Active Inference:
A process theory

Friston, FitzGerald, Rigoli, Schwartenbeck & Pezzulo (2017)

The goal

- To frame active inference in terms of neuronal dynamics and observed behavior.
- To show how Active inference encompasses other models and theories of cognition and behavior.

The ingredients of Active inference

Definition. *Active inference rests on the tuple* $(O, P, Q, R, S, T, \Upsilon)$*:*

- *A finite set of outcomes* O
- *A finite set of control states or actions* Υ
- *A finite set of hidden states* S
- *A finite set of time-sensitive policies* T
- *A generative process* $R(\tilde{o}, \tilde{s}, \tilde{u})$ *that generates probabilistic outcomes* $o \in O$ *from (hidden) states* $s \in S$ *and action* $u \in \Upsilon$
- *A generative model* $P(\tilde{o}, \tilde{s}, \pi, \eta)$ *with parameters* η*, over outcomes, states, and policies* $\pi \in T$*, where* $\pi \in \{0, \ldots, K\}$ *returns a sequence of actions* $u_t = \pi(t)$
- *An approximate posterior* $Q(\tilde{s}, \pi, \eta) = Q(s_0|\pi) \ldots Q(s_T|\pi)Q(\pi)Q(\eta)$ *over states, policies and parameters with expectations* $(\mathbf{s}_0^\pi, \ldots, \mathbf{s}_T^\pi, \boldsymbol{\pi}, \boldsymbol{\eta})$

The generative model

$$P(\tilde{o}, \tilde{s}, \pi, \eta) = P(\pi)P(\eta)\prod_{t=1}^{T} P(o_t|s_t)P(s_t|s_{t-1}, \pi)$$

$$P(o_t|s_t) = Cat(\mathbf{A})$$

$$P(s_{t+1}|s_t, \pi) = Cat(\mathbf{B}(u = \pi(t)))$$

$$P(s_1|s_0) = Cat(\mathbf{D})$$

$$P(\pi) = \sigma(-\gamma \cdot \mathbf{G}(\pi))$$

$$P(\mathbf{A}) = Dir(a)$$

$$P(\mathbf{B}) = Dir(b)$$

$$P(\mathbf{D}) = Dir(d)$$

$$P(\gamma) = \Gamma(1, \beta).$$

$$\eta = \{a, b, d, \beta\}$$

The approximate posterior

$$Q(x) = Q(s_1|\pi)\dots Q(s_T|\pi)Q(\pi)Q(\mathbf{A})Q(\mathbf{B})Q(\mathbf{D})Q(\gamma)$$

$$Q\left(s_t|\pi\right) = Cat\left(\mathbf{s}_t^\pi\right)$$

$$Q(\pi) = Cat\left(\boldsymbol{\pi}\right)$$

$$Q(\mathbf{A}) = Dir(\mathbf{a})$$

$$Q(\mathbf{B}) = Dir(\mathbf{b})$$

$$Q(\mathbf{D}) = Dir(\mathbf{d})$$

$$Q(\gamma) = \Gamma(1, \boldsymbol{\beta})$$

$$\mathbf{x} = (\mathbf{s}_0^\pi, \dots, \mathbf{s}_T^\pi, \boldsymbol{\pi}, \boldsymbol{\eta})$$

$$\boldsymbol{\eta} = (\mathbf{a}, \mathbf{b}, \mathbf{d}, \boldsymbol{\beta})$$

Behavior, action and reflexes

$$u_t = \min_u E_Q[D[P(o_{t+1}|s_{t+1})||R(o_{t+1}|s_t, u)]]$$
$$= \min_u \mathbf{o}_{t+1} \cdot \varepsilon^u_{t+1}$$

$$\varepsilon^u_{t+1} = \widehat{\mathbf{o}}_{t+1} - \widehat{\mathbf{o}}^u_{t+1}$$
$$\mathbf{o}_{t+1} = \mathbf{A}\mathbf{s}_{t+1}$$
$$\mathbf{o}^u_{t+1} = \mathbf{A}\mathbf{B}(u)\mathbf{s}_t$$
$$\mathbf{s}_t = \sum_\pi \boldsymbol{\pi}_\pi \cdot \mathbf{s}^\pi_t .$$

Free energy and expected Free energy

$$Q(x) = \arg\min_{Q(x)} F$$
$$\approx P(x|\tilde{o})$$

$$F = E_Q[\ln Q(x) - \ln P(x, \tilde{o})]$$
$$= E_Q[\ln Q(x) - \ln P(x|\tilde{o}) - \ln P(\tilde{o})]$$
$$= E_Q[\ln Q(x) - \ln P(\tilde{o}|x) - \ln P(x)]$$

$$= \underbrace{D[Q(x)||P(x|\tilde{o})]}_{relative\,entropy} - \underbrace{\ln P(\tilde{o})}_{log\,evidence}$$

$$= \underbrace{D[Q(x)||P(x)]}_{complexity} - \underbrace{E_Q[\ln P(\tilde{o}|x)]}_{accuracy},$$

$$G(\pi) = \sum_\tau G(\pi, \tau),$$

$$G(\pi, \tau) = E_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln P(s_\tau, o_\tau|\tilde{o}, \pi)]$$
$$= E_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln P(s_\tau|o_\tau, \tilde{o}, \pi) - \ln P(o_\tau)],$$

$$\approx \underbrace{E_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau, \pi)]}_{(-ve)\,mutual\,information} - \underbrace{E_{\tilde{Q}}[\ln P(o_\tau)]}_{expected\,log\,evidence}$$

$$= \underbrace{E_{\tilde{Q}}[\ln Q(o_\tau|\pi) - \ln Q(o_\tau|s_\tau, \pi)]}_{(-ve)\,epistemic\,value} - \underbrace{E_{\tilde{Q}}[\ln P(o_\tau)]}_{extrinsic\,value}$$

$$= \underbrace{D[Q(o_\tau|\pi)||P(o_\tau)]}_{expected\,cost} + \underbrace{E_{\tilde{Q}}[H[P(o_\tau|s_\tau)]]}_{expected\,ambiguity},$$

