Biol. Cybern. 84, 401–410 (2001)

Biological Cybernetics

A computational model of action selection in the basal ganglia. I. A new functional anatomy

K. Gurney, T. J. Prescott, P. Redgrave

Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK

Received: 16 February 2000 / Accepted in revised form: 30 October 2000

Abstract. We present a biologically plausible model of processing intrinsic to the basal ganglia based on the computational premise that action selection is a primary role of these central brain structures. By encoding the propensity for selecting a given action in a scalar value (the salience), it is shown that action selection may be recast in terms of signal selection. The generic properties of signal selection are defined and neural networks for this type of computation examined. A comparison between these networks and basal ganglia anatomy leads to a novel functional decomposition of the basal ganglia architecture into 'selection' and 'control' pathways. The former pathway performs the selection per se via a feedforward off-centre on-surround network. The control pathway regulates the action of the selection pathway to ensure its effective operation, and synergistically complements its dopaminergic modulation. The model contrasts with the prevailing functional segregation of basal ganglia into 'direct' and 'indirect' pathways.

1 Introduction

An important task for the vertebrate nervous system is the resolution of conflicts between functional units that are physically separated within the brain but are in competition for behavioural expression. Stated informally, it is the problem of how we decide 'what to do next'. This situation is particularly acute if several such units are competing for a common resource as, for example, with the neural systems involved in feeding and drinking, both of which require the use of oral muscles. In other cases, multiple simultaneous actions may be permitted as, for example, when an animal walks and chews at the same time. Competition can also arise in modalities where behavioural expression is indirect, for instance between systems attempting to gain control of cognitive resources. In general, therefore, we suppose that the brain is processing a large number of sensory and cognitive streams or channels, each of which may be requesting some action to be taken. The task demanded of the animal, in order for appropriate behaviour to occur, is to suppress the majority of these requests while allowing the expression of only a limited number (in some cases just one). This problem of *action selection* is clearly crucial to our understanding human behaviour (Duncan 1995) but also arises within ethological analysis of animal behaviour (McFarland 1989) (where it is often referred to as the problem of 'behavioural switching' or 'decision-making') and is a central concern in the control of behaviour-based robots (Arkin 1995). We have recently argued that there are many common features in the analysis of the action-selection problem, irrespective of the disciplinary context (Prescott et al. 1999). However, from the biological perspective we can ask the following questions: (1) what is the neural substrate performing action selection? and (2) how might this substrate operate functionally?

In order to identify a neural substrate for action selection, it is helpful to be aware of the general computational issue of how actions – requested by spatially separated functional systems – might be selected in principle. From the architectural perspective there are two major possibilities (Snaith and Holland 1990). First, competing modules could work directly with each other and resolve any conflicts by inter-module processing. Alternatively, it is possible that competitors communicate requests to a central arbitrating mechanism which then selects, from these, a subset for behavioural expression. Recently we have reviewed these possibilities (Redgrave et al. 1999) and argued that, in terms of connectivity and metabolic efficiency, effective action selection favours centralised switching devices. In addition we proposed that in the vertebrate brain the basal ganglia (a group of functionally related, central brain structures) are well suited to play this role. The basal ganglia have been implicated in an extensive range of processes including perception and cognition (Brown

Correspondence to: K. Gurney
(e-mail: k.gurney@shef.ac.uk)

et al. 1997), working memory (Levy et al. 1997), and many aspects of motor function (Marsden and Obeso 1994; Graybiel 1995). However, one recurring theme in the literature (see for example Cools 1980; Mink 1996; Kropotov and Etlinger 1999), is that the basal ganglia are implicated in the problem of action selection. Our recent work (Prescott et al. 1999; Redgrave et al. 1999) has developed the idea of selection as a major unifying hypothesis of basal ganglia function, showing how it relates to known anatomy and physiology, and how it meets several high-level computational requirements.

In this paper and in Gurney et al. (2001), henceforth referred to as GPR2, we address the second question given above: how might the neurobiological substrate for action selection operate? Our account takes the form of a biologically inspired model of processing intrinsic to the basal ganglia. In developing the model we had four main aims. First, to articulate quantitatively the notion of 'selection' and show how it may be applied in the context of basal ganglia modelling. Second, to reinterpret the basal ganglia anatomy in terms of a set of neural mechanisms specialised for selection. Third, to give a quantitative explanation of the functional contribution to selection made by different basal ganglia nuclei. Fourth, given the importance of dopamine in regulating basal ganglia function (see for example Mink 1996), to begin an analysis of the role of dopamine in the context of selection.

This paper deals primarily with the first two of these aims; the latter two are considered in GPR2. Quantitative analysis allows contact to be made with experimental data and is used to bolster our claim for a new functional architecture of the basal ganglia. However, the line of reasoning leading to this functional scheme is not contingent on detailed mechanistic specifics, and hence it stands on its own merits.

After completing a review of basal ganglia anatomy physiology and function, Sect. 2 shows how a model of processing intrinsic to the basal ganglia may be couched in terms of the more tractable problem of signal selection, and goes on to formalise this process. Next, we proceed to explore the functional architecture of the basal ganglia using a methodology grounded in the following observation: given that the selective functions of basal ganglia are instantiated in neural circuits, the mechanisms they utilise must belong to the repertoire available to such systems. Conversely, an examination of this set of mechanisms will help direct an interpretation of the possible functional configurations of the basal ganglia anatomy. Section 3 therefore looks at general neural network mechanisms for selection, and Sect. 4 shows how these may be mapped onto the basal ganglia anatomy. The result is a reinterpretation of the functional anatomy of the basal ganglia into two new pathways: one devoted nominally to the selection process per se, and another given over to regulating or modulating the operation of this primary selection pathway.

1.1 Basal ganglia anatomy and physiology

The anatomical and physiological data which form the basis of our model are well known, and are described in several recent reviews (Gerfen and Wilson 1996; Mink 1996; Wickens 1997; Redgrave et al. 1999). The primate basal ganglia sub-nuclei and their intrinsic connections are shown in schematic form in Fig. 1a. The main components of the primate basal ganglia are the striatum, the globus pallidus (GP) and the subthalamic nucleus (STN) in the forebrain, and the substantia nigra (SN) in the midbrain. The globus pallidus contains two sub-divisions – the internal and external segments (GPi and GPe, respectively) – while the substantia nigra contains distinct areas designated compacta (SNc) and reticulata (SNr). Homologous structures (though often with different names) are found in the nervous systems of other vertebrate classes (Medina and Reiner 1995).

