

# Calculating Consequences: Brain Systems That Encode the Causal Effects of Actions

Saori C. Tanaka,<sup>1,3</sup> Bernard W. Balleine,<sup>4</sup> and John P. O'Doherty<sup>1,2</sup>

<sup>1</sup>Division of the Humanities and Social Sciences and <sup>2</sup>Computation and Neural Systems Program, California Institute of Technology, Pasadena, California 91125, <sup>3</sup>Advanced Telecommunications Research Institute International Computational Neuroscience Laboratories, Kyoto 619-0288, Japan, and

<sup>4</sup>Department of Psychology and the Brain Research Institute, University of California, Los Angeles, California 90024

The capacity to accurately evaluate the causal effectiveness of our actions is key to successfully adapting to changing environments. Here we scanned subjects using functional magnetic resonance imaging while they pressed a button to earn money as the response–reward relationship changed over time. Subjects' judgments about the causal efficacy of their actions reflected the objective contingency between the rate of button pressing and the amount of money they earned. Neural responses in medial orbitofrontal cortex and dorsomedial striatum were modulated as a function of contingency, by increasing in activity during sessions when actions were highly causal compared with when they were not. Moreover, medial prefrontal cortex tracked local changes in action–outcome correlations, implicating this region in the on-line computation of contingency. These results reveal the involvement of distinct brain regions in the computational processes that establish the causal efficacy of actions, providing insight into the neural mechanisms underlying the adaptive control of behavior.

**Key words:** goal-directed behavior; contingency; causality; free-operant conditioning; fMRI; prefrontal cortex

## Introduction

The capacity of humans and other animals to detect the causal effect of their actions on environmental events is a critical determinant of adaptive behavior allowing the acquisition and performance of new and existing behavioral strategies to be regulated by their consequences (Dickinson and Balleine, 1993; Balleine and Dickinson, 1998, 2000). The on-line detection of changes in the causal efficacy of actions relies on the computation of temporal correlations between the rate of performance and the occurrence of environmental events, particularly the relationship between actions and motivationally significant events such as rewards (Baum, 1973; Dickinson, 1994; Dickinson and Balleine, 1994). Because responding is effortful, it is of considerable advantage for animals to encode the likelihood that an action will result in a valued consequence and so increase performance when that likelihood is high or reduce performance if it is low. Contingency is the term used by behavioral psychologists to describe the relationship between an action and its consequences or outcome, which is defined in terms of the difference between two probabilities: the probability of the outcome given the action is performed and the probability of the outcome given the action is not performed (Hammond, 1980; Beckers et al., 2007). Sensitivity to contingency has been shown to be a key feature of goal-directed

behavior in rodents, and, at a neural level, recent evidence suggests that this capacity is mediated by a corticobasal ganglia circuit involving the prefrontal region of rat medial prefrontal cortex and the dorsomedial striatum (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Yin et al., 2005). Analyses based on a constellation of deficits observed in patients with damage to the frontal lobe have generated the suggestion that particularly the prefrontal cortex plays a general role in planning and perhaps the encoding of goal-directed action (Milner, 1982; Bechara et al., 1994; Rolls et al., 1994), although whether these structures are in fact involved in the on-line encoding of the contingent or indeed the causal relation between action and outcome has not been addressed.

The goal of this investigation was to determine the neural substrates of contingency detection in humans, together with its subjective concomitant: the subject's judgment of the causal effectiveness of his/her own actions. To achieve this, we forsook the traditional trial-based approach, typically used in experiments using humans and nonhuman primates, in which subjects are cued to respond at particular times in a trial, for the unsignaled, self-paced approach more often used in studies of associative learning in rodents in which the subjects themselves choose when to respond. This approach allowed us to assay a subject's behavioral sensitivity to changes in the contingency between responding and rewards, and to measure [with blood oxygen level-dependent (BOLD) functional magnetic resonance imaging (fMRI)] neural responses related to detection of these contingencies.

