Higher Level Cognition: What’s Missing

- Planning, problem solving, reasoning, complex decision-making
- What do all of these have in common?
- Top-down control of behavior: Instead of reacting in a bottom-up fashion to stimuli, behavior is driven (controlled) by an actively maintained representation of what we are supposed to be doing...
- Allows us to behave in contextually appropriate fashion instead of just giving the strongest, most dominant response
- Also gives us the ability to link events across time points, and to carry out behaviors that are extended across time...

Higher Level Cognition: What We Know

Frontal damage impairs planning, reasoning, decision-making, self-initiated actions, self-awareness, social interaction...

The Range of Frontal Functions

Activation-based Working Memory

- The prefrontal cortex (PFC) can maintain information over time as firing of neurons (activation-based memory).
- Top-down vs bottom-up PFC
  - Buschman & Miller, 2007, Science
    - parietal act for target location precedes pfc act for pop-out
    - pfc act precedes parietal for search

Activation-based working memory

- The PFC can maintain information over time as
  - Parietal act for target location precedes PFC act for pop-out
  - PFC act precedes parietal for search
Top-down vs bottom-up PFC

Buschman & Miller, 2007, Science

• greater low freq pfc-parietal synchronization for top-down
• greater high freq synchrony for bottom-up


The Range of Frontal Functions

Activation-based working memory

Inhibition

Stroop: Difficulty inhibiting prepotent response.

Flexibility

Continue with same response after task changes, perseveration.

Fluency

Difficulty generating variety of responses.

Executive control

Probs w/ goal-directed planning, coordinating. (shopping)

Monitoring/evaluation

e.g., Error-monitoring.

The Range of Frontal Functions

The Stroop Task

The Stroop Effect:

GREEN

Possible explanation: differential pathway strength

• two pathways: word reading and color naming

• these compete to generate response

• word reading pathway is much stronger than color naming

• two pathways word reading and color naming


Red

Green

Yellow

Pink

The Range of Frontal Functions

Shopping (spending)

Execute control, focus w/ goal-directed planning, coordinating.

Fluency

Difficulty generating variety of responses.

Flexibility

Continue with same response after task changes,

Inhibition

Stroop: Difficulty inhibiting prepotent response.

Activation-based working memory

The Range of Frontal Functions
Stroop Effect:

Puzzle: If the color naming pathway is weaker than word reading, how do we manage to name the color of the word "green" above?

Solution: Prefrontal cortex actively maintains a representation of the task you are supposed to be doing (color naming or word reading). This actively maintained task representation biases processing in the posterior cortex by activating units in the appropriate pathway. For example, the color naming task representation in the prefrontal cortex sends activation to the units in the color naming pathway...

Model of the Stroop Task (Corwin, Donker, & Mecklheinrich, 1990):

Effects of Training:
The Stroop Task: Model Data


Color Naming

Word Reading

Stroop Data (Dunbar & MacLeod, 84)

Control
Conflict
Congruent

Pathway Strength vs Processing Speed Theories

- Model predicts key difference between word reading and color naming is pathway strength (reading > color naming). This results in:
  - Word reading being faster than color naming
  - Asymmetric interference effects

- Other (verbal) theories posit Stroop effects resulting from a horse race rather than direct competition

Horse Race Theories

- Color does not affect word reading because the word reading process runs to completion before color is processed
- Conversely, word identity does affect color naming because word reading process completes before color response is generated
- This theory, stated as such, implies that it should be possible to get color to interfere with word info if the color naming process has a head start
Stroop Accounts: Not a Horse Race

• Color naming will interfere with new task.
• New task will have no effect on color naming.
• Well-learned than color naming we will find that
  Predictions: If we could come up with a task that is
  Even Less
  related to the Stroop paradigm, with less overlap
  between the two tasks, we would find that
  on absolute pathway strength
  Stroop (whether a process is controlled or automatic) depends

Stroop Accounts: Automaticity

• Early accounts of Stroop focused on automatic vs controlled
  processing.
• According to these theories, word reading is automatic and
  color naming is controlled.
  • Automatic processes don’t suffer from interference (they
    proceed automatically), but controlled processes do.
  • In contrast, model focuses on relative pathway strength.
    Stronger pathways interfere with weaker pathways.
  • Status (whether a process is controlled or automatic) depends
    on absolute pathway strength.
  • Prediction: If we could come up with a task that is
    even less well-learned than color naming, we will find that:
    • New task will have no effect on color naming.
    • Color naming will interfere with new task.
As new task is practiced repeatedly, effects should reverse. Stroop Accounts: Continuum, not a Dichotomy.

