Separate Brain Regions Code for Salience vs. Valence During Reward Prediction in Humans

Jimmy Jensen,[1] Andrew J. Smith,[2] Matthäus Willeit,[1] Adrian P. Crawley,[3] David J. Mikulis,[3] Irina Vitcu,[1] and Shitij Kapur[1,4]*

[1]*Schizophrenia Program and the PET Centre, Centre for Addiction and Mental Health, Toronto, Canada*
[2]*Department of Psychology, McMaster University, Hamilton, Canada*
[3]*Department of Medical Imaging, University of Toronto, Toronto, Canada*
[4]*Department of Psychiatry, University of Toronto, Toronto, Canada*

◆ ════════════════════════ ◆

Abstract: Predicting rewards and avoiding aversive conditions is essential for survival. Recent studies using computational models of reward prediction implicate the ventral striatum in appetitive rewards. Whether the same system mediates an organism's response to aversive conditions is unclear. We examined the question using fMRI blood oxygen level-dependent measurements while healthy volunteers were conditioned using appetitive and aversive stimuli. The temporal difference learning algorithm was used to estimate reward prediction error. Activations in the ventral striatum were robustly correlated with prediction error, regardless of the valence of the stimuli, suggesting that the ventral striatum processes salience prediction error. In contrast, the orbitofrontal cortex and anterior insula coded for the differential valence of appetitive/aversive stimuli. Given its location at the interface of limbic and motor regions, the ventral striatum may be critical in learning about motivationally salient stimuli, regardless of valence, and using that information to bias selection of actions. *Hum Brain Mapp 28:294–302, 2007.*
© 2006 Wiley-Liss, Inc.

Key words: computational models; ventral striatum; reward; fMRI

◆ ════════════════════════ ◆

INTRODUCTION

An animal's survival depends on its ability to predict and respond to appetitive as well as aversive stimuli. Anticipation of an appetitive stimulus facilitates approach while anticipation of an aversive stimulus facilitates avoidance. It has been suggested that the mesolimbic dopaminergic system, especially the ventral striatum, plays a critical role in these behaviors. The ventral striatum receives dopaminergic inputs from the ventral tegmental area, afferents from the basolateral amygdala and orbitofrontal cortex [Haber et al., 1995; Schoenbaum et al., 2003], and projects to the ventral pallidum [Mogenson et al., 1993; Mesulam, 2000]. Thus, it is ideally positioned to serve as a gateway from "motivation to action" in anticipation and response to rewarding stimuli [Mogenson et al., 1993].

While most of the data on this issue derive from animal behavioral experiments, studies in humans using fMRI show the ventral striatum to be activated in anticipation of monetary, gustatory, and olfactory rewards [Elliott et al., 2000; Berns et al., 2001; Breiter et al., 2001; Knutson et al., 2001; Gottfried et al., 2002; O'Doherty et al., 2002] and it has been suggested that the striatum's role in reward processing is dependent on the saliency associated with reward rather than the value itself [Zink et al., 2004]. Recently, studies [McClure et al., 2003a; O'Doherty et al., 2003] have shown that the pattern of these brain activations is consistent with the concept of "reward prediction," as formalized by the

temporal difference (TD) learning algorithm, a machine learning approach to reinforcement learning. The TD algorithm [Sutton and Barto, 1990] learns to predict the future rewards by calculating the difference between actual and predicted reward, called the prediction error (δ). As an example, McClure et al. [2003a] used an appetitive reward learning paradigm (delivery of juice) and showed that fMRI blood oxygen level-dependent (BOLD) responses in the left striatum are consistent with a prediction error pattern as generated by the TD model. A negative prediction error caused by the absence of juice when it was expected correlated with a decreased BOLD response, while a positive prediction error (receiving juice when unexpected) correlated with an increased BOLD response. O'Doherty et al. [2003] showed that over the course of learning the relationship between a visual stimulus and appetitive reward, the prediction error signal in the ventral striatum and the orbitofrontal cortex shifted in time from the unconditioned stimulus (US) to the conditioned stimulus (CS), again as predicted by the TD model. These findings fit nicely with nonhuman primate findings by Schultz and colleagues, who showed that the firing of midbrain dopamine neurons, which project to the ventral striatum, also conforms to the behavior of the TD learning algorithm [for discussions of phasic dopamine signaling prediction error, see, e.g., Montague et al., 1996; Schultz et al., 1997; Schultz, 2002].

Conventionally, TD represents primary rewards with positive values and primary punishments with negative values. A positive prediction error is elicited when the environment behaves better than expected (unexpected reward or omission of expected punishment) and a negative signal for worse than expected (unexpected punishment or omission of expected reward). The outcome of a missed reward and a punishment are thus similar in this model. We call this a fully signed model as the information contained in the prediction error signal is sufficient for achieving optimal behavior (i.e., behavior that maximizes reward) [Sutton and Barto, 1990].

In contrast, and within recent fMRI contexts [Jensen et al., 2003; Seymour et al., 2004], unexpected aversive events have been shown to recruit areas traditionally associated with TD-like prediction errors in human fMRI experiments, e.g., the ventral striatum. This raises the possibility of whether certain parts of the reward prediction circuitry treat rewards and punishments in a similar manner in as much as both of them are motivationally salient events. In other words, is the prediction error fully signed (sufficient for acquiring optimal behavior), or what we term "partially signed"? A partially signed model treats appetitive and aversive events equivalently and is effectively responding to the presence of an unpredicted significant event, and not due to whether that event is better or worse than expected. Such a signal still carries information about an event's magnitude and the degree to which it was expected, but not about its valence. As a result, the information from this model would be insufficient to learn optimal behavior and would require additional processing of valence-specific information else-

where. In the partially signed case, a positive prediction error indicates the presence of any unexpected significant stimulus, while a negative prediction error denotes the absence of an expected and significant stimulus.

