INSTITUTE OF PHYSICS PUBLISHING NETWORK: COMPUTATION IN NEURAL SYSTEMS

Network: Comput. Neural Syst. **13** (2002) 131–156 PII: S0954-898X(02)32557-0

The role of intra-thalamic and thalamocortical circuits in action selection

M D Humphries and K N Gurney

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

E-mail: m.d.humphries@shef.ac.uk and k.gurney@shef.ac.uk

Received 28 June 2001
Published 21 January 2002
Online at stacks.iop.org/Network/13/131

Abstract

We previously proposed that the basal ganglia (BG) play a crucial role in action selection. Quantitative analysis and simulation of a computational model of the intrinsic BG demonstrated that its output was consistent with this proposition. Here we build on that model by embedding it into a wider circuit containing the motor thalamocortical loop and thalamic reticular nucleus (TRN). Simulation of this extended model showed that the additions gave five main results which are desirable in a selection/switching mechanism. First, low salience actions (i.e. those with low urgency) could be selected. Second, the range of salience values over which actions could be switched between was increased. Third, the contrast between the selected and non-selected actions was enhanced via improved differentiation of outputs from the BG. Fourth, transient increases in the salience of a non-selected action were prevented from interrupting the ongoing action, unless the transient was of sufficient magnitude. Finally, the selection of the ongoing action persisted when a new closely matched salience action became active. The first result was facilitated by the thalamocortical loop; the rest were dependent on the presence of the TRN. Thus, we conclude that the results are consistent with these structures having clearly defined functions in action selection.

1. Introduction

1.1. The action selection problem

Within the vertebrate nervous system there is a selection problem whenever significant stimuli, whether external or internal, co-occur or overlap in time. At this juncture a decision must be made about which stimulus to respond to. For example, an antelope, whilst grazing due to hunger (an internal stimulus), sees what could be a lioness, an external stimulus that causes fear. It is faced with a choice: do I stay and graze or do I flee from a potential predator? The key point is that it cannot do both actions as they require the same motor resource: the leg

muscles. To graze the animal must maintain a stable rigid posture while it lowers its neck to feed, but fleeing requires leg muscles to be used in locomotion. Hence a decision must be taken between two competing actions that require the same motor resource, and presumably some neural substrate must be implicated in making that decision. The selection problem is just as relevant in cognitive aspects of behaviour. Again, some mechanism (action selector) must be arbitrating between competing choices. This paper, while postulating the same underlying neural substrate for both cognitive and motor decision making, will discuss only the motor component as this is a more tractable problem.

An analysis of the basal ganglia's (BG) internal circuitry and external connectivity led us to propose that they played a crucial role in solving the action selection problem (Redgrave *et al* 1999, Prescott *et al* 1999). This led to the development of a computational model of the intrinsic BG (Gurney *et al* 2001a, 2001b). Our primary aim in this paper is to determine the affect on the action selection capabilities of this model by its extension to structures extrinsic to the BG. The plan of the paper is as follows: we first review the particular aspects of BG anatomy which may be pertinent to action selection. We then go on to introduce our functional interpretation of this anatomy and the quantitative model that resulted. This is followed by an introduction to the extrinsic structures (a thalamocortical loop) and a discussion of the primary characteristics of a switching mechanism. Finally, we report the results of testing the original computational model against two versions of the extended model.

1.2. Basal ganglia anatomy and physiology

The BG are a set of interconnected nuclei, predominantly located in the forebrain. Their connections and relative locations in the rat brain are shown in figure 1. The rat BG is principally comprised of the striatum, the subthalamic nucleus (STN), the entopeduncular nucleus, the external segment of the globus pallidus (GPe) and the two sub-divisions of the substantia nigra: pars reticulata (SNr) and pars compacta (SNc). Homologous nuclei exist within most vertebrate brains, for example, the entopeduncular nucleus is equivalent to the globus pallidus internal segment (GPi) in cats and primates.

The striatum receives massive excitatory input from all of cortex (Gerfen and Wilson 1996), from intralaminar thalamic nuclei (Jones 1985, Price 1995), and from the amygdala and hippocampal formation (Kelley *et al* 1982). It is this disparate set of input structures which identified the BG as being a potential neural substrate of the action selection mechanism because it allows striatum to receive sensory, motor, contextual (and other memory forms) and proprioceptive information. Striatum also receives dopaminergic input from the SNc (Gerfen and Wilson 1996), the effect of which is determined by the receptor type on the post-synaptic striatal projection neuron. Dopamine has a predominantly excitatory affect on neurons with D1-type receptors *in vitro* (Umemiya and Raymond 1997) and *in vivo* (Hernandez-Lopez *et al* 1997, Gonon 1997). Conversely, dopamine has an inhibitory affect on D2-type receptor neurons (Delgado *et al* 2000). Output from striatal projection neurons is GABAergic and phasically active and, therefore, provides inhibitory input to their targets within the BG. The distinction between D1 and D2 type receptor neurons is further reinforced by their projection targets: D1 neurons project predominantly to SNr and GPi; D2 neurons project predominantly to GPe (Gerfen *et al* 1990).

A feature of medium spiny projection neurons, which may be crucial to action selection, is that they are silent in their normal state (DOWN-state). In this state, the neuron's membrane potential is far below the firing threshold, and is very stable. Massive co-ordinated input from cortex and other afferent connections is required to push the neuron into its firing-ready UP state. Thus, striatum can act as a filter to prevent low-level signals from reaching its afferent nuclei within the BG.

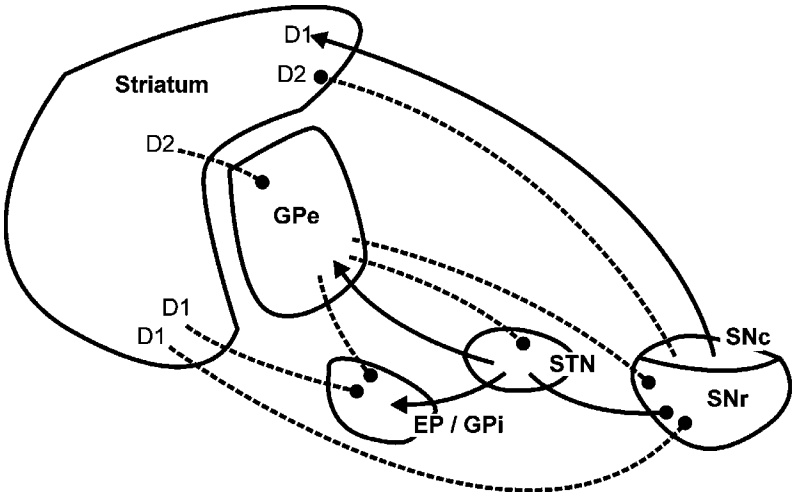

Figure 1. The connectivity, relative position, and relative size of the nuclei that comprise the rat BG. Note the separate projection targets of the D1 and D2 receptor striatal neurons. Excitatory pathway: ——; inhibitory pathway: - - - -. GPe: external segment of the GPe. STN: subthalamic nucleus. SNc: substantia nigra pars compacta. SNr: substantia nigra pars reticulata. EP(GPi): entopeduncular nucleus (internal segment of the globus pallidus).

The STN receives widespread excitatory cortical input, in addition to excitatory input from intralaminar thalamic nuclei (Bevan *et al* 1995). Its output is excitatory and tonically active, with a spontaneous firing rate between 10–30 Hz (Fujimoto and Kita 1993, Wichmann *et al* 1994). Hence, it is the major source of excitation within the BG, driving GPe, GPi and SNr. STN also receives inhibitory (GABAergic) input from GPe neurons, forming a negative feedback loop. Besides the excitatory input from STN, GPe also receives inhibitory input from D2-type striatal neurons (see above). The GPe neurons project to SNr, SNc and GPi, providing inhibitory input.

GPi and SNr are the output nuclei of the BG. They tonically inhibit their thalamic and hindbrain targets, in part because they are driven by the tonic excitatory input from STN. Tracing studies have shown that output projections from these nuclei are segregated into separate, discrete channels which have clearly delineated targets (Hoover and Strick 1999, 1993). The anatomical channels found in the output nuclei appear to be maintained throughout the BG (Gerfen and Wilson 1996). This is further delineated in motor areas, where there is a clear somatotopic organization maintained throughout the BG (Hoover and Strick 1999, Brown and Sharp 1995, Brown *et al* 1998).

1.3. Functional anatomy of the BG

While much is known about the anatomy and physiology of the BG researchers have had difficulty clearly defining its functional role. The dominant model of BG functional anatomy, illustrated in figure 2(*a*), was proposed by Albin and colleagues (Albin *et al* 1989). They split the internal connections of the BG into two pathways: the direct pathway containing the D1-type striatal cells and the output nuclei, and the indirect pathway containing the D2-type striatal cells, the STN and the GPe. The direct pathway expressed the output of the BG through the inhibition of the GPi/SNr and, therefore, the disinhibition of their target structures; the indirect pathway had a modulatory influence on the direct pathway. This

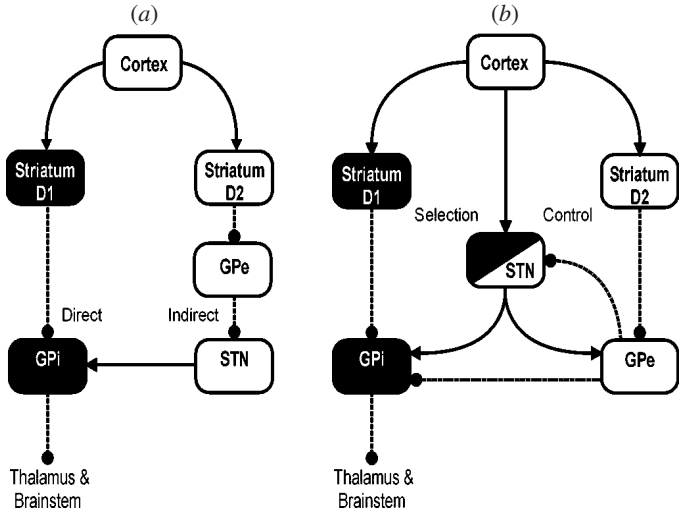

Figure 2. (*a*) The direct and indirect pathway model of BG anatomy proposed by Albin *et al* (1989). (*b*) The proposed new functional anatomy of the BG, from Gurney *et al* (2001a). The selection pathway (shaded boxes) causes the disinhibition of the BG's projection targets. One of the actions of the control pathway is to scale the level of activity in the selection pathway. The STN is half-shaded as it plays a role in both pathways. Excitatory: ——. Inhibitory: - - - -.

model, though still widely used (Onla-or and Winstein 2001, Ni *et al* 2001), has been criticized for failing to include numerous important anatomical pathways (Parent and Cicchetti 1998, Parent *et al* 2001).

We proposed a new functional anatomy (see figure 2(*b*)) based on our hypothesis that the BG plays a crucial role in action selection (Gurney *et al* 2001a). This implies that the BG component structures should be described in terms of their ability to form neural circuits for selecting signal inputs. In the new model the 'selection' pathway (which performs selection per se) and 'control' pathway (which has a modulatory influence on the selection pathway) supersede the 'direct' and 'indirect' pathways of Albin *et al*'s model. We go on here to outline how the BG could operate as an action selector, based on this functional anatomy.

