Strategic regulation of cognitive control by emotional salience: A neural network model

Bradley Wyble [a]; Dinkar Sharma [b]; Howard Bowman [b]
[a] Massachusetts Institute of Technology, Cambridge, MA, USA [b] University of Kent at Canterbury, Canterbury, Kent, UK

PLEASE SCROLL DOWN FOR ARTICLE

Strategic regulation of cognitive control by emotional salience: A neural network model

Bradley Wyble

Massachusetts Institute of Technology, Cambridge, MA, USA

Dinkar Sharma and Howard Bowman

University of Kent at Canterbury, Canterbury, Kent, UK

We present a neural network model of stimulus processing, which uses a mechanism of adaptive attentional control to regulate the moment to moment deployment of attention according to both the demands of the current task, and the demands of emotionally salient information. This mechanism allows negative emotional information to reduce cognitive control to aid in the detection of threats, which produces a momentary withdrawal from the current task set to allow unbiased processing of available information. The combination of cognitive and emotional regulation of task set allows this model to address inter-trial aspects of emotional interference in colour naming. In particular, we focus on the nature of the emotional interference in colour naming (McKenna & Sharma, 2004) as well as in word reading (Algom, Chajut, & Lev, 2004) and show how this form of interference is functionally distinct from the classic Stroop effect. Our model addresses a range of findings in colour naming and word reading tasks and is informed by recent neuroimaging data concerning the interaction between the anterior cingulate and prefrontal cortices. The model is used to explore the interface between cognition and emotion with a series of predictions, including a qualitative distinction between state and trait forms of anxiety.

INTRODUCTION

The human cognitive system is excellent at prioritising competing processing demands in real time. However, the process of pursuing long-term goals and the need to respond optimally and immediately to incoming environmental

stimuli are mutually incompatible. There is thus a trade-off between the need to meet goals and the need to respond optimally according to the salience level of environmental stimuli.

This tradeoff may exist between cognitive control, thought to regulate effort in response to task difficulty (Botvinick, Braver, Barch, Carter, & Cohen, 2001), and emotional effects, thought to reorient processing resources to potentially threatening information, such as angry faces (Öhman, Lundqvist, & Esteves, 2001b) or dangerous organisms (Öhman, Flykt, & Esteves, 2001a).

In this paper we address the putative relationship between task-based cognitive control and emotionally salient information by presenting a computational model of adaptive attentional control. This model replicates key data regarding both the Stroop effect and emotional interference in an effort to elucidate the functional distinction between these effects. The intent of this paper is to show how cognitive and emotional priorities trade off in a manner that adapts to changing task and environmental demands from one moment to the next.

Stroop and emotional interference

Stroop effects (interference or facilitation of identification for multi-dimension stimuli) are thought to result from the semantic relationship of one dimension of a stimulus, e.g., a word, and another, e.g., ink colour (Stroop, 1935; MacLeod & MacDonald, 2000).

Emotional stimuli also interfere in cognitive tasks, such as colour naming and word reading. This effect is typically measured as a difference in response latencies when colour naming negative emotional words (e.g., "REJECTED", "FAILURE", "SUICIDE") compared to neutral words. This effect is most clearly identified when the emotional word has a particularly pronounced significance for the subject, e.g., the word SPIDER for a spider phobic (Watts, McKenna, Sharrock, & Trezise, 1986).

The time course of emotional interference

One important difference between Stroop and emotional interference concerns their respective time courses. Emotional interference during colour naming has a prolonged time course, producing a slowing of response times that carries from one trial to a following trial. In McKenna and Sharma (2004), participants performed seven colour-naming trials in one of two particular sequences. In an emotional sequence, the first trial was a negative word followed by six trials of neutral words. Subjects were not slowed at naming the colour of the negative word but were slowed on naming the colour of the following neutral word. Based on these findings, McKenna and Sharma argued for the existence of fast (intra-trial) and slow (inter-trial)

components in this type of task. The slow component corresponds to interference that can only be observed across trials.

These findings have important implications because most experimental investigations have been predicated upon the idea that there is no such inter-trial effect. Interference effects that operate across trials can explain why larger interference for colour naming emotionally negative stimuli is found when these stimuli are presented in separate blocks compared to when they are mixed randomly with neutral stimuli (Holle, Neely, & Heimberg, 1997; Richards, French, Johnson, Naparstek, & Williams, 1992).

Based on a meta analysis of 70 emotional Stroop studies Phaf and Kan (2007) have provided further support for the presence of a slow emotional effect. They conclude that "the emotional Stroop effect seems to rely more on a slow disengagement process than on a fast, automatic, bias" (p. 184). The slow effects have also been generalised to a wider range of stimuli, in particular addiction-related words (Albery, Sharma, Niyazi, & Moss, 2006; Waters, Sayette, Franken, & Schwartz, 2005).

Emotional interference in word reading

Another key difference between Stroop and emotional interference is that the latter occurs even when the emotional stimulus is task relevant. This can be demonstrated as a slowing in the reading of emotional words (Algom, Chajut, & Lev, 2004).

MODELLING STROOP AND EMOTIONAL INTERFERENCE

Computational models can serve a valuable role in explaining these forms of interference. Our own work builds on a tradition of prior models including Cohen, Dunbar, and McClelland (1990), Cohen and Huston (1994), Matthews and Harley (1996), and Botvinick et al. (2001).

Such models involve parallel processing of word and colour information in pathways that are connected to corresponding nodes in a response layer. When ink colour and word conflict, interference, implemented by inhibitory interactions, delays the point at which response nodes cross threshold. Facilitation effects from congruent trials result from convergence of both streams onto a single response node.

Another feature of these models is that the word pathway has stronger connections to response nodes than the ink-colour pathway, producing a default tendency to read the word over naming its ink colour. This is argued to result from an asymmetry in practice; we spend more time reading than colour naming. A task-demand system can intervene to counteract this tendency and allow the system to name ink colour over words, but the stronger connections in one pathway provide faster performance than

the other. Another common feature of Stroop studies, that facilitation effects may be weaker than interference, is simulated by competition between response units (Stafford & Gurney, 2004).

More recent research has demonstrated that experimental context can change the size of the Stroop effect as well as the relative magnitudes of both Stroop interference and Stroop facilitation (cf. Melara & Algom, 2003; Melara & Mounts, 1993). Although we do not simulate such effects it would be possible to do so by varying either the connection weights between the word form and colour form pathways or the strength of input.

Computational accounts of emotional interference

Previous models have simulated interference from negative emotional stimuli by assuming that their representations in the word pathway have an excitatory bias, giving them a greater ability to interfere with the naming of ink colour than neutral words (Matthews & Harley, 1996). Battye (2004) described a model designed around a similar set of mechanisms that exhibits emotional interference arising from either state (a threat monitoring node) or trait (increased weights from emotional nodes) anxiety. However, attentional bias towards negative emotional stimuli is incapable of explaining interference in the slow component (i.e., an inter-trial effect) and emotional interference for task-relevant emotional stimuli. To accommodate these findings we turn to an account of cognitive control.

Addressing the slow component

There is compelling evidence that the brain does not treat each Stroop trial as an independent event. Data from Tzelgov, Henik, and Berger (1992) show that the relative frequency of incongruent, congruent and neutral trials within a block has an effect on the magnitude of Stroop interference. As the proportion of colour trials (equally divided among congruent and incongruent) increases, Stroop interference decreases. This point was investigated in depth by Dishon-Berkovits and Algom (2000), Sabri, Melara, and Algom (2001), and Melara and Algom (2003).

Cognitive control

A model that can accommodate these block composition effects is described by Botvinick et al. (2001) through the addition of a cognitive control mechanism, which modulates the strength of top-down control from trial to trial based on the cumulative difficulty of previous trials. In their work, cognitive control is mediated by a feedback loop that measures task difficulty as a function of conflict (i.e., cross-talk between competing colours) in the response layer. Control is increased when conflict is high,

which is characteristic of incongruent trials. When the task is colour naming, an increase in cognitive control directs attention to the colour-naming pathway at the expense of word reading. As a result of this inter-trial carryover of control settings, this model can explain the data from Tzelgov et al. (1992).

A model by Melara and Algom (2003) described a means by which unintended variations in the correlation between colour and word dimensions can force changes in attentional allocation that cause trial block effects similar to Tzelgov et al. (1992) as well as Dishon-Berkovits and Algom (2000). This is a competitive account to that of cognitive control, but it is critical to note that our account of emotional interference is not incompatible with the Melara and Algom model.

A new account: Adaptive attentional control

Here we describe a model that simulates both Stroop and emotional forms of interference as distinct phenomena. In our account, prioritisation of processing is mediated in a way that adapts to the current context of the subject, including task difficulty, but also environmentally salient stimuli.

Emotional interference in tasks like colour naming and word reading results from a reduction of cognitive control following exposure to emotional stimuli. Thus, the degree of task involvement is adaptively regulated by environmental information, particularly that which might be associated with threats. Such adaptation is a defensive mechanism, allowing an organism to efficiently locate the source of an environmentally salient signal by relaxing the current task-specific processing bias. In effect, a potentially dangerous stimulus causes the attentional system to withdraw, momentarily, from the current task.

