A Biologically Plausible Model of the Wisconsin Card Sorting Test

Sean R. Aubin

April 1, 2015

1 Introduction

The Wisconsin Card Sorting Test (WCST) is a psychological test used to evaluate executive function. It has been used to diagnose a wide variety of mental dysfunctions including Attention Deficit Hyperactivity Disorder [8], Substance Dependence [3], Autism [13], Parkinson's disease, Huntington's disease, and Schizophrenia (with and without tardive dyskenisia) [1]. The task is simple. The patient must match one single varying trial card to one of a set of four constant matchable cards as seen below in figure 1. The only feedback given is whether the match is correct or incorrect. The rules are simple, consisting of colour, shape or number matching. After a certain number of correct trials the rule shifts. The patient is warned about the potential for rule shifting when presented the task. Common measure for differentiating performance on the task are the number of perservative errors and categories achieved. Perservative errors refer to how much an old rule is applied after a new rule is required. Categories achieved refers to how many categories the patient experiences, since each category/rule requires ten cards to be matched successfully before a new one is presented.

In this report, a biologically plausible model of the WCST will be presented. All components except for routing and control have been implemented in spiking neurons using the Neural Engineering Framework (NEF) [5]. These components manipulated symbols as enabled by the Semantic Pointer Architecture (SPA) [6] [12]. The feasibility of the control and routing segments has been shown previously.

Although the individual components of the network work as intended and their parameters explored in a meaningful manner, the overall network does not yet match psychological data, due to time constraints on development.

2 System Description

Here I will briefly describe the neuroanatomical mappings as described by the most recent literature, before describing the methods used to model the aforementioned components.

Being a test for executive function, it is not surprising that components of the frontal lobe and the basal ganglia are known to be the two major contributors of processing for the WCST. In particular. it has been determined that

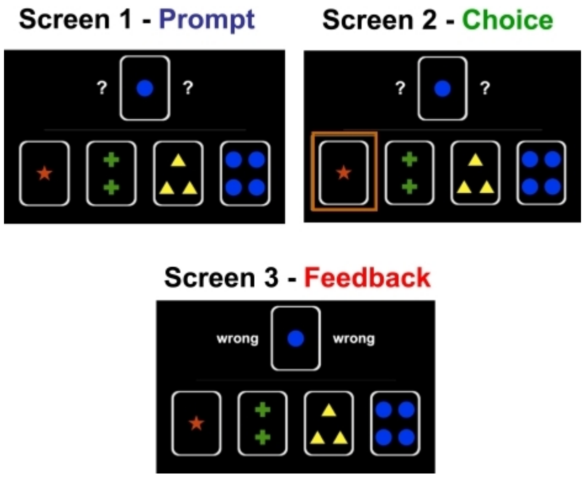


Figure 1: Example of task environment and action flow [7].

the prefrontal cortex, the caudate nucleus and the mediodorsal thalamus are activated during set shifting after receiving negative feedback [9]. Consequently, the control and gating parts of the model will be mapped onto the thalamus and caudate nucleus, while all transforms and comparisons will be considered to be cortical functions.

2.1 Representation

SPA was used for symbolically representing the cards as vectors. The vocabulary used consisted of the four colours, numbers and shapes that vary across cards, as well as binding vector for each attribute. Each item of the vocabulary was assigned a random unit vector of d dimensions. These vectors were combined to create a card by binding with circular convolution (represented here by $\circledast$) each attribute of the card to it's corresponding category. Thus, $NUMBER \circledast ONE + COLOUR \circledast RED$ would be an example card expression.

The vectors bound to $NUMBER$ are actually *unitary* vectors, as this enabled the realistic representation of numbers as seen in Spaun [6]. Where $ONE \circledast ONE = TWO$ and $ONE \circledast TWO = THREE$ and so on.

2.2 Learning Transformation

Previous computation models of the WCST have focused on probabilistically loading pre-determined rules either in a purely ACT-R-like symbolic manner [3] or a connectionist mutual inhibition model [1]. Instead of those previous approaches, but with a similar goal as Rougier et al. [11], this model attempts instead to learn the rules dynamically while being presented the stimulus. Although pursuing the same goal, this model takes a more biologically plausible route, given that the model of Rougier et al. is based on Leabra. Leabra is known to be biologically implausible [4], whereas each component of this model has been previously implemented in neurons.

