Neuron 51, 381–390, August 3, 2006 ©2006 Elsevier Inc. DOI 10.1016/j.neuron.2006.06.024

Neural Differentiation of Expected Reward and Risk in Human Subcortical Structures

Kerstin Preuschoff,[1,*] Peter Bossaerts,[1,2] and Steven R. Quartz[1,2]
[1] Computation and Neural Systems Program
Social Cognitive Neuroscience Laboratory
[2] Division of Humanities and Social Sciences
California Institute of Technology
1200 East California Boulevard, 228-77
Pasadena, California 91125

Summary

In decision-making under uncertainty, economic studies emphasize the importance of risk in addition to expected reward. Studies in neuroscience focus on expected reward and learning rather than risk. We combined functional imaging with a simple gambling task to vary expected reward and risk simultaneously and in an uncorrelated manner. Drawing on financial decision theory, we modeled expected reward as mathematical expectation of reward, and risk as reward variance. Activations in dopaminoceptive structures correlated with both mathematical parameters. These activations differentiated spatially and temporally. Temporally, the activation related to expected reward was immediate, while the activation related to risk was delayed. Analyses confirmed that our paradigm minimized confounds from learning, motivation, and salience. These results suggest that the primary task of the dopaminergic system is to convey signals of upcoming stochastic rewards, such as expected reward and risk, beyond its role in learning, motivation, and salience.

Introduction

When faced with decision-making in an uncertain world, it is fundamental to evaluate both expected rewards and risks. Higher expected rewards are usually preferred over lower expected rewards. But sensitivity to risk is also ubiquitous. For instance, when an investor has the option of either opening a simple savings account (low expected reward but a known outcome) or investing all of her money into a particular stock (higher expected reward but an uncertain outcome), she may prefer the option with the lower expected reward because of the higher risk of the alternative. Economic studies (Bossaerts and Plott, 2004; Holt and Laury, 2002) have confirmed that risk considerations, in addition to expected reward, indeed play a role in decision-making under uncertainty and in the valuation of risky gambles. This sensitivity to both expected reward and risk is not unique to financial situations. It is also observed in nonhuman primates facing uncertain rewards (Fiorillo et al., 2003; McCoy et al., 2003) and in bees choosing among different flowers (Real, 1991).

In neuroscience, evidence has accumulated that brain activation correlates with expected reward. Human fMRI studies have found that subcortical dopaminergic structures such as striatum are involved in reward-related processes. Activity in these structures correlates with reward value of a variety of stimuli, including primary rewards, such as gustatory stimuli (Berns et al., 2001; O'Doherty et al., 2002) and abstract stimuli, such as money (Breiter et al., 2001; Elliott et al., 2000, 2003; Knutson et al., 2000, 2001, 2003). Studies of nonhuman primate conditioning document a monotonically increasing relationship between phasic activity of midbrain dopamine neurons and reward probability or expected reward (Fiorillo et al., 2003; Tobler et al., 2005).

The correlations found between risk and activation in *cortical* regions are unambiguous (McCoy et al., 2003; Huettel et al., 2005, 2006). However, correlations found between risk and activation in *subcortical* regions are not. In the nonhuman primate brain, delayed firing of dopaminergic neurons was positively correlated with risk when risk was modulated by changing reward probabilities (Fiorillo et al., 2003; Tobler et al., 2005). Correlation was also positive in caudate neurons when risk was modulated by manipulating problem complexity (stimulus recognition uncertainty; Lauwereyns et al., 2002; Takikawa et al., 2002). For the human brain, however, the findings are inconsistent. When risk was modulated by altering the degree to which one knows the probability of reward, i.e., when manipulating knowledge of the reward probability rather than the probability itself, activity in striatum correlated negatively with risk (Hsu et al., 2005). In contrast, when risk concerned problem complexity (categorization uncertainty; Grinband et al., 2006), striatal activation correlated *positively* with risk.

Consequently, to date we are aware of no studies that have examined whether and how subcortical dopaminoceptive regions in the human brain code for risk when risk is modulated by changing reward probability. This is the primary type of modulation for decision-making under uncertainty, and it needs to be understood before modulating knowledge of probabilities and studying learning of probabilities.

Here we manipulated probabilities so that not only risk changed over the full range, but also expected reward, and in such a way that expected reward and risk varied orthogonally. We then determined whether and how activation in subcortical dopaminoceptive regions correlated with expected reward and risk.

In addition, our study addresses a number of important open issues about the representation of expected reward and risk in subcortical dopaminoceptive regions. First, since neuroscientific studies have examined expected reward for a limited number of values, the precise representation, or mathematical model, of reward expectation in the brain remains unknown. As financial decision theory models reward expectation as mathematical expectation of reward (Knutson et al., 2003), it makes specific predictions regarding the form brain activation must take to represent expected reward. When reward is kept constant across rewarded trials,

*Correspondence: kerstinp@caltech.edu

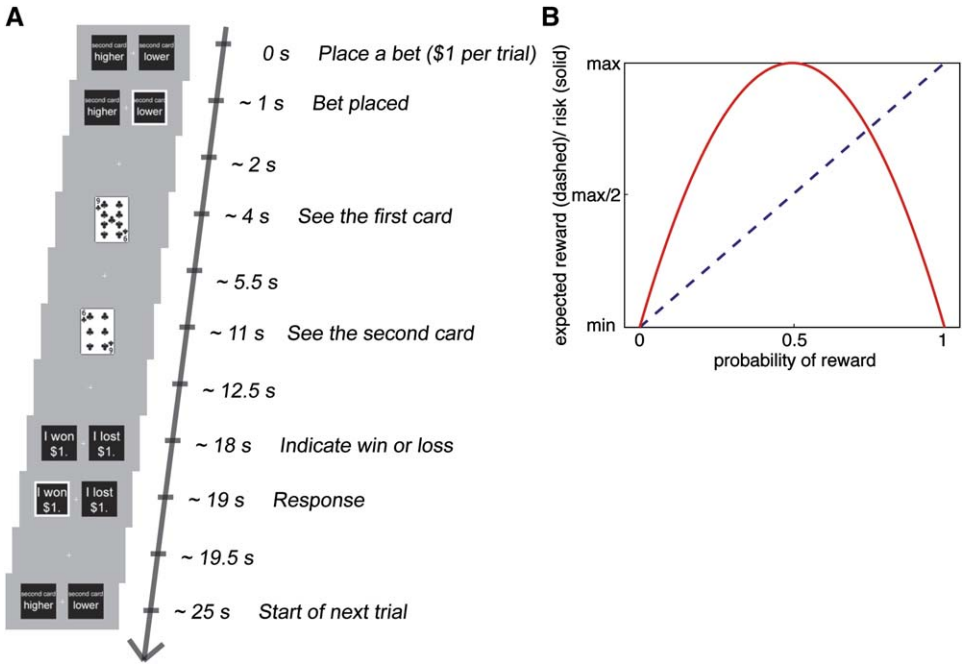

Figure 1. Experimental Design

(A) On each trial, two cards were drawn (without replacement within each trial) from a deck of ten, numbered 1 through 10. Before seeing either card, subjects first placed a $1 bet on one of two options, "second card higher" or "second card lower" (than first card shown). Subjects could earn $1 if they guessed the right card and lost $1 if they were wrong. Once the bet was placed, subjects saw card 1, followed ~7 s later by card 2. At the end of each trial, subjects had to indicate whether they won or lost on this trial. A $0.25 penalty was imposed for misreporting, independent of the outcome of the gamble. All times shown are with respect to the onset of the trial.

