Extending the Contention Scheduling Model of Routine Action Selection: The Wisconsin Card Sorting Task and Frontal Dysfunction

Richard P. Cooper (R.Cooper@bbk.ac.uk)
Department of Psychological Science, Birkbeck, University of London
Malet Street, London WC1E 7HX, UK

Abstract

We extend a previously developed model of routine action selection by incorporating functional components to support behaviour in a simple non-routine task – sorting cards according to a rule that must be discovered by the subject. A minimal extension to the previous model, consisting of an activation-based working memory/inference system in which evidence is incorporated by simply exciting or inhibiting relevant rule nodes, is demonstrated to be capable of capturing basic performance on the task. The task is commonly used in assessing frontal brain injury, and the extended model is further shown to be capable of capturing the gross behavioural characteristics of frontal patients. However, it is argued that a purely activation-based working memory cannot capture the requirements of more complex tasks. The paper thereby demonstrates 1) how the basic routine action model might be extended to more complex behaviours, but 2) that such behaviours require more than simple activation-based memory processes to structure non-routine behaviour over time.

Keywords: Cognitive architecture; contention scheduling; supervisory system; Wisconsin card sorting task; Frontal dysfunction.

Introduction

Norman and Shallice (1986) argued, on the basis of evidence from slips and lapses in naturalistic everyday action and the more severe errors of patients with frontal lesions, that action is controlled by two systems: a low-level routine system (*contention scheduling*) which is responsible for behaviour in routine or mundane situations when our attention is not focused on action, and a higher-level non-routine system (the *supervisory system*) which works by biasing contention scheduling when acting in novel situations or when it is necessary to avoid temptation. (See Shallice (2006) for an updated overview of the account.) In previous work we have developed a model of the contention scheduling component of the theory, and shown how everyday slips and lapses (Cooper & Shallice, 2000), as well as the more flagrant errors of action that occur following frontal (Cooper et al., 2005) and left parietal (Cooper, 2007) brain injury, may be accounted for in terms of damage to different parts of the contention scheduling system. Previous computational work has not, however, considered in any detail how the supervisory system might act to bias contention scheduling in non-routine situations. This paper begins to redress this omission by considering how the contention scheduling model might be extended to capture behaviour on a simple neuropsychological task that requires both inhibition of a prepotent response and generation of novel (or at least novel with respect to the task) behaviours.

The task we consider is the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948). Subjects in the task are required to sort a series of cards, presented one at a time, into four piles. Drawn on each card is a set of shapes (e.g., two red circles or four blue squares). The piles to which the cards must be sorted are indicated by "target" cards. Each target card differs with respect to the number, colour and shape of items it depicts (see Figure 1). Thus subjects may sort cards to match the targets on any of the three dimensions. During the task, subjects are given feedback after sorting each card, and are required on the basis of this feedback to infer the correct sorting rule and use it for sorting subsequent cards. The trick is that once the subject correctly sorts 10 cards in sequence, the experimenter changes the sorting rule without warning. The subject must then use feedback to adjust his/her sorting rule. This is more difficult than it might at first seem, as some cards match the targets on multiple dimensions, so feedback can be ambiguous. Even so, neurologically healthy subjects have little difficulty on the task. For example, in a sample of 48 subjects tested at Birkbeck, mean sorting accuracy was over 40 correct out of 64 cards. Patients with frontal lesions, however, are known to perform poorly (see, e.g., Stuss et al., 2000), frequently successfully determining the first sorting rule but failing to change rules following negative feedback, i.e., they make perseverative errors.

