Article:
Stafford, T. and Gurney, K.N. (2004) The role of response mechanisms in determining reaction time performance: Piéron's Law revisited. Psychonomic Bulletin and Review, 11 (6). pp. 975-987. ISSN: 1531-5320

The role of response mechanisms in determining reaction time performance: Piéron's Law revisited.

Tom Stafford & Kevin N. Gurney
Department of Psychology, University of Sheffield, UK

Abstract

A response mechanism takes evaluations of the importance of potential actions and selects the most suitable. Response mechanism function is a non-trivial problem, and one which has not received the attention it deserves within cognitive psychology. We make a case for the importance of considering response mechanism function as a constraint on cognitive processes, and emphasise links with the wider problem of behavioural action selection. First, we show that, contrary to previous suggestions, a well-known model of the Stroop Task (Cohen, Dunbar & McClelland, 1990) relies on the response mechanism for a key feature of its results - the interference-facilitation asymmetry. Second, we examine a variety of response mechanisms (including that in the model of Cohen et al.) and show that they all follow a law analogous to Piéron's Law in relating their input to reaction time. In particular, this is true of a decision mechanism not designed to explain RT data, but based on a proposed solution to the general problem of action selection, and grounded in the neurobiology of the vertebrate basal ganglia. Finally, we show that the dynamics of simple artificial neural elements also support a Piéron-like law.

This is a preprint of:

> Stafford, T. & Gurney, K.N. (2004). The role of response mechanisms in determining reaction time performance: Piéron's Law revisited. *Psychonomic Bulletin & Review, 11* (6), 975-987.

It is based on the final MS submitted to the publisher and so is identical except for some formatting changes. Please feel free to circulate or cite.

Author's Note: Thanks to Stephen Want and Mark Humphries for reading drafts of this paper and to Dick Eiser and Jim Stone for advice during preparation. Tram Neill, Jennifer Stoltz, Tracy Brown and one anonymous reviewer provided invaluable comments on an earlier draft. Special thanks are due to Nicol Harper for suggesting we investigate the possible neuronal underpinning of Piéron's Law.

Correspondence concerning this article should be addressed to Kevin Gurney, Department of Psychology, University of Sheffield, Western Bank, Sheffield, S10 2TP, UK. Email: k.gurney@shef.ac.uk

1 Introduction

A response mechanism is the component of any model of human decision making which takes evaluations of the importance of potential actions and selects the most suitable action. Despite appearances, the design of a biologically plausible switching mechanism is a non-trivial problem (Redgrave, Prescott & Gurney, 1999). In this paper we argue that response mechanisms are an important part of cognitive models, that their function is an important area for investigation and the processes of the human response mechanism have consequences for performance in diverse areas of human behaviour. The importance of the study of response mechanisms is highlighted by the emergence of the topic in disciplines separate from cognitive psychology. Thus, ethology, robotics and neuroscience have come to recognise the importance of the 'action selection problem' (Prescott, Redgrave & Gurney, 1999; Tyrell 1992). From these perspectives, a response mechanism is necessary to deal with the resolution of conflicts between functional units that are in competition for behavioural expression. For example, a food deprived animal should re-evaluate the importance of feeding if a predator is detected, thereby requiring a response selection to the new stimulus (e.g. continue feeding, flee, or fight).

Appropriate behavioural selection is also clearly an issue of central importance to cognitive psychologists, but response mechanisms have generally not received the emphasis that is their due. All cognitive models must explicitly, or implicitly, contain a response mechanism. In many formal models this may be no more than a simple threshold, above which activations indicate a response. Even this minimal feature is fulfilling the role of the response mechanism, although this kind of mechanism is functionally impoverished (Ratcliff, Van Zandt & McKoon, 1999; Stafford, 2003). Thus, a response mechanism should be designed to cope with the multiple, conflicting demands of the behaving organism (Redgrave, Prescott & Gurney, 1999). Something about the way these demands are reconciled in the human case may be shown by studies of reaction times in tasks such as the Stroop task (Stroop, 1935) and other simple choice paradigms (Luce, 1986). Conversely, of course, the study of response mechanisms and of the action selection problem should illuminate facets of performance in these paradigms.

Within cognitive psychology the study of response mechanisms in their own right has largely been restricted to the modelling of decision mechanisms that are able to mimic the pattern of reaction time in simple choice paradigms (Luce, 1986; Ratcliff & Rouder, 1998). Further, while response mechanisms have been studied in this 'choice theory' context, little is known about the brain regions where they might be instantiated, nor has their been an attempt to create models which are based on neurophysiological principles or constrained by known neuroanatomy.

The significance of response mechanism has been raised by a recent discussion that has made it clear that models of decision mechanisms inspired by different research paradigms have convergent properties (Ratcliff, 2001; Reddi & Carpenter, 2000). The comparison of these different models raises general issues in study of decision making, such as the issue of exactly which stage of processing the response mechanism

is modelling (Carpenter & Reddi, 2001).

The starting point for the investigations of response mechanisms presented in this paper is the popular model of the Stroop task described by Cohen, Dunbar & McClelland (1990); henceforth 'The Cohen model'. This is a connectionist implementation of the hypothesis that performance in the Stroop task can be understood using the concepts of automatic and controlled processes. These two processes are distinguished in the model through different 'strengths of processing' within two paths with differently weighted connections. The Cohen model remains the 'standard' model of Stroop processing (e.g. Ellis & Humphreys, 1999; MacLeod & MacDonald, 2000). The model captures the essential generic features of the processing of conflicting stimuli in a connectionist framework. However, the simplicity of a model can be deceptive. The correct explanation for the behaviour of the model (and, thus, in turn for human performance) may not be immediately apparent. We present here an investigation into which mechanisms in the model may explain the successful simulation of Stroop phenomena, particularly those of interference and facilitation.

We show that the full explanation of Stroop performance by the model must involve the, hitherto neglected, response mechanism. In doing so, we demonstrate that the temporal relationship between this mechanism's input and output (reaction time) is analogous to Piéron's Law (1914; Piéron, 1920, 1952). We therefore go on to investigate the relation between other response mechanisms (Phaf, Vanderheijden & Hudson, 1990; Ratcliff, 1978; Reddi and Carpenter, 2000; Gurney, Prescott & Redgrave, 2001a) and Piéron's Law. One such mechanism is consistent with neuroscientific constraints since it can be identified with a set of specific neuroanatomical structures (the basal ganglia) and has been the subject of a computational model of action selection (Gurney, Prescott & Redgrave, 2001a; Gurney, Prescott & Redgrave, 2001b). We go on to examine the underpinning of a Piéron-like law in this mechanism and its relation to neural functionality.

2 The Stroop Task

The Stroop task (Stroop, 1935) is a popular paradigm for the investigation of the cognitive mechanisms involved in attention, automaticity, and, most importantly for us, the processing of conflicting stimuli and conflicting responses (for reviews see MacLeod, 1991; MacLeod & MacDonald, 2000). The Stroop colour-naming task involves responding to the colour of a coloured word string which can itself be the name of a colour. There are three possible general classes of stimuli. For congruent stimuli the word and the colour match (e.g. the word 'red' in red ink), for conflicting stimuli the word and the colour are at odds (e.g. the word 'red' in green ink), and for control stimuli the irrelevant dimension is, at least nominally, neutral with regard to the target dimension (for example, the string 'XXXX' in green ink, or the word 'chair' in green ink). Typical response times for all conditions when using non-word control stimuli (such as 'XXXX') are shown in figure 1 (Dunbar & MacLeod, 1984). This baseline condition is

