

Distinct Contributions of Ventromedial and Dorsolateral Subregions of the Human Substantia Nigra to Appetitive and Aversive Learning

Wolfgang M. Pauli,¹ Tobias Larsen,¹ Sven Collette,¹  Julian M. Tysza,¹ Ben Seymour,^{2,3} and John P. O'Doherty¹

¹Division of Humanities and Social Sciences and Computation and Neural Systems Program, California Institute of Technology, Pasadena, California 91125,

²Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom, and ³National Institute for Information and Communications Technology, Osaka, 565-0871, Japan

The role of neurons in the substantia nigra (SN) and ventral tegmental area (VTA) of the midbrain in contributing to the elicitation of reward prediction errors during appetitive learning has been well established. Less is known about the differential contribution of these midbrain regions to appetitive versus aversive learning, especially in humans. Here we scanned human participants with high-resolution fMRI focused on the SN and VTA while they participated in a sequential Pavlovian conditioning paradigm involving an appetitive outcome (a pleasant juice), as well as an aversive outcome (an unpleasant bitter and salty flavor). We found a degree of regional specialization within the SN: Whereas a region of ventromedial SN correlated with a temporal difference reward prediction error during appetitive Pavlovian learning, a dorsolateral area correlated instead with an aversive expected value signal in response to the most distal cue, and to a reward prediction error in response to the most proximal cue to the aversive outcome. Furthermore, participants' affective reactions to both the appetitive and aversive conditioned stimuli more than 1 year after the fMRI experiment was conducted correlated with activation in the ventromedial and dorsolateral SN obtained during the experiment, respectively. These findings suggest that, whereas the human ventromedial SN contributes to long-term learning about rewards, the dorsolateral SN may be particularly important for long-term learning in aversive contexts.

Key words: appetitive learning; aversive learning; brainstem; dopaminergic midbrain; fMRI

Significance Statement

The role of the substantia nigra (SN) and ventral tegmental area (VTA) in appetitive learning is well established, but less is known about their contribution to aversive compared with appetitive learning, especially in humans. We used high-resolution fMRI to measure activity in the SN and VTA while participants underwent higher-order Pavlovian learning. We found a regional specialization within the SN: a ventromedial area was selectively engaged during appetitive learning, and a dorsolateral area during aversive learning. Activity in these areas predicted affective reactions to appetitive and aversive conditioned stimuli over 1 year later. These findings suggest that, whereas the human ventromedial SN contributes to long-term learning about rewards, the dorsolateral SN may be particularly important for long-term learning in aversive contexts.

Introduction

It is well established that the substantia nigra (SN) and ventral tegmental area (VTA) of the midbrain play an important role in

facilitating learning and updating of stimulus–reward associations, especially via the phasic activity of dopamine neurons that have been found to encode a reward prediction error (RPE) signal (Mirenowicz and Schultz, 1996; Montague et al., 1996; Schultz et al., 1997; Schultz, 1998; Day et al., 2007; Tsai et al., 2009; Flagel et al., 2011; Takahashi et al., 2011; Lammel et al., 2012).

Received June 13, 2015; revised Aug. 17, 2015; accepted Sept. 8, 2015.

Author contributions: W.M.P., T.L., B.S., and J.P.O. designed research; W.M.P. and T.L. performed research; W.M.P. and J.M.T. contributed unpublished reagents/analytic tools; W.M.P., T.L., and S.C. analyzed data; W.M.P. and J.P.O. wrote the paper.

This work was supported by the National Institute of Mental Health Caltech Conte Center for the Neurobiology of Social Decision Making. We thank all members of the O'Doherty Human Reward and Decision Making laboratory for help with preprocessing, statistical analysis of pupillometry and fMRI data, and helpful discussions.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Wolfgang M. Pauli, MC 228-77, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125. E-mail: pauli@caltech.edu.

T. Larsen's current address: Interdepartmental Centre for Mind/Brain Sciences, Università degli studi di Trento, 38060 Trento, Italy.

DOI:10.1523/JNEUROSCI.2277-15.2015

Copyright © 2015 the authors 0270-6474/15/3514220-14\$15.00/0

One important question is about the role that these midbrain nuclei might play in aversive learning, over and above their role in appetitive learning. Early reports in nonhuman primates, which focused on dopamine neurons in those areas, found that the majority of these neurons exhibited response properties consistent with a predominantly selective involvement in appetitive learning (Mirenowicz and Schultz, 1996). More recently, it has been reported that, although dopamine neurons in the ventromedial SN are involved selectively in appetitive learning, as previously reported, a subpopulation of neurons in the dorsolateral SN are involved in learning about aversive events, as well as about appetitive events. These neurons are described as increasing in activity in response to surprising occurrences of both appetitive and aversive outcomes (Matsumoto and Hikosaka, 2009). Others have suggested that the apparent responsiveness of some dopamine neurons to aversive stimuli does not reflect prediction error coding per se but rather the effects of generalization (Schultz, 2010), initial sensory-driven responses, and/or overshadowing arising from the presence of overlapping neutral cues for appetitive and aversive learning (Fiorillo et al., 2013; Fiorillo, 2013). Furthermore, it has also been proposed that, although aversive learning signals are indeed present in the SN, such signals are carried by GABAergic neurons rather than dopamine neurons (Cohen et al., 2012). Regardless of the specific neurotransmitter system involved, these findings suggest the possibility that multiple types of learning-related signals are present in these midbrain nuclei.

Much less is known about the response properties of these midbrain nuclei in the human brain. Neuroimaging methods have widely reported prediction error activity in the brain, although typically in target areas of dopamine neurons, such as the striatum as opposed to directly in the midbrain structures (O'Doherty et al., 2003; Delgado, 2007). Measurement of activity in midbrain structures with functional neuroimaging has been limited by an increased susceptibility of the brainstem as a whole to the effects of physiological noise, and a lack of spatial resolution. Nevertheless, some studies have reported activity in SN and/or VTA in response to prediction errors during reward presentation or omission (O'Doherty et al., 2002; Wittmann et al., 2005; D'Ardenne et al., 2008). A recent study has also reported activity in the SN in response to anticipation of aversive outcomes, although it is not known from this experiment whether this region is encoding a prediction error for aversive outcomes per se, or a pure anticipatory value signal (Hennigan et al., 2015).

Here, we used a high-resolution imaging approach optimized for the brainstem while participants underwent higher-order appetitive and aversive Pavlovian learning, where a distal conditioned stimulus was followed by a proximal conditioned stimulus. Our aims were threefold: (1) to determine whether human SN and VTA regions are exclusively involved in encoding a RPE, or whether these structures are also involved in encoding prediction errors during learning about aversive events; (2) to establish whether there exists regional functional specialization within the SN and VTA in appetitive and aversive learning (Matsumoto and Hikosaka, 2009); and (3) to test whether we would find evidence for an unsigned prediction error (UPE) signal that shows increases in activity in response to unexpected presentations of neutral, appetitive, or aversive stimuli (D'Ardenne et al., 2013), or whether instead activity would always decrease in response to aversive stimuli, similar to the way in which an unexpected reward omission results in a decrease in activity to below baseline as found in single neurons encoding RPE signals (Mirenowicz and Schultz, 1996; Roesch et al., 2007).

Materials and Methods

Participants. Eight participants (4 females) with a mean age of 26 years (SD: 2.27 years) participated in a behavioral pilot study. Written informed consent was obtained from all subjects, according to a protocol approved by the Human Subjects Protection committee of the California Institute of Technology (Pasadena, CA).

Twenty-nine right-handed participants (10 females) with a mean age of 33.65 years (SD: 4.8 years) participated in the fMRI study. All subjects were free of neurological or psychiatric disorders and had normal or corrected-to-normal vision. Written informed consent was obtained from all subjects, according to a protocol approved by the Human Subjects Protection committee of the California Institute of Technology (Pasadena, CA). Six subjects had to be excluded from analysis due to a technical problem with equipment in the scanner room that resulted in the presence of systematic noise in the fMRI data.

Task description. Human volunteers participated in a higher-order Pavlovian conditioning paradigm, in which they learned to associate two sequentially presented conditioned stimuli (fractal images), and a pleasant (apple or white grape juice, Trader Joe's), affectively neutral (artificial saliva made of 25 mM KCl and 2.5 mM NaHCO₃), or unpleasant (salty tea made of 2 black tea bags and 29 g of salt per liter) flavor liquid. Before the beginning of the experiment, participants were given samples of white grape juice and apple juice and got to choose which one they would like to receive as a pleasant liquid.

Details of the trial structure are shown in Figure 1A. To mitigate against effects of swallowing-related movement on measurements of midbrain activity, we used a two-stage sequential conditioning paradigm in which one visual cue probabilistically predicted another, which in turn deterministically predicted the delivery of either an aversive, appetitive, or neutral liquid outcome (Fig. 1A). Critically, in this paradigm, the deterministic proximal cue (CSp) was sometimes not delivered as predicted by the distal cue (CSd), thereby inducing both positive and negative prediction errors in both appetitive and aversive learning contexts (Fig. 1B), and capturing the core feature of the temporal difference algorithm: learning via prediction errors induced by sequential predictors (Seymour et al., 2004). Specifically, in 30% of trials, the expectation evoked by the distal cue would be reversed by the proximal cue (Fig. 1B). This higher-order trial structure enabled us to study the neural representations of prediction errors at the onsets of CSd and CSp, rather than at the time of delivery of liquids, and was designed so to avoid confounding motion artifacts elicited by the delivery of liquids.

The distal cue was presented randomly at 1 of 8 possible locations around the fixation cross, and the proximal cue was presented randomly at 1 of the remaining 7 locations. We chose to vary cue locations for several reasons: to ensure that learning occurred with respect to the cue identity as opposed to being based on (1) spatial location or (2) a specific saccade direction, which could have involved a type of instrumental cue-saccade learning. By ensuring that the constant variable across trials is the cue identity as opposed to position, this helps to ensure that the main learning component we are measuring is Pavlovian. Furthermore, (3) varying spatial position maximizes the salience of the cues, by requiring participants to move their attention. Importantly, because spatial position of cues is randomized across trials, there is no systematic relationship between spatial position and reward and aversive prediction errors. Thus, although it is possible that variance in spatial position might induce additional variance into the prediction error signal, this is very unlikely to have confounded our main conclusions regarding a difference in prediction error signals in the reward and aversive conditions. It is also important to note that randomization of spatial position is a very pervasive experimental manipulation in human learning studies, for the reasons described above (O'Doherty et al., 2004; Seymour et al., 2005; Tobler et al., 2006).