PAPER OVERVIEW

The purpose of this paper is two fold. First, in describing both Stroop and emotional forms of interference within the same model, it is demonstrated how they can be viewed as functionally distinct effects. Second, having developed a computational model that integrates cognitive and emotional function, we demonstrate its ability to generate predictions for novel tasks that explore the interface between cognitive and emotional forms of interference.

In the following section we explain the general function of the model, with details and parameter settings left for the Appendix. The model will be shown to replicate interference and facilitation within the Stroop task (Dunbar & MacLeod, 1984) including inter-trial effects (Tzelgov et al., 1992). The same model will be compared with recent neuroimaging data

from Kerns et al. (2004) to provide constraints on the duration of inter-trial effects. We then turn to emotional interference; using the same parameters, we replicate the inter-trial effect from McKenna and Sharma (2004), and the effect of emotional stimuli on word reading. After describing the model's ability to replicate these findings, it is put to use in generating predictions about how cognitively challenging and negative emotional stimuli affect the balance of control in the deployment of attention. Furthermore, the model is used to explore a putative distinction between trait and state forms of anxiety. Finally, we conclude with a discussion of the model's relation to previous theoretical and experimental work regarding emotion, attention, and recent findings in neuroimaging.

MODEL IMPLEMENTATION

The model we propose simulates the reaction times of subjects attempting to name the colour or read the word of a stimulus. Accuracy is generally very good in these experiments so parameters are set to provide nearly perfect performance.

Our adaptive attentional control model, pictured in Figure 1, inherits both the activation equations and the feed forward pattern of the connectionist architecture for Stroop models proposed by Cohen et al. (1990), that consists of parallel word and ink-colour processing streams.

Emotional Regulation of Cognitive Control
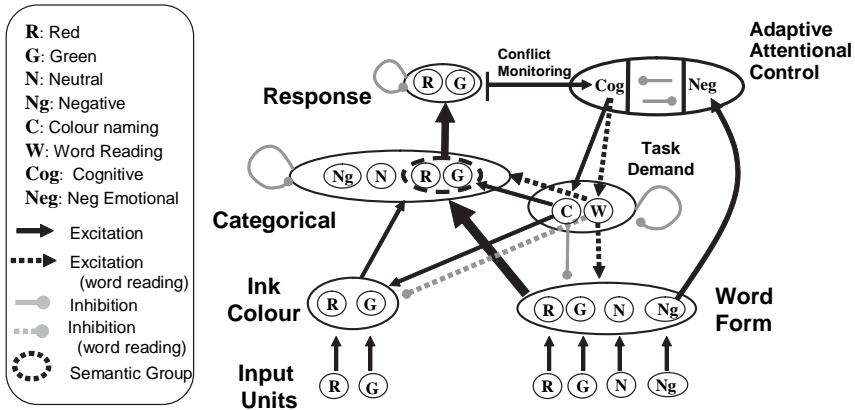
Figure 1. Processing layers are denoted by large ovals. Circles indicate neural nodes, labelled according to type. Connections touching a layer project to the entire layer, either as an identity matrix between neurons of similar types (e.g., "R" to "R"), or from one node to multiple nodes (only in the case of task demand). Note that this version of the model is configured for colour naming. For word-reading tasks, the dashed lines to and from the word ("W") would be activated.

A task-demand system configures the model to produce one response or the other and is driven by an adaptive attentional control circuit that allows both task difficulty and negative emotional input to regulate the amount of top-down control. We use primarily feed-forward connections between layers, a modification that eliminates the recurrent dynamics of the Botvinick et al. (2001) and Cohen and Huston (1994) models, and avoids the need to suppress the response layer while a response is settled on.

Architecture

Our model is organised into layers of processing in accordance with local processing similarity. Nodes in the colour and word-form layers are activated directly by colour and word stimuli. The word-naming stream also contains nodes representing neutral and negative emotional words.

These colour and word-form nodes project to corresponding nodes in the category layer, intended to represent activity in a semantic space. This categorical level of representation enables us to separate motor response and conceptual interference. The separation allows our model to replicate Stroop interference for stimuli lacking response competition, such as semantic associates (Klein, 1964), and also plays a role in simulating emotional interference in word reading.

Interference occurs within the category and response layers through lateral inhibitory connections between nodes. If two nodes are activated, they will inhibit one another, ultimately increasing the time for response nodes to cross output threshold.

Responses

Nodes of the category layer project to the response layer, which ultimately determines the response of the network. Activation of the response nodes builds over the trial until one of them crosses threshold, at which point a reaction time is calculated. In models of this sort, processing cycles are seen as an analogue of response time (Cohen et al., 1990). Furthermore, the model is capable of committing an error if the incorrect response node crosses threshold first. Errors in the sort of task that we model here are generally rare under normal instructions, so parameters were set to produce zero error rates, but this is by no means critical to our implementation.

Task demand

A pair of task-demand nodes configures the system to perform colour-naming or word-reading tasks. The colour-naming task-demand node provides excitatory bias to the colour pathway and suppresses word-form processing, while the word-naming task-demand node does the opposite.

Task demand also biases activation in the category layer for the valid responses; colours for colour naming, and all nodes for word reading.

Adaptive attentional control

The model configures which of the task-demand nodes is active at the start of a trial block and then regulates the moment to moment deployment of task demand using the following two mechanisms.

Cognitive control

Cognitive control provides excitatory input to one of the two task-demand nodes according to experimental instructions. In the absence of input, the cognitive node has an output of 0.5, providing a tonic level of activation to the correct task-demand node, which sustains the task set.

However, the presence of conflict at the response layer (defined as two nodes being coactive) increases the activation of the cognitive control node, which in turn excites the appropriate task-demand unit and further biases processing towards one of the processing pathways to resolve the conflict.

As in the Botvinick et al. (2001) model, a Hopfield energy equation (1) over the two response nodes provides a measure of conflict.

$$E_{i,j,t} = \sum_{i,j} out_{i,t} out_{j,t} \tag{1}$$

In this equation, $out_{i,t}$ is the output of response node i at time t. The output of this equation is passed through a logistic equation (see Appendix) and the result is multiplied by a weight before becoming an excitatory input to the cognitive component of the adaptive attentional control mechanism.

Unlike Botvinick et al. (2001), our model updates all nodes concurrently. There are no trial-wise steps, activation is just carried over from the end of one trial to the next.

Emotional regulation of cognitive control

The negative word-form input has an excitatory link to the negative emotional node ("Neg") of attentional control, which, in turn, inhibits the cognitive node. Thus, presentation of a negative emotional word can inhibit cognitive control, thereby reducing excitation of the appropriate task-demand node and producing a withdraw from the task set. This effect is slow relative to the performance of a single trial, in that changes in task demand produced by negative words are generally not rapid enough to affect the current trial; the response node has already reached threshold by the time this feedback circuit can strongly affect the system. These excitatory

connections to the negative emotional node are specific to words of a negative emotional valence, which we assume are associated with threat.

Modelling without noise

Unlike many previous accounts, our model avoids the use of artificial noise; a given series of trials always produces the same output. Variance arises due to inter-trial effects from randomly ordered stimuli in mixed blocks. To compute average reaction times for trial types in a mixed-block design, the model is repeatedly run for blocks of 10 trials randomly selected from the chosen proportion of trial types (e.g., 33.3% congruent, 33.3% incongruent, 33.3% neutral for Dunbar & MacLeod, 1984). To demonstrate that simulated outputs are reliable for such mixed-block paradigms, the model is repeatedly run to simulate 10 subjects. From these results, the mean and standard error are computed. These means are shown in the figures, including error bars that reflect standard error (standard deviation divided by square root of 10).

For experiments involving a pure block or a specific trial sequence (as in simulating the slow component of emotional interference in colour naming; McKenna & Sharma, 2004) it is only necessary to simulate a single series of trials, as the model's behaviour would be identical during repetitions.

Parameters

Many of the Stroop effects simulated here have been replicated previously (Botvinick et al., 2001; Cohen et al., 1990) but we include them because they provide strong constraints on the allowable range of parameters. In particular, the Stroop effect requires weights that heavily favour word reading over colour naming by default, and task demand compensates by biasing activation towards colour naming.

Parameters were fixed at a single set of values for all of the simulations reported below. Then, reaction time of the model was measured in terms of the number of cycles between trial onset and crossing of threshold by any of the response nodes. Cycles were converted to simulated milliseconds by applying a slope and intercept as described by Cohen et al. (1990). These parameters are reported in the figures, and serve to address the effect of methodological differences and other sources of variance between experiments, such as response production (e.g., word pronunciation vs. button pressing), inter-stimulus interval, and presentation format. Critically, while the slope and intercept vary, the slope was not permitted to be negative, ensuring that this parameter fitting cannot qualitatively reverse a trend in the model's output.