MacLeod & Dunbar, 1988

But how do PFC units come to represent task rules? Stroop model is a nice simple account of PFC function, but it somehow assumes that PFC ‘knows’ how to maintain a rule for color naming and to magically bias color-naming hidden units. Interesting question is how these rule-like representations develop in the first place?

Can PFC learn to assign abstract rule-like representations that code for stimulus dimensions (e.g., color) by experience (with multiple colors)? Can PFC learn to assign abstract rule-like representations that code for stimulus dimensions (e.g., color) by experience (with multiple colors)? How do PFC units come to represent task rules?

Developing PFC Reps

PFC Specializations — Rule-Like Abstract Reps

Stimuli Network

color
size
shape
position
texture

Hidden (83 units)

Task Hidden (16 units)

Left Stimulus Right Stimulus Task

NF MF SF LF

A1 A2 A3 A4
B1 B2 B3 B4
C1 C2 C3 C4
D1 D2 D3 D4
E1 E2 E3 E4

A1 A2 A3 A4
B1 B2 B3 B4
C1 C2 C3 C4
D1 D2 D3 D4
E1 E2 E3 E4

AG

Dimension Cue

Response

Network

Input/Target

left right task

left right task

left right task

Correct

Response

Key: Do repeated trials of same task – continuous attention to shapes, etc.
The Range of Frontal Functions

Monitoring/evaluation
Executive control (top-down vs. goal-directed programming)
Inhibition
Flexibility/continuity
Activation-based working memory
Prospective memory

The Rule-Like Abstract Reps

Posterior Full PFC
Stimuli: color, size, shape, position, texture
Rule = One stimulus dimension (row) relevant at a time. (e.g., card-sorting tasks)
Abstraction derives from sustained maintenance over trials!
Posterior net 'memorizes' specific combinations of features/responses for each task, doesn't develop systematic representations.

Adaptive gating is key:
Within block of trials feature changes but gating mech learns to maintain constant PFC rep (until rule switches, performance goes down → update).

As is breadth of experience (same stimuli across different tasks)
(increasing pressure to use same PFC reps across tasks → systematicity); with small # tasks can get by with memorizing.

Rule-Like Abstract Reps

Generalization
Training Regimen
Network Configuration
Rule Likeness Predicts Generalization
r = .97
Abstraction → better generalization across tasks (accuracy on stimuli not seen in particular task).
Interaction of nature (PFC mechanisms) and nurture (breadth of experience).

Stroop Performance

Stroop Task
Neutral Conflict

Stroop Task: Lesion Data
Same network & parameters: PFC control representations developed entirely through learning from random initial weights!
LF = left frontal (DLPFC) lesions in people and model
devlop and rely on top-down control from anterior frontal regions.

The Range of Frontal Functions

Activation-based working memory
Prospective memory

Monitoring/evaluation
Executive control (top-down vs. goal-directed programming)
Inhibition
Flexibility/continuity
Activation-based working memory
Prospective memory

Rule-Like Abstract Reps

Posterior Full PFC
Stimuli: color, size, shape, position, texture
Rule = One stimulus dimension (row) relevant at a time. (e.g., card-sorting tasks)
Abstraction derives from sustained maintenance over trials!
Posterior net 'memorizes' specific combinations of features/responses for each task, doesn't develop systematic representations.

Adaptive gating is key:
Within block of trials feature changes but gating mech learns to maintain constant PFC rep (until rule switches, performance goes down → update).

As is breadth of experience (same stimuli across different tasks)
(increasing pressure to use same PFC reps across tasks → systematicity); with small # tasks can get by with memorizing.

Rule-Like Abstract Reps

Generalization
Training Regimen
Network Configuration
Rule Likeness Predicts Generalization
r = .97
Abstraction → better generalization across tasks (accuracy on stimuli not seen in particular task).
Interaction of nature (PFC mechanisms) and nurture (breadth of experience).

Stroop Performance

Stroop Task
Neutral Conflict

Stroop Task: Lesion Data
Same network & parameters: PFC control representations developed entirely through learning from random initial weights!
LF = left frontal (DLPFC) lesions in people and model
devlop and rely on top-down control from anterior frontal regions.
Dynamic Categorization Tasks

Wisconsin Card Sort
A B C D

Experimental task (like Stroop), but captures some essential aspects of higher level cognition.

Frontal patients perseverate with the first rule.

→ weight based tendencies build up when categorizing according to rule and you need to actively maintain the new rule to counteract these weight-based tendencies.