Belief updating

$$\left.\begin{aligned}
\mathbf{s}_\tau^\pi &= \sigma(\widehat{\mathbf{A}} \cdot \mathrm{o}_\tau + \widehat{\mathbf{B}}_{\tau-1}^\pi \mathbf{s}_{\tau-1}^\pi + \widehat{\mathbf{B}}_\tau^\pi \cdot \mathbf{s}_{\tau+1}^\pi) \\
\boldsymbol{\pi} &= \sigma(-\mathbf{F} - \boldsymbol{\gamma} \cdot \mathbf{G}) \\
\boldsymbol{\beta} &= \beta + (\boldsymbol{\pi} - \boldsymbol{\pi}_0) \cdot \mathbf{G}
\end{aligned}\right\} \text{Inference,}$$

$$\left.\begin{aligned}
\widehat{\mathbf{A}} &= \psi(\mathbf{a}) - \psi(\mathbf{a}_0) \quad & \mathbf{a} &= a + \textstyle\sum_\tau o_\tau \otimes \mathbf{s}_\tau \\
\widehat{\mathbf{B}} &= \psi(\mathbf{b}) - \psi(\mathbf{b}_0) \quad & \mathbf{b}(u) &= b(u) + \textstyle\sum_{\pi(\tau)=u} \boldsymbol{\pi}_\pi \cdot \mathbf{s}_\tau^\pi \otimes \mathbf{s}_{\tau-1}^\pi \\
\widehat{\mathbf{D}} &= \psi(\mathbf{d}) - \psi(\mathbf{d}_0) \quad & \mathbf{d} &= d + \mathbf{s}_1
\end{aligned}\right\} \text{Learning.}$$

Gradient descent and neuronal dynamics

$$\dot{\widehat{\mathbf{s}}}_{\tau}^{\pi} = \partial_{\widehat{\mathbf{s}}}\mathbf{s}_{\tau}^{\pi} \cdot \varepsilon_{\tau}^{\pi}$$

$$\mathbf{s}_{\tau}^{\pi} = \sigma(\widehat{\mathbf{s}}_{\tau}^{\pi})$$

$$\dot{\boldsymbol{\beta}} = \boldsymbol{\gamma}^{2}\varepsilon^{\gamma}$$

Explicit representation of future and past states.

$$\varepsilon_{\tau}^{\pi} = (\widehat{\mathbf{A}} \cdot o_{\tau} + \widehat{\mathbf{B}}_{\tau-1}^{\pi}\mathbf{s}_{\tau-1}^{\pi} + \widehat{\mathbf{B}}_{\tau}^{\pi} \cdot \mathbf{s}_{\tau+1}^{\pi}) - \widehat{\mathbf{s}}_{\tau}^{\pi}$$

$$\varepsilon^{\gamma} = (\beta - \boldsymbol{\beta}) + (\boldsymbol{\pi} - \boldsymbol{\pi}_{0}) \cdot \mathbf{G}.$$

Prediction error responses as the change in Free energy or belief updating.

Reward prediction error

- If planned actions minimize expected Free Energy (G), then policies become more precise.
- Dopaminergic firing reflects this increase in precision or expected reward.
- Precision is related to reward prediction error.

Functional neuroanatomy

(solutions to) Belief updating

Action selection (and Bayesian model averaging)

$$u_t = \min_u \mathbf{o}_{t+1} \cdot \varepsilon^u_{t+1}$$
$$\varepsilon^u_{t+1} = \ln \mathbf{A}\mathbf{s}_{t+1} - \ln \mathbf{A}\mathbf{B}(u)\mathbf{s}_t$$
$$\mathbf{s}_t = \sum_\pi \boldsymbol{\pi}_\pi \cdot \mathbf{s}^\pi_t$$

State estimation (planning as inference)

$$\mathbf{s}^\pi_\tau = \sigma(\bar{\mathbf{A}} \cdot o_\tau + \bar{\mathbf{B}}^\pi_{\tau-1}\mathbf{s}^\pi_{\tau-1} + \bar{\mathbf{B}}^\pi_\tau \cdot \mathbf{s}^\pi_{\tau+1})$$

Policy selection

$$\boldsymbol{\pi} = \sigma(-\mathbf{F} - \gamma \cdot \mathbf{G})$$
$$F(\pi,\tau) = \mathbf{s}^\pi_\tau \cdot (\bar{\mathbf{s}}^\pi_\tau - \bar{\mathbf{A}} \cdot o_\tau - \bar{\mathbf{B}}^\pi_{\tau-1}\mathbf{s}^\pi_{\tau-1})$$
$$G(\pi,\tau) = \mathbf{o}^\pi_\tau \cdot (\bar{\mathbf{o}}^\pi_\tau - \mathbf{U}_\tau) + \mathbf{s}^\pi_\tau \cdot \mathbf{H}$$

