Behavioral/Systems/Cognitive

Differential Engagement of the Ventromedial Prefrontal Cortex by Goal-Directed and Habitual Behavior toward Food Pictures in Humans

Sanne de Wit,[1,2] Philip R. Corlett,[2,3] Mike R. Aitken,[4] Anthony Dickinson,[4] and Paul C. Fletcher[2]

[1]Amsterdam Center for the Study of Adaptive Control in Brain and Behavior (Acacia), Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands, [2]Department of Psychiatry, University of Cambridge, Cambridge CB2 2QQ, United Kingdom, [3]Department of Psychiatry, Connecticut Mental Health Center, Yale University, New Haven, Connecticut 06510, and [4]Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom

According to dual-system accounts, instrumental learning is supported by both a goal-directed and a habitual system. Although behavioral control by the goal-directed system, through outcome–action associations, dominates with moderate training, stimulus–response associations are thought to form concurrently in the habit system. It is therefore challenging to isolate the neural substrate of the goal-directed system in neuroimaging research with healthy human volunteers. Recently, however, de Wit et al. (2007) developed an instrumental discrimination task that distinguishes between goal-directed and habit-based responding. In this task, cues are congruent, unrelated, or incongruent with subsequent outcomes. Whereas performance on congruent and control trials can be supported by both the goal-directed and habitual system, performance on the incongruent discrimination relies solely on the habit system. In the present study, we used this task with healthy participants undergoing functional magnetic resonance imaging to demonstrate that engagement of the goal-directed system during learning is reflected in increased activity in the ventromedial prefrontal cortex. Moreover, using a subsequent outcome devaluation manipulation, we show that this area is involved in guiding decision making when goal values change, even in the absence of external cues to guide performance. We can therefore exclude a purely Pavlovian account of ventromedial prefrontal function and unequivocally demonstrate its involvement in the acquisition as well as deployment of goal-directed knowledge.

Introduction

Instrumental learning can be supported by a goal-directed and a habitual system. Animal research provides evidence for this associative dual-system account (see Fig. 1) (Thorndike, 1931; Sutton and Barto, 1981; Dickinson and Balleine, 1994; Killcross and Coutureau, 2003) [for review, see Dickinson and Balleine (1993), Dickinson (1994), and de Wit and Dickinson (2009)]. In the goal-directed system, cues in our environment (stimuli; S) make us think of our goals (outcome; O), which in turn remind us of the responses that have yielded these in the past (response; R). In associative terms, stimuli activate actions via S→O→R associative chains (James, 1890; Pavlov, 1932; Asratyan, 1974; Hommel, 2003). Knowledge of the consequences of one's behavior allows the goal-directed agent to perform a given action only when those consequences are currently desirable, or in other words, when the behavioral consequences constitute a goal (Adams and Dickinson,

1981a). In contrast, when the habit system takes over, behavior becomes directly driven by contextual cues through S→R associations and thereby loses its immediate sensitivity to goal value (Thorndike, 1911; Adams and Dickinson, 1981b). The increased efficiency that is attained with habitual responding comes therefore at the price of decreased flexibility. According to the dual-system account, the habit system tends to take over with extensive practice, at least in part because S→R associations provide a more direct route to action selection than goal-directed S→O→R associative structures (Adams, 1982; Dickinson, 1985).

Because associations are formed concurrently in the goal-directed and habitual systems, it is difficult to dissociate their contributions in healthy volunteers without brain lesions. Recently, however, we developed a conflict task that allows us to isolate the contribution of the goal-directed system to instrumental discrimination learning (Dickinson and de Wit, 2003; de Wit et al., 2006, 2007; de Wit and Dickinson, 2009). In the current experiment, we contrasted brain responses to discriminations that could be supported by both the goal-directed and the habit system (cue–outcome congruent and control discriminations) with one in which the S–R habit system should predominate (cue–outcome incongruent discrimination; see Fig. 2 for a grayscale representation of the instrumental contingencies). Goal-directed responding is rendered disadvantageous in the incongruent discrimination because it creates conflict between the response engendered by an event acting as a discriminative stimulus and the
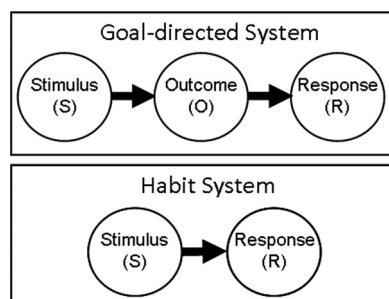
Figure 1. Dual-system account of instrumental learning. Instrumental actions are mediated by associative structures that build up concurrently in a goal-directed and habit system. In the goal-directed system, the outcome is represented in the associative structure, allowing incentive evaluation of the outcome to impact on action selection. We should note that there are several competing associative accounts of goal-directed action (see, e.g., Balleine and Ostlund, 2007; de Wit and Dickinson, 2009). The research presented in this paper does not aim to distinguish between these different accounts, and we will therefore, for the sake of simplicity, focus on the S→O→R account.
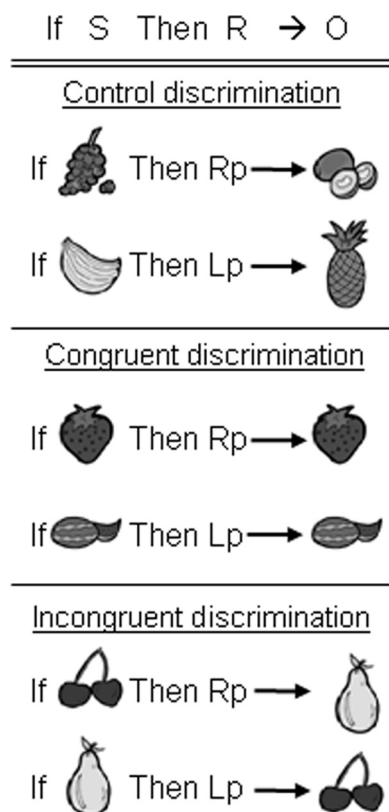

Figure 2. Illustrative grayscale examples of the contingencies in the control, congruent, and incongruent discriminations (see Materials and Methods for elaborate description of the contingencies).

response engendered when the same event has the status of an outcome [see Fig. 5 for an illustration of the associative structures; and see de Wit et al. (2007) for a more elaborate discussion of the associative theory]. Given the putative role of ventromedial prefrontal cortex (vmPFC), and adjacent medial OFC (mOFC), in the deployment of goal-directed knowledge (Valentin et al., 2007), we focused on activity in that area during the learning of these discriminations. Following learning, we used an outcome-devaluation task to assess purely outcome-based responding. A key advantage of our design was that cues were absent at this stage

to ensure that subjects used their knowledge of the action–outcome relationships to decide which action to choose.