The functional anatomy contains a channel architecture, based on the tracing studies described above. Every possible appropriate action is represented in a separate channel. The level of neural activity input to a channel represents the urgency attached to that action, which we term 'salience'. Salience levels are calculated in the striatum from sensory, proprioceptive and contextual information. The striatum's extensive afferent connections, described above, make it ideally placed to perform this function. The salience level provides the basis for action selection; the action(s) with the highest salience are selected and executed.

Within the selection pathway, the more salient an action is, the more the D1 striatal neurons inhibit the output nuclei neurons of the same channel. The tonic inhibitory output of each channel in the GPi/SNr, therefore, decreases with increasing inhibition provided by the striatum (and, by definition, with increases in the salience level). With a sufficiently high level of salience the output channel can effectively be turned off, removing its tonic inhibition. This causes complete disinhibition of the targets of that channel (a subset of cells in the BG target nuclei and/or a subset of those nuclei). Any subsequent excitatory input reaching these targets such as, for example, a motor command from the cortex, could then cause the target cells to fire. This mechanism of selection has been termed 'selective disinhibition' (Chevalier

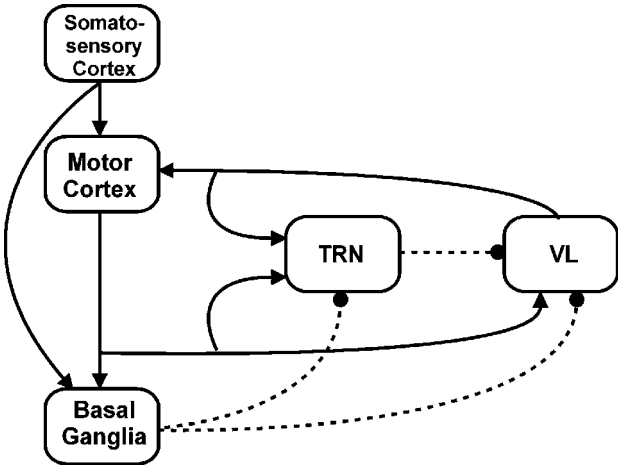

Figure 3. The connection scheme of the extended computational model. The box labelled 'Basal Ganglia' contains the functional anatomy shown in figure 2. Sensory signal levels are input to the model from the somatosensory cortex. TRN: thalamic reticular nucleus. VL: ventrolateral thalamus. Excitatory: ——. Inhibitory: - - - -.

et al 1985). Furthermore, the diffuse projections of STN neurons to GPi/SNr across multiple channels allows STN to increase the output level of non-selected channels. This emphasizes the difference in output of the selected and non-selected channels.

However, because each GPi output channel receives input from all STN channels, a mechanism for scaling the level of excitation is necessary. The control pathway has several functions, one of which is to limit, via the GPe–STN negative feedback loop, the overall level of activity (Gurney *et al* 1998, 2001b). This limit is kept roughly constant as BG recruits more channels so that the selection process can continue properly.

Analysis and simulation of a quantitative model of the functional anatomy showed that the BG were capable of outputting signals consistent with action selection (Gurney *et al* 2001b). This model will henceforth be denoted as the *intrinsic* model. Further, the ability of the model to perform action selection was critically dependent on the level of dopamine. When dopamine levels were too high, multiple channels were selected too easily: this may correspond to many actions being executed and is consistent with behavioural states associated with attention deficit hyperactive disorder (ADHD) (Swanson *et al* 1998). Very low levels of dopamine resulted in no selection occurring. This may correspond to immobility and the inability to initiate actions, as observed in Parkinson's disease patients. Thus, direct parallels could be drawn between the model's behaviour under abnormal dopamine conditions and disorders known to be caused by dysfunction of the BG.

2. Thalamocortical interactions

2.1. The motor thalamocortical loop

To extend the intrinsic model, we have embedded it into a thalamocortical loop comprising the ventrolateral (VL) thalamus, the motor cortex, and the thalamic reticular nucleus (TRN). This extended model is shown in figure 3.

The VL thalamus is normally considered to be the principal motor thalamic nucleus (Price 1995, Jones 1985). GPi neurons project densely to the VL thalamus in all vertebrates (Uno *et al*

1978) and the VL thalamic neurons project in turn to the primary motor cortex (Zarzecki 1991). Completing the loop, primary motor cortex neurons project to the striatum (Turner and DeLong 2000, Gerfen and Wilson 1996) and reciprocally to the VL thalamus (Price 1995). Collaterals from the corticothalamic and thalamocortical projections converge on neurons in the motor sector of the TRN. These TRN neurons then send inhibitory input back to the VL thalamus. Hence there are three loops in operation: the extended motor cortex–BG–VL thalamus–motor cortex loop; the positive feedback loop formed by the motor cortex and VL thalamus; and the negative feedback loop formed by the TRN and VL thalamus which is modulated by cortical input.

In extending the BG functional anatomy to incorporate the motor cortex, VL thalamus, and TRN, we retained the channel-based architecture of the intrinsic model for several reasons. First, as already noted, the BG output nuclei send somatotopically organised projections to their targets. Second, all cortically projecting thalamic nuclei contain topographic maps and these maps are maintained in the projections to the cortical areas (Adams *et al* 1997). By extension, the somatotopic projections from the GPi are maintained in the VL thalamus projections. Third, primary motor cortex contains a well defined somatotopic map, with separate areas for vibrissae, eyes, tongue, lips and so on (Hall and Lindholm 1974). Fourth, the projection from primary motor cortex to striatum is also somatotopically organized (Brown *et al* 1998, Brown and Sharp 1995). Thus, there is substantial neuroanatomical evidence that discrete channels are maintained outside the BG, at least at the level of separate body parts.

Input to the model comes from somatosensory cortex and represents pre-processed sensory data. A copy of the input is sent to the motor cortex. This would occur via the extensive intracortical connections (Farkas *et al* 1999, Izraeli and Porter 1995, Miyashita *et al* 1994). The original input signal and the copy (via motor cortex) are summed in the striatum, forming the salience level. This operation would be achieved via the overlapping projection targets of somatosensory cortex and motor cortex neurons in the striatum (Flaherty and Graybiel 1994).

2.2. The thalamic reticular nucleus

Although the TRN has been attributed a variety of roles related to attention (Weese *et al* 1999, Crick 1984, Guillery *et al* 1998, Newman *et al* 1997), the functional design of the TRN used in the current model did not explicitly address attentional processing. Rather, it followed from anatomical data on the TRN interpreted in the channel-based scheme. The design's basic form is shown in figure 4 for three channels.

The TRN can be divided into multiple sensory, motor and 'association' sectors defined by their input–output connections (Guillery *et al* 1998, Lozsádi 1994). Each sector receives input from the appropriate thalamic nuclei and cortical areas, which are collaterals of the reciprocal projections between the thalamus and cortex. In turn, the neurons in a TRN sector project back solely to the appropriate thalamic nucleus. For example, in the rat, the vibrissal somatosensory sector of the TRN receives input from the 'barrel' cortex (Hoogland *et al* 1987) and ventrobasal thalamus, and projects back to the ventrobasal thalamus (Shosaku *et al* 1984). This level of topographic organization within the TRN allows each sector and its associated thalamocortical loop to be regarded as a distinct circuit.

TRN neurons rarely project back to the thalamic neurons from which they receive synaptic contact. When a thalamocortical neuron is driven by excitatory input, the TRN neuron(s) it synapses on are driven to inhibit other thalamocortical neurons in the same nucleus. Hence, the output of one thalamocortical neuron causes the inhibition of surrounding thalamocortical neurons, via a distal mechanism (Pinault and Deschênes 1998).

We have interpreted this anatomical organization in terms of our channel architecture (see figure 4). The thalamocortical neurons in a channel project only to the TRN neurons in the same channel; these TRN neurons then project predominantly back to VL thalamus neurons in

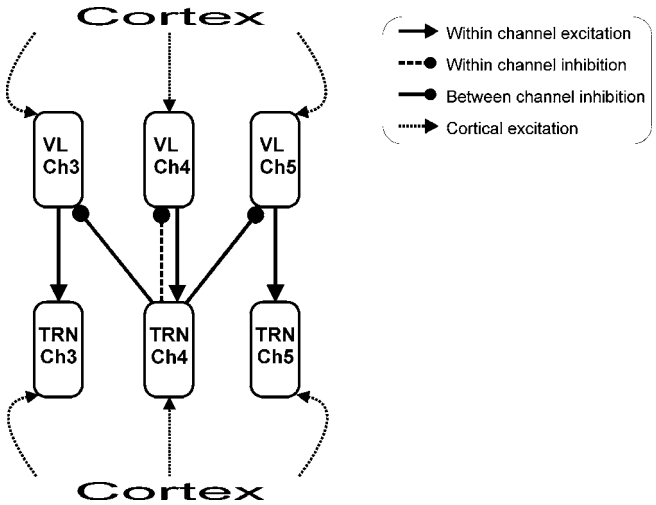

Figure 4. The interconnectivity of the VL thalamus and the TRN. Full connections are only shown for channel 4. BG inputs are omitted for clarity.

other channels. We have, however, included a sparse inhibitory input from the TRN channel back to the corresponding channel in the VL. This models the minority of TRN neurons which *do* synapse back on to the VL thalamus neurons which project to them (Kim and McCormick 1998), forming a closed loop. The TRN design we have adopted means that there is both within- and between-channel inhibition in VL thalamus, but only within-channel excitation in the TRN. Therefore, the TRN acts as a distal lateral inhibition mechanism.

While this channel-based interpretation of the TRN-dorsal thalamus connectivity remains to be confirmed anatomically in the motor pathway, there is evidence for it in other pathways. The ascending sensory pathway of the rat vibrissal system contains discrete cell groups in every structure. In the ventrobasal thalamus, these discrete cell groups (barreloids) respond maximally to the stimulation of a specific whisker (Diamond 1995, Friedberg *et al* 1999). Thalamocortical neurons in one barreloid all project to the same area of TRN, forming a corresponding discrete cell cluster of maximally responsive neurons in the TRN (Sumitomo and Iwama 1987). Sumitomo and Iwama also found that stimulation of a single whisker activated the cells in the corresponding barreloid and, as expected, these excited the corresponding area of the TRN. However, the activity in other barreloids was inhibited. Hence, these TRN neurons appeared to project back to thalamocortical neurons in other barreloids. Our channel-based scheme is, therefore, consistent with mechanisms of between-channel inhibition in place in the rat TRN-dorsal thalamus complex.

Finally, there is evidence for a sparse inhibitory input from the BG output nuclei to the TRN. Projections from the SNr synapse in TRN sectors which project to VL and VM thalamus (Kolmac and Mitrofanis 1998, Cornwall *et al* 1990, Gandia *et al* 1993). In the model, this input has been assumed to be channel specific. This connection will be discussed in more detail in section 4.8.

2.3. The requirements of a switching mechanism

Integral to the BG's postulated role as the vertebrate brain action selector is the ability to switch between actions after an initial action has been selected. For this role, we laid out

two computational requirements and three desirable characteristics that a central switching mechanism must fulfil (Redgrave *et al* 1999). The first computational requirement has been outlined above, namely that selection depends on the relative salience levels of the competing actions. Therefore, the action with the highest salience is selected and expressed, with the caveat that the level of GPi/SNr output must be below some selection threshold to allow sufficient disinhibition of the target structures.

