

**Exploring the brain mechanisms involved in reward
prediction errors using conditioned inhibition**

by

Jessica Mollick

B.A., U.C. Berkeley 2011

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Master of Philosophy
Department of Psychology

2013

This thesis entitled:
Exploring the brain mechanisms involved in reward prediction errors using conditioned inhibition
written by Jessica Mollick
has been approved for the Department of Psychology

Randall O'Reilly

Prof. Tor Wager

Prof. Matt Jones

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Mollick, Jessica (M.S. Psychology)

Exploring the brain mechanisms involved in reward prediction errors using conditioned inhibition

Thesis directed by Prof. Randall O'Reilly

Many previous studies of the brain areas involved in reward prediction errors have not accounted for the downstream projections of dopamine areas when interpreting these results. We propose that paradigms like conditioned inhibition, which involves pairing a rewarded CS with an inhibitor that always cancels the reward, can reduce this confound and allow for further specification of the computational role different brain regions play into the RPE signal. Further predictions of the role of different brain areas in reward learning and how positive and negative valence learning interact in the brain are inspired by the PVLV model, a more biologically plausible alternative to TD learning that uses two parameters, learned value and primary value, compared to the single RPE parameter used by TD. To test the predictions of the PVLV model and compare activity across different conditions that allowed us to examine the roles of different regions in RPE computation, we ran a conditioned inhibition fMRI study using juice rewards, with a particular focus on examining the brain activity in several regions of interest in the PVLV model, including the ventral striatum, central nucleus of the amygdala and lateral habenula. We found that the rewarded CS activated the VTA/SN, consistent with many other studies of reward learning, as well as several regions in the basal ganglia. There was also overlap between activations for the CS, CS+Inhibitor and Inhibitor in the prefrontal cortex, insula and basal ganglia. Also, better than expected rewards activated the medial OFC; while worse than expected rewards activated the lateral OFC. In the amygdala, we found increased activity for the juice reward compared to the neutral solution, but we did not find increased activity for the rewarded CS as predicted by the PVLV model.

Contents

Chapter	
1	Introduction 1
1.0.1	Introduction 1
1.0.2	Conditioned inhibition 2
1.0.3	Overview of PVLV model 4
1.0.4	Alternative accounts 9
1.0.5	Specific predictions for conditioned inhibition 10
1.0.6	Results summary 15
2	Biological data 17
2.0.7	CeA and BLA 17
2.0.8	RMTg and habenula 21
2.0.9	Dopamine and negative valence areas 22
2.0.10	Brain systems for positive and negative valence 23
3	Methods 25
3.0.11	Behavioral methods 25
3.0.12	Data analysis 29
3.0.13	Analysis methods 29
3.0.14	Computational model 31

4	Results	34
4.0.15	Juice reward analysis	34
4.0.16	CS Type analysis	37
4.0.17	Rewarded CS activity	38
4.0.18	Expectation analysis	42
4.0.19	Prediction errors in SN/VTA and habenula	44
5	Discussion	46
5.0.20	SN/VTA and lateral habenula	46
5.0.21	Amygdala	48
5.0.22	Basal ganglia	49
5.0.23	Prefrontal Cortex and Insula	50
5.0.24	Conclusion	52
	Bibliography	55
	Appendix	
A	Modeling equations: PVLV and Rescorla-Wagner	63
A.0.25	Rescorla-Wagner model of reward learning	63
A.0.26	PVLV model: simulating positive and negative valence	63
B	Simulation diagrams of conditioned inhibition	65
B.0.27	LVe (central amygdala) at time of CS	66
B.0.28	PVi learning for rewarded CS	67
B.0.29	VTA learning for rewarded CS	68
B.0.30	Conditioned inhibition learning	69

Chapter 1

Introduction

1.0.1 Introduction

Reinforcement learning has been a central issue in psychology for at least 100 years, and has inspired a multidisciplinary tradition of research on the behavioral, computational and biological mechanisms involved in learning associations with reward. A major breakthrough in the field occurred when the functional role of dopamine neurons was found to fit with formal learning theories. [93, 95]. The key concept in reinforcement learning is a reward prediction error (RPE), the difference between the amount of reward expected and the reward actually received, which is directly encoded by dopamine neurons [94, 93]. When an unexpected positive valence reward is received dopamine neurons respond with a burst when the reward is presented [57, 93]. Once the reward becomes expected, dopamine neurons respond with a burst to the CS that predicts the reward. If this reward is unexpectedly omitted, dopamine neurons will respond with a dip when the reward was expected to occur. [93]

Recently, there have been many fMRI studies which have found this reward prediction signal directly in the brain. However, this RPE signal has been observed in several regions, including the amygdala [54, 56, 6], nucleus accumbens, [39, 55], ventral tegmental area [16] and ventromedial prefrontal cortex [4, 89]. As these brain areas receive projections from dopamine neurons in the midbrain, there is a possibility that the observed activity could be a downstream broadcast of the RPE signal encoded by dopamine neurons. In this study, we got around this problem by using a paradigm that minimizes this dopamine signal, while still maintaining reward associations. We did

this by pairing a conditioned inhibitor with a positive valence CS. As a conditioned inhibitor is defined by the ability to reduce the behavioral conditioned response in the opposite direction of the excitatory CS, the pairing of the inhibitor with the positive CS effectively cancels out the positive reward association, and inhibits dopamine neurons [105].

Conditioned inhibitors are interesting because they act similarly to negatively valenced stimuli by reducing the overall conditioned response to an excitatory CS. [62], and there has been a lot of interest in understanding the relationship between positive and negative valence pathways. A good deal of evidence suggests that negative valence stimuli activate distinct aversive pathways in the brain. Does a conditioned inhibitor activate the same negative valence pathways?

The diversity of neural pathways that contribute to a reward prediction error signal contrasts with the simplicity of the prominent reward learning computational framework, temporal difference learning (TD). There is an alternative model to TD, PVLV[71, 30], which explains the role of differential neural pathways, motivated by existing data, in the computation of the RPE signal. PVLV stands for primary value, learned value, which captures the model's fundamental distinction between the LV system, which controls anticipatory responding for CSes, and the PV system, which controls responding for USes, at the time of reward presentation. Crucially, the PVi part of the system decreases responses to expected rewards, accounting for the data showing that dopamine firing at the time of US presentation decreases for expected rewards.

While this model provides one set of hypotheses about how these brain areas work to compute reward prediction errors, we can also test these specific predictions in human fMRI studies. This study provided one such test by using a conditioning paradigm called conditioned inhibition, which can minimize the confounding influence of the widespread dopamine projections, hopefully allowing us to more fully elucidate the role of these areas in the computation of reward prediction error.

1.0.2 Conditioned inhibition

Throughout life, we face many cases where a particular expected reward is not received. For example, imagine a child in the kitchen who knows that his favorite cereal box is associated with a

delicious sugar rush, but his mom doesn't allow him to eat it. How does this affect the child? Does the cereal become even more desirable when the mother isn't around? And how does the child feel about the mother (or father) figure that doesn't allow the cereal to be eaten? This example describes the situation of conditioned inhibition, which occurs in everyday life anytime a rewarded CS is paired with something that makes the reward impossible to attain.

Conditioned inhibition involves first pairing a CS with reward, then introducing the conditioned inhibitor, which always predicts cancellation of the reward when paired with an excitatory CS. In monkey VTA and substantial nigra pars compacta (SNc), dopamine neurons responds with an overall depression to the conditioned inhibitor, while they continue to respond with a burst to the excitatory CS. [19]. While an excitatory CS is defined in terms of the ability to elicit a conditioned response, a conditioned inhibitor is defined by the ability to reduce the behavioral conditioned response caused by the excitatory CS [82]. A stimulus fits the definition of a conditioned inhibitor if the combination of the excitatory CS and the inhibitor causes a reduced conditioned response compared to the excitatory CS alone [82, 75]. Therefore, it is the laboratory equivalent of the mother with cereal example described earlier, and the mother can be described as a conditioned inhibitor if the child's conditioned response to seeing the cereal box (say, reaching a hand inside and grabbing some tasty froot loops) is reduced when the inhibitor (the mother) is around, but the child still reaches into the froot loops box when alone.

Conditioned inhibition can address the concern that activity in many regions playing a role in the computation of reward prediction error is confounded by prediction error signals because the conditioned inhibition paradigm includes several conditions where dopamine signaling is actually inhibited. Similarly, we chose the paradigm because it allowed us to test several specific hypotheses about the brain systems involved in learning for both better than expected and worse than expected trials, and how these signals train associations for the associated CSes. As described in further detail in Section 1.0.5, the PVLV model generates a set of specific hypotheses about the brain areas involved in each phase of learning for the conditioned inhibition task.

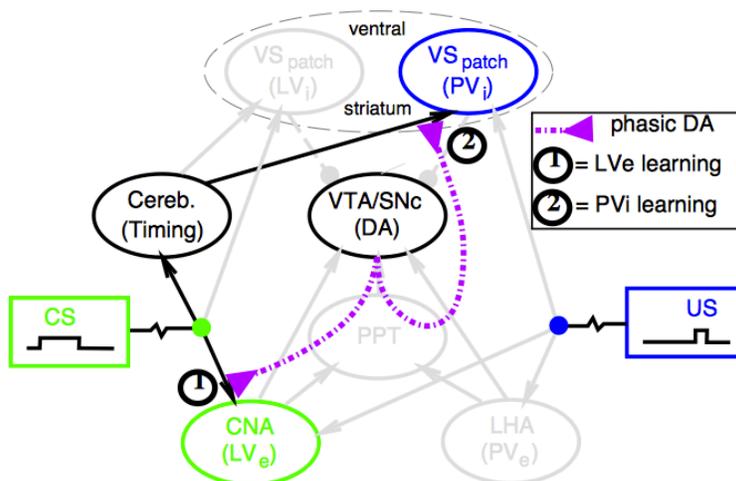
1.0.3 Overview of PVLV model

Providing an alternative to temporal differences learning, the PVLV model [71, 30] describes the role of distinct neural pathways in the computation of a reward prediction error signal, motivated by existing data. In comparison to TD learning, which relies on the single mechanism of reward prediction error, the PVLV model uses two mechanisms, but both models compute similar, but not identical, values corresponding to dopamine release.

In the PVLV model, the fundamental distinction is between two different learning systems - the LV system controls anticipatory responding to a CS that reliably predicts a reward, and the PV system which regulates responses at the time of reward. In PVLV, the PV system regulates learning at the time of reward presentation using the Rescorla-Wagner learning rule, which calculates the difference between the expected reward value and the actual outcome, and then changes the synaptic weights for the associated sensory input to reflect this difference.

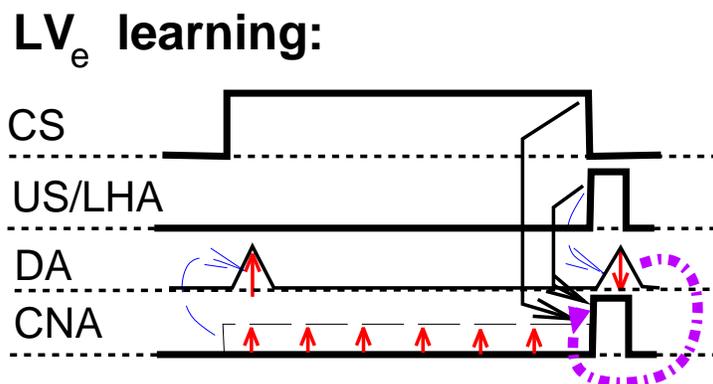
How does the basic sequence of reward learning work in PVLV? We begin with a typical reward learning trial, where a conditioned stimulus (CS) is followed by a primary positive valence reward, which sends an input to the excitatory component of the PV system (PVe), corresponding to neurons in the lateral hypothalamic area (LHA) that receive direct projections from sensory areas associated with primary rewards and send an excitatory projection to the midbrain dopamine nuclei [30]. As shown in Figure 1.1, this causes dopamine release, which causes learning in two PVLV subsystems, LVe and PVi, which will regulate later responses to this primary reward and conditioned stimulus. Once an association between a conditioned stimulus and a reward is learned, the LVe and PVi systems work together to regulate dopamine signaling at different parts of a reward trial.

Figure 1.1: Adapted from [30], this shows how learning occurs in PVLV when a primary reward is presented, causing dopamine release through projections of the lateral hypothalamic area to VTA/SNc. Learning occurs at two sites, the LVe projections in the CeA, and the PVi projections in the ventral striatum.



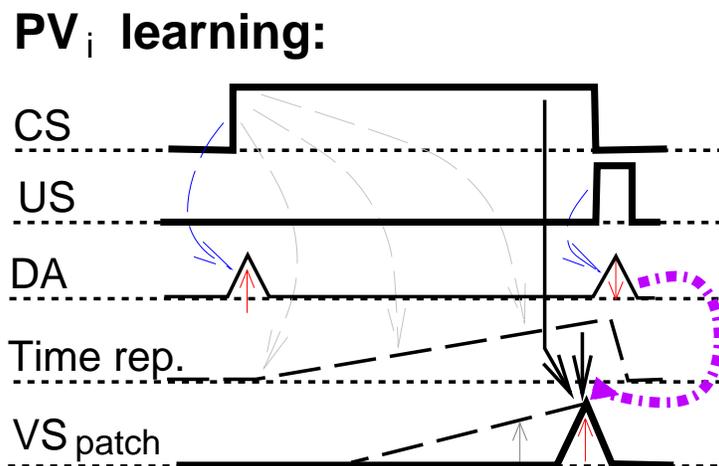
Over the course of learning, the LVe system changes the associative weights for a CS based on the discrepancy between the current LV weight for that CS and the current reward value (PV), as described in Equation A.4. While LVe learning occurs at the time of primary reward presentation, these LVe weights only influence dopamine release at the time of CS presentation, influencing anticipatory responses to the rewarded CS. This is illustrated in Figure 1.2, which shows how dopamine release at the time of reward presentation increases anticipatory weights in LVe/CNA that will drive excitatory dopamine release at the time of CS presentation. Biologically, the LV system is represented by neurons in the central amygdala, which plays an important role in controlling anticipatory responding for positive valence CSes, and has excitatory projections to dopamine neurons in the midbrain, as will be reviewed in Section 2.

Figure 1.2: An illustration of LVe learning in PVLV from [30], showing how dopamine release at the time of reward increases LVe weights in the CNA. These LVe weights will then influence excitatory dopamine release when the CS is presented.



The second important component in PVLV that regulates dopamine release is the PVi system. Similarly to the LVe system, the PVi system learns at the time of reward presentation. However, the PVi system plays a functionally different role than the LVe system, and regulates dopamine responses at the time of reward presentation once a reward is expected. As shown in Figure 1.3, the PVi component of the model is actually represented by inhibitory neurons (patch cells) in the ventral striatum which show excitatory responses in advance of expected rewards. Similarly to the LVe system, this PVi system learns based on the occurrence of unconditioned stimuli (PVs), and will change the associated PVi weights for the current CS depending on how different the current reward value is from the current PVi weight for that CS. Once an association of a CS with reward is learned, the inhibitory neurons in the PVi system then send shunting inhibition onto dopamine neurons in the VTA/SNc, preventing an excitatory dopamine burst at the time when reward presentation is expected. Therefore, phasic dopamine release at the time of reward is reflected by the difference between PVe, the reward value of the stimulus that drives dopamine neurons, and PVi, the inhibitory influence from ventral striatal neurons that shunts firing to the reward. This phasic dopamine value is equivalent to a what would be computed by a standard reward prediction error signal.

Figure 1.3: An illustration of PV_i learning in PVLV from [30], showing how dopamine release at the time of reward increases PV_i weights in patch neurons in ventral striatum. These PV_i neurons will then showing ramping excitation, finally sending shunting inhibition to dopamine neurons in the VTA/SNc at the time of reward presentation for an expected reward, inhibiting a dopamine burst.