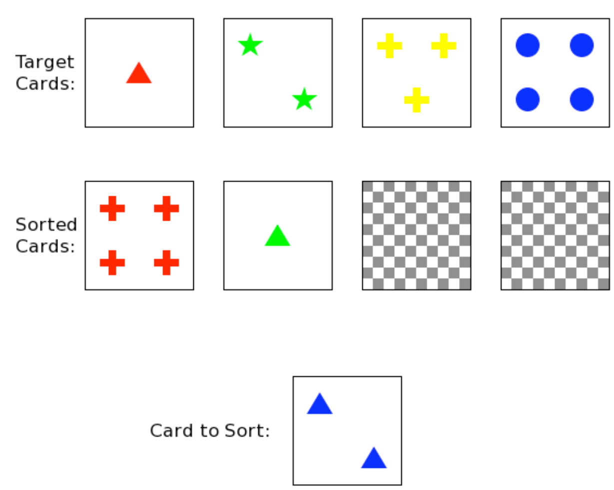

Figure 1: The Wisconsin Card Sorting Test, after two cards have been sorted according to the colour of their symbols and as preparing to sort the third card.

Extending the CS Model: A Naïve Model of WCST

We consider first a naïve and somewhat minimal extension of the contention scheduling model that is capable of completing the WCST at levels comparable to neurologically healthy adults.

Model Assumptions and Description

As discussed above, we assume that behaviour is the product of a simple scheduling system capable of effecting routine sequential behaviour (*contention scheduling*) regulated or biased by a more complex system capable of planning, reasoning and structuring behaviour in the pursuit of intentions (the *supervisory system*). The contention scheduling system has been described in detail elsewhere (e.g., Cooper & Shallice, 2000; Cooper et al., 2005; Cooper, 2007). At its heart is a hierarchically structured interactive activation network in which schemas that encode familiar goal-directed action sequences compete for the control of behaviour, with competition implemented through lateral inhibition between sets of schemas that correspond to alternate ways of achieving a desired goal or sets of schemas that share cognitive or effective resource requirements. The schema network is complemented by further interactive activation networks in which nodes represent objects (with separate object representation networks for different abstract object functional roles). The networks interact, such that schema nodes may excite object representation nodes and vice versa. These interactions encode actions that may be facilitated or afforded by the state of the environment (e.g., that a card on the table might be picked up, or that a card in hand might be placed on the table).

The naïve model of WCST assumes that the contention scheduling system includes schemas for sorting cards according to the different criteria (i.e., *sort by colour*, *sort by number* and *sort by form*), and supplements it with a minimal supervisory (or control) system capable of biasing a specific sorting schema on the basis of positive or negative feedback obtained during the task. The key component of the minimal supervisory system is an activation-based working memory system that contains nodes corresponding to the different schemas that might be used for sorting the cards. It is assumed that when a card is presented for sorting, the most active working memory element biases the corresponding schema within the contention scheduling system, resulting in the card being sorted according to the corresponding criterion (assuming that the scheduling system is functioning correctly). Positive feedback from the experimenter (indicating that the card was sorted correctly) results in excitation of all working memory nodes consistent with the attempt, while negative feedback (if the card was sorted incorrectly) results in inhibition of all working memory nodes consistent with the attempt. Thus, if the card to be sorted depicts one green triangle, and the card is placed under the left-most target card (which in the standard test shows one red triangle), positive feedback will result in excitation of both the sort-to-number and the sort-to-form working memory nodes, while negative feedback will result in inhibition of both of these nodes.

In order to give behaviour coherence over time, we assume that the activation of nodes within working memory persists over time, but that this persistence is imperfect (i.e., activation decays). We also assume that the activation of nodes is subject to noise. For simplicity we adopt for the working memory component the same activation-update equations used in the interactive activation networks, namely:

$$A_{t+1} = \overline{\sigma}\left(\sum_{i=0}^{t} P^i . I_{t-i}\right)$$

where A_t is the activation of a node on card sorting step t, I_t is the net input (excitation or inhibition plus noise) to the node on card sorting step t, P is a persistence parameter (see below) and $\overline{\sigma}$ is a sigmoidal squashing function bounded between 0 and 1 whose output, with zero net input, is 0.1.

