On Framing Effects in Decision Making: Linking Lateral versus Medial Orbitofrontal Cortex Activation to Choice Outcome Processing

Sabine Windmann[1,2,3], Peter Kirsch[4], Daniela Mier[4], Rudolf Stark[4], Bertram Walter[4], Onur Güntürkün[2], and Dieter Vaitl[4]

Abstract

■ Two correlates of outcome processing in the orbitofrontal cortex (OFC) have been proposed in the literature: One hypothesis suggests that the lateral/medial division relates to representation of outcome valence (negative vs. positive), and the other suggests that the medial OFC maintains steady stimulus–outcome associations, whereas the lateral OFC represents changing (unsteady) outcomes to prepare for response shifts. These two hypotheses were contrasted by comparing the original with the inverted version of the Iowa Gambling Task in an event-related functional magnetic resonance imaging experiment. Results showed (1) that (caudo) lateral OFC was indeed sensitive to the steadiness of the outcomes and not merely to outcome valence and (2) that the original and the inverted tasks, although both designed to measure sensitivity for future outcomes, were not equivalent as they enacted different behaviors and brain activation patterns. Results are interpreted in terms of Kahneman and Tversky's prospect theory suggesting that cognitions and decisions are biased differentially when probabilistic future rewards are weighed against consistent punishments relative to the opposite scenario [Kahneman, D., & Tversky, A. Choices, values, and frames. *American Psychologist, 39,* 341–350, 1984]. Specialized processing of unsteady rewards (involving caudolateral OFC) may have developed during evolution in support of goal-related thinking, prospective planning, and problem solving. ■

INTRODUCTION

The processing of reward and punishment is known to crucially involve the orbitofrontal cortex (OFC). Studies with humans and animals show that neurons in the OFC encode the quality, quantity, probability, and timing of anticipated rewards (Kalenscher et al., 2005; Kirsch et al., 2003; Hikosaka & Watanabe, 2000; Schultz, Tremblay, & Hollerman, 2000; Tremblay & Schultz, 2000; Watanabe, 1999; Quintana & Fuster, 1992). In addition, OFC neurons respond to primary as well as secondary rewards and punishments of various sensory modalities (Gottfried, O'Doherty, & Dolan, 2003; O'Doherty, Critchley, Deichmann, & Dolan, 2003; Rolls, Kringelbach, & de Araujo, 2003; Rolls, O'Doherty, et al., 2003; Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Critchley, Mathias, & Dolan, 2001; Knutson, Fong, Adams, Varner, & Hommer, 2001; Elliott, Friston, & Dolan, 2000; Small, Zatorre, Dagher, Evans, & Jones-Gotman, 2000; Rogers, Owen, et al., 1999; Elliott, Frith, & Dolan, 1997). Theoretical accounts of these findings suggest that the OFC mediates the acquisition of stimulus–reinforcement associations as well as the flexible adjustment of these associations when contingencies change to bias decision making in the context of goal-directed behavior (Pickens, Saddoris, Gallagher, & Holland, 2005; Bechara, 2004; Schoenbaum, Setlow, Saddoris, Gallagher, 2003; Milad & Quirk, 2002; Baxter, Parker, Lindner, Izquierdo, & Murray, 2000; Damasio, 2000; Quirk, Russo, Barron, & Lebron, 2000; Schultz et al., 2000; Rolls, 1996, 2002).

Patients with lesions in the OFC tend to perseverate in decision-making tasks with alternating or probabilistic stimulus–outcome relations. They keep making disadvantageous or erroneous responses despite negative feedback even though they can sometimes recognize and verbalize the currently valid contingencies (Fellows & Farah, 2005; Bechara, Damasio, Tranel, & Damasio, 1998; Rolls, Hornal, Wade, & McGarth, 1994; Iversen & Mishkin, 1970). This suggests that their deficit does not stem from intellectual or memory-related dysfunctions, but rather from emotional or executive impairments (for discussions, see Bechara, 2004; Bechara, Tranel, & Damasio, 2000; Rolls, 1996, 2002). Although the patients are affectively responsive, and sometimes display disinhibited affect, they do not seem to link their emotional experiences to the relevant predictive situations or cues to ensure adaptive decision making.

The most widely distributed task used to examine such affect-related deficits in decision making is the Iowa Gambling Task (Bechara, Tranel, & Damasio, 2000; Bechara, Damasio, Trane, et al., 1998; Bechara, Damasio, Damasio, & Anderson, 1994). On this task, subjects can win points (or play money) by drawing cards from four decks that are associated with different amounts and probabilities. On the original version of the task, subjects are rewarded 100 points for each card of decks A and B, and 50 points for each card of decks C and D, respectively, which they quickly learn. However, what is learned only slowly as subjects continue playing the game is that some of the cards are, in addition to the steady gains, associated with unexpected losses. The amounts or probabilities of the losses are relatively high for decks A and B, resulting in a negative netto balance (expectancy value per card −25) when chosen consistently. By contrast, the amounts or probabilities of the losses associated with decks C and D are relatively low so that they are more than outweighed by the steady wins (expectancy value per card +25).

An inverted version of this task also exists, where small, steady punishments conflict with unexpected large rewards (Bechara, Tranel, & Damasio, 2000). Subjects lose 100 points with each card drawn from decks E and G, and 50 points on decks F and H. The unsteady wins outweigh those punishments for decks E and G (expectancy value per card +25), but not for decks F and H (expectancy value per card −25). Thus, the crux of both task versions is that subjects have to choose from the decks of cards associated with the relatively disadvantageous ''immediate'' (steady) outcomes to profit from the associated probabilistic (unsteady) outcomes that ensure positive netto outcomes in the long run.

Using both versions of the Iowa Gambling Task, Bechara and his colleagues have shown numerous times that patients with lesions in the medial part of the OFC (i.e., the ventromedial prefrontal cortex) are unable to overcome the impulse to choose the immediately advantageous cards to maximize long-term outcomes (e.g., Bechara, Tranel, & Damasio, 2000; Bechara, Damasio, Trane, et al., 1998; Bechara, Damasio, Damasio, et al., 1994). The cognitive component of this dysfunction has been termed myopia for the future (Bechara, Tranel, & Damasio, 2000). Bechara and colleagues hypothesized that the patients fail to generate and perceive the somatic states associated with risk perception (where risk means high probability of netto punishment) that normally biases decision making in healthy individuals (Bechara, 2004; Damasio, 2000; Bechara, Damasio, Trane, et al., 1998). Other researchers more or less explicitly deny the contribution of somatic markers and instead suggest that the patients are unable to extract and flexibly update environmental reward predictors from their experiences of reward and punishment (Fellows & Farah, 2005; Heims, Critchley, Dolan, Mathias, & Cipolotti, 2004;

Maia & McClelland, 2004; Schultz et al., 2000; Rolls 1996, 2002).

Alternative views notwithstanding, the common assumption is that the OFC needs to assemble information on outcome probabilities over time to be able to extract the long-term expectancy value of choice options, whether these are positive or negative. Although Bechara and colleagues have focused primarily on medial parts of the OFC to explain these abilities, there is reason to believe that lateral parts may play an even more crucial, and functionally distinct, role in this process. The precise nature of this subregional specialization within the OFC is not well understood at present, but is clearly suggested by neuroanatomical findings in primates (Öngür, Ferry, & Price, 2003; Carmichael & Price, 1995; Morecraft, Geula, & Mesulam, 1992).

