Models of the motor system

Fall 2021

Opher Donchin

Part 1

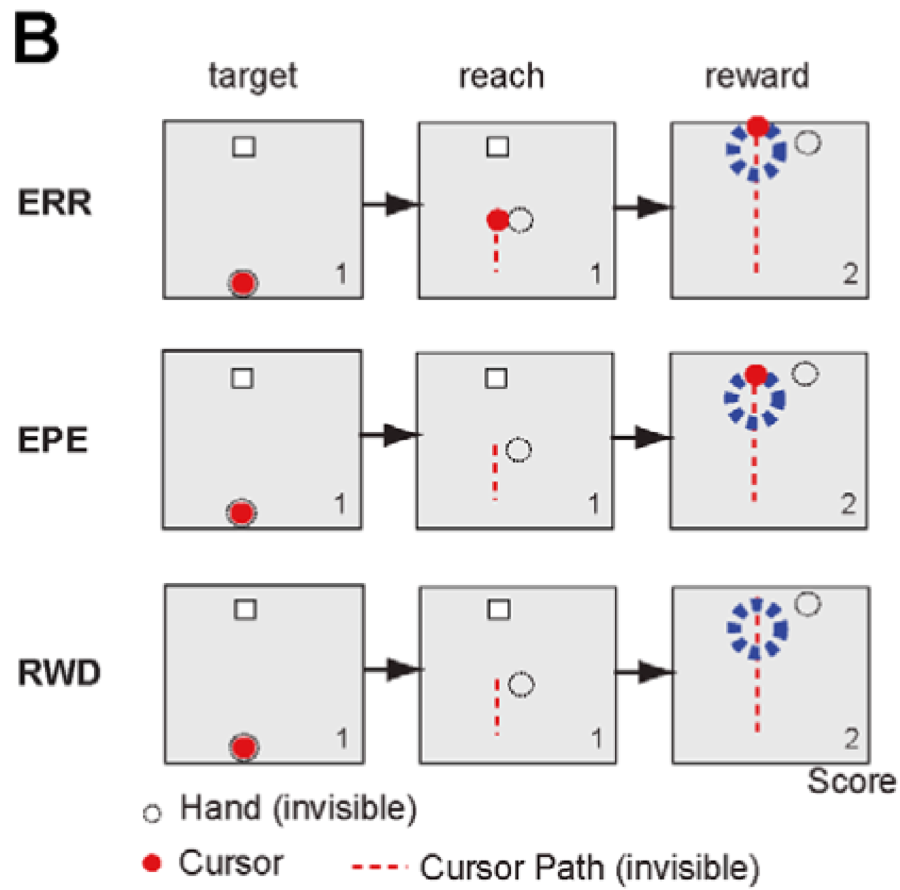
- Reminder reward learning

Teaching with errors and rewards

Error feedback

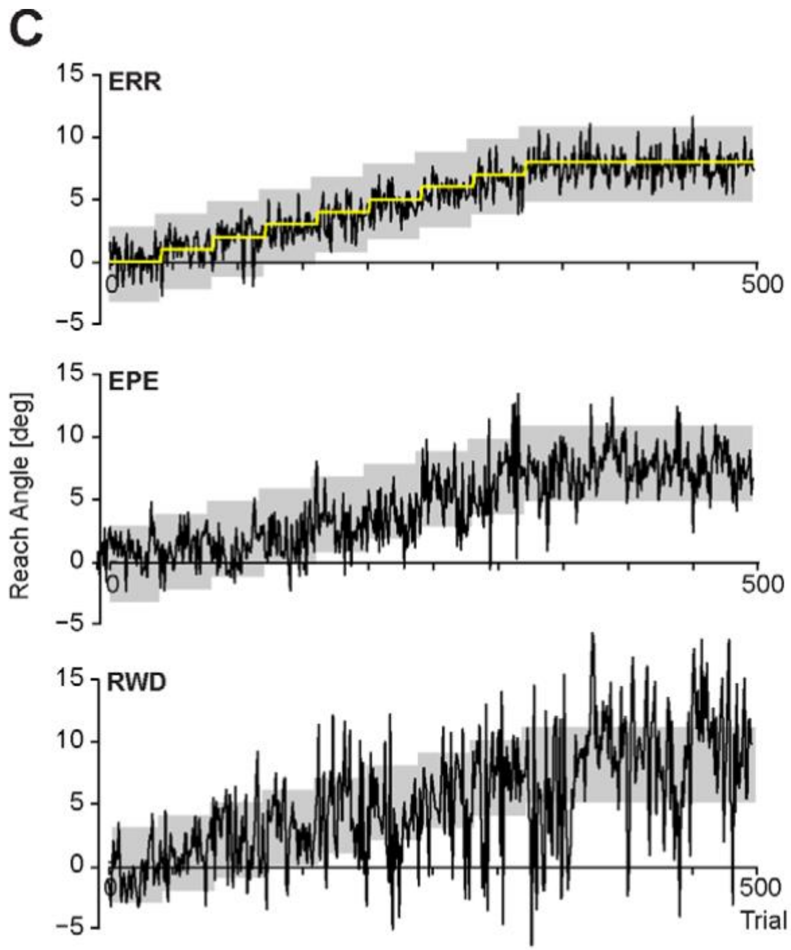
Endpoint feedback

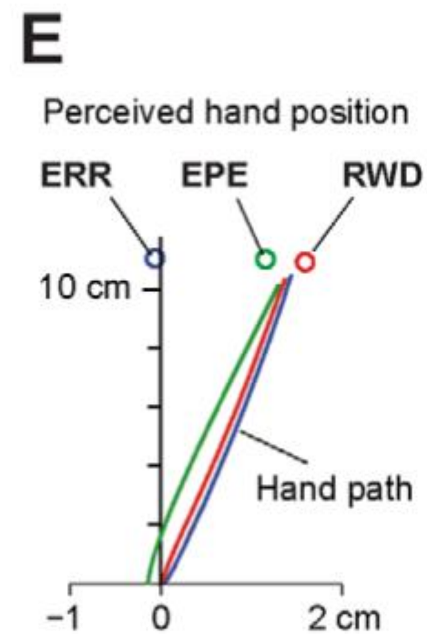
Reward feedback



Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Learning from errors and rewards





Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Model the movements

Hand and cursor

$$h^{(k+1)} = u^{(k)} + n_h^{(k)}$$
$$c^{(k)} = h^{(k)} + p^{(k)}$$
$$y^{(k)} = c^{(k)} + n_y^{(k)}$$

h: hand

u: motor command

c: cursor

p: perturbation

y: sensed position

n: noise

Internal estimates

$$\hat{h}^{(k+1)} = \hat{p}^{(k)} + u^{(k)}$$
$$\hat{p}^{(k+1)} = a\hat{p}^{(k)} + n_p^{(k)}$$

a: forgetting factor

Vectorization

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$
$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

$\mathbf{x}^{(k)} = \begin{pmatrix} p^{(k)} \\ h^{(k)} \end{pmatrix}$: state vector

$$\mathbf{n}_x^{(k)} = \begin{pmatrix} n_p^{(k)} \\ n_h^{(k)} \end{pmatrix}$$

$$A = \begin{pmatrix} a & 0 \\ 1 & 0 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Kalman filter as error based learning

State equations

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$
$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

$$\mathbf{x}^{(k)} = \begin{pmatrix} p^{(k)} \\ h^{(k)} \end{pmatrix} \quad \mathbf{n}_x^{(k)} = \begin{pmatrix} n_p^{(k)} \\ n_h^{(k)} \end{pmatrix}$$
$$A = \begin{pmatrix} a & 0 \\ 1 & 0 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Kalman filter

$$\hat{\mathbf{x}}^{(k|k-1)} = A\hat{\mathbf{x}}^{(k-1|k-1)} + \mathbf{b}u^{(k)}$$
$$\hat{\mathbf{x}}^{(k|k)} = \hat{\mathbf{x}}^{(k|k-1)} + K^{(k)}\left(y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}\right)$$

$y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}$: sensory prediction error

$K^{(k)}$: Kalman gain

Kalman gain

$$K^{(k)} = P^{(k|k-1)}C^T\left(CP^{(k|k-1)}C^T + \sigma_y^2\right)^{-1}$$

$P^{(k|k-1)}$: Variance of state estimate

Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

State variance

State equations

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$
$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

Variance propagation

$$P^{(k|k-1)} = AP^{(k-1|k-1)}A^T + Q^{(k)}$$
$$P^{(k|k)} = \left(I - K^{(k)}C\right)P^{(k|k-1)}$$

Kalman filter

$$\hat{\mathbf{x}}^{(k|k-1)} = A\hat{\mathbf{x}}^{(k-1|k-1)} + \mathbf{b}u^{(k)}$$
$$\hat{\mathbf{x}}^{(k|k)} = \hat{\mathbf{x}}^{(k|k-1)} + K^{(k)}\left(y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}\right)$$
$$K^{(k)} = P^{(k|k-1)}C^T\left(CP^{(k|k-1)}C^T + \sigma_y^2\right)^{-1}$$

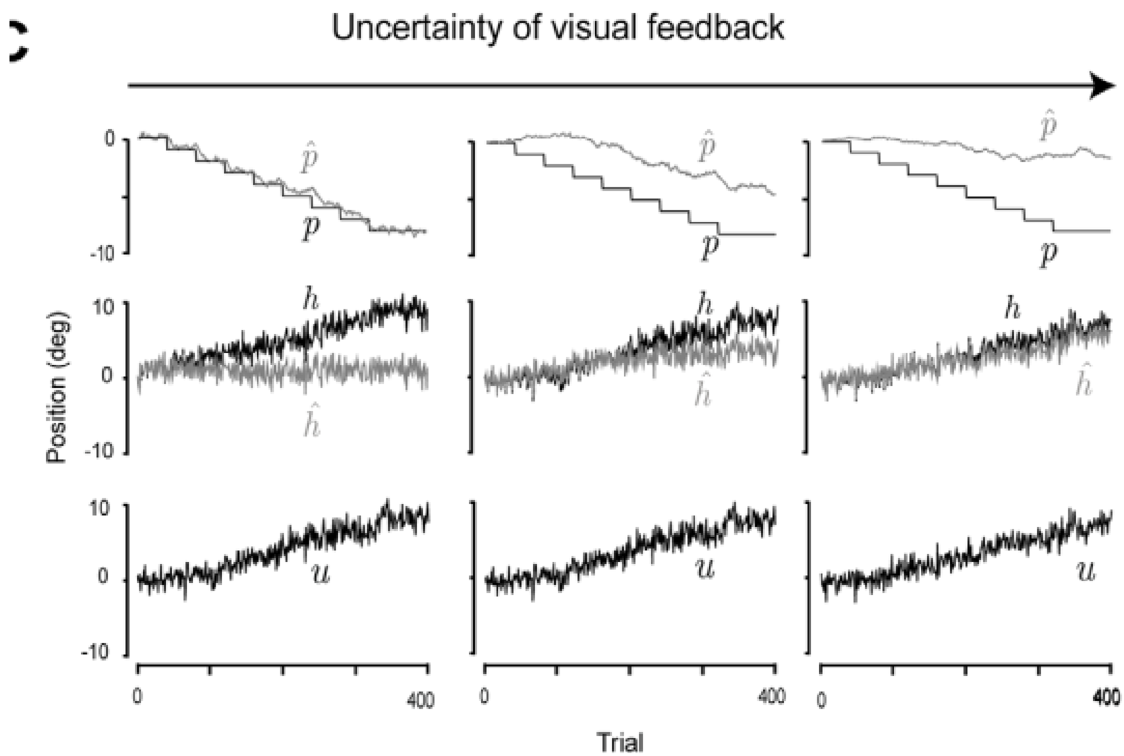
$$Q^{(k)} = \begin{pmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_p^2 \end{pmatrix}\text{: variance of state variables}$$

Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Credit assignment problem

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$
$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

$$\hat{\mathbf{x}}^{(k|k-1)} = A\hat{\mathbf{x}}^{(k-1|k-1)} + \mathbf{b}u^{(k)}$$
$$\hat{\mathbf{x}}^{(k|k)} = \hat{\mathbf{x}}^{(k|k-1)} + K^{(k)}\left(y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}\right)$$
$$K^{(k)} = P^{(k|k-1)}C^T\left(CP^{(k|k-1)}C^T + \sigma_y^2\right)^{-1}$$



Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Deciding on the movement

State equations

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$

$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

Kalman filter

$$\hat{\mathbf{x}}^{(k|k-1)} = A\hat{\mathbf{x}}^{(k-1|k-1)} + \mathbf{b}u^{(k)}$$

$$\hat{\mathbf{x}}^{(k|k)} = \hat{\mathbf{x}}^{(k|k-1)} + K^{(k)}\left(y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}\right)$$

$$K^{(k)} = P^{(k|k-1)}C^T\left(CP^{(k|k-1)}C^T + \sigma_y^2\right)^{-1}$$

Variance propagation

$$P^{(k|k-1)} = AP^{(k-1|k-1)}A^T + Q^{(k)}$$

$$P^{(k|k)} = \left(I - K^{(k)}C\right)P^{(k|k-1)}$$

Control policy

$$u^{(k)} = -\hat{p}^{(k)} + w_r^{(k)} + n_u^{(k)}$$

$w_r^{(k)}$: reward-maximizing policy

Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Reward based learning

State equations

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} + \mathbf{b}u^{(k)} + \mathbf{n}_x^{(k)}$$

$$y^{(k)} = C\mathbf{x}^{(k)} + n_y^{(k)}$$

Kalman filter

$$\hat{\mathbf{x}}^{(k|k-1)} = A\hat{\mathbf{x}}^{(k-1|k-1)} + \mathbf{b}u^{(k)}$$

$$\hat{\mathbf{x}}^{(k|k)} = \hat{\mathbf{x}}^{(k|k-1)} + K^{(k)}\left(y^{(k)} - C\hat{\mathbf{x}}^{(k|k-1)}\right)$$

Control policy

$$u^{(k)} = -\hat{p}^{(k)} + w_r^{(k)} + n_u^{(k)}$$

Policy learning

$$w_r^{(k+1)} = w_r^{(k)} + \alpha_r \delta^{(k)}$$

$$\delta^{(k)} = r^{(k)} + \gamma \hat{V}^{(k+1)} - \hat{V}^{(k)}$$

$\delta^{(k)}$: reward prediction error

$V^{(k)}$: Discounted accumulated reward

$\hat{V}^{(k)}$: Estimate of accumulated reward

γ: Discounting factor

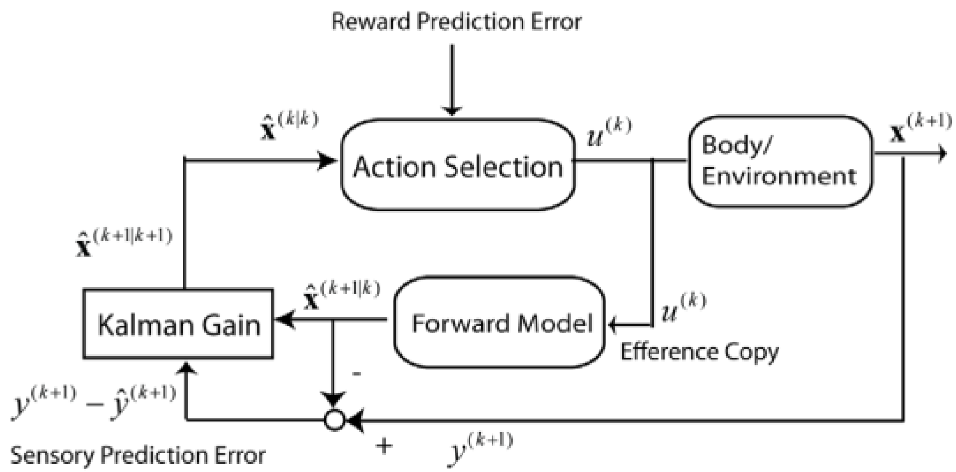
Reward learning

$$V^{(k)} = E\left[r^{(k)} + \gamma r^{(k+1)} + \gamma^2 r^{(k+2)} + \cdots\right]$$

$$\hat{V}^{(k+1)} = \hat{V}^{(k)} + \alpha_V \delta^{(k)}$$

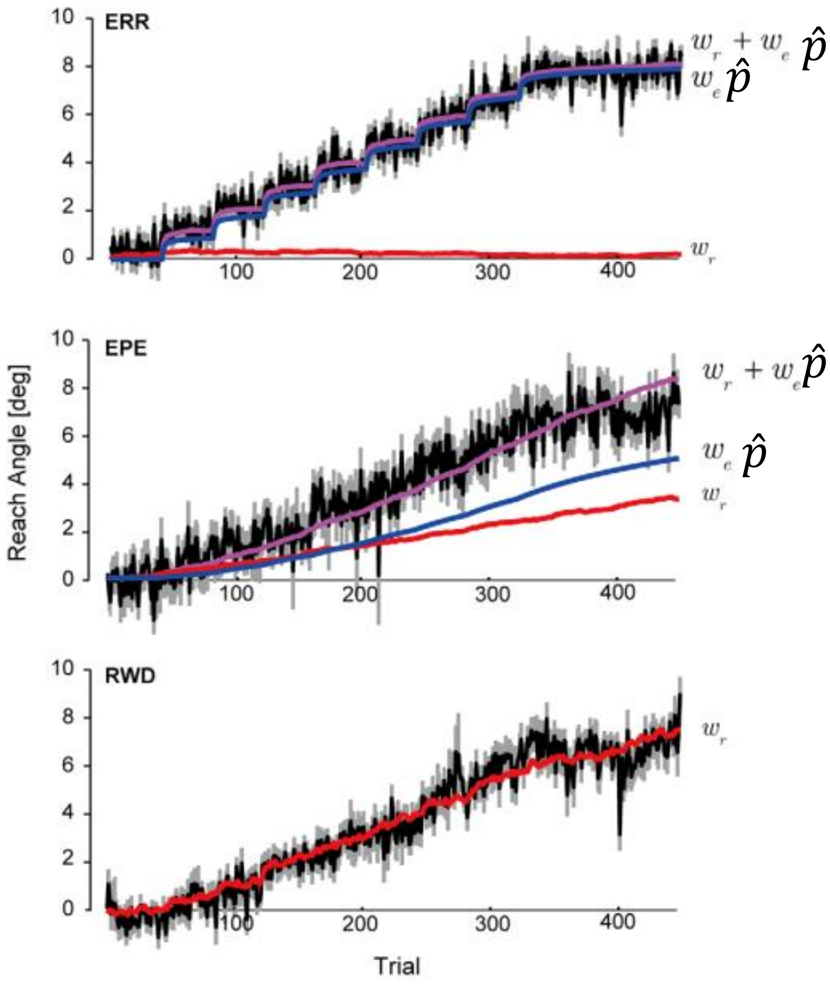
Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Box and arrow diagram of the model



Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Model successfully learns from error or reward

