frontiers

The neural correlates of cued reward omission

Jessica A. Mollick[1*], Luke J. Chang[2], Anjali Krishnan[3], Thomas E. Hazy[4], Kai A. Krueger[4], Guido K. Frank[5], Tor D. Wager[6], Randall C. O'Reilly[7]

[1]Department of Psychiatry, Yale University, United States, [2]Department of Psychological and Brain Sciences, Dartmouth College, United States, [3]Department of Psychology, Brooklyn College (CUNY), United States, [4]eCortex, Inc., United States, [5]UCSD Eating Disorder Center for Treatment and Research, University of California, San Diego, United States, [6]Department of Psychological and Brain Sciences, Dartmouth College, United States, [7]Department of Psychology and Computer Science, Center for Neuroscience, University of California Davis, United States

Submitted to Journal:
Frontiers in Human Neuroscience

Specialty Section:
Cognitive Neuroscience

Article type:
Original Research Article

Manuscript ID:
615313

Received on:
08 Oct 2020

Revised on:
14 Jan 2021

Journal website link:
www.frontiersin.org

Conflict of interest statement

The authors declare a potential conflict of interest and state it below

Randall C. O'Reilly is CSO, and Jessica A. Mollick, Thomas E. Hazy, and Kai A. Kruger are researchers at eCortex, Inc., Boulder, Colorado, which may derive indirect benefit from the work presented here.

Author contribution statement

J.A.M., T.D.W., R.C.R, and G.K.F. contributed to the conception and design of the study. J.A.M. performed statistical analyses of the data. A.K., L.J.C., K.A.K, T.H, T.D.W, G.K.F. and R.O.R. contributed to interpretation and discussion of neuroimaging and behavioral results. J.A.M. wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Keywords

conditioned inhibition, Habenula, fMRI, negative, Prediction error, Reward, Learning

Abstract

Word count: 230

Compared to our understanding of positive prediction error signals occurring due to unexpected reward outcomes, less is known about the neural circuitry in humans that drives negative prediction errors during omission of expected rewards. While classical learning theories such as Rescorla-Wagner or temporal difference learning suggest that both types of prediction errors result from a simple subtraction, there has been recent evidence suggesting that different brain regions provide input to dopamine neurons which contributes to specific components of this prediction error computation. Here, we focus on the brain regions responding to negative prediction error signals, which has been well-established in animal studies to involve a distinct pathway through the lateral habenula. We examine the activity of this pathway in humans, using a conditioned inhibition paradigm with high-resolution functional MRI. First, participants learned to associate a sensory stimulus with reward delivery. Then, reward delivery was omitted whenever this stimulus was presented simultaneously with a different sensory stimulus, the conditioned inhibitor. Both reward presentation and the reward-predictive cue activated midbrain dopamine regions, insula and orbitofrontal cortex. While we found significant activity at an uncorrected threshold for the conditioned inhibitor in the habenula, consistent with our predictions, it did not survive correction for multiple comparisons and awaits further replication. Additionally, the pallidum and putamen regions of the basal ganglia showed modulations of activity for the inhibitor that did not survive the corrected threshold.

Contribution to the field

This manuscript advances our understanding of the neural mechanisms involved in reward learning, specifically unexpected reward omissions, which result in negative reward prediction errors (RPEs), and how the brain responds to cues predicting these omissions. While many computational models of learning, such as Rescorla-Wagner and temporal-difference learning, suggest that both positive RPEs for unexpected rewards and negative RPEs for omissions result from a simple subtraction, recent evidence suggests that different brain areas contribute to specific components of this computation. To further examine brain areas involved in computing negative RPEs, we adapted a conditioned inhibition task to human fMRI. In conditioned inhibition, a cue reliably associated with reward omissions, a conditioned inhibitor, acquires the ability to reduce reward expectations when paired with a conditioned reward cue. We found evidence that the lateral habenula responded to the reward omission cue. There are several important applications of this research. For example, depression has been associated with enhanced encoding of negative RPEs in the lateral habenula, and negative RPEs also contribute to extinction, where unexpected reward omissions reduce reward expectations. This research further supports the role of the habenula in negative RPEs, and demonstrates the utility of translating animal paradigms to human fMRI.

Funding statement

Supported by: NIH IBSC Center 1-P50-MH079485 and NIH R01GM109996.

Ethics statements

Studies involving animal subjects

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: The studies involving human participants were reviewed and approved by Institutional Review Board of the University of Colorado, Boulder. The patients/participants provided their written informed consent to participate in this study.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable human images or data is presented in this study.

Data availability statement

Generated Statement: The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: All data used for the ROI analysis figures and behavioral rating data can be found in the Open Science Framework: https://osf.io/njbmf/. Neuroimaging data used for the figures of second-level analysis results can be found on Neurovault: https://neurovault.org/collections/8676/.

The neural correlates of cued reward omission

Jessica A. Mollick[1], Luke J. Chang[2], Anjali Krishnan[4], Thomas E. Hazy[6], Kai A. Krueger[6], Guido K. W. Frank[5], Tor D. Wager[2], Randall C. O'Reilly[3]

[1]Department of Psychiatry, Yale University, New Haven, United States

[2]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, United States

[3]Department of Psychology and Computer Science Center for Neuroscience, University of California Davis, Davis, CA, United States

[4]Department of Psychology, Brooklyn College of the City University of New York, Brooklyn, United States

[5]UCSD Eating Disorder Center for Treatment and Research, University of California, San Diego, United States

[6]eCortex, Inc, Boulder, United States

*** Correspondence:**
Corresponding Author
Jessica.Mollick@yale.edu

Keywords: conditioned inhibition, habenula, fMRI, negative, prediction error, reward, learning

Abstract

Compared to our understanding of positive prediction error signals occurring due to unexpected reward outcomes, less is known about the neural circuitry in humans that drives negative prediction errors during omission of expected rewards. While classical learning theories such as Rescorla-Wagner or temporal difference learning suggest that both types of prediction errors result from a simple subtraction, there has been recent evidence suggesting that different brain regions provide input to dopamine neurons which contributes to specific components of this prediction error computation. Here, we focus on the brain regions responding to negative prediction error signals, which has been well-established in animal studies to involve a distinct pathway through the lateral habenula. We examine the activity of this pathway in humans, using a conditioned inhibition paradigm with high-resolution functional MRI. First, participants learned to associate a sensory stimulus with reward delivery. Then, reward delivery was omitted whenever this stimulus was presented simultaneously with a different sensory stimulus, the conditioned inhibitor. Both reward presentation and the reward-predictive cue activated midbrain dopamine regions, insula and orbitofrontal cortex. While we found significant activity at an uncorrected threshold for the conditioned inhibitor in the habenula, consistent with our predictions, it did not survive correction for multiple comparisons and awaits further replication. Additionally, the pallidum and putamen regions of the basal ganglia showed modulations of activity for the inhibitor that did not survive the corrected threshold.

1 Introduction

While the field of reinforcement learning has generally focused on the role of reward prediction errors in training reward expectations, the mechanisms involved in learning about omission of expected reward delivery are less well understood. Classical models of learning such as Rescorla-Wagner and TD models suggest that prediction errors result from a simple subtractive computation, which also has been shown to match the firing of dopamine neurons. However, there is also recent evidence suggesting that brain areas projecting to dopamine neurons may provide input which contributes to specific parts of this computation, for example, some regions may encode the level of expected reward (Cohen et al., 2012), while others may respond specifically to worse than expected outcomes. Here, we focus on the latter computation, which has been well-established in animal studies, showing that neurons in the lateral habenula respond both to aversive outcomes and the omission of an expected reward, and further drive an inhibition of dopamine neurons, leading to the "dip" component of prediction error encoding how much worse something was than expected (Matsumoto and Hikosaka, 2009b).

In appetitive Pavlovian conditioning, individuals learn expectations about stimuli that are reliably paired with rewards. This conditioning procedure causes the previously neutral cue to drive a conditioned response. In conditioned inhibition, a conditioned stimulus (CS) associated with reward is presented simultaneously with a conditioned inhibitor (CI), which causes the expected reward not to occur. Conditioned inhibition occurs because the unexpected omission of reward causes a negative reward prediction error. By learning theories like Rescorla-Wagner, if another sensory stimulus is reliably present during these unexpected omissions, the accumulation of negative prediction errors causes the conditioned inhibitor to acquire negative value. This results in inhibitory conditioning, and a reduction of the conditioned response. For example, imagine that you enjoy drinking tea, but cannot

make it when your kettle is broken. Over time, the broken kettle becomes a conditioned inhibitor because it reliably predicts the omission of tea.

Computationally, conditioned inhibition is an interesting problem, because it relies on the negative prediction errors that occur when the CS+ is unexpectedly followed by a reward omission in the presence of the inhibitor, which causes the conditioned inhibitor to acquire negative value, even though the conditioned inhibitor has never been paired with an aversive stimulus. Once inhibition is acquired, the inhibitor can pass the summation test, meaning there is a reduced conditioned response to a CS paired with the inhibitor compared to the CS alone (Rescorla, 1969a). Further, we chose the paradigm based on the potential to dissociate the mechanisms of reward prediction at the time of the CS from those controlling reward predictions at the time of the unconditioned stimulus (US). In the trials where the inhibitor is presenting concurrently with the CS+, there may be a representation of the CS+ linked with an expectation of reward, along with a representation of the inhibitor linked with a reward omission. Interestingly, Tobler et al. (2003) showed a combined burst and dip to the CS+ paired with the Inhibitor, which may reflect these two associations. In contrast, at the time of the US, the conditioned inhibition procedure leads to an expectation of no reward, evidenced by the ability of the conditioned inhibitor to transfer inhibition to a novel CS+, and the enhanced dopamine burst when the conditioned inhibitor is unexpectedly followed by reward (Tobler et al., 2003). This account was recently simulated in a computational model of conditioned inhibition and other conditioning phenomena incorporating separate learning mechanisms for the control of dopamine responses at the time of the CS and US (Mollick et al., 2020). However, this theoretical account does not incorporate the idea that there might be learning for the combined stimulus of CS+ and Inhibitor as well, signaling a new context of reward omissions, drawing on ideas of state-splitting that may also occur in extinction (Redish et al., 2007), or as a conjunctive representation, possibly represented in the hippocampus (Rudy and O'Reilly, 2001).

The prediction error response in dopamine neurons includes both increases in firing for better than expected outcomes and decreases in firing, or dopamine dips, for worse than expected outcomes. However, few studies have focused on understanding the role of certain brain areas in the processes driving dopamine dip signals for worse than expected outcomes and how these areas are involved in learning about stimuli that predict reward omissions. In particular, an unanswered question remains about the extent to which brain areas involved in learning about reward omissions overlap with those involved in learning about aversive stimuli.

Theories about how the positive and negative valence learning systems interact have proposed that something that stops a negative state leads to positive emotions, while the omission of a positive reward leads to negative emotions (Mowrer, 1956; Solomon and Corbit, 1974; Seymour et al., 2007b; Maia, 2010). However, human fMRI studies have generally focused on the neural correlates of positive prediction errors for reward outcomes, though some have also begun to examine whether regions like the lateral habenula (Hennigan et al., 2015) and periaqueductal grey (PAG) (Roy et al., 2014) encode prediction error signals for aversive outcomes. While these studies have greatly advanced our understanding of the brain areas involved in learning about reward and aversion, they do not examine whether the same brain areas that are involved in associations of conditioned stimuli with rewards are the same regions that drive prediction errors if reward expectations are violated. Further, the diffuse modulatory effects of dopamine release make it difficult to tell whether the brain areas that encode reward prediction errors are providing inputs to the dopamine system or reflecting downstream effects of dopamine release. We ran a conditioned inhibition paradigm to look specifically at the negative prediction error mechanisms associated with learning about a predictor of reward omissions and compare those with learning signals for positive reward predictors.

While previous fMRI studies have looked at the neural mechanisms involved in monetary losses and the presentation of aversive stimuli, no human fMRI studies have focused on the learning about

predictors of reward omissions in a conditioned inhibition experiment. Dopamine neurons respond to a conditioned inhibitor with an inhibition, or pause in tonic firing, the same pattern of dopamine release in the substantia nigra seen to an aversive stimulus (Tobler et al., 2003; Schultz, 2007). Intriguingly, recent research has shown that this inhibition of dopamine neurons, or dip, is driven by the lateral habenula, which has been found to be activated during aversive processing and reward omissions (Matsumoto and Hikosaka, 2009b). In this study, we examined if the same signals that have been reported for reward omissions in monkey studies, particularly an increase in lateral habenula activity accompanied by a reduction in the firing of dopamine neurons, could be observed in human fMRI. These signals also occur for a CS associated with reward omission, so we predicted a strong habenula signal for the conditioned inhibitor that was associated with reward omission.