(B) Expected reward and risk as a function of the probability of reward. Expected reward, measured as mathematical expectation of reward, increases linearly in the probability of reward p (dashed line). Expected reward is minimal at $p = 0$ and maximal at $p = 1$. Risk, measured as reward variance, is an inversely quadratic function of probability that is minimal at $p = 0$ and $p = 1$ and maximal at $p = 0.5$ (solid line). As such, expected reward and risk are orthogonal over the full range of probabilities, p in [0,1]. When subjects place their bet, the reward probability p is 0.5. After display of card 1, the reward probability changes, depending on whether the subject bet that the second card is higher or lower, and depending on the number on card 1. If the subject bet that the second card is going to be lower, then p increases linearly in the number on card 1; otherwise p decreases linearly in the number on card 1.

mathematical expectation of reward increases linearly in the probability of reward. We therefore hypothesized that brain activation increases *linearly* in reward probability if it is to reflect expected reward. To test this representation hypothesis stemming from financial decision theory, the probability of reward p needs to be varied over all probabilities (ranging from $p = 0$ to $p = 1$) with a sufficient number of intermediate values.

Second, the specific form of risk representation in the brain is unknown. As with reward expectation, financial decision theory suggests a specific metric for measuring risk, namely, variance, the mean squared deviation from the expected outcome (Markowitz, 1952). When reward magnitude is kept constant across rewarded trials, reward variance is quadratic in reward probability p; variance attains a maximum at $p = 0.5$, and minimums at the extremes, $p = 0$ and $p = 1$ (Figure 1B). Because variance is monotonically increasing for $p < 0.5$ and monotonically decreasing for $p \geq 0.5$, care has to be exercised that p varies sufficiently to ensure that the effects of changes in risk and expected reward can be disentangled. Otherwise, an increase in risk may be confounded with a change in expected reward (Critchley et al., 2001; Dreher et al., 2005; Knutson et al., 2003). From a statistical point of view, variance and expected reward become orthogonal if they are varied over the full range of reward probabilities and are mean corrected (i.e., after subtracting their average values).

Third, previous neuroscience studies have focused on the learning aspect of reward anticipation (Hollerman and Schultz, 1998; Mirenowicz and Schultz, 1994; Romo and Schultz, 1990), leaving it unclear whether activation related to reward expectation and risk in subcortical structures requires learning and motivation (Knutson et al., 2001) to be present. Many of these studies have been guided by Temporal Difference (TD) models of learning (Sutton, 1988). In the case of risk encoding in dopaminergic neurons, it too has been interpreted in terms of reward learning (Fiorillo et al., 2003; Tobler et al., 2005). The controversy regarding the interpretation of subcortical dopaminergic activation is compounded by studies that suggest that such activation may represent salience (Zink et al., 2004) or nonspecific forms of uncertainty (Aron et al., 2004; Berns et al., 2001). Our hypothesis is that the primary task of the dopaminergic system is to convey signals of upcoming stochastic rewards, like expected reward and risk, while learning, salience, and motivation constitute only secondary, albeit important, tasks. Testing this hypothesis requires disassociating the signaling task from learning, salience, and motivation, which requires a perceptual experimental paradigm, unlike the previously employed

conditioning tasks. Elimination of learning confounds is especially important because the correlation of sustained activation of dopaminergic neurons with uncertainty (Fiorillo et al., 2003) has been interpreted as the effect of backpropagation of reward prediction errors during learning (Fiorillo et al., 2005; Niv et al., 2005).

Finally, if a single brain system, the dopaminergic system, is to represent two parameters (expected reward and risk) of a single phenomenon (a gamble), the issue of discrimination arises. Discrimination could be achieved spatially, in which different regions could specialize in encoding the different parameters or distinct neural populations within the same region of the brain could encode different parameters. Another possibility is that discrimination could be achieved temporally, in which the same subregion sequentially encodes the two parameters.

To test how subcortical dopaminergic structures encode these two parameters, it is necessary to utilize an experimental design that allows for distinguishing these alternative encoding strategies. Based on recent evidence of activation of dopaminergic neurons in the nonhuman primate brain when risk is modulated by varying probability (Fiorillo et al., 2003), we expected to find an early-onset activation in subcortical dopaminoceptive regions that correlated positively with expected reward, while a late-onset activation would correlate positively with risk. We therefore allowed for sufficient time between stimulus and outcome and used a statistical analysis of the imaging data that is able to capture potential temporal differentiation.

Nineteen subjects played a gamble where two cards were drawn (without replacement within each trial) from a deck of ten, numbered 1 through 10 (Figure 1A). Before seeing either card, subjects first placed a $1 bet on whether the first or the second card would be higher. Once the bet was placed, subjects saw card 1, followed ~7 s later by card 2. We refer to the time interval between display of card 1 and card 2 as the anticipatory period. Upon display of card 1, the probability of winning changes as a function of the number on card 1. For instance, if the subject bet on "second card higher," the probability of winning is given by the number of cards initially in the deck (always 10) minus the number displayed on the first card (C) and divided by the number of cards remaining in the deck: $p = (10-C)/9$.

Since a new deck was used on every trial, subjects had no prior information about the outcome of the gamble, so that on any given trial the initial probability of winning at the time of bet was $p = 0.5$ with maximal risk. As a result, gains and losses were independent of the strategy the subject chose. In addition, there is no role for learning, as any strategy is optimal. Following the presentation of card 2, subjects were asked to report whether they won or lost. In case of an incorrect response subjects lost $0.25, independent of whether their gamble had paid off. As such, motivation (the degree to which one is willing to work to report whether one won or lost) during the anticipation period should not depend on expected reward or risk.

