A COMPUTATIONAL MODEL OF THE INTERACTION OF NEUROBIOLOGICAL CIRCUITS FOR CATEGORY LEARNING

by

Li Xin Lim

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Psychological Sciences

West Lafayette, Indiana

December 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. Sebastien Hélie, Chair

Department of Psychological Sciences

Dr. Anne B. Sereno

Department of Psychological Sciences

Dr. Richard Schweickert

Department of Psychological Sciences

Approved by:

Dr. Margo J. Monteith

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

COVIS	Competition between Verbal and Implicit Systems
HT	Hypothesis-testing System
P	Procedural Learning System
II	Information-integration
RB	Rule-Based
fMRI	Functional magnetic resonance imaging
STN	Subthalamic Nucleus
GPi	The internal segment of the Globus Pallidus
preSMA	Pre-supplementary Motor Area
PMd	Dorsal Premotor Cortex
C	Capacitance
v	Membrane potential
v_r	Resting membrane potential
v_t	Instantaneous spiking threshold
I	Input to Izhikevich neurons
u	Membrane recovery variable
a	Recovery time constant
k	Rheobase resistance to Izhikevich neurons
b	Input resistance to Izhikevich neurons
c	Voltage reset value
d	The total difference between outward and inward currents during a spike
K^+	Potassium ions
Na^+	Sodium ions
I_{preSMA}	Input to preSMA
I_{STN}	Input to STN
I_{PMd}	Input to PMd
α_L	Lateral inhibition from other PMd units
α_{preSMA}	α-function of preSMA

α_{STN}	α-function of STN
α_{PMd}	α-function of PMd units
t_{α}	Time since the cell voltage reached a peak for α-function
Λ	Constant that determines the duration of signal propagation in the synapse
t	Time
v_i	The expected value of the chosen option
r	Experienced outcome
ACC_{HT}	Accuracy probability of HT
$ACC_{HT_{II}}$	Accuracy probability of HT in II task
$ACC_{HT_{RB}}$	Accuracy probability of HT in RB task
ACC_{P}	Accuracy probability of P
$ACC_{P_{II}}$	Accuracy probability of P in II task
$ACC_{P_{RB}}$	Accuracy probability of P in RB task
F_S	Accuracy feedback of each learning system
F_{HT}	Accuracy feedback of HT
F_P	Accuracy feedback of P
$HT_{response}$	Response of HT
$HT_{confidence}$	Confidence of HT
$P_{response}$	Response of P
$P_{confidence}$	Confidence of P
$P_{confidence_{max}}$	Maximum value of $P_{confidence}$
A_{HT}	The activation of the learning systems HT with the presence of cued RB task
A_P	The activation of the learning systems P with the presence of the cued II task
$R_{HT_{Task}}$	Working memory proactive interference to the accuracy probability of HT of the current task
$R_{HT_{RB}}$	Proactive interference affecting the accuracy probability of HT in RB task
$R_{HT_{II}}$	Proactive interference affecting the accuracy probability of HT in II task
R_S	Proactive interference perceptual cue's decay effects on preSMA input
S	The association of task perceptual cues to the strategy or learning system

S_{max}	The maximum value S can carry
S_{HT}	The association of RB perceptual cues to the RB strategy or HT
S_P	The association of II perceptual cues to the procedural strategy or P
vc	Perceptual cue
vc_{HT}	Perceptual cue for RB strategy
vc_P	Perceptual cue for procedural strategy
γ	Learning rate
γ_{vc}	Learning rate for perceptual cue
γ_S	The learning rate for the association of task perceptual cues to the strategy or learning system
γ_{HT}	Learning rate for HT
γ_P	Learning rate of $P_{confidence}$
$\gamma_{P_{dist}}$	Distribution for participants in a group as a factor of deriving γ_P
$\gamma_{P_{dist}mean}$	Mean for $\gamma_{P_{dist}}$ to determine the distribution of γ_P for different participants in a group
D	The exponential decay rate for proactive interference
D_{dist}	The distribution for participants in a group as a factor of deriving D
$D_{dist_{mean}}$	The mean for D_{dist} to determine the distribution of D for different participants
N	Number of trials since the last task switch
PCT	Perceptual categorization task
DBM	Decision Bound Models
GLC	General linear classifier
BIC	Bayes information criterion

ABSTRACT

The goal of this proposal is to design a neurobiologically-based model that describes the switching mechanism in category learning based on existing category learning theory and model. COVIS is a neurobiologically-detailed theory of multiple systems in category learning. COVIS postulates two systems that compete throughout learning—a frontal-based declarative hypothesis-testing system that uses logical reasoning and depends on working memory and executive attention, and a basal ganglia-mediated system that uses procedural learning. However, no biological mechanism has been proposed to account for the interaction between the COVIS systems. We propose a model that employs a neurobiological-based circuit that describes the interaction and switching between the hypothesis-testing system and the procedural learning systems in COVIS. With the hypothesis-testing system and procedural learning system modeled as black boxes, the model focuses on the network that facilitates switching. In COVIS, both learning systems generate a response signal in each trial based on the stimuli given. Our model incorporates the Izhikevich firing model that represents the activity of the neuronal cells from the hyperdirect pathway of the cortico-basal ganglia network. The hyperdirect pathway acts as a gate for the response signal of the procedural learning system to reach the premotor units for action selection. We propose that the procedural learning system's response is inhibited from approaching the premotor units when the hypothesis-testing system is in control of the response. However, if rule-based strategies fail, inhibition to the procedural system's response is reduced. The reduction in inhibition results in the acceptance of responses from both learning systems in the premotor units. To validate the proposed model, we fit the model to two groups of participants in a perceptual category-learning task. One group of participants used the optimal procedural strategy in the task and the other used a suboptimal rule-based strategy. The categorization task was an information-integration task, whereby participants had to switch away from rule-based strategies and learn to integrate the stimulus dimensions to be able to perform optimally. We were able to differentiate the switchers from the non-switchers by adjusting the parameters in the model. In addition, we fitted another task to the model in which participants from different age groups with or without Parkinson's disease were asked to switch between rule-based and procedural strategies on a trial-by-trial basis. We were able to match the learning curve, accuracy switch cost, and proportion of switchers of the different groups of participants.

INTRODUCTION

In the past decades, attention has been drawn to numerous psychological and biological systems to explain category learning. There is abundant evidence that suggests learning and memory in humans are mediated by multiple systems (Eichenbaum & Cohen, 2001; Schacter, Wagner, & Buckner, 2000; Squire, 2004). With the increased popularity of multiple learning systems, researchers have begun questioning how these systems interact (Ashby & Crossley, 2010; Poldrack & Packard, 2003; Schroeder, Wingard, & Packard, 2002). An example of a cognitive model with multiple learning systems is the Competition between Verbal and Implicit Systems (COVIS) model (Ashby et al.,1998). COVIS describes category learning with an explicit hypothesis-testing system and an implicit procedural-learning system. This thesis built on COVIS and sought to understand the interaction between two learning systems, namely the prefrontal cortex-based hypothesis-testing system (HT) and the striatal-based procedural learning system (P). The HT uses executive attention and working memory to learn through declarative reasoning, while the P learns through reinforcement learning. We were particularly interested in the neural mechanism behind the interaction between the two systems that facilitates the switch from one system to another in different category learning tasks.

Multiple Systems Theory in Category Learning

A vast amount of research had been focusing on a multiple-systems theory in category learning (Ashby et al., 1998; Ashby & Valentin, 2017; Erickson & Kruschke, 1998; Smith & Grossman, 2007). The theory was formed based on the idea that different aspects of the environment were learned through different categorization systems. The dual-process theory is an example of using multiple systems in category learning. The main feature of the theory was built on the idea that there are at least two systems in the brain (Kahneman, 2012). The information processing in humans has multiple levels and may be interactive, which is comprised at a minimum of an explicit declarative rule learning mode and a gradual implicit tuning that reflects intuitive learning (Kahneman, 2012; Sun, 2002).

The explicit system in the dual-process theory is unique and evolutionary recent to primates, including humans (Evans, 2003). The explicit system is executed in the central working memory

system, and this leads to a restricted capacity of processing and makes it slower when compared to the implicit system which uses fewer resources (Tsujii & Watanabe, 2009). According to COVIS, rule-based categorization is explicit. The system categorizes an element to a class by determining if it fits a membership rule that expresses the category (Smith & Grossman, 2007). The cognitive operations outlining the rule-based categorization start with selectively attending to each diagnostic feature of the test element. For each attended feature, the value that the element holds is matched and compared to the feature(s) specified in the rule. The final categorization is achieved by combining the outcomes of each feature evaluation. With that, rule-based categorization entails selective attention and temporary storage for rules testing and outcomes (Smith, Patalano, & Jonides, 1998). One example is a conjunctive rule that allows categorization based on the criterion from two separable stimulus dimensions. For instance, for stimuli in the form of sine-wave gratings that vary across trials in rotation angle and bar width can be categorized with a rule saying "respond A if the bars are tilted with a greater angle and the bars are thick, and respond B otherwise".

The implicit system in the dual-process theory is often noted when there is a lack of conscious awareness of the memories' details or if memory has even been stored. This makes humans unaware of the learning, such that they can either be unaware of the stimulus, how the stimulus is interpreted or the effect of the stimulus itself on related actions or judgments (Ashby & Casale, 2002; Bargh, 1994). There may be multiple implicit category learning systems, and two possibilities are a procedural learning-based system and a perceptual representation system. The perceptual representation system relies on perceptual learning in the visual cortex. On the other hand, the procedural learning system refers to the acquisition of new knowledge of motor skills through the repeated performance of a task and is mediated, in part, by the basal ganglia (Cantwell et al., 2017). The optimal strategy for this learning is difficult or impossible to be described verbally (Ashby & Valentin, 2017).

System Switching in Category Learning

It has been gaining recognition that a situation that involves only one type of learning is rare (Slusarz & Sun, 2001). While it is possible to manipulate conditions to bias learning towards one or the other type of learning system, in most situations, both types of learning are involved to a certain degree, with varying amounts of contributions from each system (Slusarz & Sun, 2001).

Separable and distinct neural mechanisms have been found to facilitate explicit and implicit learning, with the hippocampus and temporal-parietal cortex mediating explicit learning and representation of declarative knowledge (Cohen et al., 1985; Eichenbaum, 1999). The implicit learning circuit is presumably mediated by a cortical-subcortical circuit (Heindel et al., 1989; Knowlton, 2002; P. J. Reber & Squire, 1994). The medial temporal lobe and the basal ganglia are differentially involved in the explicit and implicit system, respectively (Yang & Li, 2012). The learner's reliance on the medial temporal lobe reduces swiftly as learning progresses (Yang & Li, 2012).

In general, it is difficult for individuals to switch between categorization learning systems, even for healthy young adults (Helie & Fansher, 2018). Erickson (2008) reported that only 37% of the participants were able to switch categorization learning systems on a trial-by-trial basis. Also, only around 40% of undergraduate students were identified as switchers with continuous-dimension stimuli by Crossley, et al. (2018). Two sessions of training and preparation time for system switching before presenting the categorization stimulus were required for a higher proportion of switchers (Hélie, 2017).

Computational models can be used to further understand the multiple-system theory in category learning and switching between categorization learning systems. COVIS is an existing computational model that employs the dual-process theory (Ashby et al., 1998; Hélie, Paul, & Ashby, 2012). COVIS works by implementing category learning that includes an explicit hypothesis-testing system and an implicit procedural-learning system. The explicit hypothesis-testing system learns through declarative memory, which is done by selecting and testing simple expressible hypotheses. The implicit procedural-learning system employs non-declarative memory whereby learning is mediated by reinforcement learning as the system gradually assigns motor responses to regions of perceptual space. On each trial, the model compares the confidence in both the hypothesis-testing system and procedural-learning systems and produce one response, either from the hypothesis-testing system or the procedural-learning systems (Hélie et al., 2012).

Two types of category structures have been substantially studied with COVIS, namely rule-based (RB) and information-integration (II) categories (Maddox & Ashby, 2004). In RB tasks, the task can be solved with verbalizable rules that maximize the accuracy as the optimal strategy. The RB category learning is shown to be regulated by an explicit hypothesis-testing system (Filoteo et al., 2005; Maddox et al., 2010). On the other hand, implicit procedural based learning

facilitates II learning. The procedural strategy developed in II learning is said to be the optimal strategy for II tasks. II tasks are associated with tasks where the strategy that maximizes accuracy is difficult to describe verbally. Performance in II tasks is maximized when information from two or more stimulus dimensions are integrated at some pre-decisional stage, whereby learning involves the incremental acquisition of stimulus-response associations (Ashby et al., 1998; Filoteo et al., 2005; Maddox et al., 2010). Often, RB strategies can be applied in II tasks, but it will bring suboptimal performance. COVIS assumes that participants start by guessing or formulating simple rules through hypothesis-testing. Participants might abandon the rule-based strategies if these rules have failed, and they may adopt an alternative, more intuitive, and less verbal method of categorization (Ashby et al., 1998).

Each system is fit to learn particular categorization tasks (Paul & Ashby, 2013). However, the systems' interaction remains unknown. Our aim in this study is to seek to understand how switching from one system to another occurs and to propose a neural mechanism that facilitates such switching capabilities. In COVIS, both explicit systems and implicit systems can learn through feedback. Switching from a system to another is based on confidence and trust in each system. The confidence in each system varies on each trial according to the success of individual responses generated by each system.

Independent Learning in Systems

From Paul and Ashby (2013), there are at least three ways of describing single response selection when two learning systems are simultaneously active. One possibility is that the outputs of the two systems are mixed for the generation of the final output. The categorization response of such a model shows a composition of declarative and procedural processes, with different weights put on the declarative output and procedural output. In the RB task, a declarative output might be weighed higher than the procedural output, but in an II task, the procedural output may be weighed higher than the declarative output. Another possibility of response selection involves soft switching, where only one system controls the response of each trial, and that control alternates between the two systems on a trial-by-trial basis. As opposed to the soft switch, a third possibility is the use of a hard switch, where one system is used solely and only a single switch is involved to shift control to the other system throughout the task, as the task demands it. COVIS predicts learning starts with simple explicit rules. The original version of COVIS assumes trial-by-trial

switching between explicit and procedural systems in people as learning progresses (Ashby et al., 1998). For the same stimulus, the explicit system and procedural system might come up with different (or same) responses, which is dependent on the system's learning. The two systems compete for behavioral control, leaving the individual settling with one final response in general (M. J. Crossley & Ashby, 2015).

Ashby & Crossley (2010) suggested that procedural system's response is inhibited during the execution of explicit strategies. Ashby & Crossley (2010) performed a category learning experiment that used sine-wave gratings as stimuli and hybrid category structures where optimal accuracy was achievable using a simple rule in conditions with steep bar orientation and procedural strategy in cases when the orientation was moderate. It was reported that trial-by-trial switching was observed in 3 out of 53 participants in different experiments. Almost all participants used a suboptimal strategy, following either a simple suboptimal rule-based strategy or suboptimal procedural strategy on all trials. The number of participants that used rule-based strategy outnumbered the number of participants that used the procedural strategy.

Furthermore, an antagonistic relationship between neural activation in medial temporal lobes and striatum during category learning was shown with increased striatal activation and decreased medial temporal lobe activation in category learning (Moody, Bookheimer, & Knowlton, 2004; Nomura et al., 2007; Poldrack et al., 1999). Also, numerous fMRI studies showed negative correlations between the activation in medial temporal lobe and striatum in skill learning (Dagher et al., 2001; Jenkins et al., 1994).

However, studies on active inhibition of behavioral output may not accurately reflect learning. A persistent striatal activation was reported during a declarative task in neuroimaging work by Foerde, Knowlton, and Poldrack (2006). Moreover, there is evidence from behavioral studies suggesting the learning of the procedural system even when the explicit system controls behavior. An example is a result from Crossley and Ashby (2015) where participants were trained with stimuli that required learning a unidimensional rule first, and the same stimuli were then used in an II categorization task. In the experiment, a manipulation was introduced to interfere with procedural learning but not RB learning during the RB training. The introduction of the manipulation reduced transfer performance in the II task, which suggests procedural learning during RB training where the explicit system is in control of the overall behavior.