To examine the brain areas involved in each of these computations, we ran a novel fMRI study, adapting the conditioned inhibition paradigm from Tobler et al. (2003) to human participants. Using a taste pump apparatus, participants learned to associate previously neutral visual stimuli with the presentation of orange juice rewards. In an initial conditioning block, participants learned associations of a CS+ with the orange juice reward and a CS- with the neutral solution. Importantly, this was then followed by a conditioned inhibition procedure, where the originally rewarded stimulus was paired with another cue that deterministically lead to the neutral solution instead of the expected orange juice reward. Due to the disappointment (and negative reward prediction errors) resulting from omission of the expected orange juice, the cue that predicts omission becomes a conditioned inhibitor. Importantly, once a cue has acquired inhibitory properties, it should both reduce the value of the CS+ during the predictive phase, and lead to an expectation of no reward at the time of the unconditioned stimulus. Further, these inhibitory properties can be tested by unexpectedly following the conditioned inhibitor with a juice reward. If it has acquired inhibition, then the unexpected presentation of reward after the inhibitor should lead to a prediction error signal. Further, the

prediction error for the inhibitor followed by an unexpected reward should be larger than the prediction error that occurs when a neutral control stimulus is unexpectedly followed by reward, due to the inhibitory properties acquired by the inhibitor during conditioned inhibition.

See **Figure 1** for a schematic of the conditioned inhibition fMRI design.

[Figure 1 about here.]

2 Materials and Methods

2.1 Participants

19 participants (13 female) ranging between 19 and 55 years old, from the University of Colorado, Boulder, and the local community volunteered for the study. All participants were right-handed and generally in good health. Participants were screened for MRI contraindications and provided informed written consent for protocols approved by the Institutional Review Board of the University of Colorado, Boulder. Participants were paid $48 for completing the study in addition to earnings from the task.

2.2 Experimental Procedures

The functional imaging was divided into 6 scanning runs, with an average length of 9 minutes, with brief 1-2 minute breaks between blocks. The first 10 volumes of each run were discarded to account for equilibration of the scanner`s magnetic field. The experimental design is shown visually in **Figure 2**, and **Table S8** provides a schematic overview of trial types in each block.

[Figure 2 about here.]

In the first conditioning block of 48 trials, which lasted 8.4 minutes, participants were exposed to the initial CS - US contingencies, with equal number of CS+ and CS- trials. During Pavlovian

conditioning, one fractal stimulus (CS+) was associated with reward (orange juice) 75% of the time, while another fractal (CS-) was deterministically associated with a neutral outcome (artificial saliva, .0116g KCl, .0105g NaHCO3 per 500ML/water) (O'Doherty et al., 2003; Frank et al., 2012). This first conditioning block lasted 8.4 minutes.

This was followed by four conditioned inhibition blocks (blocks 2 -5), which consisted of 200 trials total, where the CS+ was paired with another stimulus, the conditioned inhibitor, in 22% of these trials (44 trials). Presentation of the conditioned inhibitor was deterministically associated with the presentation of the neutral solution, negating the reward prediction elicited by the CS+. In another 22% of total trials, participants continued to experience the initial CS - US pairing, with reward presented in 75% of these trials in order to keep the reward association from being extinguished by the conditioned inhibition procedure. The remaining trials consistent of several different neutral control stimuli (50% of total trials), and the conditioned inhibitor viewed alone (6% of total trials).

This conditioned inhibition training was broken into several different blocks, based on the same sequence of trials as the original Tobler et al. (2003) paper. The first block of conditioned inhibition lasted 5.7 minutes and consisted of 32 trials, 8 each of CS+, CS- and the CS+ paired with the Inhibitor, and an additional CS- control (consisting of two fractal images).

In blocks 3-5, which each lasted 10.4 minutes, and consisted of 168 total trials, participants saw 5 different stimuli, the CS+, CS-, CS+ paired with the Inhibitor, the CS- control consisting of two fractal images, the Inhibitor, and another CS- control consisting of a single fractal image. Each of these blocks consisted of 56 trials, 12 trials each of the CS+, CS-, CS+ paired with the Inhibitor, and the CS- control consisting of two images. Each block also included 4 trials where the Inhibitor was viewed alone and 4 trials of the CS- control consisting of a single fractal image.

The inhibitor was shown less frequently alone (1:3 ratio compared to other trials) to minimize learning about the inhibitor in isolation, which would have reduced the strength of the inhibitory procedure, as done in a previous conditioned inhibition study in monkeys (Tobler et al., 2003) which also included a block of conditioned inhibition before the inhibitor was viewed alone. The order and type of trials in each block was based off of the design in this original study.

In the final block, we ran an inhibition test block, which lasted 8.4 minutes and consisted of 48 trials. In the test block, we followed the conditioned inhibitor with an unexpected juice reward 75% of the time, in order to test for positive reward prediction errors. A control CS- was also paired with an unexpected juice reward 75% of the time in the conditioned inhibition test block, and we expected less of a prediction error signal for the CS- paired with reward than the inhibitor since it did not develop an association with reward omission.

In each trial, there was a presentation of a fractal conditioned stimulus for 2 seconds. This was followed by 2 mL of orange juice (Tropicana brand) or the neutral-tasting solution, which was delivered by a taste pump connected to the stimulus computer. The onset of taste delivery was logged, and delivery of solution after receiving the trigger took about 3 seconds. It was a delay conditioning paradigm, where the CS remained onscreen until the US delivered was completed. After each trial of CS and US presentation, there was an ITI randomly sampled to be between 4 and 8 seconds. We selected participants who reported a preference for orange juice in the prescreening interview. Further, presentation of orange juice has been associated with higher pleasantness ratings than artificial saliva in prior studies (Takemura et al., 2011). Visual stimuli were presented with a projector inside the fMRI head coil.

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 (McClure et al., 2003). 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 (Kronforst-Collins and Disterhoft, 1998), and we wanted to focus on the role of subcortical regions in conditioning. In addition, it is worth noting that the final block of conditioned inhibition, which we call an "inhibition test", was designed to replicate a specific condition in the Tobler et al. (2003). study which compared the prediction errors for the conditioned inhibitor followed by an unexpected reward with the prediction errors for a control stimulus. However, notably, behavioral tests of conditioned inhibition have suggested that two additional tests are important for assessing conditioned inhibition, an inhibitor should suppress responding to a CS+ when presented together (summation test), and also acquire conditioned excitatory properties more slowly when paired with a US in a retardation test (Rescorla, 1969b; Sosa and Ramírez, 2019).

2.3 Data Acquisition

Magnetic-resonance imaging (MRI) data were acquired at the Center for Innovation and Creativity at CU Boulder using a 3T Siemens Trio scanner and a 32-channel receive-only head coil. To guide the functional imaging, a structural volume of the entire brain was acquired first using a T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence (repetition time (TR): 2530 ms, echo time (TE1: 1.64ms, TE2: 3,5ms) , flip angle (FA): 7°, voxel: 1x1x1-mm isotropic, field of view (FOV): 2.29 x 2.29 x 2mm).

High-resolution functional images were acquired with a blood-oxygen-level-dependent (BOLD) contrast using a T2*-weighted gradient-echo echo-planar imaging (EPI) sequence (TR: 1300 ms, TE:

25 ms, 75%, acceleration factor: 2, 22cm FOV, in-plane voxel size: 2.29 mm, slice thickness 2mm, no gap (voxel-size: 2.29 x 2.29 x 2 mm). With these parameters, 24 contiguous slices were collected in interleaved-ascending order for each volume. Slices were aligned parallel to the base of the OFC. Due to the focus of our study on subcortical areas, we acquired limited coverage, which included the amygdala, insula, midbrain, thalamus, striatum, and ventral prefrontal cortex.

The functional imaging was divided into 6 scanning runs, with an average length of 9 minutes. The first 10 volumes of each run were discarded to account for equilibration of the scanner's magnetic field.

2.4 Monterary reward task: Follow up study

In order to investigate whether there were behavioral effects of conditioned inhibition, we ran a follow-up study to look at how the conditioned inhibition procedure affected reward expectation. The study had an identical design, but used monetary rewards, and allowed us to look at the behavioral effects of conditioned inhibition by having participants rate their expectation of reward at the end of each training block. As our behavioral measure of conditioning, we assessed the reward expectation for each stimulus at the end of each block using a continuous rating scale for reward expectation, ranging from No Expectation to Strongest Expectation. Subject saw each CS in succession, followed by the rating scale depicted in **Figure S5**, which was adapted from rating scales that have previously been validated for affective ratings across many modalities (Bartoshuk et al., 2004).

2.5 Preprocessing

The preprocessing pipeline followed a well-validated preprocessing pipeline that has been used in several other studies (Wager et al., 2013; Woo et al., 2014), and is available online (https://github.com/canlab/), but with a distinct warping step. The first 10 images were discarded to account for the stabilization of the BOLD signal. Then, the functional images were motion-corrected

using the realignment procedure in SPM 8, using a rigid-body, affine (6 parameter) registration that helps correct for head movement during scanning. To identify outliers, we computed the mean and standard deviation across voxels for each image for all slices, and then calculated the Mahalanobis distance of each mean and standard deviation value, considering any volumes with a significant χ^2 value as outliers, per the procedure described in Wager et al. (2013).

Next, these motion-corrected functional images were co-registered to the structural images using FSL's epi_reg script, an affine co-registration that improves registration by segmenting the structural and functional images (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Each structural T1 image was warped to standard space using the Advanced Normalization Toolbox (ANTs) (Avants et al., 2014). We then combined the transformation matrix from the functional to structural transformation with the warping matrix from the transformation of the structural to standard space to warp the functional data into standard space. After the transformation, a 4mm FWHM Gaussian smoothing kernel was applied to the images.

The functional images were corrected for slice timing to account for acquiring slices at slightly different timepoints and then motion corrected using the realignment procedure in SPM8. Each outlier image detected by the Mahalanobis distance method was modeled as a nuisance covariate, by inserting a dummy code variable of 1 where the spike occurred.

In addition, we calculated several regressors of non-interest, which included an intercept for each run, dummy regressors for outlier images calculated by the spike detection method above, and motion-related covariates, which included 6 mean-centered motion parameter estimates, their squared values, successive differences and squared successive differences. Additional nuisance regressors were calculated by determining the first five principal components from the signal in the ventricles in the warped functional images with a 4mm smoothing kernel.

2.6 fMRI analysis

The fMRI analysis involved separate regressors for each of the different stimuli in the experiment, including separate regressors for each of the stimuli in the conditioning, conditioned inhibition and inhibition test phases, to allow us to assess the effects of Pavlovian conditioning and conditioned inhibition on brain activity. To this end, we generated separate first-level model task regressors for the CS+ and the CS- in the first conditioning block. In the three following conditioned inhibition blocks, we created first-level model task regressors for the CS+, CS+ paired with the Inhibitor, and Inhibitor viewed alone, as well as the two other CS- stimuli. In the conditioned inhibition test blocks, we generated separate regressors for the Inhibitor and the CS-.

In the same first-level model, we also modeled each of the different outcomes following the CS with separate regressors, to allow us to examine the effects of expectation on outcome activity. Therefore, we generated separate regressors for the expected presentation of reward following the CS+, the unexpected reward omission resulting from presentation of the neutral solution following the CS+ (omission), and presentation of the neutral solution following each CS- (as a control stimulus for the effects of taste stimulation). In each case, the duration of each CS event was set to 2 seconds, while the duration of each US event was set to 3 seconds. The fixation cross, which was presented between each CS-US trial, was explicitly not modeled and considered the implicit baseline.

Further, to assess the effectiveness of the conditioned inhibition procedure, we ran a conditioned inhibition test where the inhibitor and the CS- control were unexpectedly paired with reward in the last block of the experiment. Therefore, the first level model also included separate regressors for both cue and outcome activity in this inhibition test block. If conditioned inhibition was successful and caused the inhibitor to acquire negative value, positive prediction errors should result when it is unexpectedly followed by a reward.

To specifically examine this, we looked separately at outcome activity when the Inhibitor was unexpectedly paired with reward and the trials where the Inhibitor was paired with no reward. Similarly, we also modeled the trials in the inhibition test block where the CS- was unexpectedly paired with reward separately from the trials where the CS- was followed by the expected neutral solution (no reward).

For the group level GLM analysis, we used a robust regression procedure, which has been shown to decrease sensitivity to outliers (Wager et al., 2005). Whole brain results were corrected for multiple comparisons with $q < .05$, FDR (false-discovery rate).

2.7 ROI analysis

We defined ROIs according to probabilistic atlases, whenever possible. For each anatomical atlas, we used a threshold to include only voxels that had 75% or higher probability. For the habenula ROI, we used the habenula ROI from a high-resolution atlas of the thalamus based on histological data (Krauth et al., 2010). Recent papers on defining the habenula in human fMRI suggest that total habenula volume (medial and lateral) is around 31-33mm, approximately the size of a single voxel in standard fMRI protocols (Lawson et al., 2013). For this reason, we cannot differentiate between medial and lateral habenula in the ROI analysis. For the SNc and VTA ROIs, we used a binary mask created from an anatomically specified ROI based on single-subject structural scans (Pauli et al., 2018), but not the structural images of the current sample.

