

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/286636889>

Prefrontal executive function and adaptive behavior in complex environments

Article in *Current Opinion in Neurobiology* · April 2016

DOI: 10.1016/j.conb.2015.11.004

CITATIONS

75

READS

577

1 author:



[Etienne Koechlin](#)

Ecole Normale Supérieure de Paris

98 PUBLICATIONS 9,978 CITATIONS

SEE PROFILE



Prefrontal executive function and adaptive behavior in complex environments

Etienne Koechlin

The prefrontal cortex (PFC) subserves higher cognitive abilities such as planning, reasoning and creativity. Here we review recent findings from both empirical and theoretical studies providing new insights about these cognitive abilities and their neural underpinnings in the PFC as overcoming key adaptive limitations in reinforcement learning. We outline a unified theoretical framework describing the PFC function as implementing an algorithmic solution approximating statistically optimal, but computationally intractable, adaptive processes. The resulting PFC functional architecture combines learning, planning, reasoning and creativity processes for balancing exploitation and exploration behaviors and optimizing behavioral adaptations in uncertain, variable and open-ended environments.

Address

INSERM, Département d'Etudes Cognitives, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France

Corresponding author: Koechlin, Etienne (etienne.koechlin@upmc.fr)

Current Opinion in Neurobiology 2016, **37**:1–6

This review comes from a themed issue on **Neurobiology of cognitive behavior**

Edited by **Alla Karpova** and **Roosbeh Kiani**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 11th December 2015

<http://dx.doi.org/10.1016/j.conb.2015.11.004>

0959-4388/Published by Elsevier Ltd.

Introduction

Adaptive behavior is critical for organisms to survive in real-world situations that are often changing. Basal ganglia in vertebrates are subcortical nuclei including the striatum that are thought to implement basic adaptive processes akin to what is usually referred to as (temporal-difference) Reinforcement Learning [1–4]. RL consists of adjusting online stimulus–action associations to the rewarding/punishing values of action outcomes. Importantly, RL is both a very simple and robust process endowing the animal with the ability to learn optimal behavioral strategies even in complex and uncertain situations [5•]. In mammals, basal ganglia further form loop circuits with the prefrontal cortex (PFC) [6] to further the flexibility and complexity of the behavioral repertoire, in essence overcoming the critical limitations of the RL processes. Here, we review recent findings from both empirical and

computational studies and outline a general theoretical framework describing the PFC function as implementing adaptive processes devoted to overcoming key RL adaptive limitations.

From reinforcement learning to adaptive planning

A first critical limitation in basic RL (also named *model-free* RL) is that behavior cannot adjust to internal changes in subjective values of action outcomes [7,8]. Consider, for instance, action A in a given situation leads to water and action B leads to food. If you are thirsty but replete, RL will reinforce action A relative to B in this situation. When the situation reoccurs, you will then select action A rather than B. If you are now hungry rather than thirsty, however, this is certainly a maladaptive behavior. The problem arises because basic RL make no distinctions between rewarding values of action outcomes and action outcomes *per se*.

Overcoming this limitation requires learning an internal model that specifies the outcomes resulting from actions, regardless of rewarding values. Learning this model is simply based on outcome likelihoods given actions and current states. This *predictive* model is thus learned besides the stimulus–action associations learned through RL (collectively named the *selective* model here). The predictive model especially enables to internally emulate RL without physically acting [5•]: This model predicts the outcomes of actions derived from the selective model, so that their rewarding values may be internally experienced according to the current motivational state of the agent (e.g. thirsty or hungry). Stimulus–action associations are then adjusted accordingly through standard RL algorithms. This emulation is commonly referred to as *model-based* RL [5•]. Behavior is thus adjusted to the agent's motivational state before acting and reflects internal planning. Model-based RL also enables to generally adapt faster than RL to external changes in action–outcome contingencies and/or outcome values [5•].

