Motor adaptation

Rob van Beers

VU UNIVERSITY AMSTERDAM
Department of Human Movement Sciences

Donders Institute for Brain, Cognition and Behaviour

Motor learning

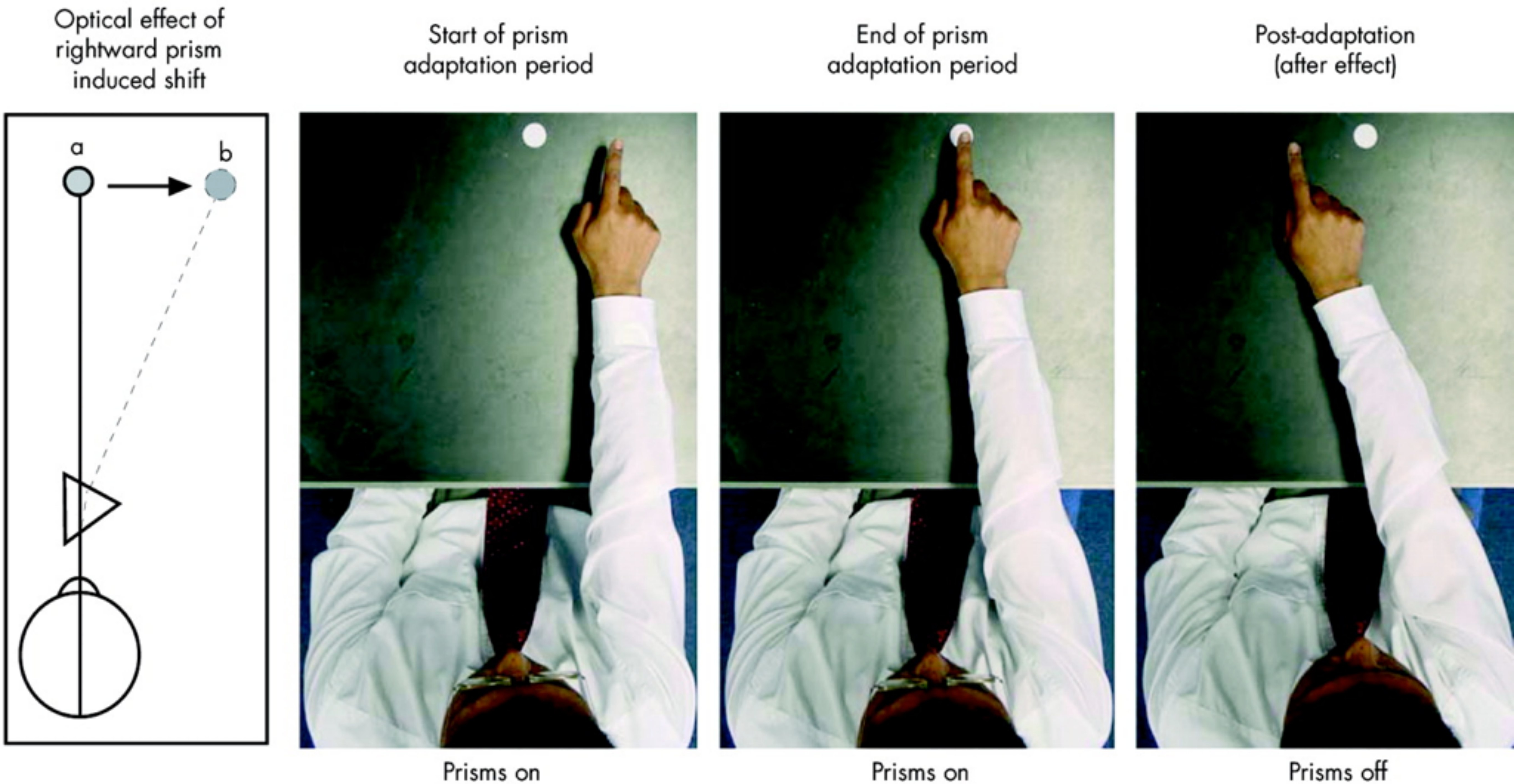
- **Motor learning** includes:
  - motor skill learning
  - motor adaptation
- **Motor skill learning**:
  - learning a new skill, like riding a bicycle
  - indices of learning: reduced variability, improved speed
  - learning can go on for years
- **Motor adaptation**:
  - adaptation of motor performance to a change in the environment
  - goal: reduction of perturbation-induced systematic error
  - mechanism: adjustment of existing forward model
  - time scale: minutes

Movement types

- **Motor adaptation** studied for various types of movement:
  - reaching
  - saccades
  - locomotion
  - throwing
  - …

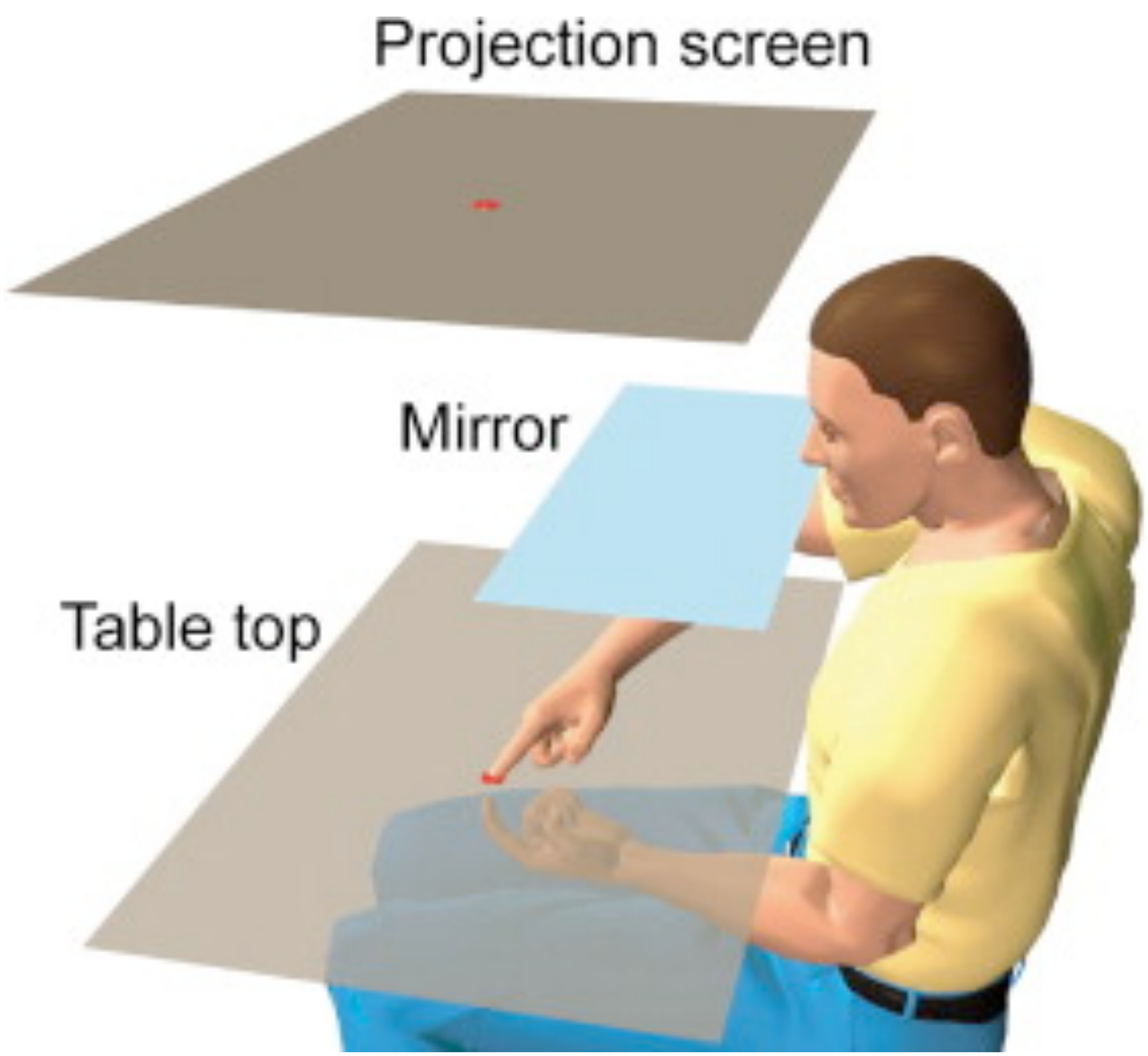
- **Here: focus on reaching**
  - Reaching modelled most extensively
  - Generally comparable results for other movement types

Prism adaptation



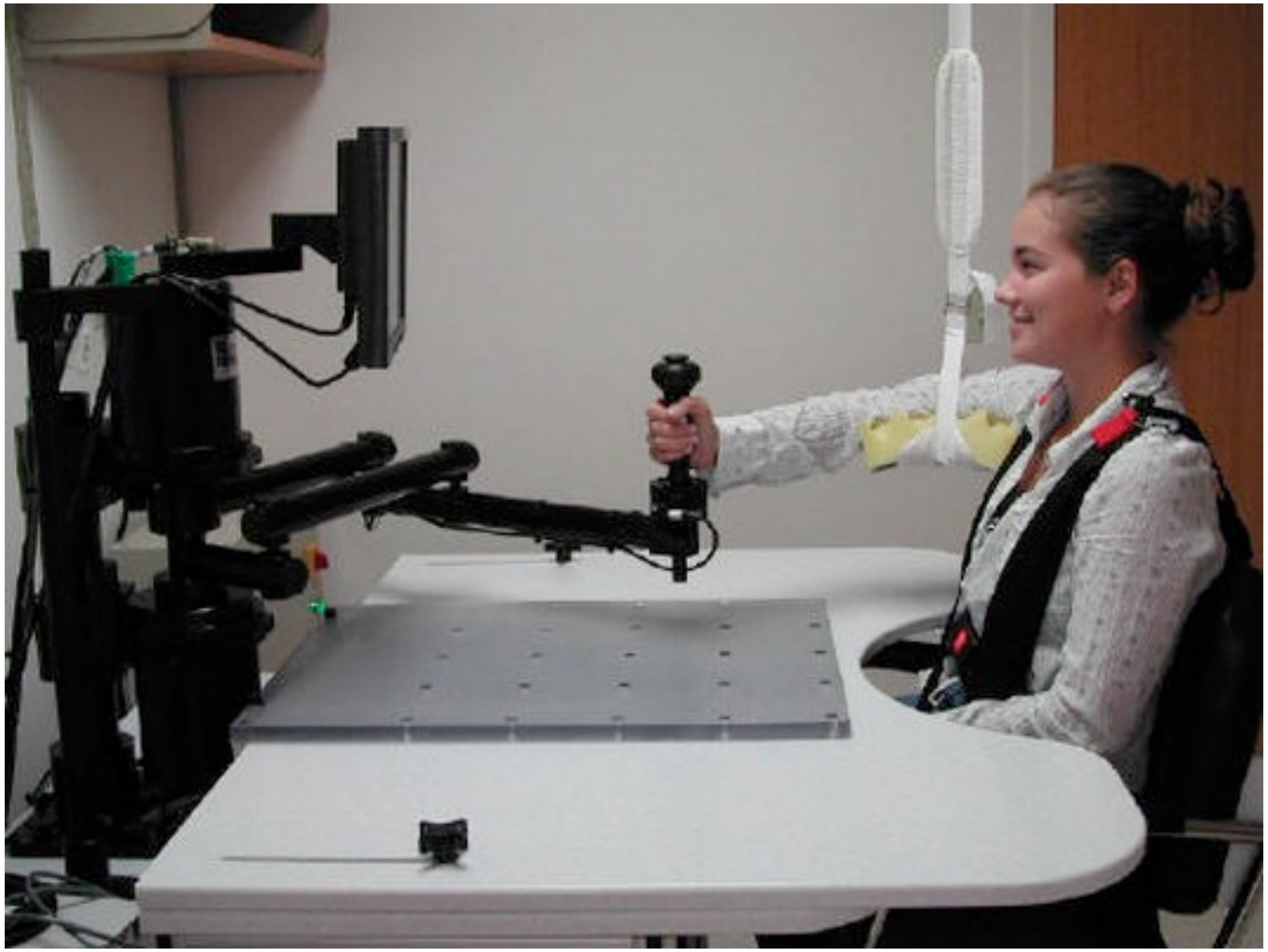
Visuomotor adaptation in VR

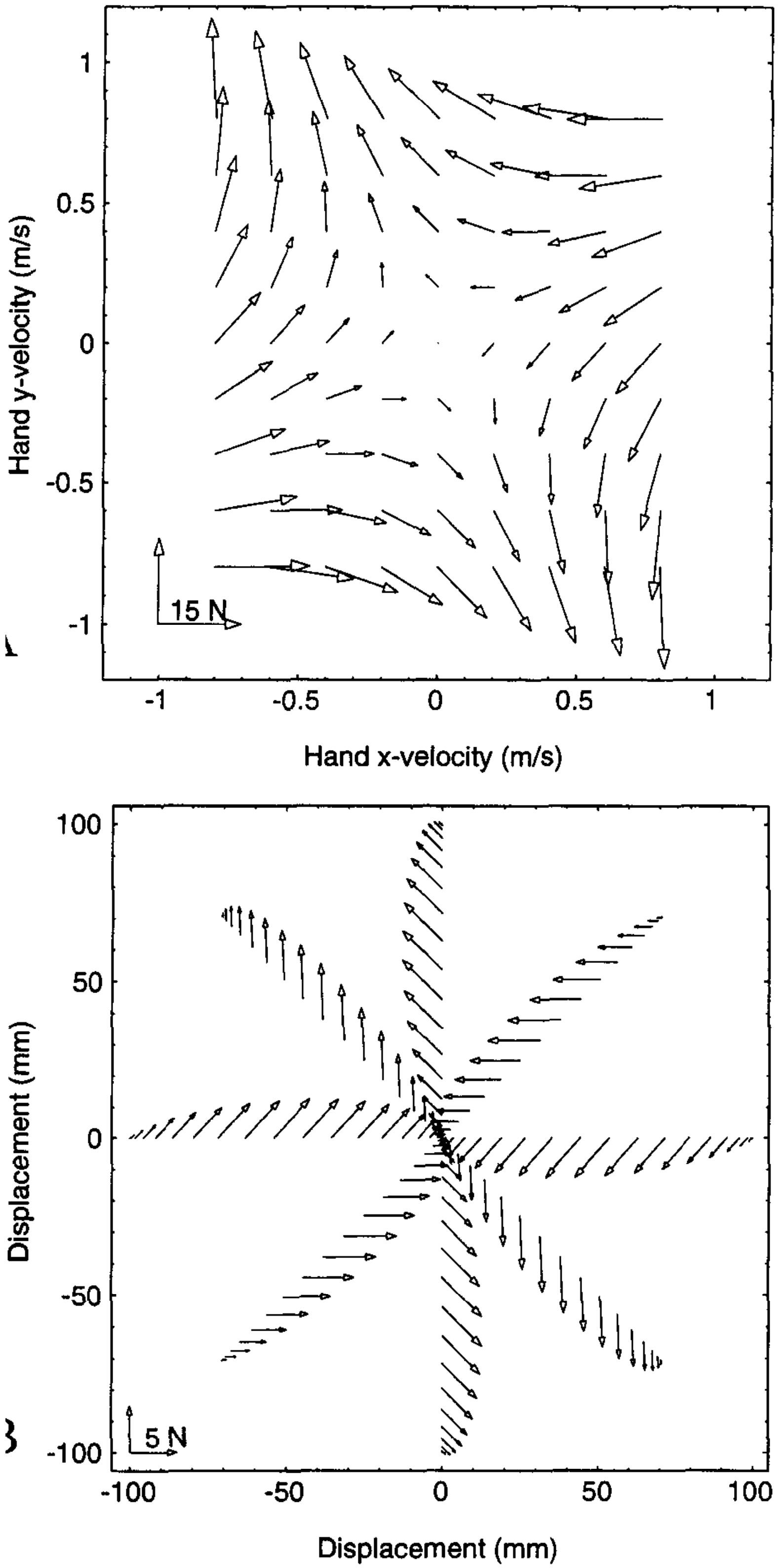
- In 2D:



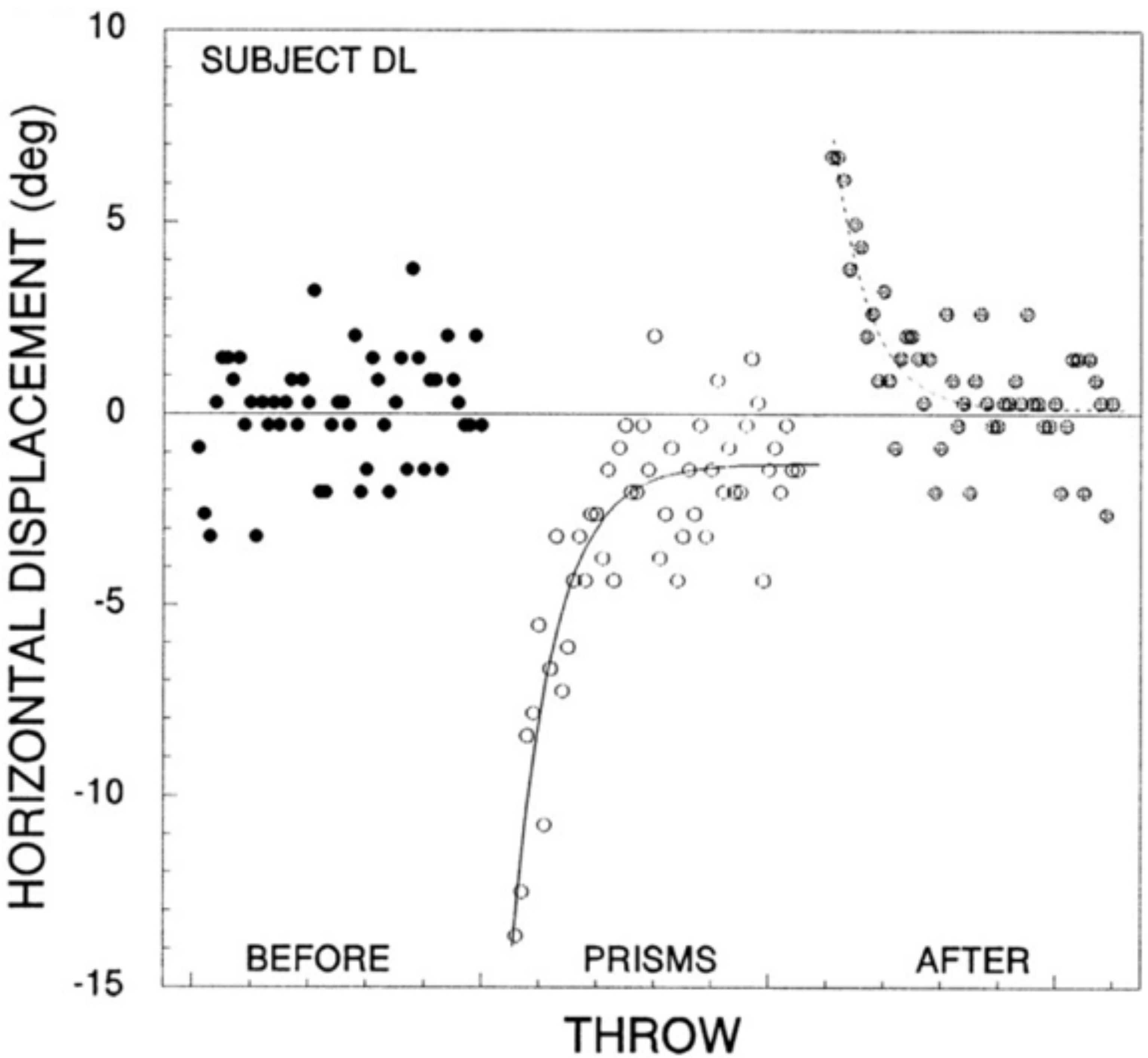
- In 3D:

Force fields





Prism adaptation: typical result

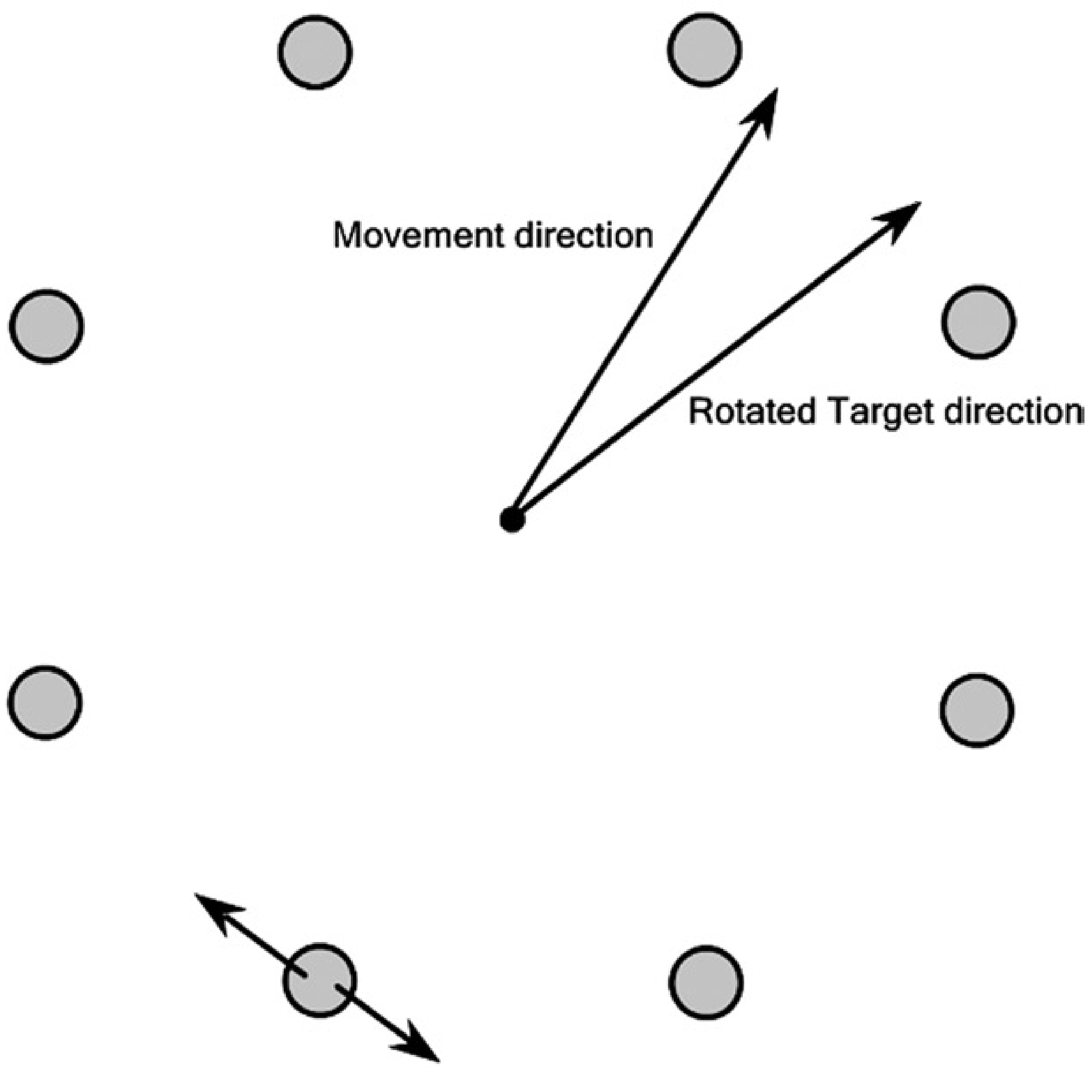


- Perturbation on:
  - initially large errors
  - errors decrease approximately exponentially
- Perturbation off:
  - initially large errors in opposite direction (*after effect*)
  - errors decrease approximately exponentially

