A Natural History of Skills

Thomas Boraud, Arthur Leblois, Nicolas P. Rougier

HAL Id: hal-01874690

https://hal.inria.fr/hal-01874690

Submitted on 14 Sep 2018

A Natural History of Skills

Thomas Boraud[1,2,3,4,*], **Arthur Leblois**[1,2,3,*], **Nicolas P. Rougier**[2,5,6,*]

[1]CNRS, UMR 5293, IMN, 33000 Bordeaux, France
[2]University of Bordeaux, UMR 5293, IMN, 33000, Bordeaux, France
[3]CNRS, French-Israeli Neuroscience Lab, 33000 Bordeaux, France
[4]CHU de Bordeaux, IMN Clinique, 33000 Bordeaux, France
[5]INRIA Bordeaux Sud-Ouest, 33405, Talence, France
[6]LaBRI, University of Bordeaux, IPB, CNRS, UMR 5800, 33405 Talence, France
[*]These authors contributed equally to this work.

Abstract

The dorsal pallium (*a.k.a.* the cortex in the mammals) makes a large loop circuit with the basal ganglia and the thalamus known to control and adapt behavior but the who's who of the functional roles of these structures is still debated. Influenced by the Triune brain theory that was proposed in the early sixties, many current theories propose a hierarchical organization on the top of which stands the cortex to which the subcortical structures are subordinated. In particular, habits formation has been proposed to reflect a switch from conscious on-line control of behavior by the cortex, to a fully automated subcortical control. In this review, we propose to revalue the function of the network in light of the current experimental evidence concerning the anatomy and physiology of the basal ganglia-cortical circuits in vertebrates. We briefly review the current theories and show that they could be encompassed in a broader framework of skill learning and performance. Then, after reminding the state of the art concerning the anatomical architecture of the network and the underlying dynamic processes, we summarize the evolution of the anatomical and physiological substrate of skill learning and performance among vertebrates. We then lay out our hypothesis that the development of automatized skills relies on the BG teaching cortical circuits and is actually a late feature linked with the development of a specialized cortex or pallium that evolved in parallel in different taxa. We finally propose a minimal computational framework where this hypothesis can be explicitly implemented and tested.

1. Introduction

The frontal cortex, the basal ganglia (BG) and the thalamus are associated into a tripartite functional loop known as the cortex-basal ganglia loop (CBG loop). Since the early 90s, we roughly subdivide this broad network into 3 subparts, from the central sulcus to the frontal pole: a motor loop, a cognitive loop and a limbic loop. This terminology has been coined by human/primate anatomists, physiologists and clinicians according to the functions associated to the corresponding cortical area. The motor loop (a.k.a. the extrapyramidal loop for the anatomists) drives voluntary movement and learning, the cognitive loop performs planning and decision making, while the limbic loop is in charge of emotions and mood. Despite different functional roles, the similarities in the architecture of these loops bring to the idea that they share the same dynamical properties and process neural information along similar mechanisms. However, if everybody agreed that these processes lie ultimately upon a focal activation of a specific cortical area that triggers an action/decision/emotion (at least in primates), there are many discrepancies in the literature about the who's who (what's what?) of the relationship between neural substrate and functions (Daw *et al.*, 2005; Graybiel, 2008; Pasupathy and Miller, 2005; Samejima and Doya, 2007).

The question is insidiously overshadowed by the triune brain theory (1973) proposed by Paul D. MacLean in the 1970s. Briefly, this theory assumes that the brain evolved in three stages, each of them adding new behavioral features to the ethogram (the inventory of behaviors) of the species granted with this upgrade. At first, appeared a reptilian brain (grossly the diencephalon and the BG) that brings instinctual behaviors (aggression, dominance, territoriality, ritual displays, etc.). Then, the paleo-mammalian complex (limbic system) involved in feeding, reproductive and parental behaviors in animals granted with the ability to display motivation and emotion. Finally, species at the top of the evolutionary tree were gifted with a neo-mammalian brain that allowed them to perform abstract thinking, planning, and for the most advanced of them language. Each level inhibited the lower one and the key to human neuropsychiatric pathology stems from defects in these inhibition processes. The triune brain theory was widely diffused because of its elegant adequacy between structure and function and vague similarities with Freud's psychoanalysis theories. It became familiar to a broad popular audience thanks to authors such as Carl Sagan or Arthur Koestler. Even if the triune brain theory is long considered as outdated (Hodos and Butler, 1997; Striedter, 2005), it is still influential in psychology. For our concern, it influenced the neurobiology of decision making, by establishing 3 dogmas: i) there is a hierarchical classification of higher brain functions; ii) each function is underlain by a specific anatomic structure and iii) structures compete against each other in order to produce a

behavior. All the theories that have been proposed to explicit the respective role of the different components of the CBG loop (see below) encompassed these 3 features and most (but not all of them) considered that the cortex must be in charge of the more elaborated brain functions and sometime even is the ultimate umpire of the competition. This conceptual framework gave birth to a flourishing literature with partially overlapping concepts opposing automatism to voluntary action, habits to goal oriented behavior, model free *vs* model-based decision making, etc.

In this review, we propose to step aside from this conceptual background in order to revalue the function of the CBG loop. We will first briefly review the current theories and show that they could be encompassed in a broader framework of skill learning and performance. Then, after reminding the state of the art concerning the anatomical architecture of the network and the underlying dynamic processes, we will summarize the evolution of the anatomical and physiological substrate of skill learning and performance among a few classes of vertebrate. We will lay out our hypothesis that the development of automatized skills relies on the BG teaching cortical circuits and is actually a late feature linked with the development of a specialized cortex or pallium that evolved in parallel in different taxa. We will confront it to other current hypotheses on the function of CBG circuits. Finally, we present a minimal computational framework were this hypothesis can be explicitly implemented and tested.

1.1. The function(s) of the forebrain

As we stated above, most of the theories proposed lie in the contrast between a deliberative behavior that necessitates constant updating and an automatic process. Influenced by psychology, this automatic process has been coined a habit.

The first theory stands that the cognitive CBG loop supports deliberative/planned behavior, while the motor one support habits (Belin *et al.*, 2009; Yin and Knowlton, 2006; Yin *et al.*, 2005). The former dominates the latter acting in new context and relying secondary on the more rustic motor system when things become boringly repetitive.

In the second theory, the opposition is between the prefrontal cortex (PFC) in charge of deliberative processes and the BG supporting habitual behaviors (Daw *et al.*, 2005; Daw *et al.*, 2006).

The third hypothesis is more innovative as it reverses the hierarchy and emphasizes learning rather than habits *per se*. It posits that the BG drive learning in the cortical areas and become increasingly less

engaged in action performance decision making as the task is learned (Graybiel, 2008; Helie *et al.*, 2015; Piron *et al.*, 2016; Samejima and Doya, 2007). Our theory, which is derived from this latter one, encompasses the concept of habits in to the broader concept of skills as it'll be explained below.

1.2. Habits, automaticity & skills

When the literature about the neural anatomy of the decision-making process comes under scrutiny, it is important to keep in mind that authors coming from different background use behavior with different meaning. Those coming from a behaviorist background have a very precise definition for habits but consider goal-oriented decision making as a kind of default mode response (Balleine and Dickinson, 1992; Balleine and O'Doherty, 2010; Dickinson, 1985). For the ones inspired by economics (a trend often dubbed neuroeconomics), goal-oriented decision making implies deliberative choice while habits are unconscious automatisms acquired slowly after intensive training that is less systematically operationalized (Graybiel, 2008; Seger and Spiering, 2011).

Instead of focusing on the competition process between habitual/automatic and goal-oriented behaviors during a decision at a given time, we propose to reframe the question around the dynamics of skill acquisition. Each decision reflects the degree of skill expertise acquired by a subject in a specific field. According to this framework, explorative, vicarious behaviors are the trace of an early stage of skill acquisition while automatic stereotyped ones appear once skills are consolidated. But the process is continuous rather than binary and the border between the two extreme forms of behavior is fuzzy. Another important consequence is that we can identify cases in which the skill is never automatized by lack of time or because the anatomical architecture doesn't allow it.

We propose then i) to substitute for habits the broader concept of skills and ii) to include an evolutionary dimension.

Motor skills, as opposed to habits, do not necessarily occur repeatedly over time (Graybiel, 2008), but involve many features of habits: (i) they are learned, and they become (ii) automatic, (iii) stereotyped, and (iv) inflexible after extensive training. We thus voluntarily step away from the dependence of habits on reward devaluation and its definition as opposed to goal-directed behavior. We finally propose to broaden the definition of skills to the cognitive and limbic network, encompassing the same properties as illustrated on the Table 1.

	Flexible learning phase (acquisition/exploration)	**Inflexible stereotyped phase** (automatization/exploitation)
Motor	Skill acquisition Playing behavior Motor Babbling	Stereotyped gestures Motor repertoire
Cognitive	Deliberative decision Causality experience Exploratory behavior	Heuristics Cognitive bias & expertise Attention & Memory
Limbic	Hedonistic exploration Taste learning Emotional lability	Preferences & sexual orientation "Tastes" Empathy

Table 1: classification of skills according to functional domains and phase.

The migration from the flexible learning phase to the inflexible stereotyped phase can be broken down into three stages (Fitts, 1964): the cognitive stage (early acquisition), the associative stage (smoothing performances) and the autonomous stage (continuous improvement). From our perspective, the two first phases correspond to the flexible learning phase that is characterized by an initial high level of error (early acquisition) and correspond to the early exploration of the parameter space, until a gradual reduction of errors occurs (smoothing performances). The late phase (continuous improvement) corresponds to the full and automated exploitation of the initial learning characterized by a very low level of errors in the absence of any feedback. This also corresponds to the proposal by (Anderson, 1982) where the author considers only two stages, a declarative stage and a procedural stage for the acquisition of cognitive skills.