With this activation-update equation, activation of working memory nodes tends to 0.1 in the absence of any net excitation or inhibition. Net excitation pushes the activation of a node towards 1, while net inhibition suppresses the activation of node towards 0. Given this formulation, the behaviour of the supervisory aspects of the model is determined by four parameters:[1]

- P: The persistence of working memory representations across card sorting steps.
- N: Standard deviation of noise added to the input of working memory representations on each card sorting step.
- F_e: Excitatory activation of matching working memory representations following positive feedback – a non-negative real number.
- F_i: Inhibitory activation of matching working memory representations following negative feedback – a non-negative real number.

Behaviour of the Model

As anticipated, with appropriate parameter settings the model is capable of performing the WCST with relatively few errors. Thus, in a typical run with $P = 0.85$, $N = 0.05$, $F_e = 0.25$ and $F_i = 0.75$, the model succeeds in correctly sorting approximately 55 cards out of 64, with all errors occurring following a change in sorting category. This corresponds to the upper limit of normal performance.

A full explanation of the model's behaviour requires explanations at the level of both working memory and contention scheduling. We begin with working memory. Figure 2 shows the activation profiles of working memory elements over the complete duration of one administration of the WCST (64 cards) with the above parameter settings.

[1] Additional parameters govern the behaviour of the contention scheduling component of the model. For all simulations reported in this paper we fix those parameters to the values used in other recent work (e.g., Cooper et al., 2005; Cooper, 2007).

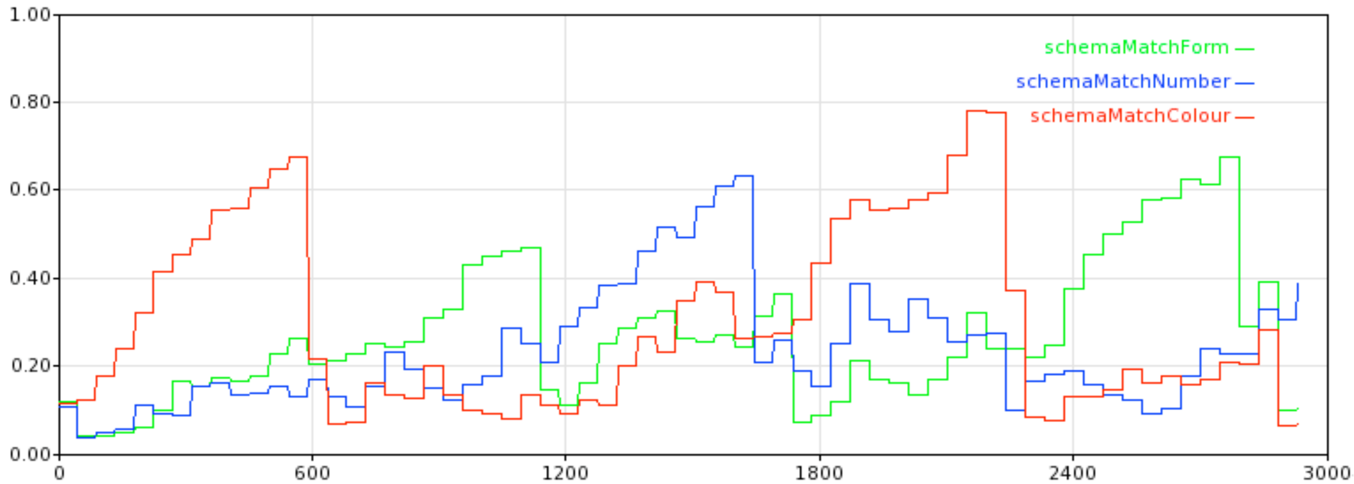

Figure 2: Activation profiles of working memory elements over the duration of the WCST. Activation is plotted on the vertical axis with processing cycles plotted on the horizontal axis.