The experiment consisted of two sessions, lasting ~16 min each. Each session was composed of 70 trials, yielding a total of 140 trials. In one of the sessions, only pleasant and neutral flavors were presented, whereas in the other, aversive session, conditioned stimuli predicted the subsequent delivery of either the unpleasant flavor stimulus, or the affectively neutral stimulus. The order of presentation of the two sessions was fully coun-

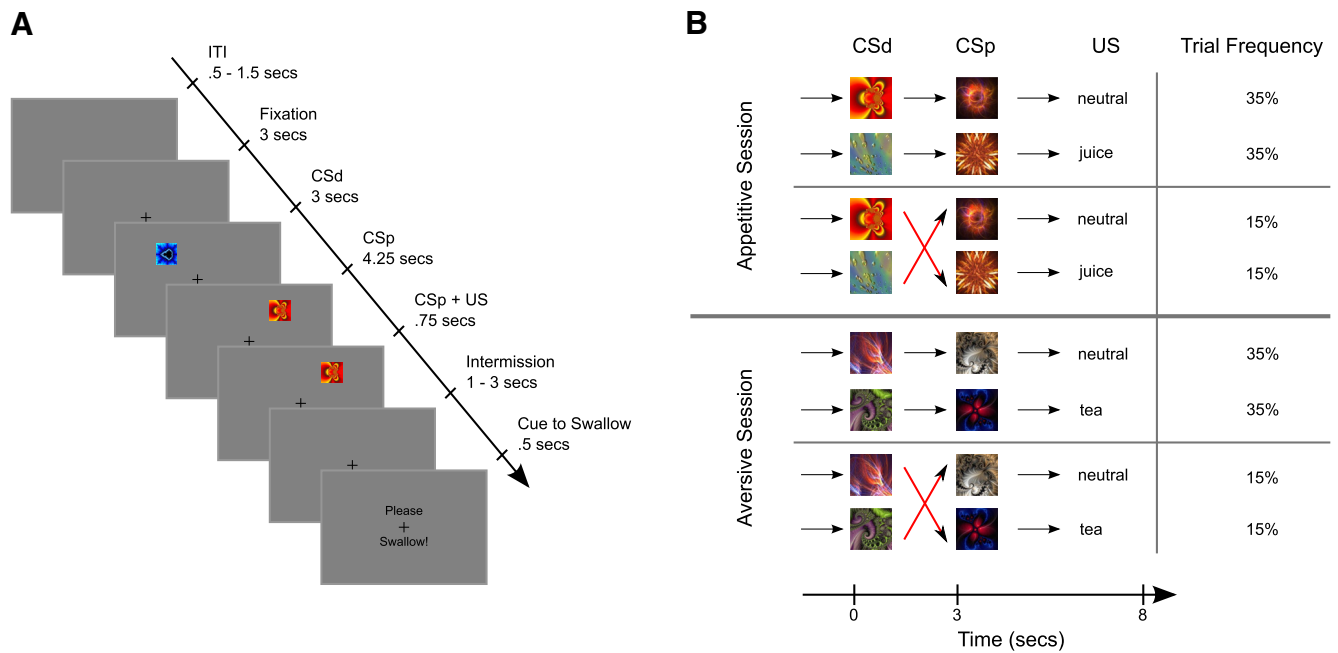


Figure 1. **A**, Illustration of the sequence of events within a trial. At the beginning of each trial, a cue stimulus (CSd) was presented randomly in 1 of 8 locations around the fixation cross, followed by a second cue (CSp) presented randomly at 1 of the remaining 7 locations 3 s later. After another 4.25 s, liquid was presented over the course of 0.75 s. The presentation of the proximal cue and liquid delivery coterminated. After another 1–3 s (uniformly distributed), a visual cue appeared on the screen for 0.5 s, instructing participants to swallow the liquid. A further 1–5.5 s later, the next trial was triggered. Subjects were told that the visual stimuli would allow them to predict which liquid would be delivered at the end of each trial. **B**, The experimental design expressed as a Markov chain, giving four separate trial types. In 70% of trials, the CSd validly predicted the following CSp and liquid. However, in 30% of trials, the CSp would reverse the expectation evoked by the CSd.

terbalanced across participants. The rationale for including the appetitive and aversive conditioning procedures in separate sessions as opposed to including both conditions intermixed within the same sessions was to avoid contrast effects observed in a prior study (Seymour et al., 2005), which reported that cues signaling the aversive liquid tended to overwhelm cues signaling the pleasant liquid such that both the pleasant and the neutral cue stimuli were viewed as similarly experienced as a pleasant relief when contrasted against the aversive outcome. Having separate sessions for appetitive and aversive conditioning ensured robust behavioral conditioning in both the appetitive and aversive cases and largely avoided contrast effects between the appetitive and aversive conditions. It should be noted that this choice might have had the adverse effect that neutral outcomes were considered aversive in the appetitive condition, and appetitive in the aversive condition (Kim et al., 2006).

Participant instructions. Before the conditioning session, subjects received the following verbal task instructions: “In each trial, an image will appear on the screen, followed by a second image, which will be followed by the delivery of a liquid. Each image will help you predict what kind of liquid will be delivered. There will be two sessions, in the first session you will either receive juice [salty tea] or a neutral solution, in the second you will either receive salty tea [juice] or a neutral solution.”

Apparatus. The pleasant, neutral, and unpleasant tasting liquids were delivered by means of three separate electronic syringe pumps. These pumps pushed 0.75 ml of liquid to the subject’s mouth via clear PVC plastic tubes (<http://www.freelin-wade.com>; outside diameter, 8 mm; inside diameter, 4.8 mm), the other ends of which were held between the subject’s lips like a straw, while they lay supine in the scanner, or placed their chins on a chin rest in the pilot study.

Stimulus ratings. One year after the last volunteer had participated in the MRI study, participants of the MRI study were contacted again and invited to participate in a follow-up survey. In this follow-up survey, participants were asked to rate the different conditioned stimuli they had experienced during the experiment. Rating scales ranged from –5 (very unpleasant) to 5 (very pleasant). All 23 participants included in the fMRI data analysis were invited to participate in the follow-up survey, of which 17 responded, corresponding to a response rate of 70%. The responses of 2 participants were excluded because they had responded indiscriminately only with extreme ratings for all stimuli.

To evaluate how well participants remembered stimulus–outcome association in the follow-up survey, we calculated the difference in rating of the most aversive or appetitive conditioned stimulus and the neutral stimulus of the aversive and appetitive session, respectively.

Pupil dilation and eyeblink. To obtain behavioral measures of conditioning, we measured participants’ pupillary and eyeblink responses, both of which have been found to be reliable indexes of Pavlovian conditioning in prior studies. Indeed, recently we have found that, whereas pupil dilation is a good indication of learning with appetitive juice rewards, eyeblink responses are a more robust indicator for learning about aversive events (Prévost et al., 2013). To obtain these measurements, an infrared camera continuously recorded a video of the participants’ pupils at 60 frames per second. Pupil diameter was extracted using the open-source eye tracking software MrGaze. Before statistical analysis, pupil data were down-sampled to 20 frames per second and baseline corrected to the pupil size at the onset of each conditioned stimulus presentation. Statistical analysis was conducted using the cumulative percentage change in pupil diameter between 0.5 and 2 s after stimulus onset as this is the time window after stimulus presentation we have found previously to be responsive during conditioning (Prévost et al., 2013). For the statistical analysis of eyeblink rate, we counted the number of eyeblinks detected by the eye tracking software during the duration of each conditioned stimulus, then mean-centered each participant’s eyeblink rates across the two sessions, to account for individual differences.

For a model-free analysis of the pupil and blink data, we used linear mixed-effects models with random factor participant and fixed factor stimulus valence (i.e., CS⁺ vs CS[−]). The effective degrees of freedom were calculated using the Welch–Satterthwaite approximation (Satterthwaite, 1946), to achieve a more conservative estimate of *p* values.

Fluctuations in respiration and heart rate. In the fMRI experiment, peripheral pulse and respiration were recorded using a pulse oximeter positioned on the left index finger of subjects’ left hand and a pressure sensor placed on the umbilical region. The time courses derived from these measures were used to derive a regressor of no interest in the fMRI data analysis using the RETRO-ICOR algorithm (Glover et al., 2000).

Additional motion regressors. In addition to the rigid body motion regressors during the realignment step of data processing of fMRI data, a camera continuously recorded the position of the tip of participant’s

nose. The time course derived from this measure was used as a regressor of no interest in the fMRI data analysis.

Statistical analysis of behavioral data. Behavioral data (i.e., stimulus ratings, pupil and eyeblink data) were analyzed using a linear mixed-effects model approach (Pinheiro and Bates, 2000), using the R statistics package lme4 (Bates et al., 2008). Linear mixed-effects models were chosen because they allow specification of random effects (Fisher, 1919) in addition to fixed (experimental) effects to account for repeated measurements made on the same participants.

Computational model analysis. The temporal difference (TD) learning algorithm (Sutton and Barto, 1998), with a temporal discounting parameter, and identical learning rates for the CS_p and CS_d time points was used to predict pupil dilation (juice session) and blink rate (salty tea session).

The value of a distal cue was updated according to the following:

$$V(\text{CS}_d) \leftarrow V(\text{CS}_d) + \alpha[\lambda V(\text{CS}_p) - V(\text{CS}_d)]$$

The value of the proximal cue was updated according to the following:

$$V(\text{CS}_p) \leftarrow V(\text{CS}_p) + \alpha[r + V(\text{CS}_p)]$$

In these equations, α represents the learning rate, and λ the temporal discounting factor. The deliveries of pleasant and aversive liquids were coded as $r = 1$ and the neutral liquid was coded as $r = 0$. Cue values were initialized with 0 at the beginning of each session. Value and prediction error estimates of the TD algorithm were used as regressors in a linear mixed-effects model, with participants as the random effect factor, and the TD value or prediction error estimates, as well as their interaction, at the onsets of CS_d and CS_p as fixed effects. To determine the best-fitting learning rates, we performed a complete 2D grid search (50 equidistant steps from 0.001 to 0.999) for each combination of learning rate and temporal discounting, and recorded the log-likelihood of the population data, given the model and the learning rates.

We conducted a permutation test to evaluate the fit of pupil and eyeblink responses to conditioned stimulus onsets by the temporal difference model (for details on the temporal difference model, see Computational model analysis). Specifically, we permuted each participant's sequence of pupil/blink responses to eliminate any contingency between conditioned stimulus type and these responses. A full grid search over the free parameters of the temporal difference model was conducted for each permutation. To create a robust baseline distribution, we created 1000 permutations of each participant's data. The effective degrees of freedom of the linear mixed-effects models were calculated using the Welch–Satterthwaite approximation (Satterthwaite, 1946), to achieve a more conservative estimate of p values.

fMRI data acquisition. Functional imaging was performed on a 3 tesla MRI system (Magnetom Tim Trio, Siemens Medical Solutions) located at the Caltech Brain Imaging Center (Pasadena, CA) with a 32-channel head receive array for all the MR scanning sessions. To reduce involuntary head motion, participants' heads were securely positioned with foam pads.

