

TEL AVIV UNIVERSITY
DEPARTMENT OF PSYCHOLOGY

REINFORCEMENT LEARNING
AND THE BASAL GANGLIA

SEMINAR PROJECT SUBMITTED TO
PROF. INA WIENER
AS PART OF AN
INSTRUCTED READING SEMINAR
SUPERVISED BY DR. DAPHNA JOEL

by

Yael Niv

December 2001

Contents

1	Introduction	3
2	Anatomy of the Basal Ganglia	4
3	Functional Anatomy and Physiology of the Striatum	6
3.1	Physiology of the Striatum	6
3.2	Three Types of Striatal Aspyiny Interneurons	9
3.3	Spiny Projection Cells are also Heterogenous	9
3.4	Segregation of the Striatal Afferents	11
4	What are the Basal Ganglia Doing?	
	From Structure to Function	13
5	What Happens When Things Go Wrong?	
	Basal Ganglia Disorders	17
6	Learning in the Striatum:	
	The Double Role of Dopamine	21
6.1	The Role of Dopamine in Motor Activation	21
6.2	The Role of Dopamine in Reward-Mediated Learning	23
6.2.1	The Dopaminergic Reinforcement Signal	24
6.2.2	The Connectivity of the Dopaminergic System: Afferent Connections	25
6.2.3	The Connectivity of the Dopaminergic System: Efferent Connections	27
6.2.4	Influence of Striatal Activation on Dopaminergic Neurons	28
6.2.5	Reward-Mediated Learning: A Cellular Mechanism	29
6.3	Two Roles Subservd by Two Dopamine Receptors?	32
7	Computational Models of the Basal Ganglia	33

7.1	Models of the Striatum	34
7.1.1	Is the Striatum a Competitive Network?	34
7.1.2	Parallels Between the Striatum and the Cerebellum	34
7.1.3	Cortical Assemblies and Learning in the Striatum	36
7.2	Actor-Critic Models of the Basal Ganglia	38
7.2.1	The Computational Actor-Critic Model	38
7.2.2	Examining the Basal Ganglia through the Actor-Critic Framework	40
8	Summary: Reinforcement Learning and the Basal Ganglia	42
	References	44

1 Introduction

The basal ganglia are an extensively studied group of subcortical neuronal structures. Much research has been devoted to elucidating the nature of the neural processing in these structures, and the role of the basal ganglia in normal behavior. The basal ganglia probably hold an important role in contextual analysis of the environment and adaptive use of the acquired information in order to plan and execute intelligent behaviors (Houk, 1995). The striatum (the input stage of the basal ganglia) has been recognized as a critical structure in learning of stimulus-response habits as well as motor, perceptual and cognitive skills (Joel & Wiener, 1994; Suzuki, Miura, Mishimura, & Aosaki, 2001). The striatum is also a major contributor of basal ganglia input to the dopaminergic system, and the major recipient of dopaminergic afferents (Joel & Wiener, 2000).

Reinforcement learning (RL) is a paradigm which describes learning in an agent embedded in its environment. The agent does not learn via a supervisor or a supervisory stimulus, but through its interactions with its environment. The capacity of the environment to provide rewards to the organism as a result of its actions, is the basis for this type of learning (Sutton & Barto, 1998). The core ideas of modern reinforcement learning come from classical and instrumental conditioning theories in psychology (although, as pointed out by Barto (1995), psychologists do not use the term reinforcement learning). Behavioral research shows that reinforcement learning is a fundamental process by which both vertebrates and invertebrates learn to achieve goals from their interactions with the environment, as most natural learning processes are conducted in the absence of an explicit supervisory stimulus.

Several brain regions have been implicated in reinforcement learning, including the midbrain dopaminergic neurons of the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA) in rats and primates, and their target areas in the basal ganglia (e.g. Graybiel & Kimura, 1995; Houk & Wise, 1995; Schultz, 1998). Reinforcement learning models of basal ganglia functioning regard the dopaminergic input to the striatum as a reward signal, which influences action-selection learning in the striatum.

In the following I will provide a short review of the basal ganglia and their relationship to reinforcement learning. The account is by no means comprehensive, and is also somewhat general and does not only deal with issues relevant to reinforcement learning, as it has stemmed from a general instructed-reading seminar on the basal ganglia. Thus one goal of this report is to summarize the general features of the basal ganglia organization scanned in the seminar, while another goal is to try to find possible connections between this data and computational models of reinforcement learning.

The general account will begin with the anatomy of the basal ganglia, with a more detailed inspection of the anatomy and physiology of the striatum. In order to complete the general picture of the basal ganglia, will then very briefly summarize some views regarding the function of the basal ganglia in normal behavior, and their influence on behavior as can be elucidated by the various symptoms of basal ganglia disorders. This account will be very brief and certainly does not encompass even a small percentage of the vast literature regarding these issues. I will start focusing on learning in the basal ganglia with a biologically-based account of the two roles of dopamine in motor activation and in reward mediated learning in the striatum. This discussion will lead to the more formal computational models of the basal ganglia, and finally to the analogy between the basal ganglia and the actor-critic reinforcement learning architecture. I will summarize with a few highlights as to my personal view of reinforcement learning and its connection to the basal ganglia.

2 Anatomy of the Basal Ganglia

The basal ganglia are comprised of the striatum (consisting of most of the caudate and the putamen, and of the nucleus accumbens), the internal (medial) and external (lateral) segments of the globus pallidus (GPe and GPi respectively), the subthalamic nucleus (STN), the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc) and pars reticulata (SNr). The basal ganglia as a whole provide a feedback loop to the cortex, which is the main source of afferents, as well as the target of most of the

basal ganglia efferents (via basal ganglionic influence on the thalamus). The general organization of the basal ganglia is that of a feed-forward network (Bergman et al., 1998). The input stage of the basal ganglia is the striatum, which is innervated by excitatory (glutamatergic) pyramidal neurons from all areas of the neocortex, via a massive converging corticostriatal projection. The striatum is a relatively homogenous structure composed mainly (90-95%) of medium sized spiny cells which are GABAergic projection neurons, while the rest are interneurons.

The striatum sends its efferents via two major pathways, a direct one and an indirect one (Alexander & Crutcher, 1990). Via the direct pathway, the striatum inhibits the output structures of the basal ganglia, the SNr and the GPi (the entopeduncular nucleus in rodents). Although these two structures are separated by the fibres of the cerebral pedunculus in most mammals, they contain cytologically similar neurons, and are thus commonly regarded as one functional structure (GPi/SNr). In the absence of striatal inhibition, the GABAergic GPi/SNr cells fire tonically and inhibit midbrain motor areas and thalamic nuclei (including the mediodorsal and ventral tier nuclei, the intralaminar thalamic nuclei, the superior colliculus and the pedunculopontine nucleus). Thus, through two inhibitory serial connections, the direct influence of the striatum on the thalamus is that of release from inhibition. The disinhibited thalamus, in turn, enhances cortical activity via an excitatory glutamatergic projection to frontal cortical areas.

The polysynaptic indirect pathway to the GPi/SNr is an opposing pathway of activity. This pathway originates in the striatum with GABAergic projections to the GPe. The GPe, in turn, inhibits the STN through GABAergic connections, and the latter excites the GPi/SNr through glutamatergic connections. Thus the indirect pathway exerts an overall opposing excitatory effect on the output structures of the basal ganglia, which results in inhibition of the thalamus. In addition to this feed-forward scheme, dopaminergic neurons in the midbrain, which are located in the VTA and in the SNc, also receive afferent projections from the striatum (Albin, Young, & Penney, 1989), and provide feedback to the striatum and to the prefrontal cortex (Parent & Hazrati, 1993; Schultz et al., 1995; Schultz, 1998). Fig-

ure 1 depicts the general anatomy of the basal ganglia.

This general scheme of connections does not portray the whole picture, and there are numerous other connections and feedback loops in the basal ganglia, which complicate the flow of information. In addition to its cortical afferents, the striatum is also innervated by the intralaminar thalamic nuclei of the thalamus (Gerfen, 1992). Striatal projection neurons are also interconnected via the striatal interneurons, as well through axon collaterals to neighboring projection neurons. The STN, which provides a major input to the pallidum and substantia nigra (pars reticulata and pars compacta), also receives a prominent topographically organized direct input from somatosensory (Parent & Hazrati, 1993), motor and premotor (Albin et al., 1989) cortices. The STN also projects back to the GPe in what could be a negative feedback loop (Albin et al., 1989). In addition, in cats and rodents, the GPe has been shown to send substantial projections directly to the reticular nucleus of the thalamus, which then projects to other thalamic nuclei. The dashed lines in figure 1 depict these additional projections. Nevertheless, the basic structure of the basal ganglia is that of a feed-forward network, as the lateral connections and the interconnections through interneurons in the striatum are weak in comparison to the large feed-forward pathways from the striatum to the thalamus (Bergman et al., 1998).

3 Functional Anatomy and Physiology of the Striatum

3.1 Physiology of the Striatum

The striatum is comprised of inhibitory GABAergic spiny projection neurons and aspiny interneurons. Spiny projection neurons are densely covered with dendritic spines, and receive converging inputs from as many as 10,000 cortical afferents, a convergence which is similar to that found in Purkinje cells in the cerebellum (Houk, 1995). On a very large scale, the input from the cortex into the striatum is organized in a general topographical man-

Figure 1: **The basic basal ganglia circuitry**

ner, which to some extent functionally defines regions in the striatum based on their cortical inputs as motor areas, association areas and limbic areas, although there are no discrete boundaries between these regions. Nevertheless, there is a considerable overlap of innervated regions, such that many cortical regions innervate one region in the striatum, and a specific cortical area usually innervates a more widespread area in the striatum, compared to the size of the originating cortical area (Gerfen, 1992). On a smaller scale, the cortical afferents are dispersed in a non-topographical manner, and adjacent cortical neurons do not innervate the same, or adjacent striatal neurons (Wilson, 1995).

