Neural Mechanisms of Hierarchical Planning during Navigation

Balaguer J[1], Hassabis D[2], Spiers H[2], Summerfield C[1]

1 Dept. Experimental Psychology, University of Oxford, Oxford, UK.
2 Dept. Psychology, University College London, London, UK.

Correspondence : jdobalaguer@gmail.com

1 Introduction

Reinforcement Learning (RL) is a general framework of Machine Learning that also describes mammalian behaviour [1].
Hierarchical Reinforcement Learning (HRL) [2] is an extension of RL for temporal abstraction [3,4].
To achieve this, HRL exploits the concept of **options** (or contexts), each defined by its **subgoal state**.

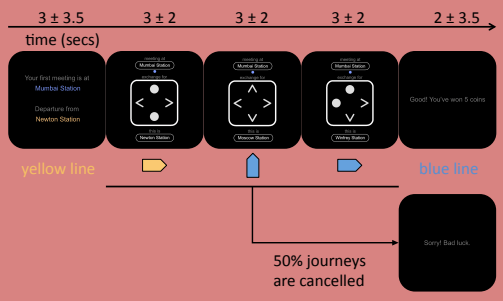
Prediction error signals for interim goals have been observed in the striatum [5], but it remains unknown whether **representations of context switching** exist in the human brain.

We predicted that subgoal states would be associated with activation in the **dorsal Anterior Cingulate Cortex (dACC)**, a region previously associated with task-switching [6] and tracking the reward value of alternative contexts [7].

Methods

N = 19 participants (age 00±00, 00 males) performed a navigation task while undergoing **fMRI**. There was a preliminary **training session**.

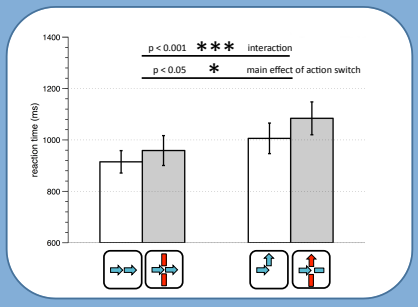
2 Design



Participants navigated in a **virtual subway network** (stations and lines) with which they were familiar. The instruction was to minimise the length of their journey. The **map was not shown** during navigation.
Possible directions (actions) were North/South/East/West.
Regular, **exchange** and **elbow** stations allowed us to disentangle between effects due to number of directions, direction switch and line switch.

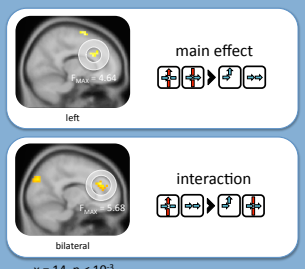
3 Results

Behavioural



Neural

General Linear Model of BOLD signal



dACC was significantly more activated in **exchange stations** than in regular stations.

Critically, this activation was strongest when **switching line**.

x = 14, $p < 10^{-3}$

4 Conclusions

This results provide evidence that a unique **network is activated when participants reach an interim goal** during navigation. This network bears a striking resemblance to that previously implicated in response inhibition during speeded decisions[8]. **The dACC**, which forms a prominent part of this network, may be key for **switching to a new context** in pursuit of reward[7].
Our findings suggest that human dACC and/or interconnected structures signal when a subgoal has been reached.

References

1. **Sutton, Barto (1998)**, *Reinforcement Learning: An Introduction*, The MIT Press
2. **Sutton, Precup, Singh (1999)**, *Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning*, Artificial Intelligence
3. **Schapiro et al. (2013)** *Neural representations of events arise from temporal community structure*, Nature Neuroscience
4. **Botvinick, Niv, Barto (2009)**, *Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective*, Cognition
5. **Ribas-Fernandes et al. (2011)** *A Neural Signature of Hierarchical Reinforcement Learning*, Neuron
6. **Hyafil, Summerfield, Koechlin (2009)**, *Two mechanisms for task switching in the prefrontal cortex*, Journal of Neuroscience
7. **Kolling et al. (2012)** *Neural mechanisms of foraging*, Science
8. **Aron et al. (2007)** *Triangulating a Cognitive Control Network Using Diffusion-Weighted Magnetic Resonance Imaging (MRI) and Functional MRI*