The basolateral and centromedial amygdala ROIs were derived from the CIT atlas, which is in the same standard space as the functional images (Tyszka and Pauli, 2016). The anatomical ROIs of the caudate, palldium and putamen were derived from the Harvard-Oxford Subcortical Atlas (Smith et al., 2004).

To visualize results in our a-priori ROIs, results were corrected for multiple comparisons with $q < .05$, FDR (false-discovery rate) across a merged mask of all ROIs (OFC, Insula, Amygdala, Accumbens, Caudate, Pallidum and Putamen, SNc, VTA and Habenula). To create this mask, we included voxels with 75% or higher probability from each probabilistic subcortical atlas and all voxels from the bilateral Insula and bilateral Orbital Frontal Cortex atlas.

Additionally, we conducted comparisons of mean activity across different conditions in the a-priori ROIs. When performing ROI analyses, we looked at mean activity in each ROI across subjects, calculated based on the individual subject-level beta images for each condition from the first-level analysis. For tests of ROI activity, $p < .005$ (Bonferroni corrected for comparisons across 10 ROIs) was considered significant, and for each test, we report whether it exceeded the Bonferroni correction.

Results that did not exceed the FDR threshold across the mask of all ROIs or survive Bonferroni correction are reported for information only. For visualization purposes for ROI results in the basal ganglia, substantia nigra and habenula, we plot the results at an uncorrected threshold of $p < .001$ or $p < .005$. An additional result from the habenula ROI is shown at $q < .05$, FDR, small volume corrected with a binary mask of the ROI.

3 Results

3.1 Behavioral Results: Monetary Reward Task

In order to investigate the behavioral effects of conditioned inhibition, we ran a follow-up behavioral study using monetary rewards to look at how the conditioned inhibition procedure affected reward expectation. This study allowed us to examine the behavioral effects of conditioned inhibition by having participants rate their expectation of reward at the end of each training block. As outlined above, conditioned inhibition occurs if a conditioned inhibitor presented concurrently with a CS+ is

able to elicit a reduced conditioned response compared to the CS+ alone. To investigate whether there was a behavioral effect, we asked participants to rate how strongly they would expect reward on a continuous rating scale, ranging from No Expectation to Strongest Expectation after viewing the fractal stimulus. This behavioral test revealed that our conditioning procedure was successful, as mean ratings across blocks showed a significantly higher rating for the CS+ than the CS- [$t(18) = 11.84$, $p < .0001$]. We also found that the ratings for the CS+ when presented concurrently with the Inhibitor were significantly lower than the ratings for the CS+ alone [$t(18) = 7.07$, $p < .0001$], indicating that the conditioned inhibition procedure had significantly reduced reward expectations, demonstrating conditioned inhibition. See **Figure 2** for an illustration of the behavioral ratings for the CS+, CS- and CS+ paired with the Inhibitor.

[Figure 2 about here.]

3.2 fMRI Results

3.2.1 BOLD Responses to reward delivery

Based on prior studies, we expected that presentation of the juice reward would lead to activity in sensory regions associated with gustatory sensations, such as the insula, along with regions associated with reward outcomes, including juice rewards, such as the amygdala, OFC, midbrain, and striatum (O'Doherty et al., 2001; Kringelbach et al., 2003; O'Doherty et al., 2003; D'Ardenne et al., 2008; Frank et al., 2008; Metereau and Dreher, 2013; Pauli et al., 2015). We conducted a whole brain analysis to look at the effects of juice reward presentation compared to the neutral control solution. This comparison compared juice presentation (the expected presentation of reward following the CS+) with the presentation of the neutral control solution following each CS- cue. However, there were no significant voxels at $q < .05$ across the whole-brain mask.

We also conducted ROI analyses in a set of a-priori ROIs including the OFC, amygdala, insula and striatum. To visualize these results and correct for multiple comparisons across ROIs, we show activity that survived correction across the ROI mask in **Figure 3A and 3B**. Activity in the OFC [x=-28, y=36,z=-8, t=7.94, k=3] and insula ROIs [mm center x=38,y=-4,z=6, t=6.51, k=3] for juice compared to neutral solution survived FDR correction at $q < .05$ corrected across a mask of all ROIs, as shown in **Figure 3A and Figure 3B**. For visualization purposes, we also show activity at $p < .005$ and $p < .001$ uncorrected in these regions.

[Figure 3 about here.]

These regions are summarized in **Table 1**, among other key contrasts.

There was activity at a whole-brain uncorrected threshold of $p < .001$ in regions expected from prior studies of rewarding outcomes, including the insula, orbitofrontal cortex, basolateral amygdala and putamen for juice compared to the neutral solution, as shown in **Table S2**. Additional results from ROI analyses that did not exceed the Bonferroni-corrected p-value are reported in the Supplementary Information. However, we present this for information only, noting that it did not survive correction for multiple comparisons.

We further examined activity to an unexpected reward omission, examining the 25% of trials where the CS+ was unexpectedly followed by the neutral solution compared to neutral solution presentation following control trials (where it was expected). This revealed two peaks in the orbital frontal cortex [x=40,y=24,z=-10, k=8, t =5.83] and insula [x=48,y=18,z=-4, k=8, t=6.16] surviving FDR correction at $q < .05$ correction across the whole brain.

We also conducted analyses to compare the mean activity of voxels in several ROIs for juice reward presentation compared to presentation of the neutral solution, correcting for the number of ROIs used.

An ROI analysis of mean activity averaged across the caudate ROI showed significant activity for juice compared to neutral solution [$p = .0121$, Bonferroni corrected $p = .121$, $t(18)=2.79$ mm center L=14,12,10 and R=-12,10,10], which did not survive the Bonferroni correction. The activity in the caudate ROI did not survive FDR correction across the mask of ROIs at $q < .05$, FDR. There were two peaks within the caudate at an uncorrected threshold of $p < .005$, one located within dorsal caudate [mm center = 10,20,4, k=19, t=7.68, and the other in more ventral caudate [mm center 12,12,14, k=8, t=7.24], shown in **Figure 3C**. We present this for information only but not for making inference, as the caudate cluster did not survive multiple-comparisons correction.

Additionally, there was significant ROI activity in the central amygdala ROI for the juice reward compared to neutral solution [$p = .0033$, Bonferroni corrected $p = .033$, $t(18)=3.378$, mm center L=24,-8,-10, R=-24,-10,-12], which survived Bonferroni correction across all ROIs. For visualizaion purposes, this is shown at an uncorrected threshold of $p < .001$ and $p < .005$ in **Figure 3D**.

An ROI analysis of mean activity in the habenula ROI showed a significant deactivation for juice presentation compared to the neutral solution [$p = .0452$, $t(18)= -2.15$, mm center L= 4,-24,2, R=-2,24,2], however, this did not survive the Bonferroni corrected threshold. The activity in the habenula ROI did not survive FDR correction across the mask of ROIs at $q < .05$, FDR. For visualizaion purposes, this is shown at an uncorrected threshold of $p < .001$ and $p < .005$ in **Figure 3E**.

The substantia nigra and VTA ROIs did not show significant activity for juice compared to the neutral solution. All other comparisons of ROI activity that did not survive Bonferroni correction are reported in the supplementary material (**Section 1.1**), and summarized in **Table S1**. Results from the

ROI analysis for the juice compared to neutral solution contrast, along with ROI results from other contrasts, are summarized in **Table 1.**

[Table 1 about here.]

3.2.2 BOLD responses to CS presentations

We expected that a conditioned stimulus associated with reward would increase BOLD signals in the orbitofrontal, insular and ventromedial prefrontal cortical regions (Kim et al., 2011; Diekhof et al., 2012). As expected, we found activity in the bilateral insula for the CS+ compared to the CS- at $q < .05$, FDR, $k > 5$, corrected across the whole brain [x=30, y=26, z=0, t=8.31, k=11 and x=-30, y=22, z=4, k=5], consistent with other studies that have found activity in the insula for food reward cues (Tang et al., 2012). This is shown in **Figure 3A**.

Further, we conducted a focused ROI analysis of activity to the CS+ compared to the CS+, correcting across a merged mask of all ROIs. While there was activity in the orbital frontal cortex for a CS+ compared to a CS- [x=-28, y=18, z=-20, t=10.97, k=26, and x=28, y=18, z=-16, t=8.2, k=16], this activity was significant at $p < .001$, uncorrected, but did not survive the whole-brain corrected FDR threshold. However, activity in the OFC ROI survived FDR correction at $q < .05$, $k > 5$ voxels within the mask of all ROIs [x=-24, y=18,z=-20, t=5.91, k=8 and x=-34,y=20,z=-20, t=5.46, k=8], as shown in **Figure 3B**, along with activity at an uncorrected threshold of $p < .001$ and $p < .005$ for visualization. Additionally, there were two peaks in the insula at $q < .05$, FDR, corrected across the all-ROI mask [x=32,y=24,z=0, k=53, t=8.31 and x=-30, y=22, z=4, k = 37, t=11.8].

The regions that survived FDR correction across the mask of all ROIs for the CS+ > CS-, along with other contrasts, are summarized in **Table 1**.

Further peaks from the whole-brain threshold of $p < .001$ uncorrected for the CS+ compared to the CS- include insula, thalamus and midbrain as described in **Table S4**. At a threshold of $p < .005$ uncorrected, there was a cluster including the striatum (caudate) and extending to the pallidum [x=-12,y=8,z=2, t=9.71, k=25] for the CS+ compared to the CS-. However, this did not survive FDR correction across the mask of all ROIs, and is mentioned only for information only, noting that it did not survive correction for multiple comparisons.

We expected activity for conditioned stimuli associated with reward in the midbrain, amygdala and striatum (Breiter et al., 2001; O'Doherty et al., 2002; O'Doherty et al., 2006; Pauli et al., 2015). We next conducted an ROI analysis based on prior studies which found responses in the midbrain for predictors of a positive valenced reward (Adcock et al., 2006; O'Doherty et al., 2006; Pauli et al., 2015). As predicted, we found more activity in SNc [$t(18) = 3.17$, Bonferroni corrected $p = .053$, $p = .0053$, mm center L=8,-18,-14, R=-8,-20,-14] and VTA [$t(18) = 2.58$, $p = .0189$, Bonferroni corrected $p = .189$ mm center=0,-20,-16] for the CS+ than the CS-, as shown in **Figure 5**. While the CS+ > CS- effect did not exceed the Bonferroni corrected p-value threshold in the SNc or VTA ROIs, it was just below the margin of significance in the SNc. For visualization purposes, the CS+ > CS- effect in substantia nigra is shown in **Figure 4C** at $p < .005$. However, only a single voxel in this region [x=8,y=-14,z=-12, t=6.72, k=1] survived correction for multiple comparisons at $q < .05$, FDR across the mask of all ROIs. The CS+> CS- effect was only visible an a lower, uncorrected threshold of $p < .05$ in the VTA ROI, which is shown in **Figure S6** for visualization, but we note that it did not survive correction for multiple comparisons.

There was not significant ROI activity in the amygdala, nucleus accumbens, caudate, pallidum or putamen for the CS+ compared to the CS-.

[Figure 4 about here.]

3.2.3 BOLD responses to the conditioned inhibitor

We expected that the conditioned inhibitor would recruit activity in regions that have been shown to respond to predictors of reward omissions. However, we are unaware of other fMRI studies using a conditioned inhibition design with rewards, so it is unclear whether the same regions that have been shown to respond to predictors of monetary loss and aversive stimuli also respond to conditioned inhibitors, or predictors of reward omission. Conditioned inhibitors have never explicitly been followed by a negative valence outcome, but acquire negative value by reliably signaling a reward omission. Based on computational theories of learning such as TD, Rescorla-Wagner and PVLV(Rescorla, 1969a; Sutton and Barto, 1990; O'Reilly et al., 2007; Mollick et al., 2020), this occurs because the negative reward prediction errors that occur when a predicted reward is unexpectedly omitted cause the conditioned inhibitor that predicts reward omission to acquire negative value. We conducted a whole brain analysis for regions responding to the conditioned inhibitor compared to control stimuli, but did not find any regions that survived correction for multiple comparisons at $q < .05$, FDR.

While there is little research on brain areas that encode conditioned inhibitors in humans (though see Meyer et al. (2019) for a negative valence version), previous studies have shown that predictors of monetary loss are associated with BOLD activity in the insula (Samanez-Larkin et al., 2008).

Consistent with this data, we also saw significantly more mean activity in the bilateral insula ROI for the inhibitor than the control stimuli [$p = .0386$, $t(18)=2.23$, mm center L=38,4,0, R=-36,2,0]. However, this activity did not survive correction at FDR $q < .05$ across the mask of ROIs or Bonferroni correction for multiple comparisons. In the whole-brain analysis for the inhibitor compared to controls, there was activity in the insula at $p < .005$ uncorrected, as shown in **Figure S2**. We present this for information only, noting it did not survive correction for multiple comparisons.