Each step in the graphs (corresponding to approximately 40 processing cycles, see below) corresponds to the sorting of one card. On the first step working memory elements corresponding to all three sorting schemas have activations close to resting levels, with only noise differentiating them. In this example, the most active element is that which corresponds to matching to *form*. This is therefore selected as the initial rule. The corresponding schema within contention scheduling then receives top-down excitation from the supervisory system, resulting (as discussed below) in the first card being placed under the target card that shares the *form* feature. The first card depicts one green triangle, so matching to form involves matching this card with the left-most target card, which depicts one red triangle. This is incorrect – colour is initially the correct sorting criterion – so negative feedback is provided. This results in inhibition of the working memory representations of all schemas that are consistent with the current sorting attempt. Note though that this attempt matched against two criteria, sorting by form and sorting by number. Hence, the working memory representations of both receive inhibition. The working memory representation corresponding to sorting by colour is the only one not to receive inhibition, and hence is the representation that is most active when the second card is presented. The second card is therefore sorted by colour. Positive feedback results in excitation of this working memory representation, ensuring that it remains the most active, while the activations of the other nodes begin to return to their resting levels.

The model continues sorting by colour, with feedback occasionally providing support for multiple working memory representations (when a card matches against more than one criterion). Only when the criterion changes (after ten successful sorts to the colour criterion) does sorting to colour result in negative feedback. The representation of sorting to colour in working memory is rapidly inhibited, while the representation of sorting to form is excited (through positive feedback when a card matches against the form criterion). Once the activation of the representation of sorting to form exceeds that of sorting to colour (and sorting to number) the model switches to sorting to form (i.e., on presentation of a card, top-down excitation is passed to the schema that corresponds to sort-by-form within the contention scheduling system).

We turn now to the contention scheduling component. Figure 3 shows the activation profile of schema nodes within this component of the model over the first two sorting events. On presentation of the first card, top-down excitation is passed to the *sort-by-form* schema as described above. This results in that schema's activation rising to its maximum level during the first few processing cycles. The *sort-by-form* schema activates in turn the subschemas corresponding to *pick-up card* and *put-down card*. It also activates representations of cards in the object representation networks (which are not shown in the figure). Thus, the presented card (rather than, e.g., the target card) is activated as the card to be picked-up and, once the presented card is held, the target key card which matches this on the form feature is activated as the destination for the *put-down card* schema. The first card is therefore placed under the left-most key card.

Processing is similar during sorting of the second card

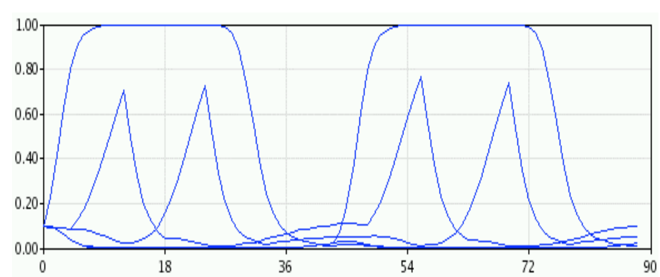

Figure 3: Activation profiles of schema nodes within contention scheduling during two consecutive sorting events. The vertical axis shows activation while the horizontal axis shows processing cycles. The first peak within each sorting event (cycles 12 and 56) corresponds to picking up a card while the second corresponds to placing it in the appropriate target pile (cycles 24 and 69).

