CALL FOR PAPERS | Decision Making: Neural Mechanisms

Value and probability coding in a feedback-based learning task utilizing food rewards

Elizabeth Tricomi and Karolina M. Lempert
Department of Psychology, Rutgers University, Newark, New Jersey
Submitted 27 January 2014; accepted in final form 2 October 2014

Tricomi E. Lempert KM. Value and probability coding in a feedback-based learning task utilizing food rewards. J Neurophysiol 113: 4–13, 2015. First published October 22, 2014; doi:10.1152/jn.00086.2014.—For the consequences of our actions to guide behavior, the brain must represent different types of outcome-related information. For example, an outcome can be construed as negative because an expected reward was not delivered or because an outcome of low value was delivered. Thus behavioral consequences can differ in terms of the information they provide about outcome probability and value. We investigated the role of the striatum in processing probability-based and value-based negative feedback by training participants to associate cues with food rewards and then employing a selective satiety procedure to devalue one food outcome. Using functional magnetic resonance imaging, we examined brain activity related to receipt of expected rewards, receipt of devalued outcomes, omission of expected rewards, omission of devalued outcomes, and expected omissions of an outcome. Nucleus accumbens activation was greater for rewarding outcomes than devalued outcomes, but activity in this region did not correlate with the probability of reward receipt. Activation of the right caudate and putamen, however, was largest in response to rewarding outcomes relative to expected omissions of reward. The dorsal striatum (caudate and putamen) at the time of feedback also showed a parametric increase correlating with the trialwise probability of reward receipt. Our results suggest that the ventral striatum is sensitive to the motivational relevance, or subjective value, of the outcome, while the dorsal striatum codes for a more complex signal that incorporates reward probability. Value and probability information may be integrated in the dorsal striatum, to facilitate action planning and allocation of effort.

fMRI; striatum; caudate; nucleus accumbens; reward processing

Decision making is guided by information about the value and probability of the outcomes of one’s choices. Value is a subjective measure that varies with internal motivational states and goals. For example, ordering chocolate almond cake may be rewarding, but not to someone with a nut allergy or to someone who is too full for dessert. Estimates of outcome probability, on the other hand, are derived by learning the contingencies that exist between actions and outcomes. For some actions, the contingency between action and outcome is very strong, such as when one turns on a faucet and water pours out. Other actions, such as playing a slot machine, have less predictable consequences.

The subjective value of an outcome, defined here as the desirability of the outcome, if obtained, can be altered through changes in motivational state. In both animals and humans, a selective satiety procedure has been used to reduce outcome value; individuals are fed a food reward to satiety, so that the value of that outcome decreases (Balleine and Dickinson 1998; Tricomi et al. 2009). Further delivery of that food constitutes an outcome of low subjective value. For instance, the receipt of an extra slice of cake when you are sated would be an outcome of low value, even though the first slice may have had a high subjective value.

In contrast, the omission of an expected positive outcome constitutes a violation of a learned action-outcome contingency, without resulting in a change in outcome value. For example, turning on a faucet and having no water come out indicates a change in outcome probability; however, the value of the outcome itself does not change. Thus although delivery of an unpleasant outcome and omission of an expected rewarding outcome are both less favorable than receipt of a reward, they differ in terms of the information they provide about outcome value and probability.

A network consisting of the striatum and its afferents from midbrain, prefrontal, and limbic structures has been implicated in various aspects of feedback-based decision making (Doya 2008; Rushworth and Behrens 2008). The head of the caudate nucleus has been shown to be necessary for acquiring action-outcome associations. Its activity in humans is modulated by perceived action-outcome contingency (Tricomi et al. 2004; Yin et al. 2005). The nucleus accumbens, on the other hand, is required for the modulation of action vigor by motivational signals and is active during presentations of rewarding stimuli (de Borchgrave et al. 2002; O’Doherty et al. 2004). It is unclear, however, whether these striatal structures differentially process different forms of negative feedback.

In this fMRI experiment, we investigated the role of the striatum in processing value-based and probability-based feedback. Participants were initially trained to associate fractal cues with food rewards. Then, value was altered through a selective satiety procedure, which devalued one food outcome. Probability of outcome receipt was altered by omitting the outcome on a majority of trials presented in a second phase of the experiment. This allowed us to examine brain activity related to receipt and omission of expected rewards and devalued outcomes.
Table 1. Fractal-outcome probabilities in training phase and test phase

<table>
<thead>
<tr>
<th></th>
<th>Outcome 1</th>
<th>Outcome 2</th>
<th>No Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Fractals 1–4</td>
<td>75%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>Fractal 5</td>
<td>4%</td>
<td>8%</td>
</tr>
<tr>
<td>Test</td>
<td>Fractals 1 and 3</td>
<td>75%</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>Fractals 2 and 4</td>
<td>9%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>Fractal 5</td>
<td>3%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Outcome 1 = M&Ms for fractals 1 and 2, Goldfish for fractals 3 and 4; vice versa for outcome 2 (see Fig. 1). One outcome is devalued prior to the test phase.