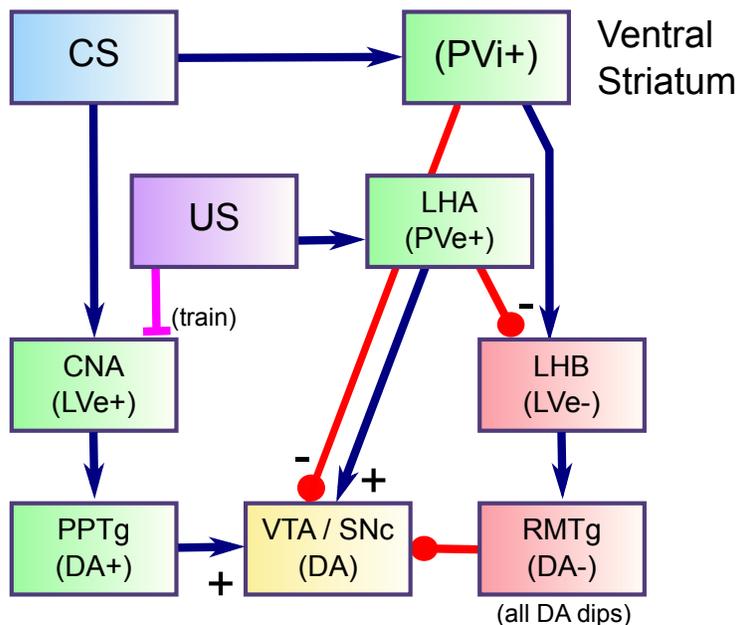


In PVLV, the combination of the LV weights that regulate anticipatory dopamine release at the time of CS presentation and PV_i weights that inhibit dopamine release for expected rewards is able to capture the finding that dopamine release moves from the time of the US to the time of the CS for expected rewards [93]. This same finding is captured by the backwards chaining of prediction error in time in TD learning, but one disadvantage of TD learning is that it only works well if the relationship between the CS and US is reliable. In contrast, the existence of two separate computational systems in PVLV allows it to be more robust to these changes, as well as allowing us to better capture existing biological data on the brain mechanisms behind the computation of the RPE [71].

Finally, a recent addition to PVLV has allowed us to provide an explanatory account of the neural mechanisms involved in unexpected reward omissions, which are crucial for learning the negative associations for conditioned inhibitors that reduce excitatory responses to CSes. While dopamine neurons respond with a dip to the omission of expected rewards, the source of this

dopamine dip was unknown until recent data revealed that the lateral habenula, a small structure located in the epithalamus near the pineal body, sends an inhibitory signal to dopamine neurons for reward omissions and negative valence, by projecting to GABAergic neurons in the RMTg (located nearby traditional dopamine neurons in the VTA) [52, 35]. Along with excitations for reward omissions and negative valence, the lateral habenula fires for CSes that predict these outcomes, providing evidence that it acts as a negatively valenced learned value signal (LVe-) which can then inhibit dopamine firing for CSes that have been associated with reward omission.

Figure 1.4: Full PVLV model for positive valence learning. Excitatory connections are shown in blue, inhibitory connections in red. The pink connection represents US representations in the CeA that train LVe CS representations over learning.



Bringing all of this together, Figure 1.4 shows the PVLV model with excitatory and inhibitory connections, providing an illustration of the biological circuitry which regulates responding to positively valenced conditioned stimuli. As described earlier, the PVe system in the lateral hypothala-

mus influences dopamine release to a positively valenced US. Over repeated pairings of a CS with this reward, the LVe weights from the CeA to VTA/SNc learn to drive anticipatory dopamine firing to the CS, and the PVi weights from the ventral striatum to the VTA/SNc inhibit dopamine release at the time of an expected reward. Using information from the PVi system through projections from the ventral striatum, the lateral habenula can learn about unexpected reward omissions and drive dopamine dips for inhibitory CSes. The computational roles of each brain area in PVLV in reward learning are supported by a wide range of biological data, reviewed in Section 2.

1.0.4 Alternative accounts

However, considering modeling of the conditioned inhibition paradigm, it is important to compare and contrast the predictions of several computational models of reward learning, particularly because they make some differential predictions.

As described earlier in the PVLV model, the Rescorla Wagner learning rule (see appendix A.1) that determines changes in associative strength of cues depending on the difference between an expected value and what actually occurs, has been shown to fit the function of dopamine neurons well. This model can actually also account for conditioned inhibition, because it allows the expected value to depend on the summed associative strength of all cues. Since the CS has a prior positive value, this leads to an overall positive expectation when the CS is paired with the inhibition. When this is followed by a reward omission, this discrepancy between the expected outcome and the actual outcome causes the inhibitor to acquire a negative associative value.

One other potential way of viewing the role of the inhibitor for the CS-Inhibitor pair involves viewing it as a contextual cue that influences the processing of the CS, potentially inhibiting conditioned responses if the inhibitory properties are well-learned. If this was the case, one could view the inhibitor as an "occasion setter" that indicates whether the association between the CS-Inhibitor pair and no-reward is currently active.[9]. However, this does not explain the dopamine dip that occurs when the inhibitor is viewed alone [105], which may signal a negative association for the inhibitor.

Another related theory posits that conditioned responses to the CS-Inhibitor trials might be different than to the rewarded CS due to the assignment of a different latent cause or internal model of the environment to the CS-Inhibitor trials compared to those of the CS-alone, learned by repeated pairings of the CS-Inhibitor with no reward [28]. However, if only few CS-Inhibitor trials are observed, they tend to be assigned to the same latent cause as the rewarded CS trials, but as more evidence accumulates they tend to be assigned to a new latent cause that predicts reward omission.

While both of these theoretical accounts can provide potential explanations for the occurrence of conditioned inhibition, they are not specific about the neural mechanisms behind the learning of conditioned inhibition. This is a similar issue that is faced by the TD framework, which can capture a wide range of learning phenomena with the simple reward prediction error (RPE) parameter, but cannot address the specific computational role of different brain areas in this computation. In contrast, the PVLV model provides a more complete explanation of the neural mechanisms behind learning of conditioned inhibition, and can account for the activity of dopamine neurons in different parts of the task through biological mechanisms that are well supported by existing data, such as the role of the lateral habenula in learning about unexpected reward omissions. To test these predictions, we used a model based fMRI approach, particularly by simulating a neural network of the different brain areas involved in the conditioned inhibition task.

1.0.5 Specific predictions for conditioned inhibition

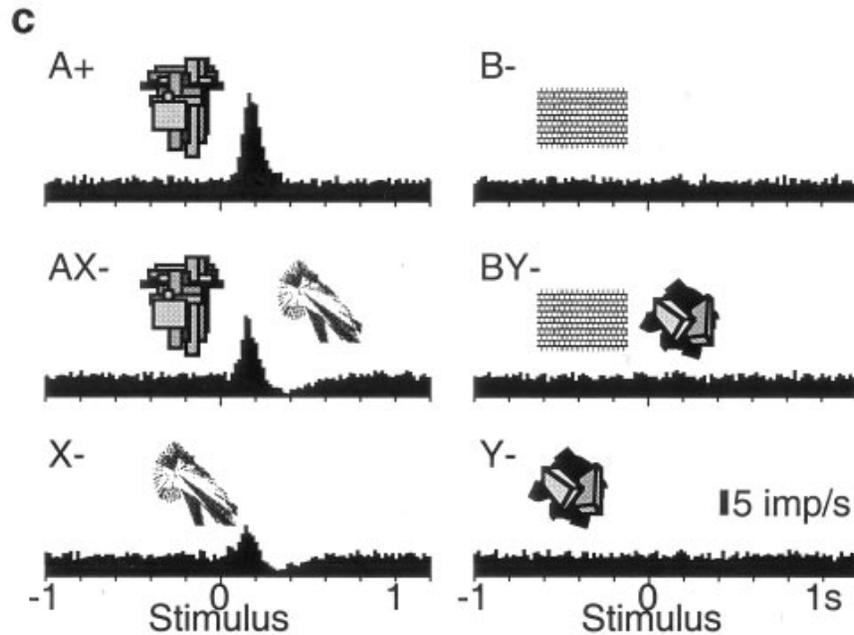
Given the described computational role of the brain areas included in the PVLV model, we were able to make specific predictions about the brain areas involved in conditioned inhibition. The ability to test these specific predictions was a major motivation for running the conditioned inhibition paradigm as an fMRI study.

1.0.5.1 Predictions motivated by previous versions of conditioned inhibition

As shown in Figure 1.5, which shows electrophysiological data from a conditioned inhibition paradigm ran in monkeys[105], neurons in the SNc/VTA show a specific activity profile for each of the CS Types in a conditioned inhibition task after the learning block. This enabled us to make specific predictions about the type of activity we would predict for each condition in SN/VTA. We had three major predictions about activity in the SN/VTA during presentation of each of the CSes.

- After learning, the rewarded CS should show the most overall BOLD activity in SNc/VTA compared to other stimuli. However, given what we know about dopamine neurons, activity in the SNc/VTA should first be strongest at the time of US presentation, and should then move to the time of the rewarded CS, once the pairing between the CS and reward is well-learned.
- The CS-Inhibitor shows a burst-dip profile in the SNc/VTA. Given the limited temporal resolution of traditional fMRI, it is unclear whether this would be result in significant activity in SNc/VTA.
- We would not expect the SNc/VTA to activate to the Inhibitor because neurons in SNc/VTA showed an overall depression to the Inhibitor

Figure 1.5: Adapted from [19]. The rewarded CS shows a burst at the time of CS presentation, while the CS+inhibitor shows depressed activity with a burst-dip profile, and the Inhibitor shows the least overall activity.



The electrophysiological recordings in [19] also enable us to make specific predictions about how activity changes in the SNc/VTA for the the CS, and CS+Inhibitor, over the course of learning. In particular, this enables us to make two predictions about the brain activity during the course of learning:

- The CS+Inhibitor (AX) may show a dopamine dip, possibly reflected as reduced brain activity in the SNc/VTA, at the time of reward omission in the first block, potentially due to disappointment of the reward expectation caused by the rewarded CS stimulus.
- After the first block, the CS+Inhibitor shows a reduced activity profile, and less of a dip at the time of reward omission. This may be due to the CS+Inhibitor pair decreasing the overall association with reward, leading to a lower reward expectation.

1.0.5.2 Predictions motivated by PVLV model

Due to the computational role of the central amygdala in encoding associations of CSes with reward, and the role of the lateral habenula in learning about worse than expected rewards and the CSes that predict those outcomes, we can make specific predictions about activity in these regions during the conditioned inhibition task.

Predictions for the lateral habenula:

- The lateral habenula shows a negatively valenced prediction error profile. Therefore, when the CS+Inhibitor is followed by no reward, the lateral habenula should show BOLD activity, especially in the first block. As reward omission becomes more predictable, this activity should move to the time of the CS that predicts reward omission.
- Similarly, as neurons in SNc/VTA showed an overall depression to the Inhibitor, we would expect the lateral habenula to activate to the Inhibitor.
- It is unclear how the lateral habenula would act to the CS+Inhibitor. Given the computational model, we would expect the habenula to activate to any trials showing the Inhibitor, but it is unclear whether the positive reward associations for the CS, which is viewed at the same time, would reduce the habenula activation to the inhibitor.

We also had specific predictions about how the central nucleus of the amygdala would act during the conditioned inhibition task. However, the function of the central amygdala is still somewhat controversial, as reviewed in Section 2.0.7, so the below predictions are limited to positive valence representations in the amygdala only.

- Given that we think the central amygdala encodes CS-Reward associations, we expect that it will be active when the rewarded CS is presented, after the association with reward is well-learned.
- The PVLV model also proposes that the central amygdala (LVE) will also increase its weights at the time of presentation of a primary reward (PV), so we would predict that the

central amygdala would activate at the time of reward presentation, especially during the first learning block where the reward association is first being learned.

- We would expect the central amygdala to show activity in any condition where a rewarded CS is viewed. Therefore, we would expect activity in the central amygdala for the CS-Inhibitor. However, there are several hypotheses that say that the amygdala also represents worse than expected and negative valence rewards, so we may expect amygdala activity for the Inhibitor as well.

Relating this back to the overall goal of discovering the specific computational role of different brain regions in the calculation of reward prediction error, the PVLV computational model provides an overall theoretical framework which allows us to flesh out the specific roles of the amygdala, ventral striatum, SNc/VTA and lateral habenula in the reward learning process. Importantly, we can make predictions about how the information encoded by each of these regions plays distinct roles in the computation of reward prediction error, and use these computational hypotheses to predict the timecourse of each region at the time of CS presentation and the period where reward is (or is not) presented.

In addition, we have included conditions in the conditioned inhibition paradigm, particularly the CS-Inhibitor and Inhibitor conditions, where dopamine signaling is reduced overall. This feature of the experimental paradigm relates to our second goal, which involves reducing the confounding influence of the reward prediction error signal (RPE), encoded by dopamine neurons, on downstream activity of the brain regions of interest, especially the ventral striatum, amygdala and prefrontal cortex. In this case, we should be able to examine the observed activity in each region, and how it changes over learning, to see whether our hypotheses about the particular role of each region hold out. In particular, an important test is whether the amygdala shows activity for the rewarded CS when paired with the inhibitor, because in that case the CS is still associated with reward, but the conjunction of these two stimuli always leads to inhibited firing in dopamine neurons. Overall, a major goal of using this paradigm was to allow us to examine whether there were dissociations

between the brain areas involved in driving anticipatory responses to CSes with reward associations, thought to involve the central amygdala, and regions that primarily regulate processing at the time of reward, such as the ventral striatum, as that is one of the central claims of the PVLV model.

1.0.6 Results summary

In the conditioned inhibition study, we found activity in the SNc/VTA when the rewarded CS was presented, consistent with the finding that dopamine is released to a rewarded CS. However, this was not accompanied by excitatory activity in the central amygdala, contradicting the PVLV model's prediction that the CeA drives phasic dopamine bursting for a rewarded CS. We also found activity in the SNc/VTA when comparing the activity for the juice reward to the activity for the neutral solution, suggesting that the juice reward was rewarding in comparison to the neutral solution, consistent with the finding that SNc/VTA neurons fire in response to rewarding stimuli. We also examined activity in the lateral habenula, and surprisingly found significant activity during the presentation of all conditioned stimuli, inconsistent with the prediction that activity in the lateral habenula would be selective to the Inhibitor and CS+Inhibitor which had associations with reward omission. In addition, the trials where the juice reward was unexpectedly omitted did not activate the lateral habenula.

While there was a lack of activity for the reward CS or any of the other stimuli in the amygdala, there was significantly more activity when the juice reward was presented compared to the neutral solution. While amygdala activity during presentation of a juice reward could be consistent with the PVLV model if this activity reflected learning of LVe weights that would later drive phasic dopamine to the associated CS, there was not activity in the amygdala for the rewarded CS, discounting this interpretation. Also, the amygdala activated during presentation of the neutral solution, but this activity was less strong than the activity to the juice reward. In addition, we saw activity in the ventral striatum during presentation of the rewarded CS, but only during the first conditioning block, and there was not significant activity in the accumbens during reward presentation. This was also inconsistent with the PVLV model's prediction that

PVi representations in the ventral striatum would shunt activity to an expected reward. Finally, we also examined whether the amount a particular reward was better or worse than expected activated different brain regions, and found that activity in mOFC and amygdala reflected the amount a reward was better than expected, while activity in IFG, insula and lateral OFC reflected the amount a reward was worse than expected.

Chapter 2

Biological data

Building on the described role of the regions involved in the computation of reward prediction error in the PVLV model, it is useful to provide an outline of the biological evidence supporting the functional role of these regions. We begin by describing the evidence supporting the role of the central nucleus of the amygdala (CeA) in representing associations of CSes with positive valence rewards, and how learning in the amygdala supports this function. Next, we describe the evidence for the role of the CeA in facilitating anticipatory responses to CSes and the role of projections from the CeA to VTA/SNc in controlling these responses. After this clear evidence for the role of the CeA in facilitating learned responses to positive valence, we summarize data on amygdala representations of positive and negative valence and how this relates to conditioned inhibition. Finally, we describe the evidence supporting a role of the lateral habenula in conditioned inhibition, and relate this to distinctions between the brain systems for positive and negative valence learning, providing evidence for segregated representations of positive and negative valence along with evidence for non-specific "saliency" representations to any motivationally relevant stimulus.

2.0.7 CeA and BLA

One of the crucial assumptions of PVLV is that the central nucleus of the amygdala (CeA) encodes positive associations of a CS with reward, and controls motivational responding through the release of dopamine at the time of CS presentation. Several pieces of evidence show that the central amygdala contains cells that have the necessary properties to support this crucial part of

reinforcement learning.

First, the CeA appears to be crucial for positive valence conditioning [3, 70]. In particular, there are multi-modal neurons in the amygdala that show sustained responses to primary rewards, and also learn to fire for an associated CS [70]. In addition, evidence for L-LTP in the form of c-fos early gene expression has been observed for positively valenced CSes after conditioning [38, 44]. The CeA also sends a direct excitatory projection to the midbrain dopamine nuclei, as well as the PPT and LHA, which also project to midbrain DA cells. Additional evidence for the connection between CeA and dopamine release comes in the form of electrophysiological studies showing the stimulation of CeA neurons can cause dopamine release. [77, 29].

