Brief Communications

Double Dissociation of Stimulus-Value and Action-Value Learning in Humans with Orbitofrontal or Anterior Cingulate Cortex Damage

Nathalie Camille, Ami Tsuchida, and Lesley K. Fellows
Department of Neurology & Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada

Adaptive decision making involves selecting the most valuable option, typically by taking an action. Such choices require value comparisons, but there is debate about whether these comparisons occur at the level of stimuli (goods-based) value, action-based value, or both. One view is that value processes occur in series, with stimulus value informing action value. However, lesion work in nonhuman primates suggests that these two kinds of choice are dissociable. Here, we examined action-value and stimulus-value learning in humans with focal frontal lobe damage. Orbitofrontal damage disrupted the ability to sustain the correct choice of stimulus, but not of action, after positive feedback, while damage centered on dorsal anterior cingulate cortex led to the opposite deficit. These findings argue that there are distinct, domain-specific mechanisms by which outcome value is applied to guide subsequent decisions, depending on whether the choice is between stimuli or between actions.

Introduction

Recent neuroscience models of economic decision making propose processes for valuing "goods," in turn informing selection of the actions needed to obtain the desired option (Padoa-Schioppa, 2011). Converging evidence argues that orbitofrontal cortex (OFC) plays an important role in linking stimuli (i.e., goods) to their subjective, relative values (Kable and Glimcher, 2007; Peters and Büchel, 2010). In contrast, the neural substrates of action value have yet to be definitively established. Dorsal anterior cingulate cortex (dACC) seems to encode information that could support the selection of action based on value (Gläscher et al., 2009; Hayden and Platt, 2009; Wunderlich et al., 2009), but whether this region is necessary for such choices has not been tested in humans. Further, it is not clear whether decisions proceed stepwise, with choice resting on the transformation of stimulus value to action. Indeed, fMRI evidence argues that action value representations in dACC are not needed for stimulus-based decision making (Wunderlich et al., 2010). An alternative view is that value influences choices between stimuli and between actions through distinct, perhaps parallel, mechanisms. This possibility is supported by work on striatal value representations (e.g., Lau and Glimcher, 2008), and by lesion studies showing that stimulus-value and action-value learning can be dissociated in rats and nonhuman primates (Kennerley et al., 2006; Ostlund and Balleine, 2007; Rudebeck et al., 2008).

Reinforcement learning paradigms have proven useful in defining the neural processes supporting optimal choices under dynamic conditions. There is increasing evidence that regions within the frontal lobes are involved in adaptive decision making in such contexts. Most of this work has focused on OFC, which has been shown to play a necessary role in stimulus-value learning, particularly when reinforcement contingencies are changing, as occurs in reversal learning or when feedback is probabilistic (Fellows, 2007; Murray et al., 2007; Walton et al., 2010). Work in monkeys argues that dACC plays an analogous role in learning the value of actions (Kennerley et al., 2006; Rudebeck et al., 2008; Hayden and Platt, 2010).

There are few data in direct support of the claim that dACC is involved in action-value learning in humans (Gläscher et al., 2009; Wunderlich et al., 2009), although this region has been implicated in potentially related processes, including error monitoring, response inhibition, and self-generated action selection (Gehring and Willoughby, 2002; Rushworth et al., 2007b; Mansouri et al., 2009). Given that OFC neurons have also been shown to encode action-value information in animal models (Tsujimoto et al., 2009; van Wingerden et al., 2010), that dACC neuronal activity can reflect stimulus value (Amiez et al., 2006), and that both OFC and dACC activations have been reported in action-value tasks in humans (Gläscher et al., 2009), there is a need to test whether these regions play specific or general roles in stimulus-based and action-based choice.

Here, we examined these two forms of learning in humans with focal frontal lobe damage. Trial-by-trial analysis showed that OFC damage disrupted the ability to sustain the correct choice of stimulus, but not of action, after positive feedback, while damage centered on dACC led to the opposite deficit.

