Orbitofrontal cortex distributes reinforcement to the decision that caused it

Kay H Brodersen[1,2,3], Laurence T Hunt[3], Ekaterina I Lomakina[1,2], Matthew F S Rushworth[3], Timothy E J Behrens[3,4]

[1] Department of Computer Science, ETH Zurich, Switzerland
[2] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland
[3] Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, United Kingdom
[4] Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom

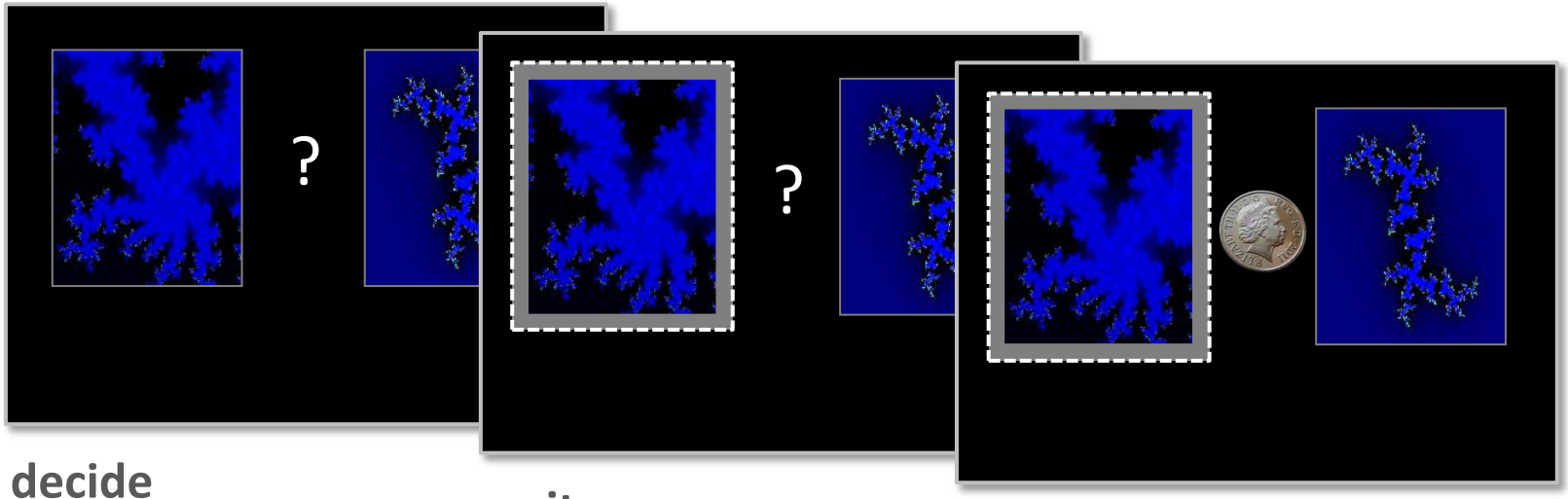
How do we distribute reinforcement to the correct decision?

- Reward-maximizing behaviour is based on the ability to associate an observed outcome with the decision that caused it.
- A lesion to the lateral orbitofrontal cortex (LOFC) keeps reward processing intact. But it disrupts the ability to correctly associate rewards with preceding decisions.

  Walton et al., 2010, *Neuron*
- Using fMRI in humans, we set out to explain this effect by examining the role of the LOFC in reinforcement learning.

Experimental design

We examined the role of the LOFC in reversal learning using a simple decision-making task. Subjects had to learn, by trial and error, the reward probabilities of two options.



