The Role of the Lateral Habenula in inhibitory-driven action selection

A thesis submitted in fulfillment of the requirements for the degree of

Master of Philosophy

By

Felix Lok Tin Wong

Sydney Medical School

University of Sydney

2016

Abstract

In order to behave adaptively, animals are required to use cues that predicts the presence or the absence of the desired outcome to guide their selection of actions. Previous studies of how cue influences on choice, using the Specific Pavlovian Instrumental Transfer (S-PIT) paradigm, have yielded considerable progress in our understanding of the underlying mechanism. However, most of them focused on the influence from the excitatory association, where the cue is signalling the presence of outcome. While some pioneering studies have demonstrated the possibility of a cue predicting the absence of outcome affecting the animal's choice behaviour, the neural mechanism that specific to this effect is still largely unexplored. Therefore, the aim of the current thesis was to investigate the role of the Lateral Habenula (LHb) in this type of inhibitory-driven action selection process, as there is reasonable evidences suggesting that this region involved heavily in processing cue that signal the absence of outcome. Our result has shown that the LHb lesion i) weakened the effect of conditioned inhibition, ii) abolished the reversed S-PIT effect that caused by the negative predicting cue, but iii) not affected the normal S-PIT that elicited by the positive predicting cue nor the choice bias that based on the value of the outcomes. Overall speaking, it suggested that the LHb is essential for stimulus-based, and not value-based, choice in situations where the stimuli have been trained as negative, but not positive, predictors of their associated outcomes.

Statement of Authentication

I, Felix Lok Tin Wong, declare that this thesis is submitted in fulfillment of the requirement for the award of Master of Philosophy, in the School of Medicine, University of Sydney.

To the best of my knowledge, the work presented in this thesis is original unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Felix Lok Tin Wong

31st March 2016

Care and use of animals

The research presented in this thesis was approved by the Animal Ethics Committee at the University of Sydney. The care and use of animals complies with the Rules Governing the Use of Animals in Research and Teaching at the University of Sydney with the Australia Code of Practice for the Care and Use of Animals for Scientific Purposes Act (1985 and its subsequent amendments).

Acknowledgements

I would like to express my greatest gratitude to my supervisors, Prof. Bernard Balleine and Dr. Vincent Laurent for their guidance. Without their support, I would never be able to have the opportunity to be a Mphil student and finish the thesis.

I would also want to thank all my colleagues. Thanks to Ash, Billy, Genevra, Serena and Shauna who taught me a lot of technique in preforming animal surgeries and in various lab works.

Last but not least, thanks to my family for their unconditional support and encouragement.

Table of Contents

List of Figures

Chapter 5

Appendixes

List of Tables

Chapter 1

List of Abbreviations

5-HT - Serotonin

A - Action

AESOP - Affective Extension of the Sometimes Opponent Process

A-O - Action-outcome association

BLA - Basolateral Amygdala

CeA - Central nucleus of amygdala

ChAT - Choline acetyltransferase

CIN - Cholinergic interneuron

CRF - Continuous reinforcement

D1R - Dopamine D1 receptors

D2R - Dopamine D2 receptors

DA - Dopamine

DBB - Diagonal band of Broca

Different - Responding on the lever associated with the different outcome as the conditioned stimulus

DLS - Dorsal lateral striatum

DMS - Dorsal medial striatum

DOR - Delta opioid receptor

DRN - Dorsal raphe nucleus

DREADDs - Designer receptors exclusively activated by designer drugs

eGFP - Enhanced Green Fluorescent Protein

GABA - Gamma-aminobutyric acid

GABAergic - Gamma-aminobutyric acidergic

Glut - Glutamate

GPb - Border region of the globus pallidus internal segment

GPe - Globus pallidus external segment

GPi - Globus pallidus internal segment

ITI - Inter-trial interval

IPN - Interpeduncular nucleus

LHA - Lateral hypothalamic area

LHb - Lateral habenula

LHb-l - Lateral part of lateral habenula

LHb-m - Medial part of lateral habenula

LPO - Lateral preoptic area

LS - Lateral septum

MDT - Mediodorsal thalamus

MHb - Medial habenula

mPFC - Medial prefrontal cortex

MRN - Median raphe nucleus

MS - Medial septum

MSN - Medium spiny neurons

NAc - Nucleus accumbens

NAc-C - Nucleus accumbens core region

NAc-S - Nucleus accumbens shell region

OFC - Orbital frontal cortex

O - Outcome

PBS - Phosphate buffered saline

PFA - Paraformaldehyde

PFC - Prefrontal cortex

PIT - Pavlovian-instrumental transfer

Pre S - Period before presentation of stimulus

PS - Posterior septum

RMTg - Rostromedial tegmental nucleus

RPE - Reward prediction error

RR - Random ratio

S - Stimulus

S-O - Stimulus-outcome association

S-PIT - Specific Pavlovian-instrumental transfer

Same - Responding on the lever associated with the same outcome as the conditioned stimulus

S.E.M. - Standard error mean

SN - Substantia nigra

SNc - Substantia nigra pars compacta

SNr - Substantia nigra pars reticulata

STN - Subthalamic nucleus

VP - Ventral pallidum

VTA - Ventral tegmental area

Chapter 1: Background and preliminary data.

Everyday we make numerous choices, from what food to eat to what goods to buy. Usually, these choices are determined by the expectation and the appraisal of the outcomes that they procure in relation to our current needs and desires. However, studies in animals and humans have revealed that this decision-making process is heavily controlled by environmental stimuli, such that stimuli associated with specific outcomes bias choice toward actions procuring those same outcomes. This control is adaptive as it encourages animals to obtain commodities the availability of which is signalled by the environment. On the other hand, this control can lead to maladaptive behaviour in situations where the predictive power of the stimuli outweighs the evaluation of the outcome in guiding choice. For example, it is well established that stimuli associated with food can motivate feeding behaviour even in the absence of hunger, providing some insight in the mechanisms that could lead to obesity. Similarly, compelling evidence indicates that stimuli present during drug consumption can trigger craving and relapse after long periods of abstinence despite the negative consequences associated with drug taking.

Given its relevance to obesity and drug addiction, numerous studies have aimed at describing how environmental stimuli influence choice between different courses of actions. Specifically, most, if not all of these studies have focused on the control that stimuli that positively predict a particular outcome (i.e., excitatory stimuli) exert over action selection. However, recent findings have shown that environmental stimuli negatively predicting an outcome (i.e., inhibitory stimuli) are also able to guide choice between actions. In fact, the ability to use inhibitory stimuli may help to optimize behaviour by avoiding the selection of actions that are unlikely to earn any reward. For example, fruit picking in a field where there are signs this it has recently been explored by others is less likely to produce a good yield and one should turn to harvest elsewhere. However,

how inhibitory stimuli regulate our choices remains largely unknown. Therefore, the aim of this thesis was to investigate inhibition-driven decision-making processes and some of the neural circuitry involved. In the coming sections, I will review previous work on stimulus guided decision-making, and will then explore the underlying circuitry that contributes to this decision-making process.

Goal-directed responding

Animals can readily learn about the contingency between a specific action and its outcome (A-O), and they also learn to perform this action when the value of its outcome satisfies their current needs and desires (de Wit & Dickinson, 2009). Accordingly, when the delivery of an outcome becomes independent from the action that the animal has performed, such actions will be less likely to be performed in the future as its association with the outcome has been degraded (Balleine & Dickinson, 1998). Likewise, if the outcome is devalued, either by pairing it with illness (Adams & Dickinson, 1981; Colwill & Rescorla, 1985; Dickinson & Balleine, 1994) or by inducing sensory-specific satiety with free access to the outcome (Balleine and Dickinson, 1998), the associated action will be less likely to be carried out as it now procures a less desirable outcome.

Choice guided by excitatory cues: Pavlovian-Instrumental transfer

Although the outcome is always the consequence of the action in most instrumental learning paradigms, the relationship between action and outcome has been suggested to be bidirectional, meaning that not only A-O but also O-A associations are formed during the learning stage. Ostlund and Balleine (2007) have in fact demonstrated that such O-A associations can dominate A-O ones to guide response selection. In their study, animals were trained with different food rewards preceding and/or following specific actions. In total two pairs of Outcome-Action-Outcome (O-A-O) relationships were trained, i.e. O2-A1-O1 and O1-A2-O2. After extinguishing responding

on both actions, delivery of O1 re-invoked A2 while delivery of O2 promoted A1, suggesting that action selection was primarily driven by O1-A2/O2-A1 rather than A1-O1/A2-O2. This implies that reinstatement of instrumental responding is mostly driven by the signaling relationships between the stimulus properties of an outcome (e.g., taste, odour and so on) and the action associated with that outcome.

Given the aforementioned role of O-A in response selection, it is reasonable to assume that any association involving the representation of a particular outcome could also influence choice. One example of such associations is that formed between a stimulus and an outcome during Pavlovian conditioning (i.e., S-O). And indeed, Pavlovian-Instrumental transfer (PIT) has provided ample evidence that S-O associations can influence choice. This paradigm consists of three stages. The first stage is Pavlovian conditioning during which subjects learn that two stimuli predict two distinct outcomes (i.e., S1-O1 and S2-O2). Instrumental training is then conducted and involved learning that the two outcomes can be earned by performing two distinct actions (i.e., A1-O1 and A2-O2). Finally, the subjects received a PIT test that assesses choice between the two instrumental actions in the absence or presence of either stimulus. This test constantly shows that a stimulus triggers higher responding on the action with which it shares a common outcome (i.e., the “Same” action: A1 during S1 and A2 during S2) than on the other action (i.e., the “Different” response: A2 during S1 and A1 during S2) or than in the absence of stimuli (i.e., baseline performance). In other words, a stimulus predicting a particular outcome biases choice towards actions delivering that same outcome.

Recent studies have suggested that the elevation of responding observed in PIT could be caused by the excitation of the outcome representation through two different processes: a general motivational process and an outcome sensory specific process (Corbit & Balleine, 2003; Corbit,

Muir & Balleine, 2001). The first process refers to the affective properties of the outcome – i.e., appetitive or aversive – while the second process deals with the perceptible properties of the outcome – e.g., taste, olfactory or texture. Studies by Corbit and colleagues (Corbit & Balleine, 2005; 2011; Corbit, Janak, & Balleine, 2007) have shown how the two processes can drive choice behaviour in different ways. In their studies (**see Table 1**), animals were trained with three pairs of stimulus-outcome associations (S1-O1, S2-O2, S3-O3) followed by instrumental training of A1-O1 and A2-O2. In a subsequent transfer test, they found that S1 and S2 only elevated above baseline responding on the action with which they shared a common outcome (S1: A1>A2; S2: A2>A1). In contrast, S3 increased performance of the two trained instrumental actions above baseline. However, if the test was carried out when the animal was sated, S3 failed to produce any increase in performance while S1 and S2 remain capable of: (i) biasing choice towards the action earning the outcome that they used to predict and (ii) increasing responding on that action above baseline. Accordingly, the authors suggested that the elevation effect by S3, which is termed general PIT, was driven by the general motivational process of O3 to boost responding on all actions associated with an outcome of the same motivation category, e.g. increasing all actions linked with an appetitive outcome. Thus, general PIT disappeared when the animal was sated, as the outcome had lost its positive/appetitive value. On the other hand, they proposed that the effects exerted by S1 and S2, which was outcome-specific and therefore termed as specific PIT, was driven by the sensory specific properties of O1 and O2 and elevated responding on the actions through the corresponding O-A association. Therefore, specific PIT was able to resist reduction of appetite or even outcome devaluation (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994).

Instrumental training	Pavlovian training	PIT test - Hungry	PIT test – Sated
A1-O1 A2- O2	S1-O1 S2-O2 S3-O3	S1: A1 v A2 S2: A1 v A2 S3: A1 v A2	S1: A1 v A2 S2: A1 v A2 S3: A1 v A2

Table 1 – The experimental design of Experiment 1 in Corbit, Janak and Balleine, (2007). Animals were first trained with two pairs of A-O and three pairs of S-O. They were then tested for PIT with the three stimuli. The first test was conducted when the animals were hungry, whereas the second test was carried out after the animals were sated on their maintenance diet for 24 hours. Table adopted from Corbit et al. (2007).

One interesting but puzzling issue in the latter study remains in the inability of S1 and S2 to trigger general PIT. That is, although it was predicted that S1 would elevate A1 above A2, it could have been expected that it would also increase A2 above baseline, just like S3 did. This was, however, not the case. To account for this finding, the authors suggested that some form of inhibitory processes might also contribute to specific PIT. Specifically, a stimulus predicting a particular outcome elevates performance on an action with which it shares a common outcome while somehow inhibiting performance on actions earning a different outcome. Although clearly speculative, such a suggestion is particularly interesting when considering the effects that inhibitory stimuli may exert on choice between actions. Accordingly, the next section will summarize current understanding of the effects that inhibitory stimuli exert on action selection.

Choice guided by inhibitory cues

Very few studies have investigated the influence that inhibitory stimuli – i.e., those that negatively predict the occurrence of a particular outcome – exert on choice between actions. The first convincing study was that conducted in rats by Delamater, SoSa and LoLordo (2003). In this study, the authors generated inhibitory stimuli through backward conditioning, which involves the

repeated delivery of an outcome just before a particular stimulus (i.e., O then S). Under such condition, past research has shown that the stimulus gains some inhibitory properties. Thus, Delamater et al. trained two stimuli in a backward manner with two distinct outcomes (O1-S1 and O2-S2) while also arranging that these two outcomes could be earned by performing two distinct actions (A1-O1 and A2-O2). The authors then gave the rats a PIT test in order to assess choice between the two actions in the absence or the presence of either stimulus. The results showed that the backward stimuli reduced performance on the action that delivered the outcome with which the stimuli had been associated. Thus, S1 lowered A1 below baseline but left A2 relatively unaffected and similar to baseline. In contrast, S2 reduced A2 below baseline while leaving performance on A1 similar to that of baseline. Importantly, the authors also used a conventional inhibitory test to determine whether their stimuli were predicting the absence of their associated outcome. This test known as retardation refers to the observation that an inhibitory stimulus takes longer to acquire conditioned responding than a neutral stimulus. The two backwardly trained stimuli did display such a property; S1 was a potent inhibitor of O1 while S2 was an inhibitor of O2. This study therefore constituted the first evidence that inhibitory stimuli, just like excitatory ones, could guide action selection. Specifically, they selectively suppress the performance of actions earning the outcome that they predict will not be presented. Such inhibitory association was further verified by Laurent and Balleine (2015, Experiment 3), and interestingly, they showed that the temporal proximity between the backward stimulus and the outcome is critical to create the inhibitory association, as a short delay (10 seconds) is required or otherwise an excitatory one is formed instead.

Using a similar approach in mice, Laurent, Wong and Balleine (2015) recently examined the influence of forwardly and backwardly trained stimuli on action selection (**Figure 1A**). Thus, one

group of mice (group Forward) learned that two stimuli predicted the delivery of two distinct outcomes (i.e., S then O; S1-O1 and S2-O2). In another group of mice (group Backward), we used backward conditioning to train S1 and S2 to predict the absence of O1 and O2 (i.e., O then S; O1-S1 and O2-S2). All mice then received instrumental conditioning during which one action earned O1 (A1-O1) while another action delivered O2 (A2-O2). The two groups of mice were subsequently administered a PIT test that assessed choice between the two actions in the absence or the presence of either stimulus. As expected, mice in the Forward group exhibited specific PIT (**Figure 1D**) as the stimuli biased choice towards the action with which they shared a common outcome and responding on that action was well above baseline. Interestingly, choice was reversed in the Backward group. That is, the stimuli guided choice away from the action delivering the outcome that they had been associated with, towards the action earning the different outcome. Importantly, responding on the latter action was higher than that of baseline. These findings are therefore distinct from that of Delamater et al. (2003), as these authors only reported the former inhibitory effect but not the latter excitatory one. Although the reason for this discrepancy remains unclear, the two experiments differed at least in one important respect. Our study employed a transfer test procedure that extinguished instrumental responding prior to testing the effect of the stimuli. This was not case in the original study from Delamater et al. (2003). Earlier experiments on the effects of excitatory stimuli on response selection have revealed the importance of instrumental extinction. In the absence of such extinction, specific PIT has often been observed as a reduction of performance below baseline on the different action (i.e., the one delivering an outcome different from the one predicted by the stimulus) and the maintenance of performance of the same action (i.e., the one delivering the outcome predicted by the stimulus) at baseline level. It is then possible that Delamater et al. (2003) would have obtained a strict reversal of the PIT

effect (i.e., an increase above baseline of performance on the different action) if they had extinguished instrumental responding. Regardless, our experiments indicated that inhibitory stimuli influence choice between actions in an opposite manner to that of excitatory stimuli.