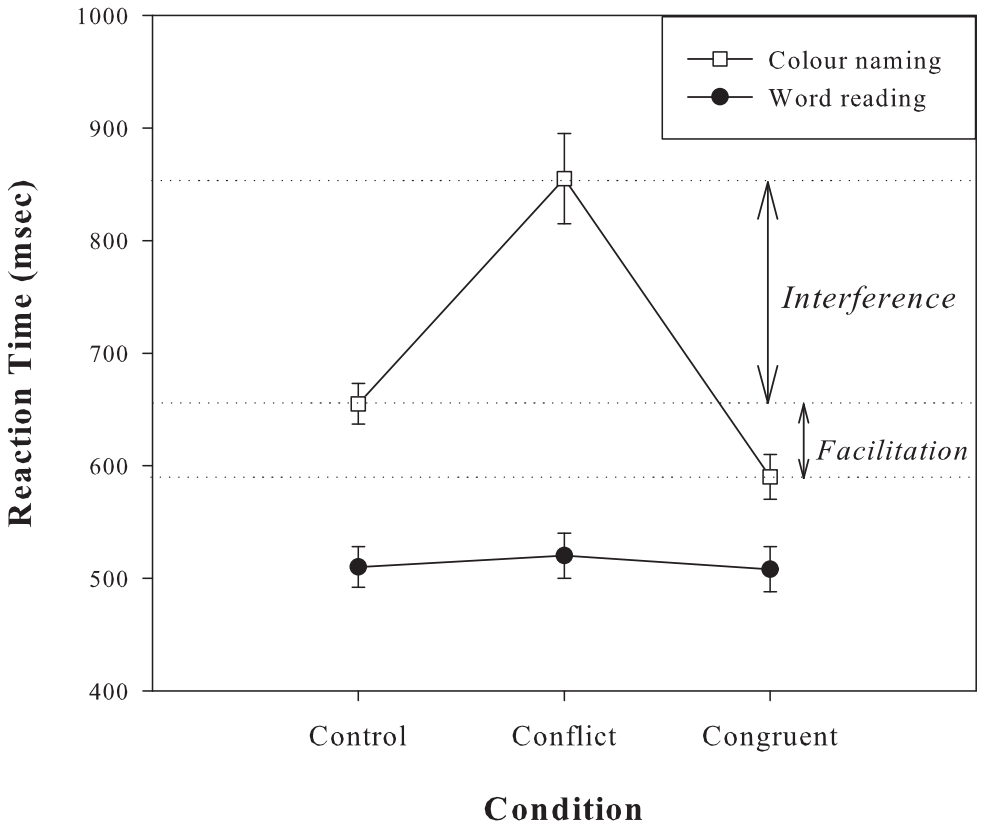

Figure 1: The basic Stroop effect for colour-naming with non-word control. N.B. Interference is greater than facilitation. Data from Dunbar & MacLeod (1984, p.630).

chosen to provide data because it the one used in the Cohen model.

The main feature of Stroop task response time data is that the word dimension of the stimulus affects the speed of colour naming although, in the complementary word-reading task, the colour dimension does not significantly affect the speed of word-reading. This has traditionally been interpreted within an automaticity framework (e.g. Posner & Snyder, 1975) in which the reading of the words occurs automatically, despite the influence of attention, and affects the naming of colours. Recent results show that word reading is not automatic, and words can be successfully ignored if the task conditions are right (Besner, Stolz & Boutilier, 1997; Besner, 2001; Durgin, 2000). Results such as these support a general deconstruction of the notion of automaticity (Bargh, 1989; Duncan, 1986; Logan, 1988, Pashler, 1998; Ryan, 1983). One novel aspect of the Cohen model was that it showed how notions of the conditional and quantitative nature of what have previously been considered automatic processes may be naturally incorporated within a connectionist framework.

A second feature in Stroop data with non-word control stimuli, is the asymmetry of the influence of congruent versus incongruent information in the colour-naming task. The interference effect on reaction times produced by conflicting colour-words is greater than the facilitation effect produced by congruent colour-words. The Cohen model provides an explanation of this asymmetry and it is this explanation which is the point of departure for our analysis of response mechanisms.

The preceding discussion indicates the wide-ranging extent of the explanatory power of the Cohen model, which may account for its receiving considerable attention. This model therefore deserves close scrutiny in order to understand the mechanisms responsible for its properties.

3 A Model Of The Stroop Effect

The model of processing in the Stroop task advocated by Cohen, Dunbar & McClelland (1990) consists of a Parallel Distributed Processing (PDP) network (Rumelhart, McClelland & The PDP Research Group, 1986) which processes the raw stimulus characteristics and then provides outputs for a response mechanism. The model simulates a two-colour Stroop task with non-word control (e.g. 'XXXX' in the relevant colour) and has two output nodes corresponding to the two possible responses ('red' and 'green' say). The response mechanism is based on an accumulator model of the kind used to simulate reaction times in simple choice situations (Luce, 1986). The relative evidence in favour of each competing responses is accumulated in its associated 'bin' (in our example, one for 'red', one for 'green') until the crossing of a threshold by the value of one bin signals a response. Although, as we will show, the response mechanism is an important functional part of the model, it is rarely discussed in reviews of the model (Ellis & Humphreys, 1999; MacLeod, 1991; MacLeod & MacDonald, 2000).

Cohen et al (1990) base their explanation of the ratio of interference to facilitation on the properties of the function relating input to output in the individual units of the network. Like many other PDP models, the units in the Cohen model use a logistic activation function to relate the weighted sum of inputs ('net input') in a unit to its output. The logistic function limits each unit's output y to lie between 0 and 1 according to the rule

$$y = \frac{1}{1 + e^{-x}},$$

where x is the net input. According to Cohen et al (1990), this function is source of the asymmetry between interference and facilitation effects. The basis of their explanation (shown graphically in figure 2A) is that, relative to a control condition baseline greater than zero, *decreasing* the input by certain amount Δx, produces a larger change in output than a similar *increase* in input of Δx.

This explanation is also used to support the view that interference and facilitation are products of the same mechanism (Cohen, Dunbar & McClelland, 1990; Cohen, Servan-Schreiber & McClelland, 1992) - a matter we return to in the discussion. This explanation accompanies the exposition of the model

