

## Brain regions involved in the learning and application of reward rules in a two-deck gambling task

E. Hartstra<sup>a,c,d,\*</sup>, J.F.E. Oldenburg<sup>a,c,d</sup>, L. Van Leijenhorst<sup>a,c</sup>, S.A.R.B. Rombouts<sup>a,b,c</sup>, E.A. Crone<sup>a,c</sup>

<sup>a</sup> Institute of Psychology, Leiden University, The Netherlands

<sup>b</sup> Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands

<sup>c</sup> Leiden Institute for Brain & Cognition (LIBC), Leiden, The Netherlands

<sup>d</sup> Department of Experimental Psychology, Ghent Institute for Functional and Metabolic Imaging, Ghent University, Ghent, Belgium

### ARTICLE INFO

#### Article history:

Received 26 February 2009

Received in revised form 16 January 2010

Accepted 18 January 2010

Available online 25 January 2010

#### Keywords:

Decision-making

Rule learning

fMRI

### ABSTRACT

Decision-making involves the ability to choose between competing actions that are associated with uncertain benefits and penalties. The Iowa Gambling Task (IGT), which mimics real-life decision-making, involves learning a reward–punishment rule over multiple trials. Patients with damage to ventromedial prefrontal cortex (VMPFC) show deficits learning these rules, although this performance deficit is not exclusively associated with VMPFC damage. In this study, we used functional Magnetic Resonance Imaging (fMRI) to study the roles of prefrontal cortex regions involved in rule learning and rule application in healthy adults using an adapted version of the Iowa Gambling Task. Participants ( $N=20$ ) were asked to infer rules over series of 16 trials in a two-deck card game. Rewards were given on each trial and punishment was unpredictable. For half of the series, those decks that gave high rewards were also better decks in the long run. For the other half of the series, the decks that gave low rewards were better decks in the long run. Behaviorally, participants started to differentiate between advantageous and disadvantageous decks after approximately four/six trials, and learning occurred faster for high-reward decks. Lateral PFC (lat-PFC) and Anterior Cingulate Cortex (ACC)/pre-supplementary motor area (pre-SMA) were most active for early decisions, whereas medial orbital frontal cortex (med-OFC) was most active for decisions made later in the series. These results suggest that lat-PFC and ACC/pre-SMA are important for directing behavior towards long-term goals, whereas med-OFC represents reward values towards which behavior should be directed.

© 2010 Elsevier Ltd. All rights reserved.

### 1. Introduction

Decisions in daily life are often made based on incomplete information or in situations in which there is a chance for unfavorable outcomes. Learning from the outcomes of prior choices and using this information to guide decisions is therefore important. In prior research, Bechara, Damasio, Damasio, and Anderson (1994) introduced a card gambling task, the Iowa Gambling Task (IGT), that mimics real-life situations in the way it factors uncertainty, reward and punishment (see Bechara et al., 1994, for paradigm details). In this task, money can be won or lost by selecting cards from decks; each card gives both a reward and an occasional punishment. Participants need to learn over the course of the task that selecting cards from decks which result in high immediate reward are disadvantageous in the long run because they also result in occasional high punishment. Moreover, they need to learn that

decks which result in low immediate reward are advantageous in the long run because these are associated with even lower occasional punishment. Healthy participants increasingly select cards from the advantageous decks and learn to overcome their initial preference for the disadvantageous decks. In contrast, neuropsychological studies have demonstrated that patients with damage to the ventromedial prefrontal cortex (VMPFC) fail to learn which of the decks are advantageous in the long run and keep selecting from the disadvantageous decks (Bechara et al., 1994; Bechara, Damasio, Tranel, & Damasio, 1997; Bechara, Tranel, Damasio, & Damasio, 1996). This suggests that VMPFC is important for decision-making under uncertainty.

Recent neuroimaging studies have confirmed the involvement of VMPFC in decision-making using risk-taking paradigms in healthy adults. These studies have shown that VMPFC, but also extended regions within the lateral orbitofrontal cortex (OFC) are sensitive to different aspects of risk-taking (Cohen, Heller, & Ranganath, 2005; Kringelbach & Rolls, 2004; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). Even though VMPFC and OFC are often used interchangeably, VMPFC refers only to the medial region of the OFC (BA 10/11) whereas the OFC also includes

\* Corresponding author at: Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium. Tel.: +32 9 264 94 18.

E-mail address: [Egbert.Hartstra@UGent.be](mailto:Egbert.Hartstra@UGent.be) (E. Hartstra).

lateral regions, such as BA 47. For this reason, we will refer to the specific OFC regions, which can be separated into medial and lateral OFC (med-OFC and lat-OFC respectively).

Prior studies indicated that different regions of the OFC are sensitive to reward and punishment information. Specifically, activity in lat-OFC correlates with the magnitude of a received punishment, and increased activation in med-OFC correlates with the magnitude of a received reward (O'Doherty, 2007; O'Doherty et al., 2001). Recent evidence indicates that these regions are also active in anticipation of loss and gain (Tom, Fox, Trepel, & Poldrack, 2007). In addition, when healthy participants were asked to make decisions that involve either high- or low risk, a region in med-OFC is more active for the high-risk decisions than for the low risk decisions (see Van Leijenhorst, Crone, & Bunge, 2006).

Animal research has shown that the OFC is important for learning and updating associations between stimuli and their associated reward or punishment (Kringelbach & Rolls, 2004; Rolls, 2000; Walton, Bannerman, Alterescu, & Rushworth, 2003). Furthermore, a study in which rats learn that specific odors predict future access to either a reward or punishment, showed that the neurons in the OFC show increased activity over time (Schoenbaum, Chiba, & Gallagher, 2000). Additionally, patient studies demonstrated that damage to med-OFC impairs learning in the IGT, but not explicit knowledge about contingencies (Bechara et al., 1997). These findings support the hypothesis that the OFC is involved in learning in decision-making under uncertainty. Frank and Claus (2006) explained the behavior patterns observed in patients with damage to the med-OFC using a modeling approach. They argued that the key function of the OFC is to hold representations of reinforcement magnitudes in working memory. Specifically, they argue that the OFC links stimuli and corresponding reinforcement values, thereby employing top down control over the basal ganglia and premotor areas (Frank & Claus, 2006).

