

A Model of Cognitive Control in the Wisconsin Card Sorting Test: Integrating Schema Theory and Basal Ganglia Function

Andrea Caso
(andrea@andreacaso.com)

Richard P. Cooper
(R.Cooper@bbk.ac.uk)

Centre for Cognition, Computation and Modelling
Department of Psychological Sciences, Birkbeck, University of London
Malet Street, London, WC1E 7HX, United Kingdom

Abstract

We present a schema-based model of a classic neuropsychological task, the Wisconsin Card Sorting Task (WCST), where competition between motor and cognitive schemas is resolved using a variation of a neuroanatomically detailed model of the basal ganglia (Gurney et al., 2001). We show that the model achieves a good fit with existing data at the group level, and correctly identifies two distinct cognitive mechanisms held to underlie two distinct types of error. However, at the individual level, the correlations amongst other error types produced by the model differ from those observed in the human data. To address this, we cluster participant performance into distinct groups and show, by fitting each group separately, how the model can account for the empirically observed correlations between error types. Methodologically, this demonstrates the importance of modelling participant performance at the sub-group or individual level, rather than modelling group performance. We also discuss implications of the model for the WCST performance of elderly participants and Parkinson's patients.

Keywords: schema theory; contention scheduling; basal ganglia; Wisconsin Card Sorting Task; modelling individual performance

Introduction

Schema theory is a framework based on the idea that behaviour in many areas depends on abstractions over instances, i.e., schemas. In these abstract terms, schema theory is very general. It has been applied in domains ranging, for example, from event memory (Bartlett, 1932) to motor control (Schmidt, 1976). Norman and Shallice (1980) applied the theory in the domain of routine sequential action. Their theory proposes that action schemas work in a cooperative or sequential fashion, but also compete with each other for activation. While schema theory is helpful in representing functional interactions in the action-perception cycle, it is not committed to a specific neural implementation. However, at the neural level the basal ganglia have been proposed as a good candidate for resolving competition between schemas in order to carry out action selection (Redgrave et al., 2001). In part this is because of their recurrent connections with the cortex.

In this paper we present a model of the Wisconsin Card Sorting Task (WCST) where competition between motor and cognitive schemas is resolved using a variation of a neuroanatomically detailed model of the basal ganglia. We

use a genetic algorithm to search the model's parameter space and obtain a good fit for the data. Further analysis of correlations between error types, however, suggests the need to model individual participant data. Yet for reasons of computational efficiency this is impractical. We therefore cluster participant performance into a small number of distinct groups (5) and run separate genetic algorithms to fit the groups individually. The results capture both group performance and correlations between error types across individuals.

The Task and the Model

In the WCST, participants are required to sort a series of cards into four categories based on binary (i.e., correct / incorrect) feedback. Each card shows one, two, three or four shapes, printed in one of four colours, and there are four shapes (triangle, star, cross, circle). It is therefore possible to sort cards according to colour, number or shape. To succeed, participants must match each successive card with one of four target cards (One Red Triangle, Two Green Stars, Three Yellow Crosses, Four Blue Circles), and use the subsequent feedback to discover the appropriate rule, but once they have discovered the rule (as indicated by a succession of 10 correct sorts), the experiment changes the rule without notice. The task yields a number of dependent measures, including the number of rules obtained (with a deck of 64 cards), the number of cards correctly sorted, the number of perseverative errors (where negative feedback is ignored) and the number of set-loss errors (where the participant fails to stick with a successful rule).

The model comprises three cognitive schemas and four motor schemas (see Fig. 1).¹ Cognitive schemas represent the selection rules (Sort by Colour, Sort by Number, Sort by Shape) while the four motor schemas represent the acts of putting the stimulus card below each of the four target cards. Each schema has an activation level that varies over time as a function of input from various sources. Cognitive schemas are fed by an external channel that changes by a fixed amount according to external positive/negative feedback. Motor schemas are fed by cognitive schemas, and this signal is rule-dependent. If, for instance, the stimulus card displays

¹Source code for the simulation, including a complete list of parameters and their values, is available from the first author on request.

