

THESIS

MANY HATS: INTRA-TRIAL AND REWARD-LEVEL DEPENDENT BOLD
ACTIVITY IN THE STRIATUM AND PREMOTOR CORTEX

Submitted by

Erik J Peterson

Department of Psychology

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2012

Master's Committee:

Advisor: Carol A Seger

Lucy J Troup

Charles W Anderson

ABSTRACT

MANY HATS: INTRA-TRIAL AND REWARD-LEVEL DEPENDENT BOLD ACTIVITY IN THE STRIATUM AND PREMOTOR CORTEX

Lesion, drug, single-cell recording, as well as human fMRI studies, suggest dopaminergic projections from VTA/SNc (ventral tagmental area/substantia nigra pars compacta) and cortically driven striatal activity plays a key role in associating sensory events with rewarding actions both by facilitating reward processing and prediction (i.e. reinforcement learning) and biasing and later updating action selection. We, for the first time, isolated BOLD signal changes for stimulus, pre-response, response and feedback delivery at three reward levels. This design allowed us to estimate the degree of involvement of individual striatal regions across these trial components, the reward sensitivity of each component and allowed for a novel comparison of potential (and potentially competing) reinforcement learning computations.

Striatal and lateral premotor cortex regions of interest (ROIs) significant activations were universally observed (excepting the ventral striatum) during stimulus presentation, pre-response, response and feedback delivery, confirming these areas importance in all aspects of visuomotor learning. The head of the caudate showed a precipitous drop in activity pre-response, while in the body of the caudate showed no significant changes in activity. The putamen peaked in activity during response. Activation in the lateral premotor cortex was strongest during stimulus presentation, but

the drop off was followed by a trend of increasing activity as feedback approached. Both the head and body of the caudate as well as the putamen displayed reward-level sensitivity only during stimulus, while the ventral striatum showed reward sensitivity at both stimulus and feedback. The lack of reward sensitivity surrounding response is inconsistent with theories that the head and ventral striatum encode the value of actions. Which of the three examined reinforcement learning models correlated best with BOLD signal changes varied as a function of trial component and ROI suggesting these regions computations vary depending on task demand

TABLE OF CONTENTS

INTRODUCTION	1
METHODS	12
RESULTS	20
DISCUSSION	25
CONCLUSION	30
REFERENCES	41

1. INTRODUCTION

How agents - human or otherwise - learn to associate sensory events with arbitrary actions has been extensively studied in the context of stimulus-response (S-R) tasks, in which an image becomes associated with a small number of possible responses, generally two, based on trial-by-trial feedback. The input structures of the striatum (i.e., the ventral striatum, the head and body of the caudate and the putamen), in conjunction with the dopaminergic projections from the ventral tagmental area/substantia nigra pars compacta (VTA/SNc) into the striatum, are important in facilitating this learning process. While each of the striatal input structures themselves subserve distinct computational roles (reviewed briefly below), direct neuronal recordings in non-human animals show they also display distinct intra-trial firing patterns centered around stimulus presentation, pre-response, during and after response as well as before and during feedback delivery. These differential patterns suggest that even in the context of a single trial, striatal subregions carry out unique computations. No human work to date has analyzed these patterns in the context of whole brain fMRI imaging, nor has any work of which the authors are aware considered how these patterns are affected by reward-level. Reward-level is relevant as much previous work has implicated the striatum and its dopamine efferents as playing a key role in reward valuation and processing (M R Delgado, L E Nystrom, Fissell, Noll, & J A Fiez, 2000; John O'Doherty et al., 2004; Bray & J P O'Doherty, 2007; Rodriguez, Aron, & R.A Poldrack, 2006).

The first goal of this study was to examine how the different striatal regions and premotor cortex are recruited during stimulus, pre-response, response, feedback and how

activity is modulated by different reward values (verbal feedback, low monetary, and high monetary) during these periods . By employing a rapid 1 sec TR and limiting the imaging window we were able to isolate BOLD signal changes for each of these periods. Combined with anatomically based regions of interest (ROIs) this approach allowed for a global quantitative assessment of striatal and premotor involvement throughout visuomotor learning. The putamen peaked in activity during response. Activation in the lateral premotor cortex was strongest during stimulus presentation, but the drop off was followed by a trend of increasing activity as feedback approached. Both the head and body of the caudate as well as the putamen displayed reward-level sensitivity only during stimulus, while the ventral striatum showed reward sensitivity at both stimulus and feedback. The lack of reward sensitivity surrounding response is inconsistent with theories that the head and ventral striatum encode the value of actions.

Secondly, as discussed in detail below, the putamen plays an important role in response selection and while the head and body of the caudate are important for the more abstract category selection. However nearly all work to date (however see (Xue, Ghahremani, & Russell A Poldrack, 2008)) has treated response and category selection a singular events. As a result, we employed a task design that decoupled category selection (i.e. the response labels or names) from the response itself. Via a reinforcement learning model based analysis, for the first time this, allowed us to directly measure the sensitivity of striatal regions both to both stimulus-response and stimulus-category associations, as well as to consider a novel alternative model that responses are associated only with feedback independent of the initial stimulus. This last model is simplified theoretical interpretation of the action-value tunings in the dorsal striatum that have been previously

reported (Lau & Glimcher, 2007). Surprisingly, we observed that correlations with either of the three models was not limited to either intra-trial or specific striatal subregions. These results suggest that multiple associations are encoded in the striatum simultaneously and that striatal computations shift dynamically based on task demands.

Functional striatal neuroanatomy.

Striatal regions of interest include the ventral striatum (nucleus accumbens) the putamen (dorsolateral striatum) the caudate (dorsomedial striatum). The latter was further subdivided into anterior (head) and posterior (body) regions consistent with previous work (Seger & C. M. Cincotta, 2005; Seger, 2007; Seger, E. J. Peterson, Corinna M Cincotta, Lopez-Paniagua, & Anderson, 2010). Our premotor region of interest was identified on the basis of studies that observed bilateral (BA6) activity during stimulus-response tasks requiring finger responses (Seger & C. M. Cincotta, 2005; Seger et al., 2010; Sun, Miller, & D'Esposito, 2005). This region, we presume, should present activity similar to dorsal premotor cortex typically examined in the monkey literature (Brasted & Wise, 2004; Buch, Brasted, & Wise, 2006)

The contributions of each striatal region to S-R learning are distinct, as each is thought to belong to independent cortico-striatal loops. The motivational loop connects the orbital frontal cortex (OFC) and the anterior cingulate to the ventral striatum. Studies in both animal models (direct neuronal recordings) and in human subjects (fMRI and PET) have implicated the ventral striatum in both the processing and expectation of feedback as well as perhaps encoding the value of actions (M R Delgado, Locke, Stenger, & J A Fiez, 2003; Roesch, T. Singh, Brown, Mullins, & Schoenbaum, 2009; John P

O'Doherty, Peter Dayan, Karl Friston, Critchley, & Raymond J Dolan, 2003; Seger & C. M. Cincotta, 2005). This role is subserved through connections to the orbital frontal cortex (OFC) and anterior cingulate (ACC), which encode the absolute value of sensory information and error or conflict monitoring respectively (J. O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Seger, 2008; Quilodran, Rothé, & Procyk, 2008; Grafton, Schmitt, Horn, & Diedrichsen, 2008; Rudebeck et al., 2008).

The head of the caudate participates in executive associative loop, where it is thought to play role both feedback anticipation and processing, especially of the cognitive type (e.g. money as opposed to food rewards), given its connections to dorsolateral prefrontal cortex (Shohamy et al., 2004; Seger & C. M. Cincotta, 2005). It and may also play a role in gating working memory contents (Randall C O'Reilly, 2006; Mars & Grol, 2007) and in causal inference (Tanaka, Balleine, & John P O'Doherty, 2008). Unlike the head, the body of the caudate acts in the context of the visual loop, connecting the ventral visual path to both the visual and pre-motor cortices where it thought to play a role in both linking visual information to motor actions as well as in visual categorization and visual learning tasks (Ashby & Maddox, 2005; Seger & C. M. Cincotta, 2005).

The putamen, as part of the motor loop, interacts with premotor cortex, the SMA (supplementary motor area) and portions of the parietal cortex to select both motor plans and/or specific actions (Shohamy et al., 2004; Seger, 2008). In this circuit the pre-motor cortex acts (in another network with both frontal and parietal regions) to represent potential movements as well as to prepare motor plans (Toni, Thoenissen, & Zilles, 2001; Ueda & Kimura, 2003).

Striatal recruitment across trial phases and reward levels.

Single-unit recordings of phasically active neurons (PANs) in non-human animals as well as human neuroimaging suggest that striatal activity increases at several distinct times inside single trials. Activity changes are generally observed following stimulus presentation, pre-response, at and after response and during feedback delivery. It is thought these bursts subserved distinct (often controversial) functional roles which may be subserved by semi-independent neuronal populations (Seger, 2008; Houk et al., 2007; Barnes, Kubota, Hu, Jin, & Graybiel, 2005; Bar-Gad, Morris, & Bergman, 2003). For example, reward sensitive neurons were shown by Lau and Glimcher (2008) to be distinct from the pre and post-response neurons. While Schmitzer-Torbert *et al* (2004) reported both spatial feature and stimulus-response mappings and Xue *et al* (2008) reported stimulus-label mappings along side stimulus-response mapping, though the former elicited greater activation.

During stimulus presentation bilateral burst firing has been consistently reported across all regions of the striatum (both dorsal and ventral), as well as in dorsal pre-motor cortex, (Brasted & Wise, 2004; Hadj-Bouziane, Meunier, & Boussaoud, 2003; Buch et al., 2006) but not without exception; Buch *et al* (2006) observed no stimulus related putamen activity. The degree of activity during stimulus presentation in these same regions has been positively correlated with prior reward history in some, but not all, laboratories (Lau & Glimcher, 2008; Hollerman, Tremblay, & W. Schultz, 1998; Brasted & Wise, 2004; Buch et al., 2006). Human fMRI experiments have additionally shown that stimulus-related BOLD changes in both ventral and dorsal striatum correlate with prior reward history as well as reward valence or event salience or both (Knutson, C M

Adams, G W Fong, & D. Hommer, 2001; M R Delgado et al., 2003; Zink, Pagnoni, Martin, Dhamala, & Berns, 2003; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004; Cooper & Knutson, 2008). Activity at stimulus has also been associated with novelty or in other studies increasing behavioral performance (Wittmann, Nathaniel D Daw, Seymour, & Raymond J Dolan, 2008; Toni et al., 2001). However dorsal activity, as opposed to ventral, is contingent on the presented stimuli having behavioral relevance (Tanaka et al., 2008; John O'Doherty et al., 2004; Yin, Ostlund, & Balleine, 2008).

