Behavioral/Systems/Cognitive

Testing the Reward Prediction Error Hypothesis with an Axiomatic Model

Robb B. Rutledge,[1] **Mark Dean,**[2] **Andrew Caplin,**[2] **and Paul W. Glimcher**[1,2]

[1]Center for Neural Science and [2]Department of Economics, New York University, New York, New York 10003

Neuroimaging studies typically identify neural activity correlated with the predictions of highly parameterized models, like the many reward prediction error (RPE) models used to study reinforcement learning. Identified brain areas might encode RPEs or, alternatively, only have activity correlated with RPE model predictions. Here, we use an alternate axiomatic approach rooted in economic theory to formally test the entire class of RPE models on neural data. We show that measurements of human neural activity from the striatum, medial prefrontal cortex, amygdala, and posterior cingulate cortex satisfy necessary and sufficient conditions for the entire class of RPE models. However, activity measured from the anterior insula falsifies the axiomatic model, and therefore no RPE model can account for measured activity. Further analysis suggests the anterior insula might instead encode something related to the salience of an outcome. As cognitive neuroscience matures and models proliferate, formal approaches of this kind that assess entire model classes rather than specific model exemplars may take on increased significance.

Introduction

Our understanding of the natural world progresses through the development of explanatory models designed to capture compact descriptions of physical events. Within neuroscience, these explanatory models tend to develop through a process of competitive evolution in which highly specified models are tested against one another. Other disciplines, including physics and economics, often employ an alternative approach, dividing the space of all possible models into subdomains and then attempting to falsify the hypothesis that one or more members of an entire class of models can account for a set of empirical observations. These model classes are typically defined by sets of testable rules called axioms. Popper (1959) argued that the most powerful test of any theory derives from formal efforts aimed at falsification. In this tradition, the axiomatic approach attempts to falsify entire model classes.

Dopamine neurons are thought to encode a reward prediction error (RPE) signal, the difference between experienced and predicted rewards. Numerous studies have fit specific parameterized RPE models to measurements of dopamine neuron activity (Schultz et al., 1997; Hollerman and Schultz, 1998; Nakahara et al., 2004; Bayer and Glimcher, 2005; Joshua et al., 2008; Matsumoto and Hikosaka, 2009) and functional magnetic resonance imaging (fMRI) measurements of neural activity in dopamine target areas (McClure et al., 2003; O'Doherty et al., 2003, 2004; Seymour et al., 2004; Abler et al., 2006; Li et al., 2006; Pessiglione et al., 2006; Behrens et al.,

2008; D'Ardenne et al., 2008; Hare et al., 2008). Model competitions have shown that parameterized temporal difference approaches (Sutton and Barto, 1990) account for electrophysiological data better (Schultz et al., 1997) than RPE models related to the approach of Rescorla and Wagner (1972). Unfortunately, comparing correlation coefficients for different RPE models cannot tell us whether key features of dopamine-related activity are fundamentally incompatible with specific critical features of the entire class of possible RPE models. The regression approach cannot, in principle, falsify the hypothesis that dopamine-related activity encodes some kind of RPE signal and therefore cannot formally test this hypothesis.

Caplin and Dean recently examined the necessary and sufficient properties of any RPE signal (Caplin and Dean, 2008a,b), finding that any such signal must possess three critical features. They showed that if any one of these features is absent, the observed signal cannot represent an RPE signal regardless of whether it is correlated with specific parameterized RPE models. If all of these features are present, then the measured signal meets criteria of both necessity and sufficiency for representing an RPE signal. By empirically testing these formal mathematical axioms, it is possible to test the entire class of RPE models for a neural signal measured from any brain area.

To axiomatically test the hypothesis that specific neural signals can encode RPEs, we used fMRI to measure blood oxygen level-dependent (BOLD) activity as subjects played monetary lotteries for real money. We asked whether BOLD responses in specific candidate RPE areas satisfied the necessary and sufficient criteria for encoding RPEs. Any neural signal falsifying one or more axioms cannot in principle encode any type of RPE signal. Such a signal cannot be accounted for by any model in the RPE model class. We also tested whether any candidate RPE area might alternatively encode the absolute value of the RPE signal, a quantity related to saliency.

Materials and Methods

Subjects. Fourteen paid volunteers participated in the experiment (nine female, all right-handed, mean age = 26.0 years). All subjects participated in two scanning sessions. Two subjects were excluded from further analysis due to excessive head motion during the scanning sessions. Participants gave informed consent in accordance with the procedures of the University Committee on Activities Involving Human Subjects of New York University (New York, NY).

Experimental task. Before scanning, subjects were endowed with $100 in cash. Subjects also received a show-up fee of $35 at the end of each scanning session, regardless of task earnings. On each trial, subjects chose between two monetary lotteries, an "observation lottery" and a "decoy lottery," where the probability of each prize was represented by the area of that prize's slice (Fig. 1A). To test the axiomatic model, it was necessary to collect data with two prizes available (+$5, −$5) at a variety of probabilities (0–100% in 25% increments). Thus, the observation lottery set consisted of five lotteries that yielded eight possible outcomes and thus eight possible trial types. To ensure that subjects usually chose from the observation set, the decoy lottery always had a lower mathematical expected value (ranging from $1.25 to $5 lower). The decoy set also included prizes (+$0, −$10) not available in any observation lottery. Subjects were required to choose between an observation lottery and a decoy lottery to ensure that they were actively engaged in the task. After a 12.5 s fixation period, options were presented for 5 s. The fixation cross was extinguished, indicating that the subject had 1.25 s to make their selection by button press. After a 7.5 s delay period, the prize was revealed for 3.75 s as a change in the color of that prize's slice in the chosen lottery. If a subject failed to make a button press in the required time window, the subject lost $10. Out of a total of 3024 trials completed, subjects missed 21 trials and chose the decoy lottery in 28 trials, completing 2975 trials to the observation set.

Imaging. Imaging data were collected with a Siemens Allegra 3 tesla head-only scanner equipped with a head coil from Nova Medical. T2*-weighted images were collected using an echo planar imaging sequence. We collected 23 slices oriented parallel to the anterior commissure-posterior commissure (AC-PC) plane [repetition time (TR) = 1.25 s, echo time (TE) = 30 ms, ascending interleaved order, 3 × 3 × 3 mm, 64 × 64 matrix in a 192 mm field of view (FOV)]. This volume provided coverage of the subcortical, frontal, and midbrain regions of interest while omitting part of the parietal lobe and the crown of the skull in all subjects. Each scan consisted of 396 sequential sets of images. The first four images were discarded to avoid T1 saturation effects. There were 16 trials during each scan. Each trial lasted 30 s (Fig. 1A). Each subject completed 13–16 scans over two sessions, with most subjects (n = 9) completing eight scans in each session. The dataset consisted of 74,844 volumes, with an average of 130 min of functional data per subject. We also collected high-resolution T1-weighted anatomical images using a MP-RAGE pulse sequence (144 sagittal slices, TR = 2.5 s, TE = 3.93 ms, inversion time = 900 ms, flip angle = 8°, 1 × 1 × 1 mm, 256 × 256 matrix in a 256 mm FOV) for coregistration of functional data.

Data analysis. Functional imaging data were analyzed using BrainVoyager QX (Brain Innovation), with additional analyses performed in MATLAB (MathWorks) and STATA (StataCorp). We sinc interpolated functional data in time to adjust for staggered slice acquisition. We cor-
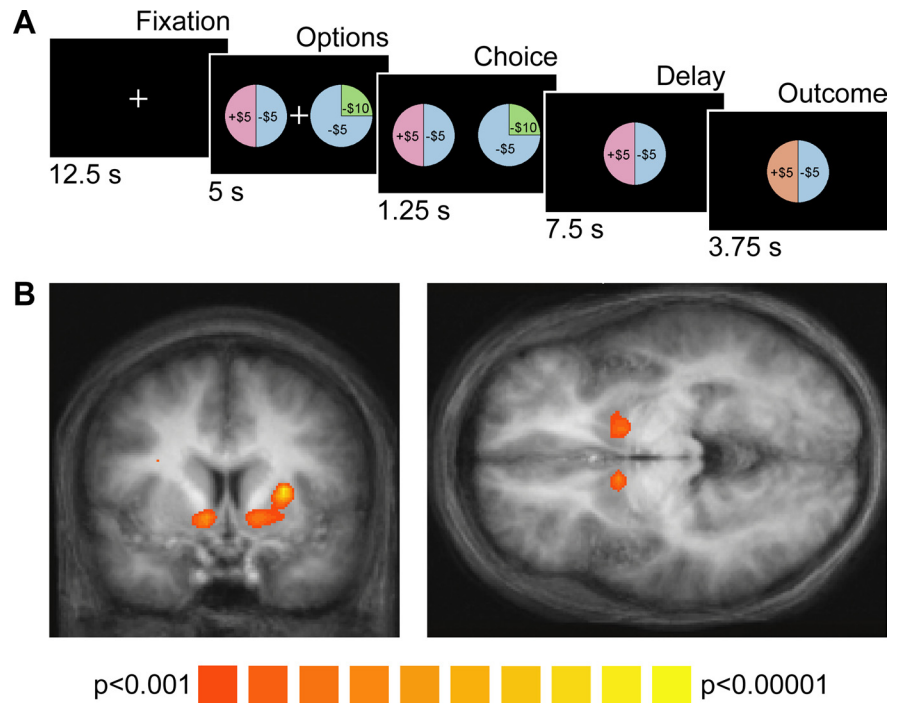
Figure 1. Experimental task and group reward prediction error analysis. **A**, Experimental task design with timing indicated. On each trial, subjects were presented with two options, lotteries with the probability of each prize indicated by the area of the prize's slice. After 5 s, the fixation cross was extinguished and the subject had 1.25 s to indicate their decision by pressing a button. After a delay period, the prize was revealed by a change in the color of the associated slice, here winning $5 from a lottery with a 50% chance of winning $5. **B**, Areas in which neural activity was correlated with the "predicted RPE" in a random-effects group analysis. At a threshold of $p < 0.001$ (uncorrected), areas of correlation were found in the bilateral nucleus accumbens (coronal and axial images at $y = +5$ and $z = −4$, respectively), left putamen (coronal image), and right caudate. Predicted RPE was defined as the mathematical difference in dollars between the prize received and the lottery's expected value. The color scale indicates the t value of the contrast testing for a significant effect of predicted RPE during the outcome period. Data are overlaid on the mean normalized image and shown in radiological convention, with the right hemisphere on the left.

rected for any head movement by realigning all images to the first volume of the session using six-parameter rigid body transformations. We detrended and high-pass filtered (cutoff of three cycles per scan) to remove low-frequency drift in the signal. We then coregistered images to each subject's high-resolution anatomical scan, rotated into the AC-PC plane, and normalized into Talairach space using piecewise affine Talairach grid scaling with trilinear interpolation. Data were spatially and temporally unsmoothed, except for the group random-effects analysis.