Because the focus of our study was the midbrain, we only acquired T2*-weighted echo planar images with coverage limited to the midbrain while subjects were performing the task. This coverage included the ventral posterior part of the prefrontal cortex, the striatum and globus pallidus, the insula, the amygdala, and the upper part of the cerebellum (among other regions). A total of 40 slices were acquired with an interleaved-ascending order for each T2*-weighted EPI volume, with an isotropic resolution of 1.5 mm. Phase encoding oversampling with controlled foldover was used to eliminate signal from anterior and posterior brain regions, achieving an in-plane volume localization ("zoomed EPI"). Other imaging parameters included the following: TR, 2770 ms; TE, 30 ms; flip angle, 81 degrees; field of view, 160 mm × 160 mm × 100 mm; matrix, 64 × 64; a whole-brain high-resolution T1-weighted structural scan (voxel size, 0.77 × 0.77 × 1.0 mm); and a partial coverage high-resolution T2-weighted structural scan (T2-weighted 3D SPACE, isotropic voxel size, 0.75 mm). Dual-echo gradient echo field maps were acquired to allow geometric correction of the EPI data in the midbrain. Spatial distortion is not that pronounced in the mid-brain but is worth

doing when the data are high resolution. We discarded the first 3 EPI volumes before data processing and statistical analysis to allow for magnetization equilibration.

fMRI data analysis. The SPM8 software package was used to analyze the fMRI data (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London), the FSL feat pipeline was adapted to preprocess the high-resolution fMRI data. Slice-time correction was applied to the functional images to adjust for the fact that different slices within each image were acquired at slightly different points in time. Images were corrected for participant motion, and gradient fieldmaps were applied to correct distortions of the zoomed EPI images. For the purpose of mid-brain analysis, images were smoothed using a 2 mm FWHM 3D Gaussian kernel to account for residual motion of participants.

For the first-level analysis, the event-related fMRI data were analyzed by constructing time-series of prediction errors at the onset of CS_d and CS_p stimuli. These time-series were estimated by fitting pupil and eyeblink responses (see Computational model analysis), and used as parametric regressors in a GLM. Separate GLMs were run for signed and unsigned prediction errors. The following regressors of no interest were included in the model in the following order, orthogonalized to the first regressor of no interest. We included 10 regressors to account for physiological fluctuations (2 related to respiration, and 4 high pass filtered and 4 low pass filtered photoplethysmography measures of heart beat), which were estimated using the RETRO-ICOR algorithm (Glover et al., 2000). Six scan-to-scan motion parameters derived from the affine realignment procedure were also included as regressors of no interest. To account for motion of subjects at a high temporal resolution, we added an additional regressor of no interest (see Additional motion regressors). Finally, we performed an independent components analysis (FSL melody) and included the time course of those 5 components (of 90) as additional regressors of no interest, which loaded the highest on the interpeduncular cistern, located just anteriorly to the SN and VTA.

Before the second-level analysis, participants' zoomed EPIs were coregistered to a multimodal T1/T2 optimal template, created iteratively using the Advanced Normalization Toolbox (Avants et al., 2010). The costs and benefits of different image registration pipelines were discussed previously (Limbrick-Oldfield et al., 2012). Results of second-level analyses were affinely transformed into MNI space.

Evaluation of parameter estimates. To gain a detailed insight into how BOLD responses scaled with prediction errors, we exported the subject β values of the first-level GLM and performed a stepwise regression. To avoid double dipping (Kriegeskorte et al., 2009), we defined participant ROIs using a leave-one-participant-out cross-validation procedure. Although this procedure did not limit ROIs to be localized to the same hemisphere as was found in the overall analysis, we found that ROIs were largely overlapping and all localized to the same hemisphere (data not shown). Similar asymmetries have been reported in previous studies (D'Ardenne et al., 2013; Hennigan et al., 2015). For each participant, we created 5 bins of consecutive ranges of TD prediction errors, so that each bin contained the same number of trials (14 trials). For example, the first bin may include all CS onsets in which prediction errors were between -1 and -0.8 , the second bin all trials with prediction errors between -0.8 and -0.3 , and so on. The borders between bins varied between participants because each participant experienced a different order of trial types, and the size of prediction errors is a function of the specific sequence of events. The stimulus onsets of the trials within a bin were used to form an onset regressor. This resulted in a GLM with 5 onset regressors, corresponding to the 5 bins of prediction errors. For each participant, we calculated the z-scored median β value within the ROIs and then calculated the truncated population mean and SEs (excluding extreme values from the top 5% and bottom 5% of the distributions). This analysis was performed separately for the two conditioned stimuli (CS_d and CS_p), and the ventromedial and dorsolateral SN ROIs in the appetitive and aversive session, respectively.

Functional gradient within the SN. To investigate whether it was possible to identify a functional gradient within the SN, we tested for a systematic change in fitted parameter estimates along a ventromedial to dorsolateral axis within the SN. Because the above evaluation of parameter estimates suggested that the UPE regressor was not adequately rep-

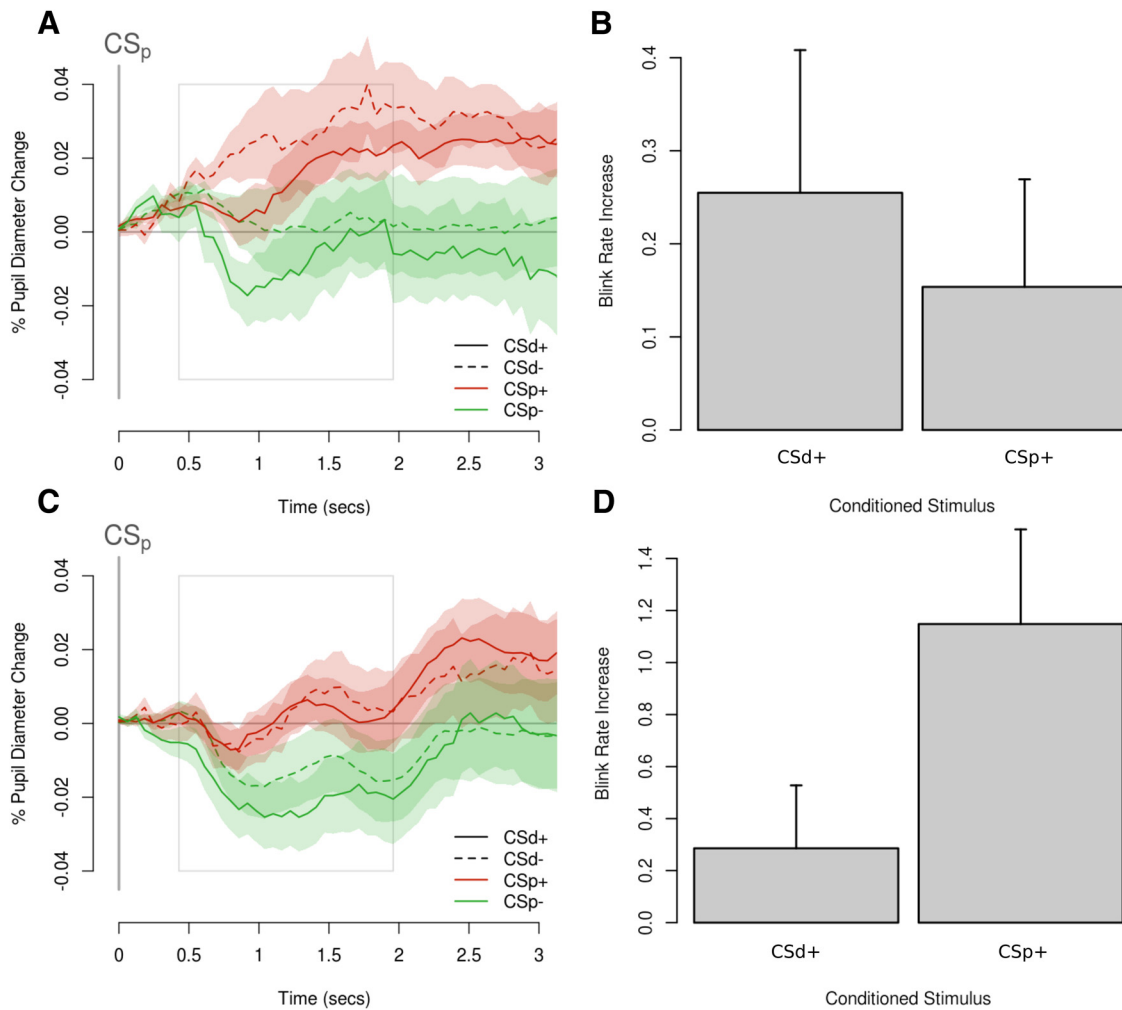


Figure 2. Behavioral effects of conditioning. **A**, Plot of averaged pupil responses from the appetitive session of the behavioral pilot study aligned to the time of onset of the proximal cue (CSp). Solid lines indicate when the distal cue (CSd) predicts the delivery of the juice reward. Dashed lines indicate when the CSd predicts the delivery of the neutral outcome. Red represents when the CSd predicts the delivery of the juice reward. Green represents when the CSp predicts the delivery of the neutral outcome. Both the value of the expected reward and violations in those expectations when the CSp predicts a different outcome to the CSd (i.e., prediction error) influence changes in pupil size in the appetitive session of the pilot study. **B**, The presentation of both CSd⁺ and CSp⁺ resulted in an increase in blinking relative to the blink rate after a CSd⁻ and CSp⁻ in the aversive learning session of the pilot study. **C**, **D**, Analogous results for the fMRI study. Shaded areas and error bars indicate standard errors.

resenting activation dynamics during the aversive session, in this analysis we focused on the fitted parameter estimates for the RPE regressor in the appetitive session and the combined regressor for aversive value signals at the distal cue and RPEs at the proximal cue during the aversive session. To determine the location of each SN voxel along the ventromedial to dorsolateral axis, we calculated its Manhattan distance from an origin placed at the most ventromedial and central, along the rostrocaudal axis, voxel of the SN. The resulting distance of each voxel was rounded to whole millimeters; and for each participant and regressor, the parameter estimates were averaged across equidistant voxels. To test for statistical significance, we performed a linear mixed-effects model analysis, with the random effect participant, and the fixed effects distance and parametric fMRI regressor.

Results

Behavioral pilot experiment

Appetitive conditioning leads to conditioned pupil dilation at CS onsets

We recorded pupil diameter changes in response to conditioned stimulus onsets as an autonomic measure of appetitive conditioning (Bitsios et al., 2004; Seymour et al., 2007; Bray et al., 2008; Prévost et al., 2013). We expected that the presentation of a CS

would result in a pupil dilation response correlated with the value of the CS (O'Doherty et al., 2006; Bray et al., 2008), and to the degree to which the CS presentation changes the participants' reward expectations (Nassar et al., 2010; Preuschoff et al., 2011).

Although we did not find any direct evidence of changes in the pupil responses to the distal CS, we found clear evidence for higher-order learning at the time of presentation of the proximal CS, as the pupil response at that time reflects a mix of the expected reward and violations of expectations (i.e., prediction errors) (Fig. 2A). Trials in which the distal cue predicted the neutral outcome, but the proximal cue now predicted the juice outcome elicited the stronger pupil dilation relative to trials in which both the distal cue and the proximal cue predicted the reward outcome. Furthermore, trials in which the reward outcome was expected at the time of the distal cue, but the proximal cue now signaled a neutral outcome resulted in a stronger pupil constriction than trials in which both the distal and proximal cue predicted the neutral outcome.