In the pre-response period -- after stimulus onset leading up to responding sometimes including presentation of a pre-response cue and/or a “go” signal -- activity in the dorsal caudate is tightly clustered around response, much more so than either putamen or premotor (Buch et al., 2006). Others have reported that caudate activity is correlated with the reward history of the selected action (Lau & Glimcher, 2008), but again this is not always observed (Hadj-Bouziane & Boussaoud, 2003; Brasted & Wise, 2004). Putamen activity reached its maximum value following response, and showed a much broader overall tuning than the dorsal premotor cortex which increased throughout this period as well, reaching its maximum just prior to response followed by a rapid decline. The single neuroimaging (fMRI) study of which are aware which attempted to isolated pre-response and response-related activity found no significant striatal activity, however increases in dorsal premotor cortex (pre-response) and motor cortex (at response) were observed (Toni, Schluter, Josephs, K. Friston, & Passingham, 1999). The lack of observed striatal activity was likely due to inadequate (repetition time = 6 sec). All reports show activity leading up to response is tightly correlated with correct responding with incorrect responses showing no significant, above baseline, activity (Buch et al.,

2006; Hadj-Bouziane & Boussaoud, 2003; Brasted & Wise, 2004; Lau & Glimcher, 2008). There are however conflicting reports on how activity changes with learning during this period. Hadj-Bouziane *et al* (2003), reported no significant striatal or premotor activity until behavioral responses had plateaued, unlike Lau and Glimcher (2008) who showed the opposite pattern.

In the post-response phase (defined as the time from approximately half a second after response execution to reward delivery, also known as pre-reward) dorsal putamen and caudate activity steadily increase, likely due to reward anticipation, whereas dorsal premotor cortex typically shows no activity (Buch et al., 2006; Hadj-Bouziane & Boussaoud, 2003; Brasted & Wise, 2004).

Cellular recordings of reward delivery and the subsequent period are characterized by sharp spiking activity throughout premotor cortex and the striatum, followed by an rapid decline in dorsal premotor, and a comparatively extended decline in the putamen and caudate (Buch et al., 2006; Hadj-Bouziane & Boussaoud, 2003; Brasted & Wise, 2004; Barnes et al., 2005). Human imaging experiments have confirmed both dorsal and ventral striatal response for a variety of rewards: verbal, monetary, and social (Seger & C. M. Cincotta, 2005; Rodriguez et al., 2006; Cooper & Knutson, 2008; Izuma, Saito, & Sadato, 2008; Montague, King-Casas, & J. D. Cohen, 2006). Additionally head of the caudate activity has been shown to be modulated by reward valence and magnitude (M R Delgado et al., 2003) with activity in only that striatal region increasing in the presence of feedback when compared to observational learning (Corinna M Cincotta & Seger, 2007). The bulk of studies on feedback delivery have been focused on how reward is processed, with the dominate model being the reward “prediction error

hypothesis”. Consistent with human fMRI work on reward processing (Haruno & Kawato, 2006), all reports show a (sometimes weak) correlation in firing rates with either subjective value, or the difference between subjective value and the actual reward, known as a reward prediction error (RPE).

Reinforcement learning and the striatum.

The second goal of this study is to further refine our understanding of the type and nature of the computations carried out in the striatum and lateral premotor cortex through the use of reinforcement learning models. We fit these models to the subjects behavioral data and then regressed them onto the BOLD data. Previous research has found that two reinforcement learning measures, prediction error and reward prediction, describe activity in the ventral and dorsal striatum, respectively (Seger et al., 2010). The current design allows for two extensions of previous work. First, we constructed separate models that could, for the first time, distinguish response selection from category selection. Model SO (stimulus-outcome) estimated the value (“Q”) of the two category label options associated with each of the eight stimuli. Model SR (stimulus-response) provided an estimate of the relative value of the two possible motor responses (right hand button press vs left hand button press) associated with each of the stimuli. Model RF (response-feedback) estimated the value of the two responses independent of the stimuli. Second by examining these models within the stimulus, pre-response, response and feedback periods we can gain novel insight into the dynamic computations performed within striatal and premotor regions.

As discussed in above PANs in striatum are active intermittently throughout stimulus presentation, before after and during response as well as before during and after feedback. During feedback presentation, and under some circumstances during stimulus presentation, the dopamenergic projections from the substantial nigra pars compacta (SNc) and ventral tagmental area (VTA) lead to widespread but rapid (phasic) changes in dopamine concentration in both the ventral and dorsal striatum. Schultz (1992) initially observed that these phasic firing patterns were similar to the scalar reward prediction errors (RPE) produced in models of reinforcement learning. In the intervening years many of these similarities were confirmed in both animal models and human fMRI: an expected reward that met expectation was met with no change to the base firing rate, while a greater than expected reward lead to burst of activity, while a less the expected reward lead to depression in activity (Mirenowicz & W. Schultz, 1994; Mirenowicz & Wolfram Schultz, 1996; Montague, P. Dayan, & Sejnowski, 1996). Additionally, SNc/VTA phasic activity back-propagates from reward delivery to stimulus presentation -- a key feature of the RPE in temporal difference (TD) reinforcement learning algorithms (D'Ardenne, McClure, Leigh E Nystrom, & Jonathan D Cohen, 2008; Roesch, Calu, & Schoenbaum, 2007). There is also limited data suggesting RL models are predictive of non-human animals' choice behavior (Hampton & John P O'Doherty, 2007).

The RPE hypothesis has also been extended to incorporate SNc/VTA phasic activity observed following presentation of novel stimuli (Kakade & P. Dayan, 2002; Guitart-Masip, Bunzeck, Stephan, R. J. Dolan, & Duzel, 2010; Wittmann et al., 2008) as well as to explain reward anticipatory firing via an average reward prediction error (Knutson, Charles Adams, Grace Fong, & Daniel Hommer, 2001; Knutson & Wimmer,

2007). Another variation allowed for the observation of simultaneous neural implementation of model-free and model-based reinforcement learning (Gläscher, Daw, Peter Dayan, & John P O'Doherty, 2010). Alternative but reconcilable accounts have also been offered that allow for dissociation of first and second order conditioning as well as pavlovian to instrumental transfer (R. O'Reilly, Frank, Hazy, & Watz, 2007). The RPE hypothesis has also been incorporated into theoretical accounts of motivation (Peter Dayan, 2007) and addiction (Redish, 2004). However observed phasic changes in SNc/VTA related to reward uncertainty (Christopher D Fiorillo, Tobler, & Wolfram Schultz, 2003) and optimistic firing, where the best choice not the selected choice is signaled (Roesch et al., 2007) have yet to be theoretically incorporated. The classical definition of reward also seems inadequate to describe the observed responses during goal achievement in humans (Tricomi & Julie A Fiez, 2008) and the phasic response when rats view action movies (Blatter & Wolfram Schultz, 2006).

The RPE hypothesis has been directly challenged. Tracer studies in non-human animals suggest limited direct input from visual areas prevents the VTA/SNC from rapidly receiving the detailed visual information necessary for reward processing; for a RPE to be calculated precise timing is required (Dommett et al., 2005). On the basis of this data critics have argued for a saliency interpretation of dopamine phasic activity (P. Redgrave, Gurney, & Reynolds, 2007). A saliency detection system (how relevant is current sense information relevant to behavior) it is argued requires less detailed visual information. This saliency interpretation is supported by a phasic response in SNc/VTA and the ventral striatum in response to salient but non-rewarding visual stimuli (Zink et al., 2003; Zink et al., 2004; Zink, Pagnoni, Chappelow, Martin-Skurski, & Berns, 2006).

However saliency fails to account for activity in these same regions when either positive or negative rewards are omitted, equally salient events that lead to opposite neural responses (D'Ardenne et al., 2008; John P O'Doherty, Buchanan, Seymour, & Raymond J Dolan, 2006). The concept of saliency itself requires further refinement. How to define and objectively measure the saliency of events remains unclear. Additionally, recent work has shown that while there is indeed a tight time window in which individual SNc/VTA cells fire, these windows are distributed across a temporal spectrum which may alleviate the theoretical need for a precise singular temporal response (C. Fiorillo, Newsome, & W. Schultz, 2008).

Besides the RPE the other output from a reinforcement learning model is an estimate of total future reward (denoted here as Q) for a given action given the current context (known in the RL literature as a state). Q or subjective value correlated activity has been reported in putamen, the head of the caudate, and the ventral striatum during stimulus presentation as opposed to feedback delivery when this region displays RPE like activity (Haruno & Kawato, 2006; Seger et al., 2010; Knutson, Taylor, Kaufman, R. Peterson, & Glover, 2005; Kable & Glimcher, 2007; Yacubian et al., 2006), Dorsolateral prefrontal cortex activity has also been correlated with (diminishing) value in a delayed reward task (Kim & Fukuda, 2008).

2. METHODS

10 subjects (6 female, 4 male, all right handed, 23-41 years old) were scanned. Subject prescreening included those fluent in English, and excluded those with any history of psychiatric or neurological conditions, magnet-incompatible implants, claustrophobia, brain injury, or knowledge of written Japanese.

Task

Trials were divided into 4 parts: stimulus (stimulus/category-label presentation), pre-response, response (which included the post-response period) and feedback (see Fig 1). Stimuli were black and white Japanese kanji characters presented in white on a black background (see Fig 1 for an example). Stimulus presentation was followed by the category labels - an image of yellow and blue cartoon fish resting in the bottom corners. Participants were instructed to learn to associate each of the eight kanji images with one of the two category labels. Category label presentation was followed by two response screens. The first, a pre-response cue consisting of two yellow lines in the center of the screen, served to isolate response preparation/planning from responding and post-response processing and/or reward anticipation. Responding took place during the response cue (two green arrows presented at the same location as the pre-response cue). Responses were made by a single button press using the response pads resting on the participants thighs placed under each hand. For example, if the yellow fish was thought to be correct and it appeared on the right-hand side the subject would press the button with their right index finger on the pad held under their right hand. However, the location of the category labels would switch randomly from one trial to the next to prevent the

subject from learning the motor response (e.g., left or right hand) rather than the abstract category label (blue or yellow). Response was followed by feedback then a fixation cross and the start of the next trial. Image presentation times and a task diagram can be found in Fig 1.

On each trial subjects were given one of four possible forms of feedback or reward: three positive reward levels and one negative. Positive reward was either verbal with no monetary reward (“correct”), a small (\$0.10) or a large (\$0.50) monetary reward. Negative feedback was always verbal (“incorrect”); there were no monetary penalties. The feedback type depended on the stimulus: for each subject, of the 8 stimuli, 2 were associated with the high monetary reward, 2 with the low monetary reward, and 4 with verbal feedback. Previous work, in another related task, has demonstrated these levels lead to detectable graded changes in the striatal BOLD response with higher activity for larger monetary rewards (Knutson et al., 2001; M R Delgado et al., 2000).

