Biologically constrained action selection improves cognitive control in a model of the Stroop task

Tom Stafford* and Kevin N. Gurney

Adaptive Behaviour Research Group, Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK

The Stroop task is a paradigmatic psychological task for investigating stimulus conflict and the effect this has on response selection. The model of Cohen *et al.* (Cohen *et al.* 1990 *Psychol. Rev.* **97**, 332–361) has hitherto provided the best account of performance in the Stroop task, but there remains certain key data that it fails to match. We show that this failure is due to the mechanism used to perform final response selection—one based on the diffusion model of choice behaviour (Ratcliff 1978 *Psychol. Rev.* **85**, 59–108). We adapt the model to use a selection mechanism which is based on the putative human locus of final response selection, the basal ganglia/thalamo-cortical complex (Redgrave *et al.* 1999 *Neuroscience* **89**, 1009–1023). This improves the match to the core human data and, additionally, makes it possible for the model to accommodate, in a principled way, additional mechanisms of cognitive control that enable better fits to the data. This work prompts a critique of the diffusion model as a mechanism of response selection, and the features that any response mechanism must possess to provide adaptive action selection. We conclude that the consideration of biologically constrained solutions to the action selection problem is vital to the understanding and improvement of cognitive models of response selection.

Keywords: Stroop; response selection; action selection; diffusion model; basal ganglia

1. INTRODUCTION

The Stroop task provides a thoroughly explored experimental framework for investigating cognitive aspects of selection. In this task, subjects have to name the ink colour of word strings which can themselves spell out the name of a colour. When the ink colour and the word name contradict each other response selection is slowed and is more prone to error (compared to conditions where the word name is neutral or complementary with respect to the ink colour). This is 'the Stroop effect'. A simple reversal of the task, that of reading the word name and ignoring the ink colour, does not produce an opposite effect (a 'reverse Stroop' effect).

The asymmetrical interaction of the colour- and word-naming processes can be interpreted within an automaticity framework (Posner & Snyder 1975; MacLeod 1991). Here, word reading is an 'automatic', or 'overlearnt', response which is triggered on stimulus presentation and difficult to interrupt, and colour naming is a controlled process which is not automatic and is liable to interference from word reading. Variations on the basic Stroop task have been successful in clarifying the nature of automatic processing (Besner *et al.* 1997; Besner & Stolz 1999; Dishon-Berkovits & Algom 2000; Durgin 2000).

Here, however, we wish to focus on the Stroop task as defining a process of selection. The Stroop task has a long history of use in the investigation of aspects of response selection at a cognitive level (MacLeod 1991) and, more recently, at the neural level (MacLeod & MacDonald 2000). In particular, while early processing of stimulus information is clearly important to an understanding of the Stroop task, the final response uttered on each trial is subject to the constraints imposed by a response or decision mechanism, which must translate internal cognitive states into motor action.

Much progress has been made in investigating decision mechanisms in simple two-alternative choice tasks. Mathematical models of such simple decisions are able to accurately predict the patterns of reaction times (RTs) and errors across task variations, and there is a considerable history in psychology of their development and refinement (Luce 1986; Ratcliff & Smith 2004). More recently, it has been possible to connect these models with neurophysiological data (Ratcliff *et al.* 2003; Reddi *et al.* 2003) and with an information theoretic foundation for optimal decision making (Bogacz *et al.* 2007*a*,*b*). These developments promise an exciting period of cross-fertilization between neuro-biological and psychological perspectives on simple decisions (Platt 2002; Smith & Ratcliff 2004; Opris & Bruce 2005). The current work investigates how one instance of this class of model serves selection in a model of the more complex Stroop task.

An additional perspective on decision making is supplied by workers in neuroscience, animal behaviour, ethology and robotics who have defined, and explored solutions to, the problem of *action selection*: the resolution of conflicts between competing requests for

behavioural expression through a final common motor path (Redgrave *et al.* 1999).

The aim of the present work was to determine whether a biologically plausible model of the putative locus of action selection in humans (the basal ganglia) could work as the response mechanism in a model of a cognitive task (the Stroop task). Therefore, this is a first step in making links between possible neural substrates for action selection, neural correlates of decision making and cognitive processes of selection.

2. MODELLING THE STROOP TASK

In a seminal paper, Cohen *et al.* (1990) described a model of processing in the Stroop task and variants.

This simple connectionist model (hereafter 'The Cohen model') involves the translation of a localist input representation into a response representation, via a feedforward two-layer network trained with standard backpropagation. The architecture is shown in figure 1. The main features of this network are as follows:

(i) Differential training of the network: word inputs are presented at ten times the frequency of colour inputs during training. This results in a stronger weighting of signals representing this aspect of the stimulus.

(ii) Attentional sensitization: the network implements attention as an additional input which offsets a bias (in effect a default inhibition) on all hidden units. This interacts with the sigmoidal output function of the units so that moderately sized colour- or word-input signals do not result in a commensurate increase in output, unless presented in combination with attentional input. Signals in the word-processing pathway are, however, large enough to partially overcome the default inhibition without the aid of attentional input.

(iii) RTs are generated by the dynamics of a response mechanism that works with evidence accumulation: the two output units of the network are taken to indicate, at each time-step, the evidence favouring each response. This evidence is compared and the difference accumulated, until the total crosses a threshold, when a response is said to have been made.

The current work focuses on the third element above: the response mechanism and its role in determining overall model behaviour.

The Cohen model matches the basic Stroop data very well (figure 2). Not only does the model capture the quantitative difference that word reading is faster than colour naming, and unaffected by the word information, but it also matches the asymmetry between the size of the interference effect (the slowing of colour naming due to contradictory word information) and the facilitation effect (the speeding of colour naming due to compatibility with the word information). All the simulations presented, both of our model and our replication of Cohen's model, are shown run without added noise since this does not affect the mean results.
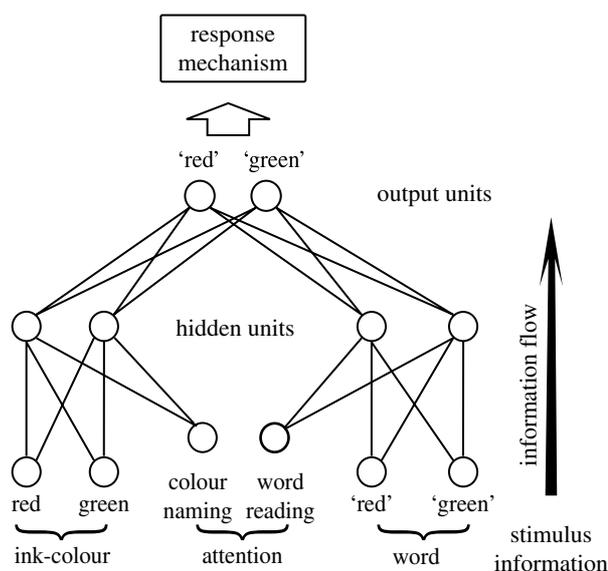
Figure 1. Architecture of the Cohen model.

In addition to matching the fundamental data, the model gives an implementational definition of automaticity: automaticity arises from greater strength of processing. In a connectionist framework, this means stronger weightings between stimulus and response (as in the Cohen model), or additional connections between modules involved in stimulus response translation (as in other connectionist models of Stroop processing, Phaf *et al.* 1990; Zhang *et al.* 1999). Either way, the implication is that there is no sharp dichotomy between 'automatic' processes and 'controlled' processes, and, additionally, that other quantitative differences, such as response time differences, arise out of this single fundamental mechanistic difference.

A plausible alternative theory of Stroop processing—and of automatic processing in general—is that more automatic processes are those in which pre-response processing is faster. This theory suggests that Stroop interference is due to the response evoked by the (contradictory) word element of the stimulus arriving at some response bottleneck earlier, creating slower selection of the opposite (and correct) response when it arrives there (we can posit that in a connectionist network this would be reflected by faster transference times between stimulus input and the model response mechanism which arbitrates action selection). This theory may be tested in so-called stimulus-onset asynchrony (SOA) experiments. These involve the two elements of the conventional Stroop stimulus, the word and colour, being presented asynchronously. This manipulation allows either element to appear before the other and thus, it is assumed, to be given a 'headstart' in perceptual processing.

