Dissociable Neural Responses in Human Reward Systems

Rebecca Elliott, Karl J. Friston, and Raymond J. Dolan

Received March 30, 2000; revised May 23, 2000; accepted May 23, 2000.

Copyright © 2000 Society for Neuroscience 0270-6474/00/206159-07$15.00/0

Most adaptive behavior is driven by basic survival needs such as food, drink, and sex that are experienced as rewards and by avoiding aversive situations that are experienced as punishments. The reinforcing effects of these behavioral outcomes are mediated through distinct neural mechanisms. In animals, ascending dopaminergic systems have been shown to be critically involved in responses to various reinforcing stimuli, including food and drugs of abuse. The ventral striatum, particularly the nucleus accumbens, is probably the structure most reliably linked to reward-related processes (Wise, 1980; Robbins and Everitt, 1992; Schultz et al., 1993, 1996; Stern and Passingham, 1996), but other structures are also involved, including midbrain regions of ventral tegmental area (VTA) and substantia nigra (Ljungberg et al., 1992; Schultz, 1997), the amygdala (Cadot et al., 1989; Everitt and Robbins, 1992), and regions of the basal forebrain (Arvanitogiannis et al., 1996; Panagis et al., 1997).

Reward in humans has been studied less extensively. However, functional imaging studies using infusions of nicotine (Stein et al., 1998) or cocaine (Breiter et al., 1997) have associated the rewarding effects of these drugs with neural responses in regions including nucleus accumbens, brainstem, amygdala, and prefrontal cortices. More abstract reinforcers also exert powerful motivational effects in humans, as societal preoccupation with gambling and other risk-taking behavior testifies. Other functional imaging studies have associated financial reward with activation of ventral striatum (Koepp et al., 1998), midbrain, thalamic, and prefrontal regions (Thut et al., 1997). In humans, the nonspecific excitement engendered by risk-taking behavior may be as important in maintaining these behaviors as the potential rewards, and it has not been clearly established how these nonspecific effects are expressed in the human brain. Neurepsychological studies suggest that ventral prefrontal regions may be an important interface between cognitive and emotional components of risk-taking behaviors (Damasio, 1994; Bechara et al., 1994). Furthermore, patients with lesions to ventromedial prefrontal regions show pronounced impairments on gambling tasks and fail to show normal task-related autonomic changes (Bechara et al., 1996).

In this experiment, we used functional magnetic resonance imaging (fMRI) to measure neural responses to rewards while subjects performed a simple gambling task. Correct and incorrect responses were associated with financial rewards and penalties, and we assessed the relationship between the level of accumulated gain or loss and regional hemodynamic response. The design also allowed us to consider how this response was modulated by the psychological context in which rewards or penalties were experienced. Our general hypothesis, based on animal studies (Koob, 1992; Robbins and Everitt, 1992, 1996; Schultz et al., 1993; Aosaki et al., 1994) and previous imaging studies of financial reward (Thut et al., 1997; Koepp et al., 1998) was that interconnected regions of the midbrain, striatum, limbic system, and prefrontal cortices would show reward-related activity. Specifically, we predicted responses in ventral striatum, VTA, substantia nigra, amygdala, basal forebrain, prefrontal cortex, or some subset of these regions.

MATERIALS AND METHODS

Experimental paradigm. Subjects were presented with pairs of stimuli depicting playing cards, one red and one black, and were told that on half the trials the red card was correct, and on the other half the black card was correct. The task was to guess the correct card on each trial and respond with a button press. Without the subject’s knowledge, feedback (correct or incorrect) was provided according to a prespecified pseudorandom sequence, irrespective of the actual choices made. At the side of the presentation screen a bar displayed a cumulative “reward” score across all the trials. The height of this bar had direct financial implications; subjects began the experiment with a £10 “stake”. Every correct response was associated with an increase in height of one increment, representing £1, whereas every incorrect response was associated with a decrease of one increment (Fig. 1a). There is no meaningful performance measure in this task because subjects are simply guessing in the absence of information. The variable of interest is the reward level; the task merely provides a realistic context in which rewards and penalties are experienced. At debriefing, subjects all described feelings of pleasure and disappointment in response to rewards and penalties, respectively, although we acknowledge that there were probably individual differences in the extent to which subjects were motivated by reward and penalty, and this potential variation is not explicitly addressed in the present design.