SIMULATION RESULTS

We now describe critical experimental results and how the model addresses each of them.

Simulating the Stroop effect

We simulate classic Stroop effects as do previous computational accounts. One critical effect is that incongruence produces a larger change in reaction time than congruence. Another is that word-reading tasks are uniformly faster than colour-naming tasks, and exhibit minimal interference or facilitation. This pattern is demonstrated by the data of Dunbar and MacLeod (1984) in Figure 2. The model replicates the following features of their data reported in Experiment 1a. Incongruent colour naming is slower than neutral and congruent colour naming. Colour naming is slower than word reading in all conditions. Word reading of congruent stimuli is slightly faster than reading neutral and incongruent trials (by 2.5 ms, the effect, though small in simulation, is reliable). In our simulations, neutral trials involved activation of a colour node in one pathway and no corresponding colour node in the other pathway.

Simulating semantic neighbourhood Stroop effects

Interference within the category layer can produce a semantic gradient in Stroop interference as described by Klein (1964). Words semantically associated with specific colours (e.g., "sky" or "banana") are capable of producing Stroop interference despite not having an associated response node (Klein, 1964).
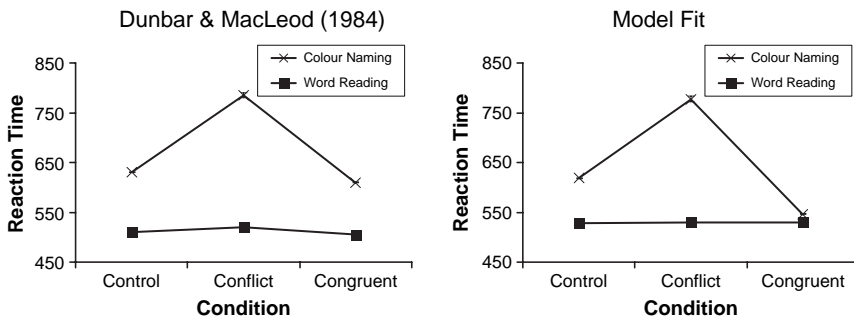

Figure 2. Comparison of model results to empirical data for classic Stroop interference exhibiting stronger interference than facilitation for colour naming and a lack of effects in word reading. Data replotted from Figure 1 of Dunbar and MacLeod (1984), with permission. Standard error for the simulated conditions are in the range [0.2 to 5 ms]. RT = Cycles *1.5 + 320.

In our implementation, semantic interference is provided by virtue of the task-demand system, which, we argue, is incapable of biasing a task-relevant category of words (e.g., colour words) without that activation spilling over to semantic neighbours. We simulate this by connecting the colour-naming task-demand node to a category-layer node representing a semantic associate of a colour, which has no corresponding response node. Consequently, a semantic associate of a colour word is capable of interfering with the naming of ink colour, but not as much as a colour word, as shown in Figure 3.

The ability to replicate semantic Stroop effects is one benefit of including a categorical level of representation. This is a distinguishing feature of our model compared with those of Botvinick et al. (2001) and Cohen and Huston (1994), which lack an intermediate layer of nodes organised by semantic similarity.

Simulating block-composition effects

The composition of a trial block affects how the trials within that block are processed. In Tzelgov et al. (1992), subjects were shown blocks containing different proportions of colour-neutral words in steps of 25%. Of the colour-word trials, half were congruent and half were incongruent. So, a 75% colour-word trial block was composed of 37.5% congruent, 37.5% incongruent, and 25% neutral trials. Their results demonstrated that as the proportion of colour words is increased within a block, the amount of interference caused by an incongruent trial decreases, while facilitation remained unaltered as shown in Figure 4 (i.e., the difference in facilitation across blocks was not significant in their experiment).

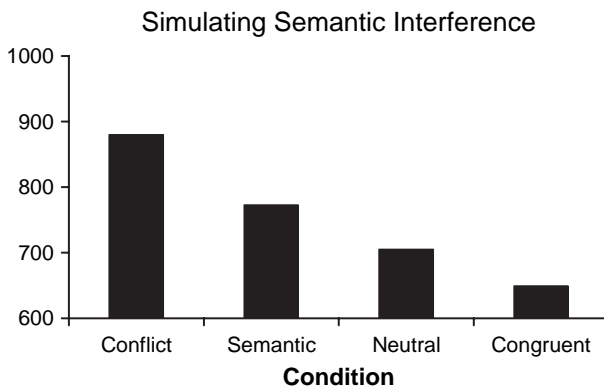**Figure 3.** Simulation of semantic interference by a word node that receives a bias input in the category layer but has no corresponding response node, representing a semantic associate of a colour (e.g., banana). The corresponding reaction time is elevated relative to a neutral word due to competition within the category layer. These simulations are deterministic; there is no variance. RT = Cycles *1.0+500.
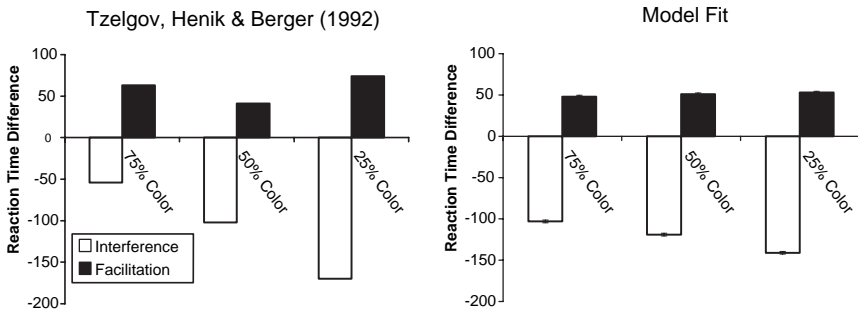
Figure 4. Demonstrating a qualitative fit of data from Tzelgov et al. (1992), in which varying the relative proportion of colour and neutral trials in blocks can parametrically affect the amount of interference observed. The interference and facilitation bars each represent differences in reaction times for incongruent and congruent trials, respectively, in comparison with control trials. Data replotted from Figure 1 of Tzelgov et al. (1984), with permission. Error bars are provided for the simulated data. RT = Cycles *1.0 + 500.

Conflict-driven cognitive control replicates this pattern of results. Each incongruent trial produces conflict, which ultimately excites task demand. This effect temporarily increases the bias towards the colour pathway and reduces interference in future incongruent trials.

Therefore, as the proportion of incongruent trials within a block increases, the probability of any incongruent word being preceded by another incongruent word increases and, consequently, the average amount of interference over all incongruent trials declines. The facilitation produced by congruent trials does not vary appreciably in simulated results.

The time course of cognitive control

Evidence from trial-based analyses of brain-imaging data can provide critical constraints on modelling. The anterior cingulate cortex (ACC) is part of a network thought to participate in the mediation of attentional allocation in cognitively demanding tasks. For example, increased activation of the ACC is found during incongruent Stroop trial blocks relative to congruent blocks using PET (Pardo, Pardo, Janer, & Raichle, 1990). More recent work has suggested that the ACC is involved in maintaining the task set, as well as responding to variation in task difficulty or error commission (Dosenbach et al., 2006; Weissman, Gopalakrishnan, Hazlett, & Woldorff, 2004). This is precisely the role that the cognitive control node in our model reflects; it maintains a baseline level of excitation to the task-demand node, which is elevated in the presence of difficulty.[1] Other research suggests that a region of dorsolateral prefrontal cortex (DLPFC) plays a role in imposing

---

[1] Error trials in our model would also be associated with elevations in the conflict measure.

the task set (MacDonald, Cohen, Stenger, & Carter, 2000), which is broadly analogous to the task-demand nodes in our model (see also Cohen & Servan-Schreiber, 1992).

Under the assumption that the ACC and PFC perform analogous functions to cognitive control and task demand, we used imaging data to both constrain and test our implementation. Kerns et al. (2004) provided data that looked at ACC and PFC activation in a Stroop task for the four possible pairings of congruent and incongruent trials within a mixed block (i.e., iI, iC, cI, cC, where iC refers to a congruent trial preceded by an incongruent trial). Both ACC activation and interference (i.e., a longer reaction time) were larger for an incongruent trial that followed a congruent trial (i.e., cI) than for an incongruent trial that followed an incongruent trial (i.e., iI).

These data informed the time course of the nodes involved in adaptive attentional control (both cognitive and emotional were assumed to have the same course) such that the influence of conflict built up and decayed within a single trial while the task-demand nodes changed more slowly, allowing activation in one trial to affect a subsequent trial. This is unlike the Botvinick et al. (2001) model, in which cognitive control accumulates activation over the course of multiple trials. Note that no parameters (apart from slope and intercept to match RTs) were changed for any of the experiments reported within this paper; this fitting was done initially.