In addition to the OFC, the dorsal anterior cingulate cortex (ACC)/pre-supplementary motor area (pre-SMA) and the lateral prefrontal cortex (lat-PFC), have been implicated in risk-taking and future oriented decision-making. The ACC/pre-SMA is thought to have a selective function in the process of decision-making, and controls and selects behavior based on stimulus-reward associations (Cohen et al., 2005). Furthermore, it is thought to be important for representations of risky decisions in which there is an increased potential for failure to obtain a reward (Cohen et al., 2005). A prior study which tested the dissociative roles of the med-OFC and ACC/pre-SMA in decision-making and risk-taking demonstrated that both regions are sensitive to uncertainty, but that the ACC is sensitive to which of two options results in less punishment, whereas med-OFC is sensitive to which of two options results in more reward (Blair et al., 2006).

To date, the role of lat-PFC in decision-making under uncertainty is not well understood. Lat-PFC receives input from various brain areas, including the OFC and the ACC. Therefore, it is well situated to play a role in rule learning and decision-making (Cohen et al., 2005; Rolls, 2000). Even though damage to lat-PFC does not hamper IGT performance, damage to both lat-PFC and the med-OFC results in worse performance on the IGT compared to damage to the med-OFC alone (Bechara, Damasio, & Damasio, 2000). IGT performance requires working memory, which is thought to be regulated by lat-PFC (Barch et al., 1997). Interestingly, a prior neuroimaging study showed that lat-PFC is more active when individuals choose for long-term gain, relative to short-term gain, suggesting that this region may keep track of future incentives (McClure, Laibson, Loewenstein, & Cohen, 2004). Thus, lat-PFC may be especially important in the initial learning phase of the IGT, because different outcomes have to be maintained in working memory and participants have to anticipate long-term goals. Indeed, a prior study which focused on rule learning in a cognitive selection task

demonstrated that lat-PFC is active at the onset of rule learning (Seger & Cincotta, 2006).

Few fMRI studies investigated how OFC, ACC/pre-SMA and lat-PFC are involved in affective rule learning in healthy individuals. One prior study by Fukui, Murai, Fukuyama, Hayashi, and Hanakawa (2005) examined brain activation in healthy adults while they performed the IGT. They showed that the medial prefrontal cortex exhibits increased activation during disadvantageous decisions as compared to advantageous decisions. Interestingly, performance on the IGT was significantly correlated with activation in this area, especially for disadvantageous decisions (Fukui et al., 2005). Even though the study by Fukui et al. (2005) provides an important starting point for understanding the involvement of prefrontal cortex regions in IGT performance, it does not address the question as to how these areas contribute to rule learning. Here, we examine this question using event-related fMRI, and we aim to disentangle different phases of learning.

In the current fMRI study we investigated the contribution of OFC, ACC/pre-SMA and lat-PFC when learning and applying reward rules. To this end, we used an adaptation of the IGT to examine different phases of rule learning and rule application. We distinguished between two types of reward rules; rules in which high reward resulted in long-term gain, and rules in which low reward resulted in long-term gain (Tomb, Hauser, Deldin, & Caramazza, 2002). Based on prior studies, we hypothesize that during decision-making, lat-PFC and the ACC/pre-SMA would be more active in the initial phases of learning, whereas OFC was expected to be more active during later phases, after reward associations were learned.

## 2. Methods

### 2.1. Participants

Twenty healthy, right-handed participants (nine men) took part in the study (mean age = 20.45 years, SD = 2.31). All participants reported to be healthy and had no history of neurological disorders. All participants gave written informed consent, and the study was approved by the Medical Ethical Review Board at the Leiden University Medical Center. Upon completion of the experiment, all participants received 30 euros as compensation for participation.

### 2.2. Behavioral assessment

In order to obtain an estimate of IQ, two subtests of the Wechsler Adult Intelligence Scales (WAIS-III), Similarities and Block Design, were administered following the scan session. Analyses revealed that all participants had average or above average estimated IQs (mean IQ = 113, SD = 8.20, ranging between 100 and 130).

### 2.3. Task instructions

Participants were instructed to select cards from one of the two decks during several series of 16 trials. At each new series, they had to infer which of the two decks was better in terms of long-term profit. Each selected card led to a reward or a combination of a reward and punishment. After the card was chosen, participants were asked to indicate how certain they were that their choice led to the best result on a three-point scale. Following the certainty decision, the outcome of the card choice was presented. At the beginning of each new series of 16 trials, participants were presented with a 3-s instruction screen that stated that two new decks of cards would be presented ("New series"). In total, participants completed 16 series during the experiment.

### 2.4. Task conditions

Card selection resulted in a fixed reward on each trial. The amount of reward differed between the two decks; one deck always resulted in a high reward (e.g. 100 on each deck) and the other deck always resulted in a lower reward (e.g. 50 on each deck). Additionally, a variable loss occurred randomly on 8 out of 16 trials. Thus, reward trials and reward + punishment trials both had a .5 probability. Reward/punishment contingencies were manipulated in such a way that at the end of each series of 16 trials the sum of reward-punishment for one deck resulted in a net reward, whereas for the other deck it resulted in a net loss.

Two different rule schemes were used. In the low reward rule scheme, the deck that gave lower rewards resulted in a net reward (Low rule condition, comparable to the original IGT). In the high-reward rule scheme the deck that gave higher rewards resulted in a net reward (High rule condition). Table 1 depicts an example of the

**Table 1**  
Example of two rule schemes used in the High rule condition or in the Low rule condition. In both cases deck A is the advantageous choice.