Materials and Methods

Subjects
Sixteen healthy right-handed volunteers were recruited via advertisement in the local community. Two volunteers did not learn to perform the task above chance level and were therefore excluded, leaving eight females and six males (mean age, 24.7; SD = 3.5). The mean correlate of intelligence quotient (IQ) provided by the National Adult Reading Test (Nelson, 1982) was 35.4 (SD = 8.6; corresponding estimated verbal IQ = 116). All subjects gave written consent before the experiment and received an honorarium for their participation. The study was approved by the Peterborough and Fenland Local Research Ethics Committee. The subjects had normal structural MR brain scans, as confirmed by neuroradiological assessment. A telephone screening interview established that they did not have a history of psychiatric or physical illness (particularly cardiovascular or neurological disorders), head injury, or any history of substance abuse. Finally, subjects were without contraindications for functional magnetic resonance imaging (fMRI) scanning.

Stimuli
The stimuli consisted of two sets of colored icons, the first representing 11 different fruits: strawberry, orange, pineapple, pear, bananas, cherries, grapes, kiwifruit, melon, lemon, and coconut (see also de Wit et al., 2007), the second representing 11 different junk foods: popcorn, pizza, cake, hotdog, lollypops, ice cream, chips, donut, chocolate, hamburger, and sweets. Two different food sets were used so that we could boost power by running the tasks twice in each participant. Order of set presentation was counterbalanced across subjects. All pictures were presented on a standard PC monitor and responses on a left or right key were recorded on a buttonbox using a program written in Visual Basic 6.0.

Conflict task
The task was adapted for fMRI from the version used by de Wit et al. (2007). The main changes were that subjects received a demonstration of the task before going into the scanner, and that inside the scanner all subjects received training and testing with two sets of food pictures in succession. Finally, the trial structure was adapted for imaging purposes.

Demonstration of conflict task and instructions outside the scanner
All subjects received a demonstration of the conflict task outside the scanner, using the following instructions on the computer screen:
  "In this game, you will get the chance to earn points by collecting items from inside a box on the screen by opening the box by pressing either the right or the left key. If you press the correct key, the box will open to reveal a drink inside and points will be added to your total score. However, if you press the incorrect key, the box will be empty and no points will be added to your total. Your task is to learn which is the correct key to press. Sometimes it will be the left key and sometimes the right key. The picture on the front of the door should give you a clue about which is the correct response. To give you an impression of the game you will be asked to play later on, we will first give you some demonstration trials. Just follow the instructions on the screen."

  Having read these instructions, subjects were instructed to operate a left and right key on a button box with their index and middle finger of the right hand. On the computer screen, they were shown a picture of a closed box with a picture of a glass of beer on the front door. At the bottom of the screen we showed them the instructions "Press Left." Pressing the left key led to a picture of an open empty box. On the following screen subjects were again shown a picture of a glass of beer on the front door of a box, but this time with the instruction "Press Right." Pressing the right key was rewarded with a picture of a glass of champagne and one point. Subjects were then shown in the same manner that a glass of soda signaled that pressing the right key would not be rewarded, while pressing the left key was rewarded with a glass of wine and one point. Subjects were then given the following instructions:

"You have had a chance to learn which was the correct key to press for two different pictures. In the following demonstration, you will no longer be told which response to make, and your task is to press the correct key. From now on, each box will remain on the screen for a fixed time, and if you fail to make a response during that time the trial will end and you will gain no points. Only the first key press on each trial will count and the quicker a correct response is made the more points will be added to your total, so try to respond as quickly as possible!"

Subsequently, subjects received four practice trials with the beer stimulus and four trials with the soda drink stimulus, randomly intermixed. Pressing the correct key for the beer and the soda was rewarded with points, and with either champagne or a glass of wine inside the box, respectively. Pressing the incorrect key was always followed by an empty box. As in the subsequent scanner task, the faster a response was made, the more points were earned: 0–1 s, five points; 1–1.5 s, four points; 1.5–2 s, three points. If no response was made within the 2 s stimulus presentations, subjects were shown the message "Too slow! No points gained!" and the trial was aborted. The total score was always displayed at the top of the screen. The outcomes (another drink/empty box/"too slow" message) were always shown for 1 s. The intertrial interval varied randomly between 0.5 and 2.5 s. Following the training demonstration, subjects received instructions for the outcome-devaluation test:

"In the next phase, two open boxes will appear on the screen with different drinks inside them. One drink is earned by a left response in the first stage and the other by a right response. Although both drinks were valuable previously, one of them is now devalued and earns no points, whereas the other is still valuable and gains points. The devalued drink will have a cross on it. You should respond by pressing the key that earns a valued drink. The points you earn now will not be shown on the screen but you will see your final total at the end of the game. As in the training phase, only your first response will count."

The subjects were then shown two open boxes on the screen (one above the other), one containing a glass of wine and one containing a glass of champagne. On the first trial, the wine had a red cross superimposed on it, signifying that the left response associated with it no longer earned any points, while on the second trial the champagne was shown with a cross, signifying that the right response was no longer rewarded. During the 5 s test trials, subjects did not receive any feedback about their performance, but at the end of the test they were shown their total score, followed by some final instructions.

"The actual game will be very similar to this. However, it will be a lot harder, because you will be asked to learn the correct responses to many different food pictures (instead of drinks). Try to collect as many points as possible. You should pay attention to the types of foods that are found inside the boxes following each response, because later on you will be asked to gather some types of foods but not others. Remember to respond quickly, as quicker correct responses earn you more points. This is the end of the demonstration. If anything in these instructions is unclear, please ask the experimenter. If not, you're ready to go! Please tell the experimenter when you are ready to play the game. Good luck!"

Conflict task inside the scanner
Once settled in the scanner, participants were reminded of the key requirements of the task and that their objective was to try to gain as many food pictures (and points) as possible, by pressing the correct key on a button box. They were also instructed to look attentively at any fixation crosses that would appear intermixed with the experimental trials. We gave each subject discrimination training and testing with two sets of pictures (fruits and junk foods) with a short rest period in between.

As with the demonstration phase, participants were shown boxes bearing a food and were required to use this information to select the left or right key press. A correct response led to another food picture and points. The experimental design comprised four discriminations (see also de Wit et al., 2007): common outcomes, cue–outcome incongruent, cue–outcome congruent, and control (the latter three are illustrated in grayscale in Fig. 2). Each of the eight different discriminative stimuli was presented twice during each of six blocks, as well as two 3 s fixation crosses, amounting to a total of 96 training trials and 12 fixation crosses during each session. Trials were presented in a random order. Each subject was

run with a different assignment of the food events to the four discriminations (with a total of 12 permutations).