Independent learning with multiple systems is also seen in sequence learning, as suggested by the interplay of distinct one-dimensional and multidimensional systems (Keele et al., 2003). In sequence learning, the two systems learn independently and compete to control behavior. For starters, the one-dimensional system is implicit and takes in the perceptually raw stimulus or response features (M. J. Crossley & Ashby, 2015). The one-dimensional system involves neurobiological systems from dorsal regions, which include the supplementary motor, primary motor, and parietal cortices. The multidimensional system, on the other hand, can be either implicit or explicit and receives perceptually abstract stimulus or response features. The multidimensional system involves ventral regions of the brain, which include prefrontal, lateral premotor, temporal, and occipital lobes.

In cases of simple motor-learning tasks, for instance, visuomotor adaptation learning is seen through the interaction of multiple systems (M. J. Crossley & Ashby, 2015). As an example, two systems or processes mediate learning in motor tasks, namely an implicit process that is regulated by the cerebellum, and a separate prefrontal-based learning process. The former implements a sensory prediction error signal, which results in the difference between expectation and received feedback. The latter learns through distinct error signals, which results in the difference between feedback and outcome goal (Taylor & Ivry, 2012). These two processes learn and operate independently, which reflects consistency with the concept of inhibition of procedural systems to access motor outputs by the explicit system. However, the process does not prevent learning within the procedural system.

System Switching Neurobiological Interactions

There have been neuroimaging and behavioral evidence that suggest antagonistic relationships between explicit and procedural memory (Ashby & Crossley, 2010; Poldrack & Packard, 2003; Schroeder et al., 2002). However, Foerde et al. (2006) observed a persistent striatal activation even during explicit tasks. Moreover, Crossley and Ashby (2015) suggested that the inhibition between explicit and procedural systems does not operate at the level of learning, but instead occurs at the level of expression. The procedural learning system is meditated by the sensorimotor basal ganglia and dorsolateral striatum, and the explicit system by the prefrontal cortex and medial temporal lobes (Ashby & Ennis, 2006; Fletcher et a., 1998; Mishkin, Malamut, & Bachevalier, 1984; Tulving & Markowitsch, 1998). There might exist many possible anatomical

circuits that could serve to mediate the interaction between the two learning systems. One of these many possibilities is through the hyperdirect pathway as suggested by Ashby & Crossley (2010).

To adapt to the changing environment, processes which annul irrelevant or unsuitable behaviors are important to aid in achieving goal-directed behavior (Kenner et al., 2010). Such processes may act like an "emergency brake" which gives nonselective inhibition to suppress all responses (Kenner et al., 2010). In our case, all responses from a learning system (i.e., the procedural learning system) are inhibited. Present studies showed evidence of the involvement of the hyperdirect pathway regions in response stopping and response switching situations (Kenner et al., 2010). The hyperdirect pathway served a role in response inhibition and response switching through nonselective inhibition (Kenner et al., 2010).

The hyperdirect pathway as shown in Figure 1 acts to either allow or inhibit signals from the striatum to the cortex. The hyperdirect pathway starts with direct excitatory glutamate projections from the frontal cortex to the subthalamic nucleus (STN). Through excitatory glutamate projections, the connection from STN is directed to the internal segment of the globus pallidus (GPi) (Joel & Weiner, 1997; Parent & Hazrati, 1995). This excitatory input offsets the inhibitory input from the striatum to the GPi, which reduces the influence from striatum to the cortex (through the thalamus).

By reducing the subthalamic activity, the hyperdirect pathway permits signals from the striatum to the cortex; by increasing the subthalamic activity, it prevents the influence. Thus, through the hyperdirect pathway, the influence from the procedural system to the motor output system in cortex can be prevented, without having a direct interaction with the striatum itself. Since the striatum is presumably the learning site of the procedural system, the procedural system could still learn while the declarative system dominates responding (M. J. Crossley & Ashby, 2015). In the hyperdirect pathway, the computational function of STN is to generate pauses for stopping and conflict processing. These pauses allow time for the accumulation of evidence for the 'right actions' (Aron, Herz, Brown, Forstmann, & Zaghloul, 2016).

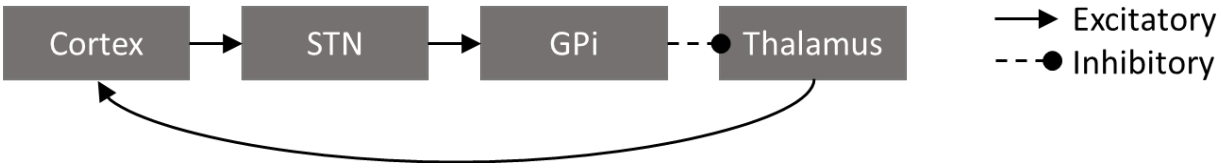

Figure 1. The hyperdirect pathway of the cortico-basal ganglia network.

THESIS STATEMENT

Previous works on the computational models of the multiple systems have focused on the learning systems. The interaction between systems, particularly the neurobiological basis of the switching interaction, is rarely discussed. We aim to build the model in hopes to see how different brain areas and cells are connected to give rise to the system switching mechanism. Combining experimental and theoretical information, our goal is to reconstruct the part of the brain structure and functions to be used to examine and propose new theories regarding learning system switching.

Our model is designed to show the neural circuit underlying the switching between the learning systems. The model shares the same concept with COVIS in terms of the dual-process learning system involved and describing the switch from the hypothesis-testing system to the procedural learning system. In addition, the model includes the Izhikevich (2006) neuron firing models and integrates the hyperdirect pathway into system-switching. Our goal is to show how switching occurs from decreasing the inhibition on procedural learning system's response, and how the independent responses generated from the two learning systems are transmitted through the model and result in a single outcome.

In this thesis, we describe the system-switching model at three levels: neuronal level, interneuron interaction with synaptic plasticity, and the model's circuit. The model was applied to simulate a perceptual categorization task, where parameters of the model were adjusted to fit the data of two groups of participants in the task: one group of participants that were able to switch between category learning systems to perform optimally in the task and another group of participants that stuck to a category learning system with suboptimal performance. In addition, we included another task to simulate, which is to fit published data to see system-switching changes in participant groups with different age conditions.

By analyzing the changes in parameters used to distinguish participant groups with different switching capabilities, we sought to understand the implication of the changes and differences in the adjustment of the parameters. By doing so, we might be able to sort out the sources that cause the incapability in switching, which could be revalidated later with clinical tests. This can help in understanding the reduction in system switching flexibility with aging and conditions such as Parkinson's disease.

METHODS

Model Overview - Computational Model of System Switching

The model simulates the switching mechanism between the hypothesis-testing system (HT) and the procedural learning system (P) of COVIS in a category-learning task. The model involves neurons from three areas of the brain: the pre-supplementary motor area (preSMA), the subthalamic nucleus (STN), and the dorsal premotor cortex (PMd), which were modeled with Izhikevich firing models, and three plastic synapses.

The model focused on the simulation of category learning, wherein each trial of the simulation, the model had to assign a given 'stimulus' to one of the available groups. For instance, the stimuli may be a circular sine-wave gratings, as shown in Figure 2, which varies from trial to trial in terms of orientation and bar width. For this example, to categorize the stimulus, the model had to read the value of orientation or bar width (or both). Both the HT and P generated their response, and the model 'selected' which system to follow for each trial as the responses from HT and P compete at the PMd units level.

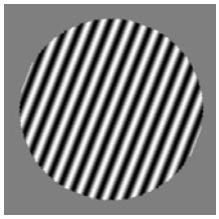

Figure 2. An example of circular sine-wave gratings that were shown to participants in a category learning task. The sine wave gratings differed in terms of orientation and bar width from trial-to-trial in a given task, where the orientation means the rotation angle and the bar width is the thickness of the black bars.

On each trial, both HT and P gave their own categorical decision as the output response to the stimulus. The systems responded independently to one stimulus on each trial. The responses from HT and P were fed into PMd so that the PMd units that reflect the actions resulting from the categorical decisions were activated accordingly. For instance, in a two-choice categorization task that requires sorting stimuli into category 'A' or 'B', if HT chooses 'A' as the categorical decision, the PMd unit that signals the pressing of the button 'A' for category 'A' is activated.

Initially, the PMd units should only receive input from the HT, according to the bias in COVIS that participants will learn with an RB strategy in the beginning. Thus, responses from P were inhibited from reaching the PMd units. Switching was depicted as the reduction in inhibition of the response of P to the PMd units, until a point where the activity induced in PMd units by the response of P is higher than that of HT. The neuronal firing models were incorporated to show the process of inhibition (or reduction in inhibition) of the response of P. As an output, the model gave the categorical action, for instance, either categorizing the stimulus as category A or B. The model then learned through the feedback for accuracy of choice. Figure 3 shows the flow of the circuit of the model.

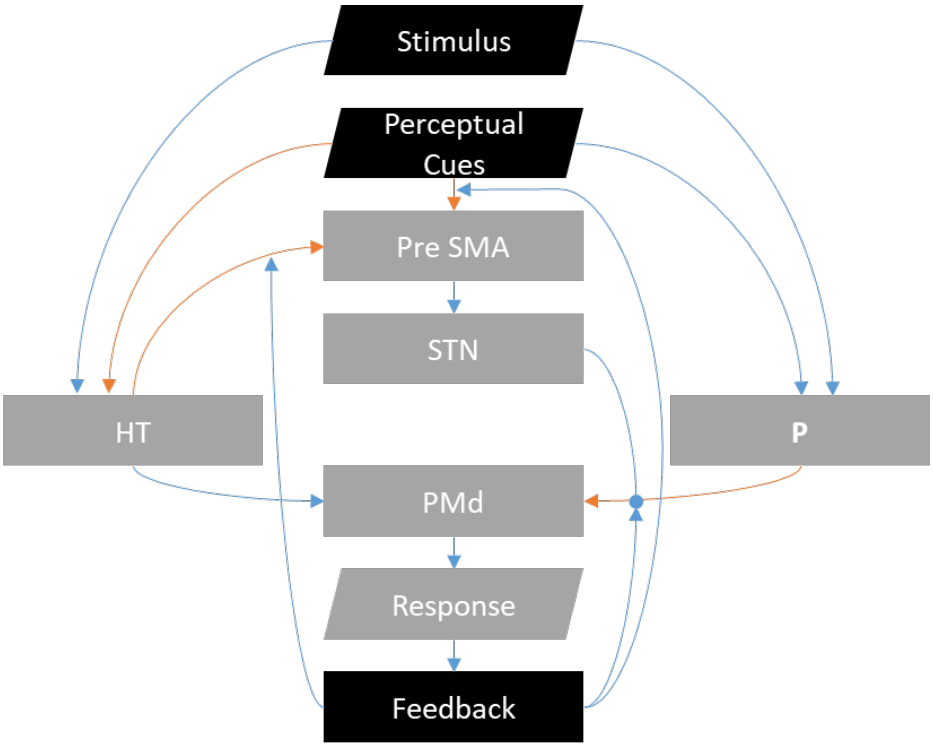

Figure 3. Block diagram representing the model for system switching. Orange lines indicate a plastic connection, while blue lines indicate a fixed connection. Arrow ends indicate activation while dot ends indicate inhibition. Grey boxes indicate processes in the agent while black boxes indicate processes outside the agent.

Brain Areas in the Model

Ashby and Crossley (2010) suggested the control of learning system interactions by the frontal cortex and the STN through the hyperdirect pathway in the basal ganglia. The hyperdirect pathway starts with the excitatory glutamate projections from the frontal cortex through the preSMA to the STN (Paul & Ashby, 2013). preSMA plays a role in facilitating the switch between learning systems' responding by inhibiting competing motor plans when response conflict occurs (Hikosaka & Isoda, 2010; Nachev, Wydell, O'Neill, Husain, & Kennard, 2007; Paul & Ashby, 2013). preSMA sends the switch-related signals to STN, leading STN to suppress ongoing behavior that is no longer relevant, to promote the execution of the new behavior (Hikosaka & Isoda, 2010).

In procedural learning, the striatum should first inhibit the globus pallidus so that the basal-ganglia-recipient thalamus can be disinhibited to initiate response goals (Ashby, Ennis, & Spiering, 2007). To suppress the response in procedural learning, STN increases activation in the internal segment of globus pallidus (GPi), strengthening the inhibition of thalamus. COVIS assumes a one-to-one connection from the GPi to the thalamus (Ashby, Paul, & Maddox, 2011). In our model, the output from the thalamus is considered as the response of P. Inhibition of thalamus' output is modeled as the inhibition of P's output with increase STN activity. Responses from the learning systems are then directed to the PMd accordingly to allow planning and preparation for appropriate movement accordingly.

Modeling of Computational Neuronal Cells

The nervous system can be seen as a multilevel organization that segregates in terms of spatial scales, which range from the molecular level to the whole nervous system. Neurons serve an important role in the nervous system and are a type of cells that are specialized to process information and signal. The neurons generate electrical potentials to transmit information to other cells to which they are connected. Information processing capabilities of the neurons are facilitated through the subcellular level mechanism, through cascades of biochemical reactions on a molecular level. As we developed such brain models consisting of spiking neurons, we must meet the requirement of having computationally simple neurons that are capable of generating firing patterns exhibited by actual biological neurons (Izhikevich, 2003). Such properties of the neurons

can be modeled by a system of coupled nonlinear ordinary differential equations, for example, the Hodgkin-Huxley model (Hodgkin & Huxley, 1952). It employs a dynamical system theory with varying neuronal parameters to shape the neuronal activities. For instance, the Hodgkin-Huxley equation can be turned into repetitive activity by maintaining depolarizing current at the neuronal membrane (Schierwagen, 2009). By reducing the maximal K^+ conductance or by moving the Nernst potential for the K^+ in the depolarizing direction, one can obtain large amplitude periodic solutions (Schierwagen, 2009).

However, the intention of the model was not quantitatively fitting neural activation but to seek for a qualitative agreement. The Izhikevich model proves to be computationally efficient as compared to the Hodgkin-Huxley equations in producing rich spiking and bursting dynamics as exhibited by neurons when simulating multiple neurons in real-time while maintaining a certain level of biological plausibility. Each neuron in the Izhikevich model is implemented as follows:

$$\begin{gathered} C\dot{v} = k(v - v_r)(v - v_t) - u + I + noise \\ \dot{u} = a[b(v - v_r) - u] \\ if\ v \geq v_{peak}, then\ \ v \leftarrow c, u \leftarrow u + d \end{gathered} \tag{1}$$

where C represents the membrane capacitance, v is the membrane potential, v_r is the resting membrane potential, v_t is the instantaneous spiking threshold, I is the input, u is the abstract term that describes membrane recovery variable for Na^+ and K^+ ion channel gatings, a is the recovery time constant, k and b represent the rheobase and input conductance respectively, c is the voltage reset value, d is the total difference between outward currents and inward currents during a spike, and noise follows a normal distribution of N(0,1).

The Izhikevich model can reproduce spiking and bursting behavior of known types of cortical neurons by varying the parameters in Equation (1). The model follows the bifurcation theory and normal form reduction (Ermentrout & Kopel, 1986; Hoppensteadt & Izhikevich, 1997). Eq. 1 generalizes the quadratic integrate-and-fire more, with the usual form $\dot{v} = v^2 + I$ being a special case of Eq. 1 (Izhikevich, 2003). Through bifurcation, the biophysically accurate Hodgkin-Huxley model is reduced to a two-dimensional system of ordinary differential equations, which was obtained by fitting the spike activation dynamics of several types of neurons (Izhikevich, 2007). This enables the resetting of auxiliary after-spike when the membrane potential reaches a

threshold and allows resetting both the membrane potential and the membrane recovery when the membrane potential reaches the peak (Izhikevich, 2003).

Variable v represents the membrane potential of the neuron, while variable u represents membrane recovery for the activation of K^+ ionic currents and inactivation of Na^+ ionic currents when the membrane potential reaches the peak, as negative feedback to v. t is the parameter for time. The membrane potential operates in the mV scale while time is measured in millisecond (ms). Line 3 of Equation (1) describes the reset of the membrane potential and recovery variable after the spike reaches the peak. I serves as the input of dc currents injected to the neuron (Izhikevich, 2003).

