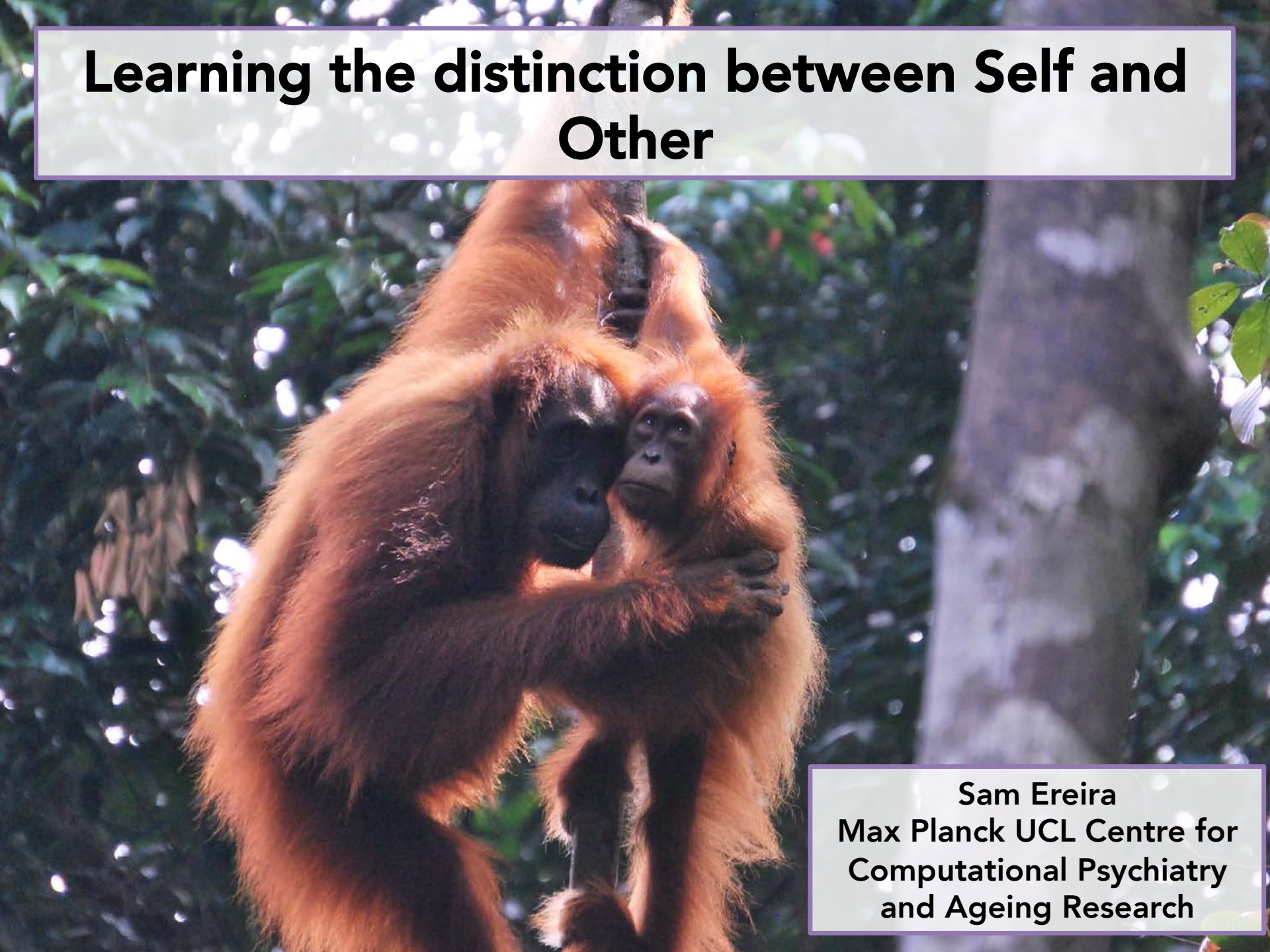
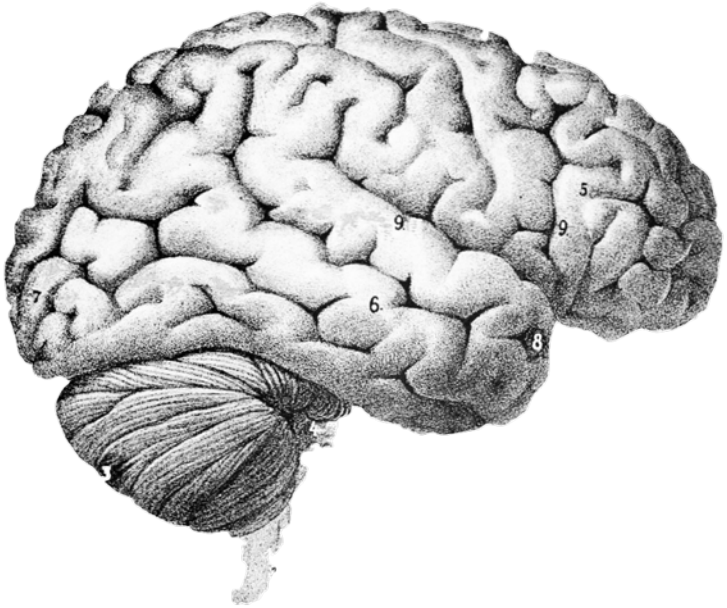
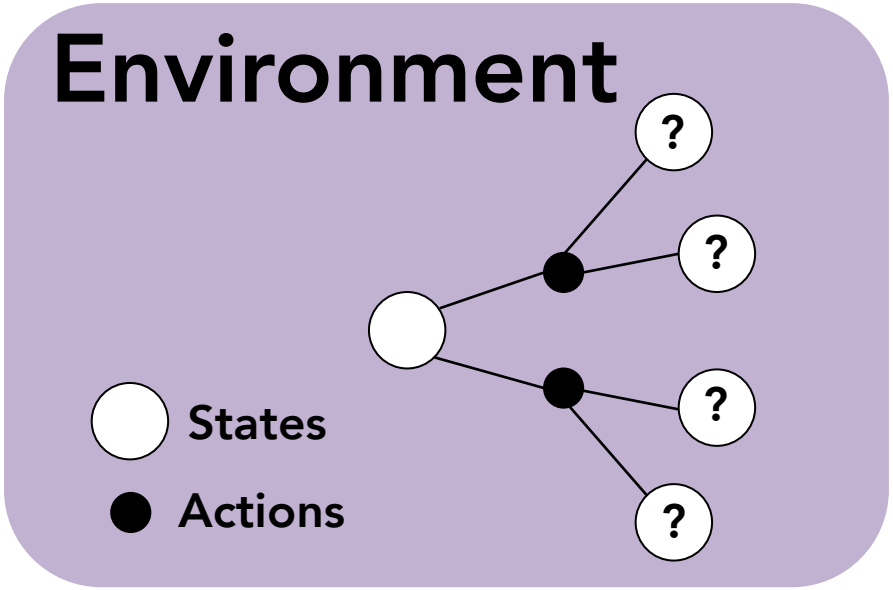
Learning the distinction between Self and Other



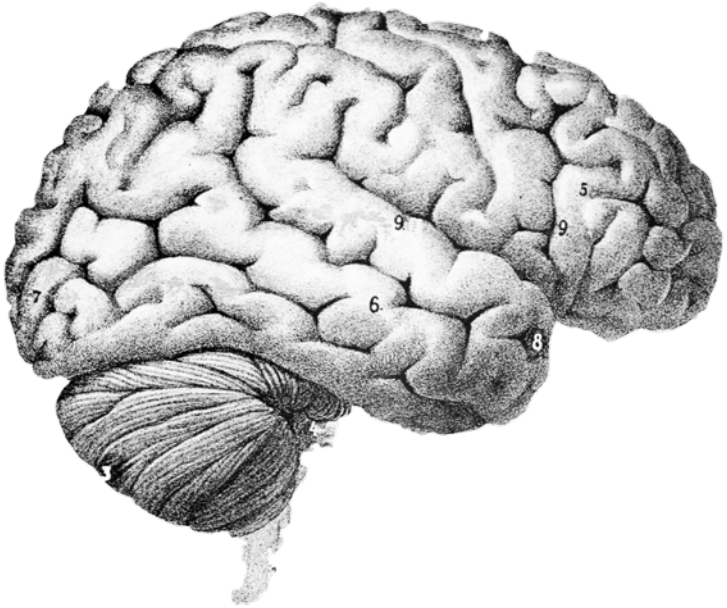
Sam Ereira
Max Planck UCL Centre for Computational Psychiatry and Ageing Research