In order to achieve adequate separation of the trial into its components the sluggish BOLD response requires a much longer acquisition window than the cellular recordings that inspired this work. Post-pilot study interviews suggested that when the task featured only long (4-12 sec) delays between each trial components the task appeared subjectively artificial, as well as was boring to the point of frustration, leading to day dreaming and/or disinterest. To counter this trials were divided into two temporal types - long and short. Short trials were designed to follow a rapid time course (2.5-3 sec), fast enough to require continuous attention. The randomly interspersed long trials allowed for temporal separation of the BOLD signal into discrete trial components by introduction of variable delays, or “jitter”, between components. Long trials contained a

total intra-trial jitter (the variable time between each component of the trial) of longer than 1500 ms, while short have an intra-trial jitter totaling less than 1500 ms. Long trials comprised 45% of all trials. The remaining 55% were short. Both long and short trials were evenly distributed over the three reward levels for a total 80 trials per subject for low and high money conditions and 160 trials for verbal-only trials. The duration of intra-trial (long and short) and inter-trial jitter (the length of the fixation cross) was sampled from three separate exponential distributions. Mean jitter times were 3.6 s (ITI), 2.6 s (long) and 0.3 s (short). While the short condition was too short to allow for separation of trial components its inclusion prevented participants from developing confounding precise temporal expectation. The distributions decay constants were chosen so that the resulting distribution's closely resembled the ideal exponential distributions (Serences, 2004). Only trials in which the participant responded correctly were included in the univariate fMRI analyses.

Subjects were pre-trained in order to familiarize them with the task and limited response window. This task was identical to the task employed during fMRI imaging with the following modifications: a single, non-monetary, reward level (correct/incorrect) was used. If subject responded outside the response window (the green arrows) the trial immediately ceased. Prior to beginning the training subjects were told any interruption in the trial sequence was the result of a response outside the assigned window. During training, subjects completed only 10 trials of eight different stimuli. The stimuli were kanji characters and didn't reappear during fMRI acquisition.

Data acquisition and fMRI preprocessing.

fMRI data collection was performed on a GE 3T magnet, with the following functional acquisition parameters: TE of 26 ms, a whole brain BOLD sampling rate (TR) of 1000 ms, resulting in a slice thickness of 4mm using a standard pulse sequence (gradient-recalled echo-planar imaging, abbreviated SPR-EPI). In addition, whole brain high-resolution anatomical data was acquired at a resolution of 1.2 mm³, again using a standard pulse sequence (spoiled gradient recalled - SPGR). Due to the memory limitations of the scanner the experiment was divided into 4 sections or scans (of about 80 trials each) ranging in length from 12 to 14 minutes. E-prime (Psychology Software Tools, Pittsburgh, PA, version 2) was used to control both stimulus presentation and behavioral recording, using rear-projection inside the scanner and an LCD monitor outside. In scanner responding occurred via two magnet compatible responses boxes. Responses during pre-training were made using the arrow keys on a standard QWERTY keyboard.

In order to achieve the fast 1 sec TR, only partial brain functional data was collected, encompassing ventral striatum and extending dorsally to dorsal lateral premotor cortex (Fig. 2).

fMRI data was preprocessed using standard methods, including rigid body and elastic (12) sub-volume motion correction, slice time normalization (spline fit), and temporal and spatial data smoothing using a high pass filter of 4 Hz and a Gaussian kernel full width at half maximum of 4.0 mm, respectively. Finally, each participant's high resolution anatomical image and functional data was normalized into Talairach space (Talairach & Tournoux, 1988). The hemodynamic response model (i.e. the “design

matrix”) was subjected to the same low-pass filtering as the functional data. The BOLD data was subjected to auto-regressive (AR(1)) smoothing prior to GLM calculation to minimize false positives due to non-white noise (Smith, K. D. Singh, & Balsters, 2007).

Region of interest (ROI) analysis.

Two hand drawn sets of ROIs covering the ventral striatum (“ventral str”), the head of the caudate (“head”), the body of the caudate (“body”), the putamen, and the lateral pre-motorcortex (“lat PreMotor”) were created based on the subjects combined Talarach normalized high-resolution anatomical data. Due to the imperfect nature of the alignment process ROI definitions were conservative, including only voxels that clearly overlapped across subjects. To ensure adequate separation of the head and body of the caudate they were separated by two voxels. Initially unilateral ROIs were created but early analyses showed strong consistency between the two hemispheres, as a result all analyses were carried out bilateral ROIs combining homologous regions in both right and left hemispheres. Changes in activity across trial components were statistically assessed with pair-wise Bonferroni-corrected *t*-tests, unequal variance assumed.

Reinforcement learning model construction.

Using the behavioral and fMRI data we constructed three distinct temporal difference (SARSA - (Sutton & Barto, 1998)) reinforcement learning models that considered the distinct associative relationships. Each model is identical in overall mathematical form (Eq. 1 and 2 discussed below) and differed only in which theoretical associations were reinforced. Said another way, the behavioral data input into each

model is constant but the data was differentially factored based on the proposed associative relation (i.e. the state-space). Model 1 (denoted as SO) proposed that the relationship between the abstract stimulus and the category label is reinforced. Model 2 (denoted as SR) proposed that the relationship between the abstract stimulus and the response is reinforced. Model 3 (denoted as RO) proposed that only the value of the two possible responses are reinforced independent of proceeding abstract stimuli (i.e. independent of the state or context). Each model was constructed using trials from all three feedback levels but treating the received rewards as equivalent (i.e. binary). Equivalency among rewards is a common assumption in all reinforcement learning models (Sutton & Barto, 1998). An control analysis incorporating only verbal feedback trials lead to a nearly identical pattern of results. Two of the regression constants in this control analysis trended towards significance ($p=0.11$) and became significant when all trials were used, suggesting incorporating all trials in this fashion lead only to increased power.

$$\delta = r_t + Q_{t1} - Q_{t0} \quad (1)$$

$$Q_{t0} \leftarrow Q_{t0} + \alpha\delta \quad (2)$$

In Eq 1 Q_{t1} represents the current estimate, Q_{t0} is the estimate from the previous time-step, r_t is the reward (either 0 or 1) and δ is the difference between the previous and current expectations plus the current reward. Eq 2. the old value estimate is updated at the end of the current trial by adding current estimate to the δ multiplied by the learning constant (α). Alpha was set, on subject-by-subject and model-by-model basis, using log-

likelihood parameter optimization where choice probability was estimated using the sigmoidal “soft-max” function (Pessiglione, Seymour, Flandin, R J Dolan, & Frith, 2006; Sutton & Barto, 1998). The soft-max function has a single free parameter (β). Average and standard deviation of parameters for each model: SO - $\alpha=0.17$ (0.08), $\beta=4.48$ (0.40); SR: $\alpha=0.32$ (0.17), $\beta=4.33$ (1.15); RO: $\alpha=0.34$ (0.07) , $\beta=4.47$ (.26). As the average parameters varied between models, and fits varied between subjects, it is possible that differences between models described below were the result of parameter selection. To control for this, each model was rerun for each subject using the model average α value, which controlled for inter-subject variability, and the average α for all models and subjects. Neither altered the significant findings, however both did decrease the value of several regression constants suggesting the individual fit approach lead to increased power. For typical model output see Fig 7.

Parametric regression.

The Q and RPE (i.e., δ) values from the SO, SR, and RO were mapped into separate HRF models of the BOLD response inside BrainVoyager (v2.1.3). Each HRF estimate for each of the three models included regressors for the Q value matching stimulus, pre-response, and post-response, while the RPE was regressed against the feedback period. Like the univariate models described in analysis 1 these models were the subjected to a 4 Hz high-pass filter to match preprocessed BOLD data. These HRF estimates for each model and component (e.g. Q(SO), Q(SR) and Q(RO)) were orthogonalized with respect to each other and the univariate HRF estimates described in above (Matlab R2007a using BVQXtools v0.8c). All general linear model (GLM)

analysis was done using a fixed effects (FFX) model as opposed to the random effects (RFX) model employed in the univariate analysis. The RFX model, with its reduced sensitivity but capacity for generalization beyond the subject pool, showed no significant effects in these analysis. Additionally due to the rapid learning observed in this task, only BOLD data from scan 1 was analyzed as once learning has plateaued the Q and RPE values approach one (which is identical to the univariate analysis regressors) and zero respectively.

3. RESULTS

The results section is comprised of three subsections. The behavioral results, the univariate fMRI analyses which examined both the BOLD response for whole brain and the striatal and cortical regions of interest (ROI) across intra-trial components and reward-levels and the reinforcement learning model based (i.e., multivariate) analyses.

Behavioral results

Behavioral data was divided into 4 blocks, matching the fMRI data acquisition periods (scans). Response accuracy was averaged across subjects and stimulus images for each scan (Fig 3, right). Above chance learning was observed during scan 1 (78.8 %). By scan 3 learning had plateaued at 96%. Each individual's performance was consistent with the group mean, except for a single subject who showed somewhat delayed learning (48% accuracy in scan 1, but reaching 97% by scan 4). Overall reaction times were nearly constant across scans (Fig. 3, left). This was expected, as the 700 ms response window was near the approximately 370 ms floor observed during pilot studies. A 4.4% response miss-rate (i.e. when a subject failed to respond during the response window) was observed. Missed response trials were excluded from both behavioral and fMRI analyses.

Univariate fMRI: whole brain.

A whole-brain analysis controlling for inter-subject variability (i.e. random effects, “RFX”) over both reward-levels and trial components (denoted as all>base) revealed activation across bilateral striatum, lateral premotor cortex, posterior cingulate

cortex (PCC), anterior cingulate cortex (ACC), insula, lateral dorsal prefrontal cortex (PFC), insula, among others (all contrasts employed a threshold of $p < 0.01$, FDR corrected). Reward-level independent activations (HLV>base) for the individual trial components was similarly broad. The stimulus, pre- and post response periods showed significant activation of the dorsal striatum, PPC, insula, lateral premotor cortex, posterior and anterior parietal cortex, and deactivation in dorsomedial PFC; Activity did appear to decrease pre-response in the head and posterior parietal. Post-response activation in the lateral PFC was observed in addition to the areas listed above. During feedback delivery ACC, dorsolateral PFC, ventral as well as dorsal striatal activity was observed. Reward level dependent contrasts for stimulus presentation were nearly identical to the reward-independent contrast for this period while no reward sensitivity was displayed pre- or post-response. Reward-level sensitivity was displayed during feedback robustly in the ventral striatum, with small clusters (<10 voxels) also appearing the head of the caudate and putamen. Overall, the observed patterns of activity are highly consistent with previous fMRI studies (Seger & C. M. Cincotta, 2005; Knutson et al., 2001) suggesting the task and fMRI data are sound.

Univariate regions of interest analysis.