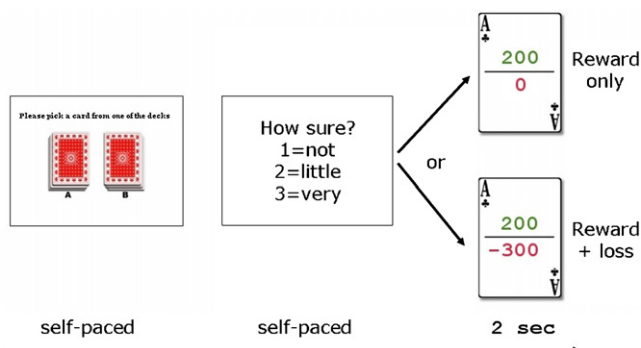
Trial number	High rule condition				Low rule condition			
	Deck A		Deck B		Deck A		Deck B	
	Lose	Win	Lose	Win	Lose	Win	Lose	Win
1	0	700	0	350	0	50	0	100
2	-420	700	-1400	350	0	50	0	100
3	0	700	0	350	-50	50	-200	100
4	-350	700	-2100	350	0	50	0	100
5	-420	700	-1750	350	0	50	0	100
6	0	700	0	350	-50	50	-300	100
7	-280	700	-1750	350	-50	50	-300	100
8	0	700	0	350	-40	50	-250	100
9	0	700	0	350	0	50	0	100
10	0	700	0	350	0	50	0	100
11	-350	700	-1750	350	-60	50	-250	100
12	0	700	0	350	-60	50	-200	100
13	0	700	0	350	-50	50	-250	100
14	-280	700	-1750	350	-40	50	-250	100
15	-350	700	-1400	350	0	50	0	100
16	-350	700	-2100	350	0	50	0	100
Total	-2800	11,200	-14,000	5600	-400	800	-2000	1600
Sum (win-lose)		8400		-8400		400		-400

two rule schemes used in this experiment. Participants were notified in advance that in each new series, one of the two decks would be more advantageous in terms of long-term profit.

The reason for using two different rule schemes was to ensure that participants had to learn in each series of 16 trials which deck was advantageous without using an explicit strategy (e.g. “always go for low reward decks”). Using series of 16 trials ensured that we had enough trial repetitions over the course of the experiment to compare brain activation during the first trials of a series with that during later trials (and vice versa). Two decks were used instead of the four originally used in the IGT to increase the likelihood of successful learning in relatively few trials. Pilot testing confirmed that participants were able to perform above chance level after approximately five card selections.

### 2.5. Experimental details

Once the series began, each trial started with the presentation of a fixation cross in the centre of the screen for 250 ms, followed by the presentation of two virtual card decks with the text: “please pick a card from one of these decks” at the top of the screen, and the letters A and B at the bottom of the decks (see Fig. 1). Participants were instructed to pick a card from one of the decks by pressing one of the two buttons that corresponded to the letters A and B using the index and middle finger of their right hand. Following the response, a blank screen was presented for 500 ms, followed by the presentation of the question: “how certain are you of your choice?” and three response options: (1) unsure, (2) a little sure and (3) very sure, that corresponded to the index, middle and ring finger respectively. Following this choice, a feedback screen was presented for 2000 ms showing an image of a turned



**Fig. 1.** On each trial, participants were shown two decks of cards that required a two-choice decision. Each deck could be associated with an advantageous rule or a disadvantageous rule, which was kept fixed during a 16-trial sequence. Each decision was followed by a 500 ms blank screen, followed by a screen asking participants about their choice certainty on a scale of 1–3. This response was followed by a 2000 ms feedback display, indicating the number of points won (in green) or lost (in red and preceded by a “-” sign) and the total number of points won during the task. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

card displaying the amount of points won printed in green at the top of the card and the amount of points lost printed in red at the bottom. The numbers were separated by a black line. The total points earned in the experiment so far were depicted at the bottom part of the screen. Blank screens jittered in increments of 0, 500, 1000 or 1500 ms separated the trials.

Following instructions and prior to the scan session, participants performed two practice series; one in the High rule condition, and one in the Low rule condition. At the beginning of the experiment participants were given 500 points to start with. The total number of points that was won was presented on each trial and at the end of each series. The experimental session was separated in four blocks to give participants breaks during the task. Each task block consisted of four series of 16 trials, and each block lasted between 6 and 8 min (partly self-paced). In total, the task lasted approximately 30 min.

### 2.6. Data acquisition

Images were acquired using a Philips Achieva 3T whole body MRI scanner using a standard head-coil at the Leiden University Medical Centre (LUMC). Functional data were acquired using a T2\*-weighted gradient-echo echo-planar pulse sequence (EPI, 38 2.75 mm axial slices, using sequential acquisition, 10% interslice gap, TR = 2.211 s, TE = 30 ms, field of view (FOV) = 220 mm, 80 × 80 voxels data matrix, 2.75 mm × 2.75 mm inplane resolution). The first two volumes of each scan were discarded to allow for T1-equilibration effects. High resolution T1-weighted anatomical images were collected at the end of the scan session sequence (TR = 9.76 s, TE = 4.6 ms, field of view (FOV) = 224 mm, 192 × 192 voxels data matrix). Head motion was restricted using a pillow and foam inserts that surrounded the head. The maximum amount of motion allowed was 3 mm and the maximum observed was ±1.8 mm. Using E-prime software (Psychology Software Tools, Inc., 2001; E-prime version 1.1 SP3), stimuli were projected onto a screen which was visible to the participants through a mirror attached to the head-coil. Responses were made with the right hand using a response box attached to the upper leg, responses and reaction times were recorded on each trial.

### 2.7. fMRI data analysis

Data pre-processing and analysis were conducted using SPM5 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>). First, all functional images were realigned using rigid body transformation, after which they were slice time corrected using the middle slice as a reference. The structural image of each subject was co-registered with his or her mean functional image after which all functional images were normalized to the Montreal Neurological Institute (Montreal, Quebec, Canada) T1 template. The images were resampled into 3 mm<sup>3</sup> voxels and spatially smoothed with a Gaussian kernel (8 mm, full-width at half maximum). A high pass filter of 128 Hz was applied during fMRI data analysis. Statistical analyses were performed on individual subjects' data using the GLM in SPM5.