The imaging literature suggests two hypotheses. The first states that lateral versus medial OFC subregions are differentially involved in the representation of positive versus negative outcome values. Medial parts of the OFC are thought to signal reward and pleasant stimulation, whereas lateral parts encode punishment and unpleasant stimulation. The empirical support for this presumption is manifold. O'Doherty, Kringelbach, Rolls, Hornak, and Andrews (2001) reported that in a reinforcement reversal task where subjects could symbolically win or lose money, the functional magnetic resonance imaging (fMRI) blood oxygen level dependent (BOLD) signal increased in the medial OFC in response to rewarding feedback and decreased in response to punishment, whereas the opposite was observed in lateral parts of the OFC. In fact, the BOLD signal in these two regions was linearly correlated with the amounts of reward and punishment received, respectively. Similar results with a primary reinforcer were reported by Small et al. (2000), who used positron emission tomography (PET) to investigate brain activity associated with eating chocolate before and after satiation. These researchers found that activity in caudomedial parts of the OFC decreased gradually with satiation, whereas caudolateral activity increased gradually with the growing aversion against eating more chocolate. Likewise, Rolls and colleagues found unpleasant smell (Rolls, Kringelbach, et al., 2003, see also Anderson et al., 2003) and unpleasant touch (Rolls, O'Doherty, et al., 2003) to be represented more laterally in the OFC compared with pleasant smell and touch that activated medial OFC subregions. Finally, Markowitsch, Vandekerckhovel, Lanfermann, and Russ (2003) found that happy as compared to sad episodes retrieved from autobiographic memory activated medial orbitofrontal regions (among other structures), whereas the inverse comparison activated lateral orbitofrontal regions. Results of numerous other studies with stimuli from various modalities conform with this picture of a valence-specific distinction of lateral versus medial OFC activation (Ursu & Carter, 2005; Kawabata & Zeki, 2004;

Gottfried et al., 2003; Zald & Pardo, 2002; Breiter et al., 2001; Goel & Dolan, 2001; Knutson et al., 2001; Blair, Morris, Frith, Perret, & Dolan, 1999).

On the other hand, researchers have shown that lateral parts of the OFC are more involved than medial parts in inhibiting inappropriate or overlearned behavioral responses in situations characterized by ambiguity, uncertainty, and conflict (McClure, Laibson, Loewenstein, & Cohen, 2004; O'Doherty, Kringelbach, et al., 2003; Critchley et al., 2001; Elliott, Dolan, & Frith, 2000; Dias, Robbins, & Roberts, 1996). In an overview of functional imaging findings, Elliott, Dolan, and Frith (2000) were among the first to explicitly propose this functional segregation along the lateral–medial division. According to this view, the lateral OFC, in particular caudal parts (Brodmann's are [BA] 47), is recruited primarily when responses previously associated with reward have to be suppressed during ongoing task performance, a function that specifically calls on executive and inhibitory control processes (Aron, Robbins, & Poldrack, 2004; Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003). Evidence in support of this suggestion is provided by studies that found caudolateral OFC activation in response to situations involving invalid, probabilistic, misleading, or otherwise ambiguous or conflicting cues. Among them are guessing and gambling tasks as well as memory and semantic retrieval tasks, with the latter showing more left-sided activation than nonverbal tasks (Goel & Vartanian, 2005; Aron, Robbins, et al., 2004; Cardillo, Aydelott, Matthews, & Devlin, 2004; Horn, Dolan, Elliott, Deakin, & Woodruff, 2003, Schnider, 2003; Cools, Clark, Owen, & Robbins, 2002; Berns, McClure, Pagnoni, & Montague, 2001; Shimamura, 2000; Gabrieli, Poldrack, & Desmond, 1998; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). In contrast, medial parts of the OFC are thought to merely monitor and hold online currently relevant reinforcement contingencies, especially when the outcomes are overlearned, steady, and familiar (O'Doherty, Critchley, et al., 2003; Elliott, Dolan, & Frith, 2000; Elliott & Dolan, 1999; Nathaniel-James, Fletcher, & Frith, 1997).

Taken together, it is at present unclear whether the valence or the steadiness of the outcomes determine medial versus lateral OFC activation. In any case, it seems surprising that reduced performance in the original version of the Iowa Gambling Task has been linked primarily with dysfunctions of *medial* OFC (ventromedial prefrontal cortex). This original task requires devaluation of decks that appear advantageous at first blush due to their relatively high steady (''immediate'') reward values. Successful performance on this task thus requires a shift from preacquired, initially reinforced responses to alternative options that are associated with less long-term punishment. The underlying processes, extraction of punishment values as well as representation of inconsistent outcomes likely to induce response shifts, have both been linked primarily with *lateral* por-

tions of the OFC (Fellows & Farah, 2005; O'Doherty, Critchley, et al., 2003).

By contrast, the inverted Iowa Gambling Task requires appreciation of probabilistic future rewards delivered in the context of steady punishments. On this task, subjects have to learn to tolerate relatively high levels of sure punishments to obtain even higher rewards delivered unsteadily. Adaptive decision making on this task does not involve extraction of future punishment values, nor does it require control of the impulse to choose immediate rewards. What would seem helpful instead is some form of reduced sensitivity to immediate punishment in combination with the ability to extract valid predictors of future rewards (e.g., Oya et al., 2005; McClure et al., 2004; Schultz, 2004; Zald et al., 2004; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Montague & Berns, 2002; Berns et al., 2001; Schultz et al., 2000).

Thus, although formally equivalent (i.e., equivalent with regard to the probabilities, absolute amounts, and expectancy values involved in the advantageous and disadvantageous choice options) and commonly considered psychologically equivalent (in terms of addressing the sensitivity for future outcomes), the two versions of the Iowa Gambling Task might actually recruit different cognitive mechanisms and brain structures during outcome processing. As only Bechara, Tranel, and Damasio (2000) in their study of patients with ventromedial prefrontal lesions have applied both tasks, a direct comparison of the cognitive and neural mechanisms that may be differentially involved is currently lacking. In effect, this questions why only the original task is commonly used to assess risk-seeking behavior, impulse control, and sensitivity for long-term consequences as an index of OFC function and related processes (e.g., Evans, Kemish, & Turnbull, 2004; Heims et al., 2004; Hooper, Luciana, Conklin, & Yarger, 2004; Overman, 2004; van Honk et al., 2004; Tomb, Hauser, Deldin, & Caramazza, 2002).

To address this issue, the present study measured lateral versus medial orbitofrontal activation with event-related fMRI during processing of rewards and punishments delivered in the two versions of the Iowa Gambling Tasks. We asked, specifically, whether the relative involvement of lateral versus medial subregions of the OFC in the processing of rewards and punishments would be the same for both tasks. If positive versus negative valence is truly the main factor determining the relative contribution of medial versus lateral OFC regions, as stated by the valence hypothesis outlined above, then both tasks should show this valence effect despite the different underlying stimulus–response contingencies. By contrast, if lateral versus medial OFC activation depends on the contingency of the outcomes (steady vs. unsteady), then the two tasks should yield different activation patterns in association with reward versus punishment processing.

To decide between these two possibilities we first performed analyses of the influence of valence for the two tasks separately, and then performed direct comparisons of the valence effects observed in the two tasks. As for the latter, we compared the reward > punishment contrast of the original task with the reward > punishment contrast of the inverted task. We reasoned that this analysis would bring out the effects of the different contingencies (steady vs. unsteady) because reward is steady in the original task and unsteady in the inverted task, whereas the opposite holds for punishment.

