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

TOM STAFFORD and KEVIN N. GURNEY
University of Sheffield, Sheffield, England

A response mechanism takes evaluations of the importance of potential actions and selects the most suitable. Response mechanism function is a nontrivial problem that has not received the attention it deserves within cognitive psychology. In this article, we make a case for the importance of considering response mechanism function as a constraint on cognitive processes and emphasized links with the wider problem of behavioral 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., 1990) 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 neurons also support a Piéron-like law.

A response mechanism is the component of any model of human decision making that 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 nontrivial problem (Redgrave, Prescott, & Gurney, 1999). In this article, we argue that response mechanisms are an important part of cognitive models, that their function is an important area for investigation, and that the processes of the response mechanism have consequences for performance in diverse areas of human behavior. The importance of the study of response mechanisms is highlighted by the emergence of the topic in disciplines apart from cognitive psychology. Thus, ethology, robotics, and neuroscience have come to recognize the importance of the "action selection problem" (Prescott, Redgrave, & Gurney, 1999; Tyrrell, 1992). From these perspectives, a response mechanism is necessary to deal with the resolution of conflicts between functional units that are in competition for behavioral expression. For example, a food-deprived animal should reevaluate 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 behavioral selection is also clearly an issue of central importance to cognitive psychologists, but response mechanisms have generally not received their due emphasis. 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 such a minimal feature fulfills the role of the response mechanism, although this kind of mechanism is functionally impoverished (Ratcliff, Van Zandt, & McKoon, 1999; Stafford, 2003). A response mechanism should be designed to cope with the multiple, conflicting demands of the behaving organism (Redgrave et al., 1999). Something about the way these demands are reconciled in humans 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 modeling of decision mechanisms that can mimic the pattern of reaction time in simple choice paradigms (Luce, 1986; Ratcliff & Rouder, 1998). Furthermore, although response mechanisms have been studied in this "choice theory" context, little is known about the brain regions where they might be instantiated, nor has there been an attempt to create models based on neurophysiological principles or constrained by known neuroanatomy.

The significance of response mechanisms has been raised by a recent discussion that 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 dif-

ferent models raises general issues in the study of decision making, such as the issue of exactly which stage of processing the response mechanism is modeling (Carpenter & Reddi, 2001).

The starting point for the investigation of response mechanisms presented in this article is the popular model of the Stroop task described by Cohen, Dunbar, and Mc-Clelland (1990), henceforth the "Cohen model." This is a connectionist implementation of the hypothesis that performance in the Stroop task can be understood with the concepts of automatic and controlled processes, which 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, but its simplicity can be deceptive. The correct explanation for the behavior 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 Cohen 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 (Piéron, 1914, 1920, 1952). We therefore go on to investigate the relation between other response mechanisms (Gurney, Prescott, & Redgrave, 2001a; Phaf, Van der Heijden, & Hudson, 1990; Ratcliff, 1978; Reddi & Carpenter, 2000) and Piéron's law. One such mechanism is consistent with neuroscientific constraints because 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 et al., 2001a, 2001b). We go on to examine the underpinning of a Piéron-like law in this mechanism and its relation to neural functionality.