To demonstrate the standard regression approach, we performed group random-effects analysis using the summary statistics approach. For this analysis we spatially smoothed all data with an 8 mm full width at half-maximum Gaussian kernel. The regression model consisted of a single regressor of interest encoding the "predicted RPE" on each trial during the outcome period, defined for these purposes as the difference between the reward received in dollars and the expected value of the lottery. Three additional regressors modeled the option onset, button press, and outcome onset for all trials. All four regressors were convolved with the canonical two-gamma hemodynamic impulse response function ($\tau_1 = 6$ s, $\tau_2 = 15$ s, ratio of peak to undershoot = 6). A statistical map was then generated for the regressor of interest using one-sample t tests. This map is shown for demonstration purposes only, without any minimum cluster threshold or corrections for multiple comparisons (Fig. 1B).

For our principal analysis, we independently defined anatomical regions of interest (ROIs) in individual subjects for 11 brain areas: the nucleus accumbens, anterior insula, caudate, putamen, medial prefrontal cortex, amygdala, posterior cingulate cortex, thalamus, ventral tegmental area, substantia nigra, and habenula. These regions were chosen because they have been found to have activity consistent with specific RPE models in previous neuroimaging and neurophysiological studies.

Criteria for these structural definitions, primarily those established by the Center for Morphometric Analysis (Charlestown, MA) (Rademacher et al., 1992; Caviness et al., 1996), are described in the supplemental data, and distributions for these definitions across subjects are shown (see Fig. 3 and supplemental Figs. 1 and 3, available at www.jneurosci.org as supplemental material).

Our ROIs were largely located in subcortical and midbrain areas. The amygdala and posterior cingulate cortex ROIs were located near the boundaries of our acquisition volume, making these ROIs particularly susceptible to artifacts from the motion correction algorithm. To limit these artifacts, we excluded from our ROIs any voxel from a given scan for which the standard deviation of percentage signal change exceeded 2%, a degree of variance incompatible with a continuous BOLD signal (supplemental Table 1, available at www.jneurosci.org as supplemental material). In practice, this excluded on average <5% of voxels for all structures except the amygdala (12%) and posterior cingulate cortex (23%). We further limited the effects of motion on BOLD activity in individual voxels by using a regression model that included the six motion predictor regressors and their temporal derivatives. We then averaged data across each anatomical ROI to produce a mean time course for each ROI that was converted to percentage signal change using two baseline TRs as indicated.

We made no assumptions about the shape of the hemodynamic response functions in our anatomical ROIs, but removed correlations between time points at the subject and trial-type level by using an AR4 autoregressive model while maintaining consistent time point averages. We then averaged activity within the 5-TR window, weighting each time point equally. We computed parameter estimates by ordinary least squares for each of the eight trial types for the 2975 trials in the observation set controlling for subject-level differences in activity. We evaluated the axiomatic RPE model and RPE absolute value model described below, testing for differences between parameter estimates using Wald tests of linear restriction. We tested the robustness of our results by performing these analyses while systematically varying the baseline measurement time and the starting time for the 5-TR analysis window.

Axiomatic RPE model. To determine whether the BOLD signal measured in the striatum and other candidate RPE areas meets the criteria of necessity and sufficiency for encoding an RPE signal, we formally tested the RPE hypothesis using the axiomatic model developed by Caplin and Dean (2008a). This approach makes no specific assumptions about the precise form of subjective variables like "reward" and "expectation," which are not part of the RPE hypothesis but are required to be fully specified by the traditional regression approach. Using this approach, we can thus explicitly test whether a given neural signal falsifies or satisfies the three conditions of necessity and sufficiency for the entire class of RPE models (subject to some technical restrictions) (Caplin and Dean, 2008a,b).

For example, all RPE models assume that an RPE signal responds similarly to any fully anticipated outcome, whether it be winning or losing $5 or winning an apple or an orange, and the axiomatic model's third axiom formally captures that intuition. Surprisingly, this assumption has never been tested on dopamine neurons for prizes with similar sensory properties, like apple juice and orange juice. The critical feature of this approach is that if any neural signal does not satisfy this axiom or either of the other two axioms, then it cannot, in principle, represent an RPE signal. For the two-prize case we tested, the three axioms are necessary and sufficient criteria for the RPE model class (Caplin et al., 2010).

We tested our measurements of neural activity against the three axioms: consistent prize ordering (axiom 1), consistent lottery ordering (axiom 2), and no surprise equivalence (axiom 3). The three axioms are as follows, where $\delta(z, p)$ is neural activity associated with receiving prize z (e.g., winning $5) from lottery p (e.g., 50% probability of winning $5). $\delta(z)$ is the one-prize "lottery," where prize z has 100% probability:

Axiom 1: $\delta(z, p) > \delta(z', p) \Rightarrow \delta(z, p') > \delta(z', p')$;
Axiom 2: $\delta(z, p) > \delta(z, p') \Rightarrow \delta(z', p) > \delta(z', p')$;
Axiom 3: $\delta(z') = \delta(z)$.

Axiom 1: consistent prize ordering. When expectations (e.g., lotteries) are fixed and prizes are varied, any difference in activity between prizes for an RPE signal reflects the different rewards associated with those prizes. Ranking prizes by neural activity captures those differences, and these rankings must be the same for all lotteries.
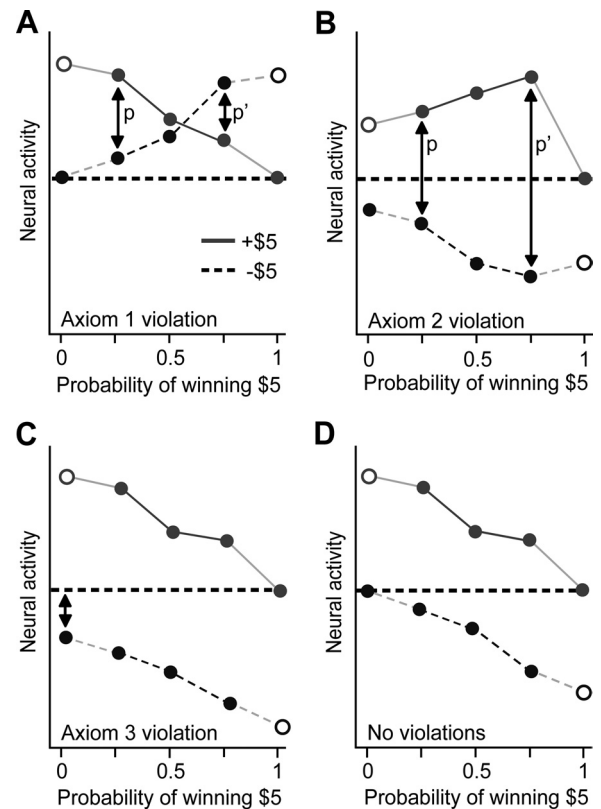
Figure 2. The axiomatic RPE model. Hypothetical neural activity is shown for two prizes (winning and losing $5) received from five lotteries with probabilities of winning from 0 to 100%. Only two prizes are possible, so, for example, the lottery with a 50% probability of winning $5 also has a 50% probability of losing $5. **A**, Example of a violation of axiom 1. **B**, Example of a violation of axiom 2. **C**, Example of a violation of axiom 3. **D**, A pattern of activity with no axiomatic violations.

Consider two different lotteries, p and p', which have 25 and 75% probabilities of winning $5, respectively (Fig. 2A). Caplin and Dean (2008a) demonstrated that if an RPE signal responds with higher activity to winning than losing $5 from lottery p, then it must be the case that the signal also responds with higher activity to winning than losing $5 from lottery p'. Alternatively, if it responds with lower activity for lottery p, then it must also respond with lower activity for lottery p'. Figure 2A shows a hypothetical result that would falsify this first criterion. Hypothetical neural activity (for example, BOLD activity from some brain area) is plotted against the probability of winning $5; each point represents activity associated with receiving a particular prize (winning or losing $5) from one of the five lotteries in the observation set. Open circles represent unobservable outcomes; for example, observing the activity associated with losing $5 when the probability of winning $5 is 100% is impossible. If more activity is associated with more reward, higher activity for winning than for losing $5 from lottery p implies that winning $5 has the higher experienced reward (Fig. 2A). Higher activity for losing than winning $5 from lottery p' implies the opposite, and this contradiction violates the first axiom. Any crossing of the lines thus contradicts consistent prize ordering and proves that the activity under study cannot, in principle, encode any form of RPE signal. This must be true for any two prizes, for example, comparing apples and oranges; if the activity is higher for apples than oranges for one lottery, it must be higher for all lotteries between apples and oranges for any RPE representation.

