

Reward-Related Activity in Ventral Striatum Is Action Contingent and Modulated by Behavioral Relevance

Thomas H. B. FitzGerald,^{1*} Philipp Schwartenbeck,^{1,2,3*} and Raymond J. Dolan¹

¹The Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom, ²Centre for Neurocognitive Research and Department of Psychology, University of Salzburg, 5020 Salzburg, Austria, and ³Neuroscience Institute and Centre for Neurocognitive Research, Christian-Doppler-Klinik, Paracelsus Medical University Salzburg, 5020 Salzburg, Austria

Multiple features of the environment are often imbued with motivational significance, and the relative importance of these can change across contexts. The ability to flexibly adjust evaluative processes so that currently important features of the environment alone drive behavior is critical to adaptive routines. We know relatively little about the neural mechanisms involved, including whether motivationally significant features are obligatorily evaluated or whether current relevance gates access to value-sensitive regions. We addressed these questions using functional magnetic resonance imaging data and a task design where human subjects had to choose whether to accept or reject an offer indicated by visual and auditory stimuli. By manipulating, on a trial-by-trial basis, which stimulus determined the value of the offer, we show choice activity in the ventral striatum solely reflects the value of the currently relevant stimulus, consistent with a model wherein behavioral relevance modulates the impact of sensory stimuli on value processing. Choice outcome signals in this same region covaried positively with wins on accept trials, and negatively with wins on reject trials, consistent with striatal activity at feedback reflecting correctness of response rather than reward processing per se. We conclude that ventral striatum activity during decision making is dynamically modulated by behavioral context, indexed here by task relevance and action selection.

Key words: action value; multisensory; policy selection; reward; ventral striatum

Introduction

In a laboratory environment it is common that reward contingencies depend upon a single feature of the environment (for example, the pitch of a tone). In more ecological contexts multiple features of the environment potentially signal reward, and the relative importance of these features varies according to context. Despite its importance, the mechanisms by which task relevance modulate behavior are poorly understood (Wilson and Niv, 2011). One possibility, suggested by theoretical models, is that behavioral relevance is controlled by selective attention (Dayan et al., 2000; Gershman et al., 2010). Here a single value prediction is generated, based upon a combination of stimulus features weighted by their behavioral importance, and this drives behavior (Fig. 1). Alternatively, different stimulus features might be automatically evaluated (Pessiglione et al., 2008), with a relevance modulation reflected in the strength of the connections between value representations and effector regions (Fig. 1). Func-

tional neuroimaging can enable a discrimination between the above accounts. If relevance modulates the influence of stimulus features on reward representations, then only the value of relevant features should be signaled in reward-related areas, such as the ventral striatum (Schultz et al., 1992; Knutson et al., 2001; O'Doherty et al., 2004; Kable and Glimcher, 2007). If, in contrast, behavioral relevance exerts an effect on the links between reward signaling and effector areas, then the value of all stimulus features should be simultaneously represented.

We devised a paradigm in which subjects were presented simultaneously with one of two visual and auditory cues. The (explicitly signaled) reward contingencies for each trial were determined either by the visual cue, the auditory cue, or a combination of the two (the “cross-modal” condition). The relevant cue indicated whether the subject was presented with either a “good” or a “bad” offer, and subjects then indicated a choice whether to accept this or not. Feedback was provided on each trial, but only if the subject accepted the offer did outcomes contribute to their winnings. Using this design we were able to test a hypothesis that activity in ventral striatum reflects the value only of behaviorally relevant stimuli, consistent with an effect of relevance on evaluation processes themselves (Dayan et al., 2000).

In addition, our task allowed us to test whether outcome signals in the ventral striatum indexed an updating of action policies (Klein-Flügge et al., 2011; Li and Daw, 2011), or rather an updating of the value assigned to particular actions (Watkins and Dayan, 1992). This is important for understanding how human subjects learn (Dayan and Daw, 2008; Friston et al., 2009). Crit-

Received Oct. 14, 2013; revised Nov. 25, 2013; accepted Dec. 5, 2013.

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

This work was supported by a Wellcome Trust Senior Investigator Award 098362/Z/12/Z to R.J.D. and The Wellcome Trust Centre for Neuroimaging is supported by core funding from Wellcome Trust Grant 091593/Z/10/Z. We thank the Functional Imaging Laboratory radiographers for their time and patience.

*T.H.B.F. and P.S. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Thomas H.B. FitzGerald, Wellcome Trust Centre for Neuroimaging, 12 Queen Square, London WC1N 3BG, UK. E-mail: thomas.fitzgerald@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.4389-13.2014