In testing the model, a comparison between two-trial sequences produces a similar pattern both for behaviour (Figure 5a) and output of the cognitive control node (Figure 5b). The conflict produced by one incongruent trial increases task demand such that an immediately following incongruent trial produces minimal behavioural interference, and minimal conflict. During a congruent trial, task demand decays sufficiently so that a following incongruent trial again produces interference and conflict. Thus, our simulation suggests that a more rapid time course of task monitoring, which considered primarily the current trial, is best suited to fit ACC activation in Stroop tasks. Note that in the simulated results (Figure 5b) the cognitive control node is always partially active during performance of a task.

Having set the time constants of cognitive control and task demand, we attempted to simulate the second set of imaging data in Kerns et al. (2004), to determine if these two nodes interact in a similar way as the ACC and PFC regions do in the neuroimaging data. In their analysis, activation of the PFC in each trial was compared to the activation of the ACC in the previous trial, and a nearly linear relationship was observed (Figure 5c). Thus, the state of PFC seems to be determined, in part, by the difficulty of the immediately preceding trial, as indexed by ACC activation.

To perform this same analysis in the model, the averaged outputs of the cognitive-control node and the task-demand nodes were computed for each
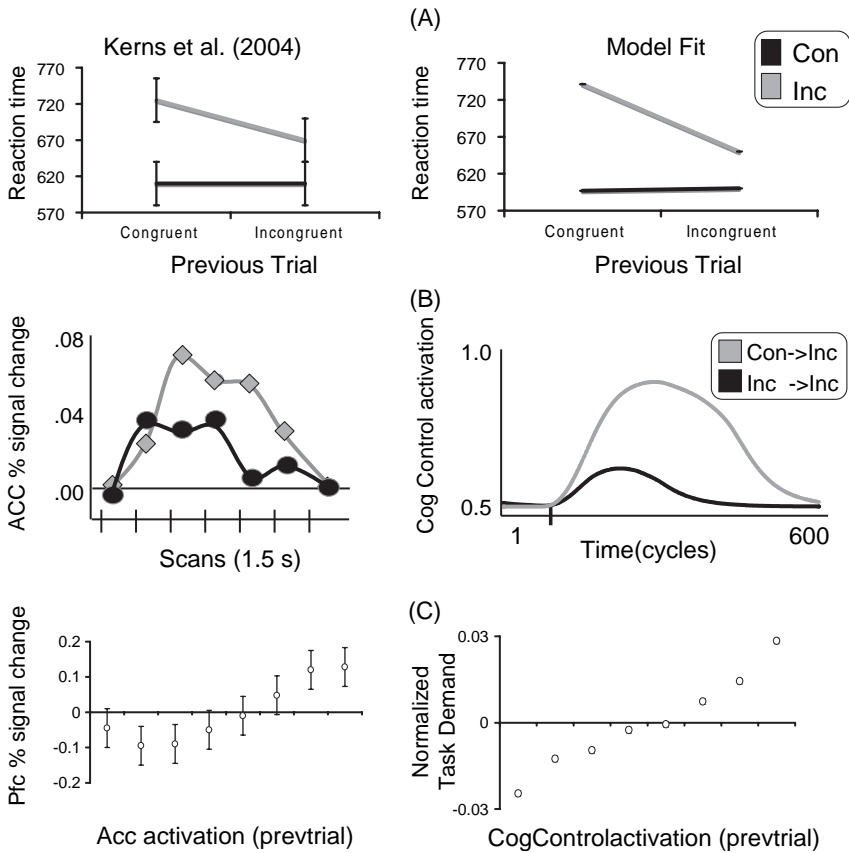
Figure 5. Replication of behavioural and imaging data from Kerns et al. (2004). Trial pairs demonstrate the ability of a single incongruent trial to reduce interference on a subsequent incongruent trial (a), which is accompanied by a dramatic reduction in ACC activation in imaging data, and the mean activation of the cognitive control node within the model (b). Finally, increased activation of the ACC is correlated with increased activation of the PFC (c) on the subsequent trial, suggesting that levels of conflict within the ACC increase enforcement of the attentional set by frontal mechanisms. In the model, the same temporal relationship is observed between activation of the cognitive-control node and the task-demand node. Data replotted from Figures 1 and 2d of Kerns et al. (2004), with permission. Error bars in (a) and (c) are too small to be clearly observed. RT = Cycles *0.75 + 485.

trial in a colour-naming simulation for trial blocks containing 70% congruent and 30% incongruent trials, as in the Kerns et al. study. Then, the average activation level of the task-demand node in each trial was placed into one of eight bins according to increasing levels of cognitive-control node output during the previous trial. The results show a similar linear trend (Figure 5c), confirming that the predicted temporal relationship between

cognitive control on one trial and task-demand activity in the following trial is obtained.

Simulating emotional effects in colour naming

We now turn to explaining emotional interference effects, which can result from the interplay between cognitive and affective aspects of adaptive attentional control.

The fast component in colour naming

Negative emotional effects during the fast component (i.e., intra trial) affect the reaction time during the same trial in which the negative word is presented. No fast effects emerge with these parameters, as the model represents a subject with no form of anxiety. However, in the prediction section, we investigate a qualitative distinction that may be drawn between state and trait anxiety through consideration of fast and slow component effects.

The slow component in colour naming

In our model, emotional interference during the slow component (i.e., inter trial) is the hallmark of the effect of emotional salience on cognitive control. Through its connection to the negative-affect component of control, a negative emotional word causes cognitive control to be temporarily suppressed. As a result, the tonic excitation of task demand is temporarily reduced; the system exhibits a temporary task withdraw response.

The reduction in task demand relaxes the bias towards colour naming in the input pathways, as well as the colour-word bias in the category layer. These effects make the system less effective at naming colours. However, because it takes time for these effects to percolate through the chain of nodes (i.e., from negative word to negative affect to cognitive control to task demand), the current response has already been generated by the time these changes occur.

Consequently, there is an increase in reaction time to name the colour of the neutral word following an emotional trial, producing interference in the slow component. Over the course of following neutral trials, cognitive control will return to baseline level, and the effect of the negative word will dissipate. The performance of the model in a simulation of McKenna and Sharma (2004), Experiment 4, is shown in Figure 6 alongside the original data. In their results, only the difference at position 2 was found significant.

It is important to note that the time constants used to fit the data of Kerns et al. (2004), as described in the previous section, are perfectly suited to explaining the data of McKenna and Sharma. Thus, there is a remarkable
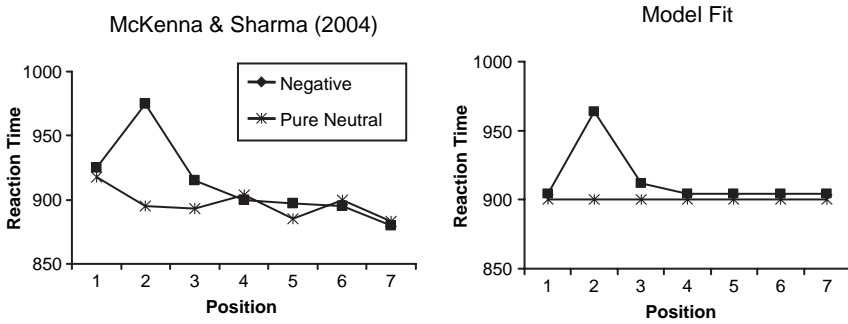
Figure 6.   Replication of the temporal profile of interference during the slow component as described in McKenna and Sharma (2004). Position 1 corresponds to the negative emotional input for negative trial sets. All other positions are neutral words. Data replotted from Figure 2 of McKenna and Sharma (2004), with permission. These simulations are deterministic; there is no variance. RT = Cycles *4 + 80.

convergence of evidence from these emotional and cognitive paradigms, suggesting that the influence of either a negative emotional word, or an incongruent trial should decay back to baseline over the course of one or two trials, rather than shifting gradually over the course of many trials.

Figure 7 depicts the trial-wise time course of changes in attentional bias in response to either an incongruent or an emotional word. In this figure, the line graph reflects the summed absolute value of task demand inputs to colour, word and category layers for a colour-naming task. Thus, larger values of task bias (Figure 7a), produced by an incongruent trial, represent a stronger excitation of the colour pathway, stronger suppression of the word pathway and increased bias excitation of colour nodes in the category layer. Reduced values of task bias (Figure 7b) represent the opposite effect.
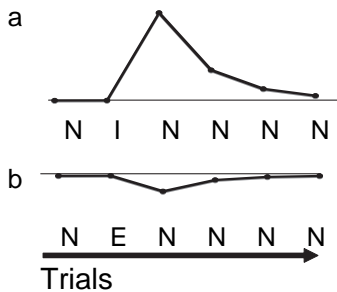
Figure 7.   The task-dependent bias of the network in the colour-processing task is increased for trials following incongruent (denoted "I") trials and suppressed for trials following negative-emotional (denoted "E") trials in a sequence of neutral trials. The upward shift indicates that bias inputs from task demand push the system to faster colour naming performance with less interference. The downward shift indicates a change towards slower colour naming.

Simulating emotional effects in word reading

The influence of negative emotional words in word reading is an important contrast to colour-naming tasks because the emotional input occurs in the same processing pathway as the stimulus feature that has to be processed (i.e., the emotional stimulus is in a task-relevant dimension).