The Stroop Task

The Stroop task (Stroop, 1935) is a popular paradigm for investigation of the cognitive mechanisms involved in attention, automaticity, and, most important for us, the processing of conflicting stimuli and conflicting responses (for reviews, see MacLeod, 1991, and MacLeod & MacDonald, 2000). The Stroop color-naming task involves responding to the color of a colored word string which can itself be the name of a color. There are three possible general classes of stimuli. For congruent stimuli, the word and the color match (e.g., the word *red* in red ink), for conflicting stimuli, the word and the color 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
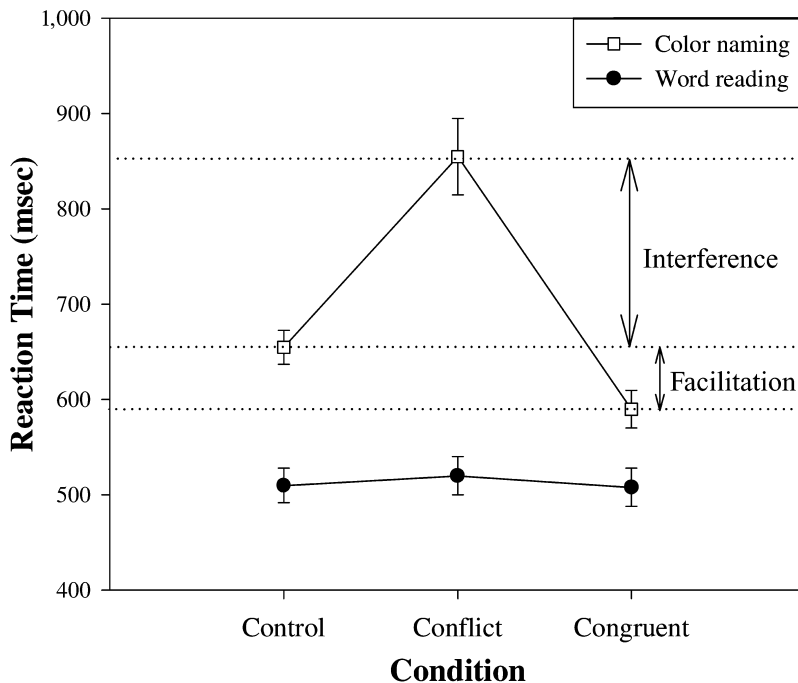

Figure 1. The basic Stroop effect for color naming with nonword control. Note: Interference is greater than facilitation. (Data are from Dunbar & MacLeod, 1984, p. 630.)

when using nonword control stimuli (such as XXXX) are shown in Figure 1 (Dunbar & MacLeod, 1984). This baseline condition is chosen to provide data because it is 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 color naming although, in the complementary word-reading task, the color 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 colors. Recent results show that word reading is not automatic, and words can be successfully ignored if the task conditions are right (Besner, 2001; Besner, Stolz, & Boutilier, 1997; 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 nonword control stimuli is the asymmetry of the influence of congruent versus incongruent information in the color-naming task. The interference effect on reaction times produced by conflicting color-words is greater than the facilitation effect produced by congruent color-words. The Cohen model provides an explanation of this asymmetry, and this explanation 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.

A Model of the Stroop Effect

The model of processing in the Stroop task advocated by Cohen et al. (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-color Stroop task with nonword control (e.g., XXXX in the relevant color) 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 favor of each competing response 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 Cohen 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 the 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 a 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 et al., 1990; Cohen, Servan-Schreiber, & McClelland, 1992)—a matter we return to in the Discussion section. This explanation accompanies the exposition of the model in textbooks (e.g., Ellis & Humphreys, 1999; Sharkey & Sharkey, 1995) and critical reviews (notably, MacLeod, 1991). However, close examination of the explanation illustrated 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.

The True Locus of the Interference–Facilitation Asymmetry Lies in the Response Mechanism