ROI analyses consisted of two parts - (1) reward-level independent (all the reward levels compared to the baseline condition) and (2) reward-level dependent (the two monetary reward levels contrasted to verbal-only).

Univariate reward-level independent activity (ROI).

Significant activation across trial components was observed within all ROIs (Fig. 4, left), excepting the ventral striatum (which however did show reward-level dependent activity, see below). Across the trial components several of the ROIs showed distinct patterns of recruitment (Fig. 4, right). The head of the caudate (“head”) showed a significant ($t(4) = 3.26, p < .05$) drop in activity in the response period as compared to the other component. The body of the caudate was recruited to a consistent degree across all trial components. The putamen was significantly elevated in activity post-response ($t(4) = 5.77, p < .05$). Lateral premotor cortex was most significantly more active during stimulus onset compared to pre-response ($t(4) = 4.23, p < .05$) or response ($t(4) = 2.81, p < .05$).

To examine any learning related changes, reward-level independent activity was examined separately for each of the four scans. The overall pattern across ROIs was that activity tended to decrease with learning stimulus, pre-response and feedback periods, whereas activity in the post response periods was relatively constant (Fig 5). Additionally, the rate of decline appeared to vary by region, especially during stimulus presentation, with a stronger decline in the striatal regions than in the premotor ROI.

Univariate reward-level dependent activity (ROI).

Across trial components, all ROIs showed significant reward level dependent activity, with greater activity for monetary than verbal rewards (Fig 6, left). Analysis of the individual trial components across ROIs (Fig 6, right) revealed that reward level only affected activity at some points during the trial. No ROIs showed reward-level dependent

activity in either the pre or post-response periods. All striatal ROIs showed significant activity at stimulus onset. In addition, the ventral striatum was also recruited during feedback. Although the premotor cortex showed an overall reward level activity effect, when examined across trial components none of the effects reached significance.

Overall reward-level dependent effects were weaker for each ROI when compared to the reward-level independent effects. However, the limited amount of detectable reward-level dependent activity is likely not due to insufficient power. In the reward level independent condition stimulus related activity in both the head and body of the caudate was as strong as that observed pre-response. Additionally activity in the was strongest in the putamen post-response and during feedback in lateral premotor region, therefore these time periods should be most sensitive to reward-level dependent changes. There was insufficient power to examine possible learning (i.e. by scan activity) and reward level interactions.

Model-based regions of interest (ROI) analysis.

During stimulus presentation only model SO positively correlated with BOLD changes in any striatal region; SO correlated with activity in both the body of the caudate and the lateral premotor area. Pre-response, model SR and RF positively correlated with activity in the lateral premotor cortex, while model SO displayed a negative correlation with activity in the head and body of the caudate, the putamen and the lateral premotor ROI. Pre-response there was also a trend ($p=0.073$) towards a positive relation between model SR and the activity in the body of the caudate. Post-response, both models SO and RF correlated positively with putamen activity (each accounting for approximately

equal variance). Both the body and lateral premotor were negatively correlated with model SR (Fig 8.).

During feedback delivery, the RPE from model SO significantly explained variance in the ventral striatum and lateral premotor cortex (Fig 9.). No other region or model reached or approached the significance threshold.

4. DISCUSSION

We scanned subjects while they learned to associate 8 abstract images with two category labels (blue fish or yellow fish) via the feedback received at the end of each trial. By employing a jittered spacing between putative trial components (stimulus, pre-response, post-response, feedback) and a rapid 1 sec sampling time (TR) we were able to separately model, and thus isolate, the BOLD response for each trial component and measure the change in that response across three reward levels. We then estimated the average activity for several regions of interest (ROIs) known to be important in visuomotor learning in order to assess how involved that brain region is (on average) at during each trial component and how that activity changes with learning. We also performed a model-based GLM analysis comparing the output of three reinforcement learning models to the BOLD data for each of the intra-trial components.

Head of the Caudate

When all reward levels (“correct”, \$0.10 and \$0.50 abbreviated as HL>V) were compared to the implicit baseline, activity in the head of the caudate dropped significantly at response, rising again during feedback delivery suggesting that perhaps reward anticipation (Knutson et al., 2001) or response working memory (Hadj-Bouziane & Boussaoud, 2003; Brasted & Wise, 2004) plays a more limited role in overall head function, or at least requires less neural activity to implement. The observed reward-level dependent (“HL>V”) response during stimulus is consistent with the value prediction account (Hollerman et al., 1998; Lau & Glimcher, 2007) as well as reports of

motivational modulation (Mauricio R Delgado, Li, Schiller, & Phelps, 2008). However the lack of reward sensitivity pre-response is inconsistent with action value accounts as is the lack of a significant correlations between any of the reinforcement learning models and BOLD signal changes (Lau & Glimcher, 2007, 2008; Seger et al., 2010). However as correct responses changed on a trial by trial basis while the category responses remained constant, these data raise the possibility that action value tuned PANs require direct and differential reinforcement to discriminate among competing actions. The lack of reward-level sensitivity in the post-response/pre-feedback period in both the head of the caudate and the ventral striatum as well as the reward-level sensitivity observed during feedback delivery is the inconsistent with a reward anticipation account striatal function (Knutson et al., 2001; Knutson, G W Fong, C M Adams, Varner, & D. Hommer, 2001), but is consistent with previous reports of an interaction between reward magnitude and feedback delivery (M R Delgado et al., 2000). The lack or reward sensitivity pre and post response is consistent with a salience account of head function (Zink et al., 2003; 2004; 2006). The reward level sensitivity at stimulus throughout the striatum and at feedback presentation in the ventral striatum does not necessarily contradict a saliency account as responding to both saliency and reward-level during feedback delivery has been previously observed (Cooper & Knutson, 2008).

Body of the caudate

Unlike the head the body of the caudate exhibited a nearly constant level of activity in reward-level independent contrasts, suggesting perhaps a consistent computational role. Like the head of the caudate, this region exhibited a reward-level

sensitivity only at stimulus onset (for a discussion of this finding see the head of the caudate section above). The body was however sensitive to model SO at stimulus -- consistent with previous research in our lab indicating that body of the caudate in the visual loop links stimuli with appropriate category based responses (Seger & C. M. Cincotta, 2005; Seger et al., 2010). A trend towards pre-response activity for model SR in the body is intriguing as it hints that the body might associate stimuli directly with responses and not only with the abstract level category labels.

Putamen

Reward independent contrasts in the putamen show a strong increase in pre-response compared to all other trial components highlighting this region's role in action selection. However no reward-level response was observed in this pre-response period. The putamen positively correlated pre-response with both model SO and model RF indicating both a context dependent response selection and context independent overall biasing towards a particular response. Note that Haruno and Kawato (2006) and Seger *et al* (2010) found putamen associated with Q also. However due to inadequate temporal resolution in these studies they could not localize the temporal origin of the BOLD signal changes. The lack of reward-level response combined with the significant model derived correlation paint a more nuanced view of putamen function. In this view the overall value of the response is not signaled, but the relative value of the category and the context-independent response are.

Ventral striatum

The reward-level dependent activity in the ventral striatum at both stimulus and feedback is consistent with previous reports of this regions importance in pavlovian and affective valuation (M R Delgado, Stenger, & J A Fiez, 2004), reward anticipation (Hollerman et al., 1998; Knutson et al., 2001) and feedback processing (Shohamy et al., 2004; John P O'Doherty et al., 2003). The ventral striatum was sensitive the Q value if model RF at response, and the RPE of model SO during feedback. This indicates that ventral striatum is sensitive to reward value of particular actions (Lau & Glimcher, 2008; Roesch et al., 2009) as well as confirming previous research stating that it is sensitive to RPE integrating stimulus and outcome (Montague et al., 1996; D'Ardenne et al., 2008; Seger et al., 2010).

Lateral premotor

The significant increase in reward-level independent activity observed as stimulus onset where the abstract categories must be mapped onto the variable is consistent with the lateral pre-motor cortex's established role in motor planning. However following the drop off pre-response there was a trend toward increasing activity as the trial progressed, implying this region may play an unappreciated role in responding, feedback anticipation and processing. However while it is unclear the function of this activity, it is likely not involved in estimating or processing relative monetary value as no reward-level dependent responses were observed. Lateral premotor cortex's correlations to the learning models was complex. Model SO correlated positively at Stimulus -- indicating need to integrate stimulus / whole context information into response selection and the SO

RPE at feedback, indicating possible updating of values. The other two models (SR and RF) correlated in the pre-response period, indicating that as the response is closer to being selected, other considerations come into play in place of the high level context.

Significant negative correlations in the multivariate analyses.

In addition to the positive correlations model SO at stimulus and model SR at response were negatively correlated with bold signal changes. It difficult to interpret a negative correlation in a parametric regression model. That several ROIs, many of which showed otherwise different patterns of activity, demonstrated this negative correlation suggest a common computational may occur in each. Though the nature of this computation is unclear.

5. CONCLUSIONS

The patterns of activity across the two analysis approaches deepen our understanding of the contributions of each striatal region and the premotor cortex to visuomotor learning. Ventral striatal activity was characterized by reward-level sensitivity at stimulus and feedback, and by its sensitivity to reward prediction error in the stimulus category (SO) model. These patterns of activity are consistent with the putative role of the ventral striatum in reward (value) prediction, affect, and reward processing. Activity in the ventral striatum decreased across time faster than in any other region suggesting its importance diminishes with learning.

Activity in the dorsal caudate (head and body) extended across all phases of stimulus-response learning. Reward level modulated activity in both of these caudate regions at stimulus onset, suggesting a common role in predicting the immediate value of a stimulus. However the body, but not the head, correlated with SO model's category contingent value estimate during stimulus presentation, implying a role in category selection or visual categorization for the body not shared with that of the head. In addition, reward independent activity in the head was significantly decreased during the pre-response time period, unlike the body's sustained activation.

The putamen was active across all phases of the stimulus-response trial, but peaked in activation in the response phase. This suggests that the putamen plays a stronger role in response selection or response processing than category selection or feedback processing. This response related activity displayed no sensitivity to reward magnitude, though the putamen was sensitive to reward magnitude at the time of stimulus

presentation. Activity in the response phase in the putamen had a positive correlation with both model SO and RF suggesting that the putamen tracks both the value of the current category and the average value of the two responses independent of previous context.

Lateral premotor cortex displayed the greatest range of activities; its functional role may be rather complex. Lateral premotor cortex was significantly active across all phases of the trial. Activity was greatest at the time of stimulus presentation: in addition activity at this time was correlated with the SO model suggesting either the value of the selected category is incorporated into this regions motor planning capabilities or this region plays a previously unappreciated role in category selection. However, pre-response activity was best described by two the models of response value (SR and RF), and was negatively related to the category model (SO), while activity at feedback was well modeled by the SO model's RPE. Whether this reflects reward or response-outcome processing is unclear.