Received June 22, 2011; revised Aug. 24, 2011; accepted Sept. 4, 2011.

Author contributions: N.C., A.T., and L.K.F. designed research; N.C., A.T., and L.K.F. performed research; A.T. contributed unpublished reagents/analytic tools; N.C., A.T., and L.K.F. analyzed data; N.C., A.T., and L.K.F. wrote the paper.

This work was supported by operating grants from the Canadian Institutes of Health Research (CIHR; MOP 77583, 97821), a CIHR Clinician-Scientist award to L.K.F., a CIHR postdoctoral fellowship to N.C., and a Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Award to A.T.

Correspondence should be addressed to Lesley K. Fellows, Department of Neurology & Neurosurgery, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, Quebec H3A 2B4, Canada. E-mail: lesley.fellows@mcgill.ca.

DOI:10.1523/JNEUROSCI.3164-11.2011

Copyright © 2011 the authors 0270-6474/11/3115048-05$15.00/0

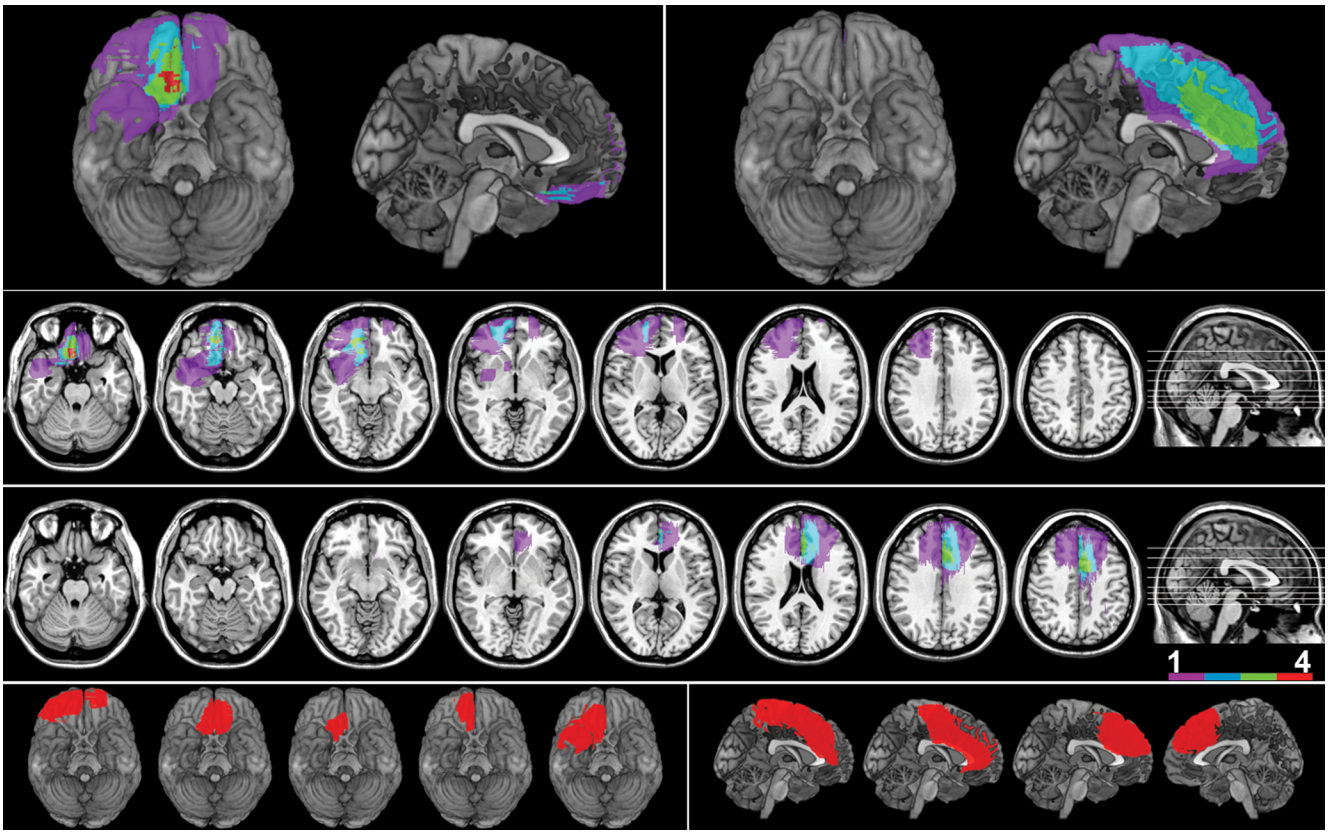

Figure 1. Lesion overlap images superimposed on 3D reconstructions of the MNI brain (top left panel OFC group, top right panel dACC group) and on the same axial slices (middle panel upper row, OFC group; lower row, dACC group). Colors show degree of overlap, as indicated in the legend. The bottom panel shows a representative image for each individual subject, with those in the OFC group on the left, and the dACC group on the right.

Materials and Methods

Patients with fixed focal lesions centered on either OFC ($N = 5$) or dACC ($N = 4$) were compared to 17 healthy participants (CTL). Participants were recruited through the McGill Cognitive Neuroscience Research Registry based solely on the location and extent of brain injury. All had fixed, circumscribed lesions of at least 6 months duration. Lesions were due to ischemic stroke, tumor resection, or aneurysm rupture. Lesions were traced from the most recently available MR or CT imaging onto the standard MNI brain by a neurologist experienced in this procedure, and blind to performance.

Age- and education-matched healthy participants were recruited by advertisement in the Montreal area. This control group was free from neurological or psychiatric illness that might be expected to affect cognition and scored >26 on the Montreal Cognitive Assessment and <12 on the Beck Depression Inventory. The sample as a whole included 10 women and 16 men. All participants provided written, informed consent according to the principles outlined in the Declaration of Helsinki. The local research ethics board approved the study. All participants were compensated for their time and inconvenience. All subjects participated in a larger study of stimulus-value learning that has been reported previously (Tsuchida et al., 2010).