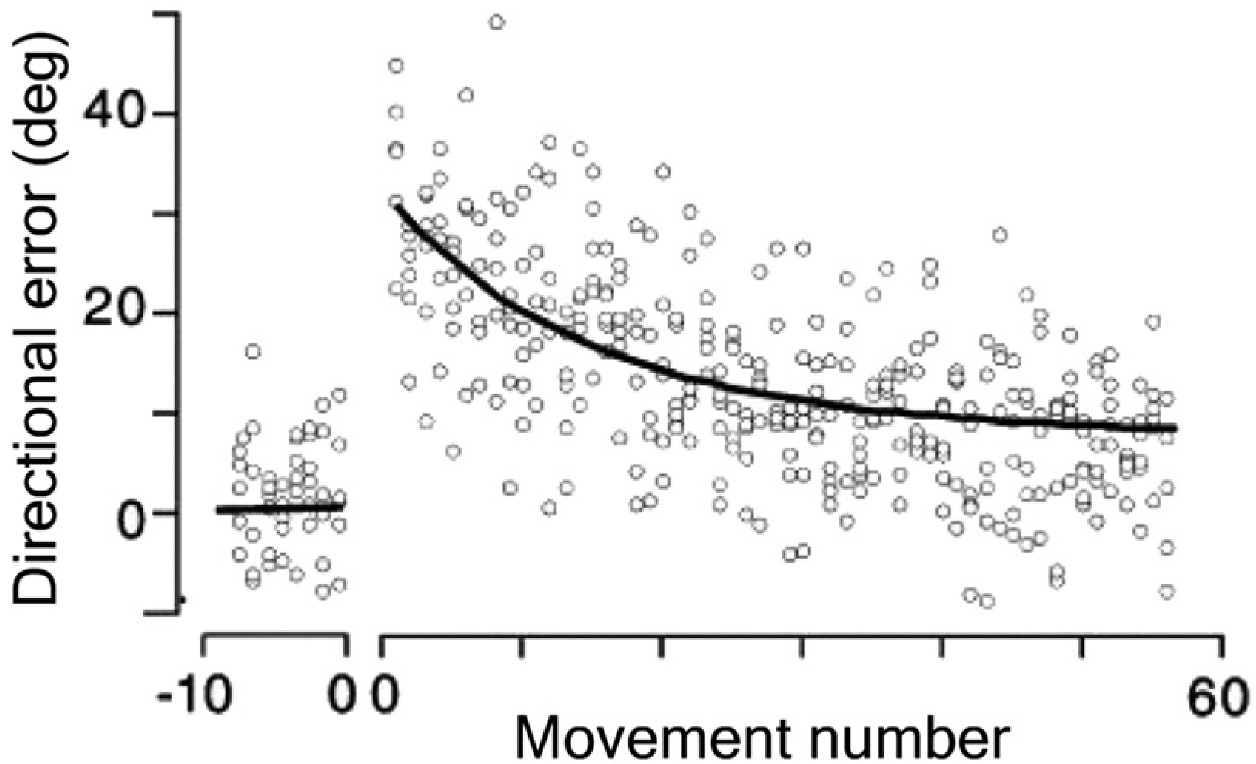
Martin et al. (2002) J Neurophysiol 88: 1685-1694

Visuomotor rotation paradigm

- Perturbation:



- Typical result:



  - initially large errors
  - errors decrease approximately exponentially

Experimental results so far

- Stepwise introduction of perturbation:
  - approximately exponential reduction of errors

- **How can this be modelled?**
- Desired: model that predicts movement direction in each movement, based on error in previous movement (learning from errors)
- Neglect biomechanics, muscle mechanics etc, assume accurate forward model

- **Important**:
  - this is a model in **discrete time**
  - model should be described by a **difference equation** (as opposed to *differential equation*)

A very simple model

- Define:
  - $x(i)$: movement direction in movement i
  - $e(i)$: movement error in movement i

- We look for relation between $x(i + 1)$, $x(i)$: and $e(i)$:

$$x(i+1) = x(i) + Be(i)$$

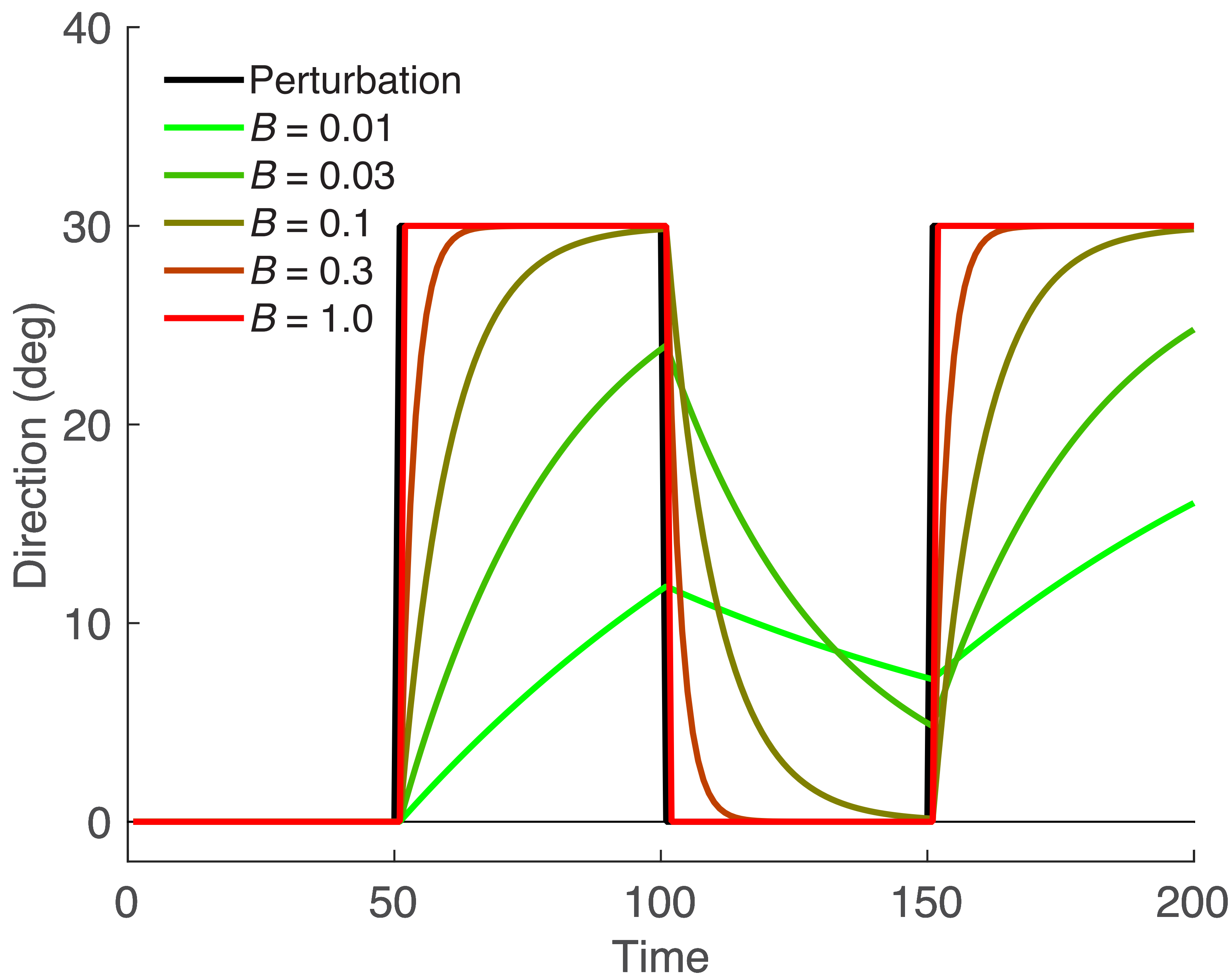
- with:
  - B: learning rate (proportion of error that is corrected for)

A very simple model

$$x(i+1) = x(i) + Be(i)$$
$$x(1) = 0$$

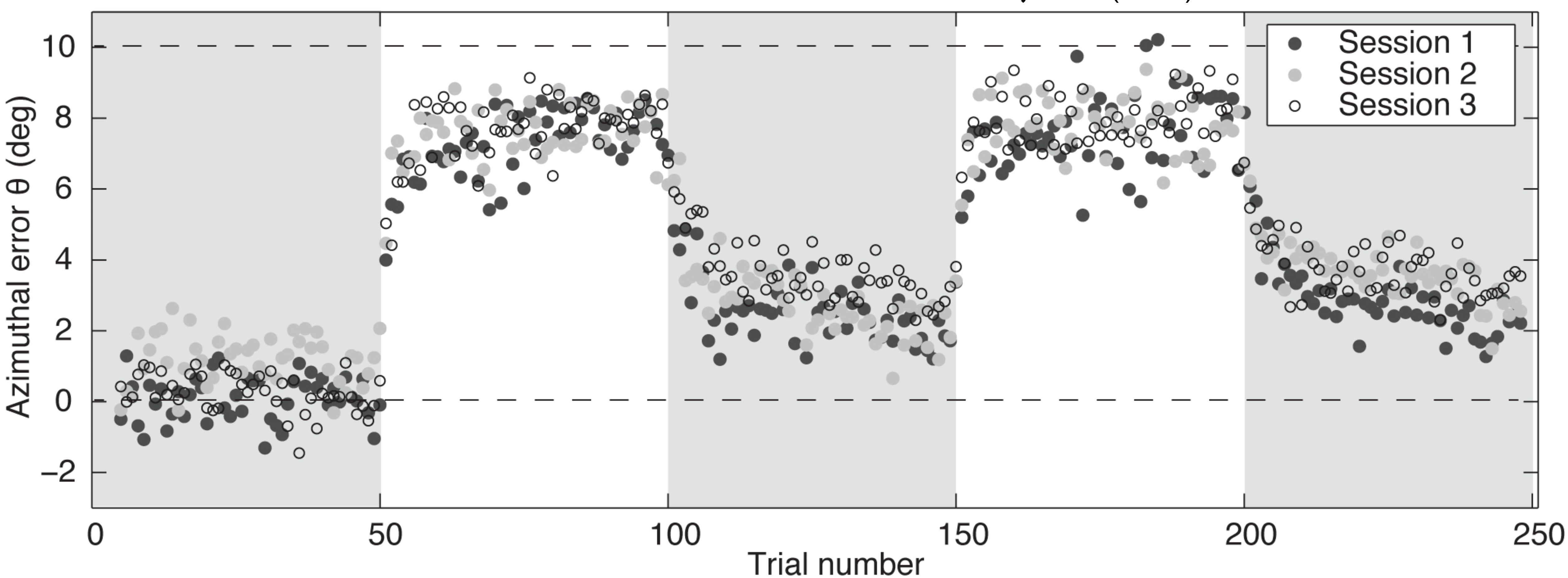
- This model produces exponential error reduction.
- B determines adaptation rate

- Prediction:
  - adaptation will be complete if the perturbation block is long enough



Data on this paradigm

van der Kooij et al. (2015) PLoS ONE 10: e0117901



- Prediction:
  - adaptation will be complete if the perturbation block is long enough
- Observation:
  - adaptation is often **incomplete**
- How should we adjust the model?

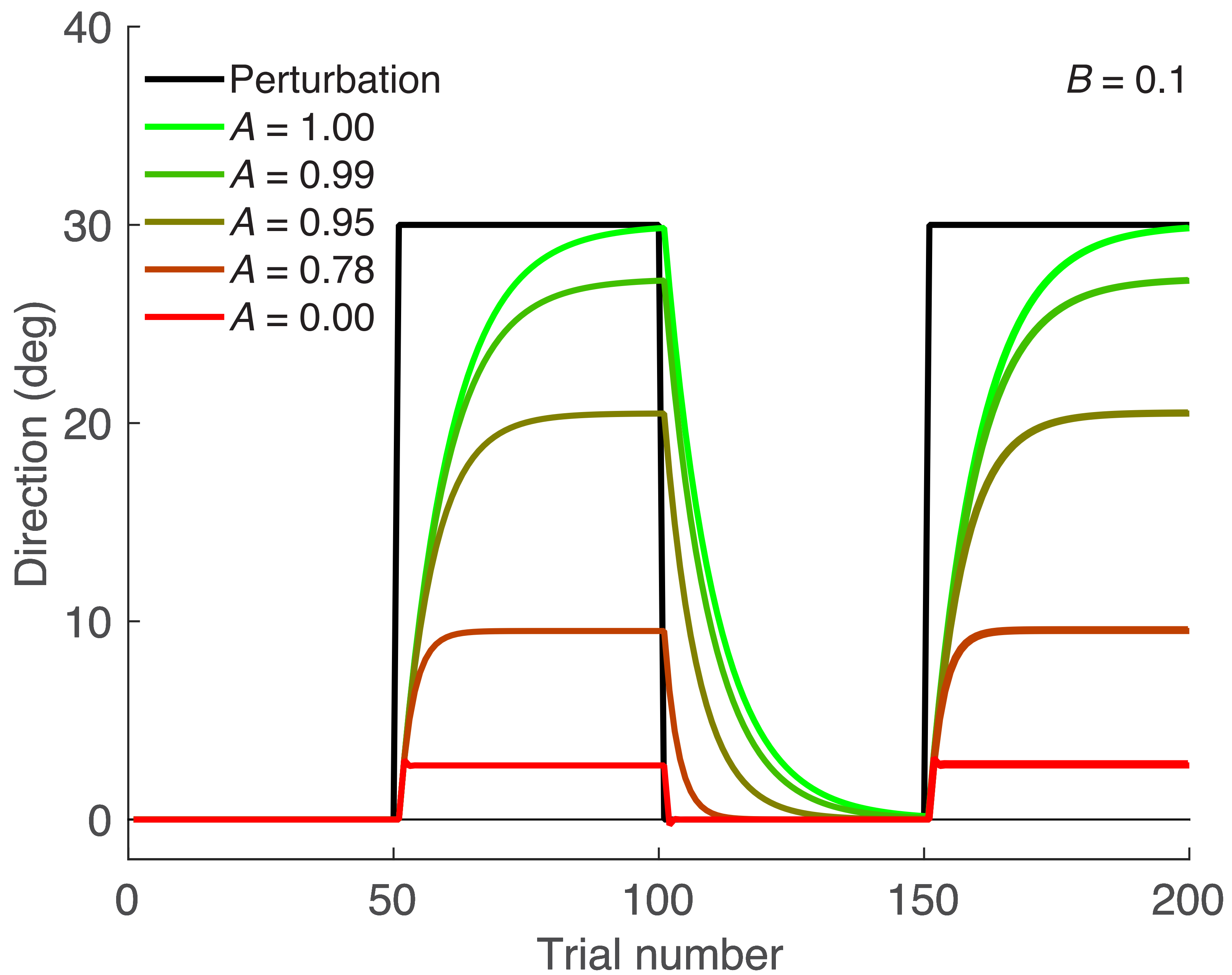
Data on this paradigm

- Add to the model:
  - A: retention factor

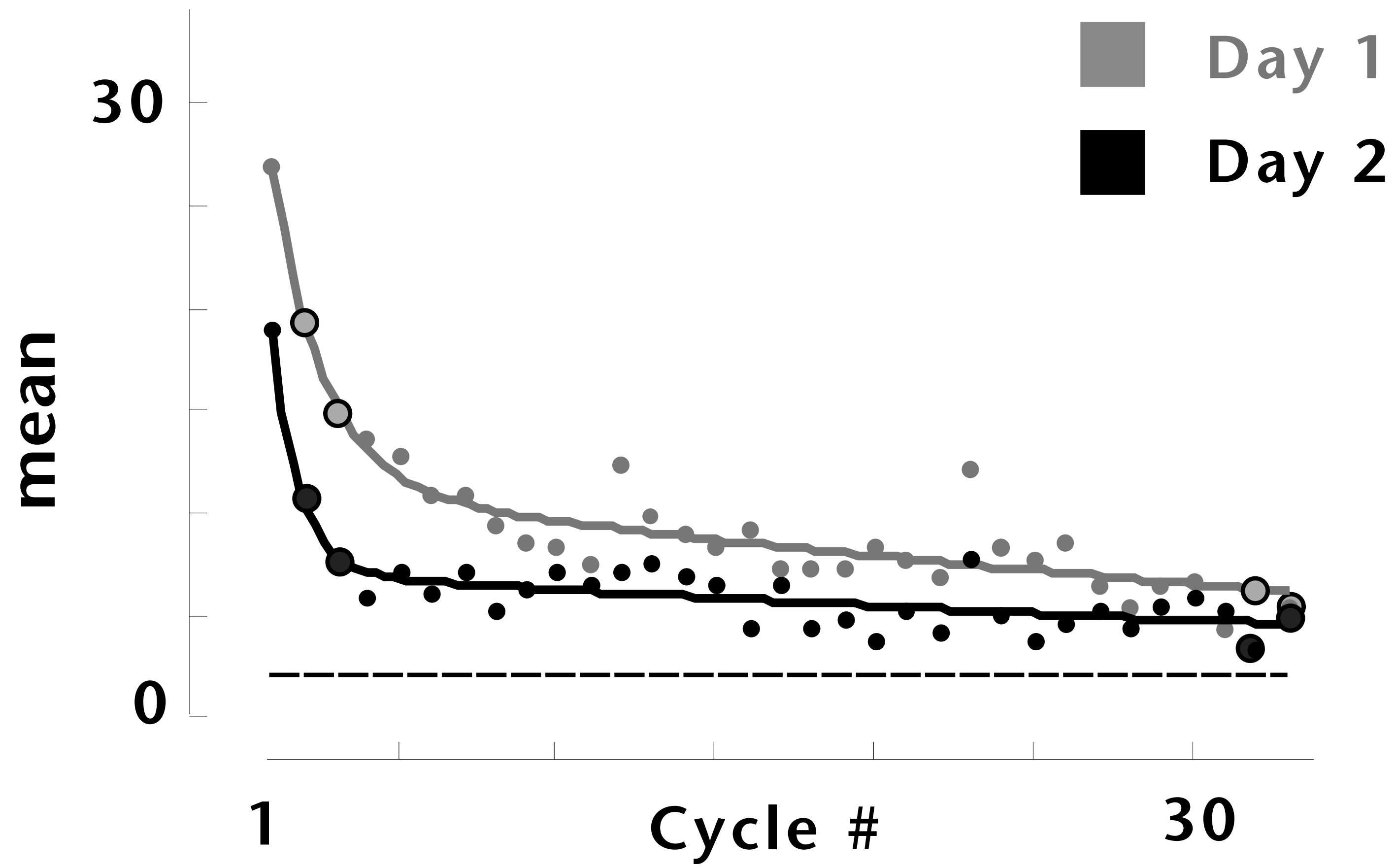
$$x(i+1) = Ax(i) + Be(i)$$

- Retention factor can account for incomplete adaptation.

- Prediction:
  - same time constant of exponential at every stepwise perturbation change



Data on this paradigm



Krakauer et al. (1999) Nat Neurosci 2: 1026-1031

- Relearning is generally faster than initial learning: **savings**
- This is incompatible with model considered so far.
- How should we adjust the model?

The two-state model

OPEN ACCESS Freely available online PLOS BIOLOGY

Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning

Maurice A. Smith[1*], Ali Ghazizadeh[2], Reza Shadmehr[3]

1 Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, United States of America, **2** Department of Bioengineering, University of California Berkeley, Berkeley, California, United States of America, **3** Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

Multiple processes may contribute to motor skill acquisition, but it is thought that many of these processes require sleep or the passage of long periods of time ranging from several hours to many days or weeks. Here we demonstrate that within a timescale of minutes, two distinct fast-acting processes drive motor adaptation. One process responds weakly to error but retains information well, whereas the other responds strongly but has poor retention. This two-state learning system makes the surprising prediction of spontaneous recovery (or adaptation rebound) if error feedback is clamped at zero following an adaptation-extinction training episode. We used a novel paradigm to experimentally confirm this prediction in human motor learning of reaching, and we show that the interaction between the learning processes in this simple two-state system provides a unifying explanation for several different, apparently unrelated, phenomena in motor adaptation including savings, anterograde interference, spontaneous recovery, and rapid unlearning. Our results suggest that motor adaptation depends on at least two distinct neural systems that have different sensitivity to error and retain information at different rates.

Citation: Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol 4(6): e179. DOI: 10.1371/journal.pbio.0040179

The two-state model

$$x(i+1) = Ax(i) + Be(i)$$

- Idea:
  - there are two (hidden) states:
    - x_1 (fast state) that learns quickly but has poor retention
    - x_2 (slow state) that learns slowly but has good retention
    - output is sum of both states

$$x_1(i+1) = A_f x_1(i) + B_f e(i)$$

$$x_2(i+1) = A_s x_2(i) + B_s e(i)$$

$$x(i) = x_1(i) + x_2(i)$$

$$B_f > B_s,\ A_f < A_s$$

The two-state model

$$x_1(i+1) = A_f x_1(i) + B_f e(i)$$

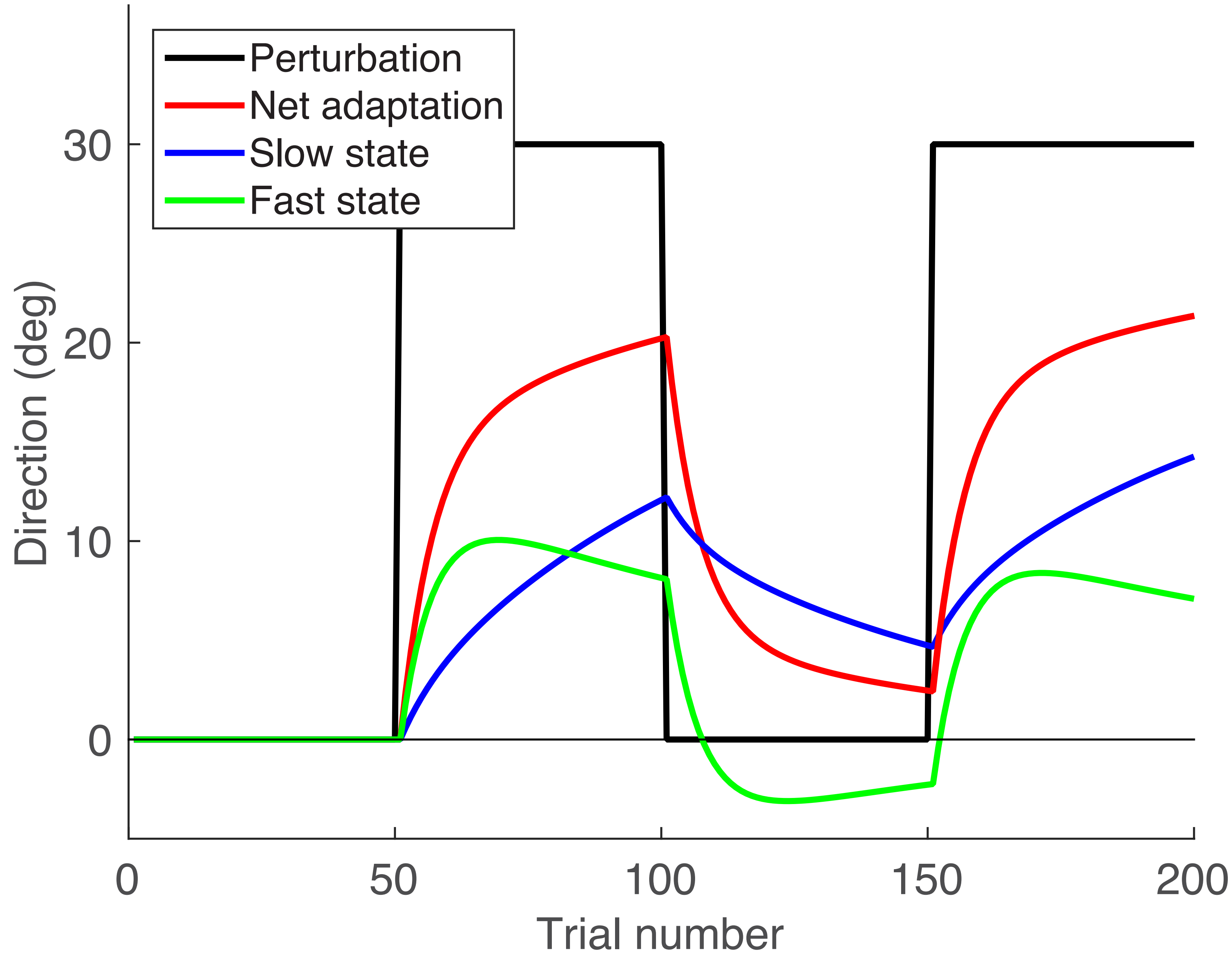
$$x_2(i+1) = A_s x_2(i) + B_s e(i)$$

$$x(i) = x_1(i) + x_2(i)$$

$$A_f = 0.92,\ B_f = 0.06$$

$$A_s = 0.995,\ B_s = 0.02$$

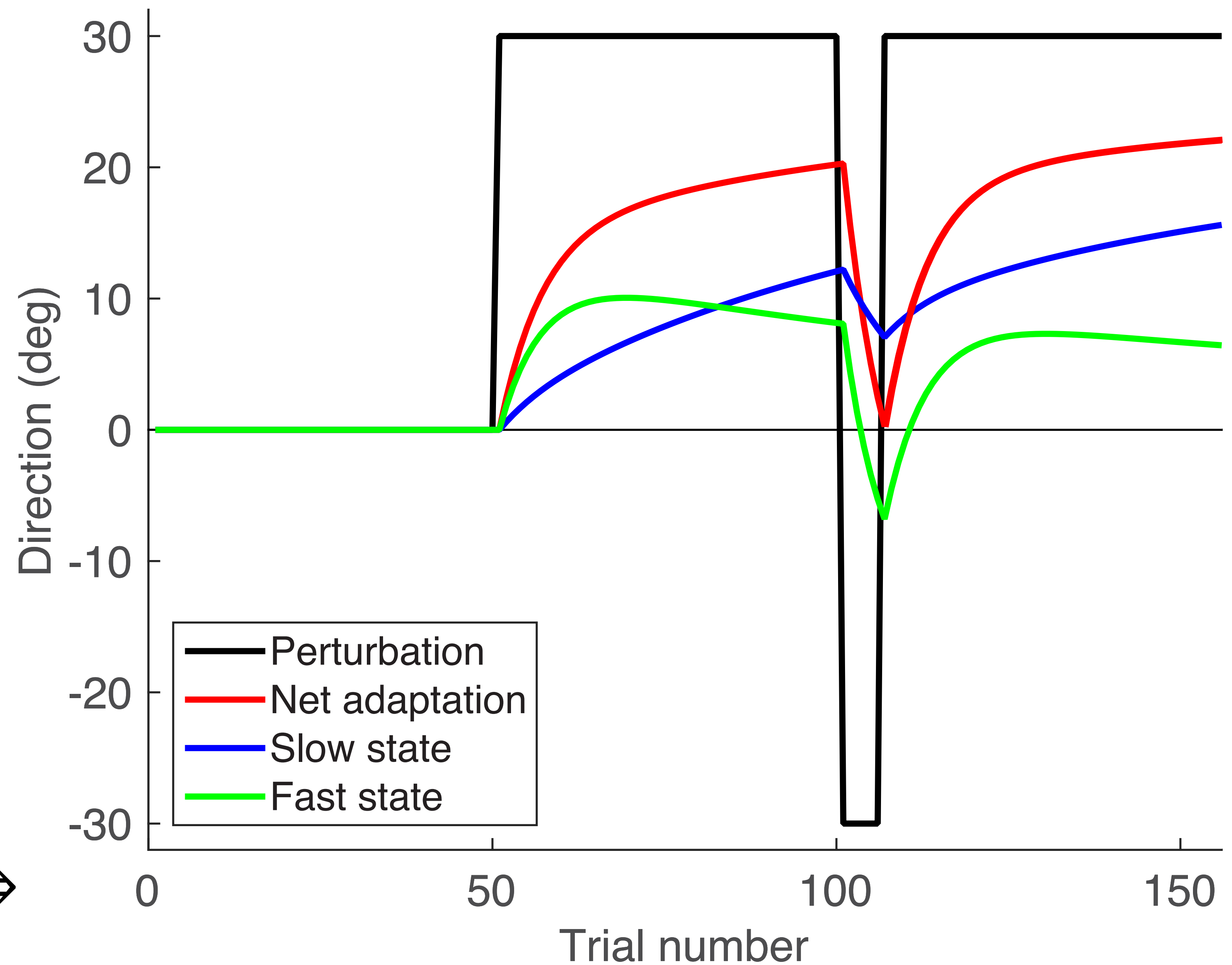
- Can explain savings.
- Or is this simply because the net adaptation differs between the start of the two adaptation blocks?



