RAPPORT DE STAGE D'OPTION SCIENTIFIQUE

Rewards, Punishment and the Brain

Y-Lan Boureau
Ecole Polytechnique, Promotion X2002

Directeur de stage : Peter Dayan

Gatsby Computational Neuroscience Unit
University College London

· 2005 ·

NON CONFIDENTIEL

Option : Biologie
Directeur d'Option : Sylvain Blanquet

Directeur de stage : Peter Dayan
Gatsby Computational Neuroscience Unit
Alexandra House
17 Queen Square
LONDON, WC1N 3AR, UK

Résumé

Ce travail s'intéresse aux interactions entre facilitation et inhibition des réponses comportementales. Généralement, les récompenses suscitent l'approche et sont associées à une facilitation des réponses, tandis que les punitions provoquent l'évitement et inhibent les réponses comportementales autres que la fuite ou la lutte.

La dopamine (DA) et la sérotonine (5HT) ont été largement associées à la régulation facilitatrice et inhibitrice, respectivement, du comportement, ou encore, à l'approche et au retrait. La dopamine facilite l'action, et renforce les associations action-résultat, stimulus-résultat, et stimulus-réponse. La sérotonine a généralement une action inhibitrice sur les réponses liées à la récompense, mais peut aussi les faciliter ; l'existence d'au moins 14 types de réceptors de la sérotonine donne lieu à une vaste gamme d'effets complémentaires (et parfois apparemment contradictoires). La sérotonine et la dopamine interagissent au niveau de leurs nombreuses cibles communes dans le cerveau, et la sérotonine régule aussi la dopamine au niveau des corps cellulaires des neurones. L'influence ambiguë de la sérotonine sur le système dopaminergique rappelle son influence sur le comportement, en ce qu'elle est généralement inhibitrice mais parfois excitatrice.

La punition constitue un contexte d'interactions accrues entre approche et retrait, aussi bien qu'entre sérotonine et dopamine. En conséquence, l'étude de ce qui se passe dans ce cadre peut permettre de clarifier le sens général de ces interactions.

Nous proposons un modèle du comportement pendant la punition, ainsi qu'une interprétation en terme de neuromodulateurs. Au coeur du modèle se trouve l'hypothèse que le meilleur moyen de se sortir de la punition est d'étouffer la plupart des récompenses, tout en amplifiant les effets de celles qui sont passées à travers le filtre ; les stratégies pour faire face seraient ainsi mieux appliquées et apprises. Le modèle attribue à la dopamine sa fonction habituelle facilitant l'action, et la sérotonine est présentée comme régulant l'approche et l'inhibition à un niveau hiérarchiquement supérieur, en sculptant la fonction dopaminergique.

Abstract

This work is about interactions between facilitation and inhibition of behavioural responses. Generally, rewards elicit approach and are associated with facilitation of responses, whereas punishments elicit avoidance and inhibit behavioural responses other than fight or flight.

Dopamine (DA) and serotonine (5HT) have been widely associated with respectively facilitatory and inhibitory regulation of behaviour, or approach and withdrawal. Dopamine facilitates action (drive effect), and reinforces action-outcome, stimulus-outcome and stimulus-response associations (learning effect). Serotonin generally inhibits reward-related responding, but can also facilitate it ; there being at least 14 subtypes of serotonin receptors allows for a vast range of complementary (and sometimes seemingly contradictory) effects. Serotonin and dopamine interact at the level of their numerous shared target areas in the brain, and serotonin also regulate dopamine at the level of cell bodies. The ambiguous influence of serotonin on the dopaminergic system parallels its effects on behaviour ; that is, generally inhibitory but sometimes excitatory.

Punishment provides a context of increased interactions between approach and withdrawal, as well as serotonin and dopamine. Therefore, studying what happens during punishment can provide insight on the general meaning of these interactions.

Here, we provide a model of behaviour during punishment and propose an interpretation in terms of neuromodulators. The core of the model is the hypothesis that the best way to cope during punishment is to dampen most rewards, while overemphasizing the effect of the spared ones ; coping strategies would thus be better implemented and learnt. Dopamine is attributed its usual facilitatory role on action, and serotonin is pictured as regulating approach and inhibition on a higher hierarchical level by carving dopamine function.

Acknowledgments

I am very grateful to Peter for hosting me at Gatsby, providing ready and insightful advice, and reading my messy emails full of receptors.

I would also like to thank all people at Gatsby, especially Quentin and Mátè for everyday support, Peggy for helping me in my attempts to turn Gatsby into a French-speaking place, Yael for her good-humour and skills for organizing things, Boris for buying the sofa-chair in Nathaniel's office, Nathaniel for allowing me to use it, the other Peter too because he's good company, and Iain for ready advice about how to use a computer. And Ben even if he's not at Gatsby.

I am also grateful to Peter Shizgal for kindly sending me a poster with his results.

Thanks to my parents too for assuring delivery of this report.

Contents

Chapter 1

Introduction

> A horse ! a horse ! my kingdom for a horse !
>
> *King Richard III,* Act V, Scene IV
> SHAKESPEARE

Is a horse worth a kingdom ?

Sometimes it is, and knowing when is critical for adaptive behaviour.
Rules for organizing behaviour have to be flexible, as organisms that use them live in a perpetually and unpredictably changing environment, which results in a tradeoff between the efficacy of rigid rules, and the adaptiveness of flexible ones. Behaviour has to be at the same time reliable and versatile.

Grossly, behaviour is organized along two broad directions : avoiding all that reduces the fitness of the organism, and seeking what increases it. Animals accordingly build a bipolar mapping of their environment : the attractive pole of the rewards, and the repulsive one of the punishments. Primary rewards are directly connected to fitness, and roughly boil down to sex, food and water ; primary punishments like pain signal a decline in the access to these rewards.

But in order to maximize their access to primary rewards, complex organisms like animals learn about responses that can improve or worsen this access, and consequently extend the appetitive properties of rewards to all stimuli that can somehow lead to them, as well as avoiding stimuli that have led to primary punishments in the past. Thus, the range of rewards and punishments is considerably broadened, from water to its sight, and then the sound of a spring ; primary rewards elicit *consummatory* behaviour (eating, drinking, fornicating), secondary rewards elicit approach - that is, *preparatory* behaviour. Conversely, primary punishment triggers flight, fight, or freezing, and secondary punishment elicits withdrawal.

Complexity arises from the fact that all the instructions determined by rewards and punishment are superimposed on top of one another, and often in a contradictory fashion ; should a button that delivers both an electrical shock to your finger (punishment) and a squirt of sugar water (reward) in your mouth be pressed ? The intricate causal mapping of actions to their consequences over time makes decision making no simple matter.

Neuromodulators (acetylcholine, norepinephrine, dopamine, serotonin) are deeply involved in the processes that underlie this kind of decision making, critically influencing both action and learning. Here, we focus on dopamine and serotonin, which are critically implicated in a number of neurological, psychiatric and psycho-social disorders (eg. depression, schizophrenia), and are still poorly understood despite active and abundant research.

Dopamine has long been attributed a tight relationship with reward (reviewed in (Wise, 2004; Berridge & Robinson, 1998)), and this proved robust against various attacks (Salamone *et al.*, 2005; Horvitz, 2000). Serotonin, beside its involvement in a wide range of processes (Lucki, 1998), has been hypothesized to play an important part in adapting inborn punishment responding (Deakin & Graeff, 1991). In mammalian brains, they share common target areas and interact directly ; a number of findings show that serotonergic function can inhibit dopaminergic function, which suggests dopamine and serotonin behave as opponents in reward and punishment processing (Daw *et al.*, 2002).

A hasty interpretation of these theories might lead one to equate dopamine with reward and serotonin with punishment, with the coexistence of both systems simply making up for the impossibility for neurons to fire negatively. However, this simplistic approach does not accommodate experimental evidence (eg. the elevated dopaminergic function during punishment) - nor is it adequate in the first place to consider punishment and reward as merely, the negative and positive parts of a common scale(Solomon & Corbit, 1974).