METHODS

Participants

Twenty-five participants were initially recruited for this study (16 women, 9 men; mean age = 22.4 yr, SD = 2.4 yr). Four participants were excluded from analysis because of excessive motion during scanning (n = 1), signal-to-noise ratio or mean intensity abnormality (n = 2), or equipment malfunction during session (n = 1). The final behavioral and neuroimaging analyses were conducted on 21 participants (14 women, 7 men; mean age = 22.3 yr, SD = 2.5 yr). Participants responded to a posted advertisement, and all gave written informed consent. All participants were prescreened to ensure that they were not dieting and that they enjoyed eating chocolate and cheddar crackers. Average weight for female subjects was 123.4 lb (SD = 17.42 lb, range = 105–166 lb) and for male subjects was 160.33 lb (SD = 24.94 lb, range = 130–200 lb). Since the experiment involved food consumption, the eating attitudes test (EAT-26; Garner et al. 1982) was administered prior to the experiment, which indicated no eating disorders in any of the participants [mean score: 4.2 (SD); range: 0–10]. All scores were under the 11-point cutoff, a threshold for screening that ensures the exclusion of anyone at high risk for an eating disorder (Orbitello et al. 2006). Participants were asked to fast for at least 4 h prior to the experiment, although they were allowed to drink water. The experiment was approved by the Institutional Review Board of the University of Medicine and Dentistry of New Jersey as well as the Institutional Review Board of Rutgers University.

Experimental Procedure

At the beginning of the experiment, participants were asked to rate on a scale from 0 to 5 how pleasant they would find eating M&Ms (Mars, McLean, VA) and cheddar-flavored Goldfish crackers (Pep-peridge Farm, Norwalk, CT). Then they performed the training phase of a computerized learning task. During each trial, a fractal image was shown on the screen for 1.5 s, along with a row of three squares directly below it. One of the three squares was highlighted in yellow, indicating which button press was the correct one for that particular trial (1, 2, or 3). Subjects were instructed that if they pressed this response button on the keyboard within the 1.5-s time interval, they would have the chance to earn a food reward. After a jittered interstimulus interval of 1–8 s, if the button was pressed, either a picture of an M&M or a Goldfish cracker would be shown (indicating a food reward of the corresponding type, to be consumed after the training phase) or a Ø symbol was shown, indicating no reward. These outcomes were probabilistically associated with the five fractal stimuli, in the proportions shown in Table 1. Specifically, two of the fractal images were more likely to be associated with Goldfish rewards and two were more likely to yield M&Ms. In all cases in which no button was pressed in the allotted time, the Ø symbol was shown. A jittered intertrial interval (ITI) of 1–8 s followed each trial.

The total number of M&Ms and Goldfish earned was in proportion to the number of times these outcomes were displayed, and after 120 training trials (24 for each fractal), the subject was given his/her earnings to eat.

After this training phase, participants consumed their earned food rewards (generally, between 5 and 7 Goldfish crackers and M&Ms). In addition, one of the two foods was randomly chosen to be devalued through a selective satiety procedure (e.g., Gottfried et al. 2003; Tricomi et al. 2009), in which subjects were asked to eat that food until it was no longer pleasant to them. Thus participants ate only a small amount of one food, so that its subjective value remained high, whereas they ate much more of the other food, so that its subjective value would decrease through satiation. To encourage subjects to eat a large amount of the food to be devalued, they were provided with ~1 cup of this food in a bowl in front of them. Most subjects consumed this amount, although a few consumed more or less than this. In every case, participants ate more than twice as much of the devalued food than they ate of the still-rewarded food. After the subjects finished eating, they rated the pleasantness of the two foods again. Then, they performed the test phase of the experiment during acquisition of fMRI data. In this phase, the same basic task was utilized; however, the outcome for two of the fractals was omitted on a majority of trials (see Table 1). A jittered ITI of 1–8 s between trials was used to aid in the estimation of the BOLD signal produced on each trial. Participants were instructed to maintain responding for all conditions, and since the fractal-outcome associations were probabilistic, there was a possibility of reward on every trial. Trials were divided into four runs of ~7.5 min each, and presentation of the five trial types was randomized (32 trials of each type; 160 trials overall). The BOLD responses for the following conditions could then be compared: 1) receipt of expected reward, 2) omission of expected reward, 3) receipt of devalued outcome, 4) omission of devalued outcome, and 5) expected omission of outcome (see Fig. 1). Participants were told that at the end of the study they would be asked to eat the amount of each food reward earned during the task.

fMRI Data Acquisition

A 3-T Siemens Allegra head-only scanner and a Siemens standard head coil were used for data acquisition at the University of Medicine and Dentistry of New Jersey. Anatomical images were acquired with a T1-weighted protocol (256 × 256 matrix, 176 1-mm sagittal slices)
Functional images were acquired with a single-shot gradient echo EPI sequence (TR = 2,500 ms, TE = 25 ms, flip angle = 80°, FOV = 192 × 192 mm, slice gap = 0 mm). Forty-three contiguous oblique-axial slices (3 mm × 3 mm × 3 mm voxels) were acquired in an oblique orientation of 30° to the anterior commissure-posterior commissure (AC-PC) axis, which reduces signal dropout in the ventral prefrontal cortex relative to AC-PC aligned images (Deichmann et al. 2003).