Cue–outcome incongruent. In this discrimination, stimulus pairs reversed their status as cues or outcomes across different trials. For example, cherries signaled that pressing the right key would be rewarded with a pear, whereas a pear signaled that pressing the left key would be rewarded with cherries. As illustrated in the top panel of Figure 5, a goal-directed approach to this discrimination should cause response conflict. When cherries acted as the discriminative stimulus, the correct right key press should become activated via a cherries→pear→right (S→O→R) associative chain, but because cherries also functioned as an outcome for the opposite left key press, the latter incorrect response should be activated directly via a cherries→left (O→R) associative link. As a result, it should be hard, if not impossible, to solve the cue–outcome incongruent discrimination in a goal-directed manner. Instead, subjects were forced to rely on solely cherries→right and pear→left (S→R) associations encoded by the habit system for discriminative support.

Cue–outcome congruent. In this discrimination, the same events acted as discriminative stimuli and outcomes for the same responses. For example, a strawberry signaled that right key presses would be rewarded with another strawberry, while a melon signaled that left key presses would be rewarded with another melon. This discrimination should be soluble by both the goal-directed and the habit system (see top panel Fig. 5).

Control. Here, two foods acted as discriminative stimuli signaling which response was correct on each trial, and two other foods acted as outcomes. For example, in one component of this discrimination, grapes signaled that pressing the right key would be rewarded with kiwifruit, while left key presses would not be rewarded. In the other component, bananas signaled that left key presses would be rewarded with a pineapple, whereas right key presses would not be rewarded. According to associative accounts, performance on this control discrimination could be supported by both the goal-directed and the habit system (see top panel Fig. 5). This discrimination can be solved in a goal-directed manner by forming two S→O→R associative chains: grapes→kiwifruit→right and bananas→pineapple→left. We would expect behavioral control through direct grapes→right and bananas→left associations to build up concurrently in the habit system. With only limited training performance should, however, be predominantly controlled by the goal-directed system.

Common outcomes. Here, left and right key presses were rewarded with the same event. For example, an apple signaled that right key presses would be rewarded with a lemon, whereas a coconut signaled that left key presses would be rewarded with a lemon. Performance on this discrimination cannot be supported by an S→O→R associative structure because the common outcome representation should activate both responses, and should therefore entirely rely on the habitual system (for review, see Urcuioli, 2005). However, because we did not find evidence for inferior performance on this task relative to the control discrimination in the present study, we feel that this final manipulation did not produce the desired "differential-outcomes effect," leaving us uncertain about the nature of the associative structures mediating this task. Given that we cannot use an outcome-devaluation task to clarify matters in this condition, we decided not to subject these data to further fMRI analysis.

Following training with each set of pictures, subjects received a reminder of the instructions for the outcome-devaluation test, which was then performed during scanning. On each test trial, subjects were shown two fruits or two junk foods (inside open boxes) that belonged to one particular discrimination (congruent/control/incongruent). One fruit/junk food was therefore previously earned by a right key press and the other by a left key press. In the test phase, one of the fruits/junk foods was now shown with a cross superimposed, symbolizing that, on this trial, this food was no longer worth any points. Subjects were required to press the key for the still-valuable food. Each outcome-devaluation test consisted of a total of 16 test trials, with each of the eight outcomes being devalued twice (once shown at the top of the screen, and once at the bottom) and two 3 s fixation crosses. At the end of each game, participants were shown their final score on the screen.

Questionnaires outside the scanner
Subjects were asked to indicate on a printed questionnaire for each fruit or junk food that had functioned as a discriminative stimulus, whether the right or left response had been correct, and which fruit/junk food was presented inside the box following a correct response for that discriminative stimulus.

fMRI data acquisition
We used a Siemens Trio scanner operating at 3 tesla. A total of 250 gradient echo T2*-weighted echo-planar images depicting blood oxygenation level-dependent contrast were acquired for each subject. The first six images were treated as "dummy" scans and discarded to avoid T1 equilibration effects. Images were positioned at 30° to the anterior commissure–posterior commissural line and comprised 49 slices, each of 2 mm with a 0.5 interslice gap. A repetition time of 3000 ms was used with an echo time of 30 ms and 90° flip angle. The scanner has a 192 mm field of view with a 64 × 64 data matrix.

Data were analyzed using statistical parametric mapping in the SPM5 program (www.fil.ion.ucl.ac.uk). Images were realigned then spatially normalized to a standard template and spatially smoothed with a Gaussian kernel (6 mm full width at half-maximum). The time series in each session were high-pass filtered (with cutoff frequency 1/120 Hz) and serial autocorrelations were estimated using an AR(1) model. Events were modeled using a canonical hemodynamic response function (plus first derivative) convolved with a 3 s boxcar function placed at the onset of each trial (that is we analyzed brain responses to the trial as a whole). In addition, a parametric function was applied to each condition to model effects of time. These functions were used as covariates in a general linear model and a parameter estimate was generated for each voxel for each event type. The parameter estimate, derived from the mean least-squares fit of the model to the date, reflects the strength of the covariance between the data and the canonical response function for a given condition. The responses to each condition were modeled separately compared with fixation and parameter estimates taken forward to a group analysis treating intersubject variability as a random effect.

To maximize sensitivity without unacceptable type I error, we confined our imaging analyses to regions of interest (ROIs) selected specifically for each contrast on the basis of prior studies. Below we summarize the contrasts used to the regions of interest. Full details of the precise regions of interest used for each key contrast are given in the section below. In brief, for the analysis of goal-directed learning (contrast 1 below), we focused on vmPFC. The analysis of the subsequent use of goal-directed knowledge during the outcome-devaluation test (contrast 2 below) was confined to a 10-mm-radius sphere around the focus of maximal activation within vmPFC identified by the goal-directed learning contrast. This sphere also formed the basis for extraction of subject-specific activity to establish whether the level of activation during goal-directed learning predicted goal-directed performance at subsequent test (contrast 3). The analysis of response conflict activity (contrast 4) was confined to dorsomedial, dorsolateral, and ventrolateral PFC. Each contrast was corrected for multiple comparisons using the false discovery rate (FDR) (Genovese et al., 2002). For completeness, we ran whole-brain comparisons (contrast 5) to establish whether there were any brain regions outside the ROIs that showed experimental effects (corrected for multiple comparisons across the whole brain). Finally, given the absence of any significant activations outside the ROIs, we ran an exploratory analysis of striatal regions comparing all conditions to baseline (see contrast 6 below).

(1) Investigating vmPFC activation in association with goal-directed learning. Our experimental design was based on the hypothesis that goal-directed behavior (and hence corresponding brain activity) would be attenuated in association with the cue–outcome incongruent condition. We therefore directly compared activation in both the control and the cue–outcome congruent conditions with the cue–outcome incongruent condition, predicting that regional responses reflecting goal-directed behavior would be greater in the two former conditions.