2. Multiple systems at work

Before describing the natural history of skills, it may be useful to start with a summary of the brain structures involved in primates (including humans) and to describe their evolution starting from the most ancient vertebrates.

2.1. The mammalian cortex

In mammals, the brain circuits underlying skill learning and execution encompass orbito-frontal cortex (OFC), Prefrontal Cortex (PFC), Anterior Cingular Cortex (ACC), and motor cortical areas: Premotor Cortex (PMC), Supplementary Motor Area (SMA) and Motor Cortex (MC) *per se*. They are organized following a hierarchical manner: OFC is involved in valuation and manipulation of abstract concepts (Daw *et al.*, 2006; Samejima and Doya, 2007), PFC and ACC are involved in deliberative processes and the motor areas are involved in the expression of behaviors (Alexander and Crutcher, 1990b) and are the major interface between the brain and the musculoskeletal system. However, this interfacing is highly variable from one species to another. For example, concerning the motor areas, pyramidal neurons reach directly spinal motoneurons in a significant proportion only in primates, but not in rodents or cats (Lemon and Griffiths, 2005). They indeed act more as a modulatory system than a driving system in the latter; decorticated rats (Whishaw *et al.*, 1981) or cats (Bjursten *et al.*, 1976) can still behave quite normally, unlike decorticated primates (Travis and Woolsey, 1956).

The cortex is subdivided into 6 layers of neurons and comprises dozens of different types of cell types. The cortex is traditionally divided into multiple areas, which vary based on their cytoarchitecture, their inputs, their neurochemical signatures, the target of their projections and their function. At the microscopic level, each area can be divided into functional units, the cortical columns, which contain neurons that share the same receptive fields (for sensory areas) or behavioral tuning (for motor and cognitive areas). Briefly, the output of each column is provided by the output neurons, the glutamatergic pyramidal cells. On one hand, direct lateral excitation between pyramidal cells from the same or different columns provide the cortical columns with a mechanism for self and mutual excitation respectively. On the other hand, feedforward inhibition mediated by GABAergic inhibitory interneurons allows for lateral inhibition within and between columns (Horton and Adams, 2005). As the spread of lateral inhibition is thought to overcome that of the lateral excitation, the net effect of lateral interactions between columns likely tends toward mutual inhibition, allowing competition mechanisms between the columns that can drive behavior (Adesnik and Scanziani, 2010).

2.2. The thalamus and the basal ganglia

The sub-cortical territories of the skill learning network include the anterior thalamus and the BG. Along a canonical model that was formalized back in the late 1980s, cortical and subcortical structures are organized into functional feedback loops (Figure 1) that influence the dynamic properties of the cortical areas (Albin *et al.*, 1989; Alexander and Crutcher, 1990a; Alexander and Crutcher, 1990b; Parent and Hazrati, 1995a, b). The pyramidal neurons of the cortex send axons to the Striatum and the Subthalamic Nucleus (STN), which are the two main input structures of the BG. The former is made of several different populations of neurons of which the GABAergic medium spiny neurons (MSN), which represent more than 90% of the striatal neural population, are the outputs. The latter is made of at least one population of glutamatergic neurons. The main output structures of the BG are the internal segment of the Globus Pallidus (GPi) and the pars reticulata of the Substantia Nigra (SNr). Both are GABAergic structures. They send their inhibitory outputs mainly to the anterior thalamic nuclei (ventro-lateral and ventral anterior in primates) and also to the brainstem nuclei. The thalamus connects back to the cortex hence closing the loop with glutamatergic connections. The dynamic properties of the network are controlled by three pathways. The direct pathway conveys the cortical inputs through the striatum and the GPi/SNr. The indirect pathway is tri-synaptic, encompassing the striatum, the external part of the globus pallidus (GPe), the STN and then the GPi/SNr. The hyperdirect pathway corresponds to cortical input directly to the STN and then to the GPi. The direct pathway exerts a positive feedback on cortical area and the hyper-direct and indirect pathways exert a negative feedback (Mink, 1996; Nambu *et al.*, 2000; Nambu *et al.*, 2002).

Despite a topographic organization into at least three parallel functional loops (Alexander *et al.*, 1986), controlling, respectively, limbic, cognitive and motor processes, the system also shows converging and diverging features. The reduction of the number of neurons from cortex to BG output in primates, for example, implies that one GPi neuron is influenced by more than one million cortical neurons (Parent and Hazrati, 1995a, b). This high convergence allows for the integration of a multitude of features in order to generate an appropriate behavior. On the contrary, the STN connections to the GPi are sparse and divergent. Because the STN is the common feature of the two negative feedback loops (hyperdirect and indirect), it makes this small structure the bottleneck of a powerful lateral inhibition system (Leblois *et al.*, 2006a; Nambu *et al.*, 2002). Indeed, even far-apart cortical areas may inhibit each other through the divergent cortico-subcortico-cortical pathway, and the spatial extent of the

resulting mutual inhibition likely overcomes by several orders of magnitude the extent of lateral inhibition provided by cortical inhibitory interneurons. This classical vision should be updated with further connections inside BG (see Figure 1, in grey): the GPe also sends direct connections to the SNr, the GPi and with a specific population recently identified back to the striatum (Mallet *et al.*, 2012; Mallet *et al.*). The anatomy of the striatum is also rendered further complex by various populations of GABAergic and cholinergic interneurons that make it more an integrative structure than a simple input layer (Parent and Hazrati, 1995a; Silberberg and Bolam, 2015). We already demonstrated (Box 1) that the CBG network encompasses all the necessary features to allow for the selection of motor programs out of the dynamic competition between the feedback loops (Leblois *et al.*, 2006a). We also show that intrinsic noise is enough to shape motor output in new conditions without the necessity of pre-existing bias (Guthrie *et al.*, 2013), making the system more efficient for exploration in new environments (Guthrie *et al.*, 2013; Topalidou *et al.*, 2015).

Outside the BG, there are other features that are often underrated and may significantly contribute to the dynamics of the CBG loop. Direct connections between the thalamic nuclei and the BG shortcut the cortical-subcortical loop. This second feedback loop possesses similar dynamic properties as the longer one (see Figure 1). The cortex also exerts a positive feedback to the thalamus creating another positive feedback loop outside the BG. It is extremely difficult to understand the dynamics of neuronal activity within these closed loop circuits with classical word models or "box and arrow" models. As an example, we have shown in the past that the closed loop structure within the BG-thalamo-cortical networks underlies the normal and pathological dynamics of neuronal activity and can only be revealed using the tools of non-linear dynamical theory (Leblois et al., 2006). For this reason, we will develop a theoretical framework to study the cooperation of the cerebellum and BG.
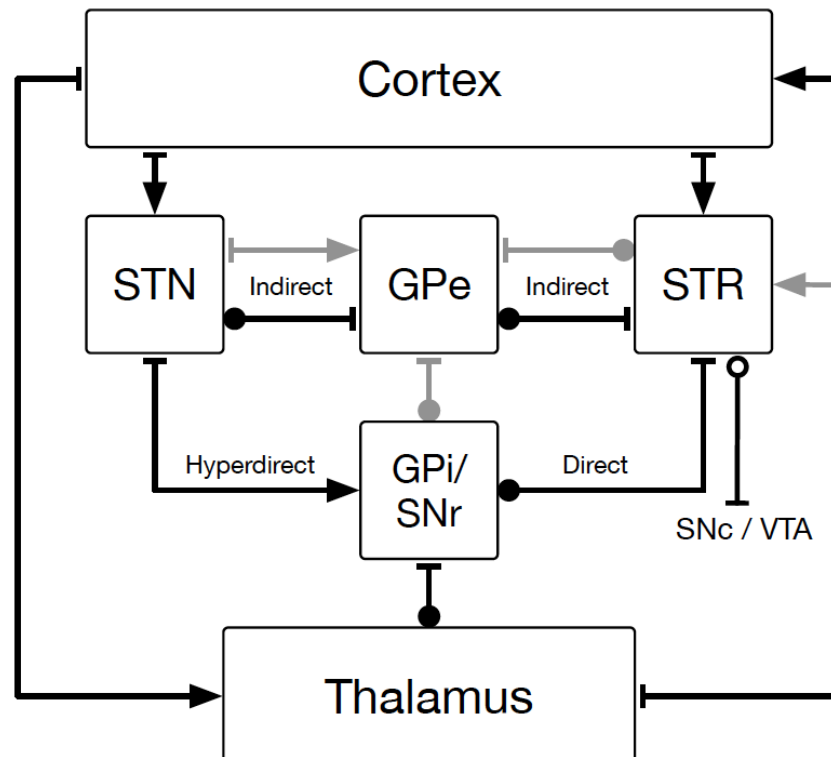
Figure 1. Schematic view of the cortico - basal ganglia - thalamus network (CBG) in primates. In black are showed the connections considered in the "classical" model: the BG works a three layers network - with 2 inputs (STN and STR), 2 outputs (GPi/SNr) and 1 intermediary layer (GPe) - encompassed in a broad feedback loop to the cortex through the thalamus. In grey are showed connections that have been evidenced or emphasized in the last decade. GPe: Globus Pallidus pars Externa; GPi: Globus Pallidus pars Interna; SNr: Substantia Nigra pars Reticulata; STN: Subthalamic nucleus; STR: Striatum.