Copyright © 2014 the authors 0270-6474/14/341271-09\$15.00/0

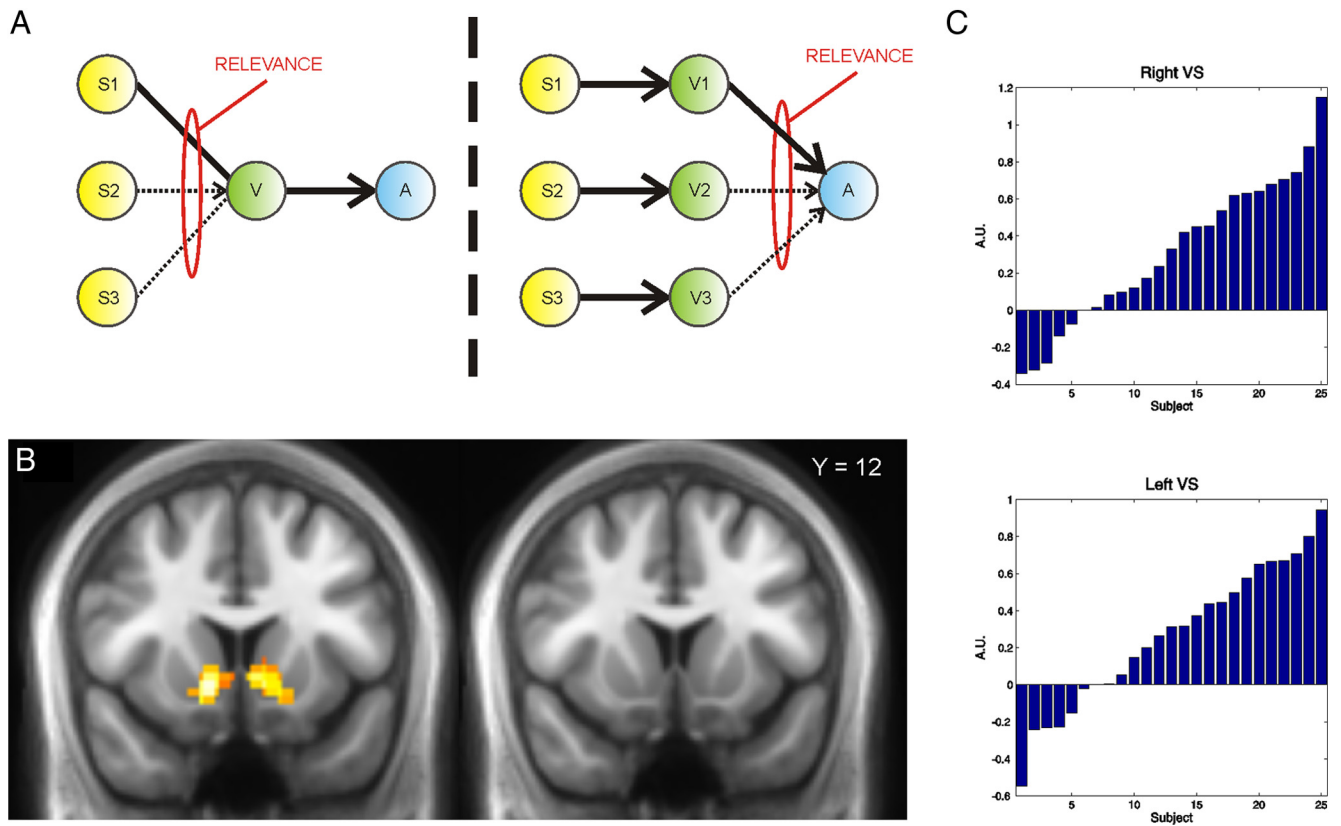


Figure 1. *A*, Diagram illustrating different models for the effect of task relevance on value processing. Task relevance can modulate the input of sensory signals (S1–S3) to value-processing regions, leading to a single relevance-modulated value signal (V) that is then used for action selection and planning (A; left). Alternatively, task relevance can modulate the output of value-processing regions, which predicts the simultaneous representation of the value of all motivationally valenced stimulus features (right). *B*, Activity in bilateral ventral striatum reflects the value of behaviorally relevant features of the environment (left) but not behaviorally irrelevant ones (right). The difference between these contrasts (the interaction effect) was significant in both hemispheres. This supports models where task relevance modulates inputs to value-processing regions, leading to the generation of a single relevance-modulated value prediction, as illustrated in the diagram on the left side of Figure 2*A*. (Image thresholded at $p < 0.005$ uncorrected for display purposes. Color bar indicates the voxelwise T statistic; $Y = 11$ mm.) *C*, Single-subject parameter estimates for the difference between responses to behaviorally relevant and irrelevant value, averaged across the whole of the right (top) and left (bottom) anatomically defined ventral striatum ROIs (positive values reflect greater activation for relevant than irrelevant value). Significantly greater activation was observed to relevant than irrelevant value.

ically, the distinct accounts outlined above make opposite predictions about trials where subjects choose to reject an offer. If striatal signals reflect a direct updating of policies, as previously suggested (Li and Daw, 2011), then a foregone win represents a mistaken action and we would then expect to see a negative response in ventral striatum. If, on the other hand, striatal activity reflects an updating of action (or stimulus) values themselves then this predicts foregone rewards should be associated with a positive signal.

Materials and Methods

Subjects. Twenty-five (17 female) right-handed subjects, age range 19–48, all free of psychiatric or neurological disease, participated in the study. The study was approved by the Joint National Hospital for Neurology and Neurosurgery (University College London Hospitals NHS trust) and Institute of Neurology (University College London) Ethics Committee. Subjects were paid according to their performance during the task (receiving from £17.40–39.80).

Stimuli and task. On each trial of the experiment, subjects were asked to decide either to accept or reject an offer made to them (Fig. 2). They were instructed that whatever they decided, the outcome of the trial (either a win or a loss) would be shown, but only if they had chosen to accept the offer would it impact on their winnings in the session. The task comprised three experimental conditions: a “visual” condition in which offer value was determined solely by the visual cue (one of two colored boxes), an “auditory” condition in which it was determined solely by the auditory cue (one of two short clips of synthesizer pads), and a “cross-modal” condition where the combination of the auditory and visual cues

dictated the offer value. Each trial consisted of a concurrent presentation of both auditory and visual cues. Stimuli were thus identical between conditions, and the individual trials differed only according to which stimulus features were behaviorally relevant for that trial. Conditions were presented in a pseudorandomized order, and were explicitly signaled to the subject by text presented for 1500 ms before stimulus presentation (Fig. 2).

Offers could be either good or bad, with good offers indicating an 80% win probability and a 20% probability of loss, and vice versa for bad offers. Two visual cues were used, one of which was fixed as the good cue throughout the experiment, with a similar arrangement for the auditory cues. This meant that there were four possible cue combinations in the cross-modal condition. For each subject we specified that good offers in the cross-modal condition were indicated either by congruent stimuli (both cues good or both cues bad) or incongruent stimuli, and this arrangement was counterbalanced across subjects so as to decorrelate the effects of value and congruence in the cross-modal condition. In subjects for whom congruent stimuli indicated good offers, the text for the cross-modal condition read “Congruent,” while for the other subjects it read “Incongruent.”

Stimuli were presented for 2000 ms (Fig. 2) and subjects were instructed to make an accept or reject response via a functional magnetic resonance imaging (fMRI)-compatible button box (response keys were counterbalanced across subjects). Outcomes (text reading “Win” or “Lose”) were presented visually for 1200 ms after a delay, which varied between 2000 and 8000 ms. The outcome of each trial was presented regardless of whether subjects chose to accept or reject an offer. In trials where they had chosen accept, a bar at the bottom of the screen indicating their cumulative earnings either increased or decreased in length by equal