The two-state model

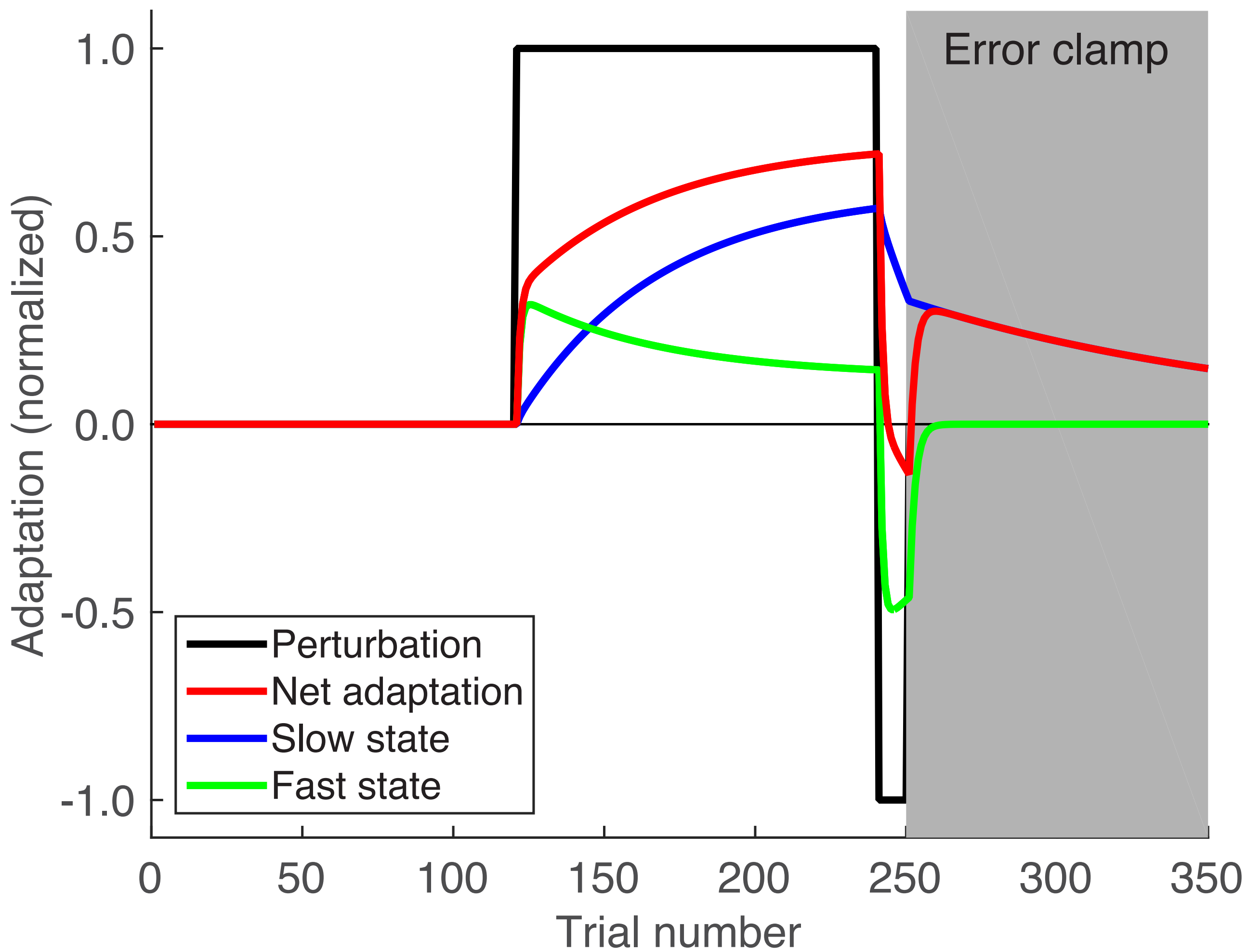
$$x_1(i+1) = A_f x_1(i) + B_f e(i)$$
$$x_2(i+1) = A_s x_2(i) + B_s e(i)$$
$$x(i) = x_1(i) + x_2(i)$$
$$A_f = 0.92,\ B_f = 0.06$$
$$A_s = 0.995,\ B_s = 0.02$$

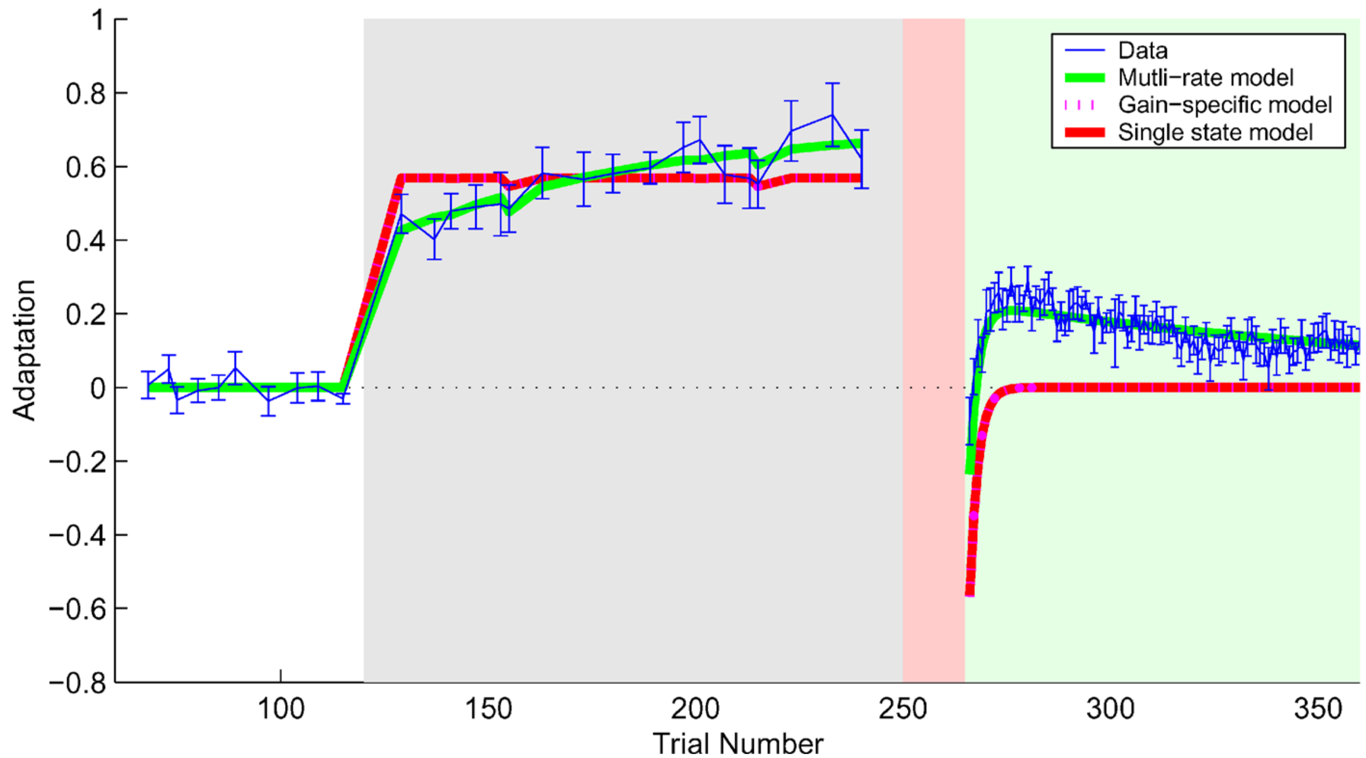
- Savings still present!
- Interpretation:
  - internal states different at start of 2nd block
  - fast state initially < 0 $\Longrightarrow$ adapts more (forgets less) $\Longrightarrow$ savings



The two-state model

> Counterintuitive prediction: adaptation rebound

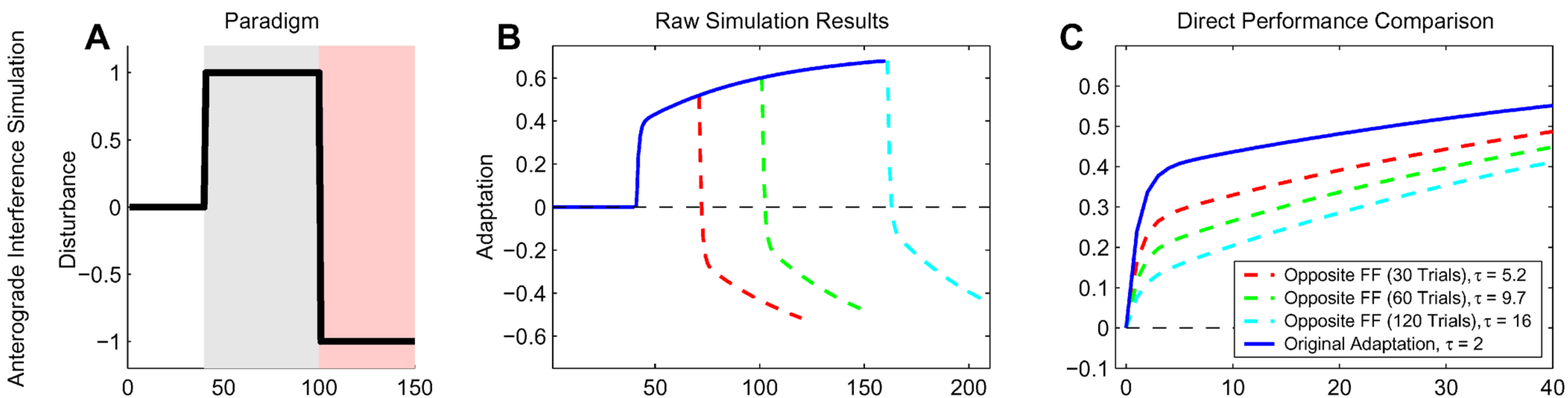




Smith et al. (2006) PLoS Biol 4: e179

The two-state model

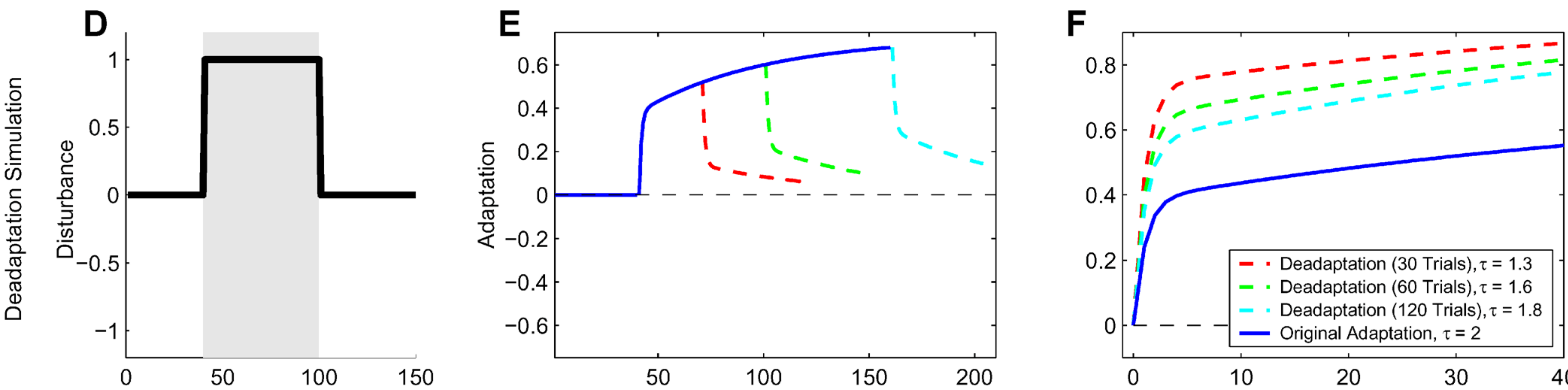
- Other effects this model can explain:
  - anterograde interference: initial adaptation is more rapid (shorter time constant) than subsequent adaptation in opposite direction



Smith et al. (2006) PLoS Biol 4: e179

The two-state model

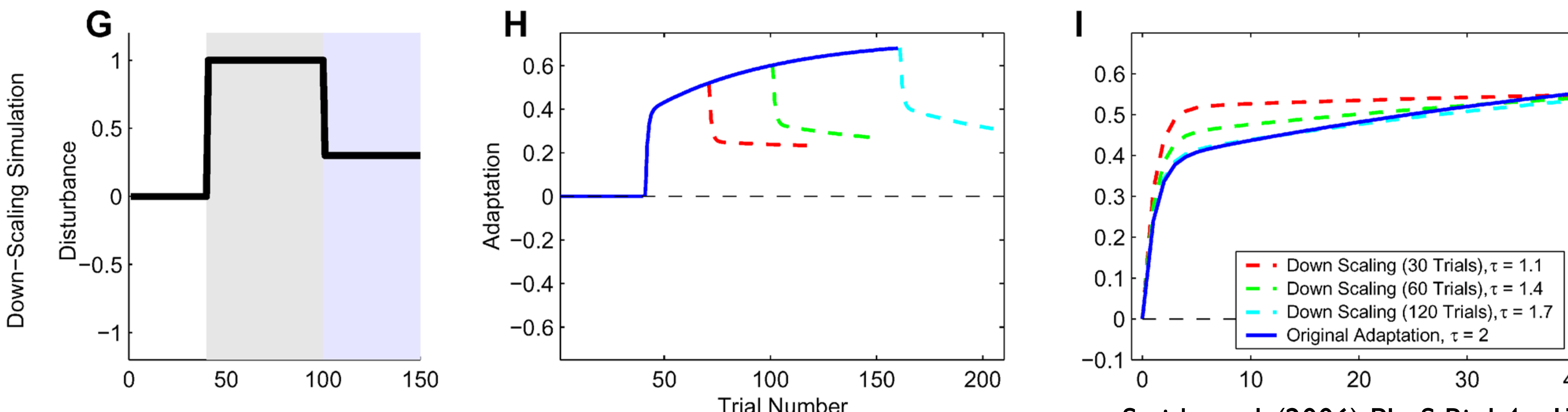
- > Other effects this model can explain:
  - > rapid unlearning: initial adaptation is slower (longer time constant) than subsequent washout



Smith et al. (2006) PLoS Biol 4: e179

The two-state model

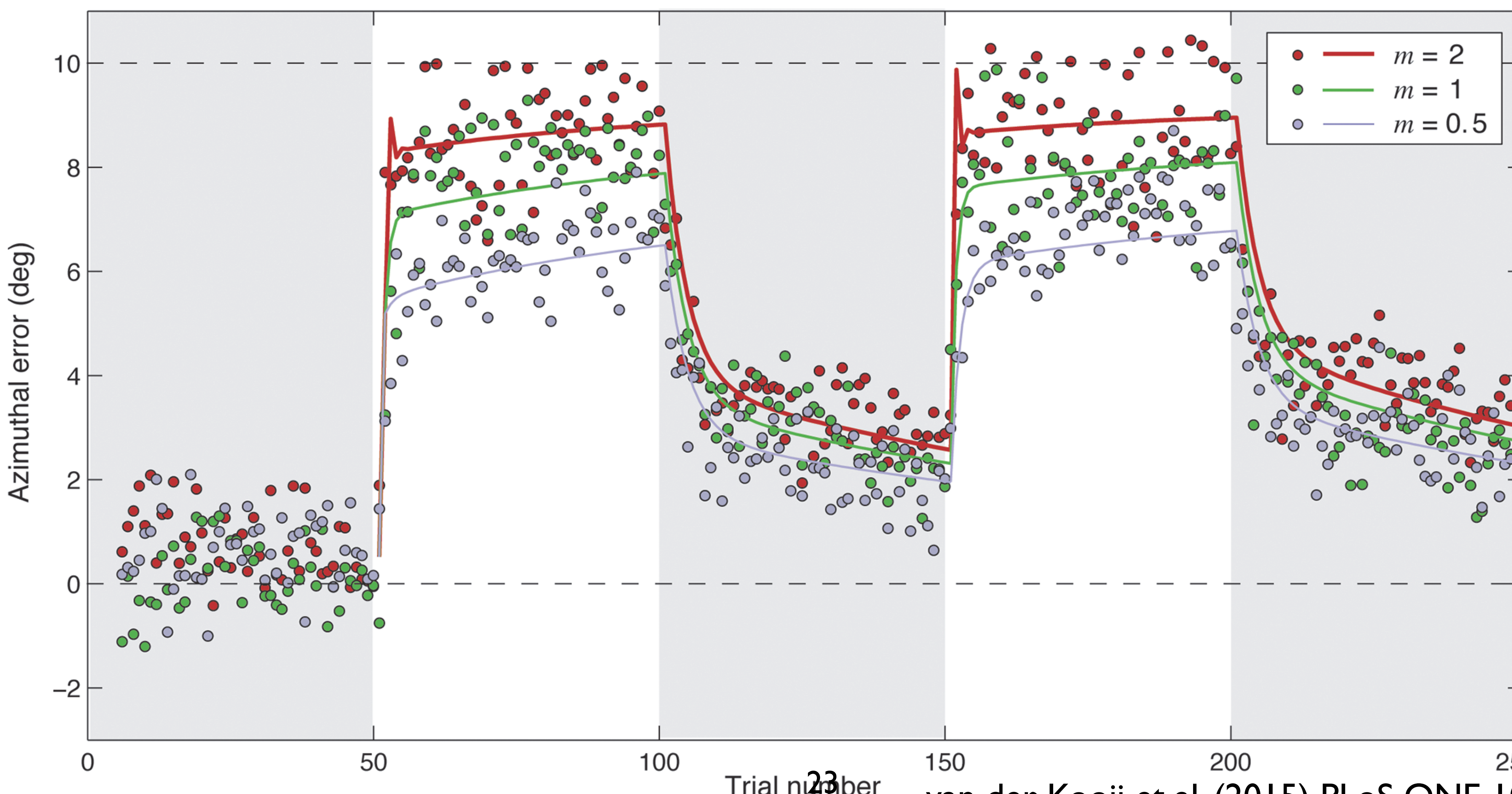
- Other effects this model can explain:
  - rapid downscaling: initial adaptation is slower (longer time constant) than subsequent adaptation to a scaled-down version of the initial perturbation



Smith et al. (2006) PLoS Biol 4: e179

The two-state model

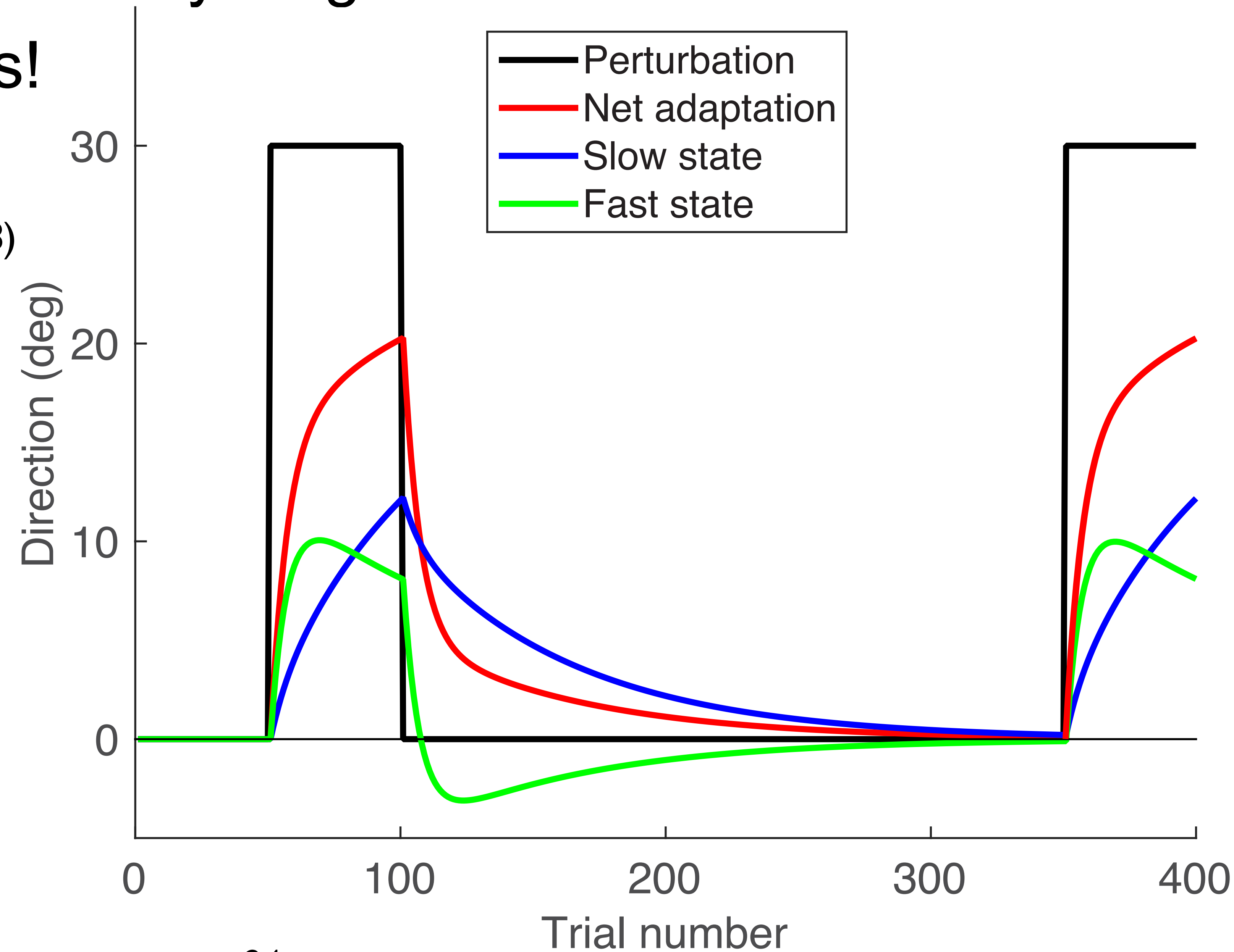
- Other effects this model can explain:
  - adaptation to visuomotor rotation when error feedback is scaled (up or down)



van der Kooij et al. (2015) PLoS ONE 10: e0117901

The two-state model

- > Other effects:
  - > what if washout block is very long?
  - > prediction: no savings!
  - > but savings still observed (Zarahn et al. (2008) J Neurophysiol 100: 2537-2548)
  - > so two-state model fails here!