2.3. Skill learning and synaptic plasticity

Considering the organized cortex on one side and the powerful and versatile cortex-BG-thalamus loop on the other side, we have two systems that are able to drive actions and skills on their own. The former is able to perform well learned skills even if disconnected from subcortical influence from the BG (Andalman and Fee, 2009; Desmurget and Turner, 2010; Piron *et al.*, 2016; Turner and Desmurget, 2010), while the latter can also provide a sufficient substrate for many complex behaviors (Bartus and Levere, 1977; Jaldow *et al.*, 1989; Kukleta and Libouban, 1967; Oakley, 1981). Why has such redundancy been selected through the evolutionary process? What are the advantages? What are the costs? A first hint may be provided by the plasticity underlying the learning process in each of the two systems.

2.3.1 Hebbian learning. Even if there is a very diffuse DA innervation of the prefrontal cortex, it is supposed that at this level, plasticity relies mainly on Hebbian learning (HL) principles (Keysers and Perrett, 2004; Song *et al.*, 2000). If two synaptically connected cortical neurons are co-activated, they reinforce the synaptic weight of the connection between them. It follows that repetition of co-activation will favor the genesis of neuronal assemblies that can be re-activated together later. If this connection is made through inhibitory neurons, it reinforces competition mechanisms. This system is simple and robust, but because of its insensitivity to the outcomes, it lacks flexibility.

2.3.2. Dopamine and reinforcement learning. Since the seminal works of Wolfram Schultz and his colleagues, it is known that dopamine (DA) is a neural substrate of the reward prediction error signal in the reinforcement learning (RL) process (Apicella *et al.*, 1992; Fiorillo *et al.*, 2003; Schultz, 1992). Briefly the dopaminergic neurons of the pars compacta of the Substantia Nigra and the ventral tegmental area (VTA) are able to encode the difference between the outcome of an action and the expectation of the subject. This signal is used by the target of the dopaminergic nuclei to modify their synaptic weight and therefore increase or decrease the probability to choose the same action in the next occurrence (Schultz, 2006; Schultz *et al.*, 1997). This synaptic plasticity explains why performance increases by trial and error. The synapses between the pyramidal neurons and the MSN of the striatum are the major target of DA output in the brain and therefore it is where this plasticity is hypothesized to occur (Schultz, 2006; Schultz *et al.*, 1997).

3. Parallel emergence across evolution

A picture starts to emerge with a cortical system that is able to create robust neuronal assemblies that drive behavior, displays the necessary dynamics to select a single behavioral output and decide on its own, but relies on a rather simple and unsupervised plasticity rule, and a subcortical system allowing more flexible behavioral adaptation thanks to more complex learning rules, but which involves a much larger and therefore slower and more energy demanding network. We are already far from a ventral deliberative system and an automatic dorsal one. Beyond the anatomical and functional organization of the brain, the triune brain hypothesis relies on a simplistic conception of evolution as a ladder (*scala*

naturae), with the vertebrate brains going from "primitive" brains in lower vertebrates to "evolved" brains in mammals and ultimately humans, and the progressive and linear acquisition of more and more complex cognitive traits and behaviors along evolution. In the following, we will argue that in many vertebrate taxa, brains have evolved alongside to give rise to complex behavior. We will hereafter gather evidence that while an increase in the diversity and complexity of behaviors appears to require a growth of the dorsal pallium (DP), the control and learning of complex skills always rely on the interaction between the growing pallium and the "reptilian brain", namely the BG loops.

A look at vertebrate evolution and the parallel development of the various brain areas involved in motor control and skill learning may indeed shed light on the division of labor at play during complex behaviors. Vertebrate orders with quite distant common ancestors such as fishes, birds and mammals display contemporary species with complex skills such as avoiding predators, catching prey, communication, tool use and design, cooperation, etc. (Bird and Emery, 2009a, b; Gunturkun, 2012; Patton and Braithwaite, 2015). These behaviors require planning and the combination of multiple elementary movements into complex motor skills. Interestingly, the size of the pallium has increased in all three aforementioned orders, and the most gifted contemporary species display a larger pallium/cortex to brain mass ratio (Figure 2). This high ratio is somehow correlated to the number of neurons in the DP but the relationship is far from linear (Herculano-Houzel, 2011a, b; Herculano-Houzel *et al.*, 2006; Wullimann and Vernier, 2007).
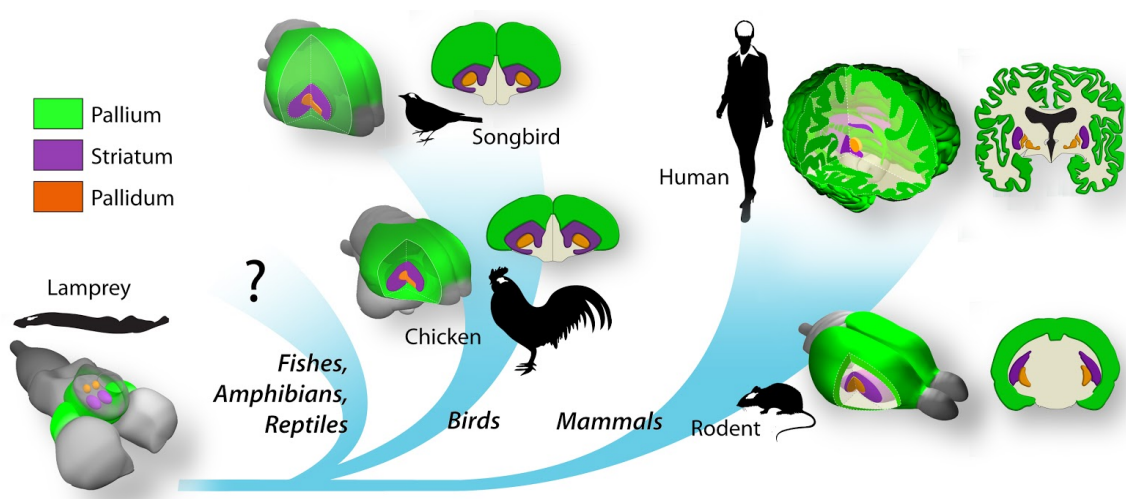

Figure 2. Comparative evolution of the striatum and pallium in vertebrates. the ratio of the brain mass devoted to the pallium increase in parallel in various vertebrates' taxa.

3.1. Lamprey

The lamprey, the oldest of all the vertebrates, appeared about 560 million years ago. Its relatively simple nervous system makes it a model of choice for studying networks involved in locomotor activity. The lamprey is devoid of fins and moves by undulating its body composed of a series of myotomes. Each is controlled by a couple of glutamatergic generators. When a motor command is sent into the cord from the Mesencephalic Locomotion Region (MLR) of the animal, a rhythmic activity is propagated in the different segments of the animal with a rostro-caudal shift generating a sinusoidal movement of the body of the animal which allows its displacement through the water. The MLR is somatotopically organized, thus, if a differential stimulus is exerted on the two sides, the animal will turn in the direction of the one stimulated most effectively. The animal can thus be piloted from this structure that receives, among other things, inputs from the diencephalon and especially from the thalamus. The latter allows for the interfacing between the sensory stimuli (visual, auditory, olfactory) and the motor system and receives massive inputs from the BG.

Even if BG are not as individualized as in the mammals, they are nonetheless composed of the same populations of neurons all intermingled in a single basal ganglia structure: MSN-like, GPi-like and GPe-like GABAergic cells, STN-like glutamatergic neurons and various interneurons (Grillner *et al.*, 2013). The functional organization is also very similar to the one of the mammals: MSN-like and STN-like populations receiving inputs from the thalamus and GPi-like sending output to the thalamus and the MLR. Between the input layer and the output, there are the three pathways described above: the direct, the indirect and the hyper-direct. Therefore, in lamprey, the thalamus and the BG are organized into a feedback loop very similar to the one described in more recent vertebrates (see Box 1). From this overall anatomical and physiological similarity, we can assume that the dynamical properties of the network are fundamentally the same. It implies that the competition mechanisms that can elicit the selection of motor programs (Leblois *et al.*, 2006a) already exist in the lamprey and can drive motor behavior of the animal. Like mammals, the lamprey brain possesses dopaminergic nuclei that innervate thalamic-MSN synapses and an associated rudimentary limbic system able to inform the dopaminergic system about the rewarding value of actions. Even if it has not been demonstrated yet, it implies that lamprey already possesses the components of the RL process (Stephenson-Jones *et al.*, 2012).

The most crucial difference between lamprey and mammals is the lack of a cortex. Instead they have a DP that represents only a very small fraction of the brain total mass. It is maybe already involved in locomotion behaviors as recently suggested (Ocana *et al.*, 2015), and sends numerous input to the BG. However, the role of the pallium in behavior in the lamprey is likely limited. Beside its small side, the DP of the lamprey is not as well organized as the mammalian cortex. It consists mostly of excitatory neurons organized in 3 layers and it is devoid of the columnar organization and therefore it is highly

unlikely that this rudimentary pallium can offer a consistent selection mechanism, in contrast with mammalian cortex (Pombal *et al.*, 2009; Pombal and Puelles, 1999).

3.2. Fishes, frogs & reptiles

Lampreys belong to the informal group of the anamniotes (a.k.a. lower-vertebrates) that also includes the fishes and the amphibians. For many species in this group, the ethograms are quite limited: foraging, mating, fight and flight. There are differences between species in nuptial behavior or the mode of foraging (some are herbivores, others are carnivores) but, more importantly, some of the more recently appeared species have developed much more elaborated motor and cognitive skills, including nest building (Kawase *et al.*, 2013), predation or escape strategies, or communication (Patton and Braithwaite, 2015). The emergence of such complex behaviors correlates with the growth of the pallium in several fish clades (Vernier, 2017). Interestingly, the development of the pallium during embryogenesis follows a different path in different fish clades: in one case (actinopterygians), the pallium develops by eversion of the cerebral hemispheres, while in another (sarcopterygians, from which tetrapods and then reptiles and mammals evolved) the cerebral hemispheres are shaped by an invagination process. These two strategies for increasing pallial size appeared independently during evolution, supporting the idea of parallel development of the DP in various phylogenetic branches that independently allowed the emergence of complex behaviors. The increase in pallial size leads to a wide distribution of the coefficient of encephalization among fishes, significantly overlapping with that of birds and mammals, even if the main part remains lower (Wullimann and Vernier, 2007). As in the case of the lamprey, the central role of the Thalamus-BG loop in foraging has been confirmed in some of these species (Finkenstadt, 1989; Patton and Grobstein, 1998). Its role in other behaviors remains to be demonstrated, but it is highly likely.