## Materials and Methods

**Subjects.** Fourteen healthy right-handed volunteers (seven males and seven females) participated in the study. The volunteers were preassessed

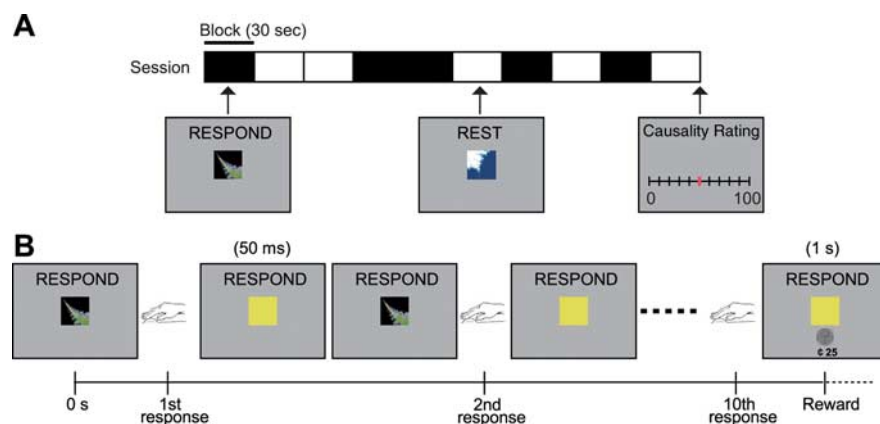
Received Nov. 28, 2007; accepted May 20, 2008.

This work was supported by a grant from National Institute of Mental Health to J.P.O.D. and by grants from the Gordon and Betty Moore Foundation to J.P.O.D. and the Caltech Brain Imaging Center. S.C.T. is funded by research fellowships from the Japan Society for the Promotion of Science for Young Scientists.

Correspondence should be addressed to John P. O'Doherty at the above address. E-mail: joherty@hss.caltech.edu.

DOI:10.1523/JNEUROSCI.1808-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/286750-06\$15.00/0



**Figure 1.** Experimental task. **A**, Example of the structure of a single session. Each experiment consisted of four sessions lasting 5 min each. A single session includes five “RESPOND” blocks, in which the subject made button presses, and five “REST” blocks. At the end of each session, subjects rated how causal their button presses were in earning money on a scale from 0 to 100. **B**, Example of the event schedule of the RESPOND block. When the subject pressed the button, the stimulus on the screen turned yellow for 50 ms. Reward (25 cents) was delivered at the first button press after the 10th response (on average) on the VR10 schedule.

to exclude those with a previous history of neurological or psychiatric illness. All subjects gave informed consent, and the study was approved by the Institutional Review Board of the California Institute of Technology. One subject was later excluded from the analysis because of a complete lack of responding on one of the schedules.

**Experimental procedures.** To maximize experimental variability in the response–reward contingencies experienced by our subjects, we used two different types of reward schedule: variable-ratio (VR) schedules, in which subjects were rewarded according to the number of responses performed, and variable-interval (VI) schedules, wherein subjects were rewarded not in proportion to the number of responses made, but according to the interval between successive rewards. Because of methodological constraints imposed by the fMRI method, we randomly interspersed 30 s blocks of responding on these different schedules with rest periods (“REST” block) (Fig. 1A), which were explicitly cued to the subjects. Otherwise, responding was self-paced and unconstrained during the “active” periods (“RESPOND” block). The order of presentation of the blocks was randomized throughout.

Within each block of responding, subjects were invited to freely press a button as often as they liked to obtain monetary rewards, which were in 25 cent units distributed according to four schedules of reinforcement (Fig. 1B): a VR10 schedule, in which subjects are rewarded on average for every 10 responses; a VI4 schedule, in which subjects are reinforced on average every 4 s; a VR-yoked schedule, in which the number of responses made for reinforcement during performance of the VI4 schedule from the preceding subject; and a VI-yoked schedule, in which subjects were reinforced according to the intervals to reward experienced by the preceding subject during performance of the VR10 schedule.

At the end of each session, subjects were asked to rate how causal their actions were, i.e., whether making a response caused them to receive money, using a scale from 0 to 100, where 0 indicated not causal and 100 indicated strongly causal. Subjects completed four sessions of 5 min, each associated with a specific schedule the order of which was counterbalanced across subjects.