METHODS

Participants

Twenty-two healthy subjects (11 men) participated (mean age 26.3 years, $SD = 6.66$, all but one male subject right-handed). A mixed group of male and female subjects was chosen, as we intended to additionally test for sex differences in OFC activation (Overman, 2004; van Honk et al., 2004; Reavis & Overman, 2001). Subjects were undergraduate students, mostly freshmen of psychology (90%). They were paid €15 for participation plus the money they had gained in the two gambling tasks (if any). Subjects gave fully informed consent prior to participation. Procedures were approved by the local ethics committee at the University of Bochum, Germany.

Data Acquisition

A 1.5-T Siemens Symphony MRI scanner with a Quantum gradient system (Siemens, Erlangen, Germany) was used to acquire T2*-weighted images based on the BOLD contrast evoked by single-shot gradient-echo EPI sequences (TA = 100 msec, TE = 60 msec, TR = 3 sec, flip angle = 30°, FOV = 192 mm). Each volume contained 30 slices (5 mm thickness, no gap, 64 × 64 matrix, axial oblique orientation), similar to earlier fMRI studies on OFC activation (O'Doherty, Dayan, et al., 2003; Berns et al., 2001; Critchley et al., 2001). Slices were acquired interleaved in ascending order. The first three volumes were meant to allow for T1 calibration effects and were discarded before analysis. Prior to the functional measurement, a T1-weighed anatomical MRI scan was taken from each subject to exclude any structural brain abnormalities.

Materials and Procedures

Subjects were scanned while performing on computerized versions of the two Iowa Gambling Tasks developed by Bechara, Tranel, and Damasio (2000) and Bechara, Damasio, Damasio, et al. (1994). We computerized the manual version of these tasks based on the amounts and sequences given in Figure 1 of Bechara et al. (2000), with no increments, and punishment values of −1250 instead of −125 for deck B (as this is a typo in the original article). The procedures were controlled by the software package Presentation (Neurobehavioral Systems, Albany, CA).

Four virtual card decks (blue on black background) were presented on the screen. A trial began when the ''traffic light'' displayed above the decks turned green. Using a keypad with four buttons, subjects chose a card from one of the four decks by pressing the corresponding button. Subsequently, the traffic light as well as the chosen card turned red and no further keypress was allowed on this trial. After 3 sec of anticipation, two numbers were displayed on the card (indicating the outcome of this trial).

Original Task

A positive number with a plus sign appeared to inform subjects how many points they had had won (e.g., +100). Below this number was a negative number or a zero that informed subjects how many points they had lost (e.g., −300). Two decks contained cards on which the first (positive) value was always +100 (= steady rewards), but was sometimes accompanied by an even higher negative value (= unsteady punishments). In the long run, the expectancy values of cards from these decks were −25. The other two decks contained cards with a steady reward value of only +50, but these values were associated with relatively mild unsteady punishments. As a consequence, rewards more than outweighed the punishments in the long run for cards from these decks (expectancy value of +25), which made them ''good'' decks.

Inverted Task

The card values were the same as in the original task but with inverted signs. The upper number was negative and informed subjects how many points they had lost (e.g., −100). The lower number was positive and indicated how many points they had won, if any (e.g., +300). The punishments were small and steady (−50 or −100), although the rewards were relatively high and irregular (= unsteady). The rewards more than outweighed the punishments on the −100 decks (the ''good'' decks, expectancy value per card +25), but not on the other two decks (the ''bad'' decks, expectancy value per card −25).

Reward and punishment values were displayed for 3 sec, after which there was an intertrial interval varying randomly between 0.5 and 1.5 sec to allow for desynchronization of the events from repetition time (TR). Hence, the time window for obtaining the fMRI scans in the phase of interest varied between 3.5 and 4.5 sec, consistent with (or even extending) the intervals used in previous studies (e.g., Cools et al., 2002; Knutson et al.,

2001). A new trial began when the outcomes disappeared from the screen, the card turned blue again, and the traffic light above the decks turned green to indicate a new trial. One hundred trials were performed on each task (Bechara, Tranel, & Damasio, 2000). Half of the subjects started with the original task, and the other half with the inverted task (balanced gender proportions). There was a break of a few minutes between the two tasks during which subjects did not leave the scanner and were told to relax. The positions of the four card decks were randomized between subjects and tasks. The entire recording session took about 30 min.

Subjects were given 2000 points start credit, but were told that this credit would have to be returned before the final balance was determined and paid off. Their actual balance during the game was continuously indicated by two ''gold barrels'' displayed on the screen below the four decks of cards. These barrels became smaller or larger depending on the rewards and punishments obtained. When the wins (losses) accumulated to more than 1000 points, another barrel was added (eliminated). When subjects had lost their entire start credit, they were informed that they would receive another credit of 2000 points, and two new barrels were displayed on the screen. By the end of the game, they were given one cent in exchange for each point they had won.

In all other respects, procedures and instructions were identical with the manual version of the gambling tasks described by Bechara, Tranel, and Damasio (2000). After the experimental sessions, subjects filled out some personality and mood questionnaires for future comparison with a clinical sample in a planned follow-up study (which has not been carried out yet). These included German versions of the State-Trait Anxiety Inventory, Beck Depression Inventory, Eysenck's Impulsivity Scale (I7), the Barratt Impulsiveness Scale, and the Sensitivity to Reward and Punishment Scale (Torrubia, Ávila, Moltó, & Caseras, 2001). These measures did not reveal any particularly relevant information, and no consistent correlations with performance scores or brain activation patterns, presumably due to the low variance of the quite homogeneous sample of students, and are therefore not detailed.

Data Analysis

Performance on the two gambling tasks was determined by subtracting the number of cards drawn from the good decks from the number of cards drawn from the bad decks for blocks of 20 trials (Bechara, Tranel, & Damasio, 2000).

Functional imaging data were analyzed using SPM2 (Welcome Department of Cognitive Neurology, London, UK). EPI images were first corrected for sequential slice timing and then realigned to the first image to adjust for head movements. These realigned images were then spatially normalized to a standard EPI template. Finally, a 6-mm isotropic Gaussian kernel filter (full width at half maximum) was applied for smoothing.

The BOLD response function was modeled for trials with netto rewards and punishments for each task separately using the synthetic hemodynamic response function and its temporal and spatial derivatives. Linear contrasts were computed for each subject at the first level of a random effects analysis. Subsequent analyses of the group data at the second level focused on lateral and medial subregions of the OFC using small volume corrections (Worsley et al., 1996) as implemented in SPM2. Lateral OFC was defined as the orbital part of the middle and the inferior frontal gyrus. The medial OFC was defined as the left and right gyrus rectus and the medial orbital part of the superior frontal gyri. The required masks for these analyses were designed using the software program MARINA (Walter et al., 2003). The creation of masks is based on the anatomical parcellation of the brain as reported by Tzourio-Mazoyer et al. (2002). One-sample t tests were computed to determine the effects of reward alone, punishment alone, and the contrasts of reward versus punishment and punishment versus reward. Differences between the two tasks were determined by comparing the reward > punishment contrasts of the original task with that of the inverted task. Note that this analysis is equivalent to comparing the punishment > reward contrast for the inverted task with that of the original task. Gender differences in brain activation were not observed and are therefore not detailed. Activations outside the region of interest (ROI) are depicted in the figures but are not analyzed in detail and are not interpreted because they were not hypothesized beforehand.