Participants performed two computerized value-driven learning tasks, on two separate occasions. The reinforcement structure of the tasks was identical: one choice was associated with a win 6 of every 7 trials (and so a loss 1 trial in 7), while the other was associated with the opposite contingencies: a loss in 6/7 trials, and a win in 1/7 trials. Once subjects had selected the currently better option 13 of any 14 trials, the contingencies were reversed. The tasks continued for 70 trials after the initial reversal, with further reversals occurring whenever the learning criterion was met. The stimulus-value task involved choosing between two decks of cards of different colors, presented on a computer monitor. Responses were self-paced. Each choice resulted in a win or a loss of $50 of play money shown on the screen for 2 s, followed by a 500 ms intertrial interval. The position of the decks was randomly assigned to the right or left of the screen on each trial. A running total score was displayed at the top of the screen, and updated after each trial. The task is described in more detail by Tsuchida et al. (2010).

The action-value task had the same instructions and format, except that participants were advised that they had to choose between two possible movements, supination or pronation of the wrist (i.e., an internal or external twisting motion in the same plane). Care was taken in the instructions to refer to the choice as between wrist movements, demonstrated in a brief on-screen video clip, rather than to spatial or verbal labels for those movements. A "go" cue presented centrally on the computer monitor signaled the start of each trial, and responses were self-paced and followed by win or loss feedback displayed for 2 s, exactly as in the stimulus-value task. There was a 500 ms intertrial interval. Six practice trials preceded the main task to ensure participants understood the required movements. Participants completed stimulus-value and action-value tasks on separate days, at least 2 months apart, in both cases as part of larger, unrelated task batteries. The action-value task was performed second. An additional 8 healthy subjects were tested on the action-value task alone. Action-value performance in this group was indistinguishable from the control group that completed both tasks ($p > 0.2$ for all measures), ruling out any substantial transfer of learning across tasks under these conditions.

