Integration of appetitive and aversive reinforcers and the neuromodulation of reward seeking and pain avoidance

Dissertation

Zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften

der Otto-von-Guericke-Universität Magdeburg

von: **M.Sc, Anton Ilango Micheal**

geb. am 03. Mai 1978 in Tirunelveli, Indien

Gutachter: Prof. Dr. Frank W. Ohl

                     Prof. Dr. Holger Schulze

eingereicht am: 21. Juni 2010

verteidigt am: 27. Oktober 2010

Acknowledgements

Abbreviations

ANOVA   analysis of variance

BSR   brain stimulation reward

CR   conditioned response

CS   conditioned stimulus

DA   dopamine

DOPAC   dihydroxyphenylacetic acid

FS   footshock

FSCV   fast scan cyclic voltammetry

GABA   gamma-aminobutyric acid

ICSS   intracranial self-stimulation

LH   lateral hypothalamus

LHb   lateral habenula

MFB   medial forebrain bundle

mPFC   medial prefrontal cortex

NAc   nucleus accumbens

OFC   orbitofrontal cortex

PBP   parabrachial pigmented nucleus

PFR   parafasciculus retroflexus

PIF   parainterfasciculus retroflexus

PN   paranigral

PPTg   pedunculopontine tegmental nucleus

REM   rapid eye movement

RMTg   rostromedial tegmental nucleus

RLi   rostral linear nuclei

RT   reaction time

S-R   stimulus-response

SEM   standard error of the mean

SNc   substantia nigra pars compacta

SNr   substantia nigra pars reticulata

SRN   serotonergic raphe nuclei

TH   tyrosine hydroxylase

US   unconditioned stimulus

VTA   ventral tegmental area

VP   ventral pallidum

i.e.   *id est* (that is to say)

e.g.   *exempli gratia* (for example)

6-OHDA   6-hydroxydopamine

Abstract

New behaviours in animal and man can be acquired, in principle, by either reward- or punishment-reinforced learning. But as popular wisdom maintains, learning may be most efficient if "carrot and stick" reinforcements are combined. In spite of its high theoretical, clinical and educational relevance, neither the general nature nor the detailed dynamics of the direct interaction of reward and punishment nor its dynamics during learning are understood.

Midbrain dopamine system, especially the ventral tegmental area (VTA) plays a vital role in motivated behaviour. Electrical stimulation of this system has a positively reinforcing effect on behaviour. Using this feature of this widely projecting reward system, we first studied the acquisition and extinction of the tone conditioned hurdle crossing in shuttle-box. In a similar way, we studied the same learning motivated by avoidance of aversive footshock. After studying the learning driven by either positively reinforcing stimulation of the ventral tegmental area or by negatively reinforcing footshock, we integrated both reinforcers. The boosted learning observed for the combination of reward and punishment in the same session demonstrated a putatively dopamine-dependent convergent effect. Subsequently, omission procedures were employed to clarify the respective roles of appetitive and aversive reinforcers previously observed in the interaction scenario. Further clarification was achieved by comparing results from continuous reinforcement and partial reinforcement protocols. Taken together the results demonstrate that, reward and punishment operate differently during fully predicted continuous and partially predicted reinforcement conditions. The results further imply that instrumental learning mechanisms vigorously rely on dopamine signal that is associated with response. Consequently, dopamine plays discernible but important roles in both reward seeking and pain avoidance.

Contents

List of figures and tables

1. Introduction

1.0.1. Overview and framework

This dissertation presents the unified understanding gained over the years on *"the action of appetitive reinforcers (electrical stimulation of the ventral tegmental area), aversive reinforcers (electrical footshock), and their interaction during auditory learning"*. Dopamine (DA) transmission into the efferent regions is associated not only with natural rewards such as food, water and sex but also with consumption of drugs of abuse. Dopaminergic neurons are associated with motor execution, goal-directed behaviour, working memory, associative learning especially reward processing and prediction. Impairment of DA system can cause neurological and psychiatric disorders. This thesis takes advantage of available information on the role of DA for encoding the reward and punishment and how they contribute to the motivated behaviour. In this broad field of research concerning the role of DA on learning, different concepts were proposed and have been growing. Thus, it is necessary to have a broad introduction to establish the framework for understanding how appetitive and aversive reinforcement interaction has been addressed.

Section 1.1 in the introduction, describes the basic anatomy of midbrain DA system. Due to the heterogeneous nature and wide projection of the midbrain DA system, I have summarized the essential points in this section. **Section 1.2** focuses on the role of DA for motivated behaviour and change in the striatum during learning is in order, hence the DA role can be made clear. The chapter focuses on nucleus accumbens (NAc) DA changes due to the massive projection from ventral tegmental area (VTA), though the prefrontal DA release is important as NAc DA. Special emphasis is given to the DA role in reward processing and avoidance learning. **Section 1.3** starts with the history of intracranial self-stimulation (ICSS) and explains the mechanisms of ICSS which was used extensively in my research to optimize the reward. Also, it adds insight to the self-stimulation behaviour supported by different brain systems and the underlying mechanisms. **Section 1.4**, reviews previous studies that have addressed the interaction of appetitive and aversive reinforcement. Since DA plays vital role in pain avoidance and reward seeking, I outlined the questions answered by the present study using positively reinforcing VTA stimulation as a tool. The methods part describes the details

of the experiments and explains the training procedures for different groups and the experimental manipulations.

More specifically, the methods part explains the methodological details of the same instrumental behaviour driven by appetitive and aversive reinforcer using brain stimulation reward (VTA stimulation) as appetitive and footshock (FS) as aversive stimuli. The result part describes mainly the analysis of conditioned response rate and latency in different experimental conditions. Also, the follow-up experiments focusing on the nature of their interaction during continuous and partial reinforcement procedures are described. In light of the data presented, the role of dopamine in opponent processes and the possible brain systems underlying the integration of these processes are discussed. Finally, from the knowledge I gained, I have added insights about the questions I want to explore in the future.

1.1. Midbrain dopamine system-anatomy and connections

Dopaminergic neurons from the Retrorubral Field (A8), Substantia Nigra (A9), Ventral Tegmental Area (A10) provide a major ascending pathway to various brain regions. These structures are located in the ventral mesencephalon. The long axon DAergic neurons develop embryologically from a single cell group. The ascending projection fields from A8, A9 and A10 towards the various forebrain targets are overlapping each other. Depending on the target structures, different dopamine systems (i.e. mesocortical, mesolimbic and nigrostriatal) play different roles in behavioural activation.
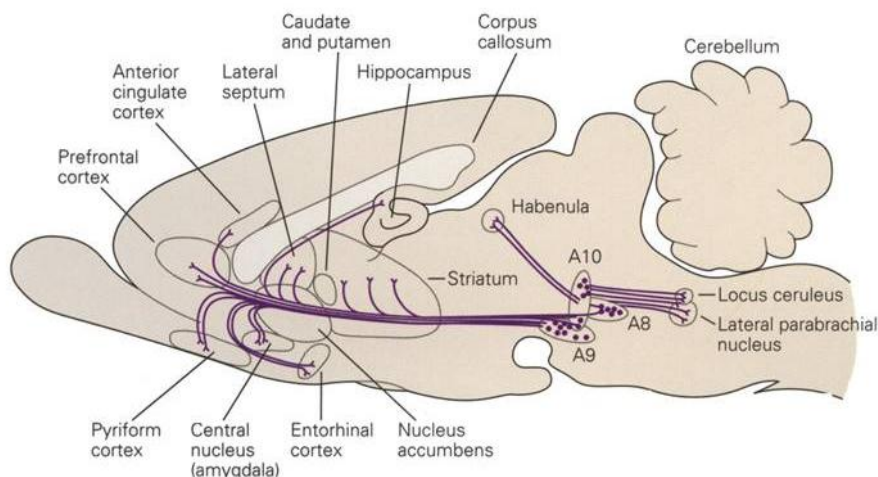
Figure 1. <u>The ascending pathways of dopaminergic system.</u> *Adapted from 'Principles of neuroscience' by Kandell, Schwartz and Jessell (2000).*

1.1.1. Ventral tegmental area (VTA)

Tsai (1925) has coined the terms "nucleus tegmenti ventralis" or "ventral tegmental nucleus". They derived from his work on opossum brain. Subsequent studies in mammals led to the term VTA-"ventral tegmental area of Tsai" which was introduced by Nauta (1958). Due to the heterogenous cytoarchitectonic nature of the region and the lack of precise boundaries, the term ventral tegmental area seems to be more appropriate.

Ventral tegmental area (VTA) and its dopaminergic projections are referred to as mesocortico-limbic system. The term "mesolimbic" is used to designate the fibers which terminate in the ventral striatum (nucleus accumbens and part of the olfactory tubercle) and

the other limbic structures such as septum, amygdala and hippocampus. While the term "mesocortical" refers to the fibers that run further rostrally to the prefrontal cortex (PFC) (Oades and Halliday 1987).
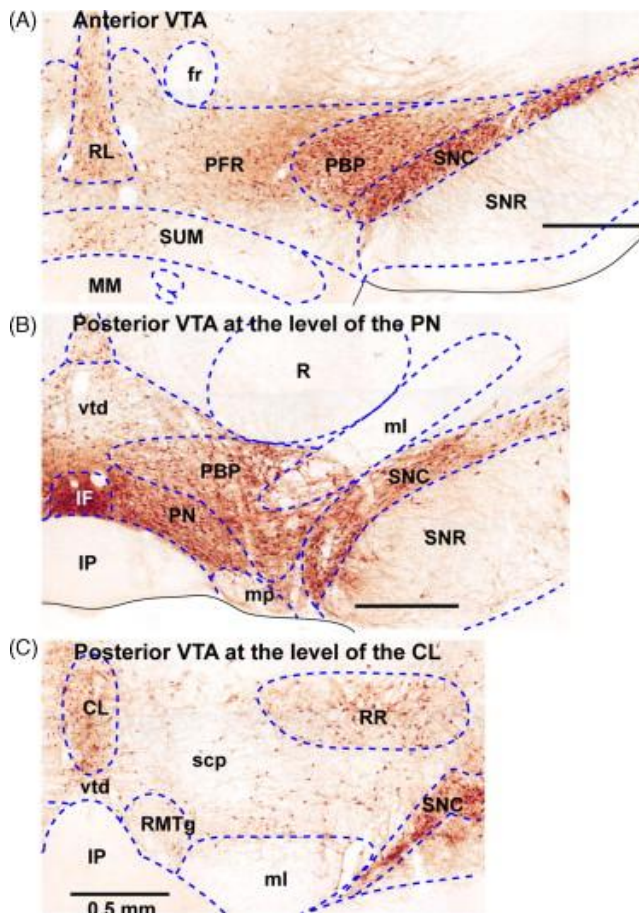
Figure 2. <u>Cytoarchitechtonic features of the VTA.</u> *Three coronal sections stained with tyrosine hydroxylase at the level of VTA showing the different cytoarchitectonic features of VTA and the distribution of dopaminergic neurons. Abbreviations: CL-caudal linear nucleus; fr-fasciculus retroflexus; IF-interfasciculus nucleus; IP-interpeduccular nucleus; ml-medial lemniscus; PBP-parabrachial pigmented area; PFR-parafasciculus retroflexus area; PN-paranigral nucleus; R-red nucleus; RL-rostral linear nucleus raphe; RR-retrorubral nucleus; scp-superior cerebellar peduncle; SNc-substantia nigra pars compacta; SNr-substantia nigra pars reticulate; SUM-supramammiliary nucleus; vtd-ventral tegmental decusssation. Adapted from Ikemato (2010).*

Based on the projections to the ventral striatum, the VTA has been divided into a caudomedial and a lateral part. The caudo medial part of the VTA sends massive DAergic projections to the ventromedial striatum including NAc shell, and the lateral part to the ventrolateral striatum including NAc core (Ikemoto 2007). DAergic and non-DAergic projections from VTA are arranged topographically along the medio-lateral lines. The medial part of the VTA project to the medial prefrontal, orbito prefrontal, cingulate and perirhinal cortex (mesocortical DA system). VTA consists of several nuclei. Classical VTA consists of paranigral, parainterfascicular, and parabrachial pigmented nuclei. However, this view has been changed (Paxinos and Watson 2004). Currently VTA consists of many sub-nuclei

(Fig.2). They are, the parafasciculus retroflexus (PFR), the parabrachial pigmented (PBP), the parainterfascicular (PIF), the paranigral (PN), the interfascicular (IF), as well as the rostral linear (RLi) and caudal linear (CLi) subnuclei (Oades and Halliday 1987, Paxinos and Watson 2004, Ikemato 2007). PBP, PIF and PN consist of DAergic cell body rich zones (Nair Roberts et al 2008). The cytoarchitechtonic studies suggest that, 60-65% of the cells in the VTA are DAergic, while 32-35% of the cells are GABAergic and 2-3% are glutamatergic. Apart from projecting to nucleus accumbens (NAc), medial prefrontal cortex (mPFC) and somatosensory cortex, the glutamatergic projections from rostro-medial VTA, make local synaptic contacts with the DAergic and non DAergic VTA neurons (Dobi et al 2010). GABAergic projections to NAc and PFC are also notable.

VTA also receives projections from many brain systems. Among them are habenula, ventral pallidum (VP), serotonergic raphae nuclei (SRN), laterodorsal tegmental nucleus (LDTg), rostromedial tegmental nucleus (RMTg) probably inhibits the VTA in many different ways. The ventral pallidum (VP) receives GABAergic inputs from NAc, in turn it sends GABAergic input to the VTA (Wu et al., 1996). Both, VTA and SNc receive afferent projections from lateral habenula (LHb) and limbic striatum (Oades and Halliday 1987; Joel and Weiner 2000). Lateral habenula inhibits VTA and SNc directly by gltamatergic projections or indirectly through serotonergic raphe nuclei (SRN) and GABAergic projection system called rostromedial tegmental nucleus (RMTg) (Jhou et al., 2009, Ikemoto 2010, Brinschwitz etal., 2010). The hindbrain region pedunculopontine tegmental nucleus (PPTg), sends glutamatergic and cholinergic projections to dopamine cell bodies of the VTA. Also, laterodorsal tegmental nucleus (LDTg) sends glutamatergic projections to VTA. The tail region GABAergic RMTg receives more input from LHb and project heavily to midbrain DAergic neurons. Due to the current status of the VTA anatomy, previously defined some posterior part of the VTA (called retro VTA or posterior tail of the VTA or caudal pole of the VTA) should be considered as the rostral tip of the RMTg (Jhou et al., 2009).

From the anatomical studies described above, it is clear that VTA is a heterogenous brain system. The present study did not aim at assessing the potentially different contributions of individual nuclei within the VTA. The electrodes were aimed towards the PBP (See methods 2.2).

1.1.2. Substantia nigra pars compacta (SNc) and Substantia nigra pars reticulata (SNr)

Substantia nigra (SN) was first described in the 18th century followed by its compact and reticular divisions in 1888. It was in 1925, that the ventral tegmental area (VTA) was suggested to be a distinct nuclear entity. The SN contains neurons whose axons project to the caudate nucleus and putamen. Degeneration of dopaminergic neurons that connect the substantia nigra pars compacta (SNc) with the caudate nucleus causes Parkinson's disease. The name "reticulata" derived due to its reticulated appearance and "compacta" because of its densely backed cells. SNc is divided into two regions. The medial part and the loosely compacted lateral part called SN pars lateralis (SNl). SNc comprise of two layers or tiers namely dorsal and ventral tier. Ventral tier is also referred as "densocellular layer".

Neurons from SNc preferentially projects to caudate and putamen (called dorsal striatum) known as nigrostriatal projection, while the more medial part of the DAergic system ie, VTA sends massive projections to ventral striatum (nucleus accumbens and part of the olfactory tubercle). The nigrostriatal pathway also courses through the lateral hypothalamus without synapsing. DAergic axons from the SNc innervates the SNr. SNc dopaminergic neurons receives GABAergic inputs from striatum and globus pallidus. Medium spiny neurons in the caudate and putamen receive convergent inputs from cortical pyramidal neurons, DAergic neurons from SNc. The output of the medium spiny neurons targets the globus pallidus and substantia reticulata. So, the firing patterns of the nigral DAergic neurons are modulated by GABAergic inputs from striatum, globus pallidus and from reticulata. SNr receive abundant afferrences from the striatum and pallidum. Subthalamic nucleus sends glutamatergic projections to SNr. GABAergic pars reticulata output neurons project to thalamus and /or superior colliculus.

Together, midbrain DA system constitutes a complicated network because of its afferent and efferent connections with various brain systems. Their contributions to the behaviour are only starting to emerge. In rat and primate PFC, the DAergic innervation differs. It shows the layer-specific patterns of innervation. In the rat, the deep layers V-VI receive the strongest DA input while in primates the upper layers I-III are at least as densely innervated (Goldman-Rakic et al., 1992). While SNc mainly consists of DAergic neurons, the neurons in VTA are heterogenous. Overall, in primates and rats, DA neurons in dorsolateral SNc projects to dorsal striatum while ventromedial SNc and VTA projects to ventral striatum. Due to its

projection pattern to dorsal vs ventral striatum, some speculate that SNc contributes to stimulus-response learning while VTA regulates the information about reward and motivation, hence playing a vital role in addiction.

1.2. Dopamine – role in motivation and reinforcement learning

Many hypotheses have been developed over the years concerning the role of dopamine in motivated behaviour. In the following sections, I explain some influential hypotheses and the current state.

1.2.1. Wanting vs liking

For many years, it was believed that dopamine is responsible for the subjective pleasures that are inherently associated with natural rewards. The role of dopamine for pleasure (anhedonia hypothesis), arose from studies which examined the role of DA in addiction or reward. Later the notion was discarded. The current consensus on the role of dopamine for incentive motivation was proposed by Robinson and Berridge (1998). Incentive motivation explained the need of dopamine for the procurement of reward or reward seeking. Selective destruction of DAergic neurons of VTA inhibits the tendency to working for food and cocaine seeking behaviour. Animals received DA receptor antagonists failed to learn the appetitive instrumental conditioning tasks. Usage of low doses of DA antagonists affects the reinforcement strength without affecting locomotor output. DA blocking agent such as pimozide treated animals failed to inject amphetamine into their veins (Wise et al., 1980; 2004). DA antagonists blocked the development of conditioned place preference by food or drugs of abuse. Mice born with the inability to synthesize DA ('DA deficient': tyrosine hydroxylase gene was inactivated) starve to death rather than eat the readily available food and water, However, they still have the intact tendency to eat the food and water if delivered inside their mouth which is similar to the DA-depleted rats (Szczypka et al., 1999). Also, DA deficient mice prefer more rewarding sweet solutions than unpalatable water (Zhou et al., 2003; Szczypka et al., 1999; Cannon and Palmiter 2003). Hyperdopaminergic mutant mice (DA transporter was knocked out, so synaptic DA level was elevated) showed higher incentive motivation to learn the runway task because of wanting for sucrose reward but lacks to elicit orofacial liking (Pecina et al., 2003). These behavioural phenotyping studies on genetically altered mouse strains provided compelling evidence that DA is necessary to search, seek out the reward ("wanting") rather than to consume or eat ("liking") (Salamone et al., 1994; Salamone and Correa 2002).

The role of DA in sexual motivation falls along the lines of incentive motivation. Copulatory cues progressively increased the NAc DA. For example, NAc DA release in male rat was increased vigorously during the anticipatory phase in which the receptive female was covered with a wire mesh cage. The DA level in the male NAc reached its peak after the removal of wire mesh and at the beginning of copulation, and it dropped down during the copulation especially after repeated bouts with her (Pfaus et al., 1990; Wensksten et al., 1993; Robinson et al., 2001). However, again the introduction of new female led to the renewal of DA release and copulation (Coolidge effect) (Fiorino 1997). This DA dynamics during the stages of copulation is very specific to the NAc but was absent in dorsal striatum suggesting the role VTA plays. So, the DA dynamics in NAc are consistent with Robinson's suggestion that DA mediates "wanting" of salient rewards, rather than the subjective hedonic experience ie "liking". Instead of mediating the hedonic aspect of appetitive reward such as food, water and sex, neuronal activity in the DA system represents the "wanting" of reward after learning (Berridge and Robinson 1998; Roitman et al., 2004). Therefore, the role of DA in incentive motivation was dependent on the reinforcement history derived from the organism's liking (for, e.g., taste) of primary reinforcer and its evoked DA release.

Furthermore, DA -depleted animals take the easy route to obtain small reward rather than obtain big reward which requires effort such as climbing over the barrier. Also, DA blockade led to choosing the small immediate reward over long delayed rewards (Cardinal et al., 2000). This impulsive choice can be induced by lesions of the NAc core (Cardinal et al., 2001).

1.2.2. Dopamine and reward

As reviewed above, dopamine is necessary for appetitive motivational processes. Here, I like to emphasize the neurophysiological findings obtained from midbrain DAergic neurons during reward dependent learning.

The midbrain DAergic neurons are spontaneously active. They are even awake (spontaneously active) during REM sleep. DAergic neurons show two predominant patterns of firing called tonic and phasic. Tonic activity consists of a regular spike firing pattern of ~1-6 Hz (Grace and Bunny 1984a). This pattern maintains the basal extracellular levels of DA in the striatum and prefrontal cortex. Phasic activation of DAergic neurons increases their firing rate upto 24 HZ. This burst firing of DA neurons leads to the transient increases in

extracellular DA concentrations in the efferent brain regions. This change also evoke different range of effects on efferent neurons by altering the levels of DA (Grace and Bunny 1984a; 1984b).

The slow, steady (tonic) firing in the resting animal changes to brief, abrupt (phasic) in discharge for meaningful stimuli such as attention generating sensory stimuli for eg, loud clicks or bright flashes of light (Horvitz et al., 1997; Horvitz 2000), or biologically more efficient stimuli such as reward. However, the phasic mode evoked by neutral sensory stimuli habituates quickly, when the subject experience the repetition (Ljungberg et al., 1992). In the following paragraphs, I narrate some of the experimental evidences obtained by Wolfram Schultz and his colleagues.