The question of interest is whether the neural substrates consistent with the different prediction error models (fully signed or partially signed) can be identified and dissociated with fMRI, and if so, which one better characterizes the processing performed in the ventral striatum.

In the current report, we will reserve the term "reward prediction error" for the fully signed model and will use the term "salience prediction error" for the partially signed prediction error as it reports an unexpected biologically significant stimulus.

The first prominent finding linking prediction error in living systems to TD was the finding by Schultz et al. [1997], who showed that midbrain dopamine neurons, in primates offered a juice reward, follow a pattern of firing consistent with the reward prediction error. While the firing of these neurons by themselves cannot as easily be measured in humans, studies in humans have sought an analogous signal related to learning in the brain using fMRI [McClure et al., 2003a; O'Doherty et al., 2003; Seymour et al., 2004]. The successful human studies in this regard indicate that the BOLD signal in the ventral striatum shows a pattern analogous to TD, when humans are learning about primary reinforcers such as juice delivery and pain. However, positive and negative reinforcers have not been used within the same paradigm in earlier studies. So, going into this study, we had three major questions of interest. Do the same brain regions drive reward learning for appetitive and aversive events? Does the pattern of learning conform to the reward prediction error model or the salience prediction error model? And which regions show BOLD responses consistent with different elements of the prediction models? To investigate these, we examined fMRI BOLD responses to classical Pavlovian learning involving mixed appetitive, aversive, and neutral events within the same run.

MATERIALS AND METHODS

Subjects

Twenty subjects (13 women) aged 34 ± 9 years gave written informed consent and participated in the study according to the guidelines of the local ethics review board. In an initial separate session before the scanning, all subjects underwent a structured interview concerning their physical and psychiatric health history. Only healthy subjects were included.

Experimental Protocol

The paradigm was based on classical Pavlovian conditioning using a 33% partial reinforcement schedule with cutaneous electrical stimulation to the left index finger as aversive unconditioned stimulus ($US_{Aversive}$) and $5 as appetitive ($US_{Appetitive}$). The intensity of the 200-ms

US$_{Aversive}$ was titrated individually until it reached a level at which the subject said it was "unpleasant but tolerable." The US$_{Appetitive}$ consisted of a \$5 bill projected to the screen for 1 s and the subjects were told before the experiment that they would gain an additional \$5 for each \$5 bill they saw. Three conditioned stimuli, yellow, blue, and red circles, with a duration of 5 s were used. The US$_{Aversive}$ immediately followed the offset of one of the colored circles (CS$_{Aversive}$) in 33% of the trials. Similarly, the US$_{Appetitive}$ followed another colored circle (CS$_{Appetitive}$) in 33% of the trials while the third colored circle (CS$_{Neutral}$) had no programmed consequences. A fixation cross was presented between trials and a fixed intertrial interval of 8 s was used. The experiment consisted of a total of 120 randomized trials; 30 nonreinforced trials of each type (CS$_{Aversive}$, CS$_{Appetitive}$, and CS$_{Neutral}$) and 15 reinforced trials of the two affectively significant types (reinforced CS$_{Aversive}$ and reinforced CS$_{Appetitive}$).

Apparatus

The US$_{Aversive}$ was delivered by a stimulating bar electrode (30 mm electrode spacing; Chalgren Enterprises, Gilroy, CA) placed on the left index finger using a gel as electrolyte. The electrode was attached to a Grass Instruments SD-9 stimulator (Grass-Telefactor, West Warwick, RI) via well-isolated coaxial cable leads through a waveguide. The subjects used an adjustable mirror located above their eyes to view the back-projected images on a screen placed at the foot of the scanner bed. The E-prime software (Psychology Software Tools, Pittsburgh, PA) controlled the stimulus presentations and triggered the stimulator. Galvanic skin response (GSR) was continuously monitored by PowerLab 2/20 (AD Instruments, Castle Hill, Australia) via long, well-isolated cables through a waveguide. MRI-compatible Ag/AgCl electrodes attached to the terminal phalanx on the left middle and ring finger respectively were used.

Image Acquisition

MRI scans were acquired by a GE Signa 1.5 T scanner (General Electric, Waukesha, WI) equipped with a standard head coil. In a single session, 700 volumes (28 contiguous axial 4.4-mm–thick slices) covering the whole brain were acquired using a T2*-sensitive spiral sequence [repetition time (TR) = 2,240 ms; echo time (TE) = 25 ms; flip angle = 85°; matrix = 64 × 64; field of view (FOV) = 200 × 200 mm]. For localization purposes, IR-Prepped 3D FSPGR T1-weighted anatomical images (124 contiguous axial 1.5 mm thick slices) were acquired (TR = 12 ms; TE = 5.4 ms; flip angle = 20°; matrix = 256 × 256; FOV = 200 × 200 mm).