While human fMRI studies have found activity in the habenula to predictors of aversive stimuli (Lawson et al., 2014) as well as aversive outcomes (Hennigan et al., 2015), and negative reward prediction error signals associated with reward omissions (Salas et al., 2010), it is unclear whether the habenula shows activity for predictors of reward omission in humans.

Based on animal studies, we predicted that the habenula would show an increase in activity for the conditioned inhibitor, as it showed an increase in activity for a CS that predicted omission of reward, accompanied by a reduction in SN/VTA activity for the Inhibitor (Tobler et al., 2003; Matsumoto and Hikosaka, 2009b). Consistent with this prediction, there was significant activity in the habenula for the conditioned inhibitor viewed alone compared to the mean activity for all control stimuli, including the CS- (B), the single stimulus neutral cue (Y-) and the compound stimulus (BY) neutral cue [$t(18) = 2.22$, $p = .0397$, Bonferroni corrected $p = .397$, mm center L= 4,-24,2, R=-2,-24,2]. However, this activity for the conditioned inhibitor did not survive the Bonferroni corrected threshold.

Further, this was not significant when the inhibitor was compared to only the neutral (Y-) control (consisting of a single cue shown at a similar rate to the inhibitor) that was always followed by the neutral solution [$t(18) = 1.119$, $p = .2779$, Bonferroni corrected > 1, mm center L= 4,-24,2, R=-2,-24,2]. Habenula activity for the conditioned inhibitor compared to the controls is shown at an uncorrected threshold of $p < .005$ and $p < .001$ in **Figure 5A**, along with 2 voxels surviving FDR correction at $q < .05$ [x=-2,y=-24,z=0, t=3.76, k=2], small-volume corrected with the habenula mask. However, this region did not appear when FDR correction was done across the mask of all ROIs, and therefore we strongly qualify this result, which awaits further replication before inference can be made.

[Figure 5 about here.]

To compare the role of the habenula and substantia nigra in our learning task, we compared activity in both ROIs, as shown in **Figure 6**. Consistent with the hypothesis that the substantia nigra encodes positive valence, activity in the substantia nigra increased for the CS+ paired with reward compared to the CS-, but not for the Inhibitor compared to control stimuli. We further found that there was significantly more activity in the substantia nigra for the CS+ > CS- effect than Inhibitor > Control comparison in the substantia nigra [p = .003452, Bonferroni corrected p = .03452, t=3.13]. Further, limited evidence pointed towards the habenula encoding negative valence, as it significantly increased for an Inhibitor paired with a reward omission, but not the positively valenced CS+, though this comparison did not survive Bonferroni correction. However, there was not a significant difference between the Inhibitor > Control and the CS+ > CS- effect in the habenula [p = .5281, Bonferroni corrected p > 1, t=-0.6371].

[Figure 6 about here.]

Along with the habenula, we also expected that regions of the basal ganglia would respond to the conditioned inhibitor associated with reward omissions. The ventral striatum has been shown to be activated by predictors of aversive stimuli (Jensen et al., 2003), and a cue associated with monetary loss activated a more posterior region of the ventral striatum (Seymour et al., 2007a). Further, animal studies have shown that pallidum communicates aversive expectations to habenula (Hong and Hikosaka, 2008), and studies showed that basal ganglia stimulation influenced habenula activity (Hong and Hikosaka, 2013). There was activity in the putamen region of striatum [x=20,y=4,z=-6, k=53, t(18)=11.29], which extended into the pallidum, for the inhibitor compared to control stimuli at p < .005, uncorrected, as shown in **Figure 5B**. We provide this for information only, but not for making inference as it did not survive correction for multiple comparisons. There was an increase in the mean ROI activity in the pallidum and putamen ROIs for the inhibitor compared to control

stimulus but this did not reach the significance threshold (see Supplementary Table S1 and Section 1.1).

Additionally, shown in **Figure 5C**, we ran a contrast comparing activity for the CS+ to that for the Inhibitor, which showed activity in the substantia nigra at $p < .005$ uncorrected, but this activity did not survive correction for multiple comparisons. Only a single voxel in the SNc survived correction across the mask of all ROIs at $q < .05$, FDR [x=12,y=-18,z=12, k=1, t= 6.72]. Further, an ROI analysis of mean ROI activity in the SNc showed more activity for the CS+ than the conditioned inhibitor [$p = .01$ uncorrected, Bonferroni corrected $p = .1$, $t(18)= 2.88$, mm center L=8,-18,-14, R=-8,-20,-14], but this did not survive Bonferroni correction for the number of ROIs.

If the inhibitor acquired a negative association, there should be a prediction error when the inhibitor is paired with a reward, resulting in activity in dopamine regions. However, an analysis looking at the mean activity in the VTA, SNc or Accumbens ROIs for the inhibitor followed by a reward did not show significant activity in the trials where the Inhibitor was followed by an unexpected reward.

There was a significant increase in activity in the putamen ROI during taste presentation when the inhibitor was unexpectedly followed by reward which survived whole-brain FDR correction at $q < .05$, FDR [x=32,y=-16.z=-4, t=7.16, k=5]. Further, an ROI analysis of mean activity in the putamen ROI showed an increase in the putamen ROI during taste presentation when the inhibitor was unexpectedly followed by reward, compared to when the inhibitor was followed by a neutral solution [$p = .0417$, $t(18)= 2.19$, mm center L=26,2,0, R=-26,2,0], but this did not exceed the threshold for Bonferroni correction.

An additional, more sensitive test of conditioned inhibition may be the comparison of responses to the inhibitor followed by reward to the control stimulus followed by reward, as the Tobler et al. (2003) study showed a larger response to the inhibitor followed by reward than the control stimulus

followed by reward. This may reflect a greater prediction error resulting from the unexpected reward presentation following the conditioned inhibitor compared to the control stimulus. Greater prediction errors when the inhibitor is followed by reward may occur because computational models such as the Rescorla-Wagner model suggest that the inhibitor acquired negative value through the conditioned inhibition procedure, compared to the control stimulus which has no inhibitory association and thus a smaller prediction error when unexpectedly followed by reward. However, when we compared these two conditions, there was not a significant difference in putamen mean ROI activity when inhibitor was unexpectedly followed by reward compared to when the control stimulus was unexpectedly followed by reward [$p = .3305$, Bonferroni corrected $p > 1$, $t=1.00$]. The regions that survived whole brain correction for the Inhibitor compared to Controls and the Inhibition test are summarized in **Table 1**. Notably, only the signal for the inhibitor paired with reward survived correction for multiple comparisons across the whole brain.

Based on our discussion of the potential of conditioned inhibition to dissociate between representations of associations of a CS+ with reward and the representation of an inhibitor with reward omissions, we compared the activity for the CS+ and Inhibitor to that of the Inhibitor viewed alone. This should reveal regions that selectively reflected the associations of the CS+ with reward. 3 clusters in visual cortex, including the lingual gyrus, showed more activity for the CS+ and Inhibitor compared to the Inhibitor alone, as described in **Table S7.** However, there were also important visual differences between the CS+ and Inhibitor, as the CS+ was represented by a single visual cue and the Inhibitor was represented by two visual cues. Therefore, it is difficult to interpret whether the visual cortical regions reported above reflect visual differences between the cues or the association of the CS+ with reward. Further, it is possible that activity in this region may also reflect the conjunction of the combined stimulus consisting of the CS+ and the Inhibitor.

4 Discussion

563 Recent research has suggested that the lateral habenula drives dopamine dips for aversive stimuli and
564 reward omissions, so we expected a selective activation of the habenula for the conditioned inhibitor,
565 paired with a reduction in SN/VTA activity. Consistent with these predictions, we found significant
566 activity at an uncorrected threshold in habenula for a conditioned inhibitor associated with reward
567 omission compared to the mean of all control stimuli (but not when compared to the second control
568 stimulus). However, the activity in the habenula for the conditioned inhibitor did not survive FDR
569 correction across the mask of all ROIs, and the test of mean ROI activity for the inhibitor compared
570 to controls in habenula did not exceed the Bonferroni-corrected p-value threshold, and thus should
571 not be strongly interpreted. While other studies have found activity in habenula for predictors of
572 aversive outcomes, such as shock (Lawson et al., 2014), or an aversive bitter juice outcome
573 (Hennigan et al., 2015), and one study found habenula activity during the omission of an expected
574 reward (Salas et al., 2010), none have shown that a reward omission stimulus, or conditioned
575 inhibitor, drives habenula activity in humans. The habenula signals we found for a conditioned
576 inhibitor are consistent with a recent animal study (Laurent et al., 2017), which found that projections
577 from the habenula to the RMTg, the tail region of the VTA, which sends inhibitory connections to
578 dopamine neurons (Bourdy and Barrot, 2012) were crucial for the effects of a conditioned inhibitor
579 on choice. One limitation of our study is that, due to the resolution of standard fMRI data, our use of
580 smoothing, and the small size of habenula (Lawson et al., 2013), we cannot differentiate medial from
581 lateral habenula. This limits our ability to directly relate to the lateral habenula signals observed in
582 animal studies. Further, while there was activity in the habenula for the inhibitor compared to the
583 mean of all control stimuli, this was not significant when the inhibitor was compared to a single
584 control stimulus, and the Inhibitor > Controls effect was not significantly different than the CS+ >
585 CS- effect. These may speak to a lack of power due to our small sample size, and the small effect size
586 of the habenula findings await future replication before inferences can be made.

Though increased activity in the habenula is associated with a dip, or pause in tonic dopamine firing (Matsumoto and Hikosaka, 2009b), we did not see a significant reduction in SN/VTA activity during presentation of the conditioned inhibitor. While we did not see a significant decrease in substantia nigra or VTA activity for the conditioned inhibitor compared to a neutral CS-, few studies have actually shown a significant decrease in BOLD in dopaminergic areas during a negative reward prediction error. For example, D'Ardenne et al. (2008) did not see a significant decrease in SN/VTA activity when an expected reward was omitted, and Rutledge et al. (2010) similarly did not find signals consistent with reward prediction error encoding in these midbrain areas.

One potential reason that we did not see significant reductions in BOLD signals for the conditioned inhibitor could be related to the physiology of the midbrain dopamine system. For example, inhibitory synaptic input has been shown to increase BOLD signals (Logothetis, 2008), and it is possible that inhibitory signals during reward omissions are conveyed from the lateral habenula to GABAergic neurons in the RMTg (which inhibit the SN/VTA). Further, these inhibitory neurons are spatially close to dopaminergic neurons and may not be spatially resolvable with the resolution of fMRI (Düzel et al., 2009). Such signals could potentially explain cases where SN/VTA activity increased for an aversive stimulus, for example, Pauli et al. (2015) found that neurons in the SN showed an aversive value signal, and Hennigan et al. (2015) found activation of the SN for a shock stimulus. Another potential explanation is that the inhibitor signaled the omission of the expected reward, which was a salient event, and activity in dopamine neurons as well as BOLD signals in human studies have been associated with salient and novel events (Horvitz, 2000; Matsumoto and Hikosaka, 2009a; Bromberg-Martin et al., 2010b; Krebs et al., 2011; Richter et al., 2020).

We also replicated previous studies, which showed activation in the SN/VTA area for a rewarding outcome, and studies showing activity in the SN/VTA area for a CS+ paired with reward (O'Doherty et al., 2002). While we expected signals in the striatum and amygdala during anticipation of the juice

611 reward (O'Doherty et al., 2002), we did not see significant amygdala signals for the CS+ compared to
612 CS-, and the striatal activity during reward anticipation did not survive the whole-brain corrected
613 threshold.

614 While some studies have found activations in ventral striatum for pleasant taste presentation (Frank
615 et al., 2008), other studies found more dorsal regions of striatum (O'Doherty et al., 2001; McClure et
616 al., 2003; Frank et al., 2012; Hennigan et al., 2015), or did not observe striatal activity for taste
617 presentation (O'Doherty et al., 2002). We found that regions of the basal ganglia were involved in
618 learning about reward, as the caudate showed activity during presentation of the juice reward
619 compared to a neutral solution, consistent with other studies (O'Doherty et al., 2002), but this activity
620 did not survive FDR correction across the mask of ROIs. We also observed activity in the putamen
621 for juice compared to the neutral solution, but this did not survive correction for multiple
622 comparisons. Activity in the dorsal striatum, including the dorsal caudate, has been correlated with
623 pleasantness ratings (Small et al., 2003), and putamen activity has also been associated with the
624 subjective feeling of appetite (Porubská et al., 2006).