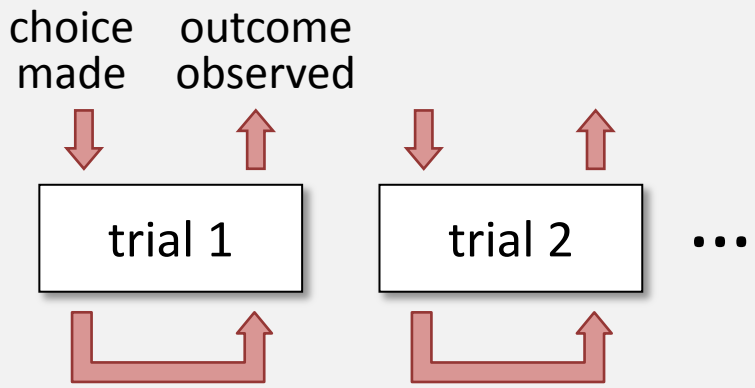
decide
(delay until response)

wait
(jittered, mean 5 s)

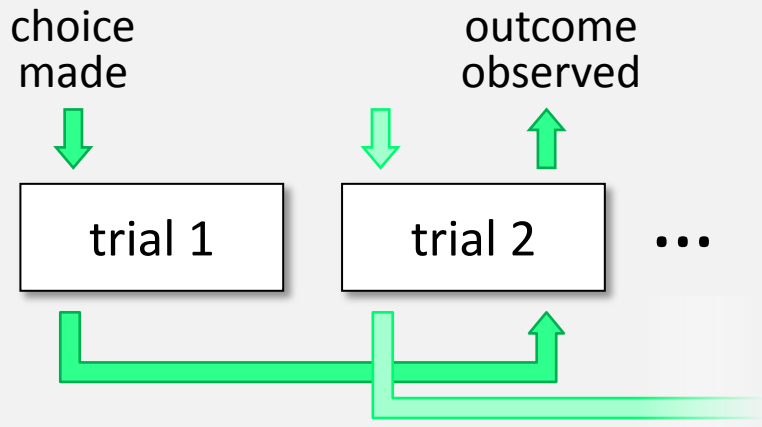
outcome
(fixed delay, 1 s)

Experimental design

The experiment consisted of 120 trials, grouped into 8 blocks. Each block followed one of the following two instruction types:



In a **DIRECT** block, whenever a rewarded option was chosen, its reward was presented on the *same* trial.

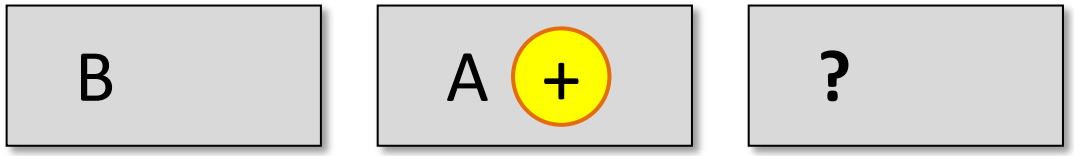


In an **NBACK** block, rewards were delayed by one trial such that a reward had to be linked to the choice made on the *previous* trial.

Imaging analysis

To understand the neural basis of association learning, which events should we look at?

Example:



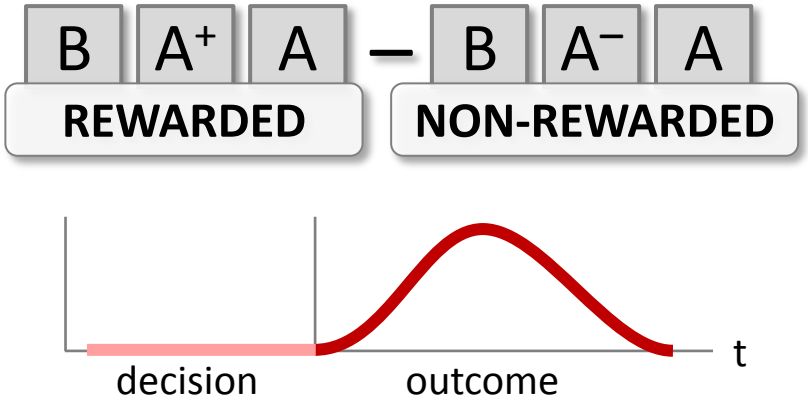
DIRECT condition: should stick with A!
NBACK condition: should switch back to B!