Axiom 2: consistent lottery ordering. When rewards are fixed and expectations (e.g., lotteries) are varied, any difference in activity between lotteries for an RPE signal reflects the different predicted rewards of those lotteries. Ranking lotteries by neural activity captures those differences, and these rankings must be the same for all prizes (e.g., for both winning and losing $5).

Consider again lotteries p and p', again with 25 and 75% probabilities of winning $5, respectively (Fig. 2$B$). If more activity is associated with more reward, an RPE signal that responds with lower activity to losing $5 from lottery p' than from lottery p implies that p' has higher predicted reward. Therefore, it must also respond with lower activity to winning $5 from lottery p' than from lottery p. Figure 2B shows a violation of this axiom. Lower activity for losing $5 from lottery p' than lottery p implies that lottery p' (the lottery with a 75% chance of winning $5) has the higher predicted reward. Higher activity for winning $5 from lottery p' than from lottery p implies the opposite. The downward slope of the line for losing $5 implies that lotteries with a higher probability of winning $5 have higher predicted reward. The upward slope of the line for winning $5 implies the opposite. For any two lotteries, any difference in the signs of slopes between the lines for the two prizes contradicts consistent lottery ordering and proves that the activity under study cannot, in principle, encode any form of RPE signal.

Axiom 3: no surprise equivalence. The final criterion of necessity and sufficiency identified by Caplin and Dean (2008a) was that RPE signals must respond identically to all fully predicted outcomes, conditions under which the reward prediction error is zero. If there is no reward prediction error, the signal must always generate the same response regardless of the prediction.

Consider the two one-prize "lotteries" shown as the filled endpoints of the lines in Figure 2C. If, as shown in the plot, the signal responds with less activity to losing than winning $5 when both outcomes are fully anticipated, this violates the third axiom and proves that the activity under study cannot, in principle, encode any form of RPE signal.

These three representational constraints must be obeyed by any member of the class of RPE models (Caplin and Dean, 2008a,b; Caplin et al., 2010), whether a Rescorla–Wagner model, a temporal-difference model, an RPE model with a high or low learning rate, or an RPE model with any arbitrary utility function. If an observed neural signal fails to meet any of these criteria, then the proposition that it can encode an RPE signal can be considered formally falsified. In contrast, a neural signal that demonstrates all three properties is one that, in the two-prize case, meets the sufficient criteria for encoding an RPE signal (as proven for the two-prize case by Caplin et al., 2010). A pattern of activity satisfying all three axioms is shown in Figure 2D.

To test the axioms empirically, neural activity estimates are compared to determine whether any axioms are violated. To give an example of how these tests can be performed, consider a situation in which the number of units of neural activity simply equals the difference between the prize and the lottery expected value in dollars. Winning $5 when the probability of winning $5 was only 25% would thus be associated with 7.5 units of activity. To test the first axiom, consider the lottery with a 25% probability of winning $5. The activity is higher for winning than losing $5 from this lottery ($7.5 > -2.5$). The first axiom is satisfied if the activity is also higher for winning than losing $5 from the 50% lottery and the 75% lottery, as it is ($5 > -5$; $2.5 > -7.5$). We see that the activity is lower for losing $5 from the 75% than the 25% lottery ($-7.5 < -2.5$). The second axiom is satisfied if the activity is also lower for winning $5 from the 75% than the 25% lottery, as it is ($2.5 < 7.5$). The two other pairwise comparisons must also be checked. Finally, the third axiom is satisfied because the activity is the same for a fully anticipated win or loss of $5 ($0 = 0$). Such a signal thus obeys all three axioms.

Critically, when we test the axioms we do not make any assumptions about the magnitude of experienced and predicted rewards for prizes or lotteries, nor about the hemodynamic response function of subjects or brain areas. However, since our analysis is performed at the group level, we do assume that, to the degree that measurements from individual subjects provide similar degrees of statistical power, these subjects have the same direction of ordering over prizes and lotteries. For example, we assume that all subjects either prefer winning to losing $5 or alternatively prefer losing to winning $5. Dopamine-related activity is thought to increase with experienced reward and decrease with predicted reward. Although the theory itself does not require this, if we assume that subjects prefer winning to losing $5 and prefer lotteries with a higher probability of winning $5, then we predict that the axioms will be satisfied specifically in the way indicated in the leftmost column in Table 1, which we refer to

Table 1. Axiomatic RPE model statistical tests

| Axiom	NAcc	AI	Caud	Put	MPFC	Am	PCC	Thal
1.1 +	+	=	+	+	+	+	+	+
1.2 +	+	=	+	+	+	+	=	=
1.3 +	+	−	+	+	+	+	+	=
2.1 −	−	=	−	−	−	=	=	=
2.2 −	−	+	−	−	+	−	+	=
2.3 −	−	−	−	−	−	−	−	−
2.4 −	−	−	−	=	−	=	−	−
2.5 −	−	−	−	−	−	−	−	−
2.6 −	−	=	−	−	−	−	−	−
3 =	=	−	=	−	=	=	=	−

Testing the three axioms of the axiomatic RPE model on our data requires 10 statistical tests. Wald tests of linear restriction were performed on parameter estimates computed with a baseline of TR 9–10 and an analysis window of TR 26–30 (parameter estimates for the nucleus accumbens and anterior insula are shown in Figure 4) with the sign of all significant tests indicated ($p < 0.05$). We predicted that RPE signals would satisfy the axioms in the way indicated by the signs in the leftmost column. At $p < 0.05$, the nucleus accumbens and caudate each satisfy all three axioms. The anterior insula and thalamus falsify all three axioms. The amygdala and medial prefrontal cortex each satisfy two axioms and the putamen and posterior cingulate cortex each satisfy one axiom. Axiomatic statistical test 1.1: $\{+\$5, 25\% \text{ probability of winning } \$5\} - \{-\$5, 25\%\}$; 1.2: $\{+\$5, 50\%\} - \{-\$5, 50\%\}$; 1.3: $\{+\$5, 75\%\} - \{-\$5, 75\%\}$; 2.1: $\{+\$5, 50\%\} - \{+\$5, 25\%\}$; 2.2: $\{-\$5, 50\%\} - \{-\$5, 25\%\}$; 2.3: $\{+\$5, 75\%\} - \{+\$5, 50\%\}$; 2.4: $\{-\$5, 75\%\} - \{-\$5, 50\%\}$; 2.5: $\{+\$5, 75\%\} - \{+\$5, 25\%\}$; 2.6: $\{-\$5, 75\%\} - \{-\$5, 25\%\}$; 3: $\{+\$5, 100\%\} - \{-\$5, 0\%\}$. NAcc, Nucleus accumbens; AI, anterior insula; Caud, caudate; Put, putamen; MPFC, medial prefrontal cortex; Am, amygdala; PCC, posterior cingulate cortex; Thal, thalamus.

as "strongly satisfying" the axioms. They could also be satisfied in many other ways, including if all the signs in the leftmost column were reversed. Such a pattern might in fact be predicted for the habenula. In contrast to dopamine neurons, electrophysiological data suggests that habenula neuron activity decreases with experienced reward and increases with predicted reward (Matsumoto and Hikosaka, 2007).

Finally, we refer to the axioms as being "weakly satisfied" if all pairwise comparisons are consistent with the axioms, but not all tests have the predicted signs. The most trivial example of this kind of weak satisfaction would be a pure noise signal for which all pairwise comparisons yielded equal values. If all pairwise comparisons were identical (and hence all lines in a plot like in Fig. 2 were horizontal and overlapping), no axiom would be falsified, but the axioms would only be weakly satisfied. All comparisons would also be equal and, hence, the axioms only weakly satisfied for an RPE signal if the subject is indifferent between the two prizes, for example, if both prizes are winning $5. This is another degenerate case of weak satisfaction that would be of no interest.

We looked for signals that strongly satisfied the axioms by counting the number of tests with the predicted sign at $p < 0.05$ for a wide range of baselines and analysis windows. Baselines were selected around the end of the fixation period and the end of the delay period. BOLD activity in all areas was observed to be relatively similar across trial types during these periods. A range of analysis windows was tested starting before the outcome period and lasting into the next trial. The axioms are thus satisfied in a meaningful way for dopamine-related activity (in the strong sense) if and only if all of our statistical tests have the predicted sign. Because this requires the conjunction of 10 statistical tests, the probability of the axioms being satisfied with the predicted signs at $p < 0.05$ for a given signal for all tests is approximately one in a billion (for seven tests, excluding tests 2.5 and 2.6, which are not independent of the other tests). This fact renders the spurious identification of an RPE representation that strongly satisfies the axioms highly unlikely. For each signal with all of the predicted signs, we thus used the largest p value for seven tests for the first and second axioms to compute a conjunction p value and corrected for multiple comparisons (baselines, analysis windows, and ROIs). For the habenula, we also tested the axiomatic RPE model with the signs of all statistical tests reversed to search for a sign-reversed RPE signal (Matsumoto and Hikosaka, 2007).