In total these results are consistent with the notion that the both the striatum and lateral premotor cortex contribute separate, context dependent, calculations throughout visuomotor learning. As such the competing views of striatal function may be reconcilable. What you find depends on when you look.

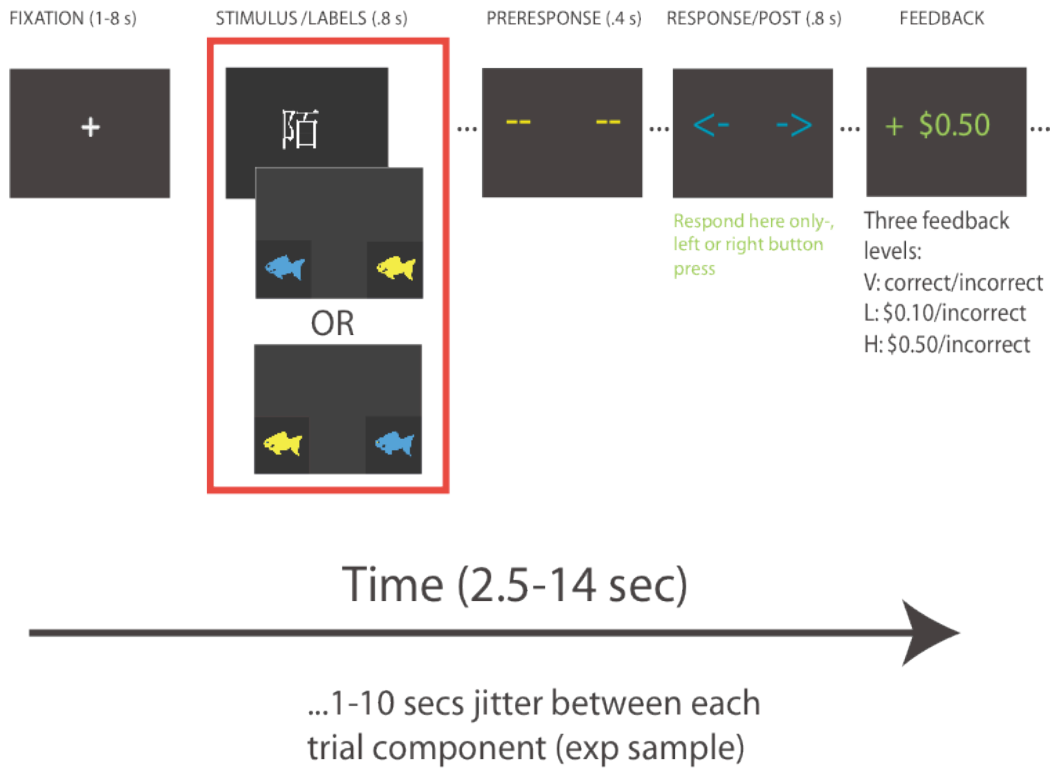


Fig 1. Diagram of the behavioral task.

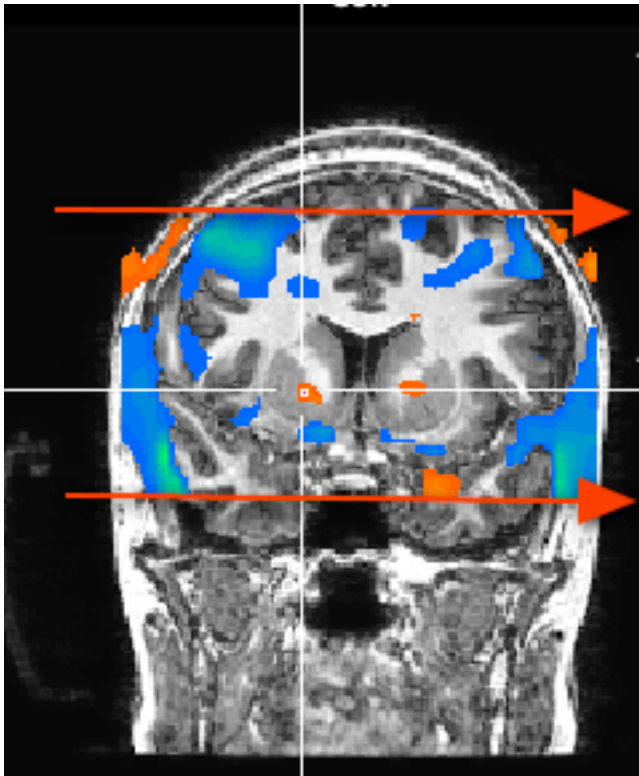


Fig 2. Functional acquisition window. Example data taken from Subject 1, scan, contrast: stim>base.

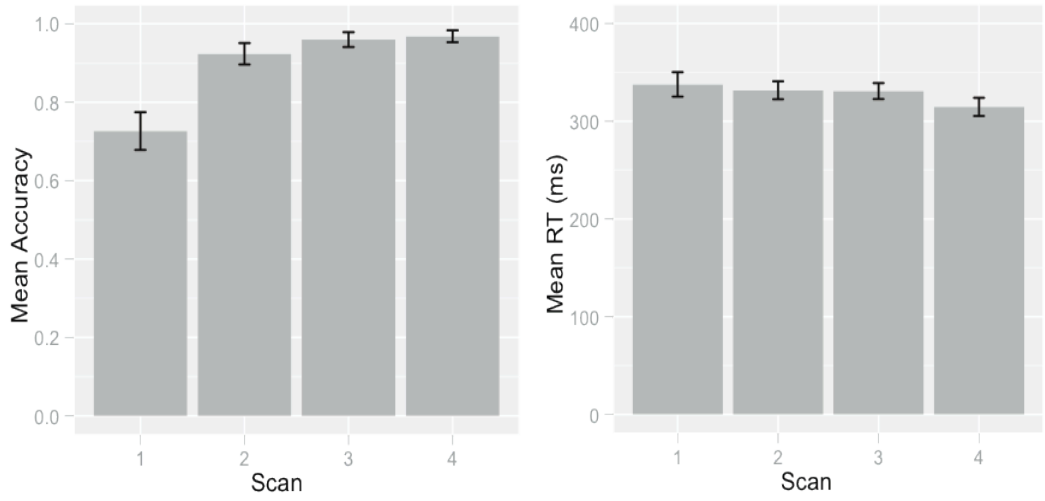


Fig 3. Average accuracy (left) and reaction times (right) across fMRI scanning blocks (80/each) for all subjects (error bar: SE).

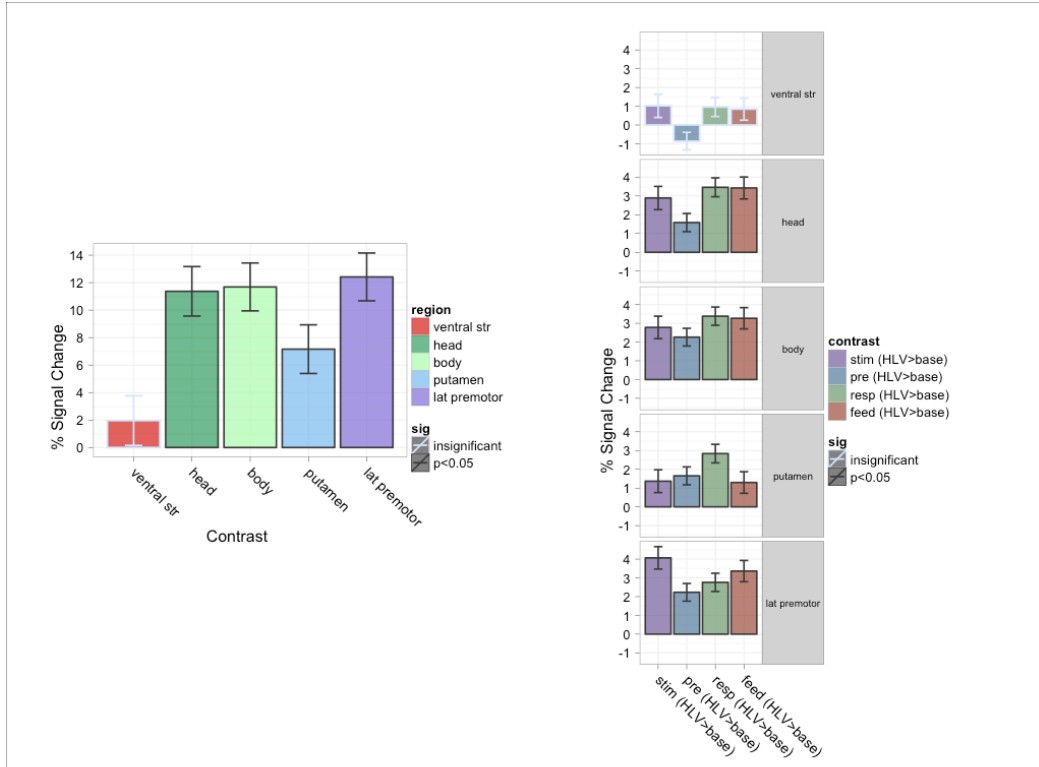
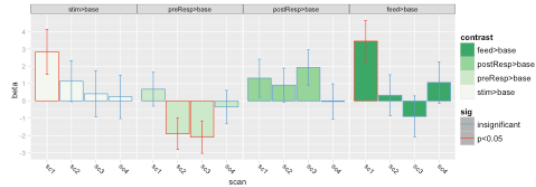
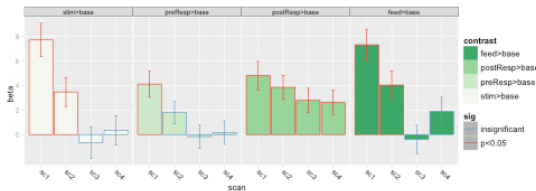


Fig 4: Reward-level independent ROI activation patterns abbreviated as “HLV>base”. **Left.** Overall activity, that is all components combined (error bars: SE). **Right.** Intra-trial component activity (error bars: SE). Units are the absolute arbitrary regression constants (i.e. Beta values) that reflect the degree of linear relation between the BOLD data and the HRF convolved model. Significance in figure legends is for the regression constants (Beta value).

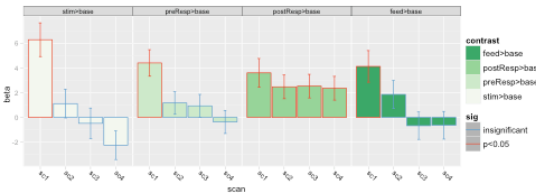
A. Ventral Striatum



B. Head



C. Putamen



D. Lat. premotor

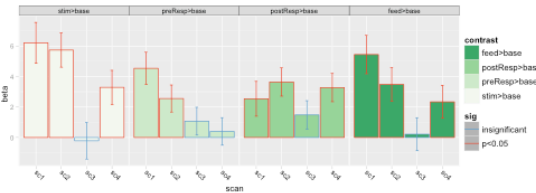


Fig 5. Time-course of BOLD signal changes for select ROIs at stimulus, pre and post-response, and feedback. Activity patterns in body of the caudate were identical to those observed in the head (**B.** above). Significance in figure legends is for the regression constants (Beta value).