The role of the CeA in conditioned orienting responses also supports the PVLV model. Orienting responses involve a rearing or a startle response to a novel CS, but the orienting response habituates after repeated exposure to a particular CS. However, after a CS is repeatedly paired with a US, animals tend to acquire the orienting response to the CS. The acquisition of such a conditioned orienting response crucially depends on the CeA, which is not crucial for the expression of the conditioned orienting response after acquisition. [26, 33].

Another set of behaviors that depend on an intact CeA are autoshaping responses, a set of behaviors which involve approach and orientation to a CS, as well as often some sort of manipulation of the CS. These responses often occur naturally at a low rate, but are reinforced through continuous pairings of the CS with the US, and also rely of the CeA for acquisition, but not expression. In the PVLV model, a potential explanation for these behaviors involves the CeA/LVe component of the model driving phasic dopamine burst at CS onset that reinforce these conditioned responses.

Both the dependence of these behaviors on an intact CeA & the timing of these responses at the time of the CS supports the potential role of the CeA in CS-triggered dopamine proposed by the PVLV model, distinct from the PV-triggered dopamine which mediates responses at the time of the US. [30]. In particular, the CeA can reinforce such behaviors because they are naturally triggered by the CS alone, and then become further reinforced by a phasic dopamine burst at the time of CS onset once a reward association is learned. Both of these pieces of data provide

important support for the role that the CeA plays in PVLV for dopamine release at the time of a conditioned stimulus, and illustrate the general role of the CeA for reinforcing any conditioned responses natively associated with CSes that are modulated by phasic dopamine release. [30].

While this evidence is convincing support for the role of the CeA in positively valenced learning, it remains unclear how the CeA encodes conditioned stimuli with less positive associations. This is an important question because the degree of overlap between the positive and negative valence representations in the CeA can shed light on the degree of overlap between the learning mechanisms for the two valences.

Interestingly, much recent data on the amygdala supports the existence of multiple types of representations in the amygdala, with some cells encoding representations selective to either positive or negative valence outcomes, while other cells responded similarly to outcomes of both valences. [5]. This outcome-specific coding was also influenced by expectation, though the selective cells only showed an expectation modulation for their preferred valence, while the non-selective cells showed an excitation to unexpected outcomes of either valence. A similar pattern occurs for positive and negative conditioned stimuli, with some in cells in basolateral amygdala showing selective firing for either positive and negative valence CSes, while other cells were excited by CSes of either valence [97]. Given that the amygdala shows this pattern of both valence-specific responses and valence general responses for any motivationally relevant stimuli, how would we predict that it responds to conditioned inhibitors?

Considering the case of conditioned inhibition, it would be useful to know how cues that signal omission of an outcome are processed, and how these representations relate to positive and negative valence cue representations. Several pieces of evidence suggest that the pattern of both segregated valence-specific and overlapping general arousal representations in the amygdala also applies to the conditioned inhibition paradigm.

For example, a recent study suggests that negatively valenced CSes may be represented in a separate population of neurons from cues that encode omission of a food US.[81]. While few studies have looked at conditioned inhibition in the positive valence case, we do know that CeA lesions do

not significantly impair the acquisition of conditioned inhibition, while they do impair orienting and autoshaping responses as discussed above. [32]. Recent work on safety signals, which are paired with a fear-conditioned cue to indicate omission of a negatively valenced outcome, also supported the view that the amygdala contains both segregated valence and general salience representations. In this negative valence version of conditioned inhibition, they found two different populations of neurons in the basolateral amygdala that fired to the fear cue paired with the safety signal, including some that were selective only to the fear+safety combination, and others which fired to the fear+safety cue and to a reward cue. [91]. This study is interesting because it suggests that there may be overlap between brain representations for omission of negative valence and reward in the amygdala, along with more specialized representations specific to either valence. While it is difficult to make conclusions without more evidence, these initial studies suggest that conditioned inhibition may operate differently depending on the valence of the US, which is not surprising given the relatively greater evolutionary importance of avoiding negatively valenced outcomes.

Given this general pattern of results in the amygdala, it becomes important to consider whether CeA and BLA play distinct roles in controlling different aspects of conditioned responding. While basolateral (BLA) and central amygdala (CeA) are generally thought to function in parallel, some evidence suggests that they regulate distinct aspects of conditioned responding. In particular, the BLA appears to mediate associations between stimuli and their sensory properties, while the CeA plays a role in connecting these stimuli with their affective and emotional properties. [1]. Similarly, lesion data in a devaluation paradigm suggested that the BLA may be more involved in mediating the effects of motivation on instrumental responses for specific USes, while the CeA is crucial for non-specific motivational properties[13].

Consistent with the discovery of general salience representations in basolateral amygdala, recent proposals have suggested that the basolateral amygdala may represent an associability signal [46, 85, 6], which dynamically controls learning rates for the CS such that the learning rate increases under surprising conditions and decreased when predictions are reliable. This type of representation is not specific to a particular valence. However, given the diversity of data reviewed above, it

seems clear that there are both valence-specific and general salience representations in both nuclei of the amygdala. In our view, this further enhances the importance of discovering the specific computational role of each of these representations. While we have a clear hypothesis about the role of the CeA projections to the VTA/SNc in the modulation of conditioned responding for positive valence CSes, the role of other types of representations in the amygdala, including those specialized for negative valence, remains to be fully specified. To better understand the amygdala circuitry for negative valence learning, it seems crucial to consider how the dopamine system and lower brainstem circuitry responds to negative outcomes and how this information relates to the representations in the amygdala.

2.0.8 RMTg and habenula

Recently, there has been a surge of research on the neural mechanisms behind the dopamine dip, or pause in tonic dopamine firing, that occurs when an expected reward does not appear. This is crucial for understanding the mechanisms behind the conditioned inhibition task because electrophysiological data shows that dopamine neurons respond to the combination of the CS+Inhibitor, and the Inhibitor viewed alone, with a dopamine dip [105].

Recent research has revealed that one of the brain areas driving the dopamine dip is the lateral habenula, located in the epithalamus near the pineal body, which responds to the occurrence of negatively valenced stimuli, as well as the omission of an expected reward. [51]. The anatomical circuit in the lateral habenula that responds to reward omissions receives projections from the globus pallidus internal (GPi) in the basal ganglia [34]. Information about rewards, encoded by distinct neural populations of reward positive neurons, excited by rewards, and reward negative neurons, excited by reward omissions, is then passed onto the lateral habenula and onto a GABAergic population of neurons in the RMTg or tail of the VTA. The RMTg is primarily composed of inhibitory GABAergic neurons, and activation of the RMTg leads to inhibition of dopamine neurons. Therefore, excitation of reward omission neurons in the RMTg due to inputs from the lateral habenula provides an explanatory mechanism for the dopamine dip observed when

an expected reward is omitted.

Along with coding for reward omissions, the lateral habenula and RMTg are also excited by the occurrence of negatively valenced USes, supported by data showing that the lateral habenula and RMTg activate during painful stimulation [98, 100, 52]. Along with showing this pattern of excitation to negatively valenced USes and reward omissions, the lateral habenula and RMTg also respond to CSes which predict these outcomes. [52, 35]. Therefore, the data supports the idea that the lateral habenula and RMTg, probably through their inhibitory effect on dopamine firing, can encode aversive teaching signals for reward omissions or negatively valenced stimuli and train associations for the CSes that predict these outcomes. Interestingly, the same circuitry in the lateral habenula and RMTg is also recruited during extinction learning tasks for positively valenced USes such as drug or food rewards. [50, 24].

Bringing this back to the conditioned inhibition task, we can take two pieces of evidence as potential support for a role for the LHb/RMTg circuit in conditioned inhibition- including the data pointing to involvement of the this circuit in learning about omission of a positively valenced reward, and the occurrence of a dopamine dip during presentation of the Inhibitor cue that predicts cancelation of an upcoming reward [105]. Given the lateral habenula encodes an excitatory signal for a CS predicting a reward omission that is then passed on the RMTg to inhibit dopamine neurons, we predict involvement of the lateral habenula in the conditioned inhibition task, particularly to learn about reward omission outcomes in the initial learning phase and use this to encode a negative association for the Inhibitor that predicts the reward omission, which will later be reflected as a dopamine dip to presentation of the Inhibitor.

2.0.9 Dopamine and negative valence areas

Unfortunately, much less is known about how the dopamine system responds to negative valence compared to positive valence. While the majority of "classical" dopamine neurons, meaning those with standard electrophysiological profiles in lateral VTA or SNc, respond with a dip to the occurrence of negatively valenced stimuli, such as an airpuff or footshock [12, 52], a distinct

population of dopamine neurons located mostly in the medial VTA respond with an excitation to negatively valenced stimuli. [41]. These neurons receive a projection from the RMTg, which integrates inputs from the lateral habenula and many other negatively valenced areas, and send to the medial PFC, while inhibiting dopamine release in lateral accumbens shell.

In addition, the offset of a negatively valenced stimulus also appears to be rewarding, and the same dopamine neurons that were inhibited during receipt of a shock show a phasic burst at the offset of the shock [10]. One of the major areas that the RMTg receives negatively valenced information from is the parabrachial nucleus (PBN), a nuclear complex in the brainstem which receives information about pain, as well as visceral and gustatory information from the nucleus of the solitary tract.[25]. The PBN also sends inputs to the central and basolateral amygdala, particularly to the lateral capsular division (CeC) or "nocioceptive amygdala". [18]. The PBN also sends nocioceptive information to dopamine neurons in the SNc and VTA; and electrical stimulation of PBN neurons results in a short-latency dip in dopamine neurons, likely due to this information passing through the RMTg which then inhibits the DA neurons, or a direct projection from the PBN to inhibitory interneurons. [12]. Further understanding of the distinct functional roles of dopamine neurons that are excited vs inhibited to negative valence outcomes, and their downstream projections to different populations of neurons in regions such as the amygdala and ventral striatum, will be crucial for understanding positive and negative learning.

2.0.10 Brain systems for positive and negative valence

In general, the brain has distinct circuits in several areas for positive and negative valence learning, though general mechanisms for both valences also exist. In particular, negative valence learning recruits a wide range of structures in the brainstem and amygdala. Recent electrophysiological work has shown that many of the brain areas involved in reward learning have distinct representations for positive and negative valence. This pattern of "reward positive" and "reward negative" neurons, which are sometimes located in distinct nuclei but may also be represented by a heterogenous population of neurons, has been shown in the RMTg [35], amygdala [97, 91],

ventromedial prefrontal cortex [58] and VTA/SNc [52, 41].

Bringing all of this data together, there are some answers to our central question about the brain representations for positive and negative valence. While it is clear that the central amygdala plays a crucial role in controlling motivational responding for both valences, a lot of data supports the idea that the brain has clearly segregated representations for positive and negative valence, particularly in the dopamine system, where there is a clear segregation between the lateral habenula system/RMTg for negative valence and reward omissions, and the classical dopamine neurons in the VTA/SNc involved in positive reinforcement.

However, this is not to say that there are no ambiguities in the dopamine representation of negative valence learning. As discussed previously, there are dopamine neurons excited by the occurrence of negative valence rewards, but the computational role of these neurons and their anatomical projections is poorly understood. Also, recent data suggests that midbrain dopamine nuclei may represent motivational value and motivational salience in parallel circuitry, with motivational value representations relying more on the VTA and lateral habenula pathway, while salience representations may rely more on SNc projections to dorsolateral prefrontal cortex and dorsal striatum[53, 11].

In general, the biological data suggests that while there are separate populations of neurons responding to positive and negative valence, many regions contain neurons of both types, such as the amygdala [97, 5] and vmPFC [58]. This may sometimes be accompanied by general salience representations non-selective to valence. While this might seem like a confusing picture, this overall heterogeneity of dopamine pathways actually underscores the importance of using experimental paradigms that make it possible to examine the role of different brain areas in the computation of the reward prediction error, such as conditioned inhibition.

Chapter 3

Methods

3.0.11 Behavioral methods

The conditioned inhibition paradigm began with a conditioning block that involved pairing a fractal image CS (A) with an orange juice reward, and a control CS, signaled by a different fractal image (B) was paired with a neutral solution. These two trial types used in this first conditioning are shown in Figure 3.2. Importantly, reward presentation is always probabilistic for the rewarded CS, such that it is followed by the juice reward 75% of the time and the neutral solution 25% of the time. In the second block, we begin the training trials for conditioned inhibition by showing the rewarded CS (A) with the Inhibitor (X), signaled by a different fractal image, together, and this CS-Inhibitor conjunction is always followed by the neutral solution, as shown in Figure 3.3. In this way, the Inhibitor is supposed to develop a negative connotation because it always cancels the reward the would have been received for the reward CS alone. We continue to show the reward CS (A) followed by the juice reward, and the control CS (B) in these conditioned inhibition blocks, as well as a control pair (BY), both followed by the neutral solution, as shown in Figure 3.4.

In the conditioned inhibition portion of the study (blocks 3-5), we show twelve trials each of the reward CS (A), CS-Inhibitor (AX), and control CSes (B,BY), and also show the Inhibitor (X) alone, as well as another neutral control (Y), but only show 4 trials of these per block to avoid creating a strong conditioned response to the Inhibitor alone. In the final block, the Inhibitor (X) and neutral control (Y) are paired with the juice reward with a 75% reward, 25% no reward contingency, based on previous work that used a prediction error signal to the Inhibitor when paired

with reward to assess whether it formed inhibitory properties [105]. The number of trials for each CS in each block and reward probability for each CS in each block is shown in Figure 3.1.

Figure 3.1: This table illustrates the number of trials in the study in each block and the reward contingencies in each block.

All trials	CS (A)	Control (B)	CS+Inhib (AX)	Control (BY)	Inhibitor (X)	Control 2 (Y)
Block 1	24 (75% R, 25% NR)	24 (100% NR)				
Block 2	8 (75% R, 25% NR)	8 (100% NR)	8 (100% NR)	8 (100% NR)		
Blocks 3-5	12 (75% R, 25% NR)	12 (100% NR)	12 (100% NR)	12 (100% NR)	4 (100% NR)	4 (100% NR)
Block 6	24 (75% R, 25% NR)	24 (75% R, 25% NR)				

Figure 3.2: The rewarded CS(A) is paired with an orange juice reward with a 75%/25% contingency, while the neutral CS (B) is always paired with the neutral solution. For all stimuli, the CS stays onscreen for 2 seconds and is immediately followed by the US.

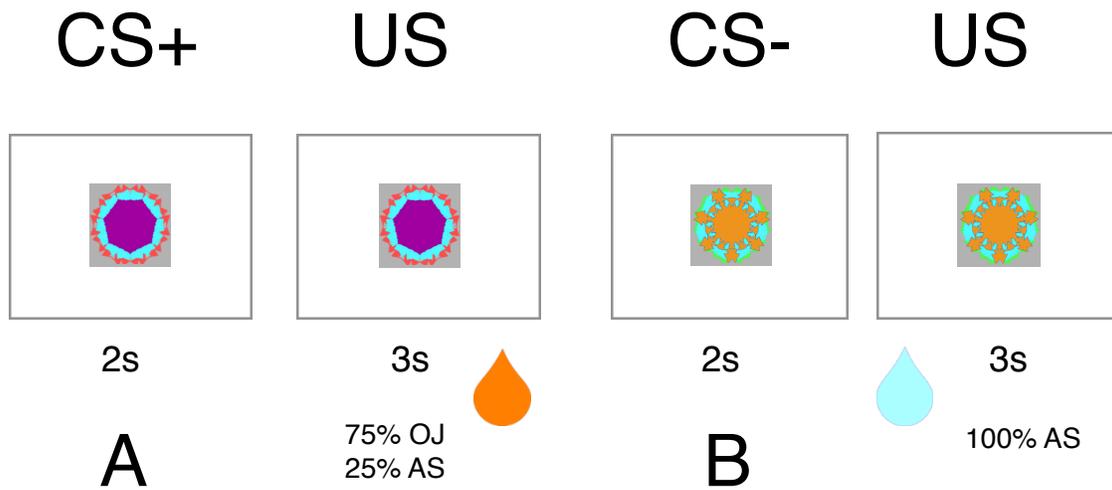


Figure 3.3: After the conditioning block, we continue to show the rewarded CS and control, but introduce the CS-Inhibitor pairing. When the inhibitor is shown on the screen with the rewarded CS, it causes cancellation of the expected reward for the CS+, and instead is followed by the neutral solution. The inhibitor is also shown alone, paired with the neutral solution.