The major source of excitation within the basal ganglia is the STN, while connections between most other nuclei are inhibitory. The SNc provides dopaminergic input to the striatum and, depending on the receptor type of the post-synaptic cells, may exert an inhibitory or excitatory effect. Thus, dopaminergic transmission is primarily excitatory when mediated by D1-type receptors (Akkal et al. 1996) but it can exert an inhibitory effect when activating D2 receptors (Gerfen et al. 1990; Harsing and Zigmond 1997) The overwhelming pro-

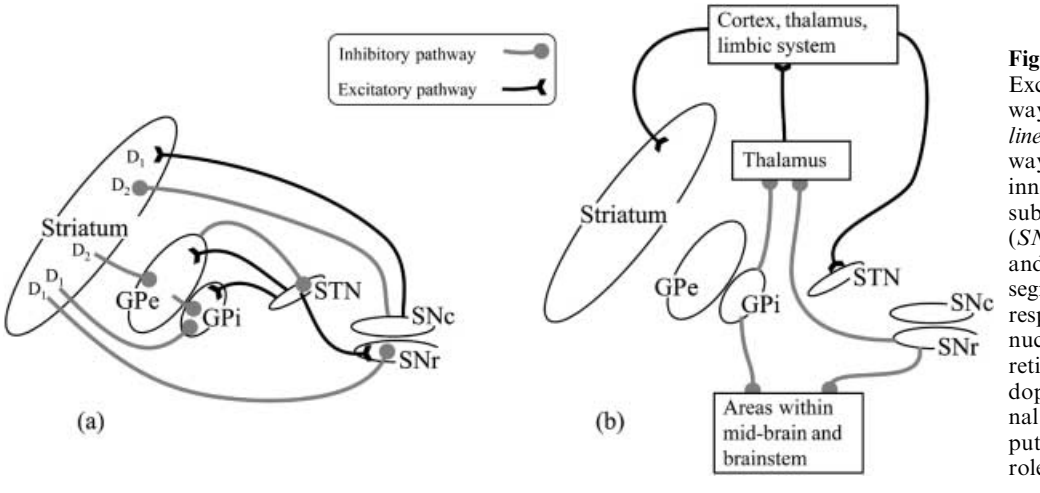

Fig. 1a,b. Basal ganglia anatomy. Excitatory and inhibitory pathways are denoted by *solid* and *grey lines*, respectively: **a** internal pathways, including dopaminergic innervation of striatum from substantia nigra pars compacta (*SNc*). Key to other nuclei: *GPe* and *GPi*, external and internal segments of the globus pallidus, respectively; *STN*, subthalamic nucleus; *SNr*, substantia nigra pars reticulata. *D*1 and *D*2 refer to dopamine receptor types; **b** external pathways showing major inputs and outputs and the central role of the thalamus

portion (90%) of neurons in striatum are GABAergic medium spiny cells. These cells receive major excitatory inputs to the basal ganglia from a wide range of brain structures and also dopaminergic input from SNc. They provide phasic inhibitory output to the GPi, GPe and SNr. The differential effects of SNc innervation are propagated through the basal ganglia via striatal efferents. Thus, GPi/SNr receives projections exclusively from medium spiny cells associated with D1-type receptors, while GPe receives projections mainly from cells with D2-type receptors (Gerfen et al. 1990). The projections to GPi/SNr are also associated with collaterals to GPe, but these tend to be much less heavily branched – and fill a smaller volume – than those that originate on cells with D2 receptors (Kawaguchi et al. 1990). We take this as evidence that striatal cells with D1-type receptors provide a much weaker projection to GPe than their D2-type counterparts.

In their default state, medium spiny neurons are largely silent and do not respond to low levels of input. However, on receiving substantial levels of coordinated excitatory input, these cells yield a significant output whose magnitude may be subsequently affected by low-level inputs which are ineffective when presented in isolation. This dichotomous behaviour is described by using the terms 'down state' and 'up state', respectively, for these two modes of operation (Wilson 1995).

The anatomical context of the basal ganglia is shown in Fig. 1b. Its principle input nuclei (striatum and STN) receive afferents from virtually the entire cerebral cortex, from the mid-line and intralaminar nuclei of the thalamus, and from the limbic system (principally the amygdala and hippocampus). The basal ganglia are therefore capable of processing an enormous variety of information, which highlights its possible significance as a central selection mechanism. Moreover, this basal ganglia input occurs via a series of afferent parallel processing streams or channels (Hoover and Strick 1993; Alexander et al. 1986) and, where motor areas are concerned, displays a somatotopic organisation (DeLong et al. 1983; Flaherty and Graybiel 1993). The main basal ganglia output nuclei are the GPi and SNr, which provide extensively branched GABAergic efferents to functionally related zones of the ventral thalamus (which in turn projects back to the cerebral cortex), superior colliculus and other brain stem areas.

1.2 Basal ganglia functional mechanisms

While much is known of the anatomy and physiology of the basal ganglia, comparatively little is known of the functional architecture and its intrinsic computations. Such knowledge is crucial for understanding how the basal ganglia might mediate selective functions. In general terms, the basal ganglia enables or 'gates' actions by the release of inhibition (Chevalier et al. 1985; Deniau and Chevalier 1985). The default function of the basal ganglia output nuclei is to exert a widespread tonic inhibitory control over target structures. The basal ganglia are then able to promote actions via disinhibition of their associated target structures while maintaining inhibitory control over others.

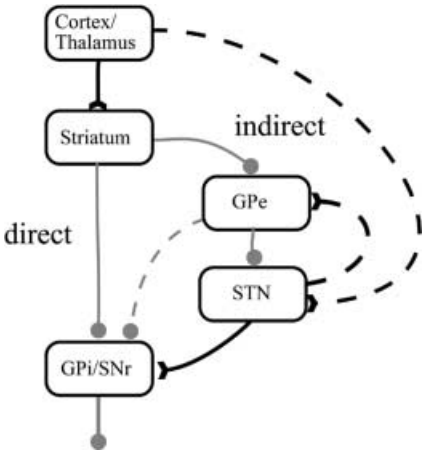

Fig. 2. Prevailing functional interpretation of the basal ganglia. Excitatory and inhibitory pathways are denoted by *solid* and *grey lines*, respectively. In addition, pathways which receive less emphasis in this model are shown as *broken lines*

The prevailing model of the functional architecture intrinsic to the basal ganglia was originally proposed by Albin et al. (1989), and is shown in Fig. 2. Two routes are available for striatal modulation of the output nuclei: (1) a 'direct pathway' in which GABAergic output from striatum is able to directly inhibit elements in GPi and SNr, and (2) an 'indirect pathway' involving a double synaptic route from GPe to STN and thence to GPi/SNr, which is able to exert an overall excitatory influence on the output nuclei. While this model has helped to direct research in the past decade, several shortcomings have emerged. First, this proposal remains to be developed as a full computational model, so that several aspects of its operation remain unclear including, for example, the precise manner of interaction between the direct and indirect pathways (Alexander and Crutcher 1990). Second, there are now several well-established pathways – e.g. the external input to STN, the GPe innervation of GPi/SNr and the STN excitation of GPe (see Fig. 1) – which have yet to be fully integrated into the model, although Smith et al. (1998) have recently acknowledged the importance of these pathways. Third, it has become apparent that the model is unable to accommodate certain functional data; for example, the role of GPe in parkinsonian animals (Chesselet and Delfs 1996; Parent and Cicchetti 1998). These shortcomings have been critically assessed by the authors of the model who look forward to 'the destruction of the model and its resurrection in a more realistic form' (Albin et al. 1995).