625 We also predicted that regions of the basal ganglia send signals to the lateral habenula encoding the
626 level of reward expectation, allowing it to drive a dopamine dip if an expected reward is not received.
627 While there was activity pallidum and putamen for the inhibitor compared to a control stimulus that
628 was significant at a whole-brain uncorrected threshold, this activity did not survive correction for
629 multiple comparisons and should not be strongly interpreted. As the behavioral ratings from the
630 monterary reward task demonstrated that the inhibitor in that study acquired negative value, and the
631 imaging study found that the inhibitor led to activity in the pallidum, this is consistent with another
632 study that observed pallidal activity increasing with the negative value of a shock cue (Lawson et al.,
633 2014). While animal studies have shown that the pallidum encodes both positive and negatively
634 valenced outcomes (Tachibana and Hikosaka, 2012) and signals about reward and punishment pass

through the globus pallidus border region to drive activity in the habenula (Hong and Hikosaka, 2008), future studies are needed to understand how these computations are reflected in BOLD signals during reward omission learning, particularly given that the striatal peaks for the inhibitor did not survive correction for multiple comparisons. As with the habenula results, the striatal peaks for the inhibitor await further replication before inferences can be made.

While we also saw a non-significant increase in BOLD activity for the CS+ in the habenula, such differences could potentially be explained by the complexities of mapping neuronal spiking in this region to BOLD signals, and reduced power due to the sample size. Several studies have shown that a CS+ associated with reward decreases neural firing in habenula neurons compared to cues associated with reward omissions (Matsumoto and Hikosaka, 2009b; Bromberg-Martin et al., 2010a). Further, stimulating the output pathway from the habenula led to a decrease in motivational salience to a CS+, indexed by approach behaviors (Danna et al., 2013), while decreasing habenula output lead to an increase in motivational salience, consistent with the idea that activity in habenula projection neurons decreases for reward cues. However, as discussed in our interpretation of midbrain signals, inhibitory synaptic input has been shown to increase BOLD signals (Logothetis, 2008), and in some cases, inhibitory neurotransmission may also lead to increases in metabolic activity that could increase BOLD. If the reward CS+ led to activity in inhibitory input regions projecting to habenula such as the basal ganglia or pallidum (Hong and Hikosaka, 2008; 2013), or inhibitory neurotransmission in habenula neurons inhibited by reward led to an increase in metabolic activity, this could potentially cause increases in BOLD signals to a CS+. Additionally, Bromberg-Martin et al. (2010a) showed increases in habenula activity to both appetitive and aversive cues at the start of a trial, though these same neurons clearly differentiated between a CS+ and the CS- associated with no reward during other parts of the trial.

Further, we observed activity in the putamen surviving correction for multiple comparisons when the conditioned inhibitor was unexpectedly followed by a juice reward in the inhibition test at the end of the experiment. This may reflect a prediction error if the conditioned inhibition procedure caused the inhibitor to acquire negative value, consistent with other studies that have found putamen regions respond to prediction error signals (O'Doherty et al., 2003; Seymour et al., 2004). However, when we conducted an additional test comparing the magnitude of putamen ROI signals for the conditioned inhibitor unexpectedly followed by reward compared to the Control stimulus followed by reward, we did not find a significant difference, even though the Tobler et al. (2003) study observed a stronger response in dopamine neurons to the conditioned inhibitor followed by reward that the control stimulus followed by reward. This may be related to a lack of temporal resolution in our study, as the cues occurred 2 seconds before the responses to outcomes and may have been difficult to resolve from the outcome activity. In addition, by a prediction error encoding account, responses to the cues may have driven the opposite response, with the inhibitor resulting in less activity than the control stimulus.

Along with subcortical regions, we found that the orbital frontal cortex and anterior insula showed involvement in the reward learning task. We replicated prior studies showing activation of the anterior insula for taste stimuli (Nitschke et al., 2006; O'Doherty et al., 2006; Frank et al., 2012). We also observed activity in orbitofrontal cortex for the receipt of the taste stimulus, consistent with other studies (O'Doherty et al., 2001; O'Doherty et al., 2002). Further, we saw activity surviving the whole-brain corrected threshold in the orbitofrontal cortex for a conditioned stimulus associated with a reward, consistent with other studies that have shown activity in orbitofrontal cortex for a conditioned stimulus that predicted reward presentation (Gottfried et al., 2002; Kim et al., 2006). We further saw a signal in the orbital frontal cortex for the negative reward prediction error condition resulting when the CS+ was unexpectedly followed by neutral solution. This finding is consistent

682 with animal data which has shown that the orbital frontal cortex may be particularly important for
683 driving dopamine dip signals for worse than expected outcomes, as dopamine neurons no longer
684 showed that a reduction in firing for an unexpected reward omission when OFC was lesioned
685 (Takahashi et al., 2011). Additionally, a study applying conditioned inhibition in a negatively
686 valanced domain found that children with anxiety disorders represented safety signals (conditioned
687 inhibitors of fear) differently in the vmPFC than children without anxiety (Harrewijn et al., 2020).

688 We also observed activity for the conditioned inhibitor in the anterior insula, but only at a whole-
689 brain uncorrected threshold. This is interesting due to other papers which have suggested a role for
690 the insula in safety signal processing in the aversive domain (Christianson et al., 2008). Activity in
691 the insula has also been related to loss anticipation, as it increases to predictors of loss (Samanez-
692 Larkin et al., 2008) and loss aversion in decision making (Fukunaga et al., 2012), and is related to
693 individual differences in avoidance learning (Paulus et al., 2003). As the CS+ also showed activity in
694 the insula surviving whole-brain correction, there was not selective activation in this region for the
695 inhibitor, and insula has also been associated with positive valence food and drug cues (Tang et al.,
696 2012). The increases in insula activity we saw for both the positively valenced CS+ and the
697 negatively valenced conditioned inhibitor are in agreement with papers that have found evidence of
698 "salience" or an unsigned prediction error in the insula (Rutledge et al., 2010). Notably, activity in
699 the insula also showed a significant increase surviving whole-brain correction during trials where the
700 CS+ was unexpectedly followed by the neutral solution, trials which lead to negative reward
701 prediction errors.

702 Conditioned inhibition provides an interesting way to examine the functioning of the dopamine
703 system and can be used to look at how different brain regions are involved in learning about a
704 conditioned inhibitor that predicts not getting reward. It allows comparing the brain activity for
705 stimuli associated with reward predictors to those associated with reward omissions and is interesting

for examining how the subcortical areas projecting to the dopamine system, which have been primarily studied in animal learning studies, translate to humans in an fMRI task. Additionally, understanding the brain areas that drive this frustration signal for reward omissions can be translated to understand how disorders that involve persistent negative predictions, such as depression, may involve distortions in these systems. For example, recent research suggests that punishment prediction errors in the lateral habenula correlates with symptoms of depression (Kumar et al., 2018), and future studies could examine whether this relationship extends to reward omission cues. Generally speaking, the neural mechanisms of disappointment or frustration signals involved in conditioned inhibition are understudied relative to rewards, but further understanding of these signals has great translational and clinical relevance; for example, recent animal data indicates that cocaine use impairs the ability of dopamine neurons to suppress firing during omission of an expected reward (Takahashi et al., 2019), and recent human studies have found changes in negative reward prediction error signals in cocaine addiction (Parvaz et al., 2015).

5 Conflict of Interest

Randall C. O'Reilly is CSO, Seth A. Herd is CEO, and Jessica A. Mollick, Thomas E. Hazy, Ananta Nair, and Kai A. Kruger are researchers at eCortex, Inc., Boulder, Colorado, which may derive indirect benefit from the work presented here.

6 Author Contributions

J.M., T.D.W., R.O.R, and G.F. contributed to the conception and design of the study. J.M performed statistical analyses of the data. A.K., L.C., K.K, T.H, T.D.W, G.F. and R.O.R contributed to interpretation and discussion of neuroimaging and behavioral results. J.M. wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

7 Funding

Supported by: NIH IBSC Center 1-P50-MH079485 and NIH R01GM109996.

8 Acknowledgments

Thanks to Wolfgang Pauli for helpful comments.

9 Data Availability Statement

All data used for the ROI analysis figures and behavioral rating data can be found in the Open Science Framework: https://osf.io/njbmf/. Neuroimaging data used for the figures of second-level analysis results can be found on Neurovault: https://neurovault.org/collections/8676/.

10 References

Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J.D.E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50(3), 507-517.

Avants, B.B., Tustison, N.J., Stauffer, M., Song, G., Wu, B., and Gee, J.C. (2014). The Insight ToolKit image registration framework. *Frontiers in Neuroinformatics* 8. doi: 10.3389/fninf.2014.00044.

Bartoshuk, L.M., Duffy, V.B., Green, B.G., Hoffman, H.J., Ko, C.W., Lucchina, L.A., et al. (2004). Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiology & Behavior* 82(1), 109-114. doi: 10.1016/j.physbeh.2004.02.033.

Bourdy, R., and Barrot, M. (2012). A new control center for dopaminergic systems: pulling the VTA by the tail. *Trends in neurosciences*.

Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron* 30(2), 619-639. doi: 10.1016/S0896-6273(01)00303-8.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010a). Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron* 67(1), 144-155.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010b). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68(5), 815-834.

Christianson, J.P., Benison, A.M., Jennings, J., Sandsmark, E.K., Amat, J., Kaufman, R.D., et al. (2008). The Sensory Insular Cortex Mediates the Stress-Buffering Effects of Safety Signals But Not Behavioral Control. *Journal of Neuroscience* 28(50), 13703-13711. doi: 10.1523/JNEUROSCI.4270-08.2008.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482(7383), 85-88. doi: 10.1038/nature10754.

D'Ardenne, K., McClure, S.M., Nystrom, L.E., and Cohen, J.D. (2008). BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science* 319(5867), 1264-1267. doi: 10.1126/science.1150605.

Danna, C.L., Shepard, P.D., and Elmer, G.I. (2013). The habenula governs the attribution of incentive salience to reward predictive cues. *Frontiers in human neuroscience* 7, 781.

Diekhof, E.K., Kaps, L., Falkai, P., and Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – An activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia* 50(7), 1252-1266. doi: 10.1016/j.neuropsychologia.2012.02.007.

Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B.H., and Tobler, P.N. (2009). Functional imaging of the human dopaminergic midbrain. *Trends in neurosciences* 32(6)**,** 321-328.

Frank, G.K.W., Oberndorfer, T.A., Simmons, A.N., Paulus, M.P., Fudge, J.L., Yang, T.T., et al. (2008). Sucrose activates human taste pathways differently from artificial sweetener. *NeuroImage* 39(4)**,** 1559-1569. doi: 10.1016/j.neuroimage.2007.10.061.

Frank, G.K.W., Reynolds, J.R., Shott, M.E., Jappe, L., Yang, T.T., Tregellas, J.R., et al. (2012). Anorexia Nervosa and Obesity are Associated with Opposite Brain Reward Response. *Neuropsychopharmacology* 37(9)**,** 2031-2046. doi: 10.1038/npp.2012.51.

Fukunaga, R., Brown, J.W., and Bogg, T. (2012). Decision making in the Balloon Analogue Risk Task (BART): Anterior cingulate cortex signals loss aversion but not the infrequency of risky choices. *Cognitive, Affective, & Behavioral Neuroscience* 12(3)**,** 479-490. doi: 10.3758/s13415-012-0102-1.

Gottfried, J.A., O'Doherty, J., and Dolan, R.J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *Journal of Neuroscience* 22(24)**,** 10829-10837.

Harrewijn, A., Kitt, E.R., Abend, R., Matsumoto, C., Odriozola, P., Winkler, A.M., et al. (2020). Comparing neural correlates of conditioned inhibition between children with and without anxiety disorders–A preliminary study. *Behavioural Brain Research***,** 112994.

Hennigan, K., D'Ardenne, K., and McClure, S.M. (2015). Distinct midbrain and habenula pathways are involved in processing aversive events in humans. *The Journal of Neuroscience* 35(1)**,** 198-208.

Hong, S., and Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral habenula. *Neuron* 60(4)**,** 720-729. doi: 10.1016/j.neuron.2008.09.035.

Hong, S., and Hikosaka, O. (2013). Diverse sources of reward value signals in the basal ganglia nuclei transmitted to the lateral habenula in the monkey. *Frontiers in human neuroscience* 7(November)**,** 778-778. doi: 10.3389/fnhum.2013.00778.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96(4)**,** 651-656.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage; Amsterdam* 17(2)**,** 825-841. doi: http://dx.doi.org/10.1006/nimg.2002.1132.

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 5(2)**,** 143-156. doi: 10.1016/S1361-8415(01)00036-6.

Jensen, J., McIntosh, A.R., Crawley, A.P., Mikulis, D.J., Remington, G., and Kapur, S. (2003). Direct Activation of the Ventral Striatum in Anticipation of Aversive Stimuli. *Neuron* 40(6)**,** 1251-1257. doi: 10.1016/S0896-6273(03)00724-4.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2006). Is Avoiding an Aversive Outcome Rewarding? Neural Substrates of Avoidance Learning in the Human Brain. *PLOS Biology* 4(8)**,** e233. doi: 10.1371/journal.pbio.0040233.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2011). Overlapping Responses for the Expectation of Juice and Money Rewards in Human Ventromedial Prefrontal Cortex. *Cerebral Cortex* 21(4)**,** 769-776. doi: 10.1093/cercor/bhq145.