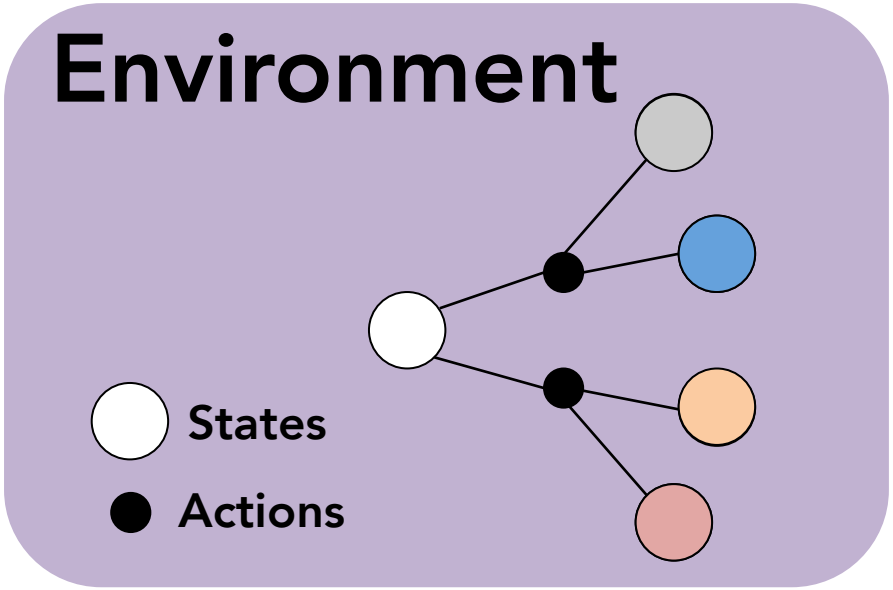
Levels of abstraction in the brain



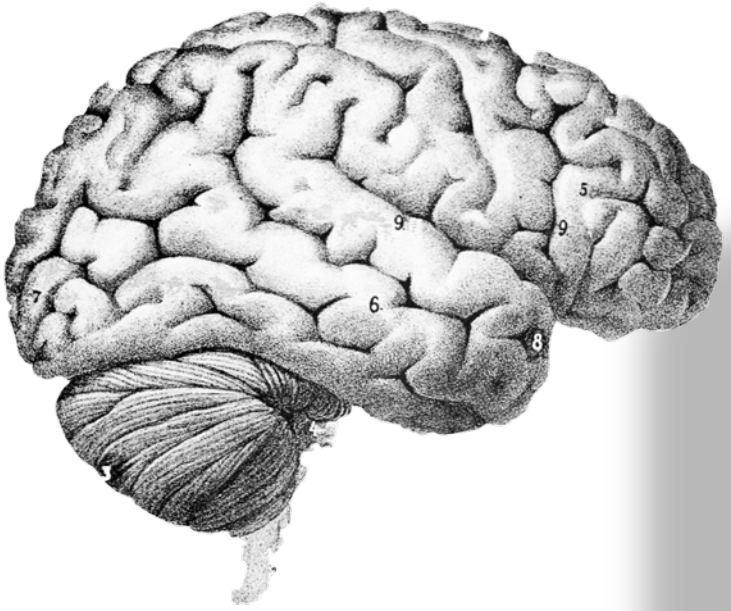


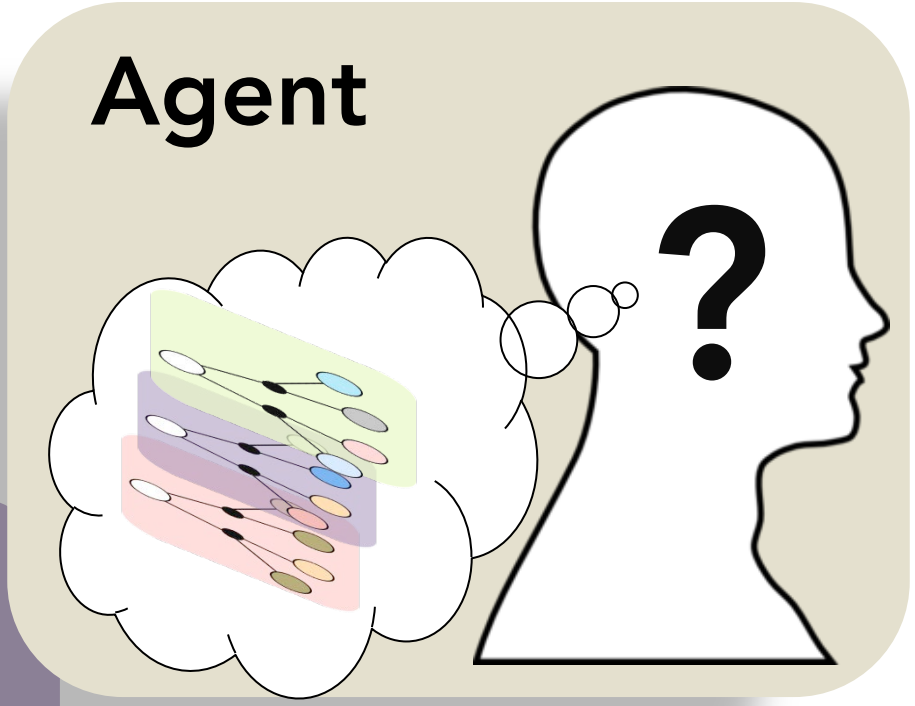
Levels of abstraction in the brain

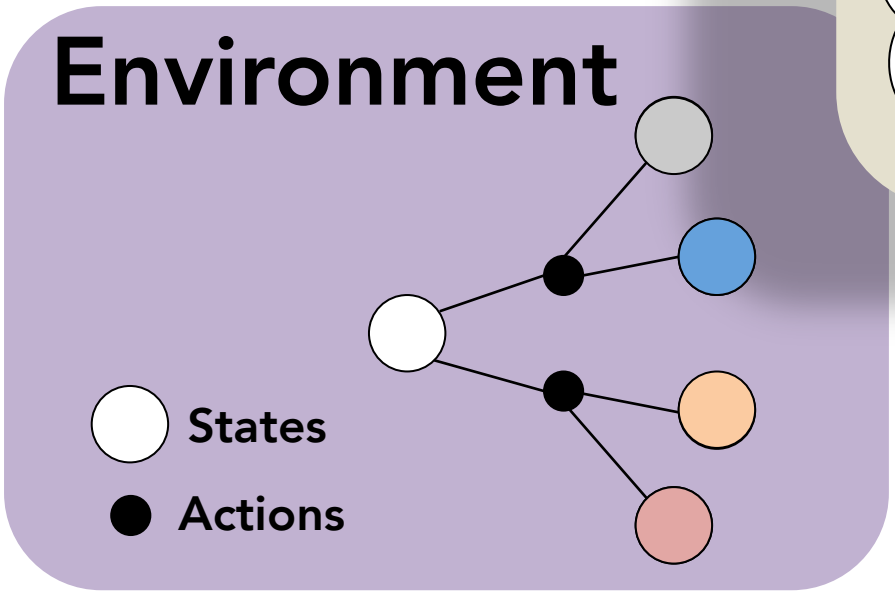




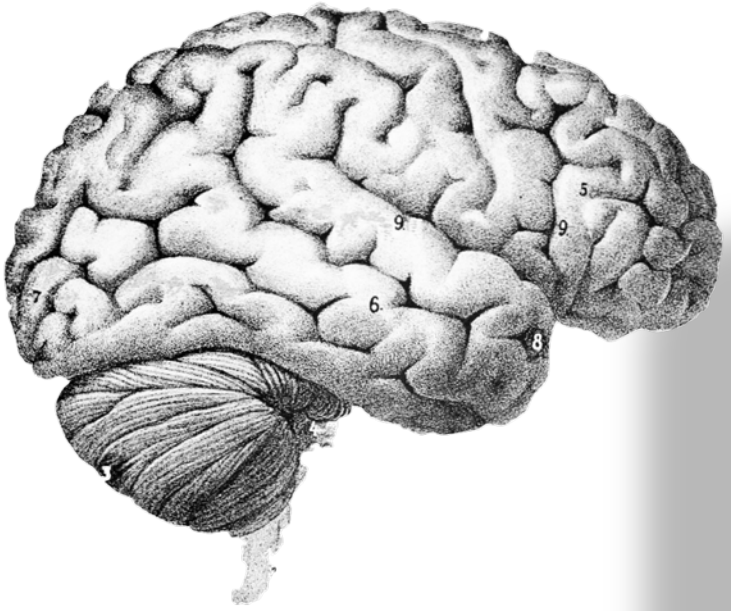
Levels of abstraction in the brain



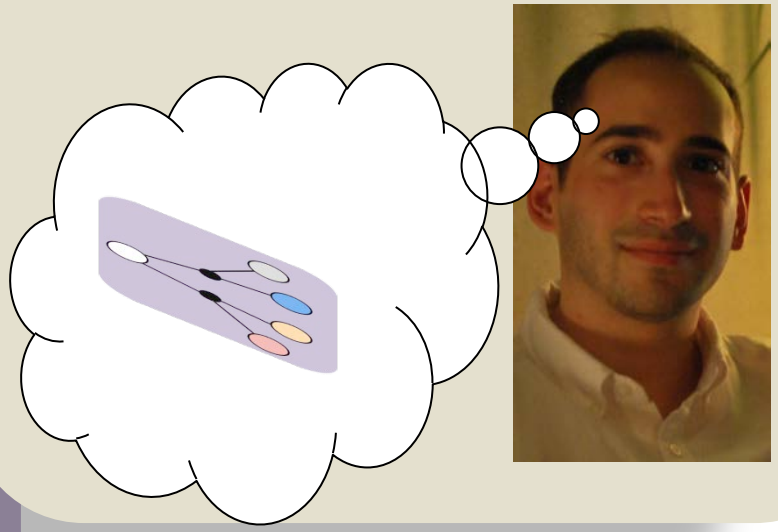




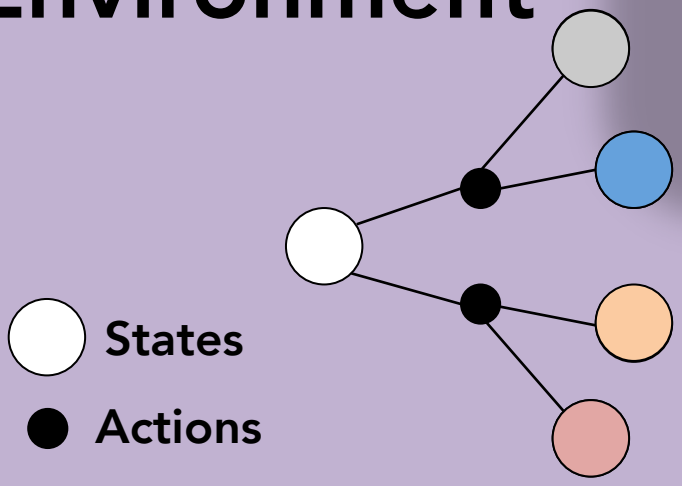
Levels of abstraction in the brain



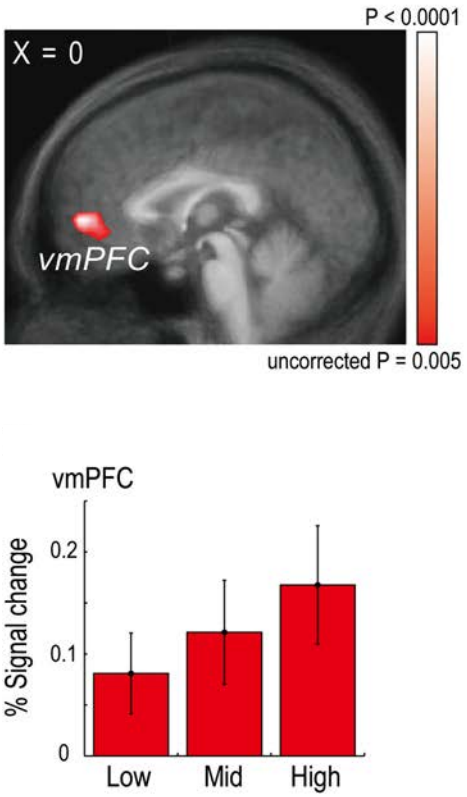
Agent



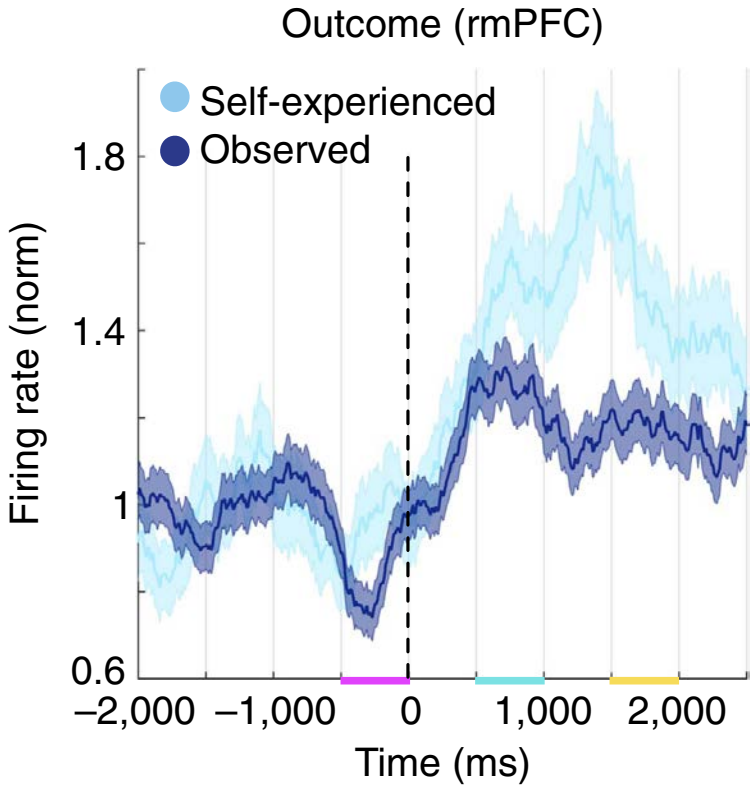
Environment



My brain can simulate your computations…

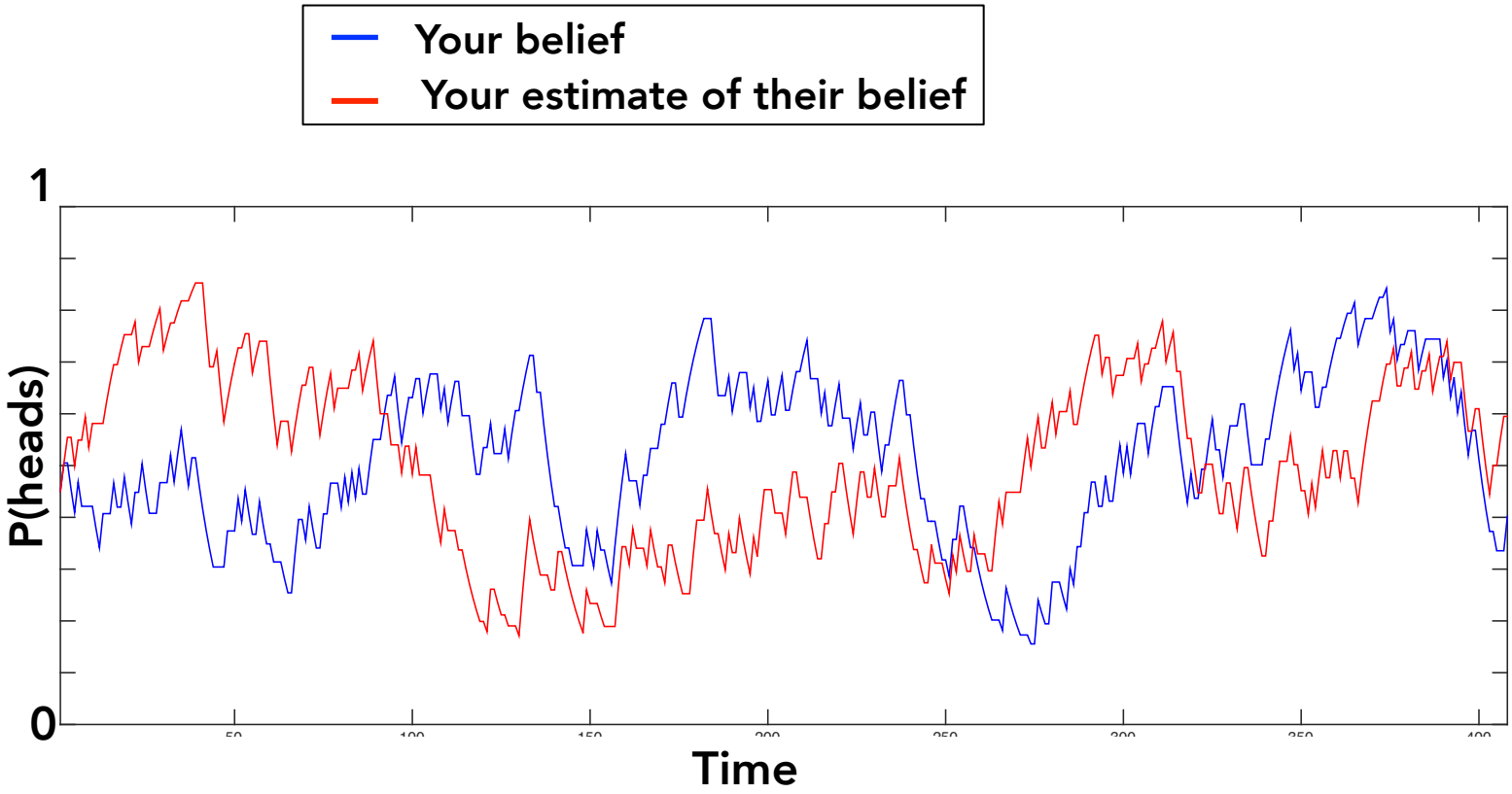


