10-14-2020 10:00 AM

The Impact of Levodopa Administration on Learning from Short-term and Long-term Action Consequences: a Paradigm Validation.

Masood Rezaei, *The University of Western Ontario*

Supervisor: Morton, J. Bruce, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Neuroscience

Abstract

Behavioral and neuroimaging studies have identified two valuation systems in the human brain for controlling behavior known as model-free (MF) and model-based (MB). MF is based on immediate evaluation and MB is based on long-term evaluation of the outcome of our decisions. Previous studies suggest that dopamine baseline activity may play an important role in the balance between the two systems and determine how they compete or interact in controlling our actions. The overarching aims of this study is to investigate the impact of levodopa administration on learning from immediate and long-term action consequences, and to dissociate the role of striatal subregions in learning and action selection. Here, an fMRI fast-event related paradigm is designed and validated which enables to computationally model the integration of MF and MB learning and decision making on both behavioral and neural levels.

Keywords

model-free, reinforcement learning, model-based, Bayesian, dopamine, Levodopa, learning, decision making.

Summary for Lay Audience

Humans use two strategies for evaluating decisions at hand and choosing the most appropriate action with better payoff. These strategies arise from two separate valuation systems in the brain called model-free (MF) and model-based (MB) systems that often compete, but other times interact with each other to control our behavior. MF behavior involves considering the immediate reward even though the long-term consequences may not be favorable (e.g., unhealthy food consumption). On the other hand, MB behavior involves considering long-term outcome even if an action is not associated with short-term reward such as working hard for distant goals in the future. Dopamine is a neurotransmitter in the brain that is involved in several cognitive functions such as reward-based learning and action selection. Previous studies have shown that administering levodopa, the primary medication for Parkinson's disease, would increase the dopamine availability in the brain and would bias the choices toward the long-term goals. Here, I have designed and validated a paradigm that is the foundation for a pharmacological manipulation study of dopamine using levodopa combined with functional magnetic resonance imaging (fMRI) to investigate the balance between MF and MB systems and to identify the neural correlates of learning from short-term and long-term action consequences.

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. J. Bruce Morton for his countless support in every step of this project. I truly appreciate the time and effort you put to keep the wheels of this project turning. Also, I would like to thank my advisors Dr. Penny MacDonald, Dr. Bobby Stojanoski, Dr. Kasey Van Hedger and Dr. Julio Martinez-Trujillo. You have helped me with insightful advises and thoughts.

I would like to thank Ms. Bea Goffin our lab manager and all of my lab members in the Cognitive Development and Neuroimaging Laboratory. Thank you all for your support, mentorship and times that we shared. I learned many lessons from you. Also, I would like to thank Macdonald lab members who helped me with important aspect of this study.

Finally, I would like to thank my family and friends for all of their supports. Special thanks to my lovely wife Mohadese Khosravi Najafabadi who is my cheerleader in every challenge I face in life. Thank you for giving me motivation for success.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

1 Introduction

As human beings, we learn from our interaction with environment to maximize rewards and minimize losses. Making the best choices relies on many factors such as availability of information, valuation of immediate and distant rewards/punishments, and also how we preferably weigh the short-term and long-term consequences of our actions (Fischer et al., 2017). In some cases, we do things such as strenuous physical exercise, knowing that despite the short-term discomfort, such activities are likely to be beneficial in the long run. In other cases, such as recreational drug use, our actions are driven by short-term pleasures and occur even though such actions may have detrimental consequences in the future. During last decades, scientists have tried to shed light on how we make decisions in the changing environment. By integrating knowledge from psychology, neuroscience, computer science and economics, a number of fundamental questions about the underlying mechanism of learning and decision making have been answered and has led us to deeper questions regarding the complexity of the system.

Behavioral and neuroimaging studies have provided evidence with regard to the presence of two systems in the brain for guiding actions. These systems are often named with different terminologies across fields such as reflexive versus reflective, retrospective versus prospective, automatic versus deliberative, habitual versus goal-directed, or model-free (MF) versus model-based (MB) (Dayan & Berridge, 2014). MF learning is based on evaluation of immediate outcomes through trial and error. It drives habitual behavior arising from slow accumulation of rewards through cached estimate of iterative updates of expectation. Basically, a summary of experience associated with a situation or an action provides information that can act as a basis for future choices. This type of information is fast and computationally simple but at the cost of inflexibility (Daw et al., 2005). MB learning on the other hand drives goal-directed behavior based on deliberative and prospective consideration of the future outcomes associated with a situation or an action. Unlike MF, MB learning chains every action with a corresponding outcome and keeps the history of experience from every action in a decision tree. Although it is

computationally demanding in terms of energy, time, and memory, it is less susceptible to error (Daw et al., 2005). Previous studies have shown that the two systems often compete or cooperate in guiding our actions (Balleine & O'Doherty, 2010).

Generally, dopamine (DA) has an important role in encoding short-term and long-term values as well as integration of the two systems. Previous studies show that DA baseline activity is a major contributing factor determining the balance between MF and MB behavior (Wunderlich et al., 2012; Deserno et al., 2015). However, our current understanding of the exact role of DA in the arbitration between the two systems is very limited. Additionally, there is not enough evidence showing where MF and MB information are integrated in the brain. Generally, there is a consensus view that ventral striatum (VS) encodes MF learning (Huang et al., 2020). Recently, one study has reported an integrated representation of both MF and MB learning along the striatum. It specially shows the involvement of dorsal striatum (DS) in processing MB learning (Fischer et al., 2017) which is only reported in animal literature (Ballein & O'Doherty, 2010). Fischer et al. (2017) suggests that value-related learning in the striatum is not only limited to the VS. Also, it opens the possibility of an integrated representation of both learnings and cooperation of the two systems along the striatum. Our current understanding of the role of DA in MF and MB learning and decision-making as well as our knowledge of brain areas related to the processing of these information roots in the animal studies. Therefore, in the following, some basics of neurophysiology of reward-based learning and decision making is reviewed. Then, the animal models of habitual and goal-directed behavior as well as the transition from animal to human studies is discussed. Finally, the integration of MF and MB systems and the pharmacological manipulation of DA using levodopa is reviewed which forms the foundation of the current study.

1.1 Dopamine and Reward-based Learning

DA is a neurotransmitter that is extensively linked to different aspects of reward-based learning and decision making (Bayer & Glimcher, 2005). In 1990s, a major shift happened in our understanding of the role of DA in reward-based learning. Schultz et al.

(1992) found that the midbrain DA neurons encode unexpected outcome associated with a stimulus (as well as cues leading to unexpected outcome) through a brief burst of activity. This was an important observation as it led to formation of reward prediction error (RPE) theory. RPE is an estimation of the difference between the value of experienced reinforcer (reward or punishment) and what was expected (Glimcher & Fehr, 2013). This estimation is used by an organism to maximize future rewards. Several findings suggest that DA bursts of activity in the VS matches the RPE signal (Schultz et al., 1992; Berke, 2018) and since RPE is naturally a learning signal, it became apparent that DA activity plays a fundamental role in learning (Berke, 2018). Later, Steinberg et al. (2013) found a causal link between DA activity and RPE by using optogenetic technique. A further investigation was to dissociate encoding of reward and punishment. Bayer and Glimcher (2005) used single neuron recording and found that the DA neurons respond to positive error through a brief burst of firing. In a later study, Bayer et al. (2007) further dissociated negative and positive prediction error and found that the negative prediction error is encoded by a brief pause in firing of dopaminergic neurons.

The response profile of midbrain DA neurons can be characterized in two ways which are called phasic and tonic. On cellular level, phasic DA activity is a fast and transient stimulation of post-synaptic neuron arising from a behavior or an environment (Crockett & Fehr, 2014). Phasic DA activity encodes action learning by facilitating intended action through promoting long-term potentiation in D1 receptors and obstructing unintended actions through promoting long-term depression in D2 receptors in the striatum (Maia & Frank, 2011). On the other hand, tonic DA activity is a slow constant stimulation of the DA baseline level which maintains back-ground extracellular DA concentration (Crockett & Fehr, 2014). Tonic DA activity can influence the sensitivity of post-synaptic neuron in the detection of phasic bursts of firing (Grace, 1991). The tonic DA activity was long believed to convey the signal of motivation (Salamone & Correa, 2012) and action selection (Maia & Conceição, 2017). However, this idea has recently been challenged and some believe that the interpretation of tonic DA activity as the motivational signal needs to be reconsidered (for a review: Berke, 2018).

1.2 Dopamine and Striatum

DA neurons are mainly located in ventral tegmental area (VTA) and substantia nigra pars compacta (SNpc). The axons of these neurons project mainly to the frontal cortex (orbital frontal cortex (OFC), dorsal lateral prefrontal cortex (dlPFC), and ventral lateral prefrontal cortex (vlPFC)) and the striatum (Schultz, 1999). The striatum is the primary input of several subcortical structures collectively called the basal ganglia. Beside the striatum, the basal ganglia include Globus pallidus (external and internal), Subthalamic nucleus (STN) and Substantia nigra (pars compacta and reticulata). The striatum receives glutamatergic input from nearly the entire neocortex, project to different intrinsic nuclei of the basal ganglia, and at the end, send information back to the cortex through thalamus (Alexander et al., 1986). Additionally, it receives input from the midbrain dopaminergic neurons and is an important site in the processing of reward-based learning and action selection as discussed above (for a review of functional anatomy: Maia & Frank, 2011).

The striatum can further be subdivided into the DS and the VS. The DS includes caudate nucleus and putamen. From a functional standpoint, the DS can be further subdivided into dorsolateral and dorsomedial striatum and each have separate connections to the dopaminergic neurons in midbrain and are thought to play different roles in the process of learning and action selection (for a review: Cox & Witten, 2019). On the other hand, the VS comprises of Nucleus accumbens which has core and shell regions. All of the striatal subregions are involved in different cortico-basal ganglia-thalamo-cortical loops but can receive and relay information between the sub circuits as well (Hans, 2011). These loops and their connections with the dopaminergic system is recognized as the brain circuitry for acquisition and regulation of habitual and goal-directed behavior.

It has been more than a century from the classic Thorndike (1911) research that scientists have been investigating the underlying mechanism of associative learning and adaptive behavior. Today, we are building on the literature that was initially adapted from animal studies to model habitual and goal-directed behavior. Below, related animal studies are reviewed which led us to investigate the integration of MF and MB behavior in humans.

1.3 Animal Models of Habits and Goals

It had been long believed that control over adaptive behavior relies solely on the regions of PFC. However, studies have shown that the striatum plays an important role in action learning and the control of executive functions (Tanaka et al, 2006). Using behavioral tasks that are comparable between human and rats, studies have found homologous regions for processing two types of learning. In particular, habit formation is governed by the activity of prelimbic cortex and dorsomedial striatum in rats (homologous to medial PFC, medial OFC and anterior caudate nucleus in humans). On the other hand, goal-directed behavior relies on the neural activities in dorsolateral striatum in rats (homologous to posterior lateral putamen in humans) (Ballein & O'Doherty, 2010).

These studies are based on the idea that habitual and goal directed behaviors rely on the distinct associations between stimulus, response and outcome (S-R-O). Specifically, goal-directed behavior arises from making association between the representation of "response" with its corresponding "outcome" (R-O). On the other hand, habitual behavior is driven by the association between the representation of stimulus with a particular response (S-R) regardless of the magnitude of the outcome. In this regard, action control in goal-directed behavior is based on evaluating action consequences, whereas in habitual behavior it is based on the presence of the stimulus itself (Ballein & O'Doherty, 2010). Based on the animal literature, goal-directed and habitual behaviors are distinguishable in two ways: one is sensitivity to outcome devaluation and the other is sensitivity to contingency degradation.

During the first initial trials in a behavioral paradigm in rats (e.g., instrumental conditioning), action selection relies on evaluating the action consequences after making a response (R-O association). Here, an action that is associated with a rewarding outcome is reinforced and the likelihood of repeating the same action is increased. During this stage, the behavior is goal-directed and is sensitive to outcome devaluation, i.e., the choice preference will change if an action is no longer associated with a reward. As an illustration, Adams and Dickinson (1981) trained rats to press a lever for receiving a reward. Devaluing the reinforcer by pairing it with an illness (lithium chloride injection)

resulted in a significant reduction in the lever pressing, suggesting that rats are capable of goal-directed behavior and are adapting their behavior according to the outcome. Subsequent studies showed that the execution of goal-directed or habitual behavior depends on the amount of training. In particular, Adams (1982) showed that after a period of overtraining, lever pressing becomes independent of the reward delivery and insensitive to devaluation. This suggests that during the early phase of training, performance is dependent on making the R-O association. But after a period of overtraining, when performance becomes habitual, it relies on making the S-R association regardless of outcome devaluation.

Another dissociation between goal-directed and habitual behavior is the sensitivity to the contingent relationship between an action and its consequence. This means that if an action is no longer associated with a reward, then goal-directed control will show a subsequent reduction in that particular action whereas the behavior will not be affected if it is a habit (Dolan & Dayan, 2013). Dickinson (1998) showed that over-trained rats have response persistency in an instrumental contingency degradation paradigm whereas under-trained rats were sensitive to outcome contingency. Using cellular recording, studies have shown that the R-O association is related to the activity of dorsomedial striatum which arises goal-directed behavior and neurons in the dorsolateral striatum make the S-R association which is the foundation of habitual behavior (Yin et al., 2009). Furthermore, a transition from the R-O to the S-R association in behavior is associated with a transition of activity from dorsomedial to dorsolateral striatum (for a review: Cox & Witten, 2019).

