

**REPORT****Dorsal striatum is necessary for stimulus-value but not action-value learning in humans**Khoi Vo,<sup>1</sup> Robb B. Rutledge,<sup>2</sup> Anjan Chatterjee<sup>3</sup> and Joseph W. Kable<sup>1</sup>

1 Department of Psychology, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104, USA

2 Wellcome Trust Centre for Neuroimaging, UCL, London WC1N 3BG, UK, and Max Planck UCL Centre for Computational Psychiatry and Ageing Research, London WC1B 5EH, UK

3 Department of Neurology, University of Pennsylvania, 3 West Gates, 3400 Spruce St., Philadelphia, PA 19104, USA

Correspondence to: Joseph W. Kable,  
3720 Walnut St., Philadelphia,  
PA 19104, USA  
E-mail: kable@psych.upenn.edu

Several lines of evidence implicate the striatum in learning from experience on the basis of positive and negative feedback. However, the necessity of the striatum for such learning has been difficult to demonstrate in humans, because brain damage is rarely restricted to this structure. Here we test a rare individual with widespread bilateral damage restricted to the dorsal striatum. His performance was impaired and not significantly different from chance on several classic learning tasks, consistent with current theories regarding the role of the striatum. However, he also exhibited remarkably intact performance on a different subset of learning paradigms. The tasks he could perform can all be solved by learning the value of actions, while those he could not perform can only be solved by learning the value of stimuli. Although dorsal striatum is often thought to play a specific role in action-value learning, we find surprisingly that dorsal striatum is necessary for stimulus-value but not action-value learning in humans.

**Keywords:** striatum; reinforcement learning; action-value; stimulus-value

**Abbreviation:** BIC = Bayesian information criterion

**Introduction**

A wealth of research links the striatum to learning on the basis of positive and negative feedback (Poldrack and Packard, 2003; Squire, 2009). Theoretical proposals framed around reinforcement learning suggest that the dorsal and ventral striatum are used to learn the value of actions and states, respectively (Houk *et al.*, 1995; Niv, 2009), and detailed computational models describe how such quantities could be learned from specific dopaminergic inputs to the direct and indirect pathways (Cohen and Frank, 2009; Morita *et al.*, 2012). Neurophysiological studies in animal models identify action-value and value-learning signals, frequently in dorsal striatum (Samejima *et al.*, 2005; Lau and Glimcher, 2008;

Kable and Glimcher, 2009; Balleine and O'Doherty, 2010; Ding and Gold, 2010; Haber and Knutson, 2010; Cai *et al.*, 2011).

Functional imaging studies provide converging evidence in humans, finding value and feedback signals in the striatum during reward learning, category learning and decision-making (Poldrack *et al.*, 2001; Poldrack and Packard, 2003; Kable and Glimcher, 2009; Balleine and O'Doherty, 2010; Haber and Knutson, 2010; Davis *et al.*, 2012). However, these studies only demonstrate the involvement of the striatum in learning, not the necessity of the striatum for learning. The necessity of striatum for learning in humans has been inferred from studies of individuals with Parkinson's disease, who are impaired at implicit category learning and learning from positive feedback (Knowlton *et al.*,

1996; Frank *et al.*, 2004; Shohamy *et al.*, 2004; Rutledge *et al.*, 2009). However, Parkinson's disease does not affect the striatum directly but instead affects the midbrain dopamine neurons that innervate the striatum. Brain damage that is both extensive and restricted to the striatum is extremely rare. Here we test an individual (Patient XG) with a rare pattern of brain damage, which was restricted to the dorsal striatum bilaterally, to evaluate the necessity of the dorsal striatum for learning in humans.

## Materials and methods

### Participants

At the time of testing, Patient XG was a 52-year-old, right-handed male who had suffered from hypoxic ischaemic damage following cardiac arrhythmia 7 years previously. MRI revealed damage in the caudate and putamen bilaterally (Fig. 1). Ventral sections of the striatum were spared, including the nucleus accumbens, but no sparing was apparent in dorsal striatum.

One of the authors (A.C.) has followed Patient XG clinically. At initial presentation, Patient XG was cortically blind. His vision started to recover within days and was completely normal within a few months. He did not show signs of amnesia, aphasia, impairments in attention or emotion, or changes in personality. He has demonstrated prominent signs of basal ganglia dysfunction since his original injury. These signs include rigidity, marked dystonic posturing (Supplementary Fig. 1), akinesia (worse in his left arm), splaying of his left toes, hypomimia, and slow and rigid gait.