The transform learning mechanism is the same that has been previously applied to the Raven's Progressive Matrices (RPM) task sequence solver [10].

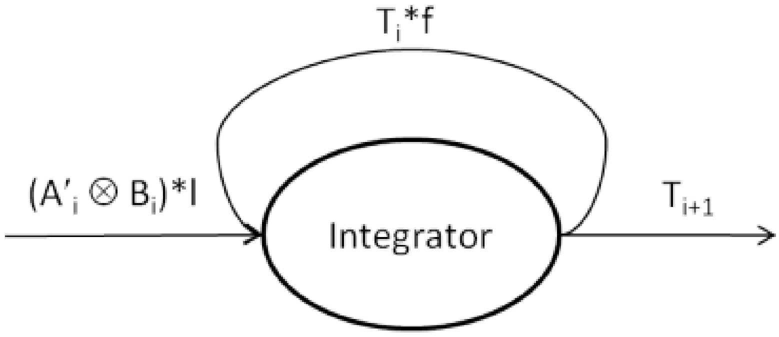


Figure 2: A modified integrator for taking the running average of the transforms [10].

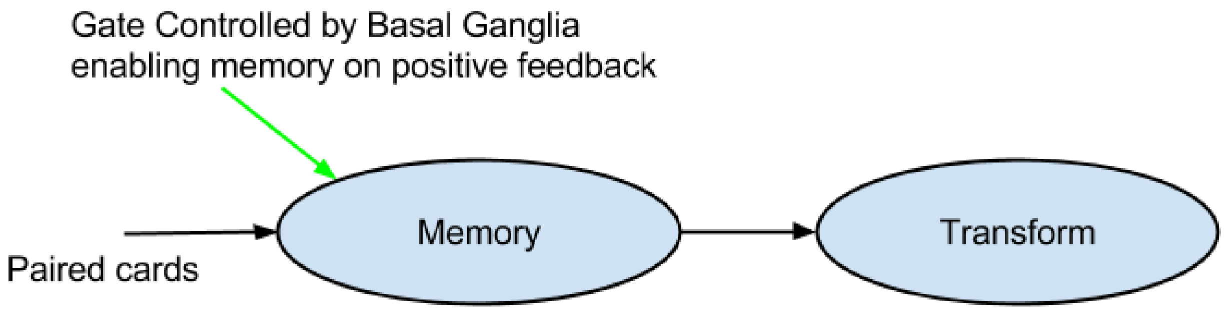


Figure 3: The basic transform network.

Essentially, given two cards that we've matched successfully (that we've received positive feedback on) we want to find the relation between them such that $card_1 \circledast T = card_2$ giving $T = card_1' \circledast card_2$ where the prime indicates an inverse. Each pair of matched cards can contribute to the transform by a averaging over the transforms of each pair.

Note however that for the application of WCST there is no implied order of the matched cards as there is for finding transforms between the cells of the RPM. To compensate for this, instead of calculating the transform as $T = card_1' \circledast card_2$, it is calculated as $T = card_1' \circledast card_2 + card_2' \circledast card_1$.

Given that the WCST only gives intermittent feedback, we must be able to hold a value over a period of time and to be able to average the transform sequentially (as opposed to the previous question which assumed we had access to all the transforms at once). Calculating the transform sequentially with an integrator as seen in figure 2. Keeping a value over time means coupling an input gated memory network (provided by Nengo [2]) to the input of the integrator which is gated by the basal ganglia (discussed in the next session) to only save inputs on positive feedback.

