

# Accepted Manuscript

Parts to Principles: Anatomical Origins of prefrontal organization

Christopher H. Chatham, David Badre

PII: S0010-9452(12)00174-8

DOI: [10.1016/j.cortex.2012.05.013](https://doi.org/10.1016/j.cortex.2012.05.013)

Reference: CORTEX 838

To appear in: *CORTEX*

Received Date: 12 December 2011

Revised Date: 11 May 2012

Accepted Date: 11 May 2012

Please cite this article as: Chatham CH, Badre D, Parts to Principles: Anatomical Origins of prefrontal organization, *CORTEX* (2012), doi: 10.1016/j.cortex.2012.05.013.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



## PARTS TO PRINCIPLES

## TITLE:

Parts to Principles: Anatomical Origins of Prefrontal Organization

## AUTHORS:

Christopher H. Chatham\* and David Badre

\*corresponding author

Department of Cognitive, Linguistic and Psychological Sciences

Brown University

190 Waterman Street

Providence, RI 02906

David\_Badre@brown.edu

Phone: 401-863-9563

Fax: 401-863-1300

## AFFILIATION:

Department of Cognitive, Linguistic and Psychological Sciences, Brown University

RUNNING HEAD: PARTS TO PRINCIPLES

## PARTS TO PRINCIPLES

**KEYWORDS:** corticostriatal loop; nigrothalamic pathway; policy abstraction; state abstraction; working memory gating

## 1. Introduction

In “The Organisation of the Mind,” Shallice and Cooper (2011; henceforth S&C) identify “isolable cognitive subsystems” of higher-level cognition that are definable computationally, dissociable neuroanatomically, and functionally important for domains as diverse as semantic and episodic memory, working memory maintenance and manipulation, reasoning and planning, and supervisory processes and consciousness. Hence, the book provides an integrative view of the neurocognitive underpinnings of higher-level cognition, as gleaned from the cognitive neuroscience approach.

The isolable subsystems in PFC identified by S&C undoubtedly reflect differences in its intrinsic and extrinsic cortico-cortical connectivity (e.g., Yeterian et al., in press). However, as discussed below, the focal loops that interlink PFC with striatum (e.g., Middleton and Strick, 2002) are central to PFC function. Thus, the anatomical organization of these loops is also a key determinant of the functional organization of PFC subsystems.

## 2. Policy Abstraction and Dorsal PFC

As reviewed by S&C, abundant neuroimaging and neuropsychological data indicate that increasingly rostral subregions of the dorsal PFC are recruited when task demands become more abstract (e.g., Koehlin et al., 2003; Badre and D’Esposito, 2007; S&C p350-360). Although controversy surrounds the precise definition of abstraction (S&C, p.358), one framework for understanding these effects invokes the concept of policy abstraction (Badre et al., 2010; Badre and Frank, 2012; Frank and Badre, 2012).

By this view, rostral areas of dorsal PFC represent abstract state-to-action mappings – i.e., policies – specifying optimal *collections* of actions given a state or context (e.g., on Monday morning [state], make coffee [action]; Figure 1A). Caudal areas of dorsal PFC represent concrete state-to-action mappings selecting specific actions (e.g., when using the French press, grind coffee beans coarsely). This hierarchical decomposition of large state-to-action mapping problems improves learning (fewer actions/action classes are considered in solving subproblems) and generalization (subproblem solutions can be reused independently) (Botvinick, 2008).

Policy abstraction may emerge from dynamics in a series of nested, topographically-organized frontostriatal loops. Such loops are a hypothesized mechanism for selective updating of working memory: the basal ganglia “gate” task-relevant information into PFC-based working memory and keep irrelevant information out (e.g., O’Reilly and Frank, 2006). Thus, for a concrete policy where color but not shape of a stimulus determines the correct keypress, this loop would gate color (but not shape) information into working memory to select a response. Nesting these loops can allow for abstract policy: information maintained in rostral PFC can influence gating in more caudal frontostriatal loops (e.g., when shape, as maintained in rostral PFC, determines whether color or orientation is gated into the more caudal response-selection loop). Thus, the observed rostrocaudal gradient in PFC could be an emergent property of nested frontostriatal loops.

## PARTS TO PRINCIPLES

Indeed, simulations of these frontostriatal loops predict behavioral and neuroimaging results from hierarchical tasks (Badre and Frank, 2012), and hierarchical nesting in these loops accelerates learning (Frank and Badre, 2012). Moreover, high-definition diffusion spectrum tractography has provided initial evidence for this nested anatomical pattern in frontostriatal connectivity (Verstynen et al., in press). Thus, the isolable nature of dorsal PFC subsystems may emerge from their hierarchical interactions with highly topographic striatal circuits (Middleton and Strick, 2002).

But what about the ventrolateral prefrontal cortex, where rostrocaudal gradients are less commonly observed (e.g., S&C, p358)?

