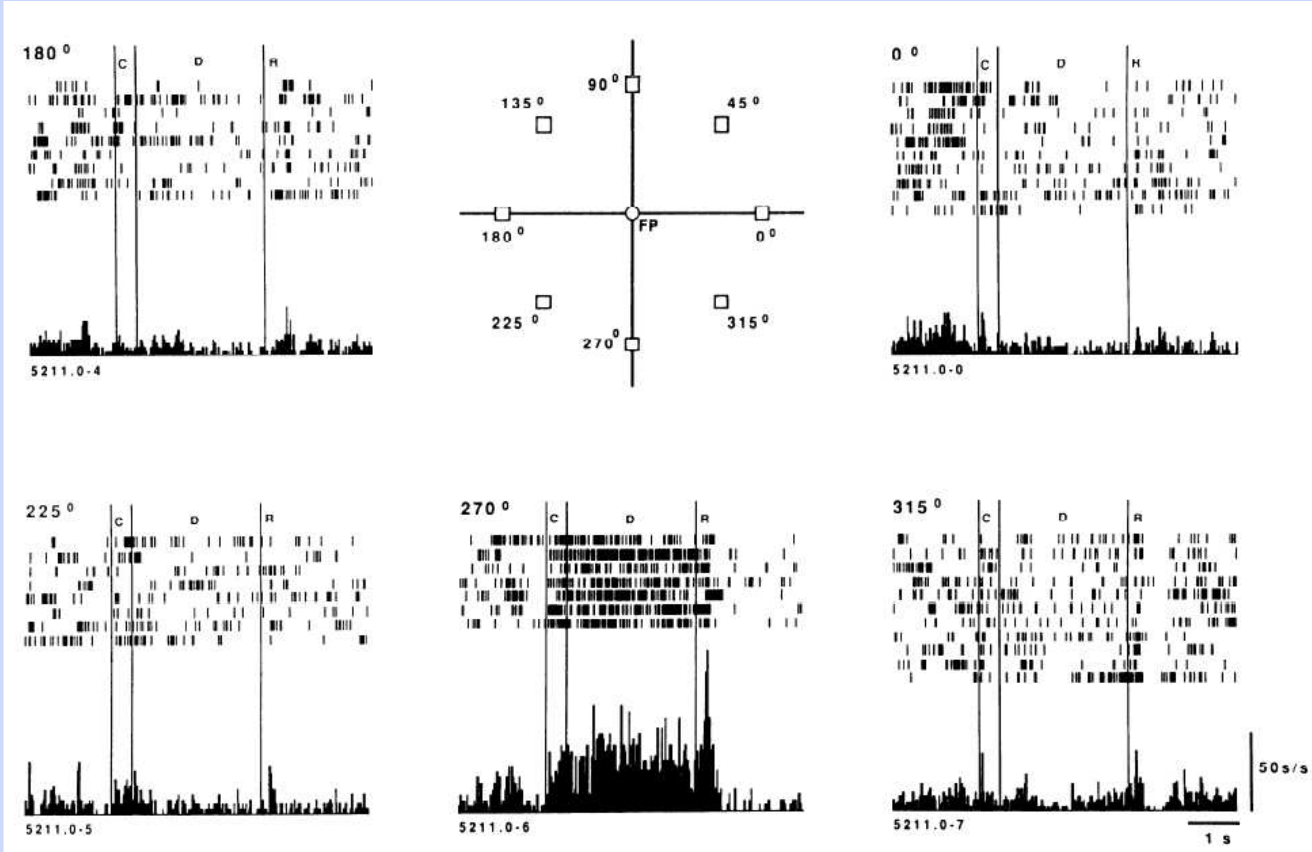
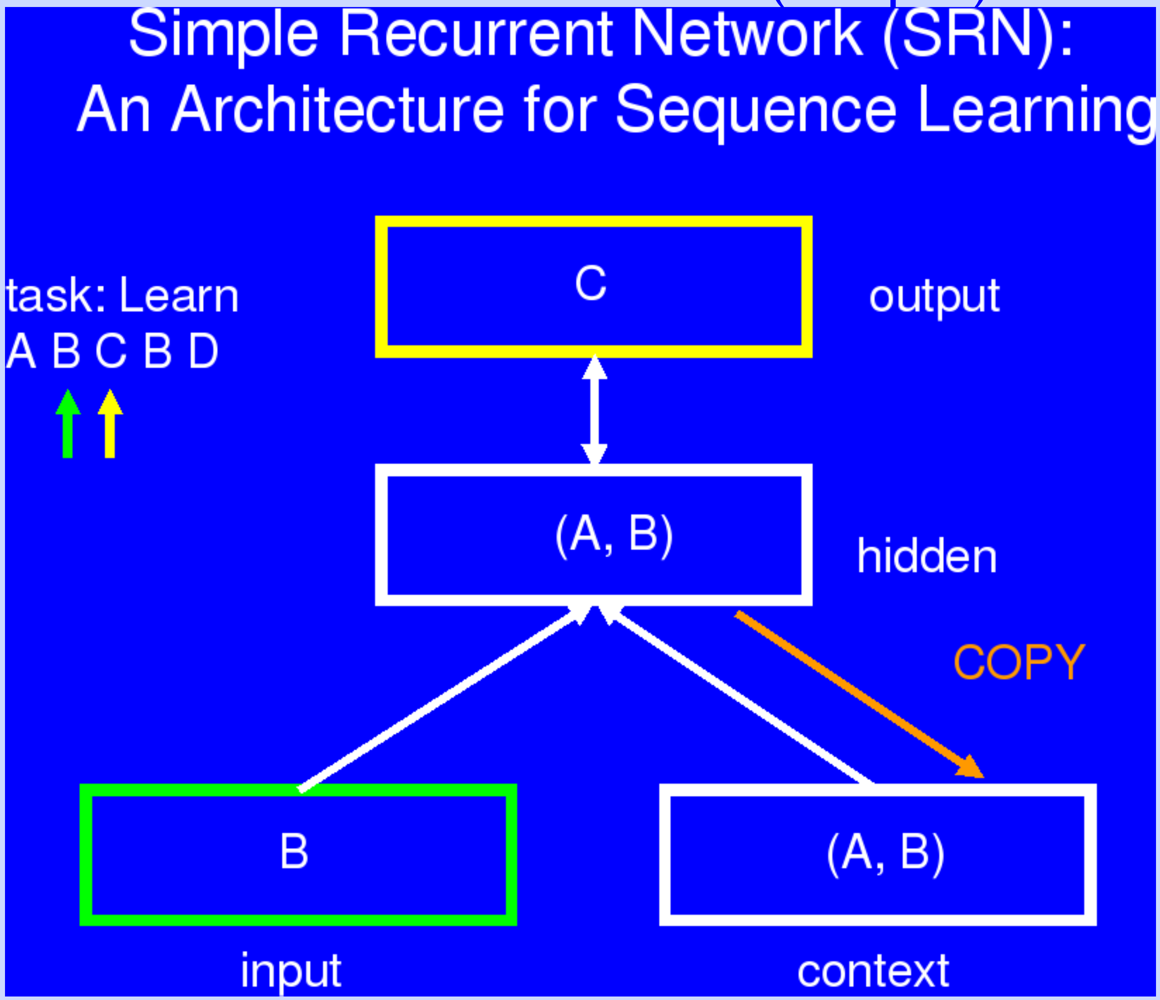
Prefrontal Cortex: Delay-related activity

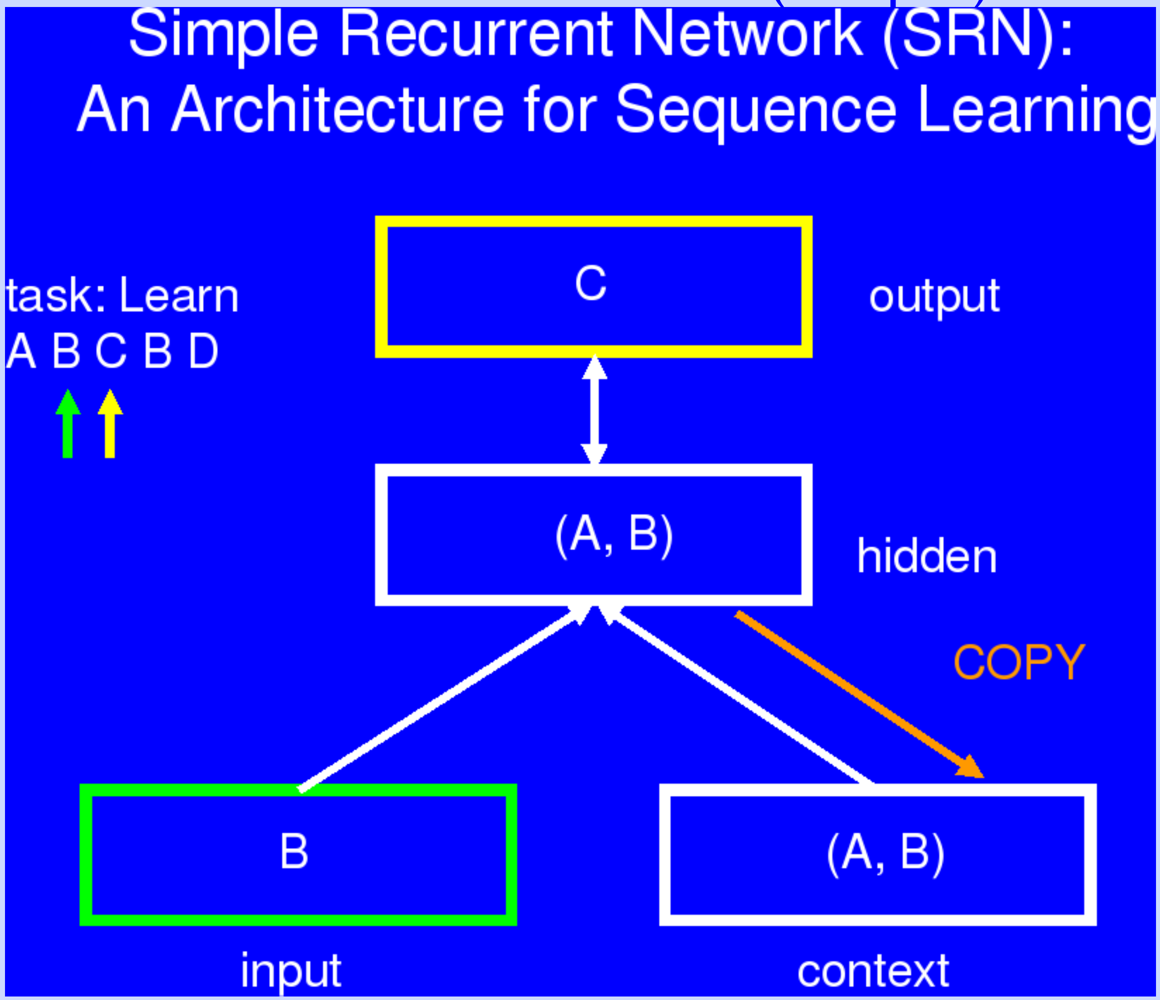


Spatial delayed-response task; Funahashi et al, 1989

Remember the SRN? (chap 6)



Remember the SRN? (chap 6)



this is a *gating* network: context only updated at discrete timepoints

Simple SRN story is not flawless

- How is hidden→ “copy” function implemented biologically?
- During settling, context must be *actively maintained* (ongoing hidden activity has no effect on context).
- Assumes all context is relevant: What if distracting information presented in middle of sequence? Want to only hold on to *relevant* context.
- What if want to hold on to more than one piece of information in WM at a time?? Or to selectively update one part of WM while continuing to robustly maintain others?
- And what if the decision of whether or not to update information *depends* on currently internal WM state?

Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.

Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.

- But PFC patients don't have delusions, hallucinations, psychosis. Also some imaging studies show *increase* in PFC activity in SZ (e.g., Manoach, 2003; Calicott et al, 2003)

Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.
- But PFC patients don't have delusions, hallucinations, psychosis. Also some imaging studies show *increase* in PFC activity in SZ (e.g., Manoach, 2003; Calicott et al, 2003)
- Psychosis thought to stem from increased DA in BG (Weinberger, O'Donnell, Grace, etc)

Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.

- But PFC patients don't have delusions, hallucinations, psychosis. Also some imaging studies show *increase* in PFC activity in SZ (e.g., Manoach, 2003; Calicott et al, 2003)

- Psychosis thought to stem from increased DA in BG (Weinberger, O'Donnell, Grace, etc)

- Other disorders with BG/DA dysfunction are associated with frontal-like cognitive deficits (PD, addiction, ADHD, etc); these deficits correlate w/ DA in BG (Muller et al, 2000; Remy et al, 2000)

Widely Accepted Role of Prefrontal Cortex (PFC)

PFC helps keep us on task, promotes cognitive flexibility via ability to rapidly switch state (e.g., task switching):

- Robust maintenance of neural activity (*working memory*)

Widely Accepted Role of Prefrontal Cortex (PFC)

PFC helps keep us on task, promotes cognitive flexibility via ability to rapidly switch state (e.g., task switching):

- Robust maintenance of neural activity (*working memory*)
- Flexibility from *adaptive gating* (via Basal Ganglia, BG) that switches between maintenance and rapid updating

Widely Accepted Role of Prefrontal Cortex (PFC)

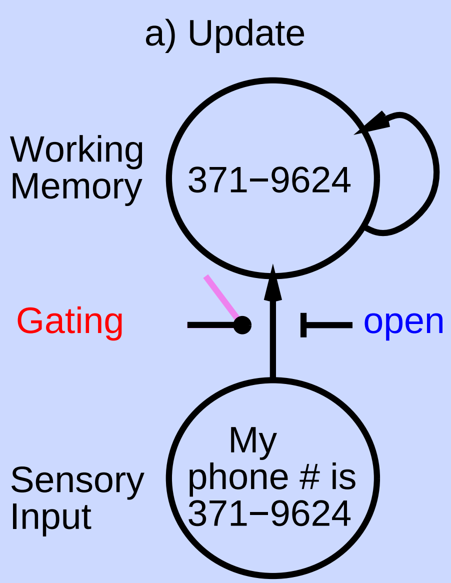
PFC helps keep us on task, promotes cognitive flexibility via ability to rapidly switch state (e.g., task switching):

- Robust maintenance of neural activity (*working memory*)
- Flexibility from *adaptive gating* (via Basal Ganglia, BG) that switches between maintenance and rapid updating
- Extensive interconnectivity, allowing *top-down biasing* of task-relevant info in other cortical / subcortical areas.

In short, PFC is the "central executive!"

Working Memory Demands: Updating & Maintenance

Hochreiter & Schmidhuber, 1997; Braver & Cohen, 1999; Frank et al, 2001

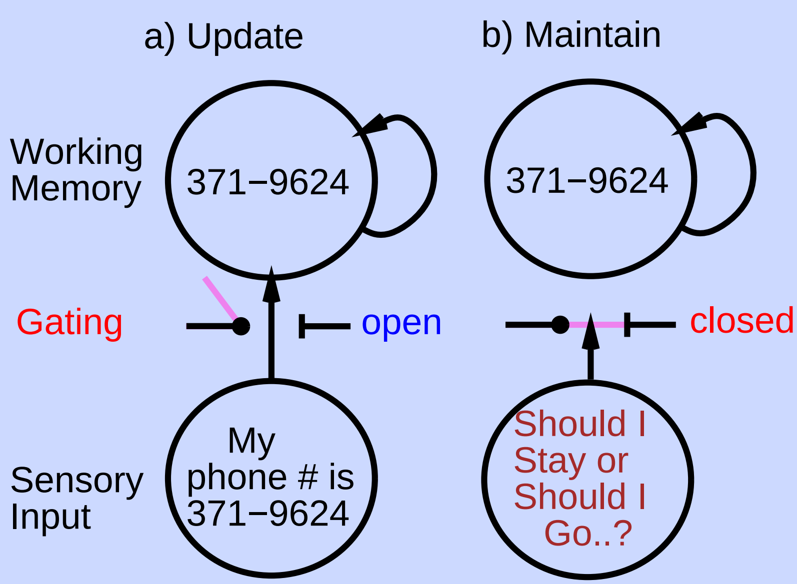


- Working memory: robust maintenance of information, but must also have ability to be rapidly updated — requires *gating*.

- You've got to know when to hold 'em, know when to fold 'em.

Working Memory Demands: Updating & Maintenance

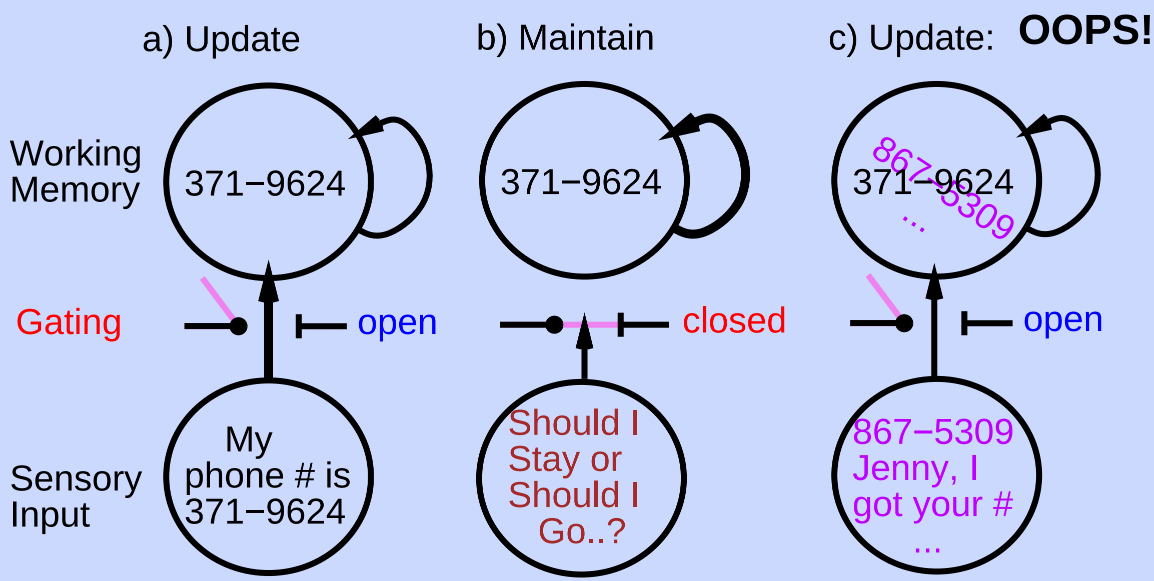
Hochreiter & Schmidhuber, 1997; Braver & Cohen, 1999; Frank et al, 2001



- Working memory: robust maintenance of information, but must also have ability to be rapidly updated — requires *gating*.
- You've got to know when to hold 'em, know when to fold 'em.

Working Memory Demands: Updating & Maintenance

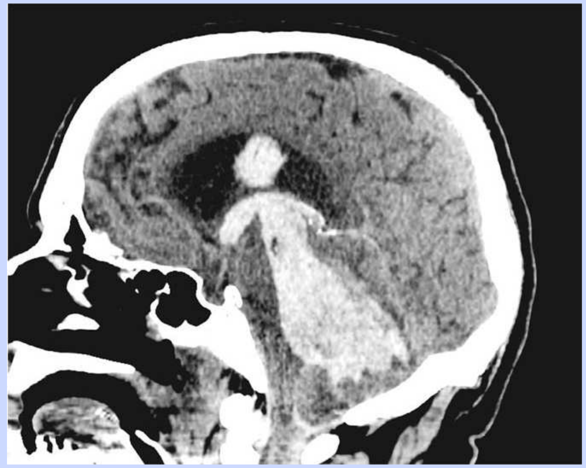
Hochreiter & Schmidhuber, 1997; Braver & Cohen, 1999; Frank et al, 2001



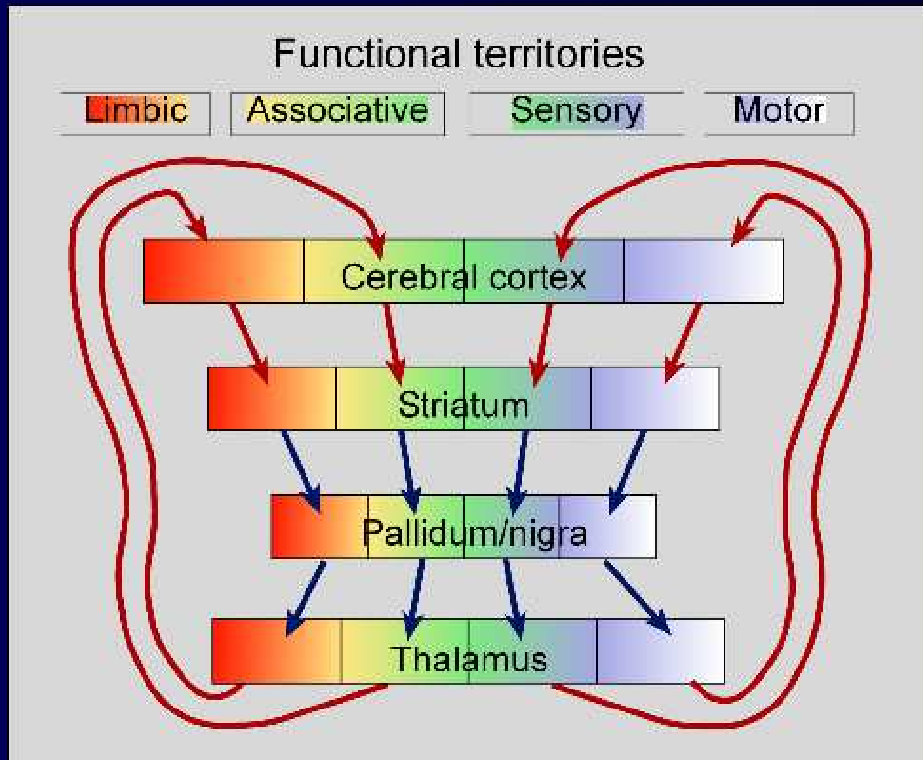
- Working memory: robust maintenance of information, but must also have ability to be rapidly updated — requires *gating*.

- You've got to know when to hold 'em, know when to fold 'em.

But who controls the controller?



Basal Ganglia Architecture: Cortically based loops



Alexander, G. E., et al. (1986). "Parallel organization of functionally segregated circuits linking basal ganglia and cortex." Ann. Rev. Neurosci. **9**: 357-381.

BG damage ⇒ deficits in motor, learning, motivation, working memory, cognitive control

BG Model Extension to Working Memory and Attention

- In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996)

BG Model Extension to Working Memory and Attention

- In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996)
- In parallel circuits, the *BG* may reinforce the *updating* of PFC working memory representations (Alexander et al, 1986; Frank, Loughry & O'Reilly, 2001).

BG Model Extension to Working Memory and Attention

- In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996)
- In parallel circuits, the *BG* may reinforce the *updating* of PFC working memory representations (Alexander et al, 1986; Frank, Loughry & O'Reilly, 2001).
- Dopamine in *PFC* supports robust *maintenance* over time (Lewis & O'Donnell, 2000; Durstewitz & Seemans, 2002)

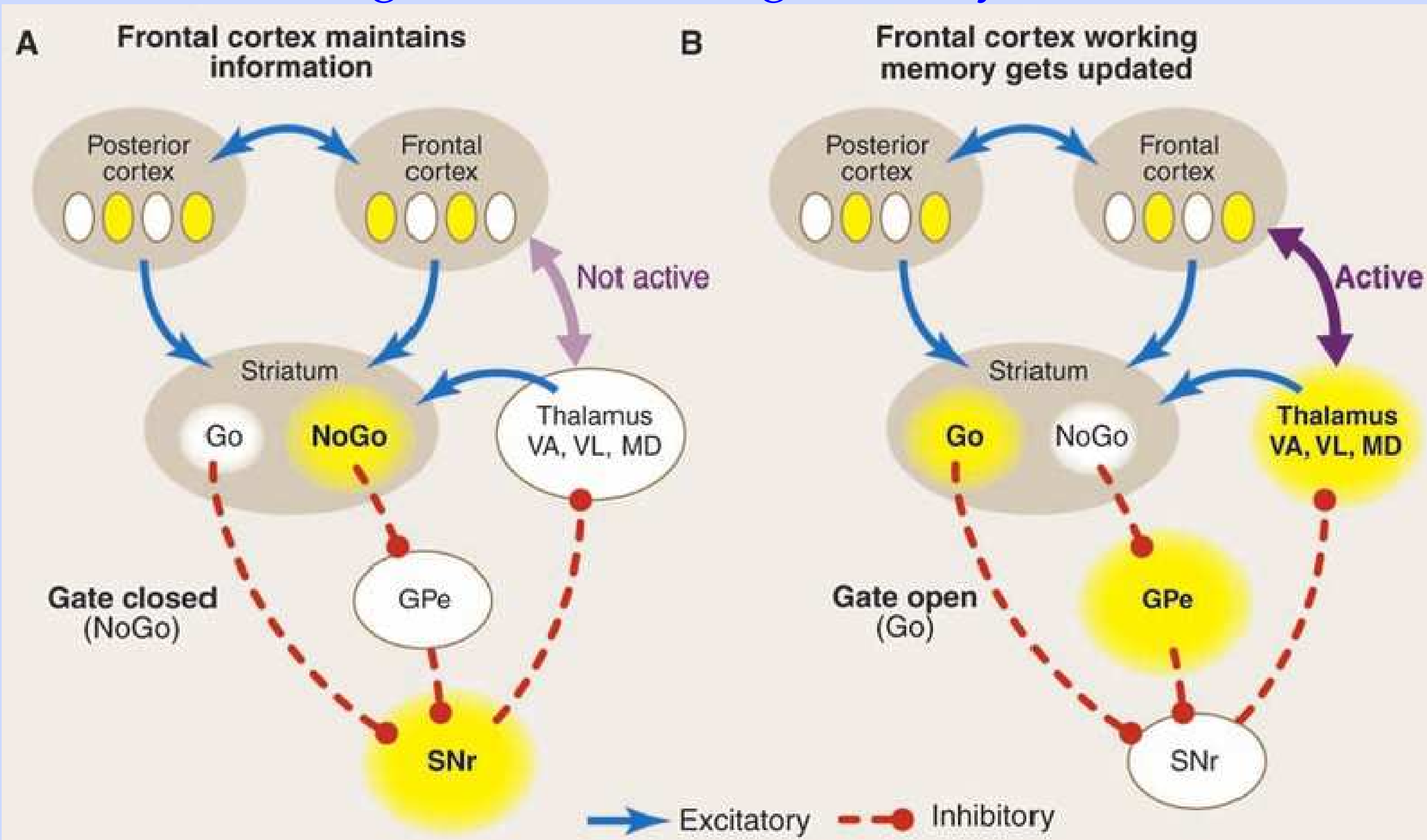
BG Model Extension to Working Memory and Attention

- In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996)
- In parallel circuits, the *BG* may reinforce the *updating* of PFC working memory representations (Alexander et al, 1986; Frank, Loughry & O'Reilly, 2001).
- Dopamine in *PFC* supports robust *maintenance* over time (Lewis & O'Donnell, 2000; Durstewitz & Seemans, 2002)
- Phasic DA bursts thought to occur for task-relevant ("positive") information, reinforcing BG updating signals (O'Reilly & Frank, 2006; Frank & O'Reilly, 2006).

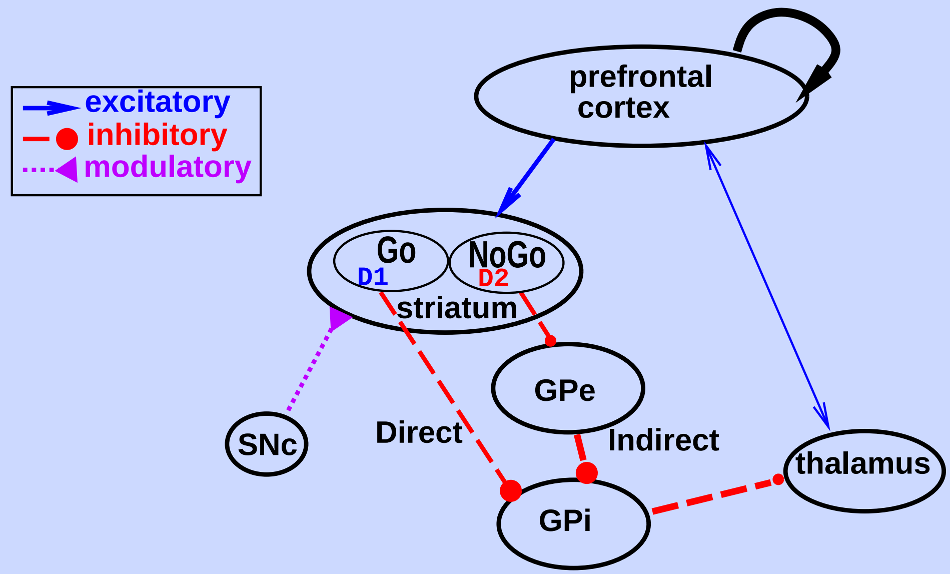
BG Model Extension to Working Memory and Attention

- In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996)
- In parallel circuits, the *BG* may reinforce the *updating* of PFC working memory representations (Alexander et al, 1986; Frank, Loughry & O'Reilly, 2001).
- Dopamine in *PFC* supports robust *maintenance* over time (Lewis & O'Donnell, 2000; Durstewitz & Seemans, 2002)
- Phasic DA bursts thought to occur for task-relevant ("positive") information, reinforcing BG updating signals (O'Reilly & Frank, 2006; Frank & O'Reilly, 2006).
- Time course of DA activity: maintenance in PFC, updating thru BG.

BG Gating of PFC Working Memory / Attention

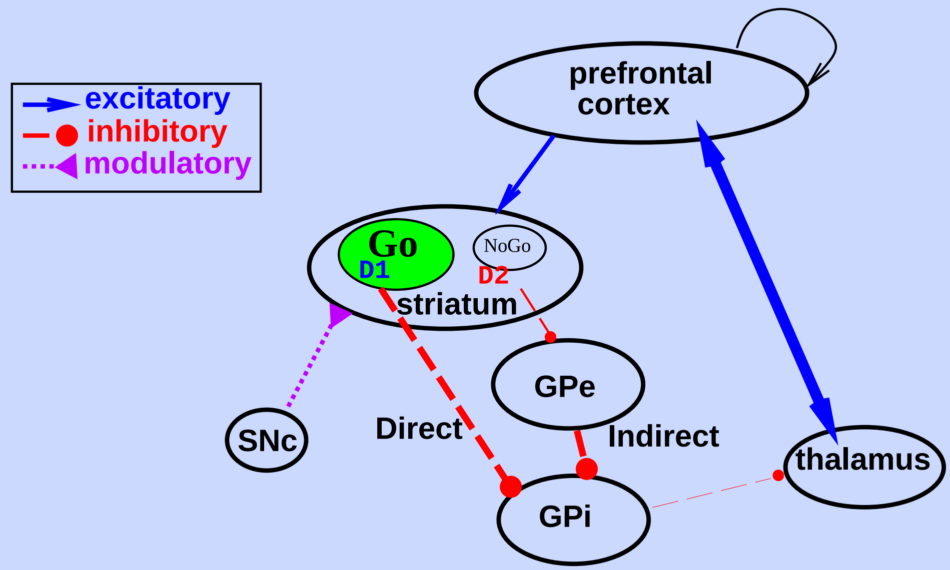


BG Disinhibition of PFC = Gating of Working Memory / Attention



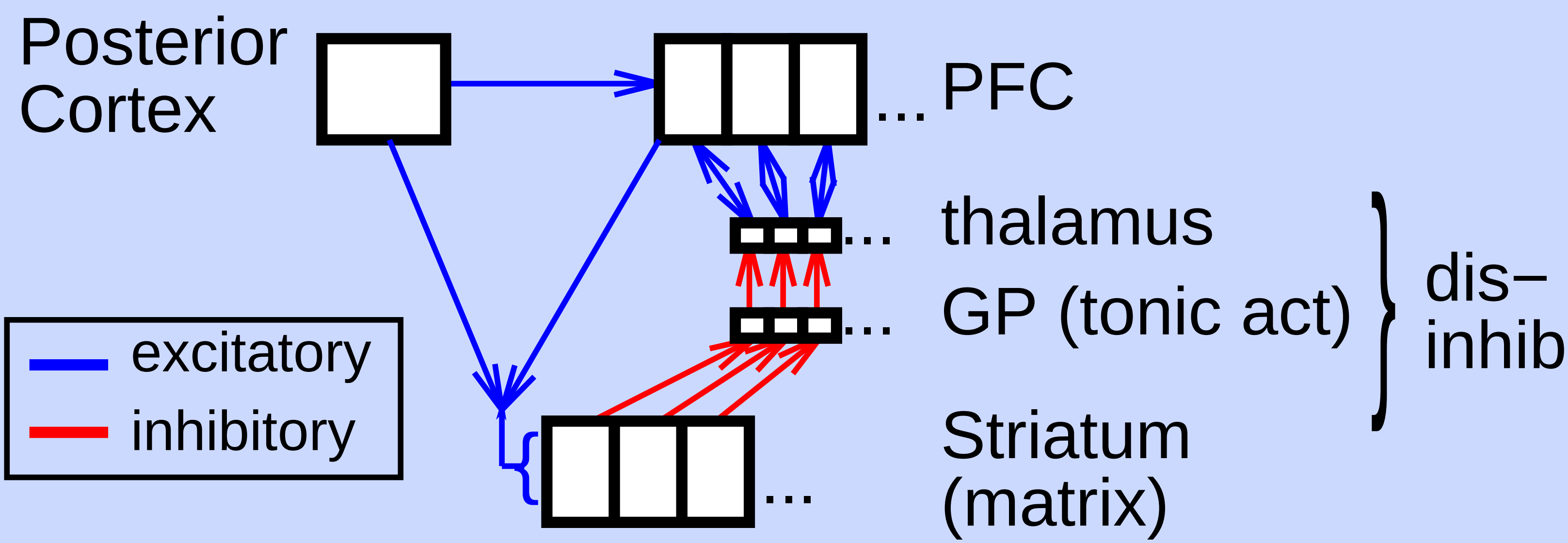
- Base state: Thalamus inhibited = Gate is closed
  PFC robustly maintains prior states.

BG Disinhibition of PFC
= Gating of Working Memory / Attention



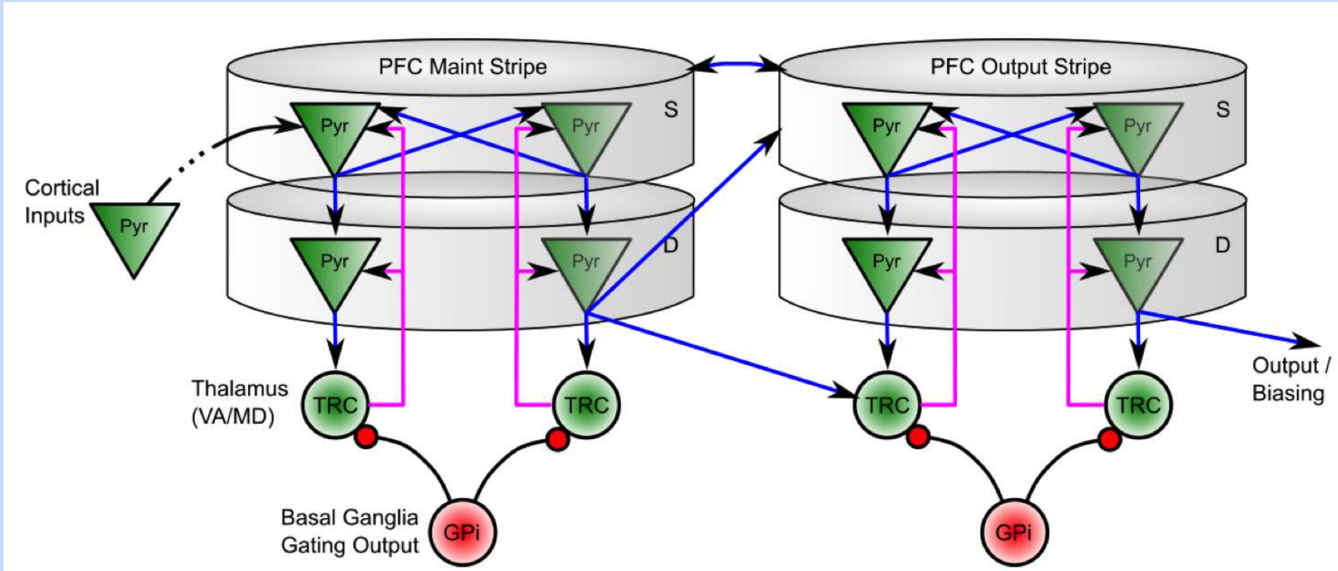
- Base state: Thalamus inhibited = Gate is closed
  PFC robustly maintains prior states.
- Striatum fires: Thalamus disinhibited = Gate opened
  PFC rapidly updated to maintain new information.