1.2.2.1. *Reward prediction error*

During associative learning, the phasic mode elicited by the unexpected reward gradually vanishes after many trials and shifted to the predictive stimulus onset after learning. However, slow or cessation of activity observed during the time when the expected reward was omitted (Schultz 1998).

The shift from the nature of reward coding to predictive stimulus encoding led to the proposed reward-prediction error hypothesis. That means, the phasic DAergic signals code the difference between the expected reward and the experienced or obtained reward (based on reinforcer history). From his influential work on primates, Wolfram Schultz suggested that reward-prediction error by mid brain dopamine system is acting as a teaching signal (Teacher (DA) – Student (Striatum)). This error signal modulates the other basal ganglia neurons for appropriate response in the future. Electrophysiological evidences originated from his work elucidated the midbrain DAergic neuron's role in reward processing during learning. Subsequently, reinforcement learning modelers placed the DA system's action under the temporal difference models.

Similar studies have been undertaken by other labs which measured the phasic DA signal in the NAc by voltammetry measurements. In instrumental conditioning tasks, phasic DA signal in NAc elicited by the natural reward/brain reward shifted to reward predictive stimuli during the course of learning (Stuber et al., 2008; Owesson-White et al., 2008). This phasic DA signal for reward predictors suggests the "cue-induced wanting" nature of the system.
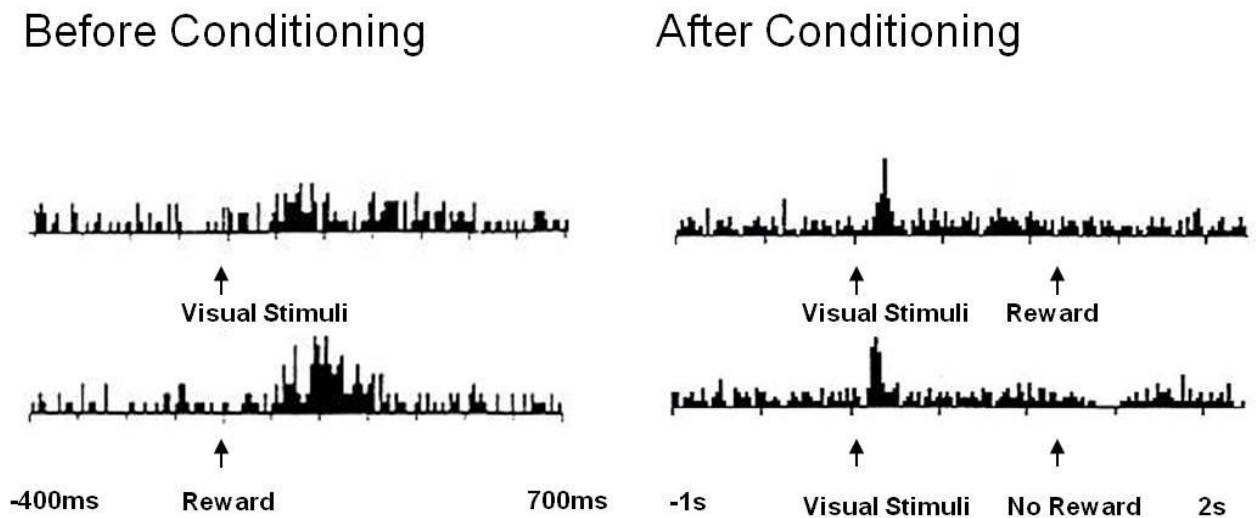
Figure 3. Phasic burst firing of a dopaminergic neuron before and after conditioning. *Before conditioning, light evoked responses were observed but unexpected reward elicitted more bursting. After conditioning, CS onset evokes burst, while no bursting was evident during the time of reward delivery. However, when the expectation was violated by reward omission, a decrease in baseline firing (inhibition) occurs at the time of predicted reward. Adapted from Schultz, Apicella and Ljungberg (1993).*

1.2.2.2. *Reward prediction error signal: alternative arguments*

Alternative arguments remain over the phasic shift of DAergic neurons for the conditioned stimulus and the proposed reward prediction error hypothesis. The prediction error hypothesis was originally derived from recordings of midbrain DAergic neurons in behaving monkeys. In these experiments, visual stimuli were employed. Peter Redgrave and his colleagues argued that this short latency (70-100 ms after the stimulus onset), short duration (100-200 ms) burst of phasic response as proposed by Schultz is too short to encode the error in reward prediction. These authors provided alternative interpretations which stems from the anatomical and electrophysiological experiments (reviewed in Redgrave et al., 1999; 2006; 2008). The following points evolved from their experiments, deserve attention: (1) A direct tectonigral projection which connects the deep layers of the superior colliculus to the caudal pole of the SNc was reported in rats, cats and monkeys (Comoli et al., 2003; McHaffie et al., 2006; May et al., 2009). (2) The latency of phasic DA response obtained from the SNc lies in

between the early sensory response and pre saccadic motor response of the superior colliculus. (3) In rats, the visually evoked activity in SNc can be seen with the absence of visual cortex but not without the visual layers of superior colliculus (Dommett et al., 2005). From the latency and duration of the DAergic phasic signal they refer the phasic dopamine signal as "sensory prediction error" instead of "reward prediction error". They proposed, this might be a signal for pre attentive sensory processing which translate the motivation to action (behavioural 'Go' signal) and prepare the organism for biologically relevant events for better learning.

1.2.2.3. Explicit reflection of reward value by dopaminergic neuron

The phasic dopaminergic signals also reflect the information about the reward value such as magnitude, probability and delay. Using three different stimuli, responses of DAergic neurons for different magnitude of juice rewards were investigated. It was found that the neural coding by DAergic neurons represented the associated reward magnitude by coding sensitivity towards stimuli which signalled the larger reward (Tobler et al., 2005). Subsequent investigation on reward probability coding gave clear insights. In this experiment, five different visual stimuli associated with different probability of juice reward occurrences were used. They were 0 (no reward), 0.25, 0.50, 0.75 and 1 (continuous or certain reward). Interestingly, midbrain DAergic neurons encoded the probability and uncertainty which are inherently linked to each other. The predictive stimulus encoding by DA system showed positive correlation with the probability of reward occurrences. With the maximum uncertainty (ie., P=0.5), the DAergic neurons showed gradual increase and sustained firing (tonic and phasic) during the visual cue that lead to reward delivery or no reward. This uncertainty coding was absent during certain situations (ie., P=0 or P=1) and higher than what they have seen when the P=0.25 or P=0.75 (Fiorillo et al., 2003). In another study, different stimuli associated with different delay of reward ranging from 2 s to 16 s were investigated. DAergic neuron activity showed more sensitivity towards the stimuli which signalled the smaller delay of reward (2 s). But, with the longer delay (16 s) condition, the activity was more sensitive after the reward, when compared to CS onset probably reflects the temporal uncertainty (Kobayashi and Schultz 2008).

The above mentioned electrophysiological data from primates indicated that after learning DAergic neurons implicitly reflect some aspect of reward information namely magnitude,

probability and delay (Tobler et al., 2005; Fiorillo et al., 2003; Kobayashi and Schultz 2008). Indeed, these three are not independent factors to constitute the reward value. It is necessary to study the response of DAergic neuron in tasks such as temporal discounting which effectively integrates the factors such as delay and magnitude to understand their role in decision. Probably, midbrain DAergic neurons pass the reward related information to nucleus accumbens (NAc) during the time when no explicit motor responses were evident, and due to its convergent input from other brain systems NAc is in the center position to mediate the reinforcement learning processes.

In rodents, VTA DAergic neurons response was investigated during instructed and free choice trials between an immediate small reward vs a delayed larger reward. When the delay time was increased, the cue evoked phasic DA signal decreased slowly. The DAergic neurons preferred immediate reward rather than delayed. In free choice trials, though the decision led to small reward or no reward, the neuronal response preferred the potential value of the cue (Roesch et al., 2007). The role of DAergic neuron's role in reward prediction was explored for few decades. In saccadic decision making task, the DAergic neuron response for the informative cue was investigated. In this task, the choice of target led to the presentation of informative cue. Interestingly, the informative cue which successfully predicted the reward excited the DAergic neuron's firing while random cue inhibited it. The DAergic neuron which signaled the expectation of reward also signaled preference for the expectation of advance information about reward. This suggests the seeking nature of the DA system (Bromberg Martin and Hikosaka 2009).
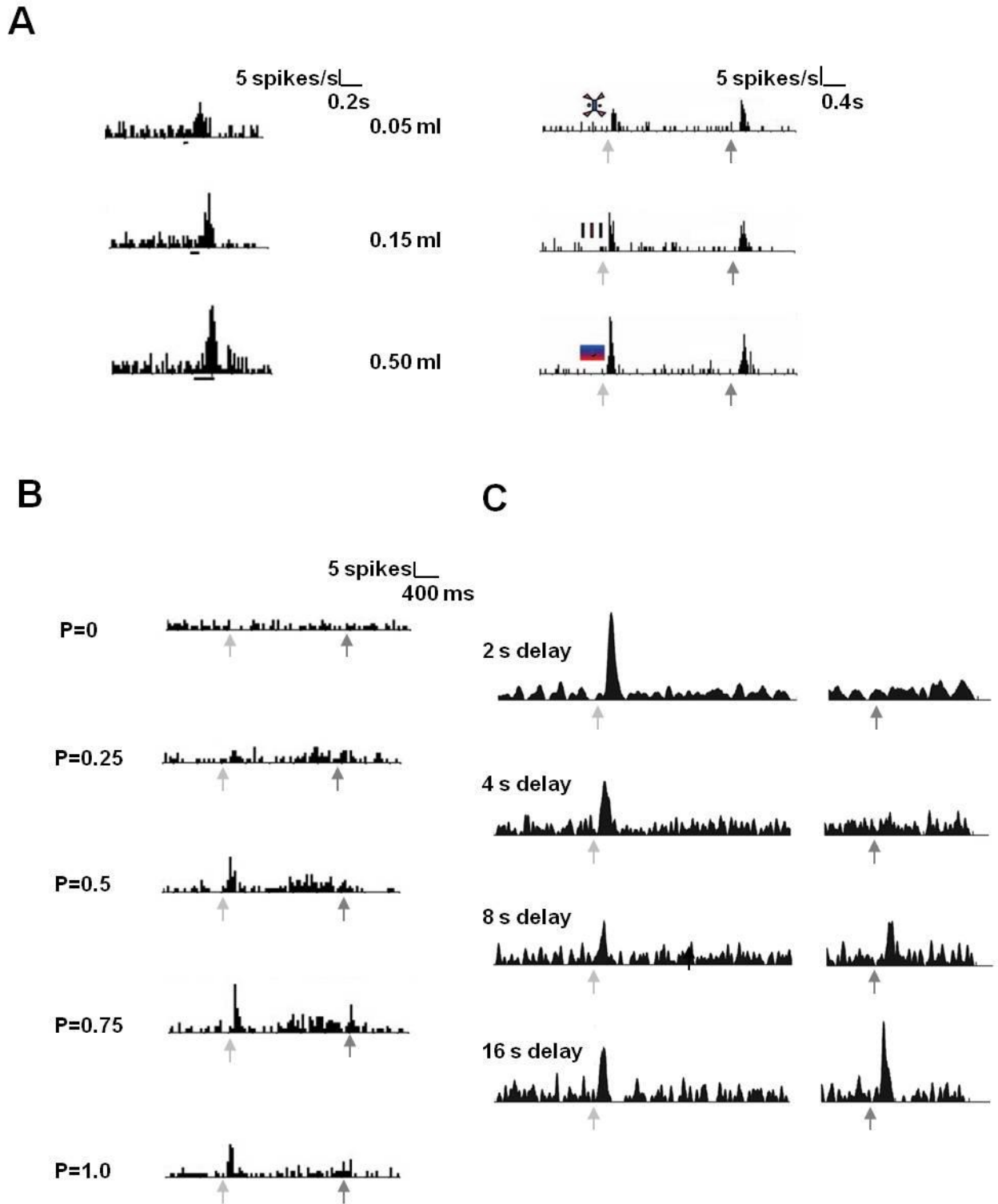
Figure 4. <u>Phasic burst firing of dopaminergic neuron in response to reward magnitude, probability and delay.</u> *Light grey arrow indicates stimulus onset and dark grey indicates reward onset or stimulus offset. (A) Left panel: A single dopamine neuron in response to different magnitude of reward in the absence of any predictive stimuli. Right panel: Single neuron activity conditioned to three different stimuli associated with different magnitude of reward. The time of reward delivery*

indicated by the thick bars below the raster plot. The frequency of firing after stimulus onset, increased with their reward value. (B) Dopaminergic neuron response to conditioned stimuli predicted different propability (P) ranging from 0.0 to 1.0 of juice reward. For stimulus predicted low probability (P=0.25) or omission of reward (at P=0), dopaminergic neuron did not show a significant level of phasic burst. When the uncertainty is high (at P=0.5), the dopaminergic neuron showed gradual increase and sustained firing during the whole CS duration. (C) Activity of a dopamine neuron with variable delay ranging from 2 s to 16 s. The left panel illustrates the single dopamine neuron activity which was aligned to stimulus and the right panel to reward onsets for each experimental condition. For longer delay, evoked response by the conditioned predictor was smaller (Figures were modified from Tobler, Fiorilla and Schultz (2005); Fiorilla, Tobler and Schultz (2003); Kobayashi and Schultz (2008)).

1.2.3. Dopamine and punishment

A large number of studies have suggested that DA is released in response to appetitive reinforcers and plays a vital role in reinforcement learning based on reward (Wise 2004; Schultz and Dickinson 2000). However, DA is also released in response to aversive unconditioned stimuli, CSs that predict them and also for other salient stimuli such as novel stimuli (Salamone 1994; Young et al., 1998; Horvitz 2000; Stark et al., 2001; Young 2004) which would be consistent with a more general motivational role. Also, rewarding and aversive lateral hypothalamic stimulation evoked NAc dopamine release (Hernandez et al., 2006; Rada et al., 1998). In many studies, slow elevation of extracellular DA level in NAc and PFC was reported for stressful stimuli such as inescapable intensive footshock (>0.55 mA), tailshock and tailpinch. The previous postmortem studies which measured the DA metabolites from efferent regions also reported similar finding. DA depletion in PFC increased the stress (mild footshock) evoked DA efflux in NAc shell (King et al., 1997). Within the NAc, the shell and core showed different patterns of extracellular DA level for offset of aversive stimuli. In one study, immediately after the offset of 20 min presentation of footshock (0.35 mA for 200 ms/s), DA level was elevated in NAc shell while no change was observed in NAc core (Kalivas and Duffy 1995).

However, the DAergic neurons response for aversive stimuli sang a different tune. The general suggestion is that increase in firing occurs for appetitive stimuli but not for aversive. Aversive stimuli gradually change the firing rate of DAergic neurons or local modulation of

extracellularly measured DA release in target structures dissociates dopamine release from firing was possible (reviewed by Salamone 1994; Horvitz 2000).

Even though it was clear that, omission of reward and the conditioned inhibitor (tone which signalled the omission of reward) decreased the firing rate of DAergic neurons (Tobler et al., 2003), little is known about their responses to aversive stimuli. Peripheral nociceptive stimuli inhibited most of the 78% of presumed DAergic neurons in the SNc of rats (Gao et al., 1990; Tsai et al., 1980). Also intense noxious stimuli inhibited the firing rate of 51% of presumed nigrostriatal DAergic neurons (Schultz and Romo 1987). The heterogenous VTA DAergic neurons showed mixed responses to conditioned aversive stimuli during differential fear conditioning task in awake rabbits (Guarraci and Kapp 1999). Wolfram Schultz and his colleagues reported that, VTA/SNc neurons preferably increased the rate of firing by appetitive juice reward rather than mild air puff to the hand or hypertonic saline to the mouth (Mironewicz and Schultz 1996). Application of 2-3 s tailpinch induced endogenous DA in the striatum was studied with the combination of single unit and voltammetric recording. It increased the frequency of firing in most of the recorded neurons, but the phasic DA signal was not elevated (Williams and Miller 1990).

Good evidence for a response of DAergic neurons to aversive stimuli was recently obtained. In anesthetized rats, Ungless et al. (2004) addressed the issue by showing the response of VTA neurons for aversive footpinch. After recording the extracellular unit activity from VTA neurons subsequent labeling was done by juxtacellular injection of Neurobiotin. Interestingly, those neurons which showed inhibitory responses (reduction in firing rate and bursting activity) were TH-positive therefore they are DAergic (Ungless et al., 2004).
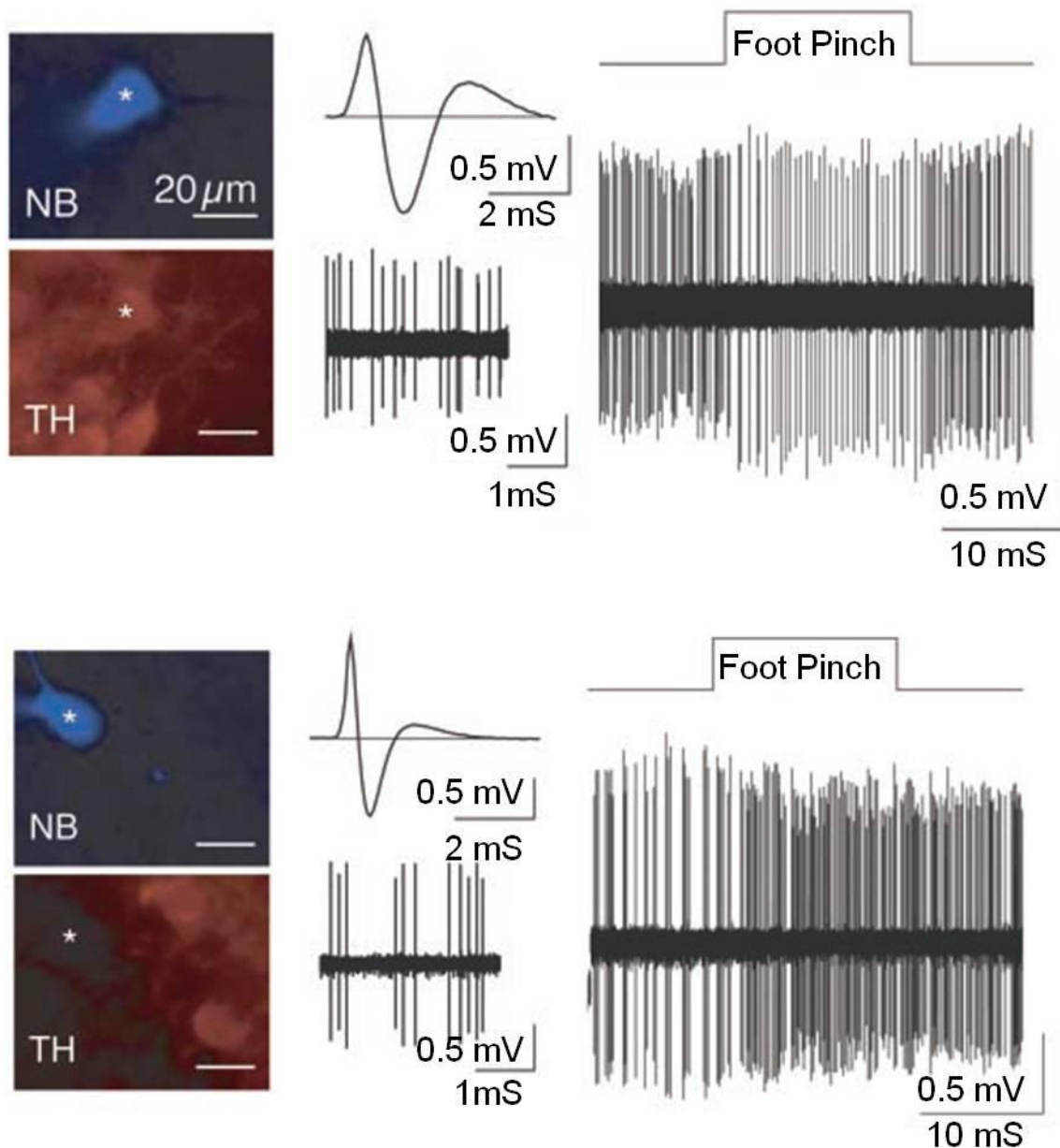
Figure 5. <u>Aversive stimuli inhibit the dopaminergic neurons.</u> *The excited neurons are non DAergic. The upper panel shows a neurobiotin (NB) labeled tyrosine hydroxylase (TH)-positive neuron and, therefore dopaminergic. The broad action potential (upper trace) and slow firing (lower trace) of the extracellularly recorded properties of identified DAergic neurons. Aversive footpinch decreased the firing rate of DAergic neuron.*

The lower panel shows a neurobiotin (NB) labeled tyrosine hydroxylase–negative neuron and, therefore non-DAergic. Examples of broad action potential (upper trace) and slow firing rate (lower trace) is shown by this non dopaminergic neuron. Aversive stimuli increased the firing rate of this neuron. Adapted from Ungless et al., 2004.