Other models

- Models developed to explain also savings after long washout block:
  - ‘Variations of two-state model’:
    - Single-state varying-parameter model (Zarahn et al. 2008)
    - Parallel one-fast, multiple-slow process model (Lee & Schweighofer 2009)
    - Error-dependent learning rate model (Herzfeld et al. 2014)
  - Conceptually very different model:
    - Relevance-estimation model (Berniker & Kording 2011)

Zarahn et al. (2008)

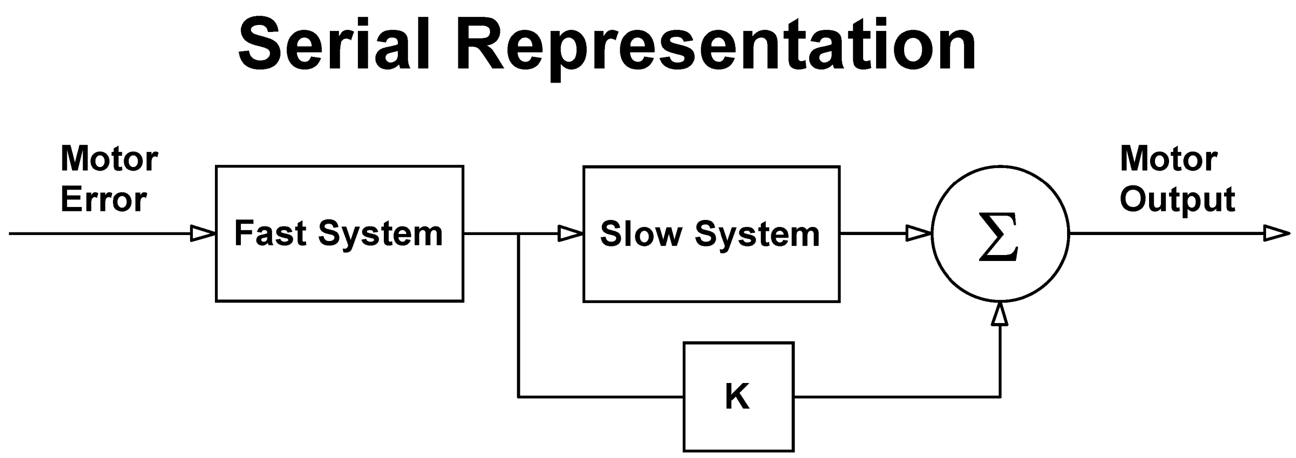
- Single-state varying-parameter model (Zarahn et al. 2008)
- Is essentially the single-state model in which A and B are not constant, but may vary between phases in an experiment.

$$x(i+1) = A(i)x(i) + B(i)e(i)$$

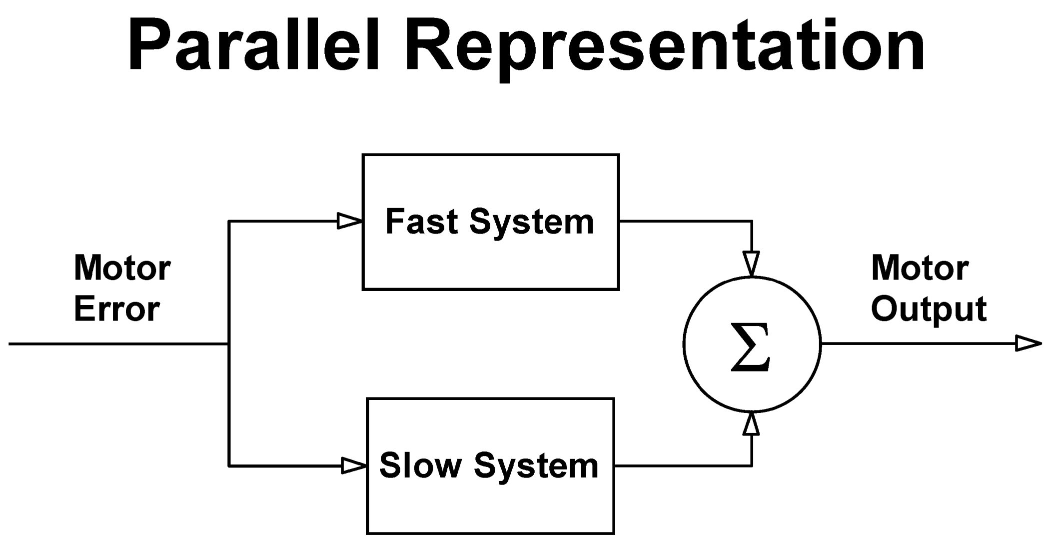
  - unclear what determines the parameter values

Lee & Schweighofer (2009)

- Problems with the two-state model:
  - cannot explain savings after long washout block
  - cannot account for learning multiple tasks simultaneously
  - unclear whether fast and slow processes are arranged serially or in parallel



Smith et al. (2006) PLoS Biol 4: e179

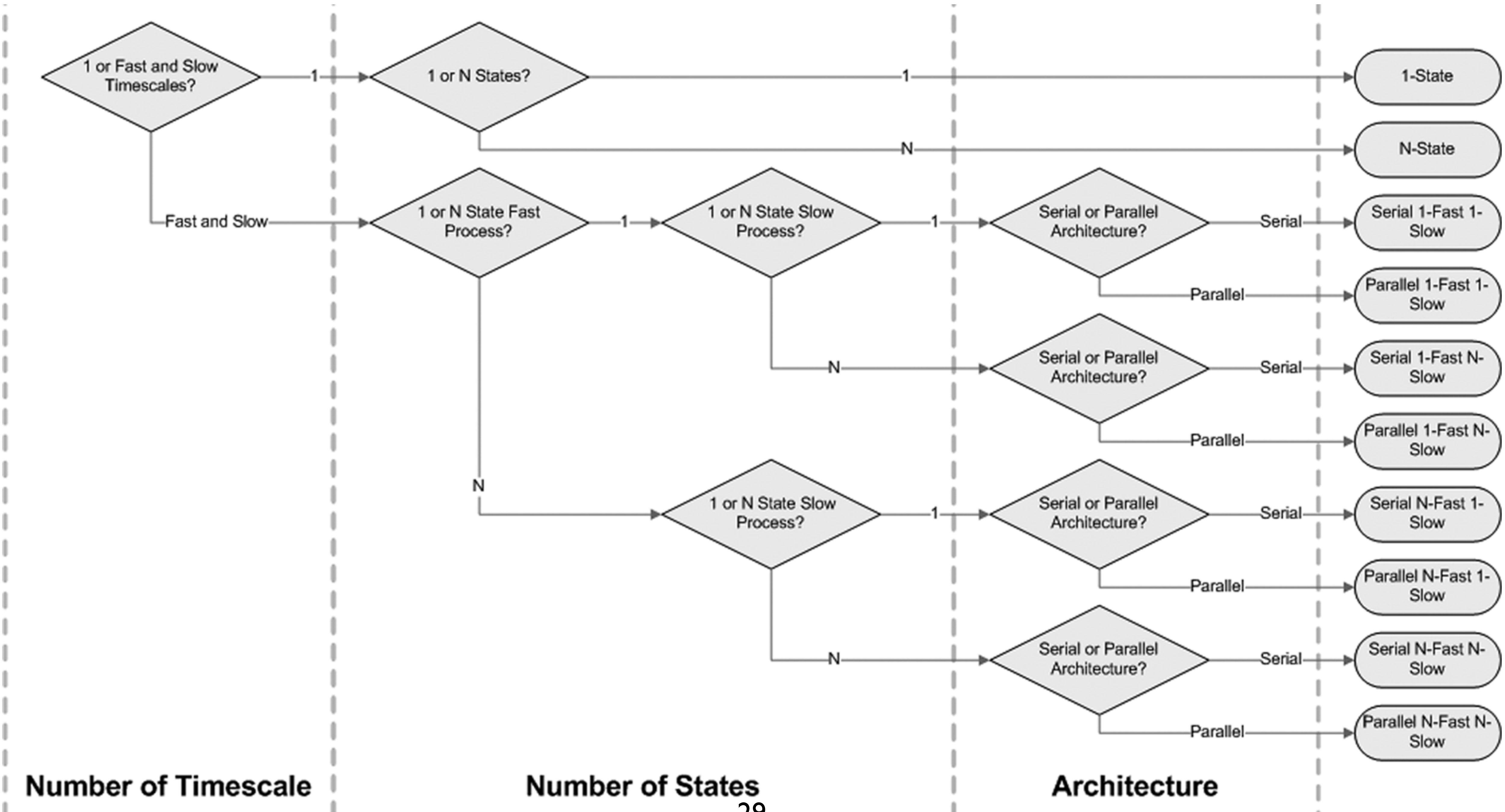


Lee & Schweighofer (2009)

- Lee & Schweighofer (2009) considered various possibilities:
  - are there fast and slow timescales, or just a single one?
  - is there one state or more than one for each time scale?
    - multiple states could account for learning multiple tasks simultaneously, with the state determined by the context
  - are the fast and slow states arranged serially or in parallel?

Lee & Schweighofer (2009)

> Overall, 10 possible models:



Lee & Schweighofer (2009) J Neurosci 29: 10396-10404

Lee & Schweighofer (2009)

- Model with fast and slow timescales and multiple states for each:

$$\mathbf{x}_f(i+1) = A_f \mathbf{x}_f(i) + B_f e(i)\mathbf{c}(i)$$

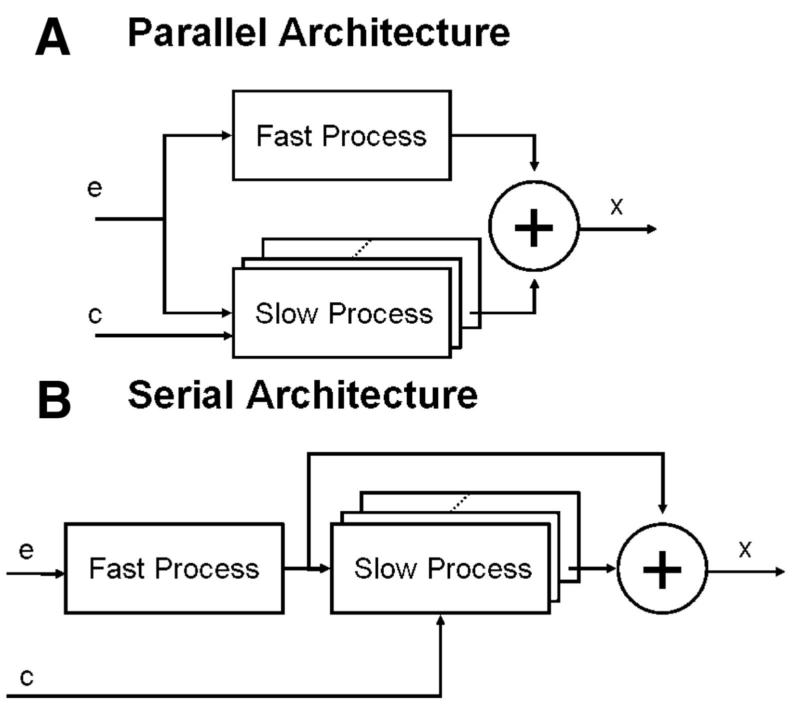
$$\mathbf{x}_s(i+1) = A_s \mathbf{x}_s(i) + B_s e(i)\mathbf{c}(i)$$

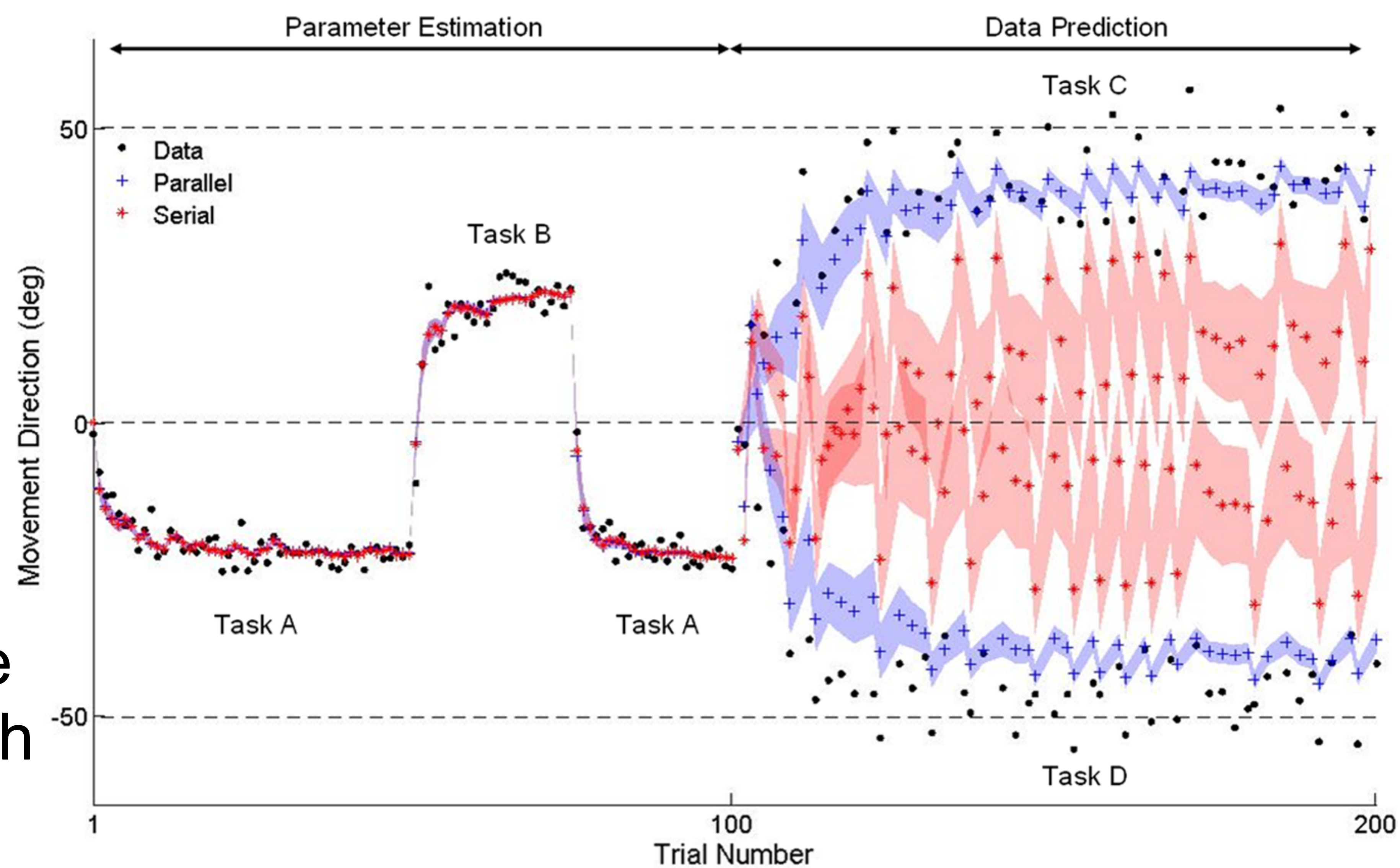
$$x(i) = \mathbf{x}_f(i)^T \mathbf{c}(i) + \mathbf{x}_s(i)^T \mathbf{c}(i)$$

- $\mathbf{x}_f$ and $\mathbf{x}_s$ are vectors (containing multiple states)
- $\mathbf{c}(i)$: contextual cue (vector with 0's and a single 1 for the relevant state)
- Only learning for the relevant state and only this state contributes to the output; all states forget.

Lee & Schweighofer (2009)

- They tested these models on existing data and on data from a new dual visuomotor adaptation paradigm:



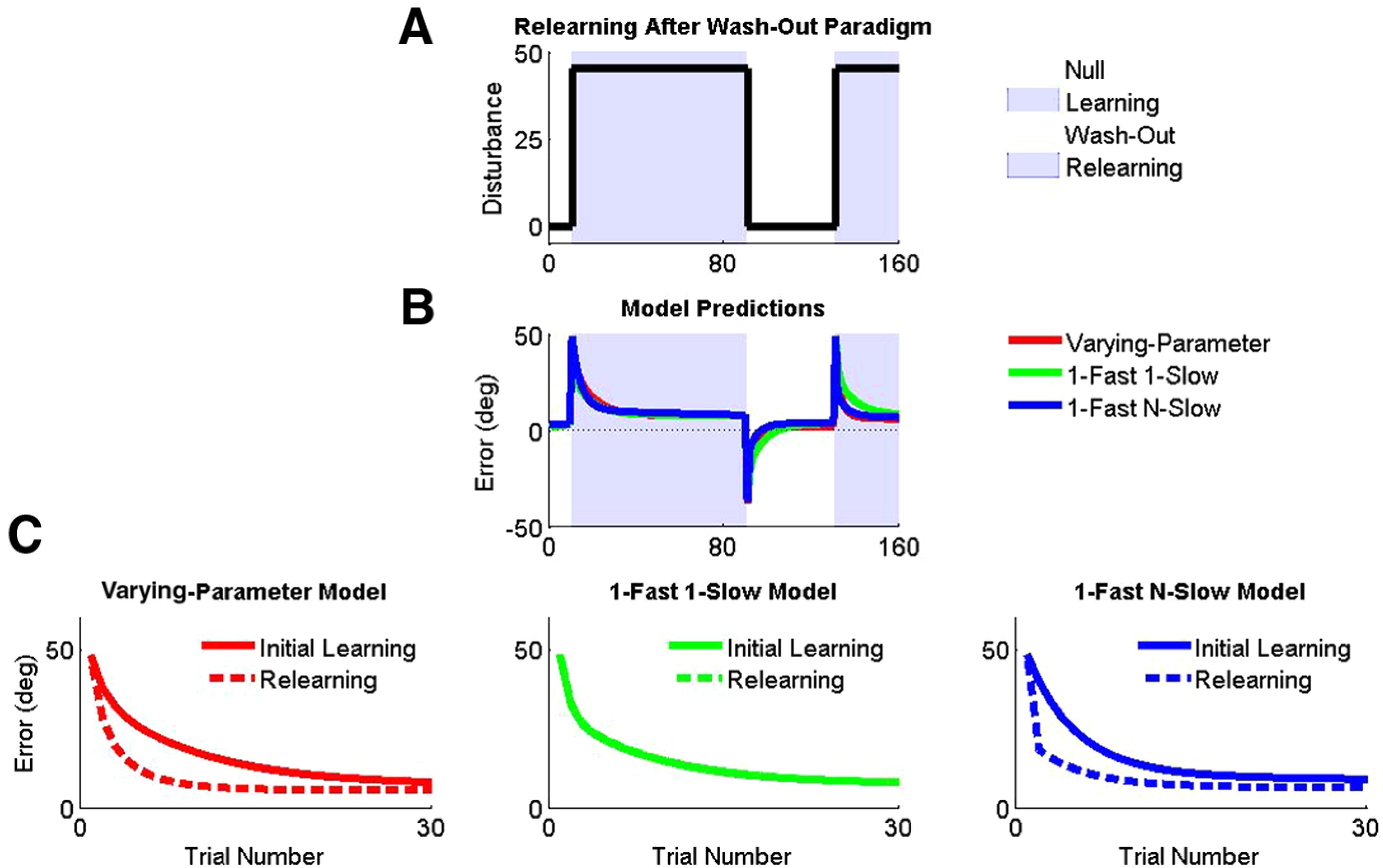


- 'Best model': fast process with single state in parallel with slow process with multiple states

Lee & Schweighofer (2009) J Neurosci 29: 10396-10404

Lee & Schweighofer (2009)

> Savings after long washout block



Lee & Schweighofer (2009) J Neurosci 29: 10396-10404

Herzfeld et al. (2014)

- Main equation for learning (and forgetting):

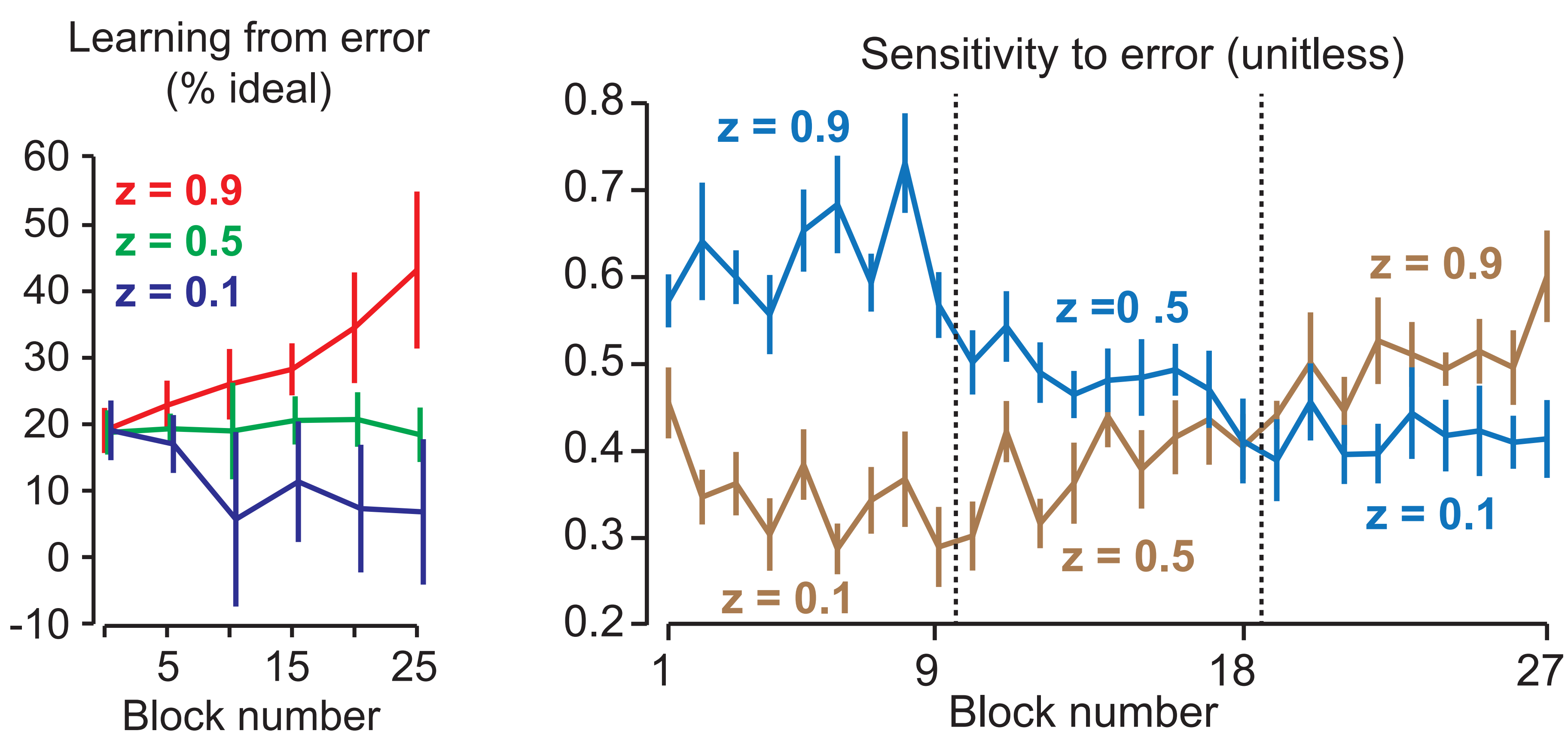
$$\hat{x}(i+1) = a\hat{x}(i) + \eta(i)e(i)$$

$\hat{x}(i)$	Subject's belief of perturbation x in movement i
a	Retention factor (A in other models)
$\eta(i)$	Error sensitivity in movement i (learning rate B in other models)
$e(i)$	Error in movement i

- Models considered so far: learning rate is a constant
- However, brain learns relatively more from small errors than from large errors (Wei & Körding (2009) J Neurophysiol 101: 655-64; Marko et al. (2012) J Neurophysiol 108: 1752-1763)
- Learning rate/error sensitivity may therefore not be a constant.