Data Analysis

Analysis of imaging data was conducted with Brain Voyager QX software, version 2.0 (Brain Innovation, Maastricht, The Netherlands). The data were initially corrected for motion and slice scan time by cubic spline interpolation. Furthermore, spatial smoothing was performed with a three-dimensional Gaussian filter (8-mm FWHM), along with voxelwise linear detrending and high-pass filtering of frequencies (3 cycles per time course). Structural and functional data of each participant were then transformed to standard Talairach stereotaxic space (Talairach and Tournoux 1988).

After preprocessing, the Talairach-transformed fMRI data were analyzed with a random-effects general linear model (GLM). The onsets of each cue (fractal presentation) and outcome event were modeled as stick functions and then convolved with a canonical hemodynamic response function to create regressors of interest for the different conditions: 1) receipt of expected reward, 2) omission of expected reward, 3) receipt of devalued outcome, 4) omission of devalued outcome, and 5) expected omission of outcome. We conducted pairwise comparisons between each of these conditions, both at presentation of cue and at presentation of outcome. Regressors of no interest were also generated using the realignment parameters from the image preprocessing to further correct for residual subject motion. Missed trials and trials in which feedback was incongruent with the dominant cue-outcome contingencies (e.g., when omission of reward followed the reward cue) were modeled as confound predictors.

A 2 × 2 ANOVA with time of task (first or second half) and receipt of reward (reward or reward omission) as factors was performed at the time of the cue, in order to check for learning of fractal-outcome contingencies over time. Additional analyses focused on identifying regions that showed differential responses to the different fractal cues and different types of outcome events. In addition to generation of whole brain statistical maps, a priori regions in the caudate and nucleus accumbens were selected for testing; parameter estimates were also extracted from these regions to graph activation across conditions. These anatomically based regions of interest (ROIs) ensure statistical independence when comparing parameter estimates across all conditions (Kriegerkorte et al. 2009), whereas the whole brain analyses show the full extent of the brain regions exhibiting significant differences between conditions. In the caudate, cubic regions of 8 mm³ were centered at x = ±11, y = 11, z = 8, based on the average of coordinates reported elsewhere (Delgado et al. 2000, 2004; Tricomi et al. 2004; Zink et al. 2004). The nucleus accumbens regions were the same size and were centered at x = ±10, y = 8, z = −4; these coordinates have been used to define nucleus accumbens ROIs in previous studies (Bischoff-Grethe et al. 2009; Cools et al. 2002), based on Talairach atlas location and peak activation coordinates from prior work (Breiter et al. 2001; Delgado et al. 2000; Knutson et al. 2001). These regions were selected because of their previously evidenced roles in the processing of rewarding feedback. For example, the nucleus accumbens region has shown a specific role for processing of the valence of a stimulus (Knutson et al. 2005), while the caudate nucleus region has been shown to encode receipt of large versus small reward (Delgado et al. 2003).

A complementary analysis included the trialwise probability of reward receipt as a parametric modulator. Thus this analysis identified brain regions that demonstrated an increase in BOLD signal at the time of feedback, correlating with the probability that the valued outcome (i.e., the still-rewarding food) would be received on any given trial. This probability value was calculated separately for each fractal, based on the reinforcement history of that fractal cue. Therefore, the probability value on any given trial was the running average of the reinforcement of the fractal cue up until that trial, with 0 indicating no receipt of reward and 1 indicating receipt of reward.

Finally, we estimated an additional model that included the expected value as a parametric modulator of the cue regressor and the prediction error as a parametric modulator of the outcome regressor. These models were calculated with the following equations: B(1) = 1 for cues associated with the still-valued food, or 0 otherwise; and for all t > 1, PE(t) = V(t) − B(t); B(t + 1) = B(t) + αPE(t), where PE is the prediction error, t is the trial number, V(t) = 1 for still-valued outcomes or 0 otherwise, α = 0.65, and B is the expected value. The expected value was calculated separately for each fractal cue. Since the task did not involve acquisition of behavioral choice data, the α value could not be calculated from behavioral data. Instead, the fMRI data from the first five subjects were modeled iteratively, using α values ranging from 0.1 to 0.7 in steps of 0.05 (cf. Hare et al. 2008); α = 0.65 was found to provide the best fit to the data from these subjects, so this value was used for the complete data set.

Throughout this report, all t-tests are two-tailed. For our whole brain analyses, all significant clusters were identified at P = 0.005 and withstood a contiguous cluster threshold of 5 voxels (i.e., >135 mm³). The cluster threshold was used to correct for multiple comparisons and was determined with the Cluster Threshold Estimator plug-in in BrainVoyager QX, which identifies the threshold at which the mapwise probability of a false detection (i.e., type I error rate, or corrected threshold) remains lower than 0.05. As this tool provides information on the type I error rate across the whole brain, it constitutes a principled correction and is compliant with recent recommendations to avoid arbitrary cluster thresholds (Bennett et al. 2009).