In a recent neuropsychological evaluation (Supplementary Table 1), Patient XG did not demonstrate evidence of dementia or general cognitive impairment. He demonstrated normal visual acuity, contrast sensitivity, and colour vision. He performed normally on tests of intermediate vision, spatial location, and visuospatial memory. He also performed normally on tests of executive function.

We tested 11 age- and education-matched controls with no history of neurological or psychiatric disorders [seven females; one left-handed; mean age  $\pm$  standard deviation (SD),  $56.4 \pm 4.0$  years; education,  $14 \pm 1.3$  years]. To control for Patient XG's motor impairments, all participants responded with their right hand only. Correspondingly, Patient XG's reaction times evidenced little slowing relative to controls (Supplementary Fig. 2).

### Tasks

Based on current theories of striatal function, we initially hypothesized that Patient XG would be impaired on all tasks that involve learning from positive and negative feedback. To test this hypothesis, we extensively characterized participants' performance on a battery of seven learning tasks in 8 h of testing across 4 days (total  $n = 5972$  trials, Supplementary Table 2, see Supplementary material for task details):

- (i) The Weather Prediction Task ( $n = 400$  trials) is a probabilistic classification task (Knowlton *et al.*, 1996; Shohamy *et al.*, 2004). Participants predict 'rain' or 'shine' based on combinations of four learned probabilistic cues.
- (ii) The Probabilistic Selection Task ( $n = 960$  training trials,  $n = 132$  test trials) tests learning from positive versus negative feedback (Frank *et al.*, 2004). In the training phase, participants see three different pairs of cues. One cue in each pair is more likely to be associated with positive feedback. In the testing phase,

participants choose without feedback between novel pairs involving either the cue most associated with positive feedback or the cue most associated with negative feedback.

- (iii) The Crab Game ( $n = 640$  trials) is a dynamic foraging task involving variable-ratio schedules (Rutledge *et al.*, 2009). Participants select in which of two 'traps' to find 'crabs'.
- (iv) The Fish Game is a probabilistic reversal learning task ( $n = 640$  trials). Participants select in which of two 'lakes' to 'fish'.
- (v) The Bait Game ( $n = 640$  trials) is a variant of the Fish Game in which participants learn to avoid losses rather than collect gains.
- (vi) Stimulus-value learning ( $n = 640$  trials) is a reversal learning task that can only be solved by learning stimulus values (Glascher *et al.*, 2009). Participants choose between two fractal stimuli that are probabilistically rewarded at different rates.
- (vii) Action-value learning ( $n = 640$  trials) is an analogous reversal learning task that can only be solved by learning action values (Glascher *et al.*, 2009). Participants choose between two actions of a trackball mouse that are probabilistically rewarded at different rates.

Three of these tasks can only be solved by learning stimulus values (Weather Prediction Task, Probabilistic Selection Task, Stimulus-value learning), one task can only be solved by learning action values (Action-value learning), and three tasks can be solved using either stimulus values or action values (Crab, Fish and Bait Games). For each task, we collected data in two separate sessions on two different days, with new stimuli for each session. As incentive, participants earned additional performance-based compensation for one randomly selected task each session. This investigation was approved by the Institutional Review Board at the University of Pennsylvania and all participants gave informed consent.

### Statistics

We compared performance against chance using binomial probability, Patient XG's performance against matched controls using a modified *t*-test specifically designed for case studies (Crawford and Howell, 1998), and fit reinforcement learning models to behaviour using optimization routines in MATLAB. See Supplementary material for details.