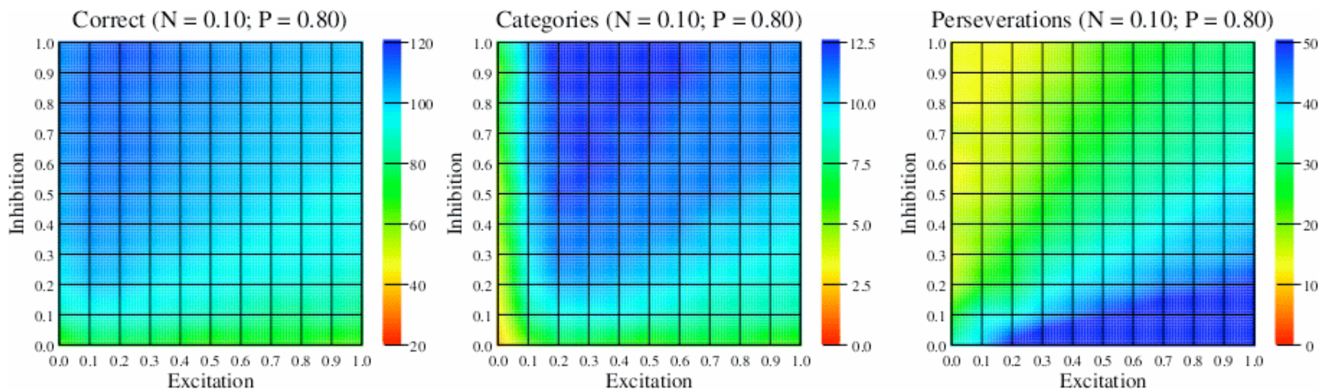

Figure 4: Contour maps showing the number of cards correctly sorted (out of 128), number of categories achieved and classical perseverations when N is 0.1, P is 0.8, and F_i and F_e vary from 0.0 to 1.0. Data are averaged over 10 attempts at each parameter combination.

(cycles 42 to 85), except that it is the *sort-by-colour* schema that is most active, and hence the card that is being sorted is placed under the target key card that matches it on the colour feature.

Parameter Study 1: "Normal" Behaviour

The behaviour of this naïve model depends upon the values of the model's four parameters. Thus, good performance requires that inhibition following negative feedback (F_i) is substantially greater than excitation following positive feedback (F_e). If not, the model will perseverate following negative feedback, as positive feedback during a run of correct responses will result in the working memory representation of the correct sorting criterion becoming highly active, and it will take several consecutive errors following a change in criterion for this activation to subside and be exceeded by that of a competing sorting criterion. At the same time, persistence must be relatively high. If it is too low, behaviour on each card sort will be based primarily on feedback from the previous trial – feedback that can be ambiguous if a card matches against multiple criteria.

Given the potential complexity of interactions between parameter values, two systematic surveys of the parameter space were conducted. In parameter study 1, the model's susceptibility to standard perseverative errors was investigated by varying F_e, F_i and P from 0.0 to 1.0 in steps of 0.1 with N at 0.1, 0.2 and 0.3. The model was run 10 times at each point in the parameter space, and three dependent variables – the number of correct sorts, categories achieved and classical perseverative errors – were recorded for each run of the model. In each case the model was required to sort 128 cards, with the simulated experimenter changing the sorting criterion whenever 10 consecutive cards were sorted correctly. Thus, following Stuss et al. (2000) but unlike most behavioural studies, the test was not terminated after 6 categories had been achieved. Scoring was automated by a separate program that implemented the scoring algorithm described by Heaton (1981).

These simulations demonstrated that, for each value of N, there are values for the other parameters that result in accurate sorting with few errors (e.g., $N = 0.1$, $P = 0.9$, $F_i = 0.1$, $F_e = 0.8$) that is similar to the behaviour of normal participants. They also demonstrated, however, that the model generates high numbers of perseverative errors (i.e., more than 1/3rd of responses) and achieves relatively few categories when P is high and F_i is low relative to F_e. Thus, Figure 4 shows contour maps for the number of cards correctly sorted, number of categories achieved, and number of perseverative errors when N is 0.1, P is 0.8, and F_i and F_e vary from 0.0 to 1.0. From the figure, it can be seen that N is 0.1, P is 0.8, F_i is 0.1 and F_e is 0.7, the model correctly sorts 60 to 80 cards (out of 128), obtains 5.0 to 7.5 categories, but produces 40 to 50 perseverative responses.