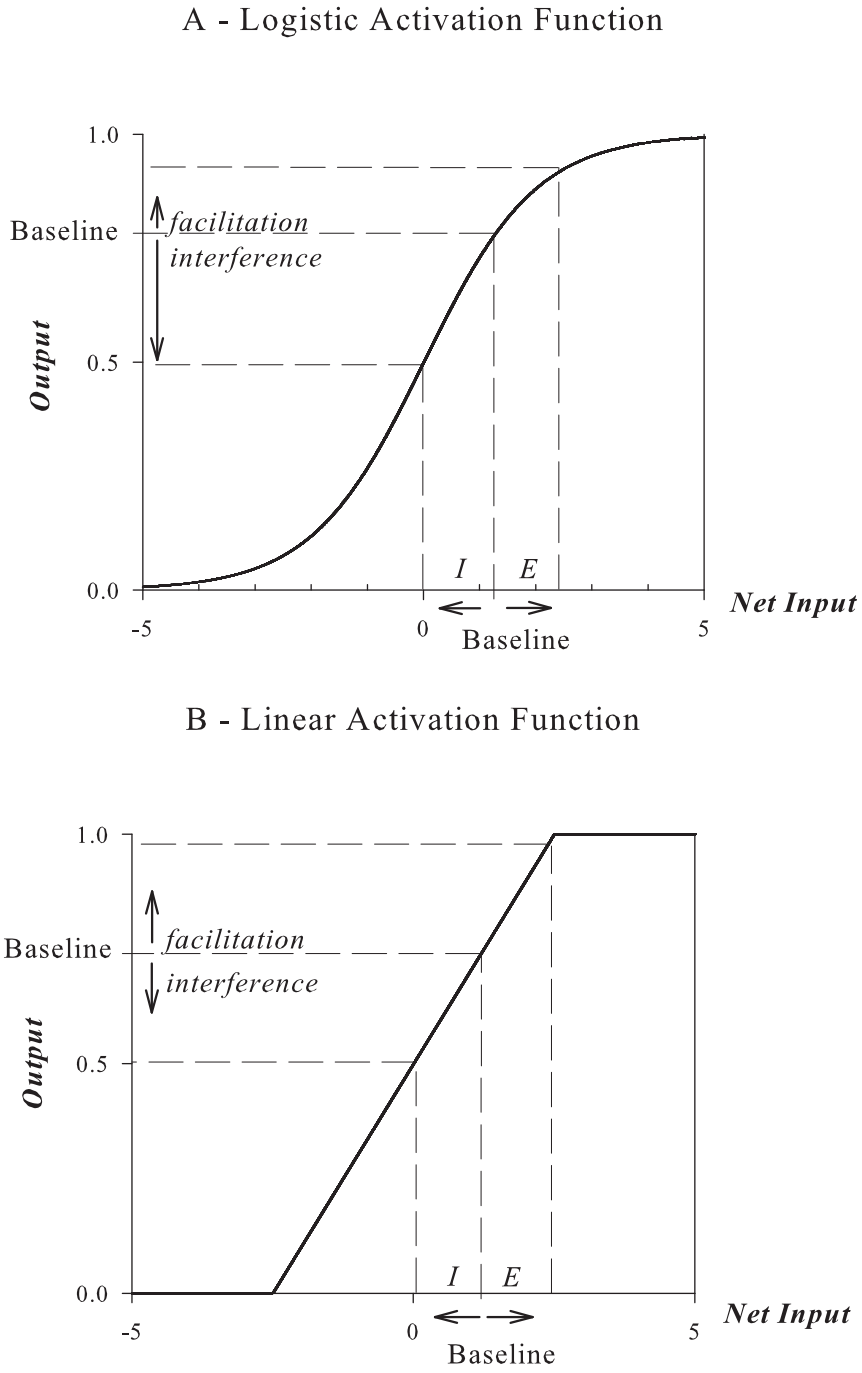

Figure 2: The logistic activation function (A) and the piecewise linear activation function (B) with annotation showing how excitation, 'E', and inhibition, 'I', of the baseline input affect output. The effect of equal excitation and inhibition is asymmetrical for the logistic function, the putative source of the difference between interference and facilitation, and symmetrical for the piecewise linear function.

in textbooks (e.g. Ellis & Humphreys, 1999, Sharkey & Sharkey, 1995) and critical reviews (notably, MacLeod, 1991). However, close examination of the explanation shown in figure 2, shows that it is contingent on the activity in the neutral condition falling above the inflection (midline) point of the curve. Cohen et al (1990) do not provide a justification for their assumption that the unit input in the neutral condition falls above the inflection point of the sigmoid function. Indeed in our replication of their reported model the unit input in the neutral condition falls just below, rather than above, the inflection point. That the simulations still produce interference effects greater than facilitation effects is an indication that the sigmoid function cannot be responsible for this effect in this particular model.

4 The True Locus Of The Interference - Facilitation Asymmetry Lies In The Response Mechanism

4.1 Simulations show irrelevance of the unit activation function non-linearity

In order to demonstrate that the non-linearity of the activation function cannot on its own explain the difference between interference and facilitation in the Cohen model, we present two simulations: one is a replication of the original model, and the other is identical, but uses a piecewise linear activation function instead of a logistic activation function (see figure 2B). The rationale for this is that, since the slope of the region of interest of the piecewise linear function is constant, it ceases to be the case that increases in a unit's net input (associated with facilitation) result in smaller output changes than decreases in the net input (associated with interference). The piecewise linear function possesses the essential 'squashing' non-linearity of the logistic function but with simpler description - it is defined by

$$y = \begin{cases} 1 & : x < \epsilon \\ 0 & : x > 1/m + \epsilon \\ m(x - \epsilon) & : otherwise \end{cases}$$

where m determines the position of the function on the x-axis and ϵ the slope. We used values $m = 0.2$ and $\epsilon = -2.5$ to approximate the shape of the logistic sigmoid used in the original formulation of the Cohen model. Both simulations (sigmoid and piecewise linear activation) used the post-training weights given in Cohen et al (1990).

The results (figure 3) show that both simulations yield essentially the same pattern of reaction times. At this stage, the argument given by Cohen et al is still potentially valid *if* the piecewise linear function saturates during the network operation. However, we observed this not to be the case and all unit outputs were exercised over the strictly increasing part of the function. The similarity of the results using the different activation functions therefore shows that the decreasing slope of the logistic function *cannot* be the source of the difference between interference and facilitation.

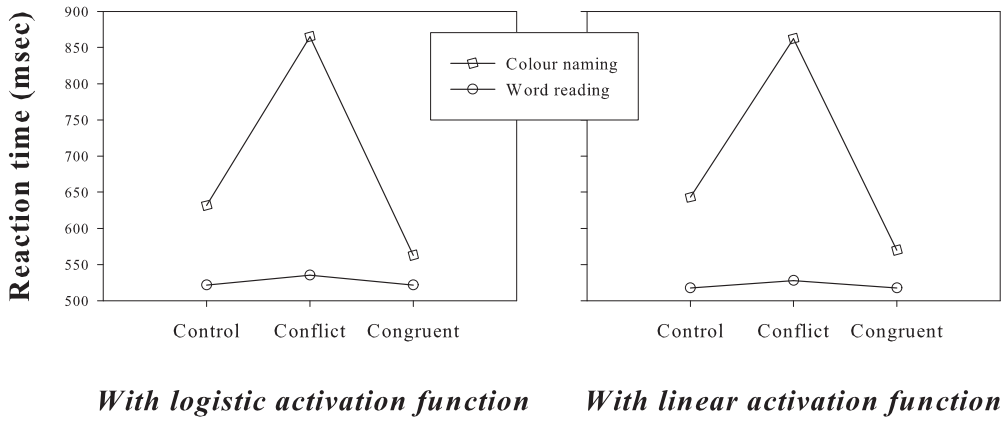

Figure 3: Reaction times on the colour naming task for our replication of the original Cohen et al (1990) model using the original logistic activation function (A) and using a piecewise linear activation function (B). The empirical data is accurately simulated with both activation functions.

Further insight is gained by inspecting the effect the different activation functions have on the unit outputs that are provided to the response mechanism. The response mechanism of the model calculates reaction times based on the relative strength of evidence, that is, the difference between the target and competing outputs of the network. The inputs to the response mechanism, calculated from the outputs of the model with the weights and parameters given in Cohen et al (1990), are shown in table 1. Two things should be noted: first, the outputs for the word-reading conditions are not shown, since they do not vary significantly between conditions or models and are not relevant to the discussion of the point at hand; second, these are the values close to their asymptotes at equilibrium we deal with dynamical issues later. These response mechanism input values are in no way chosen by us, but result from the design of the model by Cohen et al (1990) with only the activation function changed as described.

Table 1.

Two points may be made from an inspection of table 1 and figure 3. First, there is a small difference (0.04) in relative evidence for interference and facilitation with the logistic activation function, and this difference will promote the required asymmetry. However, it has a negligible effect on reaction times, since the model with piecewise linear activation has no difference in relative evidence but produces a very similar pattern of results (figure 3). Second, although the piecewise linear model produces symmetric inputs to the response mechanism, there is still an asymmetry in reaction time differences (with respect to control). This means that the equilibrium values of the network output cannot be responsible for the asymmetry; there is no network-bound 'symmetry breaking' via equilibrium outputs. Notice that the symmetry in strength of evidence relative to control here is contingent, not only on the piecewise linearity of the node output function, but also on symmetries in the patterns of input to the final stage of network processing. Thus, while the strictly increasing part of the piecewise linear activation function can (unlike

Table 1: **The strength of evidence for target response in the three colour-naming conditions for models using the two activation functions. Note: The change in strength of evidence relative to the control condition is shown in brackets for the conflict and congruent conditions.**

	Logistic function		Piecewise Linear Function	
Condition	Strength of evidence	Change	Strength of evidence	Change
Control	0.48		0.46	
Conflict	0.27	-0.21	0.24	-0.22
Congruent	0.64	+0.17	0.68	+0.22

the sigmoid) faithfully transmit symmetric differences in net-input to nodes in the output layer, such difference must be in place for this to occur. The architecture and training of the Cohen model both conspire to ensure this is the case.

Further, Cohen et al (1990) claim that the time constant for leaky integration in the artificial neurons of the model also played a role in determining the asymmetry between interference and facilitation. However, we have found that removing the neuron temporal dynamics altogether, so that their outputs change instantaneously, has a negligible effect on the simulation results. Combining the second and third observations here, implies that neither network dynamics nor equilibrium properties can be held responsible for the asymmetry in facilitation and interference. This means that the locus of the asymmetry must, for this particular model, lie in the response mechanism; it remains to be seen exactly what feature of the response mechanism does create this phenomenon.

