2016

Genetic Dissection of Neural Circuits Underlying Value Based Decision Making

Hirofumi Nakayama

GENETIC DISSECTION OF NEURAL CIRCUITS UNDERLYING VALUE BASED DECISION MAKING

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Hirofumi Nakayama

June 2016

GENETIC DISSECTION OF NEURAL CIRCUITS

UNDERLYING VALUE BASED DECISION MAKING

Hirofumi Nakayama, Ph.D.

The Rockefeller University 2016

Decision making is the fundamental process that we utilize to accomplish objectives in everyday lives. To understand the neural substrates of this process, we developed a behavioral task for mice that required repetition of the processes of action initiation, action selection, and learning. The task is a two-option choice task with stochastic reward delivery and reversals. To map brain areas involved in this type of value-based learning, we inactivated neuronal activity in the prefrontal cortex (PFC) and the nucleus accumbens (NAc) while mice performed the task.

Inactivation of the NAc resulted in altered action initiation and learning but had subtle effects on choice. To dissect underlying neural circuitry, specific cell types and inputs of the NAc were inactivated. Inactivation of two dominant cell types in the NAc, direct and indirect pathway medium spiny neurons (MSNs), showed partially overlapping but distinct behavioral effects. Inactivation of direct pathway MSNs showed the stronger effect on learning while inactivation of indirect pathway MSNs also showed the effect. In contrast, only inactivation of indirect pathway MSNs affected behavioral measures of action initiation. The contribution of specific inputs to the NAc, dopaminergic and glutamatergic inputs, were also studied. While both experiments affected behavioral measures of initiation of action, only the inactivation of dopaminergic inputs affected

learning. The effect on learning was specific to trials after reward omissions, and the effect was more prominent in trials which animals spent less time to initiate. These results provided new insights into the function of the NAc in processing information about reward values.

In contrast, inactivation of two subregions in the PFC, the anterior cingulate cortex (ACC) and the orbitofrontal cortex OFC, affected action initiation and action selection. The action initiation was affected by inactivation of both areas, but OFC inactivation affected more behavioral measures. In contrast, action selection was affected more prominently in ACC inactivation. These differential effects on action initiation and action selection suggested the functional distinction between these two areas.

In this study, we have developed a behavioral assay that allowed us to dissect different aspects of cognitive functions for decisions in mice and revealed roles of distinct circuit elements in the NAc and the PFC. Utilizing temporally precise inactivation, we found that the same circuit element was used for different cognitive processes depending on the timing. Although this type of behavioral task has been used extensively in rats and primates to understand decision making, identification of cell types and circuits required for these behaviors has been difficult in these species due to the lack of the powerful genetic methodologies. The approach we have demonstrated here is important because it enables genetic dissection of complex behaviors in mice, allowing studies of circuit properties that are executed by specific cell types in the cerebral cortex and basal ganglia. Since the approach taken in this study can be expanded to other neural circuits and behavioral paradigms, this and future studies will reveal the neural basis of decision making and, perhaps, lead to new approaches to treatments for maladaptive behaviors.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor Dr. Nathaniel Heintz. Nat let me explore what I thought the right thing, and it was a good experience for me to develop scientific skills. At the same time, my project could never be done without his support and resources in the lab. From his attitude to science, I learned how important it is to seek for great scientific discoveries without being trapped with trivial questions. I was grateful to have an opportunity to work in such an environment. I'd also like to thank my thesis committee members Drs. Winrich Freiwald and Jeffrey Friedman. It was a great pleasure to have opportunities to get feedback from such great scientists. I also like to thank Dr. Gordon Fishell for becoming my external examiner.

I would also like to thank members of Heintz lab. A special thanks to Jie Xing for invaluable help with experiments. Without her help, I could not finish my thesis. I would also like to thank Erika Andrade, for setting up operant chambers in the lab and letting me use. Without the equipment, I couldn't even start my project. I would also like to thank Eric Schmidt, Laura Kus, and Judy Walsh for reading my thesis and gave me suggestions for writing.

Outside the lab, I'd like to thank Kunal Shah of Precision Fabrication Facility for the help with 3D printing. The use of 3D printer greatly enhanced the speed of experimental cycles. Kunal was always helpful when I need to use equipment in the facility. Some experiments in my thesis can never be possible without him.

I owed a lot to Hiroshi Abe, a former postdoc in Gilbert lab. Hiroshi gave me a lot of invaluable suggestions for experimental design and data analysis. Discussion with him

was always stimulating, and I could learn a lot from his deep insights into a wide range of topics in neuroscience. That will be a great asset for my future career.

I also like to thank Akinori Ebihara, my classmate in Freiwald lab, and Miho Nakajima, a former graduate student in Heintz lab. I was happy to have fellow students who also have a passion for science. Without Akinori and Miho, my life in the Rockefeller could have been much harder.

I am deeply grateful to Nakajima Foundation for continuous support during my graduate studies.

In the end, I'd like to express my deepest gratitude to my parents and sister in Japan. They were always supportive and encouraged me to pursue my career. Without their support, I couldn't finish my work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Cognitive functions required for decision making: Action Initiation, Action Selection, and Learning

Natural environments and our daily lives are full of uncertainty. We and other animal species are always faced with decision making under uncertainty in order to obtain preferable outcomes and to avoid dangers. To achieve such adaptive decisions, animals need to exert two distinct cognitive processes. The first process is 'Action selection'. Animals need to form predictions about possible outcomes to select an appropriate action. The second process is 'Learning'. If the outcome of the action is different from the prediction, animals need to update the prediction so that they can make a better decision in the future. The process to achieve optimal decision through repeated action selection and learning is known as 'Reinforcement Learning'. This concept was originally developed in the field of engineering (Sutton and Barto 1998), and it is now widely used to explain decision making in a variety of animal species (O'Doherty et al. 2003; Samejima et al. 2005; Ito and Doya 2009; Figure1.1).

During action selection, values of available options are estimated based on experiences. Consequently, animals select an action by comparing expected reward value of each action. This process is sufficient to achieve the optimal decisions in stable environments. However, in most natural situations, the contingency between choice and outcome is occasionally changed. As a consequence, decisions based on the current action values are no longer optimal. In order to address environmental changes, the learning process becomes critical. At learning, action values are modified using environmental feedback, which is the difference between the predicted value and the actual outcome value in case

of value-based decision. Subsequent decisions are made based on the modified action value. These processes enable animals to behave adaptively even if the environment or the context frequently changes (Schultz et al. 1997, Daw and Doya 2006).

In the mammalian brain, midbrain dopamine neurons encode a reward prediction error signal when the expected reward value and the actual reward values are different (Mirenowicz and Schultz 1994). This signal meets the requirement for teaching signal in learning as was theorized in psychology (Rescorla and Wagner 1972; Pearce and Hall 1980), and it is supposed to contribute to learning by modifying neuronal activity and plasticity in the cortico-basal ganglia circuit (Montague et al. 1996).

In addition to action selection and learning, a decision about whether or not to perform an action is also a key determinant of the outcome. This process, 'Action initiation', is not explained by above processes, but it is more critical in some situations than the decision about which actions to choose. In the case of animals foraging for food, refraining from food seeking in the presence of predators during the day is more advantageous than the decision of to where they look for food. This concept was formalized in the modified reinforcement learning model (Niv et al. 2007), and recent work has begun to study the neural substrates underlying this process (Wang et al. 2013).

The decision about whether or not to perform and action, the process of action initiation, is associated with various cognitive processes in the mammalian brain. Abnormal impulsivity can be regarded as the state in which impaired action initiation is observed in pathological states, for example addiction and attention deficits hyperactivity disorder (ADHD) (Schachar et al. 1995; Fillmore and Rush 2001; Monterosso et al. 2005; Dalley et al. 2011). Another concept relevant for action initiation is motivation or response

vigor, which defines how hard animals will work to obtain an outcome (Niv et al. 2007; Wang et al. 2013). Action initiation is also involved in reward-seeking behavior (Nicola 2010). Deficits in this process lead to an inability to initiate a movement to obtain a reward even if animals or humans have an intact ability to assess reward value and control movements (Schmidt et al. 2008). Thus, the control of action initiation, together with action selection, is important for cognitive control of learning.

The mammalian brain, which performs the above behavioral processes, is a highly evolved organ composed of hundreds of different cell type (Gong et al. 2007, Nelson, Sugino et al. 2006). Precise manipulation of these circuit elements is crucial for understanding how neural circuits achieve decision making. Since action initiation, action selection, and learning are associated with each other and not easily dissociable, combining well designed behavioral paradigms and mouse genetics will provide new insights into the neural substrates of decision making.

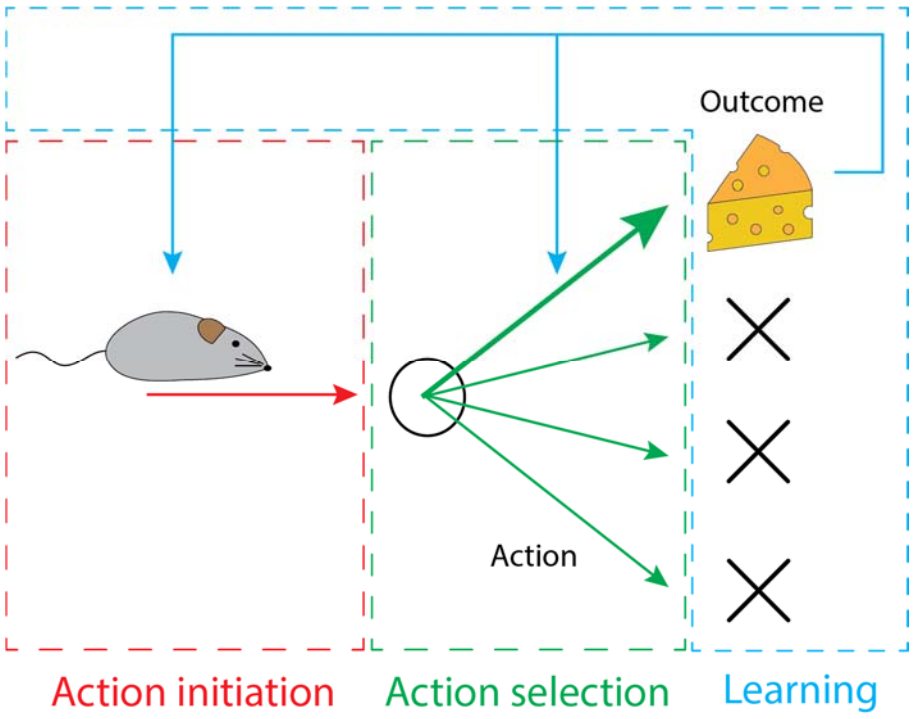


Figure 1.1 Conceptual framework of value-based decision making.

Multiple cognitive processes are necessary for adaptive decisions. (Red) Action initiation: Animals need to initiate movement (Green) Action selection: Animals need to choose one of the available actions based on outcome prediction (Action selection). (Blue) Learning: An outcomes are delivered, and that invigorate subsequent action selection and modify action selection.

1.2 Neural substrates for decision making

1.2.1 Anatomy

The basal ganglia are composed of several subcortical nuclei, including the striatum and the pallidum, and they play key roles in a variety of motor, emotional, and cognitive behaviors (Alexander, DeLong et al. 1986; Albin et al. 1989; Graybiel 2008). The proper

functioning of the basal ganglia requires the interaction with the neuromodulatory systems and the neocortex, including the prefrontal cortex (PFC).

The neocortex is both the primary input structure to the basal ganglia and the recipient of basal ganglia outputs (Parent and Hazrati 1995; Kravitz and Kreitzer 2012, Figure 1.2). The striatum, the largest basal ganglia nucleus, receives inputs from the neocortex (Eblen and Graybiel 1995; Kincaid et al. 1998). The projections from the neocortex to the striatum are organized in a topographic manner (McGeorge and Faull 1989; Berendse et al. 1992; Brown et al. 1998; Voorn et al. 2004). The striatum is roughly divided into three subregions: the dorsolateral striatum, the dorsomedial striatum, and the nucleus accumbens (NAc, also known as the ventral striatum). While the dorsal striatum receives inputs from a wide range of sensory and motor cortices, the majority of the cortical inputs to the NAc originate from the PFC (McGeorge and Faull 1989). The striatum projects to basal ganglia output structures, including the globus pallidus and the substantial nigra pars reticulata (Gimenez-Amaya and Graybiel 1990; Levesque and Parent 2005). Those output structures send information back to the neocortex via thalamic nuclei (Joel and Weiner 1994; Sakai et al. 1998).

The cortico-basal ganglia circuit both send inputs to and receive outputs from neuromodulatory systems. (Graybiel 1990). Especially, the dopaminergic modulation of the cortico-basal ganglia circuit plays significant roles in motor and reward-related behaviors (Schultz 1998; Nieoullon 2002). Most of the dopaminergic projections to the cortico-basal ganglia circuits originate from the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc) (Jimenez-Castellanos and Graybiel 1987). These

midbrain dopaminergic neurons also receive inputs from the cortico-basal ganglia circuit including the PFC and the striatum (Watanabe-Uchida et al. 2012).

For successful decision making, neural circuits need to receive the feedback of animals' actions. Thus, the loop organization of the cortico-basal ganglia circuit and its reciprocal interaction with neuromodulators have attracted a great deal of interest. In the following sections, the contribution of the cortico-basal ganglia circuit to action initiation, action selection, and learning is reviewed.

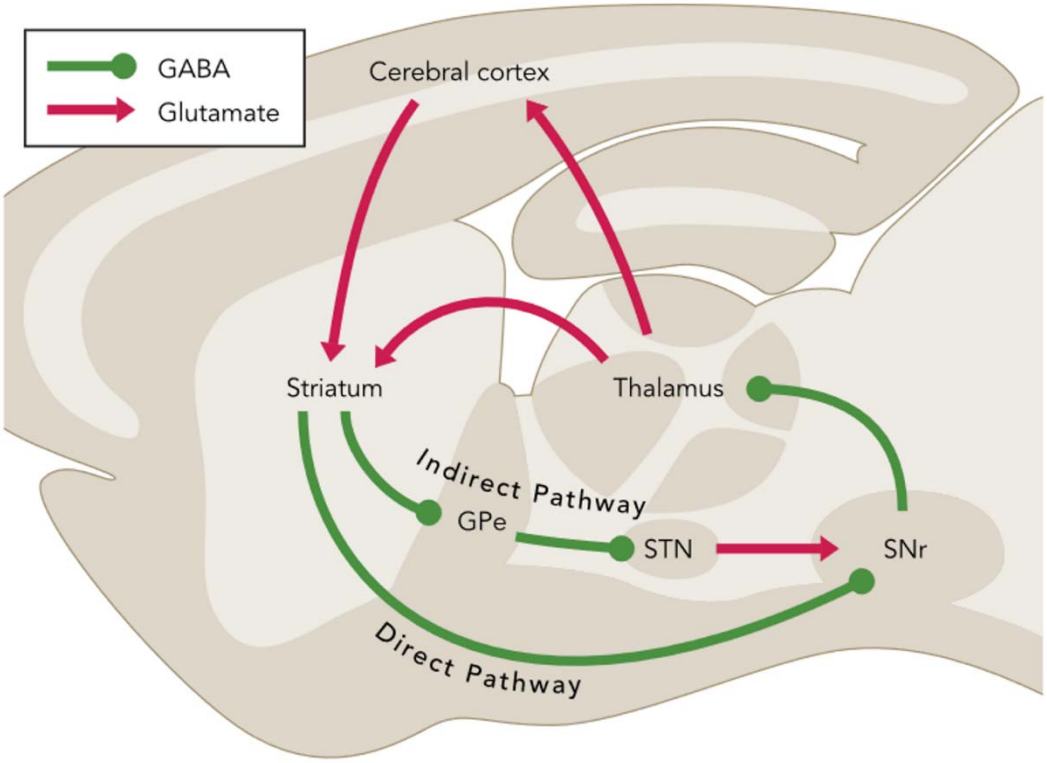


Figure 1.2 Anatomical organization of cortico-basal ganglia loop

Adapted from (Kravitz and Kreitzer 2012)

1.2.2 Action initiation

Both the cortex and the striatum are involved in the generation of voluntary actions. In both primate premotor cortex and rat secondary motor cortex, neuronal firing is coupled to the initiation of voluntary action (Tanji and Kurata 1985; Romo and Schultz 1992; Kurata and Tanji 1985; Mita et al. 2009, Murakami et al. 2014). In the primate pre-supplementary motor area, preparatory activity of a movement is observed when monkeys are trained to respond based on internally generated time estimates (Mita et al. 2009). These neural activities can be preparatory signals for specific actions (Romo and Schultz 1992; Erlich et al., 2011; Guo et al., 2014) or more general signal for urgency or impulse to act (Reddi and Carpenter 2000; Churchland et al. 2008; Thura et al. 2012), involvement of those medial motor and prefrontal cortical areas in action initiation is suggested.

On the other hand, the striatum, especially the nucleus accumbens is potentially involved in action initiation in the context of reward-seeking behavior. In rat nucleus accumbens, a pause in neuronal firing is observed preceding the onset of feeding behavior (Krause et al. 2010). This study also showed that the electrical stimulation of the NAc disrupted consummatory licking behavior. This observation indicates the causal involvement of the NAc in action initiation. In addition to consummatory behavior, the NAc is also important for reward associated sensory cues to invigorate instrumental behavior. NAc lesions in rats disrupt the performance in Pavlovian to Instrumental Transfer (PIT); the behavioral paradigm used to assess the ability of reward associated sensory cues to drive instrumental actions (Corbit and Balleine 2011). In the context of addiction studies, drug

seeking behavior induced by drug-associated cues or priming injection of drugs is associated with the action initiation process and depends on the NAc (Parkinson et al., 1999). These behavioral studies suggest that the NAc is also involved in the process of action initiation. While the cortical contribution to action initiation is often associated with specific motor responses, actions that the NAc drives are more related to specific outcomes.

Another example of cognitive control involved in action initiation is impulse control. Impulsivity is defined as the inability to inhibit inappropriate actions and the inability to wait for the outcome (Robbins 2011). Human patients with damages in the PFC or attention deficit hyperactivity disorder (ADHD) patients become impulsive and unable to inhibit inappropriate responses (Schachar et al. 1995; Aron et al., 2003). In reward based decisions, those patients have deficits in discounting reward value-based on the temporal proximity (Scheres et al. 2008). Consequently, subjects tend to choose the smaller amount of reward that can be obtained immediately, compared to the larger amount of reward obtained after delay. Lesion of rat orbitofrontal cortex (OFC) lesion affected the degree to which rats discount reward value that is obtained after delay (Winstanley et al. 2004, Rudebeck et al. 2006, Mar et al. 2011, Stopper et al. 2012) However, the increase or decrease of impulsivity by lesion varies across different studies. Impulsivity regarding response inhibition is modeled by the number of premature responses in a choice task such as 5-chocie serial reaction time task (5-CSRTT) (Chudasama et al. 2003). In 5-CSRTT, animals have to respond to a visual cues within a predetermined time window. Rats with infralimbic cortex lesion showed an increased number of premature responses (responses before stimulus onset) (Chudasama et al. 2003). OFC lesion also increased

premature responses under some task conditions. Although there is some inconsistency across studies, they strongly suggest a contribution of the PFC to the control of impulsivity.

The process of action initiation has also been studied in the context of addiction. Those who are addicted to drugs or alcohol behave prematurely without assessing the long term consequences of drug or alcohol taking behavior (Jentsch and Taylor 1999, Robinson and Berridge 2003; Fillmore and Rush 2001; Monterosso et al. 2005).

1.2.3 Action selection

Multiple areas of the cortico-basal ganglia circuits are involved in the action selection process. In both rodent and primate, neurons that encode actions are present throughout the cortico-basal ganglia circuit (Feierstein et al. 2006; Samejima et al.2005; Sul et al. 2010; Seo et al. 2014). Samejima et al. recorded caudate neurons in monkeys performing a choice task whose action values are dynamically changed. In their study, about one-third of neurons dynamically tracked the value of each action, indicating that action value is estimated in the brain as in the reinforcement learning model (Samejima et al. 2005). Neurons in the primate rostral cingulate motor areas change activity preceding action selection when a monkey changes the action after observing a reduced reward for the previous action (Shima and Tanji 1998). However, the same neurons did not respond when a reward is constantly given to the action. The contextual modulation of neuronal activity indicates that such neuronal activity reflects a selection process rather than a motor command for the specific movement.

From perturbation studies in rodents, other prefrontal and areas such as the anterior cingulate cortex (ACC), the secondary motor cortex (M2), the medial prefrontal cortex (mPFC) and the orbitofrontal cortex (OFC) have also been implicated in goal-directed action selection (Tervo et al. 2014, Sul et al. 2011, Ostlund and Balleine 2005, Gremel and Costa 2013). The information about movements, action values and outcome values is also represented in the premotor areas and mPFC both in rodents and primates (Campos et al. 2005; Matsumoto et al. 2007; Sul et al. 2011, Murakami et al. 2014). Perturbations in the M2 and the ACC were reported to affect outcome dependence of choices in rats (Sul et al. 2011; Tervo et al. 2014). In the striatum, flexible goal-directed actions depend on the dorsomedial striatum (Yin et al. 2005) and inflexible habitual actions depend on the dorsolateral striatum (Yin et al. 2004). The optogenetic activation of the dorsomedial striatum in mice affected the choice probability of options that is associated with the different probability of reward (Tai, Lee et al. 2012). These studies suggest that a wide range of cortico-striatal areas are involved in action selection.

Abnormal action selection is characteristic of the state of compulsivity that is often associated with prefrontal damage or pathological states such as obsessive-compulsive disorder (OCD). Compulsive individuals take actions that don't lead to desired consequences. Perturbations of several cortical and subcortical regions were reported to affect the level of compulsivity. Chronic photostimulation of the medial OFC or axonal projection from the medial OFC to the ventromedial striatum increased the frequency of innate grooming that is a model of compulsive behavior in rodents (Ahmari et al. 2013). Acute photostimulation of the lateral OFC in Sapa3 mutant mice rescued maladaptive grooming behavior acquired as a conditioned response (Burguière et al. 2013).

Compulsivity has also been studied in the context of addiction. Various pharmacological agents are known to affect compulsive behavior (Vanderschuren and Everitt 2004; Mundt et al., 2009; Barker et al. 2014). Even without brain damage or pharmacological agents, behavior can become compulsive as is observed in feeding behavior (Johnson and Kenny 2010; Smith et al., 2014). Deep brain stimulation or pharmacological treatment of the NAc shell reduces the compulsive eating (Halpern et al. 2013; Smith et al., 2014). These findings suggest the relevance of specific brain areas for action selection and compulsivity.

1.2.4 Learning

Because of the dense projection of midbrain dopaminergic neurons to the ventral striatum, the ventral striatum is supposed to play a key role in modifying action or stimulus value-based on reward prediction error signals. This view is supported by a human fMRI study showing that the ventral striatum displayed activity that is correlated with reward prediction error (O'Doherty et al. 2004). Learning through reward prediction error was directly tested by a Pavlovian value blocking procedure (Holland 1984). Lesion of rat ventral striatum abolished the value unblocking effect, indicating that Pavlovian learning through prediction error of value is disrupted (McDannald et al. 2011). Lesion of the NAc also impaired the acquisition of instrumental learning (Corbit et al. 2001; Atallah et al. 2006).

The cortical contribution of learning has also been studied, focusing on the mPFC. The sensitivity of instrumental actions to value is disrupted by pre-training lesion of the PL but not by the post-training lesion (Ostlund and Balleine 2005). This work suggests that

the contribution of the PL to the learning of value. In a reaction time task, in which subjects have to respond within a specific time window after sensory cue onset, post-error adjustment of the behavior is observed (Naranayan et al. 2013). A contribution of the mPFC to this aspect of behavior is suggested both in human and rodent studies (Narayanan and Laubach 2008, Narayanan et al. 2013). A reward prediction error signal observed in primate mPFC also suggests a contribution of this area to reward-based learning (Matsumoto et al. 2007).

1.3 Decision making for uncertain outcomes

Decisions that we face in our daily lives and animals' decisions in natural environments are often accompanied by changes in the rules or contexts. Such uncertainty has not been addressed in paradigms such as classical instrumental conditioning or Pavlovian conditioning, which has been used for studying cortico-basal ganglia circuit.

For decisions under uncertainty, different neural systems are recruited compared to decisions under stable conditions (Cools et al. 2002). Also, brain areas involved in decisions under stable conditions can behave differently in uncertain situations (Durstewitz, Vittoz et al. 2010; Karlsson et al. 2012). In the OFC and the posterior cingulate cortex of primates, a subset of neurons represents reward risk independent from reward values (McCoy and Platt 2005, O.Neill and Schultz 2010). The differential representation of reward values and risks indicates that an uncertain situation is not exactly same as the sequence of different deterministic events in terms of neuronal representation. Such decisions under uncertainty are more vulnerable in pathological

situations such as Schizophrenia and Parkinson's disease compared to the same decisions under stable environment (Knowlton et al. 1996).

In behavioral paradigms for value-based decision making, one approach to introducing uncertainty is to make outcome delivery stochastic and to change the action-outcome contingency within a behavioral session. If animals receive a reward only in part of the trials, they are required to learn from both positive and negative outcomes. The repeated reversal or contingency change of the action-outcome association creates uncertainty and is advantageous in making animals learn continuously without forming habitual responses. Thus, choice tasks for probabilistic reward are the useful approach to studying value-based decision making.

Several studies have conducted electrophysiological recording studies in rats and primates during probabilistic choice tasks. These studies identified cortical and striatal neurons that represent different aspects of a decision such as movements, action values, and outcome values (Sul et al. 2009, Ito and Doya 2009, 2015, Samejima et al.2005). Inactivation of the NAc shell using GABA receptor agonist muscimol reduced the probability of staying in the same choice after receipt of reward (Dalton 2014). Several studies addressed the contribution of the serotonergic system to similar tasks. However, there is inconsistency among different studies probably because of differences in experimental conditions (Bari et al. 2010, Fonseca et al. 2015). Risk-based decision making in which subjects have to choose between small certain reward and large uncertain reward is also used to study how subjects respond to stochastic outcomes. In rats, muscimol inactivation of the medial OFC caused an increased tendency to choose large uncertain reward along with the increased probability of choosing the same option

after rewarded trials (Stopper et al. 2012). Muscimol inactivation of the mPFC caused a similar effect while inactivation of the OFC did not affect risk preference (St Onge and Floresco 2010). However, another study reported opposite effect in OFC lesioned rats (Abela and Chudasama 2013).

