NeuroImage 159 (2017) 9–17

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Local striatal reward signals can be predicted from corticostriatal connectivity

Peter Smittenaar[a,*], Zeb Kurth-Nelson[a,b], Siawoosh Mohammadi[a,c], Nikolaus Weiskopf[a,d], Raymond J. Dolan[a,b]

[a] Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London, WC1N 3BG, UK
[b] Max Planck-University College London Centre for Computational Psychiatry and Ageing Research, London, WC1B 5EH, UK
[c] Department of Systems Neuroscience, Medical Center Hamburg-Eppendorf, Hamburg, Germany
[d] Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, 04103, Leipzig, Germany

ABSTRACT

A defining feature of the basal ganglia is their anatomical organization into multiple cortico-striatal loops. A central tenet of this architecture is the idea that local striatal function is determined by its precise connectivity with cortex, creating a functional topography that is mirrored within cortex and striatum. Here we formally test this idea using both human anatomical and functional imaging, specifically asking whether within striatal subregions one can predict between-voxel differences in functional signals based on between-voxel differences in corticostriatal connectivity. We show that corticostriatal connectivity profiles predict local variation in reward signals in bilateral caudate nucleus and putamen, expected value signals in bilateral caudate nucleus, and response effector activity in bilateral putamen. These data reveal that, even within individual striatal regions, local variability in corticostriatal anatomical connectivity predicts functional differentiation.

1. Introduction

The basal ganglia, the central structures in reward-guided action selection, exhibit a remarkably intricate architecture whereby inputs from cortex are topographically organized into multiple cortico-striatal loops (Alexander et al., 1986). Rather than a division into neatly segregated pathways, axons from multiple cortical regions converge in overlapping parts of the striatum (Averbeck et al., 2014; Haber, 2010). This places the striatum at a crossroads of information processing thought to drive, amongst other functions, reward-guided behaviors (Averbeck et al., 2014; Haber and Behrens, 2014). This arrangement is somewhat at odds with an otherwise rigid anatomical parcellation of the striatum into nucleus accumbens, caudate nucleus and putamen (Voorn et al., 2004). Here we ask whether knowing the corticostriatal inputs to each voxel of the striatum allows us to predict functional activity within that voxel. If we can do so even *within* classical subregions of the striatum – the caudate nucleus and the putamen – then this suggests that a detailed knowledge of structural connectivity can provide a more detailed guide to local function than does anatomy alone.

We tested our hypothesis using a methodology first reported in a study within the visual domain (Saygin et al., 2012). This previous study reported an accurate prediction of functional responses to faces versus scenes for individual voxels in the fusiform gyrus based on structural connectivity fingerprints of these very same voxels. Although this approach has been extended to visual responses in other regions of cortex (Osher et al., 2015), to the best of our knowledge it has not been applied to higher cognitive functions or to an examination of subcortical structures. Given the great diversity of inputs into the striatum—spanning most of cortex (Alexander et al., 1986)—its subregions are particularly well-suited for an examination of such structure-function relationships. Specifically, we examined the caudate nucleus and putamen during an instrumental reinforcement learning task using functional and diffusion-weighted magnetic resonance imaging (MRI). To validate our approach we examined motor effector activity related to hand and foot actions. We then used this method to predict individual intra-region variability in the expression of reward and expected value signals from individual corticostriatal connectivity profiles, finding a dependence of function on each voxel's distinct pattern of cortical connectivity.

2. Materials and methods

2.1. Participants

Twenty-four adults participated in the experiment (14 female, 10

---

[*] Corresponding author. 12 Queen Square, WC1N 3BG, London, UK.
E-mail address: petersmittenaar@gmail.com (P. Smittenaar).

http://dx.doi.org/10.1016/j.neuroimage.2017.07.042
Received 12 February 2017; Received in revised form 24 May 2017; Accepted 19 July 2017

male; age range 18–36 years; mean ± SD = 22.5 ± 4.5 years). All participants were right hand dominant, had no history of psychiatric or neurological disorder, were not taking any medication known to affect neural or cognitive function, had normal or corrected-to-normal vision and passed the safety requirements to enter a MRI scanner. All subjects provided written informed consent prior to the start of the experiment, which was approved by the Research Ethics Committee at University College London (UK). One further subject was excluded due to excessive movement (images could not be realigned successfully).

2.2. *Overview of the approach*

We tested the hypothesis that corticostriatal input into the caudate nucleus and putamen reliably predicts functional responses during instrumental learning. To do so we estimated, for each voxel in bilateral caudate nucleus and putamen, functional activation to motor responses, reward and expected value activations during a 2-armed bandit task. These same voxels were also characterized in terms of their structural connectivity to 148 cortical regions using diffusion imaging and probabilistic tractography. We could then predict functional activation from corticostriatal structural connectivity using a leave-one-out cross-validation (LOOCV) procedure (Saygin et al., 2012). All these analyses were performed in subject space, with only summary statistics for each participant taken to the group level. All reported p-values are two-tailed.

