

Adaptive coding of reward prediction errors is gated by striatal coupling

Soyoung Q. Park^{a,b,1}, Thorsten Kahnt^{b,c}, Deborah Talmi^d, Jörg Rieskamp^e, Raymond J. Dolan^d, and Hauke R. Heekeren^{a,b,f}

^aDepartment of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany; ^bBerlin School of Mind and Brain, Humboldt Universität zu Berlin, 10117 Berlin, Germany; ^cBernstein Center for Computational Neuroscience, Charité—Universitätsmedizin Berlin, 10115 Berlin, Germany; ^dWellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom; ^eDepartment of Psychology, University of Basel, 4055 Basel, Switzerland; and ^fDahlem Institute for the Neuroimaging of Emotion, 14195 Berlin, Germany

Edited* by Ranulfo Romo, Universidad Nacional Autónoma de México, Mexico City, D.F., Mexico, and approved February 1, 2012 (received for review December 5, 2011)

To efficiently represent all of the possible rewards in the world, dopaminergic midbrain neurons dynamically adapt their coding range to the momentarily available rewards. Specifically, these neurons increase their activity for an outcome that is better than expected and decrease it for an outcome worse than expected, independent of the absolute reward magnitude. Although this adaptive coding is well documented, it remains unknown how this rescaling is implemented. To investigate the adaptive coding of prediction errors and its underlying rescaling process, we used human functional magnetic resonance imaging (fMRI) in combination with a reward prediction task that involved different reward magnitudes. We demonstrate that reward prediction errors in the human striatum are expressed according to an adaptive coding scheme. Strikingly, we show that adaptive coding is gated by changes in effective connectivity between the striatum and other reward-sensitive regions, namely the midbrain and the medial prefrontal cortex. Our results provide evidence that striatal prediction errors are normalized by a magnitude-dependent alteration in the interregional connectivity within the brain's reward system.

adjustment | dopamine | normalization | context invariance | functional connectivity

From receiving a piece of chocolate to winning a lottery, the range of possible rewards in the world is immense, yet the coding range of reward-sensitive neurons is limited. An efficient way for the brain to solve this problem is by dynamically adjusting the activity range of neurons according to the momentarily available rewards. Such an adaptive coding mechanism maximizes the discriminability between different values in a given reward context, thus enabling efficient information processing.

Specifically, adaptive coding of reward prediction errors (PEs) has been suggested by a wide range of theories from economics and reinforcement learning. A PE quantifies the difference between the expected and the actually received reward. Prospect theory suggests that such changes are coded according to an individual reference outcome, such as the status quo or individual expectations (1, 2). In reinforcement learning theory, on the other hand, the PE is considered to be essential for updating the reward values associated with the predicting cue, thus acting as a teaching signal (3, 4). Adaptive coding of PEs is essential for two reasons. First, the reward magnitude (lottery or chocolate) is already encoded during expectation. Hence, in terms of effective neural coding, it is not necessary to represent the reward magnitude redundantly when computing the PEs. Second, to cover the wide range of all possible rewards, it is inevitable to optimally exploit the limited range of neural firing rates. By doing so, the neural system can represent the whole range of rewards and, at the same time, remain sensitive to change.

Indeed, animal recording studies have shown that dopaminergic midbrain neurons encode reward PEs (5, 6) according to an adaptive coding scheme (7). Specifically, these neurons increase their activity for the larger of two potential reward outcomes and decrease their activity for the smaller outcome independent of the

absolute reward magnitude (7). Human studies using functional magnetic resonance imaging (fMRI) highlight PE-related activity in the ventral striatum (8–14), activity that is often presumed to reflect a dopaminergic input from the midbrain.

Adaptive coding is a normalization process that brings different magnitudes onto the same coding scale. Although adaptive coding in reward-sensitive neurons is well documented (7, 15–17), it is unknown how the brain normalizes different ranges of rewards to enable the adaptive coding. One possible mechanism for such normalization is via modulation of connectivity with other reward-coding areas, including areas that might show sensitivity to actual reward magnitudes. The major dopaminergic innervations to the striatum originate in the ventral tegmental area (VTA) and the substantia nigra (SN) (18). Additionally, there is strong input from regions encoding reward value, notably the orbitofrontal cortex (OFC) and the ventromedial prefrontal cortex (vmPFC) (19–24). Based on these connections, we hypothesized that changes in striatal connectivity with these regions would underlie an adjustment in the coding range of striatal PEs. Specifically, when a high reward magnitude is encountered, a dynamic change in connectivity would render striatal PE coding comparable to that of a lower reward magnitude.