1.4 Approaches used in the current study

The choice task for stochastic reward is not only useful for assessing the process of action selection and learning but also for studying action initiation. Since the task is composed of temporally distinct trials, behavioral measures such as time to initiate each trial or the frequency of responding prematurely before choices becoming available can be used to study the action initiation process. These behavioral measures are used to study impulsivity in rodent 5-choice serial reaction time task (Chudasama et al. 2003). In another study, the ability of rats to refrain from responding is used to study the mechanism for initiation of voluntary actions (Murakami et al., 2014). For these reasons, we employed the trial based probabilistic choice task in the current study to dissociate cognitive functions using different behavioral measures.

We chose mice as a model animal in this study rather than rats or primates which are more commonly used in studying value-based decision making. The following reasons factored into this decision. First, the availability of genetic tools such as transgenic mice and viruses, allows the manipulation of neuronal activity in the specific elements of neural circuits. Second, the smaller brain size is more suitable for perturbation experiments, especially silencing experiments to show the causal involvement of a specific brain area. Since the brain consists of numerous different cell types that form a

complex network, the selective manipulation of cell types or specific projections is critical to elucidate neural substrates for decision making.

We utilized optogenetics, a method to alter neuronal activity using light-sensitive ion channels or pumps, to manipulate neuronal activity (Boyden et al. 2005; Han, Chow et al. 2011). This approach provides better temporal resolution and accessibility to specific components of neuronal circuits compared to other methods such as excitotoxic lesion or pharmacological perturbation. Perturbation of neuronal activity during a specific time window is necessary both to dissociate cognitive processes and to avoid compensatory changes in gene expression or synaptic plasticity in neural circuits that may occur during chronic perturbations. Although pharmacological methods have also been used for transient manipulation of neuronal activity, even 1 or 2 hours of manipulation of activity is enough to cause compensatory circuit changes (Goeshen et al. 2011). For these reasons, even if behavioral paradigms contain behavioral measures of different cognitive processes, manipulation of neuronal activity on a slower time scale can hamper interpretation of underlying mechanisms. Therefore, the temporal resolution available in optogenetics is critical for our purpose.

1.5 Unanswered questions and aims of the current study

As described in this introduction, the involvement of the PFC, the NAc, and areas connected with these areas in value-based decision making have been suggested. However, there are issues that are remained to be answered.

The contribution to action initiation has been suggested in the premotor cortex, the PFC and the NAc (Chudasama et al. 2003; Corbit and Balleine 2011; Murakami et al. 2014). However, behavioral changes observed in some behavioral paradigms such as PIT can also be induced by changes in action selection and learning. Therefore, clear dissociation of different cognitive processes needs to be performed under a single behavioral paradigm. Changes in neuronal activity during the period of action initiation were widely observed in the PFC and the NAc along with their input and output structures including the VTA (Mirenowicz and Schultz 1994; Narayanan et al. 2008; Ito and Doya 2015). However, the causal involvement of such neuronal activity has been understudied. Therefore, it is critical to perform transient perturbation of neuronal activity in the PFC and the NAc during specific time periods. Another question that is remained to be answered is how these areas interact each other. Since those areas have both direct and indirect interaction, it is crucial to investigate the involvement of specific projections or cell types in these areas.

Although the contribution of the dorsal striatum to action selection is well established (Samejima et al. 2005; Tai and Lee et al. 2012), the contribution of input structures to the dorsal striatum has not been extensively studied. The PFC is one of the primary input structures to the dorsal striatum, and its contribution to action selection has been suggested from electrophysiological recording studies (Sul et al. 2010; Kennerly et al. 2011). However, neuronal representations of task-related events in the PFC are heterogeneous and are often redundantly observed in different subregions in the PFC (Sul et al. 2010; Kennerly et al. 2011). Although the contribution of the NAc to action selection has been understudied, some neurons in the NAc did respond to specific choices

(Ito and Doya 2008; Ito and Doya 2015). Therefore, it is critical to investigate the causal involvement of different subregions of the PFC and the NAc.

The existence of reward prediction error signals in the VTA and dense projection from the VTA to the NAc indicate the involvement of these structures in learning. However, the causal involvement of specific cell types, inputs, and outputs of these structures are remained to be understood. Therefore, it is essential to test the contribution of specific cell types or specific projections to learning under a single behavioral paradigm in mice.

Utilizing approaches described in the previous section, we aimed to dissect the contribution of the PFC and the NAc to action initiation, action selection, and learning processes in value-based decision making. In Chapter 2, we established a behavioral paradigm in mice to study value-based decision making. In Chapter 3, we performed the inactivation of the different circuit elements in the NAc while mice performed the behavioral task. This chapter was intended to achieve anatomical and functional dissociation of action initiation, action selection and learning in the NAc. In Chapter 4, we performed inactivation of each of the OFC and the ACC using the same behavioral task as Chapter 3. This chapter was intended to reveal the functional distinction of PFC subregions. Through these experiments, we hoped to clarify the contribution of the PFC and the NAc to each of three cognitive steps. We also hoped to identify the circuit elements for these cognitive processes using area, projection and cell type specific manipulation of neural activity.

Chapter 2: Development of Probabilistic Reversal Task

An appropriate behavioral paradigm to study value-based decision making needs to meet following criteria. (1) Animals must be required to change actions frequently based on choice and outcome history. (2) Different aspects of the decision must be dissociable using distinct behavioral measures. Therefore, we decided to develop a trial based choice task with stochastic reward and reversal. The trial based organization of the task enables dissociation of action initiation, action selection, and learning. The stochastic delivery of reward and within session reversal contribute to choice flexibility over long behavioral sessions so that a large amount of data can be collected from individual animals.

2.1 Subjects and configuration of the operant chamber

All training sessions and test session were performed in a sound and light attenuated operant chamber (Med Associates). The operant chamber was equipped with a fan for ventilation, and it provides background white noise. Inside the chamber were illuminated using house light. One nose-port was placed in the middle of one wall, and two food magazines were placed on the left and right sides (Figure 2.1). Two food magazines were connected with food pellet dispensers that deliver a 14mg food pellet (BioServe). Control of stimuli in the chamber and data acquisition were performed using programs written for MedState Notation (Med Associates).

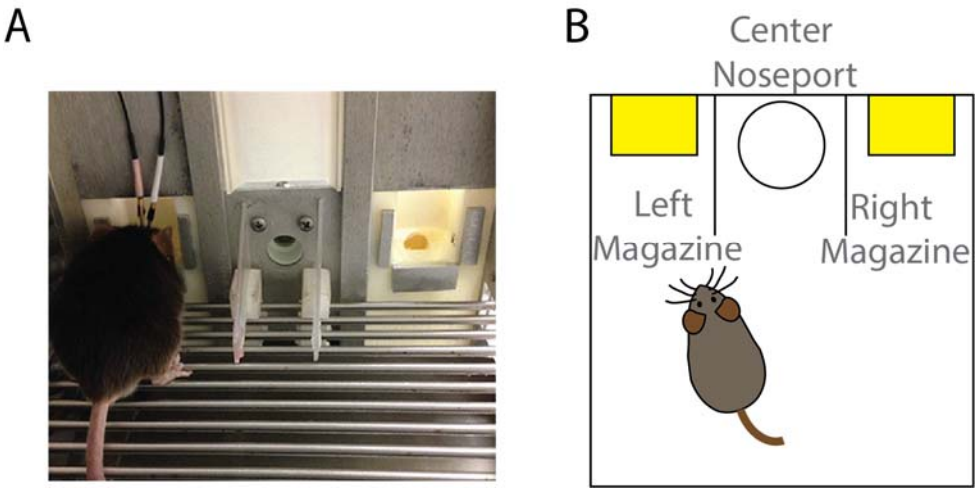


Figure 2.1 Configuration of the operant chamber.

(A) Snapshot of a mouse in the operant chamber. (B) Schematic of the operant chamber shown in A. One nose-port was placed in the center of one wall, and two food magazines were placed on the left and right sides.

2.2 Experimental design of the probabilistic reversal task

In order to test the ability of mice to learn appropriate actions through trial and error, a probabilistic reversal task was developed. In a probabilistic reversal task, mice have to choose either left or right food magazine at each trial. These choices are associated with different reward probability (75% for one side and 0% for the other), and this arrangement was switched several times in a daily session. Mice were trained to obtain the maximal amount of reward through trials and errors.

All of the behavior experiments were performed in the operant chamber. Trial start was signaled by the illumination of an LED inside the center nose-port and mice were required to make a nose-poke to the center nose-port to initiate each trial. After a center

nose-poke is made, LEDs inside left and right magazines are turned on to signal the time window of choice. Once mice break the infrared beam inside the magazine, a food pellet was delivered at the pre-determined probability. One of the two food magazines is assigned as the correct food magazine and the other side as the incorrect food magazine. Choices of the correct food magazine were rewarded at the probability of 75%, and choices of the incorrect food magazine were never rewarded. Once animals reached a performance criterion (80% correct choices over the last ten trials), the correct and incorrect sides were stochastically switched. This switch was called a 'reversal' and occurred with 15% probability at the beginning of each trial. After the reversal, the positions of the correct and incorrect sides were kept constant until animals reach the performance criterion again. The stochastic nature of reward delivery and block transition make animals explore the better options through trial and error. The absence of reward after incorrect choices or 25% of correct choices were signaled by a click sound. The left or right food magazine entry were recorded as a choice and followed by 4.5sec inter-trial interval (ITI). Left or right magazine entries after center LED onset in the next trial were recorded as 'premature responses' and punished by 3 seconds time out. Failure to make a center nose-poke within 20sec after the onset of center LED is recorded as 'Trial omission' and also punished with a 3-second time out.

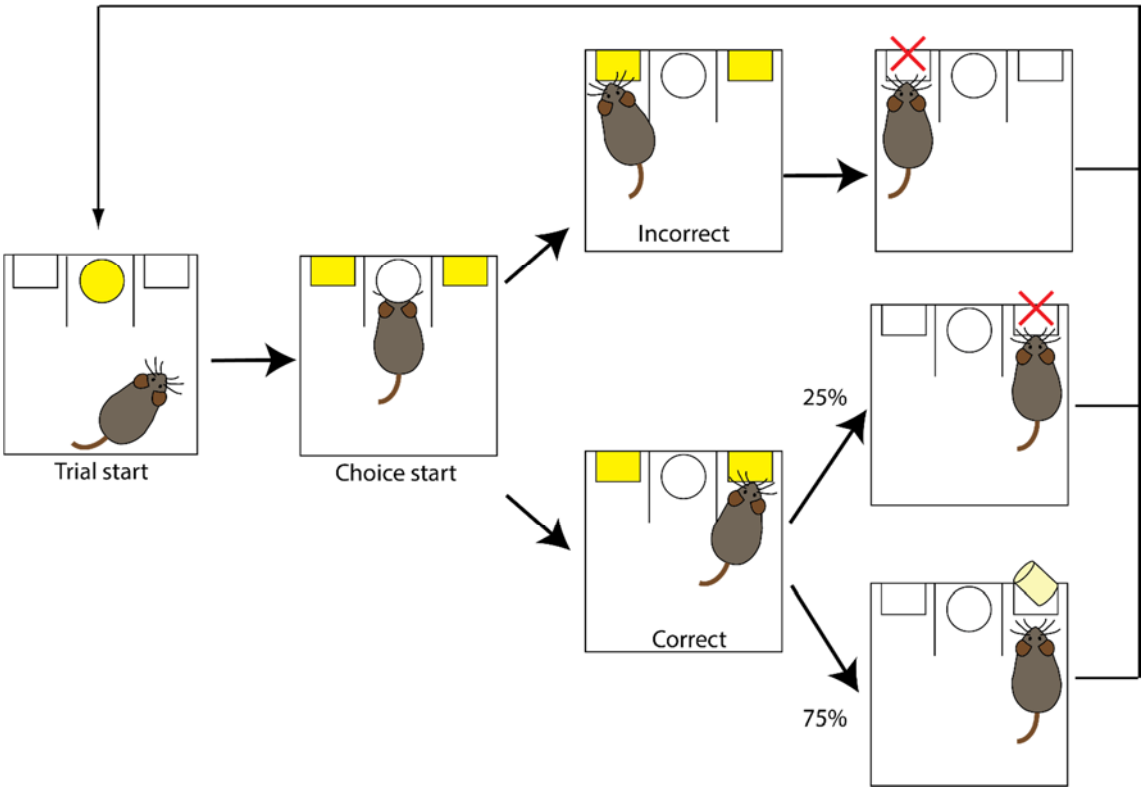


Figure 2.2 Probabilistic reversal task

The schematic shows the flow of a single trial. Trial start was signaled by the onset of a center LED, and a mouse can initiate trial by making center nose-poke. Then, LEDs in the left and right food magazine were turned on, and the mouse can choose either the left or right food magazine. Subsequently, a reward was delivered at predetermined probability, and the next trial started after an inter-trial interval. Reward probability was switched after mice reaching a performance criterion.

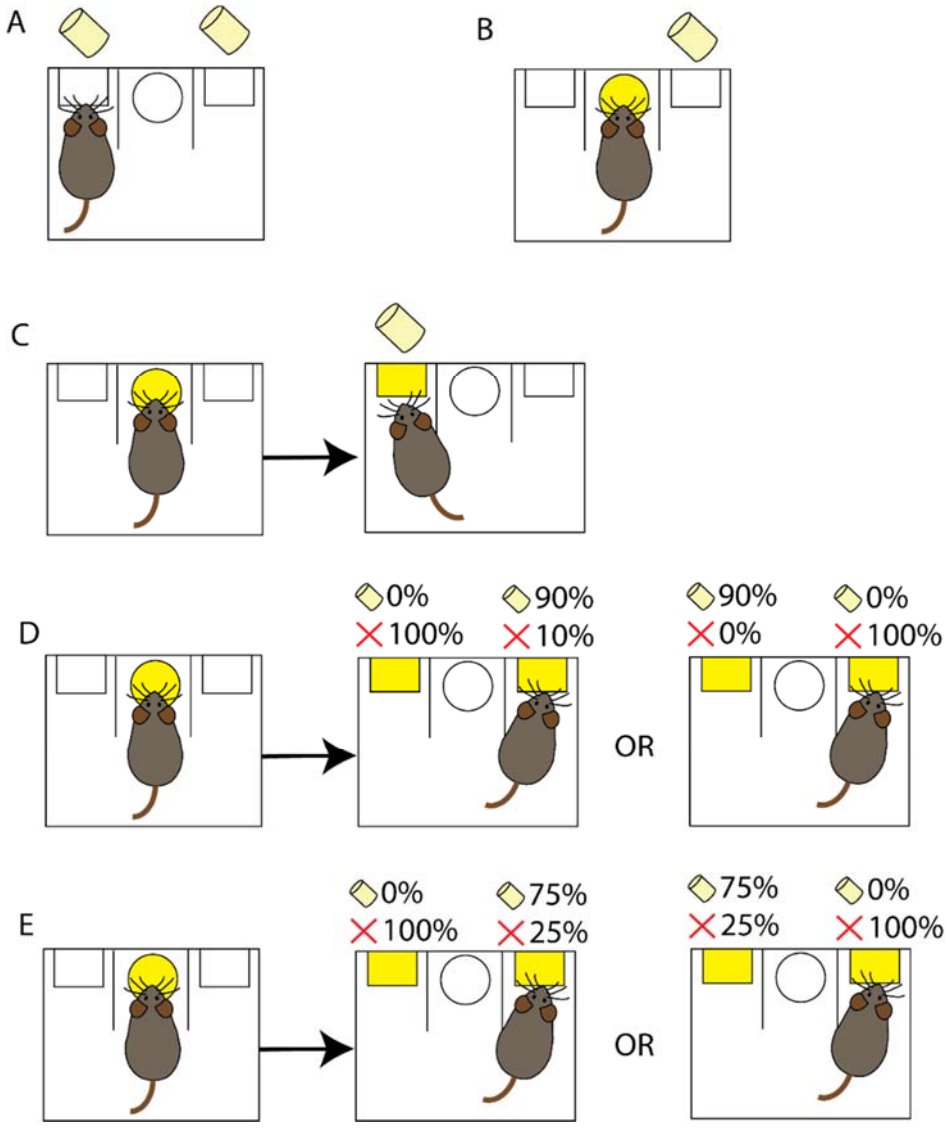
2.3 Training steps and learning efficiency

Animals were food restricted until their body weights became stable at 80-90% of their normal body weights. Animals were first trained to collect food pellet from the left and right food magazines (Magazine training, Figure 2.3A). Animals were next trained to make center nose-pokes (Step 1, Figure 2.3B). In step 1, center nose-pokes trigger food pellet delivery from either the left or right food magazine. The position of food pellet delivery was randomly determined at each center nose-poke. The next step required animals to initiate trials by making a center-nose-poke and enter food one of the two food magazines (Step 2, Figure 2.3C). In this step, center nose-pokes turned on the LED in either left or right magazine but this didn't lead to pellet delivery. Entries to the illuminated food magazine led to a food pellet delivery. Entries to the unilluminated food magazine were punished with a 3-second time-out (House light extinguished). The position of the illuminated and unilluminated food magazines were randomly determined at each center nose-poke. In the next step, behavioral requirements were the same as the probabilistic reversal task except that reward probability of correct choice was set to 90% (Step 3, Figure 2.3D). Once animals achieve three or more reversals in a single behavioral session, animals were moved to the final step. As the final step of training, animals were trained with the same condition as the final probabilistic reversal task (75% reward probability for correct choices). However, inactivation was not performed in this step (Step 4 Figure 2.4D). Once animals achieved three or reversals in a single behavioral session, animals were tested in optogenetic inactivation experiments described in Chapter 3 and Chapter 4.

Most animals completed magazine training in one day (Figure2.4A), although some animals were anxious in a novel environment or had difficulty in collecting food pellets from food magazines, they learned to collect food pellets after additional magazine training. Although some animals spent more than 10 days to pass a single step, all animals completed training steps within 30 days (Figure 2.4 B-F).

Figure 2.3 Training steps of the probabilistic reversal task.

(A) Magazine training: Food pellets were delivered to both food magazines every minute regardless of entries to food magazines. (B) Step 1: Food pellets were delivered when animals made nose-pokes to the illuminated center nose-port. The position of the food pellet delivery was randomly chosen for each center nose-poke. (C) Step 2: Center nose-pokes turned on an LED in either the left or right food magazine. Food pellets were delivered after entries to the illuminated food magazine. (D) Step 3: Center nose-pokes turned on LEDs in both left and right food magazines. A food pellet was delivered after a magazine entries based on pre-determined probability (90% vs. 0%). The probability was switched once animals learn to choose the side with the higher reward probability. (E) Step 4: The same procedure as Step3 but reward probabilities were changed to 75% and 0%.


A
B
C
D
0%
100%
90%
10%
OR
90%
0%
0%
100%
E
0%
100%
75%
25%
OR
75%
25%
0%
100%

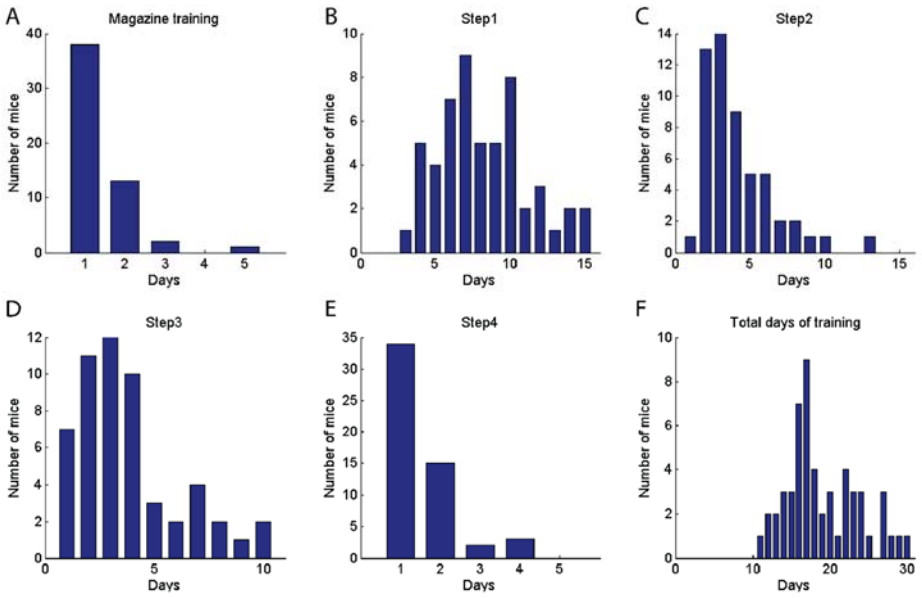


Figure 2.4 Number of days spent on each training step.

Number of days spent on each training step is displayed as histogram (A) Magazine training (B) Step 1 (C) Step 2 (D) Step 3 (E) Step 4 (F) Total training duration

2.4 Baseline behavior performance and behavioral measures to be tested

In the last session of training step 4, animals flexibly change their choices between two options based on choice and outcome history. Animals could respond to the reversal of the correct option. The probability of making a correct choice (choice with high reward probability) was high at trials preceding reversals. Since reversals were not signaled, the correct choice probability was decreased after reversal but it gradually increased over trials (Figure 2.5A). The contribution of choice and outcome history to the next choice

were investigated by the probability of making the same choice as the previous trial (Stay probability). The stay probability was higher after choice and outcome combination that predict the higher amount of reward (Figure 2.5B). Test sessions follow the last training session. The number of reversals, total number of trials of daily sessions, and the probability of correct choice were studied to see the stability of behavioral performance across test sessions. Although the number of reversal and probability of correct choice were slightly increased over sessions, these behavioral measures were relatively constant over sessions (Figure 2.5 C-E).

In addition to behavioral measures of baseline performance, several behavioral measures obtained through the probabilistic reversal task can be used to study action initiation and action selection. These behavioral measures include premature response, trial omission, start reaction time, and choice reaction time. Although these behavioral measures had an increasing or decreasing trend in first few sessions, the performance in adjacent sessions were stable (Figure 2.6).

These results showed that animals trained with the probabilistic reversal task made a choice based on choice and outcome history. In addition, their behavioral performance was stable across repeated behavioral sessions. These observations indicate that the probabilistic reversal task is an appropriate behavior paradigm for value-based decision making. In addition, the use of multiple behavioral measures obtained through the probabilistic reversal task enables functional dissociation of action initiation, action selection, and learning. In the following chapters, we selectively manipulate circuit elements in the PFC and the NAc to investigate their contribution to each cognitive process.

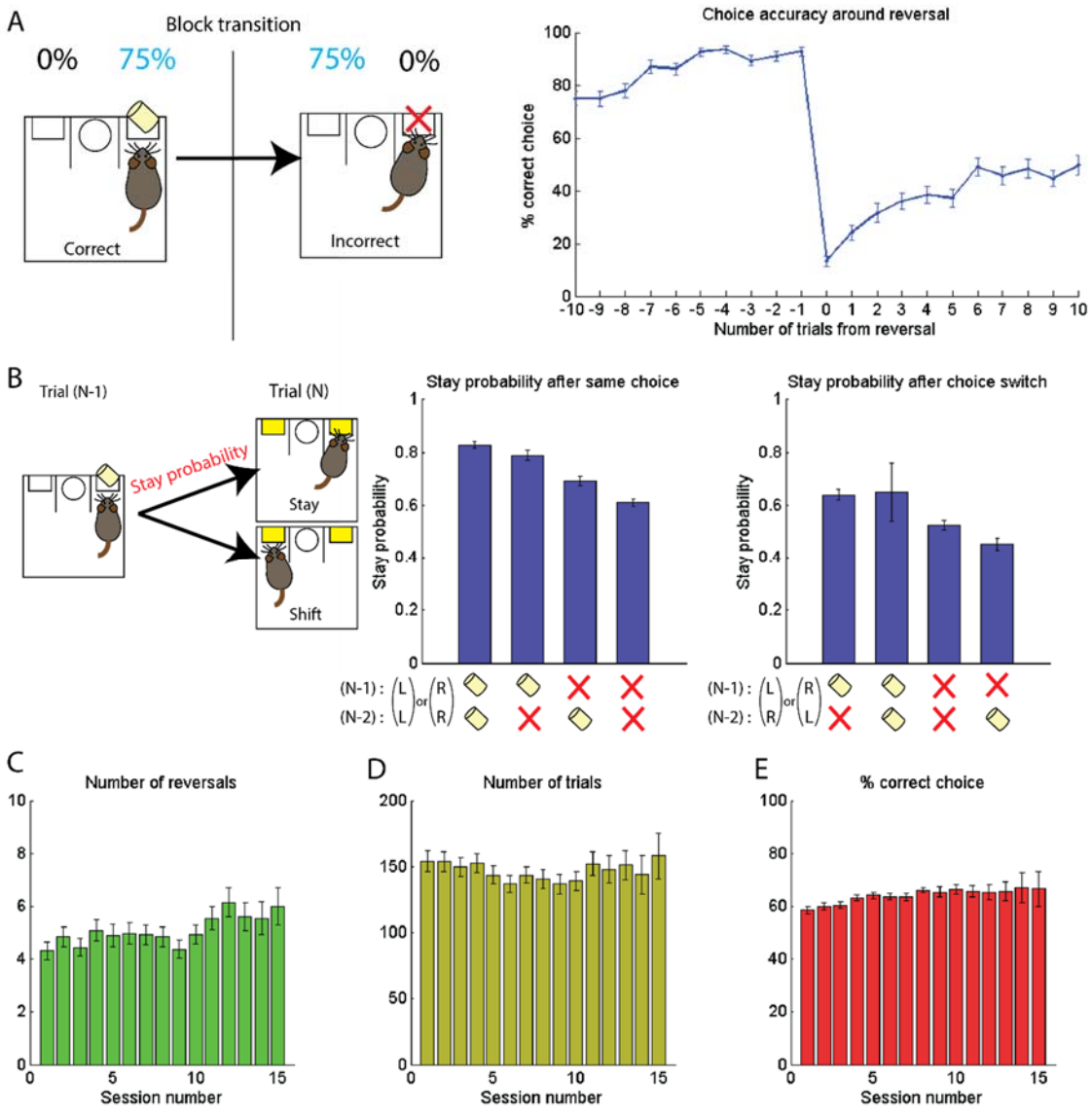


Figure 2.5 Baseline behavioral performance.