1.4 MF and MB Systems in Humans

Inspired by animal paradigms, a new line of studies combined neuroimaging techniques with the human versions of reward-based learning and instrumental conditioning tasks and captured the mutual and distinct representation of goal-directed and habitual behavior across human brain. Delgado et al. (2000) developed a paradigm with three conditions (reward, punishment, neutral) in which participants had to guess the outcome of each card. They found distinct striatal BOLD signal for reward and punishment. Particularly,

the DS (here, caudate nucleus) remains active following a reward, while its activity drops significantly following a punishment, suggesting a different neural mechanism for encoding reward and punishment. O'Doherty et al. (2003) simulated a Pavlovian conditioning task where conditioned stimuli were associated with positive, neutral, or negative outcomes. Their results showed activity in the VS in response to negative and positive prediction error. In a different paradigm, Haruno et al. (2004) show that short-term reward is associated with the activity in dorsal caudate nucleus, and the accumulation of reward is represented in the OFC.

Later, two studies mimicked the free operant paradigm from animal studies in a human functional magnetic resonance imaging (fMRI) research and found homologous results between the two species in both behavioral and neural level. Tricomi et al. (2009) examined outcome devaluation in under-trained and over-trained groups. They showed that the behavior in the under-trained group is sensitive to outcome devaluation whereas in the over-trained group it became insensitive. Interestingly, a comparison of imaging data between early and later phases of training showed task related increase of activity in right posterior putamen (homologous to dorsolateral striatum in rodents). In another study, Valentin et al. (2007) scanned participants in two sessions. In the first session, they trained participants moderately and associated actions with the delivery of different rewarding drinks (tomato juice, chocolate milk and orange juice). In the second session, after feeding the participants with one of the options to satiety, they saw that activity of the OFC drops significantly in response to the devalued in comparison to non-devalued option.

The OFC is a region that has been linked repeatedly with valuation. Specifically, ventromedial PFC (vmPFC) is a complex structure that is involved in the representation of action value (FitzGerald et al., 2012), stimulus value and outcome value (Dolan & Dayan, 2013). However, Camille et al. (2011) studied human participants who had focal damage in the OFC or the dorsal anterior cingulate cortex (dACC) damage. They showed that participants with the OFC damage have deficit in learning the stimulus value but not the action value whereas participants with the dACC damage have impairment in an

opposite direction. in another study, Kovach et al. (2012) used a four-armed bandit task in human participants with lesion in the frontopolar cortex (FPC) in which they had to choose between four available options on the screen. The task was designed in a way that required tracking the task contingency rather than a simple S-R association to maximize reward. Compared to healthy controls, lesioned participants were unable to tract the task contingency and their decisions were relied entirely on the reward history of choices, suggesting the importance of the FPC in outcome valuation. Another region that has less been linked to MB processing is the hippocampus as it is an important region in the representation of a cognitive map and the future states (Dolan & Dayan, 2013). Human subjects with the hippocampal lesion are reported to have impairment in imagining possible future states (Hassabis et al., 2007)

Altogether, human neuroimaging studies have localized regions that are involved in processing and executing MF and MB behavior. These are different regions of the PFC (more importantly FPC and OFC) as well as their connections to the striatum collectively called corticostriatal pathways. Taking into account the complexity of this network, studies have reported similar but sometimes dissimilar regions for processing the same function. The dissimilarity between the studies might be due to the small differences in their methodological approach. Below, two important paradigms are discussed that enables us to study the integration of the two systems.

1.5 Methods to Study MF and MB Integration

The paradigms to study the integration of MF and MB systems can generally be classified into two categories: sequential learning and inferential learning tasks (Doll et al., 2012). The sequential learning tasks distinguish the behavior in terms of considering the immediate reward or searching the decision tree for choosing the most appropriate action. This structure is frequently studied in two-step Markov decision task, multi-arm bandit task and mazes. The inferential learning tasks involve making inferences about the reward which is studied in paradigms such as serial reversal contingency or Binary Learning Urn Task (BLUT). Here, the Markov decision task and the BLUT will be discussed in detail as they are more relevant to the context of this study.

1.6 Markov Decision Task

Markov decision task has two steps (Figure 1a). First, participant has to choose between two options which leads to a second stage. The second stage choices lead to rewarded/unrewarded outcome with different probabilities (i.e., one option is mostly associated with reward (70%) while the other option is rarely associated with reward (30%)). The transition from first to second stage choices is based on a probabilistic rule (Figure 1.b) (Daw et al., 2011). Through repeated explorations, the participant gradually builds a mental model of predictable associations between actions and outcomes at every stage. This task has a general framework: (1) States, which represent the stimuli or the contexts; (2) Available actions at any particular state; (3) Utilities, which represent the immediate value associated with each state and can be quantified in terms of how rewarding or punishing each state would be; and finally (4) Transition from each state to another which is determined by actions. In order to successfully complete the task, a participant has to learn and adapt a policy that involves integrating information from all four aspects (Dolan & Dayan, 2013). Learning the utilities associated with each state without considering the transition rule would lead to a pure MF approach, i.e., considering only the immediate outcomes which would not lead to a profitable long-term outcome.

Using this framework, the two-step Markov decision task has been used extensively in both human and animal studies (Huang et al., 2020). This task is designed to capture how MF and MB systems use distinct mechanisms to evaluate and control actions. A model learner only considers the magnitude of reward/punishment associated with each action based on experience in the past. However, MB learner looks forward and considers the task structure to search through all the possible actions and action consequences. In the Markov decision task, a MF agent would only consider the second-stage choices with their corresponding rewards while a MB agent would consider task structure and transition probabilities from the first to second stage choices. Logically, the task is designed in a way that gives separate prediction for MF and MB strategies. As an illustration, imagine that a participant's choice in the first stage has led to the less

probable second stage and the second choice has led to the reward. From a MF standpoint, this experience would increase the likelihood of repeating the same action (choosing the same first stage choice) as it has been associated with reward. However, for a MB agent that has an internal model of the task structure, the reward associated with the second stage would increase the expected value for the other first stage choice. This is because the probability of reaching to that particular second stage is higher by choosing the alternative option. Thus, a MB agent would decrease the tendency to repeat the same action (Daw et al., 2011).

Using this task allows for distinguishing MF and MB behavior in terms of staying on the same choice or switching to the alternative option. In particular, a MF learner preferably stays on the rewarded option without considering common/rare transition rule from the first to the second stage. On the other hand, a MB learner switches to the more rewarding option by considering the common/rare transition. This distinction gives two separate predictions that allows for studying the underlying neural mechanism as well as the individual differences in MF and MB behavior. Moreover, mathematical models would allow us to assess the behavioral and neural reflection of MF and MB trial-by-trial. Despite many strengths, the two-stage Markov decision task has several important shortcomings as a method for examining the integration of MF and MB learning. This task requires active maintenance and retrieval of the information and is therefore demanding of working memory. As such, task performance may reflect individual differences in working memory function in addition to difference in reward-based learning.

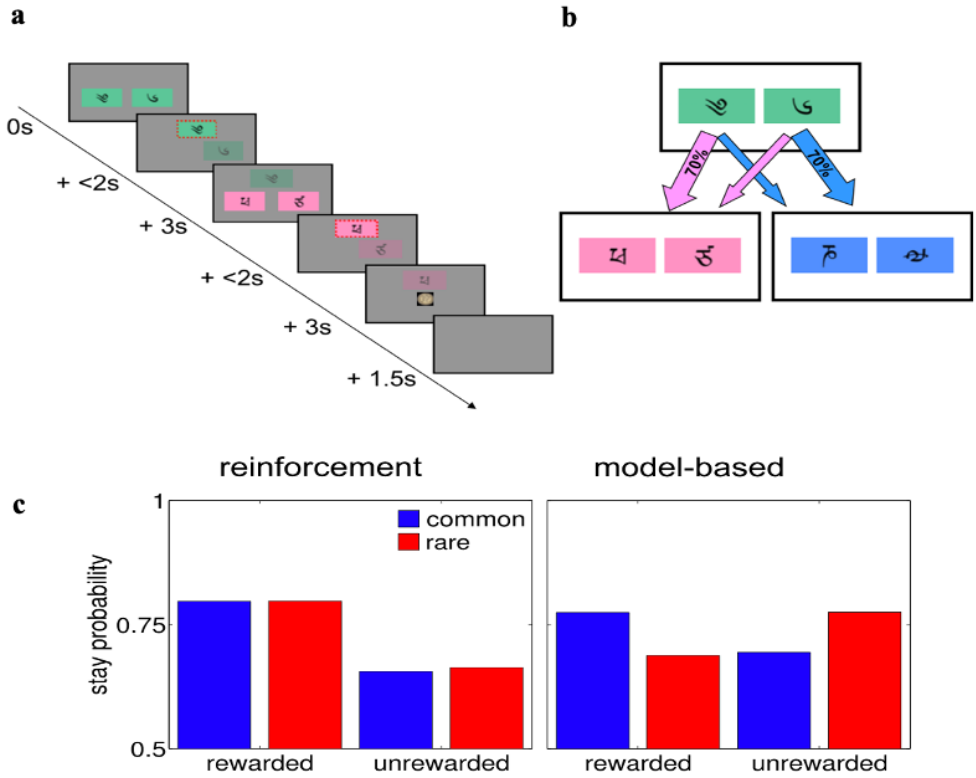


Figure 1. Schematic of the two-stage Markov task. (a) Timeline of an example trial. Two options are presented in the first stage (green boxes) leading to the second-stage choices (either between two pink or blue options), which are reinforced with monetary reward based on certain likelihoods. (b) Structure of state transition in the task. Each of the first-stage choices are associated with either of the second stage choices but with different probabilities. (c) Stay-switch probability based on MF or MB strategies. A MF reinforcement learner decides to stay on the same choice after being rewarded disregarding the common/rare transition. However, a MB learner uses a rare transition to update the value for the alternative option at the first stage and to switch subsequent choices (adapted from Daw et al., 2011; use of this is under the copyright permission defined by Creative Common license: http://creativecommons.org/licenses/by/3.0/)

1.7 Binary Learning Urn Task (BLUT)

To address the methodological issues in the Markov decision task, Fischer et al. (2017) developed a novel BLUT for examining the integration of MF and MB learning in humans. In this task, participants learn the value of two different urns by selecting between them over a number of trials and making note of the number of points earned after each choice. Selecting a particular urn yields a varying number of points from trial to trial, with any particular outcome varying from negative (e.g., -40) to positive (e.g., +40) in increments of 10 points. In each trial, after selecting an urn and making note of the number of points received, participants indicate whether their belief about the goodness (or badness) of the urn has changed. Throughout the task, participants have available to them two probability distributions, one representing the payout likelihoods for a "good" urn and one representing the payout likelihoods for a "bad" urn, where "good" and "bad" urns are urns that return positive and negative long-term payoffs respectively. Therefore, when updating their urn valuations, participants can consider either the immediate magnitude/valence of the feedback or its information content (i.e., what it signals about the long-term probability distributions). To illustrate, Figure 2a shows an example trial in which a participant has received +40 points as a result of choosing the yellow urn. Although the valence of the feedback is positive, this payout is more likely to be drawn from the bad urn since its likelihood is higher in the bad pie chart (Figure 2b). Looking into belief updates would make it possible to study how learning is driven by the magnitude of payouts or long-term inferences. That is, a MF learner only takes the valence of the feedback into account to form the belief about the chosen urn while a MB learner considers the information provided in the pie charts. At the end of each block, participant has the chance to receive bonus points by correctly identifying the true nature of both urns (Figure 2c).

The design of the task is well-suited to study how immediate and long-term action consequences shape learning. First, the pie charts displaying the probability distributions remain on the screen throughout the task. This minimizes working memory involvement and ensures the availability of both short-term and long-term information at the time of learning and action selection. Second, short-term and long-term evaluations are

orthogonalized in a way that can have opposing reflection on belief update and this reflection can be studied and computationally modeled to see how behavior is shaped by short-term or long-term action consequences.

