigmoid. We determined a threshold for each rat for each session for each zone. All preference data was measured during the task and each rat demonstrated a different preference indicated by the amount of time that rat was willing to wait for reward.
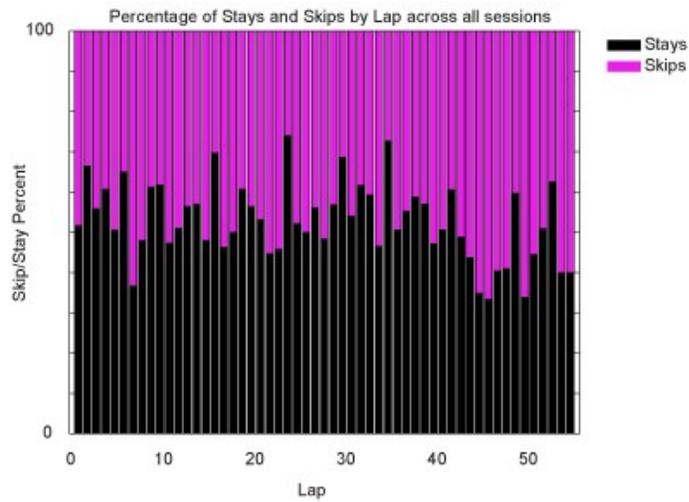
Figure 6-3 **Percentage of stays and skips on a given lap across all sessions and zones**.
The number of stays that occurred on a given lap did not significantly vary across all sessions. On each lap
the number of passes the rat stayed was compared across all sessions for that lap. For example, on lap 1,
50% of all passes resulted in a stay. [327]

VTE. Vicarious trial and error (VTE) was measured as the integrated absolute angular change in the orientation of motion of the head, as measured by sequences of head position samples (Papale , 2012; Steiner and Redish, 2012). This measure was calculated through a short algorithm sequence: first the position of the head $<x,y>$ was sampled at 60 Hz via the Cheetah Neuralynx system. Change in head position $<dx,dy>$ was calculated using the Janabi-Sharifi(Janabi-Sharifi , 2000) algorithm. Orientation of motion $<phi>$, was calculated as the arc-tangent of $<dx,dy>$. Change in orientation of motion was $<dphi>$ was calculated by applying the Janabi-Sharifi algorithm to $<phi>$. VTE was measured as the sum of the absolute value of $<dphi>$ over first two seconds of time after entering a zone $<IdPhi>$.

Time spent at delay

All Rats VTE aligned to Threshold

+/− Delay time from threshold (threshold at 0 s)

R210 Session 10
Lap 5

Zone4
Threshold 14(s)
Delay 1(s)
Stay

Zone3
Threshold 26(s)
Delay 12(s)
Stay

Zone1
Threshold 13(s)
Delay 15(s)
Skip VTE

Zone2
Threshold 13(s)
Delay 21(s)
Skip

VTE log ldPhi distribution for R210

R210 stay at zone 4 VTE 165

Threshold 14(s)
Delay 9(s)

Linear speed during behavioral sample

At Feeder
Receives Reward

R210 skip at zone 4 VTE 112

Threshold 14(s)
Delay 29(s)

R210 threshold at zone 4 VTE 1174

Threshold 14(s)
Delay 16(s)

R210 threshold at zone 4 VTE 2194

Threshold 14(s)
Delay 14(s)

At Feeder
Receives Reward

Log ldPhi value: higher = more VTE

Time from zone entry to zone exit (s)

123

Figure 6-4 **Behavioral Summary.**
First Row. In order to determine whether rats were waiting for a specific tone before leaving, we measured the time spent at each zone encounter over all rats, over all sessions. Graph shows number of seconds spent waiting as a function of the delay offer. A rat waiting out the entire delay would add into the x=y line; a rat leaving immediately would add into a cluster near the 0 duration waited. As can be seen in the histogram, rats tended to wait through the entire delay or leave after 3 seconds.
Vicarious trial and error (VTE). When rats encounter certain decisions, they sometimes pause and turn back and forth between the multiple options, as if deliberating between them(Muenzinger and Gentry, 1931; Muenzinger, 1938). In humans and other primates, a similar process can be seen in saccade-fixate-saccade (SFS) sequences (Padoa-Schioppa and Assad, 2006; Krajbich , 2010). Previous studies have found these VTE events to primarily occur during flexible (non-automated) behaviors (Muenzinger and Gentry, 1931; Johnson and Redish, 2007; van der Meer and Redish, 2009; Papale , 2012; Steiner and Redish, 2012), however, previous studies have not examined the relationship between VTE and decision difficulty.

Rats running the task showed three clear behaviors on encountering a new spoke – they sometimes just ran down the spoke to sample the food-delivery site, they sometimes skipped the spoke, and they sometimes paused and expressed VTE at the decision-point. As noted above, sampling tended to occur when the delays were below the threshold that rat had for that flavor, while skips tended to occur when the delays were above threshold. We quantified VTE through a measure of the integrated angular velocity of the head position of the animal(Steiner and Redish, 2012). We found that VTE tended to occur at the threshold, decreasing dramatically when the delay was less than threshold (generally a sample), but also decreasing when the delay was greater than threshold (generally a skip), (blue dotted lines, bottom plot; Linear Regression, R^2=0.95, P < 0.001 pre threshold; R^2=0.76, P<0.001 post threshold).
Several behaviors were typical when rats encountered a delay upon entering a zone.

Second Row. If the delay was above threshold, rats would often skip the zone relatively quickly (decrease in speed at 1 second followed by increasing speed after 2 seconds), spending little time in the current zone. VTE on these passes was typically low. Third Row. If rats decided to stay, they generally proceeded to the reward site and waited until the tone counted down and reward was delivered (as indicated by the very low average speed for the remainder of the time in zone). Again, on these passes VTE was typically quite low. Fourth Row. On close to threshold delays, rats demonstrated stronger VTE. If rats chose to sample the reward, they would proceed towards the feeder and wait through the remainder of the delay (early fluctuation in speed indicates high VTE, followed by decrease, near 0 cm/s speed indicates the rat has arrived at the feeder location where he remains until reward is received). Fifth Row. Again if rats encountered a close to threshold delay and this time chose to skip the reward, VTE remained high. Rats remained stationary for a longer period of time (from 1 to 6 seconds) before finally locomoting and leaving the current zone for the next zone. [327]

Identifying regret-inducing and control situations. On entry into a given

("current") zone, we defined the situation as <u>regret-inducing</u> if it met the following three

conditions: (1) The offer at the previous zone was a delay < threshold for that previous

zone for that rat for that session. (2) The rat skipped the previous offer. (3) The offer at the current zone was a delay > threshold for that current zone for that rat for that session.

The first control was defined using the same criteria as for regret-inducing situations, but that (2) The rat took the previous offer. This control situation keeps the sequence of offers the same, but controls for the rat's agency/choice.

The second control was defined using the same criteria as for regret-inducing situations, but that (1) The offer at the previous zone was a delay > threshold for that previous zone for that rat for that session. This control situation keeps the rat's choices the same, but makes the choice to skip the previous option the correct one.

Condition	Offer at previous	Rat's action	Offer at current
Regret-inducing	Delay < threshold	SKIP	Delay > threshold
Control 1	Delay < threshold	STAY	Delay > threshold
Control 2	Delay > threshold	SKIP	Delay > threshold
Rejoice	Delay < threshold	SKIP	Delay < threshold
Rejoice	Delay > threshold	SKIP	Delay < threshold
Regret	Delay > threshold	STAY	Delay < threshold

Table 6-2 **Summary of the different conditions**. [327]

Curvature. In order to identify the pause-and-look behavior, we measured the curvature of the path of the animal's head, and identified the point of maximum curvature and the direction of that point. Curvature was measured through the following

125

algorithmic sequence: the position of the head was measured at 60 Hz from the LEDs on the headstage via the camera in the ceiling, giving <*x,y*> coordinates, velocity <*dx,dy*> was calculated using the Janabi-Sharifi algorithm acceleration <*ddx, ddy*> was calculated by applying the Janabi-Sharifi algorithm [124] to <*dx,dy*>. Finally, the curvature at each moment was defined as[113]

$$\frac{dx * ddy + dy * ddx}{(dx^2 + dy^2)^{1.5}}$$

Neurophysiology

Cells were recorded on a 64 Channel Analog Cheetah-160 Recording system (Neuralynx, Bozeman MT) and sorted offline in MClust 3.5 (Redish, current software available at http://redishlab.neuroscience.umn.edu/). For all sessions, the position of the rat was tracked via overhead camera viewing colored LEDs on the headstage.

Reward Responsiveness. We are interested in determining how a cell modulates its activity during reward delivery. To measure this quantitatively, we compared the firing rate of the cell in the 3s after reward-delivery to 500 randomly-selected 3s intervals throughout the task. If a cell's firing rate is different (whether increased or decreased) during reward delivery, then it carries information about reward delivery. We can measure this change by determining if the activity during the 3s after reward-delivery was significantly different than the bootstrap. Because these distributions were not

126

Normal, we used a Wilcoxon to calculate significance. Responsiveness to each reward site was calculated independently.

Bayesian decoding. Any decoding algorithm consists of three parts: (1) a training set of tuning curves which defines the expected firing as a function of the variable in question, (2) a test set of spikes or firing rates, and (3) the posterior probability calculated from (1) and (2). In this manuscript, we used two decoding processes – one in which the tuning curves were defined as the neural activity in the 3 seconds after reward-delivery at the four reward-locations [*p(Reward)*], and one in which the tuning curves were defined as the neural activity in the 3 seconds after initial cue-delivery (zone entry) [*p(Zone)*]. When calculating *p(Zone)*, time after reward delivery was not included. This was only important for delays < 3 seconds.

p(Reward). Throughout the paper we refer to this measure as "*p(Reward)*", however, mathematically, it is *p(reward|spikes)*. Assuming a uniform distribution of reward priors, this equation is:

$$p(reward|spikes) = \frac{p(spikes|reward) * 0.20}{p(spikes)}$$

We defined the training set of *p(spikes|reward)* as the firing rate during the 3s after a given reward delivery (e.g. *p(spikes|banana), etc.*). In order to provide a control for unrelated activity, we also included a fifth condition in our calculation, defined as

127

average firing rate during times the animal was not in any countdown zone. Thus, the training set consisted of five expected firing rates: firing rate after reward-receipt (1) at banana, (2) at cherry, (3) at chocolate, and (4) at nonflavored, plus a fifth control of expected firing rate (5) on the rest of the maze. Because of the inclusion of the fifth (average firing rate when not at reward) condition, the normalization factor is 0.20.

p(Zone). Throughout this chapter we refer to this measure as "*p(Zone)*", however, mathematically, it is *p(zone|spikes)*. Assuming a uniform distribution of reward priors, this equation is:

$$p(zone|spikes) = \frac{p(spikes|zone) * 0.20}{p(spikes)}$$

We defined the training set of *p(spikes|zone)* as the firing rate during the 3s after entry into a given trigger zone (e.g. *p(spikes|banana-zone), etc.*). In order to provide a control for unrelated activity, we also included a fifth condition in our calculation, defined as average firing rate during times the animal was not in any trigger zone. Thus, the training set consisted of five expected firing rates: firing rate after zone-entry (1) at banana, (2) at cherry, (3) at chocolate, and (4) at nonflavored, plus a fifth control of expected firing rate (5) on the rest of the maze. Because of the inclusion of the fifth (average firing rate when not in any trigger zone) condition, the normalization factor is 0.20.

Calculating representations of previous, current, next opposite. In order to average across passes between different rewards, we first calculated the posterior probability for a given question (e.g. *p(Reward)* or *p(Zone)*) separately for each restaurant or zone. We then rotated the results based on the zone/reward in question to define a *current* zone/reward (the one the rat is currently encountering), a *previous* zone/reward (the one the rat had just left), a *next* zone/reward (the one the rat would encounter next), and an *opposite* zone/reward. By utilizing ensemble decoding, we can effectively ask what recorded neurons are representing with the highest probability, taking into account both increases and decreases in firing rate.

6.4 Results

6.4.1 Revealed preferences

Four rats were trained on the Restaurant Row task (Figure 6-1). All four rats showed similar behaviors in that they were likely to wait through the delay for delays less than a threshold, but unlikely to wait through the delay for delays greater than a threshold. Skips and stays were evenly distributed throughout each session (Figure 6-3**)**. When rats skipped an option, they left within the first ~5 seconds, independent of delay (Figure 6-4**).** The threshold between waiting and skipping tended to be different for the different flavors for a given rat (Figure 6-1**b-e**). These thresholds were consistent within

129

rat, but differed between rats (Figure 6-1**f-i**) and where consistent across zones across days (Figure 6-5, Figure 6-6, Figure 6-7, Figure 6-8, Figure 6-9**,** R210: ANOVA p = 0.99 R222: ANOVA p = 0.02 (no sessions significant after multiple comparisons) R231: ANOVA p = 0.14 R234: ANOVA p = 0.0005 (only session 2 was significantly different than sessions 5-10 but not different than 1, 3 and 4, otherwise, no other sessions were significantly different), indicating an underlying revealed preference for each flavor of food that did not change across a given session (Figure 6-5**)**. There were no differences in reward handling between delays, rats generally waited 20-25 seconds after consuming reward before leaving for the next zone (Figure 6-10).
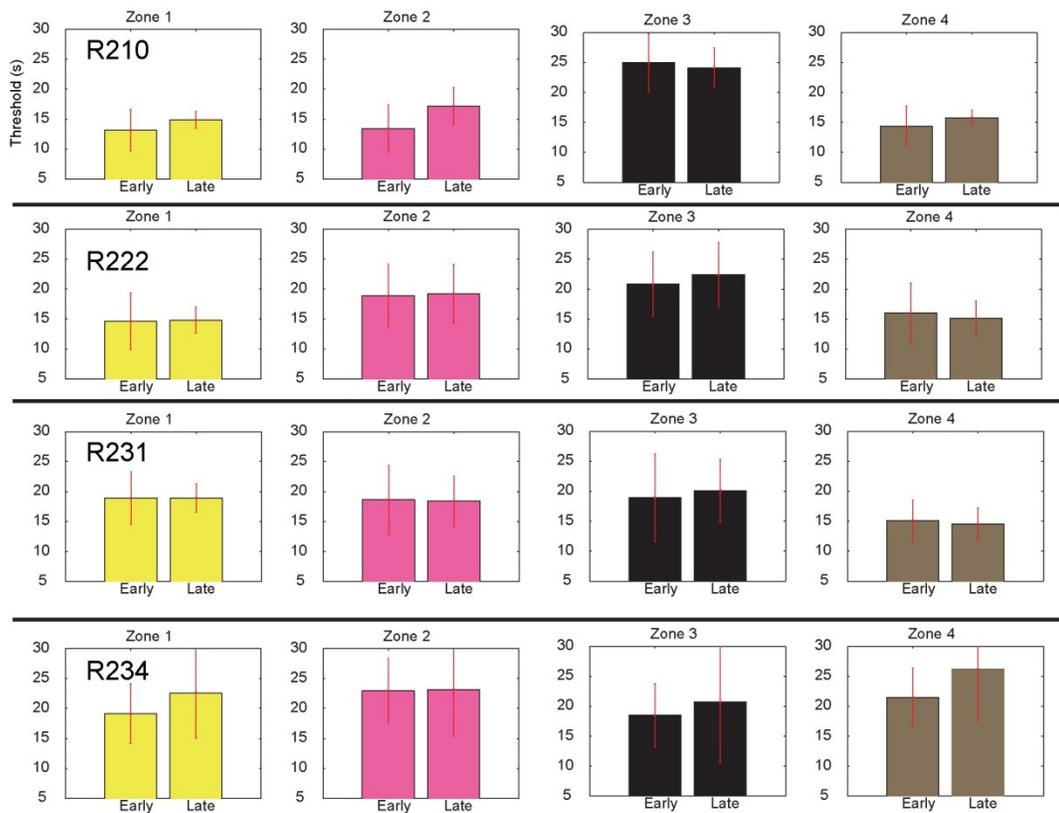
Figure 6-5 **Comparison of thresholds within session by rat**.
Thresholds were consistent within each session. If we compared the thresholds from the first half to the second half, no thresholds were significantly different between the first and second half of each session. [327]
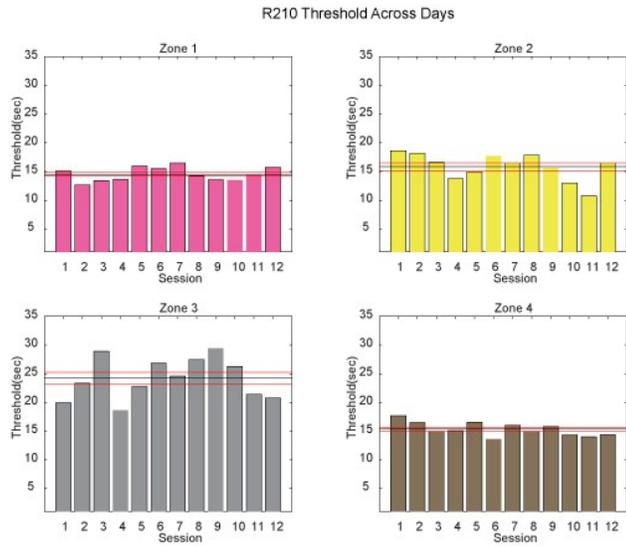
Figure 6-6 **R210 Thresholds across session by zone**.
Black lines indicate the overall average threshold across all sessions. Red lines indicate the standard error of the mean. Thresholds for each zone, for each rat were consistent across days. [327]
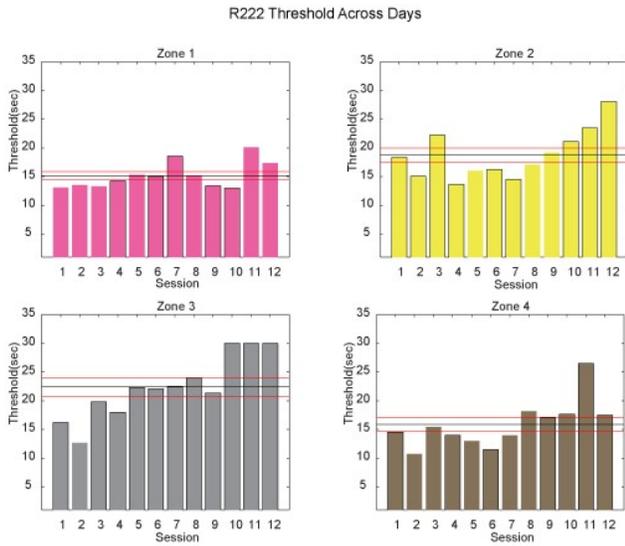
Figure 6-7 **R222 Threshold by session**. [327]

132
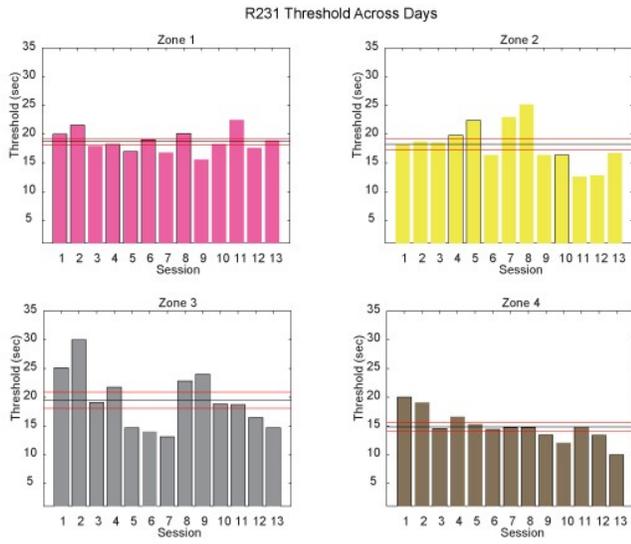
Figure 6-8 **R231 Threshold by session**. [327]
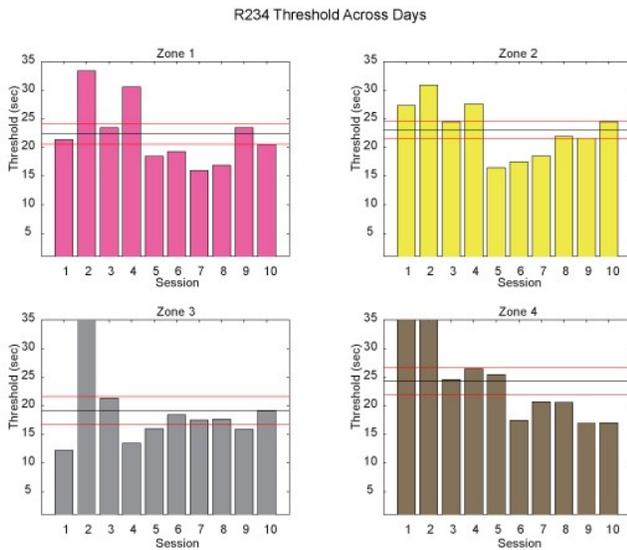
Figure 6-9 **R234 Threshold by session**. [327]
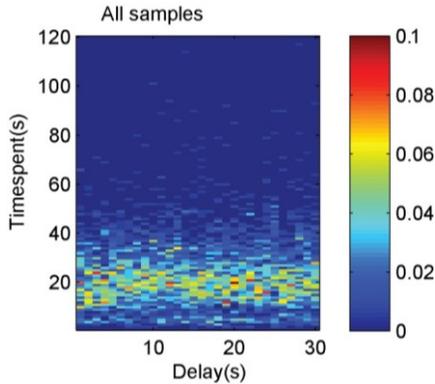
Figure 6-10 **Average time spent in zone after consuming reward.**
After consuming food, rats typically took 20-30 seconds before leaving the zone. This did not change as a function of the delay the rat had waited before receiving the food. [327]

In order to directly test whether the rats were making economic decisions (comparing value and cost), after completing the primary Restaurant Row experiment, we ran two of the rats on a variant of the task in which one reward site provided three times as much food as the other three sites. In this control task, rats were run on four 20-minute blocks, so that each site could be the large reward site for one block. (The order of which reward site provided excess reward was varied pseudo-randomly. Rats were removed to a nearby resting location for one minute between blocks.) Rats were consistently willing to wait longer for more food (Figure 6-11). All results reported here except for Figure 6-11 are from the primary Restaurant Row experiment.
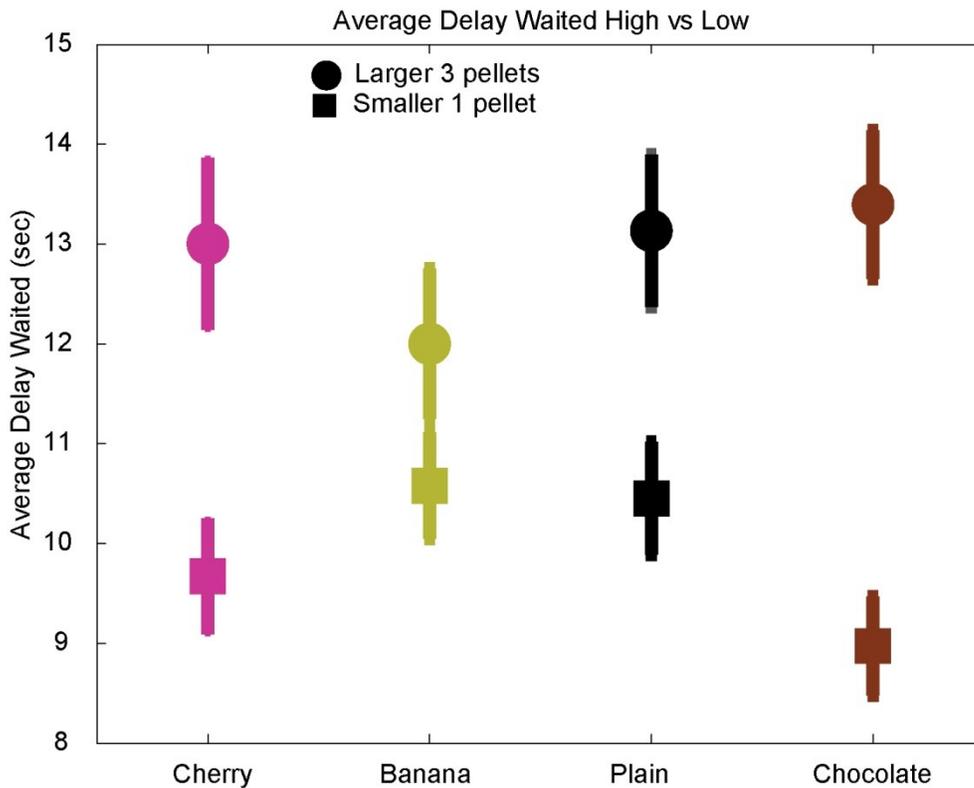
Figure 6-11 **Economic behavior on RRow.**

To determine if the rats took value into account when making decisions to stay or go (a key tenet in neuroeconomics [130,205,228]), two of the rats (R231 and R234) underwent an additional variation of the Restaurant Row task following completion on the unmodified version of the task. In this modified version, sessions consisted of four 20 minute blocks. During each 20 minute block, one reward flavor site dispensed three food pellets rather than two pellets (i.e. 3x 45 mg), while the other sites only dispensed one food pellet (i.e. 1x 45 mg). The four blocks allowed us to have each site be the "more valuable" site for one block. The order was randomly determined each day. Delays were randomly selected, as in the original task. Each 20 minute block was followed by a one minute rest, during which time the rat was removed to a small flower pot to the side. Each rat ran one complete session of four blocks per day.