RESULTS

Behavioral Results

Participants indicated the pleasantness of each food type both before and after the selective satiety procedure on a Likert scale [ranging from 0 (very unpleasant) to 5 (very pleasant); see Table 2]. A 2 × 2 ANOVA of the Likert scale ratings with time (before/after selective satiety procedure) and food type (valued/devalued) as factors revealed a significant interaction (F1,20 = 40.4, P < 0.0001). Post hoc t-tests confirm that after the devaluation procedure pleasantness ratings of the devalued food decreased significantly more than those for the valued food (t20 = 8.698, P < 0.0001). This cannot be attributed to a general satiety or decrease in pleasantness over the training phase, because the pleasantness ratings for the valued food did not change significantly over time (t20 = −0.77, P = 0.45).

Table 2. Pleasantness ratings of foods before and after selective satiety procedure

<table>
<thead>
<tr>
<th>Food Type</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presatiety procedure pleasantness ratings (0–5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reward</td>
<td>4.19</td>
<td>0.93</td>
</tr>
<tr>
<td>Devalued</td>
<td>4.10</td>
<td>0.83</td>
</tr>
<tr>
<td>Post satiety procedure pleasantness ratings (0–5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reward</td>
<td>4.33</td>
<td>0.73</td>
</tr>
<tr>
<td>Devalued</td>
<td>1.86</td>
<td>1.46</td>
</tr>
</tbody>
</table>

Values are pleasantness ratings of foods before and after the selective satiety procedure, on a scale of 0 (very unpleasant) to 5 (very pleasant).

J Neurophysiol • doi:10.1152/jn.00086.2014 • www.jn.org
Table 3. Regions showing a significant cue condition (reward cue vs. reward omission cue) by task half interaction

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size, mm³</th>
<th>Max t</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precentral gyrus</td>
<td>233</td>
<td>25.41</td>
<td>x 56 y 4 z 12</td>
</tr>
<tr>
<td>Ventral putamen/caudate nucleus</td>
<td>766</td>
<td>25.32</td>
<td>x 20 y 4 z -3</td>
</tr>
<tr>
<td>Medial orbitofrontal cortex</td>
<td>236</td>
<td>15.85</td>
<td>x -1 y 31 z -22</td>
</tr>
<tr>
<td>Posterior insula</td>
<td>229</td>
<td>34.70</td>
<td>x -34 y -14 z -6</td>
</tr>
</tbody>
</table>

Regions are those identified as showing a significant cue condition (reward cue vs. reward omission cue) by task half (1st half vs. 2nd half) interaction (P < 0.05, corrected).

fMRI Results

Results at cue. We analyzed the cue data with respect to the type of outcome each fractal predicted. We did not identify any significant effects in the striatum for any pairwise contrasts of the experimental conditions when collapsing across trials from the entire length of the scan. However, since the fractal-outcome contingencies during the scan were altered from the training phase, we expected that BOLD responses at the time of the cue might change over time, as the new contingencies were learned. We identified significant effects in the striatum (ventral putamen, extending dorsally to the caudate) and medial orbitofrontal cortex (OFC) than the reward omission cue did in the second half of the task (F_{1,20} > 9.9, P < 0.05, corrected). The activation peaks are listed in Table 3, and a cluster-threshold corrected map of this ANOVA is featured in Fig. 2.

Additionally, the regions identified as showing a significant effect of expected value at the time of cue presentation are listed in Table 4. The right putamen, as well as several cortical regions, showed this effect. We did not find significant effects in our a priori ROIs in the nucleus accumbens (left: t_{20} = 0.287, P = 0.77; right: t_{20} = 0.428, P = 0.67) or caudate (left: t_{20} = 0.068, P = 0.94; right: t_{20} = 0.699, P = 0.49), however.

Results at outcome. We identified significant effects in the striatum for three pairwise contrasts at the time of outcome presentation (namely, reward receipt vs. devalued outcome, reward receipt vs. expected omission of an outcome, and reward receipt vs. reward omission). The regions identified as showing a significant effect for each of the outcome-related contrasts (reward vs. devalued outcome, reward vs. expected omission of an outcome, and reward receipt vs. reward omission) are listed in Table 5. Of interest, the reward > devalued outcome contrast identified a region in the right ventral striatum, the reward > expected omission feedback contrast identified regions in the right caudate nucleus and right putamen, and the reward > reward omission contrast identified a region in the left putamen. We also note that an additional region in the left medial prefrontal cortex was identified as showing a reward > devalued outcome effect at P < 0.005 (peak Talairach coordinates: x = -7, y = 45, z = 5), but this region was only 4 voxels in size and therefore did not survive our cluster-threshold correction.

Figures 3 and 4 show cluster-threshold corrected maps depicting the striatal regions we identified at the time of the outcome. Figure 3 depicts a right nucleus accumbens region that showed a significant effect for the reward > devalued prediction error at outcome.

Table 4. Regions showing significant parametric modulation of cue by expected value and of outcome by prediction error

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size, mm³</th>
<th>Max t</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putamen</td>
<td>178</td>
<td>4.90</td>
<td>x 26 y 13 z 6</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>131</td>
<td>4.15</td>
<td>x -7 y -47 z 60</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>168</td>
<td>4.88</td>
<td>x -13 y 52 z 18</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>391</td>
<td>4.28</td>
<td>x -40 y 46 z 9</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>140</td>
<td>4.93</td>
<td>x -46 y -53 z 27</td>
</tr>
<tr>
<td>Prediction error at outcome</td>
<td></td>
<td></td>
<td>x -5.03 y -43 z 28</td>
</tr>
</tbody>
</table>

Regions are those identified as showing significant parametric modulation of cue by expected value and of outcome by prediction error (P < 0.05, corrected).