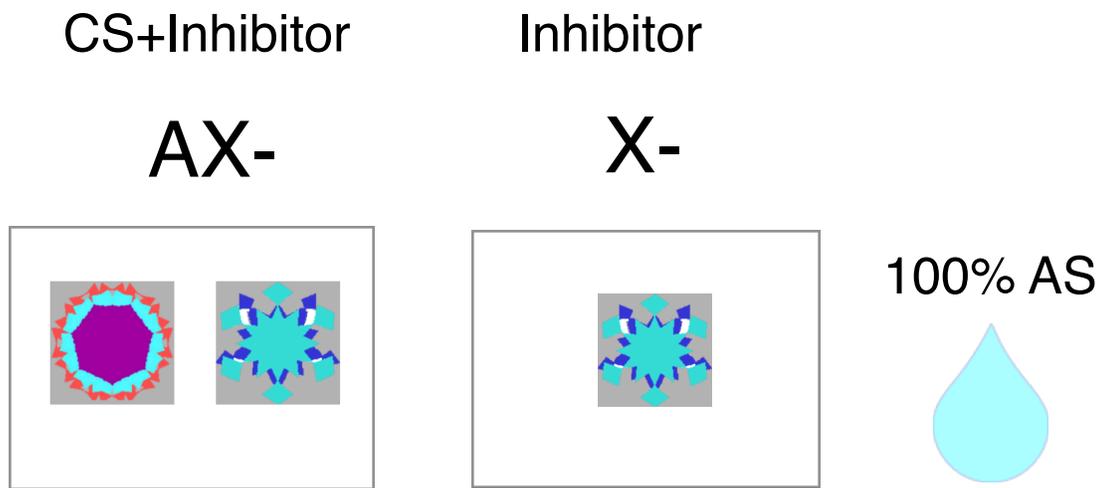
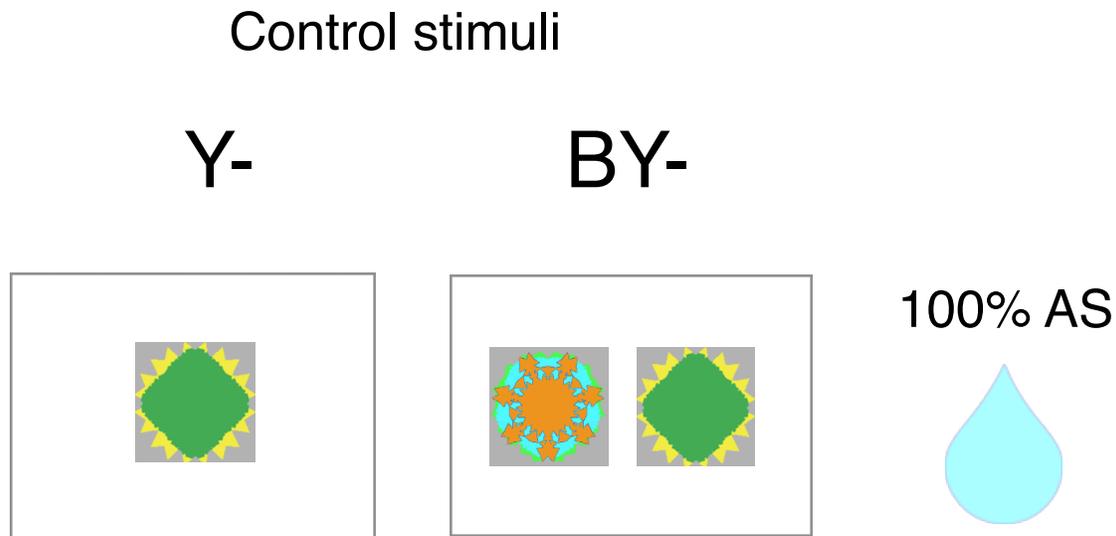


Figure 3.4: These control stimuli are also shown during the conditioned inhibition block, and are always paired with the neutral solution.



There were several features of the design that were motivated by careful consideration of the learning problem. For example, we wanted to keep the duration between the conditioned stimulus (CS) and juice reward (US) consistent because there is evidence the striatum responds to temporal prediction errors [55]. In addition, we chose a delay conditioning paradigm, where the CS remains onscreen while the US is delivered, because there has been considerable evidence that trace conditioning, which involves showing a CS that is removed before reward is delivered, depends on the integrity of the prefrontal cortex and hippocampus [40, 59]. The PVLV model can account for delay conditioning, but not for trace conditioning, because learning about the CS only can occur at the time of the US [71, 30].

3.0.12 Data analysis

3.0.13 Analysis methods

The imaging data was first preprocessed using standard preprocessing methods, then first-level models were specified for each subject in SPM8. In the first analysis, we focused on the

effects of juice reward compared to the neutral solution, and thus created one regressor to collapse across all CS types and separate regressors for juice presentation and presentation of the neutral solution. In the CS Type analysis, we created one regressor for each CS Type, concatenating the data across several runs, and created one regressor for all presentations of the neutral solution. However, the events following the rewarded US A, were modeled with separate regressors, with one for US A events that were rewarded, and a separate regressor for US A events that were not rewarded. Unless otherwise stated, all analyses involved the full set of 19 subjects. Two of these subjects were missing a single run of conditioned inhibition due to technical difficulties with juice presentation. To run the ROI analyses, we extracted the mean activation value for each subject in the voxels ROI of interest, and compared effects across different contrasts with standard statistical tests.

In addition, after running the first analysis on a subset of subjects with less movement, we ran a sanity check model, which involved comparing the brain activity for all swallowing trials (either orange juice or artificial saliva) to baseline trials. This analysis showed significant activations in white matter tracts and ventricles for any swallowing trials, and ICA analysis also showed some potential whole-brain artifacts. As many of the regions we were interested in are in regions that already have low signal to noise ratios, particularly the amygdala and nucleus accumbens, it was important to see whether the results survived controlling for this potential source of noise. To control for this, we specified an additional nuisance regressor generated from extracting the first 5 principle components from the white matter and ventricles, segmented based on each subject's individual structural, along with movement parameters taking into account movement in the x,y, and z directions, and dummy spike regressors controlling for unusual images. All results reported are including this additional regressor to control for potential confounds due to swallowing.

3.0.13.1 Better and worse than expected analysis

To model events that were better or worse than expected, we focused on the rewarded CS (A), since it was the only CS with probabilistic variability in payoff rate (the control stimuli and

inhibitor were consistently paired with the neutral solution). To calculate the amount that each event was better or worse than expected, we assumed that expectations would follow from the exact probabilities of reward. Therefore, for the first conditioning block, each time the rewarded CS was followed by a reward it was coded with a 1. Events in the first conditioning block that were not rewarded were not considered better or worse than expected. After the first conditioning block, we assumed that the reward expectation for the rewarded CS was .75, following from the actual probability of reward. Therefore, any US event with orange juice was considered .25 units better than expected, and any US event with neutral solution following the rewarded CS was considered .75 units worse than expected. The interpretation of these results is important, because both the better and worse than expected analyses involved fitting a parameter to the BOLD data at the time of US presentation calculating how much better or worse than particular trial was than the reward expectation, to determine whether there was a linear increase in BOLD activity as the amount better/worse than expected increased. To avoid collinearity issues, the worse than expected regressor was orthogonalized compared to the better than expected regressor.

3.0.14 Computational model

3.0.14.1 Model overview

As described previously, a major feature of the PVLV model are LVe representations, which capture the learned value of the CS and influences dopamine release in the CeA for a rewarded CS. This is well supported by the amygdala data showing that learning about positively valenced USes, based on these cells getting excited by the CS and US, occurs in the CeA, and data showing that conditioned responses at the time of CS crucially involve the CeA. The other important part of the PVLV model is the PV system, which includes an excitatory component, PVe, which provide low-level representations of primary rewards that drive dopamine firing, and an inhibitory component, PV_i, representing inhibitory neurons in the ventral striatum that inhibit dopamine release in the VTA for expected rewards. At the time of reward, phasic dopamine release is accounted for by the

difference between PVi and PVe, closely fitting the data showing that dopamine neurons calculate the difference between an expected and received reward. [93]

Finally, the PVLV model provides an account of the circuitry involved in control of dopamine release for positive and negative valence USes and CSes, and incorporates an explanatory mechanism for the dopamine dip that occurs when an expected reward is not received, or when a negative valence US occurs, in terms of the habenula/RMTg projection to the SNc/VTA which inhibits dopamine release, by causing a pause in the tonic dopamine firing reflected as a dip[8, 35]. By drawing on these learning systems for positive and negative valence, the PVLV model can provide an explanation of the computational and biological mechanisms involved in conditioned inhibition.

3.0.14.2 Learning of conditioned inhibition in PVLV

In the first conditioning block, the rewarded CS gets a positive LVe value, due to the LVe/CeA system learning that the CS is reliably paired with reward. The rewarded CS also increases the PVi weights for the CS after multiple pairings with reward, which have an inhibitory effect on neurons in the VTA/SNc. Both the LVe and PVi weights learn at the time of reward presentation, but influence activity during different parts of the the conditioning procedure. Increased LVe weights increase dopamine release at the time of CS presentation, while the PVi weights have an inhibitory effect on dopamine release at the time of an expected reward, and both of these mechanisms capture the observed result that dopamine bursts move from the time of the reward to the time of CS presentation for an expected reward. See Appendix B for an illustration of increases in the LVe and PVi (Figure B.0.27 and Figure B.0.28) weights over repeated reward pairings and an illustration of how activity in the VTA (Figure B.0.29) for decreases at the time of the US and increases during CS presentation for expected reward.

After the first conditioning block, we begin the conditioned inhibition procedure. While the rewarded CS has been paired with reward previously, the conditioned inhibitor has never been seen before. As PVLV sums across LVe values for all stimuli to calculate for the overall reward expectation at the time of the CS, it sums a strong positive LVe value for the rewarded CS with

no L_{Ve} value for the inhibitor, creating an overall positive expectation for the CS+Inhibitor.

When the reward fails to appear after a CS+Inhibitor presentation, this causes a dopamine dip. In PVLV, two distinct mechanisms can explain the dopamine dip. First of all, the PVi weights for the rewarded CS inhibit dopamine release at the time reward would be expected. On top of this shunting inhibition, the lateral habenula is excited by the reward omission, and sends additional inhibitory drive to dopamine neurons through the RMTg, causing the dip. Once the dip occurs, the PVLV model learns a negatively valenced L_{Ve} weight for inhibitor. Our model places these negatively valenced LV signals in the habenula, based on data showing that the habenula learns about CSes that signal omission of negative valence outcome. Over repeated pairings of the CS+Inhibitor with reward omission, the Inhibitor begins to acquire a negative value, while the reward association for the CS stays positive based on the interleaving of CS+Reward trials. Figure B.0.30 in Appendix B illustrates how the model acquires positive and negative learned values for the CS and Inhibitor.

Chapter 4

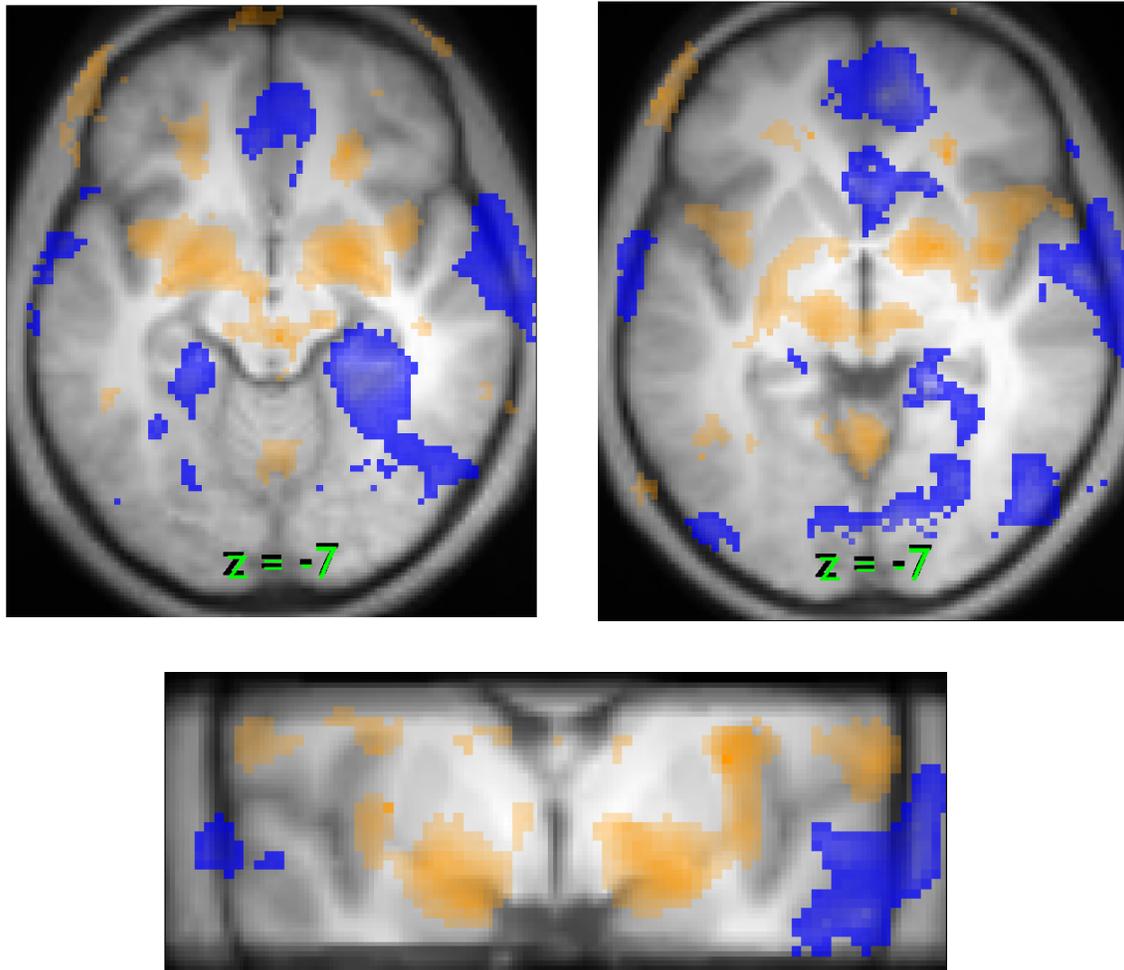
Results

4.0.15 Juice reward analysis

In particular, we were interested in comparing the effects of juice reward presentation with presentation of the neutral solution to see whether this contrast activated crucial regions in PVLV that represent reward value, particularly the SNc/VTA and lateral hypothalamus. In addition, we predicted activity in the regions in PVLV that adjust their learned weights during reward presentation, including the central amygdala and ventral striatum.

As shown in Figure 4.1, the contrast of juice reward - artificial saliva activated the insula, lateral OFC as well as the ventral pallidum and putamen, caudate nucleus, and a midbrain region corresponding to the VTA/SN ROI. In addition, the SN also showed a significant activation in an ROI analysis focusing on the juice-artificial saliva contrast ($p=.0420$, $t=2.18$). There was not significant activity in the nucleus accumbens or lateral hypothalamus.

Figure 4.1: The contrast of juice reward(orange) and the neutral solution(blue) showed that juice activated several regions including the amygdala, insula, pallidum, putamen, a midbrain area consistent with the SN/VTA and medial OFC (BA11). ($p < .05$, FDR).