4.2 A response mechanism which follows an analogue of Piéron's Law produces the asymmetry between interference and facilitation

In the Cohen model, the unit outputs for the two competing responses are passed from the connectionist part of the model to the response mechanism. This mechanism is based on evidence accumulation. In a basic version of this scheme, each of the two possible decisions ('red', 'green', say) is associated with an evidence 'bin' and, at each time-step, each bin has its value altered by an amount proportional to the difference between the network output for its corresponding decision and that of the alternative decision. Thus, introducing decision indices $i, j = 1, 2$, if μ_i is the change in evidence for decision i and y_i the associated network output then

$$\mu_i = \alpha(y_i - y_j),$$

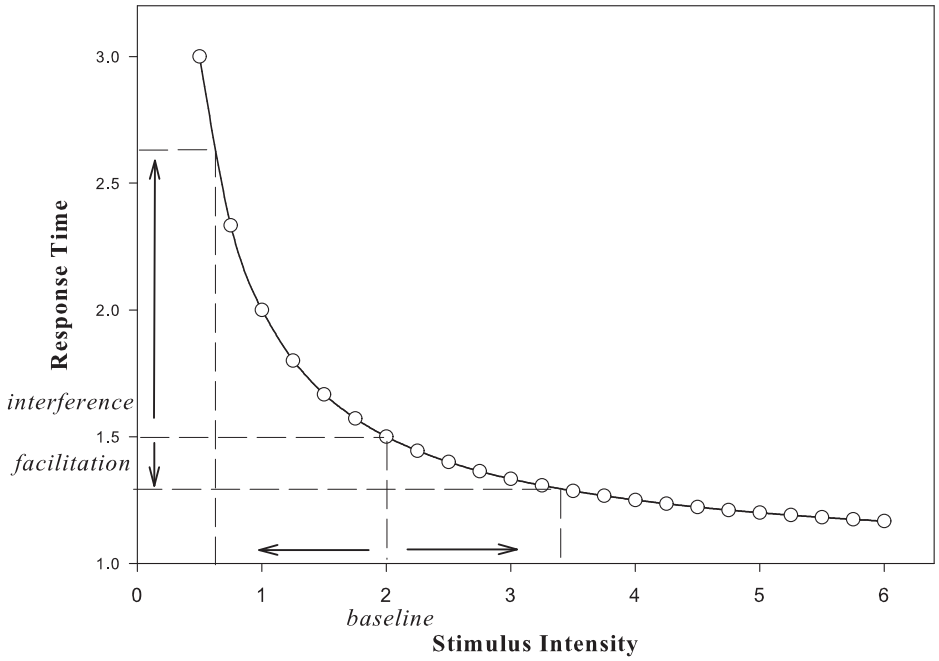

Figure 4: Response time as a function of strength of relative evidence in Cohen et al's response mechanism. The change due to increased intensity is greater that the change due to a decrease in intensity.

where $i \neq j$ and α is a scaling parameter less than 1 which determines the rate of evidence accumulation. The counters are initialised to zero at the start of each trial and a decision is signalled when the counter for either decision crosses some threshold. Cohen et al (1990, p.338) finesse this basic scheme by adding zero-mean gaussian noise to the evidence μ_i before accumulating it in each counter. In calculating response time based on relative evidence, the response mechanism in the Cohen model is similar to many other models of reaction time, including those based on mathematical models of decision processes (e.g. Luce, 1986; Ratcliff & Rouder, 1998) and on the neurophysiology of saccades (Reddi & Carpenter, 2000).

The role of the response mechanism may be elucidated by examining the functional relationship between the reaction time (RT) and the strength of evidence $E = y_i - y_j$, under the approximation that E, is fixed for the duration of the response. The resulting function is shown in figure 4, which shows that response time is a negatively accelerating function of input. As shown in figure 4, increasing the relative strength of evidence above baseline for a decision does not speed the response time as much as an equally sized decrease slows response time. This is exactly what is required to explain the fact that interference is greater than facilitation in the Cohen model.

Further insight about this function may be obtained by quantifying its analytic form. Let $b_i(n)$ be the value of evidence in bin i at timestep n. Without loss of generality, assume bin 1 forces a decision by reaching the threshold θ. At each time step, bin 1 is incremented by αE so that $b_1(n) = n\alpha E$. Let n_I

be the smallest integer n such that $b_1(n) \geq \theta$ then, if αE is much less than θ (or equivalently, n is much greater than 1), $n_I \alpha E \approx \theta$. Rearranging and taking the log of both sides

$$\log n_I \approx \log(\theta/\alpha) - \log E. \tag{1}$$

Now, n_I is proportional to the reaction time, RT, so, putting $RT = cn_I$, Equation 1 may be written

$$\log \text{RT} \approx \log k - \log E, \tag{2}$$

where, $k = c\theta/\alpha$. This is a special case of the more general form

$$\log(RT - R_0) \approx \log\, k - \beta \log E, \tag{3}$$

where $\beta = 1$, and $k = 0$. This, in turn, may be written as

$$RT \approx R_0 + kE^{-\beta} \tag{4}$$

which expresses the reaction time as an exponentially decreasing function of the strength of evidence with an asymptotic response time R_0.

If strength of evidence is replaced by stimulus intensity then equation (4) corresponds to Piéron's Law (Piéron, 1914; Piéron, 1920, Piéron, 1952) which describes an early finding from psychophysics that intensity of a stimulus is related to the latency of response by an exponentially decaying function. Piéron's Law has been found to hold for both visual and auditory stimuli (reviewed in Luce, 1986), for gustatory reaction times (Bonnet, Zamora, Buratti & Guirao, 1999) and for simple and choice reaction time tasks (Pins & Bonnet, 1996). From equation (3) the law may be expressed in an affine (linear with non-zero offset) form by plotting the log of the input against the log of RT (minus the asymptotic value); the resulting straight line has slope $-\beta$ and intercept $\log k$. Straight line plots of this kind provide (with careful interpretation — see below) a convenient method of assessing to what extent other functions follow a form analogous Piéron's Law. Such a plot for the Cohen et al (1990) response mechanism is shown in figure 5.

The asymptote for the data and Figures 5-8 were all produced using a procedure outlined below. However, to understand why we adopted this technique, it is necessary to beware as Luce (1986) notes, that fitting Piéron's Law to data provides "an estimation problem of some delicacy". An important factor is whether the fit is carried out before or after the transformation to log-log coordinates. The transformation to log-log coordinates exaggerates the discrepancy between the data and the best-fit line at lower RTs. Hence fitting in log-log space can provide the illusionary appearance of a better fit. With this in mind, a Piéron's Law-like curve, as defined in equation (4) was fitted to the data using the *fminsearch* function from MATLAB version 6.1. This is an unconstrained non-linear optimisation procedure which uses the simplex search method (Lagarais, Reeds, Wright & Wright, 1998). The asymptotic value obtained

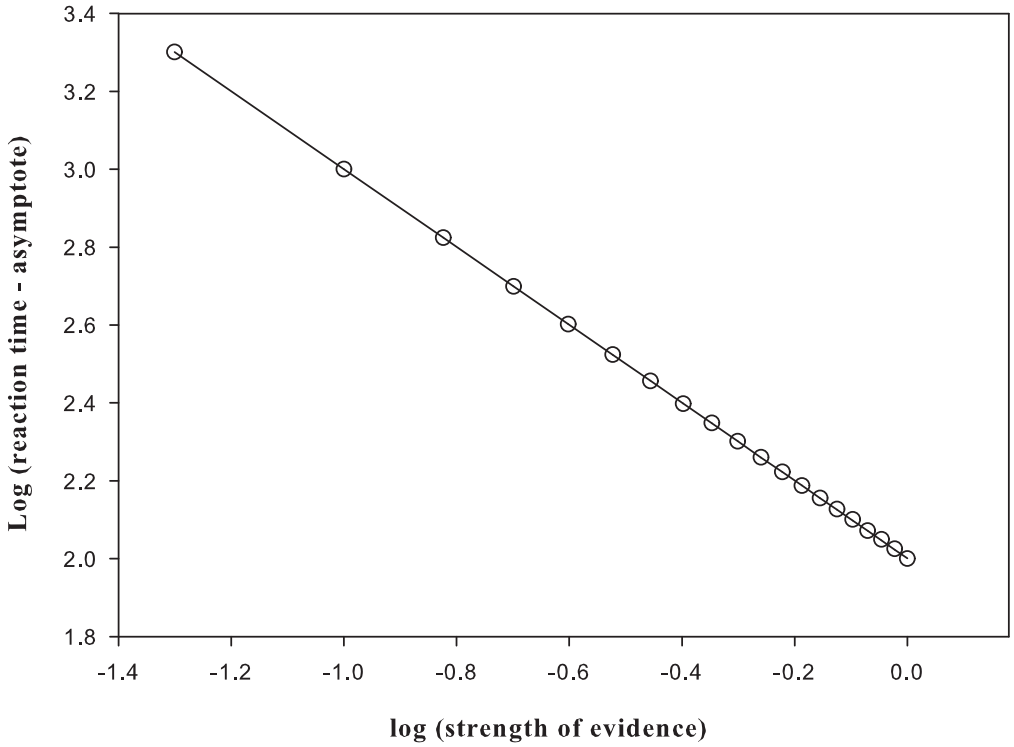

Figure 5: The input-response time function of the Cohen et al (1990) response mechanism shown on a log-log plot. The line representing the fit to a Piéron-Law-like function is shown, which in this case fits the data points exactly. The best-fit line and asymptote are derived by the standard procedure as defined in the text.