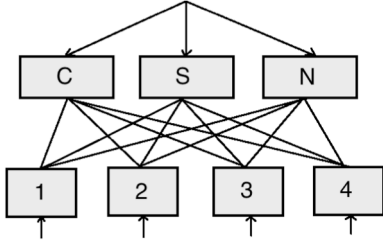


Figure 1: Schematic of the model, not showing competition between schemas. Cognitive schemas (top row) send signals to the motor schemas (bottom row)

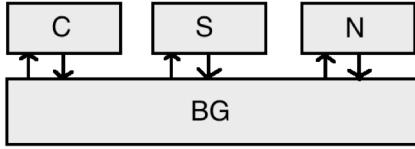


Figure 2: Schematic of the competition between schemas. The basal ganglia units compute the amount of inhibition that each schema receives given the activation of the others.

three red circles, the colour schema will excite the fourth motor schema (Four Blue Circles), the shape schema will excite the third motor schema (Three Yellow Crosses), and the colour schema will excite the first motor schema (One Red Triangle). Motor schemas are also fed by environmental cues depending on the stimulus card feature. Thus, when cognitive schemas are not strong enough to influence motor schemas, action selection may be driven by stimulus features only.

This simple model is complemented by a mechanism that implements and resolves competition between schemas within each hierarchical level: cognitive and motor schemas feed into two parallel computational mechanisms that each return a signal in the form of inhibition to the individual channels at each level (see Fig. 2 for an illustration at the cognitive level). In the brain, this competition between schemas is thought to be carried out by the basal ganglia

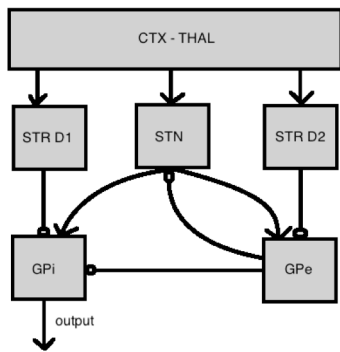


Figure 3: Schematic of the basal ganglia. Legend: Cortex-Thalamic complex (CTX-THAL), Striatum (STR), Subthalamic nucleus (STN), Globus Pallidus Internal/External Segment (GPi and GPe)

(Gurney et al., 2001). Corticobasal loops are mostly segregated (Alexander et al., 1986) and this is reflected in the model through the independence of information processed in the basal ganglia units at the two levels (cognitive and motor).

The model also implements a rudimental learning mechanism. This consists in a fixed change in signal to the cognitive schemas following a reward. Its purpose is to analyse how baseline levels of signal influence schema selection and ultimately, performance on the WCST. Manipulation of the thresholds of saturation functions in cortical units and associated basal ganglia units represent dopamine signalling in the cortex and in the basal ganglia, respectively. Therefore, the mechanism underlying cognitive control is a feedback-driven signal to the cognitive schemas.

Computation in Individual Units

The model consists of 7 cortical units, 3 of which control cognitive operations and 4 of which control motor operations (see Fig. 1). These units correspond to schemas. Cognitive and motor units send their signal to their respective striatal units (see Fig. 3). Subthalamic units connect all units at the same hierarchical level (cognitive or motor), ensuring that the basal ganglia units act as a competitive suppressor of schemas as a function of the other schemas' outputs.

Individual units are connected as shown in Fig. 3. Their computations are shown below. In all cases, u_i represents the entry signal to the unit, a_i is the result of integration along the time domain, and o_i represents the output of the individual units. The function σ computes the sigmoid function of the input, ensuring output values are bounded between 0 and 1. Sigmoid functions have a fixed slope and threshold. Varying the threshold of cortical or striatal units alters the way competition between units is carried out, and can be considered a function of tonic dopamine present in the circuit. (In a separate simulation it has been shown that the level of external dopamine from the substantia nigra pars compacta (SNpc) unit can be simulated by varying the threshold of the saturation curve in the striatum (β_{ctx}), without making use of an additional unit.)