RPE absolute value model. To test whether a signal can represent the magnitude of the RPE signal for the two-prize case we tested, we must make two assumptions about how the RPE is constructed. First, we assume that the RPE is the mathematical difference between the experienced and predicted reward. Second, we assume (for the two-prize case) that the predicted reward is equal to $p_z u_z + (1 - p_z)u'_z$ where p_z, u_z, and
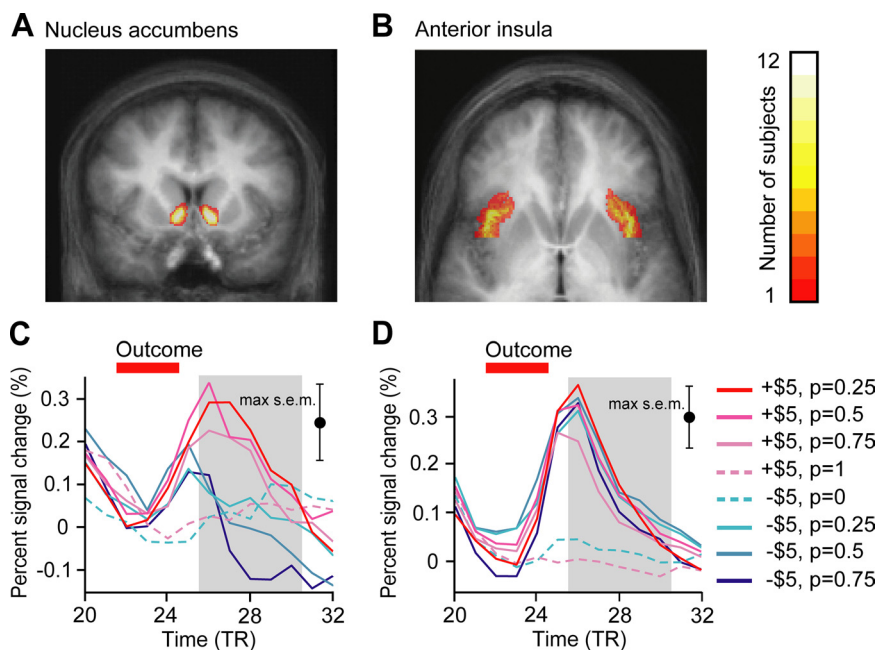
Figure 3. BOLD responses in the nucleus accumbens and anterior insula. **A, B,** ROIs were defined in individual subjects by anatomical criteria for the nucleus accumbens (coronal image) (**A**) and anterior insula (axial image) (**B**). The color scale indicates the number of subjects containing a particular voxel in the individual ROI definitions. Data are overlaid on the mean normalized image and shown in radiological convention, with the right hemisphere on the left. **C, D,** Data were averaged across all voxels in the individual anatomical ROIs and replotted as trial averages. Trial averages are color-coded by predicted RPE for each of the eight trial types. The outcome period (TR 22–24) is indicated. The window (TR 26–30) for which the axioms were tested is shown in gray. The largest SEM for any time point for any trial type is shown on the right. Anatomical ROIs and trial averages for additional areas are shown in supplemental Figures 1 and 3.

u'_z are the probability and utility of prize z and the utility of prize z', respectively. Thus, the RPE absolute value (abs) when prize z is received is $(1 - p_z)\text{abs}(u_z - u'_z)$. When prize z' is received, it is $(1 - p'_z)\text{abs}(u_z - u'_z)$. Since the second term is always the same for the two-prize case, the RPE absolute value should be a decreasing function of probability. We test whether activity decreases with prize probability with the following condition: the activity associated with receiving prize z from lottery p is higher than that for receiving prize z' from lottery p' if and only if the probability of receiving prize z from lottery p is less than the probability of receiving prize z' from lottery p'.

In this way we examine the possibility that how surprising an outcome is, a property related to salience, can be encoded by the BOLD response in a particular brain area. For the habenula we also tested for a sign-reversed RPE absolute value signal.

Results

Traditional regression-based analysis

A number of previous studies have examined the RPE hypothesis by selecting a fully parameterized member of the RPE model class and correlating some element of the model with measured BOLD activity. We first completed a standard random-effects regression analysis of this type (Fig. 1*B*) to allow comparison with the results of our axiomatic RPE model analysis. To accomplish this, we had to make several assumptions about the structure of the neural representation of concepts like "reward" and "expectation," variables which the regression model would not measure directly. We therefore assumed, as have previous studies (D'Ardenne et al., 2008), that reward, or more formally the utility function for gains and losses, is a linear function of monetary reward with no change in slope at the origin (the mathematical representation of reward proposed by Pascal). We also assume that the predicted reward is equal to the utilities of the prizes weighted by their objective probabilities (an assumption equivalent to the independence ax-

iom in expected utility theory). If these assumptions are correct, they require that the RPE signal be proportional to the difference in dollars between the outcome received and the lottery's expected value. We also assumed, as have previous studies (e.g., Li et al., 2006), that the BOLD response in all areas would follow the canonical two-gamma hemodynamic impulse response function, which has been well validated in sensory and motor cortex (Friston et al., 1998; Vazquez and Noll, 1998). With these parameterizations in hand, we found that BOLD activity in the striatum (including parts of the nucleus accumbens, putamen, and caudate) was significantly correlated ($p < 0.001$, uncorrected) with the RPE term specified in this model. This result is consistent with numerous previous studies (e.g., McClure et al., 2003; O'Doherty et al., 2003, 2004; Pessiglione et al., 2006). At a more liberal threshold ($p < 0.01$), BOLD activity in the medial prefrontal cortex was also correlated with the RPE term in this model, but not activity in other candidate RPE areas, including the anterior insula, amygdala, and posterior cingulate cortex. While these data clearly indicate that BOLD activity in the striatum is correlated with the predictions of this particular RPE model, they cannot tell us whether the data are actually compatible with the RPE hypothesis. This is because we cannot determine whether the limits to the observed correlation derive from a fundamental and insurmountable mismatch between critical properties of the signal and the model or simply from limitations in the accuracy of our measurements. To address this issue, we turned next to a test of the necessary and sufficient signal properties required for any RPE representation.

Testing the RPE hypothesis by the axiomatic method

Neuroimaging studies have identified activity in numerous brain areas that is correlated with the predictions of particular RPE models. To test the hypothesis that BOLD activity in these brain areas can, in principle, be precisely described by at least one member of the RPE model class, we first anatomically defined ROIs and then computed estimates of the average BOLD activity for each of the eight trial types from the set of observation lotteries. This allowed us to produce plots of the kind shown in Figure 2 for each brain area. We then performed statistical tests on these data in an effort to falsify one or more of the axioms of the RPE hypothesis.

We first extracted BOLD responses in all subjects from the nucleus accumbens and the anterior insula (Fig. 3*A,B*); both regions were identified as possible RPE areas in previous studies (e.g., Pessiglione et al., 2006; Voon et al., 2010). We then plotted the average BOLD responses for the eight trial types, all possible outcomes from the five observation lotteries (Fig. 3*C,D*), converted to percentage signal change relative to a baseline selected as the last two TRs of the fixation period (TR 9–10). The outcome of each trial was presented on the screen for three TRs (TR 22–24). Due to the lag in the hemodynamic response (∼5 s or four TRs), we specified our initial analysis window as TR 26–30 (anal-
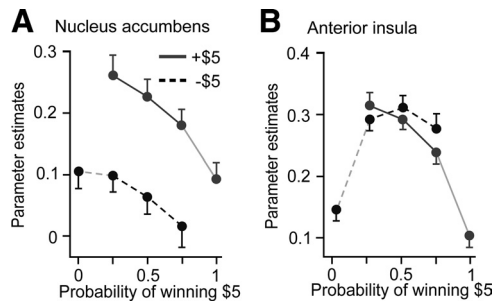
Figure 4. Testing the axiomatic RPE model. **A, B,** Parameter estimates and 95% confidence intervals are plotted for each trial type for the two prizes (winning and losing $5) against the probability of winning $5. The data from the nucleus accumbens satisfies all three axioms at $p < 0.05$ (**A**). The data from the anterior insula falsifies all three axioms at $p < 0.05$ (**B**). Test results are shown in Table 1. Results for additional areas are shown in supplemental Figure 2 (available at www.jneurosci.org as supplemental material).

yses presented below relax this assumption). For each brain area, we then estimated parameters for each of the eight trial types, averaging activity across the analysis window, weighting time points equally. Our methodology necessarily assumes a degree of consistency across subjects in the relationship between neural activity and reward. For example, we assume that the sign of the relationship between neural activity and reward is the same for all subjects. However, we make no assumptions about the shape of the hemodynamic response function in different brain areas. The resulting parameter estimates are plotted for the nucleus accumbens (Fig. 4A) and anterior insula (Fig. 4B). For each area, we then performed 10 Wald tests of linear restriction (Wald, 1943) on the relations between these parameter estimates that instantiate the three critical axiomatic criteria. Test results are shown in Table 1. Although testing the axioms requires no assumptions about the precise ordering of prizes or lotteries, we predicted that subjects would prefer winning to losing $5 and lotteries with a higher probability of winning $5 and that BOLD activity would be related to this preference. This led us to predict that the axioms would be satisfied in the manner specified in the leftmost column in Table 1.

For BOLD activity in the nucleus accumbens (Fig. 4A), a subregion of the ventral striatum, axiom 1 is satisfied with higher activity for winning than losing $5 for the three two-prize lotteries (all $p < 0.001$). Axiom 2 is satisfied with all lines significantly downward sloping (all $p < 0.05$). Finally, axiom 3 is satisfied with activity not significantly different for the two fully anticipated outcomes ($p = 0.29$). This signal thus satisfies all three necessary and sufficient conditions of the axiomatic RPE model and can unambiguously encode an RPE signal. All conditions were satisfied in exactly the manner predicted, and the conjunction p value of all the tests corrected for multiple comparisons for the nucleus accumbens is $p < 0.000005$.