Rats were willing to wait longer for the larger reward. This manipulation indicates that increasing the reward size increased the time rats were willing to wait, which implies that increasing reward size had more value, and that the rats were behaving economically.

There is no reason to expect the increase in the amount of time willing to wait for larger rewards to be linear. Subjective value depends upon the internally generated function for each reward (humans [150,169], rats **[2,26,372]**, primates[228]). Because rats, like humans, have preferences, we would expect that different rewards would have different values. However, the only real way to measure a value is by the choices that occur within a given context. By measuring the revealed preferences for each flavor we are essentially determining the subjective value of each reward.

135

Rats discount hyperbolically [187,188,231]. In addition, it has been shown that rats' preference saturates as the number of pellets increases. Thus, the amount of time a rat will wait for 4 pellets is not twice the time a rat will wait for 2 pellets [231]. We would not expect the value of 3 pellets to be exactly equal to 3x the value of 1 pellet. The amount a rat will consume at a given moment is not a linear relationship to the amount of food available. The time a rat would be willing to wait for 3 pellets should be greater than the time it would be willing to wait for 1 pellet. The time spent waiting for 3 pellets was larger than the average time spent waiting for 1 pellet. [327]

6.4.2 Reward responses

We recorded 951 neurons from orbitofrontal cortex (OFC) and 633 neurons from

ventral striatum (vStr). Neurons were identified as reward-responsive if their activity

during the 3s following reward delivery was significantly different (P<0.05, Wilcoxon)

than a bootstrapped (n=500) sample of activity during 3s windows taken randomly across

the entire session [328,354]. 81% of OFC neurons responded to reward; 86% of vStr neurons

responded to reward. Responses in both OFC and vStr often differentiated between the

four reward sites. (see Figure 6-12, Figure 6-13).
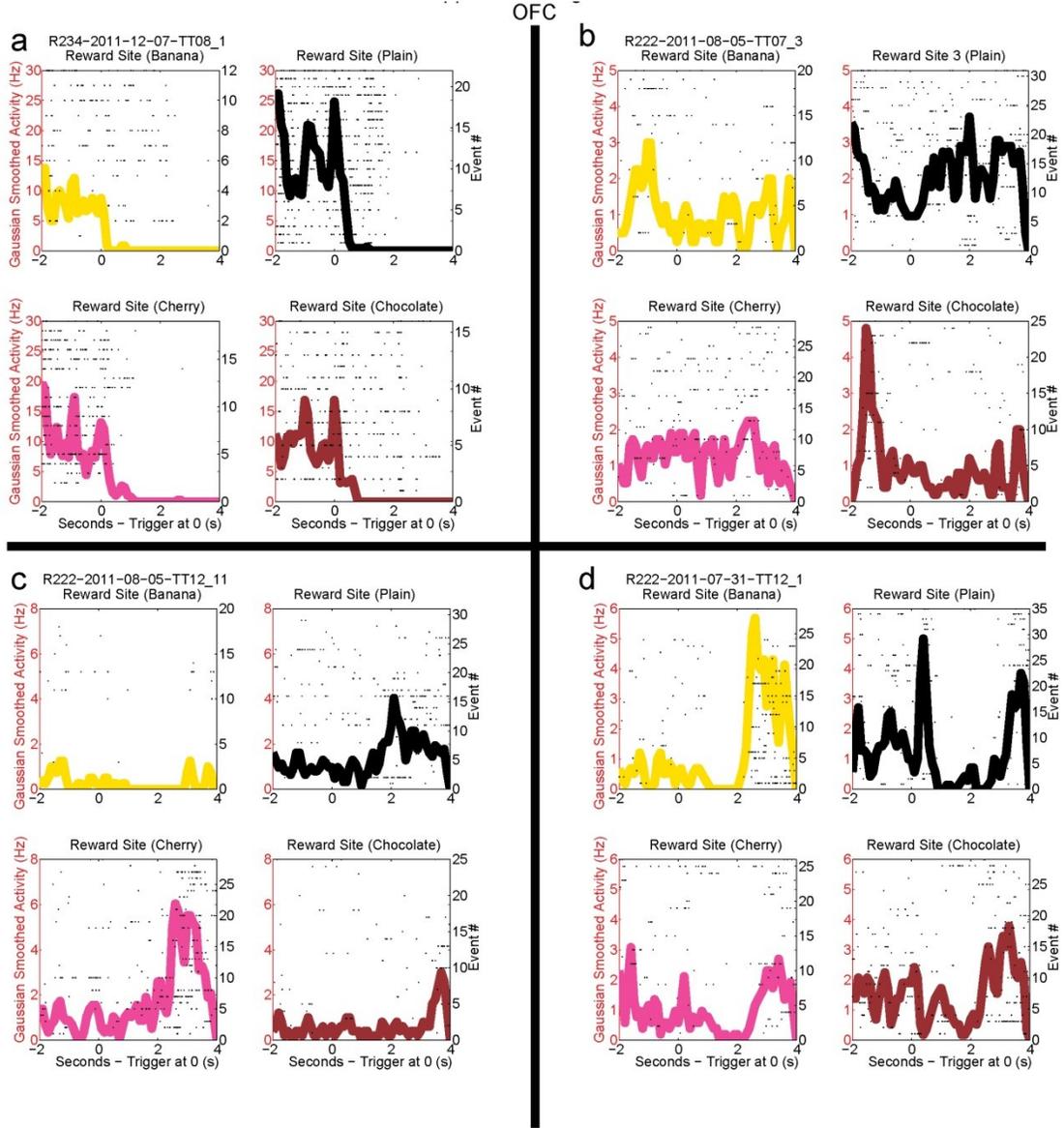
Figure 6-12 **OFC cells reliably represent different flavors of reward**.
Example reward-related cells from orbitofrontal cortex (OFC). Each super-panel (**a,b,c,d**) shows firing from a single cell. Within each super-panel, each subpanel shows that cell's response around the time of reward-delivery. The color of the trace indicates flavor (yellow = banana; black = plain/non-flavored; magenta = cherry; brown = chocolate) and the trace itself indicates the response (in spikes/sec) of the cell. Small dots indicate spikes on individual reward-delivery events. Trace shows average firing over all events, smoothed with a Gaussian window (sigma=50 ms). As can be seen in these examples, different cells responded differently (but reliably) to the different flavor-reward-sites. [327]
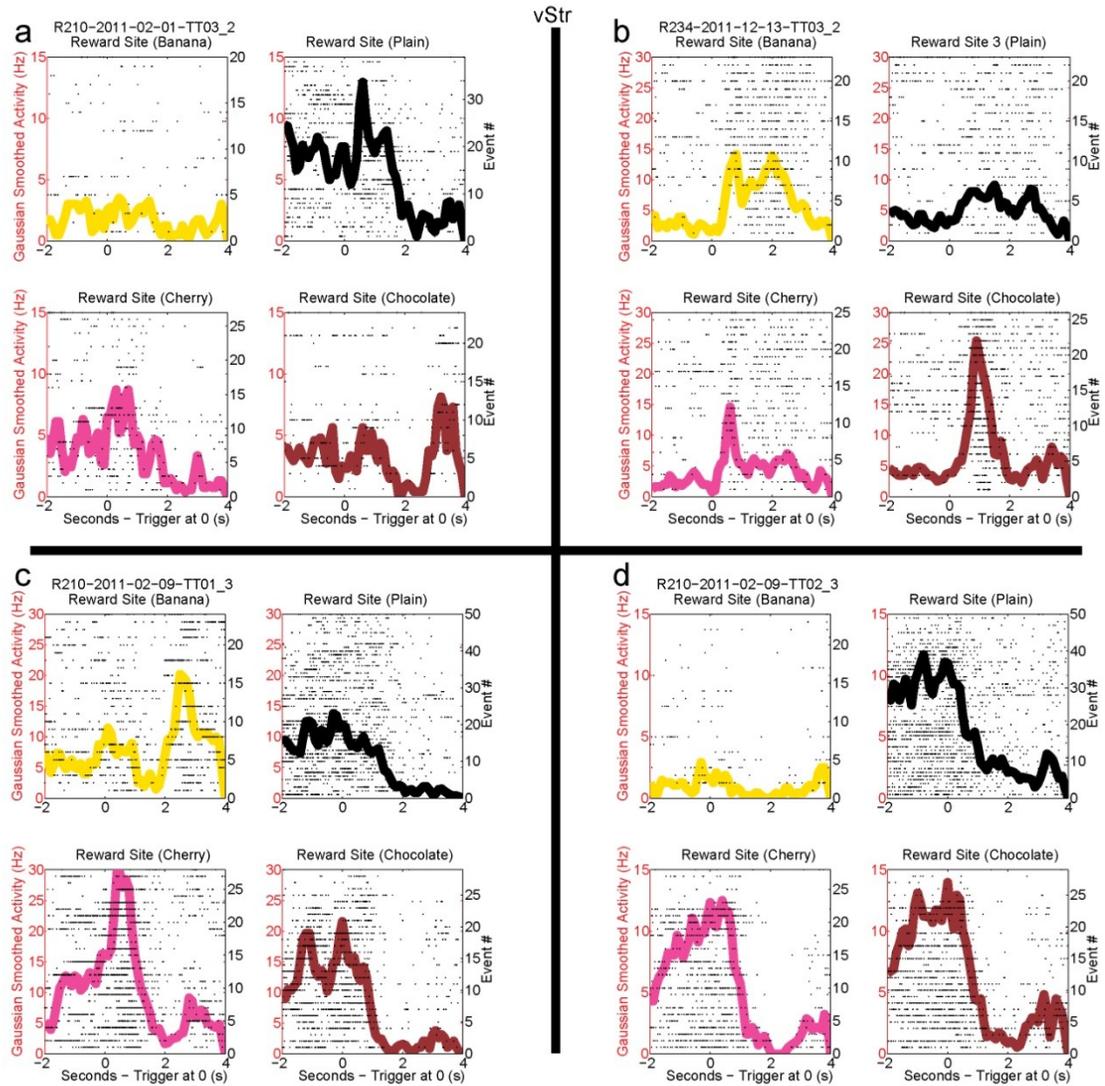
Figure 6-13 **vStr cells reliably represent different reward flavors.**
Example reward-related cells from ventral striatum (vStr). Display as in Figure 6-12.

Because responses differentiated between rewards, a decoding algorithm applied to these neural ensembles should be able to distinguish between the reward sites. We used a Bayesian decoding algorithm [380] with a training set defined by the neuronal firing rate in the 3s following delivery of reward (which we used to calculate *p(spikes|Reward)*) or a training set defined by the neuronal firing rate in the 3s following entry into a zone (which we used to calculate *p(spikes|Zone)*). In order to provide a control for unrelated activity, we also included a fifth condition in our calculation, defined as average neuronal firing rate during times the animal was not in any countdown zone. Thus, the training set consisted of five expected firing rates: firing rate after reward-receipt or zone entry (1) at banana, (2) at cherry, (3) at chocolate, and (4) at nonflavored, plus a fifth control of expected firing rate (5) on the rest of the maze. From this training set, Bayesian decoding uses the population firing rate at a given time to derive the posterior probability of the representation *p(Reward|spikes)* or *p(Zone|spikes)*. For simplicity, we will refer to these two measures as *p(Reward)* and *p(Zone)*.

In order to pool data from all four sites, we categorized and rotated each reward site based on the current position of the animal. This gave us four sites that progressed in a serial manner – the <u>previous</u> site, the <u>current</u> site, the <u>next</u> site, and the <u>opposite</u> site (

Figure 6-14**c**). All analyses were based on this categorization. Throughout all analyses, we used a leave-one-out approach so that the encounter being decoded was not included in the definition of the training set. Both OFC (

139

Figure 6-14**a**) and vStr (

Figure 6-14**b**) were capable of reliably distinguishing between the current reward

site (see also Figure 6-15) and the other sites. Shuffling the interspike intervals of the

cells removed all of these effects. *p(Reward)* and *p(Zone)* calculated from shuffled data

were consistently at 0.14 (Figure 6-16). As shown in Figure 6-17 and

Figure 6-14, the ensemble reliably differentiated entries into the different zones as

well as the different rewards. As shown in

Figure 6-14**d,e** during normal behavior, the ensemble reliably represented the

current zone on entry into it and the current reward on receipt of it.

Zone entry responses. Previous research has suggested that in simple association

tasks in which cues predict reward, both OFC and vStr cells respond to cues predictive of

reward [128,193,266,269,297,299,313]. Both OFC and vStr neural ensembles distinguished the

different zones both at the time of reward (

Figure 6-14**a-c**) and at the time of entering the zone/tone sounding (

Figure 6-14**d-f**). (for single cell differentiation see Figure 6-12 and Figure 6-13**,**

decoding differentiation Figure 6-15and Figure 6-17). These representations were related;

thus, neural activity in OFC and vStr also predicted the reward type of the current zone

during zone enter/cue onset (

Figure 6-14**g-i,** Figure 6-18**a,b**). Shuffling the interspike intervals eliminated these effects (Figure 6-18**c,d**).
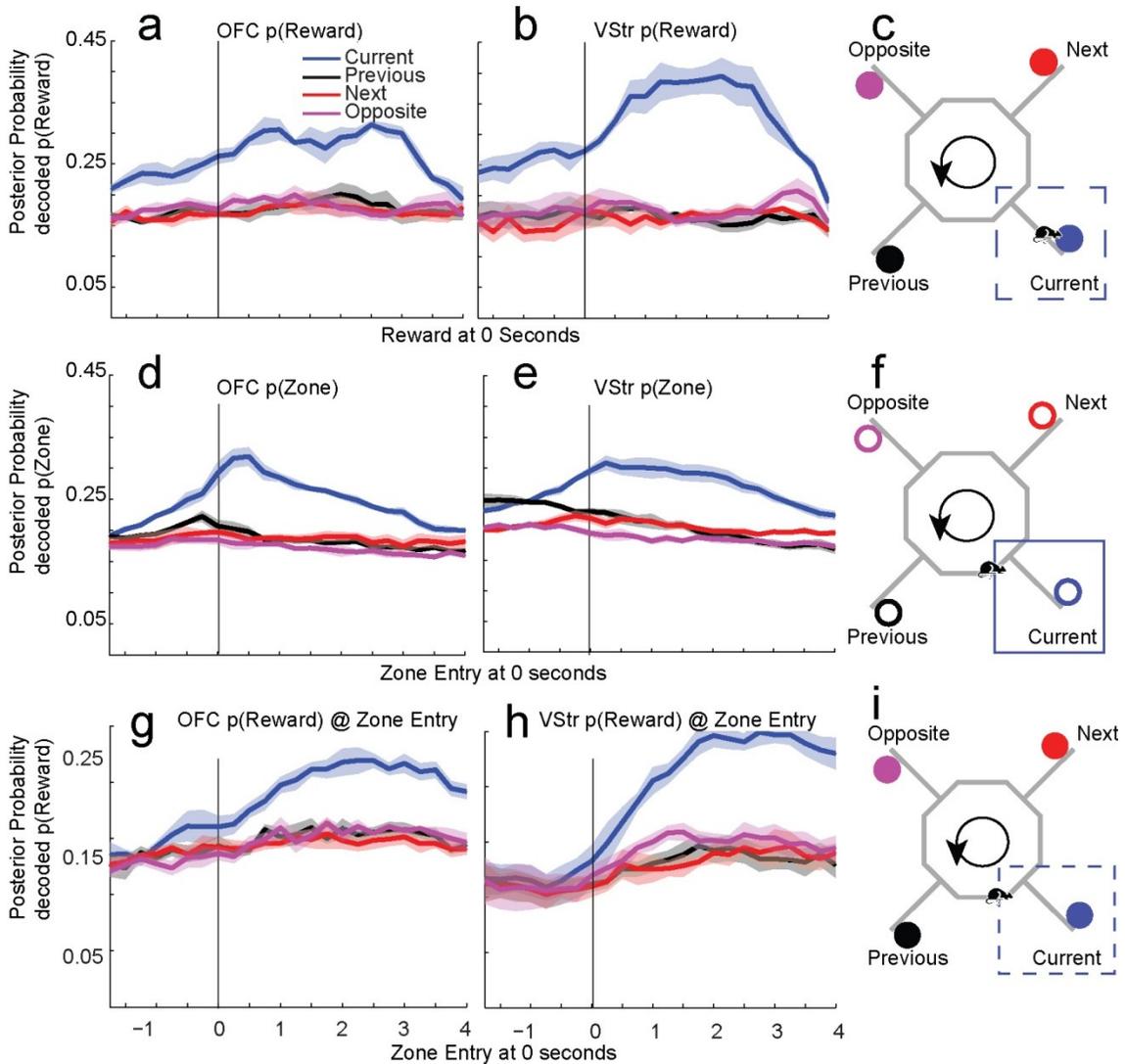
Figure 6-14 **Ensembles in OFC and vStr representing the current reward and the current zone**.
a-b, *p(Reward) @ Reward*, Defining the training set for decoding as firing at reward delivery and the test set as firing at each moment surrounding reward-delivery, the neural ensemble decoded the current reward reliably (distribution of current reward was determined to be significantly different, empirical cumulative distribution function, alpha = 0.05). *p(Reward)* is the posterior probability indicating representation of a

141

given reward flavor as calculated from a Bayesian decoding operation as per <u>Methods</u>. **c**, Cartoon indicating that the training set is the set of reward types, and the test set is firing when the rat receives reward. (A leave-one-out approach prevented the tautology.) **d-e**, *p(Zone) @ Zone*, Defining the training set for decoding as neuronal firing at zone entry and the test set as neuronal firing at each moment surrounding zone-entry, the neural ensemble decoded the current zone reliably. *p(Zone)* is the posterior probability indicating representation of a given zone entry as calculated from a Bayesian decoding operation as per <u>Methods</u>. **f**, Cartoon indicating that the training set is zone entry, and the test set is neuronal firing when the rat enters the zone, triggering the cue that signals the delay. (A leave-one-out approach prevented the tautology.) **g-h**, *p(Reward) @ Zone*, Defining the training set for decoding as neuronal firing at reward-delivery and the test set as neuronal firing at each moment surrounding zone entry, the neural ensemble at time of zone entry decoded the current reward type reliably. **i**, Cartoon indicating that the training set is the reward flavor, and the test set is neuronal firing when the rat enters the zone, triggering the cue. [327]
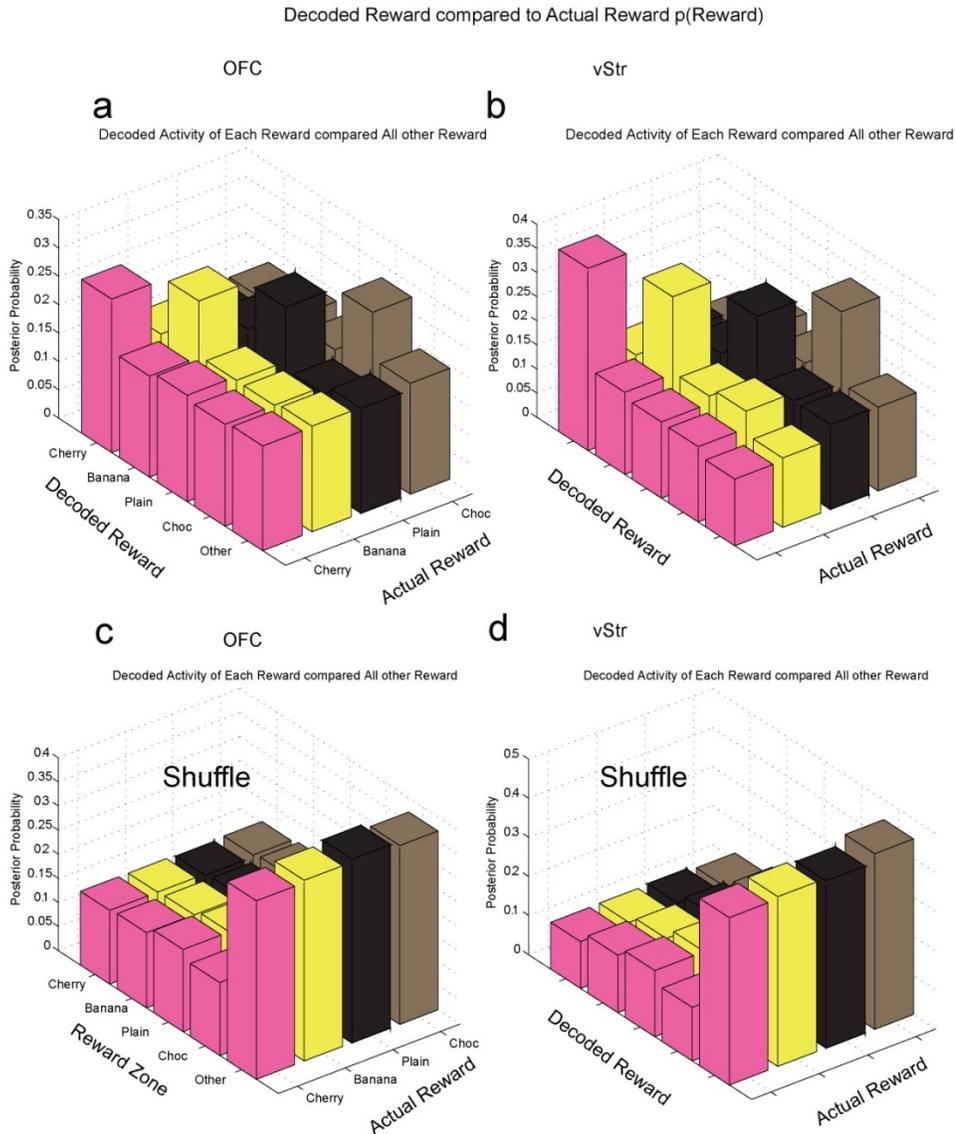
Figure 6-15 **OFC and vStr ensembles track the current reward.**
Orbitofrontal (OFC) and ventral striatal (vStr) neural ensembles accurately tracked the rewarded flavor during reward receipt as identified previously [356-358]. Panels show the confusion matrices of the decoding. We calculated p(Reward) @ Reward for each flavor, using a leave-one-out approach to avoid the tautology. Note that, as per Methods, the decoding returns five values, for each of the four flavors plus the fifth "other" condition. **a,b**, The strong increase in the identity comparison implies separate representations of each flavor-reward-site. **c,d**, Shuffling the interspike intervals of the cells removes these representations. [327]