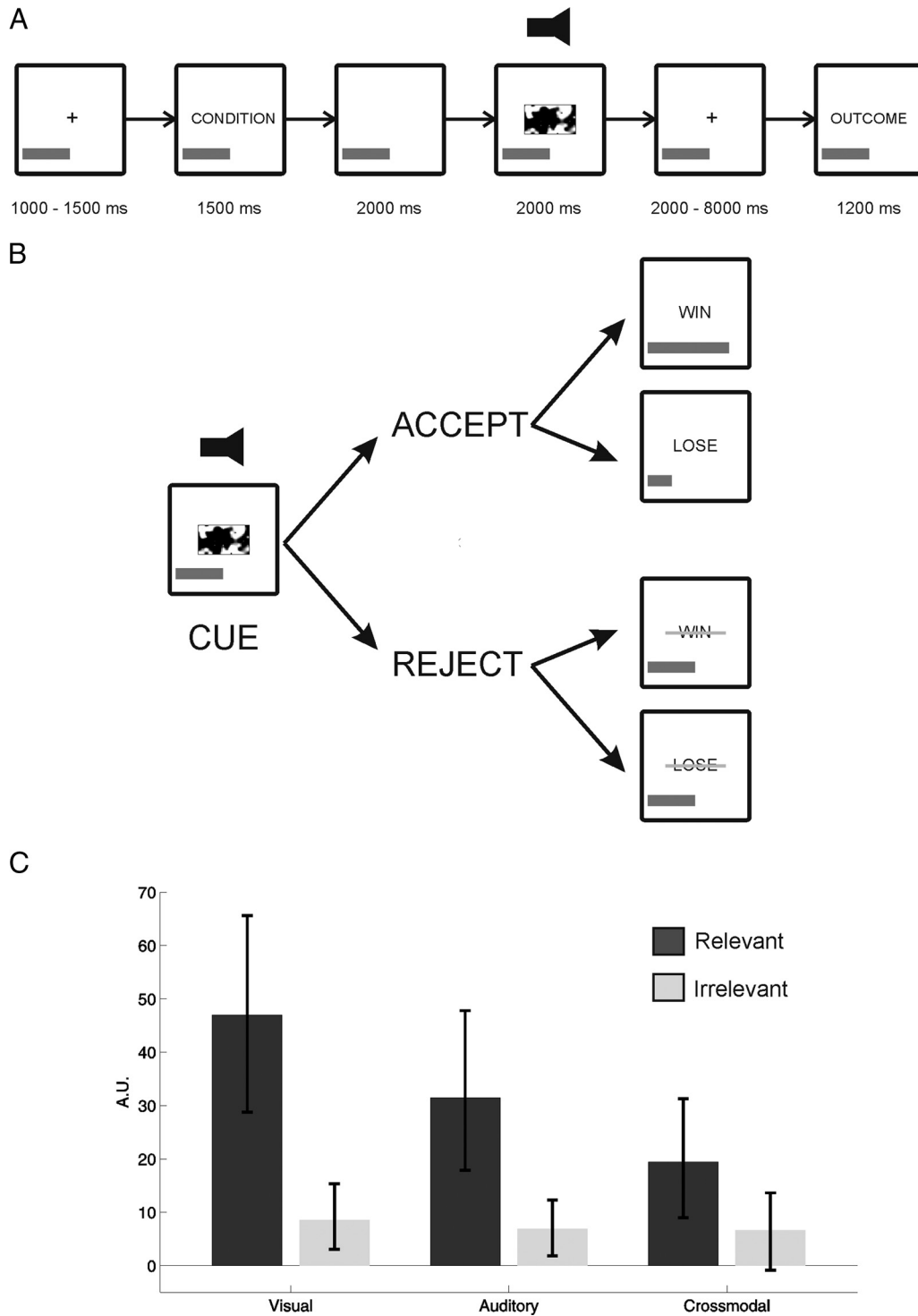


Figure 2. *A*, Time course of a single trial. After a jittered fixation interval (1000–1500 ms), subjects were presented with text indicating which properties of stimuli determined the reward contingencies on that trial (de facto which “condition” they were in). Thus, the text said “Visual” for trials where the visual cue determined the outcome of the trial, “Auditory” for trials where the auditory cue determined the outcome, and either “Congruent” or “Incongruent” where the combination of both determined the outcome. Each subject saw only one of Congruent or Incongruent, and these were counterbalanced across subjects to decorrelate value and congruence. Following a brief (2000 ms) interval, one of two patterned boxes and one of two synthesizer pads were presented to the subject for 2000 ms, who then chose to accept or reject the offer before its offset. Outcomes were presented visually for 1200 ms after a variable delay (between 2000 and 8000 ms). The subject’s current winnings during the session were indicated by the length of a bar constantly displayed at the bottom of the screen. *B*, Illustrates possible outcomes. Subjects were shown the outcome of the trial (indicated by text saying either “WIN” or “LOSE”) and whether they chose to accept or reject an offer made to them. If they chose to accept the offer, the bar at the bottom of the screen indicated whether their cumulative earnings during the session increased or decreased in length by equal amounts according to whether they won or lost. If they chose to reject the offer in contrast, the text was presented with a line through it to indicate that this choice did not affect their earnings and the cumulative earnings bar did not change in length. *C*, Behavioral results. Logistic regression analysis demonstrated that in all three task conditions, behaviorally relevant stimulus properties (dark gray) had a stronger effect on behavior than irrelevant ones (light gray). Error bars indicate bootstrapped 90% confidence intervals.

amounts according to whether they won or lost. If they had chosen to reject the text was presented with a line through it indicating that it did not impact on their earnings, and accordingly the earnings bar did not change in length. At the start of each session the indicator bar started at a length corresponding to winnings of £3.60. This was implemented to ensure losses were meaningful even in early trials of the session. On trials where subjects failed to make either an accept or reject response, the words “Too Slow” were displayed, and subjects lost an amount equal to trials where they accepted an offer and a lose outcome occurred. Note that these latter trials (1.4%) were excluded from our behavioral and neuroimaging analyses.

Before scanning, subjects were trained on the value of the auditory and visual stimuli separately using a simple instrumental learning procedure in which auditory and visual stimuli were presented alone in separate blocks of 24 trials. As in the main task, subjects chose to accept or reject offers indicated by the stimulus present on each trial, and feedback was presented in relation to the outcome of both accepted and rejected gambles. They then underwent at least one training session consisting of 60 trials of the task proper (altered to reduce the gap between action selection and offer presentation to a maximum of 2500 ms; the length of training varied between subjects according to their speed of learning). During scanning subjects performed two sessions each consisting of 120 trials. We present behavioral data acquired during scanning alone.

Behavioral analysis. To check that subjects were able to adequately acquire and maintain the reward contingencies during the task, we analyzed the mean probability of selecting the correct action (accepting on trials when a good stimulus was presented and rejecting on trials where a bad one was presented) in each of the three conditions. We compared mean accuracy rates between conditions by taking the mean rates for each subject as summary statistics, testing for differences using a two-tailed Wilcoxon signed rank test.

To test whether the valence of stimuli in task-irrelevant modalities affected behavioral responding, we performed a logistic regression for each subject with separate regressors for the valences of task-relevant and task-irrelevant stimuli in each of the three conditions (giving us a total of six regressors). These were used to predict whether subjects accepted the offer (or not) on each trial (as reflected by a positive regression coefficient). We then performed group-level statistics using the single subject regression coefficients and one-tailed signed rank tests. This reflected our strong prior hypothesis that the presence of positively valenced stimuli should make subjects more likely to accept on any given trial.

fMRI data acquisition and preprocessing. Gradient-echo T2*-weighted echo-planar (EPI) images were acquired on a 3 T Trio Siemens scanner with a resolution of 3 mm isotropic. Scanner settings (echo time, 30 ms; repetition time, 3.36 s; 48 slices acquired in descending order at an angle of 30° in the anterior–posterior axis) were designed to optimize sensitivity in the orbital frontal cortex (Deichmann et al., 2003). In each session, at least 469 images were collected (~27 min each, two per subject). The first five images from the task sessions were discarded to allow for T1 equilibration effects, and the fMRI time series realigned and unwrapped to correct for both motion-related and static distortions (Hutton et al., 2002). Whole-brain 1 mm × 1 mm × 1 mm T1-weighted structural images were acquired and coregistered with mean EPI images. Functional and structural data were then spatially normalized to Montreal Neurological Institute (MNI) space and smoothed with a 6 mm³ full-width at half-maximum (FWHM) Gaussian using the DARTEL toolbox (Ashburner, 2007). Respiration and heart rate were recorded using a breathing belt and pulse oximeter (Hutton et al., 2011).

Region of interest selection. Based on previous literature we defined regions of interest (ROIs) for our contrasts of interest in the ventral striatum (6 mm radius spheres centered at [11 11 -2] and [-11 11 -2]; Guitart-Masip et al., 2011) and the ventromedial prefrontal cortex (vmPFC; 8 mm spheres centered at [6 50 -11] and [-6 50 -11]; Wright et al., 2013). These ROIs were used for small volume correction in our fMRI analyses.

fMRI univariate analysis. We created a general linear model (GLM) containing separate events for each offer condition (visual, V_v , auditory, A , or cross-modal, C), modeled as 2 s duration boxcars. Each of these event regressors was modulated by three additional parametric regressors, reflecting the value (indicated by a zero or one) of stimuli in visual (V_v), auditory (V_a), and cross-modal (V_c) modalities. We modeled outcome pre-

sentation separately for both accept and reject conditions, using a stick function, with each of these regressors in turn modulated by an additional parametric regressor indicating whether a win or a loss was signaled on that trial, giving four regressors in total. Regressors reflecting condition presentation time (when the text indicating which sensory modality was relevant was presented to subjects); the six motion regressors produced by the realignment stage of preprocessing; and physiological noise regressors consisting of six cardiac regressors, six respiratory regressors, and two regressors for heart rate change and change in respiratory volume (Hutton et al., 2011) were included as regressors of no interest. Unless otherwise stated, we report results that were significant at a threshold of $p < 0.05$, familywise error corrected (p_{wfb}), either for the whole brain, or using small volume correction for one of our regions of interest (p_{svc}).

