Earl Miller - An Integrative Theory of Prefrontal Cortex Function (2001)

Tom Rochette <tom.rochette@coreteks.org>

July 24, 2025 — daae079c

0.1 Context

0.2 Learned in this study

0.3 Things to explore

1 Overview

2 Notes

2.1 Introduction

- The prefrontal cortex is the neocortical region that is most elaborated in primates, animals known for their diverse and flexible behavioral repertoire

2.1.1 The Role of the PFC in Top-Down Control of Behavior

- The PFC is not critical for performing simple, automatic behaviors, such as our tendency to automatically orient to an unexpected sound or movement
- The PFC is critical in situations when the mappings between sensory inputs, thoughts, and actions either are weakly established relative to other existing ones or are rapidly changing
- Two classic tasks illustrate this point: the Stroop task and the Wisconsin card sort task
  - In the Stroop task, subjects either read words or name the color in which they are written
  - In WCST, subjects are instructed to sort cards according to shape, color, or number of symbols appearing on them and the sorting rule varies periodically
- In this article, we argue that all these functions (selective attention, behavioral inhibition, working memory, rule-based or goal-directed behavior) depend on the representation of goals and rules in the form of patterns of activity in the PFC, which configure processing in other parts of the brain in accordance with current task demands
- We build on the fundamental principle that processing in the brain is competitive: Different pathways, carrying different sources of information, compete for expression in behavior, and the winners are those with the strongest sources of support

2.1.2 Overview of the Theory

- We assume that the PFC serves a specific function in cognitive control: the active maintenance of patterns of activity that represent goals and the means to achieve them
- The aggregate effect of the bias signals (provided by the PFC) is to guide the flow of neural activity along pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task

2.1.3 Minimal Requirements for a Mechanism of Top-Down Control

- The PFC must provide a source of activity that can exert the required pattern of biasing signals to other structures
- The PFC must maintain its activity robustly against distractions until a goal is achieved, yet also be flexible enough to update its representations when needed
- It must house the appropriate representations, those that can select the neural pathways needed for the task
- PFC representations must have a high capacity for multimodality and integration
- The PFC must exhibit a high degree of plasticity

2.2 Properties of the PFC

2.2.1 Convergence of Diverse Information

- One of the critical features for a system of cognitive control is the requirement that it have access to diverse information about both the internal state of the system and the external state of the world

2.2.2 Convergence and Plasticity

- Given that goal-directed behavior depends on our ability to piece together relationships between a wide range of external and internal information, it stands to reason that top-down control must come from PFC representations that reflect a wide range of learned associations
- Study of monkeys and humans with PFC damage also suggest that the PFC is critical for learning rules
  - Petrides found that following PFC damage, patients could no longer learn arbitrary associations between visual patterns and hand gestures
  - In monkeys, damage to ventrolateral area 12 or to the arcuate sulcus region also impairs the ability to learn arbitrary cue-response associations
- Learning of visual stimuli-response conditional associations is also impaired by damage to the PFC from the temporal cortex
- Passingham argues that most, if not all, tasks that are disrupted following PFC damage depend on acquiring conditional associations (if-then rules)
- In sum, these results indicate that PFC neural activity represents the rules, or mappings required to perform a particular task, and not just single stimuli or forthcoming actions
- We assume that this activity within the PFC establishes these mappings by biasing competition in other parts of the brain responsible for actually performing the task
- These signals favor task-relevant sensory inputs (attention), memories (recall), and motor outputs (response selection) and thus guide activity along the pathways that connect them (conditional associations)

2.2.3 Feedback to Other Brain Areas

- Appearance of a cue object instructed monkeys to recall and then choose another object that was associated with the cue during training
- In the intact brain, information is shared between IT (inferior temporal) cortices in the two cerebral hemispheres
- By severing the connecting fibers, each IT cortex could only see (receive bottom-up inputs from) visual stimuli in the contralateral visual field
- When Tomita et al. examined activity of single neurons in an IT cortex that could not "see" the cue, it nonetheless reflected the recalled object, albeit with a long latency
  - It appeared that visual information took a circuitous route, traveling from the opposite IT cortex (which could "see" the cue) to the still connected PFC in each hemisphere and then down to the "blind" IT cortex
  - This was confirmed by severing the PFC in the two hemispheres and eliminating the feedback, which abolished the IT activity and disrupted task performance

