Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Neurocomputing 69 (2006) 1322–1326

NEUROCOMPUTING

www.elsevier.com/locate/neucom

A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions

Nicola De Pisapia*, Todd S. Braver

Cognitive Control and Psychopathology Laboratory, Department of Psychology, Washington University, Saint Louis, Missouri (MO) 63130-4899, USA

Available online 3 February 2006

Abstract

A computational model is presented that describes dual mechanisms of cognitive control through interactions between the prefrontal cortex (PFC) and anterior cingulate cortex (ACC). One mechanism, reactive control, consists in the transient activation of PFC, based on conflict detected in ACC over a short time-scale. The second mechanism, proactive control, consists in the sustained active maintenance of task-set information in a separate PFC module, driven by long time-scale conflict detected in a separate ACC unit. The computational function of the first mechanism is to suppress the activation of task-irrelevant information just prior to when it could interfere with responding. The role of the second mechanism is to prime task-relevant processing pathways prior to stimulus-onset, in a preparatory fashion. The model provided an excellent fit to both the behavioral and brain imaging data from a previous detailed empirical study on humans performing the color-word version of the Stroop task. The model captured changes in reaction times across conditions, accuracy, and transient and sustained activity dynamics within lateral PFC and ACC.

Keywords: Prefrontal cortex; Anterior cingulate cortex; Proactive and reactive cognitive control; Conflict

1. Introduction

A great deal of convergent research has suggested that the lateral prefrontal cortex (PFC) and anterior cingulate cortex (ACC) play a critical role in human cognitive control. The relationship between these two brain areas has been studied extensively, both in neuroimaging (for example in [3,5]) and in neural-network models [1]. In particular, one focus has been on the role of ACC in detecting response conflict during the execution of a cognitive task, and the subsequent translation of this conflict into cognitive control. The basic hypothesis is that when high conflict occurs between different motor or behavioral responses, cognitive control mechanisms intervene to bias one response versus the others depending on the task requirement, thus overcoming the conflict.

In these previous studies PFC and ACC interactions have been characterized in terms of a single conflict–control loop mechanism: performance of certain task conditions leads to detection of response conflict, which in turn leads to the engagement or increase of cognitive control, that results in improved conflict resolution in subsequent performance. Here, motivated by previous experimental findings in humans, we develop a new neural-network model in which ACC-PFC interactions are described by two, rather than one, distinct conflict–control loops. The first mechanism, reactive control, is characterized by the transient activation of PFC based on conflict detected in ACC over a short time-scale (on the order of milliseconds). The second mechanism, proactive control, is characterized by the sustained active maintenance of task-set information in a separate PFC module, which is driven by long time-scale conflict detected in a separate ACC unit (on the order of several seconds or minutes). The computational function of the first mechanism is to suppress the activation of task-irrelevant information just prior to when it could interfere with responding. The role of the second mechanism is to prime task-relevant processing pathways prior to stimulus-onset, in a preparatory fashion.

We conducted several computational simulations with this model to examine how well it could account for

*Corresponding author.
E-mail address: ndepisap@wustl.edu (N. De Pisapia).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.12.100

detailed empirical data regarding human behavioral performance and brain activation. The task studied was the color-word version of the Stroop task, a benchmark experimental preparation for examining response conflict and cognitive control.

2. Methods

2.1. Simulation method

The model is a large-scale connectionist network [7], simulating rate code spiking activity of brain regions involved in the execution of the color-naming Stroop test. This task requires verbally responding with the name of the font-color in which a visually presented English word appears. For example, the words DOG (in green color), RED (in red color), GREEN (in blue color) would require as correct verbal responses, respectively, “green”, “red”, and “blue”. Trials can be of several different types: neutral, congruent, and incongruent. Neutral trials are like the first example, in which the word name does not refer to a color. Congruent trials are like the second example, in which the word name and font-color coincide. Incongruent trials are like the third example, in which the word name refers to a different color than the font-color used for it.