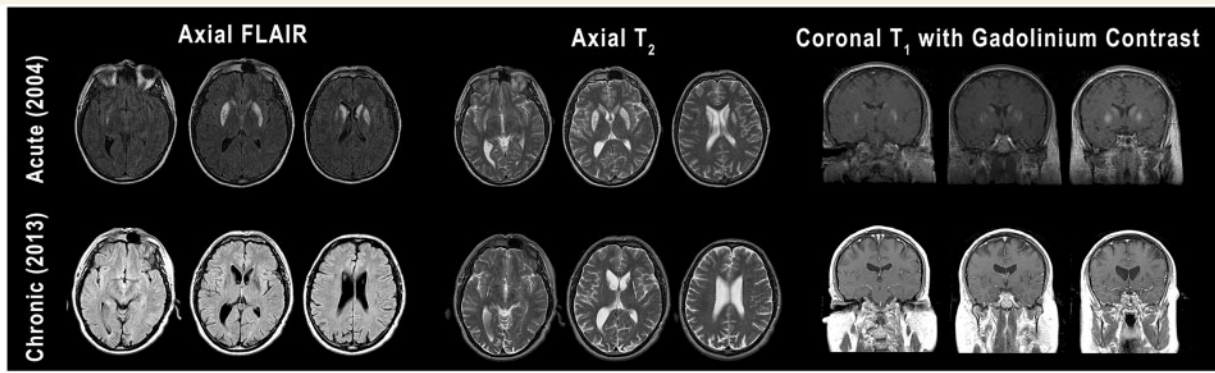
## Results

### Weather Prediction Task

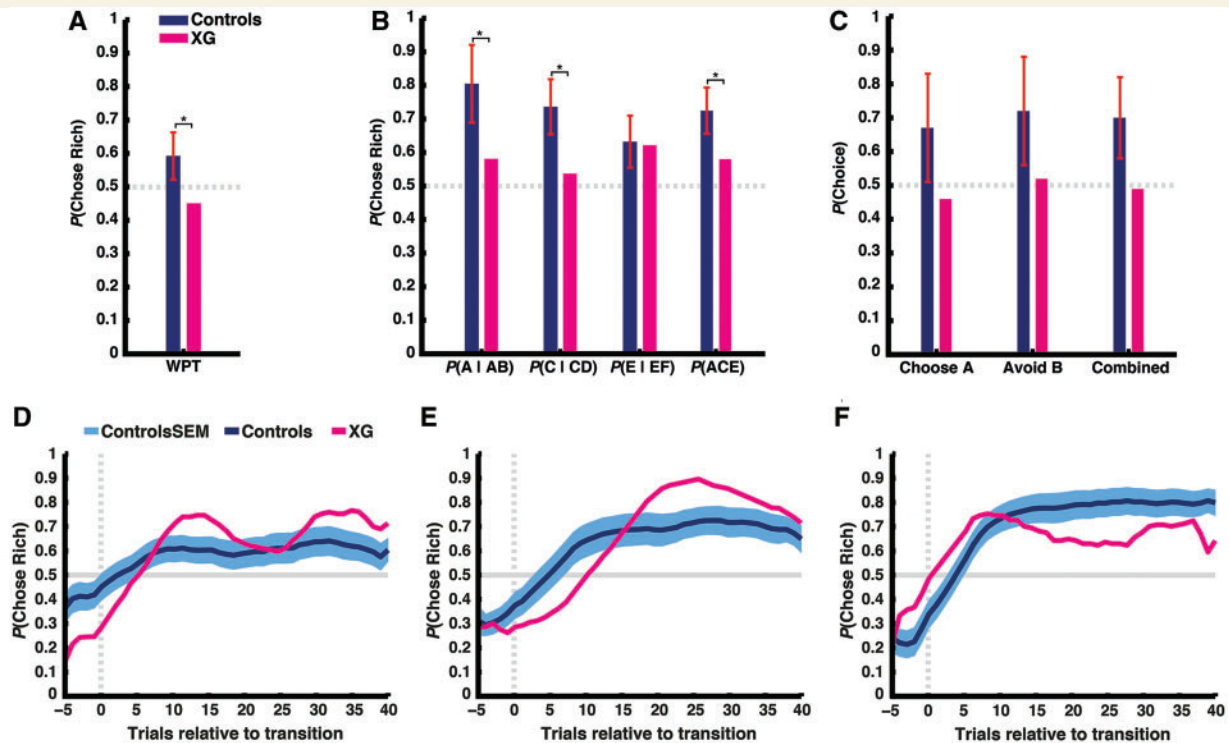
Patient XG chose the more likely option 45.0% of the time, which was slightly worse than chance performance ( $z = -1.63$ ,  $P = 0.051$ ). In contrast, controls chose the more likely option  $59.2 \pm 7.4\%$  of the time ( $z > 3.09$ ,  $P < 0.001$ ). The difference between Patient XG and controls was statistically significant [ $t(10) = -1.83$ ,  $P = 0.049$ ; Fig. 2A].