was used to plot the log of reaction time less the asymptote on the y-axis, while the log of the input to the response mechanism was plotted along the x-axis. Functions which fit Piéron's Law exactly produce straight lines when plotted like this. The line shown on the graphs is the best fit line that results from using the set of parameters derived from this optimisation procedure. For some plots the range of data used to derive the asymptote was longer than the range of data shown. This was done in order to more accurately derive the asymptotic value, which is the major influence on the straightness of the line. When this was the case, a second simplex search was done for the range of data shown on the graph with the asymptotic value fixed but the other two parameters, k and β, unconstrained.

The fit in figure 5 is very good, and is only limited by the approximation invoked to obtain equation (1). Therefore, in the limiting case of very small time steps, the Cohen et al (1990) response mechanism follows a Piéron's Law-like function exactly. We are not asserting that strength of evidence is the same as stimulus intensity, but that strength of evidence — the input to the Cohen response mechanism — fulfils an analogous role to that of stimulus intensity vis-à-vis the relation to reaction time.

To summarise, the differential magnitude of reaction time change under facilitation and interference conditions in the model may be attributed to the shape of a function determined wholly by the response mechanism. This function relates reaction time to differential evidence between the outputs of the connectionist network and fits a form described by Piéron's law.

5 Other Response Mechanisms Obey Piéron's Law

Having established that one popular model of the Stroop effect relies on the fact that its response mechanism follows a function analogous to Piéron's Law, it is natural to go on and ask: Is this an idiosyncrasy of the accumulator model utilised by Cohen et al (1990) or do other models, and other response mechanisms, also follow such a law? We show below that a number of successful response mechanisms, taken from different research fields, all follow a Piéron's Law-like form.

The SLAM model (Phaf, Vanderheijden & Hudson, 1990) also successfully simulates reaction times in the Stroop task. The type of response mechanism it uses is a 'sampling and recovery procedure' (Raaijmakers and Shiffrin, 1981). At each time step, an output unit in the network is chosen (according to a sampling distribution based on relative unit activities) as a candidate for implementing a response decision. If the chosen unit has index i and activation $a_i(t)$, then its probability P_i of forcing a response is given by

$$P_i(t) = 1 - e^{-\alpha_i(t)}. \tag{5}$$

To gain more insight we now assume that (as with the accumulator mechanism) the unit chosen for activation is fixed and that its activation is constant over the response time. There is then a fixed probability P_i that a decision is made at each time step. It is straightforward to show that the expected reaction time $\langle RT \rangle$ is then just $1/P_i$ so that

$$\langle RT \rangle = \frac{1}{1 - e^{-\alpha_i}}. \tag{6}$$

To see if this could be expressed in the form of a Piéron-like law, we sought values of R_o, k, β which allow $\langle RT \rangle$ to be expressed as function of a_i in the form given by equation (4). These were found using the non-linear function fitting routine described above and the results are shown in figure 6. Thus, the Phaf et al (1990) response mechanism follows Piéron's Law very closely if we interpret the unit activation and the *expected* reaction time $\langle RT \rangle$ as the independent and dependent variables respectively.

Arguably the most successful mathematical model of response times for two-choice decisions is the diffusion model (Ratcliff & Rouder, 1998; Ratcliff, Van Zandt & McKoon, 1999; Ratcliff, 1978). This model belongs to the general class of random walk models, which are closely related to accumulator models such as that used by Cohen et al (1990). They differ mainly in that they contain only a single counter or accumulator, which is incremented or decremented towards positive and negative thresholds representing

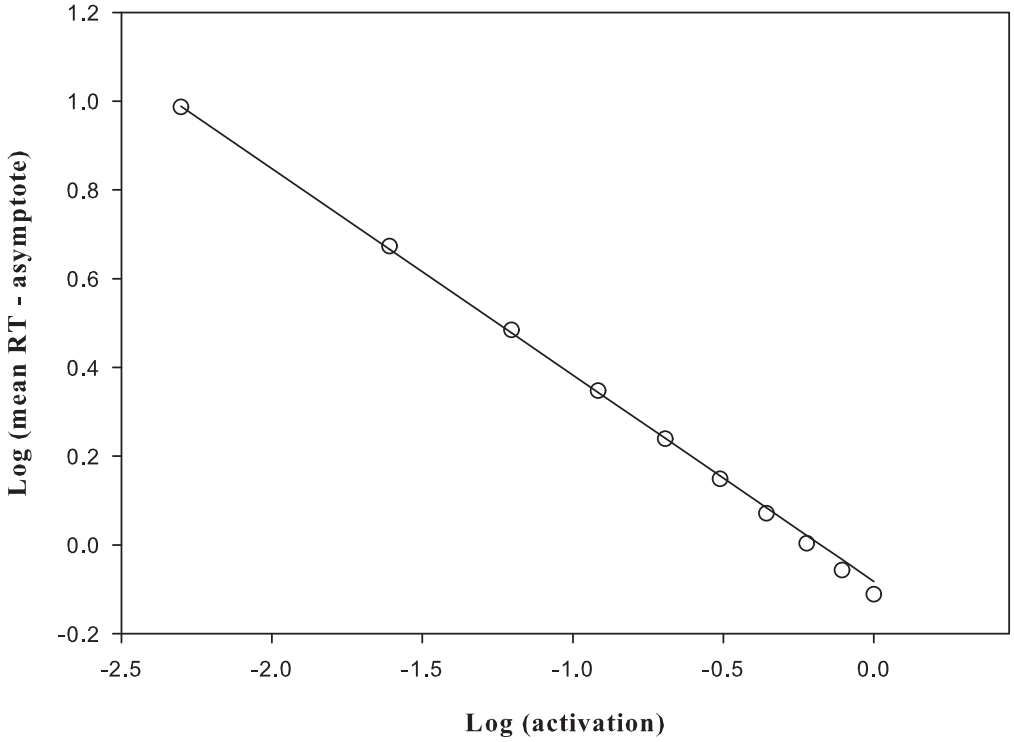

Figure 6: The input-response time function for the Phaf et al (1990) response mechanism shown on a log-log plot. The best-fit line and asymptote are derived by the standard procedure as defined in the text.

the two competing responses. Both classes of models have a long history of investigation in the context of choice reaction time studies (Luce, 1986). In the diffusion models, within each trial, the drift is stochastic. However, it is possible to define a mean drift rate as the mean rate of approach to the threshold, and which may be considered to reflect the relative strength of evidence for a response. Although a key strength of the diffusion model is that it accounts for the distribution of response latencies, we consider the model without the inclusion of noise in order to more easily derive the input—mean response time function.

A somewhat different model due to Reddi & Carpenter (2000) is termed LATER (Linear Approach to Threshold with Ergodic rate) and uses a constant drift rate within each trial but varies this rate randomly from trial to trial. This model is based on studies of saccade generation latency in humans and other primates (for reviews see Gold & Shadlen, 2001; Schall, 2001).

In both the diffusion and LATER models the mean drift rate r acts like the rate of evidence accumulation αE used for the model of Cohen et al (1990) with the result that their generic form fits Piéron's Law almost exactly with respect to r as the independent variable. Thus, these models display a Piéron Law-like relation of exactly the same form shown in figure 5.