The experimental data are shown in figure 3a. By convention SOAs which involve the to-be-ignored element being presented first are labelled negative. Clearly, no amount of headstart for colour information (i.e. at negative SOAs) allows it to interfere with word reading (Glaser & Glaser 1982), demonstrating that the automaticity of word reading is not a consequence of enhanced speed of processing. For colour naming, the word element causes interference
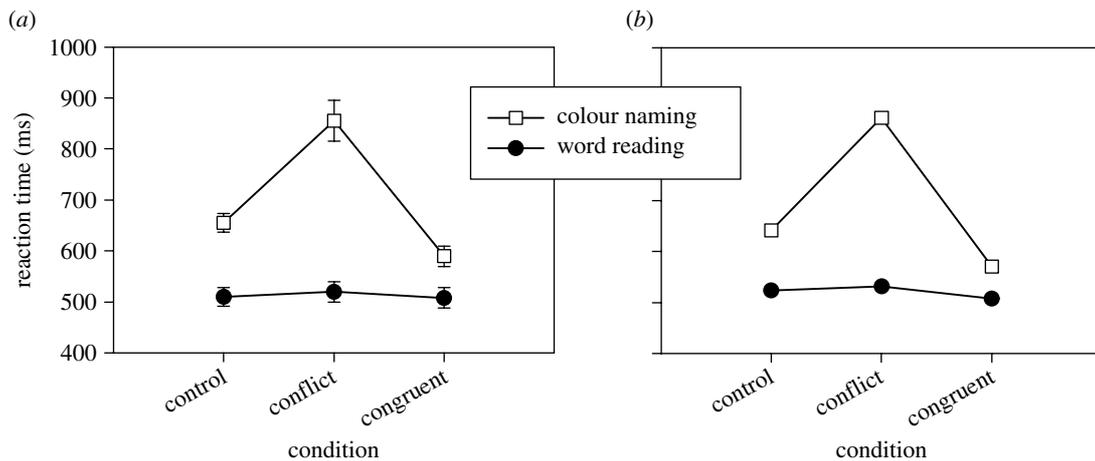
Figure 2. The fundamental pattern of RTs in the basic Stroop tasks. There are two tasks: word-naming (filled circles) and colour naming (unfilled squares). Within each task, there are three possible conditions; in the congruent condition word and the colour agree, in the conflict condition the word and the colour disagree, in the control condition the irrelevant element is neutral with respect to the target. (*a*) Empirical data from Dunbar & MacLeod (1984) for which standard error bars are shown, and (*b*) simulation data from replication of model of Cohen *et al.* (1990).

if it appears at any point before colour processing is finished (up to 300 ms after the appearance of the colour element—close to the asymptotic limit for RTs). Additionally, the appearance of the word before the colour always causes interference, however, long the subject is given to accommodate to the presence of the word. This and other results which contradict the automaticity as speed-of-processing account (Dunbar & MacLeod 1984), leave the automaticity as strength-of-processing account more preferable (this is not to say that strength-of-processing accounts do not imply that automatic processes will be faster than controlled processes—they do—rather they merely assert that speed of processing is a by-product of a more fundamental distinction between the two types of processes rather than being causative in itself).

The original simulation data for the SOA manipulation within the Cohen model are shown in figure 3*b*. The Cohen model simulates the correct relative ordering of the RTs in all conditions with respect to the empirical data. Cohen *et al.* (1990) note some discrepancy between their simulation and the model—firstly, that in the simulations colour information does interfere with word reading, albeit marginally, and that, secondly, the influence of word information on colour reading is not reduced but increases for SOAs before −200 ms. These discrepancies would not contradict the empirical data, and hence a strength of processing account, if the size of these effects was limited to that shown over the range originally tested by Cohen *et al.* (1990). The size of the interference and facilitation effects are not, however—as Cohen *et al.* (1990) suggest—asymptotic with increasingly negative SOA (as shown below, figure 4). Given this, it is of concern that the primary model of the automaticity-based account of Stroop processing, the Cohen model, is not able to simulate the primary data which falsifies the speed-of-processing account but, instead, produces a pattern of RTs which would, if true, appear to validate a speed-of-processing account.

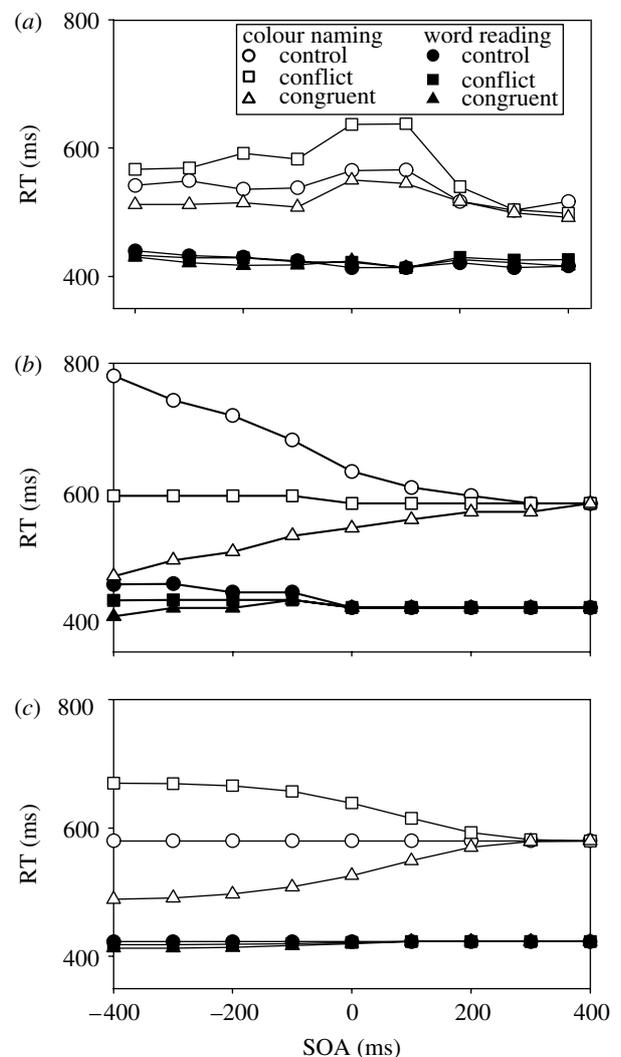Figure 3. Empirical data, (*a*), for the Stroop tasks with stimulus-onset-asynchrony (adapted from Glaser & Glaser 1982). Simulation data: (*b*), replication of Cohen *et al.* (1990). (*c*) Using the basal ganglia response mechanism. Negative SOAs represent the irrelevant element of the stimulus appearing before the target element, positive SOAs represent the target element appearing before the irrelevant element.
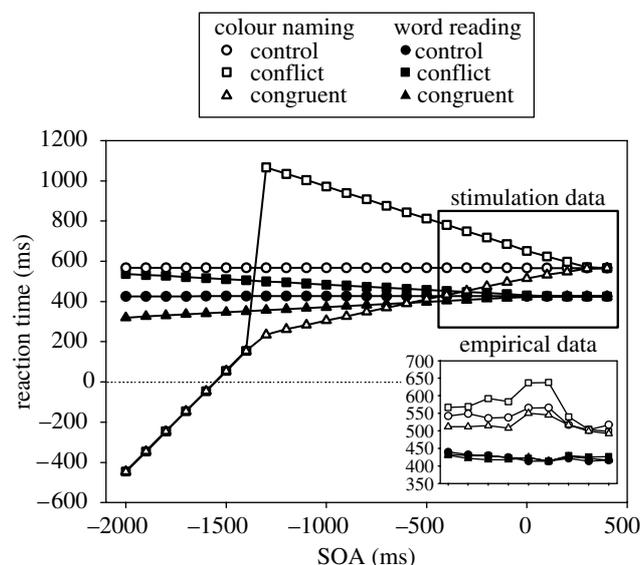
Figure 4. The SOA simulation of the Cohen model. The original range of data shown by Cohen *et al.* (1990) is demarcated by the box. The simulation data corresponds roughly to the empirical data over the this range (−400 to +400 ms) but beyond that diverges.

(a) *Limitations of the Cohen model with stimulus-onset asynchrony are failures of response selection*

Our replication of Cohen *et al.*'s (1990) model shows that, beyond the range of data of SOA values they originally present, the trends visible in the original data continue so that the model behaves inconsistently with the strength-of-processing account and consistently with the experimentally disproved speed-of-processing account of automaticity in the Stroop task (figure 4).

Consider the change in the simulated RTs as SOA gets more negative—as the to-be-ignored element of the stimulus appears increasingly before the to-be-responded-to element. For the colour-naming task in the conflict condition, the model RT increases as the word element slows selection based on the colour. Eventually, beyond −1300 ms, the word is presented early enough to prompt a response on its own. This response will be an incorrect one, since in the conflict condition the word is opposite to the colour. RTs now start to decrease with increasingly negative SOAs because the RT is defined as the time between the onset of the to-be-responded-to stimulus element and the occurrence of selection. Hence, RT eventually falls below zero because selection occurs *before* the onset of the colour (this is highlighted in figure 4 by the point at which the RT lines cross the dotted line representing zero on the RT axis). If the word is congruent to the colour information then there is comparable interference, but this reveals itself as a speeding of the correct response (which likewise falls below zero RT beyond −1500 ms). Note that from here the time between the onset of the irrelevant element and selection is constant, so that beyond this point the rate of the decrease in RT becomes a function of the decrease in SOA, not of changes in model output.