Perhaps surprisingly, the data for the anterior insula indicate a very different conclusion (Fig. 4B). Axiom 1 is falsified at $p < 0.05$; the activity is higher for losing than winning $5 from the 75% lottery ($p < 0.001$), but this is not true for the other lotteries at $p < 0.05$. Axiom 2 is also falsified at $p < 0.05$ in two different ways: activity is higher for losing $5 from the 50% than the 25% lottery ($p = 0.032$), but this is not true for winning $5, and activity is lower for winning $5 from the 75% than the 25% lottery ($p < 0.001$), but this is not true for losing $5. Finally, axiom 3 is also falsified; the activity is significantly higher for losing $5 than winning $5 for the fully anticipated outcomes ($p < 0.001$). Therefore, this signal falsifies all three necessary and

sufficient conditions of the axiomatic RPE model, only one of which is required to falsify the hypothesis. Thus, we can conclude that this signal cannot possibly encode any type of RPE signal. BOLD activity in the anterior insula, despite the fact that correlations with the predictions of specific RPE models have been observed in some studies, cannot in principle encode an RPE signal under the conditions we examined.

We also tested several other areas that previous studies suggest might encode RPE signals. Anatomical definitions and BOLD time series for six additional areas are shown in supplemental Figure 1 (available at www.jneurosci.org as supplemental material), with tests presented in Table 1. BOLD activity in the caudate also satisfies all three axioms at $p < 0.05$ and can encode an RPE signal with a corrected conjunction $p < 0.000005$. Activity in the putamen, medial prefrontal cortex, and amygdala, but not the posterior cingulate cortex, satisfies the first axiom at $p < 0.05$. However, activity in all four areas falsifies the second axiom at $p < 0.05$, so these signals cannot, in principle, represent an RPE if the representation is to arise at the time of our analysis window relative to this specific baseline. The signal in the thalamus also falsifies all three axioms and cannot encode an RPE signal under these conditions. However, because measurements of BOLD activity are noisy, whether or not a signal satisfies the axioms might depend on the baseline and analysis window used to estimate the responses; an RPE model might well describe activity at other times or against other baselines. To test this possibility and to examine the robustness of our findings, we analyzed signals for a wide range of baselines and analysis windows.

Assessing the robustness of axiomatic RPE model tests
In the preceding section, to test the axiomatic RPE model we averaged the signal across a 5-TR analysis window (TR 26–30) beginning around the expected peak of the hemodynamic response. We converted the raw signal to percentage signal change using the last two TRs of the fixation period as a baseline. This standard practice in fMRI time series analysis adjusts for magnetic field drift that detrending and high-pass filtering fail to correct. To assess the robustness of our results, we counted the number of tests that were significant at $p < 0.05$ with the predicted sign for a range of baselines and analysis windows. We plot the results of this analysis in Figure 5, with results for 11 possible baselines (including no baseline) plotted against the starting time of the 5-TR analysis window. Color indicates the number of significant tests with the predicted sign for that particular baseline and analysis window, with the number of significant tests indicated by color.

For certain baselines and analysis windows, BOLD activity in the nucleus accumbens, caudate, amygdala, and posterior cingulate cortex (Fig. 5) had the predicted sign for all 10 tests of the axiomatic model, all with a corrected conjunction $p < 0.000005$. Swaths of red in Figure 5 indicate that most tests had the predicted sign for a range of baselines and analysis windows, suggesting that the axiomatic RPE model is robustly appropriate for these areas. Signals measured from each of these areas satisfy the axioms in exactly the way predicted, and thus these signals can encode RPEs (supplemental Fig. 2, available at www.jneurosci. org as supplemental material). Although BOLD activity in the putamen and medial prefrontal cortex did not have the predicted sign for all 10 tests for any baseline or analysis window (Fig. 5D,E), there are signals for both areas that satisfy all three axioms at $p < 0.05$ (supplemental Fig. 2, available at www.jneurosci.org as supplemental material). For example, for a baseline TR 8–9 and analysis window TR 28–32, the medial prefrontal cortex
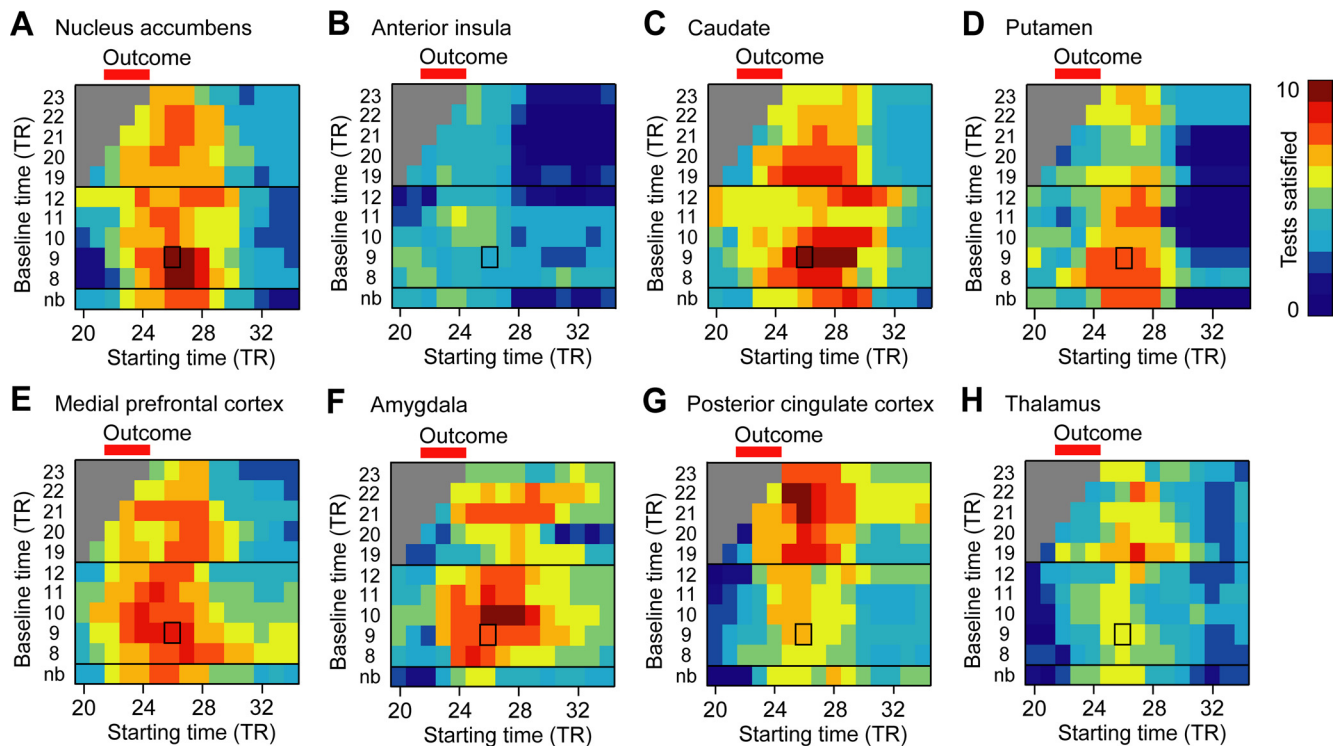
Figure 5. Assessing the robustness of axiomatic RPE model analyses. ***A–H***, Heat maps show results of the axiomatic analysis for a variety of baselines and starting times for the 5-TR analysis window. Testing the axiomatic model across areas requires 10 statistical tests. The first TR of the baseline is indicated for each 2-TR baseline. The color scale indicates the number of tests with the predicted sign at $p < 0.05$. The baseline and analysis windows used for the analyses in Figure 4 and Table 1 is indicated by rectangles. All ROIs are defined by anatomical criteria in individuals. Neural activity in the nucleus accumbens (***A***), caudate (***C***), putamen (***D***), medial prefrontal cortex (***E***), amygdala (***F***), and posterior cingulate cortex (***G***) has the predicted result for the majority of tests for a variety of baseline and analysis windows. Neural activity in the anterior insula (***B***) and thalamus (***H***) does not have the predicted sign regardless of the choice of baseline and analysis window. nb, No baseline. Dopaminergic midbrain and habenula results are shown in supplemental Figure 4 (available at www.jneurosci. org as supplemental material).

weakly satisfies all three axioms, and all tests have the predicted sign except tests 2.1 and 2.2. The signs of these two tests are both equal (rather than minus) and therefore weakly satisfy the second axiom. The putamen signal for a baseline TR 22–23 and analysis window TR 25–29, for example, also weakly satisfies all three axioms. The putamen also strongly satisfies the axioms at $p < 0.10$ with a corrected conjunction $p < 0.0005$. For all of these RPE areas, the majority of tests have the predicted sign for a wide range of baselines and analysis windows.

In contrast, the anterior insula does not appear to satisfy the criteria for an RPE representation for any baseline or analysis window. There is only a single baseline and analysis window over the entire range tested for which this area (Fig. 5*B*) even has the predicted sign for the majority (six) of the tests. There exists no baseline and no analysis window within the range of TR 22–36 for which all three axioms are satisfied for either the anterior insula or the thalamus; the signal from both areas cannot possibly encode an RPE representation under the conditions we examined, and this result is robust to choice of baseline and analysis window.

Our traditional random-effects regression analysis revealed correlations with our particular RPE model only in the striatum (Fig. 1*B*) and, at a more liberal threshold ($p < 0.01$, uncorrected), the medial prefrontal cortex. We were therefore surprised to see several other brain areas from which signals satisfied the axiomatic RPE model, some of which (amygdala and posterior cingulate cortex) are rarely identified in neuroimaging studies of the RPE hypothesis. Plotting the average BOLD response to positive and negative outcomes from the three

two-prize lotteries (the lotteries for which the outcomes were uncertain) reveals that the hemodynamic responses to outcomes in the amygdala and posterior cingulate cortex, and also the medial prefrontal cortex, appear to bear little similarity to the canonical hemodynamic response function (Fig. 6). For example, all three signals terminate at a higher level than they started. This may suggest that prior regression-based analyses have failed to identify some RPE signals due to incorrect hemodynamic assumptions. Even in the nucleus accumbens, the hemodynamic prediction appears to fit the data poorly, with the signal rising initially for all outcomes and then dipping well below the starting level. As shown here, our analysis methods circumvent these issues.