Parallel Stripes = Selective Gating



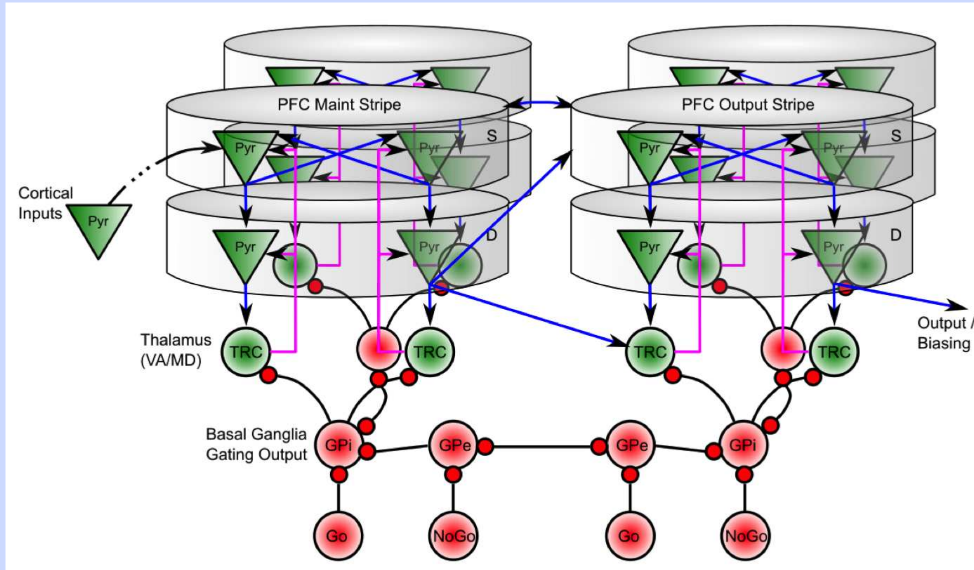
PFC/BG loops form independent stripes = *selective* gating.

BG gates flow: superficial → Deep PFC



Maintenance via Thalamocortical loops, BG disinhibits
Superficial reflects inputs and maintenance
Separate Maintenance vs. Output PFC / BG *stripes*

Multiple stripes, competition



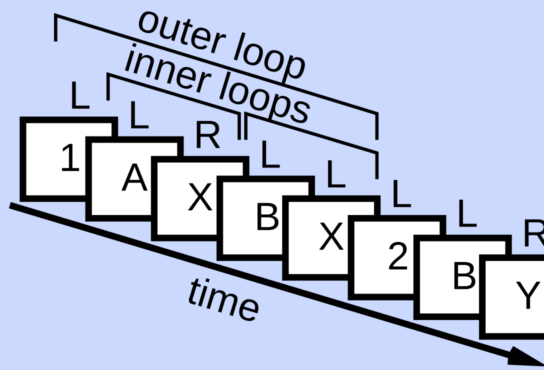
Competition in striatum/GP between Maint vs Output and diff stripes

Example Task: Specific Working Memory Demands

Target (R):

1 = A-X

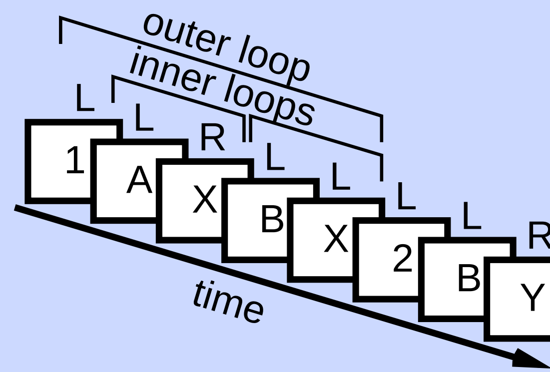
2 = B-Y



Must maintain *outer loop* (1,2) while updating *inner loop* (A,X..).

Example Task: Specific Working Memory Demands

Target (R):
1 = A-X
2 = B-Y

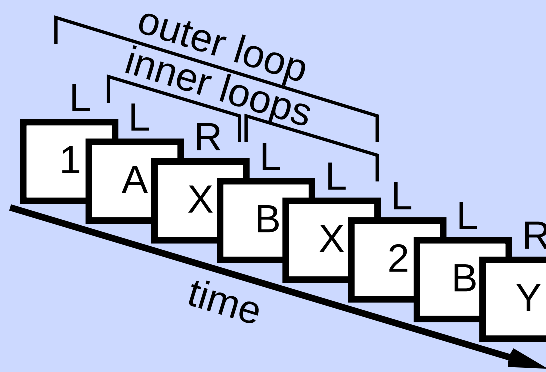


Must maintain *outer loop* (1,2) while updating *inner loop* (A,X..).

- **Robust** maintenance (over intervening stimuli; outer loop).

Example Task: Specific Working Memory Demands

Target (R):
1 = A-X
2 = B-Y

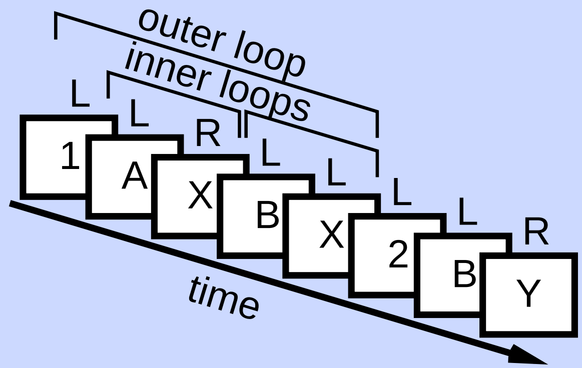


Must maintain *outer loop* (1,2) while updating *inner loop* (A,X..).

- **Robust** maintenance (over intervening stimuli; outer loop).
- **Rapid** and **selective** updating (keep 1,2 while updating A,X..).

Example Task: Specific Working Memory Demands

Target (R):
1 = A-X
2 = B-Y



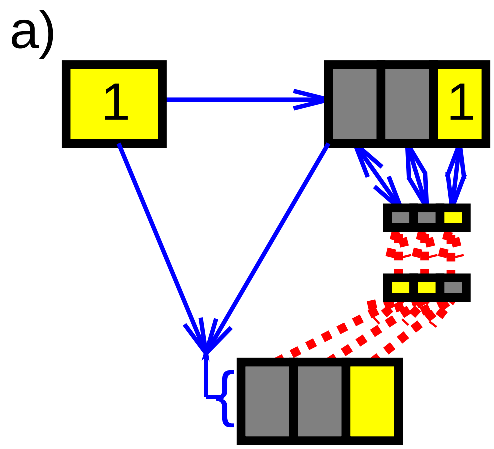
Must maintain *outer loop* (1,2) while updating *inner loop* (A,X..).

- **Robust** maintenance (over intervening stimuli; outer loop).
- **Rapid** and **selective** updating (keep 1,2 while updating A,X..).

Many real-world examples: language, planning...

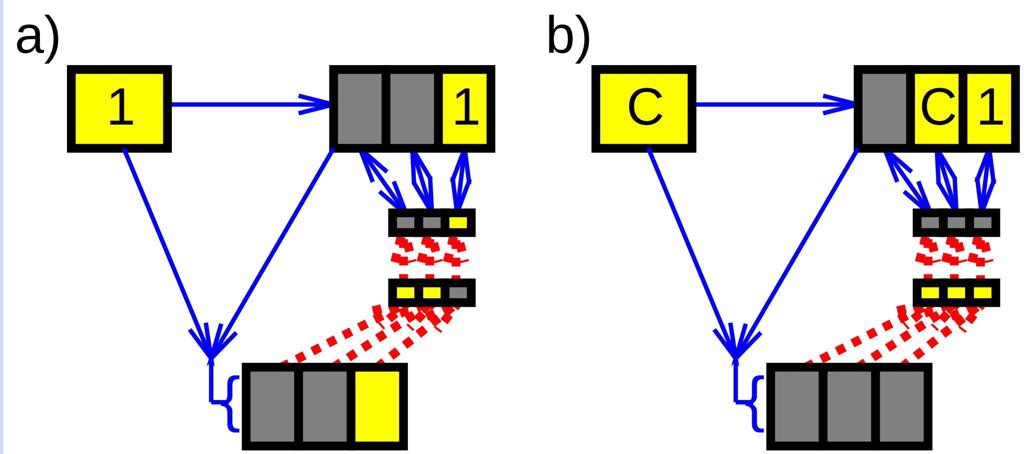
PFC/BG Model of 1-2-AX

(Frank, Loughry & O'Reilly, 2001)



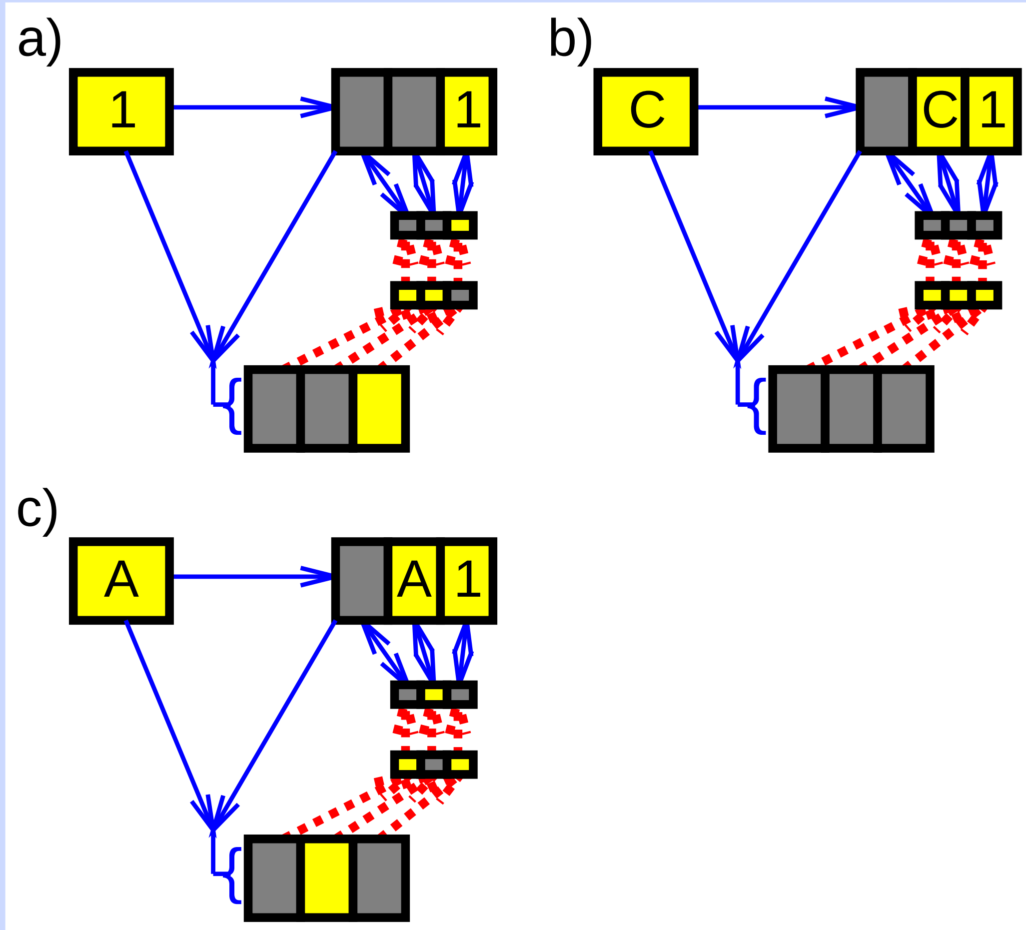
FC/BG Model of 1-2-AX

(Frank, Loughry & O'Reilly, 2001)



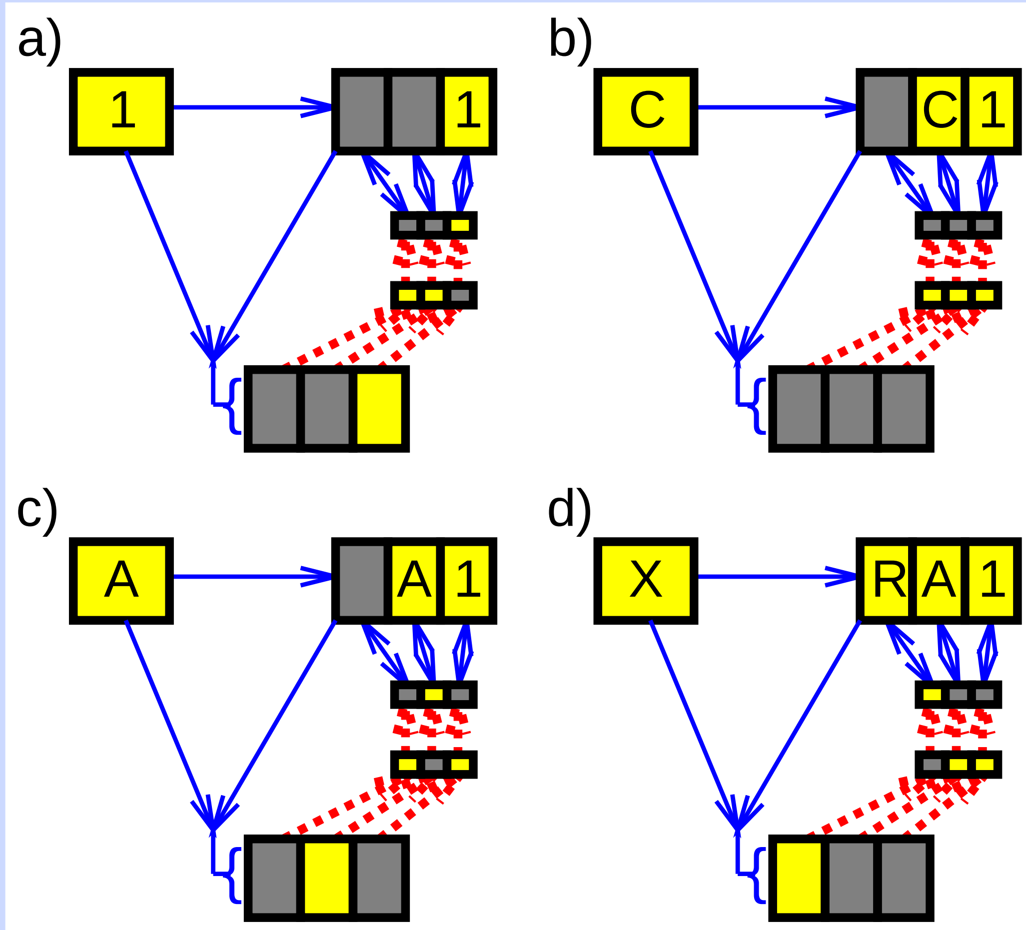
FC/BG Model of 1-2-AX

(Frank, Loughry & O'Reilly, 2001)



FC/BG Model of 1-2-AX

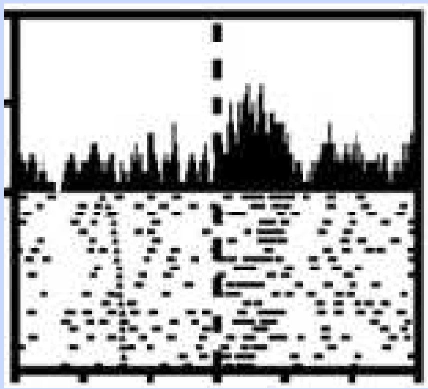
(Frank, Loughry & O'Reilly, 2001)

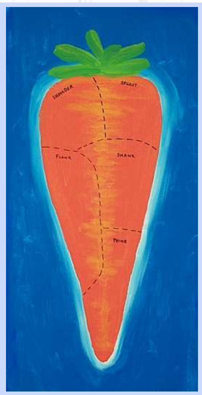


But How do BG “Know” What to Update/Ignore?

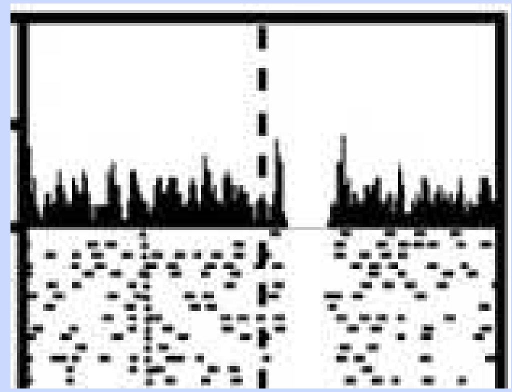
But How do BG “Know” What to Update/Ignore?

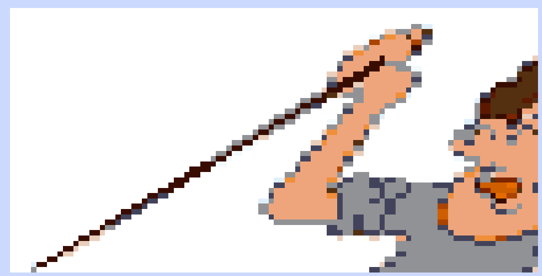
Correct:





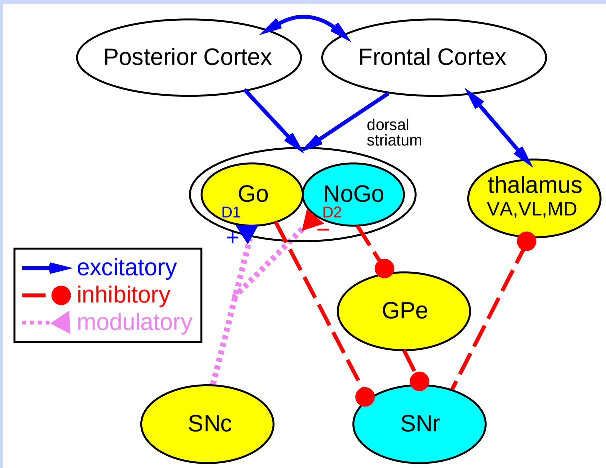
Incorrect:



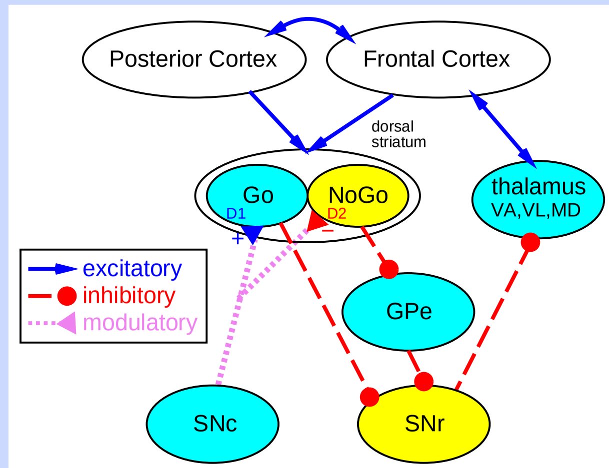


DA effects on BG Updating Of PFC

DA Burst: Increased Go



DA Dip: Increased NoGo



- DA bursts/dips modulate Go vs. NoGo firing and learning.

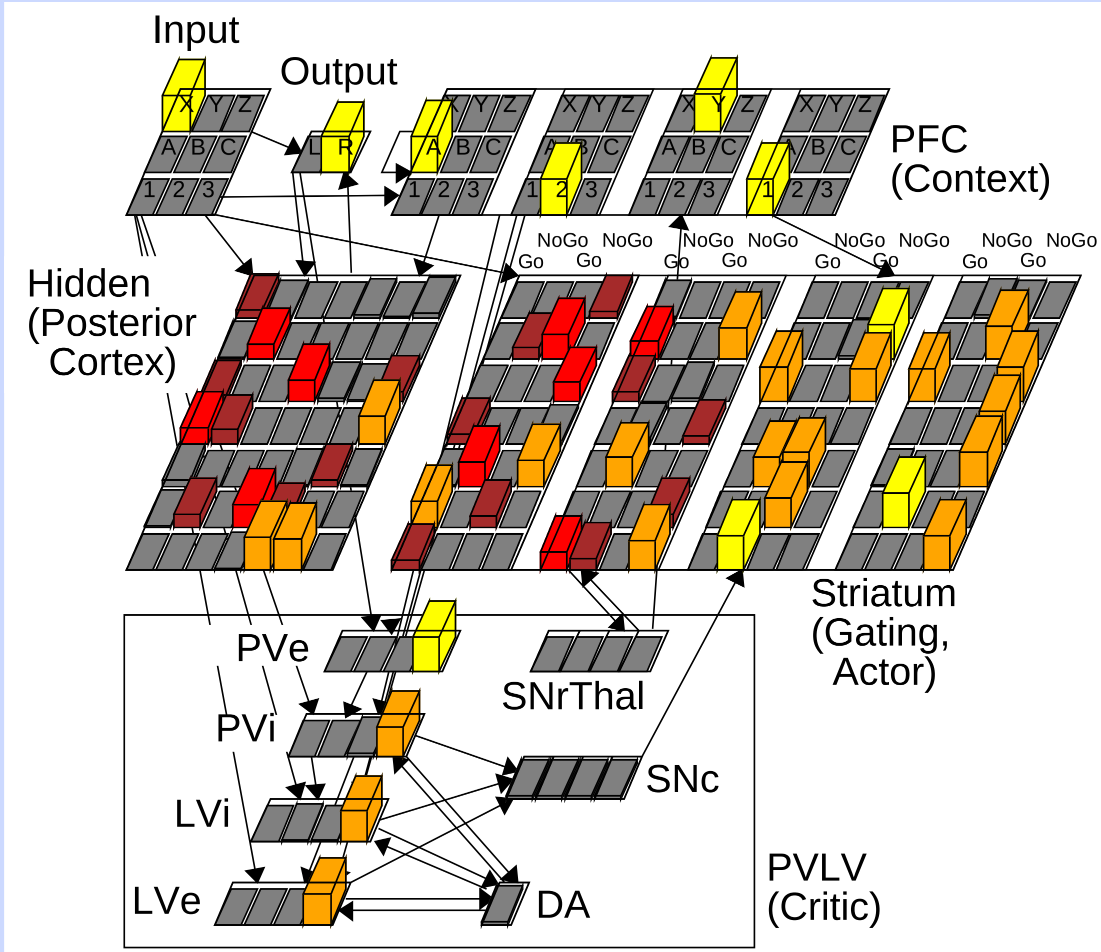
- Burst = Excitatory D1 on Go

  Dip = Release of D2 inhibition on NoGo

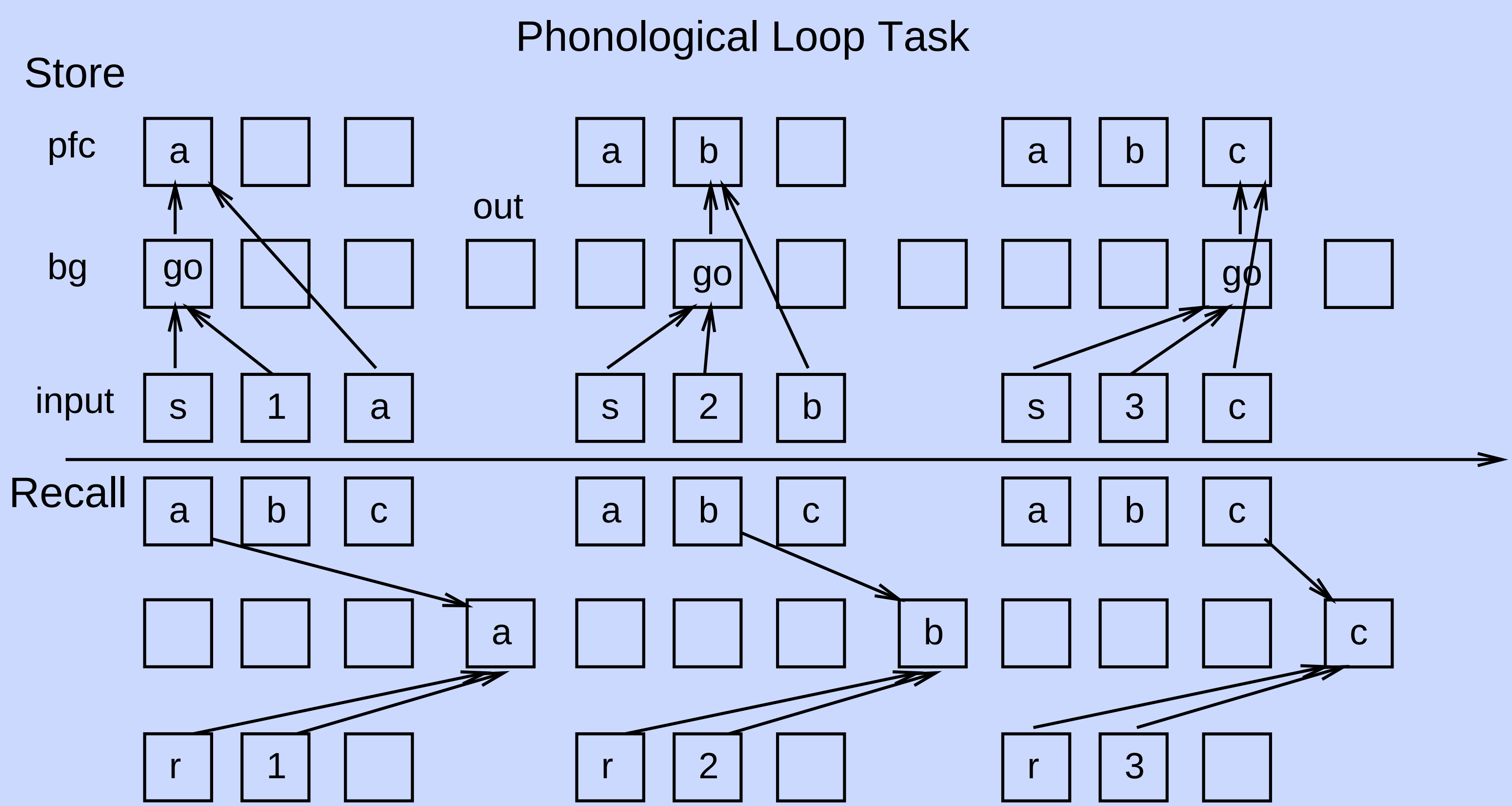
  Same mechanism as in basic motor circuitry! (Frank, 05; Gerfen, 00)

PFC/BG Model, Learns 12-AX and other WM tasks

(O'Reilly & Frank, 2006)

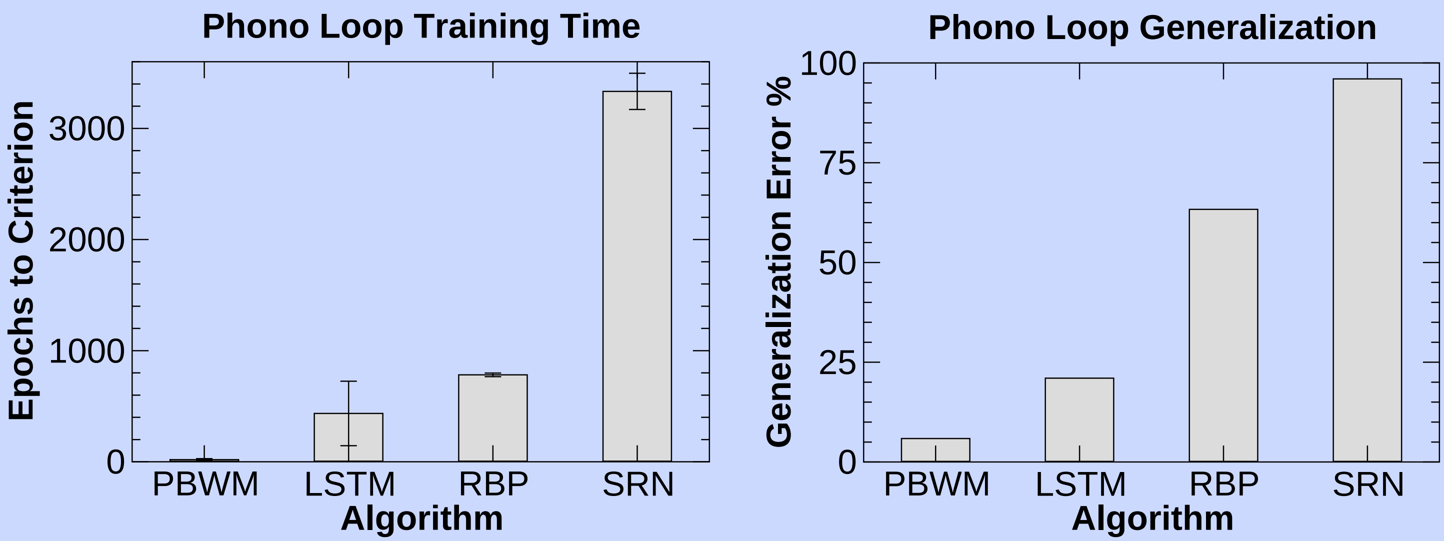


Phonological Loop Task



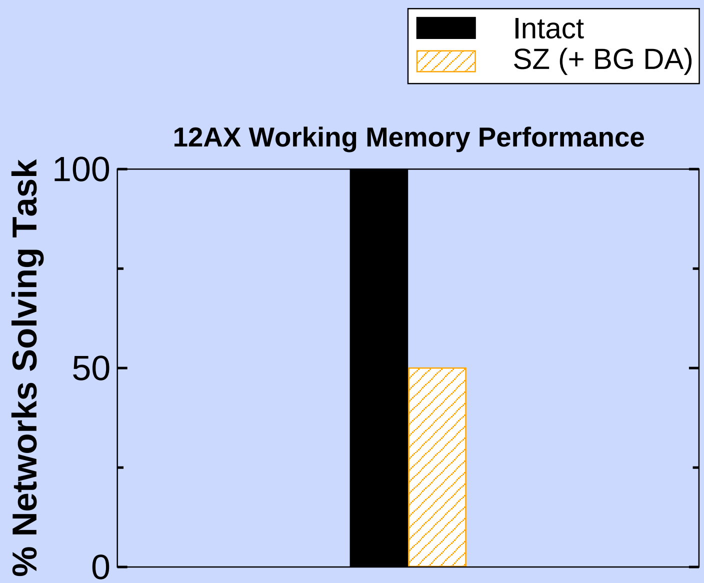
Phonological Loop Task

(O'Reilly & Frank, 2006)

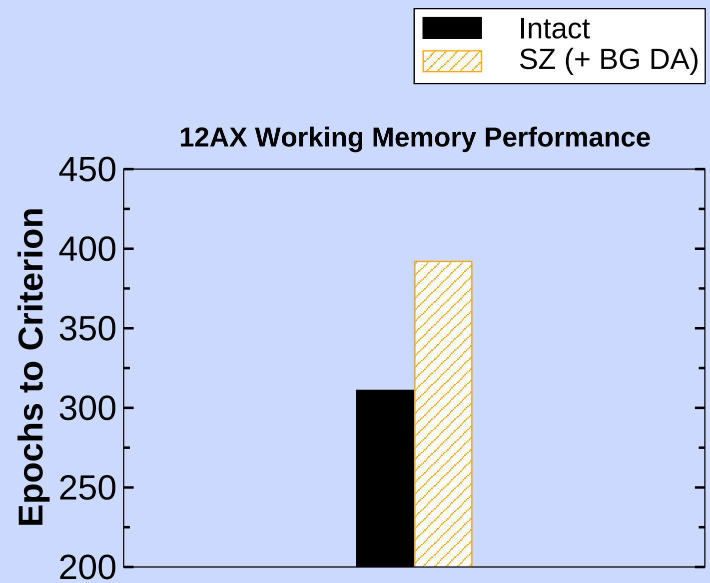


PBWM = Prefrontal Basal-ganglia Working Memory; LSTM = Long Short Term Memory (Schmidhuber et al.); RBP = Recurrent BackPropagation; SRN = Simple Recurrent Network.

Simulating BG DA increases in SZ: Effects on 12AX Working Memory Performance

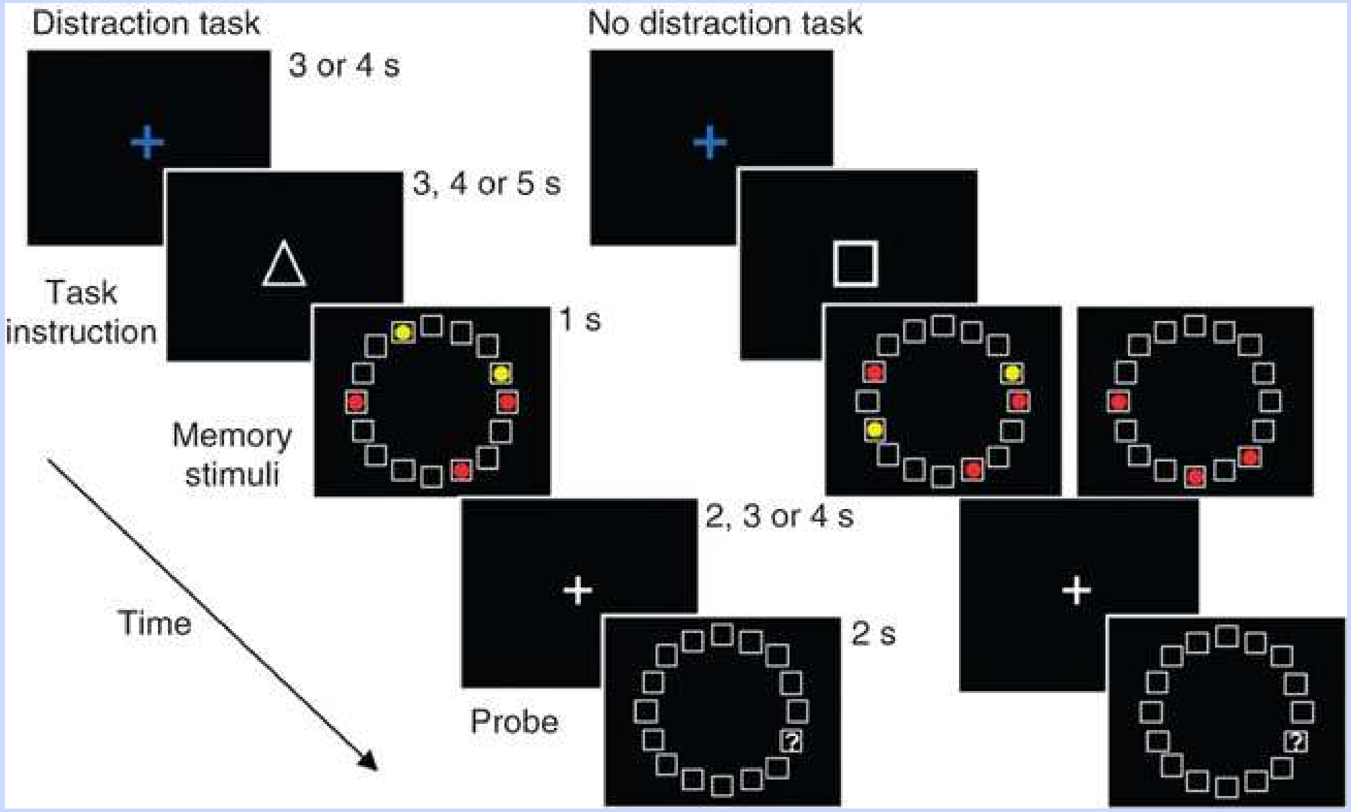


Accuracy



Training Needed to Learn

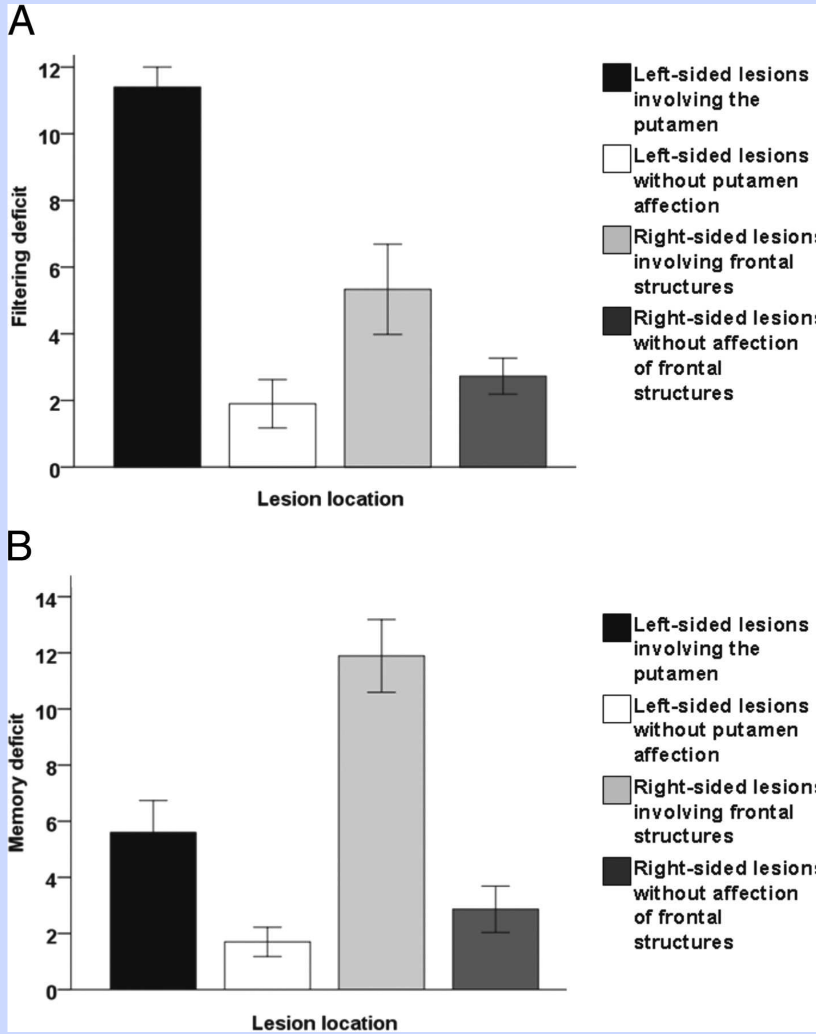
Evidence for BG gating of PFC



Instruction cue signals whether subsequent yellow information is distracting (and should not be stored in working memory).