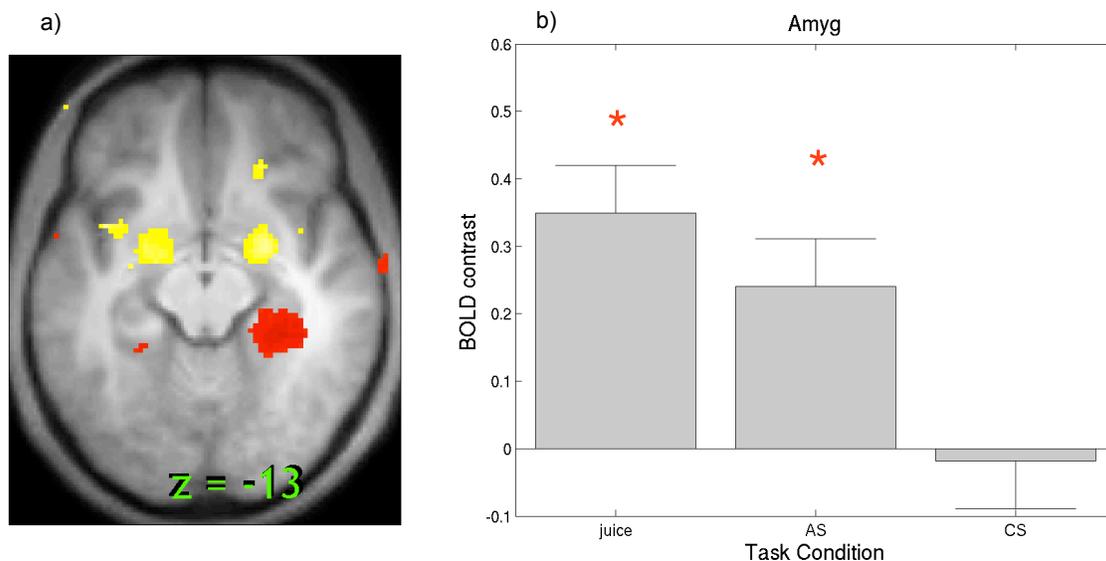


4.0.15.1 Juice reward: ROI analysis

To shed light on the question of how positive outcomes were represented in the amygdala, and see whether the central amygdala showed activity in response to the US is crucial for training learned value representation in PVLV, we further compared brain activity in the amygdala during presentation of the juice reward and neutral solution. In a first analysis, we compared the mean activation values for juice and the neutral solution in the amygdala, and found that both juice ($p = .0003$, $t = 4.54$) and the neutral solution ($p = .0012$, $t = 3.83$) significantly activated bilateral amygdala,

as shown in Figure 4.2b. However, the contrast of juice-neutral solution also activated the amygdala ($p=.0015$, $t=3.75$), indicating that amygdala showed greater activity for the rewarded solution than the neutral solution, as shown in Figure 4.2a. Similarly to the amygdala, the insula also showed activations for both juice and the neutral solution, however there were significant activations for the Juice-neutral solution contrast in the insula ($p=.0013$, $t=3.81$), and the OFC ($p=.029$, $t=2.35$), indicating that these regions had greater overall activity for the juice reward than the neutral solution. In addition, we ran an additional ROI analysis focusing on the central nucleus of the amygdala, and found a similar pattern such both juice and artificial saliva significantly activated the central amygdala, but the contrast of Juice-artificial saliva also had a significant effect. ($p = .000057$, $t=5.7$), meaning that the overall activity was stronger in the central amygdala for juice reward presentation than for the neutral solution.

Figure 4.2: (a) There was more activity in the amygdala for juice than for the neutral solution, as there was significant activity for a contrast of juice with the neutral solution ($p < .05$, FDR). (b) However, the amygdala ROI showed significant activity for both juice presentation and presentation of the neutral solution.

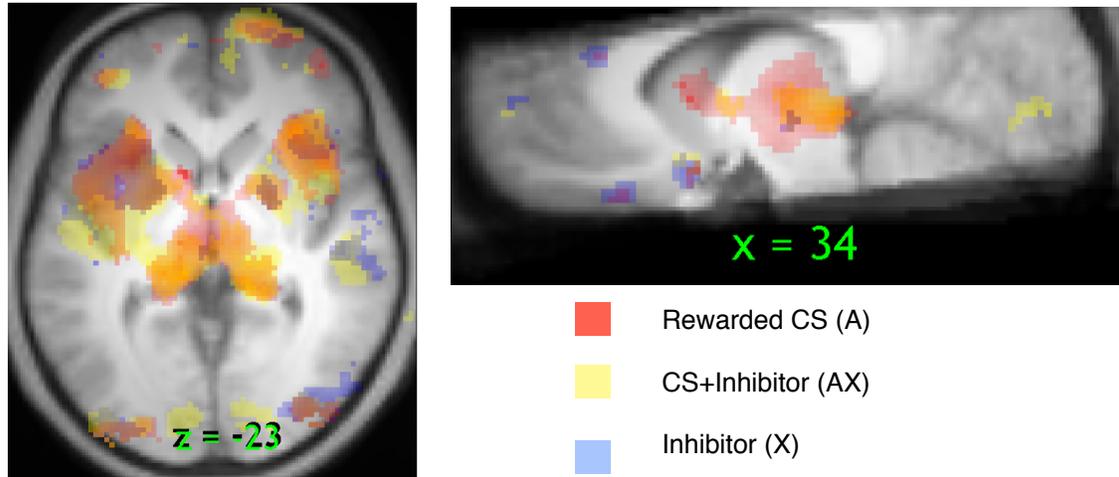


4.0.16 CS Type analysis

We were also interested in comparing the brain activity for the rewarded CS with the brain activity for the Inhibitor, to assess similarities and differences between the processing of positive and negative valence CSes, especially to see whether the positive CS was primarily represented in the central amygdala and the CS with a negative association, the Inhibitor, was presented in the lateral habenula. In addition, we were interested in comparing this activity with the activity for the CS+Inhibitor conjunction, to see how the presence of the inhibitor influences the processing of the rewarded CS, and whether it fully inhibits all reward-predictive activity. We were also interested in examining brain activity in the basal ganglia and prefrontal cortex, particularly vmPFC and OFC, because these regions also play a role in positive and negative valence representations.

As shown in Figure 4.3, which portrays activity for the rewarded CS, CS+Inhibitor and Inhibitor, the different CS types had some distinct effects across the brain but activated many of the same regions. In general, several regions of the PFC activated non-selectively to several of the CS Types. For example, the subcallosal cortex around BA 25, showed activation for the rewarded CS, inhibitor and CS+Inhibitor, as did frontopolar cortex (BA10). In addition, BA11 (OFC) also showed activity for the rewarded CS and CS+Inhibitor. There was activation for the CS, CS+Inhibitor and Inhibitor alone in lateral OFC (BA47). The pallidum, putamen and insula were also significantly activated in all three conditions across subjects. The left caudate also activated for the rewarded CS, CS+Inhibitor and inhibitor. Surprisingly, there was also significant activity in the lateral habenula ROI for all of the CS conditions, but no differential activity between the CS types. There was also no significant activity in the amygdala for any of the CS types.

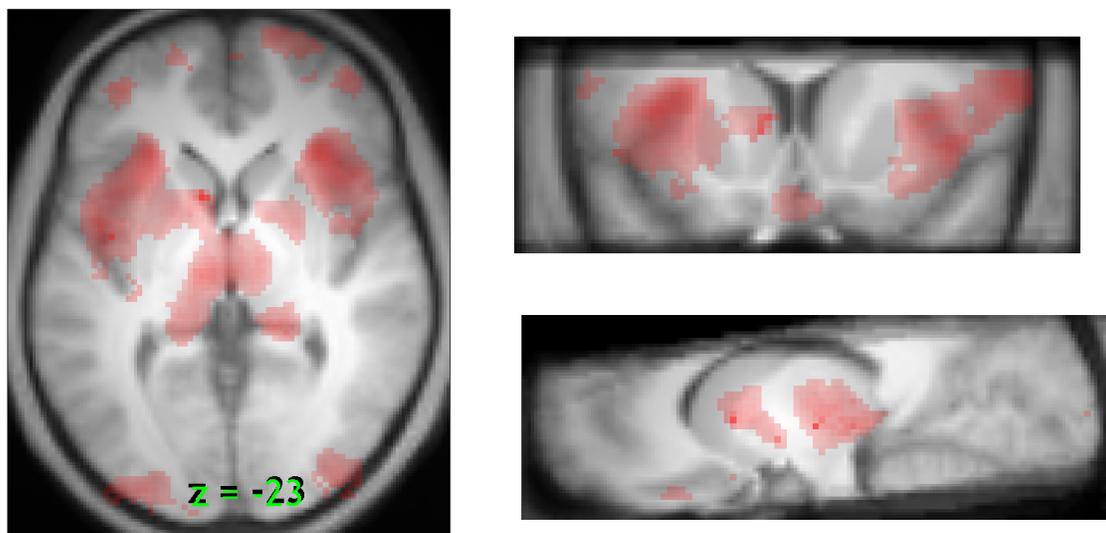
Figure 4.3: Several stimuli, including CS A (red), CS AX (yellow) and CS X (blue) showed significant overlap in vmPFC, BA 25, lateral OFC and insula at $p < .005$, uncorrected. There was also substantial overlap in parts of the BG, including the palladium and putamen.



4.0.17 Rewarded CS activity

We also wanted to look more specifically at the brain activity for the rewarded CS, particularly to see whether there was activity in the CeA and VTA/SNc as these regions are crucial for representing CS associations with reward and releasing dopamine due to this association. We were also interested to see whether there were high level representations of the learned value of the CS in prefrontal regions that encode value representations, such as the OFC and vmPFC, and other parts of the basal ganglia. As illustrated in Figure 4.4, the rewarded CS activated regions of the basal ganglia, including the palladium and putamen, the insula, and prefrontal regions such as lateral OFC and vmPFC.

Figure 4.4: The rewarded CS(A) showed activity in a wide range of regions, including the caudate, pallidum and putamen. There was also activity for the rewarded CS in lateral OFC and medial OFC, as well as subgenual ACC (BA 25).



4.0.17.1 ROI analysis

Using a ROI analysis, we were interested in further examine brain activity for the rewarded CS in the central amygdala and SNc/VTA due to this region's role in PVLV in encoding associations of CSes with reward and driving dopamine firing to the rewarded CS. However, there was no significant activity in the amygdala or central amygdala for any of the CS conditions. However, the substantia nigra and VTA both showed activity for the rewarded CS. There were significant activations in the VTA for the rewarded CS after conditioning compared to fixation ($p=.0019$, $t=3.65$), as well as during the conditioning block ($p=.0066$, $t=3.07$), as shown in Figure 4.5b which compares mean activity in a VTA ROI across conditions. The SNc also showed a similar profile of activations to the rewarded CS in the conditioning block and after conditioning, but was only significant after conditioning ($p=.0053$, $t=3.16$). Similarly, the SNc/VTA showed significant activity for the rewarded CS after conditioning when thresholded at $FDR < .05$, as shown in Figure 4.6.

Figure 4.5: The rewarded CS(A) selectively activated the VTA/SN during the conditioning block and after conditioning. (a) VTA activity plotted at $p < .005$, FDR corrected showing activity for CS A after conditioning. (b) There was significant activity in the VTA roi during the first conditioning block ($p = .0019$, $t = 3.65$) and for the rewarded CS after conditioning ($p = .0019$, $t = 3.65$).

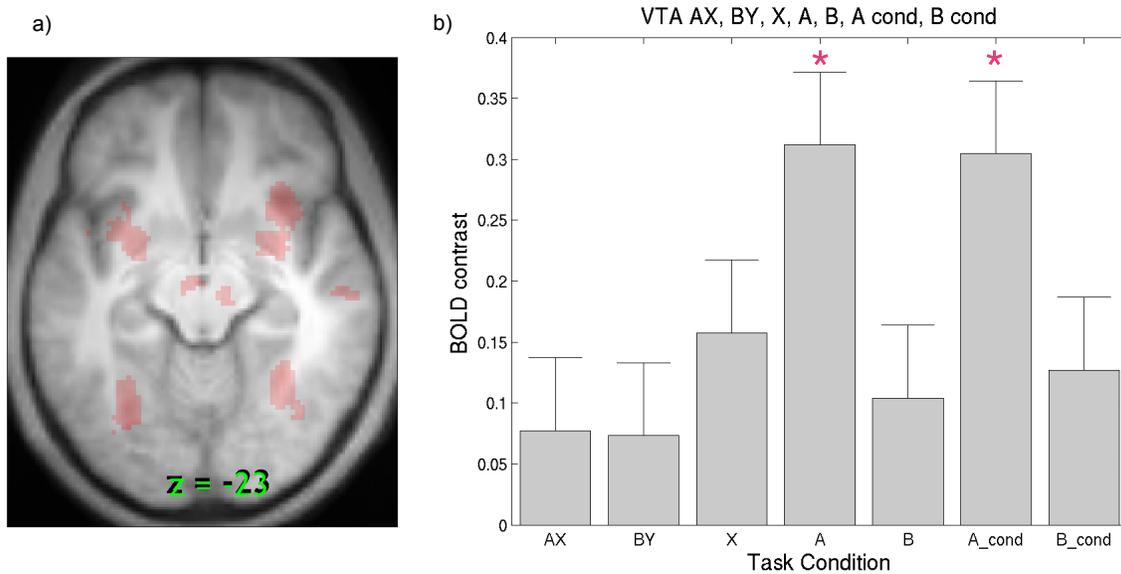
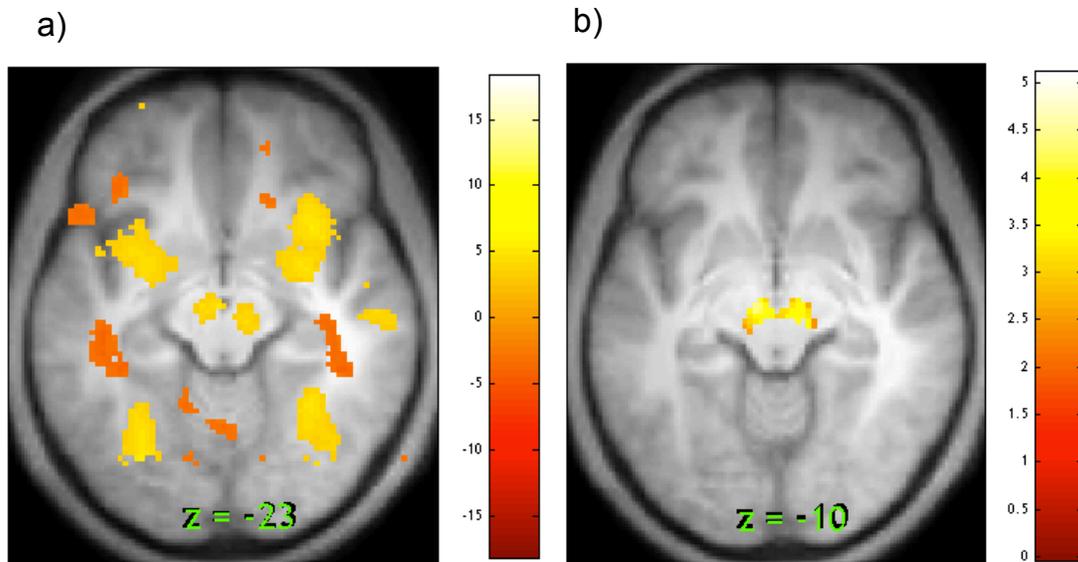
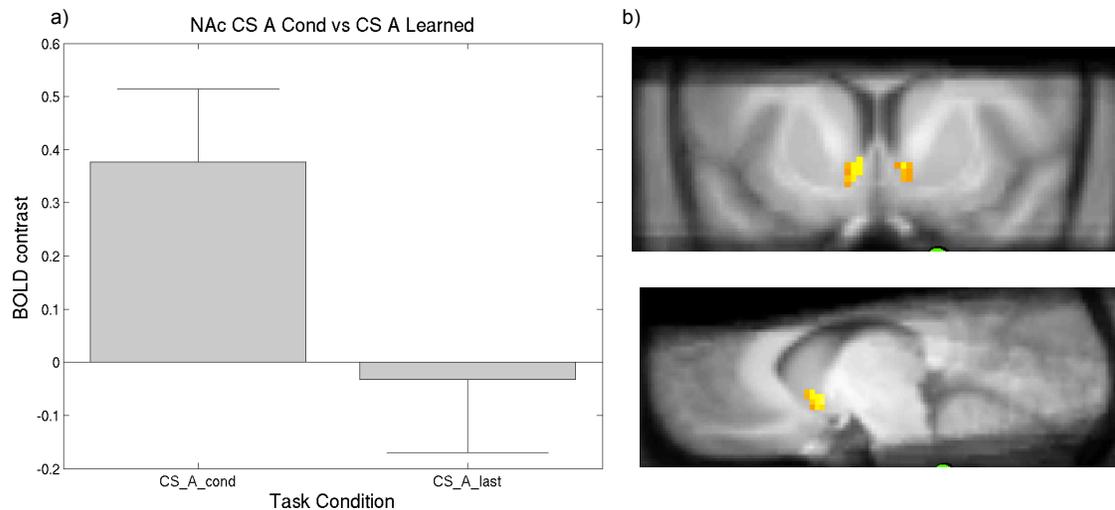


Figure 4.6: Brain activity in the SN/VTA for the rewarded CS, showing (a) full brain activation, and (b) viewed on a brain masked with an SN/VTA roi, both thresholded at ($p < .05$, FDR).



In addition, we wanted to see whether we noticed the initial signs of activity in the ventral striatum, where inhibitory PVi neurons are hypothesized to show ramping activity that will eventually inhibit the VTA/SNc at the time of reward presentation. However, as shown in Figure 4.7, there was only significant mean activation in the nucleus accumbens during the first conditioning block.