Reward level was kept constant across all rewarded trials. Because of this, expected reward and risk (variance) upon display of card 1 change only as a function of the probability of winning, as shown in Figure 1B. Altering the reward level would have potentially introduced a confounding factor, namely, varying complexity, which is known to induce activation in subcortical dopaminoceptive structures in itself (Grinband et al., 2006).

Results

In this section, we first report statistics on task performance. Using a voxel-based analysis, we subsequently document that subjects encoded the task as a reward prediction problem by replicating previously found activation patterns for reward. We then focus on the anticipatory period to find regions of interest (ROIs) whose activity is modulated by expected reward and risk. We determine how activity varies with the probability of reward within the identified ROIs to verify that activations correlate with mathematical expectation of reward and reward variance. Finally, we report on tests to determine whether the results were affected by learning, motivation, or salience.

Task Performance

Participants won on 48.60% ± 3.95% of all trials and correctly reported the outcome of their bet on 97.8% ± 2.6% of all trials, showing that the gamble was indeed random and subjects kept track of the cards displayed on the screen.

Reward Activation

The contrast between wins and losses (i.e., the difference in activation following wins versus that following losses) revealed significant activation ($p < 0.0001$) of a subcortical network including caudate, globus pallidus, thalamus, and putamen as well as midbrain and cingulate gyrus (see Table S1 in the Supplemental Data), in agreement with previous reports (Delgado et al., 2000; Elliott et al., 2003; Knutson et al., 2000, 2001, 2003). The contrast between losses and wins, i.e., "negative reward," revealed no significant activation, which is also supported by prior findings (Knutson et al., 2003). This indicated that subjects were encoding the task as a reward prediction problem and motivated our investigation of decision variables underlying this response.

Anticipatory Period Activation

Focusing on the anticipatory period, we first used a model to define ROIs that correlate with expected reward and risk (reward variance) during this period. Based on nonhuman primate evidence that temporally distinct responses of dopaminergic neurons might encode expected reward and risk respectively (Fiorillo et al., 2003), we decomposed the anticipatory period into (i) a response at the onset of card 1 (initial subperiod), followed by (ii) a response until the onset of card 2 (subsequent subperiod). The duration of the initial response (i) was set at only 1 s to allow for an onset of the delayed response as early as 1 s after the stimulus. The duration of the subsequent response (ii) was therefore longer (~6 s), in accordance with the finding in the nonhuman primate brain of sustained neuronal firing that correlates with risk. Because the total length of the anticipatory period is only ~7 s, whereas hemodynamic responses typically peak at only about 4 s, we

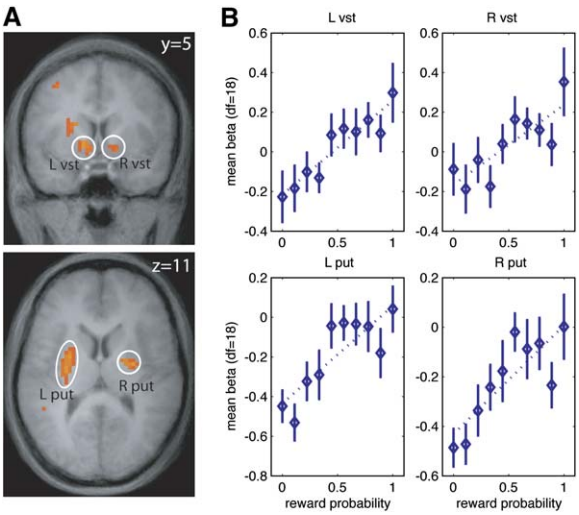

Figure 2. Immediate Neural Correlates of Expected Reward

(A) Neural activations related to expected reward (immediate response within 1 s of display of card 1). Bilateral activity in putamen (L put, R put) and ventral striatum (L vst, R vst) correlates with the probability of win, and hence, expected reward (random effects, $p < 0.001$). Neural responses are displayed in axial and coronal formats.

(B) Mean activations (parameter estimates β with standard error) for ten probabilities. In both left and right ventral striatum (vst) and putamen (put) neural responses increase with increasing probability of win. Dashed line indicates the best linear fit (L vst: $r^2 = 0.87$, $p < 0.001$; R vst: $r^2 = 0.66$, $p < 0.01$; L put: $r^2 = 0.69$, $p < 0.01$; R put: $r^2 = 0.7$, $p < 0.01$).

Error bars = SEM.

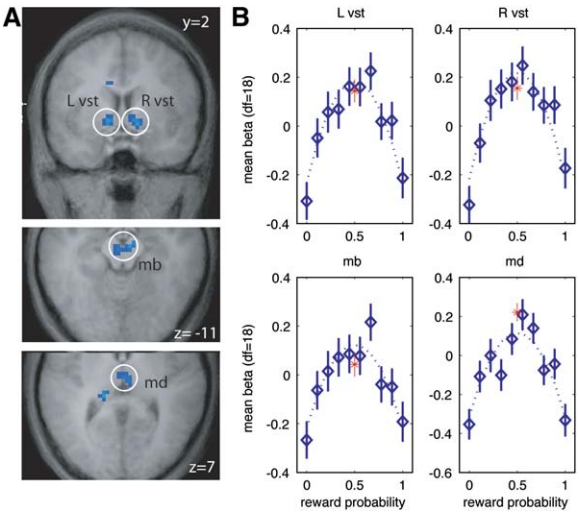

Figure 3. Delayed Neural Correlates of Risk

(A) Neural activations related to risk (delayed response, after 1 s of display of card 1 and until display of card 2). Brain regions whose activity correlates with reward variance, reflecting risk (random effects, $p < 0.001$), include left and right ventral striatum (L vst, R vst) extending into the subthalamic nucleus, midbrain (mb), and mediodorsal thalamic nucleus (md).

(B) Mean activations (parameter estimates β with standard error) for ten probabilities. Neural responses in regions displayed in (A) increase toward medium probabilities and decrease toward low and high probabilities. Dashed lines indicate best quadratic fit (L vst: $r^2 = 0.89$, $p < 0.001$; R vst: $r^2 = 0.88$, $p < 0.001$; mb: $r^2 = 0.84$, $p < 0.001$; md: $r^2 = 0.80$, $p < 0.001$). Across all four regions, the quadratic fit is significantly better than a model that predicts low activation at $p = 0,1$ and high activation for $p \neq 0,1$ ($p < 0.001$; results for individual brain regions: L vst: $p < 0.01$; R vst: $p < 0.01$; mb: $p < 0.01$; md: $p = 0.01$). Red data points (asterisks) at $p = 0.5$ indicate late-onset activation levels between the time of bet and card 1 when risk is maximal. Error bars = SEM.

did not, however, expect to be able to detect the precise length of the respective responses, and hence, to differentiate between phasic and sustained durations. We therefore only focused on differentiation of the onset of the signal: early [response (i)] versus later [response (ii)].