Empirical studies confirm that human behaviors cannot be fully explained by model-free RL, but instead have a model-based component. [9,10•,11]. Neuroimaging studies show that both the inferior parietal cortex and lateral PFC are involved in learning predictive models [12], with the former possibly encoding these models [13] and the latter, in association with the hippocampus, retrieving these models for emulating model-based RL [9]. Furthermore, empirical evidence argues that the

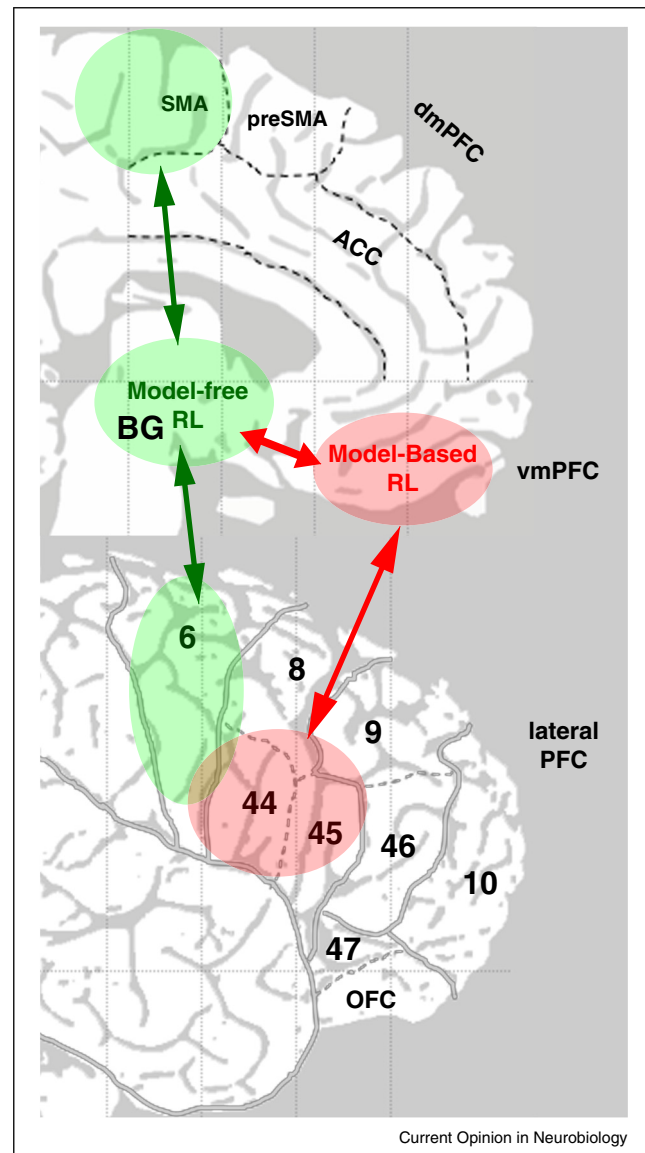
orbitofrontal cortex (ventromedial PFC in humans) in association with the striatum encode action outcomes from predictive models and their actual rewarding values [14–18,19^{*}]. Together, these studies also suggest that the ventromedial PFC may directly learn and encode simple predictive models directly mapping stimulus–action pairs onto expected valued outcomes [20^{*}], while the inferior parietal cortex and lateral PFC may be involved in implementing more complex predictive models as multi-step state–action–state maps (Figure 1).

Some authors have proposed that in the brain, model-free and model-based RL form two concurrent instrumental controllers. In this view, their relative contribution to action selection is a function of the relative uncertainty and/or reliability about reward and outcome expectations derived from selective and predictive models, respectively [21,22]. Others have proposed that model-free and model-based RL form two cooperative systems with model-free RL driving online behavior and model-based RL working off-line in the background to continuously adjust model-free RL [5^{**},10^{**},23]. Recent behavioral results support the second view [10^{**}]. As shown below, this view is also more consistent with the present theoretical framework.