The striatal spiny neurons are usually silent, but under certain conditions fire phasically (in bursts lasting tens of milliseconds to seconds) in response to massive excitatory inputs from the cortex (Gerfen, 1992; Wilson, 1995). These neurons are interconnected with inhibitory collateral connections. However, the quiescence of striatal spiny neurons is not a result of lateral inhibition per se (as, for instance, GABA antagonists do not cause tonic firing), but a result of the nonlinear properties of these neurons (Wilson, 1995). A dominant inhibitory non-inactivating potassium current which is dispersed all over the soma and the dendrites of striatal spiny neurons, shunts weak or uncorrelated synaptic inputs. This inward rectifying current ensures that only correlated synaptic inputs to wide areas of the dendritic tree, will cause sufficient depolarization. From a certain level of depolarization the electrotonic structure of the neuron rapidly collapses, the membrane resistance rises and the effective dendritic lengths shorten, thus relaxing the requirement for temporal coherence within the cortical inputs. Thus striatal spiny neurons are characterized with subthreshold UP (depolarized) and DOWN (hyperpolarized) states, and can fire readily in response to cortical inputs only when in the UP state. Lesions or deactivation of the cortex result in a constant DOWN state. Although they are mostly quiet, the spiny cells are active in maintaining their subthreshold depolarized states, and thus can be seen as constantly determining whether or not to fire in response to cortical depolarizing episodes (Kawaguchi, 1997).

3.2 Three Types of Striatal Aspiny Interneurons

The striatum seems a homogenous structure, but nevertheless, striatal neurons are heterogenous and can be categorized according to their neurochemical content and their firing patterns. The striatal aspiny interneurons can be divided into three groups (Kawaguchi, 1997): The cholinergic interneurons are large aspiny cells with 2-5 primary dendrites, that fire tonically but irregularly at 2-10Hz in response to excitatory synaptic inputs. These neurons are relatively depolarized and usually close to their firing threshold, but have long after-hyperpolarization (AHP) durations. A second group of parvalbumin containing GABAergic aspiny cells have axons with very dense collateral arborizations, and fire phasically at high frequencies in response to cortical stimulation. These neurons have more negative resting potentials and shorter duration spikes than the other interneurons. The last group of interneurons are the somatostatin/nitric-oxide-synthase containing neurons. These aspiny cells have axon collaterals that extend further than those of the other types of striatal cells. Their resting potential is close to threshold, and they release GABA, as well as nitric oxide, somatostatin and neuropeptide Y.

3.3 Spiny Projection Cells are also Heterogenous

Neurochemical and axonal tracing techniques also reveal a complex mosaic organization within the seemingly homogenous spiny projection neurons (Albin et al., 1989; Gerfen, 1992; Gerfen & Young, 1988; Kawaguchi, 1997). The striatal spiny projection neurons can be functionally divided into two groups, the patch (striosome in primates) and the matrix, which reside in spatially distinct compartments in the striatum, and have distinct projection targets. Several chemical markers distinguish between these two compartments, for instance staining for μ -opiate receptors (in rodents) and acetylcholinesterase (in cats and primates), can differentiate between patch and matrix neurons (Gerfen, 1992; Gerfen & Young, 1988).

The matrix compartment includes 85% of the striatum and projects to the GPe and GPi, and to the GABAergic SNr, while patch neurons project to the GPe and the dopaminergic SNc (Gerfen & Young, 1988). Spiny cells

in both compartments are similar in their basic electrophysical properties, and exhibit UP and DOWN states. Their inputs, however, are not similar in origin, as most cortical areas have projections to the matrix, while the patch neurons receive inputs mainly from the orbital and medial frontal cortical areas.

The cells in each compartment can be further divided into subtypes based on their output target, as some patch neurons project exclusively to the GPe, while others project to the SNc, and matrix cells project either to the SNr and to the GPi with collaterals to the GPe, or exclusively to the GPe (Gerfen & Young, 1988). The use of neurochemical markers assists in defining subsets of neuronal populations, as spiny GABAergic projection cells can contain substance P, GABA, enkephalin or dynorphin (Kawaguchi, 1997). There is also a distinction between spiny projection cells with respect to the dopamine receptors they express. Matrix neurons containing substance P or dynorphin project mainly to the GPi/SNr (direct pathway), and express the dopamine D_1 receptor, while those containing enkephalins project exclusively upon the GPe (indirect pathway) and express the dopamine D_2 receptor. Patch neurons projecting to the SNc contain mainly substance P (Albin et al., 1989; Gerfen, 1992). Substance P and enkephalin containing cells are distributed uniformly within the striatum, but thalamic afferents from the centeromedian nucleus preferentially innervate those cells which are part of the indirect pathway (Gerfen, 1992). Figure 2 summarizes the above detailed description of the striatum and its cell types.

Cell types in the direct and indirect pathways interact within the striatum through their axonal collaterals, while interaction between patch and matrix is achieved only through the striatal aspiny interneurons, as dendrites of spiny projection neurons respect the boundaries of these compartments and are restricted to their compartment only (Gerfen, 1992; Kawaguchi, 1997). Even within a compartment, long-distance interactions are presumably due to the interneurons, as the axon collaterals of the spiny cells are fairly local and are close to the cell bodies (Kawaguchi, 1997). To summarize, we have above described four types of striatal spiny projection cells (patch/ matrix, direct/indirect), which are considered to be differentially involved in basal

ganglia function (Kawaguchi, 1997).

3.4 Segregation of the Striatal Afferents

Gerfen (1992) claims that the striatal patch-matrix compartmentalization may be viewed as providing two phylogenetically distinct neuroanatomical circuits through which cortical information is processed. According to Gerfen, corticostriatal neurons in the deep parts of layer V and layer IV of the cortex provide inputs to the striatal patch compartment, while inputs to the matrix originate from superficial layer V and supragranular layers. There is also evidence that patch neurons receive cortical afferents mainly from prefrontal and limbic cortices, while the matrix receives cortical afferents from primary motor and somatosensory cortex as well as frontal, parietal and occipital cortex. Gerfen sees this as indicative of the functional significance of the striatal patch-matrix compartmentalization, which he claims appears to be related to the laminar organization of the cortex, and works to provide parallel input-output pathways through which the segregated output of sublaminae of layer V is maintained and differentially affect the dopaminergic and GABAergic neurons in the substantia nigra. He suggests that patch neurons are targeted by the allocortical areas, the older areas of the cortex which have a more diffuse and general influence, through 'non-specific' neuromodulatory feedback systems such as the cholinergic ventral pallidum and the dopaminergic neurons in the midbrain. On the other hand, the neocortex provides a more specific feedback circuitry through the matrix and its GABAergic target structures.

Similar to the striatal projection cells, the striatal aspiny interneurons also receive inputs from cortical or thalamic afferent fibers. Cholinergic interneurons receive thalamic afferents, while the other two subtypes receive only cortical inputs. The different types of glutamatergic receptors on these types of cells, as well as on the striatal projection neurons, are differentially expressed, possibly underlying variations in synaptic plasticity (Calabresi, Pisani, Mercuri, & Bernardi, 1996)(see the next section for a discussion of synaptic plasticity in the basal ganglia).

Figure 2: **Detailed description of striatal compartmentalization and cell types:** Ach - Cholinergic interneurons. Pv - Parvalbumin containing GABAergic interneurons. NO - Somatostatin/nitric oxide synthase containing interneurons which release GABA, nitric oxide, somatostatin and neuropeptide Y. P_m - Matrix projection neurons containing dynorphin or substance P (express the dopamine D_1 receptor). enk_m - Matrix projection neurons containing enkephalin (express the dopamine D_2 receptor). P_p - Patch projection neurons containing substance P (project to the SNc). enk_p - Patch projection neurons containing enkephalin (project to the GPe in the indirect pathway).

4 What are the Basal Ganglia Doing? From Structure to Function

The classical view sees the role of the basal ganglia as integrating diverse cortical inputs (indicating the state of the organism and the stimuli which it has encountered), and influencing the cortical areas which are engaged in motor planning in order to assist in choosing and executing the appropriate action. This scheme views the basal ganglia and the striatofrontal interaction as assisting in motor planning, and sequential programming of behavior (Cools, 1980; Marsden, 1986; Joel & Wiener, 1994), with the frontal cortex taking a central role in flexible behavior, planning and decision making, and the striatum controlling routine aspects of behavior (Joel & Wiener, 1994).

The large-scale anatomical features of the basal ganglia are characterized by a “funneling”, a progressive reduction in amount of cells, along the pathways leading from the cerebral cortex, through the basal ganglia and to the ventrolateral thalamus. This architecture influenced the classical view of basal ganglia functioning, which stressed the role of the basal ganglia in integrating convergent inputs from cortical association and sensorimotor areas, to common thalamic target zones. In contrast, the current view sees the functional architecture of the basal ganglia as essentially parallel in nature, and dealing not only with motor selection (Alexander & Crutcher, 1990). This view is supported by anatomical, physiological and pharmacological evidence.