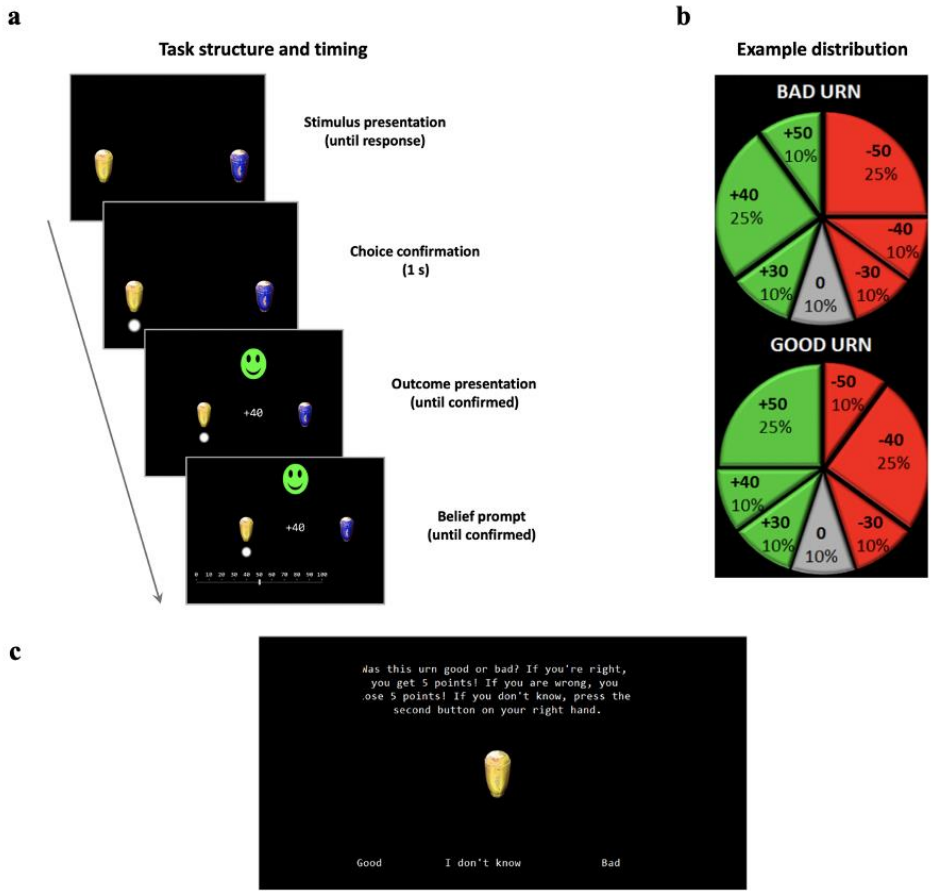


Figure 2. Schematic of the BLUT (a) An example trial with the timing of each event. Participant chooses between two urns and receives a pay-out. Considering the magnitude of the pay-out or the information it carries to infer from the pie charts, participant updates his/her belief about the chosen urn by adjusting the marker on the belief bar (b) Pie charts representing possible pay-outs with their respective likelihoods in the "good" and "bad" urn. (c) At the end of each block, participant can identify the true long-term valence of both urns and receive bonus points. Misidentifying will result in losing points and the option to avoid this gamble is also given.

1.8 The Integration of MF and MB Systems

Taking into account the presence of two valuation systems in the brain (Balleine et al., 2008), researchers started to investigate the integrated representation of MF and MB values in the brain as the two systems learn concurrently and often interact in executing adaptive behavior. Gläscher et al. (2010) found that there exists two separate learning signals in human brain being processed in distinct regions for guiding actions. They found the representation of MF learning in the VS and MB learning in the intraparietal sulcus (IPS) and the lateral PFC (lPFC). The representation of MF learning in the VS is well documented in previous studies (Pagnoni et al., 2002; for a meta-analysis: Huang et al., 2020). Later, Daw et al. (2011) tested an important question of whether the ventral striatal BOLD signal reflects purely MF learning and found that although activity in the VS (as well as the vmPFC) is associated with MF learning, but the BOLD signal is better explained if MB learning is also introduced as an explanatory variable.

Recently, Fischer et al. (2017) investigated this integration in the BLUT and showed that human behavior is not purely MF or MB, but a mixture of both (Figure 3b). Using fMRI, they found that MF and MB learning are represented in the VS and the DS respectively (Figure 3a). Although the dissociation was apparent, they found an incremental representation of both signals in an opposite direction with the degree of overlap in the medial striatum corresponding to the degree of bias toward MB belief update. Moreover, the gradient slope of MF signal representation from the VS to the DS was correlated with the degree of MF influence on belief updating; meaning that the steeper slope of MF signal was associated with a greater bias toward MF behavior on trial level (Figure 3c). This finding suggests that an integrated representation of both signals is important for adapting MB control over choices.

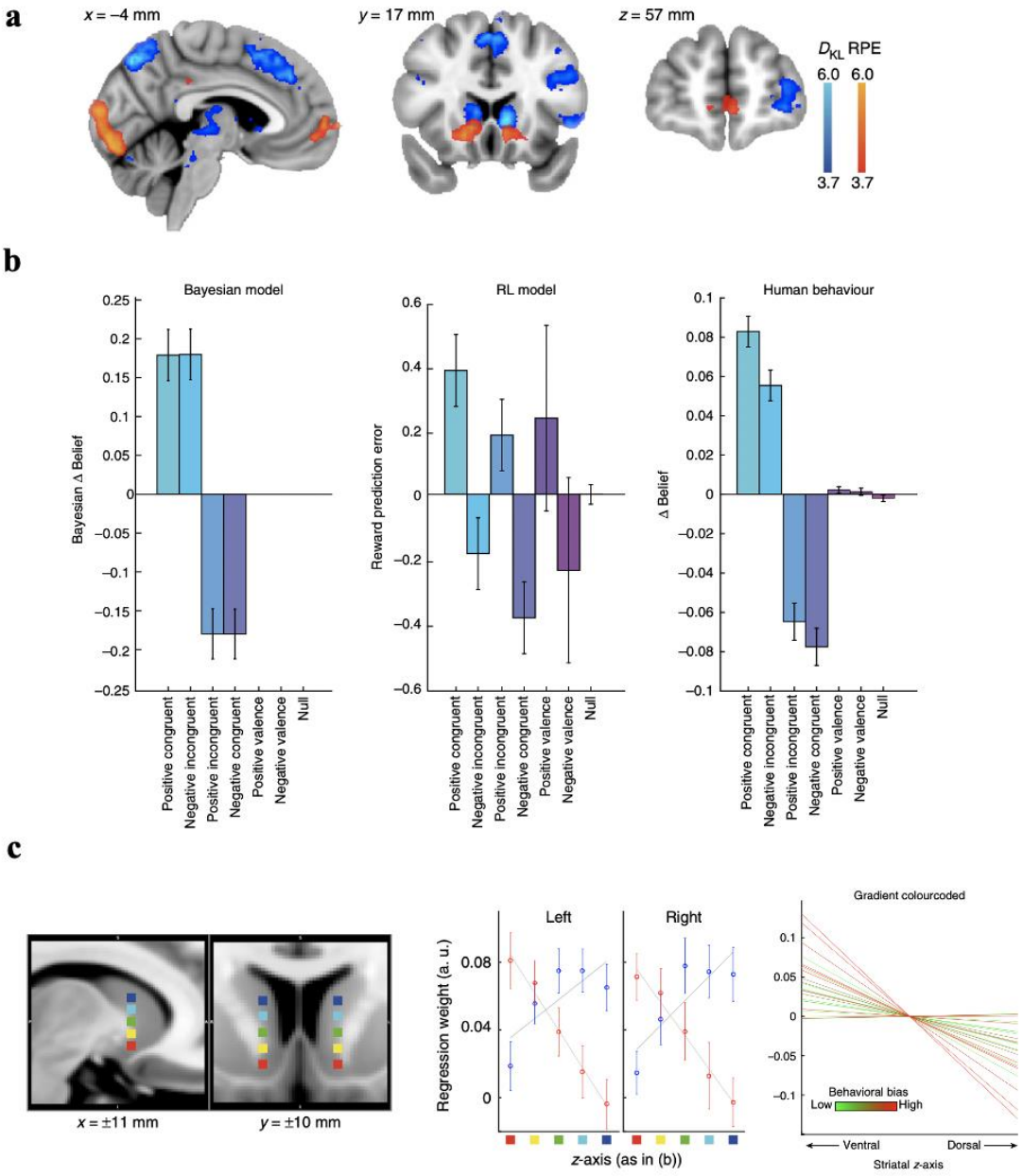


Figure 3. Summary of Fischer et al. (2017) results. (a) fMRI results of MF (RPE) and MB (Dkl) learning. (b) The Bayesian model updates the belief following each payout by only considering the information provided in the pie charts regarding the true nature of each urn. Whereas the RL model updates the belief by considering the valence of payouts regardless of the information provided in the pie charts. Human behavior is neither purely MF nor MB, but a combination of both. (c) Incremental representation of MF learning from ventral to dorsal striatum and MB

learning in an opposite direction. Color marks indicate 5 mm steps along the striatum from ventral to dorsal. The steeper gradient slope of MF learning was associated with the greater MF behavioral bias (Adapted from Fischer et al., 2017; use of this is under the copyright permission defined by Creative Common license: http://creativecommons.org/licenses/by/4.0/)

1.9 Individual Differences in MF and MB Behavior

The integration of MF and MB learning raises an important question regarding the individual differences in MF and MB behavior. In fact, what are the contributing factors in the balance between the two systems? One way to address this question is to focus on the anatomical differences between individuals who show opposite behavioral preference. de Wit et al. (2012) used diffusion tensor imaging (DTI) to highlight the individual differences in corticostriatal connectivity. Specifically, they show that vulnerability to habitual behavior is associated with the strength of white matter tracts between premotor cortex and posterior putamen as well as the gray matter density in the putamen. On the contrary, goal-directed flexible behavior is associated with the stronger white matter tracts between the vmPFC and caudate nucleus.

1.10 Pharmacological Manipulation

Another approach toward understanding the individual differences is the use of pharmacological manipulation of the DA precursor (Crockett & Fehr, 2014). A common method to alter brain DA concentration is using levodopa. Pessiglione et al. (2006) administered levodopa to human participants in an fMRI instrumental learning study and showed that levodopa modulates striatal BOLD signal of the RPE. In another study using the Markov decision task, Wunderlich et al. (2012) show that levodopa administration biases behavior toward MB choices. Also, Deserno et al. (2015) used F-DOPA (a radiolabeled variant of levodopa) in a combined positron imaging topography (PET) and fMRI study to measure striatal DA level while participants performed the Markov decision task. They show that the ventral striatal DA level is associated with the balance

between MF and MB decision making. In particular, they found that higher presynaptic DA level in the VS is associated with the behavioral bias toward MB choices.

In summary, a number of studies have identified brain regions that are involved in MF and MB learning and decision making. A recent quantitative meta-analysis on fMRI studies shows that MB learning is associated with activity of the mPFC and the OFC whereas MF learning is associated with activity of different regions of the striatum. Also, both MF and MB learnings are associated with the activity of the VS (Huang et al., 2020). In line with this, Daw et al. (2011) used the Markov decision task and found an integrated representation of MF and MB learning in the VS. On the other hand, Fischer et al. (2017) used the BLUT and found an integrated representation of both learnings along the striatum. Apart from this, the DA projections from midbrain to the striatum encode reward-based learning and previous studies have found that DA baseline level determines the balance between MF and MB behavior. One way to increase the DA baseline level in the brain is to use levodopa which is the primary medication for Parkinson's disease. Previous studies used the Markov decision task and showed that levodopa administration biases decisions toward MB choices. However, this arbitration has not been studied in the BLUT which is a new way to study and computationally model the balance between the two systems with minimum working memory involvement.

1.11 Current Study

One of the overarching aims of this study is to do a pharmacological manipulation of DA to address the effect of DA augmentation on the balance between MF and MB behavior. By using fMRI, we aim to measure changes in neural activity with respect to the medication and behavioral task administration. The second overarching aim of this study is to identify brain regions involved in the processing of MF and MB learning and action selection. Previous fMRI study of the BLUT reported the association of MF learning with activity of the VS and MB learning with activity of the PFC and the DS (Fischer et al., 2017). However, some results suggest that learning is not associated with activity of the DS (Hiebert et al. 2019). To address this, I designed an active and a passive version of the

BLUT to dissociate learning and action selection and study the contribution of striatal subregions in learning from short-term and long-term action consequences.

RESEARCH QUESTIONS:

1. Does a pharmacological augmentation of DA influence the balance between MF and MB learning and performance in the BLUT?
2. By dissociating learning and action selection in the design of the BLUT, can we find evidence that the DS is involved in learning and/or performance?

To address the first question, Levocarb will be administered which contains 100 mg of levodopa and 25 mg of Carbidopa. Levodopa is the precursor of DA that can cross blood-brain-barrier and be synthesized to DA by reaching to the dopaminergic neurons in the midbrain. Carbidopa is a peripheral decarboxylase inhibitor which inhibits DA synthesis before reaching the brain as well as reducing side effects such as nausea and lowering of blood pressure. Levocarb reaches its half-life in an hour and increases DA availability which provides the possibility for cognitive testing. Previous studies have investigated the influence of Levocarb administration using the Markov decision task. The Markov decision task recruits working memory, and DA manipulation might reflect individual differences in working memory function. However, the availability of both short-term and long-term information in the BLUT minimizes the involvement of working memory; as a result, this task might better reflect the integration of MF and MB learning and decision making regardless of working memory. To address the second question, active and passive versions of the task have been designed and validated to be administered inside an fMRI scanner. The paradigm validation is discussed further in this document.

In summary, I aim to study the effect of a pharmacological manipulation of DA in a within-subject design. In this regard, participants will complete two sessions, one on Levocarb and the other on Placebo in a double-blinded design. Moreover, in order to dissociate learning and action selection, active and passive versions of the task will be administered in a between-subject design. The passive version is the yoked control for the active version and will be administered to the matched participants. Participants in each group will complete the task inside a 3T fMRI scanner which makes it possible to

identify the brain regions involved as well as to measure the effect of medication on the BOLD signal. Although high spatial precision of fMRI allows for localizing areas involved, its low temporal resolution needs to be considered. In this regard, a fast event-related fMRI paradigm is designed and validated for the sequence of three events in each trial of the BLUT. Using fast event-related design allows for estimating the hemodynamic responses for every single event in each trial.

Chapter 2

2 Methods

2.1 Participants

The primary criteria to participate in this study is to be healthy, right-handed, and between 18-30 years old. Here, healthy means that participant is not taking any regular medication at the time of data collection, especially those that would interfere with the DA medication. Prior to the study session, participants are screened for contraindications of Levocarb administration (Appendix A&B) as well as MRI scanning (Appendix C). Moreover, participants are required to have systolic blood pressure of above 100 to be eligible to proceed with Levocarb administration (Appendix D).