Although imaging the dopaminergic midbrain structures is notoriously difficult and few studies have reported success at identifying possible RPE signals in the midbrain (but see D'Ardenne et al., 2008), we tested whether the BOLD responses we extracted at 3 tesla at standard resolution from the ventral tegmental area and substantia nigra might satisfy the axiomatic RPE model. We found no evidence of RPE signals in BOLD responses in either area (supplemental Fig. 4, available at www. jneurosci.org as supplemental material). These data cannot address whether spiking patterns in these areas are consistent with the RPE theory or whether other fMRI measurement techniques might reveal such a signal. We also tested whether BOLD responses in the habenula might encode an RPE signal or alternatively a sign-reversed RPE signal, as a recent electrophysiological study has suggested is carried by spiking activity (Matsumoto and Hikosaka, 2007). We found no evidence for either an RPE or a sign-reversed RPE BOLD signal in these data (supplemental Fig.
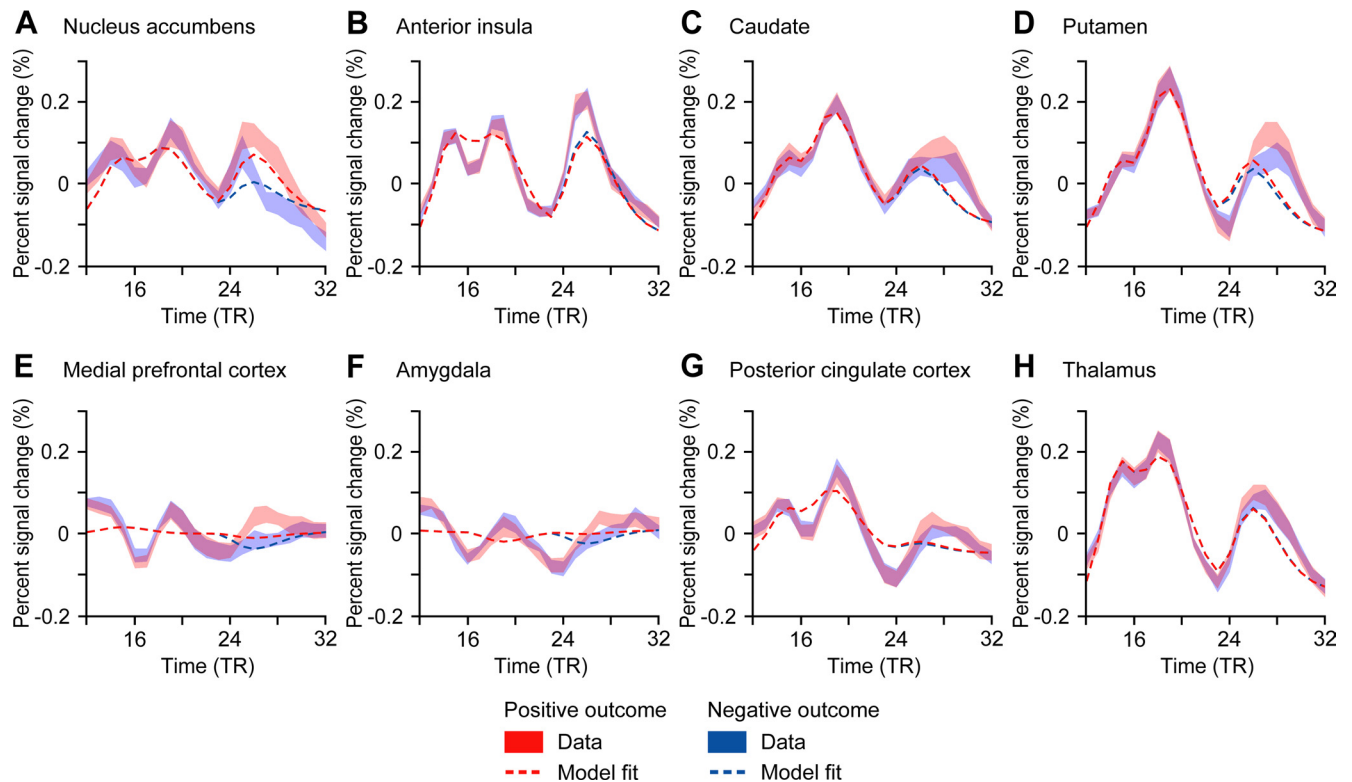
Figure 6. BOLD responses to positive and negative outcomes. *A–H,* BOLD responses for positive (red) and negative (blue) outcomes are plotted, averaged across subjects. Results are for the three two-prize lotteries. Error bars reflect ± SEM across subjects. Dotted lines represent best fits for a typical regression model with regressors for option onset, choice, and outcome onset, convolved with the canonical two-gamma hemodynamic impulse response function with fits averaged across subjects.

4, available at www.jneurosci.org as supplemental material). Supplemental Figure 3 (available at www.jneurosci.org as supplemental material) displays ROI definitions and trial averages for all three areas.

Understanding the anterior insula: RPE absolute value signals
Given previous reports indicating that BOLD activity in the anterior insula is often correlated with the predictions of specific RPE models and our finding that anterior insula activity cannot serve as an RPE signal, we examined whether the signal in the anterior insula might encode some other reward-related information. One possibility is that the signal encodes something about how surprising or salient an outcome is to a subject. Although there is little formal agreement regarding the definition of the term "salience," one natural assumption would be that an outcome is more salient if it is less likely. In our experimental setting, a greater response to an outcome with lower probability is equivalent to encoding the absolute value of the RPE signal if we assume that subjects form their expectations by linearly combining the utilities of prizes weighted by their probabilities. An RPE absolute value ("salience") model has a testable restriction that it places on our dataset. Activity associated with receiving prize z from lottery p is higher than activity for receiving prize z' from lottery p' if and only if the probability of receiving prize z from lottery p is less than the probability of receiving prize z' from lottery p'. Testing this restriction requires evaluating all 28 pairwise comparisons between outcomes.

We replotted the parameter estimates against the probability of the prize received for the nucleus accumbens and anterior insula for baseline TR 22–23 and analysis window TR 24–28 (Fig. 7). For this baseline and analysis window, we found that BOLD
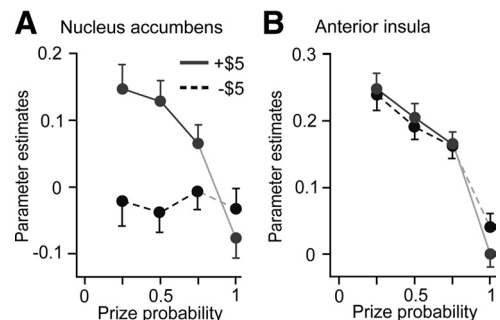
Figure 7. Testing the RPE absolute value model. *A, B,* Parameter estimates and 95% confidence intervals are plotted for the two prizes (winning and losing \$5) against the probability of receiving that prize for the anatomical ROIs shown in Figure 3. The baseline is TR 22–23 and the analysis window TR 24–28. Neural activity in the nucleus accumbens does not appear to be a decreasing function of prize probability (*A*). Neural activity in the anterior insula is a largely decreasing function of prize probability (*B*), consistent with encoding the absolute value of the RPE signal, a quantity related to salience.

activity in the anterior insula was largely a decreasing function of prize probability, as would be predicted for a salience signal (Fig. 7B). This was not the case for the nucleus accumbens (Fig. 7A). We evaluated the 28 tests of the RPE absolute value model for eight brain areas (Table 2) and found that in the anterior insula, 27 of 28 tests had the predicted sign at $p < 0.05$. In the nucleus accumbens only 15 of 28 tests had the predicted sign at $p < 0.05$.

We conducted our tests of the RPE absolute value model on neural signals estimated with a range of baseline and analysis windows, as we did for the axiomatic RPE model. For each baseline and analysis window, we counted how many of the 28 statistical tests were significant at $p < 0.05$ with the predicted sign. In

Table 2. RPE absolute value model statistical tests

Condition	NAcc	AI	Caud	Put	MPFC	Am	PCC	Thal
1.1	+	=	+	+	+	+	=	+	+
1.2	+	+	+	+	=	+	=	+	+
1.3	+	+	+	+	+	+	+	+	+
1.4	+	+	+	+	+	+	+	+	+
1.5	+	+	+	+	+	+	+	+	+
1.6	+	+	+	+	+	+	+	+	+
1.7	+	−	+	=	−	−	−	−	=
1.8	+	−	+	+	−	−	−	−	=
1.9	+	+	+	+	+	+	−	−	+
1.10	+	=	+	+	+	=	=	−	=
1.11	+	=	+	+	=	=	−	−	+
1.12	+	=	+	+	+	+	=	=	+
1.13	+	+	+	+	−	=	=	=	=
1.14	+	+	+	+	+	+	+	+	=
1.15	+	+	+	+	+	+	+	=	+
1.16	+	+	+	+	+	+	+	+	+
1.17	+	−	+	−	−	−	−	=	=
1.18	+	+	+	+	+	+	−	+	=
1.19	+	−	+	−	=	=	=	=	+
1.20	+	=	+	+	+	+	=	+	+
1.21	+	+	+	+	+	+	+	+	+
1.22	+	+	+	+	+	+	+	+	+
1.23	+	+	+	+	+	+	=	+	+
1.24	+	+	+	+	+	+	+	+	+
1.25	=	+	=	+	+	+	+	+	+
1.26	=	+	=	+	+	+	+	=	=
1.27	=	+	=	+	+	+	+	=	=
1.28	=	−	−	−	=	−	=	+	=

Testing the RPE absolute value model requires 28 statistical tests. Wald tests of linear restriction were performed on parameter estimates with an analysis window of TR 24–28 and a baseline of TR 22–23 (parameter estimates for the nucleus accumbens and anterior insula are in Figure 7) with the sign of all significant tests indicated ($p < 0.05$). The leftmost column indicates the predicted signs for an RPE absolute value signal. Tests 1.1–1.24 compare outcomes to other outcomes with lower probability. Tests 1.25–1.28 compare outcomes to other outcomes with the same prize probability. Tests are listed in supplemental data. ROI abbreviations are as in Table 1. NAcc, Nucleus accumbens; AI, anterior insula; Caud, caudate; Put, putamen; MPFC, medial prefrontal cortex; Am, amygdala; PCC, posterior cingulate cortex; Thal, thalamus.