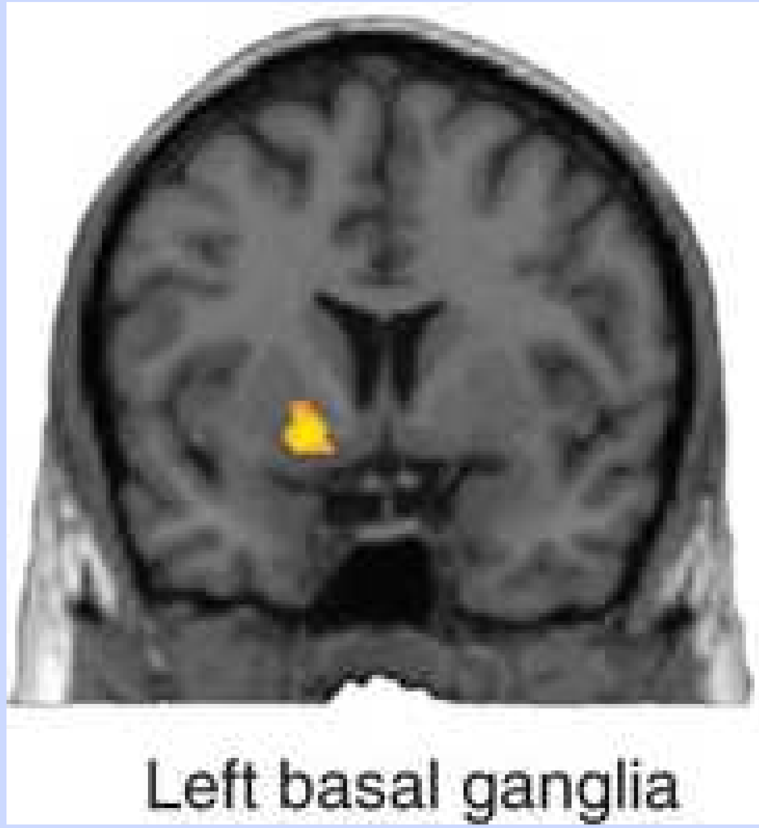
McNab & Klingberg (2008), *Nature Neuroscience*

Evidence for BG gating of PFC: Lesion patients



Baier et al (2010), *Journal of Neuroscience*

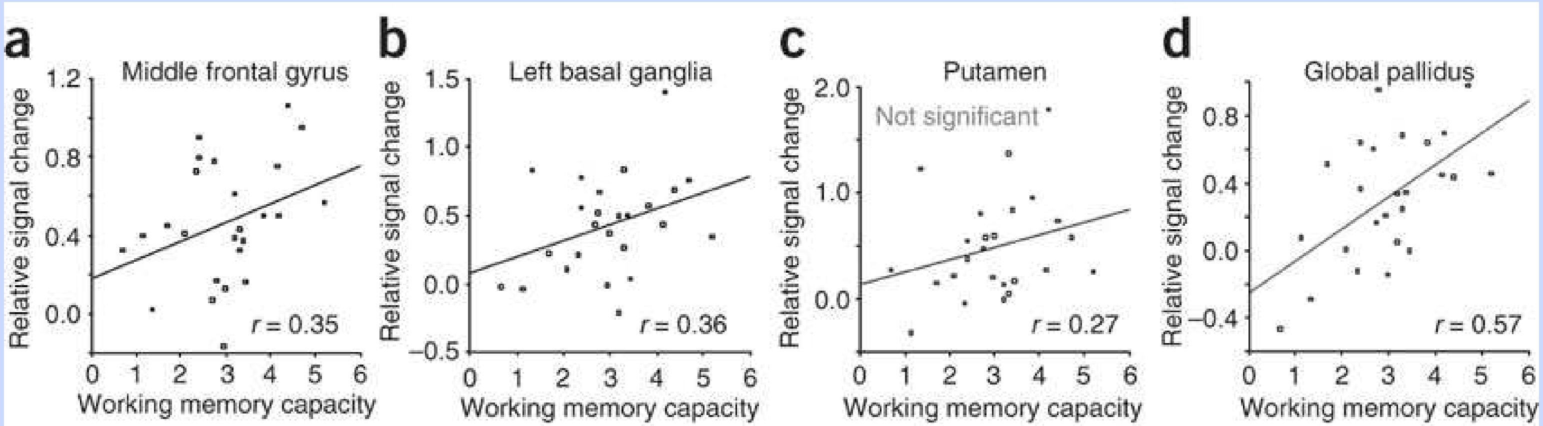
Evidence for BG gating of PFC: Neuroimaging



Distractor versus no-distractor contrast: NoGo!

McNab & Klingberg (2008), *Nature Neuroscience*

Both BG and PFC act correlate with WM capacity

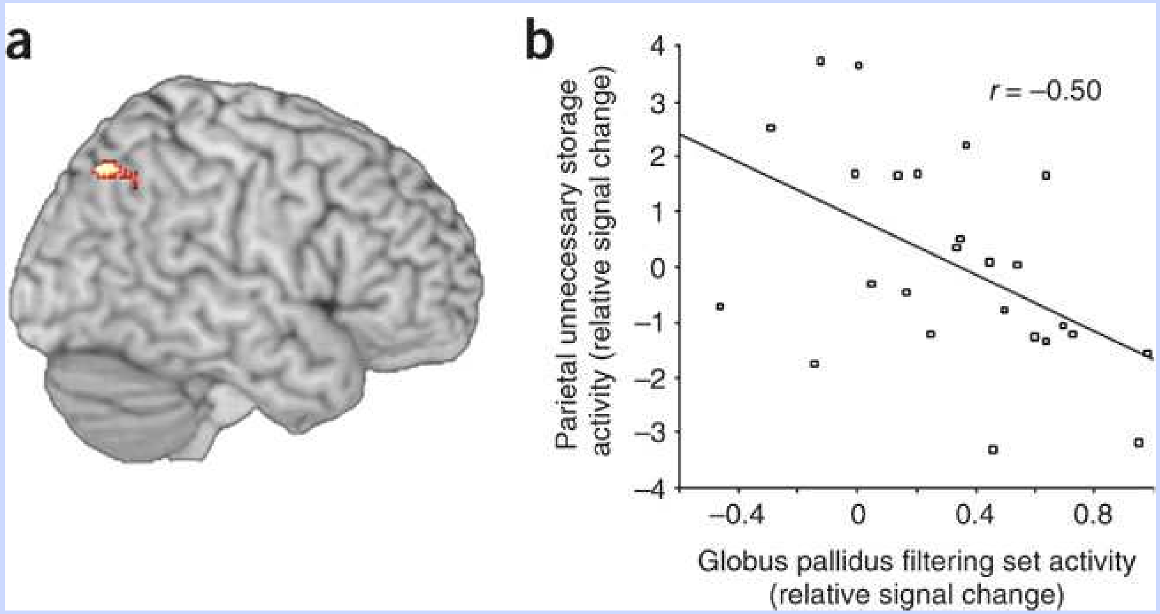


Better WM function correlates with PFC (middle frontal gyrus) and BG activity in Dist vs NoDist contrast.

Globus pallidus = output of BG; stronger → more NoGo

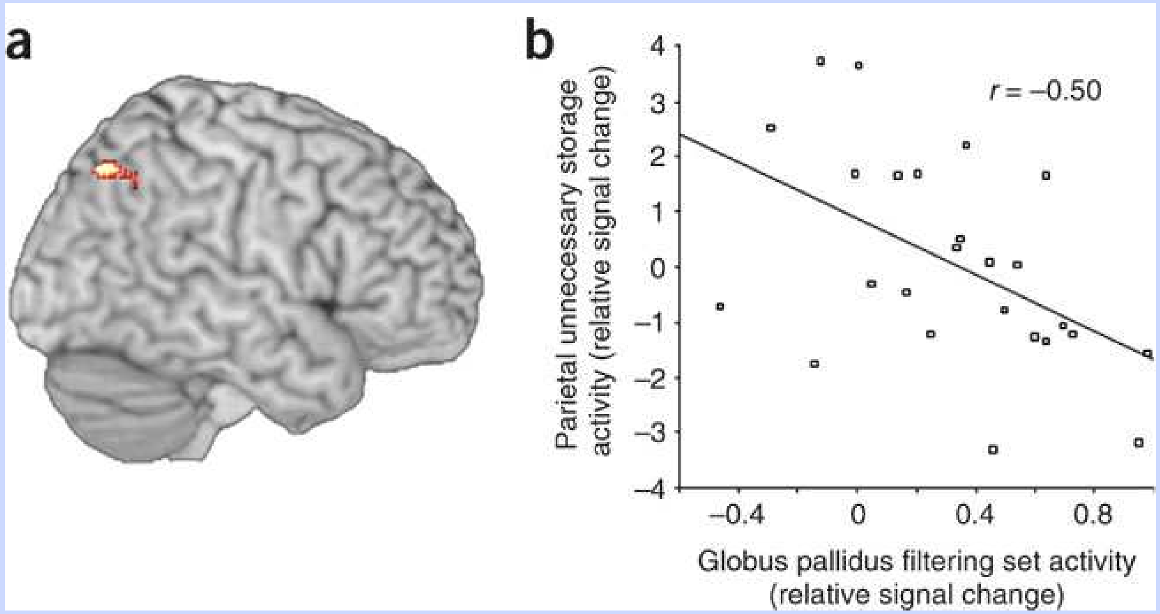
McNab & Klingberg (2008), *Nature Neuroscience*

Evidence for BG gating of PFC: Neuroimaging



More GP activity ⇒ Less “unnecessary storage” of distracting info in parietal cortex (spatial storage; this parietal region is sensitive to memory load).

Evidence for BG gating of PFC: Neuroimaging



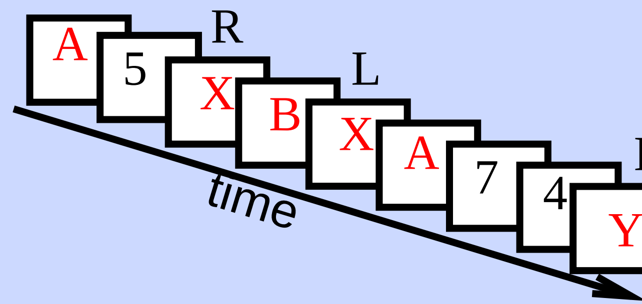
More GP activity $\Rightarrow$ Less "unnecessary storage" of distracting info in parietal cortex (spatial storage; this parietal region is sensitive to memory load).

No such correlation in PFC, suggesting that BG = filter (gate).

McNab & Klingberg (2008), *Nature Neuroscience*

Testing the Model: AX-CPT Task

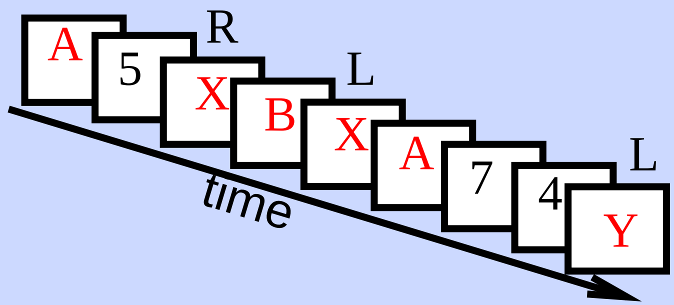
Cohen et al (1997); Barch et al (2001); Frank & O'Reilly, 2006



- target sequence is A - X; press R button.
- non-target sequences A - Y, B - X, B -Y; press L button.
- ignore all distractors (5,7)

Testing the Model: AX-CPT Task

Cohen et al (1997); Barch et al (2001); Frank & O'Reilly, 2006



- target sequence is A - X; press R button.
- non-target sequences A - Y, B - X, B -Y; press L button.
- ignore all distractors (5,7)
- Tests gating and maintenance components of working memory / attention.

B

x

A

×

A

5

x

A

7

4

Υ

A

7

5

x

A

5

7

Y

A

×

B

4

5

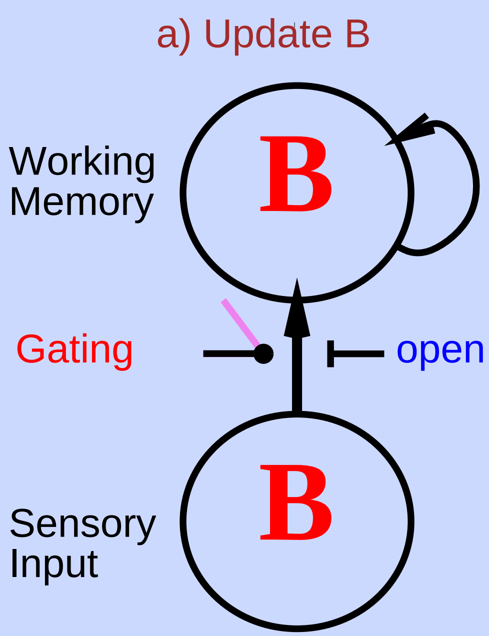
4

x

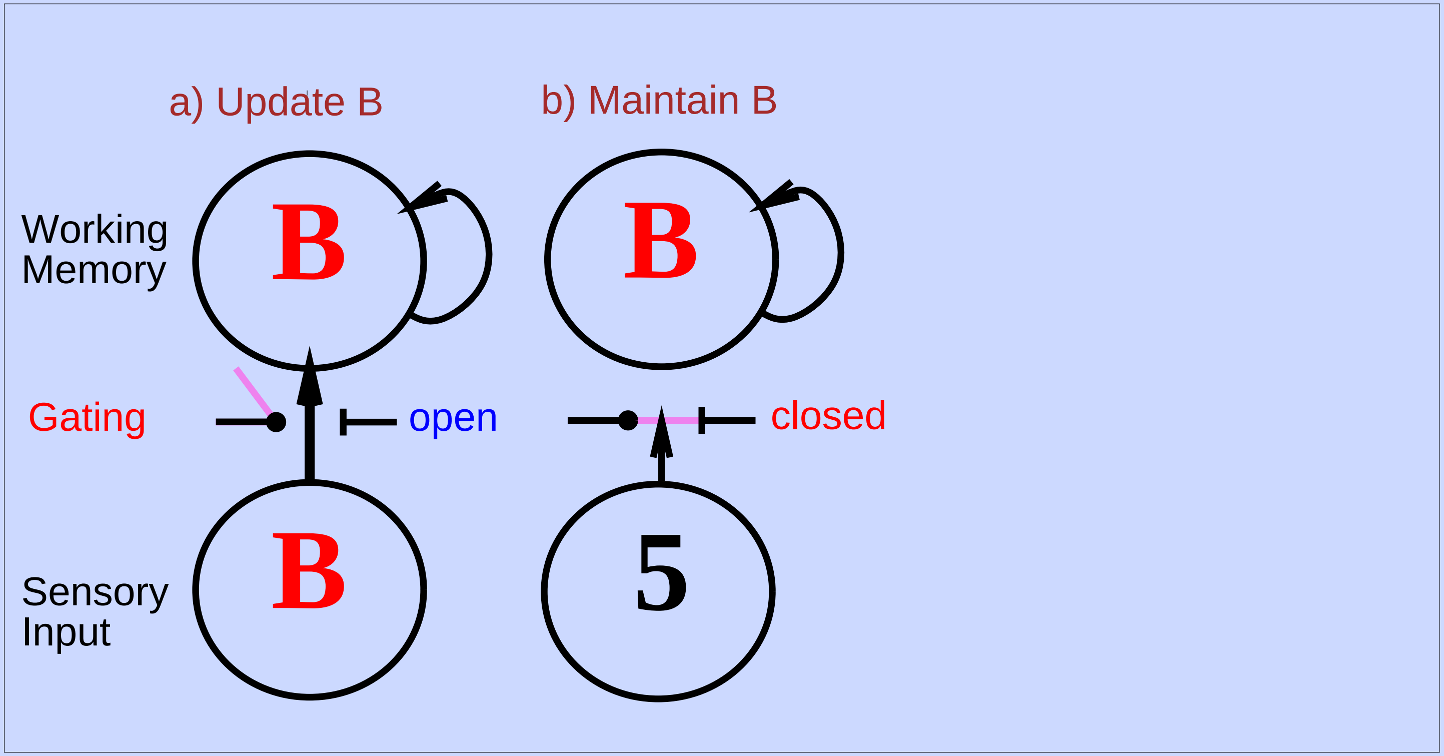
Working Memory Demands: AX-CPT Task

- A-X target sequence on 70% of trials.
- B-Y is control condition, not dependent on working memory.
- **Key conditions: B-X, A-Y**

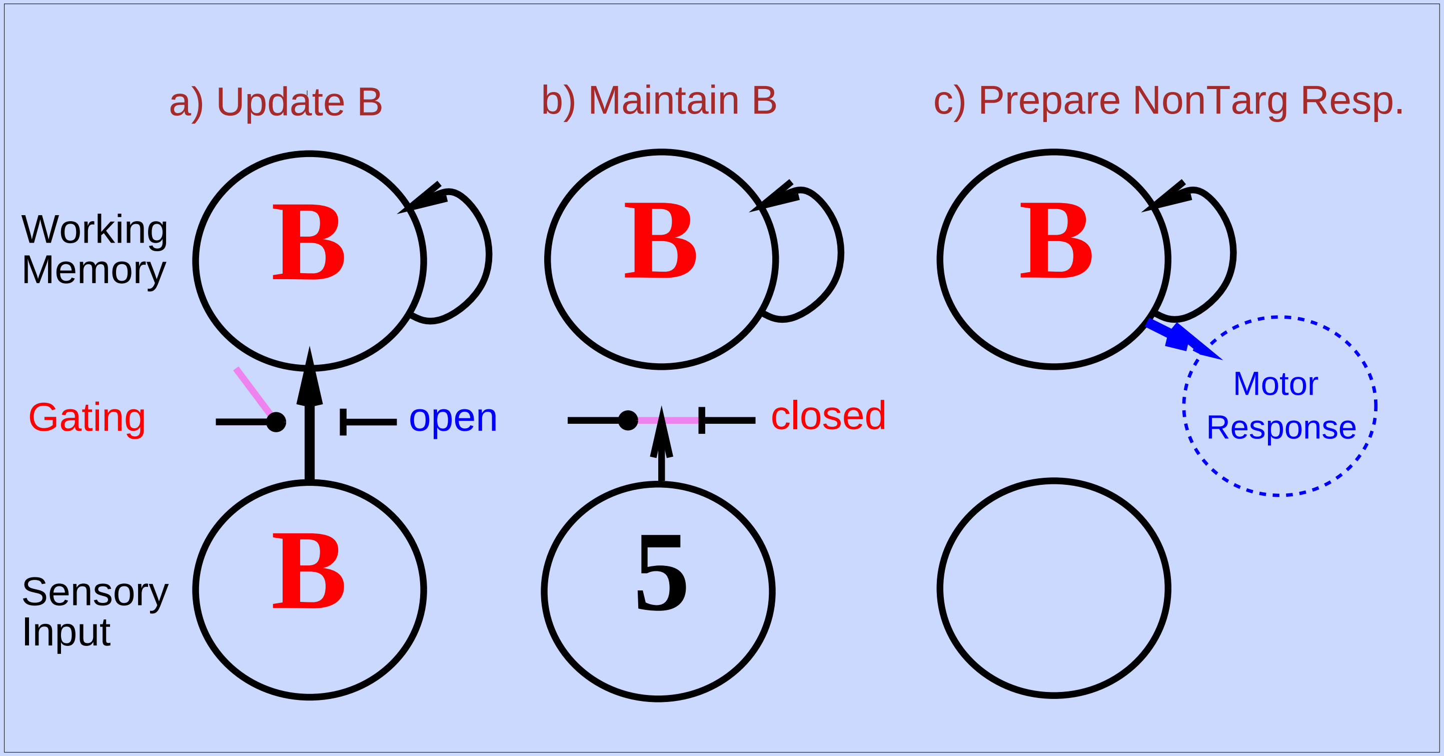
B-X: ↑ Working memory, ↑ performance



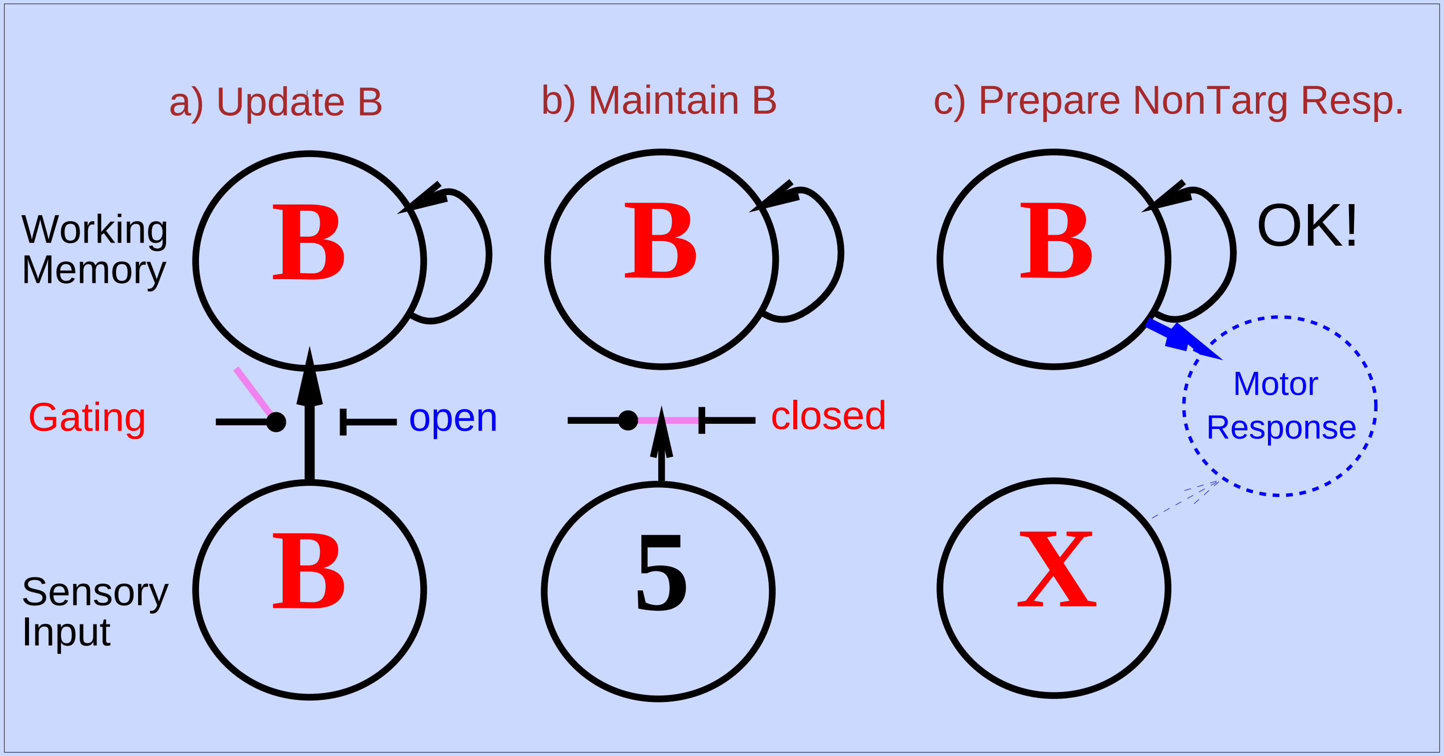
B-X: $\uparrow$ Working memory, $\uparrow$ performance



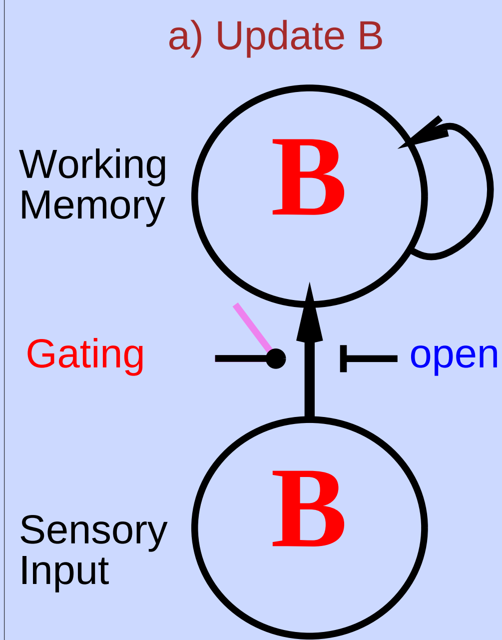
B-X: ↑ Working memory, ↑ performance



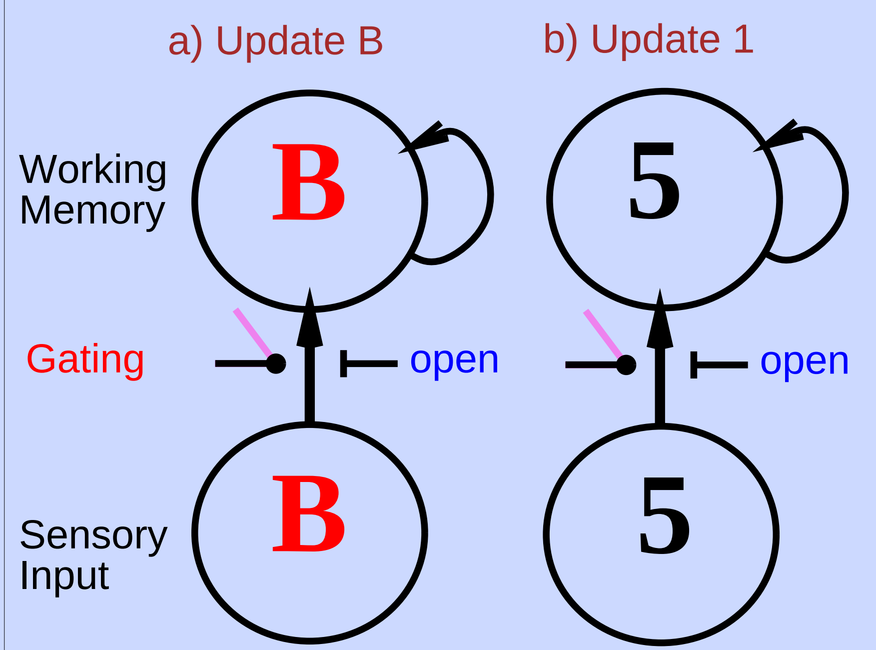
B-X: ↑ Working memory, ↑ performance



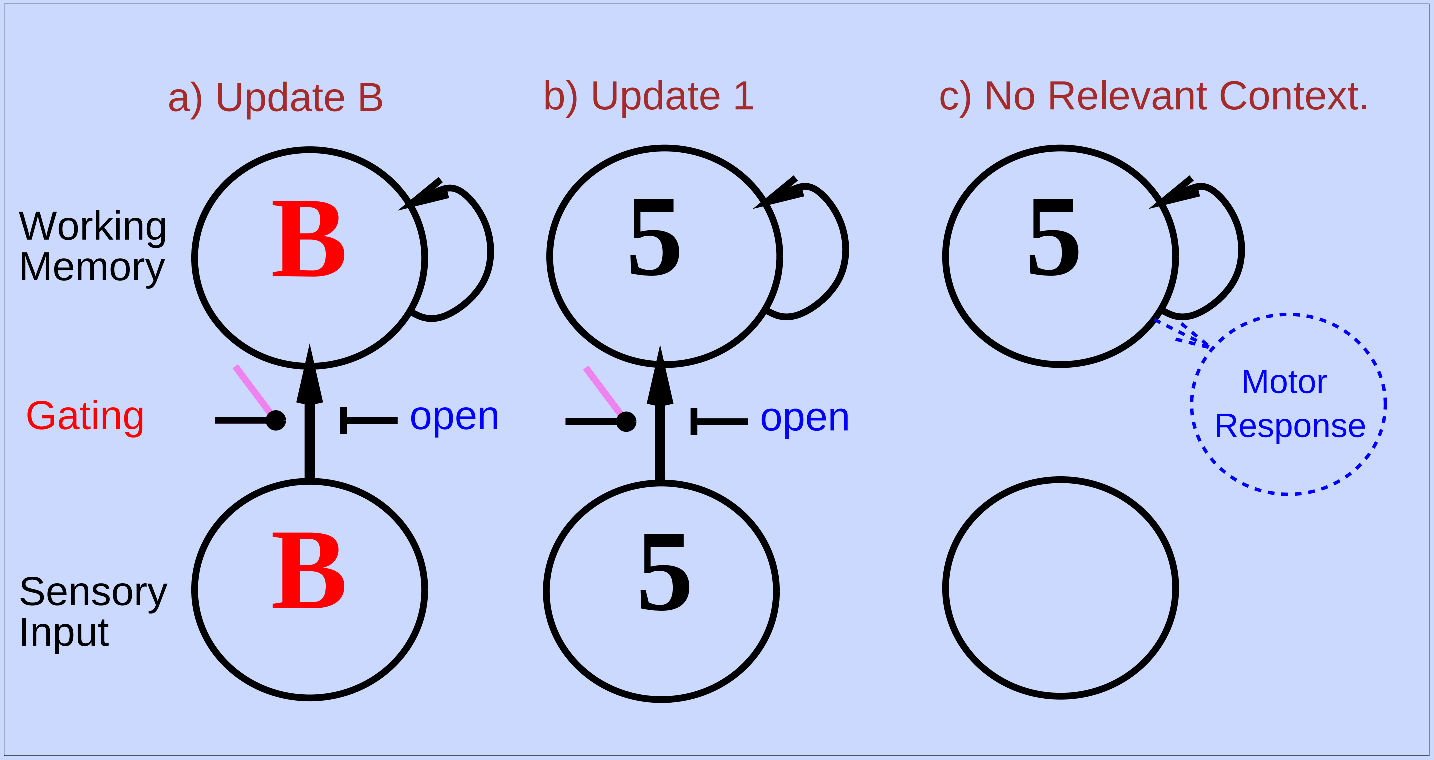
B-X: ↓ Working memory, ↓ performance



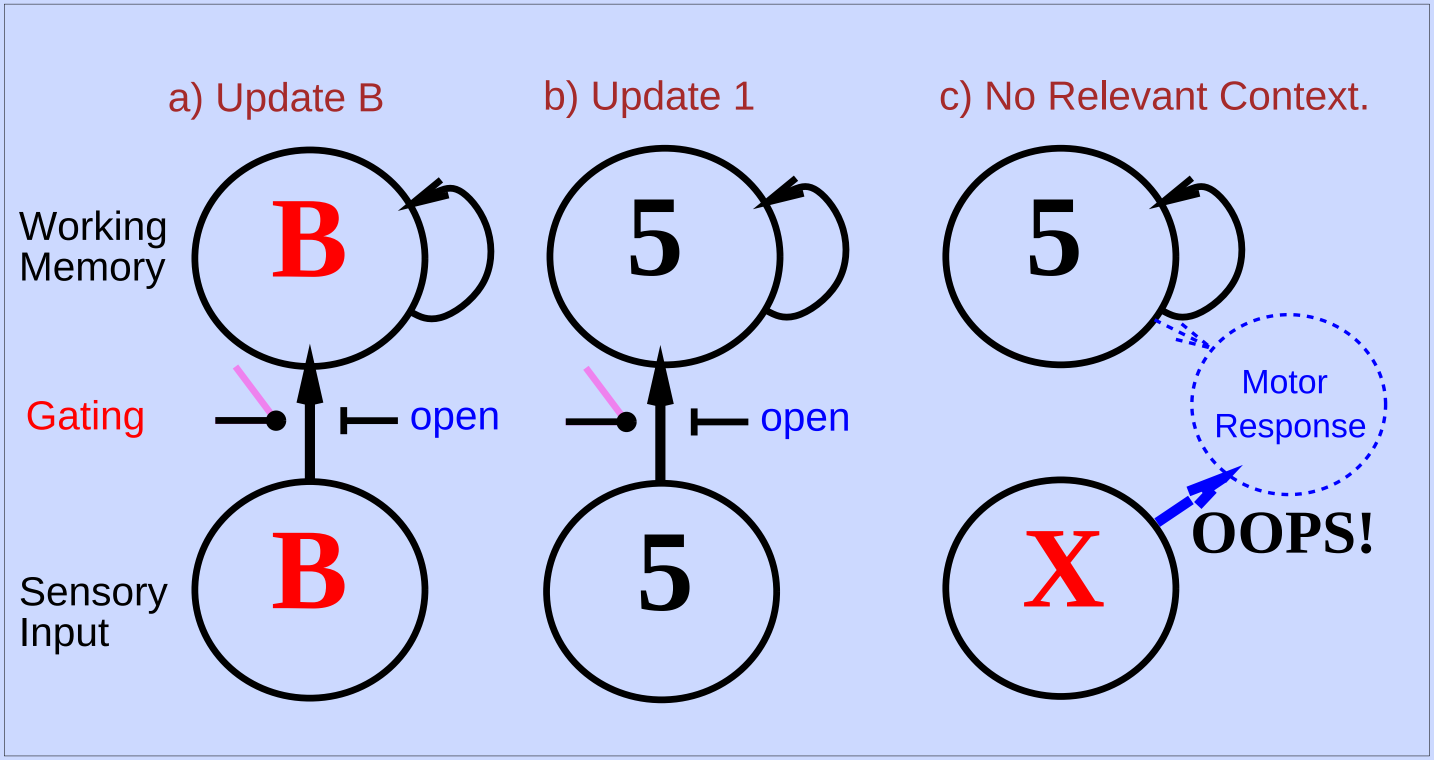
B-X: $\downarrow$ Working memory, $\downarrow$ performance



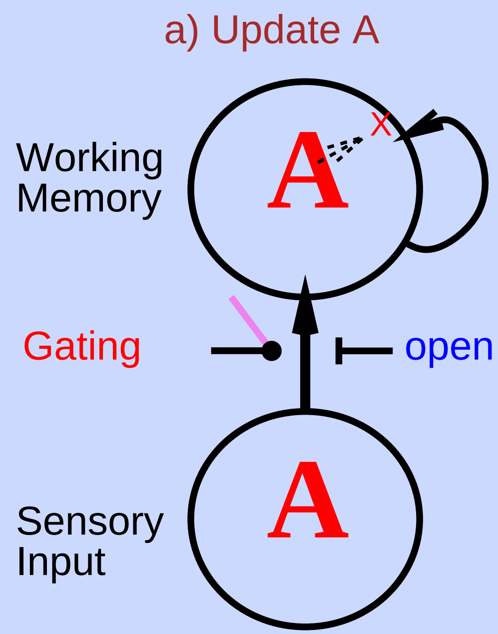
B-X: ↓ Working memory, ↓ performance



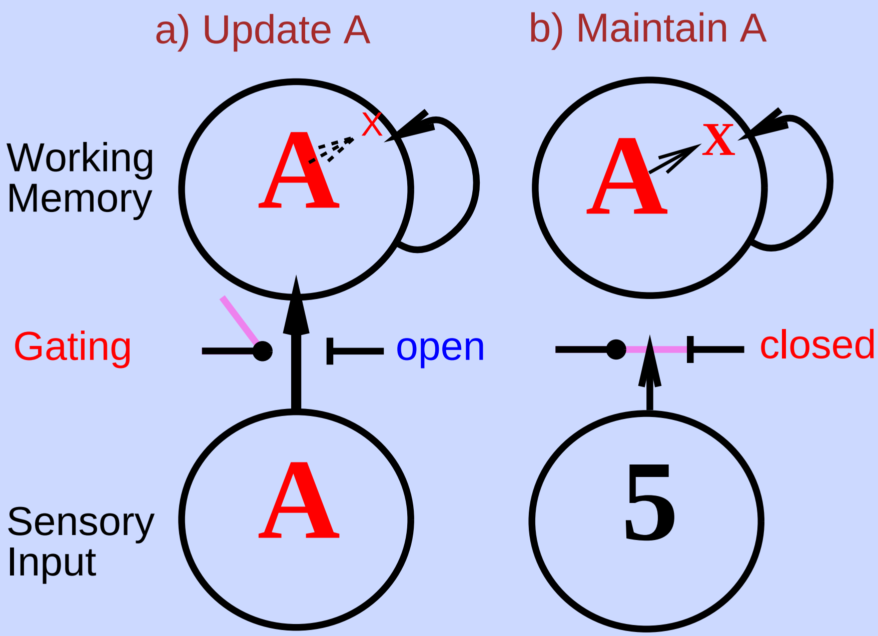
B-X: $\downarrow$ Working memory, $\downarrow$ performance