### Probabilistic Selection Task

During the training phase, which required participants to learn the values of the specific stimuli, Patient XG was able to select the option with the higher reward probability more often than chance ( $z > 3.09$ ,  $P < 0.001$ ). However, his performance was impaired relative to controls, especially for choice pairs with the most distinguishable probabilities [all trials: Patient XG = 58.0%, controls =  $72.4 \pm 7.2\%$ ,  $t(10) = -1.92$ ,  $P = 0.042$ ; AB trials, where



**Figure 1** MRI from the acute phase of Patient XG's injury (*top row*) and more recently (*bottom row*). Three different contrasts (FLAIR, T<sub>2</sub>, T<sub>1</sub> with gadolinium) are shown. Going *left to right*, axial images progress from inferior to superior and coronal images progress from posterior to anterior. In the acute images, contrast enhancement can be seen in the caudate and putamen bilaterally, indicative of recent injury (i.e. inflammation and breakdown of the blood–brain barrier). More recent images show pronounced loss of tissue in the caudate and putamen bilaterally. Both sets of images show sparing of ventral regions of the striatum, including the nucleus accumbens.



**Figure 2** (A) Patient XG and healthy controls' performance in the Weather Prediction Task. Plotted is the proportion of times participants chose the more likely option ('rain' or 'shine') given the cue combination. (B) Patient XG and healthy controls' performance in the training phase of the Probabilistic Selection Task. Plotted is the proportion of times participants chose the higher rewarded option (A, C and E) from each pair (AB, CD and EF) and across all pairs during the training phase. (C) Patient XG and healthy controls' performance in the test phase of the Probabilistic Selection Task. Plotted is the proportion of times participants choose A from novel pairs, avoid B from novel pairs, and choose the highest stimulus value across all novel pairs. In A–C, error bars denote standard deviation and asterisks denote  $P < 0.05$ . (D, E and F) Patient XG and healthy controls' performance for Crab Game, Fish Game, and Bait Game, respectively. Plotted is the probability of choosing the richer option across the course of a block (averaging over 16 total blocks per task, representing 14 total transitions). Vertical dashed line denotes block transition. Horizontal line denotes chance performance. Data smoothing kernel = 11.

A is rewarded with 80% probability: Patient XG = 58.1%, controls = 80.5 ± 12.1%,  $t(10) = -1.76$ ,  $P = 0.054$ ; CD trials, where C is rewarded with 70% probability: Patient XG = 53.8%, controls = 73.6 ± 8.6%,  $t(10) = -2.21$ ,  $P = 0.026$ ; Fig. 2B]. Critically, however, Patient XG did not solve the task the same way controls do. He was only more likely to repeat choices of the higher rewarded option when stimuli were presented in the exact same left-right configuration (Supplementary Fig. 3), a pattern consistent with action-value but not stimulus-value learning. During the test phase, which required participants to make novel choices based on learned stimulus values, Patient XG did not choose the highest reward probability option from novel pairs more often than chance (45.8%,  $z = -0.90$ ,  $P = 0.18$ ), nor did he avoid the lowest reward probability option from novel pairs more often than chance (52.1%,  $z = 0.13$ ,  $P = 0.55$ ). Although controls did choose the highest reward probability option (67.4 ± 15.6%) and avoid the lowest reward probability option (72.3 ± 15.6%) at greater than chance levels (both  $z > 3.09$ ,  $P < 0.001$ ), variability was high enough that the difference between Patient XG and controls did not reach statistical significance [all trials combined:  $t(10) = -1.70$ ,  $P = 0.06$ ; choose highest:  $t(10) = -1.33$ ,  $P = 0.11$ ; avoid lowest:  $t(10) = -1.24$ ,  $P = 0.12$ ; Fig. 2C].

## Probabilistic reversal learning tasks

In contrast to his impaired performance on the Weather Prediction and Probabilistic Selection Tasks, Patient XG performance in the Crab, Fish and Bait Games was at above-chance levels (all  $z > 3.09$ ,  $P < 0.001$ , Fig. 2D–F). For all three tasks, there was no significant difference between Patient XG and healthy control participants [Crab Game: Patient XG = 65.5%, controls = 60.0 ± 4.0%,  $t(10) = 1.30$ ,  $P = 0.89$ ; Fish Game: Patient XG = 66.1%, controls = 65.8 ± 6.3%,  $t(10) = 0.04$ ,  $P = 0.52$ ; Bait Game: Patient XG = 64.7%, controls = 73.5 ± 9.4%,  $t(10) = -0.89$ ,  $P = 0.20$ ].