It is important at this point to define the terminology we use in relation to this paradigm. The term “reward” can have various connotations; here “reward level” is defined as the height of the bar (accumulated wins). High levels are likely to induce subjective feelings of pleasure and hedonism, which are facets of reward in this experiment. Subjects reported such feelings at post-scan debriefing. Low levels of the bar (accumulated loss) do not however reflect punishment in the usual sense of the term, although loss of previously gained rewards is clearly a form of negative outcome. We
were not involved in the subsequent regression analysis. Rest blocks were used to model low frequency drift in signal and were not involved in the subsequent regression analysis.

The task was divided into 24 test blocks. Each block had twelve trials of 3.5 sec duration, so that the entire block lasted 42 sec. Test blocks were separated by 42 sec periods of rest during which subjects fixated centrally. At the start of each block, the height of the reward bar maintained the level achieved at the end of the previous block. All the subjects adopted a frequency matching strategy such that they picked each color on approximately 50% of trials (red selected on 51.2% of trials overall, black on 48.8%). In fact, the sequence of outcomes, corresponding to the actual rewards and penalties experienced, derived from the constrained binomial random walk function. This same function determined outcomes in all subjects, who consequently experienced the same sequence of rewards and penalties. The red asterisks mark the point at which blocks of rest were interspersed. These rest blocks were used to model low frequency drift in signal and were not involved in the subsequent regression analysis.

Therefore use the term “penalty” to define the removal of financial rewards. The other experimental factor is change in level of the bar, as distinct from absolute level. We use the term “increasing reward” to describe positive changes and the term “increasing penalty” to describe negative changes.

The actual statistical model used the height of the reward bar, its rate of change, and situations in which the subjects experience a series of wins or losses leading to rapid changes in reward level. It distinguishes between situations in which the subjects experience a series of wins or losses leading to rapid changes in reward level (a winning or losing streak in anecdotal terms) and situations in which the subjects experience a series of wins or losses alternate leading to a more static overall reward level. Finally, we looked at the interaction between the two functions (height and rate of change) to determine those regions sensitive both to absolute reward level and to how fast the level is changing.

The statistical model is thus based on trial-by-trial rewards and penalties, and the analysis represents an event-related characterization of the neural responses to these. The critical extension here is that the response evoked by each event is allowed to vary with reward level and rate of change of reward level, allowing us to look at responses to rewards and penalties specific to different contexts: (1) high accumulated reward or high accumulated penalty; the main effects of level; (2) rapidly increasing or decreasing reward: the main effects of rate of change; and (3) both extreme levels of reward (high or low) and rapid change: the interaction term.

\[ \frac{dH}{dt} \]

where \( H \) is height, and \( t \) is trial. Essentially this function represents the gradient of the smoothed reward level function and provides a measure, not of overall reward level, but of how fast that level is changing over successive trials. It distinguishes between situations in which the subjects experience a series of wins or losses leading to rapid changes in reward level (a winning or losing streak in anecdotal terms) and situations in which the subjects experience a series of wins or losses alternate leading to a more static overall reward level.

Neural responses were measured in nine healthy volunteers scanned using a Siemens Vision system at 2 T to acquire T1-weighted structural images and gradient echo, echoplanar T2*-weighted images with blood oxygenation level-dependent (BOLD) contrast. Functional images were acquired in two runs, each of 240 volumes comprising 48 3 mm axial slices with 3 mm in-plane resolution. For each run, six preliminary “dummy” volumes, to allow for T1 equilibration effects, were acquired and subsequently discarded. Thereafter, volumes were acquired continuously every 4.2 sec so each block of 12 behavioral trials corresponded to 10 scans.

This temporal asynchrony is important because time-locking trials to scans introduces a systematic bias in sampling over peristimulus time. A second function was also used to model the data, derived from the mathematical differential of the height of the bar with respect to trial. This convolution provided a degree of temporal smoothing, which meant that the asynchrony between trials and scans did not pose a problem. The function displayed in Figure 1a was used as a model, and the analysis involved determining in which regions the neural response was well modeled by that function. The function was convolved with the hemodynamic response to take account of the temporal properties of the BOLD signal.