Cortical Units:

$$u_i \Leftarrow \sum_j w_{i,j} \cdot u_j + o_{ext,i} + o_{thal,i}$$

$$a_i(t) \Leftarrow \delta \cdot a_i(t-1) + (1-\delta) \cdot u_i(t-1)$$

$$o_i \Leftarrow \sigma(a_i)$$

Striatum (D1 and D2):

$$u_i \Leftarrow o_{ctx,i}$$

$$a_i(t) \Leftarrow \delta \cdot a_i(t-1) + (1-\delta) \cdot u_i(t-1)$$

$$o_i \Leftarrow \sigma(a_{strD1/D2,i})$$

Subthalamic Nucleus:

$$\begin{aligned} \mathbf{u}_{\text{stn},i}(\mathbf{t}) &\leftarrow \mathbf{w}_{\text{stn}} \cdot \mathbf{o}_{\text{ctx},i} + \mathbf{w}_{\text{gpe_stn}} \cdot \mathbf{o}_{\text{gpe},i}(\mathbf{t} - 1) \\ \mathbf{a}_{\text{stn},i}(\mathbf{t}) &\leftarrow \delta \cdot \mathbf{a}_{\text{stn},i}(\mathbf{t} - 1) + (1 - \delta) \cdot \mathbf{u}_{\text{stn},i}(\mathbf{t} - 1) \\ \mathbf{o}_{\text{stn},i} &\leftarrow \sigma(\mathbf{a}_{\text{stn},i}) \end{aligned}$$

Globus Pallidus (External Segment):

$$\begin{aligned} \mathbf{u}_{\text{gpe},i} &\leftarrow \mathbf{w}_{\text{stn_gpe}} \cdot \sum_i \mathbf{o}_{\text{stn},i} + \mathbf{w}_{\text{strD2_gpe}} \cdot \mathbf{o}_{\text{strD2},i} \\ \mathbf{a}_{\text{gpe},i}(\mathbf{t}) &\leftarrow \delta \cdot \mathbf{a}_{\text{gpe},i}(\mathbf{t} - 1) + (1 - \delta) \cdot \mathbf{u}_{\text{gpe},i}(\mathbf{t} - 1) \\ \mathbf{o}_{\text{gpe},i} &\leftarrow \sigma(\mathbf{a}_{\text{gpe},i}) \end{aligned}$$

Globus Pallidus (Internal Segment):

$$\begin{aligned} \mathbf{u}_{\text{gpi},i}(\mathbf{t}) &\leftarrow \mathbf{w}_{\text{stn_gpi}} \cdot \sum_i \mathbf{o}_{\text{stn},i} + \mathbf{w}_{\text{gpe_gpi}} \cdot \mathbf{o}_{\text{gpe},i}(\mathbf{t} - 1) \\ &\quad + \mathbf{w}_{\text{strD1_gpi}} \cdot \mathbf{o}_{\text{strD1},i}(\mathbf{t} - 1) \\ \mathbf{a}_{\text{gpi},i}(\mathbf{t}) &\leftarrow \delta \cdot \mathbf{a}_{\text{gpi},i}(\mathbf{t} - 1) + (1 - \delta) \cdot \mathbf{u}_{\text{gpi},i}(\mathbf{t} - 1) \\ \mathbf{o}_{\text{gpi},i} &\leftarrow \sigma(\mathbf{a}_{\text{gpi},i}) \end{aligned}$$

Thalamus:

$$\begin{aligned} \mathbf{u}_i &\leftarrow \mathbf{o}_{\text{gpi},i} \\ \mathbf{a}_i(\mathbf{t}) &\leftarrow \delta \cdot \mathbf{a}_i(\mathbf{t} - 1) + (1 - \delta) \cdot \mathbf{u}_i(\mathbf{t} - 1) \\ \mathbf{o}_i &\leftarrow -\sigma(\mathbf{a}_i) \end{aligned}$$

Feedback

Feedback takes place after each trial. If the selected response is correct, the external signals $o_{\text{ext},i}$ to the cognitive units² that correspond to the matched features are increased by a fixed amount b_l . If the selected response is incorrect, inputs to those units that correspond to the matched features are decreased by a fixed amount b_l .

Simulation of Wisconsin Card Sorting Test

Simulation of an Individual Task

To simulate the WCST, a virtual deck of 64 cards is produced, shuffled and presented to the model. All the units perform the computation outlined in the previous section. The first motor unit to reach a fixed activation value (measured as the area under the time-curve, rather than simply as a threshold) is selected. After the selection and feedback, a new card is presented. The resulting plot for activation of the cognitive units is shown in Fig. 4.