Suzuki et al. Neuron 2012



Hill et al. Nature Communications 2016

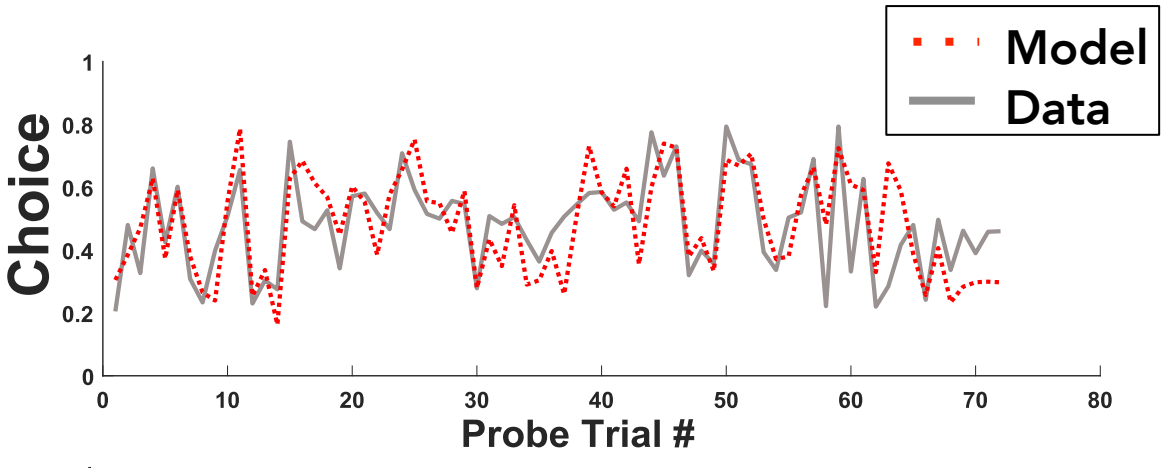
Paradigm: A probabilistic false-belief task



Participants must keep track of their own beliefs AND the other person's beliefs (probed every 4-9 coin flips)

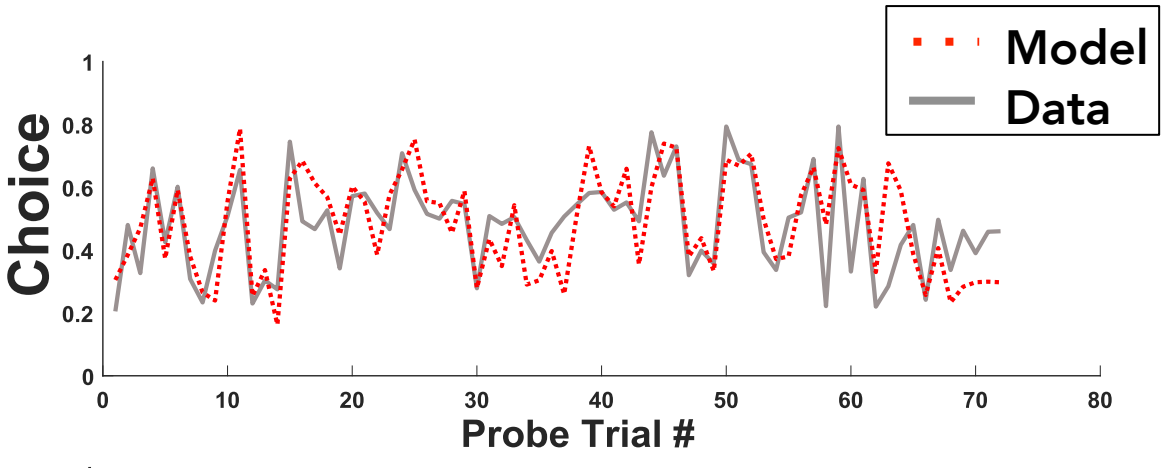
Ereira et al. *PLoS Biol, 2018*

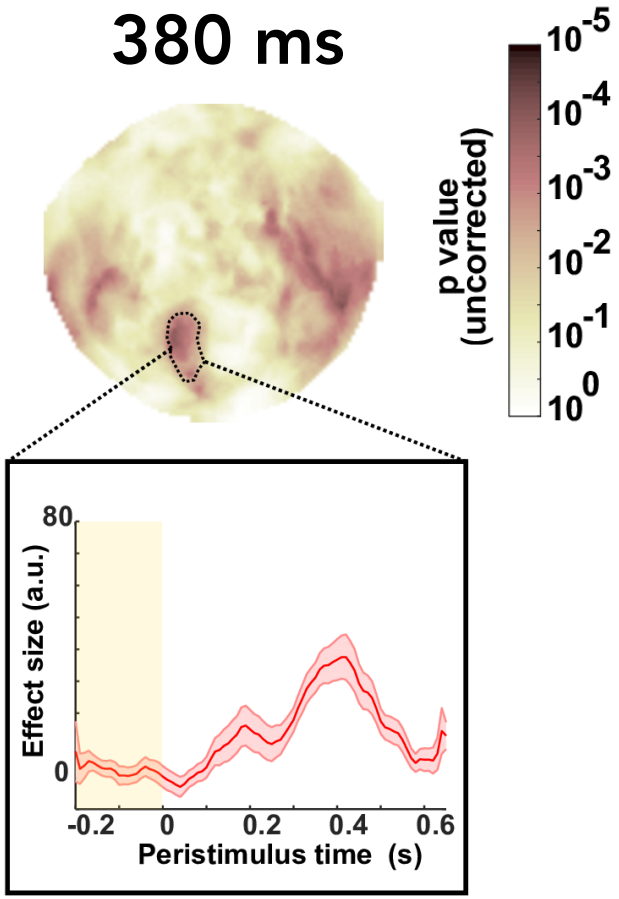
Behaviour and neural activity can be explained with simulated PEs



Ereira et al. *PLoS Biol, 2018*

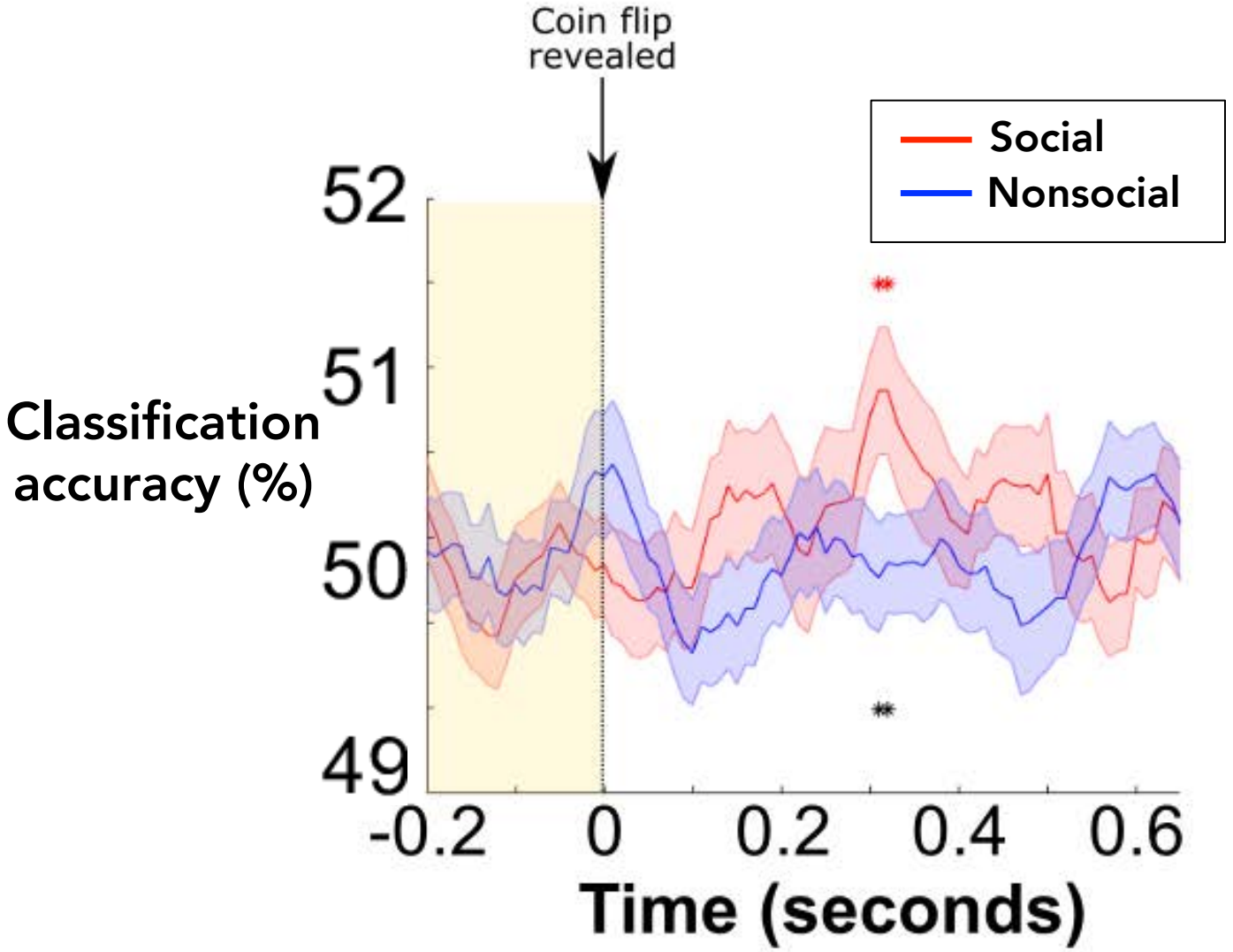
Behaviour and neural activity can be explained with simulated PEs