Different combinations of parameters $a, b, c,$ and d results in various intrinsic firing patterns of different neuronal cells. Depending on the value of b, the resting potential of the neuron model is set between -70 to -60mV. Since most neurons do not have a fixed threshold for action potential, the model's threshold potential can be as low as -55mV or -40mV, which is dependent on the history of membrane potential before the occurrence of the spike. The parameter a gives the time scale of the recovery variable, u, wherein smaller values of a reduce the speed of recovery. Parameter b accounts for the sensitivity of the recovery variable u to the membrane potential's sub-threshold fluctuations. Increasing b enhances the coupling of v and u, which results in sub-threshold oscillations and low threshold spiking dynamics. If the value of b is greater than that of a ($b > a$), the dynamic corresponds to a saddle-node bifurcation of the resting state (Izhikevich, 2000). Parameter c represents the after-spike reset value of the membrane potential, v, which depicts the effect of fast high threshold K^+ conductance. Parameter d defines the after-spike reset of the recovery variable, u, which portrays the effect of the high, slowly reacting conductance threshold for NA^+ and K^+ channels in the neurons (Izhikevich, 2003).

Table 1. Parameter Units for the Izhikevich Firing Model

Parameter	Units
C	pF
$\dot{v}$	$V.s^{-1}$
k	$pA.mV^{-1}$ = n Ω^{-1}
v, v_r, v_t, v_{peak}	mV
u	mV
I	pA
$\dot{u}$	$V.s^{-1}$
a	ms^{-1}
b	$n\Omega^{-1}$
c	mV
d	pA

As a result, the Izhikevich firing model was adjusted to reproduce cells from the preSMA, STN, and PMd areas. Both the preSMA and PMd cells were modeled with Equation (1 by adjusting the parameters to the values as shown in Table 2 (using the Izhikevich value for cortical pyramidal neurons). On the other hand, the STN cells were modeled with an adaptation of Izhikevich's firing model (Izhikevich, 2003; Michmizos & Nikita, 2011), as follows:

$$\begin{gathered} \dot{v} = 0.04v^2 + 5v + 145.5 - u + 1.3\, I_{STN} + noise \\ \dot{u} = \ 0.02[0.2v - u] \\ if\ v \geq 25, then\ \ v \leftarrow -65, u \leftarrow u + 2 \end{gathered} \tag{2}$$

where, I_{STN} is input to STN cells. The part $0.04v^2 + 5v + 145.5$ was obtained by fitting the spike initiation dynamics of the STN cells (Izhikevich, 2003). Equation (1 was transformed into Equation (2 by rescaling the variables (Izhikevich, 2010).

Table 2. Parameters in the Izhikevich Firing Model Adjusted to Fit preSMA and PMd Cells

Parameters	preSMA	PMd
v_r	-60	
v_t	-40	
v_{peak}	0.7	
C	100	
k	0.7	
I	$130\ I_{preSMA}$	$70\ I_{PMd} - 20\ \alpha_L$
a	0.03	
b	-2	
c	-50	
d	100	

I_{preSMA} and I_{PMd} are inputs to preSMA and PMd respectively. The inputs will be discussed in the latter part of this section. Apart from the input, each PMd units received lateral inhibition from the other PMd units in the form of α_L. The number of PMd units depends on the number of motor selections as required in the task. Lateral inhibition from other PMd units decreases the activity of the PMd unit.

Each unit of cells represents the activities of the respective cell groups. For each trial, each neuron model was given a time frame of 2000ms. The activity of preSMA and STN cells started at time = 0ms, while with the burn-in period, the PMd units started receiving input at time = 500ms. Figure 4 shows an example spike train for each modeled neuron.

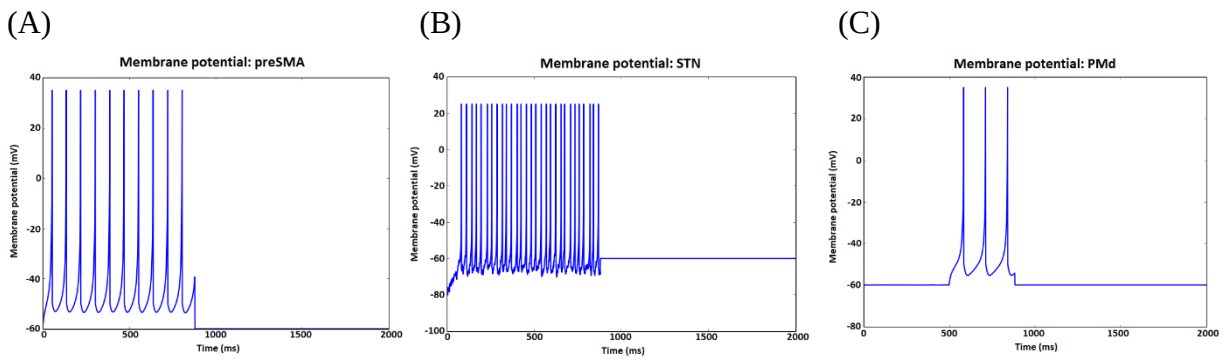

Figure 4. Firing patterns of (A) preSMA (B) STN, and (C) PMd cells. The firing patterns were generated based on the specification stated in Table 2 (preSMA and PMd) and Equation (2) (STN).

Synapse

However, single neurons could not give the whole story of producing behaviors. Neurons work in groups composing networks. A small number of interconnected neurons carry out information processing capabilities that are not seen in single neurons. The interconnection may aid in the exhibition of complex behavior. Our goal was to generate a model that mimics information processing in the brain to simulate the performance of a certain task at some level of abstraction. The model was intended to capture the abstract of the brain's computation to an arbitrary degree, and still predict some aspect of brain activity or behavior. Multilayers of interconnected neurons form brain networks, and the neurons are believed to adjust their synapses to accommodate the appropriateness of the output to the task (Lillicra et al., 2016).

Interneuron connections were simulated as synapses which are simplified by modeling the delays of spike propagation through the synaptic cleft. The simulation represented the gradual delivery of neurotransmitters from the presynaptic neuron to the postsynaptic neuron. A standard way of modeling the synapse is to use an α-function (Ashby & Helie, 2011; Rall, 1967) as follows:

$$f(t_\alpha) = \frac{t_\alpha}{\lambda} exp^{(\frac{\lambda - t_\alpha}{\lambda})} \quad (3)$$

where t_α is the time since the cell voltage reached v_{peak}, λ is a constant that determines the duration of signal propagation in the synapse. Greater λ values delay synaptic transmission. Every time the presynaptic neuron cell spikes, upon reaching the peak membrane potential, v_{peak}, the α-function is delivered to the postsynaptic cell, with spiking time starting at t =0. The α-function has a maximum value of 1.0 and it decays to .01 at $t = 7.64\lambda$ (Ashby & Helie, 2011). While the propagation of neurotransmitter is still in process if a second v_{peak} is reached, a new α-function that corresponds to the second v_{peak} is added to the first α-function, giving an integrated α-activity. The latency in a typical synapse is generally less than 0.5ms (Winlow, 1990) and this delay was approximated with $\lambda = 60$.

Model's Circuit

Response Generation in Hypothesis-Testing and Procedural Learning Systems

From Figure 3, HT was connected to PMd and preSMA, and feedback from responses affected the connection from HT to preSMA. Upon generating a response, the response from HT was fed into PMd units as input (the blue line from HT). The response was then checked for its accuracy, and feedback for both HT and P were given separately. The connection from HT to preSMA (the orange line from HT) served as the confidence of HT ($HT_{confidence}$), which was fed into preSMA as input for activation. Greater $HT_{confidence}$ increased the activity in preSMA. $HT_{confidence}$ was adjusted with feedback from responses through learning.

The response of P, on the other hand, was directed only to PMd. However, this connection was dependent on the inhibition of STN and feedback from the response. STN received its input from preSMA, which activity was heightened with greater $HT_{confidence}$. Thus, greater $HT_{confidence}$ increased activity in STN, preventing the response of P from reaching the PMd units.

Since the sole focus of the model is on the mechanism that facilitates switching between HT and P, both HT and P were modeled as black boxes in our model. It was hypothesized that HT would be preferred in the beginning with high confidence, which leads to the inhibition of P and prevent P activation from reaching the PMd cells. As discussed earlier, different categorization tasks may require a different strategy to maximize accuracy. RB tasks are coupled more with RB strategies which responses should follow mainly the response from HT, whereas II tasks are

coupled more with responses from P. In II tasks, since responses from HT are considered suboptimal, higher failure of responses from HT decreases $HT_{confidence}$. This weakens the inhibition of P, slowly allowing the response generated by P to reach PMd units.

In a behavioral experiment that involved a category learning task, participants were asked to categorize a perceptual stimulus. To simulate this situation, the model received stimuli in the form of its desired categorization response. In a two-choice categorization task, if the stimulus belongs to category 'A', then the desired categorization response of the stimulus is 'A'. The desired categorization response served as the input to the learning systems. Both HT and P have their accuracy probability (Acc_{HT}, Acc_P), and generated responses based on the accuracy probability independently. For instance, if the probability of HT accuracy is 0.8, HT has an 80% chance of producing an accurate response. If the response is accurate, the categorization decision from HT should follow the correct category of the stimulus (the desired categorization response), else the response should be the incorrect category. The same process independently applies to the P system. As a result, the response of each system is independent, one system can be accurate while the other might not.

The accuracy probability of each system is task-dependent. For instance, HT has a higher chance of generating an accurate response in an RB task where the optimal solution requires an RB strategy, as compared to its performance in an II task where the optimal solution requires a procedural strategy.

Our model was able to simulate the accuracy performance in category learning tasks that demanded changing of categorization strategy. Throughout this thesis, this situation is referred to as *task switching*. In the model, task switching could be cued by a perceptual cue. Perceptual cues gave contextual indications to the participants that inform them if they are in the same task or not. When switching from a task to another, due to the proactive interference of the previous task, performance in the new task might dip in the first few trials after a switch. However, in a learned task, the accuracy should be able to recover to a certain extent depending on the persistence of the interference. Accuracy probability of a system, ACC_{System} is given as follows:

$$ACC_{HT} = ACC_{HT_{Task}} - \left[R_{HT_{Task}}(1-D)^N\right]_{after\ task\ switch} \quad (4)$$

$$ACC_P = ACC_{P_{Task}} \quad (5)$$

where, $ACC_{System_{Task}}$ is the accuracy probability of the respective system in a given task, and $R_{HT_{Task}}$ is the proactive interference affecting the accuracy probability of HT of the current task, D is the exponential decay rate, and N is the number of trials since the last task switch.

Procedural Learning System's Inhibition

preSMA. STN exerted an inhibitory activity on P to prevent P's response from reaching the PMd units. The activity of STN was dependent on the preSMA cells. The connection from preSMA to STN was excitatory. High $\boldsymbol{HT_{confidence}}$ activated the preSMA, which led to the indirect inhibition of P's output. This in turn signified the use of HT for the overall response. $\boldsymbol{I_{preSMA}}$ fed into preSMA's Izhikevich firing model as input and activated the unit accordingly. The nonnegative input to preSMA cells in Equation (6) was in the form of square function that varied with the presence of perceptual cues and confidence of HT, $HT_{confidence}$.

$$I_{preSMA} = \left[A_{HT} - A_P + HT_{confidence}\right]^+ \quad (6)$$

where, A_{HT} is the activation of the learning systems HT with the presence of cued RB task, A_P is the activation of the learning systems P with the presence of the cued II task.

Other than the main perceptual stimulus, perceptual cues may be given in an experiment to imply which categorization task is currently carried out. In an experiment that involved more than one category learning task, the perceptual stimulus may be given along with the perceptual cue, so that the participants are aware of the change in task. For example, a stimulus can be presented on a colored background to indicate that the task changes. Different colored backgrounds may be used as perceptual cues for different tasks. Perceptual cue, vc was added to the model to simulate categorization tasks that require a trial-by-trial switch between categorization learning systems. There may be a perceptual cue for an RB task, and another for an II task. For example, a blue background may indicate that subjects are performing an RB categorization task while a green background may indicate an II task. As a result, a task given with the same perceptual cue meant the stimulus shown for that task could be categorized using the same response strategy. A change in the perceptual cue from one to another meant a change in the strategy required to categorize the

stimulus shown. Each task may require a specific strategy to achieve optimum performance (i.e., RB strategy for RB task and procedural strategy for II task).

This part of the model is especially useful with frequent task switches after the participant had gone through training in the task. It acted as a memory guide to direct the model to a learning system after learning the association of the cue to a strategy. With sufficient training, a participant may associate an RB task with an RB strategy, and II task with procedural strategy. An RB strategy signified the use of HT for the overall response, while a procedural strategy signified the use of P for the overall response. When a cued task was presented, the participant may be more likely to use a 'preferred' strategy (or even a learning system) when they have learned the association of such task and strategy. If an RB strategy was 'preferred' in a task, this part of the model increased the preSMA input to increase the inhibition of P's response to the PMd cells, allowing the use of HT's response as the overall response. If a procedural strategy was preferred, the preSMA's input was lowered to reduce the inhibition of P's response to the PMd cells, promoting the use of P's response as the overall response of the model for that trial.

A_{HT} and A_P were given as follows:

$$A_{HT} = S_{HT}[vc_{RB} + (1 - vc_{RB})R_S(1 - D)^N] \quad (7)$$

$$A_P = vc_{II} \times S_P \quad (8)$$

where, vc_{RB} and vc_{II} are the presence of perceptual cue for RB and II tasks (respectively), S_{HT} is the association of an RB strategy to an RB perceptual cue and S_P is the association of a procedural strategy to an II perceptual cue, R_S is the working memory's proactive interference decay effects on preSMA input, D is the exponential decay rate (same as in Eq. 4), and N is the number of trials since the last task switch.

vc was represented in the form of a binary input; when the perceptual cue was present, vc =1, whereas when the perceptual cue was absent, vc = 0. The presence of perceptual cue for an II task ($vc_{II} = 1$) brought down the input to preSMA cells with a factor of S_P, which in turn reduced activation of the STN, thus encouraging the use of P in the overall choice response. On the other hand, the presence of perceptual cue for an RB task ($vc_{RB} = 1$) increased the inhibition of STN with stronger input to the preSMA cells with a factor of S_{HT}. With learned vc_{RB} association to an

RB strategy, when the task switched, even in the absence of vc_{RB}, the interference or continuation of using an RB strategy might still linger with an exponential decay function even when the task has changed to an II task. Participants who failed to switch systems from task to task may have experienced a larger interference-effect. These non-switchers might abandon an optimal strategy (e.g., procedural strategy in II task) during the frequent task switching trials and stick to a suboptimal strategy (e.g., RB strategy in II task). Crossley et al. (2018) showed that in a task intermixed with of RB and II trials, some participants abandon their procedural strategies to use either guessing or an RB strategy in the II task, but optimal strategy use was unaffected in the case of the RB task in the intermixed trials task.

The $HT_{confidence}$ indicated how confident the response from HT is, which affects which system the switching mechanism follows. $HT_{confidence}$ ranged from 0 to 1, and was set to 0.99 initially. Regardless of the perceptual cues, the overall response follows that of HT at the start of the trials because $HT_{confidence}$ was initially high. This satisfied the bias to RB strategies and responses at the beginning of a category learning task.

STN. The integrated $\boldsymbol{\alpha}$-activity from preSMA, $\boldsymbol{preSMA_{\alpha}}$ served as an input to the STN. The $\boldsymbol{\alpha}$- function governing $\boldsymbol{\alpha_{preSMA}}$ is defined in Equation (3).

$$I_{STN} = \alpha_{preSMA} \tag{9}$$

where, t is the time in the trial in ms.

Category selection: PMd. With greater activation of STN, when the accumulation of its integrated $\boldsymbol{\alpha}$-activity exceeded the activity of P, the response of P was inhibited. The equation that governs the inhibition of the response of P was given by:

$$I_{PMd} = \left(H\mathrm{T}_{\mathrm{response}}\right) + \left[\left(P_{response} \times P_{confidence}\right) - \left(0.67\ \alpha_{STN}\right)\right]^{+} \tag{10}$$

where, I_{PMd} is the input to PMd unit, $HT_{response}$ is the response of HT that goes to the PMd unit, $P_{response}$ is the response of P that goes to the PMd unit, $P_{confidence}$ is the confidence of P, and α_{STN} is the integrated α-activity of STN. α_{STN} is defined in Equation (3).

Initially, when the value of $HT_{confidence}$ was high enough, the activity of STN was high enough so that the value of its integrated α-activity was greater than the value of $P_{response} \times P_{confidence}(n)$. The difference in the output from STN and the response of P as shown in Equation (10) gave a negative value. With the + sign in the equation, any difference that was smaller than 0 is set to 0. The part of the equation left only any positive values and set any negative values to 0. The addition of the difference to the response of HT and its weight gave the input to the corresponding PMd unit. However, if the integrated α-activity of STN was smaller than the value of $P_{response} \times P_{confidence}(n)$, the difference gave a positive value. Hence, P was not inhibited, and the response from P was added to the response of HT as the input into the PMd unit.