In the current study, we aimed to investigate the normalization process underlying adaptive coding by means of fMRI. First, we addressed the question of whether PEs are represented according to an adaptive coding scheme in the human striatum. Second, we investigated how the brain implements the reward rescaling for adaptive coding.

Subjects performed a simple reward prediction task that induces PEs. In each trial, subjects saw a cue indicating the possible reward; in trials with high reward magnitude, subjects saw 1€ combined with either high or low probability (66% or 33%). In trials with low reward magnitude, subjects saw 10ct with high or low probability (Fig. 1A). The reward cues appeared on either the right or the left side of the screen, and subjects were asked to indicate the position of the reward cue by pressing a button. After a variable delay, subjects saw either the corresponding coin, indicating the outcome was obtained, or the coin with a red cross superimposed, indicating the outcome was omitted (Fig. 1A). Importantly, in our task, the reward magnitudes were combined with different probabilities, allowing us to disentangle PE-related activity from outcome-related activity.

Author contributions: S.Q.P., D.T., R.J.D., and H.R.H. designed research; S.Q.P. and T.K. performed research; S.Q.P. and T.K. analyzed data; and S.Q.P., T.K., J.R., R.J.D., and H.R.H. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: soyoung.q.park@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1119969109/-DCSupplemental.

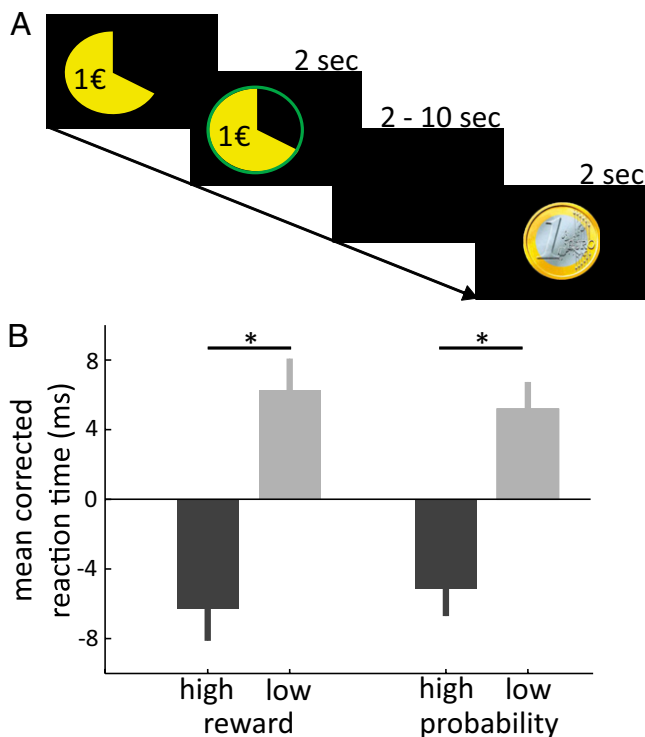


Fig. 1. Task description and behavioral data. (A) In each trial, subjects saw a visual cue indicating the reward magnitude and probability on the left or right side of the screen. Subjects indicated the location of the cue by pressing a button, after which a green circle surrounded the cue. After a variable delay, the reward outcome was shown to the subjects. (B) Mean corrected reaction time data. Left two bars show high and low reward magnitudes, and right two bars show high and low reward probabilities. The greater the reward magnitude and the higher the probability, the faster the subjects responded. (Error bars: SEM.)

Results

Behavioral Results. Subjects correctly indicated the location of the cue in $99.33\% \pm 0.02\%$ of the trials. A two-by-two ANOVA (reward magnitude \times probability) on the reaction time data revealed a significant main effect of magnitude ($F_{1,27} = 13.75$, $P < 0.001$) and a significant main effect of probability ($F_{1,27} = 12.67$, $P < 0.001$). There was no significant magnitude-by-probability interaction ($F_{1,27} = 1.73$, $P = 0.2$). Subjects responded faster in the high compared with low reward magnitude (high, 555 ms; low, 568 ms) and in the high- compared with low-probability trials (high, 557 ms; low, 567 ms), indicating that both reward magnitude and probability affected reward expectations independently (Fig. 1B).