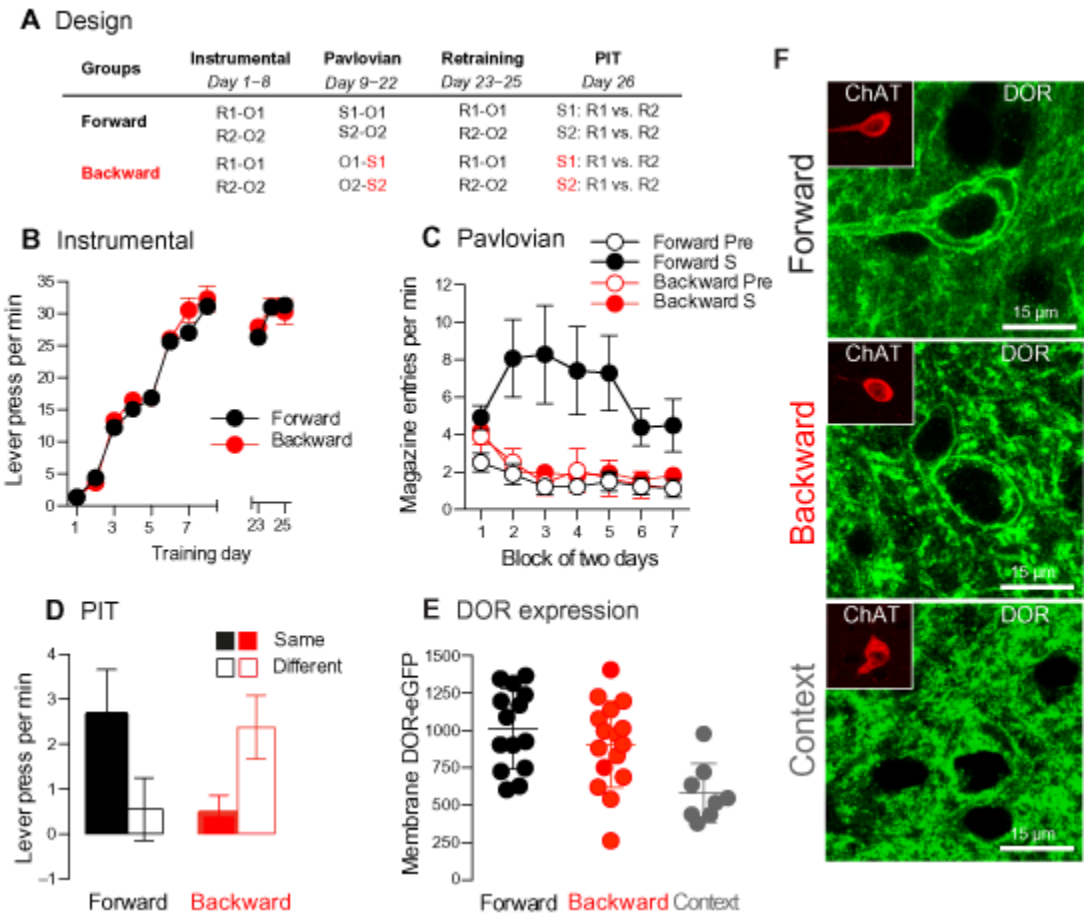


Figure 1. Inhibitory stimuli generated through backward training reverse the traditional PIT effect and produce delta-opioid receptor (DOR) accumulation on the membrane of cholinergic interneurons in the nucleus accumbens shell (NAc-S) (Laurent, Wong & Balleine. 2015). **A**) All animals initially received instrumental training of A1-O1 and A2-O2 (NB: A1 and A2 are termed R1 and R2), followed by Pavlovian conditioning. For one group (Forward), this conditioning involved learning that two stimuli (S1 and S2) predicting two distinct outcomes (O1 and O2). For another group (Backward), conditioning of the stimuli was administered in a backward manner. That is, the delivery of O1 and O2 occurred 10s before the presentation of S1 and S2, respectively. **B**) Instrumental training was successful as both groups increased their lever press rates across days.

C) Conditioning responding gradually increased across days in Group Forward, indicating successfully excitatory training. In contrast, this increase was absent in the Backward Group and the levels of responding were identical in the presence or absence of the stimuli, suggesting successful inhibitory learning. **D**) During the transfer test, the Forward group displayed specific PIT, as evidenced by the ability of the stimuli to increase performance on the action (Same) with which they shared a common outcome. In contrast, choice was reversed in the Backward group such that the stimuli biased choice away from the action (Same) delivering the same outcome as the one they predicted, towards the action (Different) earning the other outcome. E) Quantification of membrane DOR expression in NAc-S CINs revealed no difference between the forward and the backward groups, and both groups showed higher expression than a control group only exposed to the conditioning chamber. F) Confocal images showing that the DOR expression on the membrane of NAc-S CINs in mice in forward, backward training and context only group. Figure adopted from Laurent et al. (2015).

Our laboratory has recently replicated in rats the reversal of choice produced by backward stimuli using other procedures to generate specific inhibitors. These procedures included conditioned inhibition, over-expectation as well as a replication of backward conditioning (Laurent & Balleine, 2015). In all cases, it was found that a stimulus predicting the omission of a particular outcome biases choice away from an action delivering that outcome, towards an action earning a different outcome. In addition to demonstrating the strong influence that inhibitors can exert over action selection, it is essential to note that this finding has important implication for contemporary theories of instrumental conditioning. These theories assume that choice is governed by i) the relationship that an action shares with its outcome and ii) the desirability of this outcome relative to the current circumstances. Embedded in this assumption is that choice is uniquely driven by the excitatory relationships that are formed between the actions and the outcomes they procure (i.e., A-O associations). The reversal of choice produced by inhibitory stimuli constitutes therefore a challenge for such excitatory-based theories. Although they successfully predict a reduction of performance on the action procuring the outcome of which the omission is predicted by the inhibitory stimuli, they cannot explain the elevation of performance on the action delivering another outcome. To explain this latter effect, they have proposed that instrumental conditioning produces both excitatory and inhibitory relationships. That is, as subjects learn that an action earns a particular outcome (e.g., A1-O1 and A2-O2), they also learn that this action does not deliver an outcome earned by another response (e.g., A1-noO2 and A2-noO1). Under such a scheme, the authors argued that the reversal of choice produced by inhibitory stimuli could be explained. Specifically, a stimulus predicting the absence a particular outcome (i.e., S1-noO1) would evidently reduce performance on action earning that outcome (i.e., A1-O1) but it would also elevate performance on an action that has also a negative relationship with that outcome (i.e., A2-

noO1). Indeed, both the stimuli (i.e., S1) and the response (i.e., A2) lead to the same prediction (i.e., noO1). Although considerable work is required to further test such a proposal, it suggests that the effect of inhibitory stimuli on choice constitutes a general phenomenon because it is the product of the associations that are inherently established across instrumental conditioning. As such, it logically follows that the ability of excitatory and inhibitory conditioning to influence action selection should be mediated, at least partially, by common neural substrates.

The neural circuitry underlying choice guided by excitatory stimuli

Pavlovian-instrumental transfer is a powerful paradigm to study how excitatory stimuli can influence choice between actions. PIT essentially allows studying the manner in which Pavlovian and instrumental conditioning interact with each other. It is, then, not surprising that PIT has been found to require activity in brain regions underlying those two forms of conditioning. These include the orbitofrontal cortex (OFC), the amygdala, the dorsal and ventral striatum, the ventral tegmental area (VTA) and the mediodorsal thalamus (MDT). The current section will briefly summarize our understanding of the roles played by these various structures in choice between actions driven by excitatory stimuli.

The Orbitofrontal Cortex

Compelling evidence suggests that the OFC is involved in processing stimulus-outcome expectancies, possibly functioning in concert with other regions such as the basolateral complex of the amygdala (BLA, see below) (Holland & Gallagher, 2004; Picken, Saddoris, Gallagher & Holland, 2005; Balleine, Leung & Ostlund, 2011). However, the exact role played by the OFC in action-outcome processing remains highly debated, with studies revealing a critical role while others failed to do so. One possible explanation for this discrepancy may originate from the distinct subregions of the OFC that were targeted in these studies. Regardless, work in our laboratory has

found that rats with OFC lesions are able to appropriately select actions according to the value of the outcome they deliver (Balleine et al., 2011); i.e., the rat favours responding on an action delivering a valued outcome over an action earning a devalued outcome. However, these rats are unable to express specific PIT, indicating a failure to specifically use Pavlovian expectancies to guide action selection. This specific role of the OFC in S-O associations is consistent with the inability of OFC-lesioned animals to reduce conditioned responding to a stimulus predicting a devalued outcome (Picken et al., 2005).

The Amygdala

As mentioned above, the amygdala also plays an important role in S-O associations. However, this region is composed of anatomically and functionally distinct nuclei and each of them appears to process different outcome-related information. Specifically, it has been proposed that the central nucleus of the amygdala (CeA) deals with the general affective properties (i.e., aversive or appetitive) of an outcome (Balleine & Killcross, 2006) while the BLA is involved in processing the sensory specific properties of this same outcome (Cador, Robbins & Everitt, 1989; Balleine & Killcross, 2006; Parker & Balleine, 2013). Consistent with this proposal, pharmacological inactivation of the CeA abolishes general PIT but it leaves intact specific PIT. In contrast, a similar manipulation in the BLA removes specific PIT but leaves general PIT unaffected (Corbit & Balleine, 2005; Hall et al, 2001; Holland & Gallagher, 2003). Further confirmation of the role of the BLA in processing sensory specific information and S-O associations comes from a recent study (unpublished data from our laboratory) showing that the BLA lesion prevents the reduction of conditioned responding to a stimulus of which the association with its outcome has been degraded. In addition, the same lesion has been found to prevent value-based choice. That is, rats with such lesions are unable to choose an action according to the value the outcome that it procures.

Once again, these findings are all consistent with the view that the BLA processes the sensory specific properties of motivationally significant outcomes.

The Dorsal Striatum

Considerable progress has been made in determining the role played by the dorsal striatum in instrumental conditioning. This role depends on the subregion considered and the nature of the learning produced by instrumental conditioning. For instance, it is well accepted that instrumental conditioning is initially goal-directed; it relies on A-O associations as performance depends on the contingency between the action and the outcome as well as on the value attributed to that outcome. With further training, however, the performance becomes habitual and is no longer sensitive to changes in action-outcome contingency or in outcome value (Dickinson, 1985). Goal-directed and habitual behaviour have been shown to recruit the dorsomedial striatum (DMS) and the dorsolateral striatum (DLS) respectively. Accordingly, rats that have received moderate instrumental training fail to display outcome devaluation in the absence of activity in the DMS (Yin, Knowlton & Balleine, 2005; Yin, Ostlund, Knowlton & Balleine, 2005; Shiflett, Brown & Balleine, 2010). Animals that have received extensive instrumental training also fail to show such outcome devaluation under the saline infused conditions, indicating that their behaviour is habitual. However, the sensitivity to outcome value can be recovered by disrupting neural activity in the DLS, suggesting that this disruption rendered the behaviour goal-directed again (Yin, Knowlton & Balleine, 2006). Contrasting with this clear dissociation of DLS and DMS manipulations on choice based on outcome value, the roles of these two regions on choice based on the predictive stimuli present at test remains more uncertain. This is possibly due to the fact that PIT can be observed in situations where instrumental responding is either habitual or goal-directed. Regardless, both the DLS and DMS appear to be required for stimulus-based choice although their

role appears distinct. For instance, DLS inactivation leaves unaffected the ability of a stimulus to bias choice towards an action with which it shared a common outcome but it severely diminishes responding overall, bringing performance close to baseline level. In contrast, DMS inactivation spares overall responding but removes the ability of a stimulus to produce a bias (Corbit & Janak., 2007). Thus, the DMS appears particularly important to retrieve specific A-O associations in the context of the S-O associations in order for those stimuli to bias choice.

The Ventral Striatum

Similar to the dorsal striatum, the role of the ventral striatum in choice between actions depends on the subregion that is considered. Thus, the nucleus accumbens core (NAc-C) has been found to be critical for choices based on outcome value or those based on the motivational properties of stimuli present at the time of test (i.e., general PIT), as indicated by the inability of animals to successfully perform these choices when neuronal activity is locally disrupted (Shiflett & Balleine., 2010). In contrast, general PIT and value-based choices are left unaffected by disruption of neuronal activity in the nucleus accumbens shell (NAc-S). Yet, such disruption abolishes specific PIT (Corbit & Balleine., 2011).

The role of NAc-S in specific PIT is particularly interesting as, unlike other brain regions, its role appears to be restricted to expressing choice bias. For instance, rats with NAc-S lesion remain able to stop responding to a stimulus when the contingent relationship with its outcome is degraded (unpublished data from our laboratory), implying that this brain region plays little role in S-O associations. As mentioned, the inability of NAc-S manipulation to disrupt value-based choice also suggests that this brain region is not essential for processing A-O associations. This highly specific involvement of the NAc-S in expressing specific PIT has lead our laboratory to investigate the local molecular and cellular processes that were necessary for this expression. Initial

pharmacological manipulation revealed that local activation of delta-opioid receptors (DOR), but not mu-opioid receptors, was necessary for the ability of a stimulus to bias choice towards an action with it shared a common outcome (e.g., **Figure 2F**) (Laurent, Leung, Maidment & Balleine, 2012; Laurent, Bertran-Gonzalez, Cheing & Balleine 2014; Laurent et al., 2015). Further studies were then conducted using DOR-eGFP knock-in mice, which express a functional but fluorescent form of DOR, in order to determine the cellular localization of these receptors in the NAc-S. It was found that specific PIT performance is positively correlated with an increase in DOR expression on the membrane of cholinergic interneurons (CINs) within the NAc-S (e.g., **Figures 1E-F**) (Laurent et al., 2014a and Bertran-Gonzalez et al., 2013). Interestingly, however, it was the learning of the specific and contingent S-O associations that triggered DOR accumulation on NAc-S CINs, although this accumulation does not appear necessary for processing these associations. Indeed, recent work in our laboratory has shown that DOR-eGFP mice receiving Pavlovian training under a DOR antagonist treatment are as able as control mice to subsequently cease responding to a stimulus of which the outcome is no longer valued (Laurent et al., 2012). Thus, these findings indicate that Pavlovian conditioning produces an increase of DOR expression on NAc-S CINs that is subsequently recruited for expressing specific PIT.

Although identifying the critical role played by DOR on NAc-S CINs has significantly improved our understanding of the cellular mechanisms underlying specific PIT, it has only provided information about the local changes occurring in the NAc-S. Further work was therefore conducted to determine how these local changes could modulate overall NAc-S activity and its influence on output structures to promote specific PIT. As in the rest of the striatum, medium spiny neurons (MSNs) constitute the only output neurons of the NAc-S (Bertran-Gonzalez, Herve, Girault & Valjent, 2010; Gerfen & Surmeier, 2011). These GABAergic neurons can be segregated in two

distinct populations according to the dopamine receptors that they express and the structures that they target. Thus, striatonigral MSNs usually express the dopamine D1 receptors (D1R) while striatopallidal MSNs generally express the dopamine D2 receptors (D2R). Our laboratory found that D1R blockade in the NAc-S, but not local D2R blockade, removes specific PIT. Interestingly, unilateral NAc-S blockade of D1R combined with contralateral DOR blockade has the same effect, suggesting that DOR and D1R cooperate in the NAc-S to promote specific PIT (Laurent et al., 2014). In fact, this cooperation was further confirmed by showing that activation of DOR on CINs can modulate activity of D1R-containing MSNs via muscarinic M4 receptors that are uniquely expressed on this population of MSNs. Taken together, these findings suggest that delta-opioid and D1-dopaminergic processes cooperate in the NAc-S to promote the influence of excitatory stimuli on choice between actions.

The mediodorsal thalamus (MDT)

The MDT has received substantial attention as it is reciprocally connected with most of the regions identified as playing a critical role in PIT. These include the dorsal and ventral striatum, the OFC, the BLA and the VTA. This pattern of connections has led to the suggestion that the MDT could act as a hub that integrates information about A-O and S-O associations (Balleine, Morris & Leung, 2014). Initial work revealed that MDT lesions abolish value-based choice as indicated by the inability of animals to preferentially select an action of which the outcome is valued (Corbit, Muir & Balleine, 2003). However, a subsequent study contradicted this finding by obtaining the traditional decrease of performance on an action delivering a devalued outcome even though animals had received MDT lesions (Ostlund & Balleine, 2008). The origin of this discrepancy appears to originate from the timing of lesion. Thus, animals in the former study had received MDT lesions prior to Pavlovian and instrumental training whereas animals in the latter

study had been given the lesion after the two training stages. With respect to stimulus-based choice, however, the timing appears to be of less importance. Indeed, post-training lesions were found to abolish specific PIT as shown by the inability of stimuli to bias choice towards an action with which they shared a common outcome (Ostlund & Balleine, 2008; Parnaudeau et al., 2015). This finding suggests that the MDT may be particularly important to process S-O associations during specific PIT.

The Ventral Tegmental Area (VTA)

Midbrain dopamine neurons located in the VTA have received considerable attention in reward related behaviours. Indeed, many theories attribute a critical role for dopamine in such behaviours, although the exact nature of this role remains debated. For instance, some argue that dopamine is necessary to compute prediction error (Cohen et al., 2012; Glimcher., 2011; Rescorla & Wagner., 1972), to control effort allocation (Assadi, Yucel & Pantelis, 2009; Floresco, Maric & Ghods-Sharifi, 2008), to modulate motivation (Ikemoto., 2007) or to attribute incentive salience to reward-related stimuli (Berridge & Robinson., 1998). Regardless of which of these roles VTA dopamine neurons actually play, most of these theories would expect these neurons to have some involvement during stimulus-based choice. This view is reinforced by the fact that systemic administration of dopamine antagonists reduces PIT (Ostlund & Maidment, 2011) while infusion of dopamine agonist directly in the NAc-S, which receives dopamine projection from the VTA, increases PIT (Lex & Hauber., 2008). To further assess the role of the VTA, Corbit et al. (2007) performed local inactivation in rats that were then submitted to general or specific PIT. They found that VTA inactivation reduces overall performance but leaves intact the ability of a stimulus to selectively bias choice towards an action with which it shared a common outcome (i.e., specific PIT). It also left unaffected the capacity of a stimulus to elevate responding overall (i.e., general

PIT). Together, these data therefore suggest that the VTA mostly exerts an effect on performance, increasing vigour and action initiation. However, it is essential to note that the use of pharmacological inactivation to study the role of VTA midbrain dopamine neurons remains highly problematic. Indeed, the VTA is composed of other neuronal types (e.g., GABA and glutamate neurons) (Hikosaka, 2014) and there has been distinction made between the roles played by tonic and phasic release of dopamine from VTA neurons. This highlights the need for additional experiments that would evaluate more specifically the roles of VTA midbrain dopamine neurons.

The neural circuitry underlying choice guided by inhibitory stimuli

Given that the study of the influence of inhibitory stimuli on choice between actions remains in its infancy, we know very little about the underlying neural circuitry. However, and as mentioned before, the theoretical explanation of inhibition-driven choice proposed by Laurent and Balleine (2015) suggests that excitatory and inhibitory stimuli may guide action selection through common neurobiological mechanisms. Indeed, their proposal is centred on the existence of inhibitory action-outcome associations that naturally develop across instrumental training. Given this proposal, we have recently examined whether excitatory and inhibitory stimuli would bias choice between actions via similar mechanisms in the NAc-S, a region that we have seen is essential for the expression of specific PIT. Consistent with the assumption that common mechanisms are at play, we found that NAc-S blockade of DOR disrupted choice driven by both excitatory and inhibitory stimuli (**Figure 2**), although the two types of stimuli influenced choice in an opposite manner. Further, we also revealed that the two forms of choice were associated with an increase in DOR expression on NAc-S CINs (**Figure 1**). These findings are therefore consistent with the view that inhibitors and excitors guide choice via a common neural circuitry. Yet, this commonality is likely to be restricted to brain regions essential for choice expression, such as the

NAc-S. Information about the nature and the predictive properties (i.e., excitatory or inhibitory) of the stimuli being presented is likely to be provided by distinct brain structures. Although substantial evidence suggests that the BLA may provide such information about excitatory stimuli, potentially via direct projections to the NAc-S, we know little about the neural circuitry underlying learning about inhibitory stimuli. Neverthless, recent findings suggest that the lateral habenula (LHb) could be essential.