To compare the overall effect of behaviorally relevant properties of the stimuli with behaviorally irrelevant ones, we created appropriately weighted contrast images for each subject encoding the mean activity for relevant values ($V_{\text{rel}} = ((V_a|A) + (V_v|V) + (V_c|C))/3$), irrelevant values ($V_{\text{irr}} = ((V_v|A) + (V_c|A) + (V_a|V) + (V_c|V) + (V_a|C) + (V_v|C))/6$), and their difference ($V_{\text{rel}} - V_{\text{irr}}$). We performed a second-level analysis using a summary statistics approach. To test for whether the overall effects we observed were present in each of the three task conditions, we also performed similar analyses for each task conditions (V , A , or C) separately.

To test for intersubject correlations between activity in the ventral striatum and the effects of task-relevant and task-irrelevant value signals on behavior, we extracted the average parameter estimates from our striatal ROIs for relevant and irrelevant value signals (averaged across modalities), and compared these with parameter estimates from our logistic regression (averaged across modalities).

For approximately half of the subjects in our study ($n = 12$) incongruent cross-modal stimuli were specified as good, while for the remaining subjects ($n = 13$) this was the case for congruent stimuli. To rule out the possibility that additional computations in one of these groups, for example, the configural demands in the incongruent group, altered neuronal representations of value in such a way as to eliminate behaviorally irrelevant value signals that would have otherwise been present, we tested for relevant and irrelevant value effects separately in both congruent and incongruent groups, using the mean parameter estimates for each contrast extracted from our striatal ROIs.

To assess outcome processing, for each subject we created separate contrasts for the parametric regressors encoding wins or losses in both the accept (W_{acc}) and reject (W_{rej}) conditions. The mean activity across accept and reject conditions reflecting correct outcomes was calculated as $W_{\text{acc}} - W_{\text{rej}}$, reflecting the fact that losses in the reject condition signaled that a subject had made the correct decision. (Error signaling across conditions was thus reflected by $W_{\text{rej}} - W_{\text{acc}}$.)

fMRI multivariate decoding analysis. To explore brain regions that might contain information about stimulus properties even on task-irrelevant trials, we applied a multivariate searchlight decoding approach to our data using a linear support vector machine (SVM; Kamitani and Tong, 2005; Norman et al., 2006; Kriegeskorte and Bandettini, 2007; Chadwick et al., 2012). For each subject, using unsmoothed native space data, we first estimated a GLM containing a separate 2 s boxcar regressor encoding stimulus presentation for each trial, together with regressors encoding outcomes, condition presentation, movement, and physiological noise as described above. We then calculated a single T statistic image for each trial, and used these for decoding analysis using a linear C-SVM implemented in LIBSVM (Chang and Lin, 2011). Analysis was based on T statistic images rather than contrast images as these downweight the effects of noisy voxels and have been shown to be advantageous for multivariate analysis (Misaki et al., 2010).

For each stimulus property (visual, auditory, cross-modal), we attempted to decode which stimulus (good or bad) was present on each trial, separating trials where the stimulus property was relevant from those where it was irrelevant giving a total of six separate decoding analyses (for the cross-modal condition we attempted to classify between congruent and incongruent combinations alone, not between all four individual pairings, since these depended on the visual and auditory stimuli themselves). To assess classification accuracy at each voxel we first extracted T values for each voxel within a spherical searchlight with

6 mm radius centered on it (31 voxels), and then performed classification with 10-fold cross-validation. Trials were randomly separated into 10 partitions, one partition was removed from the dataset, the classifier was then trained on the remaining nine, and then accuracy was assessed on the tenth. This was repeated 10 times, using a different partition each time, and the resulting estimates were averaged to give a single decoding accuracy value for each voxel. Classification accuracy images were normalized to MNI space and smoothed with an 8 mm³ FWHM Gaussian using DARTEL, and second-level inference performed using SPM.

Modality-specific responses. We also examined whether activity reflecting offer values or trial outcomes varied depending upon the task condition (whether visual, auditory, or cross-modal stimuli were task relevant), effects we hypothesized might be observed in visual, auditory, and multisensory areas for the three experimental conditions based on recent demonstrations that rewarding feedback activates sensory cortices (Pleger et al., 2008, 2009; Weil et al., 2010; FitzGerald et al., 2013). Accordingly we compared relevant value signals between conditions in the model described above, and created an additional model in which separate regressors were used for outcome events in the three conditions. Despite performing a number of different analyses using different functional and anatomical ROIs, we found no clear evidence of offer or trial outcome-related activity that differed between conditions, and we do not discuss the results of these analyses any further below.

Results

Behavior

For all three conditions, subjects showed a strong preference for selecting the correct action (accepting on trials with a positive expected utility and rejecting on trials with a negative expected utility; Visual: $\mu = 0.97$, $\sigma = 0.03$; Auditory: $\mu = 0.93$, $\sigma = 0.06$; Cross-modal: $\mu = 0.89$, $\sigma = 0.10$). Accuracy was significantly higher in the visual compared with both the auditory ($p = 0.004$ signed rank test) and cross-modal conditions ($p < 0.001$ signed rank test). Accuracy was significantly higher in the auditory than the cross-modal condition ($p = 0.032$ signed rank test). No significant differences were observed for rates of correct responding to good and bad stimuli, suggesting accuracy was unaffected by whether subjects were making a decision about positively or negatively valenced stimuli.

The results of our logistic regression analysis suggest the effect of stimulus valence in nonrelevant modalities was much smaller than in relevant modalities. However, these effects were significantly greater than zero for all three sensory modalities, reflecting a positive effect of value on acceptance likelihood (Visual relevant: $\mu = 47.0$, $p < 0.001$; Visual irrelevant: $\mu = 8.58$, $p = 0.004$; Auditory relevant: $\mu = 31.4$, $p < 0.001$; Auditory irrelevant: $\mu = 6.89$, $p = 0.008$; Cross-modal relevant: $\mu = 19.4$, $p < 0.001$; Cross-modal irrelevant: $\mu = 6.62$, $p = 0.020$; all signed rank test). There were no significant differences in the effect of task-irrelevant valence between the distinct modalities.

Offer value signals in the ventral striatum

Across all three conditions, the value of behaviorally relevant stimulus features correlated positively with activity in bilateral ventral striatum (Right: [12 9 -6], $Z = 4.31$, $p_{\text{svc}} = 0.0003$; Left: [-9 12 -6], $Z = 3.90$, $p_{\text{svc}} = 0.002$; Fig. 1). No other region showed significant positive or negative correlations with behaviorally relevant value, although activity in vmPFC correlated positively with value, but this did not survive correction for multiple comparisons (peak voxel: [3 45 -6], Z -score: 2.53). Focusing on activity in ventral striatum alone we were unable to show either a significant positive or negative correlation with behaviorally irrelevant value (Fig. 1) offer value signals.