(A) Percentage of correct performance around reversal (Last session of 75% vs. 0%) (B) Choice and outcome dependence of stay probability (Last session of 75% vs. 0%) (C) Number of total trials during test sessions (D) Number of reversals during test sessions (E) Performance changes over training sessions during test sessions (n=59 mice)

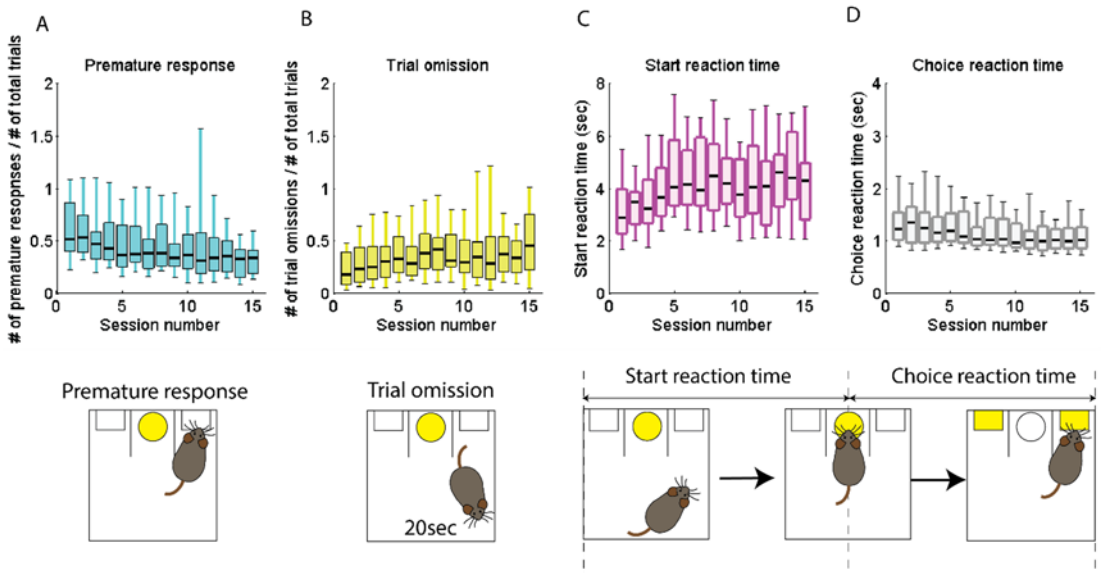


Figure 2.6 Across session changes of behavioral measures obtained from individual trials.

Daily change in behavioral measures. These behavioral measures were obtained for a daily session of each animal and plotted in box plots (A) The frequency of premature responses (# of premature responses / # of total trials) (B) The frequency of trial omissions (# of trial omissions / # of total trials) (C) Median start reaction time (D) Median choice reaction time (n=59 mice)

Chapter 3: The Nucleus Accumbens Is Required for Action Initiation and Learning

The NAc is implicated in a wide range of reward dependent behaviors. However, inactivation of neuronal activity in the NAc at fine temporal resolution has not widely studied. Since neurons in the NAc respond to a wide range of task-related parameters, inactivation experiments are especially meaningful to understand what these neural correlates mean. Therefore, we conducted optogenetic inactivation of the NAc while mice were performing the probabilistic reversal task to understand causal roles of the NAc in decision making and to give new insights into its anatomical basis.

3.1 Non-specific inactivation of the nucleus accumbens

In the probabilistic reversal task, each trial was divided into three different time periods (pre-start, pre-choice and post-choice period). The pre-start period begins from the onset of the center LED and ends at center nose-poke. The pre-choice period follows the pre-start periods and ends when either left or right food magazine is chosen. The post-choice period is 3 seconds following left or right choice. A 532nm laser was used to deliver green light through optical fibers during each of three time periods at 10% of trials in an interleaved manner. (Figure 3.1).

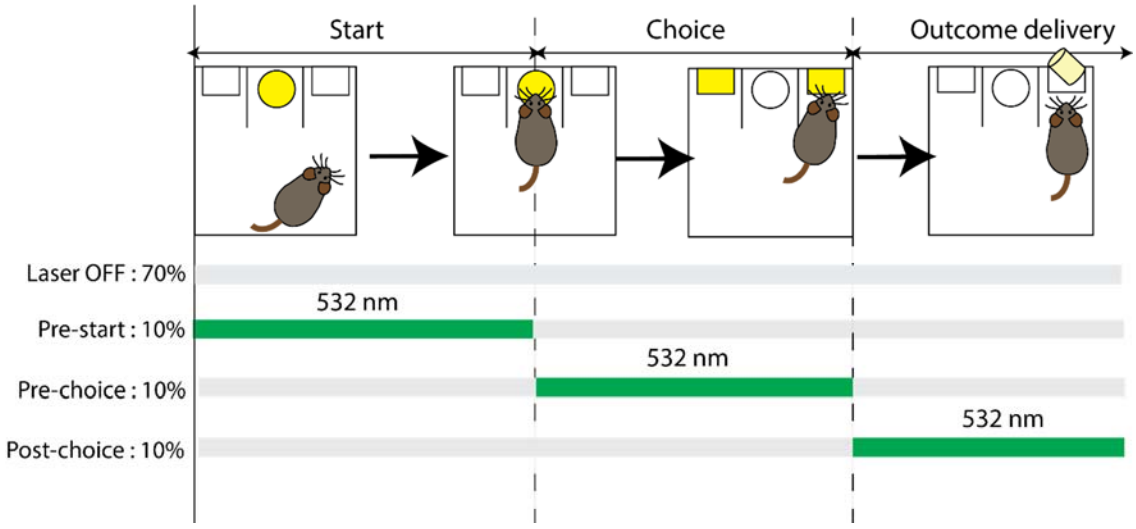


Figure 3.1 Timings of inactivation during each trial type.

During pre-start inactivation trials, the laser was turned on at the onset of center LED and turned off when animals nose poke to the center port. During pre-choice inactivation trials, the laser was turned on at the center nose poke and turned off when animals choose either left or right food magazine. During post-choice inactivation trials, the laser was turned on at the choice of left or right food magazine and last for 3 seconds. Each inactivation condition consisted of 10% of entire trials and placed at randomly interleaved order.

In the beginning, we tried to inactivate all cell types in the NAc together. In order to inactivate the NAc, we utilized archaerhodopsin (ArchT), a light gated proton pump driven by green and yellow light. Neurons in the NAc were transduced with Adeno-associated virus (AAV) that expresses ArchT under ubiquitous CAG promoter so that

ArchT is expressed in the NAc regardless of cell types. For the delivery of light to drive ArchT, optical fibers were implanted bilaterally into the core subregion of the NAc (Figure 3.2 A, B).

Inactivation of the NAc at these three time periods led to distinct behavioral changes. Pre-start inactivation affected a behavioral measure of action initiation. The frequency of premature responses was increased (Figure 3.2 C). However, the frequency of trial omissions and start reaction time were not significantly affected (Figure 3.2 D, E). The effect of pre-start and pre-choice inactivation on the following choice was quantified using stay probability as in Figure 2.5 B (Figure 3.2 F-H). First, stay probability was calculated for all trials, and pre-start inactivation showed the decreased tendency of stay probability (Figure 3.2 G). Since pre-start inactivation also increased the frequency of premature responses, this decrease can be the secondary effect of premature responses. To address this possibility, stay probability was calculated from the trials without premature responses in order to avoid the effect of time-outs given after premature responses. Stay probability was not decreased when trials with premature responses and trial omissions were excluded from the analysis (Figure 3.2 H). Pre-choice inactivation did not affect choice itself (Figure 3.2 G, H).

In contrast, post-choice inactivation affected the choice of the following trial (Figure 3.2 I, J). Stay probability was increased after unrewarded trials by post-choice inactivation (Figure 3.2 J) while same manipulation did not affect stay probability after rewarded trials. This effect was specific to post-choice inactivation. Pre-start inactivation or post-choice inactivation did not increase stay probability after unrewarded trials (Figure 3.2 G, H). The effect of inactivation on premature response also depended on the timing of the

inactivation. While pre-start inactivation caused increased the frequency of premature responses, post-choice inactivation decreased the frequency of premature responses (Figure 3.2 G).

In summary, non-specific inactivation of the NAc during pre-start and post-choice inactivation led to distinct behavioral changes. Increase in the frequency of premature responses induced by pre-start inactivation indicates the impact on action initiation. Increased stay probability after unrewarded trials with post-choice inactivation indicates the impact on learning in an outcome-specific manner. On the other hand, control animals that have fiber implanted into the NAc and received the same illumination of laser did not show changes in premature responses or stay probability (Figure 3.3)

Figure 3.2 Non-specific inactivation of the NAc.

(A) The target of inactivation. (B) A coronal section of a representative mouse brain. Circles illustrate the position of the NAc. (C) Effect of pre-start inactivation on premature response. (D) Effect of pre-start inactivation on trial omission. (E) Effect of pre-start inactivation on start reaction time. (F-H) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (G) Stay probability calculated from the entire trials. (H) Stay probability calculated from trials without premature responses and trial omissions. (I-J) Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (K) The frequency of premature responses measured after laser off or post-choice inactivation trials.

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (C), (D), (E), (K) Wilcoxon signed-rank test. (G), (H), (J) Chi-squared test. (n=9 mice, 107 sessions)

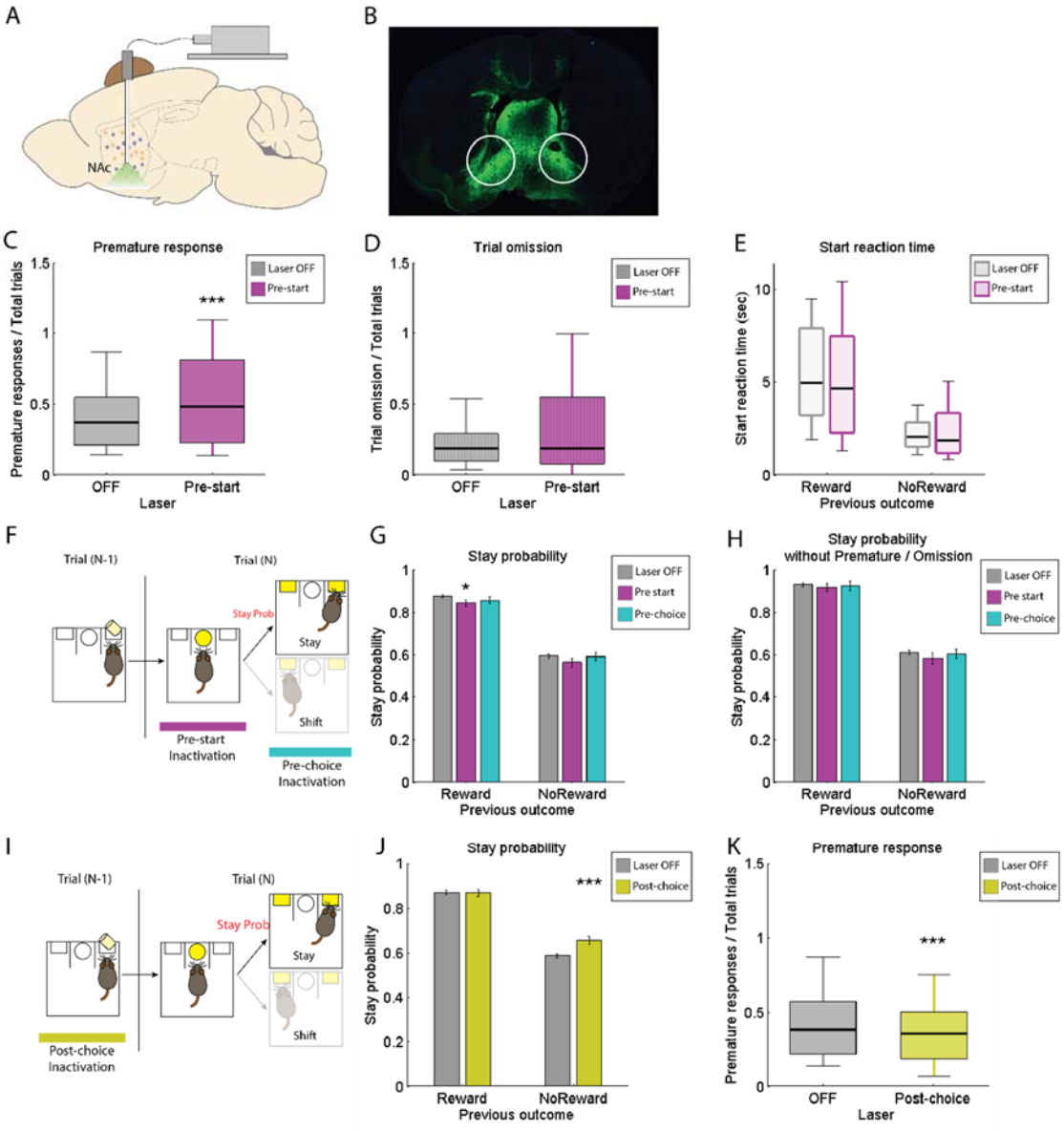
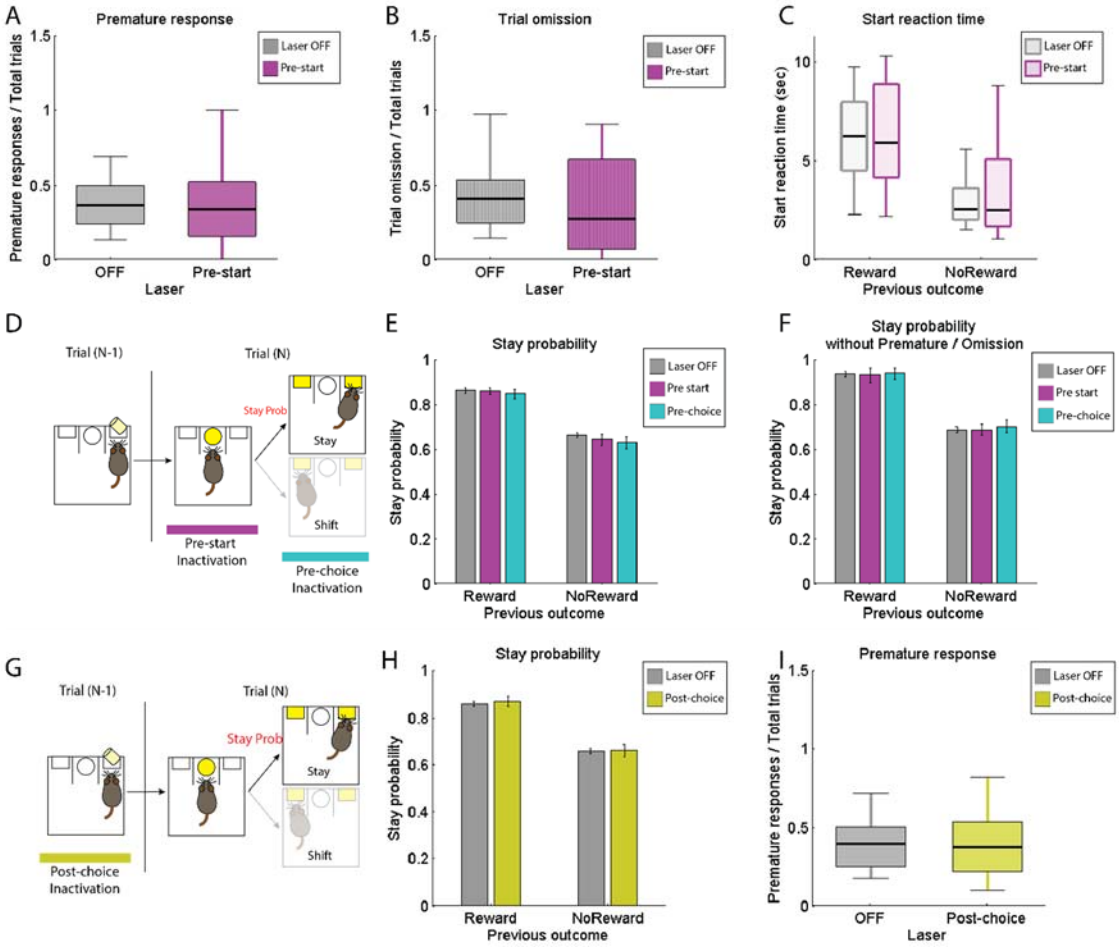

A
NAc
B
C
Premature response
Premature responses / Total trials
Laser OFF
Pre-start
***
OFF
Pre-start
Laser
D
Trial omission
Trial omission / Total trials
Laser OFF
Pre-start
OFF
Pre-start
Laser
E
Start reaction time
Start reaction time (sec)
Laser OFF
Pre-start
Reward
NoReward
Previous outcome
F
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start
Inactivation
Pre-choice
Inactivation
G
Stay probability
Stay probability
Laser OFF
Pre start
Pre-choice
*
Reward
NoReward
Previous outcome
H
Stay probability
without Premature / Omission
Stay probability
Laser OFF
Pre start
Pre-choice
Reward
NoReward
Previous outcome
I
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Post-choice
Inactivation
J
Stay probability
Stay probability
Laser OFF
Post-choice
***
Reward
NoReward
Previous outcome
K
Premature response
Premature responses / Total trials
Laser OFF
Post-choice
***
OFF
Post-choice
Laser

Figure 3.3 Control animals with optical fiber implant only.

(A) Effect of pre-start inactivation on premature response. (B) Effect of pre-start inactivation on trial omission. (C) Effect of pre-start inactivation on start reaction time. (D-F) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (E) Stay probability calculated from entire trials. (F) Stay probability calculated from trials without premature responses and trial omissions. (G-H)Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (I) The frequency of premature responses measured after laser off or post-choice inactivation trials.

$p>0.05$, (A), (B), (C), (I) Wilcoxon signed-rank test. (E), (F), (H) Chi-squared test. (n=5 mice, 73 sessions)


A
Premature response
Premature responses / Total trials
Laser OFF
Pre-start
OFF
Pre-start
Laser
B
Trial omission
Trial omission / Total trials
Laser OFF
Pre-start
OFF
Pre-start
Laser
C
Start reaction time
Start reaction time (sec)
Laser OFF
Pre-start
Reward
NoReward
Previous outcome
D
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start
Inactivation
Pre-choice
Inactivation
E
Stay probability
Stay probability
Laser OFF
Pre start
Pre-choice
Reward
NoReward
Previous outcome
F
Stay probability
without Premature / Omission
Stay probability
Laser OFF
Pre start
Pre-choice
Reward
NoReward
Previous outcome
G
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Post-choice
Inactivation
H
Stay probability
Stay probability
Laser OFF
Post-choice
Reward
NoReward
Previous outcome
I
Premature response
Premature responses / Total trials
Laser OFF
Post-choice
OFF
Post-choice
Laser

3.2 Inactivation of specific cell types in the NAc

Results of non-specific inactivation of the NAc motivated us to perform cell-type specific inactivation experiments. Since the direct pathway and indirect pathway medium spiny neurons (MSNs) are two dominant cell types in the striatum, those cell types were targeted.

3.2.1 Inactivation of direct pathway MSNs

Direct pathway MSNs were targeted using the Drd1-Cre (EY266) mouse line. Cre-dependent AAV expressing ArchT was injected into the NAc (Figure 3.4 A, B). Contrary to non-specific inactivation of the NAc, inactivation pre-start inactivation did not cause significant changes in behavioral measures of action initiation (Figure 3.4 C-E). Both pre-start and pre-choice inactivation did not cause significant changes in a choice of the current trial although there were increasing trends in stay probability after unrewarded trials (Figure 3.4 F-H).

On the other hand, post-choice inactivation caused a similar behavioral change in choice of the following trial as was observed in non-specific inactivation experiments. Stay probability after unrewarded trials was increased by post-choice inactivation of the previous trial (Figure 3.4 I, J). Post-choice inactivation was also accompanied by the reduced premature response in the following trial (Figure 3.4K).

3.2.2 Inactivation of indirect pathway MSNs

Indirect pathway MSNs were targeted using A2A-Cre mouse line. In order to cover large number of neurons, this mouse line was crossed with ArchT reporter mouse line (Ai40 line, Allen Institute), and optical fibers were implanted into the NAc (Figure 3.5 A, B). Pre-start inactivation led to a large increase in the frequency of premature responses and trial omissions (Figure 3.5 C, D). Pre-start inactivation had differential effects on start reaction time depends on the previous outcome. Pre-start inactivation decreased start reaction time after rewarded trials while it increased start reaction time after unrewarded trials (Figure 3.5 E). Stay probability was decreased by pre-start inactivation (Figure 3.5 F-G). Since stay probability of the trials without premature response and trial omission were affected less by pre-start inactivation, this effect was partly due to the increased number of premature response and trial omission rather than effect on choice itself (Figure 3.5 H). Stay probability was not affected by pre-choice inactivation, which were more proximate to the timing of choice (Figure 3.5 F-G).

Post-choice inactivation increased stay probability after the unrewarded trials (Figure 3.5 I-J). In contrast to post-choice inactivation of direct pathway MSNs, post-choice inactivation of indirect pathway MSNs did not affect premature responses on the next trial (Figure 3.5K).

In summary, behavioral measures of action initiation were differentially affected by inactivation of the direct pathway and indirect pathway MSNs. During pre-start period, only the inactivation of indirect pathway MSNs affected premature response and trial omission (Figure 3.4C-E, 3.5 C-E) suggesting the differential role of the direct and

indirect pathway for action initiation. Although a similar trend was observed in stay probability change after post-choice inactivation, the inactivation of direct pathway MSNs caused a larger increase (Figure 3.4J, 3.5J). The larger effect of pre-start inactivation of indirect pathway MSNs and the larger effect of post-choice inactivation of direct pathway MSNs suggest functional segregation of these cell types in the NAc.

Figure 3.4 Inactivation of direct pathway MSNs in the NAc.

(A) The target of inactivation. (B) A coronal section of a representative mouse brain. Circles illustrate the position of the NAc. (C) Effect of pre-start inactivation on premature inactivation on start reaction time. (F-H) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (G) Stay probability calculated from the entire trials. (H) Stay probability calculated from trials without premature responses and trial omissions. (I-J) Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (K) The frequency of premature responses measured after laser off or post-choice inactivation trials.

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (C), (D), (E), (K) Wilcoxon signed-rank test. (G), (H), (J) Chi-squared test. . (n=6 mice, 74 sessions)

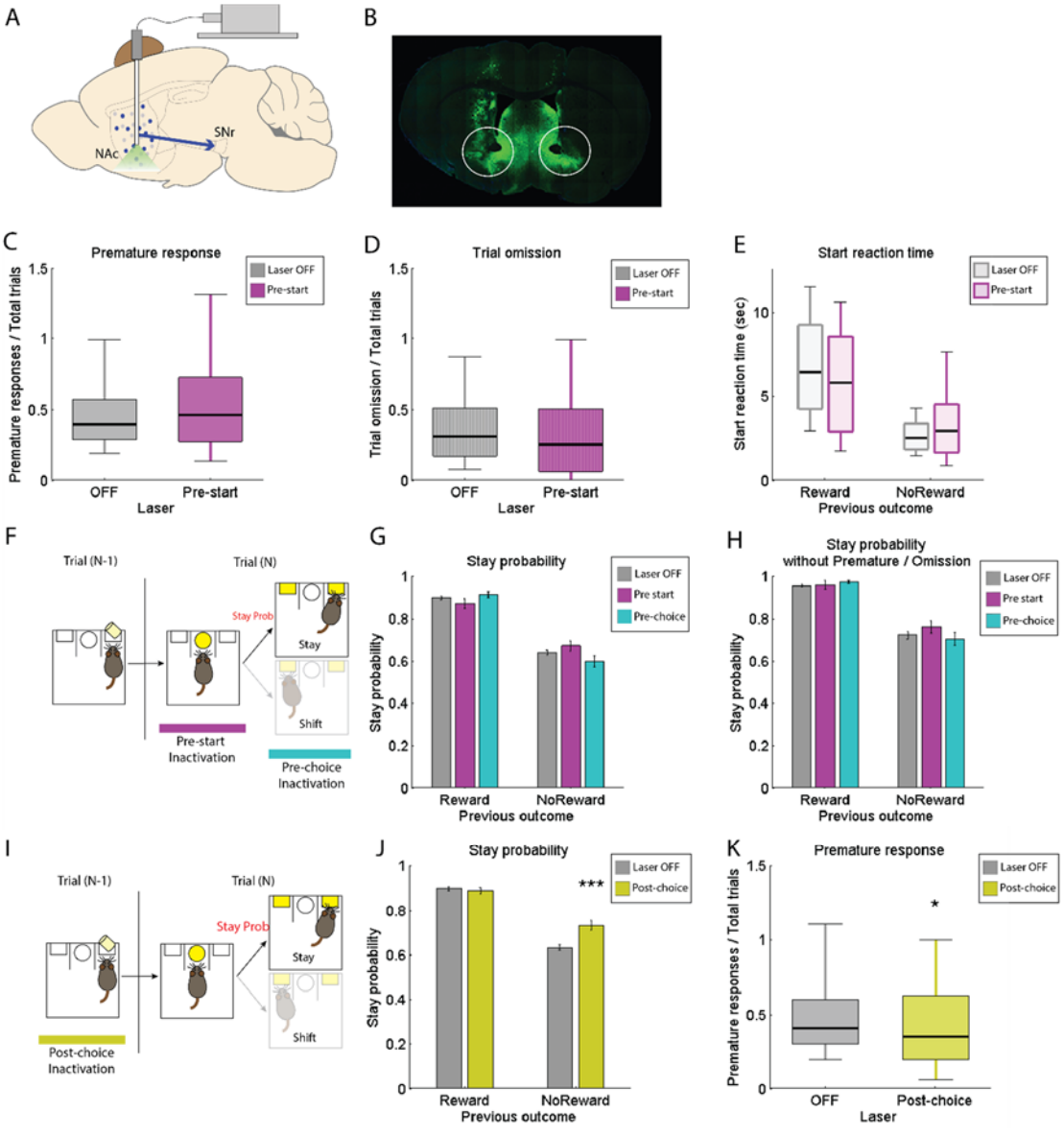
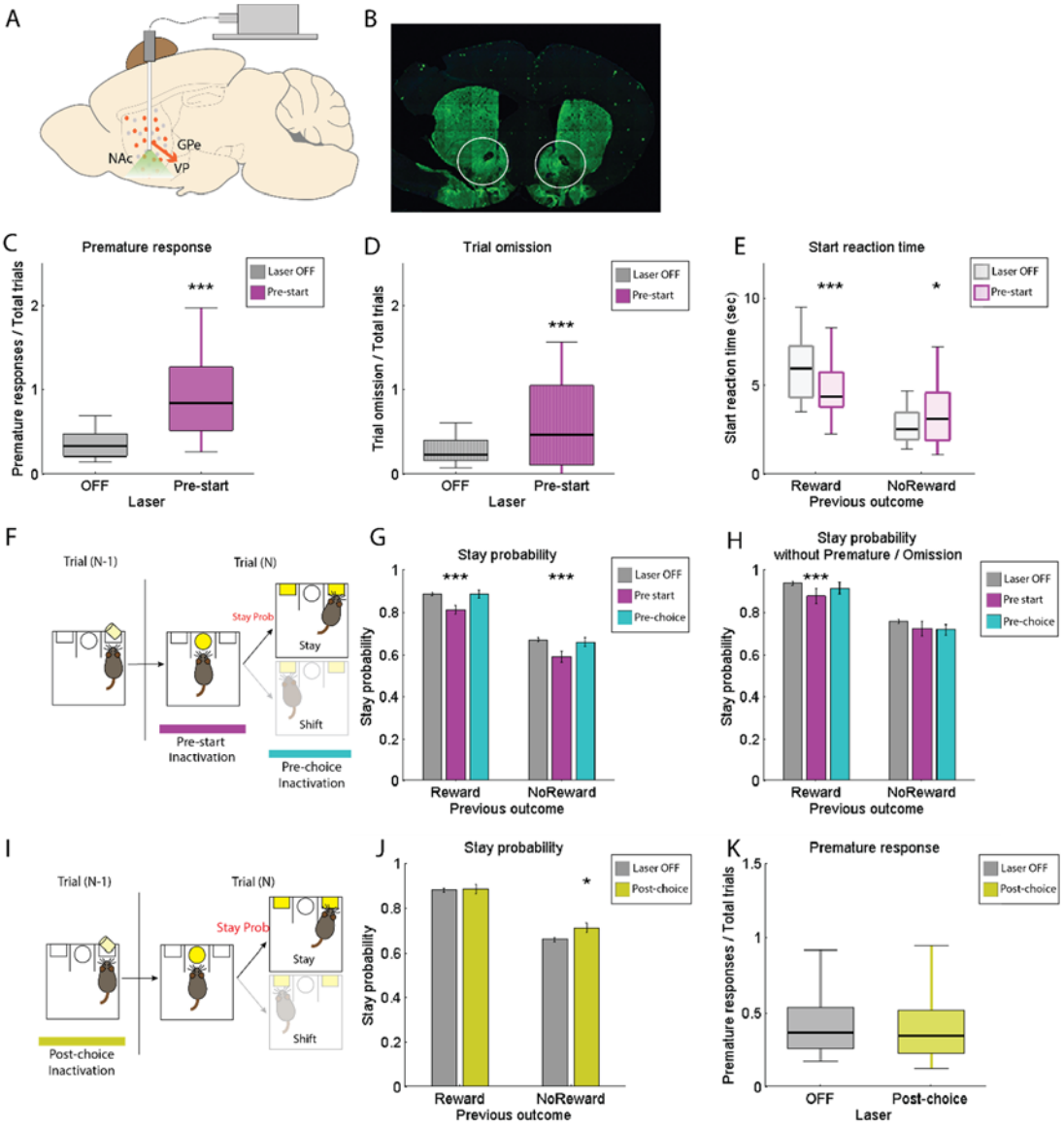

A
NAc
SNr
B
C
Premature response
Laser OFF
Pre-start
Premature responses / Total trials
OFF
Pre-start
Laser
D
Trial omission
Trial omission / Total trials
E
Start reaction time
Start reaction time (sec)
Reward
NoReward
Previous outcome
F
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start
Inactivation
Pre-choice
Inactivation
G
Stay probability
Pre start
Pre-choice
H
Stay probability
without Premature / Omission
I
Post-choice
Inactivation
J
***
Post-choice
K
*

Figure 3.5 Inactivation of indirect pathway MSNs in the NAc.