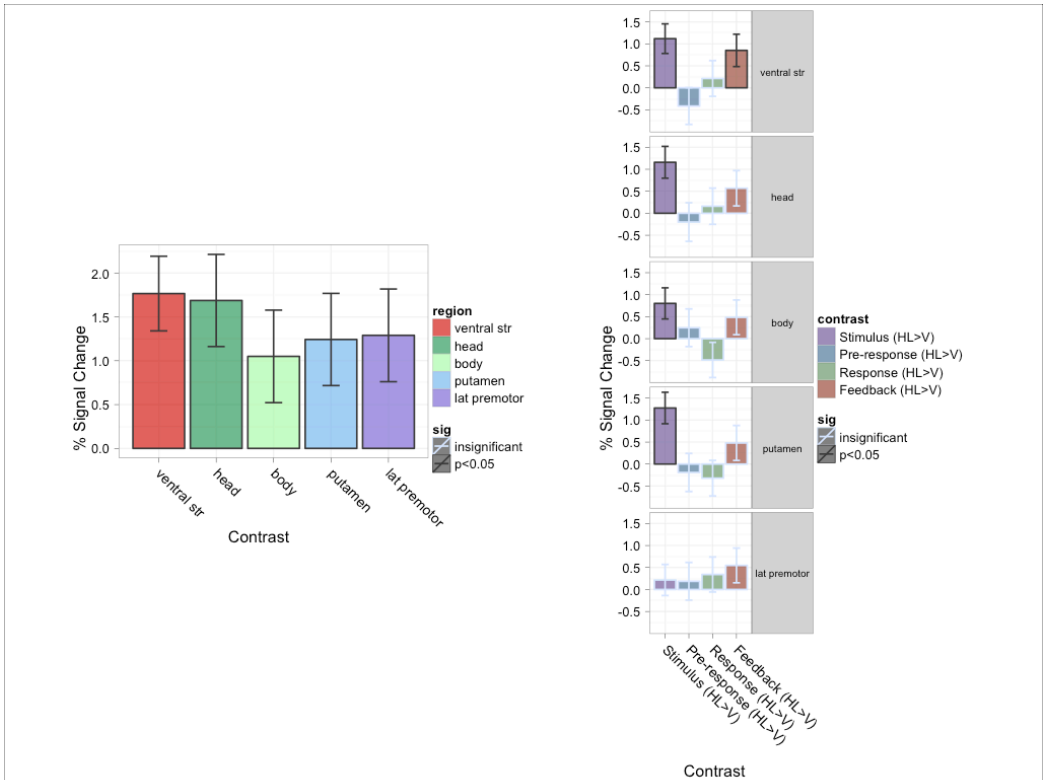


Fig 6: Reward-level dependent ROI activation patterns abbreviated as “HL>V”. **Let.** Overall activity, that is all components combined (error bars: SE). **Right.** Intra-trial component activity (error bars: SE). Units are the absolute arbitrary regression constants (i.e. Beta values) that reflect the degree of linear relation between the BOLD data and the HRF convolved model. Significance in figure legends is for the regression constants (Beta value).

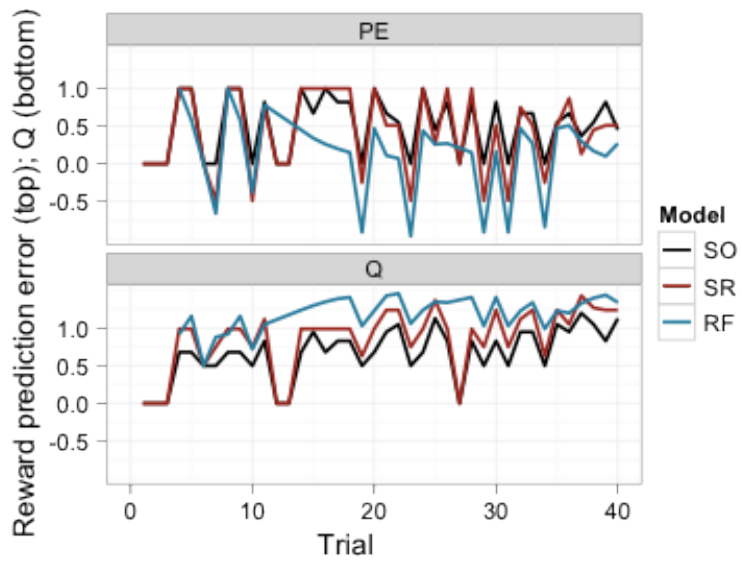


Fig 7. Typical out from the three reinforcement learning model (Subject 2) prior to HRF (hemodynamic response function) convolution and orthogonalization.

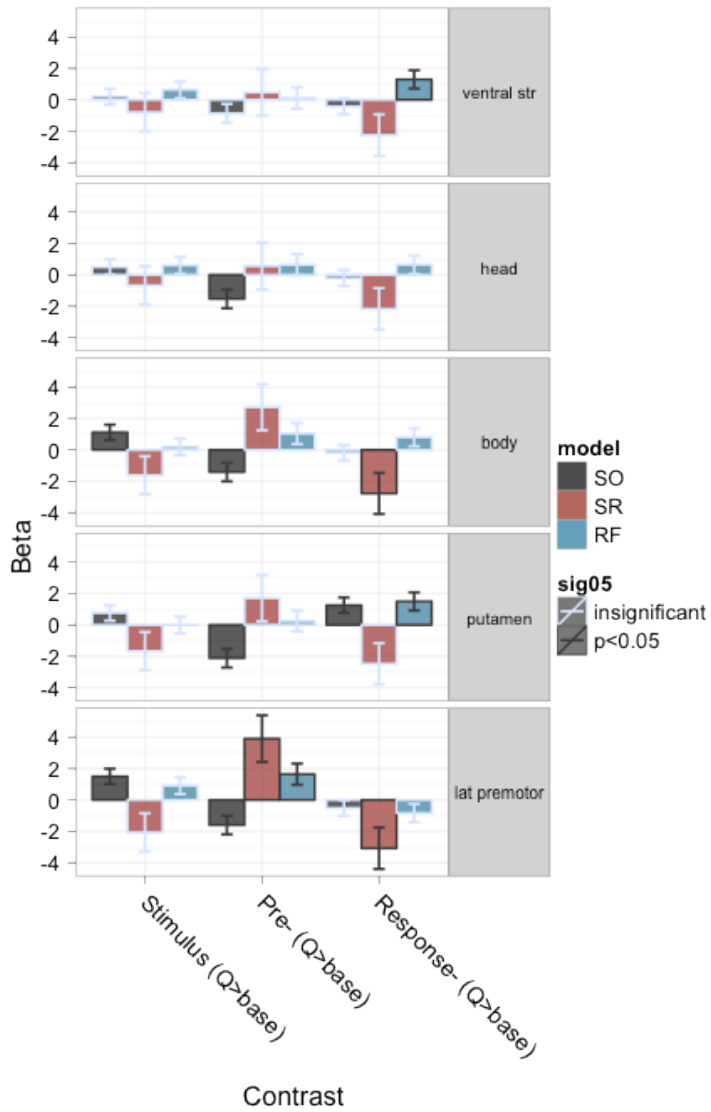


Fig. 8. Regression analysis of Q values (the three reinforcement learning models (SO, SR, and RF) estimate of total future rewards) for stimulus, pre-response and post-response BOLD data (scan 1).

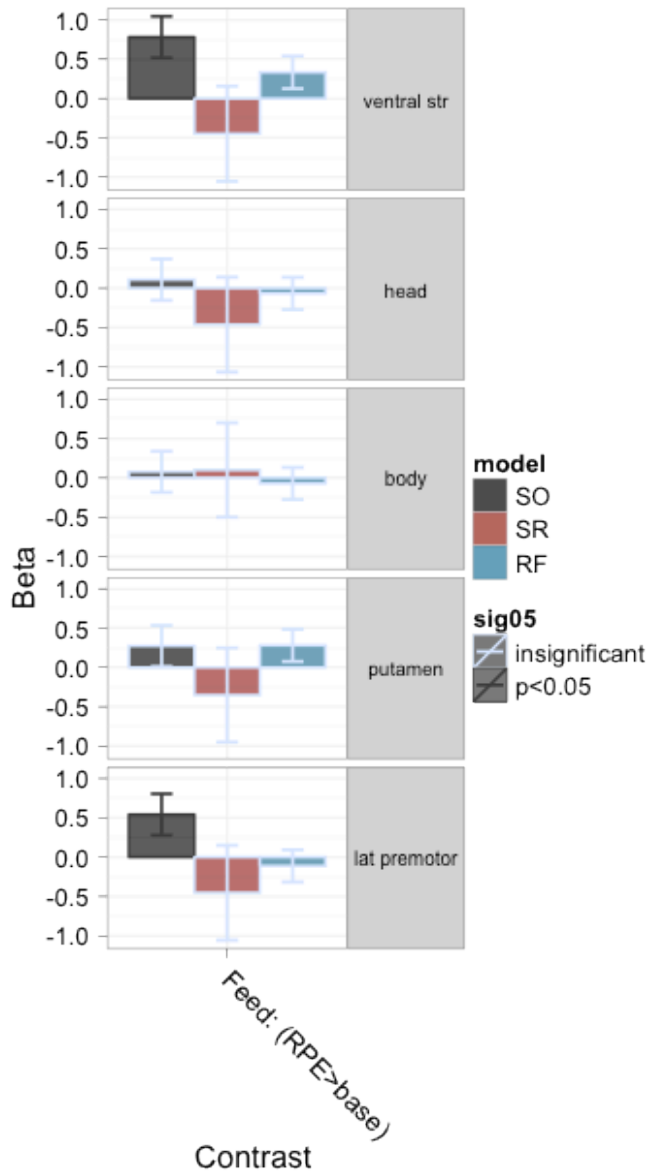


Fig 9. Regression analysis of the reward prediction error (RPE) for each of the three reinforcement learning models (SO, SR, and RF) for feedback related BOLD changes in scan 1.