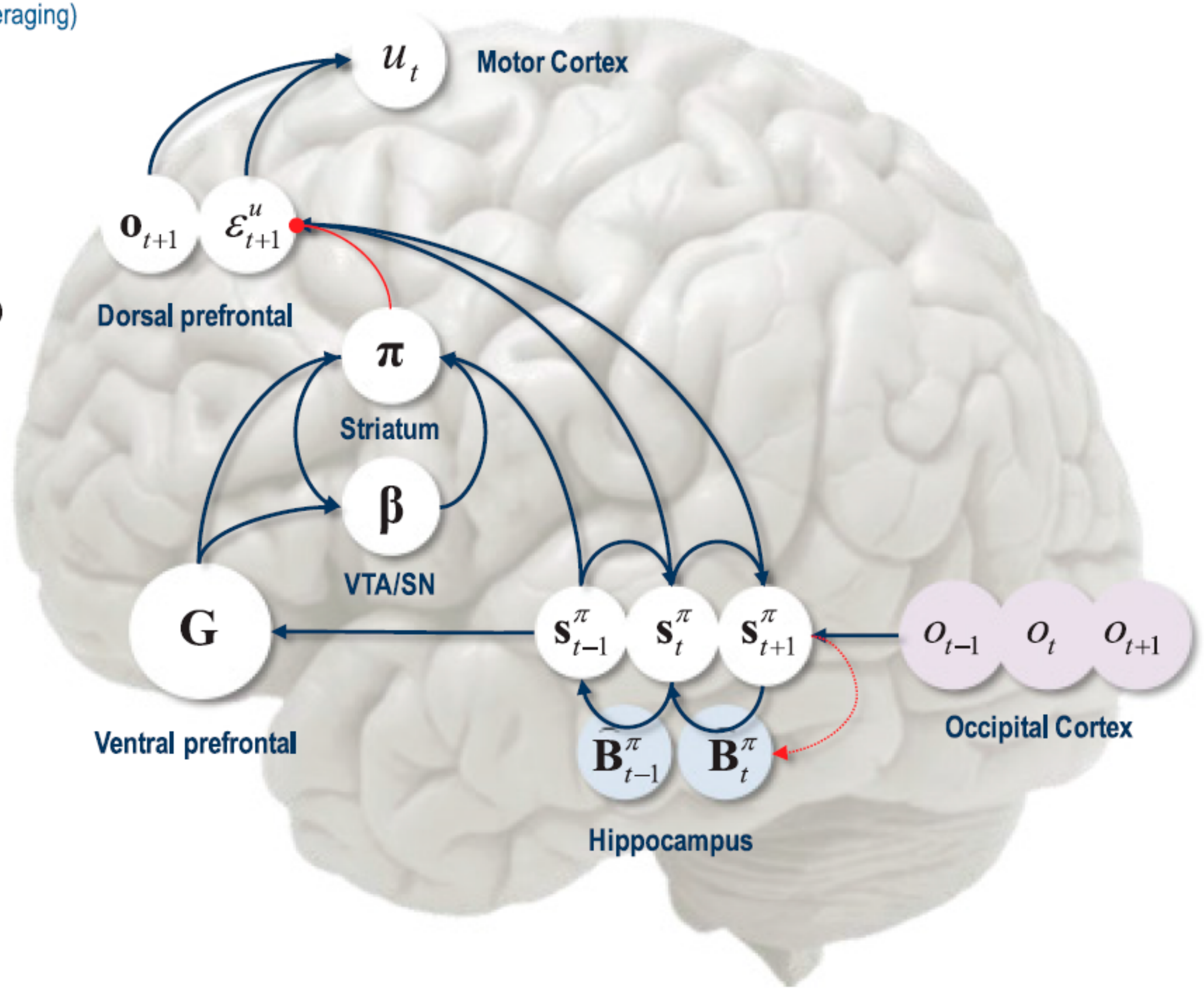
Precision (incentive salience)

$$\boldsymbol{\beta} = \beta + (\boldsymbol{\pi} - \boldsymbol{\pi}_0) \cdot \mathbf{G}$$