Simulations show irrelevance of the unit activation function nonlinearity. In order to demonstrate that the nonlinearity of the activation function alone cannot 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
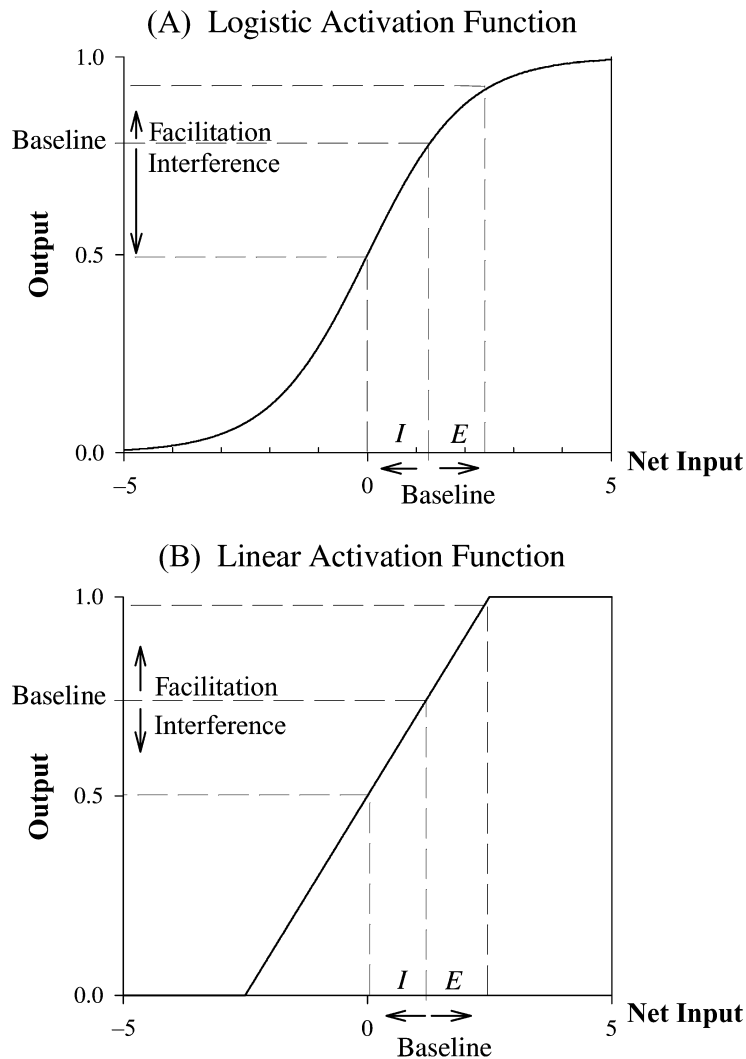
(A) Logistic Activation Function

(B) Linear Activation Function

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.

the essential "squashing" nonlinearity of the logistic function but with a simpler description; it is defined by

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

where ε determines the position of the function on the x-axis, and m the slope of the strictly increasing part. We used values $m = 0.2$ and $\varepsilon = -2.5$ 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 posttraining 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. (1990) 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.

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
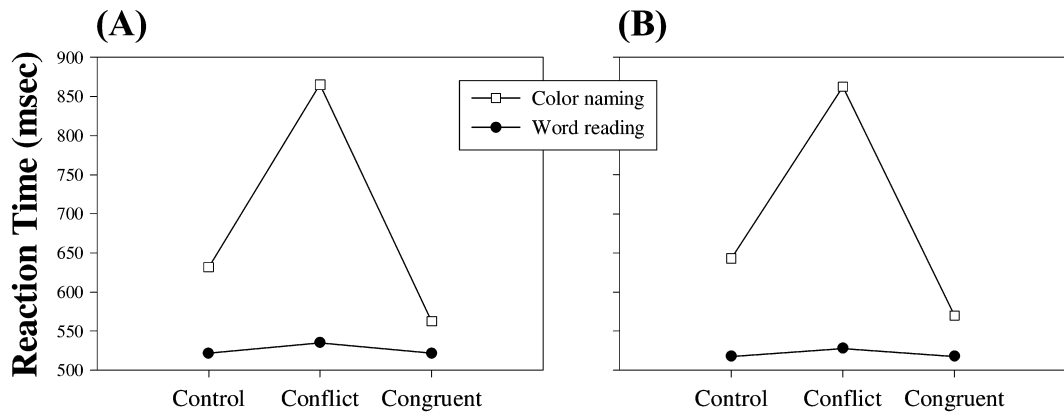
Figure 3. Reaction times on the color-naming task for our replication of the original Cohen, Dunbar, and Mc-Clelland (1990) model, using the original logistic activation function (A) and a piecewise linear activation function (B). The empirical data is accurately simulated with both activation functions.

mechanism of the model calculates reaction times based on the relative strength of evidence—namely, 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, because 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.

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

Furthermore, 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. Given the above observations, 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 creates this phenomenon.