RESULTS

Behavioral Results

For the original task, the ANOVA with repeated measures indicated that performance changed significantly across the five blocks of 20 trials, $F(4,84) = 6.16$, $p < .001$. Post hoc analyses indicated that performance increased significantly between the first and second blocks; $F(1,21) = 6.16$, $p < .001$, then tended to increase further between the second and the third block where it peaked, $F(1,21) = 3.68$, $p < .07$, so that the difference was maximal between first and third block, $F(1,21) = 19.25$, $p < .001$. However, performance stagnated thereafter in the fourth and fifth block (see Figure 1), both of which were not associated with any further performance increases. This nonlinear pattern across the five blocks was due to the fact that some subjects who had learned to perform well on the task finished off one or two of the good decks before the end of the game, and from then on had to continue by drawing cards from the remaining decks of which two
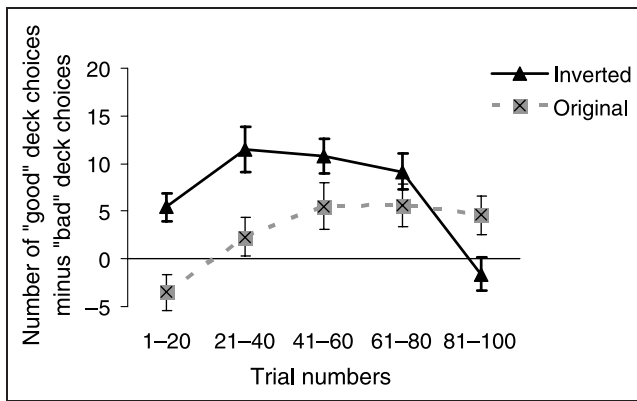
Figure 1. Decision-making performance on the original and the inverted Iowa Gambling Tasks. Performance declined towards the end of the session, as some subjects had finished off one of the advantageous decks (good decks) and from then on had to continue drawing from the remaining decks, of which two were disadvantageous (bad) in the long run.

were bad. This was expected based on pilot testing, and was preferred to providing endless decks, first because a sufficiently high number of trials was needed in all conditions for fMRI signal analysis, and second because unlike previous studies (e.g., Fukui, Murai, Fukuyama, Hayashi, & Hanakawa, 2005), we were not interested in anticipatory processes during gambling, but in mere outcome encoding. The steadiness and the values of the outcomes were the same, whether they occurred during three-deck or four-deck playing. We nevertheless repeated all fMRI analyses using data from only those trials for which all decks were still available (see the section Reanalysis below).

Performance on the inverted task also changed significantly across blocks, $F(4,84) = 13.037, p < .001$. As can be seen from Figure 1, performance peaked earlier than in the original task, namely, in the second block. This difference between the tasks was significant, as indicated by a significant task by block interaction, $F(4,84) = 8.73$, $p < .01$. In the inverted task, performance began at a relatively high level (compared to the original task) and then increased significantly to the second block, $F(1,21) = 22.35, p < .001$, from which point on it remained stable up to the fourth block, $F(2,42) = .36$, ns. Performance then decreased significantly between the fourth and fifth block, $F(1,21) = 17.28, p < .001$, again because some subjects had finished off one of the good decks.

For both tasks, performance peaked earlier than in the original study by Bechara, Tranel, and Damasio (2000) probably due to transfer effects between the two tasks. Unlike in the present study, control subjects in the study of Bechara, Tranel, and Damasio performed on only one of the Iowa Gambling Tasks.

Overall performance scores indicated that the inverted task was significantly easier than the original task,

$F(1,21) = 10.07, p < .01$. Although these differences in performance require interpretation, they do not provide any particular problem for the subsequent analysis of the imaging data, as the trials are sorted into conditions according to the netto values of the chosen cards. The primary effect of the differential performance is only that the reward condition contains more trials for the inverted task compared to the original task, whereas the opposite holds for the punishment condition. This might reduce signal-to noise ratios in simple reward and punishment contrasts, but should not affect the reward > punishment comparisons in any disadvantageous way.

Brain Imaging Data

Original Task

Reward and punishment alone produced no significant activation in lateral or medial OFC that survived small volume correction. However, the reward > punishment contrast yielded bilateral activation mainly in medial OFC (BA 10/11), and in the anterior part of the left lateral OFC (BA 10) as indicated by Figure 2 (top row) and Table 1. The reversed contrast (punishment > reward) yielded significant bilateral activation specifically in caudal parts of the lateral OFC (BA 47), see Figure 2 (second row) and Table 1.

Inverted Task

Reward and punishment alone again produced no significant activation in lateral or medial OFC that survived small volume correction. The punishment > reward contrast also produced no significant clusters. However, the reversed contrast (reward > punishment) yielded strong bilateral activation in both, caudal portions of medial and lateral OFC (BA 47), as indicated by Figure 2 (third row) and Table 1.

Comparison of the Original and Inverted Tasks

The direct comparison between the original and the inverted task yielded no significant effect for reward alone and punishment alone. The reward > punishment contrast activated left and right caudolateral OFC (BA 47) and, to a smaller degree, caudomedial subregions (BA 34) in the inverted task more than it did in the original task (Figure 2 bottom row). The reversed comparison (original > inverted) yielded no significant differences between the two tasks.

It should be noted that OFC activation is often subject to susceptibility artifacts in functional magnetic imaging leading to potential signal loss and distortion in these areas. Our results may therefore have underestimated the actual amount of orbitofrontal activation. However, there is no reason to believe that this problem applies differentially to the two versions of the tasks, especially