## Action-based and stimulus-based reversal learning tasks

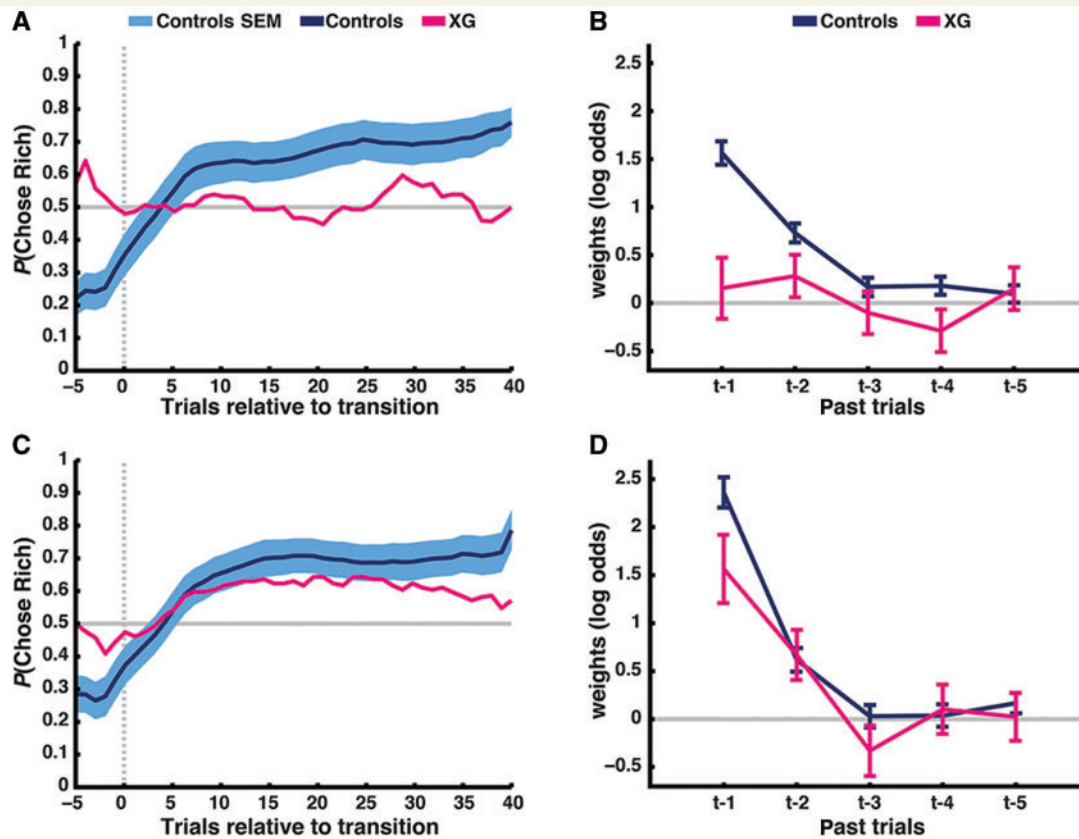
One hypothesis, which reconciles Patient XG's impaired performance on the Weather Prediction and Probabilistic Selection Tasks with his intact performance on the Crab, Fish and Bait Games, is that he can learn the value of actions but not the value of stimuli. (An alternative, that Patient XG has problems on tasks without reversals, is less consistent with the data, Supplementary Fig. 4.) This hypothesis can be evaluated directly by comparing Patient XG's performance in two tasks that can only be solved by learning stimulus values or action values, respectively. In stimulus-value learning, Patient XG's performance was not different from chance ( $z = 0.36$ ,  $P = 0.64$ , Fig. 3A) and was significantly lower than control participants [Patient XG = 51.1%, controls = 65.7 ± 6.9%,  $t(10) = -2.03$ ,  $P = 0.035$ ]. In contrast, Patient XG chose the richer action at above-chance levels in action-value learning ( $z > 3.09$ ,  $P < 0.001$ , Fig. 3C) and his performance was not significantly different from control participants

[Patient XG = 59.2%, controls = 67.4 ± 6.9%,  $t(10) = -1.14$ ,  $P = 0.14$ ].