Data Quality

The images were visually inspected for signal dropout due to magnetic susceptibility in the region of ventral striatum. Volumes acquired during shocks were discarded for all subjects as some of the images showed artifacts in the slices obtained during the delivery of the US$_{Aversive}$. Five subjects'

fMRI data could not be used. One subject withdrew due to feelings of claustrophobia before the experimental scan had started, one scan was not successfully reconstructed due to technical problems, one subject's images showed signal dropout in the region of the ventral striatum, one subject's images contained artifacts, and one subject moved more than the allowed 2.5 mm during the scan. Further, five subjects could not correctly report the contingency pattern of the CSs-USs. These subjects reported random contingencies and could not explain any relationship after the experiment according to their written self-reports or when these were reviewed verbally. GSRs were obtained for three of these five subjects and showed no signs of learning. Thus, based on self-reports and GSRs, 5 subjects did not learn the contingencies and were dropped from analysis and subsequently 10 subjects' data were used in the analyses.

SPM99 Analysis

All volumes were realigned to the first volume [Friston et al., 1995b] and the anatomical image was coregistered to the mean functional image to ensure that they were aligned. Finally, the images were spatially normalized [Friston et al., 1995a] to a standard EPI template [Evans et al., 1993], resampled at 3 × 3 × 3 mm, and smoothed using a 10-mm full-width half-maximum (FWHM) isotropic kernel. Data were high-pass-filtered using a cutoff value of 128 s and low-pass-filtered using a hemodynamic response function.

fMRI Data Analyses Using Subtraction Analyses

The data were analyzed by modeling five event types as stick functions convolved with a synthetic hemodynamal response function (HRF). The five events consisted of the three CS onsets but with the reinforced events modeled separately as two separate regressors. The latter two regressors were not used in any contrasts. The individual contrast images were moved up to a second-level random-effects model. The data were thresholded at $P < 0.01$ (uncorrected) and small volume corrections for regions of interest (ventral striatum, amygdala, anterior insula, and orbitofrontal cortex) were used based on coordinates from a previous study from our group [Jensen et al., 2003] and the coordinates reported by Anderson et al. [2003]. Spheres with a radius of 10 mm were used for the ventral striatum and amygdala while 15 mm in the other regions of interest.

fMRI Data Analyses Using TD Model-Generated Regressors for Prediction Error at CS

Following a number of existing accounts, the temporal difference learning model was used to generate reward prediction errors for comparison with the fMRI data. We adopt the standard approach [for example, see O'Doherty et al., 2003], which effectively starts with the assumption of a discrete number of states for representing the CS, the US, and the interval between them. The learning rate (α), the discount factor (γ) that determines the extent to which rewards that arrive earlier are more important than rewards
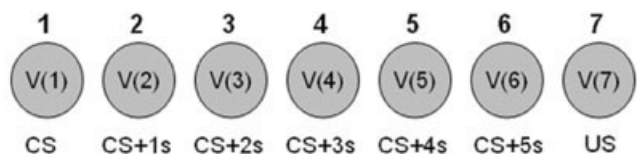
$$\text{For current state, } n, \text{ prediction error } \delta(n) = \underline{r(n)} + \gamma\underline{V(n)} - V(n-1)$$

$$\text{Update } V(n\text{-}1) := V(n\text{-}1) + \alpha\ \delta(n)$$

Figure 1.
The standard tapped delay line assumption effectively yields a set of states for representing a trial. Each state is given a unique index (written above the circles), and each state maintains a value estimate (inside each circle). The value estimate is central to all temporal difference methods and represents, for each state, the future expected reward from this state. Each state has an intrinsic reward value, r, which is supplied by the environment on activation of that state. The reward value of the US state was set to 1 or -1, while the reward value of all other states was set to 0. Each time a state is entered, a prediction error signal is generated. This signal, which is positive for "better than expected" and negative for "worse than expected," is then used to update the previous value estimate. The value estimate of each state is updated once per trial, and over successive trials, the prediction error signal effectively moves from the US to the CS via the intermediate states. As learning progresses, the value estimates represent with increasing accuracy the future reward associated with each state. An implicit assumption is made that the appropriate state is recognized based on the current position within the current trial type. The above representation was maintained separately for the three different CS types ($CS_{Appetitive}$, $CS_{Aversive}$, and $CS_{Neutral}$).

that arrive later on, and the number of intermediate timing states between CS and US (N) are all parameters that must be set by hand. A brief ad hoc search for a good set of parameters yielded $\alpha = 0.2$, $\gamma = 1$, $N = 5$. However, it is important to note that the qualitative behavior of the model is robust to a range of such parameters, as is the model's match to the fMRI data. It is also possible that better parameters may have existed. Other studies have used similar parameters [O'Doherty et al., 2003, 2004], although $N < 5$ is more common. The appetitive US was arbitrarily assigned a reward value of 1, and the aversive US was assigned a reward value of -1, i.e., $r(US_{appetitive}) = 1$ and $r(US_{aversive}) = -1$. All other states were assigned a reward value of 0, and all value estimates were initialized to 0. The first state was assumed to be unpredictable and therefore always generated a prediction error that was equal to its value estimate. Figure 1 shows an overview of the model.

The first prediction error vector was created based on the values of $\delta(1)$ for each appetitive trial. A second prediction error vector was similarly created from the values of $\delta(1)$ for each aversive trial. A third vector was created based on $\delta(1)$ for neutral trials (all these values were 0). There is no interaction between any of the different trial types because only

the appetitive trials influenced the appetitive value estimates, etc.

The three vectors were then combined into one vector, which reflected the order of the trials as they were actually presented to subjects. The values of the vector were resampled using linear interpolation to yield values at the start of volume acquisitions where CS onsets occurred.

The vector was then used in two ways. One, in order to examine which brain region activities correlated with the reward prediction error (fully signed), the vector was used as described above. Two, for the salience prediction error (partially signed), all prediction error values for the aversive trials were inverted. The model could have been used to generate these values directly by setting $r(US_{Aversive}) = 1$.