However, their follow up study argues that most of the previous studies that have measured the DAergic neuron response to aversive stimuli aimed the parabrachial pigmented nucleus (PBP) of the dorsorostral VTA. The ventromedial DAergic neurons from the paranigral nucleus (PN) were mostly neglected. In a recent study, these authors found DAergic neurons from the dorsal VTA (PBP) to be inhibited by aversive footshock while the ventral VTA (PN) DAergic neurons were excited. Therefore, one can hypothesize that the DA release evoked by aversive stimuli might be due to the action of DAergic neurons in the PN. It should be noted that PN sends projections to NAc shell and mPFC, both regions showed DA release for aversive stimuli (Brischoux et al., 2009). So it is evident that even within the VTA, two functionally distinct DA systems are available.
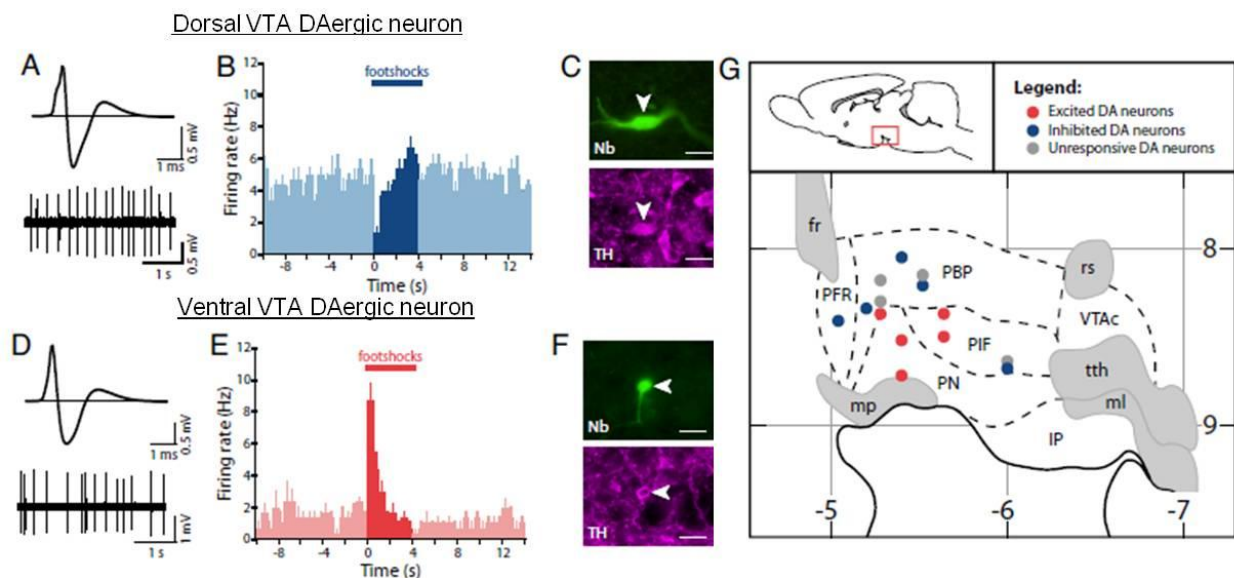
Figure 6. Two functionally distinct dopaminergic neurons in the VTA. *DAergic neurons in the dorsal VTA are inhibited by noxious stimuli whereas ventral VTA DAergic neurons excited. (A and D) shows the averaged extracellular waveform and baseline firing activity from a neuron. (B and E) shows the Inhibitory and excitatory response to aversive footshock. (C and F) shows the immunohistochemically identified DAergic neuron. (G) A parasagittal view of the VTA showing the anatomical segregation of functional subgroups. fr-fasiculus retroflexus; ml-medial lemniscus; mp-mammillary peduncle; rs-rubrospinal tract; tth-trigeminothalamic tract; IP-interpeducncular nucleus; PBP-parabrachial pigmented nucleus; PFR-parafasciculus retroflexus area; PIF-parainterfascicular nucleus; PN-paranigral nucleus and VTAc-ventral tegmental area caudal. Adapted from Brischoux et al., 2009.*

Recent, electrophysiological studies in primates suggests that positive and negative motivational states are conveyed by two different subsets of DAergic neurons situated in the VTA+SNc medial part and lateral part of the SNc.
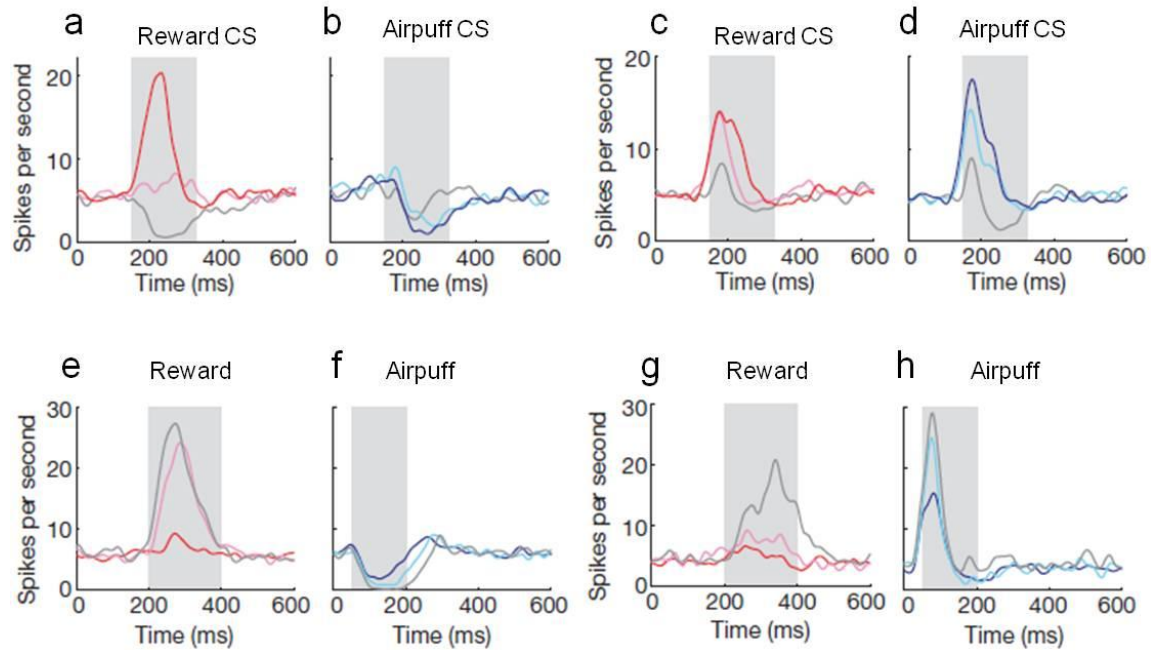
Figure 7. Distinct dopaminergic neurons convey positive and negative motivational signals.

Averaged activity of airpuff conditioned stimulus (ACS) inhibited (a and b) and excited neurons (c and d). Spike density functions are shown for 100% reward CS (red), 50% reward CS (pink), and 0% reward CS (grey) in the appetitive block (a and c); 100% airpuff CS (dark blue), 50% airpuff CS (light blue), and 0% airpuff CS (grey) in the aversive block (b and d).

Averaged activity of airpuff unconditioned stimulus (AUS) inhibited (e and f) and excited neurons (g and h). Spike density functions are shown for 100% reward (red), 50% reward (pink), and 0% reward (grey) in the appetitive block (e and g); 100% airpuff CS (dark blue), 50% airpuff CS (light blue), and 0% airpuff CS (grey) in the aversive block (f and h). Adapted from Matsumoto and Hikosaka 2009.

1.2.4. Role of dopamine in avoidance learning

Eventhough much information is available on the role of dopamine in avoidance learning, there is no evidence concerning the firing of DAergic neuron during active avoidance learning. Low dose of DA antagonists impaired the active avoidance responses without affecting the motor function (e.g. Cooper et al., 1973). Also, depletion of DA in the striatum by 6-hydroxydopamine (6-OHDA) injection into the substantia nigra pars compacta (SNc) led to the impairment of active avoidance responses in rats. However, the experimentally induced reductions of DA levels in the striatum also disrupted other functions like feeding, drinking and sensorimotor functions (Jackson et al., 1977; Salamone 1994). In an operant conditioning task of avoidance learning, the effect of nucleus accumbens (NAc) dopamine depletion was investigated. NAc DA depletion severely disrupted the task in which rats were trained to delay the arrival of footshock for the next 30 s, by pressing the lever for every 5 s. From this experiment, it was concluded that 6-hydroxydopamine (6-OHDA) DA depletion in NAc disrupted the lever press avoidance responses at least during the initial stage of learning (McCullough et al., 1993). The number of avoidance, but not escape was correlated with NAc DA increase (Sokolowski et al., 1994).

D2 antagonist sulpride injected into the NAc was found to inhibit shuttle-box footshock avoidance learning. At the same time, prefrontal cortex, amygdala and caudate putamen injections led to no effect (Wadenberg et al., 1990). Furthermore, active avoidance learning was disrupted by alpha-methyl-p-tyrosine and further DA injection into the NAc reversed the avoidance performance (Bracs et al., 1982). A detailed understanding about the role of DA during shuttle-box active avoidance learning originated from the studies by Holger Stark and colleagues. In auditory cortex, DA release was observed during auditory avoidance learning (Stark et al., 1997). Active avoidance learning progressed with the increase of DA release in medial prefrontal cortex (mPFC). These findings suggest the role of DA for the initial stage of avoidance learning (Stark et al., 1999; 2001). In another study, Mongolian gerbils were trained with two tones associated with footshock. For both tones, the animal needs to cross the hurdle to avoid footshock. In principle, the two tones signalled the same meaning (Go signal). After learning was well established, one tone signalled Go response, while the other signalled No Go response. Enhancement of DA release was observed in mPFC while the animal learns this new strategy (Stark et al., 2004). From this study, it was an evident that cortical DA release can generally be associated with the aversive reinforcement learning.

1.3. Reinforcing dopamine system-"seeking"

1.3.1. Intra cranial self-stimulation (ICSS) - overview

Fifty years ago, the concept of brain's reward system that reinforces pleasure seeking behaviour was first proposed. James Olds and Peter Milner discovered this phenomenon when their electrode missed their target "midbrain reticular formation". Instead of seeing the aversive reaction and further avoidance, the rat with misplaced electrode preferred the place associated with electrical brain stimulation. Then, James Old came up with the idea of providing them a lever to get the brain reward which was similar to the operant conditioning (Olds & Milner 1954). These experiments led to the finding called intracranial self-stimulation (ICSS) of the brain, in which rats pressed the lever for brain stimulation reward with reinvigorating search strategy. The appetitive nature of brain stimulation reward is more powerful than natural reinforcers such as food and water. Ever since this historic finding emerged, research on brain mechanisms of motivation, reward, and addiction have gained much attention and researchers have explored brain systems which support self-stimulation (Olds and Fobes 1981). The finding of ICSS also enabled the anatomical mapping of the reward systems and understanding the reinforcement mechanisms of the brain. They found that many sites that support ICSS situated along the course of dopaminergic innervation. It is clear that, DA plays a vital role for learning and performing the ICSS. However, other neurotransmitters also regulate the rewarding aspect of ICSS obtained from midbrain DAergic system and their terminal regions (Cheer et al., 2005). For example, animals are able to selfadminister cholinergic agonists and GABA antagonists into the VTA (Panksepp 1998; Rolls 2005).

More technical studies have been done to provide the information carried out by the stimulated axon to various brain systems. Furthermore, studies on reward properties of drugs of abuse were also followed. Classical experiments on BSR also opened new avenues for exploring the selective self activation of brain systems within appetitive states such as thirsty and hunger (Gallistel and Beagley 1971). Enormous amount of work on ICSS have been focused on the rodent medial forebrain bundle of the lateral hypothalamus, thus it was considered as a sheet of motivation for many years. Rats ran faster for brain stimulation reward when compared to food reward in maze runway test. They were able to cross the

electrified grid floor to obtain the brain reward (Olds 1958). Subsequent studies employed the brain stimulation reward in instrumental sensory conditioning.

Many studies applied medial forebrain bundle (MFB) reward in learning paradigms. From simple place conditioning to sensory conditioning paradigms, rewarding MFB stimulation was successfully integrated. Beyond the simple conditioning tasks, using electrical stimulation of somatosensory cortical (SI) and MFB as a cue and reward, rat was navigated over three-dimensional route (Talwar et al., 2002).

1.3.2. Brain systems which support self-stimulation

Intracranial self-stimulation (ICSS) can be conducted in a number of brain systems ranging from midbrain dopamine system to orbitofrontal cortex (OFC) of primates. In rodents, ICSS in certain brain regions such as lateral hypothalamus (medial forebrain bundle), mid brain DAergic system (substantia nigra pars compacta and ventral tegmental area) and nucleus accumbens provide a very strong incentive to restimulate, creating a feedback loop that reinforces the animals to stimulate again. Most of the self-stimulation supporting brain systems are lying on the ascending DAergic pathway. Electrical stimulation of midbrain DA system and lateral hypothalamus usually evoke exploratory search behaviour. In septum, amygdala, hippocampus and prefrontal cortex, the self-stimulation rates were found to be lower and behavioural activation was less evident. ICSS responses obtained from medial septal area and locus coeruleus were slow and rhythmic. Some prefrontal areas also support self-stimulation. In primates, self-stimulation can be obtained from electrodes in posterior part of the OFC (Mora et al., 1980). Hunger and thirst can modify the ICSS responses obtained from lateral hypothalamus and in primate OFC. Other types of behaviours are elicited by self-stimulation of certain brain systems. Stimulus bound feeding from lateral hypothalamus stimulation (Margulus and Stein 1962), drinking from zona inserta stimulation (Mogenson and Stevenson 1966) and periodic shivering which resembles sexual behaviour from stimulation of medial septum was reported. Eventhough ICSS can be obtained from number of DA terminal regions, reliable and more vigorous rates of self-stimulation can be obtained from MFB of the lateral hypothalamus, SNc and VTA. And altering the stimulation strength in a time dependent fashion, can lead to inverted U shaped response rate in these regions (See Figure 10).

1.3.3. Some general properties and mechanisms of ICSS

Free application of brain stimulation reward (BSR) can elicit exploratory and seeking behaviour. Rapid extinction is one of the main aspects of ICSS. However, hungry rats take more time to extinguish, suggesting the presence of common drive (Deutsch and Di Cara 1967). Single "priming dose" (free non contingent electrical pulse) can reinstate the extinguished self-stimulation behaviour. Application of priming dose can be used at the beginning of ICSS training to facilitate the learning. Hungry animals prefer to self-stimulate rather than eating food. This repetition of behaviour is persistent over hours until they are physically exhausted. Many addictive, psychoactive drugs such as from cocaine to cannabis summate with ICSS, provided the general reward potentiating effects. Parallel leftward shifts in the rate-frequency and rate-intensity curves of the ICSS can be obtained by using DA agonists, and rightward shifts with low dose of antagonists (Miliaressis et al., 1986; Wise 1996; Wise 2002). Animals deprived of REM sleep stimulated their LH vigorously with low current intensity, suggesting to view this effective behaviour from multiple directions ranging from stress, common arousal to enhanced motivation. Based on pharmacological considerations, Roy Wise suggested that DA antagonists affect the synaptic transmission of the reward signal across the DAergic synapse or synapse where DA provides a modulatory signal (Wise 1980).

Even though the research on the mechanism of ICSS is growing over the last fifty years, still the precise timing of DA release into the NAc and the how DA governs the vigorous rate of ICSS ie., seeking out and reinstate is unclear. Precise timing of DA release and the regulation of ICSS is still ongoing discussion in the extensive amount of literature. Some controversies and confusion are remaining on the mechanisms of the ICSS from medial forebrain bundle (MFB). The DA fibers from VTA are branching off at the level of MFB before ascending to the NAc. MFB lesions failed to structurally and functionally disconnect the VTA from forebrain structures in contrast to the LH lesions where lesions were made rostral to MFB. However, rostral MFB lesions increase the threshold for VTA self-stimulation (Simmons et al., 1998). ICSS of both MFB and VTA increases extracellular concentration of DA in NAc and PFC (Hernandez et al., 2006; Fibiger et al., 1987; Simmons et al., 1998). Increase in acetylcholine was observed in VTA for MFB self-stimulation. Self-stimulation of MFB was greatly reduced after the blockade of muscarinic receptors in the VTA.

Electrophysiological studies suggest ICSS of the MFB depends on the activation of the terminals of the DA afferents but not on midbrain DAergic neurons. And MFB stimulation activates the descending, non DAergic, myelinated axons which courses through the descending direction (Gallistel et al., 1981; Bielajew and Shizgal 1986). The GABAergic projection neurons from the VTA (~36%) ascend to the NAc can be activated by MFB stimulation (Steffensen et al., 2001). The unit responses in the NAc core for the noncontingent MFB stimulation were preferentially responsive to GABA (Cheer et al., 2005). Both MFB and VTA ICSS can occur without continuous DA release into the NAc (Hernandez et al., 2006; Hernandez and Shizgal 2009).

Even though phasic DA response evoked by ICSS of the MFB disappear rapidly in the NAc shell, it reappears for every stimulation train when the inter reward interval is set to 10s (Cheer et al., 2005). Measured by microdialysis techniques, the stimulation of the MFB induced elevation of DA in the NAc, remains stable even though the inter train interval (ITI) was set upto 12s. Even though decrease in ITI to 1.5s increased the peak, the DA concentration falls as stimulation continues. However, it stayed above baseline for long period (Hernandez et al., 2006). Recently, the cellular mechanisms of learning the ICSS were elucidated using SNc self-stimulation. The rate of learning the ICSS of SNc is correlated with the degree of potentiation of synapses made by cortical afferents onto medium spiny neurons of the dorsal striatum, a potentiation that needs DA receptors (Reynolds et al., 2001). In the following paragraph, I focus on VTA self-stimulation induced changes and their mechanism since I used VTA stimulation in my studies.

1.3.4. Electrical stimulation by the experimenter or Self-stimulation of VTA

Electrical stimulation of VTA evokes exploratory search behaviour. The vigorous self-stimulation obtained from VTA can continue over hours. Earlier metabolic mapping studies which compared the local rates of cerebral glucose utilization between response contingent (ICSS) and non contingent (experimentor delivered) stimulation of VTA found higher level of glucose utilization in the motor cortex of ICSS group. Also higher level of glucose utilization was observed in the NAc and mPFC of ICSS group (Porrino et al., 1984). Studies using fast scanning cyclic voltammetry suggest that phasic DA transients can be reliably evoked in NAc by certain stimulation parameter (24 biphasic pulses, 60 Hz, 125–150 µA, 2 ms per phase) to VTA (Day et al., 2007; Owesson-White et al., 2008). DA release by VTA

stimulation attenuated the synaptically evoked glutamatergic inputs to the NAc from other regions such as hippocampus or amygdala. Also, 20 Hz/2 s of VTA stimulation led small rise of DA in the prefrontal cortex (PFC), eliminated within a few seconds. Moreover, VTA stimulation elicited short-losting inhibitory effect on the spontaneous firing of pyramidal neurons in the PFC (Lavin et al., 2005). Temporally precise pairing of VTA stimulation with auditory stimuli modified the cortical representation of the paired tone. Repeated stimulation of VTA (10 biphasic pulses, 100 Hz, 100-200 µA, 0.1 ms per phase) between the 4 kHz and 9 kHz, increased the receptive field of the preceding 4 kHz tone but reduced the 9 kHz trailing tone (Bao et al., 2001; Bao et al., 2003).

Learning the ICSS of VTA depends on the phasic DA release in the NAc (Garris et al., 1999). Measured by micro dialysis and voltammetry, electrical stimulation of the VTA increased the extracellular DA level and evoked phasic transients in the NAc by increasing the population activity and bursting (Fibiger et al., 1987; Fiorino 1993; Garris et al., 1999; Phillips et al., 2003). Also, DA level in other efferent systems like olfactory tubercle and PFC are altered by VTA stimulation. The VTA self-stimulation rate was substantially reduced after DA receptor antagonist spiroperidol injections or 6-hydroxydopamine lesions into the ipsilateral NAc (Mogenson et al., 1979; Fibiger et al., 1987). Recent studies suggested that the rise of DA level in the NAc dialysate sample remained elevated for 2 hours of ICSS of the VTA (Hernandez and Shizgal 2009). Repeated stimulation (ICSS or experimenter delivered) of the MFB or VTA produced prolonged elevation of the DA tone in the NAc (Hernandez et al., 2006; Hernandez and Shizgal 2009). Even though an earlier study measured the phasic DA change by FSCV, and reported the dissociation between the increased DAergic transmission and continued ICSS of the VTA, recent dialysis studies clarified the clear correlation between the elevated DA and ICSS for atleast 2 hours (Garris et al., 1999; Hernandez and Shizgal 2009). Owesson-White et al.,(2008) trained rats to lever press after the cue onset. The lever press was followed by VTA stimulation delivered with different time intervals. Interestingly, at the beginning of learning, DA transients were observed for the electrical stimulation. But, when the trial progressed, the cue evoked DA in NAc increased while electrically evoked DA decreased. During extinction, the electrical stimulation evoked DA was disappeared and this was followed by decrease in amplitude of cue-evoked DA (Owesson-White et al., 2008). The above said experiments matched the DA neuron responses for reward prediction in which the shift in increase of firing from reward onset to cue onset were reported (Owesson-White et

al., 2008; Schultz et al., 1997; Schultz 1998). It seems to be the rapid DA signaling detected by FSCV in the NAc while learning the ICSS is dynamic. ICSS of the VTA induced changes in firing patterns was reduced for D1 receptor blockade but not D2. Also D1 receptor blockade into the NAc shell abolished lever pressing for ICSS (Cheer et al., 2007).

ICSS-supporting systems suggest that functionally connected reward centers are operating like networks both serially and in parallel. The imminent of this chapter is to emphasize that even though ICSS behaviour is effectively elicited along the DAergic projection regions, other systems also play a role (cf. Panksepp 1998).

In the next chapter, we will see the summarized classical experiments on appetitive and aversive reinforcement and how we utilized the application of brain stimulation reward during learning experiments to address the interaction of appetitive and aversive reinforcers.

1.4. Appetitive and aversive reinforcement

According to Skinner, events that strengthen or increase the likelihood of preceding responses are called positive reinforcers, and events whose removal strengthens preceding responses are called negative reinforcers (Skinner 1938). Thorndike called those stimuli as "satisfiers" and "annoyers". Generally, we refer to these two types of reinforcers as "reward" and "punishment". Based on the affective attributes that determine the reinforcing nature of unconditioned stimulus, we can give the operational definition and classify them as appetitive and aversive reinforcers.

Conditioning involves the association of neutral stimuli with appetitive or aversive reinforcers. Animals direct their behaviours, in both natural and laboratory situations (e.g. instrumental conditioning experiments), in such a way as to obtain appetitive reinforcers ("rewards") and avoid aversive reinforcers ("punishments"). In most animal conditioning experiments, behavioural measures of conditioning and of brain systems have been studied with one type of reinforcer (appetitive or aversive) only. Hence, the nature of the interaction between appetitive and aversive reinforcers during associative learning in the same experimental situation is not well understood. Scrutinizing this interaction experimentally meets with substantial difficulties (see Dickinson 1976; Mackintosh, 1983; Magoon and Critchfield 2008 for an overview of the underlying theoretical problems). On the procedural side there has been a lack of learning paradigms that train the same behaviour using both appetitive and aversive reinforcers delivered with the same temporal contingency and titrated to achieve comparable effects, such that their combinatorial influence can be quantified. Consequently, most classical work on the subject has relied on indirect methods, typically utilizing sequential interaction between reward-driven and punishment-driven tasks.