Mink and Thach (1993) have proposed an alternative interpretation of a subset of basal ganglia intrinsic circuitry which could, in principle, mediate a selective function. Their organisation emphasises the STN as a major input nucleus and depends on the diffuse nature of STN projections to the output nuclei GPi/SNr (Parent

and Hazrati 1993, 1995). The current study incorporates a modification of this scheme as one of its central components. However, we go on to explore its function quantitatively, suggest a novel role for the GPe, and demonstrate how the dopaminergic modulation of striatum may operate in the context of the selection hypothesis of basal ganglia function.

2 The process of selection

Competing candidate actions are assumed to be represented in networks distributed widely throughout the central nervous system. The overall activity level of the neural representation of a given action may determine its *salience* or propensity to be selected for execution, as proposed by Koechlin and Burnod (1996). Rather than dealing directly with the neural codes for each action, we propose that the basal ganglia works with their (scalar-valued) saliences. Actions are therefore represented at the input stages of the basal ganglia in terms of the 'common currency' of salience (mechanistic details are discussed further in Sect. 4.2). At the basal ganglia output, actions are mediated by the release of inhibition, and the degree to which this takes place defines another scalar-valued signal. Thus, in a model which treats only processing intrinsic to the basal ganglia, we are able to reframe the process of 'action selection' as one of 'signal selection', in which large salience-signal inputs at striatum and STN select for low signal outputs at GPi/SNr. For our purposes, an action is defined as the behavioural expression determined by a specific population of cells within basal ganglia. These populations are what we refer to as 'channels', and may be associated with anything from an elemental motor act through to an extensive behavioural strategy (Redgrave et al. 1999). It remains a subject of further work to determine the repertoire of actions that the basal ganglia mediates, and how the selection of low-level movements is related to that of higher-level behaviours.

2.1 Defining signal selection

Although the reduction to signal selection is a significant conceptual simplification, it is important to be clear as to exactly what this implies. Signal selection may be viewed as an input–output transformation which dichotomises a set of inputs X into output sets S and $\bar{S}$ of 'selected' and 'non-selected' signal, respectively. The criterion for membership of S might be that the transformed signal be greater than some threshold θ and that all transformed signals less than this threshold are in $\bar{S}$. However, it is possible that signals of each class have values which are arbitrarily close to θ, so that selection will potentially be prone to disruption by noise (Horowitz and Hill 1989).

We therefore refine our definition and partition the transformed signal interval into three intervals, determined by two *selection thresholds* θ_1 and θ_2, allowing transformed signals to be in S, $\bar{S}$ or an indeterminate set Y_0, where $y \in Y_0 \Rightarrow \theta_1 < y < \theta_2$. In addition we need to allow for the possibility that selected signals are deemed to be those in either the upper or lower intervals, so that we refer to a *large-signal encoding* if $y \in S \Rightarrow y \geq \theta_2$ and a *small-signal encoding* if $y \in S \Rightarrow y \leq \theta_1$. To proceed further, it is convenient to assign labels to input and output signals so that the selection transformation is defined by a set of mappings $x_i \rightarrow y_i$, where $i = 1 \ldots n$, $n = |X|$, and $x_i \in X$, $y_i \in S \cup Y_0 \cup \bar{S}$. Selection also implies that ordering relations are preserved. Thus, if all inputs and outputs are contained in some interval I, a *selection mechanism* is now defined to be a mapping $G : I^n \rightarrow I^n$ and two thresholds θ_1 and θ_2, where $G(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)$, and which obeys one of the order-preserving relations

$$\begin{array}{lll} x_i \leq x_j \Rightarrow y_i \leq y_j & : & \text{large-signal selection} \\ x_i \leq x_j \Rightarrow y_i \geq y_j & : & \text{small-signal selection} \end{array} \quad (1)$$

where we have specialised to the case in which it is the larger inputs that get mapped into S.

Clearly any order-preserving function G is formally a candidate for a selection mechanism, and it remains to establish criteria that might make it truly useful in a selection context. The *decisiveness* $D(x)$ of the mechanism applied to X is given by

$$D(x) = 1 - \frac{|Y_0|}{|X|} \quad (2)$$

where $0 \leq D(x) \leq 1$. $D(x) = 1$ signifies an extremely decisive or 'clean' selection with respect to X in which all signals are transformed into one of S or $\bar{S}$. $D(x) = 0$ means that no selection has taken place and all signals are in the indeterminate set Y_0. Next we specify what fraction $\phi(x) = |S|/|X|$ of the signals are selected and refer to this as the selection *promiscuity*; alternatively we may talk of the *selectivity* $1 - \phi(x)$. A promiscuous mechanism is one for which $\phi(x)$ is large for all x. If only one input is selected ($\phi(x) = 1/n$) for all x, we say that the mechanism instantiates *hard switching*. If however, $\phi(x) > 1/n$ for some x we say the mechanism allows for *soft switching*.

Consider now a small-signal encoding mechanism. If the quiescent state ($x_i = 0\ \forall i$) satisfies $y_i \in \bar{S}\ \forall i$ then, in order for inputs to become selected, we must have $\mathrm{d}y_i/\mathrm{d}x_i|_x < 0$ for some x. If additionally, $\mathrm{d}y_i/\mathrm{d}x_j|_x > 0, i \neq j$ then we say there is *competition* between i and j. It is clear that competition promotes increased selectivity since previously selected inputs may become de-selected at the expense of selecting new ones.

3 Neural networks for selection

Signal selection in a neural network may be accomplished by associating each node with a spatial pattern of synaptic contacts that consists of two concentric zones with opposite polarity of influence. This scheme may be configured using intra-layer, lateral recurrent links, or as a feedforward net. If the central synaptic zone is excitatory then we obtain an on-centre off-surround

network. These nets have proved popular in modelling self-organizing feature detectors in visual cortex, most notably using their recurrent instantiation (see for example Von der Malsburg 1973) but also in their feedforward form (Grossberg 1976). In a biological context, recurrent nets will necessarily be confined to a particular sub-nucleus. While these may be important in basal ganglia computation, we are primarily interested in interpreting inter-nuclear pathways and so focus, at this stage, on feedforward nets.