Figure 2. Activation patterns
in the two versions of the Iowa
Gambling Task. Top row:
reward > punishment contrast
of the original task. Second
row: punishment > reward
contrast of the original task.
The former contrast shows
more medial OFC activation,
whereas the latter shows more
lateral OFC activation. Third
row: OFC activation in the
reward > punishment contrast
of the inverted task. This
contrast activates both medial
and lateral subregions of the
OFC. Bottom row: reward >
punishment activation in the
inverted task contrasted with
the reward > punishment
activation in the original task.
This contrast activates the
lateral and (to a lesser degree)
caudomedial OFC in the
inverted task more than it
does in the original task.
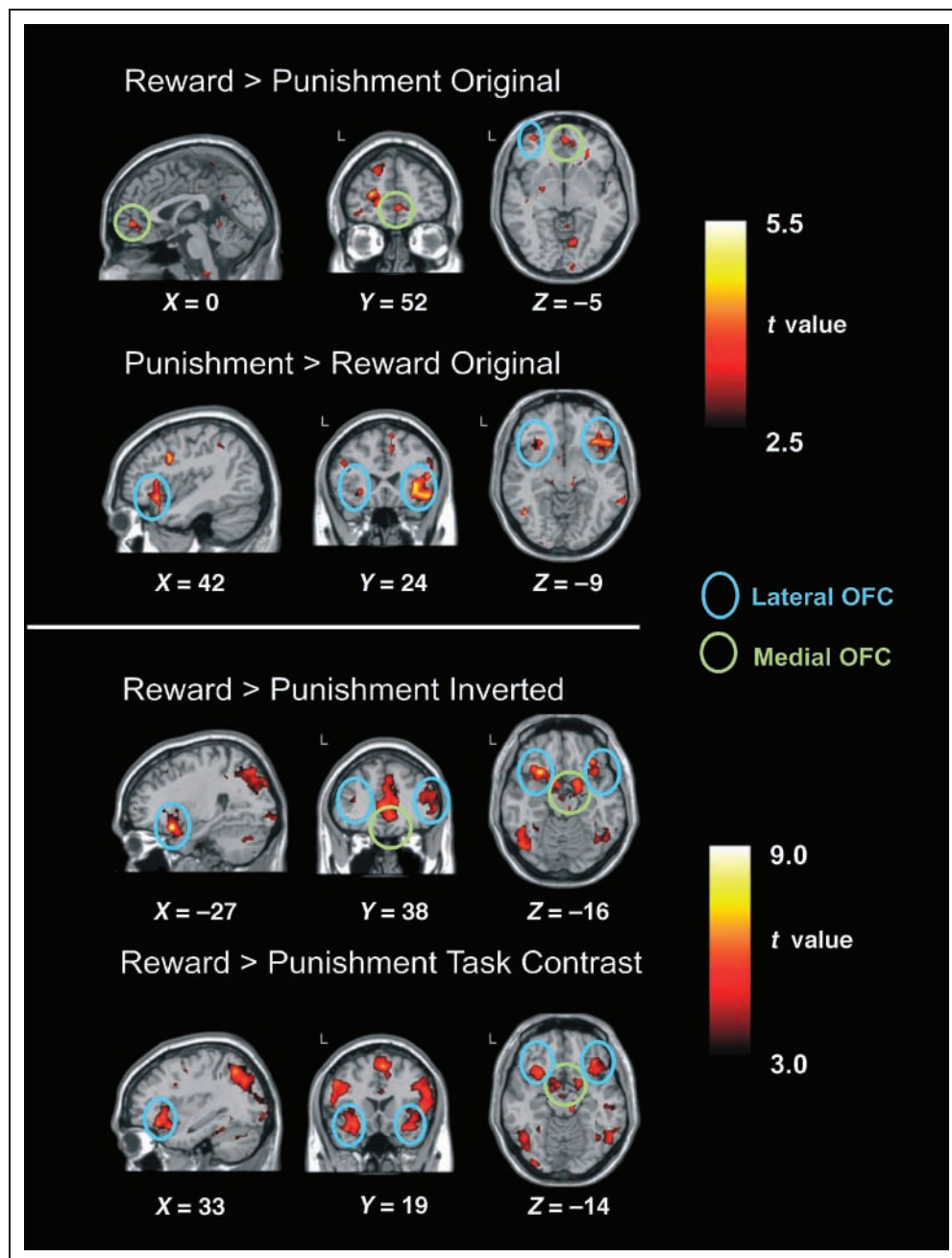
because we used a within-subjects design in a single
session with identical slice orientations. Therefore, com-
parisons between the two tasks should be valid and
revealing.

Reanalysis

To exclude the possibility that our results pattern was
significantly influenced by the fact that a reduced num-
ber of choice options was available on a fraction of trials
for subjects who had finished off one of the decks, we
reanalyzed the data using only trials during which all
four choice options were available (74.2%). Results were
somewhat weaker than in the analysis shown here, but

the main significance pattern was essentially unchanged
so that our conclusions would have been the same.

DISCUSSION

Using the original and inverted versions of the Iowa
Gambling Task, the present study investigated activation
of lateral and medial OFC subregions in response to
unsteady punishments and unsteady rewards received in
the context of steady outcomes of the opposite valence,
respectively. Numerous previous studies suggested that
the medial OFC encodes reward, whereas the lateral
OFC encodes punishment, an account referred to here

Table 1. Results of ROI Analyses Performed for the Two Iowa Gambling Tasks

Contrast	BA	X	Y	Z	T_{max}	p(T)	k	p(k)
Reward > punishment original								
Medial OFC left	10	0	52	−5	4.41	.07	66	.003
Medial OFC right	11	18	26	−14	4.66	.04	11	ns
Lateral OFC left	10	−36	52	−8	4.14	ns	45	.02
Lateral OFC right	–	–	–	–	–	–	–	–
Punishment > reward original								
Medial OFC left	–	–	–	–	–	–	–	–
Medial OFC right	–	–	–	–	–	–	–	–
Lateral OFC left	47	−30	26	6	3.83	ns	35	.03
Lateral OFC right	47	45	22	−11	4.78	.05	177	.001
Reward > punishment inverted								
Medial OFC left	32	0	38	−7	5.53	.005	71	.002
Medial OFC right	32	3	38	−7	6.38	.001	112	.001
Lateral OFC left	47	−27	17	−16	9.03	.001	404	.001
Lateral OFC right	47	33	31	−14	6.7	.001	384	.001
Punishment > reward inverted								
Medial OFC left	–	–	–	–	–	–	–	–
Medial OFC right	–	–	–	–	–	–	–	–
Lateral OFC left	–	–	–	–	–	–	–	–
Lateral OFC right	–	–	–	–	–	–	–	–
Reward–punishment inverted > reward–punishment original								
Medial OFC left	–	–	–	–	–	–	–	–
Medial OFC right	34	18	8	−13	4.69	.04	7	ns
Lateral OFC left	47	−39	19	−19	7.1	.001	323	.001
Lateral OFC right	47	33	28	−14	6.93	.001	422	.001
Punishment–reward inverted > punishment–reward original								
Medial OFC left	–	–	–	–	–	–	–	–
Medial OFC right	–	–	–	–	–	–	–	–
Lateral OFC left	–	–	–	–	–	–	–	–
Lateral OFC right	–	–	–	–	–	–	–	–

Displayed are the Brodmann's areas (BA), the Talairach coordinates (X, Y, Z), and the *t* value of the most significant voxel as well as the cluster size (k). p(T) indicates the error probability of the *t* value, p(k) the error probability of cluster size. *p* values are corrected for multiple testing (familywise error).

as the valence hypothesis (cf. Markowitsch et al., 2003; Rolls, Kringelbach, et al., 2003; Rolls, O'Doherty, et al., 2003; Small et al., 2000). Other studies suggested that the lateral OFC (in particular caudal portions, BA 47) represents uncertain, unexpected, and conflicting outcomes likely to prompt behavioral changes and suppression of previously rewarded responses (O'Doherty, Critchley, et al., 2003; Elliott, Dolan, & Frith, 2000). By

directly comparing the reward > punishment contrasts obtained in the original and the inverted tasks, we found that it is the steadiness of the outcomes that influences the pattern of OFC subregion activation, above and beyond any valence effects, in conjunction with significant behavioral performance differences between the two task variants. These findings may help to explain affect-related asymmetries in decision-making biases (Trepel, Fox, & Poldrack, 2005; Sanfey, Hastie, Colvin, & Grafman, 2003; De Brabander, Declerck, & Boone, 2002; Windmann & Krüger, 1998; Kahneman & Tversky, 1984).

Our analyses were performed in three steps. First, by investigating lateral and medial OFC activation in response to reward and punishment alone, we attempted to find confirmatory evidence for the valence hypothesis. However, we found that rewards and punishments alone did not produce any significant OFC activation in either of our analyses. Reliable OFC activation was found only when rewards and punishments were contrasted with each other. This observation could be taken as a first hint that outcome valence per se does not determine OFC activation; what seems crucial instead is either the valence *difference* between expected outcomes and actual outcomes (in accord with evidence from primate electrophysiology, cf. Oya et al., 2005; Tobler, Fiorillo, & Schultz, 2005; Holroyd, Larsen, & Cohen, 2004; Schultz et al., 2000), or the different *contingencies* of the contrasted outcomes, as discussed below.