To formally test this result, we computed a linear mixed-effects model in which we regressed value and prediction error

signals from the computational model as well as the interaction between these two variables against pupil responses to the onset of either cue (with random effects factor participant, see Materials and Methods). This model fit pupil responses best with a learning rate of 0.26 and a temporal discounting rate of 0.79. We found a significant interaction effect of value and prediction error on the pupil responses ($t = 3.40$, $p < 0.001$, $df = 1294.1$), indicating that the pupil response reflects a mix of these two computational variables.

An analogous analysis of the pupil dilation in response to CS onsets was run for the aversive conditioning session but did not yield any statistically significant effects, possibly because of the high blink rate during aversive conditioning (see below).

Aversive conditioning elicits eyeblink reflex responses at CS onsets

A prominent conditioned response found during conditioning with an aversive flavor outcome reported previously by our group is the presence of blinking responses elicited by a conditioned stimulus associated with an aversive flavor (Prévost et al., 2013). To test for this in the present data, we recorded eyeblinks in response to the onset of conditioned stimuli as a measure of conditioning in the aversive session (Fig. 2B). We predicted that the presentation of a CS⁺ would result in a larger eyeblink response, compared with a neutral CS⁻, averaged across proximal and distal cues. Our statistical analyses suggest that this was indeed the case ($t = 2.06$, $p = 0.04$, $df = 1110.7$).

Using a statistical approach analogous to the one used for the pupil response above, we found that the best fit of blink rate changes was achieved with a learning rate of 0.53 and essentially no temporal discounting ($\lambda = 0.99$). An investigation of the associations among different TD variables and eyeblink responses for the best fit yielded a main effect of both value ($t = 2.71$, $p = 0.007$, $df = 1306.2$) and prediction error ($t = -2.26$, $p = 0.024$, $df = 1306.2$).

An analogous analysis of the blink response at CS onset was run for the appetitive conditioning session but did not yield any statistically significant results. This result therefore confirms our previous finding that the blink response appears to be specific to aversive conditioning (Prévost et al., 2013).

fMRI experiment

Appetitive conditioning leads to conditioned pupil dilation at CS onsets

As in the pilot study, we predicted that pupil responses to the presentation of a CS would correlate with TD value and RPEs. To test this prediction, we used the identical analysis approach as in the behavioral pilot study to investigate the correlation of TD value and prediction error with pupil and eyeblink responses during the fMRI experiment.

We found that the best fit of pupil diameter changes with the TD model was achieved with a learning rate of 0.22 and a temporal discount rate of 0.89. The log-likelihood of this best fit (-4530.41) was significantly better than the log-likelihoods of fits of the permuted data (mean \pm SD, -4552.2 ± 1.43 , z -score = 15.3 , $p = 2.1e-53$).

To avoid overfitting in the next step, we did not use the best fitting parameters for learning rate and temporal discount rate estimated from the behavioral data in these same fMRI subjects but instead used the parameters estimated from the behavioral pilot study. This way, our statistical results would be unbiased. We found a significant interaction effect of value and prediction error on pupil responses to CS onsets at the time of the proximal cue ($t = -2.99$, $p = 0.003$, $df = 3200.4$), but also a main effect of

prediction error ($t = 6.86$, $p < 0.001$, $df = 3200.3$). These findings provide further evidence of successful higher-order conditioning (Fig. 2C) but also confirm our observations of an interaction between stimulus value and prediction errors on pupil responses to the onset of the proximal cue, as found in the behavioral pilot study.

An analogous out-of-sample test of the pupil dilation at CS onset was run for the aversive conditioning session. Our statistical analyses suggest a positive correlation between pupil diameter and value ($t = -2.99$, $p = 0.003$, $df = 3191.0$) and prediction error ($t = 4.38$, $p < 0.001$, $df = 3189.0$). However, as in the behavioral pilot study, the pupil dilation results from the aversive conditioning session should be treated with caution, as changes in pupil dilation could be influenced directly by an increased eyeblink rate during aversive conditioning.

Aversive conditioning elicits eyeblink reflex responses at CS onsets

As in the pilot study, we predicted that the presentation of a CS⁺ would result in a larger eyeblink response, compared with a neutral CS⁻ (Fig. 2D). Our statistical analyses suggest that this was indeed the case ($t = 5.07$, $p < 0.001$, $df = 3190.0$).

We found that the best fit of eyeblink responses was achieved with a learning rate of 0.64 and essentially no temporal discounting (i.e., $\lambda = 0.99$). The log-likelihood of this best fit ($-11,392.98$) was significantly better than the log-likelihoods of the permuted data (mean \pm SD, $-11,407.43 \pm 1.34$, z -score = 10.81 , $p = 7.9e-28$).

Regressing value and prediction error against eyeblink responses using the learning rate and temporal discount parameters derived from the behavioral pilot study (again to avoid overfitting) revealed a positive correlation between CS value and blink response ($t = 4.83$, $p < 0.001$, $df = 3188.0$). An analogous analysis of the blink reflex at CS onset was run for the appetitive conditioning session but did not yield any significant results.

Together, these results therefore replicate our findings from the behavioral pilot study (Fig. 2A,B) on the effect of appetitive and aversive conditioning on pupil and blink responses, respectively. In both datasets, fitting the pupil and eyeblink responses led to a lower learning rate in the appetitive session, compared with the aversive session. It should be noted that it is possible that this difference in learning rates may merely be the result of fitting different modalities of conditioned responses, which could potentially differ in how their expression evolves over time, without any differences in how quickly the associations have been acquired. At the same time, it has also been demonstrated that fMRI results are relatively robust to variance in learning rates when, for example, the temporal difference algorithm is applied for creating statistical regressors in computational fMRI (Wilson and Niv, 2015). Exploratory analyses also revealed this to be the case in the current dataset. Thus, differences in the learning rates across conditions do not substantively account for the differences in results found between conditions.

Reward TD prediction errors in the ventral striatum and amygdala

Previous research has found BOLD responses consistent with a TD RPE in the ventral striatum (O'Doherty et al., 2002, 2004; Schönberg et al., 2007) and the amygdala (Prévost et al., 2013). To validate our paradigm, we first tested whether BOLD responses in the ventral striatum correlated with TD RPEs. For this purpose, we used the prediction error estimates of the TD model at the onsets of CSd and CSp as a parametric regressor, applying the learning rates discovered through the analyses of eyeblink and pupil reflex data. The TD model makes different predictions

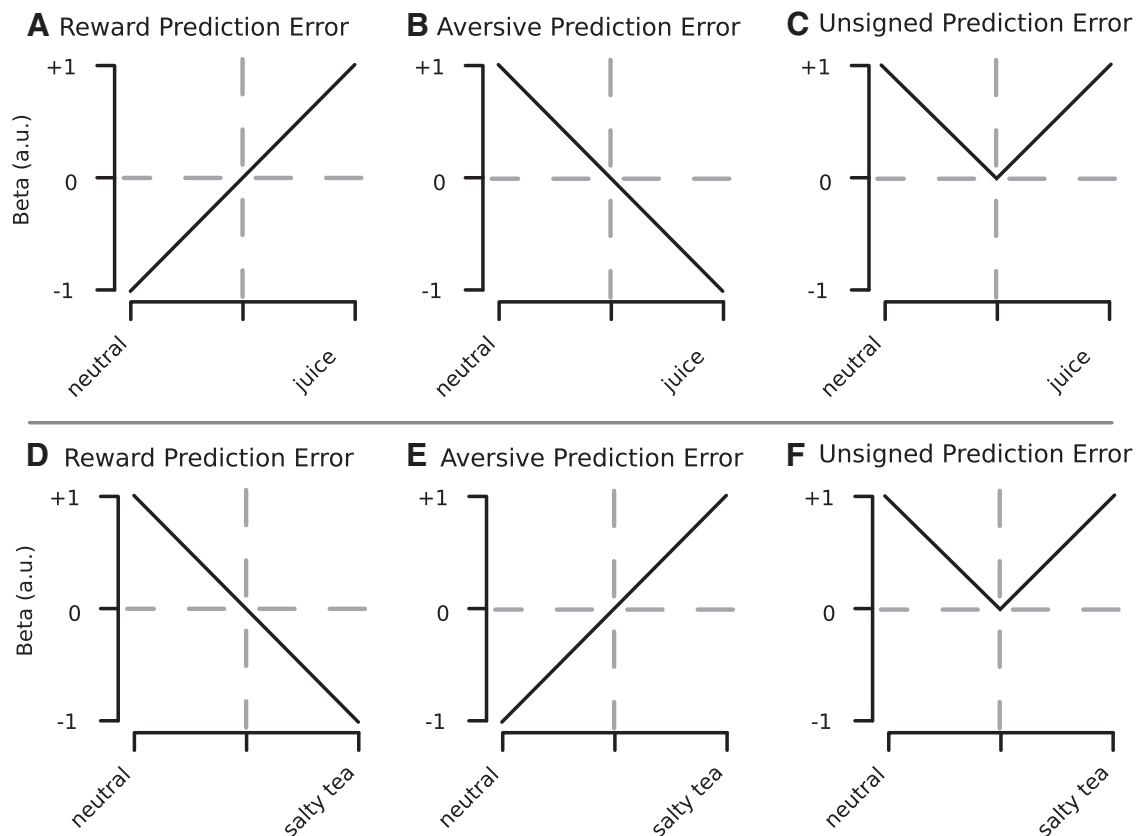
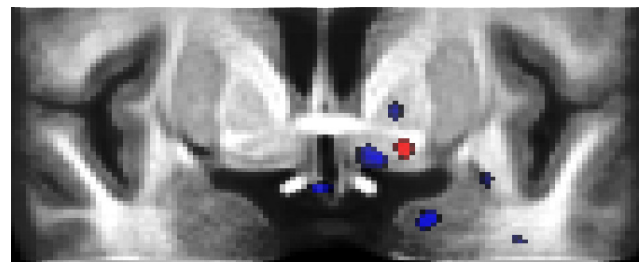


Figure 3. Top panel, Hypothetical BOLD responses for voxels signaling RPE (**A**), aversive prediction error (**B**), and unsigned prediction (**C**) in the appetitive session. **D**, **E**, Hypothetical BOLD responses for the same types of prediction errors in the aversive session.

about the prediction error response at the two cue presentation time-points (Fig. 3A). At the time of the most distal cue (CSD), the response is predicted to be synonymous with an expected value (EV) signal because the TD error signal computes the difference between the expected future reward predicted at that time point less the predicted reward at the previous time point (before the onset of CSD). Because the predicted reward is zero before the onset of the distal cue, the prediction error response at the time of the distal cue is simply the expected future reward at that time point. However, at the time of presentation of the more proximal cue (CSp), the response is predicted to depict the difference between the prior expectation (signaled by the CSD) and the new revised expectation elicited by the CSp. If, for example, the CSD predicting the delivery of the neutral liquid is followed by the CSp predicting the delivery of the appetitive liquid, then a positive prediction error will be elicited, whereas if the CSD cue predicting the onset of the reward stimulus is first presented followed by the CSp predicting the onset of the neutral stimulus, a negative prediction error will be elicited.