Algom et al. (2004) have compared blocked and mixed arrangements of neutral and emotional words, and found that interference is observed during the blocked condition. This suggests that word-reading interference is produced exclusively in the slow component. Although there is evidence for facilitatory effects from other research: lower perceptual thresholds and faster lexical decisions for negative emotional words (Dijksterhuis & Aarts, 2003; Matthews & Southall, 1991; Small & Robins, 1988), this evidence has been disputed (see, e.g., Labiouse, 2004).

Our model simulates emotional interference in word reading in the same manner as colour naming. Negative emotional input suppresses input to the task-demand system, reducing the excitatory bias applied to both the word-form and category layers. This reduction in the task-based bias increases the time required for a word-reading response node to reach threshold on the trial following an emotional word. Thus, in a blocked format, every negative emotional word (except the last) can delay performance on the following emotional word. However, when the emotional and neutral trials are randomly mixed, the slow component of interference produced by a negative emotional word has an equal chance of delaying performance on emotional as neutral trials. Figure 8 illustrates our replication of the sixth experiment of Algom et al. (2004). In their results, the difference between conditions was significant for blocked, but not mixed, presentation.

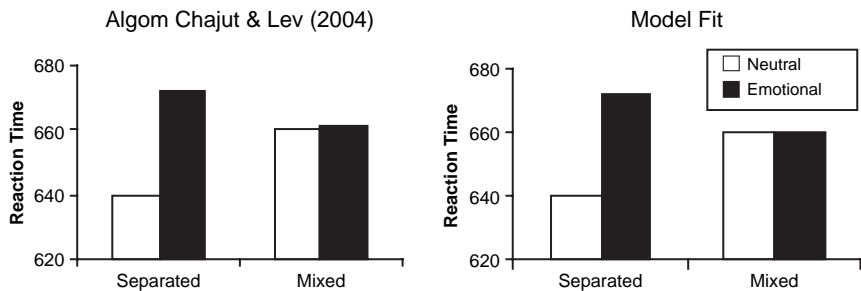**Figure 8.** The difference in the latency to read emotional words is strongly present only in blocked format. In the model, this interference is on the slow component and largely washes out when the trial blocks contain both emotional and neutral words. Data replotted from Figure 6 of Algom et al. (2004), with permission. Standard error for these simulations is less than 0.5 ms. RT = Cycles *2 + 360.

PREDICTIONS

Having developed an account of the interaction between cognitive and emotional influences on attentional control, we attempt to predict how Stroop effects and emotional interference interact in novel experiments. All of the effects modelled here use the same parameter settings as for the simulations above, apart from state and trait anxiety, which necessarily involve additional parameters.

Emotional effects increase Stroop interference

Emotional interference is hypothesised to decrease task demand, exacerbating the interference effects obtained from incongruent Stroop trials. We use a paradigm similar to the Tzelgov et al. (1992) block-composition experiment but which incorporates emotional stimuli. Subjects are asked to colour-name stimuli presented in blocks of varying composition of colour and emotional words. In blocks containing few emotional words (75% colour words), interference effects are relatively small, as task demand remains consistently elevated due to the frequent occurrence of incongruent trials. However, with increasing numbers of negative emotional stimuli inserted into the block (25% colour words), task demand is consistently suppressed by negative emotional input. Thus interference for incongruent trials should be elevated as the system spends a larger percentage of time in a task withdrawal state.

Figure 9 illustrates predicted reaction times for experimental conditions, one being the previously described simulation of the Tzelgov et al. (1992) results, and the other being a novel experimental paradigm, in which the neutral words are replaced by negative emotional stimuli.

The model predicts that replacing neutral words with emotional words will increase the amount of interference observed during incongruent trials, and
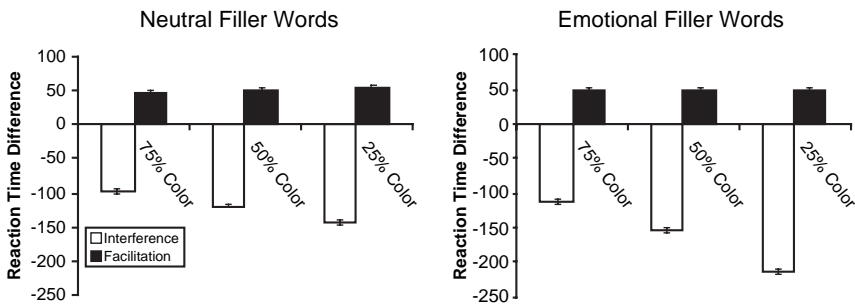

Figure 9. Simulated reaction times for the previously described experiment of Tzelgov et al. (1992) using neutral filler words, and for a version of the same experiment that uses emotional filler words. Incongruent trials during the emotional-filler condition show a larger increase in interference. Error bars are provided for the simulated data. RT = Cycles *1.0 + 500.

increasingly so as the percentage of such emotional filler words increases. Congruent trials will not be affected by the emotional manipulation.

The model also makes qualitative predictions concerning imaging results in comparing the results between the two experiments. Relative to the conventional paradigm of Tzelgov et al., the emotional variant will exhibit increased dorsal ACC activation following incongruent trials (because interference is increased), and decreased activation following emotional trials (because of suppression of cognitive control). PFC activation during the emotion filler blocks will be generally suppressed compared to the original Tzelgov condition. However, following an incongruent trial there will be greatly enhanced PFC activation, larger than what would be observed in the original design.

Stroop incongruence attenuates emotional interference

Our account also predicts that an incongruent trial reduces the impact of a following negative emotional stimulus by suppressing the emotional node of adaptive control. Thus, there is an important asymmetry when Stroop and emotional forms of interference interact; emotional interference enhances Stroop interference, but Stroop interference decreases emotional interference.

This can be illustrated with the following arrangement of stimuli. In the emotional interference paradigm described by McKenna and Sharma (2004), subjects see a sequence of neutral trials preceded by a negative word, and the emotional effect is observed in colour-naming latencies for the following words. The model predicts that an incongruent Stroop trial immediately prior to a negative emotional word will mitigate emotional interference, on following trials (Figure 10).

In fact, an incongruent trial produces a facilitatory effect on responding to neutral words, as increased task demand augments colour processing.

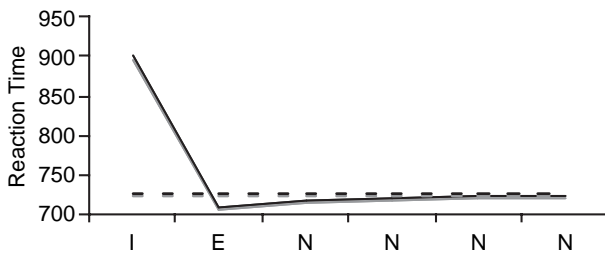**Figure 10.** Our model predicts that the inter-trial effect of an incongruent stimulus can prevent a subsequent emotional stimulus from producing interference during colour naming. The increase in task bias due to the incongruent stimulus diminishes the influence of the word processing. These simulations are deterministic; there is no variance. RT = Cycles *1.0 + 500, in accord with the simulations of Tzelgov et al. (1992).

This predicted effect is small compared to Stroop interference and would be difficult to observe in mixed-block paradigms, such as Tzelgov et al. (1992).

Here the model makes a speculative prediction concerning activation of the more rostral portions of the ACC, thought to be involved in the processing of emotional stimuli (see the discussion below). An emotional word in a stream of neutral stimuli should produce an increase in rostral ACC activation. However, if that emotional word is immediately preceded by a Stroop incongruent trial, rostral ACC activation will be attenuated.

Individual differences in state and trait anxiety

Current explanations for interference with emotional stimuli highlight the role of individual differences. In particular, it is suggested that high anxiety (both clinical and subclinical) can lead to shifts in processing priorities towards threatening stimuli (e.g., Mathews & Mackintosh, 1998; Mathews & MacLeod, 2002; Mogg & Bradley, 2005; Williams, Mathews, & MacLeod, 1996; Williams, Watts, MacLeod, & Mathews, 1988).

One of the classic distinctions in the domain of individual differences regarding anxiety is described as a difference between state anxiety, in which subjects are placed into an anxious state, and chronic conditions in which subjects are perpetually anxious about a particular concern, such as arachnophobia, or any form of negative input—trait anxiety (Williams et al., 1997). As suggested by Matthews and Harley (1996), these forms of anxiety can be implemented differently, leading to distinct predictions concerning their relative impact on fast and slow components of emotional interference. Using the simplest implementation of state and trait anxiety that can be described with this model, trait anxiety can be modelled as strengthened connections from early layers of processing to later stages. This configuration is theorised to be the result of increased practice effects from years of focusing on the topic of anxiety (Figure 11). State anxiety can be implemented as biases in early stages of processing that cause increased baseline activation of negative emotional word nodes. Thus, trait anxiety reflects a permanent structural change to the network while state anxiety is the result of transient changes in the activation state of input nodes.