The WSCT requires multiple transforms to be maintained in memory so

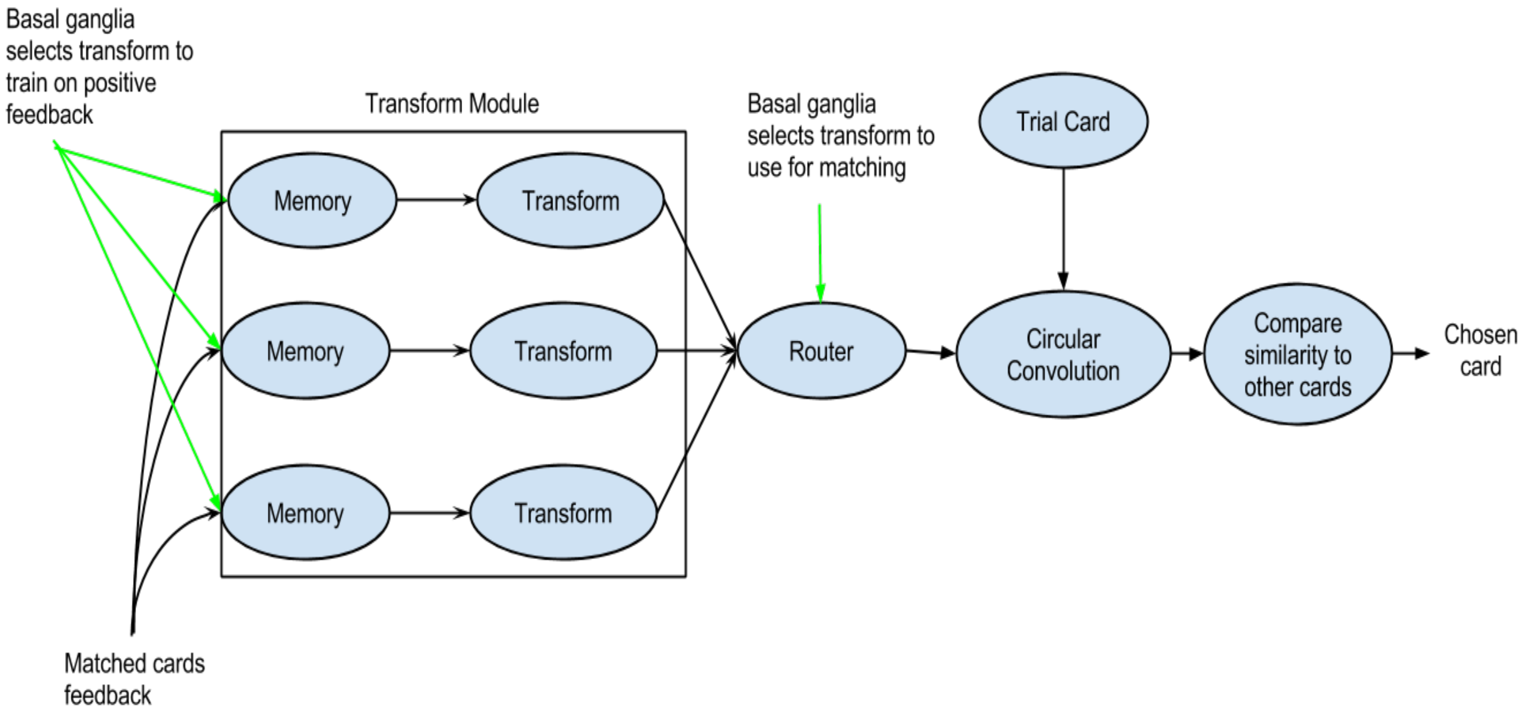


Figure 4: The learning module of the WCST network.

that rules can be revisited. To accomplish this, three transforms networks are then combined into one module gated by the basal ganglia as shown in figure 4.[1]

To determine which pile to sort the trial card into, the transform (chosen by the basal ganglia) was binded to the trial card using circular convolution. After being passed through a clean-up memory with the valid card formats, the binded term is then compared to the matchable cards using the dot product. The matchable card with the highest dot product was chosen and evaluated by the card simulation.

2.3 Control

For controlling and routing the information through the model two networks where artificially created. First, an artificial basal ganglia was created to monitor the accumulating reward and to control the gates. The basal ganglia was programmed such that after a certain reward threshold was reached, any subsequent negative reward would be greatly amplified. After a negative reward was received, the reward would drop past the threshold causing the basal ganglia to switch transform modules. Adding these two nodes resulted in the model in figure 5.