2.3. *Task*

The task required participants track stimulus-specific action values and this enabled us to probe how these action values are represented and updated in neural structures during feedback. Participants had to learn two separate two-armed bandits which were distinguished by their color (red or blue; see Fig. 1). On each trial, one of these two slot machines was presented to the participant, and on half the trials a response was required using either right index finger or right ball of the foot on a force-sensitive sensor. Binomial feedback was then presented which indicated a reward or no-reward. The probability of reward given a bandit s and action a, $p(r|s_i,a_j)$ where $i \in \{1,2\}$ and $j \in \{1,2\}$, changed slowly over trials, forcing participants to continue to explore throughout the experiment so as to maximise the total reward obtained.

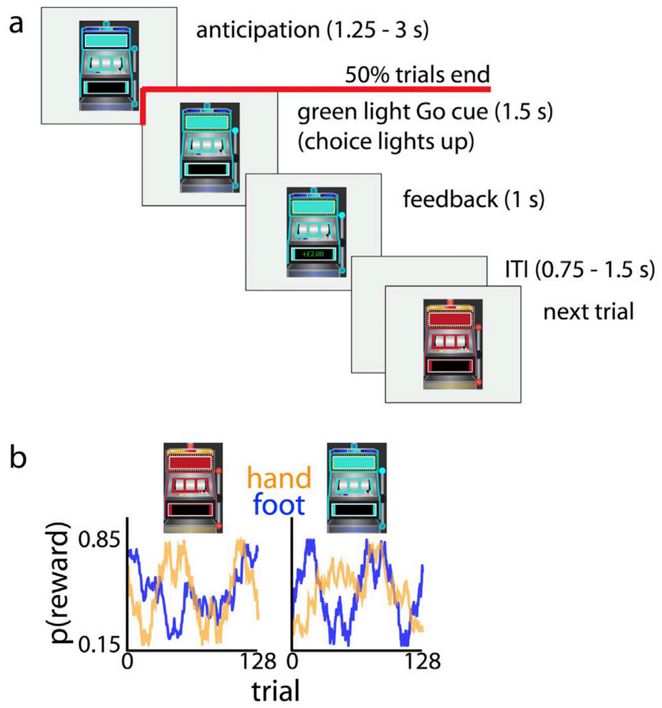

Fig. 1. Reinforcement learning task involving right hand and right foot responses. (a) Task design. On half the trials ('abort' trials) the slot machine disappeared before the Go signal and the next trial started; on the other half ('response' trials) lights on the slot machine would turn green, serving as a Go signal; participants responded by pressing force-sensitive buttons with either their right hand or foot. Feedback was then presented consisting of either "+ £2.00" in green, or "+ £0.00" in red. (b) The probability of obtaining the reward varied over time per response, and per slot machine. This meant participants were required to track 4 random walks that varied between p (reward) of 0.15 and 0.85.

Participants came to the laboratory for a practice session before the scanning session. The interval between practice and scanning session ranged between 1 and 20 days (mean ± SD = 7 ± 4.4 days). At the practice session, participants performed a full set of 512 trials to accustom themselves with the task and force buttons. A different set of reward probabilities was used each day but otherwise the parameters of the experiment were identical. In the experiment proper, participants performed 512 trials (approximately 42 min) consisting of 128 red-abort, red-response, blue-abort, and blue-response trials each (Fig. 1). The order of these four trial types was randomly determined and only constrained such that no trial type occurred for more than 3 trials in a row.

2.3.1. *Reward probabilities*

The $p_t(r|s_i,a_j)$, where t indicates trial number, was generated by a Gaussian random walk for each action a and stimulus s as follows:

$$p_{t+1}(r|s_i,a_j) = p_t(r|s_i,a_j) + N(0,0.01)$$

where for the first trial the probability was randomly drawn from *U(0.15,0.85)*. The walks were not generated anew for each participant—rather, one set of two pairs was used for each participant's practice, and one set was used for each participant's scanning session. However, the assignment of these two pairs to the red and blue slot machine was randomized, and the subsequent assignment of random walk to the two available actions was also randomized. This meant that volatility and availability of reward were matched between participants. The walks were constrained in their upper (0.85) and lower (0.15) values and in their mean value (between 0.4 and 0.6). The highest correlation between any two of the four walks was 0.38, forcing participants to learn about the value of each option through trial-and-error rather than inferring the value of one option based on changes in the other.

2.3.2. *Trial design*

Examining value representations in the BOLD signal at both choice and outcome phase is challenging due to the sluggishness of the BOLD response. We considered two trial designs to alleviate this issue: a slow design where choice and feedback events are separated by at least 8 s (Behrens et al., 2008), and a fast design in which half the trials are cancelled at any point between choice and feedback phase (Guitart-Masip et al., 2012). Pilot data with both designs (data not shown) suggested participants were more accurate at learning reward probabilities in the fast design, possibly due to task disengagement when participants are faced with long pauses. Also, a slow design might lead to non-striatal learning mechanisms dominating behavior, whereas we were specifically interested in such striatal mechanisms (Foerde et al., 2012). We thus opted for the fast design. In this paper we do not report correlates of action values during choice as we were unable to reliably observe its neural correlates in the striatal regions; we only examine motor responses, expected value at outcome and reward responses.

2.4. *Reinforcement learning models*

We used temporal difference (TD) reinforcement learning models to model participants' behavior and estimate quantities that might be represented in the BOLD signal in the striatum, most notably rewards and action values. Each slot machine i defines a state s_i where two actions a_j are available. The reward r on trial t can be either 0 or 1. The value of action j in state i is updated after feedback by: