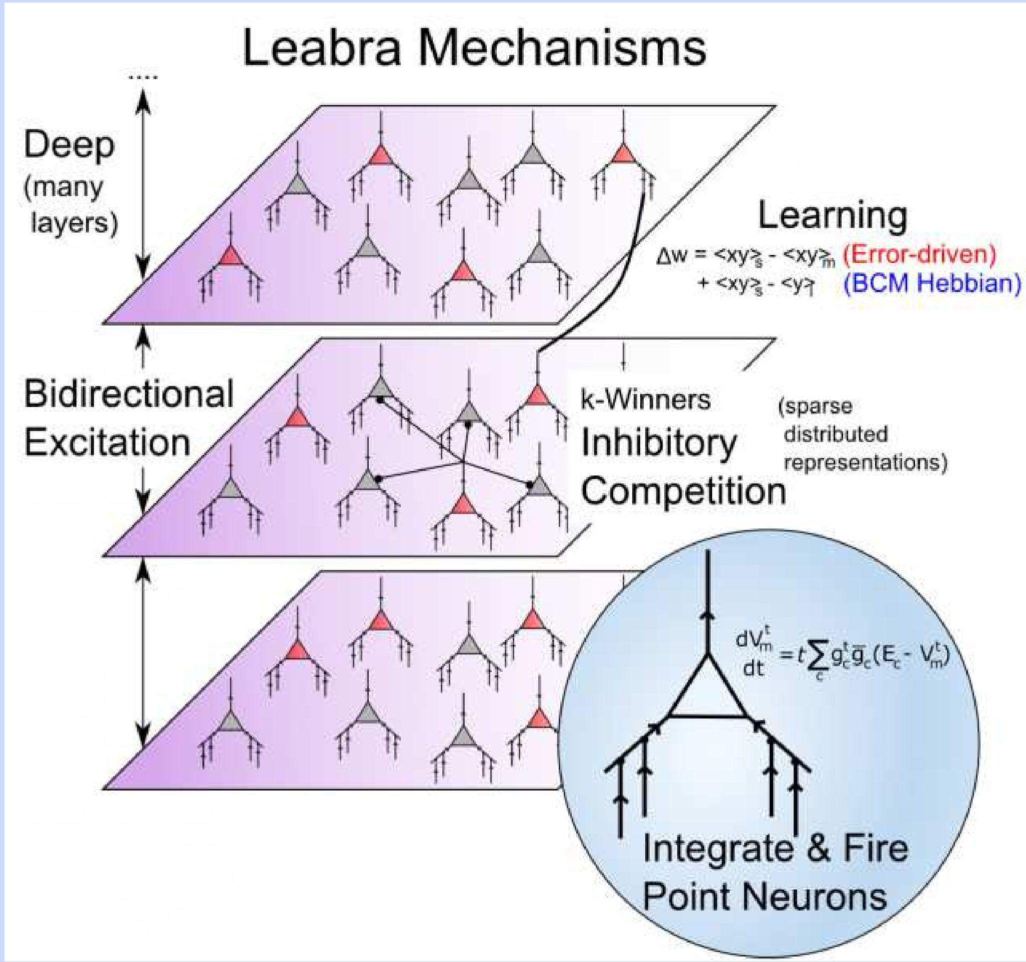
In Transition

- from Part I: Basic Mechanisms.
- to Part II: Perception, Attention, Memory, Language, Higher Level Cognition

Summary of Part I: Basic Mechanisms



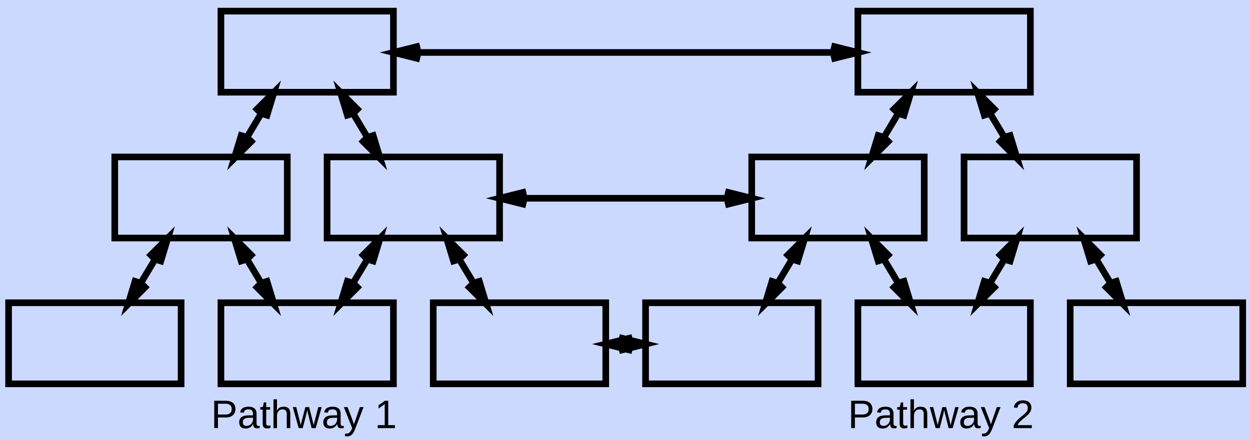
Micro and Macro-Neurocomputomics

Micro = basic mechanisms common across brain areas.

Macro = organization, differentiation, interactions of brain areas.

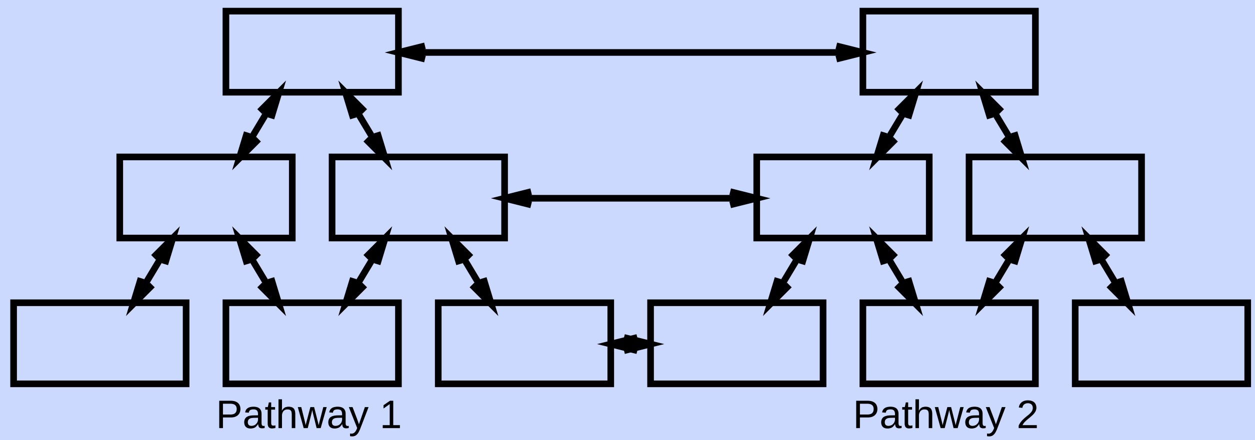
Need to consider general principles for macro organization before we can think about larger cognitive functions.

Macro Structural Principles



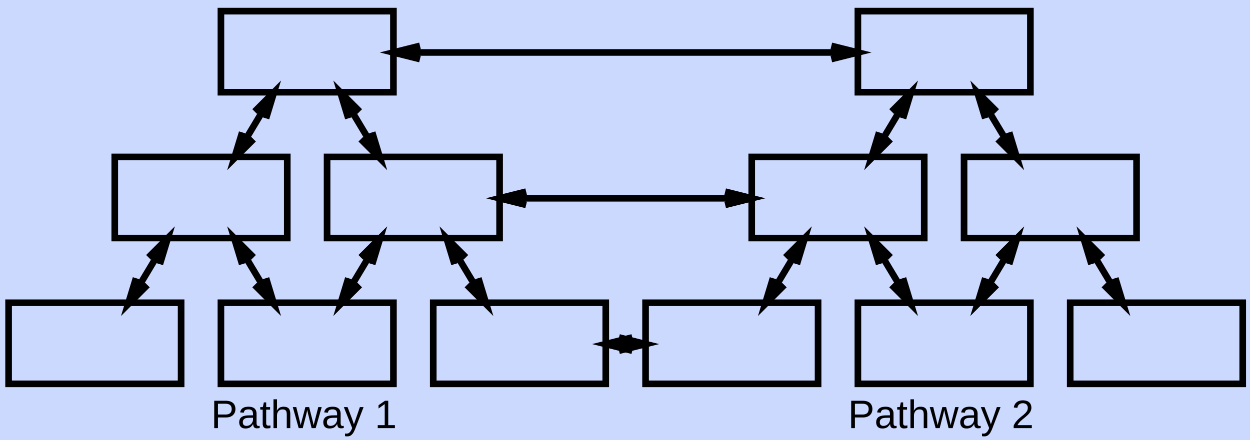
- Hierarchical sequence of transformations.
  - Emphasize some distinctions, ignore others
  - For object recognition you want to ignore differences in location, lighting, size, rotation
  - When reaching for objects, you want to emphasize location, size, and ignore object identity

Macro Structural Principles



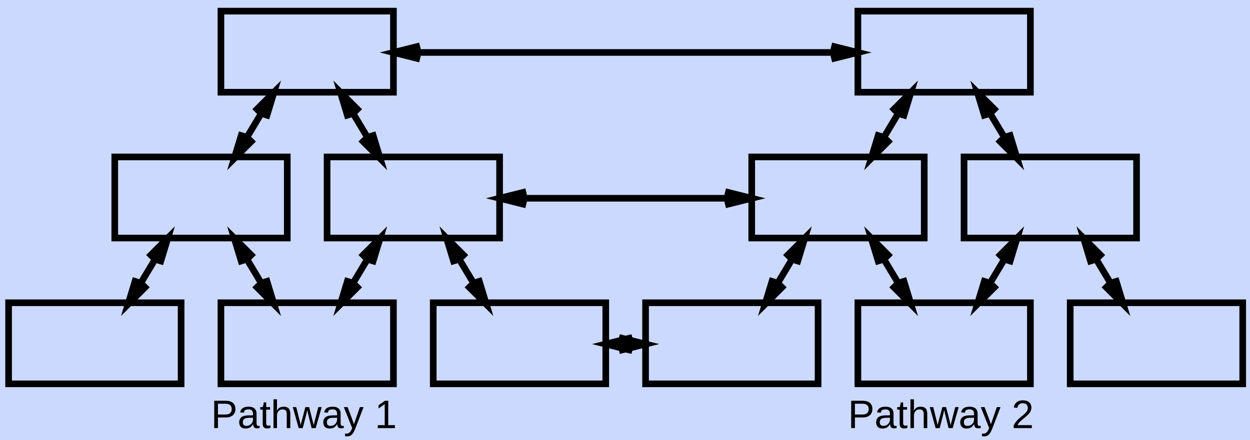
- Specialized pathways.
  - Location-invariant object recognition vs. recognizing orientation & location for actions (seeing for identifying and seeing for action)
  - patients with ventral stream damage have *blindsight* (e.g Milner & Goodale 1995): they can reach and grasp objects at different locations/orientations but cannot perceive them!

Macro Structural Principles



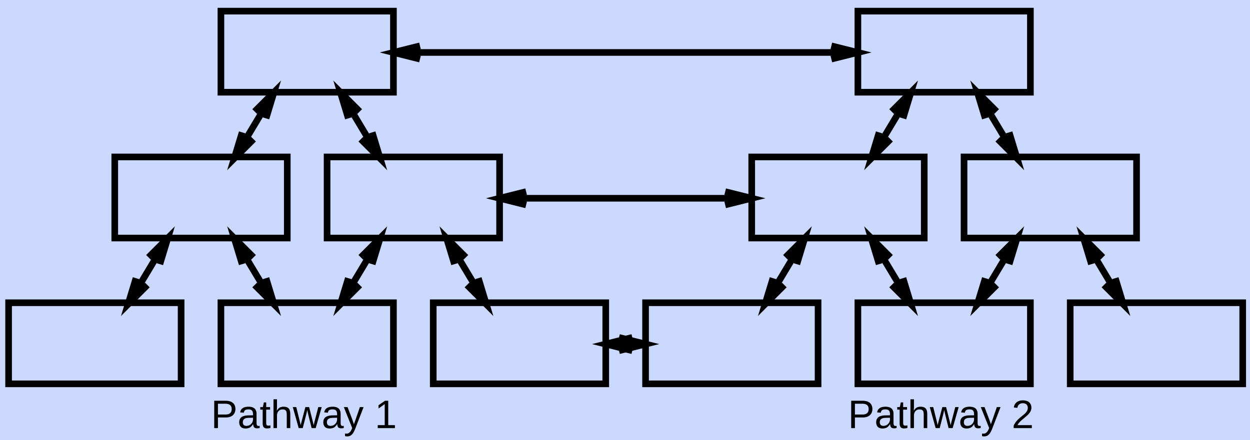
- Inter-pathway interactions.
  - Visual attention is an emergent property of interactions between object identification & spatial pathways

Macro Structural Principles



- Higher-level association areas
  - Integration of e.g., visual and auditory information
  - At extreme, thought to underlie *synesthesia*

Macro Structural Principles



- Large-scale Distributed Representations
  - Knowledge is distributed across multiple brain areas
  - Multiple areas participate in representing a given thing (e.g., apple)
  - Each area represents multiple things
  - Same idea as distributed representation among units for individual items, but just now across multiple areas/modalities, etc

Macro Dynamic Principles

- Processing as **multiple constraint satisfaction**
- Attractors, settling dynamics, amplification: active memory
- Inhibitory competition: attention.
- Where do constraints come from?

Macro Dynamic Principles

- Where do constraints come from?
  - perceptual inputs ("bottom-up" constraints)
  - Also, we have the ability to maintain firing of neurons even in the absence of bottom-up stimulation
  - Make use of bidirectional excitatory connections
    $\rightarrow$ *Active memory* – constitutes an *inner mental context*

Macro Dynamic Principles

- Where do constraints come from?
  - perceptual inputs (“bottom-up” constraints)
  - Also, we have the ability to maintain firing of neurons even in the absence of bottom-up stimulation
  - Make use of bidirectional excitatory connections
    $\rightarrow$ *Active memory* – constitutes an *inner mental context*

She swam from the overturned canoe to the bank.
She walked from the post office to the bank.