**Figure 1. Cognitive activation paradigm. a shows the task that was used; the pairs of black and red cards to one side of the screen and the cumulative reward bar to the other. b shows the sequence of outcomes, corresponding to the actual rewards and penalties experienced, derived from the constrained binomial random walk function.**

This experiment conforms to a factorial design with the level of the bar and rate of change of level as independently varying factors. The interaction between the two provides a measure of how responses to reward level are modulated when that level occurs in the context of increasing or decreasing reward. This design was motivated by the important psychological consideration that the experience of reward or penalty is a function of the context in which it occurs. Our paradigm enabled us to distinguish activity associated with high reward levels during a “winning streak”, in which reward level is increasing, from activity associated with the same reward levels in situations in which reward level is static or decreasing. The analogous distinction can also be made for the context of “losing streaks.”
In mathematical terms, the statistical parametric maps (SPMs) were based on a multiple-regression analysis, which can be thought of as testing for partial correlations between the neurophysiological time series and the regressors in question. This approach (Friston et al., 1995b;c; Worsley and Friston, 1995) models observed changes at each voxel in terms of the linear sum of a number of continuously varying stimulus functions or regressors. Subject-specific low-frequency confounds were removed (Friston et al., 1995d) in the regression, using the rest blocks to model out these confounds, and global differences were controlled by proportional scaling (Friston et al., 1995a). The significance of the association between the observed time series and one, or a linear combination, of these regressors is tested with the $T$ statistic to give a SPM($T$). In this instance the regressors were the height of the bar, the rate of change of height, and their interaction. The level of the bar and its rate of change are, by definition, orthogonal factors.

Statistical inferences, corrected for the volume analyzed, were based on the theory of random Gaussian fields (Friston et al., 1995c) and used a fixed effects analysis, because we were only making inferences about the normal subjects studied. Two levels of statistical inference were used, in accordance with the established SPM conventions discussed in detail elsewhere (Friston et al., 1995b,c,e). The data were initially thresholded at $p < 0.001$ uncorrected, and regions about which we had an a priori hypothesis were reported at this threshold. For regions about which there was no a priori hypothesis, a more stringent threshold of $p < 0.005$ corrected for multiple comparisons was used, and regions were only reported if they survived at this threshold. The exception to this is in the case of bilateral regions. If a region about which there was no a priori hypothesis was activated at $p < 0.05$ corrected on one side and $p < 0.001$ uncorrected on the other, both are reported. These are the only two thresholds used and all voxels are reported at one or the other threshold, in line with accepted practice. For the purposes of clarity, we use $p < 0.001$ to indicate the uncorrected threshold and *$p < 0.05$ to indicate the corrected threshold throughout this manuscript.

Finally, to facilitate the interpretation of interaction, we used a masking procedure to look separately at interactions in voxels showing (1) a positive relationship with reward level, (2) a negative relationship with reward level, and (3) no relationship with reward level. For masking purposes only, these main effects were thresholded at an inclusive threshold of $p < 0.05$ uncorrected. It should be noted that the voxels reported in the interaction terms were significant at $p < 0.001$ uncorrected or $p < 0.05$ corrected as usual: the masking procedure is descriptive rather than statistical, allowing us to determine which of the voxels significantly activated in the interaction are associated with which main effects. Anatomical localizations were determined by reference to the atlas of Duvernoy (1991) and to structural MRIs of the group.

**RESULTS**

**Responses to high and low reward levels**

Reward level, reflected by the overall height of the bar, was significantly predictive of activity in a region of the right midbrain ($^*p < 0.05$) (Fig. 2a, Table I). With the resolution of our data (smoothed to facilitate intersubject averaging), it was not possible to attribute this activation to a particular structure. As can be seen in Figure 2a, the activation is somewhat above the substantia nigra, below the thalamic nuclei, and lateral to hypothalamic nuclei. There was also a significant activation ($p < 0.001$) of right ventral striatum, lateral to nucleus accumbens (Fig. 2b, Table I). These regionally specific responses are to high levels of reward and can be
formulated in terms of a positive reward signal. A negative relationship between reward level and activity was observed in both hippocampi (\( p < 0.05 \), left; \( p < 0.001 \), right) and parahippocampal gyri (BA 35; \( p < 0.05 \)) (Fig. 2c; Table 1). These responses may reflect a negative reward signal or a positive signal to financial penalty.

There were no regions showing a significant neural response associated with a main effect of change in height.