From adaptive planning to Bayesian inference

A second critical limitation of RL systems described above is that adapting and learning new external contingencies gradually erases previously learned ones. This again leads to maladaptive behavior in environments exhibiting periodically recurring external contingencies (i.e. recurrent situations): RL systems have no memory and need to entirely relearn previously encountered situations. In uncertain and open-ended environments where new situations may always arise (i.e. the environment corresponds to an infinite-multidimensional space), overcoming this limitation requires solving a nonparametric probabilistic inference problem [24^{*}] for constantly arbitrating between continuing to adjust ongoing behavior through (model-free and/or model-based) RL, switching to previously learned behaviors and even creating/learning new behaviors. Previously learned behaviors along with the ongoing behavior thus form a collection of discrete entities stored in long-term memory and referred to as *task sets* [25]. Task sets are abstract instantiations of the situations the agent inferred to have encountered so far and comprises the selective and predictive model learned when the task set guided behavior [26]. Task sets further comprise an additional internal model — the *contextual* model of the likelihood of any external cues — learned when the task set guided behavior in the past [27,28], and likely encoded in lateral PFC regions [29,30]. The aforementioned arbitration problem has optimal statistical solutions based on Dirichlet process mixtures [31,32] which in practical cases, are actually

Figure 1



Reinforcement learning in the human frontal lobes. Schematic diagram showing main subcortical and cortical structures involved in reinforcement learning (RL). Green: brain regions involved in model-free RL. BG: basal ganglia. SMA: supplementary motor area. Numbers indicate broadmann's area. BA 6: premotor cortex. Red: Brain regions involved in model-based RL. vmPFC: ventromedial PFC. dmPFC: dorsomedial PFC. OFC: orbitofrontal cortex. ACC: anterior cingulate cortex. Arrows indicate critical interregional connectivity presumably underpinning RL.

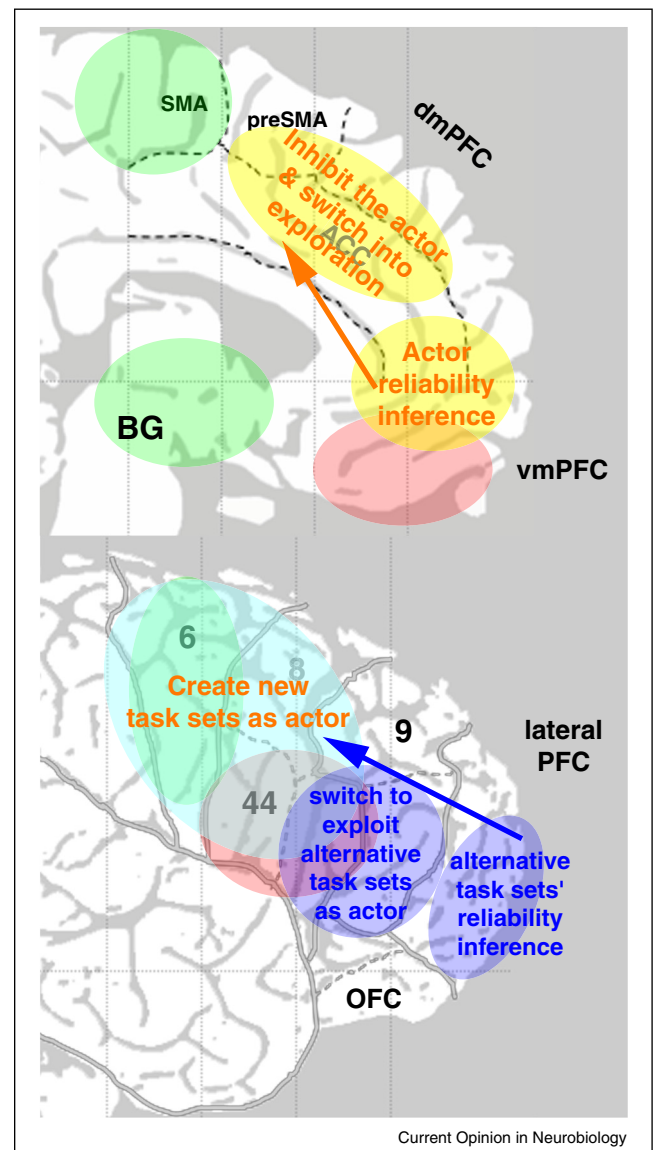
computationally intractable and consequently, biologically implausible.