We next examined whether the activation we observed conformed to the model of expected reward and risk as specified in financial decision theory. We tested the hypothesis that activation levels relate to reward probability in the way that mathematical expectation of reward and reward variance relate to reward probability (Figure 1B). We changed the specification of our general linear model to compare activation levels at different probabilities within the identified ROIs. To do this, in the new model, one predictor for each individual probability level replaces the predictors for expected reward and risk in the old model.

Modulation of Anticipatory Period Activation by Expected Reward

Over the initial subperiod (1 s) of the anticipatory period, expected reward was highly correlated with activation in putamen, ventral striatum, globus pallidus, anterior cingulate cortex, midbrain, and other regions (Figure 2A; Table S2). We also detected significant activation to expected reward during the subsequent subperiod (6 s) in several foci in the cerebellum and medial temporal gyrus. Although our imaging sequence was not optimized for frontal regions, we also found activation in medial orbital gyrus and gyrus rectus (Table S3).

Based on our a priori hypothesis that subcortical structures encoded expected reward as mathematical expectation of reward, and hence, that activation increased linearly in reward probability, we compared the responses in ventral striatum and putamen for each of the ten reward probabilities that arose as a result of the number on card 1 (Figure 2B). Activation in bilateral ventral striatum (L vst, R vst) and putamen (L put, R put) showed a linear increase with increasing reward probability; the best linear fit is highly significant and explains a large proportion of the variance of the mean activation levels (L vst: $r^2 = 0.87$, $p < 0.001$; R vst: $r^2 = 0.66$, $p < 0.01$; L put: $r^2 = 0.69$, $p < 0.01$; R put: $r^2 = 0.7$, $p < 0.01$).

Modulation of Anticipatory Period Activation by Risk

During the second (6 s) subperiod of the anticipatory period, risk was highly positively correlated with activation in an area extending posterior to and bilateral from the ventral striatum to the subthalamic nucleus as well as mediodorsal thalamic nucleus, midbrain, and bilateral anterior insula (Figure 3A). Risk was not significantly correlated over the initial (1 s) subperiod with activation in any of the subcortical regions of interest except for midbrain. Instead, risk correlated significantly over this subperiod with activation in the anterior insula and orbitofrontal cortex (Table S4). As we are focusing on subcortical structures, we do not elaborate here on the latter finding.

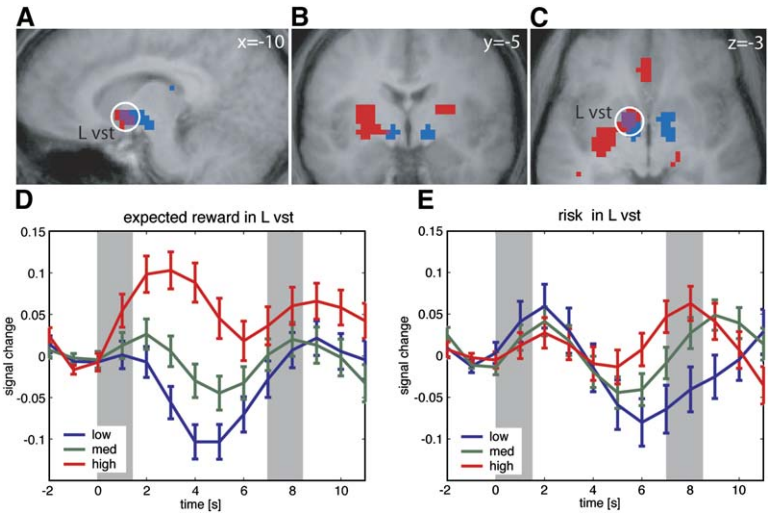

Figure 4. Temporal Encoding of Expected Reward and Risk

(A to C) Immediate (within 1 s of display of card 1) activations related to expected reward (probability of win; red) and delayed (after 1 s of display of card 1) activations related to risk (blue) superimposed on a mean anatomical image. Activations are identical to those shown in Figure 2 and Figure 3 but are not pseudo-color-coded in this map. (A and C) Spatial relationship between encoding of expected reward and risk include overlapping regions (displayed in purple) in ventral striatum (vst) and spatially contiguous areas. (D and E) Averaged adjusted time courses showing different temporal patterns for representations of expected reward and risk during the anticipatory period (t = 0, 1, …7 s) in the same subregion of left ventral striatum (purple region in [A] and [C]). Separation of time courses for low, medium, and high expected reward trials peaks early in the anticipatory period. Separation of time courses for low, medium, and high risk trials starts later and peaks around the time card 2 is shown. Gray bars indicate the presentation of card 1 (t = 0 s) and card 2 (t = 7 s). Error bars = SEM.

To determine whether the risk-related activation over the second (6 s) subperiod of the anticipatory period reflected reward variance, we used the same approach as for expected reward and studied activation separately for each of the ten different reward probability levels (Figure 3B). If the activations reflected reward variance, then their relationship with reward probability should be quadratic, with maximum at $p = 0.5$, and minima at $p = 0$ and 1. We found that responses in ventral striatum (L vst, R vst), midbrain (mb), and thalamic nucleus (th) are indeed maximal at intermediate probabilities and minimal at both minimal ($p = 0$) and maximal ($p = 1$) probabilities. Furthermore, the responses in all four regions of interest were shown to correlate with a function that is inversely quadratic (inversely u-shaped) in the probability of winning, with a maximum at p = 0.5; the best quadratic fit is highly significant and explains a large proportion of the variance of the mean activation levels (L vst: $r^2 = 0.89$, $p < 0.001$; R vst: $r^2 = 0.88$, $p < 0.001$; mb: $r^2 = 0.84$, $p < 0.001$; md: $r^2 = 0.80$, $p < 0.001$). To ensure that the close quadratic fit did not merely occur because activation is low at $p = 0$ and 1 while high elsewhere, we performed a standard nonnested hypothesis test (Davidson and MacKinnon, 1981) that determined whether a simple model with low activation at $p = 0,1$ and high for $p \neq 0,1$ should be rejected in favor of the quadratic model. Across the four brain regions, the quadratic fit was found to be significantly better ($p < 0.001$; results for individual brain regions: L vst: $p < 0.01$; R vst: $p < 0.01$; mb: $p < 0.01$; md: $p = 0.01$).