To unambiguously demonstrate that ventral striatal activity reflects the value of behaviorally relevant, more than behaviorally

irrelevant, stimuli it is necessary to show not just a difference in significance between conditions but also a significant difference as reflected in the ($V_{\text{rel}} - V_{\text{irr}}$) contrast. This was exactly what we observed in bilateral ventral striatum (Right: [12 9 -6], $Z = 4.66$, $p_{\text{svc}} < 0.0001$; Left: [-12 9 -6], $Z = 3.84$, $p_{\text{svc}} = 0.002$), again consistent with this region being preferentially engaged by the value of behaviorally relevant stimulus features. We also tested whether this difference, in our analysis pooled across V , A , and C , was also evident in each of these conditions considered separately. For the visual condition, the ($V_{\text{rel}} - V_{\text{irr}}$) contrast showed a significant correlation with activity in bilateral striatum (Right: [12 12 -6], $Z = 3.74$, $p_{\text{svc}} = 0.003$; Left: [-12 12 -9], $Z = 2.89$, $p_{\text{svc}} = 0.043$). Similar effects were seen in the cross-modal condition (Right: [12 6 -3], $Z = 4.44$, $p_{\text{svc}} = 0.0001$; Left: [-9 9 -6], $Z = 3.47$, $p_{\text{svc}} = 0.008$) while in the auditory condition a significant correlation was seen in the right striatum alone ([9 9 -3], $Z = 2.86$, $p_{\text{svc}} = 0.044$), with activity in the left striatum ROI not surviving correction for multiple comparisons ([-9 12 0], $Z = 2.20$, $p_{\text{svc}} = 0.158$, $p = 0.0137$ uncorrected). These data are consistent with the idea that activity in ventral striatum reflects the value of behaviorally relevant stimuli alone, regardless of whether these features are visual, auditory, or conjoint visual and auditory modalities.

Between-subject effects

Although we found strong positive correlations between the size of effect of task-relevant value on behavior and activity in bilateral ventral striatum (Right: $r = 0.581$, $p = 0.002$; Left: $r = 0.552$, $p = 0.004$), we failed to observe a similar relationship for task-irrelevant value (Right: $r = -0.243$, $p = 0.241$; Left: $r = -0.183$, $p = 0.382$). This indicates that the effects of task-irrelevant value on behavior are not mediated by signals in the ventral striatum, but we also acknowledge the alternative possibility that it could reflect the fact these signals are small in amplitude.

Both groups of subjects, namely those who performed the task under conditions where congruent stimuli represented good offers and those for whom this was the case for incongruent stimuli, showed significantly greater striatal responses to relevant than irrelevant value (Congruent right: $\mu = 0.289$, $p = 0.003$; Congruent left: $\mu = 0.266$, $p = 0.004$; Incongruent right: $\mu = 0.339$, $p = 0.021$; Incongruent left: $\mu = 0.266$, $p = 0.039$; all signed rank test). No significant responses to irrelevant value were observed in either condition, though a trend was observed in left ventral striatum for the Incongruent group (Congruent right: $\mu = -0.097$, $p = 0.971$; Congruent left: $\mu = -0.058$, $p = 0.916$; Incongruent right: $\mu = 0.043$, $p = 0.235$; Incongruent left: $\mu = 0.079$, $p = 0.055$; all signed rank test). This suggests that differential representation of task-relevant and -irrelevant value was largely unaffected by the valence of congruent stimuli in the cross-modal condition.

Offer value signals in the rest of the brain

To examine whether other brain regions might represent the value of behaviorally irrelevant stimuli, we generated whole-brain activation maps. No regions showed activity that survived correction for multiple comparisons, and even at a very liberal threshold ($p < 0.01$ uncorrected, minimum cluster size 5 voxels), and we did not observe activation either in sensory cortex or in regions typically associated with value, such as the vmPFC or striatum (Table 1). This is consistent with the hypothesis that behavioral relevance gates the flow of information into value-sensitive regions, but like all negative findings it should be interpreted with caution.

Table 1. Whole-brain results from behaviorally irrelevant value contrast

Location	Side	p (FWE-corr)	T	Z	p (unc)	X	y	z
White matter	Left	0.973	4.2	3.6	0.0002	-27	-36	27
Fusiform gyrus	Right	0.994	4	3.47	0.0003	36	-51	-18
Cerebellum	Right	>0.999	3.63	3.21	0.0007	-3	-57	-24
White matter	Left	>0.999	3.61	3.2	0.0007	-21	-15	27
Amygdala	Right	>0.999	3.56	3.16	0.0008	-33	3	-24
Thalamus	Midline	>0.999	3.45	3.08	0.001	0	-24	-3
Insula	Left	>0.999	3.44	3.07	0.0011	-39	-18	-9
White matter	Left	>0.999	2.66	2.47	0.0068	-42	-24	0
Middle frontal gyrus	Right	>0.999	3.39	3.04	0.0012	48	3	54
Precentral gyrus	Right	>0.999	2.86	2.63	0.0043	48	-6	57
Cerebellum	Right	>0.999	3.37	3.02	0.0013	12	-60	-24
Pallidum	Left	>0.999	3.31	2.98	0.0015	-15	0	3
Olfactory cortex	Right	>0.999	3.15	2.85	0.0022	21	9	-15
White matter	Left	>0.999	3.15	2.85	0.0022	-24	-27	42
Supplementary motor area	Right	>0.999	3.08	2.8	0.0026	9	-9	51
Postcentral gyrus	Right	>0.999	2.86	2.62	0.0044	60	-6	39

No activations were significant after correction for multiple comparisons across the whole brain.

In addition, in an exploratory analysis we performed a multivariate decoding analysis to see whether stimulus representations not evident in localized mean signal changes were present when they were task irrelevant (because we used only two visual and auditory stimuli, our decoding analysis is unable to distinguish between representations of stimulus value, and representations of the stimulus per se, if indeed these are different; nonetheless, it can provide information about whether and where some stimulus features are represented). Visual stimulus properties could be decoded from visual cortex on both task-relevant and task-irrelevant trials (Relevant: [30 -84 -6], $Z = 5.57$, $p_{wb} = 0.001$; Irrelevant: [-36 -81 -6], $Z = 5.19$, $p_{wb} = 0.006$), and conjunction analysis at a threshold of $p < 0.001$ uncorrected revealed a large overlap in bilateral visual cortex (Fig. 3), consistent with the hypothesis that the flow of information from sensory regions representing individual stimuli to the ventral striatum is gated by task relevance (Fig. 1). No significant differences between relevance conditions were found, even using a small volume correction for the results of the conjunction analysis.

For auditory and cross-modal stimuli no regions showed decoding accuracy that survived correction for multiple comparisons in either the task-relevant or -irrelevant conditions. This may reflect properties of the stimuli themselves, or else of the neuronal responses in regions processing auditory and cross-modal stimuli.

Outcome signals in the ventral striatum

In the accept condition, activity in bilateral ventral striatum correlated positively with rewarded outcomes (Right: [9 9 -6], $Z = 2.82$, $p_{svc} = 0.045$; Left: [-9 9 -6], $Z = 2.95$, $p_{svc} = 0.033$), as predicted by previous findings (Seymour et al., 2004). In the reject condition, striatal activity showed the opposite pattern, manifesting a significant negative pattern of responding (Right: [9 12 -6], $Z = 2.85$, $p_{svc} = 0.047$; Left: [-12 12 -9], $Z = 3.09$, $p_{svc} = 0.027$; Fig. 4). This supports the idea that activity in this region at outcome presentation time is best explained in terms of a signal needed for implementation of a successful behavioral policy (Klein-Flügge et al., 2011; Li and Daw, 2011) rather than one used in updating an action value using a fictive reward signal (Watkins and Dayan, 1992; Lohrenz et al., 2007).