(A) The target of inactivation. (B) A coronal section of a representative mouse brain. Circles illustrate the position of the NAc. (C) Effect of pre-start inactivation on premature response. (D) Effect of pre-start inactivation on trial omission. (E) Effect of pre-start inactivation on start reaction time. (F-H) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (G) Stay probability calculated from the entire trials. (H) Stay probability calculated from trials without premature responses and trial omissions. (I-J) Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (K) The frequency of premature responses measured after laser off or post-choice inactivation trials.

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (C), (D), (E), (K) Wilcoxon signed-rank test. (G), (H), (J) Chi-squared test. . (n=6 mice, 85 sessions)


A
NAc
GPe
VP
B
C
Premature response
Premature responses / Total trials
Laser OFF
Pre-start
***
OFF
Pre-start
Laser
D
Trial omission
Trial omission / Total trials
Laser OFF
Pre-start
***
OFF
Pre-start
Laser
E
Start reaction time
Start reaction time (sec)
Laser OFF
Pre-start
***
*
Reward
NoReward
Previous outcome
F
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start Inactivation
Pre-choice Inactivation
G
Stay probability
Stay probability
Laser OFF
Pre start
Pre-choice
***
***
Reward
NoReward
Previous outcome
H
Stay probability without Premature / Omission
Stay probability
Laser OFF
Pre start
Pre-choice
***
Reward
NoReward
Previous outcome
I
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Post-choice Inactivation
J
Stay probability
Stay probability
Laser OFF
Post-choice
*
Reward
NoReward
Previous outcome
K
Premature response
Premature responses / Total trials
Laser OFF
Post-choice
OFF
Post-choice
Laser

3.3 Inactivation of specific inputs to the NAc

Projections from multiple brain areas are also the determinant of function heterogeneity of the NAc. Therefore, we tested the effect of inactivation of distinct inputs to the NAc. Dopaminergic input and glutamatergic inputs were selectively inactivated in the NAc. The dopaminergic projection originates from midbrain dopaminergic neurons, and glutamatergic inputs are mainly from the neocortex and the hippocampus.

3.3.1 Inactivation of dopaminergic input to the NAc

In order to target dopaminergic projection, the Slc6a3-Cre (SG62) mouse line was used. Cre dependent AAV expressing ArchT was injected into the ventral tegmental area (VTA), and optical fibers were bilaterally implanted into the NAc (Figure 3.6 A, B). The localization of ArchT in axon terminals of dopaminergic neurons enabled inactivation of the dopaminergic input to the NAc. Pre-start inactivation increased the frequency of premature responses (Figure 3.6 C) while it did not affect start reaction time and the frequency of trial omissions (Figure 3.6 D-E). Pre-start and pre-choice inactivation did not cause significant changes in choice (Figure 3.6 F-H).

On the other hand, post-choice inactivation increased the stay probability after unrewarded trials (Figure 3.6 I-J). In contrast to pre-start inactivation, post-choice inactivation did not affect the premature response in the following trial (Figure 3.6 K).

3.3.2 Inactivation of glutamatergic inputs to the NAc

Glutamatergic inputs to the NAc were targeted using the Emx1-Cre mice line in which Cre recombinase is expressed in the majority of cortical pyramidal neurons and hippocampal neurons. The Emx1-Cre line was crossed with an Ai40 line so that all cre positive neurons express ArchT. Optical fibers were implanted bilaterally into the NAc to inactivate glutamatergic terminals in the NAc (Figure 3.7 A, B). Pre-start inactivation increased the frequency of premature responses as is observed in the inactivation of the dopaminergic input (Figure 3.7 C), and it did not affect start reaction time (Figure 3.7 E). Both pre-start and pre-choice inactivation did not affect choice (Figure 3.7 F-H).

For other behavioral measures, inactivation of glutamatergic inputs to the NAc showed distinct behavioral changes compared to inactivation of dopaminergic inputs. Pre-start inactivation decreased the frequency of trial omissions (Figure 3.7 D). Contrary to other inactivation experiments described so far, post-choice inactivation of glutamatergic inputs to the NAc did not increase stay probability following both rewarded and unrewarded trials (Figure 3.7 I-J).

In summary, selective inactivation of inputs to the NAc with different neurotransmitter caused differential behavioral changes in behavioral measures of action initiation and learning. Post-choice inactivation increased stay probability after unrewarded trials only in dopaminergic projection (Figure 3.6 J, 3.7J). Premature response was affected by pre-start inactivation of both dopaminergic and glutamatergic inputs (Figure 3.6 C, 3.7 C) while only inactivation of glutamatergic inputs to the NAc affected trial omission (Figure 3.6 D, 3.7 D).

Figure 3.6 Inactivation of dopaminergic inputs to the NAc.

(A) The target of inactivation. (B) A coronal section of a representative mouse brain. Left: Circles illustrate the position of the NAc. Right: Injection sites in the VTA. Circles illustrate the position of the VTA. (C) Effect of pre-start inactivation on premature response. (D) Effect of pre-start inactivation on trial omission. (E) Effect of pre-start inactivation on start reaction time. (F-H) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (G) Stay probability calculated from the entire trials. (H) Stay probability calculated from trials without premature responses and trial omissions. (I-J) Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (K) The frequency of premature responses measured after laser off or post-choice inactivation trials. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (C), (D), (E), (K) Wilcoxon signed-rank test. (G), (H), (J) Chi-squared test. (n=6 mice, 63 sessions)

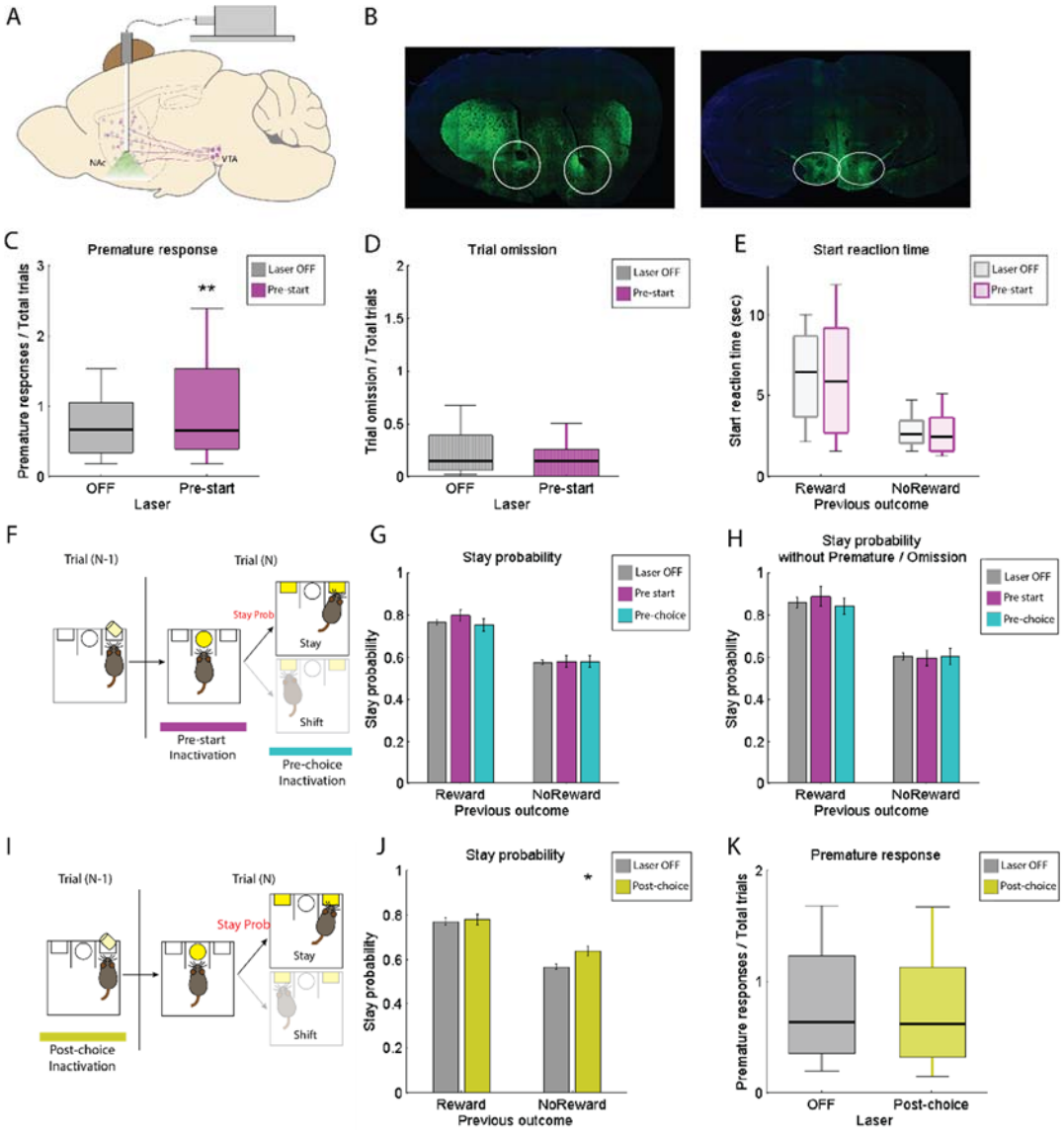
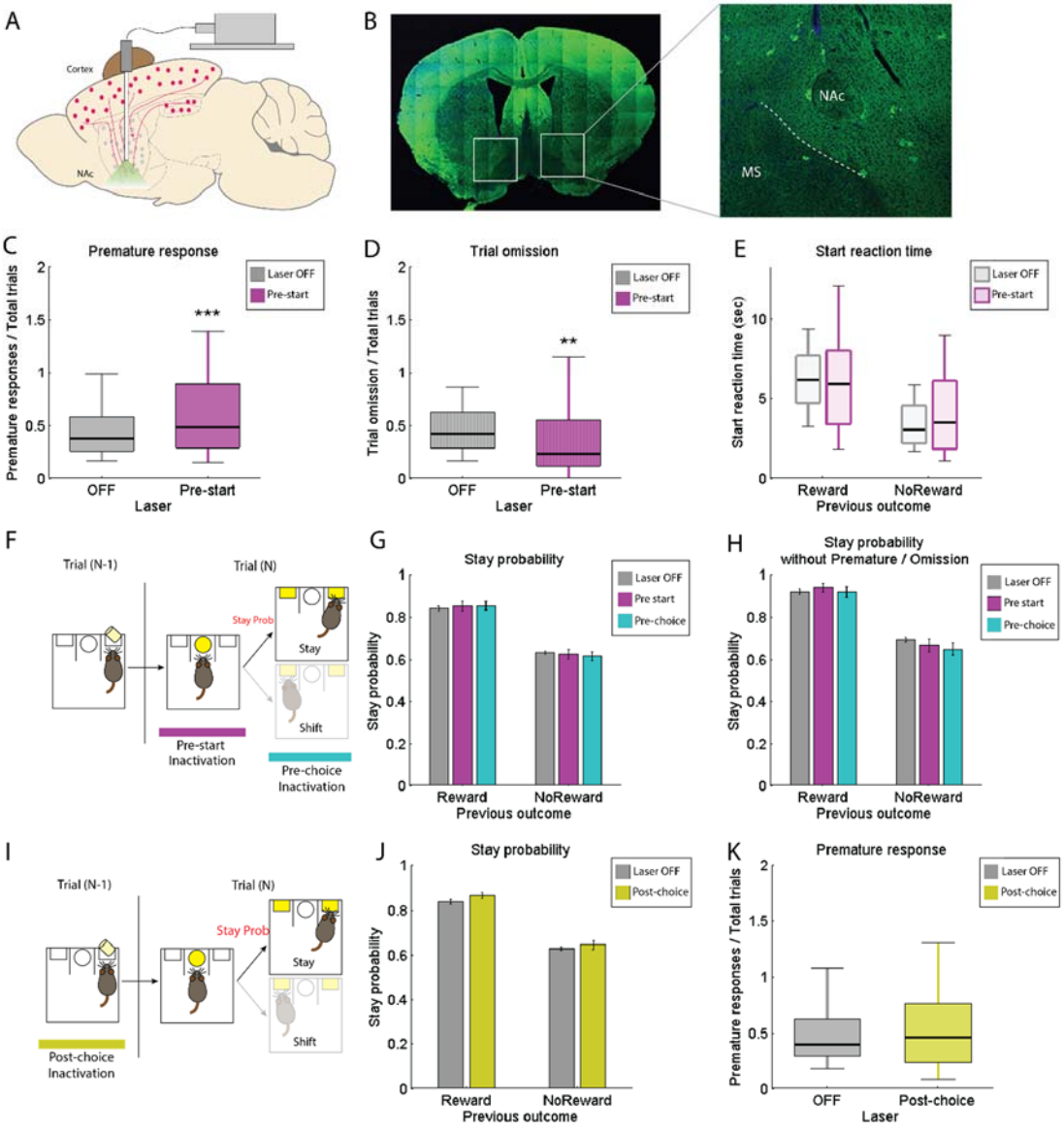

A
NAc
VTA
B
C
Premature response
Laser OFF
Pre-start
**
Premature responses / Total trials
OFF
Pre-start
Laser
D
Trial omission
Laser OFF
Pre-start
Trial omission / Total trials
OFF
Pre-start
Laser
E
Start reaction time
Laser OFF
Pre-start
Start reaction time (sec)
Reward
NoReward
Previous outcome
F
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start Inactivation
Pre-choice Inactivation
G
Stay probability
Laser OFF
Pre start
Pre-choice
Stay probability
Reward
NoReward
Previous outcome
H
Stay probability without Premature / Omission
Laser OFF
Pre start
Pre-choice
Stay probability
Reward
NoReward
Previous outcome
I
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Post-choice Inactivation
J
Stay probability
*
Laser OFF
Post-choice
Stay probability
Reward
NoReward
Previous outcome
K
Premature response
Laser OFF
Post-choice
Premature responses / Total trials
OFF
Post-choice
Laser

Figure 3.7 Inactivation of glutamatergic inputs to the NAc.

(A) The target of inactivation. (B) Left: A coronal section of a representative mouse brain. Squares illustrate the position of the NAc. Right: A magnified picture shows increased axonal density in the NAc compared to the medial septal nucleus (MS) that is located in the more ventral and medial location. (C) Effect of pre-start inactivation on premature response. (D) Effect of pre-start inactivation on trial omission. (E) Effect of pre-start inactivation on start reaction time. (F-H) Effect of pre-start and pre-choice inactivation on stay probability of the current trial. (G) Stay probability calculated from the entire trials. (H) Stay probability calculated from trials without premature responses and trial omissions. (I-J) Effect of post-choice inactivation at the previous trial on stay probability of the current trial. (K) The frequency of premature responses measured after laser off or post-choice inactivation trials. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (C), (D), (E), (K) Wilcoxon signed-rank test. (G), (H), (J) Chi-squared test. (n=7 mice, 104 sessions)


A
Cortex
NAc
B
NAc
MS
C
Premature response
Premature responses / Total trials
Laser OFF
Pre-start
***
OFF
Pre-start
Laser
D
Trial omission
Trial omission / Total trials
**
E
Start reaction time
Start reaction time (sec)
Reward
NoReward
Previous outcome
F
Trial (N-1)
Trial (N)
Stay Prob
Stay
Shift
Pre-start Inactivation
Pre-choice Inactivation
G
Stay probability
Stay probability
Pre start
Pre-choice
H
Stay probability without Premature / Omission
I
Post-choice Inactivation
J
Post-choice
K

3.4 Reaction time dependence of stay probability change

Post-choice inactivation increased stay probability specifically after unrewarded trials. This outcome-specific effect was not easily explained by general deficits in learning. In order to investigate the mechanisms that generated the outcome-specific effect, data from post-choice inactivation trials were analyzed more in detail.

In reinforcement learning models, values of behavioral events such as sensory stimuli, actions and outcomes were discounted over the passage of time (Montague et al. 1996; Sutton and Burto 1998). The temporal discounting of values were introduced into the models so that recent events or events that would occur in the near future affected more on the future decisions..

Since start reaction time after unrewarded trials was shorter than that after rewarded trials (Figure 3.2 E, 3.4 E, 3.5 E, 3.6 E, 3.7 E), we speculated that the difference in temporal discounting effects between these trial types were the cause of the outcome-specific effect of post-choice inactivation. To test this idea, trials with short start reaction time and long start reaction time were compared for laser off trials and post-choice inactivation traials. Therefore, in order to see the effect of temporal proximity of the timing of post-choice inactivation and the timing of choice at the following trial, we compared the effect of post-choice inactivation between trials with short and long start reaction time.

The median start reaction time of laser off trials was used as a threshold and trials with reaction time shorter than the threshold was classified as short reaction time trials. The rest of trials were classified as long start reaction time trials. When stay probability after unrewarded trials was calculated separately for short and long reaction time trials, only stay probability calculated from short reaction time trials showed significantly increased

by post-choice inactivation (Figure 3.8 A-D). Since some conditions showed as increased trend of stay probability even in long reaction time trials, the effect of inactivation on the stay probability was assessed using permutation test. Inactivation effect difference was quantified using the following formula.

Inactivation effect difference = $(StayP_{(ON, Short)} - StayP_{(OFF, Short)}) - (StayP_{(ON, Long)} - StayP_{(OFF, Long)})$

StayP : Stay probability after unrewarded trials

ON/OFF : Laser ON/OFF of the previous trial

Short/Long RT : Short or long reaction time of the current trial

Probability distribution of the inactivation effect difference was generated using permuted behavioral data in which all choice and outcome conditions were kept the same, and only labeling of inactivation conditions were randomly assigned. In non-specific inactivation and inactivation of indirect pathway MSNs, the inactivation effect was significantly stronger in trials with short start reaction time. (Figure 3.8 E-H). The other two conditions also showed trends of increasing inactivation effect. These results suggest that post-choice inactivation tends to be more effective when the timing of inactivation and next trial is proximal. In contrast, such difference was not observed in stay probability after rewarded trials (Figure 3.8 I-L).

One interpretation of this result is that inactivation of the NAc transiently change the behavioral state of mice, and the effect decayed over the passage of time. In order to test this possibility, trials in which mice receive pre-start inactivation after unrewarded trials were compared. Although the timing of the pre-start inactivation is more proximate to the

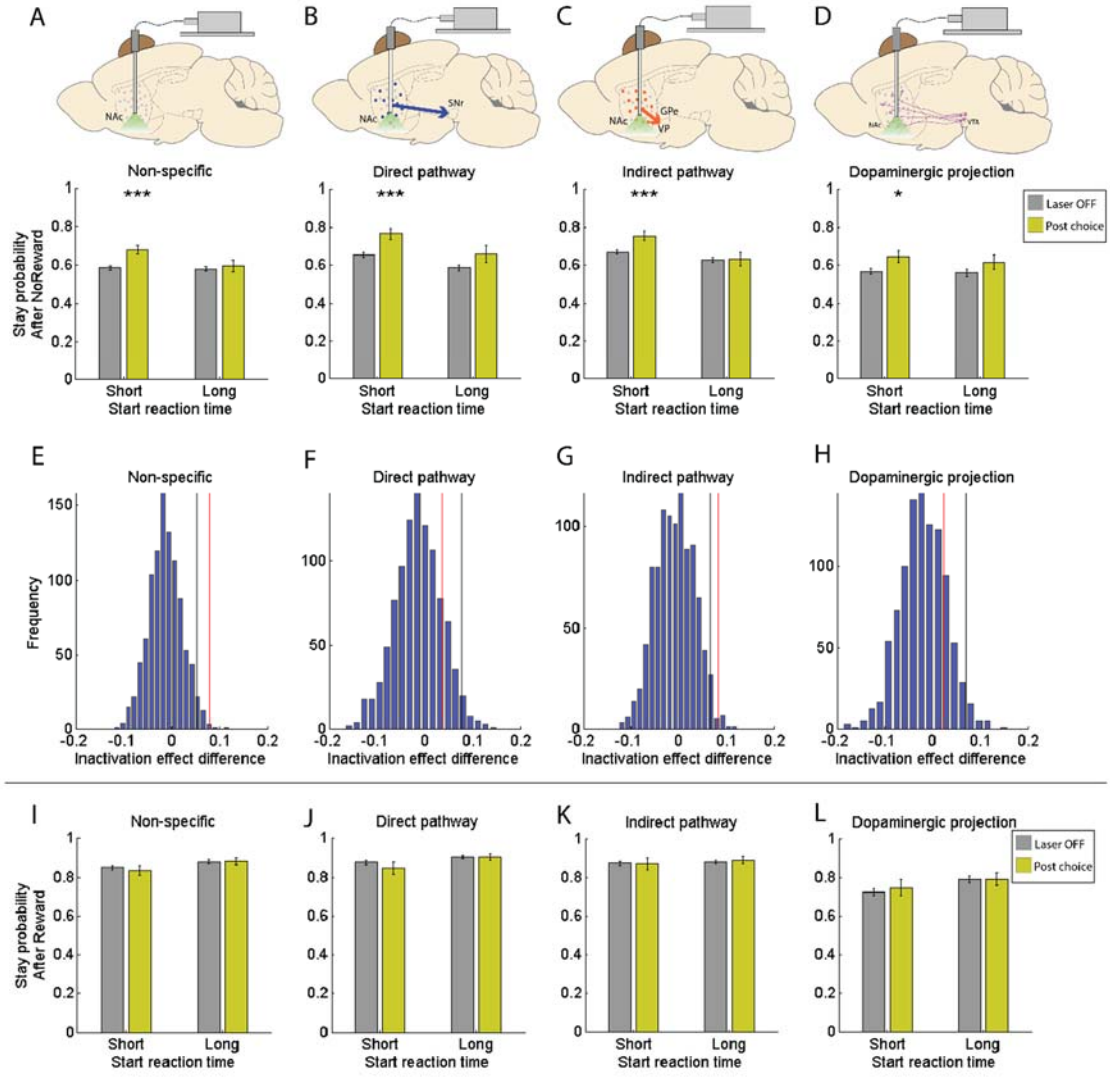
timing of choice compared to the timing of post-choice inactivation in the previous trials, pre-start inactivation did not increase stay probability after unrewarded trials (Figure 3.2E, 3.4E, 3.5E, 3.6E). Therefore, both the temporal proximity of inactivation to the next trial and the coincidence of inactivation and outcome delivery (the absence of food pellet) were necessary for the effect of post-choice inactivation.

Figure 3.8 Start reaction time dependence of post-choice inactivation effect on stay probability after unrewarded trials

(A-D) Stay probability after unrewarded trials was calculated separately from trials with short or long start reaction time. (E-H): Difference of inactivation effect (StayP(Laser ON)-StayP(Laser OFF)) between short and long start reaction time trials. Histograms were obtained by permutations (1000 times). Black vertical lines showed the threshold for p=0.05. Red vertical lines showed inactivation effect difference obtained from experimental data that was calculated from bars in (A-D). (I-J) Stay probability after rewarded trials were calculated separately from trials with short or long start reaction time.

(E) Non-specific inactivation: p=0.004 (F) Direct pathway inactivation: p=0.229 (G) Indirect pathway inactivation: p=0.018 (H) Dopaminergic input inactivation: p=0.303, permutation test

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (A-D), (I-L) Chi-squared test. (A, E, I) n=9 mice, 107 sessions. (B, F, J) n=6 mice, 74 sessions. . (C, G, K) n=6 mice, 85 sessions. (D, H, L) n=6 mice, 63 sessions.