The neural-network model (see Fig. 1 and Table 1) developed to simulate this task—built upon an earlier model described in [1]—consisted of two input layers, one representing stimulus features of the color dimension (3 units, one for red, one for green, one for neutral), and the other representing lexical features of the word dimension (3 units, as for the color dimension). A response layer (2 units, one for red response, one for green response) coded the output of the network. The activation function of each unit simulates rate code spiking activity of large brain regions [7]. As a first qualitative approximation, we assume a monotonic relationship between these activations and percentage changes in blood oxygenation level dependent (BOLD) signals, as measured in fMRI experiments. One-to-one weights coming from the word red and green units onto the response red and green units were greater in strength than the corresponding one-to-one weights coming from the color layer onto the same response layer. This weight strength asymmetry captures

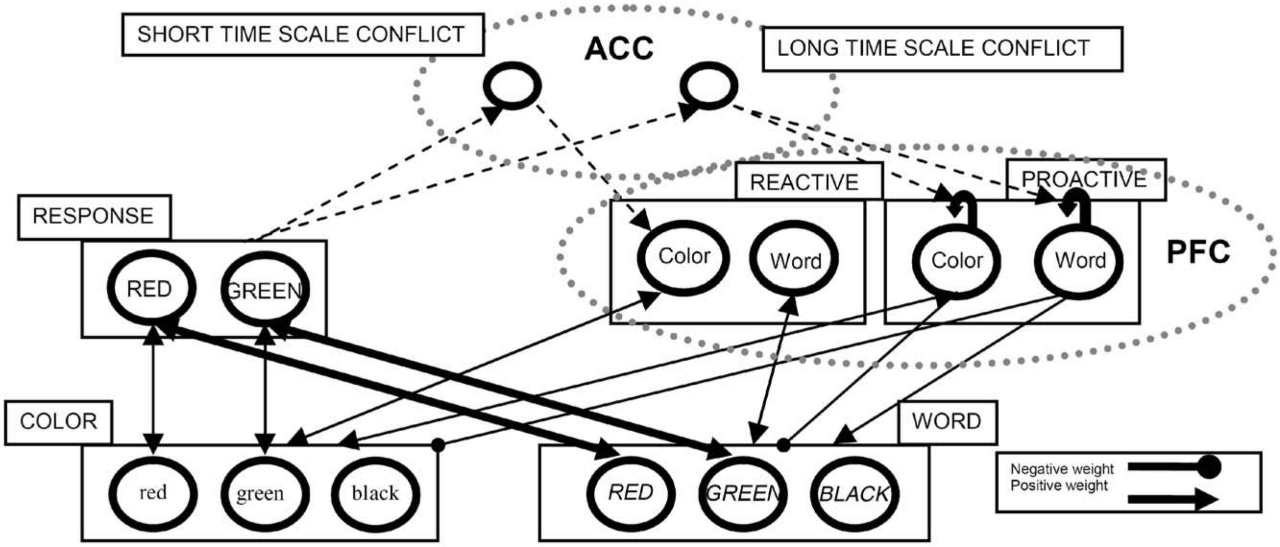

Fig. 1. A model of dual mechanisms in cognitive control in the color-naming Stroop test. Excitatory connections (arrowheads) impinging on units represent unit-specific inputs; connections impinging on network layers (represented by rectangles) represent inputs to the entire layer. Inhibition (circle-heads) is also present within each layer. For detailed description and equations see main text and Table 1.

Table 1
Key equations used in the model

Activation of a unit j is a logistic function of the net-input net_j (see below) into unit j at time t.	$a_j = \frac{1}{1+\mathrm{e}^{-\mathrm{net}_{j(t)}}}$
Raw net input, *input* is binary, and *noise* are normally distributed values used to induce variability (mean = 0, standard deviation = 1).	$rawnet_j(t) = \sum_i a_i(t)w_{ij} + input + bias + noise$
Cascade net input, used to simulate continuous time dynamics as $\tau = 0.0115$ (see Ref. [6]).	$net_j(t) = ((1-\tau)\,net(t-1)) + (\tau\,rawnet_j(t))$
Short term scale conflict of response units i and j, with weight w_{ij}, measured at time t. It is an average of Hopfield energy (see Ref. [4]) in the last 200 simulation time steps.	$shortconf(t) = \frac{1}{200}\sum_{t-200}^{t}\sum_i\sum_j a_i(t)a_j(t)w_{ij}$
Long time-scale conflict, measured and changed only after response. $\alpha = 0.95$	$longconf(t_{\mathrm{CurrTrial}}) = \alpha\, longconf(t_{\mathrm{LastTrial}}) + ((1-\alpha)\, shortconf(t_{\mathrm{response}}))$