To support an on-centre off-surround architecture requires a source of diffuse inhibition in basal ganglia, for which there is no evidence. On the other hand, it is possible that the diffuse excitation provided by STN may contribute to an off-centre on-surround scheme. Figure 3a shows an example of a feedforward net with off-centre on-surround connectivity. Assuming an interpretation of the inputs x_i in terms of normalised rate-coding, $0 \leq x_i \leq 1, \forall i$. Each input node transmits its value x_i to all output neurons but sends inhibition to only one output neuron which is assigned the same index i. We refer to this input and output node combination as a channel, although no interpretation in terms of basal ganglia circuitry is implied at present. The nodes in the output layer are simple semi-linear units whose activation a_i is the weighted sum of the unit's inputs. Thus, if w^+, w^- are the magnitudes of the excitatory and inhibitory weights respectively, then

$$a_i = -w^- x_i + w^+ \sum_{j \neq i}^{n} x_j \tag{3}$$

For two arbitrary channels p and k, it then follows $a_p - a_k = (w^+ + w^-)(x_k - x_p)$. That is, if $x_k \geq x_p$ then $a_k \leq a_p$. Given a monotonic increasing relation between the output y_i and the activation a_i, there is therefore an order-preserving map between inputs and outputs as required for the net to instantiate a small-signal encoding selection mechanism. However, for the net to perform a useful function (possess non-zero decisiveness and promiscuity), the balance between inhibition and excitation must be chosen carefully. To investigate this, put $w^+ = \delta w^-$, and let $\langle x \rangle_{/i}$ be the mean value of the inputs excluding i. The expression for the activation (3) may then be rewritten as

$$a_i = w^- [\delta(n-1)\langle x \rangle_{/i} - x_i] \tag{4}$$

where n is the number of channels. Now suppose δ is fixed, then a_i is unbounded with increasing n unless $\langle x \rangle_{/i}$ is a function of n and $\langle x \rangle_{/i} = O(1/n)$, where $f(n) = O(h(n))$ is used to indicate that $f(n)/h(n)$ is bounded as $n \to \infty$. Thus, given a monotonic squashing relation $y = y(a)$ between the output y and the activation, then unless the inputs are conditioned in this way, all outputs will eventually saturate at (or close to) their maximum value. The condition on the inputs may be relaxed if, instead, $\delta = O(1/n)$. We refer to the input-independent setting of an appropriate ratio between excitation and inhibition as *capacity scaling*. Putting $\delta = O(1/n)$ offers a synaptic mechanism for 'hardwiring' this property into the network.

The selection mechanism description is completed by supplying the two thresholds θ_1 and θ_2, and a specification of $y(a)$. Consider the piecewise linear squashing function

$$y_i = \begin{cases} 0 & : a < \epsilon \\ m(a - \epsilon) & : \epsilon \leq a \leq 1/m + \epsilon \\ 1 & : a > 1/m + \epsilon \end{cases} \tag{5}$$

Since the y_i are normalised, $0 \leq \theta_1 < \theta_2 \leq 1$. Under quiescent conditions, $a_i = 0 \; \forall i$ and, if $\epsilon \geq 0$ then $y_i = 0 \; \forall i$. Since the selection criterion is that $y_i \leq \theta_1$, this means that all channels are selected when there is no input to the network. This is clearly not admissible and so we put $\epsilon < 0$, which gives a resting or 'tonic' output value $y_0 > 0$ and ensures that $0 \leq \theta_1 < \theta_2 \leq y_0$. Finally, we note that there is the potential for inter-channel competition because $\mathrm{d}a_i/\mathrm{d}x_j = w^+ > 0$, $i \neq j$ so that $\mathrm{d}y_i/\mathrm{d}x_j > 0$ if $\epsilon \leq a \leq 1/m + \epsilon$.

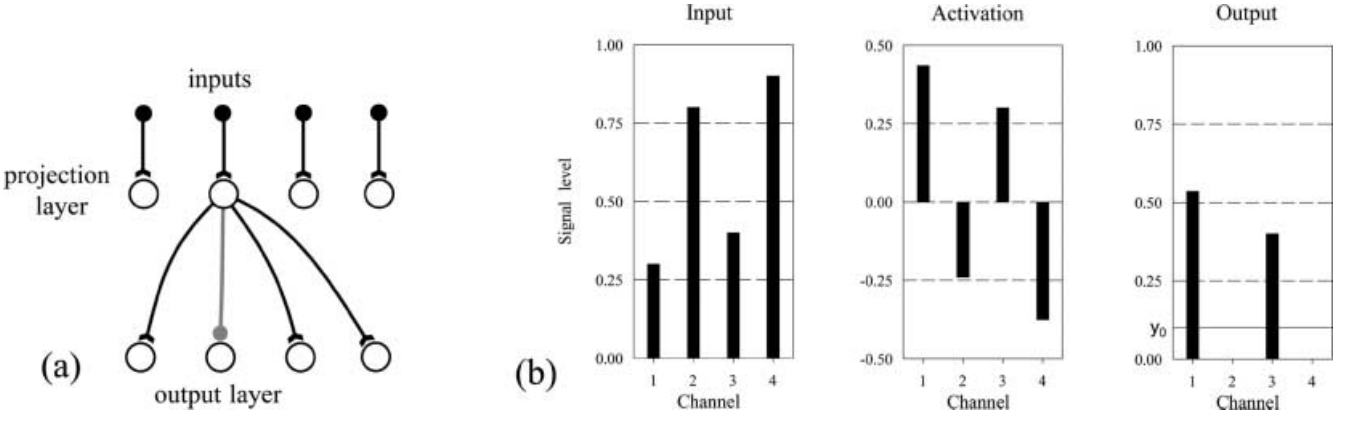

Fig. 3a,b. Off-centre on-surround neural network: **a** architecture of such a net with four channels. Excitatory and inhibitory links are denoted by *solid* and *grey lines*, respectively; **b** simulation of the network in (**a**) demonstrating the silencing of outputs on channels 2 and 4 which have the largest inputs. The excitatory and inhibitory weights were 0.45 and −1.35, respectively

Figure 3b shows simulation results of a 4-channel network of the type discussed above, with $w^- = 1.35, w^+ = 0.35, \epsilon = -0.1$ and $m = 1$. The channels with the smaller inputs (channels 1 and 3) have outputs greater than tonic, while the channels with the largest inputs (channels 2 and 4) have suppressed their outputs to zero. This has resulted in the selection of channels 2 and 4 with the other two channels being decisively not-selected. The possible selection of more than one channel means that the feedforward nets may implement soft switching. In summary, a small-signal encoding selection mechanism may be implemented in a feedforward, off-centre on-surround network as long as it possesses capacity scaling and non-zero tonic output.