Consistent with previous findings, we identified a region in the ventral striatum (peak MNI coordinates: $-15, 1, -9$ mm; $t = 3.875, p = 4.1e-4, df = 22$) in which the BOLD response correlated positively with a TD RPE in the appetitive session (Fig. 4).

In addition to this previously reported finding, we also tested for a region correlating with the TD RPE during the aversive learning session, that is, a signal that indicates, for example, that an already expected aversive outcome will no longer be delivered. Specifically, we tested for a region in which, at the time of presentation of the more distal stimulus, activity is predicted to show increases the more a cue is associated with the absence of the



$y = 2$

Figure 4. In the appetitive session, the BOLD response in the ventral striatum correlated with the RPE estimates of the TD algorithm (red). In the aversive session, the BOLD response in the ventral striatum and the medial amygdala correlated with the RPE estimates of the TD algorithm (blue); that is, a deactivation for the unexpected presentation of a CS that indicated the delivery of an aversive liquid (Thresholded at $p = 0.005$).

aversive predicting stimulus, whereas at the time of the proximal cue presentation, activity is predicted to increase in response to an unexpected omission of the aversive predicting stimulus, and to decrease in response to the unexpected delivery of the aversive predicting stimulus (for illustration, see Fig. 3D). We found an adjacent region of ventral striatum (peak MNI coordinates: $-12, 3, -10$ mm; $t = -3.862, p = 4.2e-4, df = 22$) in which the BOLD response correlated positively with an RPE in the aversive session.

Next, we tested for a region in the medial amygdala that showed a pattern of BOLD responses in the aversive session that is also consistent with a TD RPE (Prévost et al., 2013). Confirming this previous finding, we identified a region in the medial amygdala (peak MNI coordinates: $-20, -3, -23$ mm;

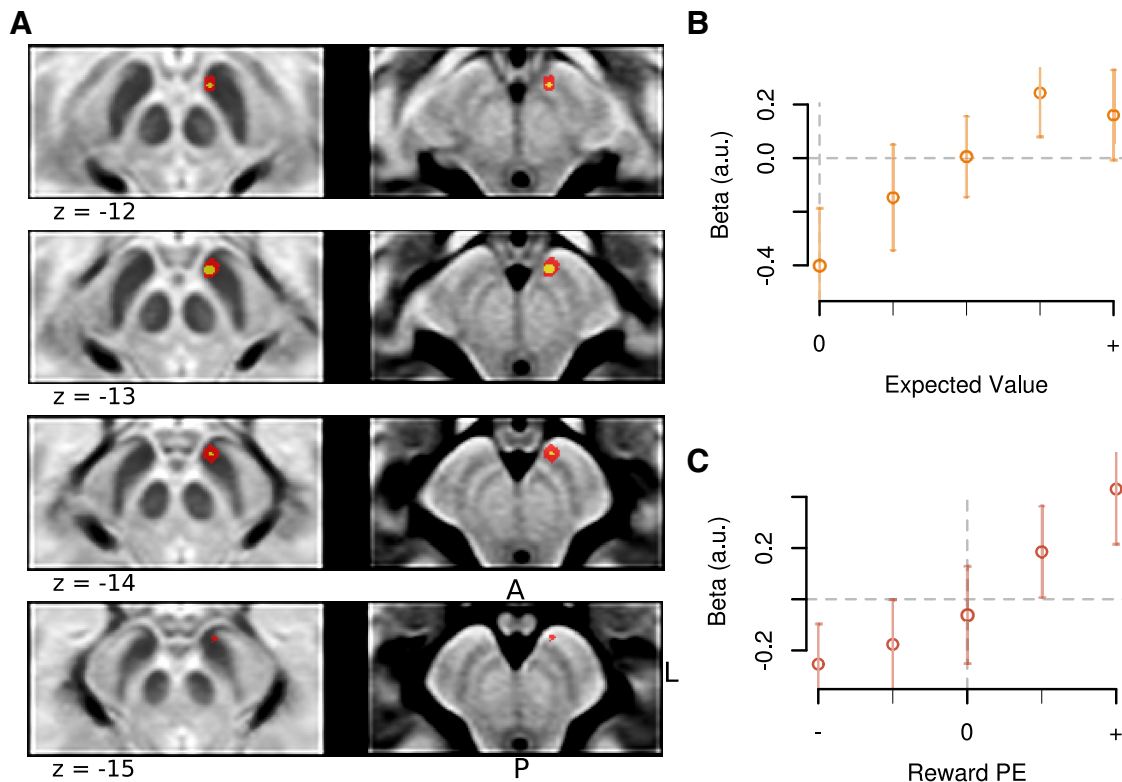


Figure 5. *A*, In the appetitive session, the BOLD response in the ventromedial SN correlated with the RPE estimates of the TD algorithm (red; $p < 0.005$; yellow; $p < 0.001$). Mean subject β values demonstrate how BOLD responses scaled linearly with TD RPEs in the appetitive session. *B*, At the time of the distal cue, the TD error response is predicted to be isomorphic to the expected value of the cue. *C*, Proximal cue. Subject β values were extracted from individual ROIs identified through a leave-one-subject cross-validation approach. In *B* and *C*, error bars indicate standard errors.

$t = -4.012$, $p = 2.9 \times 10^{-4}$, $df = 22$) in which the BOLD signal correlated positively with the TD RPE in the aversive session.

TD RPEs in the substantia nigra

Next, we examined whether different regions within the SN were selectively involved in encoding an appetitive or aversive temporal difference prediction error signal during appetitive and aversive learning. We first tested for a TD RPE signal. We identified a ventromedial region (peak MNI coordinates: $-5, -12, -12$ mm; $t = 3.99$, $p = 3.3 \times 10^{-4}$, $df = 22$) correlating with TD RPE during the appetitive conditioning session (Fig. 5*A*). The region we identified corresponds to the ventromedial midbrain SN (dopamine area A8), close to the border to the VTA (A9) (Eapen et al., 2011). To validate our model-based regression analysis and to establish the extent to which the identified cluster conformed to a TD error signal, we split the model-predicted prediction error signal into 5 bins for each time point (see Materials and Methods); and in a *post hoc* nonparametric regression analysis on the extracted time-series from that cluster, we estimated parameter estimates for these separate bins (using a leave-one-out analysis approach; see Materials and Methods). At the time of the distal cue presentation, the TD RPE should resemble a signal that would scale with the magnitude of the expected predicted future reward, which is exactly what we observed in the midbrain cluster (Fig. 5*B*). At the time of the proximal cue presentation, the activation should scale from negative to positive, depending on whether reward expectancies were positively or negatively violated following the presentation of the proximal cue, which is once again what we observed (Fig. 5*C*). Thus, the signal we found in the ventromedial SN conforms closely to the expected profile of a TD RPE signal during the appetitive learning session.

TD aversive prediction error

Next, we tested for regions correlating with a TD aversive prediction error signal during the aversive conditioning session. An aversive prediction error would show an increase in activity in response to increasing predictions of an aversive outcome in response to the most distal cue, and at the time of the proximal cue would show an increase in activity in response to the unexpected delivery of an aversive predicting stimulus or outcome, whereas the unexpected omission of an aversive predicting stimulus or outcome would result in a decrease in activity (Fig. 3*E*). However, we did not find any significant correlations with an aversive going prediction error signal anywhere in the SN (dopamine area A8). This was true not only in the aversive conditioning session, but also in the appetitive session (in which at the time of the proximal cue such a signal would show an increase in activity to the unexpected omission of reward; Fig. 3*B*).

Unsigned prediction errors

A previous electrophysiological study reported that a subpopulation of neurons in the SN are involved in a “saliency” type code, in which both unexpected aversive stimuli and unexpected appetitive stimuli produce an increase in neural activity (Matsumoto and Hikosaka, 2009). To investigate this question, we tested for a region within the SN that simultaneously expressed an increased BOLD response to unexpected aversive and appetitive conditioned stimuli in the aversive and appetitive session, respectively. However, no region within the SN showed a significant response profile of this sort.

Because a recent fMRI study reported UPE signals in midbrain nuclei (D’Ardenne et al., 2013), we next investigated whether any area in the SN showed a pattern of BOLD responses consistent

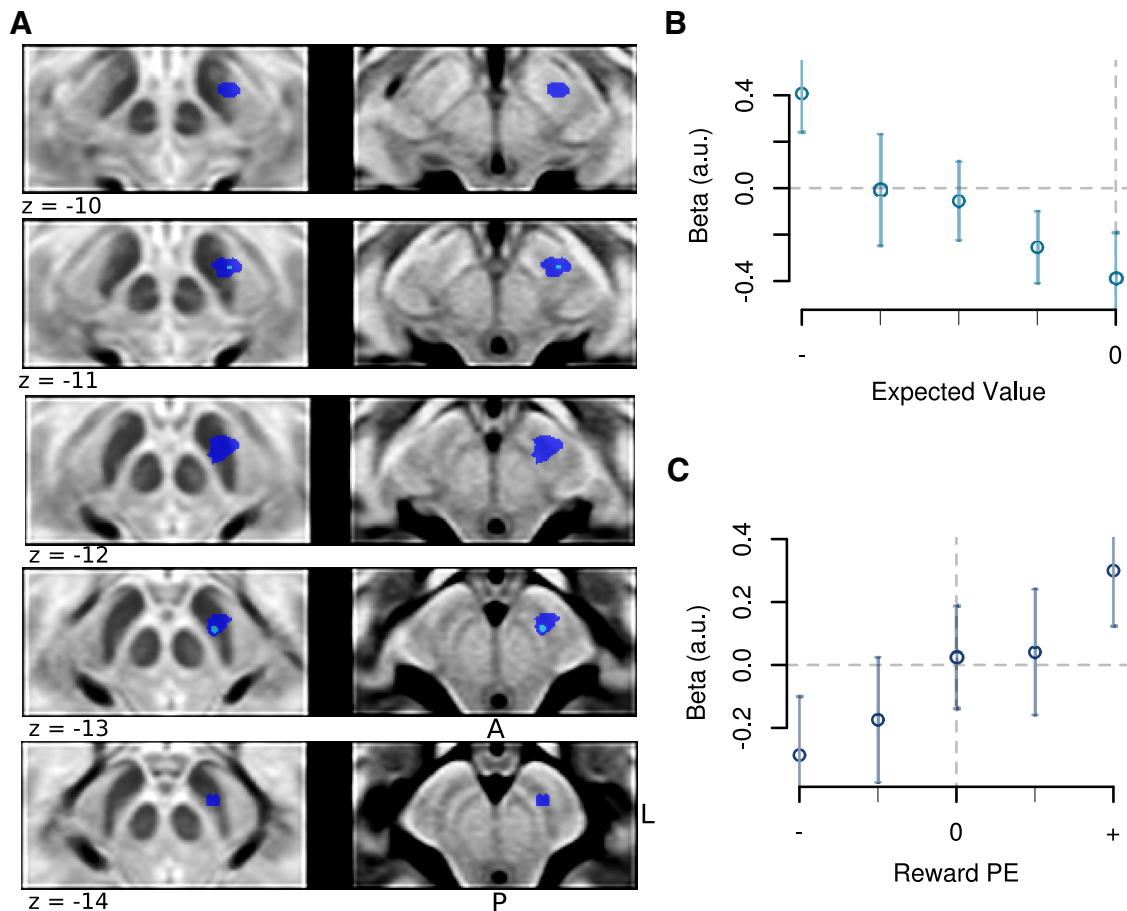


Figure 6. BOLD responses in the aversive session in the dorsolateral SN. **A**, In the aversive session, the BOLD response in the dorsolateral SN correlated with an UPE signal (dark blue; $p < 0.005$; light blue; $p < 0.001$). However, an interrogation of the underlying response profile in this region by plotting mean subject β values demonstrates that the BOLD response in this region reflects a more nuanced signal. **B**, At the time of the most distal cue, the BOLD responses scaled linearly with an aversive value signal. **C**, At the time of the proximal cue, activity correlated with an RPE.