In this study, we used both variable-ratio and short variable-interval schedules to allow subjects to sample across a broad contingency space to create variance in both subjective causality judgment and objective contingency. The decision to use both variable-ratio and variable-interval schedules was based on previous findings in rodents suggesting that these schedules produce the greatest variation in the experienced action–outcome contingency (Dickinson et al., 1983; Dawson and Dickinson, 1990; Dickinson, 1994). This variability in contingency is partly explained by the fact that interval schedules typically decouple response rate and reward rate, leading to a lower contingency on those schedules, particularly compared with ratio schedules, which, by virtue of the tight coupling

between response rate and rewards, can lead to a higher contingency estimates. However, in the present study we did not find either consistent or significant differences in the degree of contingency or in causality judgment across these schedules, largely because we used rather short intervals in the VI schedules (VI4), so as to match the inter-reinforcement interval as closely as possible to the low VR (VR10) schedules. Therefore, instead of comparing across schedules, we took advantage of intrinsic variability in the degree of contingency across schedules within subjects, by comparing responses elicited on schedules with high and low objective contingencies within each subject regardless of which schedule fell into these categories between subjects.

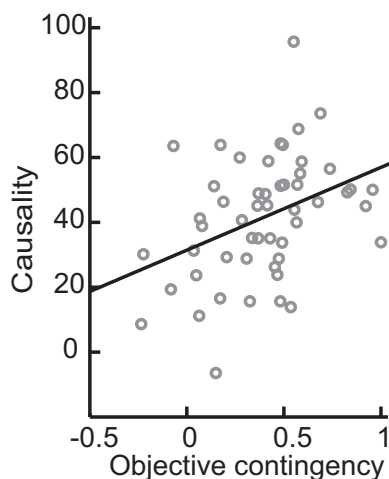
**Imaging procedures.** A 3 Tesla scanner (MAGNETOM Trio; Siemens) was used to acquire both structural T1-weighted images and T2\*-weighted echoplanar images (repetition time = 2.81 s; echo time = 30 ms; flip angle = 90°; 45 transverse slices; matrix = 64 × 64; field of view = 192 mm; thickness = 3 mm; slice gap = 0 mm) with BOLD contrast. To recover signal loss from dropout in the medial orbitofrontal cortex (mOFC) (O’Doherty et al., 2002), each horizontal section was acquired at 30° to the anterior commissure–posterior commissure axis.

**Data analysis.** We used SPM2 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) for preprocessing and statistical analyses. The first four volumes of images were discarded to avoid T1 equilibrium effects. The images were realigned to the first image as a reference, spatially normalized with respect to the Montreal Neurological Institute (MNI) echoplanar imaging template, and spatially smoothed with a Gaussian kernel (8 mm, full width at half-maximum). We used high-pass filter with cutoff = 200 s.

For each subject, we constructed an fMRI design matrix by modeling each “respond” period within a session as a 30-s-long block. To allow for initial learning and stabilization of responding, we modeled the first RESPOND block separately from the other four RESPOND blocks. Behavioral analysis confirmed no significant differences between the second and fifth blocks (using paired *t* tests) on a range of behavioral measures including response number, response rate, number of responses per reinforcer, intervals per reinforcer, and total reinforcers obtained, confirming that learning effects were stable by the end of the first RESPOND block and thereby justifying our inclusion of the last four blocks as representative of stable responding sessions. These regressors were convolved with a canonical hemodynamic response function. Motion parameters were entered as additional regressors to account for residual effects of head motion. The design matrix was then regressed against the fMRI data to generate parameter estimates for each subject. Contrasts of parameter estimates between sessions were then entered into a subsequent between-subject analysis to generate group-level random-effects statistics. All coordinates indicate the MNI coordinate system.

To compute the objective contingency for each schedule as experienced by the subjects, we divided up each session into 10 s bins and counted the number of responses performed within each 10 s bin and the number of outcomes received in each bin. We then tabulated these two variables and computed the overall correlation across bins between the number of responses made and the number of outcomes received per bin across the whole session. Those schedules with a high correlation coefficient are, therefore, those with a high contingency between the rate of change of responding over time and the rate of reward delivery, whereas schedules with a lower correlation generate a weaker response–reward contingency.