Recent studies, however, show that a biologically plausible, online algorithm approximating Dirichlet process mixtures can account for human behavior in both recurrent and open-ended environments [24^{*},33^{*},34^{**}]. This

algorithm has two key features. First, it infers online the absolute reliability of the task set driving ongoing behavior (named the *actor*), i.e. the posterior probability that the actor remains consistent with current external contingencies or equivalently, that the external situation has not changed (considering a potentially infinite range of external situations). While the actor remains reliable (more likely consistent than inconsistent: reliability > 0.5), it adjusts through RL and drives behavior. Actor reliability is inferred according to the likelihood of actual action outcomes and external cues derived from actor's predictive and contextual model, respectively. Second, when the actor passes from the reliable to the unreliable status (reliability < 0.5), the situation has presumably changed and the overall algorithm passes from the exploitation to the exploration mode. A new actor is then formed as a weighted mixture of task sets stored in long-term memory, thereby creating a new task set in memory. More specifically, the new task set is built through the Bayesian averaging of internal (selective, predictive and contextual) models across stored task sets, given current external cues and action outcomes according to task-sets' contextual and predictive models. When external situations reoccur, accordingly, new task sets are created and learned, so that the number of task sets with quasi-identical models stored in long-term memory scales with the occurrence frequencies of situations. As a result, new actors are formed according to the occurrence frequencies and current external evidence associated with previously encountered situations. When external situations are new, such new actors simply learn new external contingencies in their internal models, which register in the newly created task set. Note that the new actor may be initially inferred as being unreliable but still drives behavior. Through RL, however, the new actor will subsequently reach reliability. The overall algorithm then returns to the exploitation mode: when this actor is eventually deemed as unreliable, a new actor is again created as described above.

Neuroimaging results show that the activity in different parts of the PFC co-varies with various parameters of the proposed algorithm and related algorithmic events [34^{••}] (Figure 2): (1) the PFC appears to specifically monitor the absolute reliability of actors in the ventromedial region according to action outcomes. This finding is consistent with the involvement of ventromedial PFC in encoding predictive models (see above) and subjects' confidence in their behavioral choices [35,36]; (2) algorithmic transitions from the exploitation to the exploration mode (when reliable actors become unreliable) were associated with selective transient responses in the dorsomedial PFC (dorsal anterior cingulate cortex: dACC). Consistent with this result, other fMRI studies in human subjects have provided evidence for the involvement of dACC in controlling the switch from exploitation to exploration [37]. Furthermore, extracellular recordings in rodents have

Figure 2



Prefrontal functional architecture regulating adaptive behavior beyond RL. Schematic anatomic-functional diagram of prefrontal cortex (see Figure 1 legend for labels). Yellow: PFC regions monitoring the reliability of the task set driving ongoing behavior and learning external contingencies through RL (i.e. the actor) for detecting when it becomes unreliable and switching into exploration. Cyan: PFC regions involved in creating new actor task-sets from long-term memory for driving exploration behavior. Blue: PFC regions monitoring the reliability of alternative task sets not contributing to ongoing behavior for detecting when one becomes reliable and exploiting it as actor, thereby preventing or switching out of exploration periods (in which case the newly created task set driving exploration is simply disbanded). Arrows indicate critical information flows from anterior regions involved in inferential processes to posterior regions involved in switching processes.

shown that neuronal ensembles in dACC exhibit abrupt phases transitions when animals switch from exploitation to exploration. [38,39^{••},40[•],41^{••}]. Electrophysiological recordings further suggest that the dACC enforces

switching away from the current actor at the task set level, while the pre-SMA is involved in inhibiting its subordinate elements like stimulus–action associations composing selective models [42–44].