Punishment provides a convenient frame to study the links between aversively and appetitively motivated behaviour, as well as the interactions of dopamine and serotonin in a context that dramatically motivates both action and inhibition of action. Understanding the nature of coping behaviour during punishment should provide insight on the nature of the relation between serotonin and dopamine.

In this report, we provide a normative model for how animals should reshape their preferences when undergoing punishment (eg. making a horse worth more than a kingdom), and how this leads to a choice of coping responses (eg. flight). We then propose a mapping of features of the model onto the neuromodulatory system, and show how the sometimes confusing observed data on neuromodulatory function during punishment could reflect an optimization of coping behaviour as accounted for in the proposed behavioural model.

The first part of the report provides some background on reward, punishment, and the related roles of dopamine and serotonin, as well as their interaction. The second part introduces the model of preferences during punishment and the postulated mapping onto neuromodulators, and shows how this could reconcile a number of somewhat puzzling experimental findings. The third part states the contributions of this work to the field, and provides directions for further investigation.

Chapter 2

Some background on reward, punishment, and their links with neuromodulators

2.1 Reward

This section provide the basics about conditioning and establishes the deep involvement of DA in reward processing. The role of 5HT is also examined.

2.1.1 Behavioural data - Conditioning and reinforcement learning

The study of behavioural response to reward relies highly on conditioning experiments.
Appetitive conditioning consists in repeated pairings of an originally motivationally neutral stimulus (eg. a light), the *conditioned stimulus* (CS), with an already established reward (the *unconditioned stimulus*, or US, also called *reinforcer*).

Turning a light into a reward

Through *incentive learning*, the CS is endowed with rewarding properties - that is, the capacity to elicit approach and to be worked for. Incentive learning can be isolatedly studied in *classical* (or *Pavlovian*) conditioning, where CS and US are paired without requiring any response from the animal.

Even stimuli perceived as unconditioned rewards, like the smell of food, actually require incentive learning ; according to Pavlov (Pavlov, 1927), the corresponding unconditioned stimulus would be the "chemical properties of the food itself acting upon receptors in the mucous membrane of the mouth and tongue".

Another interesting conditioning paradigm to assess incentive learning is the *place preference conditioning* paradigm, where an animal develops a preference for a place where it has received rewards. Once learning of a CS has been achieved, the CS is capable of conditioning a place preference.

Stimulus-response strengthening - response reinforcing

Beside incentive learning, conditioning can also strengthen the link between a stimulus and a response (S-R), through *instrumental* conditioning, which lies at the core of habits. Instrumental conditioning requires a behavioural response from the animal (eg. a lever press) for the reward to be delivered ; each pairing of an S-R pair with a reward delivery strenghtens the S-R link.

Contrary to a Stimulus - Outcome link (as in incentive learning), the S-R link can be thought of as an 'outcome-blind' link, as it seems to retain no memory of the reward that reinforced it. This can be assessed by *devaluation* trials (Dickinson & Balleine, 2002), where the response is still performed at the presentation of the stimulus, while the reinforcer that was used to learn the S-R association is not worked for any more. An 'outcome-blind' association like S-R is insensitive to devaluation, whereas what could be called an 'outcome-flavoured' association like Action-outcome or stimulus-outcome has a strength that varies according to the motivational meaning of the reinforcer.

Drive and incentive motivation

Presentation of a reward (the primer) also makes an animal eager to work in subsequent trials for the *same* primary reward as the one associated with the primer, and to a lesser extent, for all kinds of rewards.

Extinction of conditioning

Conditioned rewards are rewards only insofar as they point to evolutionarily meaningful rewards. Thus, it is possible to extinguish a conditioned reward, by presenting it repeatedly without the reinforcer.

Similarly, a S-R association needs to be constantly reinforced to be maintained, so that repeated unreinforced presentations of a stimulus that has been paired with a response weaken the S-R association.

2.1.2 Dopamine and reward

Briefly, brain dopaminergic system consists of DA neurons concentrated in closely situated ventral tegmental area (VTA) and substantia nigra (SN), that project to a number of areas, notably the nucleus accumbens (NAc), the prefrontal cortex (PFC), the amygdala, the striatum. DA cells can fire in two distinct modes, by single spikes or by bursts (Grace & Bunney, 1984b,a), which gives way to a distinction between *tonic* firing, at about 4Hz, like a continuous pacemaker background, and *phasic* firing, which can be triggered by punctual stimuli ; only a subpopulation of DA cells are active (Floresco *et al.*, 2003), some of them bursting, others not. There are 5 DA receptors subtypes, the most familiar of which are D1 and D2 (Missale *et al.*, 1998).

Dopaminergic activity seems to be profoundly related to each of the reward-related behavioural features that we have listed in last paragraph. In order to examine the role of dopaminergic function in reward processing, a vast range of experimental procedures are available. It is possible to measure dopaminergic activity with electrophysiology, microdialysis, voltammetry(see Appendix 1 for an analysis of their respective advantages and downsides). Another approach consists in examining the behavioural consequences of dopaminergic manipulations, either 1)through lesions of DA cells, 2)with genetically-engineered mice (eg. DD-mice, that are incapable of synthetizing DA), or 3)by pharmacological manipulations that inhibit or enhance dopaminergic function.

All these methods converge to the conclusion that DA is deeply involved in reward-related processing. We hereafter provide some of the most striking examples.

Clues from electrophysiology - Temporal difference, phasic DA firing and the prediction of reward

One of the most dramatic pieces of evidence of an involvement of DA in reward processing comes from electrophysiology. Immediately after an unpredicted reward has been given, DA neurons have been observed to fire phasically (fig. 2.1). In the course of conditioning, as the pairings between CS and US are repeated, this peak of DA phasic firing transfers from US to CS, and this transfer correlates with the attribution of rewarding properties to the CS.

Thus, DA phasic firing seems to signal *an error in the prediction of reward* : learning of the contingency between the CS and the US implies that 1) the US comes to be predicted by the CS, and consequently becomes *less unpredicted*, and 2) the CS (which is unpredicted anyhow) turns more and more into a reward. DA phasic firing reflects exactly that evolution.

Conversely, in trials where the predicted reward is omitted, DA cells depress their firing at the time when the reward was expected.

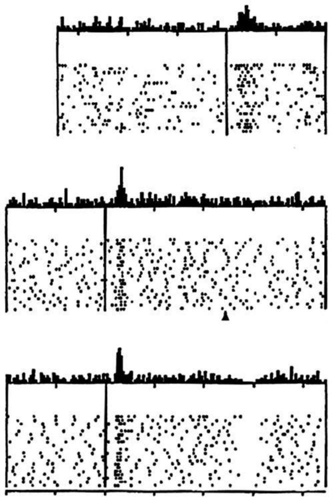


Figure 2.1: Accross trials, Dopamine cells phasic firing shifts from the reward to the predicting conditioned stimulus. If the expected reward fails to be delivered, a depression of Dopamine cells firing is observed instead at the time of the missed predicted reward.Figure from (Schultz *et al.*, 1997)

These observations have led to the suggestion that DA phasic firing encodes a reward prediction error, as used in Temporal Difference (TD) algorithms (Schultz *et al.*, 1997). Thus, it has been possible to use TD algorithms to infer DA signal and look for a correlation with this signal in fMRI studies to identify the neural substrates of reward in the brain (O'Doherty *et al.*, 2003, 2004), as the Nucleus Accumbens (NAc) and Ventral Putamen.

Dopamine and reinforcement

As reviewed in (Wise, 2004), DA seems to be essential to learn and maintain both the stimulus-outcome association and stimulus-response automatic link. Under neuroleptics, which decrease DA function, learning is impaired for both Pavlovian and instrumental conditioning to new CS : neuroleptic-treated animals do not learn to work for food, water, cocaine injections, etc, and fail to develop place preferences conditioned by these rewards. On the other hand, already learnt associations seem to hold, in that neuroleptic-treated animals continue to perform the learnt response ; however, if DA reduced function is maintained over time, the responses decrease and finally disappear.