1.4.1. Classical works on the investigation of appetitive and aversive reinforcer integration

Early works by Konorski and collaborators focusing on stimulus approach and withdrawal behaviours proposed that the interaction between appetitive and aversive reinforcers is mutually inhibitory in nature (Konorski and Szwejkowska 1956; Konorski 1967). Some of the early work started with the US-US association, by using one type of US as a CS and the other as US. Erofeeva successfully paired the painful shock as a predictor (CS) for food and initiated the investigations. Dickinson and Pierce (1979) also used the shock as a CS

signaling water reward in rabbits. Consequently, the aversive shock lost some of its properties after the preconditioning. Along these lines, cat refused to eat when blast of air on the face was shown while it was eating the food (Masserman 1943). So the counter conditioning procedures led to the conclusion that atleast one of the reinforcers lost their nature while pairing.

Subsequent studies addressed the behavioral influence of stimuli associated with one type of reinforcer on stimuli associated with the other, using summation, retardation and counter-conditioning procedures (Dickinson and Pearce 1977; Dickinson and Dearing 1979; Mackintosh 1983). Here also, an aversive stimulus was observed to suppress an appetitive response, and an appetitive stimulus was observed to suppress an aversive response (Estes and Skinner 1941; Dickinson and Pearce 1977).

The effects of an aversive stimulus on an appetitive response and of an appetitive stimulus on an aversive response were mutually inhibitory, hence causing the phenomenon of conditioned suppression. In the same operant conditioning chamber, appetitive conditioning was severely inhibited if the rats were allowed to do avoidance learning. This conditioned suppression was developed gradually with slighter recovery (Estes and Skinner 1941; Dickinson and Pearce 1977; Hineline 1972).

Scavio et al. (1974) demonstrated that preconditioning of stimuli with shock, impaired the further development of an appetitive response in rabbits. Appetitive conditioning and avoidance learning using the same stimuli indicated that appetitive and aversive conditioned motivational states interact substractively. This suggests that substractive motivational state exists between them (Bull 1970). Conditioned stimuli previously associated with shock can inhibit or block the association of a conditioned stimulus paired with the omission of expected food reward (Dickinson and Dearing 1979). Fear conditioning was greatly enhanced if the CS was previously paired with food (Dickinson 1977).

The underlying motivational states of reinforcer interaction were also investigated using stimulus pre-exposure or pre-conditioning with one reinforcer. The above studies strengthen the argument that appetitive and aversive reinforcers can indeed interact. However, once the behavior is learned, intrinsic motivation also drives and strengthens associative learning (cf. Rolls 2008). It has been demonstrated that, irrespective of reinforcer presentation, response-contingent neutral stimuli also can have intrinsic reinforcing properties (Reed et al., 1996).

Stimulus generalization gradients in appetitive and aversive reinforcement investigated with two different responses showed that gradients of effect were different for the two types of reinforcers (Hearst 1960).

Thus, Konorskian model which explored the relationship between the appetitive and aversive systems brought us four types of motivation. They are 1) Prediction of reward (Hope) 2) Prediction of aversive events (Fear) 3) Omission of expected reward (Frustration) 4) Omission of aversive events (Relief). Hence, he proposed the similarity in the motivational state by the CS associated with shock and the CS signalled the omission of food. Omission of an expected aversive outcome can be rewarding (an aversive inhibitor), and omission of expected reward can be aversive (an appetitive inhibitor).

In order to compare the potentially different roles of reward and punishment, it is desirable to develop behavioral procedures which can incorporate both types of reinforcers within the same training session (e.g. Magoon and Critchfield 2008; Morrison and Salzman 2009) as most previous experiments measured the effect of one reinforcer on the previously established conditioned response (CR) by the other reinforcer. Previous work had studied excitatory or inhibitory interactions between sequential reward- and punishment-driven learning processes (Dickinson 1976; 1977), concurrent schedules of reward and punishment without conditioned stimuli (Kelleher and Cook 1959; Olds and Olds 1962), combinations with secondary reinforcers associated with the opposite valence (Morris 1975; Baron et al., 1977), and non-contingent schedules of aversive and appetitive reinforcers (Stein, 1965; Margulus and Stein 1968; Carder 1970; Castro-Alamancos and Borrell 1992).
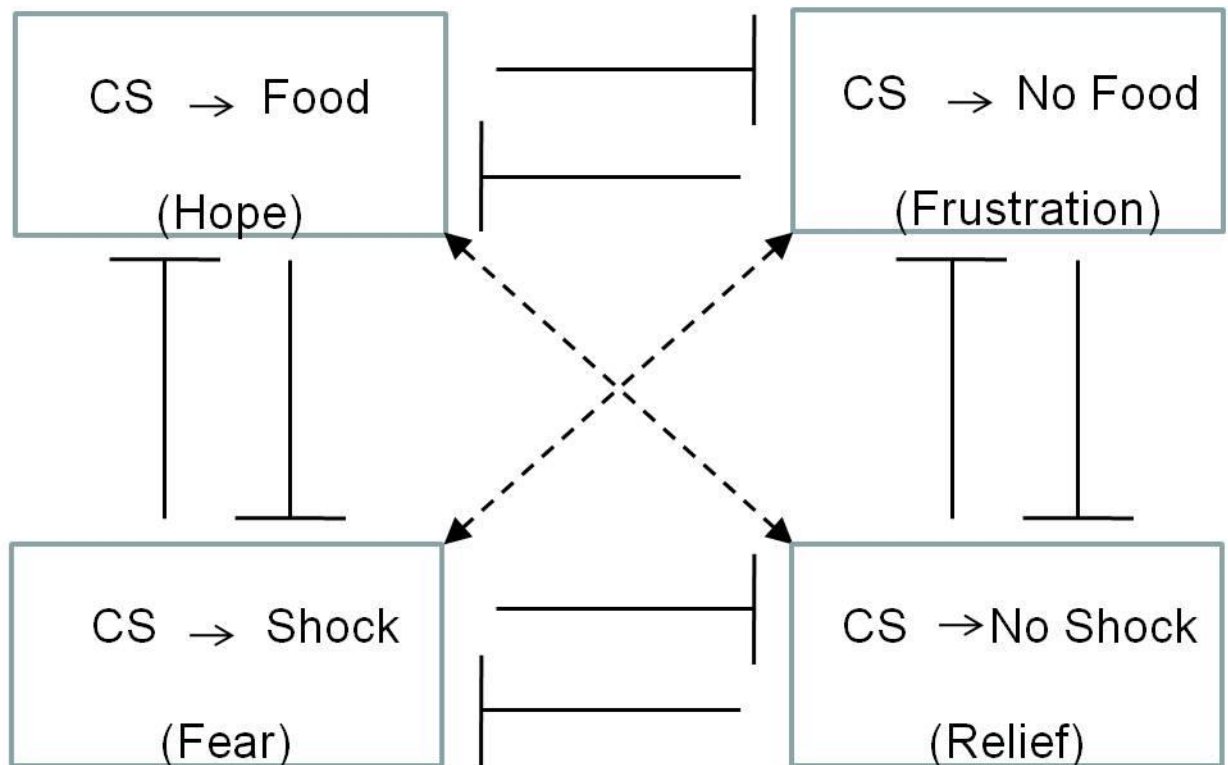
Figure 8. <u>Appetitive, aversive excitatory and inhibitory relationships.</u> *Adapted from 'Conditioning and associative learning' by Mackintosh 1983.*

1.4.2. Problems to address the Interaction through conventional reinforcers

Given the nature of conventional reinforcers such as food and footshock, they involve different behavioural contingencies which are not easily combined in the same experiment and involve different information processing. One principle difference between appetitive and aversive reinforcers is that the effect of appetitive reinforcers typically saturates with prolonged presentation while the effect of aversive reinforcers typically does not. These

circumstances make it difficult to study an identical behaviour that is driven either by reward or by punishment. On the grounds of saturatory effect or drive decay produced by natural reward, we took the alternative route ie., to use the brain stimulation reward as an appetitive reinforcer. Both brain stimulation reward and footshock can create a powerful long-lasting drive to work for, or to avoid it. Though we cannot fully equate the brain stimulation reward with all aspects of natural reward, we could investigate the incentive-motivational property of the reward (Bindra and Cambell 1967).

1.4.3. Reinforcing brain stimulation to address the interaction

It has been reported that rewarding brain stimulation reduced aversive reinforcing property of the peripheral shock when both reinforcers were paired (Cox and Valenstein 1965; Carr and Coons 1982). Some earlier investigators studied the perceptual (applied as a cue) and reinforcing nature of the rewarding brain stimulation. When positively reinforcing posterior hypothalamic brain stimulation was used as a CS, it facilitated the avoidance of aversive footshock. It also decreased the rate of ICSS during the first post conditioning sessions (Mogenson and Morrison 1962). Efforts were already made to demonstrate the facilitation of learning using non-contingent application of positively reinforcing brain stimulation on aversive avoidance behaviour during Sidman avoidance and shuttle-box avoidance learning respectively (Margules and Stein 1968; Castro-Alamancos and Borrell 1992). Non contingent (Margulus and Stein 1968) and contingent presentation of MFB stimulation for avoidance facilitated the Sidman avoidance learning (Carder 1970). After learning to avoid the punishing midbrain stimulation of periventricular grey matter, stimulation of MFB and the associated tegmental regions at the start of the trial (as a priming dose) improved the performance (Stein 1965). Non contingent stimulation of reinforcing medial septum stimulation facilitated active avoidance learning (Goldstein 1966).

Also, ICSS treatment given after the trial or session improved the avoidance learning. Delayed post trial rewarding stimulation of LH facilitated alcove avoidance which requires the inhibition of response (Huston and Muller 1978). Clearly, ICSS not only facilitated the instrumental conditioning, but also the tasks in which inhibition of motor responses is required such as alcove avoidance learning. Learning the reversal of safe compartment was improved if the animals were subjected to post trial application of rewarding LH stimulation during one way active avoidance task (Mondadori et al., 1976; Huston and Muller 1978).

1.5. Aim of the present work and our experimental scheme

The purpose of this detailed introduction is to provide an appropriate foundation and conceptual framework that is necessary for understanding how the role of the mesolimbic DA system in motivated bahaviour by contributing to reward seeking and pain avoidance. From the background of incentive motivation theory and reward prediction error hypothesis, it is clear that the output of the VTA to the ventral striatum is fundamental to any processes that allow the organism for seeking the life's basic needs and to cope up with the threat (Ikemoto 2007).

Moreover, nucleus accumbens (NAc) DA release is a crucial prerequisite for appetitive and aversive motivated behaviour. Episodes of burst firing which leads to transients of NAc DA release can be reliably evoked from intracranial stimulation or self-stimulation (ICSS) of the VTA at frequencies >60Hz. Moreover, reinforcing efficacy of the stimulation can be better studied and optimized with ICSS. Using the brain stimulation reward (BSR) (optimized through ICSS of the VTA) we addressed the relation between appetitive and aversive reinforcement learning.

If DA system is a seeking system, the application of reinforcing electrical brain stimulation of VTA on sensory conditioning should generate response for conditioned stimulus to obtain reward (VTA stimulation). Separately, with the same contingency, we can study two way active avoidance learning in which the tone followed by aversive fooshock if the animal does not perform the response. It is possible that the DA system also serves to process aversive events to increase survival chance. If avoidance learning comes under control of positively reinforcing brain systems, and DA is necessary to progress the avoidance learning, stimulation delivered at the time point of avoidance could accelerate learning process. With the above mentioned points in mind, in the first experiment, three groups of ICSS responders were trained to cross the hurdle for the presentation of the tone.

In one group, crossing the hurdle after tone onset was reinforced by optimized VTA stimulation (appetitive reinforcement). In the second group, the failure to produce the CR was punished by footshock (aversive reinforcement). In both cases the CS duration was the same. In preparatory experiments, current strengths for FS and VTA stimulation were separately calibrated to produce the same asymptotic level of behavioural performance in individual animals. In the third group, the success of avoiding the punishment was associated with

optimized VTA stimulation. In the three groups, we have studied the acquisition and extinction. FS-reinforced learning is initially dominated by aversive experience which leads to subsequent relief upon successful avoidance. Therefore, the primary question we addressed was whether effects from punishment and rewards inhibit each other, or alternatively, whether the relief from punishment and receipt of reward input facilitate the learning process (equivalence hypothesis) (Dinsmoor, 2001).

In the follow-up study, we aimed at investigating the nature of interaction of VTA stimulation with avoidance that facilitated learning. So the second experiment focused on the omission of one reinforcer in the combination experiment after animals had reached maximum performance followed by omission of the remaining reinforcer (extinction). In the third experiment, we violated the outcome expectation by means of partial reinforcement procedures to address the nature of reinforcer interaction (See Table).

Previous studies from our laboratory proved that Mongolian gerbils are a suited animal model for the investigation of learning mechanisms (Wetzel et al., 1998; 2008; Ohl et al., 1999; 2001).

Specifically, in this dissertation, I have addressed the questions raised below:

- How does learning of the same behaviour differ after motivation of it by reward (brain stimulation reward, BSR) or by avoidance of punishment (electrical footshock)?

- How is learning affected when both types of reinforcers are combined in a single training protocol?

- In animals trained with the combined reinforcers, what are the effects of later removal of either one type of reinforcer?

- What would be the nature of the reinforcer interaction, if the prediction of outcome is found to be violated?

2. Methods

2.1. Subjects

A total number of 80 adult male Mongolian gerbils (*Meriones unguiculatus*) obtained from Tumblebrook Farms, West Brookfield, MA, USA (age: 3-6 months, weight: 90 g-110 g) were used in the study. Gerbils were individually housed 3 days before experiments were conducted and were maintained on 12h light/dark cycle (light on 07:00-19:00 h) throughout the experiment. Subjects were experimentally naive, with free access to food and water. Experiments were conducted between 07:30 h to 17:00 h, with individual subjects being trained at a consistent time of day to maintain their daily session interval. All experimental procedures were approved by the Ethics Committee of the State of Sachsen-Anhalt, Germany.

2.2. Surgical procedures

Surgery and implantation of electrodes were performed under ketamine (100mg/kg) (Ratiopharm GmbH, Ulm, Germany) and xylazine (5 mg/kg) (Bayer Vital GmbH, Leverkusen, Germany) anesthesia, which was given intra-peritoneally. The skin above the brain was removed and the skull was cleaned with 3% H_2O_2 to prevent any possible infection. Due to the chronic nature of the experiment, stainless steel (00) needles were fixed around the skull for the stability of the implantation. Animals were fixed in a stereotaxic frame (David Kopf Instruments, USA) (Fig. 9a). Flat brain coordinate was obtained by setting the incisor bar at -5. Bipolar stimulation electrodes with the tips separated by ~0.2 mm were custom made from Teflon-insulated stainless steel microwires (diameter: 140 μm; Science Products GmbH, Germany) and implanted at the level of the ventral tegmental area (2.6 mm posterior to bregma, 1.3 mm lateral to the midline, 5.0 mm ventral to the brain surface) according to the stereotaxic atlas for gerbil by Loskota (1974). Our coordinates, were aimed towards the larger VTA region called "parabrachial pigmented nucleus", which lies medial to the substantia nigra pars compacta and dorsal to the paranigral nucleus. The electrode was fixed in place with dental acrylic cement, and the procedure was repeated in the opposite hemisphere (Fig. 9b). The first self-stimulation training session followed after atleast 4-day recovery period.
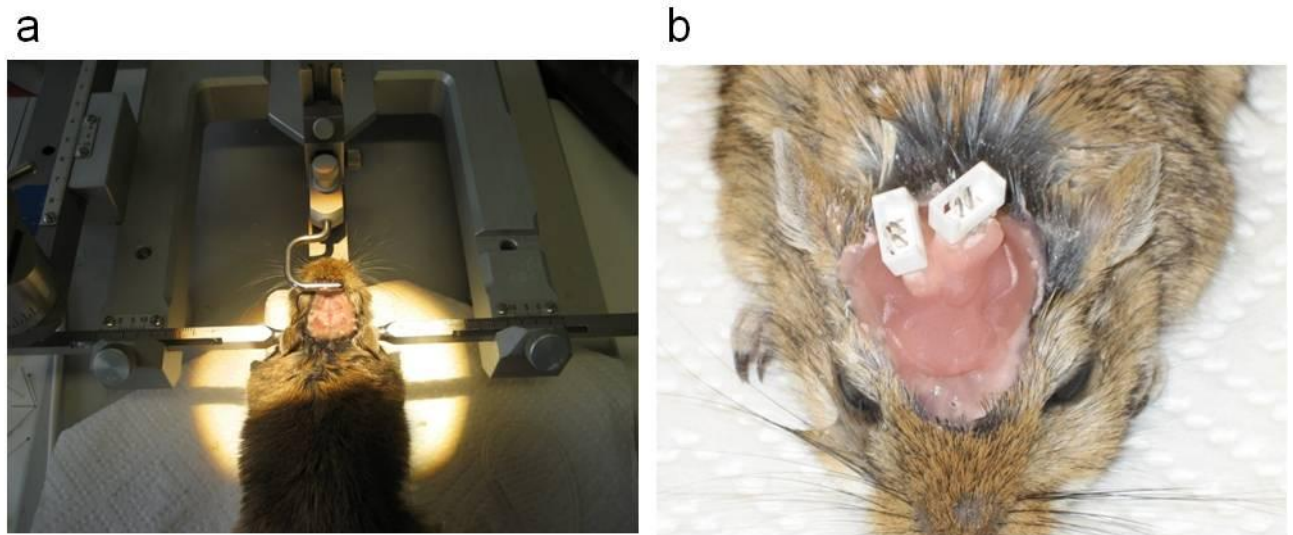
Figure 9. A) Fixation of the animal into the stereotaxic frame. B) After implantation of electrodes in both hemisphere.

2.3. Intracranial self-stimulation

Self-stimulation training was conducted in a custom-made operant chamber (18 cm × 18 cm × 23 cm) with a metal lever in the right lower corner. Gerbils were trained to press the lever for a brief VTA stimulation (200 ms train; 20 biphasic pulses of 0.2 ms duration at 100 Hz). Within 2 to 5 days, the animals learned to press the lever for VTA self-stimulation. Electrodes in both hemispheres were assessed initially. The electrode which supported maximum lever pressing performance in the absence of side effects was chosen for further optimization and shuttle-box learning.

After the gerbils showed lever pressing at stable rates, the optimum stimulus intensity was determined for each individual using ascending and descending stimulation intensities (step size 20 µA; each duration used for 5min) in two separate daily sessions. The current intensity which led to the maximum lever press rate (mean response rate between the ascending and descending stimulus intensities) was selected as the optimum intensity (Fig. 10). The final session consisted of training the subjects with that optimum intensity for 30 min duration. This optimal intensity was used throughout the learning tasks in the shuttle-box. Animals that did not reach the criterion (900 bar presses in 30 min within five consecutive days) were excluded from the study. After at least three days of rest, the intracranial self-stimulation

(ICSS) responders (n = 54) were allotted to various shuttle-box learning groups. The current strength determined in this way was found, when used as appetitive reinforcer in the conditioning experiments, to produce a similar level of final performance as in the FS-conditioned animals.
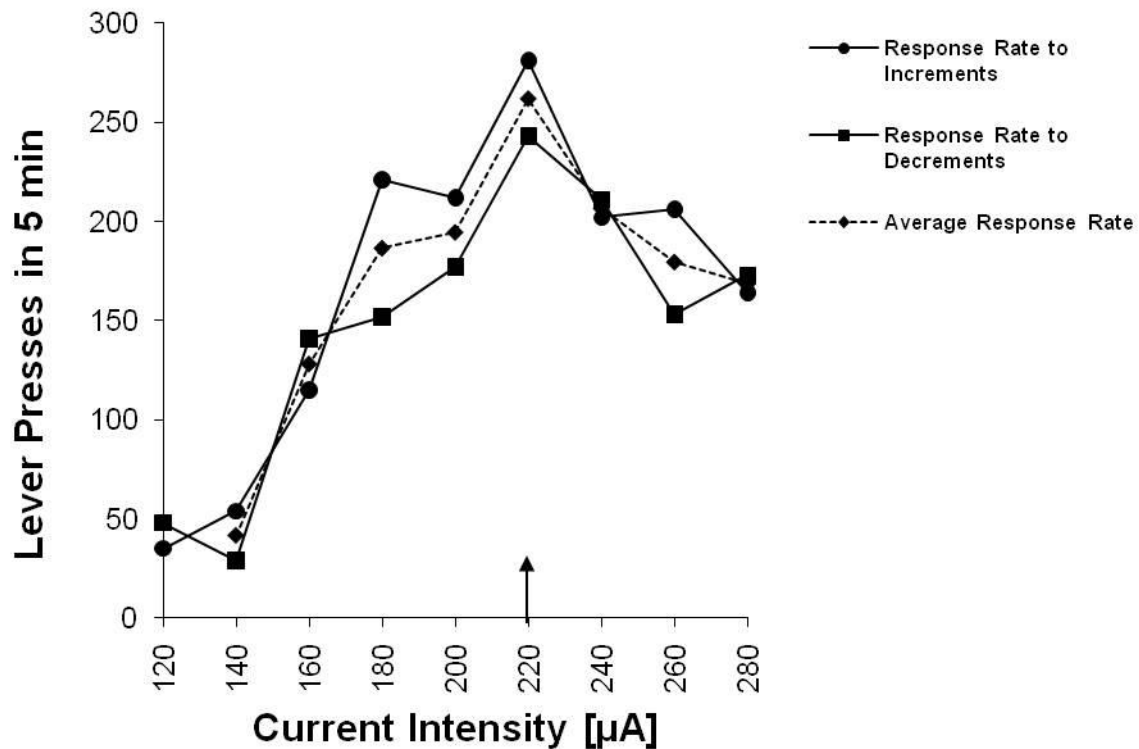
Figure 10. Optimizing the reinforcing VTA stimulation through ICSS: *Exemplary illustration of the procedure preceding shuttle-box training for determining the stimulation current strength for the brain stimulation reward. Plotted is the lever pressing rate of one gerbil against stimulation current intensity used for VTA stimulation. Stimulation current intensity was systematically increased from low levels to high levels (circles) and systematically decreased from high levels to low levels (squares). Diamond symbols indicate the mean rates of lever pressing of the increasing and decreasing series of measurements. The current intensity that produced maximal lever pressing rate (arrow) was used in the later shuttle-box experiments.*

2.4. Shuttle-box learning

Identical behaviour, i.e., tone-conditioned hurdle crossing driven by appetitive, aversive or the combination of both reinforcers was investigated using shuttle-box conditioning procedures. After that, the nature of their interaction was investigated by omission experiments (see Table 1 for overview). Animals were trained in a shuttle-box (38cm × 19cm × 22.5cm) (Hasomed GmbH, Magdeburg, Germany) which had two compartments separated by a 6cm high hurdle. Each daily session consisted of 60 trials with intertrial interval duration of 20 24s. A session began with a 3min habituation period. The conditioned stimulus (CS) was a series of 2 kHz pure tones (6s, 200ms tone duration, 300ms inter-tone interval). When footshock (FS) was used as a reinforcer, a FS of maximally 4s was applied through the grid floor at the end of the CS in case the animal did not cross the hurdle during the 6s CS presentation. The FS was switched off when the animal escaped, i.e. crossed the hurdle during FS presentation. The intensity of FS was slowly raised from 400µA to 600µA during the first training session. Since the rate of conditioning also depends on the change in unconditioned stimulus intensity, we subsequently maintained a constant FS intensity of 600µA. The appetitive reinforcer consisted of five blocks of 200ms trains of electric pulses for brain stimulation reward (BSR) with inter-stimulus interval of 300ms. This pulse train was automatically delivered without delay when the animal reached the other shuttle-box compartment by tone-conditioned hurdle crossing. In different groups, the appetitive reinforcer, the aversive reinforcer, or the combination of both reinforcers was applied. A flexible cable connected with a swivel allowed the electrical brain stimulation and easy movement during shuttle-box learning. The electrical stimulation was delivered by an isolated pulse stimulator (Model 2100, A-M Systems Inc., Carlsborg, USA). Crossing the hurdle during the CS presentation was considered as a conditioned response (CR). The CR rates (number of CR / number of trials) and mean response latencies (times of reaching the new compartment after hurdle crossing) were analyzed in each session. In the first group, the CS was followed by the aversive FS if the animal did not cross the hurdle within 6s. In the second group, a response was considered a CR if a hurdle crossing occurred within 6s of the CS period, in which case, the CR was immediately followed by BSR. In the third group, trained with both FS and BSR, each successful CR was followed by BSR and each failure to produce a CR was punished by FS. So, they received the reward for successful avoidance. The three groups mentioned above (FS: footshock alone, BSR: brain stimulation reward

alone, FS+BSR: footshock and brain stimulation reward combined), were trained in eight acquisition sessions followed by five extinction sessions in which reinforcers were not associated.