The analysis was confined to the areas of frontal cortex identified by Valentin et al. (2007): specifically, ventromedial PFC and orbitomedial PFC. We used the "Pickatlas" Tool (Maldjian et al., 2003) to select regions based on the

aal template and all contrasts were corrected for multiple comparisons on the basis of the subset of voxels examined (see supplemental Fig. 1, available at www.jneurosci.org as supplemental material).

(2) Investigating vmPFC activation in association with deployment of goal-directed knowledge when outcome values change. The purpose of this complementary analysis was to determine whether regions identified during goal-directed learning were also active during a task in which decisions were made on the basis of outcome value, without the aid of cues to direct performance. We therefore analyzed brain responses in outcome devaluation trials (see above), specifically comparing responses to outcomes from the control conditions (in which goal-directed behavior was predicted to occur) with responses to outcomes from the cue–outcome-incongruent condition (in which goal-directed behavior should be attenuated).

The analysis was confined to regions identified by comparison 1 above as being related to goal-directed learning. Note that the comparison used to identify the ROI (goal-directed learning) was independent of the comparison under test at this stage. We used Pickatlas to select a sphere (radius = 10 mm) around the maximum activation identified by the above comparison.

(3) Investigating the relationship between vmPFC activation during training and subsequent behavioral performance on the outcome-devaluation test. We created mean brain responses across the control trials during discrimination training of the two sessions and mean behavioral test performance (percentage correct) across the two sessions, to investigate whether vmPFC activation during training was predictive of behavioral performance during test.

(4) Identifying activity in brain regions associated with response conflict. Our experimental design is based on the idea that the cue–outcome incongruent condition engenders conflict between the response that would relate to a picture's status as a discriminative stimulus and the response related to its status as an outcome. It is this conflict that should lead to attenuation in goal-directed control (which the comparisons above sought to identify). The experimental manipulation therefore predicts that conflict should be maximized in this cue–outcome incongruent condition. Given that there are specific regional hypotheses about the brain activation associated with conflict: specifically, dorsomedial PFC regions (Kerns et al., 2004; Rushworth et al., 2004; Marsh et al., 2007) and lateral PFC regions (Milham et al., 2003; Kerns et al., 2004; van Veen et al., 2004), we examined the engagement of these regions by determining regional responses that were greater to the cue–outcome incongruent condition than to the cue–outcome congruent and control conditions. We used the "Pickatlas" Tool (Maldjian et al., 2003) to select the following regions: middle and inferior frontal gyri bilaterally with medially anterior and midcingulate cortex and supplementary motor area (see supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Within this set of regions, we identified those voxels in which responses were greatest for the cue–outcome incongruent condition (with correction for multiple comparisons on the basis of the subset of voxels examined).

(5) Whole-brain analysis for the key comparisons (supplemental material, available at www.jneurosci.org). A whole-brain analysis, corrected for multiple comparisons across the whole brain, was performed for each of the key comparisons.

(6) Investigating striatal activation during all discriminations relative to baseline. Although the neural supporting system for S–R habit formation is not a focus of this study, we should expect, according to our theoretical analysis, that this system shows greater activation across the three discriminations than during baseline. The striatum (especially the dorsal region) has frequently been proposed as a critical substrate for habit formation, and we therefore conducted an analysis with a "Pickatlas" Tool mask of the striatum (including ventral and dorsal regions; see supplemental Fig. 1, available at www.jneurosci.org as supplemental material) to contrast [congruent + control + incongruent] > baseline.

Results
Behavioral results
Statistical analysis was performed using SPSS 15.0. All *p* values involving repeated-measures factors are based on Greenhouse–Geisser sphericity corrections, and all significant (*p* < 0.05)

higher-order interactions involving the factor of interest (discrimination type) are reported.

Discrimination training

To assess behavioral performance, we calculated accuracy percentages (correct responses divided by total number of responses × 100), with 50% representing performance at chance level. "Missed trials" were omitted from the analysis (two common outcomes, two cue–outcome congruent, six control, and three cue–outcome incongruent; across the two sessions and all 14 subjects).

As can be seen in Figure 3, subjects rapidly learned to perform all discriminations. In line with our predictions, performance on the incongruent discrimination was inferior to that on the congruent and control discriminations. Unexpectedly, however, performance on the common-outcomes discrimination was at a similar level as that on the control discrimination. We conducted an ANOVA with three within-subject factors: session (first/second), block (1–6), and discrimination (common outcomes/congruent/control/incongruent). In line with our observations, the analysis revealed a significant effect of discrimination [$F_{(3,39)} = 5.03$, mean squared error (MSE) = 794.3, $p <$ 0.01]. *Post hoc* Tukey–Kramer analysis revealed that cue–outcome incongruent performance was worse overall than cue–outcome congruent, control, and common outcomes, whereas performance on the latter three discriminations was statistically indistinguishable. Also the effect of block was significant ($F_{(5,65)} =$ 38.06, MSE = 381.2, $p < 0.0005$), but there was no significant block × discrimination interaction ($F = 1.23$).

As can be seen in the right panel of Figure 3, responding became faster as a consequence of training. Overall, however, participants tended to react relatively slowly on cue–outcome incongruent trials. The lower accuracy on incongruent trials, as reported above, was therefore not due to a speed–accuracy trade-off. Statistical analysis yielded a significant effect of block ($F_{(5,65)} = 47.95$, MSE = 0.034, $p < 0.0005$), but more importantly, there was a significant effect of discrimination type ($F_{(3,39)} = 9.71$, MSE = 0.026, $p <$ 0.0005), as well as a significant discrimination × session interaction effect ($F_{(3,39)} = 3.69$, MSE = 0.024, $p < 0.05$), which prompted separate statistical analyses of the two sessions. Both analyses yielded significant effects of discrimination (session 1: $F_{(3,39)} = 7.97$, MSE = 0.023, $p < 0.005$; session 2: $F_{(3,39)} = 5.82$, MSE = 0.027, $p < 0.01$), which were further investigated with *post hoc* Tukey–Kramer analyses. During the first session subjects were significantly slower to respond on the cue–outcome incongruent trials than on the cue–outcome congruent and common-outcomes trials. In contrast, during the second session subjects responded slower on the cue–outcome incongruent as well as common-outcomes trials than on the cue–outcome congruent trials. Most importantly, we can exclude a speed–accuracy trade-off account of the relatively low accuracy of performance on incongruent trials.