4 Model development

4.1 Underlying assumptions

In the absence of clearly established principles of basal ganglia function, it is appropriate in the first instance to adopt a system-level approach (Churchland and Sejnowski 1994). In this 'coarse-grained' or 'low-magnification' view we are interested in the ensemble behaviour of cells and the way in which entire sub-populations within each nucleus might work with each other. Further, we believe that it is helpful to understand the mechanisms at work within the basal ganglia prior to examining models which deal with the wider anatomical context that includes afferent and efferent structures (such as cortex, thalamus and superior colliculus).

Our general functional hypothesis is that an actively competing channel will excite a population of cells within striatum leading to inhibition of a corresponding population in the output nuclei (GPi/SNr). This suppresses the tonic inhibitory control exerted by these cells over their target structures, thereby allowing them to become active and to express the action they are associated with (Chevalier et al. 1985; Deniau and Chevalier 1985).

4.2 Reinterpreting functional anatomy

Consider once again the network of Fig. 3. This may be interpreted as an action selection mechanism if we associate its inputs with the salience values on a set of discrete-action processing channels. The assumption that channels are physically distinct is consistent with the possible division of basal ganglia into disjoint, parallel-processing streams (Alexander et al. 1986; Hoover and Strick 1993) and, where appropriate, a somatotopic organisation (DeLong et al. 1983; Flaherty and Graybiel 1993). Clearly, this assumption may break down under certain circumstances, but we contend that it is good approximation for many choices of channel set. The populations of neurons in each basal ganglia sub-nucleus associated with a single channel are treated as being identical, so that we may work with a single representative in each case. This is a standard approach for investigating at the systems level and has been adopted, for example, in modelling cortex (Douglas et al. 1989).

While the network in Fig. 3 possesses the ability, in principle, to perform action selection, it is not possible to identify the net with any part of the basal ganglia circuitry because it is not biologically plausible. Thus, the projection layer innervates with both inhibition and excitation, and the diffuse excitatory zone has a highly localised discontinuity. A more realistic implementation of the architecture which overcomes these objections is shown in Fig. 4a. This implementation has separate excitatory and inhibitory input structures and uniformly distributed excitatory efferents. With suitable weights, this network can be made to behave identically to its simpler counterpart in Fig. 3. Further, the input structures have been identified with STN and striatum.

Turning now to the use of salience as input, we propose that one purpose of the medium spiny neurons is to extract the salience of action requests on their associated channels. This may take place within the extensive dendritic contacts of these cells, which have been implicated in processing patterns of widely distributed input activity and is evident, for example, in the

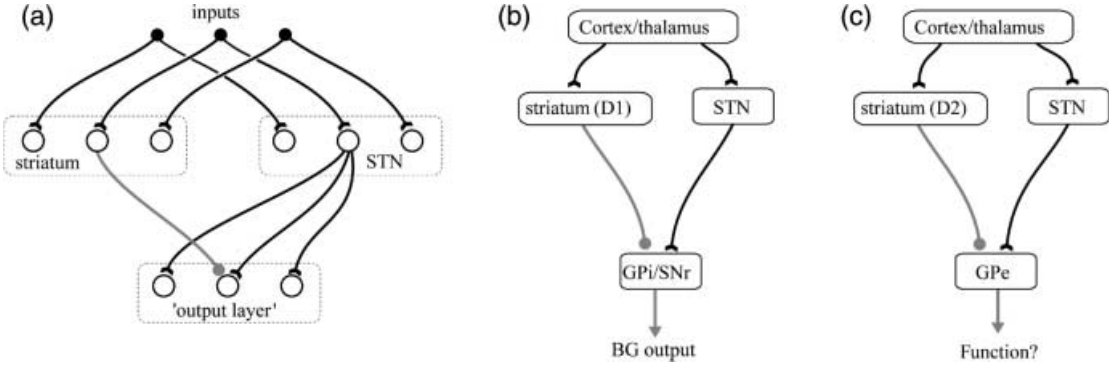

Fig. 4a–c. The component parts of the new functional architecture. Mapping of the network in Fig. 3 onto the basal ganglia anatomy: **a** separate input nuclei for excitation and inhibition but, as yet, unspecified output nucleus; **b** one specific instantiation of the system in (**a**) – the 'selection pathway' which includes the basal ganglia output nuclei; **c** the other instantiation of (**a**) constitutes a 'control pathway' and its 'output', the *GPe*, which supplies control signals to other basal ganglia nuclei

convergent pattern of cortico-striatal connections involved in motor processing (Flaherty and Graybiel 1991; Graybiel et al. 1994). The details of this mechanism reside at the subcellular level and will also involve an understanding of the neural encoding of actions in command structures external to the basal ganglia. Given that we are working at the system level and dealing with processing within basal ganglia, we assume that salience has been extracted on the dendritic arbor and that the total input to a model striatal neuron is simply the salience of the action it is currently processing.

The STN may lack the sophisticated apparatus of the medium spiny cells to extract salience, but this is not critical to the overall putative role of STN in the model. We suppose that the STN sends diffuse innervation to its target nuclei (Parent and Hazrati 1993, 1995) so that each population in these nuclei receives the sum total of STN output. Thus, individual channel contributions do not have to be maintained at the STN targets and it is sufficient for STN to supply excitation which is proportional to the total activity afferent to basal ganglia. While widespread axon-collateral branching to targets is one possible mechanism for the redistribution of excitation (and is the one we adopt here), it may also be instantiated using the widespread connectivity within STN (Kita et al. 1983; Afsharpour 1985) such that multiple efferents may be excited by a single, focused stimulus. This is consistent with the observation that there is widespread excitatory influence generated in STN following stimulation of a single site in sensorimotor cortex (Fujimoto and Kita 1992), a process which has been modelled by Gillies and Willshaw (1998). Further, the architecture we are proposing is not dependent on there being a single projective field (PF) of STN equivalent to the whole of GPi/SNr. The output nuclei may be served by multiple PFs of STN instantiating multiple copies of the architecture; the main prerequisite is that PFs of STN channels are more diffuse than their counterparts in striatum.

Finally, our model assumes that both striatum and STN are innervated by branched collaterals from a common input. While there is evidence that this does occur (Feger et al. 1991) it may not be universal. A weaker but functionally equivalent assumption is that striatum and STN receive highly correlated inputs or 'copies' of the same signals.

