Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system

Thomas E. Hazy, Michael J. Frank and Randall C. O’Reilly

Department of Psychology,
University of Colorado Boulder, USA

Phil. Trans. R. Soc. B (2007)

Seminar A, SS 09
Motivation

**Challenge:** Modeling higher level areas of human cognition, including decision making, problem solving, and executive control.

The prefrontal cortex (PFC) is crucial for maintaining current context, goals, and other information in an active state that guides ongoing behavior in a coherent, task-relevant manner.
Goal 1

Biologically based computational models can help to break open the mysteries that the PFC houses the „central executive“ by describing the underlying mechanisms in computational detail:

1. How does the PFC know what actions or plans to select?
2. How does experience influence the PFC?
3. How do the specific neural properties of the PFC enable this kind of function?
4. How do these differ from those in other non-executive areas?
Goal 2

The risk run by detailed models is that they provide an elaborate fiction, instead of facts, about how the brain actually works.

**Ansatz:**
Build models that integrate several key working memory tasks in a single instantiation of the model (set of benchmarks)
Goal 1: Core set of six functional demands

These functional demands provide a basic set of constraints for the biologically based model.

1-2-AX-CPT task:
If the subject last saw a 1, then the target sequence is A–X. If the subject last saw a 2, then the target sequence is B–Y.
1. Rapid updating

As each relevant stimulus is presented, it must be rapidly encoded in working memory.
2. Robust maintenance

Information that remains relevant should be maintained in the face of the interference from ongoing processing or other stimulus inputs.
3. Multiple, separate working memory representations

To maintain the outer loop stimuli (1 or 2) while updating the inner loop stimuli (A or B), these two sets of representations must be distinct within the PFC.
4. Selective updating

Only some elements of working memory should be updated at any given time, while others are maintained. (For example, in the inner loop, A or B should be updated while the task demand stimulus 1 or 2 is maintained).
5. Independent output-gating for top-down biasing of processing

For working memory representations to achieve controlled processing, they must be able to bias (control) processing elsewhere in the brain - and at the appropriate time.
6. Learning what and when to gate

Underlying all successful working memory task performance is the need to learn when to gate appropriately - both gating ‘in’ for maintenance and ‘out’ for biasing processing elsewhere.
Rapid updating vs. Robust maintenance (1 & 2)

Both are in direct conflict with each other?

**Ansatz:**
Dynamic gating mechanism

![Diagram showing D2 and D1 receptor activation](image-url)
Dynamic gating via basal ganglia

In the motor system, the BG are interconnected with frontal cortex through parallel loops.

Subthalamic nucleus (STN)
External segment of the globus pallidus (GPe)
Internal segment of the globus pallidus (GPi)
Substantia nigra (SNr)
How gating solves the 1-2-AX task

(a) stim  
1 → PFC 1

(b) C  
C → C 1

(c) A  
A → A 1

(d) X  
X → RA 1
Multiple, separate working memory representations (3)

... are possible owing to the ‘striped’ micro-anatomy of the PFC, which is characterized by small, relatively isolated groups of interconnected neurons (preventing undue interference between representations in different even nearby stripes).

- Functionally similar to and roughly the same size as the well described hypercolumns of the visual cortex (?).

- Estimated 20000 such stripes in human frontal cortex
Selective updating (4)

Occurs owing to the existence of independently updatable parallel loops of connectivity through different areas of the BG and frontal cortex.

• Loops are selective to the relatively fine grained level of the anatomical stripes in PFC.

• This stripe-based gating architecture has an computational advantage over the global nature of a purely dopamine-based gating signal.
Independent output-gating for top-down biasing of processing (5)

Occurs via **output-gated projections** from actively maintained representations in PFC to relevant areas (posterior cortex)

- Deep, output-generating laminae of the PFC display thresholded behaviour. They do not fire until a threshold is reached (via BG)
- Output-gating is the same mechanism as the motor gating that the BG are typically described as performing
Learning what and when to gate (6)

Solving the temporal credit assignment problem

(a)

CS
US/r
DA

(b)

CS
PFC
BG—Go
US/r
DA

(maint in PFC)
(causes updating)
(spans the delay)
(reinforces Go)
DA system learning

• Is accomplished by a dopamine-based reinforcement-learning mechanism (PVLV)

• Learning occurs in parallel for maintenance and for output.

• BG are thought to learn to facilitate the selection of the most appropriate response while suppressing all other responses.

• Increases in DA enhance BG Go firing via D1 receptors, whereas decreases in DA during negative reinforcement have enhancing NoGo firing via simulated D2 receptors.
PVLV learning model

**Primary value (PV) system:** Learns to cancel primary rewards.

Rescorla–Wagner equation

\[ \delta^t = r^t - \hat{r}^t, \quad \hat{r}^t = \sum_i w_i^t x_i^t. \quad \Delta w_i^t = \epsilon \delta^t x_i^t. \]

where \( r^t \) is the current reward value at time \( t \), \( \hat{r}^t \) is the expected or predicted reward value, and \( \delta^t \) is the discrepancy or error between the two using synaptic weights \( w_i \) from a set of sensory inputs \( x_i \).

We take the \( \delta^t \) to represent the dopamine firing deviations from baseline.
Primary value (PV) system: Learns to cancel primary rewards.

