Hierarchical Control over Effortful Behavior by Dorsal Anterior Cingulate Cortex

Clay Holroyd

Department of Psychology
University of Victoria
Proposal

• Problem(s) with existing theories of ACC function.
  – Putative functions minimally and inconsistently impaired by ACC damage in humans

• Whereas current theories of ACC focus on punctate, trial-to-trial changes in control (e.g., conflict processing, reinforcement learning), ACC may be more concerned with learning about the higher-level structure in behavior

Holroyd & Yeung (2011, 2012)
Proposal

- ACC learns value of and selects task
- ACC applies control to ensure that actions comply with task goals

Control

ACC: task

reinforcement

hunt wildebeests

left foot forward, right foot forward, etc.

Striatum: actions

Botvinick et al. (2009)
Shenhav et al. (2013)
Proposal

“hunt wildebeests”

“search for water”

“family time”
Modeling Work

• Why?
  – Formal instantiation of the HRL-ACC theory provides a means for explicitly
delineating the theory’s underpinning assumptions while simultaneously
demonstrating its internal consistency.

• What?
  – The model proposes a hierarchical relationship between rostral ACC, dorsal ACC
and the striatum, each of which control an effortful process on the basis of a
common set of principles carried out at differing levels of abstraction.

Holroyd & McClure
(under review)

Sam McClure

Stanford University
Modeling Work

• Focus on the consequences of ACC lesions

• But these effects are poorly understood in humans

• Examine an animal model
  – Carefully controlled, replicable behavioral data
  – Lesion, intracranial studies
  – Rat behavior is simpler than human behavior
Frontal-midline homologues

- dorsal ACC homologue in rats

Assumption: Rostral ACC = prelimbic cortex in rats?
Three Principles

- Hierarchy
- Reinforcement
- Control
Humans:

Meta-task selection: Go to park, go to library

Go to park

Task selection: Play basketball, barbeque

barbeque

Action selection: open grill, pour charcoal, etc.

Rats:

Meta-task selection: Do experiment

Do experiment

Task selection: Task strategy #1, #2

Strategy #2

Action selection: Move north, south, east, west, sit.
Reinforcement

- Reinforcement determines values associated with actions/tasks
  - **Striatal actions**: standard reinforcement learning (SARSA)
  - **ACC tasks**: average reward associated with each task
  - **Prelimbic cortex meta-task**: average reward across tasks
Control signal from the higher (meta-task/task) level suppresses an effortful cost by the lower (task/action) selection level.

- Striatal cost: effortful actions penalized
- ACC cost: task switches penalized

Prelimbic cortex and ACC exercise control over task/action selection according to a common algorithm.

In the absence of control, different levels can become decoupled.

- e.g., ACC can chose and execute an task to do homework in the library, but in the absence of control, striatum can leave the library to enjoy sunshine outside

Control regulated by feedback control loops
Simulations

• Experimental criteria
  – reveal effects of ACC, prelimbic cortex lesions
  – effects replicated multiple times

• Cross Maze task:

• Barrier Maze task:
Cross-maze task

- Rats first trained on one task (e.g., place task: phase 1)
- Then trained on second task (e.g., response task: phase 2)
Cross-maze task: ACC

- On each trial ACC selects either a task for executing either the place strategy or the response strategy, by weighing their learned average reward values against an assumed cost of switching from one task to the other.
Cross-maze task: **Striatum**

- **prelimbic cortex**
  - Meta-task selection
    - e.g., perform in environment A vs. environment B

- **anterior cingulate cortex**
  - Task selection:
    - e.g., execute place task vs. response task

- **striatum**
  - Action selection:
    - e.g., move north, west, east, south in the maze, or sit

- The **striatum** learns to navigate individual steps of the maze on the basis of separate sets of task-specific state-action values.
Cross-maze task: Prelimbic Cortex

- Prelimbic cortex learns the average reward of the experiment itself, across tasks.
- Applies a top-down control signal over ACC activity to overcome the switch cost to facilitate switching between tasks.
- Prelimbic cortex lesions impair strategy switches.
Cross-maze Results: Phase 2

- **Lesion**: prelimbic control $= 0$: switch cost not suppressed

![Performance following shift and following reversal graphs](image-url)

- **Sham**
- **PLC Lesion**
Cross-maze vs. Barrier Maze

• The cross maze task demonstrates a role for prelimbic cortex in hierarchical behavior