Figure 8, we plot the results of this analysis, using the conventions in Figure 5. While most tests have the predicted sign for a range of baseline and analysis windows for the anterior insula (Fig. 8B), this is not the case for the nucleus accumbens (Fig. 8A) or amygdala (Fig. 8F). Some evidence for an RPE absolute value signal was present in other areas, including the thalamus and caudate in particular (Fig. 8) and the substantia nigra (supplemental Fig. 5, available at www.jneurosci.org as supplemental material).

Discussion

Neuroimaging studies have identified brain areas where BOLD activity is correlated with highly specified RPE models. Here, we used an axiomatic model to show that BOLD activity in the nucleus accumbens satisfies necessary and sufficient conditions for the RPE model class. This signal can encode RPEs (some RPE model accurately describes this signal), as previous studies have suggested but never formally tested. Signals measured from the caudate, putamen, amygdala, medial prefrontal cortex, and posterior cingulate cortex also satisfy the axiomatic RPE model: for each area there exists some RPE model that can account for measured activity. This approach required none of the auxiliary assumptions about unobservable subjective variables like reward and expectation necessary with the traditional regression approach. Rather than looking for correlation with specific RPE models, the axiomatic approach tests the properties critical to the entire RPE model class. We show here that BOLD activity in the anterior insula falsifies the axiomatic model and cannot, in prin-

ciple, encode RPEs under these conditions, despite positive findings in previous studies with other methods. This BOLD activity may instead encode a signal related to salience.

Reward prediction error models and the anterior insula

We have identified six brain areas that strongly satisfy the axiomatic model. However, our most surprising result is that BOLD activity in the anterior insula robustly falsifies the axiomatic model. There is no way of defining or parameterizing an RPE model to account for this BOLD signal. This is a critical logical feature of the axiomatic approach, allowing us to unambiguously contradict the hypothesis that BOLD activity in the anterior insula encodes some kind of RPE signal (Seymour et al., 2004; Pessiglione et al., 2006; Wittman et al., 2008; Voon et al., 2010). There is no doubt that anterior insula activity can be correlated with RPE model predictions. However, our tests suggest that the limits of observed correlations arise from properties of the signal that are fundamentally incompatible with any RPE representation; this activity can be correlated with some RPE models rather than actually encoding some type of RPE signal.

However, it is important to note that we did not find anterior insula activity correlated with RPE regressors in a traditional correlation-based analysis. RPE model predictions may be particularly well correlated with features of a salience representation in a task other than ours. It may also be that beneath the level of the BOLD signal, some anterior insula neurons represent RPEs or some part of an RPE signal, a possibility that we cannot address here. Our data indicate that aggregate anterior insula activity measured by fMRI falsifies the axioms and thus cannot encode RPEs.

Many studies identifying RPE model correlations in the anterior insula involved pain (Seymour et al., 2004) or financial losses (Pessiglione et al., 2006; Voon et al., 2010), events that may be particularly salient. The anterior insula is also implicated in representing uncertainty (Huettel et al., 2005; Grinband et al., 2006), prediction errors related to reward variance (Preuschoff et al., 2006, 2008), and in processing salient stimuli (Jensen et al., 2007; Seeley et al., 2007). Ullsperger and von Cramon (2003) reported greater anterior insula activity for negative than for positive feedback when negative feedback is less frequent (and thus possibly more salient).

However, BOLD activity in the anterior insula almost completely satisfied an alternate model for RPE absolute value. In our task, this activity is a decreasing function of prize probability, consistent with some notions of salience. This is important, because dopamine neurons and dopamine-related activity may encode salience in addition to or instead of RPEs (Berridge and Robinson, 1998; Redgrave et al., 1999; Horvitz, 2000; Zink et al., 2003, 2004). A recent electrophysiology study identified a subpopulation of neurons in the dorsolateral substantia nigra that increases its activity in response to unexpected appetitive and aversive events (Matsumoto and Hikosaka, 2009) and, although controversial, it has been suggested that these are dopamine neurons. Another study identified a subset of dopamine neurons in anesthetized rats responsive to aversive events (Brischoux et al., 2009). One must be cautious in interpreting these results with respect to RPEs, because it may be problematic to relate activity in unconscious animals to reward predictions. Although we found that anterior insula activity may encode RPE absolute value, a quantity related to salience, we did not find activity in any other area satisfying the constraints of the RPE absolute value model.
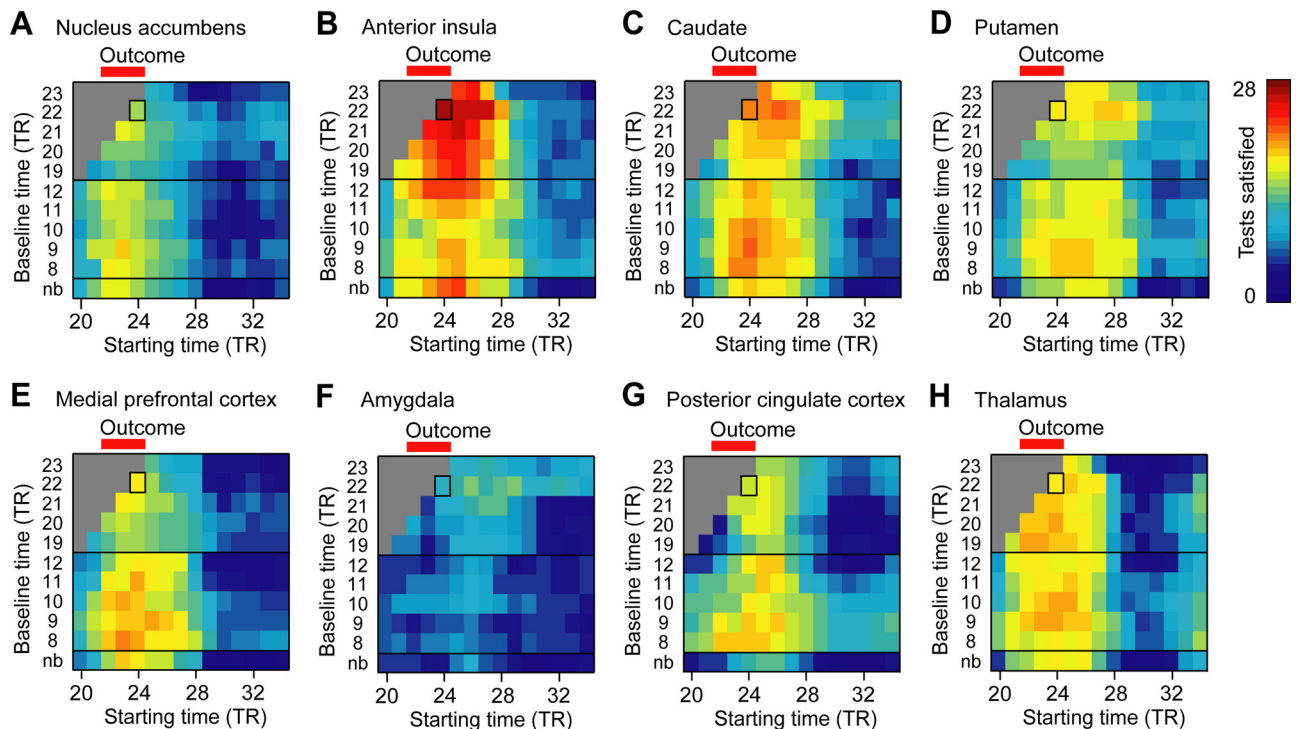
Figure 8. Assessing the robustness of RPE absolute value model analyses. **A–H,** Heat maps show results of the analysis for a variety of baselines and starting times for the 5-TR analysis window. Testing the RPE absolute value model requires 28 statistical tests. The first TR of the baseline is indicated for each 2-TR baseline. The color scale indicates the number of tests with the predicted result for an RPE absolute value signal at $p < 0.05$. The baseline and analysis window used for the analyses in Figure 7 and Table 2 is indicated by rectangles. All ROIs are defined by anatomical criteria in individuals. The neural activity in the anterior insula has the predicted result for an RPE absolute value signal for most tests for a variety of baseline and analysis windows. The nucleus accumbens and amygdala do not have the predicted result for an RPE absolute value signal regardless of the choice of baseline and analysis window. nb, No baseline. Dopaminergic midbrain and habenula results are shown in supplemental Figure 5 (available at www.jneurosci.org as supplemental material).