Parameter Study 2: "Frontal" Behaviour

It is clear from parameter study 1 that the naïve model is susceptible to perseverative behaviour, at least when persistence is high and feedback inhibition is low relative to feedback excitation. While this echoes the behaviour of certain frontal patients, the number or proportion of perseverative errors alone is a coarse measure of behaviour. Parameter study 2 therefore sought to evaluate the model's performance against a published dataset with a more fine-grained scoring system, namely the dataset and scoring system of Stuss et al. (2000).

Stuss et al. (2000) tested six groups of patients (four groups with frontal lesions centred in different areas and two non-frontal patient groups) and control participants on three versions of the WCST, with increasing instructional support on successive versions. In scoring participant behaviour, errors were subdivided into four categories: perseveration of preceding category (PPC: a response that matches the previous sorting criterion but not the current one), perseveration of preceding response (PPR: a response that matches exactly the features matched on the immediately preceding incorrect trial), set loss (an error following attainment of the current sorting category, as demonstrated by three consecutive correct responses, at least one of which was non-ambiguous) and other errors. Subtle differences between the various frontal groups were observed. For example, when participants were told the possible sorting criteria prior to the test (Stuss et al.'s 64A

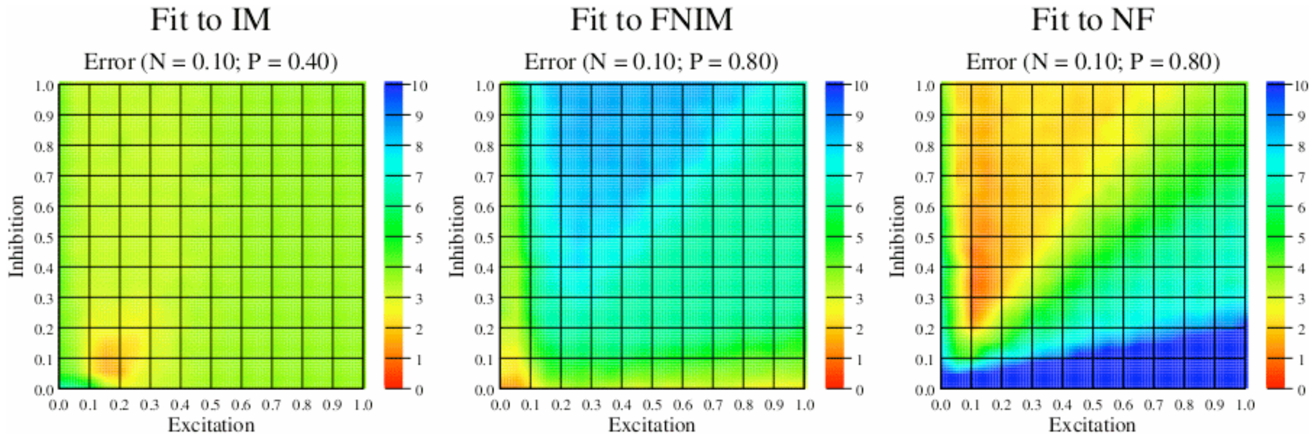

Figure 5: Goodness of fit plots for best fitting planes in $F_e \times F_i$ space for each of the three groups. The best fit to the inferior medial group (IM: left panel) occurs with $P = 0.40$. The best fits for the frontal non-inferior medial (FNIM: centre panel) and the non-frontal (NF: right panel) groups occur with P = 0.80.

condition), inferior medial patients achieved significantly fewer sorting categories and produced significantly more set loss errors than control and non-frontal patients, but they did not make significantly more PPC or PPR errors. Other frontal groups achieved even fewer categories and made fewer set loss errors than the inferior medial patients, but made significantly more PPC and PPR errors than the inferior medial, non-frontal and control groups.