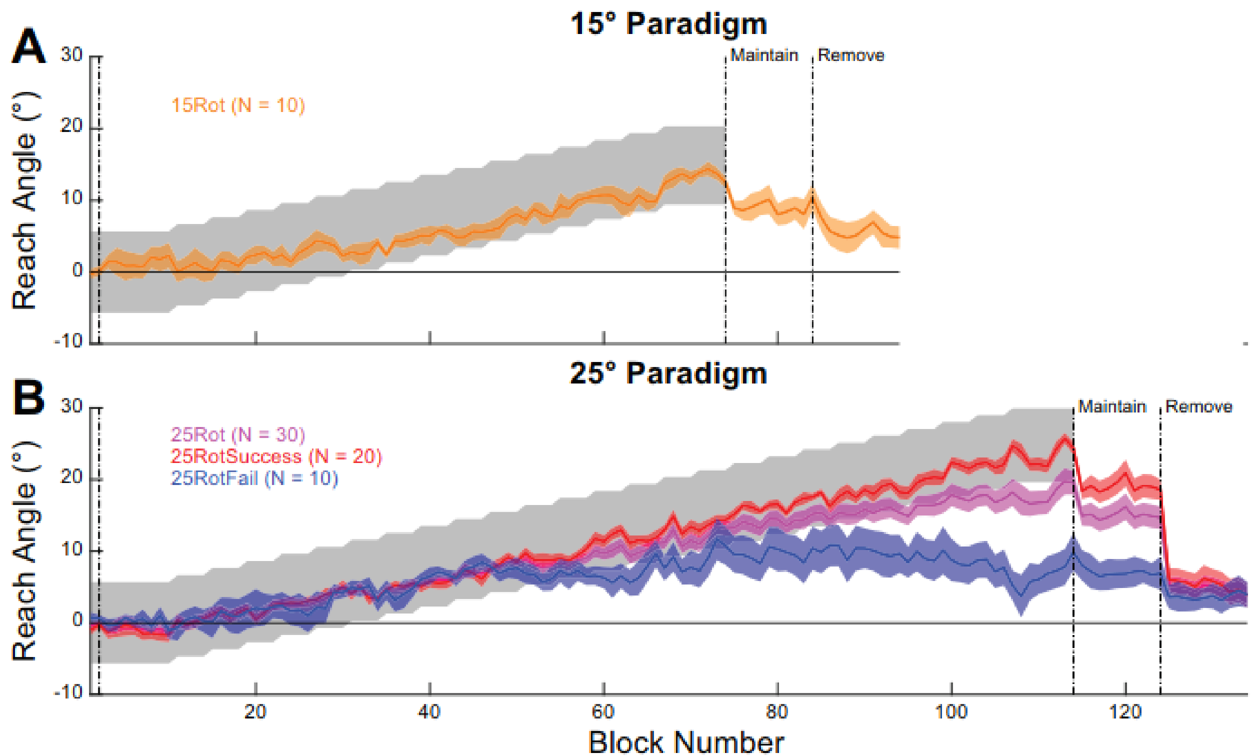


Izawa *PLOS Computational Biology* 2011 doi:10.1371/journal.pcbi.1002012

Part 2

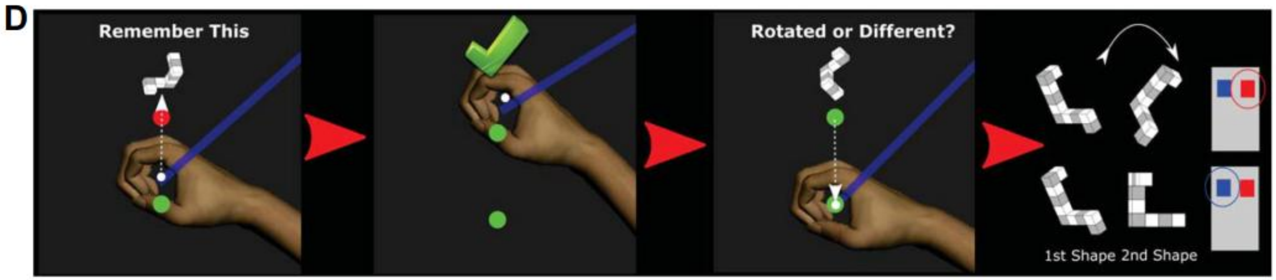
- Reinforcement and explicit learning

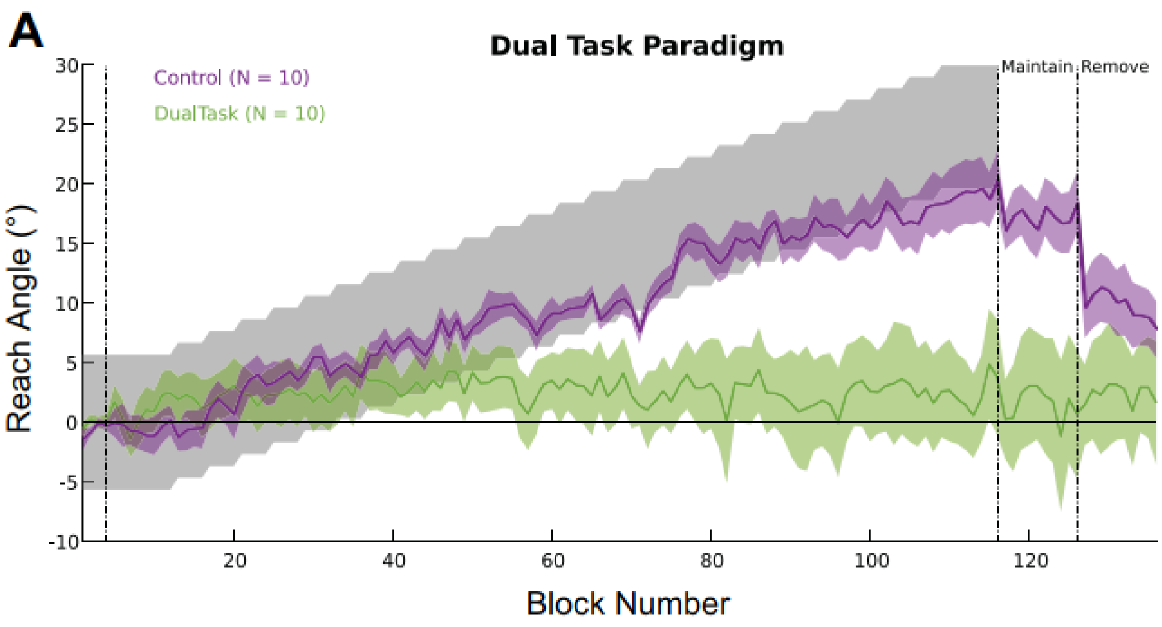
Reinforcement learning can be "turned off"



Holland *J Neurophysiol* 2018 doi:10.1152/jn.00901.2017

Dual task wipes out reinforcement learning

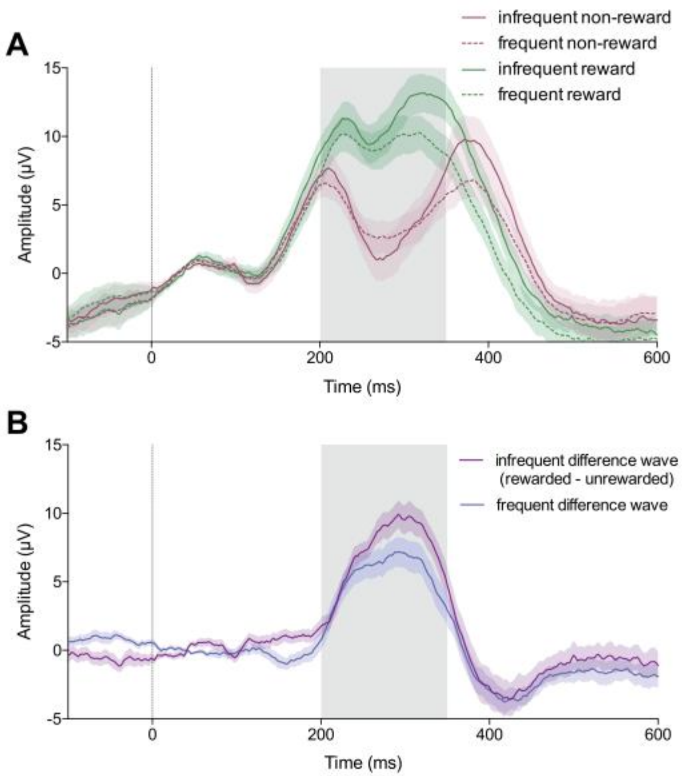




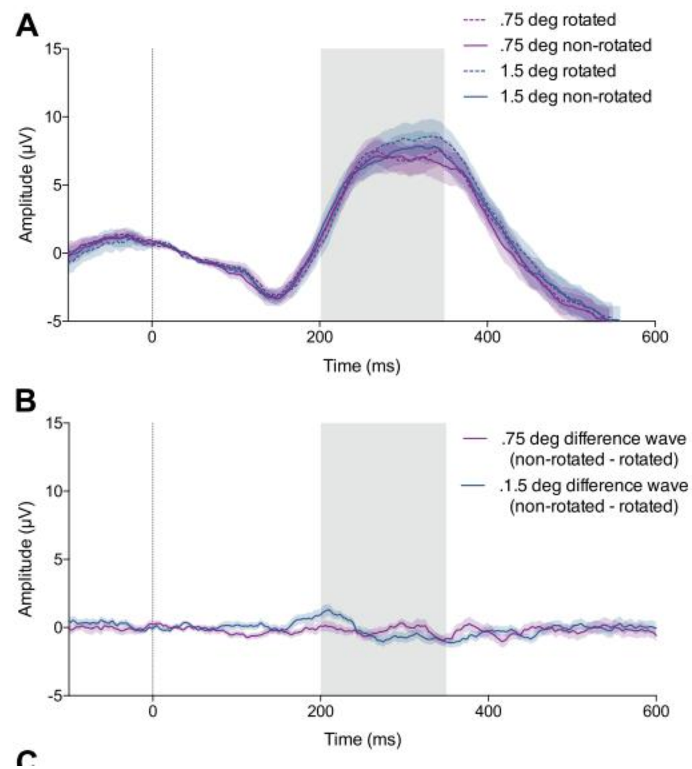
Holland *J Neurophysiol* 2018 doi:10.1152/jn.00901.2017