Figure 4.7: (a) The rewarded CS(A) selectively activated the NAc during the conditioning block. This activity was selective to the first conditioning block and did not occur to the rewarded CS after the conditioning block. This was reflected by a significant mean in the NAc ROI only for the first conditioning block ($p=.0237$, $t=2.485$). (b) Brain activity to the rewarded CS during the conditioning block, masked with a NAc ROI ($p<.05$, FDR)



4.0.18 Expectation analysis

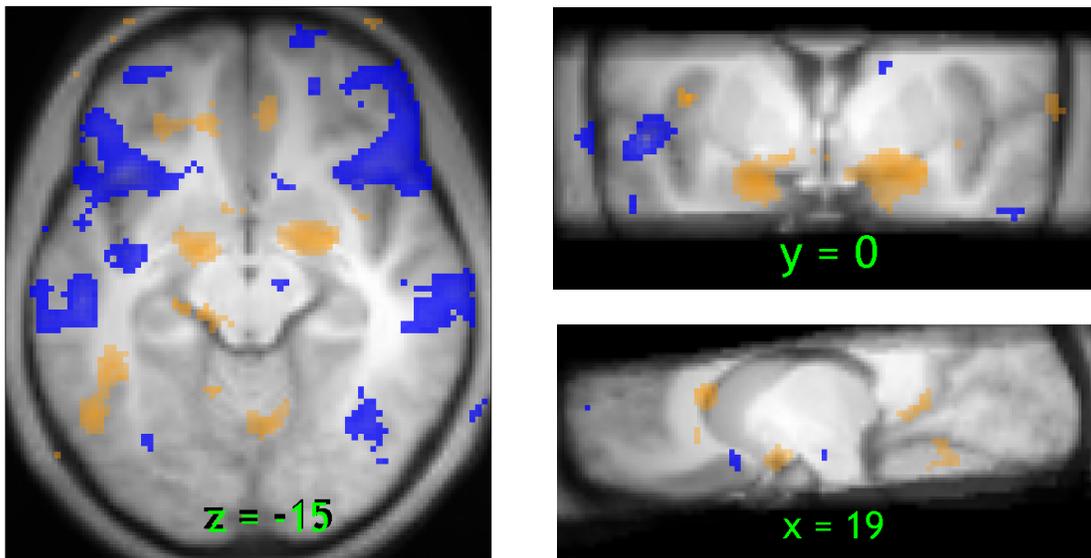
In addition, we were interested in doing an analysis of reward presentation events that were better or worse than expected to see if the brain areas involved in processing better and worse than expected rewards activated the regions specified by the PVLV model, particularly the lateral habenula for worse than expected trials and the SNc/VTA, accumbens and amygdala for better than expected trials.

To model events that were better or worse than expected, we focused on the rewarded CS (A), since it was the only CS with probabilistic variability in payoff rate (the control stimuli and inhibitor were consistently paired with the neutral solution). As described in Section 3.0.13.1, we fit a parameter to the brain data describing the amount better or worse the current stimulus was than the reward expectation, calculated based on the overall probability of reward. The interpretation of these results is important, because both the better and worse than expected analyses involved

fitting a parameter to the BOLD data at the time of US presentation calculating how much better or worse than particular trial was than the reward expectation, to determine whether there was a linear increase in BOLD activity as the amount better or worse than expected increased.

As shown in Figure 4.8, which compares brain activity for worse than expected trials with better than expected trials, we found that the better than expected regressor activated the medial OFC and amygdala. In contrast, the worse than expected contrast activated the lateral OFC and insula. However, an ROI analysis of the amygdala found that both the amount better and worse than expected activated the amygdala significantly. But, the better-worse than expected contrast also significantly activated the amygdala ($p = .049$, $t = 2.12$), and the central amygdala ROI ($p = .0157$, $t=2.68$), meaning that there was more activity in the amygdala that reflected better than expected reward outcomes than reflecting worse than expected outcomes.

Figure 4.8: After controlling for the main effects of US presentation, the amount a CS was better than expected activated medial OFC and amygdala, while worse than expected outcomes activated lateral OFC, insula and IFG.



4.0.19 Prediction errors in SN/VTA and habenula

Given our predictions of the regions involved in encoding the dopamine dip for an unexpected omission, as well as the well-established results showing that dopamine release transitions from the time of the US to the time of the CS after learning, we were interested to see whether activity in the SNc/VTA and lateral habenula was consistent with these predictions. In particular, we predicted that the lateral habenula would show increased activity when an expected reward following the rewarded CS was omitted, while the SNc/VTA would show reduced activity driven by this inhibitory signal in the habenula.

However, while there was significant activation in the SN for the juice-artificial saliva contrast, the analyses that looked at the signal in more detail were somewhat unclear. While a contrast that collapsed across all reward presentation trials for the rewarded stimulus showed significant activity, more specific analyses looking at whether a particular trial following the rewarded CS was rewarded or not did not show significant effects.

In an analysis collapsing across the entire experiment, we would only expect increasing activity in the SNc/VTA the first blocks during US presentation following the rewarded CS. After that, subjects should have formed an association between the CS and US, and increased activity should occur in the dopaminergic nuclei during presentation of the CS, which is consistent with the results that were observed.

However, the trials where the rewarded CS was unexpectedly followed by no reward, should reflect negatively valenced prediction errors when collapsing across the entire experiment. However, we did not see significant activity in the lateral habenula during these trials, and there was also no activity in the VTA or SNc, as might have been predicted by a temporal differences account that does not distinguish positive and negatively valenced prediction errors.

Similarly, we also did not see clear activity signatures in the SNc/VTA or lateral habenula for the parametric analysis looking at better or worse than expected trials. As a reminder, this parametric analysis involved calculating a level of reward expectation for the rewarded CS (A),

and then fitting a parameter to the BOLD data during all of the US A trials determined by the amount a particular US event was better or worse than expected. There was not any significant activity in the SN/VTA that increased parametrically as the amount that a particular trial was better or worse than expected increased, and there was also a lack of significant activity in the lateral habenula for worse than expected trials.

Chapter 5

Discussion

The main aim of the experiment involved further understanding of the brain areas involved in reward learning. By choosing the conditioned inhibition task, we wanted to add to existing research, building beyond reward prediction error accounts which presume to capture reward learning within a single reward prediction error parameter fitted to multiple regions of the brain. In addition, we hoped to decrease the widespread influence of dopamine projections on the brain areas involved in reward learning, allowing us to evaluate hypotheses about the computational role of different brain areas involved in the calculation of the RPE and interpret our observed activations more clearly.

5.0.20 SN/VTA and lateral habenula

The results observed in the SNc/VTA showed a highly selective pattern for the rewarded CS both during and after the conditioning period. Activations of the SN/VTA for a rewarded CS is a replication of previous studies showing that anticipation of a pleasant CS activated the SN/VTA [69]. However, there was no significant activity in the SNc/VTA for the CS paired with the Inhibitor, Inhibitor, or any of the control CSes. This result is interesting in light of theories of conditioned inhibition because it suggests that the addition of the Inhibitor in the CS+Inhibitor trials must have been able to inhibit the brain activity in the SNc/VTA for the rewarded CS. This is actually consistent with the theoretical account of conditioned inhibition proposed by the PVLV model, because it predicts that the excitatory signal for the Inhibitor cue that predicts reward omission will excite the lateral habenula, leading to an inhibition of the dopamine neurons. The

observed activity may also be consistent with the findings in the monkey electrophysiology study [105], which found a short-latency burst followed by dip to the CS+Inhibitor stimulus, but it is difficult to tell whether a lack of activity would be the expected BOLD signature of the latter dopamine dip, given the limited temporal resolution of fMRI and the lack of other data about dopamine levels in the SNc/VTA.

While the above results show that the SNc/VTA showed activity selective to the rewarded CS, the results in the SNc/VTA during reward presentation are less clear. We observed activity in an SNc ROI for the contrast comparing the juice reward presentation to the neutral solution, consistent with other studies finding increased activity in this region for rewarding stimuli [17]. In addition, the increased dopamine activity during presentation of the reward CS over the entire experiment is consistent with the observed result that dopamine cells fire to the CS that predicts reward presentation after learning [93].

However, when we further divided the activity during juice reward presentation following the reward CS across the entire experiment into rewarded and non-rewarded trials we did not see any activity in the SNc/VTA for either condition. However, this may not be that inconsistent with our proposed model after all since the CS association with reward was learned after the first few trials, and activity to an unexpected omission was expected to result in a dopamine dip rather than an excitation.

In addition, we did not find activity that was consistent with our proposed account of the role of the habenula in learning about reward omission, as there was no differentially significant activity in the lateral habenula for the Inhibitor and CS+Inhibitor, as all stimuli showed significant activations. We also did not find activity in the lateral habenula when the rewarded CS was unexpectedly followed by no reward. Similarly, we also did not see clear activity signatures in the SNc/VTA for the parametric analysis looking at better or worse than expected trials, and found that neither the lateral habenula or SNc/VTA increased their activity in a graded way depending on how much worse or better than expected a given outcome was.

5.0.21 Amygdala

Interestingly, the activations in the amygdala observed seemed to be selective to the US presentation time period, and it showed relatively little significant activity during the CS presentation period, including no significant differences between the different types of CSes. Also, even though the amygdala showed significant activity for both the presentation of the juice reward and the neutral solution, there was significant activity in the contrast comparing the juice reward to the neutral solution, showing that the juice activated the amygdala more than the neutral solution.

The lack of amygdala activity to the rewarded CS contradicts the predictions of the PVLV model, which predicted that the central amygdala was crucial to drive anticipatory activity for the rewarded CS. However, we did see activity in the SNc/VTA for the rewarded CS, indicating that there was an excitatory dopamine response to the reward anticipation. Since the CeA was thought to be crucial to drive this excitatory dopamine response, this set of results is inconsistent with the predictions of the PVLV model.

However, the observed amygdala activity to presentation of the juice reward was not inconsistent with the PVLV model, which predicts that changes in the LVe weights would occur each time the reward is presented, especially when the association between the CS and reward is first being learned. However, we also observed activity in the amygdala during the presentation of the neutral solution, which we did not expect because the neutral solution was thought to be affectively neutral. While many have shown that the amygdala activates for positive and negatively valenced events, particularly if they are motivationally salient in light of current goals [15, 56], amygdala activity is not expected for neutral stimuli. Given the amygdala activated to both solutions, this calls into question the affective valence of the neutral solution. Consistent with the proposed view of the amygdala in the detection of motivational salience, several studies have found that the amygdala also tracks behavioral relevance, showing increased activity for stimuli relevant to a behavioral response, such as go stimuli in a go-no-go task or goal-relevant stimuli [90, 73, 72]. Given the importance of responding to the reward presentation trials, which required swallowing the juice reward,

it is possible that these trials were more behaviorally relevant for subjects than the cue presentation trials. Another related possibility could be that the presentation of either solution activated the amygdala more than the CS presentation alone because the amygdala plays an important role in conveying visceral sensory inputs, primarily received from ascending sensory taste inputs in the brainstem and gustatory thalamus [80] to limbic regions in the orbital and medial prefrontal cortex [79] which integrate this visceral information. In addition, if the neutral solution was not affectively neutral and instead slightly aversive, this would be consistent with many studies which found that the amygdala activates to appetitive and aversive stimuli [5, 91, 56].

5.0.22 Basal ganglia

The striatal regions identified that showed activity to juice reward presentation, particularly the putamen and the anteroventral striatum, incorporating the ventral caudate, putamen, have been activated in several studies for the receipt of juice rewards. [56, 87, 21]. In a previous study, the ventral pallidum, along with the ventral striatum and hypothalamus, showed increased connectivity with the left insula during receipt of juice compared with a negative valence stimulus.[87].

We also observed increased activity in the nucleus accumbens for the presentation of the rewarded CS, but this only occurred in the first conditioning block. One potential explanation for this could be that the accumbens only activates when the association of the reward CS with reward is strongest in the first conditioning block, before the reward CS begins being paired with the inhibitor in the conditioned inhibition blocks. However, we continue to show the rewarded CS paired with reward in the conditioned inhibition blocks and the reward CS did drive activity in the SNc/VTA after conditioning, so the lack of activity in the accumbens after the first block is somewhat surprising. Other studies have found that the OFC and amygdala show habituation to taste stimuli over the course of the experiment[69, 64] so it could also be possible that this signal is only strong enough to drive anticipatory activity in the accumbens during the first block, but not later blocks. The general finding that the accumbens activity reflects associations of a CS with reward is one of the most prominent results in reward learning, and supported by numerous

studies[39, 48, 55, 45, 69].

5.0.23 Prefrontal Cortex and Insula

There have been several different proposals for how the brain processes better and worse than expected outcomes. In several studies of monetary reward and loss, medial OFC has been found to correlate with positive outcome anticipations and correct trials, while lateral OFC correlated with the anticipation of negative outcomes. [67, 48]. However, this distinction between lateral and medial OFC has recently been challenged by recent studies showing that positive and negatively valenced outcomes converge on the same OFC neurons. There appear to be dissociable effects of mOFC and IOFC lesions, such that mOFC lesions impaired performance when relative values were harder to distinguish, while IOFC lesions appear to be particularly important for credit assignment, as IOFC lesioned animals had difficulties attributing value to the stimulus responsible for the outcome compared with recently seen stimuli.[88].

In the conditioned inhibition study, the better and worse than expected trials occurred only after the rewarded CS, given it was the only CS in the study with probabilistic variation in reward delivery (all control stimuli were always followed by the neutral solution). The differential activation of IOFC for worse than expected rewards (neutral solution after a reward CS), may have been because credit assignment was particularly important to distinguish these trials for the neutral solution following the neutral stimuli, while there would not have been the same conflict for juice reward trials since they only ever followed the rewarded CS.

In general, we also observed that the vmPFC (area 25) activated for several of the conditioned stimuli, including the rewarded CS, CS+Inhibitor and Inhibitor. Several accounts propose that the vmPFC is involved in general purpose valuation, and vmPFC signals have consistently been shown to fit ratings of subjective values or preferences. [88, 45, 74]. In addition, some evidence indicates that distinct regions of the vmPFC encode positive and negative valence, with the dorsal region of the vmPFC (also called rdACC) thought to preferentially encode responses to punishments, while the ventral vmPFC preferentially encodes responses to rewards. [58, 86]. However, there

have also been studies showing that vmPFC can integrate multiple dimensions of reward, including positive and negative valence attributes, into a subjective value representation [74, 36, 2]. Bringing these findings back to the interpretation of the brain activity in the conditioned inhibition task, we did not observe that the CSes with different valence associations, such as the rewarded CS and Inhibitor, were represented in different regions of the vmPFC.

Similarly, we observed activity in the insula for the rewarded CS, CS+Inhibitor and Inhibitor. Since each of these cues were associated with a taste outcome, either the orange juice or neutral solution, these results are consistent with research showing that the insula can encode anticipatory information for taste stimuli. [49]. In addition, we also observed activity in the insula during the presentation of the juice reward and the neutral solution. This is consistent with several studies that have found activation of the insula during receipt of taste stimuli [65, 69, 56, 87]. One view of the role of the insula is that it represents interoceptive awareness of bodily states [14], and integrates this with computations of the risk and uncertainty of these bodily states [99]. This explanation would be consistent with the observed activity during receipt of the juice reward and the neutral solution, as well as the activity observed during cues that predict these outcomes. Finally, the insula also showed brain activity during outcome presentation that increased parametrically as amount an outcome was worse than expected increased. We did not find that anticipatory activity when viewing the CSes varied differentially depending on the valence of the CS, contradicting some studies showing that insula activity decreased during anticipation of positive valence rewards [37], but consistent with other studies showing that the insula encodes shows overlapping activity during the prediction and receipt of many different types of taste outcomes[65, 69, 56, 87]. Also, the observed activity in the insula for worse than expected outcomes is consistent with other studies that have observed insula activity during aversive prediction errors or negative valence outcomes [96, 48, 65].

5.0.24 Conclusion

By running a conditioned inhibition study that included conditions where dopamine signaling was canceled or greatly reduced, we were able to add to previous literature that has mostly focused on understanding the neural mechanisms of reward prediction error without taking into account the confounding influence of downstream dopamine projections. In addition, using this paradigm allowed us to test the predictions of our computational model, PVLV, which makes specific predictions about the brain areas involved in the computation of reward prediction errors. We were also interested in running this study because it allowed us to examine the neural mechanisms of reward omission learning, which we predicted would recruit circuitry in the lateral habenula that also plays a role in negative valence learning.

We observed several results that were consistent with the predictions of the PVLV model. We observed a selective activation of the SNc/VTA for the rewarded CS and none of the other stimuli, which is consistent with the proposed explanation of conditioned inhibition in the PVLV model as well as electrophysiological data[105] showing that dopamine neurons respond to the CS+Inhibitor with decreased activity. There was also activity in the SNc/VTA for a contrast comparing the juice reward with the neutral solution.

We also observed activity in the ventral striatum for the rewarded CS in the first conditioning block as well as activity in the amygdala during the presentation of the juice reward, both of which are results consistent with the proposed roles of these regions in learning about rewards proposed by PVLV. However, we observed several results that were inconsistent with our predictions, including a lack of activity in the amygdala for the conditions involving the reward CS. In particular, the amygdala showed activity during US presentation, which was significant for both the juice reward and the neutral solution, but there was more overall activity for the juice reward.