For the correlation analyses reported in Figure 3, *D* and *E*, we performed a more fine-grained analysis of the local computation of contingency within each 10 s interval of responding for the subjects. We divided each 10 s time window into 200 ms bins and created two vectors (of



**Figure 2.** Plot of causality value against global objective correlation.

length 50), one for responses and one for rewards. We entered the number of responses that occurred in each 200 ms bin within the interval. If no response occurred within that bin, a zero was entered. Because of the short time length of the bin, for each bin usually either no response occurred or a single response occurred. Similarly, for the reward vector, we entered for each bin the number of rewards obtained in that 200 ms interval and zero otherwise. We then computed a correlation between the number of responses and rewards obtained in each bin across the 10 s interval.

We computed the correlation between the number of responses performed and the number of rewards obtained for each 200 ms bin across every 10 s time window. We then extracted the BOLD signal averaged across those voxels in each region of interest found to show significant effects in the contrast of high – low objective contingency at  $p < 0.01$ , and averaged the resulting time series into 10 s bins. We next performed a regression analysis of the binned BOLD data against the local objective contingency, after adjusting the data for the effects of between subject variance.

## Results

### Behavioral results

We found a highly significant correlation between objective contingency and subjective causality ratings calculated across all subjects and sessions ( $R^2 = 0.63$ ;  $p = 0.5 \times 10^{-5}$ ) (Fig. 2). To assess this result further, we took the schedule with the highest and lowest objective contingency measure for each subject and compared their associated causality judgments (the specific schedules assigned to each condition for each subject are listed in supplemental Table 1, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). The high-contingency schedules ( $65.8 \pm 5.3$ ; mean  $\pm$  SE) were associated with significantly higher causality judgments than the low-contingency schedules ( $41.5 \pm 7.2$ ), across subjects ( $t = 3.771$ ;  $df = 12$ ;  $p < 0.005$ , one-tailed paired  $t$  test). These results suggest that subjects were sensitive to relative changes in contingency and that this was reflected in their subjective causality judgments. It is possible, however, that the different schedules induced other effects that could have produced the changes in causality and, to assess this, we also compared other possible differences between the contingencies that could potentially have influenced the results, including the intervals between successive rewards (high,  $4.21 \pm 0.94$ ; low,  $5.23 \pm 0.77$ ), the number of responses made per reward (high,  $9.35 \pm 0.97$ ; low,  $10.7 \pm 1.2$ ), and the total number of rewards obtained (high,  $357 \pm 45$ ; low,  $274 \pm 35$ ). None of these measures showed significant differences as a function of contingency (at  $p < 0.05$  in paired  $t$  tests), helping

to rule out confounding explanations for the subsequent imaging results. Response rates and variability in response rates across each session are shown in supplemental Figures 1 and 2, respectively (available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). Paired  $t$  tests revealed no significant difference in the overall response rates (high,  $2.74 \pm 0.30$ ; low,  $2.24 \pm 0.34$ ), nor in the variance in response rates (high,  $0.695 \pm 0.092$ ; low,  $0.685 \pm 0.094$ ) as a function of contingency (high vs low).