Outcome signals in the rest of the brain

For correct actions, significant positive correlations with obtained rewards were evident in the right caudate ([21 18 21], $Z =$

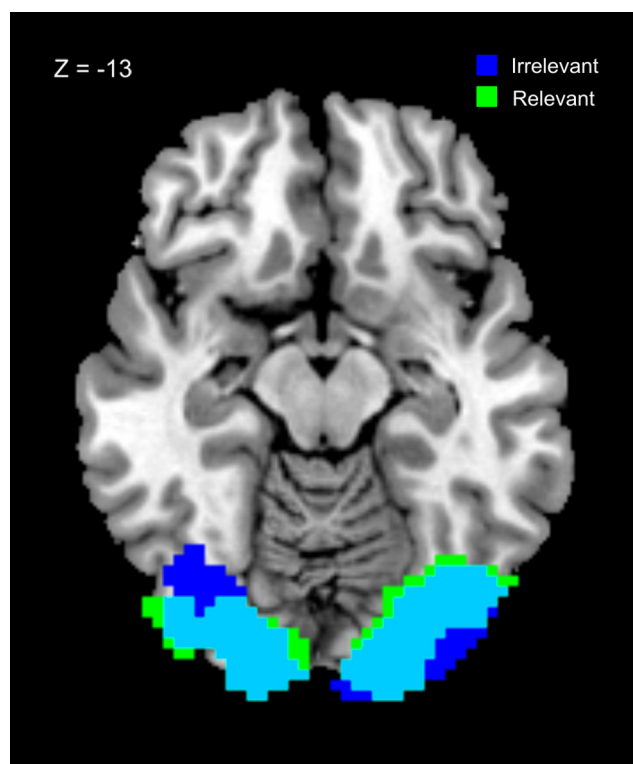


Figure 3. Results of the searchlight decoding analysis for visual stimuli in task-relevant and -irrelevant conditions. Significant decoding was possible in visual cortex for both relevance conditions, and conjunction analysis revealed a large area of overlap in visual cortex. This suggests that visual stimuli (whether task relevant or not) are represented in a similar fashion in sensory areas. (Images thresholded at $p < 0.001$ uncorrected; blue irrelevant, green relevant.)

4.86, $p_{wb} = 0.033$), the left superior parietal cortex ([-15 60 69], $Z = 4.78$, $p_{wb} = 0.050$), and vmPFC (Right: [9 51 -6], $Z = 4.08$, $p = 0.002$; Left: [-3 48 -9], $Z = 3.75$, $p_{svc} = 0.007$). Negative correlations with forgone rewards were found in the left supplementary motor area ([-6 -15 51], $Z = 5.00$, $p_{wb} = 0.016$), the right precentral gyrus ([24 -9 51], $Z = 4.92$, $p_{wb} = 0.024$), and the right vmPFC ([9 48 -9], $Z = 3.54$, $p_{svc} = 0.014$). The finding that the vmPFC responds to feedback about correct decisions in both accept and reject conditions is interesting as it suggests, that, as for the striatum, it is concerned with evaluating the quality of action outcomes rather than processing outcomes themselves,

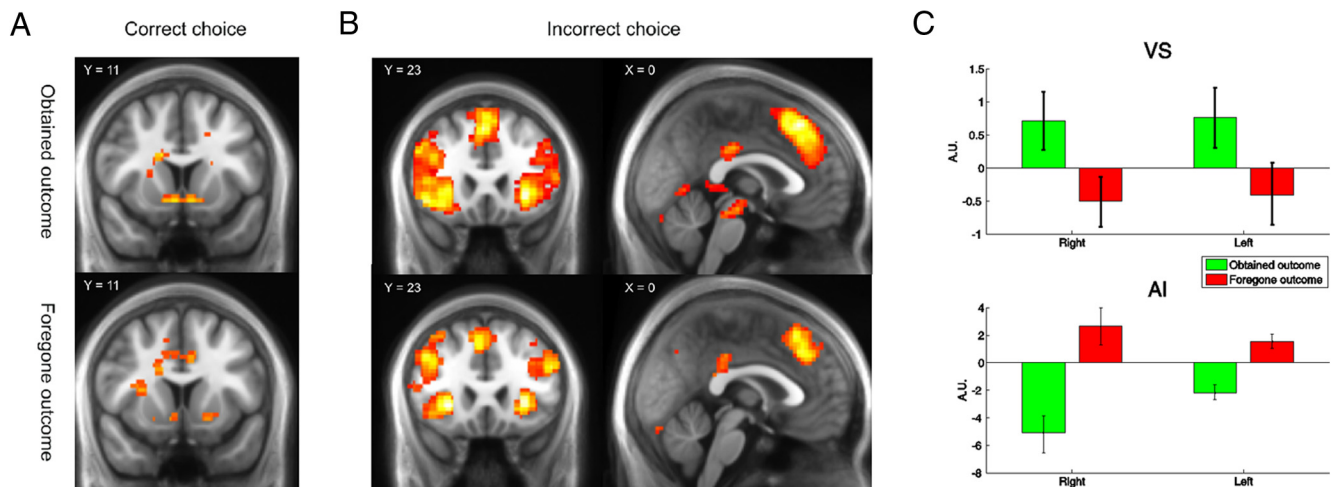


Figure 4. *A*, Outcome activity in ventral striatum was correlated positively with obtained wins (top) and negatively with foregone wins (bottom), consistent with a role in signaling whether a subject had performed the correct action or not. (Image thresholded at $p < 0.005$ uncorrected for display purposes; $Y = 11$ mm.) *B*, Activity in the dorsomedial prefrontal cortex and bilateral anterior insula correlated negatively with obtained wins (top) and positively with foregone ones (bottom), consistent with a role in error signaling. (Image thresholded at $p < 0.005$ uncorrected for display purposes; $Y = 23$ mm, $X = 0$ mm.) *C*, Illustration of outcome responses. Mean parameter estimates extracted from voxels in the left and right ventral striatum showing the strongest positive responses to obtained outcomes (top), and from voxels in the left and right anterior insula showing the strongest negative responses to obtained outcomes (bottom). Activity in as the ventral striatum and anterior insula showed opposite patterns of responding to the wins minus losses contrast in the obtained (accept, green) and foregone (reject, red) conditions. (Note that these plots are illustrative only, and all inference was performed using the standard SPM analysis.) Error bars indicate 95% confidence intervals.

consistent with the recent finding that vmPFC activity encodes information about specific actions (FitzGerald et al., 2012).

Foregone rewards were positively correlated with activity in the right lateral prefrontal cortex ([48 27 30], $Z = 5.00$, $p_{wb} = 0.016$), bilateral anterior insula (Right: [30 21 -3], $Z = 4.99$, $p_{wb} = 0.017$; Left: [-27 21 0], $Z = 4.86$, $p_{wb} = 0.033$), and the dorsomedial prefrontal cortex ([6 30 42], $Z = 4.84$, $p_{wb} = 0.036$; Fig. 4). A similar pattern of activity was observed for obtained losses, which were positively correlated with activity in the dorsomedial prefrontal cortex ([6 30 39], $Z = 5.52$, $p_{wb} = 0.001$), bilateral anterior insula (Right: [48 18 -9], $Z = 5.16$, $p_{wb} = 0.007$; Left: [-30 18 -3], $Z = 5.47$, $p_{wb} = 0.001$), right middle temporal gyrus ([60 -30 -6], $Z = 5.14$, $p_{wb} = 0.008$), and right lateral prefrontal cortex ([48 12 21], $Z = 4.81$, $p_{wb} = 0.043$; Fig. 4). These results are consistent with findings of previous studies of error monitoring, which implicate the dorsomedial (or possible anterior cingulate) and insula cortices (Ridderinkhof et al., 2004; Klein et al., 2007), as well as a recent study examining counterfactual outcome processing, which showed similar activity in the dorsomedial prefrontal cortex (Boorman et al., 2011).