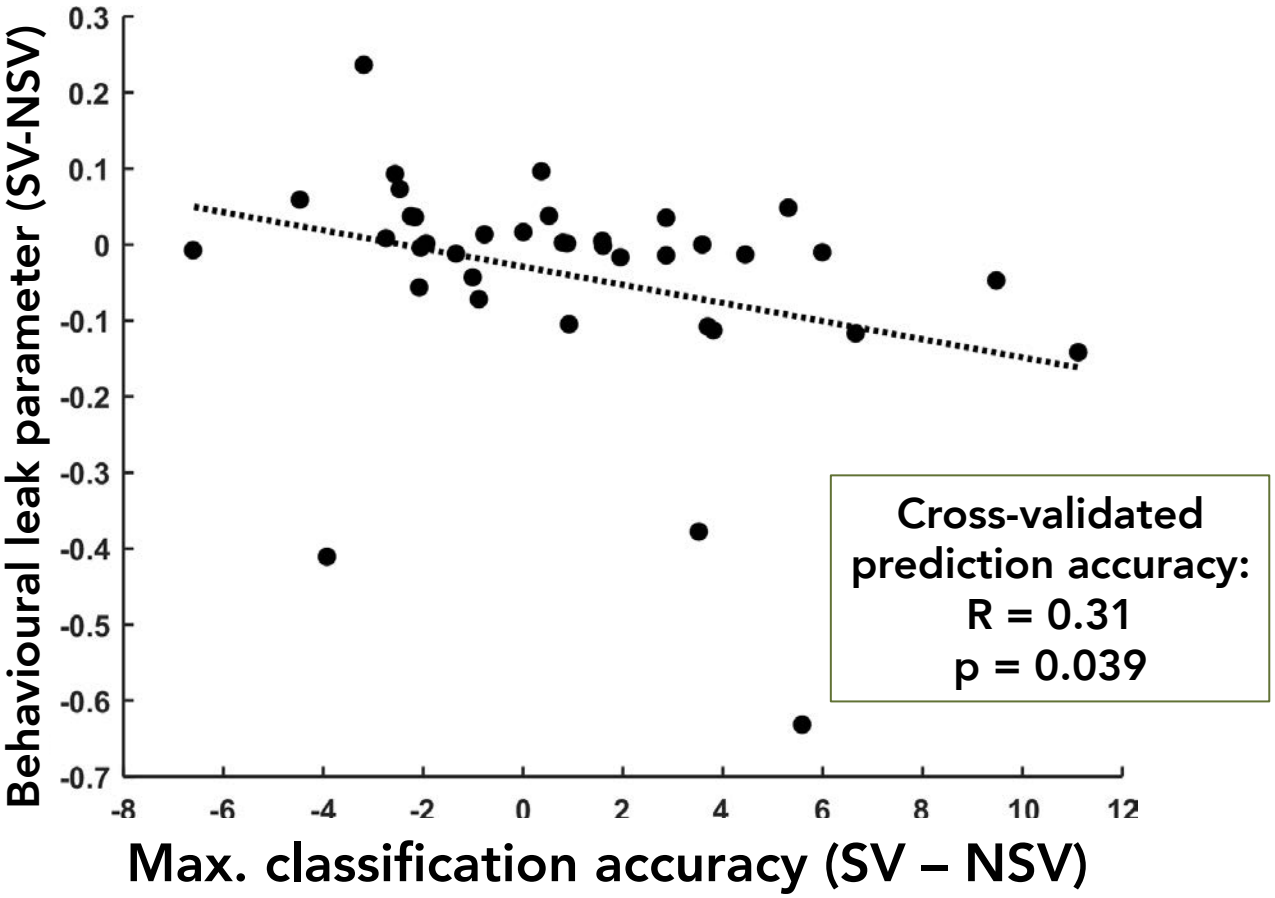
Ereira et al. *PLoS Biol, 2018*

Agent identity is decodable from low level learning signals



Ereira et al. *PLoS Biol, 2018*

Agent decodability predicts behavioural self-other distinction



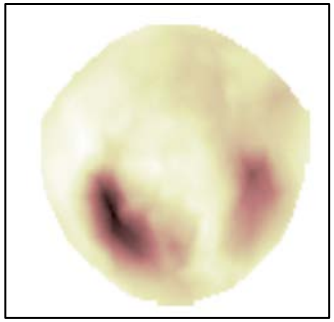
Ereira et al. *PLoS Biol, 2018*

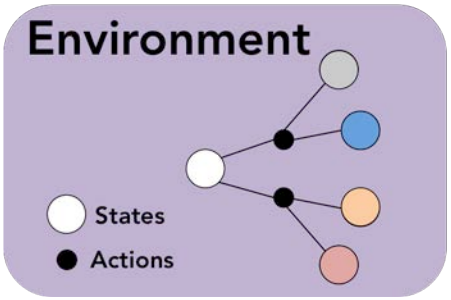
Interim summary

Self-other distinction can be achieved by expressing low-level learning signals in agent-specific neural activity patterns

Interim summary

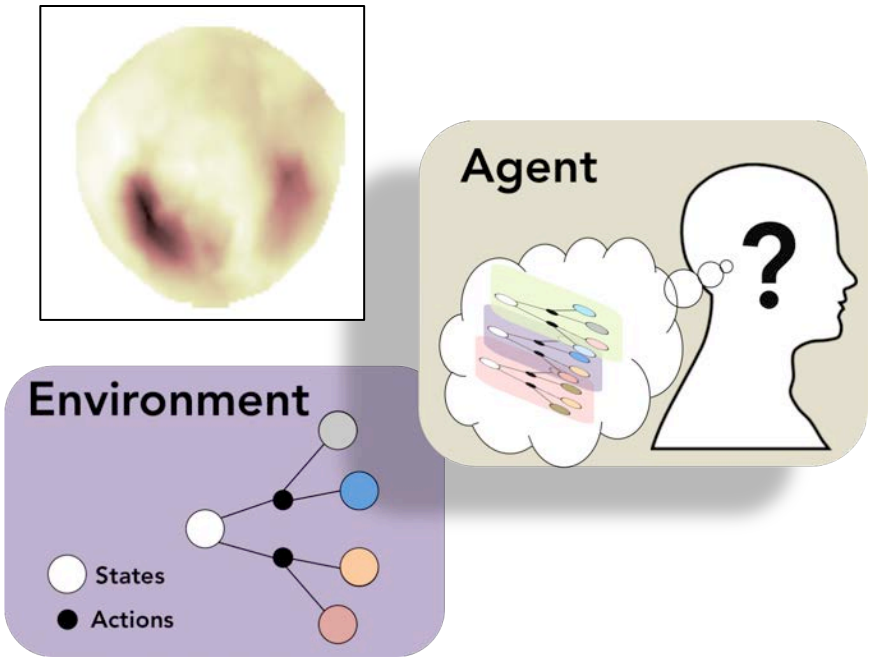
Self-other distinction can be achieved by expressing low-level learning signals in agent-specific neural activity patterns





Interim summary

Self-other distinction can be achieved by expressing low-level learning signals in agent-specific neural activity patterns



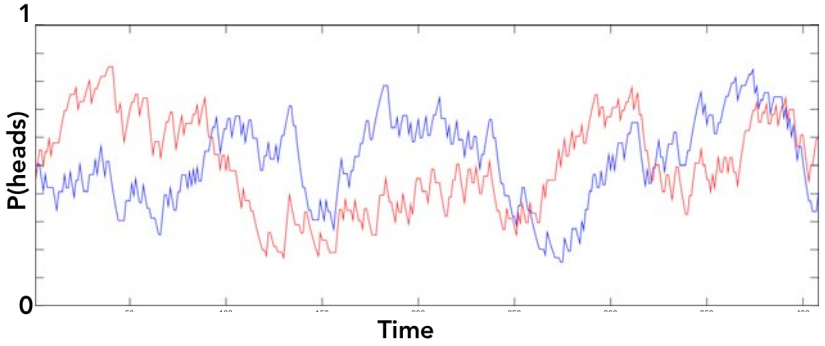
Getting a handle on self-other distinction

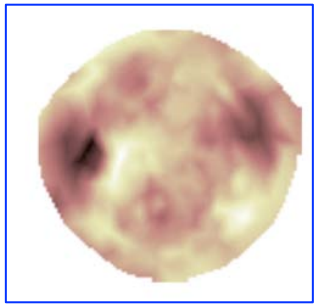
?

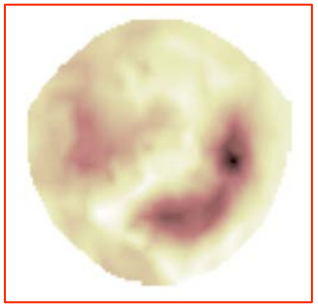
Self-other distinction

Behaviour

Brain





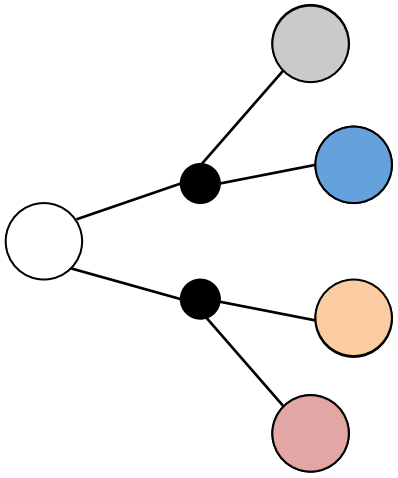


A new hypothesis

The degree of self-other distinction is learned

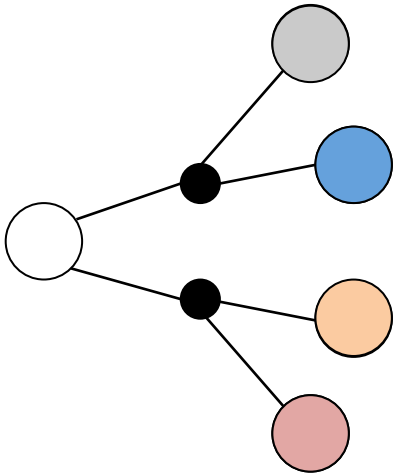
A new hypothesis

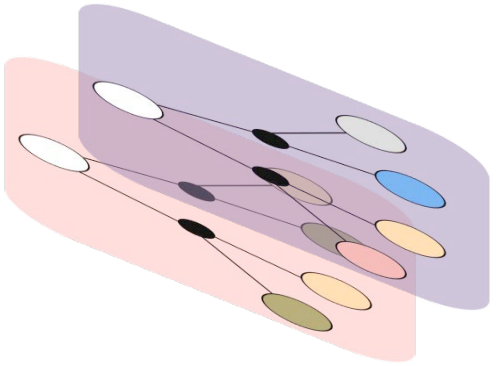
The degree of self-other distinction is learned



A new hypothesis

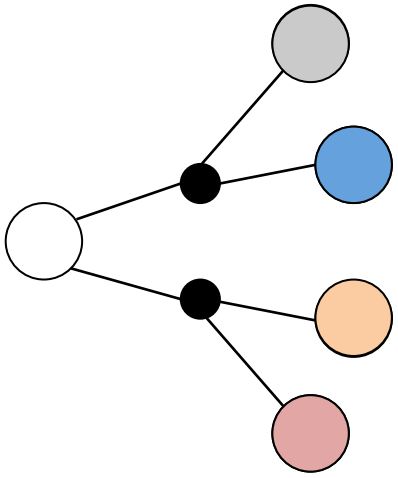
The degree of self-other distinction is learned