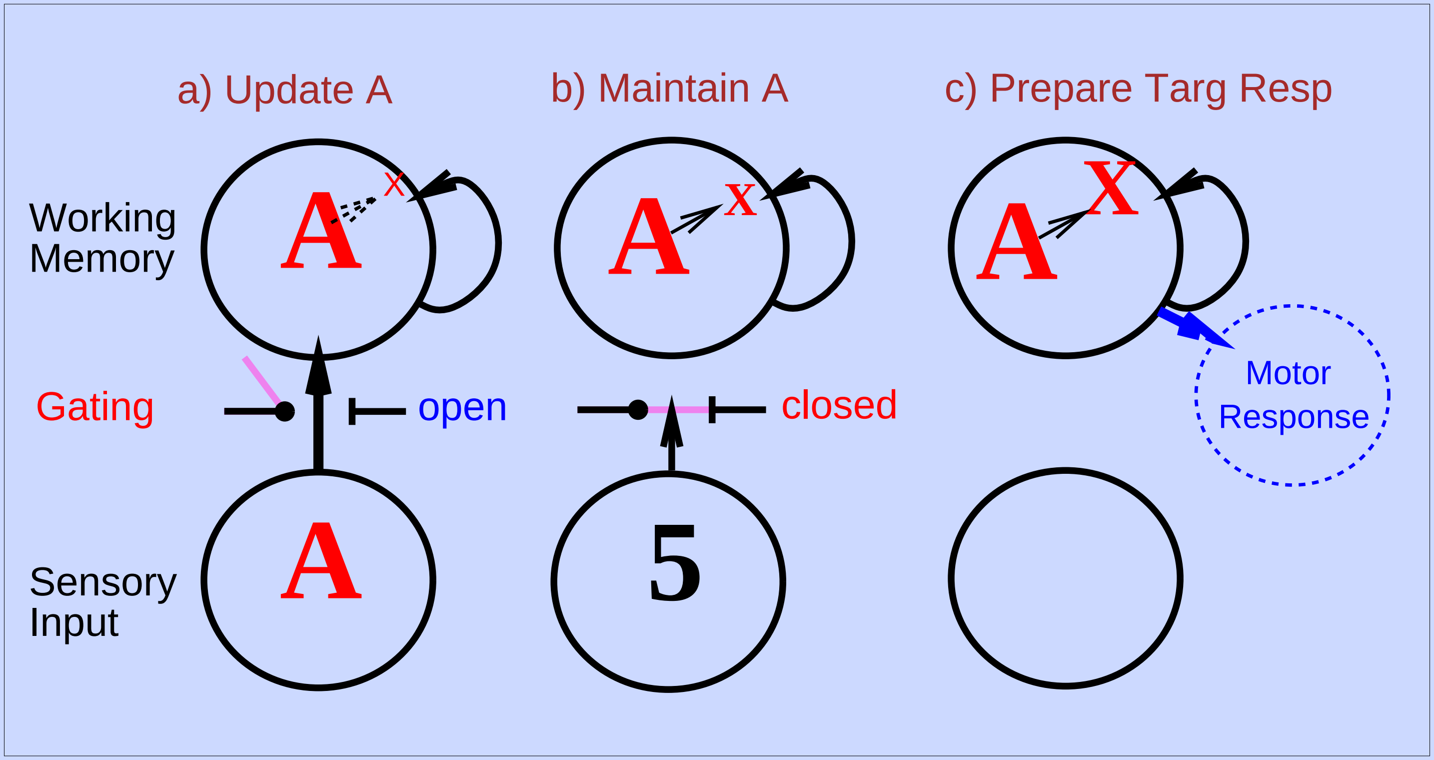
A-Y: $\uparrow$ Working memory, $\downarrow$ performance



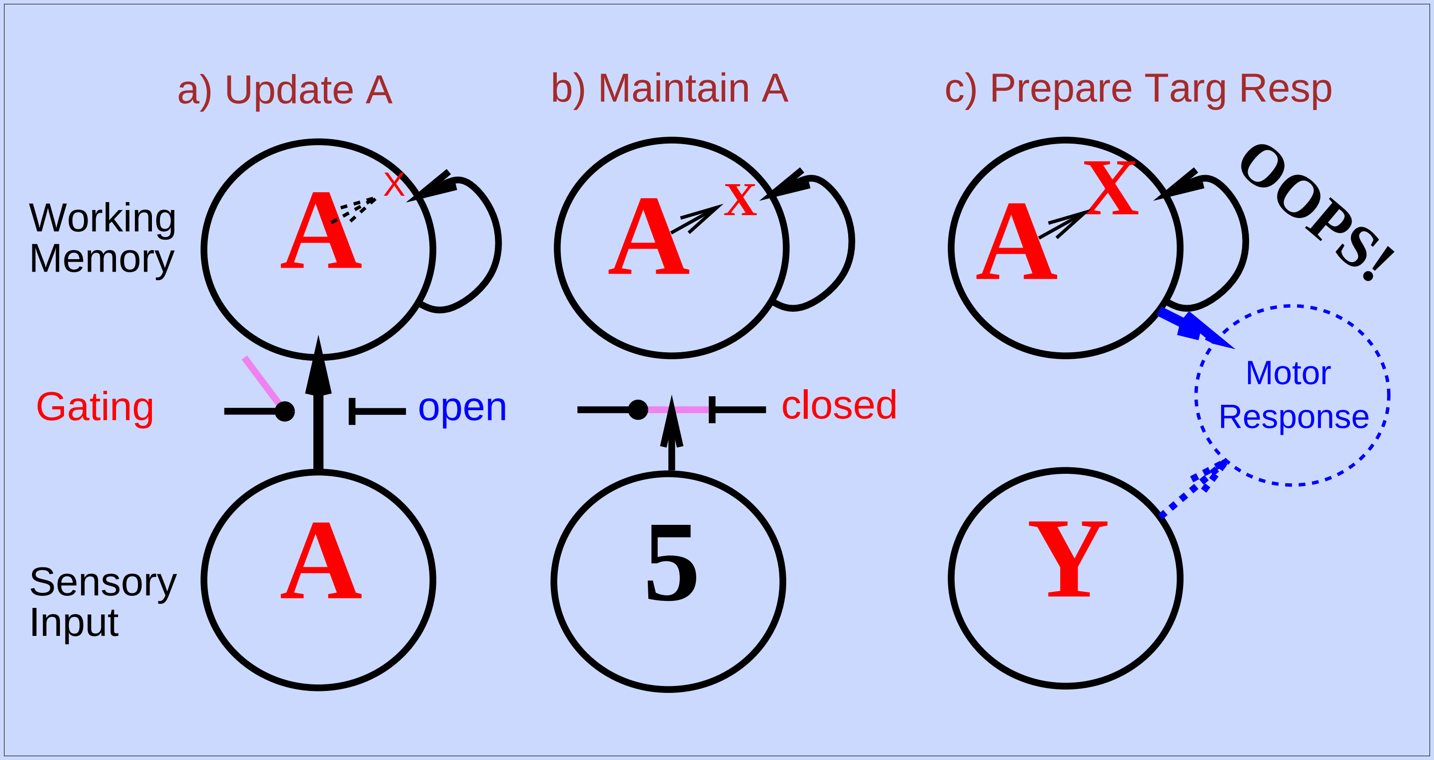
A-Y: ↑ Working memory, ↓ performance



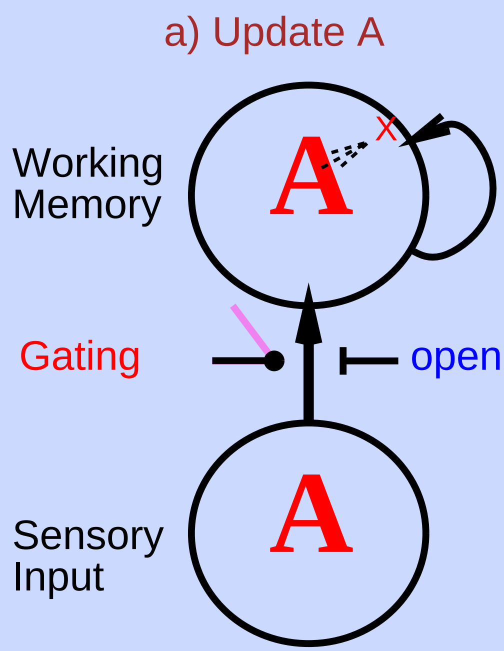
A-Y: $\uparrow$ Working memory, $\downarrow$ performance



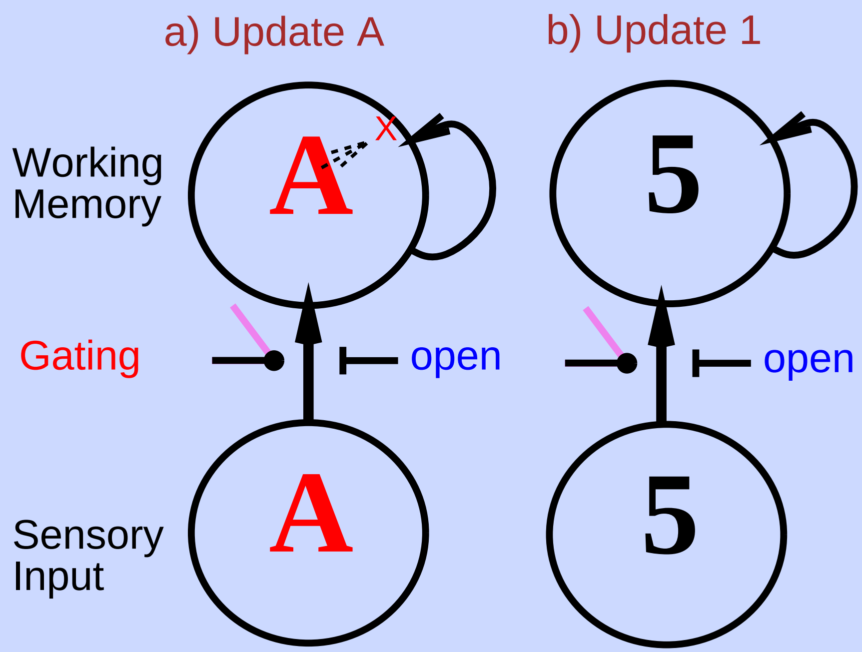
A-Y: ↑ Working memory, ↓ performance



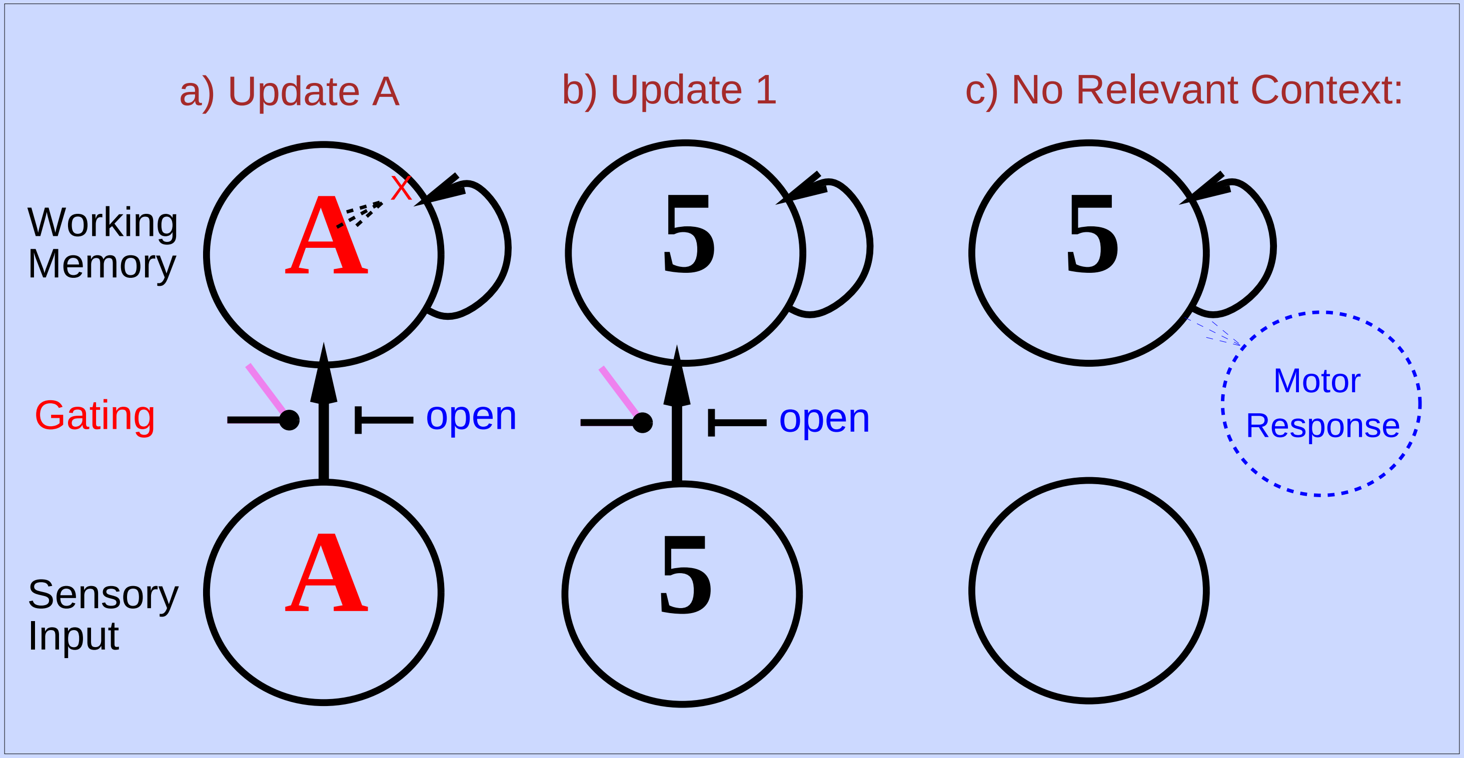
A-Y: $\downarrow$ Working memory, $\uparrow$ performance



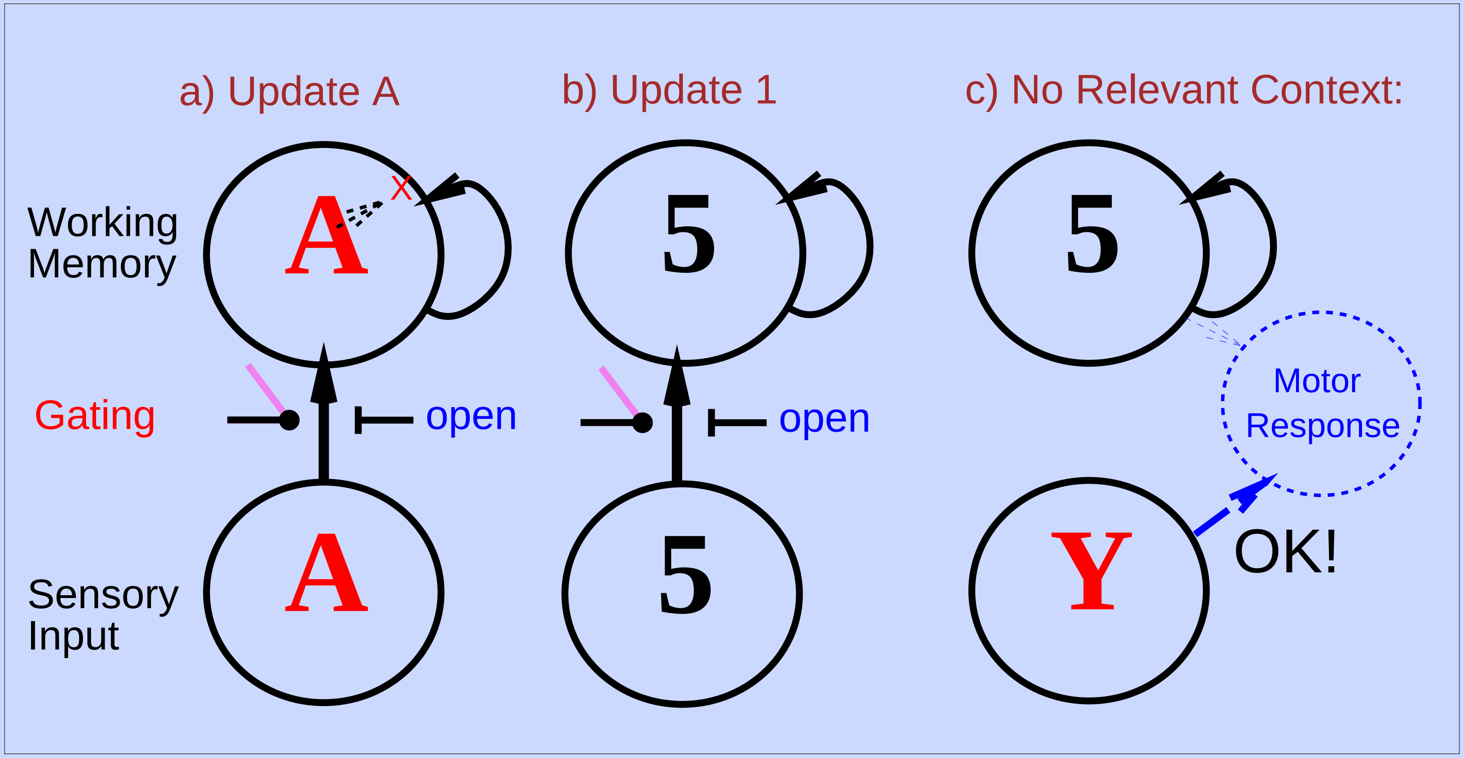
A-Y: ↓ Working memory, ↑ performance



A-Y: ↓ Working memory, ↑ performance



A-Y: $\downarrow$ Working memory, $\uparrow$ performance

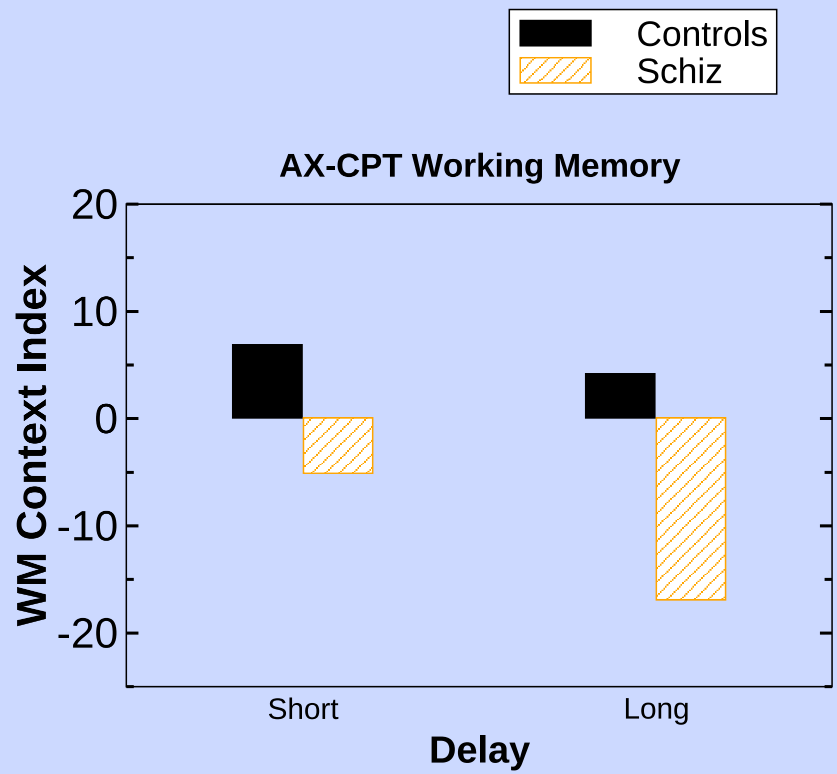


Working Memory Demands: AX-CPT Task

- **More active maintenance of task-relevant info**
  $\rightarrow$ **more A-Y false alarms, less B-X false alarms**

SZ: Impaired Working Memory

Barch et al (2001)



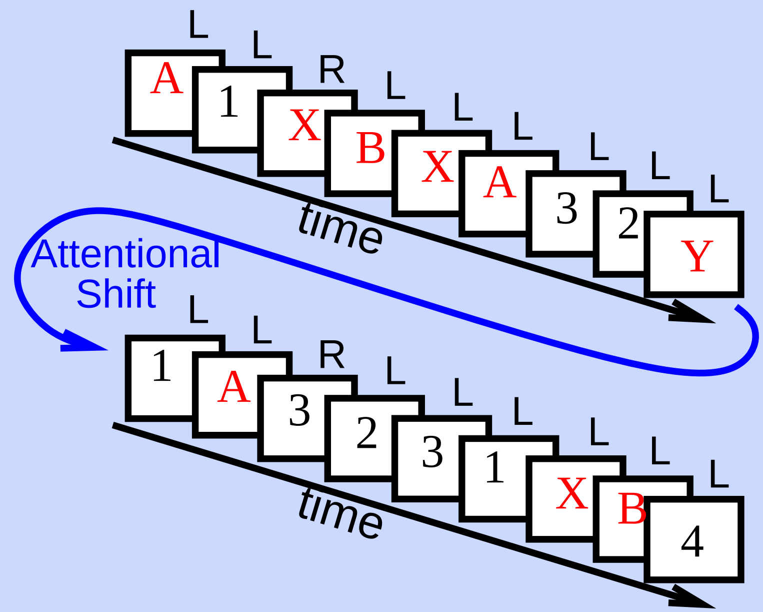
- WM context index = BX - AY accuracy

Psychopharmacological Studies

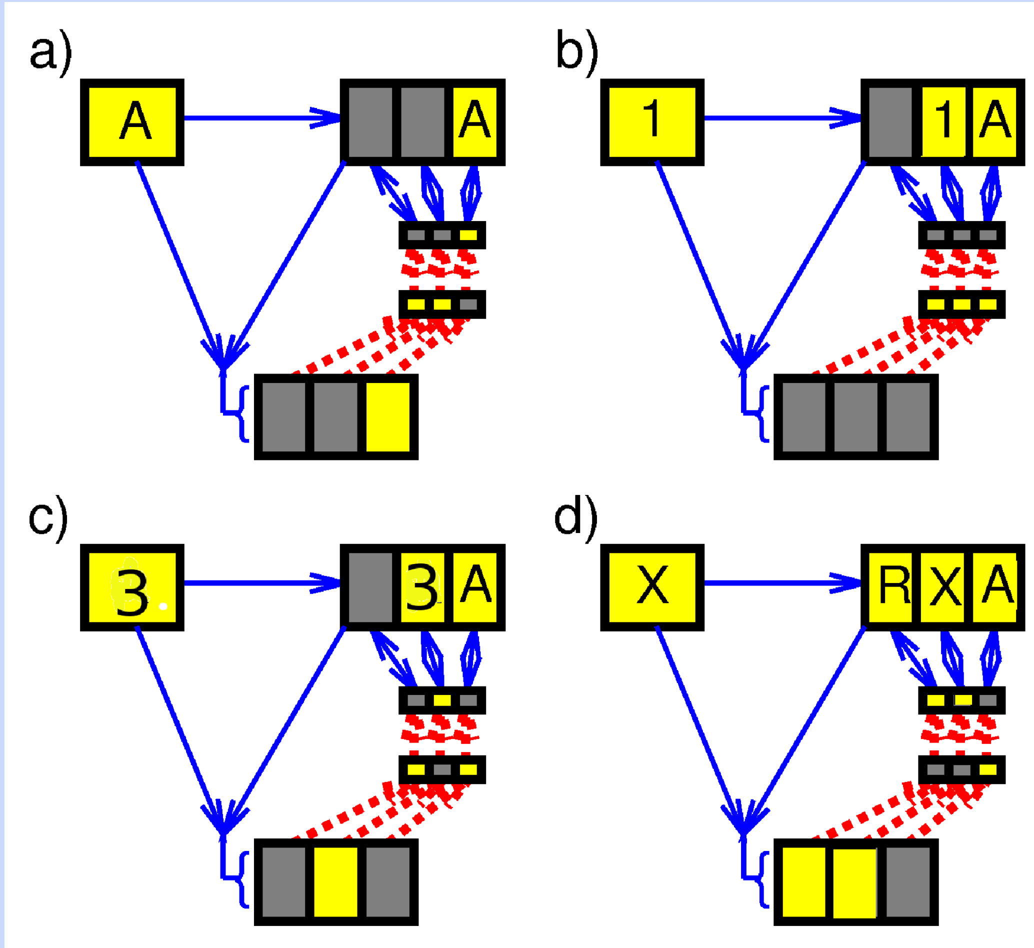
Frank & O'Reilly, 2006

- Double blind, within-subjects design (N=28).
- Cabergoline and Haloperidol: D2 agonist and D2 antagonist
- D2 agents: preferential action in BG
  (Camps et al, 1987; Moghaddam & Bunney, 1990;
  Arnsten et al., 1994; Seemans & Yang, 2004)

Working memory gating task

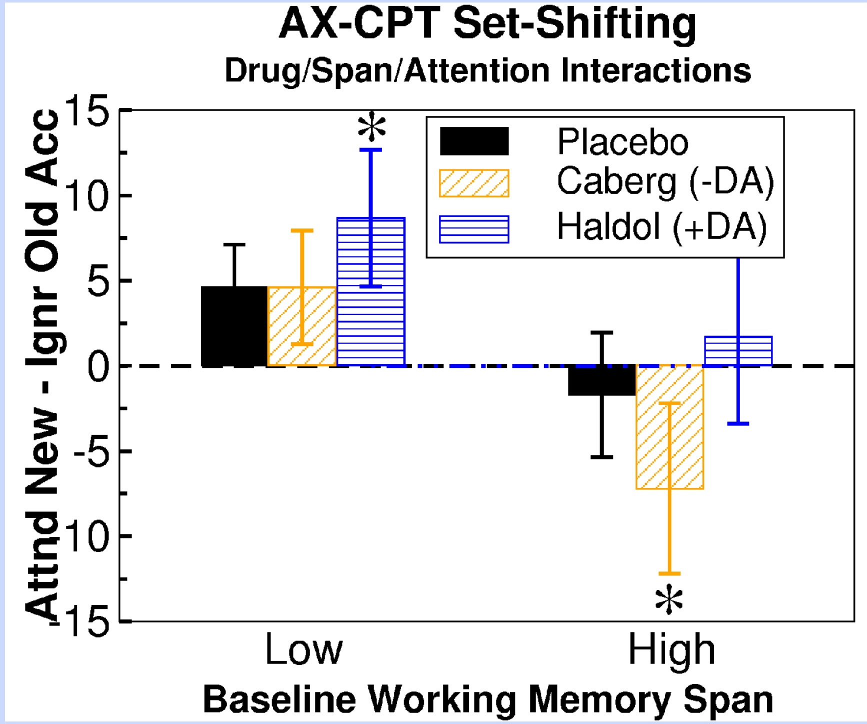


BG Gating of PFC Working Memory



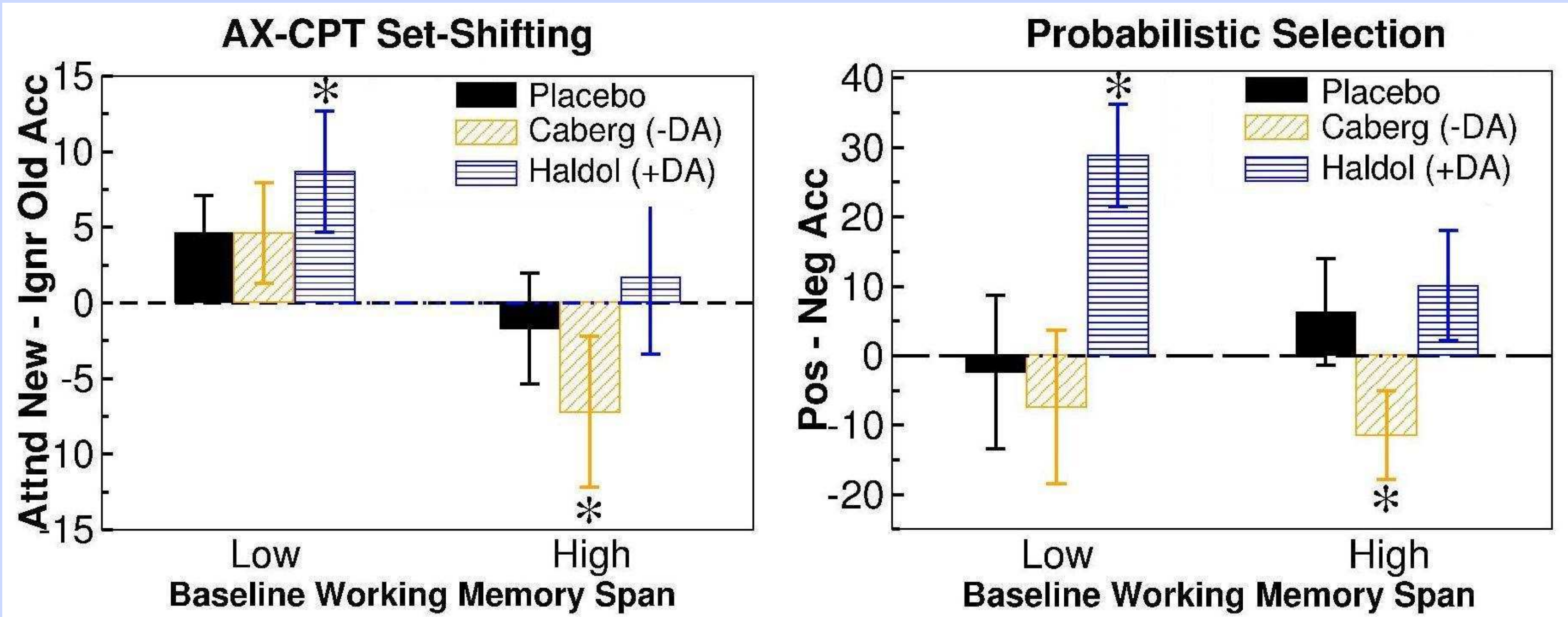
Frank et al 2001, O'Reilly & Frank 2006, Hazy et al 2007...

DA drug effects on working memory gating

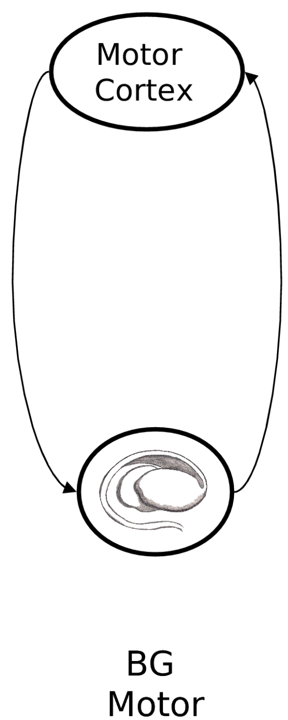


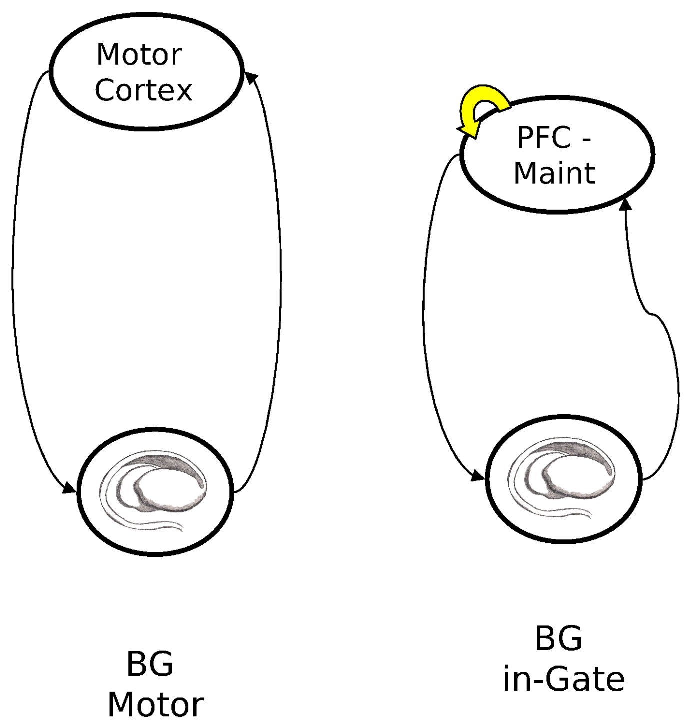
see Moustafa et al 08 and Frank et al 07 for similar drug effects in Parkinson’s and ADHD

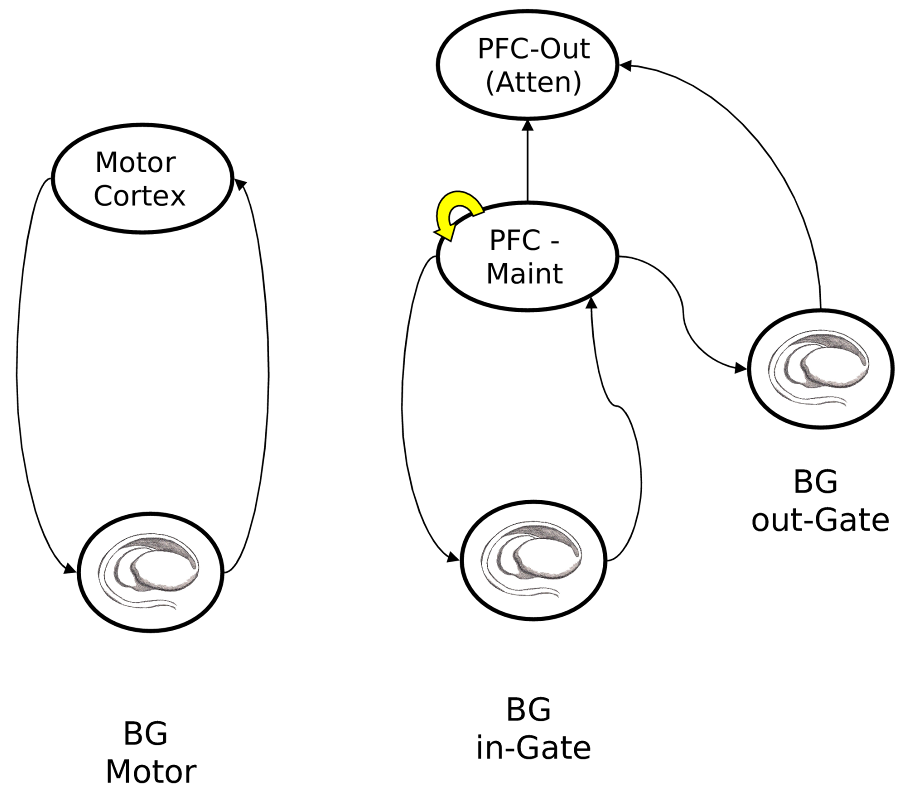
Analogous drug effects on learning and working memory



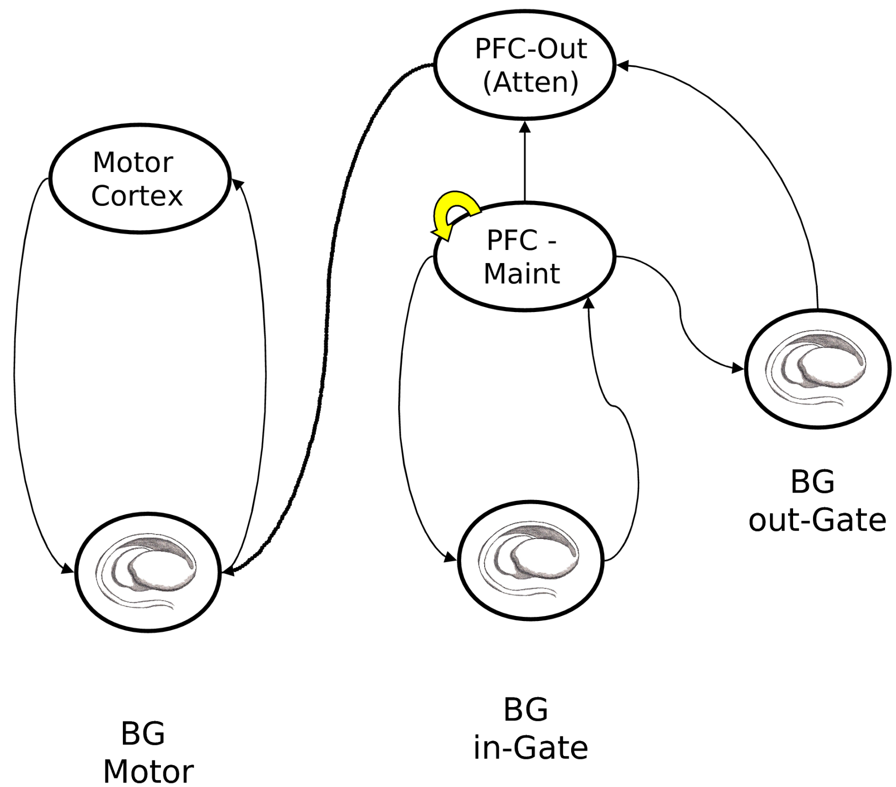
How do multiple BG-FC circuits interact in motivated behavior?