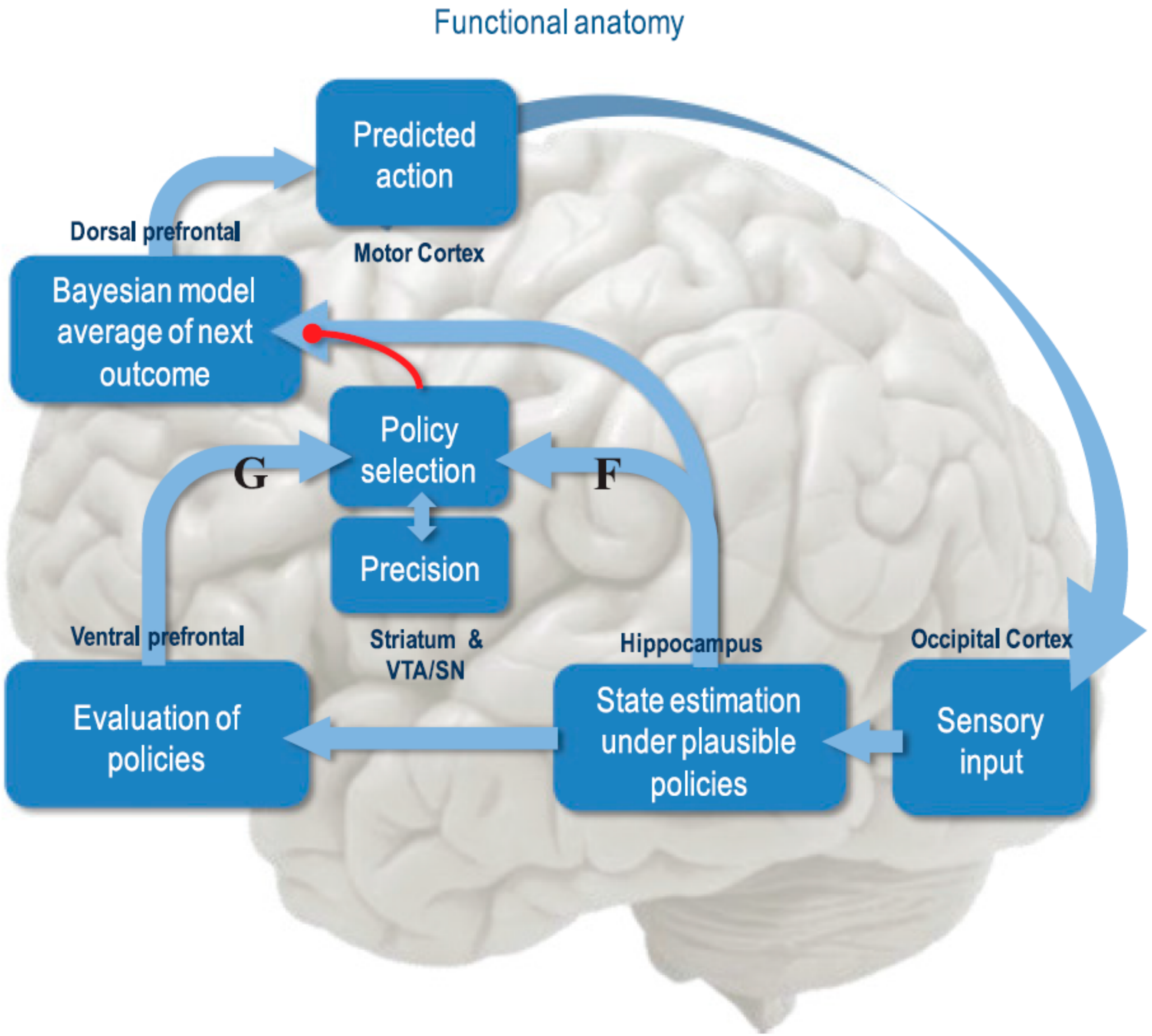
Learning

$$\mathbf{b}(u) = b(u) + \sum_{\pi(\tau)=u} \boldsymbol{\pi}_\pi \cdot \mathbf{s}^\pi_\tau \otimes \mathbf{s}^\pi_{\tau-1}$$
$$\bar{\mathbf{B}} = \psi(\mathbf{b}) - \psi(\mathbf{b}^0)$$