For the same fundamental reasons, in the word-naming task, the conflict and congruent conditions diverge in the same way (albeit over a longer time span, the point at which word-reading times fall below zero SOA is not shown here). Thus, the model behaves in accordance with the experimentally disproved speed-of-processing account: presenting colour information ahead of word information creates a reverse Stroop effect—colour information interferes with word reading. This is surprising, not least because the stated purpose of the model was to validate a strength of processing account.

Here, we trace this flaw to the response mechanism used in the model. Cohen's model of Stroop processing explicitly draws on the choice behaviour literature (Luce 1986) and adopts an exact analogue of the diffusion model (Ratcliff 1978; Ratcliff & Smith 2004) to resolve the response selection problem presented by the Stroop task. In the diffusion model, the balance of evidence regarding the two possible responses at each point in time is used to adjust a running total. The momentary balance of evidence is defined by the strength of evidence in favour of one response minus the strength of evidence in favour of the other. At each time-step, the change in the running total is drawn from a normal distribution with a mean defined by the balance of evidence (in this case, this is the difference between the output units of the connectionist front-end). When this total, which reflects the accumulated evidence, crosses either a positive threshold (indicating selection of one response) or a negative threshold (indicating selection of the other response) selection occurs. The diffusion model has been shown to be an analytically tractable form of several connectionist models of decision making, and an optimal decision algorithm for a two-choice decision situation (Bogacz *et al.* 2007*a*) where either desired accuracy or time-to-decision is specified (obviously these two mutually constrain each other). Further, potential neurobiological correspondences to the evidence accumulation processes of the diffusion model have been identified (Gold & Shadlen 2000; Ratcliff *et al.* 2003; Reddi *et al.* 2003).

The diffusion model response mechanism takes the outputs of the connectionist 'front-end' of the Cohen model as inputs. Because the model, like all connectionist models, works on graded signals, there is always some input due to the to-be-ignored stimulus, even if this is very small due to the attentional inhibition. In the case of the colour-naming task, it is integral to the model's function that some influence of the word element of the stimulus survives attentional selection and comes to influence the response stage. Without this feature, the basic effect of Stroop interference would not be present. However, in SOA conditions, this influence of the to-be-ignored element may accumulate indefinitely. This affects selection time to an extent proportional to the time it is presented multiplied by the strength of evidence conveyed. Hence, arbitrarily small amounts of evidence can provoke erroneous selection if presented for long enough, or they can massively slow correct selection (because accumulated evidence for the opposite response must be overcome).

The fact that Cohen *et al.*'s (1990) model involves a response mechanism is ignored in textbook treatments of the model (Sharkey & Sharkey 1995; Ellis & Humphreys 1999; O'Reilly & Munakata 2000) and

even overlooked in Cohen *et al*.'s (1990) own analysis of the function of the model. We argue that this reflects a regrettable, but not untypical, neglect of the action selection problem in psychology. Reinforcing this view, we have recently shown how, contrary to the original account of Cohen *et al*. (1990), it is the response mechanism, not the neuronal transfer function, which generates the important differences in RTs between conditions (Stafford & Gurney 2004), and it is the response mechanism which explains the asymmetry in the magnitudes of the interference and facilitation effects in the Cohen model (a matter about which there has been some debate, MacLeod & MacDonald 2000).

In summary, our investigation of evidence accumulation as a mechanism of selection in the Cohen model of the Stroop task will have general implications for theories of selection. The core element in this investigation is to show how a more biologically realistic response mechanism—a model of action selection in the basal ganglia—overcomes the deficiencies noted here.

3. THE BASAL GANGLIA AND THALAMIC COMPLEX AS A RESPONSE MECHANISM IN A COGNITIVE TASK

The basal ganglia are a set of subcortical nuclei that have been implicated in a range of motor and cognitive functions (Brown *et al*. 1997). Recently, we have provided a unified account of basal ganglia function by hypothesizing that they are a key element in resolving the action selection problem by serving as a central 'switch' or arbiter between action requests (Redgrave *et al*. 1999). Anatomically, this is plausible because the basal ganglia receive widespread input from all over the brain, including many areas of the cortex (Parent & Hazrati 1993) and subcortex (McHaffie *et al*. 2005). Outputs from the basal ganglia project back, directly or indirectly, to their input targets, forming closed anatomical loops (Alexander & Crutcher 1990; McHaffie *et al*. 2005). For loops including cortex, this occurs indirectly via thalamus. We focus, first, on those aspects of our decision circuitry that make use of the basal ganglia alone.

(a) *The basal ganglia and action selection*

Our model of the circuitry intrinsic to the basal ganglia is drawn directly from our earlier work (Gurney *et al*. 2001*a*). This, in turn, is based on the known anatomy and physiology of the vertebrate basal ganglia, shown in figure 5*a* and described detail in several recent reviews (e.g. Mink 1996; Smith *et al*. 1998). The main input nuclei of the basal ganglia are the striatum and the subthalamic nucleus (STN). The STN is the only source of excitation within the basal ganglia. In primates, the major output nuclei are the internal segment of the globus pallidus (GPi), and substantia nigra pars reticulata (SNr). These nuclei provide extensively branched GABAergic efferents to functionally related zones of the ventral thalamus (which in turn projects back to the cerebral cortex), the midbrain and hindbrain areas critical for movement (e.g. Kha *et al*. 2001). The external segment of the globus pallidus (GPe) is an internal source of inhibition within the basal ganglia. Two separate

Figure 5. Basal ganglia anatomy and functional architecture (*a*) basal ganglia anatomy used as the basis for the model, (*b*) new functional architecture for basal ganglia (Gurney *et al*. 2001*a*) used in the current work. See text for details.

striatal populations have been identified (Gerfen & Young 1988): (i) a population that contains the neuropeptides substance P and dynorphin, and preferentially expresses the D1 subtype of dopamine receptors and (ii) a population that contains enkephalin and preferentially expresses the D2 subtype of dopamine receptors. In most accounts of basal ganglia anatomy, the D1-preferential population is usually associated with projections to SNr and GPi alone, while its D2 counterpart is associated with projections to GPe (Gerfen *et al*. 1990).

The basic assumption underlying our model was that the brain is processing, in parallel, a large number of sensory, cognitive and motivational streams or 'channels', each of which may be requesting/promoting different actions to be taken. For effective use of limited motor resources, it is necessary to suppress the majority of these requests while allowing the expression of only a small number (in some cases

just one). This channel-based scheme is consonant with the view that the basal ganglia comprise a series of afferent and efferent parallel processing streams or loops (Alexander *et al.* 1986; Hoover & Strick 1993; Middleton & Strick 2002). At the systems level, the smallest neuronal population we needed to consider was, therefore, the set of neurons responsible for a single channel within each of the basal ganglia nuclei.

A further assumption was that, implicit in the representation of each action, there is an encoding of its *salience* or propensity to be selected for execution. In our model, we assumed that channel salience had already been extracted from phasic excitatory input by processes in the basal ganglia input nuclei. The input to the model, therefore, was simply the scalar-valued salience of each channel. The basal ganglia output is inhibitory and tonically active. Selection then occurs via selective disinhibition of target structures (Chevalier & Deniau 1990) which include (as well as thalamus) premotor areas of the brainstem. Once inhibition has been released in this way, the corresponding behaviour is enacted. In summary then, large salience signal inputs at striatum and STN select for low signal outputs at the SNr/GPi.

We used the computational premise of selection to guide our interpretation of basal ganglia anatomy in functional terms. One architectural feature that may be invoked in this respect is the diffuse excitation from STN to its targets—GPe and SNr/GPi (Parent & Hazrati 1993, 1995)—in combination with more focused inhibition from striatum to the same nuclei. This constitutes an off-centre, on-surround network that can perform a selection function, as noted by Mink & Thach (1993). However, it is not clear *a priori* what function GPe serves in this scheme, since it is not an output nucleus of the basal ganglia able to implement selection directly. We resolved this problem by observing that, while selection could be performed in principle by the complex of striatum (D1), STN and SNr/GPi alone, the relative levels of excitation and inhibition required to achieve this function were only obtained (and indeed guaranteed) by the inhibition supplied by GPe. We therefore hypothesized that the GPe acts within a control pathway (comprising striatum (D2), STN and GPe) as a source of control signals for the selection pathway (striatum (D1), STN, SNr/GPi). The new functional architecture described above (Gurney *et al.* 2001*a*) is shown in figure 5*b*. Note that it is quite different from the prevailing 'direct/indirect' pathway scheme of Albin *et al.* (1989), and hypothesizes a different role for GPe from that posited by Frank *et al.* (2007) and Hazy *et al.* (2007).