5.1 Biologically grounded decision mechanisms

In addition to models based on matching simulation results to behavioural studies we have also investigated a response mechanism which is based on the neuroanatomy of a brain system believed to play a crucial role in behavioural response selection — the basal ganglia. The basal ganglia are a subcortical complex of nuclei which we have proposed fulfill the role of a 'central switch' in mediating behavioural action selection in vertebrates (Prescott, Redgrave & Gurney, 1999; Redgrave, Prescott & Gurney, 1999). Briefly, our model of the basal ganglia (Gurney, Prescott & Redgrave, 2001a; Gurney, Prescott & Redgrave, 2001b; Humphries & Gurney, 2002) is based on the known connectivity and neurotransmitter function of the nuclei of these brain circuits and upon the hypothesis that the system functions as a central selection mechanism. Behaviours compete with each other for expression and the basal ganglia selects those that are most urgent; that is those with the largest salience. It is important to realise that, unlike the previous models described here, the basal ganglia model is not based on an attempt to simulate the pattern of reaction times in any particular context. Rather, it is founded on the proposed mechanisms of the biological systems putatively contributing to them. Despite this, the basal ganglia model also follows an approximation to Piéron's Law (with respect to input salience), as shown in figure 7. In figure 7, a line (dotted) is also shown which was found by fitting to the data in the log-log space. Although it appears better than the fit found using our non-linear optimisation technique, it is an inferior fit in the untransformed (non log-log) space (see the remarks above about fitting to Piéron's Law).

It is possible to explain the trends observed for the basal ganglia model reaction times by the functionality of the units which make up the model. Like many other neural network models (including that

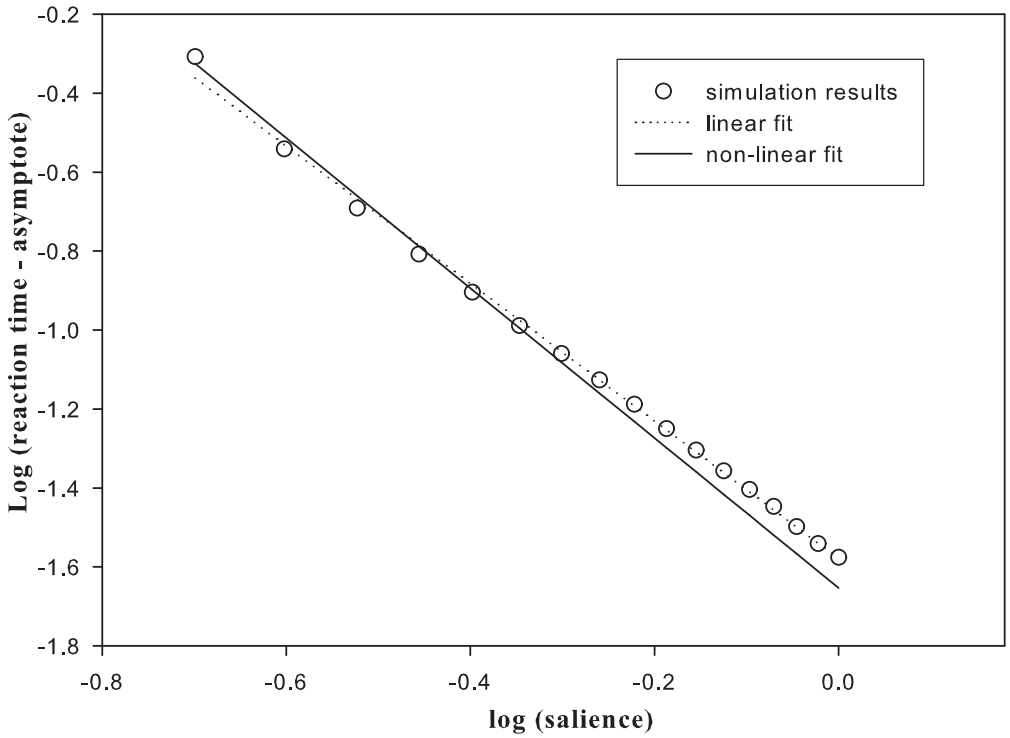

Figure 7: The input response time function for the basal ganglia selection model shown on a log-log plot. The best-fit line derived by the standard procedure (as defined in the text) is shown as a solid line. The best-fit line obtained after the transformation to log-log space is shown as a dashed line.

of Cohen et al, 1990) the basal ganglia model consists of *leaky integrator* neurons. Such model neurons represent the simplest possible approximation to a dynamic neural membrane and, in a way similar to real neurons, adjust their output gradually to be commensurate with their input. To represent the function of such a neuron formally, let $I(t)$ and $a(t)$ be the input and activation of the neuron, respectively, at time t. then

$$\frac{da(t)}{dt} = -pa(t) + qI(t) \tag{7}$$

where p determines a characteristic time constant $\tau = 1/p$, and q is a constant which affects the overall influence of the input. The dynamics mean that the neuron is continually integrating ('accumulating') information over time and therefore has some of the characteristics of the response mechanisms discussed above. We can now derive a relationship between the response time t_θ of the neuron and a constant input I, where t_θ is defined as the time for the activity a, to cross a critical threshold, θ. Suppose that the neuron is at rest and receives a step input I at $t = 0$. It is then straightforward (see for example, Kaplan, 1952) to solve equation (7) to obtain

$$a = \frac{qI}{p}\left(1 - e^{-pt}\right). \tag{8}$$

When $t = t_\theta$ then $a = \theta$. Substituting these into (8) and solving for t_θ in terms of I gives

$$t_\theta = -\frac{1}{p}\log\left(1 - \frac{p\theta}{qI}\right). \tag{9}$$

This function is shown in figure 8 together with a regression line based on fitting Piéron's Law.

The similarity to the basal ganglia input-response time function suggests that the essential characteristics of the latter may reflect the basic response properties of its component model neurons.

6 Discussion

6.1 Summary of results

We have demonstrated several main results. First, certain features of the model of Cohen et al (1990) can only be properly understood if attention is paid to the response mechanism. In particular, the asymmetry between facilitation and interference in this model is not a result of the choice of artificial neuron output function (the sigmoid). In fact, the main contribution to this asymmetry is not grounded at all in the connectionist 'front-end' (either its equilibrium or dynamic aspects), but rather in the response mechanism. Second, this mechanism follows a relation analogous to Piéron's Law and, while the input to the response mechanism is not stimulus intensity as such, it is an analogous quantity — the 'strength of evidence' supplied by the front-end. The asymmetry is then easily explained in terms of the nonlinearity of the power law that describes the Piéron-like, input-output relation of the response mechanism. Third,

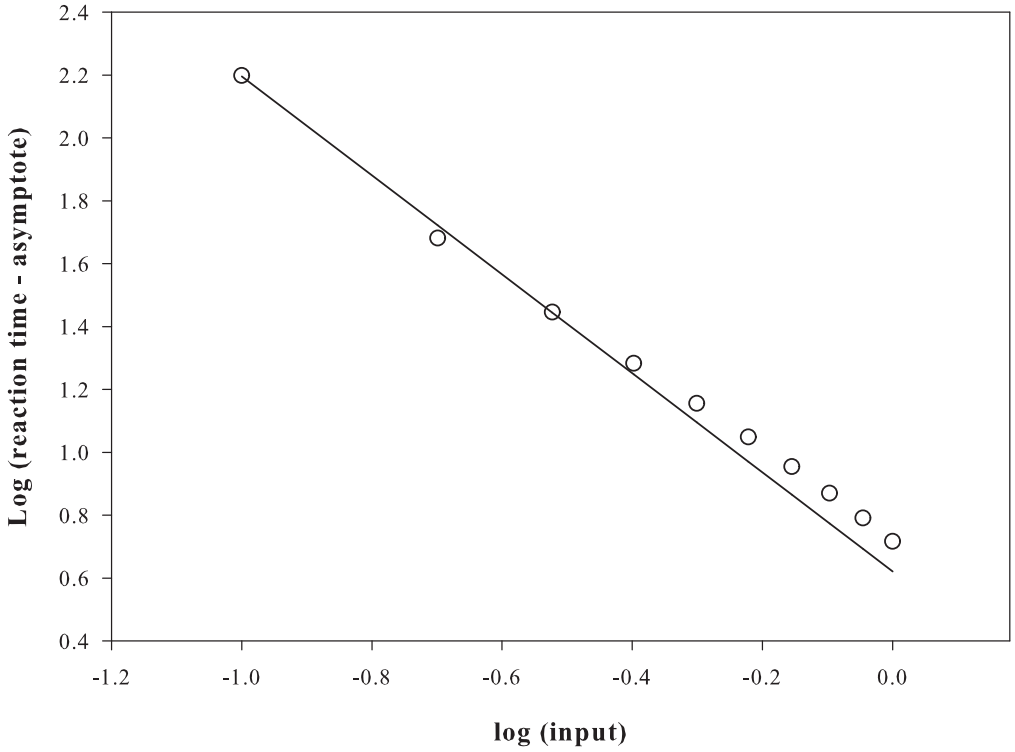

Figure 8: The input response time function for a simple model neuron derived analytically, shown on a log-log plot. The best-fit line and asymptote are derived by the standard procedure as defined in the text.

following on from this, several other response mechanisms follow at (least approximately) a Piéron-like relationship if their input ('drift rate', 'salience' etc) is interpreted as the independent variable giving rise to a reaction time. In particular, a decision mechanism based on the neurobiology of the basal ganglia also follows a Piéron-like law even though it was not designed to explain RT data. The underlying reason for this last mechanism behaving as it does appears to reside in the dynamics of its neural elements. Some of the possible consequences of these results are now explored further.