Motor
Cortex
BG
Motor

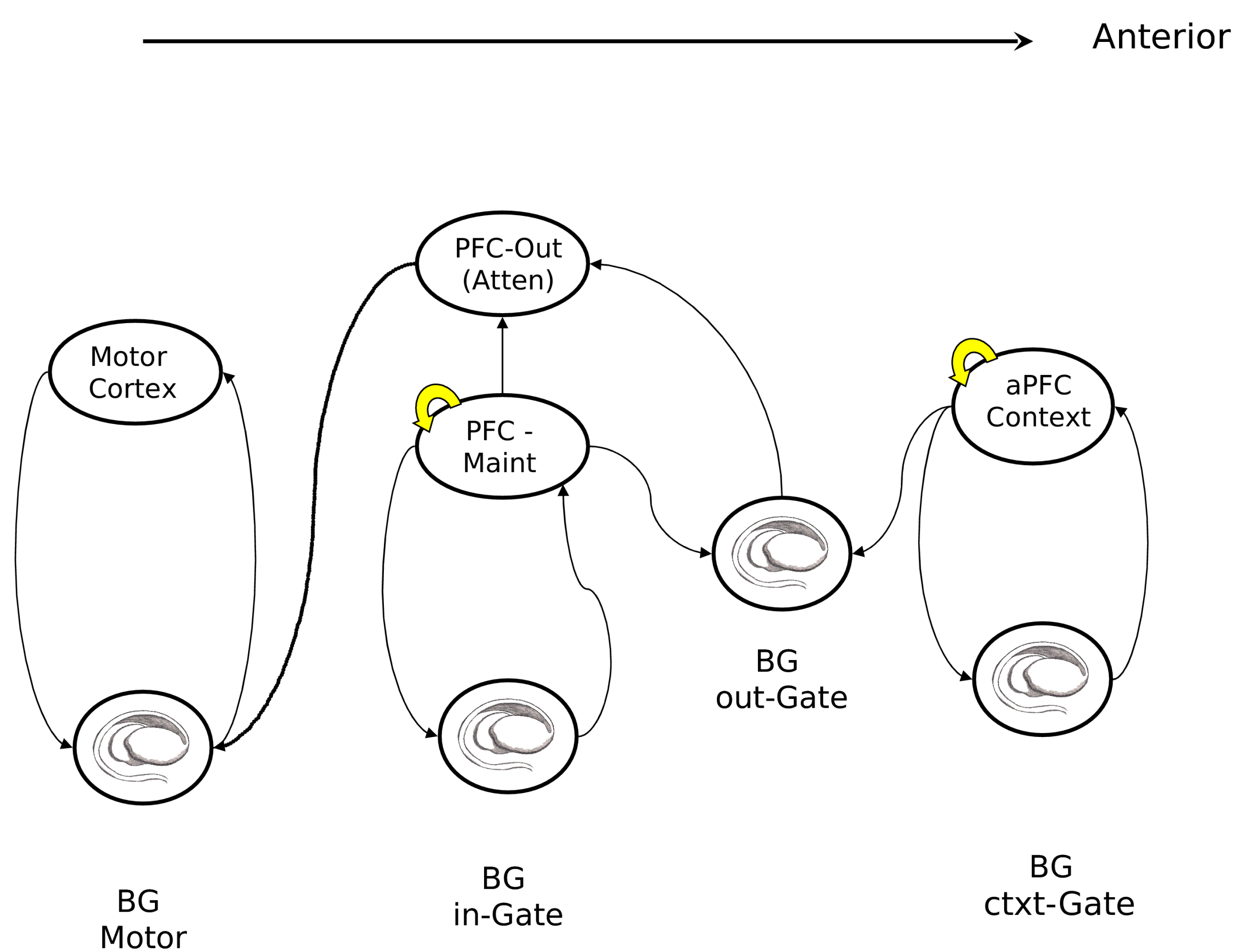

Motor
Cortex
PFC -
Maint
BG
Motor
BG
in-Gate



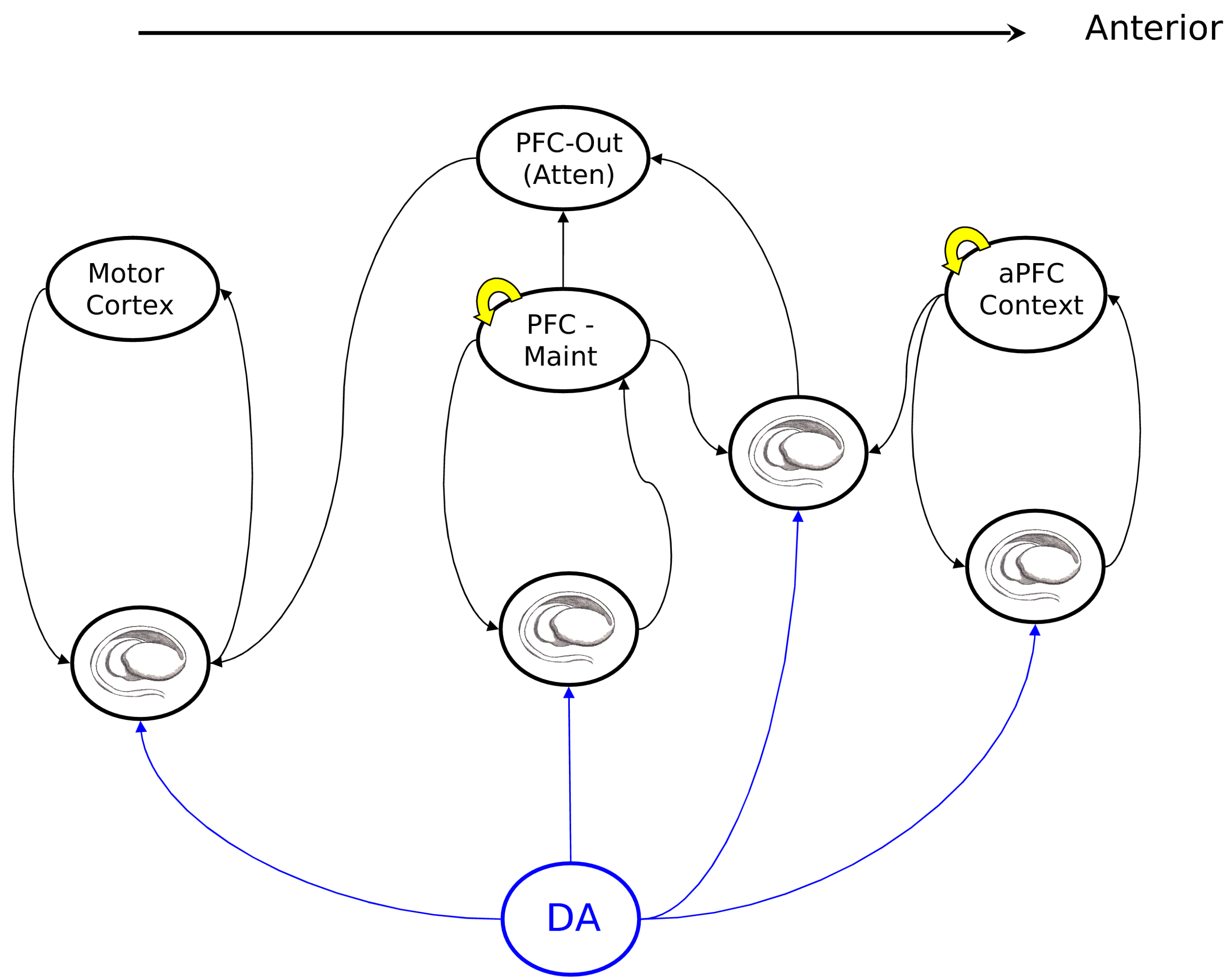
- Only some maintained PFC reps relevant for processing during intermediate stages of processing



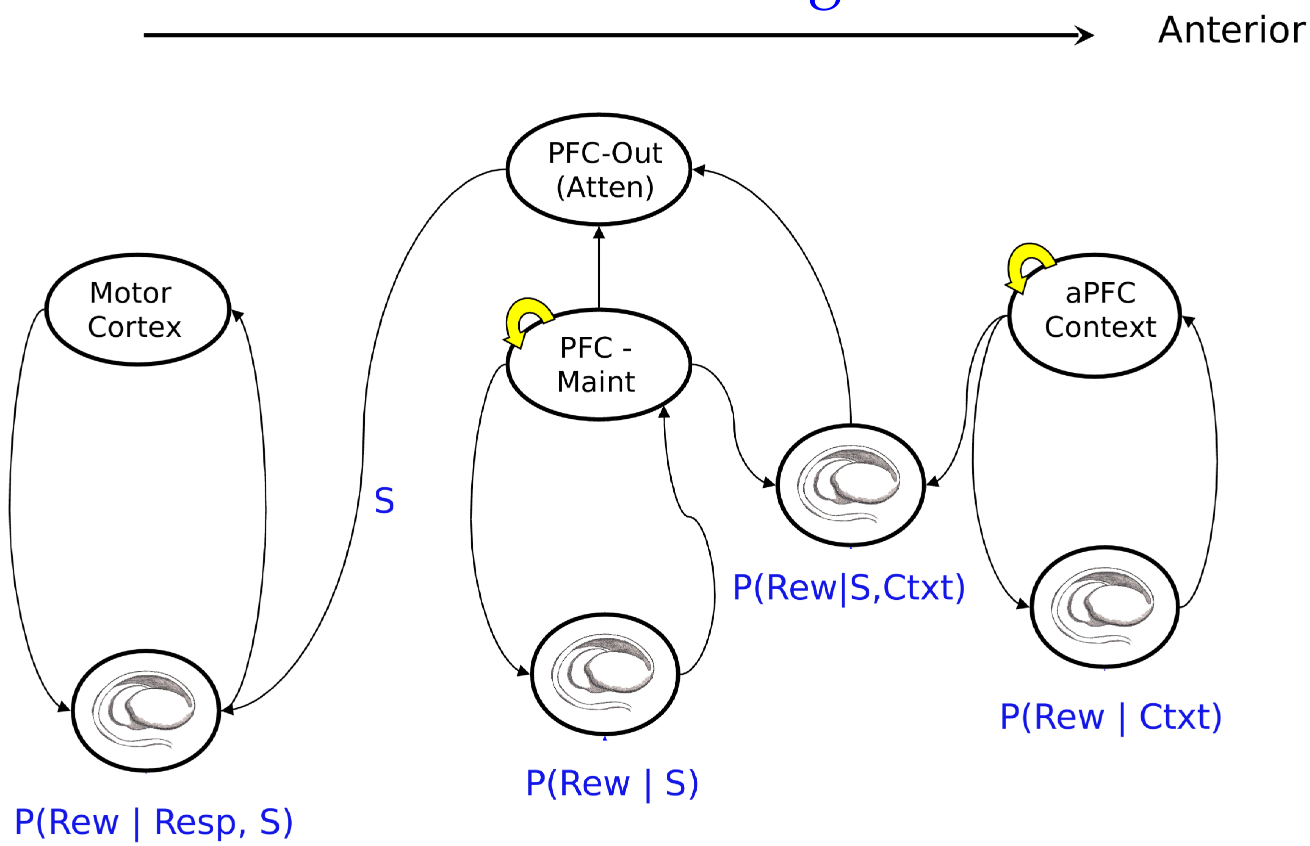
- Reduces number of S-R mappings needed to be learned by motor circuit



- Contextual dependencies for output gating (e.g., arithmetic); see also LSTM (Hochreiter & Schmidhuber 97)

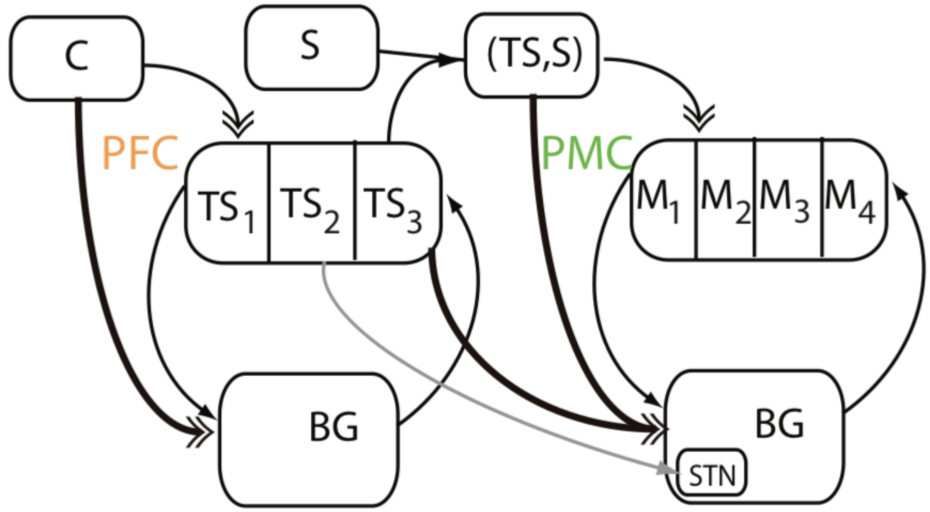

Anterior
PFC-Out
(Atten)
Motor
Cortex
PFC -
Maint
aPFC
Context
DA

Hierarchical interactions among BG-FC circuits



- BG gates frontal "actions" (motor, working memory, context)
- Anterior PFC influence posterior circuits via BG *output gating*
- BG-FC circuits learn reward probabilities via DA prediction error signals

Hierarchical interactions in BG-FC circuits: PFC influences on BG learning



Collins & Frank 2013, *Psych Rev*; Frank & Badre 2012

Broader speculations:
Why does motor control develop so slowly in humans??

- Standard story: infants born early due to large head, small birth canal
- ’Fourth trimester’

Broader speculations: Why does motor control develop so slowly in humans??

- Standard story: infants born early due to large head, small birth canal
- ’Fourth trimester’
- But 3 month old infants still pretty incompetent (from *babycenter.com*):

Broader speculations:
Why does motor control develop so slowly in humans??

- Standard story: infants born early due to large head, small birth canal
- ’Fourth trimester’
- But 3 month old infants still pretty incompetent (from *babycenter.com*):

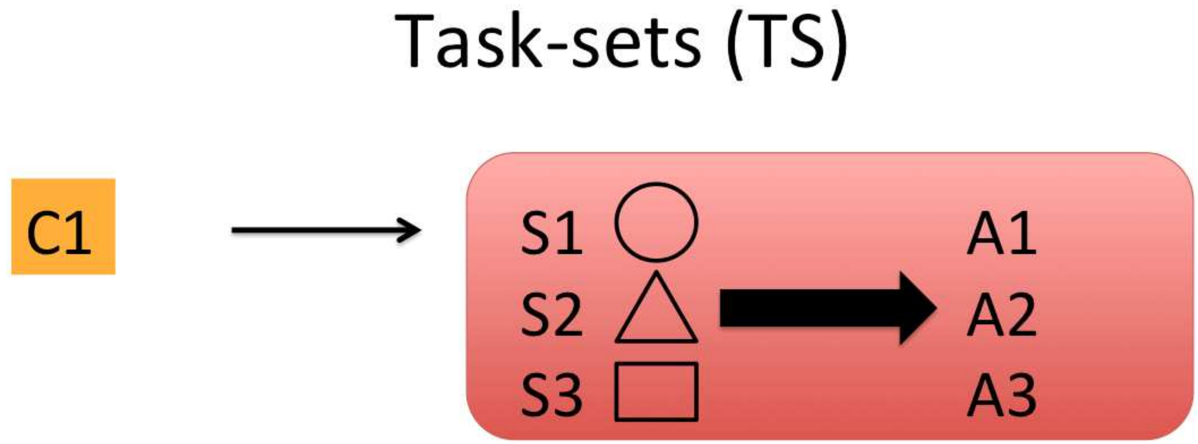
  “You no longer need to support his head. When he’s on his stomach, he can lift his head and chest. He can open and close his hands..”

Broader speculations:
Why does motor control develop so slowly in humans??

- Standard story: infants born early due to large head, small birth canal
- 'Fourth trimester'
- But 3 month old infants still pretty incompetent (from *babycenter.com*):

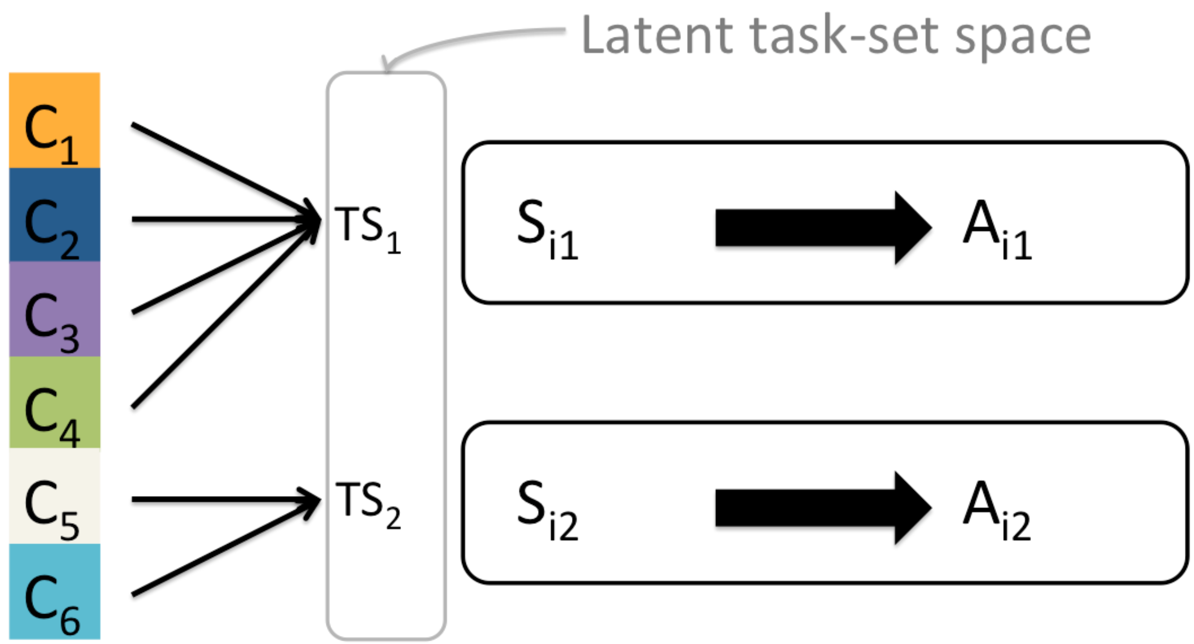
  "You no longer need to support his head. When he's on his stomach, he can lift his head and chest. He can open and close his hands.."

- Hypothesis: human brain is wired to discover generalizable structure....
  which is initially inefficient.


Task-sets (TS)
C1
S1
S2
S3
A1
A2
A3

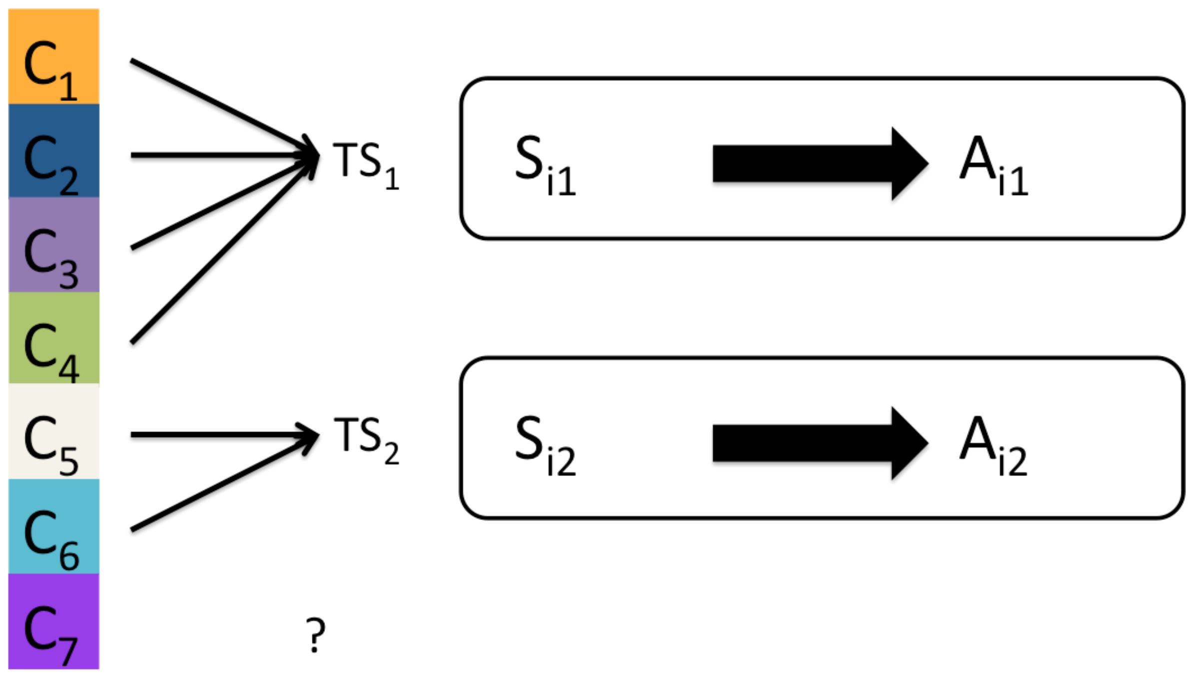
Task-sets (TS)

Abstracting Task-set rules

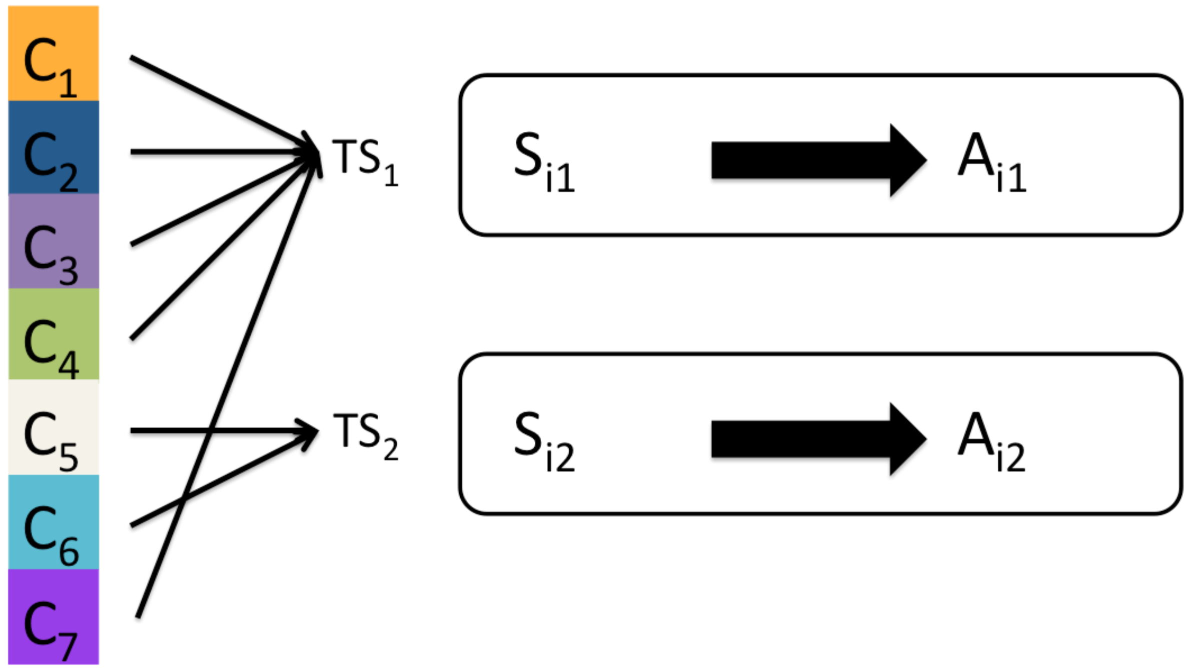


TS as abstract rule objects
Reverberi et al 2011
Woolgar et al 2011

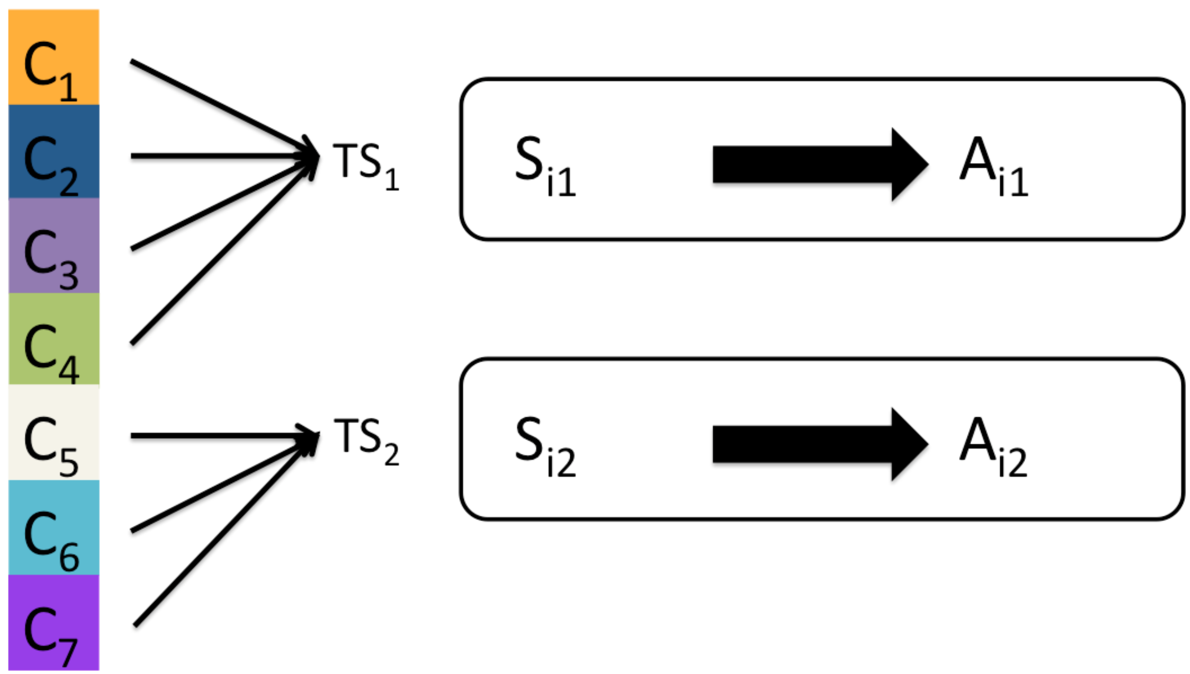
Abstracting Task-sets rules



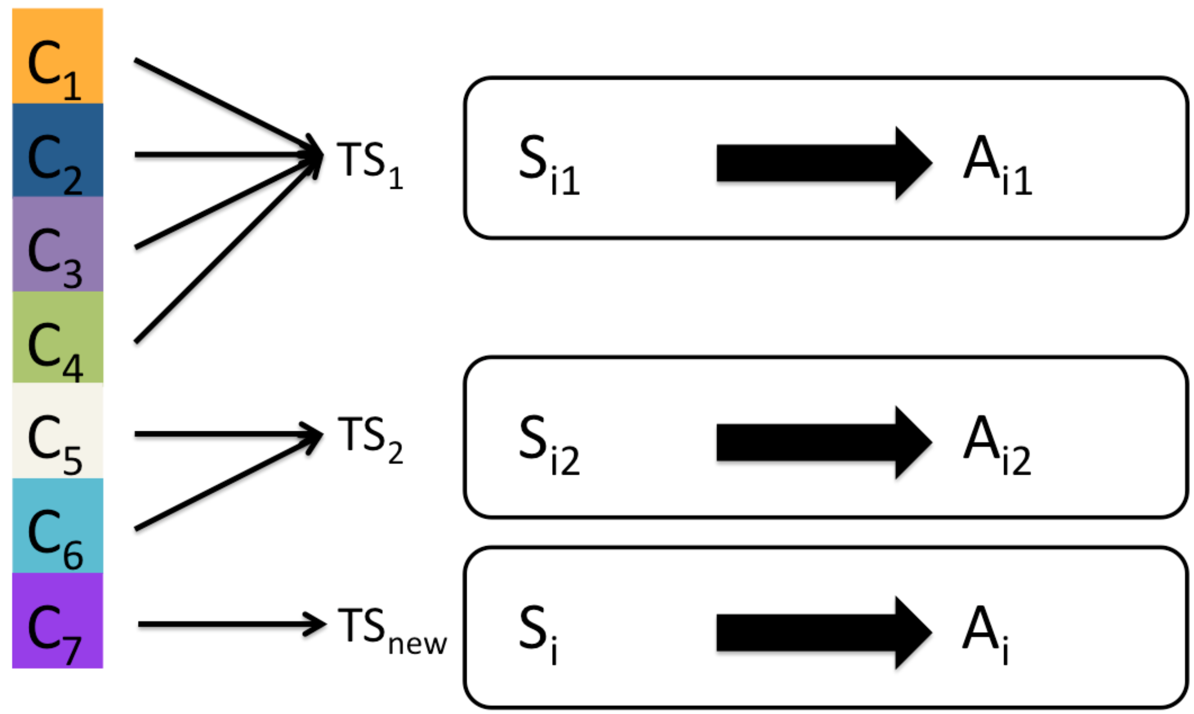
Abstracting Task-sets rules



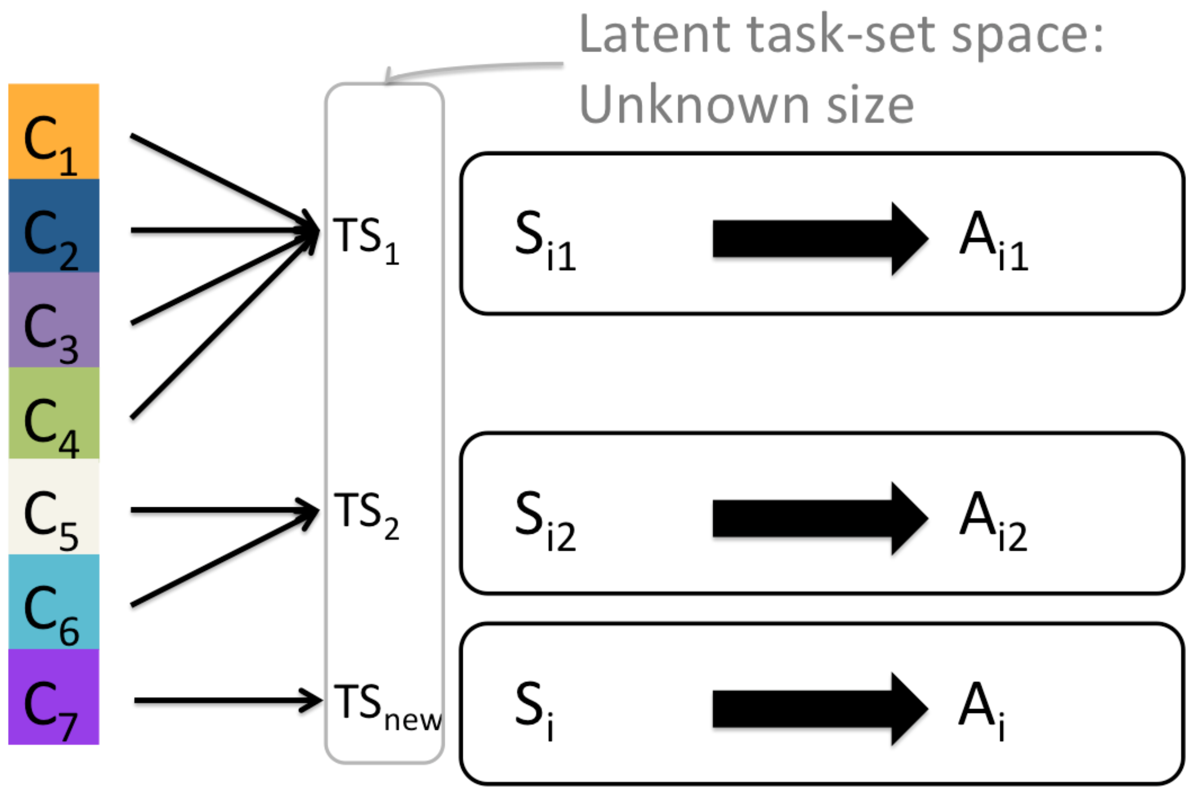
Abstracting Task-sets rules



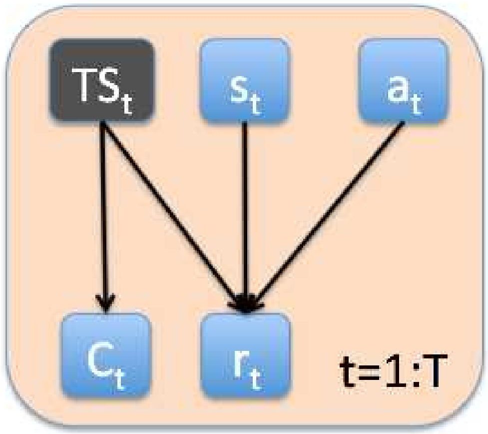
Abstracting Task-sets rules



Abstracting Task-sets rules



C-TS model



Task-sets are clustered across contexts and can be revisited in new contexts

- Prior prob on TS space given a new C:

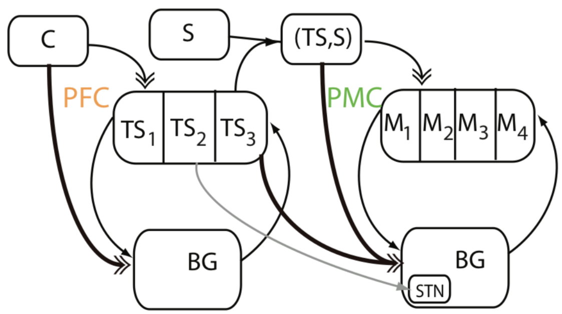
$$P(TS^* = . \,|c_{n+1}) = \begin{cases} P(TS^* = TS_{new}|c_{n+1}) & = & \alpha/A \\ \forall i \neq new,\ P(TS^* = TS_i|c_{n+1}) & = & \sum_j P(TS_i|c_j)/A \end{cases}$$

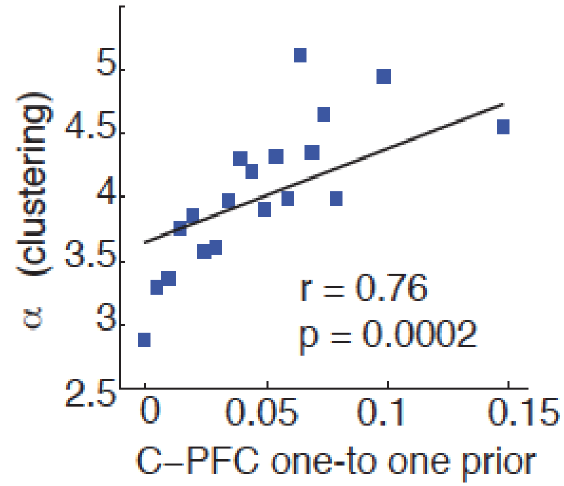
- $\alpha > 0$: Clustering parameter

- Chinese restaurant process Jordan, Blei Teh 2005

see also Gershman et al 2010

Model mimicry: C-TS and hierarchical BG-PFC network

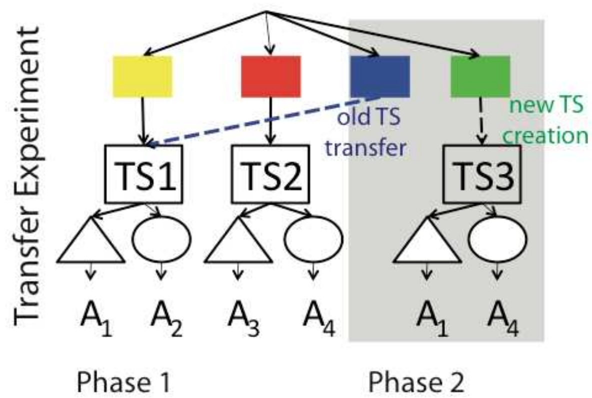


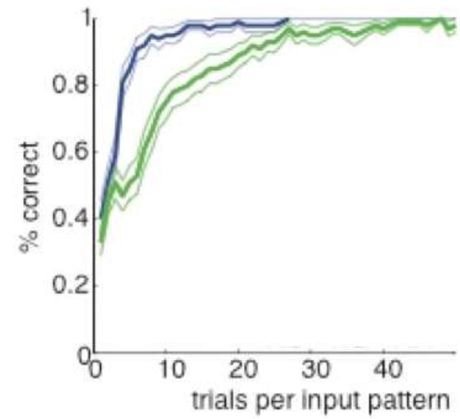


- Sparseness of context-PFC connectivity matrix is linked to α clustering
- *Both models are approximations of the same process: building TS structure*
- fMRI evidence for hierarchical PFC-BG mechanisms Badre et al 2010; Badre & Frank 2012