We were interested in the question whether different brain areas were recruited during different phases of learning. We therefore used deck presentation as the main event of interest in the GLM. A canonical hemodynamic response function (HRF) was convolved at deck presentation onset for each trial, using a duration time of zero. Trials of each series were subdivided in three phases; trials 1–5 (first phase), trials 6–10 (second phase) and trials 11–15 (third phase). The last trial of each series was modeled separately to avoid time-overlap with the next series. As a result the GLM consisted of eight regressors. We computed contrast images comparing brain

activation patterns for the first phase (trial 1–5) versus last phase (trial 11–15) of the series, and for each phase versus fixation baseline. The resulting contrast images computed for each participant were submitted to second level group analyses. At the group level, whole brain contrasts between conditions were computed by performing one-tailed *t*-tests on these images, treating participants as the random effect. Whole brain statistical maps were thresholded at  $p < .05$  corrected for multiple comparisons (FWE) or at  $p < .001$  uncorrected for multiple comparisons, with an extent threshold of 10 contiguous voxels.

### 2.8. Region of interest analysis

Using the MARSBAR toolbox for use with SPM5 (Brett, Anton, Valabregue, & Poline, 2002) we extracted the mean parameter estimates for lat-PFC, ACC/pre-SMA and med-OFC for the three phases (first, second, third) of each rule condition (High, Low) relative to fixation baseline. The ROI approach served to further characterize the patterns of activation for these regions. ROIs for lat-PFC and ACC/pre-SMA were drawn from clusters of activated voxels found in the first > third whole brain contrast. For lat-PFC we used a cluster which had its peak activation at  $x = 39, y = 36, z = 39$  (cluster size 36 voxels), and for ACC/pre-SMA the peak activation was at  $x = -6, y = 12, z = 54$  (cluster size 57 voxels). The ROI for med-OFC was based on the third > first whole brain contrast, which had its peak activation at  $x = -6, y = 51, z = 9$  (cluster size 78 voxels).

## 3. Results

### 3.1. Behavioral data

#### 3.1.1. Choice pattern

To examine whether participants differentiated between advantageous and disadvantageous decks over the course of a 16-trial series, a RULE TYPE (High rule versus Low rule)  $\times$  TRIAL (1–16) repeated measures ANOVA for the number of advantageous choices was performed. Selection from advantageous decks increased over trials (main effect TRIAL,  $F(15, 285) = 31.99, p < .001$ ) and in general participants selected cards from advantageous decks more in the High rule relative to the Low rule trials (main effect RULE TYPE,  $F(1, 19) = 78.1, p < .001$ ). A RULE TYPE  $\times$  TRIAL interaction showed that learning occurred earlier for High rule relative to Low rule trials,  $F(15, 285) = 4.60, p < .001$  (see Fig. 2A). One sample *t*-tests revealed that for High rule series participants performed significantly above chance level from trial 3 onwards ( $M = .75, t(19) = 7.16, p < .001$ , and for the Low rule from trial 4 onwards ( $M = .68, t(19) = 4.97, p < .001$ ).

#### 3.1.2. Certainty choices

A RULE TYPE (High rule, Low rule)  $\times$  TRIAL (1–16) repeated measures ANOVA was performed for the mean certainty scores (see Fig. 2B). Certainty scores increased over trials (main effect TRIAL,  $F(15, 285) = 87, p < .001$ ), and were higher for High rules than for Low rules (main effect RULE TYPE,  $F(1, 19) = 50.97, p < .001$ ). Similar to the choice results, the certainty scores increased earlier for High rule relative to Low rule trials (TRIAL  $\times$  RULE TYPE,  $F(15, 285) = 7.33, p < .001$ ).

### 3.1.3. Correlation between certainty scores and performance

The number of advantageous choices and certainty scores showed a similar pattern over the 16 trials. In order to test if better performance (more advantageous choices) was correlated with higher certainty scores, we performed a correlation analysis between mean performance and mean certainty score separately for the High rule and Low rule conditions. This analysis confirmed a high correlation for the High rule condition ( $r = .57, p < .001$ ), but not for the Low rule condition ( $r = .34, p = .15$ ). Possibly, participants had more explicit knowledge about the rule scheme used in the High rule condition.

### 3.1.4. Correlation between performance and reward

A correlation analysis was performed to examine whether more frequent selection of advantageous decks resulted in a higher total amount of points at the end of the task. Indeed, as expected by design of the task, in both rule conditions advantageous deck selections correlated positively with the total amount of points earned,  $r = .93, p < .01$  in the High rule condition and  $r = .64, p < .01$  in the Low rule condition.

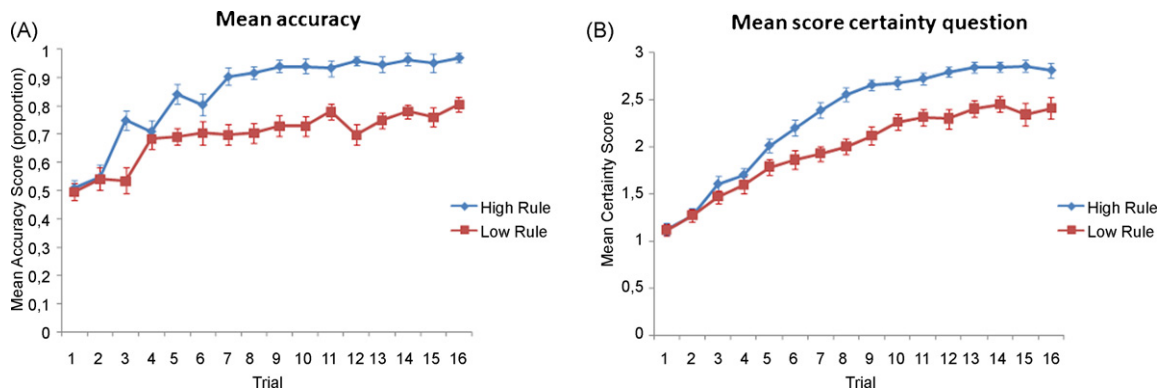
Together, the behavioral results demonstrate that the task manipulations resulted in the expected pattern of choices, and that the two-deck gambling task yielded results comparable to the original IGT (for more behavioral analyses see the supplementary material). These behavioral results set the stage for studying brain activation during the task.