Herzfeld et al. (2014)

- Herzfeld et al. (2014) found in several experiments that error sensitivity can vary over time, as a function of the probability z that a perturbation will remain the same:

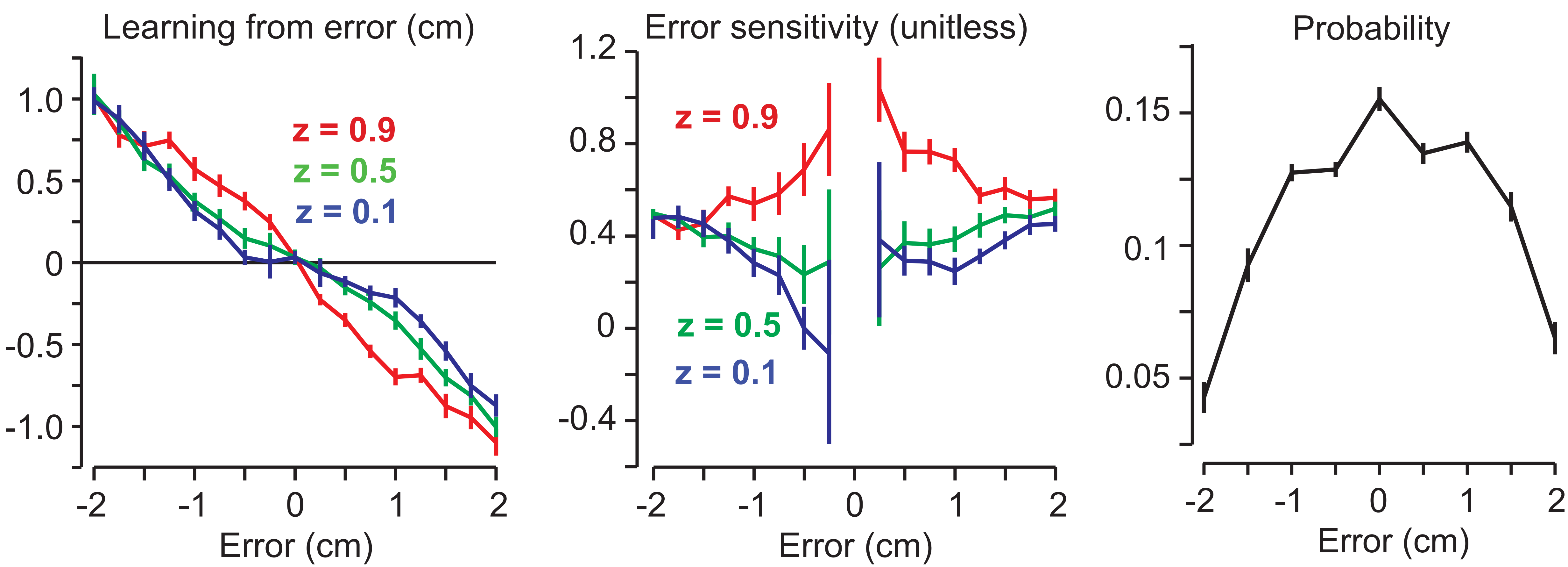


Herzfeld et al. (2014) Science 345: 1349-1353

- Error sensitivity decreases when perturbation is likely to change.

Herzfeld et al. (2014)

> Error sensitivity also changed as a function of the error size:

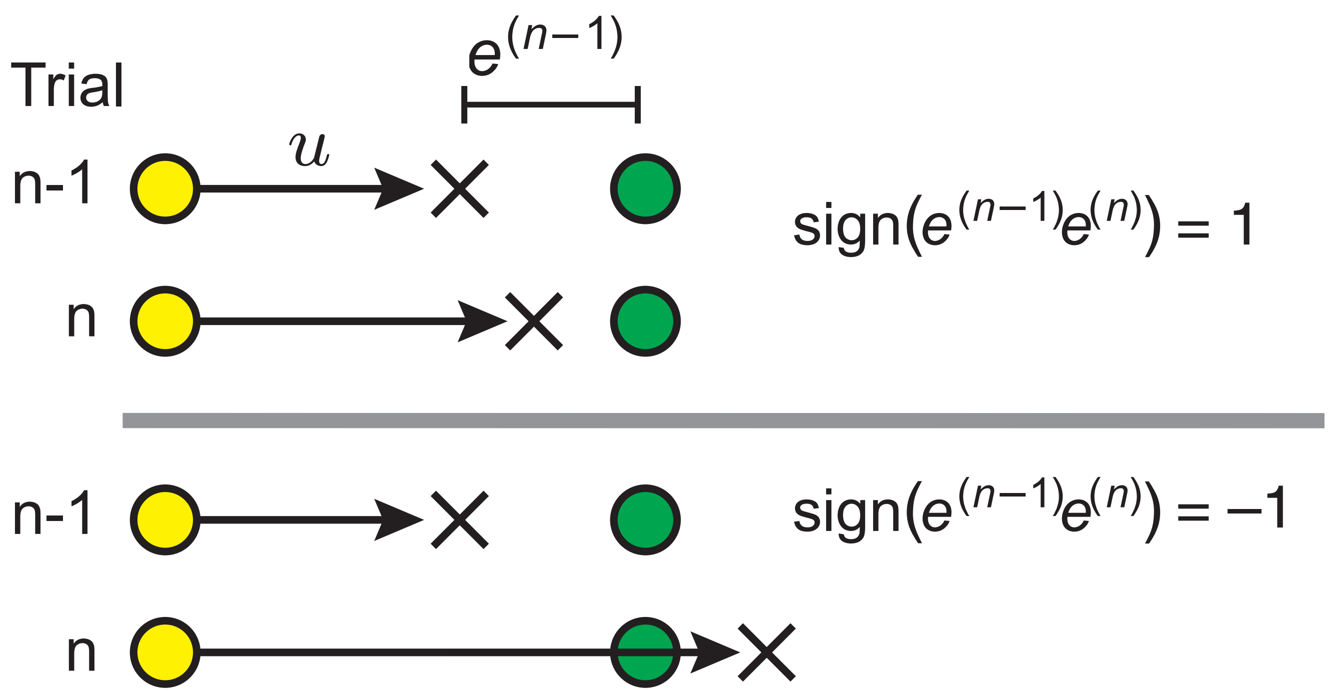


Herzfeld et al. (2014) Science 345: 1349-1353

> Error sensitivity changed most for errors that had been experienced most.

Herzfeld et al. (2014)

- Main assumptions:
  - error sensitivity is not constant, but is a function of the error size
  - error sensitivity can change locally, around the error just experienced
  - error sensitivity increases if two consecutive errors have the same signs, and decreases if they have opposite sign



Herzfeld et al. (2014) Science 345: 1349-1353

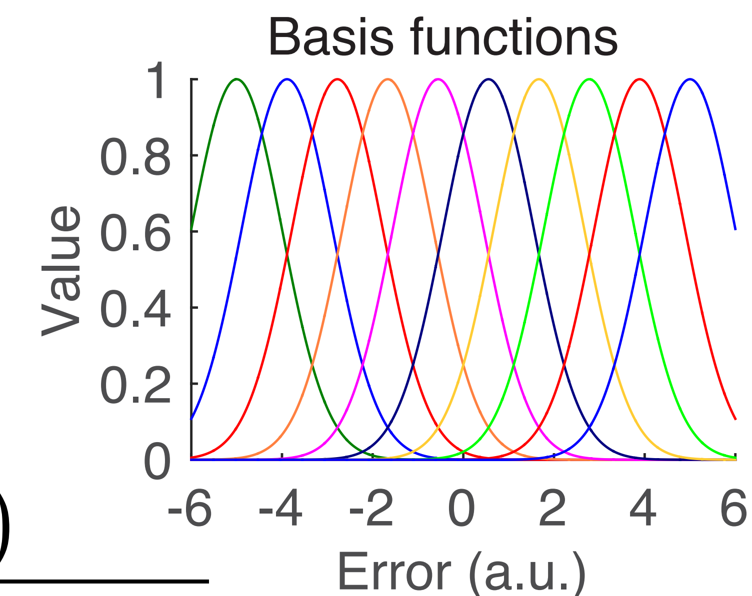
Herzfeld et al. (2014)

> Implementation of error sensitivity function using population coding:

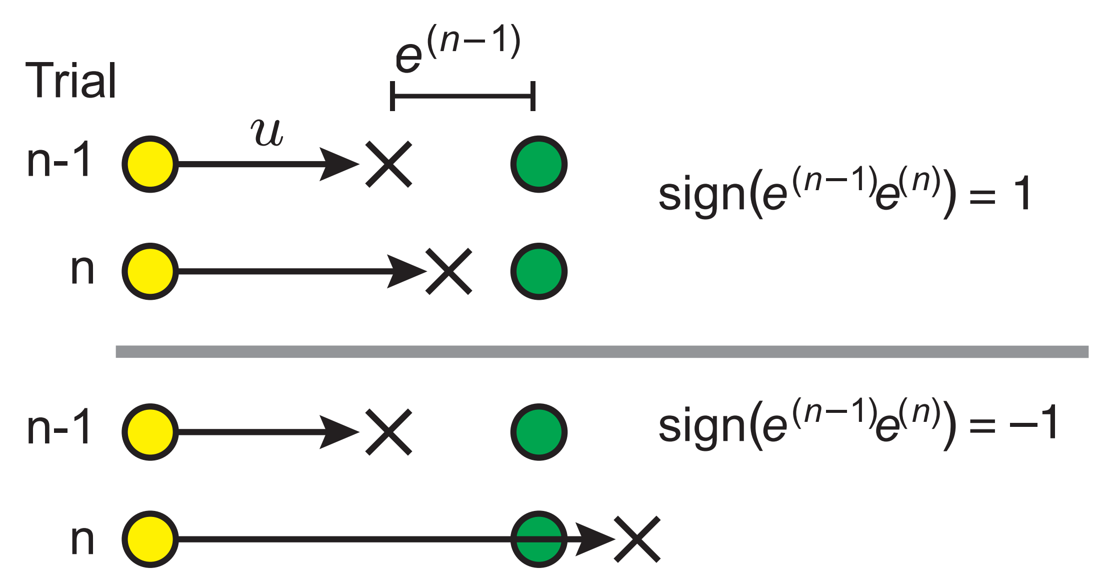
$$\eta\left(e(n)\right) = \sum_i w_i g_i\left(e(n)\right)$$

$$g_i\left(e(n)\right) = \exp\left(\frac{-(e(n)-\breve{e}_i)^2}{2\sigma^2}\right)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \beta\,\mathrm{sign}\left(e(n-1)e(n)\right)\frac{\mathbf{g}\left(e(n-1)\right)}{\mathbf{g}^T\left(e(n-1)\right)\mathbf{g}\left(e(n-1)\right)}$$



w_i	weight of basis function i
g_i	basis function i
$\breve{e}_i$	preferred error of basis function i
σ^2	variance of basis functions
β	update rate

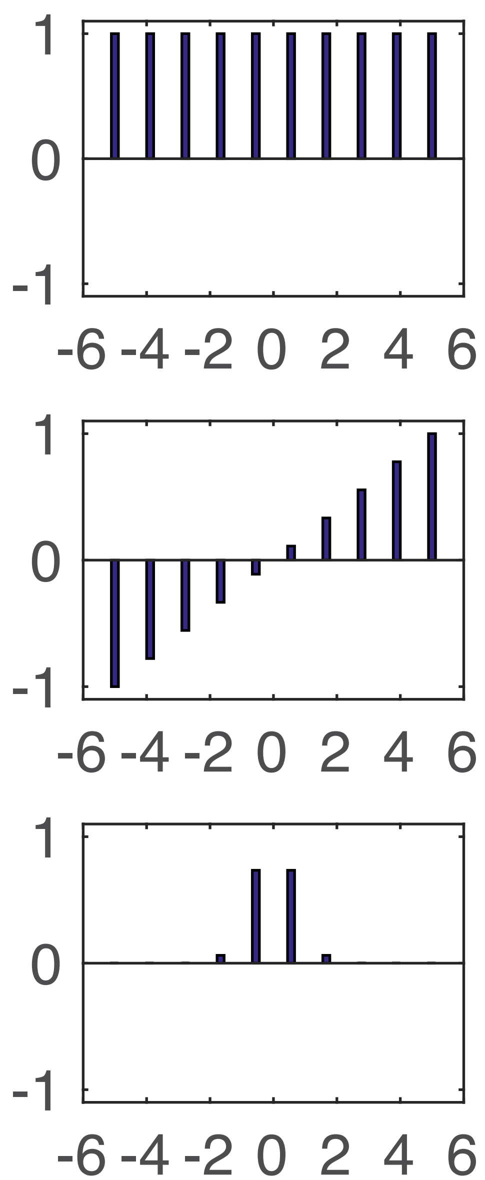


Herzfeld et al. (2014) Science 345: 1349-1353

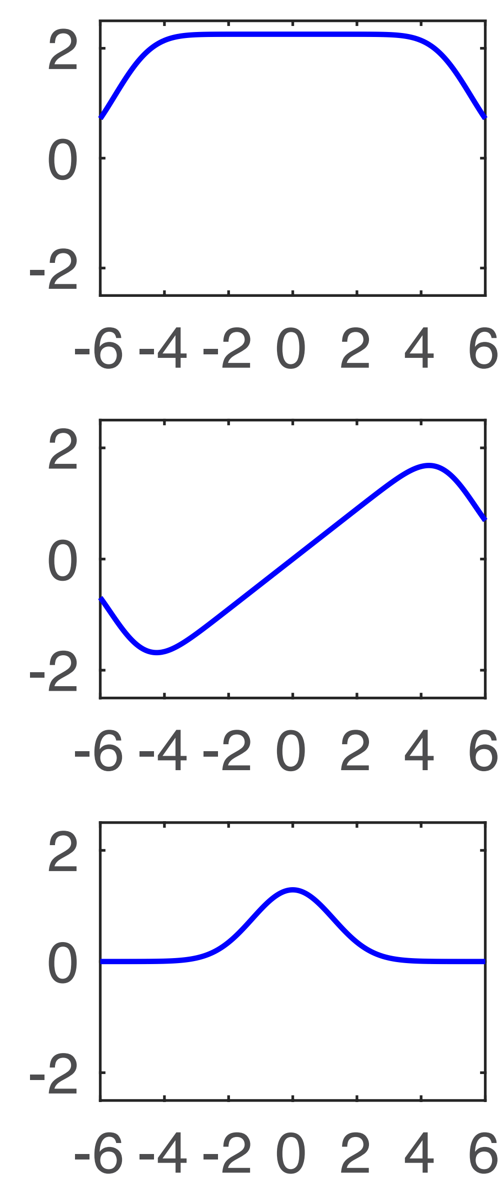
Herzfeld et al. (2014)

> Examples of error sensitivity functions:

> **w** (a.u.) plotted as a function of preferred error:

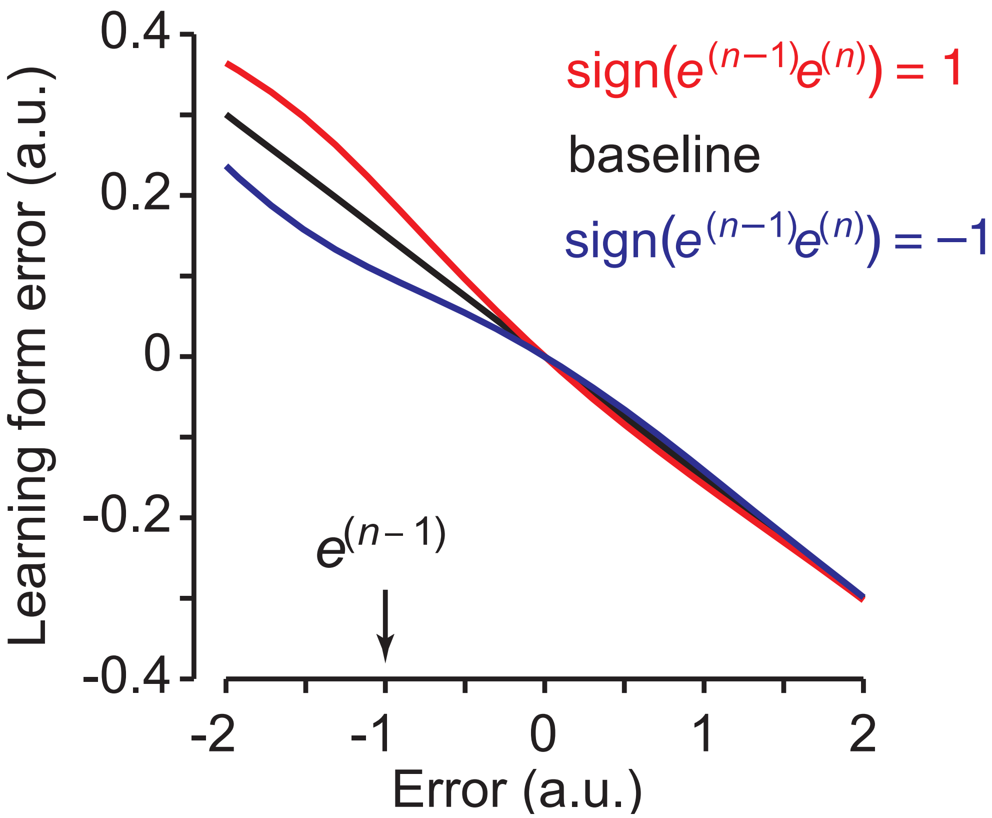


> corresponding error sensitivity (a.u.) plotted as a function of error:



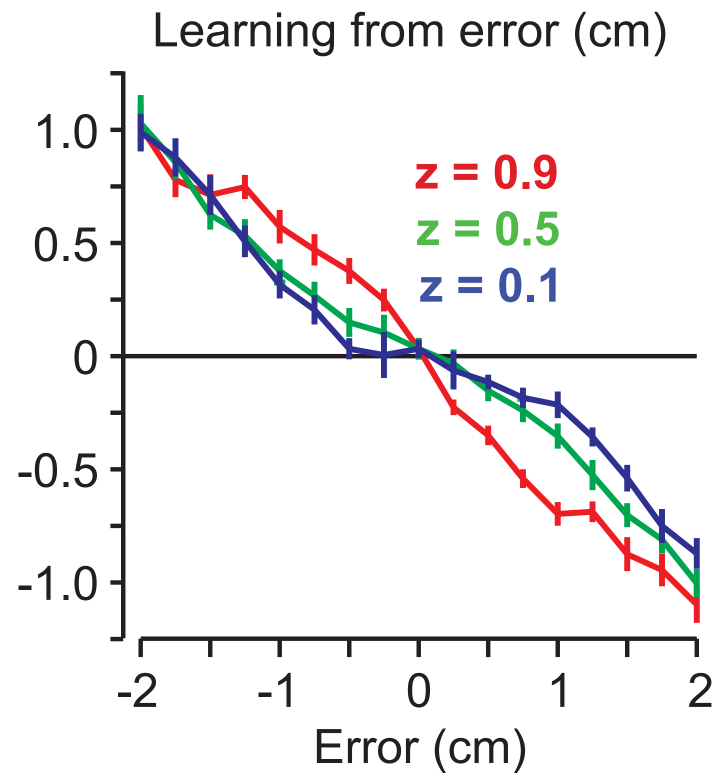
Herzfeld et al. (2014)

> Error sensitivity function is changed after every movement made:

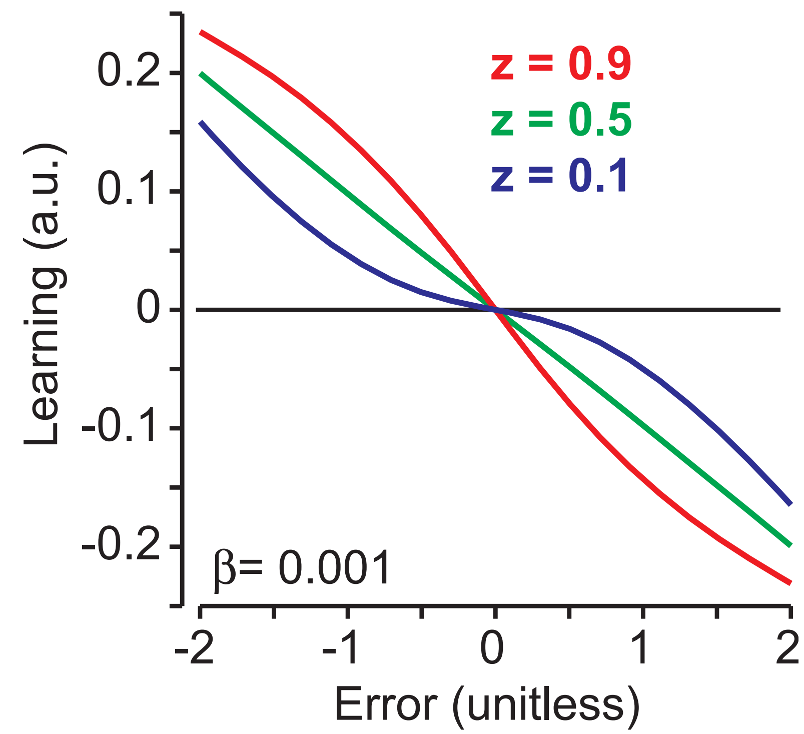


Herzfeld et al. (2014) Science 345: 1349-1353

> Observed:

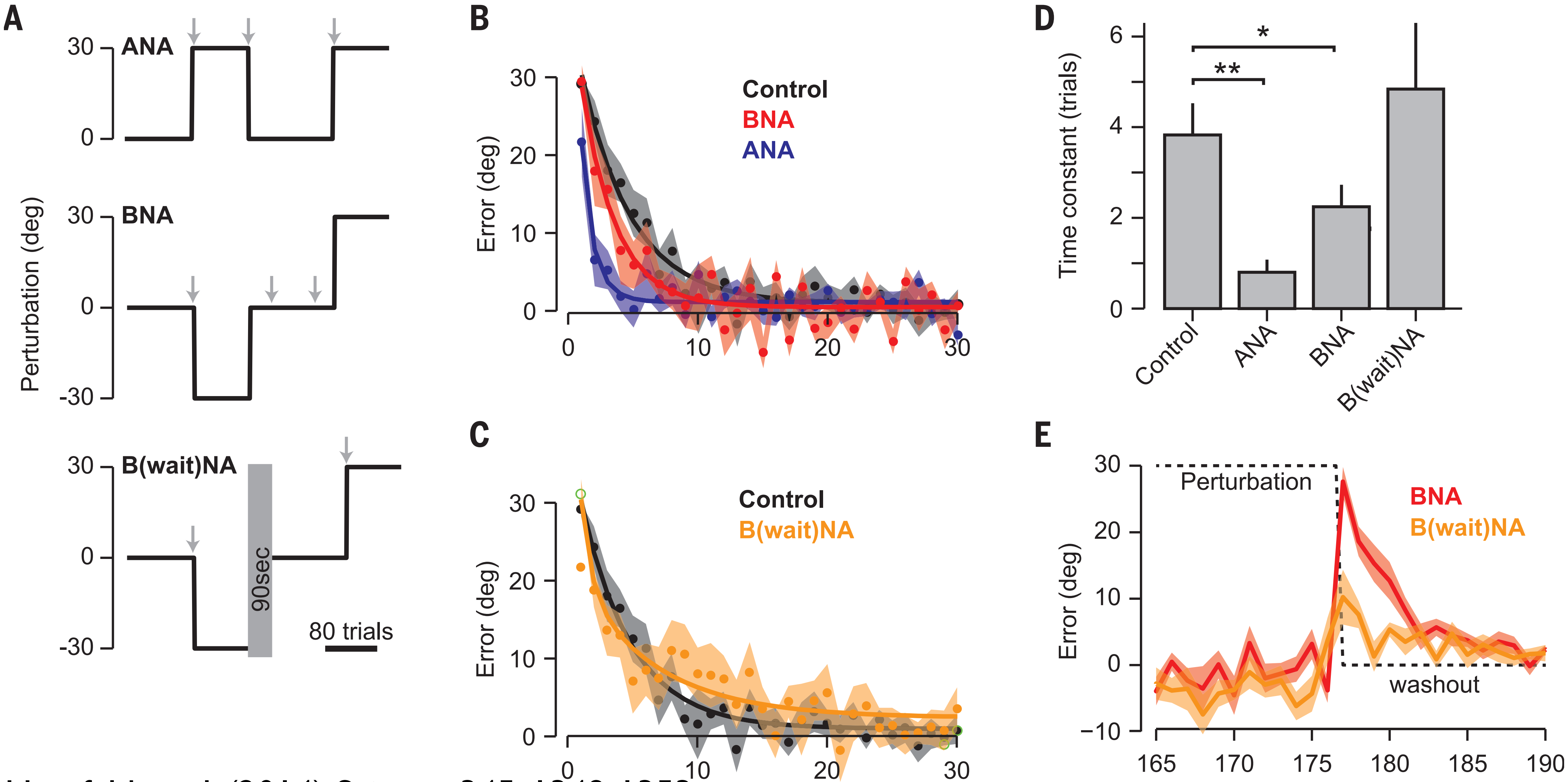


> Model:



Herzfeld et al. (2014)

> This model can explain savings, regardless of # washout trials:



Herzfeld et al. (2014) Science 345: 1349-1353

> Also savings in BNA paradigm, as predicted by this model.