The second computational requirement covered the criteria for terminating an action. Termination may occur, in a normally functioning BG, when an action is interrupted by a competing action with a higher salience. This condition ensures that the animal progresses smoothly from one action to another without a conflicting use of motor resources occurring. We previously proved that the intrinsic model was able to fulfil both of these computational requirements (Gurney *et al* 2001b).

We also set out the desirable characteristics to define how an ideal switching mechanism should operate and the behavioural affects this would have. The primary characteristic was *clean switching*: the competition between actions should be resolved rapidly and decisively. If this did not occur then an animal would be forced into lengthy periods of quiesence after the termination of an action while the next one is being selected. This would leave an animal vulnerable to predators and natural hazards.

Once an action has been successfully selected, it is undesirable that its expression be hampered or temporarily interrupted by losing competitors. The second desirable characteristic, therefore, was that there should be an *absence of distortion*. Any losing competitor should be sufficiently suppressed so that it can not interfere with the selected action. Furthermore, short-term increases in salience on a non-selected channel should also be suppressed, unless the increase is of a sufficient magnitude to warrant selection.

When the competition for selection is between actions with almost equal saliences then it is essential that the selected action continues after its salience has dropped below the level of its immediate competitors. We call this desirable characteristic *persistence* and, to illustrate its utility, consider the case in which an animal is almost equally hungry and thirsty. One of these, say thirst, is slightly more salient and so the action of drinking is selected. However, after a brief period of drinking, the salience level of thirst drops below that of hunger, so the animal begins to eat. Again, after a brief period of eating, the salience level of hunger would drop below that of thirst, and the animal would return to drinking. This could continue *ad infinitum* as neither sensation is properly satiated, with the animal constantly oscillating between two behaviours: a phenomenon known as 'dithering' (Houston and Sumida 1985). Hence, the ability for an action to persist until it has been satisfactorily completed is essential for normal function.

One feature of salience input which is pertinent to the arguments above concerning persistence is the possibility of transient increases in salience level. The suppression of transient increases on non-selected channels would allow the selected action to be successfully completed. However, there must also be allowances for switching to a sufficiently high salience action, so that the animal is not forced to ignore potentially important events.

2.4. Aims

While the output signals from the intrinsic model were consistent with action selection, we have yet to determine whether the model works effectively when embedded in a wider anatomical context; hence we constructed a model which includes the thalamus and cortex. We had two aims for this extended model. First, to ascertain whether the extended model could fulfil the desirable characteristics of a switching mechanism laid out in the previous section. Second, to show that the selection and switching capabilities of the intrinsic model could be maintained.

3. Quantitative modelling

3.1. The model neuron

The extended model uses leaky-integrator artificial neurons, just as were used in the intrinsic model (Gurney *et al* 2001b). Let u be the total afferent input and k be the constant determining the rate of activation decay. The activation a of a leaky integrator is then

$$\dot{a} = -k(a - u) \tag{1}$$

where $\dot{a} \equiv \mathrm{d}a/\mathrm{d}t$. In all that follows, we are describing the activation at equilibrium $\tilde{a}$ which is just $\tilde{a} = u$.

The output y of the neuron, corresponding to the mean firing rate, is bounded below by 0 and above by 1. In simulation, this is achieved by using a piecewise linear output function. However, as we previously demonstrated (Gurney *et al* 2001b), it is possible to ensure that y never exceeds 1 so that the output relation can be written as

$$y = m(a - \epsilon)H(a - \epsilon) \tag{2}$$

where ϵ is the output threshold, $H()$ is the Heaviside step function, and m is the slope of the output function.

3.2. Motor cortex

Motor cortex receives sensory input S_i from somatosensory cortex, and input from VL thalamus y_i^v, where i is the channel index. The strength of the synaptic connections from VL thalamus and somatosensory cortex are denoted by w_{vl} and w_s, respectively (weights are assumed to be given as absolute magnitudes throughout this section). If $\tilde{a}_i^m$ is the equilibrium activation of channel i in motor cortex then

$$\tilde{a}_i^m = w_{vl} y_i^v + w_s S_i. \tag{3}$$

Then, if ϵ_m is the output relation threshold term (see equation (2)) the output y_i^m of motor cortex is given by

$$y_i^m = m(\tilde{a}_i^m - \epsilon_m)H(\tilde{a}_i^m - \epsilon_m). \tag{4}$$

3.3. Thalamic reticular nucleus

The TRN receives input from three sources: motor cortex y_i^m, VL thalamus y_i^v and the BG output nuclei, y_i^b. Let the synaptic strengths from the three sources be w_m, w_v and w_{bg}, respectively, and let $\tilde{a}_i^t$ be the activation at equilibrium of the ith channel of TRN, then

$$\tilde{a}_i^t = w_v y_i^v + w_m y_i^m - w_{bg} y_i^b. \tag{5}$$

Thus, if ϵ_t is the output relation threshold term, then the output y_i^t of TRN is given by

$$y_i^t = m(\tilde{a}_i^t - \epsilon_t)H(\tilde{a}_i^t - \epsilon_t). \tag{6}$$

3.4. Ventrolateral thalamus

VL thalamus receives input from motor cortex y_i^m, the BG output nuclei y_i^b and the TRN y_i^t. The inhibitory input from the TRN has two distinct components, as discussed in section 2.2. Within-channel input is assigned the weight w_T^*. The between-channel input contacts all channels in

VL thalamus except the corresponding channel i. Hence, the total between-channel output of the TRN Y_i^t is given by

$$Y_i^t = m \sum_{j \neq i}^{n} (\tilde{a}_j^t - \epsilon_t) H(\tilde{a}_j^t - \epsilon_t) \tag{7}$$

where n is the total number of channels. Let the strength of the between channel connection be w_b, the motor cortical input be w_x, and BG input be w_o, then the activation at equilibrium of the ith channel in VL thalamus is

$$\tilde{a}_i^v = w_x y_i^m - (w_o y_i^b + w_T^* y_i^t + w_b Y_i^t). \tag{8}$$

Then, if ϵ_v is the output relation threshold term, the output y_i^v of VL thalamus becomes

$$y_i^v = m(\tilde{a}_i^v - \epsilon_v) H(\tilde{a}_i^v - \epsilon_v). \tag{9}$$

3.5. Striatum

Input to the striatum is a combination of sensory input (from somatosensory cortex) and motor cortical input, which we denote by S_i and y_i^m, respectively. The strength of the synaptic connections from somatosensory and motor cortex are w_{sc} and w_{mc}. Thus, the salience level c input to the ith striatal channel is given by

$$c_i = w_{sc} S_i + w_{mc} y_i^m. \tag{10}$$

We retain the disparate action of dopamine used in the intrinsic model in the two separate pathways (selection and control) by using a multiplicative factor in the synaptic weight. Thus, let λ_e and λ_g parametrize the tonic level of dopamine in the control and selection pathways respectively, where $0 \leqslant \lambda_e, \lambda_g \leqslant 1$. Then the action of dopamine in the control pathway can be characterized as a modification to the synaptic weights: $(1 - \lambda_e) w_{sc}$ and $(1 - \lambda_e) w_{mc}$. Similarly, for the selection pathway $(1 + \lambda_g) w_{sc}$ and $(1 + \lambda_g) w_{mc}$. The activation functions for D1 and D2 striatal neurons, respectively, are then

$$\begin{aligned} \tilde{a}_i^g &= c_i (1 + \lambda_g) \\ \tilde{a}_i^e &= c_i (1 - \lambda_e). \end{aligned} \tag{11}$$

The output relation for neurons in the selection pathway is

$$y_i^g = m(\tilde{a}_i^g - \epsilon) H^{\uparrow}(\lambda_g) \tag{12}$$

where $H^{\uparrow}(\lambda_g) = H(\tilde{a}_i^g - \epsilon)$. The up-arrow emphasizes that the value of ϵ indicates the difficulty of the UP/DOWN-state transition: if ϵ is given a positive value, then the neuron must receive input of at least this level to have a non-zero output. Similarly, the control pathway's output relation is

$$y_i^e = m(\tilde{a}_i^e - \epsilon) H^{\uparrow}(\lambda_e). \tag{13}$$

3.6. STN

Similar to the previous section on striatum, the equations describing STN activation and output have to be rewritten to accommodate the split input from somatosensory cortex S_i and motor cortex y_i^m, which replaces the original single salience input. Let the strength of the synaptic connections from somatosensory and motor cortex be w_{st} and w_{mt}. Then the equilibrium activation $\tilde{a}^+$ of the ith STN channel is given by

$$\tilde{a}_i^+ = w_{st} S_i + w_{mt} y_i^m + w_g y_i^p, \tag{14}$$

where w_g is the weight of the GP–STN pathway and y_i^p is the output of GP. If ϵ' is the output relation threshold term, the output y_i^+ of STN remains

$$y_i^+ = m(\tilde{a}_i^+ - \epsilon')H(\tilde{a}_i^+ - \epsilon'), \tag{15}$$

where ϵ' is given a moderate negative value to simulate the tonic output of STN. However, as STN output is diffuse across all channels in its target structures we need to consider the total STN output Y^+, which is given by

$$Y^+ = m\sum_{i=1}^{n}(\tilde{a}_i^+ - \epsilon')H(\tilde{a}_i^+ - \epsilon'), \tag{16}$$

where n is the number of channels.

3.7. GPe and GPi

For the convenience of the reader, we briefly recap the equations for the activation and output of the GPe and GPi from Gurney *et al* (2001b). We emphasize that the descriptions of the BG nuclei (striatum, STN, GPe, and GPi) given here are identical to those used in the intrinsic model, except for the split somatosensory and motor cortex input which replaces the single salience input into striatum and STN.

A GPe channel receives input from the corresponding striatal D2 population channel y_i^e (in the control pathway). It also receives input from all the STN channels Y^+. Let the striatum D2 to GPe connection strength be w_{ep} and that of the STN to GPe connection be w_{sp}. Then the activation at equilibrium $\tilde{a}_i^p$ of the ith GPe channel is given by

$$\tilde{a}_i^p = w_{sp}Y^+ - w_{ep}y_i^e. \tag{17}$$

If ϵ_p is the output threshold term, then the output y_i^p of the ith GPe channel is

$$y_i^p = m(\tilde{a}_i^p - \epsilon_p)H(\tilde{a}_i^p - \epsilon_p). \tag{18}$$

A GPi channel receives input from three sources: the corresponding striatal D1 population y_i^g and GPe channel y_i^p, and diffuse STN input Y^+. The strength of the synaptic connections from striatum D1, GPe, and STN are given by w_{gb}, w_{pb}, and w_{sb}, respectively. The equilibrium activation of the ith GPi channel $\tilde{a}_i^b$ is thus given by

$$\tilde{a}_i^b = w_{sb}Y^+ - w_{pb}y_i^p - w_{gb}y_i^e. \tag{19}$$

Letting ϵ_b be the output threshold term, the output y_i^b of a GPi channel is

$$y_i^b = m(\tilde{a}_i^b - \epsilon_b)H(\tilde{a}_i^b - \epsilon_b). \tag{20}$$

4. Simulation results

4.1. Parameter details

All simulations were conducted using a six channel model, to be consistent with the previously published model. Parameter settings from the original model (e.g. w_g) were maintained (see Gurney *et al* (2001b) for details). The slope parameter m was set to 1 throughout. The weights w_{vl}, w_s, w_x, w_m, w_o, and w_v were also set to 1. Weights w_{sc}, w_{mc}, w_{st}, and w_{mt} were set to 0.5, to model the roughly equal level of somatosensory and motor cortex input to striatum and STN, and to limit the total excitatory input to a possible maximum of 1 (in line with the synaptic strength of the direct salience connection in the intrinsic model). To model the very sparse nature of closed loop connections between VL thalamus and the TRN, w_T^* was set to 0.1. The open-loop (between channel) connection weight w_b was set to 0.7, because TRN

input to VL thalamus is significantly smaller than cortical feedback (Price 1995). Finally, w_{bg} was set to 0.2 to model the sparse nature of the connection between the BG output nuclei and the TRN.