The evidence is in support of an organization comprised of several structurally and functionally segregated “circuits” linking cortex, basal ganglia and thalamus. Alexander and Crutcher (1990) define five such loops, based on their cortical targets: the ‘motor’ circuit is focused on the precentral motor fields, the ‘oculomotor’ loop projects to the frontal and supplementary eye-fields, two ‘prefrontal’ circuits project to dorsolateral prefrontal and lateral orbitofrontal cortex, and the limbic circuit is focused on the anterior cingulate and the medial orbitofrontal cortex. Each circuit is composed of the same basic basal ganglia architecture, and includes a direct as well as an indirect pathway. Other authors tend to make a distinction between

only three loops in the striatum, for instance, Marsden (1986) distinguishes between a “motor loop” concerned with control of movement from sensory motor cortex via the putamen and back to the supplementary motor areas, a “complex loop” dealing with control of thought, linking frontal association cortex via caudate nucleus, and a “limbic loop” via the ventral striatum controlling emotion. Joel and Wiener (1994) and Parent and Hazrati (1993) also propose a tripartite anatomical subdivision into motor, associative and limbic loops in rats and primates. DeLong (1990) proposes a functional model of the basal ganglia motor circuit, which builds on the main organizational feature of a re-entrant pathway through which influences from specific areas of the cortex are returned to parts of those same areas after processing within the basal ganglia and thalamus. The loop he describes is, at least in part, a “closed” loop, as it originates and terminates in the precentral motor areas of the cortex.

This concept of parallel segregated processing loops conceives different areas of the basal ganglia as executing different functional roles, with little crosstalk between these parallel circuits. It is still under debate, however, whether the outputs from different regions of the striatum remain segregated as they pass through the GPi/SNr and thalamus. Using double-labeling anterograde tract-tracing methods in a primate brain, Parent and Hazrati (1993) showed that a single striatopallidal fiber arborizes twice in each pallidal segment, in accordance with similar evidence from rodents. This results in a multistriatal representation in which the same information is represented multiple times in the GPi/SNr, which adds a high degree of redundancy at the output level of the basal ganglia. Parent and Hazrati also showed that projections from small adjacent striatal cell groups do not overlap and show no convergence at the pallidal and nigral level. This allows for the finely tuned corticostriatal information to retain its specificity as it flows through the pallidum and the SNr. Thus, the redundancy in segregated striatal output can be used in order to project the same information to different targets (“multiple divergent output streams”), and thus serve the function of amplification and diversification of the striatal signal (Parent & Hazrati, 1993).

In contrast to the segregated and specific direct effects of the striatum on

groups of neurons in the GPi/SNr, the effect of the indirect excitatory STN input is probably more diffuse. Within a single band-like terminal field, the STN fibres can uniformly excite a rather large collection of neurons in the GPi/SNr, causing them to fire tonically. Since the indirect subthalamopallidal and direct striatopallidal projections converge onto the same SNr/GPi neurons, the striatum can specifically select a STN-driven subgroup over which it will exert its inhibitory control, thus disinhibiting their target thalamic neurons. In this way, the striatum can allow for the execution of a specific movement, through its influences on the premotor areas in the cortex. In such a scheme, the specificity of the control of the basal ganglia over motor behavior relies heavily upon the highly ordered organization of the direct striatopallidal projection (Parent & Hazrati, 1993).

In an interesting study of crosscorrelograms of spikes recorded simultaneously from pallidal cells in behaving monkeys, Bergman et al. (1998) investigated the question of whether the basic structure of the basal ganglia is indeed that of segregated loops of information processing or that of information sharing in a converging funnel architecture. Using multiple electrode recording in healthy monkeys performing a visual-spatial GO/NO-GO task, they investigated the degree of correlation between the firing patterns of simultaneously recorded individual pallidal neurons. Contrary to the considerable anatomical evidence supporting information sharing in a funneling architecture, they found flat crosscorrelograms indicating the absence of direct and indirect interactions between pairs of pallidal neurons. The independence of pairs of neurons was not a function of the distance between the neurons recorded, or whether the monkey was performing the task or resting. In contrast, in most other brain areas correlated firing was shown by 20-50% of the neuron pairs, and an even higher degree of correlation was found amongst the striatal tonically-active interneurons. Thus, despite the strong anatomical convergence of inputs from the striatum into the pallidal level, this study shows that most pallidal neurons are not driven by common inputs, and that the direct lateral interactions between pallidal neurons is very weak. Bergman et al. (1998) hypothesize that the neural activity in the basal ganglia reflects certain key elements of information which are

extracted from the more complex information contained in the activity of cortical neurons. Such a dimensionality reduction of information is facilitated by independent activity of neurons, which maximizes their information content. They further show that after treatment with MPTP (a dopaminergic neurotoxin which produces biochemical, anatomical and clinical changes resembling those found in humans suffering from Parkinson's disease), there ensues considerable correlation between pallidal neurons, which start firing in a synchronized fashion in low frequency periodic bursts. This firing pattern is speculated to cause the muscle tremor characteristic of Parkinsonian patients, and could be an effect of dopamine depletion providing that the dopaminergic modulation of the cross-connections between corticostriatal modules facilitates the independent action of striato-pallidal modules in the normal state. In the next chapter we will touch upon the role of dopamine in the basal ganglia circuitry, and its effects on the function of the basal ganglia, in more detail.

Joel and Wiener (1994) elaborate the segregated circuit account of basal ganglia functioning by adding an account of the inherent interactions between the different circuits, described as a "split-circuit" scheme. A split circuit originates from one frontocortical area, but through the diverging projections of the corresponding striatal region to both SNr and GPi terminates in two frontocortical areas. Each proposed split circuit consists of a closed circuit originating and terminating in the same cortical area, and another split circuit terminating in a different cortical area and subserving integrated processing. The associative split circuit includes a closed circuit terminating in the associative prefrontal cortex, and an open circuit traversing the associative GPi but terminating in the premotor cortex which is part of the motor circuit. The motor split circuit includes a closed circuit terminating in the primary motor cortex and supplementary motor area, and a split circuit traversing the SNr and terminating in the associative prefrontal cortex. The limbic split circuit also includes a limbic prefrontal cortex closed circuit, and a open circuit into the associative prefrontal cortex (and maybe another open circuit into the motor and premotor cortices). The proposed function of these processing loops is as follows: the limbic split circuits selects goals according

to information about the external and internal environment. The associative split circuit receives highly processed sensory information and selects motor programs which can achieve the goals conferred through the connection to the limbic split circuit. The motor circuit receives sensory and motor information and information regarding the selected motor program, and selects the actions which are needed in order to execute the program. Thus the different aspects of behavior are integrated to produce coherent behavioral output (Joel & Wiener, 1994).

Joel and Wiener (1994) propose a model of the interaction between the striatum and the frontal cortex, in which the major striatal projections to the frontal lobes do not directly influence behavior, but bias the selection of activity patterns by the cortical neurons. The striatum provides information regarding the most appropriate action, but other cortical biasing effects on the frontal cortex may override this input and induce the choice of a different action. In well-learned situations the striatal bias will be strong and will most likely correspond to the cortical bias, thus resulting in effortless production of routine behavior. In novel situations, however, the striatal bias will be weak and the action selection process will not be automatic, but will require a cortical supervisory process.

5 What Happens When Things Go Wrong? Basal Ganglia Disorders

There are a number of human disorders which are related to basal ganglia functioning, and specifically to the role of dopamine in the basal ganglia. These disorders offer a source of information as to the role of the intact basal ganglia in normal behavior. In the following is a short (and not comprehensive) discussion of several of these disorders, and their possible neural correlates.

Disorders of the basal ganglia are associated with a broad spectrum of clinical phenomena ranging from uncontrollable excess of movement to severe restriction of movement. Movement disorders of the basal ganglia can be

categorized into three categories (Albin et al., 1989):

1. **Hyperkinetic movement disorders** are characterized by an excess of movement which is uncontrollable and intrudes into the normal flow of motor activity. These disorders are suppressed by the administration of dopamine D_2 receptor antagonists. The most common disorder in this group is chorea, with *Huntington's disease*, a hereditary striatal neurodegenerative pathology, being the most prototypic choreoathetoid disorder. Choreoathetoid movements are also commonly observed in Parkinsonian patients overdosed with dopamine replacement therapy. Two other hyperkinetic disorders are ballism and tic syndromes (e.g. Tourette's syndrome). Hyperkinetic disorders also respond to cholinergic and anticholinergic medications, with cholinergic agonists alleviating choreoathetosis.
2. **Hypokinetic disorders** in which akinesia, bradykinesia and rigidity are the prominent features. The most well studied such disorder is *Parkinson's disease*, which results from impaired nigrostriatal dopaminergic transmission, as a result of degeneration of SNc neurons or from blockade of dopamine receptors. This disease can also be artificially induced with MPTP, which causes irreversible SNc degeneration. These disorders are pharmacologically the opposite of the hyperkinetic ones, and are relieved by dopamine agonists and cholinergic antagonists.
3. **Dystonia** - This disorder consists of the spontaneous assumption of unusual fixed postures lasting seconds to minutes. Little is known about the pathologic anatomy of dystonia. Dystonia also appears in advanced Huntington's disease. This type of movement disorder is also characterized by impairment of normal movements, in particular rapid or finely coordinated movements which become clumsy and slow.

Albin et al. (1989) proposed a model of the function of the basal ganglia based on basal ganglia anatomy, which stressed the division of the striatum into patch and matrix compartments, the interplay between the direct and indirect pathways, and the role of dopamine in the striatum. This model

rests on the belief that the basal ganglia act as regulators of cortical function via their influence on thalamocortical projections. According to this model, hyperkinetic movement disorders result from decreased activity of the STN, such as through the specific loss of striatal projections to the GPe (i.e. loss of enkephalin containing neurons in the striatum) which is found in the first stages of Huntington's disease, when chorea is most prominent. This hypothesis is strengthened by the fact that ablative striatal lesions in both humans and other mammals do not produce a hyperkinetic movement disorder, which is, however, reliably produced by lesions of the STN (Albin et al., 1989). Administration of D_2 antagonists increases the synthesis of enkephalins in the striatum (presumably indicating increased neuronal activity), indeed ameliorating hyperkinetic movements. These effects can be blocked by coadministration of scopolamine, an anti-muscarinic cholinergic agent, and are reversed by D_2 receptor agonists.