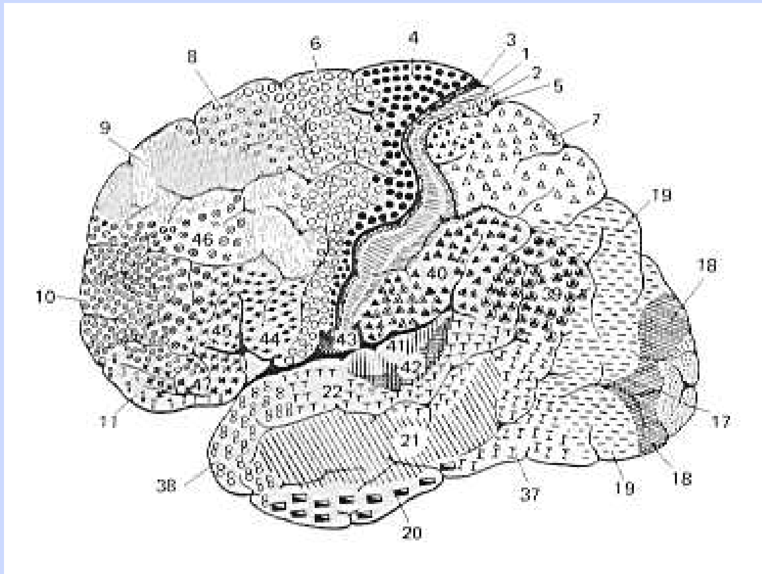
Macro Dynamic Principles

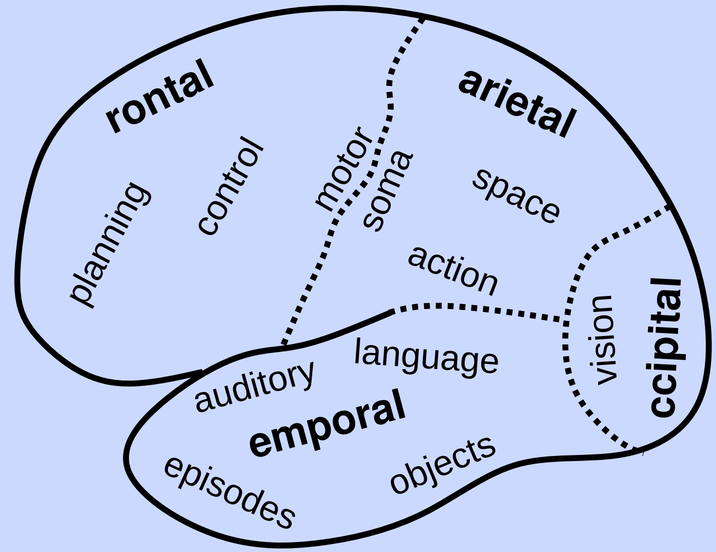
- Where do constraints come from?
  - perceptual inputs ("bottom-up" constraints)
  - Also, we have the ability to maintain firing of neurons even in the absence of bottom-up stimulation
  - Make use of bidirectional excitatory connections
    $\rightarrow$ *Active memory* – constitutes an *inner mental context*

She swam from the overturned canoe to the bank.
She walked from the post office to the bank.

Active memory can pertain to concrete stimulus representations as well as more abstract things..

General Functions of the Cortical Lobes

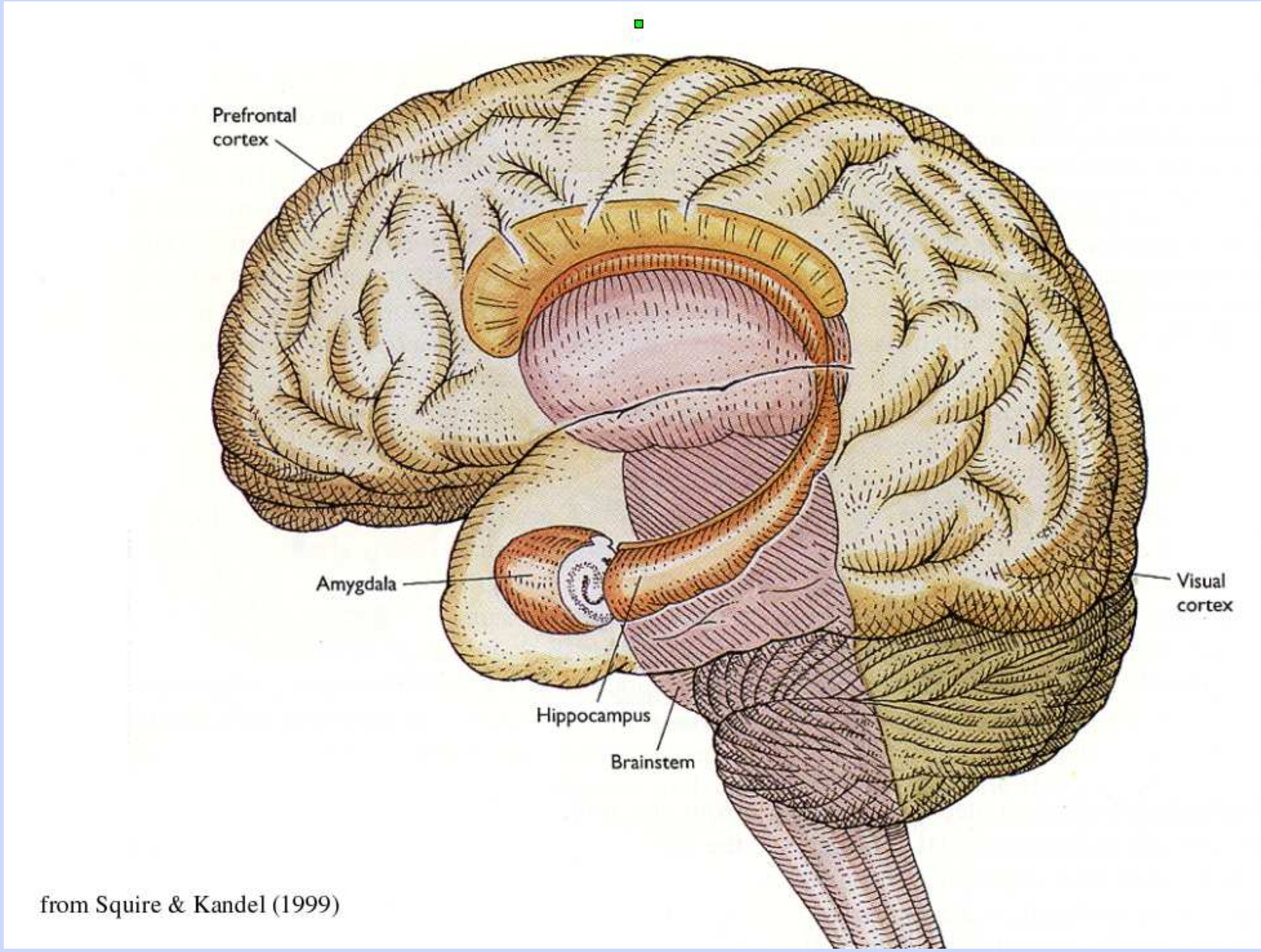




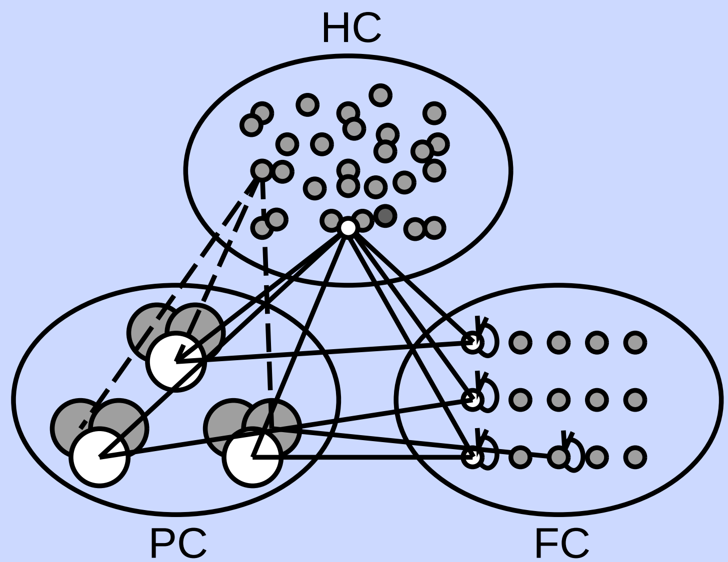
- Occipital lobe: vision
- Temporal lobe: hearing, speech perception, object recognition...
- Parietal lobe: representing body & external spaces
- Frontal lobe: Motor control, cognitive control (planning, working memory, etc)

Other Areas

- Hippocampus (rapid episodic encoding).
- Thalamus (sensory input, attention).
- Amygdala (emotion, affective associations).
- Basal ganglia (BG) (motor control, sequencing, reward learning, gating of PFC...).
- Cerebellum (motor learning, forward model? cognitive role via timing?).
- Midbrain neuromods: VTA - dopamine, raphe - serotonin, locus coeruleus - norepinephrine.


Prefrontal
cortex
Amygdala
Hippocampus
Brainstem
Visual
cortex
from Squire & Kandel (1999)

Tripartite Functional Organization



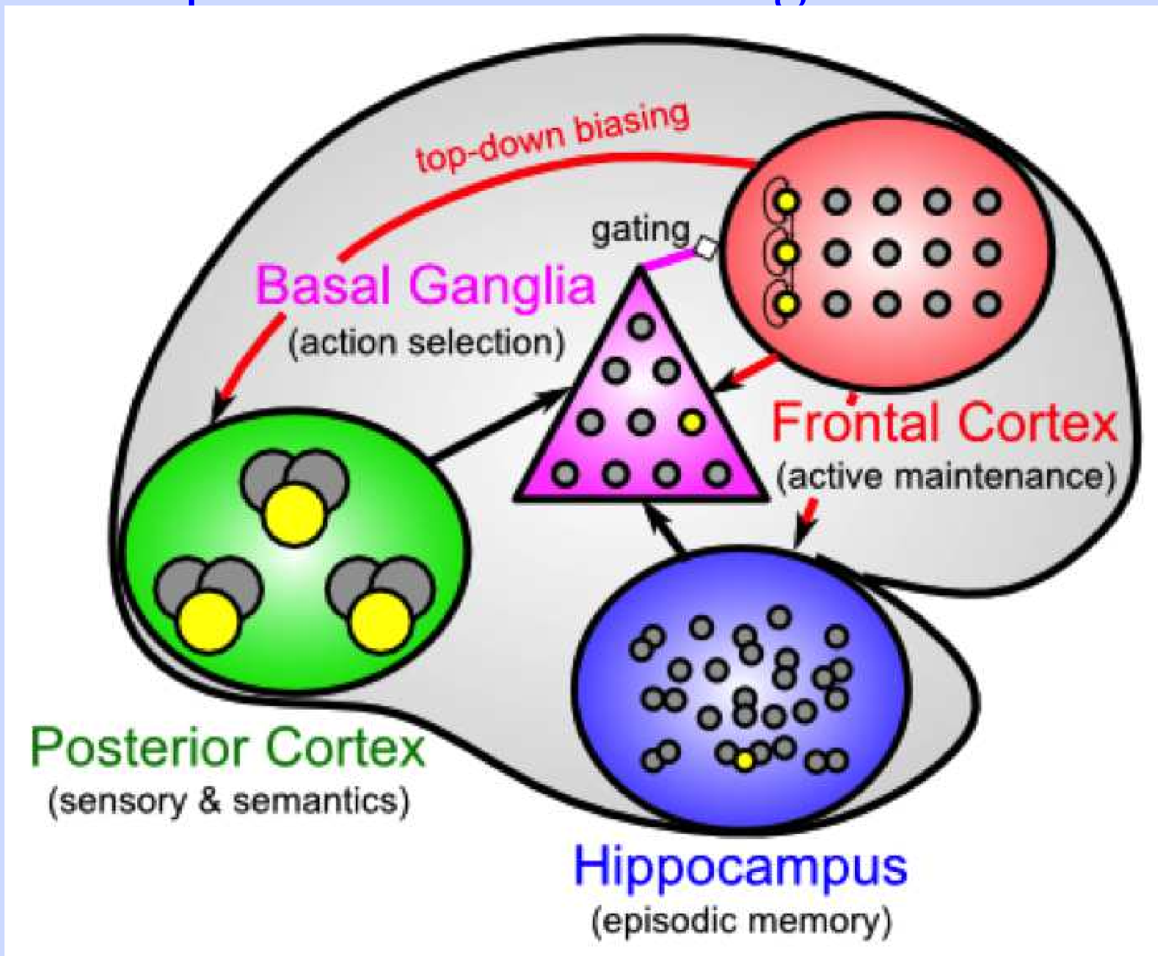
PC = posterior perceptual cortex: *slow integrative learning*

HC = hippocampus and related structures: *rapid memorization*

FC = prefrontal cortex: *active maintenance ("working memory")*

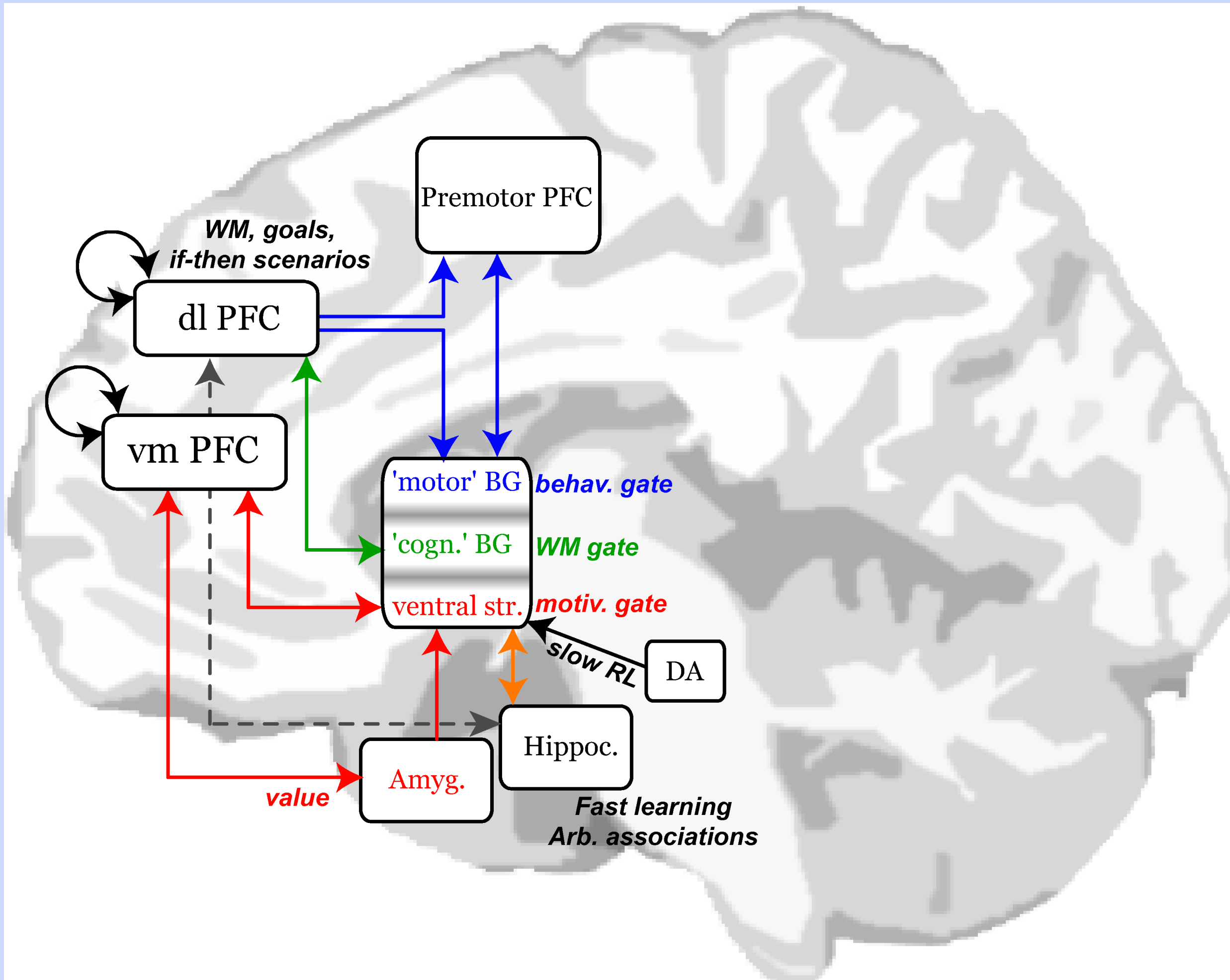
Defined by set of functional *trade-offs*.