The dopamine parameters, λ_e and λ_g, were set to 0.2, as they were for the intrinsic model. The offset (threshold) values ϵ_m, ϵ_t, and ϵ_v were set to 0. To model the difficulty of forcing striatal neurons in to their UP-state, the offset ϵ was set to 0.2, so that a significant level of excitatory input would be required before a striatal neuron gave a non-zero output. Finally, to ensure that STN units give tonic output in the absence of input, the offset value ϵ' was set to -0.25.

4.2. Example outputs

We first consider example outputs which illustrate the features analysed in the subsequent sections. Three models were tested for this paper. First, the original intrinsic model, for comparison purposes with the new models. Second, a model with the thalamocortical loop (TC model) but no TRN in which the weights w_T^* and w_b were set to 0. This allowed us to study the affect of the positive feedback loop formed by motor cortex and VL thalamus. Third, the full extended model, shown in figure 3, which we shall term the TRN model. We emphasize again that the descriptions of the BG nuclei were identical across the three models except for the instantiation of two input sources to striatum and STN in the TC and TRN models. Figure 5 shows the outputs of the GPi and motor cortex (where applicable), and the sensory input, on three channels for all three models. In each simulation, the input on channel 1 (0.4) started at $t = 1$, and the input on channel 2 (0.6) started at $t = 2$. A transient event on channel 1 began at $t = 3$ and terminated at $t = 4$. The transient was 0.2 in amplitude so that the sensory inputs on both channels temporarily became equal.

A channel i was considered selected if the GPi output y_i^b of that channel fell below the selection threshold θ_s, which was set to 0.05. This was taken to indicate that the tonic inhibition from the GPi had decreased sufficiently to allow its targets to fire. Although a zero output would demonstrate unequivocal selection, it is unrealistic to suppose that neurons have to be held in a completely silent state for a behaviourally meaningful period to indicate selection. Therefore, θ_s is given a non-zero value to allow for low level firing that may occur in real BG output neurons while under inhibition from D1 striatal neurons.

The successful switching between actions is dependent on the selection of the newly winning competitor and the deselection of the losing competitor (the desirable characteristic of absence of distortion—see section 2.3). In model terms, a successful switch occurred when the channel 2 GPi output fell below θ_s and the channel 1 GPi output exceeded θ_s after the onset of channel 2 input.

Inspection of the left-hand column of graphs in figure 5 shows that after the onset of sensory input on channel 1 the GPi output fell below θ_s in all models and, therefore, the channel became selected. After the onset of sensory input on channel 2, this channel also became selected (centre column). Simultaneously, the GPi output on channel 1 rose above the selection threshold, thereby deselecting the channel. Thus, with this input pair, successful selection of both channels, and switching between them, could take place. Of particular note is the different levels of channel 1 GPi output equilibrium reached after deselection: both the TRN and intrinsic models showed a considerable contrast between their channels 1 and 2 GPi outputs at time $t \cong 2.7$, with the TRN model showing the greatest difference. The difference in output of the TC model's channels 1 and 2 at $t \cong 2.7$ was much smaller.

The transient event on channel 1 illustrates the abilities of the three models to maintain an ongoing selection in the face of possible disruption. During the transient, the GPi output of

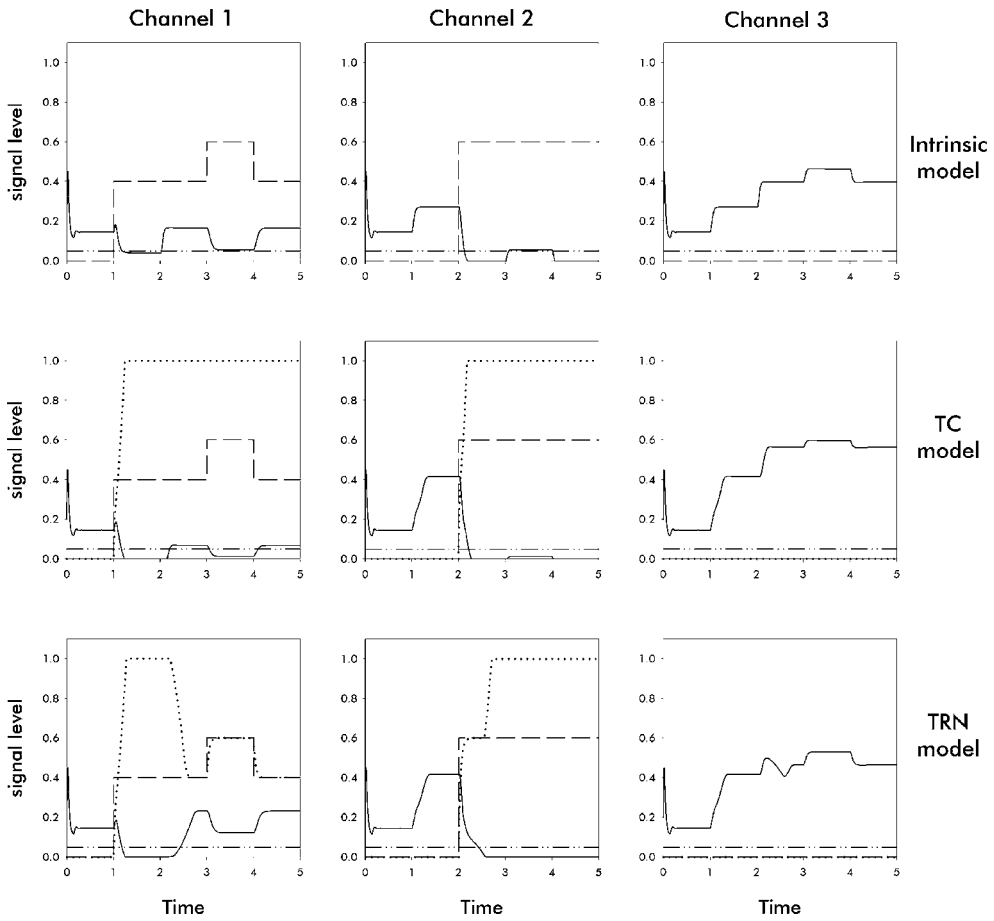

Figure 5. Example outputs from the intrinsic model (top), TC model (middle), and TRN model (bottom). They show the selection (at $t = 1$) and switching (at $t = 2$) capabilities of all three models. Note the varying levels of inter-channel differences between the GPi outputs across the three models after the onset of channel 2 salience. Only the TRN model showed successful suppression of the transient event at $t = 3$. The third column illustrates the GPi output changes in a non-active channel in response to changes in input levels on active channels. The dash–dotted line represents the selection threshold θ_s. GPi: ——. Motor cortex: ·····. Sensory input: - - - -.

channel 1 in both the intrinsic and TC models is equal to the GPi output of channel 2. As both outputs are below θ_s we must conclude that the two actions are expressed simultaneously. In the TRN model, the GPi output on channel 1 is prevented from falling below θ_s, while GPi output on channel 2 is held constant. Thus, the transient event does not adversely affect the ongoing selected action in the TRN model, but potentially could do so in the other models.

The right-hand column of graphs shows the GPi output on a third, inactive, channel in each of the three models (figure 5). They illustrate how the onset of sensory inputs to other channels causes a corresponding rise in the GPi output level of the inactive channels in all models. This rise in output level serves to strengthen the contrast between active and non-active channels, ensuring that the ongoing action selection competition is not affected.

For the TC and TRN models, the output of motor cortex is also plotted for all channels. Obviously, with no sensory input, there is nothing to drive motor cortex and so it stays silent. When there is sensory input, the models respond differently. The onset of input to a TC

model channel causes the motor cortex output to rapidly reach saturation and to remain there constantly, regardless of sensory events in the same or other channels. However, the motor cortex output in the TRN model becomes saturated in the selected channel only. When that channel is deselected, the motor cortical output falls to the level of the sensory input. Note that, in channel 1, motor cortex output in the TRN model does not increase beyond the sensory input level during the transient event but is suppressed.

4.3. Selection properties

It was first necessary to establish that the introduction of the thalamocortical loop and the TRN did not cause the loss of the established selection and switching abilities of the intrinsic BG model (Gurney *et al* 2001b). To test this, 121 simulations were run on each model, consisting of the sensory input pairs (S_1, S_2) where S_1 and S_2 range from 0 to 1 in steps of 0.1. The input to channel 1 began at time $t = 1$; the input to channel 2 began at $t = 2$. This gives us two time intervals in which GPi output on any channel may change: $I_1 = 1 \leqslant t \leqslant 2$ and $I_2 = t > 2$.

In the previous section, selection of channel i was defined as $y_i^b \leqslant \theta_s$. Using this definition, and given the onset times of the sensory input, the outcome of a simulation could be characterized by one of four states. First, *no selection*, where $y_1^b \wedge y_2^b > \theta_s \forall t$. Second, *selection*, where $y_1^b \leqslant \theta_s$ in $I_1 \wedge y_2^b > \theta_s$ in I_2; or $y_1^b > \theta_s \forall t \wedge y_2^b \leqslant \theta_s$ in I_2. Third, *no switching*, where $y_1^b \wedge y_2^b \leqslant \theta_s$ in I_2. Fourth, *switching*, where $y_1^b \leqslant \theta_s$ in $I_1 \wedge y_1^b > \theta_s$ in $I_2 \wedge y_2^b \leqslant \theta_s$ in I_2. The sets of sensory input pairs which caused these four output states in each of the three models are illustrated in figure 6.

The results are summarized in the histogram in the bottom right-hand corner. It is clear from the plots that all three models showed successful selection across a wide range of sensory input pairs. Furthermore, all three models were capable of successful switching. However, a comparison of the three models shows that the introduction of the thalamocortical loop (TC model) allowed much lower-level sensory inputs to become selected: the minimum input needed was 0.2, compared with the intrinsic model's minimum of 0.4. However, there was a large increase in the number of instances for which there was no switching in the TC model, when compared to the intrinsic model. Hence, there was a trade-off between low input selection capability and the ability to successfully switch over a wide range of pairs.

After the introduction of the TRN the low input selection capability gained by the thalamocortical loop was maintained. Furthermore, the TRN model was able to successfully switch between selected inputs over a much greater range of pairs then the TC model, and returned switching performance to that of the intrinsic model. Thus, both the TRN and intrinsic models were able to perform one aspect of clean switching in that most sensory input pair competitions were resolved decisively.