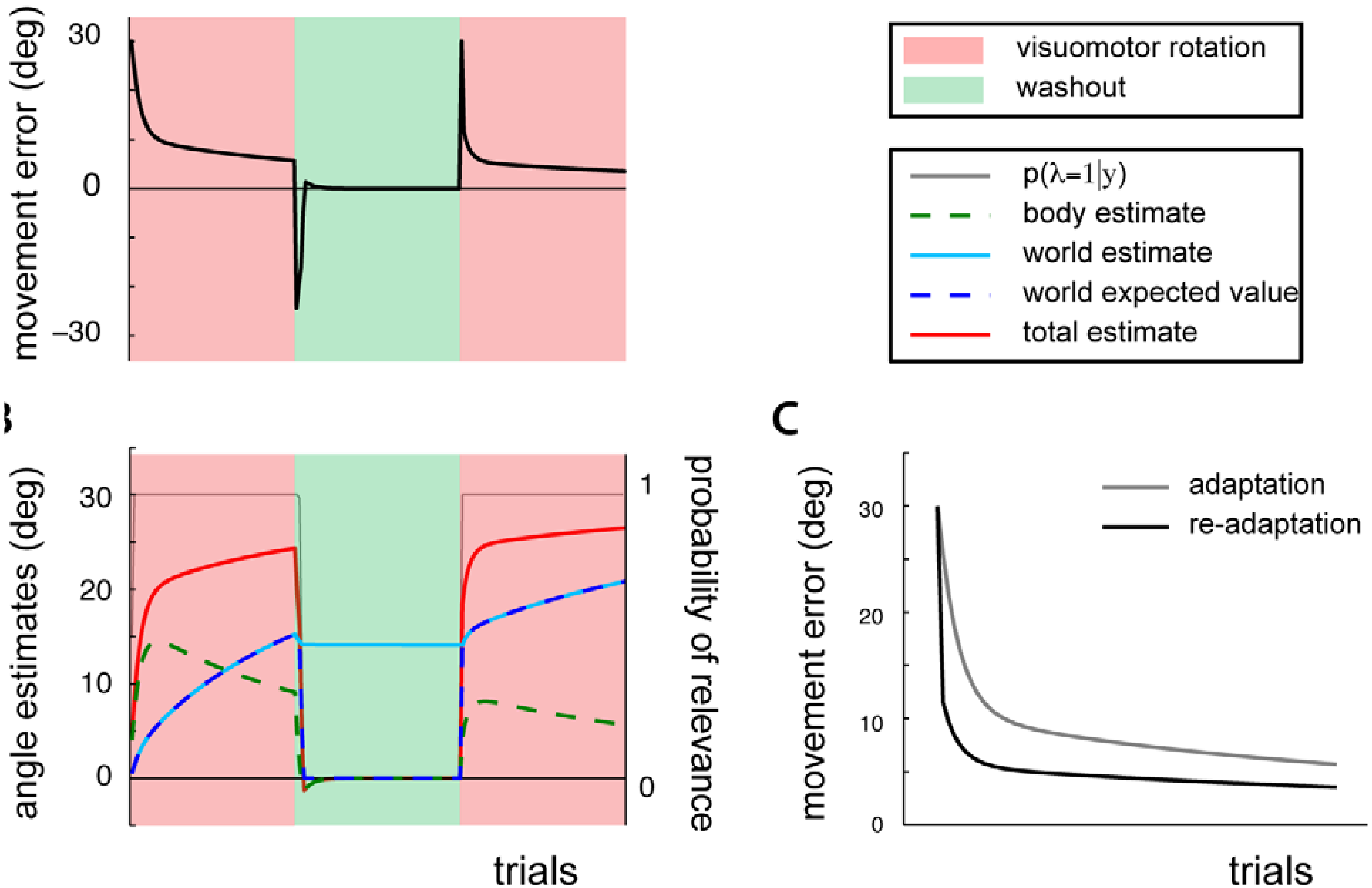
> And no savings in B(wait)NA paradigm, also as predicted.

Berniker & Kording (2011)

- Conceptually very different than previous models.
  - Previous models: how?
  - This model: why? (and also: how?)
- Ideas:
  - movement errors can be caused by:
    - world disturbances (external perturbations)
    - our own body (internal changes, noise)
  - this model probabilistically infers sources of motor errors, and their relevance to the current circumstances
  - estimates of body and world state are then updated based on the inferred relevance

Berniker & Kording (2011)

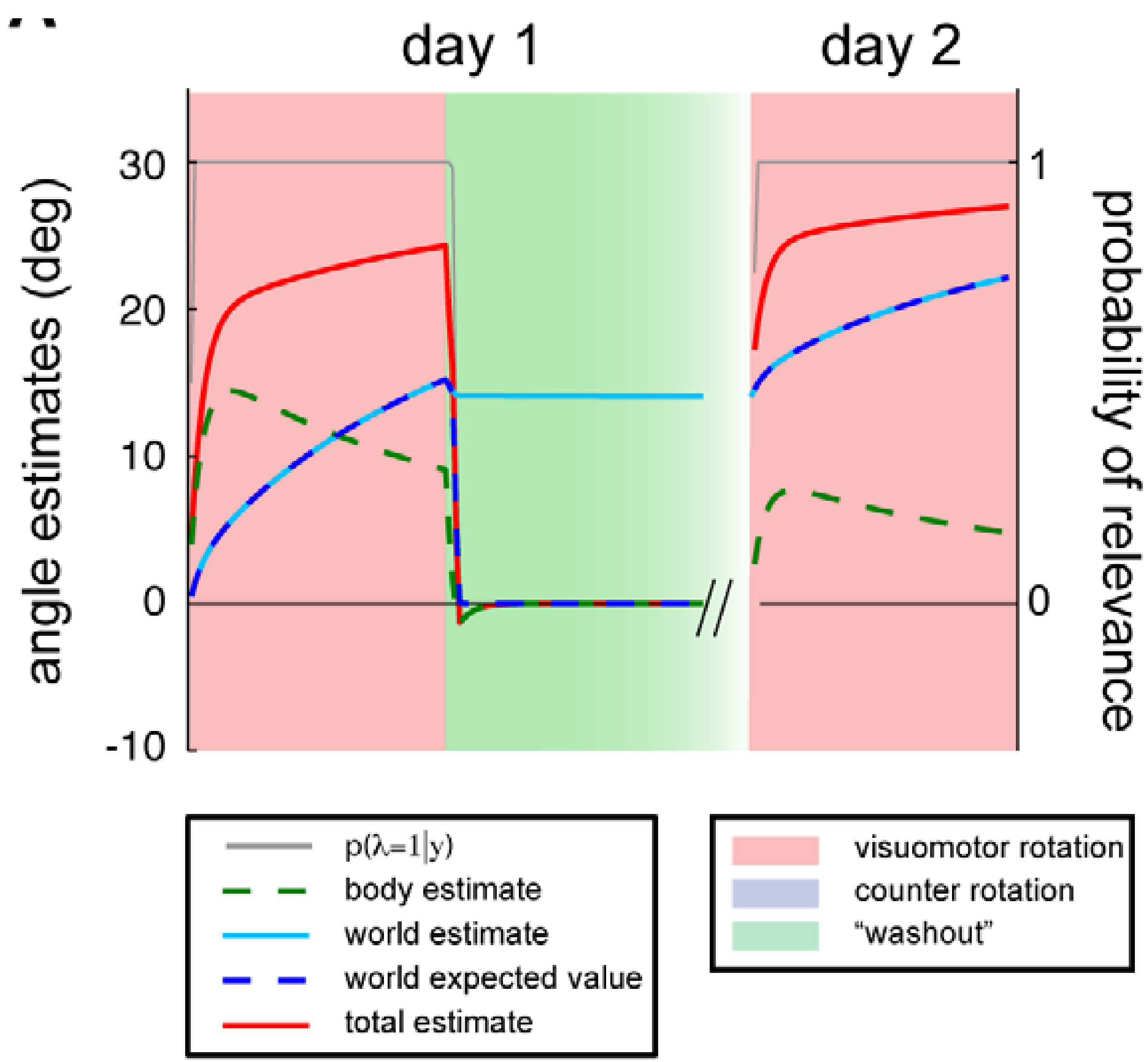
> This model can explain short-term savings:



Berniker & Kording (2011) PLoS Comp Bio 7: e1002210

Berniker & Kording (2011)

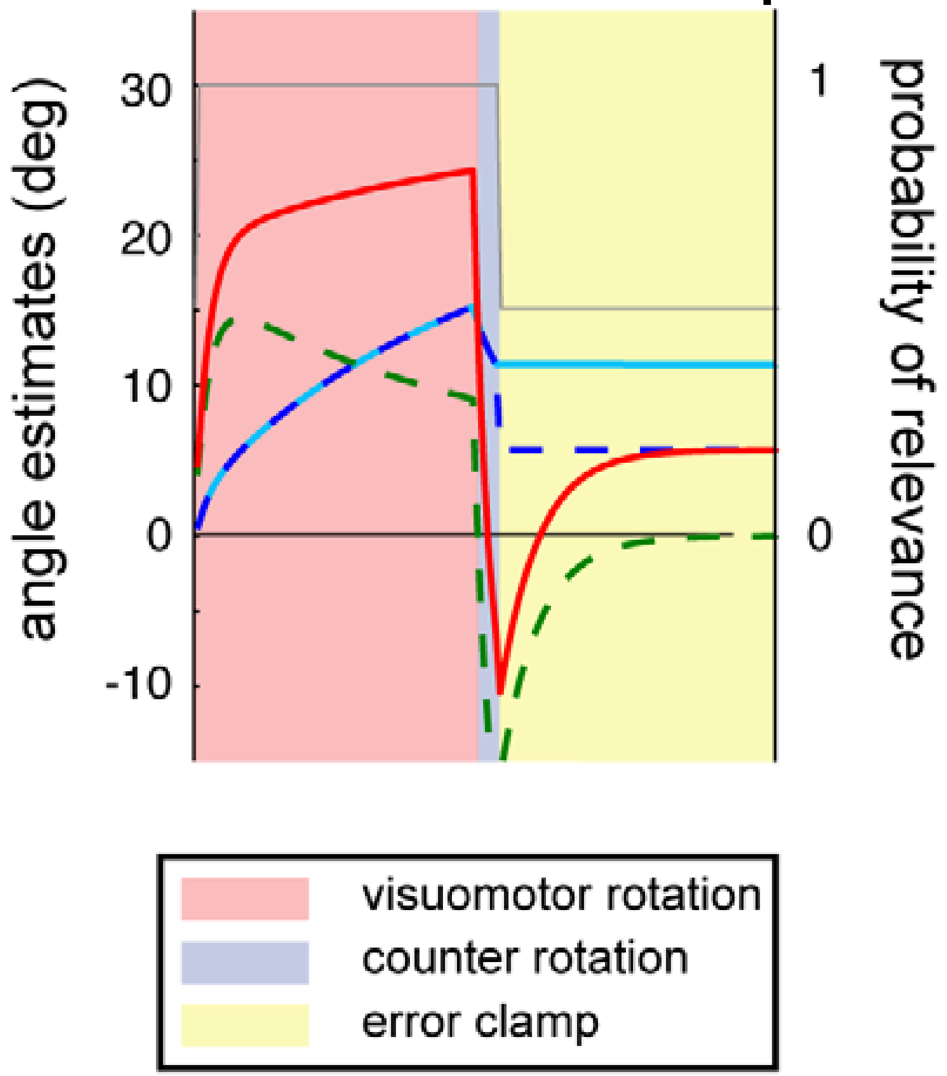
- This model can explain long-term savings:



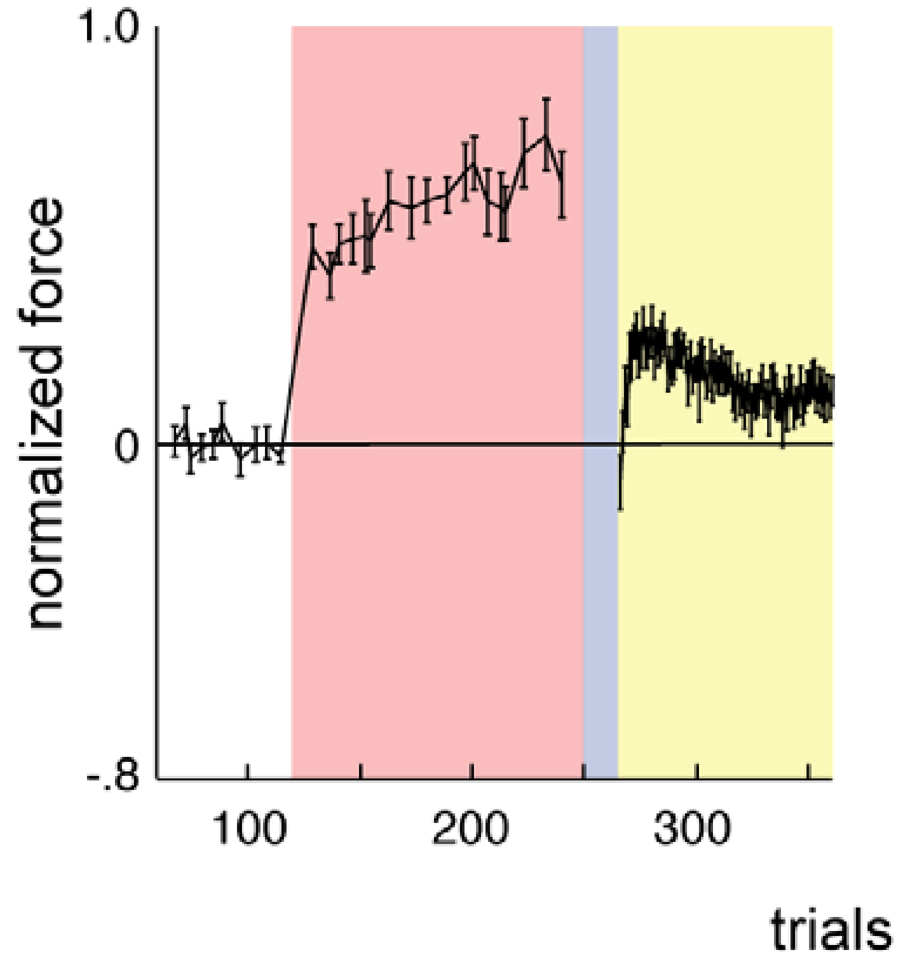
Berniker & Kording (2011) PLoS Comp Bio 7: e1002210

Berniker & Kording (2011)

> This model can explain spontaneous rebound:



Berniker & Kording (2011) PLoS Comp Bio 7: e1002210



> Other strong point: this model predicts sensory recalibration along with motor adaptation, which is also observed.

Advanced topics

- Noise
- Generalization
- Reinforcement learning
- Implicit vs. explicit learning

Noise

- All models considered so far: deterministic
- However, all neural signals are noisy.
- What role does noise play in motor adaptation?
- Two kinds of signals that contain noise:
  - sensory signals
    - error signals
  - motor signals
    - movement planning
    - movement execution

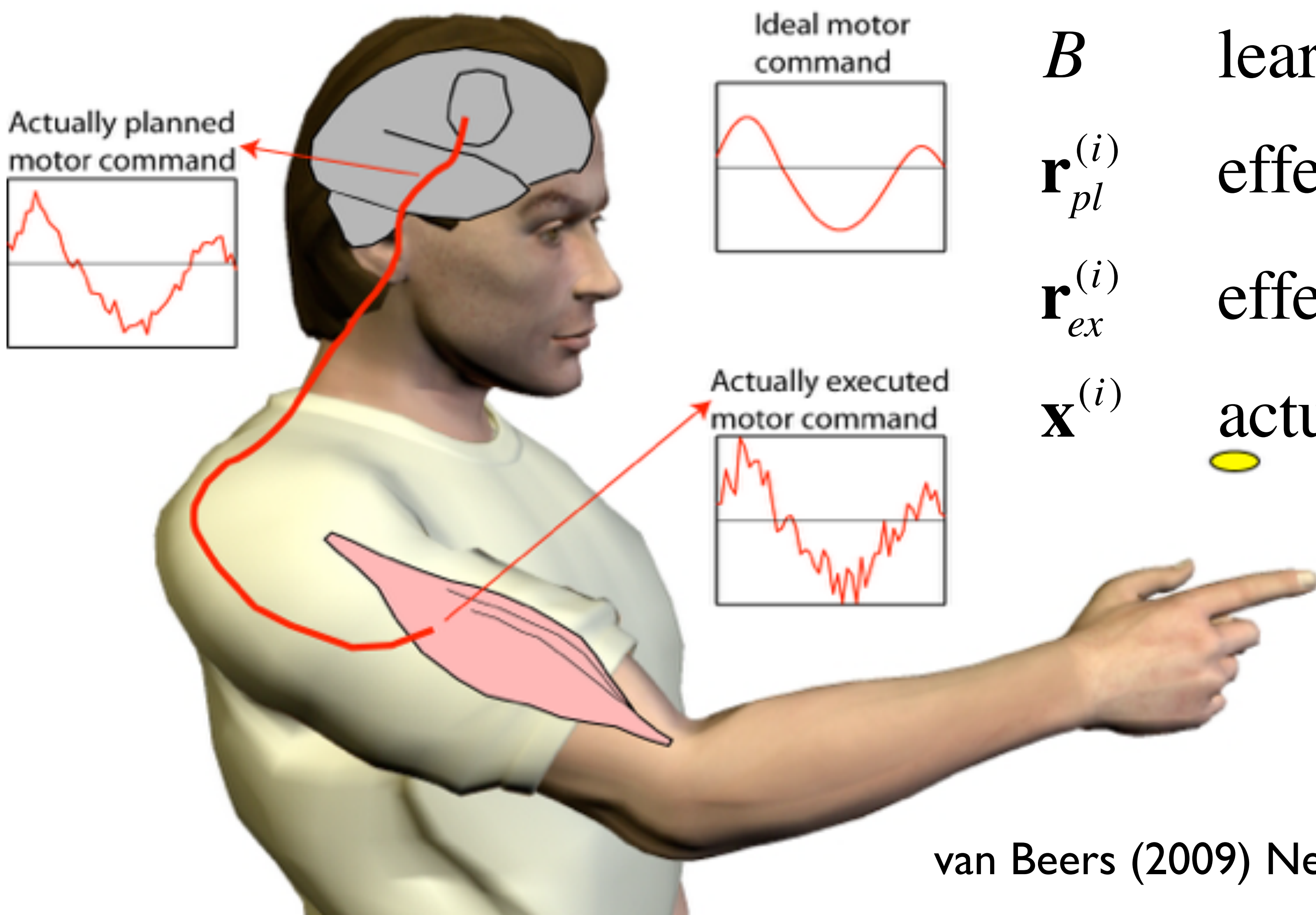
Planning and execution noise

> Model based on time-series analysis of repeated movements in absence of perturbations:

$$\mathbf{m}_{pl}^{(i+1)} = \mathbf{m}_{pl}^{(i)} - B\mathbf{e}^{(i)} + \mathbf{r}_{pl}^{(i+1)}$$

$$\mathbf{x}^{(i)} = \mathbf{m}_{pl}^{(i)} + \mathbf{r}_{ex}^{(i)}$$

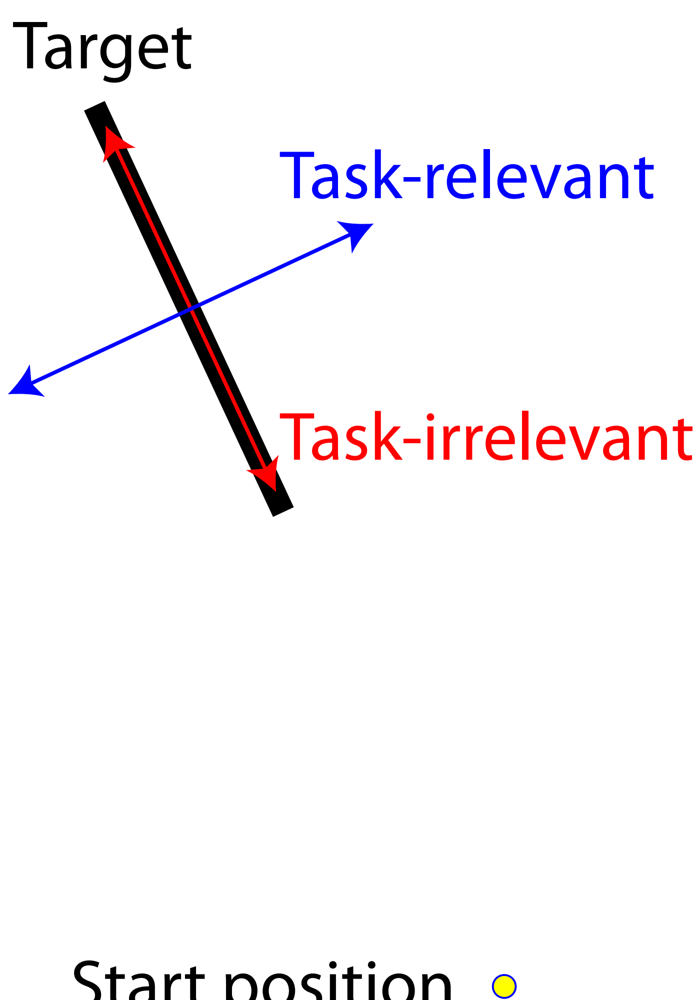
$\mathbf{m}_{pl}^{(t)}$	planned endpoint in movement i
$\mathbf{e}^{(i)}$	error in movement i
B	learning rate
$\mathbf{r}_{pl}^{(i)}$	effect of planning noise in movement i
$\mathbf{r}_{ex}^{(i)}$	effect of execution noise in movement i
$\mathbf{x}^{(i)}$	actual endpoint in movement i

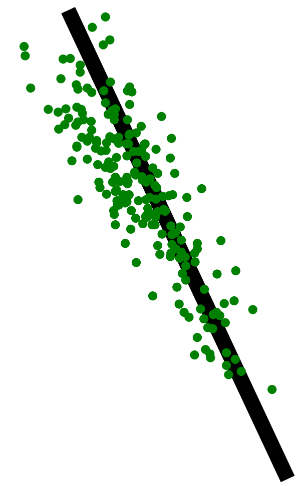


van Beers (2009) Neuron 63: 406-417

Planning and execution noise

- Effect of planning and execution noise with and without error-driven planning corrections:

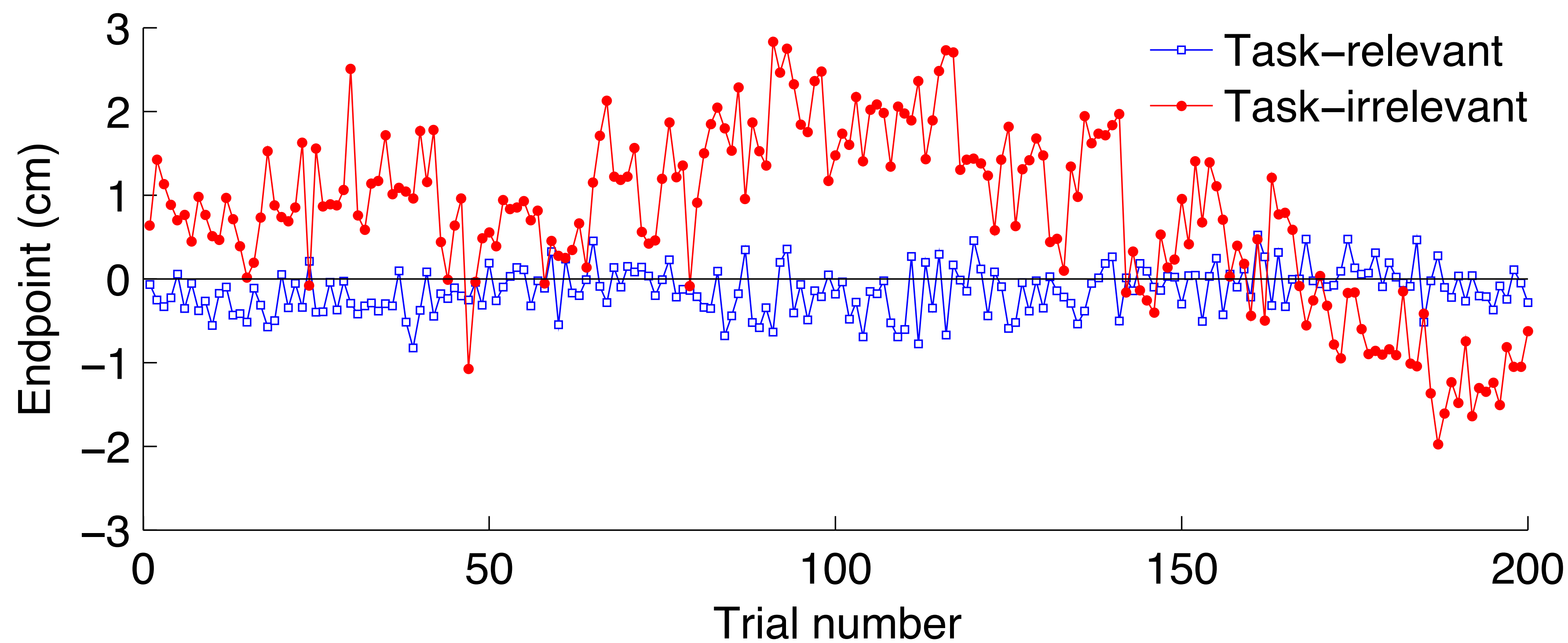




$$\mathbf{m}_{pl}^{(i+1)} = \mathbf{m}_{pl}^{(i)} - B\mathbf{e}^{(i)} + \mathbf{r}_{pl}^{(i+1)}$$

$$\mathbf{x}^{(i)} = \mathbf{m}_{pl}^{(i)} + \mathbf{r}_{ex}^{(i)}$$

- **Random walk** of motor planning in task-irrelevant dimensions.