The estimated values in the collapsed vector were then convolved with a hemodynamic response function in order to obtain the characteristics of the BOLD response. Thus, for each vector (regressor), one value per volume was obtained. Data were thresholded at 0.01 (uncorrected) searching for significant activations using P values corrected for cluster size. The individual contrast images were moved to the second level for a random-effects analysis.

fMRI Data Analyses Using TD Model-Generated Regressors for Prediction Error at US

The model described above was also used to generate a prediction error signal at the time of the US. The approach was exactly the same as above except that the vectors were constructed from the values of $\delta(7)$ rather than $\delta(1)$. These values were then treated in a similar way as in the analysis above.

GSR Recording and Analysis

The GSR was sampled at 10 Hz. GSR data for eight subjects only were available due to technical problems for two.

To correct for possible MRI-induced artifacts, the GSR signal was digitally low-pass-filtered using a cutoff value of 2 Hz. To determine the GSR, the peak amplitude value within the 10 s following the cue onset was taken and subtracted by the value at the CS onset. The frequency of values higher than 0.05 µS was calculated for each of the five trial types modeled: $CS_{Aversive}$, $CS_{Neutral}$, $CS_{Appetitive}$, $US_{Aversive}$, and $US_{Appetitive}$.

Self-Reports

The subjects were asked to rate their degree of uneasiness and excitement when the CSs were shown. This was done on scales with four anchors (relaxed to extremely uneasy and neutral to extremely excited, respectively). In comparisons, we used Wilcoxon's rank-order test.

RESULTS

Behavioral measures confirmed that CS-US learning took place in the analyzed subjects. GSRs recorded during the fMRI session showed significantly higher frequency of re-

sponses above threshold (> 0.05 μS) during both $CS_{Appetitive}$ and $CS_{Aversive}$ as compared to $CS_{Neutral}$ [$CS_{Aversive}$ (58%) vs. $CS_{Neutral}$ (24%) t(7) = 4.23; $P < 0.01$; $CS_{Appetitive}$ (31%) vs. $CS_{Neutral}$ (24%) t(7) = 3.08; $P < 0.05$]. There were more GSR activations above threshold for aversive events as compared to appetitive [$CS_{Aversive}$ (58%) vs. $CS_{Appetitive}$ (31%) t(7) = 3.77; $P < 0.01$]. In comparison, the GSR frequency above threshold for events including $US_{Aversive}$ was 97% and $US_{Appetitive}$ 66%. In a postscan self-report, subjects reported a higher degree of discomfort with the $CS_{Aversive}$ vs. $CS_{Neutral}$ ($P < 0.01$) and a greater degree of excitement with $CS_{Appetitive}$ vs. $CS_{Neutral}$ ($P = 0.06$). Taken together, the data confirm that the unconditioned stimuli were salient, of opposite valence, and the subjects learned the associations between CS and US.

To address the questions above, we analyzed the fMRI data in three complementary ways: a conventional subtraction analysis of BOLD response induced by $CS_{Appetitive}$, $CS_{Aversive}$, and $CS_{Neutral}$ events; an analysis of the BOLD responses to CS onset based on the TD model's predictions for both the reward and salience models; and analyses of the responses to US based on both models' predictions to probe the validity of the TD hypothesis as fully as possible within the limitations of the protocol by considering all expectation-violation conditions.

As our a priori focus was on the ventral striatum, we used a conventional subtraction analysis and tested the effects using a random-effects model using small volume correction (SVC). In keeping with the logic of SVCs, the locations of the spheres for regions of interests were decided a priori based on coordinates reported previously [Anderson et al., 2003; Jensen et al., 2003]. The analyses revealed no significant activations when contrasting $CS_{Appetitive}$ vs. $CS_{Neutral}$ or $CS_{Aversive}$ vs. $CS_{Neutral}$. For the salience contrast [$(CS_{Appetitive} + CS_{Aversive})/2 - CS_{Neutral}$], an activation of the right ventral striatum (Fig. 2B; peak coordinates = 9, 6, −3; Z = 3.01, $P_{SVC} < 0.05$) was obtained. Parameter estimates (Fig. 2A) suggest that the $CS_{Aversive}$ had a larger effect in this region as compared to the $CS_{Appetitive}$, although both are above the mean, while the $CS_{Neutral}$ was not. No other activations in the regions of interests or in other regions of the brain were obtained when correcting for multiple comparisons for any of the contrasts.

To identify regions that were more sensitive to aversive valence, we used the ($CS_{Aversive} - CS_{Appetitive}$) contrast and it showed activation in the right anterior insula (Fig. 2D; peak coordinates = 24, 30, −6; Z = 3.46; $P_{SVC} < 0.05$). The opposite, appetitive valence contrast ($CS_{Appetitive} - CS_{Aversive}$), yielded an activation in the medial orbitofrontal cortex (Fig. 2C; peak coordinates = 3, 30, −18; Z = 4.21; $P_{SVC} < 0.01$). The $CS_{Aversive}$ and $CS_{Appetitive}$ showed no significant differences in activations within the ventral striatum. No other significant activations were found in the regions of interests or in any other regions when corrected for multiple comparisons.