### 3. State Abstraction and Ventral PFC

In contrast to the highly-topographic striatal pathway innervated by dorsal PFC (involving the internal globus pallidus; GPi), the pathway innervated by ventral PFC (involving the substantia nigra reticulata; SNr; Kitano et al., 1998; Middleton and Strick, 2002) is far less topographic. These two pathways may also differ functionally, with the GPi-mediated pathway being more central to action-side processing, and the SNr pathway more involved in sensory-side processing (e.g., SNr's unique projections to inferotemporal cortex and sensory hallucinations arising from SNr lesions; Middleton and Strick, 2000). This implies that a more integrative and sensory-related ventral subsystem operates alongside a more segmented, hierarchical, and action-related dorsal subsystem.

Why might the brain complement a rostrocaudally-segmented policy abstraction hierarchy in dorsal PFC with a more integrative ventral system? Learning is enhanced not only by nesting actions into more abstract action classes (i.e., through policy abstraction), but also by collapsing across irrelevant distinctions in contextual state, so that only relevant features of the state are considered. For example, instant coffee may be preferable when running late; this feature of the state (i.e., timeliness) should be preferentially selected among less-relevant contextual features (e.g., the weather; Figure 1B) in setting the appropriate task. Similarly, many categorization problems involve learning about classes of inputs that can be similarly acted upon. Such "state abstraction" also requires the highly integrative monitoring of context and transitions, so that non-identical contexts can be treated equivalently only as long as that is itself contextually appropriate (Botvinick, 2008; Gureckis and Love, 2010).

A ventral PFC subsystem for state abstraction seems consistent with much neuropsychological data (e.g., the mnemonic selection functions ascribed to this region [S&C, p.375-8], as well as the ventral PFC foci of the state-related "monitoring" and "task-setting" functions [S&C, p367-8]). Neuroimaging offers additional support: activity in ventral PFC does not differentiate irrelevant distinctions between stimuli of the same abstract class, unlike dorsal PFC (Hon et al., 2012). Multivariate patterns in ventral PFC are so abstract as to be highly similar across fully distinct actions (e.g., response commission and response inhibition; Chatham et al., 2012) as though this area is tasked with identifying, selecting, or monitoring states themselves, largely independent of their precise mappings to action.

### 4. Conclusions

## PARTS TO PRINCIPLES

PFC subsystems may reflect in part their dissociable patterns of frontostriatal connectivity. Dorsal PFC participates in a set of GPi-mediated corticostriatal loops that preserve rostrocaudal topography and give rise to policy abstraction, operating alongside a more integrative SNr-mediated ventrofronto-striatal circuit for state abstraction. This framework provides a computationally-, neuropsychologically-, and neuroanatomically-grounded basis for identifying isolable subsystems in higher-level cognition.

## FIGURE LEGEND

Figure 1. The coffee-making example of policy abstraction (A) and its interaction with state abstraction (B).

## REFERENCES:

- Badre D and D'Esposito M. fMRI evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19(12):1-18, 2007.
- Badre D and Frank MJ. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Evidence from fMRI. *Cerebral Cortex*, 22 (3):509-526, 2012.
- Badre D, Kayser AS, and D'Esposito M. Frontal cortex and the discovery of abstract action rules. *Neuron*, 66 (2):315-316, 2010.
- Botvinick MM. Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12 (5):201-208, 2008.

## PARTS TO PRINCIPLES

- Chatham CH, Claus ED, Kim A, Curran T, Banich MT and Munakata Y. Cognitive Control Reflects Context Monitoring, Not Stopping, In Response Inhibition. *PLoS One*, 7(2): e31546, 2012.
- Frank MJ and Badre D. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 2: Computational Analysis. *Cerebral Cortex*, 22 (3):527-536, 2012.
- Gureckis TM and Love BC. Direct Associations or Internal Transformations? Exploring the Mechanisms Underlying Sequential Learning Behavior. *Cognitive Science*, 34 (1): 10-50, 2010.
- Hon N, Ong J, Tan R, and Yang TH. Different types of target probability have different prefrontal consequences. *Neuroimage*. 59(1):655-62, 2012.
- Koechlin E, Ody C, and Kouneiher F. The architecture of cognitive control in the human prefrontal cortex. *Science*, 302 (5648):1181-1185, 2003.
- Kitano H, Tanibuchi I, and Jinnai K. The distribution of neurons in the substantia nigra pars reticulata with input from the motor, premotor and prefrontal areas of the cerebral cortex in monkeys. *Brain Research*. 784 (1-2):228-38, 1998.
- Middleton FA and Strick PL. Basal Ganglia Output and Cognition: Evidence from Anatomical, Behavioral, and Clinical Studies. *Brain and Cognition*. 42(2):183-200, 2000.
- Middleton FA and Strick PL. Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cerebral Cortex*. 12(9):926-35, 2002.
- O'Reilly RC and Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*. 18 (2):283-328, 2006.
- Shallice T and Cooper R. *The Organisation of Mind*. Oxford University Press, UK, 2011.
- Verstynen T, Badre D, Jarbo K, and Schneider W. Microstructural organizational patterns in the human corticostriatal system. *Journal of Neurophysiology*, in press.
- Yeterian EH, Pandya DN, Tomaiuolo F and Petrides M. The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex*, in press.