Also, the ratings of both foods at the beginning of the task did not differ, demonstrating that, overall, there were no individual preference biases toward either of the food items (R = 0.491; P = 0.63). Therefore, the ratings were consistent with the interpretation that the subjects became sated on the devalued but not the valued food.
outcome contrast \( (P < 0.05, \text{corrected}) \). Figure 4A depicts a right caudate nucleus region that showed a significant effect for the reward > expected omission contrast \( (P < 0.05, \text{corrected}) \). Also shown in Fig. 5 are plots of the regression coefficients for all conditions in the caudate and nucleus accumbens, which were extracted based on a priori cubic regions of 8 mm\(^3\), centered on coordinates from previous work (Bischoff-Grethe et al. 2009; Cools et al. 2002; Delgado et al. 2000, 2004; Tricomi et al. 2004; Zink et al. 2004). Since the regression coefficients shown in these graphs were made with a priori ROIs, they are not subject to a bias that could be introduced by displaying the coefficients based on a particular statistical contrast. The graphs of the regression coefficients based on our significant clusters of activation look very similar, however.

In the a priori right nucleus accumbens region, rewarding outcomes yield significantly greater activity than devalued outcomes \( (t_{20} = 3.85; P < 0.001) \). This region also shows significantly greater activation for rewarding outcomes versus expected omissions \( (t_{20} = 4.22; P < 0.001) \). Similarly, the left nucleus accumbens shows a significant difference between rewarding and devalued outcomes \( (t_{20} = 2.87, P = 0.009) \). This effect is also significant in the left caudate nucleus \( (t_{20} = 3.32; P = 0.003) \). The right caudate, however, shows an

Table 5. *Regions showing significant differences between outcome conditions*

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster size, mm(^3)</th>
<th>Max ( t )</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reward feedback &gt; devalued feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nucleus accumbens</td>
<td>148</td>
<td>4.25</td>
<td>11 10 −3</td>
</tr>
<tr>
<td>Reward feedback &gt; expected omission feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>296</td>
<td>4.30</td>
<td>56 −14 27</td>
</tr>
<tr>
<td>Right cerebellum, posterior lobe (declive)</td>
<td>229</td>
<td>5.10</td>
<td>44 −65 15</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>200</td>
<td>5.25</td>
<td>29 58 3</td>
</tr>
<tr>
<td>Right cerebellum, posterior lobe (cerebellar tonsil)</td>
<td>381</td>
<td>4.64</td>
<td>35 −47 −36</td>
</tr>
<tr>
<td>Right cerebellum, posterior lobe (declive)</td>
<td>139</td>
<td>4.46</td>
<td>32 −71 −21</td>
</tr>
<tr>
<td>Putamen</td>
<td>188</td>
<td>4.94</td>
<td>26 13 9</td>
</tr>
<tr>
<td>Caudate nucleus</td>
<td>547</td>
<td>5.62</td>
<td>5 7 6</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>227</td>
<td>4.40</td>
<td>−4 23 36</td>
</tr>
<tr>
<td>Posterior cingulate gyrus</td>
<td>197</td>
<td>4.90</td>
<td>−1 29 24</td>
</tr>
<tr>
<td>Angular gyrus</td>
<td>188</td>
<td>5.07</td>
<td>−40 −59 36</td>
</tr>
<tr>
<td>Anterior lobe of cerebellum</td>
<td>1350</td>
<td>6.22</td>
<td>−43 −53 −27</td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>691</td>
<td>5.60</td>
<td>−49 74 −6</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>475</td>
<td>5.47</td>
<td>−49 56 −12</td>
</tr>
<tr>
<td>Reward receipt &gt; reward omission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior lobe of cerebellum</td>
<td>382</td>
<td>4.87</td>
<td>35 −47 24</td>
</tr>
<tr>
<td>Inferior occipital gyrus</td>
<td>299</td>
<td>4.88</td>
<td>26 −92 −6</td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>277</td>
<td>5.39</td>
<td>23 −77 −9</td>
</tr>
<tr>
<td>Anterior lobe of cerebellum</td>
<td>159</td>
<td>5.28</td>
<td>29 −38 −18</td>
</tr>
<tr>
<td>Putamen</td>
<td>291</td>
<td>5.38</td>
<td>−19 4 9</td>
</tr>
<tr>
<td>Precuneus</td>
<td>149</td>
<td>5.57</td>
<td>−37 71 36</td>
</tr>
<tr>
<td>Anterior lobe of cerebellum</td>
<td>540</td>
<td>4.58</td>
<td>−37 −50 21</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>246</td>
<td>4.45</td>
<td>−40 −56 24</td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>382</td>
<td>4.78</td>
<td>−49 7 33</td>
</tr>
</tbody>
</table>

Fig. 3. The right nucleus accumbens showed significantly greater activation after rewarding outcomes than devalued outcomes \( (P < 0.05, \text{corrected}) \). Image is left-right reversed.