To get better anatomical precision of the degree of overlap between aversive and appetitive activations within the ven-
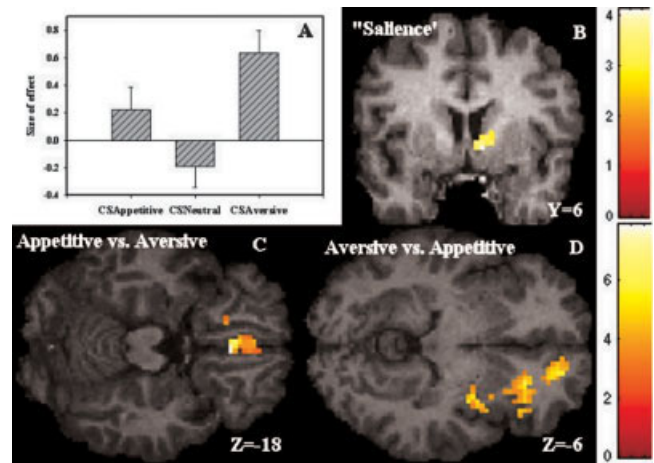

Figure 2.
A: The size of effects for CSs are shown for peak voxel in the ventral striatum (9, 6, −3) using the contrast [($CS_{Aversive}$ + $CS_{Appetitive}$)/2 − $CS_{Neutral}$]. **B:** A statistical parametric map (SPM) showing activations in the ventral striatum using the same contrast as in A. **C:** An SPM obtained with the contrast $CS_{Appetitive}$ vs. $CS_{Aversive}$ showing activations in the medial orbitofrontal cortex. **D:** An SPM with the contrast $CS_{Aversive}$ vs. $CS_{Appetitive}$ showing activations in the right anterior insula. The colors refer to *t*-values as coded in the bars to the far right. The upper bar refers to B and lower to C and D.

tral striatum, we also analyzed data from each individual without normalization or smoothing. To do this, we drew individual ROIs centered at the peak voxel coordinate obtained in the group analysis (approximately 9, 6, −3) with a spherical search volume with a radius of 10 mm and counted the number of voxels above the threshold of 0.05 uncorrected. The total search volumes consisted of 82–105 voxels. For $CS_{Appetitive}$, 9.8% ± 10.9% of the voxels were activated, whereas 13.0% ± 14.1% for $CS_{Aversive}$ were activated. Using a conjunction analysis for single subjects as implemented in SPM99 (http://www.fil.ion.ucl.ac.uk/spm), 11.1% ± 11.4% of the voxels in the total search volume were involved in both appetitive and aversive events. This value significantly differed from 0 [t(9) = 3.08; $P < 0.05$]. It should, however, be noted that the statistical thresholds for conjunction analyses used in SPM99 end up being more liberal toward each trial type as compared to regular analyses.

The foregoing suggests that certain brain regions (notably the ventral striatum) have a central role in learning about salience, while others differentially code valence. This raises two important questions: whether these regions act in conformity with the concept of prediction error, and whether the distinction between salience and valence observed in activation (subtraction) studies is also relevant in prediction error modeling.

In the traditional TD model, the prediction errors carry information not only for the salience, but also for the valence of events. Over successive learning trials, the prediction

Figure 3.

The left panel displays the cross-correlation co-efficient between the predicted signal yielded by the partially signed TD model and the MRI signal in a sphere with a 6 mm radius around peak voxel (coordinate 6, 3, −3) in the ventral striatum. The dashed line indicates the confidence limit and the x-axis shows lag number (TR = 2,240 ms). The right panel is the statistical parametric map of the activation in the ventral striatum correlating with the salience prediction error obtained at the CS onset with the partially signed model. The colors refer to t-values.
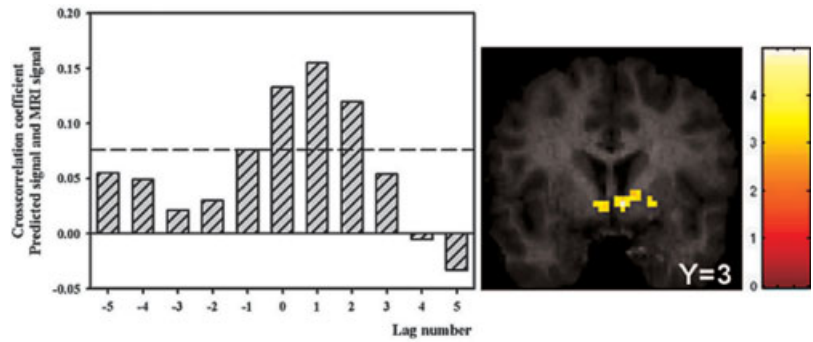
error migrates from the US to the CS as the US becomes increasingly predicted. Using this fully signed model, reward prediction error values were generated for each trial at the CS onset and these were then convolved with a hemodynamic response function and then correlated with the empirically recorded fMRI data. This model resulted in no significant clusters in the ventral striatum or in the other regions of interests. Using a more liberal statistical threshold (uncorrected $P = 0.05$) and spatial extent thresholding (i.e., cluster statistical weight), we found no cluster in the ventral striatum but an activation in the left prefrontal cortex (peak at coordinate −12, 63, 9; $Z = 3.34$; $P_{extent} < 0.05$).

To test whether ventral striatum activations conformed to a more general salience prediction error, we adapted the fully signed model to yield the partially signed model by treating appetitive and aversive events similarly, i.e., the unexpected presentations of an appetitive or aversive reinforcer (or CS predictive of either) gave rise to positive prediction error, and the omission of the same reward led to a negative prediction error. Using spatial extent thresholding, a single cluster was obtained in the entire brain peaking in the right ventral striatum (i.e., 6, 3, −3; $Z = 3.36$; $P_{extent} < 0.001$; Fig. 3), suggesting that the ventral striatum processes salience prediction error. This region was coextensive with the independently determined activation seen in the simple subtraction analysis above (Fig. 2A), thus showing that ventral striatal activity corresponded to valence insensitive salience prediction error.