### 3.2. Imaging data

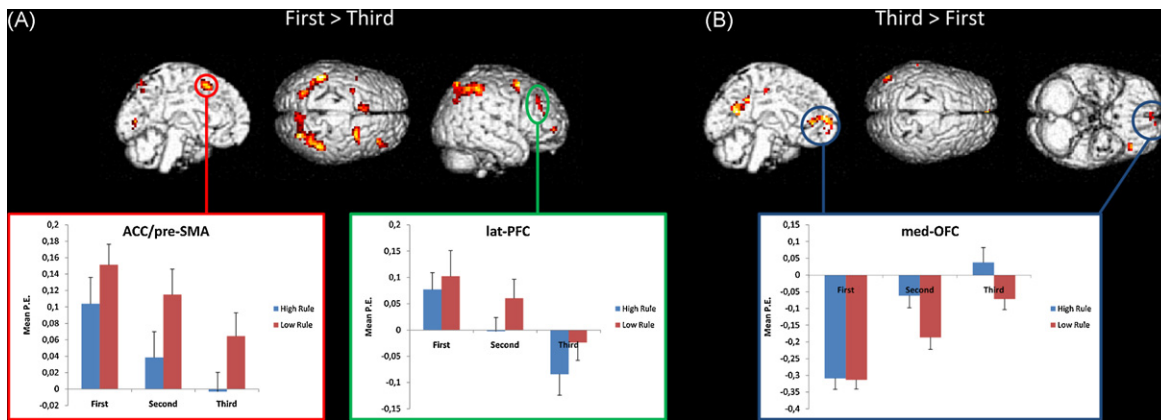
#### 3.2.1. Whole brain analysis

**3.2.1.1. First phase > third phase decisions.** The first set of analyses examined which regions showed increased activation during decisions in the first phase of learning relative to the third phase, by testing the first > third contrast (across rule types). There were no regions detected at  $p < .05$ , FWE, therefore the threshold was set at  $p < .001$  uncorrected. This analysis showed significant activation in multiple areas including right lat-PFC (39, 36, 39) and ACC/pre-SMA (–6, 12, 54) (see Fig. 3A). A whole brain two-sample *t*-test did not result in significant differences between the High and Low rule conditions for this contrast. MNI coordinates for the regions that were active in this contrast are reported in Table 2.

**3.2.1.2. Third phase > first phase decisions.** The second set of analyses examined which regions showed an increase in activation during decisions in the third phase of learning relative to the first phase, by testing the third > first contrast at  $p < .05$  FWE. As can be seen in Fig. 3B, activation in med-OFC (0, 57, –9) was increased at the end of the series relative to the beginning of the series. More-



**Fig. 2.** The mean percentage of advantageous choices (A) and mean certainty scores (B) for each deck and for each rule type were computed for the 16 trials separately and for each rule scheme separately.



**Fig. 3.** Panel A depicts whole brain results for regions which were active in the first > third contrast (thresholded at  $p < .001$ , uncorrected). The red boxed graph depicts mean parameter estimates (P.E.) extracted for the ACC/pre-SMA ( $-6, 12, 54$ ) for each IGT phase separately. The green boxed graph depicts mean P.E. extracted for the lat-PFC ( $39, 36, 39$ ) for each phase separately. Panel B depicts whole brain results for regions which were active in the third > first (thresholded at  $p < .05$ , FWE) contrast. The blue boxed graph depicts the mean P.E. extracted for the med-OFC ( $-6, 51, 9$ ) for each phase separately. In all ROI graphs blue bars represent the High rule condition whereas the red bars represent the Low rule condition. For ROI analysis details see method section.

over, the activation in med-OFC extended into the ventral ACC ( $3, 39, 6$ ). A two-sample  $t$ -test did not result in differences for the Low and the High rule conditions (see Table 2 for all regions that were active for this contrast).

Together, the results of the whole brain analyses show that lat-PFC, ACC/pre-SMA and med-OFC are differentially involved in decision-making for different phases of the task; with lat-PFC and ACC/pre-SMA showing decreasing activation and med-OFC showing increasing activation over trials.

The analyses reported above were collapsed across reward and punishment trials. One potential confound could be that these findings might have been the result of participants experiencing more punishment in the first phase and more reward in the third phase. To test this, we performed an additional whole brain analysis in which the feedback was kept constant. Trials were subdivided in reward and punishment trials. Reward trials were defined as trials in which the sum of the received points was equal or greater than zero. Punishment trials were defined as trials in which the sum of the received points was lower than zero. Both reward and punishment trials were subdivided in three phases; trials 1–5 (First phase), trials 6–10 (second phase) and trials 11–15 (third phase). The last trial was modeled separately to avoid time-overlap with the next series. As a result the GLM consisted of seven regressors. Otherwise, the analysis was identical to the previous whole brain analysis.

For reward feedback trials, the first > third contrast (modeled at the time of the decision) again resulted in increased activation in ACC/pre-SMA ( $-6, 12, 54$ ) and lat-PFC ( $42, 36, 36$ ) (see Table 2). The third > first contrast again resulted in activation in med-OFC ( $-3, 60, -9$ ). The same analyses for punishment feedback trials did not result in significant activation in the first > third contrast, but the third > first contrast again showed activation in med-OFC ( $-3, 57, 9$ ) (see Table 2). These results confirm that neural activation at the time of the decision is also observed when controlling for prior reward and punishment trials. One exception is the first > third contrast for punishment trials, but possibly this is due to fewer observations in this contrast. Whole brain two-sample  $t$ -test for both the first > third and third > first contrasts did not result in significant differences between the reward and punishment trials ( $p < .001$  uncorrected).

### 3.2.2. ROI analysis

Next, we explored the whole brain findings further by performing additional ROI comparisons for activation in ACC/pre-SMA,

lat-PFC and med-OFC during each phase of the task. A repeated measures ANOVA with factors RULE (high, low)  $\times$  PHASE (first, second, third) was computed for ACC/pre-SMA, lat-PFC and med-OFC separately. All reported post hoc tests used Bonferroni correction. The results are presented in Fig. 3.

The analysis for ACC/pre-SMA showed a decrease in activation over the three phases (main effect for PHASE,  $F(2, 38) = 24.98$ ,  $p < .001$ ) and all three phases differed significantly from each other, all  $p$ -values  $< .05$ . Activation in the Low rule condition was significantly higher than in the High rule condition,  $F(1, 19) = 8.87$ ,  $p < .01$  (see Fig. 3A red boxed graph), but the RULE  $\times$  PHASE interaction was not significant.