816 Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., Morel, A., and Székely, G. (2010). A mean three-
817 dimensional atlas of the human thalamus: generation from multiple histological data.
818 *Neuroimage* 49(3)**,** 2053-2062.

819 Krebs, R.M., Heipertz, D., Schuetze, H., and Duzel, E. (2011). Novelty increases the mesolimbic
820 functional connectivity of the substantia nigra/ventral tegmental area (SN/VTA) during
821 reward anticipation: Evidence from high-resolution fMRI. *Neuroimage* 58(2)**,** 647-655.

822 Kringelbach, M.L., O'Doherty, J., Rolls, E.T., and Andrews, C. (2003). Activation of the human
823 orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness.
824 *Cerebral cortex (New York, N.Y. : 1991)* 13(10)**,** 1064-1071.

825 Kronforst-Collins, M.A., and Disterhoft, J.F. (1998). Lesions of the caudal area of rabbit medial
826 prefrontal cortex impair trace eyeblink conditioning. *Neurobiology of Learning and Memory*
827 69**,** 147-162.

828 Kumar, P., Goer, F., Murray, L., Dillon, D.G., Beltzer, M.L., Cohen, A.L., et al. (2018). Impaired
829 reward prediction error encoding and striatal-midbrain connectivity in depression.
830 *Neuropsychopharmacology* 43(7)**,** 1581. doi: 10.1038/s41386-018-0032-x.

831 Laurent, V., Wong, F.L., and Balleine, B.W. (2017). The lateral habenula and its input to the
832 rostromedial tegmental nucleus mediates outcome-specific conditioned inhibition. *Journal of*
833 *Neuroscience***,** 3415-3416. doi: 10.1523/JNEUROSCI.3415-16.2017.

834 Lawson, R.P., Drevets, W.C., and Roiser, J.P. (2013). Defining the habenula in human neuroimaging
835 studies. *NeuroImage* 64**,** 722-727. doi: 10.1016/j.neuroimage.2012.08.076.

836 Lawson, R.P., Seymour, B., Loh, E., Lutti, A., Dolan, R.J., Dayan, P., et al. (2014). The habenula
837 encodes negative motivational value associated with primary punishment in humans.
838 *Proceedings of the National Academy of Sciences* 111(32)**,** 11858-11863. doi:
839 10.1073/pnas.1323586111.

840 Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453(7197)**,** 869.

841 Maia, T.V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learning &*
842 *Behavior* 38(1)**,** 50-67. doi: 10.3758/LB.38.1.50.

843 Matsumoto, M., and Hikosaka, O. (2009a). Two types of dopamine neuron distinctly convey positive
844 and negative motivational signals. *Nature* 459(7248)**,** 837-841.

845 Matsumoto, O., and Hikosaka, M. (2009b). Representation of negative motivational value in the
846 primate lateral habenula. *Nature Neuroscience* 12(1)**,** 77-84.

847 McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal Prediction Errors in a Passive
848 Learning Task Activate Human Striatum. *Neuron* 38(2)**,** 339-346. doi: 10.1016/S0896-
849 6273(03)00154-5.

850 Metereau, E., and Dreher, J.-C. (2013). Cerebral Correlates of Salient Prediction Error for Different
851 Rewards and Punishments. *Cerebral Cortex* 23(2)**,** 477-487. doi: 10.1093/cercor/bhs037.

852 Meyer, H.C., Odriozola, P., Cohodes, E.M., Mandell, J.D., Li, A., Yang, R., et al. (2019). Ventral
853 hippocampus interacts with prelimbic cortex during inhibition of threat response via learned
854 safety in both mice and humans. *Proceedings of the National Academy of Sciences* 116(52)**,**
855 26970-26979. doi: 10.1073/pnas.1910481116.

856 Mollick, J.A., Hazy, T.E., Krueger, K.A., Nair, A., Mackie, P., Herd, S.A., et al. (2020). A systems-
857 neuroscience model of phasic dopamine. *Psychological Review*.

Mowrer, O.H. (1956). Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. *Psychological review* 63.

Nitschke, J.B., Dixon, G.E., Sarinopoulos, I., Short, S.J., Cohen, J.D., Smith, E.E., et al. (2006). Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nature Neuroscience* 9(3)**,** 435-442. doi: 10.1038/nn1645.

O'Doherty, J., Rolls, E.T., Francis, S., Bowtell, R., and McGlone, F. (2001). Representation of Pleasant and Aversive Taste in the Human Brain. *Journal of Neurophysiology* 85(3)**,** 1315-1321.

O'Doherty, J.P., Buchanan, T.W., Seymour, B., and Dolan, R.J. (2006). Predictive Neural Coding of Reward Preference Involves Dissociable Responses in Human Ventral Midbrain and Ventral Striatum. *Neuron* 49(1)**,** 157-166. doi: 10.1016/j.neuron.2005.11.014.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron* 38(2)**,** 329-337. doi: 10.1016/S0896-6273(03)00169-7.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural Responses during Anticipation of a Primary Taste Reward. *Neuron* 33(5)**,** 815-826. doi: 10.1016/S0896-6273(02)00603-7.

O'Reilly, R.C., Frank, M.J., Hazy, T.E., and Watz, B. (2007). PVLV: the primary value and learned value Pavlovian learning algorithm. *Behavioral neuroscience* 121(1)**,** 31-49. doi: 10.1037/0735-7044.121.1.31.

Parvaz, M.A., Konova, A.B., Proudfit, G.H., Dunning, J.P., Malaker, P., Moeller, S.J., et al. (2015). Impaired Neural Response to Negative Prediction Errors in Cocaine Addiction. *Journal of Neuroscience* 35(5)**,** 1872-1879. doi: 10.1523/JNEUROSCI.2777-14.2015.

Pauli, W.M., Larsen, T., Collette, S., Tyszka, J.M., Seymour, B., and O'Doherty, J.P. (2015). Distinct Contributions of Ventromedial and Dorsolateral Subregions of the Human Substantia Nigra to Appetitive and Aversive Learning. *Journal of Neuroscience* 35(42)**,** 14220-14233. doi: 10.1523/JNEUROSCI.2277-15.2015.

Pauli, W.M., Nili, A.N., and Tyszka, J.M. (2018). A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. *Scientific Data* 5**,** 180063. doi: 10.1038/sdata.2018.63.

Paulus, M.P., Rogalsky, C., Simmons, A., Feinstein, J.S., and Stein, M.B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage* 19(4)**,** 1439-1448. doi: 10.1016/S1053-8119(03)00251-9.

Porubská, K., Veit, R., Preissl, H., Fritsche, A., and Birbaumer, N. (2006). Subjective feeling of appetite modulates brain activity: An fMRI study. *NeuroImage* 32(3)**,** 1273-1280. doi: 10.1016/j.neuroimage.2006.04.216.

Redish, a.D., Jensen, S., Johnson, A., and Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychological review* 114(3)**,** 784-805. doi: 10.1037/0033-295X.114.3.784.

Rescorla, R.A. (1969a). Conditioned inhibition of fear resulting from negative {CS-US} contingencies. *Journal of comparative and physiological psychology* 67(4).

Rescorla, R.A. (1969b). Pavlovian conditioned inhibition. *Psychological Bulletin* 72(2)**,** 77.

Richter, A., Reinhard, F., Kraemer, B., and Gruber, O. (2020). A high-resolution fMRI approach to characterize functionally distinct neural pathways within dopaminergic midbrain and nucleus accumbens during reward and salience processing. *European Neuropsychopharmacology*.

Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G.E., and Wager, T.D. (2014). Representation of aversive prediction errors in the human periaqueductal gray. *Nature Neuroscience* 17(11)**,** 1607-1612. doi: 10.1038/nn.3832.

Rudy, J.W., and O'Reilly, R.C. (2001). Conjunctive representations, the hippocampus, and contextual fear conditioning. *Cognitive, affective & behavioral neuroscience* 1(1)**,** 66-82.

Rutledge, R.B., Dean, M., Caplin, A., and Glimcher, P.W. (2010). Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience* 30(40)**,** 13525-13536. doi: 10.1523/JNEUROSCI.1747-10.2010.

Salas, R., Baldwin, P., de Biasi, M., and Montague, P.R. (2010). BOLD Responses to Negative Reward Prediction Errors in Human Habenula. *Frontiers in human neuroscience* 4.

Samanez-Larkin, G.R., Hollon, N.G., Carstensen, L.L., and Knutson, B. (2008). Individual Differences in Insular Sensitivity During Loss Anticipation Predict Avoidance Learning. *Psychological Science* 19(4)**,** 320-323. doi: 10.1111/j.1467-9280.2008.02087.x.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in neurosciences* 30(5)**,** 203-210.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007a). Differential Encoding of Losses and Gains in the Human Striatum. *The Journal of Neuroscience* 27(18)**,** 4826-4831. doi: 10.1523/JNEUROSCI.0400-07.2007.

Seymour, B., Doherty, J.P.O., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., et al. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429(June)**,** 664-667. doi: 10.1038/nature02636.1.

Seymour, B., Singer, T., and Dolan, R. (2007b). The neurobiology of punishment. *Nature Reviews Neuroscience* 8(4)**,** 300-311. doi: 10.1038/nrn2119.

Small, D.M., Jones-Gotman, M., and Dagher, A. (2003). Feeding-induced dopamine release in dorsal striatum correlates with meal pleasantness ratings in healthy human volunteers. *NeuroImage* 19(4)**,** 1709-1715. doi: 10.1016/S1053-8119(03)00253-2.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23**,** S208-S219. doi: 10.1016/j.neuroimage.2004.07.051.

Solomon, R.L., and Corbit, J.D. (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychological review* 81(2)**,** 119-145.

Sosa, R., and Ramírez, M.N. (2019). Conditioned inhibition: Historical critiques and controversies in the light of recent advances. *Journal of Experimental Psychology: Animal Learning and Cognition* 45(1)**,** 17.

Sutton, R.S., and Barto, A.G. (1990). "Time-Derivative Models of Pavlovian Reinforcement," in *Learning and Computational Neuroscience,* eds. J.W. Moore & M. Gabriel. (Cambridge, MA: MIT Press), 497-537.

Tachibana, Y., and Hikosaka, O. (2012). The Primate Ventral Pallidum Encodes Expected Reward Value and Regulates Motor Action. *Neuron* 76(4)**,** 826-837. doi: 10.1016/j.neuron.2012.09.030.

942 Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'donnell, P., Niv, Y., et al. (2011).
943 Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex.
944 *Nature neuroscience* 14(12)**,** 1590-1597.

945 Takahashi, Y.K., Stalnaker, T.A., Marrero-Garcia, Y., Rada, R.M., and Schoenbaum, G. (2019).
946 Expectancy-Related Changes in Dopaminergic Error Signals Are Impaired by Cocaine Self-
947 Administration. *Neuron* 101(2)**,** 294-306.e293. doi: 10.1016/j.neuron.2018.11.025.

948 Takemura, H., Samejima, K., Vogels, R., Sakagami, M., and Okuda, J. (2011). Stimulus-Dependent
949 Adjustment of Reward Prediction Error in the Midbrain. *PLOS ONE* 6(12)**,** e28337. doi:
950 10.1371/journal.pone.0028337.

951 Tang, D.W., Fellows, L.K., Small, D.M., and Dagher, A. (2012). Food and drug cues activate similar
952 brain regions: A meta-analysis of functional MRI studies. *Physiology & Behavior* 106(3)**,**
953 317-324. doi: 10.1016/j.physbeh.2012.03.009.

954 Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by
955 dopamine neurons in a conditioned inhibition paradigm. *Journal of Neuroscience* 23**,** 10402-
956 10410.

957 Tyszka, J.M., and Pauli, W.M. (2016). In vivo delineation of subdivisions of the human amygdaloid
958 complex in a high-resolution group template. *Human Brain Mapping* 37(11)**,** 3979-3998. doi:
959 10.1002/hbm.23289.

960 Wager, T.D., Atlas, L.Y., Lindquist, M.A., Roy, M., Woo, C.-W., and Kross, E. (2013). An fMRI-
961 based neurologic signature of physical pain. *New England Journal of Medicine* 368(15)**,**
962 1388-1397.

963 Wager, T.D., Keller, M.C., Lacey, S.C., and Jonides, J. (2005). Increased sensitivity in neuroimaging
964 analyses using robust regression. *NeuroImage* 26(1)**,** 99-113. doi:
965 10.1016/j.neuroimage.2005.01.011.

966 Woo, C.-W., Koban, L., Kross, E., Lindquist, M.A., Banich, M.T., Ruzic, L., et al. (2014). Separate
967 neural representations for physical pain and social rejection. *Nature Communications* 5**,**
968 ncomms6380. doi: 10.1038/ncomms6380.