Finally, if the encoding indeed reflects variance of reward (risk as measured in financial decision theory), we should be able to use activation levels during the anticipatory period to successfully predict activation levels *after* the bet but *before* display of card 1. At the time of bet, the probability of win is p = 0.5, and risk is maximal (Figure 1B). Therefore, the activation level in ventral striatum must be similar to the levels for p = 0.5 during the anticipatory period. Figure 3B shows the level of activation in the same ROIs after the bet and before card 1 is shown. This activation level falls into the confidence interval of when reward variance is maximal. Since the periods before card 1 and 2 are not identical in length, the additional data points have to be evaluated carefully. Nonetheless, they provide corroborating evidence supporting the hypothesis that these ROIs reflect reward variance.

Modulation of Anticipatory Period Activation by Both Expected Reward and Risk

When simultaneously mapping the activation clusters reported for expected reward (activation over the initial 1 s subperiod) and risk (activation over the subsequent 6 s subperiod), a region in left ventral striatum emerged where the clusters overlap (Figures 4A–4C). We defined this as a region of interest to determine how the different levels of expected reward and risk were reflected in the time courses. We compared average adjusted hemodynamic responses to card 1 for low, medium, and high expected reward (Figure 4D), and low, medium, and high levels of risk (reward variance; Figure 4E). Early during the anticipatory period, the hemodynamic response increased with the level of expected reward, whereas starting from about 4 s after the onset of the anticipatory period (when card 1 is displayed), the hemodynamic response increased with the level of risk.

Testing for Absence of Learning, Motivation, and Salience Confounds

As we were interested in determining whether there were subcortical activations that were related to decision-making parameters independently of their previously documented role in reward prediction learning, our task was designed so that learning would not improve the potential outcome of the gamble. Nonetheless, it may be possible that learning-related signals are generated during the task, particularly if the reward prediction learning role of these structures is primary.

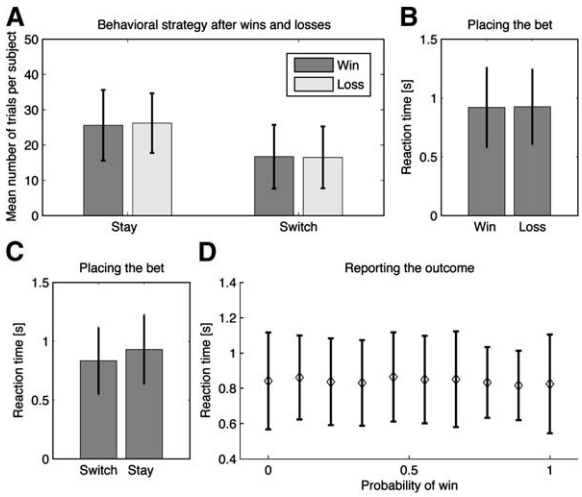

Figure 5. Relationship between Subject Behavior and Outcome History

(A) Histograms of choice behavior as a function of the outcome (win, loss) in the previous trial. "Stay" refers to a trial where the subject bets the same as in the previous trial; "Switch" refers to a trial where the subject bets differently from the previous trial (if subject bet that second card is higher in previous trial, then subject bids that second card is lower in current trial, and vice versa). Evidence of learning would emerge if subjects tend to switch more after a loss and tend to stay after a win.

(B) Mean reaction time from trial start to placement of the bet as a function of outcome (win, loss) in prior trial. Evidence of learning would emerge if reaction times tend to be shorter after gains than after losses.

(C) Mean reaction time from trial start to placement of the bet as a function of difference in choice between current and previous trial (switch, stay). Evidence of learning would emerge if reaction times tend to be shorter for stays than for switches.

(D) Mean reaction times from display of card 2 to reporting of outcome as a function of probability of reward as of display of card 1. Evidence of motivation would emerge if mean reaction times decrease in reward probability. Evidence of salience would emerge if mean reaction times are maximal for maximum risk ($p = 0.5$) and minimal for minimum risk ($p = 0,1$). Error bars = SEM.

See Table S6 for individual results.

To test this possibility, we probed our behavioral and imaging data for evidence of learning, salience, and motivation. We distinguished between *switch* and *stay* trials, defining a switch trial as one in which a subject chooses a different bet than in the previous trial; in a stay trial, the subject chooses the same bet. For instance, if a subject who chose "second card higher" on the previous trial chooses "second card lower" in the current trial, then the current trial is a switch trial.

Learning would imply that the likelihood of switching increases after a loss trial, while the likelihood of staying increases after a win trial. There was, however, no significant difference between the number of switches after loss trials versus win trials or the number of stays following win trials versus loss trials (Figure 5A). We do find a (insignificant) tendency to stay rather than switch regardless of outcome, which is consistent with previous reports on the status quo bias in decision-making under uncertainty (Samuelson and Zeckhauser, 1988; it should be added, however, that in our paradigm the status quo bias does not lead to inferior performance; any strategy is optimal, as pointed out before).

Learning would also imply that reaction times to placing the bet would be effected by previous outcomes. Reaction times to placing the bet corroborate the absence of learning effects: reaction times are approximately equal whether the previous trial generated a win or a loss (Figure 5B) and they do not differ significantly across stay and switch trials (Figure 5C).

To determine whether learning could have caused the reported activations in subcortical structures for expected reward and risk over the two subperiods of the anticipatory period, we included a variable in our general linear model that indicated whether the immediately preceding trial generated a loss or a gain. Under TD learning, activation for expected reward should be significantly correlated with this indicator variable (while activation for risk could be explained as an effect of backpropagation of prediction errors [Fiorillo et al., 2005; Niv et al., 2005]). We found no significant effect of prior-trial outcome in the regions where activation was found to be reflecting expected reward. The lack of an effect as predicted by TD learning indicates the absence of a learning confound.

We also examined reaction times to determine whether motivation or salience affected our results during the anticipatory period (Figure 5D). One could legitimately be concerned that higher expected reward induces higher motivation, while higher risk induces higher salience. The penalty for falsely reporting the outcome after the anticipatory period is independent of the outcome, while no reward is given to correct reporting of the outcome. Since both expected reward and risk (reward variance) are related to reward probability, we verified the absence of motivational and salience confounds by plotting reaction times to outcome reporting against reward probability. Figure 5D confirms that there is indeed no relationship.

While the data presented in Figure 5 is pooled over all subjects, the results reported also hold on an individual subject basis (see also Table S6). Specifically, differences in strategy (switch or stay) after wins versus losses are not significant in 17 of 19 subjects ($p > 0.05$, not corrected for multiple comparison). There are no significant differences ($p > 0.05$) for any subject for reaction times after win versus loss trials and for switch versus stay trials. No significant linear or quadratic relationship between reaction times and probability of win emerges for any subject. Where applicable, we also tested for long-term effects of learning and found no significant results. We see a general trend of decreasing reaction time over time, which does not affect the results reported in Figure 5.