The fourth and fifth groups were trained with both reinforcers like the FS+BSR group. After they had been trained to reach maximum performance, in one group we omitted the BSR and in the other group we omitted the FS during sessions 9-13. This was followed by extinction training (sessions 14-18) in which we removed the remaining reinforcer, presenting only the CS.

The abovementioned experiments were repeated under partial reinforcement conditions. To determine the optimal probability of combined reinforcer presentation for the partial conditioning procedure, pilot experiments were first conducted in non-operated animals studying the effect of probability of FS on the CR rate. Four groups of gerbils were trained with FS occurrence probabilities of 0.1, 0.15, 0.2 and 0.5, respectively. Since, in these experiments, FS probability of 0.15 produced asymptotic CR rates between 30% and 50%, and FS probability of 0.2 produced asymptotic CR rates between 50% and 70% (Fig 14), we decided to use a FS probability of 0.15 for training with the combined appetitive-aversive reinforcer. This choice accounted for the expected increase in asymptotic CR level when FS motivation would be complemented by BSR to achieve combined reinforcement. Five (out of 15) animals did not learn the task and were excluded from the study. The sixth and seventh groups of animals were trained with BSR for hits and FS for misses (FS+BSR) with the partial reinforcement schedule, i.e., in 9 out of 60 trials (P=0.15), the hits were rewarded and the misses were punished according to a pseudorandom schedule. After reaching the desired level, like in the other groups, one reinforcer was omitted during sessions 9-13. This was followed by extinction sessions (14-18) in which the remaining reinforcer was also removed.

Table 1. Experimental scheme for the three main experiments.

Experimental group	Associated reinforcers		
	Session 1-8	**Session 9-13**	**Session 14-18**
Experiment 1			
Group 1 (FS)	FS	Extinction	-
Group 2 (BSR)	BSR	Extinction	-
Group 3 (FS+BSR)	FS+BSR	Extinction	-
Experiment 2			
Group 4 (FS+BSR → BSR)	FS+BSR	BSR (omission of FS)	Extinction
Group 5 (FS+BSR → FS)	FS+BSR	FS (omission of BSR)	Extinction
Experiment 3			
Partial Reinforcement			
Group 6 (FS+BSR → BSR)	FS+BSR	BSR (omission of FS)	Extinction
Group 7 (FS+BSR → FS)	FS+BSR	FS (omission of BSR)	Extinction

2.5. Histology and post processing

2.5.1. Isolation of brain and sectioning

At the conclusion of behavioural experiments, the gerbils were deeply anesthetized with ketamine and xylazine mixture, and then the animals were decapitated. The brains were rapidly isolated and frozen in isopentane, immersed in liquid nitrogen and finally stored at -20°C. Within two weeks, we have started to section the brains. The brain was fixed using the embedding medium (Tissue-Tek) and mounted on the freezing microtome. Coronal sections of 40µm roughly posterior to bregma were made using a sliding microtome (Leica CM 3050S, Leica Microsystems Nussloch GmbH, Karlsruhe, Germany). Every second section from the anterio-posterior extension of VTA anatomical limits (~2.5 to 3.5mm posterior to bregma), mounted on gelatin-coated glass microscope slides and allowed to dry overnight. Sections were treated with 70% of alcohol for atleast two hours. Prussian blue staining followed by Nissl staining was performed to reveal the ion deposits around the electrode tips.

2.5.2. Prussian blue staining

Prussian blue staining was done to reveal the ion deposits around the electrode.

Description: With the presence of acid, ferric ion from the tissue reacts with the ferrocyanide and formed a bright blue pigment called Prussian blue, or ferric ferrocyanide. The Prussian blue staining was performed under the temperature of 37°C.

Solutions and Reagents:

- 1% Aqueous solution of potassium ferrocyanide was prepared ie, 2 g of potassium ferrocyanide, trihydrate (K4Fe(CN)6.3H2O, FW 422.4) was mixed with 200 ml of distilled water. 1% hydrochloric acid (ie, 5.405ml from 37% concentrated HCL) was added to the potassium ferrocyanide solution just before use.

- 400 ml of 0.1 M phosphate buffer solution at pH 7.4

The sections were transferred from 70% ethanol solution to distilled water and allowed for 10 min. Thereafter, it was transferred to the acidic pottasium ferrocyanide solution for 10 min. Immediately after the reaction, the sections were immersed in 0.1 M PBS (2×10 min). After which, the sections were rinsed thrice in distilled water (3×5 min).

2.5.3. Nissl staining

Crysyl violet staining can detect the Nissl substance (rough endoplasmic reticulum) found in the neurons. Therefore, we can detect the cell bodies using this method. Below, I summarize the procedure we routinely use in our laboratory.

First, sections were immersed in 0.5M acetate puffer (pH 4.0-4.2) and stained for ~5 min with 0.5% cresyl violet. The sections were then immersed in 0.05M acetate buffer (pH 4.0-4.2). Sections were later passed through a series of ethanol solutions of ascending concentrations (2 min in each of 50%, 70% and 96% v/v ethanol in distilled water). Further the sections were treated (2×5 min) in the solution consists of a mixture of isopropanol and 96% ethanol (2:1 ratio) as part of the dehydration process. After the dehydration, the sections were delipidated in xylol or roticlear (3×5 min) before being cover-slipped using merkoglas and allowed to dry.

2.5.4. Microscopy

The sections were used to verify the blue colour reaction product around the electrode site (Fig 11). Serial sections were examined under a light microscope. The intensity of the blue colour reaction product around the electrode often fades and disappears as we move away from the electrode position.

The locations of the electrode tips were determined with reference to the stereotaxic atlas of gerbils (Thiessen 1977; Loskota 1974). Due to the lack of precise Mongolian gerbil atlas, mouse atlas (Paxinos and Franklin 2001) was also referred to find the anatomical landmarks.
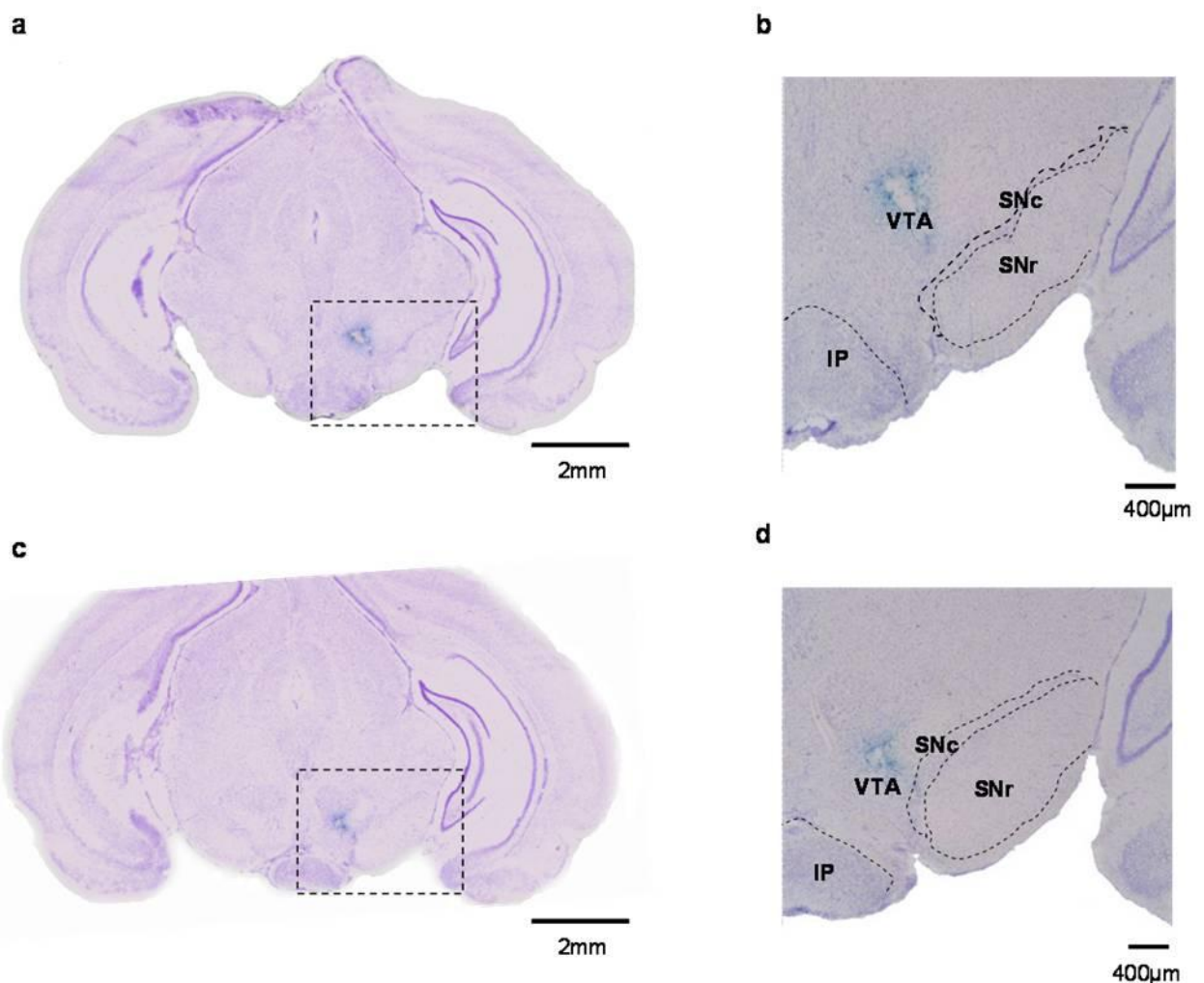
Figure 11. Localization of stimulation sites for BSR using histological analysis. *a,c. Exemplary photomicrographs of two coronal sections (40μm) from approximately 2.6mm and 2.8mm posterior to bregma. The sections were stained with Prussian Blue followed by Nissl staining. Note the blue color reaction product around the electrode location in the stimulated hemisphere. b,d.*

Magnification of the area surrounding the stimulation site in the ventral tegmental area from panels a and c (rectangular areas). Anatomical structures: IP - interpeduncular nucleus; SNc - substantia nigra pars compacta; SNr substantia nigra pars reticulate; VTA - ventral tegmental area.

2.6. Data analysis

CR rates and response latencies were evaluated. To illustrate the learning curves, mean CR rate over daily sessions in each group, during acquisition, omission of one reinforcer and extinction was plotted. Additionally, we measured the hurdle crossing rate during the intertrial interval. The behavioural data from all experiments were analyzed using repeated measures analysis of variance (ANOVA) with two within-subject factors (GROUP and SESSION), using SPSS software for windows (version 8.0). The degrees of freedom were corrected to more conservative values using the Greenhouse-Geiser correction procedure. After confirmation of significant main effect, Tukey post hoc tests were performed to assess pairwise differences among groups.

3. Results

All results were obtained in an auditory shuttle-box learning paradigm. In a 2-compartment shuttle-box, Mongolian gerbils were trained to change the current compartment by crossing a hurdle as the CR to the onset of a series of pure tones. Three experiments were conducted to investigate potentially different effects of appetitive and aversive reinforcers as well as the nature of their interaction when both types of reinforcers were combined within a single session. In Experiment 1, acquisition and extinction of the CR were studied in three experimental groups, using either reward of hits by electrical brain stimulation reward (BSR), punishment of misses by electrical footshock (FS), or a combination of both in the same session (FS+BSR). In Experiment 2, effects of the omission of one type of reinforcer (appetitive or aversive) in animals that had been trained using the combined appetitive-aversive reinforcers were studied using a continuous reinforcement schedule. To violate the expectation of predicted outcome (reward in the case of 100% performance) we further conducted partial reinforcement experiments (See Methods). In Experiment 3, the analogous omission experiment was conducted under a partial reinforcement schedule. Table 1 gives an overview of the experimental groups. In all experiments, CR rates and reaction times were analyzed using a GROUP $\times$ SESSION repeated measures analysis of variance (ANOVA) with SESSION as the repeated factor.

3.1. Experiment 1. *Effects of appetitive, aversive, or combined appetitive-aversive reinforcers on acquisition and extinction of the CR*

In three experimental groups, animals were trained by appetitive reinforcer (BSR on hits), aversive reinforcer (FS on misses), or combination of both reinforcer (BSR on hits and FS on misses) in the tone-cued shuttle-box task. Figure 12a shows the mean CR rates during acquisition and extinction of the CR.

3.1.1. Experiment 1, acquisition

Analysis of the CR rates during the acquisition sessions (1-8) indicated a significant GROUP $\times$ SESSION interaction [$F_{(6.1,45.7)} = 6.1$, $p<0.005$], and main effects of SESSION [$F_{(3,45)} = 52.56$, $p<0.001$] and GROUP [$F_{(2,15)} = 9.9$, $p<0.005$]. The differences in the CR rate in the three groups decreased over the course of learning (Fig 12a). In group FS+BSR, early

(sessions 1-3) acquisition was accelerated by the combination of both reinforcers. Tukey post-hoc comparisons confirmed significant differences between the groups during the first three sessions but not during later (sessions 4-8) stages of acquisition. Specifically, group FS+BSR reached significantly higher CR rates compared to group FS during the first session (p<0.001) and compared to group BSR during sessions 2 and 3 (p<0.001). There was no significant difference between groups BSR and FS observed during the first session of training (p>0.05). However, group FS performed better than group BSR (p<0.005) during the second and third sessions.

Analysis of reaction time (see Material and Methods) was restricted to sessions 4-8, after learning had stabilized (Fig 12b). Reaction time analysis in the three groups showed no effect of GROUP $\times$ SESSION interaction [$F(4.21,31.5) = 1.56$, $p>0.2$], but main effects of SESSION [$F(2.1,31.5) = 7.4$, $p<0.005$] and GROUP [$F(2,15) = 10.6$, $p<0.005$]. Subjects trained with the combination of both reinforcers showed significantly shorter RTs compared to group FS (p<0.05), but not compared to group BSR (p>0.05). RTs did not differ between groups FS and BSR (p>0.05), except for the last session (p<0.05). RTs in group FS tended to be longer than in groups BSR and FS+BSR, indicating a speed-up of RT with the involvement of rewarding stimuli. The evaluation of mean number of intertrial crossing across the eight acquisition sessions showed no significant differences between the three groups [$F(2,105) = 1.51$, $p=0.25$].

3.1.2. Experiment 1, extinction

The last session of acquisition (session 8) served as a baseline for extinction analysis. All groups showed a clear extinction of the conditioned hurdle crossing over sessions (Fig 12a). Main effects for factor SESSION [$F(2.6,39.1) = 55.5$, $p<0.0001$] and for factor GROUP [$F(2,15) = 6.2$, $p<0.05$] but no significant GROUP $\times$ SESSION interaction [$F(5.21,39.1) = 1.66$, $p>0.05$] was found. Tukey post hoc analysis revealed that the main effects were carried by differences between group FS+BSR on the one hand and groups FS and BSR on the other while groups FS and BSR did not differ significantly from each other (p>0.05).

a
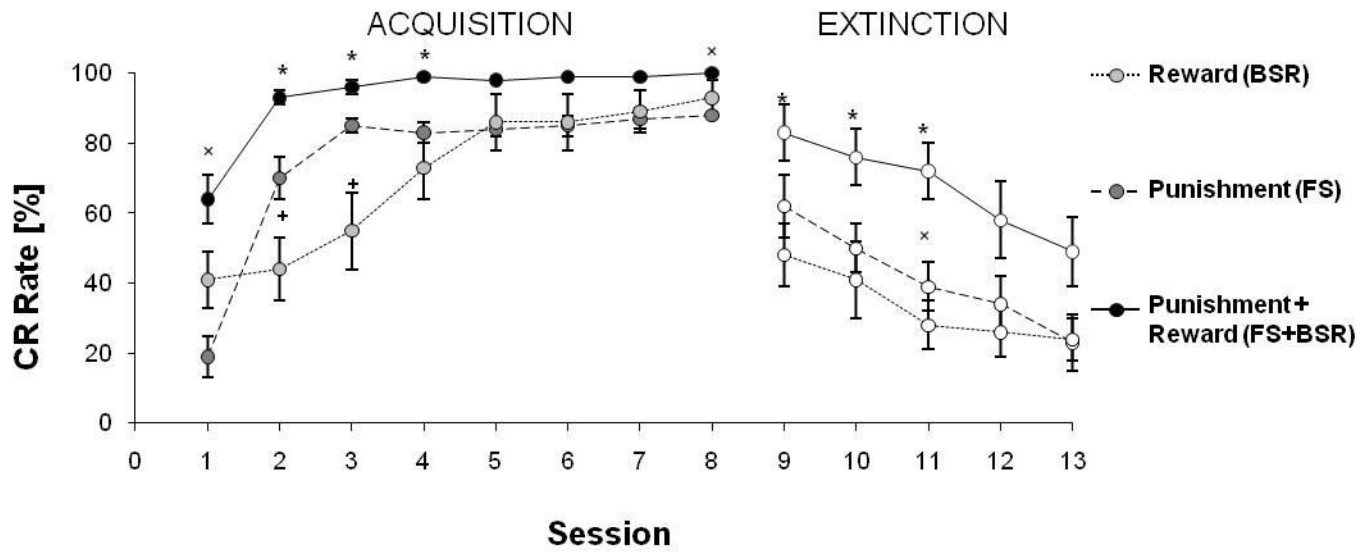
b
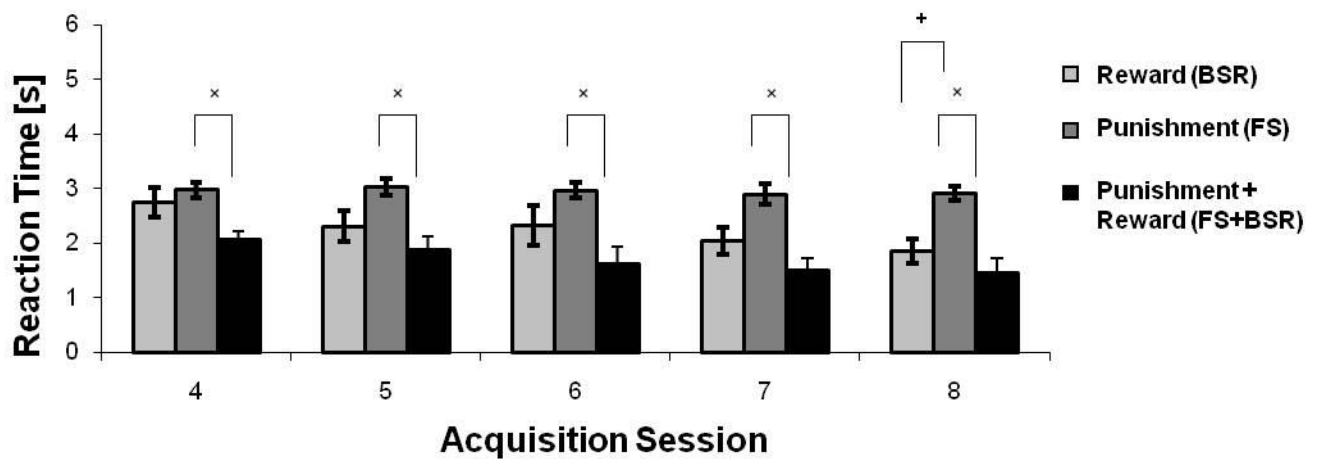
Figure 12. Effect of reinforcer type on acquisition and extinction of a conditioned response (CR). *In the three groups (n=6 in each group), learning was driven by FS punishment following misses (FS, dark grey), brain stimulation reward following hits (BSR, light grey), or both (FS+BSR, black). a. Mean CR rates (± SEM) plotted against training session. The break in the graphs indicates the transition from acquisition training (sessions 1-8) to extinction training (sessions 9-13, indicated by open symbols). Significant (p<0.05) pair wise differences are indicated by symbols: *: BSR vs. FS+BSR; ×: FS vs. FS+BSR; +: FS vs. BSR). b. Mean reaction time (± SEM) during acquisition sessions 4-8.*

Experiment 2. Effects of omission of one reinforcer after completed training with a combined appetitive-aversive reinforcer (continuous reinforcement schedule).