The term "reptiles" refers to several lines of different origins distributed in more than 9000 species. The first reptiles appeared around 340 million years (Benton and Donoghue, 2007). They exhibit a greater diversity of behavior than anamniotes and for some species such as crocodiles and lizards, a relatively high encephalization quotient (the ratio of brain weight to body weight) and more elaborate social behaviors have been described (Brumm and Zollinger, 2017; O'Connell and Hofmann, 2011). In these species, the DP still represents less than 20% of the total cerebral mass and the structure does not differ significantly from that of the anamniotes.

3.3. Mammals

Mammals appeared around 220 million years ago. There are approximately 5,400 species in this class that colonized deep water, land and air and it is among this class that the richest and most varied ethograms can be found, showing complex social behaviors (including highly hierarchical organization) and unprecedented cognitive performances. Two behavioral characteristics, linked to each other, distinguish them from other vertebrates. The first one is a period of dependence of young mammals that can last several years. The second is play. During the juvenile period, much of the time saved on the search for food is used to play (Bekoff *et al.*, 1980). Although the function of this activity is still debated, it certainly has a central role in skill-learning processes including foraging and social behaviors. There is, moreover, a correlation between the complexity of the social organization of a species, the length of the juvenile period and the time spent playing (Bekoff *et al.*, 1980).

As we already stated above, beside its increase in size, the cortex represents a very significant improvement as compared to the DP, in particular its capacity to be able to perform selection in autonomy from subcortical feedbacks thanks to its architecture that associates positive feedback with strong lateral inhibition (Helie *et al.*, 2015; Piron *et al.*, 2016). The complex organization of the cortex may also be linked to the length of the juvenile period: it takes time to be trained in order to develop its own automatisms and therefore the period of infancy of the young mammals may be partly devoted to its training (in which game plays a significant role). A very significant fraction of the ethogram is thus controlled by the cortex, which can, under some circumstances, become autonomous from the subcortical inputs. During learning, the BG would play an important guiding role in early phase of training (Piron *et al.*, 2016) while the strengthening of the cortico-cortical connections would allow the automatization and/or gradual development of habits, in which the cortex can choose by itself (albeit slower) if we disconnect the feedback after learning (Desmurget and Turner, 2008; Desmurget and Turner, 2010; Piron *et al.*, 2016). This empowerment by the cortex thus captures quite well the dynamics of the goal-oriented decision making as well as automatic skill execution. Convergent experimental findings and theoretical investigations support this idea that the "engram" of automatized skills shifts to regions outside the basal ganglia after sufficient training, including the neocortex (Graybiel, 1998; Houk and Wise, 1995; O'Reilly and Frank, 2006; Pasupathy and Miller, 2005). It can also explain the transfer of stimulus response association from the dorsomedial to the dorsolateral part of the CBG loop: once the connectivity between the cues and motor response is strong enough, inactivation of the ventral part, do not modify the response of the model. It could be noticed that as long as a sub-cortical feedback occurred, habits are reversible but it can take a very long time. However, if for one reason or another, sub-cortical feedback loops are disrupted, automatic

skills cannot be reversed and this may account for some motor or cognitive disorders, such as tic or addictive behaviors (Helie *et al.*, 2015).

3.4. Birds

Birds offer an interesting comparison of parallel evolution with mammals. They appeared 150 million years ago (later than the first mammals) on a branch that also carried the late dinosaurs (R.I.P.). Although descending all from a common ancestor, the 10,000+ extant species of birds present as much behavioral variety as mammals. It concerns courtship rituals that can take very different forms such as singing, complex displays or dances but also forms of communication concerning the establishment of a territory, defense against predators or the foraging. Many species of birds are also organized into social groups (Crook, 1965). Cognitive abilities of birds are known to be well developed. It is not a coincidence that the 2-armed bandit task, the gold standard of decision making task, was developed in pigeons (Herrnstein *et al.*, 1989). Caledonian crows are able to develop the use of tools (Emery, 2006) and are also able to solve abstract problems such as the Aesop fable task (Logan *et al.*, 2014) with as good performance as seven-year-old children. Songbirds provide also a significant example of the extent of their learning abilities (Mooney, 1995; Nottebohm *et al.*, 1990; Tchernichovski and Marcus, 2014). It is to be noticed that DP, thalamus and a part of the BG (called area X) are the main components of the neural network devoted to this ability. Playing behaviors have also been described in corvids and psittacines (Diamond and Bond, 2003). And finally, the magpie is the only non-mammal up to now that has passed successfully the mirror task for self-recognition (Prior *et al.*, 2008).

The enrichment of the ethogram of birds is correlated with the morphological and functional modifications of their encephalon. First of all, most of the birds possess a higher encephalization quotient than anamniotes and reptiles. But it is the development of the DP which is the more remarkable in these species. It represents more than 30% of the brain mass and can exceed 50% in the corvids (57% in the Caledonian crow) and some songbirds (Cnotka *et al.*, 2008; Herculano-Houzel, 2011b; Willemet, 2013). In some species, the DP organized itself in a hierarchy of specialized nuclei, with each nucleus controlling a specific behavior or interacting with others to control more complex behaviors (Pfenning *et al.*, 2014; Puelles *et al.*, 2000; Reiner *et al.*, 2004).

For social behaviors such as singing, it has been shown that in adult birds, once the song is learned, the animals can sing even if the feedback from the BG region called area X is blocked (Bottjer *et al.*, 1989; Scharff and Nottebohm, 1991). Furthermore, motor corrections implemented in the song to

correct vocal errors are initially driven through the CBG loop and later incorporated into a cortical pathway (Andalman and Fee, 2009; Warren *et al.*, 2011). On the other hand, decorticated pigeons can still learn RL tasks (Cerutti and Ferrari, 1995). It implies that in birds, BG are necessary and sufficient for RL-based learning (Fee and Goldberg, 2011), but DP nuclei can work autonomously, just as mammal cortex, for stereotypic behaviors. These observations further suggest that CBG circuits provide behavioral adaptations that are subsequently stabilized by training premotor cortical areas.
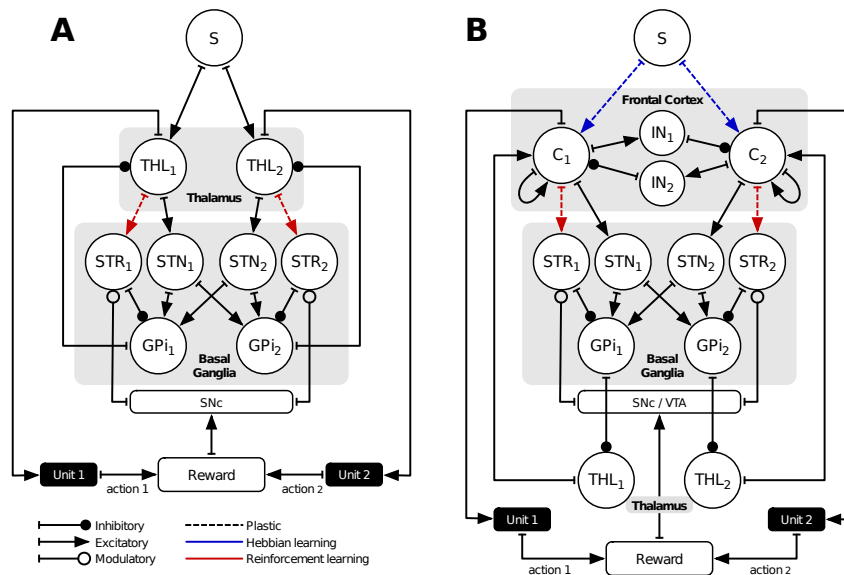
Box 1

Figure 3: A. A putative architecture for the selection of action through a primitive model of the basal ganglia and the thalamus (e.g. Lamprey). The stimulus (S) presentation results in the initial co-activation of two units in the thalamus (THL_1 and THL_2) until one of them is inhibited through the recurrent circuit (e.g. $THL_1 \rightarrow STN_1 \rightarrow GPI_2 \multimap THL_2$) and allows for the selection of the other. Reinforcement learning occurs between Thalamus and Striatum. **B.** A more elaborated model with lateral competition mechanism inside the frontal cortex (e.g. primate). The stimulus (S) presentation results in the initial co-activation of the two units in the frontal cortex (C_1 and C_2) until one of them is inhibited through either the recurrent circuit (e.g. $C_1 \rightarrow STN_1 \rightarrow GPI_2 \multimap THL_2 \rightarrow C_2$) and/or through the local competition conveyed by interneurons (e.g. $C_1 \rightarrow IN_1 \multimap C_2$) and allows for the selection of the other. Reinforcement learning occurs between the thalamus and the striatum while Hebbian learning occurs between the parietal cortex (S) and the frontal cortex.