REFERENCES

- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience and Biobehavioral Reviews*, *32*(2), 265–78. doi:10.1016/j.neubiorev.2007.07.010
- Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, *58*(3), 451–63. doi:10.1016/j.neuron.2008.02.021
- Schmitzer-Torbert, N., & Redish, A. D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J Physiol.*, *91*(5), 2259–72. doi:10.1152/jn.00687.2003
- Xue, G., Ghahremani, D. G., & Poldrack, R. A. (2008). Neural substrates for reversing stimulus-outcome and stimulus-response associations. *J Neurosci*, *28*(44), 11196–204. doi:10.1523/JNEUROSCI.4001-08.2008
- Seger, C. A., & Cincotta, C. M. (2005). The Roles of the Caudate Nucleus in Human Classification Learning. *J Neurosci*, *25*(11), 2941–2951.
- Seger, C. A. (2007). How do the basal ganglia (bg) contribute to categorization. *Neuroscience and Biobehavioral Reviews*, *inprep*.
- Seger, C. A., Peterson, E. J., Cincotta, C. M., Lopez-Paniagua, D., & Anderson, C. W. (2010). Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage*, *50*(2), 644–56. doi:10.1016/j.neuroimage.2009.11.083
- Sun, F. T., Miller, L. M., & D'Esposito, M. (2005). Measuring temporal dynamics of functional networks using phase spectrum of fMRI data. *Neuroimage*, *28*(1), 227–37. doi:10.1016/j.neuroimage.2005.05.043

- Brasted, P. J., & Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur J Neurosci*, *19*(3), 721–40.
- Buch, E. R., Brasted, P. J., & Wise, S. P. (2006). Comparison of population activity in the dorsal premotor cortex and putamen during the learning of arbitrary visuomotor mappings. *Experimental Brain Research*, *169*(1), 69–84. doi:10.1007/s00221-005-0130-y
- Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., Roy, S. A., et al. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philos Trans R Soc Lond, B, Biol Sci*, *362*(1485), 1573–83. doi:10.1098/rstb.2007.2063
- Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, *437*(7062), 1158–61. doi:10.1038/nature04053
- Bar-Gad, I., Morris, G., & Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology*, *71*(6), 439–73. doi:10.1016/j.pneurobio.2003.12.001
- Hadj-Bouziane, F., Meunier, M., & Boussaoud, D. (2003). Conditional visuo-motor learning in primates: a key role for the basal ganglia. *J Physiol Paris*, *97*(4-6), 567–79. doi:10.1016/j.jphysparis.2004.01.014
- Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Physiol.*, *80*(2), 947–63.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci*, *21*(16), RC159.
- Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to

- reward and punishment: effects of valence and magnitude manipulations. *Cognitive, affective & behavioral neuroscience*, 3(1), 27–38.
- Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., & Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *J Neurosci*, 23(22), 8092–7.
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3), 509–17.
- Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *Neuroimage*, 39(1), 538–47. doi:10.1016/j.neuroimage.2007.08.009
- Wittmann, B. C., Nathaniel D Daw, N. W., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron*, 58(6), 967–73. doi:10.1016/j.neuron.2008.04.027
- Toni, I., Thoenissen, D., & Zilles, K. (2001). Movement preparation and motor intention. *Neuroimage*, 14(1 Pt 2), S110–7. doi:10.1006/nimg.2001.0841
- Tanaka, S. C., Balleine, B. W., & O'Doherty, J. P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. *J Neurosci*, 28(26), 6750–5. doi:10.1523/JNEUROSCI.1808-08.2008
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304(5669), 452–4. doi:10.1126/science.1094285
- Yin, H. H., Ostlund, S. B., & Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur J Neurosci*, 28(8), 1437–48. doi:10.1111/j.1460-9568.2008.06422.x
- Hadj-Bouziane, F., & Boussaoud, D. (2003). Neuronal activity in the monkey striatum during

- conditional visuomotor learning. *Experimental Brain Research*, 153(2), 190–6.
doi:10.1007/s00221-003-1592-4
- Toni, I., Schluter, N. D., Josephs, O., Friston, K., & Passingham, R. E. (1999). Signal-, set- and movement-related activity in the human brain: an event-related fMRI study. *Cereb Cortex*, 9(1), 35–49.
- Rodriguez, P. F., Aron, A., & Poldrack, R. (2006). Ventral–striatal/nucleus–accumbens sensitivity to prediction errors during classification learning. *Human Brain Mapping*, 27(4), 306–313. doi:10.1002/hbm.20186
- Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, 58(2), 284–94. doi:10.1016/j.neuron.2008.03.020
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annu Rev Neurosci*, 29, 417–448.
- Cincotta, C. M., & Seger, C. A. (2007). Dissociation between striatal regions while learning to categorize via feedback and via observation. *Journal of cognitive neuroscience*, 19(2), 249–65. doi:10.1162/jocn.2007.19.2.249
- Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Physiol.*, 95(2), 948–59. doi:10.1152/jn.00382.2005
- Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci*, 12(12), 4595–610.
- Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J Physiol.*, 72(2), 1024.
- Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by

- appetitive rather than aversive stimuli. *Nature*, 379, 449–451.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*, 16(5), 1936.
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, 319(5867), 1264–7. doi:10.1126/science.1150605
- Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*, 10(12), 1615–24. doi:10.1038/nn2013
- Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *PNAS*, 104(4), 1377–82. doi:10.1073/pnas.0606297104
- Kakade, S., & Dayan, P. (2002). Dopamine: generalization and bonuses. *Neural Networks*, 15(4-6), 549–559.
- Guitart-Masip, M., Bunzeck, N., Stephan, K. E., Dolan, R. J., & Duzel, E. (2010). Contextual Novelty Changes Reward Representations in the Striatum. *Journal of Neuroscience*, 30(5), 1721.
- Knutson, B., Adams, C., Fong, G., & Hommer, D. (2001). Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens. *J Neurosci*, 21(16), 159.
- Knutson, B., & Wimmer, G. E. (2007). Splitting the difference: how does the brain code reward episodes? *Annals of the New York Academy of Sciences*, 1104, 54–69. doi:10.1196/annals.1390.020
- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable

- neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585–95. doi:10.1016/j.neuron.2010.04.016
- O'Reilly, R., Frank, M. J., Hazy, T., & Watz, B. (2007). PVLV: the primary value and learned value Pavlovian learning algorithm. *Behav. Neurosci.* Retrieved from <http://content.apa.org/journals/bne/121/1/31.pdf>
- Dayan, P. (2007). Motivated Reinforcement Learning. *NIPS*, 1–8.
- Redish, A. D. (2004). Addiction as a computational process gone awry. *Science*, 306(5703), 1944–7. doi:10.1126/science.1102384
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–902. doi:10.1126/science.1077349
- Tricomi, E., & Fiez, J. A. (2008). Feedback signals in the caudate reflect goal achievement on a declarative memory task. *Neuroimage*, 41(3), 1154–67. doi:10.1016/j.neuroimage.2008.02.066
- Blatter, K., & Schultz, W. (2006). Rewarding properties of visual stimuli. *Experimental Brain Research*, 168(4), 541–6. doi:10.1007/s00221-005-0114-y
- Dommett, E., Coizet, V., Blaha, C. D., Martindale, J., Lefebvre, V., Walton, N., Mayhew, J. E. W., et al. (2005). How visual stimuli activate dopaminergic neurons at short latency. *Science*, 307(5714), 1476–9. doi:10.1126/science.1107026
- Redgrave, P., Gurney, K., & Reynolds, J. (2007). What is reinforced by phasic dopamine signals? *Brain Research Reviews*. doi:10.1016/j.brainresrev.2007.10.007
- Zink, C. F., Pagnoni, G., Chappelow, J., Martin-Skurski, M., & Berns, G. S. (2006). Human striatal activation reflects degree of stimulus saliency. *Neuroimage*, 29(3), 977–83.

doi:10.1016/j.neuroimage.2005.08.006

Fiorillo, C., Newsome, W., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*. doi:10.1038/nn.2159

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *J Neurosci*, *25*(19), 4806–12.

doi:10.1523/JNEUROSCI.0642-05.2005

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci*, *10*(12), 1625–33. doi:10.1038/nn2007

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006).

Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci*, *26*(37), 9530–7. doi:10.1523/JNEUROSCI.2915-06.2006

Kim, S., & Fukuda, M. (2008). Lessons from fMRI about mapping cortical columns. *The Neuroscientist*, *14*(3), 287–99. doi:10.1177/1073858407309541

O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*, *49*(1), 157–66. doi:10.1016/j.neuron.2005.11.014

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J Physiol.*, *84*(6), 3072–7.

Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, *21*(4), 1690–700.

doi:10.1016/j.neuroimage.2003.12.021

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. *Thieme*,

New York.

- Smith, A. T., Singh, K. D., & Balsters, J. H. (2007). A comment on the severity of the effects of non-white noise in fMRI time-series. *Neuroimage*, *36*(2), 282–8. doi:10.1016/j.neuroimage.2006.09.044
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: an Introduction*.
- Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *J Neurosci*, *27*(52), 14502–14. doi:10.1523/JNEUROSCI.3060-07.2007
- Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond, B, Biol Sci*, *363*(1511), 3787–800. doi:10.1098/rstb.2008.0161
- Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*(17), 3683–7.
- Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain*, *127*(Pt 4), 851–9. doi:10.1093/brain/awh100
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–37.
- Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci*, *29*(42), 13365–76. doi:10.1523/JNEUROSCI.2572-09.2009
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine

- dependent prediction error underpin reward-seeking behaviour in humans. *Nature Neuroscience*, 442(31), 1042–1045.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci*, 4(1), 95–102. doi:10.1038/82959
- Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, 57(2), 314–25. doi:10.1016/j.neuron.2007.11.031
- Grafton, S. T., Schmitt, P., Horn, J. V., & Diedrichsen, J. (2008). Neural substrates of visuomotor learning based on improved feedback control and prediction. *Neuroimage*, 39(3), 1383–95. doi:10.1016/j.neuroimage.2007.09.062
- Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci*, 28(51), 13775–85. doi:10.1523/JNEUROSCI.3541-08.2008
- O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, 314(5796), 91–4. doi:10.1126/science.1127242
- Mars, R. B., & Grol, M. J. (2007). Dorsolateral prefrontal cortex, working memory, and prospective coding for action. *J Neurosci*, 27(8), 1801–2.
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual review of psychology*, 56, 149–78. doi:10.1146/annurev.psych.56.091103.070217
- Ueda, Y., & Kimura, M. (2003). Encoding of direction and combination of movements by primate putamen neurons. *Eur J Neurosci*, 18(4), 980–94.
- Bray, S., & O'Doherty, J. P. (2007). Neural coding of reward prediction error signals during