Collins & Frank 2013 *Psych Rev*

Neural Network – Results





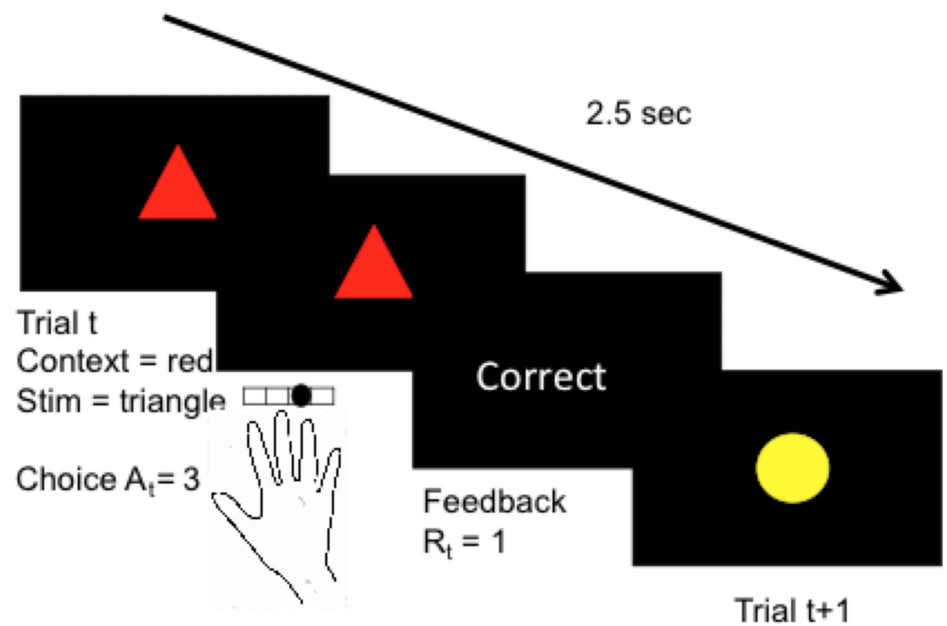
Phase 2
new contexts:
Old TS
New TS

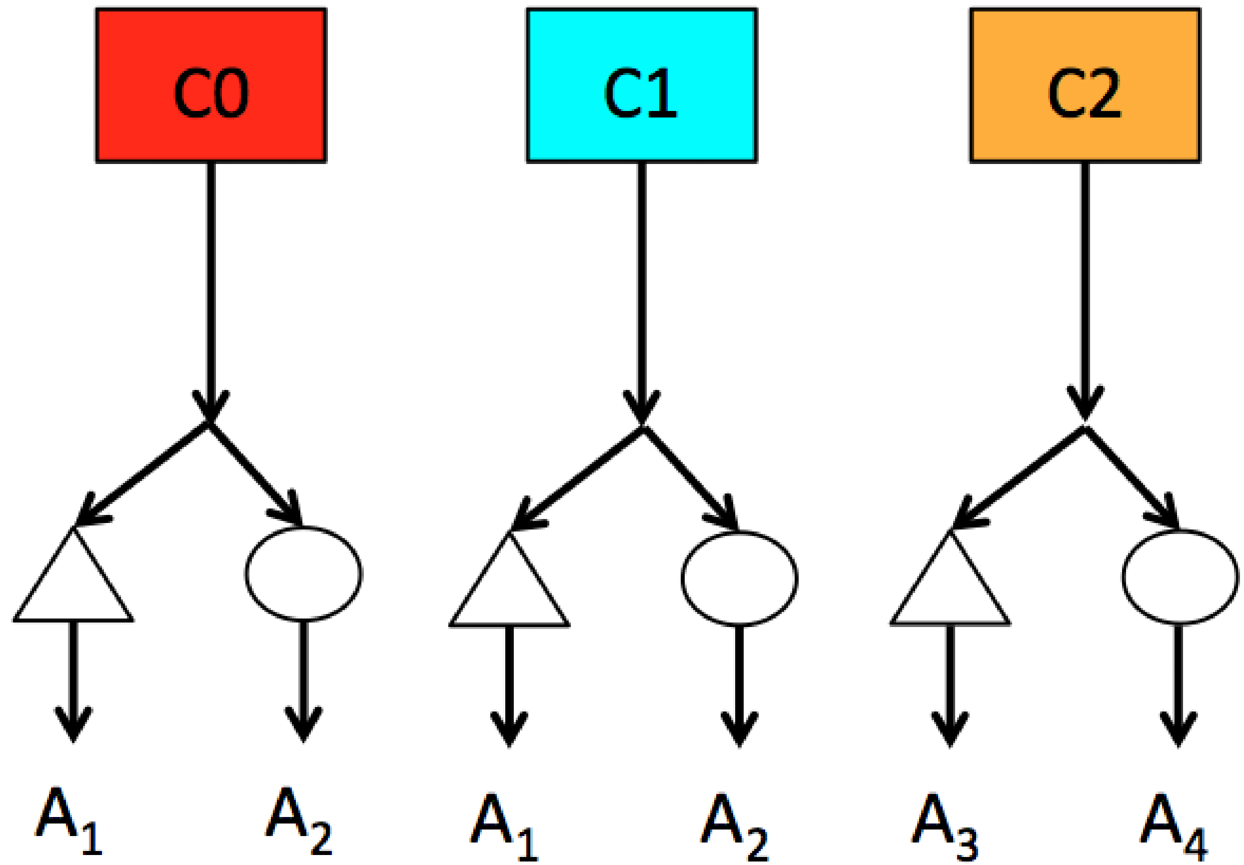
The network learns efficiently unsupervised,
Predicts positive, negative transfer

Collins & Frank, Psych Review, 2013

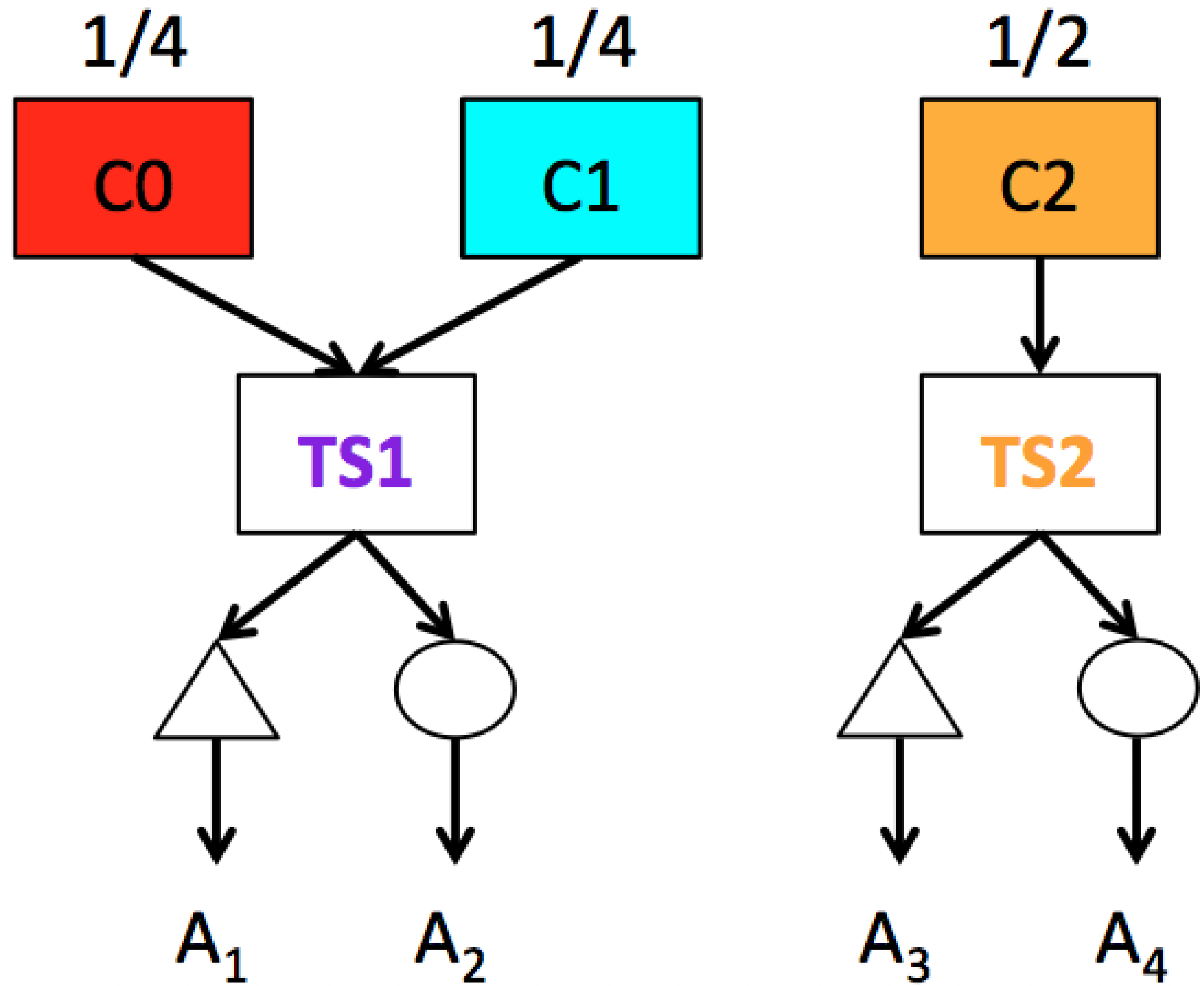
Experiment:

- Instrumental learning
  - E.g. colored shapes

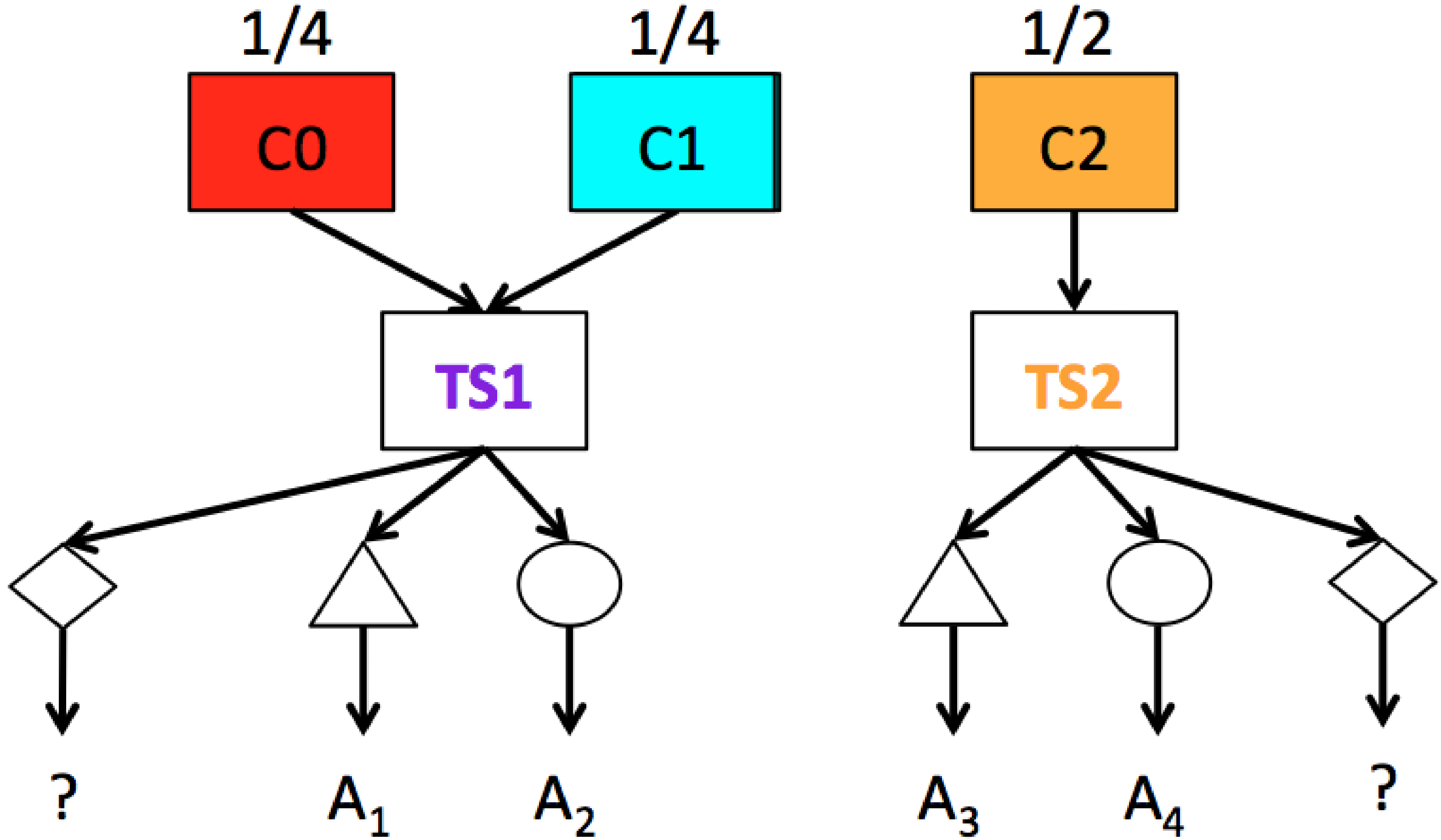




C0
C1
C2
A1
A2
A1
A2
A3
A4



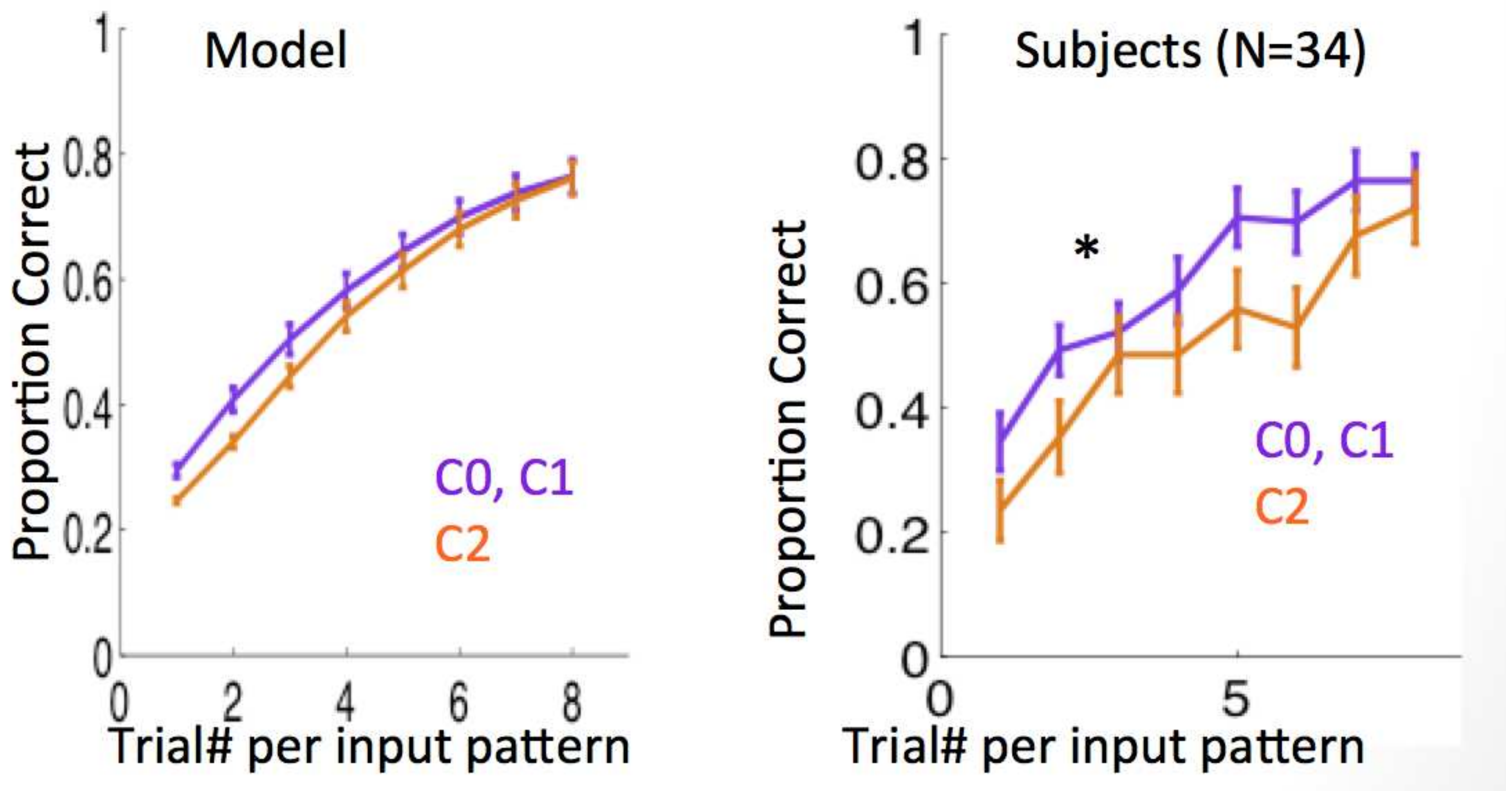
1/4
C0
1/4
C1
1/2
C2
TS1
TS2
A1
A2
A3
A4



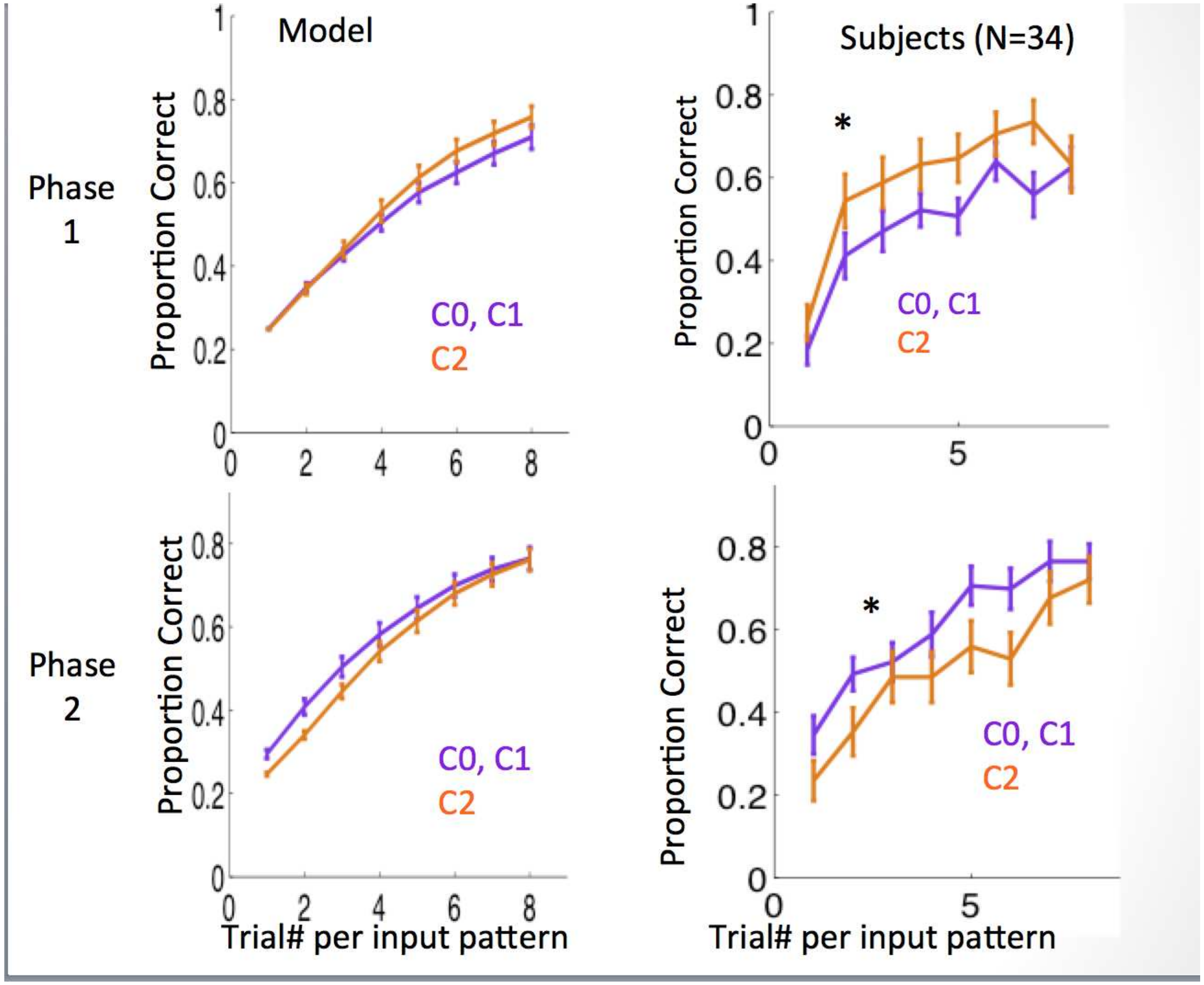
1/4
C0
1/4
C1
1/2
C2
TS1
TS2
?
A1
A2
A3
A4
?

Clustering affords faster learning within existing rule sets...

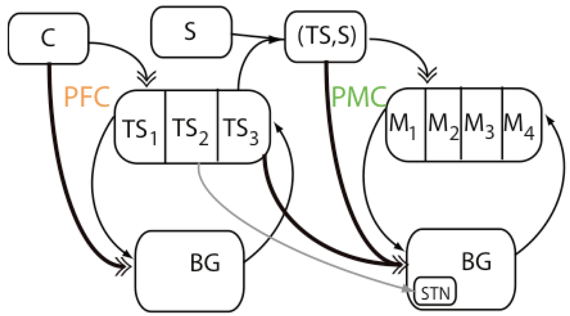
Phase 2

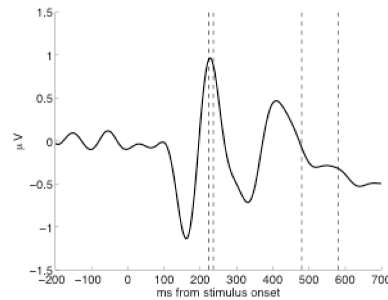


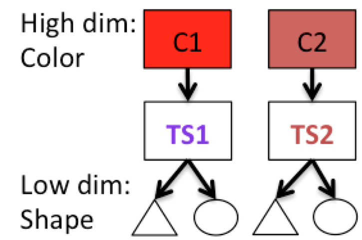
But initial (phase 1) clustering is inefficient and slow!

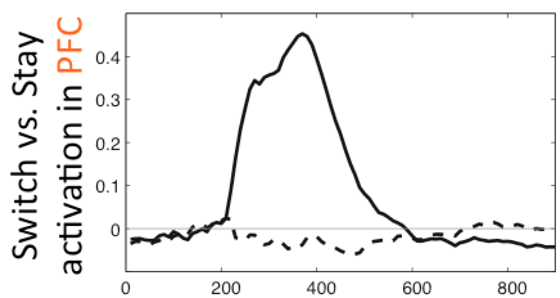


Neural model predictions and EEG

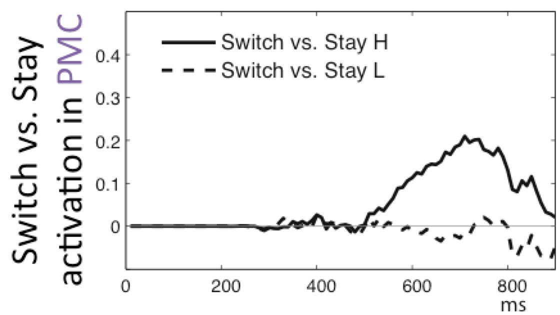




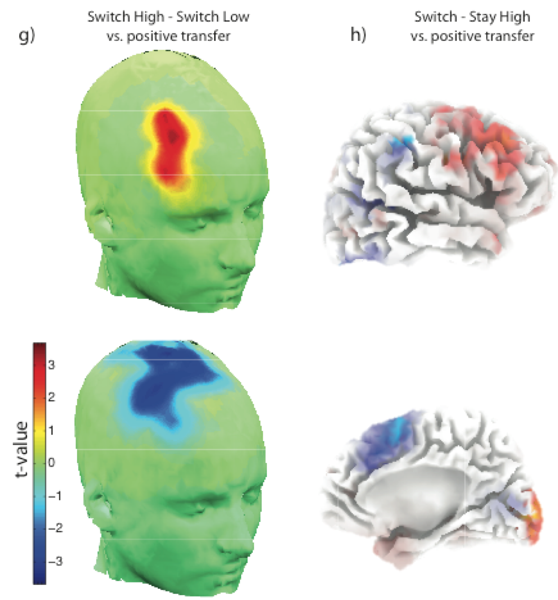




Early
TS-switch
positivity

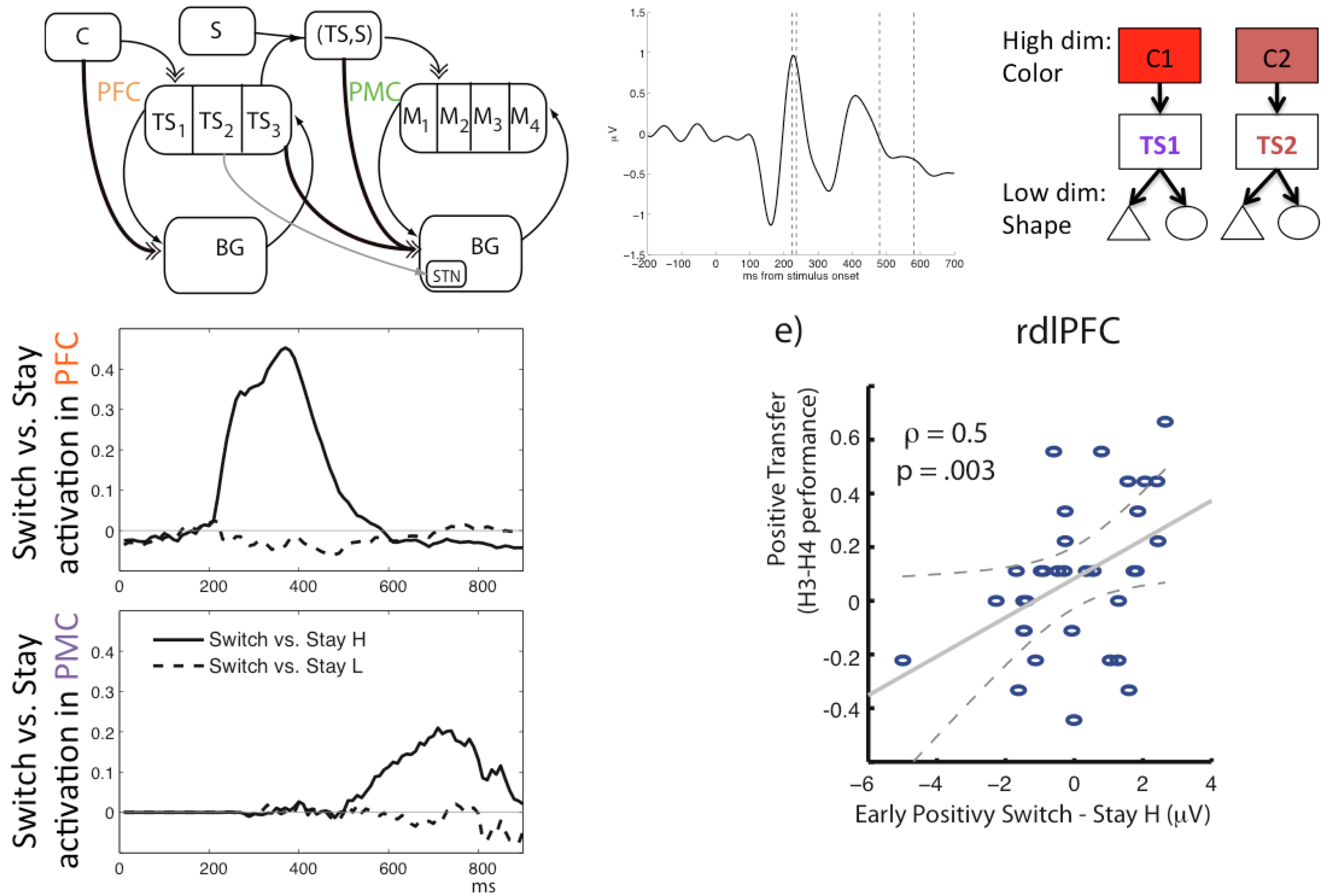


Late
TS-switch
negativity



Collins, Cavanagh & Frank, JoN, 2014

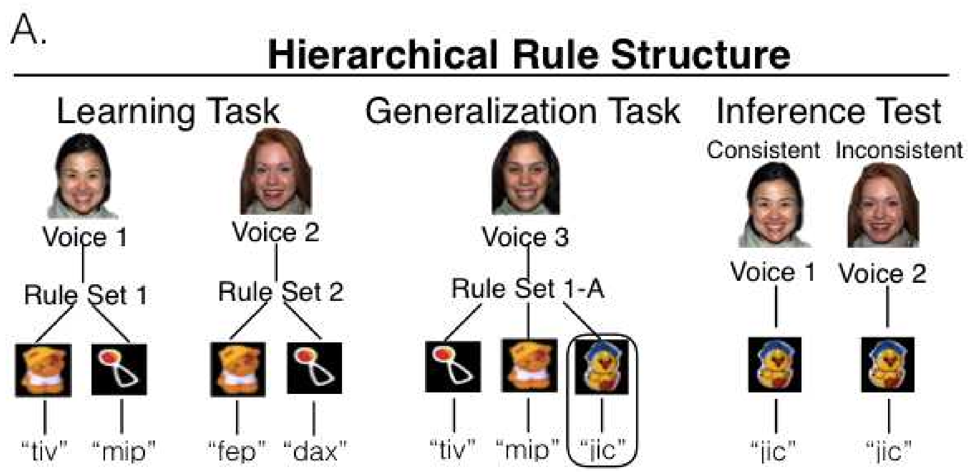
Neural model predictions and EEG

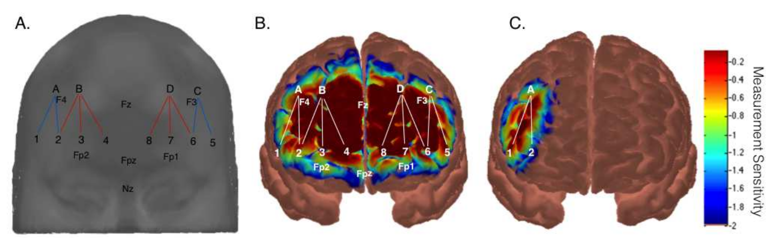


Collins, Cavanagh & Frank, JoN, 2014

Now, in Infants!

8-Month-Old Infants Spontaneously Learn and Generalize Hierarchical Rules

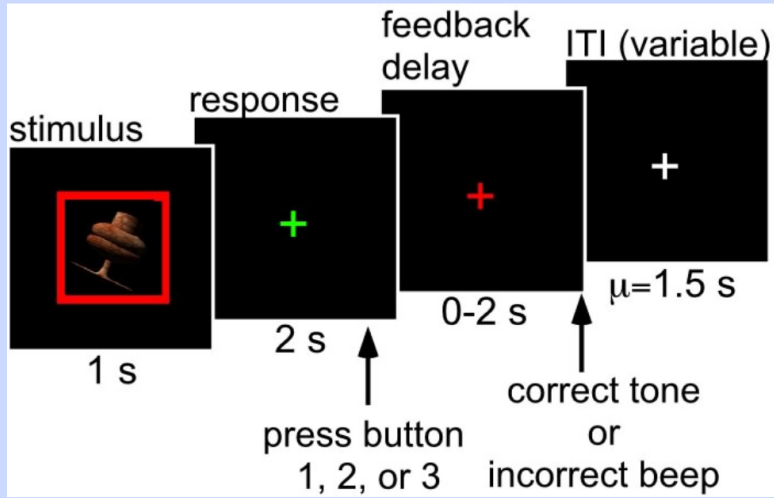


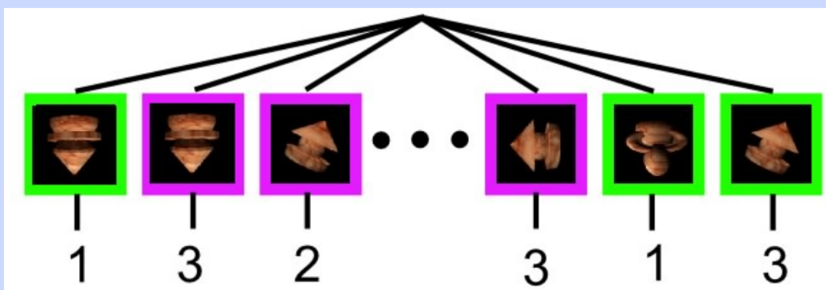


Werchan et al, 2015; *in prep*

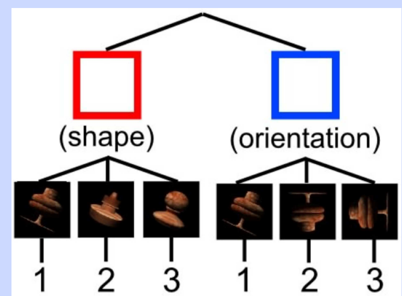
Badre Hierarchical Learning Task

Badre, Kayser & D'Esposito, 2010



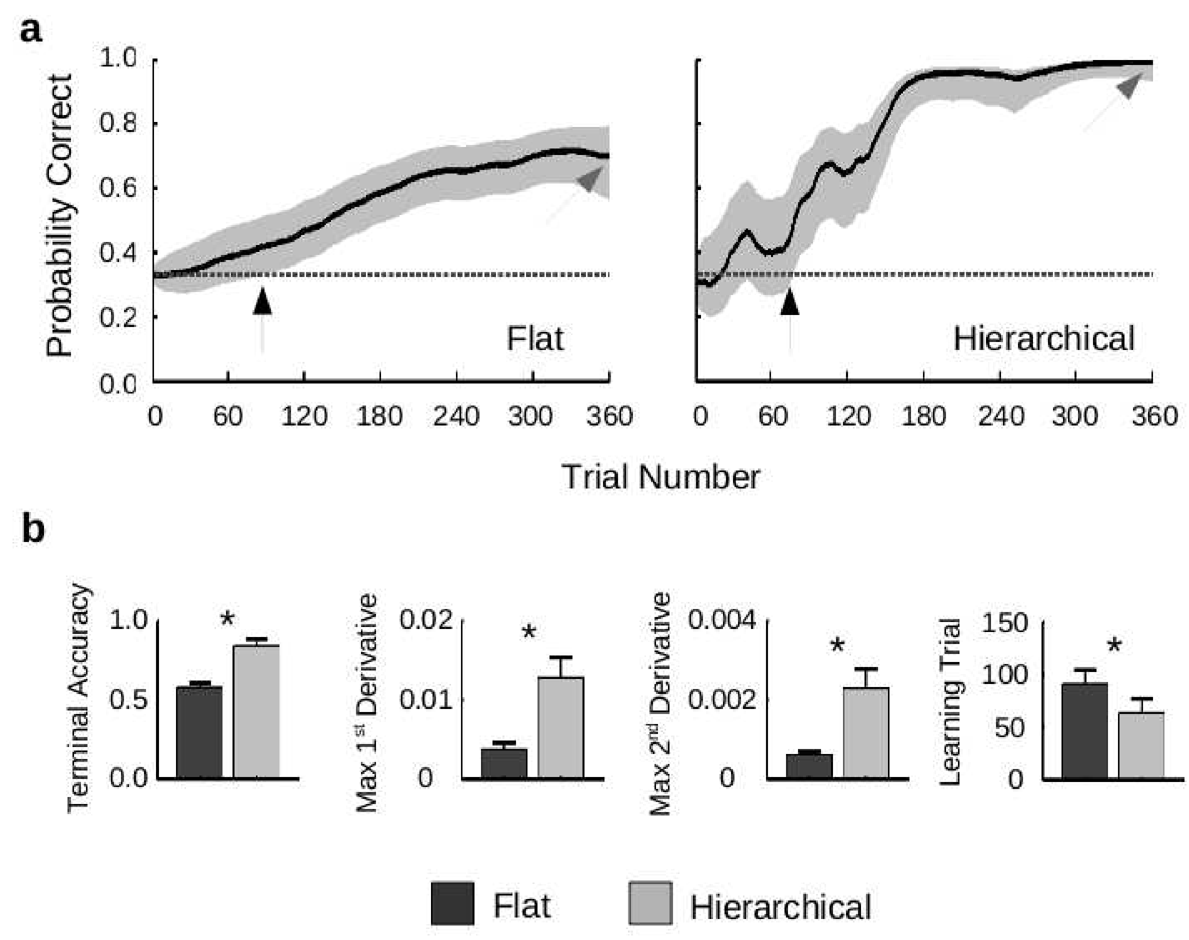


Flat (conjunctive)

