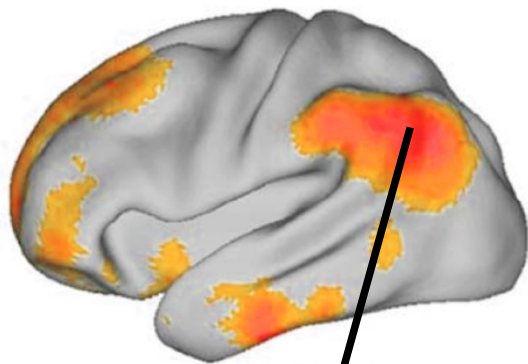**PMC:** Global monitoring and information integration

$$\delta\theta_k \propto -\mathbb{E}_{(s,a,r,s')\sim\mathcal{D}}[(\tilde{Q}(s,a|\theta_k) - y_k)\nabla_{\theta_k}\tilde{Q}(s,a|\theta_k)]$$
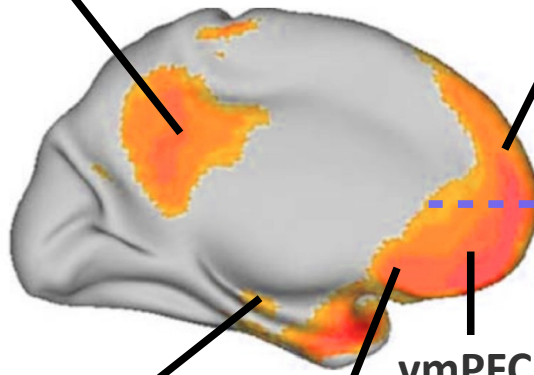
dmPFC (BA9): Action consideration

$$a_k \leftarrow \operatorname{argmax}_{a\in\mathcal{A}}\tilde{Q}(s_k,a|\theta_k)$$

Left TPJ: World semantics

$$s_0 \xrightarrow{r_1} s_1 \xrightarrow{r_2} s_2 \xrightarrow{r_3} \dots$$
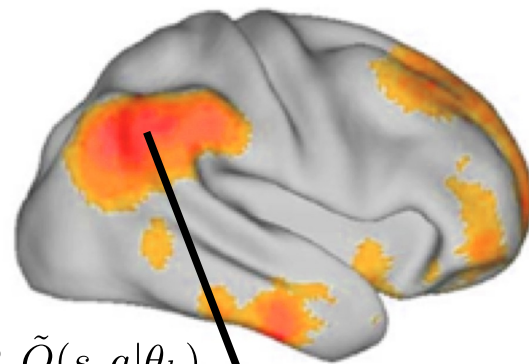
vmPFC (BA10): $\tilde{Q}(s,a|\theta_k)$
Stimulus-value association

Nucleus Accumbens:
Reward / punishment signals

Hippocampus: Memory,
space, and experience replay

$$(s,a,r,s') \sim \mathcal{D}$$

Right TPJ: Prediction error signaling
$$\tilde{Q}(s,a|\theta_k) - y_k$$