The resulting model (Gurney *et al.* 2001*b*) was able to successfully select and switch between channels based on their input salience. In addition, the model allowed dopaminergic modulation of basal ganglia function in ways compatible with disorders of dopamine function (e.g. Parkinson's disease). While, the role of dopamine is not discussed here, we note that the model is rich enough, in principle, to account for data derived from studies with relevant clinical populations.
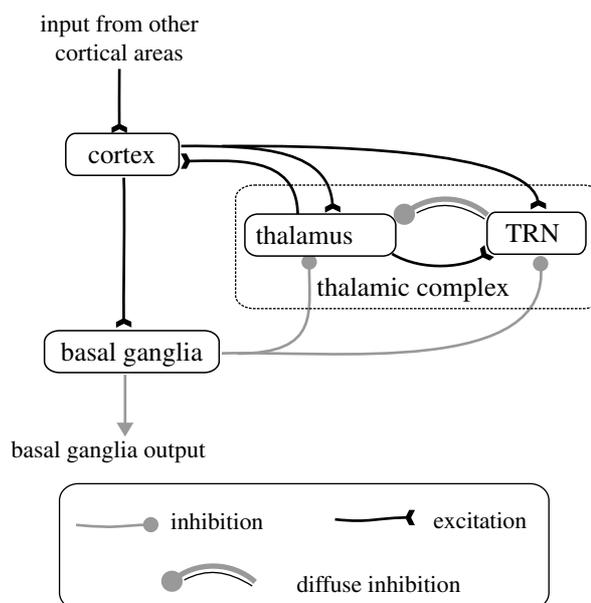
Figure 6. The model of cortico-basal ganglia-thalamo-cortical loops; adapted from Humphries & Gurney (2002).

(b) *Including the thalamic complex*

As noted above, the basal ganglia sit in a wider anatomical context comprising closed loops of cortex–basal ganglia–thalamus–cortex. In previous work, we modelled such loops by embedding the basal ganglia model (described above) into a loop incorporating motor and somatosensory cortex (Humphries & Gurney 2002). In that instance, there are well-understood anatomical relations between these cortical areas, basal ganglia and specific nuclei within thalamus. In the current work, the specific areas of cortex associated with word reading and colour processing are not well understood. We therefore adopt a simplified version of the model in Humphries & Gurney (2002) by using only a single cortical area (figure 6).

Further, whereas in the somatosensory/motor loop the thalamic nucleus is identified as the ventrolateral thalamus (Price 1995), here it is left non-specific and labelled 'thalamus' in figure 6. A component common to both the original and simplified scheme is the thalamic reticular nucleus (TRN) which sends diffuse inhibition to thalamus. The extended thalamo-cortical model retains the channel-based scheme of the basal ganglia model and reciprocal connections between thalamus and TRN imply the latter acts as a distal lateral inhibition mechanism for the former. Input to the model comes from other cortical areas and constitutes an initial representation of salience.

The original somatosensory/motor loop model displays enhanced selection capabilities in several respects when compared with the model of the basal ganglia alone (Humphries & Gurney 2002). Further, using these models in robot controllers has shown that their selection behaviour is of sufficient efficiency and sophistication to be behaviourally adequate in realistic environments (Girard *et al.* 2003; Prescott *et al.* 2006). Details of these models are to be found in Gurney *et al.* (2001*a*) and Humphries & Gurney (2002) and also in the annotated code which is provided in the accompanying electronic supplementary material.

(c) *Combining the Cohen model with the basal ganglia response mechanism*

It is natural to ask if the extended thalamo-cortical-basal ganglia model, viewed as a decision mechanism, can perform appropriate selection in a cognitive task. The model was developed with a view to accounting for action selection in the domain of systems neuroscience with no intention, originally, of being used to generate RT data. Further, the model's ability to account for such data would therefore serve to validate it, and open up possibilities for investigating biologically plausible response mechanisms in the study of cognition.

The rationales for the connectionist Cohen model and our systems neuroscience model are quite different. The Cohen model is a minimal connectionist model designed to test a high-level hypothesis about automatic and controlled processing. On the other hand, our basal ganglia models are biologically constrained, respecting the known anatomy of the target circuits, and were designed to test the hypothesis that those specific circuits could support action selection. Further, whereas learning is a key component of the Cohen model, it does not figure in our models of basal ganglia and thalamus.

There are, however, sufficient points of contact between the two models to allow them to be joined in a unified scheme. Thus, the model in figure 6 is built out of standard leaky integrator neurons (Arbib 1995)— a feature that it shares with the Cohen model—so that they both use a common signal representation denoting neuronal population responses.

The reaction-time behaviour of the model is read from the output units of basal ganglia. Recall that these represent neuron populations providing tonic (continuous background) inhibition to motor targets, and that selection occurs on those channels whose inhibitory output is sufficiently reduced. RT is then interpreted as the time to selection, which is the time from stimulus onset to reduction of basal ganglia output on the selected channel to some threshold value. Moreover, we suppose that this selection threshold may be greater than zero. Although a zero output would demonstrate unequivocal selection, it is unrealistic to suppose that a population of neurons have to be held in a completely silent state for a behaviourally meaningful period of time to allow selection.

Given these observations, a combined model was constructed using the connectionist 'front-end' of the Cohen model (figure 7). This performed initial stimulus processing to provide initial salience input to the thalamo-cortical-basal ganglia model which constituted the response mechanism. The latter uses two channels mimicking the possible outcomes in the Stroop task.

The neural network component of Cohen et al.'s (1990) model performs what is normally thought of as the cognitive elements of the task: stimulus–response translation, attentional control and learning. Only one minor change is required to this 'front-end' to make it compatible with the basal ganglia model response mechanism. The output units of the original Cohen model have resting values of 0.5, the midpoint of their output range which lies in the interval [0, 1]. This is inconsistent with our new interpretation of these
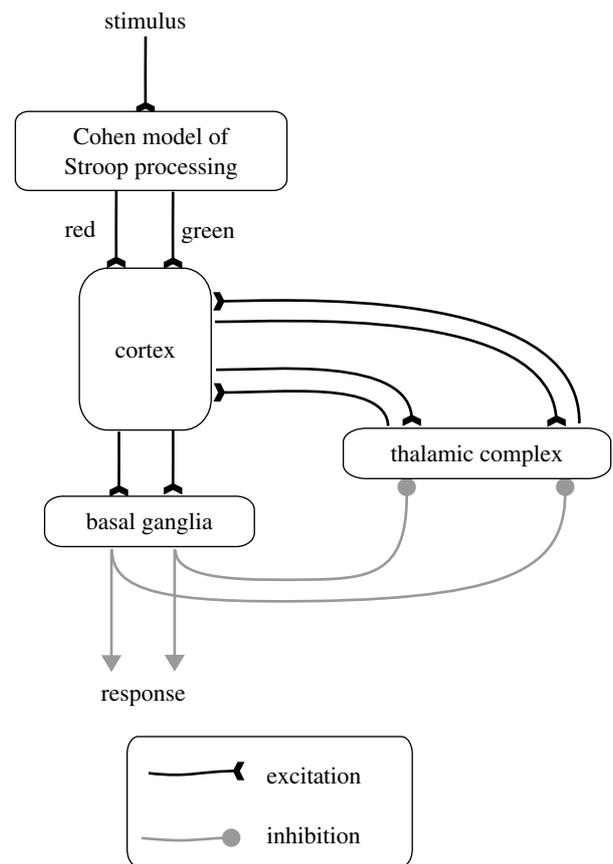

Figure 7. Architecture of the combined model. Initial inputs are processed by the connectionist 'front-end' of the Cohen model as described in Cohen et al. (1990). Outputs representing the evidence in favour of the two possible responses are interpreted as initial salience inputs to the cortical component of the thalamo-cortical-basal ganglia model in figure 6. The basal ganglia outputs determine which response occurs (the output channel which first passes a selection threshold) and the RT (the time required to reach that threshold).

signals as salience values, since it indicates that all possible responses have moderately strong saliences at rest. In the combined model, the resting values of the front-end are set to 0.1, indicating a weakly salient input to the basal ganglia (small changes in weight initialization are also required as a consequence of this manipulation; for details see Stafford 2003).

In all other respects, the combined model is exactly as published by Cohen et al. (1990), except with the basal ganglia model replacing evidence accumulation as the method of final response selection. The basal ganglia thalamo-cortical model used is exactly as published elsewhere (Gurney et al. 2001b; Humphries & Gurney 2002).

(d) *Simulations I: matching basic empirical data*

The combined model successfully replicates the basic (colour naming) Stroop task and the word-reading variation (figure 8). This shows that the model is capable of performing basic selection in a cognitive task and producing realistic RT values.