2.2 Study Session

An entire study session is approximately three hours. In order to administer Levocarb, blood pressure is measured for three times using the standard blood pressure cuff along with mood questionnaire using the Bond & Lader Visual Analogue Mood Scale (Appendix E). First to ensure the participant does not have low blood pressure before drug administration, then 45 minutes after administration to ensure that the participant is ready to go inside the scanner; and lastly, after scanning and before leaving the session. During the 45-minute waiting period, task instructions are provided, and an example block is administered to participants while they are waiting in the study room. Also, additional questionnaires (e.g., Starkstein Apathy Scale (Appendix F), Beck Depression Inventory, and Beck Anxiety Inventory) are given in order to measure symptoms that might be affected by the dopaminergic medication.

After running the pilot sessions, we aimed to reduce the scanning duration as it became apparent that it is long and cause sleepiness in participant. However, choice of the amount of data acquisition is a tradeoff between boredom and data quantity. To replicate findings of Fischer et al. (2017), we aim to collect 240 trials per participant for behavioral modeling. On the other hand, there are different types of learning which would happen in

this task: incidental learning which involves learning the task instruction and structure (e.g., likelihood of payouts in the pie charts and learning which buttons to press) that is irrelevant to the research questions and occur during the first initial blocks. The essential learnings are learning the short-term and long-term valence which would happen trial-by-trial throughout each block. In this regard, the four initial blocks are run outside the scanner which reduces the scanning duration to around 60 minutes. Beside reducing the scanning time, these four blocks will help familiarize the participant with the task prior to collecting imaging data.

2.3 Experimental Task Design

In my study, I used E-prime 3 to design and administer two versions of the task and used MATLAB to extract information generated by the task for computational modeling. To dissociate learning and action selection in the BLUT, I designed an active and a yoked passive version of the task. The passive version is administered to a participant that is matched with another person in the active group and involves performing and learning the task through other's choices.

2.4 Active Version

In the active version, participant has to make active decisions throughout the task. First, two urns are presented, and participant has to choose between them. After selecting an urn, a payout is randomly drawn from the chosen urn based on whether it is a good or a bad urn and according to the probability of different payouts presented in the pie charts (e.g., Figure 2b). Then, a belief bar will be presented on the screen in which participant has to indicate whether his/her belief about the chosen urn has changed in light of the experienced payout. The active version is similar to the BLUT used in Fischer et al. (2017) except in two parts. First, a motor response is required for the feedback section in each trial. That is, the participant has to confirm receiving the feedback by clicking a button. By adding this, every event in each trial has a motor response which gives this possibility to extract similar pattern of motoric responses further in the analyses of imaging data. Second, the marker on the belief bar will be reset to 50 (neutral) after every

trial and this would incentivize the participants to mentally keep track of their belief about each urn.

2.5 Passive Version

The purpose of designing the passive version is to examine the same sequence of events without having participants making any active decision. For this reason, the exact same sequence of events made and experienced by a participant in the active group will be administered to a matched participant in the passive group. The participant will simply observe one of the two urns being selected, what the outcome of that selection is, and how the valuation marker was updated by the participant in the active. The only responses that should be made in the passive condition is to click the button for the number of times it has been recorded from the matched participant in the active condition. This means that there will be no urn selection or belief update and we do not aim to behaviorally model belief updates in the passive group. The only active decision here is to decide whether the true nature of each urn is "good" or "bad" which is made at the end of each block. The design of the passive version would make it possible to dissociate the neural correlates of learning and action selection. Figure 4 demonstrates the two versions of the task.

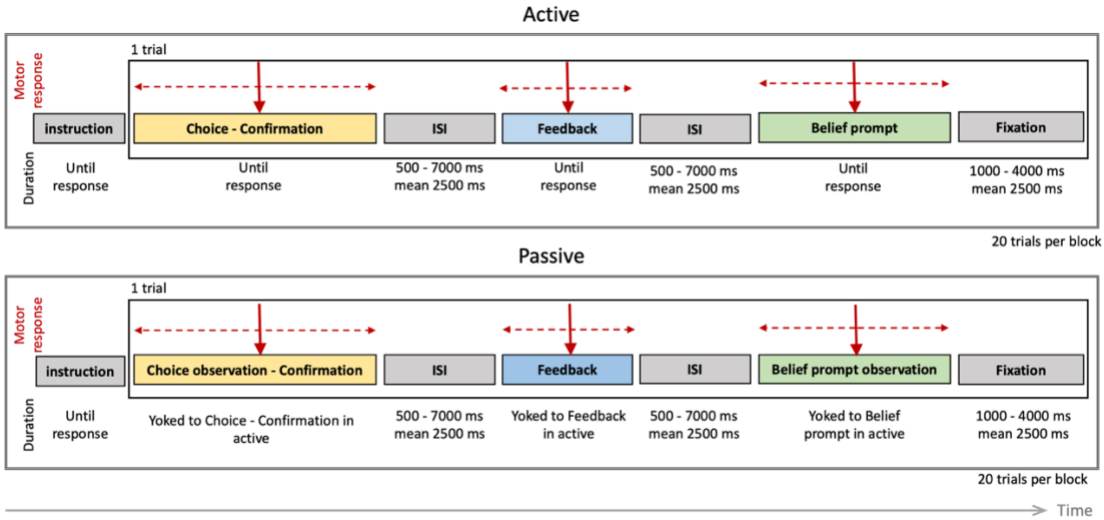


Figure 4. The active and passive BLUT. The three consecutive events in each trial include urn selection, feedback observation and belief update. The sequence of events is the same between both versions. The vertical red arrows represent the

motoric responses and the horizontal dashed arrows show the time window in which the motoric responses would happen. Inter-stimulus intervals (ISI) are randomly drawn from an exponential curve with the mean of 2500 ms.

Chapter 3

3 Paradigm Validation

An important challenge in a fast event-related design is the collinearity of parameters in close temporal proximity. In an fMRI experiment, collinearity happens when two regressors approximately share the same onset or follow each other with a fixed interval which would result in blurry hemodynamic response function (HRF) and impacts the quality of estimating variability between predictors (Mumford et al., 2015). The common method to separate predictors in close temporal proximity is to use jittered ISI so that the β parameter for events of interest can be estimated and the relative contribution of each regressor would be clear.

3.1 Creating Jittering and Design Matrix

Using random jittering as opposed to periodic jittering leads to more variance in the BOLD response and is a better way to separate events (Pernet, 2014). In Fischer et al. (2017) a uniform distribution between 3-8 s with 1 s steps was used to determine the ISI. In order to accelerate the paradigm, I have used an exponential curve ranging between 500-7000 ms with 500 ms steps and an average of 2500 (Figure 5). Changing the step-size adds more variability to the ISI distribution allowing for more randomness and unpredictability. Also, using an exponential distribution accelerates the paradigm allowing for collecting more data as well as decreasing subject boredom. Having done this, I needed to make sure the resulting predictors in the design matrix would be sufficiently orthogonalized to allow the β coefficients associated with those predictors to be estimable. One way of addressing this is by generating the design matrix and correlating the columns to check how orthogonalized they are and make sure that the correlations are not above the threshold of ±0.3, which is the standard value for evaluating whether the predictors are sufficiently orthogonalized. In order to address this, after feeding the task with an exponential ISI distribution, a set of 30 simulated data was acquired to extract the randomly drawn ISI (Figure 5), create the design matrix (Figure 6) and examine whether the drawn ISI would not result in significant overlap of HRF of closely presented events (Figure 7).

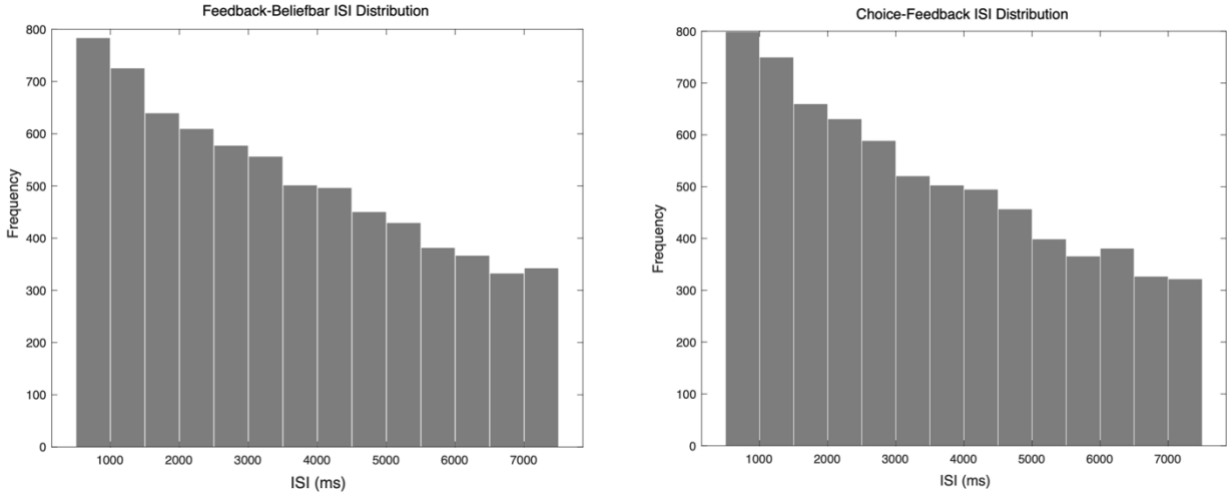


Figure 5. The distribution of randomly generated ISI between Choice-Feedback and Feedback-Belief prompt in the simulation dataset as the result of feeding the task with an exponential curve ranging 500-700 ms (mean = 2500 ms).

I used MATLAB to extract the onset and the duration of each event and create the design matrix from the E-prime output. The resulting design matrix had six different event types based on which urn was selected (good/bad), which feedback was drawn (positive/negative), and how the marker was adjusted (above/below 50) (Figure 6a).

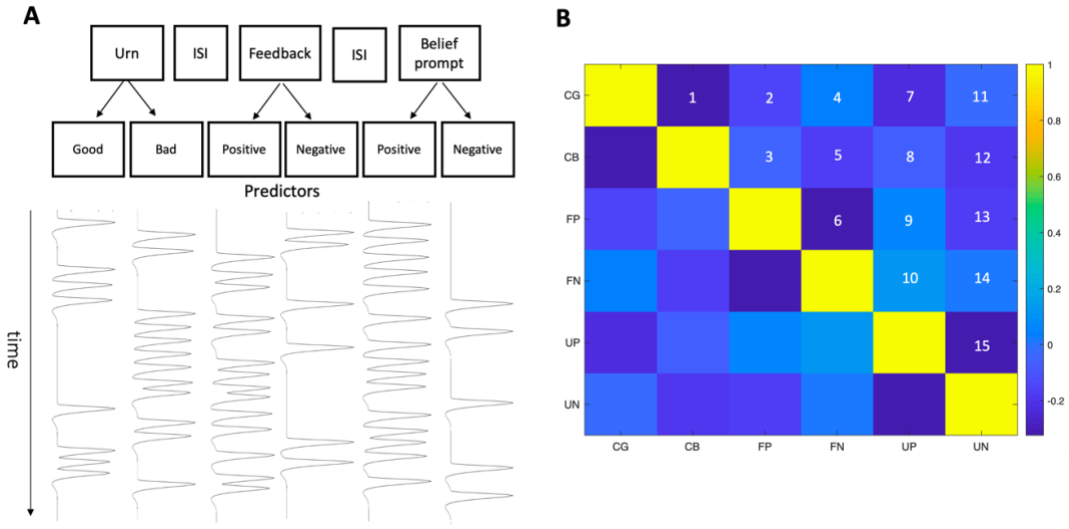
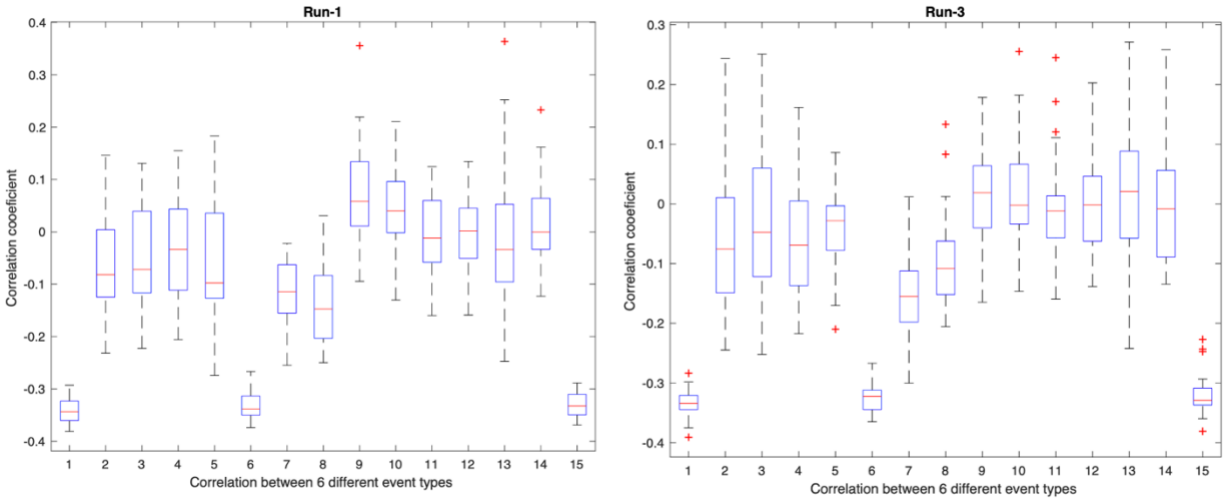
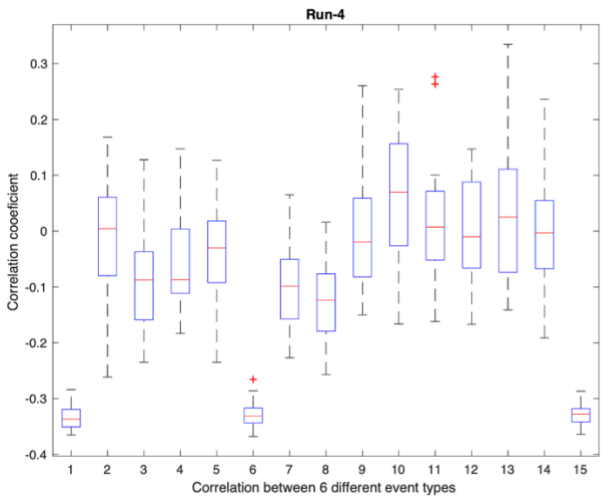