Lateral Habenula

The habenula (**Figure 3**) is a highly preserved structure throughout evolution as it can be found in lamprey, reptiles, amphibians and mammals (Herkenham & Nauta, 1977; Kemali & Guglielmotti, 1977; Engbretson, Reiner & Brecha, 1981;Stephenson-Jones, Floros, Robertson & Grillner, 2012). It can be divided into the lateral part (LHb) and the medial part (MHb) with distinct structures and functions. The MHb has been suggested to be involved in endocrine and immunological function (Silver, Silverman, Vitkovic & Lederhendler, 1996; Whilhelm, King, Silverman & Silver, 2000). The LHb can be further divided into the lateral part (LHb-l) and the medial part (LHb-m). The LHb-l receives input from the basal ganglia via the globus pallidus (GPb) (Herkeham & Nautu., 1977), and projects to VTA via the rostromedial tegmental nucleus (RMTg). When it is activated, it triggers the activation of RMTg neurons through glutamatergic connections, which can in turn inhibit dopamine neurons in the VTA via their GABAergic projections (Matsumoto & Hikosaka., 2007; Jhou et al., 2009). On the other hand, the LHb-m receives input from the limbic region (Herkeham & Nautu., 1977) and projects to the dorsal and medial raphe nuclei (DRN and MRN), which have been suggested to modulate the serotonin system in the midbrain (Herkeham & Nautu., 1979; Reisine, Soubrie, Artaud & Glowinski, 1982; Kale, Strecker, Rosengren & Bjo, 1989; Amat et al., 2001; Gonçalves, Sego & Metzger, 2012; Kim., 2009).

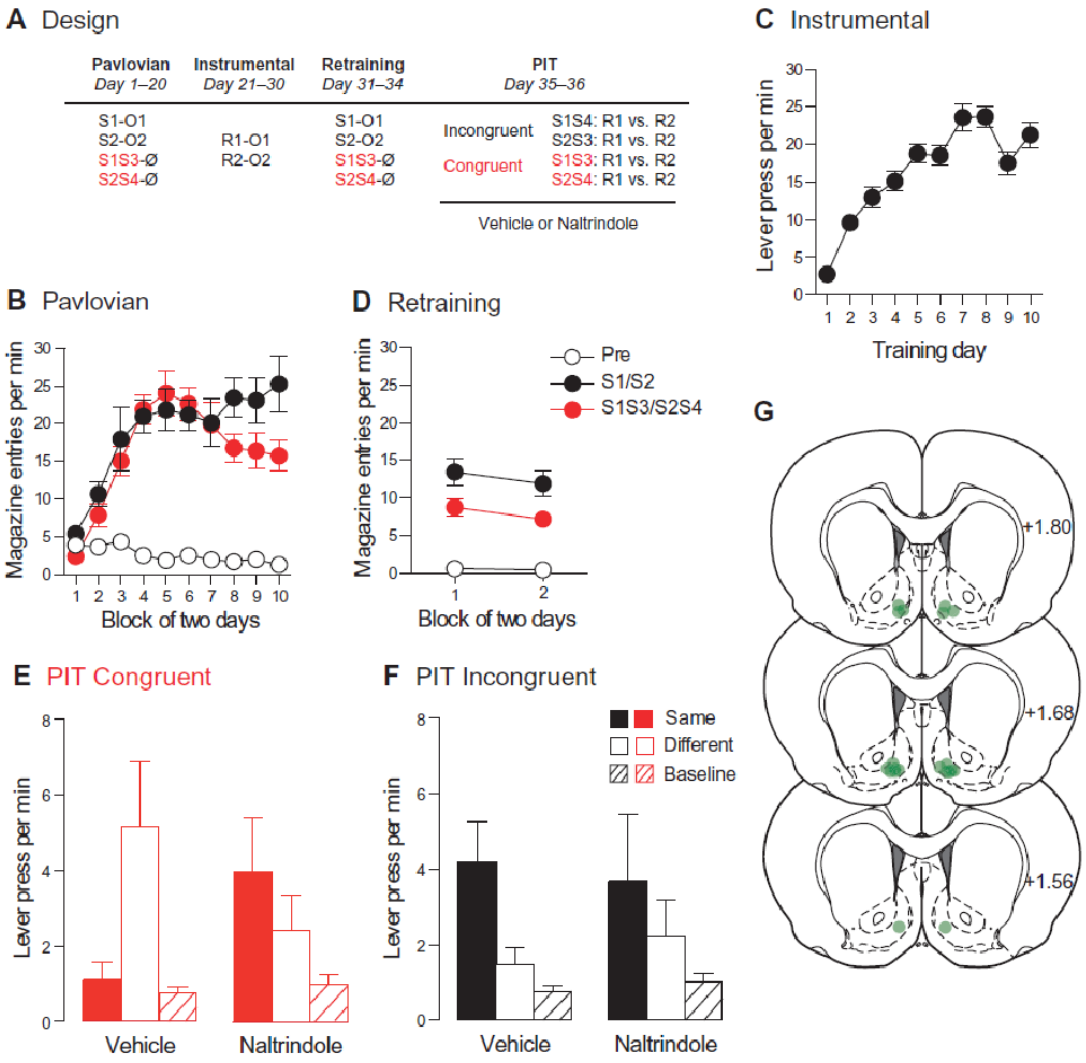


Figure 2 – The effect of excitatory and inhibitory stimuli on choice between actions is abolished by DOR blockade in the NAc-S (Laurent, Wong & Balleine, 2015). **A**) Animals initially received Pavlovian training during which S1 and S2 predicted O1 and O2, respectively. In addition, S1 and S2 were also presented in compound with two other stimuli (S3 and S4) in the absence of any outcome. The aim was to generate two conditioned inhibitors, with S3 signaling the absence of O1 and with S4 that of O2. The animals then received an instrumental training of A1-O1 and A2-O2. Following two reminder sessions, a PIT test was conducted under infusion of saline or the DOR antagonist naltrindole in the NAc-S. **B**) Pavlovian training was successful as the two excitatory stimuli (S1 and S2) elicited substantial conditioned responding. Although the compound triggered similar levels of conditioned responding at the start of training, these levels declined towards the end of the session, suggesting successful inhibitory training. **C**) Instrumental training occurred smoothly as animals increased their rate of lever presses across training. **D**) The reminder session reveals a similar pattern of performance as that observed at the end of Pavlovian training **E**) A compound composed of an excitor and an inhibitor of the same outcome produced the reversal of PIT. This reversal was abolished by DOR blockade in the NAc-S. (**F**) A compound composed of an excitor of one outcome and an inhibitor of another outcome produced specific PIT unless DOR were antagonized in the NAc-S. (G) Placement of the injection cannula tips in the NAc-S.

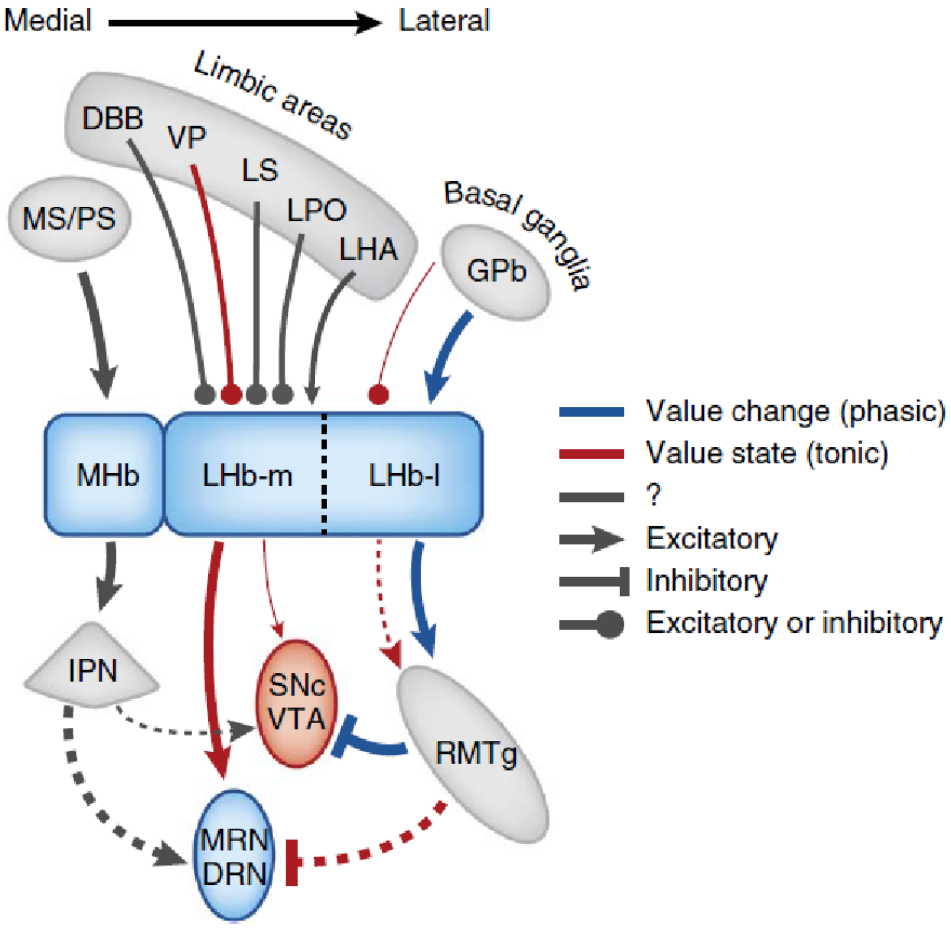
Distances on the atlas templates are indicated in millimetres relative to bregma. Figure adopted from Laurent et al. (2015).

One of the critical anatomical and physiological properties of the LHb is its ability to inhibit VTA dopamine through the RMTg. As mentioned before, these dopamine neurons have been suggested to compute reward prediction error (RPE) in the brain. RPE refers to the discrepancy between the expected and actual outcome of a particular rewarding event. Current theories hold that the larger this discrepancy, the more learning occurs and conversely, a small discrepancy will drive little learning (Rescorla & Wagner., 1972; Glimcher., 2011). Consistent with the view that midbrain dopamine neurons act as a neural correlate of reward prediction error, their activity has been shown to peak when an unexpected rewarding outcome is encountered (Cohen et al., 2012). This positive reward prediction error is thought to strengthen the associations between that rewarding outcome and the antecedent environmental events (e.g., a stimulus during Pavlovian training). It logically follows that a negative reward prediction error is produced when an expected rewarding outcome fails to occur, resulting in the development of an inhibitory association between the rewarding outcome and the antecedent environmental events. At the neural level, this negative prediction error is accompanied with a drop in activity of midbrain dopamine neurons. Importantly, substantial evidence suggests that this drop could be driven by activation of the LHb-RMTg-VTA pathway.

Electrophysiological studies (Matsumoto & Hikosaka, 2007; 2009) have found that the majority of LHb neurons are excited when animals are presented with a stimulus positively predicting an aversive outcome (i.e., the opposite of rewarding) or a stimulus negatively predicting a rewarding/appetitive outcome (i.e., an inhibitory stimulus). Importantly, the role of the LHb is not restricted to Pavlovian tasks. Indeed, local lesions impair the extinction of a previously rewarded response, consistent with its proposed role as a source of negative RPE (Friedman et al., 2010; 2011). In line with the LHb ability to control activity of midbrain dopamine neurons, local lesion

has been found to increase dopamine metabolism and dopaminergic neuronal activity in the VTA (Lisoprawski et al., 1980; Nishikawa, Fage & Scatton, 1986). Recently, several studies have employed optogenetics to determine the exact role of the LHb. It was found that activating LHb inputs (i.e., entopeduncular nucleus) or outputs (i.e., the RMTg), result in conditioned place avoidance (Shabel et al., 2012; Stamatakis & Stuber., 2012), again consistent with a role of this structure in negative RPE. Interestingly, communication between the LHb and the VTA does not appear to be unidirectional. For instance, Stamatakis et al. (2013) revealed that a unique group of VTA neurons project back to LHB and that their activation promotes GABA release in the LHb, suggesting the existence of a feedback mechanism by which VTA dopaminergic functioning to promote RPE through inhibition of LHb and RMTg activities.

Beyond its role in reward processing, Floresco and colleagues (Stopper & Floresco., 2013; Stopper et al., 2014) have attempted to investigate the role of the LHb in decision-making processes such as those involved in choice between large/risky and small/certain rewards using pharmacological inactivation and electrostimulation of LHb. They found that LHb inactivation impairs decision-making when facing uncertain reward, i.e. the animals were unable to choose the reward with higher expected value despite the fact that they were able to discriminate reward size. Moreover, the authors also examined the temporal relation between LHb activation and choice behaviour against uncertain outcomes using electro-stimulation, and revealed that while stimulation immediately before choice would lead to a shift away from the choice with higher expected value, stimulation at the time of reward delivery would result in a shift to the alternative choice on the next trial. Importantly, the effect of stimulation did not appear when the outcomes of both options were guaranteed. Thus, the authors concluded that phasic dopamine activity provides guidance in the incoming choice and feedback about the recent choice, which is consistent



Figue 3. The proposed connections around LHb. The LHb-m receives input from the limbic system and modulates the serotonin system by its projection to MRN and DRN. The LHb-l receives input from basal ganglia and modulates the dopamine system mainly by its projection to RMTg. Projections in dashed line mean that it requires more studies to be confirmed, the thickness of the projections refers to the presumed strength of connection. DBB, diagonal band of Broca; LHA, lateral hypothalamic area; LPO, lateral preoptic area; LS, lateral septum; MS, medial septum; PS, posterior septum. Figure adopted from Proulx, Hikosaka & Malinow. (2014).

with the proposed role of dopamine in RPE theory, and that the LHb could be the preference center to integrate the cost and gain when dealing with ambiguous situations by modulating the phasic dopamine signal downstream.

In summary, the main finding here is that the LHb responds to stimuli that predict the absence of an appetitive outcome. In the context of instrumental conditioning, the LHb seems to provide some feedback information with respect to the availability of the outcome. Further, the LHb is likely to be involved in biasing choice according to their cost/benefit ratio. Although the value-based decision making process does not completely share the same neural network with the cue guided decision making process, the LHb is still in a good position to process information about cues predicting the absence of a particular outcome and to influence specific PIT. The goal of the current thesis was therefore to study the role of the LHb in cue guided decision-making, especially when the cues are predicting the omission of an appetitive outcome. Moreover, to the best of our knowledge, the relationship between LHb and PIT has never been studied. Therefore, these studies can increase our understanding in how the LHb interacts the current neural framework of PIT.

Chapter 2: General method and overview of the experiments

Three experiments investigated the role played by the lateral habenula (LHb) in various inhibitory processes. In the first experiment, this role was studied by assessing the effects of electrolytic lesions of the LHb on the development of conditioned inhibition using a procedure developed by Rescorla (2002). The next two experiments explored the impact of the same lesion on specific Pavlovian-Instrumental transfer. In one experiment, the Pavlovian cues were trained as negative predictors of the trained outcomes whereas in the other experiment the cues were trained as positive predictors of these same outcomes.

Subjects

59 Naive male hooded wistar rats were obtained from the Laboratory Animal Services (University of Sydney, Australia). They were housed in plastic boxes (two to four per box) located in a climate-controlled colony room and were maintained on 12hr light/dark cycle (light from 7am). Rats were at least 12 weeks old at the start of the experiments and weighted around 390g. Three days before behavioural training, all rats were handled daily and were put on a food deprivation schedule to maintain them at ~85% of their ad libitum feeding weight. The Animal Ethics Committee at the University of Sydney approved all experimental procedures.

Apparatus

Training and testing took place in 16 Med Associates (St. Albans, VT, USA) operant chambers enclosed in sound- and light-resistant shells. Each operant chamber was equipped with a pump fitted with a syringe that could deliver 0.1 ml of a 20% sucrose solution into a recessed magazine. Each chamber was also equipped with a pellet dispenser that could deliver grain food pellets (45mg;

BioServe Biotechnologies). The chambers contained two retractable levers that could be inserted to the left and right side of the magazine. An infrared photobeam crossed the magazine opening, allowing for the detection of head entries. A 3W, 24 V house light provided illumination of the operant chamber, and each chamber contained a Sonalert that, when activated, delivered a 3 kHz pure tone, a 28 V DC mechanical relay that was used to deliver a 2Hz clicker stimulus, and a white noise generator (80 dB). A set of two microcomputers running proprietary software (Med-PC; MED Associates) controlled all experimental events and recorded magazine entries and lever presses. Outcome devaluation was conducted in a separate room that contained 16 distinct plastic boxes.

Surgery

Rats were anaesthetized with continuous flow of a mixed isoflurane and oxygen gas (5% induction; 2-3% maintenance) and were placed in a stereotaxic frame (Kopf, Tujunga, CA, USA). An incision was made on the scalp cleaned with povidone-iodine (Betadine; Virginia, QLD, Australia) to expose the skull. Bregma and lambda were aligned on the same horizontal plane by adjusting the incisor bar and holes were drilled bilaterally above the LHb at the following coordinates: AP: -3.4, ML: +/- 1.0, DV: -5.7 (all coordinates are indicated in millimetres relative to bregma, Paxinos and Watson, 1998). A lesion was induced by passing a current at 7-10 Volt for 20 seconds with an LM4 lesion maker (Grass Instruments, Quincy, Mass) using an insulated electrode that was bared 1mm from the tip. The same procedures were applied to rats in the sham group except that no current was passed. The incision was closed using wound closure clips (EZ Clip; Stoelting, Wood Dale, IL, USA). All rats were given a 0.4 ml intraperitoneal injection of procaine penicillin solution (300 mg/kg, Ilium Benicillin, Glendenning, NSW, Australia) after the surgery. Animals were allowed to recover for 7 days before the behavioural procedures.

Histology

After all the behavioural procedures, all rats were deeply anesthetized with injection of sodium pentobarbital and perfused through the heart with ice-cold 4% paraformaldehyde in 0.1 M phosphate buffer. Brains were extracted and postfixed in 4% paraformaldehyde overnight. 40 μm free-floating coronal sections were cut through the LHb by using a vibratome (Leica) and were then stained with cresyl violet. The area of lesion was determined under a microscope by a trained observer, who was unaware of the treatment groups, with boundaries defined by the atlas of Paxinos and Watson (1998). Animals with inaccurate or extensive damage at the lesion site were excluded from the statistical analysis. **Figure 4** shows a schematic reconstructions of the LHb lesion region in the animals included in the analysis of the current studies.