As can be seen in Fig. 4, when the first card is presented the system must work out that ‘colour’ is the first correct

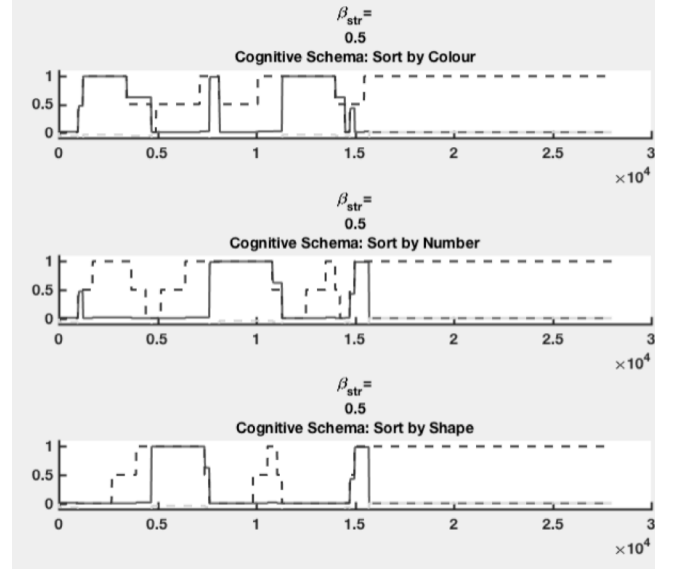


Figure 4: Activation of cognitive schemas during a complete run (involving sorting all 64 cards). Solid lines represent the actual activation while dashed lines represent the external input due to positive/negative feedback. Processing cycles are represented on the horizontal axis.

sorting criterion. Feedback alone is not sufficient, as the selected card may match more than one feature. Basal ganglia units intervene by suppressing the inappropriate cognitive schemas, enabling the correct schema to be permanently selected. When the sorting criterion changes (after 10 correct responses) the system tends to perseverate for a short period of time, before selecting the correct criterion again. Feedback-dependent external activation and resolution of competition both play a role in activating the correct cognitive schemas. Whereas the activation of cognitive schemas is regulated by feedback, the activation of motor schemas is regulated by cognitive schemas and environmental cues.

Parameters

The model has a number of parameters. One important parameter is the threshold of the saturation curve of the striatum, represented by the threshold of the sigma function applied to the striatal output (β_{str}). Extreme values of this parameter (substantially greater than or less than 0.5) disrupt the competition between schemas. When the threshold is too high schemas are driven by their input values and they undergo increasingly homogenous inhibition from the basal ganglia. This phenomenon is analogous to the Parkinson's disease dopamine depletion in the SNpc (Cooper & Shallice, 2000).

Dependent Measures

Performance was scored according to a range of measures as indicated in Heaton (1981). Completed Categories (CC) and Total Errors (TE) measure the overall performance. A Set

²That is, those cortical units that represent cognitive schemas.

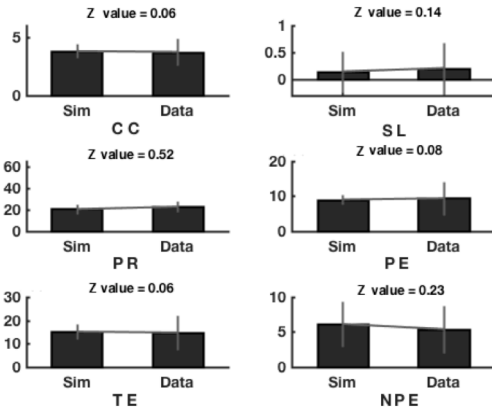


Figure 5: Comparison between Simulation and Data from neurologically healthy young participants. Z values indicate the z score of the difference between human and simulated data for each dependent measure.

Loss Error (SL) is counted whenever an incorrect response is selected after 5 or more correct responses, where at least one is unambiguous (i.e., the card matches only one feature). A Perseverative response (PR) is counted whenever a response would have been correct under the previous rule. (A subject can score a perseverative response even before completing the first category: if three consecutive responses are made selecting the same sorting rule, that rule will be the criterion that the subject can perseverate to.) Those perseverative responses that are also incorrect responses are counted as Perseverative Errors (PE). Non-perseverative errors (NPE) are calculated as the Total Errors (TE) minus Perseverative Errors (PE).

Results

Results for two sets of 48 participants (48 healthy young adults and 48 simulated participants) are depicted in Fig. 5. The figure compares the aggregate results from the simulation (Sim) with the aggregate data from the human participants (Data).