Different EEG for reward and error based learning

Reward based

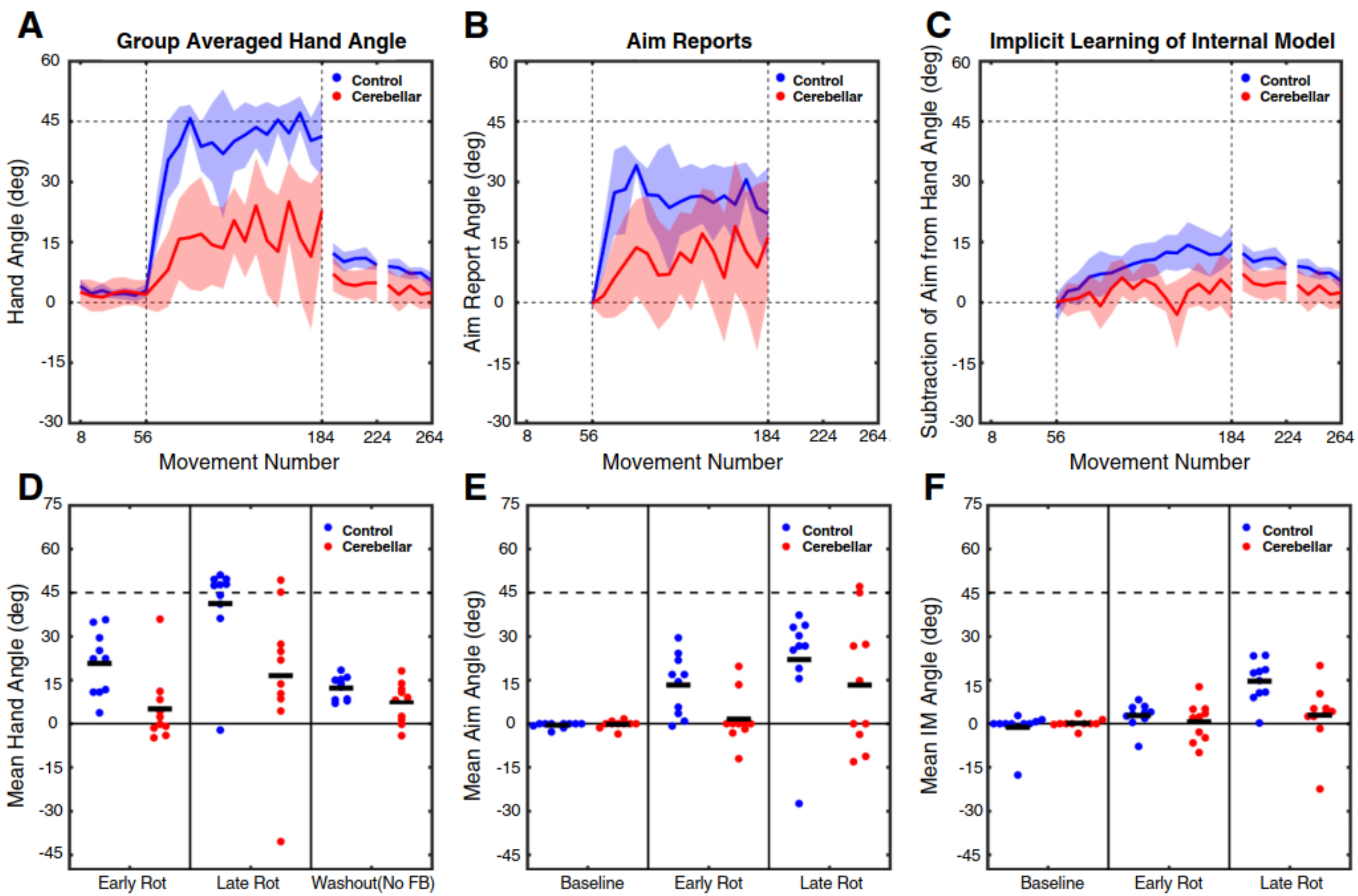


Error based



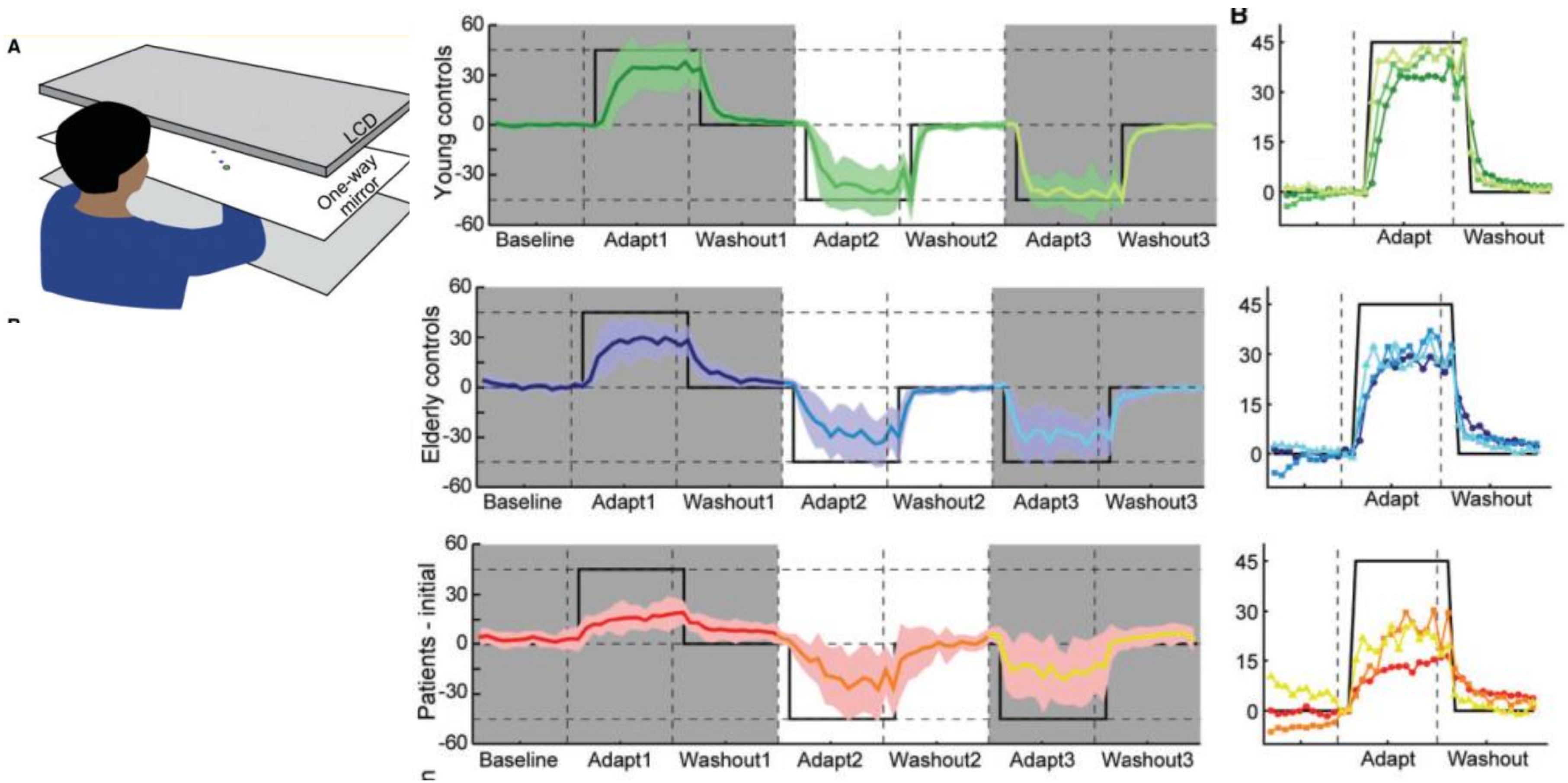
Palidis *J Neurophysiol* 2019 doi:10.1152/jn.00792.2018