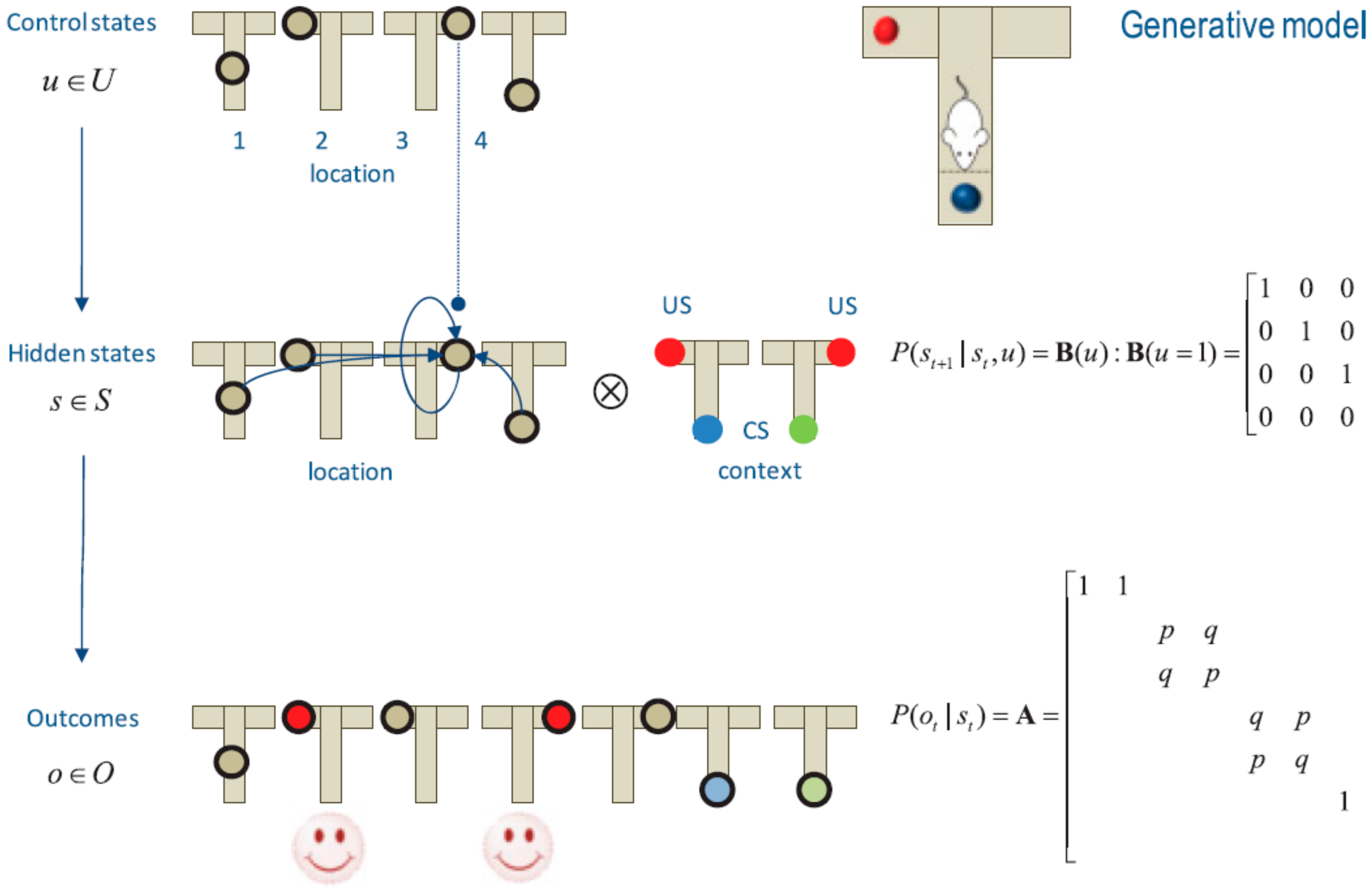
Functional anatomy



Functional neuroanatomy



Simulations of inference (T-maze)



$$P(s_{t+1} \mid s_t, u) = \mathbf{B}(u) : \mathbf{B}(u=1) = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ldots$$

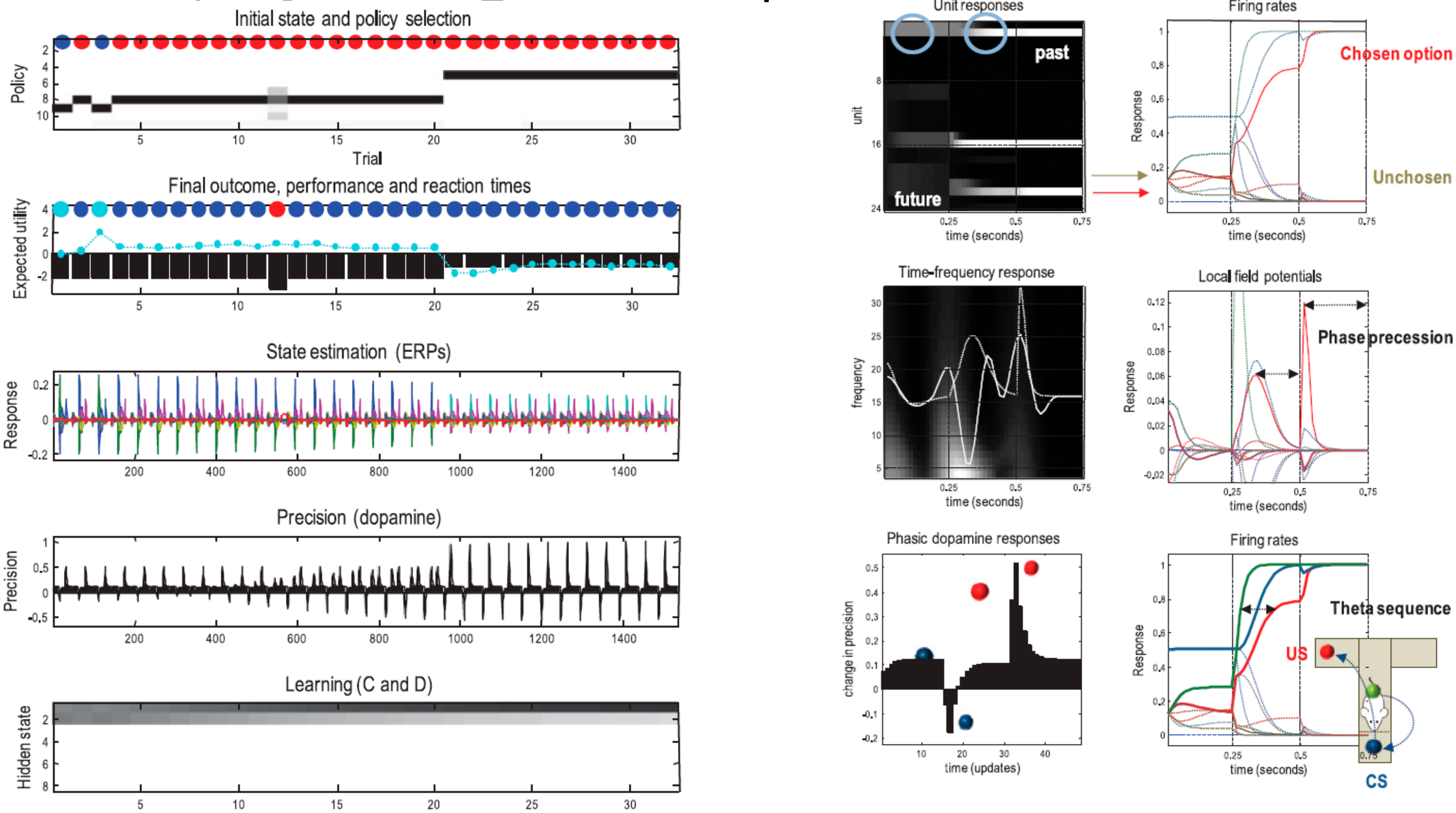
$$P(o_t \mid s_t) = \mathbf{A} = \begin{bmatrix} 1 & 1 & & & & & & \\ & & p & q & & & & \\ & & q & p & & & & \\ & & & & q & p & & \\ & & & & p & q & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix} : q = 1 - p$$