A further exploration of the TD model's ability to explain activations was made by examining prediction error at the times of US deliverances. Since we used a partial reinforcement schedule, on some occasions the predicted US appeared and on others it did not, thus causing the prediction error signal to change from above baseline to below baseline on a trial-by-trail basis (in either variant of the TD models used above). This provided a more rigorous test of the prediction error hypothesis than just looking at the response to the CS, which moves slowly and steadily toward asymptote. The prediction error pattern generated by the fully signed model yielded a cluster in the right somatosensory cortex (24, 6, 60; $Z = 3.39$; $P_{extent} < 0.05$). The prediction errors based on the partially signed model (positive = unexpected salience; negative = omission of expected salience)

robustly correlated with activations in the left ventral striatum, bilateral anterior insula, and medial orbitofrontal cortex (Fig. 4, Table I).

DISCUSSION

We confirm in these studies a central role for the ventral striatum in processing of rewarding events, whether they are appetitive or aversive. Further, we demonstrate that the BOLD fMRI signal shows deviations consistent with the TD learning prediction error. The unique contribution of these findings is that they raise the possibility that a salience (as opposed to reward) prediction error signal may better characterize how the BOLD signals change.

These data are consistent with reports in humans by McClure et al. [2003a] as well as O'Doherty et al. [2003], who found support for a role for the striatum in reward prediction error within appetitive tasks. While McClure et al. [2003a] and O'Doherty et al. [2003] did not examine aversive events, a recent report by Seymour et al. [2004] showed that fMRI BOLD signal in the striatum and anterior insula correlated with the TD model predictions in pain learning. The current study confirms these previous findings and, by using concurrent appetitive and aversive stimuli, allows us to dissociate salience and valence of rewarding events in the context of prediction error signaling. The results lead us to suggest that these previous findings can be reconciled under a more general rule: the ventral striatum activations are consistent with signaling an error in predicting the salience of a stimulus, regardless of its valence. While in the current study, salience implicates appetitive and aversive events only, it is plausible that it could also apply to merely novel events as organisms treat them as salient and react to them.

McClure et al. [2003b] and Montague et al. [2004] have recently shown how prediction error signaling and the linking of mesolimbic dopamine system to salience can be merged within the computational accounts of TD learning. Our data suggest that their ideas would apply not only to appetitive but also to aversive situations. This is in keeping with the well-replicated finding that blocking dopamine transmission, particularly in the ventral striatum (nucleus accumbens) in animals, diminishes the motivational salience
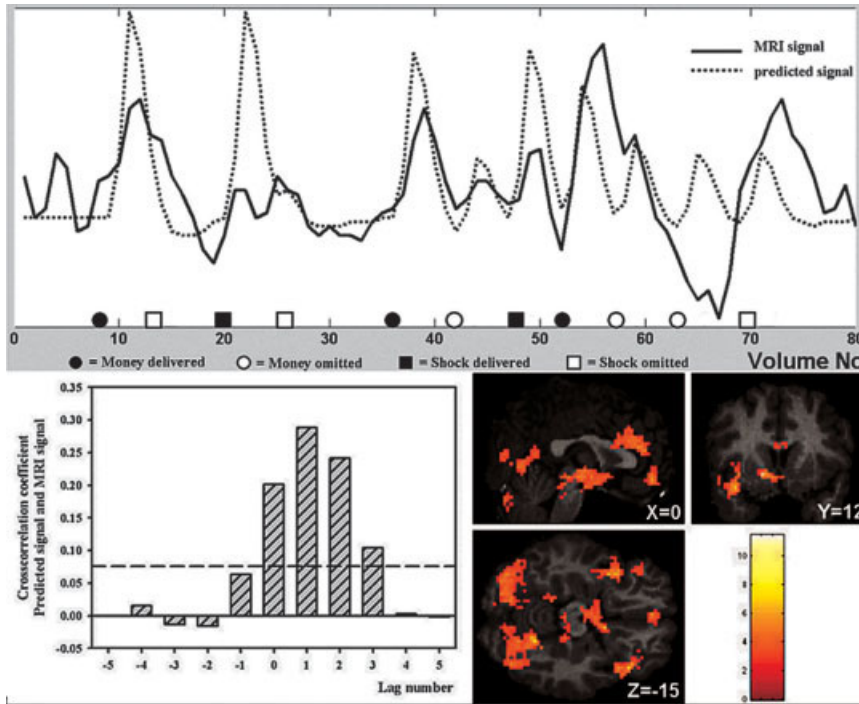
Figure 4.
The upper panel is showing predicted signal at US yielded by the partially signed TD model (dotted) and the actual MRI signal (solid) obtained from a sphere with a radius of 6 mm around the peak voxel (−15, 12, −6) in the ventral striatum. Values are mean-centered and the *x*-axis indicates volume number (TR = 2,240 ms). While only the first 80 volumes are shown for clarity, they are representative of the entire session as shown by the strong cross-correlation between the predicted signal from the partially signed model and the MRI signal (lower left panel). The dashed line indicates the confidence limit and the x-axis shows lag number where lag number reflects number of time points (TRs). The lower right is the SPM of the activation in the ventral striatum correlating with the salience prediction error obtained at the US by the partially signed version of the TD model. The colors refer to *t*-values.

of both appetitive as well as aversive reinforcers [Salamone et al., 1997].