Tripartite Functional Organization



Defined by set of functional *trade-offs*.

Multiple systems in decision making



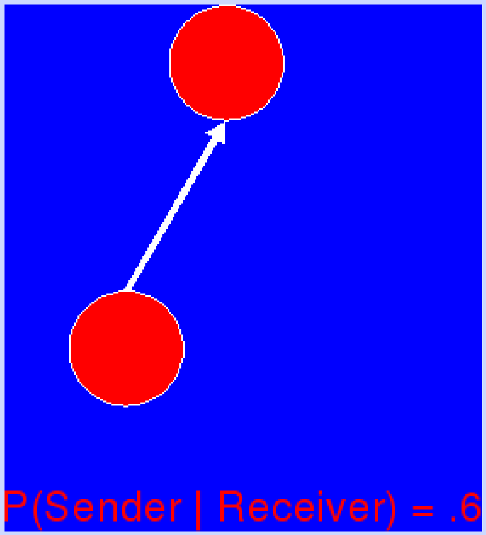
Computational Trade-offs in Learning & Memory

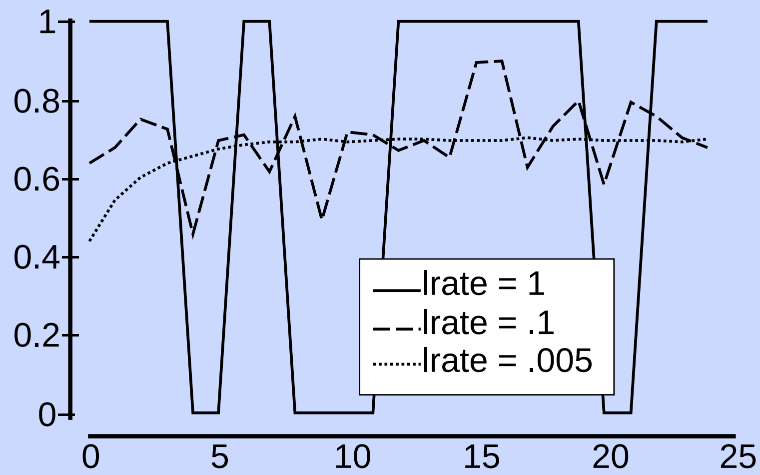
Trade-offs: *Computational objectives that are mutually incompatible and thus cannot be achieved by a single brain system*.

→ Begin to address psychological distinctions between different learning & memory processes, informed by mechanisms required.

- Learning *statistical structure* vs. *memorizing specific events*
- *Isolated maintenance* (holding in mind multiple items of info) vs. *inference (spreading activation: smoke→fire)*
- *Robust maintenance* vs. *rapid updating*

1. Slow *vs* Fast Learning



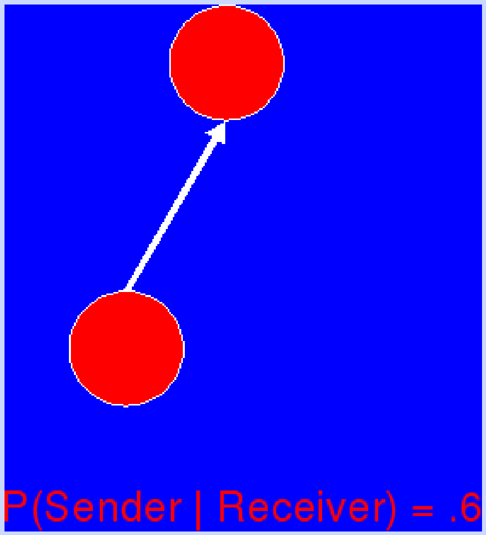


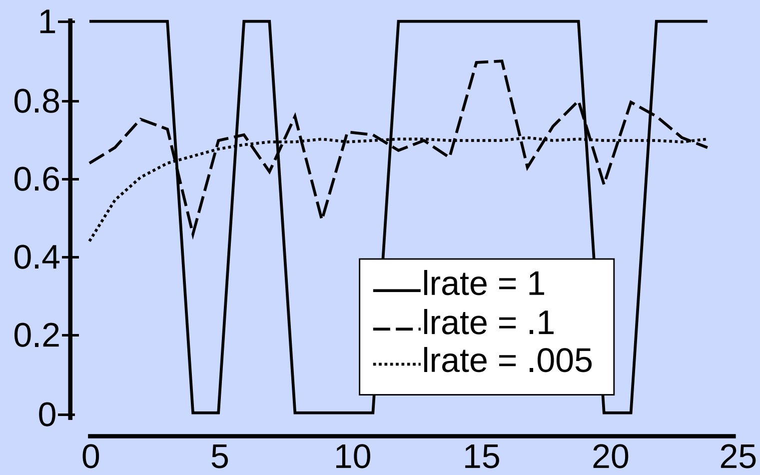
Learning must be *slow* to capture (statistical) structure (averaging).

But you also have to be able to learn rapidly.

Tradeoff solved by 2 systems: cortex learns slowly, hippo rapidly.

1. Slow *vs* Fast Learning





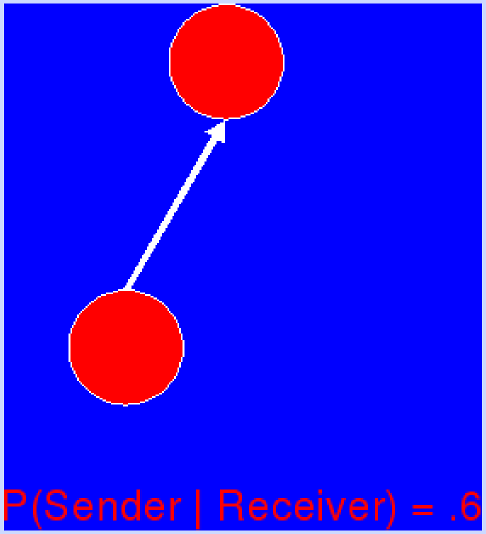
Learning must be *slow* to capture (statistical) structure (averaging).

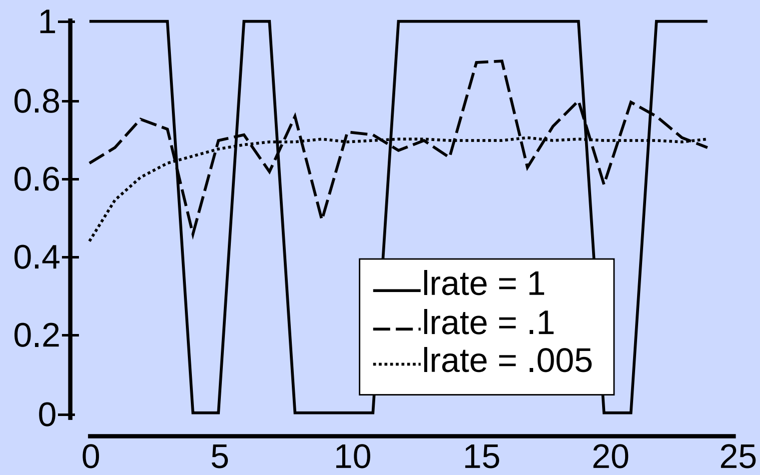
But you also have to be able to learn rapidly.

Tradeoff solved by 2 systems: cortex learns slowly, hippo rapidly.

3rd system: Active memory (prefrontal cortex) $\approx$ fastest (immediately accessible)

1. Slow *vs* Fast Learning





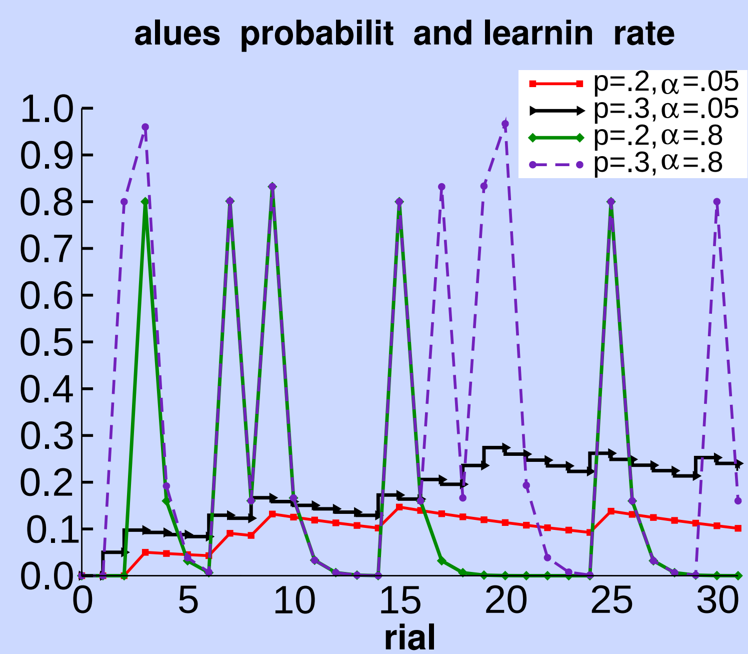
Learning must be *slow* to capture (statistical) structure (averaging).

But you also have to be able to learn rapidly.

Tradeoff solved by 2 systems: cortex learns slowly, hippo rapidly.

3rd system: Active memory (prefrontal cortex) $\approx$ fastest (immediately accessible)
but learning to develop pfc reps in first place is slow, allows abstraction.

1b. Slow *vs* Fast [Reinforcement] Learning

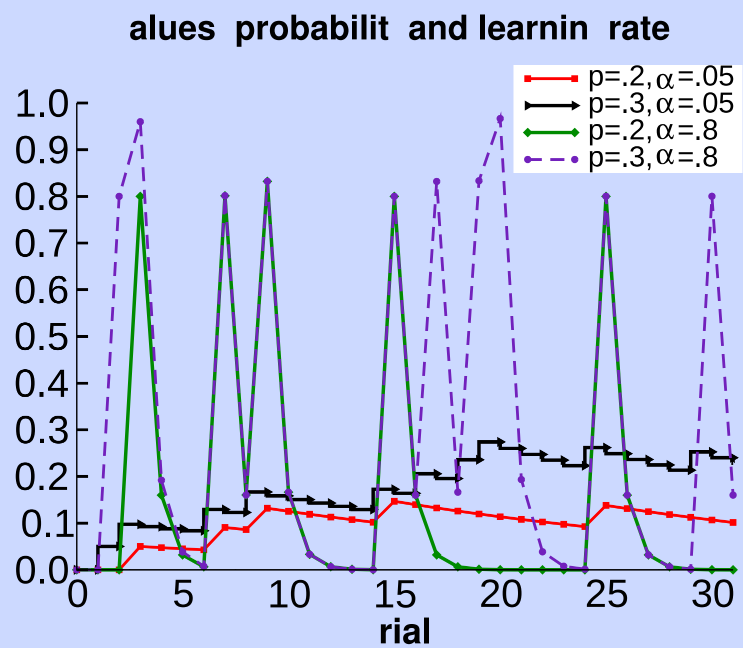


[Reinforcement] Learning must be *slow* to capture best actions that work on average.

But you also have to be able to sensitive to rapid changes in value (e.g., stock market).

Tradeoff solved by 2 systems:
BG learns slowly, PFC flexibly updates new states and can override habitual choices.

1b. Slow *vs* Fast [Reinforcement] Learning



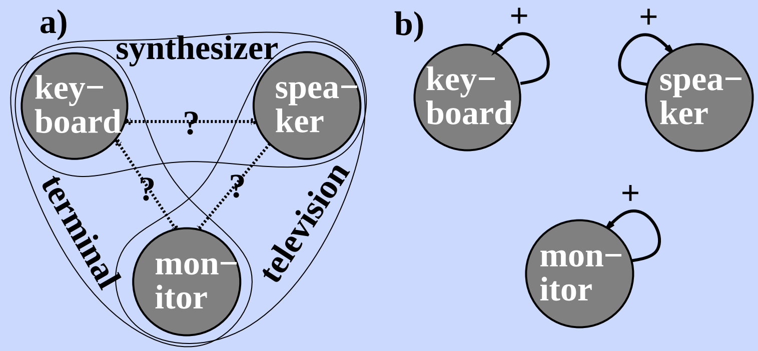
[Reinforcement] Learning must be *slow* to capture best actions that work on average.

But you also have to be able to sensitive to rapid changes in value (e.g., stock market).

Tradeoff solved by 2 systems:
BG learns slowly, PFC flexibly updates new states and can override habitual choices.

→ lots of evidence for differential BG and PFC contributions to habitual and rapid action-outcome learning, across species, methods.

2. Active Memory *vs* Overlapping Distributed Reps



Overlapping distributed representations are useful for capturing information about the world.

But overlap & interconnectivity cause spread, which is not useful for maintaining *specific* information over time.

Tradeoff solved by two systems: PC has overlapping distributed representations, FC is isolated for maintenance.

3. Active Memory: Another Trade-off

Active memory needs specialized updating & maintenance mechs.