For both tasks, we fit a regression model estimating the influence of past rewards and past choices on current choices (Fig. 3B and D). The signature of reinforcement learning in this model is an exponential decline in the influence of past rewards. While control participants exhibited this pattern in stimulus-value learning, with significant weights for rewards received in the past two trials (both  $P < 0.001$ ) and higher weights for the most recent trial than the second most recent trial ( $P < 0.001$ ), weights for Patient XG were not significantly different from zero for even the most recent trial ( $P > 0.3$ ). In contrast, both Patient XG and control participants showed this signature of reinforcement learning in action-value learning (past two weights both  $P < 0.001$  and  $t - 1 > t - 2$ ,  $P < 0.001$ ). Consistent with these regression models, a three-parameter reinforcement-learning model explained choice behaviour in the stimulus-value learning task significantly better than a model with no learning in control participants while accounting for the number of parameters in the models [likelihood ratio test, all  $\chi^2(2) > 11$ ,  $P < 0.001$ ], but not in Patient XG [likelihood ratio test,  $\chi^2(2) = 0.70$ ,  $P = 0.70$ ]. The reinforcement-learning model explained choice behaviour in the action-value learning task significantly better for both Patient XG and control participants [likelihood ratio test, all  $\chi^2(2) > 11$ ,  $P < 0.001$ ]. We also computed Bayesian Information Criterion (BIC) scores for the two models, which penalize for model complexity. The reinforcement-learning model was preferred by BIC (lower score) for all control subjects for both tasks. For Patient XG, the three-parameter reinforcement learning model was preferred for the action-value learning tasks (reinforcement learning model BIC = 753.29 versus null model BIC = 876.05) but not for the stimulus-value learning task (reinforcement learning model BIC = 913.07 versus null model BIC = 893.68). Further, Patient XG's learning rate in the three-parameter reinforcement learning model was significantly lower than controls for stimulus-value learning [Patient XG,  $\alpha = 0.06$ , controls,  $\alpha = 0.65 \pm 0.20$ ,  $t(10) = -2.28$ ,  $P < 0.01$ , Fig. 4], but within the normal range for action-value learning [Patient XG,  $\alpha = 0.98$ , controls,  $\alpha = 0.75 \pm 0.18$ ,  $t(10) = 1.29$ ,  $P = 0.11$ , Fig. 4], and the difference in learning rate between the two tasks was significantly greater in Patient XG than controls [ $t(10) = 5.19$ ,  $P < 0.001$ ].

## Discussion

In contrast to our initial hypothesis, Patient XG exhibited a surprising dissociation between stimulus-value and action-value learning. In multiple tasks that required learning the reward contingencies for different stimuli, including the Weather Prediction and Probabilistic Selection Tasks originally used to demonstrate learning deficits in Parkinson's disease (Knowlton *et al.*, 1996; Frank *et al.*, 2004; Shohamy *et al.*, 2004), Patient XG exhibited a total deficit in absolute terms, with performance indistinguishable from chance. The one exception to this pattern was the training phase of the Probabilistic Selection Task, where Patient XG's performance was at above-chance levels though still impaired relative to controls—but how he achieved this level of performance was not