Second, we inspected the pattern of OFC activation obtained from the reward > punishment contrasts in the two tasks separately. For the original task, we observed significant activation in anterior portions of the medial OFC, particularly in BA 10, extending to anterior lateral portions of this same region in the left hemisphere, whereas strong and bilateral activation in the caudolateral OFC was observed for the reversed contrast (punishment > reward), specifically in BA 47 (also called inferior frontal cortex). So far, these results appeared grossly consistent with the valence hypothesis that would have predicted medial but not lateral activation for this contrast. However, results were at this point still confounded with the different contingencies of the delivered rewards (steady) and punishments (unsteady).

The issue was clarified when the same contrasts were performed for the inverted task. The reward > punishment contrast again revealed significant medial OFC activation, albeit with a much more caudally located peak compared to the original task. In addition, the reward > punishment contrast produced strong bilateral activation in BA 47, despite the positive value of the contrasted outcomes (for which the valence hypothesis predicts medial activation). The reversed contrast (punishment > reward) produced no significant activation, neither in lateral OFC (as would have been predicted by the valence hypothesis), nor elsewhere in the brain, presumably because mild punishments are not experienced as significant events anymore when they are effectively expected with certainty. In summary, the inverted task produced activation patterns that were not only different from that of the original task but also inconsistent with the valence hypothesis.

The fact that these different patterns in OFC activation of the original and the inverted tasks were indeed significantly different was evidenced by our third analysis, the direct comparison of the reward > punishment contrasts between the two tasks. This analysis revealed much more caudolateral OFC (BA 47) activation in both hemispheres for the inverted task relative to the original one, with a striking effect size. Notably, the valence of the contrast (positive minus negative) was equally positive for both tasks; what differed were only the outcome contingencies (steady vs. unsteady). Obviously, then, these different contingencies were responsible for the strong task differences in caudolateral OFC activation, in addition to some minor differences in the right medial OFC (which missed significance at the cluster level). We note that the same significant effect would have emerged had we compared the size of the punishment > reward effect of the original task with that of the inverted task. The equivalence of these two contrasts makes it even clearer that the activation pattern is related to the contingency of the contrasted outcomes, not to their valence.

On the one hand, these results are incompatible with any simple interpretation of the valence hypotheses and seem more consistent with the conclusions of Elliott, Dolan, and Frith (2000) and O'Doherty, Critchley, et al. (2003), who stressed that (caudo)lateral OFC activation is found under conditions of inconsistent outcome delivery invoking suppression of previously reinforced responses and inducing behavioral shifting (see also Aron, Robbins, et al., 2004; Aron, Fletcher, et al., 2003). Earlier evidence seemingly supporting the valence hypothesis may have to be reinterpreted in light of these findings. Some negatively valenced stimuli (such as angry faces; Blair et al., 1999) might activate subregions of the lateral OFC not primarily because of their negative valence, but because they induce feelings of uncertainty and the need for response preparation and behavioral change. Likewise, devaluated primary reinforcers such as food and smell might activate these lateral regions not because they elicit aversive experiences, but because of their ambiguous and transient motivational qualities that call for regulation of consummatory behavior (Gottfried et al., 2003; Small et al., 2000). In general, it seems hard to tease apart negative emotional states from subjective experiences of uncertainty, ambiguity, and conflict likely to induce exploratory behavior, approach–withdrawal conflict, and behavioral shifting, and, conversely, positive emotional states from feelings of familiarity and safety that result from steady stimulus–outcome relationships (Garcia-Marques, Mackie, Claypool, & Garcia-Marques, 2004; Elliott, Newman, Longe, & Deakinet, 2003; Zajonc, 1980).

On the other hand, it must be reconsidered that along with the differential patterns in OFC activation, significant differences were found between the two tasks in behavioral performance: The inverted task was easier than the original task as evidenced by a larger proportion of choices from good decks. Naturally, this also resulted in higher credit balances that may have influenced subsequent decision making as well as OFC activation to some degree (cf. Elliott, Friston, & Frith, 2000), although probably not significantly so as our reanalysis suggested. Such secondary effects not withstanding, what really needs to be explained is why the performance differences between the two tasks emerged in the first place and how they relate to the activation differences observed in the caudolateral OFC.

We think that some of these effects might reflect different mechanisms involved in the processing of unexpected rewards relative to unexpected punishments. Such differences may have evolved during evolution due to differential behavioral implications. Behaviors directed at receiving temporally and spatially distant (and therefore genuinely uncertain) rewards such as foraging and problem solving require complex planning, strategic thinking, and flexible set-shifting and should therefore rely heavily on the ability to identify and integrate valid reward predictors from unsteady, ambiguous, and conflicting environmental cues (Montague & Berns, 2002). Protective acts, on the other hand, may be more short-lived and reactive and therefore less strategically and flexibly controlled, calling on autonomous physiological defense systems more than on prospective thinking and planning. Due to these differential cognitive–behavioral requirements, the brain may have developed special capabilities to extract the long-term value of unsteady rewards, more than for punishments (Trepel et al., 2005). Caudolateral OFC activation may be part of this functional specialization, as it has been observed specifically during reward prediction, hypothesis generation, impulse inhibition, and problem solving (Goel & Vartanian, 2005; Vartanian & Goel, 2005; Paulus, Feinstein, Tapert, & Liu, 2004). Differences between medial and lateral OFC neuromodulation by noradrenaline, serotonin (Rogers, Lancaster, Wakeley, & Bhagwagar, 2004; Rogers, Tunbridge, et al., 2003), and dopamine (Oya et al., 2005; McClure et al., 2004; Zald et al., 2004; O'Doherty, Dayan, et al., 2003; Berns et al., 2001; Schultz et al., 2000) may be of further (and perhaps crucial) importance. In effect, the brain may be better equipped to deal with the inconsistent pattern of outcomes presented in the inverted task (where rewards were unsteady and punishments steady) relative to the original task (where punishments were unsteady and rewards were steady), despite the formal equivalence of these two tasks in terms of the overall complexity of the four choice options and their associated long-term expectancy values. This predisposition may underlie the stronger and more valence-specific OFC activation as well as the superior behavioral performance in the inverted task relative to the original task.

In line with this interpretation, prospect theory has long posited that the prospect of future wins in the context of inevitable punishments induces different cognitive and decision-making biases than does the prospect of future losses in the context of steady rewards (Kahneman & Tversky, 1984), even when the expectancy values of the available choice options are the same. For example, when told in a thought experiment about an epidemic expected to kill 600 people, and asked to choose between a medical program [A] that can save the lives of 200 people and a medical program [B] that is associated with a one-third probability that 600 people will be saved plus a probability of two thirds that no one will be saved, the majority of people prefer program A (although the expectancy values of [A] and [B] are the same). Kahneman and Tversky (1984) call this behavior ''risk averse'' in a situation where a positive outcome (lives will be saved) is predicted in the context of a negative standard (inevitable loss of lives). By contrast, when given the choice between a medical program [C] by which 400 people will die, and program [D] that provides a one-third probability that no one will die plus a two-thirds probability that 600 people will die, the majority of subjects prefers program D (although the expectancy values of both options are the same as in [A] and [B]). Kahneman and Tversky call this tendency ''risk seeking'' in a situation where negative outcomes are predicted (lives will be lost) in the context of a positive reference state (all alive). Thus, despite objectively equal expectancy values of the four choice options, the framing of the future outcomes crucially determines the decision bias: The prospect of future rewards (options A and B) makes people want to minimize outcome variance and unsteadiness (i.e., minimize the prediction error; Oya et al., 2005; Schultz, 2000) more than does the prospect of future punishment (options C and D).