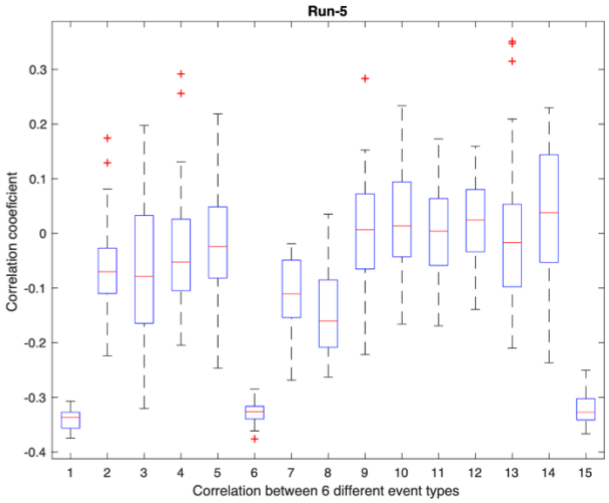
Figure 6. Steps in calculating the correlation coefficients between the six various event types. Choice good (CG) and Choice Bad (CB) are selection of good/bad urn in

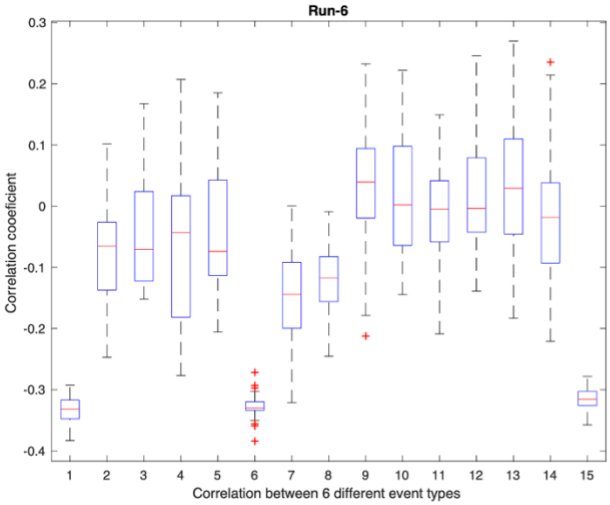
the choice stimuli. Feedback Positive (FP) and Feedback Negative (FN) are based on the valence of feedback following a choice. Update Positive (UP) and Update Negative (UN) are based on whether the belief bar is updated toward 100 or 0. (A) The onset and the duration of every event was extracted from the E-prime data file. Then, the created delta function was convolved to HRF to evaluate six predictors. (B) correlations between all the predictors were below ±0.3 showing that the predictors in close temporal proximity are not highly correlated. Note that the numbers on the upper triangle corresponds to the x-axis in Figure 7.

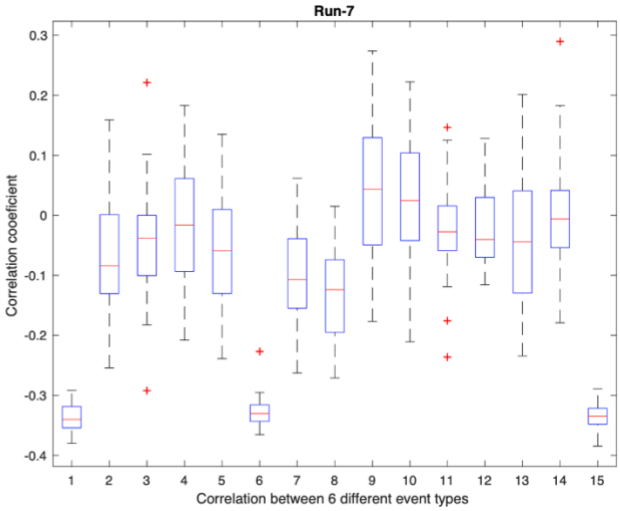
Next, the delta function was convolved to HRF to create the predictors in the design matrix and the correlation between predictors was measured resulting in 15 different correlation coefficients. As shown in Figure 7, all the coefficients are below ±0.3 showing that the jittering distribution can separate events properly. It is worth mentioning that the duration of each event was kept at 500 ms in the simulation dataset. In practice however, event durations will vary based on the participant's responses. This will add more variability to the timing of events and would help to further decorrelate sequential events (Mumford et al., 2015).

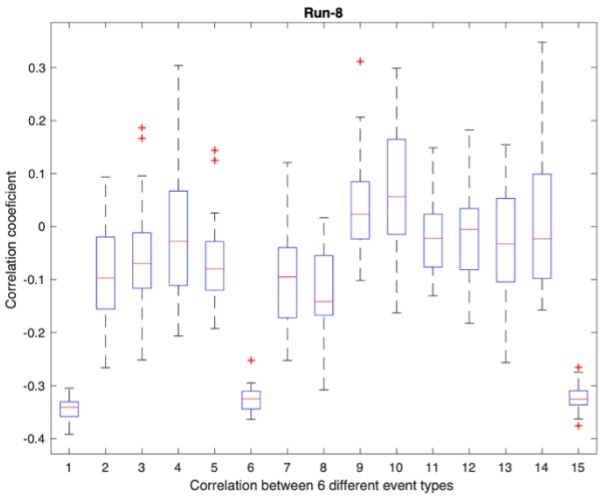


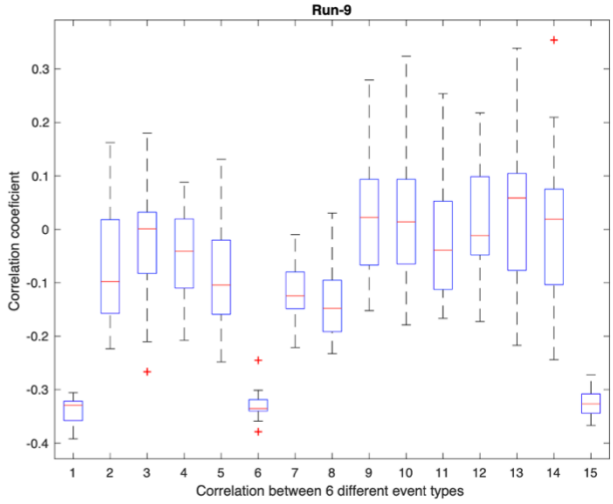

Run-4
Correlation cooeficient
Correlation between 6 different event types


Run-5
Correlation cooeficient
Correlation between 6 different event types


Run-6
Correlation cooeficient
Correlation between 6 different event types


Run-7
Correlation cooeficient
Correlation between 6 different event types


Run-8
Correlation cooeficient
Correlation between 6 different event types


Run-9
Correlation cooeficient
Correlation between 6 different event types

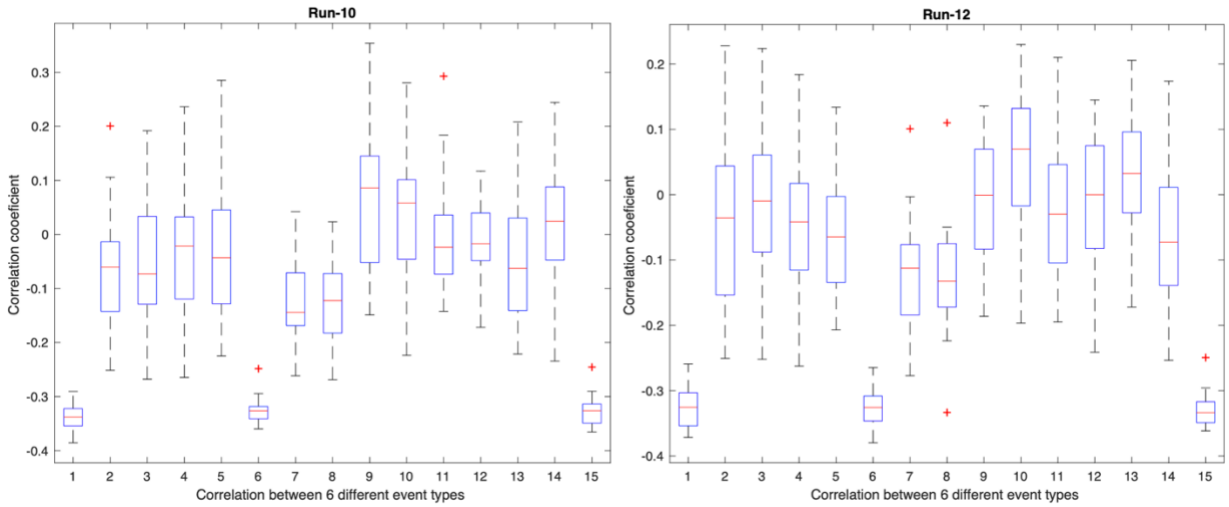


Figure 7. Correlation among six predictors (Choice Good, Choice Bad, Positive Feedback, Negative Feedback, Positive Update, Negative Update) for every block (Run 1-12) resulting in 15 different correlations. All the correlation coefficients are between ±0.3 meaning that the different events in each trial are successfully separated by the created ISI. Note that both of the urns are good in Run 2 and bad in Run 11 and the boxplots for these blocks are not included in the figure.

3.2 Preliminary Estimations

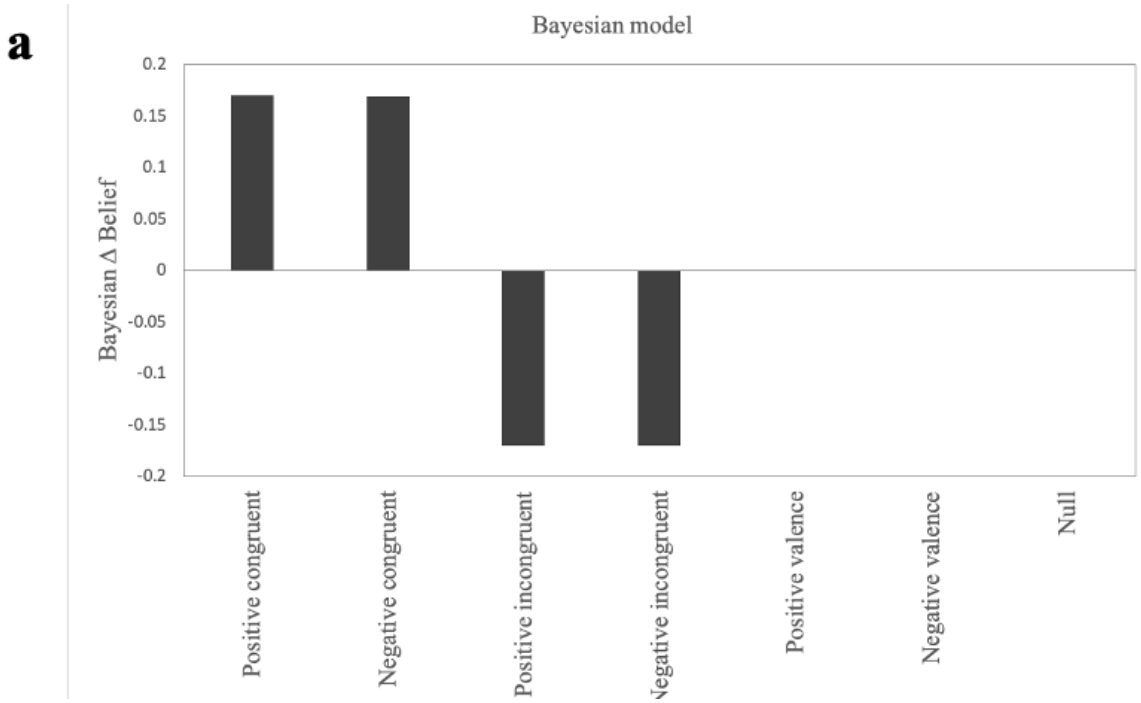
Based on whether a payout carries information about the long-term value of an urn, every possible event type in this task can be either informative or non-informative. The information content can be used to infer whether the short-term and long-term valence of a payout is congruent or incongruent. To illustrate, in Figure 2b, +40 and +50 are informative events as they have unequal probabilities between "Good urn" and "Bad urn" whereas +60, -60 and 0 are non-informative events as they have equal probability between the two. A congruent event can signal that both short-term and long-term valence of the chosen urn is positive (e.g., +50) or negative (e.g., -50) and an incongruent event (e.g., +40) would convey that the long-term outcome of the chosen urn can be negative even though it is accompanied by a positive reward. Also, beside zero that is neutral both in short-term and long-term, some payouts carry no information about the long-term consequences but still have positive or negative payouts (e.g., +30 and -30). Table 1 demonstrates all the possible payouts in different blocks in the behavioral task.