143
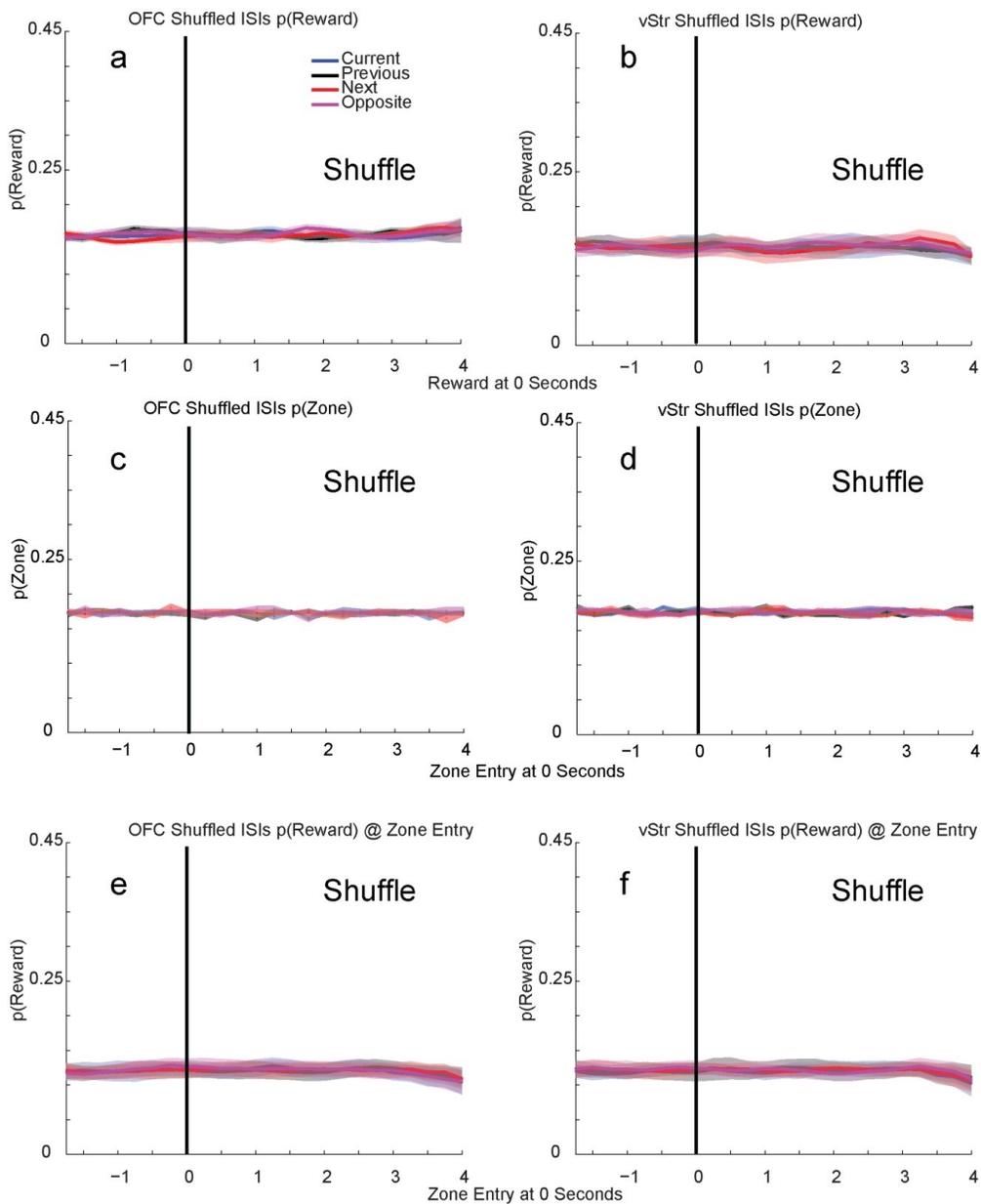
Figure 6-16 **Shuffles.**
To determine the chance level for the representations of
p(Reward) @ Reward, p(Zone) @ Zone, and *p(Reward) @ Zone*, we shuffled the interspike intervals.
Shuffling the interspike intervals preserves the firing characteristics of the cells but disrupts their alignment
to temporal events. Shuffling the interspike intervals for all cells during reward receipt produced a chance
level of ~0.14 for all conditions. [327]

Decoded Zone compared to Actual Zone p(Zone)
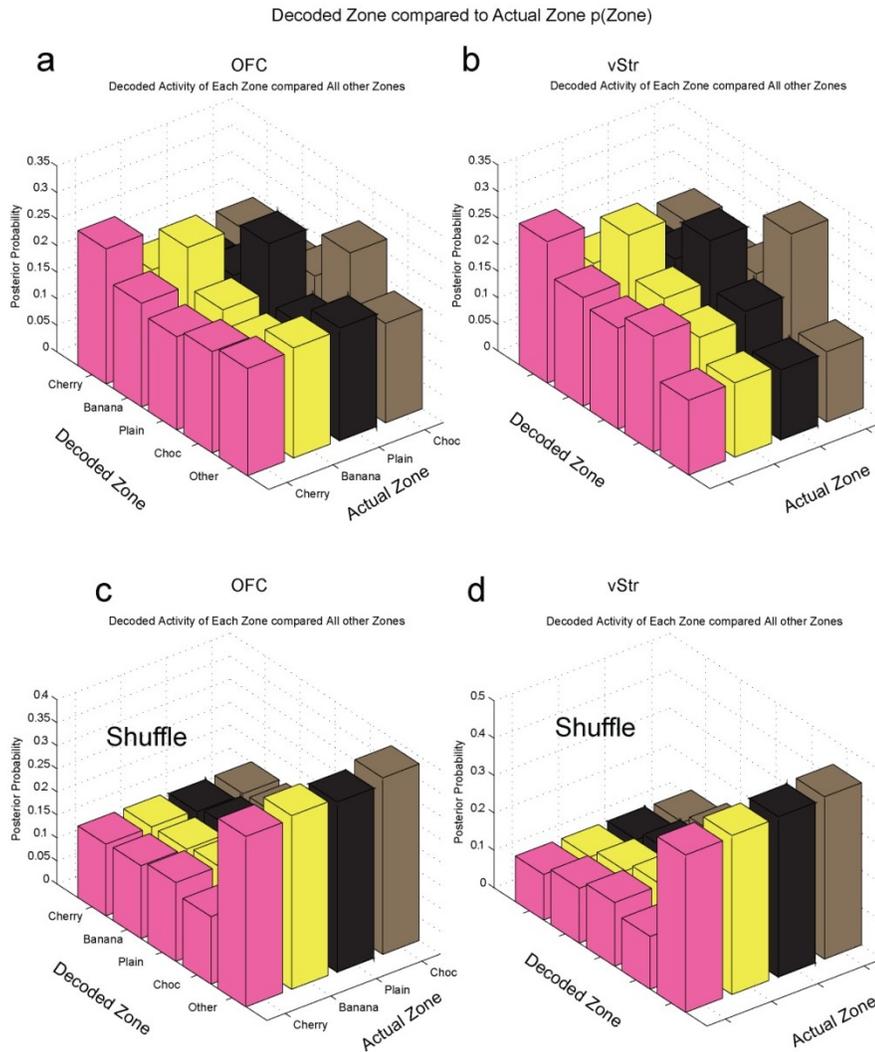
Figure 6-17 **OFC and vStr ensembles track the current zone/cue.**
Orbitofrontal (OFC) and ventral striatal (vStr) neural ensembles differentiated cue signals at entry into the different zones. As in Figure 5.11, panels show the confusion matrices of the decoding. We calculated *p(Zone) @ Zone* for each zone, using a leave-one-out approach to avoid the tautology. Note that, as per Methods, the decoding returns five values, for each of the four flavors plus the fifth "other" condition. **a,b**, The strong increase in the identity comparison implies separate representations of each trigger zone. **c,d**, Shuffling the interspike intervals of the cells removes these representations. [327]

145
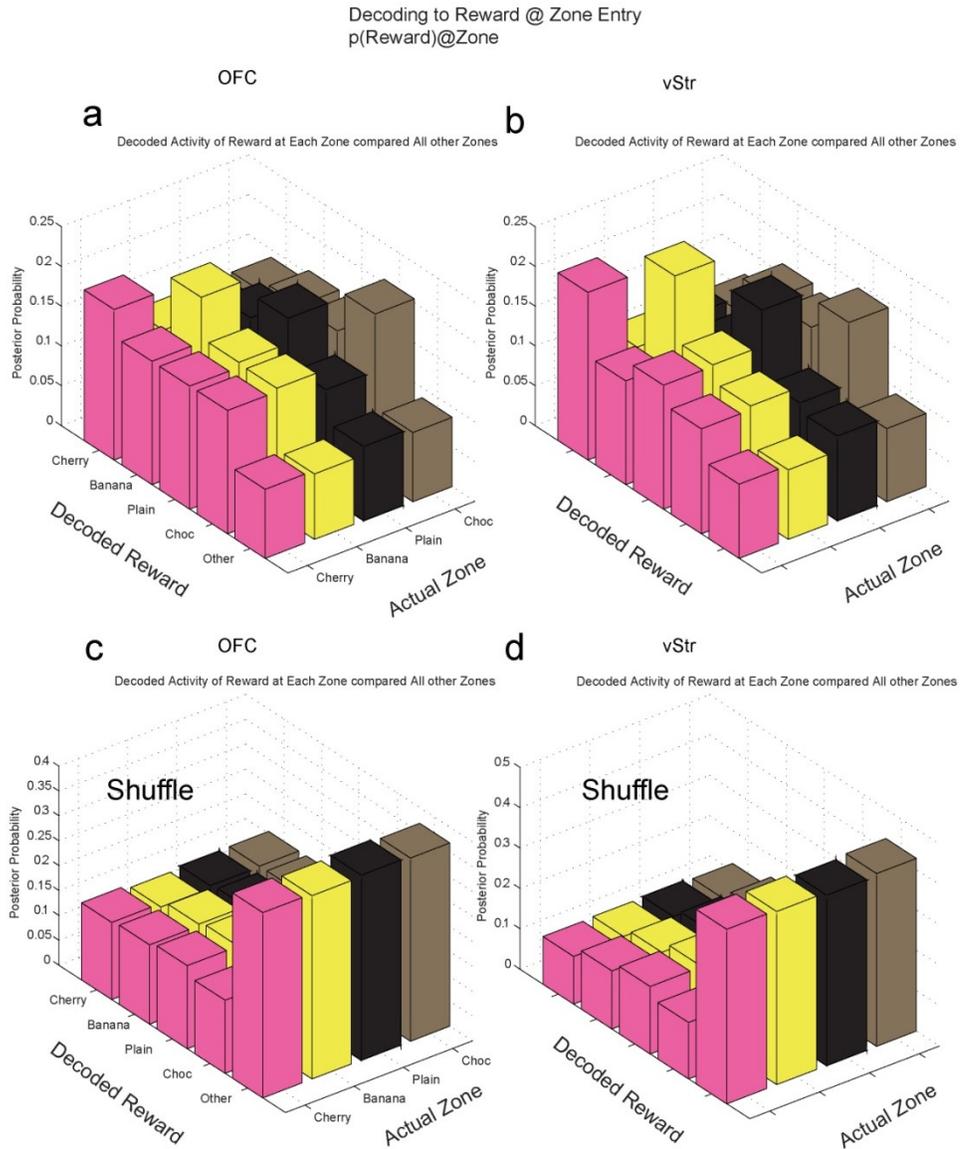
Figure 6-18 **OFC and vStr ensembles accurately track which reward is paired with each zone.**
To determine the relationship between cues and reward-related activity, we calculated the confusion
matrices for the decoding for *p(Reward) @ Zone*. **a,b**, The strong increase in the identity comparison
implies matched representations between each reward and zone. **c,d**, Shuffling the interspike intervals of
the cells removes these representations. [327]

Both OFC and vStr responded strongly under conditions in which the animal

determined the cost to be worth staying (e.g. when the delay was below threshold, Figure

6-19**a,b**). In contrast, neither structure represented expectations of reward under conditions in which the animal determined the cost to not be worth it (i.e. skips, when delay was above threshold, Figure 6-19**c,d**). This suggests that these structures were indicating expected value, and predicting future actions.

To directly test this hypothesis, we compared reward-related decoding when the rat encountered a delay near threshold (threshold +/-2 s) and either stayed to sample the feeder (Figure 6-20**a,b**) or skipped to proceed to the next reward option (Figure 6-20**c,d**). When the animal stayed (waiting for a reward), both OFC and vStr increased their representations of the current reward at the time of zone entry. In contrast, when the animal skipped the current reward, neither OFC nor vStr reliably represented the current reward/zone. Shuffling the interspike intervals of the cells removed these effects (Figure 6-21).
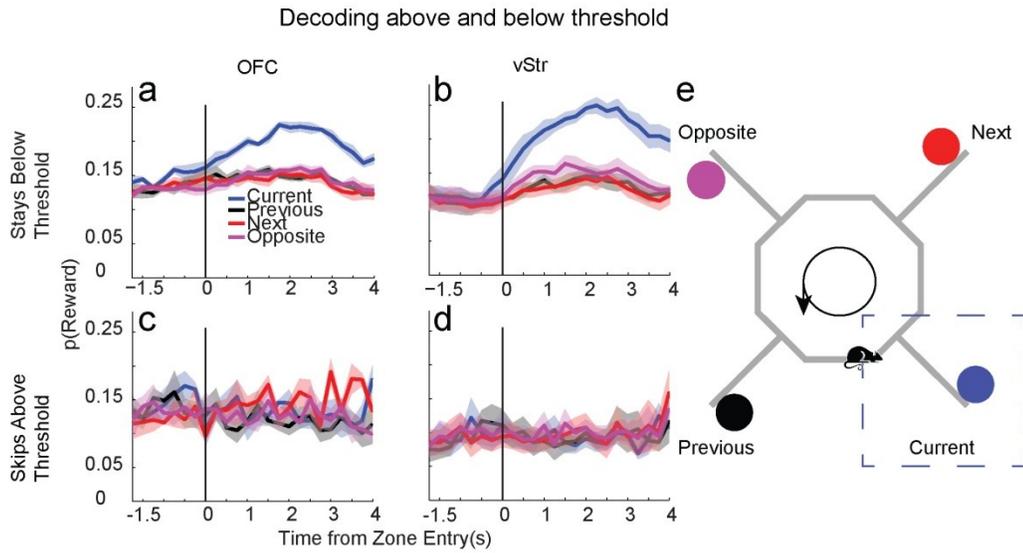
Figure 6-19 **Representations of expected reward as a function of delay and threshold**.
In order to determine whether orbitofrontal (OFC) and ventral striatal (vStr) signals predicted behavior at time of zone-entry, we measured *p(Reward) @ Zone* for all offers above and below the threshold for a given rat for a given flavor-reward-site. **a,b**, Low-cost offers in which the rat waited through the delay. **c,d**, High-cost offers in which the rat skipped out and did not wait through the full delay. **a,c**, OFC. **b,d**, vStr. **e**, Cartoon indicating that this decoding operation was based on a training set at the reward, but a test-set at zone-entry. [327]
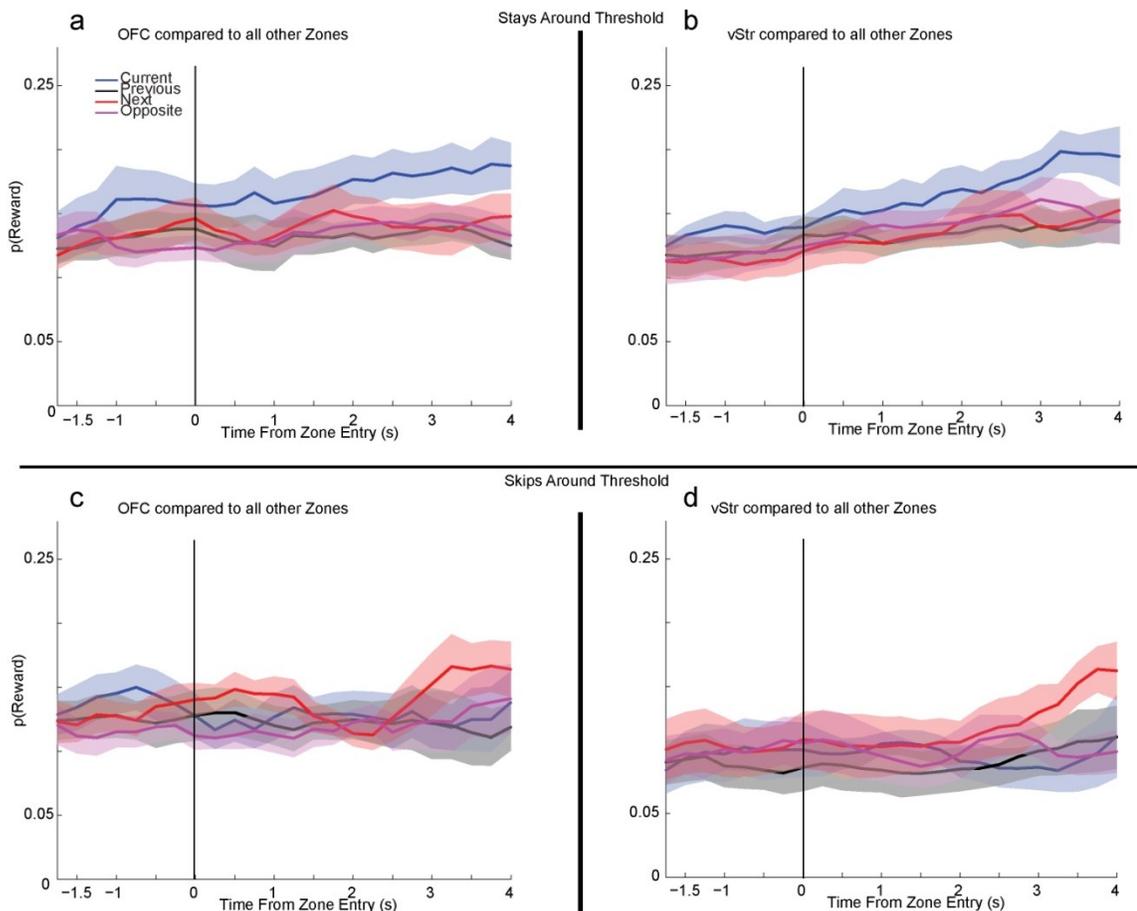
Figure 6-20 **Stay and skip decoding around threshold.**
In order to determine whether orbitofrontal (OFC) and ventral striatal (vStr) signals predicted behavior differentially for similar offers, we measured *p(Reward) @ Zone*, for all offers near threshold (delay within 2 seconds above or below threshold). **a,b**, Encounters in which the rat waited through the delay. **c,d**, Encounters in which the rat skipped out and did not wait through the full delay. **a,c**, OFC. **b,d**, vStr. Note that the current reward was better represented during stays than the other zones (**a,b**). In contrast, during skips, the current zone was not better represented; instead, the representations of the next zone began to appear after 2-3 seconds (**c,d**). [327]
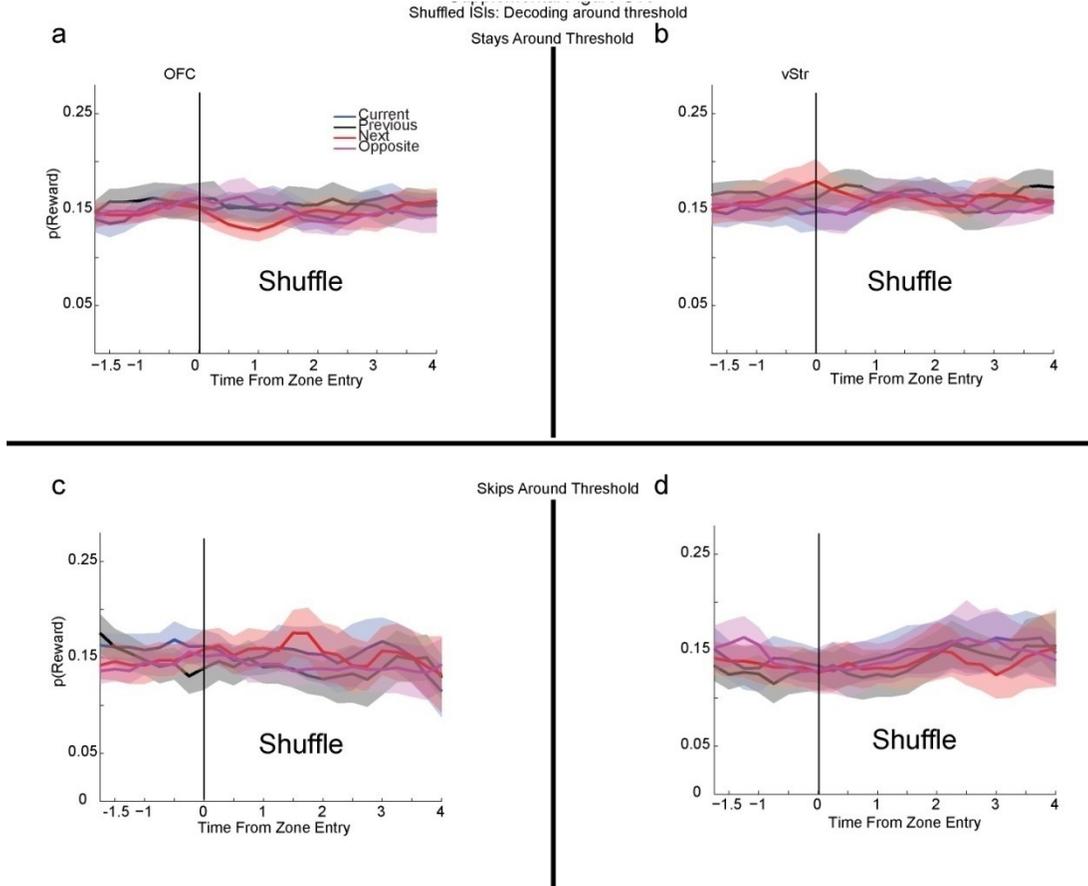
Figure 6-21 **Shuffles around threshold.**
Analysis of the same data shown in **Figure 5.16**, but with interspike intervals shuffled. Shuffling ISIs removed all effects. [327]

6.4.3 Regret

Regret entails the recognition that one has made a mistake, that an alternate action would have been a better option to take [23,174]. As noted above, a regret-inducing situation requires two properties be satisfied: (1) the undesirable outcome should be a result of the agent's previous action, and (2) following the selection of an option, the outcome/value of all options needs to be known, including the outcome/value of the unselected options.