The same analysis for lat-PFC revealed that activation did not differ between rule conditions and decreased over phases (main effect for PHASE,  $F(2, 38) = 16.22$ ,  $p < .001$ , but not for RULE,  $p > .05$ ; see Fig. 3A green boxed graph). Post hoc tests showed that activation in all three phases differed significantly from each other, with all  $p$ -values  $< .05$ .

Finally, the analysis for med-OFC showed an increase in activation over phases,  $F(2, 38) = 61.81$ ,  $p < .001$ , and all three phases differed significantly from each other, all  $p$ -values  $< .05$ . Activation in the High rule condition was higher compared to the Low rule condition,  $F(1, 19) = 4.37$ ,  $p = .05$ . However, the increase in activation over phases occurred faster in the High rule condition than in the Low rule condition, as indicated by the significant interaction effect of RULE  $\times$  PHASE,  $F(2, 38) = 4.47$ ,  $p < .05$  (see Fig. 3B blue boxed graph).

## 4. Discussion

The goal of this study was to investigate the contribution of OFC, ACC/pre-SMA and lat-PFC when learning and applying reward rules in a gambling task. Specifically, we examined which brain regions were most active for decisions made during an early phase of the task and which brain regions were most active for decisions made during a later phase of the task. For this purpose, we developed a simplified version of the Iowa Gambling Task, in which reward rules had to be learned while event-related fMRI data were acquired.

As anticipated, participants were successful in learning which of two decks was advantageous in the long run as indicated by an increasing number of advantageous choices and higher certainty scores over the course of 16 trials. Reward rules were learned more slowly for Low rule series, where decks that initially resulted in larger reward were accompanied by higher punishment and there-

**Table 2**  
MNI coordinates of regions for all tested contrasts.

Voxels	z-Value	x	y	z	BA	Region
<i>First phase &gt; third phase contrast, uncorrected <math>p &lt; .001</math> with 10 contiguous voxels</i>						
Frontal						
57	4.97	-6	12	54	6	L frontal, medial frontal gyrus
30	4.32	-42	3	36	6	L frontal, precentral gyrus
77	4.27	30	6	57	6	R frontal, sub-gyral
22	4	-27	0	54	6	L frontal, middle frontal gyrus
16	3.84	30	54	0	10	R frontal, middle frontal gyrus
36	3.73	39	36	39	8	R frontal, middle frontal gyrus
	3.44	45	42	21	9	R frontal, superior frontal gyrus
Parietal						
371	4.27	48	-42	54	40	R parietal, inferior parietal lobule
	4.06	33	-69	36	19	R parietal, precuneus
	3.82	3	-75	54	7	R parietal, precuneus
266	4.18	-42	-42	45	40	L parietal, inferior parietal lobule
	4.13	-24	-60	42	7	L parietal, precuneus
	3.88	-45	-39	57	40	L parietal, inferior parietal lobule
Occipital/sub-cortical						
25	3.86	-18	-81	3	18	L occipital, lingual gyrus
	3.27	-24	-78	-9	-	L posterior, declive
12	3.81	30	24	0	-	R sub-lobar, claustrum
	3.63	30	24	12	-	R sub-lobar, claustrum
<i>Third phase &gt; first phase contrast, <math>p &lt; .05</math> FWE with 10 contiguous voxels</i>						
Frontal						
78	5.51	-6	51	9	9	L frontal, medial frontal gyrus
	5.26	9	54	-3	10	R frontal, medial frontal gyrus
	4.93	0	57	-9	10	L frontal, medial frontal gyrus
Parietal						
16	5.32	-57	-12	12	43	L parietal, postcentral gyrus
Temporal/occipital/sub-cortical						
20	5.56	-48	-72	30	39	L temporal, middle temporal gyrus
15	5.16	-60	-30	18	42	L temporal, superior temporal gyrus
80	5.44	-9	-63	21	31	L limbic, posterior cingulate
	5.23	-12	-45	30	31	L limbic, cingulate gyrus
25	5.25	3	39	6	24	R limbic, anterior cingulate
10	4.99	-3	-24	45	31	L limbic, cingulate gyrus
24	5.48	-36	3	-12	-	L sub-lobar, claustrum
11	5.13	12	-66	-6	-	R anterior, culmen
<i>First phase &gt; third phase (reward trials only) contrast, uncorrected <math>p &lt; .001</math> with 10 contiguous voxels</i>						
Frontal						
49	4.62	-6	12	54	6	L frontal, medial frontal gyrus
	4.28	-9	18	48	6	L frontal, superior frontal gyrus
84	4.3	24	6	54	6	R frontal, sub-gyral
	4.11	33	6	54	6	R frontal, sub-gyral
	3.28	21	15	66	6	R frontal, superior frontal gyrus
53	4.02	42	36	36	8	R frontal, middle frontal gyrus
	3.41	39	33	45	8	R frontal, superior frontal gyrus
28	3.94	-42	3	30	6	R frontal, precentral gyrus
23	3.84	-27	0	54	6	L frontal, middle frontal gyrus
10	3.76	30	54	0	10	R frontal, middle frontal gyrus
Parietal						
386	4.61	45	-39	45	40	R parietal, inferior parietal lobule
	4.46	48	-42	54	40	R parietal, inferior parietal lobule
	3.88	42	-60	51	7	R parietal, superior parietal lobule
274	4.15	-27	-60	42	19	L parietal, precuneus
	4.11	-45	-42	48	40	L parietal, inferior parietal lobule
	3.76	-24	-72	33	31	L parietal, precuneus
35	3.86	6	-72	57	7	R parietal, precuneus
Occipital						
263	4.43	-9	-78	3	18	L occipital, lingual gyrus
	4.39	-24	-69	-3	19	L occipital, fusiform gyrus
	4.2	-21	-81	0	18	L occipital, lingual gyrus
<i>First phase &gt; third phase (punishment trials only) contrast, uncorrected <math>p &lt; .001</math> with 10 contiguous voxels</i>						
No regions detected						
<i>Third phase &gt; first phase (reward trials only) contrast, <math>p &lt; .05</math> FWE with 10 contiguous voxels</i>						
Frontal						
40	5.59	-6	57	6	10	L frontal, medial frontal gyrus
	5.31	-3	60	-9	10	L frontal, medial frontal gyrus
	5.12	-9	63	-3	10	L frontal, medial frontal gyrus