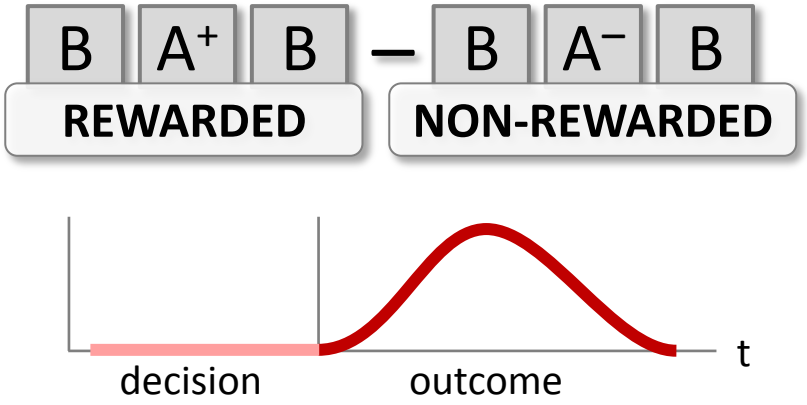
critical timepoint for fMRI analysis

Imaging hypotheses

1 Does the LOFC encode rewards?

DIRECT trials



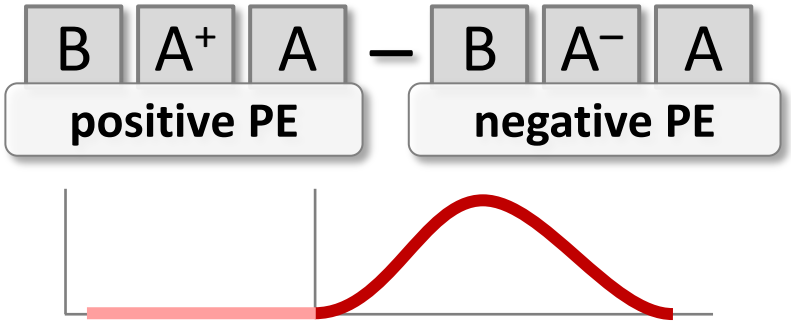


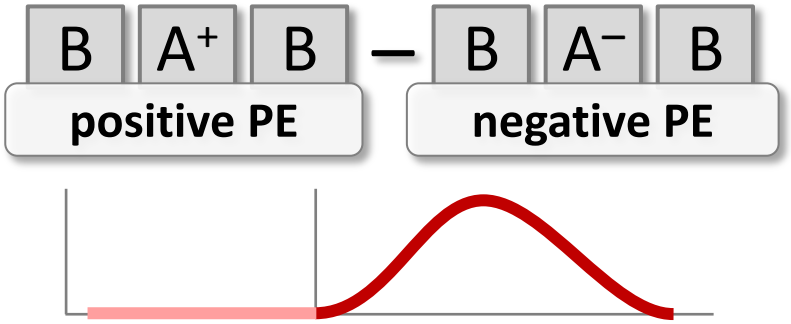
Imaging hypotheses

1 Does the LOFC encode rewards?