Protecting representations from interference (robust maintenance of working memory) vs. being receptive to update important, unexpected information

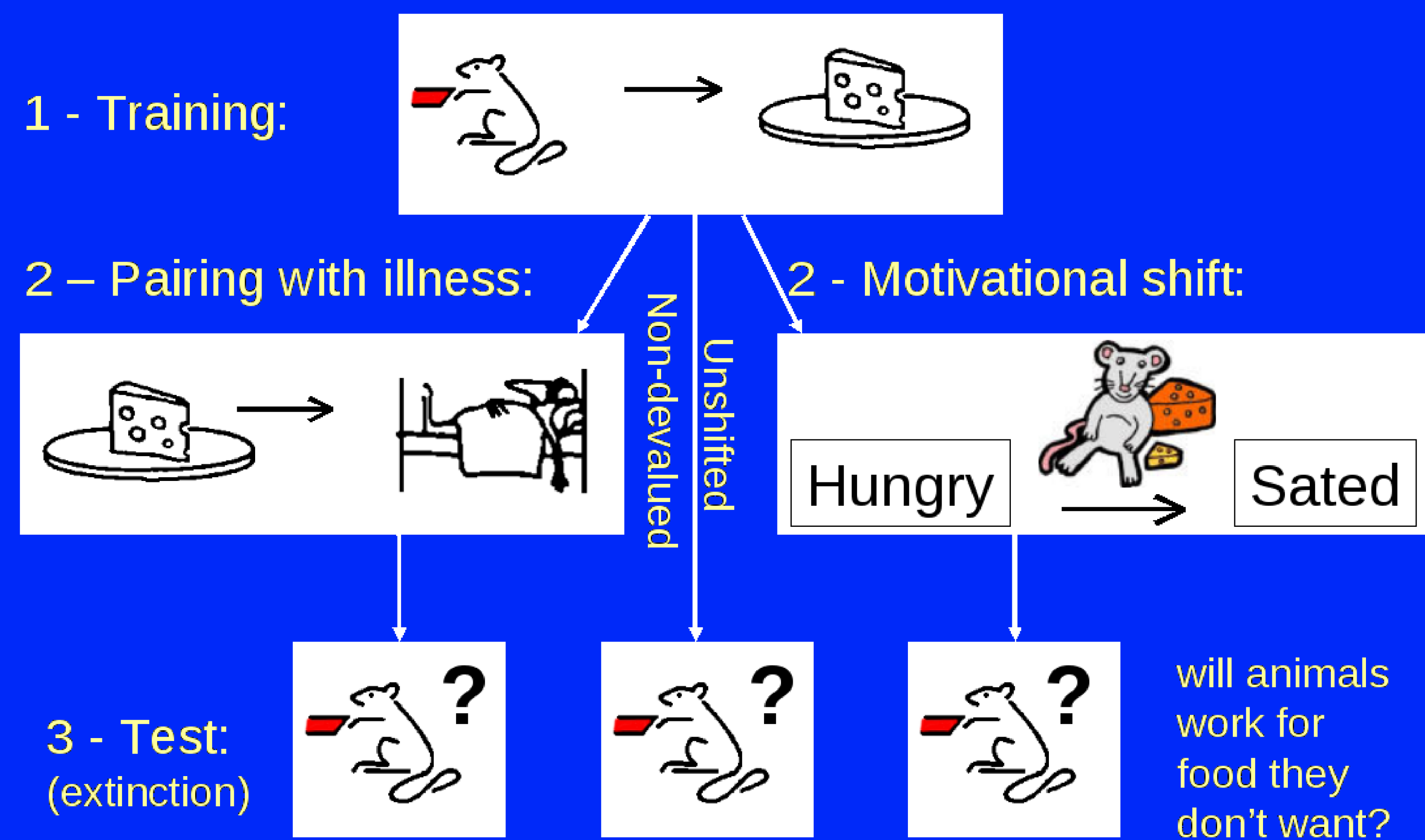
Basal ganglia may contribute to this updating function

4. Model-Based vs. Model-Free RL (not in text)

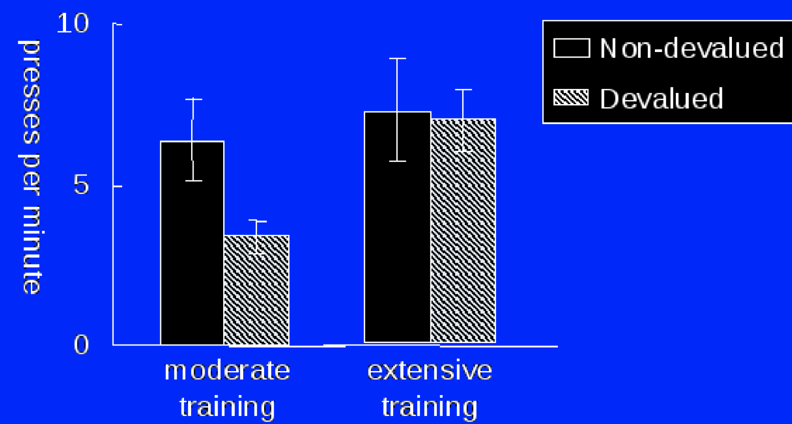
- *Model-free: (Habits)*
  - Incrementally learn to associate stimuli (states) and actions with value, using only (DA-based) reward prediction errors to update values (TD learning and variants thereof; BG model). Then just select action with highest *"Q value"* (or Go-NoGo value).

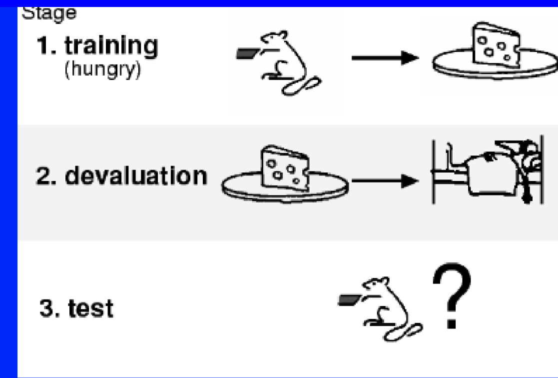
- *Model-based: (Cognitive)*
  - Actually represent the environment ("world-model") and predicted transition from one state to another, and how these are affected by our (and others') actions....

Devaluation experiment



Behavioral results





outcome-sensitive
"goal-directed"
like Tolman

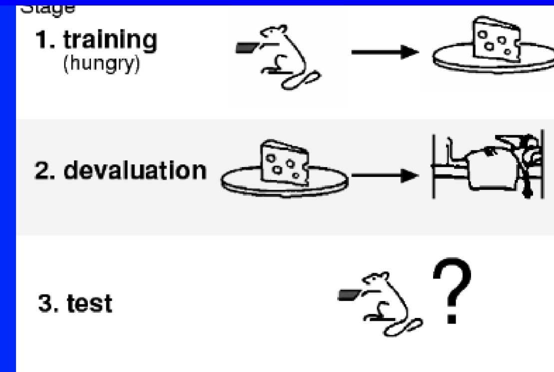
outcome-insensitive
"habitual"
like S-R (Thorndike)

Animals will *sometimes* work for food they don't want!

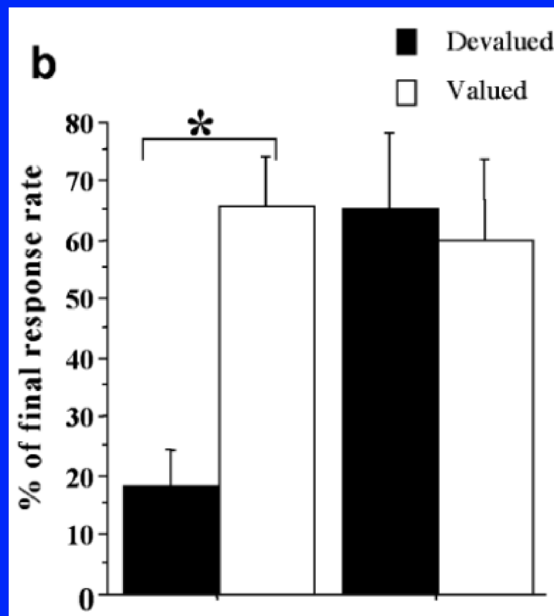
→ in daily life: actions become automatic with repetition

Holland (2004)

Lesion results I



overtrained rats



dorsolateral striatum lesion

control (sham lesion)

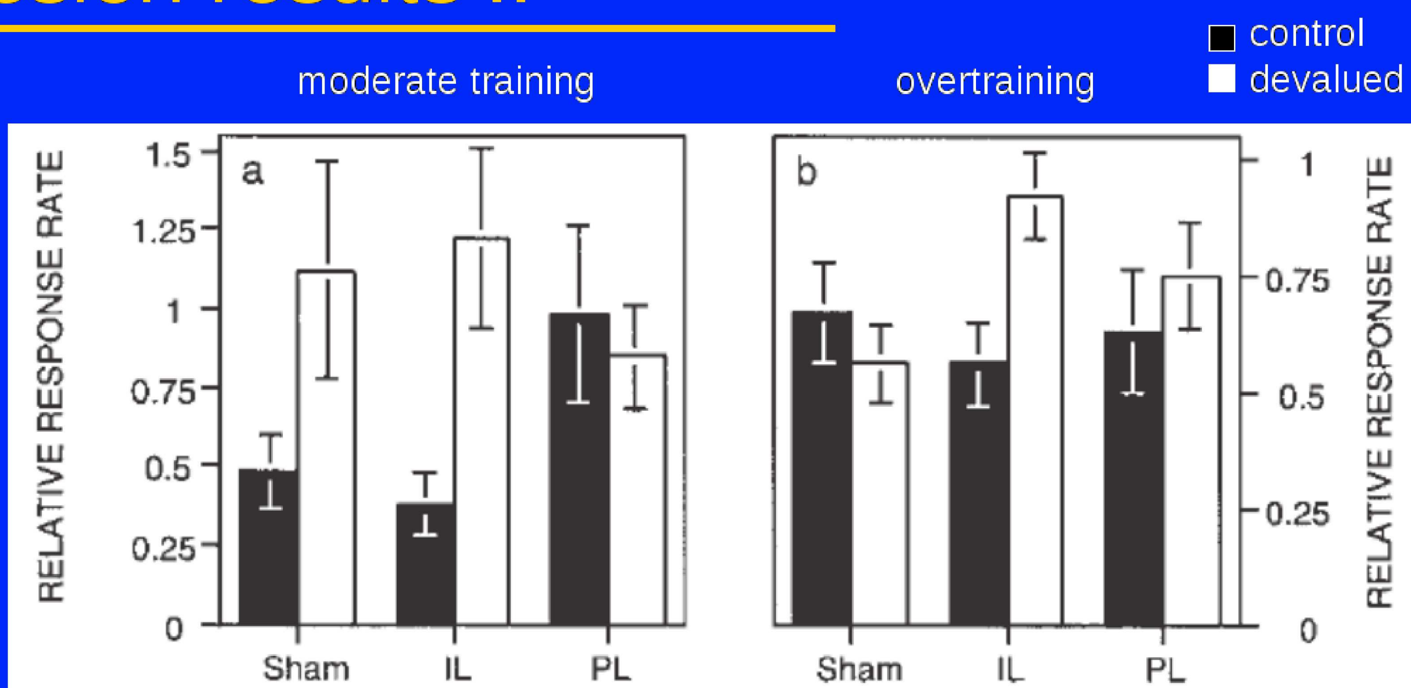
animals with lesions to dorsolateral striatum (BG) never develop habits despite extensive training

→ also treatments depleting dopamine

→ also lesions to infralimbic division of PFC (same corticostriatal loop)

Yin et al (2004)

Lesion results II



Killcross & Coutureau (2003)

Prelimbic (PL) PFC lesions cause animals to leverpress habitually even with only moderate training

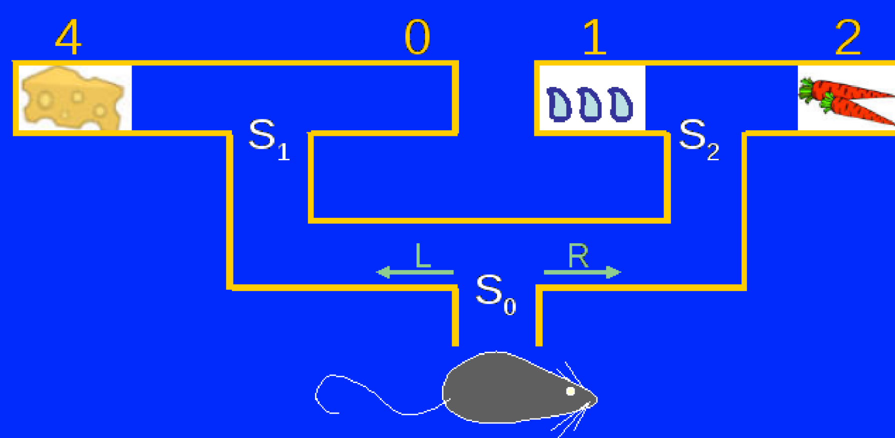
→ also dorsomedial PFC and mediodorsal thalamus (same loop)

→ double dissociation with IL PFC

What do these findings tell us?

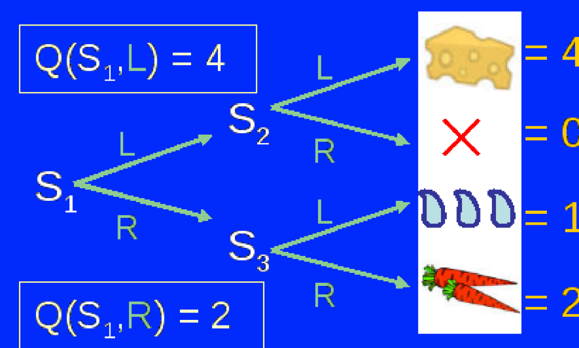
- The same action (lever-pressing) can arise from two psychologically & neurally **dissociable** pathways
  - moderately trained behavior is ***goal-directed***: dependent on outcome representation of what might happern
  - overtrained behavior is ***habitual***: apparently not dependent on outcome, like S-R learning
- S-R habits really exist (in humans too), they just don't describe all of behavior
- Lesions suggest ***two parallel systems***, in that the intact one can apparently support behavior at any stage. (see also BG vs Hippo in S-R vs cognitive map)

Strategy I: Model-based RL

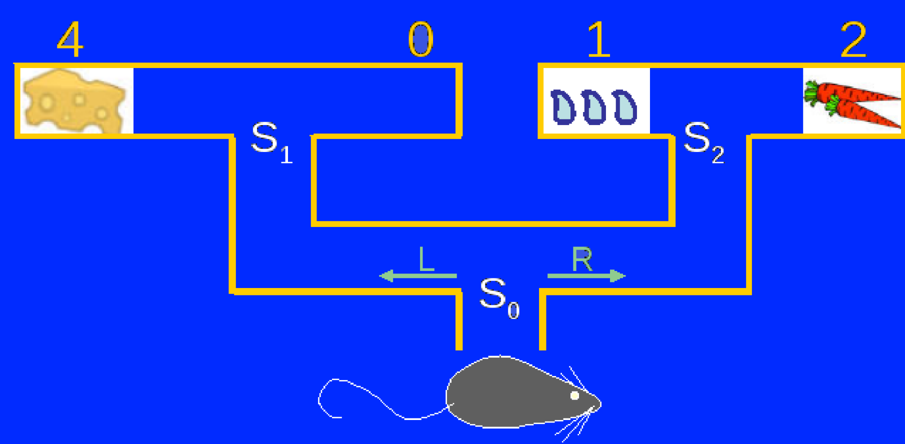


learn model of task through experience (= cognitive map)

choosing actions is hard (need to compute Q values by iterative search through map)