Fig. 4. The right caudate nucleus (A) showed significantly greater activation after rewarding outcomes than expected omission of feedback \( (P < 0.05, \text{corrected}) \). The right caudate (B) and the right central OFC (C) also show a significant parametric modulation of activation by probability of reward receipt; as the probability of reward increases, these regions are activated more. Images are left-right reversed.

Figs. 3, 4.
increase in BOLD signal at the rewarding feedback and a decrease at the expected omission of an outcome, producing a significant difference between these conditions ($t_{20} = 4.30$; $P < 0.001$).

The results of the $t$-tests above suggest that the nucleus accumbens may be more sensitive to information about outcome valence (reward vs. devalued), while the caudate may be more sensitive to reward probability. However, an expected reward omission may be interpreted as both an outcome of low value (a null sign) and a low probability of reward receipt. Furthermore, a devalued outcome both has low motivational value and is associated with a low probability of reward. Therefore, we performed a complementary analysis, identifying brain regions that showed a significant parametric increase in BOLD signal at the time of outcome correlating with the trialwise probability of reward receipt; as the probability of reward increases, these regions are activated more (Table 6). We found a significant effect in the right caudate nucleus (Fig. 4; $P < 0.05$, corrected). A region in the right central OFC was also significantly activated in this analysis. Although this analysis accounted for trialwise fluctuations in probability, the average probability of reward receipt for each condition is included in Fig. 5, to help illustrate this effect.

Finally, as shown in Table 4, the only region identified as showing an effect of prediction error at the time of feedback presentation was the left inferior frontal gyrus, which showed larger decreases in activation as the prediction error increased. No significant effects were found for our a priori regions in the caudate (left: $t_{20} = 0.027$, $P = 0.98$; right: $t_{20} = 0.344$, $P = 0.73$) and nucleus accumbens (left: $t_{20} = 0.185$, $P = 0.86$; right: $t_{20} = 0.149$, $P = 0.88$).

**DISCUSSION**

Our results provide evidence that, at the time of outcome, the ventral striatum carries information about the value of the outcome, which is dependent on the subject’s current motivational state. Meanwhile, the dorsal striatum (caudate and putamen) carries a signal that includes information about both the value of the stimulus and reward probability. These results build on previous work examining reward representations in the human brain. It is now well known that both the ventral and dorsal striatum show effects of valence, with increases in activation for rewards, such as monetary gains, and decreases in activation for punishments, such as monetary losses (Coricelli et al. 2005; Delgado et al. 2000, 2003; Yacubian et al. 2005; Delgado et al. 2003).

**Table 6. Regions showing significant parametric modulation of feedback by probability of reward receipt**

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster size, mm$^3$</th>
<th>Max $t$</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precentral gyrus</td>
<td>163</td>
<td>5.99</td>
<td>35 $-$ 5</td>
</tr>
<tr>
<td>Anterior middle frontal gyrus</td>
<td>193</td>
<td>5.00</td>
<td>32 58 9</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>334</td>
<td>5.73</td>
<td>$-$1 $-$ 29</td>
</tr>
<tr>
<td>Angular gyrus</td>
<td>343</td>
<td>4.72</td>
<td>$-$40 $-$ 59</td>
</tr>
<tr>
<td>Anterior lobe of cerebellum</td>
<td>488</td>
<td>4.94</td>
<td>$-$43 $-$ 53</td>
</tr>
</tbody>
</table>

Regions are those identified as showing a significant parametric modulation of feedback by probability of reward receipt ($P < 0.05$, corrected).

*J Neurophysiol* • doi:10.1152/jn.00086.2014 • www.jn.org
absence of a reward and the dorsal putamen showed a differ-
tial response to rewards and the expected

context-dependent subjective value). This finding builds on previ-
ous work focusing on effects of satiety on neural re-

Activation to receipt of appealing food outcomes relative to
devalued food outcomes. Many studies that have previ-
ously investigated value representations in the brain have focused on
rewards of varying magnitude, such as small and large mone-
tary rewards (Breiter et al. 2001; Delgado et al. 2003; Guitart-
Masip et al. 2011; Knutson et al. 2005; Smith et al. 2009;
Tobler et al. 2007; Yacubian et al. 2006) and in reinforcement learn-
ing processes (Daw and Doya 2006; Montague et al. 2006).

Striatal Response to Changes in Cue-Outcome Contingency

After the training phase of this experiment, the food out-
comes associated with two of the fractal cues were omitted.
Although we did not find significant cue-related effects in the
striatum when collapsing across the whole fMRI phase of the
experiment, we found evidence that the striatum adjusted its
response to the cues based on the change in cue-outcome
contingency. At the time of cue presentation, activation of the
putamen was parametrically modulated by the expected value
of the upcoming outcome. Additionally, in the second half of
the task only, cues predicting an upcoming reward produced
greater activation in the striatum than cues predicting reward
omission. These findings provide evidence that signals in the
striatum reflect the known value of an upcoming outcome as
well as the value of a received outcome. That is, information
about the value of the upcoming outcome, as well as about the
cue’s reinforcement history, is reflected in striatal activity at
the time of the cue. This fits with previous evidence indicating
a role for the striatum in anticipation of upcoming rewards
(Bjork and Hommer 2007; Ernst et al. 2004; Galvan et al.
2005; Gottfried et al. 2003; Haruno and Kawato 2006; Knutson
et al. 2005; Yacubian et al. 2006) and in reinforcement learning
processes (Daw and Doya 2006; Montague et al. 2006).