Neuroimaging Results. Our first analysis of imaging data focused on the question of adaptive coding of PEs. We applied a whole-brain general linear model (GLM) that included onset regressors for the reward cue and the outcome as well as two parametric regressors at the time of the reward outcome. The first parametric regressor coded outcome delivery as 1 for received and -1 for omitted rewards, thus accounting for the variance caused by received vs. omitted rewards. The second parametric regressor accounted for the PE-related variance, and, importantly, this regressor was orthogonalized with respect to the first parametric regressor (received vs. omitted). Hence, this PE regressor accounts for variance in the brain oxygen level-dependent (BOLD) signal that is independent of outcome-related (received vs. omitted) activity. Voxel-wise one-sample t tests on the parameter estimates of the parametric PE regressor revealed a significant correlation

between PE and activity in the ventral striatum [Fig. 2A; $P < 0.05$, family-wise error (FWE) small volume-corrected (SVC), $[-6, 18, -3]$; $t_{27} = 4.19$; see Table S1 for whole-brain results].

After having identified the region in ventral striatum coding reward PEs, we determined whether this striatal region adaptively codes PE responses. Specifically, we tested whether the striatal PE responses were invariant for different reward magnitudes. In case where PEs are coded according to an adaptive coding scheme, representations of striatal PEs for high reward should not differ from those of low reward. We applied a GLM to the striatal data, but this time all of the regressors were split for high- and low-reward trials. The onset regressors for cue and outcome, the two parametric regressors for outcome-related variance, and the two PE coding regressors were regressed against the BOLD signal in the striatum. Both PE regressors were orthogonalized with respect to the parametric modulation of reward outcome. This comparison of PE-related responses in high vs. low reward magnitudes revealed no significant difference in the striatum (Fig. 2B; $t_{27} = -0.98$, $P = 0.34$).

We further performed an additional analysis in which we directly tested whether the striatal changes in BOLD signal are significantly better predicted by an adaptive PE (PE modulated only by probability) than by a nonadaptive PE (PE modulated by probability \times magnitude). If an adaptive PE predicts striatal activity better than a nonadaptive PE, it would support adaptive PE coding in the striatum. In a case where PEs are modulated by reward magnitude, then a nonadaptive PE would provide a better prediction of striatal activity. Importantly, our analysis shows that the striatal BOLD signal is significantly better predicted by an adaptive compared with a nonadaptive PE (Fig. 2C; $t_{27} = 2.82$, $P = 0.0089$; mean parameter estimates: adaptive PE, 0.37 ± 0.014 ; nonadaptive PE, 0.29 ± 0.018). This invariance in the representation of striatal PEs is consistent with an adaptive coding scheme shown in dopaminergic neurons in primates (7).

Having shown that PEs are adaptively coded in the striatum, our next analysis sought to determine whether there were dynamic changes in striatal coupling associated with this magnitude-dependent rescaling. We hypothesized that reward-sensitive brain regions with striatal innervations, specifically the vmPFC and the VTA/SN, would modulate striatal activity as a function of reward magnitude. We performed a whole-brain psychophysiological interaction (PPI) analysis where the striatal time series of PE-related activity (Fig. 3A) was selected as a physiological variable and the reward magnitude was selected as a psychological variable (high vs. low reward). Comparing striatal connectivity between reward

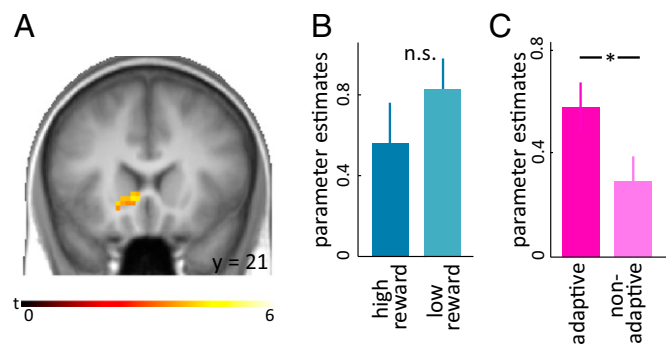


Fig. 2. Adaptive coding of PEs in the striatum. (A) Parametric modulation with trial-wise PE revealed significant correlation with BOLD responses in the striatum. (B) Across different reward magnitudes, no significant difference in PE-related responses was observed ($t_{27} = -0.98$, $P = 0.34$), confirming an adaptive coding of striatal PE. (C) Adaptive PE predicted striatal BOLD responses significantly better than nonadaptive PE did ($t_{27} = 2.82$, $P = 0.0089$). (B and C) The y axes represent the parameter estimates of PE. (Error bars: SEM.)