969

970

971

Table 1: Summary of results across contrasts; regions that survived whole-brain correction, either at whole-brain FDR corrected threshold, or with FDR correction across a mask of all ROIs or p < .001 or p < .005 in a-priori ROIs. * Note that habenula activity for Inhibitor > Controls did not survive correction across the mask of all ROIs, and was small-volume corrected within the habenula mask.

Brain region	x	y	z	t	k	p	Correction	SVC
Whole-brain, FDR q < .05								
CS+ > CS-								
Insula	30	26	0	8.31	11	< .05	FDR	N
Insula	-30	22	4	11.8	5	< .05	FDR	N
Omission > Neutral								
Orbitofrontal Cortex	40	24	10	5.83	8	< .05	FDR	N
Insula	48	18	-4	6.16	8	< .05	FDR	N
Middle Temporal Gyrus	56	-42	-4	6.12	6	< .05	FDR	N
Inhibitor + Unexpected Reward								
Putamen	32	-16	-4	7.16	5	< .05	FDR	N
Whole-brain, FDR q < .05, ROI mask								
Juice > Neutral								
Insula	38	-4	6	6.51	3	< .05	FDR	Y
OFC	-28	36	-8	7.94	3	< .05	FDR	Y
CS+ > CS-								
OFC	-24	18	-20	5.91	8	< .05	FDR	Y
OFC	-34	20	-20	5.46	8	< .05	FDR	Y
SNc	8	-14	-12	4.8	1	< .05	FDR	Y
CS+ > Inhibitor								
SNc	12	-18	-12	6.72	1	<.05	FDR	Y
p < .001, Uncorrected								
Juice > Neutral								
Amyg (L)	-20	-2	-12	20.27	34	< .001	Unc.	Y
Amyg (R)	22	0	-14	10.38	21	< .001	Unc.	Y
p < .005, Uncorrected								
Juice > Neutral								
Caud	10	20	4	7.68	19	< .005	Unc.	Y
Caud	12	12	14	7.24	8	< .005	Unc.	Y
CS+ > CS-								
SNc	8	-14	-12	8.85	7	< .005	Unc.	Y
Inhibitor > Controls								

LHb	-2	-24	-2	6.71	11	< .005	Unc.	Y
Striatum (Inc. Pallidum, Putamen)	20	4	-6	11.29	53	< .005	Unc.	Y
FDR q < .05, SVC								
Inhibitor > Controls								
LHb*	-2	-24	0	3.76	2	< .05	FDR	Y

Figure Captions

Figure 1. Experimental design: In the conditioning block, the CS+ is paired with an orange juice reward 75% of the time and a neutral solution 25% of the time, while a control CS- is always paired with neutral solution. In the conditioned inhibition block, the CS+ is paired with an inhibitor which leads to reward omission. The rewarded CS+ continues to be shown. This is followed by subsequent conditioned inhibition blocks, where the inhibitor is shown alone in a subset of trials, along with neutral controls. The experiment ends with a conditioned inhibition test where the inhibitor is unexpectedly followed by a juice reward, and the second control stimulus is also unexpectedly followed by reward.

Figure 2 Results from the Monetary Reward Task conducted as a follow-up study. Mean ratings across subjects for CS+ CS-, and CS paired with the Inhibitor across blocks, which revealed significantly lower ratings for the CS+ paired with the Inhibitor than the CS+ alone.

Figure 3 A) Juice compared to the neutral solution showed activity in the Insula ROI [x=38,y=-4,z=---6, t=6.51, k=3], corrected at FDR q < .05 within the all-ROI mask. B) Juice compared to neutral solution showed activity in OFC, corrected at FDR q < .05 with the all-ROI mask. [x=-28 y=36, z=-8, t=7.94, k=3 voxels]. Activity also shown at p < .001 and p < .005 for visualization. C) Juice compared to neutral solution showed activity in the caudate ROI [x=10,y=20,z=4, t=7.68, k=19] at p <.005, uncorrected D) Juice compared to neutral solution showed activity in the amygdala ROI [L=-20,-2,-12, t=20.27, k = 34, R=22,0,-14, t=10.4, k=21] at p < .001 uncorrected.

Figure 4:

A) Whole-brain activity for the CS+ compared to CS- showed activity in the insula at q < .05, FDR

B) Comparing the whole brain activity for the CS+ to that for the CS-, there was activity in the OFC [x=24, y=18,z=20, t=5.91, k =8 and x=-34, y=20, z=20, t=5,46], corrected across the all-ROI mask.

C) CS+ > CS- in the SN/VTA, p <.005, uncorrected. [peak: 8, -14, -12, k=7, t= 8.85]

D) The ROI analysis showed significant activity in the SNc [t(18) = 3.17, p = .0053, Bonferroni p = .053) and VTA (t(18) = 2.58, p = .0189, Bonferroni p = .189] for the CS+ compared to the CS-.

Figure 5. A) Activity for the Conditioned Inhibitor compared to the control stimuli was significant in the lateral habenula at p <.005 [x= -2, y= -24, z=-2, t=6.71, k=11]. B) Activity in the basal ganglia ROIs (Pallidum and Putamen) for the Inhibitor > Controls at p < .005 C) A contrast comparing

activity for the CS+ to the Inhibitor showed activity in the SNc at $p < .005$. [x=8, y=-16, z=-12, k=11, t=12.35].

Figure 6. A) The SNc showed a significant increase for a CS+ compared to a CS-, but not for an Inhibitor compared to a control cue. There was a significant difference between these effects. B) The LHb showed a significant increase for an Inhibitor compared to a control cue, but not a CS compared to a CS-.