A response mechanism that 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, which is based on evidence accumulation. In a basic version of this scheme, each of the two possible decisions (*red* and *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 be-

Table 1
The Strength of Evidence for Target Response in the Three Color-Naming Conditions for Models Using the Two Activation Functions

| Condition | Logistic Function | | Piecewise Linear Function | |
	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

tween 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

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

where $i \neq j$ and α is a scaling parameter less than 1 that determines the rate of evidence accumulation. The bins are initialized to zero at the start of each trial, and a decision is signaled 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 calculation of 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 seen in Figure 4, which shows that response time is a negatively accelerating function of input. 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 time step n. Without loss of generality, assume that 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 RT, so $RT = cn_I$ where c is a constant. Thus, Equation 1 may be written

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

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 $R_0 = 0$. 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.
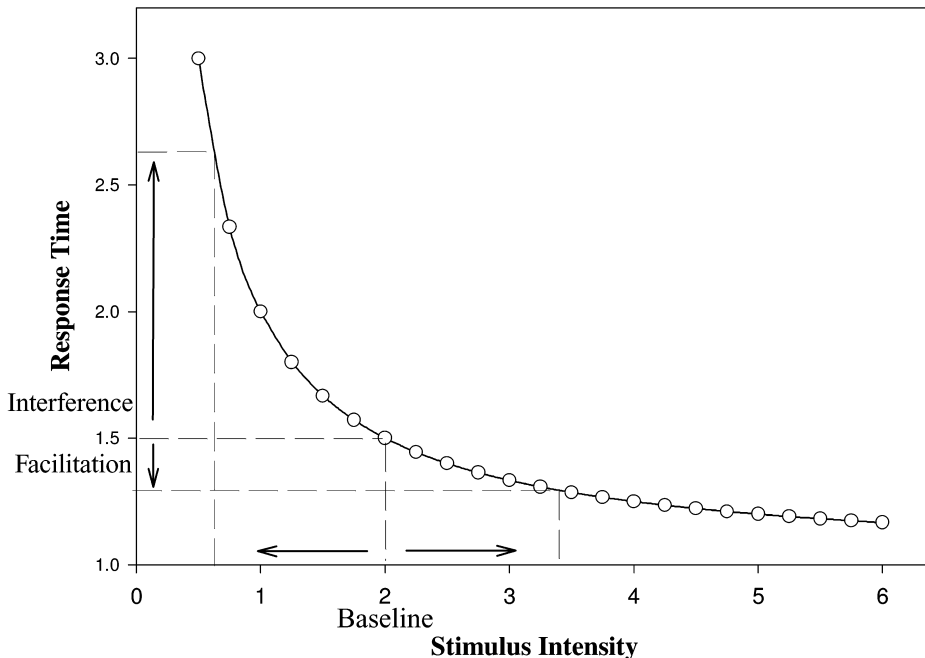

Figure 4. Response time as a function of strength of relative evidence in Cohen, Dunbar, and Mc-Clelland's (1990) response mechanism. The change due to increased intensity is greater than the change due to decreased intensity.

If strength of evidence is replaced by stimulus intensity, Equation 4 corresponds to Piéron's law which describes an early finding from psychophysics that the 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 nonzero 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 to Piéron's law. Such a plot for the Cohen et al. (1990) response mechanism is shown in Figure 5.

The asymptotes for the data and Figures 5–8 were all produced with a procedure outlined below. However, to understand why we adopted this technique, it is necessary to be aware, 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-fitting 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-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 nonlinear optimization procedure that uses the simplex search method (Lagarias, Reeds, Wright, & Wright, 1998). The asymptotic value obtained was used to plot the log of reaction time minus the asymptote on the *y* axis, while the log of the input to the response mechanism was plotted along the *x* axis. Functions that fit Piéron's law exactly produce straight lines when plotted like this. The line shown on the graphs is the best-fitting line that results from using the set of parameters derived from this optimization 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-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 a role analogous to that of stimulus intensity vis-à-vis the relation to reaction time.