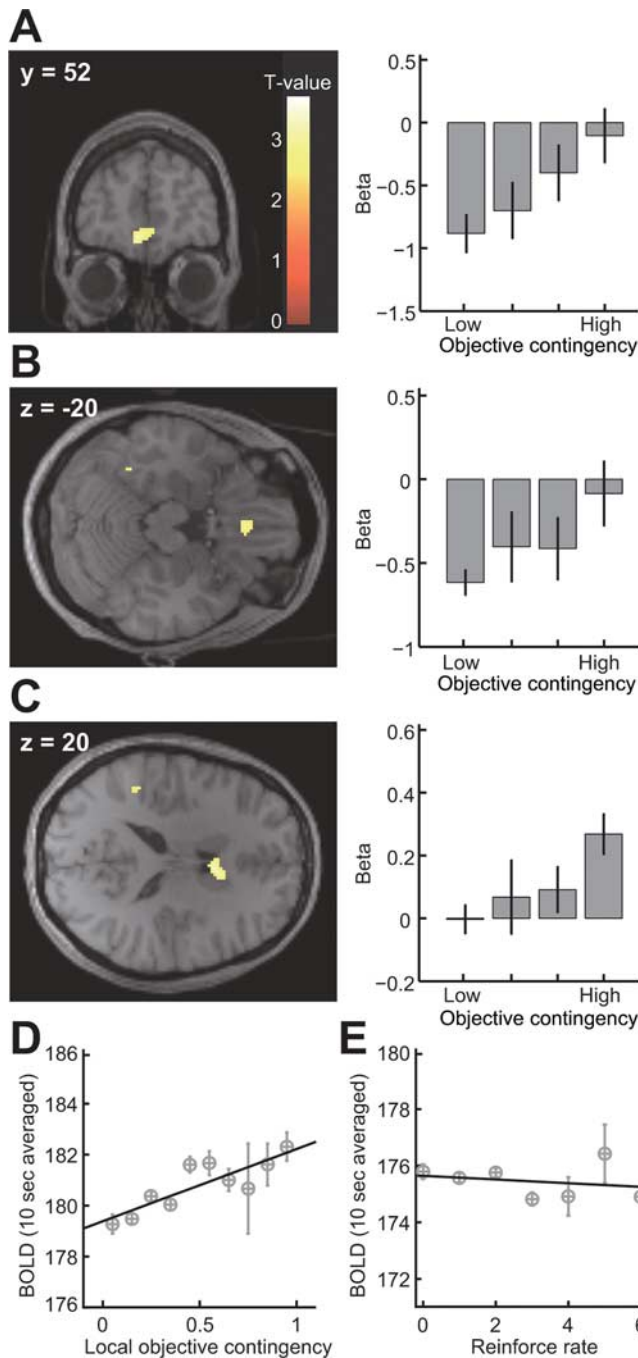
### fMRI results: effects of contingency

We next contrasted the average evoked BOLD signal during the high-contingency schedule with that elicited during the low-contingency schedule, to detect brain regions showing changes in activity as a function of differences in objective contingency. We found that three regions in particular showed significant effects of contingency: the medial prefrontal cortex (mPFC) [significant at  $p < 0.05$ , corrected for small volume (SVC)] (Fig. 3A), the mOFC ( $p < 0.05$ , SVC) (Fig. 3B), and the dorsomedial striatum (specifically anterior medial caudate nucleus;  $p < 0.001$  uncorrected) (Fig. 3C) (see supplemental Table 2, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material).

We then looked at a finer 200 ms time scale to see how neural activity in our regions found to be sensitive to contingency changed over time as a function of local fluctuations in the correlation between responses and rewards during performance. We computed the local objective contingency within each 10 s time interval of task performance, by counting the number of responses and rewards in 200 ms bins within that interval, and computing the contingency across the whole 10 s window between these variables (see Materials and Methods). We computed the correlation between average evoked BOLD signal in each 10 s window and the local objective contingency from each of the areas that we previously found to be sensitive to contingency, and we found a highly significant correlation between the local objective contingency and averaged BOLD signal in only one of three of the areas: the mPFC ( $R^2 = 0.72$ ;  $p = 0.0021$ ) (Fig. 3D). No significant correlations were found in the other two regions. Although mPFC is known to also respond to receipt of rewarding outcomes (Elliott et al., 1997; O'Doherty et al., 2001; Knutson et al., 2003), there was no significant correlation between the overall reward rate and activity in this area ( $R^2 = 0.053$ ;  $p = 0.62$ ) (Fig. 3E), ruling that out as a potential explanation for our results. These findings suggest, therefore, that the medial prefrontal cortex is involved in the on-line computation of contingency.

### fMRI results: subjective causality

Finally, we tested for areas showing changes in activity related directly to the subjects' own subjective causality ratings over sessions. A comparison between schedules with high compared with low causality ratings revealed significant effects in the mPFC (Fig. 4A). Although many other areas were also activated in this contrast (supplemental Table 3, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material), a plot of the parameter estimates for each of the activated areas revealed that mPFC was one of only three regions showing a linearly increasing response profile as a function of increasing causality judgments across all of the four sessions for each subject (the other two regions showing linear changes with causality are lateral OFC and a more dorsomedial area of prefrontal cortex shown in supplemental Fig. 3, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). This result suggests that mPFC is not only involved in computing the local objective contingency between responding and rewards, but that