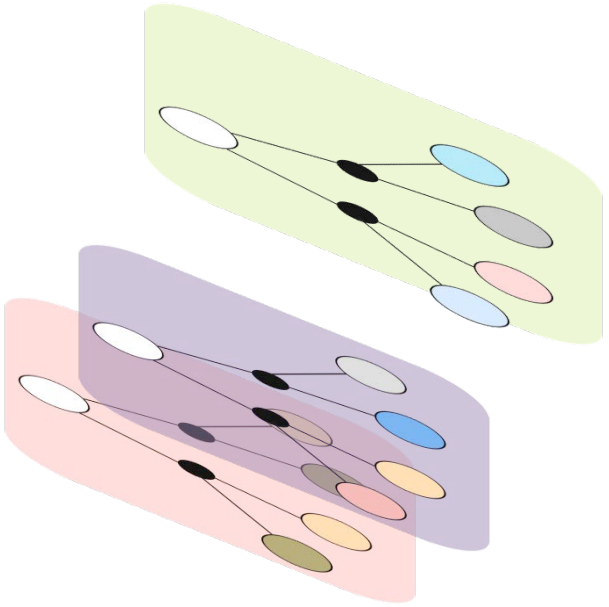




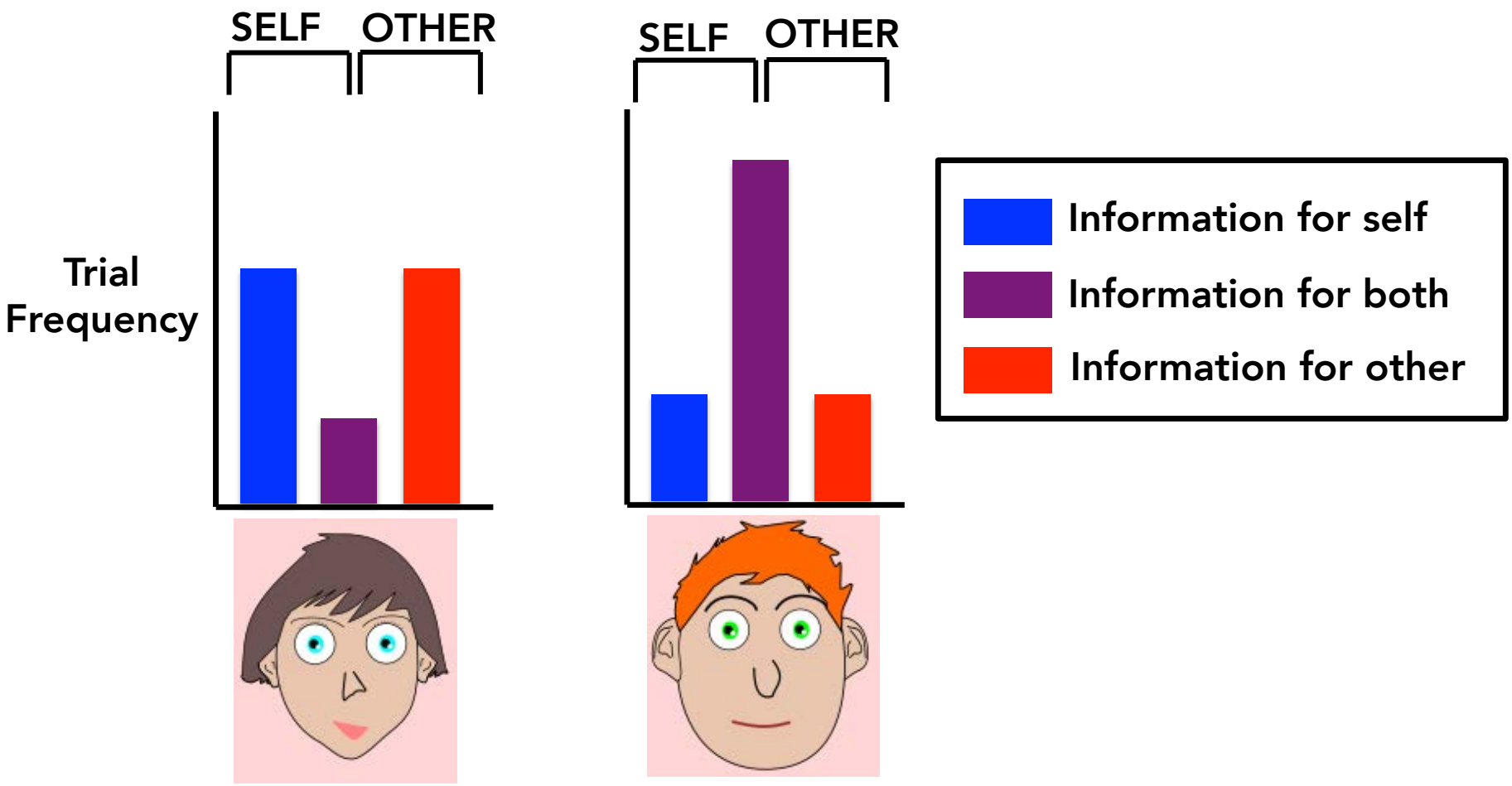
A new hypothesis

The degree of self-other distinction is learned

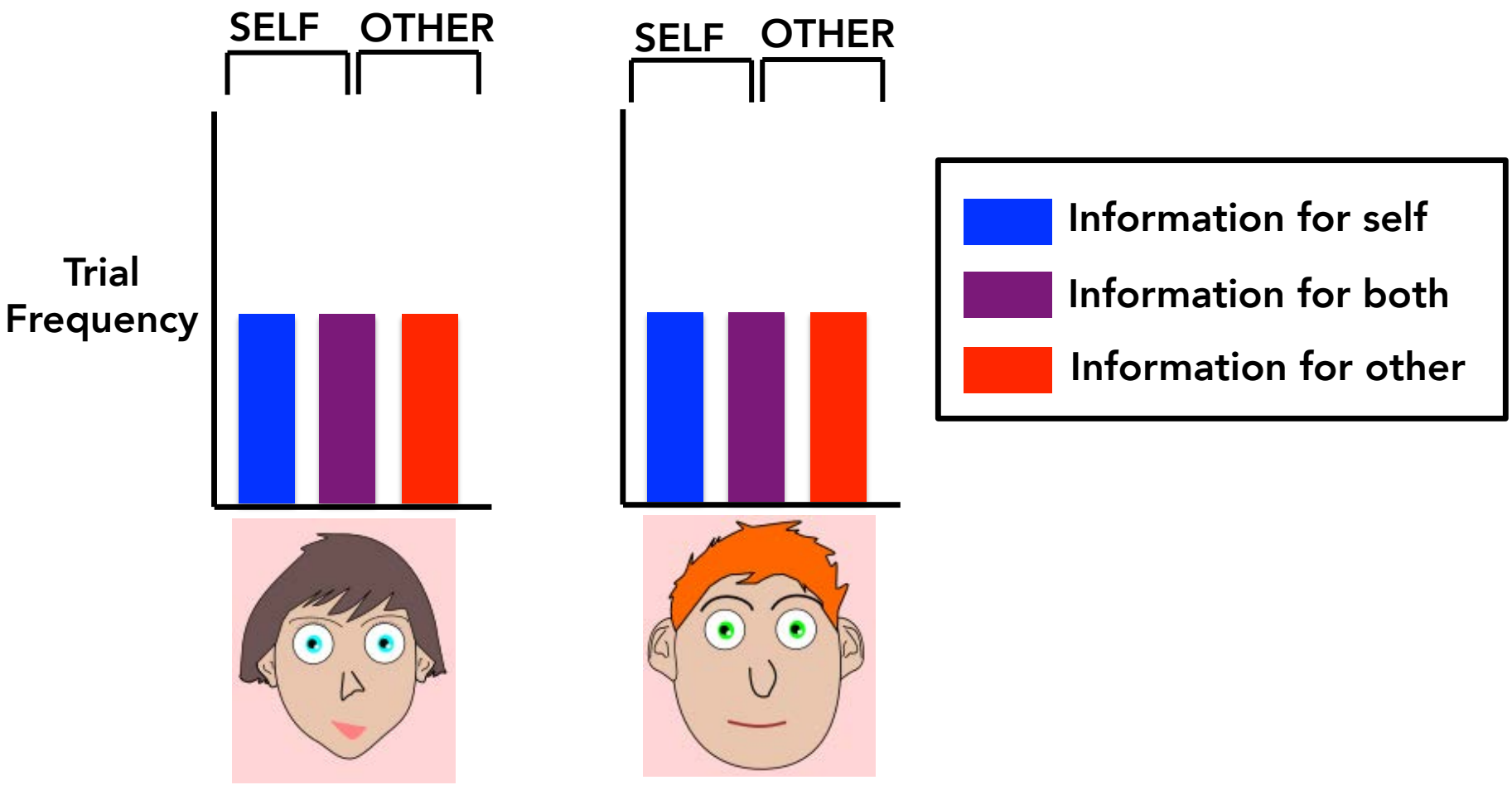




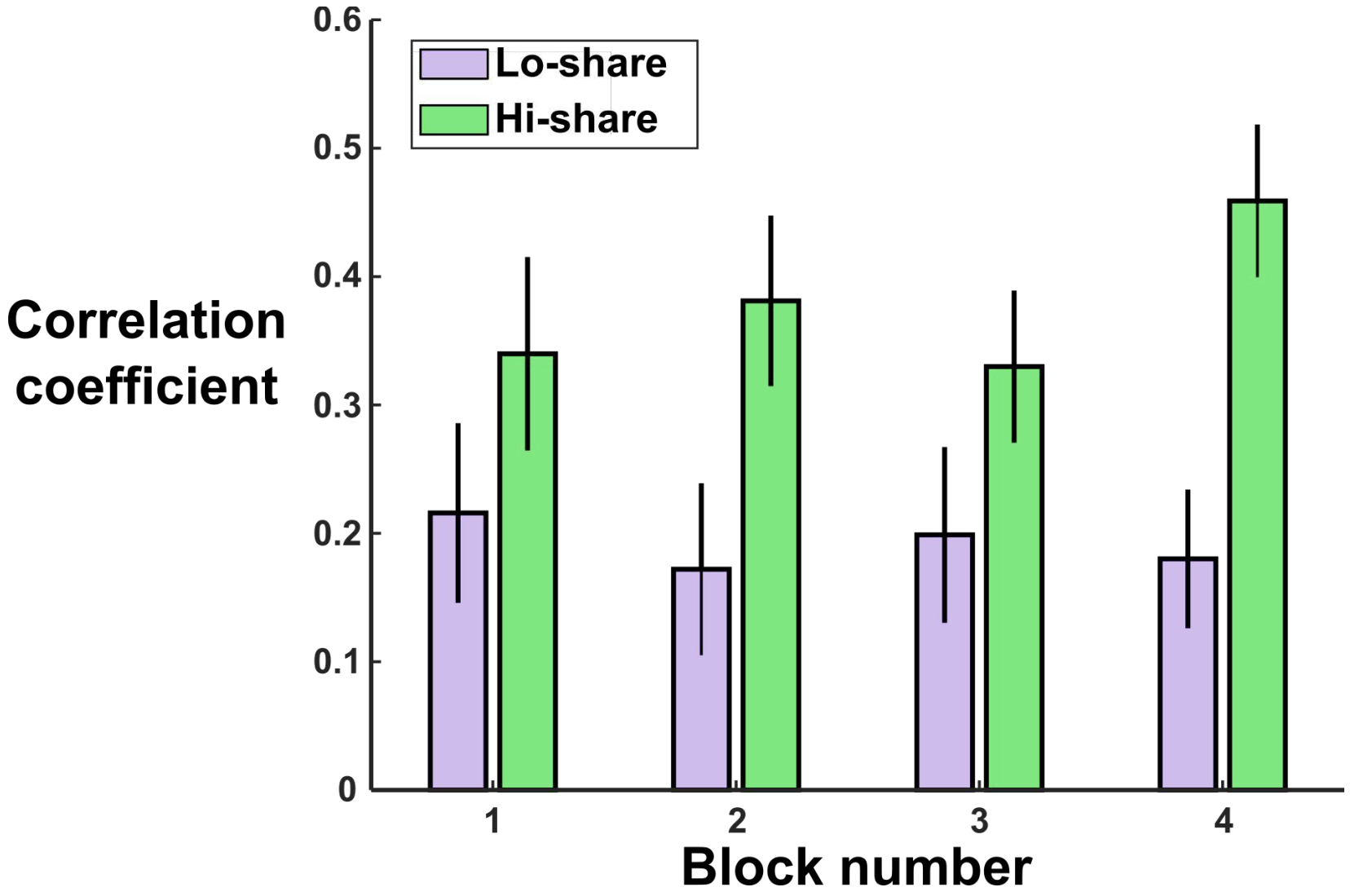
Learning the distinction between self and other