**Context-dependent responses to financial rewards and penalties**

A key aspect of rewarding situations is how reward-related responses are modulated by psychological context. Context-dependent neural responses of this nature were assessed using the interaction between the absolute height of the bar and rate of change. By testing for this interaction in regions that correlated positively with reward level, we identified neural responses uniquely associated with reward level in the context in which reward was increasing rapidly, a so-called winning streak, relative to situations in which it was not (i.e., stable or decreasing reward levels). These responses were in right anterior medial thalamus (\( p < 0.05 \), bilateral pallidum (\( p < 0.001 \)), and bilateral subgenual cingulate (\( p < 0.001 \)) (Fig. 3).

Similarly, we assessed interactions occurring in regions showing that correlated negatively with reward level. Enhanced activation of both hippocampi (\( p < 0.001 \)) was associated with a high level of penalty in the context of increasing penalty (Table 2). Thus the hippocampal response to removal of reward, or accumulated penalty, became more pronounced when that level was experienced in the context of a losing streak.

**Context-dependent responses to both rewards and penalties**

Although our data show specific neural responses to either reward or penalty that are modulated by psychological context, we also identified associated with both high reward levels during a winning streak and high penalty levels during a losing streak. This analysis identified regions responsive to emotionally salient experiences, in which the salience is congruent, but blind to the valence (i.e., good or bad) of the experience (Table 2). The regions encompassed bilateral orbitofrontal cortex (BA 47), insula, and head of caudate (all \( p < 0.05 \)) (Fig. 4).

**DISCUSSION**

Our key finding is a demonstration of neural responses to abstract financial reinforcers. Crucially, we have shown that these responses are dissociable with respect to the psychological context, determined by subjects’ recent experience of changes in reward or penalty. Thus, these data suggest that certain neural responses are associated with reward level per se, whereas others are associated with an interaction between actual reward level and changes in level.

Reward level per se was related to activity in the midbrain and ventral striatum, a crucial component of dopaminergic projection systems. Although the midbrain activation could not be localized to either the substantia nigra or the VTA, key sites of origin of ascending dopamine systems, recent evidence, suggests that dopaminergic projections to the prefrontal cortex may be more widespread in origin than has previously been believed (Williams and Goldman-Rakic, 1998). Our results are consistent with, and extend

---

**Table 1. Maximally activated voxels in areas where significant evoked activity was related to reward level**

<table>
<thead>
<tr>
<th>Region</th>
<th>Left/Right</th>
<th>Brodmann's area</th>
<th>Talairach coords</th>
<th>Z value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive correlation with reward level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midbrain</td>
<td>R</td>
<td>9  - 12</td>
<td>-3</td>
<td>4.65*</td>
</tr>
<tr>
<td>Ventral striatum</td>
<td>R</td>
<td>18 9</td>
<td>-6</td>
<td>3.50*</td>
</tr>
<tr>
<td>Negative correlation with reward level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampus</td>
<td>L</td>
<td>-24 - 18</td>
<td>-18</td>
<td>5.38*</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>33 - 12 - 15</td>
<td>4.11</td>
<td></td>
</tr>
<tr>
<td>Parahippocampal gyrus</td>
<td>L</td>
<td>35 - 15 - 30</td>
<td>-12</td>
<td>4.73*</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>35 12 - 33 - 12</td>
<td>4.59*</td>
<td></td>
</tr>
</tbody>
</table>

\( p < 0.05 \) corrected, all other regions are \( p < 0.001 \) uncorrected.
findings from, single neuron and excitotoxic lesion studies in ani-
mals. Midbrain dopamine neurons have been shown to respond in
a relatively homogeneous way to primary reinforcers (Ljungberg et
al., 1992; Mirenowicz and Schultz, 1996), whereas numerous stud-
ies have identified the ventral striatum as critical in reward-related
processing (Apicella et al., 1991; Schultz et al., 1993) (for review,
see Koob, 1992; Robbins and Everitt, 1996). The nucleus accum-
bens has probably been most consistently related to reward, but
responses have been observed throughout the ventral third of the
striatum in animals (Apicella et al., 1991; Schultz et al., 1993;
Schultz, 1997). The region observed here is lateral to the nucleus
accumbens but falls clearly within the ventral striatum. The positive