Table 2 (Continued)

Voxels	z-Value	x	y	z	BA	Region
31	5.23	-57	-6	12	43	L frontal, precentral gyrus
	5.14	-45	-9	6	13	L sub-lobar, insula
	4.83	-57	-6	3	22	L temporal, superior temporal gyrus
25	5.19	-51	24	-3	45	L frontal, inferior frontal gyrus
	5.11	-39	33	-12	47	L frontal, inferior frontal gyrus
	4.88	-45	21	-12	47	L frontal, inferior frontal gyrus
Parietal						
36	5.39	-12	-39	45	31	L limbic, cingulate gyrus
	5.05	-9	-48	36	31	L parietal, precuneus
33	5.36	-9	-60	24	31	L parietal, precuneus
Occipital/temporal/limbic						
19	5.65	12	-69	-6	-	L anterior, culmen
24	5.26	-48	-72	30	39	L temporal, middle temporal gyrus
	5.12	-39	-78	33	19	L occipital, superior occipital gyrus
11	5.16	3	3	45	24	R limbic, cingulate gyrus
<i>Third phase &gt; first phase (punishment trials only) contrast, p &lt; .05 FWE with 10 contiguous voxels</i>						
Frontal						
15	5.41	-24	36	48	8	L frontal, superior frontal gyrus
	5.06	-21	42	42	8	L frontal, superior frontal gyrus
63	5.41	6	54	18	9	R frontal, medial frontal gyrus
	5.32	-3	57	9	9	L frontal, medial frontal gyrus
	5.14	3	42	3	32	R limbic, anterior cingulate
Parietal						
33	5.71	-6	-63	27	31	L parietal, precuneus
Temporal/limbic						
10	5.32	-9	-54	9	30	L limbic, posterior cingulate
12	5.19	-42	-75	30	39	L temporal, middle temporal gyrus

fore were disadvantageous in the long run, similar to the original IGT (Bechara et al., 1994). In contrast, reward rules were learned faster and certainty scores were higher for High rule series, in which decks that initially resulted in high reward were also advantageous in the long run (see also Tomb et al., 2002). Together, the behavioral results demonstrate that performance over a 16-trial period in the current two-deck gambling task is comparable to that reported for the original IGT.

The imaging analyses resulted in three main findings: (1) lat-PFC and ACC/pre-SMA were most active during the first phase of rule learning relative to the later phase, (2) med-OFC was most active during later phases of rule learning relative to earlier phases of rule learning, (3) these differences in activation were specific for decision-making and independent of the received feedback.

#### 4.1. Regions showing decreasing activation over rule learning

Decisions made in the first phase of the task were associated with increased activation in lat-PFC and ACC/pre-SMA. ROI analysis further showed that lat-PFC did not differentiate between the reward rules used in this study and that there was a steady decrease in activation over trials. These findings can be interpreted in terms of a general working memory involvement that is needed in both rule conditions. Indeed, lat-PFC is thought to be important for keeping track of the choice consequences (Barch et al., 1997; Bechara et al., 2000; McClure et al., 2004) and holding information active in working memory (D'Esposito, 2007; Smith & Jonides, 1999). Especially in the early phases of decision-making which require active learning, it is important to keep track of the consequences of prior choices because the contingencies are still unknown.

In contrast to lat-PFC, ACC/pre-SMA showed additional sensitivity to the type of reward rule used. Decisions made in the Low rule condition were characterized by more ACC/pre-SMA activation than decisions made in the High rule condition. Possibly, this difference reflects the higher magnitude of uncertainty during decisions made in Low rule trials. This interpretation is strengthened by the observation that certainty scores were also lower in the Low rule

condition. Furthermore, this interpretation fits with prior research showing that pre-SMA is involved in guiding behavior in uncertain situations (Volz, Schubotz, & von Cramon, 2003) and making decisions under uncertainty (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

#### 4.2. Regions showing increasing activation over rule learning

As anticipated, med-OFC showed an increase in activation over the course of the task as shown by the whole brain analysis. ROI analysis further revealed that activation in med-OFC was higher for the High rule condition and the increase in activation over phases occurred faster in this condition as compared to the Low rule condition. In prior research, it was shown that the med-OFC is especially sensitive to rewards (O'Doherty et al., 2001) and is involved in representing positive outcomes (O'Doherty, 2007; Rolls, Kringelbach, & de Araujo, 2003; Ursu & Carter, 2005; Windmann et al., 2006). In line with previous work we show that the med-OFC links stimuli with their corresponding reinforcement value over time (Frank & Claus, 2006). It is unlikely that the observed med-OFC activation pattern simply reflects a neural response to the reward feedback, because in that case the neural response would be expected to be greater for high-reward rule series, even in the early phases.

It should be noted that the OFC probably has a broader role than representing stimulus–outcome associations alone. Recent studies suggest that the OFC plays a role in predicting future outcomes, in processing prediction error signals, and even in computing decisions (for a detailed review see O'Doherty, 2007). Moreover, there is an ongoing debate about whether processes other than reinforcement representation can explain the reported dissociations between med-OFC and lat-OFC. For example, Windmann et al. (2006) investigated the role of medial and lateral OFC using two versions of the IGT, the original version and an inverted version. The original IGT used steady rewards and unsteady punishments, whereas the inverted IGT used unsteady rewards and steady punishments, which was accomplished by inverting the reward/loss frequencies. They reported that not valence but steadiness of out-

comes influenced OFC activation; med-OFC was more sensitive to steady outcomes whereas lat-OFC showed more sensitivity to unsteady outcomes. In the current study, lat-OFC did not show consistent activation associated with either early rule learning or later rule application. Possibly, lat-OFC is more sensitive to immediate, but not long-term outcomes, which would be consistent with observations by Windmann et al. (2006). However, our design was not optimized for testing the steady versus unsteady outcomes hypothesis. In both the first and the third phase, punishment is given on an unsteady basis. Therefore, unlike the accumulation of reward, punishment may not be a very effective signal for long-term profit. The role of the med-OFC versus lat-OFC should be investigated in more detail in future studies.