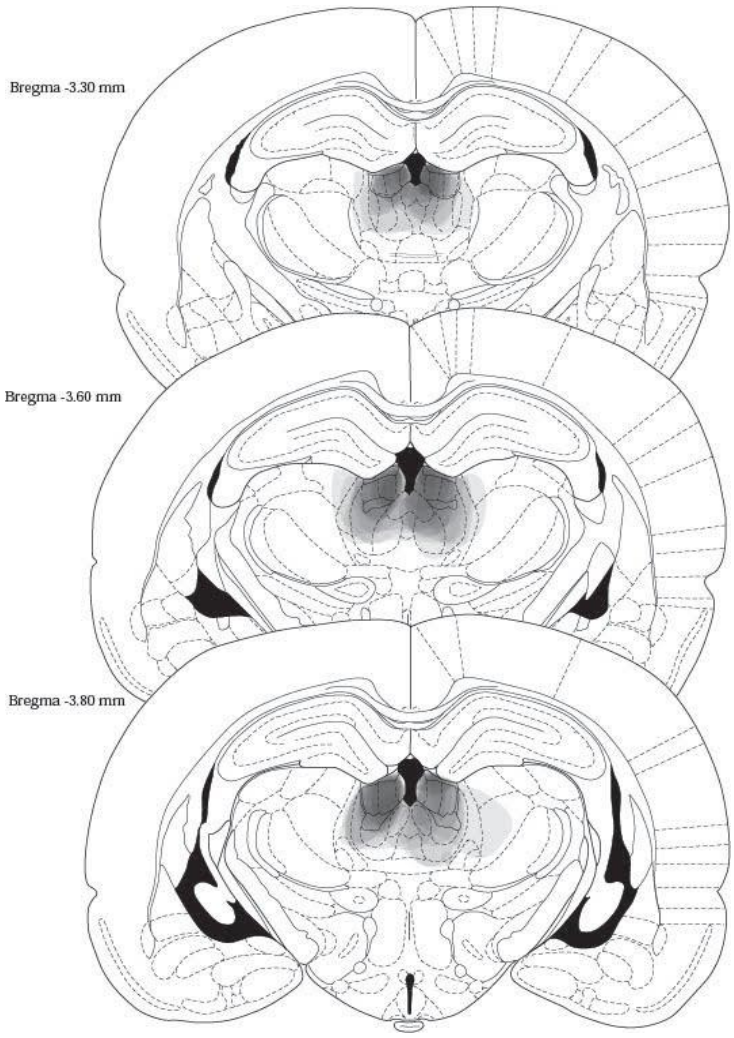


Figure 4 – Schematic reconstructions of the LHb lesion in coronal brain sections, with the lesion region of each subject represented as a separate, stacked layer.

Chapter 3: Role of LHb in conditioned inhibition

Introduction

LHb neurons have been shown to increase their activity across presentations of a stimulus that signals the absence of an appetitive outcome (Matsumoto & Hikosaka., 2007). This finding, combined with others, has led to the suggestion that the LHb acts as a neural correlate of negative reward prediction error in the brain. Specifically, it has been proposed that the LHb exerts its inhibitory effects by shutting-down activity of midbrain dopamine neurons in the VTA, neurons that have been argued to compute positive reward prediction error (Matsumoto & Hikosaka., 2007; Bromberg-Martin, Matsumoto & Hikosaka, 2010; Hong et al., 2011;Tian & Uchida., 2015). Based on these findings, the aim of the present experiment was to examine the role of the LHb in Pavlovian inhibitory learning. If the LHb computes negative reward prediction error, it logically follows that local lesion should impair the ability of a stimulus to gain inhibitory properties.

In the present experiment, rats received sham or LHb lesions and were then exposed to a behavioural paradigm developed by Rescorla (2002) (see **Figure 5**). In this paradigm, rats initially learn that one stimulus (i.e., S1) predicts a particular food outcome (i.e., O) while two other stimuli (i.e., S2 and S3) signaled its absence. This is done through a conditioned inhibition procedure during which S2 and S3 are individually presented in compound with S1 in the absence of any outcome (i.e., S1S2-nothing and S1S3-nothing). In addition, the rats are also exposed to two other stimuli (i.e., S4 and S5) in the absence of any consequence. The inhibitory properties imbued to the conditioned inhibitors are firstly assessed through a retardation test. That is, one of the conditioned inhibitors (e.g., S2) and one of the latent inhibitors (e.g., S4) are now trained to predict the occurrence of the outcome O. The rate at which this inhibitor acquires conditioned responding

is compared to that of one of the pre-exposed stimulus (i.e., S4). Given that this latter stimulus is not specifically trained to predict the absence of the outcome, it is expected to develop conditioned responding faster than the conditioned inhibitor (i.e., S4>S2). In other word, acquisition of conditioned responding to S2 will be retarded. Subsequently, the inhibitory properties of the conditioned inhibitors is again tested using a so-called subtractive summation test. This test refers to the observation that the level of conditioned responding elicited by an excitatory stimulus (e.g., S1) is reduced when this stimulus is presented in compound with an inhibitory stimulus. Thus, rats are presented with two compounds: S2S5 and S4S3. In each compound, one of the stimuli has been recently trained as a positive predictor of the outcome (i.e., S2 and S4) whereas the other stimulus is either a conditioned inhibitor of the outcome (i.e., S3) or a pre-exposed stimulus (S5). Evidence for the inhibitory properties carried by the conditioned inhibitor S3 should be revealed by lower levels of conditioned responding elicited by S4S3 as compared to S2S5. Although somewhat complex, this paradigm presents the critical advantage to compare the rate of inhibitory strengths carried by stimuli of similar familiarity: S2 vs. S4 and S3 vs. S5. It should be noted that S4 and S5 might well carry some inhibitory properties, as indicated by the fact that such pre-exposed stimuli are commonly referred to as latent inhibitors. However, latent inhibitors have typically failed to display net inhibitory properties such as those shown by conditioned inhibitors. Further, unlike conditioned inhibition, latent inhibition does not require negative reward prediction error to develop (i.e., the outcome is not expected during pre-exposure and therefore no reward prediction error occurs). Thus, the critical question here remains in the effects of lesioning the LHb. If this region is essential to Pavlovian inhibitory learning and negative reward prediction errors, animals in the lesion group should show no difference in conditioning responding to the various

stimuli/compounds during the retardation/subtractive summation test. That is, the behavioural procedure would not have turned S2 and S3 into conditioned inhibitors of the outcome.

Material and methods

*Behavioural procedures (refer **Figure 5**)*

Magazine training - Rats with either sham (n = 10) or LHb (n = 10) lesion were given a single 20-min session of magazine training during which 20 pellets were delivered using a random time 60 s schedule.

Pavlovian training stage 1 & 2 - In stage 1, all rats were given 16 daily sessions of Pavlovian training during which a tone stimulus (S1) was trained to predict a food outcome (O, grain pellets). The house light and the flashing light (i.e., house light flashing) or the noise and clicker were trained in parallel as conditioned inhibitors (S2 and S3). Each session consisted of 16 presentations of the tone, which lasted for 30 s and was followed by a pellet delivery at the end of presentation. There were also eight presentations of each compound stimulus (S1S2 and S1S3) that lasted for 30 s but no reward was delivered. Within each group, half of the animals were trained with the light and flashing light as the conditioned inhibitors while the remaining half was trained with the noise and the clicker as the conditioned inhibitors. The order of stimulus presentations was pseudorandom and the intertrial interval (ITI) varied from 90 s to 210 s with an average of 150 s. After the 16 sessions, two control stimuli, or latent inhibitors (S4 and S5), were introduced in stage 2. These stimuli were clicker and noise for the rats that had received compound presentation with the light and flashing light, while they were light and flashing light for the other rats. Each session consisted of 12 presentations of tone with reinforcement and six presentations for each of the remaining stimuli/compounds (S1S2, S1S3, S4 and S5) without reinforcement.

Pavlovian training stage 3 - All animals were then given eight daily sessions of Pavlovian training. For each animal, one of the conditioned inhibitors (S2) and one of the latent inhibitors (S4) were trained in an excitatory manner. That is, these stimuli were now followed by the delivery of the food pellets. The identity of S2 and S4 was counterbalanced within group. Thus, each session consisted of six 30-s presentations of each stimulus followed by a pellet delivery using a variable ITI that averaged to 150 s. Throughout all the sessions, magazine entries were recorded and separated into S2 and S4 period, as well as, a prior S2/S4 period with equal length (Pre/S2/S4; 30 s).

Test - On the next day, all animals received a half-session of stage 2 followed by two non-reinforced presentations of each compound stimuli (S2S4 and S3S5). Each compound consisted of one conditioned inhibitor and one latent inhibitor, yet only one of them in each pair was reinforced in phase 2 training, such that one compound was a conditioned inhibitor together with a reinforced latent inhibitor (S3S4), another compound was a reinforced conditioned inhibitor together with a latent inhibitor (S2S5). Magazine entries were recorded for each compound presentation, as well as for a period of equal length (Pre; 30 s) prior to the compound presentation.

Stage 1 (16 sessions)	Stage 2 (6 sessions)	Stage 3 (8 sessions)	Test
S1-O S1S2- S1S3-	S1-O S1S2- S1S3- S4- S5-	S2-O S4-O	S2S5 vs S3S4

Figure 5. The design and training schedule of Experiment 1

Results and discussion

Two animals from the lesion group were excluded from the analysis, as the lesion extended beyond the LHb boundaries. The data for the first stage are presented in **Figure 6** and are depicted as the number of magazine entries in the absence or presence of the stimuli/compounds. This first training stage was clearly successful as all animals displayed a gradual increase in magazine entries during the excitatory stimulus S1 while showing a lower number of such entries during the compounds S1S2 and S1S3. An ANOVA was conducted using Group (i.e., Sham vs. Lesion), Period (Pre vs. S1 vs. S1S2/S1S3) and Session as main factors. It revealed a main effect of Period ($F_{(2, 32)} = 43.563$, $p < 0.001$) and Session ($F_{(15, 240)} = 8.469$, $p < 0.001$) but no main effect of Group ($F_{(1, 16)} = 3.402$, $p = 0.084$). However, significant Period x Group ($F_{(2, 32)} = 5.081$, $p = 0.012$), Session x Group ($F_{(15, 240)} = 2.019$, $p = 0.015$) and Period x Group x Session ($F_{(30, 480)} = 2.065$, $p = 0.001$) interactions indicated that the lesion group was slower to acquire conditioned responding to S1. Yet, a significant Period x Session interaction ($F_{(30, 480)} = 22.915$, $p < 0.001$) revealed that all groups increased their level of magazine entries to S1 across training. Separate analysis within each group confirmed that the magazine entries rate during S1 was higher than the one in the Pre period (For the Sham group: $T_{(18)} = 7.091$, $p < 0.01$; For the Lesion group: $T_{(14)} = 5.411$, $p < 0.01$), and the one in the S1S2/S1S3 period (For the Sham group: $T_{(18)} = 6.702$, $p < 0.01$; For the Lesion group: $T_{(14)} = 4.900$, $p < 0.01$).

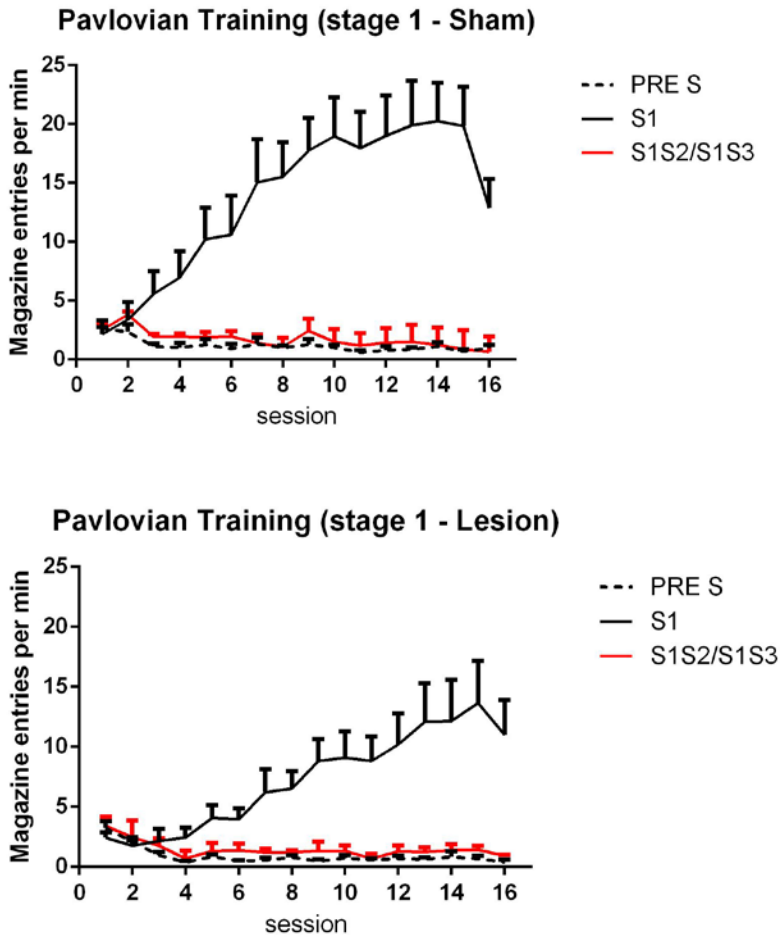


Figure 6. The rate of magazine entry for both groups across sessions in stage 1. Error bars denote s.e.m.

Performance during the second stage is presented in **Figure 7** in the same manner as before. The only difference is the inclusion of performance during the two latent inhibitors S4 and S5. This training stage was also successful as the levels of magazine entries during the excitatory stimulus S1 were significantly higher than those during the compounds, the latent inhibitors and in the absence of any stimulus. An ANOVA was conducted using the same main factors as before and it revealed a main effect of Period ($F_{(3, 48)} = 41.845$, $p < 0.001$) but no main effect of Session ($F_{(5, 80)} = 1.003$, $p > 0.1$) or Group($F_{(1, 16)} = 0.478$, $p > 0.1$). There was also a Period x Session interaction ($F_{(15, 240)} = 5.721$, $p < 0.001$), indicating that the difference between the levels of magazine entries during S1 and the other periods grew larger across training. Separate analysis within each group confirmed the Period x Session effect in both groups (For the Sham group: $F_{(15,135)} = 5.916$, $p < 0.01$; For the Lesion group: $F_{(15,105)} = 2.012$, $p = 0.0209$), and also indicated that there was no difference between the magazine entries rate during S1S2/S1S3 and the one during S4/S5 in both groups (Ts < 0.5).

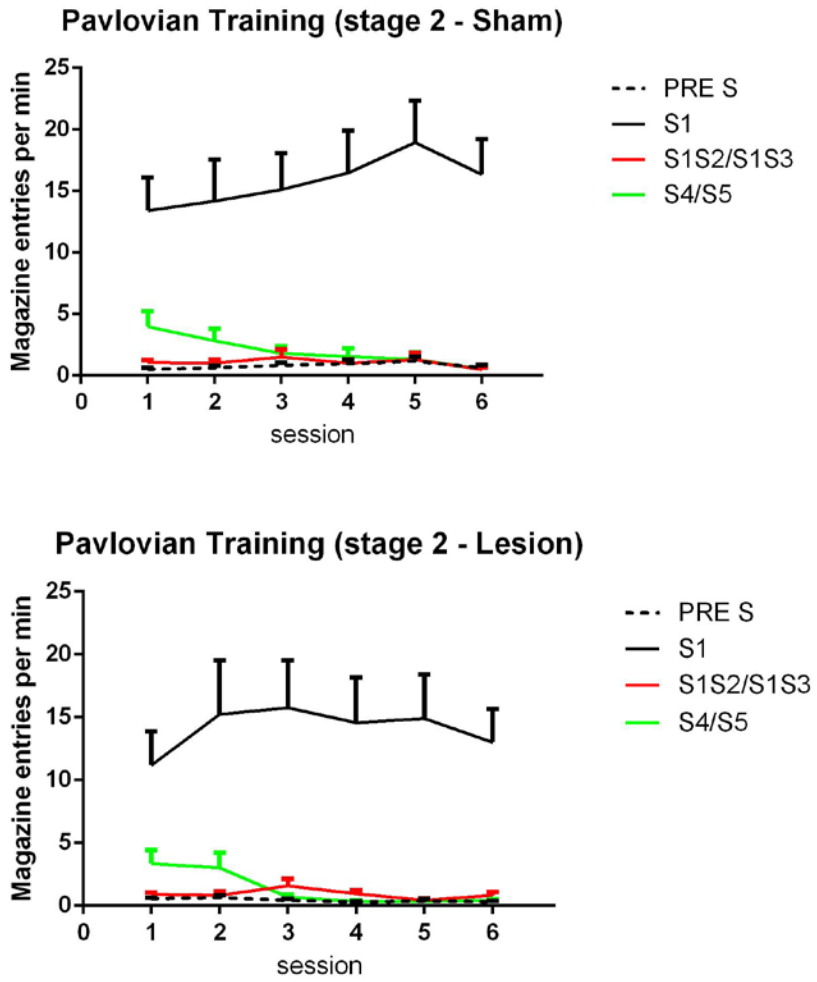


Figure 7. The rate of magazine entry for both groups across sessions in stage 2. Error bar denotes s.e.m.