But cerebellar patients have poor explicit AND implicit



Butcher *J Neurophysiol* 2017 doi:10.1152/jn.00451.2017

Can be rescued by "providing" explicit strategy

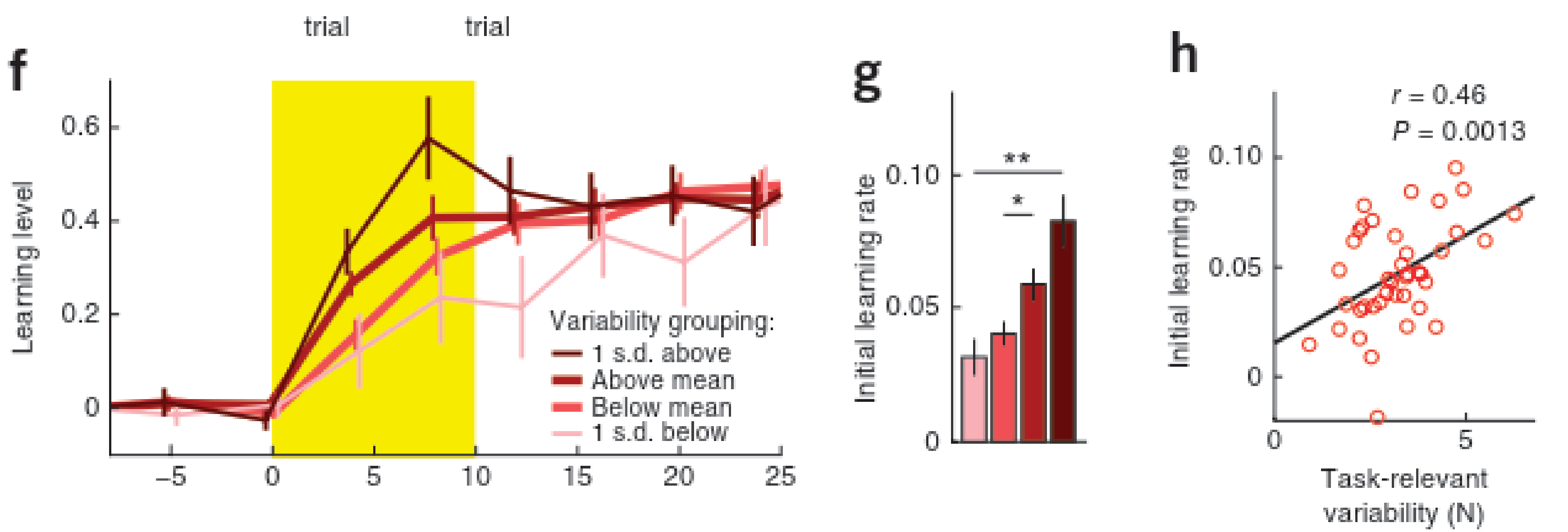


Wong *J Brain* 2019 doi:10.1093/brain/awz020

Part 3

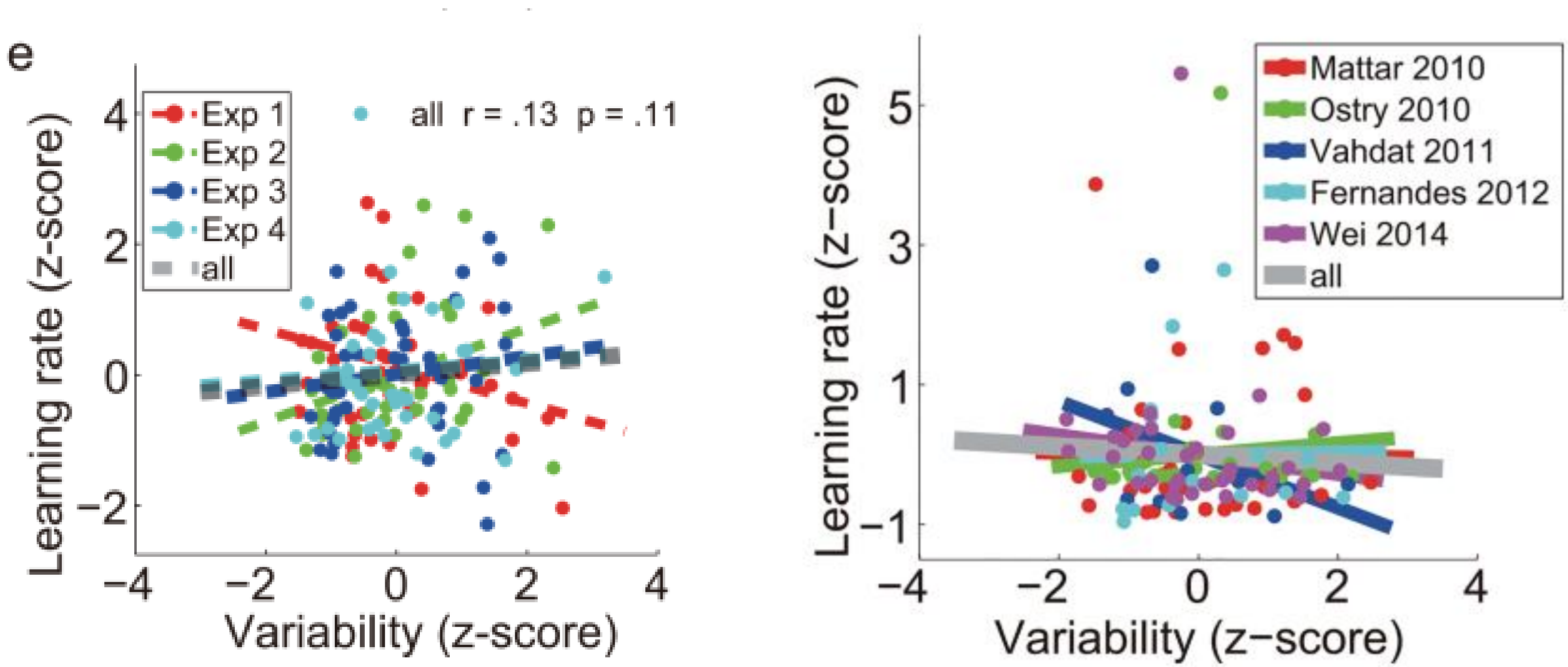
- Exploration and noise

Variability is correlated to learning



Wu et al, Nature Neuroscience, 2014

But not in all experiments



He et al, PLOS Computational Biology, 2016

Explanation: different types of noise

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k y_k$$ Learning

$$K_k = P_{k|k-1} H^T S_k^{-1}$$ Learning rate

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1}$$ Forgetting

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q$$ Planning noise

$$y_k = Z_k - H\hat{x}_{k|k-1}$$ Behavior

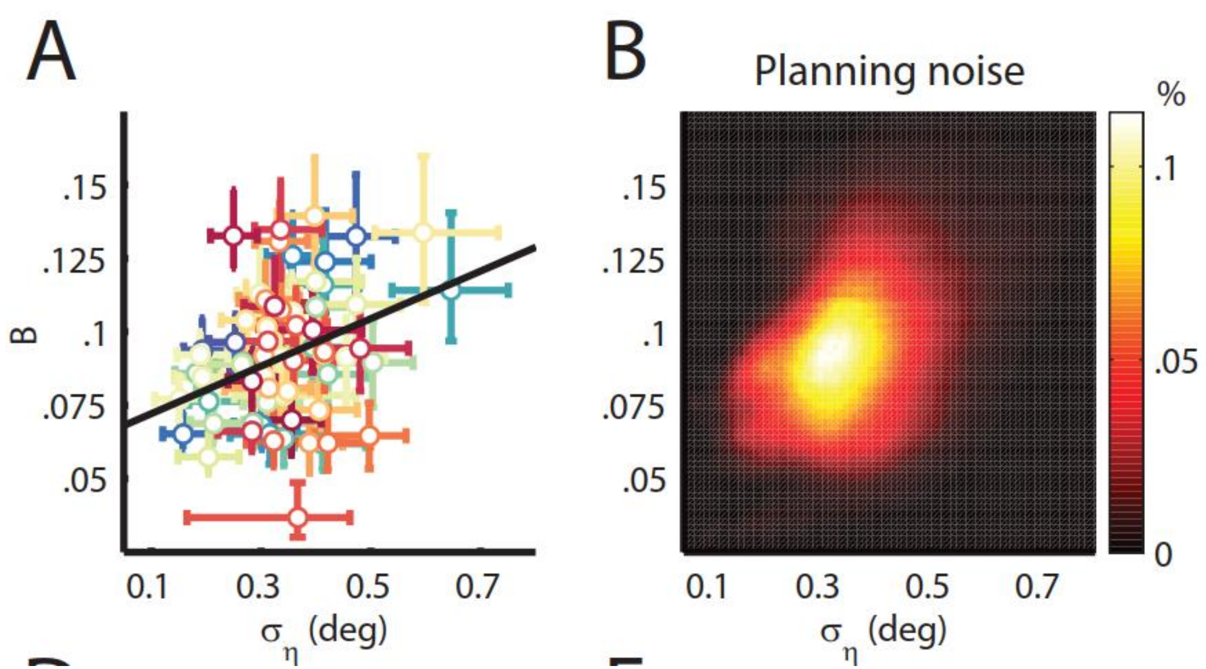
$$S_k = HP_{k|k-1}H^T + R$$ Execution noise

Kalman filter models predict learning will be:

- Correlated with planning noise
- Inversely correlated with execution nosie

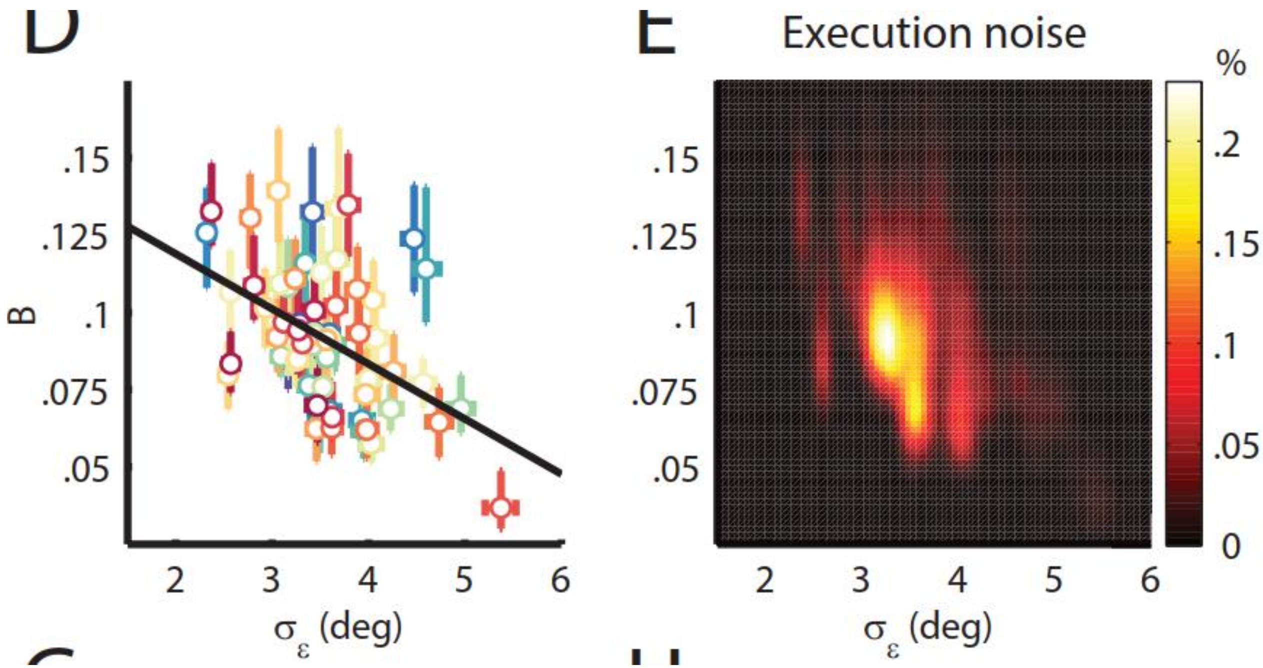
He et al, PLOS Computational Biology, 2016

Planning noise correlated with learning



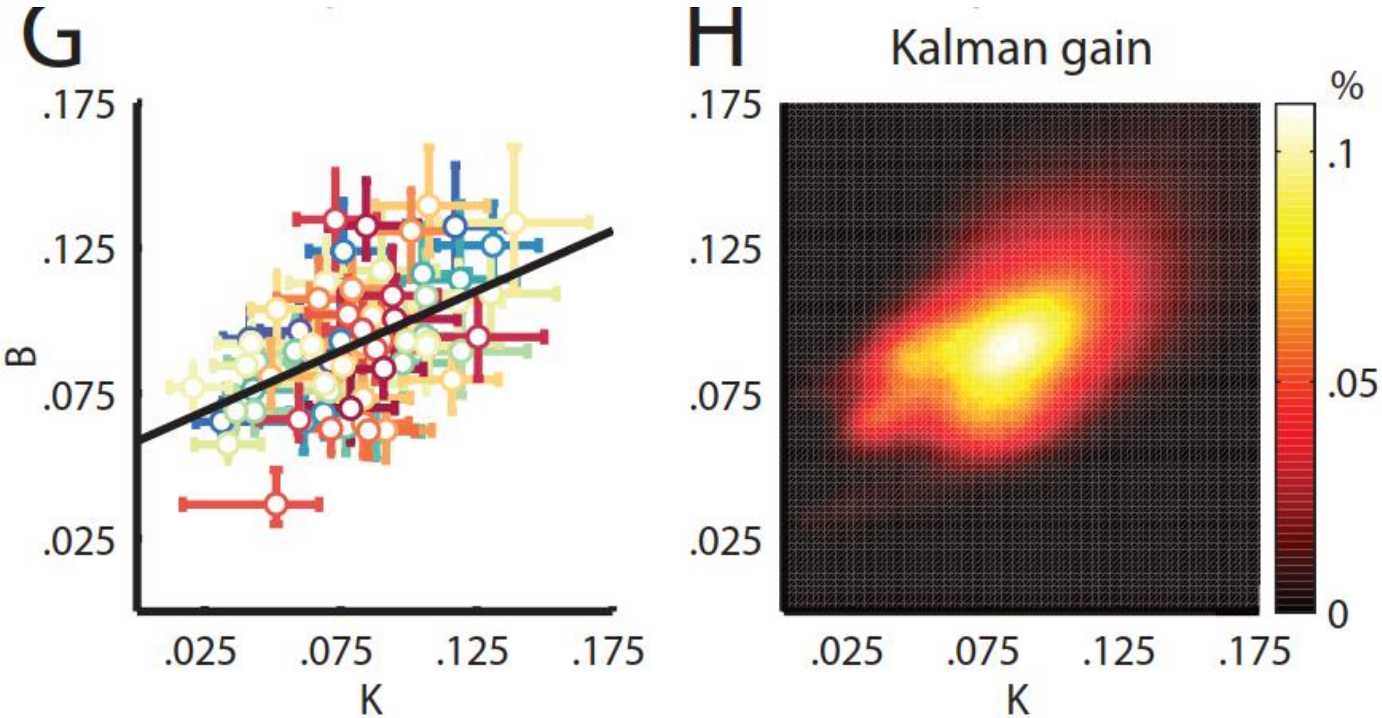
Van Der Vliet eNeuro 2018 doi:10.1523/ENEURO.0170-18.2018

Execution noise inversely correlated



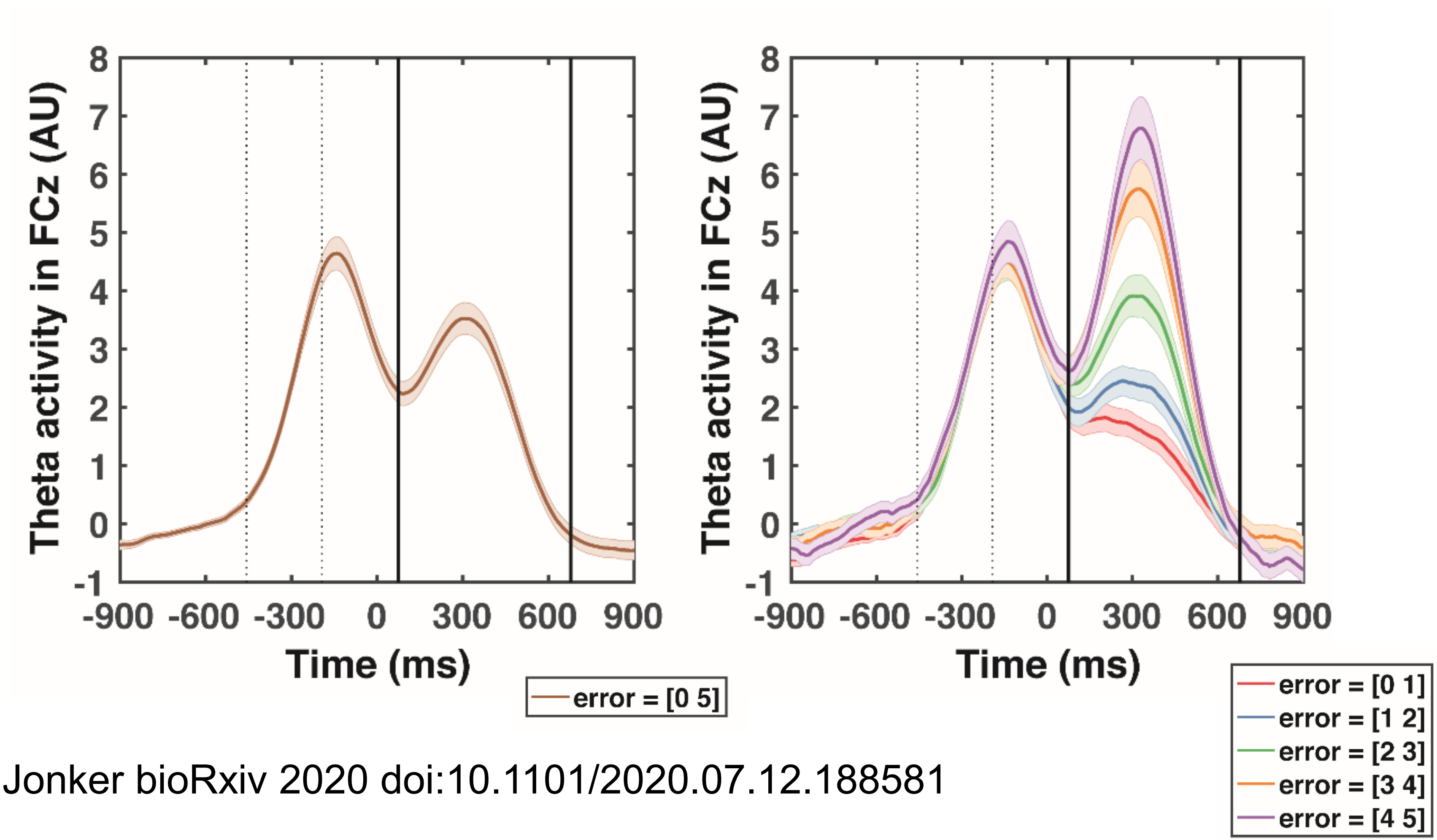
Van Der Vliet eNeuro 2018 doi:10.1523/ENEURO.0170-18.2018