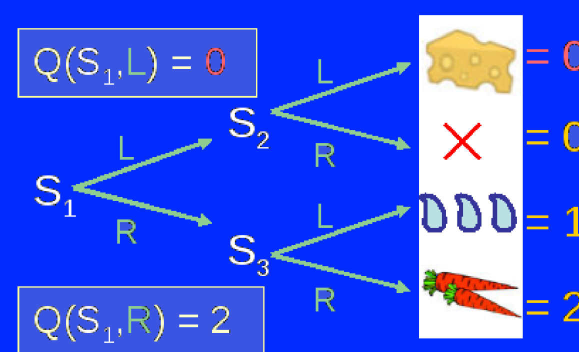


Strategy I: Model-based RL

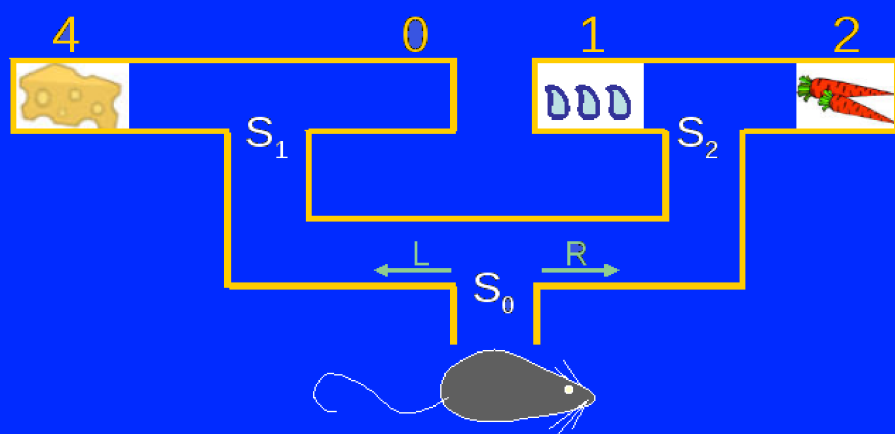


learn model of task through experience (= cognitive map)

but flexible, efficient representation (recompute values online)



Strategy II: Model-free RL



- Shortcut: store long-term values
  - then simply retrieve them to choose action
- Can learn these from experience
  - without building or searching a model
  - incremental “sampling” and prediction errors
  - dopamine dependent TD learning

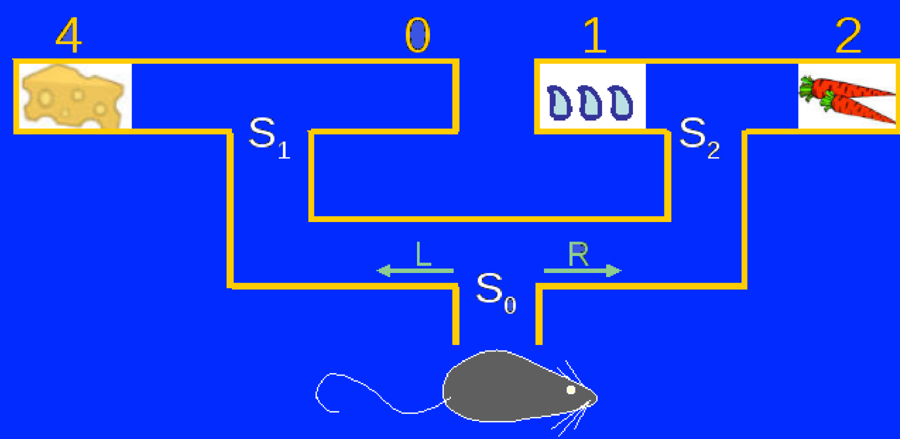
Stored:

$Q(S_0,L) = 4$
$Q(S_0,R) = 2$

$Q(S_1,L) = 4$
$Q(S_1,R) = 0$

$Q(S_2,L) = 1$
$Q(S_2,R) = 2$

Strategy II: Model-free RL



choosing actions is easy so behavior is quick, reflexive (S-R)

but inflexible, need relearning to adapt to any change (habitual)

Stored:

$Q(S_0,L) = 4$
$Q(S_0,R) = 2$

$Q(S_1,L) = 4$
$Q(S_1,R) = 0$

$Q(S_2,L) = 1$
$Q(S_2,R) = 2$

Summary: Model-based vs Model-free RL

- instrumental conditioning reveals that rats indeed have **S-R habits** (and humans, Tricomi et al, 2009)

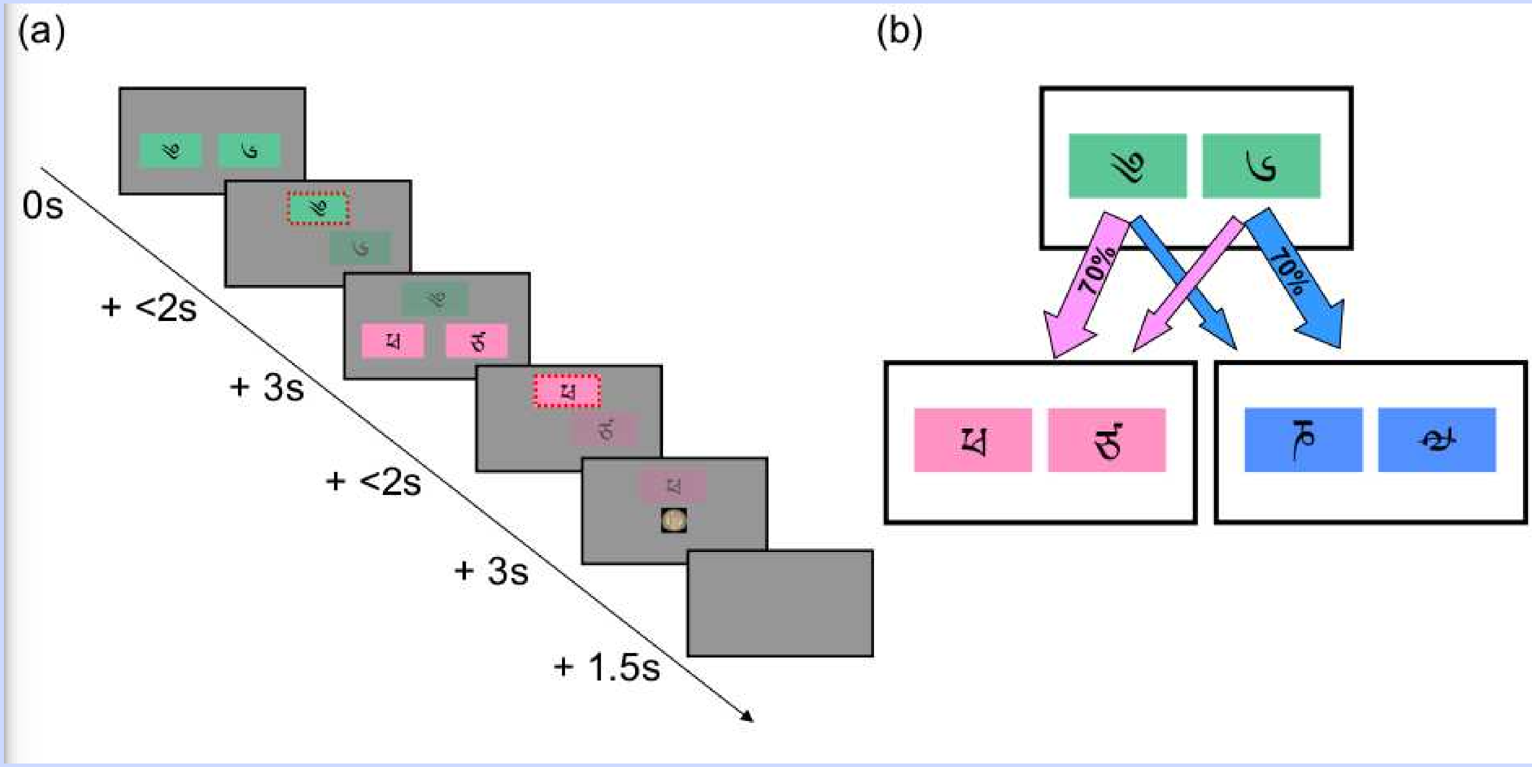
- but even humble rat is cognitive: must distinguish habits from ***goal-directed*** behaviors

- understand this distinction ***algorithmically*** in terms of different RL strategies for decision making, and ***mechanistically*** in terms of functional properties of biological systems involved (BG, PFC, HC..)

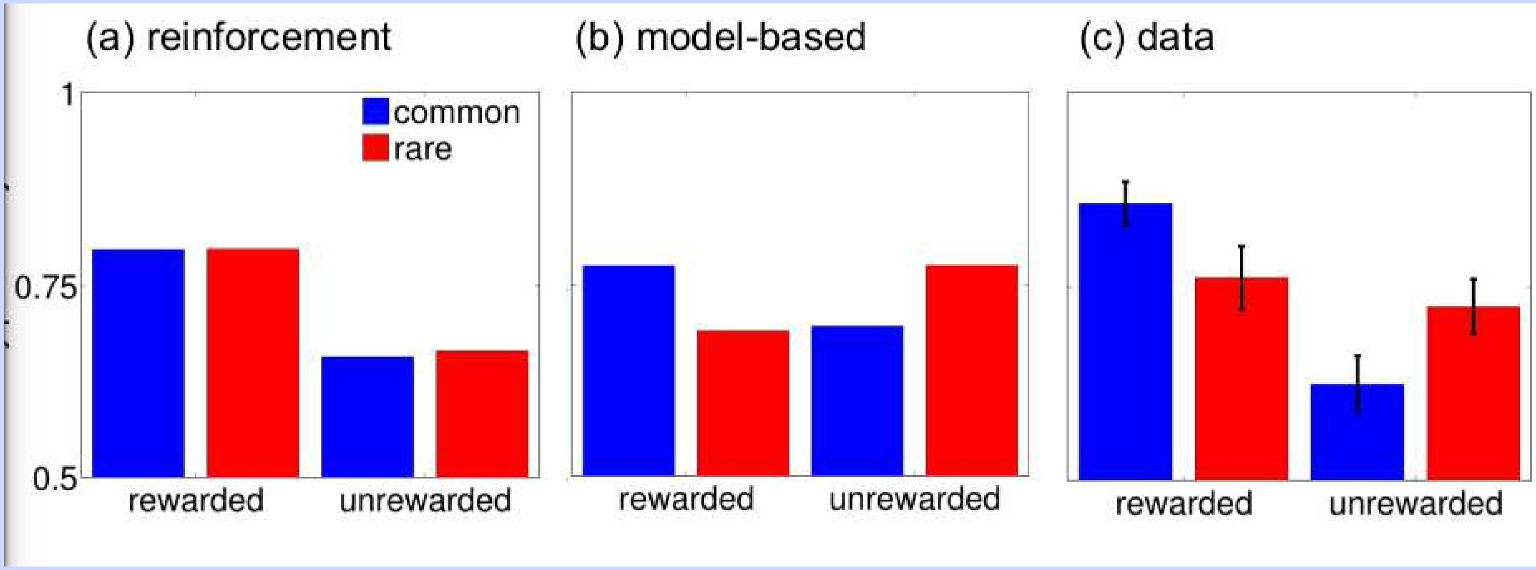
- note: ***same*** overt behavior can be the product of ***different*** neural (computational) systems (controllers)

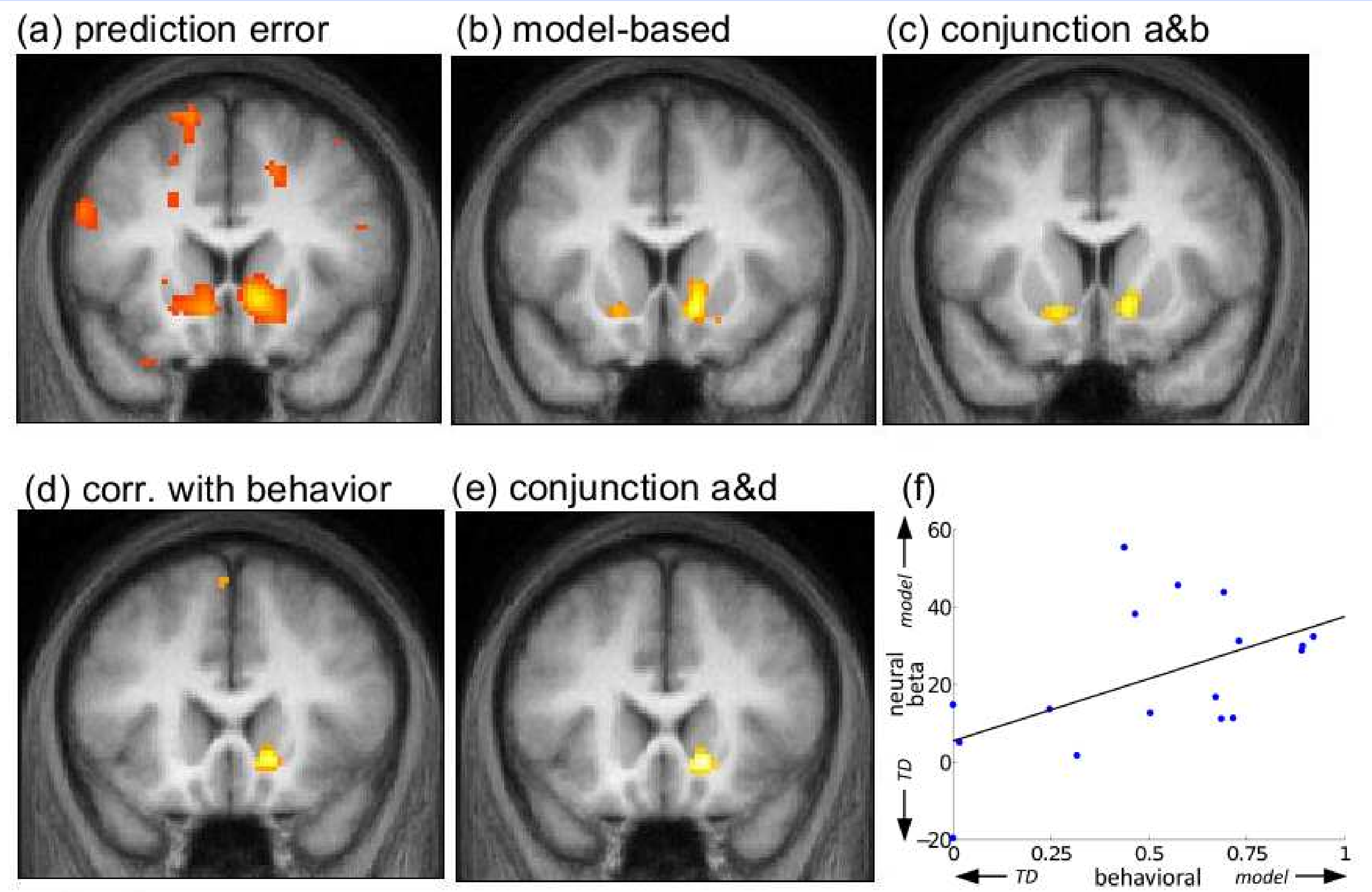
- For computational models of these and related phenomena, including how the brain might arbitrate between the two systems, see Daw, Niv & Dayan (2005) and Frank & Claus (2006)