The data from the retardation test administered after the two training stages are shown in **Figure 8** in the manner described above. Unfortunately, potential significant differences were overshadowed by large individual differences in performance during presentation of the conditioned inhibitor within each group (see Figure A1 in the appendices). To control for this individual difference, we standardize the levels of magazine entries into an elevation ratio (ER) that controls for the level of entries in the absence of the stimuli (e.g., ER_{S2}= S2/preS2+S2 and ER_{S4}=S4/preS4+S4). The corresponding data are plotted in **Figure 9** and ANOVA was conducted using Group, Period (S2 vs. S4) and Session as main factors. This analysis failed to reveal any main effect of Group ($F < 0.001$) or Period ($F_{(1, 16)} = 3.287$, $p = 0.089$), but a significant main effect of Session ($F_{(7, 112)} = 46.270$, $p < 0.001$) and Period x Session interaction ($F_{(7, 112)} = 3.039$, $p = 0.006$). However, no interaction with Group was found (Group x Session: $F_{(7, 112)} = 1.411$, $p = 0.208$; Group x Period x Session: $F_{(7, 112)} = 0.702$, $p = 0.670$). We conducted separate analysis using the main factor Period and Session within each group. In the Sham group, we found a main effect of Session ($F_{(7, 63)} = 23.606$, $p < 0.001$), no main effect of Period ($F_{(1, 9)} = 2.339$, $p > 0.1$) but a significant Session x Period interaction ($F_{(7, 63)} = 2.913$, $p = 0.011$). Subsequent pairwise comparisons with Bonferroni correction indicated higher levels of magazine entries in the presence of the latent inhibitor S4 than in the presence of the conditioned inhibitor (S2) during the first ($T_{(9)} = 3.685$, $p < 0.01$) and second session ($T_{(9)} = 4.118$, $p < 0.01$) of the retardation test. Although weak and transient, this effect suggested that the acquisition of the conditioned inhibitor S2 was retarded compared to that of the latent inhibitor S4. In contrast, a similar analysis in the Lesion group showed a main effect of session ($F_{(7, 49)} = 27.477$, $p < 0.001$), no main effect of Period ($F_{(1, 7)} = 1.152$, $p > 0.1$) and no Period x Session interaction ($F_{(7, 49)} = 0.960$, $p > 0.05$). Thus, our retardation test provided some evidence that the conditioned inhibition procedure employed in the

present experiment imbued S2 with some inhibitory properties that were removed by lesion of the LHb.

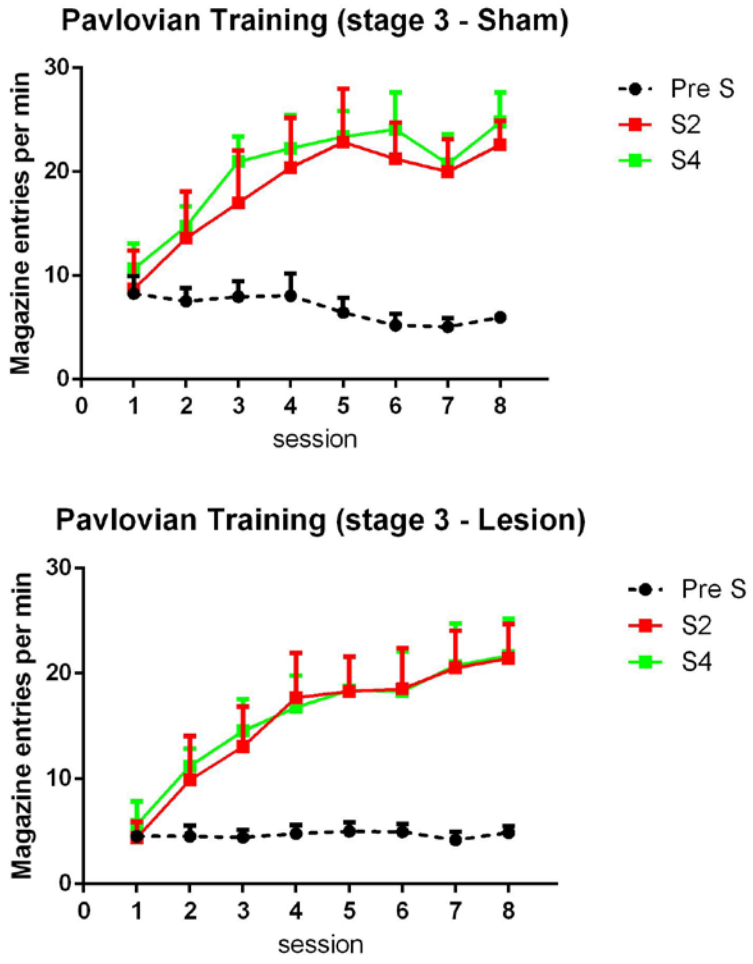
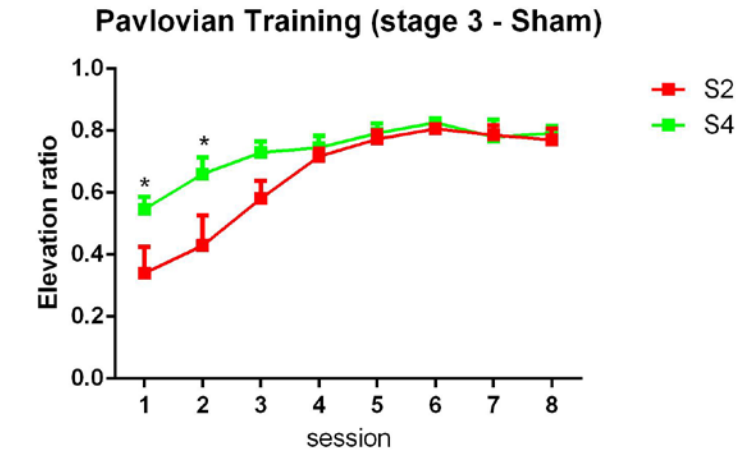


Figure 8. The rate of magazine entry of both groups across sessions in stage 3. Error bar denotes s.e.m.



.

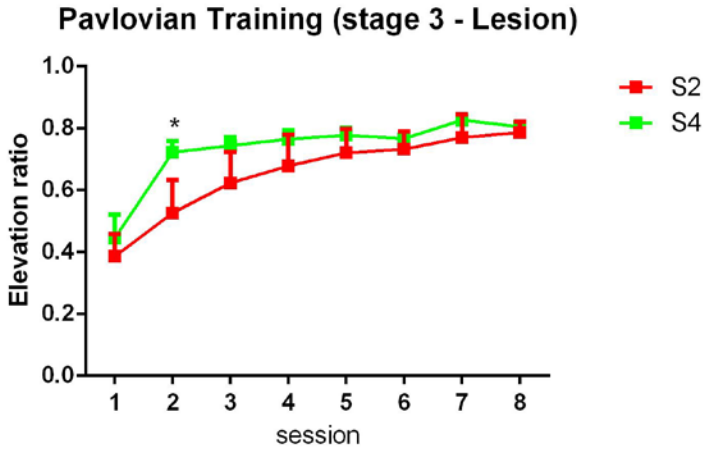


Figure 9. The elevation ratio of both groups across sessions in stage 3. Error bar denotes s.e.m. and * indicates $p < 0.05$

The results from the summation test are shown in **Figure 10**. There was no main effect of group, but an effect of Period ($F_{(2, 32)}$ = 59.65, p < 0.001) was found in preliminary analysis. Separate ANOVA on entries rate with Period (pre-stimuli vs S2S5 vs S3S4) as main factors was performed in each group. The effect of Period was statistically significant ($F_{(2, 18)}$ = 39.404, p < 0.001) in the sham group, subsequent pairwise comparison with Bonferroni correction revealed statistical significant difference between the pre-stimulus period and S2S5 period ($T_{(9)}$=7.664, p < 0.01), and between the pre-stimulus period and S3S4 period ($T_{(9)}$=7.192, p < 0.01), but not between S2S5 and S3S4 ($T_{(9)}$=0.472, p > 0.1). Similarly, the effect of period was statistically significant ($F_{(2, 14)}$ = 22.810, p < 0.001) in the lesion group and subsequent pairwise comparison with Bonferroni correction revealed a statistically significant difference between the pre-stimulus period and S2S5 period ($T_{(7)}$=6.720, p < 0.001), and between the pre-stimulus period and S3S4 period ($T_{(7)}$=5.135, p < 0.01), but not between S2S5 and S3S4 ($T_{(7)}$=1.585, p > 0.1).

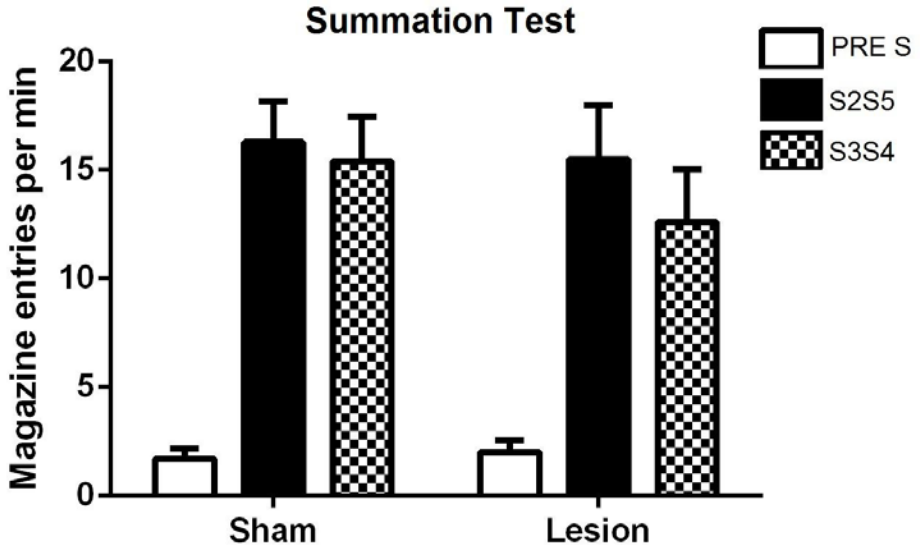


Figure 10. The rate of magazine entry of both groups in the summation test. Error bar denotes s.e.m.

The result of the present experiment failed to replicate the results obtained by Rescorla (2002). In his study, Rescorla initially observed that the conditioned inhibitor, S2, took longer to acquire conditioning responding during the retardation test relative to the latent inhibitor S4. This was consistent with the view that the conditioned inhibitor S2 reliably predicted the absence of the outcome. Rescorla then obtained a clear subtractive summation effect during which the compound composed of a conditioned inhibitor and a newly trained latent inhibitor (S3S4) elicited lower levels of conditioned responding than a compound composed of a newly trained conditioned inhibitor and a latent inhibitor (S2S5). Once again, this was taken as evidence that the conditioned inhibitor (S3) was carrying substantial inhibitory properties relative to the latent inhibitor (S5). The present experiment revealed some weak evidence of retardation in the Sham group across the first few trials that consisted in pairing the conditioned inhibitor (S2) with the outcome that it predicted would be omitted. However, this experiment clearly failed to obtain subtractive summation. Why we were unable to replicate Rescorla's finding remains unclear. This is particularly puzzling given that similar parameters (e.g., stimulus identity, amount of training) were used except for the strain of rats employed. Future experiments are therefore required to confirm whether the findings obtained by Rescorla can be replicated in another laboratory.

Given the issues just reported, it remains difficult to conclude with any certainty whether the LHb is necessary for the development of Pavlovian inhibitory learning. Yet, it is essential to note that the subtle retardation effect observed in the Sham group was absent in the Lesion group. This indicates the need for further experiments that would aim at further determining the role of the LHb in conditioned inhibition. For example, these experiments could use more traditional procedures to establish the conditioned inhibitors and involve comparisons with other stimuli than latent inhibitors. However, this is this latter aspect of Rescorla's design that motivated the present

experiment. Indeed, the use of latent inhibitors ensures that stimuli of similar familiarity are being compared. Further, latent inhibition is particularly interesting in the context of exploring the role played by the LHb in negative reward prediction error. For instance, it is commonly believed that latent inhibition can develop in the absence of reward prediction error as pre-exposure to a neutral stimulus does not break any expectancy of a particular outcome. In contrast, conditioned inhibition develops because the outcome predicted by the excitatory stimulus that is presented in compound with the conditioned inhibition does not occur. It logically follows that LHb lesion should impair conditioned inhibition but not latent inhibition.

Apart from the main finding, a trend of lower magazine entries rate in the lesion group was observed during the Pavlovian training. The trend also appeared in the Pavlovian training in the later experiment, yet there was no report of such trend in previous studies. It was unlikely due to the impairment of locomotion as the later experiment showed a similar level of instrumental response rate between the sham and the lesion group. It could be that the extensive lesion caused an impairment in reacting toward or attention to the stimuli. However, it is also possible that the lesion of LHb disrupted the processing of the incentive salience of the stimuli, which has been previously suggested by Danna, Shepard and Elmer (2013).

In summary, the present experiment only partially replicated Rescorla's finding (2002). That is, we did observe a subtle retardation effect with a conditioned inhibitor relative to a latent inhibitor. This effect was removed by lesion LHb, suggesting that this region may be involved in computing negative RPE. To confirm further this involvement, the next chapter aimed to determine the effects of LHb lesion on the ability of an inhibitory stimulus to drive choice between actions.

Chapter 4: The role of LHb on choice guided by inhibitory stimuli

Introduction

The aim of the second experiment was to examine the role of the LHb on the ability of inhibitory stimuli to drive choice between actions. Just as in the previous experiment, rats received either sham or electrolytic lesion in the LHb before being exposed to the behavioural procedure (see **Figure 11**). This procedure was identical to that developed by Laurent et al. (2015) and used backward conditioning to generate two stimuli (S1and S2) that were specifically predicting the absence of two distinct food outcomes (O1 and O2). Then, rats were administered instrumental training during which the two outcomes could now be earned by performing two distinct actions (A1 and A2). After this training, we assessed the influence of the two inhibitory stimuli on choice between actions through a traditional specific PIT test. This test was followed by an assessment of the goal-directed properties and the integrity of the instrumental associations (i.e., A-O) using an outcome devaluation procedure. Our predictions were as follows. In the control animals, we expected that the inhibitory stimuli would bias choice away from an action delivering the outcome that they predicted would be omitted and towards actions delivering other outcomes. In contrast, we expected the LHb lesion to remove the inhibitory prediction of the cues and so to reverse this latter choice, with the stimuli now biasing choice towards the action with which they shared a common outcome. Finally, we were expecting all animals regardless of lesion to display goal-directed action control following outcome devaluation and so show a reliable outcome devaluation effect.

Materials and Methods

Backward Pavlovian Training - All animals (sham: *n* = 12; lesion: *n* =12) received eight daily sessions of backward Pavlovian training. This training consisted of twenty-four outcome deliveries (twelve for each outcome) followed by presentation of one of two stimuli that was turned on 10 s after the rats entered the magazine to consume the outcome. The stimuli (S1 and S2) were clicker and tone and the two distinct outcomes (O1 and O2) were pellet and sucrose solution. Delivery of O1 was also followed by S1 whereas delivery of O2 was always followed by S2. The stimulus duration varied from 2 to 58 s with an average of 30 s and separated by an intertrial interval varied from 80 to 200 s with an average of 150 s. We applied such parameters as it has been shown to generate Pavlovian inhibitors in the past (Laurent et al., 2014) and we assumed that it would promote instrumental responding in the presence of the stimuli during subsequent test, as it prevents the animals from timing outcome delivery. The order of the stimuli was varied between three set of pseudorandom order and the stimulus-outcome associations were counterbalanced between and within groups. Throughout the session, both levers were retracted and magazine entries were recorded and separated into stimulus period and a prior outcome delivery period with equal length (Pre O; 30 s).

Instrumental training - All rats then received eight days of instrumental training during which two actions (A1 and A2; left and right lever presses) were trained to deliver the two outcomes (O1 and O2) in separate daily sessions. The order of the sessions was counterbalanced, as were the response-outcome relationships that were also counterbalanced with the stimulus-outcome relationships established during Pavlovian training. Each session ended when 20 outcomes were earned or when 30 minutes had elapsed. For the first two days, lever pressing was continuously reinforced (i.e., each action earned an outcome). Then, the probability of the outcome given a

response was gradually shifted over days using increasing random ratio schedules: a RR5 schedule (p = 0.2) was used on days 3-5 and a RR10 schedule (p = 0.1) was used on days 6-8.

Pavlovian-instrumental transfer test - After the final day of RR10 training, rats received a Pavlovian reminder session that was followed the next day by a single Pavlovian-instrumental choice test. Both levers were inserted into the box, but no outcomes were delivered. Responding was extinguished on both R1 and R2 for 8 minutes to establish a low rate of baseline performance. Then, the rats received four 2-min presentations of each stimulus in the following order: tone-clicker-clicker-tone-clicker-tone-tone-clicker. The ITI was set at 3 minutes. Magazine entries and lever pressing rate were recorded and separated into Pre-S and S period.

Outcome devaluation test - Prior to the devaluation test, both groups received two days of instrumental re-training on an RR10 schedule. On the day of test, all rats received 1-hour access to one of the outcomes before being given a choice test. This test lasted 5 min and consisted of presenting the two levers but no outcome was delivered. The lever pressing rates were recorded throughout the test. The same procedure was repeated the following day except that the other outcome was devalued. The order of outcome devaluation was counterbalanced within groups.

Pavlovian (backward, 8 sessions)	Instrumental (8 sessions)	PIT	Devaluation test
O1-S1 O2-S2	A1-O1 A2-O2	S1/S2: A1 vs A2	Free access O1/O2: A1 vs A2

Figure 11. The design and training schedule of Experiment 2

Result and discussion

The data from Pavlovian training are presented in **Figure 12** and are plotted as the mean number of magazine entries across days. The inhibition generated by backward conditioning has been revealed in the past through a lack of difference in performance displayed in the presence of the stimuli or in their absence. This effect was clearly replicated in the present experiment. An ANOVA was conducted using Group (sham vs. lesion), Period (PreO vs. S) and Session as main factors. There was no main effect of Group ($F_{(1, 22)} = 0.387$, $p > 0.1$), no effect of Period ($F_{(1, 22)} = 0.983$, $p > 0.1$) but there was a main effect of Session ($F_{(7, 154)} = 2.302$, $p = 0.29$). Further, a significant Period x Session interaction ($F_{(7, 154)} = 24.357$, $p < 0.001$) revealed that any difference between the levels of magazine entries in the presence or absence of the stimuli disappeared across time. Importantly, there was no interaction involving Group as a factor (*Fs*<1.5)

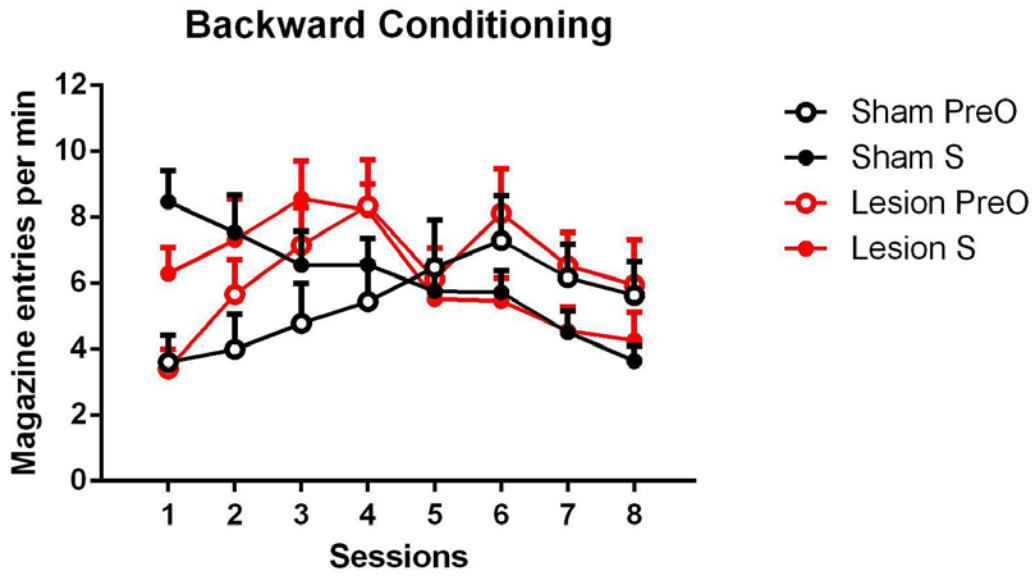


Figure 12. The rate of magazine entry across session in the backward conditioning stage. Error bar denotes the s.e.m.