To summarize, the differential magnitude of reaction time change under facilitation and interference condi-
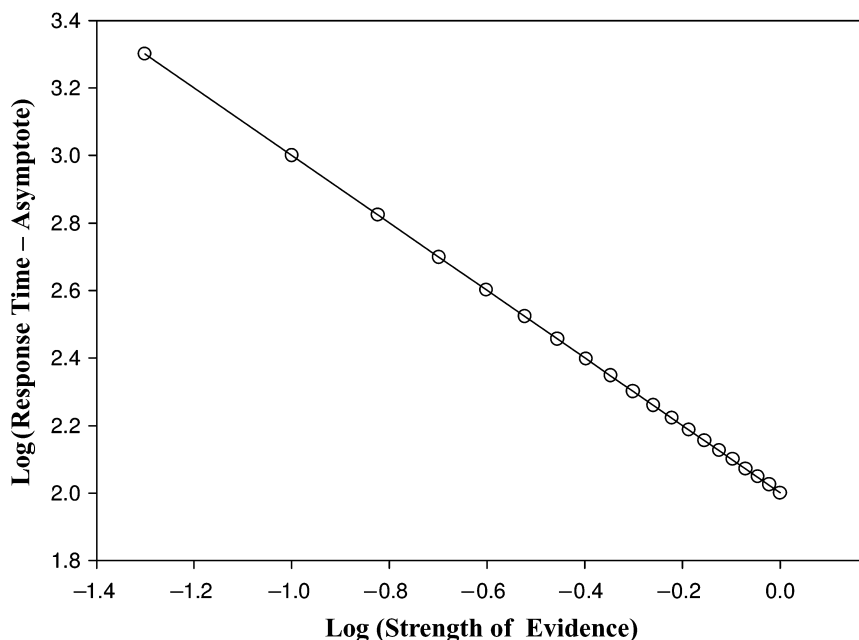

Figure 5. The input–response time function of Cohen, Dunbar, and McClelland's (1990) response mechanism shown on a log–log plot. The line representing the fit to a Piéron-like function is shown to fit the data points exactly. The best-fitting line and asymptote are derived by the standard procedure as defined in the text.
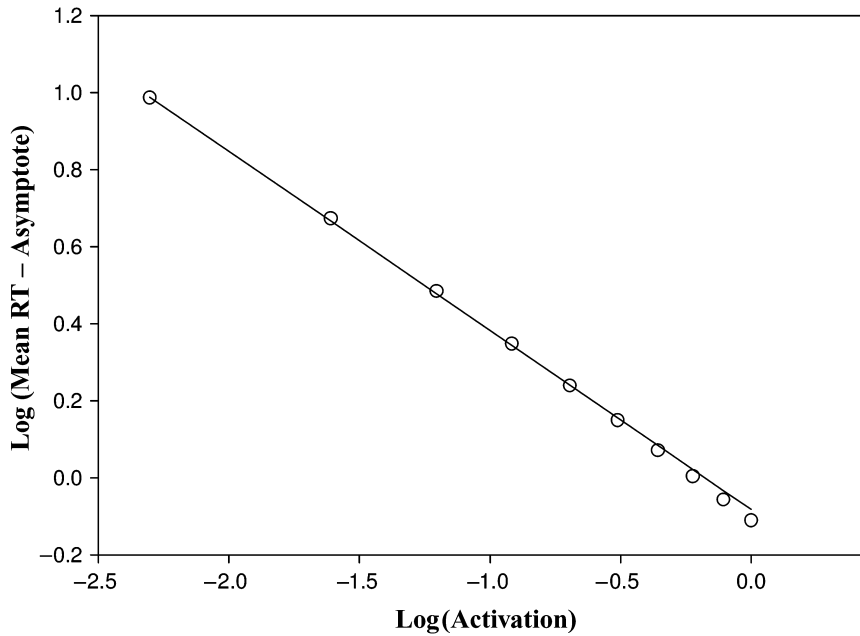
Figure 6. The input–response time function for Phaf, Van der Heijden, and Hudson's (1990) response mechanism shown on a log–log plot. The best-fitting line and asymptote are derived by the standard procedure as defined in the text.

tions 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.