A
B
C
D
NAc
SNr
GPe
VP
VTA
Non-specific
Direct pathway
Indirect pathway
Dopaminergic projection
***
***
***
*
Laser OFF
Post choice
Stay probability
After NoReward
Short
Long
Start reaction time
E
F
G
H
Frequency
Inactivation effect difference
I
J
K
L
Stay probability
After Reward

3.5 Summary

Non-specific inactivation of the NAc showed its involvement in action initiation and learning. From cell type or input specific inactivation experiments, each circuit element contributes to either one or both processes. One notable finding was the preferential contribution of indirect pathway MSNs to action initiation (Figure 3.5). Such specific effect was not predicted from the functional dichotomy of direct and indirect pathway MSNs. The learning effect shown by the increase of stay probability in the following trials were observed by inactivation of different circuit elements in the NAc (Figure 3.4, 3.5, 3.6). However, this effect was not observed in inactivation of glutamatergic inputs to the NAc (Figure 3.7). In contrast, inactivation of glutamatergic inputs to the NA only showed changes in action initiation (Figure 3.7). Another interesting observation was the outcome and reaction time dependent effects of post-choice inactivation on learning (Figure 3.8). Potential cause and underlying mechanisms will be discussed in Chapter 5.

Chapter 4: The Prefrontal Cortex Is Required for Action Initiation and Action Selection

While inactivation of multiple circuit elements in the NAc revealed their roles in action initiation and learning, inactivation of the NAc during pre-choice period only had a subtle effect on choice. The prefrontal cortex (PFC) has been suggested to play important roles in a wide range of cognitive control, including action initiation and action selection (Miller and Cohen 2001; Rudebeck et al. 2008; Murakami et al. 2014). However, temporally precise inactivation to dissect underlying neural substrates and cognitive processes was not extensively studied in mice. To investigate the contribution of the PFC to action initiation and action selection, inactivation of two PFC subregions, the ACC and the OFC was performed.

4.1 Inactivation of cortical neurons using VGAT-ChR2-EYFP mice

In order to inactivate the activity of pyramidal neurons in the PFC, the VGAT-ChR2-EYFP mouse line was used (Figure 4.1). In VGAT-ChR2-EYFP mice, the majority of cortical interneurons express Channelrhodopsin-2 (ChR2). Activation of ChR2 expressing interneurons leads to suppression of pyramidal neuron activity in the same area. The efficiency of silencing is characterized in a previous study (Guo et al. 2014).

The ACC and the OFC were inactivated at pre-start, pre-choice or post-choice period as in NAc inactivation experiments (Figure 4.2). Although the functional distinction of different prefrontal subregions has been studied, anatomical boundaries between different prefrontal subregions are not clear. Therefore, to obtain insights into anatomical location, the effect of inactivating different cortical volumes were compared. Inactivation of the

different size of areas was achieved using different laser powers as the previous study characterized the extent of inactivation in distant areas from the stimulation center (Guo et al. 2014). Three different laser powers (High:12-16mw, Medium: 4-5mW, Low:1.5-2.5mW) were used in order inactivate either large portion of the PFC or more localized areas in the PFC.

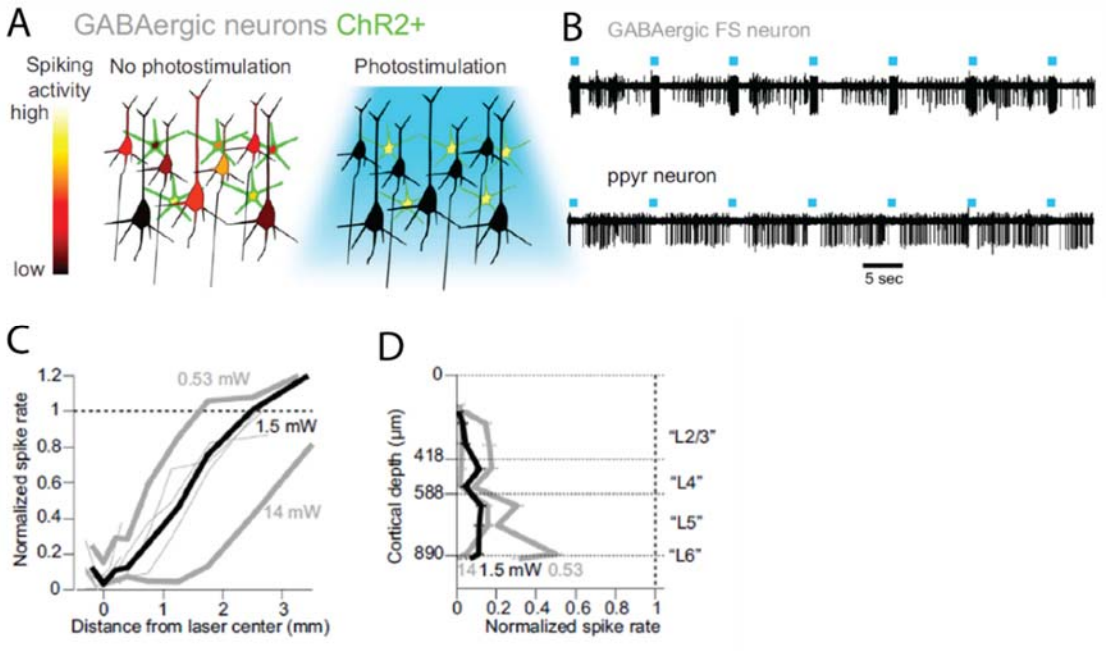


Figure 4.1 Inactivation of cortical pyramidal neurons using VGAT-ChR2-EYFP mouse line.

(A) Schematics of inactivation of pyramidal neurons by inactivation of cortical interneurons. (B) Results of slice recording in response to photostimulation. (C) Expected area that was inactivated at different laser power

(Adapted from Guo et al., 2014)

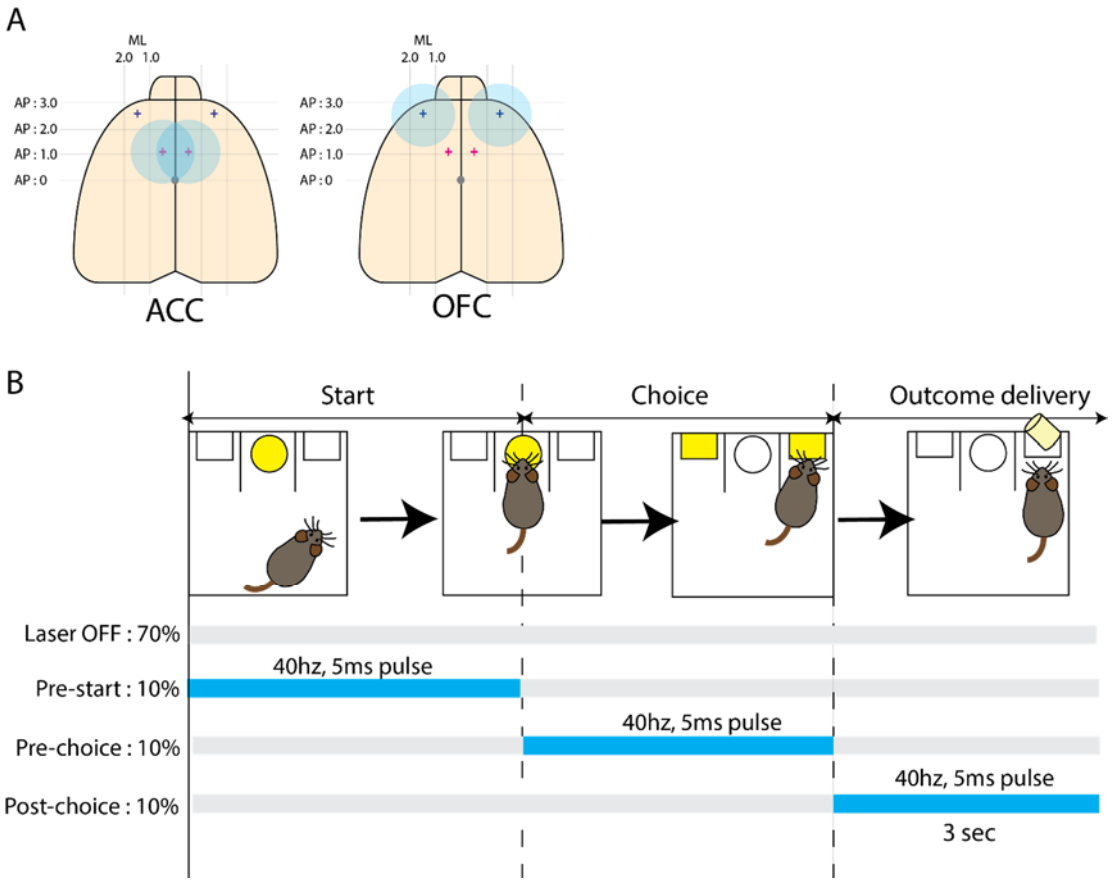


Figure 4.2 Location of inactivation and time windows of inactivation.

(A) Horizontal view of mouse brain with targeted location of the fiber implant (B) Three inactivation conditions (pre-start, pre-choice, and post-choice inactivation) was given at interleaved trials

4.2 Inactivation of the ACC and the OFC at high laser power

At highest laser power, inactivation aimed at the ACC and the OFC supposed to inactivate large overlapping areas in the PFC because of the size of the area receiving photoinhibition (Figure 4.1C-D). In order to see the effect of inactivation of the large portion of the PFC, inactivation at high laser power was performed. Consistent with the expected range of inactivation, inactivation of both the ACC and the OFC caused similar behavioral changes.

Behavioral measures of action initiation were affected in both ACC inactivation and OFC inactivation. Pre-start inactivation increased the frequency of both premature responses and trial omissions (Figure 4.3 A, B, D, E). ACC inactivation also increased start reaction time specifically after unrewarded trials (Figure 4.3 C). Pre-choice inactivation of either the ACC or the OFC caused increased choice reaction time (Figure 4.4). Although OFC control animals also showed increased reaction time probably due to strong laser power and proximity of implant sites and location of eyes, the OFC or the ACC inactivation group showed much larger increase in reaction time.

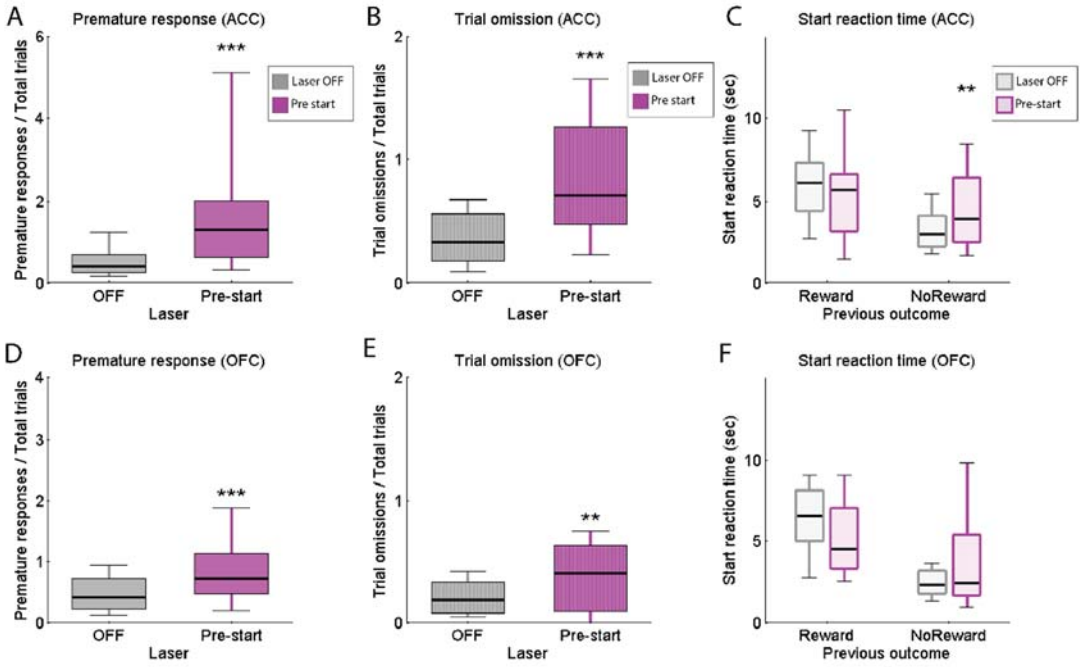


Figure 4.3 Impaired action initiation by inactivation of the ACC or the OFC at high laser power during pre-start period.

(A) Effect of ACC inactivation on premature response. (B) Effect of ACC inactivation on trial omission. (C)Effect of ACC inactivation on start reaction time. (D) Effect of OFC inactivation on premature response. (E) Effect of OFC inactivation on trial omission. (F) Effect of OFC inactivation on start reaction time. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (A)-(F) Wilcoxon signed-rank test (ACC: n=5 mice, 38 sessions, OFC: n=5 mice, 42 sessions).

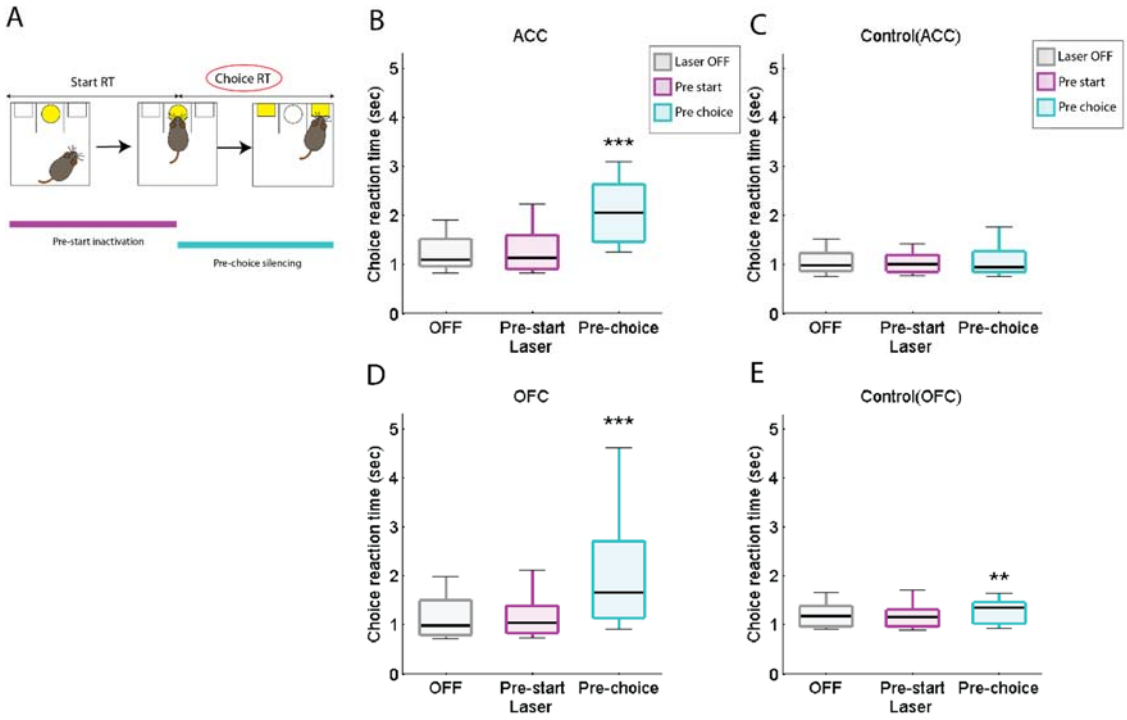


Figure 4.4 Increase choice reaction time by pre-choice inactivation of the ACC or the OFC at high laser power

(A) A schematic of choice reaction time. (B)-(D) Choice reaction time at laser off, pre-start inactivation and post-choice inactivation trials. (B) ACC (C) Control for ACC (D) OFC (E) Control for OFC. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (B)-(E) Wilcoxon signed-rank test (ACC: n=5 mice, 38 sessions, OFC: n=5 mice, 42 sessions, Control (ACC): n=5 mice, 28 sessions, Control (OFC): n=5 mice, 41 sessions).

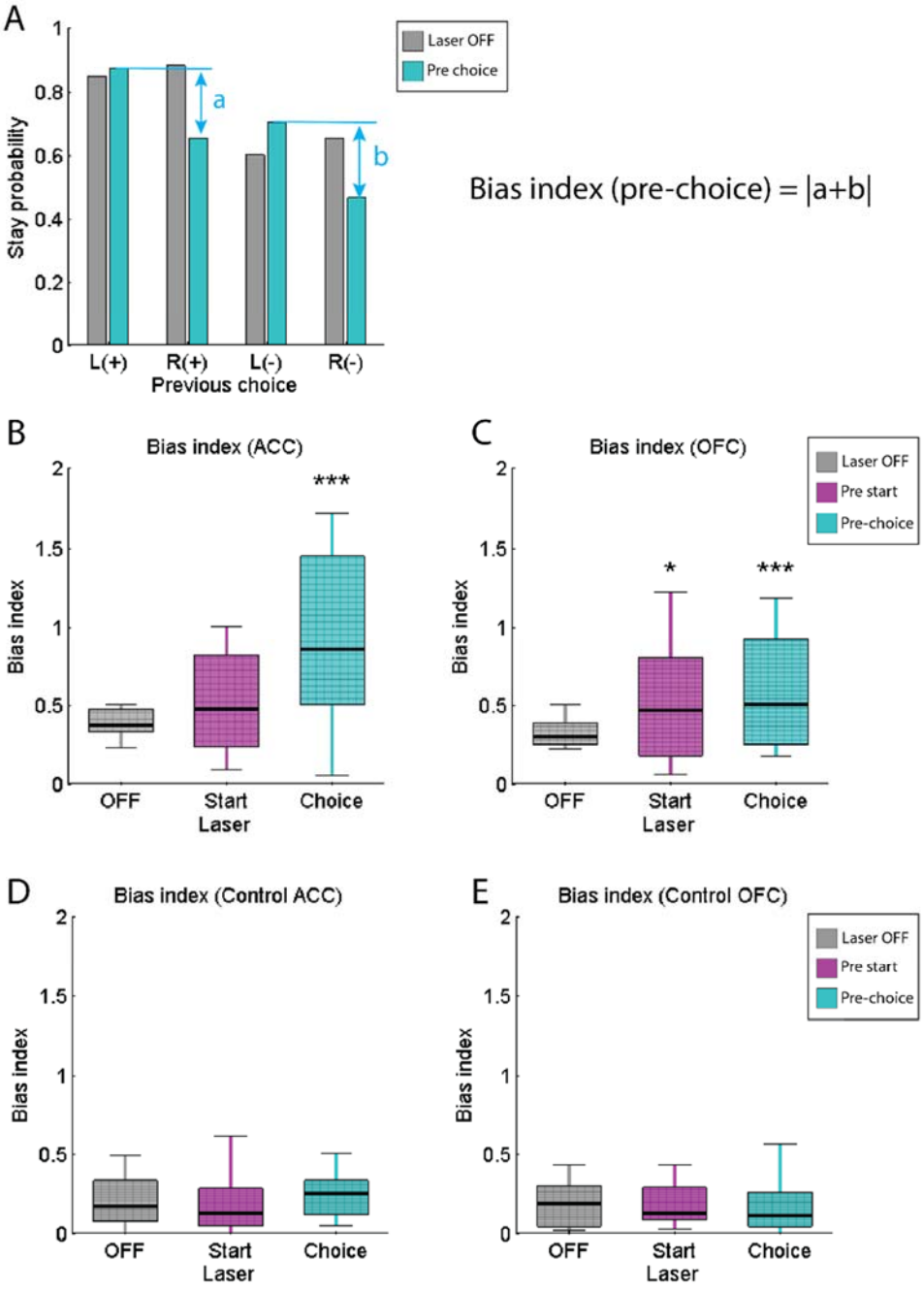
Contrary to NAc inactivation experiments, pre-choice inactivation of the ACC or the OFC affected the choice of the current trial. Pre-choice inactivation biased the current choice to either the left or right side regardless of previous choice or outcome (Figure4.5). The direction of bias varied across different animals. This bias was quantified by the difference in stay probability after left and right choice (Figure 4.5A). The bias index was calculated using following formula

Bias index = $| StayP_{L(+)} - StayP_{R(+)} + StayP_{L(-)} - StayP_{R(-)} |$

$StayP_{L(+)}$ denotes the stay probability after left choice rewarded trials, and $StayP_{L(-)}$ represents stay probability after left choice unrewarded trials. The same notation applies to the stay probability after the right choices. The absolute value was taken to calculate the bias index so that only biases occurred in the same direction between rewarded and rewarded trials gave higher values. The bias index was calculated for each inactivation condition. Both post-choice inactivation of the ACC and the OFC significantly increased bias index (Figure 4.5 B, C). Pre-start inactivation of the OFC also increased the bias index (Figure 4.5 C). Pre-start inactivation of the ACC did not cause significant increase but showed an increased trend for bias index (Figure 4.5 B).

Figure 4.5 Choice bias by pre-start or pre-choice inactivation.

(A) Stay probability from an example session. The bias index was calculated by the difference of stay probability after left and right choice. The difference was calculated separately for rewarded and unrewarded trials and combined. (B) Choice bias index from ACC inactivation experiments (C) Choice bias index from OFC inactivation experiments. (D-E) Choice bias index from control experiments. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (B)-(E) Wilcoxon signed-rank test (ACC: n=5 mice, 38 sessions, OFC: n=5 mice, 42 sessions)


A
Laser OFF
Pre choice
Stay probability
a
b
L(+)
R(+)
L(-)
R(-)
Previous choice
Bias index (pre-choice) = |a+b|
B
Bias index (ACC)
***
Bias index
OFF
Start
Choice
Laser
C
Bias index (OFC)
Laser OFF
Pre start
Pre-choice
*
***
D
Bias index (Control ACC)
E
Bias index (Control OFC)

4.3 Inactivation of the ACC at medium or low laser power

Inactivation experiments using high laser power confirmed the contribution of the PFC in the probabilistic reversal task especially for action initiation and action selection. We inactivated the ACC using medium or low laser power in order to see the effect in more localized inactivation.

Pre-start inactivation of the ACC increased the frequency of premature responses both at medium and low laser power (Figure 4.6 A, E) while trial omission was affected only at medium laser power (Figure 4.6 B, F). Pre-start inactivation did not affect reaction time except choice reaction time at low laser power (Figure 4.6 C, D, G, H). Pre-choice inactivation increased choice reaction time at both medium and low laser power (Figure 4.6 H). In addition, pre-choice inactivation at medium power increased choice bias as is observed in inactivation at high laser power (Figure 4.6 I).

4.4 Inactivation of the OFC at medium or low laser power

At both medium and low laser power, pre-start inactivation of the OFC affected behavioral measures for action initiation. The frequency of premature responses was increased by inactivation at both medium and low laser power while the frequency of trial omissions was decreased at medium laser power (Figure 4.7 A, B, E, F). In contrast to inactivation at high laser power, pre-start inactivation at medium or low laser power decreased start reaction time after rewarded trials (Figure 4.7 C, G).

Contrary to the effect of pre-start inactivation of the ACC on choice bias, both pre-start inactivation and pre-choice inactivation of the OFC did not significantly increase choice

bias at medium or low laser power (Figure 4.7 I, J) while there is an increasing tendency of bias index in pre-start inactivation at medium laser power.

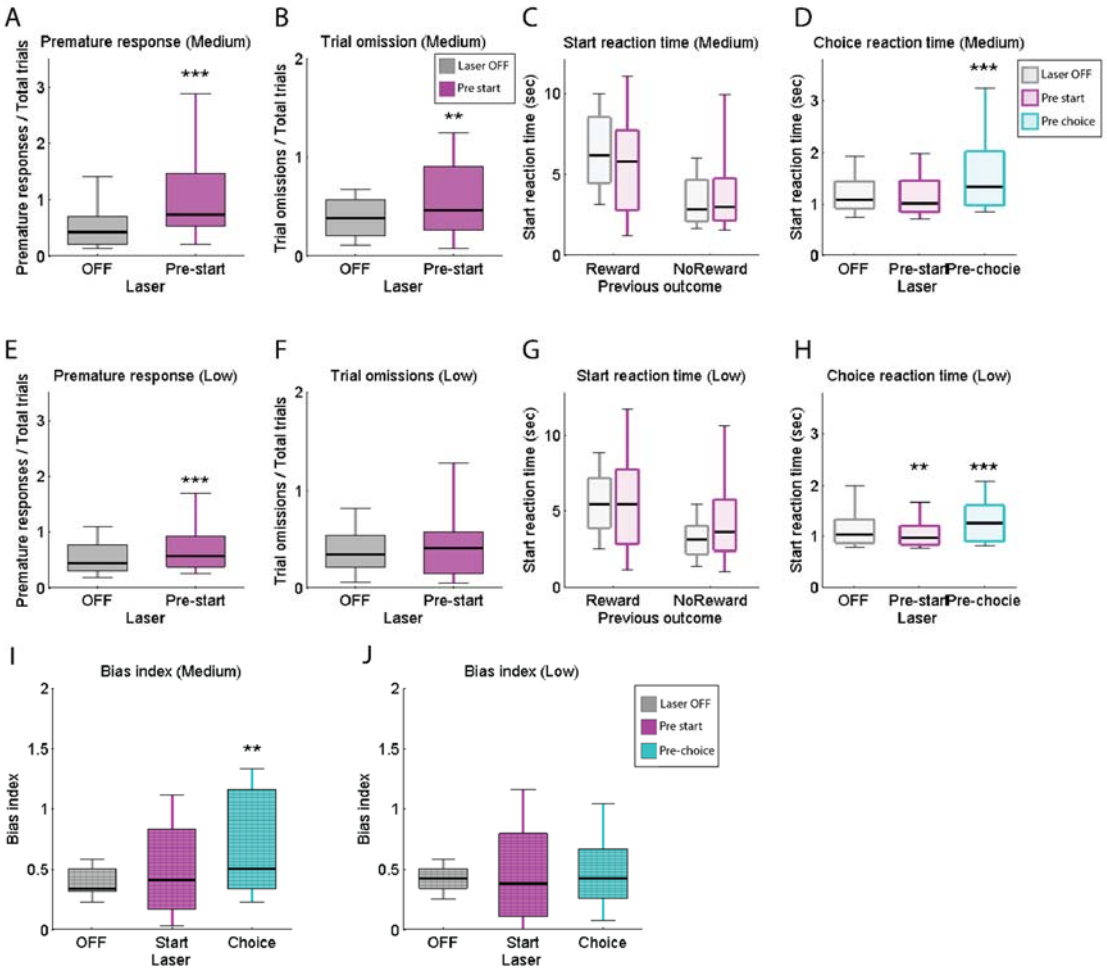


Figure 4.6 Inactivation of the ACC at medium or low laser power.