A genetic algorithm attempted to find the best parameters that produce low t statistics and low z statistics between data and simulation. Given the presence of a multitude of parameters that influence each other in a non-linear fashion, a perfect fit is unattainable. However, the model appears to do a good job in reproducing group mean and standard errors, as shown by the figure.

Correlational Analysis

Analysing aggregate data is not sufficient to assess model performance, since a model should also aim to dissociate between psychological constructs (Cassimatis et al., 2008). Therefore, correlational analysis between the most informative variables (TE, PE, SL) was also performed, using bootstrapping and sampling the mean value to obtain 1000 points. Multiple runs of the sampling algorithm produce very similar results. Fig. 6 and Fig. 7 show the

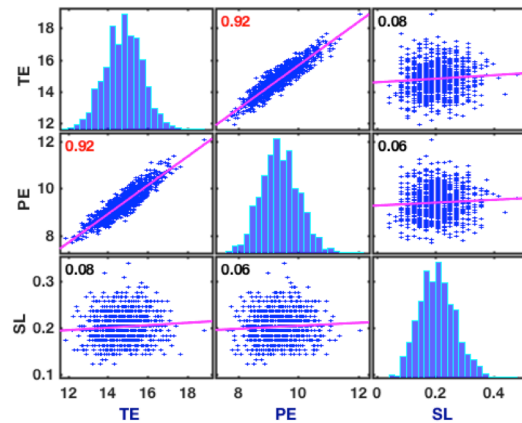


Figure 6: Correlations – Neuropsychological Data

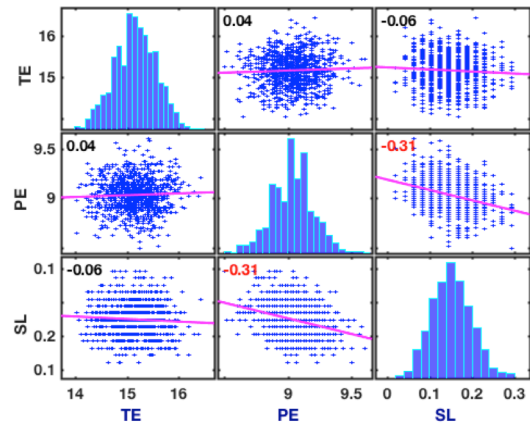


Figure 7: Correlations – Simulation

correlation matrices for these variables in both the human data and the simulation.

The correlation matrices show that the simulation correctly identifies that the mechanism that produces set loss error can be dissociated from the process that causes other types or errors. However, the simulation fails to reproduce the high correlation ($r = .91, p < .01$) between Total Errors and Perseverative Errors. In addition, it displays a weak but significant negative correlation ($r = -.31, p < .01$) that is not present in the empirical data.

Discussion

The model yields an adequate fit for young participants on the WCST. Computation in the model appears to be stable, in that minimal parameter variations do not disrupt functioning. The model also correctly reflects the independence between Set Loss Errors (SL) and Total Errors (TE) found in the human data, suggesting a dissociation in the cognitive processes that produce those errors.

However, the model is subject to several limitations. The lack of positive correlation between PE and TE in the simulation is both puzzling and concerning. One possibility,

however, is that this apparent failing reflects the implicit assumption that performance of the human participants can be modelled by a single set of parameter values (i.e., by a group of 48 virtual participants with identical cognitive characteristics). We explore this possibility in the following section.

Grouping Data

Introduction

In the light of the failure of the model to reproduce the empirically observed correlations between TE and PE, we analyse how data from young participants can be clustered into a small number of groups based on the three critical dependent variables reflecting errors (TE, PE, SL).

These three types of errors have been specifically chosen because they are most representative of performance failures. Data clustering was calculated using a k-means algorithm with $k = 5$ (purely for reasons of computational efficiency). Two points were excluded because they were outliers. The algorithm was initialised based on the observation of the spatial 3D distribution of points. The most distinctive features are the accumulation of points around the origin, the sparseness of points as total and perseverative errors increase, and an isolated cluster of points with SL equal to 1.

Fig. 8 shows how the clustering of the groups and Table 1 shows mean and standard deviation of the dependent variables in the individual groups.

Simulation

After clustering the groups, as outlined in Table 1, we run five genetic algorithms separately to determine best-fitting parameter values for each group. In each case, seven model parameters were initially randomised to values within their reasonable ranges, and model errors recorded. A t-value between the simulation's and the original experimental data was computed and its mean used as the inverse of the GA's fitness value. Table 1 shows performance errors of the simulation with the highest fitness and Fig. 9 shows a 3D representation of the individual values.