- classical conditioning with attractive faces. *J Physiol.*, 97(3036), 3036–3045.
- Delgado, M. R., Stenger, V. A., & Fiez, J. A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cereb Cortex*, 14(9), 1022–30. doi:10.1093/cercor/bhh062
- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience and Biobehavioral Reviews*, 32(2), 265–78. doi:10.1016/j.neubiorev.2007.07.010
- Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, 58(3), 451–63. doi:10.1016/j.neuron.2008.02.021
- Schmitzer-Torbert, N., & Redish, A. D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J Physiol.*, 91(5), 2259–72. doi:10.1152/jn.00687.2003
- Xue, G., Ghahremani, D. G., & Poldrack, R. A. (2008). Neural substrates for reversing stimulus-outcome and stimulus-response associations. *J Neurosci*, 28(44), 11196–204. doi:10.1523/JNEUROSCI.4001-08.2008
- Seger, C. A., & Cincotta, C. M. (2005). The Roles of the Caudate Nucleus in Human Classification Learning. *J Neurosci*, 25(11), 2941–2951.
- Seger, C. A. (2007). How do the basal ganglia (bg) contribute to categorization. *Neuroscience and Biobehavioral Reviews*, *inprep.*
- Seger, C. A., Peterson, E. J., Cincotta, C. M., Lopez-Paniagua, D., & Anderson, C. W. (2010). Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage*, 50(2), 644–56. doi:10.1016/j.neuroimage.2009.11.083
- Sun, F. T., Miller, L. M., & D'Esposito, M. (2005). Measuring temporal dynamics of functional

- networks using phase spectrum of fMRI data. *Neuroimage*, 28(1), 227–37.
doi:10.1016/j.neuroimage.2005.05.043
- Brasted, P. J., & Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur J Neurosci*, 19(3), 721–40.
- Buch, E. R., Brasted, P. J., & Wise, S. P. (2006). Comparison of population activity in the dorsal premotor cortex and putamen during the learning of arbitrary visuomotor mappings. *Experimental Brain Research*, 169(1), 69–84. doi:10.1007/s00221-005-0130-y
- Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., Roy, S. A., et al. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philos Trans R Soc Lond, B, Biol Sci*, 362(1485), 1573–83.
doi:10.1098/rstb.2007.2063
- Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, 437(7062), 1158–61. doi:10.1038/nature04053
- Bar-Gad, I., Morris, G., & Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology*, 71(6), 439–73. doi:10.1016/j.pneurobio.2003.12.001
- Hadj-Bouziane, F., Meunier, M., & Boussaoud, D. (2003). Conditional visuo-motor learning in primates: a key role for the basal ganglia. *J Physiol Paris*, 97(4-6), 567–79.
doi:10.1016/j.jphysparis.2004.01.014
- Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Physiol.*, 80(2), 947–63.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing

- monetary reward selectively recruits nucleus accumbens. *J Neurosci*, 21(16), RC159.
- Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, affective & behavioral neuroscience*, 3(1), 27–38.
- Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., & Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *J Neurosci*, 23(22), 8092–7.
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3), 509–17.
- Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *Neuroimage*, 39(1), 538–47. doi:10.1016/j.neuroimage.2007.08.009
- Wittmann, B. C., Nathaniel D Daw, N. W., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron*, 58(6), 967–73. doi:10.1016/j.neuron.2008.04.027
- Toni, I., Thoenissen, D., & Zilles, K. (2001). Movement preparation and motor intention. *Neuroimage*, 14(1 Pt 2), S110–7. doi:10.1006/nimg.2001.0841
- Tanaka, S. C., Balleine, B. W., & O'Doherty, J. P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. *J Neurosci*, 28(26), 6750–5. doi:10.1523/JNEUROSCI.1808-08.2008
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304(5669), 452–4. doi:10.1126/science.1094285
- Yin, H. H., Ostlund, S. B., & Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks.

- Eur J Neurosci*, 28(8), 1437–48. doi:10.1111/j.1460-9568.2008.06422.x
- Hadj-Bouziane, F., & Boussaoud, D. (2003). Neuronal activity in the monkey striatum during conditional visuomotor learning. *Experimental Brain Research*, 153(2), 190–6. doi:10.1007/s00221-003-1592-4
- Toni, I., Schluter, N. D., Josephs, O., Friston, K., & Passingham, R. E. (1999). Signal-, set- and movement-related activity in the human brain: an event-related fMRI study. *Cereb Cortex*, 9(1), 35–49.
- Rodriguez, P. F., Aron, A., & Poldrack, R. (2006). Ventral–striatal/nucleus–accumbens sensitivity to prediction errors during classification learning. *Human Brain Mapping*, 27(4), 306–313. doi:10.1002/hbm.20186
- Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, 58(2), 284–94. doi:10.1016/j.neuron.2008.03.020
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annu Rev Neurosci*, 29, 417–448.
- Cincotta, C. M., & Seger, C. A. (2007). Dissociation between striatal regions while learning to categorize via feedback and via observation. *Journal of cognitive neuroscience*, 19(2), 249–65. doi:10.1162/jocn.2007.19.2.249
- Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Physiol.*, 95(2), 948–59. doi:10.1152/jn.00382.2005
- Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci*, 12(12), 4595–610.
- Mirenowicz, J., & Schultz, W. (1994). Importance of unpredictability for reward responses in

- primate dopamine neurons. *J Physiol.*, 72(2), 1024.
- Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, 379, 449–451.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*, 16(5), 1936.
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, 319(5867), 1264–7. doi:10.1126/science.1150605
- Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*, 10(12), 1615–24. doi:10.1038/nn2013
- Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *PNAS*, 104(4), 1377–82. doi:10.1073/pnas.0606297104
- Kakade, S., & Dayan, P. (2002). Dopamine: generalization and bonuses. *Neural Networks*, 15(4-6), 549–559.
- Guitart-Masip, M., Bunzeck, N., Stephan, K. E., Dolan, R. J., & Duzel, E. (2010). Contextual Novelty Changes Reward Representations in the Striatum. *Journal of Neuroscience*, 30(5), 1721.
- Knutson, B., Adams, C., Fong, G., & Hommer, D. (2001). Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens. *J Neurosci*, 21(16), 159.
- Knutson, B., & Wimmer, G. E. (2007). Splitting the difference: how does the brain code reward episodes? *Annals of the New York Academy of Sciences*, 1104, 54–69.

doi:10.1196/annals.1390.020

- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–95. doi:10.1016/j.neuron.2010.04.016
- O'Reilly, R., Frank, M. J., Hazy, T., & Watz, B. (2007). PVLV: the primary value and learned value Pavlovian learning algorithm. *Behav. Neurosci.* Retrieved from <http://content.apa.org/journals/bne/121/1/31.pdf>
- Dayan, P. (2007). Motivated Reinforcement Learning. *NIPS*, 1–8.
- Redish, A. D. (2004). Addiction as a computational process gone awry. *Science*, *306*(5703), 1944–7. doi:10.1126/science.1102384
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–902. doi:10.1126/science.1077349
- Tricomi, E., & Fiez, J. A. (2008). Feedback signals in the caudate reflect goal achievement on a declarative memory task. *Neuroimage*, *41*(3), 1154–67. doi:10.1016/j.neuroimage.2008.02.066
- Blatter, K., & Schultz, W. (2006). Rewarding properties of visual stimuli. *Experimental Brain Research*, *168*(4), 541–6. doi:10.1007/s00221-005-0114-y
- Dommett, E., Coizet, V., Blaha, C. D., Martindale, J., Lefebvre, V., Walton, N., Mayhew, J. E. W., et al. (2005). How visual stimuli activate dopaminergic neurons at short latency. *Science*, *307*(5714), 1476–9. doi:10.1126/science.1107026
- Redgrave, P., Gurney, K., & Reynolds, J. (2007). What is reinforced by phasic dopamine signals? *Brain Research Reviews*. doi:10.1016/j.brainresrev.2007.10.007

- Zink, C. F., Pagnoni, G., Chappelow, J., Martin-Skurski, M., & Berns, G. S. (2006). Human striatal activation reflects degree of stimulus saliency. *Neuroimage*, *29*(3), 977–83.
doi:10.1016/j.neuroimage.2005.08.006
- Fiorillo, C., Newsome, W., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*. doi:10.1038/nn.2159
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *J Neurosci*, *25*(19), 4806–12.
doi:10.1523/JNEUROSCI.0642-05.2005
- Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci*, *10*(12), 1625–33. doi:10.1038/nn2007
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci*, *26*(37), 9530–7. doi:10.1523/JNEUROSCI.2915-06.2006
- Kim, S., & Fukuda, M. (2008). Lessons from fMRI about mapping cortical columns. *The Neuroscientist*, *14*(3), 287–99. doi:10.1177/1073858407309541
- O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*, *49*(1), 157–66. doi:10.1016/j.neuron.2005.11.014
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J Physiol.*, *84*(6), 3072–7.
- Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *Neuroimage*, *21*(4), 1690–700.

doi:10.1016/j.neuroimage.2003.12.021

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. *Thieme, New York.*

Smith, A. T., Singh, K. D., & Balsters, J. H. (2007). A comment on the severity of the effects of non-white noise in fMRI time-series. *Neuroimage, 36*(2), 282–8.

doi:10.1016/j.neuroimage.2006.09.044

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: an Introduction.*

Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *J Neurosci, 27*(52), 14502–14. doi:10.1523/JNEUROSCI.3060-07.2007

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond, B, Biol Sci, 363*(1511), 3787–800. doi:10.1098/rstb.2008.0161

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport, 12*(17), 3683–7.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain, 127*(Pt 4), 851–9. doi:10.1093/brain/awh100

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38*(2), 329–37.

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci, 29*(42), 13365–76. doi:10.1523/JNEUROSCI.2572-

09.2009

- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine dependent prediction error underpin reward-seeking behaviour in humans. *Nature Neuroscience*, *442*(31), 1042–1045.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci*, *4*(1), 95–102. doi:10.1038/82959
- Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–25. doi:10.1016/j.neuron.2007.11.031
- Grafton, S. T., Schmitt, P., Horn, J. V., & Diedrichsen, J. (2008). Neural substrates of visuomotor learning based on improved feedback control and prediction. *Neuroimage*, *39*(3), 1383–95. doi:10.1016/j.neuroimage.2007.09.062
- Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci*, *28*(51), 13775–85. doi:10.1523/JNEUROSCI.3541-08.2008
- O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–4. doi:10.1126/science.1127242
- Mars, R. B., & Grol, M. J. (2007). Dorsolateral prefrontal cortex, working memory, and prospective coding for action. *J Neurosci*, *27*(8), 1801–2.
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual review of psychology*, *56*, 149–78. doi:10.1146/annurev.psych.56.091103.070217
- Ueda, Y., & Kimura, M. (2003). Encoding of direction and combination of movements by

primate putamen neurons. *Eur J Neurosci*, *18*(4), 980–94.

Bray, S., & O'Doherty, J. P. (2007). Neural coding of reward prediction error signals during classical conditioning with attractive faces. *J Physiol.*, *97*(3036), 3036–3045.

Delgado, M. R., Stenger, V. A., & Fiez, J. A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cereb Cortex*, *14*(9), 1022–30. doi:10.1093/cercor/bhh062