[1]The number of learning modules has been artificially limited to three, since it was known beforehand that there were only three rules to learn. More realistic rule management approaches will be discussed in the conclusions of this report.

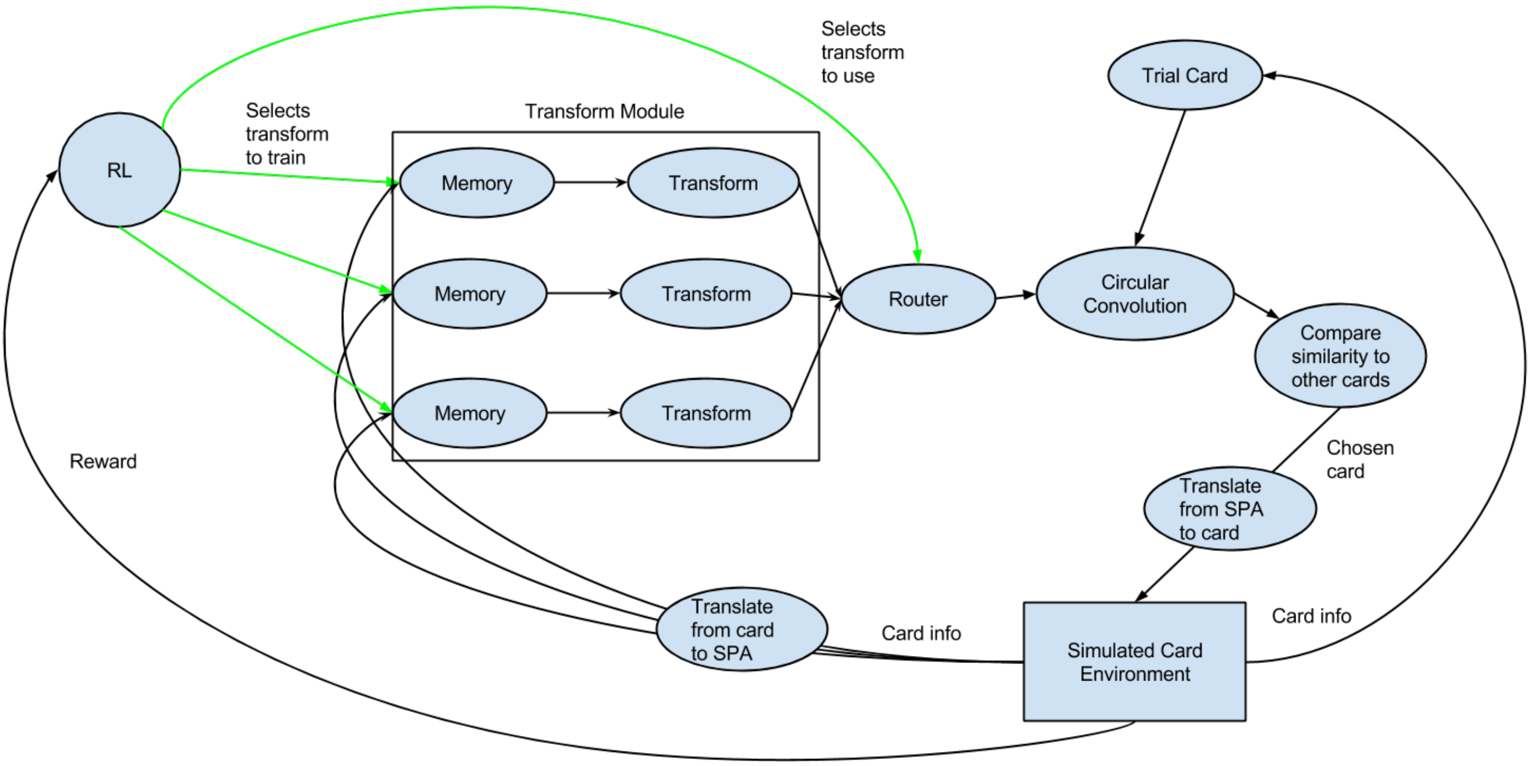
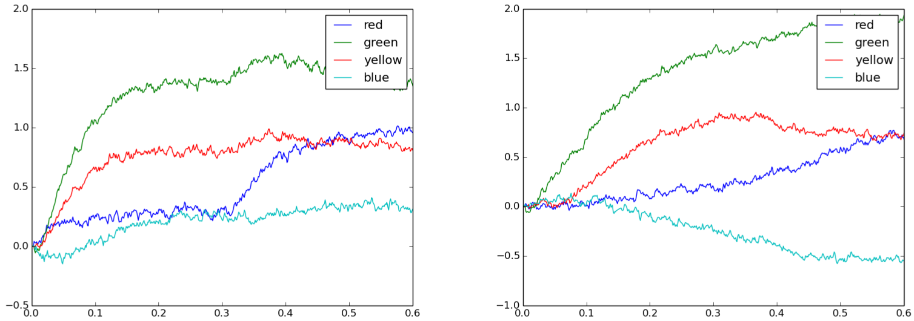