Units are indexed as i or j, while t is time, w_{ij} is the weight between unit i and j.

the automatic tendency to read the written words, rather than to name the color in which they are written. A negative bias towards the input units ensured that their activity was above baseline only during simulated stimulus presentation. Lateral inhibition within each layer ensured competitive activation dynamics. No learning took place during simulations, because in this model we did not focus on associative processes (which may also be an important component of performance). Noise was added at the net-input to induce variability.

The main modifications made to the original Botvinick et al. [1] model were on the conflict detection units (which simulate ACC function), and on the task units layers (which simulate lateral PFC function in different subregions), that exert cognitive control onto the input layers. In the original model there was one conflict unit, calculating the Hopfield energy in the response layer [4], and only at response time. In our modified version, there are two conflict units, one calculating conflict—still as Hopfield energy, but across a short time-scale, i.e., a moving window of 200 time steps (for equation, see Table 1, fourth row). The second conflict unit calculated long time-scale conflict, which remained constant for each trial, and it was computed as an average of previous short time-scale conflicts at the time of response output (Table 1, fifth row).

Furthermore, in the original model there was one task layer, with one unit for the color naming task and one unit for the word reading task (to capture the ability of participants to rapidly switch between the two tasks, as required in some versions of the Stroop). In our modified version, there are instead two task layers, the reactive one, which is modulated by short time-scale conflict input (2 units, color naming and word reading), and the proactive one, which is modulated by long time-scale conflict input (again two units). Furthermore, the proactive units have self-recurrent weights, whose values passively decay with time, but are also selectively increased or decreased following each trial, based on long time-scale conflict input. This reflects a tendency to exert more control (via active maintenance of task information) following a high level of experienced response conflict.