6.2 The Single Mechanism contention

Cohen et al (1990) make the claim that their model, unlike other theories of Stroop processing, demonstrates that the asymmetry between interference and facilitation *could* stem from the same mechanism. Our analysis of the Cohen model shows that this claim is still true *of this model*, although we would now move the locus of the source of this asymmetry to the response mechanism. Of course, whether this is validated experimentally is another question. The 'single mechanism' explanation has been criticised on empirical grounds because of evidence which shows that interference and facilitation can be differentially affected by experimental manipulations (Tzelgov, Henik & Berger, 1992; MacLeod, 1998) and that interference and facilitation are not only created by separate processes but at separate stages (Brown, Gore & Carr, 2002; MacLeod, 1998; MacLeod & MacDonald, 2000, Brown, 2003).

While this evidence appears to be at odds with the model of Cohen et (1990), it is not impossible to conceive of other models which could accommodate these data. The architecture of the Cohen model is one, specific example of a feedforward network and, in particular, its structure allows (with the piecewise linear activation function) symmetric effects at the output layer under the congruent/conflict manipulation (and which contributed to the pattern of results shown in table 1). In contrast, in order to accommodate the data cited above, a model would have to flexible enough to allow two things. First, unlike the Cohen model, the outputs alone should allow asymmetric differential evidence between congruent and conflict stimuli; this would permit the locus of asymmetry to exist at more than one site ('front-end' and response mechanism). In such a model it may also be possible, of course, to account for the asymmetry entirely in the front-end, and invoke a more linearly behaving response mechanism. Second, the outputs should allow the size of the interference effect to be manipulated independently of that of facilitation. While the Cohen model may not be able to support these features, it is the case that a 2-layer connectionist network can be constructed to give any pattern of outputs in response to its inputs that one demands (Funahashi, 1989; Hornik, Stinchcombe & White, 1989). That is, a network could be built to support the required pattern of 'evidence' relationships described above and which, in combination with a (possibly Piéron-like) response mechanism, *could* account for a complex pattern of data describing facilitation and interference in the Stroop effect. However, the unbridled application of the (essentially unlimited) computational power of networks to model psychological data in this way has not been without criticism (e.g. McCloskey, 1991). Thus, according to the critics, using unstructured

multi-layer perceptrons, one obtains a 'black-box' whose internal mechanisms are not transparent, and whose relationship to psychological processes or anatomical loci is unclear. Now, in spite of the possible shortcomings of the Cohen model (in terms of its flexibility to explain the full range of data pertaining to congruence and facilitation), it is not a black-box in the sense described above. Rather, it is founded on the principle of implementing a well defined hypothesis — that automaticity and of processing are described in terms of their relative 'strength of processing'. To this extent, the model is to be commended, for its being both constrained (by the hypothesis) and transparent in its operation.

Lindsay & Jacoby (1994) also address the single mechanism controversy. Their analysis, based on a process dissociation procedure (Jacoby, 1991), focusses on the independent but co-occuring contributions of word-reading and colour-naming processes to interference and facilitation. The two processes use different response bases and this, combined with the different contributions of the two processes in the conflict and congruent conditions, produces the asymmetry between interference and facilitation. As Lindsay & Jacoby (1994) note, their analysis is compatible with parallel processing models such as that of Cohen et al (1990) and we suggest that this focus on processes and mechanisms, grounded in quantitative models, is the most fruitful perspective for advancing the current debate.

6.3 Baseline effects

Notwithstanding the discussion above, it has been suggested that, under certain circumstances (use of noncolour-word control) interference is *not* greater than facilitation. According to this account, the use of an 'XXXX' control pattern provides a baseline which is faster than a noncolour-word control, and thus exaggerates interference effects at the expense of facilitation effects (Brown, Roos-Gilbert & Carr, 1995; Brown, Joneleit, Robinson & Brown, 2002; Brown, Gore & Carr, 2002, Brown, 2003). Baseline effects could be accounted for by a model which took input from non-colour words as well as colour words, and in which non-colour-word controls produce a strength of evidence (for input to the response mechanism) less than the XXXX controls. Clearly this is outside the remit of the Cohen model. However, the possibility that such a model could be constructed in principle is guaranteed using the same arguments about the generality of 2-layer nets invoked in the previous section. Once again, however, the provisos outlined in the last section about the utility of such a model (in terms of the transparency of its explanatory power) will apply if one adopts a critical stance vis-à-vis the kind of network that underpins it.

6.4 Piéron's Law and information integration at the neuronal level

It is intriguing that several well-known response mechanisms exhibit characteristics similar to those of Piéron's Law. This points to the possibility that any psychologically plausible response mechanism should obey a law of this kind. However, this is not a deduction from our results, and we have to admit the possibility that entirely different (non Piéron-like) relationships could exist between the input of a psychologically validated response mechanism and its output.

In the biologically grounded basal ganglia model, the particular shape of the Piéron law-like function for the BG model (shown in figure 7) resembles that of the model neuron function (shown in figure 8). This gives some support for the hypothesis that the Piéron's Law properties of the basal ganglia model are due to the properties of its fundamental units, rather than being an anomalous result of the particular connectivity of the system. We therefore conjecture that any system comprised of units with the same dynamics as model neurons such as these will follow a Piéron-Law like function. This raises the possibility that Piéron's Law itself (as observed in human data) may be based on the information integration properties of individual neurons.

It is usually acknowledged that the function of neuronal elements is substantially more complex than that expressed by the simple equation (7) we use to model the units in this paper. This would appear to argue against a simple neural element explanation of Piéron's Law. However Koch (1999) points out that the complexity of multiple non-linear intra-neuronal processes could combine to create an approximately linear input to mean-firing-rate relationship. Thus, the function of some biological neurons may be approximately equivalent to that of the units described here. This, in turn, makes it at least plausible that Piéron's Law could be grounded in the properties of biological neurons.

The neuronal explanation of Piéron's Law supposes that a systems-level property arises by preserving properties of the system's fundamental through levels of construction involving significant inter-element interaction. We call this kind of explanation 'transparent'. This can be contrasted with an explanation in which systems-level properties do not exist at the system elemental level but only emerge from the interaction of these components. Transparent explanations may be more robust to minor modifications of systems-level features of the model than emergent explanations

6.5 The value of the biologically grounded action selection perspective

The shift in locus of explanation for the relative values of facilitation and interference in the Cohen et al.'s model of Stroop performance emphasises the importance of response mechanisms for cognitive tasks. Our view is that response selection in such tasks is a special instance of the more general process of behavioural action selection conducted by all animals in a continuous fashion in their the natural state (Prescott, Redgrave & Gurney, 1999; Redgrave, Prescott & Gurney, 1999). The problem of action selection is therefore central to a study of human behaviour, and so we are not surprised to find mechanisms mediating its solution occupying a central place in understanding laboratory-based cognitive tasks. Further, while existing models of the Stroop task use abstract models of response selection crafted explicitly to model reaction time data, our model of the basal ganglia (Gurney, Prescott & Redgrave, 2001a; Gurney, Prescott & Redgrave, 2001b; Humphries and Gurney, 2002) is based on biological considerations and was constructed to solve the problem of selection. However, it is able to successfully model a similar relation between reaction times and salience as is displayed by other response mechanism models between RTs and evidence or drift rate. In addition it formed the basis for proposing a possible neurally inspired

explanation for this Piéron-like relationship.