150

Our task and behavior satisfies these conditions. Because the rats were time-limited on the Restaurant Row task, encountering a high-cost delay after not waiting through a low-cost delay means that skipping the low-cost delay was a particularly expensive missed opportunity.

In the Restaurant Row task, a rat would sometimes skip an offer that was less than that rat's threshold for that flavor on that day and then encounter an offer at the subsequent site that was greater than that rat's threshold for that flavor on that day. Because the delay is a cost and value is matched (by definition) at threshold, this sequence is one in which the rat skipped a low-cost offer, only to find itself faced with a high-cost offer. As discussed above, this sequence should induce regret [23,174]. From the economic and psychology literature, we can identify these sequences as potential "regret inducing" situations. We can compare these conditions to control conditions which we would expect to provide disappointment rather than regret.

The previous literature suggests that during regret, there should be manifest changes in the animal's behavior and neurophysiology that reflect *a recognition of the missed opportunity*, as well as subsequent behavior taking choices that one might not have made normally. Theoretically, the key to regret is a representation of the action not taken [38,95,156-158]. This implies that there should be representations of the previous choice during the regret-inducing situations, particularly in contrast to control conditions that are "merely disappointing".

151

Thus, we define a <u>regret-inducing situation</u> as one in which (1) the rat skipped a low-cost/high-value reward (delay less than measured threshold for that flavor for that day), and then (2) the rat encountered a high-cost/low-value reward (delay greater than measured threshold for that flavor for that day). In this situation, the rat has made an economic mistake: if it had taken a different action (waited for that previous reward), it would have had a more valuable session. For consistency, we will refer to the opportunity in (1) as the <u>previous zone/previous reward</u> and the opportunity in (2) as the <u>current zone/current reward</u>.

As noted above, one needs to differentiate regret-inducing sequences from sequences that are merely disappointing. By definition, a disappointing sequence occurs when one encounters a situation that is worse than expected, but not due to one's own agency. There are two controls that need to be taken into account, a control for the sequence of offers (control 1), and a control for the animal's actions (control 2).

To control for the sequence of offers, we took sequences in which the rat encountered the same sequence of offers, but took (stayed for) the first offer. This matched control should only induce disappointment (worse than expected, but not due to the fault of the rat) [22,173]. Control 1 differs from the regret-inducing situation only in that the rat took the previous offer rather than skipping it. In summary, control 1 was defined as situations in which the delay at the previous zone was below threshold and the rat waited for reward, followed by an encounter at the current zone such that delay was

above threshold. In this situation, the rat did not make a mistake (as it waited for reward at the previous zone); the delay at the current zone was merely worse than the rat was willing to wait for, making the rat (presumably) disappointed. Control 1 controls for the sequence observed by the rat.

To control for the rat's actions, we took sequences in which the offer at the previous zone was greater than threshold (and skipped), and, again, the rat encountered a higher-than-threshold offer at the current zone. In this second control condition, the rat skipped the previous offer, but that was the "correct" action to take, as the previous offer was above threshold. This second control condition should also induce disappointment because the rat has encountered two high-cost offers in a row. But this second control condition should not induce regret, because the rat's actions were consistent with its revealed preferences. Control 2 differs from the regret-inducing situation only in that the delay at the previous offer was above rather than below threshold. In summary, control 2 was defined as situations in which the delay at the previous zone was above threshold, followed by an encounter at the current zone such that delay was above threshold. In this situation, the rat did not make a mistake (since it skipped a high-cost delay at the previous zone), but the delay at the current zone was worse than the rat was willing to wait for, making the rat (presumably) disappointed. Control 2 controls for the reward sequence seen by the rat.

153

Potential regret and control instances were found within each session by comparing the delays at each of the zones to the threshold of that zone for that rat for that day. Regret instances and control instances were evenly distributed throughout each session across all rats. The distribution of the high-cost offers at the current zone did not differ between the potential regret-inducing sequences and matched controls (Figure 6-22).

a
Cumulative Distribution Function Matched Delays between Regret and Control 1

b
Cumulative Distribution Function Matched Delays between Regret and Control 2
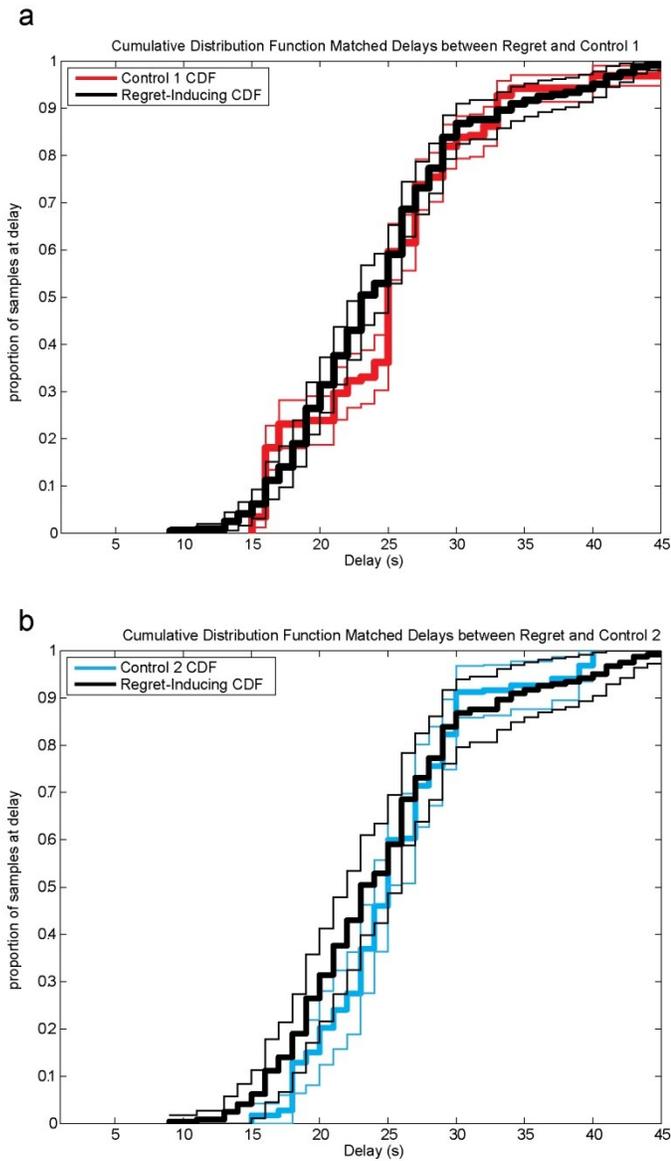
Figure 6-22 **Match samples during regret and controls.**
It is important to ensure that the current delay offers made in the matched control encounters had the same distribution as the regret-inducing instances. Graph shows the cumulative distribution function (CDF) of the "current" offers included in each condition. The distributions were closely matched, indicating that any results seen (e.g. Figure 6-26) were not a result of differences between the current offers. **a**, Regret-inducing vs. control 1. **b**, Regret-inducing vs. control 2. The thin lines on the empirical distribution plot represent the 95% confidence intervals (alpha = 0.05). Mann Whitney U tests indicated that the distribution of delays were not significantly different (vs control 1, **a**, p=0.20; vs control 2, **b**, p=0.11). [327]

155

Behaviorally, rats paused and looked backwards towards the previous option upon encountering a potentially regret-inducing sequence, but did not do so in either control condition (Figure 6-23). We identified pause-and-look events as points of high curvature and derived an orientation (see **6.3 Methods**). During potential regret-inducing sequences, rats were more likely to look backwards towards the previous option than during either of the matched control conditions ($P<0.05$, Watsons Circular U Test). In the first control condition (where the rat took a good offer and then encountered a bad offer), the rat tended to look towards the current zone, but then skip it and go on to the next zone. In the second control condition (where the rat encountered two bad offers in a row), the rat tended to look towards the next zone. Thus, there was a behavioral difference, implying that the rats recognized these three situations differently.
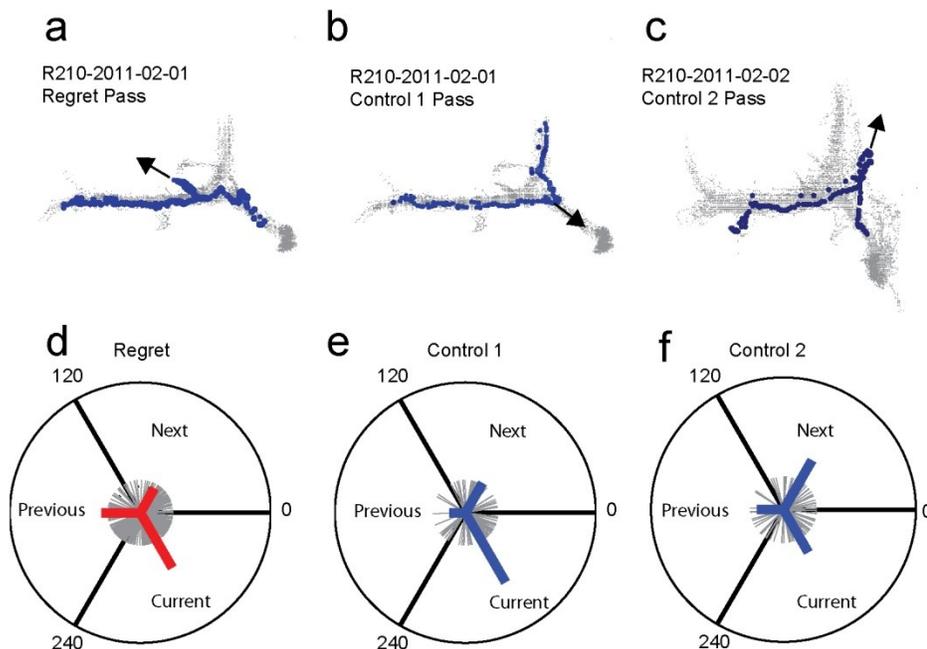
Figure 6-23 **Behavioral responses in regret-inducing and control situations.**

All passes were rotated so as to align on entry into a "current" zone. Orientation was measured using the curvature measure as per 7.3.5 Methods. **a-c**, examples of approaches for each of the three conditions: regret-inducing, control 1 (same sequence, took previous option), and control 2 (two long delays in a row). **a**, In a regret-inducing example, when the animal entered the zone, he paused and looked backwards towards the previous zone. **b,** In a control 1 example, the animal looked towards the current reward spoke, but proceeded on to the next zone. **c**, In a control 2 example, the animal looked towards the next zone, but turned back towards the current reward. **d-f**, Summary statistics. The first re-orientation event was measured as per 7.3.5 Methods. Grey traces show all pausing re-orientations over all instances within that condition. Heavy line shows vector average within each 120 degree arc. **d**, In the regret-inducing conditions, rats tended to orient towards the previous zone or current spoke. **e**, In the control 1 conditions, rats tended to orient only towards the current spoke. **f**, In the control 2 conditions, rats tended to orient towards the next zone. The distributions in **d**, **e**, and **f**, were significantly different from each other (Watson's Circular U, see text). [327]

During potential regret instances individual reward-responsive neurons in OFC and vStr showed firing patterns more consistent with the previous reward than the current one (Figure 6-24). Neural activity peaked immediately following the start of the look back towards the previously skipped, low cost reward. In order to quantify these changes in representation during regret-inducing situations and disappointment-inducing controls, we examined the population dynamics using a Bayesian decoding algorithm. Population decoding analyses offer insight into the dynamics of neural population. Ensemble activity more accurately represents the dynamics of the entire population much compared to that of a single cell. In order to determine the neural population representation during these situations, we utilized all the cells and measured the Bayesian representations of *p(Reward)* and *p(Zone)*.
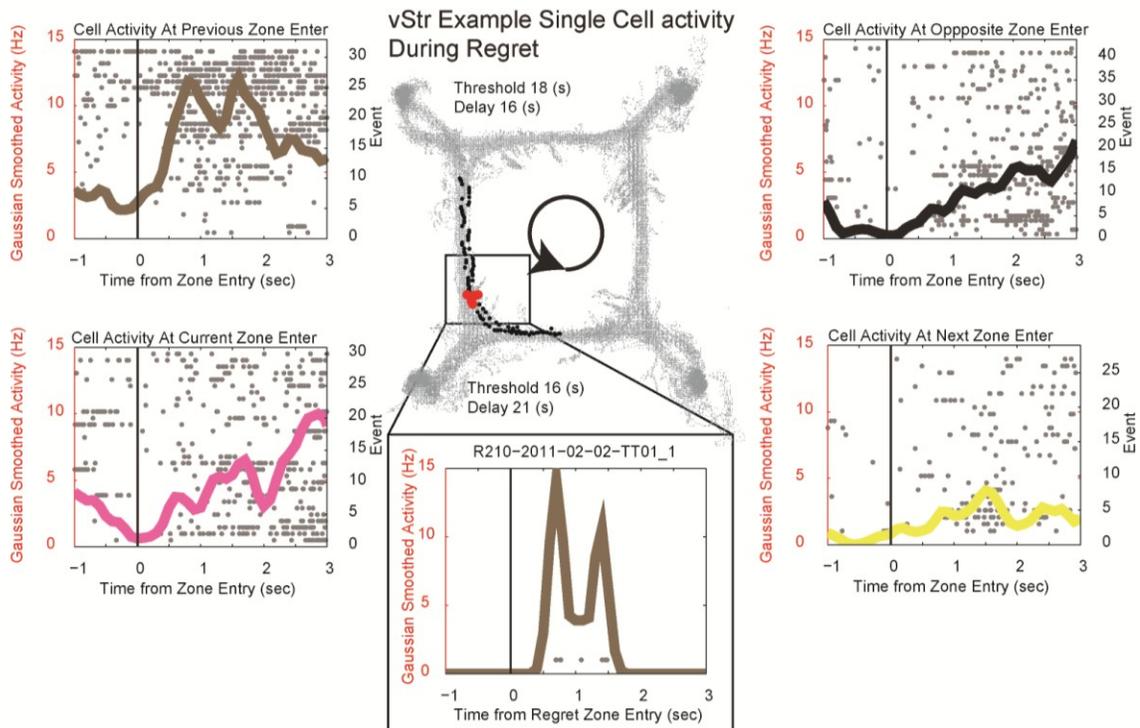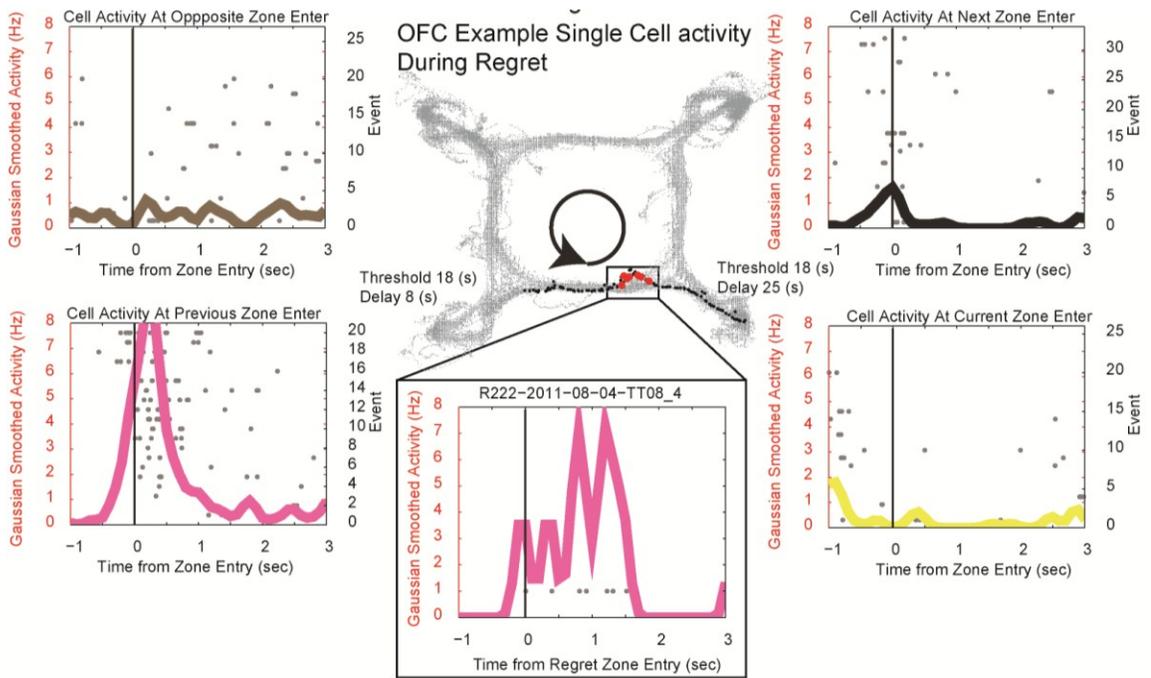
Figure 6-24 **Single reward cells in OFC and vStr during regret-inducing situations.**

Top Panel. OFC Example Cell during regret-inducing situation. Grey dots represent individual spikes. Solid colored lines indicate Gaussian smoothed activity, sigma = 50ms. Black = nonflavored pellets, pink = cherry flavored, yellow = banana flavored, brown = chocolate flavored. Black dots in the center panel represent behavioral samples during this particular instance. Red dots show spikes aligned to behavior. The rat traveled in a counterclockwise direction. The maze has been aligned so that the current zone is represented by the bottom right zone. This particular cell responded most to entry into the cherry reward zone, and little to the banana reward zone. When the rat skipped a low cost cherry zone opportunity and encountered a high cost banana zone opportunity, the rat looked back towards the previous reward; and the activity of the cell approximated that of the cherry-zone-entry response. **Bottom Panel**. Display same as top panel, vStr example cell during a regret-inducing situation, this time from the chocolate-reward zone to the cherry-reward zone. [327]

While our first inclination was to look for representations of the missed reward, human subjects self-report that they regret actions taken or not, more than they do missed outcomes [95,156-158]. We did find a weak representation of the missed reward (not significant, see Figure 6-25).
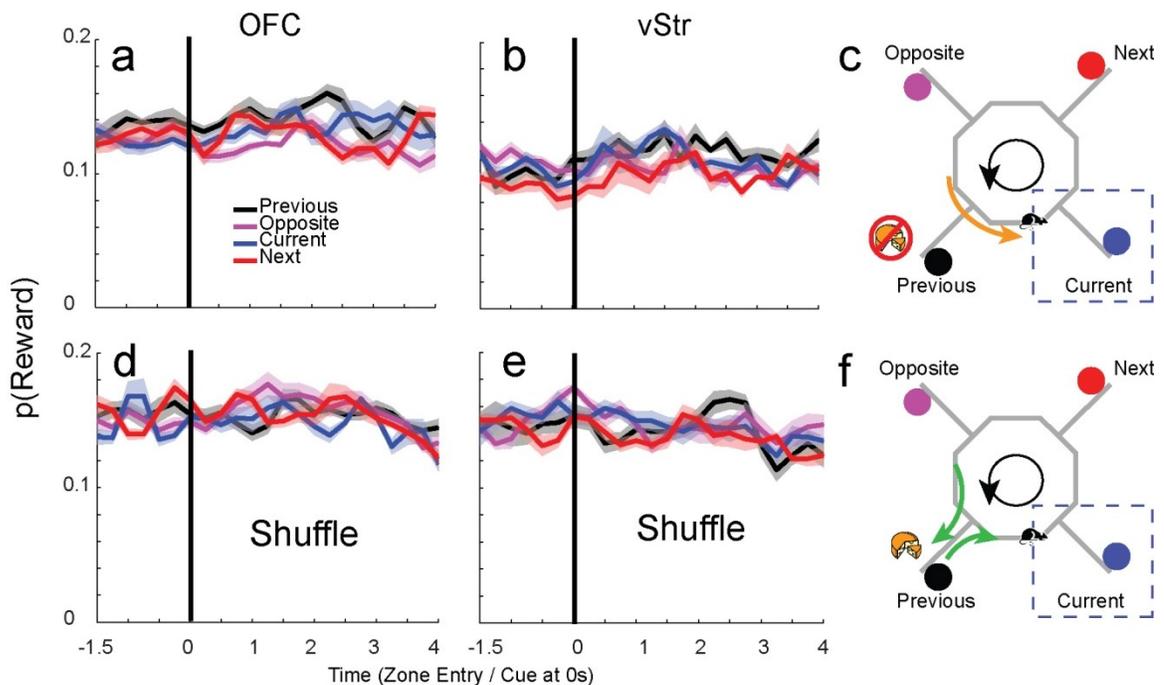
Figure 6-25 **Decoding to previous reward during regret.**
Under normal conditions the current reward is accurately represented. However during regret instances the current reward representations are drastically decreased. Instead neuronal firing rates more accurately represent the missed previous reward. The average decoding for the previous *p(Reward)* was different from

the shuffled data (ANOVA p << 0.001 for VStr and p << 0.001). However, *p(Reward)* for OFC was not significant (ANOVA p = 0.046) after controlling for multiple comparisons. *p(Reward)* for vStr was not significant after controlling for multiple comparisons. [327]

However, we found that there were strong representations of the previous decision-point (*p(Zone)*) that were significantly different than all four other zones (Figure 6-26**a-c**, outside the 95% confidence interval as determined by empirical cumulative distribution function). This differentiation of the previous zone was not observed in either control condition. In the first control (same sequence), both OFC and vStr demonstrated increased representations of the next zone (Figure 6-26**d-f**). By definition, these control instances were high-cost encounters with the current reward site (e.g. above threshold), and, thus the rats were likely to skip them. In the second control condition (two bad offers), the representation of both the current and previous zones increased and were different than the representations of other rewards (Figure 6-26**g-i**). However, this response was markedly different from that seen during potential regret instances as the increase in representation of the previous zone could not be differentiated from the increase in representation of the current zone. Shuffling interspike intervals eliminated all of these effects. Other more positive situations (rejecting a low-cost previous offer and then encountering a low-cost offer, or rejecting a high-cost offer and then encountering a low-cost offer) both led to strong representations of the current zone (Figure 6-27**)**. Other more positive situations (rejecting a low-cost previous offer and then encountering a low-cost offer, or rejecting a high-cost offer and then encountering a low-cost offer) both led