Results

Lesion overlap images are shown in Figure 1. Demographic and clinical information are reported in Table 1. All groups were of similar age, education, and estimated IQ.

Compared to healthy subjects, patients with OFC damage made significantly more errors while learning the initial associa-

Table 1. Demographic information for all three groups [mean (SD)]

Group	*N*	Sex ratio (F/M)	Age (years)	Education (years)	MoCa	BDI	ANART IQ estimate	Sentence comprehension accuracy	Verbal fluency		Digit span	
									FAS	Animals	Forward	Backward
CTL	17	6:11	61.2 (10.6)	16.1 (3.3)	27.7 (1.6)	4.7 (5.1)	124.6 (6.4)	0.96 (0.1)	12.5 (5.5)	21.4 (3.8)	7.0 (1.3)	5.2 (1.6)
dACC	4	4:0	64.0 (13.7)	16.5 (1.3)	26 (0.0)	8.5 (8.9)	118.7 (2.9)	1.00 (0.0)	11.0 (7.1)	14.8* (4.9)	5.7 (1.1)	4.3 (0.5)
OFC	5	0:5	58.2 (16.1)	14.8 (3.1)	28 (1.6)	10.6 (7.8)	123.1 (13.0)	0.98 (0.04)	11.8 (2.6)	22.6 (3.3)	6.7 (1.5)	4.6 (1.52)

*Values that differ significantly between the two groups, *t* test, $p < 0.05$. F, Female; M, male; MoCA, Montreal Cognitive Assessment; BDI, Beck Depression Inventory; ANART, American National Adult Reading Test.

Table 2. Global task measures by group and task [mean (SEM)]

Group	Number of reversals		Errors					
			Initial Learning		Reversal		Total errors	
	Action	Stimulus	Action	Stimulus	Action	Stimulus	Action	Stimulus
CTL	3.94 (0.25)	4.41 (0.19)	3.76 (0.46)	1.76 (0.54)	5.33 (0.75)	3.54 (0.85)	22.12 (1.36)	15.18 (2.00)
dACC	3.00 (1.22)	4.00 (0.71)	7.00 (3.34)*	3.25 (0.75)	5.13 (1.94)	5.21 (2.31)	28.25 (7.93)	19.25 (3.17)
OFC	4.20 (0.20)	3.00 (0.84)*	3.80 (1.24)	7.80 (2.03)*	4.78 (0.75)	4.13 (0.91)	23.40 (3.67)	20.75 (2.72)

Choices from the currently less adaptive option were considered errors, regardless of the nature of the feedback that followed. Errors in reversal are reported as a mean per reversal. *Significantly worse performance of the patient group compared to the control group, one-tailed unpaired *t* test, $p < 0.05$.

tions (unpaired one-tailed *t* test, $t = -4.2$, $p < 0.001$) and completed significantly fewer reversals ($t = 2.5$, $p < 0.05$) in the stimulus-value task, but not the action-value task (errors in initial learning $t = -0.03$, $p > 0.1$; reversals, $t = -0.5$, $p > 0.1$). In contrast, the dACC group made significantly more errors while learning the initial action-value associations ($t = -1.8$, $p < 0.05$), but did not differ from healthy subjects in learning the initial associations in the stimulus-value task ($t = -1.2$, $p > 0.1$). (Table 2).

Prior work in monkeys found that the most profound effect of circumscribed dACC lesions was on the influence of feedback on trial-by-trial performance (Kennerley et al., 2006; Rudebeck et al., 2008). We undertook a similar analysis, examining the effect of a win or a loss on the choice made on the next trial. Lesions involving dACC had striking effects on action-value learning after wins at this trial-by-trial level: a mixed ANOVA with group as between-subjects factor and the proportion of win–shift trials (out of all post-win trials) as within-subject factor demonstrated a significant task × group interaction ($F_{(2,23)} = 6.06$, $p < 0.01$). The dACC group was significantly more likely to shift to the other option after a previously winning choice (i.e., win–shift) than were healthy controls in the action-value task [Tukey's honestly significant difference (HSD) *post hoc* test, $p = 0.019$]. In contrast, the same patients were no more likely to shift after a win in the stimulus-value task than were healthy subjects ($p = 0.59$).