- Together, these findings suggest that identification of an intended stimulus relies on interactions between the PFC and the posterior cortex

2.2.4 Active Maintenance

- If the PFC represents the rules of a task in its pattern of neural activity, it must maintain this activity as long as the rule is required
- Usually this extends beyond the eliciting event and must span other intervening, irrelevant, and potentially interfering events
- The capacity to support sustained activity in the face of interference is one of the distinguishing characteristics of the PFC
- Studies have demonstrated that neurons within the PFC remain active during the delay between a transiently presented cue and the later execution of a contingent response. Such delay period activity is often specific to a particular type of information, such as the location and/or identity of a stimulus, forthcoming actions, expected rewards, and more-complex properties such as the sequential position of a stimulus within an ordered series or a particular association between a stimulus and its corresponding response
- Other areas of the brain exhibit a simple form of sustained activity
  - In many cortical visual areas, a brief visual stimulus will evoke activity that persists from several hundred milliseconds to several seconds
- What appears to distinguish the PFC is the ability to sustain such activity in the fact of intervening distractions
- Sustained activity in the PFC can maintain a sample memory across the distractors. By contrast, sustained activity in extrastriate visual areas (such as the IT and posterior parietal cortex) is easily disrupted by distractors
- Posterior cortical neurons seem to reflect the most recent input regardless of its relevance, whereas the PFC selectively maintains task-relevant information

2.2.5 Learning "Across Time" Within the PFC

- Typically, the internal representation of goals and associated rules must be activated in anticipation of the behavior they govern
- Rules often involve learning associations between stimuli and behaviors that are separated in time
- How can associations be learned between a rule or event that occurs at one point in time and contingent behaviors or rewards that occur later?
- The capacity to actively maintain representations over time is fundamental to associative learning, as it allows information about fleeting events and actions to comingle that would otherwise be separated in time
- Midbrain dopamine (DA) neurons exhibit relatively low levels of spontaneous firing but give bursts of activity to behaviorally salient events, especially the delivery of unpredicted, desirable stimuli, such as food or juice reward
- As learning progresses, however, DA neurons become activated progressively earlier in time, by events that predict reward, and cease their activation to the now-expected reward
- If the predicted reward fails to appear, activity is inhibited at the expected time of its delivery, and if the reward (or an event that has come to predict it) appears earlier than expected, it will again elicit DA neural responses
- Thus, midbrain DA neurons seem to be coding "prediction error," the degree to which a reward, or a cue associated with reward, is surprising
- The aim of the cognitive system is not only to predict reward but to pursue the actions that will ensure its procurement
- The prediction error signal could help mediate this learning by selectively strengthening not only connections among neurons that provide information about the prediction of reward, but also their connections with representations in the PFC that guide the behavior needed to achieve it

2.3 A Guided Activation Theory of PFC Function

2.3.1 A Simple Model of PFC Function

- Most neural network models that address the function of the PFC simulate it as the activation of a set of "rule" units whose activation leads to the production of a response other than the one most strongly associated with a given input

2.3.2 Guided Activation as a Mechanism of Cognitive Control

- The PFC is modulatory rather than transmissive. That is, the pathway from input to output does not "run through" the PFC. Instead, the PFC guides activity flow along task-relevant pathways in more posterior and/or subcortical areas
- This distinction between modulation vs transmission is consistent with the classic pattern of neuropsychological deficits associated with frontal lobe damage. The components of a complex behavior are usually left intact, but the subject is not able to coordinate them in a task-appropriate way

2.3.3 Updating of PFC Representations

- The mechanisms responsible for updating representations within the PFC must be able to satisfy two conflicting demands:
  - They must be responsive to relevant changes in the environment (adaptive)
  - They must be resistant to updating by irrelevant changes (robus)

3 See also

4 References