While the general scheme in Fig. 4 is now plausible, the key to our reinterpretation of the functional anatomy is the realisation that there are *two* instantiations of this architecture within basal ganglia. One of these is similar to that suggested by Mink and Thach (1993) and makes use of GPi/SNr as its 'output layer'. There is another, however, which makes use of GPe as its output. This distinction has a physiological basis in the differentiation of projection targets of medium spiny cells according to their mode of dopaminergic modulation (Fig. 1a). Thus, GPi/SNr is innervated principally by striatal cells associated with D1-type receptors, while GPe receives projections mainly from cells with D2-type receptors. The two resulting 'pathways' are shown in Fig. 4b and c, respectively. Since GPi/SNr provides the basal ganglia output, it is reasonable to suppose that the system shown in Fig. 4b instantiates the selection mechanism per se; we therefore designate it the *selection pathway*.

For the other pathway (Fig. 4c) it is not immediately clear in what sense the GPe is an 'output' nucleus, since its efferents are confined to other basal ganglia nuclei. The problem can be resolved by supposing that this subsystem forms a *control pathway* whose function is to regulate the properties of the main selection mechanism via signals supplied by GPe. The general view that GPe exerts a control influence within the basal ganglia has been alluded to previously by Parent and Hazrati (1995).

To help discover the precise role of the control pathway, we can make the following qualitative observations. First, the inhibition of STN by GPe should lead to a reduction in the excitation of GPi/SNr in the selection pathway. In GPR2 we demonstrate that this inhibitory control is exactly matched to the requirement of capacity scaling in the sense described in Sect. 3. Second, the dopaminergic modulation in the two pathways acts synergistically. In accordance with the discussion in Sect. 1.1, we assume that D1 and D2 receptors are associated with excitatory and inhibitory dopaminergic modulation, respectively. Raising the level of dopamine will then increase the strength of striatal inhibition to GPi/SNr in the selection pathway and decrease striatal inhibition of GPe in the control pathway. This, in turn, results in decreased GPi/SNr output, both directly via the GPe→GPi/SNr pathway and indirectly via reduction of STN excitation. The net effect is that both pathways act to inhibit the basal ganglia output nuclei.

Combining the two sub-systems, and incorporating the GPe efferents that supply control signals, results in the functional architecture that forms the basis of the new model (Fig. 5). The new architecture should be compared with the 'direct/indirect' pathway interpretation (Albin et al. 1989) shown in Fig. 2. The two schemes are structurally quite different. The old 'direct' pathway consists of the whole of striatum and GPi/SNr. In contrast, the new selection pathway incorporates STN as an input nucleus, and its striatal contribution consists of the medium spiny cell population which

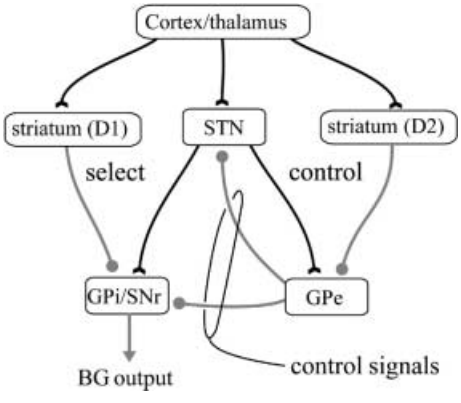

Fig. 5. The final form of the new functional architecture. The full model showing selection and control pathways combined

employs D1 dopamine receptors. The old 'indirect' pathway consists of striatum, GPe STN and GPi/SNr. The new control pathway does not contain GPi/SNr, and its striatal contribution consists of the medium spiny cell population which employs D2 dopamine receptors. Further, the proposed functional roles of the two new pathways (selection and control) are quite different from those of the old model.

4.2.1 Local striatal networks. From the functional viewpoint, striatum has been shown to support local processing mediated by inhibition (Brown and Sharp 1995), and although the anatomical basis for this is not clear (Jaeger et al. 1994), several models of striatal function (Wickens et al. 1991; Alexander and Wickens 1993; Houk et al. 1995) assume that striatum supports this style of processing. The operation of the new functional architecture is not contingent on the existence of intra-striatal inhibition but, nevertheless, the model is able to accommodate it provided the following two conditions hold. First, that a *graded* measure of salience is extracted by the maximally responsive neurons in each local network. In GPR2 we show that this holds for lateral recurrent nets that make use of inhibition only. It is not so for so-called winner-take-all (WTA) nets popular in engineering neural networks, since these nets result in all locally maximal saliences forcing the maximum output. WTA nets would appear to be less likely in striatum since they require an excitatory neighbourhood which is not consistent with the GABA inhibition expressed by medium spiny neurons. The second condition is that both striatal subsystems (defined by their dopamine receptor type) contain similar representations of highly salient inputs. While the hypothesis of local recurrent nets within each subsystem is consistent with evidence that striatal cells associated with the same receptor type innervate each other, the D1 and D2 subsystems may also interact (Smith et al. 1998). However, we suppose that this coupling does not prohibit the representation of highly salient inputs equally faithfully in both systems. This could occur because the coupling is weak, or because it is configured for cooperative processing that facilitates an equivalent representation of salience in each population. We do not consider it plausible that the maximally salient input in a local recurrent net is associated with cells of a single dopamine receptor type chosen at random from either population.

4.2.2 Hard and soft switching combined. Computationally, local striatal processing may perform a vital role. Under the somatotopic organisation of striatum, each recurrent network may process action requests associated with the same somatic area. Then, given the ability of the recurrent nets to select only a single channel (thereby performing hard switching; see Sect. 2.1) each network is ideally configured to resolve resource conflicts between motor programmes that are seeking to gain control of a single motor resource. However, channels which are selected as a result of local recurrent competition may require further selection amongst themselves. If this did not take place, undesirable and inappropriate action combinations may occur as observed, for example, in attention deficit hyperactivity disorder or Tourette's syndrome, both of which have been associated with disorders of basal ganglia function (Brito 1997; Swanson et al. 1998). On the other hand, the possibility that more than one channel be selected is a requirement of the observation that it is possible to engage in multiple, clearly separable behaviours (walking and chewing for example). These requirements may be met by a subsequent stage of soft switching, the implementation of which, we propose, is to be found in the selection pathway of the new functional architecture.