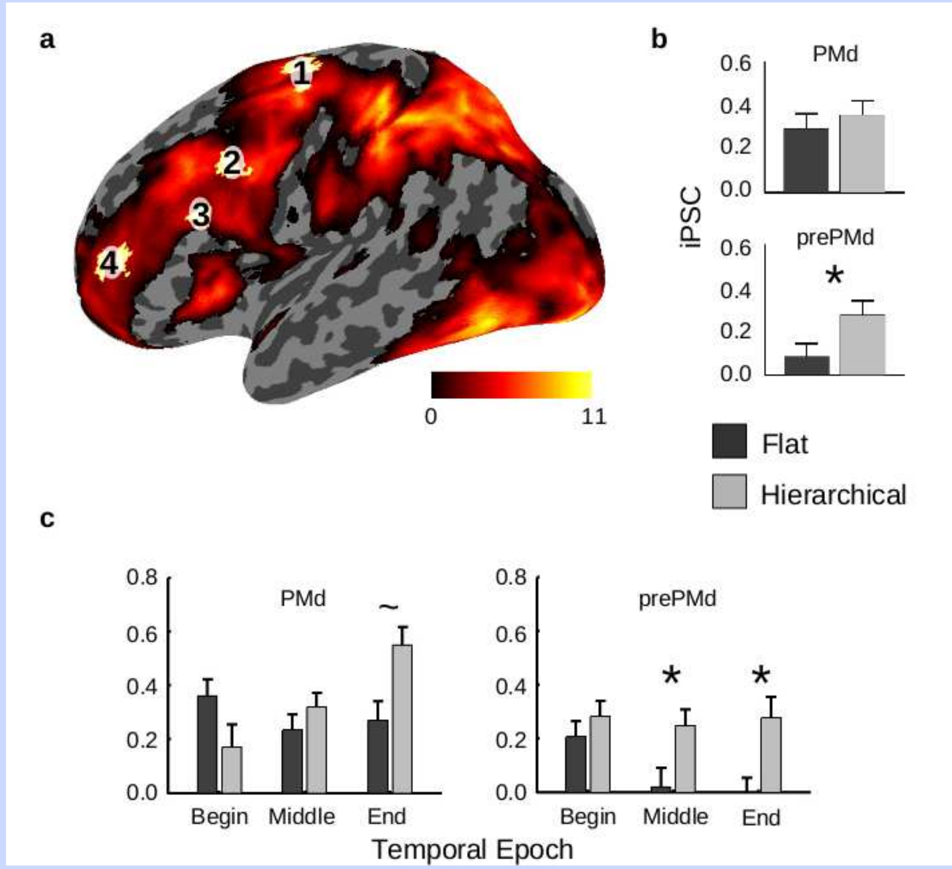


Hierarchical

Learning Curves



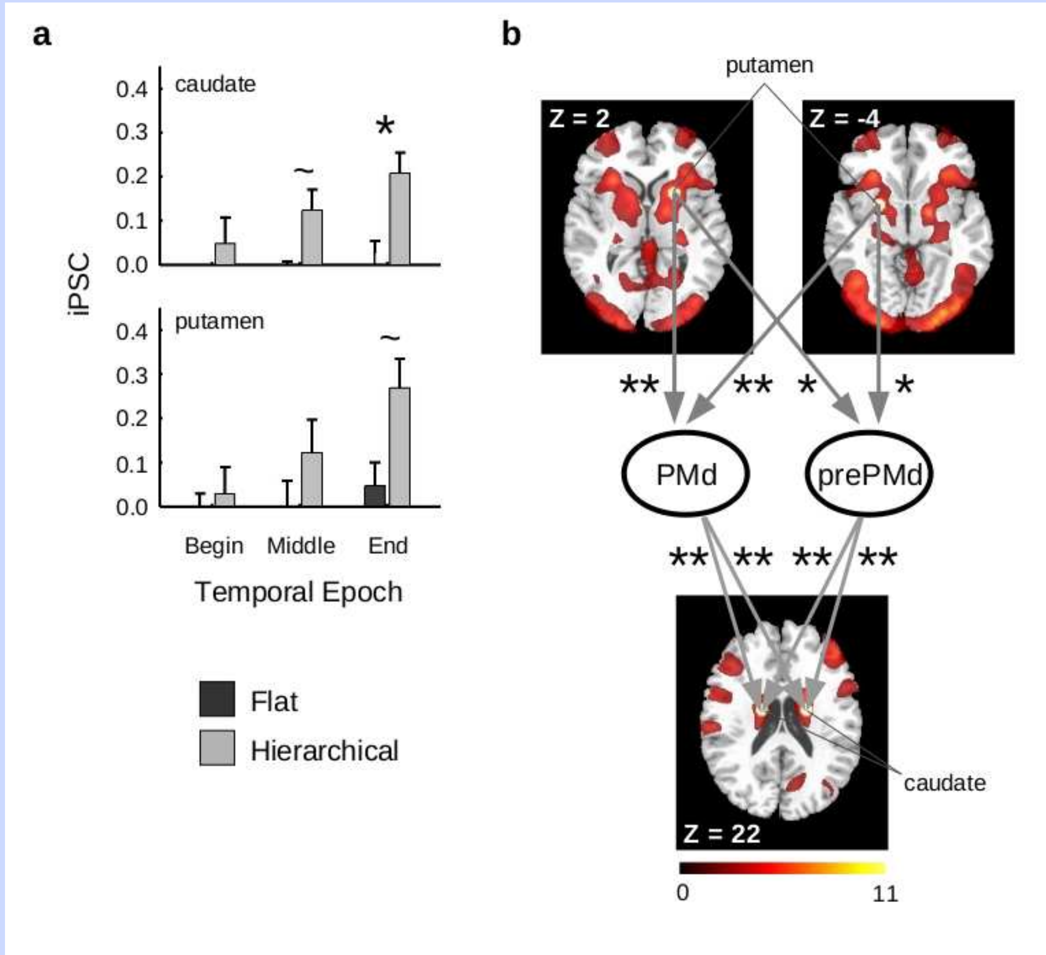
Rostro-caudal axis and hierarchical control



prePMD:

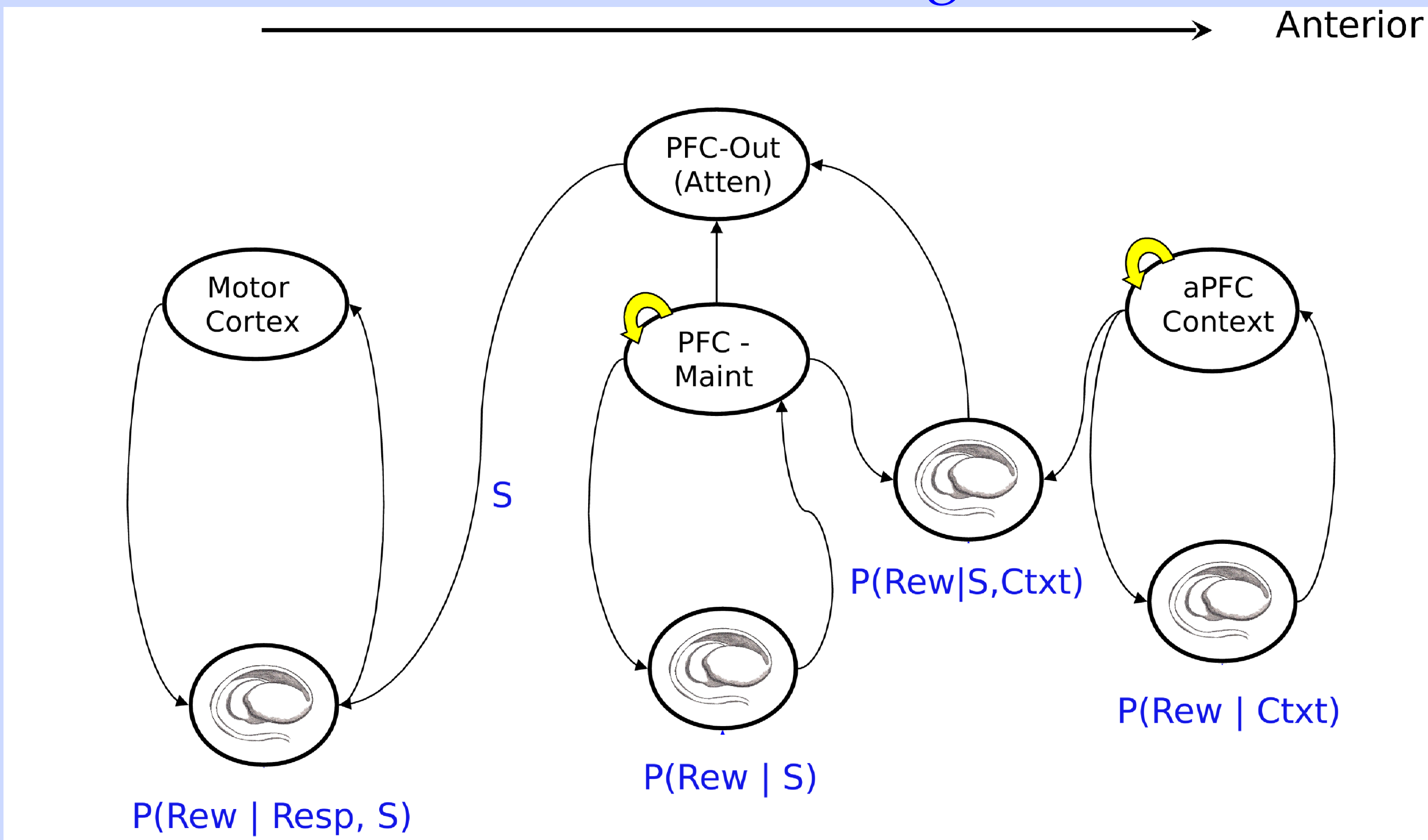
- 2nd level of hierarchical control (from Badre, Koechlin)
- activity predicts learning in 2nd level hier condition, declines in flat condition

Striatum implicated in hierarchical rule learning

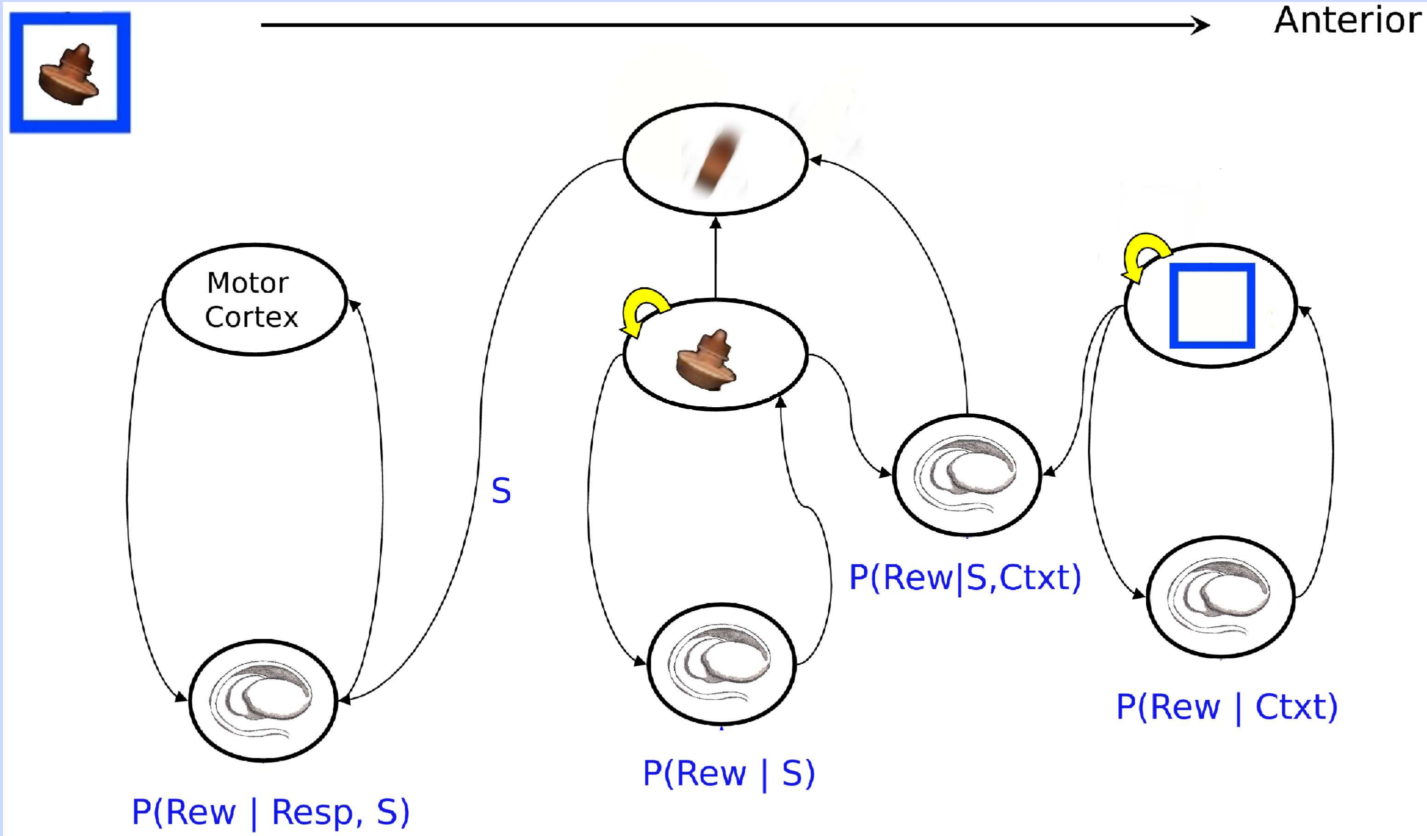


How do BG-PFC circuits interact to discover hierarchical structure?

Hierarchical interactions among BG-FC circuits

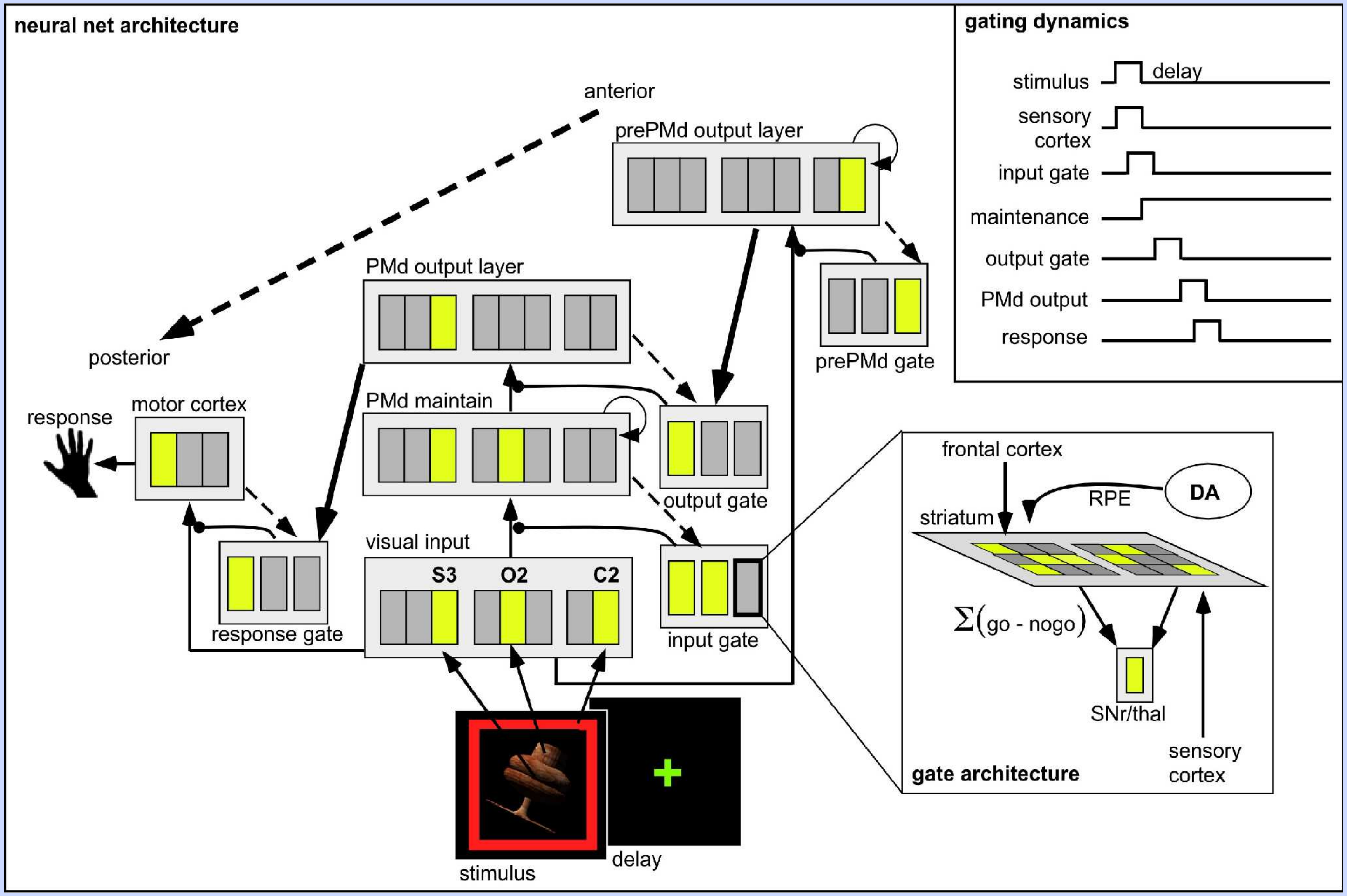


Application to Badre task

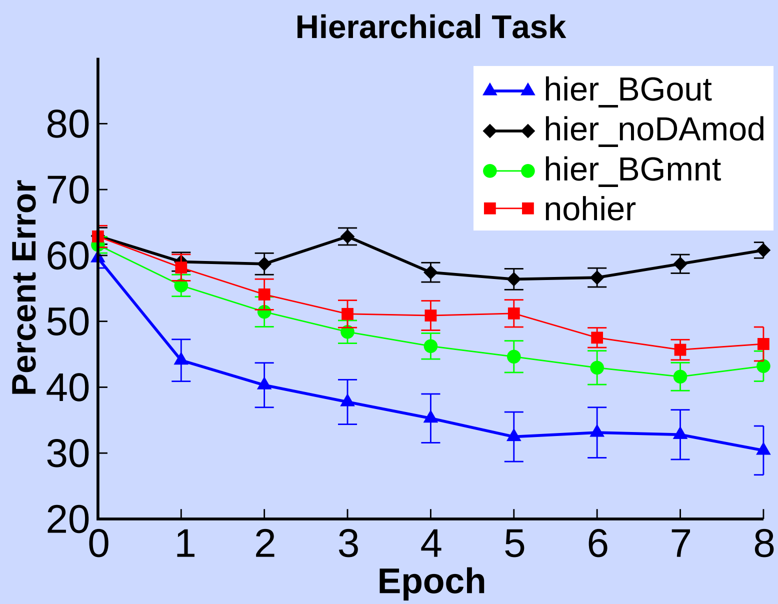


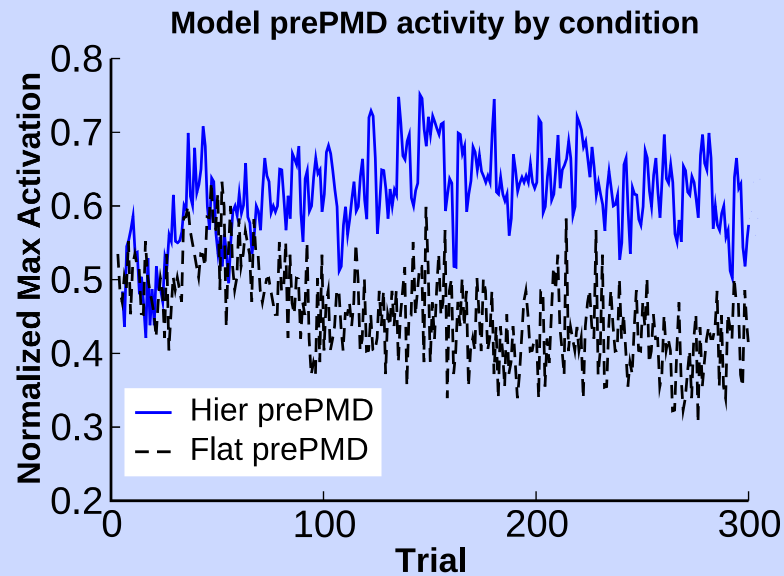
- Anterior PFC affects posterior circuits via BG *output gating*
- Reduces # of stimulus-response mappings needed to be learned in motor circuit.

Application to Badre task



Modeling BG-PFC hierarchical learning





- DA-based RL in gating networks needed for rapid learning in Hier cond
- As in fMRI data, model prePMD activity decreases in Flat relative to Hier
- This is due to reward prediction error signals punishing gating of prePMD when no hierarchical structure exists (stay tuned)

Modeling individual learners: abstract account

- Neural model makes plausible links to biology, but unsuitable for quantitative fits to individual subject behavior
- We developed abstract version of hierarchical rule learning using Bayesian mixture of experts (MoE), motivated by neural model

Modeling individual learners: abstract account

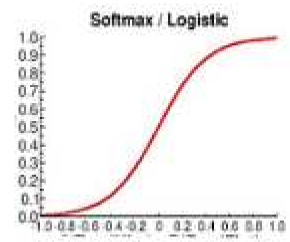
- Separate experts learn $P(Rew|Resp, Shape)$, $P(Rew|Resp, Orient)$, etc..

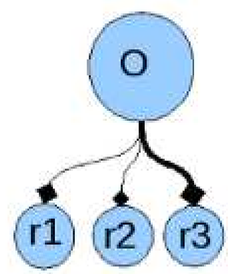
Modeling individual learners: abstract account

- Separate experts learn $P(Rew|Resp, Shape)$, $P(Rew|Resp, Orient)$, etc..
- Hierarchical experts separately learn statistics about each dimension contingent on a candidate higher order feature
  (cf. prePMD context for output gating).

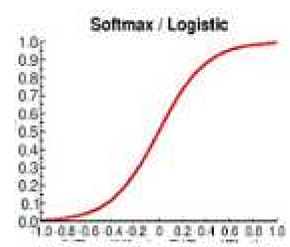
Modeling individual learners: abstract account

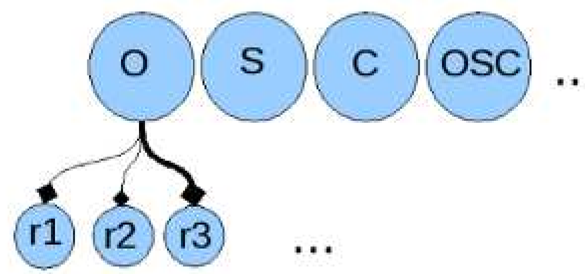
- Separate experts learn $P(Rew|Resp, Shape)$, $P(Rew|Resp, Orient)$, etc..
- Hierarchical experts separately learn statistics about each dimension contingent on a candidate higher order feature
  (cf. prePMD context for output gating).
- Credit assignment mechanism learns probability that each expert contributes to observed rewards
- Attentional mechanism selects among experts based on learned probability of success $P(Rew|e)\ \forall e$.

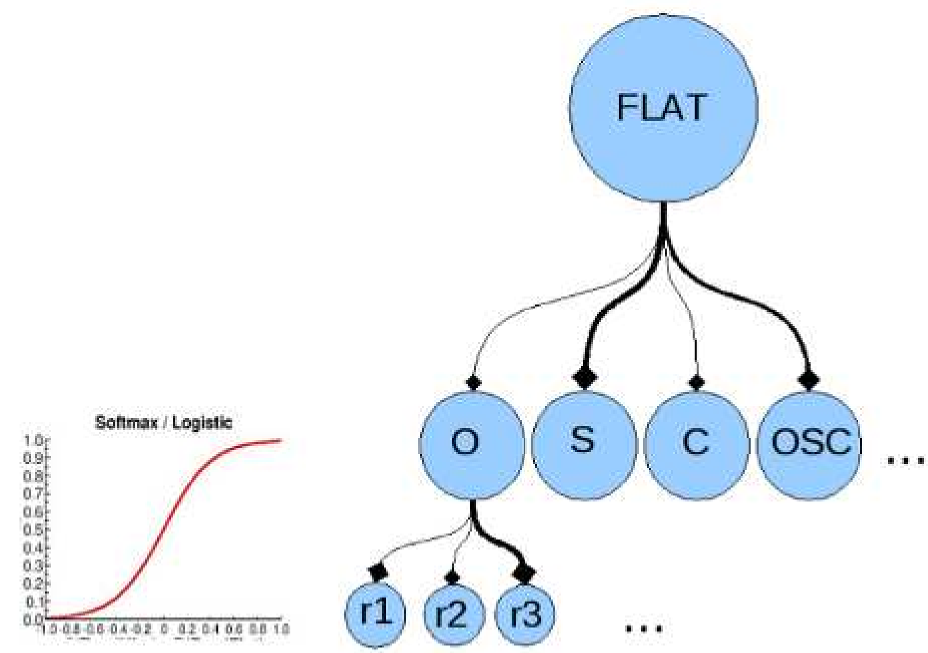




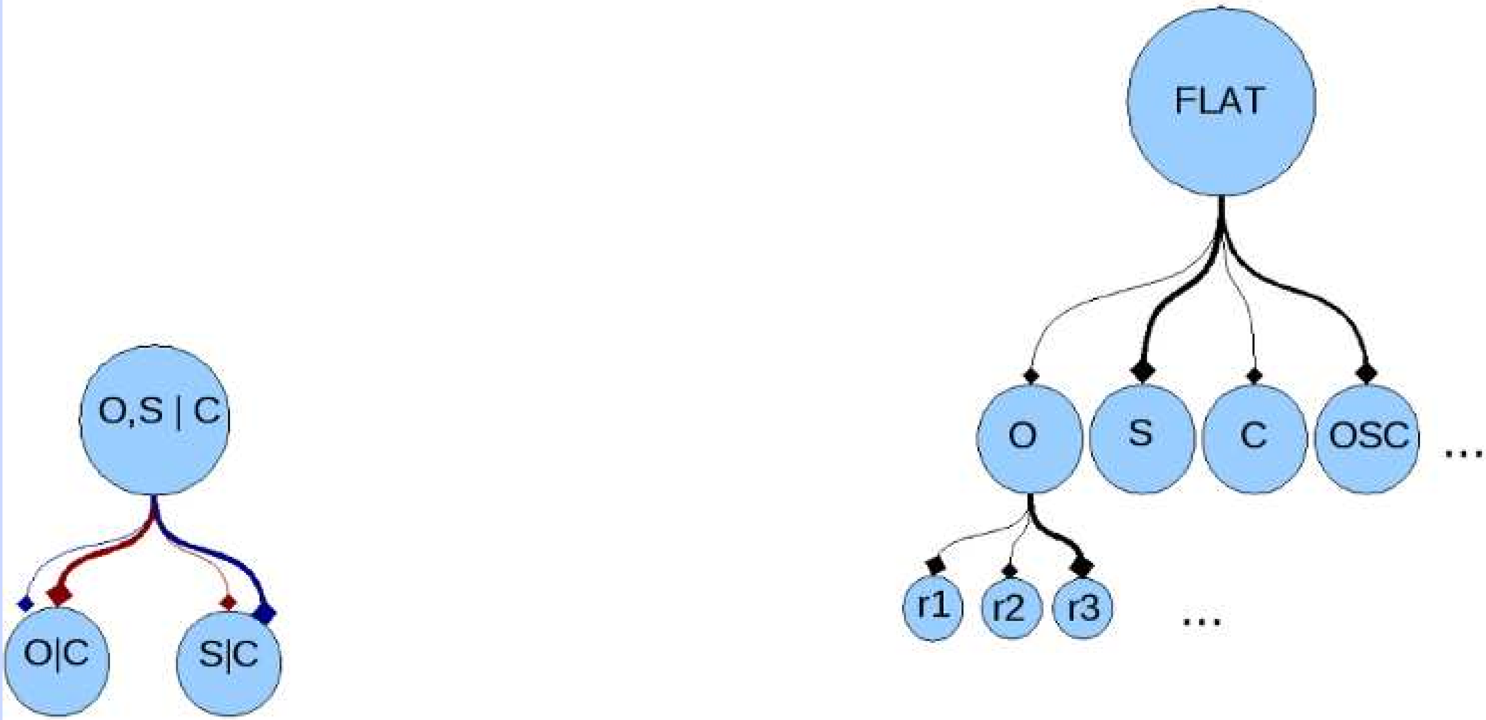
$$P(\theta_{r,O}|\delta_1 ... \delta_n) \propto P(\delta_1 ... \delta_n|\theta_{r,O}) P(\theta_{r,O})$$

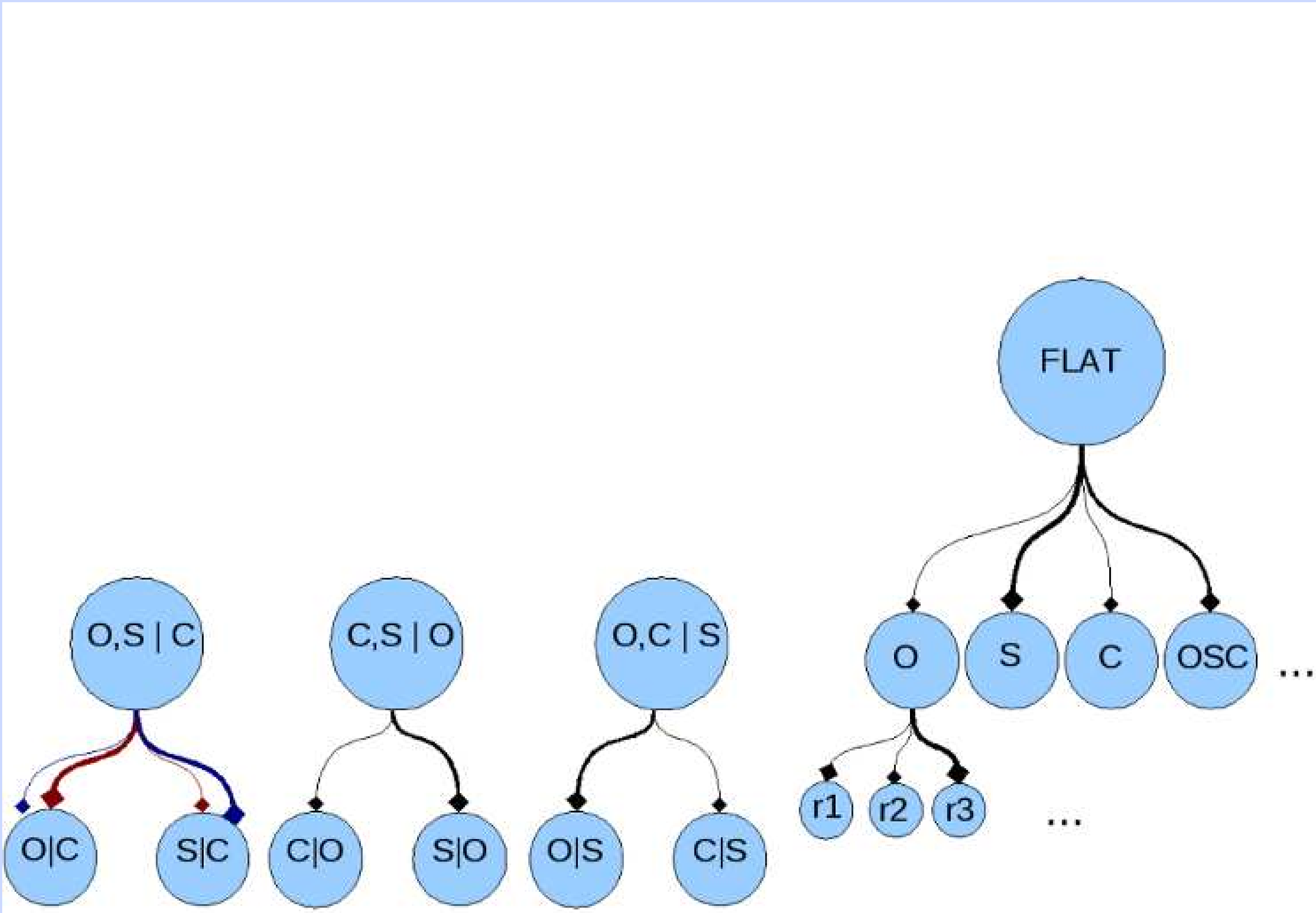

Softmax / Logistic

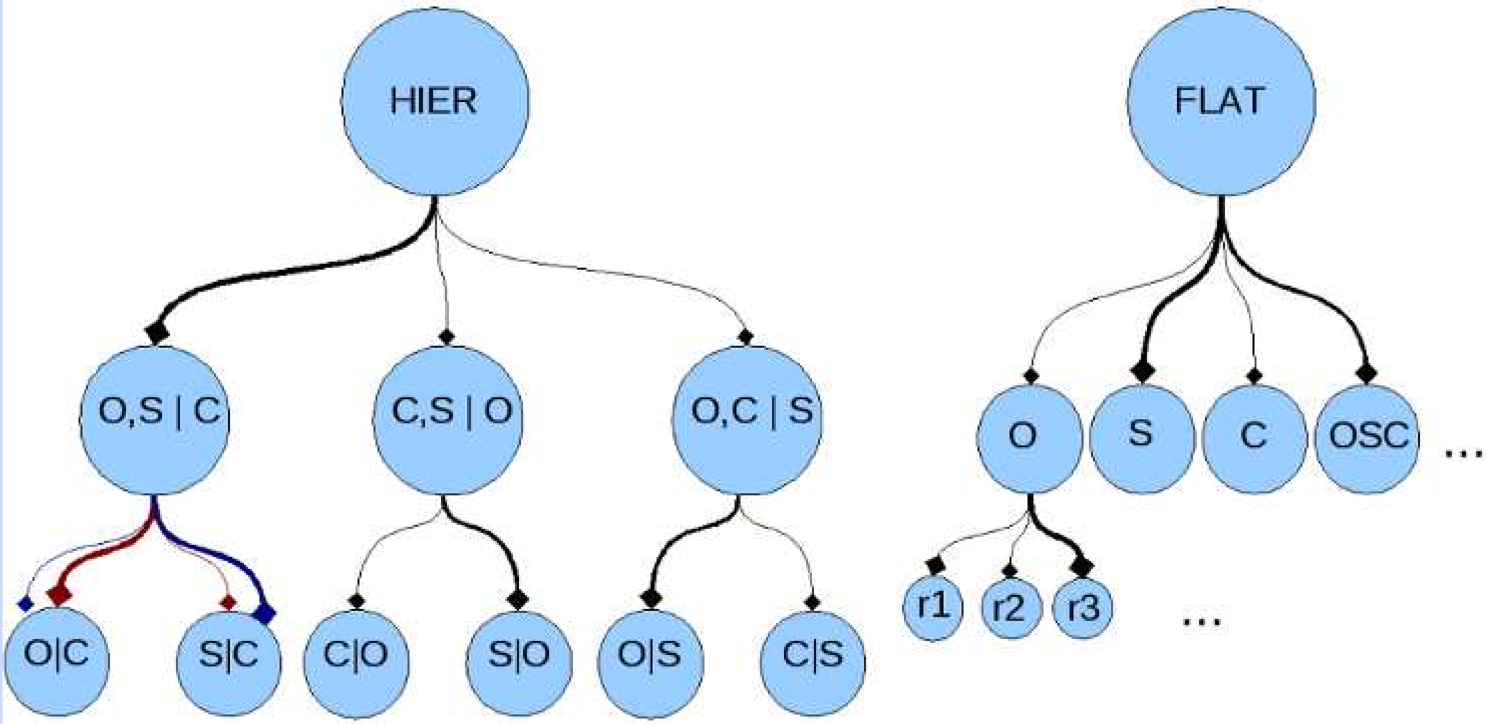

O
S
C
OSC
r1
r2
r3

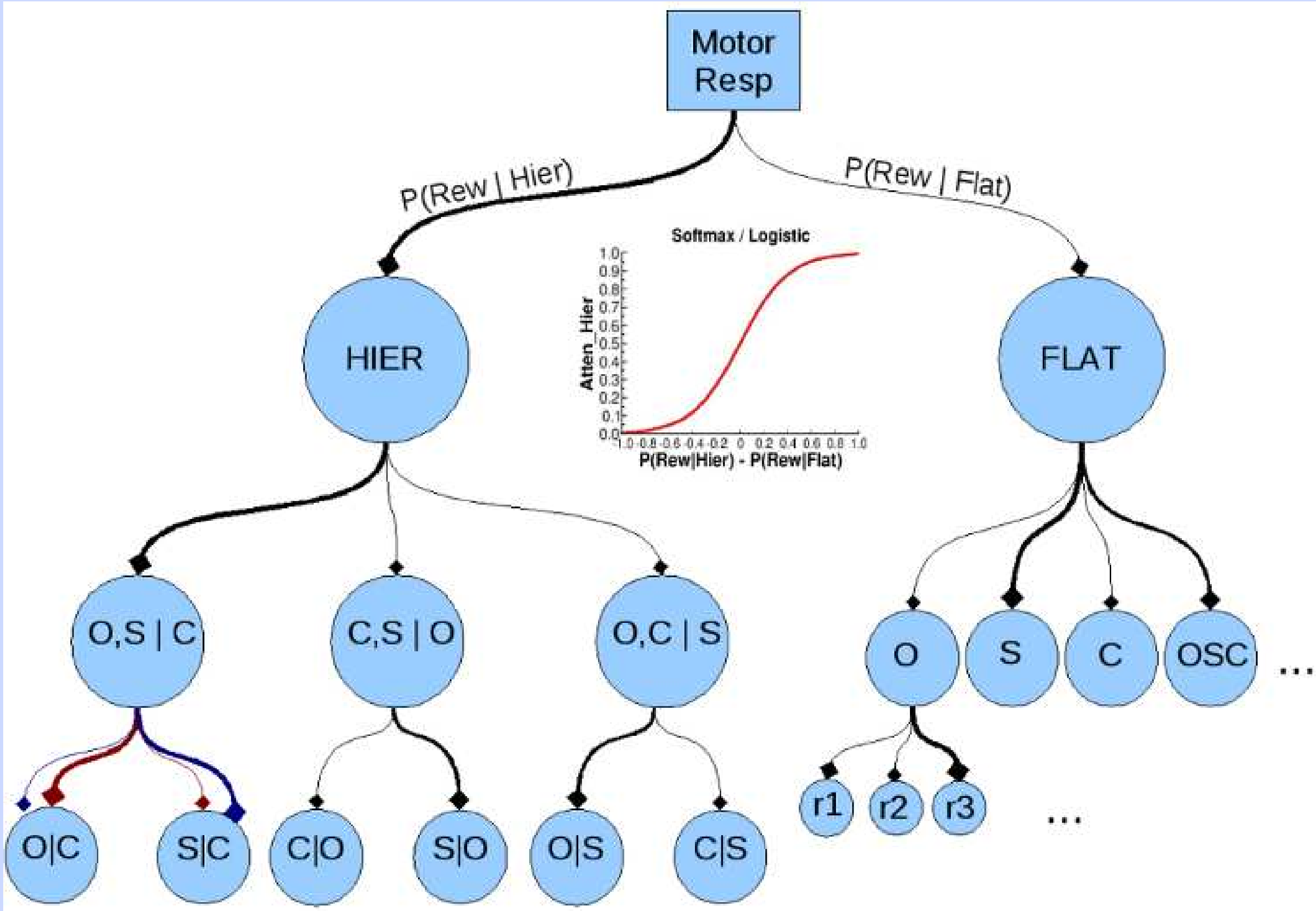


weight each expert in proportion to learned probability of expert success $P(Rew|e)\ \forall e$.