The interest in identifying brain substrates that behave like machine-learning constructs was spurred by the results of Schultz [1999, 2001, 2002], who focused on midbrain dopamine neuron firing and appetitive rewards and found no prediction error like signaling of aversive reinforcers in this region. The appetitive aspect of these findings has been tested in humans via fMRI BOLD experiments, and the earlier results suggest that the ventral striatal region shows activity consistent with the prediction error [McClure et al., 2003a; O'Doherty et al., 2003], a finding we confirm here.

TABLE I. Activations correlating with a salience prediction error regressor at the time of US deliverance using SPM99

Region	Peak coordinates	Peak Z	P (extent)
Ventral striatum			
Left	−15, 12, −6	4.86	< 0.001
Anterior insula			
Left	−39, 15, −18	3.86	< 0.01
Right	45, 21, −18	4.32	< 0.01
Medial orbitofrontal cortex			
Bilateral	−3, 48, −12	4.34	< 0.05
Lingual gyrus			
Bilateral	−15, −51, −6	4.44	< 0.001[a]
Anterior cingulate			
Bilateral	−3, 36, 18	3.79	< 0.001[a]

Data are thresholded at *P* < 0.01 (uncorrected) and corrected for cluster size.
[a] No a priori hypothesis in this region.

However, in addition, we find that aversive learning also appears to engage prediction error mechanisms and it does so with the same (rather than opposite) polarity as appetitive rewards. Several reasons can account for these differences. The initial studies by Schultz et al. recorded firing of individual dopamine neurons in the midbrain of monkeys, while our results used BOLD activations in the ventral striatum of humans as the outcome. Previous studies have shown that both appetitive and aversive events are represented in the striatum, in different neurons [Williams et al., 1993], or in the same neurons but with different responses [Ravel et al., 2003; Setlow et al., 2003], which fMRI BOLD would not be able to show. As opposed to unit recordings of single neurons, fMRI BOLD indirectly measures integrated activity of large pools of neurons, leaving many features of the underlying activity unknown, such as which transmitter is involved, type of neuron, etc. [Logothetis, 2003]. Thus, from the pattern of results we observe, we can only conclude that there are regions within the ventral striatum that overlap in appetitive and aversive events and leave open the possibility that there may be nonoverlapping populations or firing patterns that distinguish them.

One must also take into account that the fMRI signal is suggested to arise mainly from postsynaptic processes [Logothetis, 2003] and could thus be a reflection of projections to the region rather than direct firing of dopamine neurons. Given the evidence that dopamine neurons preferentially fire to appetitive events [Mirenowicz and Schultz, 1996; Ungless et al., 2004], one possibility is that dopamine release in response to aversive events may be due to direct presynaptic activation of dopamine [Joseph et al., 2003]. In

any case, the relationship between dopamine neuron firing and dopamine release is not straightforward [Garris et al., 1999], thus limiting any simple linear inferences about dopamine function from BOLD responses.

Our findings are somewhat in contrast to the fMRI findings by Delgado et al. [2000], who showed differential valence coding in the striatum, although more in the dorsal region, which is suggested to be more motor-related. Thus, the more dorsally findings by Delgado et al. [2000] might reflect the active motor task they used while a passive conditioning paradigm was used in the current study. Nonetheless, at least under the conditions of this study (mixed appetitive, aversive, and neutral events; partial reinforcement; and fMRI BOLD as the signal of interest), the ventral striatum shows a pattern of activation that is similar across appetitive and aversive conditions.

If the ventral striatum is mostly involved in the prediction of salience, the coding of valence needs to be done elsewhere. Although using different modalities for reinforcers in this study, the results implicating orbitofrontal cortex in the coding of valence is in accord with previous animal and imaging data [Anderson et al., 2003; Rolls et al., 2003; Small et al., 2003], where this region seems to have an important role in this regard. This structure has been widely suggested to guide behavior based on the anticipated value of different actions [Damasio, 1994]. This is supported by the work by Rolls et al., which suggests that the affective value of a stimulus is represented in this region [Rolls, 2000] and that dissociable regions of the human orbitofrontal cortex correlate with subjective pleasantness and unpleasantness ratings of emotional stimuli [Rolls et al., 2003].

One of the challenges of using biologically aversive stimuli like an electrical shock is that there is no simple positive equivalent of it (cf. money loss vs. money gain). As a result, the intensity of the appetitive and the aversive stimuli in the current study was not experienced as equal based on the GSR data. This is somewhat expected since it is usually easier to get GSR with negative events as compared to positive. While it would be preferable to have equated the intensity in the current study, no significant correlations were obtained between GSR and beta values in the ventral striatum for either appetitive or aversive events (data not reported), suggesting that differences in subjective intensity are unlikely to have confounded these findings.

One surprising finding in this study was the lack of amygdala recruitment. Anderson et al. [2003] reported amygdala activation to be associated to the intensity of stimulus, which led us to use it as a region of interest. The amygdala seems to be activated mainly during learning phases in aversive conditioning paradigms [Buchel et al., 1998, 1999; LaBar et al., 1998], i.e., during acquisition and extinction phases of the association. For example, the studies by Buchel et al. used a function of time as a regressor that resulted in a significant time × condition interaction for amygdala. No extinction phase was used in the current study and since we modeled for appetitive and neutral events beside the aversive ones, it might be that amygdala activation was too weak to reach statistical significance.

To conclude, we show that during classical conditioning, the ventral striatum BOLD responds consistent with the concept of reward prediction in TD learning models, but does so symmetrically for both appetitive and aversive stimuli, i.e., all salient events in the environment. Given its location at the interface of limbic and motor regions [Mogenson et al., 1993; Haber et al., 2000], the ventral striatum may be critical in learning about motivationally salient stimuli and using that information to bias selection of actions [Berridge and Robinson, 1998; McClure et al., 2003b].