#### 4.3. Limitations and future directions

An unresolved issue in the current study concerns the question as to whether decisions are made using implicit or explicit knowledge about rules schemes. In the original IGT, it was shown that participants initially apply rules based on implicit knowledge but over the course of the task, this transfers into explicit knowledge (Brand, Recknor, Grabenhorst, & Bechara, 2007). Others have confirmed this finding by showing that intact declarative memory is needed for successful IGT performance (Gupta et al., 2009). In the current study, we aimed to examine this distinction in more detail by asking participants how certain they were of their choice. As expected, certainty scores increased over trials, and increased faster for High rule trials than for Low rule trials. Notably, performance in the High rule condition was correlated with certainty scores, but this was not the case in the Low rule condition. Possibly, gaining explicit knowledge was easier in the high-reward task condition, and therefore explicit knowledge may have played a larger role in the High relative to the Low rule condition. The current study did not allow for a differentiation between choice and certainty in the brain imaging analyses, but these issues warrant further investigation in future studies. The absence of a correlation between performance and certainty in the Low reward condition suggests that the performance increase is based on implicit rule learning.

In the presented analyses we have focused on the moment of the decision. However, the moment of feedback presentation followed the decision close in time. As a consequence it was not possible to disentangle the brain signals associated with the moment of the decision and the moment of the feedback. Additional analyses which kept the value of received feedback constant yielded results to those of the main analysis, but it is possible that neural responses are differentially sensitive to information provided by the decision and the feedback. For future research it would be interesting to completely disentangle both events.

The current findings may have important implications for other studies which report impairments on the IGT for specific populations, such as substance abusers (Gonzalez, Bechara, & Martin, 2007), gamblers (Goudriaan, Oosterlaan, de Beurs, & van den Brink, 2005), delinquents (Grant, Contoreggi, & London, 2000), and developmental populations (Crone & van der Molen, 2004; Hooper, Luciana, Conklin, & Yarger, 2004). One prediction that follows from our findings is that learning difficulties can result from impairments in distinct task phases, with different underlying neural contributions. The current design sets the stage for exploring the specificity of learning impairments in specific populations.

#### Acknowledgements

The author would like to thank Kiki Zanolie for all the help with scanning and Henk Van Steenberg for sharing his programming expertise.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuropsychologia.2010.01.012.

#### References

- Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia*, 35(10), 1373–1380.
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1–3), 7–15.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10(3), 295–307.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), 1293–1295.
- Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomously to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, 6(2), 215–225.
- Blair, K., Marsh, A. A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., et al. (2006). Choosing the lesser of two evils, the better of two goods: Specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *Journal of Neuroscience*, 26(44), 11379–11386.
- Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology*, 29(1), 86–99.
- Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an spm toolbox. In *Paper presented at the 8th International Conference on Functional Mapping of the Human Brain* Sendai, Japan.
- Cohen, M. X., Heller, A. S., & Ranganath, C. (2005). Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. *Cognitive Brain Research*, 23(1), 61–70.
- Crone, E. A., & van der Molen, M. W. (2004). Developmental changes in real life decision making: Performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, 25(3), 251–279.
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 761–772.
- Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review*, 113(2), 300–326.
- 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(1), 253–259.
- Gonzalez, R., Bechara, A., & Martin, E. M. (2007). Executive functions among individuals with methamphetamine or alcohol as drugs of choice: Preliminary observations. *Journal of Clinical and Experimental Neuropsychology*, 29(2), 155–159.
- Goudriaan, A. E., Oosterlaan, J., de Beurs, E., & van den Brink, W. (2005). Decision making in pathological gambling: A comparison between pathological gamblers, alcohol dependents, persons with Tourette syndrome, and normal controls. *Cognitive Brain Research*, 23(1), 137–151.
- Grant, S., Contoreggi, C., & London, E. D. (2000). Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia*, 38(8), 1180–1187.
- Gupta, R., Duff, M. C., Denburg, N. L., Cohen, N. J., Bechara, A., & Tranel, D. (2009). Declarative memory is critical for sustained advantageous complex decision-making. *Neuropsychologia*, 47(7), 1686–1693.
- 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(6), 1148–1158.
- Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72(5), 341–372.
- McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, 306(5695), 503–507.
- O'Doherty, J. (2007). Lights, camera, action! The role of human orbitofrontal cortex in an encoding stimuli, rewards, and choices. *Linking Affect to Action: Critical Contributions of the Orbitofrontal Cortex*, 1121, 254–272.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4(1), 95–102.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306(5695), 443–447.
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10(3), 284–294.
- Rolls, E. T., Kringelbach, M. L., & de Araujo, I. E. T. (2003). Different representations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience*, 18(3), 695–703.



- Schoenbaum, G., Chiba, A. A., & Gallagher, M. (2000). Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. *Journal of Neuroscience*, *20*(13), 5179–5189.
- Seger, C. A., & Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal systems during rule learning. *Cerebral Cortex*, *16*(11), 1546–1555.
- Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*(5408), 1657–1661.
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.
- Tomb, I., Hauser, M., Deldin, P., & Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience*, *5*(11), 1103–1104.
- 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*(1), 51–60.
- Van Leijenhorst, L., Crone, E. A., & Bunge, S. A. (2006). Neural correlates of developmental differences in risk estimation and feedback processing. *Neuropsychologia*, *44*(11), 2158–2170.
- Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: Uncertainty investigated by fMRI. *Neuroimage*, *19*(2), 271–280.
- Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. S. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *Journal of Neuroscience*, *23*(16), 6475–6479.
- Windmann, S., Kirsch, P., Mier, D., Stark, R., Walter, B., Gunturkun, O., et al. (2006). On framing effects in decision making: Linking lateral versus medial orbitofrontal cortex activation to choice outcome processing. *Journal of Cognitive Neuroscience*, *18*(7), 1198–1211.