160

to strong representations of the current zone (Figure 6-27). In addition, when rats stayed for an above threshold delay, but then encountered a below threshold delay (which could be described as a potential regret inducing condition) we again found increased representations of the previous zone (Figure 6-28). The representations of the previous zone in this condition (Stay at Delay A > Threshold A to Delay B < Threshold B) were smaller when compared to the previously-described regret-inducing condition.
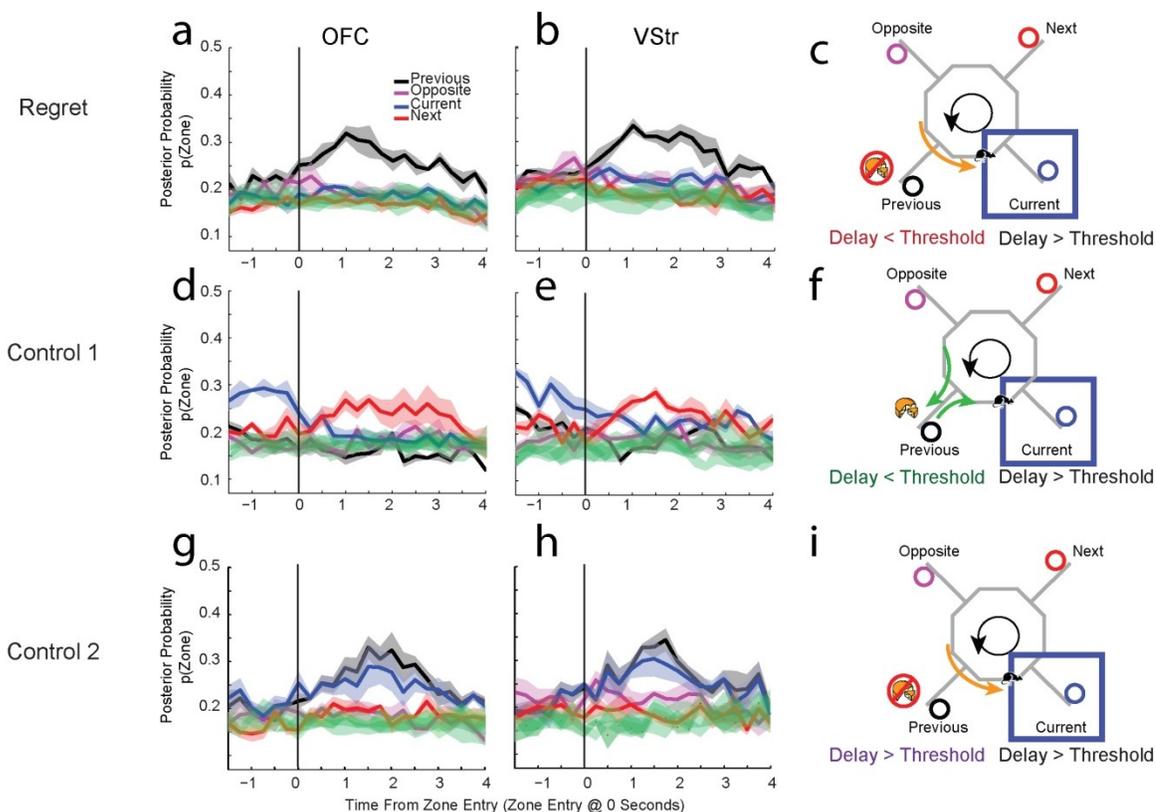
Figure 6-26 **Neural representations in OFC and vStr represent the previous zone during behavioral regret instances.**

In regret-inducing conditions, the *p(Zone)* representation of the previous encounter was high after zone entry into the current zone for both OFC (**a**) and vStr (**b**). Green traces show decoding using shuffled ISIs. Decoding to the previous zone was significantly different from all other conditions, even after controlling for multiple comparisons (ANOVA, OFC: $p \ll 0.001$; vStr: $p \ll 0.001$, distribution significantly different as determined by empirical cumulative distribution function, alpha = 0.05). Panel **c** shows a cartoon of the conditions being decoded – the rat has skipped the previous offer, even though the delay was less than

threshold for that restaurant, and has now encountered a delay greater than threshold for the current restaurant. In the control 1 condition, *p(Zone)* representation of the current zone increased until the rat heard the cue indicating a long delay, at which time, the representation changed to reflect the next zone. In control 1, *p(Zone)* representations to the current and next zones were significantly different from the other zones (ANOVA, vStr: $p \ll 0.001$; OFC: $p \ll 0.001$), although they were not different from each other after controlling for multiple comparisons (ANOVA, vStr: $p = 0.074$, OFC: $p = 0.619$). (**d**: OFC; **e**: vStr; **f**: Cartoon indicating condition.) In the control 2 condition, *p(Zone)* representation of both the current and previous zones was increased when the rat heard the cue indicating a long delay (compared to other zones ANOVA, OFC: $p \ll 0.001$, vStr $p \ll 0.001$). (**g**: OFC; **h**: vStr; **i**: Cartoon indicating condition.) Decoding to the current and previous zones in control 2 were not significantly different from each other (ANOVA, OFC: $p = 0.509$; vStr: $p = 0.268$). [327]
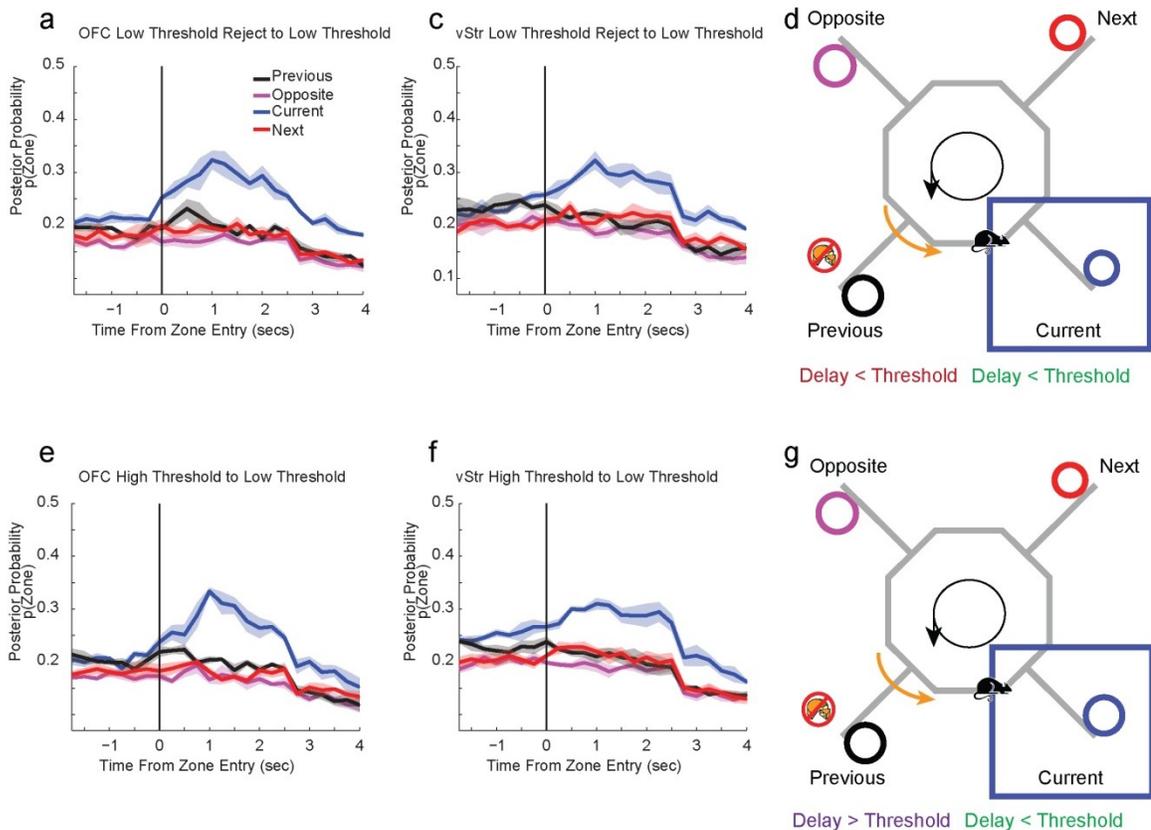
Figure 6-27 **Additional conditions, in which the rat finds a below-threshold opportunity after skipping a previous one.**
a-c If the first reward offer was lower than threshold and rats skipped then encountered a second reward lower than threshold, both OFC and vStr represent the current reward more accurately. This increase occurs immediately after the rat enters into the current, primed zone. This result is consistent with data indicating that OFC represents a given reward when a state paired with that reward has been entered. Prior to entry into the current zone, there is no difference in the representations. d-f When the rats skipped a high-threshold, high cost delay and encountered a low cost delay, both OFC and vStr ensembles accurately represented the current reward. [327]
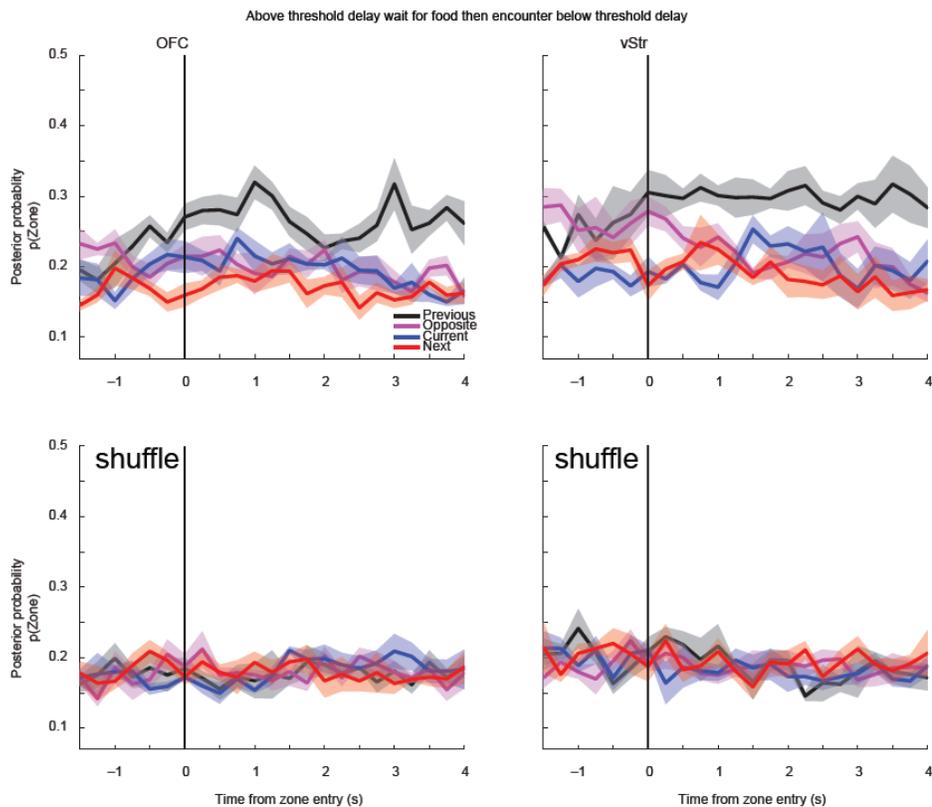
Figure 6-28 **Alternative regret situation.**
Posterior Probability p(Zone) when the rat stayed for a delay @ A > threshold @ A and encounters a delay @ B < threshold B. Both OFC and vStr ensembles increased their decoding to the previous reward. Representations of the previous zone were significant (OFC: ANOVA, p << 0.05; vStr: ANOVA, p << 0.05). However, these representations were not as strong compared to instances when the rat skipped a delay A < threshold @ A and encountered a delay B > threshold B. [327]

Thus, the rats showed different behaviors and different neurophysiological representations during regret-inducing situations, both of which reflected the information processing we would expect to see during regret. As noted above, an important role of regret in decision-making is that it changes subsequent decisions [5,6,367]. Consistent with this hypothesis, we found that rats were more likely to stay at the high cost option in a

163

regret-inducing situation than under either control condition (vs. first control condition, p=0.01; vs. second control condition p=0.06, Wilcoxon, Figure 6-29**a**). In addition, rats spent less time after eating the food before proceeding on to the next reward site following regret-inducing situations compared to non-regret conditions. (Typical handling time mean = 25.3 seconds, regret handling time mean = 15.1 seconds, control handling times are distributed as typical handling times, Figure 6-29**b**). After waiting for food through a long-delay in a regret-inducing situation, rats rushed through eating and quickly went on to their next encounter.
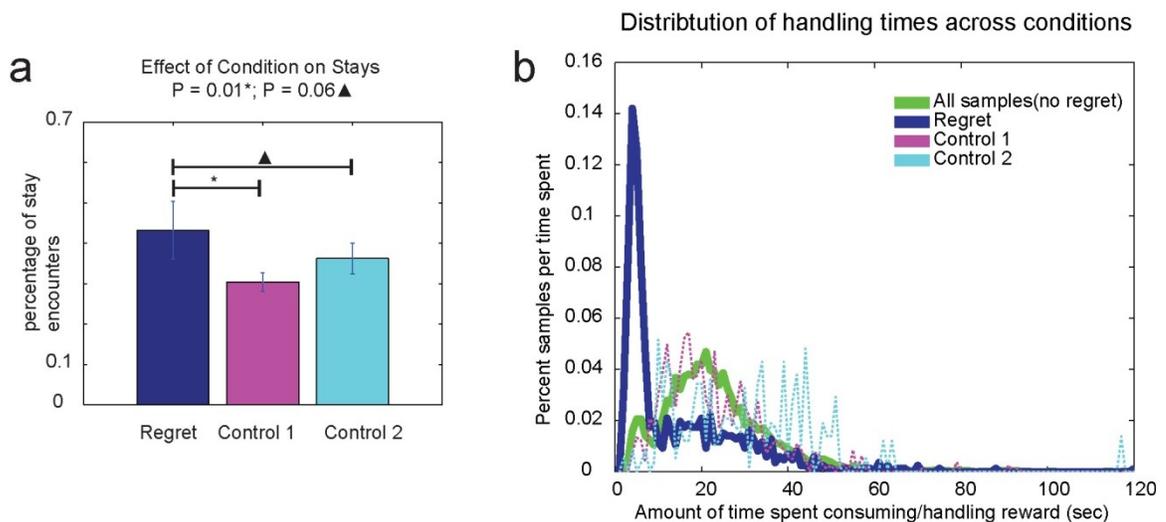
Figure 6-29 **Behavioral changes following potential regret instances.**
a, Comparing the proportion of stays to skips during each condition revealed that rats were significantly more willing to wait for reward following regret inducing instances compared to a control 1 instances (Wilcoxon, p = 0.01) or control 2 instances (Wilcoxon, p = 0.06). b, Rats spent less time consuming reward during regret than during non-regret instances. [327]

The hypothesis that the neural representation of the previous zone reflects some information processing related to regret implies that there should be a relationship between that representation of the previous zone and the animal's subsequent actions.

164

The hypothesis predicts that a stronger representation of the previous zone would lead to an increased likelihood of taking the high-cost (current) offer. To determine if there was a relationship between a rat's willingness to take the high-cost offer and the neurophysiological representations, we compared the ratio of representations of the previous and the current zones and categorized these representations by stay/skip decisions at the current zone. As shown in Figure 6-30, this ratio was increased when the animal decided to stay, but only within the regret-inducing situations. The ratio was unrelated to the decision to stay in the two control conditions. In regret-inducing situations, animals were more willing to stay on trials in which they showed an increased representation of the previous zone relative to the current zone.
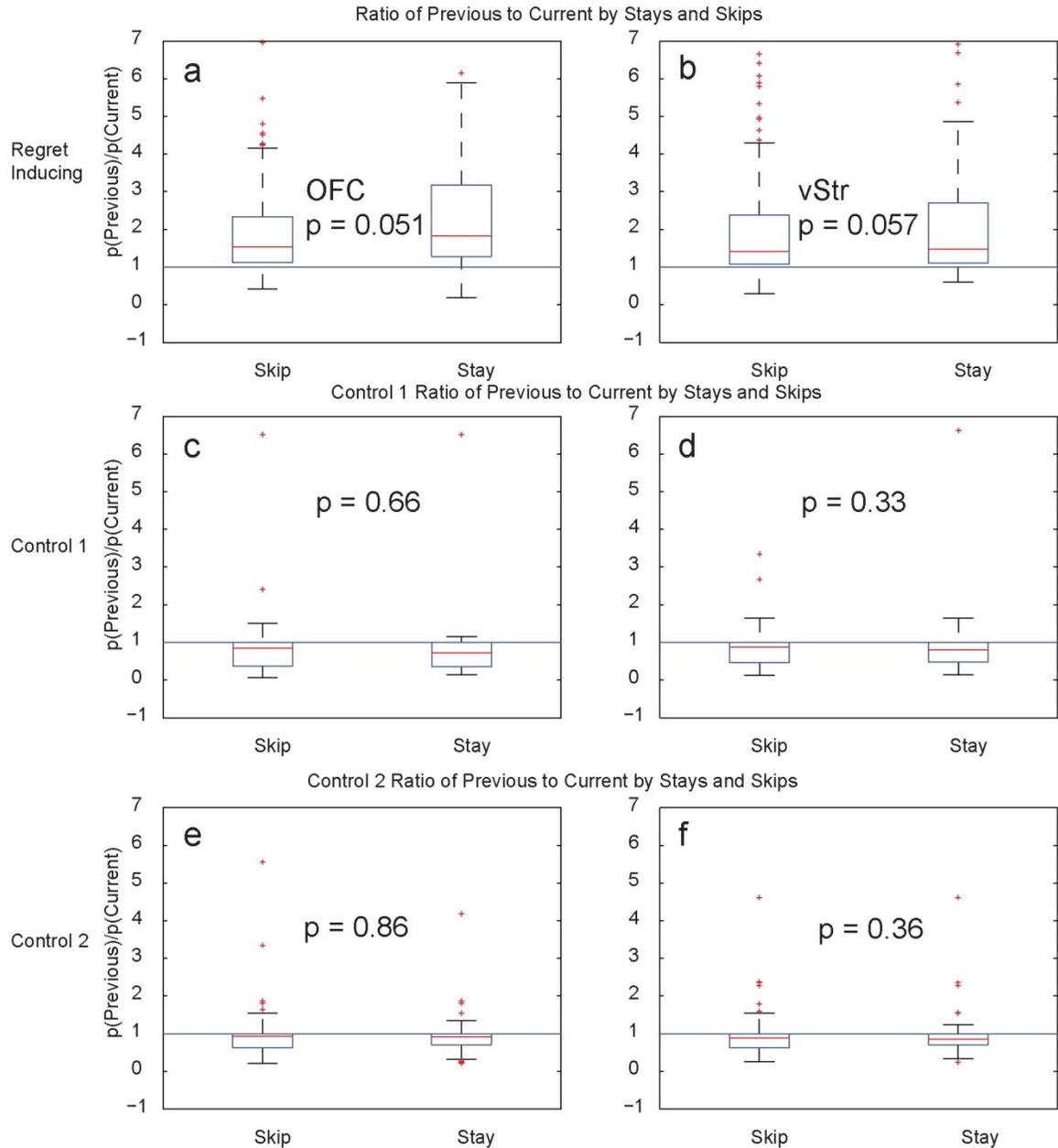
Figure 6-30 **Behavioral and neurophysiological correspondences during regret.**
In order to determine whether the representations of previous reward were different when the rat chose to stay at the high-delay (high-cost) current zone, we measured the ratio between the *p(Zone)* representation of the previous zone against the *p(Zone)* representation of the current zone from 0 to 3 seconds following zone entry for all conditions in the event that the rat skipped or stayed. Each panel shows a box plot of the distribution of $p(Zone_{previous})/p(Zone_{current})$ ratios divided between stays and skips. **a**, $p(Zone_{previous})/p(Zone_{current})$ ratios from OFC ensembles during regret-inducing conditions. **b**,

$p(Zone_{previous})/p(Zone_{current})$ ratios from vStr ensembles during regret-inducing conditions. **d**, **e**, during control 1 conditions. **f**, **g**, during control 2 conditions. Following regret inducing instances, when rats were more willing to wait for reward, *p(Zone_{previous})* was greater than *p(Zone_{current})*. [327]

6.5 Discussion

Regret is the introspective recognition that a previously chosen action led to a less desirable outcome than an alternative action would have. The two keys to identifying regret are value and agency. The Restaurant Row task, in which rats made economic (value-related, cost-dependent) decisions allowed us to identify potentially regret-inducing situations. First, the Restaurant Row task was an economic task, in which rats revealed preferences just as human and non-human primates do [109,150,228]. Second, because the rats had a limited time-budget, encountering a bad (above-threshold) offer after skipping a good (below-threshold) offer, meant that the rat had missed an opportunity. By standard economic and psychological definitions, this sequence should induce regret [23,156,174]. We were able to identify two matched sets of controls that should induce disappointment but not regret: (1) situations in which the rat encountered a similar sequence of offers but took the previous low-cost option, and (2) situations in which the rat encountered two above-threshold offers and skipped the previous high-cost option.

Our data indicate that behavioral and neurophysiological differences between the potential regret-inducing situations and the controls were consistent with a hypothesis that the rats were expressing something akin to human regret. During the regret-inducing

167

situation, rats looked backwards towards the previous (missed) goal, and the orbitofrontal cortex and ventral striatum were more likely to represent that previous goal. After it, rats were more likely to wait out the (current) high-cost offer, and rushed through handling their reward when they did. Interestingly, we found that the neurophysiological representations of counterfactual information in the regret-inducing situation were more strongly related to the missed action (activity when the action was taken, measured by *p(Zone)*) than to the missed outcome (activity when the reward was received, measured by *p(Reward)*). This is consistent with data that humans express more regret about the actions taken (or not taken) than about the missed outcomes [95,156-158].

The Restaurant Row task had three features that made it particularly well suited to the identification of regret. First, it is an economic task on which rats reveal preferences. Second, the inclusion of four "restaurants" allowed us to differentiate a general representation of other rewards from a specific representation of the mistaken choice. We found a clear and significant representation of the previous (lost) zone, not the next or opposite zones. Third, the Restaurant Row task separates the choice of waiting (staying) or going (skipping) from reward-receipt. This separation allowed us to differentiate the regret-induced representation of the previous (lost) reward (a small effect) from the regret-induced representation of the previous (mistaken) action (a large effect). Regret is more about the things you did or did not do than about the rewards you lost [38,55,58,156-158].

Prior evidence indicates that rats can combine information to form an expectation of a novel reward (imagining a particular outcome), and a role for both OFC [33,333,369] and vStr (if a model in the evaluation steps of the task exists) in this process [128,192,193]. Our data indicates that violation of an expectation initiates a retrospective comparison (regretting a missed opportunity). As with the prospective calculation of expectation, this retrospective calculation of expectation influences future behavior – rats are more willing to wait for reward following a regret instance. These two processes, the act of imagining future outcomes and the process of regretting previous, poor choices, are both necessary to modify future decisions to maximize reward.

Previous studies have identified potential representations of the counterfactual could-have-been-chosen option in rats [328], monkeys [1], and humans [56]. In humans, representations of the value of the alternative outcome activate OFC [56,57]. Abe and Lee [1] found that there were representations of an untaken alternate option in monkey OFC on a cued-decision-making task, in which the alternate option that should have been taken was cued to the monkey after the incorrect decision.