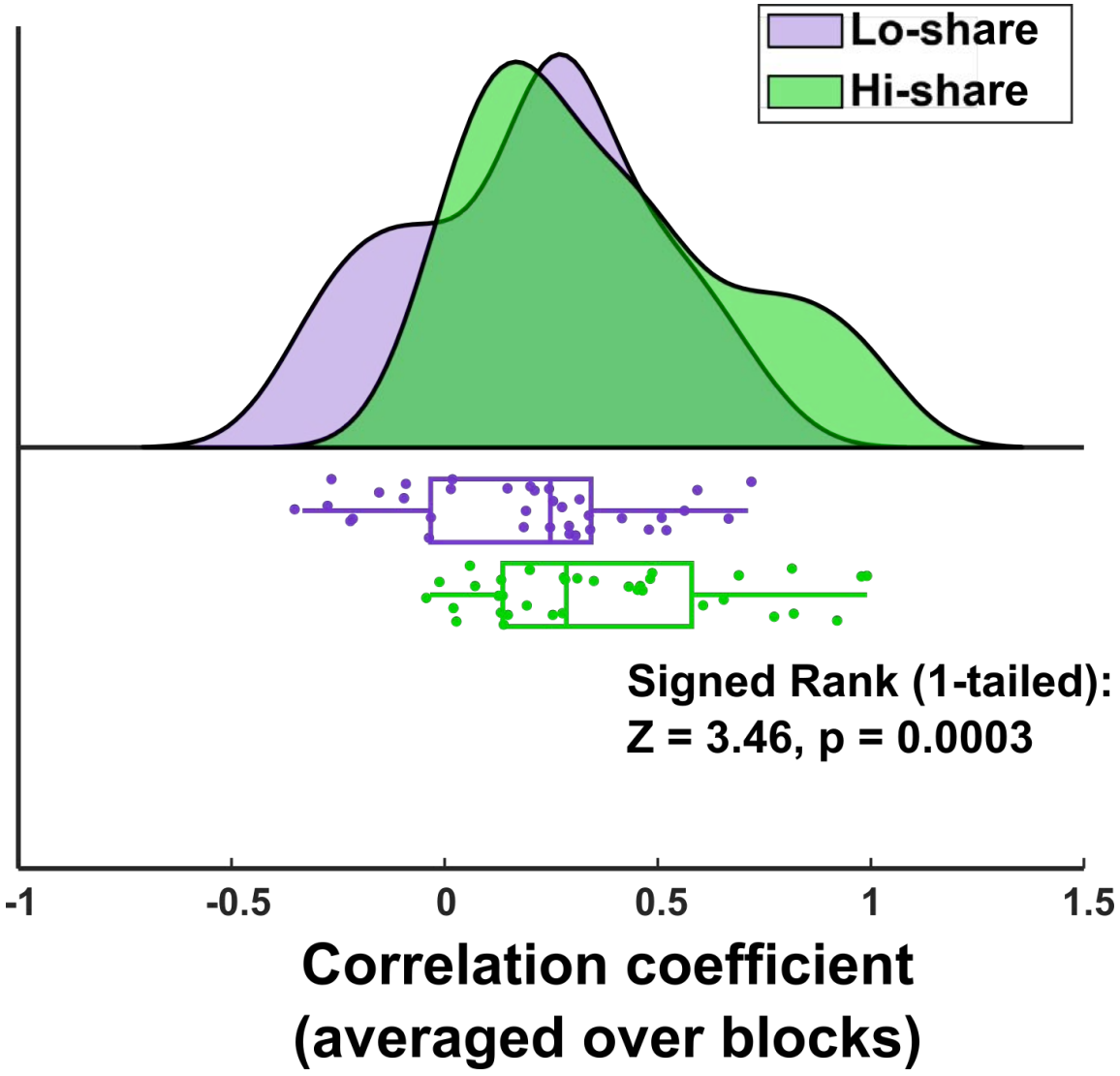
Learning the distinction between self and other



Behavioural evidence for self-other distinction learning



Behavioural evidence for self-other distinction learning



Summary

- Self-other distinction can be achieved with agent-specific learning signals (behaviour, MEG)

- The degree of self-other distinction can be learned through training (behaviour)

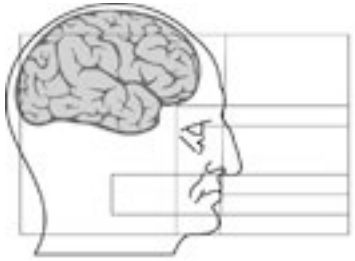
Thank you!

Supervisors:
Ray Dolan
Zeb Kurth-Nelson

Radiographers:
Yoshihito Shigihara
Clive Negus
Megan Creasey
Elisa Tebaldi

Collaborators:
Tobias Hauser
Giles Story
Sam Chamberlain

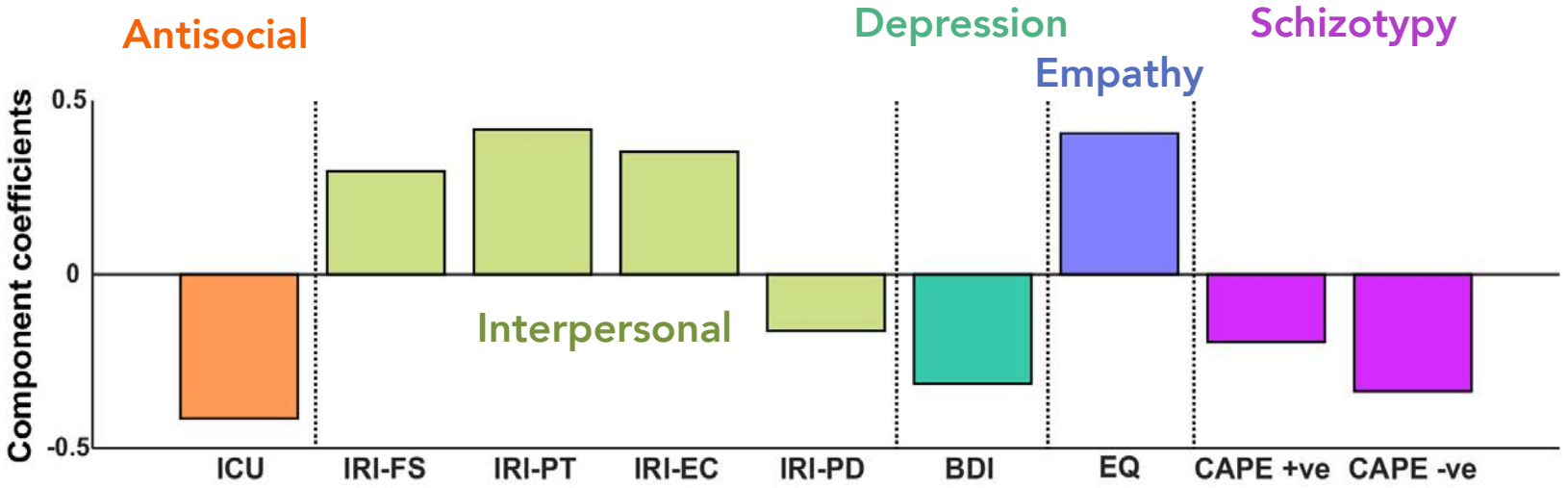
Other brilliant people:
At the Max Planck UCL Centre
At the FIL

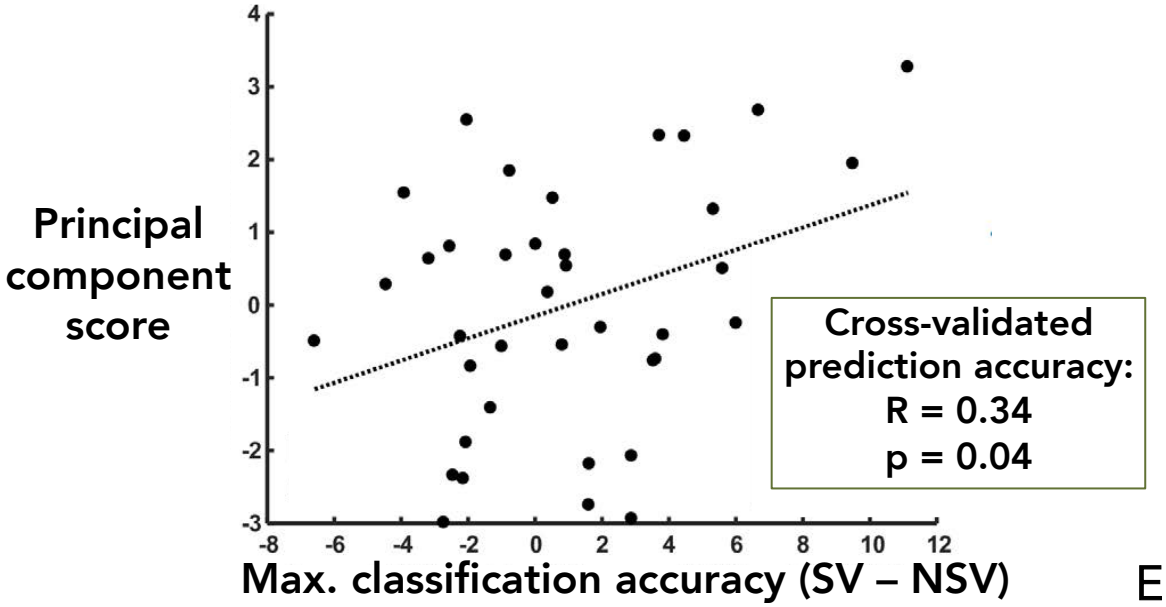


MAX PLANCK UCL CENTRE

International Max Planck Research School
COMP 2 PSYCH

Supplementary: Agent decodability predicts subclinical personality traits





Ereira et al. *PLoS Biol, 2018*

Supplementary: Modelling beliefs

$$PE^{s}_{(t)} = \begin{cases} 0 & \textit{if Decoy trial} \\ Outcome_{(t)} - \hat{B}_{(t-1)} & \textit{otherwise} \end{cases}$$

$$PE^{o} = \begin{cases} 0 & \textit{if Privileged trial} \\ Outcome_{(t)} - \widehat{B_{fb}}_{(t-1)} & \textit{otherwise} \end{cases}$$

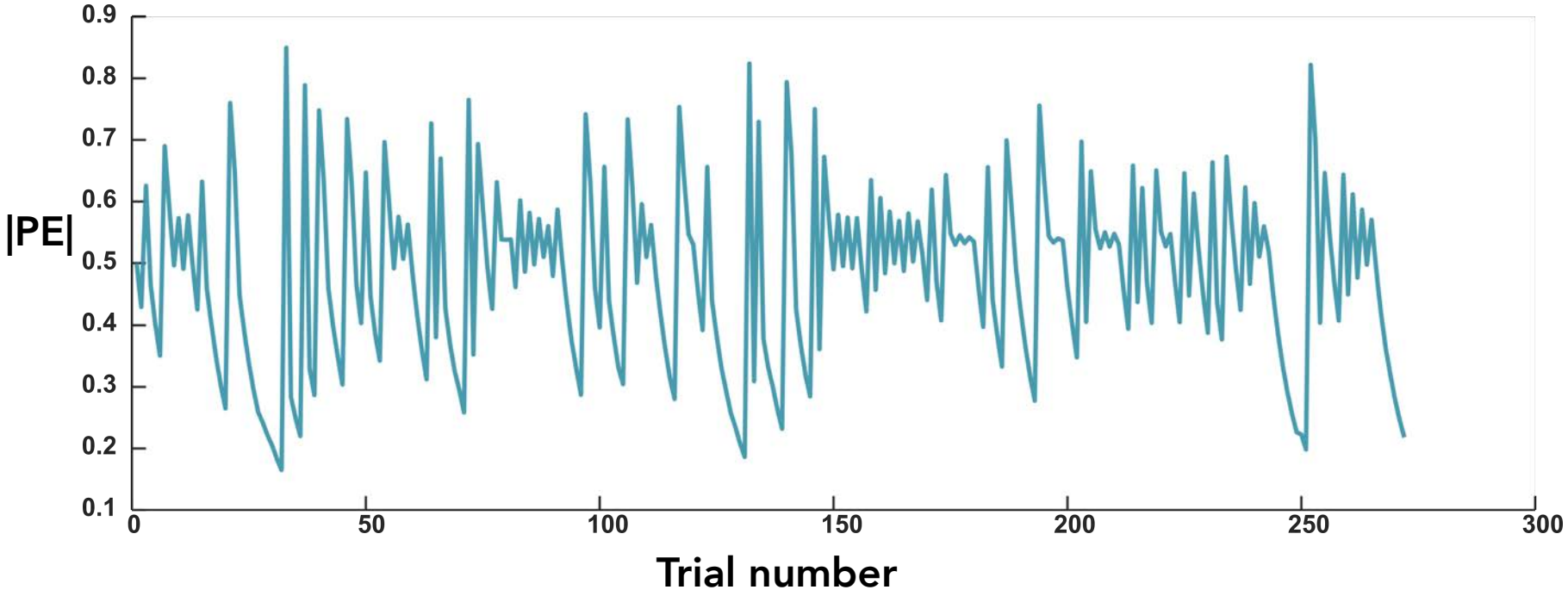
$$\hat{B}_{(t)} = \hat{B}_{(t-1)} + \alpha . PE^{s}{}_{(t)}$$
$$\widehat{B_{fb}}_{(t)} = \widehat{B_{fb}}_{(t-1)} + \alpha . PE^{o}{}_{(t)}$$

$$\hat{B}_{(t)} = \hat{B}_{(t-1)} + \alpha . PE^{s}{}_{(t)} + \delta(0.5 - \hat{B}_{(t-1)})$$
$$\widehat{B_{fb}}_{(t)} = \widehat{B_{fb}}_{(t-1)} + \alpha . PE^{o}{}_{(t)} + \delta(0.5 - \widehat{B_{fb}}_{(t-1)})$$

$$\hat{B}_{(t)} = \hat{B}_{(t-1)} + \alpha . PE^{s}{}_{(t)} + \lambda . PE^{o}{}_{(t)}$$
$$\widehat{B_{fb}}_{(t)} = \widehat{B_{fb}}_{(t-1)} + \alpha . PE^{o}{}_{(t)} + \lambda . PE^{s}{}_{(t)}$$