Subjective Value Coding in Ventral Striatum

We found that the nucleus accumbens showed selective
activation to receipt of appealing food outcomes relative to
devalued food outcomes. Many studies that have previously
investigated value representations in the brain have focused on
rewards of varying magnitude, such as small and large mone-
tary rewards (Breiter et al. 2001; Delgado et al. 2003; Guitart-
Masip et al. 2011; Knutson et al. 2005; Smith et al. 2009;
Tobler et al. 2007; Yacubian et al. 2006). Importantly, our
study extends these findings by showing that the ventral stria-
tum responds differently to the same outcome (a particular
food), depending on how pleasant the participant currently
finds that outcome to be (i.e., depending on the outcome’s
context-dependent subjective value). This finding builds on
previous work focusing on effects of satiety on neural re-

Probability Coding in Dorsal Striatum

We found that the right caudate nucleus and putamen
showed a differential response to rewards and the expected
absence of a reward and the dorsal putamen showed a differ-
etial response to rewards and omissions of reward. Addition-
ally, the caudate and the putamen both showed a parametric
increase in signal at the outcome phase as the probability of
reward receipt increased. To show this pattern of results, the
dorsal striatum would need to be sensitive both to the current
value of the outcome and its probability. Therefore, the dorsal
striatum may be an important locus for integrating value and
probability information, which would be useful for action
planning. Haber et al. (2000) have proposed a hierarchy of
information flow from the ventromedial to the dorsolateral
striatum (i.e., from nucleus accumbens to caudate to putamen),
via striatonigristriatal projections. In this way, value-related
signals from the ventral striatum may influence the dorsal
striatum, which may then integrate this information with
information about probability.

These findings are in line with previous research showing
effects of reward schedule (Tanaka et al. 2008) and expected
value (i.e., reward magnitude × probability) on activation of
the dorsal striatum (Hsu et al. 2005; Tobler et al. 2007). The
caudate has also been shown to be sensitive to action-outcome
contingency (Elliott et al. 2004; O’Doherty et al. 2004;
Tanaka et al. 2008; Tricomi et al. 2004; Zink et al. 2004). In the
expected omission condition there is a low probability that an
action will lead to reward, and thus a low action-outcome
contingency, whereas in the reward condition there is a high
contingency between the button press and the outcome. Thus
the dorsal striatum may play an important role in cementing the
learned relationship between a given action and a desired
outcome (Balleine et al. 2009; Haruno and Kawato 2006;
Samejima et al. 2005).

The representation of outcome probability is necessary to
make well-informed future decisions. For example, outcomes
of low probability may not be worth investing effort to achieve.
Alternatively, when outcome probabilities are not fixed, an
increase in the allocation of effort might be necessary to
increase the probability of obtaining a goal that is difficult to
achieve. Indeed, individual differences in dopamine responsiv-
ity (i.e., receptor availability) in the caudate, measured with
PET, have been linked to a willingness to expend effort to earn
a low-probability reward (Treadway et al. 2012). When action-
reward contingency is high, anticipated effort is also likely to
be high; that is, participants will likely put more effort into
responding when they believe that their effort will result in
increased reward. Accordingly, the putamen has been found
be active when anticipated effort is high (Kurniawan et al.
2013), and also when participants choose to expend low effort
for reward versus high effort (i.e., when their action is more
likely to produce a positive outcome; Kurniawan et al. 2010).
In addition, Croxson et al. (2009) found that the putamen
plays a specific role in encoding anticipated effort, while the
ventral striatum plays a specific role in encoding expected
reward. In our study, we did not vary levels of effort required to

J Neurophysiol • doi:10.1152/jn.00086.2014 • www.jn.org
after their omission compared with less predictable outcomes, demonstrating one way in which information about outcome probability is used to govern behavior (Dudley and Papini 1997).

It is of interest that this effect of reward probability occurred at the time of outcome presentation, rather than at the presentation of the cue. The caudate is not showing significant effects of expected value or prediction error in this experiment. Rather than coding for the difference between the actual and expected outcomes, it is instead responding more to receipt of predicted rewards than predicted lack of reward. Prediction error responses in the striatum tend to be found in the ventral striatum or the putamen, rather than the caudate (Abler et al. 2006; Hare et al. 2008; McClure et al. 2003; O’Doherty et al. 2003; Yacubian et al. 2006). Since prediction error incorporates information about how unexpected an outcome is, the ventral striatum must have access to information about outcome probability. It may be that the unexpectedness of an outcome contributes to how subjectively rewarding it is, which in turn influences ventral striatum activity, whereas the caudate signal reflects the strength of action-outcome contingency, which increases with increasing probability of reward.