On the other hand, hypokinetic disorders such as Parkinson's disease are attributed to a loss of nigral dopamine. Impairment of dopamine action increases the activity of striatal neurons projecting to the GPe, while decreasing striatal input to the GPi/SNr (as is confirmed by the decreased levels of substance P in the GPi/SNr of parkinsonian patients). Overall, in Parkinson's disease, there is an increase in basal ganglia output, and thus impaired functioning of the motor areas in the thalamus and cortex (Albin et al., 1989). It is established that the motor deficit of Parkinson's disease results from dopamine deficiency in the putamen area of the striatum, which corresponds with loss of dopaminergic cells in the lateral substantia nigra, while dopamine depletion in the medial substantia nigra and the VTA result in cognitive symptoms rather than motor symptoms (Miller & Wickens, 1991).

Regarding dystonia, the lack of post-mortem data makes it difficult to correlate anatomic changes with clinical phenomena. Albin et al. suggest that some cases of dystonia result from gross loss of basal ganglia output rather than a specific alteration in any striatal subpopulation. In these situations the motor system continues to function but without the modulating effects of the basal ganglia.

DeLong (1990) places even more emphasis on the shifts in the balance between activity in the direct and indirect pathways and the resulting alterations in GPi/SNr output, in producing the hypokinetic and hyperkinetic features of basal ganglia disorders. In general he proposes that enhanced conduction through the indirect pathway leads to hypokinesia (by increasing pallidothalamic inhibition), whereas reduced conduction through the direct pathway results in hyperkinesia (by reduction of pallidothalamic inhibition). Indeed, after MPTP treatment of monkeys, there has been observed a significant increase in tonic neuronal discharge in GPi and STN neurons, and a decrease in mean discharge rate in the GPe. DeLong notes, however, that these results are not completely consistent with those obtained from animals studied several weeks after treatment, and may be attributed in part to the acute effects of MPTP itself. Nevertheless, lesions of the STN have been shown to reverse akinesia in monkeys made parkinsonian with MPTP (Bergman, Whitman, & DeLong, 1990). There is also evidence from MPTP animals showing enhanced phasic responses to proprioceptive stimuli and voluntary movement in GPi neurons. The results show that the phasic GPi responses are increased in magnitude but decreased in selectivity. This is consistent with evidence indicating that a loss of striatal dopamine results in an increase of transmission through the indirect pathway and a reduction of transmission through the more selective direct pathway, which in overall effect increases the output of GPi/SNr but reduces its selectivity (DeLong, 1990).

Contreras-Vidal and Stelmach (1995) model this differentiated effect on the two pathways using a neural network model, and reproduce parkinsonian symptoms. Based on the hypothesis that dopamine depletion in the nigrostriatal projection causes differential expression of D_1 and D_2 receptor mRNA, they show that dopamine depletion disrupts the balance between the direct and indirect pathways, leading to smaller than normal striatopallidal activity, and a loss of dynamic range in the activity of the GPi. Thus the result is overinhibition of the thalamus by the overexcited GPi cells. The simulations also show that this effect produces the spatiotemporal handwriting deficits seen in micrografia (a progressive decrease in letter size and slowness

in writing) in parkinsonian patients.

6 Learning in the Striatum: The Double Role of Dopamine

The role of dopamine (DA) in the basal ganglia is a complex one, with many issues still unresolved. Electrophysiological studies give conflicting results about the effect of DA on striatal neurons - in some studies DA appears to be inhibitory, while in others the effect is of excitation (Albin et al., 1989). **Dopaminergic inputs to the striatum probably play a major role in two fundamental brain functions: motor activation and reward mediated learning** (Wickens, 1990; Wickens & Kötter, 1995). Motor activation is a short-term effect, while in the long term DA modulates reinforcement learning (Wickens & Kötter, 1995). These two roles of DA in the basal ganglia are not easy to separate experimentally, as it is difficult to design a behavioral measure which will not be influenced by both acquisition of a response and performance of the response, and because drugs which effect one role will usually also effect the other. Hence the difficulty in defining the exact influence of DA in each role. Moreover, the interactions between motor activation and learning are probably not accidental and in fact reflect, in some sense, the way these two operations are related in the organization of behavior (Wickens, 1990).

6.1 The Role of Dopamine in Motor Activation

Motor activation is stimulated by the action of DA on the striatum, and inhibited by DA antagonists (as in akinesia in Parkinson's disease, which can be partly relieved by treatment with DA agonist drugs - see section 5). However, at extreme levels of DA agonist activity, the motor activation takes on the form of stereotyped behavior (Wickens, 1990). L-DOPA, a DA precursor which can cross the blood-brain barrier, has also been shown to increase the size of cortical anticipatory potentials, which can be recorded before an untriggered voluntary movement or before well attended stimuli (Miller &

Wickens, 1991). This provides evidence that DA plays a part in the production of anticipatory neural activity in the frontal cortex (Wickens, 1990).

DA also acts as a neuromodulator in the striatum, influencing the effects of other neurotransmitters. DA appears to have acute modulating effects on corticostriatal glutamatergic transmission, with the sign of the effect appearing to be concentration-dependent. DA agonists applied in low concentrations facilitate release from glutamatergic terminals, while at higher concentrations release is inhibited. Additionally, a DA receptor localized to terminals of cholinergic interneurons exerts an inhibitory effect over acetylcholine release (Wickens, 1990). In general, it seems that DA increases the output of that small number of very active striatal cells, while decreasing the output of the rest of the cells (Wickens & Kötter, 1995), thus increasing the signal-to-noise ratio in the striatal activity.

Several experimental results imply that the nigrostriatal DA projections exert contrasting effects on the direct and indirect pathways (Albin et al., 1989; Alexander & Crutcher, 1990; Gerfen, 1992). Destruction of the nigrostriatal dopaminergic projection results in decreased GABA receptor density in the GPe, and an increased receptor density in the GPi/SNr, implying that DA functionally inhibits striatal projections to the GPe, and excites striatal projections to the SNr (Albin et al., 1989)(see figure 1). Dopamine depletion in the striatum also results in increased enkephalin mRNA expression and decreased expression of dynorphin and substance P mRNA (Gerfen, 1992). DA has also been found to increase the general striatal outflow (Alexander & Crutcher, 1990). Thus, it seems that the dopaminergic input serves to enhance the effect of the direct pathway, and inhibit that of the indirect one. As a result of the opposed dopaminergic modulation of the two striatal subgroups, DA depletion disrupts the delicate balance between the direct and indirect influences of the striatum onto the thalamus and the cortex (Alexander & Crutcher, 1990; Gerfen, 1992), resulting in motor control deficiencies (also see section 5).

6.2 The Role of Dopamine in Reward-Mediated Learning

DA has been implied in reward mediated learning as a result of several experimental findings: DA neurons are activated by positive rewards such as food and liquid, and by stimuli which predict such rewards (Romo & Schultz, 1990; Ljungberg, Apicella, & Schultz, 1992). Drugs such as amphetamine and cocaine exert their addictive actions in part by prolonging the influence of DA on target neurons, and DA neurons are among the best targets for electrical self-stimulation (Schultz, Dayan, & Montague, 1997). Furthermore, direct electrical stimulation of midbrain dopaminergic areas or injection of DA agonists into the neostriatum can produce conditioning effects similar to those produced by natural rewards (Wickens & Kötter, 1995), and DA antagonists have been shown to block the rewarding effects of food and water on lever pressing by rats (Wise, Spindler, & Legault, 1978). These findings implicate the nigrostriatal projection as common pathway for signalling positive reinforcement (Wickens & Kötter, 1995).

In reward mediated learning, DA appears to mediate synaptic learning in the corticostriatal pathway in an activity dependent way, such that it enhances the corticostriatal synapses specifically involved in the response that led to reward being obtained (Wickens, 1990; Wickens & Kötter, 1995; Miller & Wickens, 1991; but see Schultz et al., 1995). If a certain cortical input coincides with the occurrence of reward, these corticostriatal synapses which were active will be strengthened, thereby creating a positive feedback loop which increases the likelihood that when the same stimulus conditions occur again, the previously-rewarded response will be performed (Wickens & Kötter, 1995) (also see in this regard the discussion of the formation of corticostriatal cell assemblies in section 7). In this form of learning, DA agonists facilitate, while antagonists impair the effect of reward on learning (Wickens, 1990).

6.2.1 The Dopaminergic Reinforcement Signal

The basic pattern of dopaminergic activity in reinforcement learning situations was revealed in the studies of Schultz and colleagues (1993; Ljungberg et al., 1992; Schultz et al., 1995; Schultz, 1997). Recording extracellularly the activity of dopaminergic (DA) neurons in the VTA of monkeys during the acquisition and performance of behavioral tasks, Schultz et al. found that DA neurons respond phasically to a wide range of stimuli which can be characterized as primary rewards (Schultz, 1998), and as the experiment progresses, the response of these neurons gradually shifts back in time from the primary reward to the first phasic reward-predicting stimulus (Ljungberg et al., 1992)(see figure 3). This shift in DA activity strongly resembles the shifting of an animal's appetitive behavioral reaction from the unconditioned stimulus to the conditioned stimulus in classical conditioning experiments (Schultz et al., 1997). The phasic dopaminergic response occurs in 55-75% of the midbrain dopaminergic neurons, and occurs independently of an action or learning context (Schultz, 1998). The dopaminergic response does not distinguish between different stimuli (e.g. food, liquid, light) and different modalities of conditioning stimuli, however they discriminate between appetitive and neutral or aversive stimuli (Ljungberg et al., 1992). In all cases, the response is to the onset and not the offset of the stimulus (even if the offset predicts the reward), and only stimuli with a clear onset are effective in eliciting a dopaminergic response. Furthermore, Schultz et al. (1993; Ljungberg et al., 1992) showed that if the expected reward is omitted, there is a depression in the dopaminergic signal at the time of the expected reward, and that early reward delivery results in a dopaminergic response to the reward, and no depression at the original time of expected reward (Schultz, 1998). Midbrain dopaminergic neurons also respond to novel or very salient stimuli (even when the stimulus is not rewarding), and show a generalization response to stimuli which closely resemble reward-predicting stimuli. These responses consist of an activation and are usually followed by an immediate depression (Ljungberg et al., 1992). The generalization response is recognized as a mistake, as there is no depression when the reward fails to appear

Figure 3: **Schematic representation of the dopaminergic response to several experimental conditions.** **CS+** conditioned stimulus which predicts reward, **CS-** stimulus which resembles the reward predicting stimulus, but does not predict reward.