(A-D), (I) Inactivation of the ACC at medium laser power. (E-H), (J) Inactivation of the ACC at low laser power. (A), (E) Premature response. (B), (F) Trial omission. (C), (G) Start reaction time. (D), (H) Choice reaction time. (I), (J) Bias index. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (A)-(J) Wilcoxon signed-rank test (n=5 mice, 41 sessions (Medium power), 39 sessions (Low power))

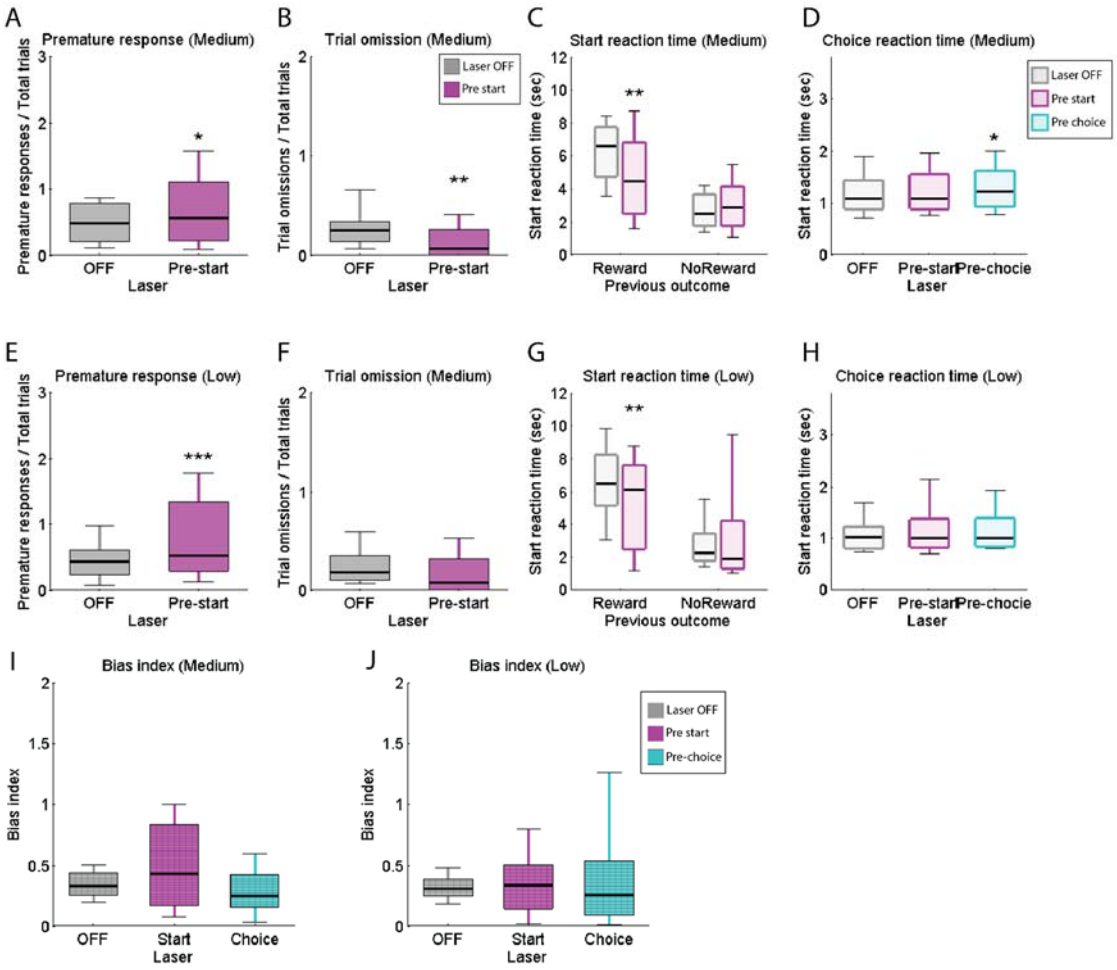


Figure 4.7 Inactivation of the OFC at medium or low laser power.

(A-D), (I) Inactivation of the OFC at medium laser power. (E-H), (J) Inactivation of the OFC at low laser power. (A), (E) Premature response. (B), (F) Trial omission. (C), (G) Start reaction time. (D), (H) Choice reaction time. (I), (J) Bias index. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (A)-(J) Wilcoxon signed-rank test (n=5 mice, 37 sessions (Medium power), 37 sessions (Low power))

4.5 Unilateral inactivation during pre-choice period

The strong choice bias observed in ACC inactivation can be caused by several different mechanisms. One possibility for the choice bias is a deficit in working memory. Inability to maintain information about a choice and an outcome of the previous trial may cause abnormal choice. However, this possibility is unlikely because pre-start inactivation showed no or weaker choice bias compared to pre-choice inactivation (Figure 4.5, 4.6,). Another possibility is that neurons in each hemisphere promote a choice to the contralateral side as was observed in the premotor cortex during the perceptual decision making tasks (Erlich et al., 2011, Guo et al., 2014). If the same argument applies to the probabilistic reversal task and the efficiency of inactivation in the left and right hemisphere is different, bilateral inactivation can also cause unilateral choice bias. To test the contribution of each hemisphere to the contralateral choice, unilateral inactivation experiments were performed using same animals.

Only pre-choice inactivation was performed, and each hemisphere was inactivated at medium laser power in 15% of trials. Unilateral pre-choice inactivation of the ACC increased choice bias when the left hemisphere is inactivated while inactivation of the right hemisphere had no effect (Figure 4.8 A). Effect of inactivation on the bias index was quantified by subtracting the bias index of laser off trials from the bias index of laser on trials. The comparison between the effect of unilateral left hemisphere inactivation (Figure 4.8 A, red bar vs. gray bar) and bilateral inactivation (Figure 4.8 B, blue bar vs. gray bar) did not lead to a significant difference (Figure 4.8). However, there were differential trends between these conditions in individual animals. The condition that gave stronger choice bias varies across individual animals (Figure 4.9). Mouse #1 and

Mouse #2 showed stronger choice bias by bilateral inactivation (Figure 4.9 A, B, F, G). In contrast, Mouse #3 showed stronger choice by unilateral inactivation (Figure 4.9 C, H).In addition, inactivation of either hemisphere did not increase choice reaction time as was observed in bilateral inactivation experiments (Figure 4.8 C, D). These results indicate that there is an effect of unilateral inactivation of the ACC, but that doesn't account for all behavioral changes in action selection observed by bilateral inactivation.

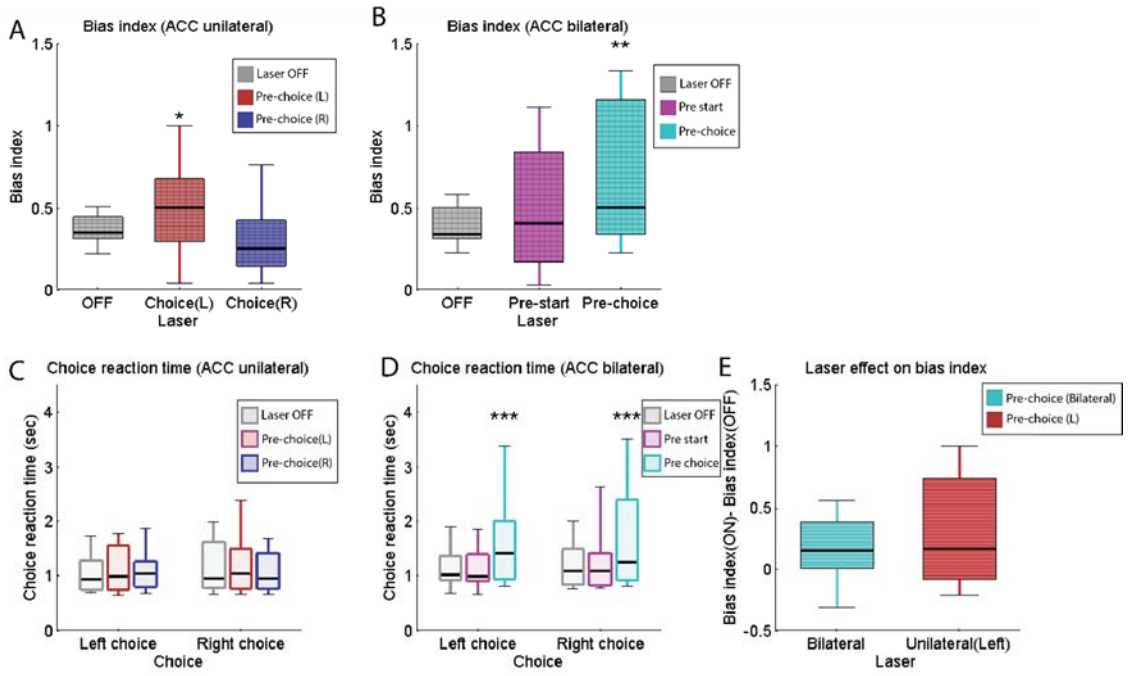


Figure 4.8 Effect of unilateral inactivation during pre-choice period.

(A) Bias index from unilateral pre-choice inactivation experiments at medium laser power. (B) Bias index from bilateral pre-start or pre-choice inactivation experiments at medium laser power (Same plot as Figure 4.6 I). (C) Choice reaction time from unilateral pre-choice inactivation experiments at medium laser power. (D) Choice reaction time from bilateral pre-start or pre-choice inactivation experiments at medium laser power. (E) Daily difference of bias indices between laser off and laser on conditions. Blue: Calculated from data in (B). The difference between 'Laser OFF' and 'Pre-choice' conditions. Red: Calculated from data in (A). The difference between 'Laser OFF' and 'Pre-choice (L)' conditions. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, (A)-(D) Wilcoxon signed-rank test, (E) Wilcoxon rank-sum test (n=5 mice, 39 sessions (Unilateral), 41 sessions (Bilateral))

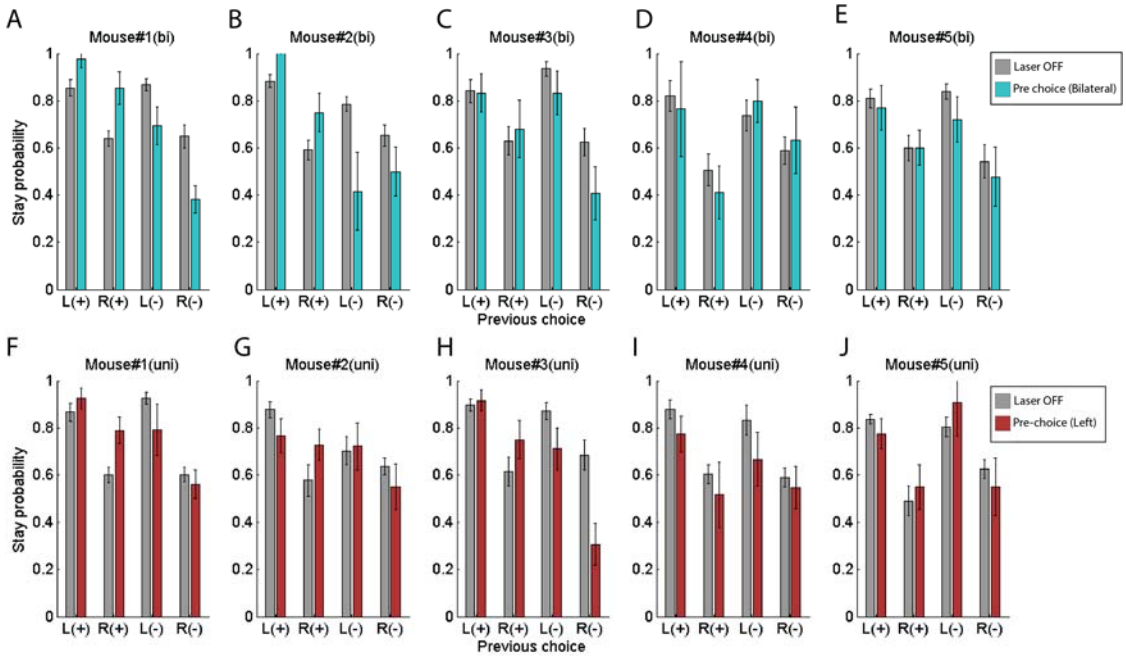


Figure4.9 Stay probability change of individual animals by unilateral or bilateral inactivation.

(A)-(E) Effect of bilateral inactivation on stay probability. (F) - (J) Effect of unilateral inactivation of the left hemisphere on stay probability.

4.6 Summary

Inactivation of the ACC or the OFC showed behavioral changes in action initiation and action selection. At high laser power, both ACC and OFC inactivation affected behavioral measures of action initiation and action selection (Figure 4.3, 4.4, 4.5).

Behavioral measures of action initiation were affected both by ACC inactivation and OFC inactivation. Both ACC and OFC inactivation increased the frequency of premature responses regardless of the laser powers tested while start reaction time after rewarded trial was reduced only by OFC inactivation (Figure 4.7). These results suggest the involvement of both the ACC and the OFC during action initiation period but in slightly different way.

At medium or low laser power, action selection was affected more strongly by ACC inactivation. ACC inactivation increased choice bias and choice reaction time increased (Figure 4.6), but OFC inactivation only showed a subtle increase in choice reaction time at medium laser power (Figure 4.7). These observations suggests the preferential involvement of the ACC in action selection process. Increased choice bias was partly explained by lateralized involvement of neurons in each hemisphere in choice (Figure4.8). However, some animals showed stronger choice bias by bilateral inactivation (Figure 4.9), and choice reaction time was not affected by unilateral inactivation (Figure 4.8). These results suggest that deficits in additional cognitive processes are necessary to explain the deficits in action selection.

Chapter 5: Discussion

5.1 Baseline behavioral performance

The behavioral performance of the probabilistic reversal task is constant over the long period of training sessions (Figure 2.5). The constant performance enabled us to collect data at different inactivation condition.

One notable difference between a previous study conducted in the similar condition using mice (Tai, Lee et al. 2012) is the baseline stay probability. In their study, stay probability after two consecutive rewarded or unrewarded trials were almost 1 or 0 respectively. This condition corresponds to the left most bar and the rightmost bar of the middle panel of Figure 2.5B. Therefore, the choice in their study depends more on recent outcomes and the choice in our study is more affected by long term integration of past outcome. One possible cause for the difference is the use of different reinforcer. We use food pellets while they use a water reward. Animals in their study peformed more trials in a single session. Since the number of trials depends on the capacity of animals to consume a reinforcer, the value of a single pellet used in the current study may be higher than that of the amount of water used in their study. In addition, the difference between the stay probability after unrewarded trials were more prominent between ours and their studies. This difference raises the possibility that the behavior in the face of aversive outcomes can be affected by overall reward rate. The potential contribution of the average reward rate will be discussed further in chapter 5.2.3.

5.2 Interpretation of behavioral changes observed during each time period

5.2.1 Trial initiation period

Pre-start inactivation led to the change in behavioral measures of action initiation (Premature response, trial omission and start reaction time) both in the PFC and the NAc. Several interpretations are possible regarding these behavioral measures.

From previous studies, the underlying cognitive processes responsible for the change in start reaction time are (1) Motivation (Wang et al. 2013), and (2) Impulsivity (Chudasama et al. 2003). Both increased motivation to perform the task and increased impulsivity can explain the observed reduction of start reaction time after rewarded trials.

The premature responses have been used as the behavioral measure of impulsivity (Chudasama et al. 2003; Pattiji et al., 2007). Although low motivation to perform the task can lead to general immobility and decrease in the frequency of premature responses, the high motivational level is unlikely to lead to increased premature response. Rather, it will increase the number of successful trials.

The trial omissions aer supposed to be affected in the opposite way as premature responses. Lower motivation or impulsivity will decrease the frequency of trial omissions, and higher motivation or impulsivity will decrease the frequency of trial omissions. In addition to these possibilities, trial omission can also be induced by disruption of preparatory activity observed in several premotor areas (Murakami et al., 2014;).

5.2.2 Choice period

The behavioral changes observed by pre-choice inactivation were the increased choice reaction time and choice bias. Change in the choice reaction time may come from perturbation in the motor system and change in subjective value estimates of choice. Since choice RT is less affected by the previous outcome, it is more likely that choice reaction time change was induced by perturbation in motor control.

On the other hand, multiple interpretations are possible for the observed choice bias in the PFC inactivation experiments. The firs possibility is a deficit in working memory. Since the probabilistic reversal task required animals to learn through trial and error, animals need to maintain information about their previous choices and outcomes. If working memory is impaired, animals will behave randomly or follow their pre-existing preference to one of the two choices. The second possibility is the lateralized contribution of each hemisphere to choice. In motor systems, some neurons encode movements or action values to the contralateral side. This lateralized function is supported by the perturbation experiments in both perceptual and value-based decision making (Tai and Lee et al. 2012, Guo et al. 2014, Hanks et al. 2015). Therefore, if an area responsible for a decision is not uniformly inactivated between the left and the right hemisphere, unilateral choice bias can be observed by bilateral inactivation experiment. The third possibility is that the inactivation exposed innate or pre-existing choice bias. At the begging of training step 3 (choice between 90% vs. 0% reward probability), animals are not aware that reversals occasionally occur, and successive unrewarded trials indicate reward availability from the other side. Therefore, it is possible that animals keep choosing one food magazine even after a prolonged period of the absence of reward if

they had experienced reward only from the food magazine. Since mice eventually experience reward from both food magazines, and such bias is supposed to be corrected after training sessions, it is still possible that such bias appeared again by inactivation of some brain areas.

5.2.3 Post-choice period

Post-choice inactivation experiments in the NAc affected the probability for mice to choose same choice following the trials with post-choice inactivation. This effect was specific to the unrewarded trials and not observed after rewarded trials. The change in the stay probability can be induced by multiple cognitive processes. The first possibility is that inactivation affected evaluation of outcome values to update future behaviors. If animals become unable to perceive outcome value correctly, subsequent choices are affected because of inaccurate feedback. The second possibility is that inactivation affected evaluation of a value of the current behavioral context. When animals learn from an outcome, the outcome value needs to be compared with the predicted or expected value. This predicted value is calculated through a variety of information regarding behavioral context, including outcome history at longer time scale, motivation to perform the task, and costs animals have to spend to make a decision. Altered calculation of the predicted value will affect subsequent choices even if the animals have intact ability to evaluate the outcome of the previous trials. Change in behavioral states including compulsivity or persistence can be regarded as deficits in this process. Changes in either or both mechanisms can affect the choice of the next trial.

5.3 NAc contribution to action initiation and learning

Behavioral changes observed by NAc inactivation experiments are summarized in Table 5.1. Underlying cognitive processes and neural substrates of these observations will be discussed in the following sections.

Table 5.1 Summary of behavioral phenotypes in NAc inactivation experiments.

	Pre-start inactivation				Post-choice inactivation	
Inactivation	Premature	Omission	Start RT (+)	Start RT (-)	StayP (+)	StayP (-)
Non-specific	↑	–	–	–	–	↑
Direct pathway	–	–	–	–	–	↑
Indirect pathway	↑	↑	↓	↑	–	↑
Dopaminergic input	↑	–	–	–	–	↑
Glutamatergic input	↑	↓	–	–	–	–

Premature: premature response, Omission: trial omission, Start RT (+): start reaction time after rewarded trials, Start RT (-): start reaction time after unrewarded trials, StayP (+): stay probability after unrewarded trials, StayP (-): stay probability after unrewarded trials

5.3.1 The NAc for action initiation

Among NAc inactivation experiments, non-specific NAc inactivation, indirect pathway MSNs inactivation, and inactivation of glutamatergic and dopaminergic inputs to the NAc increased the frequency of premature responses (Figure 3.2, 3.5, 3.6, 3.7). Inactivation of indirect pathway MSNs caused the largest effect on the frequency of premature responses and trial omission. One explanation for the increased frequency of premature responses is the increased level of general locomotion. However, this is

unlikely in the current behavioral context. Increased locomotion can't explain the increased frequency of trial omissions that were observed in inactivation of indirect pathway MSNs. Changes in premature response induced by pre-start inactivation reflect a change in impulsivity rather than a motivational change.

The strongest effect on premature responses was obtained by inactivation of indirect pathway MSNs. In addition, inactivation of indirect pathway MSNs also caused the increased frequency of trial omissions. The increased frequency of trial omissions can not be explained by the decreased level of locomotion since the frequency of premature responses was increased, and start reaction time after rewarded trial was reduced. Therefore, neither decreased motivation nor impulsivity occurred in this situation. So, disruption of the preparatory activity of goal-directed behavior or neuronal activity to trial initiation cues could be the candidate mechanisms. Previous studies showed that some neuron in the NAc increased firing rate at the initiation of new trials (Ito & Doya 2009; Atallah et al. 2014; Ito & Doya 2015). Although causal roles of such activities and underlying cell types have not been identified yet, the increased frequency of trial omissions can be the result of disruption of such activity.

Increased premature responses by inactivation of glutamatergic and dopaminergic inputs to the NAc gave new insights about how voluntary actions are initiated. Both systemic injection and intra-NAc infusion of amphetamine is known to increase general locomotion and impulsivity, and this effect is at least partially dependent on the NAc (Parkinson et al. 1999; Murphy et al. 2008). Amphetamine increases the extracellular level of dopamine. Then, the changed dopamine level is supposed to modulate other inputs to the NAc. However, it has not been clear about which inputs to the NAc are

modulated by the increased level of dopamine. Results of inactivation of glutamatergic indicate that the cortical input is one candidate that is affected by pharmacological perturbation.

5.3.2 The NAc for learning

Among NAc inactivation experiments, non-specific NAc inactivation, direct pathway MSN inactivation, indirect pathway MSNs inactivation, and inactivation of the dopaminergic input to the NAc increased stay probability after unrewarded trials with post-choice inactivation (Figure 3.2, 3.4, 3.5, 3.6). This effect was observed across the inactivation of different circuit elements in the NAc except inactivation of glutamatergic inputs to the NAc. One previous study using rats (Dalton et al., 2014) also showed the change in choice following NAc inactivation. However, the results reported there were not consistent with the observed behavioral changes in the current study. They inactivate the NAc shell using muscimol and found that stay probability after rewarded trials was reduced. In addition to the difference in inactivation methods and the anatomical location targeted, the observed difference may come from species difference between mice and rats. Both our and their studies used food pellet as a reinforcer. We used 14mg food pellet, and they used 45mg food pellet. Considering the difference in body weights of mice (20-35g) and rats (280-350g), the value of one food pellet is much smaller in rat studies. This difference may cause the difference in average subjective value of reward obtained during unit time even if number of food pellets and duration of the behavioral session is similar. It is possible that the larger reward value of a single food pellet makes

the inactivation of the NAc alone insufficient to decrease stay probability after rewarded trials.

In reinforcement learning models, values of actions are updated based on the difference between outcome values and the current prediction about action values (Montague et al. 1996; Sutton and Burto 1998; Ito& Doya, 2009). The updated action values are used for future decisions. One possibility to explain the experimental data is that post-choice inactivation directly modified the outcome values without affecting predicted or expected values of actions. If this is the case, only the outcome value of a reward omission needs to be overestimated by inactivation without affecting the outcome value of a reward. However, it is unlikely that only unrewarded trials were affected by post-choice inactivation since delivery of a reward is the more salient event compared to a reward omission.

One potential cause of the increased stay probability after unrewarded trial is that inactivation of the NAc reduced the expected or predicted action values rather than the actual outcome values they received. One study (Ito and Doya 2015) showed that subset of rat NAc neurons represent a state value during ITI and trial initiation period. The state value is the value of the current behavioral context and is calculated from a long-term outcome history during a behavioral session. Expected or predicted action values is the current probabilistic reversal also requires long-term accumulation of outcome values from previous trials. Therefore, NAc inactivation can affect expected action values or state values.

A potential mechanism for NAc inactivation to affect expected action values is to reduce 'opportunity cost'. Opportunity cost is defined as the value of unit time, or the value that

will be missed because of doing nothing during that time period (Niv et al. 2007). That is experimentally defined as the average reward value obtained during a certain time period. In the current probabilistic reversal task, animals that make random choices can still receive a reward at 37.5% of trials. Therefore, it is possible that animals perceived the passage of time during the task as a certain kind of costs, and that opportunity cost was also considered in calculation of action value.

Both theoretical work (Niv et al. 2007) and experimental work (Wang et al. 2013) suggested that the change in the opportunity cost affect how hard animals work to obtain a reward. Niv et al. proposed a modified reinforcement learning model that incorporated opportunity cost into calculation of expected values of actions, stimuli and states. In their model, the opportunity cost was calculated as the product of the average reward rate for long time scale and the time to perform the next action. This opportunity cost was subtracted from the value calculated from recent reward history, and the resultant value was used as the expected value of action for decisions. They used this model to explain the observation that animals trained under high reward rate were more motivated to perform the task in addition to the tendency to choose the option with highest value. Since the opportunity cost in the model was used to calculate expected action values, transient change in the opportunity cost can not only affect response vigor but also affect learning through a reward prediction error signal. If inactivation of the NAc decreased the opportunity cost, expected action values become larger, and stay probability will increase even after unrewarded trials. In contrast, outcome value of a reward is always higher regardless of reduction of opportunity cost. In addition, animals had already showed high

stay probability after rewarded trials. Therefore, if the effect of increased prediction error saturated, post-choice inactivation will not increase stay probability any further.

In summary, the interpretation of the increased stay probability after unrewarded trials is that the post-choice inactivation lowered the opportunity cost and led to the more persistent choice in the following trials. Situation of lower opportunity cost makes rats, pigeons, and humans behave more persistently even under aversive outcomes, such as the absence of reward (Arkes and Blumer 1985; Wikenheiser et al. 2013; Magalhães and White, 2014). This possibility can be studied more in detail by systematically modulating opportunity cost during the task by changing the duration of inter-trial interval or reward size.

Incorrect calculation of opportunity cost can be the underlying mechanisms of persistent or compulsive behaviors in the face of aversive outcomes. The same argument is also applied to drug seeking behaviors that are hard to be suppressed (Pelloux et al. 2007). These behavioral changes can be the results of altered activity in the NAc. Rats treated with D2 and D3 receptor agonist quinpirole showed compulsive checking behavior (Dvorkin et al. 2010). This study is consistent with the increased stay probability observed by inactivation of indirect pathway MSNs. Therefore, neural circuits targeted in this study can be potential targets of treatment in maladaptive behaviors such as OCD and addiction to drug or gambling.

5.3.3 Reaction time dependence of learning effect by NAc inactivation

The change in stay probability after unrewarded trials were stronger when trials with shorter start reaction time were analyzed (Figure 3.8). However, pre-start inactivation after unrewarded trials did not increase stay probability (Figure 3.2 – 3.7). Therefore, both temporal proximity to the next choice and the absence of reward are important for the increase in stay probability. The lack of effect in pre-start inactivation also implies that increased probability was not because of a change in locomotion or a change in behavioral states at longer time scale.

Rather, the results are consistent with the idea of temporal discounting of value (Montague et al. 1996; Sutton and Burto 1998). When animals face reward whose delivery is delayed, they show reduced preference for the delayed reward (Cardinal et al. 2000; Winstanley et al. 2004). This concept also applies when animals calculate value of future actions based on previous outcomes. The reinforcement learning model that incorporate temporal discounting of previously obtained reward values explains actual animal performance better than models without temporal discounting (Ito & Doya 2009). The impact of a reward is discounted if the reward was given before. Therefore, if post-choice inactivation somehow affected calculation of values, the effect can also be temporally discounted. Therefore, the underlying mechanism will be post-choice inactivation modify values during the post-choice period, and the modified value is discounted by the passage of time.