Moreover, trials under neuroleptic display striking similarities to extinction trials (where the reward is not given) (Wise, 2004), which provides further evidence that DA is needed to learn and maintain associations between stimulus and response or outcome.

Drive

Beside those reinforcing effects, DA also enhances drive and facilitates action, which links it to reward-seeking behaviour (Phillips *et al.*, 2003). Tonic levels of DA measured by microdialysis correlate with levels of activity during rewarding brain stimulation (fig. 2.3) ; transients of DA in the Nucleus Accumbens precede the lever-press response (fig. 2.2);(Stuber *et al.*, 2005; Roitman *et al.*, 2004). DA-releasing drugs like amphetamine are associated with hyperactivity.

2.1.3 Effects of Serotonin

5HT cells are located in more diffuse sources than DA cells, but mainly from midbrain raphe nuclei (that is, dorsal raphe nucleus, or DRN, and median raphe nucleus, or MRN), and project virtually

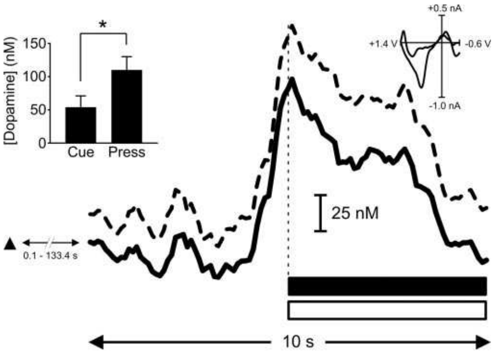


Figure 2.2: Extracellular peak of Dopamine in the core of the Nucleus Accumbens preceding the lever-press response for food. The lever press response occurs at the dotted line. Figure from(Roitman *et al.*, 2004)

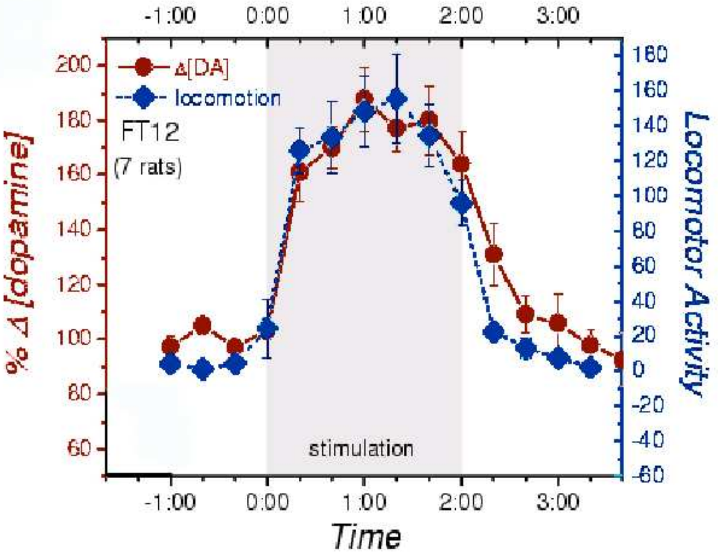


Figure 2.3: Activity correlates with Dopamine concentration during 2-hour brain stimulation (darkened zone). Figure from(Hamdani *et al.*, 2004)

everywhere. 5HT seems to be evolutionarily the oldest neurotransmitter int the brain, and uses at least 14 receptor subtypes (Barnes & Sharp, 1999).

Drugs that unspecifically enhance 5HT function, like fluoxetine (a Serotonin Specific Reuptake Inhibitor, or SSRI) generally decrease reinforced responding, with reinforcer being either natural rewards (water, sucrose) or drugs(amphetamine and cocaine administration). Conversely, 5HT depletion increases responding to water and sucrose cues.(Higgins & Fletcher, 2003)

Thus, 5HT often seems to oppose Dopamine action in reward ; more generally, Dopamine mediates approach behaviour, while 5HT seems crucial for retardation or withholding of response (Winstanley *et al.*, 2005). In cognitive tasks, 5HT depletion is associated with perseverative responding. The view of 5HT as an opponent to Dopamine is further strengthened by their sharing the same anatomical targets (notably the Nucleus Accumbens) (Daw *et al.*, 2002).

However, when infused locally into the NAc, 5HT increases responding without discriminating between unrewarded and rewarded response, and causes increased motor activity, an effect which is partially blocked by DA antagonist administration (Sasaki-Adams & Kelley, 2001).

In order to understand how contradictory effects can be associated with 5HT, it is necessary to bear in mind that 5HT acts through at least 14 receptor subtypes (Barnes & Sharp, 1999), which allows a fine regulation of 5HT effects through both the localization of particular receptors and their different sensitivity, and might explain how 5HT can be involved in such a wide number of physiological and behavioural processes (Lucki, 1998).

Thus, systemic administration of agonists of 5HT2a receptor, a particular subtype of 5HT receptor, causes hyperactivity in rats, whereas 5HT2a receptor antagonists decrease impulsive-type behaviour. The opposite pattern is observed with agonists and antagonists of 5HT2c receptor. For example, MDMA (ecstasy)-induced hyperactivity is potentiated by antagonists of 5HT2c receptor, and blocked by antagonists of 5HT2a receptor (Fletcher *et al.*, 2002). 5HT2c receptor agonists also decrease robustly cocaine self-administration, food-maintained responding, and nicotine or ethanol oral self-administration, an effect which is blocked by a specific 5HT2c receptor antagonist pretreatment (Higgins & Fletcher, 2003). Similarly, in cognitive tasks, 5HT2a agonists increase impulsive responding, whereas 5HT is generally associated with the faculty to differ responding (Koskinen *et al.*, 2003).

Another 5HT receptor subtype seems to be important in reward processing : systemic administration of an antagonist of 5HT3 receptor in rats decrease morphine- and nicotine- conditioned place preferences, as well as ethanol oral self-administration. (Higgins & Fletcher, 2003)

Locally, injection of a 5HT3 receptor antagonist in the NAc prevents the hyperactivity induced by intra-NAc infusion of exogenous DA, whereas a 5HT3 receptor agonist potentiates it.(Higgins & Fletcher, 2003)

So 5HT can locally enhance drive, seemingly indirectly by acting on DA function.

2.2 Punishment - general data

This section provides some background on behavioural response to punishment, then examines the proposed relations between 5HT and the fear response. Finally, a summary of the effects of punishment on DA function is given.

2.2.1 Behavioural data

Aversive conditioning

Aversive conditioning paradigms share quite a few properties with appetitive conditioning, only replacing 'reward' with 'punishment'- eg., water by electrical shocks ; the same vocabulary (CS, US, etc) is used.

However, some differences are worth mentioning ; learning is usually achieved far faster than with appetitive conditioning, in a few trials. Conversely, extinction is harder to achieve – that is, an animal will far more easily stop paying attention to a light that used to predict water, than one that used to predict an electrical shock.

Popular punishments

Experimenters use a wide range of artificial primary punishments. First, various electrical shocks can be used. They are efficiently aversive, and can be applied to the tail or to the paws of the rats, by electrifying the bars of the cage. (Bland *et al.*, 2004b)
Restraint (Cabib & Puglisi-Allegra, 1996), which only consists in mechanically preventing the animal from moving, is also aversive, as well as mechanical tailpinch (Ungless *et al.*, 2004).
Another useful paradigm is *forced swimming*, where the animal is put in a water-filled cylinder, and can either float, try and climb along the wall of the cylinder, or swim.
Direct stimulations of parts of the brain involved in fear, like the periacqueductal gray (PAG), are also highly aversive (Vianna *et al.*, 2001).