5 Discussion

Our initial computational premise is that one of the primary roles of the basal ganglia is to mediate action selection. We then argued that, given a common input representation in terms of salience, the selective function *intrinsic* to basal ganglia may be viewed as one of signal selection. This process was quantified as an order-preserving mapping which tended to separate signals into two sets separated by an interval $(\theta_2 - \theta_1)$. Given that the basal ganglia consist of neural circuits performing signal selection, we then explored the neural architectures that might be able to perform this kind of computation. This led to a natural interpretation of basal ganglia functional anatomy in terms of two pathways. First, a selection pathway, which contains the minimal neural infrastructure for small signal encoded selection: an off-centre on-surround feedforward network. Second, a control pathway which modulates processing within the selection pathway. Qualitative arguments led to possible specific functions of the control pathway: limiting STN excitation to foster capacity scaling, and enhanced dopaminergic modulation of selection. The new functional architecture focuses on the inter-nucleus pathways and the selection process performed by their mutual cooperation. The simulation in Fig. 3b suggests that this takes the form of soft switching. However, this may be complemented with hard switching performed by local striatal circuits which may be accommodated in the model. We propose that the basal ganglia makes use of both types of selection mechanism: hard selection to mediate conflict between channels vying for the same motor resource, and soft selection to determine appropriate combinations of motor activity.

5.1 Model limitations

Our model is inspired by biological information, both at the anatomical level (in terms of connectivity) and at the physiological level (in terms of the excitatory and inhibitory nature of each pathway). While we have endeavoured to include the major known pathways, there is evidence for a weak excitatory path from STN to striatum (Kita and Kitai 1987) and an inhibitory path

from GPe to striatum (Staines et al. 1981). While neither of these 'minor' pathways are included in other models, which emphasise the direct and indirect pathways, the existence of striatal innervation by GPe bolsters our claim for GPe having a major control function within the basal ganglia. The comparatively weak projection from D1-type medium spiny cells to GPe (Kawaguchi et al. 1990) has also been omitted. Its existence does not compromise the selection/control pathway distinction, but may serve to finesse the main functional abilities of the control pathway. Finally, while it is consistent with our systems-level approach, no effort has been made to incorporate a role for interneurons within each nucleus as are known to exist, for example, within the striatum (Kawaguchi 1997).

5.2 Comparison with other models

In comparison with some other brain areas (e.g. visual cortex), the basal ganglia have received rather scant attention from the modelling community (for a review of recent models see Beiser et al. 1997). Perhaps one impediment has been the lack of general agreement about the computational functions performed by the basal ganglia. In this respect, our recent analysis of the role of the basal ganglia in action selection (Prescott et al. 1999; Redgrave et al. 1999) was crucial in orienting the current modelling effort. One recent computational model of the basal ganglia (Berns and Sejnowski 1996) does address the problem of action selection and also makes use of widespread STN excitation. However, the Berns and Sejnowski model differs from the present one in that the GPe is not a control nucleus, but rather plays a central role in the selection process by providing a temporal delay in the 'indirect pathway'. Our model also contrasts with previous efforts which have focused on the intrinsic processing within specific sub-nuclei of the basal ganglia. The striatum in particular has attracted much attention in this respect (Wickens et al. 1991; Wilson 1995) and work has also been done on the STN (Gillies 1996; Gillies and Willshaw 1998). Clearly, these models will provide useful insights as we seek to incorporate further levels of detail into our own model.

5.3 The need for a quantitative model

The new architecture was inspired by the feedforward nets studied in Sect. 3. The heart of the proposed selection mechanism – the selection pathway – is such a net, but it remains to be shown that the entire model architecture satisfies the requirements for a selection mechanism with appropriate selectivity. These requirements include the instantiation of an order-preserving mapping and capacity scaling (the need for a tonic output level is trivially satisfied by GPi/SNr). It is not clear a priori that the architecture implements an order-preserving mapping under the influence of the control pathway (general arguments based on the monotonicity of the neuron outputs may be invoked but these will always remain less than certain). GPR2 therefore conducts a quantitative analysis showing that the model does indeed instantiate an order-preserving map and that – far from hindering selection – one of the functions of the control pathway is to provide automatic capacity scaling that does not rely on tuning the inhibitory/excitatory synaptic-weight ratio. The quantitative model also confirms the synergistic dopaminergic modulation of selection and, moreover, allows comparison between model behaviour and physiological and animal behavioural data.

References

Afsharpour S (1985) Light microscopic analysis of Golgi-impregnated rat subthalamic neurons. J Comp Neurol 236: 1–13

Akkal D, Burbaud P, Audin J, Bioulac B (1996) Responses of substantia nigra pars reticulata neurons to intrastriatal D1 and D2 dopaminergic agonist injections in the rat. Neurosci Lett 213: 66–70

Albin R, Young A, Penney J (1989) The functional anatomy of basal ganglia disorders. Trends Neurosci 12: 366–375

Albin R, Young A, Penney J (1995) The functional anatomy of disorders of the basal ganglia. Trends Neurosci 18: 63–64

Alexander G, Crutcher M (1990) Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci 13: 266–272

Alexander M, Wickens J (1993) Analysis of striatal dynamics: the existence of two modes of behaviour. J Theor Biol 163: 413–438

Alexander G, DeLong M, Strick P (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9: 357–381

Arkin R (1995) Reactive robotic systems. In: Arbib M (ed) The handbook of brain theory and neural networks. MIT Press, Cambridge, Mass., pp 793–796

Beiser D, Hua S, Houk J (1997) Network models of the basal ganglia. Curr Opin Neurobiol 7: 185–190

Berns G, Sejnowski T (1996) How the basal ganglia make decisions. In: Damasio A, Damasio H, Christen H (eds) Neurobiology of decision-making. Springer, Berlin Heidelberg New York

Brito G (1997) A neurobiological model for Tourette syndrome centered on the nucleus accumbens. Med Hypotheses 49: 133–142

Brown L, Sharp F (1995) Metabolic mapping of rat striatum: somatotopic organization of sensorimotor activity. Brain Res 686: 207–222

Brown L, Schneider J, Lidsky T (1997) Sensory and cognitive functions of the basal ganglia. Curr Opin Neurobiol 7: 157–163

Chesselet M, Delfs J (1996) Basal ganglia and movement disorders: an update. Trends Neurosci 19: 417–422

Chevalier G, Vacher S, Deniau J, Desban M (1985) Disinhibition as a basic process in the expression of striatal function. I. The striato-nigral influence on tecto-spinal/tecto-diencephalic neurons. Brain Res 334: 215–226

Churchland P, Sejnowski T (1994) Levels in nervous systems. In: The Computational Brain. MIT Press, Cambridge, Mass., pp 18–48

Cools A (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. Behav Brain Res 1: 361–378

DeLong M, Crutcher MD, Georgopoulos A (1983) Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. J Neurosci 3: 1599–1606

Deniau J, Chevalier G (1985) Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. Brain Res 334: 227–233