(Schultz, 1998).

Thus DA can be regarded as signalling the occurrence of unpredicted rewarding stimuli, or the earliest reward-predicting stimuli. The firing pattern of DA neurons also reflects information regarding the timing of delayed rewards (relative to the reward-predicting stimulus), as can be seen by the precisely timed depression of spontaneous DA activity when an expected reward is omitted. In terms of formal theories of learning, dopaminergic neurons report the difference between the occurrence and the prediction of reward (Schultz et al., 1995; Schultz, 1998).

6.2.2 The Connectivity of the Dopaminergic System: Afferent Connections

It follows from the above, that dopaminergic neurons must have access to information about the predictive value of stimuli encountered by the organism, the time of the predicted reward, and the amount of reward actually

obtained. As described in sections 2 and 3, striatal neurons provide a prominent GABAergic projection to the dopaminergic neurons in the SNc and the VTA. According to Schultz (1998) this projection may exert two opposing effects: a direct inhibition and an indirect activation via the striatal inhibition of SNr neurons which subsequently inhibit SNc neurons via their local GABAergic axon collaterals. Thus the striosomes and ventral striatum may monosynaptically inhibit and the matrix may indirectly activate DA neurons.

Wickens and Kötter (1995) propose that neural activity in the monkey ventral striatum which coincides with the delivery of the reward (Schultz, Apicella, Scarnati, & Ljungberg, 1992) may inhibit the firing of dopaminergic neurons at the time of the reward, while another source of excitatory input, possibly from the amygdala, is the one responsible for the anticipatory dopaminergic activity at the time of the first reward-predicting stimulus. However, Schultz (1998) claims that the positions of these striatal neurons relative to matrix and patch is not yet known, and striatal activations reflecting the time of expected reward have not yet been reported. Schultz (1998) hypothesizes that the polysensory reward responses of DA may be the result of feature extraction in cortical association areas, which is conveyed to the striatum, and through a double inhibition to DA neurons in the SNc. In rats there could even be a faster influence of the cortex on dopaminergic neurons, through a direct projection from the frontal cortex (which is however weak in monkeys).

Another source of inputs to midbrain dopaminergic neurons is the brainstem pedunculopontine nucleus (PPN). This nucleus is an evolutionary precursor of the substantia nigra, and in mammals it sends strong short latency excitatory glutamatergic and cholinergic influences to a large fraction of DA neurons (Schultz, 1998). Yet another massive and probably excitatory input to DA neurons arises from nuclei in the amygdala, which respond to primary rewards and reward predicting visual and auditory stimuli, without discriminating between predicted and unpredicted rewards, or between appetitive and aversive rewards (Schultz, 1998; Wickens & Kötter, 1995). A monosynaptic inhibitory projection from the dorsal raphe nucleus can contribute to

the DA response to novel and salient stimuli (Schultz, 1998).

Schultz (1998) summarizes that activation of DA neurons by primary rewards and reward predicting stimuli may be mediated through the double inhibitory striatal matrix influence, the PPN or possibly also by reward expectation-related activity in the STN. The striosomal inhibition which cancels out the matrix inhibition can explain the null-response to predicted rewards and the depression when a predicted reward is omitted.

6.2.3 The Connectivity of the Dopaminergic System: Efferent Connections

Each DA axon has $\sim 500,000$ striatal varicosities from which DA is released. As a result, there is DA input to almost every striatal neuron. Dopaminergic synapses are also found in layers I and V-VI of areas of the frontal cortex. The homogenous DA signal thus provides a parallel wave of activity which advances from the midbrain to the striatum and frontal cortex (Schultz, 1998). As a result of rapid saturation of the DA reuptake transmitter, a bursting DA response brings about much higher levels of DA compared to those resulting from a slow firing response.

This DA signal is received by two principal types of DA receptors: D_1 receptors constitute $\sim 80\%$ of DA receptors in the striatum, and are located predominantly on projection neurons belonging to the direct pathway, and D_2 receptors which are located predominantly on neurons projecting to the GPe. The D_2 receptors have a higher affinity to DA than the D_1 receptors. Schultz (1998) hypothesizes that as a result of this, the burst of DA released after rewards or reward predicting stimuli would affect the D_1 receptors near the DA release sites and all D_2 receptors, while the depression in DA as a result of omission of an expected reward would mainly reduce the tonic stimulation of D_2 receptors by ambient DA, so that a negative prediction error may only affect neurons projecting to the GPe.

6.2.4 Influence of Striatal Activation on Dopaminergic Neurons

Emanating from a close examination of the topography and relationships within the intricate connectivity of the dopaminergic system and the striatum, Joel and Wiener (2000) set forth a hypothesis pertaining to the influence of the activity of a subset of striatal neurons on the DA neurons. The active matrix neurons globally inhibit DA dendrites, but also inhibit a corresponding subset of SNr neurons which then release a subset of DA cells from inhibition. Through the indirect pathway, the striatal matrix neurons phasically inhibit a subset of GPe neurons which then release inhibition from the STN and SNr. The disinhibited STN serves to directly increase DA excitation, with the SNr acting as a limiting factor of this effect. Concomitantly active patch neurons in the striatum exert a direct inhibiting effect on DA cells, as does the “patch”-like limbic striatum. The “matrix”-like limbic striatum disinhibits a subset of DA cells through an indirect influence mediated by inhibition of the SNr. This pattern of activity results in localized DA burst firing, on the background of global inhibition of DA neurons, and an overall effect of increased signal to noise ratio in striatonigral transmission.

Joel and Wiener (1994, 2000) claim that the general organization of the connections between the striatum and the dopaminergic system also follows the open interconnected split-circuit principle (Joel & Wiener, 1994). Each nigral split circuit incorporates a closed striatum-nigra-striatum loop, while the limbic striatum is also the source of two open loops, one connecting it to the associative striatum and the other connecting it to the motor striatum.

In summary, Joel and Wiener (2000) suggest that the direct inhibitory effect of each striatal subregion on its DA input provides a mechanism which restricts striatal learning in well-learned situations (produces the null-response of DA to predicted rewards), whereas the indirect excitation maintains an optimal striatal DA level needed for the execution of well learned behaviors. Furthermore, the influence of the limbic striatum on the motor and associative regions permits it to influence learning and execution of learned behaviors so as to adjust them to the motivational state of the organism, and foster attainment of the current goal.

6.2.5 Reward-Mediated Learning: A Cellular Mechanism

The computational model of reward mediated learning consists of the selection and strengthening of particular stimulus-response (S-R) connections, which are chosen by a reward signal from a large set of alternative S-R connections. The striatum possesses the necessary anatomical features required for such a mechanism. The striatum receives inputs from large areas of the cortex, with an individual cortical neuron sending efferents to many striatal neurons. Nevertheless, a single striatal neuron receives only a relatively small influence from each cortical afferent, and the physiological properties of these neurons require a coherent input from many cortical afferents in order for the neuron to shift to an UP state in which it can fire (Wickens, 1990; Kawaguchi, 1997). Thus striatal projection neurons can be regarded as a pattern detectors, firing only when a certain pattern of cortical activity occurs.

Furthermore, the convergence of cortical and DA inputs on the same dendritic spines of striatal neurons provides a mechanism for a three factor learning rule (Wickens & Kötter, 1995). Figure 4 depicts the convergence of excitatory glutamatergic cortical afferents, and neuromodulatory midbrain dopaminergic inputs, on a striatal spiny cell. The dendritic spines of each striatal neuron are contacted by an estimated number of 10,000 cortical terminals and 1,000 DA varicosities, with DA varicosities forming synapses on the same dendritic spines of striatal neurons, that are contacted by cortical glutamate afferents (Schultz, 1998). This anatomical arrangement, together with the homogenous DA signal, provides an ideal substrate for DA-dependent synaptic changes at the spines of striatal neurons, and probably also in some (but not all) corticocortical synapses.

Dopamine has been shown experimentally to have a modulating effect on learning in the striatum, with brief exposures to DA producing long-lasting facilitation in the cortico-striatal pathway. Specifically, activation of D_1 receptors increases intracellular cAMP concentration, which can influence long-term enhancement of synaptic efficacy (LTP) (Wickens & Kötter, 1995). Wickens and Kötter summarize electrophysiological experiments which test

Figure 4: Convergence of excitatory glutamatergic cortical afferents and neuromodulatory midbrain dopaminergic inputs, on a the spine of a striatal projection neuron.

the possible combinations of activation of the presynaptic cortical neuron, the postsynaptic striatal neuron, and the dopaminergic input, to the same spine. The data, although not entirely consistent, supports the proposed three-factor learning rule, and shows that if both the pre- and postsynaptic neurons were active, in the presence of DA, there ensues LTP of the synapse, while the absence of DA reverses the effect and induces LTD. The data also suggests that the precise timing of the DA signal in relation to the pre- and postsynaptic activity is critical.