There are two parameters involved; trait anxiety is determined by the degree to which relevant connection weights are enhanced and state anxiety is determined by the strength of excitatory bias (see Appendix).

The simulated results shown in Figure 12 illustrate a prediction that there is a consistent distinction between state and trait anxiety effects. Trait anxiety allows a negative word to evoke stronger activation at the category layer, competing with colour via lateral competition, and producing interference within the emotional trial (i.e., the fast component). State anxiety causes enhanced activation of the negative word node, which allows negative
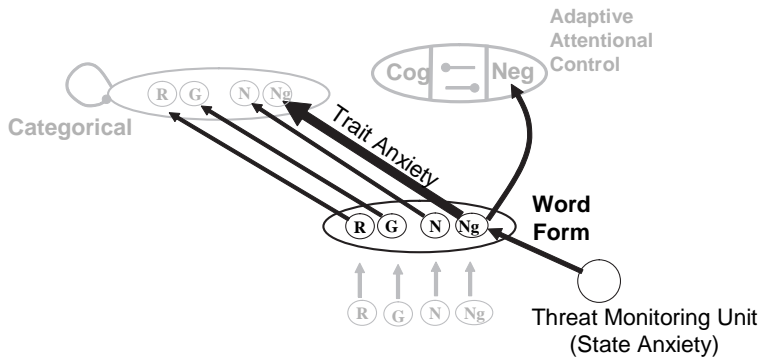
Figure 11.   Depiction of the difference between state and trait anxiety in the model. State anxiety is represented by a bias that excites the negative word-form node. Trait anxiety is represented by stronger connections to the category layer.

word input to more strongly influence adaptive attentional control in addition to an enhanced category layer activation. Accordingly, the modest fast component effect in simulated state anxiety is accompanied by an effect in the slow component.[2] Thus, the model predicts that the ratio between emotional interference observed in fast and slow components may serve as a diagnostic measure for dissociating trait and state anxiety behaviourally.

Here again the model makes useful predictions to motivate neuroimaging research. Trait anxiety should not be associated with increased rostral ACC activation, but state anxiety should. The degree of rostral ACC activation reflects emotional interference in the slow component.

DISCUSSION

This model of adaptive attentional control fits data from a broad spectrum of tasks involving both incongruent stimuli, and emotionally salient stimuli. By explaining emotional and Stroop effects within the same framework, we highlight a contrast between their functional underpinnings. Stroop interference and facilitation result from semantic similarity, as congruent or incongruent patterns of activation are evoked in the category and response layers of the network. Emotional interference, at least of the sort observed in non-anxious subjects, occurs when negative emotional stimuli produce a temporary suppression of task demand through a projection to cognitive control, resulting in a generic slow down in responding due to a task withdrawal.

---

[2] The increase in the slow component for state anxiety during word reading is reduced because the word-processing bias nears the asymptotic maximum of the logistic function in this implementation.
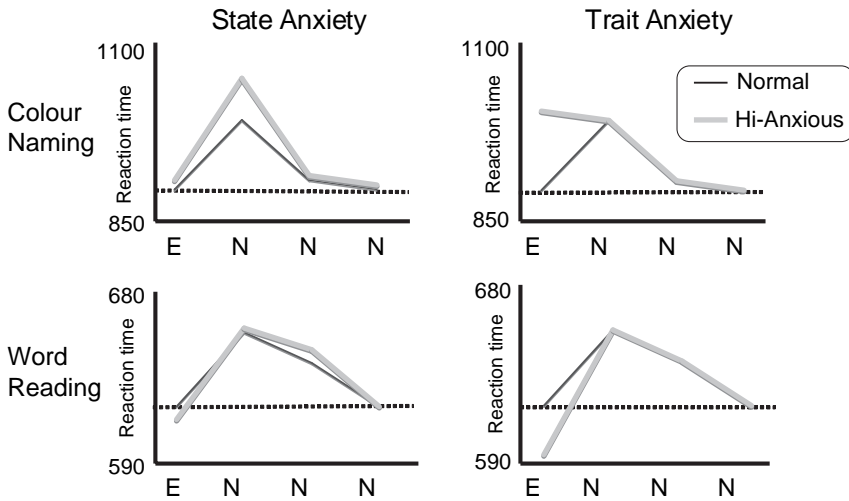
Figure 12. Simulated reaction times state and trait anxiety in colour-naming and word-reading paradigms in trial sequences. Effects that occur during the emotional trial (denoted "E") are in the fast component, while those in the following trials occur in the slow component. The dashed line indicates the baseline RT for a pure-neutral trial sequence. These simulations are deterministic, there is no variance. For colour naming: RT = Cycles *4 + 80, from the simulation of McKenna and Sharma (2004), and for word reading: RT = Cycles *2 + 360, from the simulation of Algom et al. (2004).

By describing this theory in the form of a model, it becomes easier to understand how the interaction between cognitive and emotional aspects of attentional control produce emotional interference when the emotional input is either task relevant or task irrelevant. For example, Algom et al. (2004) described emotional interference as a generic slowing of performance. This description was interpreted by Dalgleish (2005) as being difficult to distinguish from the conventional notion of the "emotional Stroop" effect. By providing a computational account of emotionally induced task withdrawal (i.e., akin to Algom et al.'s notion of a generic slow down), we isolate distinct mechanisms that cause the Stroop effect (competition between semantically incongruent representations), and emotional interference (a momentary withdrawal of attentional resources from task demand). This specificity makes it easier to contrast competing theories.

Relation to theories of attention and emotion

Our model is closely related to previous theoretical work that describes the influence of emotion on attention. For example, Williams et al. (1988) proposed a two-stage pre-attentive mechanism to account for the attentional bias from threat-related stimuli. Stimulus input is evaluated for threat by an

affective decision mechanism that is itself modulated by state anxiety. Its output is then used in a second stage to allocate attentional resources. The resource-allocation mechanism is thought to be modulated by trait anxiety, in which high trait anxious subjects allocate resources towards the source of threat whereas low trait anxious subjects allocate resources away from the location of threat. In their cognitive–motivational model Mogg and Bradley (1998) also described a two-stage affective mechanism involving an initial evaluation of valence (that is modulated by trait anxiety) followed by a mechanism that interrupts goals (to orient towards threat) or pursues goals when threat is low.

Although these models make some specific predictions about the processing of threatening stimuli, they nevertheless fail to capture the interaction between bottom-up and top-down processing. This interaction is important because it has implications for the way in which negative emotional stimuli are thought to be processed when task relevant or task irrelevant. For example, irrelevant threatening stimuli are thought to interfere with task processing because attentional resources are allocated to the threatening stimulus. This notion is similar to the within-trial effects that we describe for state and trait anxiety. It is also compatible with the notion of a hypervigilant mode of processing with threatening stimuli (e.g., Eysenck, 1992; Mathews, 1993). Such modes of processing would predict interference when threat is irrelevant to task processing but facilitation when it is task relevant. However, the idea of preferential processing of emotional stimuli cannot explain emotional interference in word reading, nor is the idea consistent with the fact that colour-naming interference is primarily restricted to the slow (i.e., inter-trial) component.

Eysenck, Derakshan, Santos, and Calvo (2007) have recently proposed an attentional control theory to explain how anxiety affects cognitive performance. Central to their model is the interaction between stimulus-driven and goal-directed attentional systems. In particular they propose that anxiety increases the processing influence of the stimulus-driven attentional system, and impairs processing efficiency of the goal-directed attentional system. These ideas are broadly similar to mechanisms in our model; state anxiety increases the ability of an emotional stimulus to influence top-down control. Accordingly, their account of attentional control may predict the same influence of state anxiety on the slow component of emotional interference.

Another theory is suggested by the data of Siegle, Steinhauer, and Thase (2004), which finds a carryover effect of pupil dilation from one trial to dilation on the next in a Stroop task. This suggests another potential factor underlying inter-trial effects; the accumulation of cognitive load, or arousal over the course of difficult or emotional trials respectively. Our own account uses a very different form of inter-trial carryover, namely fluctuations in attentional set.

Data linking pupil diameter on one trial to behaviour on the next would be ideal for discriminating between these competing ideas.

Attentional bias to negative stimuli

There is evidence to suggest a similar interaction between attention and emotion in spatial attention bias experiments. In such paradigms, emotional words or faces are flashed on the screen followed by dot probes at the same (i.e., valid) or different (i.e., invalid) positions (e.g., Fox, Russo, Bowles, & Dutton, 2001; MacLeod, Mathews, & Tata, 1986).

In particular, a number of studies have pointed out that the search for attentional bias towards emotional stimuli is complicated by increases in baseline reaction times for trials containing negative stimuli (Koster, Crombez, Verschuere, & De Houwer, 2004; Yiend & Mathews, 2001). When dot-probe reaction times are compared between trials containing a negative stimulus (e.g., a threatening picture) and trials containing only neutral stimuli, it is found that responses are generally slower in the threat-containing trial, even for responding to the location of the threatening stimulus. A finding of generally slower reaction times following presentation of a threatening stimulus is consistent with our hypothesised task withdrawal in response to negative stimuli. It is notable that, unlike in colour naming, this slowdown is observed within the same trial as the emotional stimulus. This suggests that changes in task demand occur more rapidly in such dot-probe experiments.