The interactions between neuronal assemblies in the cortex/DP, BG and thalamus are believed to allow the automatization of decision making and skills, as we describe in this review. A few essential mechanisms are postulated that allow the selection of motor programs based on their outcome and the automatization of this selection process (illustrated in the figure 3). These mechanisms are necessary for the algorithmic implementation of the selection and automatization process, and they rely on the anatomical structure and physiological properties of the network that have been revealed by decades of experimental investigation. These mechanisms are the following:

- **A competition process**: The selection of motor programs, long thought to be at the heart of the function of the BG-thalamo-cortical network (Marsden, 1982; Mink and Thach, 1996). The selection of a program, or selective activation of a given cortical population in the presence of ambiguous inputs, requires a competition mechanism by which one population can inhibit others when selected to avoid multiple activations. Such a competition mechanism can be implemented through lateral inhibition (REF cortex; Wickens et al., 2007). In the present review, we postulate that competition mechanisms exist both at the DP/cortical level, through mutual or lateral inhibition mediated by inhibitory interneurons, and in the BG-thalamo-cortical loop circuits, through the divergent STN-GPi connection that mediates an inhibitory interaction between various cortical populations (via the Cortex-STN-GPi-Thalamus-Cortex pathway, see Leblois et al., 2006). The latter mechanism could in principle mediate competition between remote cortical populations (the combination of the STN-GPi divergent connection and the feedback from GPi to frontal cortex allows inhibitory interaction between distant DP/cortical populations through the Ctx1-STN-GPi-Thalamus-Ctx2 pathway, see figure above). This competition mechanism across the BG-thalamo-cortical network is common to all vertebrates, from lamprey to primates. The underlying dynamical process rests on a symmetry breaking process between parallel CBG circuits and has been described in Leblois et al. (2006). On the contrary, the former intra-cortical competition mechanism through lateral inhibition in the DP or cortex can only exist in animals which DP/cortex contains the necessary ingredients for lateral inhibition: a powerful inhibitory population in the DP/cortex. Such mechanism (Piron et al., 2016; Topalidou, submitted) could be implemented in mammals and birds, which cortex and pallium contain sufficiently strong lateral inhibition (REF). Note that it only competition allows between neighboring or 'not-too-far' cortical populations (due to the limited range of lateral inhibition),

- **An outcome-dependent learning process**: During the initial phase of skill learning, the agent (animal, human or artificial agent for modelling purposes) must learn to select the appropriate actions to optimize their outcome. This can be achieved only if the plasticity rules applied in

the neuronal assembly driving the selection of the motor program depend on the outcome of the action. In the BG, it is long known that DA signaling denotes the rewarding value of actions performed (more exactly the reward prediction error, REF operant conditioning?), and any plasticity rule that is modulated by the dopaminergic signal will therefore qualify as an outcome-dependent learning process (see figure above).

- **A Hebbian-like learning process**: As the decision-making process or the skill is repeated by the agent multiple times, the automatization of the skill relies, in our framework, on the ability of the DP/cortex to imprint the link between a given sensory context and the optimal action to select. This can be performed with a simple Hebbian-like learning process that ensures that neuronal populations which are regularly coactivated reinforce their connections (see figure above), making their co-activation independent of any inputs from the rest of the brain.

4. The automatization of skills

4.1. Distinct but complementary roles

The observations made in the previous section provide us with several clues regarding the respective roles of the BG and DP in decision making and more generally the execution and automatization of skills:

i) There is no evidence that the telencephalon plays a decisive role in the decision-making process itself. Indeed, organisms with smaller telencephala (such as toads or pigeons) are quite capable of making goal-oriented decisions (Herrnstein *et al.*, 1989). They are also capable of learning, and demonstrate the same tradeoff between exploration and exploitation in the two-armed bandit task as the most advanced mammals (Bradshaw *et al.*, 1979; Dougan *et al.*, 1985; Gilbert-Norton *et al.*, 2009; Graft *et al.*, 1977; Herrnstein, 1974; Lau and Glimcher, 2008; Matthews and Temple, 1979; Morris *et al.*, 2006; Palminteri *et al.*, 2009; Pasquereau *et al.*, 2007).

ii) It strongly suggests a correlation between the development of the telencephalon and the complexification of the ethogram. A larger DP/cortex brings a greater diversity in the scope of the decision, not necessarily an improvement of the decision-making process itself. This diversity can be expressed in several ways: by increasing the number of options available, by sequencing the possible answers, providing a decision tree that multiplies these options even more, or by providing new fields of application of the decision (Anderson, 1982; Boraud, 2015).

iii) The parallel emergence in mammals and birds of new structures of decision that are less sensitive to outcome (*i.e.* cortical columns for the former and pallial nuclei for the latter), and therefore allowed automatic procedures, conferred a significant evolutionary advantage.

4.2. A counter-intuitive hypothesis

Facing the vast amount of anatomical and electrophysiological data collected in various species of mammals and birds, we propose to rethink the evolution of skill learning and decision making as follows. In earlier vertebrates, the BG-thalamus loop and BG-brainstem loops drove most of the animal's behaviors. Dopamine dependent striatal plasticity allowed the development of RL process. The system is fairly adaptable but cannot learn "fast" habitual reactions that shortcut the long and relatively slow BG loops. As the output neurons from the DP (pyramidal cells in mammals) increase their direct contact to the motor control systems in the brainstem or the spinal cord, this structure is able to take control progressively of more and more behaviors. In parallel, changes in the micro-organization in mammals and birds, namely local inhibitory networks, allowed for the emergence of competition processes in the DP/cortical structures (see Box 1), which, associated with HL, allowed for the development of automatic skills.

Therefore, in the mammals and birds when it is necessary to learn new skills, RL processes in the BG train the DP/cortex that strengthen connections between sensory/associative and efferent neurons with HL processes in order to develop new skills that can be expressed autonomously later as proposed by Hellie, Ell and Ashby (2015). Many experimental data, thoroughly reviewed by these authors, support their theory in humans as well. It is tantalizing to link playing behavior to this capacity: playing allows for cortical skill learning that can be used thereafter to fight, flight, forage, hunt or interact socially.

4.3. Compatibility with the alternative hypotheses

The main alternative to our hypothesis stands that the development of habits is underlain by a progressive transfer from the associative loop to the sensorimotor one (Belin *et al.*, 2009; Yin and Knowlton, 2006). As an example, a huge set of data in rats, monkeys and humans involved the

associative striatum in flexible goal-directed behavior and the sensorimotor striatum in rigid habitual one (Le Masson *et al.*, 2002; Lehericy *et al.*, 2005; Miyachi *et al.*, 2002; Poldrack *et al.*, 2005).

It has also been shown in rats that lesions of the associative striatum can make the transfer of control from the associative to the sensorimotor striatum faster (Yin *et al.*, 2005). On the other hand devaluing the food reward in rats with lesions of the sensorimotor striatum, reduces lever-pressing rates even after extended training (Yin *et al.*, 2004). It suggests that habits cannot develop when this territory is altered.

However, many results are inconsistent with this hypothesis. In Parkinsonian patients, expression of learned skills is somehow conserved (Asmus *et al.*, 2008). In primates, disconnecting the output of the basal ganglia from motor cortex doesn't disrupt the expression of automatic behaviors (Desmurget and Turner, 2010; Piron *et al.*, 2016). Interestingly, these data have been demonstrated also in songbirds (Andalman and Fee, 2009) but not in rodents.

Can we incorporate these discrepancies in our evolutionary driven hypothesis? One possibility is that "the transfer of control from the associative striatum to the sensorimotor striatum is more accurately described as a switch from a ventral-based declarative memory system to a dorsal-based procedural memory one" (Helie *et al.*, 2015). According to this hypothesis, "what is tested as goal-directed instrumental response recruits in fact networks involved in declarative memory tasks, whereas stimulus response association that is under the control of habit learning (i.e., which immediately precedes habitual behavior) depends on much the same circuitry as procedural learning" (Helie *et al.*, 2015; Khamassi and Humphries, 2012). It also suggests that evolution has segregated the dorsomedial striatum and the dorsolateral striatum into separate learning systems. If so, the results showing that activation in the associative striatum precedes activation in the sensorimotor one would reflect a faster dynamic of learning for the dorsomedial striatum than the dorsolateral one, rather than a sequential process. This segregation may be a unique feature of mammals as it has not been demonstrated yet in other species.

Concerning the segregation between a subcortical system dedicated to "habitual" behavior and a frontal cortex dedicated to "deliberative" processes, it has been shifted recently under the influence of Machine Learning into a "Model-Free" vs "Model-Based" debate (Daw *et al.*, 2006; Khamassi and Humphries, 2012). The BG are supposed to learn new associations by trial and error without prior assumptions while the cortex is able to generate representation of the world that allows it to compare each state to a predefined model onto which it can project the outcomes of future(s) action(s). In fact, we can reconcile this theory with ours by taking into account the fact that OFC and the ventral PFC, for example, receive massive inputs from the hippocampus that is involved in the building of

representation of the world. These areas can merge thalamus mediated BG inputs and hippocampal information into neural representations of state dependent action(s).

4.4 The entanglement of two systems

To formalize our hypothesis, we proposed a theoretical model (Topalidou *et al.*, submitted) that is an extension of previously published models (Guthrie *et al.*, 2013; Leblois *et al.*, 2006a; Topalidou *et al.*, 2015). This model is centered around a simple two-armed bandit task that aims to illustrate the main concepts we've introduced. Even though the task is a decision-making task, we believe the proposed architecture and the associated learning mechanisms are generic enough to be generalized to any other task that involves the basal ganglia-thalamus-cortex loop.