Differences between conditions in the processing of correct responses and errors

No region showed significantly stronger activity when comparing correct responses in the accept and reject conditions. While it is unwise to infer conclusions from a negative result, it is worth observing that this is consistent with the idea that similar processes are implemented in ventral striatum (and other areas) during the processing of feedback indicating correct or incorrect action selection, whether or not that feedback itself indicates either a positive or a negative outcome.

Because of subjects' high levels of accuracy on the task, we were unable to compare differences in activity between outcome signals when subjects chose to accept the good offer versus those when they chose to accept the bad offer (and similarly for reject trials). Future work examining this could shed useful light on the sort of outcome signals present during performance of this task.

Discussion

We show value-correlated activity in the ventral striatum reflects the behavioral relevance of stimulus features. This is consistent with the idea that selective and appropriate responding to valenced stimuli depends upon modulation that occurs before or during the process of evaluation (Dayan et al., 2000) rather than changes in the influence of automatically generated value representations on effector responses (Fig. 1). In addition, the results of our multivariate decoding analysis suggest that (at least for the visual modality) stimuli are represented in a similar manner in sensory cortex both when they are relevant and when they are irrelevant, and that task relevance is thus likely to modulate the influence of sensory areas on regions processing value. Also in accord with previous findings (Klein-Flügge et al., 2011; Li and Daw, 2011; Guitart-Masip et al., 2012), we show that activity in ventral striatal activity to rewarding outcomes is critically sensitive to features of an action (correct versus incorrect), a profile that resembles a policy update signal rather than a reward prediction error per se.

By manipulating whether a visual cue, a simultaneously presented auditory cue, or a combination of both determined the value of a trial offer we show that striatal activity was influenced solely (or only detectably) by stimulus value in the behaviorally relevant modality. Although we are unable to directly test what mechanisms are responsible for this modulation by behavioral context, our findings are in keeping with theoretical proposals, which suggest selective attention gates inputs to value-processing regions, and weight-value predictions appropriately (Dayan et al., 2000; Yu and Dayan, 2005; Gershman et al., 2010), and the recent finding that spatial attention affects value comparison signals during binary choice (Lim et al., 2011). If true, this is a clear example of a case in which selective attention, rather than reflecting limited computational resources, is in accord with the demands of optimal inference (Rao, 2005; Yu et al., 2009; Feldman and Friston, 2010; Dayan and Solomon, 2010). Indeed we see this as a promising area for future study, particularly the question of

whether diverting attention through exogenous cuing or attentional load manipulations impairs subjects' ability to focus on task-relevant stimuli, and whether there are consequential effects on ventral striatal reward signals.

Our results build on a considerable existing literature implicating ventral striatum in anticipated reward (Schultz et al., 1992; Knutson et al., 2001; O'Doherty et al., 2004; Kable and Glimcher, 2007; Bartra et al., 2013; Clithero and Rangel, 2013). Of particular relevance here is a recent finding that responses in the ventral striatum during a binary decision-making task are sensitive to exogenous manipulations of spatial attention (Lim et al., 2011), which altered the sign of a signal reflecting the difference in value between the two options. In our experiment any attentional shifts are endogenous, and occur not between spatial locations but between sensory modalities. This implies that value signals in ventral striatum are subject to flexible and dynamic modulations by attention, a finding that makes sense given the key role they are likely to play in adaptive choice.

For simplicity, we explicitly manipulated which stimulus features determined the value of the offer on a particular trial, but the issue of how subjects infer which features are currently relevant for behavior is also of considerable interest (Dayan et al., 2000). Recently, sophisticated behavioral modeling has been deployed to test hypotheses about how subjects infer which stimulus features they should attend to (Gershman et al., 2010; Wilson and Niv, 2011), but the neural substrates of this process remain to be elucidated, something for which model-based neuroimaging approaches seem ideally suited.

A question unresolved in our analysis is which brain regions mediate the (admittedly small) effects of task-irrelevant stimulus properties on behavior. One possibility is that task-irrelevant value signals are indeed present in the ventral striatum, but that these are either so small in magnitude or so transient that they were not discernible in the blood oxygenation level-dependent signal. This is suggested by a Bayesian perspective (Dayan et al., 2000), where task relevance is encoded as a context-dependent probability weight, rather than a binary on/off switch (on this account subjects are never completely certain about which context they are in, explaining the weak effect of irrelevant value behavior). Alternatively, it may be that these effects are mediated by other structures, perhaps through some sort of valence-based priming of approach (acceptance) behavior (Tucker and Ellis, 2004; Guitart-Masip et al., 2011) as, for example, seen in pavlovian-instrumental transfer (Balleine and Killcross, 2006; Bray et al., 2008; Talmi et al., 2008).

In principle, subjects could learn and maintain appropriate responding on our task either by directly learning a policy (which action they should perform given a particular state), or by basing choices on the representation of specific action values (Watkins and Dayan, 1992; Sutton and Barto, 1998; Li and Daw, 2011). These alternatives make similar predictions about the type of feedback signals we expect to see when subjects accepted the offer made to them (obtained outcomes), but critically they make opposite predictions about signals when subjects rejected the offer (foregone outcomes). If subjects solve the task by encoding and maintaining action values, this should be reflected in outcome signals with the same sign in the case of both real and foregone outcomes. If, in contrast, subjects learn policies (correct actions) directly, the real and fictive learning signals should have opposite signs, because a good foregone outcome provides evidence against the current policy (Li and Daw, 2011).

As reported previously (Li and Daw, 2011; for review, see Lohrenz et al., 2007), our data strongly suggest that activity in

ventral striatum resembles a policy update signal rather than a value prediction signal. Interestingly, this may simply be a special case of a more general role of outcome signals in the ventral striatum in signaling the accuracy or correctness of a recent action. In a recent study where participants were asked to judge the time at which rewards were delivered, striatal activity reflected the accuracy of predictions about timing, and was clearly dissociated from activity in the midbrain, which showed a pattern of activity more consistent with a reward prediction error (Klein-Flügge et al., 2011). This leads us to hypothesize that, at least in certain contexts, the key role of ventral striatum in outcome processing is to signal the success of a particular action, and hence the desirability of repeating it in the future. This may explain why outcome signals in the ventral striatum are profoundly reduced when subjects have no ability to determine the outcome through their actions (Zink et al., 2004; Coricelli et al., 2005; Nicolle et al., 2011; similar results have also been reported in the dorsal striatum (Tricomi et al., 2004).

Our results should not be interpreted as indicating outcome signals in the ventral striatum never reflect value updating (O'Doherty et al., 2004; Seymour et al., 2004; Hare et al., 2008; Rangel et al., 2008; Kim et al., 2009). Instead we take a more nuanced view that under certain circumstances subjects adopt a strategy based upon a direct updating of policies (Sutton and Barto, 1998; Li and Daw, 2011). Plausibly human subjects are able to employ different types of learning strategy depending on the precise nature of the task, and understanding when they do so has echoes with issues such as characterizing the nature of strategies used in particular environments. When considering neuronal responses, which positively covaried with outcomes indicating erroneous actions (accepted losses or forgone wins), we observed activity in similar networks of regions that have previously been associated with the error processing (Ridderinkhof et al., 2004; Klein et al., 2007). How this error signaling relates to the mirror image correctness signal we find in the ventral striatum is an interesting area for future studies to explore.

Combining the results of our two key analyses suggests that rather than simply reflecting the input of simple cue-dependent reward prediction error-like signaling from the dopaminergic midbrain, activity in ventral striatum during instrumental learning is tuned to which features of the environment are relevant for action, and the desirability or otherwise of repeating a particular action regardless of whether a subject received reward (Klein-Flügge et al., 2011). This represents a further step toward understanding the contribution the basal ganglia make to value-based decision making, and therefore human decision-processes themselves.