Neurophysiological correlates of adaptive attentional control

Recent imaging studies have made important strides in relating theories of cognitive and emotional processing to activity in specific brain regions. Our model, being a neural network, is well poised as an interface between traditional notions of emotional interference and recent neuroimaging results.

This body of work suggests that the dorsal ACC (dACC) is involved in the mediation of cognitively challenging tasks such as Stroop colour naming. Of particular note are experiments that study both reaction time and dACC activation to incongruent activity in a trial-by-trial fashion (Carter et al., 2000; Kerns et al., 2004; Mohanty et al., 2007).

There are several theoretical claims involving the specific manner in which the dACC may be involved; monitoring conflict (Botvinick et al., 2001), responding to errors (Carter et al., 1998), or providing a broader range of functionality (Milham & Banich, 2005), including both pre-onset configuration as well as dynamic adaptation to task difficulty (Weissman et al., 2004). Of these theories, our adaptive attentional control mechanism most closely corresponds to the latter. The cognitive-control node has a tonic

activation state, in a similar manner as sustained dACC activation can be observed in response to task instructions (Dosenbach et al., 2006). This cognitive control output can be temporarily increased by difficulty, such as an incongruent trial, or decreased by task withdrawal in response to emotionally salient stimuli.

The idea that attentional control is mediated by the ACC has an emotional parallel as well. There is evidence for functional heterogeneity between dorsal and rostral portions of the ACC (rACC; Vogt, Finch, & Olson, 1992). The rostral portion of this area responds primarily to emotional input in a variety of contexts (cf. Bush, Lu, & Posner, 2000; Drevets & Raichle, 1998). More direct evidence for selective activation of the dorsal and rostral ACC by cognitive and emotional versions of counting tasks has been found (Bush et al., 1998; Whalen et al., 1998).

We suggest that the ACC may represent the cortical counterpart of the amygdalar and thalamic fear circuitry (LeDoux, 2003). The putative subcortical, pre-attentive, automatic vigilance towards threatening stimuli and the generalised arousal reaction following such stimuli, may be complemented by a cortical mechanism, which acts to temporarily suppress ongoing cognitive activity to augment the ability to detect and process threatening stimuli. Imaging research confirms that there are cortical counterparts to fear responses that operate in conjunction with amygdalar mechanisms. PET imaging data from Pissiota et al. (2003) demonstrated that blood flow in the affective division of the ACC increases concurrently with amygdalar activity during periods of potentiated startle reflexes evoked by concern-related pictures presented to spider-phobic and snake-phobic subjects. More specifically, in the context of experiments such as colour naming, reading or counting, we suggest that the rACC plays a role in monitoring and responding to emotionally salient information. Furthermore, its involvement, in response to negative stimuli, may result in suppression of the ongoing task set. Such suppression could ultimately involve deactivation of the dorsal ACC, and, indirectly, the PFC.

However, this account remains controversial. Mohanty et al. (2007) examined activity in these two regions during blocked performance of negative and Stroop incongruent colour-naming trials within subjects. They did find that dACC activation accounted for variance in the PFC during the negative emotional trials, which supports our account. However, they also found no deactivation of the dACC, suggesting that the interaction of rostral and dorsal areas of this structure may be more subtle than changes in average activation. It should be stressed, however, that there was no

behavioural effect of the negative stimuli either (i.e., negative trials were not slower than neutral trials).

Inter-trial intervals and inter-trial effects

One methodological factor that has not yet been discussed is the influence of the amount of time between trials, which could modulate inter-trial effects. Tzelgov et al. (1992) used a brief response stimulus interval (RSI) that was controlled by the experimenter but was 800 ms at minimum. Imaging studies pace experiments with a considerable inter-stimulus interval (ISI) for the purpose of isolating trial-based effects in the hemodynamic response (Carter et al., 2000; Kerns et al., 2004). Both approaches find strong block composition or inter-trial effects. Thus it seems that inter-trial effects produced by an incongruent Stroop trial can carry over considerable delays. On the other hand, Sharma and McKenna (2001) found emotional interference to be sensitive to RSI.

Our model avoids the complications inherent in this issue by taking a parsimonious approach in which each trial takes a set amount of time and has a brief ISI (see Appendix), used to let network activations settle at the start of each trial. Further investigation, using parametric variations in trial timing is necessary to determine more precisely how the inter-trial effects produced by Stroop and emotional interference are affected by inter-trial intervals.

CONCLUSIONS

Our model is a quantitative and qualitative description of processes that are hypothesised to occur following perception of a potentially threatening stimulus. We explain the observed delayed interference effect as the result of an adaptive attentional control circuit that is regulated by indicators of both cognitive and affective demand. This explanation adds a temporal dimension to existing theories of emotional interference, which is critical for explaining differences observed between mixed and blocked experiment designs.

We frame the idea of emotional interference as a crucial tradeoff in our cognitive architecture. We suggest that this simulated task withdrawal response may serve as a defensive mechanism in more natural contexts. It is an automatic response to a potential stimulus, which would allow an organism to rapidly locate the source of a threat using all available sensory information by relaxing task-specific suppression of particular channels (e.g., word processing in the case of colour-naming Stroop trials).

REFERENCES

Albery, I. P., Sharma, D., Niyazi, A., & Moss, A. C. (2006). Theoretical perspectives and approaches. In M. Mafano & I. P. Albery (Eds.), *Cognition and addiction* (pp. 1–19). Oxford, UK: Oxford University Press.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, *133*(3), 323–338.

Battye, G. (2004). *Connectionist modelling of attention and anxiety.* Unpublished PhD Thesis, Cambridge University: UK.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Bush, G., Lu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Science*, *4*(6), 215–222.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging— Validation study with functional MRI. *Human Brain Mapping*, *6*, 270–282.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection and the on-line monitoring of performance. *Science*, *280*, 747–749.

Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., & Noll, D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences, USA*, *97*(4), 1944–1948.

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*(3), 332–361.

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

Cohen, J. D., & Servan-Schreiber, D. (1992). A neural network model of disturbances in the processing of context in schizophrenia. *Psychiatry Annual*, *22*, 131–136.

Dalgleish, T. (2005). Putting some feeling into it—The conceptual and empirical relationships between the classic and emotional Stroop tasks: Comment on Algom, Chajut, & Lev (2004). *Journal of Experimental Psychology: General*, *134*(4), 585–591.

Dijksterhuis, A., & Aarts, H. (2003). On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science*, *14*(1), 14–18.

Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory and Cognition*, *28*(8), 1437–1449.

Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., & Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron*, *50*(5), 799–812.

Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. *Cognition and Emotion*, *12*(3), 353–385.

Dunbar, K., & MacLeod, C. M. (1984). A horse race of a different colour: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 622–639.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Hove, UK: Lawrence Erlbaum Associates, Ltd.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. (2007). Anxiety and performance: Attentional control theory. *Emotion, 7*, 336–353.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, *130*(4), 681–700.

Holle, C., Neely, J. H., & Heimberg, H. R. (1997). The effects of blocked versus random presentation and semantic relatedness of stimulus words on response to a modified Stroop task among social phobics. *Cognitive Therapy and Research*, *21*, 681–697.

Kerns, J. G., Cohen, J. D., MacDonald, A. W. 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023–1026.

Klein, G. S. (1964). Semantic power measured through the interference of words with colour-naming. *American Journal of Psychiatry*, *77*, 576–588.

Koster, E. H., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*(10), 1183–1192.

Labiouse, C. L. (2004). Is there a real preferential detection of negative stimuli? A comment on Dijksterhuis and Aarts. *Psychological Science*, *15*(5), 364–365.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular & Molecular Neurobiology*, *23*(4–5), 727–738.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*(5472), 1835–1838.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20.

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*(10), 383–391.

Mathews, A. (1993). Biases in processing emotional information. *Psychologist*, *6*, 493–499.

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research*, *22*, 539–560.

Mathews, A. M., & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition and Emotion*, *16*, 331–354.

Matthews, G., & Southall, A. (1991). Depression and the processing of emotional stimuli—a study of semantic priming. *Cognitive Therapy and Research*, *15*(4), 283–302.

Matthews, G., & Harley, T. A. (1996). Connectionist models of emotional distress and attentional bias. *Cognition and Emotion*, *10*(6), 561–600.

McKenna, F. P., & Sharma, D. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: The role of fast and slow components. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(2), 382–392.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*(3), 422–471.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions—Effects of base-line discriminability, response-mode, and practice. *Memory and Cognition*, *21*(5), 627–645.

Milham, M. P., & Banich, M. T. (2005). Anterior cingulate cortex: An fMRI analysis of conflict specificity and functional *Human Brain Mapping*, *25*, 328–335.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy*, *36*, 809–848.