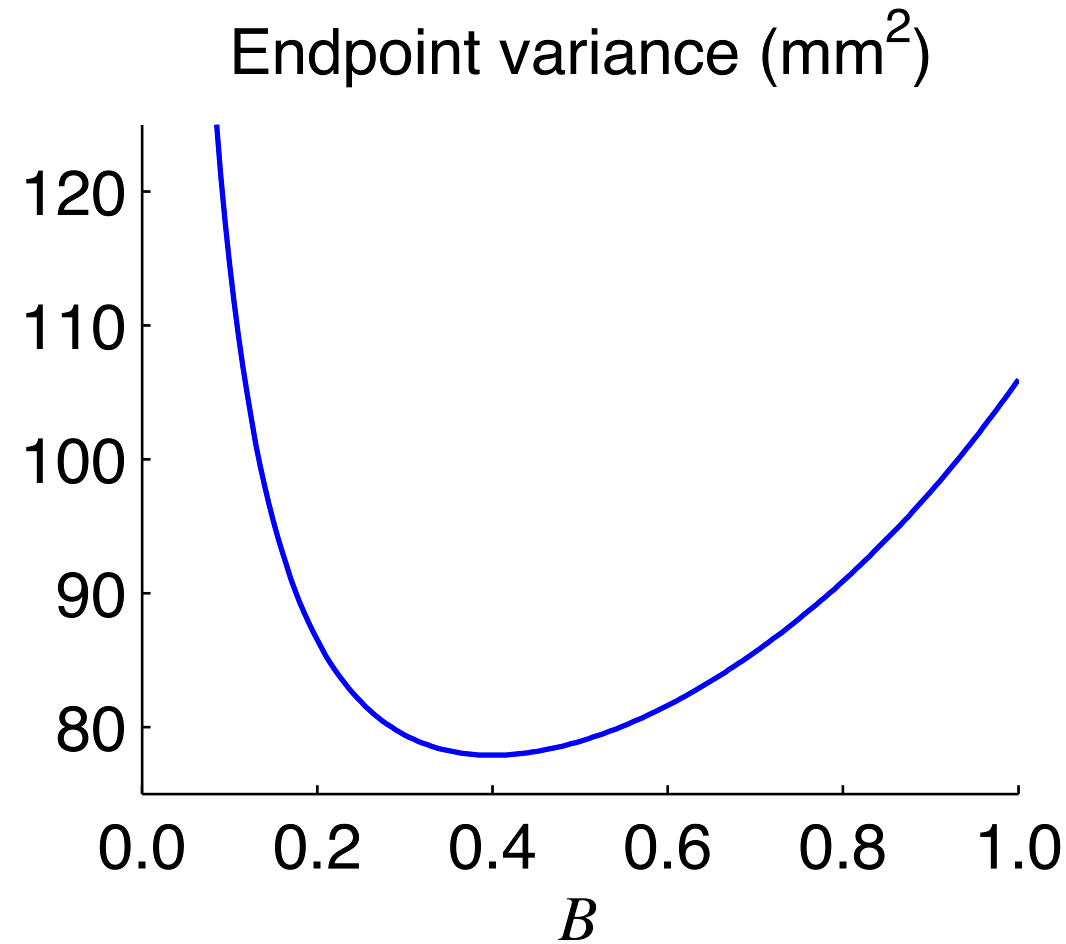


van Beers et al (2013) J Neurophysiol 109: 969-977

Planning and execution noise

- What is the learning rate in repeated, unperturbed movements?
- Data suggest: 0.4
- Why 0.4?

$$\mathrm{Var}(\mathbf{x}) = \frac{w + 2B(1-w)}{B(2-B)} \mathrm{Tr}(\boldsymbol{\Sigma}_{mot})$$



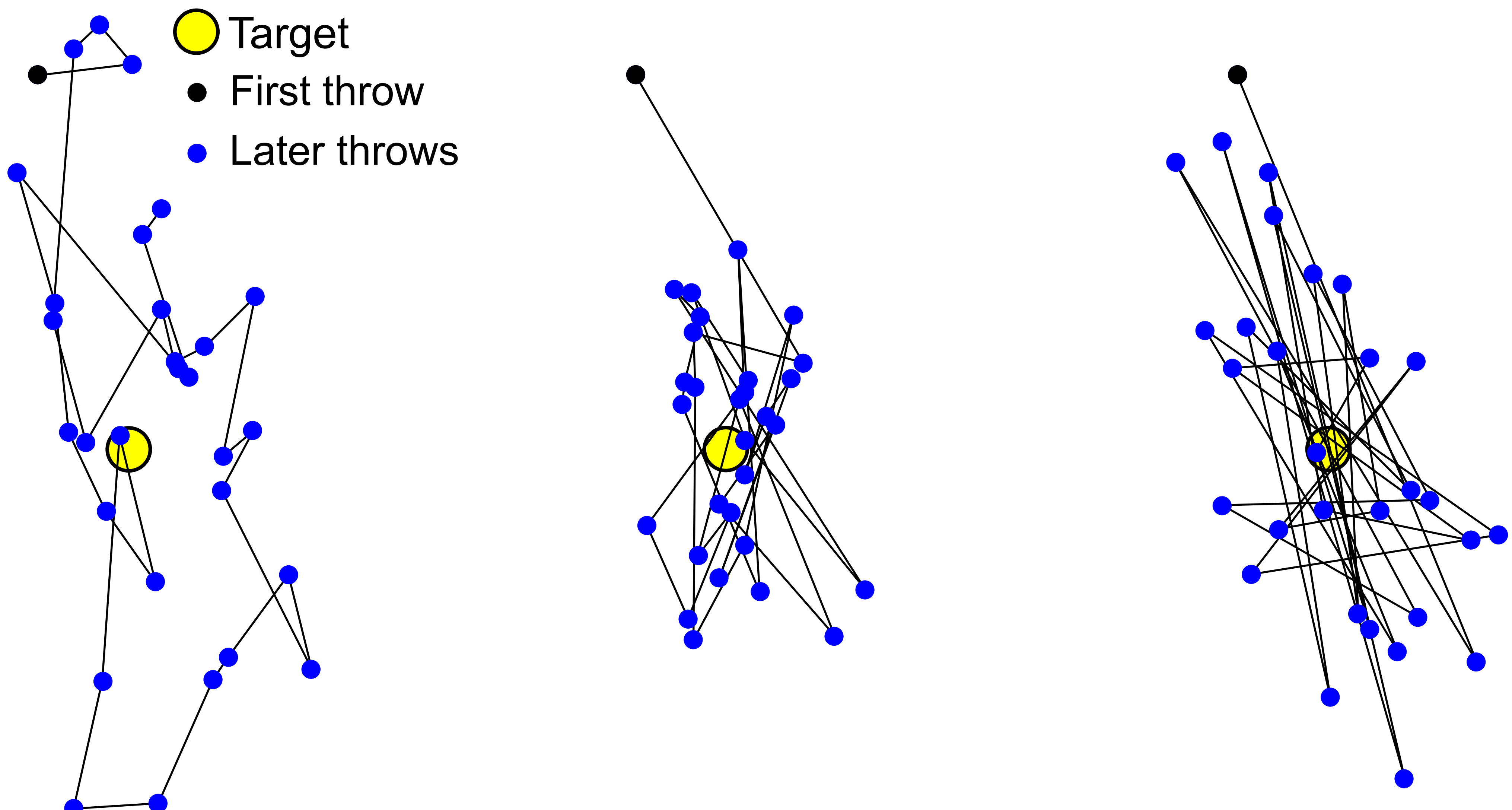
van Beers (2009) Neuron 63: 406-417

- The actual learning B rate minimizes the endpoint variance!

Planning and execution noise

- Why is a learning rate B of 0.4 optimal?

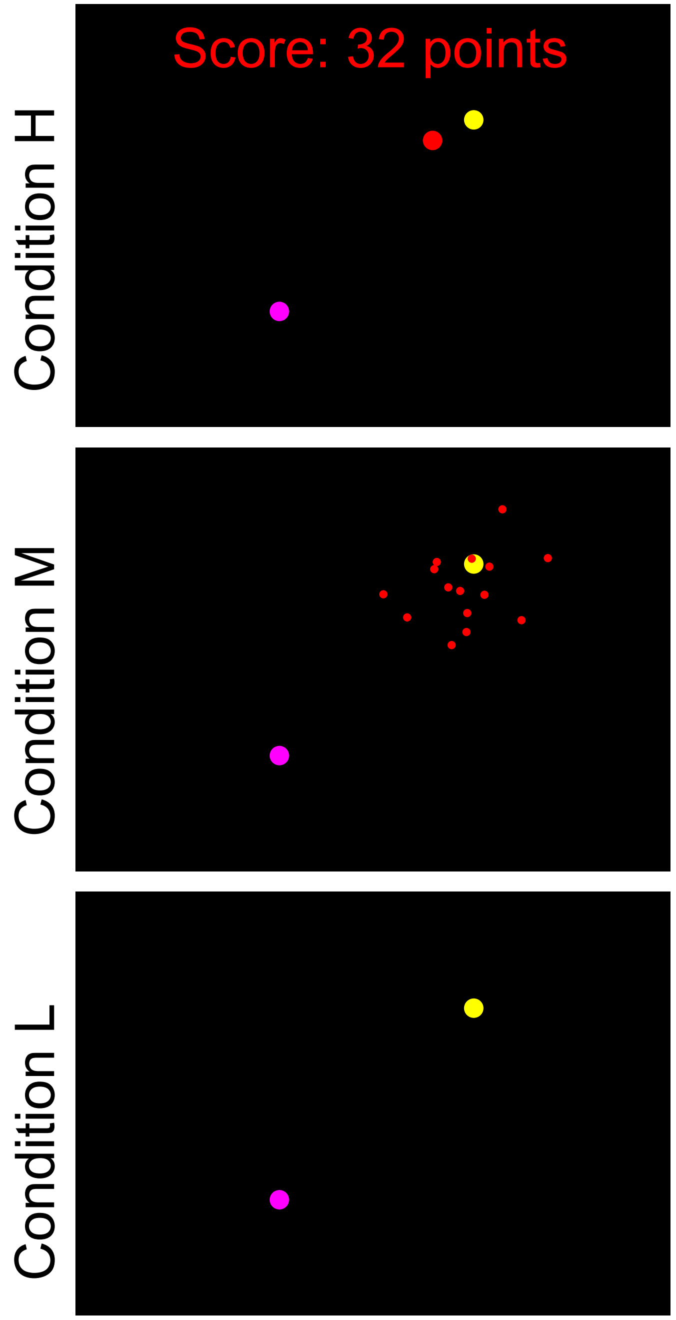
Small B: Intermediate B: Large B:



- corrections too small to reduce errors
- just right
- too much compensation for errors due to execution noise

Sensory and motor noise

> How does the learning rate change if sensory noise is increased?



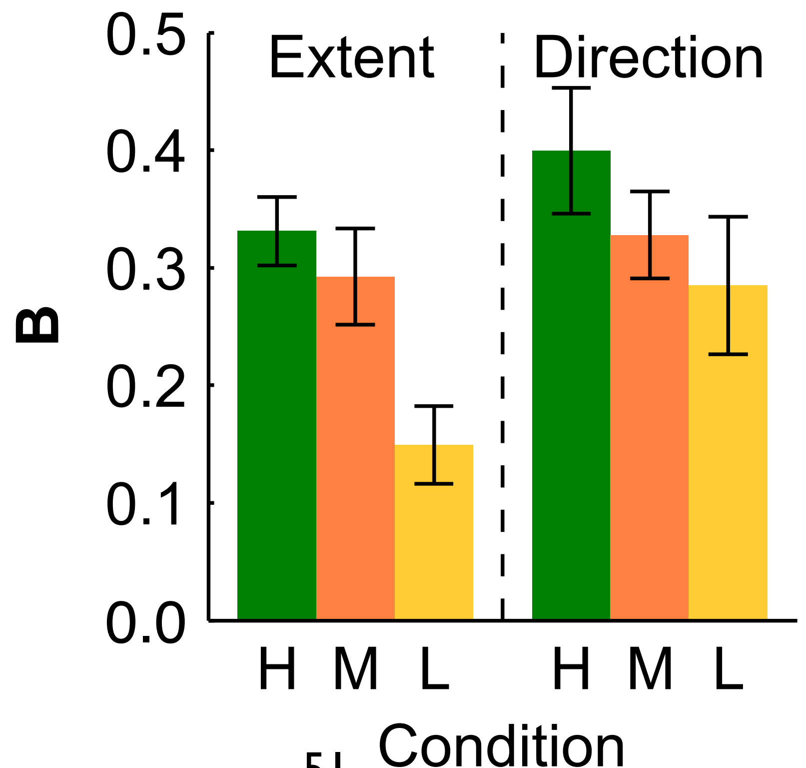
van Beers (2012) PLoS ONE 7: e49373

$$\mathbf{x}^{(t)} = \mathbf{m}_{pl}^{(t)} + \mathbf{r}_{ex}^{(t)}$$

$$\mathbf{e}^{(t)} = \mathbf{x}^{(t)} - \mathbf{x}_T$$

$$\hat{\mathbf{e}}^{(t)} = \mathbf{e}^{(t)} + \mathbf{r}_{sens}^{(t)}$$

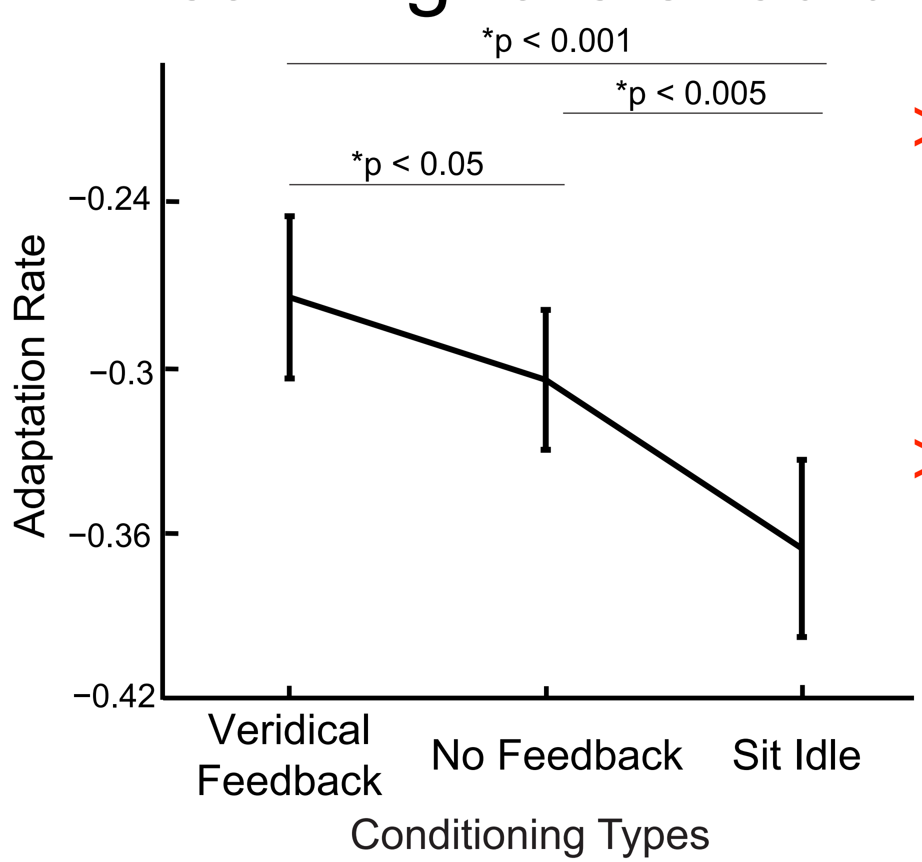
$$\mathbf{m}_{pl}^{(t+1)} = \mathbf{m}_{pl}^{(t)} - \mathbf{B}\hat{\mathbf{e}}^{(t)} + \mathbf{r}_{pl}^{(t+1)}$$



> Learning rate decreases if sensory noise is increased.

Sensory and motor noise

- Best possible strategy in presence of both sensory and motor noise: use a Kalman filter
  - optimal state estimator for linear dynamical system
  - predicts that learning rate decreases if sensory noise increases
  - other prediction of Kalman filter:
    - learning rate should increase if state estimate is more uncertain



- this has been confirmed for adaptation to a perturbation (Burge et al. (2008) J Vision 8:20; Wei & Körding (2010) Front Comput Neurosci 4:11)

- Despite this qualitative agreement with the Kalman filter, a quantitative test for the learning rate failed (van Beers (2012) PLoS ONE 7: e49373)

Conclusions regarding noise

> Noise in sensory and motor signals affect motor adaptation by determining the learning rate.
> They do so in an intuitive way to reduce variability.

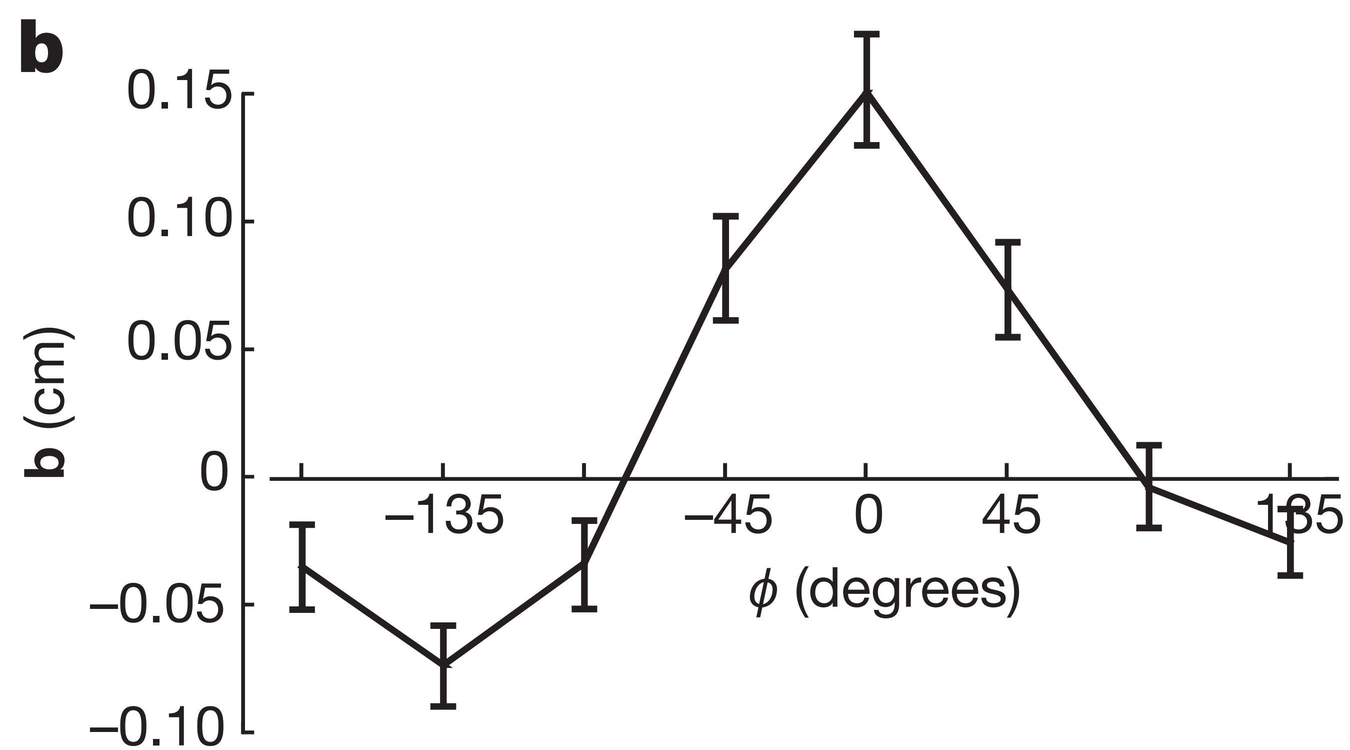
> Deterministic models are unrealistic:
> > they can be improved by adding noise terms

$$
\begin{aligned}
x_1(i+1) &= A_f x_1(i) + B_f e(i)\\
x_2(i+1) &= A_s x_2(i) + B_s e(i)\\
x(i) &= x_1(i) + x_2(i)
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
x_1(i+1) &= A_f x_1(i) + B_f e(i) + n_1(t)\\
x_2(i+1) &= A_s x_2(i) + B_s e(i) + n_2(t)\\
x(i) &= x_1(i) + x_2(i) + n_{out}(t)
\end{aligned}
$$

Generalization

- If we produce a movement error, motor planning is adjusted to avoid future errors.
- **To what extent do these adjustments generalize?**

- For reaching movements, generalization has been studied to other:
  - movement directions
  - movement amplitudes
  - movement speeds
  - spatial locations
  - arm
  - ...