4.4. Improving GPi output contrast

Further to the basic selection properties of the model, we investigated the affect of the thalamocortical loop and TRN on the differentiation of GPi outputs. This is an important aspect of maintaining an absence of distortion, for the GPi outputs of the selected and non-selected channels should be sufficiently different so that the non-selected channel outputs cannot interrupt the ongoing action. Thus, although we interpret the model neuron outputs as a mean firing rate, the noisy output of real neurons in a non-selected channel may temporarily fall below the output of the selected channel. We determine the potential for distortion by simply measuring the difference between the GPi outputs on channels 1 and 2 when they have reached equilibrium after the onset of input on both channels.

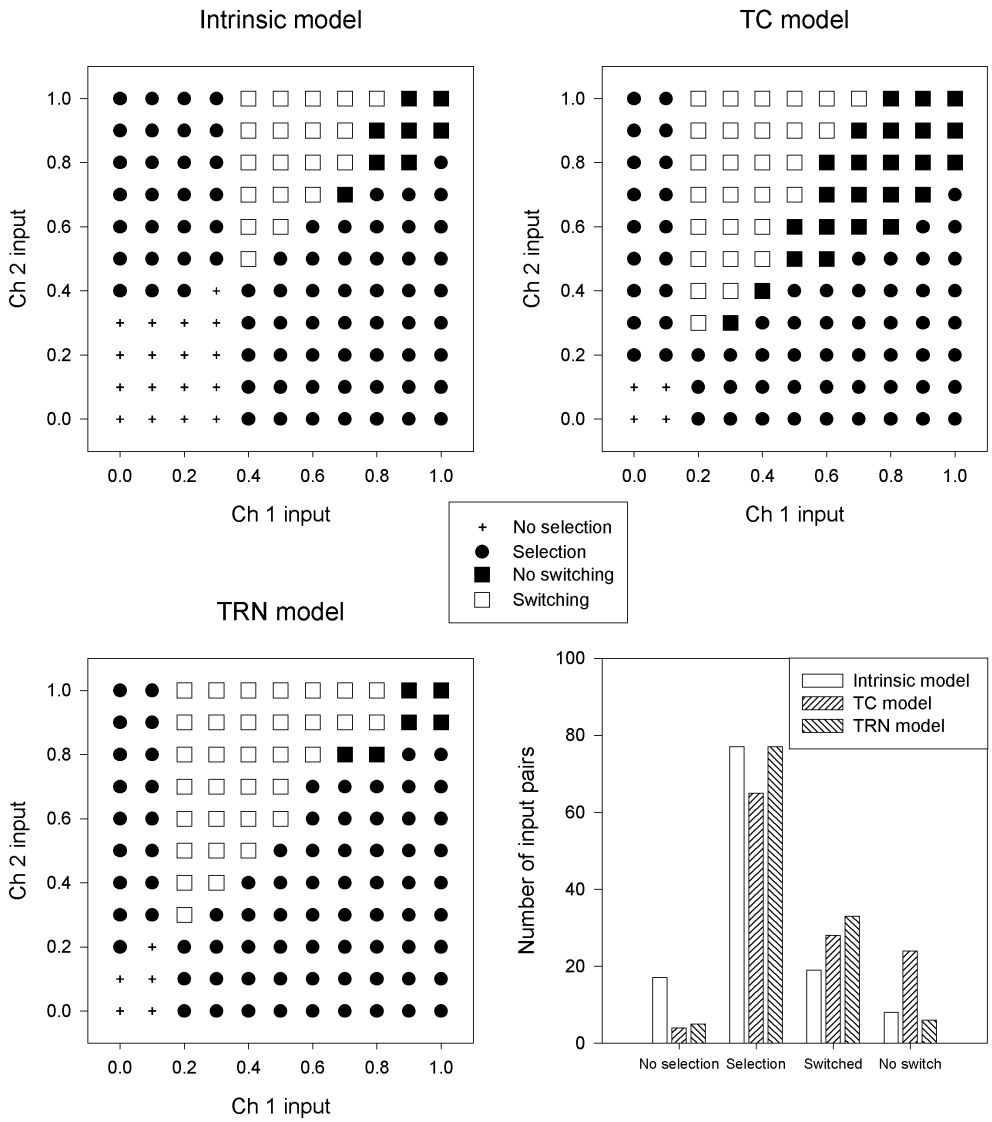

Figure 6. Selection and switching abilities of the intrinsic, TC and TRN models, in response to the input sequences described in the text. The histogram shows the number of cases that cause each of the four output states. The addition of the thalamocortical loop allowed channel selection with low level inputs. Adding the TRN maintained low level selection while also improving the switching capability.

There are two possible ways in which output contrast could be compromized. First, through the disruption of the selected channel by another channel becoming active with a low-level competitor (i.e. channel 2 after channel 1 has been selected). Second, that the switching is not clean enough, due to a high input level on the deselected channel, and so the output is not sufficiently differentiated when switching is completed (i.e. high-level input on channel 1 after channel 2 has been selected). Hence, we use the absolute difference between the GPi outputs so that both possibilities could be taken into account.

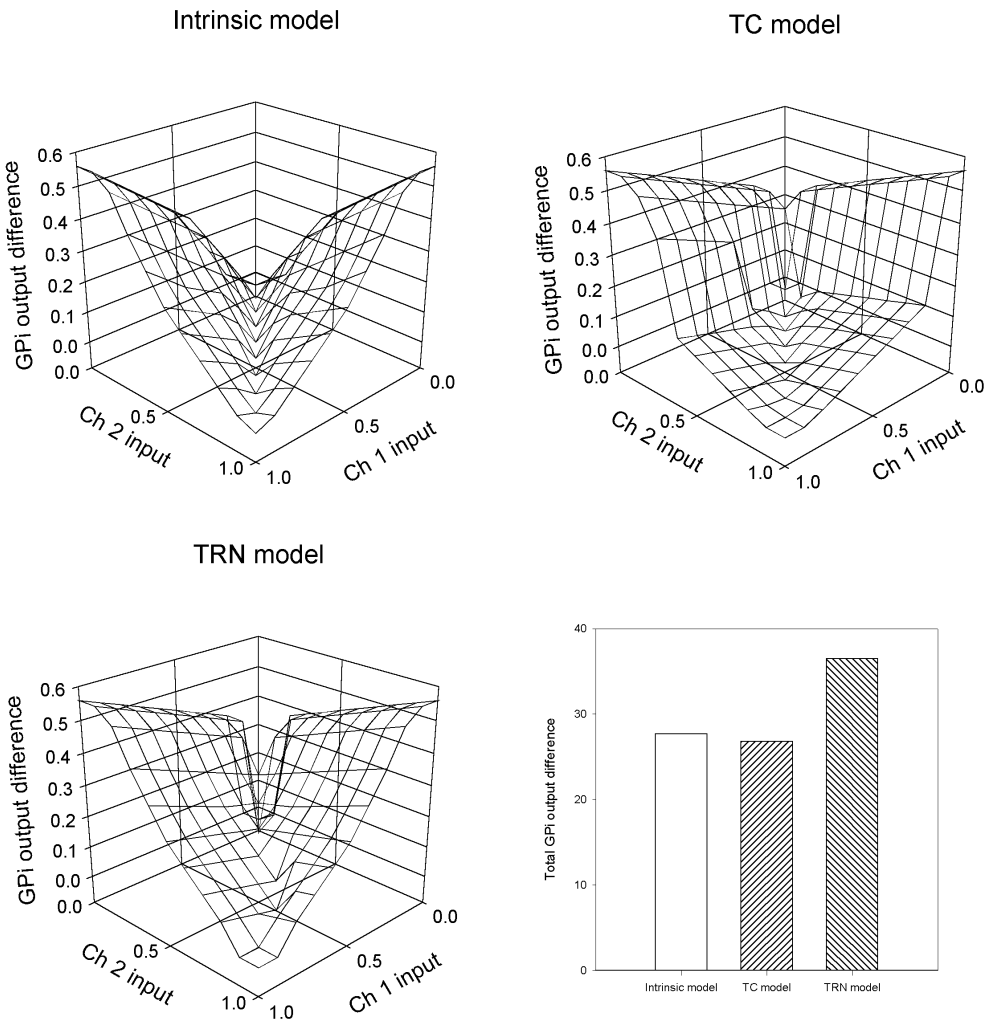

Figure 7. The changes in difference in GPi output between channels 1 and 2 after the onset of input to both channels (at equilibrium after $t = 2$). The TRN model has a greater or closely matched difference to the other models with almost any input pair combination. The histogram highlights the enhanced GPi output contrast of the TRN model, in comparison to the intrinsic and TC models.

Mesh plots of the GPi differential values for every simulation run on the three models (as detailed in the preceding section) are shown in figure 7. The intrinsic model showed small GPi output differences for low-level input pairs, and zero output differences for equal input pairs. The TC model, while showing large output difference for low-level pairs, showed small differences for all medium-to-high level pairs. By contrast, the TRN model showed relatively large output differences for all medium-level pairs and also had large differences for equal-level pairs. The histogram of total GPi output difference for each model (in figure 7) quantifies the improvement in GPi output contrast which is gained with the TRN model: the intrinsic and TC models were very similar with totals of 27.65 and 26.77, respectively, but the TRN model's total was 36.5.

4.5. Transient suppression

We noted in section 2.3 that short-term increases in salience on a non-selected channel should be suppressed so that the ongoing selection was not hampered or interrupted. To test the ability to suppress transient salience increases in non-selected channels, we ran 55 simulations on each model, one for each pair in which $S_2 > S_1$ (in the TC and TRN models, a transient salience increase is caused by an increase in sensory input). Three levels of transient increase were used: $0.5\Delta S$, ΔS and $1.5\Delta S$, where $\Delta S = S_2 - S_1$ (giving a total of 165 simulations/model). These levels correspond to transients that are less than, equal to, and greater than the level of sensory input on channel 2, respectively. An example of the stimulus used is shown in figure 5 (channel 1). The transient increase occurred on channel 1 at $t = 3$, and lasted for 1 time unit. Successful suppression required that $y_1^b > \theta_s(t > 3)$ and that, if $y_2^b \leqslant \theta_s(t = 3)$, then $y_2^b \leqslant \theta_s(t > 3)$.

The results of the simulations are shown in the plots in figure 8. We are only interested in the top left diagonal of the plots, as the effect of the salience increase on channel 1 was only tested when $S_2 > S_1$. The histogram shows the number of input pairs in each suppression category (no suppression, suppression up to $0.5\Delta S$, suppression up to ΔS, and suppression up to $1.5\Delta S$). The intrinsic model was capable of suppressing transients on channel 1 which were below the level of S_2 (i.e. $0.5\Delta S$) over the majority of input pairs (40/55). However, there was only one case ($S_1 = 0.6$, $S_2 = 1.0$) where an equal-level transient was suppressed, and no transients greater than the input on channel 2 were suppressed.

The TC model could suppress $0.5\Delta S$ transients at a lower level of input than the intrinsic model. It also showed a couple more cases of suppression with equal-level transients, and one case ($S_1 = 0.1$, $S_2 = 0.2$) of suppression of a $1.5\Delta S$ transient. However, the suppression abilities at low input levels cannot be directly compared to the intrinsic model as the intrinsic model was incapable of selection at these levels (see section 4.3). Furthermore, the TC model was capable of suppressing transients in the smallest number of cases (33) of all three models.