Figure 1.JPEG

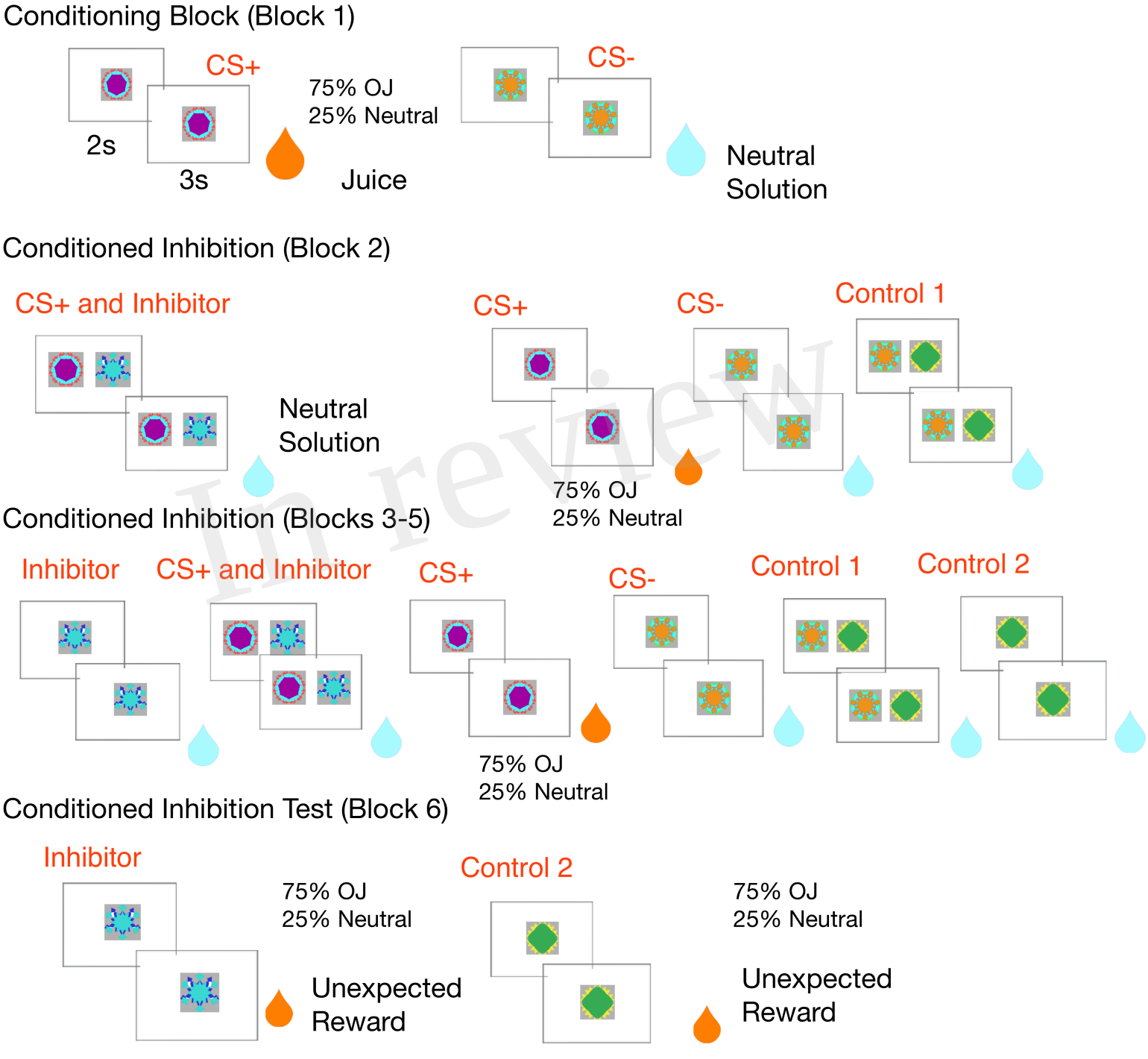


Figure 2.TIFF

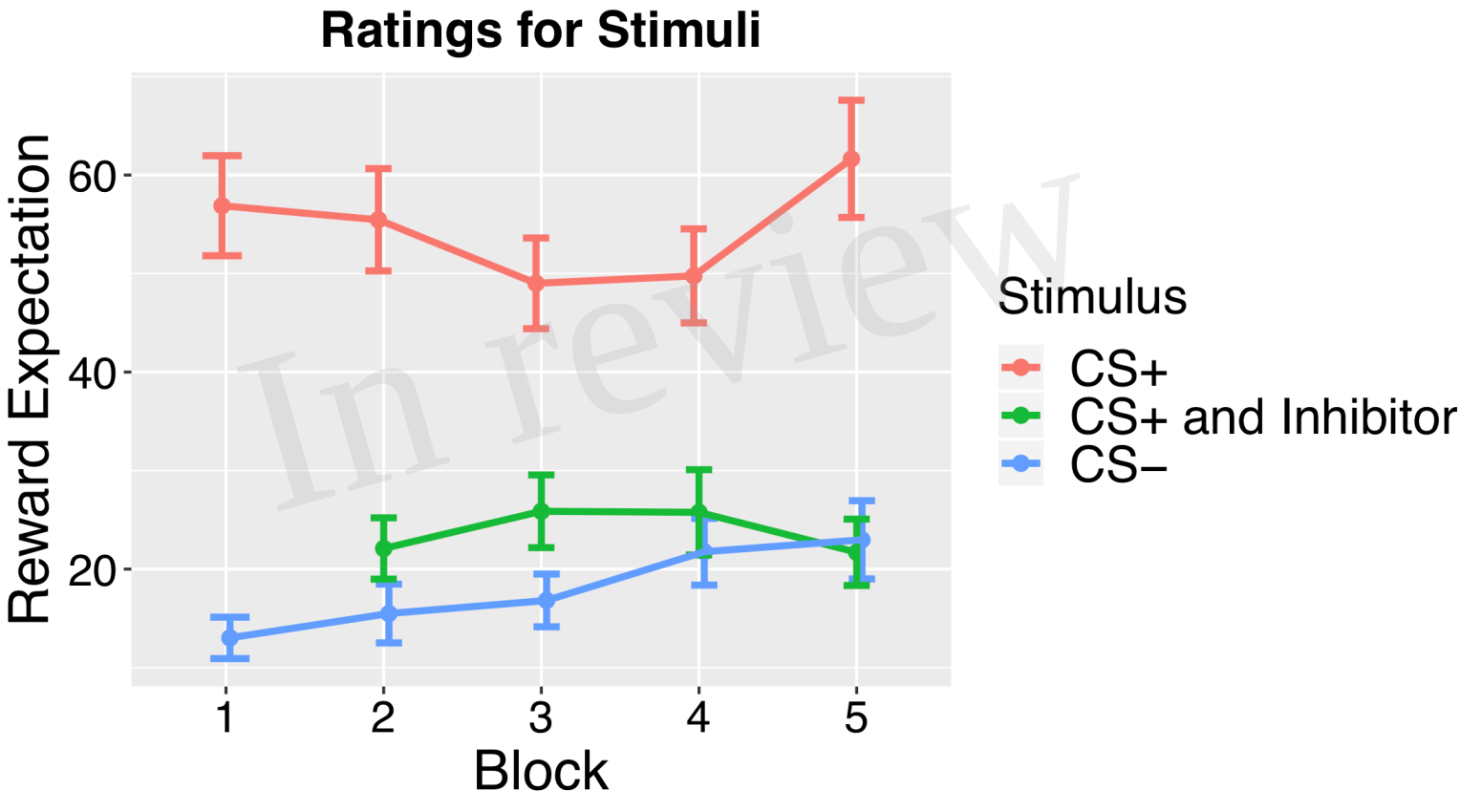
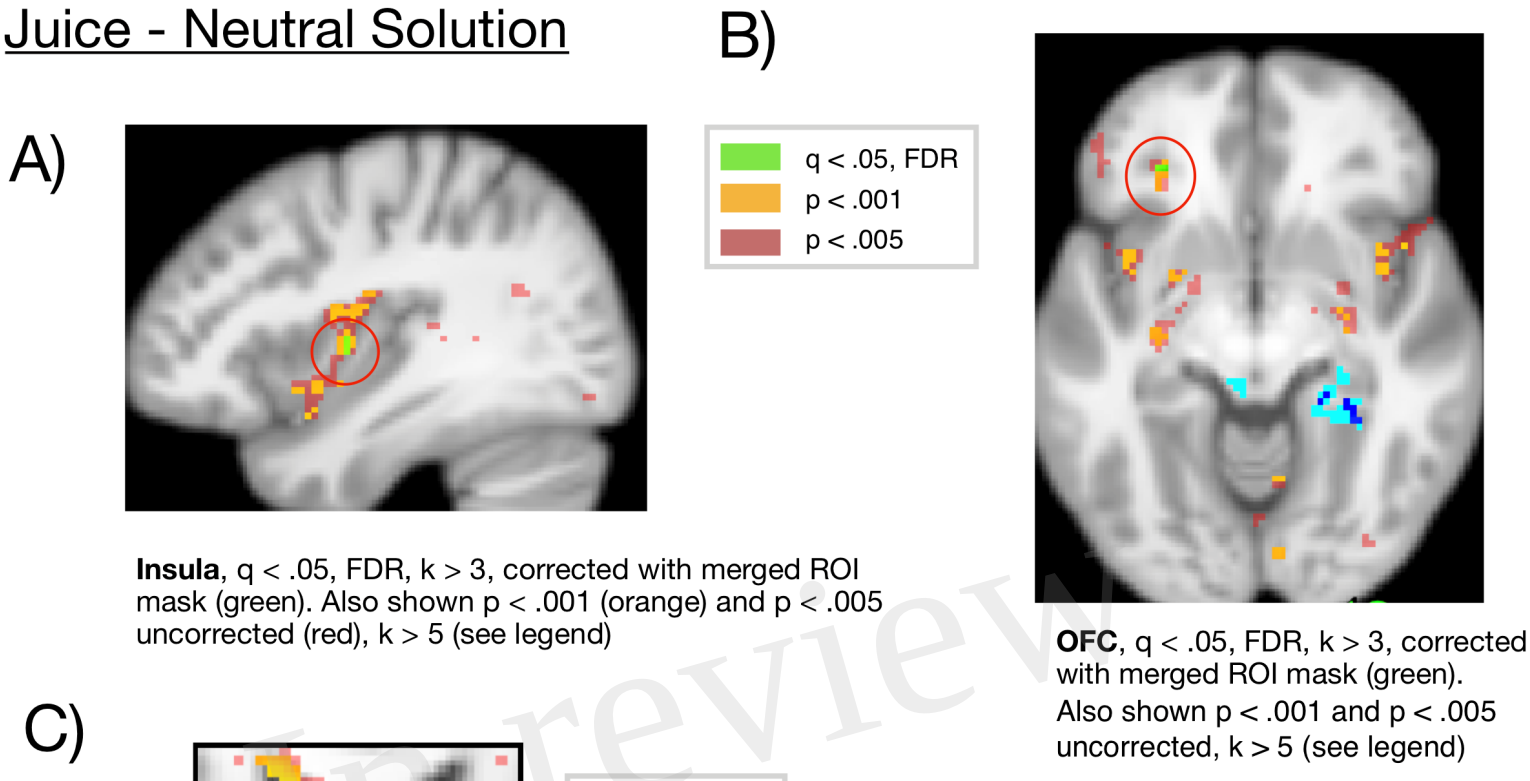


Figure 3.TIFF



Insula, q < .05, FDR, k > 3, corrected with merged ROI mask (green). Also shown p < .001 (orange) and p < .005 uncorrected (red), k > 5 (see legend)

OFC, q < .05, FDR, k > 3, corrected with merged ROI mask (green). Also shown p < .001 and p < .005 uncorrected, k > 5 (see legend)

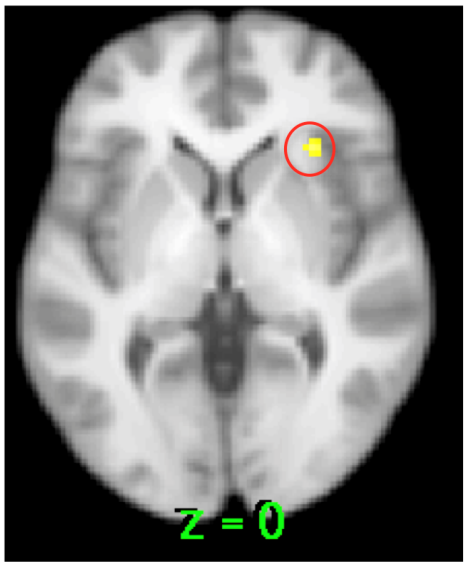
Caudate, p < .005, uncorrected, k > 5 (ROI in blue)

Habenula deactivation,
p < .001, k > 5 uncorrected (cyan)
p < .005, k > 5 uncorrected (blue)
Habenula ROI (pink)

Amygdala,
p < .001 uncorrected (orange), k > 5
p < .005, uncorrected (red), k > 5
Amygdala ROI (pink)

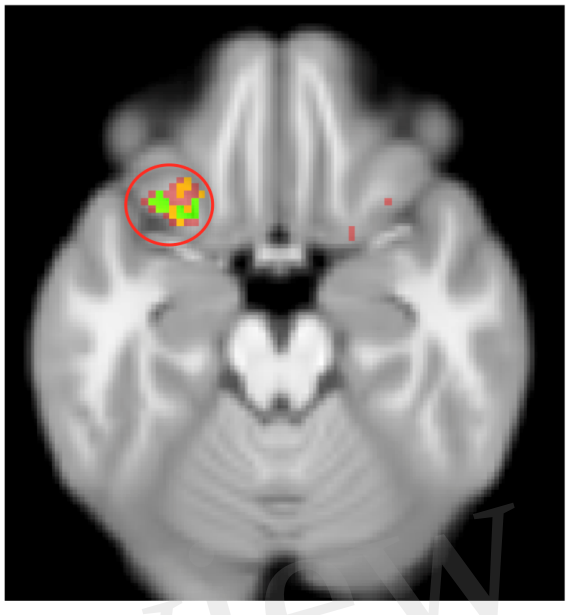
Figure 4.TIFF

A) CS+ > CS- activity



q < .05, FDR, k > 5

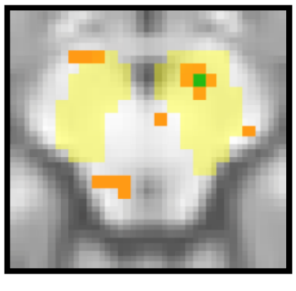
B) CS+ > CS- activity



q < .05, FDR
p < .001
p < .005

q < .05, FDR, k > 5, corrected with merged ROI mask
Also shown p < .001 and p < .005 uncorrected, k > 5 (see legend)

C) CS+ > CS- (in midbrain)



p < .005
q < .05, FDR

CS+ > CS- in SNc
p < .005, uncorrected, k > 3
(SNc ROI in yellow)

D) CS+ > CS- in midbrain ROIs

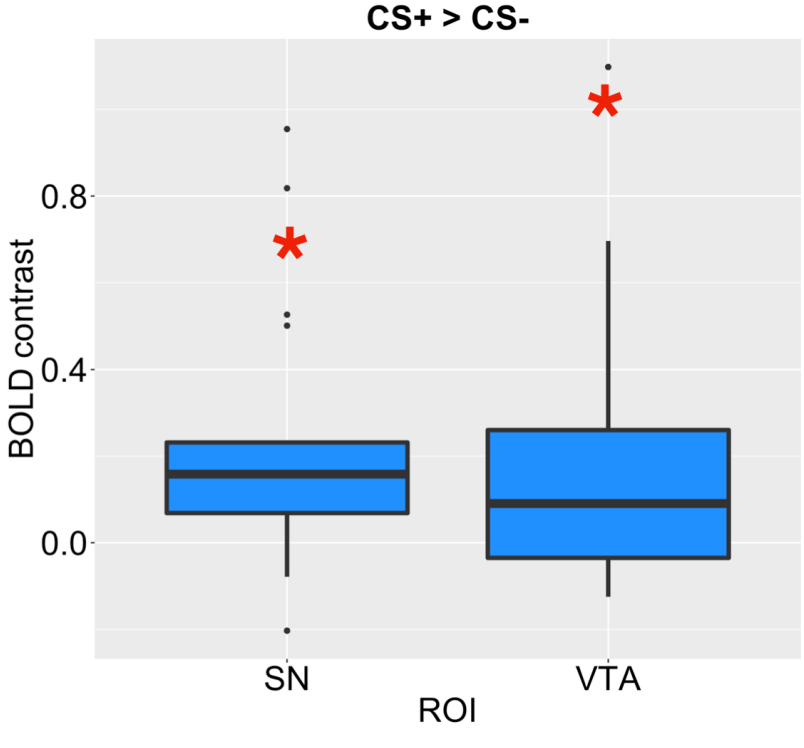


Figure 5.TIFF

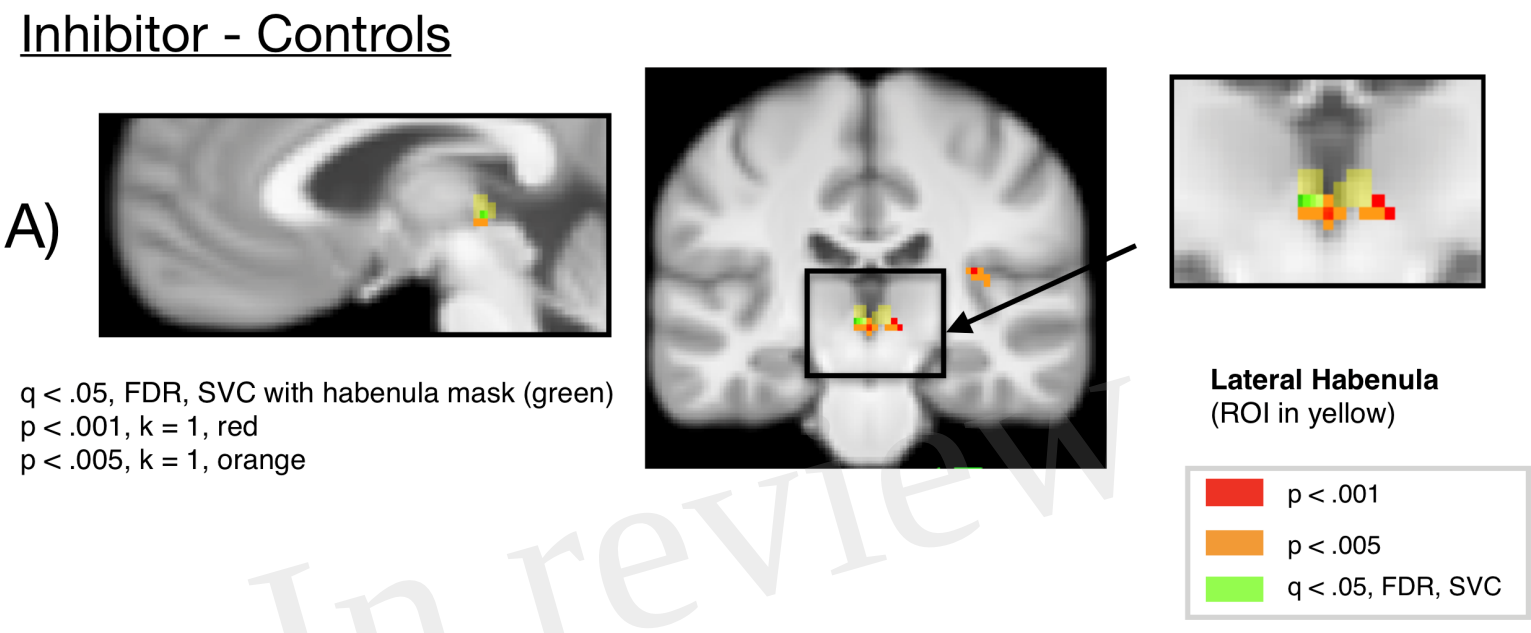


Figure 6.TIFF