From Bayesian inference to adaptive creativity

Key to the efficiency of the algorithm described above is that it gradually builds a repertoire of task sets in long-term memory. That repertoire is used to create new actors rather than adapting ongoing actors through RL when the current situation is inferred as having changed. Actor creation uses previously learned task sets optimally, because new actors are created according to both the recurrence frequencies of situations and current external evidence supporting associated task sets [24*,33*]. One limitation of this creation process, however, is that it builds selective models from previously learned selective models with no consideration for whether rewarding values of action outcomes may have changed since these selective models were learned. As explained above (first section), this may lead to create maladaptive new actors.

A direct solution to this problem is to assume that the process of actor creation triggers model-based RL before the new actor actually starts driving behavior [24*,33*]. Because new actors are assumed to be stored as new task sets in long-term memory, this model-based RL is better conceived as directly adjusting off-line the new actor's selective model before the latter drives behavior. This hypothesis has not been empirically tested yet but some empirical evidence supports it. First, the hypothesis predicts that when the situation changes, model-based RL predominates over model-free RL, while model-free RL gradually begins to dominate when the situation perpetuates. Consistently, previous studies provide evidence that behavioral performances predominantly reflect model-based RL after changes in external contingencies and subsequently, reflect model-free RL when the behavior becomes increasingly habitual [22]. Second, task set creation was associated with fMRI activations in lateral premotor and PFC regions along with activations in the striatum [34**,45] (Figure 2). Consistent with the hypothesis, fMRI studies have shown the striatum to be implicated in both model-free and model-based RL [46]. Third, the hypothesis implicates that the new actor is mentally emulated before the behavioral switch occurs, as medial PFC activations suggest in a recent human fMRI study [47].

From adaptive creativity to hypothesis testing

The algorithmic approach we describe here suggests that the core PFC function is to enable adaptive behavior to abruptly switch from RL-based adjustments of ongoing behavior to memory-based constructs of new behaviors. Such a switching/creation process is intrinsically non-parametric (i.e. discrete) given the presumed open-ended (i.e. infinitely multi-dimensional) nature of

the environment. With no additional adaptive mechanisms, the process remains irreversible and lacks flexibility. To overcome this limitation and more efficiently approximate optimal adaptive processes in uncertain environments [24*], the algorithm described in [33*] has a third key feature: it keeps monitoring the absolute reliability of task sets previously used as *reliable actors* along with the ongoing actor task set (thereby forming an *inferential buffer*). Whenever one monitored task set becomes reliable (at most one task set can be reliable at the same time), it becomes the actor. Whenever none are deemed reliable, a new actor is then created from long-term memory as described above. Critically, this newly created actor is rejected and disbanded, whenever it remains unreliable while one monitored task set becomes reliable. Conversely, it is confirmed and consolidated in long-term memory, whenever it becomes reliable before the other monitored task sets. This algorithm thus implements the notion of hypothesis testing bearing upon newly created task sets.

Behavioral studies show that this algorithm predicts human behavioral performances in protocols featuring uncertain, changing and open-ended environments [33*,34**]. Interestingly, the best prediction was obtained when the inferential buffer had a monitoring capacity of three/four task sets and only the actor in the buffer contribute to drive behavior. These findings match the notion and capacity of human working memory [48–50]. Moreover, neuroimaging results suggest that the PFC implements this algorithm [34**]: firstly, activations in the frontopolar PFC encode the absolute reliability of non-actor task sets monitored in the inferential buffer; secondly, lateral PFC transiently responds to algorithmic rejection events, i.e. when the algorithm rejects a newly created actor and retrieve one monitored task set to serve as actor; and finally, ventral striatum transiently responds to algorithmic confirmation events, i.e. when the algorithm confirms and consolidates the creation of new actors. These findings are consistent with the involvement of frontopolar PFC in holding on and monitoring the opportunity to switch to alternative courses of action [15,16,30,51,52], the involvement of lateral PFC in task switching [25,29] and the ventral striatum in reinforcement learning [53].