Model-based vs model free in humans (Daw et al, 2011)

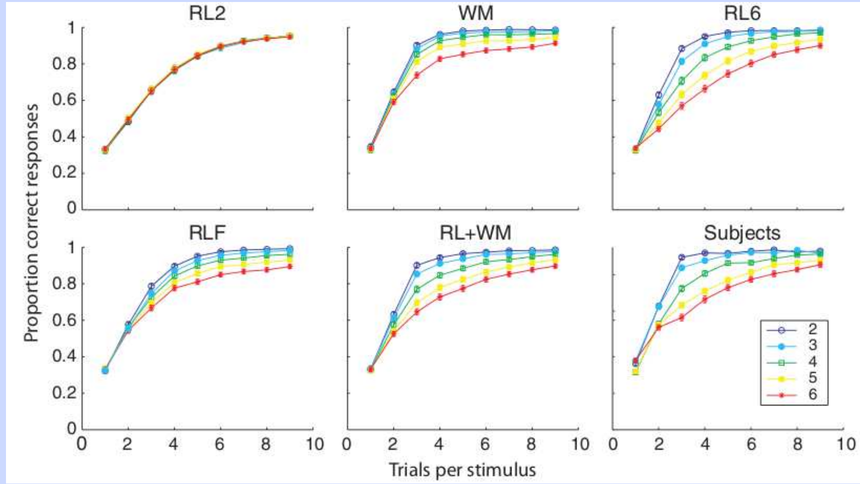


Stay probability on next trial




(a) prediction error
(b) model-based
(c) conjunction a&b
(d) corr. with behavior
(e) conjunction a&d
(f)
60
40
20
0
−20
model
TD
neural
beta
0
0.25
0.5
0.75
1
TD
behavioral
model

The economics of multiple systems: WM vs RL

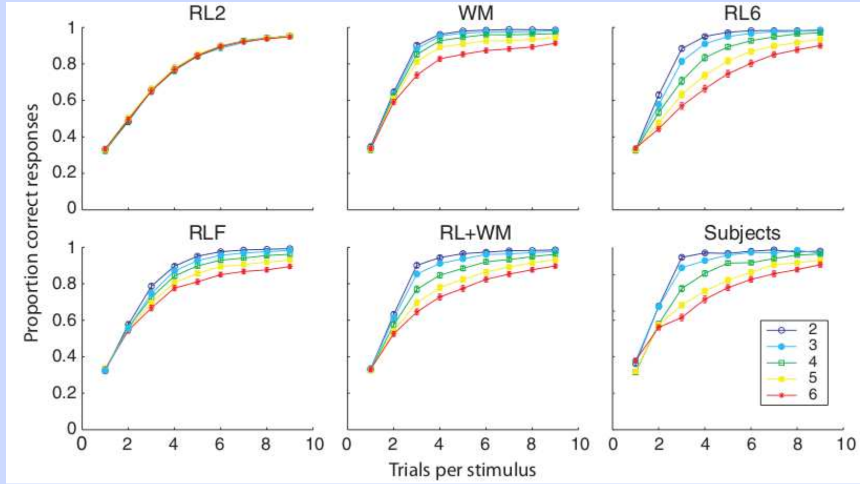


RL2 model:

$$Q_i(t+1) = Q_i(t) + \alpha\delta(t)$$

$$P_{RL}A = \frac{1}{1 + e^{-\beta(Q_A - Q_B)}}$$

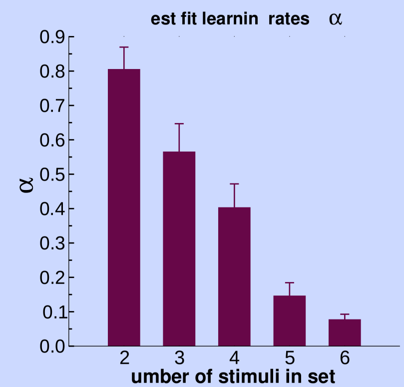
The economics of multiple systems: WM vs RL



RL2 model:

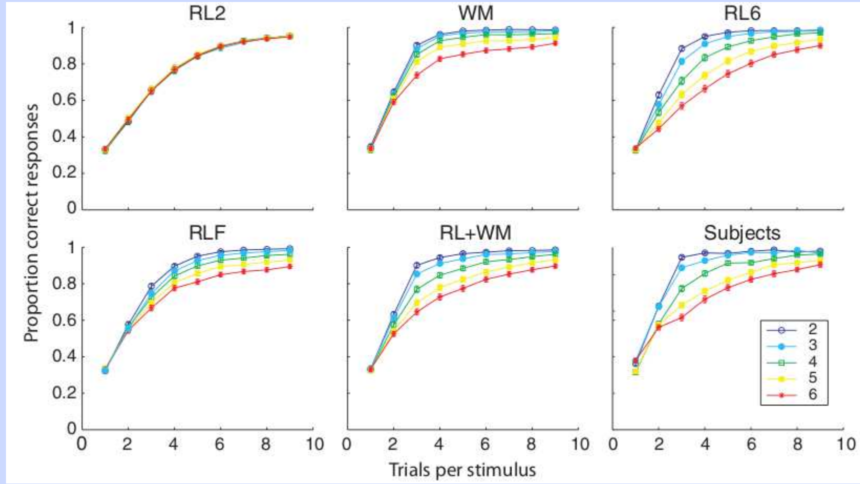
$$Q_i(t+1) = Q_i(t) + \alpha\delta(t)$$

$$P_{RL}A = \frac{1}{1 + e^{-\beta(Q_A - Q_B)}}$$



Collins & Frank 2012; 2017; 2018

The economics of multiple systems: WM vs RL

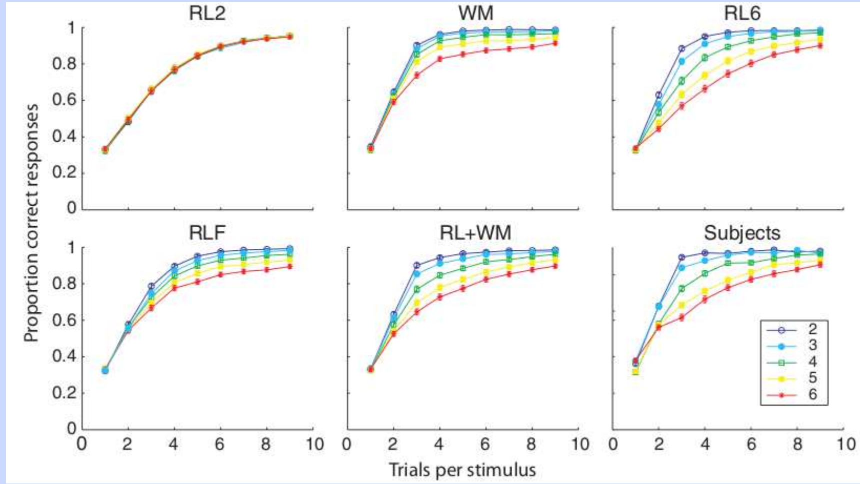


RL+WM model:

$$p(a) = [1 - w(t)] * p_{RL}(a) + w(t) * p_{WM}(a)$$

$$w_{n_S}(t+1) = \frac{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t)}{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t) + p_{RL}(r_t|s_t, a_t)(1 - w_{n_S}(t))}.$$

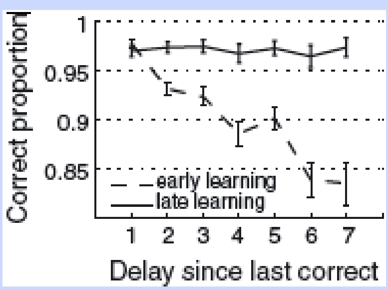
The economics of multiple systems: WM vs RL



RL+WM model:

$$p(a) = [1 - w(t)] * p_{RL}(a) + w(t) * p_{WM}(a)$$

$$w_{n_S}(t+1) = \frac{p_{WMC}(r_t|s_t, a_t)w_{n_S}(t)}{p_{WMC}(r_t|s_t, a_t)w_{n_S}(t) + p_{RL}(r_t|s_t, a_t)(1 - w_{n_S}(t))}.$$



Collins & Frank 2012; 2017; 2018

5. Exploration vs Exploitation (not in text); See Aston-Jones & Cohen, 2005, Ann Rev Neurosci

Reinforcement learning: Dopamine can reinforce rewarding actions so that they are more likely to be executed in the future.

This allows an agent to *exploit* the best possible actions in a situation that are most likely to lead to reward

But what if other possible actions are even better? How would you ever know?

5. Exploration vs Exploitation (not in text); See Aston-Jones & Cohen, 2005, Ann Rev Neurosci

Reinforcement learning: Dopamine can reinforce rewarding actions so that they are more likely to be executed in the future.

This allows an agent to *exploit* the best possible actions in a situation that are most likely to lead to reward

But what if other possible actions are even better? How would you ever know?

Norepinephrine (NE) modulates the *noise* in cortical representations, allows agent to sometimes randomly select some other action.

Exploration vs Exploitation

Reinforcement learning: Dopamine can reinforce rewarding actions so that they are more likely to be executed in the future.

This allows an agent to *exploit* the best possible actions in a situation that are most likely to lead to reward

But what if other possible actions are even better? How would you ever know?

Exploration vs Exploitation

Reinforcement learning: Dopamine can reinforce rewarding actions so that they are more likely to be executed in the future.

This allows an agent to *exploit* the best possible actions in a situation that are most likely to lead to reward

But what if other possible actions are even better? How would you ever know?

Two strategies:

- stochastic choice, where stochasticity is dynamically altered (NE)
- directed uncertainty-driven exploration (strategic information seeking)

LC and Norepinephrine

Two modes of LC firing:

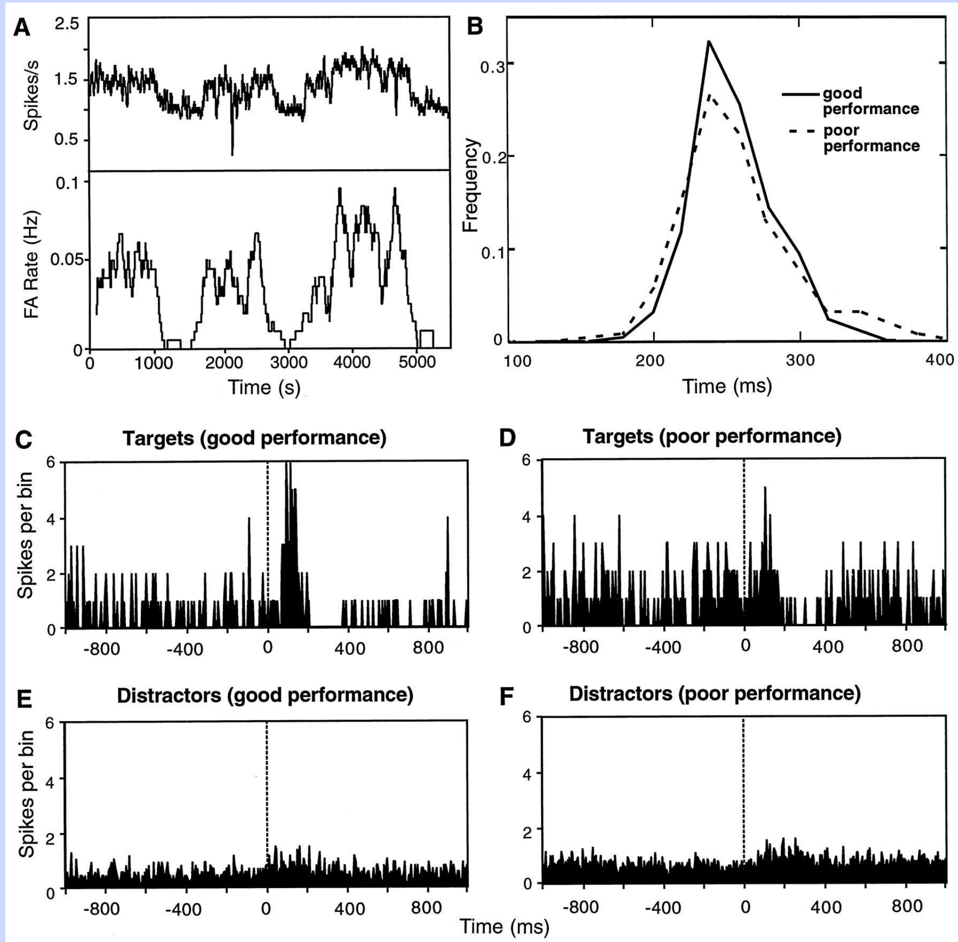
- *Tonic*: high baseline firing
  Effectively adds noise, *RT variability* (Usher et al ’99)
  Supports *exploration* of new behaviors (McClure et al 05)

- *Phasic*: low tonic, but high *evoked* firing
  Facilitates response execution and *exploitation*.