“Frontal Tasks”

• Stroop: Ability to override prepotent response (word reading)
in favor of currently relevant task (color naming) – requires top-down control.

• Prefrontal control” not just for overriding long term associations like word reading, but also for the ability to quickly change attention in an online fashion in response to changing task demands. UPDATING.
Two strategies for solving donut categorization task:

• Adjust weights to different donut types
• Actively maintain a representation of your current strategy; deactivate this representation and activate another if you receive negative feedback

Active maintenance does not strongly benefit initial learning of the rule, but greatly facilitates performance when the rule changes.

Card Sorting Tasks

In what situations do we need to consider/represent different rules in mind and have the ability to flexibly update/maintain them until one works well? Right now, thinking! I’m asking you a question you consider unanswerable (e.g., “Never: card sorting tasks are dumb”). You then evaluate the quality of what you’re holding in mind: does it make sense, is it likely to produce a good outcome?

In everyday life or just to this peculiar task? Good measure of online thinking & problem solving: The ability to flexibly consider different possibilities to guide thinking and behavior.

Dynamic Categorization Tasks: ID/ED task
Dynamic Categorization Tasks: ID/ED task

Same (Reversal)
Different (Shift)

IDR
EDR
IDS
EDS

Intra−dimensional (ID)
Extra−dimensional (ED)

Target change stimuli

Rule Change
Previous Stimuli

ID/ED and Frontal Damage
(Dias, Robbins & Roberts (1997), J Neurosci)

IDS IDR EDS

Errors to Criterion

Perseverations from Frontal Lesions

Control
Orbital
Lateral

Original interpretation: Orbital = affective inhibition, Lateral = attentional selection.

Alternative Account
(O'Reilly, Noelle, Braver & Cohen (2002), Cerebral Cortex)

Orbital PFC represents detailed features.
Lateral PFC represents abstract dimensions.

Activation-based PFC processing facilitates rule switch:
Orbital = switch to new features (IDR).
Lateral = switch to new dimension (EDS).

Perseverations = weight-based processing in absence of PFC.

ID/ED Model

Input
Hidden
PFC_Feat
Output
PFC_Dim

Dim A left, Dim B left, Dim B right.

Two dimensions, A and B (shapes & lines)

On each trial, four stims are presented:
Dim A left, Dim A right, Dim B left, Dim B right...

ID/ED Model

Input
Hidden
PFC_Feat
Output
PFC_Dim

Posterior Cortex

Activation limited in cortex: attention.
PFC provides top-down bias, with DA/updating unit.
PFC Feat = stim features indep of location
PFC Dim = abstract dimensions
PFC updating based on unexpected rewards and errors

- When there is an increase in DA activity (e.g., the model got the answer right but wasn't expecting a reward):
  - Hidden unit activity is gated into PFC
  - Connections from hidden units to DA are increased
  - PFC serves to amplify the influence of hidden units associated with correct responding

- When there is a decrease in DA activity (the model was expecting a reward but gave the wrong response):
  - PFC activity is wiped clean
  - Connections from hidden units to DA are decreased

- Also, there is some "gating noise": trial and error search

Similarities/Differences with Store-Ignore-Recall

- With S-I-R, the model had to gate the "Store" stimulus into PFC (and carry it forward in time) in order to respond correctly; S-I-R can only be solved with the help of active maintenance (working memory)

- The ID/ED task can be solved without active maintenance; but PFC can help by focusing the model's attention on useful parts of the input but is not necessary

Explanation of lesion Data: IDS

• Intradimensional shift (IDS): different stimuli pre and post-shift; the relevant dimension (A) stays the same
  - PFC is unnecessary because there are no strong, inappropriate tendencies to overcome (new stimuli)

Model Data

<table>
<thead>
<tr>
<th>IDS</th>
<th>IDR</th>
<th>EDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>16</td>
</tr>
</tbody>
</table>

Perseverations from Frontal Lesions

<table>
<thead>
<tr>
<th>Control</th>
<th>Orbital Lateral</th>
<th>ID/ED in the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>16</td>
</tr>
</tbody>
</table>

Perseverations in the Model

<table>
<thead>
<tr>
<th>Intact</th>
<th>Feat Dim</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Explanation of lesion Data: IDS

• Intradimensional shift (IDS): different stimuli pre and post-shift; the relevant dimension (A) stays the same
  - No effect of PFC lesions
  - PFC is unnecessary because there are no strong, inappropriate tendencies to overcome (new stimuli)
**Explanation of Lesion Data: IDR**

- **Intradimensional reversal (IDR):** same stimuli pre and post-shift; initially A1 = target; after the shift A2 = target.