Optimal Kalman gain predicts learning



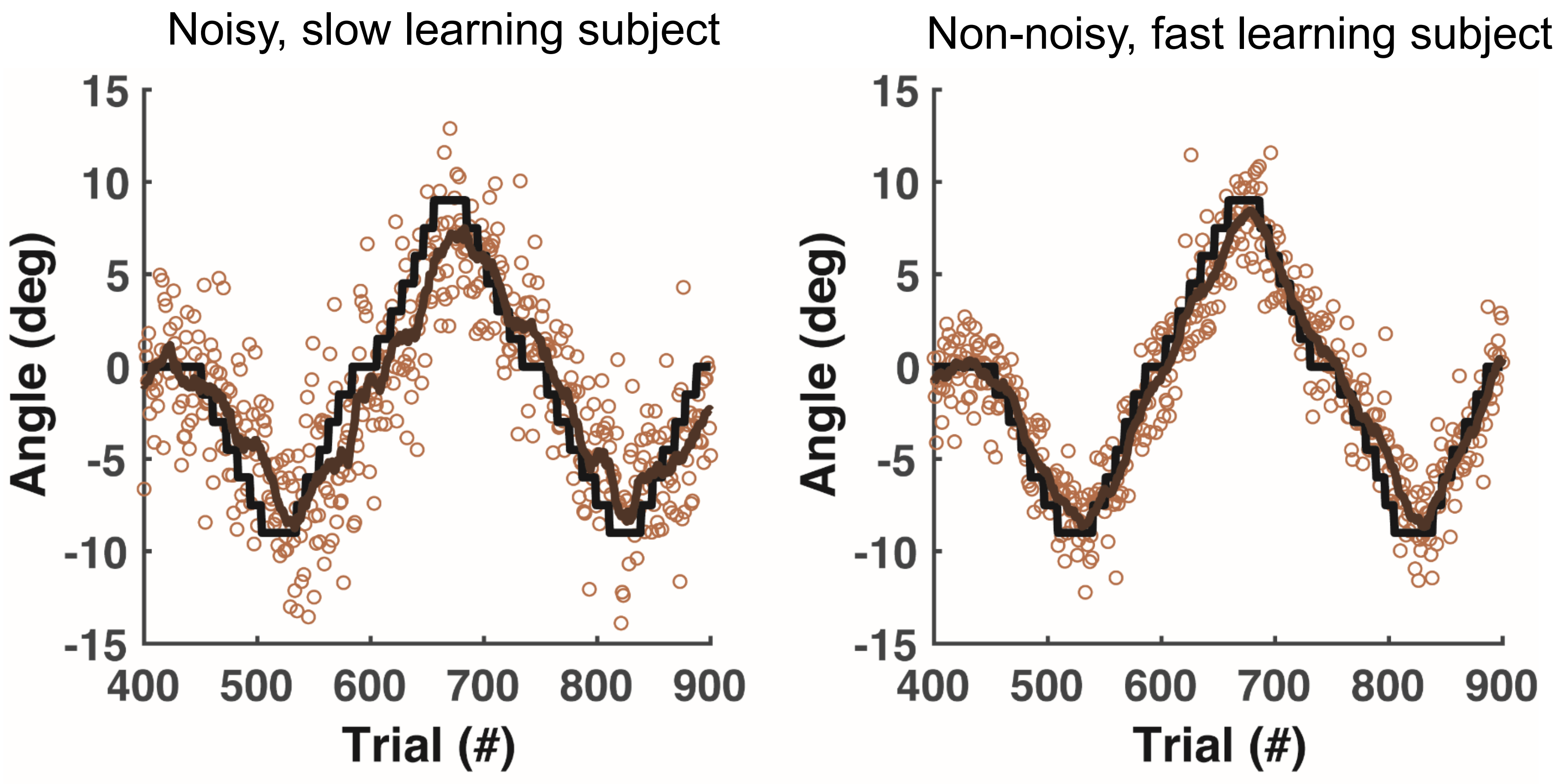
Van Der Vliet eNeuro 2018 doi:10.1523/ENEURO.0170-18.2018

Error related negativity
400 ms latency



Jonker bioRxiv 2020 doi:10.1101/2020.07.12.188581

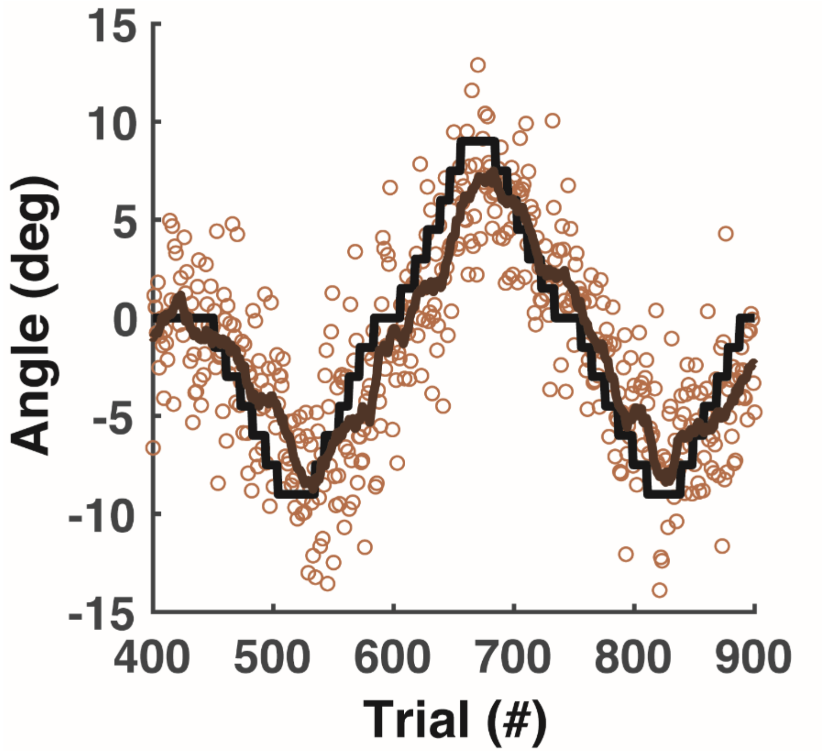
Noisier subjects learn slower



Jonker bioRxiv 2020 doi:10.1101/2020.07.12.188581

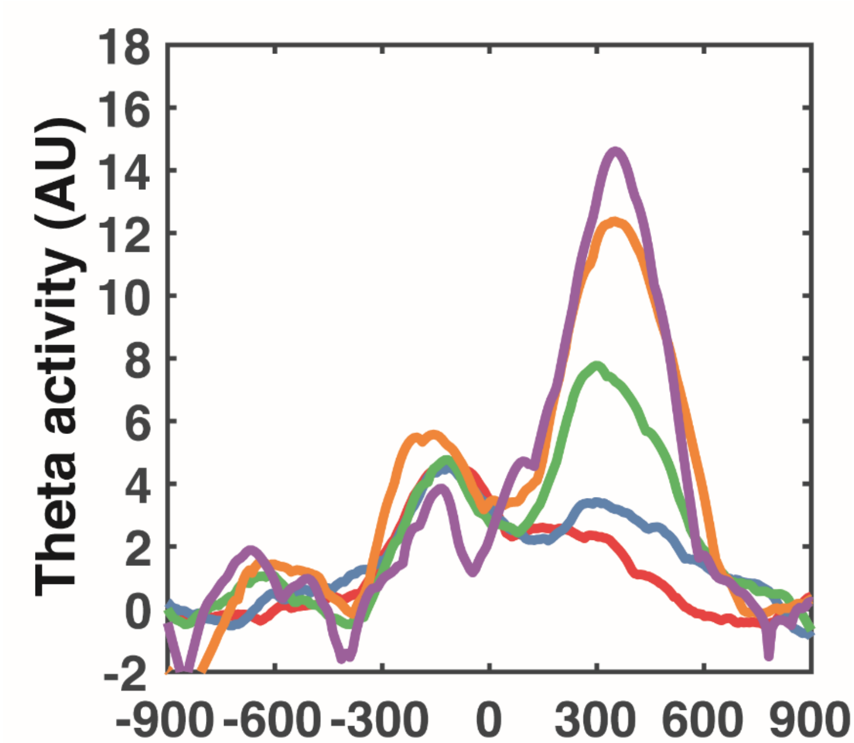
Less noisy subject has sensitivity to error