The connectivity between OFC and vStr remains highly controversial with some evidence pointing to connectivity [43,45,221] and other analyses suggesting a lack of connectivity [183,291]. The anatomical and functional mechanisms through which the OFC and vStr derive their representations of regret-related counterfactual information remains

unknown. In addition the analyses used here lack the temporal resolution necessary to determine any interactions between structures.

The Restaurant Row task introduced here allowed economic measures to identify potential regret-inducing situations, in which the rat made a decision that placed it in a less valuable situation. Because the task was time limited, any decision to wait for a reward decreased the amount of time available to receive future rewards. Human subjects self-report that they regret actions taken or not, more than they do missed outcomes [95,156-158]. Intriguingly, our decoding results showed strong representations during regret-inducing situations of the previous zone-entry where the decision was made and the action taken *p(Zone)*, but weak and non-significant representations of the missed outcome *p(Reward)*. Most hypotheses suggest that the role of regret is a revaluation of a past opportunity that drives future behavioral changes [23,174]. After making a mistake and recognizing that mistake, rats were more likely to take a high-cost option and rush through the consumption of that less-valuable option.

7. Toward an understanding of OFC function

7.1 OFC and expected outcomes

When deciding between options, most animal species have adapted to predict to a relative degree the outcome of their chosen actions. In part, this is possible because nature obeys a certain set of rules, actions have consequences, and these rules are relatively invariant over time[54,76,77,115,244]. Through trial and error animals learn to use these rules to make predictions based on previous experiences and in some cases, imagined experiences. Over time, animals no longer have to anticipate the state of nature and instead can act on stored value.

In order to represent or abstract new information about expected value a system is needed to emulate the possible outcomes and the paths to those outcomes[69,73,163,193,252,300,304,305,309,333]. OFC has been strongly implicated in the ability to store outcome information and to modify the expected outcome of states in the presence of new information, as well as formulate future approximations given the current available information[1,91,128,192,193,259,261,263,297,304,308,309,321,327,328,333-335,369]. Previous findings can be readily explained by a hypothesis that OFC represents outcome expectations.

171

An intact OFC is necessary for reversal learning in rats, humans, and non-human primates[39,53,59,62,85,91,99,122,127,179,180,189,199,274,301,307,344]. OFC associates cues with new and different rewards[37,40,215,238,239,281,284,295,301,302,305,307,322]. Reversal learning requires a revaluation of learned outcome. Without an intact OFC rats cannot learn the correct identity of a reward predicted by a cue after the cue has been reversed[128,192,193].

In imagined and hypothetical learning, OFC is necessary for the recognition of novel cues that when presented together should result in increased reward[321,333,369]. Without correct OFC function, not only do humans fail to show representation of hypotheticals and counterfactuals, they cannot report the experience of regret[41,56]. In addition, rat OFC represents the counterfactual and the neural correlates in regrets[30,328]. Rats without OFC cannot integrate multiple cues and do not show increased summation of cues for reward. In the formulation of counterfactuals and regret, humans, non-human primates and rats all show representations of the hypothetical outcome in OFC (as shown in Chapter 4 and 6)[1,56-58,327,328].

These results illustrate an important point; OFC is necessary for updating the information regarding the actions and the outcomes in a particular instance. Without OFC, previous actions cannot be modified to account for changes in outcome information and adaptive mechanisms (such as regret) are no longer experienced, limiting the ability of the decision maker to formulate a new expectation of reward.

Recent modeling evidence by Wilson et al suggests that the role of OFC is to represent outcome expectations in specific states and OFC may contribute to a 'cognitive map of space'[369], agreeing with the theory that OFC is primarily involved in outcome expectation. In reinforcement learning models, state representations are necessary to describe the abstract nature of a task[69,71-73,332,368,369]. In the context of reinforcement learning models, OFC would be responsible for formulating the expected outcome of a task and may represent the necessary actions to acquire the expected outcome. This abstract representation of the expected outcome can include space or any other variable that aids in the definition of the current state and outcome identifiers.

Recent findings are consistent with the representation of expected outcome in OFC given the situation, specifically these results support the role of OFC in the representation of both past and future outcomes and their specific properties [1,33,56-58,315,327,328,333-335]. OFC neurons only respond to a particular variable when that variable is instructive and useful in the current situation. Ramus et al showed that 77% of neurons were selective to odors when the odors were relevant to reward[251]. However, when odors were not useful, only 15% of OFC neurons responded to odors. If a cue was not informative for reward, only 25% of the neurons in OFC responded to the cue. However, if the cue provided information about reward, 50% of the neurons responded to the cue[314].

The similar results in rats, non-human primates and humans indicate that OFC serves a similar function in the calculation of expected outcomes, either imagined or retrospective. By representing specific decision outcomes and the variables relevant to those outcomes, OFC would be necessary for a wide variety of behaviors. Accessing the outcome expectation in the current situation would enable a decision maker to successfully predict any outcomes associated with the actions in that situation. If OFC function were to be compromised, the inability to successfully update outcome information in a particular situation could cause perseverance of detrimental behaviors. Errors in outcome representations would make it difficult to select the correct emotional response for a given situation (the experience of regret) or to interpret context in social cues. If one cannot learn to expect a particular outcome for a given action, it is difficult to react to an outcome since there are no preconceptions to which to compare the received outcome.

7.2 OFC and reversal learning

Early evidence suggested that OFC was necessary for preference reversals and reversal learning [19,39,85,120,122,127,189,199,274,338]. In these early experiments, one reward was paired with a unique cue and an alternative reward was paired to a separate, unique cue. Rats were trained to anticipate the rewards associated with each cue[239,308]. Following

learning of the task, OFC was lesioned and cues that predicted the rewards were reversed. Rats were unable to learn that the cues and rewards had been reversed.

How does reversal learning relate to the representation of the counterfactual and the experience of regret? If OFC is necessary for reversal learning, then one possible explanation is that the OFC is maintaining the relevant information regarding the representation of other outcomes. Destruction of the OFC eliminates the ability to successfully represent the alternative outcome. Without the ability to represent the alternative outcome, it is impossible to experience regret. Human subjects with OFC lesions are not capable of experiencing regret (according to measures of self-report and skin conductivity measures, where increased skin conductivity responses would indicate heightened emotional reactivity, wholly absent in OFC lesioned subjects) when confronted with an outcome that is much less desirable than an alternative outcome would have been[41,56-58]. This presents an interesting question; are counterfactuals necessary to successfully learn reversal tasks especially when outcome information has been changed and must be deduced from what was expected, compared to what was received? How closely related are counterfactuals and reversal learning?

Experimental evidence presented in chapter 4 strongly suggests that OFC is heavily involved in the representation of counterfactuals in rats. Upon encountering an unrewarded site, OFC neurons represented the alternative outcome, the new outcome following the reversal in the rewards. As rats approached a decision point that requires

175

representation of a learned outcome, OFC neurons represented some aspect of reward following the decision to proceed to a rewarded location. As the rat approached and received a reward, OFC neurons reliably represented the current outcome.

If a relationship exists between counterfactuals and reversal learning, then one would expect that reversal learning may be closely related to disappointment and regret through the counterfactual. Reversal learning requires that cues be paired with expected outcomes. Under these conditions, if a cue is associated with an outcome, the presentation of multiple cues generates an expectation of reward associated with each, specific cue[31,39,53,75,122,127,143,189,199,274,301,307,308]. The specificities of the cues and outcomes imply that when a rat chooses between the options, there is a specific, expected outcome. If that expected outcome is no longer predicted by the cue, the rat will modify its behavior to account for the change in the expected outcome. One question that has yet to be answered, is that during these reversal tasks when the rats first realizes that the cues have switched, does the rat experience regret, disappointment or both?

Under the framework of regret and disappointment, it may be difficult to distinguish if either is being represented in the OFC during reversal learning. Disappointment requires that the rat encounters the reversed reward, that only one reward was present, that the reward is worse than expected and importantly, that the less than expected reward is not a fault of the rat's choice. However, one key aspect of reversal learning is the presence of multiple rewards, which means that the potential for regret

could exist. This would indicate that if a rat were to encounter the reversed reward the rat may experience regret, namely the representation of the alternative reward, the realization that an incorrect choice was made and as a result, the rat has now received less reward than expected inducing behavioral change on subsequent trials; the definition of regret. On subsequent trials the rat would readily reverse the response to the now correct, reversed cue, learning from the previous mistake. The key component that may describe reversal learning as a special instance of learning from regret is the inclusion of two or more possible, expected outcomes and that the rat's action resulted in the less desirable of the outcomes.

Perhaps the best way to classify reversal learning is as a 'specific form of forced regret where an agent (such as a rat) *unknowingly* selects the wrong choice and encounters a less than desirable outcome compared to the alternative outcome'. Even in the context of uncertainty, human subjects have reported feeling regret about previous choices, even when it was not clear that their choices were necessarily incorrect at the moment of choice[95,156-158].

7.3 Value

Decision making requires that the value of each potential option be evaluated. OFC has been shown to be active during value based decision making. What aspects of value does OFC represent? Activity in OFC is correlated to multiple, positive parameters

of reward, including gustation, olfaction and general sensory aspects of reward[3,29,191,317]. When subjects reached satiation on food or water, the signal in OFC correlating to receipt of the goods decreased[217,318]. Signals in OFC are also modulated by the subjective value of goods and varied with the perceived value of wine as prices were increased or decreased[242]. Individual neurons in non-human primate OFC represent the economic value of the available options [228].

Others have postulated that along with the representation of general value, OFC is responsible for the representation of rewards in a common currency where the aspects of the reward are reduced to a single scalar value that can be compared across multiple goods[139,140,223-225,227,229]. These studies suggest that reward is reduced to a scalar quantity. Once the attributes or the reward are reduced to a scalar quantity, the common currency hypothesis states that all rewards would be compared in this scalar quantity. The common currency hypothesis suggests that OFC integrates multiple modalities of information including probability size, and time to reward[141,209,228,265,302] which contribute to the possible generation of a common currency signal[50,130,131,165,166].

There are several arguments against the coding of value in OFC. The first argument is that there are diverse population responses in OFC neurons. OFC neurons attend to many features of the situations, including cues [170,251,298] and position [81,89,250,261]. If OFC neurons were representing a single scalar value, then different populations of neurons would not respond to the different aspects of reward; rather, their neural

178

responses should only increase as the perceived value of the good increased. In reversal learning, rather than representing value, many neurons fail to reverse their signaling and instead new populations activate [91,296,362]. If OFC neurons were representing value in reversal learning, neurons would respond to the cue associated with reward both before and *after* the reversal. Because most OFC neurons do not respond to the new cue/reward pair following the reversal, they cannot be representing the value of the new cue/reward pair.

A second argument is that OFC is not necessary for initial learning, only for cue/reward pairing changes that occur after the task has already been learned[10,36,37,53,91,122,222,238,239]. Rats without an intact OFC may continue to respond a cue that predicts a reward devalued with illness; however they will not consume the reward[301,307,308]. In addition, OFC in rats and humans is not necessary for choosing between large and small rewards and is more sensitive to delays[82,261]. At least in rats, OFC lesions have only been shown to slow associative recoding of cues[91,295-297,301]; Without an intact OFC, cue/reward recoding is much less flexible while initial cue learning is unimpaired[307,308].

Other data suggests that OFC is necessary to learn from unexpected outcomes. If OFC is lesioned bilaterally, or if OFC and VTA are lesioned contralaterally, rats do not learn about unexpected outcomes[334]. In addition, neural activity in OFC is higher before

a smaller known reward than a delayed, unpredictable reward whereas spiking activity in VTA is larger after unpredictable reward and smaller after predictable[334].

Rather than signaling value, this evidence is consistent with the hypothesis that OFC signals outcome expectation in the current situation to VTA; coupling current responses with the differentiation of outcomes and impending actions. This suggests that OFC is learning outcome representations in the different states[335,369] and that the state representations of the task must be paired with the different outcomes in each state [93,94,254]. Because OFC encodes all aspects of a task [81,251,260-262,265,357,358], OFC can attribute actions to rewards forming the outcome expectation signal[344,365].

When considering the representation of the counterfactual, regret and disappointment, an outcome expectation *must* be present. The counterfactual is the representation of the alternate expected outcome. Disappointment is the recognition that the current outcome is less than the expected outcome. Most importantly, regret is the realization that the received outcome is worse than the expected outcome and that the alternative expected outcome would have been better. Without an expectation of an alternative outcome the counterfactual cannot exist.

7.4 Integrating regret theory, value, and OFC function

What does neural activity in OFC represent and what contribution does OFC provide to the process of decision making? The representations in OFC are necessary to

combine future rewards (both real and imagined) and to generate expectations as well as represent the counterfactual/regret [1,32,41,56-58,128,178,212,213,251,284,308,314,321,327,328,331,333,335,345,361,365]. Recent modeling evidence regarding OFC supports the hypothesis that OFC is responsible for the representation of the expected outcome of the current reward[335]. During initial learning VTA neurons respond to unexpected presentation of reward. Following initial stages of learning, VTA neurons will respond to the cue paired with the reward. Once learned that a specific cue predicts reward, VTA neurons will not modify their responses unless the prediction of the current reward is incorrect, neurons decrease their responses to omitted rewards and cues and increase their responses to increased rewards. OFC neurons respond to rewards as well; however they continue to respond to cues that are predictive of reward as well as the reward. Takahashi et al showed that an intact OFC was necessary for the reward prediction error response in VTA neurons[335].

OFC is more than likely involved in differentiating the different expected outcomes and their associated actions, both taken and imagined, rather than the direct representation of value. Under this framework, OFC may provide necessary information regarding the possible outcomes in the current state, comparing the outcome (including the actions that resulted in the outcome) of the expected outcome to the actual outcome.

In the context of regret, these results likely indicate that OFC acts to simulate the counterfactual, comparing what is to what could have been. The representation of each

possible outcome relies on a correct and functioning OFC as well as the existence of multiple expectations[1,41,42,56-58,327,328]. This process is necessary for the formation of the counterfactual. Without an OFC, rats, non-human primates, and humans would not be capable of representing the alternative outcome during their selection, consequently making it very difficult, if not impossible, to compare the current outcome to the desired outcome. The role of OFC in the formation of expectations and the comparison of expectations both imagined and real, agrees with the neural representation of the counterfactual during regret inducing instances in rats OFC(Chapter 6) as well as the neural representation of the counterfactual in non-human primates OFC[1,56,327].

The role of the OFC in representing expected outcomes is not limited to the outcomes received by the subject. Neural recordings in non-human primates have shown that OFC neuronal responses varied not just with the reward received by the non-human primate but depended on the amount received by a partnered non-human primate[48,49]. Other neuronal recording studies in non-human primates have also shown that OFC neurons respond to reward expectation that includes the perceived social context of the reward; namely that OFC neurons were including the information about the expected outcome (and errors) as the expected outcome was perceived by a non-human primate partner[7,371]. In humans, the size of the OFC has been correlated to the increased ability to understand social interactions and to increased interpersonal social network size[168]. This may indicate that in humans and non-human primates OFC also plays a role in the ability

to imagine not just personal counterfactuals, but also the counterfactuals of other's actions, as well as how personal actions and their counterfactuals can affect others

If an action (either performed by the subject or observed by the subject) is closely related to acquiring the desired outcome, then one can expect that the action leading to the outcome would be represented in OFC neuronal activity. Indeed this is the case, at least in rats, where the representation of the alternative action was much stronger during regret inducing situations (Chapter 6). The neural evidence agrees with prior psychological reports of regret, that humans express stronger regret for actions they should have taken rather than actions they have taken[95,156-158].

7.5 Conclusion

The evidence shown in this thesis represents the first report that rats have distinct economic preferences and that rats can experience regret. These regretful instances modify their future behavior, much like regret does in humans[96,114,355,375,377-379]. However, the economic definition of regret is incomplete and neglects several of the advances in learning from temporal difference reinforcement learning; the action and the outcome both contribute to the process of the decision [332,368,369]. Regret as it is conceptualized in economics currently ignores the actions that precede the value of the outcome. This disagrees with evidence in psychology that states human subjects regret the action that leads to the outcome/value more than the actual outcome. By assuming the action has no

value and does not contribute to the decision, economic definitions of regret are

incomplete.

If we included the action and its related value into the calculation of regret as

defined by economics, we found that the strongest representation of regret is tied to the

previously unchosen action. By including the action with the outcome in the decision,

psychological descriptions of regret agree with economic descriptions of regret (in rats

and humans). Human subjects report that they regret the action more than the outcome.

Economic descriptions of regret (regret aversion) show that actions with high regret are

avoided, implying that the actions when coupled with regret have a negative

value[23,174,355,375,379].

One key differentiation was made, disappointment and regret were distinguishable

in the neural representations of the counterfactuals. Regret was the comparison of

multiple outcomes, 'I wanted that, not this!', while during disappointment a single

outcome comparison occurred 'I wanted this and I didn't get it'. When representing

multiple outcomes, the counterfactual must have been calculated. These differences were

distinguishable in the neural representations of reward and counterfactual information

(Chapter 6) and agreed with the psychological and economic definitions of regret and

disappointment.

Consequently, can rats experience regret? According to the way regret has been

defined previously in economics[23,174], I have shown in the thesis that rats can represent

the counterfactual during regret inducing situations and that during regret, OFC ensembles represent the previously, would-have-been received reward.

In addition, this research has indicated that previous regretful experiences can modify subsequent behavior. These behavioral changes are accompanied by the representation of alternative, would have been received, reward information in the orbitofrontal cortex. However, there is no direct neural evidence that investigates how regret affects the rationality of choices or guides future behavior towards optimal economic solutions. Are there differences in the representation of regret? Do humans, rats or non-human primates always regret the action that led to the less desirable outcome more than the horrible outcome? Are there representational differences in OFC between bad decision outcomes and the process of deciding badly, the process of evaluating the potential options and acting on those decisions, both of which lead to the experience of regret[240,376,377]?

If regret plays an important role in decision processes and can explain irrational behavior under specific conditions then the evaluation of regret may be able to explain why some choices seem irrational. The orbitofrontal cortex is likely necessary for the representation of regret as well as for behavioral modifications that rely on the representation of regret. The inclusion of regret into the economics and psychology of decision making may explain perceived irrational behaviors.

This work has shown that humans and non-human primates are not the only species that rely on regret to inform future decisions or modify subsequent actions. These findings are important because they aid in our understanding how decisions are made and how the mechanism works under normal circumstances. We must first understand how decisions are made before we can understand how the system fails in case of drug addicts or traumatic brain injury.

8. References

1   Abe, H. & Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**, 731-741, (2011).

2   Ahmed, S. H. Validation crisis in animal models of drug addiction: beyond non-disordered drug use toward drug addiction. *Neuroscience and Biobheavioral Reviews* **35**, 172-184, (2010).

3   Anderson, A. K. *et al.* Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience* **6**, 196-202, (2003).

4   Arana, F. S. *et al.* Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J Neurosci* **23**, 9632-9638, (2003).

5   Arkes, H. R. & Ayton, P. The sunk cost and Concorde effects: are humans less rational than lower animals? *Psychological Bulletin* **125**, 591, (1999).

6   Aronson, E. The effect of effort on the attractiveness of rewarded and unrewarded stimuli. *J Abnorm Soc Psychol* **63**, 375-380, (1961).

7   Azzi, J. C. B., Sirigu, A. & Duhamel, J.-R. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proceedings of the National Academy of Sciences* **109**, 2126-2131, (2012).

8   Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407-419, (1998).

9   Baxter, M. G. & Murray, E. A. The amygdala and reward. *Nature reviews. Neuroscience* **3**, 563-573, (2002).

10   Baxter, M. G., Parker, A., Lindner, C. C., Izquierdo, A. D. & Murray, E. A. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *The Journal of Neuroscience* **20**, 4311-4319, (2000).

11   Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129-141, (2005).

12   Bechara, A. & Damasio, A. The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior* **52**, 336-372, (2005).

13   Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**, 7-15, (1994).

14   Bechara, A. & Damasio, H. Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia* **40**, 1675-1689, (2002).

15   Bechara, A., Damasio, H. & Damasio, A. Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex* **10**, 295-307, (2000).

16   Bechara, A., Damasio, H., Damasio, A. R. & Lee, G. P. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J Neurosci* **19**, 5473-5481, (1999).

17   Bechara, A., Damasio, H., Tranel, D. & Anderson, S. Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience* **18**, 428-437, (1998).

18   Bechara, A., Damasio, H., Tranel, D. & Damasio, A. The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *TRENDS in Cognitive Sciences* **9**, 159-162, (2005).

19   Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. Deciding advantageously before knowing the advantageous strategy. *Science* **275**, 1293, (1997).

20   Bedford, E. in *Proceedings of the Aristotelian Society.* 281-304 (JSTOR).

21   Behrens, T. E., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. Associative learning of social value. *Nature* **456**, 245-249, (2008).

22   Bell, D. Disappointment in decision making under uncertainty. *Operations Research* **33**, 1-27, (1985).

23   Bell, D. Regret in decision making under uncertainty. *Operations Research* **30**, 961-981, (1982).

24   Bell, D. E. Reply—Putting a Premium on Regret. *Management Science* **31**, 117-122, (1985).

25   Berlin, H. A., Rolls, E. T. & Kischka, U. Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain : a journal of neurology* **127**, 1108-1126, (2004).

26   Berridge, K. C. 'Liking' and 'wanting' food rewards: brain substrates and roles in eating disorders. *Physiology & behavior* **97**, 537-550, (2009).

27   Bissonette, G. B. *et al*. Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. *J Neurosci* **28**, 11124-11130, (2008).

28   Blaisdell, A. P., Sawa, K., Leising, K. J. & Waldmann, M. R. Causal reasoning in rats. *Science* **311**, 1020-1022, (2006).

29   Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences* **98**, 11818-11823, (2001).

30   Blumenthal, A., Steiner, A., Seeland, K. & Redish, A. D. Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiol Learn Mem* **95**, 376-384, (2011).

31   Bohn, I., Giertler, C. & Hauber, W. Orbital prefrontal cortex and guidance of instrumental behaviour in rats under reversal conditions. *Behav Brain Res* **143**, 49-56, (2003).

32   Boorman, E. D., Behrens, T. E. & Rushworth, M. F. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS biology* **9**, e1001093, (2011).

33   Bray, S., Shimojo, S. & O'Doherty, J. P. Human medial orbitofrontal cortex is recruited during experience of imagined and real rewards. *J Neurophysiol* **103**, 2506-2512, (2010).

34   Brendl, C. M. & Higgins, E. T. Principles of judging valence: What makes events positive or negative? *Advances in experimental social psychology* **28**, 95-160, (1996).

35   Buck, M. L. & Miller, D. T. Reactions to incongruous negative life events. *Social Justice Research* **7**, 29-46, (1994).

36   Burke, K. A., Franz, T. M., Miller, D. N. & Schoenbaum, G. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature* **454**, 340-344, (2008).

37   Burke, K. A., Takahashi, Y. K., Correll, J., Leon Brown, P. & Schoenbaum, G. Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with 'disinhibition' of responding for previously unrewarded cues. *The European journal of neuroscience* **30**, 1941-1946, (2009).

38   Burks, A. W. Laws of nature and reasonableness of regret. *Mind* **55**, 170-172, (1946).

39   Butter, C. M. Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in< i> Macaca mulatta</i>. *Physiology & behavior* **4**, 163-171, (1969).