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
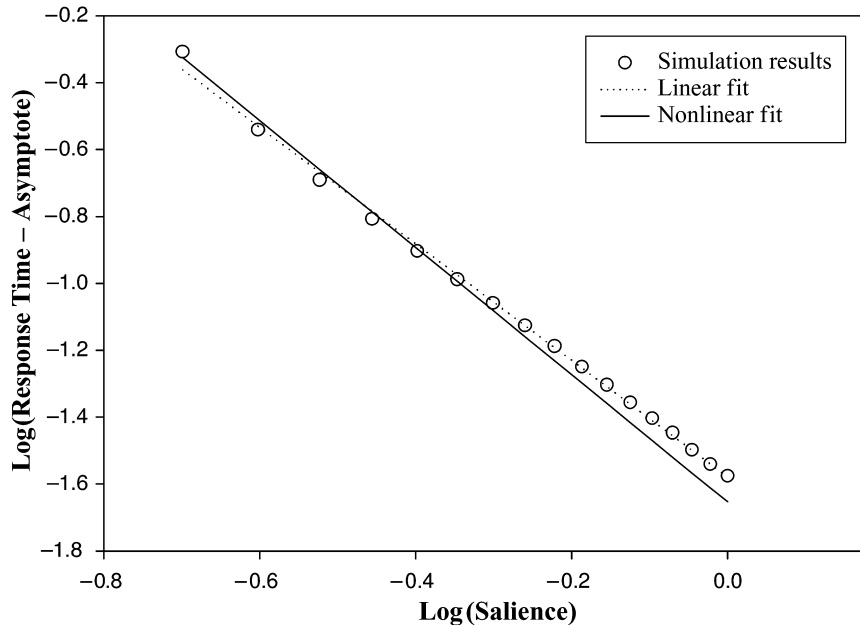

Figure 7. The input–response time function for the basal ganglia selection model shown on a log–log plot. The best-fitting line derived by the standard procedure (as defined in the text) is shown as a solid line. The best-fitting line obtained after the transformation to log–log space is shown as a dotted line.
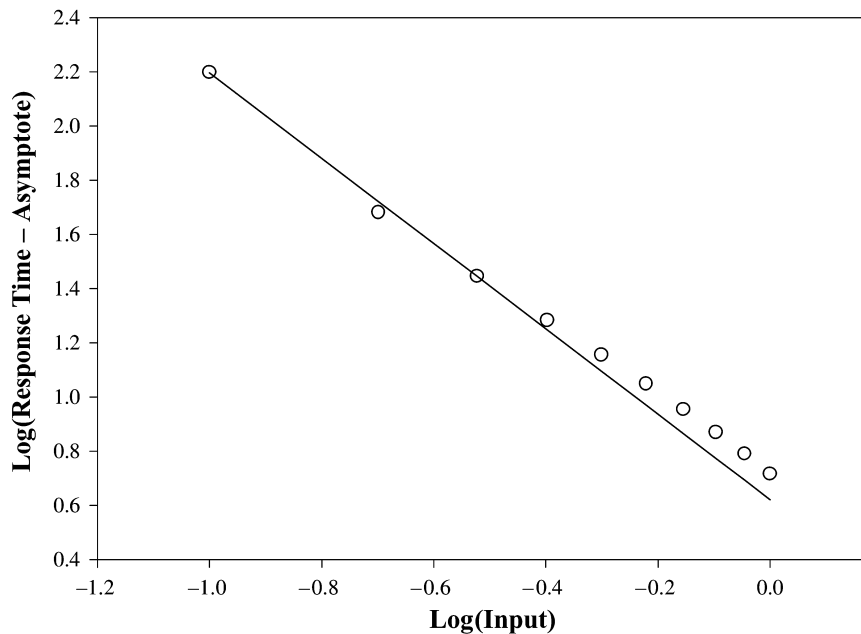
Figure 8. The input–response time function for a simple model neuron derived analytically, shown on a log–log plot. The best-fitting line and asymptote are derived by the standard procedure as defined in the text.

model utilized 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-like form.