Because of the qualitatively different nature of appetitive and aversive reinforcers, quantitative comparisons between their effects were not trivially achieved. For example, a direct comparison of the effect of reinforcer (appetitive, aversive or combined appetitive-aversive) on the acquisition of a CR in Experiment 1 was made possible by the adjustment of reinforcement parameters to values that would produce indistinguishable asymptotic CR rates. This, of course precluded an analysis of potential effects of three reinforcement types on the retention of an already acquired CR. To study potential effects of the different reinforcer type on the retention of an already acquired CR, in a second experiment, we trained two groups of animals to a high level of performance using combined appetitive-aversive training (session 1-8), and then omitted either the appetitive or the aversive reinforcer while continuing the other (session 9-13). Subsequently (session 14-18), the previously remaining reinforcer was also omitted, effectively resulting in an extinction session. This omission approach was carried out in two versions, one employing continuous reinforcement (Experiment 2) and the other one employing partial reinforcement (Experiment 3).

3.2.1. Experiment 2, Acquisition using the combination of both reinforcers

Two experimental groups were trained in an identical fashion using combined appetitive-aversive reinforcers (FS+BSR) during acquisition training (sessions 1-8; Fig. 13a). Consequently, the statistical analysis of the CR rates demonstrated that the two groups acquired the same level of performance at the same rate: The ANOVA revealed a main effect for SESSION [$F(2.9,29.2) = 114.3$, $p<0.001$], but not for GROUP [$F(1,10) = 1.28$, $p>0.05$] or the interaction GROUP $\times$ SESSION [$F(2.9,29.2) = 1.02$, $p>0.05$]. Likewise, analysis of reaction time indicated no significant differences between the experimental groups (GROUP $\times$ SESSION [$F(3.4,34.4) = 0.6$, $p>0.5$]; SESSION [$F(3.4,34.4) = 22.7$, $p<0.001$]; and GROUP [$F(1,10) = 0.005$, $p>0.5$]).

3.2.2. Experiment 2, Omission of one reinforcer

Following omission of either FS or BSR (sessions 9-13; Fig. 13a), CRs continued at high rates which remained constant across all five sessions, reflected by an absence of a SESSION

effect [F(2.6,26.5) = 0.33, p>0.5] and of a GROUP × SESSION interaction [F(2.6,26.5) = 0.70, p>0.5]. However, a main effect for factor GROUP was found [F(1,10) = 8.9, p<0.05]. The omission of BSR caused a slight drop in responding whereas the omission of FS had no effect. Paired t-tests were used to compare the last session of acquisition and the first session of omission in each group. A non-significant difference in BSR (FS omitted) (t=1.75, df=5, p=0.14) and significant difference in group FS (BSR omitted) (t=2.61, df=5, p=0.04) were found. Significant differences in reaction time (Fig. 13b) were found during the omission sessions (GROUP × SESSION [F(2.46,24.6) = 2.73, p>0.05]; SESSION [F(2.46,24.6) = 0.90, p>0.05]; and GROUP [F(1,10) = 15.04, p<0.005]). In summary, an increase in reaction time was observed for omission of reward but not for omission of punishment.

3.2.3. Experiment 2, Omission of remaining reinforcer (Extinction)

The analysis of the extinction sessions (session 14-18; Fig. 13a) revealed clear extinction (main effect of SESSION: [F(2.7,27.7) = 2.07, p<0.001]), but no differences between the two experimental groups (GROUP: [F(1,10) = 0.14, p>0.05]; GROUP × SESSION [F(2.7,27.7) = 0.93, p>0.05]). These results suggest that, after omission of one reinforcer, the strength of the conditioning was the same during extinction.

a

b
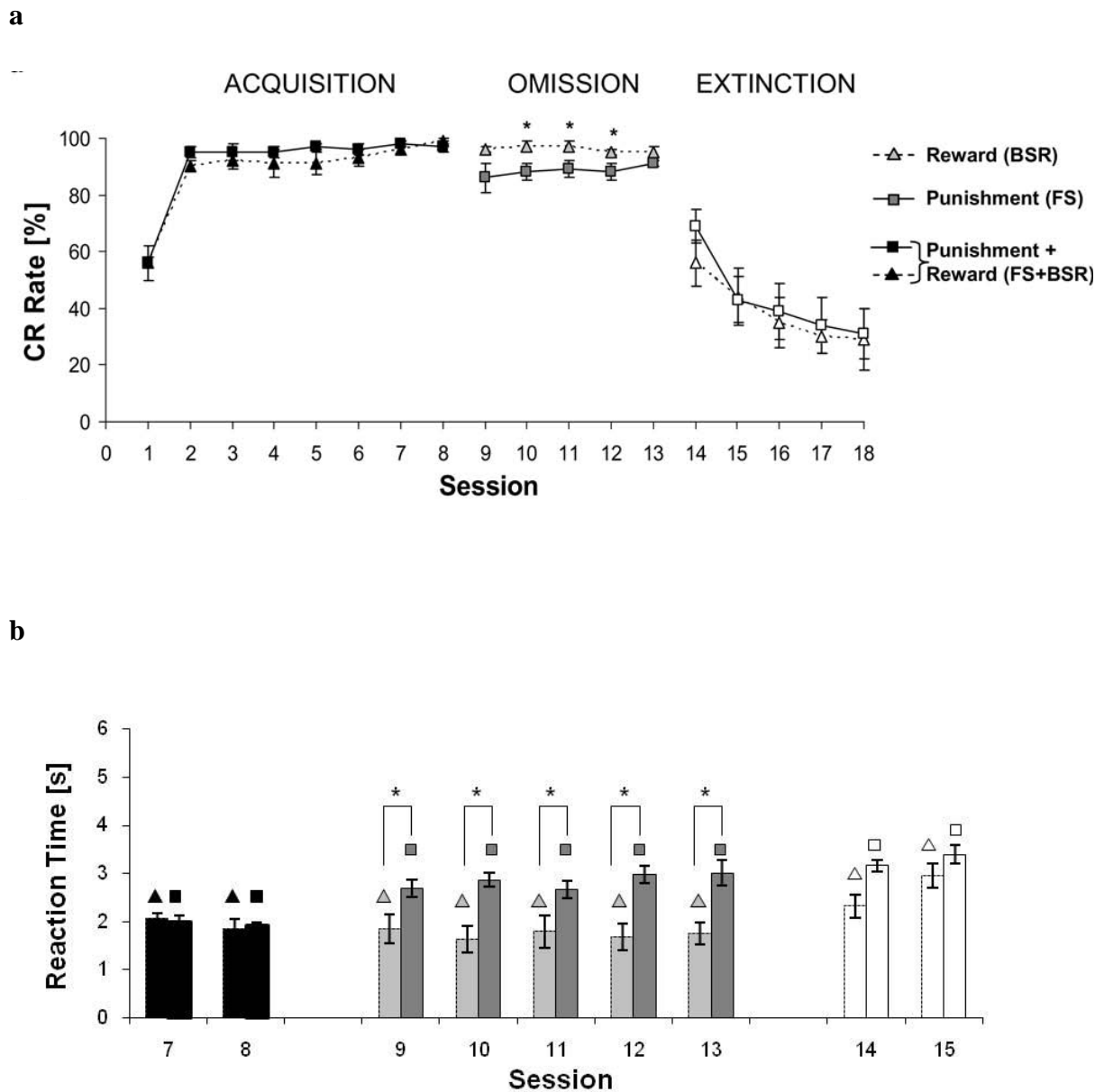
Figure 13. Effects of removing one reinforcer in animals trained with the combined appetitive-aversive reinforcement using a continuous reinforcement schedule. *After training with the combined reinforcers (black filled) during sessions 1-8, BSR was omitted in one group (retaining FS reinforcement, n=6, dark grey squares) and FS was omitted in the other (retaining BSR reinforcement, n=6, light grey triangles) in sessions 9-13. In sessions 14-18 the previously retained reinforcer was also omitted (extinction). a. Mean CR rates (± SEM) for the two experimental groups (dotted vs continuous lines) plotted against training session (* indicates significant group differences, p<0.05). b. Mean reaction time (± SEM) for acquisition sessions 7-8, omission sessions (9-13) and extinction sessions 14-15.*

3.3. Experiment 3. Effects of omission of one reinforcer after completed training with a combined appetitive-aversive reinforcer (partial reinforcement schedule).

The same type of experiments as described above was repeated using partial reinforcement because we speculated that the effect of omitting the internal reward during the continuous reinforcement procedure could be due to the reinforcer history. Specifically, the small decrease in CR rate in group FS (omission of BSR) not seen in group BSR (omission of FS) could simply reflect that, when highly trained animals showing significantly more hits than misses, omission of BSR during hits implied a greater change in reinforcement than omission of FS during misses. We therefore adjusted the probability of FS and BSR association (partially reinforced), such that the final performance of animals trained with the combined appetitive-aversive reinforcers was near 50% CR rate.
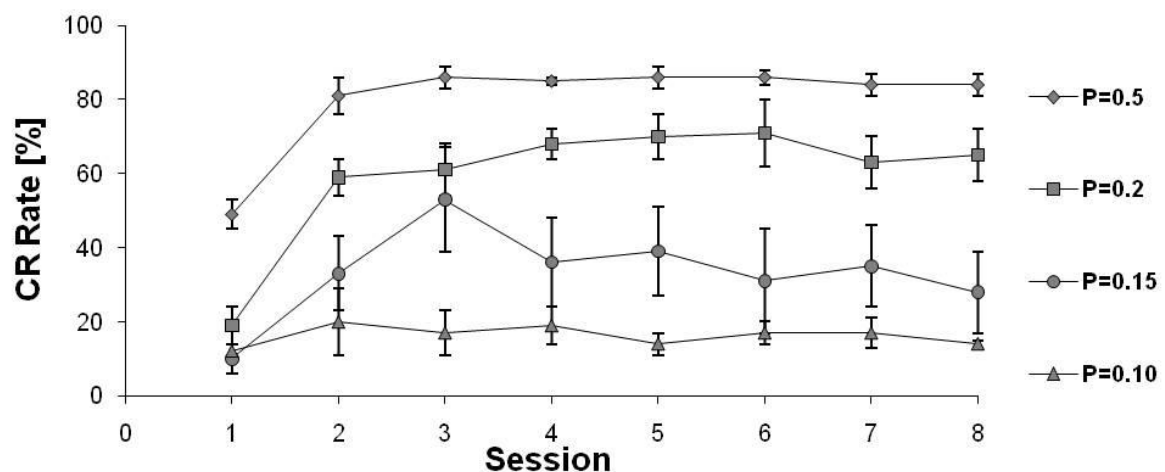
Figure 14. Effect of different probabilities of footshock presentation following misses. *The probabilities studied were P=0.5(n=5), P=0.2(n=5), P=0.15(n=4) and P=0.10(n=3). This experiment was a pilot study for Experiment 3. FS probability of 0.15 produced asymptotic CR rates between 30% and 50% and was selected as the probability for the combined appetitive-aversive reinforcement in Experiment 3.*
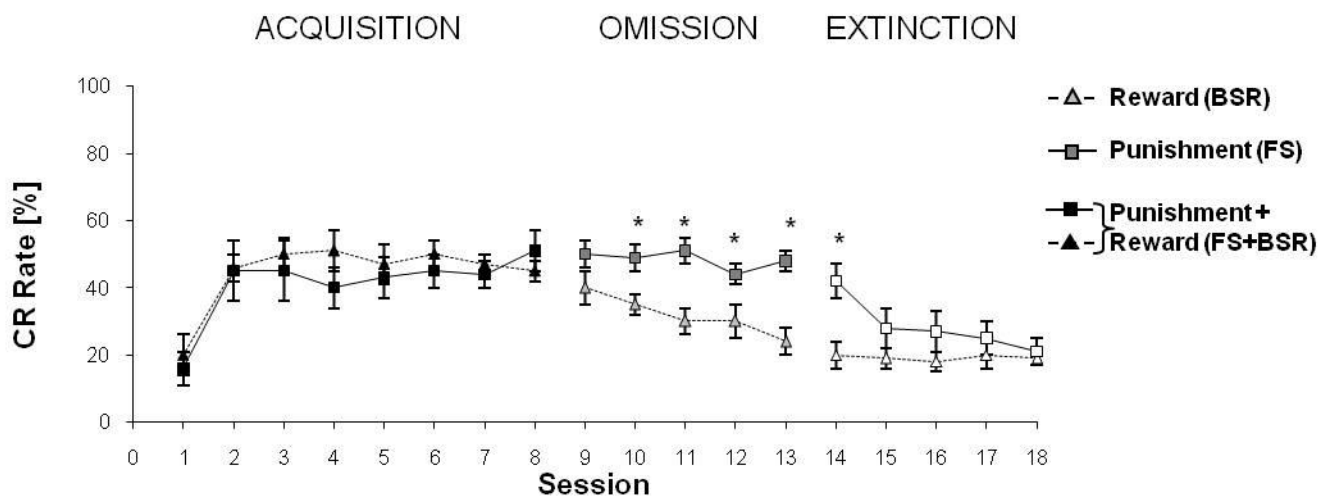
Figure 15. Effects of removing one reinforcer in animals trained with the combined appetitive-aversive reinforcers under partial reinforcement. *Mean CR rates (± SEM) for both groups were plotted against training session. After training with both reinforcers (black filled) during sessions 1-8, BSR was omitted in one group (retaining FS, n=5, dark grey squares) and FS was omitted in the other (retaining BSR, n=5, light grey triangles) in sessions 9-13. In sessions 14-18 the previously retained reinforcer was also omitted (extinction). Asterisks indicate significant group differences, p<0.05).*

3.3.1. Experiment 3, Acquisition using the combination of both reinforcers

Under the partial reinforcement condition using combined appetitive-aversive reinforcers (Fig. 15), significant acquisition (sessions 1-8) of the CR was observed (SESSION: $[F(3.0,24.1) = 12.5, p<0.001]$) that was expectedly indistinguishable for the two experimental groups (GROUP: $[F(1,8) = 0.3, p>0.5]$; GROUP $\times$ SESSION: $[F(3, 24.1) = 0.71, p>0.5]$).

3.3.2. Experiment 3, Omission of one reinforcer

Subsequent omission of one reinforcer under partial reinforcement (session 9-13) led to clear changes in the CR rate development (SESSION: $[F(2.6,21.0) = 5.8, p<0.05]$) and significant group differences (GROUP: $[F(1,8) = 11.4, p<0.05]$; GROUP $\times$ SESSION: $[F(2.6,21.0) = 3.6, p<0.05]$). These results demonstrate that under partial reinforcement conditions, omission of FS for misses decreased the CR rate while the omission of BSR for hits did not.

3.3.3. Experiment 3, Omission of remaining reinforcer (Extinction)

Analysis of the extinction curves (session 14-18; Fig. 15) revealed a main effect for SESSION $[F(2.8,23.1) = 8.7, p<0.005]$ and GROUP $\times$ SESSION remained significant $[F(2.8,23.1) = 6.54, p<0.005]$ but not for GROUP $[F(1,8) = 2.82, p>0.1]$.

4. Discussion

4.1. Experiment 1: Reinforcer integration and extinction

In essence, we found that acquisition of a new behaviour, tone-conditioned hurdle crossing, can be learned either by appetitive or aversive reinforcement and facilitated by the interaction of the two reinforcements as predicted by equivalance hypothesis. This superlearning, characterized by very fast acquisition and close to 100% performance level, is extremely unusual in a 2-way shuttle box paradigm, and was only described in rats for a particular, genetically selected, strain (Aguilar et al., 2004; Corda et al., 2005). Taken together, our results demonstrate a strong improvement of learning, the behavioural task by functional interaction between appetitive and aversive reinforcers and the relevance of a bipolar motivational state seeking for avoidance of one and gain of the other reinforcer. The acquisition of the CR under combined appetitive-aversive reinforcement (Fig.12a, sessions 1-8) suggests that the successful avoidance effect associates with the stimulation effect of intrinsic reward, therefore usage of both types of reinforcer led to faster acquisition to asymptotic level of performance. Our present findings appear to match with previous studies (Tanimoto et al., 2004; Stark et al., 2004) indicating that the termination of aversive stimuli itself is rewarding. Therefore, addition of intrinsic reward to the relief from punishment in a time dependent manner accelerated the learning. The strength of the conditioning, using combined appetitive-aversive reinforcers may be explained by an additive effect in the internal reward system.

In general, avoidance responses with aversive reinforcers have appeared to be more resistant to extinction than appetitive conditioning (Mackintosh 1983), presumably because of the underlying fear resulting from the first classical tone-shock conditioning stage of the instrumental learning. In our results (Fig. 12a), however, no significant difference in extinction was observed between the groups trained with appetitive or aversive reinforcers. This is probably due to differences in experimental procedures. In our experiment direct stimulation of the internal reward system was used instead of a natural reward, e.g. food, which nevertheless suggests that internal reward-driven learning in principle can be as resistant to extinction as avoidance learning.

As already suggested by the higher final level of performance, the strength of conditioning as revealed by extinction was much higher in the group trained with combined appetitive-aversive reinforcement (Fig. 12a). So, the gain from VTA stimulation for avoidance persisted even during extinction as indicated by the high rate of responding. Although, VTA stimulation led DA release highly improved the learning during acquisition of instrumental conditioning, the driven behaviour guided by enhanced incentive motivation can survive for long period without phasic DA signal (Wise 2004). This high resistance to extinction was probably, DA reinforcement by VTA stimulation as a response feedback for avoidance creates powerful long-lasting stimulus-response habits (Hermer-Vazquez et al., 2005; Graybiel 2008). At the behavioural level, we have less evidence concerning the comparision of extinction from appetitively and aversively reinforced learning. An earlier study in which conditioned aversive stimuli (tone-shock pairing) followed by ICSS of the LH showed greater resistance to extinction when tested with appetitive conditioning (Coulombe and White 1980).

Not many studies have so for focused on the role of DA in extinction; nevertheless it was shown that DA agonists reinstate the extinguished behaviour based on reward (Wise et al., 2004). Reinstatement after extinction elevated the DA level in NAc (Ranaldi et al., 1999). Previous studies have pointed out the fact that extinction learning also correlated with decreasing DA level in the NAc. Fast scan cyclic voltammetry measurements from NAc indicate that DAergic signals, just after lever pressing decreases during cocaine extinction, while inter response interval increases (Stuber et al., 2005).

A previous report from our laboratory has demonstrated a transitory increase of medial prefrontal DA levels during the initial stages of FS-reinforced avoidance conditioning (Stark et al., 1991; 2001; 2004). This has been interpreted as a physiological correlate of relief from the punishment. We hypothesize that, in the group trained with both reinforcers, the learning was highly facilitated because of the additive effect of footshock avoidance and brain reward. During acquisition and extinction, DA releases are different among the efferent terminal regions. In fear conditioning, DA release was elevated in mPFC during the first CS-US association while no change observed in NAc. However, when the learning of CS-US begins, NAc DA level was increased and mPFC DA level decreased. During extinction, DA level in both regions were decreased (Wilkinson et al., 1998).

Although the final level of performance shown by the groups FS and BSR was the same, the latter was slowest in developing this conditioned response level and fastest in extinction. It is interesting to note that in operant conditioning, extinction of lever pressing for brain stimulation reward requires only few trials.

4.2. Experiment 2 and 3: The nature of appetitive and aversive reinforcer interaction during continuous and partial reinforcement procedures

The purpose of the selective reinforcement omission experiments in the groups trained with combined appetitive-aversive reinforcement after reaching stable performances was to examine how the two reinforcers interact at high performance levels to maintain that level. While omission of FS did not change the maximum performance level, omission of BSR led to a drop of performance (Fig. 13a). In an earlier study, addition and omission of punishment for food reinforced responses produced opposite effect on performance ie., reduction of responses for initial addition and facilitation for later omission was reported. Even though it was not fully comparable to our study, we observed some common effect. The effect of drop in performance resembled the omission of reward effect in our study (Azrin 1960). The increase in latency followed by the omission of BSR is consistent with the assumption that attribute to the role of DA in determining the response vigor (Salamone 1994; Salomone and Corea 2002). Taken together, the asymmetric results indicate that reward is more important to maintain high levels of performance during continuous reinforcement and punishment is more important to maintain the learning in partial reinforcement (Fig. 13a, Fig. 15). It should be noted that this interpretation, while addressing the mechanistic level of events, is in accordance with Konorski's psychological model of opponent processes and the motivational states between them namely (1) Predicted appetitive stimuli (Hope) (2) Predicted aversive stimuli (Fear) (3) Absence of predicted appetitive Stimuli (Frustration) (4) Absence of predicted aversive Stimuli (Relief) (Fig. 8) (for recent review see Seymour 2007). Thus we provide the experimental framework of results which supports the current motivational hypothesis that argues for the functional relevance of aversive inhibitors (relief from punishment) and appetitive excitors (expectation of reward). The conclusion derived from this experiment suggests that punishment accelerated the initial acquisition while reward maintained the vigour of learned responses.