4.5. Computational Approach

We provided previously an original model of action selection mechanism (Leblois *et al.*, 2006b), which emerges from "the competition between a positive feedback through the direct pathway and a negative feedback through the hyperdirect pathway in the cortico-basal-thalamic loop". In order to explore the parallel organization of circuits in the BG, we then upgraded the model (Guthrie *et al.*, 2013) such as it includes all the major nuclei of the basal ganglia but GPe and is segregated into three loops (motor, associative and cognitive) that encompass also the cortex and the thalamus (Alexander and Crutcher, 1990b; Alexander *et al.*, 1986). In this model, the cortex is modeled by a single layer of excitatory neurons. We realized that, although it reproduced fairly well the behavioral and electrophysiological signature of learning and choice in a two-armed bandit task, it could not account for the development of habits in primates (Desmurget and Turner, 2010; Piron *et al.*, 2016). In fact, its architecture is closer to the decision-making network of early vertebrates, with an archaic pallium compared to that of mammals. To account for the development of cortex with its specific dynamic properties, we therefore incorporate a lateral competition mechanism in all three cortices (motor, cognitive, associative) based on short range excitation and long-range inhibition. We also added connections between the associative, cognitive and motor areas in order to allow crosstalk among these structures (Douglas and Martin, 2004). This competition process granted our cortical network with the capacity to make a decision by its own, although more slowly than when the BG are active (Figure 4).

In the early phase of learning, the system relies on intrinsic noise to select one of the options.

The effect of phasic dopamine delivery on synaptic plasticity is modeled by a RL process at the level of cortico-striatal synapse in the cognitive territory of the network. Therefore, the decision made at the cognitive level can be used to bias the decision at the motor level. Once a motor action has been performed, HL modifies the connections between the cognitive cortex and the associative cortex. The latter process *does not depend* on reward but only on the choices. It is to be noted that the cortical selection, resulting from lateral competition in the cortex, needed more trials to be strong enough to bias a decision than the cortico-basal ganglia competition process such that the cortex is initially driven by the basal ganglia output (GPi), hence it learns from the statistics provided by the BG selection. Before the development of cortico-cortical connections, the level of noise induces some irregularity that match exploratory behaviors. It is also the behavior displayed by our earlier model devoid of a proper cortical layer (Guthrie *et al.*, 2013) . Once the cortical learning has developed automatisms, the system is much less sensitive to noise and reaches almost 100% performances. An interesting prediction is therefore that lower-vertebrates such as lamprey or anamniote should be able to learn task by RL but should not be able to develop automatic skills.
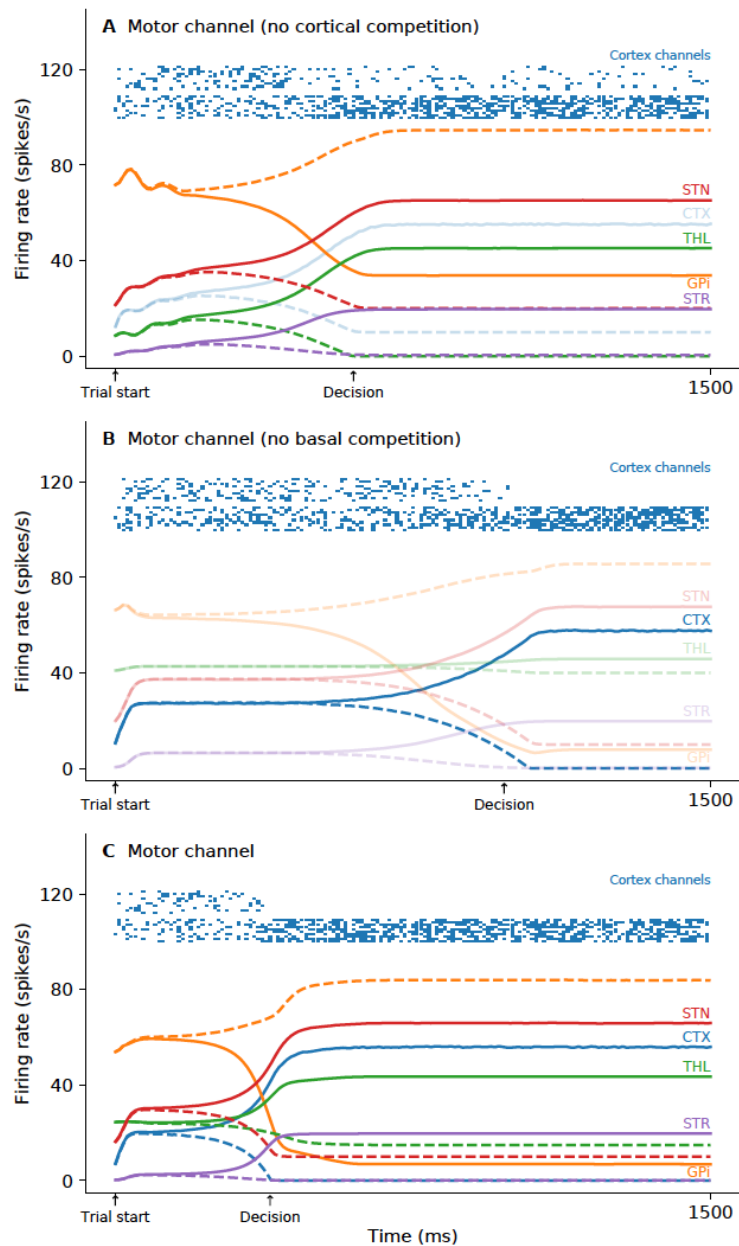
Figure 4. Competition dynamic in the model. basal ganglia and the cortex. Solid lines represent activity related to the selected population, dashed lines represent activity related to the non-selected population. Decision threshold has been set to 40 spikes/s between the two cortical populations and is indicated on the x axis. Raster plots are related to the cortical populations and has been generated from the firing rate of 10 neurons. A) Activity in the motor populations in the absence of lateral competition in the cortical populations. B) Activity in the motor populations in the absence of the feedback from the basal ganglia (GPi) to the cortical populations via the thalamus. Decision threshold is reached thanks to the direct lateral competition in both cognitive and motor cortical channels. C) Activity in the motor populations in the full model with a dual competition, one cortical, one basal. When congruent (cortical and basal decision are the same), decision time for both the motor and cortical channels are faster than in the absence of one of the competition loop (from Topalidou *et al.*, submitted).

6. Conclusion

We showed here that adopting a phylogenetic approach to the decision-making network helps to explain how it works in the species with the most complex brains, such as primates. In this taxonomic order, our conclusions are very close to the theory of the BG acting as "a general training machine for cortico-cortical connections" defended by Helie *et al.* (2015). However, we show that it accounts for a late phase of the evolution that appeared in parallel in mammals and birds and maybe some fishes that are able to develop automatic skills. In other species, subcortical loops drive most of the behavior and therefore even if they are capable of RL driven flexibility, they lack the capacity to develop strong automatic skills in order to optimize their performance.

The phylogenetic approach to understanding decision making is still in its infancy. We should gather information more comprehensively in lower vertebrates (fishes, amphibians, reptiles), mammals and birds in order to test our theory. For example, it has been demonstrated that salamanders, an amphibian with a brain architecture close to that of the lamprey, can be trained to perform a navigation task (Taylor and Adler, 1973). We just started an experimental study to test whether it relies on a classical DA dependent RL process and if we can operationalize habitual behavior in these species.

Incidentally, our theory provides a possible solution to an old conundrum concerning BG human pathology. In the classical view of the BG as a selection device (Mink, 1996), the effect of the inactivation of the STN or the GPi as a therapeutic approach to parkinsonism was paradoxical. But if we consider that such surgical treatment is practiced in a Parkinsonian patient old enough to have an over-trained cortex in the skills they use on a day-to-day basis, one can understand that it doesn't need the subcortical feedback anymore. And indeed, some experimental and clinical data have been gathered showing that in those patients learning of new skills is disrupted (Jahanshahi *et al.*, 2014; Wilkinson *et al.*, 2009).

We may also help to modify theories on addictions. Those disorders are often considered as "bad cognitive habits". They were considered hitherto as subcortical disorders. But according to our theory, they should rely on abnormal cortical processes. Recent published work seems to support this hypothesis. Guillem and Ahmed (2017) showed an abnormal ratio of neurons encoding for an addictive

substance in the OFC, but it will need more experimental evidence. It is nevertheless a stimulating thread to follow.

Finally, our theory provides an anatomical support to the think fast-think slow theory developed by Kahneman (2011). He proposed that human decision making is the result of a competition between a fast, automatic system that is prone to make mistakes and a slower, more demanding system that is more reliable. Kahneman himself never identified a neuroanatomical substrate of his theory, but many others have tried to put it in parallel with a triune brain organization: the fast system being associated to the "reptilian brain", while the slow one is identified as a product of the "neo-mammalian cortex" (Crosby, 2015). In fact, our proposition takes the opposite line. According to us, the slow system is the older RL dependent cortico-subcortical loop, while the fast one results from cortical Hebbian associations. It implies that shortcuts (heuristics) rely on cortical processes while difficult decisions and actions relying on emotional and moral motivation rather stand on a subcortical substrate. It brings a fresh view to the psychology of decision making and may help to unravel the neural correlates of cognitive bias.

Acknowledgement

Bibliography

Adesnik, H., Scanziani, M., 2010. Lateral competition for cortical space by layer-specific horizontal circuits. *Nature* 464, 1155-1160.

Albin, R.L., Young, A.B., Penney, J.B., 1989. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* 12, 366-375.

Alexander, G.A., Crutcher, M.D., 1990a. Preparation for movement: neural representations of intended direction in three motor areas of the monkey. *J. Neurophysiol.* 64, 133-150.

Alexander, G.E., Crutcher, M.D., 1990b. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13, 266-271.

Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357-381.

Andalman, A.S., Fee, M.S., 2009. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences of the United States of America* 106, 12518-12523.

Anderson, J.R., 1982. Acquisiton of cognitive skill. *Psychological review* 89, 369-403.