The PMd units took in input from Equation (10). The PMd units had a burn-in period of 500ms so that the units would not take in any inputs in the first 500ms of a trial. The burn-in period was required for the stabilization of STN's activity, and to allow HT and P to compute their responses (Buesing et al., 2011).

The PMd unit gave the output in the form of α-function, which was given by Equation (3). The PMd unit received and exerted a lateral inhibition to other PMd units with the integrated α-activity. A higher activity in one of the PMd units would suppress the activity in the other PMd unit.

The PMd unit with the integrated α-activity that reached the threshold for decision first was selected as the response of the overall system. If the α-activity of all units reached the threshold at the same time, the choice was random. If no units reached the threshold by the end of the trial, the unit with higher activity was selected as the response of the overall system. This latter case was considered an informed guess when a forced-choice was required.

Learning

In many cases, learning is seen as utilizing error signals that result from a mismatch between expectation and actual perceptions (Lillicrap et. al, 2016). In visual neuroscience, visual processing receives input in the form of image bitmaps and estimates the brain activity or

behavioral responses through deep neural networks (Kriegeskorte & Douglas, 2019). However, the training of deep neural networks relies on category-labeled images, and since there are no labeled examples in biological learning and development, such models are just visual processing models but not of development and learning (Kriegeskorte & Douglas, 2019). Reinforcement learning, on the other hand, uses environmental feedback in learning processes.

Both theoretical ideas of model-based and model-free reinforcement learning have been used to understand learning in biological systems at both the behavioral and neural systems levels (Banino, Caswell, & Kumaran, 2018; Mattar & Daw, 2018). With the success of temporal-difference reinforcement learning and Rescorla-Wagner theories for predicting dopamine-producing neuron's activities and its effect on behavior, the model-free reinforcement learning is possibly the best understood system as the neural basis of reinforcement learning in mammalian neuroscience (Frank, 2005; Houk, Adamas, & Barto, 1995; Montague, Dayan, & Sejnowski, 1996; Rescorla & Wagner, 1972b; Schultz, Dayan, & Montague, 1997). From one of the models available, the frontal cortex represented the set of available choices (Mink, 1996). Information about the choice values was encoded through the strength of the cortical synapses, whereby stronger synapses amplified striatal cells' activity (Frank, 2005). The striatal activities represented the values of choices as represented in the cortex (Lau & Glimcher, 2008; O'Doherty et al., 2004). Choice activities were driven accordingly, through a series of downstream circuitry via the basal ganglia, through the thalamus and to the cortex, or through downstream projections to brain-stem motor output areas (Neftci & Averbeck, 2019).

A necessary feature of reinforcement learning is to have a high temporal resolution for the reward signal. Such learning can be facilitated by dopamine; following a correct response, the 'immediate' release of dopamine into relevant synapses aids in the strengthening of appropriate synapses (Helie, Ell, & Ashby, 2015). Within the striatum, dopamine is quickly cleared from the synapses by dopamine activity transporters, which results in a high enough temporal resolution for dopamine to act as an effective reinforcement-learning signal. However, in the cortex, the low concentrations of dopamine activity transporters make the cortical dopamine levels change slowly (Helie et al., 2015). In a training session, the first rewarded behavior may lead to an upsurge in the cortical dopamine level, but the level tends to remain high for an extended time. This would lead to the strengthening of all synapses activated in the training regardless of the appropriateness of associated behavior. In other words, whether or not the resulting behavior was rewarded, cortical

long-term potentiation would occur throughout the training session. As suggested by Hélie, Ell, and Ashby (2015), learning in the frontal cortex may be restricted to correct responses.

Thus, such feedback learning can be represented by populations of neurons that are active when rewards and punishment are obtained (Hélie et al., 2015). Operating on the properties of synaptic depression, Rolls & Deco (2016) proposed a mechanism for non-reward computation in the orbitofrontal cortex. Expected reward triggers the activity in a pool of neurons. Their firing is maintained until the synaptic depression lessens the firing rate. With the absence of expected rewarding outcomes, the declining activity in the reward neurons releases the inhibition by inhibitory neurons, which leads to the activation of the second population of non-reward neurons (Rolls, 2016).

The Rescorla Wagner, (1972) model can describe the behavior condition response learned from the difference between expectation and environmental feedback. With a choice being made, the difference between the expected value of the chosen option, $v_{i,}$ and the experienced outcome, r is computed. If the outcome exceeds expectations, the association strength of appropriate synapses is enhanced. If the outcome is worse than expected, the strength is reduced. The Rescorla-Wagner equation (Averbeck & Costa, 2017; Rescorla & Wagner, 1972b) summarizes the process of value adjustment during learning for behavior in its simplest form, which is given by:

$$v_i(k+1) = v_i(k) + \gamma(r(k) - v_i(k)) \tag{11}$$

where γ is the learning rate of the model that controls the size of each learning updates.

Association of Task Cues to Strategy (Learning System)

Conditioned learning for each strategy to a cued task was modeled with the Rescorla-Wagner model as follows:

$$S(n+1) = S(n) + \left[vc(n) \times F_S(n) \times \gamma_S \times \left((S_{max}) - S(n)\right)\right] \tag{12}$$

where S is the association of task perceptual cues to the strategy or learning system (S_{HT} for RB strategy and S_P for procedural strategy), γ_S is the learning rate for the Rescorla-Wagner model for the association, F_S is the accuracy feedback of the corresponding learning system when compared to the actual category of the stimulus, and S_{max} is the maximum value S can carry. In the presence of a perceptual cue, the corresponding S is learned with the accuracy feedback of a particular learning system (HT for RB strategy and P for procedural strategy). F_S took either the value of 0 or 1, with 0 representing mismatch of the learning system's response with the actual stimulus category and 1 as the accurate response of the learning system. When the particular learning system gave a correct response with the presence of perceptual cue, the corresponding S increased with γ_S and the difference between S_{max} and the previous S value. As S approached its maximum value, the increment per trial decreased.

Confidence of HT

$HT_{confidence}$ was then adjusted according to the success of HT's response in each trial with Rescorla-Wagner learning:

$$HT_{confidence}(n+1) = HT_{confidence}(n) + \left[\gamma_{HT} \times \left(F_{HT}(n) - HT_{confidence}(n)\right)\right] \quad (13)$$

where, γ_{HT} is the learning rate for the Rescorla-Wagner model for HT's confidence and F_{HT} is the accuracy feedback for HT. When the HT gave a correct response, $HT_{confidence}$ increased with γ_{HT} and the difference between 1 and the previous $HT_{confidence}$. As $HT_{confidence}$ approached 1, the increment per trial decreased. When the HT gave an incorrect response, $HT_{confidence}$ decreased with γ_{HT} and the difference between 0 and the previous $HT_{confidence}$. As $HT_{confidence}$ approached 0, the decrement per trial decreased.

Confidence of P

The weight of P's response varied with the confidence of P. The latter was adjusted according to the accuracy of the response of P, which was given by the Rescorla-Wagner learning model as follows:

$$P_{confidence}(n+1) = P_{confidence}(n) + \left[K \times \gamma_P \times \left(P_{confidence_{max}} - P_{confidence}(n)\right)\right] \quad (14)$$

and $K = \begin{cases} 0 & if\ respone\ of\ P\ is\ incorrect\ or\ the\ system\ does\ not\ follow\ reponse\ from\ P \\ 1 & if\ reponse\ of\ P\ is\ correct\ and\ the\ system\ follows\ response\ from\ P \end{cases}$

where, γ_P is the learning rate of the Rescorla-Wagner model for $P_{confidence}$. When the P is in use and gives a correct response, $P_{confidence}$ increases with γ_P and the difference between the maximum value of $P_{confidence_{max}}$ and the previous $P_{confidence}$. As $P_{confidence}$ approaches $P_{confidence_{max}}$, the increment per trial decreases.

II CATEGORY LEARNING TASK

Experiment Description

The II experiment was based on previous work that was presented at the Cognitive Science Society conference (Lim & Hélie, 2019).

Participants

Fifty participants were recruited from the Purdue University undergraduate population. Each participant was given credit for participation as partial fulfillment of a course requirement. Participants gave written informed consent and all procedures were approved by the Purdue University Human Research Protection Program Institutional Review Board. All the participants underwent a Perceptual Categorization Task, PCT, and an Iowa Gambling Task, IGT. However, the current simulation focused on the PCT only.

Perceptual Categorization Task

The experiment was run on a Desktop PC equipped with a regular mouse and keyboard. Stimuli were displayed in a 21-inch monitor with 1,920 × 1,080 resolution. The experiment was controlled by in-house programs written using PsychoPy.

The stimuli used in the PCT were circular sine-wave gratings of fixed contrast and size, as shown in Figure 2. The stimuli differed in terms of orientation and bar width. The orientation is the counterclockwise rotation of the black bars from the horizontal axis in radian. The bar width was derived as the frequency of black bars in a two-dimensional space in cycles per degree. The stimuli were grouped as either A or B, with a diagonal line as a category-bound as shown in Figure 5. It was possible to have perfect accuracy and optimal performance performing the task using a procedural strategy.

The participants were informed that they were taking part in a categorization experiment and that they needed to learn to categorize the stimuli presented into either category A or B using trial-and-error. In each trial of this task, a “crosshair” was presented on the screen for one second, followed by a single stimulus presented in the center of the screen. Participants were required to choose a category for the stimulus. The stimulus was presented to the participant until a response

was made. Responses were given on a standard keyboard, with labeled stickers of “A” and “B”: whereby “s” key was labeled with sticker “A” for category A and “k” key was labeled with sticker “B” for category B. After each trial, visual feedback showing “Correct”, “Incorrect”, or “Wrong Key” was shown to the participant for one second according to their choice of categorization. The response for stimulus on each trial was recorded, as well as the response time. The participants did 600 trials grouped into six blocks of 100 trials each. The PCT took about 35 minutes to complete.

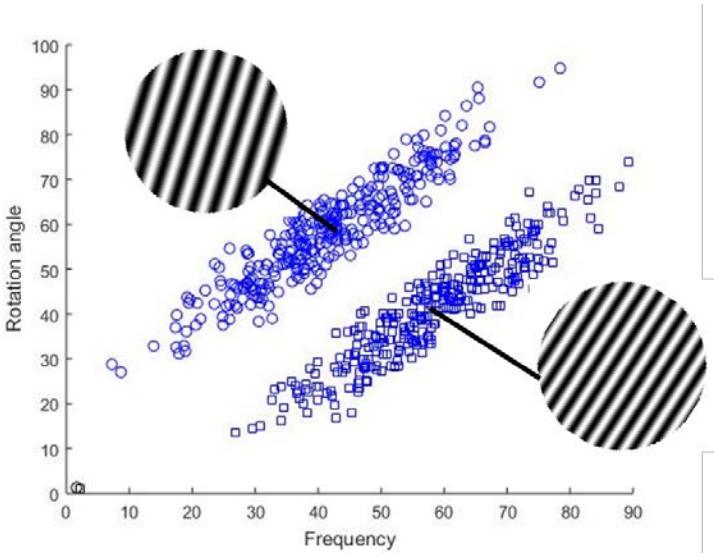

Figure 5. Category structures used in PCT (II task). “○” denote members of category “A”, and “□” denote members of category “B”. Frequency indicates bar width and is calculated in cycles per degree (cpd). Rotation angle indicates the orientation of the stimulus, and is calculated as the counterclockwise rotation from horizontal calculated in radians.

Behavioral Data Analysis

Decision Bound Models

Decision Bound Models (DBM) were applied to the data to allow the classification of participants into optimal strategy users and suboptimal strategy users. DBM was applied to the PCT to identify how participants learn to assign responses to regions of perceptual space. In DBM, it is assumed that participants determine the perceptual region where the stimulus is located and output the associated response. The decision bound is described as a partition segregating competing response regions. We used three general classes of the decision bound models to fit the

response data from the PCT (Hélie et al., 2017; Maddox & Ashby, 1993), which are: (1) the guessing models, (2) the explicit rule-reasoning models, and (3) the procedural learning models.

The guessing models assume that participants do not rely on the information of the stimulus shown on the screen and responded either "A" or "B" randomly on each trial. There were two guessing models, namely the "guessing" model and the "biased" model. The guessing model assumed an equal probability of selection for each response, and there was no free parameter in this model. The biased model assumed that participants are biased towards one of the responses, and they guessed one response with a probability p, and the other response with probability 1-p, whereby p is the free parameter of the model (Hélie et al., 2017).

With the data plotted in the Cartesian plane, the hypothesis-testing models used a describable rule-based boundary to segregate participants' responses with a vertical line, a horizontal line, or the combination of both a vertical line and a horizontal line. The hypothesis-testing models assumed participants make a separate decision for each stimulus dimension and combine these decisions depending on the criterion required. A decision bound is perpendicular to each stimulus dimension that defines the presented stimuli. For the PCT, the decision criterion may revolve around two stimulus dimensions, which are the rotation angle and bar width. An example of the rule for the rotation angle is "Choose response A if the rotation is greater than 30º, and choose response B otherwise". On the other hand, the rule for the bar width might be "Choose response A if the bars are thick, choose response B otherwise". The one-dimensional classifiers for the hypothesis-testing models had two free parameters, which were the decision criterion along the relevant perceptual dimension and the perceptual noise variance. Two-dimensional classifiers used a conjunction rule from the two stimulus dimensions and had three free parameters, which were: one decision criterion for each dimension and a common perceptual noise variance (Hélie et al., 2017).

On the other hand, the procedural learning models employ a diagonal line as the boundary, since they are incompatible with verbal rules. As a contrast to the hypothesis-testing models, in the procedural models, it was thought that perceptual information from all relevant dimensions has been integrated prior to making a decision. The information integration can be either linear or non-linear. In this experiment, we assumed a linear integration and used the general linear classifier (GLC) to divide stimulus space with a linear decision bound. GLC was governed by three parameters, which were the slope and intercept of the linear bound and perceptual noise variance.

The linear decision bound separated the plane, where one side of the bound is associated with the response "A", and the other with the response "B" (Hélie et al., 2017). For each participant's data set, the best model was selected using the Bayes information criterion (BIC):

$$BIC = r \times \ln(N) - 2 \times \ln(L) \tag{15}$$

where, r is the number of free parameters in the model, N is the size of the data block, and L is the likelihood of the data given the model.

Since the PCT was an II categorization task, participants who were able to use a procedural strategy were labeled as switchers. Switchers were able to switch away from an RB strategy to a procedural strategy as they learned throughout the task. From the DBM, the categorization performance with GLC showed categorization with integrated perceptual information from all relevant dimensions. GLC signified the use of a procedural strategy in categorization. Thus, switchers were participants with their performance indicated as GLC in the final block of the task, while non-switchers were participants without GLC.

For each participant, the categorization accuracy for each block of 100 trials was computed based on the match between the participant's responses and the actual category of the stimuli for the trials within the blocks. The accuracy learning curves with 6 blocks of 100 trials were averaged based on the categorization accuracy of all participants with the same switching capabilities (separated by switchers and non-switchers). The accuracy learning curves were presented in Figure 6 as the percentage accuracy of switchers and non-switchers.