Short-term valence	Long-term valence	Informativity	Congruency	Event distribution		Payouts			
				Good urn	Bad urn	Variant 1	Variant 2	Variant 3	Variant 4
Negative	Negative	Informative	Congruent	10%	25%	-50	-50	-30	-30
Negative	Positive	Informative	Incongruent	25%	10%	-40	-40	-20	-20
Negative	None	Non-informative	None	10%	10%	-30	-60	-10	-40
None	None	Non-informative	None	10%	10%	0	0	0	0
Positive	None	Non-informative	None	10%	10%	+30	+60	+10	+40
Positive	Negative	Informative	Incongruent	10%	25%	+40	+40	+20	+20
Positive	Positive	Informative	Congruent	25%	10%	+50	+50	+30	+30

Table 1. A list of possible event types in every block in the BLUT. The payouts vary between blocks, but they are all classified based on the same principle presented above. Informative events are those that provide information about the long-term valence of the chosen urn. That is, the probability distributions of informative events vary between the good and the bad urn which is kept consistent between blocks (10%-25%). Some events do not convey long-term information but still have positive or negative points (short-term reward) that can have reinforcing effect and be modeled using RL model. Totally, 30% of the possible payouts do not carry information about the long-term consequence and there is always a 10% probability of receiving zero points which has neither short-term nor long-term valence.

In order to study the behavioral preferences in belief updating considering the informativity and the congruency of each payout, I used MATLAB to extract and classify belief updates based on the following categories: Positive congruent (PC), negative congruent (NC), positive incongruent (PI), negative incongruent (NI), positive valence (PV), negative valence (NV) and null. Next, I used MATLAB codes to implement the assumptions of the models to calculate both RL and Bayesian predicators for every trial. Figure 8a and 8b are the RL and Bayesian simulation of belief updates in the task considering the event types. The Bayesian model updates belief only on PC, NC, PI and NI events as it considers only the information in the pie charts regarding the probability

distributions (Figure 8a). On the other hand, the RL model considers the valence of the payouts regardless of the information in the pie charts. Therefore, the RL model updates on all the event types except null (Figure 8b). Although the number of data points are very limited for statistical analysis, but as shown in Figure 8c, human behavior in our pilot data is a combination of both models meaning that participants consider the valence as well as the information content of the payouts to form their belief about the urns. Here, I am using a Rescorla-Wagner RL model to calculate the RPE and a Bayesian model to calculate the MB belief update. The hybrid model solves the task using different approach. The correlation between MF and MB estimates are tested and Figure 9 shows that the two estimates are not correlated with each other. Below, the computational modeling of belief formation and how the hybrid model calculates two estimations of belief update is discussed.



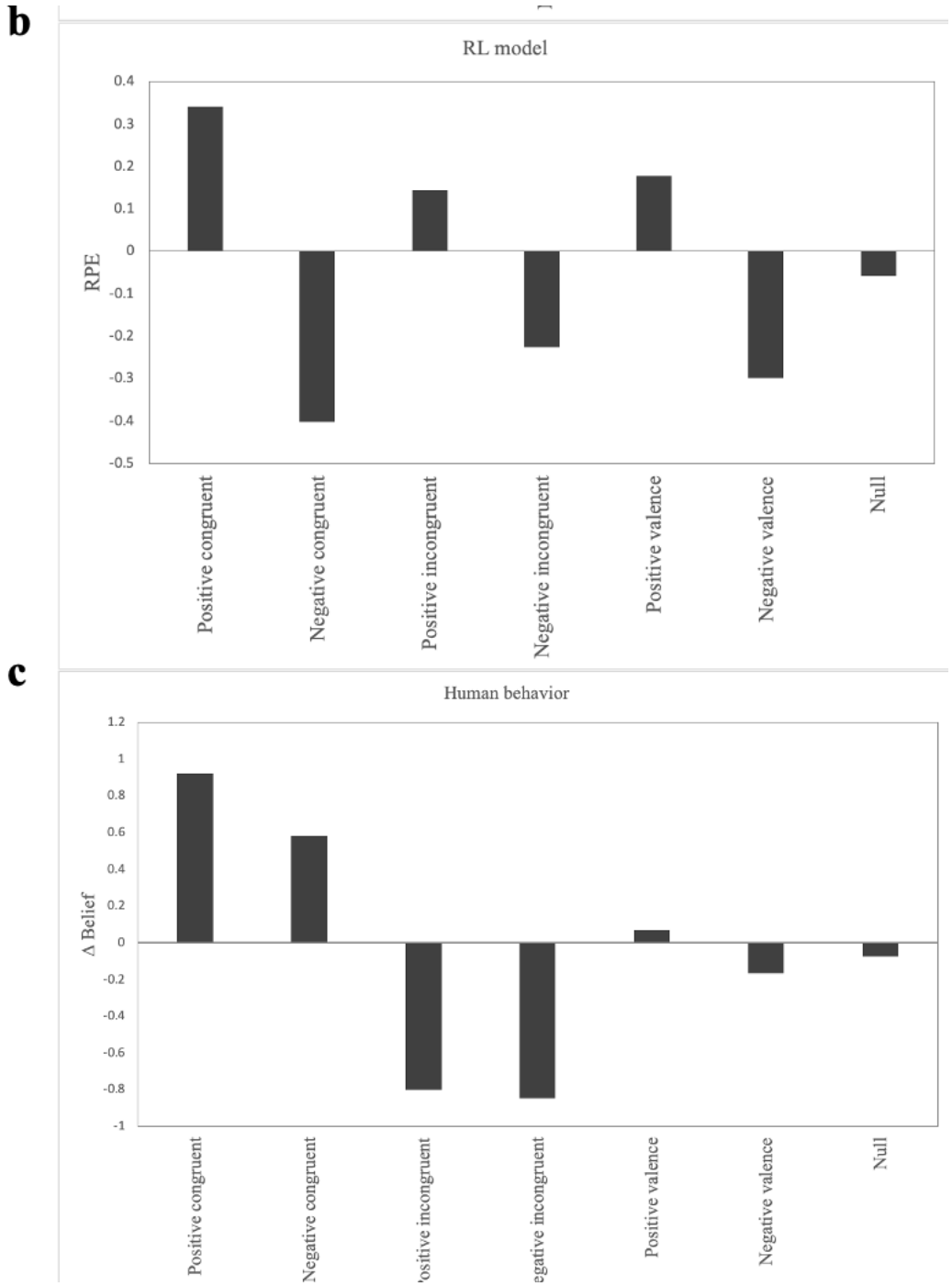


Figure 8. Belief updating by the Bayesian model, RL model and human behavior in 4 pilot datasets. (a) The Bayesian model updates the belief about the true nature of both urns by considering the information provided in the pie charts about the probability distributions. Therefore, it only updates on PC, NC, PI and NI as these are the only events that carry long-term information. (b) The RL model updates belief by considering the reward magnitude and it updates on all of the event types

in the task. (c) Our four pilot datasets show that human behavior is a combination of both the Bayesian MB and the RL MF behavior. This means that human subjects consider both the valence of the payout and the information it carries about long-term consequence.

3.3 The Rescorla-Wagner RL Model

The central concept in the Rescorla-Wagner RL model is learning from previous experience and use that experience as a basis for future choices. A MF learner considers only the evaluation of reward obtained as a consequence of an action to repeat a behavior that was successful in the past (Huang et al., 2020). To explain how this model works, let's assume that a participant has a set of stimuli (s) available to choose (here, two presented urns on the screen) and maintains a set of predictions about the value of reward obtained for every available option (V(t)). Then, upon choosing a stimulus and receiving the reward on a trial (t), the prediction error is calculated based on a comparison between how much reward is obtained in a particular trial and what was expected based on the experience from previous trials which is called Prediction Error (RPE=δ_t) (Glimcher & Fehr, 2013).

$$\delta_t = R_t - V_t(S_t) \qquad \text{Equation 3.1}$$

Worth mentioning that the RPE on every trial is only being updated for the selected stimulus (chosen urn) and predictions for the other urn remains the same until being selected. Next, the model uses RPE to update the value associated with a stimulus with a learning rate (α).

$$V_{t+1}(S_t) = V_t(S_t) + \alpha \,.\, \delta_t \qquad \text{Equation 3.2}$$

Learning rate is a number between 0 and 1 and determines the size of the update. If α is closer to 0, the update step is small and will not change the overall value much. By combining Equation 3.1 and Equation 3.2 this concept is better explained.

$$V_{t+1}(S_t) = (1-\alpha)V_t(S_t) + \alpha \,.R_t \qquad \text{Equation 3.3}$$

Equation 3.3 shows that the value update is a weighted average between the current reward ($\alpha \,.R_t$) and the previous predictions ($(1-\alpha)V_t(S_t)$). In this regard, a bigger learning rate updates the value in a way that it is influenced largely by the current reward and be more similar to what is recently experienced. On the other hand, a smaller learning rate updates the value in a form that is more biased toward older estimation of a stimulus. Further, Rescorla-Wagner model explains that the given weight to the rewards received declines exponentially from present to past trials and the steepness of the exponential curve depends on the learning rate (Glimcher & Fehr, 2013).

3.4 Bayesian MB Model

One way to explain MB learning is using Bayes' theorem. Bayesian MB model uses the information about the probability of reward/losses to compute the confidence level with respect to the rewarding nature of a set of actions. Some suggest that a Bayesian algorithm can explain MB learning since mapping its computation in the cortico-striatal circuitry is possible (Forstmann & Wagenmakers, 2015). Here, in order to update belief about the chosen urn, the posterior belief (B_{t+1}) is calculated purely MB (unbiased by MF) considering the prior belief about the urn and the likelihood of observed event (E_t) to be good based on the information provided in the pie charts regarding the probability of each event to be drawn from a good urn.

$$B_{t+1} = \frac{P(E_t|Good)\times B_t}{P(E_t|Good)\times B_t + P(E_t|Bad)\times (1-B_t)} \qquad \text{Equation 3.4}$$

The Bayesian learner calculates the difference between posterior belief and prior belief (B_t) to update the overall degree of change in the Bayesian model per observation.

$$\Delta Bt = |B(t+1) - B(t)\,| \qquad \text{Equation 3.5}$$

Then, the Kullback-Leibler Divergence (D_{KL}) is calculated based on the absolute change in the overall Bayesian belief.

$$D_{KL} = |B_{(t+1)} \times (\log B_{(t+1)} - \log B_{(t)})| \qquad \text{Equation 3.6}$$

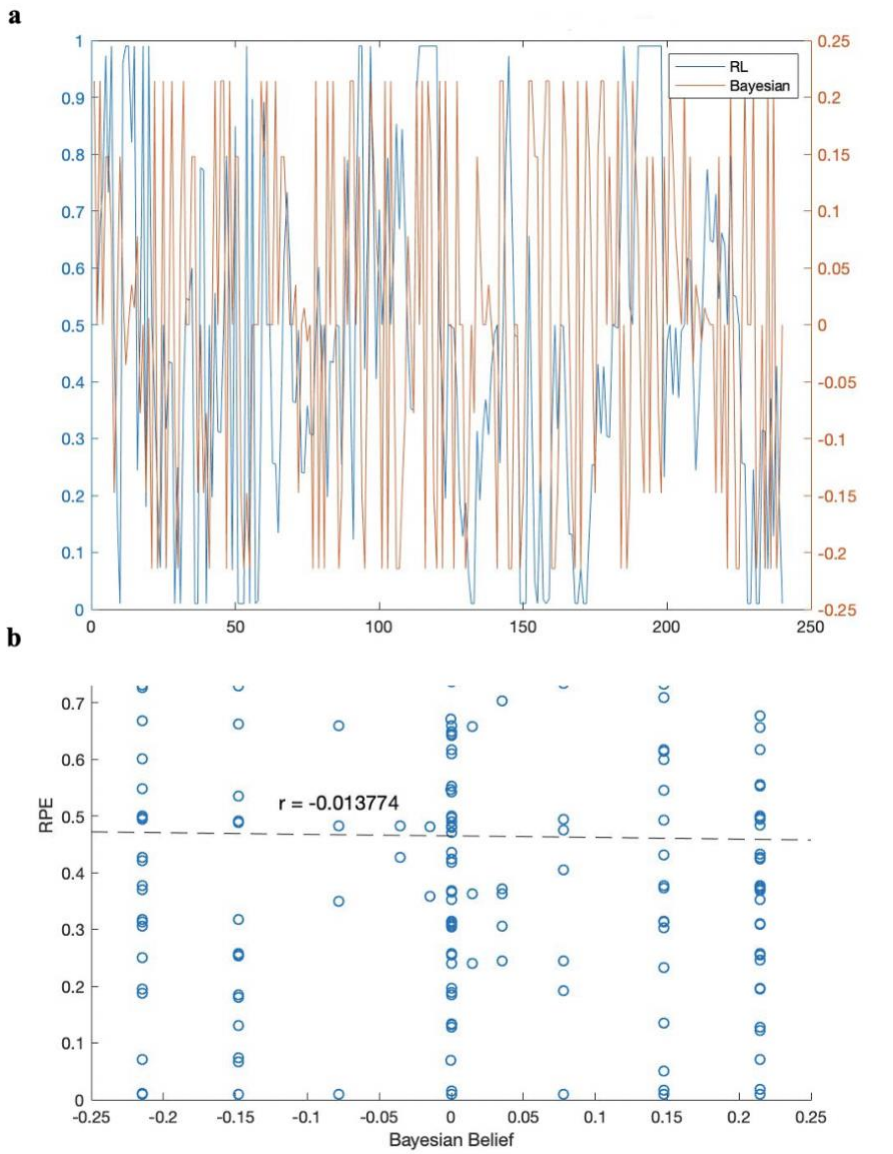


Figure 9. (a) MF and MB estimates of the belief change by the computational models. The Rescorla-Wagner RL model calculates RPE based on the valence/magnitude of the payout whereas the Bayesian model calculates DKL based on the information provided in the pie charts regarding the likelihood of observing

such event in the good and the bad urn. (b) The two products of the hybrid model give separate estimations that are not correlated with each other.