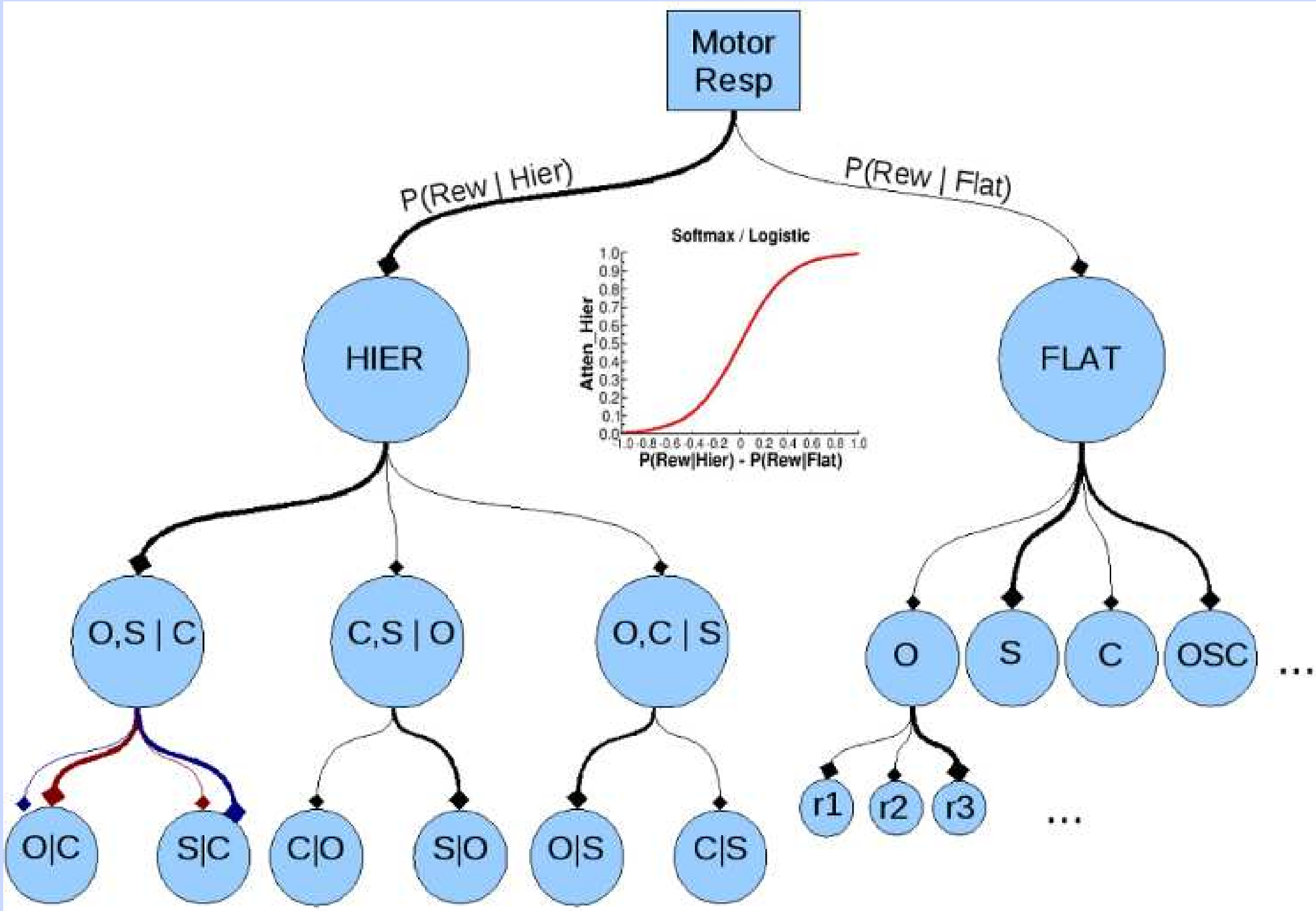

O,S | C
O|C
S|C
FLAT
O
S
C
OSC
...
r1
r2
r3
...


FLAT
O
S
C
OSC
...
r1
r2
r3
...
O,S | C
O|C
S|C
C,S | O
C|O
S|O
O,C | S
O|S
C|S

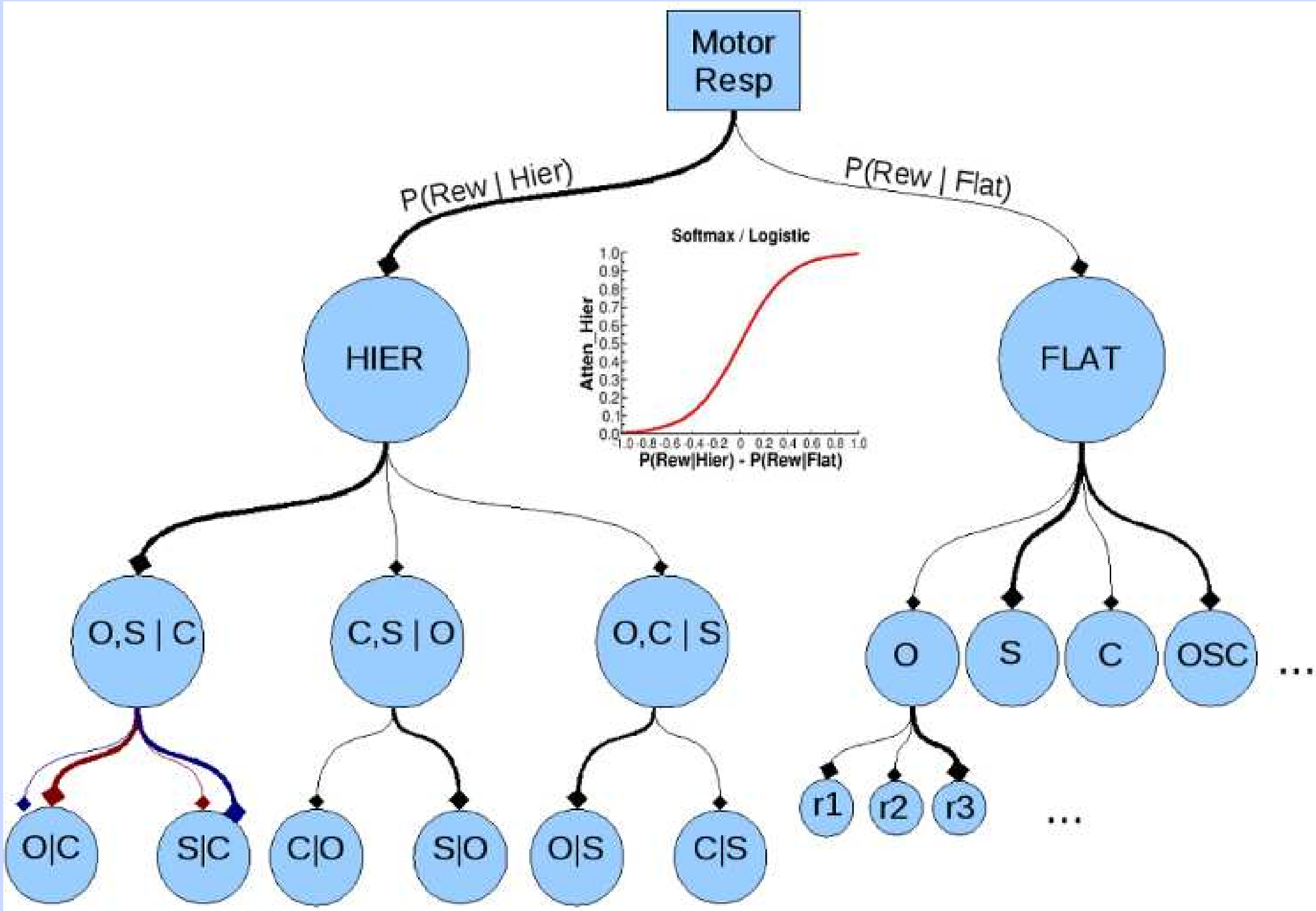

HIER
O,S | C
C,S | O
O,C | S
O|C
S|C
C|O
S|O
O|S
C|S
FLAT
O
S
C
OSC
...
r1
r2
r3
...



Motor Resp
P(Rew | Hier)
P(Rew | Flat)
HIER
FLAT
Softmax / Logistic
Atten_Hier
P(Rew|Hier) - P(Rew|Flat)
O,S | C
C,S | O
O,C | S
O|C
S|C
C|O
S|O
O|S
C|S
O
S
C
OSC
r1
r2
r3

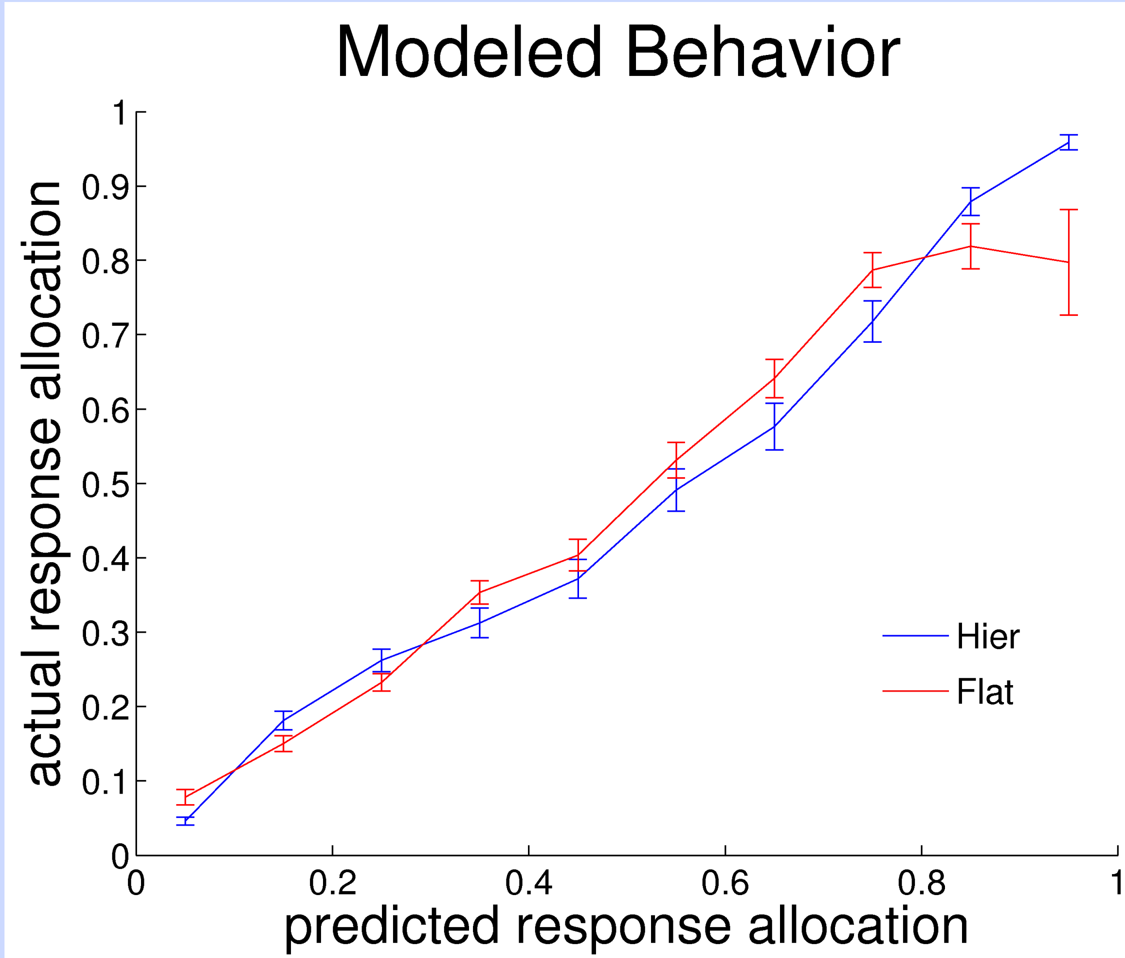


- Goal: infer latent hypotheses (attentional wts) by observing sequence of responses & rewards, maximizing likelihood of choices with few free params for each subject

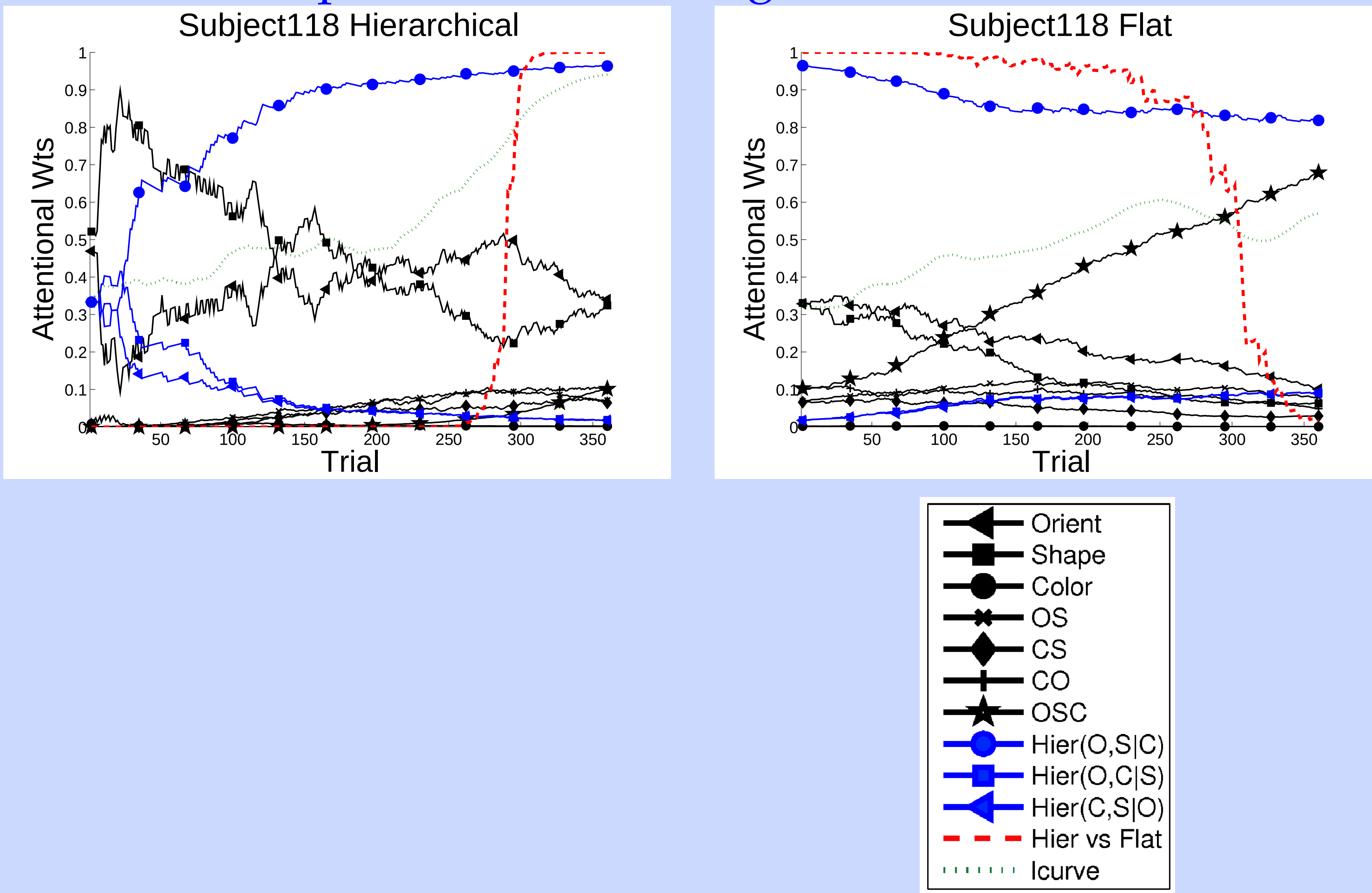


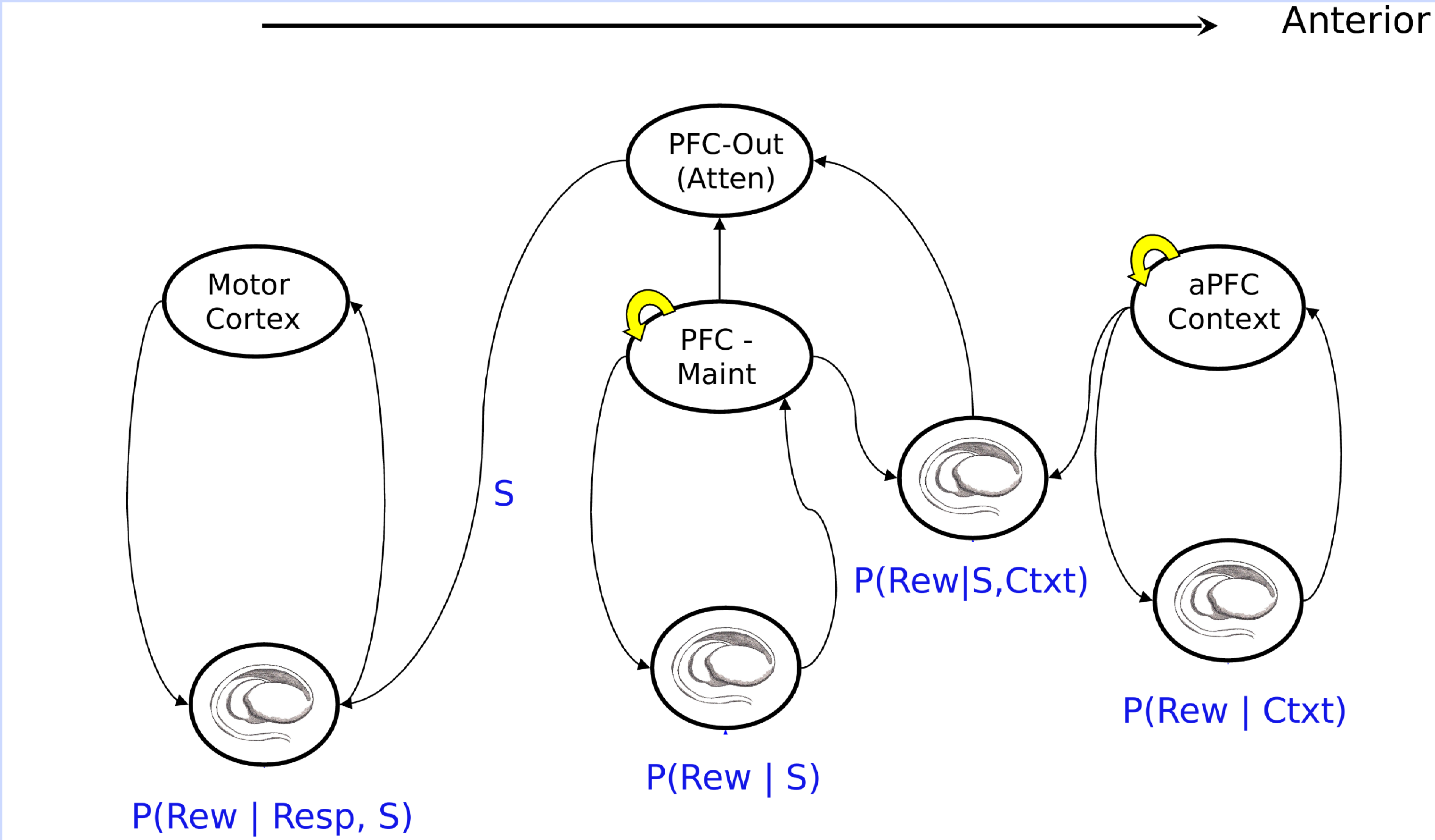
- Params: *priors* to attend to each dimension, to attend to conjunctions, to consider hierarchy. 3 softmax slopes (motor, within experts, between flat/hier).

MoE fits to humans

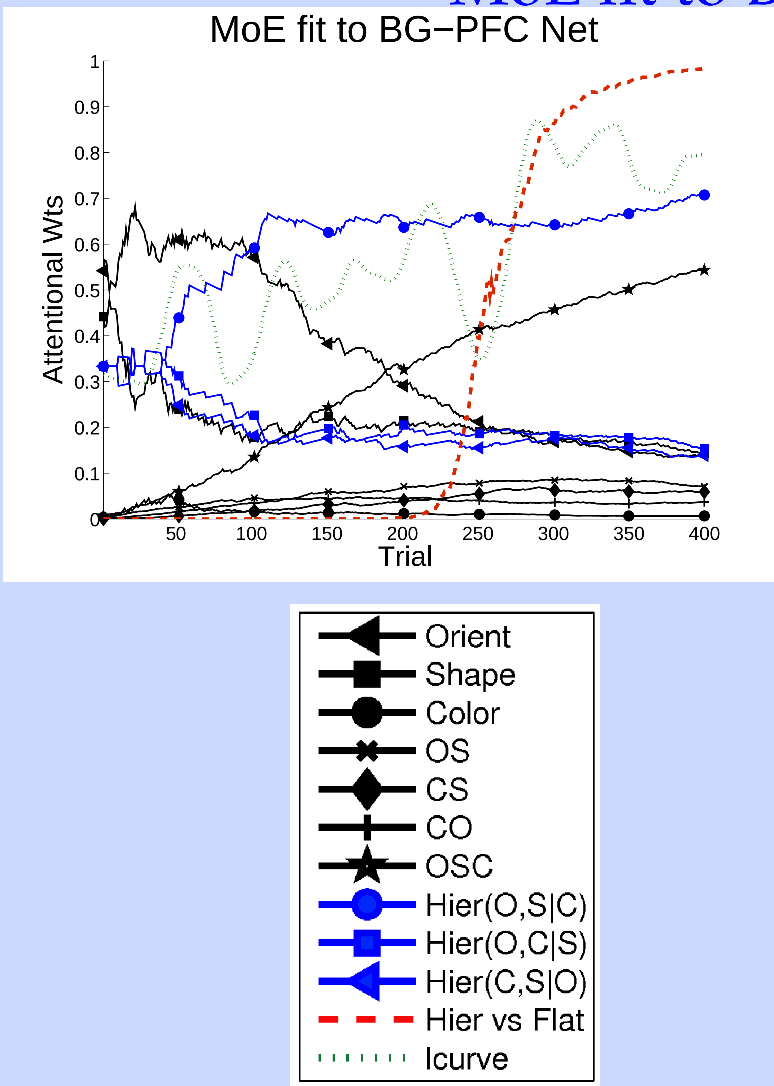


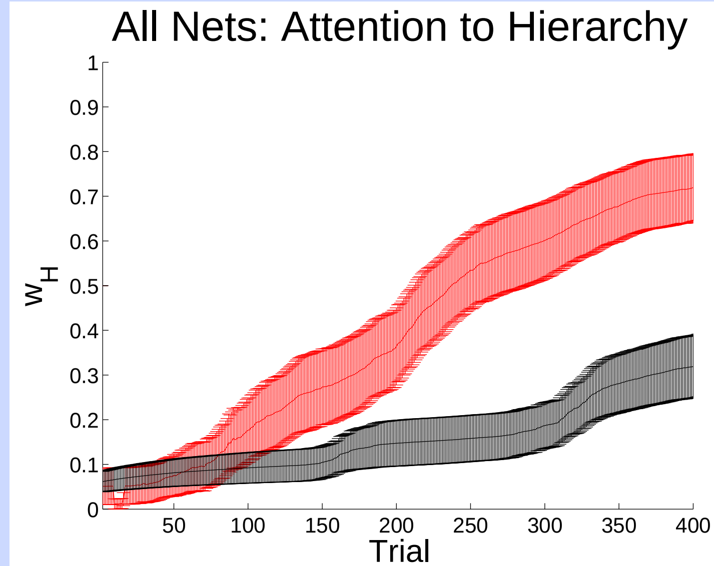
Example attentional weights: MoE fit to human



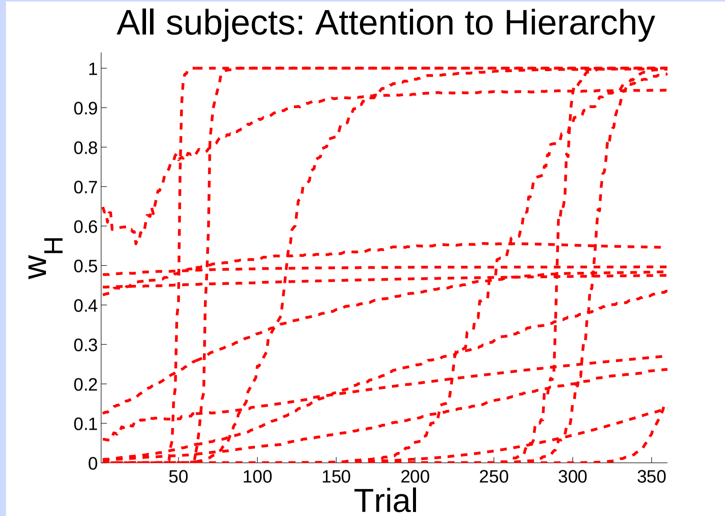

Anterior
PFC-Out
(Atten)
Motor
Cortex
PFC -
Maint
aPFC
Context
S
P(Rew|S,Ctxt)
P(Rew | Ctxt)
P(Rew | S)
P(Rew | Resp, S)

MoE fit to BG-PFC model

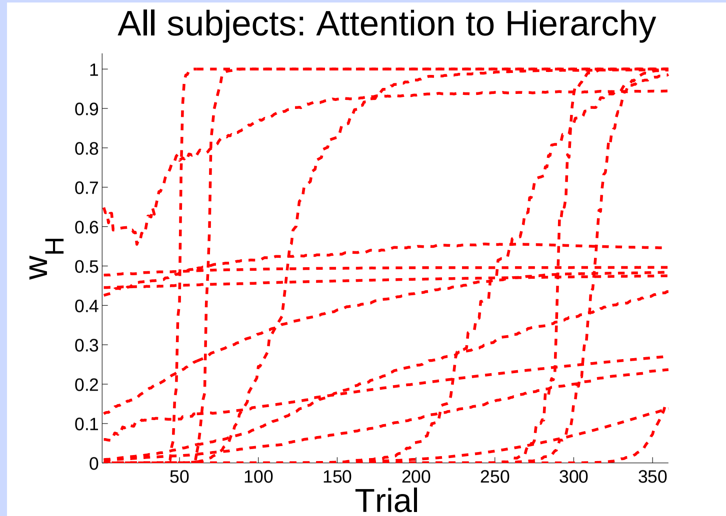


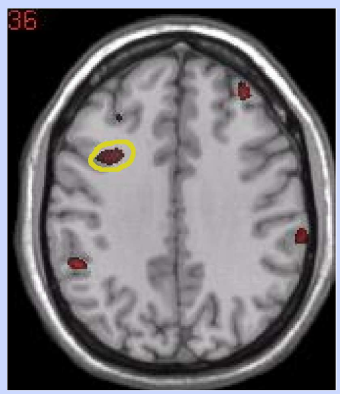


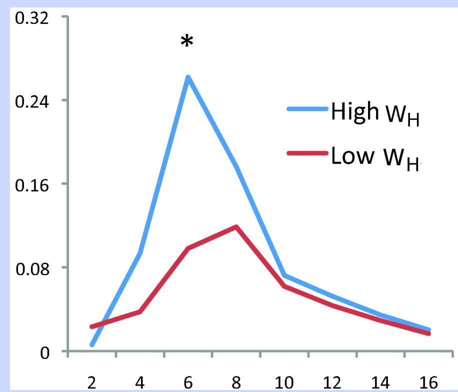
Individual differences in attention to Hierarchy



Individual differences in attention to Hierarchy

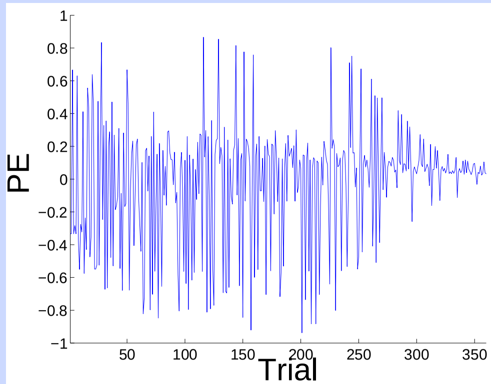






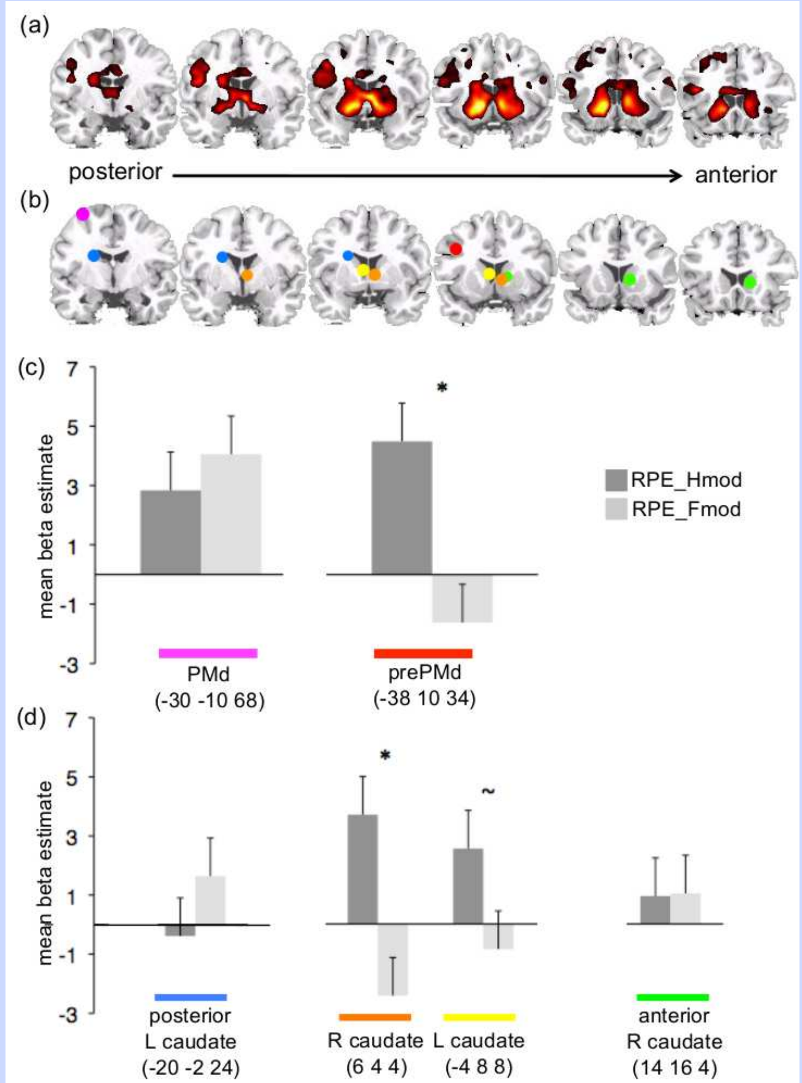
Median split on $Atten_H$ vs $Atten_F$ $\rightarrow$ more prePMD in Hier condition

Model-based fMRI: reward prediction errors

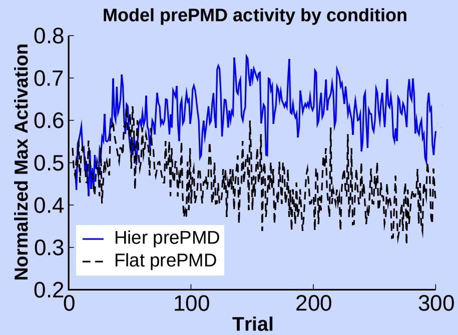


- Reward prediction errors tracked by entire striatum (cf. O'Doherty et al, 2004)

Prediction errors and attention to Hier vs Flat experts

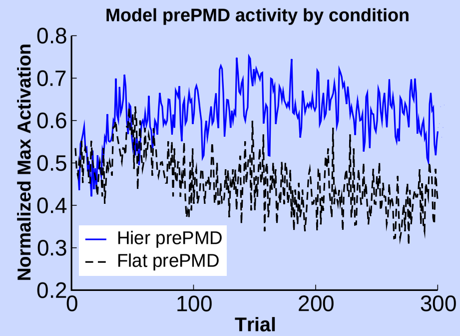


PE's predict PPMD decline in Flat condition

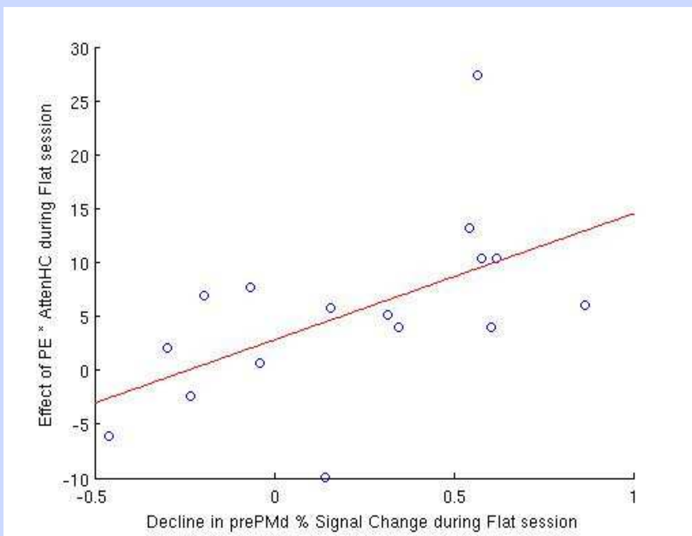


$\Rightarrow$ Decline in prePMD should be predicted by sensitivity of BG to PE's associated with hiearchical rule

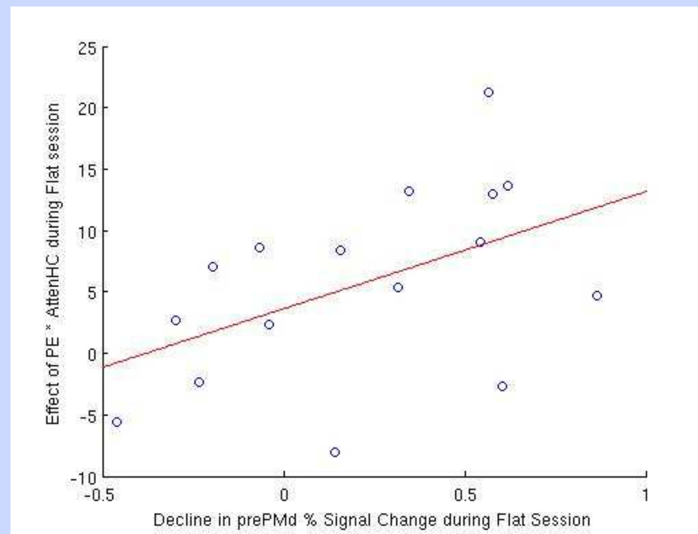
PE's predict PPMD decline in Flat condition



⇒ Decline in prePMD should be predicted by sensitivity of BG to PE's associated with hiearchical rule



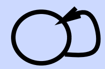
left caudate



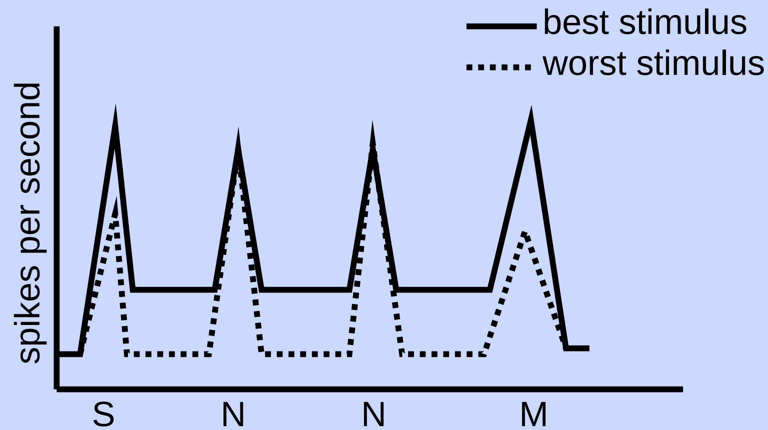
right caudate

Computational/Biological Details

Activations can be maintained in two different ways:

- Recurrent excitation: 
- Intrinsic bistability: selectively activated ion channels.

Problem with recurrent: strongest activations *at any point* determine what is maintained – *not necc. true*:



Computational/Biological Details

Our proposal:

- Thalamic disinhibition activates layer 4 FC.
- Convergence of layer 4 and cortico-cortical 2/3 projections on either 2/3 or 5/6 neurons triggers maintenance ion channels.

Other proposals:

- Thalmocortical loops themselves drive maintenance
(enough thalamic neurons to specify what is maintained?)
- Just recurrent excitation within FC
(unstable, but useful complement to ion channels).