Concluding remarks

We outlined here a theoretical framework providing a unified view of PFC function as overcoming critical limitations of model-free RL processes in basal ganglia. Beyond its role in model-based RL processes, the ventromedial PFC appears to further implement Bayesian inference processes monitoring the absolute reliability of ongoing behavior based on action outcome expectations. In contrast, the dorsomedial PFC detects when ongoing behavior becomes unreliable in order to explore and create new behaviors built from long-term memory rather

than to stay adjusting ongoing behavior through RL. Additionally, the frontopolar PFC cortex appears to infer the absolute reliability of alternative behavioral strategies. The lateral PFC detects when such alternative strategies become reliable for retrieving and exploiting them to drive behavior rather than creating new behavioral strategies. This functional architecture appears to reflect the implementation of an online algorithm approximating statistically optimal but computationally intractable processes driving adaptive behavior in uncertain, changing and open-ended environments. This theoretical framework may help future research to better understand brain and neural mechanisms regulating learning and planning, reasoning and creativity processes subserving exploitation and exploration strategies in the service of adaptive behavior.

Conflict of interest statement

Nothing declared.

Acknowledgements

Supported by a European Research Council Grant (ERC-2009-AdG#250106).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Stephenson-Jones M, Samuelsson E, Ericsson J, Robertson B, Grillner S: **Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection.** *Curr Biol* 2011, **21**:1081-1091.
2. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward.** *Science* 1997, **275**:1593-1599.
3. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ: **Dissociable roles of ventral and dorsal striatum in instrumental conditioning.** *Science* 2004, **304**:452-454.
4. Doya K: **Reinforcement learning: computational theory and biological mechanisms.** *HFSP J* 2007, **1**:30-40.
5. Sutton RS, Barto AG: *Reinforcement Learning: An Introduction*. •• Cambridge, MA: MIT Press; 1998.
A major reference in the field. A profound but easily accessible, inspirational introduction to reinforcement learning from a machine learning perspective.
6. Alexander GE, DeLong MR, Strick PL: **Parallel organization of functionally segregated circuits linking basal ganglia and cortex.** *Annu Rev Neurosci* 1986, **9**:357-381.
7. Dickinson A: **Actions and habits: the development of a behavioural autonomy.** *Philos Trans R Soc Lond B Biol Sci* 1985, **308**:67-78.
8. Balleine BW, Dickinson A: **Goal-directed instrumental action: contingency and incentive learning and their cortical substrates.** *Neuropharmacology* 1998, **37**:407-419.
9. Simon DA, Daw ND: **Neural correlates of forward planning in a spatial decision task in humans.** *J Neurosci* 2011, **31**:5526-5539.
10. Gershman SJ, Markman AB, Otto AR: **Retrospective revaluation •• in sequential decision making: a tale of two systems.** *J Exp Psychol Gen* 2014, **143**:182-194.
Behavioral study providing evidence that model-free RL controls behavior, while model-based RL trains the model-free system in the background, as originally suggested in [5*].