Chapter 4

4 Discussion

The current study designed and validated a paradigm that is well-suited for studying the integration of MF and MB learning and decision making in humans on both behavioral and neural level. The paradigm uses computational approach to model human decision-making strategies in the context of the BLUT. The paradigm validation here lays the groundwork for the subsequent steps in an fMRI data collection combined with a pharmacological manipulation. In general, this research aims to address two overarching research questions that are beyond the scope of this thesis. One is regarding the role of DA baseline level in the arbitration between MF and MB behavior and the other is to identify the brain regions that are involved in MF and MB learning and decision making. Decades of animal studies have shown that DA plays an important role in learning from experienced reward/punishments and encodes the value difference between the experienced and the expected reinforcer through a burst of firing projected from the midbrain DA neurons to the striatum (Schultz et al., 1992). One common method to investigate the psychopharmacology of learning and decision making in humans is a pharmacological manipulation of precursor availability (Crockett & Fehr, 2014). Here, we aim to use a pharmacological augmentation of DA using Levocarb. Levocarb contains mainly the DA precursor levodopa that can be synthesized to DA by reaching to the dopaminergic neurons in the midbrain as well as Carbidopa to reduce the side effects. Studies using the two-step Markov decision task have shown that levodopa administration influences the balance between MF and MB behavior biased toward MB choices (Deserno et al., 2015; Wunderlich et al., 2012). However, as behavior in the Markov task requires retrieving information from working memory about the structure of the task, and as DA plays a major role in working memory function (Cools & D'Esposito, 2011), it is difficult in this task to investigate the balance between the two systems through a pharmacological manipulation of DA regardless of working memory involvement. In this regard, I modified a paradigm developed by Fischer et al. (2017) in which the information about both short-term and long-term action consequences are available on the screen at the time of learning and decision making. The BLUT gives

orthogonal RL and Bayesian estimation on trial level which makes it possible to study the integration of the two systems. Although the BLUT is a promising paradigm in the investigation of MF and MB behavior, the design of the task makes it difficult to further dissociate the neural correlates of learning and action selection as they happen concurrently. Fischer et al. (2017) reported the representation of MF learning in the VS and MB learning in the DS (Figure 3d-f). However, some studies suggest that learning does not depend on neural activity of the DS (in rats: Atallah et al., 2007; in humans: Hiebert et al., 2019). In order to address this, the active and passive versions of the task are designed. The active version is similar to the paradigm used in Fischer et al. (2017), except that a motoric response is required in feedback section and the belief marker is reset to neutral after every trial. On the other hand, the passive version is the yoked control for the active, meaning that the matched participant in the passive group will simply observe the decisions made by the participant in the active group and learn from them. The purpose of designing a passive version of the task is to address the second overarching research question regarding the involvement of brain areas (importantly striatal regions) in learning the long-term action consequences and/or action selection. The passive version includes the exact same number of motoric responses but are not associated with active selection meaning that the actions are regardless of any decision-making preference which makes it possible to dissociate the striatal response to action selection and learning. Due to the COVID-19 pandemic, we were unable to collect data from participants who underwent the pharmacological manipulation. However, the paradigm is validated for an fMRI data collection by using 30 simulation datasets in a fast event-related design. Moreover, human belief formation based on the short-term and long-term information is classified in four behavioral datasets and the belief formation in the Rescorla-Wagner RL model and the Bayesian MB model are simulated.

4.1 Caveats and Limitations

Unlike animal studies, one big methodological challenge in the psychopharmacological investigation of human decision making is the limitation to target a certain brain region. Levodopa administration changes the DA concentration in the whole brain and targeting specifically the striatal DA level but no other region by using a pharmacological

manipulation is not possible. Moreover, DA is involved in a range of cognitive and neuronal functions and the widespread expression of DA receptors in the PFC is documented (Cohen et al., 2002). The new paradigm developed and validated here attempts to offset the working memory contribution influenced by the DA level alteration by minimizing the need for working memory information retrieval. However, levodopa administration might influence affected functions which cannot be controlled here. Furthermore, there exist individual differences in the DA baseline activity which might lead to contrasting effects on performance (Cools & D'Esposito, 2011). This means that the pharmacological manipulation would enhance or diminish performance depending on each individual's DA baseline level. Besides, learning is not ideal in the passive condition and can be a combination of both observational learning from the actions of matched participant and his/her own internal belief update regarding the value of the chosen urn. One possible solution would be measuring observational learning as well as participant's own belief by adding an additional belief prompt in each trial. However, as the aim of including the passive version is to address the second research question concerning the distinction of learning and action selection in the striatum, this adds an active decision which would interfere with the goal in designing the passive version in the first place. In this regard, in the passive condition, I aim not to model participant's own belief about the value of each urn on trial level but learning within each block can be measured by the final choices at the end of each block when participants identify the "good" and the "bad" urn.

4.2 Conclusion

The paradigm developed here is the groundwork of collecting data to study the integration of MF and MB systems in human. I used E-prime 3 to develop two versions of the BLUT that gives estimable and orthogonal parameters of the RL and the Bayesian learning by using computational models. Moreover, an exponential distribution (Figure 5) is used to randomly assign jittered ISI to separate the sequence of three events of each trial in close temporal proximity in a fast event-related design. Thirty simulated datasets are acquired and the correlation between six predictors are measured to test whether the assigned ISI would not lead to an overlap between the predictors (Figure 6&7). This

validation makes the paradigm well-suited for an fMRI investigation of the neural correlates of RL and Bayesian MB learning and the integration of the two in the context of learning from immediate rewards and long-term action consequences. Next, I aim to combine fMRI and levodopa administration to explore the arbitration between MF and MB behavior under the effect of DA precursor augmentation in healthy adults. Similar studies have addressed this using the Markov decision task and have found a MB bias as a result of levodopa administration. However, the paradigm developed here seems to be a more promising avenue as it minimizes the working memory involvement by providing all the necessary information for MF and MB valuation at the time of learning and action selection.

References

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. The Quarterly Journal of Experimental Psychology Section B, 34(2b), 77-98.

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. The Quarterly Journal of Experimental Psychology Section B, 33(2b), 109-121.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual review of neuroscience, 9(1), 357-381.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., & O'Reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature neuroscience*, *10*(1), 126.

Balleine, B. W., Daw, N. D., & O'Doherty, J. P. (2008). Neuroeconomics: decision making and the brain. Glimcher, PW, 367-388.

Balleine, B. W., & O'doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology, 35(1), 48-69.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron, 47(1), 129-141.

Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. Journal of Neurophysiology, 98(3), 1428-1439.

Berke, J. D. (2018). What does dopamine mean?. Nature neuroscience, 21(6), 787-793.

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. Journal of Neuroscience, 31(42), 15048-15052.

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. Current opinion in neurobiology, 12(2), 223-229.

Cools, R., & Robbins, T. W. (2004). Chemistry of the adaptive mind. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 362(1825), 2871-2888.

Cools, R., & D'Esposito, M. (2011). Inverted-U–shaped dopamine actions on human working memory and cognitive control. Biological psychiatry, 69(12), e113-e125.

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. Nature Reviews Neuroscience, 20(8), 482-494.

Crockett, M. J., & Fehr, E. (2014). Pharmacology of economic and social decision making. In Neuroeconomics (pp. 259-279). Academic Press.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron, 69(6), 1204-1215.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature neuroscience, 8(12), 1704-1711.

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. Cognitive, Affective, & Behavioral Neuroscience, 14(2), 473-492.

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., & Ridderinkhof, K. R. (2012). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. Journal of Neuroscience, 32(35), 12066-12075.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. Journal of neurophysiology, 84(6), 3072-3077.

Deserno, L., Huys, Q. J., Boehme, R., Buchert, R., Heinze, H. J., Grace, A. A., ... & Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. Proceedings of the National Academy of Sciences, 112(5), 1595-1600.

Dickinson, A. (1998). Omission learning after instrumental pretraining. The Quarterly Journal of Experimental Psychology: Section B, 51(3), 271-286.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. Neuron, 80(2), 312-325.

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. Current opinion in neurobiology, 22(6), 1075-1081.

Fischer, A. G., Bourgeois-Gironde, S., & Ullsperger, M. (2017). Short-term reward experience biases inference despite dissociable neural correlates. Nature communications, 8(1), 1690.

FitzGerald, T. H., Friston, K. J., & Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. Journal of Neuroscience, 32(46), 16417-16423.

Forstmann, B. U., & Wagenmakers, E. J. (Eds.). (2015). An introduction to model-based cognitive neuroscience (pp. 139-156). New York: Springer.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585-595.

Glimcher, P. W., & Fehr, E. (Eds.). (2013). Neuroeconomics: Decision making and the brain (pp. 283-298). Academic Press.

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience, 41(1), 1-24.

Hans, J. (2011). Clinical neuroanatomy: brain circuitry and its disorders (pp. 495-553). Springer Science & Business Media.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., ... & Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. Journal of Neuroscience, 24(7), 1660-1665.

Hiebert, N. M., Owen, A. M., Ganjavi, H., Mendonça, D., Jenkins, M. E., Seergobin, K. N., & MacDonald, P. A. (2019). Dorsal striatum does not mediate feedback-based, stimulus-response learning: An event-related fMRI study in patients with Parkinson's disease tested on and off dopaminergic therapy. NeuroImage, 185, 455-470.

Huang, Y., Yaple, Z. A., & Yu, R. (2020). Goal-oriented and habitual decisions: Neural signatures of model-based and model-free learning. NeuroImage, 116834.

Kovach, C. K., Daw, N. D., Rudrauf, D., Tranel, D., O'Doherty, J. P., & Adolphs, R. (2012). Anterior prefrontal cortex contributes to action selection through tracking of recent reward trends. Journal of Neuroscience, 32(25), 8434-8442.

Maia, T. V., & Conceição, V. A. (2017). The roles of phasic and tonic dopamine in tic learning and expression. Biological Psychiatry, 82(6), 401-412.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. Nature neuroscience, 14(2), 154-162.

Mumford, J. A., Poline, J. B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. PloS one, 10(4).

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. Neuron, 38(2), 329-337.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. Nature neuroscience, 5(2), 97-98.

Pernet, C. R. (2014). Misconceptions in the use of the General Linear Model applied to functional MRI: a tutorial for junior neuro-imagers. Frontiers in neuroscience, 8, 1.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature, 442(7106), 1042-1045.

Salamone, J. D., & Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. Neuron, 76(3), 470-485.

Schultz, W. (1999). The reward signal of midbrain dopamine neurons. Physiology, 14(6), 249-255.

Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. Journal of neuroscience, 12(12), 4595-4610.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nature neuroscience, 16(7), 966-973.

Sutton, R. S., & Barto, A. G. (1998). Introduction to reinforcement learning (Vol. 135). Cambridge: MIT press.

Tanaka, S. C., Samejima, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., & Doya, K. (2006). Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. Neural networks, 19(8), 1233-1241.

Thorndike, E. L. (1911). Animal Intelligence: Experimental Studies Macmillan. New York.

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. European Journal of Neuroscience, 29(11), 2225-2232.

Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. Journal of Neuroscience, 27(15), 4019-4026.

Vaillancourt, D. E., Schonfeld, D., Kwak, Y., Bohnen, N. I., & Seidler, R. (2013). Dopamine overdose hypothesis: evidence and clinical implications. Movement Disorders, 28(14), 1920-1929.

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. Neuron, 75(3), 418-424.

Yin, H. H., Mulcare, S. P., Hilário, M. R., Clouse, E., Holloway, T., Davis, M. I., ... & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nature neuroscience, 12(3), 333-341.

Appendices

Appendix A

FOR EXPERIMENTER ID: _____ Date: ____ / ____ / _____

LEVODOPA SAFETY SCREENING QUESTIONNAIRE

Please answer the following questions are accurately as possible.