with an UPE; that is, a region that exhibits a stronger BOLD response if unexpected conditioned stimuli are presented, relative to expected conditioned stimuli, independently of whether these conditioned stimuli predict an appetitive or aversive outcome (Fig. 3C,F). We identified a region in the dorsolateral SN which showed a pattern of BOLD activation that correlated with this UPE regressor (Fig. 6A; peak MNI coordinates: $-10, -17, -12$ mm; $t = 3.56, p = 8.8e-4, df = 22$).

To further investigate the response profile of this dorsolateral region, we next extracted the mean fitted parameter estimates from the voxels in this cluster from each individual subject and performed a stepwise linear regression procedure. Clusters were again defined by a leave-one-participant-out cross-validation procedure. This further analysis revealed that the signal did not indeed correspond to an UPE signal as predicted. An UPE signal should manifest as a “V-shaped” response profile at both time points (corresponding to an increased activation response both when an aversive outcome is predicted to be more and less likely to occur compared with the status quo). However, at the time of presentation of the most distal cue, activation instead showed a linearly increasing response profile as a function of increasing predictions of an aversive outcome, consistent with an aversive EV signal (Fig. 6B). On the contrary, at the time of presentation of the proximal cue, we found a signal that resembled an RPE signal, when a worse outcome than expected was revealed, activity decreased in this region, whereas when a better than expected outcome was revealed (i.e., less chance of an aversive outcome being delivered), activity increased (Fig. 6C). Thus, together,

the response profile in this region appears to exhibit a response that resembles a mix of two different signals during the aversive session: an aversive EV signal at the time of the most distal cue and an RPE signal at the time of the proximal cue presentation.

Given that our results indicate that the SN response in the aversive session reflects a combination of an aversive prediction signal at the time of the distal cue and an RPE at the time of proximal cue, in a *post hoc* analysis, we reran a new model-based fMRI analysis for the aversive session in which we set the parametric response at the time of the distal cue to be an aversive value signal and the parametric response at the time of the proximal cue to be a reward signal. Unsurprisingly, given the circular nature of this analysis (Kriegeskorte et al., 2009), these results showed a more robust effect in the dorsolateral SN than the initial unsigned error signal used in the analysis described earlier. Despite the circular nature of this analysis, to illustrate the topography of the responses in the midbrain, we plotted these results alongside the results from the RPE analysis in the appetitive session (Fig. 7A). Figure 7B shows the location of the SN with the standard MNI T1 whole brain as reference, with cross hairs indicating the location of the peak voxel representing RPEs in the appetitive session.

A functional gradient for value and prediction error signals within the SN during appetitive versus aversive learning

While the above results hint at a regional functional specialization within the SN, we next performed a direct test for this

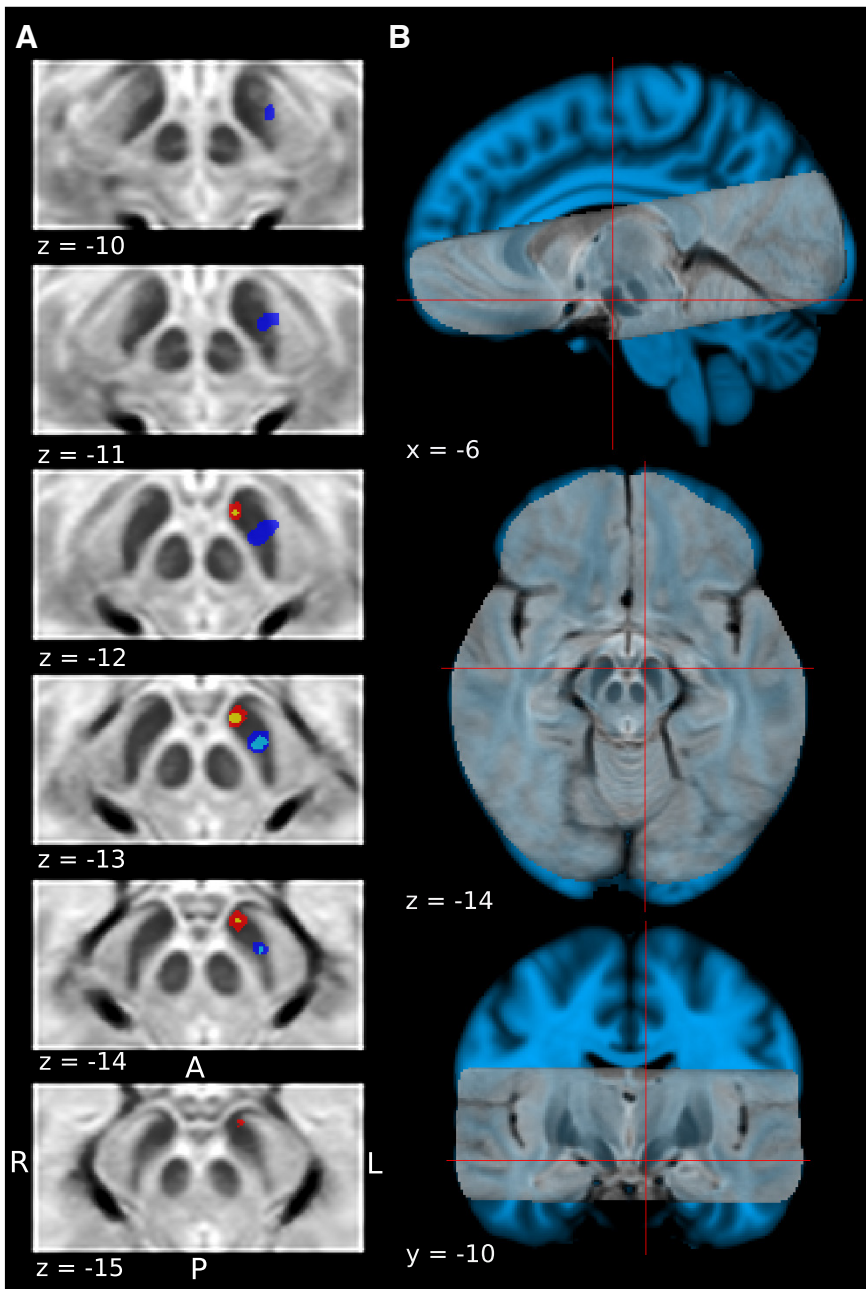


Figure 7. *A*, Comparison of extent of clusters showing a BOLD response consistent with an RPE, and a combination of expected aversive value and RPE at the onset of the distal and proximal cue, respectively. Color codes are the same as in Figures 5 and 6. *B*, High-resolution T2-weighted structural scan overlaid on top of the standard skull-stripped 1 mm MNI T1 brain for reference. Red cross hairs indicate location of peak voxel encoding an RPE signal in the appetitive learning session.

possibility. To do this, we identified the most ventromedial and caudal voxel along the rostrocaudal axis in the SN and used this voxel as an anchor point (Fig. 8*A*). We then binned the remaining voxels in the SN according to their distance from the anchor point. Voxels of the same distance were binned together. Within each voxel bin in each participant, we then estimated the average parameter estimates for the reward RPE in the appetitive condition per subject and estimated the average parameter estimates for the combined aversive value and RPE regressor in the aversive condition per subject. We then used a linear mixed-effects model to test for a significant interaction between distance and regressor (RPE in the appetitive session vs combined aversive and RPE at the distal and

proximal cue, respectively, in the aversive session) on the fitted parameter estimates. This analysis yielded a significant regressor by distance interaction (Fig. 8*B*; $t = 3.83$, $p = 1.34 \times 10^{-4}$, $df = 1287.0$), supporting the existence of a functional gradient within the SN, with voxels in the ventromedial region showing a greater tendency to be involved in responding to reward RPEs during the appetitive session, whereas voxels in the more dorsolateral region showed a greater tendency to respond to the aversive value signal (at the distal cue) and an RPE signal (at the proximal cue) during the aversive session.

Midbrain BOLD responses predict participants' affective reactions to cue stimuli 1 year later

Given that dopaminergic activity is hypothesized to be directly related to learning of stimulus–reward associations, we performed a test to determine the extent to which activity in the SN was related to the learning of long-term affective associations with the cues. One year after the fMRI experiment was completed, we invited participants to fill out a brief survey in which we presented pictures of the fractal stimuli used as cues in the experiment (1 year previously), and we asked them to provide affective ratings for those cues on a scale ranging from -5 to 5 , where -5 = very unpleasant, 0 = neutral, and 5 = very pleasant.

We first tested whether or not individuals' affective ratings for the cue were significantly different for the cues previously paired with reward, neutral outcomes, or aversive outcomes (Fig. 9*A*). The results of a linear mixed-effects model with the fixed effect objective cue value and random effect participant indicate that ratings of cue stimuli scaled linearly according to the probability with which they predicted the aversive or neutral outcomes in the aversive session ($t = 2.057$, $p = 0.046$, $df = 41$). No such effect was found for the cues of the appetitive session

($t = 0.416$, $p = 0.68$, $df = 41$).

We next tested whether the difference in affective ratings for the pleasant compared with the neutral and aversive compared with the neutral cues 1 year hence was correlated across participants with the BOLD response to the prediction error signals found in the SN 1 year previously.

For the cues used in the appetitive session, a significant correlation was found between the parameter estimates from the model-based analysis for the reward TD prediction error signal in the midbrain during the appetitive session and the difference in 1 year post affective ratings for the appetitive compared with neutral predicting cues ($\tau = 0.366$, $p =$

1.6e-4; Fig. 9B, left). Thus, even though, on average, participants did not show a significant change in their affective ratings for the reward predicting cues compared with the neutral predicting cues, those participants who did exhibit increased preference for the reward predicting cues had greater RPE responses in their midbrain when experiencing those cue-reward associations 1 year thence.