Parameter study 2 therefore explored the behaviour of the model following variation of F_e, F_i and P using the scoring system of Stuss et al. (2000). The aim was to replicate the behaviour of each of Stuss et al.'s participant groups and thereby further understand the possible nature of the deficit in each case. Note, however, that Stuss et al. found no significant differences in the pattern of behaviour between their right dorsolateral, left dorsolateral and superior medial groups – all three groups produced qualitatively similar behaviour across the four dependent variables. These frontal groups did differ, however, from the inferior medial group. Our analysis therefore merges these groups. Similarly, Stuss et al. found no significant differences between their left non-frontal, right non-frontal and control groups. Our analysis also merges these groups. This results in three groups: inferior medial (IM), frontal non-inferior medial (FNIM) and non-frontal (NF). Descriptive statistics for each group based on the 64A version of the task are shown in Table 1.

To explore the parameter space F_e and F_i were varied from 0.00 to 1.00 at intervals of 0.05 and P was varied from 0.10 to 0.90 at intervals of 0.10. N was fixed at 0.10. The model was run 50 times with 64 cards to sort at each combination of parameter values (totalling 21 × 21 × 9 × 50 = 198450 runs). Four dependent measures were collected for each run (categories achieved, PPC errors, PPR errors and set loss errors, all following definitions given in Stuss et al., 2000). For each of the three groups and for each point in parameter space, a goodness of fit measure was then calculated as the maximum of the fits to the four dependent measures, where the fit to each of the dependent measures was calculated as the difference between the simulated mean value of that dependent measure at the point in parameter space and the observed mean value of that dependent measure for the specific group divided by the observed standard deviation of that dependent measure for the group. Thus, a fit of less than one to any group would correspond to a case where each of the four dependent measures was within one standard deviation of the observed group means. Figure 5 shows plots of this goodness of fit measure for the best fits for each group in $F_e \times F_i$ space.

	Categor-ies	**PPC Errors**	**PPR Errors**	**Set Loss Errors**
NF	4.01 (0.44)	7.15 (1.09)	0.94 (0.68)	0.93 (0.48)
FNIM	1.08 (0.46)	24.27 (6.04)	11.68 (3.18)	1.14 (0.63)
IM	2.60 (0.60)	10.60 (1.70)	2.90 (0.9)	2.60 (0.70)

Table 1: Means (standard deviations) for WCST behaviour of three patient groups (derived from Stuss et al., 2000)

From Figure 5 it can be seen that the best fit to the IM group is obtained when P is 0.40, F_i is between 0.05 and 0.10 and F_e is between 0.15 and 0.20. This fit is approximately 1.5. A slightly better fit is obtained for the FNIM group, of 1.0, when P is 0.80, F_i is 0.00 and F_e is 0.05. Only for the NF group is a fit of less than one obtained, and when P is 0.80 this level of goodness of fit is obtained for a wide region of $F_e \times F_i$ space (and this result holds for other values of $P \geq 0.70$).

Discussion

The naïve model has been shown to be capable of both normal and frontal-like behaviour on the WCST (parameter study 1), but the scan of the parameter space in parameter study 2 found only modest fits for the two subgroups of frontal patients, with the best fits in each case failing to be simultaneously within one standard deviation for all dependent measures. There may be good reason for this – none of the subject groups is completely homogenous, and even if all patients in a group can be argued to have a qualitatively similar deficit, that deficit is likely to vary in degree. Nevertheless the naïve model does provide some insight into the deficits. Inferior medial frontal patients are particularly prone to PPR errors and set loss errors. These errors occur when excitation, inhibition and persistence are all low. The latter provides a clear intuitive account of set loss errors: if persistence is low it is likely that the model will frequently fail to maintain a sorting rule, even after successfully discovering the rule. If both excitation and inhibition are low the model effectively makes little use of either positive or negative feedback. This explains to some extent the existence of perseverative errors. However, the type of perseverative error depends upon maintaining some record of a sorting rule. For PPR errors, this cannot be the most recent successful sorting rule – that would result in PPC errors. Rather, it is the rule apparently used *unsuccessfully* on the previous trial. PPR errors are therefore a more accurate reflection of failure to respond to negative feedback than are PPC or classical perseverative errors.