Discussion

By utilizing a design in which expected reward and risk as measured in financial decision theory varied orthogonally and across the full range, we tested whether activation in human primary projection targets of midbrain dopaminergic neurons significantly correlated with these two critical decision-theoretic parameters. Further, the paradigm was designed to minimize potential confounds from learning, motivation, or salience, allowing us to determine whether these target areas encode expected reward and risk, the primary parameters of financial decision theory, beyond their established role in learning.

We found that during reward anticipation, initial activation in ventral striatum and other subcortical dopaminoceptive structures varied with expected reward, whereas subsequent activation in ventral striatum varied with risk. Activations correlating with expected reward and risk were thus differentiated both spatially and temporally and arose in the absence of learning, motivation, or salience confounds.

Expected Reward Is Reflected in Linear Response to Probability

The response of the ventral striatum and other subcortical structures is highly linear in reward probability. This provides strong support that the phasic responses of ventral striatum and putamen encode the expected reward parameter of financial decision theory and as such goes beyond the monotonicity of encoding expected reward shown in previous studies.

The interpretation of the early response as encoding expected reward is also consistent with theoretical TD models (Montague et al., 1996; Montague and Sejnowski, 1994), which have primarily guided the investigation of dopaminergic structures (Knutson et al., 2003), though our results reveal that this signal is generated even in the absence of learning.

Reward Variance, or Risk, Is Reflected in Quadratic Response to Probability

The quadratic relationship between reward probability and the late response in ventral striatum and other subcortical structures supports the hypothesis that risk is encoded as reward variance in the brain. Variance, however, is just one of several measures of uncertainty that are all maximal at $p = 0.5$. Within neuroscience, entropy (= minus the weighted sum of the logarithm of the probabilities of each possible outcome) is the most common measure of uncertainty and has been used extensively in information-theoretic analysis of spike trains (Bialek and Rieke, 1992). One interpretation based on nonhuman primate electrophysiology in VTA has suggested that the sustained response may be encoding entropy (Fiorillo et al., 2003), since entropy is also maximal at $p = 0.5$. Closer inspection of the data demonstrates, however, that the sustained firing of dopaminergic neurons actually correlates with magnitude (Fiorillo et al., 2005). As such, variance (which is sensitive to both probability and magnitude), not entropy, is the right measure of risk. This is consistent with financial decision theory. Financial decision theory sometimes uses additional risk metrics (skewness, kurtosis, etc.), but these appear to be unnecessary to explain valuation when risk is as small as it is in our experiments (Bossaerts and Plott, 2004).

Separation of Expected Reward and Risk through Spatial and Temporal Differentiation

Though some subcortical regions are responsive exclusively to either expected reward or risk, others are responsive to both parameters, raising the question of how these signals are differentiated. Our results indicate that the brain differentiates these signals temporally: the initial response reflects expected reward; the subsequent response reflects risk. The distinct hemodynamic responses to expected reward and reward variance in human dopaminoceptive structures follow a pattern consistent with that found in nonhuman primate ventral midbrain using electrophysiology, which report an immediate, phasic encoding of expected reward and a late-onset, sustained encoding of risk (Fiorillo et al., 2003). The late onset happened to be too late relative to the duration of the anticipation period, however, to discriminate between phasic and sustained responses as well as one has been able to do in electrophysiological studies. Our results suggest that the downstream effects of temporally differentiated activation in the ventral midbrain result in both an early onset separation of signals correlating with expected reward and a late onset separation of signals correlating with risk in the target (dopaminoceptive) structures.

It is interesting to note that, with the exception of the midbrain, we failed to find an immediate activation in subcortical structures that correlated with risk. Activation in subcortical dopaminoceptive areas that correlated with risk is invariably delayed. This raises a number of issues worth investigating in future research. Is there an immediate signal for risk elsewhere in the brain (our data suggest that insula may play a role)? If so, how are the signals of expected reward and risk combined in order to guide decisions? What is the role of the delayed risk signal in subcortical dopaminoceptive areas? Is it used to improve learning, as suggested by Tobler et al. (2005)? Interestingly, the late-onset activity in ventral striatum looks similar to delayed activity in parietal cortex reported by Huettel et al. (2005, 2006).

Brain Activation Decomposes along Basic Financial Parameters of Monetary Gambles

Our investigation was guided by the mathematical model of decision-making under uncertainty stemming from financial decision theory. This model specifies the minimal parameters that are necessary for rational choice under uncertainty (expectation and variance of reward). Our study shows that brain activity correlates with these two parameters. In financial decision theory, expectation is balanced against variance, and this trade-off has led to important insights not only about simple animal behavior (e.g., bee foraging [Real, 1991]), but also about complex human activity, such as the demand for money and its relation to yields on fixed-income securities (Tobin, 1958), or the demand for and pricing of multiple risky securities. For instance, Sharpe (1964), Lintner (1965), and Mossin (1966) demonstrated that expected returns on risky securities should increase not as a function of their own risk (variance), but only to the extent that they contribute to the risk (variance) of the securities market as a whole. Experiments confirm these predictions (Bossaerts and Plott, 2004). Later, Black and Scholes (1973) showed that prices of options (to purchase or sell securities) increase as a function of risk—again measured by variance. It is striking that brain activation at the level of subcortical dopaminergic structures reflects the separation of expected reward and risk on which financial decision theory is based.

Objective Perception Independent of Choice, Learning, and Attention

As our results are obtained under purely perceptual conditions, i.e. when no choice is to be made subsequently, the activations we report are related primarily to the

assessment of risk and reward in gambles. Many levels of processing intervene between perception and choice, so it is possible that the brain tracks expected reward and risk at the perceptual level, while additional elements, such as contextual factors (e.g., decisions by others [Abel, 1990]), modulate choice. As such, perception of reward and risk may continue even if choice is not affected (Bayer and Glimcher, 2005). Absent subsequent choice, brain activity may merely reflect information gathering for the case that a choice opportunity would suddenly and unexpectedly arise.

It is important to point out that our goal was to investigate the perception of risk and reward and whether such perception conformed to a specific mathematical model. Our finding that at the perceptual level the brain conforms to this model is entirely consistent with the fact that there may also be subjective representations of decision-making parameters that vary from this model, as they may simply be different levels of representation and/or generated under different contexts. It is important for future research to investigate when and where subjective representations of these parameters may also be generated and how these signals may be integrated or interact in the generation of choice behavior.