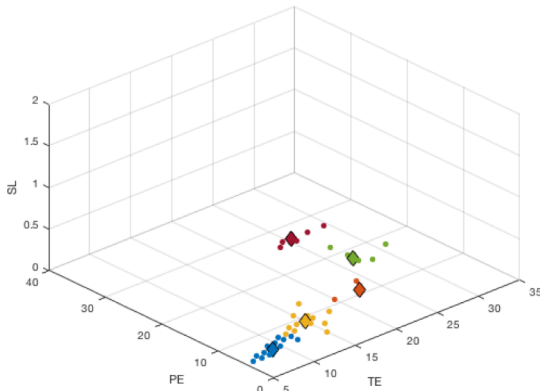


Figure 8: Clustering of experimental data

Table 1: Data Groups

G	N	TE	PE	SL
1 ●	18	8.89 ($SD = 2.03$)	6.22 ($SD = 2.03$)	0 ($SD = 0$)
2 ●	13	14.85 ($SD = 1.77$)	8.77 ($SD = 1.92$)	0 ($SD = 0$)
3 ●	5	28.00 ($SD = 1.73$)	18.40 ($SD = 2.30$)	0 ($SD = 0$)
4 ●	7	14.71 ($SD = 2.63$)	9.57 ($SD = 0.53$)	1 ($SD = 0$)
5 ●	3	22.33 ($SD = 2.08$)	11.67 ($SD = 1.15$)	0 ($SD = 0$)

Table 2: Simulation of the five clusters

G	N	TE	PE	SL
1 ●	18	8.83 ($SD = 1.38$)	5.89 ($SD = 1.08$)	0.11 ($SD = 0.32$)
2 ●	13	14.31 ($SD = 1.55$)	9.23 ($SD = 1.17$)	0.08 ($SD = 0.28$)
3 ●	3	22.00 ($SD = 5.00$)	7.33 ($SD = 0.58$)	0.67 ($SD = 1.15$)
4 ●	6	15.00 ($SD = 2.53$)	10.83 ($SD = 1.72$)	0.5 ($SD = 0.55$)
5 ●	2	18.50 ($SD = 2.08$)	9.50 ($SD = 0.71$)	0.00 ($SD = 0.00$)

Discussion

Results from the simulation are shown in Table 2. In total, 4 outliers have been excluded from the analysis (2, 1 and 1 from categories 3, 4 and 5, respectively). These outliers may conceivably have been produced by the model's unstable response to increasingly higher parameter values. Clustering the participant data into a small number of more homogenous groups greatly increases the correlation between TE and PE (r increases from .04 to .50, compared with the observed value of .92) and decreases the correlation between SL and TE/PE, improving the fit of the model in both respects. Fig. 10 displays the new correlation plots worked out combining all of the five simulations together.

General Discussion

The model we presented combines a variation of the Cooper and Shallice (2000) model of action selection and a variation of the Gurney et al. (2001) model of the basal ganglia. One of the strengths of this combined model is the possibility to generalise it to other cognitive control tasks (e.g. Stroop task, Probabilistic Reversal Learning, Eriksen Flanker Task, etc.) and to accommodate the presence of

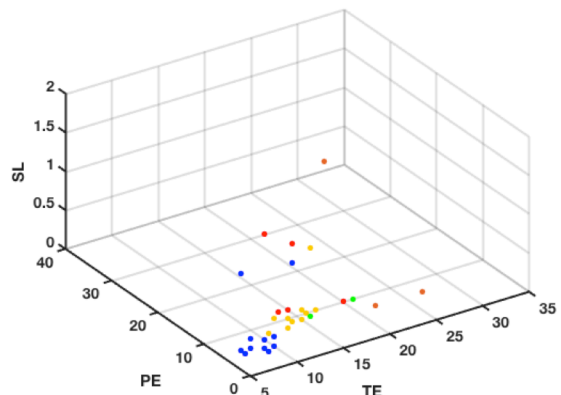


Figure 9: Simulated data with five clusters

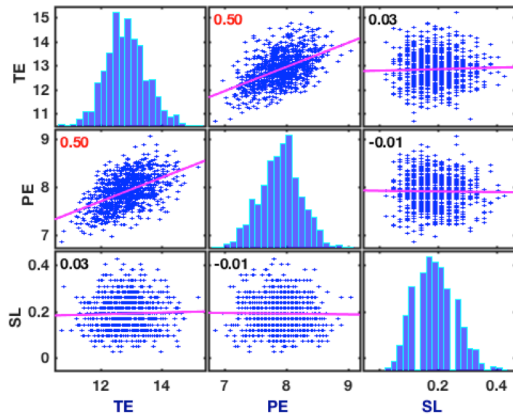


Figure 10: Correlation between performance errors aggregating the values from five different set of parameters