Inborn responses to punishment

The range of inborn responses to punishment is quite universal, and roughly reduces to the 'flight/fight/freezing system' theorised by Gray (Gray & McNaughton, 2000). Of these three responses, *Freezing* was added only recently to the Cannon's canonic 'fight or flight' couple, and is now considered a true part of fear responses. It consists in a state of hypervigilant immobility, with enhanced attention. Freezing can be thought of as a deliberational state, to determine whether the situation is aversive enough to justify a more active response, or to prepare those responses (eg. animals 'playing dead' and expecting for an opportunity to escape). Interestingly, the behavioural sequence usually observed naturally (initial freezing, and then fight) is also produced artificially by direct electrical stimulation of the PAG : stimulations first elicit freezing, and if they are made stronger, wild running and jumping (Vianna *et al.*, 2001).

Finally, a non-response to punishment is passive immobility, eg. mice in a forced swimming test floating on the water without moving. We will examine this type of failure to cope more profoundly in next section, in relation with uncontrollable punishment.

What happens to the other behavioural responses ?

Punishment has an overall disruptive effect on behaviour. Animals stop grooming and display a stressed response (increased urination and defecation) ; as for conditioned fear - that is, fear evoked by a neutral stimulus that has been trained by conditioning -, it robustly causes an interruption of ongoing behaviour, eg. a suppression of lever-pressing for food, to the point that suppression is used as a measure of learning of fear conditioning.

However, some highly appetitive behavioural responses can be boosted, and subsequently the incentive value of a stimulus can be increased, in an aversive context. Thus, sexual behaviour is increased during tailpinch : not only does the rat display more active sexual behaviour (increased mounting and intromission behaviour), but the incentive value of the female stays higher even once tailpinch is over(Leyton & Stewart, 1996).

2.2.2 Fear and Serotonin

Serotonin and adaptation of defense behaviour

Deakin and Graeff (Graeff *et al.*, 1997; Deakin & Graeff, 1991) have proposed a role for 5HT in the adaptive modulation of defense, where 5HT shapes bidirectionally innate behaviour, by inhibiting inborn flight when it is unjustified, and eliciting inhibitory avoidance when a stimulus considered neutral has been paired with punishment. In fact, administration of a 5HT releaser results in both impaired escape from an unconditionned aversive situation (inhibition of innate fear) and increased conditioned avoidance (enhancement of learned fear) (Graeff *et al.*, 1997; Zangrossi *et al.*, 2001).

5HT also fits well with Gray's theory (Gray & McNaughton, 2000) of the behavioural inhibition system (BIS) as part of a group of three complementary systems (the other two being the behavioural approach system, or BAS, and the fight-flight-freezing system, or FFFS). According to Gray, the BIS intervenes to regulate conflicts between BAS and FFFS, and arbitrate between approach and escape. In fact, 5HT has ambivalent effects on both reward and fear systems, inhibitory to some extent, but excitatory as well at times.

5HT is also involved in fear-induced analgesia following freezing and escape, which is of a non-opioid nature(Coimbra & Brandao, 1997).

Elevated Serotonin function during fear

As demonstrated by microdialysis studies (Bland *et al.*, 2004a), Serotonin extracellular concentration can be dramatically increased during punishment ; however, mild punishments sometimes fail to increase 5HT function.

Moreover, studies tracking c-fos expression have demonstrated that Serotonin neurons are activated in all the raphe nuclei.(Takase *et al.*, 2004)

Studies of electrophysiology have failed to show any perturbation to cell-firing that would correlated to heart-beat or any measure of stress, and have found only activation to very general arousing stimuli.

2.2.3 Effects of punishment on Dopamine function

Results from microdialysis

Results from microdialysis unanimously report an elevated extracellular Dopamine concentration in relation to punishment, with a wide number of punishment paradigms : restraint, tailpinch, tailshock, forced swimming.

This elevation is seen both in cortical areas and in the Nucleus Accumbens, with various results for the Striatum.

Absence of phasic activation of Dopamine neurons

Some results had previously reported that a fraction of Dopamine neurons presented the same kind of phasic activation, locked to aversive stimuli, as has been described in the 'reward' section of this report (Horvitz, 2000). However, a study by Ungless has shown that the firing from DA cells in midbrain areas is uniformly inhibited, and that the cells that increase firing during punishment were probably mistaken for DA cells, based on very similar firing properties, while they were not.(Ungless *et al.*, 2004).

2.3 Puzzling findings

Here, we give some of the main experimental findings about the interplay of 5HT and DA, in general or in the particular case of uncontrollable punishment, that we will try and explain in following sections.

2.3.1 Interregulation between Dopamine and Serotonin

Ambivalent regulation of Dopamine by Serotonin

Similarly as what has been seen with behavioural responses to reward, Serotonin regulates Dopamine in a number of ways that sometimes play against each other, depending on what 5HT receptor subtype is considered, and where. The picture that emerges is that 5HT2c receptors quite generally inhibit firing of DA cells at the level of cellular bodies in the VTA (De Deurwaerdere *et al.*, 2004; Di Matteo *et al.*, 2001; Millan *et al.*, 2000; Higgins & Fletcher, 2003; Dremencov *et al.*, 2005; Lucas & Spampinato, 2000), whereas 5HT2a (Di Matteo *et al.*, 2001; Millan *et al.*, 2000; Lucas & Spampinato, 2000) and 5HT3 (De Deurwaerdere *et al.*, 1998) receptors activation facilitates DA release at terminals, through a mechanism that is still unclear, but might involve a carrier-mediated release (Santiago *et al.*, 1998). Given that a significant fraction of DA neurons is usually silent (Floresco *et al.*, 2003), the activation of DA function might reflect an activation of silent neurons, so that the release of DA would be higher without increasing firing.

5HT1A,5HT1B and 5HT4 (De Deurwaerdere *et al.*, 2002) receptors have also been reported to increase DA release. Finally, infusion of 5HT into the NAc results in a dose-dependant elevation of extracellular DA (Parsons & Justice, 1993).

Does Dopamine exert a regulation on the serotonergic system ?

Some DA receptors are found on raphe nuclei(Ferre *et al.*, 1994), which contain most of the brain 5HT neurons, and seem to exert some weak regulation of 5HT ; however, the effect is quite weak and is far from approaching the scale of the regulation of DA release by 5HT in the VTA.

Studies that have assessed mutual monoamines regulation in the PFC have failed to find any significant local regulation of 5HT by DA, contrary to what is seen with norepinephrine (NE)(Millan *et al.*, 2000).

2.3.2 An interesting case : uncontrollable punishment

Uncontrollable punishment is detailed because it yields some of the most intriguing results in term of adapting coping and neuromodulatory responses.

Alteration of coping behaviour according to the controllability of the situation - Learned helplessness

With all the described paradigms, the experimenter can choose to either make the situation controllable, or uncontrollable. In a controllable situation, the animal can end the punishment with a behavioural response, eg. by pressing a lever, or running enough in a wheelturn, whereas in an uncontrollable situation, it has no control whatsoever(Cabib & Puglisi-Allegra, 1996; Bland *et al.*, 2004b).

When animals are submitted to uncontrollable aversive situations either repeatedly or for a prolonged time period, they subsequently display an impaired coping response (learned helplessness), in that the attempts to flee or fight will be less vigorous and even disappear in coping tests (Cabib & Puglisi-Allegra, 1996; Amat *et al.*, 2005). Their interest in rewards is also decreased (Cabib & Puglisi-Allegra, 1996; Di Chiara & Tanda, 1997) (anhedonia). Conversely, they present an enhanced sensitivity to morphine (Will *et al.*, 2002).

Uncontrollable stressful schedules are widely studied because learned helplessness and anhedonia are two features of depression, so that a better understanding of their mechanisms in animals might shed light on human depression.

Physiological effects of uncontrollable punishment

Uncontrollable punishment maintained over time or repeated results in increased activation of 5HT neurons as measured by *c-fos* expression(Takase *et al.*, 2004), increased release in medial PFC (mPFC)(Bland *et al.*, 2003a), NAc (Bland *et al.*, 2003b). On the contrary, DA stress-induced increases in the NAc are reduced (Cabib & Puglisi-Allegra, 1996) but maintained or even enhanced in mPFC(Di Chiara *et al.*, 1999; Bland *et al.*, 2003a). These findings can be related to the observation that DA function in mPFC seems to inhibit accumbal DA during stress (King *et al.*, 1997).