Likewise, our task does not involve conditioning. Behavioral data support the absence of conditioning, and statistical analysis of brain activity confirms the absence of conditioning confounds. Conditioning paradigms allow one to shed light on factors such as learning, motivation, or salience, in addition to perception (Critchley et al., 2001; Elliott et al., 2003; Ernst et al., 2004; Knutson et al., 2003; Rustichini et al., 2005; Zink et al., 2004). Here we show how expected reward and risk correlate with activation in subcortical dopaminoceptive structures when these additional elements are removed.

Implications for Pathological Decision-Making under Risk

Pathological behaviors ranging from addiction to gambling (Bechara et al., 1997), as well as a variety of mental illnesses such as bipolar disorder (Minassian et al., 2004) and schizophrenia (Shurman et al., 2005), are partially characterized by risk taking. To date, it is unknown whether such pathological decision-making under risk is due to misperception of risk or disruptions in cognitive processes, such as learning, planning, and choice. For example, a bipolar subject during a manic episode may invest in a risky business proposition either because they misperceive the risk to be lower than it actually is, or because they accurately perceive the risk to be high but may have impaired learning, attentional, working memory, or choice processes. To date, studies of pathological decision-making under risk have primarily utilized the Iowa Gambling Task (IGT) (Cavedini et al., 2002; Clark et al., 2001; Shurman et al., 2005), which was designed to assess sensitivity to future negative reinforcers (Bechara et al., 1997). Recent studies (Dunn et al., 2006; Maia and McClelland, 2004), however, suggest that impaired performance on the IGT may be due to impairments in reversal learning, working memory, attentional shift, and related high-level cognitive processes rather than misperceptions of risk per se. Since our task was designed to minimize the involvement of these high-level processes, in the future it may be utilized with clinical populations to determine whether alterations in risk perception accompany their changes in risky behavior. This may lead to a better understanding of the relative contributions of risk misperception versus cognitive impairments in these pathological cases, may suggest different treatment approaches, and may also gauge the impact on and the feedback from higher-level brain regions known to contribute to decision-making (e.g., ventro-medial prefrontal cortex [Fukui et al., 2005]).

Conclusion

In neuroscience, the investigation of the dopaminergic system in tasks involving uncertainty has emphasized reward prediction learning. Risk perception has been less well studied, yet it is central to decision-making under uncertainty, as formalized in financial decision theory. Our results show that brain activity in subcortical dopaminoceptive regions of the human brain can be separated, both spatially and temporally, into signals that correlate with (mathematical) expectation of reward and with reward variance (risk)—two fundamental parameters of financial decision theory. The role of human subcortical dopaminoceptive structures in reward-related processing is extended even further as our results suggest that these structures convey signals of upcoming stochastic rewards, like expected reward and risk, beyond these structures' role in learning, salience and motivation.

Experimental Procedures

A total of 19 subjects participated in the study (ten male, nine female; aged 18–30, mean age 21.4 years). All subjects gave full informed consent to participate in the study. The study was approved by the California Institute of Technology Institutional Review Board.

Experimental Paradigm

Each subject was given written instructions for the game and completed a brief training session outside the magnet. During scanning, trials were randomly ordered. For each session, subjects were provided with an initial endowment of $25. If no bet was placed, they lost automatically. They also lost $0.25 if they incorrectly reported the outcome of their bet or if they did not respond. Accumulated gains were shown only at the end of each session. Subjects played three sessions with 30 trials per session. At the end of the experiment, subjects selected one of the three sessions at random, which determined their final payoff.

fMRI Acquisition

Each scanning session included a localizer scan and T1-weighted MPRAGE anatomical scans (256 × 256 matrix, 176 1 mm sagittal slices) followed by the acquisition of functional images while subjects performed the gambling task. Images were acquired using a Siemens TRIO 3.0T full body MRI scanner using T2*-weighted PACE EPI (TR = 2000 ms, TE = 30 ms, 64 × 64, 3.28125 × 3.28125 mm^2, 32 3.0 mm slices, no gap, field of view = 210). For each subject, three functional runs were collected (392–400 scans each).

Data Processing and Analysis

Data were processed and analyzed using BrainVoyager v1.26. Preprocessing included motion correction (six-parameter rigid body transformation), slice timing correction, linear drift removal, high-pass filtering, normalization to Talairach space, and spatial smoothing with a full width at half-maximum Gaussian kernel of 8 mm. For each subject, a separate linear model was constructed that included the regressors described below as well as visual and motor activation. Regressors modeled the BOLD response to the specified events using a convolution kernel applied to a boxcar function.

A first-order autoregressive model was used to correct for temporal autocorrelations. For each subject, contrasts were calculated at every voxel in the brain. In a random-effects analysis, a one-sample t test determined where the average contrast value for the group as a whole (n = 19 subjects) differed significantly from zero. Statistical maps were thresholded for significance ($p < 0.001$) and cluster size (≥ 5 voxels). The model used to identify regions of interests decomposed the anticipatory period into two consecutive epochs: a short epoch (1 s from card 1) followed by a long epoch (6 s) modeling the remainder of the anticipatory period until card 2. Both epochs were modeled with three predictors, a 0th, 1st, and 2nd order term. The predictor for the 0th order term modeled the anticipatory period. The 1st and 2nd order terms modeled the same period, but predicted a hemodynamic response that scaled linearly and quadratically with the probability of win. Note that all three components were orthogonal with respect to one another. The model also included predictors for visual and motor activation as well as for wins and losses at the time of card 2. To compare activation levels at different probabilities (β estimates in Figures 2 and 3) within the identified regions of interest, the model was modified to include one predictor for each individual probability instead of the 0th, 1st, and 2nd order terms. For determining risk activation levels, a late-onset predictor between the bet and card 1 was also included to model the (maximal) risk at that time. Adjusted time courses (Figures 4D and 4E) are time courses corrected for the effects (confounds) in the reduced model (for probability time course, reduced model = full model − predictor for 1st order term [probability]; for risk time course, reduced model = full model − predictor for 2nd order term [risk]). This ensured that the effects shown were orthogonal to all effects captured by the reduced model. Any effect not included in the reduced model would show up in the adjusted data (error term), while any effect included in the reduced model should not show up. For the probability time course, adjusted data were grouped into low ($p < 0.3$), medium ($0.3 < p < 0.7$), and high ($p > 0.7$) probability trials, averaged over all trials (time-locked to card 1). This event-related average was plotted over time for each group. For the risk time course, adjusted data were grouped into low ($p = 0$ or $p = 1$), medium ($0 < p < 0.3$ or $0.7 < p < 1$), and high ($0.3 < p < 0.7$) risk trials, averaged over all trials (time-locked to card 1). This event-related average of residuals was plotted over time for each group.