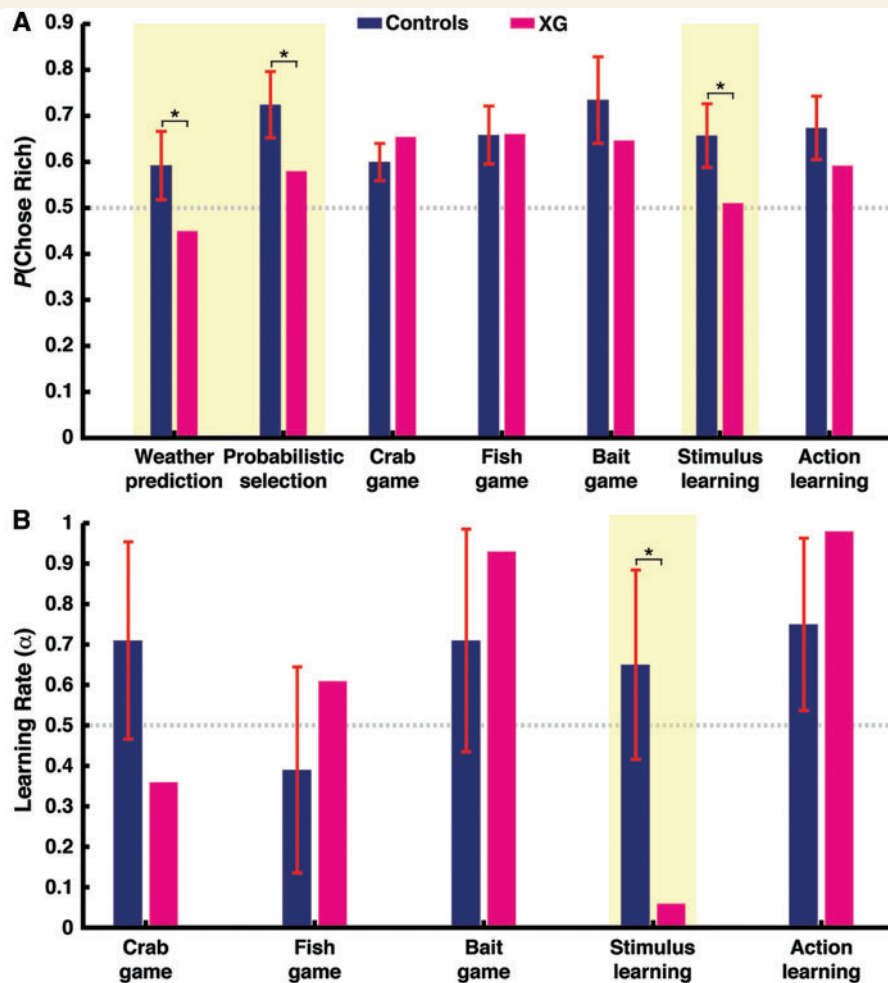
**Figure 3** (A and C) Patient XG and healthy controls' performance for the stimulus-value learning (A) and action-value learning tasks (C), respectively. Plotted is the probability of choosing the richer option across the course of a block (averaging over 16 total blocks per task, representing 14 total transitions). Vertical dashed line denotes block transition. Horizontal line denotes chance performance. Data smoothing kernel = 11. (B and D) General linear model fit showing the influences of past rewards on learning for stimulus-value learning (B) and action-value learning (D). Control fits in both tasks demonstrate the signature of reinforcement learning. However, Patient XG only shows fits consistent with reinforcement learning in action-value learning, and not in stimulus-value learning.

consistent with stimulus-value learning. In contrast to his deficit in learning stimulus values, Patient XG performed similarly to healthy controls in several tasks that could be solved by learning the reward contingencies for different actions (Glascher *et al.*, 2009; Rutledge *et al.*, 2009). Patient XG's performance was consistent with reinforcement learning in action-value learning but not stimulus-value learning, and the discrepancy in his learning rates for action values versus stimulus values was significantly greater than controls. Though we can only draw firm conclusions about Patient XG's learning in the domain of rewards/gains, his normal performance in the Bait Game suggests that action-value learning for punishments/losses is also intact.

This dissociation supports the idea that action-value and stimulus-value learning depend on distinct neural substrates (Rudebeck *et al.*, 2009; Rangel and Hare, 2010; Camille *et al.*, 2011; Padoa-Schioppa, 2011). However, that this dissociation could arise from damage concentrated in dorsal striatum is very surprising. Most models posit a role for the dorsal striatum in learning action values and the ventral striatum in learning stimulus values (Houk *et al.*, 1995; Niv, 2009). Supporting this role, neurophysiological studies have identified single neurons in dorsal striatum that encode action values or reward prediction errors that could be

used to learn action values (Samejima *et al.*, 2005; Lau and Glimcher, 2008; Kable and Glimcher, 2009; Ding and Gold, 2010; Cai *et al.*, 2011), suggesting that dorsal striatal damage would impair action-value learning and leave stimulus-value learning intact.

How can we explain this surprising result? Lesion studies have shown that ventromedial frontal regions are critical for stimulus-value learning, while dorsomedial frontal regions are critical for action-value learning (Rudebeck *et al.*, 2009; Camille *et al.*, 2011). One possible explanation for the dissociation we observe is that information from ventromedial frontal regions must pass through frontal-striatal circuits to affect behaviour, while information from dorsomedial frontal regions can affect behaviour directly through connections to motor circuits. This idea is consistent with anatomical evidence for an ascending flow of information through frontal-striatal circuits (Haber and Knutson, 2010) and for the existence of a hyperdirect pathway from dorsomedial prefrontal cortex to the subthalamic nucleus (Nambu *et al.*, 2002). Alternatively, a second possible explanation is that there is redundancy in the systems available for learning action values that is not present in the systems for learning stimulus values. Cortical pathways involving premotor cortex and posterior parietal cortex may have access to motor and learning signals necessary to solve the



