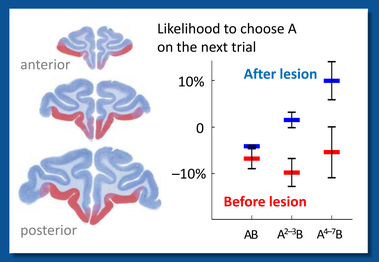
Decoding choice from prefrontal cortex during reinforcement learning

Kay H Brodersen[1] · Laurence T Hunt[1] · Mark E Walton[2] · Matthew F S Rushworth[1,2] · Timothy E J Behrens[1,2]

[1] Oxford Centre for fMRI of the Brain, University of Oxford [2] Department of Experimental Psychology, University of Oxford {kbroders,lhunt,behrens}@fmrib.ox.ac.uk, mark.walton,matthew.rushworth}psy.ox.ac.uk

1 Introduction

At the core of reinforcement learning lies the ability of the brain to form associations between choices and outcomes. Different regions of the prefrontal cortex (PFC) have been hypothesized to play different roles in this process. For example, neurons in the orbitofrontal cortex (OFC) are known to encode information about the sensory properties of a chosen stimulus at the time when a decision is made (Rolls et al. 2003). Consistent with this notion, lesions to the OFC prevent choices from being associated with specific types of outcome (Murray et al. 2007). In addition, recent OFC lesion data in macaques demonstrate that the OFC is essential in associating an outcome with the specific choice immediately preceding it.

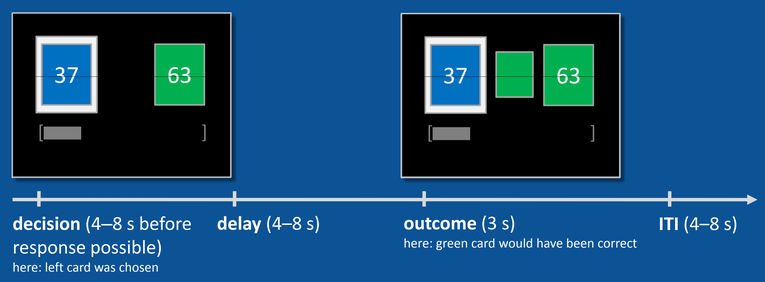


After a sequence of A choices and a single rewarded B choice, rather than trying B again, lesioned monkeys were more likely to go for A instead. They were associating the reward on B with their history of A choices. The longer the sequence of A choices, the stronger the effect (Walton et al., in preparation).

It might be expected that regions in PFC that encode choice during the decision also code for choice when the outcome is witnessed, in order that a particular outcome may reinforce the choice that caused it. We therefore investigated whether different regions of PFC contained information about choice, either at the time of the decision or at the time of the outcome.

2 Experimental paradigm

In an event-related probabilistic decision-making paradigm, participants had to choose, on each trial, between a blue and a green card. The two cards had different reward probabilities which had to be learned over time and which varied during the course of the experiment. Potential rewards were displayed on the cards. Left/right positions were randomized. Choices were indicated by pressing buttons with the left/right index finger.



3 Multivariate pattern analysis

Data from 16 subjects x 120 trials were acquired on a 3T fMRI scanner. The scanning sequence was optimized for signals in the OFC, with field-map correction and a voxel size of 2 x 2 x 2 mm^3 (TR 3s).