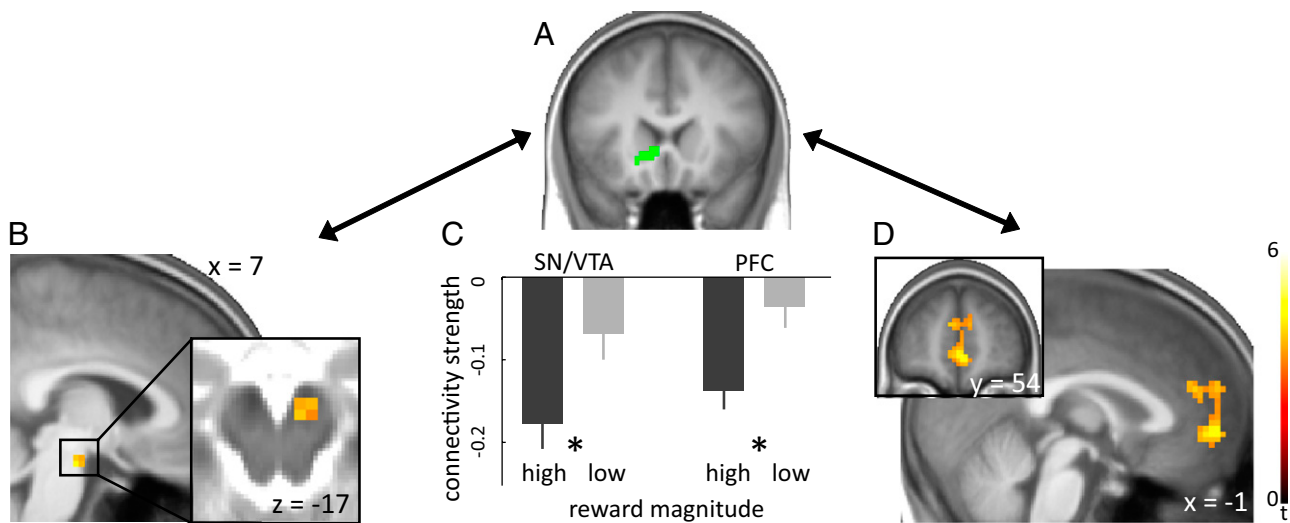


Fig. 3. Striatal coupling gates adaptive coding of PEs. (A) The striatum, showing adaptive coding of PEs, was used as seed region in the functional connectivity analysis. (B and D) VTA/SN (B) and vmPFC (D) showing significant magnitude-dependent connectivity modulation with the striatum. (B Inset) Activity superimposed on a T2-weighted image. (D Inset) A coronal view. (C) Bar graph depicts significantly less midbrain–striatal and fronto–striatal connectivity during high compared with low reward magnitudes. (Error bars: SEM.)

magnitudes revealed significant modulations in the VTA/SN ([6, -8, -21]; $P < 0.05$, FWE SVC, $t_{27} = 4.74$; Fig. 3B) and the vmPFC ([3, 54, -3], BA10/BA9; $P < 0.05$, FWE SVC, $t_{27} = 4.85$; Fig. 3D, and see Table S2 for the whole-brain results). Specifically, the coupling was significantly less during the high compared with the low reward magnitude between the striatum and both VTA/SN and vmPFC, respectively (Fig. 3C).

Discussion

From visual neurons (25, 26) to reward-coding dopaminergic neurons, scale invariance is a ubiquitous encoding principle in the brain. Retinal neurons rapidly adapt to an enormous range of light to guarantee high visual discriminability. Dunn et al. (27) have shown that this adaptation is implemented via a relay from cone bipolar cells to ganglion cells, demonstrating that such a rapid rescaling of range occurs via influences of innervating neurons. Analogously, we demonstrate that PE-related activity in the human striatum adapts to the momentarily available reward magnitude, and this effect is driven by changes in neuronal dynamics associated with different reward magnitudes.