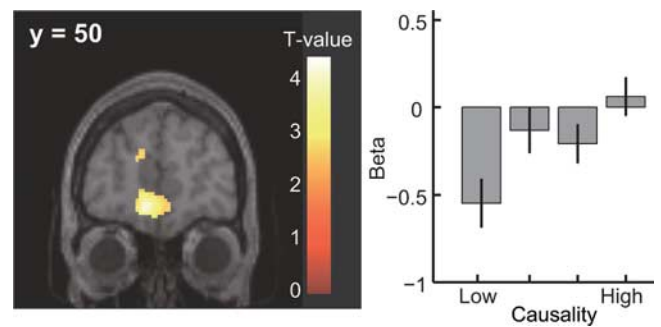


**Figure 3.** *A–C*, Voxels showing significant correlations with the global objective correlation in the mPFC [ $(x, y, z) = (-6, 52, -10)$ ;  $p < 0.05$ , SVC with a 5 mm sphere centered at  $(6, 57, -6)$  (Hampton et al., 2006); *A*], mOFC [ $(x, y, z) = (-2, 30, -20)$ ;  $p < 0.05$ , SVC with a 5 mm sphere centered at  $(-3, 36, -24)$  (Valentin et al., 2007); *B*], and dorsomedial striatum [anterior caudate nucleus;  $(x, y, z) = (6, 10, 20)$ ;  $p < 0.001$ , uncorrected]; *C*]. Bar plots show parameter estimated at each peak voxel sorted by global objective correlation. *D*, Plot of averaged BOLD signal extracted from mPFC, against the local objective correlation and regression slope ( $R^2 = 0.72$ ;  $p = 0.0021$ ). *E*, Plot of averaged BOLD signal extracted from mPFC against reinforcement rate and regression slope ( $R^2 = 0.053$ ;  $p = 0.62$ ). Error bars = 1 SEM.

activity in this area also tracks subjective judgments about the causal effectiveness of a subject's own behavior.

### Discussion

Our findings implicate a network of brain regions involving the medial prefrontal cortex, medial orbitofrontal cortex, and dorsal



**Figure 4.** Voxels showing significant activation in the high-causality condition in the mPFC [ $(x, y, z) = (-10, 50, -10)$ ;  $p < 0.001$ , uncorrected], and parameter estimates sorted at the peak voxel of the mPFC by causality value. Error bars = 1 SEM.

striatum (specifically anterior medial caudate nucleus) in computing the causal effectiveness of an individual's own behavior (Balleine, 2005; Balleine and Ostlund, 2007). These findings suggest that this network of brain regions may be responsible for the adaptive control of action selection in situations in which the temporal relationship between actions performed and rewards obtained vary over time. Sensitivity to the contingency between actions and reward delivery is indicative of goal-directed or action–outcome learning in rats (Balleine and Dickinson, 1998). Thus, the areas identified in the present study are also candidate regions for mediating goal-directed action selection in humans.

The results of the present study also demonstrate the utility of using a free-operant paradigm to study human instrumental learning. Typically in the human and indeed nonhuman primate literature, action selection is studied in a trial-based manner, in which after the onset of a cue, a single response is triggered. However, in the free-operant case, responding is unsigned and self-generated, thereby allowing us to explore the means by which subjects can modulate their responses as a function of changes in reward contingencies over time, an issue not easily addressable through standard trial-based approaches. Furthermore, the degree of similarity between the free-operant approach used here and that typically used in rodents makes it possible to build bridges between these two literatures and establish the degree of homology between the brain systems mediating instrumental learning in rodents and humans.

Our results suggest distinct contributions for different parts of prefrontal cortex and striatum in implementing goal-directed behavior. Whereas mOFC and dorsomedial striatum were more engaged by situations with a high compared with a low contingency, suggestive of a role for these regions in mediating control of behavior by the goal-directed system, the mPFC, was also found to be sensitive to changes in local contingency between responding and reward delivery, suggesting that this region may play a direct role in the on-line computation of contingency. These findings raise the interesting possibility that the corticostriatal circuitry involved in computing the causal efficacy of actions may be anatomically distinct from those circuits involved in using that knowledge to select and implement a course of action. The fact that mPFC contained representations of on-line causality, whereas dorsomedial striatum did not contain these representations but nevertheless was modulated by contingency, suggests that in this case, signals in mPFC might be used to guide activity in its dorsomedial striatal target area. Similarly, an interaction has been described previously between prefrontal cortex and dorsomedial striatum in a rather different task context, albeit

running in the converse direction to that proposed here (Pasupathy and Miller, 2005).