The TRN model could suppress transients over the greatest number of pairs (44), but this is partly due to its ability to do so with low-level inputs which are not comparable to the intrinsic model. More pertinently, the TRN model was able to suppress equal-level transients over a wide range of instances (21 cases) as well as a couple of $1.5\Delta S$ transients. This indicates that the addition of the TRN prevents the ongoing channel selection from being interrupted by a transient increase in input on another channel, the level of which is closely matched to the selected channel's input. Furthermore, it should be noted that the pairs which did not show suppression of equal transients had very high sensory input levels on channel 2. This may be a useful feature as it makes sense for a transient competitor to be able to interrupt the selected action if it is very urgent.

4.6. Closely matched saliences

A further aspect of persistence (discussed in section 2.3 with reference to the behavioural phenomenon of dithering) is that a selected action should not be interrupted by a closely matched competing action. That is, the ongoing action should persist. To test the ability of the models to perform this function, we ran simulations using closely matched sensory input pairs.

The inputs consisted of ten levels of channel 1 input ($c_1 = 0, 0.1, \ldots, 0.9$). Each level was matched with eleven levels of input on channel 2 that were equal to or larger than the channel 1 input. Thus, if $\Delta S_2 = S_2 - S_1$, then ΔS_2 had the values $0, 0.01, \ldots, 0.1$. Input to channel 1 began at $t = 1$; channel 2 salience input began at $t = 2$. Persistence of selection occurred if $y_1^b \leqslant \theta_s \wedge y_2^b > \theta_s(t > 2)$.

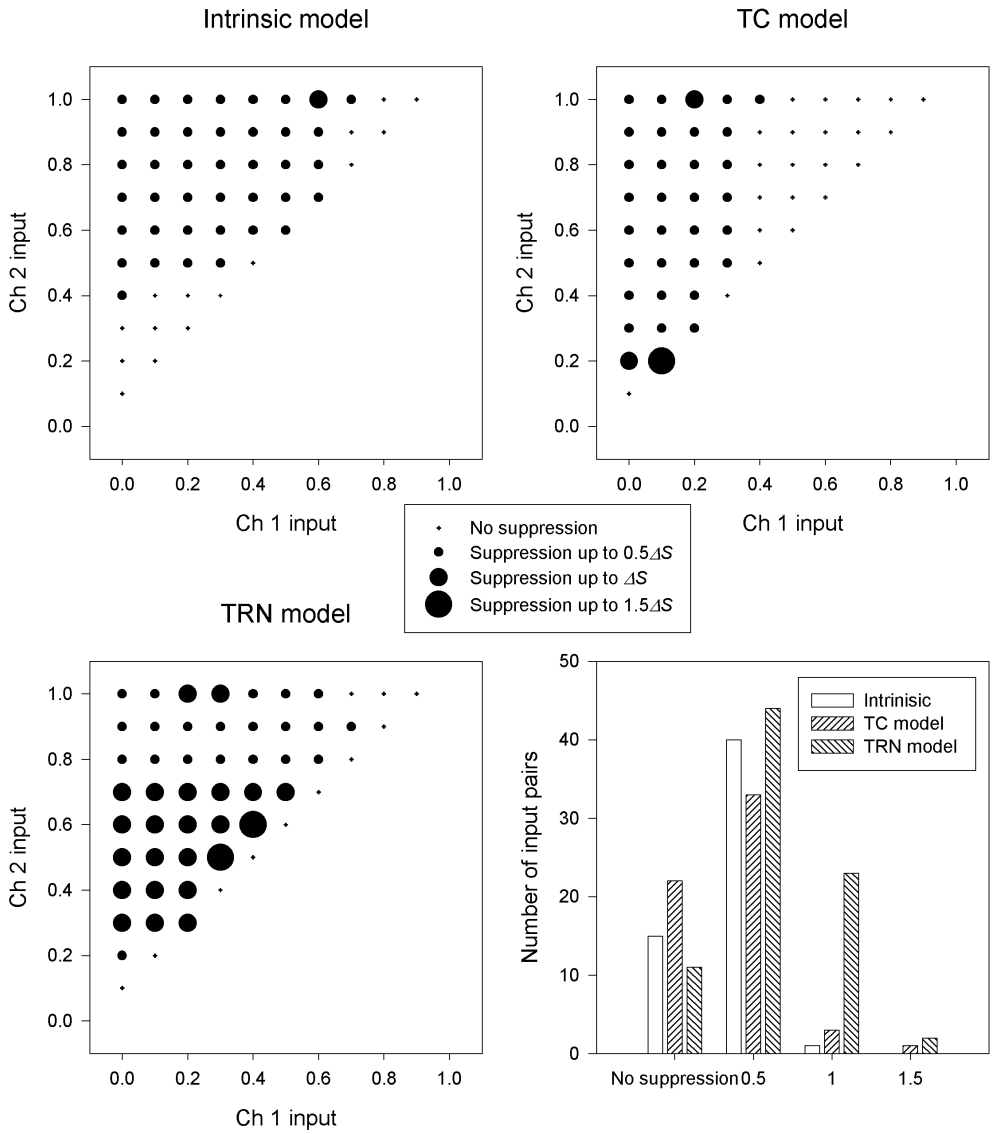

Figure 8. Transient suppression ability of the intrinsic, TC and TRN models. The intrinsic model could suppress transients in channel 1 which were lower than the channel 2 salience level for most values of channel 1 salience that were tested. The TRN model expanded on this capability by suppressing many transients which were equal to the salience input on channel 2.

The plots in figure 9 show the responses of the three models (in terms of the four output states detailed in section 4.3) to each of the input pairs. Persistence is, therefore, indicated by single-channel selection. The intrinsic model showed continued selection of channel 1 with greater levels of channel 2 input at $S_1 = 0.4, 0.5$. Similarly, the TC model only showed persistence of selection for two levels of S_1 (0.1 and 0.2). For most selection-capable values of S_1 the TC model showed dual channel selection for all values of ΔS_2. In contrast, the TRN model showed continued selection of channel 1 after channel 2 onset for six levels of S_1.

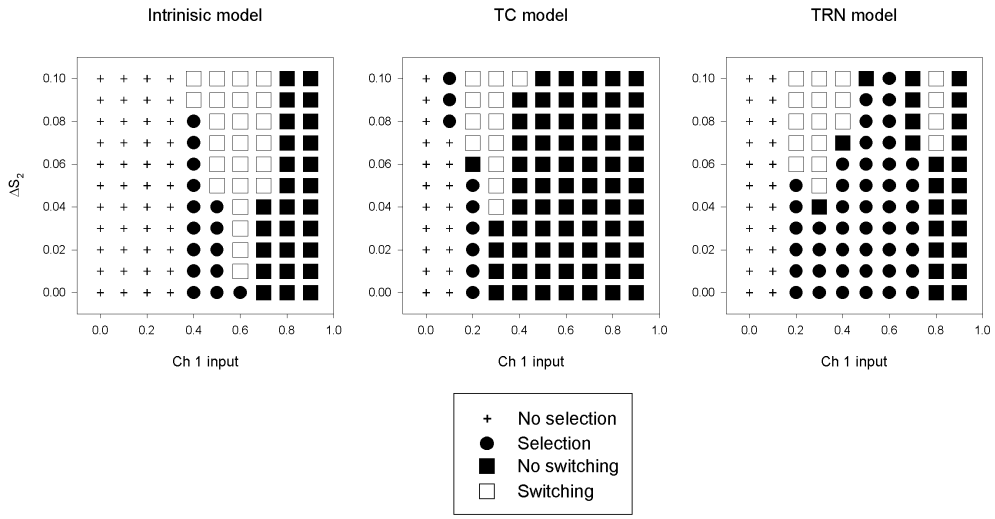

Figure 9. The relative ability of the intrinsic, TC and TRN models to maintain the selection of channel 1 after the onset of a closely matched salience on channel 2. The TRN model was able to maintain the selection of channel 1 for six levels of salience input.

This suggests that the TRN model is able to prevent the ongoing channel selection from being interrupted by a closely matched input on another channel for most selectable input values. It should be noted that, for the TRN model, there is not a monotonic relationship between S_1 and the level of ΔS_2 at which selection of channel 1 ceases, even though there is a suggestion of such a relationship in the intrinsic model.

4.7. Dopamine modulation of selection

We have explored the affect of changes in dopamine levels on the selection and switching capabilities of the intrinsic model (Gurney *et al* 2001b) and, as outlined in section 1.3, the behaviour of the model showed parallels with BG related dysfunctions. Here we investigate the affects of dopamine on the models that include the thalamocortical loop and TRN.

The central column of plots in figure 10 shows the selection and switching results for the three models that were described in section 4.3. The same simulation used to generate those results was re-run for each model using two different levels of dopamine: $\lambda_e, \lambda_g = 0$ and 0.6.

With no tonic dopamine (left-hand column of plots), all the models showed no selection or switching capability. High levels of tonic dopamine (right-hand column of plots) resulted in increased dual channel selection and a reduction in switching. This was particularly prevalent in the TC model which completely lost its switching ability. The TRN model showed the most robust response to high dopamine levels as it retained the greatest number of successful switches.

4.8. BG to TRN projections

In this paper, we have maintained the use of the GPi as the sole example of the BG output nuclei to be consistent with our previous papers (Gurney *et al* 2001a, 2001b). Although there is good evidence for a projection from SNr to the TRN (see section 2.2), there is only sparse