Outcome-devaluation tests

As can be seen in Figure 4, performance was better on the cue–outcome congruent and control trials than on the cue–outcome incongruent trials during the outcome-devaluation test. This was confirmed by a statistical analysis with the within-subject factors
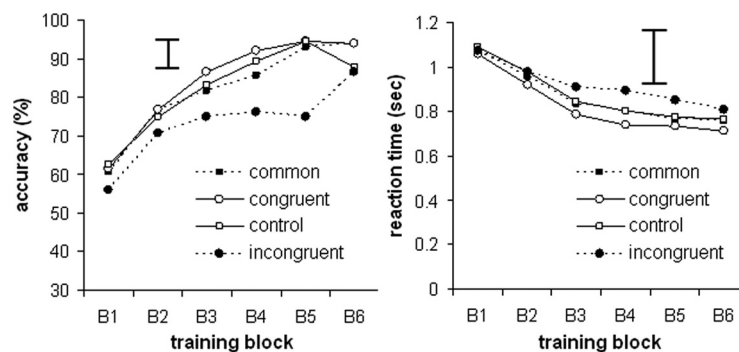

Figure 3. Instrumental discrimination training—behavioral performance. Left, Average percentages of correct responses on the six blocks of common-outcomes, congruent, control, and incongruent discrimination training collapsed across the two sessions (inside the scanner). Right, Response latencies. Error bars represent estimates of 2 standard errors of the difference (values were 4.1 for accuracy and 0.13 for response latencies).
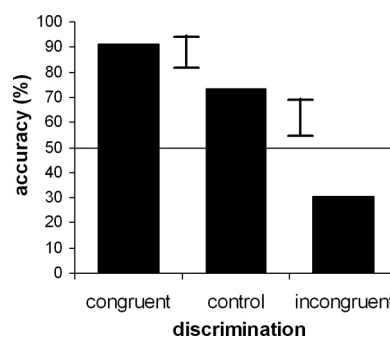

Figure 4. Outcome-devaluation test—behavioral performance. Average percentages of correct responses on congruent, control, and incongruent trials of the outcome-devaluation tests collapsed across the two sessions (inside the scanner). Error bars represent standard error of the difference (values were 6.4 for congruent − control and 7.6 for control − incongruent).

session (first/second) and discrimination (congruent/control/incongruent). This analysis yielded a significant effect of discrimination ($F_{(2,26)} = 34.97$, MSE = 779.5, $p < 0.0005$), and *post hoc* Tukey–Kramer analysis revealed that performance on the cue–outcome incongruent trials was inferior to that on the cue–outcome congruent and control trials, whereas there was no significant difference between the latter two.

A separate statistical analysis excluded the possibility of a speed–accuracy trade-off. Participants responded with average latencies of 1.2 s on cue–outcome congruent trials, and 1.6 s on both control and cue–outcome incongruent trials. A significant effect of discrimination ($F_{(2,26)} = 6.24$, MSE = 0.22, $p < 0.05$) was further investigated with *post hoc* Tukey–Kramer analysis, which revealed that performance was significantly faster on the congruent trials than on the control and incongruent trials, whereas response latencies on the latter two trial types were statistically indistinguishable.

Questionnaires

Questionnaire data are not available for one subject due to an oversight. Statistical analysis is therefore based on performance of 13 participants. The questionnaires revealed that participants remembered the common-outcomes/congruent/control/incongruent stimulus–response relationships equally well ($F = 1.13$), with average scores of 1.8 for the congruent, and 1.7 for the other three discriminations. In contrast, memory of the outcome pertaining to each component of the four biconditional discrimina-

tions did depend on discrimination type ($F_{(3,36)}$ = 6.20, MSE = 0.38, $p < 0.005$). *Post hoc* Tukey–Kramer analysis revealed that the participants were better at remembering the relationships between discriminative stimuli and outcomes when these were congruent (with an average score of 1.6) than when these were common outcomes (1.0), incongruent (1.0), or control (1.2), while memory of the latter three did not differ significantly.

Table 1. Neural correlates of goal-directed action during training and outcome devaluation test, and of response conflict during training

	Region	Coordinates (x, y, z)	Z score
Activation related to goal-directed behavior			
Control > incongruent[a]	Ventromedial PFC (right)	18, 40, 0	4.1*
		10, 18, 0	2.3**
Congruent > incongruent[a]	Ventromedial PFC	10, 32, −8	4.0*
		−2, 10, 30	2.6*
		0, −6, 30	2.3*
Control > incongruent[b]	Ventromedial PFC (right)	14, 44, 0	3.5*
		14, 42, 8	3.2*
Activation related to response conflict			
Incongruent > congruent[a]	Dorsomedial PFC	−2, 18, 28	4.5*
	Lateral PFC (right)	52, 28, 24	3.5**
		30, 28, −2	3.3**
		32, 14, 44	3.2**
	Lateral PFC (left)	−48, 26, 38	3.7*
		−42, 44, 2	3.3**
Incongruent > control[a]	No significant differences	N/A	N/A
Control > congruent[a]	Dorsomedial PFC	4, 20, 50	3.3**
		−34, 6, 40	3.4**
		−54, 18, 26	3.2**
		−44, 28, 38	3.0**

*$p < 0.05$ (FDR), functionally defined region of interest; **$p \leq 0.001$ (uncorrected), functionally defined region of interest.
[a]Contrast during learning.
[b]Contrast during outcome devaluation.

Neuroimaging results

(1) Investigating vmPFC activation in association with goal-directed learning

As above, this comparison identified voxels within the regions of interest that showed significantly greater response to the control and outcome-congruent conditions compared with the outcome-incongruent condition. Significant effects were found in a number of foci within ventromedial prefrontal cortex. The parameter estimates were 0.25 for congruent minus incongruent and 0.30 for control minus incongruent (SEMs = 0.07). The significant vmPFC activations are summarized in Table 1 and Figure 5.

(2) Investigating vmPFC activation in association with deployment of goal-directed knowledge when outcome values change

This analysis was confined to spheres of interest (radius = 10 mm) centered around the foci identified by the initial analysis above. It showed a significant effect of goal-directed responding in right vmPFC (see Table 1 and left panel of Fig. 6).

(3) Investigating the relationship between vmPFC activation during training and subsequent behavioral performance on the outcome-devaluation test

The right panel of Figure 6 illustrates the significant positive correlation ($r = 0.6$) between vmPFC activation within our region of interest during control trials of discrimination training, and subsequent behavioral performance on the outcome-devaluation test ($p < 0.05$; two-tailed).