40   Calu, D. J. *et al.* Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. *Learn Mem* **14**, 325-328, (2007).

41   Camille, N. *et al.* The involvement of the orbitofrontal cortex in the experience of regret. *Science* **304**, 1167-1170, (2004).

42   Camille, N., Griffiths, C. A., Vo, K., Fellows, L. K. & Kable, J. W. Ventromedial frontal lobe damage disrupts value maximization in humans. *The Journal of Neuroscience* **31**, 7527-7532, (2011).

43   Carmichael, S. & Price, J. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology* **371**, 179-207, (1996).

44   Carmichael, S. T. & Price, J. L. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *The Journal of comparative neurology* **346**, 366-402, (1994).

45   Carmichael, S. T. & Price, J. L. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *The Journal of comparative neurology* **363**, 615-641, (1995).

46   Carmichael, S. T. & Price, J. L. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *The Journal of comparative neurology* **363**, 642-664, (1995).

47   Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R. J. & Reinoso-Suarez, F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb Cortex* **10**, 220-242, (2000).

48   Chang, S. W. *et al.* Neuroethology of primate social behavior. *Proceedings of the National Academy of Sciences* **110**, 10387-10394, (2013).

49   Chang, S. W. C., Gariepy, J.-F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci* **16**, 243-250, (2013).

50   Chib, V. S., Rangel, A., Shimojo, S. & O'Doherty, J. P. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci* **29**, 12315-12320, (2009).

51   Chisholm, R. M. I.—THE CONTRARY-TO-FACT CONDITIONAL1. *Mind* **55**, 289-307, (1946).

52   Chudasama, Y., Izquierdo, A. & Murray, E. A. Distinct contributions of the amygdala and hippocampus to fear expression. *The European journal of neuroscience* **30**, 2327-2337, (2009).

53   Chudasama, Y. & Robbins, T. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *Journal of Neuroscience* **23**, 8771-8780, (2003).

54   Clayton, N. S., Bussey, T. J. & Dickinson, A. Can animals recall the past and plan for the future? *Nature Reviews Neuroscience* **4**, 685-691, (2003).

55   Connolly, T. & Butler, D. Regret in economic and psychological theories of choice. *Journal of Behavioral Decision Making* **19**, 139-154, (2006).

56   Coricelli, G. *et al.* Regret and its avoidance: a neuroimaging study of choice behavior. *Nat Neurosci* **8**, 1255-1262, (2005).

57   Coricelli, G., Dolan, R. J. & Sirigu, A. Brain, emotion and decision making: the paradigmatic example of regret. *Trends Cogn Sci* **11**, 258-265, (2007).

58   Coricelli, G. & Rustichini, A. Counterfactual thinking and emotions: regret and envy learning. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365**, 241-247, (2009).

59   Critchley, H. D. & Rolls, E. T. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J Neurophysiol* **75**, 1673-1686, (1996).

60   Cromwell, H. C. & Schultz, W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* **89**, 2823-2838, (2003).

61   Croxson, P. L. *et al.* Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *The Journal of Neuroscience* **25**, 8854-8866, (2005).

62   Damasio, A. Descartes'error: Emotion, reason, and the human brain. *NewYork, Putnam's Sons*, (1994).

63   Damasio, A., Everitt, B. & Bishop, D. The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex [and Discussion]. *Philosophical Transactions: Biological Sciences*, 1413-1420, (1996).

64   Damasio, A. R. The brain binds entities and events by multiregional activation from convergence zones. *Neural computation* **1**, 123-132, (1989).

65   Damasio, A. R. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition* **33**, 25-62, (1989).

66   Damasio, A. R., Tranel, D. & Damasio, H. Somatic markers and the guidance of behavior: Theory and preliminary testing. *Frontal lobe function and dysfunction*, 217-229, (1991).

67   Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M. & Damasio, A. R. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* **264**, 1102-1105, (1994).

68   Davis, M. The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience* **15**, 353-375, (1992).

69   Daw, N. D., & Doya, K. The computational neurobiology of learning and reward. *Current opinion in neurobiology* **16**, 199-204, (2006).

70   Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**, 1704-1711, (2005).

71   Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879, (2006).

72   Dayan, P. Simple substrates for complex cognition. *Front Neurosci* **2**, 255-263, (2008).

73   Dayan, P. & Daw, N. D. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience* **8**, 429-453, (2008).

74   de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A. & Fletcher, P. C. Differential engagement of the ventromedial prefrontal cortex by goal-directed

and habitual behavior toward food pictures in humans. *J Neurosci* **29**, 11330-11338, (2009).

75   Dias, R., Robbins, T. W. & Roberts, A. C. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* **380**, 69-72, (1996).

76   Dickinson, A. Associative learning and animal cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 2733-2742, (2012).

77   Dickinson, A. Causal learning: An associative analysis. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, (2001).

78   Eichenbaum, H., Schoenbaum, G., Young, B. & Bunsey, M. Functional organization of the hippocampal memory system. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13500-13507, (1996).

79   Elliott, R., Agnew, Z. & Deakin, J. F. Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. *The European journal of neuroscience* **27**, 2213-2218, (2008).

80   Everitt, B. J. & Robbins, T. W. Amygdala-ventral striatal interactions and reward-related processes. (1992).

81   Feierstein, C., Quirk, M., Uchida, N., Sosulski, D. & Mainen, Z. Representation of spatial goals in rat orbitofrontal cortex. *Neuron* **51**, 495-508, (2006).

82   Fellows, L. Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Annals of the New York Academy of Sciences* **1239**, 51-59, (2011).

83   Fellows, L. & Farah, M. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral cortex* **15**, 58-63, (2005).

84   Fellows, L. K. & Farah, M. J. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral cortex* **17**, 2669-2674, (2007).

85   Fellows, L. K. & Farah, M. J. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain : a journal of neurology* **126**, 1830-1837, (2003).

86   Ferry, A. T., Ongur, D., An, X. & Price, J. L. Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. *The Journal of comparative neurology* **425**, 447-470, (2000).

87   Foote, A. L. & Crystal, J. D. Metacognition in the rat. *Curr Biol* **17**, 551-555, (2007).

88   Frost, R. *Mountain Interval: The road not taken*. (Holt, 1916).

89   Furuyashiki, T., Holland, P. & Gallagher, M. Rat Orbitofrontal Cortex Separately Encodes Response and Outcome Information during Performance of Goal-Directed Behavior. *Journal of Neuroscience* **28**, 5127, (2008).

90   Gaffan, D. & Murray, E. A. Amygdalar interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. *J Neurosci* **10**, 3479-3493, (1990).

91   Gallagher, M., McMahan, R. W. & Schoenbaum, G. Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci* **19**, 6610-6614, (1999).

92   George, D. N., Duffaud, A. M., Pothuizen, H. H. J., Haddon, J. E. & Killcross, S. Lesions to the ventral, but not the dorsal, medial prefrontal cortex enhance latent inhibition. *European Journal of Neuroscience* **31**, 1474-1482, (2010).

93   Gershman, S. & Niv, Y. Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, (2010).

94   Gershman, S. J., Blei, D. M. & Niv, Y. Context, learning, and extinction. *Psychological review* **117**, 197-209, (2010).

95   Gilovich, T. & Medvec, V. H. The experience of regret: what, when, and why. *Psychological review* **102**, 379, (1995).

96   Giorgetta, C. *et al.* Waves of regret: A meg study of emotion and decision-making. *Neuropsychologia* **51**, 38-51, (2013).

97   Glimcher, P. W. *Neuroeconomics: decision making and the brain*. (Academic Pr, 2008).

98   Goodman, N. *Fact, fiction, and forecast*. (Harvard University Press, 1983).

99   Gottfried, J. A., O'Doherty, J. & Dolan, R. J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**, 1104-1107, (2003).

100   Groenewegen, H., der Zee, E., Te Kortschot, A. & Witter, M. Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of< i> Phaseolus vulgaris</i> leucoagglutinin. *Neuroscience* **23**, 103-120, (1987).

101   Groenewegen, H. J. Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *Neuroscience* **24**, 379-431, (1988).

102   Grout, P. On minimax regret and welfare economics. *Journal of Public Economics* **9**, 405-410, (1978).

103   Gupta, A. S., van der Meer, M. A., Touretzky, D. S. & Redish, A. D. Hippocampal replay is not a simple function of experience. *Neuron* **65**, 695-705, (2010).

104   Haber, S., Kunishio, K., Mizobuchi, M. & Lynd-Balta, E. The orbital and medial prefrontal circuit through the primate basal ganglia. *Journal of Neuroscience* **15**, 4851-4867, (1995).

105   Hampshire, S. *Morality and conflict*. Vol. 4 (Harvard University Press Cambridge, MA, 1983).

106   Hampshire, S. *Thought and action*. (Chatto and Windus London, 1959).

107   Hampton, A. N., Bossaerts, P. & O'Doherty, J. P. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* **26**, 8360-8367, (2006).

108   Hare, T. A., Camerer, C. F., Knoepfle, D. T. & Rangel, A. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J Neurosci* **30**, 583-590, (2010).

109   Hare, T. A., Camerer, C. F. & Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* **324**, 646-648, (2009).

110   Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W. & Rangel, A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* **28**, 5623-5630, (2008).

111   Hart, S. Adaptive heuristics. *Econometrica* **73**, 1401-1430, (2005).

112   Hart, S. & Mas-Colell, A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* **68**, 1127-1150, (2000).

113   Hart, W. E., Goldbaum, M., Cote, B., Kube, P. & Nelson, M. R. Measurement and classification of retinal vascular tortuosity. *Int J Med Inform* **53**, 239-252, (1999).

114   Hetts, J. J., Boninger, D. S., Armor, D. A., Gleicher, F. & Nathanson, A. The influence of anticipated counterfactual regret on behavior. *Psychology & Marketing* **17**, 345-368, (2000).

115   Heyes, C. & Dickinson, A. The intentionality of animal action. *Mind & Language* **5**, 87-103, (1990).

116   Hikosaka, K. & Watanabe, M. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cerebral cortex* **10**, 263-271, (2000).

117   Hikosaka, K. & Watanabe, M. Long- and short-range reward expectancy in the primate orbitofrontal cortex. *The European journal of neuroscience* **19**, 1046-1054, (2004).

118   Holland, P. C. & Gallagher, M. Amygdala circuitry in attentional and representational processes. *Trends Cogn Sci* **3**, 65-73, (1999).

119   Hornak, J. *et al.* Reward-related Reversal Learning After Surgical Excisions in Orbito-frontal or Dorsolateral Prefrontal Cortex in Humans. *Journal of cognitive neuroscience* **16**, 463-478, (2004).

120   Hornak, J. *et al.* Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of cognitive neuroscience* **16**, 463-478, (2004).

121   Ito, M. & Doya, K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J Neurosci* **29**, 9861-9874, (2009).

122   Izquierdo, A., Suda, R. K. & Murray, E. A. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci* **24**, 7540-7548, (2004).

123   Izquierdo, A., Suda, R. K. & Murray, E. A. Comparison of the effects of bilateral orbital prefrontal cortex lesions and amygdala lesions on emotional responses in rhesus monkeys. *J Neurosci* **25**, 8534-8542, (2005).

124   Janabi-Sharifi, F., Hayward, V. & Chen, C. S. J. Discrete-time adaptive windowing for velocity estimation. *IEEE Transactions on Control Systems Technology* **8**, 1003-1009, (2000).

125   Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci* **27**, 12176-12189, (2007).

126   Johnson, A., van der Meer, M. A. & Redish, A. D. Integrating hippocampus and striatum in decision-making. *Current opinion in neurobiology* **17**, 692-697, (2007).

127   Jones, B. & Mishkin, M. Limbic lesions and the problem of stimulus--reinforcement associations. *Experimental neurology* **36**, 362-377, (1972).

128   Jones, J. L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* **338**, 953-956, (2012).

129   Jungermann, H. The two camps on rationality. *Decision making under uncertainty*, 63-86, (1983).

130   Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* **10**, 1625-1633, (2007).

131   Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733-745, (2009).

132   Kahneman, D. & Miller, D. T. Norm theory: Comparing reality to its alternatives. *Psychological review* **93**, 136, (1986).

133   Kahneman, D. & Tversky, A. Choices, values, and frames. *American Psychologist* **39**, 341, (1984).

134   Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47**, 263-291, (1979).

135   Kahneman, D. & Tversky, A. The psychology of preferences. *Scientific American*, (1982).

136   Kahneman, D. & Varey, C. A. Propensities and counterfactuals: The loser that almost won. *Journal of personality and social psychology* **59**, 1101, (1990).

137   Kahnt, T., Chang, L. J., Park, S. Q., Heinzle, J. & Haynes, J. D. Connectivity-based parcellation of the human orbitofrontal cortex. *J Neurosci* **32**, 6240-6250, (2012).

138   Kazama, A. & Bachevalier, J. Selective aspiration or neurotoxic lesions of orbital frontal areas 11 and 13 spared monkeys' performance on the object discrimination reversal task. *The Journal of Neuroscience* **29**, 2794-2804, (2009).

139   Kennerley, S. & Wallis, J. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *European Journal of Neuroscience* **29**, 2061-2073, (2009).

140   Kennerley, S. W., Dahmubed, A. F., Lara, A. H. & Wallis, J. D. Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of cognitive neuroscience* **21**, 1162-1178, (2009).

141   Kennerley, S. W. & Wallis, J. D. Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. *J Neurophysiol* **102**, 3352-3364, (2009).

142   Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227-231, (2008).

143   Kim, J. & Ragozzino, M. E. The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiol Learn Mem* **83**, 125-133, (2005).

144   Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J. & Behrens, T. E. Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J Neurosci* **33**, 3202-3211, (2013).

145   Klüver, H. & Bucy, P. C. Preliminary analysis of functions of the temporal lobes in monkeys. *Archives of Neurology & Psychiatry* **42**, 979-1000, (1939).

146   Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M. & Hommer, D. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage* **18**, 263-272, (2003).

147   Kondo, H., Saleem, K. S. & Price, J. L. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *The Journal of comparative neurology* **493**, 479-509, (2005).

148   Kondo, H., Saleem, K. S. & Price, J. L. Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *The Journal of comparative neurology* **465**, 499-523, (2003).

149   Kondo, H. & Witter, M. P. Topographic organization of orbitofrontal projections to the parahippocampal region in rats. *The Journal of comparative neurology* **522**, 772-793, (2014).

150   Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci* **13**, 1292-1298, (2010).

151   Krettek, J. & Price, J. Amygdaloid projections to subcortical structures within the basal forebrain and brainstem in the rat and cat. *Journal of Comparative Neurology* **178**, 225-253, (1978).

152   Krettek, J. E. & Price, J. L. The cortical projections of the mediodorsal nucleus and adjacent thalamic nuclei in the rat. *The Journal of comparative neurology* **171**, 157-191, (1977).

153 Krettek, J. E. & Price, J. L. Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *The Journal of comparative neurology* **172**, 687-722, (1977).

154 Kringelbach, M., O'Doherty, J., Rolls, E. & Andrews, C. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral cortex* **13**, 1064-1071, (2003).

155 Kripke, S. A. *Naming and necessity*. (Springer, 1980).

156 Landman, J. Regret: A theoretical and conceptual analysis. *Journal for the Theory of Social Behaviour* **17**, 135-160, (1987).

157 Landman, J. *Regret: The persistence of the possible*. (Oxford University Press, 1993).

158 Landman, J. & Manis, J. D. What might have been: Counterfactual thought concerning personal decisions. *British Journal of Psychology* **83**, 473-477, (1992).

159 Lavoie, A. M. & Mizumori, S. J. Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats. *Brain research* **638**, 157-168, (1994).

160 LeDoux, J. Rethinking the emotional brain. *Neuron* **73**, 653-676, (2012).

161 LeDoux, J. E. *The Emotional Brain*. (Simon and Schuster, 1996).

162 Lee, D. Neural basis of quasi-rational decision making. *Current opinion in neurobiology* **16**, 191-198, (2006).

163 Lee, D., Seo, H. & Jung, M. W. Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci* **35**, 287-308, (2012).

164 Leonard, C. M. The prefrontal cortex of the rat. I. Cortical projection of the mediodorsal nucleus. II. Efferent connections. *Brain research* **12**, 321-343, (1969).

165 Levy, D. J. & Glimcher, P. W. The root of all value: a neural common currency for choice. *Current opinion in neurobiology*, (2012).

166 Levy, I., Snell, J., Nelson, A., Rustichini, A. & Glimcher, P. Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology* **103**, 1036-1083, (2010).

167 Lewis, D. Causation. *The journal of philosophy*, 556-567, (1973).

168 Lewis, P. A., Rezaie, R., Brown, R., Roberts, N. & Dunbar, R. I. Ventromedial prefrontal volume predicts understanding of others and social network size. *NeuroImage* **57**, 1624-1629, (2011).

169 Lichtenstein, S. & Slovic, P. *The construction of preference*. (Cambridge University Press, 2006).

170 Lipton, P. A., Alvarez, P. & Eichenbaum, H. Crossmodal associative memory representations in rodent orbitofrontal cortex. *Neuron* **22**, 349-359, (1999).

171 Loomes, G. Further evidence of the impact of regret and disappointment in choice under uncertainty. *Economica*, 47-62, (1988).

172 Loomes, G., Starmer, C. & Sugden, R. Are preferences monotonic? Testing some predictions of regret theory. *Economica*, 17-33, (1992).

173 Loomes, G. & Sugden, R. Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies* **53**, 271-282, (1986).

174 Loomes, G. & Sugden, R. Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal* **92**, 805-824, (1982).

175 Loomes, G. & Sugden, R. Testing for regret and disappointment in choice under uncertainty. *Economic Journal* **97**, 118-129, (1987).

176 Lucantonio, F., Stalnaker, T. A., Shaham, Y., Niv, Y. & Schoenbaum, G. The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci* **15**, 358-366, (2012).

177 Luhmann, C. C., Chun, M. M., Yi, D. J., Lee, D. & Wang, X. J. Neural dissociation of delay and uncertainty in intertemporal choice. *J Neurosci* **28**, 14459-14466, (2008).

178 Luk, C.-H. & Wallis, J. D. Choice Coding in Frontal Cortex during Stimulus-Guided or Action-Guided Decision-Making. *The Journal of Neuroscience* **33**, 1864-1871, (2013).

179 Machado, C. J. & Bachevalier, J. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *The European journal of neuroscience* **25**, 2885-2904, (2007).

180 Machado, C. J. & Bachevalier, J. Measuring reward assessment in a semi-naturalistic context: the effects of selective amygdala, orbital frontal or hippocampal lesions. *Neuroscience* **148**, 599-611, (2007).

181 Mackey, S. & Petrides, M. Architectonic mapping of the medial region of the human orbitofrontal cortex by density profiles. *Neuroscience* **159**, 1089-1107, (2009).

182 Maia, T. V. & McClelland, J. L. A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16075-16080, (2004).

183 Mailly, P., Aliane, V., Groenewegen, H. J., Haber, S. N. & Deniau, J. M. The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units. *J Neurosci* **33**, 5718-5727, (2013).

184 Mainen, Z. F. & Kepecs, A. Neural representation of behavioral outcomes in the orbitofrontal cortex. *Current opinion in neurobiology* **19**, 84-91, (2009).

185 Mar, A. C., Walker, A. L., Theobald, D. E., Eagle, D. M. & Robbins, T. W. Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J Neurosci* **31**, 6398-6404, (2011).

186 Mariano, T. Y. *et al.* Impulsive choice in hippocampal but not orbitofrontal cortex-lesioned rats on a nonspatial decision-making maze task. *European Journal of Neuroscience* **30**, 472-484, (2009).

187 Mazur, J. E. Hyperbolic value addition and general models of animal choice. *Psychological review* **108**, 96-112, (2001).

188 Mazur, J. E. & Biondi, D. R. Delay-amount tradeoffs in choices by pigeons and rats: hyperbolic versus exponential discounting. *J Exp Anal Behav* **91**, 197-211, (2009).

189 McAlonan, K. & Brown, V. J. Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behav Brain Res* **146**, 97-103, (2003).

190 McCauley, C. & Graham, N. Influence of values in risky decision making: A formalization. *Representative Research in Social Psychology* **2**, (1971).

191 McClure, S. M. *et al.* Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**, 379-387, (2004).

192 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y. & Schoenbaum, G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* **31**, 2700-2705, (2011).

193 McDannald, M. A. *et al.* Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *The European journal of neuroscience* **35**, 991-996, (2012).

194 McDonald, A. J. Organization of amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience* **44**, 1-14, (1991).

195 McGonigle, B. O. & Chalmers, M. Are monkeys logical? , (1977).

196 McNamee, D., Rangel, A. & O'Doherty, J. P. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat Neurosci* **16**, 479-485, (2013).

197 Mellers, B., Schwartz, A., Ho, K. & Ritov, I. Elation and disappointment: Emotional responses to risky options. *Psychological science* **8**, 423-429, (1997).

198 Mellers, B., Schwartz, A. & Ritov, I. Emotion-based choice. *Journal of experimental psychology General* **128**, 332-345, (1999).

199 Meunier, M., Bachevalier, J. & Mishkin, M. Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia* **35**, 999-1015, (1997).

200 Miller, D. T. & Gunasegaram, S. Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of personality and social psychology* **59**, 1111, (1990).

201 Mishkin, M., Vest, B., Waxler, M. & Rosvold, H. E. A re-examination of the effects of frontal lesions on object alternation. *Neuropsychologia* **7**, 357-363, (1969).

202   Mobini, S. *et al.* Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **160**, 290-298, (2002).

203   Montague, P., Dolan, R., Friston, K. & Dayan, P. Computational psychiatry. *TRENDS in Cognitive Sciences* **16**, 72-152, (2012).

204   Montague, P. & Lohrenz, T. To detect and correct: norm violations and their enforcement. *Neuron* **56**, 14-18, (2007).

205   Montague, P. R. & Berns, G. S. Neural economics and the biological substrates of valuation. *Neuron* **36**, 265-284, (2002).

206   Montague, P. R., King-Casas, B. & Cohen, J. D. Imaging valuation models in human choice. *Annu Rev Neurosci* **29**, 417-448, (2006).

207   Mora, F., Avrith, D. B., Phillips, A. G. & Rolls, E. T. Effects of satiety on self-stimulation of the orbitofrontal cortex in the rhesus monkey. *Neuroscience letters* **13**, 141-145, (1979).

208   Mora, F., Avrith, D. B. & Rolls, E. T. An electrophysiological and behavioural study of self-stimulation in the orbitofrontal cortex of the rhesus monkey. *Brain research bulletin* **5**, 111-115, (1980).

209   Morrison, Sara E., Saez, A., Lau, B. & Salzman, C. D. Different Time Courses for Learning-Related Changes in Amygdala and Orbitofrontal Cortex. *Neuron* **71**, 1127-1140, (2011).

210   Muenzinger, K. F. Vicarious trial and error at a point of choice: I. a general survey of its relation to learning efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology* **53**, 75-86, (1938).

211   Murray, E. A., Baxter, M. G. & Gaffan, D. Monkeys with rhinal cortex damage or neurotoxic hippocampal lesions are impaired on spatial scene learning and object reversals. *Behav Neurosci* **112**, 1291-1303, (1998).

212   Murray, E. A., O'Doherty, J.P., & Schoenbaum, G. What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci* **27**, 8166-8169, (2007).