A similar analysis was performed for the aversive session and revealed a significant correlation between the parameter estimates from the model-based analysis for the unsigned TD prediction error signal in the midbrain during the aversive session and the difference in affective ratings for the aversive compared with neutral predicting cues during the aversive session ($\tau = -0.371$, $p = 1.4e-4$; Fig. 9B, right). Because, as reported earlier, we found that the UPE signal in the dorsolateral SN was indeed a combination of an aversive stimulus value signal at the onset of the distal cue and an RPE at the onset of the proximal cue, we next asked whether the corresponding parameter estimates of the two stimuli correlated differentially with how well participants remembered the stimulus–outcome associations in the follow-up survey. We found that only the parameter estimates of the aversive stimulus value signal at the time of the distal cue presentation correlated with participants' subsequent differential affective ratings of the stimulus–outcome associations ($\tau = -0.438$, $p = 2.1e-05$; Fig. 9C, left), but a similar correlation was not found for the RPE signal at the onset of the time of the proximal cue ($\tau = -0.011$, $p = 0.45$; Fig. 9D, left). An analogous analysis for the appetitive session yielded similar results, with a correlation between participants' ratings of conditioned stimuli and the EV signal at the onset of the distal stimulus ($\tau = 0.275$, $p = 0.002$; Fig. 9C, right), and no significant correlation between the RPE signal at the time of the proximal stimulus and the ratings ($\tau = 0.023$, $p = 0.39$; Fig. 9D, right). Interestingly, a similar analysis for the ventral striatum did not yield any significant correlation between fitted parameter estimates for the RPE cluster, in which the BOLD response scaled linearly with an RPE, and stimulus ratings in this follow-up survey.

Together, these results suggest that valuation responses in these two distinct regions of the SN during acquisition of appetitive and aversive conditioned associations exert an enduring influence on the expression of affective responses to those conditioned cues.

Discussion

The present study used computational high-resolution fMRI tailored for imaging the human SN to investigate the involvement of this region in encoding prediction errors during appetitive and aversive higher-order learning, as well as to determine potential regional specializations within the midbrain dopamine system in appetitive and aversive learning. We found that activity in the ventromedial SN, bordering with the VTA, during appetitive learning was consistent with an RPE signal: within the context of appetitive conditioning, activity in this region increased in response to the delivery of an unexpected conditioned stimulus predicting the delivery of reward. On the other hand, activity in

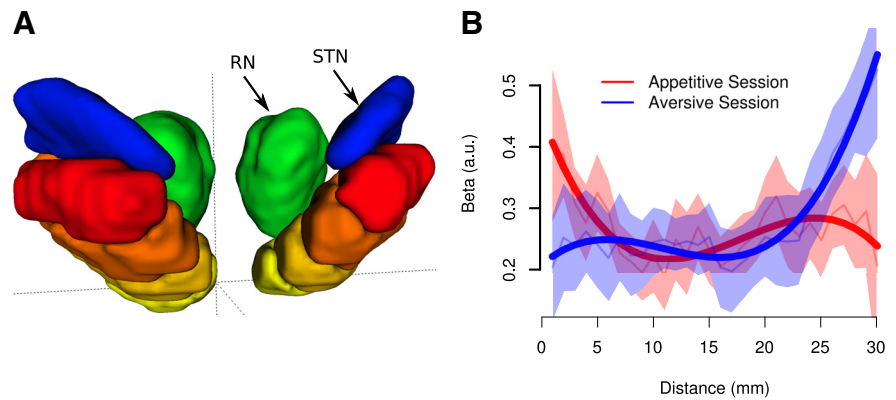


Figure 8. A direct test for a functional gradient within the SN revealed an increased tendency for the ventromedial SN to encode RPEs in the appetitive session, as well as an increased tendency for voxels in the dorsolateral SN to encode an aversive value signal and an RPE signal in the aversive session. **A**, Anterior view of SN, with voxel location along the ventromedial to dorsolateral axis indicated by heat color palette (binned for illustration purposes only). Location was defined in terms of distance from origin (crosshairs). Subthalamic nucleus (STN) and red nucleus (RN) were added for reference. **B**, Statistical analyses revealed that fitted parameter estimates show a significant location by condition interaction, confirming the existence of the proposed functional gradient along a ventromedial to dorsolateral axis. Thin lines indicate population mean for each set of equidistant voxels. Shaded areas indicate standard errors. Bold lines indicate predictions of the linear mixed-effects model.

the dorsolateral SN showed a more complex pattern of activity: an increase in activity at the time of the most distal cue predicted an aversive liquid, consistent with an aversive EV signal, whereas activity at the time of the more proximal cue correlated with an RPE signal, increasing in activity if the aversive outcome was deemed less likely than first expected. This combined signal may reflect a single unitary computational function, or else reflect the corepresentation of two distinct signals.

A striking feature of our results is that the ventromedial and dorsolateral areas of the SN differed functionally both with regard to their involvement in appetitive and aversive conditioning, respectively, and with regard to the types of signals found therein. The topographical variation in the functions of the SN that we found in the present study relates to that reported in an earlier neurophysiological investigation in nonhuman primates in which dopamine neurons in the ventromedial SN were found to predominantly encode RPEs, whereas those in the dorsolateral SN were suggested to be predominantly involved in encoding a “saliency” response, in which both unexpected aversive outcomes and unexpected reward outcomes were found to produce an increase in neural activity (Matsumoto and Hikosaka, 2009). However, in our case, the BOLD responses we found in the dorsolateral SN were indeed not a saliency signal but, on further inspection a mix of two distinct signals, an aversive value signal and an RPE. We tested for, but did not find, any regions within the SN that simultaneously expressed an increased BOLD response to unexpected aversive stimuli and appetitive stimuli when pooling across both appetitive and aversive sessions. Thus, while the results of our study in humans and that of Matsumoto and Hikosaka (2009) in monkeys agree in that there appears to be a topography in responses across the SN, the precise nature of the signal found in the dorsolateral SN does not correspond to that reported by Matsumoto and Hikosaka (2009). Naturally, as we are measuring BOLD responses and not single neuron activity specifically, we would not necessarily expect the signals reported in these studies to align precisely. One other possible reason for the difference between studies is that, in our study, the midbrain activity might reflect a contrast effect whereby, in the appetitive session, a neutral predicting stimulus is regarded as aversive because it is associated with missing out on the reward, whereas, in

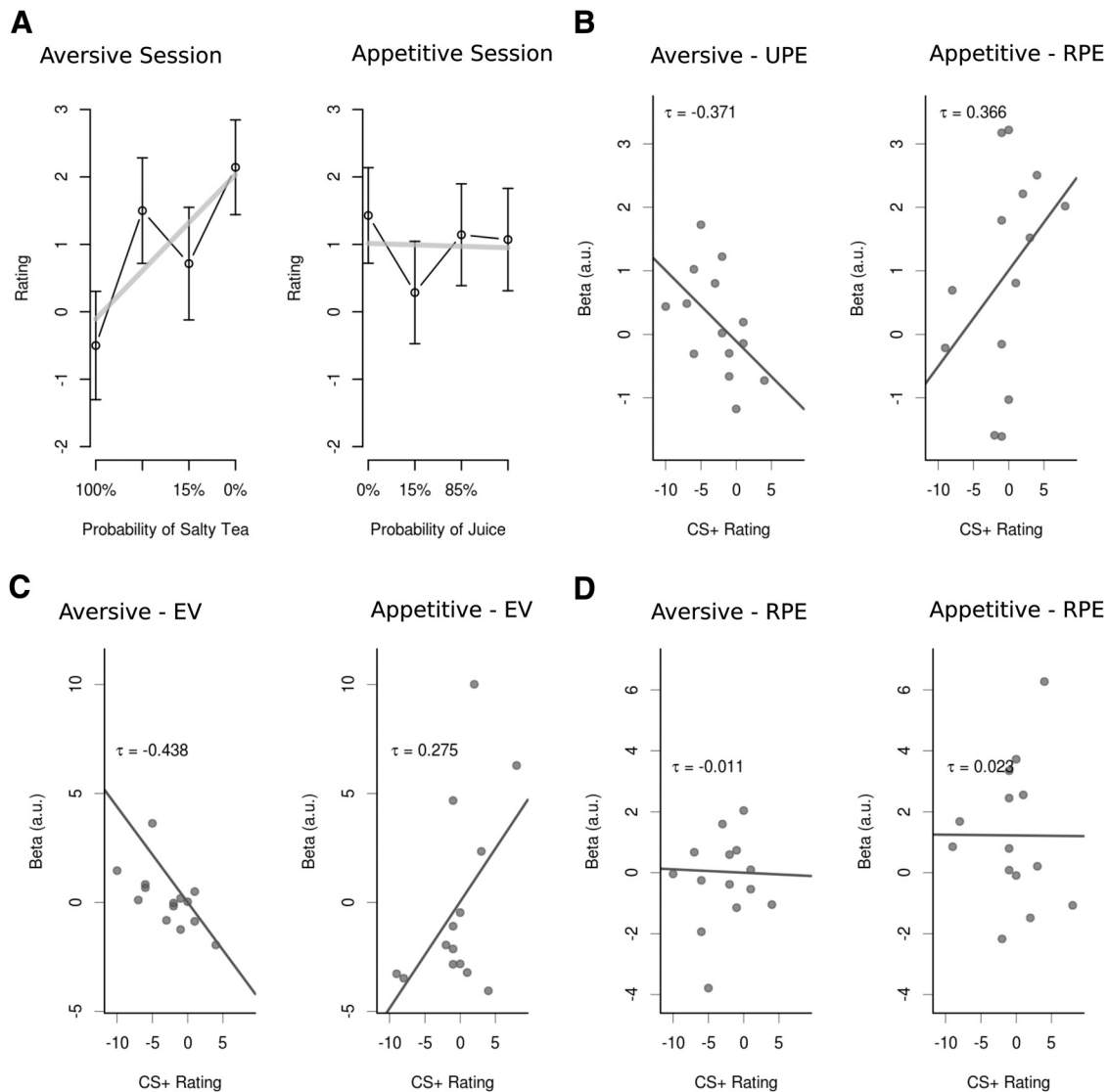


Figure 9. *A*, Participants correctly rated cue stimuli from the aversive learning session according to how accurately they predicted salty tea but failed to rate cue stimuli from the appetitive session appropriately. *B*, β values in the dorsolateral cluster for UPE correlated negatively with participants' ratings of CSs from the aversive learning session (left), and β values in the ventromedial cluster for RPEs correlated positively with participants' ratings of cue stimuli from the appetitive session (right). *C*, Participants' stimulus ratings correlated with dorsolateral β values for the EV signal at the time of the distal CS in the aversive session and showed a significant correlation with ventromedial β values in the appetitive session. No such correlation was found for the β values of to the RPE at the time of the proximal cue (*D*).

the aversive session, a neutral predicting stimulus is treated as appetitive because it is associated with avoidance of the aversive outcome (Kim et al., 2006). Flexible range adaptation has been reported in midbrain dopamine neurons for encoding of RPE signals (Tobler et al., 2005), so it is possible that such an adaptive coding mechanism is present more extensively within the SN.