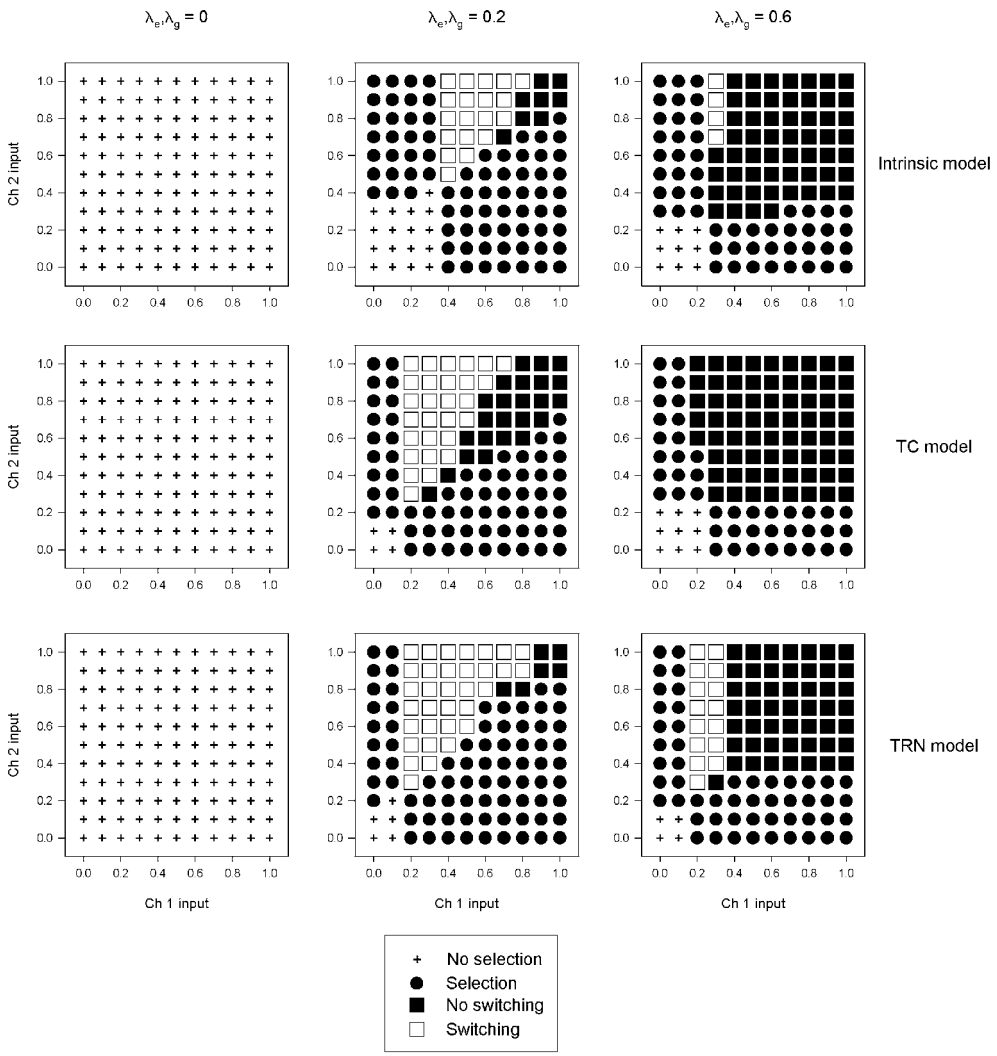

Figure 10. The effect on the selection and switching abilities of the three models due to changes in the tonic dopamine level. The centre column shows the normal responses, as illustrated in figure 6. Removing dopamine (left-hand column) resulted in no selection at any salience level in all three models. With high levels of dopamine the majority of salience inputs caused simultaneous channel selection. The TRN model was the most robust as it maintained the highest number of successful switching cases with increased dopamine.

evidence for a GPi to TRN projection (Moon Edley and Graybiel 1983). Numerous tracing studies which have looked at GPi/entopeduncular nucleus targets failed to report a projection to TRN (Kha *et al* 2000, Mengual *et al* 1999). However, as illustrated in figure 11, the removal of the GPi to TRN connection (by setting $w_{bg} = 0$) made no difference to the behaviour of the TRN model. Further, there is also good evidence for a GPe to TRN connection (Hazrati and Parent 1991, Gandia *et al* 1993), although it is not clear which TRN sectors this projection reaches. Regardless, adding this connection in place of the GPi to TRN connection and with the same weight (0.2) did not change the behaviour of the TRN model.

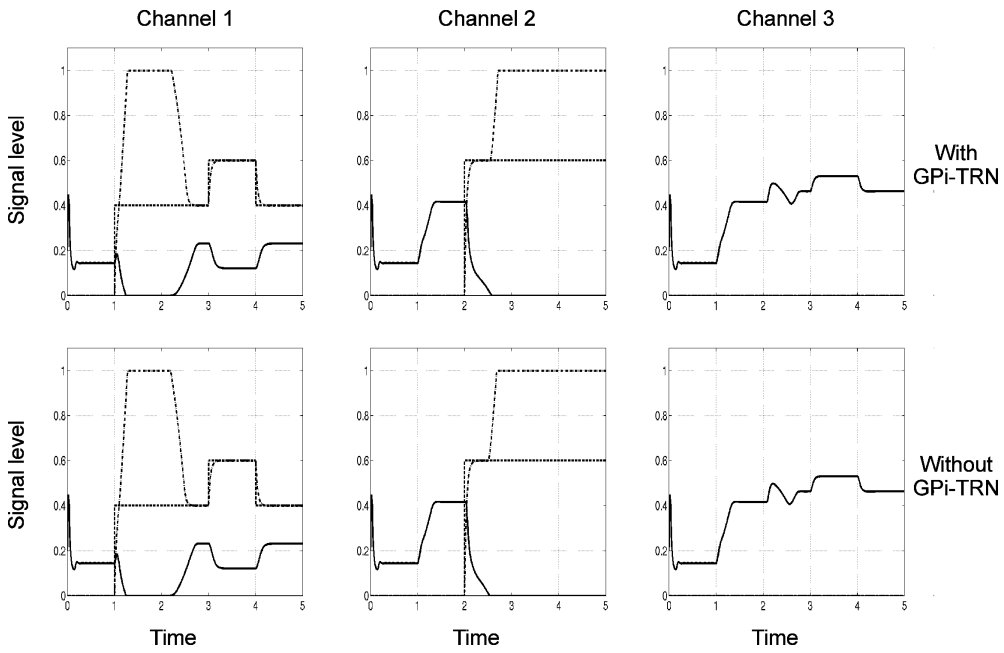

Figure 11. Variants of the TRN model. The top row shows the output of GPi and motor cortex with the GPi to TRN pathway intact. The bottom row shows that, with this pathway removed, there is no effect on the outputs. GPi: ——. Motor cortex: — · —. Sensory input: - - - -.

The ineffectiveness of the GPi–TRN and GPe–TRN pathways cannot necessarily be attributed simply to their low weight as there are other weak connections in the model which alter the model's behaviour if their weight is changed. For example, figure 12 shows the change in the selection and switching abilities of the TRN model when the within-channel TRN inhibitory connection to VL thalamus is removed ($w_T^* = 0$). Even though the normal strength of this connection is only 0.1, its removal substantially affected switching: there were six fewer successful switches, and three more cases of dual channel selection (no switching).

Thus, given that changing the within-channel TRN connection could have an affect on the BG output, we speculate that the ineffectiveness of either BG-to-TRN pathway on the BG output of active channels had two related causes which were not solely based on the low weights of the pathways. First, in a selected channel, the large excitatory input from motor cortex and VL thalamus to the TRN would have overwhelmed the small inhibitory input from the BG to the TRN. Second, in a non-selected channel, the TRN output would have no impact on the corresponding channel in VL thalamus, because that channel is already inhibited by GPi output, and the TRN output would have no impact on the selected channel for the reason outlined above.

5. Discussion

We have shown that the addition of a thalamocortical loop and appropriately designed TRN to a computational model of the BG has enhanced its selection capabilities. Furthermore, these additions to the model gave five main results which are desirable in behavioural terms: low sensory inputs (i.e. low salience actions) can become selected; the range of input values for

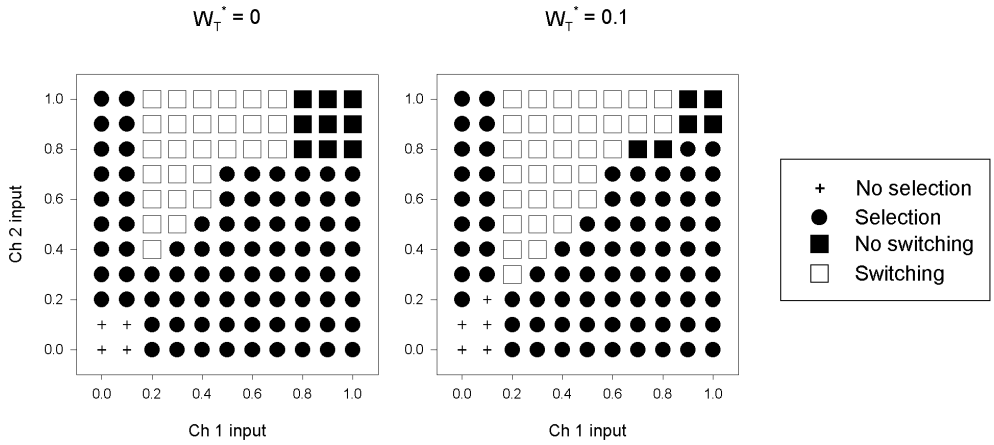

Figure 12. The removal of the TRN within-channel inhibition of VL thalamus, in the TRN model, caused minor changes in the switching and selection patterns. This illustrates that this pathway has a discernible affect on the model's behaviour, even though it has a very low magnitude weight.

which switching can take place has increased; responses to large transient events on a non-selected channel are suppressed; inputs which have closely matched levels to the currently selected input are also suppressed; and the GPi output contrast between selected and non-selected channel is enhanced. Finally, we have shown that the TRN model possesses all of the desirable characteristics of a switching mechanism that were outlined in the introduction. We discuss these features in greater detail below.

Although substantial use of the three different models was made in this paper, our intention was not to use the intrinsic or TC models as 'straw men' against which we could compare the improved abilities of the TRN model. Our overriding aim was to use the three different models to show how the addition of new, anatomically constrained, connections and structures could add new abilities while maintaining existing selection and switching abilities. By decomposing the circuit additions into two phases we were able to determine which additional component was responsible for which ability.

5.1. Thalamocortical loop

Analysis and simulation of the intrinsic model had already shown that the intrinsic connectivity of the BG could produce output signals (from GPi) which were consistent with action selection. We have shown here that, with the addition of a thalamocortical loop, low-level sensory inputs (representing low salience actions) also now result in output signals which are consistent with the selection of an action. Low salience selection may be behaviourally important because, when an animal is faced with numerous relatively unimportant choices of subsequent action, it must be able to pick and execute one of them. The alternative is to remain in a quiescent state until an event of sufficiently large salience occurs. The minimum sensory input level required for selection (0.2) in the TC and TRN models corresponds to the value of ϵ which modelled the threshold for inducing an UP-state transition in a striatal neuron. Thus, as noted in the introduction, very low-level inputs may be filtered out by the striatum, which prevents pure noise from having an affect on action selection. As previously noted, the ability of the TRN model to successfully resolve competitions over a wide range of sensory input values shows that it has the decisive aspect of clean switching.

We attribute the low input selection to the positive feedback loop formed by VL thalamus and motor cortex. Amplification of the sensory input in this loop means that the striatum receives a greater input than just two identical copies of the sensory signal. However, this thalamocortical feedback is also responsible for the simultaneous channel selection seen at many input levels in the TC model. The positive feedback loop causes motor cortex output to saturate and, when added to the sensory input in striatum, the combined total forces the BG to select both actions. Clean switching was recovered by the introduction of the TRN.

5.2. A functional role for the thalamic reticular nucleus

Besides clean switching, the TRN model also had the desirable characteristics of persistence and absence of distortion. It was able to prevent interruption of an ongoing selection by a closely matched competitor, to suppress a wide range of transient salience events, and to enhance GPi output contrast. These abilities can all be attributed to the presence of the TRN, as they were not as evident in the other two models. In addition, the low-level input selection capability of the TC model was maintained, while also improving on its selection capabilities.

A possible explanation for these abilities centres on the role of the between-channel inhibition. Losing competing inputs are suppressed because the between-channel inhibition prevents these channels from increasing their activity through the thalamocortical positive feedback loop. This could account for the TRN model's switching and persistence abilities. Transients and closely matched competing inputs would also be prevented from causing an increase in activity on that channel through the thalamocortical loop. Therefore, the input to striatum on the transient's/competitor's channel would remain at a lower level than on the currently selected channel. It is also clear that the TRN acted to help switching by reducing the level of activity in the thalamocortical loop of the deselected channel (e.g. the reduction of motor cortex output; see figure 5). The within-channel inhibition may also have influenced the switching ability, as removing this connection reduced the number of successful switching cases.