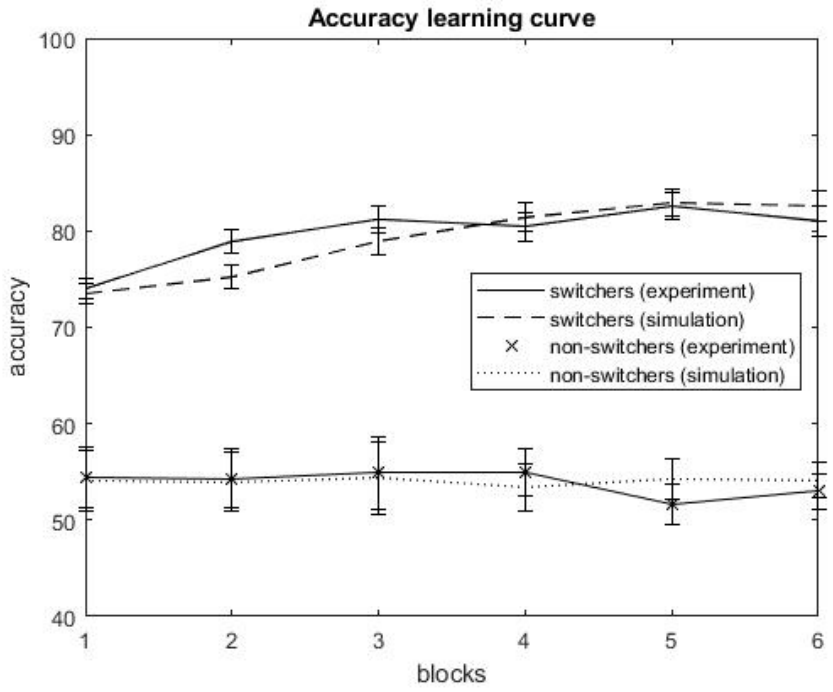

Figure 6. Average accuracy in PCT and model's simulation results for each block of 100 trials. Solid and dashed lines show data for participants who used an optimal strategy (switchers), whereas solid (with 'x' markers) and dotted lines indicate participants who used a suboptimal strategy (non-switchers). Error bars for experimental data are the standard error of the mean, while error bars for simulation data are the standard deviation.

Model Adaptation

We ran 100 simulations with the model, representing 100 'participants'. The input to the model was the correct category of each stimulus as shown to the participant in PCT, which was given as either "A" or "B". Each simulation consisted of 6 blocks of 100 trials of categorization. Throughout the tasks, there was no perceptual cue for the use of a particular category learning strategy, thus the perceptual cues were set to 0.

During each trial, the HT and P systems in the model generated a response according to the actual category of the stimulus and the system's accuracy probability. Since there were two categories (A and B) in the experiment, the response of HT and P served as the inputs for PMd units' activation for selection of Category A or B. The responses from the two systems were fed into the neuron-based circuit as described in the Methods section. The circuit was modeled as a small scale representation of the respective brain areas to hold a unit of STN cell, a unit of preSMA cell, and two units of PMd cells.

The two PMd units were included to represent category selection, one to account for pressing the button corresponding to Category A, and the other for Category B. From the output of HT and P, the PMd units A and B were activated accordingly. Since the PMd units incorporate lateral inhibition, PMd unit A received and exerted lateral inhibition on PMd unit B, and vice versa.

Table 3. Parameter Values Used in the Simulation of PCT That Were Common for Both Switchers and Non-Switchers

Model-Level	Parameter	Value
P's response generation	Acc_P	0.84
Perceptual cues	vc	0 for all trials
$HT_{confidence}$ learning	γ_{HT}	0.008
	$P_{confidence_{max}}$	5.1
Category selection	PMd integrated α-activity threshold	1.33

Note. These parameters were found using a rough grid search.

For each trial, when the output of PMd units (in the form of integrated α-activity) reached the threshold, the PMd unit that reached the threshold first was the winner of response selection. The category it represented was selected as the overall outcome of the model for that trial. If none of the α-activity of the two PMd units reached the threshold in the given time of 2000ms, the PMd unit that had a higher α-activity at time=2000ms was selected as the winner. This latter case happened in about 3% of cases in the simulations. Furthermore, if both units reached the threshold at the same time, the response was selected at random.

Accuracy Learning Curve

One category outcome, either category A or B was chosen in each trial. For each trial, if the suggested category outcome from the model matched with the actual category of the stimulus, the model received feedback with the value 1, and 0 if the two categories did not match. For each simulation, the categorization accuracy for each block of 100 trials was computed based on the

match between the category outcome suggested by the model and the actual category of the stimulus for the trials within the blocks. Thus, the accuracy learning curves with 6 blocks of 100 trials were shown in Figure 6. The parameters used in the simulation are shown in Table 3.

The goal of the simulation was to fit the data according to the two groups of participants, GLC, and non-GLC group, where participants in GLC are switchers for the task while participants in the non-GLC group are non-switchers. The parameters in the model were adjusted to fit the accuracy of categorizing the stimuli based on the two participant groups. The changes in parameters that resulted in the difference between switchers and non-switchers are shown in Table 4. As can be seen, only 2 free parameters were needed to account for group differences. From Figure 6, the accuracy performance of the two groups was simulated with an excellent fit (RMSD of 1.69).

Table 4. The difference in Parameter Values of the Model to Separate Switchers From Non-Switchers

	Value	
Parameters	Switchers	Non-switchers
Acc_{HT}	0.745	0.54
γ_P	0.07	0.003

The accuracy for HT of switchers was set higher than that of the non-switchers. Acc_{HT} accounted for the probability accuracy of HT, which specified how likely it is for HT to produce an accurate answer for each trial. With a higher value of Acc_{HT}, HT had a higher probability of producing a correct response for each trial; with a lower Acc_{HT}, HT had a lower probability of producing a correct response for each trial. Since the accuracy of HT was based on the stimulus dimensions from the PCT, assuming the participants picked the optimal rule to categorize the stimuli, a one-dimensional rule would give an accuracy of about 0.76 (or 76%). If participants used a conjunction rule that takes into consideration both the varying stimulus dimensions, the accuracy would rise to about 0.91 or (91%), assuming the best rule is used. However, due to past results from experiments, it is highly likely that participants used a one-dimensional rule in such

categorization tasks. As a result, Acc_{HT} in both switchers and non-switchers was constrained to be lower than the accuracy of 0.76.

The γ_P of switchers was higher than that of the non-switchers. Higher γ_P signified a higher learning rate for the procedural system's confidence for each update. If the P produced a correct response, the increment for the learning update would be higher. Lower γ_P resulted in a lower increment of learning updates. Hence, the model simulation suggests that switchers develop better RB strategies and learn about the efficiency of the P system more quickly when compared to non-switchers.

Switchers

With an increment of HT's accuracy and learning rate of confidence in P, the model was able to distinguish switchers from non-switchers. With higherAcc_{HT}, participant's accuracy, even when they are using RB strategies, was higher than that of the non-switchers. High Acc_{HT} might render switching harder, but the inability to switch was overcome by having higher γ_p to encourage P's confidence at a higher rate. Thus, even if the confidence of HT was high, P's high confidence allowed system switching in categorization when more trials were carried out.

Figure 7 shows the ratio of activity in P to that of HT for switchers. In the beginning, the ratio was 0, since the response from P was not fed into PMd units, due to the inhibition from STN. As P gained more confidence by producing correct responses, the ratio increased. When the ratio was below 1, if the two system's responses conflicted, the overall response of the model followed the response from HT. When the ratio was around 1, the competition between the systems was higher, and the model's overall response may follow either one of the systems (or both when the suggested responses from both systems were the same). When the ratio was much greater than 1, the overall response of the model followed the response of P when the two system's responses differed. Thus, Figure 7 shows the switch from HT to P in a categorization task, which implied the use of RB strategy to a procedural strategy. The switch occurred approximately after 200 trials of the task. From the behavioral data from PCT, about 31 out of 34 switchers used the procedural strategy by the second block of the task (trial = 200).

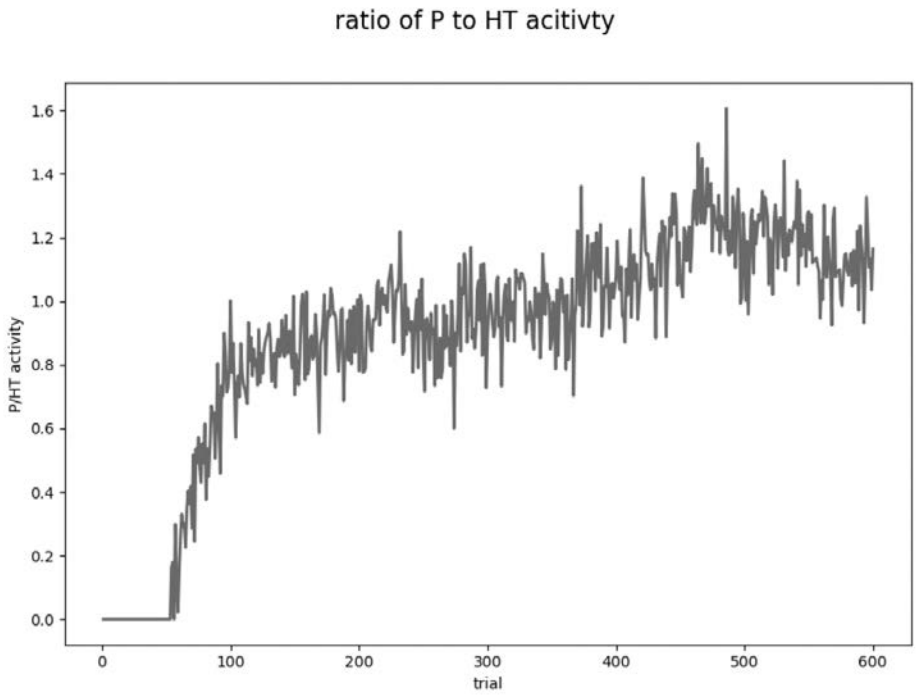

Figure 7. The ratio of activity in P to HT for every trial for switchers.

Non-Switchers

Non-switchers had lower HT accuracy and slower learning rate for confidence in P. The non-switchers comprised of those that used only RB strategies throughout the task, and those that could not even find a general strategy in categorization. From Rabi and Minda (2014), people that relied on guessing when completing the categorization task were believe to be struggling with identifying the correct rule or did not apply a rule consistently. The average accuracy of participants in the non-switchers was generally lower than that of the switchers, even in the first block where switchers still used RB strategy. Thus, the Acc_{HT} for non-switchers was lower than that of the switchers.

Switching can occur both ways when HT fails to give accurate responses over time or when P gained enough confidence. HT's consistent failure of giving a correct response should, in time, decrease the STN's activity to suppress P's outcome from reaching the PMd. However, even with sufficiently low STN activity, to suppress switching, the weight of P's response should be still lower than the weight of STN's activity. Thus, γ_p was set low enough to slow the confidence gaining process of P.

Figure 8 showed the ratio of activity in P to that of HT for non-switchers. The ratio was less than 1 throughout the task, which indicated the use of HT's response as the overall response of the model, especially when the response from the two systems differed. This implied the failure to switch from HT to P, and the constant usage of RB strategy.

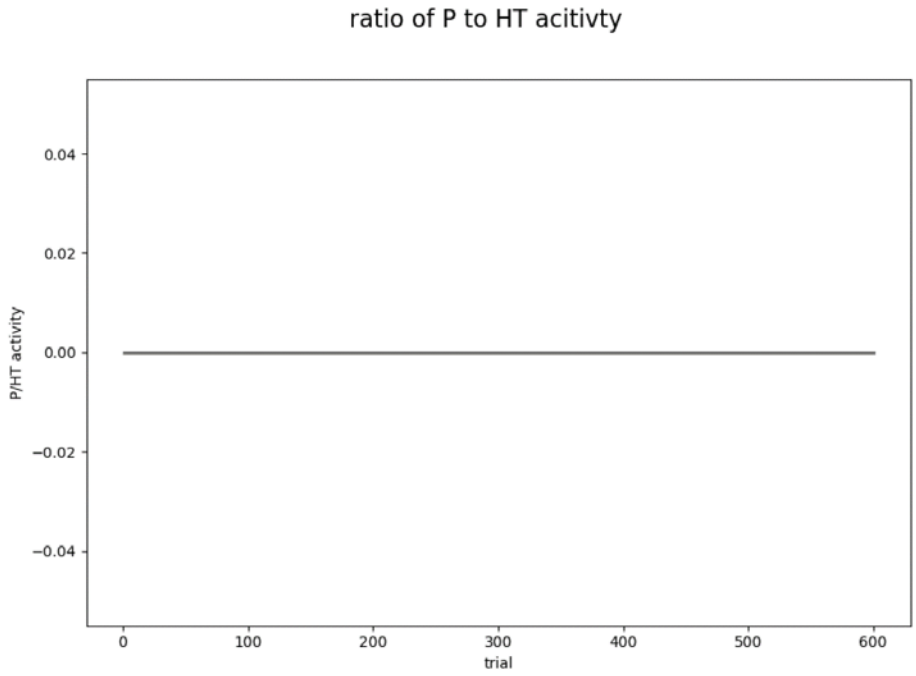

Figure 8. The ratio of activity in P to HT for every trial for non-switchers.

Discussion

In short, the ability to search for a better strategy drives the switching capabilities of switchers and non-switchers apart. At the beginning of the task, even when using an RB strategy, switchers can categorize the stimulus better than the non-switchers. This could mean that it might be difficult for non-switchers to find a good RB strategy for the task. Some of the non-switchers could revert to guessing the answers most of the time. Switchers can tackle the task with a better RB strategy and eventually learn about the efficiency of the P system more quickly.

CATEGORIZATION SYSTEM-SWITCHING DEFICITS IN TYPICAL AGING AND PARKINSON'S DISEASE

Experiment Description

Introduction

The work was based on the results of the categorization system-switching comparison for participants from different age groups and those with Parkinson's disease from Hélie and Fansher (2018). Past research had shown an association of aging with cognitive flexibility, where older adults are often worse at adapting to shifting situational demands when compared to younger adults (Wilson et al., 2018). The work from Hélie and Fansher (2018)was carried out to see if people of different age groups and different tonic dopamine levels can switch between different categorization systems flexibly on a trial-by-trial basis. The research explored the deficits in categorization system-switching in older adults and patients with Parkinson's disease, for whom the loss of dopamine-producing neurons is greater.

The experiment involved two types of category learning task: a RB category task and an II category task. Throughout the experiment, participants were asked to categorize circular sine gratings. In the II task, the stimuli differed in terms of bar width and orientation, while in the RB task, the stimuli differed only in terms of bar width. The category structures used in the RB and II tasks are depicted in Figure 9.

From Hélie and Fansher (2018), the participants took a categorization experiment with separated training and testing phases. Young adults, older adults, and people with Parkinson's disease were first trained in an RB categorization task, followed by training in an II categorization task. During the training phase, the two tasks were done separately. Color cues were provided as an indication of the type of tasks for the particular trial (RB or II). The stimulus was presented on a blue screen in RB tasks, while the screen's color was green in II tasks. After the training phase, the participants went through the testing phase. In the testing phase, a block of trial-by-trial switching was presented, where RB and II categorization tasks were intermixed[1]. The data from

[1] The experiment from Hélie and Fansher (2018) consisted of seven blocks of 100 trials. Only the first six blocks were simulated in this thesis. The last block was the same categorization task as the sixth block with the addition of a change in the button location. The button-switch was only used as a sanity check for strategy identification and is beyond the scope of the simulation. Thus, the button-switch block was not included in the simulation.

the testing phase were used to compute the switching cost in terms of the accuracy of categorization and response time changes.

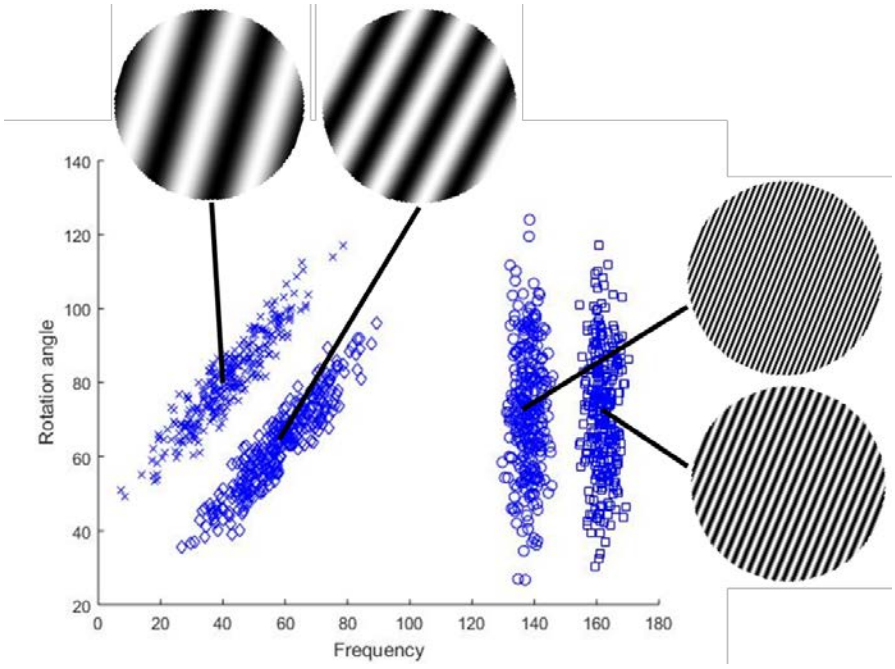

Figure 9. Category structure used in RB and II task. 'x' denote members of category "A", "◊" denote members of category "B", "○" denote members of category "C", "□" denote members of category "D". "A" and "B" are II categories, while "C" and "D" are RB categories. Frequency indicates bar width and is calculated in cycles per degree (cpd). Rotation angle indicate the orientation of the stimulus, and is calculated as the counterclockwise rotation from horizontal calculated in radians.