At this higher level of performance, there were occasional footshock (FS) and the learning was driven by contingent presentation of reward. Hence, omission of FS led to no effect on performance. The decreased CR rate observed after omission of brain stimulation reward (BSR) in animals trained with the combined appetitive-aversive reinforcers might be attributable to the reinforcer history during continuous reinforcement namely the predictability of reward. Therefore, partial conditioning allowed us to manipulate the effect of reinforcer history by violating the reinforcer expectation in an unpredictable fashion. The comparison of results from omission of one reinforcer in continuous and partial reinforcement highlights the fact that the strength of reinforcers during fully predicted continuous and partial reinforcement conditions are different (Fig. 13a, Fig. 15). Under partial reinforcement condition, omission of FS decreased the performance and omission of BSR had no effect in contrast to the continuous reinforcement where the opposite effect was found. Since, under the partial reinforcement regime, non-reinforced trials also influence the performance. The conclusions derived from omitting one reinforcer suggest that in a fully predicted, continuously reinforced environment, reward is more vital to maintain a high level of performance but, in an unpredictable, partially reinforced situation punishment is more effective-in each case due to different frequencies of occurrences of rewards and punishments.

4.3. Possible brain systems underlying the integration of reward and punishment

Reward (appetitive reinforcer) and punishment (aversive reinforcer) are in principle both potent drives for the acquisition and maintenance of behaviours. The underlying mechanisms to learn about apptitive and aversive reinforcers, the emotional states elicited by them and the brain systems that govern reinforcement integration are of high interest (Rolls 2000). It is now clear that multiple brain systems are coding the appetitive and aversive reinforcers. But, it is still unclear whether they are operating in parallel or in series during learning for pleasure seeking and pain avoidance.

During associative learning, the dopamine system has been considered to play a crucial role in encoding reinforcement occurrences. Tonically active midbrain dopaminergic neurons phasically increased their frequency of firing by an unexpected appetitive reward during the initial stage of learning (Mirenowicz and Schultz 1996) and shifted their responses towards the onset of CS after learning (Schultz et al., 1997). Their responses to punishment were not

fully explored, even though they are inhibited by omission of reward (Tobler et al., 2003) and also by aversive stimuli (Ungless et al., 2004). Beside reflecting differences of phasic firing of DAergic neurons and the more tonic activation changes presumably detected by DA microdialysis in the brain, different roles of the DAergic signals in learning processes are started to emerge (Gallistel 2006; Hernandez et al., 2006; Day et al., 2007; Stark et al., 2008). Midbrain DAergic neurons, effectively code for different aspects of reward namely magnitude (amount), probability and delay. With the available information we are not sure whether it does the same for aversive stimuli. Recent studies in primates pointed out the fact that the other basal ganglia neuromodulator ie., striatal cholinergic interneurons can efficiently encode the difference between the omission of reward and punishment (Joshua et al., 2008). Undoubtedly, single neurotransmitter DA cannot govern all the motivational processes, thus the relative engagement of other systems should be explored. Such as, interaction of serotonergic raphe nuclei (Daw 2002) and lateral habenula (Matsumoto et al., 2007) with the midbrain DAergic system for encoding the motivational states during the opponent processes such as appetitive and aversive (Seymour et al., 2005) would shed light on this issue. That would help us to understand the interaction of brain systems for efficient learning and learned helplessness and depression (Shumake et al., 2003).

Despite substantial progress in our understanding of appetitive and aversive reinforcement, the old question of which brain systems integrate them (e.g. Valenstein and Valenstein 1964), needs further exploration (Leknes and Tracey 2008). Subsequent to a demonstration with event-related fMRI that medial parts of orbitofrontal cortex (OFC) in humans are activated by an appetitively conditioned CS and lateral parts by an aversively conditioned CS (Gottfried et al., 2002), results of a visual conditioning task showed that appetitively and aversively reinforced activations (gain of reward and loss avoidance) map on the same area in medial OFC suggesting their place of interaction (Kim et al., 2006). Both medial and lateral parts of the OFC are functionally well connected with basolateral amygdala (BLA) and nucleus accumbens (NAc). BLA as well as OFC encode expected aversive and appetitive outcomes during learning (Schonenbaum et al., 1998). Similarly, most of the recorded NAc neurons are innately tuned for appetitive sucrose with decrease in firing rate and increase in firing rate for aversive quinine and are linked to motor output (Roitman et al., 2005). Given the complexity of the nervous system, it should not be surprising that information about reward and punishment occurs in many brain systems. But how the integration of reward and punishment

information signals integrated for successful learning remains unclear. Dynamic understanding between the type of neurons in different brain systems to process the meaning of opponent processes underlying the appetitive and aversive states would be highly interesting to explore.

4.4. Modulatory signals by dopamine system

It is apparent that dopamine modulates learning. The DA signal from the midbrain substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA) contributes to associative learning processes in which the exact timing of the reinforcement is vital (Schultz 1997). Consistent with the role of DA as an encoder of stimulus and response associations which drives an instrumental act, DA medications in neuropsychiatric patients showed improvement of learning from positive outcome and deficits in avoiding the aversive outcome (Frank et al., 2004).

The modulatory signal provided by VTA DAergic neurons are fast, phasic or slow, spatially diffuse signal. GABAergic input from ventral pallidum (VP) control the tonic DA level by acting across the population level and glutamatergic input from pedunculopontine tegmentum (PPTg), altering the phasic DA release by acting specifically on DA neuron (Floresco et al., 2003). Indeed, many other systems also contribute to the regulation of the tonic and phasic activity of DAergic neurons. The phasic bursts (called as "teaching signal" after learning) provide the information about saliency of the stimuli and expectation of reward to the ventral striatum. Corticostriatal, hippocampal glutamatergic afferents also modulate the tonic DA level in the striatum. Enormous amount of studies suggested the modulation of synaptic plasticity by DA. D1 and D2 also differently modulate the Go and No Go pathway.

4.5. Conclusions and future directions

The dissertation provides insight in the association of DA signal as a tool to investigate the nature of reinforcement learning processes. Also, behavioural evidences concerning the role of dopamine signal for motivated behaviour has been demonstrated. I have achieved to address the positive reinforcing nature of the DA system and explained how I used contingent activation of this system as a supplement for appetitive reinforcer thereby addressing the nature of interaction between appetitive and aversive reinforcer. I also addressed the mechanisms involved in appetitive and aversive reinforcement learning in Mongolian gerbils

(*Meriones unguiculatus*) using footshock (FS) as aversive reinforcer or brain stimulation reward (BSR) ie., electrical stimulation of the ventral tegmental area or both, during auditory learning in the shuttle-box. After studying the effect of both reinforcers separately, I addressed reinforcer integration using FS as an aversive reinforcer upon each unsuccessful trial and BSR as an appetitive reinforcer upon each successful trial. This matched power of the reinforcers, in principle, allowed to determine any possible type of interaction in combined experiments. Acquisition of a new behaviour, tone conditioned hurdle crossing, can be learned by using either appetitive BSR or FS and the FS avoidance is facilitated by the addition of VTA stimulation (BSR). Association of DA signal to the avoidance event boosted the learning in several ways especially, it shortened the training time to reach higher performance. That proves, training by "carrot and stick" combined is a superior method to facilitate learning and to counteract extinction. Extensive training and the DA feedback for every successful response also created the development of stimulus-response habits. Both VTA stimulation and the memory of initial footshock occurrences operated together to drive learning, but their effects were still dissociable. Omission of rewarding VTA signal after learning led to small drop in performance suggesting that the organism is again ready to be governed by pain processing systems. When the expectancies were violated, in an unpredictable situation, punishment plays a major role to drive learning in contrast to fully predicted environment. It is noteworthy that, in the combination of both reinforcers, reward and relief from punishment are only equivalent in the sense of a convergent learning effect but cannot replace each other for reaching maximal learned performance. Our finding of converging and dissociable effects on acquisition and maintenance of learning has implications for optimizing educational and rehabilitational strategies.

In discrimination studies, DAergic neurons profoundly showed more activation towards the Go stimuli. It was hypothesized that DA provides excitatory input to execute the Go response during Go trial, and inhibitory to withhold the response during No Go trial (Frank et al., 2004). Temporally appropriate/inappropriate DA signal as a feedback during Go, No Go task may alter the ability of cue recognition, action selection and decision. These temporally precise response contingent DA feedback will either facilitate the discrimination learning as in detection task or can be used to set the bias towards selecting that particular reinforced response. Using the knowledge we gained from the detection task, we would like to change the focus on Go, No Go discrimination task. Execution/suppression of behaviours leading to

reward or avoidance with the DA feedback would further clarify our understanding on decision making and how it modulated by dopamine release. Upcoming years, the reinforcing midbrain DAergic feedback will teach us their ability/inability to make a better decision.

Looking through the subcortical structures, positive and negative reward signals and the motivational value are successfully encoded by midbrain DAergic and lateral habenula (LHb) neurons. Even though much is known about the event dependent inhibition of midbrain dopamine system, the mechanism underlying this is started to emerge (Matsumoto and Hikosaka 2007). Electrical stimulation studies suggest that single pulse stimulation of LHb, suppressed almost all the activity of recorded DA neurons (Christoph et al., 1986; Ji and Shepard 2007). With the renewed interest in the LHb, it is time to investigate the mutual understanding and dialogue between these two systems. In reward dependent learning, LHb neurons teach the predictive information about the omission of reward to DAergic neurons as opposed to reward trials (Matsumoto and Hikosaka 2007). Perhaps, LHb receives the fully processed information about reward loss and transfers to VTA/SNc directly or indirectly (Hong and Hikosaka 2008; Jhou et al., 2009). LHb not only codes the predictive information of omission of reward during appetitive conditioning but also the arrival of aversive punishment during aversive conditioning (Matsumoto and Hikosaka 2008; Hikosaka et al., 2008). Also our recent study suggests opponent neuromodulation of learning during acquisition by electrical brain stimulation of LHb and VTA in conjunction with events such as avoidance (Shumake et al., 2010). More experimental work is needed to further understand the inhibition of DAergic neurons by LHb neurons.

Recent anatomical studies indicate, VTA is a heterogenous structure with many sub-nuclei. More studies should be done to address the less explored ventromedial part of the VTA DAergic neurons. Phasic excitation of DAergic neurons from this region may evoke DA release in the NAc and mPFC which may subserve the learning of classical tone-shock pairing at the initial stage of avoidance learning. Many electrical stimulation studies or ICSS studies (like ours) focussed VTA as a whole entity and the electrodes were aimed towards the VTA region lies medial to the SNc. From our knowledge, there is no electrical stimulation study which addresses the effect of electrical stimulation of ventromedial region of the VTA called paranigral nucleus. Also, a recent study uncovered the differences in firing rate of DAergic neurons of the primates with response to appetitive and aversive events. Positive and negative motivational states coded by the DAergic neurons were distributed differently.

Aversive reinforcer-excited type DAergic neurons were found in the dorsolateral region of the SNc and inhibited type in the ventromedial region in the SNc as well as the VTA. Due to the difference in projection towards dorsal and ventral striatum, it could be that lateral SNc provides information about salient stimuli and VTA together with the medial SNc supplies reward value (Matsumoto and Hikosaka 2009) (Fig. 7). So, the current consensus that DA neuron only transfer the reward related information needs to be revisited.

In a recent study with real-time measurement, DA release events and pH were measured in the nucleus accumbens (NAc) shell for the appetitive and aversive taste stimuli with opposite hedonic valence. DA concentration was increased for appetitive sucrose and decreased for aversive quinine, a trend which was different from pH changes (Roitman et al., 2008). Perhaps, after learning, DAergic neurons pass the information about saliency of the stimuli (based on the reinforcer value) to NAc. It could be, NAc is acting like an interface to transfer the information about appetitive reward and aversive punishment to higher brain areas like orbitofrontal cortex (OFC). This information processing can be different for appetitive and aversive motivational processes. How OFC underlies the integration of reward and punishment should be explored further (Morrison and Salzman 2009). From this point of view, OFC and NAc are especially important targets to investigate in the future.

From the behavioural point of view, this thesis has tried to integrate a vast majority of information that has emerged in the field of research "on the role of DA in motivated behaviour". The usage of VTA stimulation in my research prompted me to look the extensive literature on DA release into the NAc and its contribution towards associative learning. It will likely prove a good starting point for future work to elucidate further the neural circuit responsible for the convergence of appetitive and aversive information. It will be interesting to see how the future experiments evolve as that understanding develops by following a systems level approach.

References

Aguilar R, Gil L, Fernández-Teruel A, Tobena A (2004). Genetically-based behavioural traits influence the effects of shuttle box avoidance overtraining and extinction upon intertrial responding: a study with the Roman rat strains. *Behavioural Processes* **66**: 63-72.

Azrin NH (1960). Sequential effects of punishment. *Science* **131**: 605-606.

Bao S, Chan VT, Merzenich MM (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* **412**: 79-83.

Bao S, Chan VT, Zhang LI, Merzenich MM (2003). Suppression of cortical representation through backward conditioning. *Proceeding of the National Academy of Sciences* **100**: 1405-1408.

Berridge KC, Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning or incentive salience? *Brain Research Review* **28**: 309-369.

Bielajew C, Shizgal P (1986). Evidence implicating descending fibers in self-stimulation of the medial forebrain bundle. *Journal of Neuroscience* **6**: 919-929.

Bindra D, Campbell JF (1967). Motivational effects of rewarding intracranial stimulation. *Nature* **215**: 375-376.

Bracs PU, Jackson DM, Gregory P (1982). α-Methyl-p-tyrosine inhibition of a conditioned avoidance response: reversal by dopamine applied to the nucleus accumbens. *Psychopharmacology (Berlin)* **77**: 159-163.

Brinschwitz K, Dittgen A, Madai VI, Lommel R, Geisler S, Veh RW (2010). Glutamatergic axons from the lateral habenula mainly terminate on GABAergic neurons of the ventral midbrain. *Neuroscience* **168**: 463-476.

Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceeding of the National Academy of Sciences* **106**: 4894-4899.

Bromberg-Martin ES, Hikosaka O (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**: 119-126.

Bull JA (1970). An interaction between appetitive pavlovian CSs and instrumental avoidance responding. *Learning and Motivation* **1**: 18-26.

Cannon CM, Palmiter RD (2003). Reward without dopamine. *Journal of Neuroscience* **23**: 10827-10831.

Carder B (1970). Lateral hypothalamic stimulation and avoidance in rats. *Journal of Comparative Physiological Psychology* **71**: 325-333.

Carder B (1971). Effects of septal stimulation on active avoidance in rats. *Physiology & Behavior* **6**: 503-506.

Cardinal RN, Robbins TW, Everitt BJ (2000). The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology (Berlin)* **152**: 362-375.

Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* **292**: 2499-2501.

Castro-Alamancos MA, Borrell J (1992). Facilitation and recovery of shuttle box avoidance behavior after frontal cortex lesions is induced by a contingent electrical stimulation in the ventral tegmental nucleus. *Behavioural Brain Research* **50**: 69-76.

Cheer JF, Aragona BJ, Heien ML, Seipel AT, Carelli RM, Wightman RM (2007). Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. *Neuron* **54**: 237-244.

Christoph GR, Leonzio RJ, Wilcox KS (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *Journal of Neuroscience* **6**: 613-619.

Cheer JF, Heien ML, Garris PA, Carelli RM, Wightman RM (2005). Simultaneous dopamine and single-unit recordings reveal accumbens GABAergic responses: implications for intracranial self-stimulation. *Proceeding of the National Academy of Sciences* **102**: 19150-19155.

Comoli E, Coizet V, Boyes J, Bolam JP, Canteras NS, Quirk RH, Overton PG, Redgrave P (2003). A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nature Neuroscience* **6**: 974-980.

Cooper BR, Breese GR, Grant LD, Howard JL (1973). Effects of 6-hydroxydopamine treatments on active avoidance responding: evidence for involvement of brain dopamine. *Journal of Pharmacology And Experimental Therapeutics* **185**: 358-370.

Corda MG, Piras G, Lecca D, Fernández-Teruel A, Driscoll P, Giorni O (2005). The psychogenetically selected Roman rat lines differ in the susceptibility to develop amphetamine sensitization. *Behavioural Brain Research* **157**: 147-156.

Coulombe D, White N (1980). The effect of post-training lateral hypothalamic self-stimulation on aversive and appetitive classical conditioning. *Physiology and Behavior* **2**: 267-272.

Cox VC, Valenstein ES (1965). Attenuation of aversive properties of peripheral shock by hypothalamic stimulation. *Science* **149**: 323-325.

Daw ND, Kakade S, Dayan P (2002). Opponent interactions between serotonin and dopamine. *Neural Networks* **15**: 603-616.

Day JJ, Roitman MF, Wightman RM, Carelli RM (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature Neuroscience* **10**:1020-1028.

Deutsch JA, Di Cara L (1967). Hunger and extinction in intracranial self-stimulation. *Journal of Comparative and Physiological Psychology* **63**: 344-347.

Dickinson A (1976). Appetitive-aversion interactions: Facilitation of aversive conditioning by prior appetitive training in the rat. *Animal Learning and Behaviour* **4**:416-420.

Dickinson A, Pearce JM (1977). Inhibitory interactions between appetitive and aversive stimuli. *Psychological Bulletin* **84**: 690-711.

Dickinson A (1977). Appetitive and aversive interactions: Superconditioning of fear by an appetitive CS. *Quarterly Journal of Experimental Psychology* **29**:71-83.

Dickinson A, Mackintosh NJ (1978). Classical conditioning in animals. *Annual Review of Psychology* **29**: 587-612.

Dickinson A, Dearing MF (1979). Appetitive-aversive interactions and inhibitory processes. In *Mechanisms of Learning and Motivation* (Dickinson A, Boakes RA eds), pp. 203-231. Erlbaum, Hillsdale, New Jersey.

Dinsmoor JA (2001). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *Journal of the Experimental Analysis of Behavior* **75**: 311-333.

Dobi A, Margolis EB, Wang HL, Harvey BK, Morales M (2010). Glutamatergic and nonglutamatergic neurons of the ventral tegmental area establish local synaptic contacts with dopaminergic and nondopaminergic neurons. *Journal of Neuroscience* **30**: 218-229.

Dommett E, Coizet V, Blaha CD, Martindale J, Lefebvre V, Walton N, Mayhew JE, Overton PG, Redgrave P (2005). How visual stimuli activate dopaminergic neurons at short latency. *Science* **307**: 1476-1479.

Estes WK, Skinner BF (1941). Some quantitative properties of anxiety. *Journal of Experimental Psychology* **29**: 390-400.

Fibiger HC, LePiane FG, Jakubovic A, Phillips AG (1987). The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *Journal of Neuroscience* **7**: 3888-3896.

Fiorino DF, Coury A, Fibiger HC, Phillips AG (1993). Electrical stimulation of reward sites in the ventral tegmental area increases dopamine transmission in the nucleus accumbens of the rat. *Behavioural Brain Research* **55**: 131-141.

Fiorino DF, Coury A, Phillips AG (1997). Dynamic changes in nucleus accumbens dopamine efflux during the Coolidge effect in male rats. *Journal of Neuroscience* **17**: 4849-4855.

Fiorillo CD, Tobler PN, Schultz W (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**: 1898-1902.

Floresco SB, West AR, Ash B, Moore H, Grace AA (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience* **6**: 968-973.

Frank MJ, Seeberger LC, O'Reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**: 1940-1943.

Gallistel CR, Beagley G (1971). Specificity of brain-stimulation reward in the rat. *Journal of Comparative and Physiological Pshchology* **69**: 713-721.

Gallistel CR, Shizgal P, Yeomans JS (1981). A portrait of the substrate for self-stimulation. *Psychologival Review* **88**: 228-273.

Gallistel CR. Dopamine and reward: comment on Hernandez et al (2006). *Behavioural Neuroscience* **120**: 992-994.

Gao DM, Jeaugey L, Pollak P, Benabid AL (1990). Intensity-dependent nociceptive responses from presumed dopaminergic neurons of the substantia nigra pars compacta in the rat and their modification by lateral habenula inputs. *Brain Research* **529**: 315-319.

Garris PA, Kilpatrick M, Bunin MA, Michael D, Walker QD, Wightman RM (1999). Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature* **398**: 67-69.

Geller I (1970). Effects of punishment on lever pressing maintained by food reward or brain stimulation. *Physiology and Behavior* **5**: 203-206.

Goldstein R (1966). Facilitation of active avoidance behavior by reinforcing septal stimulation in the rat. *Physiology and Behavior* **1**: 335-339.

Grace AA, Bunney BS (1984a). The control of firing pattern in nigral dopamine neurons: burst firing. *Journal of Neuroscience* **4**: 2877-2890.

Grace AA, Bunney BS (1984b). The control of firing pattern in nigral dopamine neurons: single spike firing. *Journal of Neuroscience* **4**: 2866-2876.

Graybiel AM (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience* **31**: 359-387.

Goldman Rakic PS, Lidow MS, Smiley JF, Williams MS (1992). The anatomy of dopamine in monkey and human prefrontal cortex. *Journal of Neural Transmission* **36**: 163-177.

Guarraci FA, Kapp BS (1999). An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behavioural Brain Research* **99**: 169-179.

Hearst E (1960). Simultaneous generalization gradients for appetitive and aversive Behavior. *Science* **132**: 1769-1770.

Hermer-Vazquez L, Hermer-Vazquez R, Rybinnik I, Greebel G, Keller R, Xu S, Chapin JK (2005). Rapid learning and flexible memory in "habit" tasks in rats trained with brain stimulation reward. *Physiology and Behavior* **84**: 753-759.

Hernandez G, Hamdani S, Rajabi H, Conover K, Stewart J, Arvanitogiannis A, Shizgal P (2006). Prolonged rewarding stimulation of the rat medial forebrain bundle: neurochemical and behavioral consequences. *Behavioral Neuroscience* **120**: 888-904.

Hernández G, Shizgal P (2009). Dynamic changes in dopamine tone during self-stimulation of the ventral tegmental area in rats. *Behavioural Brain Research* **198**: 91-97.

Hikosaka O, Sesack SR, Lecourtier L, Shepard PD (2008). Habenula: crossroad between the basal ganglia and the limbic system. *Journal of Neuroscience* **28**: 11825-11829.