Adaptive coding of PEs enables an enhanced discriminability to remain sensitive for changes of all sizes, which optimizes efficient coding in neural circuits with given limitations, namely, its firing ranges (28). Previous animal studies have provided evidence of neuronal adaptation in coding different aspects of reward. Whereas midbrain dopaminergic neurons adaptively code reward PEs (7), OFC neurons show adaptive coding of reward preference (15). OFC neurons also adapt their firing range according to the momentarily available reward range and distribution of rewards (16, 17). Furthermore, recently it has been shown that the primate lateral intraparietal cortex represents saccade values depending on other available choice options. This context dependency is precisely predicted by the divisive normalization mechanism (29). In humans, fMRI studies have also shown that BOLD signals in reward-sensitive areas reveal magnitude adaptation across possible rewards (14, 30–33).

In our data, we observe significantly less connectivity between the striatum and the mPFC/midbrain in high-reward compared with low-reward conditions (Fig. 3C). In the PPI analysis, non-specific correlations across the brain were removed by regressing out the global mean from every voxel. However, this global-mean normalization shifts the correlation distribution to have a mean

near zero and forces negative correlations to appear (34–36). Therefore, it is important to interpret only the difference in connectivity between task conditions.

Our PPI results are in accord with previous studies investigating the dynamics of neuronal activation in this anatomic network. The primate striatum is tightly interconnected with the midbrain as well as with cortical areas (18, 37–39). Specifically, the striatum receives dopaminergic input from midbrain regions, creating an ascending midbrain–striatal loop (13, 40). Furthermore, PFC activation modulates striatal dopamine release via inhibitory midbrain neurons (41). More specifically, PFC neurons activate GABAergic cells in the midbrain that in turn inhibit neighboring dopaminergic neurons projecting to the striatum (42). Thus, one possible pathway underlying our connectivity result is that high magnitudes of reward activate PFC neurons, thereby increasing midbrain GABA inhibition, which in turn results in reduced dopamine release in the striatum (42). This pathway may underlie the adjustment of PE signals during high-reward outcomes and explains the relative decrease in connectivity between the mPFC/midbrain and the striatum in high vs. low magnitudes.

Although our connectivity result is in line with the neuroanatomical and neurochemical systems, we acknowledge that it is not possible to identify the underlying neurotransmitters by using functional connectivity analyses with BOLD responses. Future studies are needed to investigate the neurochemical nature of this mechanism by means of positron emission tomography or pharmacological interventions.

When a specific reward magnitude is expected, dopaminergic neurons are not sensitive to the magnitude when coding the PE (7). Analogously, in the present study, all rewards were cued to ensure that rewards are expected before the outcome was presented. Also, we show that the striatal PE was invariant for high and low reward magnitudes, which is in line with other studies showing that PEs are insensitive to different reward magnitudes when cues signal the possible reward magnitude (7, 14, 32, 43). In contrast, once a reward outcome is presented unexpectedly, activity in the striatum is different for high vs. low reward magnitudes (44). Our connectivity results suggest that, with cued magnitude, the context-dependent regulation of striatal activity takes place via striatal coupling with the PFC and the midbrain. However, in the absence of a cue, when the outcome is delivered unexpectedly, no magnitude-

dependent adaptation can occur, and the striatal activity might remain sensitive to different reward magnitudes.

Altogether, our results provide evidence that adaptive coding of PEs in humans is gated by striatal coupling. In line with recordings from primate dopamine neurons (7), our results show that striatal PEs do not differ for high and low reward magnitudes. Using an effective connectivity analysis, we show that mPFC and midbrain significantly modulate their coupling with the striatum in the face of a high reward magnitude, rendering striatal PEs comparable to the lower reward magnitude. Our results suggest an elaborate neural mechanism that facilitates the accurate representation of the reward value. The adaptive coding of PEs enables the brain to dynamically allocate its limited neural firing range to better discriminate among momentarily available reward outcomes. This mechanism provides a rich framework with which mechanistic hypotheses for pathologies that include abnormal value processing, such as drug addiction, can be examined.