Figure 5: The complete WCST task network. Green arrows indicate control signals.



(a) No memory attached (b) Witch memory attached

Figure 6: Rule acquisition of "COLOUR*GREEN" given training data at $t = 0.0$ and $t = 0.3$.

3 Design Specification

For modeling parts of the cortex, the same neurotransmitters (resulting in the same post-synaptic time-constants), membrane time constants, firing rates and refractory periods as in Jan Gosmann's n-back task (in press) were used. Both of our tasks make extensive use of working memory and the prefrontal cortex. The only exception to this is the input to the transform integrator which uses a synaptic time constant of 7ms. This was taken from Dr. Dan Rasmussen's RPM code and modifying it was not investigated.

4 Implementation

The model was implemented using Nengo 2.0 [2] as described in the previous sections. 128 dimensions were chosen for representing the SPA vectors. For neurons, all networks were given $128\ dimensions \times 50\ neurons\ per\ dimension = 6400\ neurons$, with the exception of the circular convolution network which used 200 neurons per dimension.

As shown in figure 6, the transformation network acquires rules with ease, with or without the attached memory module.

Despite the transformation network satisfying it's desired function. the model (simulated with direct neurons) performed poorly, rarely acquiring rules and not switching rules when needed. It acquired only two of six possible categories and generally did not accumulate much reward as seen in figure 7. This is most likely due to a bug in the control system. Further analysis is needed.

Given this poor performance and a lack of time for analysis of the failures, the parameters affecting the ability to learn a transform were investigated instead. Dr. Rasmussen has already investigate the effects of increasing the dimensions of the vectors representing the cards and the number of neurons [10], so I will instead focus on adjusting the forgetting rate and the differential gain of the memory input.

The comparison setup was to input a green card into the trial card input

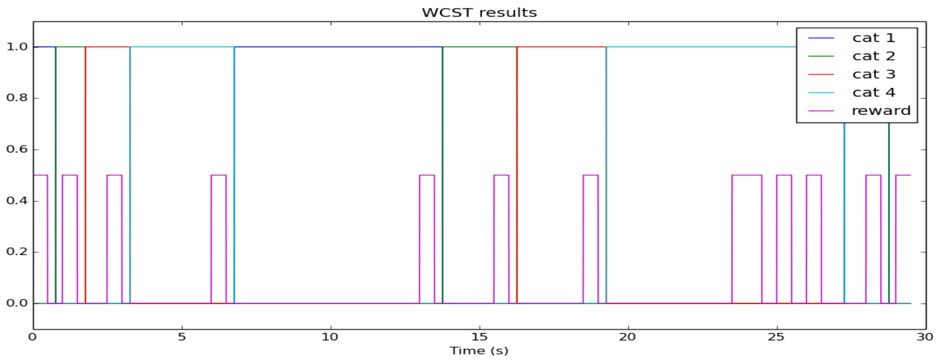
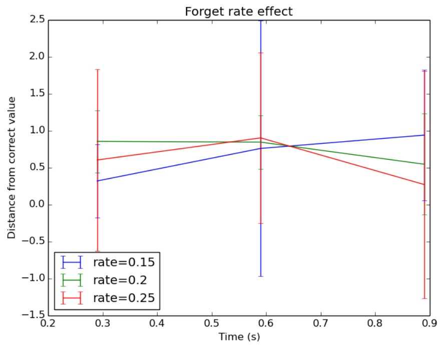
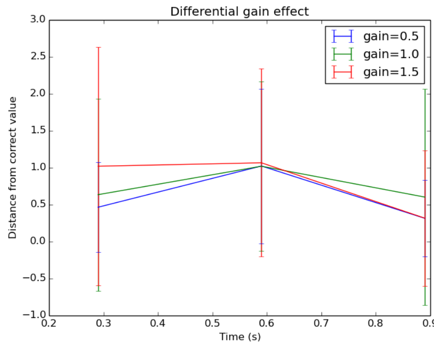