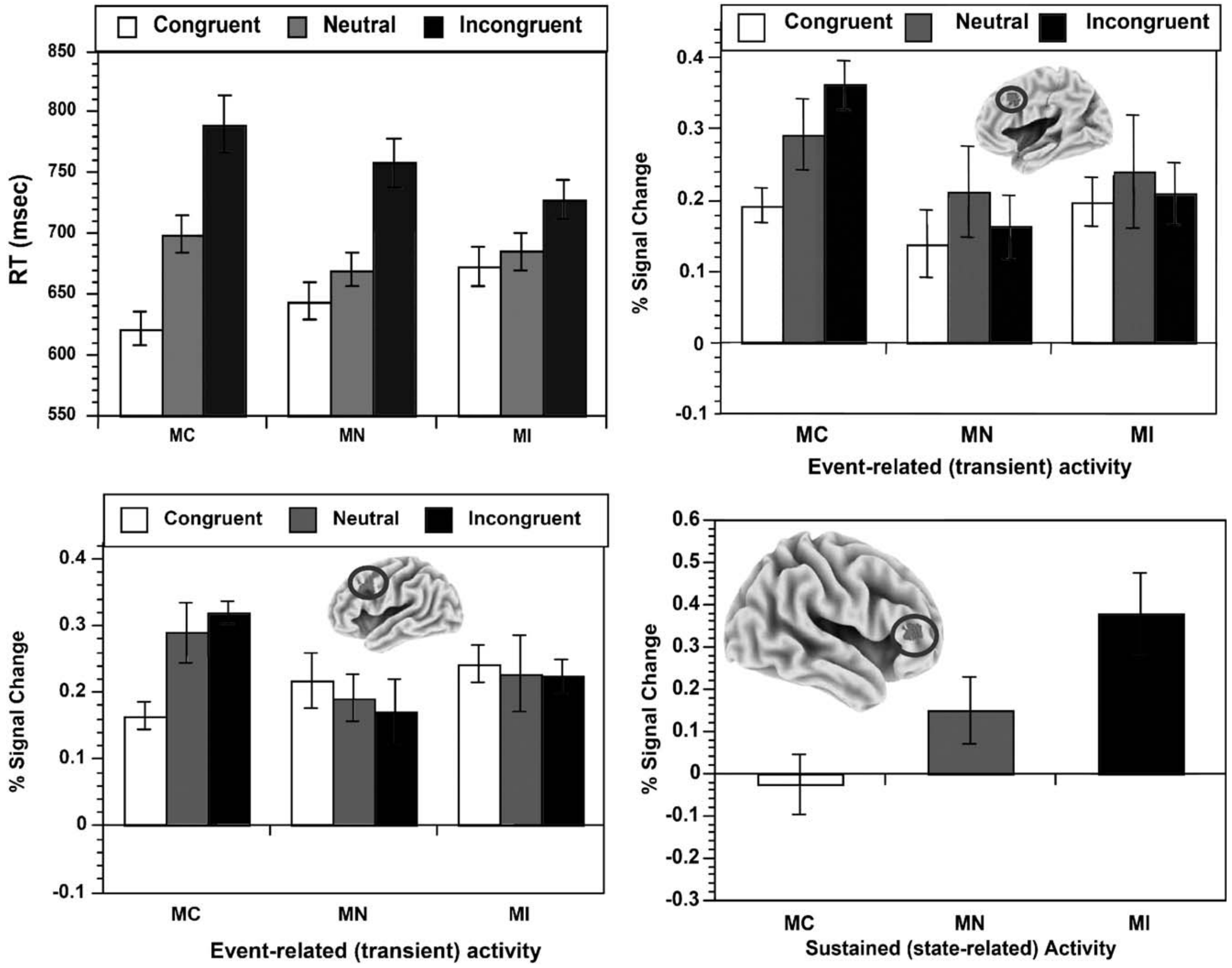

Fig. 2. Human data. Reaction times (upper left), transient percentage change in BOLD signals in ACC (upper right) and in lateral PFC (lower left), and overall sustained percentage signal change in more anterior lateral PFC (lower right) in the different conditions.

For both task layers, color task units have positive connections towards the color feature input units, whereas word task units have positive connections towards the word feature input units. Finally, when an error occurs in the response, the self-recurrent weights are reset to zero, as at the beginning of each block.

2.2. Experimental method

In a previous study [2], we examined behavioral performance from 32 participants, and fMRI BOLD data from an additional 11 participants, that performed the color-naming Stroop in three different conditions. Conflict associated with processing each trial can vary according to the information that is automatically activated by the word-name feature. Thus, response conflict will be the lowest for congruent trials (since the word name will activate the same response as the font-color), and the highest for incongruent trials (since the word name will activate a different response than the font-color). Three task conditions were presented in separate 80 trial blocks (2 each per condition): mostly congruent (MC: 70% congruents, 15% neutrals, and 15% incongruent), mostly incongruent (MI: 70% incongruents, 15% neutrals, and 15% congruents) and mostly neutral (MN: 70% neutrals, 15% incongruents, and 15% congruents).

The brain imaging data were analyzed to extract transient brain activation evoked from the performance of each task trial, as well as any residual sustained activation maintained across task blocks. Importantly, in the empirical studies, participants were not given explicit information regarding the trial frequency manipulations present across the task conditions, nor were they given any strategy instructions regarding how to optimally engage cognitive control in each condition. Likewise, they were for the most part unable to report on these manipulations at the end of the experiment. Thus, any observed changes were likely to be implicit, and in direct response to ongoing experience of task performance and conflict.