**Figure 4** (A) Summary of Patient XG and healthy controls' performance across all learning tasks. (B) Summary of Patient XG and healthy controls' learning rates. Learning rates are estimated in a basic reinforcement learning model fit to behaviour. Tasks that are solvable only by learning stimulus values are highlighted in yellow. Asterisks denote  $P < 0.05$ .

learning of action values (Romo and Salinas, 2003; Kable and Glimcher, 2009). The cerebellum may also have access to the necessary signals (Doya, 1999; Medina, 2011). In either case, our data show that structures outside the dorsal striatum are sufficient to learn to discriminate between two simple actions.

Of course, a limitation of the current study is that Patient XG is only a single case, though striking dissociations in single neuropsychological cases have often advanced cognitive neuroscience (Caramazza and McCloskey, 1988). In this case, we find evidence for the necessity of the dorsal striatum for stimulus-value but not action-value learning. The striking ability of Patient XG to learn action values despite bilateral dorsal striatal damage should prompt a careful reconsideration of current theories of striatal function in reinforcement learning.

## Funding

This work was supported by National Institutes of Health grant DA029149 to Joseph W. Kable. Robb Rutledge was supported by the Max Planck Society.

## Supplementary material

Supplementary material is available at *Brain* online.

## References

- Balleine BW, O'doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 2010; 35: 48–69.
- Cai X, Kim S, Lee D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 2011; 69: 170–82.
- Camille N, Tsuchida A, Fellows LK. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate damage. *J Neurosci* 2011; 31: 15048–52.
- Caramazza A, McCloskey M. The case for single-patient studies. *Cogn Neuropsychol* 1988; 5: 517–27.
- Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 2009; 199: 141–56.
- Crawford JR, Howell DC. Comparing an individual's test score against norms derived from small samples. *Clin Neuropsychol* 1998; 12: 482–6.

- Davis T, Love BC, Preston AR. Striatal and hippocampal entropy and recognition signals in category learning: simultaneous processes revealed by model-based fMRI. *J Exp Psychol Learn Memory Cogn* 2012; 38: 821–39.
- Ding L, Gold JI. Caudate encodes multiple computations for perceptual decisions. *J Neurosci* 2010; 30: 15747–59.
- Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw* 1999; 12: 961–74.
- Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science* 2004; 306: 1940–3.
- Glascher J, Hampton AN, O'Doherty JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 2009; 19: 483–95.
- Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 2010; 35: 4–26.
- Houk JC, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG, editors. *Models of information processing in the Basal Ganglia*. Cambridge, MA: MIT Press; 1995. p. 249–70.
- Kable JW, Glimcher PW. The neurobiology of decision: consensus and controversy. *Neuron* 2009; 63: 733–745.
- Knowlton BJ, Mangels JA, Squire LR. A neostriatal habit learning system in humans. *Science* 1996; 273: 1399–402.
- Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. *Neuron* 2008; 58: 451–63.
- Medina JF. The multiple roles of Purkinje cells in sensori-motor calibration: to predict, teach and command. *Curr Opin Neurobiol* 2011; 21: 616–22.
- Morita K, Morishima M, Sakai K, Kawaguchi Y. Reinforcement learning: computing the temporal difference of values via distinct corticostriatal pathways. *Trends Neurosci* 2012; 35: 457–67.
- Nambu A, Tokuno H, Takada M. Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. *Neurosci Res* 2002; 43: 111–17.
- Niv Y. Reinforcement learning in the brain. *J Math Psychol* 2009; 53: 139–54.
- Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. *Ann Rev Neurosci* 2011; 34: 333–59.
- Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, et al. Interactive memory systems in the human brain. *Nature* 2001; 414: 546–50.
- Poldrack RA, Packard MG. Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 2003; 41: 245–51.
- Rangel A, Hare T. Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 2010; 20: 262–70.
- Romo R, Salinas E. Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews Neuroscience* 2003; 4: 203–18.
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 2009; 28: 13775–85.
- Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW. Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J Neurosci* 2009; 29: 15104–14.
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science* 2005; 310: 1337–40.
- Shohamy D, Myers CE, Onlaor S, Gluck MA. Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? *Behav Neurosci* 2004; 118: 676–86.
- Squire LR. Memory and brain systems: 1969–2009. *J Neurosci* 2009; 29: 12711–6.