Overall, there are several limitations to this study that should be addressed, and may have influenced some of the contradictory predictions we observed. In particular, there was a temporal correlation between the CS and US in the design, which was done to avoid introducing temporal

prediction errors into the learning, but makes it difficult to resolve whether the observed activity seen in temporally contiguous continuous, such as CS and US presentation, can truly be attributed to those stimuli. In addition, we were interested in the brain activity in small subcortical ROIs, including several regions such as the SNc/VTA, lateral habenula and amygdala that can be highly influenced by physiological noise as well as signal dropout. In addition, many of these regions are often difficult to localize with standard ROI analysis approaches. There were also some limitations with the behavioral study, as we did not collect subjective ratings of reward pleasantness during the study and therefore cannot be sure that the juice was truly rewarding for all subjects, though we did screen for people who disliked orange juice. Similarly, the lack of ratings for the neutral solution makes it difficult to tell whether that solution was truly neutral or may have been slightly aversive.

To address some of these limitations, we have designed a follow up study that extended this conditioned inhibition paradigm to monetary rewards, following several studies showing substantial overlap between brain mechanisms involved in processing primary taste rewards and monetary rewards [47, 56, 45]. We also collected subjective ratings of reward expectation at the end of each block to screen for subjects who did not appear to be learning the reward associations. To account for the shortcomings of the current study, we varied the jitter between events more to decrease the correlation between the regressors for the CS and US. We also used a shorter TR (about 500 ms) and collected physiological data, to better filter out the noise in brainstem areas. However, due to the particular features of the non-standard sequence used for this new study, the data is slightly noisier and needs further analysis. To hopefully aid in this endeavor, this study includes additional tasks that should aid with quality control, including a breath hold task and functional localizer.

We have described how the current state of the field in reward learning can benefit from studies that allow us to further examine the computational role of different areas in the reward prediction error. In addition, lots more research is needed on the mechanisms involved in appetitive and aversive learning, particularly as this remains relatively unclear in several areas such as the amygdala and insula that show activity across both valences. A lot of this work can benefit from

drawing on animal models, and drawing on detailed mapping of the circuitry involved to make detailed predictions about the computational role of different areas. In our view, a productive way forward involves joint development of computational models of the brain areas involved in these learning tasks, along with experimental paradigms that allow us to test and refine these models by making it possible to dissociate positive and negative valence representations, either behaviorally as in conditioned inhibition or at the analysis stage using advanced techniques such as machine learning. By moving beyond the single parameter of the reward prediction error common in TD learning, and further creating and testing biological models of limbic circuitry such as PVLV, we can better understand how the emotional brain works for both healthy individuals and in dysfunctional brains, moving the field forward scientifically and expanding our current understanding of neurological and psychological disorders.

Bibliography

- [1] Bernard W Balleine and Simon Killcross. Parallel incentive processing: an integrated view of amygdala function. Trends in neurosciences, 29(5):272–9, May 2006.
- [2] Ulrike Basten, Guido Biele, Hauke R Heekeren, and Christian J Fiebach. How the brain integrates costs and benefits during decision making. Proceedings of the National Academy of Sciences, 107(50):21767–21772, 2010.
- [3] M G Baxter and E A Murray. The amygdala and reward. Nature Reviews Neuroscience, 3:563–572, December 2002.
- [4] Timothy EJ Behrens, Laurence T Hunt, Mark W Woolrich, and Matthew FS Rushworth. Associative learning of social value. Nature, 456(7219):245–249, 2008.
- [5] Marina A Belova, Joseph J Paton, Sara E Morrison, and C Daniel Salzman. Expectation Modulates Neural Responses to Pleasant and Aversive Stimuli in Primate Amygdala. Neuron, 55(6):970–984, 2007.
- [6] Sabrina Boll, Matthias Gamer, Sebastian Gluth, Jürgen Finsterbusch, and Christian Büchel. Separate amygdala subregions signal surprise and predictiveness during associative fear learning in humans. The European journal of neuroscience, (October):1–10, December 2012.
- [7] Erie D Boorman, Timothy E J Behrens, Mark W Woolrich, and Matthew F S Rushworth. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron, 62, June 2009.
- [8] Romain Bourdy and Michel Barrot. A new control center for dopaminergic systems: pulling the VTA by the tail. Trends in neurosciences, pages 1–10, July 2012.
- [9] Mark E Bouton. Learning and the persistence of appetite: extinction and the motivation to eat and overeat. Physiology & behavior, 103(1):51–8, April 2011.
- [10] Frédéric Brischoux, Subhojit Chakraborty, Daniel I Brierley, and Mark A Ungless. Phasic excitation of dopamine neurons in ventral vta by noxious stimuli. Proceedings of the National Academy of Sciences, 106(12):4894–4899, 2009.
- [11] Ethan S Bromberg-Martin, Masayuki Matsumoto, and Okihide Hikosaka. Dopamine in motivational control: rewarding, aversive, and alerting. Neuron, 68(5):815–834, 2010.

- [12] V Coizet, E J Dommett, E M Klop, P Redgrave, and P G Overton. The parabrachial nucleus is a critical link in the transmission of short latency nociceptive information to midbrain dopaminergic neurons. Neuroscience, 168(1):263–72, June 2010.
- [13] Laura H Corbit and Bernard W Balleine. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. The Journal of neuroscience, 25(4):962–970, 2005.
- [14] Hugo D Critchley, Stefan Wiens, Pia Rotshtein, Arne Ohman, and Raymond J Dolan. Neural systems supporting interoceptive awareness. Nature neuroscience, 7(2):189–95, February 2004.
- [15] W. a. Cunningham and T. Brosch. Motivational Saliency: Amygdala Tuning From Traits, Needs, Values, and Goals. Current Directions in Psychological Science, 21(1):54–59, January 2012.
- [16] K D’Ardenne, S M McClure, L E Nystrom, and J D Cohen. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science, 319(5867):1264–1267, 2008.
- [17] Kimberlee D’Ardenne, Samuel M McClure, Leigh E Nystrom, and Jonathan D Cohen. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science (New York, N.Y.), 319(5867):1264–7, February 2008.
- [18] Yu-Lin Dong, Yugo Fukazawa, Wen Wang, Naomi Kamasawa, and Ryuichi Shigemoto. Differential postsynaptic compartments in the laterocapsular division of the central nucleus of amygdala for afferents from the parabrachial nucleus and the basolateral nucleus in the rat. The Journal of comparative neurology, 518(23):4771–91, December 2010.
- [19] Christopher D Fiorillo, Philippe N Tobler, and Wolfram Schultz. Discrete coding of reward probability and uncertainty by dopamine neurons. Science, 299(5614):1898–1902, 2003.
- [20] Guido K W Frank, Tyson a Oberndorfer, Alan N Simmons, Martin P Paulus, Julie L Fudge, Tony T Yang, and Walter H Kaye. Sucrose activates human taste pathways differently from artificial sweetener. NeuroImage, 39(4):1559–69, February 2008.
- [21] Guido K W Frank, Jeremy R Reynolds, Megan E Shott, Leah Jappe, Tony T Yang, Jason R Tregellas, and Randall C O’Reilly. Anorexia nervosa and obesity are associated with opposite brain reward response. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 37(9):2031–46, August 2012.
- [22] Guido K W Frank, Jeremy R Reynolds, Megan E Shott, and Randall C O’Reilly. Altered Temporal Difference Learning in Bulimia Nervosa. Biological psychiatry, July 2011.
- [23] J P Frankmann and J a Adams. Psychological bulletin. Psychological bulletin, 59(2):257–72, July 1962.
- [24] Alexander Friedman, Elad Lax, Yahav Dikshtein, Lital Abraham, Yakov Flaumenhaft, Einav Sudai, Moshe Ben-Tzion, and Gal Yadid. Electrical stimulation of the lateral habenula produces an inhibitory effect on sucrose self-administration. Neuropharmacology, 60(2-3):381–7, 2011.

- [25] Carl E Fulwiler and Clifford B Saper. Subnuclear organization of the efferent connections of the parabrachial nucleus in the rat. Brain Research Reviews, 7(3):229–259, 1984.
- [26] Michela Gallagher, Phillip W Graham, and Peter C Holland. The amygdala central nucleus and appetitive pavlovian conditioning: lesions impair one class of conditioned behavior. The Journal of neuroscience, 10(6):1906–1911, 1990.
- [27] Samuel J Gershman, David M Blei, and Yael Niv. Context, learning, and extinction. Psychological review, 117(1):197–209, January 2010.
- [28] Samuel J Gershman and Yael Niv. Exploring a latent cause theory of classical conditioning. Learning & behavior, 40(3):255–68, September 2012.
- [29] S N Haber, J L Fudge, and N R McFarland. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. The Journal of neuroscience : the official journal of the Society for Neuroscience, 20(6):2369–82, March 2000.
- [30] Thomas E Hazy, Michael J Frank, and Randall C O’Reilly. Neural mechanisms of acquired phasic dopamine responses in learning. Neuroscience and biobehavioral reviews, 34(5):701–20, April 2010.
- [31] L Heimer, St Louis, Comparative Anatomy, and De Olmos. Specificity in the projection patterns of accumbal core and shell in the rat. 41:89–125, 1991.
- [32] Peter C Holland, Yolanda Chik, and Qui Zhang. Inhibitory learning tests of conditioned stimulus associability in rats with lesions of the amygdala central nucleus. Behavioral neuroscience, 115(5):1154, 2001.
- [33] Peter C Holland, Jung-Soo Han, and Helena M Winifield. Operant and pavlovian control of visual stimulus orienting and food-related behaviors in rats with lesions of the amygdala central nucleus. Behavioral neuroscience, 116(4):577, 2002.
- [34] Simon Hong and Okihide Hikosaka. The globus pallidus sends reward-related signals to the lateral habenula. Neuron, 60(4):720–9, November 2008.
- [35] Simon Hong, Thomas C Jhou, Mitchell Smith, Kadharbatcha S Saleem, and Okihide Hikosaka. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. The Journal of neuroscience, 31, August 2011.
- [36] Thorsten Kahnt, Jakob Heinzle, Soyoung Q Park, and John-Dylan Haynes. Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. Neuroimage, 56(2):709–715, 2011.
- [37] Hackjin Kim, Shinsuke Shimojo, and John P O’Doherty. Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. Cerebral cortex (New York, N.Y. : 1991), 21(4):769–76, April 2011.
- [38] Ewelina Knapska, Grazyna Walasek, Evgeni Nikolaev, Frieder Neuhäusser-Wespy, Hans-Peter Lipp, Leszek Kaczmarek, and Tomasz Werka. Differential involvement of the central amygdala in appetitive versus aversive learning. Learning & memory (Cold Spring Harbor, N.Y.), 13(2):192–200, 2006.

- [39] B Knutson, G W Fong, C M Adams, J L Varner, and D Hommer. Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport, 12(17):3683–7, December 2001.
- [40] Michelle A Kronforst-Collins and John F Disterhoft. Lesions of the caudal area of rabbit medial prefrontal cortex impair trace eyeblink conditioning. Neurobiology of learning and memory, 69(2):147–162, 1998.
- [41] Stephan Lammel, Byung Kook Lim, Chen Ran, Kee Wui Huang, Michael J. Betley, Kay M. Tye, Karl Deisseroth, and Robert C. Malenka. Input-specific control of reward and aversion in the ventral tegmental area. Nature, October 2012.
- [42] M E Le Pelley. The role of associative history in models of associative learning: a selective review and a hybrid model. The Quarterly journal of experimental psychology. B, Comparative and physiological psychology, 57(3):193–243, July 2004.
- [43] Hongjoo J Lee, Michela Gallagher, and Peter C Holland. The central amygdala projection to the substantia nigra reflects prediction error information in appetitive conditioning. Learning & memory (Cold Spring Harbor, N.Y.), 17(10):531–8, October 2010.
- [44] Hongjoo J Lee, Frank Groshek, Gorica D Petrovich, Joseph P Cantalini, Michela Gallagher, and Peter C Holland. Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. The Journal of neuroscience : the official journal of the Society for Neuroscience, 25(15):3881–8, April 2005.
- [45] Dino J Levy and Paul W Glimcher. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. The Journal of neuroscience : the official journal of the Society for Neuroscience, 31(41):14693–707, October 2011.
- [46] Jian Li, Daniela Schiller, Geoffrey Schoenbaum, Elizabeth a Phelps, and Nathaniel D Daw. Differential roles of human striatum and amygdala in associative learning. Nature Neuroscience, 14(10):1250–1252, September 2011.
- [47] Seung-Lark Lim, John P O’Doherty, and Antonio Rangel. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. The Journal of neuroscience : the official journal of the Society for Neuroscience, 31(37):13214–23, September 2011.
- [48] Xun Liu, David K Powell, Hongbin Wang, Brian T Gold, Christine R Corbly, and Jane E Joseph. Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. The Journal of neuroscience : the official journal of the Society for Neuroscience, 27(17):4587–97, April 2007.
- [49] Arianna Maffei, Melissa Haley, and Alfredo Fontanini. Neural processing of gustatory information in insular circuits. Current opinion in neurobiology, 22(4):709–716, 2012.
- [50] Stephen V Mahler and Gary S Aston-Jones. Fos Activation of Selective Afferents to Ventral Tegmental Area during Cue-Induced Reinstatement of Cocaine Seeking in Rats. The Journal of neuroscience : the official journal of the Society for Neuroscience, 32(38):13309–25, September 2012.

- [51] M Matsumoto and O Hikosaka. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature, 447:1111–1115, 2007.
- [52] Masayuki Matsumoto and Okihide Hikosaka. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature, 459(7248):837–41, June 2009.
- [53] Masayuki Matsumoto and Masahiko Takada. Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. Neuron, 2013.
- [54] Jonathan A McCabe, Ryan K Jessup, Peter Bossaerts, John P O Doherty, and Charlotte Pre. Differentiable contributions of human amygdalar subregions in the computations underlying reward and avoidance learning. Brain, pages 1–12, 2011.
- [55] Samuel M McClure, Gregory S Berns, and P Read Montague. Temporal prediction errors in a passive learning task activate human striatum. Neuron, 38:339–346, 2003.
- [56] Elise Metereau and Jean-Claude Dreher. Cerebral Correlates of Salient Prediction Error for Different Rewards and Punishments. Cerebral cortex, February 2012.
- [57] Jacques Mirenowicz and Wolfram Schultz. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature, 379(6564):449–451, 1996.
- [58] Ilya E Monosov and Okihide Hikosaka. Regionally Distinct Processing of Rewards and Punishments by the Primate Ventromedial Prefrontal Cortex. The Journal of Neuroscience, 32:10318–10330, 2012.
- [59] James R Moyer, Richard A Deyo, and John F Disterhoft. Hippocampectomy disrupts trace eye-blink conditioning in rabbits. Behavioral neuroscience, 104(2):243, 1990.
- [60] B Myers-Schulz and M Koenigs. Functional anatomy of ventromedial prefrontal cortex: implications for mood and anxiety disorders. Molecular psychiatry, 17(2):132–41, February 2012.
- [61] Kae Nakamura, Masayuki Matsumoto, and Okihide Hikosaka. Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. The Journal of neuroscience : the official journal of the Society for Neuroscience, 28(20):5331–43, May 2008.
- [62] Helen M Nasser and Gavan P McNally. Appetitive–aversive interactions in pavlovian fear conditioning. Behavioral neuroscience, 126(3):404, 2012.
- [63] J B Nitschke, G E Dixon, I Sarinopoulos, S J Short, J D Cohen, E E Smith, S M Kosslyn, R M Rose, and R J Davidson. Altering expectancy dampens neural response to aversive taste in primary taste cortex. Nature neuroscience, 9(3):435–442, 2006.
- [64] Jack B Nitschke, Issidoros Sarinopoulos, Kristen L Mackiewicz, Hillary S Schaefer, and Richard J Davidson. Functional neuroanatomy of aversion and its anticipation. Neuroimage, 29(1):106–116, 2006.
- [65] Jack B. Nitschke, Issidoros Sarinopoulos, Kristen L. Mackiewicz, Hillary S. Schaefer, and Richard J. Davidson. Functional neuroanatomy of aversion and its anticipation. NeuroImage, 29(1):106 – 116, 2006.