2.3.3 Clues to neural substrate of learned helplessness

A recent study has shown that whether a situation is deemed controllable or uncontrollable (with all the sequelae that have been described before), depends crucially on mPFC ; in (Amat *et al.*, 2005), they show that when mPFC is inactivated by local injections of muscimol (a GABA agonist), controllable stress elicits both the same neuromodulatory and behavioural pattern as uncontrollable stress : impaired coping and increased release of DA and 5HT. This tends to prove that the uncontrollable stress response is the default evolution of stress response.

The learned helplessness response is blocked by opioid antagonist naltrexone (Cabib & Puglisi-Allegra, 1994) and depends on opioids that activate the DRN (Will *et al.*, 2004; Bland *et al.*, 2004b) and 5HT in mPFC (Bland *et al.*, 2003b). Thus, it probably involves an activation of 5HT by opioids, with a long-term sensitization to opioids.

Chapter 3

Coping in punishment

3.1 A model of coping behaviour

3.1.1 Objectives of the model

As we have seen in previous chapter, there are accumulating data about 5HT and DA function during punishment, that are still isolated and are not integrated into a unified theory. In particular, there seem to be few attempts to link biophysical recordings and behavioural facts.

This work was undertaken in the view that an intuitively reasonable model of coping behaviour could give some meaning to those scattered findings, by viewing them as mechanisms of adaptation rather than physiological curiosities, and provide a plausible account of global response to punishment.

In order to try and bridge the gap between behavioural observations and biopysical measurements, the model comes into two parts : a behavioural account of coping response, and a mapping to neuromodulators.

The behavioural part proposes an account of fear response as a tradeoff between maximizing the chance to escape by appropriately overemphasizing the value attributed to coping responses (attempt to change the environment, while keeping the functioning of the organism unchanged), and minimizing the disruptive effect of the situation on normal functioning of the organism, by attenuating discomfort and adapting to punishment (attempt to take the environment as given, while acting on the organism itself).

The rationale for the tradeoff is that the animal has full control over its own organism, but only partial and both variable and unpredictable control over its environment ; however, considering that the punishment state is a deteriorated environment with respect to the animal's baseline, the optimal solution would be for the animal to alter the environment, so as to restore it to baseline, instead of altering its own baseline functioning to adapt to the new environment.

3.1.2 The model, principles

In this section, we shall examine how to model the two antagonist goals of the animals : reducing discomfort and getting out of the aversive context.

To introduce the hypotheses of the model, one can wonder what comes into account to efficiently retrieve a coping response : if the punishment is not unpleasant because of an over-efficient analgesia, the subjective need for coping is weak. If drive is uniformly elevated, coping response incurs the risk of being disrupted by irrelevant distracting behavioural responses, that should better be replaced by a freezing response to look for an answer, even if it is quite uncertain.

Hypotheses :

1. The objective level of punishment is fixed by the experimenter : B_0 (the higher B_0, the worse it is ; unpunished baseline :$B_0 = 1$).
2. The animal can choose both an action a and a level of analgesia $k \in [0, 1]$, so that the subjective badness of the situation is $B = k.B_0$, and the discomfort of the animal $B_0.k.S_0$, where S_0 is an index of its sensitivity to pain (we assume that the animal doesn't like pain).
   If the punishment is not unpleasant because of an over-efficient analgesia, the subjective urge to cope is weak : hence the need for a tradeoff between comfort (settling into punishment) and coping (struggling and non-acceptance of punishment).
3. An action a has an intrinsic value $v(a)$ and a fixed cost $c(a)$. Specifically, the estimated value of a coping response is taken to be proportional to the estimated probability that it will end punishment.
   The cost $c(a)$ can be thought of as an energetical or attentional cost, so that typically fighting is more costly than running away, which is more costly than freezing, which is more costly than passively waiting.
4. In the course of punishment, the rewarding value of any available coping responses is set 'unreasonably high' (relatively to the satisfaction it would provide to the animal in an unpunished context ; eg. running is expected to get a value remarkably higher than a taste for exercise would justify). However, the range of values is bounded ; that is, values cannot skyrocket. Therefore, to achieve a satisfying discrimination and ensure that usual distracting rewards (like S-R automatic associations) are not chosen, it is necessary to crush them down so as to expand the upper layers of the values range, so that only rewards above a threshold are considered ; the higher the subjective punishment, the higher the threshold : $v_{min}.B_0.k$, where v_{min} is the baseline threshold for a reward to be considered a reward. Thus, behaviour is focused on high-value actions and not distracted by small-value ones that would not help coping (like lever-pressing for food).
5. For each action a, we compute an index of preference :

$$P(a) = B_0.k.[v(a) - v_{min}.B_0.k] - c(a) - B_0.k.S_0$$

6. the animal picks the action and k that yield the highest preference ; once it is engaged in an action, if it is a coping response which turns out to be ineffective, the intrinsic value of the selected coping response shrinks. When the resulting index of preference sinks sufficiently below the next best one, allowing for some perseverative bias, the animal shifts response.

An example of preference curves is provided on fig. 3.1.

Coping strategies - the aversive context

We categorize the available behavioural responses as follows (only the F- responses are true coping responses) :

F1 - Freezing : Tonic freezing in hypervigilance, looking for clues to start another response

F2 - Fleeing : Escape response, or attempt to escape (eg. climbing the wall or swimming in forced-swim tests, running around during tailshock administration). Attempt to *change environments.*

F3 - Fighting : Trying to face punishment and to end it, that is, to *change the current environment* (eg. pressing levers, running in wheels).

I - Passive immobility : Passive waiting

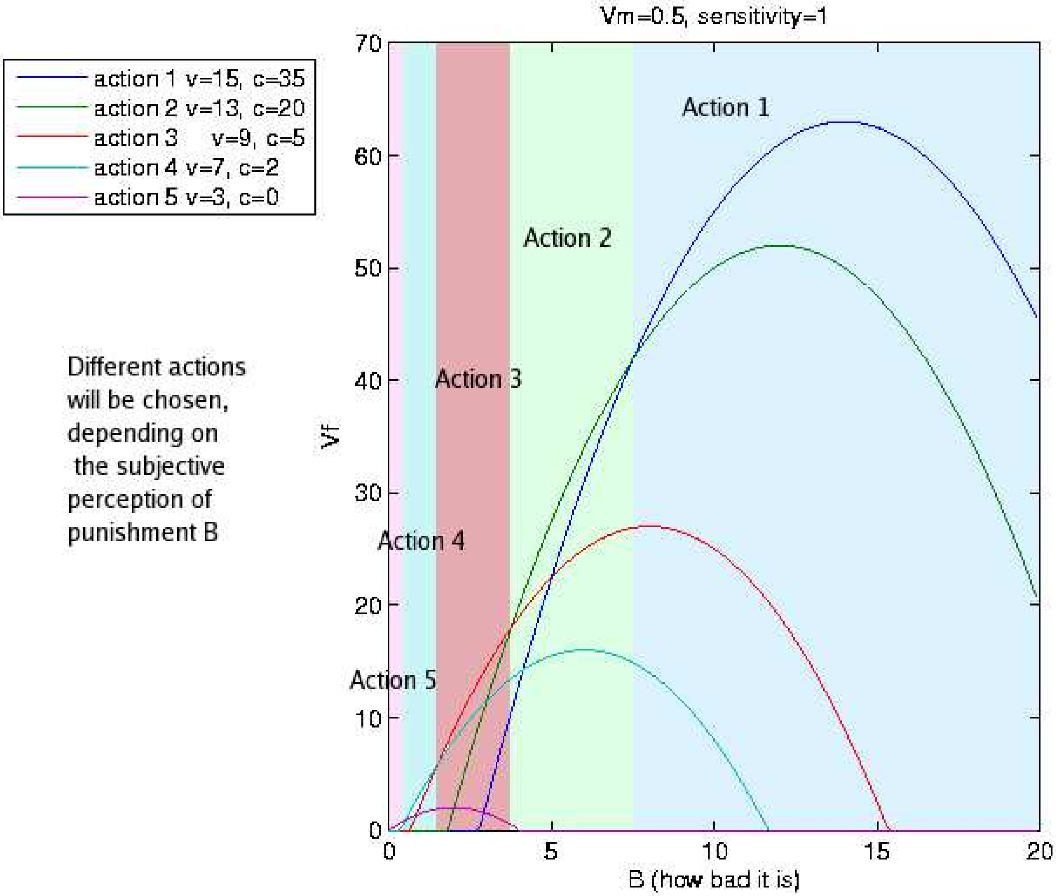
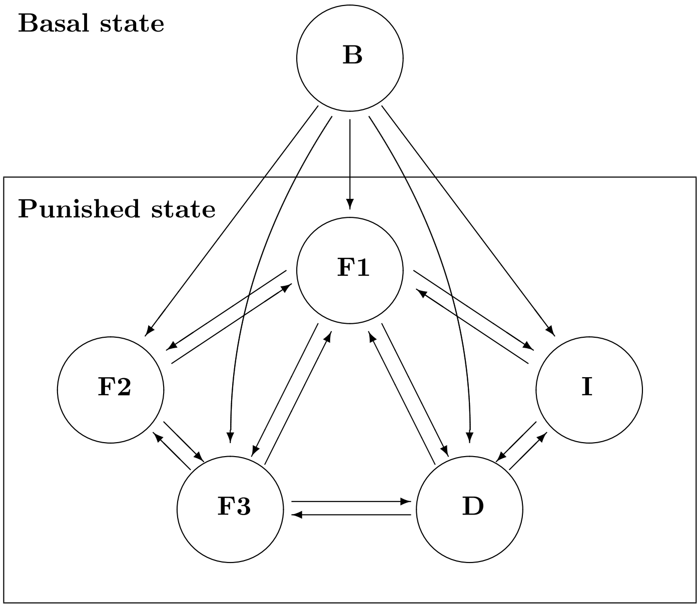