The study was done with the hypothesis that aging and lesser tonic dopamine levels would impair categorization system-switching. This impairment would be worse for people with Parkinson's disease when compared to the same age controls. The impairment would result in a higher switch cost, which lowers the proportion of participants that can switch on a trial-by-trial basis, and reduces the overall accuracy in the intermixed testing phase.

From the results, participants in all groups performed well in the training phase, but the group difference was shown in the intermix testing phase. The young adult group had higher accuracy in categorization as compared to the Parkinson's disease group, however, the difference in the older adult group and the Parkinson's disease group was not significant. The proportion of switchers in the young adult group was greater than that in the older adult group and Parkinson's

disease group. The accuracy switch cost was significant for the young adult switcher and non-switcher. For older adults, the accuracy switch cost was significant for the switchers and trending for non-switchers. For Parkinson's disease group, the accuracy switch cost was significant for non-switchers and non-significant for switchers. In short, the results suggest that aging decreases the ability to switch between categorization systems.

Model Adaptation

For each group of participants, we ran 100 simulations with the model, representing 100 'participants'. The model took in the actual category of each stimulus as shown to the participant in PCT, which was given as either "A", "B", "C" or "D". Each simulation consisted of 6 blocks of 100 trials of categorization. The task flow of the model follows the order of:

100 RB training task → 400 II training task →100 intermixes of RB and II tasks

The stimuli given to the participant were sine wave gratings similar to the stimuli from the previous task. Each sine wave gratings differs in terms of bar width and orientation. Also, perceptual cues were given to the participants along with the stimulus: a blue background that signifies RB task and green background for II task. In the model, both perceptual cue parameters vc_{HT} and vc_P gave binary output (0 or 1). During an RB task, vc_{HT} is set to 1 while vc_P to 0, and the reverse is true for II task. With the relationship of RB task to HT and II task to P, presence of vc_{HT} enhances activation to preSMA cell and thus inhibiting P's responses from reaching the PMd cells, whereas the presence of vc_P should decrease such activation to preSMA and release the inhibition of P's response. However, vc_P should not have as much effect on the preSMA units at the beginning of the trials and will only start influencing responses after undergoing more training trials. It took longer to acquire a procedural strategy as compared to an RB strategy (Helie & Fansher, 2018). This was achieved by setting the learning rate of vc_P association to a lower value as compared to that of vc_{HT}.

With two different tasks (RB and II) simulated, the accuracy probability of HT and P is task-dependent. The accuracy probability of HT and P for RB and II tasks are tabulated in Table 5. RB strategies give a more optimal performance in the RB task, while procedural strategies are

the optimal strategy for an II task. The accuracy probability of each system is given in Equations (16 and (17. The accuracy probability for each system changes with the task. In the model, the change in the task was hinted with the change in the perceptual cues. In an RB task, $vc_{HT} = 1$ and $vc_P = 0$, ACC_{HT} followed its accuracy probability in a RB task, $ACC_{HT_{RB}}$, and ACC_P followed its accuracy probability of P in a RB task, $ACC_{P_{RB}}$. In II task, $vc_{HT} = 0$ and $vc_P = 1$, ACC_{HT} followed its accuracy probability in a II task, $ACC_{HT_{II}}$, and ACC_P followed its accuracy probability in a II task, $ACC_{P_{II}}$. Upon switching from one task to another, the accuracy of the response of HT is influenced by the proactive interference, which may cause a slide in the performance at the beginning of a switch. The influence was modeled as an exponential decay and was dependent on the previous task type (RB or II). The interference was induced by the task switch.

Table 5. Accuracy Probability of HT and P for RB and II Tasks

Accuracy Probability	Value
$ACC_{HT_{RB}}$	0.90
$ACC_{HT_{II}}$	0.78
$ACC_{P_{RB}}$	0.73
$ACC_{P_{II}}$	0.88

$$ACC_{HT} = \left(vc_{HT} \times ACC_{HT_{RB}}\right) + \left(vc_P \times ACC_{HT_{II}}\right) - \left[R_{HT_{RB}}(vc_{HT}) + R_{HT_{II}}(vc_P)\right](1 - D)^N \quad (16)$$

$$ACC_P = \left(vc_{HT} \times ACC_{P_{RB}}\right) + \left(vc_P \times ACC_{P_{II}}\right) \quad (17)$$

The individual learning systems HT and P generate category responses based on their accuracy probability independently. With two tasks modeled, two sets of category options were provided: A-B (RB task) and C-D (II task). In each ask, the model can only choose from the two options (A and B, or C and D) depending on the task. Both HT and P held four units of responses

each, representing categorical choices of A, B, C, and D. The response from each system was fed into the PMd units (depending on the inhibition for P's response).

Similar to the previous task, the circuit was modeled as a small scale representation of the respective brain areas to hold a unit of STN cell and a unit of preSMA cell. With four options altogether, the model had four PMd units, one for each category options, representing groups A, B, C, and D. Each unit of cells represented the activities of the respective cell groups. For each trial, each neuron model was given a time frame of 2000ms. The activity of preSMA and STN started at time = 0ms, while with the burn-in period, the PMd units started receiving input at time = 500ms.

The model was adjusted to accommodate the trial-by-trial switching task for three groups of participants: young adults, old adults, and Parkinson's disease patient groups. Importantly, switchers and non-switchers were not separately simulated as in the previous task. The three groups were simulated with mixed switchers and non-switcher where the proportion of switchers to non-switchers was the model's outcome based on the manipulation of parameters to achieve the simulation of the specific participant group. To include both switchers and non-switchers in a group, the model was built so that the parameters γ_P and D varied from 'people-to-people', while all other parameters were fixed for simulations in the same group. The 'participants' were normally distributed and the parameters γ_P and D were computed in a logistic function. Since both parameters should be modeled independently, 'participant's' distribution was obtained twice, one for each parameter. The parameters used in the simulation of the experiment that were common for all three groups of participants were tabulated in Table 6.

$$\gamma_P = \frac{0.05}{1 + e^{-5\left(\gamma_{P_{dist}}\right)}}, \qquad \gamma_{P_{dist}} = N(\gamma_{P_{dist\,mean}}, 0.6) \tag{18}$$

$$D = \frac{0.99}{1 + e^{-1.65(D_{dist})}}, \qquad D_{dist} = N(D_{dist_{mean}}, 1.5) \tag{19}$$

where, $\gamma_{P_{dist}}$ is the distribution for participants in a group mainly to capture the value γ_P, $\gamma_{P_{dist\,mean}}$ is the mean for $\gamma_{P_{dist}}$ to determine the distribution of γ_P for different participants, D_{dist} is the distribution for participants in a group as a factor for the exponential decay rate for the

proactive interference (D), and $D_{dist_{mean}}$ is the mean for D_{dist} to determine the distribution of D for different participants. γ_P ranged from 0 to 0.05, distributed based on $\gamma_{P_{mean}}$. D ranged from 0 to 0.99, distributed based on $D_{dist_{mean}}$.

Table 6. Parameter Values Used in the Simulation of Trial-by-Trial PCT That Were Common for all Three Groups of Participants

Model-Level	Parameter	Value
HT and P	$R_{HT_{RB}}$	0.42
	$R_{HT_{II}}$	0.2
Perceptual cues	γ_{vc} for RB strategy	0.1
	γ_{vc} for procedural strategy	0.006
	$vc_{weights_{max}}$ for RB strategy	0.3
	$vc_{weights_{max}}$ for procedural strategy	0.03

Note. Parameters that were not mentioned here are those that share the same values from the previous experiment.

The goal of the simulation was to fit the data according to the three groups of participants. We fitted the model to find the parameters that govern switching between the systems that are affected by the change in aging. The changes in such parameters were reflected in the differences in the accuracy learning curve, accuracy switch cost, and switching proportions in the three groups of participants.

Differences in the Three Groups

The three groups of participants were differentiated by adjusting the value of $\gamma_{P_{dist_{mean}}}$, $D_{dist_{mean}}$, and R_{VC}. The value of the three parameters that set apart the three groups are tabulated in Table 7. Generally, the combination of higher $\gamma_{P_{dist_{mean}}}$ and $D_{dist_{mean}}$ values, and a lower R_{VC}, value facilitates the switching capabilities.

Table 7. The Difference in Parameter Values of the Model to Simulate the Three Groups of Participants

Parameters	Value		
	Young Adults	Old Adults	Parkinson's Disease Patient
$\gamma_{P_{dist_{mean}}}$	0.08	-0.21	-0.35
$D_{dist_{mean}}$	0.95	0.23	0.12
R_S	0.4	0.7	1.2

Note. Both the parameters $\gamma_{P_{dist_{mean}}}$ and $D_{dist_{mean}}$ go into a sigmoid function. A negative value in the $\gamma_{P_{dist_{mean}}}$ do not yield a negative value for γ_P

$\gamma_{P_{dist_{mean}}}$ is the mean of the normal distribution of 'participants' for $\gamma_{P_{dist}}$ in a given group. $\gamma_{P_{dist}}$ is dimensionless and gives the distribution for participants in a group as a factor of deriving the learning rate of $P_{confidence}$, γ_P . $\gamma_{P_{dist_{mean}}}$ determined the range of γ_P for participants in a group. From the simulation, $\gamma_{P_{dist_{mean}}}$ for young adults is the highest, followed by old adults and people with Parkinson's disease. A higher $\gamma_{P_{dist_{mean}}}$ leads to a higher fixed range of $\gamma_{P_{dist}}$ for the given group. This set of higher $\gamma_{P_{dist}}$ range contributes to a set of γ_P with greater values.

Both $D_{dist_{mean}}$ and R_S are related to the resistance in proactive interference. Similar to $\gamma_{P_{dist_{mean}}}$, $D_{dist_{mean}}$ is the mean of the normal distribution of 'participants' for D_{dist} in a given group. D_{dist} is dimensionless and gives the distribution for participants in a group as a factor of deriving the exponential decay rate for proactive interference, D. $D_{dist_{mean}}$ determines the range of D for participants in a group. A lower $D_{dist_{mean}}$ in the given group led to a fixed range of D that is skewed towards a lower value. This implies that more participants in the group have a lower decay of proactive interference.

Accuracy Learning Curve

One category outcome, either category A, B, C, or D was chosen in each trial. For each trial, if the suggested category outcome from the model matched with the actual category of the stimulus, the model received feedback with the value 1, and 0 if the two categories did not match. The model only simulated the choice option of A and B for RB tasks and C and D for II tasks. The model does not support the cross selection of choice in different tasks (i.e., choosing option C or D in an RB task, or choosing option A or B in an II task). For each simulation, the categorization accuracy for each block of 100 trials was computed based on the match between the category outcome suggested by the model and the actual category of the stimulus for the trials within the blocks. The accuracy learning curves with 6 blocks of 100 trials were shown in Figure 10 with an excellent fit (RMSD of approximately 1.8).

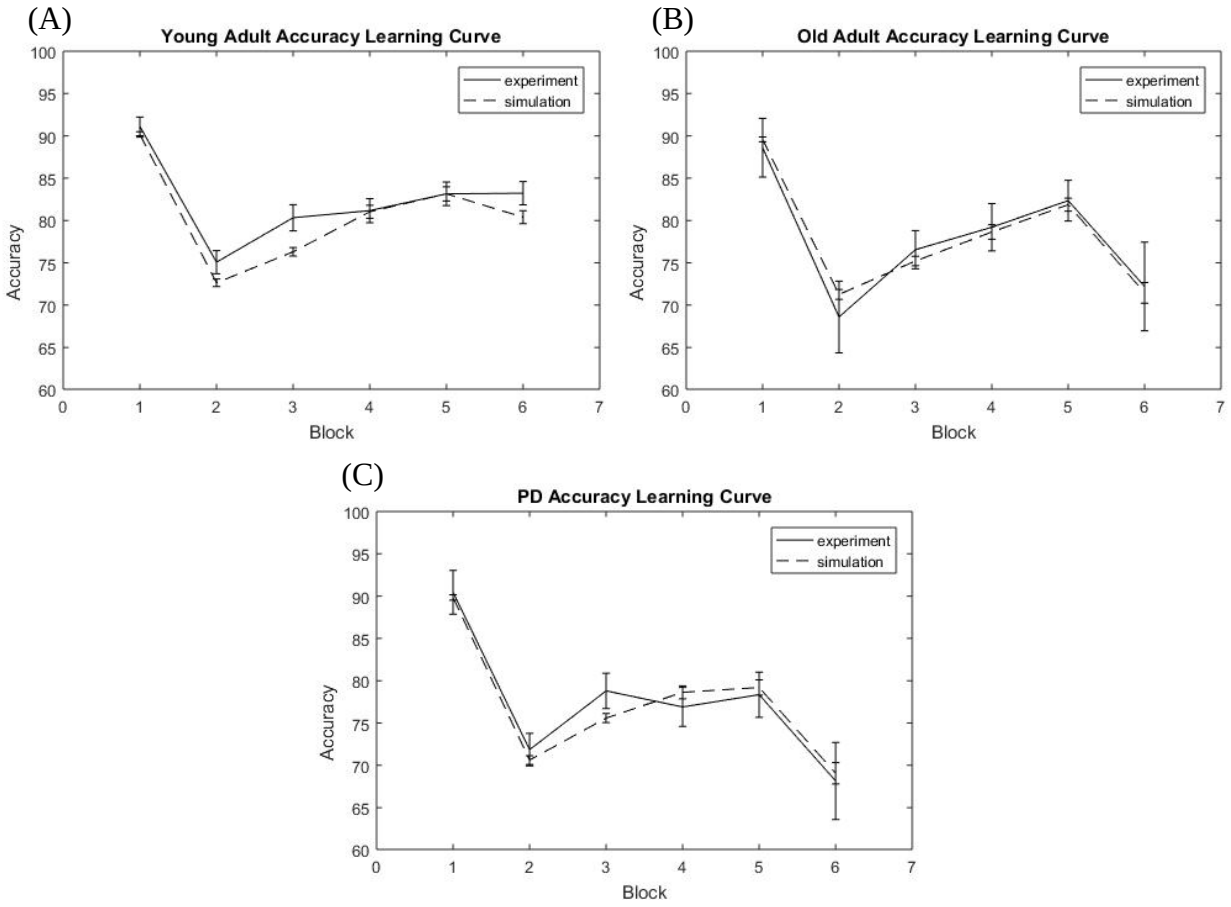

Figure 10. Mean accuracy per block in the experiment for (A) young adults, (B) old adults, and (C) Parkinson's disease patients. The solid line indicated data obtained from the simulation while the dashed line indicated data from the behavioral experiment conducted by Hélie and Fansher (2018). Error bars are the between-subject standard error of the mean.

Proportion of Switchers

Unlike the previous task simulation where switchers and non-switchers were simulated separately, the simulation of the trial-by-trial task switch experiment required simulation of mixed switchers and non-switchers in a group of participants. The proportion of switchers reflects the property of the participant group as the model's output by varying parameters. The experiment goes from a block of RB training task to four blocks of II training task, and finally a block of RB and II intermix tasks. To compute the proportion of switchers, only the last block of the experiment (the intermix trials) was taken into account. With the participants trained first in the RB and II tasks, we were interested in investigating the cause that affects the ability to switch from task to task for all three groups of participants.

Switchers were identified as 'participants' that could switch between the learning systems in a switching task. In the model, the switching task was associated with different learning system in obtaining the optimum performance (HT for RB task and P for II task). Thus, we compared the overall output response of the model and the responses of the learning systems for each trial in the last block of the experiment. In the II task, the majority of the overall response of the switchers should follow the response of P. In the RB task, the majority of the overall response of the switchers should follow the response of HT. If the 'participant' failed to do so, they were labeled as non-switcher. The proportion of switchers was obtained by computing the ratio of the total number of switchers in a given group to the total number of 'participants' in that group (i.e., 100). The proportion of switchers for the three groups is shown in Figure 11. The RMSD was 0.038.