1. Do you currently suffer or have you previously suffered from the following:

Yes	No		Yes	No	
☐	☐	Heart or coronary artery disease	☐	☐	Glaucoma
☐	☐	Atrial, nodal, or ventricular arrhythmias (i.e., irregular heart beat)	☐	☐	Asthma, chronic obstructive pulmonary disease (COPD), emphysema
☐	☐	Myocardial infarction (i.e., heart attacks)	☐	☐	Melanoma (skin cancer) or a skin growth that has not been diagnosed
☐	☐	High blood pressure	☐	☐	Phenylketonuria (PKU)
☐	☐	Liver or kidney disease	☐	☐	Endocrine (hormonal) disease
☐	☐	Diabetes	☐	☐	Stomach or intestinal ulcers
☐	☐	Mental Illness	☐	☐	Allergies to levodopa (Larodopa)/carbidopa (Lodosyn)

2. Are you currently pregnant, plan to become pregnant, or are breast-feeding? ☐ Yes ☐ No

3. Are you taking any of the following medications (within past two weeks, i.e., 14 days):

Yes	No		Yes	No	
☐	☐	**Monoamine oxidase (MAO) inhibitors** such as ixocarboxazid (Marplan), phenelzine (Nardil), tranylcypromine (Parnate)			
☐	☐	**Antidepressants** ('mood elevators') such as amitriptyline (Elavil), amoxapine (Asendin), clomipramine (Anafranil), desipramine (Norpramin), doxepin (Adapin, Sinequan), imipramine (Tofranil), nortriptyline (Aventyl, Pamelor), protriptyline (Vivactil), and trimipramine (Surmontil)			
☐	☐	Haloperidol (Haldol)	☐	☐	Metoclopramide (Reglan)
☐	☐	Risperidone (Risperdal)	☐	☐	Papaverine (Pavabid)
☐	☐	Rasagiline (Azilect)	☐	☐	Isoniazid (INH, Nydrazid)
☐	☐	Phenytoin (Dilantin)	☐	☐	Ipratropium (Atrovent)
☐	☐	Antihistamines	☐	☐	Iron pills and vitamins containing iron

Appendix A: Levodopa Safety Screening Questionnaire

Appendix B

For administrator's use only	Date (dd/mm/yy):	
	Subject #:	Session #:
	Medication:	Time:

Health and Demographic Questionnaire

Please print and fill out this form as accurately as possible and bring it with you to your first appointment session. If you are attending your appointment with another participant, please ensure you both have your own personal copies filled out.

1. Basic Demographic Information

Date of Birth: _____ Age: _____

Weight: _____ Height: _____

Handedness: _____

First language: _____ Other languages: _____

Level of Education and total years (e.g. 4 years high school, 4 years university, etc.)

_____

Occupation (If Retired, what was your previous profession): _____

2. Health-Related Information

A. Smoking History (please circle): Never Smoker Ex-Smoker Current Smoker

If current smoker, indicate how many years and how many cig/day: _____

If ex-smoker, indicate year that you quit; how many years smoking; how many cig/day:

_____

B. Alcohol History

Average number of drinks per week: _____

Has there ever been heavy alcohol consumption? (please circle) Yes No

If yes, when, for how long, and estimate your weekly alcohol consumption during that time:

_____

C. Other Drug History

Have you ever taken street drugs or other drugs that were not prescribed by a physician (please circle)? Yes No

If yes, when, what drugs, how frequently and over what period of time?

_____

D. Eye Glasses (only if applicable)

What is the prescription of your eye glasses? _____

Without the aid of glasses are you able to see near objects well (please circle)? Yes No

Without the aid of glasses are you able to see far objects well (please circle)? Yes No

E. Parkinson's Disease (only if applicable)

What year were you diagnosed with Parkinson's disease? _____

Which side of the body is *more* affected? _____

F. Obsessive-compulsive disorder (OCD; only if applicable)

What year were you diagnosed with OCD? _____

Are you currently taking medication to treat your OCD? _____

G. Substance Dependence (i.e. addiction; only if applicable)

What year were you diagnosed with Substance Dependence? _____

H. REM-Sleep Behaviour Disorder (RBD; only if applicable)

What year were you diagnosed with RBD? _____

I. Parkinson's Plus Syndromes (only if applicable)

Have you been diagnosed with (Please indicate with a ✓):

- ☐ Progressive Supranuclear Palsy (PSP)
- ☐ Multiple Systoms Atrophy (MSA)
- ☐ Lewy Body Dementia
- ☐ Cortico-basal ganglionic Degeneration (CBGD)

What year were you diagnosed with the disorder? _____

3. Previous Medical Problems

Have you had any major health problems or do you have any chronic, ongoing medical conditions such as high blood pressure, high cholesterol, diabetes, thyroid problems, multiple sclerosis or epilepsy? Have you had any strokes, heart attacks/ heart surgeries, significant head trauma, or cancer? If you've had cancer, what kind and what treatments did you receive (e.g. chemotherapy)? Have you ever had more than one seizure? Answer in the space below.

4. Family Medical Problems

Is there anyone in your family with a neurological or serious psychiatric illness such as PD, Huntington's, epilepsy, strokes at a young age (< 50 for men and < 60 for women)? Is there anyone who had trouble walking or with balance, needing a wheelchair or a walker at a young age? Any family members with dementia (such as Alzheimer's), schizophrenia, bipolar/manic depression, or severe depression or anxiety requiring hospitalization or close follow up by a psychiatrist? Answer in the space below.

5. Current Medication

Please list any medications you are currently taking, what they are treating for specifically, and the prescribed dosage.

Appendix B: Health and Demographic Questionnaire

Appendix C

MAGNETIC RESONANCE ENVIRONMENT SCREENING QUESTIONNAIRE

This MR system has a very strong magnetic field that may be hazardous to individuals entering the magnet room if they have certain metallic, electronic, magnetic, or mechanical implants, devices or objects. Therefore, all individuals are required to fill out this form BEFORE entering the magnet room. Be advised, the magnet is ALWAYS ON.

1. Have you had prior surgery or an operation (eg. arthroscopy, endoscopy, etc) of any kind? ❑Yes ❑No
   If yes, please provide: Date(s):_____ Type(s) of surgery: _____
2. Have you had an injury to the head or eye involving a metallic object (e.g. metallic slivers, foreign body)? ❑Yes ❑No
   If yes, please describe: _____
3. Have you ever been injured by a metallic object (e.g. BB, bullet, shrapnel, welding accident, etc.)? ❑Yes ❑No
   If yes, please describe: _____
4. Are you pregnant, suspect you may be pregnant or attempting to conceive? ❑Yes ❑No
5. Have you had a previous contrast dye reaction? ❑Yes ❑No

WARNING: Certain implants, devices or objects may be hazardous to you in the MR environment or the magnet room. DO NOT ENTER the MR environment or the magnet room if you have any of the following implants, devices or objects.

Please indicate if you have any of the following:

❑Yes ❑No Aneurysm clip(s)
❑Yes ❑No Cardiac pacemaker, pacemaker wires, or stents
❑Yes ❑No Implanted cardioverter defibrillator (ICD)
❑Yes ❑No Electronic or magnetically-activated implant or device (electrodes, wires, metallic filter or coil)
❑Yes ❑No Neurostimulation system, spinal cord stimulator
❑Yes ❑No Implanted or transcutaneous bio-stimulator (spinal cord, bone growth/bone fusion, tens unit, etc.)
❑Yes ❑No Cochlear, otologic, or other ear implant
❑Yes ❑No Insulin or other infusion pump
❑Yes ❑No Implanted drug infusion device
❑Yes ❑No Any type of prosthesis (heart valve, eyelid spring/wire, penile, limb, etc.)
❑Yes ❑No Shunt (spinal or intraventricular)
❑Yes ❑No Vascular access point and/or catheter
❑Yes ❑No Radiation seeds or implants
❑Yes ❑No Swan/Ganz or thermodilution catheter
❑Yes ❑No Medication patch (Nicotine, Nitroglycerine)
❑Yes ❑No Any metallic fragment or foreign body
❑Yes ❑No Wire mesh implant
❑Yes ❑No Surgical staples, clips or metallic sutures
❑Yes ❑No Joint replacement (hip, knee, etc.)
❑Yes ❑No Bone/joint pin, screw, nail, wire, plate, etc.
❑Yes ❑No Tissue expander (e.g. breast)
❑Yes ❑No IUD, diaphragm, or pessary
❑Yes ❑No Tattoo or permanent makeup
❑Yes ❑No Body piercing jewelry
❑Yes ❑No Dentures/partial plates, dental work other than fillings (*Must be removed*)
❑Yes ❑No Hearing aid (*Must be removed before entering the magnet room*)
❑Yes ❑No **Other implant** _____
❑Yes ❑No Breathing problem or motion disorder
❑Yes ❑No Do you have claustrophobia?

IMPORTANT INSTRUCTIONS

Remove all metallic objects before entering the MR environment or magnet room including:

❑ Hearing aids	❑ Paperclips
❑ Dentures	❑ Coins
❑ Beeper	❑ Credit/Debit cards
❑ Cell phone	❑ Magnetic strip cards
❑ Keys	❑ Pens
❑ Eyeglasses	❑ Pocket knife
❑ Hair pins	❑ Nail clippers
❑ Watch	❑ Steel-toed boots/shoes
❑ Safety pins	❑ Tools

❑ Jewelry (including body piercing jewelry)

Loose metallic objects are especially prohibited in the magnet room and MR environment.

I attest that the above information is correct to the best of my knowledge. I have read and understand the entire contents of this form and have had the opportunity to ask questions regarding the information on this form.

Person Completing Form: _____ _____ Date: _____/_____/_____
Print Name Signature

Form Reviewed By: _____ _____ Date: _____/_____/_____
Print Name Signature

PRIVACY POLICY
Western respects your privacy. The personal information collected on this form is collected under the authority of the University of Western Ontario Act, 1982, as amended, and is used for the purpose of ensuring safety within the MRI environment. Please direct questions about this collection, use, or disclosure of personal information to Joe Gati, CFMM Facility Manager: jgati@imaging.robarts.ca

Appendix C: MRI Safety Screening Questionnaire

Appendix D

Physiological Measures

Session #1:	Date:	Time:
	Subject #:	Medication:

Sitting **Pre-administration**	**Sitting** **45 min post-administration**	**Sitting** **Post-study**
Blood Pressure = mmHg	Blood Pressure = mmHg	Blood Pressure = mmHg
Heart rate = bpm	Heart rate = bpm	Heart rate = bpm
Standing **Pre-administration**	**Standing** **45 min post-administration**	**Standing** **Post-study**
Blood Pressure = mmHg	Blood Pressure = mmHg	Blood Pressure = mmHg
Heart rate = bpm	Heart rate = bpm	Heart rate = bpm

Session #2:	Date:	Time:
	Subject #:	Medication:

Sitting **Pre-administration**	**Sitting** **45 min post-administration**	**Sitting** **Post-study**
Blood Pressure = mmHg	Blood Pressure = mmHg	Blood Pressure = mmHg
Heart rate = bpm	Heart rate = bpm	Heart rate = bpm
Standing **Pre-administration**	**Standing** **45 min post-administration**	**Standing** **Post-study**
Blood Pressure = mmHg	Blood Pressure = mmHg	Blood Pressure = mmHg
Heart rate = bpm	Heart rate = bpm	Heart rate = bpm

Participant Prediction (circle)				
Session #1:	levodopa	placebo	**Session #2:**	levodopa	placebo

Appendix D: Physiological Measurement

Appendix E

For administrator's use only	Date (dd/mm/yy):	
	Subject #:	Session #:
Score:	Medication:	Time:

Bond & Lader Visual Analogue Mood Scale

Instructions: For each line below, put a vertical mark at the point that represents how you feel at this moment. Ensure to draw your line all the way through the horizontal line. The ends of each scale are to present the "most" that you have ever felt in your life.

ALERT	DROWSY	_____ mm
CALM	EXCITED	_____ mm
STRONG	FEEBLE	_____ mm
MUZZY	CLEAR HEADED	_____ mm
WELL COORDINATED	CLUMSY	_____ mm
LETHARGIC	ENERGETIC	_____ mm
CONTENTED	DISCONTENTED	_____ mm
TROUBLED	TRANQUIL	_____ mm
MENTALLY SLOW	QUICK WITTED	_____ mm
TENSE	RELAXED	_____ mm
ATTENTIVE	DREAMY	_____ mm
CONTENTED	PROFICIENT	_____ mm
HAPPY	SAD	_____ mm
ANTAGONISTIC	FRIENDLY	_____ mm
INTERESTED	BORED	_____ mm
WITHDRAWN	SOCIABLE	_____ mm

Appendix E: Bond & Lader Visual Analog Mood Scale

Appendix F

For administrator's use only	Date (dd/mm/yy):	
	Subject #:	Session #:
Score:	Group:	Time:

Starkstein Apathy Scale

Instructions: For each question, indicate as "Not at all", "Slightly", "Some", or "A lot" with an 'X' while leaving the other spaces blank.

Questions	Not at all	Slightly	Some	A lot
1. Are you interested in learning new things?				
2. Does anything interest you?				
3. Are you concerned about your condition?				
4. Do you put much effort into things?				
5. Are you always looking for something to do?				
6. Do you have plans and goals for the future?				
7. Do you have motivation?				
8. Do you have the energy for daily activities?				
9. Does someone have to tell you what to do each day?				
10. Are you indifferent to things?				
11. Are you unconcerned with many things?				
12. Do you need a push to get started on things?				
13. Are you neither happy nor sad, just in between?				
14. Would you consider yourself apathetic?				

Appendix F: Starkstein Apathy Scale

Curriculum Vitae

Name: Masood Rezaei

Post-secondary Education and Degrees:

Shomal University
Amol, Mazandaran, Iran
2008-2011 A.S.
Electronics

Islamic Azad University
Tehran, Iran
2011-2013 B.Sc.
Electrical Technology Engineering

Institute for Cognitive Science Studies
Tehran, Iran
2014-2018 M.Sc.
Cognitive Psychology

Western University
London, Ontario, Canada
2018-2020 M.Sc.
Neuroscience

Western University
London, Ontario, Canada
2020 Ph.D. Candidate

Honours and Awards:

Western Graduate Research Scholarships (WGRS)
2018-2020

Related Work Experience

Teaching Assistant
Western University
2018-2020