Patients with OFC lesions showed the opposite pattern: they were significantly more likely than healthy controls to shift away from a particular choice after a win in the stimulus-value task (Tukey's HSD *post hoc* test, $p = 0.017$), but did so no more often than controls in the action-value task ($p = 0.92$). These group effects were clearly evident in most individual subjects: Three of four in the dACC group had more frequent action-value win–shift responses than were observed in 95% of controls, while none in the OFC group exceeded this threshold. Four of five OFC patients made more win–shift choices in the stimulus-value task than were made by 95% of controls, while only one dACC subject met this criterion. These results are summarized in Figure 2.

There were no significant differences in choice behavior after losses in either task, for either lesion group. A mixed ANOVA with group as between-subjects factor and frequency of lose–shift behavior as within-subjects factor did not reveal any effect of group ($F_{(2,23)} = 0.34$, $p = 0.71$). Direct contrasts of patient groups with control performance after losses also showed no significant effects: unpaired one-tailed *t* tests, OFC group versus CTL $t_{(20)} = -1.1$, $p > 0.1$ for the stimulus-value task and $t_{(20)} = 0.38$, $p > 0.1$ for the action-value task; ACC group versus CTL, $t_{(19)} = 0.67$, $p > 0.1$ for the stimulus-value task and $t_{(19)} = 0.01$, $p > 0.1$ for the action-value task.

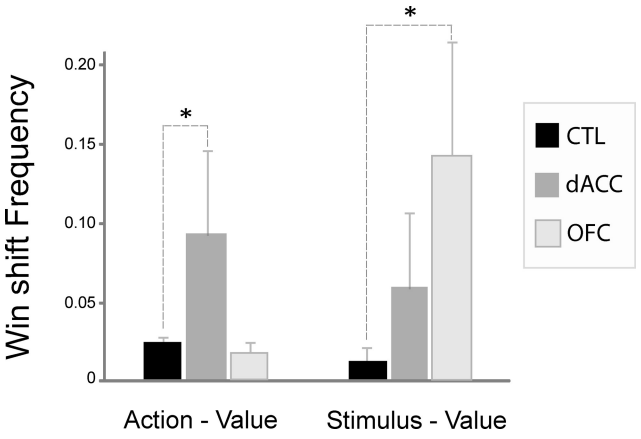

Figure 2. Mean proportion of win–shift behavior by task and group. The frequency of win–shift behavior was calculated by counting the number of response shifts immediately after congruent positive feedback and dividing it by the total number of trials following congruent positive feedback. Error bars indicate SEM. *$p < 0.05$.

Discussion

These findings extend recent work in macaques with focal lesions, by demonstrating dissociable patterns of impairment in the effects of feedback on subsequent choice in humans after discrete frontal lobe damage. We show that OFC plays a critical role in value-based choices between stimuli, but not between actions, whereas dACC seems to play an analogous role in choices between actions, but not stimuli. These patterns were particularly striking when performance was examined on a trial-by-trial basis, with dACC damage disrupting the ability to make an adaptive choice between actions (but not stimuli) following a win on the previous trial, and OFC damage similarly disrupting choices between stimuli, but not actions.

This clear demonstration of dissociability is important in constraining the potential roles of these two regions in models of learning and decision making. In particular, these findings argue against a model of decision making in which stimulus value represented in OFC is necessary for generating the representation of

action value in dACC, or in which dACC plays a very general "critic" role in learning about both stimuli and actions (Holroyd and Krigolson, 2007; Cohen and Frank, 2009).