It seems that the latter situation resembles more the original version of the Iowa Gambling Task where subjects anticipate unsteady future punishments in the context of a (positive) reference state defined by steady rewards. Conversely, in the inverted task, subjects anticipate unsteady future rewards in the context of a negative reference state (= steady punishments). Obviously, the two paradigms are different because the Iowa Gambling Task relies on gradual learning (by experience) of long-term punishment and reward contingencies, whereas Kahneman and Tversky (1984) spell the parameters of the available choice options out explicitly to assess decision-making biases at once. However, the tasks are also comparable in that they evoke affective conflict through probabilistic payoffs with equivalent expectancy values but inverted framings (cf. Trepel et al., 2005; Sanfey et al., 2003). To the degree that a parallel can be drawn, Kahneman and Tversky's theory would

suggest that subjects tend towards the bad decks in the original task more than they do in the inverted task, as they are more easily attracted by the choice options with the higher outcome variance when trying to defend a positive standard against prospective losses (risk seeking). Presumably, this tendency makes them stay on the bad decks for longer and activate the medial OFC more than in the inverted task despite accumulating losses, consistent with Elliott, Dolan, and Frith's (2000) account of medial OFC involvement in maintaining a behavioral strategy. Conversely, the inverted task might initially bias subjects towards choice options with low outcome variance as they attempt to compensate for sure losses with gains (risk averse; Sanfey et al., 2003). In contrast to the original task, these low-variance options are associated with low fluctuating negative long-term balances in the inverted task. However, owing to higher activation of lateral OFC subregions supporting behavioral shifting (O'Doherty, Critchley, et al., 2003; Elliott, Dolan, & Frith, 2000) subjects are better able to overcome this natural bias and shift towards the high-variance choice options where they eventually experience positive netto outcomes (perhaps most saliently so on deck E where the third card already wins 1150 points, see Bechara, Tranel, & Damasio, 2000, p. 2193). This view would support the notion that the ability to shift from the initially preferred choice option to alternative options is the relevant variable determining lateral OFC activation (O'Doherty, Critchley, et al., 2003; Elliott, Dolan, & Frith, 2000) as well as performance on the Iowa Gambling Task, not the ability to look into the future (Fellows & Farah, 2005).

Finally, it is noteworthy that we did not find any asymmetric activation patterns between the two hemispheres that would map onto the view taken by Davidson and colleagues implicating the left frontal cortex in the experience of positive, approach-related emotions and the right frontal cortex in negative, withdrawal-related emotions (e.g., Davidson, 2003; Davidson & Irwin, 1999). Furthermore, our failure to find any such hemispheric differences seems inconsistent with the proposal of Aron and colleagues that only right inferior prefrontal cortex is involved in cognitive control and response inhibition (Aron, Robbins, et al., 2004; Aron, Fletcher, et al., 2003). We think that these two approaches may in fact be related in that negative emotions naturally produce response suppression more than approach behavior, although we would suggest that unilateral right-sided as opposed to left-sided processing might be observed primarily when no complex language is involved (cf. Cardillo et al., 2004; Gabrieli et al., 1998; Thompson-Schill et al., 1997). The fact that we did not find such asymmetries despite the clearly nonverbal nature of the Iowa Gambling Task may be related to the fact that the outcomes were either small and highly anticipated (in the case of steady outcomes), and therefore not intense enough, or contained affective conflict

in the form of a mixture of rewards and punishments (in the case of unsteady outcomes), so that there was emotional ambiguity instead of valence-pure affect induction on any given trial. In addition, the task required constant shifting between response alternatives whose long-term values had to be gradually acquired rather than suppression of an outright incorrect response. Besides, the BOLD signal measurement was taken during outcome processing, not during decision making, which might have blurred the hemispheric differences proposed by Aron, Robbins, et al. (2004) and Aron, Fletcher, et al. (2003).

In summary, the OFC activation patterns and the decision-making behavior observed in this study suggest that the cognitive and neural processes tackled by the original and the inverted versions of the Iowa Gambling Task are not the same. Due to the different underlying contingencies, activation in the lateral OFC subregion BA 47 differentiated between the inverted and the original tasks. At the same time, subjects achieved higher performance scores in the inverted task, presumably because prediction of future rewards as compared to punishments induces different decision biases. As Kahneman and Tversky (1984) would put it, it is easier to sell humans a lottery ticket than an insurance policy, even if the expectancy values of the associated outcomes are the same. More imaging and lesion studies comparing decisions made on choice options with inverted framings would be needed to further elucidate the cognitive and neural basis of such affect-related asymmetries in prospective thinking and decision making.

REFERENCES

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D. E., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6,* 196–202.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8,* 170–177.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience, 6,* 115–116.

Baxter, M. G., Parker, A., Lindner, C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital frontal cortex. *Journal of Neuroscience, 20,* 4311–4319.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition, 55,* 30–40.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50,* 7–15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1998). Deciding advantageously before knowing the advantageous strategy. *Science, 275,* 1293–1295.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123,* 2189–2202.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience, 21,* 2793–2798.

Blair, R. J., Morris, J. S., Frith, C. D., Perret, D. I., & Dolan, D. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122,* 883–893.

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

Cardillo, E. R., Aydelott, J., Matthews, P. M., & Devlin, J. T. (2004). Left inferior prefrontal cortex activity reflects inhibitory rather than facilitatory priming. *Journal of Cognitive Neuroscience, 16,* 1552–1561.

Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology, 363,* 615–641.

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22,* 4563–4567.

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

Damasio, A. R. (2000). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London, Series B, Biological Science, 31,* 1413–1420.

Davidson, R. J. (2003). Seven sins in the study of emotion: Correctives from affective neuroscience. *Brain and Cognition, 52,* 219–132.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Science, 3,* 11–21.

De Brabander, B., Declerck, C. H., & Boone, C. (2002). Tonic and phasic activation and arousal effects as a function of feedback in repetitive-choice reaction time tasks. *Behavioral Neuroscience, 116,* 397–402.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380,* 69–72.

Elliott, R., & Dolan, R. J. (1999). Differential neural responses during performance of matching and nonmatching to sample tasks at two delay intervals. *Journal of Neuroscience, 19,* 5066–5073.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence from human neuroimaging findings. *Cerebral Cortex, 10,* 308–317.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses associated in human reward systems. *Journal of Neuroscience, 20,* 6159–6165.

Elliott, R., Frith, C. D., & Dolan, R. J. (1997). Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia, 35,* 1395–1404.

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

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

Evans, C. E., Kemish, K., & Turnbull, O. H. (2004). Paradoxical effects of education on the Iowa Gambling Task. *Brain and Cognition, 54,* 240–244.

Fellows, L. K., & Farah, M. J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex, 15,* 58–63.

Gabrieli, J. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of the left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 906–913.

Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality and Social Psychology Bulletin, 30,* 585–593.

Goel, V., & Dolan, R. (2001). The functional anatomy of humor: Segregating cognitive and affective components. *Nature Neuroscience, 4,* 237–238.