<table>
<thead>
<tr>
<th>Region</th>
<th>Left/Right</th>
<th>Brodmann's area</th>
<th>Talairach coords</th>
<th>Z value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reward during run of success</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior medial thalamus</td>
<td>R</td>
<td>6</td>
<td>−12</td>
<td>0</td>
</tr>
<tr>
<td>Pallidum</td>
<td>R</td>
<td>15</td>
<td>−3</td>
<td>6</td>
</tr>
<tr>
<td>Subgenual cingulate</td>
<td>L</td>
<td>24</td>
<td>−9</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>9</td>
<td>36</td>
<td>3</td>
</tr>
<tr>
<td>Penalty during run of failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampus</td>
<td>L</td>
<td>−30</td>
<td>−12</td>
<td>−15</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>33</td>
<td>−21</td>
<td>−15</td>
</tr>
<tr>
<td>Both above situations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral orbitofrontal cortex</td>
<td>L</td>
<td>47</td>
<td>−33</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>51</td>
<td>15</td>
<td>−12</td>
</tr>
<tr>
<td>Insula cortex</td>
<td>L</td>
<td>−36</td>
<td>−6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>42</td>
<td>9</td>
<td>−6</td>
</tr>
<tr>
<td>Head of caudate</td>
<td>R</td>
<td>21</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>−9</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

*p < 0.05 corrected, all other regions are p < 0.001 uncorrected.

Figure 4. Responses to generic risk-taking processes. Activations of right (x = 51; y = 15; z = −12; p < 0.001) and left (x = −33; y = 15; z = −15; *p < 0.05) orbitofrontal cortex, right (x = 42; y = 9; z = −6; *p < 0.05) and left (x = −36; y = −6; z = 6; *p < 0.05) insula, and right (x = 21; y = 24; z = 0; *p < 0.05) and left (x = −9; y = 12; z = 12; p < 0.001) caudate associated with both reward during a winning streak and pen-
alties during a losing streak. These activations may represent responses to risk-taking, independent of outcome.
period.

account is that amygdala response to reward habituates very rapidly therefore the role of this region remains unclear. Another possible 1999) and functional imaging (Zalla et al., 2000) studies have 1997; Koepp et al., 1998). However, other lesion (Bechara et al., 1997; Stein et al., 1998), whereas previous studies of financial ings in amygdala or basal forebrain that have been associated with reward-related processes in animals (Cador et al., 1989; Everitt and a “behavioral inhibition system” (Gray, 1982, 1995).

Further context-dependent neural responses were observed in regions closely connected to reward-related striatal and limbic structures. These responses were seen both to high levels of reward and to high levels of penalty in the congruent contexts of winning and losing steaks, respectively. These responses were thus specific to congruent situations but blind to the valence of outcomes. They may reflect more generic processes associated with risk-taking, such as the excitement that is an essential component in maintaining risky behaviors. Subjects described the situations in which the height of the bar was at an extreme, either positive or negative, and also changing rapidly, as the most exciting; this was when they experienced the “buzz” of gambling. It is thus possible that the neural response seen under these conditions was associated with this subjective experience of excitement. Both the insula (Casey et al., 1995; Buechel et al., 1998) and orbitofrontal cortex (Bechara et al., 1996, 1997) have been previously implicated in representing changes in body state associated with emotional response. The orbitofrontal cortex is also a key region in mediating emotional influences on decision-making behavior in humans (Damasio, 1994; Elliott et al., 1997) and adapting responses to different behavioral contingencies (Rolls et al., 1994; Rolls, 1996), components of both winning and losing situations. It may also, therefore, subserve the processes of decision making and behavioral guidance that are critical in gambling situations, such as that used here.

The findings we report here provide evidence for dissociable functional systems within an extended human reward system. Our results suggest that activity in ascending dopaminergic systems projecting from midbrain to ventral striatum reflects the overall level or value of reward. However, key projection sites of this system (globus pallidus, thalamus, and subgenual cingulate) respond to reward level in a context sensitive way, showing activation only when reward is both high and increasing. By contrast with these dissociable responses to reward, responses to financial penalties are not anatomically differentiated, although they do show context-dependent enhancement. Finally, we demonstrated certain neural responses that may mediate generic processes maintaining risk-taking behavior regardless of outcome. Overall the findings, particularly those suggesting context dependency in reward systems, have implications for the development of theoretical models of reward-dependent behavior in humans.

REFERENCES