2 Does the LOFC encode reward prediction errors?

■ **DIRECT** trials



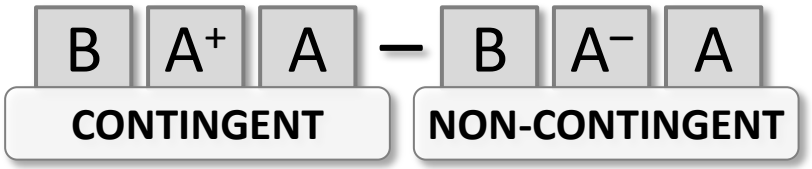


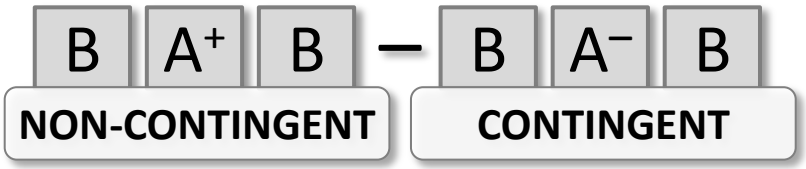
Imaging hypotheses

1 Does the LOFC encode rewards?

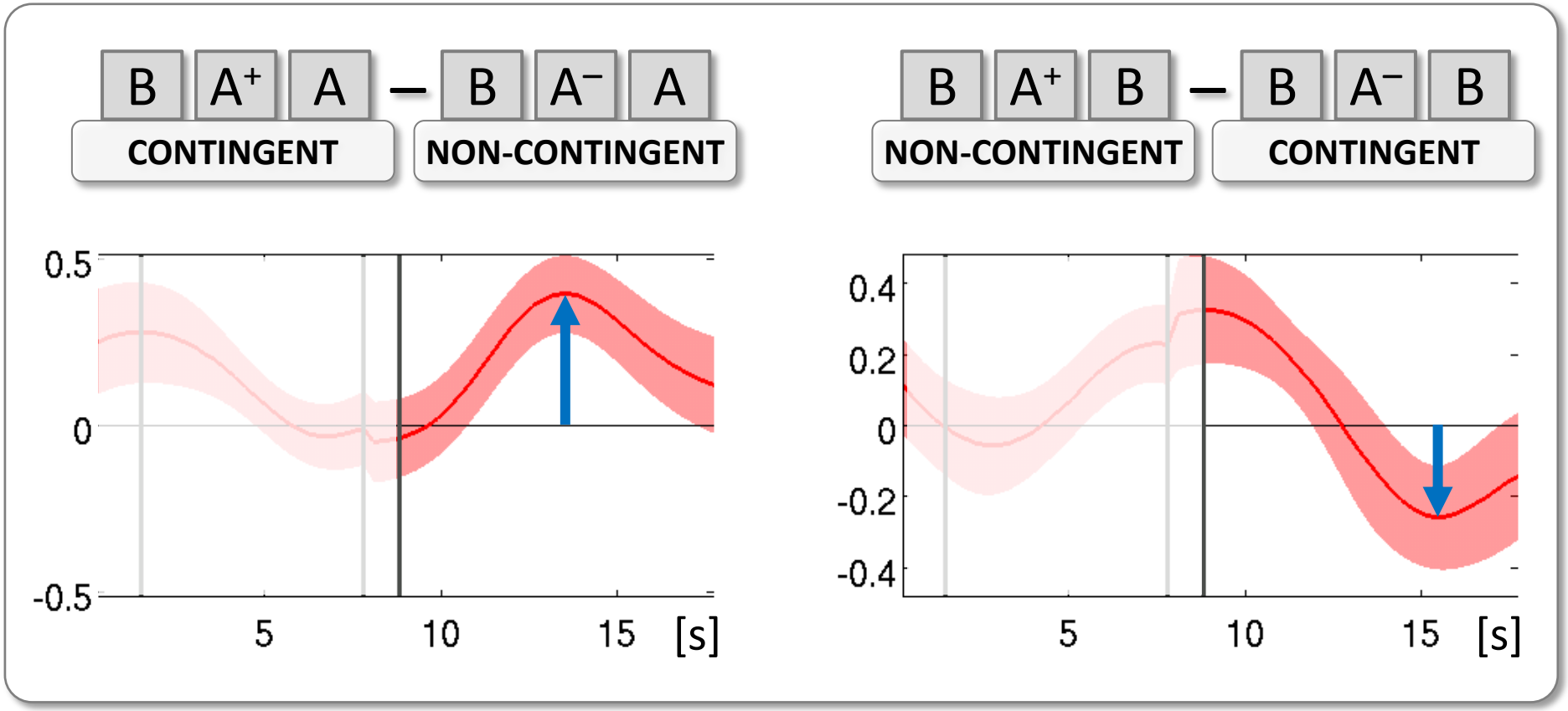
2 Does the LOFC encode reward prediction errors?