$$\ln P(o_t) = \mathbf{U} = [0 \quad 3 \quad -3 \quad 3 \quad -3 \quad 0 \quad 0]^T$$

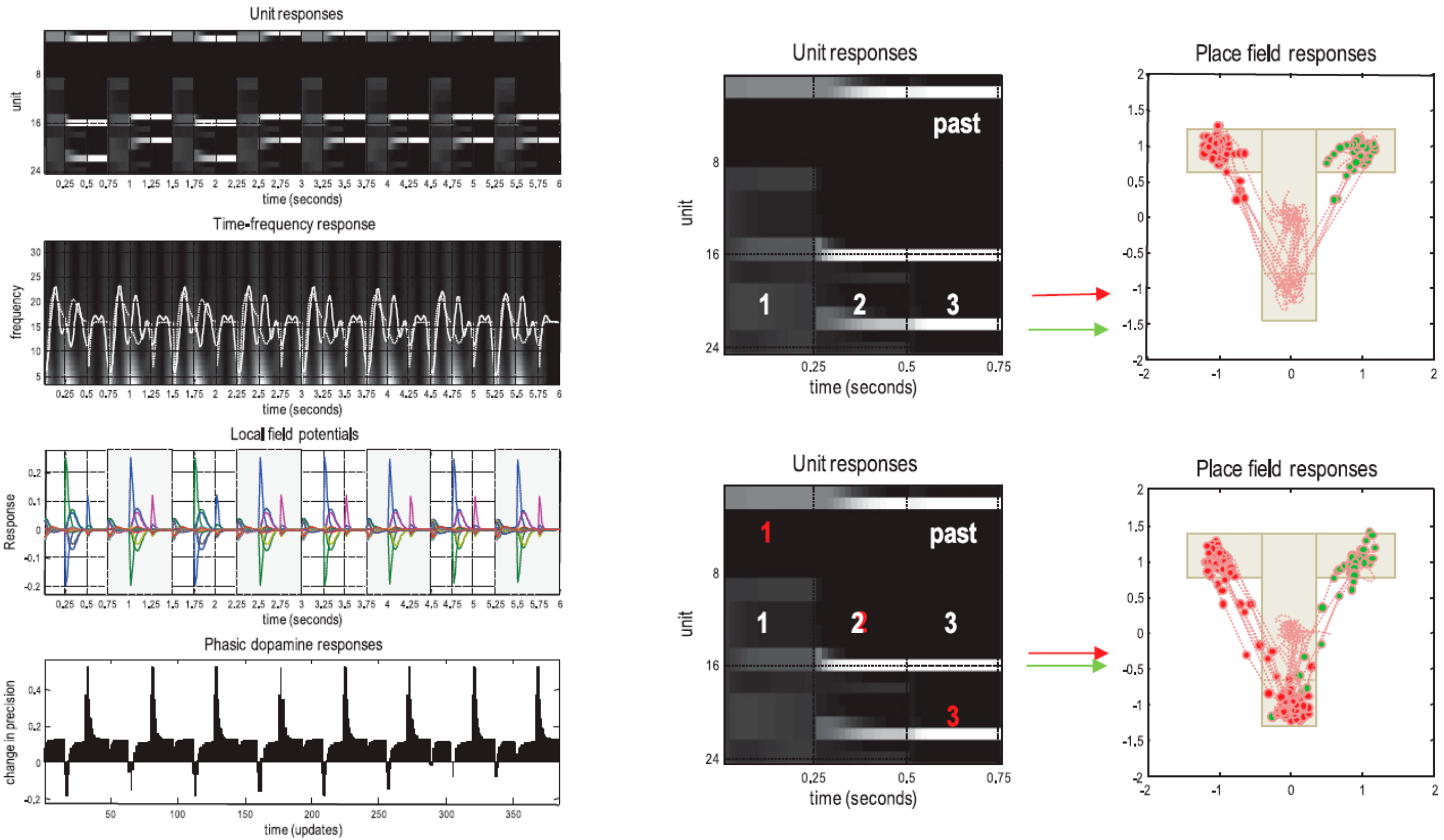
10 policies:

- Stay and move (1-4)
- Go to reward (5-6)
- go to cue and then move (7:10)