Supplemental Data

The Supplemental Data for this article can be found online at http://www.neuron.org/cgi/content/full/51/3/381/DC1/.

Acknowledgments

This work was supported by NSF Grant 0093757, the David and Lucile Packard Foundation, and the Gordon and Betty Moore Foundation. We thank John O'Doherty and members of the Social Cognitive Neuroscience Laboratory for helpful comments. We thank Steven Flaherty for technical assistance and three anonymous reviewers whose comments resulted in considerable improvement of the paper.

Received: December 15, 2005
Revised: May 3, 2006
Accepted: June 26, 2006
Published: August 2, 2006

References

Abel, A.B. (1990). Asset prices under habit formation and catching up with the Joneses. Am. Econ. Rev. *80*, 38–42.

Aron, A.R., Shohamy, D., Clark, J., Myers, C., Gluck, M.A., and Poldrack, R.A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. J. Neurophysiol. *92*, 1144–1152.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129–141.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. Science *275*, 1293–1295.

Berns, G.S., McClure, S.M., Pagnoni, G., and Montague, P.R. (2001). Predictability modulates human brain response to reward. J. Neurosci. *21*, 2793–2798.

Bialek, W., and Rieke, F. (1992). Reliability and information transmission in spiking neurons. Trends Neurosci. *15*, 428–434.

Black, F., and Scholes, M. (1973). The Pricing of options and corporate liabilities. J. Polit. Econ. *81*, 637–654.

Bossaerts, P., and Plott, C. (2004). Basic principles of asset pricing theory: evidence from large-scale experimental financial markets. Rev. of Finance *8*, 135–169.

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

Cavedini, P., Riboldi, G., Keller, R., D'Annucci, A., and Bellodi, L. (2002). Frontal lobe dysfunction in pathological gambling patients. Biol. Psychiatry *51*, 334–341.

Clark, L., Iversen, S., and Goodwin, G. (2001). A neuropsychological investigation of prefrontal cortex involvement in acute mania. Am. J. Psychiatry *158*, 1605–1611.

Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron *29*, 537–545.

Davidson, R., and MacKinnon, J.G. (1981). Several tests for model specification in the presence of alternative hypotheses. Econometrica *49*, 781–793.

Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., and Fiez, J.A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. J. Neurophysiol. *84*, 3072–3077.

Dreher, J.C., Kohn, P., and Berman, K.F. (2005). Neural coding of distinct statistical properties of reward information in humans. Cereb Cortex. *16*, 561–573.

Dunn, B.D., Dalgleish, T., and Lawrence, A.D. (2006). The somatic marker hypothesis: a critical evaluation. Neurosci. Biobehav. Rev. *30*, 239–271.

Elliott, R., Friston, K.J., and Dolan, R.J. (2000). Dissociable neural responses in human reward systems. J. Neurosci. *20*, 6159–6165.

Elliott, R., Newman, J.L., Longe, O.A., and Deakin, J.F. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. J. Neurosci. *23*, 303–307.

Ernst, M., Nelson, E.E., McClure, E.B., Monk, C.S., Munson, S., Eshel, N., Zarahn, E., Leibenluft, E., Zametkin, A., Towbin, K., et al. (2004). Choice selection and reward anticipation: an fMRI study. Neuropsychologia *42*, 1585–1597.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. Science *299*, 1898–1902.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2005). Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. Behav. Brain Funct. *1*, 7.

Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., and Hanakawa, T. (2005). Functional activity related to risk anticipation during performance of the Iowa Gambling Task. Neuroimage *24*, 253–259.

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

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

Holt, C.A., and Laury, S.K. (2002). Risk aversion and incentive effects. Am. Econ. Rev. *92*, 1644–1655.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision making. Science *310*, 1680–1683.

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

Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., and Platt, M.L. (2006). Neural signatures of economic preferences for risk and ambiguity. Neuron *49*, 765–775.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. Neuroimage *12*, 20–27.

Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J. Neurosci. *21*, RC159.

Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., and Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. Neuroimage *18*, 263–272.

Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. Nature *418*, 413–417.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev. Econ. Stat. *47*, 13–37.

Maia, T.V., and McClelland, J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. Proc. Natl. Acad. Sci. USA *101*, 16075–16080.

Markowitz, H. (1952). Portfolio selection. J. of Finance *7*, 77–91.

McCoy, A.N., Crowley, J.C., Haghighian, G., Dean, H.L., and Platt, M.L. (2003). Saccade reward signals in posterior cingulate cortex. Neuron *40*, 1031–1040.

Minassian, A., Paulus, M.P., and Perry, W. (2004). Increased sensitivity to error during decision-making in bipolar disorder patients with acute mania. J. Affect. Disord. *82*, 203–208.

Mirenowicz, J., and Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. J. Neurophysiol. *72*, 1024–1027.

Montague, P.R., and Sejnowski, T.J. (1994). The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms. Learn. Mem. *1*, 1–33.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. *16*, 1936–1947.

Mossin, J. (1966). Equilibrium in a capital asset market. econometrica. J. of the Econometric Society *34*, 768–783.

Niv, Y., Duff, M.O., and Dayan, P. (2005). Dopamine, uncertainty and TD learning. Behav. Brain Funct. *1*, 6.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. Neuron *33*, 815–826.

Real, L.A. (1991). Animal choice behavior and the evolution of cognitive architecture. Science *253*, 980–986.

Romo, R., and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. J. Neurophysiol. *63*, 592–606.

Rustichini, A., Dickhaut, J., Ghirardato, P., Smith, K., and Pardo, J.V. (2005). A brain imaging study of choice procedure. Games Econ. Behav. *52*, 257–282.

Samuelson, W., and Zeckhauser, R. (1988). Status quo bias in decision making. J. Risk Uncertain. *1*, 7–59.

Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. J. of Finance *19*, 425–442.

Shurman, B., Horan, W.P., and Nuechterlein, K.H. (2005). Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the Iowa Gambling Task. Schizophr. Res. *72*, 215–224.

Sutton, R.S. (1988). Learning to predict by the methods of temporal differences. Mach. Learn. *3*, 9–44.

Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2002). Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. J. Neurophysiol. *87*, 508–515.

Tobin, J. (1958). Liquidity preference as behavior towards risk. Rev. Econ. Stud. *25*, 65–86.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. Science *307*, 1642–1645.

Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C., and Berns, G.S. (2004). Human striatal responses to monetary reward depend on saliency. Neuron *42*, 509–517.