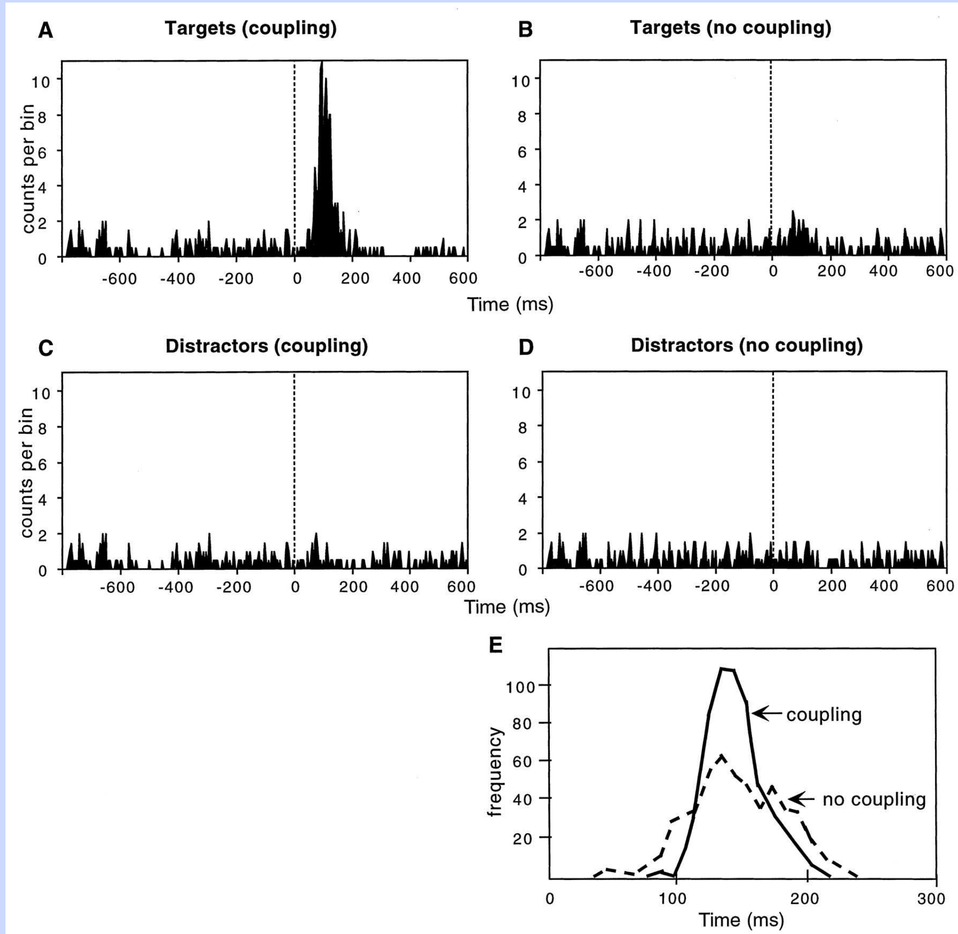
- Phasic mode observed: focused attention, infrequent target detection, good task performance

- High tonic mode during poor performance

Usher et al, 1999

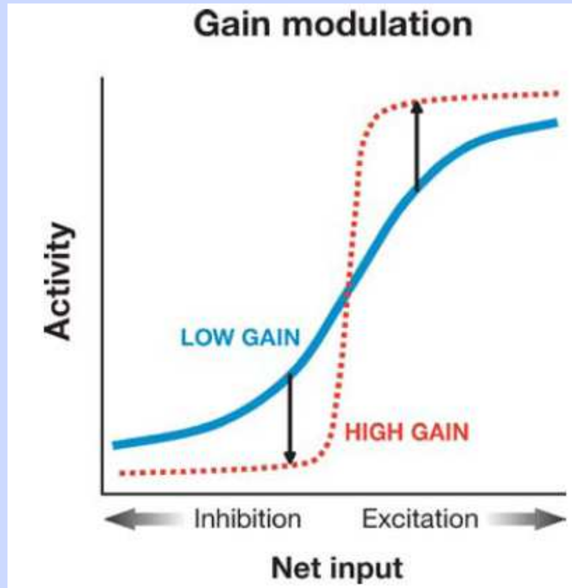


Usher et al, 1999 Model of LC



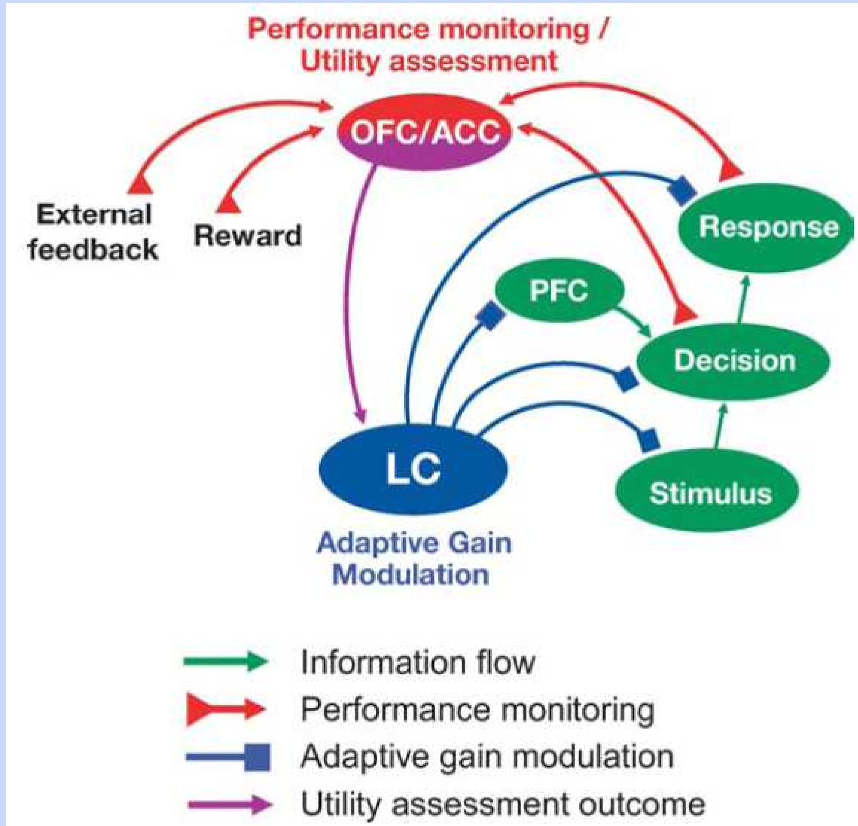
LC/NE effects: Adaptive Gain

Aston-Jones & Cohen (2005)



- Phasic NE facilitates response execution
- Tonic NE enhances noise, reps of multiple actions for exploration.

LC tonic/phasic mode under top-down control

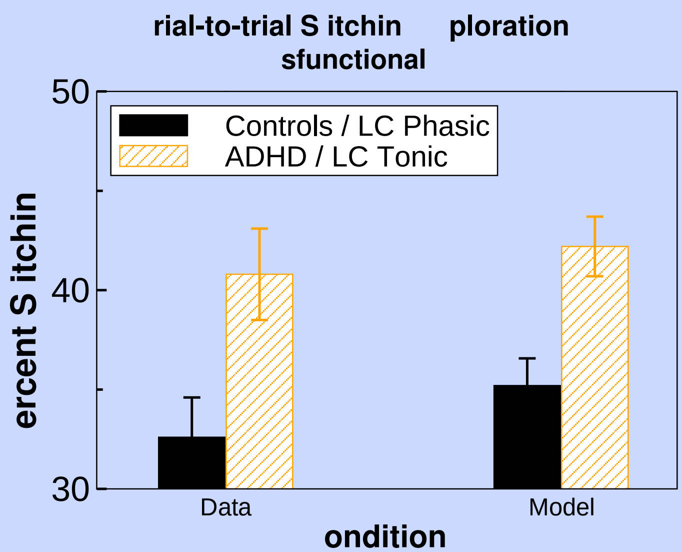


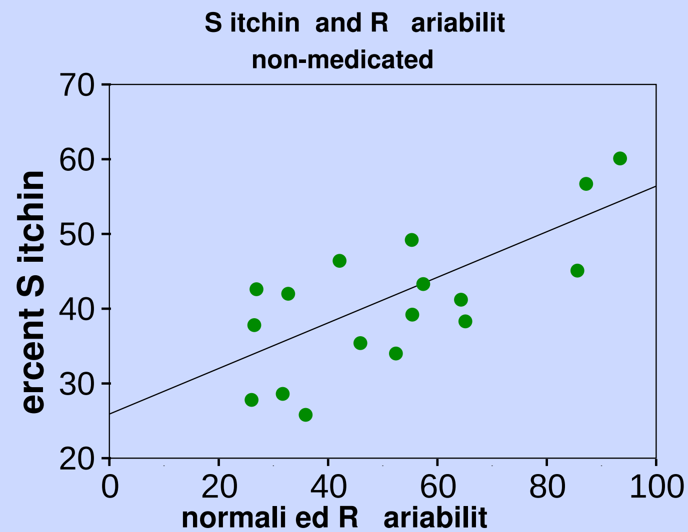
When actions no longer rewarding, NE system responds by increasing noise and exploration of new actions. See McClure et al 2005 for model

ADHD: NE dysfunction?

- Consistent finding of increased RT variability in ADHD
- Also exploration?
- Responsive to medications that modulate NE
- (Also lots of evidence for reduced BG/DA)

ADHD: NE dysfunction?





Frank, Santamaria, O'Reilly & Willcutt (2007)

Also: NE (but not DA) metabolites in urine correlated w/ RT variability in ADHD (Llorente et al, 2006)

Exploration

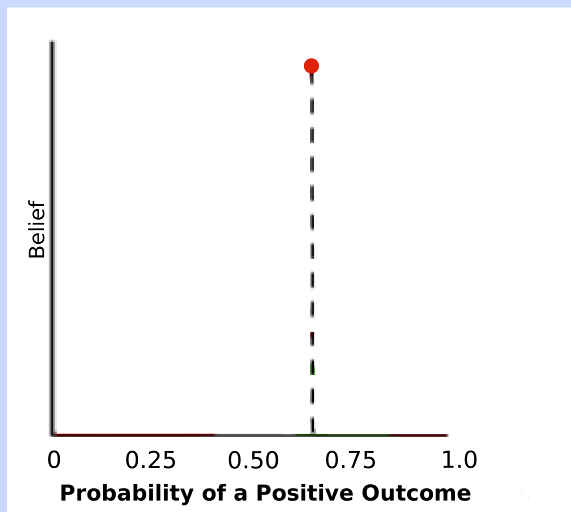
- By exploiting learned strategies, we can get a certain amount of reward

- But when to explore?

- Theory: Explore based on relative *uncertainty* about whether other actions might yield better outcomes than status quo (Dayan & Sejnowksi 96)

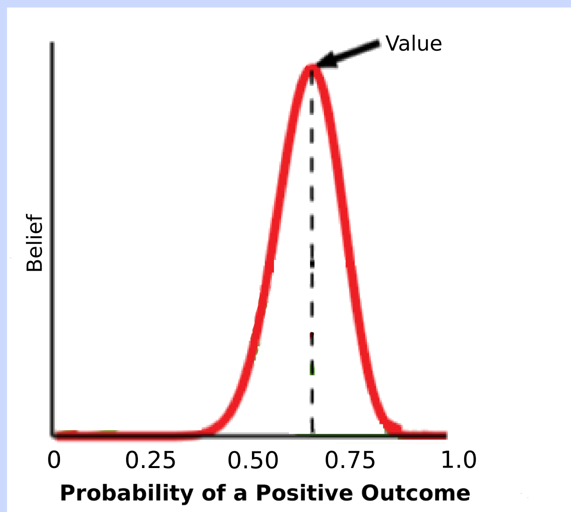
Exploration

- By exploiting learned strategies, we can get a certain amount of reward
- But when to explore?
- Theory: Explore based on relative *uncertainty* about whether other actions might yield better outcomes than status quo (Dayan & Sejnowksi 96)



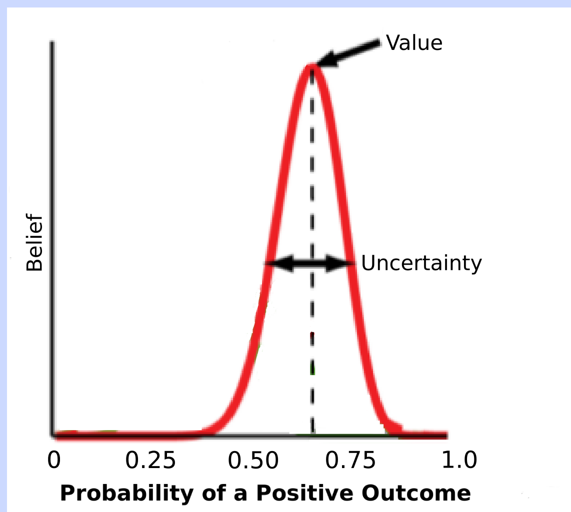
Exploration

- By exploiting learned strategies, we can get a certain amount of reward
- But when to explore?
- Theory: Explore based on relative *uncertainty* about whether other actions might yield better outcomes than status quo



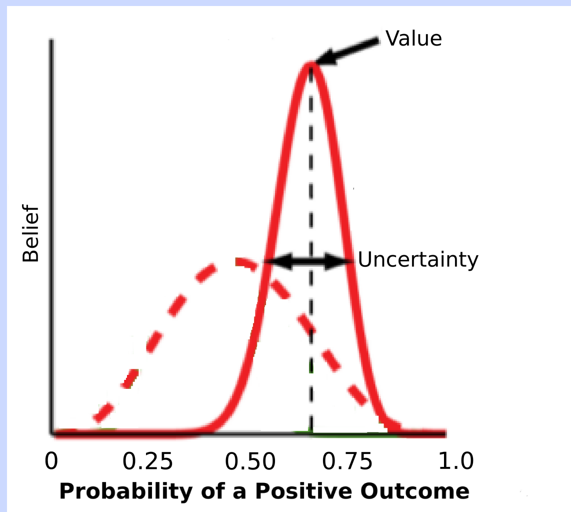
Exploration

- By exploiting learned strategies, we can get a certain amount of reward
- But when to explore?
- Theory: Explore based on relative *uncertainty* about whether other actions might yield better outcomes than status quo



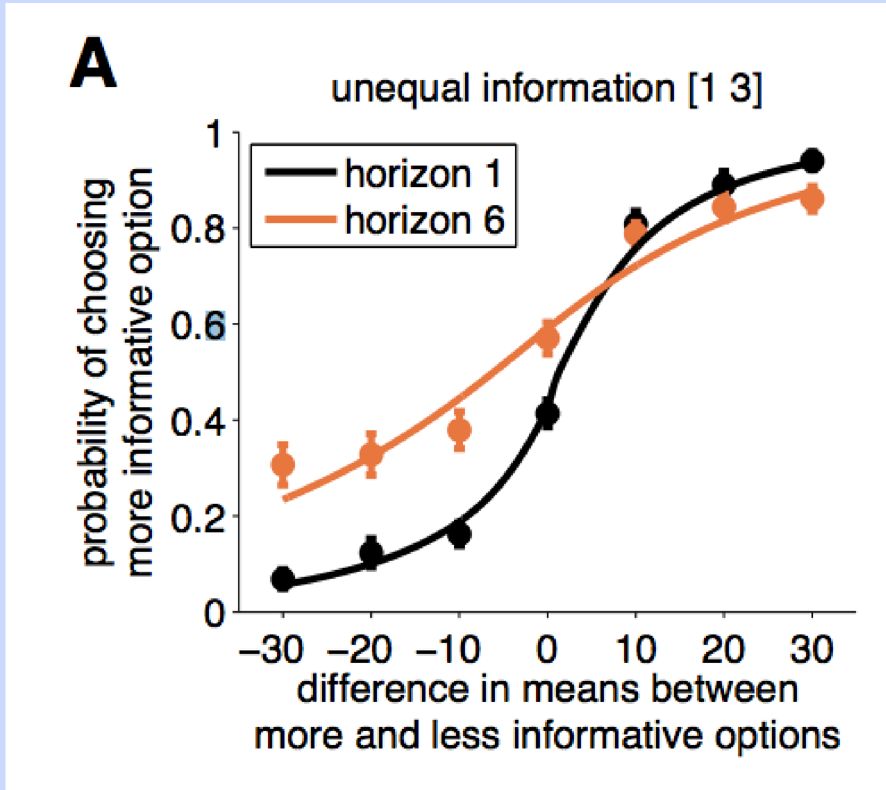
Exploration

- By exploiting learned strategies, we can get a certain amount of reward
- But when to explore?
- Theory: Explore based on relative *uncertainty* about whether other actions might yield better outcomes than status quo