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research*, *29*, 29–45.

Mogg, K., McNamara, J., Powys, M., Rawlinson, H., Seiffer, A., & Bradley, B. P. (2000). Selective attention to threat: A test of cognitive models of anxiety. *Cognition and Emotion*, *14*(3), 375–399.

Mohanty, A., Engels, A. S., Herrington, J. D., Heller, W., Ringo Ho, M. H., Banich, M. T., et al. (2007). Differential engagement of anterior cingulated subdivisions for cognitive and emotional function. *Psychophysiology*, *44*(3), 343–351.

O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. *Neural Computation*, *13*(6), 1199–1241.

Öhman, A., Flykt, A., & Esteves, F. (2001a). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478.

Öhman, A., Lundqvist, D., & Esteves, F. (2001b). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences*, *87*, 256–259.

Phaf, R. H., & Kan, K. (2007). The automaticity of emotional Stroop: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 184–199.

Pissiota, A., Frans, O., Michelgard, A., Appel, L., Langstrom, B., Flaten, M. A., et al. (2003). Amygdala and anterior cingulate cortex activation during affective startle modulation: A PET study of fear. *European Journal of Neuroscience*, *18*(5), 1325–1331.

Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology*, *83*(4), 479–491.

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 515–537.

Sharma, D., & McKenna, F. P. (2001). The role of time pressure on the emotional Stroop task. *British Journal of Psychology*, *92*(3), 471–481.

Siegle, G. J., Steinhauer, S. R., & Thase, M. E. (2004). Pupillary assessment and computational modeling of the Stroop task in depression. *International Journal of Psychophyics*, *52*, 63–76.

Small, S. A., & Robins, C. J. (1988). The influence of induced depressed mood on visual recognition thresholds—Predictive ambiguity of associative network models of mood and cognition. *Cognitive Therapy and Research*, *12*(3), 295–304.

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

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

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for colour words. *Memory and Cognition*, *20*(6), 727–735.

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. *Cerebral Cortex*, *2*(6), 435–443.

Waters, A. J., Sayette, M. A., Franken, I. H. A., & Schwartz, J. E. (2005). Generalizability of carry-over effects in the emotional Stroop task. *Behaviour Research and Therapy*, *43*, 715–732.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology, 77*(Pt. 1), 97–108.

Weissman, D. H., Gopalakrishnan, A., Hazlett, C. J., & Woldorff, M. G. (2004). Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cerebral Cortex*, *15*, 229–237.

Whalen, P. J., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., et al. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry*, *44*, 1219–1228.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. M. (1988). *Cognitive psychology and emotional disorders.* Chichester, UK: Wiley.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. M. (1997). *Cognitive psychology and emotional disorders* (2nd ed.). Chichester, UK: Wiley.

Yiend, J., & Mathews, A. (2001). Anxiety and attention to threatening pictures. *Quarterly Journal of Experimental Psychology*, *54A*, 665–681.

APPENDIX

Equations and simulations

Nodes in our model are designed to capture the excitatory and inhibitory dynamics of neural elements. Activation functions are a combination of bias, excitation, inhibition and time averaging. Connections between elements are excitatory or inhibitory and are not modifiable.

The activation dynamics of nodes in the model are governed by the same equations used in the model of Cohen et al. (1990), as shown in Equation 2.[3]

$$a_{j,t} = a_{j,t-1}(1 - \tau) + input_{j,t-1}\tau \tag{2}$$

The term $a_{j,t}$ represents the activation of node j at time t which approaches the $input_{j,t}$ (Equation 3) asymptotically at a rate governed by time constant τ.

$$input_{j,t} = \sum_i out_{i,t}w_{i,j} + bias_j \tag{3}$$

In this equation, $out_{i,t}$ is the output of a node. The weight between node i and j is $w_{i,j}$ and bias is a constant that determines the resting activation level of node j.

The output of a node $out_{i,t}$ is computed with a sigmoid function (Equation 4) restricting it to the range of values (0,1).

$$out_{i,t} = \frac{1}{1 + e^{-\gamma_i(a_{i,t}-\theta_i)}} \tag{4}$$

This equation scales all possible activation values to the range 0 to 1. The slope of the curve is controlled by the scaling parameter: γ and offset θ. All nodes are given a constant bias of -1.0 to make them inactive (i.e., an output of near 0) by default, with the exception of the cognitive control node.

Each trial consisted of 600 simulated time steps. Several nodes maintained activation from one trial to the next: cognitive and negative nodes of

---

[3] This same model has also been implemented with the more biologically plausible equations described by O'Reilly (2001). Both versions of the model are available in Matlab code by e-mail request from the authors or online at: http://www.bradwyble.com/research/models/cogcontrol/.

adaptive attentional control as well as task-demand nodes. These inter-trial nodes began each experimental block at the asymptotic level of activation that would be reached when the network was permitted to settle in the absence of incongruent or emotional input (i.e., neutral or congruent trials), after which their activation was determined by stimulus input. All other nodes were reset to a starting value of -1.0 on each trial. Stimulus onset occurred at 50 cycles and was sustained until the end of a simulated trial.

Parameters

The table shows the parameters used to simulate the data and predictions described in the body, with the exception of the slope/intercept values that were adjusted to convert cycles in milliseconds. Fitting involved setting all of the parameters to produce a general fit of the results. Then, several key parameters (marked by asterisks) were varied to produce the best possible fit of all simulated results simultaneously. The same parameter settings, except slope and intercept, were used for all of the simulations described in the text.

Identical across Layers, except where noted below:

Logistic $\gamma = 3$; Logistic $\theta = 0$; BIAS $= -1$; $\tau = .15$.[4]

	$w_{i,j}$	τ	BIAS
Time constants and biases			
Colour task demand		0.002*	
Word task demand		0.002*	
Response nodes		0.003	-2
Cognitive		0.02*	0
Negative		0.02*	
Excitatory connection strength			
Input to colour/word form	1.0		
Colour to category	1.0		
Word to category	3.5		
Category to response	6.0		
Task demand excites colour/word	0.5*		
Task demand excites category	0.63*		
Cognitive attn ctrl to task demand	2.0		
Negative word to negative attn ctrl	1.7*		
Inhibitory connection strength			
Task demand (lateral inhibition)	-2.0		
Category (lateral inhibition)	-3.0		
Response (lateral inhibition)	-5.0		
Attn ctrl (lateral inhibition)	-1.0		

---

[4] These are given as default values for those layers not mentioned below.

	$w_{i,j}$	τ	*BIAS*
Task Demand suppression of colour and word pathways	−1.0*		
Miscellaneous parameters			
Response threshold	0.0[5]		
Task demand excites category γ	100.0		
Colour task demand to category θ	−0.065*		
Word task demand to category θ	−0.12*		
Category layer (lateral inhibition) θ[6]	1		

Conflict

Conflict was computed at each time step from the current activation of the response nodes. Their activation values were made non-negative by adding an offset and then flooring the result at zero. The product of the resultant values was computed and then transformed by the logistic equation. Finally, 0.5 was subtracted and the result floored at zero so that resting levels of activation produced no conflict. This total was scaled by conflict strength and added to the cognitive control node.

Conflict parameters	
Conflict activation offset	0.95
Conflict γ	1.0
Conflict strength	40*

Principles of parameter fitting

The key parameters (marked by asterisks above) that were adjusted to fit the data were as follows:

Strength of task demand to word/colour and category layers was adjusted so that incongruent trials produced similar activation for the two competing responses. Time constants of attentional control and task demand were adjusted so that fluctuations dissipate over the course of two neutral trials as suggested by the data of Kerns et al. (2004) and McKenna and Sharma

---

[5] Negative bias pushes nodes to negative values by default.

[6] This elevated threshold value for inhibition within the category layer ensured that there was minimal inhibitory competition even when category nodes were excited by bias input.

(2004). The influence of conflict on attentional control was adjusted so that a single incongruent trial excites cognitive control sufficiently to eliminate interference for an immediately following incongruent trial. The settings of these parameters were not critical for producing the qualitative pattern of results, but did allow us to most closely match the specific time course of imaging and behavioural data from Kerns et al. (2004) and McKenna and Sharma (2004). The strength of negative word influence on attentional control was adjusted so that an emotional word attenuated cognitive control sufficiently to increase interference but not sufficiently to cause errors during colour naming, to reflect the fact that these tasks focus on reaction time, rather than accuracy.

Theta parameters of the logistic function from task demand to category layer were adjusted so that the bias to the category layer from task demand was saturated at the top of the sigmoid curve to prevent an increase in task demand from pushing response nodes to threshold in the absence of input.

State anxiety was modelled by providing a constant input of .05 to the negative node in the word input layer. Trait anxiety involved multiplying the strength of the connection from the negative node in the word input layer to the corresponding node in the category layer by a factor of 1.6. These values are not critical; the effects of anxiety vary in a continuous fashion as these parameters are adjusted.