Figure 3.1: Preference curves for 5 actions ($V_m = 0.5$, $S_0 = 1$, B_0 up to 20). Vf=preference.

D - Distractor action : Any other action (chewing bars, lever-pressing, fornicating, etc.)

We model the aversive context as a 5-state world, as represented on the following diagram, where each state is labelled according to the action that is estimated to be the best at a given moment.



For clarity, some of the arrows have not been represented.

Each transition is represented by an arrow. Transitions from one state to another can be of two types :

1. **abrupt change in the environment**, that triggers either entry into the punished state from the basal state (eg. at the presentation of a CS that has been paired with footshock), or a reevaluation of preferences with a discontinuous change (eg. value of a flight is boosted if a locked door opens, allowing escape from the aversive environment).
2. **erosion of the value of a coping response**, when the chosen response turns out to be incapable of ending the punishment as was hoped when it was picked.

We will now explain how the 'trajectory' between states is built.

3.1.3 Trajectory

Transition from basal to punished state

Some stimuli in the environment are interpreted as clues that punishment is impending ; anticipated punishment (here considered part of the punished state), with elicitation of a preparatory coping response in order to try and avoid the expected punishment, lies at the core of anxiety.

This is best studied experimentally with fear conditioning ; depending on the type of punishment to be expected, a CS can elicit different behavioural responses. If the punishment is inescapable, the CS will elicit freezing (transition from B to F1). If the punishment is escapable by a flight response (eg. the CS predicts a shock that is administered only if the rat fails to run into a safe compartment in less than 5s) that has been previously learnt, the CS will trigger a transition from B to F2.

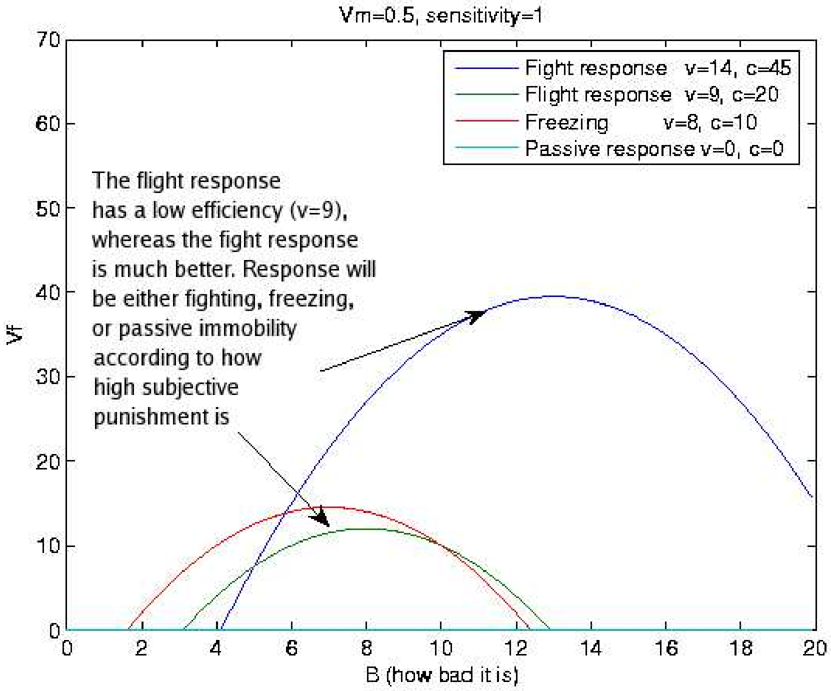


Figure 3.2: Example of higher fight response (Vf=preference)

Which state is chosen

More generally, the response that is first chosen depends on the evaluation of the situation by the animal. The animal evaluates the probability of each response to end punishment; this evaluation depends on previous punished experience.

The role of freezing at the beginning of a punishment schedule could be a rapid examination of the environment, to determine which response is better adapted.

Fig 3.2 illustrates the hypothesized preferences of a rat in an experiment where it is submitted to aversive loud noise that can be turned off by leverpressing.

Fig 3.3 illustrates a case where a rat can turn off a loud noise by giving many lever presses, or simply escape noise by hiding inside its nest.

Distracting behaviours In the model, the other possible behaviours are considered if their value is above the threshold $B.v_{min}$, to that mild punishment can result in a global enhancement of drive with all behaviours facilitated, whereas strong punishment can drastically reduce normal behaviour. Indeed, it has been reported that pinch and shock can activate feeding(Morley & Levine, 1980), aggression (Caggiula, 1972) and copulation (Caggiula, 1972; Leyton & Stewart, 1996).

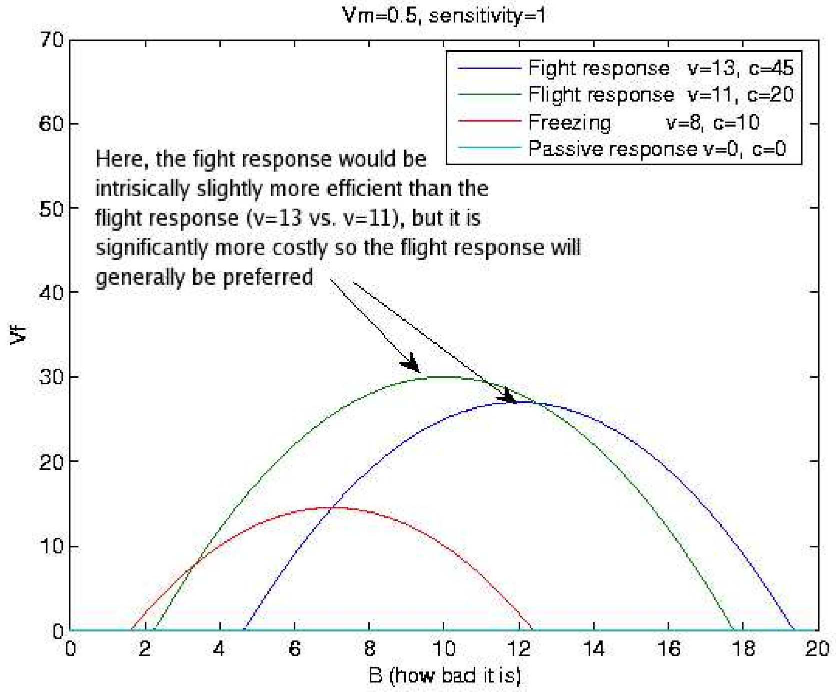


Figure 3.3: Example of higher flight response (Vf=preference)

Effect of the strength of punishment

We have hypothesized that the rat can choose a level of analgesia ; that is, the rat can choose to *diminish* subjective punishment, consequently moving leftwards on our curves. However, subjective punishment is bounded by the value of objective punishment, that is, there is no moving to the right of the value of objective punishment on the plot.