A computational requirement for selective reward mediated learning is that the synapses that participated in eliciting the action that brought about the reward, will be somehow singled out for enhancement. As the reward may be delayed, and in any case follows the action and does not arrive simultaneously with the activation of the relevant synapses, there is a need for some sort of trace that will be left in these synapses, and will mark them as eligible for enhancement upon arrival of the reward signal (Wickens, 1990). according to operant conditioning experiments, as well as intracranial

self-stimulation tests, the duration of this eligibility trace must be on the order of seconds. This computational concept of a synapse-specific “eligibility trace” may be implemented in the striatum through elevated intracellular calcium Ca^{++} levels. In wide areas of the brain, elevated intracellular calcium levels appear to be a necessary and sufficient condition for LTP to ensue. In the striatum, however, it seems that while elevated calcium levels are necessary, they are not sufficient, and there is also a need for elevated DA concentration. Under this assumption, and regarding the dendritic spine head as a chemically isolated compartment, the elevated calcium level (which can persist for some time), in a spine head containing a synapse that has been active, can serve as an “eligibility trace” for that synapse (Wickens, 1990; Miller & Wickens, 1991; Wickens & Kötter, 1995). Dopamine, in turn, acts to diffusely signal the occurrence of an unexpected reward, and thus enables the strengthening of the synapses located on the eligible spines. This allows for a retroactive effect of the reward in strengthening corticostriatal synapses, in instrumental conditioning.

Such a framework implies a multiplicative interaction between Ca^{++} and DA. For this to happen, the sites of Ca^{++} influx must be located sufficiently close to the dopamine receptor complexes, and there is a need for a dopamine *and* calcium dependent pathway which leads to synaptic enhancement. A possible mechanism proposed by Wickens (1990; Wickens & Kötter, 1995) is the phosphorylation of DARPP-32 (dopamine and cAMP regulated phosphoprotein 32) which is regulated by DA acting through cAMP, and may mediate specific interactions between DA and glutamate, acting through Ca^{++} . In this scheme, glutamate leads to increased Ca^{++} levels (through voltage dependent channels), and activation of D_1 receptors increases intracellular cAMP. This activates a cAMP dependent protein kinase which brings about the phosphorylation of DARPP-32. When phosphorylated, DARPP-32 inhibits protein phosphatase, which is known to reverse the effect of some protein kinases thought to be important in calcium-activated synaptic modification, such as calcium- and calmodulin-dependent kinase II. Thus the elevated Ca^{++} concentration in the spine functions as a “state of readiness” for enhancement by DA.

6.3 Two Roles Subserved by Two Dopamine Receptors?

The two different, and sometimes contradicting roles of DA in motor activation and reward mediated learning, are maybe mediated through different dopaminergic receptors (Wickens, 1990). In low concentrations DA agonists have been shown to facilitate release from glutamatergic terminals in the striatum, an overall excitatory effect which can be blocked by D_2 antagonists. At higher concentrations of DA, however, release from glutamatergic terminals is inhibited, an effect which can be blocked by D_1 antagonists. Different actions of the two populations of receptors involved could explain this dual effect of directly applied DA. Thus dopamine D_1 receptors may mediate long-term reward-related enhancement of corticostriatal synapses, while D_2 receptors on cholinergic terminals may mediate a short-term indirect inhibitory effects of DA on striatal neurons (Wickens, 1990).

Motor activation effects of DA agonists have been reported to last for several minutes. Some of the motor activating effects may be mediated by the D_2 inhibitory effect of DA on cholinergic activity in the striatum, as anticholinergics also produce motor activation (although they do not produce the full picture of DA motor activation). In particular, D_2 agonists inhibit acetylcholine release via a DA receptor localized on terminals of cholinergic interneurons. Inhibition of striatal projection neurons due to D_2 activation would presumably affect those neurons which are part of the indirect pathway more strongly (see figure 2), thus inhibiting the GPi/SNr, releasing inhibition from the thalamus, and producing motor activation (Wickens, 1990).

In contrast, D_1 selective DA agonists mimic the increased striatal outflow brought about by apomorphine and L-Dopa. Thus the D_1 receptors which are located on the striatal projection neurons themselves, seem to be the ones involved in increased striatal outflow. As these receptors affect mainly the neurons belonging to the direct pathway, the overall effect is again one of disinhibition of the thalamus. Nevertheless, to reproduce the full syndrome of motor activation brought about by DA agonist drugs, both enhancement of corticostriatal synapses (D_1 effect), and a decrease in interneuron cholinergic

activity (D_2 effect) are required (Wickens, 1990). It should be noted, however, that in reward mediated learning, synaptic enhancement is contingent upon reward being obtained, whereas in motor activation brought about by DA agonists, the continued high level of DA enhance the involved synapses indiscriminately, maybe resulting in the long-term inhibitory effects of high concentrations of DA.

The location of the two kinds of DA receptors in the striatum is not yet clear. There is a general consensus regarding the location of D_1 receptors which are localized to intrinsic neurons of the striatum (Wickens, 1990). The situation regarding D_2 receptors is not as clear: although there is controversial evidence for a presynaptic location on the cortical afferents, these receptors are assumed to be located presynaptically on the terminals of intrinsic cholinergic interneurons (Wickens, 1990). Although the majority of striatal projection neurons express only one of the DA receptors, there is also a considerable synergistic effect of D_1 and D_2 receptors, which may mean that either there is a subset of neurons which express both, or that there are interactions between the axon collaterals of the two types of neurons (Gerfen, 1992).

7 Computational Models of the Basal Ganglia

Several computational models of the basal ganglia have been proposed. These models build on the basic features of the basal ganglia: the segregated loops of processing, the specialization of striatal spiny neurons for cortical pattern detection and the basic division of the striatum to patch and matrix compartments (see Beiser, Hua, & Houk, 1997, for a short review).

7.1 Models of the Striatum

7.1.1 Is the Striatum a Competitive Network?

Emphasis on mutual competition and pattern recognition features, have led to the traditional computational view of the striatum as a competitive pattern classification network (e.g. Miller & Wickens, 1991; Wickens & Kötter, 1995). Such models which rely heavily on winner-takes-all dynamics have been challenged by Wilson, who claims that there is no evidence for a competitive network structure in the striatum. A competitive “winner-takes-all” network architecture requires that the neurons receive approximately the same inputs (anatomically; physiologically they will have different synaptic weights), and that they will compete amongst themselves for representation of the input via inhibitory interconnections. The striatum does not fit into this computational structure, as it is very sparsely connected, with each striatal neuron receiving a unique set of cortical inputs not common to many (if any) other striatal neurons (Kincaid, Zheng, & Wilson, 1998), and the lateral inhibition between two striatal neurons presumably responsible for the competition dynamics, is so weak that it is very difficult to demonstrate. In particular, both direct depolarization and antedromic excitation of striatal neurons do not elicit lateral inhibitory post-synaptic potentials, probably due to a lack of GABA receptors in the local axon collaterals, which may only operate through enkephalin, dynorphyn and substance P (Jaeger, Kita, & Wilson, 1994). This suggests that the fast inhibition seen in the striatum after afferent stimulation is not the result of lateral inhibition between neighboring neurons, but likely arises from a feed-forward pathway through a GABAergic interneuron (Wilson, Kita, & Kawaguchi, 1989).

7.1.2 Parallels Between the Striatum and the Cerebellum

A model of the striatum which does not rely on lateral inhibition is that of Houk (1997; Houk & Wise, 1995). These authors stress the similarities between the modular architecture of the basal ganglia and the cerebellum and the similar organization of their connections to the prefrontal cortex, and suggest that these similarities reflect similarities in information process-

ing. Both structures have been traditionally assumed to play a role in motor learning and control, but have recently also been implicated in cognitive processing, suggesting that ideas from the role of the segregated loop organization in motor control and initiation of action, may also be extended to cognitive processing and initiation of thought. Purkinje cells in the cerebellum and spiny neurons in the striatum are specially suited for parallel search pattern recognition as a result of: 1. the high convergence of diverse input patterns on these neurons (200,000 parallel fibres on each Purkinje cell and 10,000 cortical afferents to each striatal spiny neuron, organized in a way that maximizes the probability that each input configuration is different), 2. the existence of a specific training signal for modifying their many synaptic weights (through the climbing fibers in Purkinje cells and through the nigrostriatal DA input in the striatum), and 3. their sharp thresholds for state transitions in membrane potential, which can create clean decision surfaces for distinction between patterns.

Area 46 of the prefrontal cortex, a major candidate for the location of working memory, receives prominent projections from both the cerebellum and the basal ganglia. Houk thus hypothesizes that initiation of working memory discharge in the frontal cortex is controlled by burst discharge in striatal spiny neurons (analogous to the contribution of the basal ganglia to the initiation of motor actions such as saccadic eye movements). He further hypothesizes that spiny neurons, which are known to respond to sensory and motor inputs in a context-dependent manner, are trained by the reward-predicting DA input signal, to detect particular contextual events which are important for planning and controlling movement, or for controlling thinking. This hypothesis could explain the deficiency in initiation of movements and thoughts in parkinsonian patients, which could be a result of an impaired ability to recognize the relevant contexts for such initiation, as a result of DA depletion (Houk, 1997).