(4) Identification of brain regions associated with response conflict

These contrasts highlighted activation in dorsomedial PFC and lateral PFC (see Table 1 and Fig. 7). These regions were activated more during the incongruent and control discriminations than during the congruent discrimination. The parameter estimates were 0.81 for incongruent minus congruent (SEM = 0.22) and 0.51 for control minus congruent (SEM = 0.19). In contrast, we did not find any significant activations within these regions in the contrast between the incongruent and control trials.
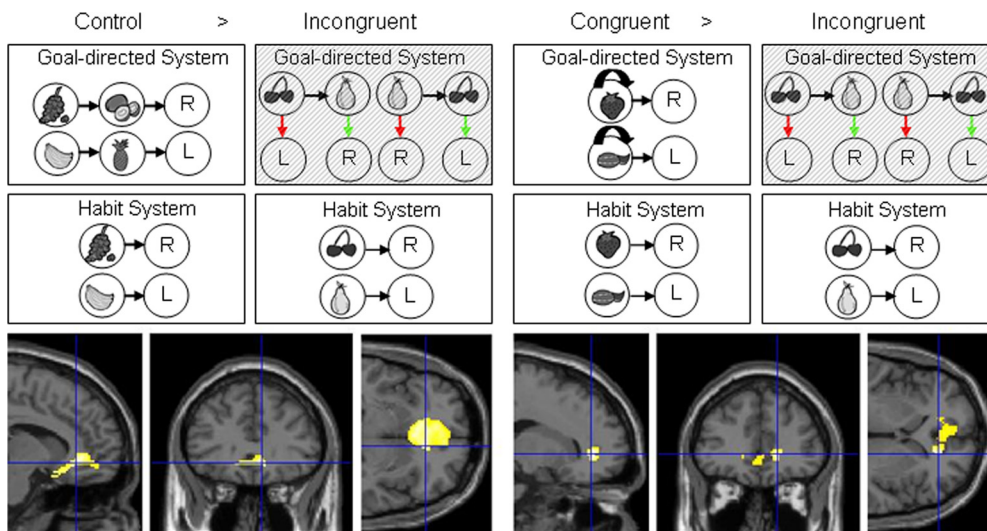

Figure 5. Instrumental discrimination training—vmPFC activations. The top panel illustrates that performance on the control (left) and congruent (right) discriminations can be supported by both the habit system and the goal-directed system, while performance on the incongruent discrimination can only be supported by the habit system. The middle panel shows the contrasts between cue–outcome congruent and control trials and cue–outcome incongruent trials, and implicates the ventromedial PFC in goal-directed learning. These orthogonal sections were taken at (10, 32, −8) and (18, 40, 0), respectively, and are thresholded at $p < 0.005$ uncorrected and at a cluster threshold of 15 voxels.

(5) Whole-brain analysis for the key comparisons (supplemental material, available at www.jneurosci.org)
No regions showed significant effects (corrected for multiple comparisons across the whole brain). For completeness, we do report, in supplemental Table 1 (available at www.jneurosci.org as supplemental material), those regions surviving an uncorrected threshold ($p < 0.001$, voxel extent = 15 voxels). The purpose of this supplemental table is to provide information to the interested reader, though we are reluctant to make any interpretation of regions listed in this table as they are outside the regions of interest and do not survive whole-brain correction.

(6) Investigating striatal activation during all discriminations relative to baseline (supplemental material, available at www.jneurosci.org)
According to our theoretical analysis the neural structures supporting habit formation should be activated equally during control, congruent, and incongruent discriminations. Indeed, we failed to find significant activation in the candidate region of dorsal striatum, in the whole-brain contrasts [incongruent > control] and [incongruent > congruent]. We should, however, predict that this region is more active during training of all three discriminations than during baseline (fixation). Although the main aim of this study was to investigate the neural substrate of goal-directed action, we wished to confirm that the striatum was involved in discrimination learning per se. To this end, we conducted an ROI analysis with a striatal mask (see supplemental Fig. 1, available at www.jneurosci.org as supplemental material) comparing activation during discrimination training with baseline. In keeping with the idea that habitual support was common to all discriminations, we found modest but significant activations within left dorsal striatum ($-20, -2, 18; Z = 3.9, p_{FDR} = 0.05$) and right dorsal striatum ($20, -10, 20; Z = 3.2; p_{uncorrected} < 0.001$).

Discussion

Under conditions in which goal-directed responding predominates, ventromedial PFC activation is significantly higher than when performance is purely habitual. Activation in this area during outcome-based responding in a devaluation test provides additional support for the role of the vmPFC in goal-directed control. The data are thus consistent with previous studies that have implicated the vmPFC in the deployment of goal-directed knowledge (Valentin et al., 2007; Tanaka et al., 2008; Gläscher et al., 2009).

According to the dual-system account, instrumental discriminations are learnt through the concurrent build-up of behavioral control in a goal-directed and a habitual system. Initially the goal-directed system exerts dominant behavioral control, but with extensive practice the habit system takes over. This dual-system view is supported by animal lesion studies demonstrating neural dissociations (Corbit and Balleine, 2003; Yin et al., 2004, 2005). Moreover, previous research has shown that humans, as well as animals, are able to circumvent behavioral control by the goal-directed system when this is required to prevent conflict and thereby perform successfully (de Wit et al., 2007), a finding that was replicated in the present study.

In recent years the discovery of homologous brain areas in animals and humans has provided the impetus for translational research in the field of human decision making. However, whereas in animals lesion work can aid the dissociation of the neural substrates of goal-directed versus habitual learning, such an analysis is more challenging in neuroimaging research with healthy volunteers. In a recent study, Valentin et al. (2007) showed that instrumental choice of a nonprefed over a prefed liquid is reflected in vmPFC activation. On the basis of activation of this same area during the initial acquisition phase, the authors argued that this area is not only important for performance on the satiety test but also for the acquisition of goal-directed knowledge or learning. However, because habit reinforcement may have taken place concurrently during the acquisition phase, their analysis does not allow one to isolate the neural substrate of goal-directed learning.

In the present study, we were able to circumvent this issue by training participants on (congruent and control) discriminations that can be solved by both systems, as well as on a (incongruent) discrimination that relies predominantly on habitual control, as confirmed with an outcome-devaluation test following on train-
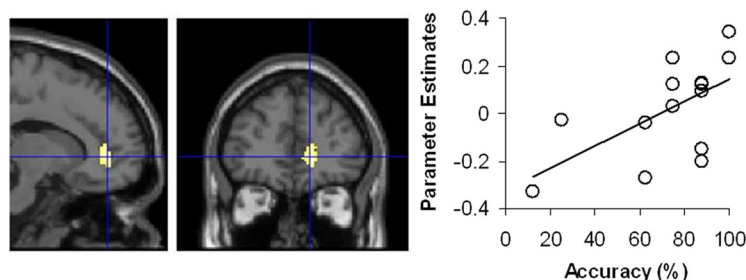

Figure 6. Outcome-devaluation test—vmPFC activations. Left, Contrast between goal-directed control performance and habitual cue–outcome incongruent performance during the test phase. The ventromedial PFC is implicated in goal-directed performance on the outcome-devaluation test. These orthogonal sections were taken at (14, 44, 0), and are thresholded at $p < 0.005$ uncorrected and at a cluster-threshold of 15 voxels. Right, Positive correlation between ventromedial PFC activation during training and behavioral performance during the subsequent outcome-devaluation test.
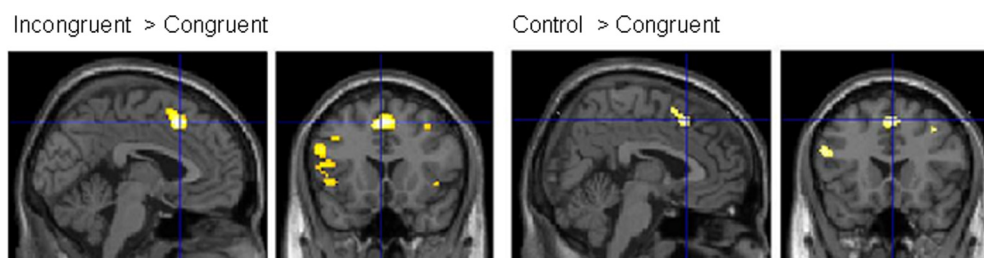