Noisy, slow learning subject



Non-noisy, fast learning subject

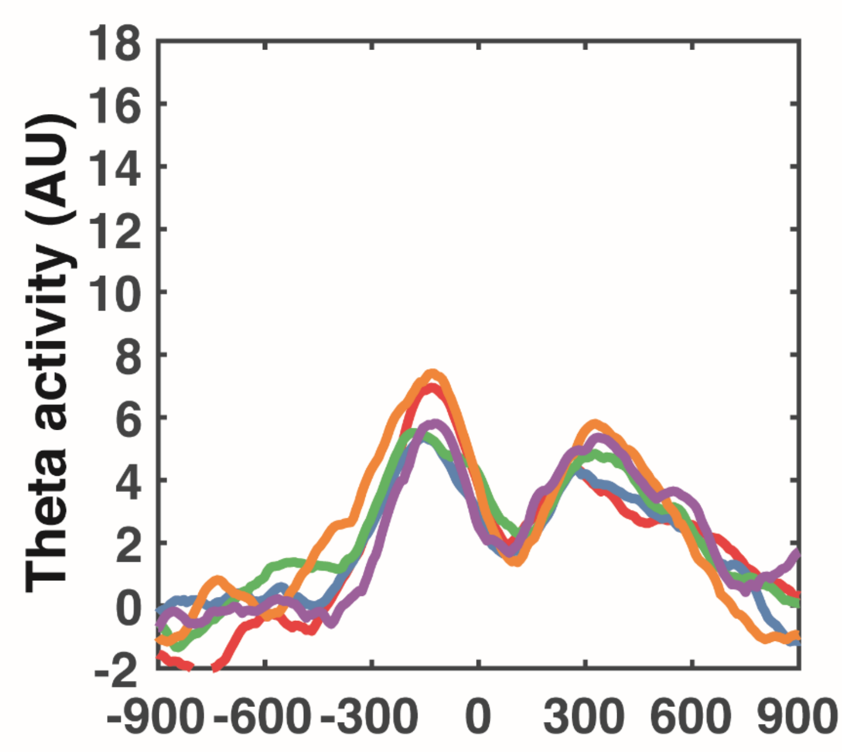
EEG highly sensitive to error



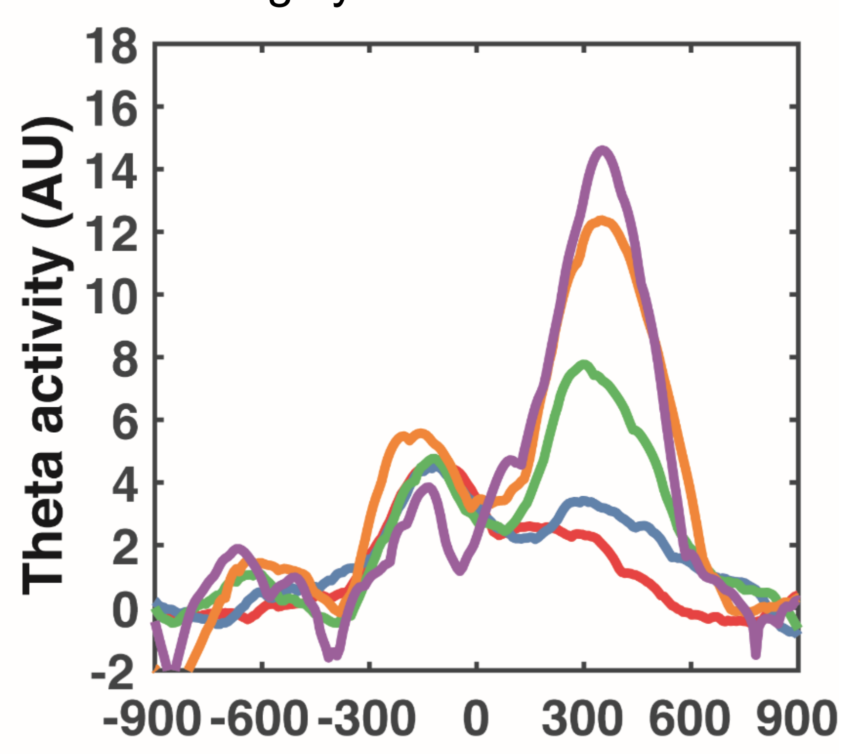
Jonker bioRxiv 2020 doi:10.1101/2020.07.12.188581

Noisier subject has no error sensitivity

Noisy, slow learning subject
EEG insensitive to error



Non-noisy, fast learning subject
EEG highly sensitive to error

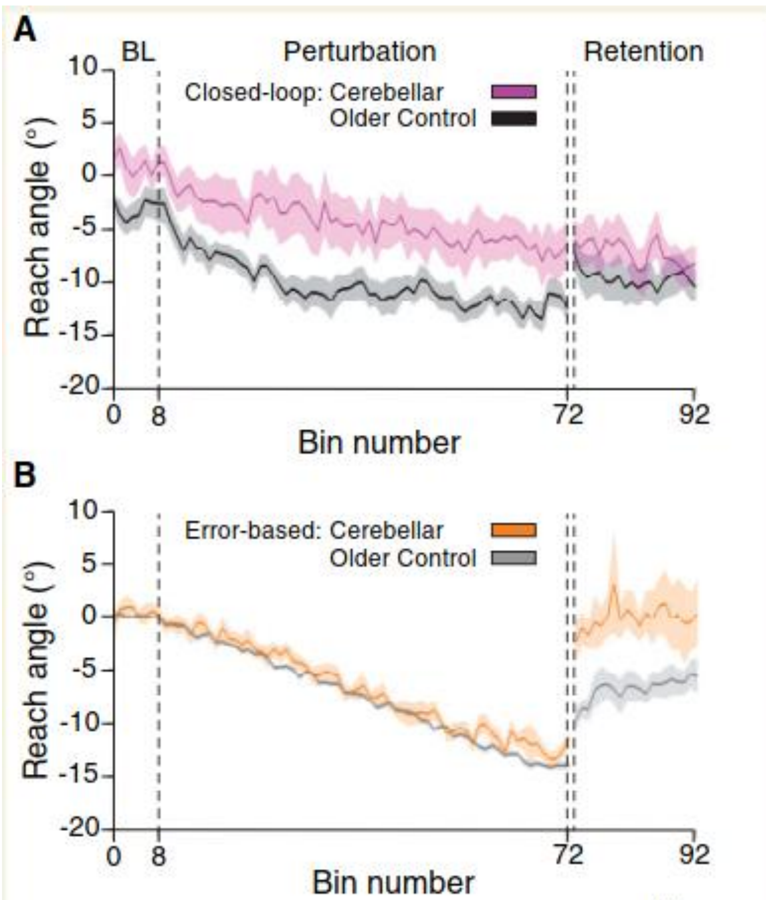


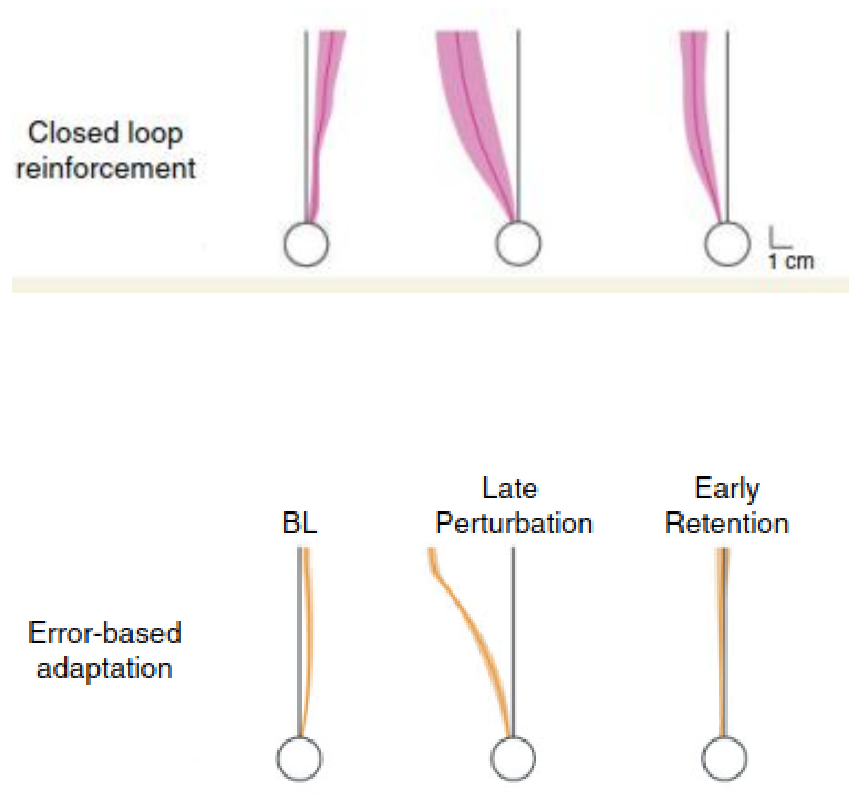
Jonker bioRxiv 2020 doi:10.1101/2020.07.12.188581

Part 3

- Noise, exploration and reinforcement learning

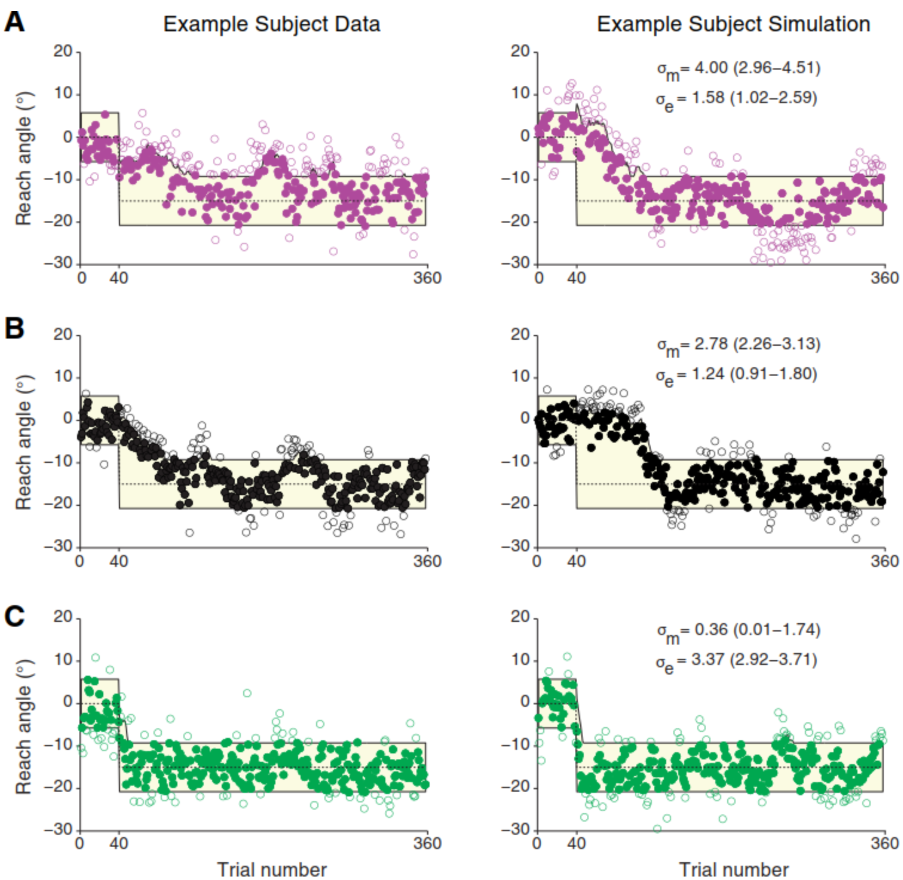
Cerebellar subjects forget error based learning