This can provide an account of the already mentioned experimental finding that aversive stimulation of the PAG elicits first freezing, then running and jumping (Vianna *et al.*, 2001), depending on how intensive the stimulation is. Thus, on Fig. 3.4, if the objective punishment is less than 3, the response will be freezing, whereas an objective punishment between 3 and 8 will elicit an escape response, and above 8, a more violent fight response.

Transitions between states in punishment

Abrupt changes Fig 3.5 provides an example of a transition triggered by an abrupt change in the environment. A rat is submitted to mild intermittent tailshocks, and has to perform a high number of lever presses to end punishment (left part) ; at some point, the door of a safe compartment unexpectedly opens, making the flight response a lot more attractive, and triggering a transition from F3 to F2 (right part).

The value of freezing *after* the onset of punishment can be understood as an active searching of clues of such abrupt changes, that is, environmental hints that a fight or flight response can unexpectedly become highly efficient.

Erosion of the value of a coping response However, if nothing new happens, the transition from a state to another will happen without the values being discontinuously updated.

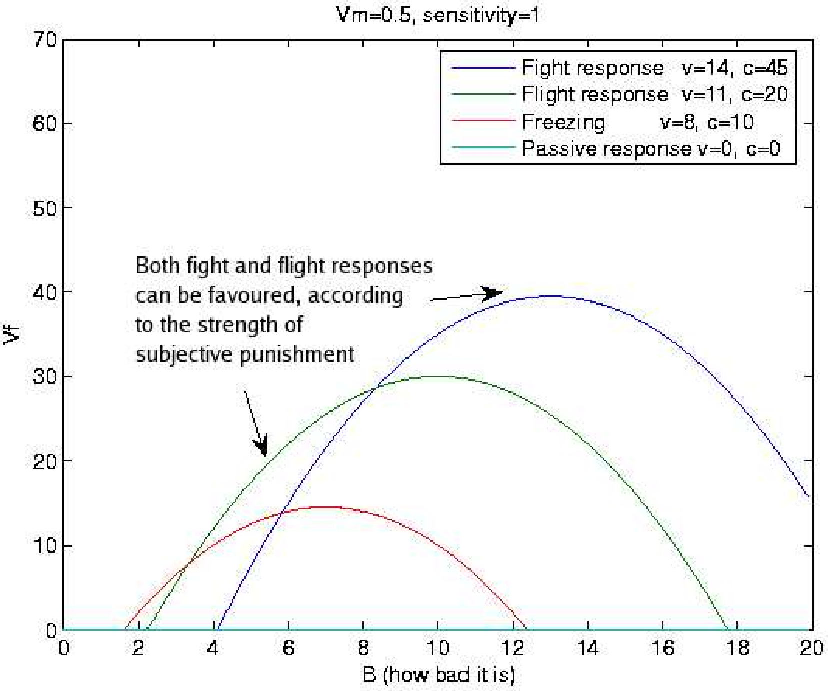


Figure 3.4: Whether an aversive stimulation elicits freezing, running away, or an attempt to fight depends on the higher bound set by objective punishment

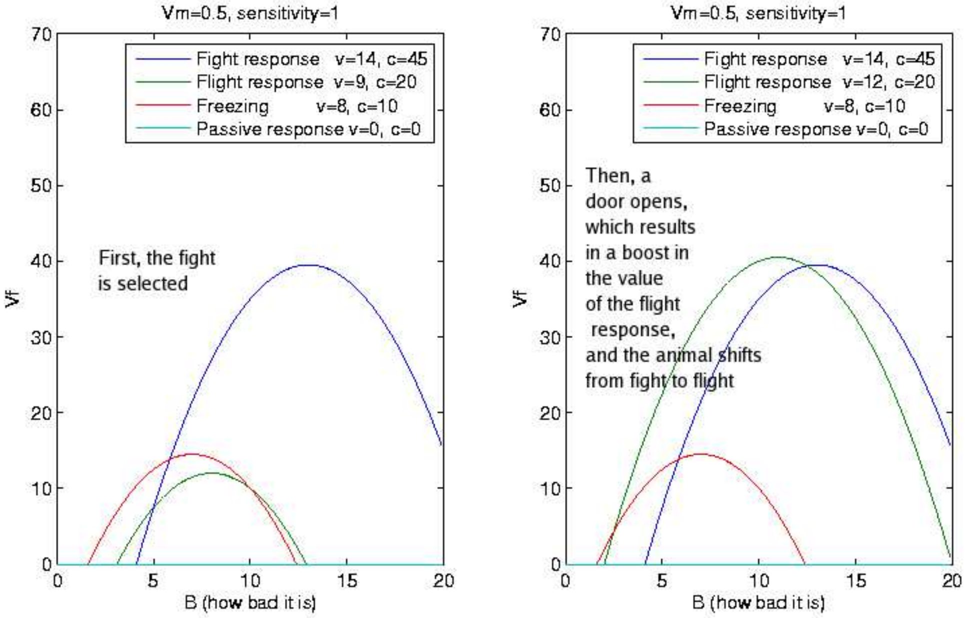
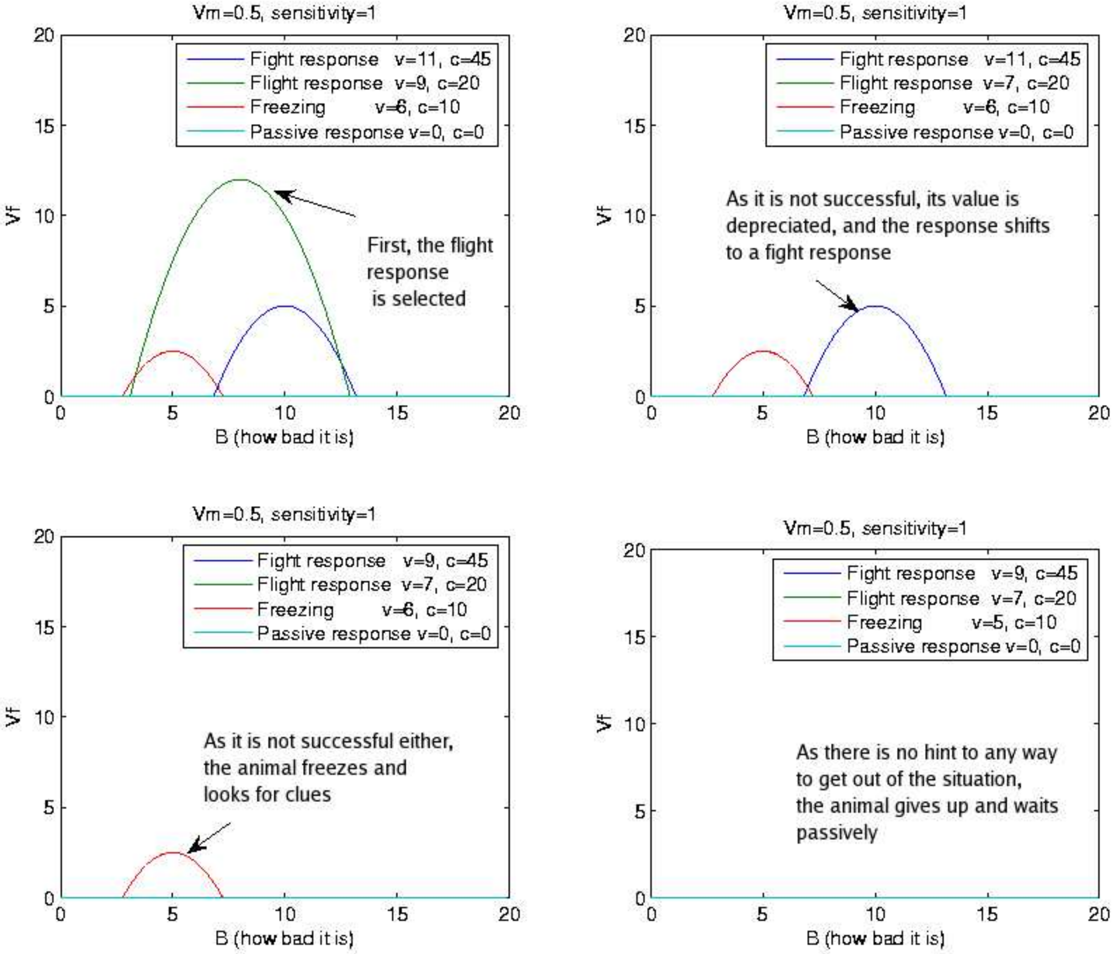