units representing other brain areas where different computation is performed (e.g., amygdala, cerebellum), enabling the simulation of cognitive tasks in broader contexts (e.g. Emotional Stroop Task, WCST in cerebellar patients). In principle, this enhances the contention scheduling theory with neuroanatomical detail, allowing a more precise localisation of processes in a particular task, and integration with functional neuroimaging data. In addition, this implementation allows for the inclusion of two distinct learning mechanisms in the cortex and the basal ganglia: the current model can potentially be updated to a learning-based model by developing these mechanisms.

With respect to cortical learning, in the model as it stands, the supervisory system that controls how subjects respond to positive and negative feedback is fixed and consequently performance tends to be too robust to basal unit dysfunctions. This might be addressed by incorporating dynamic learning that allows supervisory control to vary according to the schemas' activations, resulting in low or high baseline levels of dopamine in the striatum having a greater impact on cognitive performance.

The present paper makes the case for modelling subgroup data (or, whenever possible, individual data), instead of aggregate results, and presents evidence of how data clustering improves the model overall fit. Clustering is especially advisable for models of higher-order cognition, where subjects tend to have variable attention and may use qualitatively different cognitive strategies.

A final conclusion emerges from two joint observations: First, fitting clusters with increasingly extreme error values becomes increasingly more problematic. Second, another set of simulations (not reproduced here) shows that damaging the cortical and subcortical units threshold does not seem to produce the level of decline in performance found in Parkinson's disease patients without dementia (Paolo et al., 1996). Since healthy older controls have a different performance profile than the younger controls against which the current model was assessed, the loss of dopaminergic

cells in SNpc does not alone explain the inferior performance in the elderly and PD patients³.

These two joint findings suggest that the cognitive mechanisms producing perseverative and set loss errors might be independent only for a small number of errors. As that number increases, these two mechanisms might be correlated and possibly causally related. New experimental data to confirm this hypothesis is warranted.

References

- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9(1), 357-381.
- Bartlett, F. C. (1932). *Remembering. A Study in Experimental and Social Psychology*. Cambridge University Press, Cambridge.
- Bezard, E., Gross, C. E., & Brotchie, J. M. (2003). Presymptomatic compensation in Parkinson's disease is not dopamine-mediated. *Trends in Neurosciences*, 26(4), 215-221.
- Cassimatis, N. L., Bello, P., & Langley, P. (2008). Ability, breadth, and parsimony in computational models of higher order cognition. *Cognitive Science*, 32(8), 1304-1322.
- Cham, R., Studenski, S. A., Perera, S., & Bohnen, N. I. (2008). Striatal dopaminergic denervation and gait in healthy adults. *Experimental Brain Research*, 185(3), 391-398.
- Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, 17(4), 297-338.
- Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, 84(6), 401-410.
- Hall, H., Sedvall, G., Magnusson, O., Kopp, J., Halldin, C., & Farde, L. (1994). Distribution of D1- and D2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacology*, 11(4), 245-256.
- Heaton, R. K. (1981). *A manual for the Wisconsin card sorting test*. Western Psychological Services.
- Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behavior (UCSD CHIP Report No. 99).
- Paolo, A. M., Tröster, A. I., Blackwell, K. T., Koller, W. C., & Axelrod, B. N. (1996). Utility of a Wisconsin Card Sorting Test short form in persons with Alzheimer's and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, 18(6), 892-897.
- Schmidt, R. A. (1976). The schema as a solution to some persistent problems in motor learning theory. In G. E. Stelmach (Ed.), *Motor Control: Issues and Trends* (pp. 41-65). Academic Press: New York.

³Note also that the dopamine system is vulnerable to change in normal ageing (Cham et al., 2008).