Materials and Methods

Experimental Design. In each trial of the task, a reward-predicting cue was presented on either the left or right side of fixation. Subjects were asked to press the corresponding button on a response box as fast as possible. Each cue contained information about both the probability (67% or 33%) and the magnitude (1€ or 10ct) of the possible reward. Accordingly, there were four different cues, indicating high probability of 1€, low probability of 1€, high probability of 10ct, and low probability of 10ct (Fig. 1A). After the button press, a green circle highlighted the stimulus, and after a variable interstimulus interval, the outcome cue was presented for 2 s. In high-reward trials, the outcome was an image of a 1€ coin (for received reward) or a red cross superimposed on the 1€ coin (for omitted reward). Analogously, in low-reward trials, subjects saw either a 10ct coin in received trials or a red cross superimposed on the 10ct coin in omitted trials. The actual reception or omission of reward was determined by the probability indicated by the cue. Subjects performed five sessions with 60 trials each. A total of 33 healthy subjects were tested in the study. Five subjects were excluded from the sample [one was removed because of extreme head movement (more than 3 mm or 3°) during scanning, three subjects aborted the scanning because they felt sick, and one subject was left-handed], resulting in a final sample size of $n = 28$ (13 females and 15 males; mean \pm SD age, 25.04 \pm 2.5 y).

Behavioral Data Analysis. To monitor subjects' attention to the task, we analyzed the percentage of correct responses (indicating the location of the cue). All following analyses included only the correctly responded trials. We examined whether subjects established reward expectations during the reward-predicting cue by testing whether reaction times are influenced by both reward magnitude and probability. For this test, we computed a two-by-two (magnitude \times probability) ANOVA with repeated measures on reaction time data.

fMRI Data Acquisition and Preprocessing. Functional imaging was conducted on a 3-T Siemens Trio scanner with a 12-channel head coil. In each of the five runs, 366 T2*-weighted gradient-echo echoplanar images containing 37 slices (3 mm thick) separated by a gap of 0.75 mm were acquired. Imaging parameters were as follows: repetition time (TR), 2,000 ms; echo time (TE), 30 ms; flip angle, 70°; matrix size, 64 \times 64; field of view, 192 mm; and voxel size, 3 \times 3 \times 3.75 mm. T1-weighted and T2-weighted structural datasets were collected for the purpose of anatomical localization. The parameters for the T1-weighted dataset were as follows: TR, 1,900 ms; TE, 2.52 ms; matrix size, 256 \times 256; field of view, 256 mm; 176 slices (1 mm thick); and flip angle, 9°. The parameters for the T2-weighted dataset were as follows: TR, 8,170 ms; TE, 0.93 ms; matrix size, 256 \times 256; field of view, 256 mm; 48 slices (3 mm thick); and flip angle, 120°.

Functional data were analyzed with SPM5 (Wellcome Department of Imaging Neuroscience, University College London Institute of Neurology, London, UK). Images were slice timing-corrected, realigned, spatially nor-

malized to a standard echoplanar image template of the Montreal Neurological Institute (MNI), resampled to 3-mm isotropic voxels, and spatially smoothed with an 8-mm full width at half maximum Gaussian kernel. All included subjects moved less than the size of a single voxel (3 mm).

Model-Based fMRI Data Analysis. We computed a GLM with a parametric design (45) to identify brain regions coding PEs in an adaptive fashion. In each trial t , the PE δ was defined as

$$\delta = r_t - p_t \quad [1]$$

where r_t is the reward outcome (1 for received and 0 for omitted outcomes) and p_t is the expected probability of the reward (0.33 or 0.66). Note that the task is not a learning task because the reward probability and magnitudes were explicitly shown and did not change over the experiment. Four regressors were included in the GLM in the following order: (i) onset of the cue, (ii) onset of the outcome, (iii) parametric modulation of the outcome (coded as 1 when received and -1 when omitted), and (iv) parametric modulation of the trial-wise PEs. The PE regressor was created by parametrically modulating the stimulus function of the outcome by the normalized (mean = 0, SD = 1) trial-wise PEs. Importantly, the PE regressor (the fourth regressor) was orthogonalized with respect to the outcome-related (received vs. omitted) parametric regressor (the third regressor). All regressors were convolved with a canonical hemodynamic response function. Individual contrast images were computed for PE-related responses and taken to a second-level mixed-effect analysis using voxel-wise one-sample t tests.