The ability to realistically model learning phenomena is a key benefit of connectionist models. The combined model mimics the power-law function of learning (figure 9), just as the original Cohen model
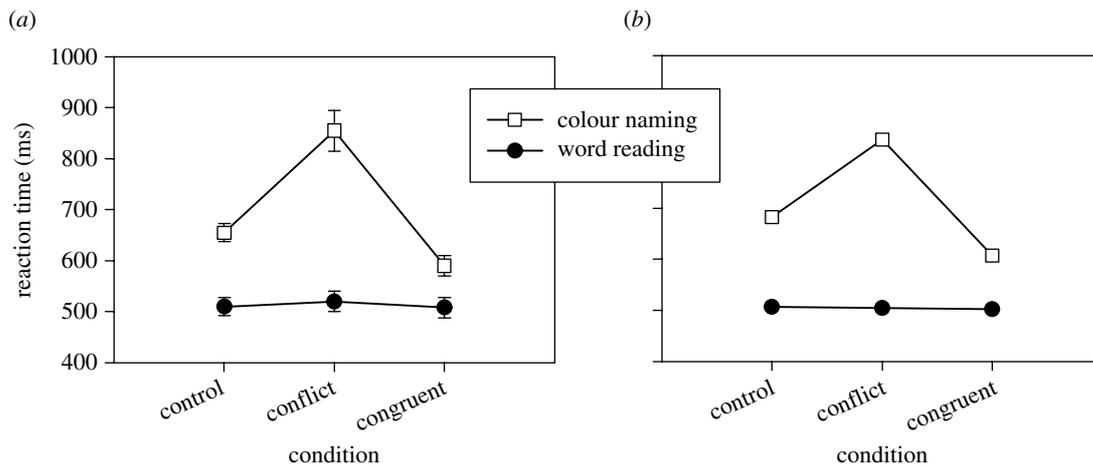
(*a*)

(*b*)

Figure 8. Comparing (*a*) empirical and (*b*) simulation RTs when using the basal ganglia model as the response mechanism for the basic Stroop. Empirical data is adapted from Dunbar & MacLeod (1984), for which standard error bars are shown.
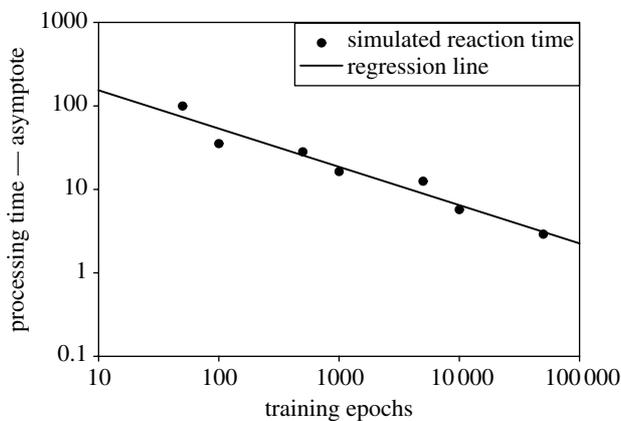

Figure 9. The model conforms to the power law of practice (Logan 1988). Both axes use a log scale. Simulation results are shown as dots. The simple regression for the data is shown as a straight line and follows the form $\log(\text{processing time}) = 2.65 - 0.46 \times \log(\text{Epochs})$. $R^2 = 0.948$.

does (note that no learning takes place in the response mechanism component in either model). This demonstrates that the learning dynamic captured by the connectionist front-end is not interfered with by the use of the basal ganglia response mechanism; graded changes in the signals from the front-end are converted into appropriately graded changes in RTs.

The SOA task reveals that using the basal ganglia as a response mechanism provides a superior fit to the data than when using the original response mechanism (figure 3*c*). Within the original range of the SOA values, the simulation data more closely matches the empirical data. Running the model at extended SOA values (figure 10) confirms that RTs using the basal ganglia response mechanism are stable. At negative SOAs, the salience output caused by the to-be-ignored element of the stimulus is not sufficient to cause selection. Thus, using the basal ganglia response mechanism, the model makes the correct selection at all SOA values. In addition, the amount of interference and facilitation it causes is limited. This is reflected in the stabilization of RTs at SOAs below −400 ms.

Figure 10. Simulation SOA data at an extended range when using the basal ganglia model as the response mechanism.

(e) *Simulations II: dynamic attentional inhibition*

Providing the model with stability under SOA conditions makes possible further model manipulations which bring the model up to date, in a principled way, with developments in our understanding of automatic processing and cognitive control. It has been suggested that selection in the Stroop task is dynamically controlled by a process that monitors for conflicts (located in the anterior cingulate cortex) and increases attentional control in response (Botvinick *et al.* 2001, 2004). Two additional simulations presented here demonstrate that using the basal ganglia model as the response mechanism allows the use of dynamic-attentional modulation to enhance the match to the empirical data.

Here, we neither propose an account of conflict monitoring nor tie it to any specific anatomical location. Instead, we implement solely the essential feature that the appearance of the to-be-ignored
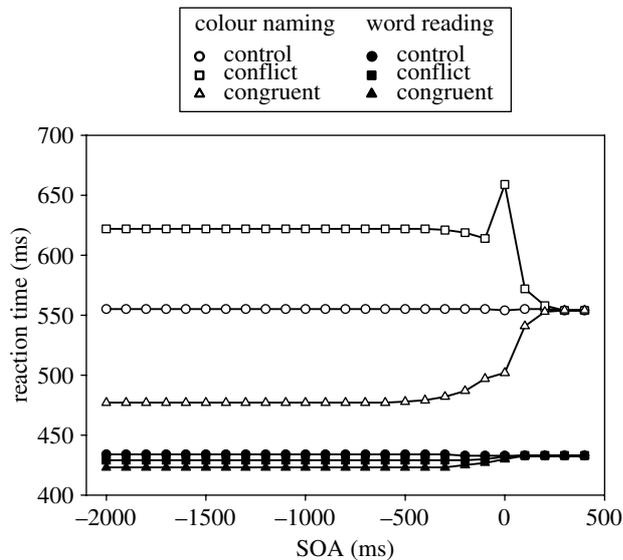
Figure 11. Simulation SOA data when using the basal ganglia model as the response mechanism and with the addition of 'dynamic attentional modulation'.

Figure 12. Simulation SOA data when using the original Cohen model with the diffusion model as the response mechanism and with the addition of 'dynamic attentional modulation'.

element provokes, after some delay, an increase in the attentional inhibition acting on it. The length of the delay used here is 100 ms which accords well with the time-scale of phenomena such as negative priming, (see May *et al.* 1995, for a review) and neurophysiological recordings of activity suppression due to attentional processes (Chelazzi *et al.* 1998). See Usher & McClelland (2001) for a discussion of the time-course of activity during choice selection.

The implementation of attentional modulation in our model is achieved in the following way. After 100 ms (simulated), the inhibition on the relevant hidden units of the Cohen model is increased in magnitude from the default value of -4 to -4.9, the value used by Cohen *et al.* (1990) in their simulations of the SOA task (at all values of the SOA). Thus dynamic attentional modulation is a modification of the mechanism that already exists in the model for implementing attentional selection, using parameters that have already been established. The parameterization of the attentional modulation could have been finessed, but we sought to test the validity of the idea without such *ad hoc* modifications.

Figure 11 shows the simulation results for the model with this dynamic attentional modulation. RTs in the colour-naming conflict condition now peak around the 0 ms SOA point, and flatten-off at a lower level, as occurs in the empirical results (figure 3*a*)—this is an improvement over both the Cohen model and the combined model without dynamic attentional modulation. This simulation both solves the stability problem and matches the peak and decline in RTs that the empirical data shows.

In contrast, with dynamic attentional modulation, the original Cohen model does not successfully match the empirical data (figure 12). Because the stability problem is not resolved, the to-be-ignored stimulus element still provokes erroneous selection at long enough SOAs, and causes unrealistic amount of response-time interference before that.
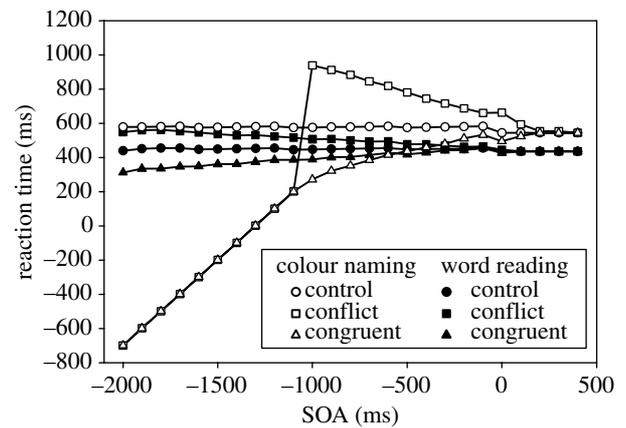
4. DISCUSSION

Our primary result is that a neurobiologically plausible model of action selection allows the successful simulation of RTs in the Stroop task, despite the fact that the model construction was structurally and functionally guided by quite different principles. Although the front-end of the model was explicitly designed to do Stroop processing, it is the response mechanism which is responsible for converting signal outputs into RTs. Structurally, the model was constrained by the known functional neuroanatomy of the basal ganglia; functionally, it was a quantitative interpretation of our action selection hypothesis (Redgrave *et al.* 1999). The basal ganglia model was neither explicitly designed to simulate RTs nor was it constrained by human cognitive performance, yet when processing outputs from the front-end of the Cohen *et al.* (1990) model it has advantages over the diffusion model, which was explicitly designed to simulation RTs, in simulating RTs.