Apicella, P., Scarnati, E., Ljungberg, T., Schultz, W., 1992. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J. Neurophysiol.* 68, 945-960.

Asmus, F., Huber, H., Gasser, T., Schols, L., 2008. Kick and rush: paradoxical kinesia in Parkinson disease. *Neurology* 71, 695.

Balleine, B., Dickinson, A., 1992. Signalling and incentive processes in instrumental reinforcer devaluation. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology* 45, 285-301.

Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 35, 48-69.

Bartus, R.T., Levere, T.E., 1977. Frontal decortication in rhesus monkeys: a test of the interference hypothesis. *Brain research* 119, 233-248.

Bekoff, M., Byers, J.A., Bekoff, A., 1980. Prenatal motility and postnatal play: functional continuity? *Developmental psychobiology* 13, 225-228.

Belin, D., Jonkman, S., Dickinson, A., Robbins, T.W., Everitt, B.J., 2009. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behavioural brain research* 199, 89-102.

Bird, C.D., Emery, N.J., 2009a. Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Sciences of the United States of America* 106**,** 10370-10375.

Bird, C.D., Emery, N.J., 2009b. Rooks use stones to raise the water level to reach a floating worm. *Current biology : CB* 19**,** 1410-1414.

Bjursten, L.M., Norrsell, K., Norrsell, U., 1976. Behavioural repertory of cats without cerebral cortex from infancy. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale* 25**,** 115-130.

Boraud, T., 2015. *Matière à décision*. CNRS éditions: Paris.

Bottjer, S.W., Halsema, K.A., Brown, S.A., Miesner, E.A., 1989. Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *The Journal of comparative neurology* 279**,** 312-326.

Bradshaw, C.M., Szabadi, E., Bevan, P., Ruddle, H.V., 1979. The effect of signaled reinforcement availability on concurrent performances in humans. *J Exp Anal Behav* 32**,** 65-74.

Brumm, H., Zollinger, S.A., 2017. Vocal plasticity in a reptile. *Proceedings. Biological sciences / The Royal Society* 284.

Cerutti, S.M., Ferrari, E.A., 1995. Operant discrimination learning in detelencephalated pigeons (Columba livia). *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas* 28**,** 1089-1095.

Cnotka, J., Gunturkun, O., Rehkamper, G., Gray, R.D., Hunt, G.R., 2008. Extraordinary large brains in tool-using New Caledonian crows (Corvus moneduloides). *Neuroscience letters* 433**,** 241-245.

Crook, J.H., 1965. The adaptive significance of avian social organisations. *Symp. zool. Soc. Lond.* 14**,** 181-218.

Crosby, G., 2015. *Fight, Flight, Freeze : Taming Your Reptilian Brain and Other Practical Approaches to Self-Improvement*. Crosbyod Publishing.

Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience* 8**,** 1704-1711.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441**,** 876-879.

Desmurget, M., Turner, R.S., 2008. Testing basal ganglia motor functions through reversible inactivations in the posterior internal globus pallidus. *Journal of neurophysiology* 99**,** 1057-1076.

Desmurget, M., Turner, R.S., 2010. Motor Sequences and the Basal Ganglia: Kinematics, Not Habits. *J. Neurosci.* 30**,** 7685-7690.

Diamond, J., Bond, A.B., 2003. A Comparative Analysis of Social Play in Birds. *Behaviour* 140, 1091-1115.

Dickinson, A., 1985. Actions and Habits: The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 308, 67.

Dougan, J.D., McSweeney, F.K., Farmer, V.A., 1985. Some parameters of behavioral contrast and allocation of interim behavior in rats. *J Exp Anal Behav* 44, 325-335.

Douglas, R.J., Martin, K.A., 2004. Neuronal circuits of the neocortex. *Annu Rev Neurosci* 27, 419-451.

Emery, N.J., 2006. Cognitive ornithology: the evolution of avian intelligence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 361, 23-43.

Fee, M.S., Goldberg, J.H., 2011. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* 198, 152-170.

Finkenstadt, T. 1989. Visual associative learning: Searching for behaviorally relevant brain structures in toads. In: *Visuomotor coordination.* pp. 799-832. Eds. J.P. Ewert, M.A. Arbib. Plenum Press: New York.

Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898-1902.

Fitts, P. 1964. Perceptual-motor skill learning. In: *Categories of human learning.* pp. 243-285. Ed. A. Melton. Academic Press: New York.

Gilbert-Norton, L.B., Shahan, T.A., Shivik, J.A., 2009. Coyotes (Canis latrans) and the matching law. *Behavioural processes* 82, 178-183.

Graft, D.A., Lea, S.E., Whitworth, T.L., 1977. The matching law in and within groups of rats. *J Exp Anal Behav* 27, 183-194.

Graybiel, A.M., 1998. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70, 119-136.

Graybiel, A.M., 2008. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31, 359-387.

Grillner, S., Robertson, B., Stephenson-Jones, M., 2013. The evolutionary origin of the vertebrate basal ganglia and its role in action selection. *The Journal of physiology* 591, 5425-5431.

Guillem, K., Ahmed, S.H., 2017. Preference for Cocaine is Represented in the Orbitofrontal Cortex by an Increased Proportion of Cocaine Use-Coding Neurons. *Cereb Cortex.*

Gunturkun, O., 2012. The convergent evolution of neural substrates for cognition. *Psychol Res* 76, 212-219.

Guthrie, M., Leblois, A., Garenne, A., Boraud, T., 2013. Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study. *Journal of neurophysiology* 109, 3025-3040.

Helie, S., Ell, S.W., Ashby, F.G., 2015. Learning robust cortico-cortical associations with the basal ganglia: an integrative review. *Cortex; a journal devoted to the study of the nervous system and behavior* 64, 123-135.

Herculano-Houzel, S., 2011a. Brains matter, bodies maybe not: the case for examining neuron numbers irrespective of body size. *Annals of the New York Academy of Sciences* 1225, 191-199.

Herculano-Houzel, S., 2011b. Not all brains are made the same: new views on brain scaling in evolution. *Brain, behavior and evolution* 78, 22-36.

Herculano-Houzel, S., Mota, B., Lent, R., 2006. Cellular scaling rules for rodent brains. *Proceedings of the National Academy of Sciences of the United States of America* 103, 12138-12143.

Herrnstein, R.J., 1974. Formal properties of the matching law. *J Exp Anal Behav* 21, 159-164.

Herrnstein, R.J., Vaughan, W., Jr., Mumford, D.B., Kosslyn, S.M., 1989. Teaching pigeons an abstract relational rule: insideness. *Perception & psychophysics* 46, 56-64.

Hodos, W., Butler, A.B., 1997. Evolution of sensory pathways in vertebrates. *Brain, behavior and evolution* 50, 189-197.

Horton, J.C., Adams, D.L., 2005. The cortical column: a structure without a function. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 837.

Houk, J.C., Wise, S.P., 1995. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb Cortex* 5, 95-110.

Jahanshahi, M., Obeso, I., Baunez, C., Alegre, M., Krack, P., 2014. Parkinson's Disease, the Subthalamic Nucleus, Inhibition, and Impulsivity. *Movement disorders : official journal of the Movement Disorder Society*.

Jaldow, E.J., Oakley, D.A., Davey, G.C., 1989. Performance of Decorticated Rats on Fixed Interval and Fixed Time Schedules. *The European journal of neuroscience* 1, 461-470.

Kahneman, D., 2011. *Thinking, fast and slow*. Farrar, Straus and Giroux: New York.

Kawase, H., Okata, Y., Ito, K., 2013. Role of huge geometric circular structures in the reproduction of a marine pufferfish. *Sci Rep* 3, 2106.

Keysers, C., Perrett, D.I., 2004. Demystifying social cognition: a Hebbian perspective. *Trends in cognitive sciences* 8, 501-507.

Khamassi, M., Humphries, M.D., 2012. Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Frontiers in behavioral neuroscience* 6**,** 79.

Kukleta, M., Libouban, S., 1967. [Acquisition of instrumental conditioning in functionally decorticated rats]. *Journal de physiologie* 59**,** 442.

Lau, B., Glimcher, P.W., 2008. Value representations in the primate striatum during matching behavior. *Neuron* 58**,** 451-463.

Le Masson, G., Renaud-Le Masson, S., Debay, D., Bal, T., 2002. Feedback inhibition controls spike transfer in hybrid thalamic circuits. *Nature* 417**,** 854-858.

Leblois, A., Boraud, T., Meissner, W., Bergman, H., Hansel, D., 2006a. Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26**,** 3567-3583.

Leblois, A., Meissner, W., Bezard, E., Bioulac, B., Gross, C.E., Boraud, T., 2006b. Temporal and spatial alterations in GPi neuronal encoding might contribute to slow down movement in parkinsonian monkeys. *The European journal of neuroscience* 24**,** 1201-1208.

Lehericy, S., Benali, H., Van de Moortele, P.F., Pelegrini-Issac, M., Waechter, T., Ugurbil, K., Doyon, J., 2005. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proceedings of the National Academy of Sciences of the United States of America* 102**,** 12566-12571.

Lemon, R.N., Griffiths, J., 2005. Comparing the function of the corticospinal system in different species: organizational differences for motor specialization? *Muscle Nerve* 32**,** 261-279.

Logan, C.J., Jelbert, S.A., Breen, A.J., Gray, R.D., Taylor, A.H., 2014. Modifications to the Aesop's Fable Paradigm Change New Caledonian Crow Performances. *PloS one* 9**,** e103049.

MacLean, P.D., 1973. *An evolutionary approach to the investigation of psychoneuro-endocrine functions*. Plenum Press: New York,.

Mallet, N., Micklem, B.R., Henny, P., Brown, M.T., Williams, C., Bolam, J.P., Nakamura, K.C., Magill, P.J., 2012. Dichotomous organization of the external globus pallidus. *Neuron* 74**,** 1075-1086.