A number of previous studies have reported a role for dorsal striatum in processes related to contingency learning in humans. Delgado et al. (2005) used a trial-based approach to changes in neural activity over time while subjects learned instrumental associations. Activity in caudate at the time of choice was found to be present during initial learning of contingencies, but decreased over time as subjects learned the contingent relationship between responses and outcomes. Tricomi et al. (2004) reported an increase in activity in this area while subjects perceived an instrumental contingency compared with when no such contingency was perceived, even though subjects were in actuality always in a noncontingent situation. The present study demonstrates that caudate is directly modulated as a function of the degree of objective contingency, that is, in situations in which contingency is high, activity in this region is increased, compared with situations in which contingency is low.

Another important feature of our data is that we found both commonalities and differences in the brain systems exhibiting sensitivity to objective contingency and those responding to subjective causality judgments. Although the same region of medial prefrontal cortex was found to respond to both, areas such as dorsolateral prefrontal cortex and lateral orbitofrontal cortex that were found to be active in relation to subjective causality judgments did not show significant objective contingency effects, whereas dorsal striatum and medial orbitofrontal cortex found in the objective contingency contrast did not show up in the subjective causality contrast. The differences in the areas engaged in these two contrasts may relate to the fact that although subjective contingency is significantly correlated with objective contingency behaviorally, the correlation is by no means perfect, and thus the differences in the results obtained may highlight differences in the network of brain regions responsible for evaluating subjective awareness of causality from those involved in computing objective contingencies. These findings suggest that the brain systems involved in mediating subjective awareness of contingencies may be at least partly dissociable from brain systems involved in using knowledge of those contingencies to guide behavior.

Another notable feature of our data is the overall decrease in activation in the RESPOND phase compared with the REST phase in mOFC and mPFC (but not in striatum). This effect might relate to the suggestion that ventral mPFC is part of a network of brain regions that increase in activation when subjects are at rest, the so called “default” network (Gusnard et al., 2001). However, although this effect may account for the overall differences in activation between RESPOND and REST periods in these regions, it is unlikely that differences observed in these areas as a function of contingency within the RESPOND period across sessions could also be explained by this phenomenon: no significant differences were found in overall response rates or responses per reinforcer in high- compared with low-contingency conditions, suggesting that the degree of task-related effort exerted is equivalent across these conditions.

Although neural responses in a number of brain regions, including orbitofrontal cortex, but in addition amygdala and ventral striatum, have previously been found to be related to expected future reward in relation to the presentation of particular cues or stimuli, these studies are likely to be probing brain systems involved in stimulus–outcome learning, in which associations between a given context and the reward presented in that context are learned, regardless of whether an action is performed or not and, even if an action is performed, whether or not that

action is contingent on reward delivery (Schoenbaum et al., 1998; Gottfried et al., 2002, 2003; Paton et al., 2006). Such stimulus–outcome processes may be always present during instrumental conditioning alongside action–outcome and stimulus–response learning components. However, the results of the present study are unlikely to be attributable to encoding of stimulus–outcome relationships; no discriminative stimuli were used to signal whether or not an outcome would be delivered at any given point in time, other than the performance of the actions themselves. Although in principle the interval between rewards could act as a form of temporal cue to reward delivery, the fact that no significant difference was found in the mean intervals between rewards in the high- and low-contingency conditions helps to rule out that explanation for the difference in activation observed between these two conditions.

Habitual or stimulus–response learning processes are also known to be engaged during instrumental conditioning (Dickinson and Balleine, 1993). However, when behavior is under control of the habitual system, rats become insensitive to changes in contingency between actions and outcome, such that responding persists on an action even if the outcome is no longer contingent on that action (Balleine and Dickinson, 1998). Thus, the areas identified in the present study most likely pertain to associative learning processes related to the encoding of action–outcome and not stimulus–response associations. This possibility is also supported by previous studies implicating neurons in these areas in discriminating between different action–outcome associations (Matsumoto et al., 2003; Schultz et al., 2003), exhibiting sensitivity to reinforcer devaluation during reward-based action selection (Valentin et al., 2007), and showing increased activity during the perception of a response–reward contingency compared with when no contingency is perceived (Tricomi et al., 2004).