(a) *Why the basal ganglia model successfully simulates reaction times*
The model captures the basic Stroop (figure 8) and learning (figure 9) phenomena because, for moderately sized saliences, selection time is based on the relative difference between the to-be-selected salience and the competing salience (if any). To understand the emergence of RT differences in the basal ganglia model, consider figure 13. Figure 13*a* shows traces (directly from the simulation) of the output signals corresponding to the correct response, in a control and a conflict condition; these signals cross the selection threshold and therefore produce a behavioural response. Note that the output signal in the conflict condition falls to a lower level than in the control condition. It is this final level to which the signal drops which defines the rate at which the signal drops and hence the time to selection. The final signal level is in turn, dependent upon the relative difference between the to-be-selected salience and the competing salience (if any).

The schematic diagram of signal time courses (figure 13*b*) clarifies the way in which final equilibrium output governs selection and RT. The rate of decrease of the output signal has the same relation to time-
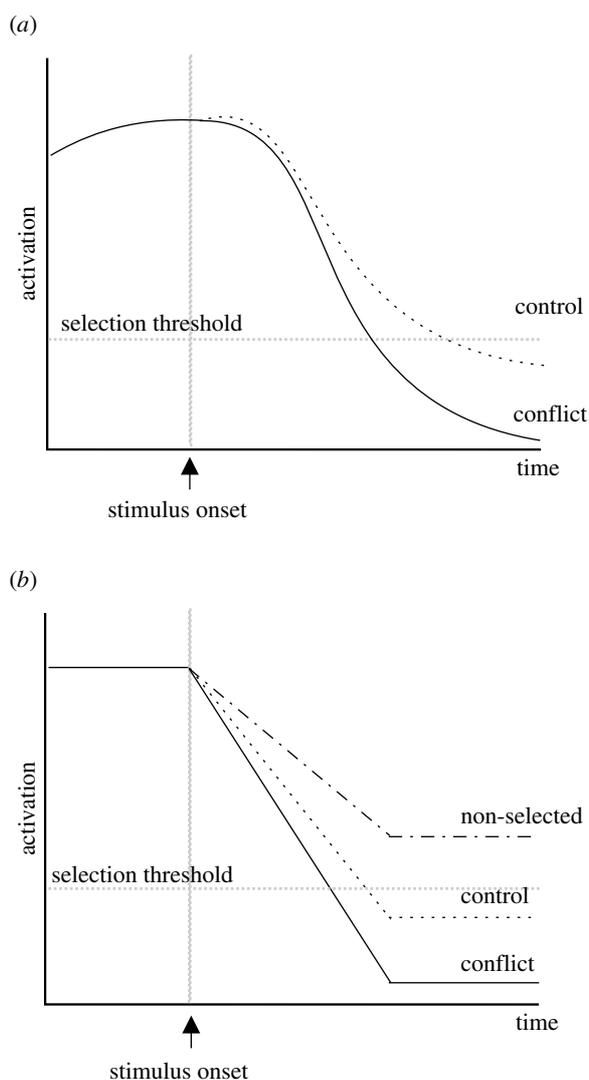
(a)

(b)

Figure 13. Selection in the basal ganglia (a) output signals from example runs of the model in the conflict (solid line) and control (dotted line) conditions (b) schematic illustration of the way in which final equilibrium output governs selection and RT. A selected signal with fast RT (solid line), a selected signal with larger RT (dotted line) and a non-selected signal (dot–dash line) are shown.

to-selection as the drift rate (strength-of-evidence) does to mean RT in the diffusion model. However, because the rate of decrease in the basal ganglia model is ultimately determined by the final output signal resting level, selection does not always occur. In particular, small saliences—which might result from a to-be-ignored stimulus—do not drive the output down beyond the selection threshold.

It is because the basal ganglia model is designed to operate continuously that it has equilibrium final states. Thus, in the idealized situation of unchanging inputs, all patterns of input eventually produce unchanging output states. In particular, for some patterns of input the final output state indicates that no action is selected. In more realistic situations, with noisy input the basal ganglia-thalamo-cortical model is stable to small transient fluctuations in salience (Humphries & Gurney 2002). It is with small saliences, and when dealing with successive rather than simultaneous inputs, that the advantages of using a selection mechanism which has non-selection equilibrium states

is revealed. Both of these cases are revealed by comparison of the SOA simulations (figures 4 and 10).

(**b**) *Weaknesses of the diffusion model*
Our simulations show a situation in which simple evidence accumulation is a non-adaptive choice process. The failure of the Cohen model on the SOA simulations is due to a model feature which is neither trivial nor irrelevant. The empirical existence of the basic Stroop interference effect demonstrates that response activation from the to-be-ignored word element of the stimulus must, at least to some extent, 'break through' any initial attentional inhibition. This activity, arriving at the response mechanism before the response activation of the colour element, is enough, in the Cohen model, to cause selection. The erroneous selection produced at long SOAs shows that a response mechanism must not make selections based on inconsequentially low inputs.

The Cohen model evidence accumulation mechanism has no minimal threshold on inputs, and no decay of accumulated evidence. This means that there are no equilibrium states and it is constantly being driven to enforce selection, no matter how long this takes. By extension, the diffusion model, the general form of the evidence accumulation mechanism used, contains no capacity for not making a selection. This is a serious flaw. At a minimum, it indicates that the context within which the diffusion model of selection is used cannot be ignored or assumed.

(**c**) *Alternative solutions*
We have considered how the choice of response mechanism affects performance in simulation of the Stroop task. Other mechanisms for matching the core empirical data could be envisaged. Cohen & Huston (1994) adapt the model of Cohen et al. (1990) to provide a better match to the SOA data. They do this by removing the diffusion model response mechanism entirely and having selection triggered by activation on the output units crossing a fixed threshold. This solves the problem of selection by arbitrarily small activations, since they do not reach the selection threshold.

This approach allows a fit to the data, but the removal of an explicit response mechanism raises some additional questions. The Cohen & Huston (1994) model, in this respect, bears some similarity to the Usher & McClelland (2001) model of perceptual choice. Both models use a single network for processing stimuli and selecting responses using a simple threshold. Bogacz et al. (2007b) have provided extensions and discussions of the optimality of the Usher & McClelland (2001) model. This model considers mechanisms of choice comprising neuron-like elements but removed from a realistic cognitive or biological architecture. Although models without explicit response mechanisms can fit behavioural data (Cohen & Huston 1994) or be shown to make optimal decisions (Usher & McClelland 2001; Bogacz et al. 2007b), two issues remain unaddressed. Firstly, which neural structures implement the model? Secondly, how is the optimal decision making provided by the model adaptively controlled?

Our approach has been to consider action representation and response selection separately, as in the original Cohen model, and to provide an account of response selection based upon the basal ganglia, as the proposed vertebrate solution to the selection problem. The benefits of using a centralized rather than distributed selection mechanism are discussed in Prescott *et al.* (1999). Among these benefits is the greater theoretical ease of coordinating between multiple competing neural loci—both in terms of lower wiring cost and in the ability to centrally mediate the equivalent of thresholds.

(d) *Benefits of the basal ganglia model*

The simulation of the SOA paradigm highlights two properties which the basal ganglia as a selection mechanism brings to the combined model to improve the possible account of the data. The first, as already discussed, is the lack of incorrect selection for arbitrarily small saliences. The second is the limit on the maximum possible influence of concurrently or consecutively active inputs. Priming of response times, whether positive or negative, occurs because activity on other channels alters the basal ganglia output signals, at a subselection level, thereby affecting the time it takes for outputs to drop below the selection threshold. A similar process occurs in the diffusion model, but accumulated evidence is not limited—and can ultimately lead to incorrect selection (as discussed earlier). Figure 14 shows the geometry of selection interference in both response mechanisms. In the diffusion model (figure 14a), the increase in RT due to a preceding to-be-ignored stimulus is a function of the size of that signal multiplied by time—the longer the to-be-ignored stimulus is presented, the greater the size of the interference effect. If the to-be-ignored stimulus is presented for long enough and the accumulated evidence reaches the selection threshold then an incorrect response is made. In the basal ganglia model (figure 14b), the amount of increase in RT due to a to-be-ignored stimulus is solely a function of the magnitude of the salience that the to-be-ignored stimulus provokes. Because, as discussed above, the basal ganglia model has equilibrium final states, some of which do not indicate selection, the rise in the output signal associated with the correct response is limited and does not increase with time after a certain point. The increase in RT result is commensurately limited, and thus the correct response is selected efficiently.