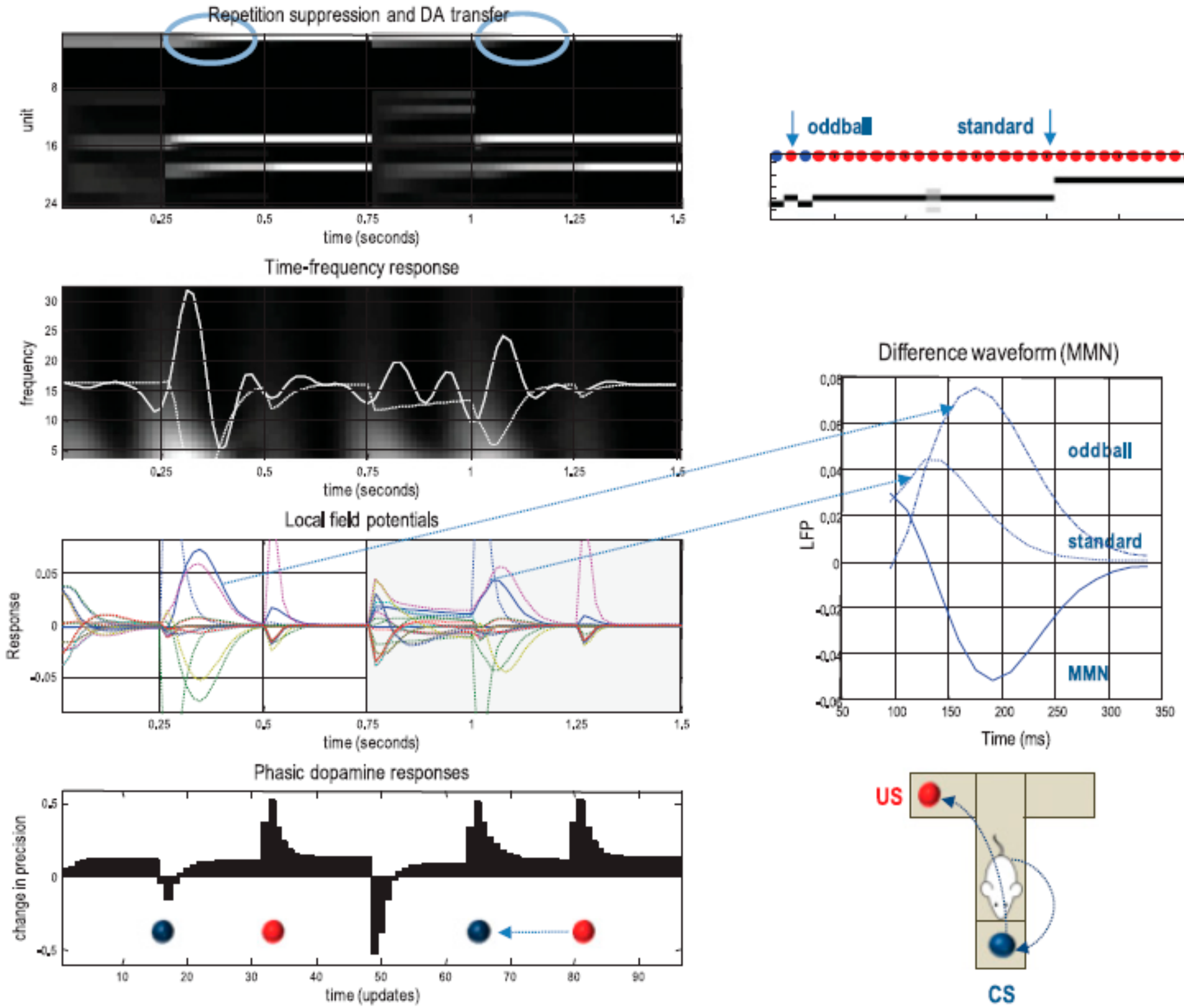
Neurophysiological responses



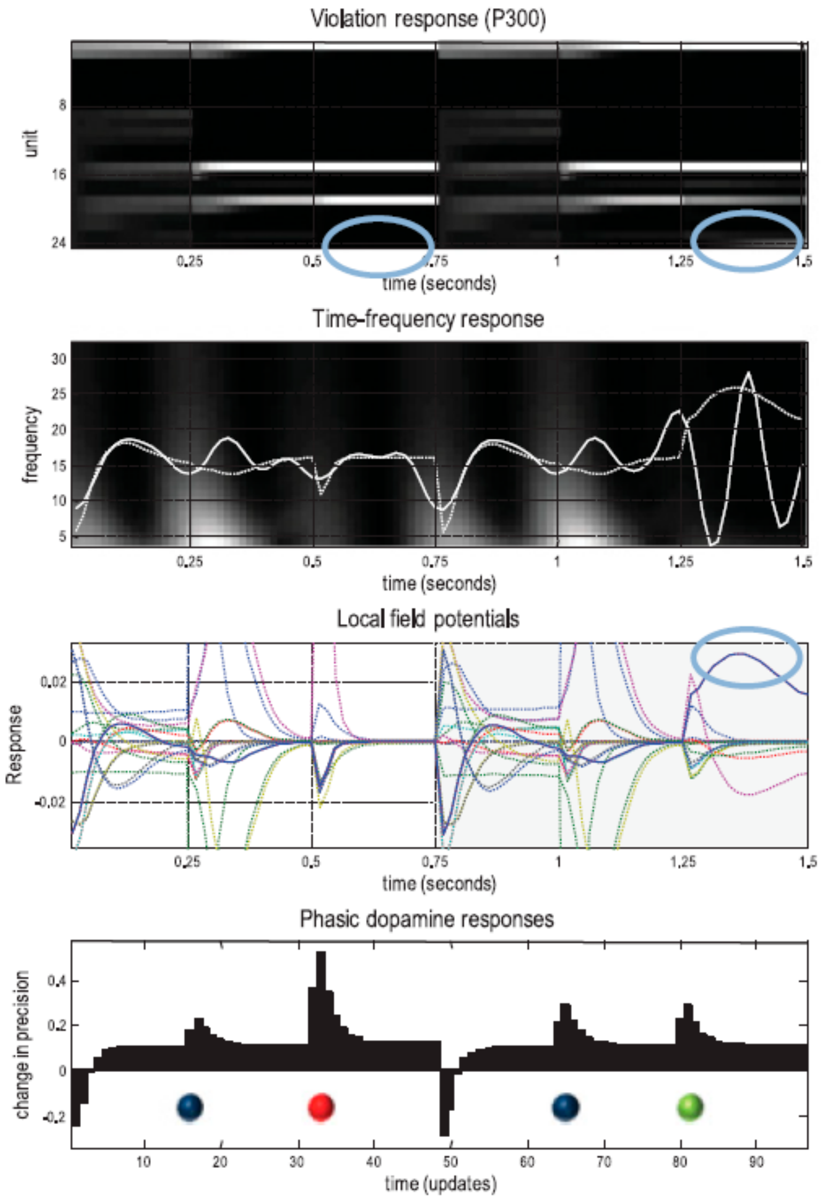
Theta-gamma coupling

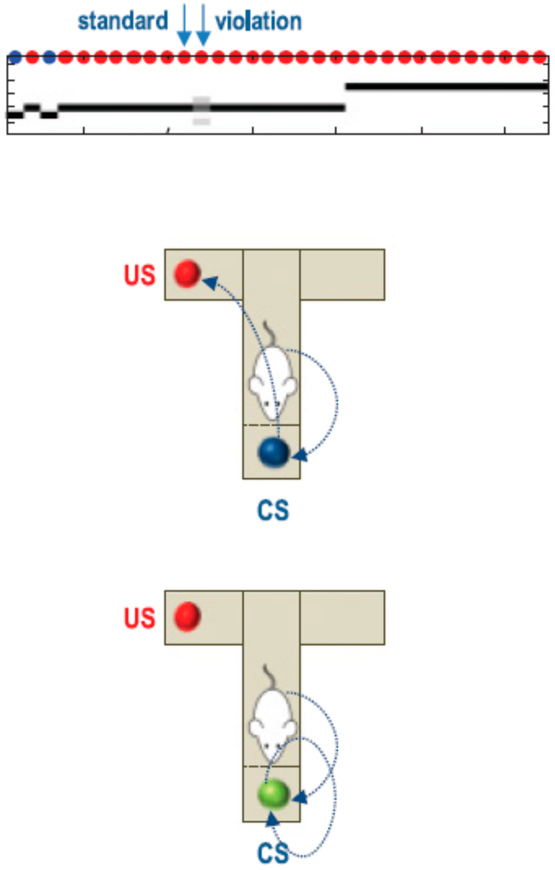