Figure 7: Performance of a run of the complete WCST network. Reward is rarely accumulated and categories almost never preserved.



(a) Effect of modifying the forget rate of the transform network



(b) Effect of modifying the differential gain of the memory network

Figure 8: Exploration of differential gain and forget rate parameters in relation to rule acquisition.

of the network, while three training values intended to promote learning of the colour rule were inserted into the memory input of the transform. Further details of this testing setup can be found in **data_collection.py** in the source code submitted with this project. The method of comparison was to measure how far the green coloured matchable card (the intended rule to learn) was from the top choice. These were the negative values in the data. If green was the top choice, the measure was the difference between the green matchable card and the second top choice. The results of this comparison are shown in figure 8 for running with 5 different random seeds.

Both the differential gain and the forgetting rate modifications have such large confidence intervals that perhaps the methodology for learning transforms should be re-evaluated. However, there does seem to be a general trend indicating that a differential gain of 1.0 is ideal for the memory and that a lower forget rate gives better results for the transform.

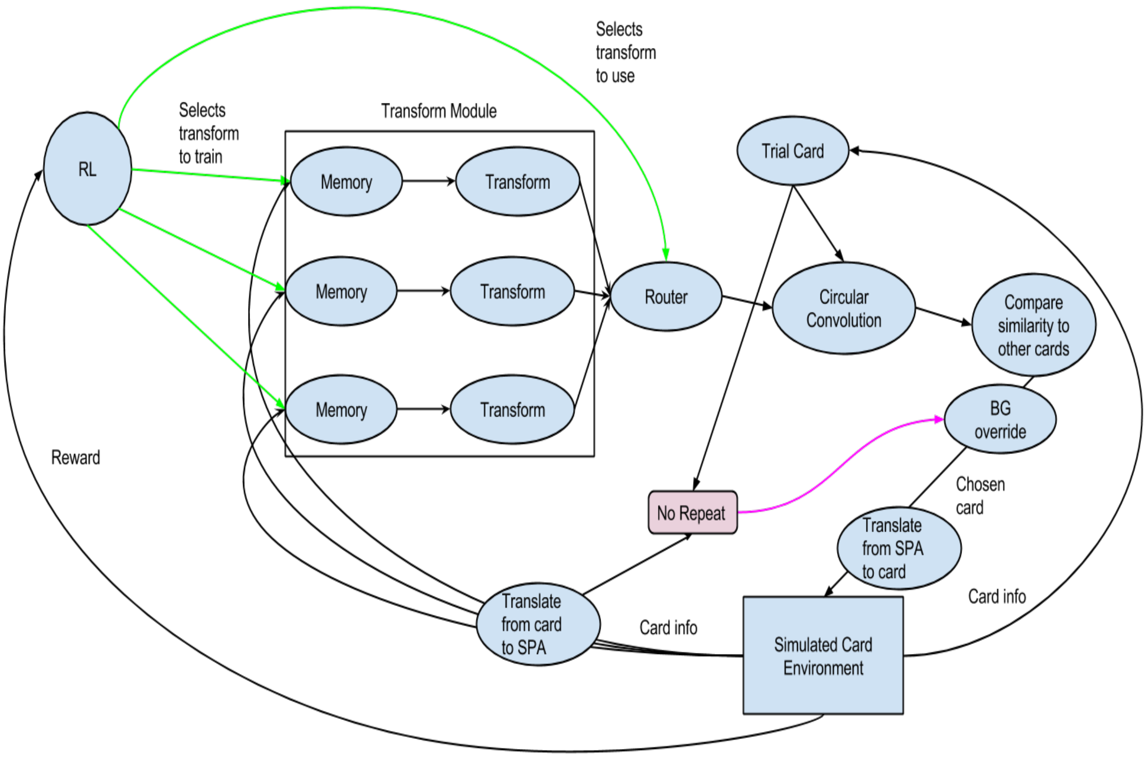


Figure 9: The complete WCST task network with added learning from negative feedback.