- Performance is impaired after PFC Dim lesions but not PFC Feat lesions.

- It's clear why PFC Dim is not important here: It involves a shift of attention within a dimension, not across dimensions.

- How does PFC Feat help performance?

  - Before shift, some hidden units learn to generate the A1 response.
  - After shift, these hidden units point to the wrong response.
  - PFC helps the model focus on other hidden units, which can then be associated with the new response.
  - This way, the model avoids having to fully unlearn the association between the original hidden units and A1 response.

**Explanation of Lesion Data: EDS**

- **Extradimensional shift (EDS):** different stimuli pre- and post-shift; initially A1 = target; after the shift, B3 = target.

- Performance is impaired after PFC Dim lesions but not PFC Feat lesions.

- It's clear why PFC Dim is important: It helps focus attention on the newly relevant dimension.

- Why can't PFC Feat serve the same function? It should be able to...

  - Without PFC Dim, PFC Feat has no sense of what constitutes a "dimension," just updates to random new pattern of features from both A and B dimensions...

**Advantages of ID/ED model**

- PFC reps are not clamped as in Stroop — updated in response to changing task demands.

- Nice fit and explanation of complex monkey data.

- Shows how working memory and cognitive control may be two sides of the same coin: activation-based memory is not just memory, but also biases activity elsewhere in the brain.

- Shows how PFC Dim and cognitive control may be.

- PFC and Feat are damaged by ID/EDs — updated in response to changing task demands.

**Performance of ID/ED model**

- Performance is improved after PFC lesion but not PFC deficit.

- Performance is impaired after PFC lesion but not PFC deficit.
Limitations of ID/ED model

• Reps not clamped, but still not learned – one to one connectivity from HL.
• Distinction between OFC = features, DLPFC = dimensions may be too convenient: observed dissociation; not much evidence of OFC-features (see Frank & Claus, 2006).

• Newer models address the issue of how PFC representations can develop in childhood and lead to higher level abstraction and generalization to new tasks (Rougier et al, 2005, PNAS).

Limitations of ID/ED model

• Doesn’t distinguish between updating and maintenance systems.
• Goal/Subgoal requires selective updating with concurrent maintenance of task relevant info.
• Updating system thought to involve the BG and DA, damaged in PD, SZ and lead to “frontal-like” impairments in Stroop, WCST, etc.
• Newer BG models address these issues in more complex tasks (eg. O’Reilly & Frank, 2006).

Goal/Subgoal Hierarchical Structure

1. Open fridge.
2. Get food items.
3. Close fridge.
4. Get bread from cupboard.

Update these subgoals to guide actions, but to guide the ordering of subgoals themselves, need to maintain overall goal of task (Make sandwich).

A Unified Activation-based Account

Executive control: Maintain top-down plans / goals over time.

Flexure: Can problem / novel categories of responses — need weight-based update to overcome prepotent categories.

Flexibility: Dynamics of activation-based more rapid than weight-based.

Inhibition: Need to maintain top-down activation for weight-based plan

Cue/Goal Hierarchical Structure

1. Open fridge.
2. Get food items.
3. Close fridge.
4. Get bread from cupboard.

Update these subgoals to guide actions, but to guide the ordering of subgoals themselves, need to maintain overall goal of task (Make sandwich).

A Unified Activation-based Account

Central frontal mechanisms:

• Activation-based working memory. Frontal neurons maintain activity over delays.
• Monitoring/evaluation e.g., Error-monitoring critical for dopaminergic modulation.

Executive control: Maintain top-down plans / goals over time.

Flexure: Can problem / novel categories of responses — need weight-based update to overcome prepotent categories.

Flexibility: Dynamics of activation-based more rapid than weight-based.

Inhibition: Need to maintain top-down activation for weight-based plan

Higher Level Cognition: What’s Missing

• Planning
• Reasoning
• Decision-making
• Emotion
• Consciousness, sense of self
• Free will

Executive control: Maintain top-down plans / goals over time.

Flexure: Can problem / novel categories of responses — need weight-based update to overcome prepotent categories.

Flexibility: Dynamics of activation-based more rapid than weight-based.

Inhibition: Need to maintain top-down activation for weight-based plan
Beyond the PFC

Bias & Binding in the PFC and Hippocampus:

**Bias**
- Familiar (Hippocampus Independent)
- Novel (Hippocampus Dependent)

**Binding**
- Automatic
  - Weak/Sustained (PFC Dependent)
  - Strong/Transient (PFC Independent)
- Novel, transient
- Routine, transient
- Novel, sustained or weak
- Familiar, sustained or weak

Controlled