To conclude, the present results highlight the brain systems involved in the adaptive control of behavior in humans. Activity in a network of brain regions including medial prefrontal cortex, medial orbitofrontal cortex, and dorsomedial striatum was found to track changes in objective contingency. These findings in humans show remarkable parallels to previous results implicating medial frontal and dorsomedial striatum in mediating similar functions in the rodent brain (Balleine and Dickinson, 1998; Killcross and Coutureau, 2003; Yin et al., 2005; Balleine et al., 2008). Indeed, this similarity between species appears to lead to the important conclusion that the brain systems involved in controlling goal-directed action selection are heavily conserved across mammalian species.

## References

- Balleine BW (2005) Neural bases of food seeking: affect, arousal and reward in corticostriatal limbic circuits. *Physiol Behav* 86:717–730.
- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Balleine BW, Dickinson A (2000) The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. *J Neurosci* 20:8954–8964.
- Balleine BW, Ostlund SB (2007) Still at the choice point: action selection and initiation in instrumental conditioning. *Ann NY Acad Sci* 1104:147–171.
- Balleine BW, Daw ND, O’Doherty J (2008) Multiple forms of value learning and the function of dopamine. In: *Neuroeconomics: decision making and the brain* (Glimcher P, Camerer C, Fehr E, Poldrack R, eds). New York: Academic, in press.
- Baum WM (1973) The correlation-based law of effect. *J Exp Anal Behav* 20:137–153.

- Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15.
- Beckers T, De Houwer J, Matute H, eds (2007) *Human contingency learning: recent trends in research and theory*. London: Psychology.
- Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* 146:145–157.
- Dawson GR, Dickinson A (1990) Performance on ratio and interval schedules with matched reinforcement rates. *Q J Exp Psychol B* 42:225–239.
- Delgado MR, Miller MM, Inati S, Phelps EA (2005) An fMRI study of reward-related probability learning. *Neuroimage* 24:862–873.
- Dickinson A (1994) Instrumental conditioning. In: *Animal cognition and learning* (Mackintosh NJ, ed), pp 4–79. London: Academic.
- Dickinson A, Balleine BW (1993) Actions and responses: the dual psychology of behaviour. In: *Spatial representation* (Eilan N, McCarthy R, Brewer MW, eds), pp 277–293. Oxford: Basil Blackwell.
- Dickinson A, Balleine BW (1994) Motivational control of goal-directed action. *Anim Learn Behav* 22:1–18.
- Dickinson A, Nicholas DJ, Adams CD (1983) The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol B* 35:35–51.
- Elliott R, Baker SC, Rogers RD, O'Leary DA, Paykel ES, Frith CD, Dolan RJ, Sahakian BJ (1997) Prefrontal dysfunction in depressed patients performing a complex planning task: a study using positron emission tomography. *Psychol Med* 27:931–942.
- Gottfried JA, O'Doherty J, Dolan RJ (2002) Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J Neurosci* 22:10829–10837.
- Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301:1104–1107.
- Gusnard DA, Raichle ME, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694.
- Hammond LJ (1980) The effect of contingency upon the appetitive conditioning of free-operant behavior. *J Exp Anal Behav* 34:297–304.
- Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26:8360–8367.
- Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13:400–408.
- Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D (2003) A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage* 18:263–272.
- Matsumoto K, Suzuki W, Tanaka K (2003) Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* 301:229–232.
- Milner B (1982) Some cognitive effects of frontal-lobe lesions in man. *Philos Trans R Soc Lond B Biol Sci* 298:211–226.
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102.
- O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. *Neuron* 33:815–826.
- Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433:873–876.
- Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439:865–870.
- Rolls ET, Hornak J, Wade D, McGrath J (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry* 57:1518–1524.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159.
- Schultz W, Tremblay L, Hollerman JR (2003) Changes in behavior-related neuronal activity in the striatum during learning. *Trends Neurosci* 26:321–328.
- Tricomi EM, Delgado MR, Fiez JA (2004) Modulation of caudate activity by action contingency. *Neuron* 41:281–292.
- 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.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* 22:513–523.