We have investigated the use of our basal ganglia-based response mechanism in models of the Stroop task whose front-end networks are modified versions of those in Cohen et al's original model. (Stafford & Gurney 2000, Stafford, 2003). These hybrid models can successfully replicate the basic patterns of Stroop data that were simulated by the Cohen model, but they are also able to deal more successfully with data from variable duration inter-stimulus-interval experiments.

We believe that now is the time to look at how choice theories might usefully be applied to more complex tasks, how choice processes might work in conjunction with other cognitive processes and how existing models might fit with models of other processes and be constrained by neuroscientific evidence and the requirements of the action selection problem. This work was unwittingly begun by Cohen et al (1990) who incorporated a standard choice theory decision mechanism at the 'back-end' of their connectionist model of automatic and controlled processing in the Stroop task. We argue that it is timely to continue this vein of investigation and broaden the scope to include neuroscientific and ethological considerations.

7 References

Bargh, J.A. (1989). Conditional automaticity: varieties of automatic infleunce in social percpetion and cognition. In Bargh, J.A. and Ullman, J. (Eds.) *Unintended Thoughts* (pp 3-51). Guildford Press

Besner, D., Stolz, J.A. and Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin and Review 4*: 221-225.

Besner, D. (2001). The myth of ballistic processing: evidence from Stroop's paradigm. *Psychonomic Bulletin and Review 8*: 324-330.

Bonnet, C., Zamora, M.C., Buratti, F. and Guirao, M. (1999). Group and individual gustatory reaction times and Piéron's law. *Physiology & Behaviour 66*: 549-558.

Brown, T.L., Gore, C.L. and Carr, T.H. (2002). Visual attention and word recognition in stroop color naming: is word recognition "automatic"? *Journal of Experimental Psychology: General 131*: 220-240.

Brown, T.L., Joneleit, K., Robinson, C.S. and Brown, C.R. (2002). Automaticity in reading and the Stroop task: testing the limits of involuntary word processing. *American Journal of Psychology 115*: 515-543.

Brown, T.L., Roos-Gilbert, L. and Carr, T.H. (1995). Automaticity and word perception: evidence from Stroop and Stroop dilution effects. J*ournal of Experimental Psychology: Learning, Memory and Cognition 21*: 1395-1411.

Brown, T.L. (2003). *The relationship between facilitation and interference in the Stroop Task: baselines, asymmetry, inverse correlations and the question of a common mechanism.*(unpublished manuscript)

Carpenter, R.H.S. and Reddi, B.A.J. (2001). Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience 4*: 337-337.

Cohen, J.D., Dunbar, K. and McClelland, J.L. (1990). On the control of automatic processes — a parallel distributed-processing account of the Stroop effect. *Psychological Review 97*: 332-361.

Cohen, J.D., Servan-Schreiber, D. and McClelland, J.L. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology 105*: 239-269.

Dunbar, K. and MacLeod, C.M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance 10*: 622-639.

Duncan, J. (1986). Consistent and varied training in the theory of automatic and controlled information processing. *Cognition 23*: 279-284.

Durgin, F.H. (2000). The reverse Stroop effect. Psychonomic Bulletin and Review 7: 121-125. Ellis, R. and Humphreys, G. (1999). *Connectionist Psychology: A Text with Readings*, Psychology Press, Hove, UK.

Funahashi, K. (1989). On the approximate realisation of continuous mappings by neural networks. *Neural Networks 2*: 183-192.

Gold, J.I. and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences 5*: 10-16.

Gurney, K., Prescott, T.J. and Redgrave, P. (2001a). A computational model of action selection in the basal ganglia I: A new functional anatomy. *Biological Cybernetics 84*: 401-410.

Gurney, K., Prescott, T.J. and Redgrave, P. (2001b). A computational model of action selection in the basal ganglia II: Analysis and simulation of behaviour. *Biological Cybernetics 84*: 411-423.

Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks 2*: 359-366.

Humphries, M.D. and Gurney, K. (2002). The role of intra-thalamic and thalamocortical circuits in action selection. *Network: Computation in Neural Systems 13*: 131-156.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language 30*: 513-541.

Kaplan, W. (1952). *Advanced Calculus*, Addison-Wesley, Reading Mass.

Koch, C. (1999). *The Biophysics of Computation: Information Processing in Single Neurons*, Oxford University Press, New York.

Lagarais, J.C., Reeds, J.A., Wright, M.H. and Wright, P.E. (1998). Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization 9*: 112-147.

Lindsay, D.S. and Jacoby, L.L. (1994). Stroop process dissociations: the relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception & Performance 20*: 219-234.

Logan, G.D. (1988). Towards an instance theory of automatization. *Psychological Review 95*: 492-527.

Luce, R.D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organisation*, Clarendon Press, New York.

MacLeod, C.M. and MacDonald, P.A. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences 4*: 383-391.

MacLeod, C.M. (1991). Half a century of research on the Stroop effect — an integrative review. *Psychological Bulletin 109*: 163-203.

MacLeod, C.M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition 26*: 201-211.

McCloskey, M. (1991). Networks and Theories — the Place of Connectionism in Cognitive Science. *Psychological Science 2*: 387-395.

Pashler, H.E. (1998). *The Psychology of Attention*, MIT Press, Cambridge MA.

Phaf, R.H., Vanderheijden, A.H.C. and Hudson, P.T.W. (1990). Slam — a connectionist model for attention in visual selection tasks. *Cognitive Psychology 22*: 273-341.

Piéron, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensites excitatrices. *L'Annee Psychologique 20*: 17-96.

Piéron, H. (1920). Nouvelles recherches sur l'analyse du temps de latence senorialle et sur la loi qui relie de temp a l'intensite d'excitation. *L'Annee Psychologique 22*: 58-142.

Piéron, H. (1952). *The Sensations: their Functions, Processes and Mechanisms*, Frederick Muller, London.

Pins, D. and Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Perception and Psychophysics 58*: 390-400.

Posner, M.I. and Snyder, C.R.R. (1975). Attention and cognitive control. In Solso, R. (Eds.) *Information Processing and Cognition: The Loyola symposium* (pp 55-86). Erlbaum

Prescott, T.J., Redgrave, P. and Gurney, K. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behaviour 7*: 99-127.

Raaijmakers, J.G.W. and Shiffrin, R.M. (1981). Search of Associative Memory. *Psychological Review 85*: 59-108.

Ratcliff, R. and Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science 9*: 347-356.

Ratcliff, R., Van Zandt, T. and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review 106:* 261-300.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review 85*: 59-108.

Ratcliff, R. (2001). Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience 4:* 336-336.

Reddi, B.A. and Carpenter, R.H. (2000). The influence of urgency on decision time. *Nature Neuroscience 3*: 827-830.

Redgrave, P., Prescott, T.J. and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience 89*: 1009-1023.

Rumelhart, D.E., McClelland, J.L. and The PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, MA.

Ryan, C. (1983). Reassessing the automaticity-control distinction: item recognition as a paradigm case. *Psychological Review 90*: 171-178.

Schall, J.D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience 2*: 33-42.

Sharkey, A.J.C. and Sharkey, N.E. (1995). Cognitive modelling: psychology and connectionism. In Arbib, M. (Eds.) *The Handbook of Brain Theory and Neural Networks First edition* (pp 200-203). MIT Press

Stafford, T. and Gurney, K. (2000). *A computational model of response selection in the Stroop task.* Paper Presented at BPS Cognitive Psychology Section 17th Annual Conference (abstract in proceedings of the British Psychological Society Vol 9, No 2, 2001), Essex, UK.

Stafford, T. (2003). *Integrating psychological and neuroscientific constraints in models of stroop processing and action selection.* PhD thesis, Department of Psychology, University of sheffield, Sheffield, UK.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology 18*: 643-662. *Tyrell, T. (1992).* Defining the action selection problem. Paper Presented at Annual Conference of the Cognitive Science Society,

Tzelgov, J., Henik, A. and Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory and Cognition 20*: 727-735.