Figure 7. Instrumental discrimination training— dmPFC and lPFC activations. The dorsomedial and lateral PFC are engaged less during the congruent discrimination than during the incongruent and control discriminations. These orthogonal sections were taken at ($-2, 18, 48$) and (4, 20, 50), respectively, and are thresholded at $p < 0.005$ uncorrected and at a cluster-threshold of 15 voxels.

ing. In line with the goal-directed account of vmPFC function, the contrast of activity during the goal-directed discriminations with that during the habitual discrimination yielded significant activations in this region. This analysis therefore allowed us to demonstrate that the vmPFC is recruited more when goal-directed learning takes place than when performance relies solely on an S→R reinforcement mechanism.

This analysis of instrumental discrimination learning does not, however, allow us to rule out yet another competing account of vmPFC, namely that it is involved in Pavlovian learning. It is generally recognized that embedded within instrumental discriminations are Pavlovian S→O relationships brought about by the simple pairing of the discriminative stimulus and instrumental outcome. Consequently, the differential PFC activation may reflect a purely Pavlovian contribution to instrumental discriminative control, rather than learning about the (R→O) relationships between actions and goals. If participants ignored the reward pictures in the incongruent discrimination, this may well have reduced Pavlovian learning relative to the other discriminations. In fact, neuroimaging research has so far used tasks with a distinct Pavlovian component. For example, in the study by Valentin et al. (2007) a purely Pavlovian account could not be excluded because Pavlovian cues were present during the satiety extinction test (as acknowledged by the authors). This is particularly problematic because we know that Pavlovian learning is susceptible to outcome devaluation (Colwill and Motzkin, 1994). Moreover, the vmPFC has been implicated in Pavlovian conditioning (O'Doherty et al., 2002) as well as in devaluation of Pavlovian outcomes (Gottfried et al., 2003).

With the aim of establishing that the vmPFC region identified in this study is implicated in goal-directed action selection through O→R associations, rather than simply Pavlovian S→O learning, we used an outcome-devaluation procedure that forced participants to choose between two actions on the basis of current outcome value, in the absence of any cues to guide performance. When we contrasted performance during goal-directed (control) trials with that during the habitual (incongruent) trials, we found significant activation in the vmPFC. Moreover, we found that vmPFC activation during training on the control discrimination predicted subsequent behavioral performance in the control trials of the outcome-devaluation test. These results are consistent with demonstrations that the rodent prelimbic cortex—which has been suggested to be homologous to parts of the vmPFC in humans (Rushworth et al., 2007) (but see Seamans et al., 2008)—is critical for goal-directed learning (Ostlund and Balleine, 2005).

In addition to providing insights into the neural substrates of goal-directed action, the behavioral observations in this study replicate a previous demonstration of flexible behavioral control in humans(de Wit et al., 2007). So far, most research on flexible control has focused on the management of conflict that arises as a consequence of competing S→R associations [e.g., Stroop task (Stroop, 1935); Go/No-Go task; flanker test (Eriksen and Eriksen, 1974)]. A noteworthy aspect of the current demonstration is that response conflict was evoked in the goal-directed system. The ability to switch control away from the goal-directed system can be of crucial importance, because habitual behavior has the advantage of requiring relatively little cognitive effort generally, and because it allows one to prevent conflict due to conflicting O→R associations. In the latter case, we have shown that humans will switch control to the habit system to allow for successful performance.

The question arises whether there is an active arbitrator between the goal-directed and habit system. Daw et al. (2005) developed a computational model of instrumental behavior that resembles the associative dual-system account. In their model the goal-directed and habitual pathways compete for behavioral control, and the brain appropriately selects the pathway that is expected to be most accurate. It is beyond the scope of the present study to identify the arbitrator as we should expect all discriminations to engage this mechanism, but we did inspect activation in conflict-related areas because we predicted that the incongruent discrimination would give rise to response conflict in the goal-directed system. Although we replicated an earlier finding that the control condition engaged lateral PFC more than the congruent condition (Roelofs et al., 2006), we failed to find evidence for stronger engagement of this area during the incongruent relative to the control condition. We should be cautious in interpreting this null effect, but one possibility is that we did not replicate this finding of related previous studies, because in the present study conflict arose as a consequence of O→R associations rather than competing S→R associations. Alternatively, conflict may have been prevented rather than resolved, through a shift toward habitual control, possibly by an online arbitrator between the goal-directed and habitual system. The lack of conflict-related activation in the incongruent-control contrast does therefore not speak against our account of incongruent performance, according to which the participants successfully adopted a habitual strategy to solve the incongruent discrimination.

Interestingly, previous studies with a rodent version of the paradigm showed that temporary inactivation of the dmPFC selectively impaired incongruent, but not congruent and control performance (de Wit et al., 2006, 2009). Although these findings may appear to be at variance with the brain activation we observed in humans, it may well reflect different behavioral strategies to resolve the conflict inherent in the incongruent discrimination. Whereas humans adopted a habit strategy, rats used a complex, goal-directed strategy that appeared to crucially depend upon active cognitive control by the dmPFC. The choice of strategy possibly depends upon the types of S/O events used [see de Wit et al. (2007) for a more elaborate discussion].

The importance of the ability to prioritize the most appropriate system, habitual or goal directed, becomes particularly clear when flexible control is impaired. An inability to switch to habits is thought to render even simple everyday activities effortful for Parkinson's disease patients, whereas the ability to shift toward goal-directed control may be impaired in drug abusers (Dickinson et al., 2002; Miles et al., 2003; Yin and Knowlton, 2006; Everitt et al., 2008; Grahn et al., 2008; Rangel et al., 2008; Redish et al., 2008) and patients with obsessive-compulsive disorder (Evans et al., 2004). The vmPFC has been implicated in both addiction and in obsessive-compulsive disorder (Everitt et al., 2007; Menzies et al., 2008), and further insights into the role of this area in goal-directed and habitual mechanisms may therefore further our understanding of adaptive as well as compulsive, maladaptive decision making.

In conclusion, we used a novel conflict task that forces participants to rely on habitual control to show unequivocally that the vmPFC is involved in goal-directed action. This is the first demonstration in humans that the vmPFC is engaged more during the acquisition of goal-directed behavior than that of habits. These findings therefore make an important contribution to our understanding of the neural mechanisms of goal-directed action in humans.