**Laterality of Effects**

The results of our whole brain analyses were significant only on the right side for both the nucleus accumbens and the caudate nucleus. In the nucleus accumbens, this seems to be due to a thresholding issue, since a similar pattern of results was observed in the left nucleus accumbens region that was defined a priori. However, we note that our results were markedly different in the right and left caudate nuclei. Our a priori defined left caudate region showed a pattern more similar to that observed in the nucleus accumbens than in the right caudate, with differential activation to rewarding and devalued outcomes. If there is a continuum with the ventromedial striatum more sensitive to subjective value and the dorsolateral striatum more sensitive to probability, it is possible that since the caudate is in the middle of this continuum we happened to find more evidence for value coding in the left caudate and probability coding in the right caudate. It is also interesting to note that, as shown in Fig. 5, the differential effect of rewarding and devalued outcomes seems to be driven by a decrease in activation to devalued outcomes for both left nucleus accumbens and left caudate, whereas this effect seems to be driven by an increase in activation to rewarding outcomes for the right nucleus accumbens. Previous work on lateralization of reward-related activity in the striatum has shown mixed results, with some studies showing higher dopamine binding and synthesis and a stronger dopaminergic reward response in the right striatum than the left (Cannon et al. 2009; Martin-Soelch et al. 2011; van Dyck et al. 2002; Vernaalen et al. 2007) and other work suggesting stronger left-lateralized activity supporting approach processes (Murphy et al. 2003; Tomer et al. 2014). Therefore, additional research will be necessary to determine the nature of potential laterality differences in the striatum, and especially whether there are laterality differences in processing reward value versus probability.

**Role of Orbitofrontal Cortex**

Previous work has suggested that the prefrontal cortex, especially the medial OFC, plays an important role in representing reward value in the brain (Grabenhorst and Rolls 2011; Kringelbach et al. 2003; O’Doherty 2007; Peters and Buchel 2010; Rudenga and Small 2013). The OFC has been shown to track decision value, goal value, and outcome value (de Wit et al. 2009; Hare et al. 2008; Peters and Buchel 2010). In this experiment, we found that the medial OFC, like the striatum, adjusted its activity to the fractal cues such that in the second half of the task cues predicting reward produced greater activation than cues predicting reward omission. Although we did not observe significantly greater activation in the OFC for reward receipt relative to devalued food outcomes, it is possible that this null result may be due to a thresholding issue, since we did identify a region in the left medial prefrontal cortex that showed an effect of reward versus devalued outcome receipt that did not withstand our correction for multiple comparisons. Additionally, we observed a significant effect of reward probability in the right central OFC. Previous research has indicated that the central OFC may play a role in stimulus-outcome learning (O’Doherty 2007; Valentin et al. 2007). Its sensitivity to reward outcome probability in our experiment may thus reflect a role in learning how strongly each of the fractal stimuli was associated with a rewarding outcome.

**Conclusion**

To adaptively make decisions, we must be able to track the subjective value of an outcome in a way that is sensitive to changes in motivational state. In addition, we must be able to register the probability of a given outcome occurring. Our results suggest that the ventral striatum is especially sensitive to the motivational relevance of an outcome, and it does not seem to track information about the probability of an outcome occurring. Activity in the dorsal striatum, however, appears to correlate with the probability of reward receipt at the time of outcome.

By using food rewards, we were able to change the subjective value of the outcome without changing the outcome itself; the magnitude of the reward did not change, but its value did. Additionally, food outcomes allow for the delivery of an outcome of low value without taking away something of high value. For example, receipt of a low-value food is an experience different from loss of money. Thus our experiment emphasizes the subjective nature of valuation and its dependence on motivational state and shows that the brain’s reward circuitry is sensitive to these subjective factors.

The reduction of an outcome’s value and the omission of a previously well-predicted reward can both be thought of as “negative” outcomes, or punishments. Our results suggest, however, that these two types of negative outcomes are not isomorphic but rather that they affect brain activity in distinct ways. Thus impairments in value and probability representation might result in distinct types of behavior. For example, an overweighting of reward value relative to probability might lead to compulsive gambling in an effort to achieve a reward of high value that has a very low probability of occurrence, whereas an overweighting of probability might lead one to avoid taking risks necessary to achieve a desired goal because of a low probability of success. A successful integration of
information about value and probability, however, allows us to make effective decisions and allocate an appropriate amount of effort to achieve our desired goals.

ACKNOWLEDGMENTS

The authors thank Mauricio Delgado and Mike Shiflett for their valuable comments.

Present address of K. Lempert: Dept. of Psychology, New York University, New York, NY 10003.

GRANTS

This work was supported by National Institute on Drug Abuse Grant R15 DA-029544 to E. Tricomi. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse or the National Institutes of Health.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: E.T. and K.M.L. interpreted results of experiments; E.T. and K.M.L. prepared figures; E.T. and K.M.L. edited and revised manuscript; E.T. and K.M.L. approved final version of manuscript; K.M.L. performed experiments; K.M.L. analyzed data.

REFERENCES


J Neurophysiol • doi:10.1152/jn.00086.2014 • www.jn.org