11. Otto AR, Gershman SJ, Markman AB, Daw ND: **The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive.** *Psychol Sci* 2013, **24**:751-761.
12. Glascher J, Daw N, Dayan P, O'Doherty JP: **States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning.** *Neuron* 2010, **66**:585-595.
13. Liljeholm M, Wang S, Zhang J, O'Doherty JP: **Neural correlates of the divergence of instrumental probability distributions.** *J Neurosci* 2013, **33**:12519-12527.
14. Hampton AN, Bossaerts P, O'Doherty JP: **The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans.** *J Neurosci* 2006, **26**:8360-8367.
15. Boorman ED, Behrens TE, Woolrich MW, Rushworth MF: **How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action.** *Neuron* 2009, **62**:733-743.
16. Boorman ED, Behrens TE, Rushworth MF: **Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex.** *PLoS Biol* 2011, **9**:e1001093.
17. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ: **Model-based influences on humans' choices and striatal prediction errors.** *Neuron* 2011, **69**:1204-1215.
18. Zhu L, Mathewson KE, Hsu M: **Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning.** *Proc Natl Acad Sci U S A* 2012, **109**:1419-1424.
19. Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, Schoenbaum G: **Orbitofrontal cortex supports behavior and learning using inferred but not cached values.** *Science* 2012, **338**:953-956.
The study demonstrates that in rats, the OFC is required for inferring rewarding values based on internal predictive models.
20. Wilson RC, Takahashi YK, Schoenbaum G, Niv Y: **Orbitofrontal cortex as a cognitive map of task space.** *Neuron* 2014, **81**:267-279.
The authors review empirical evidence supporting that the OFC/vmPFC represents predictive models of action outcomes associated with ongoing behavior.
21. Daw ND, Niv Y, Dayan P: **Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control.** *Nat Neurosci* 2005, **8**:1704-1711.
22. Lee SW, Shimojo S, O'Doherty JP: **Neural computations underlying arbitration between model-based and model-free learning.** *Neuron* 2014, **81**:687-699.
23. Pezzulo G, Rigoli F, Chersi F: **The mixed instrumental controller: using value of information to combine habitual choice and mental simulation.** *Front Psychol* 2013, **4**:92.
24. Koechlin E: **An evolutionary computational theory of prefrontal executive function in decision-making.** *Philos Trans R Soc Lond B Biol Sci* 2014:369.
This study describes the evolution of PFC function from rodents to humans as the gradual addition of new inferential capabilities for dealing with a computationally intractable, basic decision problem: exploring and learning new behavioral strategies versus exploiting and adjusting previously learned ones through RL.
25. Sakai K: **Task set and prefrontal cortex.** *Annu Rev Neurosci* 2008, **31**:219-245.
26. Domenech P, Koechlin E: **Executive control and decision-making in the prefrontal cortex.** *Curr Opin Behav Sci* 2015, **1**:101-106.
27. Collins AG, Frank MJ: **Cognitive control over learning: creating, clustering, and generalizing task-set structure.** *Psychol Rev* 2013, **120**:190-229.
28. Collins AG, Cavanagh JF, Frank MJ: **Human EEG uncovers latent generalizable rule structure during learning.** *J Neurosci* 2014, **34**:4677-4685.