- [66] M P Noonan, R B Mars, and M F S Rushworth. Distinct roles of three frontal cortical areas in reward-guided behavior. The Journal of neuroscience, 31(40):14399–14412, October 2011.
- [67] John O’Doherty, Morten L Kringelbach, Edmund T Rolls, Julia Hornak, and Caroline Andrews. Abstract reward and punishment representations in the human orbitofrontal cortex. Nature neuroscience, 4(1):95–102, 2001.
- [68] John P O’Doherty, Tony W Buchanan, Ben Seymour, and Raymond J Dolan. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. Neuron, 49(1):157–66, January 2006.
- [69] John P O’Doherty, Ralf Deichmann, Hugo D Critchley, and Raymond J Dolan. Neural responses during anticipation of a primary taste reward. Neuron, 33(5):815–26, March 2002.
- [70] Taketoshi Ono, Hisao Nishijo, and Teruko Uwano. Amygdala role in conditioned associative learning. Progress in neurobiology, 46(4):401–422, 1995.
- [71] Randall C O’Reilly, Michael J Frank, Thomas E Hazy, and Brandon Watz. PVLV: the primary value and learned value Pavlovian learning algorithm. Behavioral neuroscience, 121(1):31–49, March 2007.
- [72] Olga Therese Ousdal, Greg E. Reckless, Andres Server, Ole A. Andreassen, and Jimmy Jensen. Effect of relevance on amygdala activation and association with the ventral striatum. NeuroImage, 62(1):95 – 101, 2012.
- [73] O.T. Ousdal, J. Jensen, A. Server, A.R. Hariri, P.H. Nakstad, and O.A. Andreassen. The human amygdala is involved in general behavioral relevance detection: Evidence from an event-related functional magnetic resonance imaging go-nogo task. Neuroscience, 156(3):450 – 455, 2008.
- [74] Soyoung Q Park, Thorsten Kahnt, Jörg Rieskamp, and Hauke R Heekeren. Neurobiology of value integration: when value impacts valuation. The Journal of neuroscience : the official journal of the Society for Neuroscience, 31(25):9307–14, June 2011.
- [75] Ivan Petrovich Pavlov. Conditioned reflexes. DoverPublications. com, 1927.
- [76] John M Pearson and Michael L Platt. Dopamine: Burning the candle at both ends. Neuron, 79(5):831–833, 2013.
- [77] Anthony G Phillips, Soyoon Ahn, and John G Howland. Amygdalar control of the mesocorticolimbic dopamine system: parallel pathways to motivated behavior. Neuroscience & Biobehavioral Reviews, 27(6):543–554, 2003.
- [78] JL Price and DG Amaral. An autoradiographic study of the projections of the central nucleus of the monkey amygdala. The Journal of Neuroscience, 1(11):1242–1259, 1981.
- [79] Joseph L Price. Prefrontal cortical networks related to visceral function and mood. Annals of the New York Academy of Sciences, 877(1):383–396, 1999.
- [80] Joseph L Price. Comparative aspects of amygdala connectivity. Annals of the New York Academy of Sciences, 985(1):50–58, 2003.

- [81] Robert J Purgert, Daniel S Wheeler, Michael a McDannald, and Peter C Holland. Role of amygdala central nucleus in aversive learning produced by shock or by unexpected omission of food. The Journal of neuroscience : the official journal of the Society for Neuroscience, 32(7):2461–72, February 2012.
- [82] Robert A Rescorla. Pavlovian conditioned inhibition. Psychological Bulletin, 72(2):77, 1969.
- [83] Robert A Rescorla. Conditioned inhibition and extinction. Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski, pages 83–110, 1979.
- [84] Matthew R Roesch, Donna J Calu, Guillem R Esber, and Geoffrey Schoenbaum. Neural correlates of variations in event processing during learning in basolateral amygdala. The Journal of neuroscience : the official journal of the Society for Neuroscience, 30(7):2464–71, February 2010.
- [85] Matthew R. Roesch, Guillem R. Esber, Jian Li, Nathaniel D. Daw, and Geoffrey Schoenbaum. Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. European Journal of Neuroscience, 35(7):1190–1200, April 2012.
- [86] Mathieu Roy, Daphna Shohamy, and Tor D Wager. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends in cognitive sciences, 16(3):147–156, 2012.
- [87] K Rudenga, B Green, D Nachtigal, and D M Small. Evidence for an integrated oral sensory module in the human anterior ventral insula. Chemical senses, 35(8):693–703, October 2010.
- [88] Matthew F S Rushworth, Maryann P Noonan, Erie D Boorman, Mark E Walton, and Timothy E Behrens. Frontal cortex and reward-guided learning and decision-making. Neuron, 70(6):1054–69, June 2011.
- [89] Matthew FS Rushworth and Timothy EJ Behrens. Choice, uncertainty and value in prefrontal and cingulate cortex. Nature neuroscience, 11(4):389–397, 2008.
- [90] David Sander, Jordan Grafman, and Tiziana Zalla. The human amygdala: an evolved system for relevance detection. Reviews in the Neurosciences, 14(4):303–316, 2003.
- [91] Susan Sangha, James Z Chadick, and Patricia H Janak. Safety encoding in the basal amygdala. The Journal of neuroscience : the official journal of the Society for Neuroscience, 33(9):3744–51, February 2013.
- [92] Susan Sangha, James Z Chadick, and Patricia H Janak. Safety encoding in the basal amygdala. The Journal of neuroscience : the official journal of the Society for Neuroscience, 33(9):3744–51, February 2013.
- [93] W. Schultz. A Neural Substrate of Prediction and Reward. Science, 275(5306):1593–1599, March 1997.
- [94] Wolfram Schultz. Getting formal with dopamine and reward. Neuron, 36(2):241–63, October 2002.
- [95] Wolfram Schultz, Paul Apicella, and Tomas Ljungberg. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. The Journal of Neuroscience, 13(3):900–913, 1993.

- [96] Ben Seymour, John P O'Doherty, Martin Koltzenburg, Katja Wiech, Richard Frackowiak, Karl Friston, and Raymond Dolan. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nature neuroscience, 8(9):1234–1240, 2005.
- [97] Steven J Shabel and Patricia H Janak. Substantial similarity in amygdala neuronal activity during conditioned appetitive and aversive emotional arousal. Proceedings of the National Academy of Sciences of the United States of America, 106(35):15031–6, October 2009.
- [98] L Shelton, L Becerra, and D Borsook. Unmasking the mysteries of the habenula in pain and analgesia. Progress in neurobiology, 96(2):208–19, February 2012.
- [99] Tania Singer, Hugo D Critchley, and Kerstin Preusschoff. A common role of insula in feelings, empathy and uncertainty. Trends in cognitive sciences, 13(8):334–340, 2009.
- [100] Alice M Stamatakis and Garret D Stuber. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. Nature neuroscience, 15(8), June 2012.
- [101] Elizabeth E Steinberg, Ronald Keiffin, Josiah R Boivin, Ilana B Witten, Karl Deisseroth, and Patricia H Janak. A causal link between prediction errors, dopamine neurons and learning. Nature Neuroscience, 16(7):966–973, May 2013.
- [102] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction, volume 1. Cambridge Univ Press, 1998.
- [103] Larry W Swanson and Gorica D Petrovich. What is the amygdala? Trends in neurosciences, 21(8):323–331, 1998.
- [104] Kelly R Tan, Cédric Yvon, Marc Turiault, Julie J Mirzabekov, Jana Doehner, Gwenaël Labouèbe, Karl Deisseroth, Kay M Tye, and Christian Lüscher. {GABA} Neurons of the {VTA} Drive Conditioned Place Aversion. Neuron, 73(6):1173–1183, March 2012.
- [105] Philippe N Tobler, Anthony Dickinson, and Wolfram Schultz. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. The Journal of neuroscience : the official journal of the Society for Neuroscience, 23(32):10402–10, December 2003.
- [106] Ruud vanZessen, Jana L. Phillips, Evgeny A. Budygin, and GarretD. Stuber. Activation of VTA GABA Neurons Disrupts Reward Consumption. Neuron, 73(6):1184–1194, March 2012.
- [107] P Waelti, a Dickinson, and W Schultz. Dopamine responses comply with basic assumptions of formal learning theory. Nature, 412(6842):43–8, July 2001.

Appendix A

Modeling equations: PVLV and Rescorla-Wagner

A.0.25 Rescorla-Wagner model of reward learning

$$\Delta V = \alpha\beta(\lambda - V) \tag{A.1}$$

V represents the associative strength of the current cue, α represents the associability of A and depends on the cure's intensity or salience, β is learning rate parameter reflecting the intensity of the US and λ represents the maximum amount of conditioning supported by that US (Bush and Mosteller as cited in [42]). Therefore, associative change in Rescorla Wagner learning is determined by the error term capturing how well the US is predicted by the CS.

A.0.26 PVLV model: simulating positive and negative valence

In all equations, the ϵ parameter captures the current learning rate.

- (1) Primary value/change in PV: The PVLV model calculates the change in PV by looking at the difference between PVe (the current PV) and PVi (the expected PV).

$$\delta_{pv} = (PVe - PVi) \tag{A.2}$$

- (2) PV learning: Changes in PV weights for CSes are calculated based on the difference between PVe (the current PV) and PVi (the expected PV), causing weight changes if PVe is substantially different than PVi . Similarly, the PVi system can learn positive and negative

weights for CSes to account for positive and negative valence PVes. $W(i,g)$ represents the sending weight for the current unit with activation x_i .

$$\Delta w_{ig} = \epsilon (PVe - PVi) x_i \quad (\text{A.3})$$

- (3) LV learning: Similarly, the PVLV model learns a change in LVe weights for a CS only when a PVe is present and figures out how to change the LVe weights based on the difference between the current PVe (or value of the current US) and the previous LVe weight for that CS. $W(i,g)$ represents the sending weight for the current unit with activation x_i .

$$\Delta w_{ig} = \epsilon (PVe - LVe) x_i \text{ iff } PV > 0 \quad (\text{A.4})$$

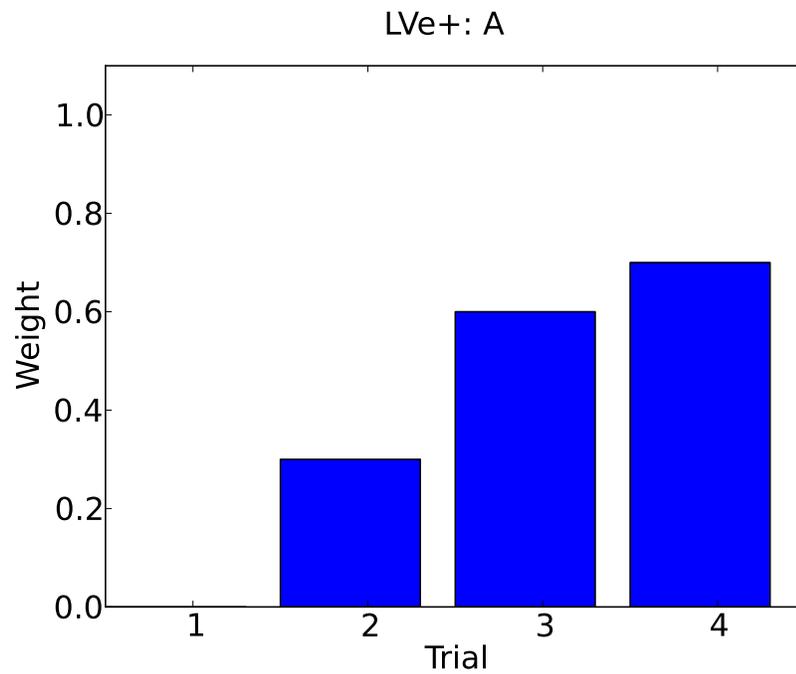
Appendix B

Simulation diagrams of conditioned inhibition

While learning about a positively valenced rewarded CS in PVLV, the model increases the PVI weights (which learn to inhibit anticipated rewards at the time of US), and the LVe weights (which learn positive and negative associations with reward (LVe+ and LVe-, respectively)).

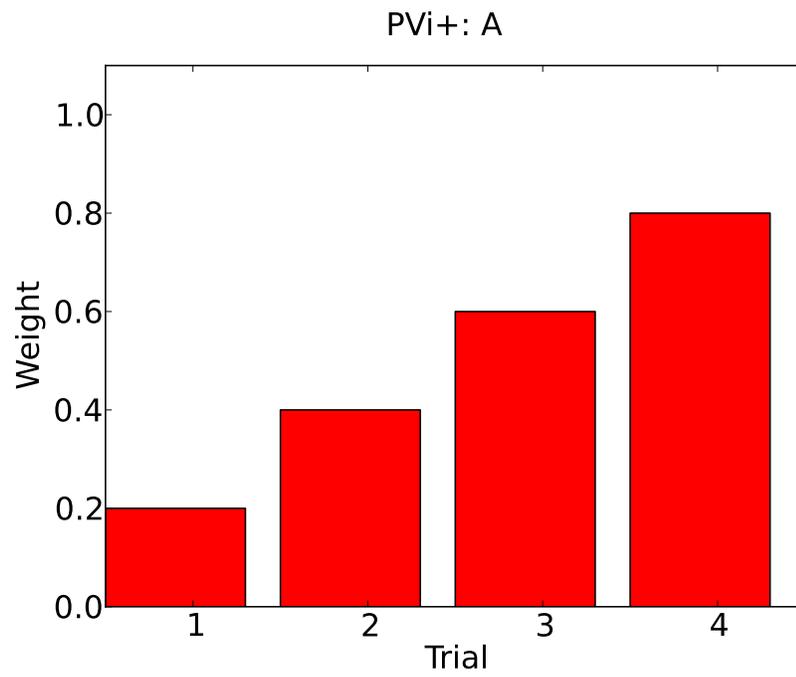
B.0.27 LVe (central amygdala) at time of CS

Figure B.1: Repeated pairings of the CS (A) with reward causes the LVe+ weights for the CS to increase. In this diagram, the learning rate was increased to simulate the effect of learning over several trials in one trial.



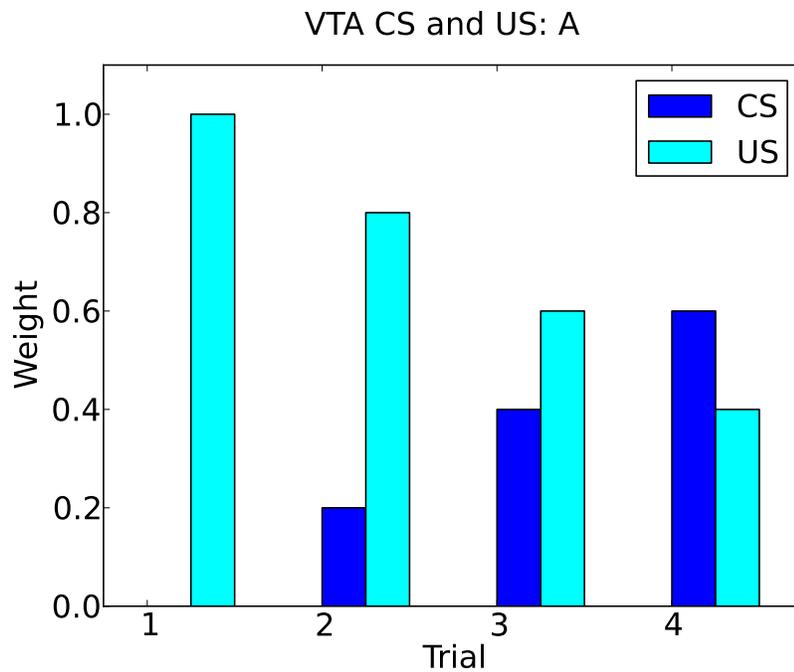
B.0.28 PVi learning for rewarded CS

Figure B.2: Repeated pairings of the CS+Inhibitor pair followed by no reward cause PVLV to learn a negatively valenced weight for the CS+Inhibitor.



B.0.29 VTA learning for rewarded CS

Figure B.3: Over repeated pairings of the CS with reward, the VTA first shows activity to a positively valenced US, due to P_{Ve} weights influencing dopamine. After repeated learning, this dopamine release moves from the time of the reward US to the CS that reliably predicts the reward. This is a joint function of the P_{Vi} weights, which inhibit dopamine firing at the time of the US for predicted rewards (see Figure B.0.28), and the L_{Ve} weights (see figure B.0.27), which influence dopamine release at the time of the CS for CSes that predict reward.



B.0.30 Conditioned inhibition learning

Figure B.4: Similarly, once the model has learned the LVe and PVi weights for the rewarded CS, to cause DA release to the CS, and PVi weights to inhibit the dopamine firing for the predicted reward, the model can then learn conditioned inhibition. Once the model has a positive LV for the reward CS, when it sees the CS+Inhibitor it begins to learn a negatively valenced prediction (LVe-) for the Inhibitor, due to differences between the predicted reward and the received reward (see A.3). In this diagram, the learning rate was also increased to simulate the effect of learning over several trials in one trial.