REFERENCES

Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N (2003): Dissociated neural representations of intensity and valence in human olfaction. Nat Neurosci 6:196–202.

Berns GS, McClure SM, Pagnoni G, Montague PR (2001): Predictability modulates human brain response to reward. J Neurosci 21:2793–2798.

Berridge KC, Robinson TE (1998): What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28:309–369.

Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001): Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron 30:619–639.

Buchel C, Morris J, Dolan RJ, Friston KJ (1998): Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20:947–957.

Buchel C, Dolan RJ, Armony JL, Friston KJ (1999): Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. J Neurosci 19:10869–10876.

Damasio AR (1994): Descarte's Error. New York: Grosset/Putnam.

Delgado MR, Nystrom LE, Fissel C, Noll DC, Fiez JA (2000): Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 84:3072–3077.

Elliott R, Friston KJ, Dolan RJ (2000): Dissociable neural responses in human reward systems. J Neurosci 20:6159–6165.

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM (1993): 3D statistical neuroanatomical models from 305 MRI volumes. Proc Inst Electric Electron Engineer Nucl Sci Symp Med Imaging 3:1813–1817.

Friston KJ, Ashburner J, Frith CD, Poline JB, Heather J, Frackowiak RS (1995a): Spatial registration and normalization of images. Hum Brain Mapp 2:1–25.

Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RS (1995b): Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189–210.

Garris PA, Kilpatrick M, Bunin MA, Michael D, Walker QD, Wightman RM (1999): Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature 398:67–69.

Gottfried JA, O'Doherty J, Dolan RJ (2002): Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. J Neurosci 22:10829–10837.

Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995): The orbital and medial prefrontal circuit through the primate basal ganglia. J Neurosci 15(7 Pt 1):4851–4867.

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.

Jensen J, McIntosh AR, Crawley AP, Mikulis DJ, Remington G, Kapur S (2003): Direct activation of the ventral striatum in anticipation of aversive stimuli. Neuron 40:1251–1257.

Joseph MH, Datla K, Young AM (2003): The interpretation of the measurement of nucleus accumbens dopamine by in vivo dialysis: the kick, the craving or the cognition? Neurosci Biobehav Rev 27:527–541.

Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001): Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12:3683–3687.

LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA (1998): Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron 20:937–945.

Logothetis NK (2003): The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23:3963–3971.

McClure SM, Berns GS, Montague PR (2003a): Temporal prediction errors in a passive learning task activate human striatum. Neuron 38:339–346.

McClure SM, Daw ND, Montague PR (2003b): A computational substrate for incentive salience. Trends Neurosci 26:423–428.

Mesulam MM (2000): Behavioral neuroanatomy: large-scale networks, association cortex, frontal syndromes, the limbic system, and hemispheric specializations. In: Mesulam M-M, editor. Principles of Behavioral and Cognitive Neurology, 2nd ed. New York: Oxford University Press. p 1–120.

Mirenowicz J, Schultz W (1996): Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature 379:449–451.

Mogenson GJ, Brudzynski SM, Wu M, Yang CR, Yim CCY (1993): From motivation to action: a review of dopaminergic regulation of limbic, nucleus accumbens, ventral pallidum, pedunculopontine nucleus circuittries involved in limbic-motor integration. In: Kalivas PW, Barnes CD, editors. Limbic Motor Circuits and Neuropsychiatry. Boca Raton: CRS Press. p 193–236.

Montague PR, Dayan P, Sejnowski TJ (1996): A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947.

Montague PR, Hyman SE, Cohen JD (2004): Computational roles for dopamine in behavioural control. Nature 431:760–767.

O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002): Neural responses during anticipation of a primary taste reward. Neuron 33:815–826.

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.

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.

Ravel S, Legallet E, Apicella P (2003): Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. J Neurosci 23:8489–8497.

Rolls ET (2000): The orbitofrontal cortex and reward. Cereb Cortex 10:284–294.

Rolls ET, Kringelbach ML, de Araujo IE (2003): Different representations of pleasant and unpleasant odours in the human brain. Eur J Neurosci 18:695–703.

Salamone JD, Cousins MS, Snyder BJ (1997): Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. Neurosci Biobehav Rev 21:341–359.

Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M (2003): Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem 10:129–140.

Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. Science 275:1593–1599.

Schultz W (1999): The reward signal of midbrain dopamine neurons. News Physiol Sci 14:249–255.

Schultz W (2001): Reward signaling by dopamine neurons. Neuroscientist 7:293–302.

Schultz W (2002): Getting formal with dopamine and reward. Neuron 36:241–263.

Setlow B, Schoenbaum G, Gallagher M (2003): Neural encoding in ventral striatum during olfactory discrimination learning. Neuron 38:625–636.

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.

Small DM, Gregory MD, Mak YE, Gitelman DR, Mesulam M-M, Parrish TB (2003): Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 39:701–711.

Sutton RS, Barto AG (1990): Time-derivative models of pavlovian reinforcement. In: Gabriel M, Moore J, editors. Learning and Computational Neuroscience: Foundations of Adaptive Networks. Cambridge, MA: MIT Press. p 497–537.

Ungless MA, Magill PJ, Bolam JP (2004): Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science 303:2040–2042.

Williams GV, Rolls ET, Leonard CM, Stern C (1993): Neuronal responses in the ventral striatum of the behaving macaque. Behav Brain Res 55:243–252.

Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004): Human striatal responses to monetary reward depend on saliency. Neuron 42:509–517.