The SLAM model (Phaf et al., 1990) also successfully simulates reaction times in the Stroop task. The type of response mechanism it uses is a "sampling and recovery procedure" (Raaijmakers & 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 $a_i(t)$, its probability P_i of forcing a response is given by

$$P_i(t) = 1 - e^{-a_i(t)}.$$

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 $\langle \mathrm{RT} \rangle$ is then just $1/P_i$, so that

$$\langle \mathrm{RT} \rangle = \frac{1}{1 - e^{-a_i}}.$$

To see whether this could be expressed in a form of a Piéron-like law, we sought values of R_0, k, and β that would allow $\langle \mathrm{RT} \rangle$ to be expressed as function of a_i in the form given by Equation 4. These were found using the nonlinear 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 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, 1978; Ratcliff & Rouder, 1998; Ratcliff et al., 1999). 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 toward positive and negative thresholds representing 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, 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 is termed LATER (linear approach to threshold with ergodic rate; Reddi & Carpenter, 2000) 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-like relation of exactly the same form shown in Figure 5.

Biologically grounded decision mechanisms. In addition to models based on matching simulation results to behavioral studies, we have also investigated a response mechanism that is based on the neuroanatomy of a brain system believed to play a crucial role in behavioral response selection—the basal ganglia. The basal ganglia are a subcortical complex of nuclei that we have proposed fulfill the role of a "central switch" in mediating behavioral action selection in vertebrates (Prescott et al., 1999; Redgrave et al., 1999). Briefly, our model of the basal ganglia (Gurney et al., 2001a, 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. Behaviors 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 realize 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 based 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 dotted line 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 nonlinear optimization technique, it is an inferior fit in the untransformed (nonlog–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 that make up the model. Like many other neural network models (including that 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 that 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 of the neuron t_θ 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 Equation 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.

Discussion

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, 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.

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 criticized on empirical grounds because of evidence that shows that interference and facilitation can be differentially affected by experimental manipulations (MacLeod, 1998; Tzelgov, Henik, & Berger, 1992) and that interference and facilitation are not only created by separate processes but at separate stages (Brown, 2003; Brown, Gore, & Carr, 2002; MacLeod, 1998; MacLeod & MacDonald, 2000).

Although this evidence appears to be at odds with the model of Cohen et al. (1990), it is not impossible to conceive of other models that could accommodate these data. The architecture of the Cohen model is a 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 be 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. Although the Cohen model may not be able to support these features, a two-layer connectionist network can be constructed to give any pattern of outputs in response to its inputs (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 multilayer 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 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 and Jacoby (1994) also address the single-mechanism controversy. Their analysis, based on a process dissociation procedure (Jacoby, 1991), focuses on the independent but co-occurring contributions of word-reading and color-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 and Jacoby 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.

Baseline effects. The discussion above notwithstanding, it has been suggested that, under certain circumstances (use of noncolor-word control) interference is *not* greater than facilitation. According to this account, the use of an XXXX control pattern provides a baseline that is faster than a noncolor-word control and thus exaggerates interference effects at the expense of facilitation effects (Brown, 2003; Brown, Gore, & Carr, 2002; Brown, Joneleit, Robinson, & Brown, 2002; Brown, Roos-Gilbert, & Carr, 1995). Baseline effects could be accounted for by a model that took input from noncolor words as well as color words, and in which noncolor-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 two-layer networks 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 underlies it.

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-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-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 Equation 7 which 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 nonlinear intraneuronal processes could combine to create an approximately linear input to the 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.

This neuronal explanation of Piéron's law supposes that a system-level property (Piéron's law) arises by preserving a property of the system's fundamental units of construction involving significant interelement interaction. We call such an 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 are emergent explanations.

The value of the biologically grounded action selection perspective. The shift in locus of explanation for the relative values of facilitation and interference in Cohen et al.'s (1990) model of Stroop performance emphasizes 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 behavioral action selection conducted by all animals in a continuous fashion in their natural state (Prescott et al., 1999; Redgrave et al., 1999). The problem of action selection is therefore central to a study of human behavior, and so we are not surprised to find mechanisms mediating its solution occupying a central place in understanding laboratory-based cognitive tasks. Furthermore, although 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 et al., 2001a, 2001b; Humphries & Gurney, 2002) is based on biological considerations and was constructed to solve the problem of selection. However, it can 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 (1990) original model (Stafford, 2003; Stafford & Gurney, 2000). 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 interstimulus 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 to broaden the scope to include neuroscientific and ethological considerations.