Thoroughman & Shadmehr (2000) Nature 407: 742-747

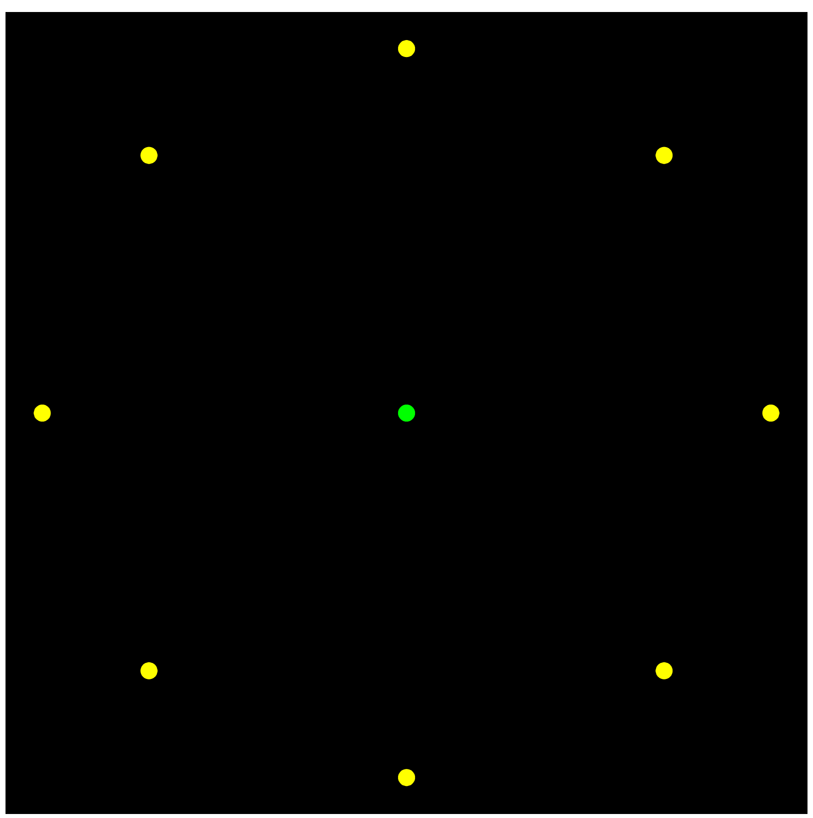
Modeling generalization

- Approach 1: straightforward extension of original two-state model

(Tanaka (2012) Neural Comput 24: 939-966):

$$x_1(i+1) = A_f x_1(i) + B_f e(i)$$
$$x_2(i+1) = A_s x_2(i) + B_s e(i)$$
$$x(i) = x_1(i) + x_2(i)$$

$$\mathbf{x}_1(i+1) = \mathbf{A}_f \mathbf{x}_1(i) + \mathbf{B}_f \mathbf{H}(i)^T e(i)$$
$$\mathbf{x}_2(i+1) = \mathbf{A}_s \mathbf{x}_2(i) + \mathbf{B}_s \mathbf{H}(i)^T e(i)$$
$$\mathbf{x}(i) = \mathbf{x}_1(i) + \mathbf{x}_2(i)$$



- Assume 8 targets:

$\mathbf{x}_1(i), \mathbf{x}_2(i)$	8×1
$\mathbf{A}_f, \mathbf{A}_s$	8×8 diagonal matrices
$\mathbf{B}_f, \mathbf{B}_s$	8×8 non-diagonal matrices
$\mathbf{H}(i)$	1×8 vector with 0's and one 1

- Off-diagonal elements of **B**'s: generalization from one direction to the others

Generalization

- Approach 2: use basis functions (Thoroughman & Shadmehr (2000) Nature 407: 742-747; Donchin et al. (2003) J Neurosci 23: 9032-9045):
- Write motor output $\mathbf{f}$ as linear combination of basis functions:

$$\mathbf{f} = \sum_i w_i g_i(\mathbf{x}) = \mathbf{w}^T \mathbf{g}(\mathbf{x})$$

- Learning from error at state $\mathbf{x}_1$ changes the weights:

$$\Delta \mathbf{w}_1 = -\eta \mathbf{g}(\mathbf{x}_1) \left| \mathbf{f} - \hat{\mathbf{f}} \right|$$

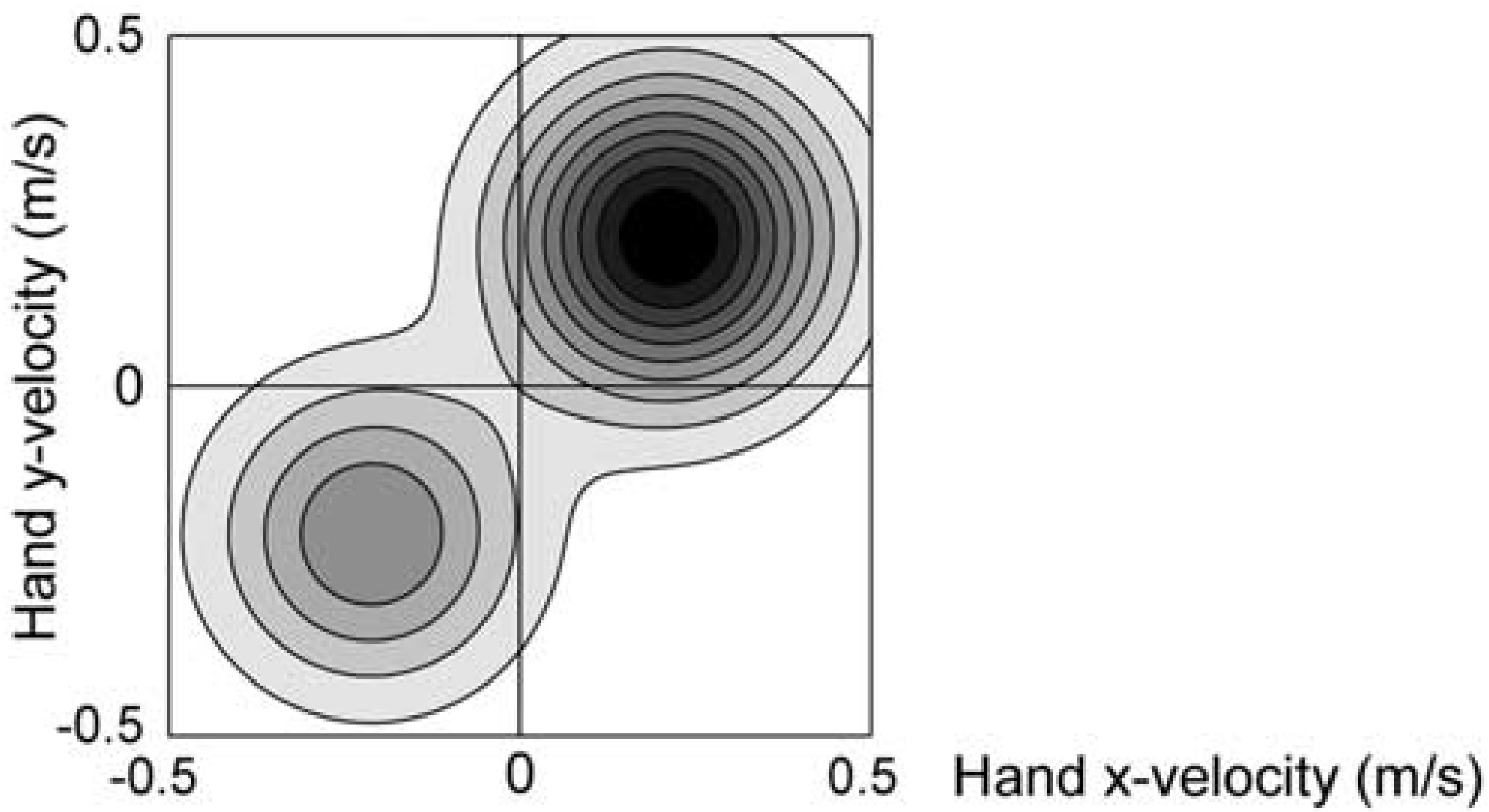
- Motor output in direction 2 is than changed by amount:

$$\mathbf{f}_2(\mathbf{w} + \Delta \mathbf{w}_1) - \mathbf{f}_2(\mathbf{w}) = (\mathbf{w} + \Delta \mathbf{w}_1)^T \mathbf{g}(\mathbf{x}_2) - \mathbf{w}^T \mathbf{g}(\mathbf{x}_2) = -\eta \mathbf{g}(\mathbf{x}_1)^T \mathbf{g}(\mathbf{x}_2) \left| \mathbf{f} - \hat{\mathbf{f}} \right|$$

- Generalization is thus determined by matrix: $\mathbf{g}(\mathbf{x}_1)^T \mathbf{g}(\mathbf{x}_2)$

Generalization

- Generalization can therefore be used to estimate the basis functions.
- Example of estimated basis function (in velocity space):

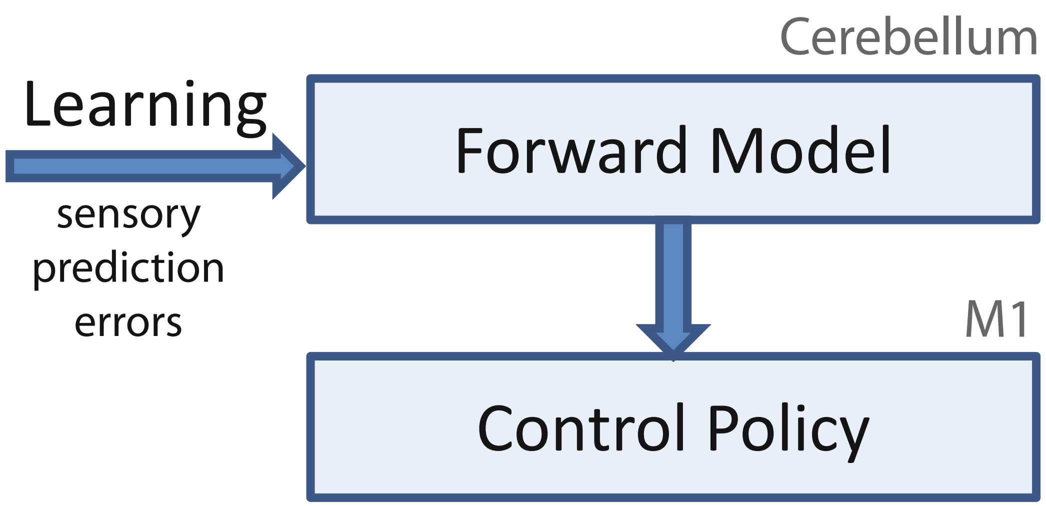


Shadmehr (2004) Hum Mov Sci 23: 543-568

Reinforcement learning

> Assumed so far:

a ***Model-based learning***



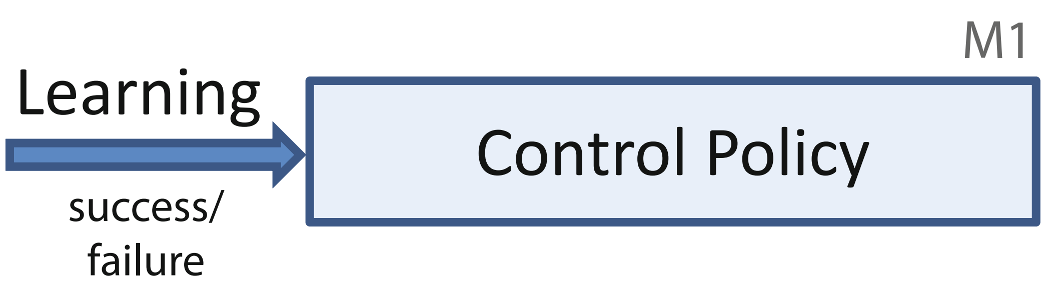
Indirect:

- Identify perturbation
- Plan movement

Haith & Krakauer (2013) Adv Exp Med Biol 782: 1-21

> Alternative learning mechanism:

b ***Model-free learning***



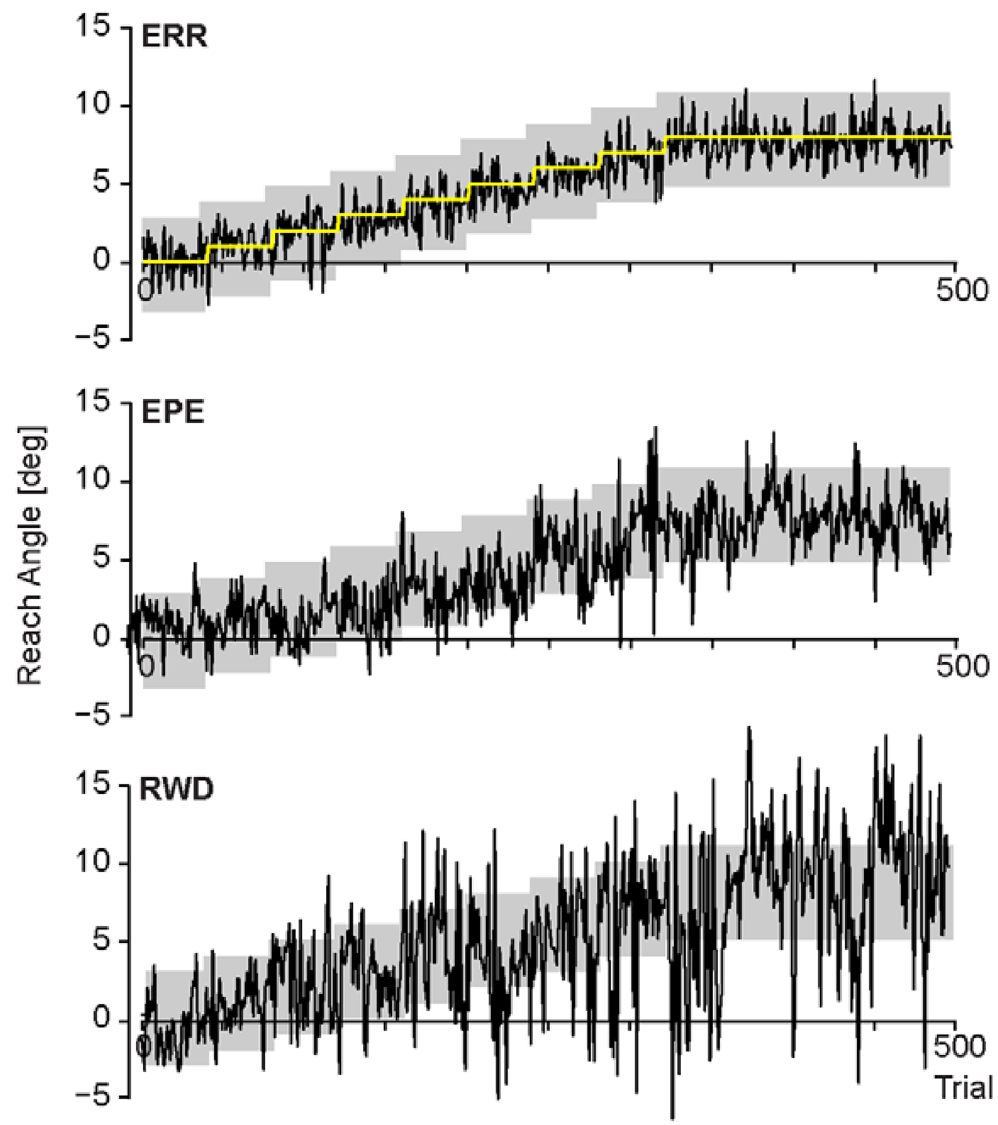
Direct:

- Remember and repeat actions that led to success

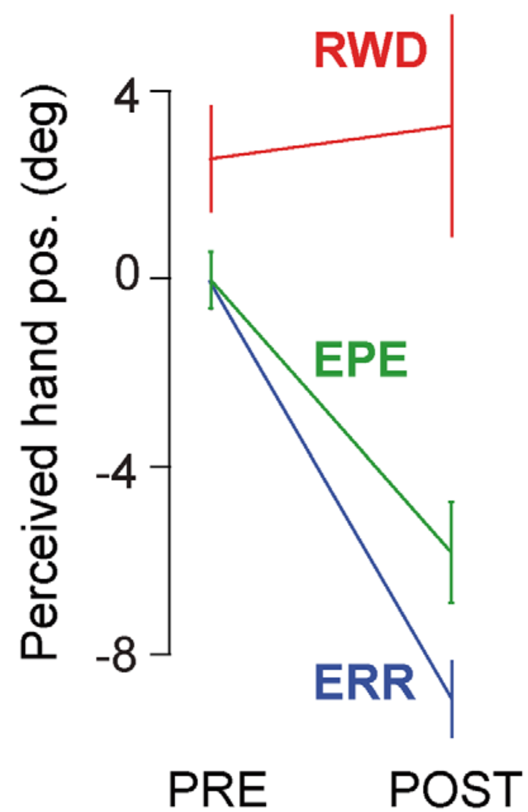
> **Reinforcement learning**

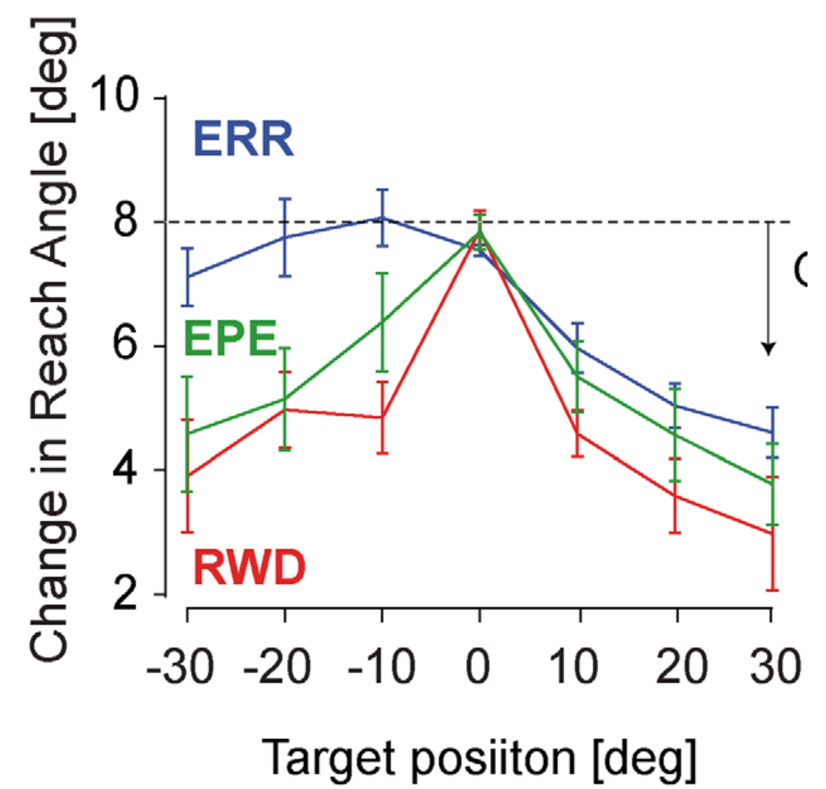
Reinforcement learning

- > To what extent contribute model-based and model-free learning to motor adaptation?
- > Is it possible to adapt on the basis of reward signals only?



- > Yes, but sensory recalibration and generalisation are different:

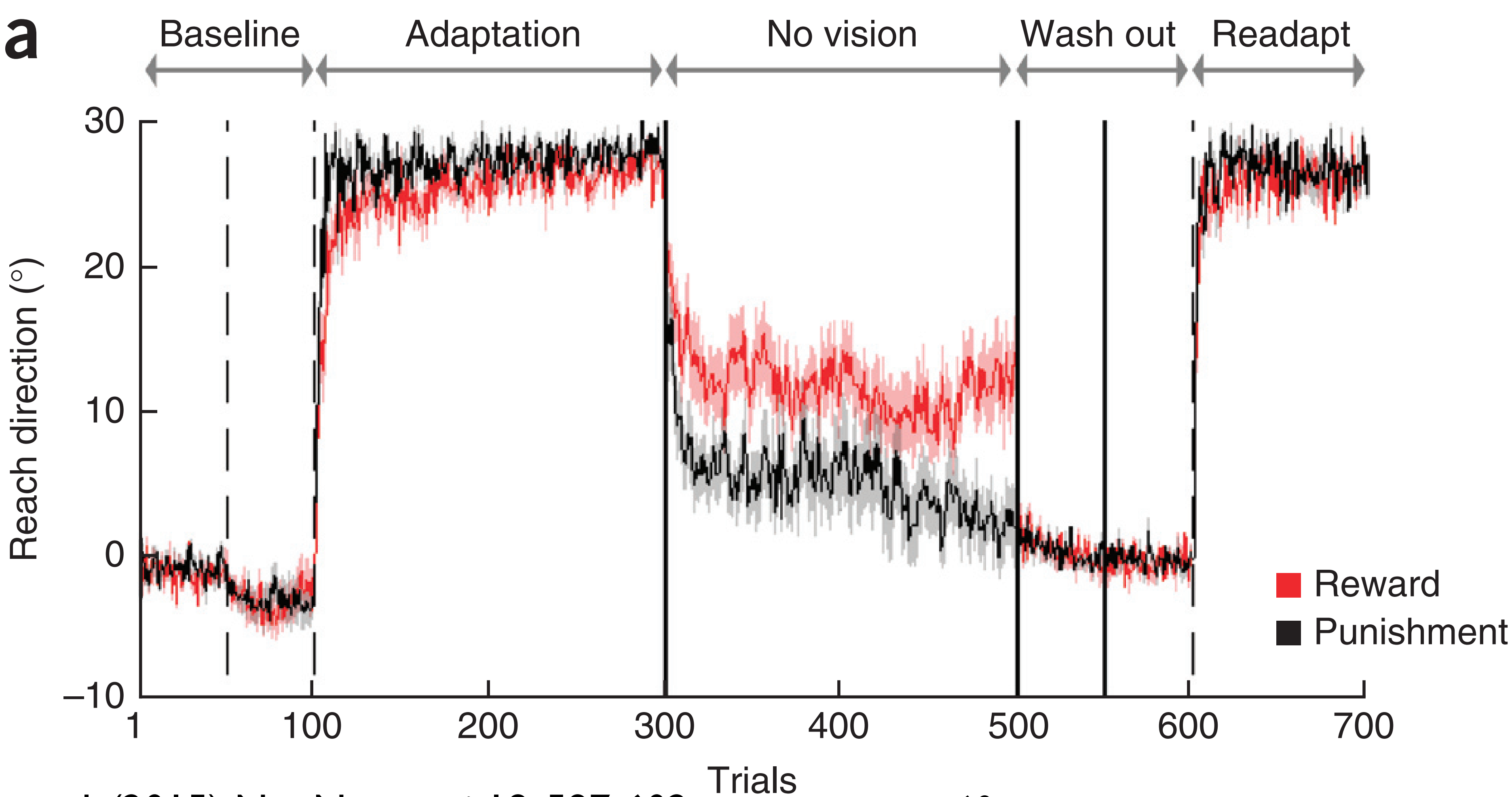




Izawa & Shadmehr (2011) PLoS Comput Biol 7: e1002012

Reinforcement learning

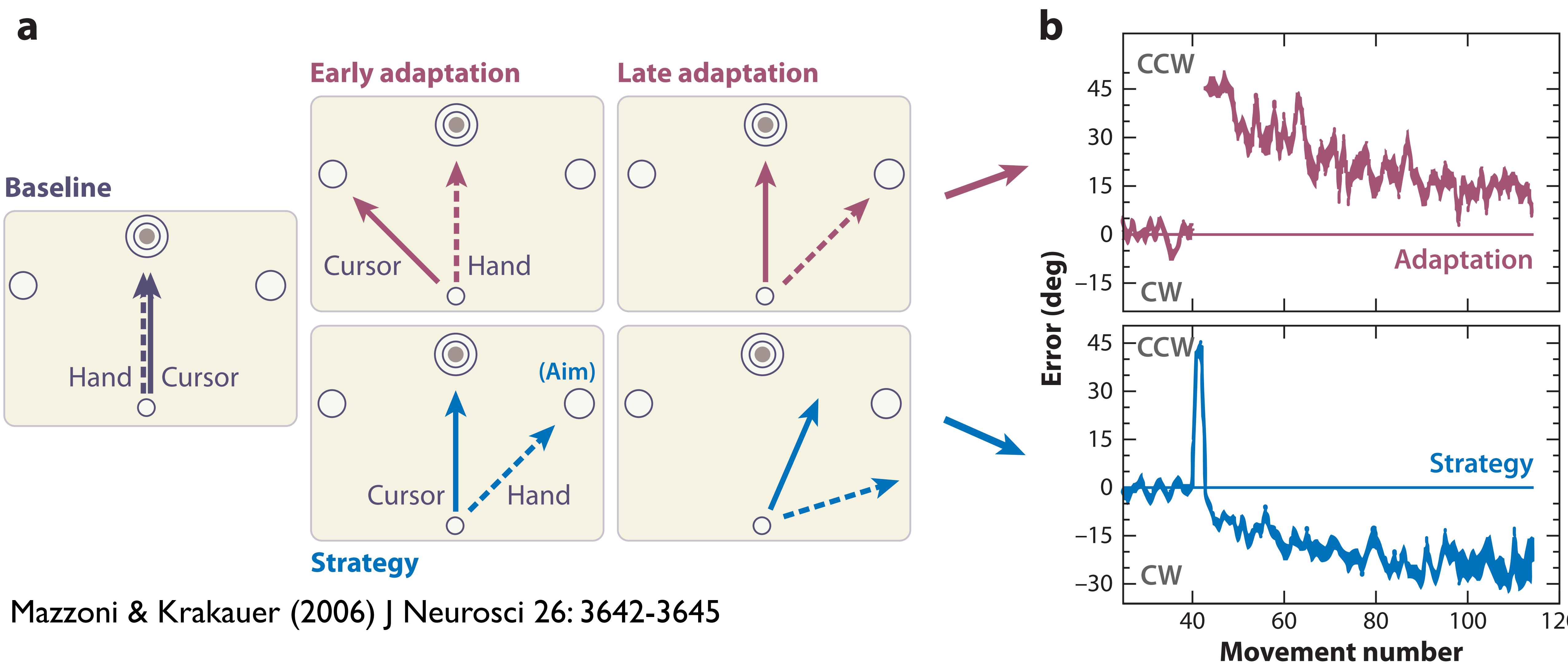
- More recent research suggests that reinforcement learning plays little or no role if error-based learning is possible.
- Other recent finding: adaptation and retention are different if subjects are rewarded vs punished.



Galea et al. (2015) Nat Neurosci 18: 597-602

Implicit vs. explicit learning

- > Another complication:
  - > implicit and explicit mechanisms play different roles



Mazzoni & Krakauer (2006) J Neurosci 26: 3642-3645

Summary

- Motor adaptation is reasonably well understood.
- Models that can explain many observations:
  - two-state model (Smith et al. 2006)
  - parallel one-fast, multiple-slow process model (Lee & Schweighofer 2009)
  - error-dependent learning rate model (Herzfeld et al. 2014)
  - relevance-estimation model (Berniker & Kording 2011)
- These models can also explain generalization.
- Consequences of noise can explain learning rates.
- Issues that also play a role:
  - reinforcement learning
  - implicit vs. explicit learning

Exercises

- Program some of the major models in Matlab.
- Analyze their behaviour.