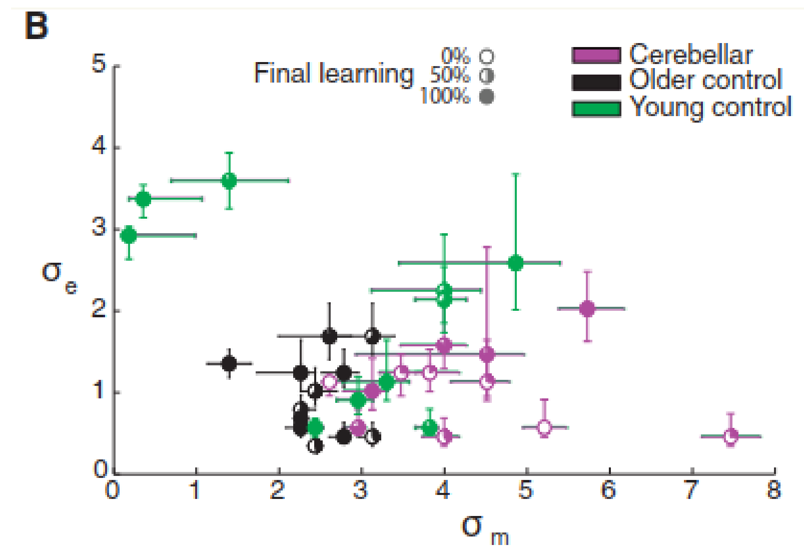




Therrien *Brain* 2016 doi:10.1093/brain/awv329

Cerebellar subjects: noise and reinforcement





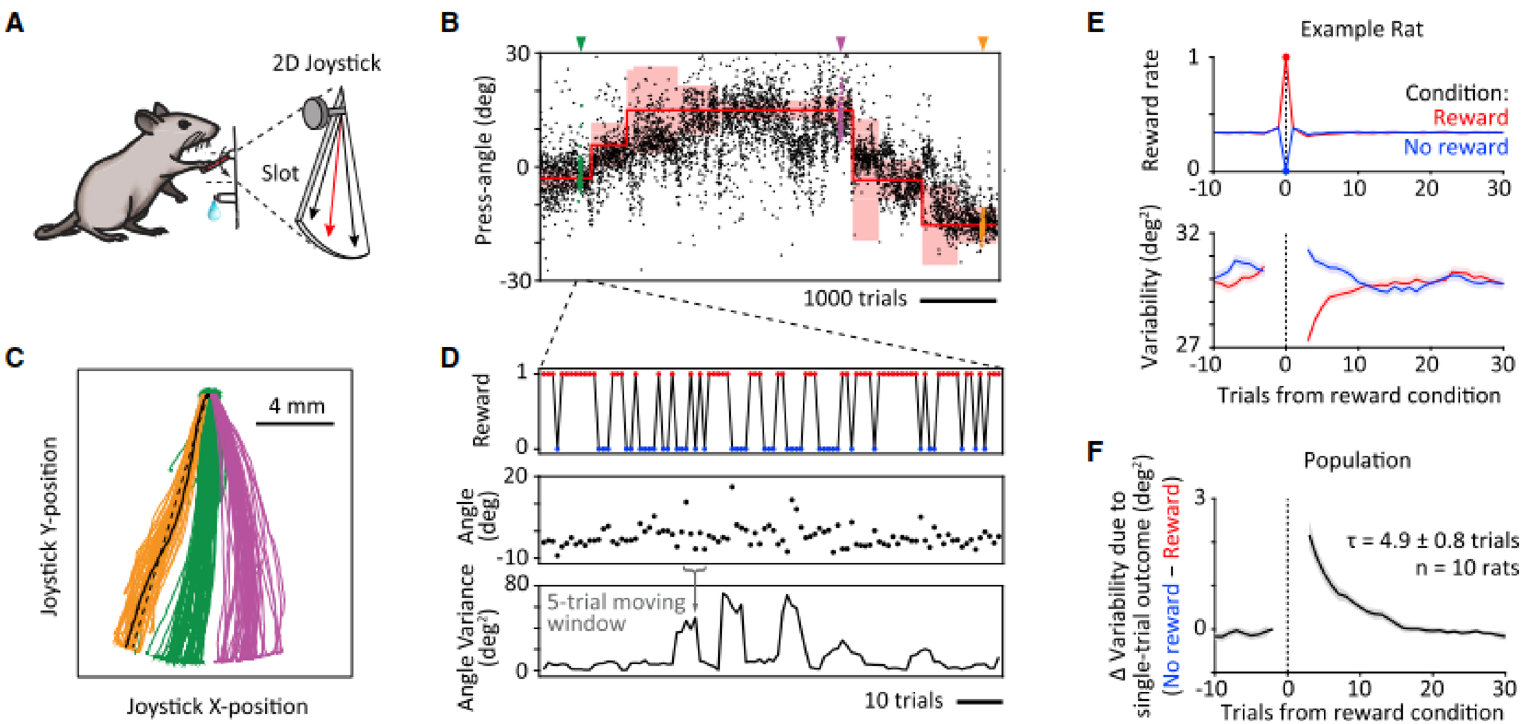
Therrien *Brain* 2016 doi:10.1093/brain/awv329

Reinforcement and variability

- Link to video

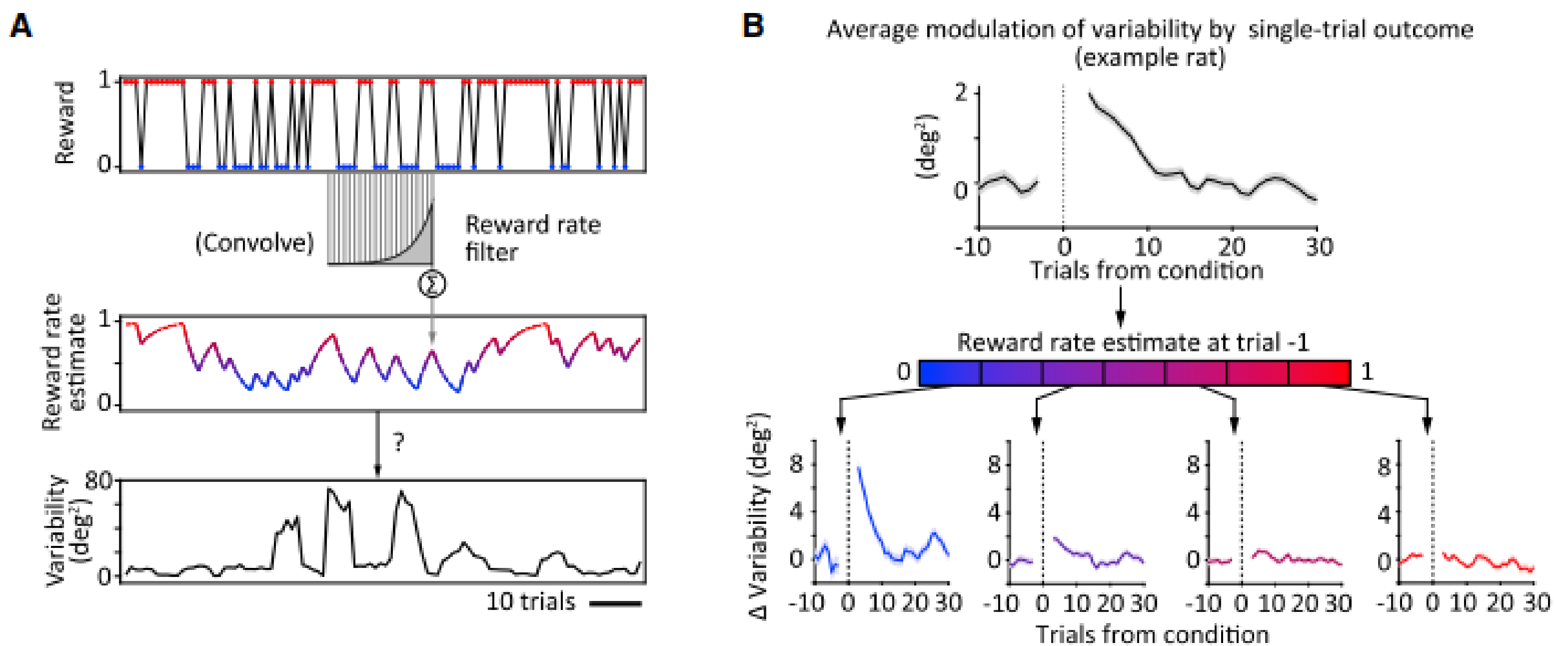
Dhawale *Curr Biol* 2019 doi:10.1016/j.cub.2019.08.052

Reward drives changes in variability



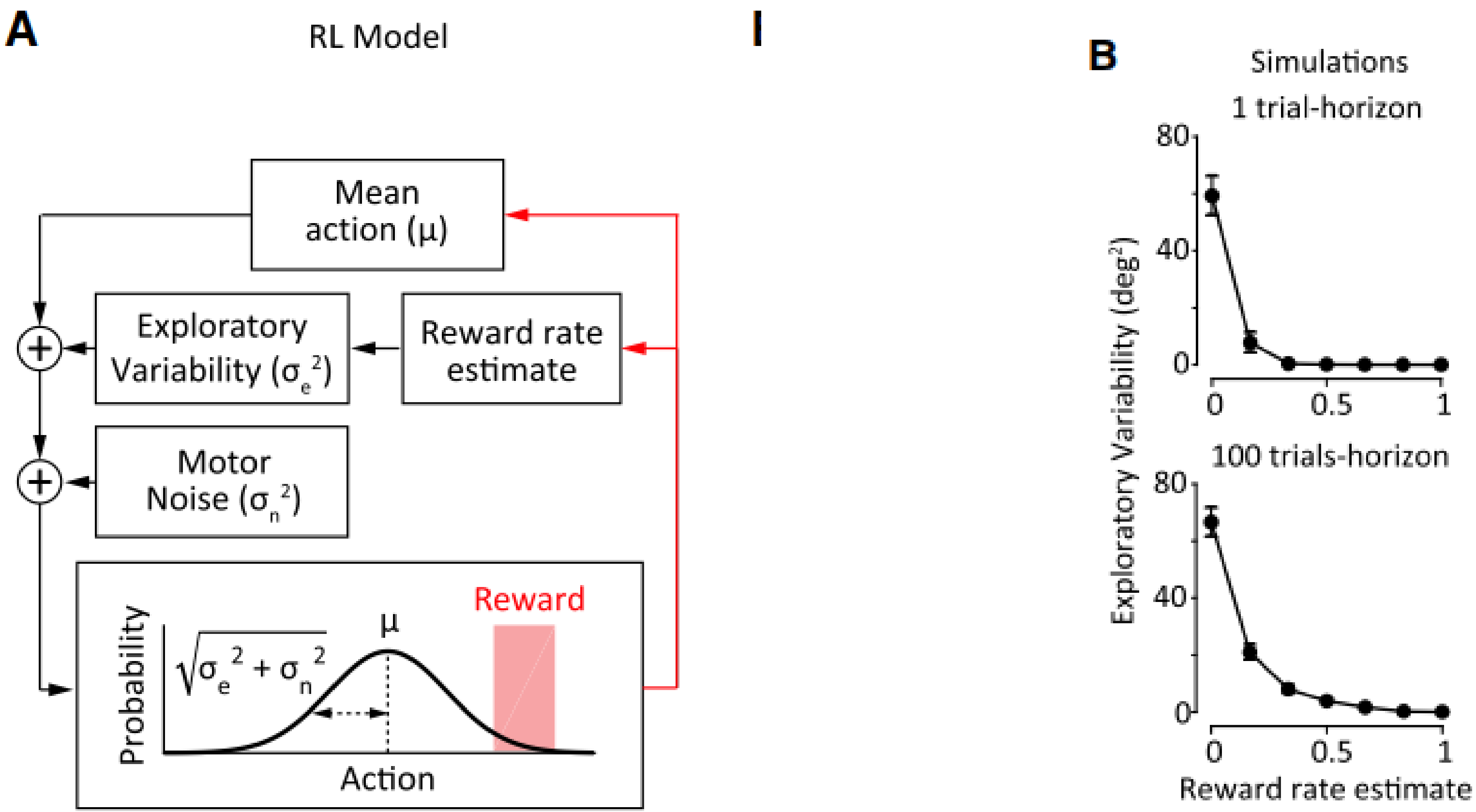
Dhawale *Curr Biol* 2019 doi:10.1016/j.cub.2019.08.052

Reward rate reduces changes in variability



Dhawale *Curr Biol* 2019 doi:10.1016/j.cub.2019.08.052

Making a model



Dhawale *Curr Biol* 2019 doi:10.1016/j.cub.2019.08.052