7.1.3 Cortical Assemblies and Learning in the Striatum

From a somewhat different perspective, Miller and Wickens (1991) hypothesize a role for the basal-ganglia feedback to the cortex in establishment and stabilisation of cortical cell assemblies. Cortical cell assemblies are a concept originally postulated by Hebb, and consist of mutually excitable distributed groups of neurons which supposedly serve to encode meaningful groupings of information in a diffuse and redundant way. Cortical cell assemblies can also serve as a working memory, and show anticipatory neural activity. The formation of cortical assemblies is dependent upon positive feedback loops, which eventually have a tendency to grow exponentially. As a result, there is a need for a negative feedback which will stabilize these assemblies and maintain a balance of their activity so as to prevent them from exploding into maximal activity of their excitatory connections, and losing the specificity of their representation. Miller and Wickens postulate that the basal ganglia provide such a mechanism that ensures overall stability of neural activity levels. They postulate that each cortical cell assembly also includes a small number of striatal cells, with winner-takes-all competition dynamics between the striatal neurons ensuring that only one cell assembly will be active at each time. In this way, the inhibitory effect on the cortical cell assemblies is a result of the competitive inhibition within the striatum, such that the striatal neurons which are active provide a positive feedback to their assembly, while through the inhibition of their neighboring striatal neurons they inhibit other competing cell assemblies.

On the other hand, Miller and Wickens argue that DA plays a role in establishing these neural cell assemblies, through reward-mediated learning when a sequence of events occurs in a regular predictable association. Miller and Wickens portray a scheme in which DA serves to strengthen the corticostriatal connections which were active as a result of context/event E_1 , which was reliably followed by a different context E_2 . As a result of the strengthening of corticostriatal connections, the striatum reliably produces activation of a different set of cortical cells, which through Hebbian plasticity, become associated with the firing of the initial cortical neurons, thus forming

an assembly which is quite independent of the striatal activity. This scheme, however, requires that DA neurons be activated by significant, although motivationally neutral events, such that, as the authors assume, “the occurrence of an event which can be predicted from a previous event is, in some sense, rewarding” (p. 83). They argue that this could be achieved through the influence of the cortex on midbrain dopaminergic neurons either directly or through the amygdala.

After a cell assembly which is relevant to context E_1 will be strengthened, and after it will precede event E_2 many times (and thus precede the activity of the E_2 cortical assembly), through Hebbian learning, and assuming that the striatal neurons participating in the E_1 assembly also innervate some of the E_2 assembly cortical neurons, the E_1 assembly will learn to activate the E_2 assembly even before E_2 occurs, hence producing anticipatory cortical activity. However, Miller and Wickens note that this does not provide a sufficient explanation for anticipatory neural activity, which should also precede E_1 (or any self initiated action).

This elaborate model proposed by Miller and Wickens, is posed as an alternative to the traditional S-R view of basal ganglia learning. However, their model suffers from a number of flaws. First, their cortical regulation model relies heavily on competitive dynamics in the striatum, an assumption which has been challenged severely. They further propose that the competitive dynamics between cortical cell assemblies rely on very few striatal neurons which participate in each assembly, maybe even one cell. This is reminiscent of “grandmother cells” which are not a very favorable concept from the practical aspect. The assumption that dopaminergic neurons fire not only in response to unpredicted rewards, but also in response to the occurrence of a predicted event, is supported by scanty evidence, if at all, as also noted by the authors. Finally, their account of the production of a new response in order to achieve a learned goal, which is the basis of their objection to the traditional S-R framework, is incomplete and unconvincing.

7.2 Actor-Critic Models of the Basal Ganglia

Several researchers find a strong resemblance between the architecture and properties of the basal ganglia, and the computational framework of actor-critic models (Barto, 1995; Houk, Adams, & Barto, 1995). Actor-critic models are a subset of computational models of reinforcement learning. In such models, an actor sub-network learns to perform actions so as to maximize the weighted sum of future rewards, which is computed at every timestep by a critic sub-network (Barto, 1995). The critic is adaptive, in that it learns to predict the weighted sum of future rewards based on the current sensory input and the actor's policy, by means of an iterative process in which it compares its own predictions to the actual rewards obtained by the acting agent. The learning rule used by the adaptive critic is the temporal difference (TD) learning rule (Sutton, 1988) in which the error between two adjacent predictions (the TD prediction error) is used to update the critic's weights. The analogy between the basal ganglia and actor-critic models builds on the strong resemblance between DA activity (as described in section 6.2.1) and the prediction-error signal in TD models (Sutton, 1988), and between between DA-dependent long-term synaptic plasticity in the striatum (as described in section 6.2) and learning guided by a prediction error signal in the actor (Barto, 1995).

7.2.1 The Computational Actor-Critic Model

Barto (1995) puts forth an concise account of the actor-critic model. An actor (a "controller" network) chooses actions to be performed by a system embedded in an environment. The environment provides the system with feedback signals, which can influence the actor's choices. A critic network observes the consequences of the actor's behavior, and iteratively learns to predict the consequences of different actions in different contexts. In most artificial reinforcement learning systems, the critic's output at any time is a number scoring the previous action. This score is a measure of the weighted sum of the expected future primary rewards, with the weights preferring immediate rewards to more distant ones. This evaluative signal is fed into

the actor, and used in subsequent choice making. This way, the actor learns according to the tactical objective of maximizing the immediate evaluative signal, which encompasses in it the strategic objective of maximizing a long-term measure of rewards. The system as a whole performs actions and learns to evaluate them simultaneously, in an iterative process in which the actor constantly learns to better choose its actions based on the evaluation signal, and in the same time the critic improves the accuracy of its evaluations, based on the consequences of the chosen actions.

The difficulty in such a framework is in learning to predict the future rewards from the information of the current timestep, without requiring an infinite amount of storage. the solution to this problem, however, is quite simple and is based on a consistency condition according to which, given the correct predictions, the prediction in the current timestep should equal that of the previous timestep less the currently obtained primary reward. The error by which two adjacent predictions fail to satisfy this condition (i.e. the difference between the current prediction and reward, and the previous prediction), is the temporal difference error, which is used by the critic network to update its weights. This TD-error based learning is in agreement with the classical conditioning Rescorla-Wagner learning rule. Barto (1995) mentions several guidelines for a neural implementation of such a learning architecture: all the modifiable synapses need a local eligibility mechanism so as to permit updating of the active synapses in subsequent timesteps; the actor units should compete between themselves so that only one unit (action) wins the competition; the modifiable synapses of the prediction unit must use a two-factor learning rule, whereas those of the actor units must use a three-factor learning rule, and finally, some mechanism is required to compute the difference between subsequent predictions (“a kind of neural differentiator” - p.229)

7.2.2 Examining the Basal Ganglia through the Actor-Critic Framework

In accordance with the basic assumption of this formal learning theory, Waelti, Dickinson, and Schultz (2001) showed that learning in dopaminergic neurons (as well as behavioral learning) occurred predominantly when DA neurons registered a reward prediction error at the time of the reward. Using the blocking paradigm they showed that when paired to a previously learned stimulus which predicts the reward, the reward predicting nature of the new stimulus is not learned by the dopaminergic neurons. It appears that learning in DA neurons is indeed driven by the reward prediction error.

Houk et al. (1995) propose that the activity of DA plays the dual roles of TD error and evaluation signal in a neural implementation of the actor-critic architecture in the basal ganglia. The homogenous nature of the dopaminergic response to unpredicted rewards and reward predicting stimuli (Schultz, 1998) can serve as a scalar evaluation signal to its targets in the striatum and frontal cortex. Houk et al. introduce a neural model which addresses the two problems of reinforcement learning: the spatial and temporal credit assignment problems of providing the reinforcement signal to the right synapses at the right time. The model stresses the inputs from the striatal striosome compartment to DA neurons, attempting to explain how such inputs can account for the transition of the DA response to the earliest reward predicting stimulus. The model is based on several anatomical and physiological features: 1. The direct influence of the striosome neurons on the SNc and VTA which is inhibitory, while the indirect influence through the STN has a net excitatory effect. 2. The electrical properties and the anatomical connections of striatal spiny neurons, as well as their involvement in reward-mediated learning, which make them ideally suited for the recognition of complex patterns of cortical afference. This specialization allows them to detect particular contexts which are valid predictors of reinforcement. 3. The reciprocal nature of projections between clusters of striatal spiny neurons and DA neurons, which is considered by Houk et al. a well-established aspect of striosomal neurons.

In addition to the striatal inputs to the DA neurons, Houk et al. hypoth-

esize another DA afferent signalling the occurrence of primary reinforcement, which may originate in the hypothalamus. Assuming that the striatal neurons have already learned to fire a burst of discharge when the predicting context occurs, and assuming that the striatal excitation of DA neurons precedes the inhibition, Houk et al. (1995) show that the resulting dopaminergic firing pattern displays the observed shift backwards in DA response. They note, however, that the model assumes a time course of inhibition which is slower than that observed in electrophysiological experiments. This proposed model of a striosome module fulfills the main functions of the adaptive critic in the actor-critic architecture.

Similarly Montague and colleagues (Montague, Dayan, Person, & Sejnowski, 1995; Montague, Dayan, & Sejnowski, 1996; Schultz et al., 1997) use fluctuations in DA activity (around a certain baseline activity) as both a supervisory signal for synaptic weight changes and as a signal which directly influences the choice of actions in models of bee foraging, rat navigation and human choice-making. In this model, the weights between sensory inputs (which report the difference between the amount of a certain sensory stimulus in the current and the previous timesteps) and the error prediction neuron, are updated only when a prediction error occurs. Thus the weights change until the sensory stimuli fully predict the rewards. Using this simple learning network combined with a very simple probabilistic action function which is dependent on the small DA fluctuations, Montague et al. (1995) simulate the foraging behavior of a bumblebee with close resemblance to the behavior of real bumblebees foraging for nectar.

In order to further model the precise timing information which is represented in the dopaminergic reward prediction signal, Montague et al. (1996; Schultz et al., 1997) use a complete serial-compound-stimulus representation in which the time interval between the stimulus and the reward is divided into timesteps, and an input component (and adjustable synaptic weight) is dedicated to each timestep. The learning in the model is Hebbian-like, and the authors show that the resulting behavior of the model dopaminergic neuron is remarkably similar to the activation patterns recorded from DA neurons in the midbrain.