The mean number of lever presses across instrumental training is plotted in **Figure 13**. This training occurred smoothly as the lever pressing rates gradually increased across days for both group. An ANOVA revealed no main effect of Group ($F_{(1, 22)} = 0.075$, $p > 0.1$), no Group x Session interaction ($F_{(7, 154)} = 0.655$, $p > 0.1$) but a main effect of Session ($F_{(7, 154)} = 269.987$, $p < 0.001$).

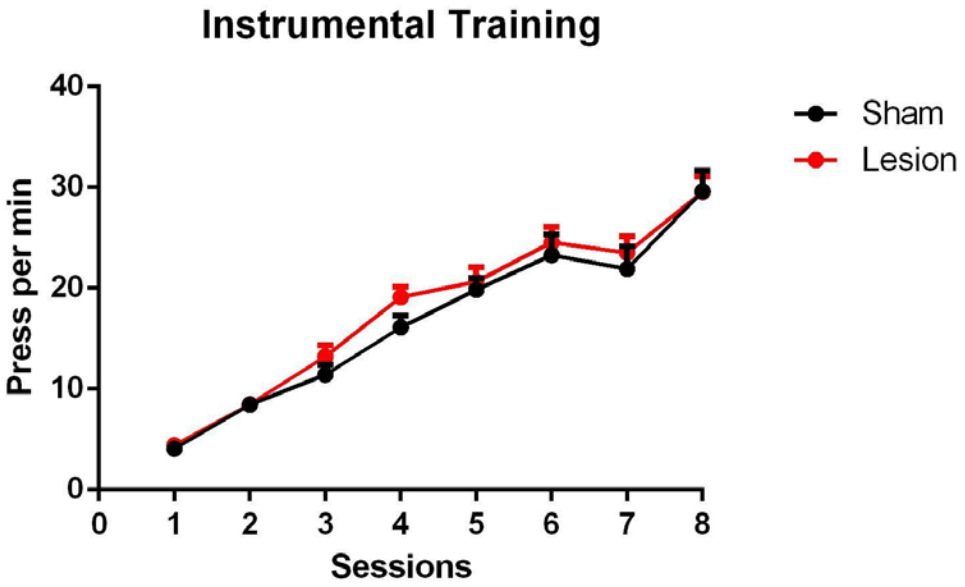


Figure 13. The lever pressing rate across sessions in the instrumental training stage. Error bar denotes s.e.m.

The data of most interest are those from the PIT test in **Figure 14** and they are plotted as the mean number of lever presses per minute when the stimulus predicted the absence of the same outcome as the response (Same), when the stimulus predicted the absence of the different outcome from the response (Different), or when no stimulus was present (Baseline). Choice between actions was clearly influenced by LHb lesion. An ANOVA was conducted using Group and Period (i.e., Baseline vs. Same vs. Different) as main effects, revealed a main effect of the latter ($F_{(2, 44)} = 6.305$, $p = 0.004$) but not of the former ($F_{(1, 22)} = 0.323$, $p > 0.1$). Critically, however, the analysis revealed a Group x Period interaction ($F_{(2, 44)} = 4.123$, $p = 0.023$). Subsequent pairwise comparison using the Bonferroni correction procedure revealed that the Sham group displayed higher responding on the Different action than during baseline ($T_{(11)} = 3.115$, $p = 0.029$). There was however no difference between responding on the Different and Same action ($T_{(11)} = 1.483$, $p > 0.1$) or between the Same action and Baseline ($T_{(11)} = 0.542$, $p > 0.1$). In contrast, the Lesion group displayed higher responding on the Same action than during baseline ($T_{(11)} = 3.551$, $p = 0.014$). However, this responding was no higher than that during the Different action ($T_{(11)} = 1.718$, $p > 0.1$). Finally, performance during the Different action was marginally higher than that during baseline ($T_{(11)} = 2.860$, $p = 0.047$). Thus, the present test replicated the traditional effect of inhibitory stimuli on choice between actions in the Sham group. That is, in this group, the stimuli biased choice towards the action delivering the outcome that they did not predict would be absent. This effect was removed by LHb lesion. In fact, in this group, the stimuli acted as if they were predicting their associated outcome, guiding choice towards the action earning that outcome.

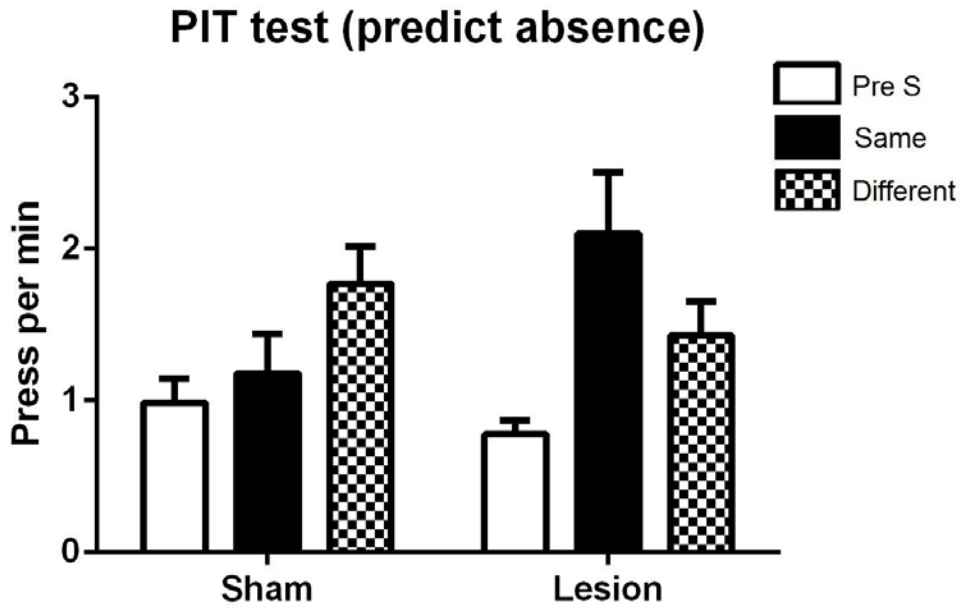


Figure 14. The lever pressing rate of both groups in the PIT test stage. The error bar represents s.e.m.

The result of the devaluation test is presented in **Figure 15** as the mean number of lever presses on the action delivering the outcome that had been devalued (Devalued) or on the action earning the outcome that remained valued (Non-devalued). Importantly, the data of the animals included in Chapter 5 (see below) were included in the analysis as these animals received a similar procedure. Inspection of the figure clearly indicates that the animals were able to choose an action based on the desirability of the outcome that it procured. An ANOVA using Group and Action (Non-devalued vs. Devalued) as main factors, revealed a main effect of the latter ($F_{(1, 47)} = 29.675$, $p < 0.001$) but not the former ($F_{(1, 47)} = 1.818$, $p > 0.1$). Further, the Group x Action interaction was not significant ($F_{(1, 47)} = 0.236$, $p > 0.1$). Thus, the LHb lesion did not impair the learning of specific action-outcome associations nor did it impair the ability of rats to use outcome value to guide choice between actions.

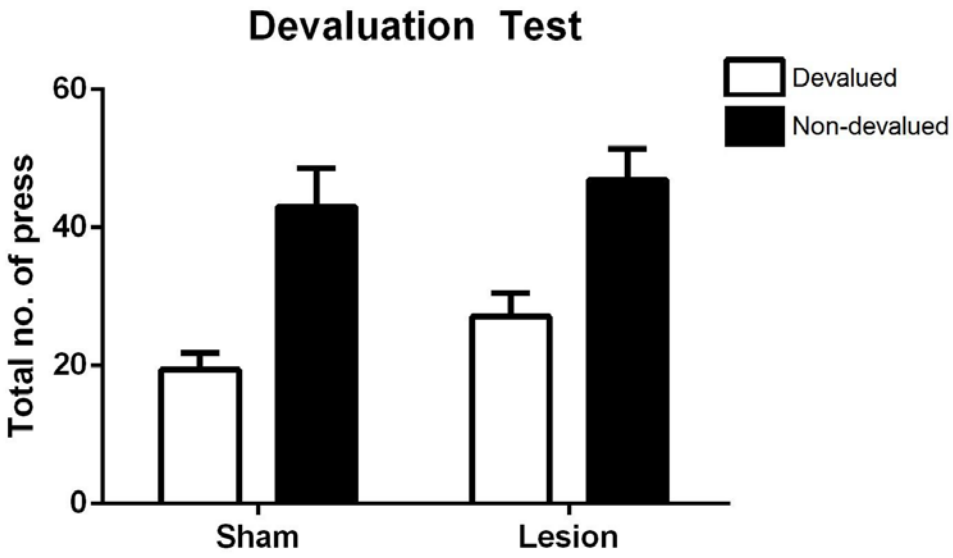


Figure 15. The total number of lever presses in both groups (included animals in Experiment 3) in the devaluation test stage. The error bar represent s.e.m.

The main finding from this experiment was that LHb lesion removed the usual influence that inhibitory stimuli exert on choice between actions. This usual influence was observed in control animals; i.e., a stimulus predicting the absence of a particular outcome was found to bias choice towards an action earning a different outcome. It should be noted, however, that we did observe a small but significant elevation of responding on the action delivering the outcome that was predicted to be absent by the inhibitory stimulus. This was unexpected and was in contrast with previous results obtained in our laboratory. The reason for such discrepancy remains unclear but it may be explained by some residual excitatory associations between the backwardly trained stimulus and its associated outcome. Indeed, it has previously been shown that backward conditioning can produce both excitatory and inhibitory associations between the stimulus and the outcome (Experiment 3, Laurent & Balleine, 2015). This analysis is supported by the effect of the LHb lesions: evidence of inhibitory stimulus-outcome association was lost as a consequence of LHb lesion. However, in lesioned animals, the stimuli that had been trained to predict the absence of a particular outcome biased choice towards the action earning that same outcome. Thus, not only did LHb lesions remove the usual influence of inhibitory stimuli on choice, it reversed it. In other words, the stimuli that had been trained as inhibitors appeared to act as positive predictors of their associated outcome. Importantly, the effect of LHb lesion was specific to the relationships that were established between the stimuli and their respective outcomes. For instance, control and lesion animals were able to select an action according to the value of the outcome that it procured: when this value was reduced, animals chose to perform the other action, the one earning a non-devalued outcome.

Taken together, the present results are therefore consistent with the view that the LHb is essential for encoding negative prediction error and, accordingly, for learning inhibitory stimulus-outcome

associations. This was evidenced during a specific PIT test during which LHb lesion was found to turn an inhibitory stimulus into an excitatory one. Yet, the present experiment did not provide any evidence that LHb lesion only disrupts inhibitory Pavlovian associations as opposed to all forms of Pavlovian associations, such as excitatory ones. The experiment conducted in the next chapter will address this issue.

Chapter 5: The role of LHb on choice guided by excitatory stimuli

Introduction

The experiment reported in the previous chapter revealed that LHb activity is critical for the bias that inhibitory stimuli exert on choice between actions. In the absence of such activity, this bias was found to be removed and even reversed, suggesting that the inhibitory stimuli were in fact acting as excitatory ones. However, it remains possible that the LHb altered the influence that Pavlovian associations may exert on choice between actions, whether these associations are inhibitory or excitatory in nature. Despite the apparent excitatory transfer in the lesion group observed in Experiment 2, the latter comparison was not made against an appropriate control. The present experiment addressed this issue by examining the effects of LHb lesion on the influence of excitatory Pavlovian associations on choice between actions. The procedure was similar to that previously used except that the stimulus-outcome associations were initially trained in a forward manner rather than a backward one (see **Figure 16**). Thus, two stimuli (S1 and S2) were trained to predict two distinct outcomes (O1 and O2). Instrumental training was then conducted during which the two outcomes could be earned by performing two distinct responses. A specific PIT test was then administered and assessed choice between the two instrumental actions in the absence or presence of either forward stimulus. If the LHb is essential for processing any kind of Pavlovian associations, we should observe a deficit in choice following local lesion.

Material and Methods

Forward Pavlovian training - All rats (group sham: n = 12 ; group lesion: n =13) were given eight daily sessions of Pavlovian training during which two stimuli (S1 and S2; clicker and tone) were paired with two distinct outcomes (O1 and O2; pellet and sucrose solution). The stimulus-outcome

relationships were fully counterbalanced between and within groups. During each session, each stimulus was presented four times in a pseudorandom order and every presentation lasted for 2 min followed by variable intertrial interval that averaged 5 min. The outcomes were delivered during each stimulus using a random time 30 s schedule. Throughout the session, both levers were retracted and magazine entries were recorded and separated into S period and a prior S period with equal length (Pre S; 2 min).

Instrumental training - All rats then received eight days of instrumental training during which two actions (A1 and A2; left and right lever presses) were trained to deliver the two outcomes (O1 and O2) in separate daily sessions. The order of the sessions was counterbalanced, as were the response-outcome relationships that were also counterbalanced with the stimulus-outcome relationships established during Pavlovian training. Each session ended when 20 outcomes were earned or when 30 minutes had elapsed. For the first two days, lever pressing was continuously reinforced (i.e., each action earned an outcome). Then, the probability of the outcome given a response was gradually shifted over days using increasing random ratio schedules: a RR5 schedule ($p = 0.2$) was used on days 3-5 and a RR10 schedule ($p = 0.1$) was used on days 6-8.

Pavlovian-instrumental transfer test - After the final day of RR10 training, rats received a Pavlovian reminder session that was followed the next day by a single Pavlovian-instrumental choice test. Both levers were inserted into the box, but no outcomes were delivered. Responding was extinguished on both R1 and R2 for 8 minutes to establish a low rate of baseline performance. Then, the rats received four 2-min presentations of each stimulus in the following order: tone-clicker-clicker-tone- clicker-tone-tone-clicker. The ITI was set at 3 minutes. Magazine entries and lever pressing rate were recorded and separated into Pre-S and S period.

Outcome devaluation test - Prior to the devaluation test, both groups received two days of instrumental re-training on an RR10 schedule. On the day of test, all rats received 1-hour access to one of the outcome before being submitted to a choice test. This test lasted 5 min and consisted of presenting the two levers but no outcome was delivered. The lever pressing rates were recorded throughout the test. The same procedure was repeated the following day except that the other outcome was devalued. The order of outcome devaluation was counterbalanced within groups. **The data from this test were presented in the previous chapter.**

Pavlovian (forward, 8 sessions)	Instrumental(8 sessions)	PIT	Devaluation test
S1-O1 S2-O2	A1-O1 A2-O2	S1/S2: A1 vs A2	Free access O1/O2: A1 vs A2

Figure 16. The design and training schedule of Experiment 3

Result and discussion

The data from Pavlovian training are presented in **Figure 17** and are plotted as the mean number of magazine entries across days. This training was successful as all groups discriminated between the stimulus period and the pre-stimulus period and this discrimination grew larger across session. An ANOVA was conducted using Group, Period and Session as main factors. It revealed a main effect of Period ($F_{(1, 23)} = 299.357$, $p < 0.001$) and Session ($F_{(7, 161)} = 8.427$, $p < 0.001$) but no main effect of Group ($F_{(1, 23)} = 0.005$, $p > 0.1$). A significant Period x Session interaction ($F_{(7, 161)} = 28.302$, $p < 0.001$) confirmed that animals were better at discriminating between the stimulus and pre-stimulus periods as training progressed. The analysis also revealed significant Group x Session ($F_{(7, 161)} = 2.559$, $p = 0.016$) and Group x Session x Period ($F_{(7, 161)} = 3.279$, $p = 0.003$) interactions,

revealing a slight decrease in performance of the lesion group towards the end of training. However, training was clearly successful in that group and the level of magazine entries prior to the stimulus and during the stimulus presentation were 1.63±0.30 and 16.37±1.71 per minute for the sham group, and were 2.00±0.57 and 12.80±1.97 per minute for the lesion group with respectively.

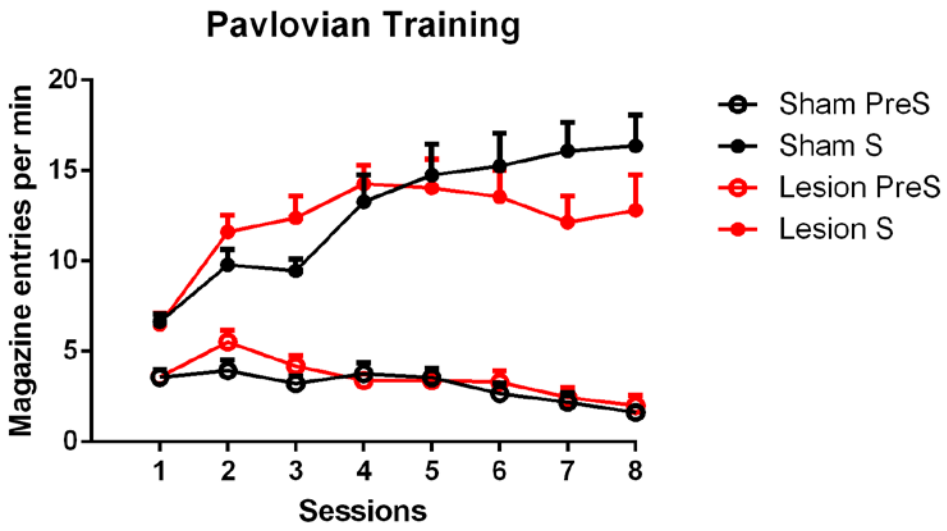


Figure 17. The rate of magazine entry of both groups across sessions in the Pavlovian training stage. The error bar represents the s.e.m.