• But according to the theory the task itself is selected and supported by ACC, not prelimbic cortex

• Barrier Maze task illustrates the role for ACC in supporting execution of a particular task
Barrier-maze task

- Rats required to choose between a small reward with no barrier, vs. a large reward with a large barrier
Barrier-maze task

• ACC selects task (either response or place works)

• Applies top-down control over striatum to overcome effortful costs of action selection.

prelimbic cortex

Meta-task selection
  e.g., perform in environment A vs. environment B

control

anterior cingulate cortex

Task selection:
  e.g., execute place task vs. response task

control

striatum

Action selection:
  e.g., move north, west, east, south in the maze, or sit

ACC:
  “Just do it!”

striatum:
  “No way!”
• **Lesion**: ACC control = 0: No control to climb barrier.
• The ACC encodes a course of action at the level of task representation and applies sufficient control to ensure its successful completion.

• Prediction: High effortful control over one obstacle will generalize to new obstacles in the same task when ACC is intact, but not when ACC is lesioned
  – e.g., moving the barrier will confuse the striatal mechanism for action selection, but not ACC
Thank you

Sam McClure, Stanford University
Nick Yeung, University of Oxford
Stroop Task

Cohen et al (1990)
**Control**

Option Selection ↓

\[ Q_L(o_{L+1}, o_L, s) = V_L(o_{L+1}, o_L, s) - \frac{C_L(o_L, s)}{1 + \varepsilon_{L+1}} \]

Q: overall value of option  
V: ACC value of option  
C: effort of switch cost  
\( \varepsilon \): prelimbic control signal

Q: overall value of action  
V: striatal value of action  
C: effort of action  
\( \varepsilon \): ACC control signal

Action Selection ↑
Reinforcement: Values

- **Learning:** Average reward (across options)
  - Prelimbic cortex
    - Meta-option selection
      - e.g., perform in environment A vs. environment B
    - Control

- **Learning:** Average reward (separately for each option)
  - Anterior cingulate cortex
    - Option selection:
      - e.g., execute place task vs. response task
    - Control
    - Cost: effort incurred by switching between options

- **Learning:** SARSA
  - Striatum
    - Action selection:
      - e.g., move north, west, east, south in the maze, or sit
    - Cost: effort incurred by executing physical actions
• **Option selection**: select and maintain the task
• Each task comes with its own set of option-specific action values

• **Action selection**: Move north, south, east, west, sit in “gridworld”
• values learned by standard principles of reinforcement learning (SARSA)

---

Caudal/dorsal ACC

Striatum and other brain areas (e.g., hippocampus)
• “Meta-Option” selection: select and maintain the task context
  • Each meta-option comes with its own set of options
• Option selection: select and maintain the task
  • Each task comes with its own set of option-specific action values
• Action selection: Move north, south, east, west, sit in “gridworld”
  • values learned by standard principles of reinforcement learning (SARSA)

Hierarchy

Task 1

Task 2

Meta 1

Meta 2

Prelimbic cortex (rostral ACC)

Caudal/dorsal ACC

Striatum and other brain areas (e.g., hippocampus)
Cross-maze Results: Phase 2

- **A: Control Group:**
  - After rule shift, error rates increase. Because reward received is less than the average, prelimbic cortex boosts control over ACC, which reduces the switch cost and facilitates the shift to the new task.

- **B: Reversible lesion during phase 2:**
  - Prelimbic control = 0
  - Prelimbic lesions eliminate top-down control over switch cost, which stymies the switch to the new task.

- **C: Reversible lesion during phase 1:**
  - No effect over the control group.

Errors are due to perseverations on the previous task.

- Prelimbic inactivation does not impair reversal learning.
  - Mappings are re-learned within the option
  - Perseverative errors are avoided because no switch to a new task
Control

- Control is regulated by a feedback control loop
  - Control values for prelimbic cortex, ACC begin the task high
  - If rewards > value for option/meta-option, control decreases
  - If rewards < value for option/meta-option, control increases

- Feedback control loop allows system to explore state space and determine what level of reward to expect, and withdraw control when no control is needed
Hierarchical Reinforcement Learning

- Sutton et al. (2003)
  - Hierarchical representations can enhance computational efficiency for problems with hierarchical structure
  - “options” = high level actions (i.e., the task)

- Botvinick et al. (2009)
  - First to explore neural mechanisms of hierarchical reinforcement learning