3 Does the LOFC encode correct contingencies?

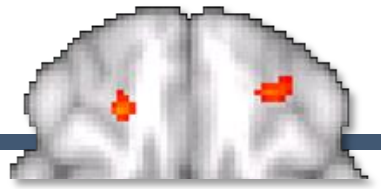




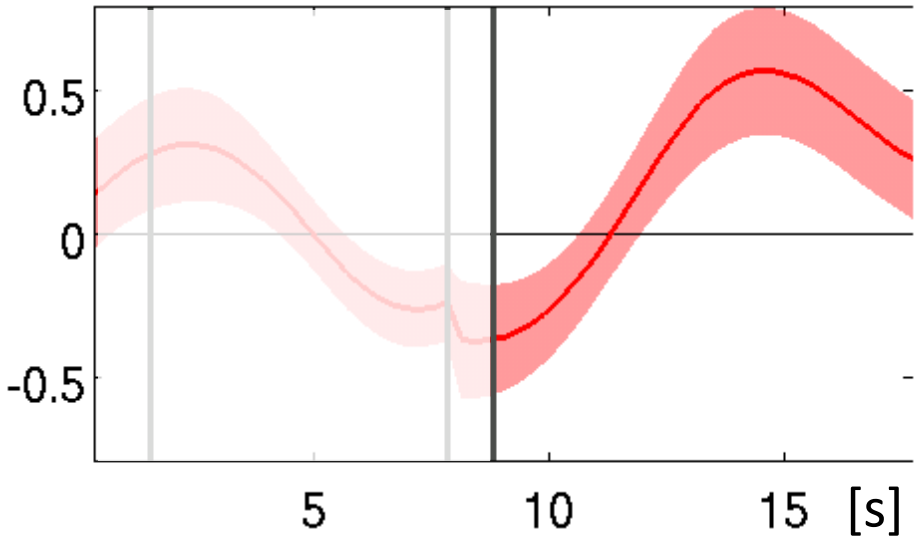
Imaging results



Imaging results

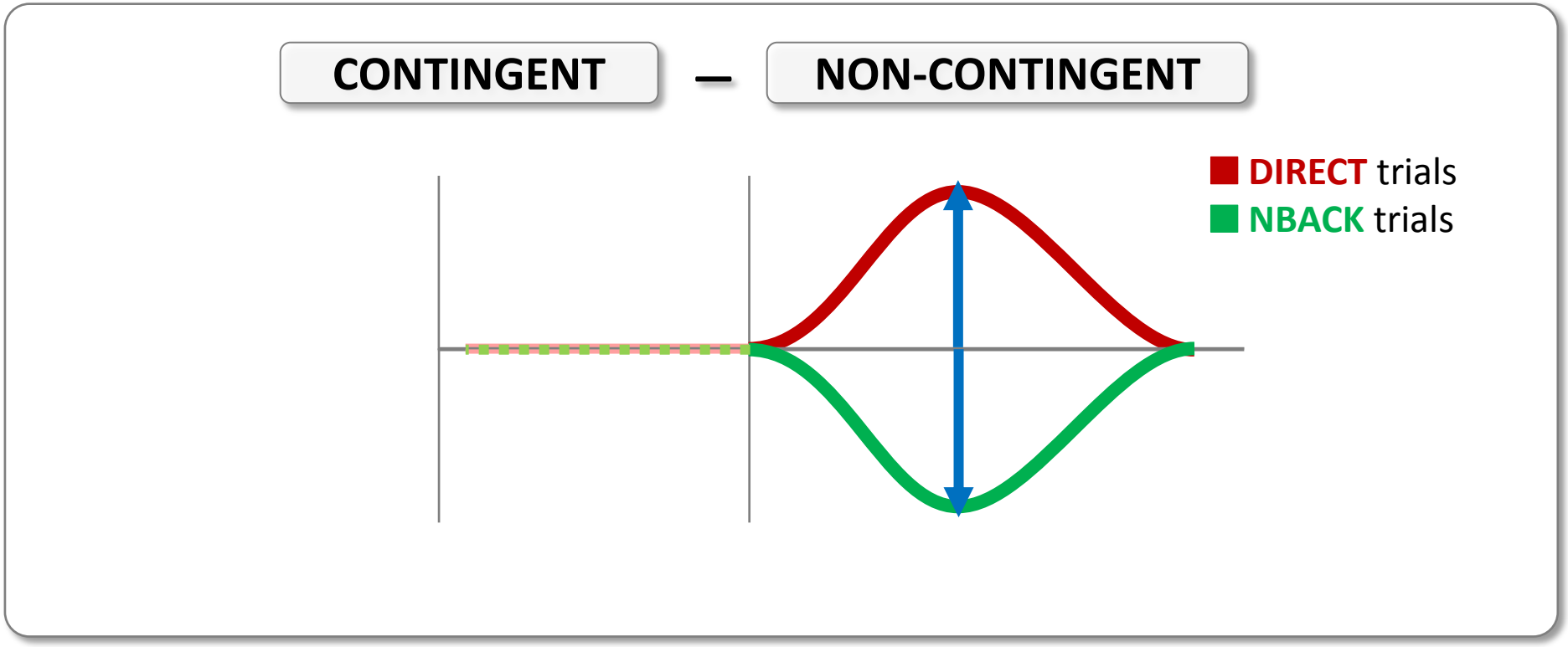


CONTINGENT — **NON-CONTINGENT**

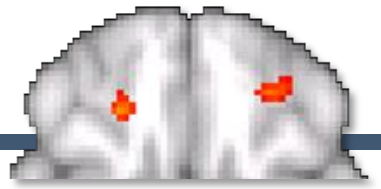


DIRECT trials

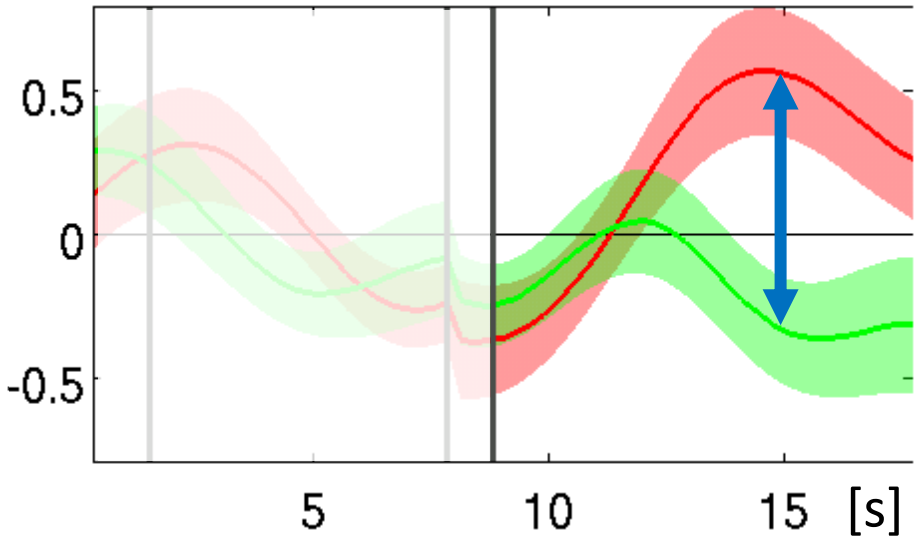
Imaging hypotheses



Imaging results



CONTINGENT — **NON-CONTINGENT**



DIRECT trials
NBACK trials

Hypothesis 1: ~~LOFC encodes reward.~~
Hypothesis 2: ~~LOFC encodes reward prediction errors.~~
Hypothesis 3: LOFC encodes the application of correct contingencies.

Summary

1. A lesion to the LOFC impairs association learning in monkeys. We set out to explain this effect using fMRI in humans.

2. We found that LOFC is neither simply driven by rewards and losses, nor by reward prediction errors, nor by switches and stays.

3. Rather, activity in LOFC specifically encodes whether a correct contingency is being applied. Contingencies are indicated irrespectively of which previous choice should be reinforced.