It is worth noting that a previous fMRI study did report BOLD responses in the SN during aversive conditioning (Hennigan et al., 2015), whereas other studies reported a variety of reward-related prediction error-related responses in the same structure (Wittmann et al., 2005; D'Ardenne et al., 2013). However, no previous study to date has reported direct comparisons between responses during appetitive and aversive conditioning. Hennigan et al. (2015) did indeed include a reward-learning condition, but they did not find any response in the midbrain in that condition, perhaps because the aversive and appetitive cues were delivered in an intermixed fashion in that study (as opposed to in separate sessions as done here). With an intermixed paradigm, contrast

effects between the cues may have resulted in the aversive cue dominating (see also Kim et al., 2011). As a consequence, the present study is the first to definitely show responses in this region in the human brain during both appetitive and aversive learning, allowing us for the first time to detect a functional topography within the human midbrain related to these two distinct types of learning.

The discovery of a functional topography in humans is consistent with earlier studies in monkeys, which suggest that ventromedial and dorsolateral areas of the SN may indeed differ functionally (Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010), reflected by dissociable afferent and efferent connectivity patterns (Haber et al., 2000; Halliday, 2004). These distinct areas of the SN are associated with a different profile of neuro-anatomical connections. Whereas dopaminergic neurons in the ventromedial SN send afferents to the shell of the nucleus accumbens (Haber et al., 2000), ventromedial prefrontal cortex (Williams and Goldman-Rakic, 1998), and orbital frontal cortex

(Porrino and Goldman-Rakic, 1982), dopaminergic projections from the dorsolateral SN, on the other hand, have been found to largely target the putamen (Haber et al., 2000) as well as the dorsolateral prefrontal cortex (Williams and Goldman-Rakic, 1998; Haber and Knutson, 2010). Thus, the different types of signals being encoded in these two different parts of the SN may be reflected in turn in differences in the output of the dopaminergic neurons in those areas, which can correspondingly impact differentially on at least partly distinct upstream neural circuits.

The present findings also have implications for studies using model-based fMRI methods more generally (O'Doherty et al., 2007). When testing for areas correlating with an unsigned TD prediction error, we found robust correlations with this signal in dorsolateral SN. Yet on closer inspection, this signal actually reflected a blend of two distinct components: an EV signal and an RPE signal. Perhaps because the averaged combination of those two signals might resemble an unsigned surprise signal, BOLD activity in this region loaded significantly onto that regressor. As a consequence, the present findings do highlight the fact that it is important to interrogate further the response properties of voxels found to correlate with a computational fMRI regressor (indeed, as would be the case with any parametric regressor) to determine precisely which type of response profile is indeed being exhibited.

Remarkably, we also found that BOLD responses in the SN during both appetitive and aversive conditioning are significantly correlated with differential subjective affective ratings for the cue stimuli more than 1 year after the fMRI data were first acquired. Those individuals with a more robust ventromedial SN prediction error response during appetitive conditioning were more likely to exhibit increased liking for the cues paired with the appetitive outcome (compared with the neutral cues), whereas those individuals with a more robust dorsolateral SN response to expectation of an aversive outcome exhibited a greater tendency to rate the aversive predicting cues as less likable than the neutral predicting cues encountered in the aversive learning session 1 year earlier. One previous study also reported that visual stimuli that predicted monetary reward were associated with a stronger BOLD response in the SN, compared with non-reward-predicting pictures, and were also associated with better recollection and source memory 3 weeks later (Wittmann et al., 2005). Our findings extend on these previous results in that they demonstrate an analogous effect even 1 year later. More importantly, however, our results are the first demonstration that the strength of a long-term conditioned affective reaction is associated with the degree of activation in the SN during aversive learning as well.

In conclusion, we found evidence for dissociable contributions of the human SN during appetitive and aversive learning. Whereas activity in the ventromedial SN reflects a reward-related TD prediction error signal during appetitive learning, the dorsolateral SN reflects a mix of an aversive EV signal and an RPE. Furthermore, activity in the SN during both appetitive and aversive learning was ultimately predictive of the degree of affective reactivity to the cues that could be measured 1 year later. Our findings provide important additional insights into the unique functions of distinct parts of SN in different forms of learning and provide an important link between the functional contributions of the SN to learning in both human and nonhuman animals.

References

Avants BB, Yushkevich P, Pluta J, Minkoff D, Korczykowski M, Detre J, Gee JC (2010) The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* 49:2457–2466. [CrossRef Medline](#)

- Bates D, Maechler M, Dai B (2008) *Lme4* package. Available at <http://lme4.r-forge.r-project.org>.
- Bitsios P, Szabadi E, Bradshaw CM (2004) The fear-inhibited light reflex: importance of the anticipation of an aversive event. *Int J Psychophysiol* 52:87–95. [CrossRef Medline](#)
- Bray S, Rangel A, Shimojo S, Balleine B, O'Doherty JP (2008) The neural mechanisms underlying the influence of pavlovian cues on human decision making. *J Neurosci* 28:5861–5866. [CrossRef Medline](#)
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68:815–834. [CrossRef Medline](#)
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482:85–88. [CrossRef Medline](#)
- D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319:1264–1267. [CrossRef Medline](#)
- D'Ardenne K, Lohrenz T, Bartley KA, Montague PR (2013) Computational heterogeneity in the human mesencephalic dopamine system. *Cogn Affect Behav Neurosci* 13:747–756. [CrossRef Medline](#)
- Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* 10:1020–1028. [CrossRef Medline](#)
- Delgado MR (2007) Reward-related responses in the human striatum. *Ann N Y Acad Sci* 1104:70–88. [CrossRef Medline](#)
- Eapen M, Zald DH, Gatenby JC, Ding Z, Gore JC (2011) Using high-resolution MR imaging at 7T to evaluate the anatomy of the midbrain dopaminergic system. *AJNR Am J Neuroradiol* 32:688–694. [CrossRef Medline](#)
- Fiorillo CD (2013) Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science* 341:546–549. [CrossRef Medline](#)
- Fiorillo CD, Yun SR, Song MR (2013) Diversity and homogeneity in responses of midbrain dopamine neurons. *J Neurosci* 33:4693–4709. [CrossRef Medline](#)
- Fisher RA (1919) The correlation between relatives on the supposition of Mendelian inheritance. *Earth Environ Sci Trans R Soc Edinb* 52:399–433. [CrossRef](#)
- Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus–reward learning. *Nature* 469:53–57. [CrossRef Medline](#)
- Glover GH, Li TQ, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44:162–167. [CrossRef Medline](#)
- Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35:4–26. [CrossRef Medline](#)
- Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20:2369–2382. [Medline](#)
- Halliday G (2004) Substantia nigra and locus coeruleus. In: *The human nervous system* (Paxinos G, Mai JK, eds), Ed 2, pp 451–503. London: Elsevier Academic.
- Hennigan K, D'Ardenne K, McClure SM (2015) Distinct midbrain and habenula pathways are involved in processing aversive events in humans. *J Neurosci* 35:198–208. [CrossRef Medline](#)
- Kim H, Shimojo S, O'Doherty JP (2006) Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol* 4:e233. [CrossRef Medline](#)
- Kim H, Shimojo S, O'Doherty JP (2011) Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cereb Cortex* 21:769–776. [CrossRef Medline](#)
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540. [CrossRef Medline](#)
- Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, Malenka RC (2012) Input-specific control of reward and aversion in the ventral tegmental area. *Nature* 491:212–217. [CrossRef Medline](#)
- Limbrick-Oldfield EH, Brooks JC, Wise RJ, Padormo F, Hajnal JV, Beckmann CF, Ungless MA (2012) Identification and characterisation of midbrain nuclei using optimised functional magnetic resonance imaging. *Neuroimage* 59:1230–1238. [CrossRef Medline](#)
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron dis-

- tinctly convey positive and negative motivational signals. *Nature* 459: 837–841. [CrossRef Medline](#)
- Mirenovic J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–451. [CrossRef Medline](#)
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947. [Medline](#)
- Nassar MR, Wilson RC, Heasly B, Gold JJ (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366–12378. [CrossRef Medline](#)
- O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. *Neuron* 33:815–826. [CrossRef Medline](#)
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337. [CrossRef Medline](#)
- O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ (2006) Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49:157–166. [CrossRef Medline](#)
- O'Doherty JP, Hampton A, Kim H (2007) Model-based fMRI and its application to reward learning and decision making. *Ann N Y Acad Sci* 1104: 35–53. [CrossRef Medline](#)
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454. [CrossRef Medline](#)
- Pinheiro JC, Bates DM (2000) *Mixed-effects models in S, S-PLUS*. New York: Springer.
- Porrino LJ, Goldman-Rakic PS (1982) Brainstem innervation of prefrontal and anterior cingulate cortex in the rhesus monkey revealed by retrograde transport of HRP. *J Comp Neurol* 205:63–76. [CrossRef Medline](#)
- Preusschoff K, t'Hart BM, Einhäuser W (2011) Pupil dilation signals surprise: evidence for noradrenergic role in decision making. *Front Neurosci* 5:115. [CrossRef Medline](#)
- Prévost C, McNamee D, Jessup RK, Bossaerts P, O'Doherty JP (2013) Evidence for model-based computations in the human amygdala during pavlovian conditioning. *PLoS Comput Biol* 9:e1002918. [CrossRef Medline](#)
- Roesch MR, Calu DJ, Schoenbaum G (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci* 10:1615–1624. [CrossRef Medline](#)
- Satterthwaite FE (1946) An approximate distribution of estimates of variance components. *Biometrics* 2:110–114. [CrossRef Medline](#)
- Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J Neurosci* 27:12860–12867. [CrossRef Medline](#)
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27. [Medline](#)
- Schultz W (2010) Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct* 6:24. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667. [CrossRef Medline](#)
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 8:1234–1240. [CrossRef Medline](#)
- Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007) Differential encoding of losses and gains in the human striatum. *J Neurosci* 27:4826–4831. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) *Reinforcement learning*. Cambridge, MA: Massachusetts Institute of Technology. Available at: <http://webdocs.cs.ualberta.ca/~sutton/book/ebook/the-book.html>.
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, Schoenbaum G (2011) Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci* 14:1590–1597. [CrossRef Medline](#)
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645. [CrossRef Medline](#)
- Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2006) Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* 95:301–310. [CrossRef Medline](#)
- Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* 324:1080–1084. [CrossRef Medline](#)
- Williams SM, Goldman-Rakic PS (1998) Widespread origin of the primate mesofrontal dopamine system. *Cereb Cortex* 8:321–345. [CrossRef Medline](#)
- Wilson RC, Niv Y (2015) Is model fitting necessary for model-based fMRI? *PLoS Comput Biol* 11:e1004237. [CrossRef Medline](#)
- Wittmann BC, Schott BH, Guderian S, Frey JU, Heinze HJ, Düzel E (2005) Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron* 45:459–467. [CrossRef Medline](#)