Repetition suppression and dopamine transfer

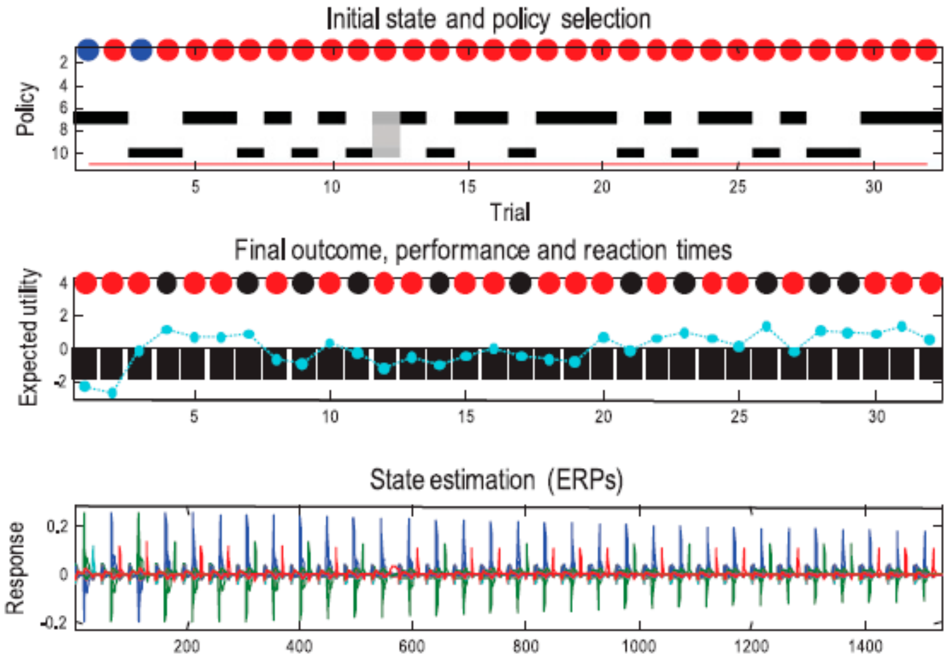


Violation responses and P300





Epistemic foraging



Epistemic foraging