This is an example of the general 'clean switching' property which has been identified as a desirable feature of any selection mechanism (Redgrave *et al.* 1999). A response mechanism needs to work in real-time, continuously, dealing with the successive selection of actions and interruption of old actions by new. The SOA paradigm shows just one situation where human action selection demonstrates clean switching. The benefits the basal ganglia model brings to modelling the Stroop task demonstrates the value of considering the constraints of natural action selection within cognitive models.
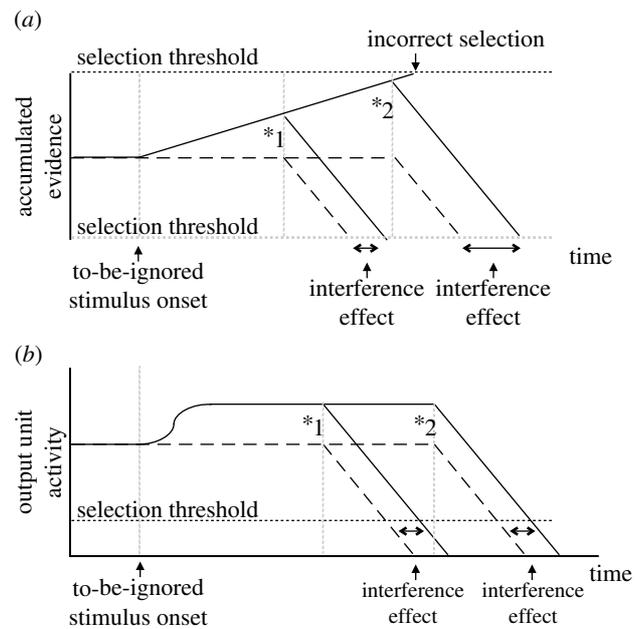
Figure 14. Interference in (*a*) the diffusion model and (*b*) the basal ganglia model response mechanisms. Signals in the diffusion model represent the accumulated evidence in favour of two possible responses, selection is indicated by crossing either the positive or negative evidence thresholds. Signals in the basal ganglia represent the activity on the to-be-selected action channel, selection is indicated by activity on that channel dropping below the selection threshold. Solid lines show signals subject to interference from a preceding input, dashed lines show signals without this competition. The signal courses for the early or later appearance of the to-be-responded-to stimulus are shown (indicated by points *1 and *2, respectively), and the corresponding size of the interference effects is indicated.

(e) *The diffusion model in the context of action selection*

That the evidence accumulation response mechanism, on the other hand, has only one type of final state (that of selecting an action) and that it continuously moves towards this state, has implications for the diffusion model as a model of response selection. The diffusion model embodies the inevitable progression towards selection because all inputs are integrated into a running total of activity, without any decay of that activity. This 'perfect integration' is actually a requirement of the proof that the diffusion model performs optimally (Smith & Ratcliff 2004; Bogacz *et al.* 2007*a*), at least for a restricted class of choices. The simulation of the SOA experiments reveals that selection by perfect integration can be unadaptive in at least some circumstances. This particular case of the general problem of clean-switching shows that adaptive action selection involves criteria beyond those which have been used to define decision optimality (i.e. criteria beyond those pertaining to the kind of simple choices which have hitherto been the main focus of analysis of choice behaviour). This is not to say that the diffusion model, or diffusion-like processes, is inappropriate for selection. Indeed, it has recently been proposed that the basal ganglia architecture is able to perform optimal decision making in a manner akin to the diffusion model, but between

multiple alternatives (Bogacz & Gurney 2007*a*). The diffusion model reflects an optimal way of integrating information if the possible choices are defined, the sources of evidence static and if the point at which the choice process begins is a given. Our claim is only that evidence accumulation and the diffusion model alone cannot provide a full account of adaptive action selection. For this wider problem mechanisms are required which signal the appropriate initialization of the accumulation process, and which reset it or effectively overcome previous accumulation of evidence. The basal ganglia thalamo-cortical model provides a first step towards the integration of the decision–optimal diffusion model into the wider context of adaptive action selection.

(f) Summary and future work

This work validates our model against the basic Stroop phenomena. Use of the basal ganglia model as the response mechanism improves the fit that can be made to the empirical data and highlights necessary features response mechanisms should contain, the lack of which was overlooked in the previous account by Cohen *et al.* (1990). Using an adaptive, action-selection based response mechanism in the model of Stroop task, allows the principled addition to the model of dynamic attentional modulation (Botvinick *et al.* 2001, 2004). Use of the basal ganglia model also extends the account of Stroop processing to connect with the neurobiology of selection.

From a wider perspective, there is a 'theoretical purity' to testing models outside of the domain that they were developed in. Firstly, the basal ganglia model, while not designed to account for RTs, successfully managed to do so. Secondly, the biologically grounded model of the basal ganglia also deals appropriately with signals provided by a more abstract connectionist model of a cognitive task. This depended upon a common signal interpretation at the interface between the two model components in terms of population rate codes. We therefore suggest that this offers a useful tactic in any high-level cognitive modelling that would enable the gradual replacement of abstract model components with more biologically realistic counterparts. Note, however, that this approach does not undermine the principled use of connectionist modelling in quantitative testing of cognitive hypotheses. Thus, in our present context, the model proposed by Cohen *et al.* (1990) was a test of the hypothesis that the Stroop effect could be accounted for in a framework in which the 'strength of processing' devoted to a perceptual or cognitive process determined its status as more or less automatic (or controlled) in relation to other processes. Our work does not challenge the validation of this particular hypothesis since the 'front-end' of the model still tests it perfectly adequately.

Finally, the ability of the model to deal with an arbitrary number of inputs will provide opportunities for future modelling investigations of additional selection paradigms. Making connection to the underlying neurobiology enriches the account possible of Stroop processing. In particular, we anticipate that the existing provision for dopaminergic modulation of signal processing in our model will allows future tests of the model against various pathologies, such as schizophrenia. Our account will also need to be broadened to account for learning within the basal ganglia. Developing a full account of the interaction of plasticity with decision making will be an important test of all existing models of action selection.

REFERENCES

Albin, R. L., Young, A. B. & Penney, J. B. 1989 The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12**, 366–375. (doi:10.1016/0166-2236(89)90074-X)

Alexander, G. E. & Crutcher, M. D. 1990 Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* **13**, 266–272. (doi:10.1016/0166-2236(90)90107-L)

Alexander, G. E., Delong, M. R. & Strick, P. L. 1986 Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381. (doi:10.1146/annurev.ne.09.030186.002041)

Arbib, M. 1995 Introducing the neuron. In *The handbook of brain theory and neural networks* (ed. M. Arbib), pp. 266–272. Cambridge, MA: MIT Press.

Besner, D. & Stolz, J. 1999 Context dependency in stroop's paradigm: when are words treated as nonlinguistic objects? *Can. J. Exp. Psychol.* **53**, 374–380.

Besner, D., Stolz, J. A. & Boutilier, C. 1997 The stroop effect and the myth of automaticity. *Psychon. Bull. Rev.* **4**, 221–225.

Bogacz, R. & Gunery, K. 2007 The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput.* **19**, 442–447.

Bogacz, R., Brown, E., Moehlis, J., Hu, P., Holmes, P. & Cohen, J. D. 2007*a* The physics of optimal decision making: a formal analysis of models of performance in two alternative forced choice tasks. *Psychol. Rev.* **113**, 700–765.

Bogacz, R., Usher, M., Zhang, J. & McClelland, J. L. 2007*b* Extending a biological model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Phil. Trans. R. Soc. B* **362**, 1655–1670. (doi:10.1098/rstb.2007.2059)

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. 2001 Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652. (doi:10.1037/0033-295X.108.3.624)

Botvinick, M. M., Cohen, J. D. & Carter, C. S. 2004 Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* **8**, 539–546. (doi:10.1016/j.tics.2004.10.003)

Brown, L. L., Schneider, J. S. & Lidsky, T. I. 1997 Sensory and cognitive functions of the basal ganglia. *Curr. Opin. Neurobiol.* **7**, 157–163. (doi:10.1016/S0959-4388(97)80003-7)

Chelazzi, L., Duncan, J., Miller, E. K. & Desimone, R. 1998 Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* **80**, 2918–2940.

Chevalier, G. & Deniau, J. 1990 Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci.* **13**, 277–281. (doi:10.1016/0166-2236(90)90109-N)

Cohen, J. D. & Huston, T. A. 1994 Progress in the use of interactive models for understanding attention and performance. In *Attention and performance*, vol. XV (eds C. Umilta & M. Moscovitch), pp. 453–476. Cambridge, MA: MIT press.

Cohen, J. D., Dunbar, K. & McClelland, J. L. 1990 On the control of automatic processes—a parallel distributed-processing account of the stroop effect. *Psychol. Rev.* **97**, 332–361. (doi:10.1037/0033-295X.97.3.332)

Dishon-Berkovits, M. & Algom, D. 2000 The stroop effect: it is not the robust phenomenon that you have thought it to be. *Mem. Cognit.* **28**, 1437–1449.