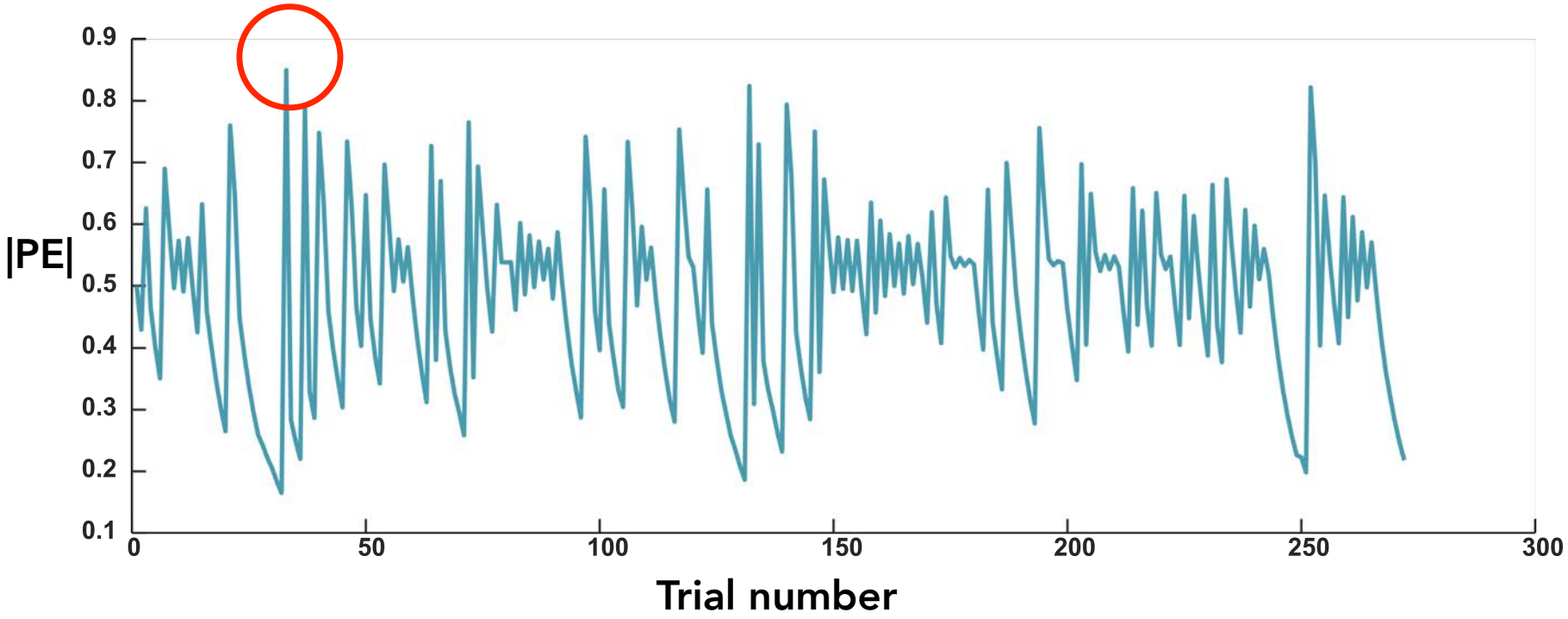
Ereira et al. *PLoS Biol, 2018*

Supplementary: Classification analysis



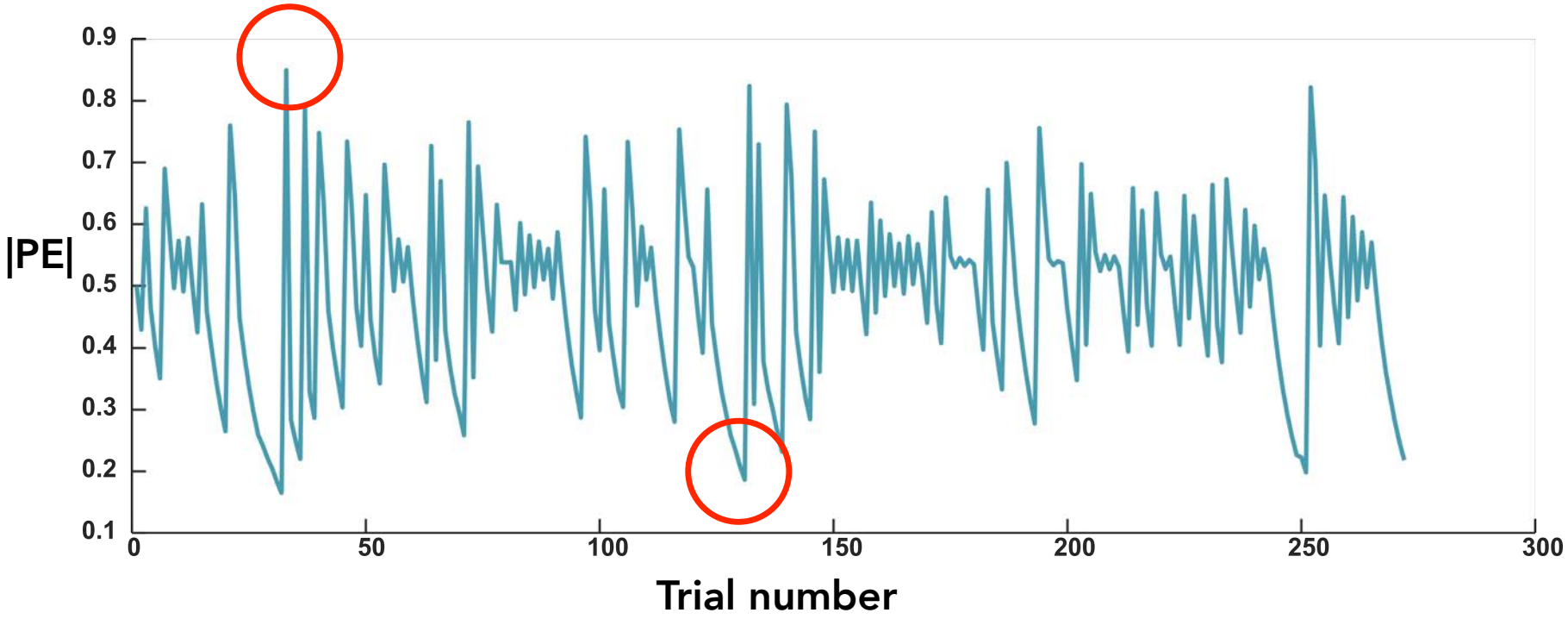
Ereira et al. *PLoS Biol, 2018*

Supplementary: Classification analysis



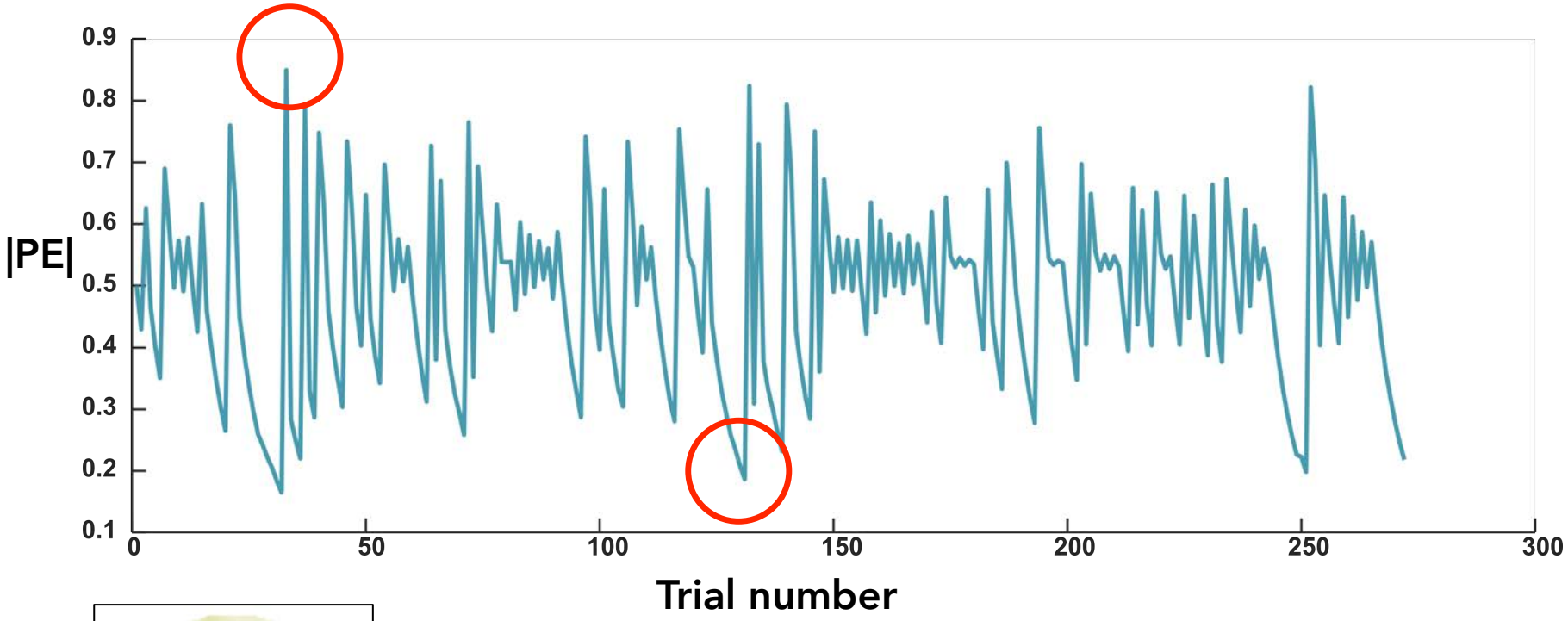
Ereira et al. *PLoS Biol, 2018*

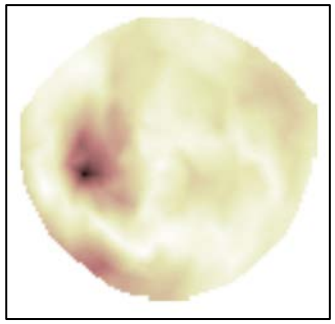
Supplementary: Classification analysis



Ereira et al. *PLoS Biol, 2018*

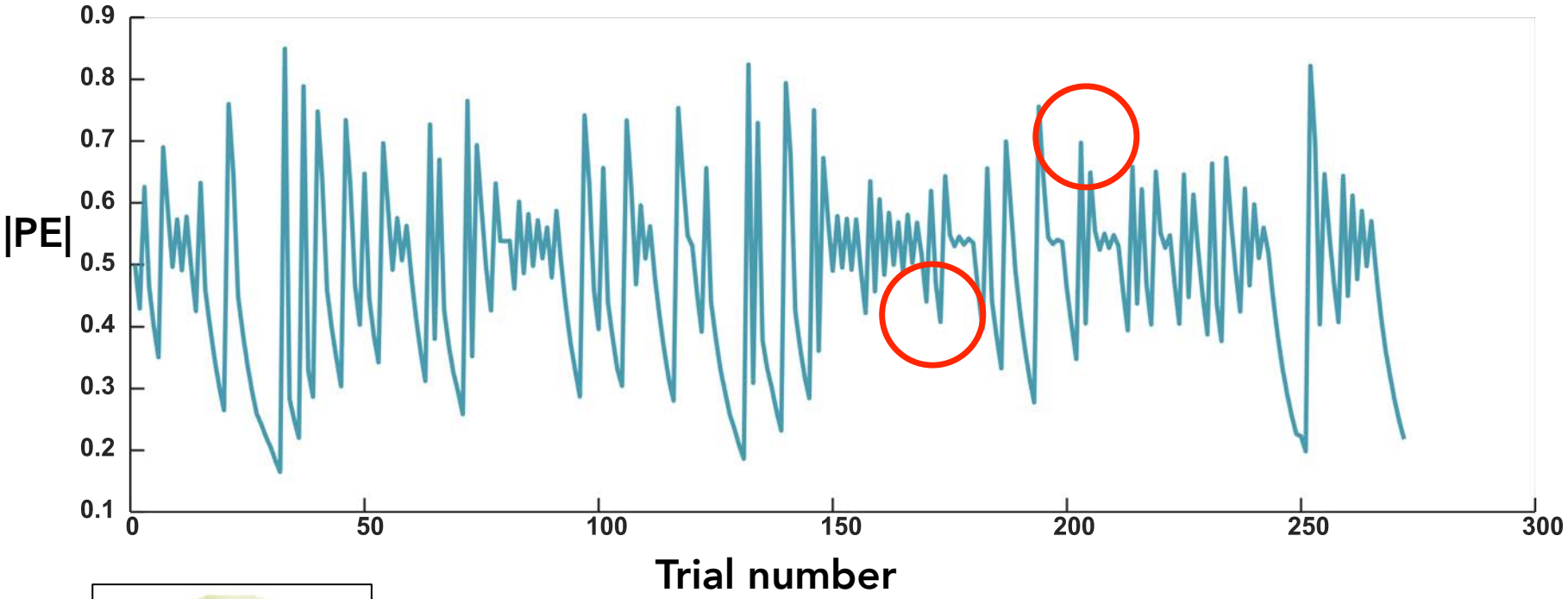
Supplementary: Classification analysis

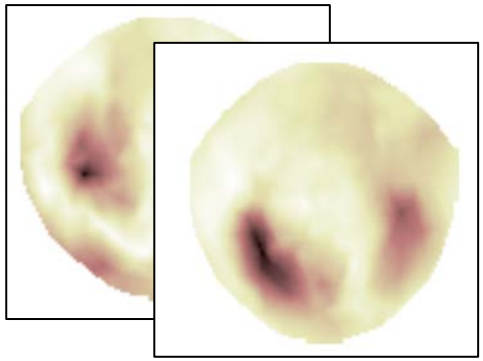