8 Summary: Reinforcement Learning and the Basal Ganglia

The basal ganglia are a very intricate and complex structure, with wide variations in neuronal types and properties, and vast and diverse connectivity within the different substructures and to other regions of the brain. The classical feedforward processing models of the basal ganglia have indeed been shown to be simplistic, but still may correctly capture some of the essence of the information processing which takes place in the basal ganglia. More elaborate computational models and formal learning theory have considerably advanced our understanding of basal ganglia functioning, illustrating again the beneficial interplay between computational modelling and rigorous experimental research. Especially in light of the complexity of the underlying structure and mechanisms, it is extremely important that the theory drive the experiments, and the experimental data drive the models in order to improve the theory.

Nevertheless, the research into the computational role of the striatum also reminds us to treat the computational paradigms with caution, and not fall into the trap of only finding what we are looking for, in order to corroborate a well understood computational model. This point is illustrated nicely by the classical view of the striatum as a competitive winner-takes-all network, which, if not only for its simplistic beauty and elegance, was dominant for several years, but has lately been challenged severely by experimental data. This should also be taken into account when viewing the models of reinforcement learning in the basal ganglia. Although a very capturing concept, the basal ganglia may not fit entirely into the actor-critic framework, and care should be taken not to force-fit the data to the models.

The most promising analogy is between the dopaminergic neural activity and the formal notion of a temporal difference reward prediction error. The similarity between the formal concept and the experimental data here is truly intriguing, and should be explored in further research. The question of how the basal ganglia generate the characteristic pattern of dopaminergic activity has yet to be fully answered. The fact that very simple models such as

that of Montague et al. (1995) can capture the essence of the dopaminergic response seems to imply that the neuronal mechanism responsible for the dopaminergic signal need not be very elaborate, and can rest on a relatively simple and robust architecture. Nevertheless, finding the neural correlates to the abstract computational notions is not straightforward. For instance, it is still not clear where the “neural differentiator” described by Barto (1995) and used in all the adaptive critic models, resides.

Other issues such as the interplay between the novelty response of dopaminergic neurons and reinforcement learning have not been tackled at all. The fact that the robust widespread dopaminergic signal conveys information about rewarding stimuli and novel unrewarding stimuli alike, may prove to be very important to the process of reinforcement learning. The nature of such an interaction may be quite difficult to assess, as is the nature of the interaction between the two roles of dopamine. Indeed the prediction error signal in the formal actor critic model serves the dual role of influencing action selection as well as learning, but the fact that these two aspects of dopaminergic information processing are almost inseparable experimentally deems the full elucidation of the role of dopamine in the basal ganglia circuitry a formidable task.

In this work I have only very briefly touched upon the issues of how an organism learns via reinforcement learning, and how is this process connected to the neural substrates in the basal ganglia. Reinforcement learning is such a fundamental learning process, that understanding the mechanisms involved in this process can considerably advance our understanding of brain functioning. As this learning process is so widely used, I would assume that not only one brain area is specialized for reinforcement learning, and furthermore, I would expect the underlying architecture and principles to be relatively simple and robust. Thus, to my opinion, it should prove very worthwhile to further investigate and exploit the similarities between learning in the basal ganglia and learning in other structures such as the cerebellum, and even the cortex.

References

- Albin, R., Young, A., & Penney, J. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neuroscience*, *12*(10), 366-375.
- Alexander, G., & Crutcher, M. (1990). Functional architecture of basal ganglia circuits: Neural substrates of parallel processing. *Trends in Neuroscience*, *13*(7), 266-271.
- Barto, A. G. (1995). Adaptive critic and the basal ganglia. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 215-232). Cambridge: MIT Press.
- Beiser, D. G., Hua, S. E., & Houk, J. C. (1997). Network models of the basal ganglia. *Current Opinion in Neurobiology*, *7*, 185-190.
- Bergman, H., Feingold, A., Nini, A., Raz, A., Slovin, H., Abeles, M., & Vaadia, E. (1998). Physiological aspects of information processing in the basal ganglia of normal and parkinsonian primates. *Trends in Neuroscience*, *21*(1), 32-38.
- Bergman, H., Whitman, T., & DeLong, M. R. (1990). Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science*, *249*, 1436-1438.
- Calabresi, P., Pisani, A., Mercuri, N. B., & Bernardi, G. (1996). The corticostriatal projection: From synaptic plasticity to dysfunctions of the basal ganglia. *Trends in Neuroscience*, *19*, 19-24.
- Contreras-Vidal, J. L., & Stelmach, G. E. (1995). A neural model of basal ganglia-thalamocortical relations in normal and parkinsonian movement. *Biological Cybernetics*, *73*, 467-476.
- Cools, A. R. (1980). Role of the neostriatal dopaminergic activity in sequencing and selecting behavioural strategies: Facilitation of processes involved in selecting the best strategy in a stressful situation. *Behavioural Brain Research*, *1*, 361-378.

- DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neuroscience*, 13(7), 281-285.
- Donahoe, J. W., & Packard-Dorsel, V. (Eds.). (1997). *Neural network models of cognition: Biobehavioral foundations*. Elsevier Science.
- Gerfen, C. R. (1992). The neostriatal mosaic: Multiple levels of compartmental organization. *Trends in Neuroscience*, 15(4), 133-139.
- Gerfen, C. R., & Young, W. S. (1988). Distribution of striatonigral and striato pallidal peptidergic neurons in both patch and matrix compartments: An in situ hybridization histochemistry and fluorescent retrograde tracing study. *Brain Research*, 460, 161-167.
- Graybiel, A. M., & Kimura, M. (1995). Adaptive neural networks in the basal ganglia. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 103-116). Cambridge: MIT Press.
- Houk, J. C. (1995). Information processing in modular circuits linking basal ganglia and cerebral cortex. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 3-10). Cambridge: MIT Press.
- Houk, J. C. (1997). On the role of the cerebellum and the basal ganglia in cognitive signal processing. In C. I. de Zeeuw, P. Strata, & J. Voogd (Eds.), *Progress in brain research* (Vol. 114, p. 543-552). Elsevier Science.
- Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 249-270). Cambridge: MIT Press.
- Houk, J. C., & Wise, S. P. (1995). Distributed modular architecture linking basal ganglia, cerebellum and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex*, 5, 95-111.

- Jaeger, D., Kita, H., & Wilson, C. J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *Journal of Neurophysiology*, *72*(5), 2555-2558.
- Joel, D., & Wiener, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience*, *63*, 363-379.
- Joel, D., & Wiener, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: An analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience*, *96*(3), 451-474.
- Kawaguchi, Y. (1997). Neostriatal cell subtypes and their functional roles. *Neuroscience Research*, *27*, 1-8.
- Kincaid, A. E., Zheng, T., & Wilson, C. J. (1998). Connectivity and convergence of single corticostriatal axons. *The Journal of Neuroscience*, *18*(12), 4722-4731.
- Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopaminergic neurons during learning of behavioral reactions. *The Journal of Neurophysiology*, *67*, 145-163.
- Marsden, C. D. (1986). Movement disorders and the basal ganglia. *Trends in Neuroscience*, 512-515.
- Miller, R., & Wickens, J. R. (1991). Corticostriatal cell assemblies in selective attention and in representation of predictable and controllable events. *Concepts in Neuroscience*, *2*(1).
- Montague, P. R., Dayan, P., Person, C., & Sejnowski, T. J. (1995). Bee foraging in uncertain environments using predictive hebbian learning. *Nature*, *377*, 725-728.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *The Journal of Neuroscience*, *16*, 1936-1947.

- Parent, A., & Hazrati, L. N. (1993). Anatomical aspects of information processing in primate basal ganglia. *Trends in Neurosciences*, *16*(3), 111-116.
- Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey mid-brain: Contingencies of responses to active touch during self-initiated arm movements. *The Journal of Neurophysiology*, *63*, 592-606.
- Schultz, W. (1997). Adaptive dopaminergic neurons report the appetitive value of environmental stimuli. In J. W. Donahoe & V. Packard-Dorsel (Eds.), *Neural network models of cognition: Biobehavioral foundations* (p. 317-335). Elsevier Science.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1-27.
- Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience*, *13*, 900-913.
- Schultz, W., Apicella, P., Scarnati, P., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of Neuroscience*, *12*, 4595-4610.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.
- Schultz, W., Romo, R., Ljungberg, T., Mirenowicz, J., Hollerman, J. R., & Dickinson, A. (1995). Reward-related signals carried by dopamine neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 233-248). MIT Press.
- Sutton, R. S. (1988). Learning to predict by the method of temporal difference. *Machine Learning*, *3*, 9-44.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.

- Suzuki, T., Miura, M., Mishimura, K., & Aosaki, T. (2001). Dopamine-dependent synaptic plasticity in the striatal cholinergic interneurons. *The Journal of Neuroscience*, *21*(17), 6492-6501.
- Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*, 43-48.
- Wickens, J. (1990). Striatal dopamine in motor activation and reward-mediated learning: Steps towards a unifying model. *Journal of Neural Transmission*, *80*, 9-31.
- Wickens, J., & Kötter, R. (1995). Cellular models of reinforcement. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 187-214). MIT Press.
- Wilson, C. J. (1995). The contribution of cortical neurons to the firing patterns of striatal spiny neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 29-50). Cambridge: MIT Press.
- Wilson, C. J., Kita, H., & Kawaguchi, Y. (1989). GABAergic interneurons, rather than spiny cell axon collaterals, are responsible for the IPSP responses to afferent stimulation in neostriatal spiny neurons. *Soc. Neurosci. Abstr.*, *15*, 907.
- Wise, R. A., Spindler, J., & Legault, L. (1978). Major attenuation of food reward with performance-sparing doses of pimozide in the rat. *Canadian Journal of Psychology*, *32*, 77-85.