Goel, V., & Vartanian, O. (2005). Dissociating the roles of right ventral lateral and dorsal lateral prefrontal cortex in generation and maintenance of hypotheses in set-shift problems. *Cerebral Cortex, 15,* 1170–1177.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301,* 1104–1107.

Heims, H. C., Critchley, H. D., Dolan, R., Mathias, C. J., & Cipolotti, L. (2004). Social and motivational functioning is not critically dependent on feedback of autonomic responses: Neuropsychological evidence from patients with pure autonomic failure. *Neuropsychologia, 42,* 1979–1988.

Hikosaka, K., & Watanabe, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cerebral Cortex, 10,* 263–271.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology, 41,* 245–253.

Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: Implications for the development of decision making and ventromedial prefrontal cortex. *Developmental Psychology, 40,* 1148–1158.

Horn, N. R., Dolan, M., Elliott, R., Deakin, J. F., & Woodruff, P. W. (2003). Response inhibition and impulsivity: An fMRI study. *Neuropsychologia, 41,* 1959–1966.

Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkey following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research, 11,* 376–386.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39,* 341–350.

Kalenscher, T., Windmann, S., Diekamp, B., Rose, J., Güntürkün, O., & Columbo, M. (2005). Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice paradigm. *Current Biology, 15,* 594–602.

Kawabata, H., & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology, 91,* 1699–1705.

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., Ott, U., Burkart, J., & Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: An event-related fMRI study. *Neuroimage, 20,* 1086–1095.

Knutson, B., Fong, G. W., Adams, C. M., Varnerm, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport, 12,* 3683–3687.

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa Gambling Task. *Proceedings of the National Academy of Sciences, U.S.A., 101,* 16075–16080.

Markowitsch, H. J., Vandekerckhovel, M. M., Lanfermann, H., & Russ, M. O. (2003). Engagement of lateral and medial prefrontal areas in the ecphory of sad and happy autobiographical memories. *Cortex, 39,* 643–665.

McClure, S., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306,* 503–507.

Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature, 420,* 70–74.

Montague, R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36,* 265–284.

Morecraft, R. J., Geula, C., & Mesulam, M. M. (1992). Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. *Journal of Comparative Neurology, 323,* 341–358.

Nathaniel-James, D. A., Fletcher, P., & Frith, C. D. (1997). The functional anatomy of verbal initiation and suppression using the Hayling test. *Neuropsychologia, 35,* 559–566.

O'Doherty, J. P., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience, 23,* 7931–7939.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 28,* 329–337.

O'Doherty, J. P., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4,* 95–102.

Öngür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *Journal of Comparative Neurology, 460,* 425–449.

Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain & Cognition, 55,* 134–147.

Oya, H., Adolphs, R., Kawasaki, H., Bechara, A., Damasio, A., & Howard, M. A., III (2005). Electrophysiological correlates of reward prediction error recorded in the human prefrontal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 8351–8356.

Paulus, M. P., Feinstein, J. S., Tapert, S. F., & Liu, T. T. (2004). Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection. *Neuroimage, 21,* 733–743.

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal lesions impair use of cue–outcome associations in a devaluation task. *Behavioral Neuroscience, 119,* 317–322.

Quintana, J., & Fuster, J. M. (1992). Mnemonic and predictive functions of cortical neurons in a memory task. *NeuroReport, 3,* 721–724.

Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of the ventromedial prefrontal cortex in the recovery of extinguished fear. *The Journal of Neuroscience, 20,* 6225–6231.

Reavis, R., & Overman, W. H. (2001). Adult sex differences on a decision-making task previously shown to depend on the orbital prefrontal cortex. *Behavioral Neuroscience, 115,* 196–206.

Rogers, R. D., Lancaster, M., Wakeley, J., & Bhagwagar, Z. (2004). Effects of beta-adrenoceptor blockade on components of human decision-making. *Psychopharmacology (Berlin), 172,* 157–164.

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B., & Robbins, T. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience, 20,* 9029–9038.

Rogers, R. D., Tunbridge, E. M., Bhagwagar, Z., Drevets, W. C., Sahakian, B. J., & Carter, C. S. (2003). Tryptophan depletion alters the decision-making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology, 28,* 153–162.

Rolls, E. T. (1996). The orbitofrontal cortex. *Philosophical Transactions of the Royal Society of London, Series B, 298,* 199–209.

Rolls, E. T. (2002). The functions of the orbitofrontal cortex. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 354–365). Oxford, UK: Oxford University Press.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery and Psychiatry, 57,* 1518–1524.

Rolls, E. T., Kringelbach, M. L., & de Araujo, I. E. (2003). Different representations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience, 18,* 695–703.

Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., & McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cerebral Cortex, 13,* 308–317.

Sanfey, A. G., Hastie, R., Colvin, M. K., & Grafman, J. (2003). Phineas gauged: Decision-making and the human prefrontal cortex. *Neuropsychologia, 41,* 1218–1229.

Schnider, A. (2003). Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Reviews Neuroscience, 4,* 662–671.

Schoenbaum, G., Setlow, B., Saddoris, M. P., & Gallagher, M. (2003). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron, 39,* 855–867.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience, 1,* 199–207.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and

behavioral ecology. *Current Opinion in Neurobiology, 14,* 139–147.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex, 10,* 272–283.

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology, 28,* 207–218.

Small, D., Zatorre, R., Dagher, A., Evans, A. C., & Jones-Gotman, M. (2000). Changes in brain activity related to eating chocolate: From pleasure to aversion. *Brain, 124,* 1720–1733.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.

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

Tomb, I., Hauser, M., Deldin, P., & Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience, 5,* 1103–1104.

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The sensitivity to punishment and sensitivity to reward questionnaire as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences, 31,* 837–862.

Tremblay, L., & Schultz, W. (2000). Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *Journal of Neurophysiology, 83,* 1877–1885.

Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research, 23,* 34–50.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.

Ursu, S., & Carter, C. S. (2005). Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: Implications for neuroimaging studies of decision-making. *Cognitive Brain Research, 23,* 51–60.

van Honk, J., Schutter, D. J., Hermans, E. J., Putman, P., Tuiten, A., & Koppeschaar, H. (2004). Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. *Psychoneuroendocrinology, 29,* 937–943.

Vartanian, O., & Goel, V. (2005). Task constraints modulate activation in right ventral lateral prefrontal cortex. *Neuroimage, 27,* 927–933.

Walter, B., Blecker, C., Kirsch, P., Sammer, G., Schienle, A., Stark, R., & Vaitl, D. (2003). MARINA: An easy to use tool for the creation of masks for region of interest analyses [Abstract]. Presented at the 9th International Conference on Functional Mapping of the Human Brain, June 19–22, 2003, New York, NY. Available on CD-ROM in *Neuroimage,* Vol. 19, No. 2.

Watanabe, M. (1999). Attraction is relative not absolute. *Nature, 398,* 661–662.

Windmann, S., & Krüger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and Cognition, 7,* 603–633.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping, 4,* 58–73.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Zald, D. H., Boileau, I., El-Dearedy, E., Gunn, R., McGlone, F., Dichter, G. S., & Dagher, A. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *Journal of Neuroscience, 24,* 4105–4112.

Zald, D. H., & Pardo, J. V. (2002). The neural correlates of aversive auditory stimulation. *Neuroimage, 16,* 746–753.