Rescorla–Wagner equation

\[ \delta_{pv}^t = PV_e^t - PV_i^t = r^t - \hat{r}^t, \]

The excitatory PV system represents the value implicitly hardwired into a primary reward (US)

\[ PV_e^t = r^t \]

whereas the inhibitory system learns to cancel out these rewards,

\[ PV_i^t = \hat{r}^t \]
**PVLV learning model**

**Learned value (LV) system:** Signals reward associations.

Rescorla–Wagner equation

\[ \delta_{lv}^t = LV_e^t - LV_i^t. \]

LVe system only learns when primary rewards are present or expected. Therefore, unlike the PV system, it is able to signal the reward association of a CS

\[ \Delta w_i^t = \begin{cases} \varepsilon (PV_e^t - LV_e^t) x_i^t, & \text{if } PV_{filter} \\ 0 & \text{otherwise.} \end{cases} \]

\[ PV_{filter} = PV_i^t > \theta_{pv} \text{ or } PV_e^t > \theta_{pv}, \]
PVLV learning model

**Learned value (LV) system:** Signals reward associations.

Rescorla–Wagner equation

\[ \delta_{lv}^t = LV_e^t - LV_i^t. \]

The LVi system is essentially the same as the LVe system, except that it uses a slower learning rate (\(\epsilon\)).

Because of its slower learning rate, LVi slowly learns which CSs are reliably associated with reward and decreases the dopamine bursts for such CSs relative to those that have more recently become associated with reward.
PVLV learning model

(Timing, adaptation)
2\textsuperscript{nd} order conditioning

Learning in the hidden and sensory layers is implemented as

**Contrastive Hebbian learning**

\[ \Delta w_{ij} = (x_i^+ y_j^+) - (x_i^- y_j^-) \]

+ … DA modulated firing
- … without DA signal

\( x_i \) … activation of pre.syn neuron \( i \)
\( y_j \) … activation of post.syn neuron \( j \)
Implementation of the gate

state 1: successful Recall

PFC

\[ S \quad S \quad S \]

Input \[ R \]

Striatum

\[ g_n \quad g_n \]

DA

prev gate = S stored
\[ \Rightarrow \text{rew on } R, \text{ assoc } S \text{ (in PFC!) w/ rew in DA sys} \]

state 2: storing S

\[ S \quad X \quad X \quad S \]

GO = store S, DA likes \[ \Rightarrow \text{GO firing reinforced} \]

state 3: ignoring I

\[ S \quad S \quad S \]

NOGO = ignore I, DA stopped firing fm syn. depress

Randomly stored S
Random selected (gate?) correct recall when input R is present
LV learns rewarded S in PFC (DR)

Randomly gating S into PFC
S in PFC detected by LV
LV triggers reward and learns the gating of S (DR)

Input I is ignored
(otherwise negative reward)
DA stops firing because of adaptation
Algorithm pseudo code

For each event:
1. Iterate over minus (−), plus (+), and update (++) phases of settling
   (a) At start of settling:
      i. For non-PFC/BG units, initialize state variables (Vm).
      ii. Apply external patterns (clamp input in minus, input and output, external reward based on minus-phase outputs).
   (b) During each cycle of settling, for all nonclamped units:
      i. Compute excitatory net input
      ii. For Striatum Go/NoGo units in ++ phase, compute additional currents based on DA inputs from SNc
      iii. Compute kWTA inhibition for each layer
      iv. Compute point-neuron activation combining E and I currents
2. Change weights with delta rule: PV, LV and striatum
   CHL: All other units
The full model
Goal 2: Benchmark tasks

Core tasks (already modelled or nearly finished):
1. Stroop task
2. Wisconsin card sort task
3. ID/ED task
4. Erikson Flanker task

Paradigms to be addressed in the extended models:
1. ABCA/ABBA task
2. Sternberg task
3. N-back task
For color naming, negative SOA means word preceded color.
Wisconsin card sort task

Sort multi-attribute cards into piles according to an unknown and changing rule.
ID/ED task

1. The subject is presented with two shapes and must learn which of them is always correct.
2. The other dimension (lines) is then added which the subject must learn is irrelevant.
3. Completely new examples of the dimensions are then introduced, and the subject must learn that one of the new shapes (i.e., same dimension) is always correct (ID shift).
4. Shapes then become irrelevant, requiring the subject to attend to the previously suppressed dimension (lines) and learn which of the lines is correct (ED shift).
5. In addition, after each of the above stages (3-4), a reversal learning test is interposed, where the correct exemplar of the dimension is swapped without warning.
Erikson Flanker task

Classify an item surrounded or flanked by distracting symbols.

SSSSS
SSHSS
HHSHH
HHHHH

When reaction times are obtained at a range of durations from short to long, accuracy in conflict conditions initially goes below chance before rising to relatively high levels.
ABCA vs ABBA

- Delayed match to sample task

- Monkeys were presented with a sequence of stimuli, and trained to respond whenever the 1st stimulus repeated, eg. ABCA

- But, ABCA task was solved via a weight-based familiarity signal encoded in inferior-temporal (IT) cortex and not PFC.

- ABBA version of the task, which contains an embedded repeat of the B stimulus, this familiarity-based system was fooled and required activation of PFC

- Familiarity vs. active maintenance (IT cortex vs PFC)
Sternberg task

Suggests a process where people sequentially scan memory items, using an exhaustive (non-self-terminating) search for both positive and negative cases.

Sternberg, 1966
N-back task

Detect a repetition of any stimulus that occurred n steps earlier in the sequence (where n can be varied from 1-5; most people cannot perform well above 3) at each time step.
Conclusions

Neurons build up strong, complex “relationships”; a neuron can only function by learning which of the other neurons it can trust to convey useful information.

In contrast, a digital computer functions like the post office, routing arbitrary symbolic packages between passive memory structures through a centralized processing unit, without consideration for the contents of these packages.

The dynamic gating mechanisms work more like a post-office, with the basal ganglia reading the zip code of which PFC stripe to update, whereas the PFC cares more about the contents of the package.