3. Results

In the empirical data (Fig. 2), behavioral performance patterns (upper left panel) indicated that interference (incongruent minus neutral reaction times) and facilitation (neutral minus congruent) effects in response times were reduced in the MI condition compared to MC, with an intermediate effect in the MN condition. This is consistent with a shift from reactive to proactive control when comparing MC versus MI conditions, since proactive control should result in a tonically reduced influence of the irrelevant word name information. The brain imaging

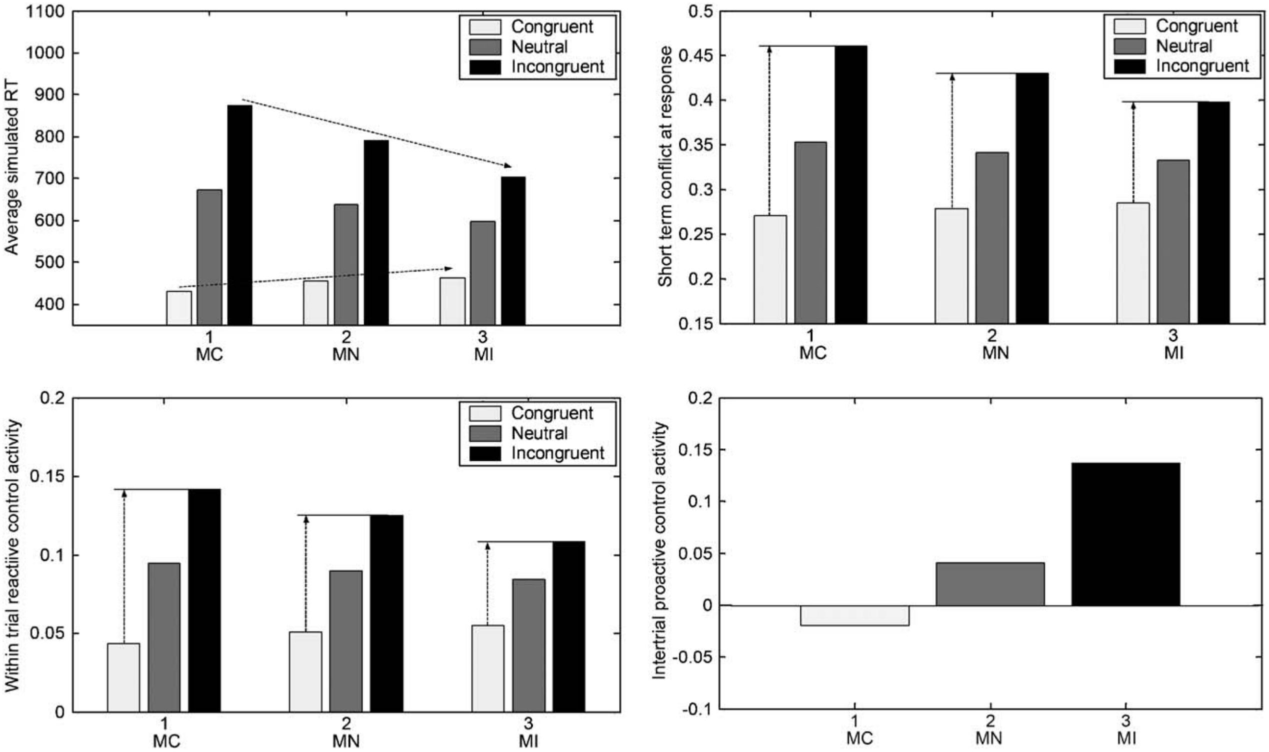

Fig. 3. Model data. Simulated reaction times (upper left), transient activity in short-time scale conflict unit (simulating ACC; upper right) and in reactive task-set units (simulating lateral PFC; lower left), and sustained overall activity in proactive task-set units (simulating anterior lateral PFC; lower right) in the different conditions. Arrows in the upper left panel indicate the increase in average reaction time for congruent trials in MI versus MN and MC, and vice versa a decrease for incongruent trials. As a consequence, interference and facilitation are reduced in MI compared to MC, as observed in human data. Arrows in the upper right and lower left panels indicate a decrease in the difference of incongruent versus congruent average activations (respectively of simulated ACC and lateral PFC) in the MI condition compared to the MN and MC condition, as in the corresponding human data reported in Fig. 2.