Dunbar, K. & MacLeod, C. M. 1984 A horse race of a different color—Stroop interference patterns with transformed words. *J. Exp. Psychol. Hum. Percept. Perform.* **10**, 623–639.

Durgin, F. H. 2000 The reverse Stroop effect. *Psychon. Bull. Rev.* **7**, 121–125.

Ellis, R. & Humphreys, G. 1999 *Connectionist psychology: a text with readings*. Hove, UK: Psychology Press Ltd.

Frank, M. J., Scheres, A. & Sherman, S. J. 2007 Understanding decision making deficits in neurological conditions: insights from models of natural action selection. *Phil. Trans. R. Soc. B* **362**, 1641–1654. (doi:10.1098/rstb.2007.2058)

Gerfen, C. R. & Young, W. S. I. I. I. 1988 Distribution of striatonigral and striatopallidal peptidergic neurons in both patch and matrix compartments: an *in situ* hybridization histochemistry and fluorescent retrograde tracing study. *Brain Res.* **460**, 161–167. (doi:10.1016/0006-8993(88)91217-6)

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J. & Sibley, D. R. 1990 D1 and D2 dopamine receptor regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250**, 1429–1431. (doi:10.1126/science.2147780)

Girard, B., Cuzin, V., Guillot, A., Gurney, K. N. & Prescott, T. J. 2003 A basal ganglia inspired model of action selection evaluated in a robotic survival task. *J. Integr. Neurosci.* **2**, 179–200. (doi:10.1142/S0219635203000299)

Glaser, M. O. & Glaser, W. R. 1982 Time course analysis of the Stroop phenomenon. *J. Exp. Psychol. Hum. Percept. Perform.* **8**, 875–894. (doi:10.1037/0096-1523.8.6.875)

Gold, J. I. & Shadlen, M. N. 2000 Representation of a perceptual decision in developing oculomotor commands. *Nature* **404**, 390–394. (doi:10.1038/35006062)

Gurney, K., Prescott, T. & Redgrave, P. 2001a A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol. Cybern.* **85**, 401–410. (doi:10.1007/PL00007984)

Gurney, K., Prescott, T. & Redgrave, P. 2001b A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol. Cybern.* **85**, 411–423. (doi:10.1007/PL00007985)

Hazy, T. E., Frank, M. J. & O'Reilly, R. C. 2007 Toward an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Phil. Trans. R. Soc. B* **362**, 1601–1613. (doi:10.1098/rstb.2007.2055)

Hoover, J. E. & Strick, P. L. 1993 Multiple output channels in the basal ganglia. *Science* **259**, 819–821. (doi:10.1126/science.7679223)

Humphries, M. D. & Gurney, K. N. 2002 The role of intra-thalamic and thalamocortical circuits in action selection. *Network Comput. Neural Syst.* **13**, 131–156. (doi:10.1088/0954-898X/13/1/305)

Kha, H. T., Finkelstein, D. I., Tomas, D., Drago, J., Pow, D. V. & Horne, M. K. 2001 Projections from the substantia nigra pars reticulata to the motor thalamus of the rat: single axon reconstructions and immunohistochemical study. *J. Comp. Neurol.* **440**, 20–30. (doi:10.1002/cne.1367)

Logan, G. 1988 Toward an instance theory of automatization. *Psychol. Rev.* **95**, 492–527. (doi:10.1037/0033-295X.95.4.492)

Luce, R. 1986 *Response times: their role in inferring elementary mental organisation*. New York, NY: Clarendon Press.

MacLeod, C. 1991 Half a century of research on the Stroop effect—an integrative review. *Psychol. Bull.* **109**, 163–203. (doi:10.1037/0033-2909.109.2.163)

MacLeod, C. & MacDonald, P. 2000 Interdimensional interference in the stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends Cogn. Sci.* **4**, 383–391. (doi:10.1016/S1364-6613(00)01530-8)

May, C. P., Kane, M. J. & Hasher, L. 1995 Determinants of negative priming. *Psychol. Bull.* **118**, 35–54. (doi:10.1037/0033-2909.118.1.35)

McHaffie, J. G., Stanford, T. R., Stein, B. E., Coizet, V. & Redgrave, P. 2005 Subcortical loops through the basal ganglia. *Trends Neurosci.* **28**, 401–407. (doi:10.1016/j.tins.2005.06.006)

Middleton, F. A. & Strick, P. L. 2002 Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cereb. Cortex* **12**, 926–935. (doi:10.1093/cercor/12.9.926)

Mink, J. W. 1996 The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425. (doi:10.1016/S0301-0082(96)00042-1)

Mink, J. W. & Thach, W. T. 1993 Basal ganglia intrinsic circuits and their role in behavior. *Curr. Opin. Neurobiol.* **3**, 950–957. (doi:10.1016/0959-4388(93)90167-W)

Opris, I. & Bruce, C. J. 2005 Neural circuitry of judgment and decision mechanisms. *Brain Res. Brain Res. Rev.* **48**, 509–526. (doi:10.1016/j.brainresrev.2004.11.001)

O'Reilly, R. & Munakata, Y. 2000 *Computational explorations in cognitive neuroscience: understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

Parent, A. & Hazrati, L. N. 1993 Anatomical aspects of information processing in primate basal ganglia. *Trends Neurosci.* **16**, 111–116. (doi:10.1016/0166-2236(93)90135-9)

Parent, A. & Hazrati, L. N. 1995 Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res. Rev.* **20**, 91–127. (doi:10.1016/0165-0173(94)00007-C)

Phaf, R. H., Vanderheijden, A. H. C. & Hudson, P. T. W. 1990 Slam—a connectionist model for attention in visual selection tasks. *Cognit. Psychol.* **22**, 273–341. (doi:10.1016/0010-0285(90)90006-P)

Platt, M. 2002 Neural correlates of decisions. *Curr. Opin. Neurobiol.* **12**, 141–148. (doi:10.1016/S0959-4388(02)00302-1)

Posner, M. & Snyder, C. 1975 Attention and cognitive control. In *Information processing and cognition* (ed. R. Solso). Hillsdale, NJ: Erlbaum.

Prescott, T. J., Redgrave, P. & Gurney, K. 1999 Layered control architectures in robots and vertebrates. *Adapt. Behav.* **7**, 99–127.

Prescott, T. J., Montes Gonzalez, F. M., Gurney, K., Humphries, M. D. & Redgrave, P. 2006 A robot model of the basal ganglia: behavior and intrinsic processing. *Neural Netw.* **19**, 31–61. (doi:10.1016/j.neunet.2005.06.049)

Price, J. 1995 Thalamus. In *The rat nervous system* (ed. G. Paxinos), pp. 629–648, 2nd edn. New York, NY: Academic Press.

Ratcliff, R. 1978 A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108. (doi:10.1037/0033-295X.85.2.59)

Ratcliff, R. & Smith, P. 2004 A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* **111**, 333–367. (doi:10.1037/0033-295X.111.2.333)

Ratcliff, R., Cherian, A. & Segraves, M. 2003 A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* **90**, 1392–1407. (doi:10.1152/jn.01049.2002)

Reddi, B., Asrress, K. & Carpenter, R. 2003 Accuracy, information, and response time in a saccadic decision task. *J. Neurophysiol.* **90**, 3538–3546. (doi:10.1152/jn.00689.2002)

Redgrave, P., Prescott, T. J. & Gurney, K. 1999 The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* **89**, 1009–1023. (doi:10.1016/S0306-4522(98)00319-4)

Sharkey, A. & Sharkey, N. 1995 Cognitive modeling: psychology and connectionism. In *The handbook of brain theory and neural networks* (ed. M. Arbib), pp. 200–203. Cambridge, MA: The MIT Press.

Smith, P. & Ratcliff, R. 2004 Psychology and neurobiology of simple decisions. *Trends Neurosci.* **27**, 161–168. (doi:10.1016/j.tins.2004.01.006)

Smith, Y., Bevan, M. D., Shink, E. & Bolam, J. P. 1998 Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* **86**, 353–387. (doi:10.1016/S0306-4522(97)00608-8)

Stafford, T. 2003 Integrating psychological and neuroscientific constraints in models of stroop processing and action selection. PhD thesis, University of Sheffield. See http://www.abrg.group.shef.ac.uk/.

Stafford, T. & Gurney, K. 2004 The role of response mechanisms in determining reaction time performance: Pieron's law revisited. *Psychon. Bull. Rev.* **11**, 975–987.

Usher, M. & McClelland, J. L. 2001 The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592. (doi:10.1037/0033-295X.108.3.550)

Zhang, H. Z. H., Zhang, J. & Kornblum, S. 1999 A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognit. Psychol.* **38**, 386–432. (doi:10.1006/cogp.1998.0703)