Hineline PN (1972). Avoidance sessions as aversive events. *Science* **176**: 430-432.

Hong S, Hikosaka O (2008). The globus pallidus sends reward-related signals to the lateral habenula. *Neuron* **60**: 720-729.

Horvitz JC, Stewart T, Jacobs BL (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Research* **759**: 251-258.

Horvitz JC (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96**: 651-656.

Ilango A, Wetzel W, Scheich H, Ohl FW (2010). The combination of appetitive and aversive reinforcers and the nature of their interaction during auditory learning. *Neuroscience* **166**: 752-762.

Ikemoto S, Panksepp J (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews* **31**: 6-41.

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

Ikemoto S (2010). Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory. *Neuroscience and Biobehavioral Reviews*. 2010 Feb 10. [Epub ahead of print]

Jackson DM, Ahlenius S, Anden NE, Engel, J (1977). Antagonism by locally applied dopamine into the nucleus accumbens or the corpus striatum of a-methyltyrosine-induced disruption of conditioned avoidance behaviour. *Journal of Neural Transmission* **41**: 231-239.

Jhou TC, Fields HL, Baxter MG, Saper CB, Holland PC (2009). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61**: 786-800.

Ji H, Shepard PD (2007). Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *Journal of Neuroscience* **27**: 6923-6930.

Joel D, Weiner I (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* **96**: 451-474.

Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H (2008). Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *Journal of Neuroscience* **28**: 11673-11684.

Kalivas PW, Duffy P (1995). Selective activation of dopamine transmission in the shell of the nucleus accumbens by stress. *Brain Research* **675**: 325-328.

Kamin LK (1969). Predictability, surprise, attention and conditioning. In *Punishment and Aversive behavior* (Campell BA, Church RM, eds), pp. 279-296. Appleton Century Crofts, New York.

Kandel ER, Schwartz JH, Jessell TM (2000). *Principles of Neural Science*, McGraw-Hill, New York.

King D, Finlay JM (1997). Loss of dopamine terminals in the medial prefrontal cortex increased the ratio of DOPAC to DA in tissue of the nucleus accumbens shell: role of stress. *Brain Research* **767**: 192-200.

Kobayashi S, Schultz W (2008). Influence of reward delays on responses of dopamine neurons. *Journal of Neuroscience* **28**: 7837-7846.

Konorski J, Szwejkowska G (1956). Reciprocal transformations of heterogenous conditioned reflexes. *Acta Biologicae Experimentalis* **16**: 95-113.

Konorski J (1967). *The integrative activity of the brain: An interdisciplinary approach.* Chicago University Press, Chicago.

Lavin A, Nogueira L, Lapish CC, Wightman RM, Phillips PE, Seamans JK (2005). Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. *Journal of Neuroscience* **25**: 5013-5023.

Leknes S, Tracey I (2008). A common neurobiology for pain and pleasure. *Nature Review Neuroscience* **9**: 314-320.

Ljungberg T, Apicella P, Schultz W (1992). Responses of monkey dopamine neurons during learning of behavioural reactions. *Journal of Neurophysiology* **67**: 145-163.

Loskota WJ, Lomax P, Verity MA (1974). *A Stereotaxic atlas of the Mongolian gerbil brain (Meriones Unguiculatus).* Ann Arbor Science Publishers, Michigan, USA.

Mackintosh NJ (1983). *Conditioning and associative learning.* Oxford University Press, New York.

Magoon MA, Critchfield TS (2008). Concurrent schedules of positive and negative reinforcement: differential-impact and differential-outcomes hypotheses. *Journal of Experimental Analysis of Behavior* **90**:1-22.

Margulus DL, Olds J (1962). Identical "feeding" and "rewarding" systems. *Science* **135**: 374-375.

Margules DL, Stein L (1968). Facilitation of sidman avoidance behavior by positive brain stimulation. *Journal of Comparative Physiological Psychology* **66**:182-184.

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

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

Matsumoto M, Hikosaka O (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**: 837-841.

May PJ, McHaffie JG, Stanford TR, Jiang H, Costello MG, Coizet V, Hayes LM, Haber SN, Redgrave P (2009). Tectonigral projections in the primate: a pathway for pre-attentive sensory input to midbrain dopaminergic neurons. *European Journal of Neuroscience* **29**: 575-587.

McCullough LD, Sokolowski JD, Salamone JD (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* **52**: 919-925.

McHaffie JG, Jiang H, May PJ, Coizet V, Overton PG, Stein BE, Redgrave P (2006). A direct projection from superior colliculus to substantia nigra pars compacta in the cat. *Neuroscience* **138**: 221-234.

Miliaressis E, Rompre PP, Laviolette P, Philippe L, Coulombe D (1986). The curve-shift paradigm in self-stimulation. *Physiology and Behavior* **37**: 85-91.

Mirenowicz J, Schultz W (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**: 449-451.

Mogenson GJ, and Morrison MJ (1962). Avoidance responses to reward stimulation of the brain. *Journal of Comparative Physiological Psychology* **55**: 691-694.

Mogenson GD, Stevenson JAF (1966). Drinking and self-stimulation with electrical stimulation of the lateral hypothalamus. *Physiology and Behavior* **1**: 251-254.

Mogenson GJ, Takigawa M, Robertson A, Wu M (1979). Self-stimulation of the nucleus accumbens and ventral tegmental area of Tsai attenuated by microinjections of spiroperidol into the nucleus accumbens. *Brain Research* **171**: 247-259.

Mondadori C, Ornstein K, Waser PG, Huston JP (1976). Post-trial reinforcing hypothalamic stimulation can facilitate avoidance learning. *Neuroscience Letters* **2**: 183-187.

Morrison SE, Salzman CD (2009). The Convergence of information about rewarding and aversive stimuli in single neurons. *Journal of Neuroscience* **29**: 11471-11483.

Mora F, Avrith DB, Rolls ET (1980). An electrophysiological and behavioural study of self-stimulation in the orbitofrontal cortex of the rhesus monkey. *Brain Research Bulletin* **5**: 111-115.

Nair-Roberts RG, Chatelain-Badie SD, Benson E, White-Cooper H, Bolam JP, Ungless MA (2008). Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience* **152**: 1024-1031.

Nauta WJ (1958). Hippocampal projections and related neural pathways to the midbrain in the cat. *Brain* **81**: 319-340.

Oades RD, Halliday GM (1987). Ventral tegmental (A10) system: neurobiology. Anatomy and connectivity. *Brain Research* **434**: 117-165.

Ohl FW, Wetzel W, Wagner T, Rech A, Scheich H (1999). Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learning and Memory* **6**: 347-362.

Ohl FW, Scheich H, Freeman WJ (2001). Change in pattern of ongoing cortical activity with auditory category learning. *Nature* **412**: 733-736.

Olds J, Milner PM (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative Physiology and Psychology* **47**: 419-427.

Olds J (1958). Self-stimulation of the brain. *Science* **127**: 315-324.

Olds ME, Fobes JL (1981). The central basis of motivation: intracranial self-stimulation studies. *Annual Review of Psychology* **32**: 523-574.

Owesson-White CA, Cheer JF, Beyene M, Carelli RM, Wightman RM (2008). Dynamic changes in accumbens dopamine correlate with learning during intracranial self-stimulation. *Proceeding of the National Academy of Sciences* **105**: 11957-11962.

Panksepp J (1998). Seeking systems and anticipatory states of the nervous system. In *Affective Neuroscience*. pp. 144-163. Oxford University Press, Oxford.

Paxinos G, Franklin K (2001). The Mouse Brain in Stereotaxic Coordinates, second edition, Academic Press, San Diego.

Paxinos G and Watson C (2004). The Rat Brain in Stereotaxic Coordinates, fifth edition, Academic Press, New York.

Pecina S, Cagniard B, Berridge KC, Aldridge JW, Zhuang X (2003). Hyperdopaminergic mutant mice have higher wanting but not liking for sweet rewards. *Journal of Neuroscience* **23**: 9395-9402.

Pfaus JG, Damsma G, Nomikos GG, Wenkstern D, Blaha CD, Phillips AG, Fibiger HC (1990). Sexual behavior enhances central dopamine transmission in the male rat. *Brain Research* **530**: 345-348.

Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* **422**: 614-618.

Porrino LJ, Esposito RU, Seeger TF, Crane AM, Pert A, Sokoloff L (1984). Metabolic mapping of the brain during rewarding self-stimulation. *Science* **224**: 306-309.

Rada PV, Mark GP, Hoebel BG (1998). Dopamine release in the nucleus accumbens by hypothalamic stimulation- escape behavior. *Brain Research* **782**: 228-234.

Ranaldi R, Pocock D, Zereik R, Wise RA (1999). Dopamine fluctuations in the nucleus accumbens during maintenance, extinction, and reinstatement of intravenous D-amphetamine self-administration. *Journal of Neuroscience* **19**: 4102-4109.

Redgrave P, Prescott TJ, Gurney K (1999). Is the short-latency dopamine response too short to signal reward error? *Trends in Neuroscience* **22**: 146-151.

Redgrave P, Gurney K (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nature Review Neuroscience* **7**: 967-975.

Redgrave P, Gurney K, Reynolds J (2008). What is reinforced by phasic dopamine signals? *Brain Research Reviews* **58**: 322-339.

Reed P, Mitchell C, Nokes T (1996). Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Animal Learning and Behavior* **24**: 38-45.

Reynolds JN, Hyland BI, Wickens JR (2001). A cellular mechanism of reward-related learning. *Nature* **413**: 67-70.

Robinson DL, Phillips PE, Budygin EA, Trafton BJ, Garris PA, Wightman RM (2001). Subsecond changes in accumbal dopamine during sexual behavior in male rats. *Neuroreport* **12**: 2549-2552.

Roesch MR, Calu DJ, Schoenbaum G (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience* **10**: 1615-1624.

Roitman MF, Stuber GD, Phillips PEM, Wightman RM, Carelli RM (2004). Dopamine operates as a subsecond modulator of food seeking. *Journal of Neuroscience* **24**: 1265-1271.

Roitman MF, Wheeler RA, Carelli RM (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**: 587-597.

Roitman MF, Wheeler RA, Wightman RM, Carelli RM (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nature Neuroscience* **12**: 1376-1377.

Rolls ET (2000). Memory systems in the brain. *Annual Review of Psychology* **51**: 599-630.

Rolls ET (2005). Brain stimulation reward. In *Emotion Explained.* pp. 288-306. Oxford University Press, Oxford.

Sadowski B, Marek P, Panocka I (1984). Enhancement of performance for brain stimulation reward after footshock in rats. *Acta Neurobiol Exp (Wars)* **44**: 51-59.

Salamone JD, Cousins MS, McCullough LD, Carriero DL, Berkowitz RJ (1994). Nucleus accumbens dopamine release increases during instrumental lever pressing for food but not free food consumption. *Pharmacology Biochemistry and Behavior* **49**: 25-31.

Salamone JD (1994). The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. *Behavioural Brain Research* **61**: 117-133.

Salamone JD, Correa M (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research* **137**: 3-25.

Scavio MJ (1974). Classical – classical transfer: Effects of prior aversive conditioning upon appetitive conditioning in rabbits. *Journal of Comparative Physiological Psychology* **86**: 107-115.

Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience* **1**:155-159.

Schultz W, Romo R (1987). Responses of nigrostraital dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *Journal of Neurophysiol*ogy **57**: 201-217.

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

Schultz W (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* **80**: 1-27.

Schultz W, Dickinson A (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience* **23**: 473-500.

Shumake J, Ilango A, Scheich H, Wetzel W, Ohl FW (2010). Differential neuromodulation of acquisition and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. *Journal of Neuroscience* **30**: 5876-5883.

Simmons JM, Ackermann RF, Gallistel CR (1998). MFB lesions fail to structurally and functionally disconnect the VTA from many ipsilateral forebrain nuclei: Implications for the neural substrate for brain stimulation reward. *Journal of Neuroscience* **18**: 8515-8533.

Skinner BF (1938). *The Behavior of organisms: an experimental analysis.* Appleton, New York.

Sokolowski JD, McCullough LD, Salamone JD (1994). Effects of dopamine depletions in the medial prefrontal cortex on active avoidance and escape in the rat. *Brain Research* **65**: 293-299.

Stark H, Scheich H (1997). Dopaminergic and serotonergic neurotransmission systems are differentially involved in auditory cortex learning: a long-term microdialysis study of metabolites. *Journal of Neurochemistry* **68**: 691-697.

Stark H, Bischof A, Scheich H (1999). Increase of extracellular dopamine in prefrontal cortex of gerbils during acquisition of the avoidance strategy in the shuttle-box. *Neuroscience Letters* **264**: 77-80.

Stark H, Bischof A, Wagner T, Scheich H (2001). Activation of the dopaminergic system of medial prefrontal cortex of gerbils during formation of relevant associations for the avoidance strategy in the shuttle-box. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **25**: 409-426.

Stark H, Rothe T, Wagner T, Scheich H (2004). Learning a new behavioral strategy in the shuttle-box increases prefrontal dopamine. *Neuroscience* **126**: 21-29.

Steffensen SC, Lee RS, Stobbs SH, Henriksen SJ (2001). Responses of ventral tegmental area GABA neurons to brain stimulation reward. *Brain Research* **906**: 190-197.

Stein L (1965). Facilitation of avoidance behaviour by positive brain stimulation. *Journal of Comparative and Physiological Psychology* **60**: 9-19.

Stuber GD, Klanker M, de Ridder B, Bowers MS, Joosten RN, Feenstra MG, Bonci A (2008). Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. *Science* **321**: 1690-1692.

Stuber GD, Wightmann RM, Carelli RM (2005). Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. *Neuron* **46**: 661-669.

Szczypka MS, Rainey MA, Kim DS, Alaynick WA, Marck BT, Matsumoto AM, Palmiter RD (1999). Feeding behavior in dopamine-deficient mice. *Proceeding of the National Academy of Sciences* **96**: 12138-12143.

Talwar SK, Xu S, Hawley ES, Weiss SA, Moxon KA, Chapin JK (2002). Rat navigation guided by remote control. *Nature* **417**: 37-38.

Tanimoto H, Heisenberg M, Gerber B (2004). Event timing turns punishment to reward. *Nature* **430**: 983.

Thiessen D (1977). *The gerbil in behavioral investigations*. University of Texas Press, USA.

Tobler PN, Dickinson A, Schultz W (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *Journal of Neuroscience* **23**: 10402-10410.

Tobler PN, Fiorillo CD, Schultz W (2005). Adaptive coding of reward value by dopamine neurons. *Science* **307**: 1642-1645.

Tsai CT, Nakamura S, Iwama K (1980). Inhibition of neuronal activity of the substantia nigra by noxious stimuli and its modification by the caudate nucleus. *Brain Research* **195**: 299-311.

Ungless MA, Magill PJ, Bolam JP (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* **303**: 2040-2042.

Valenstein ES, Valenstein T (1964). Interaction of positive and negative reinforcing neural systems. *Science* **145:** 1456-1458.

Wadenberg ML, Ericson E, Magnusson O, Ahlenius S (1990). Suppression of conditioned avoidance behavior by the local application of (−) sulpiride into the ventral, but not the dorsal, striatum of the rat. *Biological Psychiatry* **28**: 297-307.

Wenksten D, Pfaus JG, Fibiger HC (1993). Dopamine transmission increases in the nucleus accumbens of male rats during their first exposure to sexually receptive female rats. *Brain Research* **618**: 41-46.

Wetzel W, Wagner T, Ohl FW, Scheich H (1998). Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbils. *Behavioural Brain Research* **91**: 29-39.

Wetzel W, Ohl FW, Scheich H (2008). Global versus local processing of frequency modulated tones in gerbils: an animal model of lateralized auditory cortex functions. *Proceeding of the National Academy of Sciences* **105**: 6753-6758.

Wilkinson LS, Humby T, Killcross AS, Torres EM, Everitt BJ, Robbins TW (1998). Dissociations in dopamine release in medial prefrontal cortex and ventral striatum during the acquisition and extinction of classical aversive conditioning in the rat. *European Journal of Neuroscience* **10**: 1019-1026.

Williams GV, Millar J (1990). Concentration-dependent actions of stimulated dopamine release on neuronal activity in rat striatum. *Neuroscience* **39**: 1-16.

Wise RA (1980). The dopamine synapse and the notion of pleasure centers in the brain. *Trends in Neuroscience* **3**: 91-94.

Wise RA (1996). Addictive drugs and brain stimulation reward. *Annual Review of Neuroscience* **19**: 319-340.

Wise RA (2002). Brain reward circuitry: insights from unsensed incentives. *Neuron* **36**: 229-240.

Wise RA (2004). Dopamine, learning and motivation. *Nature Review Neuroscience* **6**: 483-494.

Wolgin DL, Wade JV (1990). Effect of lithium chloride-induced aversion on appetitive and consummatory behavior. *Behavioral Neuroscience* **104**: 438-440.

Wu M, Hrycyshyn AW, Brudzynski SM (1996). Subpallidal outputs to the nucleus accumbens and ventral tegmental area: anatomical and electrophysiological studies. *Brain Research* **740**: 151-161.

Young AM, Ahier RG, Upton RL, Joseph MH, Gray JA (1998). Increased extracellular dopamine in the nucleus accumbens of the rat during associative learning of neutral stimuli. *Neuroscience* **83**: 1175-1183.

Young AM (2004). Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1min microdialysis in rats. *Journal of Neuroscience Methods* **138**: 57-63.

Zhou QY, Palmiter RD (1995). Dopamine-deficient mice are severely hypoactive, adipsic, and aphagic. *Cell* **83**: 1197-1209.

Thema: Die Integration appetitiver und aversiver Verstärker und die Neuromodulation von Belohnungsstreben und Schmerzvermeidung

Zusammenfassung der Dissertation

Neues Verhalten – bei Tieren wie auch beim Menschen – kann beim Lernen sowohl durch Belohnung als auch durch Bestrafung erworben werden. Besonders gut gelernt wird, wenn Belohnung und Bestrafung ("Zuckerbrot und Peitsche") kombiniert werden. Dieser Zusammenhang ist zwar allgemein geläufig, von großem theoretischen Interesse und großer praktischer Bedeutung, im Detail sind aber die den Wechselwirkungen von Belohnung und Bestrafung (Verstärkung) zugrunde liegenden Prozesse erst wenig bekannt. Ein wesentliche Rolle beim verstärkungsinduzierten Lernen wird dem dopaminergen System, insbesondere der ventralen tegmentalen area (VTA) im Mittelhirn zugeschrieben. Stimulation dieses Kerngebietes hat einen positiven, d.h. belohnungswirksamen, Effekt auf das Verhalten. In der vorliegenden Arbeit wurde beim auditorischen Lernen von Rennmäusen in der Shuttle-Box die elektrische VTA-Stimulation als appetitive Verstärkung (Belohnung) und die Applikation elektrischer Fußreize als aversive Verstärkung (Bestrafung) verwendet. Es zeigte sich, daß die Kombination beider Verstärkungsarten im gleichen Experiment, d.h. Belohnung für richtige und Bestrafung für falsche Antworten, einen deutlich stärkeren Effekt auf das Lernen hatte ("Superlearning") als die appetitive bzw. aversive Verstärkung allein. Weitere Untersuchungen an Tieren, die mit kombinierter Verstärkung gelernt hatten, ergaben, daß die Lernleistung durch Weglassen jeweils eines Verstärkers in unterschiedlichem Maße verändert wird und dass dies von der Häufigkeit der Verstärkungsereignisse (kontinuierliche bzw. partielle Verstärkung) abhängig ist. Diese Ergebnisse zeigen, dass beim Erlernen und Aufrechterhalten bzw. Auslöschen des Gelernten die Verstärkung durch Belohnung und Bestrafung in unterschiedlicher Weise zusammenwirken und dass in beiden Fällen, d.h. beim

Streben nach Belohnung bzw. bei der Vermeidung von Bestrafung, dem dopaminergen System eine entscheidende Bedeutung zukommt.

List of publications

Ilango A, Wetzel W, Scheich H, Ohl FW (2010). The combination of appetitive and aversive reinforcers and the nature of their interaction during auditory learning. *Neuroscience* **166**: 752-762.

Shumake J, **Ilango A**, Scheich H, Wetzel W, Ohl FW (2010). Differential neuromodulation of acquisition and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. *Journal of Neuroscience* **30**: 5876-5883.

Goldschmidt J, Wanger T, Engelhorn A, Friedrich H, Happel M, **Ilango A**, Engelmann M, Stuermer IW, Ohl FW, Scheich H (2010). High-resolution mapping of neuronal activity using the lipophilic thallium chelate complex TlDDC: protocol and validation of the method. *Neuroimage* **49**: 303-315.

Curriculum Vitae

Personal Details

Name	Anton Ilango Micheal
Address	Weitling Strasse 25, 39104, Magdeburg
Date of birth	03. May 1978
Place of birth	Tirunelveli, India
Marital status	Married
Nationality	Indian

Academic Details

Since Jan 2004	Doctoral student in Leibniz Institute for Neurobiology, Magdeburg under Otto-von-Guericke Universitat.
June 1999 to May 2001	Studied 'Master of Science (Bio-Technology)' in Loyola college, University of Madras, Chennai, Tamilnadu, India.
June 1996 to Apr 1999	Studied 'Bachelor of Science (Zoology)' in St.Xavier's college, Manonmanium Sundaranar University, Palayamkottai, Tamilnadu, India.
June 1993 to Mar 1995	Higher secondary school education
Apr 1993	Completed secondary school education

Professional Details

March 2003 to Nov 2003	Worked as a 'Wissenschaftlicher Mitarbeiter' in Institute of Physiology and Pathophysiology, Johannes Gutenberg University, Mainz.

Scholarship

Selected for Tamilnadu state government aid for four years of schooling by talent search examination.

<div align="right">Anton Ilango Micheal</div>

Erklärung

Hiermit erkläre ich, dass ich die von mir eingereichte Dissertation mit dem Thema

"**Integration of appetitive and aversive reinforcers and the neuromodulation of reward seeking and pain avoidance**"

selbständig verfasst, nicht schon als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.

Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Anton Ilango Micheal

Magdeburg, 22.06.2010   Unterschrift