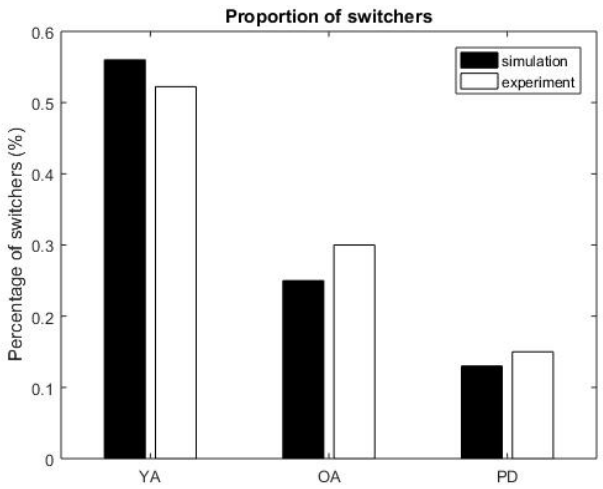

Figure 11. The proportion of switchers in each condition. Black bars indicated data from simulation while white bars indicated data from the behavioral experiment conducted by Hélie and Fansher (2018).

Accuracy Switch Cost

Accuracy switch cost was also computed based on the last block of the experiment: the intermix RB and II trials. Since it involves task switching, the first trial after a task switch was labeled a switch trial, while the remaining trials were labeled stay trials. The overall response of the trials was compared with the corresponding actual category of the stimulus for the same trials. This comparison gave the accuracy for each trial. The accuracies for the stay and switch trials were separated, and the mean for the accuracies for stay and switch trials were computed separately. The accuracy switch cost was obtained by subtracting the mean accuracy of the switch trials from the mean accuracy of the stay trials. The accuracy switch cost for switchers and non-switchers for all three groups of participants as shown in Figure 12. The RMSD was 0.015.

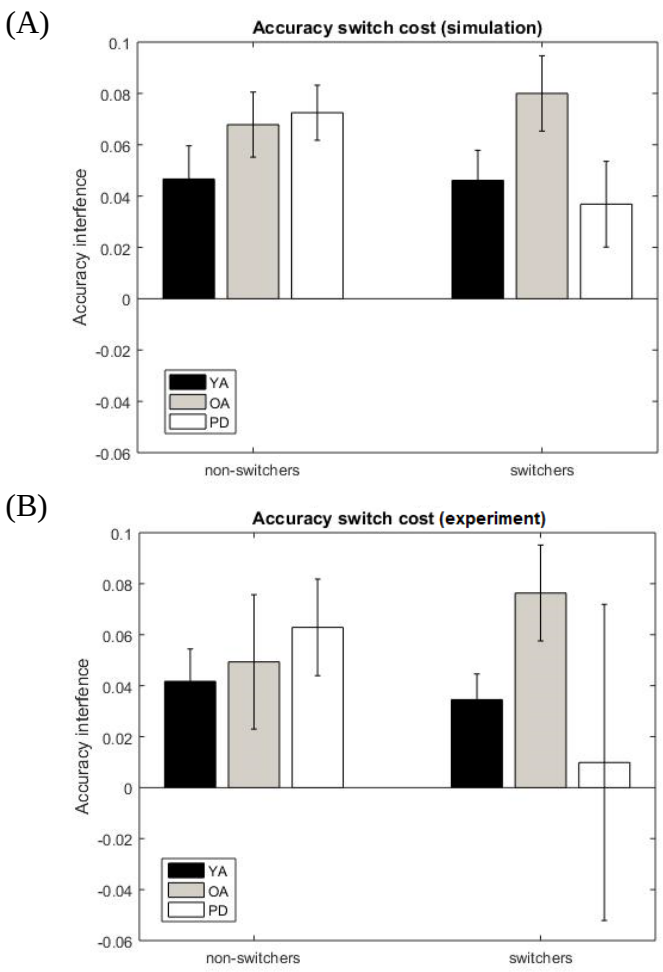

Figure 12. Accuracy switch cost for switchers and non–switchers in each group for (A) data obtained from simulation and (B) data obtained from the behavioral experiment conducted by Hélie and Fansher (2018). Black bars indicate accuracy interference for young adults, grey bars indicate accuracy interference for old adults, and white bars indicate accuracy interference for Parkinson's group. Error bars are the between-subject standard error of the mean.

In the simulation, when compared to the experiment from Hélie and Fansher (2018), a similar trend for accuracy switch cost for all three groups was observed. In young adults, the accuracy switch cost for switchers and non-switchers were about the same. Older adults had higher accuracy switch costs when compared to young adults for both switcher and non-switcher. Older adults switchers had higher accuracy switch cost when compared to the non-switchers of the same group. The accuracy switch cost of Parkinson's group was the highest among non-switchers, however, the accuracy switch cost of Parkinson's group was the lowest among switchers. The same reversal effect was observed in the experimental data from Hélie and Fansher (2018).

Discussion

With a greater chance of having a higher γ_P, young adults have a faster learning rate for P's confidence with positive reinforcement. P of people with higher γ_P gained confidence at a higher rate, and therefore increased the possibility for system switching at a higher pace. Groups with higher γ_P had a greater proportion of switchers as shown in Figure 11. The effect of higher γ_P was also reflected in the steeper increment of accuracy in the II training blocks in Figure 10. Since the majority of young adults have higher γ_P, they were faster in picking up the optimal strategy – procedural strategy as compared to the other older age groups.

Given the structure of reinforcement learning for γ_P and how it is situated in the model as a learning rate of a feedback function for P, the changes to γ_P may be dopamine-dependent. Past research has shown the association of cognitive flexibility with tonic dopamine levels (Price, Filoteo, & Maddox, 2009). In normal aging, the loss of dopamine-producing neurons is about 5% to 10% per decade of life (Karrer, Josef, Mata, Morris, & Samanez-Larkin, 2017). In addition to that, Parkinson's disease is associated with the accelerated death of dopamine-producing neurons in the substantia nigra pars compacta, which leads to the significantly low level of dopamine within the dorsal striatum (Hélie et al., 2012). The low $\gamma_{P_{dist_{mean}}}$ value in the simulation of category learning performance in people with Parkinson's disease might reflect the association of the group's performance with the low dopamine level in the dorsal striatum. Comparing to the healthy group of matching age, the tonic dopamine level was not reduced to the extent of the case in Parkinson's disease. Thus, the performance can be simulated with a higher $\gamma_{P_{dist_{mean}}}$ value which is dopamine-related. Similarly, young adults have the highest $\gamma_{P_{dist_{mean}}}$ value among the three groups.

In people with Parkinson's disease, PD-related neurochemical changes often cause impairment of several cognitive processes which include dysfunction in selective attention, working memory, and cognitive flexibility (Brown & Marsden, 1990). These impairments are especially observable in RB category learning. In category learning tasks, categorization starts with the process of rule generation, which requires the initial activation of the representations of one or more rules in working memory (Price et al., 2009). In the beginning, rule selection might be at random, and feedbacks may alter the activation of a rule representation in the working memory (Price et al., 2009). With greater positive reinforcement, the associated rule should be

maintained in the working memory and is protected from the interference of irrelevant rules of information (Price et al., 2009). Similarly, if the rule is not associated with recent positive reinforcement, the rule will be abandoned. In switching tasks, people with Parkinson's disease may have decent rule maintenance for the first few rules, however, errors increase with more rule or task switches (Price et al., 2009). Parkinson's disease is often associated with the excessive inhibition of thalamic projections to cortical regions (premotor and prefrontal regions) (DeLong, 1990; Scatton, Worms, Lloyd, & Bartholini, 1982). With the prefrontal dysfunction in Parkinson's disease, the patients are more likely to be susceptible to proactive interference, whereby additional inhibition is required to prevent previously reinforced rules from disrupting maintenance of the rules that are currently relevant (Feredoes, Tononi, & Postle, 2006; Rouleau, Imbault, Laframboise, & Bedard, 2001).

In episodic memory tasks, older adults gave more impaired performance when compared to the performance of the younger age group (Wahlheim, 2014). Similar to the Parkinson's disease deficit, one of the explanations for this performance impairment is that older adults are more susceptible to proactive interference or interferences from competing memories, where the memory of new information is hindered by related information that was learned earlier (Campbell, Hasher, & Thomas, 2010; Hasher & Zacks, 1988; Healey, Hasher, & Campbell, 2013). Thus, older adults have a greater chance of having this inhibition deficit that makes them less likely to suppress learned related information (Hasher & Zacks, 1988). In switching tasks, older adults might have a higher chance of having a recollection deficit of task sets which makes it harder to retrieve new information and at the same time, prevent previously learned related information from interfering (Hay & Jacoby, 1999). Similarly, this switching ability to abandon one mental set of information in favor of another could also be important in young adults in explaining the ability to switch between learning rules and even learning systems (Emery, Hale, & Myerson, 2008).

From the simulation, $D_{dist_{mean}}$ for young adults is the highest, followed by old adults and people with Parkinson's disease. This indicates that more young adults have higher exponential decay rates for proactive interference, D. Thus, interference from previously relevant information or rules dies down at a higher rate. Seemingly, a larger number of older adults and people with Parkinson's disease have a lower decay rate which indicates increased resistance to previously relevant information, therefore the two groups had a harder time switching to a 'new' rule that is more relevant to the current task.

On the other hand, the reverse trend was observed for R_S, whereby the R_S of young adults is lower than the older groups, and people with Parkinson's disease have a greater R_S. With R_S being the perceptual cue's decay effects on preSMA input, higher R_S indicates greater proactive interference effects in the case of task switching to the preSMA cells. In the intermix trials, task switching was frequent (a switch happened after a few trials), and rapid rule abandon from one task to another was required for optimal performance. The combination of lower D and higher R_S made switching of system from HT to P harder. The previously relevant information was harder to be suppressed in the 'new' task that requires a completely different set of rules for optimal performance. Thus, the accuracy performance as shown in block 6 of Figure 10 is lower. The number of switchers in such groups was lower compared to the groups with higher $D_{dist_{mean}}$ and lower R_S.

The varying level of proactive interference has a huge influence on the accuracy switch cost. D ranged from 0 to 0.99. In general, people with higher (0.9 to 0.99) and lower (0 to 0.1) D had lower accuracy switch costs. Proactive interference in people with high D diminished quickly after a task switch in the intermix trials. These 'people' can switch strategies and systems flexibly in the intermix trials. On the other hand, proactive interference in people with low D lingered for a long time. Thus, they might use the same strategies for the switch and most of the stay trials for the same task in the intermix trials. This may cause a lower accuracy in both switch and stay trials, and the difference yields a lower switch cost. Those with D that fell under the intermediate range (0.1 to 0.9) have a higher accuracy switch cost. The accuracy drop in the switch trials are far more significant than the accuracy drop in the stay trials. Along with the distribution of γ_P and the level of R_S, the distribution of D sort 'participants' into switchers and non-switchers in the intermix trials, which is as shown in Figure 12.

CONCLUSION

We were interested in building a computational model that simulated a neurobiological circuit that facilitates learning system switching in categorization tasks. The model was designed based on the hyperdirect pathway of the cortico-basal ganglia network. It incorporated the Izhikevich's firing model for the neuronal cells in the brain area involved. HT and P were able to generate their responses for a specific stimulus. One response for categorization would be selected as the output of the model.

The model was applied to two situations to perform different category learning tasks. The model was able to reproduce learning curves, and predict the proportion of switchers for different age groups of participants and the accuracy switch cost in trial-by-trial switching tasks. The next section is a review of the most important features and accomplishments of the model.

Key Feature

Different from other available category learning models, the model describes the biological learning system switching mechanism from HT to P in category learning tasks. Most of the available category learning models highlighted the mechanism of response generation and selection. The emphasis on the neurobiological simulation enables a deeper understanding of how switching between two learning systems works.

The model simulated system switching facilitated by the hyperdirect pathway in the basal ganglia. The brain areas involved are preSMA, STN, and PMd. HT and P were modeled as black boxes that gave responses depending on the respective probability accuracy in a certain task. The responses from the two systems were independent of each other.

The system switching was depicted as the reduction of inhibition of the response of P into PMd units. STN acts as the gate for the P's response to entering the PMd units. If the activity of P is lower than the inhibition of STN, P's response is prevented from entering the PMd units, however, if the activity of P is higher than that of STN, the response of P is allowed into the PMd units.

Accomplishment

Our model was able to show learning system switching through the reduction in STN's inhibition of P's response into PMd units. System switching was seen when the following criteria were met:

1. Low confidence in HT:

   If an RB approach was not the optimal strategy in the category learning task, when HT failed to give accurate responses over time, the confidence in HT decreased with increasing inaccurate attempts at categorizing the stimulus. This was fed into the preSMA as feedback which lowered the activation of preSMA and thus reducing the activity in STN. With reduced activity, P's response's inhibition to the PMd units by the STN faded, allowing the response of P to enter the PMd units. This increased the chance of selecting the response of P as the overall response. This process showed system switching when HT (RB strategy) failed and P took over to respond with a procedural strategy.

2. High confidence in P:

   With the success of categorizing the stimuli, after numerous trials, P gained enough confidence whereby its activity exceeded the inhibition emitted by STN. This allowed P's response to be transmitted to PMd units for action selection. This process showed system switching from HT to P after a series of training, even if the RB strategy is the optimal strategy for the task. Also, the model showed the difficulty in switching back to HT from P after extensive trials of performing a certain task. It showed the domination of P after considerable training and learning.

System switching is speculated regardless of the presence of perceptual cues even in the trial-by-trial switching tasks if the two criteria aforementioned were met. The ability to switch between systems was seen when parameters were adjusted to cater to low confidence in HT and high confidence in P as task demanded.

In addition, the model was able to show how people of the same age group, and different age groups, differed in terms of the capability to switch systems or strategies in a category learning

task. The model was tested against experimental results conducted and obtained from published data. We fitted the model to experimental results from a young adult group in a category learning task. An II categorization task was simulated.

On each trial of the simulation, each system produced its own categorical decision as the response to the stimulus. The model had to generate a single output for the categorization at the end of each trial. Through the simulation, the initial use of an RB strategy was observed when the output of the model followed the responses of HT. The model was fit to two different participant groups from a behavioral experiment. Participants who were able to switch between category learning strategies, namely the switchers, were modeled to switch from an RB strategy to a procedural strategy. The model's output would eventually follow the responses of P at the end of each simulation. Participants who were stuck with the suboptimal strategy, namely the non-switchers, were modeled to stick to the RB strategy throughout the simulation. The model's output followed the response of HT throughout the simulation. We were able to simulate this difference in the capability of switching by adjusting the parameters that governed the learning rate of P's confidence and the accuracy probability of HT of giving a correct response for a given trial. Switchers were found to have a higher learning rate of P's confidence and a higher accuracy probability of HT. Both of these parameters aided in satisfying the switching criterion which geared towards a high P confidence. High ACC_{HT} in switchers indicated switchers produced more accurate responses with RB strategies as compared to the non-switchers, especially at the beginning of the task. However, this leads to high HT confidence. To overcome the switching inhibition, high γ_P sped up the rate of increment in $P_{confidence}$ when P gave accurate responses, especially in II tasks whereby procedural strategy is the optimal strategy.

Next, we were able to fit the model to participants from different age groups in a category learning task that involved a trial-by-trial switch of category task. We were able to show how aging affects categorization learning system switching. Our model was able to reproduce the accuracy learning curve, switch cost, and proportion of switchers that were shown in the experimental results. Through our model, we simulated the aforementioned results for young adults, old adults, and people with Parkinson's disease. The difference in the switching capabilities in the three groups of participants was simulated by adjusting the parameters that governed the learning rate of P's confidence and the decay in proactive inhibition. We found that fewer participants in the older age group (old adults and people with Parkinson's disease) have a high learning rate of P's

confidence which makes system switching hard. Furthermore, more participants in the older age group suffered from proactive inhibition interference when tasks switched frequently. Proactive interference makes it difficult to switch strategies (within RB strategies or from RB strategies to procedural strategies) when the tasks demanded different strategies as the optimal strategies.

Implications and Predictions

The model was able to show the significance of the basal ganglia's hyperdirect pathway in facilitating the switch between learning systems. At the beginning of a task, the model showed the domination of HT's response as the overall response in categorizing a stimulus. This domination occurred due to the inhibition of P's response to the PMd units by the STN. The reduction in the inhibition allowed system switching.