The mean number of lever presses across instrumental training is plotted in **Figure 18**. This training occurred smoothly as the lever pressing rates gradual increased across days for both group. An ANOVA revealed no main effect of Group ($F_{(1, 23)} = 0.607$, $p > 0.1$), no Group x Session interaction ($F_{(7, 161)} = 0.559$, $p > 0.1$) but a main effect of Session ($F_{(7, 161)} = 147.554$, $p < 0.001$).

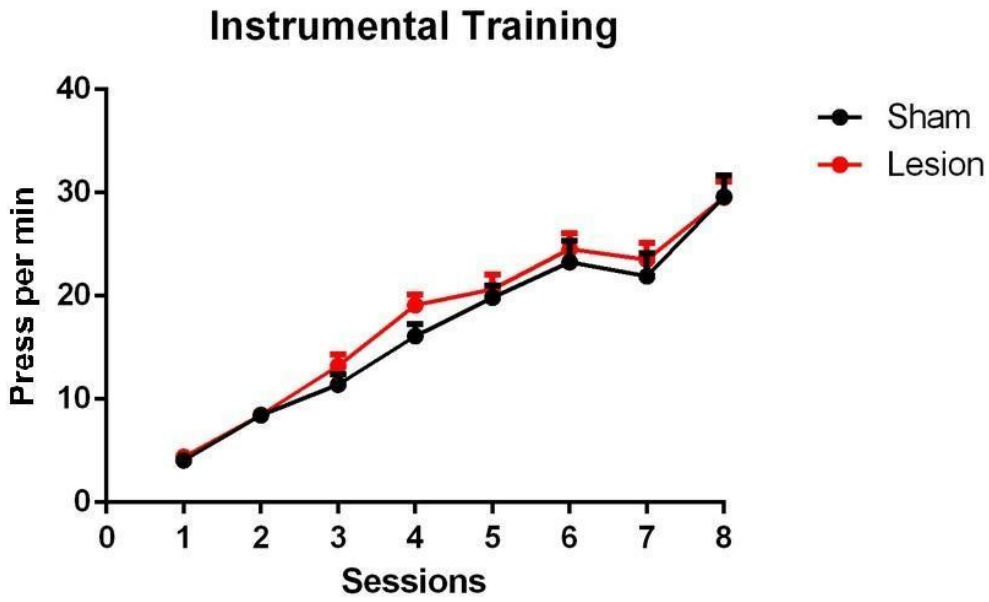


Figure 18. The lever pressing rate of both groups across sessions in the instrumental training stage. The error bar represents the s.e.m.

The data of most interest are those from the PIT test in **Figure 19** and they are plotted as the mean number of lever presses per minute when the stimulus predicted the same outcome as the response (Same), when the stimulus predicted the different outcome from the response (Different), or when no stimulus was present (Baseline). LHb lesion had clearly no effect on the influence exerted by excitatory stimuli on choice between actions. In all groups, a stimulus predicting a particular outcome was found to bias choice towards an action earning that same outcome. An ANOVA revealed a main effect of Period ($F_{(2, 46)} = 35.118$, $p < 0.01$) but no main effect of Group ($F_{(1, 23)} = 0.049$, $p > 0.1$) and no Group x Period interaction ($F_{(2, 46)} = 0.269$, $p > 0.1$). Subsequent pairwise comparison using the Bonferroni correction revealed that in both groups performance on the Same action was higher than performance on the Different action (Sham: $T_{(11)} = 4.209$, $p = 0.004$; Lesion: $T_{(12)} = 5.118$, $p = 0.001$) or than during baseline (Sham: $T_{(11)} = 4.098$, $p = 0.005$; Lesion: $T_{(12)} = 4.976$, $p = 0.001$).

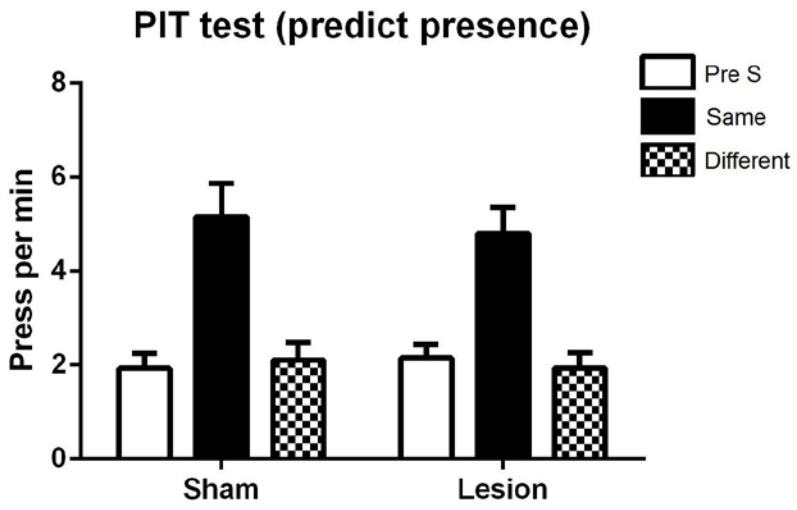


Figure 19. The lever pressing rate of both groups in the PIT test stage. The error bar represents the s.e.m.

The main finding from the present is that LHb lesion had no effect on the ability of excitatory stimuli to guide choice between actions. In both the control and lesion group, a stimulus predicting a particular outcome was found to bias choice towards an action delivering that same outcome. The presence of this bias indicates that the LHb is not necessary for establishing excitatory stimulus-outcome relationships or for promoting that these relationships exert on choice. It also shows that the LHb is not required for learning specific action-outcome associations, consistent with the ability of animals with LHb lesion to select an action according to the value of the outcome that it procures (see previous chapter). Taken with the result provided in the previous chapter, it appears that the LHb is specifically involved in establishing inhibitory stimulus-outcome associations and therefore, in promoting the influence that such associations have on choice between actions.

Chapter 6: General Discussion

Overview of the results

The current studies aimed at determining the role of the LHb in stimulus-based choice, especially in situations where the stimuli had been trained to predict the absence of a food outcome (i.e., inhibitory stimuli). Our assessment was completed by performing electrolytic lesion of the LHb and then, by submitting animals to two distinct behavioural paradigms. One involved conditioned inhibition whereas the other was specific PIT. The former paradigm was used to evaluate how LHb lesion would affect Pavlovian inhibitory learning, whereas the latter tested how Pavlovian inhibitors influence choice between actions.

The first experiment (Chapter 3) used a conditioned inhibition design initially described by Rescorla (2002) to evaluate the role of the LHb in the formation of condition inhibition. The selection of this design was motivated by the presence of appropriate control (i.e., latent inhibitor) to assess the development of inhibitory learning. Although we used parameters almost identical to as those of Rescorla (2002), we obtained little evidence of Pavlovian inhibition. Thus, the conditioned inhibitor displayed weak retardation and failed to pass a subtractive summation test. However, it was interesting to note that the weak trend for retardation was removed by LHb lesion.

The next experiment (Chapter 4) went further and examined the effects of LHb lesion on stimulus-based choice. To do so, it adopted a PIT paradigm from Delamater et al. (2003) and Laurent et al. (2015b). Thus, sham and lesioned animals were initially given backward conditioning in order to generate two specific Pavlovian inhibitors. The animals then received instrumental training followed by a PIT choice test. This test revealed a different pattern of choice in the two experimental groups. The sham animal showed a reversal of the traditional PIT effect: the stimuli biased choice away from the action delivering the outcome they predicted the absence

of, towards the action earning a different outcome. Remarkably, this reversal of choice was abolished in LHb lesioned animals. In fact, those animals displayed the traditional PIT effect: the stimuli biased choice towards the action delivering the outcome that they predicted would be absent. Thus, the inhibitory stimuli acted as excitatory ones when the LHb had been damaged prior to training. This finding provides convincing evidence that the LHb is important in mediating the inhibitory processes that allow the reversal of choice between actions.

In the last chapter (Chapter 5), we examined whether the role of LHb also extended to the influence that excitatory stimuli exert on choice between actions. To do so, sham and lesioned animals were initially given forward conditioning in order to generate two specific Pavlovian excitors. Following instrumental conditioning, a PIT test was administered to assess the influence of the two excitatory stimuli on choice between actions. As expected, control/sham animals displayed the traditional PIT effects: the stimuli biased choice towards the action with which they shared a common outcome. Importantly, this bias was left unaffected by LHb lesion, a result contrasting with the effects obtained in Chapter 4. Further, we also found LHb lesion failed to impair value-based choice. That is, sham and lesion animals were able to choose an action based on the value of the outcome it procures. Taken together, these findings indicate that the LHb is critical for stimulus-based, and not value-based, choice in situations where the stimuli have been trained as negative, but not positive, predictors of their associated outcomes.

The role of LHb and the neural basis of backward conditioning

Although our first experiment failed to provide conclusive evidence for a role of the LHb in Pavlovian inhibitory learning, the remaining experiments did support such a role. For instance, LHb lesion had no effect on value-based choice but it did disrupt stimulus-based choice. However, this disruption was highly specific, as it was only present when the stimuli were inhibitory and not

excitatory. Taken together, these results suggest that the LHb may be involved in the inhibitory processes that develop across backward conditioning. This form of conditioning constitutes a challenge for most contemporary learning theories based on prediction error, as the variation in RPE after the US and its potential role on learning is rarely discussed. One notable exception is the so-called *A*ffective *E*xtension of the *S*ometimes *O*pponent *P*rocess (AESOP) theory developed by Wagner and Brandon (1989). This theory holds that perception of a particular stimulus triggers its sensory and motivational features to be represented in an active state called A1. As time passes, the representation decays into another state called A2 before it returns to an inactive state. The theory then argues that if the representation of two stimuli (e.g., a tone and a food outcome) are concurrently represented into A1, then an excitatory relationship can be established between the two events. This is essentially what is happening during forward conditioning. In contrast, during backward conditioning, the food outcome is likely to be into the A2 state by the time the tone is encountered and represented in the A1 state. This is because of the delay inserted in between food delivery and the tone presentation. The consequence of the food being in A2 and the tone in A1 is the establishment of an inhibitory relationship between the two. Given the absence of such inhibition in the present experiments, it is possible that LHb lesion affects the representation of the stimuli and the food outcomes in the A1 and A2 states. Specifically, the PIT test results suggested that the lesion group had formed an excitatory association between the two backward stimuli and their associated outcomes. One potential explanation is that LHb lesion could have extended the time during which the outcome was represented into the A1 state, allowing the establishment of an excitatory association between this outcome and the following stimulus. Interestingly, in-vivo electrophysiology and other data provide some support to this explanation. For example, Chang and Kim (2004) have revealed that some neurons in the LHb display a biphasic response following

stimulation of the stria medullaris. This biphasic response involves an initial hyperpolarization followed by a depolarized state that then triggers a long-lasting and intense action potential. The depolarization was found to occur around 1 second after the hyperpolarization and to last more than 10 seconds. Here, we propose that this biphasic pattern is triggered in the LHb following delivering of the outcome and would be equivalent to a shift in the representation of this outcome from the A1 into A2 state. As explained, this shift would be essential to establish the inhibitory association with the incoming stimulus but it would be absent in the rats that had received the LHb lesion. This proposal is clearly speculative and further work is required to assess its validity. However, it is interesting to note that artificially triggering an LHb biphasic response has been shown to induce place avoidance (Jhou et al., 2013). If one assumes that avoidance is somewhat similar to learning that a stimulus signals the absence of reward, such data appears consistent with our proposal.

Theoretical and methodological concerns

The influence of serotonin

This thesis has focused on the effects of LHb lesions on inhibition by listing evidence for its role in reward prediction error and its action on midbrain dopamine neurons. Particularly, it has emphasized the suppressive effect exerted by the LHb on these dopamine neurons, a suppression of activity that is believed to occur via the RMTg. However, we cannot exclude the possibility that the effects reported in the current thesis were due to a disinhibition of serotonergic activities in DRN and MRN resulting from LHb lesion. Although the role played by serotonin in aversive learning has received substantial attention, we know very little about the influence that it may exert on other forms of learning such as Pavlovian appetitive conditioning or PIT. Yet, several studies do suggest a critical role. For example, it has been shown that mice carrying knockout of the

dopamine and serotonin transporters are unable to show conditioned place preference (Hall, Sora & Uhl, 2004). Further, several findings have suggested that the modulation of the serotonin system could be rewarding either by reducing synthesis of serotonin in the MRN (Fletcher, Ming & Higgins, 1993; Fletcher, Tampakeras & Yeomans, 1995) or by activation of DRN serotonin neurons (Liu et al., 2014, but also see Fonseca, Murakiami & Mainen, 2015). The role of the DRN in reward processing could be quite significant as it projects to both the VTA and the NAc directly (Zhao et al., 2015). Interestingly, several studies have pointed to a direct role for serotonin in PIT, although this role was argued to be limited to the aversive form of the PIT paradigm (Geurts, Huys, den Ouden & Cools, 2013; Hebart & Gläscher., 2015). Taken with other studies regarding the influence of Pavlovian cues on instrumental responding (Crockett, Clark & Robbins, 2009; Crockett et al., 2012; Faulkner & Deakin., 2014), it was revealed that serotonin depletion reduces behavioural inhibition towards an aversive stimulus. Provided that an aversive stimulus and an appetitive inhibitor can be considered as being similar at a motivational level, we cannot exclude a potential role for serotonin in the experiments presented here. However, the effects of the lesion conducted here should have led to an overall increase in serotonin activity. So far, the behavioural consequences of such an increase remain unknown, highlighting the need for further work.

The effect of permanent electrolytic lesion

All our experiments have employed electrolytic lesions to damage the LHb and evaluate the consequence of such damage on Pavlovian inhibition and choice between actions. Initially, we had intended to use chemogenetic or pharmacological approaches to investigate the role played by the LHb. Unfortunately, the proximity of the third ventricle made it very difficult to accurately target the structure of interest. One issue with using electrolytic lesions is that it is hard to contain the extent of the lesion. As a result, some of the animals included in the present experiments displayed

some partial damages in the MDT. As mentioned in the introduction, the MDT has been shown to be involved in specific PIT and outcome devaluation (Corbit et al., 2003; Ostlund & Balleine, 2008; Parnaudeau et al., 2015). Yet, our experiments did not reveal any impairment in either outcome devaluation or specific PIT when the influence of excitatory stimuli was being assessed. Thus, it is unlikely that partial damages of the MDT could explain the results presented in this thesis.

Another disadvantage of employing electrolytic lesions is the potential for compensatory mechanisms that occur as a result of permanent brain damage. However, a large effect of compensatory processes appears unlikely. Indeed, the effects of the lesion were found to be highly specific, impairing choice when driven by inhibitory stimuli but not when produced by excitatory stimuli. It is difficult to explain how compensatory mechanisms may have promoted the latter but not the former.

Future directions

The present experiments indicate that the LHb is necessary for the ability of inhibitory stimuli to drive choice between actions. However, it remains uncertain whether this brain region is specifically involved in such choice or in learning about the inhibitory stimulus. For instance, the first experiment described in this thesis failed to provide convincing evidence that the LHb is necessary to the formation of conditioned inhibition and, at least behaviorally, backward conditioning had much the same effect in lesioned and unlesioned rats in Chapter 4. Thus, experiments are required to further explore the role of this structure in Pavlovian inhibition. These experiments could employ other forms of inhibitory learning than backward conditioning and use other methods than electrolytic lesion to avoid extensive damage. Such methods could for example

include in-vivo electrophysiology or optogenetics. It would also be particularly interesting to examine whether the LHb plays a general role in inhibitory processes such as those involved in the extinction of Pavlovian conditioning. In addition, further experiments are required to determine the potential role played by serotonin in choice between actions, as the lesion used in our experiments is likely to have disrupted serotonin functioning in the DRN. It would also be essential to precisely delineate the neural circuitry that allows the LHb to modulate specific PIT in the presence of inhibitory stimuli. It would be particularly useful to examine the role of the RMTg, as this structure is assumed to drive the changes in midbrain dopamine activity that are driven by the LHb.

Conclusions

Successful adaptation requires the capacity to extract excitatory and inhibitory information from the environment to guide choice between courses of action. Excitatory information increases the chance of gathering commodities whereas inhibitory information can help to save effort and time by shifting choice away from actions associated with outcomes that are unlikely to occur. The current study aimed at investigating some of the neural substrates that allow inhibitory information to guide choice. Specifically, it assessed the potential role of the LHb. The present experiments indicate that this structure is involved inhibitory-driven choice and provides evidence that the LHb may play a general role in inhibitory processes. Presumably, the LHb is able to play such a role by modulating activity of midbrain dopamine neurons that are believed to compute reward prediction error. Further experiments will be necessary to validate this proposal.

References.

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly journal of experimental psychology, 33*(2), 109-121.

Amat, J., Sparks, P., Matus-Amat, P., Griggs, J., Watkins, L., & Maier, S. (2001). The role of the habenular complex in the elevation of dorsal raphe nucleus serotonin and the changes in the behavioral responses produced by uncontrollable stress. *Brain research, 917*(1), 118-126.

Assadi, S. M., Yücel, M., & Pantelis, C. (2009). Dopamine modulates neural networks involved in effort-based decision-making. *Neuroscience & Biobehavioral Reviews, 33*(3), 383-393.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*(4), 407-419.

Balleine, B. W., & Dickinson, A. (1998). The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Animal Learning & Behavior, 26*(1), 46-59.

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in neurosciences, 29*(5), 272-279.

Balleine, B. W., Leung, B. K., & Ostlund, S. B. (2011). The orbitofrontal cortex, predicted value, and choice. *Annals of the New York Academy of Sciences, 1239*(1), 43-50.

Balleine, B. W., Morris, R. W., & Leung, B. K. (2014). Thalamocortical integration of instrumental learning and performance and their disintegration in addiction. *Brain research, 1628,* 104-116.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28*(3), 309-369.

Bertran-Gonzalez, J., Hervé, D., Girault, J.-A., & Valjent, E. (2010). What is the degree of segregation between striatonigral and striatopallidal projections. *Front Neuroanat, 4*(pii), 136.