Ereira et al. *PLoS Biol, 2018*

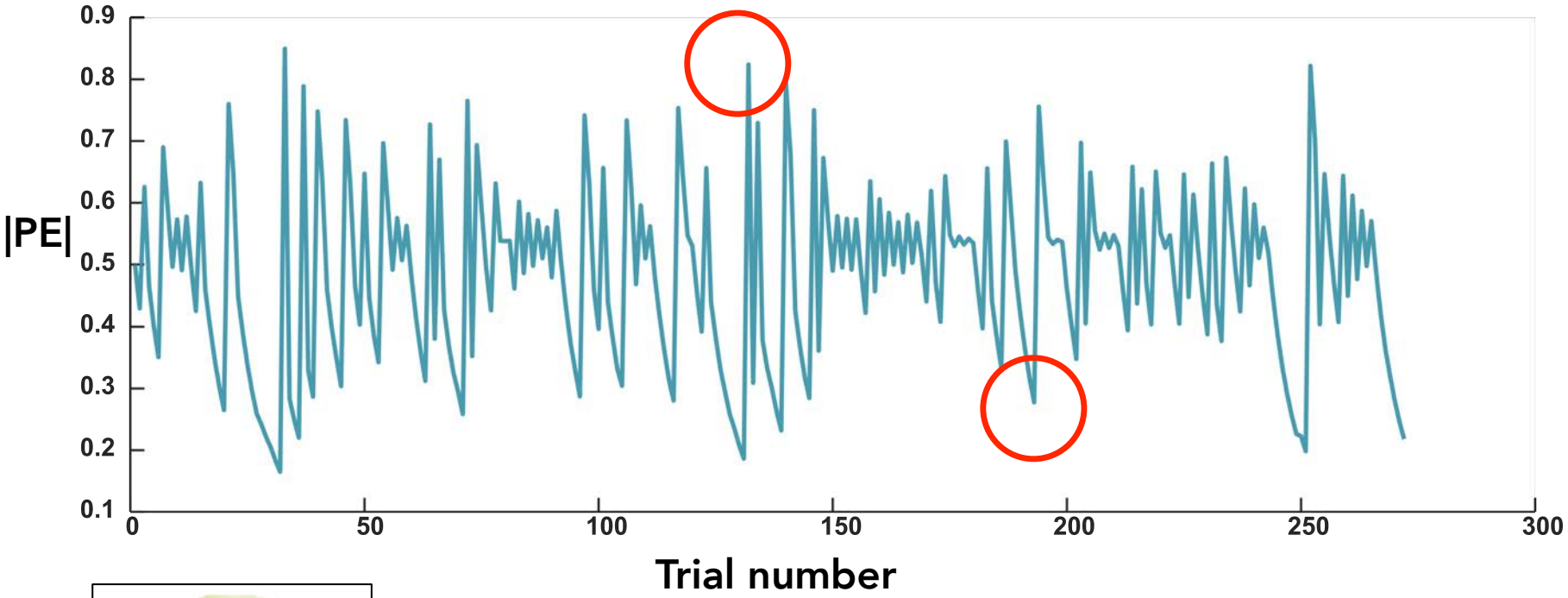
Supplementary: Classification analysis

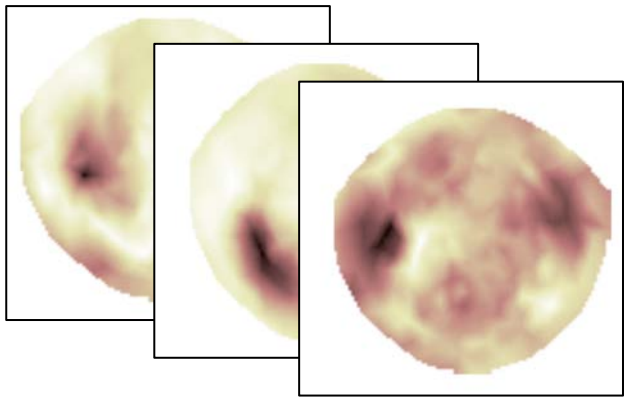




Ereira et al. *PLoS Biol, 2018*

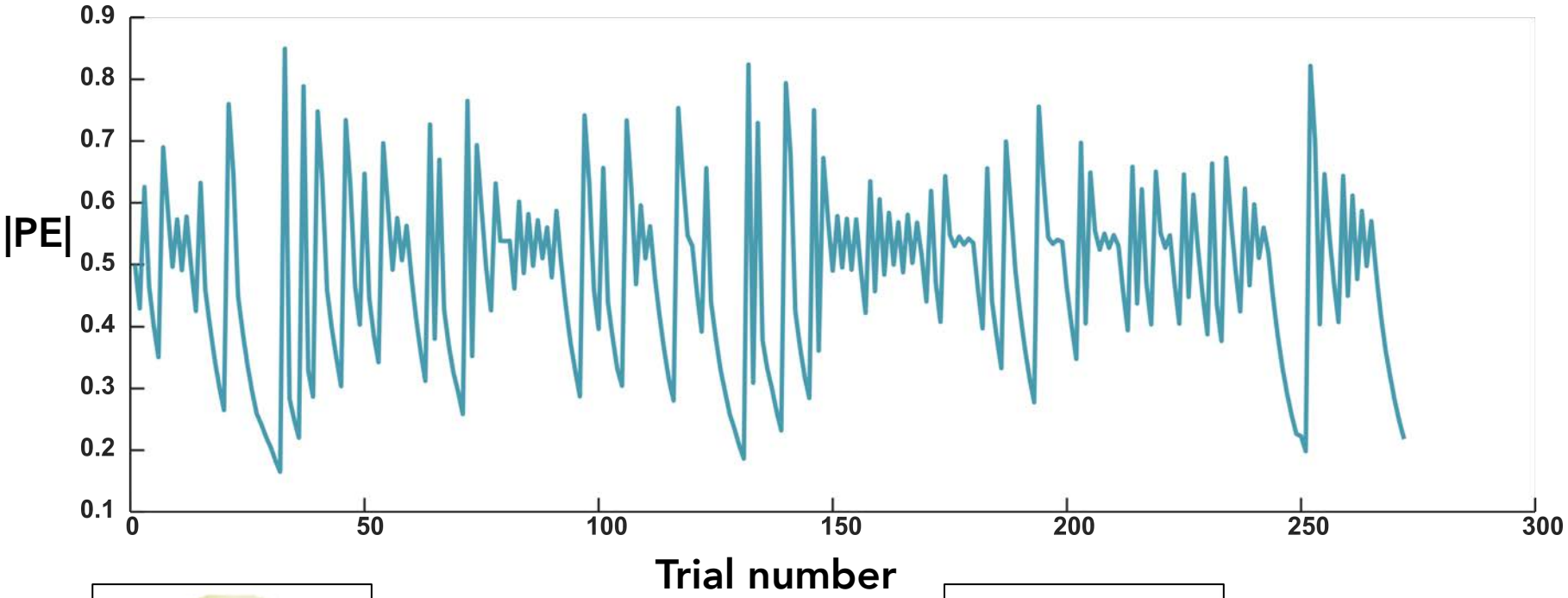
Supplementary: Classification analysis

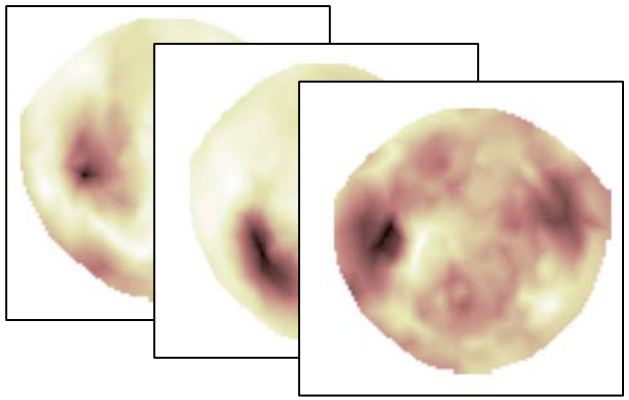


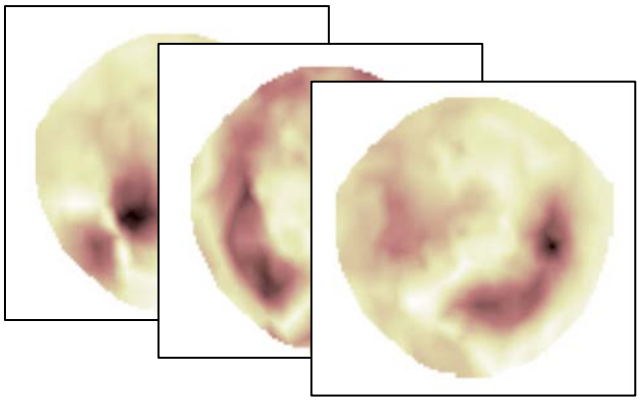


Ereira et al. *PLoS Biol, 2018*

Supplementary: Classification analysis







Ereira et al. *PLoS Biol, 2018*

Supplementary: Association between self-other distinction and delay discounting