General Discussion

In comparison with previous work, the model shares a family resemblance with models inspired by the operation of prefrontal cortex (e.g., Dehaene & Changeux, 1991; Amos, 2000; Rougier et al., 2005). Like these models, behaviour in the extended contention scheduling model is a function of bias operating on a routine system that, in the case of card sorting, embodies simple stimulus-response links. The work presented here differs from the above, however, in considering the behaviour of different frontal subgroups as revealed by Stuss et al (2000).

The extended contention scheduling model does moderately well at accounting for both normal and impaired performance, but there are severe limitations to the working memory module. Both basic assumptions – that working memory comprises nodes corresponding to atomic symbols and that evidence accrues only through processes of excitation and inhibition – are problematic. Thus, the approach does not generalize well to other non-routine behaviours such as solving Tower of Hanoi problems, which appear to require both the storage and manipulation of structured information within working memory and the manipulation of that information according to operations more complex than simple excitation or inhibition.

Indeed, in an alternative extension of the contention scheduling model to be reported elsewhere working memory has been modelled as a collection of feature-value pairs (similar to production system approaches). Space limitations prevent a full description of the model. However, as with the naïve model presented here the alternative model was able to capture normal and impaired performance on the WCST. More critically, the working memory structures of the alternative model allow it to be applied to other non-routine tasks, including solving Tower of London problems and generating random sequences of numbers – both non-routine tasks that have frequently been discussed in the literature on cognitive control. In these tasks, autonomous functioning of the lower-level system supports the solution of one-move tower problems and the generation of sequences of associated numbers (e.g., digits increasing by two). Again, the role of the supervisory system is to modulate behaviour. The system allows, in the first case, the solution of tower problems where intermediate states are required, and in the second, detection and inhibition of stereotyped responses before they are produced. This is achieved through operations on the content of working memory which depend on relations between working memory elements. It is unclear how the working memory mechanisms of the naïve model (or of other models such as those mentioned above, and also the recent influential working memory model of O'Reilly and Frank (2006)) might meet such a challenge.

References

Amos, A. (2000). A computational model of information processing in the frontal cortex and basal ganglia. *Journal of Cognitive Neuroscience*, *12*, 505–519.

Cooper, R. P. (2007): Tool Use and Related Errors in Ideational Apraxia: The Quantitative Simulation of Patient Error Profiles. *Cortex*, *43*, 319–337.

Cooper, R.P., Schwartz, M.F., Yule, P., & Shallice, T. (2005). The simulation of action disorganisation in complex activities of daily living. *Cognitive Neuropsychology*, *22*, 959–1004.

Cooper, R.P., & Shallice, T. (2000). Contention Scheduling and the control of routine activities. *Cognitive Neuropsychology, 17*, 297–338.

Dehaene, S. & Changeux, J-P. (1991). The Wisconsin Card Sorting Test: Theoretical analysis and modelling in a neuronal network. *Cerebral Cortex*, *1*, 62–79.

Grant, D.A. & Berg, E.A. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*, 404–411.

Heaton R. (1981). *A manual for the Wisconsin Card Sorting Test*. Psychological Assessment Resources. FL: Odessa.

Norman, D.A. & Shallice, T. (1986). Attention to action: willed and automatic control of behaviour. In R. Davidson, G. Schwartz & D. Shapiro (eds.) *Consciousness and Self Regulation, Volume 4*. NY: Plenum.

O'Reilly, R.C. & Frank, M.J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328.

Rougier, N.P., Noelle, D.C., Braver, T.S., Cohen, J.D. & O'Reilly, R.C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the USA*, *102*, 7338–7343.

Stuss, D.T., Levine, B., Alexander, M.P., Hong, J., Palumbo, C., Hamer, L., Murphy, K.J. & Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*, 388–402.