5 Conclusion

A model for the WCST was presented, showing the ability to acquire rules, but not to match psychological data.

Simple additions to the network that may allow more psychologically plausible performance could be the addition of a network which prevents the network from re-using a rule once negative feedback has been given.

In other words, once negative feedback has been given, the invalid transform could be saved and could be used to inhibit the choice indicated by it's convolution with the trial card as shown in figure 9.

A more complex and ambitious addition to the network could be a transform management system. The interaction between feedback in this model was simple, but limited. Only a limited number of transformation could be learned and the model could not be made aware if a transformation being learned was similar to one already saved. I'm not sure how to overcome these challenges, but I imagine almost requiring method of saving transforms almost equivalent to a serial memory [6] or perhaps more accurately to the system used to simulate the n-back task by Jan Gosmann (in press). Further research is needed.

References

[1] Andrew Amos. A computational model of information processing in the frontal cortex and basal ganglia. *Journal of Cognitive Neuroscience*, 12(3):505–519, 2000.

[2] Trevor Bekolay, James Bergstra, Eric Hunsberger, Travis DeWolf, Terrence C Stewart, Daniel Rasmussen, Xuan Choo, Aaron Russell Voelker, and Chris Eliasmith. Nengo: A python tool for building large-scale functional brain models. *Frontiers in Neuroinformatics*, 7(48), 2014.

[3] Anthony J Bishara, John K Kruschke, Julie C Stout, Antoine Bechara, David P McCabe, and Jerome R Busemeyer. Sequential learning models for the wisconsin card sort task: Assessing processes in substance dependent individuals. *Journal of mathematical psychology*, 54(1):5–13, 2010.

[4] Chris Eliasmith. *How to build a brain: A neural architecture for biological cognition*. Oxford University Press, New York, NY, 2013.

[5] Chris Eliasmith and Charles H Anderson. *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT Press, 2004.

[6] Chris Eliasmith, Terrence C Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *science*, 338(6111):1202–1205, 2012.

[7] Santosh A Helekar, Jae C Shin, Brandi J Mattson, Krystle Bartley, Milena Stosic, Toni Saldana-King, P Read Montague, and George J Hutton. Functional brain network changes associated with maintenance of cognitive function in multiple sclerosis. *Frontiers in human neuroscience*, 4, 2010.

[8] Vivienne Lawrence, Stephen Houghton, Graham Douglas, Kevin Durkin, Ken Whiting, and Rosemary Tannock. Executive function and adhd: A comparison of children's performance during neuropsychological testing and real-world activities. *Journal of Attention Disorders*, 7(3):137–149, 2004.

[9] Oury Monchi, Michael Petrides, Valentina Petre, Keith Worsley, and Alain Dagher. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 21(19):7733–7741, 2001.

[10] Daniel Rasmussen. A neural modelling approach to investigating general intelligence. Masters thesis, University of Waterloo, Waterloo, ON, 2010.

[11] Nicolas P Rougier, David C Noelle, Todd S Braver, Jonathan D Cohen, and Randall C O'Reilly. Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20):7338–7343, 2005.

[12] Terrence C Stewart, Xuan Choo, and Chris Eliasmith. Symbolic reasoning in spiking neurons: A model of the cortex/basal ganglia/thalamus loop. In *Proceeding of the 32nd Annual Meeting of the Cognitive Science Society*, pages 1100–1105. Cognitive Science Society Austin, TX, 2010.

[13] Lien Van Eylen, Bart Boets, Jean Steyaert, Kris Evers, Johan Wagemans, and Ilse Noens. Cognitive flexibility in autism spectrum disorder: Explaining the inconsistencies? *Research in Autism Spectrum Disorders*, 5(4):1390–1401, 2011.