Bertran-Gonzalez, J., Laurent, V., Chieng, B. C., Christie, M. J., & Balleine, B. W. (2013). Learning-related translocation of δ-opioid receptors on ventral striatal cholinergic interneurons mediates choice between goal-directed actions. *The Journal of neuroscience, 33*(41), 16060-16071.

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron, 68*(5), 815-834.

Cador, M., Robbins, T., & Everitt, B. (1989). Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. *Neuroscience, 30*(1), 77-86.

Cazala, P., Norena, A., Le Merrer, J., & Galey, D. (1998). Differential involvement of the lateral and medial divisions of the septal area on spatial learning processes as revealed by intracranial self-administration of morphine in mice. *Behavioural brain research, 97*(1), 179-188.

Chang, S.-y., & Kim, U. (2004). Ionic mechanism of long-lasting discharges of action potentials triggered by membrane hyperpolarization in the medial lateral habenula. *The Journal of neuroscience, 24*(9), 2172-2181.

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature, 482*(7383), 85-88.

Colwill, R. M., & Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of experimental psychology: animal behavior processes, 11*(1), 120.

Colwill, R. M., & Rescorla, R. A. (1990). Effect of reinforcer devaluation on discriminative control of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 16*(1), 40.

Corbit, L. H., & Balleine, B. W. (2003). Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *Journal of Experimental Psychology: Animal Behavior Processes, 29*(2), 99.

Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *The Journal of neuroscience, 25*(4), 962-970.

Corbit, L. H., & Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *The Journal of neuroscience, 31*(33), 11786-11794.

Corbit, L. H., & Janak, P. H. (2007). Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *The Journal of neuroscience, 27*(51), 13977-13981.

Corbit, L. H., Janak, P. H., & Balleine, B. W. (2007). General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *European Journal of Neuroscience, 26*(11), 3141-3149.

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *The Journal of neuroscience, 21*(9), 3251-3260.

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2003). Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *European Journal of Neuroscience, 18*(5), 1286-1294.

Crockett, M. J., Clark, L., Apergis-Schoute, A. M., Morein-Zamir, S., & Robbins, T. W. (2012). Serotonin modulates the effects of Pavlovian aversive predictions on response vigor. *Neuropsychopharmacology, 37*(10), 2244-2252.

Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *The Journal of neuroscience, 29*(38), 11993-11999.

Danna, C. L., Shepard, P. D., & Elmer, G. I. (2013). The habenula governs the attribution of incentive salience to reward predictive cues. *Frontiers in human neuroscience, 7,* 781.

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychological Research PRPF, 73*(4), 463-476.

Delamater, A. R., Sosa, W., & LoLordo, V. M. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Animal Learning & Behavior, 31*(4), 393-402.

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 308*(1135), 67-78. doi:10.1098/rstb.1985.0010

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior, 22*(1), 1-18.

Engbretson, G. A., Reiner, A., & Brecha, N. (1981). Habenular asymmetry and the central connections of the parietal eye of the lizard. *Journal of Comparative Neurology, 198*(1), 155-165.

Faulkner, P., & Deakin, J. W. (2014). The role of serotonin in reward, punishment and behavioural inhibition in humans: Insights from studies with acute tryptophan depletion. *Neuroscience & Biobehavioral Reviews, 46,* 365-378.

Fletcher, P. J., Ming, Z.-H., & Higgins, G. A. (1993). Conditioned place preference induced by microinjection of 8-OH-DPAT into the dorsal or median raphe nucleus. *Psychopharmacology, 113*(1), 31-36.

Fletcher, P. J., Tampakeras, M., & Yeomans, J. S. (1995). Median raphe injections of 8-OH-DPAT lower frequency thresholds for lateral hypothalamic self-stimulation. *Pharmacology Biochemistry and Behavior, 52*(1), 65-71.

Floresco, S. B., Maric, T., & Ghods-Sharifi, S. (2008). Dopaminergic and glutamatergic regulation of effort-and delay-based decision making. *Neuropsychopharmacology, 33*(8), 1966-1979.

Fonseca, M. S., Murakami, M., & Mainen, Z. F. (2015). Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing. *Current Biology, 25*(3), 306-315.

Friedman, A., Lax, E., Dikshtein, Y., Abraham, L., Flaumenhaft, Y., Sudai, E., . . . Yadid, G. (2010). Electrical stimulation of the lateral habenula produces enduring inhibitory effect on cocaine seeking behavior. *Neuropharmacology, 59*(6), 452-459.

Friedman, A., Lax, E., Dikshtein, Y., Abraham, L., Flaumenhaft, Y., Sudai, E., . . . Yadid, G. (2011). Electrical stimulation of the lateral habenula produces an inhibitory effect on sucrose self-administration. *Neuropharmacology, 60*(2), 381-387.

Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annual review of neuroscience, 34*, 441.

Geurts, D. E., Huys, Q. J., den Ouden, H. E., & Cools, R. (2013). Serotonin and aversive Pavlovian control of instrumental behavior in humans. *The Journal of neuroscience, 33*(48), 18932-18939.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences, 108*(Supplement 3), 15647-15654.

Goncalves, L., Sego, C., & Metzger, M. (2012). Differential projections from the lateral habenula to the rostromedial tegmental nucleus and ventral tegmental area in the rat. *Journal of Comparative Neurology, 520*(6), 1278-1300.

Hall, F. S., Sora, I., & Uhl, G. R. (2004). Molecular Mechanisms Underlying Behavioral Effects of Cocaine: Lessons From Knockout Mice. *Current Pharmacogenomics, 2*(2), 185-202.

Hall, J., Parkinson, J. A., Connor, T. M., Dickinson, A., & Everitt, B. J. (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *European Journal of Neuroscience, 13*(10), 1984-1992.

Hebart, M. N., & Gläscher, J. (2015). Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. *Psychopharmacology, 232*(2), 437-451.

Herkenham, M., & Nauta, W. J. (1977). Afferent connections of the habenular nuclei in the rat. A horseradish peroxidase study, with a note on the fiber-of-passage problem *Neuroanatomy* (pp. 277-299): Springer.

Herkenham, M., & Nauta, W. J. (1979). Efferent connections of the habenular nuclei in the rat. *Journal of Comparative Neurology, 187*(1), 19-47.

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes, 30*(2), 104.

Holland, P. C., & Gallagher, M. (2003). Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer. *European Journal of Neuroscience, 17*(8), 1680-1694.

Holland, P. C., & Gallagher, M. (2004). Amygdala–frontal interactions and reward expectancy. *Current opinion in neurobiology, 14*(2), 148-155.

Hong, S., Jhou, T. C., Smith, M., Saleem, K. S., & Hikosaka, O. (2011). Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *The Journal of neuroscience, 31*(32), 11457-11471.

Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens–olfactory tubercle complex. *Brain research reviews, 56*(1), 27-78.

Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B., & Holland, P. C. (2009). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron, 61*(5), 786-800.

Jhou, T. C., Good, C. H., Rowley, C. S., Xu, S.-p., Wang, H., Burnham, N. W., . . . Ikemoto, S. (2013). Cocaine drives aversive conditioning via delayed activation of dopamine-responsive habenular and midbrain pathways. *The Journal of neuroscience, 33*(17), 7501-7512.

Kale, P., Strecker, R. E., Rosengren, E., & Bjo, A. (1989). Regulation of striatal serotonin release by the lateral habenula-dorsal raphe pathway in the rat as demonstrated by in vivo microdialysis: role of excitatory amino acids and GABA. *Brain research, 492*(1-2), 187-202.

Kemali, M., & Guglielmotti, V. (1977). An electron microscope observation of the right and the two left portions of the habenular nuclei of the frog. *Journal of Comparative Neurology, 176*(2), 133-147.

Kim, U. (2009). Topographic commissural and descending projections of the habenula in the rat. *Journal of Comparative Neurology, 513*(2), 173-187.

Laurent, V., & Balleine, B. W. (2015). Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats. *Current Biology, 25*(8), 1074-1079.

Laurent, V., Bertran-Gonzalez, J., Chieng, B. C., & Balleine, B. W. (2014). δ-opioid and dopaminergic processes in accumbens shell modulate the cholinergic control of predictive learning and choice. *The Journal of neuroscience, 34*(4), 1358-1369.

Laurent, V., Leung, B., Maidment, N., & Balleine, B. W. (2012). μ-and δ-opioid-related processes in the accumbens core and shell differentially mediate the influence of reward-guided and stimulus-guided decisions on choice. *The Journal of neuroscience, 32*(5), 1875-1883.

Laurent, V., Wong, F. L., & Balleine, B. W. (2015). δ-Opioid receptors in the accumbens shell mediate the influence of both excitatory and inhibitory predictions on choice. *British journal of pharmacology, 172*(2), 562-570.

Le Merrer, J., Gavello-Baudy, S., Galey, D., & Cazala, P. (2007). Morphine self-administration into the lateral septum depends on dopaminergic mechanisms: Evidence from pharmacology and Fos neuroimaging. *Behavioural brain research, 180*(2), 203-217.

Lecourtier, L., DeFrancesco, A., & Moghaddam, B. (2008). Differential tonic influence of lateral habenula on prefrontal cortex and nucleus accumbens dopamine release. *European Journal of Neuroscience, 27*(7), 1755-1762.

Lex, A., & Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learning & Memory, 15*(7), 483-491.

Lisoprawski, A., Herve, D., Blanc, G., Glowinski, J., & Tassin, J. (1980). Selective activation of the mesocortico-frontal dopaminergic neurons induced by lesion of the habenula in the rat. *Brain research, 183*(1), 229-234.

Liu, Z., Zhou, J., Li, Y., Hu, F., Lu, Y., Ma, M., . . . Zeng, J. (2014). Dorsal raphe neurons signal reward through 5-HT and glutamate. *Neuron, 81*(6), 1360-1374.

Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature, 447*(7148), 1111-1115.

Matsumoto, M., & Hikosaka, O. (2009). Representation of negative motivational value in the primate lateral habenula. *Nature neuroscience, 12*(1), 77-84.

Nishikawa, T., Fage, D., & Scatton, B. (1986). Evidence for, and nature of, the tonic inhibitory influence of habenulointerpeduncular pathways upon cerebral dopaminergic transmission in the rat. *Brain research, 373*(1), 324-336.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology, 47*(6), 419.

Ostlund, S. B., & Balleine, B. W. (2007). Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Animal Learning & Behavior, 35*(1), 43-52.

Ostlund, S. B., & Balleine, B. W. (2008). Differential involvement of the basolateral amygdala and mediodorsal thalamus in instrumental action selection. *The Journal of neuroscience, 28*(17), 4398-4405.

Ostlund, S. B., & Maidment, N. T. (2011). Dopamine receptor blockade attenuates the general incentive motivational effects of noncontingently delivered rewards and reward-paired cues without affecting their ability to bias action selection. *Neuropsychopharmacology, 37*(2), 508-519.

Parkes, S. L., & Balleine, B. W. (2013). Incentive memory: evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. *The Journal of neuroscience, 33*(20), 8753-8763.

Parnaudeau, S., Taylor, K., Bolkan, S. S., Ward, R. D., Balsam, P. D., & Kellendonk, C. (2015). Mediodorsal thalamus hypofunction impairs flexible goal-directed behavior. *Biological psychiatry, 77*(5), 445-453.

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behavioral neuroscience, 119*(1), 317.

Proulx, C. D., Hikosaka, O., & Malinow, R. (2014). Reward processing by the lateral habenula in normal and depressive behaviors. *Nature neuroscience, 17*(9), 1146-1152.

Reisine, T. D., Soubrie, P., Artaud, F., & Glowinski, J. (1982). Involvement of lateral habenula-dorsal raphe neurons in the differential regulation of striatal and nigral serotonergic transmission cats. *The Journal of neuroscience, 2*(8), 1062-1071.

Rescorla, R. A. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior, 22*(1), 27-33.

Rescorla, R. A. (2002). Savings tests: separating differences in rate of learning from differences in initial levels. *Journal of Experimental Psychology: Animal Behavior Processes, 28*(4), 369.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory, 2*, 64-99.

Shabel, S. J., Proulx, C. D., Trias, A., Murphy, R. T., & Malinow, R. (2012). Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin. *Neuron, 74*(3), 475-481.

Sheehan, T. P., Chambers, R. A., & Russell, D. S. (2004). Regulation of affect by the lateral septum: implications for neuropsychiatry. *Brain Research Reviews, 46*(1), 71-117.

Shiflett, M. W., & Balleine, B. W. (2010). At the limbic–motor interface: disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation. *European Journal of Neuroscience, 32*(10), 1735-1743.

Shiflett, M. W., Brown, R. A., & Balleine, B. W. (2010). Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *The Journal of neuroscience, 30*(8), 2951-2959.

Silver, R., Silverman, A.-J., Vitković, L., & Lederhendler, I. I. (1996). Mast cells in the brain: evidence and functional significance. *Trends in neurosciences, 19*(1), 25-31.

Stamatakis, A. M., Jennings, J. H., Ung, R. L., Blair, G. A., Weinberg, R. J., Neve, R. L., . . . Deisseroth, K. (2013). A unique population of ventral tegmental area neurons inhibits the lateral habenula to promote reward. *Neuron, 80*(4), 1039-1053.

Stamatakis, A. M., & Stuber, G. D. (2012). Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nature neuroscience, 15*(8), 1105-1107.

Stephenson-Jones, M., Floros, O., Robertson, B., & Grillner, S. (2012). Evolutionary conservation of the habenular nuclei and their circuitry controlling the dopamine and 5-hydroxytryptophan (5-HT) systems. *Proceedings of the National Academy of Sciences, 109*(3), E164-E173.

Stopper, C. M., & Floresco, S. B. (2013). What's better for me [quest] Fundamental role for lateral habenula in promoting subjective decision biases. *Nature neuroscience, 17*(1), 33-35.

Stopper, C. M., Maric, T., Montes, D. R., Wiedman, C. R., & Floresco, S. B. (2014). Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron, 84*(1), 177-189.

Tian, J., & Uchida, N. (2015). Habenula lesions reveal that multiple mechanisms underlie dopamine prediction errors. *Neuron, 87*(6), 1304-1316.

Wagner, A. R., Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning. In S.B. Klein and R.R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theories*. (pp. 149–189). Hillsdale, NJ: Erlbaum.

Wilhelm, M., King, B., Silverman, A.-J., & Silver, R. (2000). Gonadal Steroids Regulate the Number and Activational State of Mast Cells in the Medial Habenula 1. *Endocrinology, 141*(3), 1178-1186.

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience, 22*(2), 505-512.

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behavioural brain research, 166*(2), 189-196.

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience, 22*(2), 513-523.

Zhao, H., Zhang, B.-L., Yang, S.-J., & Rusak, B. (2015). The role of lateral habenula–dorsal raphe nucleus circuits in higher brain functions and psychiatric illness. *Behavioural brain research, 277*, 89-98.

Appendix

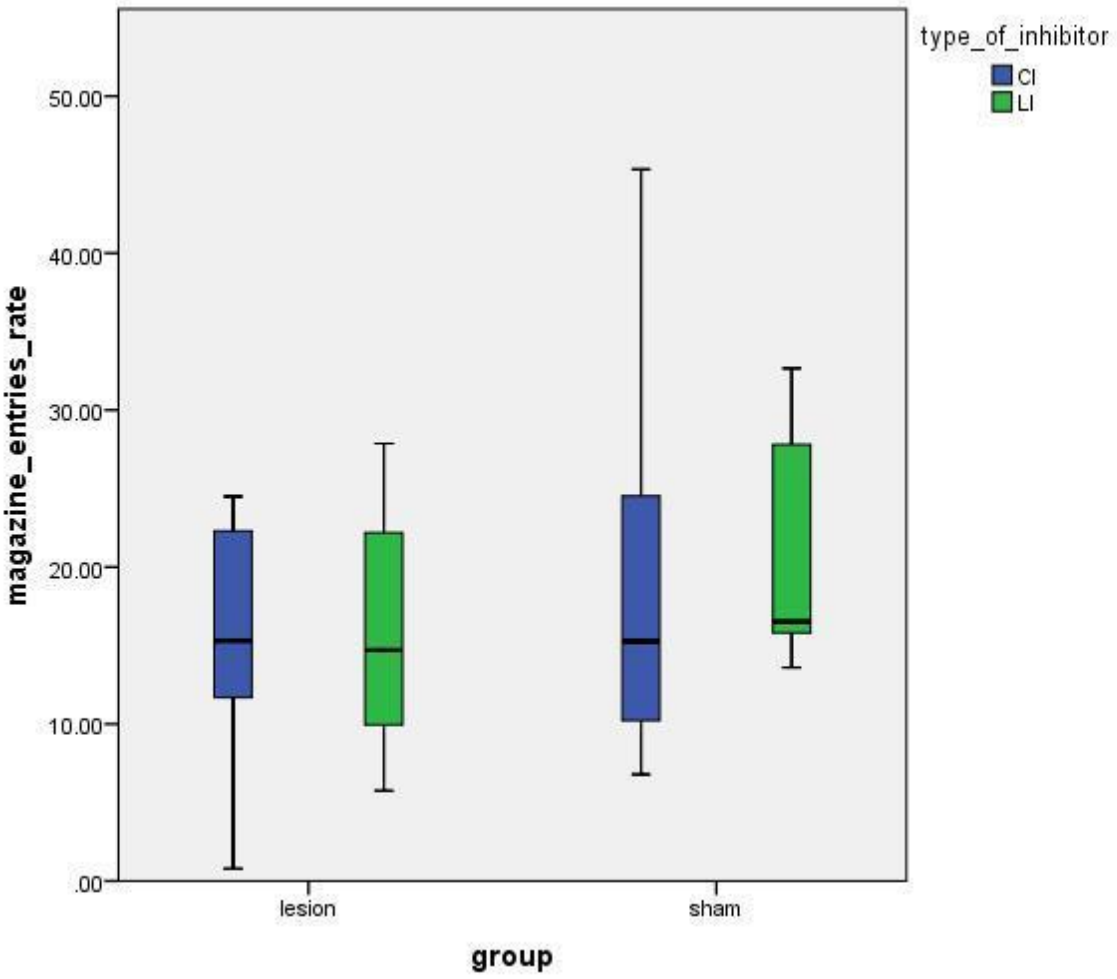


Figure A1. The box plot of the mean magazine entries rate when the S2(CI)/S4(LI) is presented of both groups during the acquisition test in experiment 1.