References

- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113. [CrossRef Medline](#)
- Balleine BW, Killcross S (2006) Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci* 29:272–279. [CrossRef Medline](#)
- Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76:412–427. [CrossRef Medline](#)
- Boorman ED, Behrens TE, Rushworth MF (2011) Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol* 9:e1001093. [CrossRef Medline](#)
- Bray S, Rangel A, Shimojo S, Balleine B, O'Doherty JP (2008) The neural mechanisms underlying the influence of pavlovian cues on human decision making. *J Neurosci* 28:5861–5866. [CrossRef Medline](#)
- Chadwick MJ, Bonnici HM, Maguire EA (2012) Decoding information in the

- human hippocampus: a user's guide. *Neuropsychologia* 50:3107–3121. [CrossRef Medline](#)
- Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 2:27.
- Clithero JA, Rangel A (2013) Informatic parcellation of the network involved in the computation of subjective value. *Soc Cogn Affect Neurosci*. Advance online publication. Retrieved May 10, 2013. doi:10.1093/scan/nst106. [CrossRef Medline](#)
- Coricelli G, Critchley HD, Joffily M, O'Doherty JP, Sirigu A, Dolan RJ (2005) Regret and its avoidance: a neuroimaging study of choice behavior. *Nat Neurosci* 8:1255–1262. [CrossRef Medline](#)
- Dayan P, Daw ND (2008) Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* 8:429–453. [CrossRef Medline](#)
- Dayan P, Solomon JA (2010) Selective Bayes: attentional load and crowding. *Vision Res* 50:2248–2260. [CrossRef Medline](#)
- Dayan P, Kakade S, Montague PR (2000) Learning and selective attention. *Nat Neurosci* [3 Suppl]:1218–1223. [Medline](#)
- Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19:430–441. [CrossRef Medline](#)
- Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. *Front Hum Neurosci* 4:215. [Medline](#)
- FitzGerald TH, Friston KJ, Dolan RJ (2012) Action-specific value signals in reward-related regions of the human brain. *J Neurosci* 32:16417–16423a. [CrossRef Medline](#)
- FitzGerald TH, Friston KJ, Dolan RJ (2013) Characterising reward outcome signals in sensory cortex. *Neuroimage* 83:329–334. [CrossRef Medline](#)
- Friston KJ, Daunizeau J, Kiebel SJ (2009) Reinforcement learning or active inference? *PLoS One* 4:e6421. [CrossRef Medline](#)
- Gershman S, Cohen J, Niv Y (2010) Learning to selectively attend. 32nd Annual Conference of the Cognitive Science Society, Portland, OR.
- Guitart-Masip M, Fuentemilla L, Bach DR, Huys QJ, Dayan P, Dolan RJ, Duzel E (2011) Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J Neurosci* 31:7867–7875. [CrossRef Medline](#)
- Guitart-Masip M, Huys QJ, Fuentemilla L, Dayan P, Duzel E, Dolan RJ (2012) Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* 62:154–166. [CrossRef Medline](#)
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28:5623–5630. [CrossRef Medline](#)
- Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R (2002) Image distortion correction in fMRI: a quantitative evaluation. *Neuroimage* 16:217–240. [CrossRef Medline](#)
- Hutton C, Josephs O, Stadler J, Featherstone E, Reid A, Speck O, Bernarding J, Weiskopf N (2011) The impact of physiological noise correction on fMRI at 7 T. *Neuroimage* 57:101–112. [CrossRef Medline](#)
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633. [CrossRef Medline](#)
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685. [CrossRef Medline](#)
- Kim H, Sul JH, Huh N, Lee D, Jung MW (2009) Role of striatum in updating values of chosen actions. *J Neurosci* 29:14701–14712. [CrossRef Medline](#)
- Klein TA, Endrass T, Kathmann N, Neumann J, von Cramon DY, Ullsperger M (2007) Neural correlates of error awareness. *Neuroimage* 34:1774–1781. [CrossRef Medline](#)
- Klein-Flügge MC, Hunt LT, Bach DR, Dolan RJ, Behrens TE (2011) Dissociable reward and timing signals in human midbrain and ventral striatum. *Neuron* 72:654–664. [CrossRef Medline](#)
- Knutson B, Adams CM, Fong GW, Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159. [Medline](#)
- Kriegeskorte N, Bandettini P (2007) Analyzing for information, not activation, to exploit high-resolution fMRI. *Neuroimage* 38:649–662. [CrossRef Medline](#)
- Li J, Daw ND (2011) Signals in human striatum are appropriate for policy update rather than value prediction. *J Neurosci* 31:5504–5511. [CrossRef Medline](#)
- Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J Neurosci* 31:13214–13223. [CrossRef Medline](#)
- Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci U S A* 104:9493–9498. [CrossRef Medline](#)
- Misaki M, Kim Y, Bandettini PA, Kriegeskorte N (2010) Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage* 53:103–118. [CrossRef Medline](#)
- Nicolle A, Bach DR, Driver J, Dolan RJ (2011) A role for the striatum in regret-related choice repetition. *J Cogn Neurosci* 23:845–856. [CrossRef Medline](#)
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430. [CrossRef Medline](#)
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454. [CrossRef Medline](#)
- Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD (2008) Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59:561–567. [CrossRef Medline](#)
- Pleger B, Blankenburg F, Ruff CC, Driver J, Dolan RJ (2008) Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J Neurosci* 28:8161–8168. [CrossRef Medline](#)
- Pleger B, Ruff CC, Blankenburg F, Klöppel S, Driver J, Dolan RJ (2009) Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol* 7:e1000164. [CrossRef Medline](#)
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556. [CrossRef Medline](#)
- Rao RP (2005) Bayesian inference and attentional modulation in the visual cortex. *Neuroreport* 16:1843–1848. [CrossRef Medline](#)
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. *Science* 306:443–447. [CrossRef Medline](#)
- Schultz W, Apicella P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci* 12:4595–4610. [Medline](#)
- Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.
- Talmi D, Seymour B, Dayan P, Dolan RJ (2008) Human pavlovian-instrumental transfer. *J Neurosci* 28:360–368. [CrossRef Medline](#)
- Tricomi EM, Delgado MR, Fiez JA (2004) Modulation of caudate activity by action contingency. *Neuron* 41:281–292. [CrossRef Medline](#)
- Tucker M, Ellis R (2004) Action priming by briefly presented objects. *Acta Psychol* 116:185–203. [CrossRef Medline](#)
- Watkins CJCH, Dayan P (1992) Q-learning. *Mach Learn* 8:279–292. [CrossRef](#)
- Weil RS, Furl N, Ruff CC, Symmonds M, Flandin G, Dolan RJ, Driver J, Rees G (2010) Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *J Neurophysiol* 104:1746–1757. [CrossRef Medline](#)
- Wilson RC, Niv Y (2011) Inferring relevance in a changing world. *Front Hum Neurosci* 5:189. [Medline](#)
- Wright ND, Symmonds M, Dolan RJ (2013) Distinct encoding of risk and value in economic choice between multiple risky options. *Neuroimage* 81:431–440. [CrossRef Medline](#)
- Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46:681–692. [CrossRef Medline](#)
- Yu AJ, Dayan P, Cohen JD (2009) Dynamics of attentional selection under conflict: toward a rational Bayesian account. *J Exp Psychol Hum Percept Perform* 35:700–717. [CrossRef Medline](#)
- Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004) Human striatal responses to monetary reward depend on saliency. *Neuron* 42:509–517. [CrossRef Medline](#)