The concept of opportunity cost is also consistent with the reaction time dependence of stay probability. Since opportunity cost is the product of average reward rate and duration of time which animals spend without performing a task, shorter start reaction time means

smaller opportunity cost. This lead to higher expected action value of previously chosen action and it increased stay probability.

5.3.4 Partially overlapped effect of direct pathway vs. indirect pathway MSNs inactivation in the NAc

Inactivation of direct pathway MSNs and indirect pathway MSNs caused a similar change in stay probability (Figure 3.5 J, 3.6 J). On the other hand, several studies in the striatum reported opposing behavioral effects by inactivation of direct and indirect pathway MSNs. Kravitz et al.(2010) activated these two population in the dorsal striatum using optogenetics reported the opposite effect on locomotion. They observed increased locomotion by the activation of direct pathway MSNs and decreased locomotion by the activation of indirect pathway MSNs. Tai et al.(2012) activated the dorsal striatum while mice were a decision making task similar to ours. They reported that unilateral activation of direct pathway MSNs increased a choice probability of the contralateral side and the opposite effects by activation of indirect pathway MSNs. Along with the difference in areas studied, there are several explanations for the lack of opposing effects. First, activation experiments, not inactivation experiments, were conducted in above studies. Interpretation of activation experiments is often confounded by stimulation parameters and leads to undesirable side effects through the activation or inactivation of areas to which neurons project. Activation experiments are useful to model the situation in which a given area or cell type receive too much excitation as the results of Kravitz et al. mimicked behaviors observed during the parkinsonian state. However, in order to study the function of a circuit during natural behaviors, it is necessary to confirm the

endogenous activity during the stimulation context and need to confirm the absence of undesirable side effects. For these reasons, behavioral impacts of activation and inactivation experiments do not always become opposite or complementary. Supporting this view, both direct pathway and indirect pathway MSNs are active during natural movements according to *in vivo* calcium imaging of the dorsal striatum in mice (Cui and Jun et al., 2012). If concurrent activation of both direct and indirect pathway MSNs are important for the successful execution of a behavior, inactivation of either population can lead to similar behavioral changes.

When dopaminergic and glutamatergic inputs were selectively inactivated, both manipulation caused increased premature response. The common behavioral effects is reasonable considering roles of these neurotransmitters. Dopamine is a neuromodulator that modulate inputs from other structure such as the neocortex to the NAc (Bamford et al. 2004). Therefore, the probable cause of increased premature response is the impaired information transmission from glutamatergic neurons either through direct perturbation of these neurons or abnormal modulation of glutamatergic inputs by abnormal dopaminergic transmission.

5.3.5 Dual role of the NAc in learning and action initiation

Two distinct models have been proposed to explain the function of the NAc in reward-based behavior. One model proposes the involvement of the NAc in learning. This view is supported by the fact that the NAc is the primary recipient of VTA dopaminergic neurons in reward prediction error signal is observed (Corbit et al. 2001; Atallah et al. 2006, McDannald et al. 2011). Another view is that the NAc is more involved in the

process of action initiation such that reward associated sensory cues drive behavior (Hall et al. 2001; Nicola 2010; Smith et al. 2011). This view describes the function of the NAc as mediating excitatory effects of reward associated sensory cues on reward-seeking behavior rather than learning.

We got behavioral changes by both pre-start inactivation and post-choice inactivation. This finding raises the possibility that these two views are not mutually exclusive and rather reflect different functioning of the NAc at different timing during a behavioral task. Such dissociation has never been possible without combining the perturbation with fine temporal resolution and trial based behavioral paradigm.

5.4 PFC contribution to action initiation and action selection

Behavioral changes observed by PFC inactivation experiments are summarized in Table 5.1. Underlying cognitive processes and neural substrates of these observations will be discussed in the following sections.

Table 5.2 Summary of behavioral phenotypes in PFC inactivation experiments.

	Pre-start inactivation				Pre-choice inactivation	
Inactivation	Premature	Omission	Start RT(+)	Start RT (-)	Choice RT	Bias index
ACC(High)	↑	↑	–	↑	↑	↑
ACC(Medium)	↑	↑	–	–	↑	↑
ACC(Low)	↑	–	–	↑	↑	–
ACC(Left)	na	na	na	na	–	↑
ACC(Right)	na	na	na	na	–	–
OFC(High)	↑	↑	–	–	↑*	↑
OFC(Medium)	↑	↓	↓	–	↑	–
OFC(Low)	↑	–	↓	–	–	–

Premature: premature response, Omission: trial omission, Start RT (+): start reaction time after rewarded trials, Start RT (-): start reaction time after unrewarded trials, Choice RT: choice reaction time. *: Same change was also observed in control animals although the effect was weaker.

5.4.1 Cortical areas covered by the inactivation at medium to high laser power

In the current study, the tip of the optical fiber was placed in the OFC or the ACC. Inactivation of these areas led to partially overlapped behavioral phenotypes, especially at medium or high laser power. The common effects were probably because of the overlap in the area that receive inactivation at higher laser power. Therefore, to understand the potential contribution of areas other than the ACC and the OFC, the cortical areas that are supposed to be inactivated along the OFC and the ACC need to be considered.

The location of the ACC targeted in this study is next to the M2 with similar anterior-posterior and dorso-ventral coordinates. The proximity to the M2 might be the reason unilateral inactivation caused choice bias. The ACC is also located close to the mPFC with similar anterior-posterior and medial-lateral coordinates. Therefore, the contribution of those areas also need to be considered to interpret the results of inactivation experiments using medium or high laser power.

The coordinates used for targeting the OFC more anterior, lateral and ventral to the ACC. The areas close to these coordinates include the insular cortex, the medial OFC, and the IL. The dorsal part of the OFC targeted in this study is the anterior lateral motor cortex (ALM), the putative premotor cortex (Komiyama et al. 2010; Guo et al. 2014). Therefore,

the potential contribution of these areas to behavioral phenotypes also need to be considered.

5.4.2 PFC for action initiation

At inactivation experiments using high laser power, both premature response and trial omission were increased by pre-start inactivation (Figure 4.3). In the prefrontal and premotor areas, neuronal activity that precedes specific motor responses has been identified (Romo and Schultz 1992; Narayanan et al. 2008; Mita et al. 2009; Erlich et al. 2011; Guo et al. 2014; Murakami et al., 2014). Neuronal activity preceding motor response can be important for both action initiation and action selection. Some of them identified neuronal activity whose peak is tightly coupled to the movement onset (Mita et al., 2009; Murakami et al., 2014). This suggests the contribution of such activity to action initiation. An increased frequency of premature responses or trial omissions can be the result of disruption of such activity.

Increased premature response and increased trial omission may look opposite effect in terms of locomotion. However, both can be originated from cognitive processes related to action initiation. Both initiating any movements and performing an appropriate action (center nose-poke not left or right magazine entries) are necessary for successful trial initiation. Deficits in the first process lead to trial omission and deficits in the second process lead to premature responses. Since animals were food restricted during the entire training and test sessions, they have a relatively strong drive to initiate movements to initiate trials, and this drive may be more difficult to be suppressed completely. But precise control to perform an appropriate action (center nose-poke) may need more

precise representation of activity and more vulnerable to perturbation. This may be why only premature response are increased in many conditions, and both premature response and trial omission were increased in some conditions that showed a large increase in premature responses.

Increase premature response can also be interpreted as deficits to behavioral inhibition or the impulsive behavioral state. This is consistent with the previous studies showing that excitotoxic lesion or pharmacological inactivation of several PFC subregions, including the OFC, led to increased premature responses (Chudasama et al. 2003; Naranayan et al. 2006), and that OFC lesioned rats showed altered preference to delayed reward (Winstanley et al. 2004, Rudebeck et al. 2006, Mar et al. 2011, Stopper et al. 2012). Increased premature responses by OFC inactivation at the current study is consistent with those previous studies. The impulsive behavioral state is also consistent with the combination of increased premature response and decreased trial omission as is observed in OFC inactivation at medium power and inactivation of glutamatergic inputs to the NAc. These manipulations probably induced relatively stronger deficits in conducting appropriate center nose-pokes with intact ability to initiate movements.

Although both ACC inactivation and OFC inactivation increased the frequency of premature responses (Figure 4.6 A, E, 4.7 A, E), start reaction time was mostly affected by OFC inactivation. Premature responses can be induced by partial disruption of the motor system. This may be the case for inactivation of ACC and inactivation of OFC at high laser power because of spatial proximity with the motor system. This may also explain why OFC inactivation at high laser power did not significantly change start reaction time while inactivation of the OFC at low and medium power induced reduction

in start reaction time (Figure 4.3 F, 4.7 C, G). Since OFC inactivation at low or medium power did not affect choice, the deficits in action initiation observed in OFC inactivation reflect more cognitive aspects rather than disruption in the motor system.

In addition, the reduction of start reaction time was observed only after unrewarded trials (Figure 4.7). The reduced start reaction time may reflect the altered evaluation of the previously obtained outcome as was observed in delay discounting tasks. This possibility can be easily tested using a modified version of the probabilistic reversal task with delayed delivery of food pellets. Interesting observation about the reduction of start reaction time is that an outcome-specific effect was induced by pre-start inactivation rather than post-choice inactivation in which animals received outcomes and consume them. It is possible that some of outcome or action-related activity observed at trial initiation period (Feierstein et al. 2006) is used to guide behavior. Since previous studies conducted inactivation with longer duration, it was not clear whether those behavioral deficits were because of the disruption of neuronal activity at the time of behavioral responses or a change in behavioral states induced by long duration inactivation of the PFC. Results of the current study support the first possibility, considering behavioral changes during specific time periods.

The increased premature response was also observed in inactivation of glutamatergic inputs into the NAc and the NAc itself. This common effect raises the possibility PFC drives the NAc in terms of action initiation. Compared to the dorsal striatum that receive inputs from most cortical regions, the NAc receive inputs from more restricted areas such as the OFC, the insular cortex, and the IL. Although further study is necessary to define

the exact cortical area involved in the process, these results indicate that the direct input from the PFC to the NAc control action initiation and impulsivity.

The increased frequency of premature responses is also observed in patients with ADHD, drug addiction, and prefrontal damage (Schachar et al. 1995; Fillmore and Rush 2001; Aron et al., 2003; Monterosso et al. 2005). In ADHD patients, decreased neuronal activity in the NAc was observed while they were anticipating rewards (Scheres et al. 2007). These studies support our results in inactivation of glutamatergic and dopaminergic inputs to the NAc. Thus, these projections to the NAc can be potential targets for the treatment of impulsive behaviors.

5.4.3 PFC for action selection

Pre-choice inactivation of the PFC increased choice reaction time and choice bias. There are several mechanisms that can affect choice as observed in the current experiments. The first possibility is the working memory deficits. The second possibility is the lateralized contribution of each hemisphere. The third possibility is the transition between goal-directed and habitual system. Each possibility is addressed in this section.

Since animals have to maintain the information about past choices and outcomes, the perturbation to working memory system is also potential mechanisms that cause choice bias observed in the ACC or OFC inactivation. Several studies reported persistent activity during the delay period of working memory in the PFC (Funahashi et al. 1989). Perturbation of activity in the PFC during this period disrupt the performance of working memory task (Rossi et al. 2012; Liu et al. 2014). However, working memory deficits

can't account for the choice bias observed in the current study. Although choice bias was observed in pre-choice inactivation of the ACC and the OFC, pre-start inactivation only showed increased trend of bias in some conditions. (Figure 4.5 B, C, 4.6 I, J, 4.7 I, J). If working memory deficits made mice unable to perform flexibly in the probabilistic reversal task, all of the three inactivation conditions should lead to similar changes in choice bias. The lack of contribution to working memory system is probably because of the difference in task requirements. Most working memory tasks in rodents and primates utilize sensory cues to dictate the direction of choice. On the other hand, in the probabilistic reversal task, visual cues (center LED for trial start and left or right LED for choice) only have a permissive role for action and did not contain information about correct choice direction. Since animals' choice did depend on past choice and outcome, the result indicates that the information is stored outside of the PFC.

In unilateral pre-choice inactivation of the ACC increased choice bias in one of two conditions (Figure 4.8 A). This result suggests that lateralized choice in the probability task is partially conducted through prefrontal, premotor or motor cortices. Because of the spatial proximity of the ACC and the secondary motor cortex, this result is consistent with the previous studies (Erlich et al. 2011; Guo et al. 2014) showing that inactivation of premotor or prefrontal areas in one hemisphere increased choice bias to the ipsilateral side. However, if this accounts for all the choice biases observed in the current experiments, the effect of unilateral inactivation should be stronger not weaker than that of bilateral inactivation since inactivation of each hemisphere supposed to counteract each other in bilateral inactivation. In addition, only bilateral inactivation of the ACC increased choice reaction time (Figure 4.8). In addition, data from individual animals

indicates that the relative strength of inactivation effects between unilateral and bilateral experiments varies across individual animals (Figure 4.9). These results suggest contribution by other systems to the process of action in addition to lateralized effect.

One potential mechanism that accounts for the increased choice bias is the shift of relative dominance of goal-directed and habitual behavioral control (Daw et al., 2005, Smith et al., 2012). The goal-directed control system is more outcome dependent and flexible but needs more cognitive loads. On the other hand, the habitual control system is less outcome dependent and flexible but needs less cognitive loads. The gradual shift from the goal-directed system to the habit system has been used to explain the change of behavior over training in multiple behavioral paradigms and model animals (Yin et al. 2004; Yin et al. 2005; Valentin et al. 2007; Balleine and O'Doherty 2011).

Excitotoxic lesion of the PL and chemogenic or optogenetic inactivation of the OFC make reward dependent instrumental actions more habitual, insensitive to the change in outcome values (Ostulund and Balleine 2005, Gremel and Costa 2013). Although goal-directed and habit systems appear at different training phase during instrumental conditioning, both theoretical and experimental work suggested these two systems work in parallel and one of them are dominant depending on the situation (Daw et al., 2005, Smith et al., 2012). Daw et al., proposed that prefrontal dependent tree-search (goal-directed) system and dorsolateral striatum dependent cache (habit) system work in parallel. In case these two systems led to contradict output, one of these systems is chosen based on the reliability of the prediction that they generate. This model explains results of experimental works that studied instrumental conditioning in rats (Killcross and Coutureau 2003, Holland 2004). Smith et al., optogenetically inactivated the infralimbic

cortex (IL) after rats' behavior becoming habitual. They showed that transient inactivation of the IL suppressed the habit and made rats' behavior sensitive to outcome value again. Their results support two systems co-exist in the brain even after a long period of training. The increased choice bias observed in our experiments can be interpreted as the emergence of habit system, and this is the opposite change compared to what was observed in Smith et al. This is the first experimental support that the transient manipulation can switch behavioral control from a goal-directed manner to habitual manner.

Although the switch between left and right choices in the probabilistic reversal task is not the acquisition of novel instrumental responses as is studied in the previous study, it is possible that suppression of goal-directed control of behavior led to choice bias induced by pre-choice inactivation of the ACC or the OFC. Considering the lack of flexibility in choice shown by insensitivity to the previous outcome support the view that the control of goal-directed choice is disrupted by inactivation of the ACC or the OFC. This idea is also supported by the recent study that conducted muscimol inactivation of the ACC (Tervo et al., 2014). Tervo et al., trained rats for matching pennies task in which rats are required to behave randomly regardless of previous choices and outcomes. Even under such requirement, control animals showed some dependency on the outcome. However, once the ACC is inactivated rats such outcome dependency disappeared.

Thus, both lateralized control of choice and the emergence of habitual behavior can account for behavioral effects on action selection. Deficits in the lateralized control of choice are relevant to contralesional neglect in patients with unilateral damages in the PFC or premotor cortex (Kerkhoff 2001). The imbalance of activity between the left and

right hemisphere in these cortical areas can lead to similar behavioral deficits. The transition from goal-directed to habitual behavioral control can be the cause of pathological behaviors such as compulsion and drug addiction (Baline and Everitt 2007; Gillan and Robbins 2014). The possibility of rapid transition between goal-directed and habitual behavior raise the potential treatments for pathological behaviors by manipulating neuronal activity in prefrontal areas.

5.5 Neural circuits for action initiation, action selection, and learning

In the current study, we developed a behavioral paradigm in mice to investigate different cognitive processes required for value-based decision making. Through temporally precise inactivation of genetically defined populations, we could successfully dissociate action-initiation, action-selection, and learning components of value-based decision making. The underlying neural circuits of each process revealed in the current study were summarized in Figure 5.1.

We revealed the contribution of multiple circuit elements in the PFC and the NAc to action initiation (Figure 5.1 A). In the PFC inactivation experiments, inactivation of each of the ACC and the OFC increased the frequency of premature responses regardless of the laser power. From the NAc inactivation experiments, indirect pathway MSNs, dopaminergic input to the NAc, and glutamatergic input to the NAc were shown to be involved in action initiation. Since the PFC is one of the primary source of glutamatergic input to the NAc, these results indicate the significance of the direct projection from the PFC to the NAc in action initiation. Inactivation of dopaminergic input to the NAc also

affected action initiation probably through altered dopaminergic modulation of neuronal activity or plasticity in corticostriatal synapses (Wang et al. 2006).

Action selection was affected only by inactivation of the PFC, and inactivation of each of the ACC and the OFC increased the choice bias (Figure 5.1 B). However, OFC inactivation affected action selection only at high laser power while ACC inactivation was effective even at medium laser power. This result suggests more preferential contribution of the ACC to action selection. Since, the choice bias was not increased by NAc inactivation experiments, the projections from the PFC to other subcortical structures are the candidate projection targets involved in action selection. The potential targets include the dorsal striatum, superior colliculus, and brainstem motor areas.

We also revealed the contribution of the NAc to learning specifically after the absence of a reward. This effect was observed in inactivation of direct pathway MSNs, indirect pathway MSNs or dopaminergic input to the NAc (Figure 5.1 C). Midbrain dopaminergic neurons not only project to the NAc but also receive inputs from the striatum both directly and indirectly (Watanabe-Uchida et al. 2012). This anatomical characteristic implies that the reciprocal interaction between the NAc and midbrain dopaminergic neurons is necessary for the outcome-specific learning effect observed in the current study.

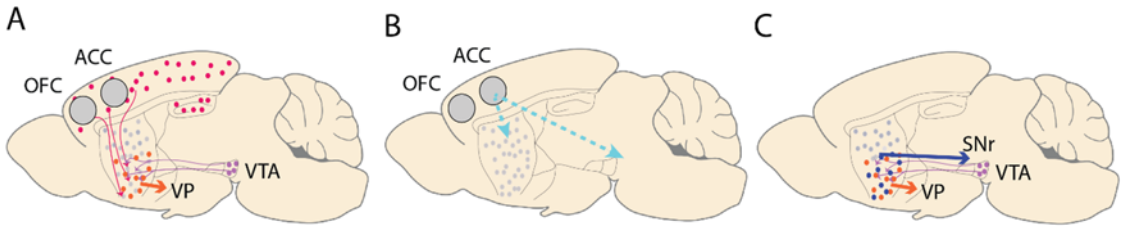


Figure 5.1 Contribution of the NAc and the PFC to cognitive processes required for decision making.

(A) Neural circuits for action initiation. PFC, two inputs to the NAc, and indirect pathway MSNs were implicated in action initiation process. (B) Neural circuits for action selection. Only the ACC (and the OFC at high power) are affected. In contrast NAc inactivation did not affect action selection. Projection from the ACC to subcortical structures other than the NAc may be involved in action selection. (C) Neural circuits for learning. Inactivation of both direct and indirect pathway MSNs and projection from the VTA to the NAc affected the learning.

5.6 Summary and Directions

The significance of the current study lies in the fact that some of the genetic targeting conducted in the current study are so far not possible in other model animals such as rats and primates. Thus, results obtained from cell type specific and projection specific inactivation experiments provided new insights about neural substrates of value-based decision making. The direction that further takes advantage of the strength of mouse genetics is to utilize other specific Cre mouse lines, especially in the PFC. Although PFC pyramidal neurons were non-specifically inactivated in the current study, use of Cre mouse lines that target specific subclasses of cortical neurons will further advance our understanding of neural circuits for decision making. The candidate cell types of the future studies include specific subpopulations of cortical interneurons and pyramidal neurons with distinct projection targets such as corticostriatal projection neurons. Activation of specific classes of interneurons can be used for inactivating subsets of pyramidal neurons. Corticostriatal neurons are not homogeneous populations and include subpopulations that project to different substructures of the striatum such as the patch and matrix compartments (Gerfen 1989, Gerfen et al. 2013). Cre mouse lines were indispensable to dissociate these populations, and inactivation experiments using distinct corticostriatal Cre mouse lines will contribute to define novel functional units in the basal ganglia circuit. Along with the investigation of roles of specific cell types, the probabilistic reversal with different size, probability and delay of reward will also assist further dissociate cognitive processes that were not tested in the current study. For example, insertion of delays to reward delivery will help us understand how premature responses or impulsivity was induced. The premature responses observed in the current

study can be induced by both the overrepresentation of past outcomes and the expectation of future outcomes. These possibilities can be dissociated by investigating the ability of animals to wait for delayed reward. If the expectation of future outcomes is disrupted, animals will have difficulties in waiting for the delayed reward while disruption of the representation of the past outcomes will not affect this process. Since the behavioral paradigms can be modified to see different cognitive components, further study will expand the findings in the current study to different behavioral paradigms, cell types, and brain areas.

Chapter 6: Materials and Methods

6.1 Mice

All procedures were approved by The Rockefeller University Institutional Animal Care and Use Committee (IACUC). All mice used in the current study had a C57BL/6J background. Drd1-Cre (EY266), Adora2a-Cre (A2A-Cre) (KG139), and Slc6a3-Cre (SG62) lines were generated at GENSAT project and maintained at The Rockefeller University. Emx1-Cre (Stock # 005628), VGAT-ChR2-EYFP (Stock # 014548), and Ai40D (Stock # 021188) lines were obtained from The Jackson Laboratory. For non-specific inactivation and dopaminergic input specific inactivation of the NAc, only male mice were used. For other experimental conditions, both male and female mice were used. Mice were at least six weeks old at the time of surgery. The number of animals for each experiment and number of test sessions used for analyzes are described in Table 6.1 and Table 6.2.

Table 6.1. Animals used for NAc inactivation experiments.

Area	Inactivation	Mouse ID	# of sessions
NAc	Non-specific	184	10
		203	7
		213	15
		216	16
		221	7
		257	13
		258	12
		259	12
		285	15
NAc	Direct pathway	264	12
		265	12
		266	15
		326	11
		327	11
		328	13
NAc	Indirect pathway	323	15
		324	12
		331	17
		330	15
		333	13
		334	16

Area	Inactivation	Mouse ID	# of sessions
NAc	Dopaminergic inputs	251	11
		273	11
		340	8
		344	11
		345	11
		348	11
NAc	Glutamatergic inputs	243	16
		269	15
		270	16
		280	13
		312	16
		314	14
		315	14
NAc	Control	253	10
		279	16
		286	16
		320	16
		332	15

Table 6.2 Animals used for PFC inactivation experiments.

Area	Mouse ID	Condition	# of sessions
ACC	260	High	9
		Medium	10
		Low	8
		Unilateral	10
	294	High	7
		Medium	8
		Low	7
		Unilateral	8
	296	High	7
		Medium	8
		Low	7
		Unilateral	7
	307	High	8
		Medium	8
		Low	10
		Unilateral	7
	325	High	7
		Medium	7
		Low	7
		Unilateral	7
Control (ACC)	262	High	1
		Medium	3
		Low	1
		Unilateral	0
	292	High	7
		Medium	7
		Low	7
		Unilateral	7
	308	High	6
		Medium	9
		Low	7
		Unilateral	6
	309	High	6
		Medium	7
		Low	8
		Unilateral	7
	310	High	8
		Medium	7
		Low	7
		Unilateral	7

Area	Mouse ID	Condition	# of sessions
OFC	261	High	9
		Medium	7
		Low	7
		Unilateral	10
	293	High	9
		Medium	8
		Low	7
		Unilateral	7
	295	High	9
		Medium	8
		Low	9
		Unilateral	9
	306	High	7
		Medium	7
		Low	7
		Unilateral	7
	311	High	8
		Medium	7
		Low	7
		Unilateral	6
Control (OFC)	197	High	9
		Medium	11
		Low	9
		Unilateral	6
	198	High	9
		Medium	11
		Low	8
		Unilateral	7
	202	High	7
		Medium	12
		Low	6
		Unilateral	7
	204	High	7
		Medium	8
		Low	4
		Unilateral	5
	317	High	9
		Medium	5
		Low	5
		Unilateral	4

6.2 Surgery and virus injection

Adeno-associated viruses were obtained from the UNC Vector Core. On arrival, 10 ul aliquots were made and stored at -80°C until use. AAV2.9-CAG-ArchT virus was used for non-specific inactivation of the NAc, and AAV2.9-FLEX-ArchT virus was used for cell-type specific or input specific inactivation of the NAc.

A mixture of ketamine (100 mg) and xylazine (1 mg) was used for anesthesia. After the animals were anesthetized, a bilateral craniotomy was performed, and viral solutions were delivered stereotaxically using a Hamilton syringe (7647-01, Hamilton Company) with a 33 gauge metal needle (7803-05, Hamilton Company). Viruses were injected into the NAc (AP: 1.3, ML: +/- 1.05, DV: -3.7 mm from bregma) or the VTA (AP: 3.2, ML: +/- 0.5, DV: -4.1 and -4.6 mm). Following virus injection, fiber optic cannulas (CFML12L05, Thorlab) were implanted bilaterally 0.3 mm above the injection sites in the NAc. For PFC inactivation experiments, fiber optic cannulas (CFML12L02, Thorlab) were implanted bilaterally in the ACC (AP: 1.2, ML: 0.4, DV: -0.85 mm)or the OFC (AP: 2.6, ML: 1.5, DV: -1.4 mm). Control animals received the same craniotomy and fiber optic cannula implant at each coordinate but did not receive an injection of viruses.

6.3 Configuration of the operant chamber

All behavioral experiments were performed in an operant chamber (ENV-307A, Med Associates). The four-sided operant chamber had a stainless steel grid floor and, the inside of the chamber was illuminated using house light. The operant chamber was put in a sound and light attenuating cubicle equipped with a fan for ventilation and providing background noise. A nose-port (ENV-313M, Med Associates) was placed in the center of

one wall, and two food magazines were placed on the left and right sides of the center nose-port. The food magazines were made using a 3D printer and equipped with an infrared sensor (ENV-303HDA, Med Associates) and a yellow LED (ENV-321DM, Med Associates). The food magazines were connected with a food pellet dispenser (ENV-203-14P, Med Associates) that released 14mg food pellets (F05684 BioServ). A tone generator (ENV-323AM, Med Associates) and a click sound generator (ENV-135M, Med Associates) were used to provide auditory feedback for outcome delivery.