Thalamocortical feedback and the between-channel inhibition could account for the improved GPi output contrast. The higher levels of input to striatum (when compared to the intrinsic model) from the combined somatosensory and motor cortex input drive the STN to excite GPi, leading to elevated GPi output levels. On the winning channel, the high levels of cortical input specific to that channel enables the striatal inhibition to exceed this massive excitatory drive and cause the channel to become selected. The prevention of feedback in the thalamocortical loop by the action of the TRN ensures that the other channels in GPi receive no major inhibitory input, leading to enhanced contrast between the selected and non-selected channels.

The previous discussion highlights two functional roles for the TRN in the model, which can be encapsulated as follows. First, the TRN acts as a gain control and effectively sets a ceiling level of total activation in the thalamocortical loop. Second, the TRN acts to 'clean-up' BG output, allowing the selection of, and the switching between, a wider range of input levels. Thus, the results suggest that the TRN may be another selection mechanism whose action is complementary to that of the BG.

The findings presented here are also consistent with Crick's (1984) searchlight hypothesis, as the action of the TRN on focusing selection could be interpreted as the focusing of attention on that channel.

5.3. The action of dopamine

We have also shown that setting the level of tonic dopamine to extreme values was detrimental to the selection and switching abilities of all three models. The removal of dopamine completely

prevented selection of any input, which is consistent with the difficulty of voluntary movement in Parkinsonian patients. Excess dopamine increased the promiscuity of selection, as all three models showed simultaneous channel selection with the majority of inputs. The TRN model was the more robust (with high dopamine values) as it was capable of successful switching in a substantial number of cases and also maintained its low-level input selection capability.

Work using a more detailed model neuron has demonstrated that dual channel selection in the mean-field models (i.e. those presented in this paper) may be equivalent to rapid alternation of selection (Humphries and Gurney 2001b). Thus, as noted in the introduction, widespread simultaneous channel selection may underly disorders characterized by rapid behavioural changes such as ADHD.

5.4. Modelling issues and further work

There are two major issues specific to the modelling work presented in this paper that need to be addressed. First, the VL thalamus is modelled identically to all other nuclei in the circuit, and its output is continuously variable between 0 and 1. This fails to capture the well known two-state output of thalamocortical neurons: tonic firing mode (which is effectively what is modelled here), and burst firing mode (Sherman 2001). It has been recently demonstrated that burst firing in thalamocortical neurons can take place *in vivo* in awake animals (Swadlow and Gusev 2001). Hitherto, burst firing in thalamocortical neurons had been mostly observed *in vitro* (Kim *et al* 1997, Kim and McCormick 1998) or in sleeping animals. The functional significance of this mode of firing is unclear, although it is possibly a robust 'wake-up' signal to the cortex. To study the effect of two-state firing in the thalamus on the behaviour of the TRN model, it would be necessary to construct a lower level model which made explicit use of membrane properties. Initial work on developing such a model has been carried out, and it has been tested in a detailed model of the STN-GPe loop (Humphries and Gurney 2001a).

The second major issue is that the model is primarily constrained by data from rat based studies. For example, there are no interneurons modelled in the VL thalamus because the rat motor thalamus is devoid of them (Sawyer *et al* 1991). However, the primate motor thalamus has an extensive network of interneurons (Jones 1985). The rat-oriented nature of the model was intentional for two reasons. First, the data available on the rat BG and thalamocortical connections is extensive and, second, the model presented here has been adapted for use in a *Khepera* mobile robot which is explicitly designed to mimic behaviour patterns specific to the rat (Montes-Gonzalez *et al* 2001).

6. Conclusions

The models presented here have shown that the addition of anatomically constrained extrinsic pathways to a computational model of the BG improves the previously demonstrated selection abilities and brings new features. The positive feedback loop formed by the motor cortex and VL thalamus allowed low salience actions to be selected. The addition of the TRN made the model capable of fulfilling the desirable characteristics of a switching mechanism. We conclude, therefore, that our models provide evidence that the TRN-dorsal thalamic complex and that the thalamocortical loop have a functional role in action selection.

Acknowledgments

M D Humphries is supported by a studentship from the EPSRC, UK. We would like to thank Peter Redgrave, Tony Prescott and Tom Stafford for many illuminating discussions and for their support. M D Humphries would also like to thank Nick Davis for further discussion and suggestions.

References

Adams N C, Lozsádi D A and Guillery R W 1997 *Eur. J. Neurosci.* **9** 204–9
Albin R, Young A and Penney J 1989 *Trends Neurosci.* **12** 366–75
Bevan M D, Francis C M and Bolam J P 1995 *J. Comp. Neurol.* **361** 491–511
Brown L L and Sharp F R 1995 *Brain Res.* **686** 207–22
Brown L L, Smith D M and Goldbloom L M 1998 *J. Comp. Neurol.* **392** 468–88
Chevalier G, Vacher S, Deniau J and Desban M 1985 *Brain Res.* **334** 215–26
Cornwall J, Cooper J D and Phillipson O T 1990 *Exp. Brain. Res.* **80** 157–71
Crick F 1984 *Proc. Natl Acad. Sci. USA* **81** 4586–90
Delgado A, Sierra A, Querejeta E, Valdiosera R F and Aceves J 2000 *Neuroscience* **95** 1043–8
Diamond M E 1995 *Cerebral Cortex* vol 11, ed E G Jones and I T Diamond (New York: Plenum) pp 189–219
Farkas T, Kis Z, Toldi J and Wolff J-R 1999 *Neuroscience* **90** 353–61
Flaherty A W and Graybiel A M 1994 *J. Neurosci.* **14** 599–610
Friedberg M H, Lee S M and Ebner F F 1999 *J. Neurophysiol.* **81** 2243–52
Fujimoto K and Kita H 1993 *Brain Res.* **609** 185–92
Gandia J A, De Las Heras S, Garcia M and Giménez-Amaya J M 1993 *Brain Res. Bull.* **32** 351–8
Gerfen C, Engber T, Mahan L, Susel Z, Chase T, Monsma F and Sibley D 1990 *Science* **250** 1429–32
Gerfen C and Wilson C 1996 *Handbook of Chemical Neuroanatomy: Integrated Systems of the CNS* vol 12, ed L Swanson *et al* part III (Amsterdam: Elsevier) pp 371–468
Gonon F 1997 *J. Neurosci.* **17** 5972–8
Guillery R W, Feig S L and Lozaádi D A 1998 *Trends Neurosci.* **21** 28–32
Gurney K, Prescott T J and Redgrave P 2001a *Biol. Cybern.* **85** 401–11
Gurney K, Prescott T J and Redgrave P 2001b *Biol. Cybern.* **85** 411–23
Gurney K, Redgrave P and Prescott A 1998 *Technical Report* vol AIVRU131 Department of Psychology, University of Sheffield
Hall R D and Lindholm E P 1974 *Brain Res.* **66** 23–38
Hazrati L-N and Parent A 1991 *Brain Res.* **550** 142–6
Hernández-López S, Bargas J, Surmier D J, Reyes A and Galarraga E 1997 *J. Neurosci.* **17** 3334–42
Hoogland P V, Welker E and Van der Loos H 1987 *Exp. Brain. Res.* **68** 73–87
Hoover J E and Strick P L 1993 *Science* **259** 819–21
Hoover J E and Strick P L 1999 *J. Neurosci.* **19** 1446–63
Houston A and Sumida B 1985 *Anim. Behav.* **33** 315–25
Humphries M D and Gurney K N 2001a *Neural Netw.* **14** 845–63
Humphries M D and Gurney K 2001b *British Neuroscience Association Abstracts* vol 16 p 71
Izraeli R and Porter L L 1995 *Exp. Brain Res.* **104** 41–54
Jones E G 1985 *The Thalamus* (New York: Plenum)
Kelley A E, Domesick V B and Nauta W J H 1982 *Neuroscience* **7** 615–30
Kha H T, Finkelstein D I, Pow D V, Lawrence A J and Horne M K 2000 *J. Comp. Neurol.* **426** 366–77
Kim U and McCormick D A 1998 *J. Neurosci.* **18** 9500–16
Kim U, Sanchez-Vives M V and McCormick D A 1997 *Science* **278** 130–4
Kolmac C I and Mitrofanis J 1998 *J. Comp. Neurol.* **396** 531–43
Lozsádi D A 1994 *J. Comp. Neurol.* **341** 520–33
Mengual E, de la Heras S, Erro E, Lanciego J L and Giménez-Amaya J M 1999 *J. Chem. Neuroanat.* **16** 187–200
Miyashita E, Keller A and Asanuma H 1994 *Exp. Brain Res.* **99** 223–32
Montes-Gonzalez F, Prescott T J, Gurney K, Humphries M and Redgrave P 2001 *From Animals to Animats 6: Proc. 6th Int. Conf. on Simulation of Adaptive Behaviour* ed J-A Meyer *et al* pp 157–66
Moon Edley S and Graybiel A M 1983 *J. Comp. Neurol.* **217** 187–215
Newman J, Baars B J and Cho S-B 1997 *Neural Netw.* **10** 1196–1206
Ni Z-G, Bouali-Benazzouz R, Gao D-M, Benabid A-L and Benazzouz A 2001 *Brain Res.* **899** 142–7

Onla-or S and Winstein C J 2001 *Brain Res. Cogn. Brain Res.* **10** 329–32
Parent A and Cicchetti F 1998 *Mov. Disord.* **13** 199–202
Parent A, Lévesque M and Parent M 2001 *Parkinsonism and Related Disorders* **7** 193–8
Pinault D and Deschênes M 1998 *Eur. J. Neurosci.* **19** 3462–9
Prescott T, Redgrave P and Gurney K 1999 *Adapt. Behav.* **9** 99–127
Price J L 1995 *The Rat Nervous System* 2nd edn, ed G Paxinos (New York: Academic) pp 629–48
Redgrave P, Prescott T J and Gurney K 1999 *Neuroscience* **89** 1009–23
Sawyer S F, Martone M E and Groves P M 1991 *Neuroscience* **42** 103–24
Sherman S M 2001 *Nat. Neurosci.* **4** 344–6
Shosaku A, Kayama Y and Sumitomo I 1984 *Brain Res.* **311** 57–63
Sumitomo I and Iwama K 1987 *Brain Res.* **415** 389–92
Swadlow H A and Gusev A G 2001 *Nat. Neurosci.* **4** 402–8
Swanson J, Castellanos F, Murias M, LaHoste G and Kennedy J 1998 *Curr. Opin. Neurobiol.* **8** 263–71
Turner R S and DeLong M R 2000 *J. Neurosci.* **20** 7096–108
Umemiya M and Raymond L A 1997 *J. Neurophysiol.* **78** 1248–55
Uno M, Ozawa N and Yoshida M 1978 *Exp. Brain Res.* **33** 493–507
Weese G D, Phillips J M and Brown V J 1999 *J. Neurosci.* **19** 10 135–9
Wichmann H, Bergman H and DeLong M 1994 *J. Neurophysiol.* **72** 494–506
Zarzecki P 1991 *Somatosens. Mot. Res.* **8** 313–25