Reward Magnitude-Dependent Changes in Striatal Connectivity. We performed a whole-brain PPI analysis (12, 13, 46) with the striatum as a seed region. Here, the entire time series over the experiment was extracted from each subject in the clusters of the striatum, in which activity significantly correlated with PE on the group level. To create the PPI regressor, we multiplied the normalized time series with two condition vectors containing ones for six TRs after each reward-magnitude type (one regressor for high and one for low magnitudes) and zeros otherwise. The method used here relies on correlations in the observed BOLD time-series data and makes no assumptions about the nature of the neural event contributing to the BOLD signal (13). The time window of six TRs (12 s) was selected to capture the entire hemodynamic response function, which peaks after three TRs and is back at baseline at approximately eight TRs after stimulus onset. These PPI regressors were used as covariates in a separate PPI-GLM, in which the following regressors were included: (i) cue onset, (ii) psychological regressor accounting for high-reward outcome, (iii) psychological regressor accounting for low-reward outcomes, (iv) physiological regressor (i.e., the entire time series of the seed region over the whole experiment), (v) the PPI regressor for high-reward outcomes, and (vi) the PPI regressor for low-reward outcomes. The onset regressors were convolved with an hemodynamic response function. The resulting parameter estimates of the two PPI regressors represent the extent to which activity in each voxel correlates with activity in the striatum for each condition. Individual contrast images for functional connectivity during high vs. low reward magnitude were then computed and entered into a one-sample t test. We then identified voxels with significant connectivity difference during low vs. high reward magnitude.

We applied an omnibus threshold for all whole-brain analyses of $P < 0.001$, uncorrected with a cluster extent threshold of $k = 10$ (whole-brain results are shown in Tables S1 and S2). Correction for multiple comparisons ($P < 0.05$, FWE correction) was then performed for clusters surviving this threshold by using 12-mm spheres around previously reported peak voxels (SVC): for the ventral striatum and midbrain, $[-8, 8, -4]$ and $[8, -18, -20]$, respectively (6); for mPFC, $[3, 54, -15]$ (25). All reported coordinates (x, y, z) are in MNI space.

ACKNOWLEDGMENTS. This work was supported by Excellence Initiative of the German Federal Ministry of Education and Research, Deutsche Forschungsgemeinschaft, Grants GSC86/1-2009 and EXC 302, the Max Planck Society, and Swiss National Science Foundation Grant 100014_130352.

1. Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–1131.
2. Kahneman D, Tversky A (1979) Prospect theory: Analysis of decision under risk. *Econometrica* 47(2):263–291.
3. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton-Century-Crofts, New York), pp 64–99.
4. Sutton R, Barto A (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
5. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
6. Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47(1):129–141.
7. Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645.