REFERENCES

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. A. Bargh & J. S. Uleman (Eds.), *Unintended thoughts* (pp. 3-51). New York: Guilford.

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

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

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

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.

Brown, T. L., Gore, C. L., & 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., & 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., & Carr, T. H. (1995). Automaticity and word perception: Evidence from Stroop and Stroop dilution effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1395-1411.

Carpenter, R. H. S., & Reddi, B. A. J. (2001). Reply to "Putting noise into neurophysiological models of simple decision making." *Nature Neuroscience*, **4**, 337.

Cohen, J. D., Dunbar, K., & 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., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology*, **105**, 239-269.

Dunbar, K., & MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception & 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 & Review*, **7**, 121-125.

ELLIS, R., & HUMPHREYS, G. W. (1999). *Connectionist psychology: A text with readings*. Hove, U.K.: Psychology Press.

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

GOLD, J. I., & SHADLEN, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, **5**, 10-16.

GURNEY, K., PRESCOTT, T. J., & 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., & REDGRAVE, P. (2001b). A computational model of action selection in the basal ganglia: II. Analysis and simulation of behavior. *Biological Cybernetics*, **84**, 411-423.

HORNIK, K., STINCHCOMBE, M., & WHITE, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359-366.

HUMPHRIES, M. D., & GURNEY, K. (2002). The role of intra-thalamic and thalamocortical circuits in action selection. *Network*, **13**, 131-156.

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

KAPLAN, W. (1952). *Advanced calculus*. Reading, MA: Addison-Wesley.

KOCH, C. (1999). *The biophysics of computation: Information processing in single neurons*. New York: Oxford University Press.

LAGARIAS, J. C., REEDS, J. A., WRIGHT, M. H., & 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., & 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 organization*. New York: Oxford University Press, Clarendon Press.

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.

MACLEOD, C. M., & 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.

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*. Cambridge, MA: MIT Press.

PHAF, R. H., VAN DER HEIJDEN, A. H. C., & 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 intensités excitatrices. *L'Année Psychologique*, **20**, 17-96.

PIÉRON, H. (1920). Nouvelles recherches sur l'analyse du temps de la-

tence sensorielle et sur la loi qui relie de temp à l'intensité d'excitation. *L'Année Psychologique*, **22**, 58-142.

PIÉRON, H. (1952). *The sensations: Their functions, processes, and mechanisms*. London: Müller.

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

POSNER, M. I., & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-86). Hillsdale, NJ: Erlbaum.

PRESCOTT, T. J., REDGRAVE, P., & GURNEY, K. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behavior*, **7**, 99-127.

RAAIJMAKERS, J. G. W., & SHIFFRIN, R. M. (1981). Search of associative memory. *Psychological Review*, **88**, 93-134.

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.

RATCLIFF, R., & ROUDER, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, **9**, 347-356.

RATCLIFF, R., VAN ZANDT, T., & MCKOON, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, **106**, 261-300.

REDDI, B. A., & CARPENTER, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience*, **3**, 827-830.

REDGRAVE, P., PRESCOTT, T. J., & GURNEY, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, **89**, 1009-1023.

RUMELHART, D. E., MCCLELLAND, J. L., & THE PDP RESEARCH GROUP (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (2 vols.). Cambridge, MA: MIT Press.

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., & SHARKEY, N. E. (1995). Cognitive modeling: Psychology and connectionism. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 200-203). Cambridge, MA: MIT Press.

STAFFORD, T. (2003). *Integrating psychological and neuroscientific constraints in models of Stroop processing and action selection*. Unpublished doctoral dissertation, University of Sheffield.

STAFFORD, T., & GURNEY, K. (2000, September). *A computational model of response selection in the Stroop task*. Poster presented at the BPS Cognitive Psychology Section 17th Annual Conference, Exeter.

STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

TYRRELL, T. (1992). *Defining the action selection problem*. Paper presented at the Fourteenth Annual Conference of the Cognitive Science Society.

TZELGOV, J., HENIK, A., & BERGER, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, **20**, 727-735.