Douglas R, Martin K, Whitteridge D (1989) A canonical microcircuit for neo-cortex. Neural Comput 1: 480–488
Duncan J (1995) Attention, intelligence, and the frontal lobes. In: Gazzaniga M (ed) The cognitive neurosciences. MIT Press, Cambridge, Mass., pp 721–734
Feger J, Robeldo P, Renwart N (1991) The subthalamic nucleus: new data, new questions. In: Bernardi G, Carpenter M, Chiara G, Morelli M, Stanzione P (eds) The basal ganglia III. Plenum, New York
Flaherty A, Graybiel A (1991) Corticostriatal transformation in the primate somatosensory system: projections from physiologically mapped body-part representations. J Neurophysiol 66: 1249–1263
Flaherty A, Graybiel A (1993) Two input systems for body representation in the primate striatal matrix: experimental evidence in the squirrel monkey. J Neurosci 13: 1120–1137
Fujimoto K, Kita H (1992) Responses of rat substantia-nigra pars-reticulata units to cortical stimulation. Neurosci Lett 142: 105–109
Gerfen C, Wilson C (1996) The basal ganglia. In: Swanson L, Bjorklund A, Hokfelt T (eds) Handbook of chemical neuroanatomy. (vol 12) Elsevier, Amsterdam, pp 371–468
Gerfen C, Engber T, Mahan L, Susel Z, Chase T, Monsma F, Sibley D (1990) D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. Science 250: 1429–1432
Gillies A (1996) Electrical properties of subthalamic nucleus projection neurons. In: Bower J (ed) Computational neuroscience. Academic, New York, pp 65–70
Gillies A, Willshaw D (1998) A massively connected subthalamic nucleus leads to the generation of widespread pulses. Proc R Soc London B Biol Sci 265: 2101–2109
Graybiel A (1995) Building action repertoires: memory and learning functions of the basal ganglia. Curr Opin Neurobiol 5: 733–741
Graybiel A, Aosaki T, Flaherty A, Kimura M (1994) The basal ganglia and adaptive motor control. Science 265: 1826–1831
Grossberg S (1976) Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. Biol Cybern 23: 121–134
Gurney K, Prescott TJ, Redgrave P (2001) A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biol Cybern 84: 411–423
Harsing LG Jr, Zigmond MJ (1997) Influence of dopamine on GABA release in striatum: evidence for D1-D2 interactions and non-synaptic influences. Neurosci 77: 419–429
Hoover J, Strick P (1993) Multiple output channels in the basal ganglia. Science 259: 819–821
Horowitz P, Hill W (1989) The art of electronics, 2nd edn. Cambridge University Press, Cambridge, Sect. 1.10
Houk J, Adams J, Barto A (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk J, Davis J, Beiser D (eds) Models of information processing in the basal ganglia. MIT Press, Cambridge, Mass., pp 249–270
Jaeger D, Kita H, Wilson C (1994) Surround inhibition among projection neurones is weak or nonexistent in the rat neostriatum. J Neurophysiol 72: 2555–2558
Kawaguchi Y (1997) Neostriatal cell subtypes and their functional roles. Neurosci Res 27: 1–8
Kawaguchi Y, Wilson J, Emson P (1990) Projection subtypes of rat neostriatal matrix cells revealed by intracellular injection of biocytin. J Neurosci 10: 3421–3436
Kita H, Kitai S (1987) Efferent projections of the subthalamic nucleus in the rat: light and electron microscopic analysis with the PHA-L method. J Comp Neurol 260: 435–452
Kita H, Chang H, Kitai S (1983) The morphology of intracellularly labelled rat subthalamic nucleus: a light microscopic analysis. J Comp Neurol 215: 245–257
Koechlin E, Burnod Y (1996) Dual population coding in the neocortex: a model of interaction between representation and attention in the visual cortex. J Cogn Neurosci 8: 353–370
Kropotov J, Etlinger S (1999) Selection of actions in the basal ganglia thalamocortical circuits: review and model. Int J Psychophysiol 31: 197–217
Levy R, Friedman H, Davachi L, Goldman-Rakic P (1997) Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. J Neurosci 17: 3870–3882
Marsden C, Obeso J (1994) The function of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. Brain 117: 877–897
McFarland D (1989) Problems of animal behaviour. Longman Scientific and Technical, London
Medina L, Reiner A (1995) Neurotransmitter organization and connectivity of the basal ganglia in vertebrates: implications for evolution of basal ganglia. Behav Brain Evol 46: 235–258
Mink J (1996) The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol 50: 381–425
Mink J, Thach W (1993) Basal ganglia intrinsic circuits and their role in behaviour. Curr Opin Neurobiol 3: 950–957
Parent A, Hazrati L (1993) Anatomical aspects of information processing in primate basal ganglia. Trends Neurosci 16: 111–116
Parent A, Hazrati L (1995) Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Res Rev 20: 128–154
Parent A, Cicchetti F (1998) The current model of basal ganglia organization under scrutiny. Mov Disord 13: 199–202
Prescott T, Redgrave P, Gurney K (1999) Layered control architectures in robots and vertebrates. Adap Behav 7: 99–127
Redgrave P, Prescott T, Gurney K (1999) The basal ganglia: a vertebrate solution to the selection problem? Neurosci 89: 1009–1023
Smith Y, Bevan M, Shink E, Bolam J (1998) Microcircuitry of the direct and indirect pathways of the basal ganglia. Neurosci 86: 353–387
Snaith S, Holland O (1990) An investigation of two mediation strategies suitable for behavioural control in animals and animats. In: Meyer J, Wilson S (eds) From animals to animats: proceedings of the first international conference on simulation of adaptive behaviour. MIT Press, Cambridge, Mass., pp 255–262
Staines W, Atmadja S, Fibiger H (1981) Demonstration of a pallidostriatal pathway by retrograde transport of HRP-labelled lectin. Brain Res 206: 446–450
Swanson J, Castellanos F, Murias M, LaHoste G, Kennedy J (1998) Cognitive neuroscience of attention deficit hyperactivity disorder and hyperkinetic disorder. Curr Opin Neurobiol 8: 263–271
Von der Malsburg C (1973) Self-organization of orientation sensitive cells in striate cortex. Kybernetik 14: 85–100
Wickens J (1997) Basal ganglia: structure and computations. Netw Comput Neural Syst 8: R77–R109
Wickens J, Alexander M, Miller R (1991) Two dynamic modes of striatal function under dopaminergic–cholinergic control: simulation and analysis of a model. Synapse 8: 1–12
Wilson C (1995) The contribution of cortical neurons to the firing patterns of striatal spiny neurons. In: Houk J, Davis J, Beiser D (eds) Models of information processing in the basal ganglia. MIT Press, Cambridge, Mass., pp 29–50