data indicated that there was increased sustained activity in right-lateralized PFC during the MI condition (lower right panel), but increased transient activity in a different PFC region (primarily left lateralized—lower left panel) and the ACC (lower right panel) specifically on incongruent trials in the MC condition. These effects are consistent with the idea that short time-scale conflict on incongruent trials engaged reactive control in the MC condition, but that protracted occurrences of conflict in the MI condition engaged sustained proactive control (via increased active maintenance of task-set information). The increase in proactive control led to a corresponding decrease in the demand for reactive control.

Simulation results (Fig. 3) demonstrate that the model provided an excellent fit to both the behavioral and brain imaging data. In particular, the model closely matched the changes in reaction times observed across conditions (upper left panel). Reaction time was simulated as the number of cycles necessary to the response units to reach a threshold after stimuli presentation [6]. The model also fit additional features of the behavioral data, including condition effects on accuracy and the shape of the reaction time distribution (not shown) effects and the pattern of reaction time distribution. Likewise, the model also showed changes in both transient (within-trial interval) (upper right and lower left panels) and sustained (inter-trial interval) activation (lower right panel) of conflict and task-units that corresponded well to the empirical pattern observed.

4. Conclusions

The model describes a non-homuncular mechanism that may explain the dramatic shift in human behavior and brain activity during different conditions of the Stroop cognitive control task. Specifically, the model suggests that local and sustained experience of conflict during performance might lead to a shift in the neural mechanisms of cognitive control engaged to perform the task. Importantly, since individuals may not have been explicitly aware of changes in task context across condition, the adjustments in cognitive control had to occur implicitly, and without conscious intervention. The model provides an explicit set of computational mechanisms by which such cognitive control adjustments might be achieved in the brain.

Acknowledgments

This research was supported by ONR N00014-04-1-0004 and R01 (NIH) MH06607803. Christine Hoyer provided invaluable assistance in the collection and analysis of empirical data.

References

[1] M. Botvinick, T.S. Braver, D. Barch, C. Carter, J. Cohen, Conflict monitoring and cognitive control, Psychol. Rev. 108 (2001) 625–652.

[2] T.S. Braver, C.M. Hoyer, Neural mechanisms of proactive and reactive cognitive control, submitted.

[3] C.S. Carter, T.S. Braver, D.M. Barch, M. Botvinick, D.C. Noll, J.D. Cohen, Anterior cingulated cortex, error detection, and the online monitoring of performance, Science 280 (1998) 747–749.

[4] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Nat. Acad. Sci. 79 (1982) 2555–2558.

[5] J.G. Kerns, J.D. Cohen, A.W. MacDonald III, R.Y. Cho, V.A. Stenger, C.S. Carter, Anterior cingulate conflict monitoring and adjustments in control, Science 303 (2004) 1023–1026.

[6] J.L. McClelland, On the time-relations of mental processes: an examination of systems of processes in cascade, Psychol. Rev. 86 (1979) 287–330.

[7] R.C. O'Reilly, Y. Munakata, Computational Explorations in Cognitive Science, MIT Press, Cambridge, MA, 2000.

Nicola De Pisapia is a research associate in the Cognitive Control and Psychopathology laboratory at the Washington University in Saint Louis. He received a Laurea (equivalent to a Master) in Philosophy and Artificial Intelligence from the University of Napoli in 1995, a Master in Philosophy and Computation from Carnegie Mellon University in 2000, and a Ph.D. in Computer Science from the University of Edinburgh in 2003. His research interests include neurocomputational models of prefrontal cortex activity, behavioral planning and cognitive control.

Todd S. Braver is an associate professor and director of the Cognitive Control and Psychopathology laboratory at the Washington University in Saint Louis. He received his Ph.D. in Cognitive Neuroscience from Carnegie Mellon University in 1997. His research interests include understanding the neural and computational mechanisms of cognitive control, their breakdown in different populations, and their interaction with individual differences and emotion.