References

Adams CD (1982) Variations in the sensitivity of instrumental responding to reinforcer devaluation. Q J Exp Psychol B 34:77–98.

Adams CD, Dickinson A (1981a) Instrumental responding following reinforcer devaluation. Q J Exp Psychol B 33:109–121.

Adams CM, Dickinson A (1981b) Actions and habits: variations in associative representations during instrumental learning. In: Information processing in animals: memory mechanisms (Spear NE, Miller RR, eds), pp 143–165. Hillsdale, NJ: Erlbaum.

Asratyan EA (1974) Conditional reflex theory and motivated behavior. Acta Neurobiol Exp (Wars) 34:15–31.

Balleine BW, Ostlund SB (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. Ann N Y Acad Sci 1104:147–171.

Colwill RM, Motzkin DK (1994) Encoding of the unconditioned stimulus in pavlovian conditioning. Anim Learn Behav 22:384–394.

Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. Behav Brain Res 146:145–157.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.

de Wit S, Dickinson A (2009) Associative theories of goal-directed behaviour: a case for animal–human translational models. Psychol Res 73:463–476.

de Wit S, Kosaki Y, Balleine BW, Dickinson A (2006) Dorsomedial prefrontal cortex resolves response conflict in rats. J Neurosci 26:5224–5229.

de Wit S, Niry D, Wariyar R, Aitken MRF, Dickinson A (2007) Stimulus-outcome interactions during conditional discrimination learning by rats and humans. J Exp Psychol Anim Behav Process 33:1–11.

de Wit S, Ostlund SB, Balleine BW, Dickinson A (2009) Resolution of conflict between goal-directed actions: outcome encoding and neural control processes. J Exp Psychol Anim Behav Process 35:382–393.

Dickinson A (1985) Actions and habits: the development of behavioural autonomy. Philos Trans R Soc Lond B Biol Sci 308:67–78.

Dickinson A (1994) Instrumental conditioning. In: Animal learning and cognition (Mackintosh NJ, ed), pp 45–79. San Diego: Academic.

Dickinson A, Balleine B (1993) Actions and responses: the dual psychology of behaviour. In: Spatial representation: problems in philosophy and psychology (Eilan N, McCarthy RA, eds), pp 277–293. Malden, MA: Blackwell.

Dickinson A, Balleine B (1994) Motivational control of goal-directed action. Anim Learn Behav 22:1–18.

Dickinson A, de Wit S (2003) The interaction between discriminative stimuli and outcomes during instrumental learning. Q J Exp Psychol B 56:127–139.

Dickinson A, Wood N, Smith JW (2002) Alcohol seeking by rats: action or habit? Q J Exp Psychol B 55:331–348.

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept Psychophys 16:143–149.

Evans DW, Lewis MD, Iobst E (2004) The role of the orbitofrontal cortex in normally developing compulsive-like behaviors and obsessive-compulsive disorder. Brain Cogn 55:220–234.

Everitt BJ, Hutcheson DM, Ersche KD, Pelloux Y, Dalley JW, Robbins TW (2007) The orbital prefrontal cortex and drug addiction in laboratory animals and humans. Ann N Y Acad Sci 1121:576–597.

Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW (2008) Review. Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. Philos Trans R Soc Lond B Biol Sci 363:3125–3135.

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15:870–878.

Gläscher J, Hampton AN, O'Doherty JP (2009) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb Cortex 19:483–495.

Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301:1104–1107.

Grahn JA, Parkinson JA, Owen AM (2008) The cognitive functions of the caudate nucleus. Prog Neurobiol 86:141–155.

Hommel B (2003) Planning and representing intentional action. Scientific-WorldJournal 3:593–608.

James W (1890) The principles of psychology. New York: Dover.

Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. Science 303:1023–1026.

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex 13:400–408.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19:1233–1239.

Marsh AA, Blair KS, Vythilingam M, Busis S, Blair RJ (2007) Response options and expectations of reward in decision-making: the differential roles of dorsal and rostral anterior cingulate cortex. Neuroimage 35:979–988.

Menzies L, Chamberlain SR, Laird AR, Thelen SM, Sahakian BJ, Bullmore ET (2008) Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. Neurosci Biobehav Rev 32:525–549.

Miles FJ, Everitt BJ, Dickinson A (2003) Oral cocaine seeking by rats: action or habit? Behav Neurosci 117:927–938.

Milham MP, Banich MT, Claus ED, Cohen NJ (2003) Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. Neuroimage 18:483–493.

Nelson HE (1982) National adult reading test (nart) test manual. Windsor, UK: NFER-Nelson.

O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. Neuron 33:815–826.

Ostlund SB, Balleine BW (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J Neurosci 25:7763–7770.

Pavlov IP (1932) The reply of a physiologist to psychologists. Psychol Rev 39:91–127.

Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9:545–556.

Redish AD, Jensen S, Johnson A (2008) A unified framework for addiction: vulnerabilities in the decision process. Behav Brain Sci 31:415–437; discussion 437–487.

Roelofs A, van Turennout M, Coles MG (2006) Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. Proc Natl Acad Sci U S A 103:13884–13889.

Rushworth MF, Walton ME, Kennerley SW, Bannerman DM (2004) Action sets and decisions in the medial frontal cortex. Trends Cogn Sci 8:410–417.

Rushworth MF, Buckley MJ, Behrens TE, Walton ME, Bannerman DM (2007) Functional organization of the medial frontal cortex. Curr Opin Neurobiol 17:220–227.

Seamans JK, Lapish CC, Durstewitz D (2008) Comparing the prefrontal cortex of rats and primates: insights from electrophysiology. Neurotox Res 14:249–262.

Stroop JR (1935) Studies of interference in serial verbal reactions. J Exp Psychol 18:643–662.

Sutton RS, Barto AG (1981) An adaptive network that constructs and uses an internal model of its world. Cogn Brain Theory 4:217–246.

Tanaka SC, Balleine BW, O'Doherty JP (2008) Calculating consequences: brain systems that encode the causal effects of actions. J Neurosci 28:6750–6755.

Thorndike EL (1911) Animal intelligence: experimental studies. New York: Macmillan.

Thorndike EL (1931) Human learning. New York: Century.

Urcuioli PJ (2005) Behavioral and associative effects of differential outcomes in discrimination learning. Learn Behav 33:1–21.

Valentin VV, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. J Neurosci 27:4019–4026.

van Veen V, Holroyd CB, Cohen JD, Stenger VA, Carter CS (2004) Errors without conflict: implications for performance monitoring theories of anterior cingulate cortex. Brain Cogn 56:267–276.

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464–476.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189.

Yin HH, Knowlton BJ, Balleine BW (2005) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505–512.