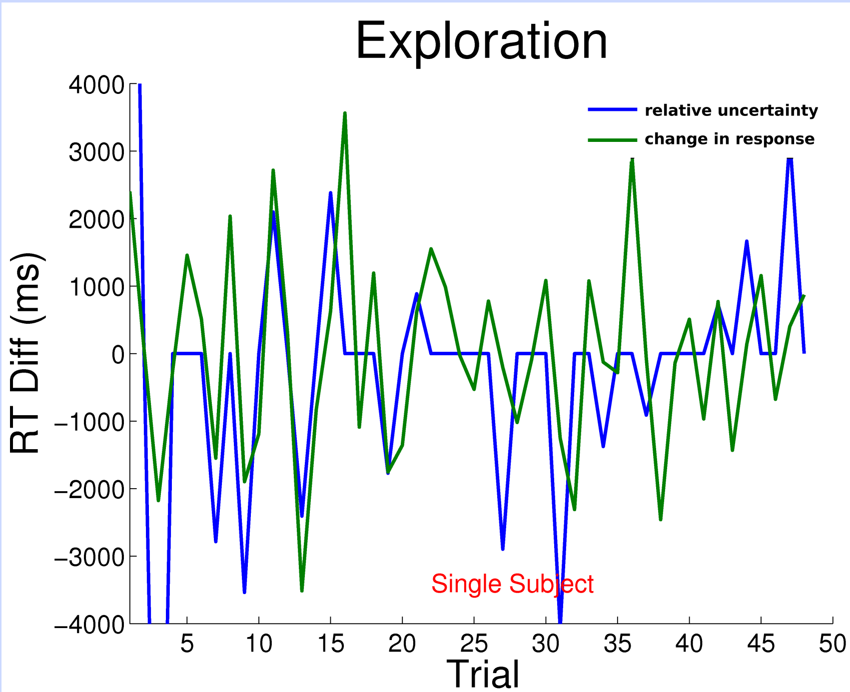
* $\mathbf{Explore(s,t)} = \epsilon\left[\sigma_{\delta|s,a=slow} - \sigma_{\delta|s,a=fast}\right]$

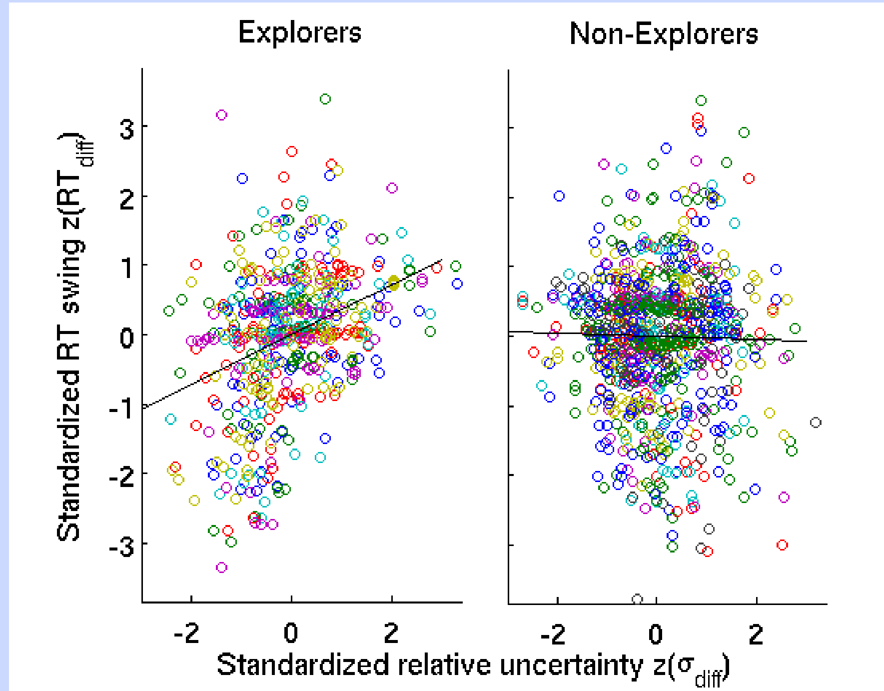
Humans combine both stochastic and uncertainty-driven exploration



Wilson et al, 2014

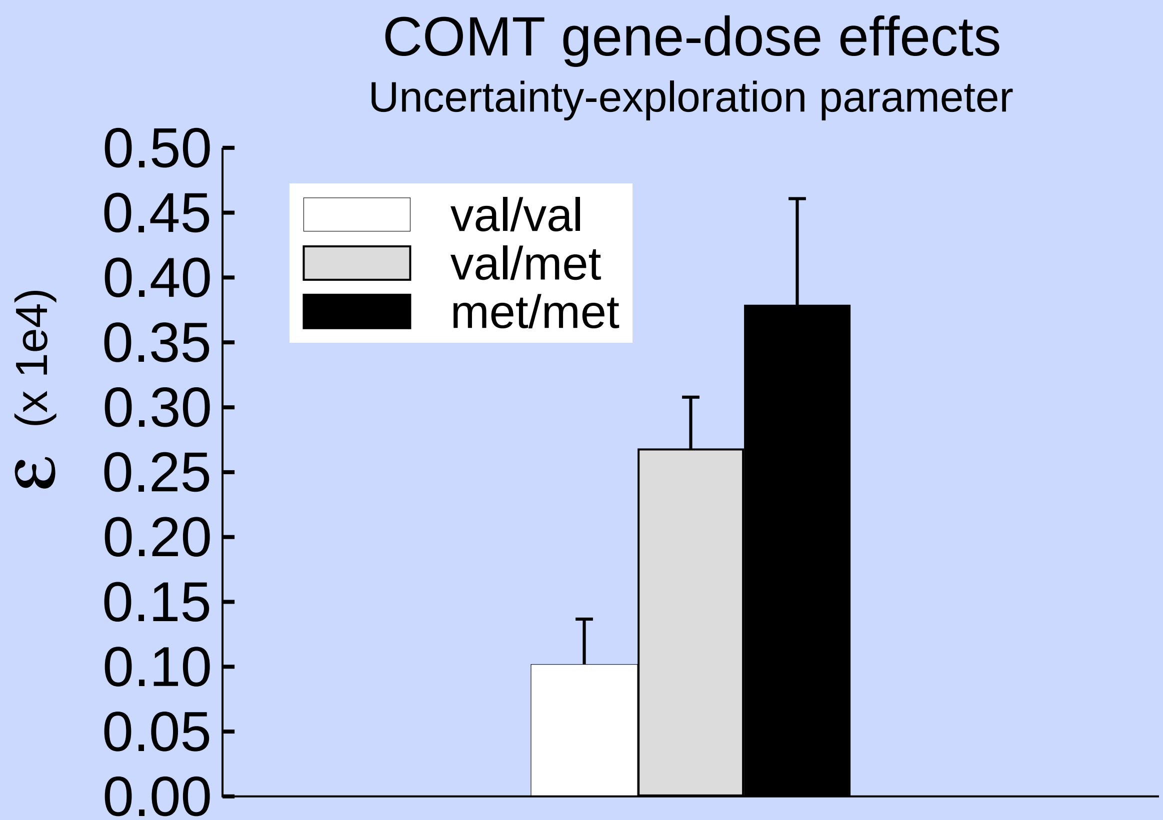
Uncertainty-driven exploration





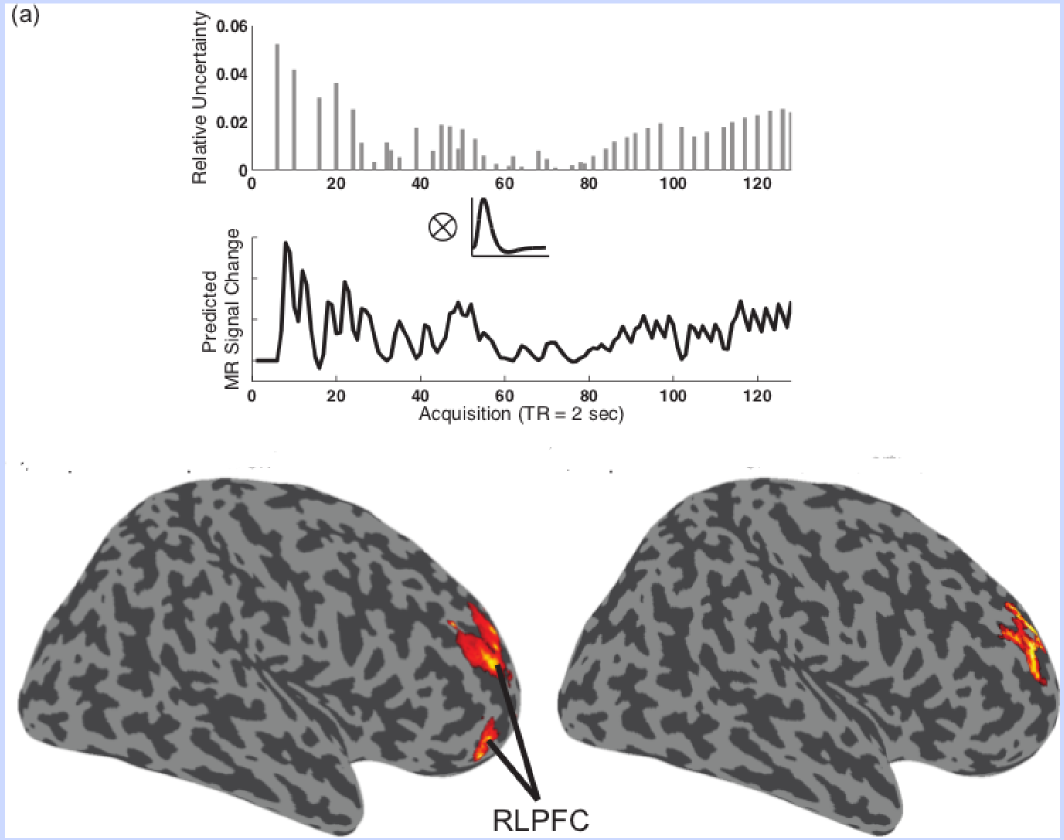
Frank et al ’09, Badre et al 2012

Prefrontal gene effect on uncertainty-driven exploration



Does the brain track relative uncertainty for exploration?

Does the brain track relative uncertainty for exploration?



$\epsilon > 0$ (*'explorers'*) *explorers > non-explorers*

Badre et al 2012, Neuron

Summary

- The functional architecture of the brain reflects the need to simultaneously achieve multiple, computationally incompatible objectives

- To avoid making trade-offs we have evolved specialized structures

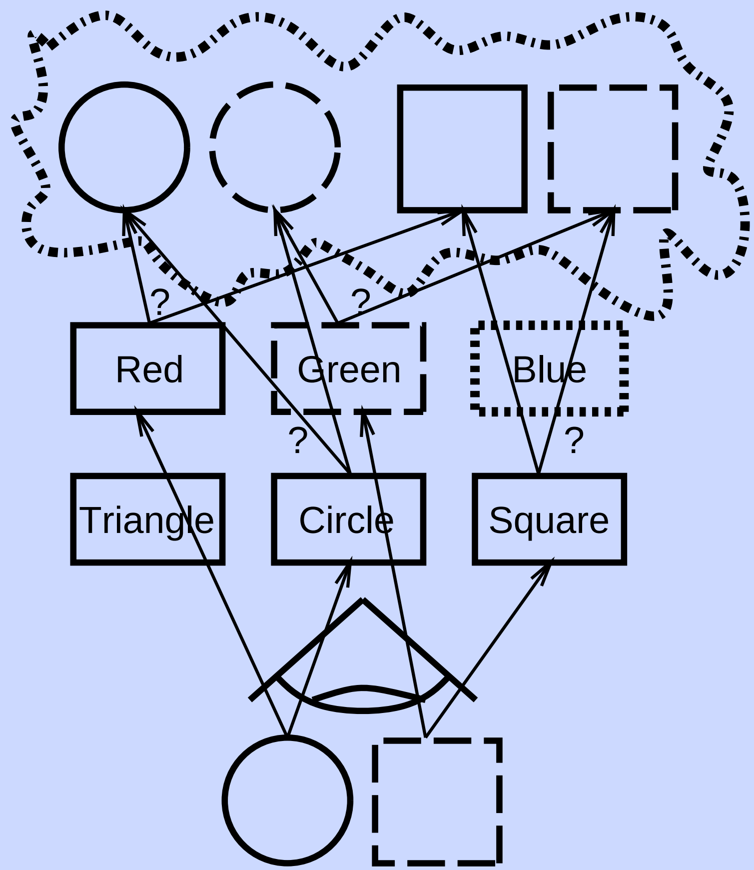
- The process of trying to build computational models (that are compatible with neurobiological and behavioral data) helps us identify these trade-offs

Challenges

Networks are good at some things, but have problems with others..

- Nobody's perfect: People tend to be bad at same things networks are..
- Don't throw the baby out w/the bathwater!

The Binding Problem



The Binding Problem: Potential Solutions

- Attention: only focus on one item.
- Encode conjunctions: no need to have all possible conjunctions separately represented.
- Dynamic synchrony: things that fire together go together.
- Nobody's perfect: people make tons of binding errors..

Other General Problems

- Representing multiple instances of the same thing
  (attention + counting, location)

- Comparing representations
  (overlap – multiple digits, settling in shared weights – goodness, PMC-PFC)

- Nobody's perfect...

Recursion and Subroutine-like processing

- In middle of processing, need to perform same processing (recursion) or different processing (subroutine)
- Easy in standard serial computer (store current state, call subroutine w/appropriate arguments)
- Harder when data and processing not separated!
- HCMP, PFC
- Nobody's perfect...
  The mouse the cat the dog bit chased squeaked.

Generalization

How to recognize new inputs given dedicated, specialized reps?

- Distributed representations: combinations of existing features.
- Abstraction: learn that all dogs might bite, not just that spike bit me..
- Nobody's perfect: Transfer is not good at all..

Important Distinctions

- Controlled *vs* Automatic Processing.
- Declarative/Procedural *vs* Explicit/Implicit.

Consciousness = influence (on Constraint Satisfaction):

- Centrality: more influence on other areas.
- Duration: longer = more influence.
- Intensity: higher = more influence.

A Cognitive Architecture