213   Noonan, M. P., Kolling, N., Walton, M. E. & Rushworth, M. F. Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *The European journal of neuroscience* **35**, 997-1010, (2012).

214   Noonan, M. P., Mars, R. B. & Rushworth, M. F. Distinct roles of three frontal cortical areas in reward-guided behavior. *J Neurosci* **31**, 14399-14412, (2011).

215   Noonan, M. P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20547-20552, (2010).

216   O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452-454, (2004).

217   O'Doherty, J. *et al.* Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport* **11**, 893-897, (2000).

218   O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329-337, (2003).

219   O'Doherty, J. P., Deichmann, R., Critchley, H. D. & Dolan, R. J. Neural responses during anticipation of a primary taste reward. *Neuron* **33**, 815-826, (2002).

220   O'Donnell, P. Ensemble coding in the nucleus accumbens. *Psychobiology* **27**, 187-197, (1999).

221   Ongur, D. & Price, J. L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral cortex* **10**, 206-219, (2000).

222   Ostlund, S. & Balleine, B. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J Neurosci* **27**, 4819, (2007).

223   Padoa-Schioppa, C. Neurobiology of economic choice: a good-based model. *Annu Rev Neurosci* **34**, 333-359, (2011).

224   Padoa-Schioppa, C. Neuronal Origins of Choice Variability in Economic Decisions. *Neuron* **80**, 1322-1336, (2013).

225   Padoa-Schioppa, C. Orbitofrontal cortex and the computation of economic value. *Ann N Y Acad Sci* **1121**, 232-253, (2007).

226   Padoa-Schioppa, C. Range-adapting representation of economic value in the orbitofrontal cortex. *J Neurosci* **29**, 14004-14014, (2009).

227   Padoa-Schioppa, C. & Assad, J. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat Neurosci* **11**, 95-102, (2008).

228   Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223-226, (2006).

229   Padoa-Schioppa, C. & Cai, X. The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. *Ann N Y Acad Sci* **1239**, 130-137, (2011).

230   Pan, W. X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J Neurosci* **25**, 6235-6242, (2005).

231   Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S. & Redish, A. D. Interactions between deliberation and delay-discounting in rats. *Cognitive, affective & behavioral neuroscience* **12**, 513-526, (2012).

232   Paton, J. J., Belova, M. A., Morrison, S. E. & Salzman, C. D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865-870, (2006).

233 Pears, A., Parkinson, J., Hopewell, L., Everitt, B. & Roberts, A. Lesions of the orbitofrontal but not medial prefrontal cortex disrupt conditioned reinforcement in primates. *Journal of Neuroscience* **23**, 11189-11201, (2003).

234 Pearson, J. M., Watson, K. K. & Platt, M. L. Decision Making: The Neuroethological Turn. *Neuron* **82**, 950-965, (2014).

235 Peters, J. & Buchel, C. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* **66**, 138-148, (2010).

236 Peters, J. & Büchel, C. Neural representations of subjective reward value. *Behavioural brain research* **213**, 135-141, (2010).

237 Phillips, A. G., Mora, F. & Rolls, E. T. Intracranial self-stimulation in orbitofrontal cortex and caudate nucleus of rhesus monkey: effects of apomorphine, pimozide, and spiroperidol. *Psychopharmacology* **62**, 79-82, (1979).

238 Pickens, C. L., Saddoris, M. P., Gallagher, M. & Holland, P. C. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behavioral Neuroscience* **119**, 317, (2005).

239 Pickens, C. L. *et al.* Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *The Journal of Neuroscience* **23**, 11078-11084, (2003).

240 Pieters, R. & Zeelenberg, M. On bad decisions and deciding badly: When intention–behavior inconsistency is regrettable. *Organizational Behavior and Human Decision Processes* **97**, 18-30, (2005).

241 Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* **27**, 9984-9988, (2007).

242 Plassmann, H., O'Doherty, J., Shiv, B. & Rangel, A. Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences* **105**, 1050-1054, (2008).

243 Platt, M. L. & Hayden, B. Learning: not just the facts, ma'am, but the counterfactuals as well. *PLoS biology* **9**, e1001092, (2011).

244 Premack, D. Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences* **104**, 13861-13867, (2007).

245 Preuschoff, K., Bossaerts, P. & Quartz, S. R. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381-390, (2006).

246 Preuss, T. Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *Cognitive Neuroscience, Journal of* **7**, 1-24, (1995).

247 Price, J. L. Beyond the primary olfactory cortex: olfactory-related areas in the neocortex, thalamus and hypothalamus. *Chemical senses* **10**, 239-258, (1985).

248 Price, J. L. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci* **1121**, 54-71, (2007).

249 Ragozzino, M. E. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann N Y Acad Sci* **1121**, 355-375, (2007).

250 Ramus, S., Davis, J., Donahue, R., Discenza, C. & Waite, A. Interactions between the orbitofrontal cortex and the hippocampal memory system during the storage of long-term memory. *Ann NY Acad Sci* **1121**, 216-231, (2007).

251 Ramus, S. J. & Eichenbaum, H. Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. *J Neurosci* **20**, 8199-8208, (2000).

252 Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience* **9**, 545-556, (2008).

253 Rangel, A. & Hare, T. Neural computations associated with goal-directed choice. *Current opinion in neurobiology* **20**, 262-270, (2010).

254 Redish, A. D., Jensen, S., Johnson, A. & Kurth-Nelson, Z. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychological review* **114**, 784-805, (2007).

255 Redish, A. J., A. A computational model of craving and obsession. *Annals of the New York Academy of Sciences* **1104**, 324-339, (2007).

256 Rescorla, R. Pavlovian conditioning. *American Psychologist* **43**, 151-160, (1988).

257 Rescorla, R. A. "Configural" conditioning in discrete-trial bar pressing. *Journal of comparative and physiological psychology* **79**, 307-317, (1972).

258 Ritov, I. Probability of regret: Anticipation of uncertainty resolution in choice. *Organizational Behavior and Human Decision Processes* **66**, 228-236, (1996).

259 Roesch, M., Calu, D., Esber, G. & Schoenbaum, G. All That Glitters... Dissociating Attention and Outcome Expectancy From Prediction Errors Signals. *Journal of Neurophysiology* **104**, 587, (2010).

260 Roesch, M. & Schoenbaum, G. From associations to expectancies: orbitofrontal cortex as gateway between the limbic system and representational memory. *The Orbitofrontal Cortex. DH Zald & SL Rauch, Eds*, 199–236, (2006).

261 Roesch, M., Taylor, A. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51**, 509, (2006).

262 Roesch, M. R., & Olson, C.R. Does It Represent Value or Reflect Motivation? *Annals of the New York Academy of Sciences* **1121**, 431-446, (2007).

263   Roesch, M. R., Calu, D. J., Burke, K. A. & Schoenbaum, G. Should I stay or should I go? Transformation of time-discounted rewards in orbitofrontal cortex and associated brain circuits. *Ann N Y Acad Sci* **1104**, 21-34, (2007).

264   Roesch, M. R., Calu, D. J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci* **10**, 1615-1624, (2007).

265   Roesch, M. R. & Olson, C. R. Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J Neurophysiol* **94**, 2457-2471, (2005).

266   Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E. & Schoenbaum, G. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci* **29**, 13365-13376, (2009).

267   Roese, N. J. Counterfactual thinking. *Psychological bulletin* **121**, 133, (1997).

268   Roese, N. J. & Olson, J. M. in *This book grew out of the symposium" Counterfactual Research: Current Issues and Recent Research," held at the convention of the American Psychological Assn, Toronto, Canada, 1993.* (Lawrence Erlbaum Associates, Inc).

269   Roitman, M. F., Wheeler, R. A. & Carelli, R. M. Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**, 587-597, (2005).

270   Rolls, E. The orbitofrontal cortex and reward. *Cerebral cortex* **10**, 284-294, (2000).

271   Rolls, E., Sienkiewicz, Z. & Yaxley, S. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *European Journal of Neuroscience* **1**, 53-60, (1989).

272   Rolls, E. T. The orbitofrontal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **351**, 1433-1443; discussion 1443-1434, (1996).

273   Rolls, E. T. The orbitofrontal cortex and reward. *Cereb Cortex* **10**, 284-294, (2000).

274   Rolls, E. T., Hornak, J., Wade, D. & McGrath, J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of neurology, neurosurgery, and psychiatry* **57**, 1518-1524, (1994).

275   Romanski, L. M., Bates, J. F. & Goldman-Rakic, P. S. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *The Journal of comparative neurology* **403**, 141-157, (1999).

276   Rosati, A. G. & Hare, B. Chimpanzees and bonobos distinguish between risk and ambiguity. *Biology letters* **7**, 15-18, (2011).

277   Rose, J. & Woolsey, C. The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep and cat. *Research Publications-Association for Research in Nervous and Mental Disease* **27**, 210-232, (1947).

278 Roseman, I. J. in *87th Annual Convention of the American Psychological Association, New York.*

279 Rudebeck, P. H., Bannerman, D. M. & Rushworth, M. F. The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cognitive, affective & behavioral neuroscience* **8**, 485-497, (2008).

280 Rudebeck, P. H. *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience* **28**, 13775-13785, (2008).

281 Rudebeck, P. H. & Murray, E. A. Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *The Journal of Neuroscience* **31**, 10569-10578, (2011).

282 Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S. & Murray, E. A. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci* **16**, 1140-1145, (2013).

283 Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M. & Rushworth, M. F. Separate neural pathways process different decision costs. *Nat Neurosci* **9**, 1161-1168, (2006).

284 Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054-1069, (2011).

285 Rustichini, A. Neuroeconomics: what have we found, and what should we search for. *Current opinion in neurobiology* **19**, 672-679, (2009).

286 Sage, A. P. & White, E. B. Decision and information structures in regret models of judgment and choice. *Systems, Man and Cybernetics, IEEE Transactions on*, 136-145, (1983).

287 Samuelson, P. A. FOUNDATIONS OF ECONOMIC ANALYSIS. (1947).

288 Samuelson, P. A. A Note on the Pure Theory of Consumer's Behaviour. *Economica* **5**, 61-71, (1938).

289 Savage, L. J. The theory of statistical decision. *Journal of the American Statistical association* **46**, 55-67, (1951).

290 Saver, J. L. & Damasio, A. R. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia* **29**, 1241-1249, (1991).

291 Schilman, E. A., Uylings, H. B., Galis-de Graaf, Y., Joel, D. & Groenewegen, H. J. The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neuroscience letters* **432**, 40-45, (2008).

292 Schmahmann, J. D. *et al.* Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain : a journal of neurology* **130**, 630-653, (2007).

293   Schmitzer-Torbert, N. & Redish, A. D. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J Neurophysiol* **91**, 2259-2272, (2004).

294   Schoeffler, M. S. Prediction of some stochastic events: A regret equalization model. *Journal of experimental psychology* **64**, 615, (1962).

295   Schoenbaum, G., Chiba, A. A. & Gallagher, M. Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. *J Neurosci* **20**, 5179-5189, (2000).

296   Schoenbaum, G., Chiba, A. A. & Gallagher, M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J Neurosci* **19**, 1876-1884, (1999).

297   Schoenbaum, G., Chiba, A. A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* **1**, 155-159, (1998).

298   Schoenbaum, G. & Eichenbaum, H. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J Neurophysiol* **74**, 733-750, (1995).

299   Schoenbaum, G. & Eichenbaum, H. Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. *J Neurophysiol* **74**, 751-762, (1995).

300   Schoenbaum, G. & Esber, G. R. How do you (estimate you will) like them apples? Integration as a defining trait of orbitofrontal function. *Current opinion in neurobiology* **20**, 205-211, (2010).

301   Schoenbaum, G., Nugent, L. S., Saddoris, M. P. & Setlow, B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* **13**, 885-890, (2002).

302   Schoenbaum, G. & Roesch, M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron* **47**, 633-636, (2005).

303   Schoenbaum, G., Roesch, M. & Stalnaker, T. Orbitofrontal cortex, decision-making and drug addiction. *Trends in Neuroscience* **29**, 116-124, (2006).

304   Schoenbaum, G., Roesch, M. R., Stalnaker, T. A. & Takahashi, Y. K. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature reviews. Neuroscience* **10**, 885-892, (2009).

305   Schoenbaum, G., Saddoris, M. P. & Stalnaker, T. A. Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann N Y Acad Sci* **1121**, 320-335, (2007).

306   Schoenbaum, G. & Setlow, B. Integrating orbitofrontal cortex into prefrontal theory: common processing themes across species and subdivisions. *Learn Mem* **8**, 134-147, (2001).

307   Schoenbaum, G., Setlow, B., Nugent, S., Saddoris, M. & Gallagher, M. Lesions of Orbitofrontal Cortex and Basolateral Amygdala Complex Disrupt Acquisition of Odor-Guided Discriminations and Reversals. *Learning & Memory* **10**, 129-140, (2003).

308   Schoenbaum, G., Setlow, B., Saddoris, M. & Gallagher, M. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* **39**, 855-867, (2003).

309   Schoenbaum, G., Takahashi, Y., Liu, T.-L. & McDannald, M. Does the orbitofrontal cortex signal value? *Ann NY Acad Sci* **1239**, 87-186, (2011).

310   Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599, (1997).

311   Schultz, W., Tremblay, L. & Hollerman, J. Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral cortex* **10**, 272-283, (2000).

312   Schwartz, O. & Simoncelli, E. P. Natural signal statistics and sensory gain control. *Nat Neurosci* **4**, 819-825, (2001).

313   Setlow, B., Schoenbaum, G. & Gallagher, M. Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* **38**, 625-636, (2003).

314   Simmons, J. M. & Richmond, B. J. Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. *Cereb Cortex* **18**, 93-103, (2008).

315   Simon, D. A. & Daw, N. D. Neural correlates of forward planning in a spatial decision task in humans. *J Neurosci* **31**, 5526-5539, (2011).

316   Singh, T. *et al.* The role of the nucleus accumbens in knowing when to respond. *Learning & Memory* **18**, 85, (2011).

317   Small, D. M. *et al.* Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* **39**, 701-711, (2003).

318   Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C. & Jones-Gotman, M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain : a journal of neurology* **124**, 1720-1733, (2001).

319   Spitzer, M., Fischbacher, U., Herrnberger, B., Gron, G. & Fehr, E. The neural signature of social norm compliance. *Neuron* **56**, 185-196, (2007).

320   St Onge, J. R. & Floresco, S. B. Prefrontal cortical contribution to risk-based decision making. *Cereb Cortex* **20**, 1816-1828, (2010).

321   Stalnaker, T. A. *et al.* Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nat Commun* **5**, (2014).

322   Stalnaker, T. A., Franz, T. M., Singh, T. & Schoenbaum, G. Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron* **54**, 51-58, (2007).

323	Stalnaker, T. A. *et al.* Neural correlates of inflexible behavior in the orbitofrontal-amygdalar circuit after cocaine exposure. *Ann N Y Acad Sci* **1121**, 598-609, (2007).

324	Stalnaker, T. A., Roesch, M. R., Franz, T. M., Burke, K. A. & Schoenbaum, G. Abnormal associative encoding in orbitofrontal neurons in cocaine-experienced rats during decision-making. *The European journal of neuroscience* **24**, 2643-2653, (2006).

325	Stalnaker, T. A. *et al.* Cocaine-induced decision-making deficits are mediated by miscoding in basolateral amygdala. *Nat Neurosci* **10**, 949-951, (2007).

326	Stalnaker, T. A., Takahashi, Y., Roesch, M. R. & Schoenbaum, G. Neural substrates of cognitive inflexibility after chronic cocaine exposure. *Neuropharmacology* **56 Suppl 1**, 63-72, (2009).

327	Steiner, A. P. & Redish, A. D. Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat Neurosci* **17**, 995-1002, (2014).

328	Steiner, A. P. & Redish, A. D. The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front Neurosci* **6**, 131, (2012).

329	Stopper, C. M., Green, E. B. & Floresco, S. B. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cerebral cortex* **24**, 154-162, (2014).

330	Stott J. J. , R. A. D. A functional difference in information processing between orbitofrontal cortex and ventral striatum during decision-making behavior. *Philosophical Transactions of the Royal Society B* **In press**, (2014).

331	Sul, J. H., Kim, H., Huh, N., Lee, D. & Jung, M. W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**, 449-460, (2010).

332	Sutton, R. S., & Barto, A.G. *Reinforcement Learning: An Introduction.* (MIT Press, 1998).

333	Takahashi, Y. K. *et al.* Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron* **80**, 507-518, (2013).

334	Takahashi, Y. K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269-280, (2009).

335	Takahashi, Y. K. *et al.* Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci* **14**, 1590-1597, (2011).

336	Taylor, A., Hunt, G., Medina, F. & Gray, R. Do New Caledonian crows solve physical problems through causal reasoning? *Proceedings of the Royal Society B: Biological Sciences* **276**, 247-254, (2009).

337	Thornton, A., Clayton, N. S. & Grodzinski, U. Animal minds: from computation to evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 2670-2676, (2012).

338   Thorpe, S., Rolls, E. & Maddison, S. The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Experimental Brain Research* **49**, 93-115, (1983).

339   Timbie, C. & Barbas, H. Specialized Pathways from the Primate Amygdala to Posterior Orbitofrontal Cortex. *The Journal of Neuroscience* **34**, 8106-8118, (2014).

340   Tobler, P. N., O'Doherty J, P., Dolan, R. J. & Schultz, W. Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* **95**, 301-310, (2006).

341   Tolman, E. C. Prediction of vicarious trial and error by means of the schematic sowbug. *Psychological review* **46**, 318, (1939).

342   Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704-708, (1999).

343   Tremblay, L. & Schultz, W. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *Journal of Neurophysiology* **83**, 1864-1876, (2000).

344   Tsuchida, A., Doll, B. B. & Fellows, L. K. Beyond Reversal: A Critical Role for Human Orbitofrontal Cortex in Flexible Learning from Probabilistic Feedback. *The Journal of Neuroscience* **30**, 16868-16875, (2010).

345   Tsujimoto, S., Genovesio, A. & Wise, S. Monkey orbitofrontal cortex encodes response choices near feedback time. *Journal of Neuroscience* **29**, 2569, (2009).

346   Tversky, A. & Kahneman, D. The framing of decisions and the psychology of choice. *Science* **211**, 453-458, (1981).

347   Ursu, S. & Carter, C. Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Cognitive Brain Research* **23**, 51-60, (2005).

348   Uylings, H., Van Eden, C., De Bruin, J., Corner, M. & Feenstra, M. The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. *The Prefrontal Cortex: Its Structure, Function and Pathology: Its Structure, Function and Pathology* **85**, 95, (1991).

349   Valentin, V. V., Dickinson, A. & O'Doherty, J. P. Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci* **27**, 4019-4026, (2007).

350   van der Meer, M., Kurth-Nelson, Z. & Redish, A. D. Information processing in decision-making systems. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* **18**, 342-359, (2012).

351   van der Meer, M. & Redish, A. Expectancies in decision making, reinforcement learning, and ventral striatum. *Front Neurosci* **4**, 6, (2010).

352   van der Meer, M. & Redish, A. Low and high gamma oscillations in rat ventral striatum have distinct relationships to behavior, reward, and spiking activity on a learned spatial decision task. *Frontiers in Integrative Neuroscience* **3**, 9, (2009).

353   van der Meer, M. A., Johnson, A., Schmitzer-Torbert, N. C. & Redish, A. D. Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* **67**, 25-32, (2010).

354   van der Meer, M. A. & Redish, A. D. Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points. *Frontiers in integrative neuroscience* **3**, 1, (2009).

355   Van Der Schalk, J., Bruder, M. & Manstead, A. Regulating emotion in the context of interpersonal decisions: The role of anticipated pride and regret. *Frontiers in psychology* **3**, (2012).

356   van Duuren, E. *et al.* Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task. *Learn Mem* **14**, 446-456, (2007).

357   van Duuren, E., Lankelma, J. & Pennartz, C. M. Population coding of reward magnitude in the orbitofrontal cortex of the rat. *J Neurosci* **28**, 8590-8603, (2008).

358   van Duuren, E. *et al.* Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. *J Neurosci* **29**, 8965-8976, (2009).

359   Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43-48, (2001).

360   Wallis, J. D. Evaluating apples and oranges. *Nat Neurosci* **9**, 596-598, (2006).

361   Wallis, J. D. Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* **30**, 31-56, (2007).

362   Wallis, J. D., Dias, R., Robbins, T. W. & Roberts, A. C. Dissociable contributions of the orbitofrontal and lateral prefrontal cortex of the marmoset to performance on a detour reaching task. *The European journal of neuroscience* **13**, 1797-1808, (2001).

363   Wallis, J. D. & Miller, E. K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *The European journal of neuroscience* **18**, 2069-2081, (2003).

364   Walton, M. E., Bannerman, D. M. & Rushworth, M. F. The role of rat medial frontal cortex in effort-based decision making. *J Neurosci* **22**, 10996-11003, (2002).

365   Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H. & Rushworth, M. F. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* **65**, 927-939, (2010).

366 West, E. A., DesJardin, J. T., Gale, K. & Malkova, L. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci* **31**, 15128-15135, (2011).

367 Wikenheiser, A. M., Stephens, D. W. & Redish, A. D. Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 8308-8313, (2013).

368 Wilson, R. C. & Niv, Y. Inferring relevance in a changing world. *Front Hum Neurosci* **5**, 189, (2011).

369 Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267-279, (2014).

370 Woolhouse, R. VI.—COUNTERFACTUALS, DISPOSITIONS, AND CAPACITIES. *Mind* **82**, 557-565, (1973).

371 Yoshida, K., Saito, N., Iriki, A. & Isoda, M. Social error monitoring in macaque frontal cortex. *Nat Neurosci* **15**, 1307-1312, (2012).

372 Young, P. T. Relative food preferences of the white rat. *Journal of Comparative Psychology* **14**, 297, (1932).

373 Zald, D. H. *et al.* Meta-analytic connectivity modeling reveals differential functional connectivity of the medial and lateral orbitofrontal cortex. *Cereb Cortex* **24**, 232-248, (2014).

374 Zeeb, F. D., Robbins, T. W. & Winstanley, C. A. Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **34**, 2329-2343, (2009).

375 Zeelenberg, M. Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making* **12**, 93-106, (1999).

376 Zeelenberg, M. & Pieters, R. Looking backward with an eye on the future. *Judgments over time: The interplay of thoughts, feelings, and behaviors*, 210-229, (2006).

377 Zeelenberg, M. & Pieters, R. A theory of regret regulation 1.0. *Journal of Consumer Psychology* **17**, 3-18, (2007).

378 Zeelenberg, M. & van Dijk, E. 9 On the comparative nature of regret. *The psychology of counterfactual thinking*, 147, (2007).

379 Zeelenberg, M. *et al.* Emotional Reactions to the Outcomes of Decisions: The Role of Counterfactual Thought in the Experience of Regret and Disappointment. *Organ Behav Hum Decis Process* **75**, 117-141, (1998).

380 Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J Neurophysiol* **79**, 1017-1044, (1998).

381   Zhou, W. & Crystal, J. D. Evidence for remembering when events occurred in a rodent model of episodic memory. *Proceedings of the National Academy of Sciences* **106**, 9525-9529, (2009).