6.4 Behavioral training procedures

After recovery from surgery, animals were food restricted to gradually decrease body weight. During food restriction, about 2 g of home cage chow were provided. In addition, a small amount of food pellets that were used during behavior training were provided to familiarize animals with them. Once the body weights reach 80-90% of their free-feeding body weights, behavioral training started.

Magazine training was 15 minutes long. One pellet was delivered to both left and right food magazines every minute (15 pellets total for each food magazine). If animals did not consume all pellets at the end of a session, they were left in the chamber for up to 1hr. If animals consumed more than half of pellets after 1 hour, they were moved to step 1. Otherwise, magazine training was repeated on the following day.

In step 1, animals were trained to poke their noses into the center nose-port. At the start of a training session, an LED inside the center nose-port was turned on. Once animals nose-poked to the illuminated center nose-port, the LED was turned off, and one food

pellet was delivered into one of the two food magazines. Detection of a head entry to the magazine resulted in the onset of the center LED after 4 seconds interval. Once mice received 30 or more pellets in a 30 minutes daily session for two days, animals were moved to Step 2. If animals failed to nose-poke to the center nose-port within a 30 minutes session, training duration was extended up to 3 hours.

In step 2, animals were trained to choose either the left or right food magazine following the center nose-poke. At the start of each trial, the center LED was turned on as in step 1. Once a nose-poke was made, the center LED was extinguished, and an LED inside either the left or right food magazine was turned on. An entry to the illuminated food magazine resulted in food pellet delivery while an entry to the other food magazine resulted in a time-out (house light was turned off for 3 seconds). Four seconds after the magazine entry, the center LED was turned on again, and a new trial began. In step 2, entries to an unilluminated center nose-port or food magazines resulted in a 3-second time-out to prevent indiscriminative nose-pokes or magazine entries. Once animals received 30 or more food pellets in a 30 minutes session for two days, they were moved to step 3. If animals did not receive any food pellet in a 30 minutes session, training duration was extended up to 3 hours.

In step 3, animals were required to initiate trials by making a center nose-poke as in the previous step. Left or right magazine entries during the illumination of the center LED were recorded as premature responses that was followed by a 3-second time-out and ITI. Failure to make a center nose-poke within 20 seconds from center LED onset was recorded as trial omission that is also followed by a 3-second time-out and ITI. After a center nose-poke, both left and right LEDs were turned on, and animals could freely

choose either the left or right food magazine. At the beginning of each behavioral session, one of the two magazines was assigned as the correct side, and the other food magazine was assigned as the incorrect side. The probability of food pellet delivery was set to 90% for the correct side and 0% for the incorrect side. Delivery or absence of a pellet was followed by distinct auditory feedback. Choice history of animals was recorded during a training session. Once animals reached a performance criterion (eight or more correct choices over the last ten trials), the correct and incorrect side were stochastically switched. This switch of reward probability was called a 'reversal', and the reversal can occur at the beginning of each trial. There was a 15% chance that the correct and incorrect sides would be switched. Once the reversal occurred, the position was kept constant until animals reach the performance criterion again. If animals achieved 3 or more reversals in a daily 60-minute session, they were moved to step 4.

In step 4, reward probability of each food magazine was set to 75% and 0%. The rest of conditions were same as in step 3. Once animals achieved 3 or more reversals in a daily session, they were moved to testing sessions.

6.5 Optogenetic inactivation procedures

For inactivation of both the PFC and the NAc, animals were tested in the same condition as training step 4. 532 nm continuous green light (6-10mW) was used for NAc inactivation, and 473 nm light (1.5-2.5, 4-5, 12-15 mW) was used for PFC inactivation. The 473 nm light pulses were delivered in 5 ms, 40 Hz trains. In both NAc and PFC inactivation experiments, light was delivered at three different time periods including pre-start period, pre-choice period and post-choice period, and the light was delivered at each

time period at 10% of trials. During pre-start inactivation trials, light was delivered from the center LED onset to the center nose-poke. If premature responses or trial omissions occurred during pre-start inactivation trials, the light was turned off during ITI, and the pre-start inactivation trial was repeated. During pre-choice inactivation trials, light was delivered from the center nose-poke to the either left or right magazine entry. During post-choice inactivation trials, light was delivered at left or right magazine entry and lasted for 3 seconds.

6.6 Histology

After behavioral experiments, animals were deeply anesthetized and transcardially perfused with 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS). Dissected brains were put in 4% PFA/PBS solutions overnight for post fixation and moved to 30% sucrose solution. Once brains sunk on the bottom of the tube, they were embedded in OCT compound, and 35 um free-floating coronal sections were made on a cryostat.

For immunohistochemistry, free-floating sections were blocked with 5% normal goat serum and 0.25% TritonX-100 in PBS. Sections were incubated with an anti-GFP antibody (1:1000, Chicken polyclonal, ab13970, Abcam) overnight at room temperature. Next day, sections were rinsed with PBS three times with PBS for 5 minutes and incubated with a secondary antibody conjugated with Alexa Fluor® 488 dye (1:500, Goat anti-chicken IgG, A-11039, Invitrogen) for 90 minutes. After rinsing with PBS three times, sections were mounted on slides and coverslipped with mounting media (ProLong Gold Antifade Reagent with DAPI, P-36931, Invitrogen). Images were acquired with a confocal microscope.

6.7 Statistical analysis

For statistical analysis of premature response, and trial omission, the frequency of each event was defined as the number of events divided by the number of trials performed. The frequency was calculated separately for laser off and inactivation conditions. From an individual animal, the pair of values was obtained for each daily session. Pairs of values from different animals were pooled, and the difference between the laser off condition and each inactivation condition was tested with Wilcoxon signed-rank test.

Statistical analysis of reaction time was performed in a similar way. The median reaction time of each inactivation condition was used as the representative value of a daily session. The difference in performance between the laser off and each inactivation condition was tested with Wilcoxon signed-rank test. In case reaction time from multiple trial conditions (rewarded and unrewarded trials) or three inactivation conditions (Laser off, pre-start and pre-choice inactivation) were compared together, p-values were corrected with Bonferroni method.

For stay probability, the number of stay choices and shift choices were counted for each inactivation condition. The number of events was summed over all behavioral events, and significance was tested using chi-squared test. The comparison was performed between laser off condition and each inactivation condition. In case, different trial types (rewarded and unrewarded trials) or three inactivation conditions (Laser off, pre-start and pre-choice inactivation) were compared together, p-values were corrected with Bonferroni method.

Choice biases observed in PFC inactivation experiments were quantified with bias index defined with the following formula.

Bias index = $|StayP_{L(+)} - StayP_{R(+)} + StayP_{L(-)} - StayP_{R(-)}|$

$StayP_{L(+)}$: Stay probability after left choice rewarded trials

$StayP_{L(-)}$: Stay probability after right choice rewarded trials

$StayP_{R(+)}$: Stay probability after left choice unrewarded trials

$StayP_{R(-)}$: Stay probability after right choice unrewarded trials

Bias index was calculated for each behavioral session. Since the absolute value was taken to calculate bias index and stay probability can have extreme values with small number of trials, the scaling of bias index tend to be larger with smaller number of trials. In order to avoid the effect of difference in the number of trials, the bias index of laser off condition was calculated using 1 / 7 of laser off trials randomly sampled from each behavioral session so that approximate trial number of each condition became same. To avoid the effect of the sampling bias of laser off trials, bias index of laser off was calculated 30 times, and the median of them was used as the bias index of laser off trials. This sampling doesn't affect the scaling of the bias index. The bias indices of each inactivation condition was calculated for daily behavioral sessions. The bias index of each inactivation condition was compared with that of the laser off condition using Wilcoxon signed-rank test. P-values were corrected with Bonferroni method. Laser effects shown in Figure 3.8 E were calculated using the following formula.

Laser effect = Bias index $_{(Laser\ on)}$ – Bias index $_{(Laser\ off)}$

(Laser on) in the above calculation means either pre-choice inactivation in bilateral inactivation experiments or pre-choice left inactivation in unilateral inactivation experiments. The value was calculated for each daily session. The values were pooled separately for unilateral or bilateral inactivation, and the difference was tested with Wilcoxon rank-sum test.

To test whether inactivation effect on stay probability depends on start reaction time, inactivation effect difference was calculated using the following formula.

Inactivation effect difference = $(StayP_{(ON, Short)} - StayP_{(OFF, Short)}) - (StayP_{(ON, Long)} - StayP_{(OFF, Long)})$

StayP : Stay probability after unrewarded trials

ON/OFF : Laser ON/OFF of the previous trial

Short/Long RT : Short or long reaction time of the current trial

The inactivation effect difference was compared with the distribution obtained by permuted data. For each behavioral session, number of trials with each inactivation condition was counted. Permuted data had the same sequence of choice, outcome, and reaction time, but labeling of inactivation condition was randomly assigned so that total number of trials with each inactivation condition matched with behavioral data. Using the permuted data, stay probability and inactivation effect difference were calculated. This process was repeated 1000 times to obtain the distribution of inactivation effect difference. Using the distribution, the rank order of inactivation effect difference calculated from the behavioral data was divided by the number of bootstrapping and used as a p-value.

Chapter 7: References

1. Abela, A. R. & Chudasama, Y. Dissociable contributions of the ventral hippocampus and orbitofrontal cortex to decision-making with a delayed or uncertain outcome. *Eur. J. Neurosci.* **37,** 640–647 (2013).

2. Ahmari, S. E. *et al.* Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science* **340,** 1234–9 (2013).

3. Albin, R. L., Young, a B. & Penney, J. B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12,** 366–375 (1989).

4. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9,** 357–381 (1986).

5. Arkes, H. R. & Blumer, C. The psychology of sunk cost. *Organ. Behav. Hum. Decis. Process.* **35,** 124–140 (1985).

6. Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J. & Robbins, T. W. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.* **6,** 115–116 (2003).

7. Atallah, H. E., McCool, A. D., Howe, M. W. & Graybiel, A. M. Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron* **82,** 1145–1156 (2014).

8. Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W. & O'Reilly, R. C. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* **10,** 126–31 (2007).

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

10. Bamford, N. S. *et al.* Dopamine modulates release from corticostriatal terminals. *J. Neurosci.* **24,** 9541–9552 (2004).

11. Bari, A. *et al.* Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* **35,** 1290–1301 (2010).

12. Barker, J. M. *et al.* Epigenetic and pharmacological regulation of 5HT3 receptors controls compulsive ethanol seeking in mice. *Eur. J. Neurosci.* **39,** 999–1008 (2014).

13. Belin, D. & Everitt, B. J. Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum. *Neuron* **57,** 432–441 (2008).

14. Berendse, H. W., Galis-de Graaf, Y. & Groenewegen, H. J. Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J. Comp. Neurol.* **316,** 314–347 (1992).

15. Brown, L. L., Smith, D. M. & Goldbloom, L. M. Organizing principles of cortical integration in the rat neostriatum: Corticostriate map of the body surface is an ordered lattice of curved laminae and radial points. *J. Comp. Neurol.* **392,** 468–488 (1998).

16. Burguière, E., Monteiro, P., Feng, G. & Graybiel, A. M. Optogenetic Stimulation of Lateral Orbitofronto-Striatal Pathway Suppresses Compulsive Behaviors. *Science (80-.).* **340,** 1243–1246 (2013).

17. Campos, M., Breznen, B., Bernheim, K. & Andersen, R. a. Supplementary motor area encodes reward expectancy in eye-movement tasks. *J. Neurophysiol.* **94,** 1325–1335 (2005).

18. Cardinal, R. N., Robbins, T. W. & Everitt, B. J. The effects of d -amphetamine, chlordiazepoxide, α-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology (Berl).* **152,** 362–375 (2000).

19. Chudasama, Y. *et al.* Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: Differential effects on selectivity, impulsivity and compulsivity. *Behav. Brain Res.* **146,** 105–119 (2003).

20. Churchland, A. K., Kiani, R. & Shadlen, M. N. Decision-making with multiple alternatives. *Nat. Neurosci.* **11,** 693–702 (2008).

21. Corbit, L. H., Muir, J. L. & Balleine, B. W. The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* **21,** 3251–3260 (2001).

22. Corbit, L. H. & Balleine, B. W. The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J. Neurosci.* **31,** 11786–94 (2011).

23. Cui, G. *et al.* Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* **494,** 238–42 (2013).

24. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* **69,** 680–694 (2011).

25. Dalton, G. L., Phillips, A. G. & Floresco, S. B. Preferential involvement by nucleus accumbens shell in mediating probabilistic learning and reversal shifts. *J. Neurosci.* **34,** 4618–26 (2014).

26. Daw, N. D. & Doya, K. The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* **16,** 199–204 (2006).

27. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8,** 1704–1711 (2005).

28. Durstewitz, D., Vittoz, N. M., Floresco, S. B. & Seamans, J. K. Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* **66,** 438–448 (2010).

29. Eblen, F. & Graybiel, a M. Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J. Neurosci.* **15,** 5999–6013 (1995).

30. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72,** 330–343 (2011).

31. Feierstein, C. E., Quirk, M. C., Uchida, N., Sosulski, D. L. & Mainen, Z. F. Representation of Spatial Goals in Rat Orbitofrontal Cortex. *Neuron* **51,** 495–507 (2006).

32. Fillmore, M. T., Fillmore, M. T., Rush, C. R. & Rush, C. R. Impaired inhibitory control of behavior in chronic cocaine users. *Drug Alcohol Depend.* **66,** 265–73 (2002).

33. Fonseca, M. S., Murakami, M. & Mainen, Z. F. Activation of Dorsal Raphe Serotonergic Neurons Promotes Waiting but Is Not Reinforcing. *Curr. Biol.* 1–10 (2015). doi:10.1016/j.cub.2014.12.002

34. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic Coding of Visual Space in the Monkey ' s Dorsolateral Prefrontal Cortex. *J. Neurophysiol.* **61,** 331–349 (1989).

35. Gerfen, C. R. The neostriatal mosaic: striatal patch-matrix organization is related to cortical lamination. *Science* **246,** 385–388 (1989).

36. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron* **80,** 1368–1383 (2013).

37. Gimenez-Amaya, J. M. & Graybiel, a. M. Compartmental origins of the striatopallidal projection in the primate. *Neuroscience* **34,** 111–126 (1990).

38. Gong, S. *et al.* Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci.* **27,** 9817–9823 (2007).

39. Goshen, I. *et al.* Dynamics of retrieval strategies for remote memories. *Cell* **147,** 678–689 (2011).

40. Graybiel, a M. Neurotransmitters and neuromodulators in the basal ganglia. *Trends Neurosci.* **13,** 244–254 (1990).

41. Graybiel, A. M. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31,** 359–387 (2008).

42. Gremel, C. M. & Costa, R. M. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat. Commun.* **4,** 2264 (2013).

43. Guo, Z. *et al.* Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81,** 179–194 (2014).

44. Hall, J., Parkinson, J. a., Connor, T. M., Dickinson, a. & Everitt, B. J. Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating pavlovian influences on instrumental behaviour. *Eur. J. Neurosci.* **13,** 1984–1992 (2001).

45. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* (2015). doi:10.1038/nature14066

46. Holland, P. C. Unblocking in Pavlovian appetitive conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **10,** 476–497 (1984).

47. Hwang, J., Kim, S. & Lee, D. Temporal discounting and inter-temporal choice in rhesus monkeys. *Front. Behav. Neurosci.* **3,** 9 (2009).

48. Ito, M. & Doya, K. Distinct Neural Representation in the Dorsolateral, Dorsomedial, and Ventral Parts of the Striatum during Fixed- and Free-Choice Tasks. *J. Neurosci.* **35,** 3499–3514 (2015).

49. Ito, M. & Doya, K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* **29,** 9861–9874 (2009).

50. Jentsch, J. D. & Taylor, J. R. Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of behavior by reward-related stimuli. *Psychopharmacology (Berl).* **146,** 373–390 (1999).

51. Jimenez-Castellanos, J. & Graybiel, a M. Subdivisions of the dopamine-containing A8-A9-A10 complex identified by their differential mesostriatal innervation of striosomes and extrastriosomal matrix. *Neuroscience* **23,** 223–242 (1987).

52. Joel, D. & Weiner, I. The organization of the basal ganglia-thalamocortical circuits: Open interconnected rather than closed segregated. *Neuroscience* **63,** 363–379 (1994).

53. Karlsson, M. P., Tervo, D. G. R. & Karpova, a. Y. Network Resets in Medial Prefrontal Cortex Mark the Onset of Behavioral Uncertainty. *Science (80-.).* **338,** 135–139 (2012).

54. Kennerley, S., Behrens, T. & Wallis, J. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* **14,** 1581–1589 (2011).

55. Kerkhoff, G. Spatial hemineglect in humans. *Progress in Neurobiology* **63,** 1–27 (2001).

56. Killcross, S. & Coutureau, E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* **13,** 400–408 (2003).

57. Kincaid, a E., Zheng, T. & Wilson, C. J. Connectivity and convergence of single corticostriatal axons. *J. Neurosci.* **18,** 4722–4731 (1998).

58. Knowlton, B. J., Mangels, J. a & Squire, L. R. A neostriatal habit learning system in humans. *Science* **273,** 1399–1402 (1996).

59. Komiyama, T. *et al.* Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464,** 1182–1186 (2010).

60. Krause, M., German, P. W., Taha, S. a & Fields, H. L. A pause in nucleus accumbens neuron firing is required to initiate and maintain feeding. *J. Neurosci.* **30,** 4746–4756 (2010).

61. Kravitz, a V *et al.* Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466,** 622–626 (2010).

62. Kravitz, a. V. & Kreitzer, a. C. Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. *Physiology* **27,** 167–177 (2012).

63. Kurata, K. & Tanji, J. Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys. II. Responses to movement triggering vs. nontriggering sensory signals. *J. Neurophysiol.* **53,** 142–152 (1985).

64. Lévesque, M. & Parent, A. The striatofugal fiber system in primates: a reevaluation of its organization based on single-axon tracing studies. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 11888–11893 (2005).

65. Magalhães, P. & White, K. G. Persistence in extinction: The sunk time effect. *J. Exp. Psychol. Anim. Learn. Cogn.* **40,** 38–54 (2014).

66. Mar, A. C., Walker, A. L. J., Theobald, D. E., Eagle, D. M. & Robbins, T. W. Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J. Neurosci.* **31,** 6398–404 (2011).

67. Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* **10,** 647–656 (2007).

68. McCoy, A. N. & Platt, M. L. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8,** 1220–1227 (2005).

69. Mcdannald, M. a, Lucantonio, F., Burke, K. a, Niv, Y. & Schoenbaum, G. Ventral Striatum and Orbitofrontal Cortex Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *J. Neurosci.* **31,** 2700–2705 (2011).

70. McGeorge, a J. & Faull, R. L. The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* **29,** 503–537 (1989).

71. Meeson, a, Palmer, M., Calfon, M. & Lang, R. A relationship between apoptosis and flow during programmed capillary regression is revealed by vital analysis. *Development* **122,** 3929–3938 (1996).

72. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72,** 1024–1027 (1994).

73. Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y. & Tanji, J. Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat. Neurosci.* **12,** 502–507 (2009).

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

75. Monterosso, J. R., Aron, A. R., Cordova, X., Xu, J. & London, E. D. Deficits in response inhibition associated with chronic methamphetamine abuse. *Drug Alcohol Depend.* **79,** 273–277 (2005).

76. Murakami, M., Vicente, M. I., Costa, G. M. & Mainen, Z. F. Neural antecedents of self-initiated actions in secondary motor cortex. *Nat. Neurosci.* (2014). doi:10.1038/nn.3826

77. Murphy, E. R., Robinson, E. S. J., Theobald, D. E. H., Dalley, J. W. & Robbins, T. W. Contrasting effects of selective lesions of nucleus accumbens core or shell on inhibitory control and amphetamine-induced impulsive behaviour. *Eur. J. Neurosci.* **28,** 353–363 (2008).

78. Narayanan, N. S., Horst, N. K. & Laubach, M. Reversible inactivations of rat medial prefrontal cortex impair the ability to wait for a stimulus. *Neuroscience* **139,** 865–876 (2006).

79. Narayanan, N. S., Cavanagh, J. F., Frank, M. J. & Laubach, M. Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* **16,** 1888–95 (2013).

80. Narayanan, N. S. & Laubach, M. Neuronal correlates of post-error slowing in the rat dorsomedial prefrontal cortex. *J. Neurophysiol.* **100,** 520–525 (2008).

81. Nelson, S. B., Sugino, K. & Hempel, C. M. The problem of neuronal cell types: a physiological genomics approach. *Trends Neurosci.* **29,** 339–345 (2006).

82. Nicola, S. M. The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *J. Neurosci.* **30,** 16585–16600 (2010).

83. Nieoullon, A. Dopamine and the regulation of cognition and attention. *Prog. Neurobiol.* **67,** 53–83 (2002).

84. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl).* **191,** 507–20 (2007).

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

86. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304,** 452–454 (2004).

87. O'Neill, M. & Schultz, W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68,** 789–800 (2010).

88. Onge, J. R. S. & Floresco, S. B. Prefrontal cortical contribution to risk-based decision making. *Cereb. Cortex* **20,** 1816–1828 (2010).

89. Ostlund, S. B. & Balleine, B. W. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.* **25,** 7763–70 (2005).

90. Parkinson, J. a, Olmstead, M. C., Burns, L. H., Robbins, T. W. & Everitt, B. J. Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.* **19,** 2401–2411 (1999).

91. Pattij, T., Janssen, M. C. W., Vanderschuren, L. J. M. J., Schoffelmeer, A. N. M. & Van Gaalen, M. M. Involvement of dopamine D1 and D2 receptors in the nucleus accumbens core and shell in inhibitory response control. *Psychopharmacology (Berl).* **191,** 587–598 (2007).

92. Pearce, J. M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87,** 532–552 (1980).

93. Pelloux, Y., Everitt, B. J. & Dickinson, A. Compulsive drug seeking by rats under punishment: Effects of drug taking history. *Psychopharmacology (Berl).* **194,** 127–137 (2007).

94. Reddi, B. a & Carpenter, R. H. The influence of urgency on decision time. *Nat. Neurosci.* **3,** 827–830 (2000).

95. Rescorla, R. a & Wagner, a R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Class. Cond. II Curr. Res. Theory* **21,** 64–99 (1972).

96. Robinson, T. E. & Berridge, K. C. Addiction. *Annu. Rev. Psychol.* **54,** 25–53 (2003).

97. Roesch, M. R., Taylor, A. R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51,** 509–20 (2006).

98. Romo, R. & Schultz, W. Role of Primate Basal Ganglia and Frontal-Cortex in the Internal Generation of Movements .3. Neuronal-Activity in the Supplementary Motor Area. *Exp. Brain Res.* **91,** 396–407 (1992).

99. Rudebeck, P. H. *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* **28,** 13775–85 (2008).

100. Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M. & Rushworth, M. F. S. Separate neural pathways process different decision costs. *Nat. Neurosci.* **9,** 1161–8 (2006).

101. Sakai, S. T., Inase, M. & Tanji, J. Pallidal and cerebellar inputs to thalamocortical neurons projecting to the supplementary motor area in Macaca fuscata: A triple-labeling light microcopic study. *Anat. Embryol. (Berl).* **199,** 9–19 (1999).

102. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310,** 1337–40 (2005).

103. Schachar, R., Tannock, R., Marriott, M. & Logan, G. Deficient inhibitory control in attention deficit hyperactivity disorder. *J. Abnorm. Child Psychol.* **23,** 411–436 (1995).

104. Scheres, a., Lee, a. & Sumiya, M. Temporal reward discounting and ADHD: Task and symptom specific effects. *J. Neural Transm.* **115,** 221–226 (2008).

105. Scheres, A., Milham, M. P., Knutson, B. & Castellanos, F. X. Ventral Striatal Hyporesponsiveness During Reward Anticipation in Attention-Deficit/Hyperactivity Disorder. *Biol. Psychiatry* **61,** 720–724 (2007).

106. Schmidt, L. *et al.* Disconnecting force from money: Effects of basal ganglia damage on incentive motivation. *Brain* **131,** 1303–1310 (2008).

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

108. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science (80-.).* **282,** 1335–8 (1998).

109. Smith, K. S., Berridge, K. C. & Aldridge, J. W. Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proc. Natl. Acad. Sci. U. S. A.* **108,** E255–E264 (2011).

110. Stopper, C. M., Green, E. B. & Floresco, S. B. Selective Involvement by the Medial Orbitofrontal Cortex in Biasing Risky, But Not Impulsive, Choice. *Cereb. Cortex* (2012). doi:10.1093/cercor/bhs297

111. Sul, J. H., Jo, S., Lee, D. & Jung, M. W. Role of rodent secondary motor cortex in value-based action selection. *Nat. Neurosci.* **14,** 1202–8 (2011).

112. Sul, J. H., Kim, H., Huh, N., Lee, D. & Jung, M. W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66,** 449–460 (2010).

113. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. (A Bradford Book, 1998).

114. Tai, L.-H., Lee, a M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* **15,** 1281–9 (2012).

115. Tanji, J. & Kurata, K. Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys. I. Responses to instructions determining motor responses to forthcoming signals of different modalities. *J. Neurophysiol.* **53,** 129–141 (1985).

116. Tervo, D. G. R. *et al.* Behavioral Variability through Stochastic Choice and Its Gating by Anterior Cingulate Cortex. *Cell* **159,** 21–32 (2014).

117. Thura, D., Beauregard-Racine, J., Fradet, C.-W. & Cisek, P. Decision-making by urgency-gating: theory and experimental support. *J. Neurophysiol.* (2012). doi:10.1152/jn.01071.2011

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

119. Voorn, P., Vanderschuren, L. J. M. J., Groenewegen, H. J., Robbins, T. W. & Pennartz, C. M. a. Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* **27,** 468–474 (2004).

120. Wang, A. Y., Miura, K. & Uchida, N. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat. Neurosci.* **16,** 639–47 (2013).

121. Wang, Z. *et al.* Dopaminergic Control of Corticostriatal Long-Term Synaptic Depression in Medium Spiny Neurons Is Mediated by Cholinergic Interneurons. *Neuron* **50,** 443–452 (2006).

122. Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A. & Uchida, N. Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. *Neuron* **74,** 858–873 (2012).

123. Wikenheiser, A. M., Stephens, D. W. & Redish, a D. Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 8308–13 (2013).

124. Winstanley, C. a, Theobald, D. E. H., Cardinal, R. N. & Robbins, T. W. Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* **24,** 4718–22 (2004).

125. Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22,** 513–523 (2005).

126. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19,** 181–189 (2004).