Relating BOLD activity to dopamine

Electrophysiological results suggest that dopamine neurons encode RPEs (Schultz et al., 1997; Hollerman and Schultz, 1998; Nakahara et al., 2004; Bayer and Glimcher, 2005; Joshua et al., 2008; Matsumoto and Hikosaka, 2009; Zaghloul et al., 2009). Most of the regions we identified in which BOLD activity could encode RPEs receive direct dopaminergic projections. Although dopaminergic drugs influence learning rates (Rutledge et al., 2009; Voon et al., 2010) and modulate BOLD activity for putative striatal RPE signals (Pessiglione et al., 2006; Voon et al., 2010), we cannot conclude that the signals we measured are due to dopaminergic activity. We also cannot determine whether the signals we measured arise from multiple sources. We conclude only that aggregate neural activity measured by fMRI satisfies the axioms in six brain areas.

Although we did not find that BOLD activity in midbrain dopamine structures encodes RPEs, imaging these particular structures is difficult. Whether this reflects a discrepancy between BOLD and spiking activity or the limitations of our imaging protocol is unclear. D'Ardenne et al. (2008) found positive correlations with specific RPE models in ventral tegmental area BOLD activity using high-resolution imaging and midbrain-specific alignment algorithms. The habenula is another difficult-to-image structure perhaps encoding a sign-reversed RPE signal (Matsumoto and Hikosaka, 2007). We were unable to find evidence for this here using a standard imaging protocol.

Comparing axiomatic and correlation-based approaches

Our finding of RPE signals in the striatum is unsurprising. However, fewer reports have found possible RPE signals in the medial

prefrontal cortex (Behrens et al., 2008), amygdala (Yacubian et al., 2006), and posterior cingulate cortex (de Bruijn et al., 2009), although electrophysiological studies have found evidence for RPE signals in these areas (McCoy et al., 2003; Belova et al., 2007; Matsumoto et al., 2007). Our analysis using a typical RPE model revealed correlations in striatum and, at a liberal threshold, medial prefrontal cortex, but not in amygdala or posterior cingulate cortex. Our data suggest that the hemodynamic response functions (HRFs) in these areas (Fig. 6) bear little similarity to the canonical HRF most commonly used in standard regression analyses. This suggests that regression-based studies of reward areas using HRFs appropriate for these regions might produce significant correlations. We additionally increased our sensitivity to possible RPE signals by testing multiple baselines and analysis windows. Correlation-based approaches could increase their sensitivity with similar methods.

Unlike the axiomatic approach, the standard approach also assumes that responses to outcomes received from one-prize "lotteries" (like a tone followed by juice reward) are intermediate between responses to uncertain positive and negative outcomes. Inspection of trial averages reveals that this is not precisely the case in our data (Fig. 3C and supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Although our axiomatic methodology makes no assumptions about how these responses relate to two-prize lotteries, these qualitatively different signals, which might differ in magnitude, timing, or both, identify another weakness of some common approaches.

One disadvantage of the axiomatic approach is that it requires independently identified ROIs. Regression-based methods can identify candidate RPE areas with whole brain analyses. For

example, Behrens et al. (2008) identified brain regions where activity increased with reward magnitude and decreased with predicted reward in a conjunction analysis. These separate regressors distinguish key features of RPE signals closely related to the first and second axioms. Such an analysis does not require a ROI-based approach, but does require additional assumptions unnecessary in our approach.

Another advantage of correlation-based approaches is that they can be used to separate multiple components of a neural signal. We cannot determine whether subelements of BOLD activity encode some part of an RPE signal. Multiple regressors (particularly when appropriately orthogonalized) can separate signal components in a way that the axiomatic approach cannot attempt. This feature of the traditional approach is particularly useful in complex tasks where a single brain area might, for example, receive diverse inputs.

While our axiomatic model can determine whether a signal can or cannot encode RPEs, there are clearly situations in which correlation-based analyses might more usefully explain neural data. The axiomatic methodology presented here complements correlation-based analyses, with each methodology best suited for addressing particular questions.

The axiomatic methodology is of particular interest because it adds an additional tool to neuroscience. Where we can falsify at least one of the axioms we can reject the entire RPE model class. Where the axioms are satisfied, additional axioms refine the model. Future research could establish whether the quantity of dopamine release in these areas measured with electrochemical methods (Phillips et al., 2003; Day et al., 2007) satisfies the axiomatic model, directly testing the linkage between RPE representations and dopamine. Axiomatizing the salience hypothesis would allow testing conditions of necessity and sufficiency on BOLD activity in the anterior insula.

Conclusion

This study introduces axiomatic modeling to neuroscience. We formally tested the RPE hypothesis, showing both that activity from six brain areas satisfies the axioms of an RPE representation and that anterior insula activity falsifies the axioms and cannot possibly encode RPEs under the conditions examined. In contrast, the standard regression approach relies on highly parameterized models with specific assumptions about reward, beliefs, and learning when it examines a theory like the hypothesis that dopamine-related activity encodes RPEs. Such an analysis yields a correlation coefficient but no direct test of the actual hypothesis under scrutiny. The axiomatic approach provides a powerful alternative in the Popperian tradition of testing a hypothesis by attempting to falsify it. By breaking hypotheses down into their basic assumptions, entire model classes can be tested. These assumptions identify the ways a model can be proven false. When data falsify an axiom, new theoretical approaches are suggested, unlike when low correlations are observed in traditional analyses. In this sense, the axiomatic approach offers novel benefits that complement existing approaches to the analysis of brain function.

References

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. Neuroimage 31:790–795.

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141.

Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of social value. Nature 456:245–249.

Belova MA, Paton JJ, Morrison SE, Salzman CD (2007) Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. Neuron 55:970–984.

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

Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci U S A 106:4894–4899.

Caplin A, Dean M (2008a) Dopamine, reward prediction error, and economics. Q J Econ 123:663–701.

Caplin A, Dean M (2008b) Axiomatic methods, dopamine and reward prediction error. Curr Opin Neurobiol 18:197–202.

Caplin A, Dean M, Glimcher PW, Rutledge RB (2010) Measuring beliefs and rewards: a neuroeconomic approach. Q J Econ 125:923–960.

Caviness VSJ, Meyer J, Makris N, Kennedy DN (1996) MRI-based topographic parcellation of the human neocortex: an anatomically specified method with estimate of reliability. J Cogn Neurosci 8:566–587.

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.

Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associate learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci 10:1020–1028.

de Bruijn ERA, de Lange FP, von Cramon DY, Ullsperger M (2009) When errors are rewarding. J Neurosci 29:12183–12186.

Friston KJ, Josephs O, Rees G, Turner R (1998) Nonlinear event-related respones in fMRI. Magn Reson Med 39:41–52.

Grinband J, Hirsch J, Ferrera VP (2006) A neural representation of categorization uncertainty in the human brain. Neuron 49:757–763.

Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of orbitofrontal cortex and striatum in the computation of goal values and prediction errors. J Neurosci 28:5623–5630.

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309.

Horvitz JC (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience 96:651–656.

Huettel SA, Song AW, McCarthy G (2005) Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. J Neurosci 25:3304–3311.

Jensen J, Smith AJ, Willeit M, Crawley AP, Mikulis DJ, Vitcu I, Kapur S (2007) Separate brain regions code for salience vs. valence during reward prediction in humans. Hum Brain Mapp 28:294–302.

Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H (2008) Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J Neurosci 28:11673–11684.

Li J, McClure SM, King-Casas B, Montague PR (2006) Policy adjustment in a dynamic economic game. PloS ONE 1:e103.

Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447:1111–1115.

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837–841.

Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. Nat Neurosci 10:647–656.

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

McCoy AN, Crowley JC, Haghighian G, Dean HL, Platt ML (2003) Saccade reward signals in posterior cingulate cortex. Neuron 40:1031–1040.

Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopamine neurons can represent context-dependent prediction error. Neuron 41:269–280.

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.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006)

Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045.

Phillips PEM, Stuber GD, Heien MLAV, Wightman RM, Carelli RM (2003) Subsecond dopamine release promotes cocaine seeking. Nature 422:614–618.

Popper K (1959) The logic of scientific discovery. New York: Basic Books.

Preuschoff K, Bossaerts P, Quartz SR (2006) Neural differentiation of expected reward and risk in human subcortical structures. Neuron 51:381–390.

Preuschoff K, Quartz SR, Bossaerts P (2008) Human insula activation reflects risk prediction errors as well as risk. J Neurosci 28:2745–2752.

Rademacher J, Galaburda AM, Kennedy DN, Filipek PA, Caviness VS Jr (1992) Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. J Cogn Neurosci 4:352–374.

Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward error? Trends Neurosci 22:146–151.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Classical conditioning II: Current research and theory (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton.

Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW (2009) Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. J Neurosci 29:15104–15114.

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

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349–2356.

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.

Sutton RS, Barto AG (1990) Time-derivative models of Pavlovian reinforcement. In: Learning and computational neuroscience: foundations of adaptive networks (Gabriel M, Moore J, eds), pp 497–537. Cambridge, MA: MIT.

Ullsperger M, von Cramon DY (2003) Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci 23:4308–4314.

Vazquez AL, Noll DC (1998) Nonlinear aspects of the BOLD response in functional MRI. Neuroimage 7:108–118.

Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010) Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. Neuron 65:135–142.

Wald A (1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans Am Math Soc 54:426–482.

Wittmann BC, Daw ND, Seymour B, Dolan RJ (2008) Striatal activity underlies novelty-based choice in humans. Neuron 58:967–973.

Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006) Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. J Neurosci 26:9530–9537.

Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, Kahana MJ (2009) Human substantia nigra neurons encode unexpected financial rewards. Science 323:1496–1499.

Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003) Human striatal response to salient nonrewarding stimuli. J Neurosci 23:8092–8097.

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.