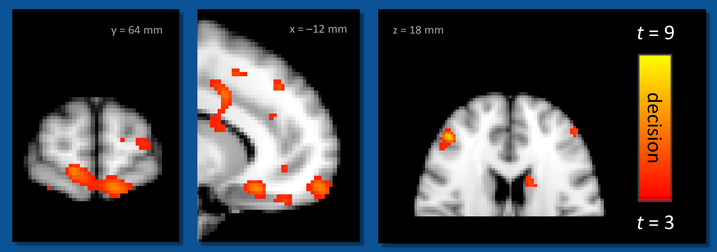
We designed a classification algorithm to decode chosen stimuli from single-trial brain activity. The classifier operated on Beta maps that were derived separately for the decision phase and the outcome phase of each trial by deconvolving the measured BOLD response with a canonical haemodynamic response function. Based on a whole-brain leave-one-trial-out cross-validation scheme, we used a Gaussian Naïve Bayes classifier to assess the discriminative information contained in local searchlights (r = 2 voxels) centered on each voxel in turn (Kriegeskorte et al. 2006). In each cross-validation fold, the training set was balanced. The two trials immediately preceding and following a test trial were excluded from the training set.

In standard space, using nonlinear registration, we assessed at which voxels the classifier performed better than chance (i.e., above 50% accuracy). In this way, we obtained maps of t scores at the group level.

4 PFC regions code for choice during a decision

In order to investigate regions encoding choice at the time of the decision, we trained and tested our classifier on brain activity during the decision phase of each trial.

Unsurprisingly, many regions in PFC were found to contain sufficient discriminative information (as were many other regions in the brain not shown here). In particular, regions coding for choice during the decision included OFC, medial polar cortex, anterior cingulate cortex (ACC), and dorsolateral PFC.



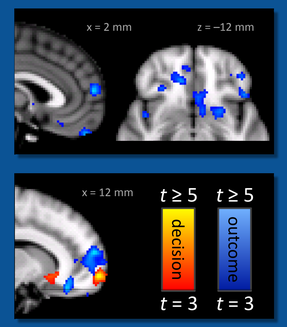
References

- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Nature Neuroscience, 10, 1214-1221.
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). PNAS, 103(10), 3863-3868.
- Rolls, E. T., Verhagen J. V., Kadohisa, M. (2003). Journal of Neurophysiology, 90(6), 3711-3724.
- Walton, M. E., Behrens, T. E. J., Rudebeck, P. R., Buckley, M. J., Rushworth, M. F. S. (in preparation).

This research was funded by UK MRC and the Wellcome Trust.

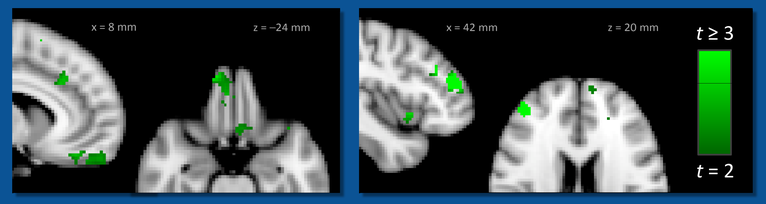
5 Some prefrontal regions also code for choice at the time of the outcome

Crucially for learning, prefrontal regions also coded for the chosen stimulus after the decision, when the outcome was observed. These areas are much more focused than those encoding choice at the time of the decision. Within PFC, they are predominantly in the OFC and medial polar cortex. Notably, within the medial polar cortex, a more dorsal region coded choice at the time of the outcome, and a more ventral region at the time of the decision.



6 DLPFC and OFC code for choice at both stages; their activity patterns generalize across stages

In order to test whether any prefrontal regions encode choice in the same way both at the time of the decision and at the time of the outcome, we trained our classifier on the decide phase and tested it on the outcome phase. The resulting generalization map overlapped with the two previous maps in dorsolateral prefrontal and orbitofrontal cortex.



7 Conclusions

Using multivariate decoding, we found prefrontal regions to encode a chosen stimulus not only at the time of the decision but also at the time of the outcome. In addition, activity patterns in DLPFC and some parts of the OFC appear to generalize across the stages. This suggests that prefrontal regions play a key role in reinforcement learning by means of a distributed encoding of chosen stimuli. The fact that, in particular, widespread regions in OFC encode choice at the time of the outcome, potentially explains the deficits in learning observed as a result of orbitofrontal lesions.