While these regions may interact, our findings indicate that there is no necessary link between these two forms of adaptive, feedback-driven choice. These results suggest that decisions can be taken in either stimulus-based or action-based decision space, depending on the context, and that these decisions engage different neural substrates. These findings argue for material-specific contributions of dACC to action selection, and OFC to stimulus selection, based on value (Rushworth et al., 2007a; Buckley et al., 2009; Hayden and Platt, 2010; Wunderlich et al., 2010).

While we took care to recruit subjects with lesions selectively affecting ventromedial and dorsomedial frontal lobes, the resulting small samples mean that the anatomical specificity of the conclusions must be treated with caution. Our own work in a substantially larger sample allows us to conclude that OFC, and not other prefrontal regions (notably including the entire dorsomedial frontal lobe), has a critical role in stimulus-value learning (Tsuchida et al., 2010). This claim is further bolstered by extensive converging evidence (Murray et al., 2007; Rangel et al., 2008). While the lesions in the group that we refer to here as "dACC" spared ventromedial prefrontal cortex, further work will be needed to be sure about the critical area for action-value learning within the human dorsomedial PFC. Damage extended into the pre-supplementary motor and supplementary motor areas in some individuals in this group, and it is likely that white matter connections between dACC and adjacent regions were disrupted even in those subjects with more focal cortical damage. That said, existing evidence from fMRI studies using action-value tasks similar to the one used here, from single-unit recordings in humans undergoing cingulotomy, and from selective dACC lesion studies in monkeys, suggest that dACC itself may be the crucial area for action-value learning within this dorsomedial region (Williams et al., 2004; Kennerley et al., 2006; Rudebeck et al., 2008; Gläscher et al., 2009; Wunderlich et al., 2009).

We have previously demonstrated changes in trial-by-trial performance after losses, as well as after wins, in the same stimulus-value task in a larger sample with OFC damage (Tsuchida et al., 2010). In the present sample, no effects after losses were detectable. This could be a valence-specific effect, but is more likely due to the smaller sample size coupled with the particular adaptive advantage of win–stay choices in these tasks. Healthy participants vary in their behavior after losses in both tasks, but only rarely shift after wins, making the anomalous win–shift behavior in the patient groups readily detectable even in small samples.

It is worth noting that the effects of either dACC or OFC damage on value-based learning are relatively mild. Subjects with frontal injury do learn the tasks, but consistently make more errors than do healthy subjects, particularly early in learning, and show maladaptive trial-by-trial strategies, notably shifting to the other option on trials that immediately follow a win. These difficulties are remarkably similar to those seen in macaques with highly selective lesions (Rudebeck et al., 2008). Clearly these prefrontal regions are not fundamental for learning, whether about stimuli or actions, both of which likely rely principally on the basal ganglia, but rather contribute "fine tuning" to produce optimal behavior under these dynamic feedback conditions. In both action- and stimulus-value cases, it seems likely that these cortical regions are engaged in integrating feedback over several trials to predict the value of the choice made on a given trial, allowing a more sophisticated, context-specific interpretation of that outcome beyond a simple "good" or "bad." This may be important for linking outcome value to a particular stimulus in a stream of choices performed over time, a role proposed for OFC (Walton et al., 2010).

dACC might play a similar role for action value. It is tempting to speculate that predicting the specific value of a response may have something in common with anticipating errors in more conventional cognitive tasks, linking the error processing literature (Taylor et al., 2007) to the current findings. In both cases, responses are adjusted based on their anticipated outcomes. We have previously demonstrated that dACC damage selectively disrupts this anticipatory stage of error processing (Modirrousta and Fellows, 2008), and other work has shown parallels between error- and value-related processes in the dorsomedial frontal lobe (Gehring and Willoughby, 2002).

The dorsomedial frontal lobe has been implicated in voluntary (i.e., self-generated) action more generally, or in conditional action selection (as in action sequence learning) specifically (Nachev et al., 2008). Again, there are points of contact between these ideas and the current study. The action-value task requires a self-generated action, in the sense that there is no external specific target, but only a general "go" cue. Adaptive choice in this task can be seen as inferring an action selection rule based on a complex feedback history. Clearly, further work will be needed to fully understand the overlap between these various constructs. The present study argues that a neuroeconomic framework is likely to be fruitful for probing the functions of response-related prefrontal regions such as dACC, and represents a step toward understanding how learning and choice mechanisms may differ for stimuli and for actions.

References

Amiez C, Joseph JP, Procyk E (2006) Reward encoding in the monkey anterior cingulate cortex. Cereb Cortex 16:1040–1055.

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, Tanaka K (2009) Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science 325:52–58.

Cohen MX, Frank MJ (2009) Neurocomputational models of basal ganglia function in learning, memory and choice. Behav Brain Res 199:141–156.

Fellows LK (2007) The role of orbitofrontal cortex in decision making: a component process account. Ann N Y Acad Sci 1121:421–430.

Gehring WJ, Willoughby AR (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295:2279–2282.

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.

Hayden BY, Platt ML (2009) Gambling for Gatorade: risk-sensitive decision making for fluid rewards in humans. Anim Cogn 12:201–207.

Hayden BY, Platt ML (2010) Neurons in anterior cingulate cortex multiplex information about reward and action. J Neurosci 30:3339–3346.

Holroyd CB, Krigolson OE (2007) Reward prediction error signals associated with a modified time estimation task. Psychophysiology 44:913–917.

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633.

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. Nat Neurosci 9:940–947.

Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior. Neuron 58:451–463.

Mansouri FA, Tanaka K, Buckley MJ (2009) Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. Nat Rev Neurosci 10:141–152.

Modirrousta M, Fellows LK (2008) Dorsal medial prefrontal cortex plays a necessary role in rapid error prediction in humans. J Neurosci 28:14000–14005.

Murray EA, O'Doherty JP, Schoenbaum G (2007) What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. J Neurosci 27:8166–8169.

Nachev P, Kennard C, Husain M (2008) Functional role of the supplementary and pre-supplementary motor areas. Nat Rev Neurosci 9:856–869.

Ostlund SB, Balleine BW (2007) Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J Neurosci 27:4819–4825.

Padoa-Schioppa C (2011) Neurobiology of economic choice: a good-based model. Annu Rev Neurosci 34:333–359.

Peters J, Büchel C (2010) Neural representations of subjective reward value. Behav Brain Res 213:135–141.

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

Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. J Neurosci 28:13775–13785.

Rushworth MF, Behrens TE, Rudebeck PH, Walton ME (2007a) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn Sci 11:168–176.

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

Taylor SF, Stern ER, Gehring WJ (2007) Neural systems for error monitoring: recent findings and theoretical perspectives. Neuroscientist 13:160–172.

Tsuchida A, Doll BB, Fellows LK (2010) Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. J Neurosci 30:16868–16875.

Tsujimoto S, Genovesio A, Wise SP (2009) Monkey orbitofrontal cortex encodes response choices near feedback time. J Neurosci 29:2569–2574.

van Wingerden M, Vinck M, Lankelma JV, Pennartz CM (2010) Learning-associated gamma-band phase-locking of action-outcome selective neurons in orbitofrontal cortex. J Neurosci 30:10025–10038.

Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron 65:927–939.

Williams ZM, Bush G, Rauch SL, Cosgrove GR, Eskandar EN (2004) Human anterior cingulate neurons and the integration of monetary reward with motor responses. Nat Neurosci 7:1370–1375.

Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making in the human brain. Proc Natl Acad Sci U S A 106:17199–17204.

Wunderlich K, Rangel A, O'Doherty JP (2010) Economic choices can be made using only stimulus values. Proc Natl Acad Sci U S A 107:15005–15010.