29. Koechlin E, Ody C, Kouneiher F: **The architecture of cognitive control in the human prefrontal cortex.** *Science* 2003, **302**: 1181-1185.
 30. Koechlin E, Summerfield C: **An information theoretical approach to prefrontal executive function.** *Trends Cogn Sci* 2007, **11**:229-235.
 31. Ferguson TS: **A Bayesian analysis of some non-parametric problems.** *Ann Stat* 1973, **1**:209-230.
 32. Teh YW, Jordan MI, Beal MJ, Blei DM: **Hierarchical Dirichlet processes.** *J Am Stat Assoc* 2006, **101**:1566-1581.
 33. Collins AG, Koechlin E: **Reasoning, learning, and creativity: frontal lobe function and human decision-making.** *PLoS Biol* 2012, **10**:e1001293.
- Detailed and comprehensive computational account of human adaptive behavior in uncertain, variable and open-ended environments in terms of statistical reasoning, and its variations across individuals.
34. Donoso M, Collins AG, Koechlin E: **Human cognition. Foundations of human reasoning in the prefrontal cortex.** *Science* 2014, **344**:1481-1486.
- This model-based fMRI study, central to the present account, characterizes the algorithmic architecture of the PFC combining RL, Bayesian inference, and hypothesis testing in regulating adaptive behavior. See also [33*].
35. De Martino B, Fleming SM, Garrett N, Dolan RJ: **Confidence in value-based choice.** *Nat Neurosci* 2013, **16**:105-110.
 36. Lebreton M, Abitbol R, Daunizeau J, Pessiglione M: **Automatic integration of confidence in the brain valuation signal.** *Nat Neurosci* 2015, **18**:1159-1167.
 37. Kolling N, Behrens TE, Mars RB, Rushworth MF: **Neural mechanisms of foraging.** *Science* 2012, **336**:95-98.
 38. Durstewitz D, Vittoz NM, Floresco SB, Seamans JK: **Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning.** *Neuron* 2010, **66**:438-448.
 39. Hayden BY, Pearson JM, Platt ML: **Neuronal basis of sequential foraging decisions in a patchy environment.** *Nat Neurosci* 2011, **14**:933-939.
- This electrophysiological monkey study reports that in foraging, ACC neurons encoded the relative value of staying for exploiting the current depleting patch, versus switching away for exploring a new patch. This decision variable increased up to a threshold triggering exploration. The study reveals neuronal mechanisms triggering behavioral switches.
40. Karlsson MP, Tervo DG, Karpova AY: **Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty.** *Science* 2012, **338**:135-139.
- This study shows abrupt transitions in neural ensembles in rat medial PFC associated with switching away from exploitation. Switches triggered exploration phases consisting of active resampling of external contingencies.
41. Tervo DG, Proskurin M, Manakov M, Kabra M, Vollmer A, Branson K, Karpova AY: **Behavioral variability through stochastic choice and its gating by anterior cingulate cortex.** *Cell* 2014, **159**:21-32.
- Using circuit perturbations in transgenic rats, the authors demonstrate that switching between exploitation versus exploration behavioral modes is controlled by locus coeruleus input into ACC.
42. Isoda M, Hikosaka O: **Switching from automatic to controlled action by monkey medial frontal cortex.** *Nat Neurosci* 2007, **10**:240-248.
 43. Matsuzaka Y, Akiyama T, Tanji J, Mushiake H: **Neuronal activity in the primate dorsomedial prefrontal cortex contributes to strategic selection of response tactics.** *Proc Natl Acad Sci U S A* 2012, **109**:4633-4638.
 44. Bonini F, Burle B, Liegeois-Chauvel C, Regis J, Chauvel P, Vidal F: **Action monitoring and medial frontal cortex: leading role of supplementary motor area.** *Science* 2014, **343**:888-891.
 45. Badre D, Kayser AS, D'Esposito M: **Frontal cortex and the discovery of abstract action rules.** *Neuron* 2010, **66**:315-326.
 46. Wunderlich K, Dayan P, Dolan RJ: **Mapping value based planning and extensively trained choice in the human brain.** *Nat Neurosci* 2012, **15**:786-791.
 47. Schuck NW, Gaschler R, Wenke D, Heinze J, Frensch PA, Haynes JD, Reverberi C: **Medial prefrontal cortex predicts internally driven strategy shifts.** *Neuron* 2015.
 48. Cowan N: **Working-memory capacity limits in a theoretical context.** In *Human Learning and Memory: Advances in Theory and Applications*. Edited by Izawa C, Ohta N. Erlbaum; 2005: 155-175.
 49. Oberauer K: **Declarative and procedural working memory: common principles, common capacity limits?** *Psychol Belg* 2010, **50**:277-308.
 50. Risse S, Oberauer K: **Selection of objects and tasks in working memory.** *Q J Exp Psychol* 2010, **63**:784-804.
 51. Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J: **The role of the anterior prefrontal cortex in human cognition.** *Nature* 1999, **399**:148-151.
 52. Koechlin E, Hyafil A: **Anterior prefrontal function and the limits of human decision-making.** *Science* 2007, **318**:594-598.
 53. Liljeholm M, O'Doherty JP: **Contributions of the striatum to learning, motivation, and performance: an associative account.** *Trends Cogn Sci* 2012, **16**:467-475.