8. 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.
9. Pagnoni G, Zink CF, Montague PR, Berns GS (2002) Activity in human ventral striatum locked to errors of reward prediction. *Nat Neurosci* 5(2):97–98.
10. McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
11. 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.
12. Park SQ, et al. (2010) Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. *J Neurosci* 30:7749–7753.
13. Kahnt T, et al. (2009) Dorsal striatal–midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *J Cogn Neurosci* 21:1332–1345.
14. Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30:619–639.
15. Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.
16. Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. *J Neurosci* 29:14004–14014.
17. Kobayashi S, Pinto de Carvalho O, Schultz W (2010) Adaptation of reward sensitivity in orbitofrontal neurons. *J Neurosci* 30:534–544.
18. Haber SN, McFarland NR (1999) The concept of the ventral striatum in nonhuman primates. *Ann N Y Acad Sci* 877:33–48.
19. Kahnt T, Heinzele J, Park SQ, Haynes JD (2010) The neural code of reward anticipation in human orbitofrontal cortex. *Proc Natl Acad Sci USA* 107:6010–6015.
20. Philiastides MG, Biele G, Heekeren HR (2010) A mechanistic account of value computation in the human brain. *Proc Natl Acad Sci USA* 107:9430–9435.
21. Ferry AT, Ongür D, An X, Price JL (2000) Prefrontal cortical projections to the striatum in macaque monkeys: Evidence for an organization related to prefrontal networks. *J Comp Neurol* 425:447–470.
22. Ongür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10(3):206–219.
23. Park SQ, Kahnt T, Rieskamp J, Heekeren HR (2011) Neurobiology of value integration: When value impacts valuation. *J Neurosci* 31:9307–9314.
24. Kahnt T, Heinzele J, Park SQ, Haynes JD (2011) Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage* 56:709–715.
25. Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. *Nature* 412:787–792.
26. Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. *Neuron* 26:695–702.
27. Dunn FA, Lankheet MJ, Rieke F (2007) Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature* 449:603–606.
28. Preuschoff K, Bossaerts P (2007) Adding prediction risk to the theory of reward learning. *Ann N Y Acad Sci* 1104:135–146.
29. Louie K, Grattan LE, Glimcher PW (2011) Reward value-based gain control: Divisive normalization in parietal cortex. *J Neurosci* 31:10627–10639.
30. Elliott R, Agnew Z, Deakin JF (2008) Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. *Eur J Neurosci* 27:2213–2218.
31. Nieuwenhuis S, et al. (2005) Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage* 25:1302–1309.
32. Bunzeck N, Dayan P, Dolan RJ, Duzel E (2010) A common mechanism for adaptive scaling of reward and novelty. *Hum Brain Mapp* 31:1380–1394.
33. Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci USA* 104:9493–9498.
34. Van Dijk KR, et al. (2010) Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *J Neurophysiol* 103:297–321.
35. Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA (2009) The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *Neuroimage* 44:893–905.
36. Fox MD, Zhang D, Snyder AZ, Raichle ME (2009) The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol* 101:3270–3283.
37. Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci* 15:4851–4867.
38. Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20:2369–2382.
39. Haber SN, Knutson B (2010) The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4–26.
40. Ikemoto S (2007) Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens–olfactory tubercle complex. *Brain Res Brain Res Rev* 56(1):27–78.
41. Frankle WG, Laruelle M, Haber SN (2006) Prefrontal cortical projections to the mid-brain in primates: Evidence for a sparse connection. *Neuropsychopharmacology* 31:1627–1636.
42. Karreman M, Moghaddam B (1996) The prefrontal cortex regulates the basal release of dopamine in the limbic striatum: An effect mediated by ventral tegmental area. *J Neurochem* 66:589–598.
43. Yacubian J, et al. (2006) Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci* 26:9530–9537.
44. Delgado MR, Locke HM, Stenger VA, Fiez JA (2003) Dorsal striatum responses to reward and punishment: Effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci* 3(1):27–38.
45. Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. *Neuroimage* 8(2):140–148.
46. Friston KJ, et al. (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6(3):218–229.

Supporting Information

Park et al. 10.1073/pnas.1119969109

Table S1. Significant whole-brain correlation with trial-wise prediction error (PE) ($P < 0.001$, uncorrected; $k = 10$)

Hemisphere	Region name	Brodmann area	MNI (x, y, z)*			t
Right	Ventral striatum		3	15	-6	4.86
Right	Inferior frontal gyrus	44	42	12	18	4.37
Left	Inferior frontal gyrus	44	-45	12	21	4.01
Left	Precentral gyrus	6	-39	0	60	3.66
Left	Precuneus		-3	-51	48	4.34
Left	Lingual gyrus	19	-21	-60	-3	4.10
Right	Lingual gyrus	19	15	-48	-6	4.17
Left	Temporal lobe	37	-36	-66	6	4.67
Right	Temporal lobe	22	63	-42	15	3.96
Left	Occipital lobe	18	-6	-87	-9	5.03
Right	Occipital lobe	19	6	-87	33	4.37
Left	Cerebellum		-39	-48	-30	5.15
Right	Cerebellum		39	-42	-30	6.52

*Coordinates are in Montreal Neurological Institute (MNI) space.

Table S2. Significant whole-brain functional connectivity modulation with striatum in the comparison of low- vs. high-reward context ($P < 0.001$, uncorrected; $k = 10$)

Hemisphere	Region name	Brodmann area	MNI (x, y, z)*			t
Right	Dorsomedial PFC	32	12	48	21	4.92
Left	Dorsomedial PFC	10	-6	57	21	4.35
Right	Medial OFC	10	3	54	-3	4.85
Left	Mid-OFC	11	-24	36	-12	4.09
Right	Midbrain		6	-18	-21	4.74
Right	Dorsolateral PFC	46	30	36	27	4.54
Right	Parietal cortex	40	51	-51	27	4.40
Right	Temporal lobe	21/37	60	-57	-6	4.08
Left	Occipital lobe	19	-36	-78	24	3.95

No single voxel survived the threshold for the opposite comparison of high- vs. low-reward context. OFC, orbitofrontal cortex; PFC, prefrontal cortex.

*Coordinates are in MNI space.