Mallet, N., Schmidt, R., Leventhal, D., Chen, F., Amer, N., Boraud, T., Berke, Joshua D., Arkypallidal Cells Send a Stop Signal to Striatum. *Neuron* 89**,** 308-316.

Matthews, L.R., Temple, W., 1979. Concurrent schedule assessment of food preference in cows. *J Exp Anal Behav* 32**,** 245-254.

Mink, J.W., 1996. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50**,** 381-425.

Miyachi, S., Hikosaka, O., Lu, X., 2002. Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale* 146, 122-126.

Mooney, R., 1995. Behavioral learning. The illuminated songbird. *Current biology : CB* 5, 609-611.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., Bergman, H., 2006. Midbrain dopamine neurons encode decisions for future action. *Nature neuroscience* 9, 1057-1063.

Nambu, A., Tokuno, H., Hamada, I., Kita, H., Imanishi, M., Akazawa, T., Ikeuchi, Y., Hasegawa, N., 2000. Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *Journal of neurophysiology* 84, 289-300.

Nambu, A., Tokuno, H., Takada, M., 2002. Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. *Neurosci Res* 43, 111-117.

Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Kirn, J., Ling, C.Y., Nottebohm, M., Suter, R., Tolles, A., Williams, H., 1990. Song learning in birds: the relation between perception and production. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 329, 115-124.

O'Connell, L.A., Hofmann, H.A., 2011. The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. *The Journal of comparative neurology* 519, 3599-3639.

O'Reilly, R.C., Frank, M.J., 2006. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural computation* 18, 283-328.

Oakley, D.A., 1981. Performance of decorticated rats in a two-choice visual discrimination apparatus. *Behavioural brain research* 3, 55-69.

Ocana, F.M., Suryanarayana, S.M., Saitoh, K., Kardamakis, A.A., Capantini, L., Robertson, B., Grillner, S., 2015. The lamprey pallium provides a blueprint of the mammalian motor projections from cortex. *Current biology : CB* 25, 413-423.

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., Pessiglione, M., 2009. Brain hemispheres selectively track the expected value of contralateral options. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, 13465-13472.

Parent, A., Hazrati, L.-N., 1995a. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res. Rev.* 20, 91-127.

Parent, A., Hazrati, L.-N., 1995b. Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Res. Rev.* 20, 128-154.

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., Gross, C.E., Boraud, T., 2007. Shaping of motor responses by incentive values through the basal ganglia. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 27, 1176-1183.

Pasupathy, A., Miller, E.K., 2005. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873-876.

Patton, B.W., Braithwaite, V.A., 2015. Changing tides: ecological and historical perspectives on fish cognition. *Wiley Interdiscip Rev Cogn Sci* 6, 159-176.

Patton, P., Grobstein, P., 1998. The effects of telencephalic lesions on visually mediated prey orienting behavior in the leopard frog (Rana pipiens). I. The effects of complete removal of one telencephalic lobe, with a comparison to the effects of unilateral tectal lobe lesions. *Brain, behavior and evolution* 51, 123-143.

Pfenning, A.R., Hara, E., Whitney, O., Rivas, M.V., Wang, R., Roulhac, P.L., Howard, J.T., Wirthlin, M., Lovell, P.V., Ganapathy, G., Mouncastle, J., Moseley, M.A., Thompson, J.W., Soderblom, E.J., Iriki, A., Kato, M., Gilbert, M.T., Zhang, G., Bakken, T., Bongaarts, A., Bernard, A., Lein, E., Mello, C.V., Hartemink, A.J., Jarvis, E.D., 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346, 1256846.

Piron, C., Kase, D., Topalidou, M., Goillandeau, M., Orignac, H., N'Guyen, T.H., Rougier, N., Boraud, T., 2016. The globus pallidus pars interna in goal-oriented and routine behaviors: Resolving a long-standing paradox. *Movement disorders : official journal of the Movement Disorder Society* 31, 1146-1154.

Poldrack, R.A., Sabb, F.W., Foerde, K., Tom, S.M., Asarnow, R.F., Bookheimer, S.Y., Knowlton, B.J., 2005. The neural correlates of motor skill automaticity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 25, 5356-5364.

Pombal, M.A., Megias, M., Bardet, S.M., Puelles, L., 2009. New and old thoughts on the segmental organization of the forebrain in lampreys. *Brain, behavior and evolution* 74, 7-19.

Pombal, M.A., Puelles, L., 1999. Prosomeric map of the lamprey forebrain based on calretinin immunocytochemistry, Nissl stain, and ancillary markers. *The Journal of comparative neurology* 414, 391-422.

Prior, H., Schwarz, A., Gunturkun, O., 2008. Mirror-induced behavior in the magpie (Pica pica): evidence of self-recognition. *PLoS biology* 6, e202.

Puelles, L., Kuwana, E., Puelles, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., Rubenstein, J.L., 2000. Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. *The Journal of comparative neurology* 424, 409-438.

Reiner, A., Perkel, D.J., Mello, C.V., Jarvis, E.D., 2004. Songbirds and the revised avian brain nomenclature. *Annals of the New York Academy of Sciences* 1016, 77-108.

Samejima, K., Doya, K., 2007. Multiple representations of belief states and action values in corticobasal ganglia loops. *Annals of the New York Academy of Sciences* 1104, 213-228.

Scharff, C., Nottebohm, F., 1991. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 11, 2896-2913.

Schultz, W., 1992. Activity of dopamine neurons in the behaving primate. *Seminars in the neurosciences* 4, 129-138.

Schultz, W., 2006. Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57, 87-115.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593-1599.

Seger, C.A., Spiering, B.J., 2011. A critical review of habit learning and the Basal Ganglia. *Frontiers in systems neuroscience* 5, 66.

Silberberg, G., Bolam, J.P., 2015. Local and afferent synaptic pathways in the striatal microcircuitry. *Current opinion in neurobiology* 33, 182-187.

Song, S., Miller, K.D., Abbott, L.F., 2000. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience* 3, 919-926.

Stephenson-Jones, M., Ericsson, J., Robertson, B., Grillner, S., 2012. Evolution of the basal ganglia: dual-output pathways conserved throughout vertebrate phylogeny. *The Journal of comparative neurology* 520, 2957-2973.

Striedter, G.F., 2005. *Principles of brain evolution*. Sinauer Associates: Sunderland, Mass.

Taylor, D.H., Adler, K., 1973. Spatial orientation by Salamanders using plane-polarized light. *Science* 181, 285-287.

Tchernichovski, O., Marcus, G., 2014. Vocal learning beyond imitation: mechanisms of adaptive vocal development in songbirds and human infants. *Current opinion in neurobiology* 28c, 42-47.

Topalidou, M., Kase, D., Boraud, T., Rougier, N., submitted. Dual Competition between the Basal Ganglia and the Cortex: from Action-Outcome to Stimulus-Response.

Topalidou, M., Leblois, A., Boraud, T., Rougier, N.P., 2015. A long journey into reproducible computational neuroscience. *Frontiers in computational neuroscience* 9, 30.

Travis, A.M., Woolsey, C.N., 1956. MOTOR PERFORMANCE OP MONKEYS AFTER BILATERAL PARTIAL AND TOTAL CEREBRAL DECORTICATIONS. *American Journal of Physical Medicine & Rehabilitation* 35.

Turner, R.S., Desmurget, M., 2010. Basal ganglia contributions to motor control: a vigorous tutor. *Current opinion in neurobiology* 20, 704-716.

Vernier, P. 2017. The Brain of Teleos Fishes. In: *Evolution of the Nervous System*. pp. 59-76. Ed. J.H. Kaas. Academic Press: London.

Warren, T.L., Tumer, E.C., Charlesworth, J.D., Brainard, M.S., 2011. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *Journal of neurophysiology* 106**,** 1806-1821.

Whishaw, I.Q., Schallert, T., Kolb, B., 1981. An analysis of feeding and sensorimotor abilities of rats after decortication. *J Comp Physiol Psychol* 95**,** 85-103.

Wilkinson, L., Khan, Z., Jahanshahi, M., 2009. The role of the basal ganglia and its cortical connections in sequence learning: evidence from implicit and explicit sequence learning in Parkinson's disease. *Neuropsychologia* 47**,** 2564-2573.

Willemet, R., 2013. Reconsidering the evolution of brain, cognition, and behavior in birds and mammals. *Frontiers in psychology* 4**,** 396.

Wullimann, M.F., Vernier, P. 2007. Evolution of the Nervous System in Fishes In: *Evolution of Nervous Systems*. pp. 39-60. Ed. J.H. Kaas. Academic Press: Oxford.

Yin, H.H., Knowlton, B.J., 2006. The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7**,** 464-476.

Yin, H.H., Knowlton, B.J., Balleine, B.W., 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *The European journal of neuroscience* 19**,** 181-189.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., Balleine, B.W., 2005. The role of the dorsomedial striatum in instrumental conditioning. *The European journal of neuroscience* 22**,** 513-523.

Abbreviations

ACC: anterior cingulate cortex
BG: basal ganglia
CBG: cortex-basal ganglia (loop)
DA: dopamine
DP: dorsal pallium
GABA: gamma-aminobutyric acid
GPe: external segment of the globus pallidus
GPi: internal segment of the globus pallidus
HL: Hebbian learning
MC: motor cortex
MLR: mesencephalic locomotor region
MSN: medium spiny neurons
OFC: orbitofrontal cortex
PFC: prefrontal cortex
PMC: premotor cortex
RL: reinforcement learning
SNr: substancia nigra pars reticulata
STN: subthalamic nucleus
VTA: ventral tegmental area