From the simulations, with the intact hyperdirect pathway, the differences in switching capabilities can be shown with the deficiency in the external systems, such as the individual learning system and prefrontal cortex. The influence of P's confidence's learning rate was significant in both within-group and between-group switching capabilities. This suggests the importance of the effect of dopamine in facilitating system switching. A lower level of dopamine within the dorsal striatum might be the key to the poorer switching capabilities of non-switchers from young adult groups as well as older adult groups.

Also, the ability to switch between the learning system in the simulation was affected by proactive interference's decay rate and its influence on the preSMA cell. From the trial-by-trial task-switching experiment, greater proactive interference impact prevented the change of strategies, as well as the change in the learning system used, as tasks switched frequently. This showed the importance of working memory attrition in strategy (Fleischer & Hélie, 2020) and learning system switching when shifting from one task to another.

Extensions, Improvements, and Future Works

The model was able to match and predict system switching through neurobiological circuits but it is by no means perfect. The model lacks the ability to simulate accurate response time for each cell. If the model can match the response time, we will be able to investigate the time-related switch cost that is accompanied by the trial-by-trial task switch. From Hélie and Fansher (2018),

switch cost was more prominent in response time as compared to just measuring accuracy. The result could give a better reflection of the different switching capabilities of the different groups of participants.

With our focus on the switching mechanism, while it is convenient to simulate the individual learning system as black boxes, the simulation of the response generation of the learning systems was restrained. The black boxes generated output based on a single source of accuracy probability (one for each task and learning system). However, even in a task, the accuracy probability should vary throughout the session. Initially, without grasping a rule, participants should select the category by guessing or by chance. Over time, more rules are generated and each rule has its accuracy probability. The model can be improved by operating under a maximum accuracy probability for each task. Depending on the accuracy feedback of the learning system, its accuracy probability can be adjusted from guessing (i.e., 50% for a two-choice category task) to the maximum value.

The model cannot perform strategy switch in un-cued task switching. Simulating an un-cued switching task is important to model experiments like an internally-cued categorization system-switching paradigm. In the trial-by-trial switching task, the model relied on environmental perceptual cues to indicate switched tasks. The accuracy probability depended on the cue for a task switch. Without the cues, it would be difficult for the model to cope with the change in the task, especially when the learning system had no feedback-dependent accuracy probability adjusting ability.

With the implication of significant regions or systems in switching capabilities, we can run participants in category experiments that requires varying working memory adaptation. For instance, buffers can be added in between task switches to see how proactive interference decay affects switching (Fleischer & Hélie, 2020). These experiments should be carried out to re-validate the influence of the manipulative parameters in the model in system switching.

The model can also broaden into animal studies. Although some aspects of the model have to be adjusted to rescale the cognitive limitations of animal subjects, the idea of the hyperdirect pathway can still be implemented to simulate the system switching capabilities of other primates. This would allow us to match single-cell recordings in the affected brain areas of the animals.

REFERENCES

Aron, A. R., Herz, D. M., Brown, P., Forstmann, B. U., & Zaghloul, K. (2016). Frontosubthalamic circuits for control of action and vognition. *Journal of Neuroscience*, *36*(45), 11489-11495.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442-481.

Ashby, F. G., & Casale, M. B. (2002). The cognitive neuroscience of implicit category learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 109-142). Philadelphia, PA: John Benjamin Publishing Company.

Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of Learning and Memory, 94*, 1-12.

Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. *The Psychology of Learning and Motivation, 46,* 1-36.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review*, *114*(3), 632-656.

Ashby, F. G., & Helie, S. (2011). The neurodynamics of cognition: A tutorial on computational cognitive neuroscience. *Journal of Mathematical Psychology*, *55*(4), 273-289.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 65-87). Cambridge, England: Cambridge University Press.

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 157-188). Cambridge, MA: Elsevier Academic Press.

Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience, 20*, 505-512.

Banino, A., Caswell, B., & Kumaran, D. (2018). Vector-based navigation using grid-like representations in artificial agents. *Nature*, *557*, 429-433.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). R. Wyer & T. Srull. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brown, R., & Marsden, C. (1990). Cognitive function in Parkinson's disease: From description to theory. *Trends in the Neuroscience, 13*, 21-29.

Buesing, L., Bill, J., Nessler, B., & Maass, W. (2011). Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology, 7*(11), e1002211.

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science*, *21*, 399-405.

Cantwell, G., Riesenhuber, M., Roeder, J., & Ashby, F. G. (2017). Perceptual category learning and visual processing: An exercise in computational cognitive neuroscience. *Neural Networks*, *89*, 31-38.

Cohen, N. J., Eichenbaum, H., Deacedo, B. S., & Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. *Annals of the New York Academy of Sciences, 444*, 54-71.

Crossley, M. J., & Ashby, F. G. (2015). Procedural learning during declarative control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1388-1403.

Crossley, M. J., Helie, S., Roeder, J. L., & Ashby, F. G. (2018). Trial-by-trial switching between procedural and declarative categorization systems. *Psychological Research*, *82*(2), 371-384.

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (2001). The role of the striatum and hippocampus in planning: A PET activation study in Parkinson's disease. *Brain*, *124*, 1020-1032.

DeLong, M. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neurosciences*, *13*, 281-285.

Eichenbaum, H. (1999). Conscious awareness, memory and the hippocampus. *Nature Neuroscience*, *2*, 775-776.

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain.* Oxford, England: Oxford University Press.

Emery, L., Hale, S., & Myerson, J. (2008). Age differences in proactive interference, working memory, and abstract reasoning. *Psychology and Aging*, *23*(3), 634-645.

Erickson, M. A. (2008). Executive attention and task switching in category learning: Evidence for stimulus-dependent representation. *Memory & Cognition*, *36*(4), 749-761.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*(2), 107-140.

Ermentrout, G. B., & Kopel, N. (1986). Parabolic bursting in an excitable system coupled with a slow oscillation. *Society for Industrial and Applied Mathematics Journal on Applied Mathematics*, *46*, 233-253.

Evans, J. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454-459.

Feredoes, E., Tononi, G., & Postle, B. (2006). Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proceedings of the National Academy of Science*, *103*, 19530-19534.

Filoteo, J. V., Maddox, W. T., Ing, A. D., Zizak, V., & Song, D. D. (2005). The impact of irrelevant dimensional variation on rule-based category learning in patients with Parkinson's disease. *Journal of International Neuropsychological Society*, *11*(5), 503-513.

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (2005). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology*, *19*(2), 212-222.

Fleischer, P., & Hélie, S. (2020). A unified model of rule-set learning and selection. *Neural Networks*, *124*, 343-356.

Fletcher, P., Shallice, T., Frith, C., Frackowiak, R., & R., D. (1998). The functional roles of prefrontal cortex in episodic memory. II. Retrieval. *Brain*, *121*, 1249-1256.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the USA*, *103*(31), 11778-11783.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*, 51-72.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning and Motivation: Advances in Research and Theory*, *22*, 193-225.

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology and Aging*, *14*, 122-134.

Healey, M. K., Hasher, L., & Campbell, K. L. (2013). The role of suppression in resolving interference: Evidence for an age-related deficit. *Psychology and Aging*, *28*, 721-728.

Heindel, W. C., Salmon, D. P., Shults, C. W., Walicke, P. A., & Butters, N. (1989). Neuropsychological evidence for multiple implicit memory systems: A comparison of Alzheimer's, Huntington's, and Parkinson's disease patients. *Journal of Neuroscience*, *9*, 582-587.

Hélie, S. (2017). Practice and preparation time facilitate system-switching in perceptual categorization. *Frontiers in Psychology*, *8*, 1964.

Helie, S., Ell, S. W., & Ashby, F. G. (2015). Learning robust cortico-cortical associations with the basal ganglia : An integrative review. *Cortex*, *64*, 123-135. https://doi.org/10.1016/j.cortex.2014.10.011

Hélie, S., Ell, S. W., Filoteo, J. V., & Maddox, W. T. (2015). Criterion learning in rule-based categorization: Simulation of neural mechanism and new data. *Brain and Cognition*, *95*, 19-34.

Helie, S., & Fansher, M. (2018). Categorization system-switching deficits in typical aging and Parkinson's disease. *Neuropsychology*, *32*, 724-734.

Hélie, S., Paul, E. J., & Ashby, F. G. (2012). Simulating the effects of dopamine imbalance on cognition: From positive affect to Parkinson's disease. *Neural Networks*, *32*, 74-85.

Hélie, S., Turner, B. O., Crossley, M. J., & Ell, S. W. (2017). Trial-by-trial identification of categorization strategy using iterative decision bound modeling. *Behaviour Research Method*, *49*, 1146-1162.

Hikosaka, O., & Isoda, M. (2010). Switching from automatic to controlled behavior: Cortico-basal ganglia mechanisms. *Trends in Cognitive Sciences*, *14*(4), 154-161.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, *117*, 500-544.

Hoppensteadt, F. C., & Izhikevich, E. M. (1997). *Weakly connected neural networks*. New York, NY: Springer-Verlag.

Houk, J. C., Adamas, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generates and uses neural signals that predict reinforcement. In J. C. Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249-270). Cambridge, MA: MIT Press.

Izhikevich, E. M. (2000). Neural excitability, spiking, and bursting. *International Journal of Bifurcation and Chaos, 10*, 1171-1266.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks, 14*(6), 1569-1572.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: The geometryof excitability and bursting*. Cambridge, MA: The MIT Press.

Izhikevich, E. M. (2010). Hybrid spiking models. *Philosophical Transactions of the Royal Society A, 368*, 5061-5070.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience, 14*, 3775-3790.

Joel, D., & Weiner, I. (1997). The connections of the primate subthalamic nucleus: Indirect pathways and the open-interconnected scheme of basal ganglia–thalamocortical circuitry. *Brain Research Reviews, 23*, 62-78.

Kahneman, D. (2012). *Thinking, fast and slow*. London, England: Penguin Books.

Karrer, T. M., Josef, A. K., Mata, R., Morris, E. D., & Samanez-Larkin, G. R. (2017). Reduced dopamine receptors and transporters but not synthesis capacity in normal aging adults: A meta-analysis. *Neurobiology of Aging, 57*, 36-46.

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review, 110*, 316-339.

Kenner, N. M., Mumford, J. A., Hommer, R. E., Skup, M., Leibenluft, E., & Poldrack, R. A. (2010). Inhibitory motor control in response stopping and response switching. *Journal of Neuroscience, 30*(25), 8512-8518.

Knowlton, B. J. (2002). The role of the basal ganglia in learning and memory. In S. D. Squire LR (Ed.), *Neuropsychology of memory* (pp. 143-153). New York, NY: Guilford Press.

Kriegeskorte, N., & Douglas, P. K. (2019). Cognitive computational neuroscience, *Nature Neuroscience, 21*(9), 1148-1160. https://doi.org/10.1038/s41593-018-0210-5.

Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron, 58*, 451-463.

Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Backpropagation for deep learning. *Nature Communications, 7*, 1-10. https://doi.org/10.1038/ncomms13276

Lim, L. X., & Hélie, S. (2019). Exploration and exploitation reflect system-switching in learning. In *CogSci 2019* (pp. 2154-2160).

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, *66*(3), 309-332.

Maddox, W. T., Pacheco, J., Reeves, M., Zhu, B., & Schnyer, D. M. (2010). Rule-based and information-integration category learning in normal aging. *Neuropsychologia*, *48*(10), 2998-3008.

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*, 1609-1617.

Michmizos, K. P., & Nikita, K. S. (2011). Local field potential driven Izhikevich model predicts a subthalamic nucleus neuron activity. In *33rd Annual international conference of the IEEE engineering in medicine and biology society* (pp. 5900-5903). Boston, MA: IEEE. https://doi.org/10.1109/IEMBS.2011.6091459

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology, 50*, 381-425.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In *Neurobiology of learning and memory* (pp. 65-77). New York, NY: Guilford.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*, 1936-1947.

Moody, T. D., Bookheimer, Z. V., & Knowlton, B. J. (2004). An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. *Behavioral Neuroscience, 118*, 438-442.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage*, *36*(3-3), T155-T163.

Neftci, E. O., & Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence, 1*, 133-143. https://doi.org/10.1038/s42256-019-0025-4

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., & Al, E. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17*, 37-43.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science2*, *304*, 452-454.

Parent, A., & Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. II. The place of the subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Research Reviews*, *20*, 128-154.

Paul, E. J., & Ashby, F. G. (2013). A neurocomputational theory of how explicit learning bootstraps early procedural learning. *Frontiers in Computational Neuroscience*, *7*.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*, 245-251.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, *13*, 564-574.

Price, A. C., Filoteo, J. V., & Maddox, W. T. (2009). Rule-based category learning in patients with Parkinson's disease. *Neuropsychologia*, *47*(5), 1213-1226.

Rabi, R., & Minda, J. P. (2014). Rule-based category learning in children: The role of age and executive functioning. *PLoS One*, *9*(1), e85316.

Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. *Journal of Neurophysiology*, *30*(5), 1138-1168.

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning and Memory*, *1*, 217-229.

Rescorla, R. A., & Wagner, A. R. (1972a). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton- Century-Crofts.

Rescorla, R. A., & Wagner, A. R. (1972b). *Classical Conditioning II: Current Research and Theory*. (W. F. Black, A. H. & Prokasy, Ed.). Appleton- Century-Crofts, New York.

Rolls, E. T. (2016). *Cerebral cortex: Principles of operation*. Oxford, United Kingdom: Oxford University Press..

Rolls, E. T., & Deco, G. (2016). Non-reward neural mechanisms in theorbitofrontal cortex. *Cortex*, *83*, 27-38.

Rouleau, I., Imbault, H., Laframboise, M., & Bedard, M. (2001). Pattern of intrusions in verbal recall: Comparison of Alzheimer's disease, Parkinson's disease, and frontal lobe dementia. *Brain & Cognition*, *46*, 15-17.

Scatton, B., Worms, P., Lloyd, K., & Bartholini, G. (1982). Cortical modulation of striatal function. *Brain Research*, *232*(2), 331-342.

Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. T. & F. I. M. Craik (Ed.), *Oxford handbook of memory* (pp. 627-643). New York, NY: Oxford University Press.

Schierwagen, A. (2009). Mathematical and computational modeling of neurons and neuronal ensembles. In *Computer Aided Systems Theory - EUROCAST 2009* (pp. 159-166). Berlin, Germany: Springer.

Schroeder, J. A., Wingard, J., & Packard, M. G. (2002). Post-training reversible inactivation of the dorsal hippocampus reveals interference between multiple memory systems. *Hippocampus*, *12*, 280-284.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.

Slusarz, P., & Sun, R. (2001). The interaction of explicit and implicit learning : An integrated model. In *Cognitive Science Society*.

Smith, E. E., & Grossman, M. (2007). Multiple systems of category learning. *Neurosceince and Biobehavorial Reviews, 32*(2), 249-264.

Smith, E. E., Patalano, A., & Jonides, I. (1998). Alternative mechanisms of categorization. *Cognition*, *65*, 167-196.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*, 171-177.

Sun, R. (2002). *Duality of the mind*. Mahwah, NJ: Lawrence Erlbaum Associates.

Taylor, J. A., & Ivry, R. B. (2012). The role of strategies in motor learning. *Annals of the New York Academy of Sciences*, *1251*(1), 1-12.

Tsujii, T., & Watanabe, S. (2009). Neural correlates of dual-task effect on belief-bias syllogistic reasoning: a near-infrared spectroscopy study. *Brain Research*, *1287*, 118-125.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, *8*, 198-204.

Wahlheim, C. N. (2014). Proactive effects of memory in young and older adults: The role of change recollection. *Memory & Cognition*, *42*, 950-964.

Wilson, C. G., Nusbaum, A. T., Whitney, P., & Hinson, J. M. (2018). Age-differences in cognitive flexibility when overcoming a preexisting bias through feedback. *Journal of Clinical and Experimental Neuropsychology*, *40*(6), 586-594.

Winlow, W. (1990). No Title. In M. U. Press (Ed.), *Proceedings of the 22th Annual Newfoundland Electrical and Computer Engineering Conference*.

Yang, J., & Li, P. (2012). Brain networks of explicit and implicit -earning. *PLOS ONE*, *7*(8), e42993.