Figure 3.5: Abrupt change in the environment resulting in a boosting of the value of a flight response



First, the flight response is selected

As it is not successful, its value is depreciated, and the response shifts to a fight response

As it is not successful either, the animal freezes and looks for clues

As there is no hint to any way to get out of the situation, the animal gives up and waits passively

Figure 3.6: Erosion of value when responses fail to terminate punishment

On Fig. 3.6 is provided an illustration of how the coping response could shift from fight to flight, then freezing, then cessation of attempts at ending punishment. We hypothesized that the value of a coping response decreases with the probability that it might result in ending punishment (the animal 'looses hope' of succeeding with the selected response) ; as time passes, the likelihood that the previously second best response is actually quite better than the current one increases. When it gets higher than some threshold, a transition occurs.

A difficulty here is the meaning of *quite better*. If the animal shifts behaviour as soon as the current response is likely to be slightly worse than another, we might end up with shifting behaviour every second or half second. So we have hypothesized a *perseverative bias*, which confers an advantage to the current behaviour over its competitors, so that the difference between preference indices of current behaviour and of its nearest competitor has to be above a threshold to trigger a switch.

3.1.4 Learning from one trial to the next

So far, the effects we have described all relate to a general notion of drive. But what are the consequences of punishment on learning, ie what adaptation should occur from one trial to the next ?

Learning about coping responses

One of the most crucial determinants of behaviour in the model is the estimated probability that a response will lead to termination of punishment. The value of each response is learnt for the specific punishment applied, as each sort of punishment has a distinct winning strategy. As one type of punishment trial becomes more familiar to the animal, each value is known with better certainty, and does not need to be guessed any more. As a consequence, the transition from basal state at the onset of punishment leads directly to the learnt best response.

But the new beliefs about the value each response are not restricted to the specific punishment type, and extend to the general pattern of coping. Thus, animals repeatedly submitted to uncontrollable stress display poor coping behaviour in a subsequent forced-swim test (Cabib & Puglisi-Allegra, 1996), or shuttlebox escape (escape from electrical shock by two crossings of the experimental chamber) (Amat *et al.*, 2005).

An animal that has experienced many unsuccessful coping attempts seems to develop a belief that coping is hopeless, which translates in the model into : intrisic values of F responses are low, so low that the passive waiting with full analgesia is finally selected.

However, it should be noted that what is always described as a failure of coping can actually be the best response ; in the forced-swim test, animals usually have no control whatsoever on the termination of punishment, so that attempts at climbing the wall or swimming are useless, and the response selected by the 'helpless' animals is actually the best one. Animals submitted to uncontrollable stress paradigms perhaps learn about the fact that the experimenter always frees them from punishment at some point, so that passive waiting reliably results in relief from punishment. Experimenters say that mice demonstrate a clear awareness that experimental trials are ended by the experimenter (which is one of the downsides of artificial settings like laboratories), for instance when they start swimming towards her when she comes near the pool to pick them when the trial is over.

Attributing incentive value

It has been reported that when behaviours are activated by punishment, not only are they higher during punishment, but their increased value remains even subsequently, in the absence of punishment (Leyton & Stewart, 1996).

This learning about boosted incentive values could be needed to achieve a fast and efficient learning about the winning strategy to face an aversive situation, so that coping responses could be learnt in a limited number of trials, compared to learning under basal conditions. Thus, increased attribution of incentive value to irrelevant stimuli (with respect to coping) like a female could be a diversion of extremely favourable conditions of learning, intended to enhance coping. Consistent with that view, superior procedural learning has been found on anxious individuals under punishment (Corr *et al.*, 1997)

3.2 Proposed links with neuromodulators

3.2.1 Overview : who does what

Analgesia

Fear-induced analgesia depends on 5HT and not opioids (Coimbra & Brandao, 1997). However, given that :

- ’uncontrollable stress’ response pattern involves opioid activation (it is blocked by naltrexone)(Cabib *et al.*, 1989)
- ...and activation of 5HT cells in the DRN (Bland *et al.*, 2004b; Takase *et al.*, 2004; Will *et al.*, 2004) and of 5HT in mPFC (Bland *et al.*, 2004a)
- opioids can enhance 5HT release at the level of the DRN (Tao & Auerbach, 2002),

analgesia of uncontrollable stress might be an opioid-mediated enhancement of immediate 5HT analgesia.
Thus, we propose that :

- $1 - k$ measures opioid activation of 5HT analgesia.
- B_0 (which is the objective strength of punishment) measures ’basal’ punishment-activated 5HT function, that is, without the additional enhancement by opioids.
- resulting 5HT function is accounted for by : $B_0.(1 + (1 - k))$

  Thus, the corresponding pattern of activation of 5HT in punishment would be : a general increase, and if the punishment is uncontrollable, a further boosting. This pattern is consistent with microdialysis results (Bland *et al.*, 2004a, 2003a).

Filtering of behaviour

Given the links of DA in NAc with facilitation of behaviour, as presented in chapter 2, DA function in NAc should be linked with the global processing of intrinsic values of responses, that is, the ’shape’ of the equation that gives preference indices from intrinsic values.

In our model, the dampening of irrelevant behavioural response is represented by the threshold $B.vmin$, below which a response is not considered. Such a dampening of responses should correspond to an *inhibition* of DA ; in view of all the evidence that 5HT2c receptors constantly inhibit DA release at the level of cellular bodies (De Deurwaerdere *et al.*, 2004; Di Matteo *et al.*, 2001; Millan *et al.*, 2000; Higgins & Fletcher, 2003; Dremencov *et al.*, 2005; Lucas & Spampinato, 2000), we propose that $k.B_0.v_{min}$ reflects the activation of 5HT2c by punishment ; that is, the inhibition of DA firing (which is consistent with results of electrophysiology (Ungless *et al.*, 2004)).

As for the activation ($k.B_0$ in the model)that compensate this inhibition downstream to it, it is an indiscriminate amplification of selected responses, for which we propose that it corresponds to the activation of 5HT2a and 5HT3 receptors, as these receptor subtypes have been constantly associated with a terminal facilitation of DA release (Di Matteo *et al.*, 2001; Millan *et al.*, 2000; Lucas & Spampinato, 2000; De Deurwaerdere *et al.*, 1998). Alone, they would mediate a global activation of behaviour, as can be seen when 5HT is infused directly inside the NAc (Sasaki-Adams & Kelley, 2001) ; but upstream activation of 5HT2c receptors gate the activation to avoid an amplification of all behaviours.

But if 5HT concentration in the NAc is increased in uncontrollable stress (Bland *et al.*, 2003b) relatively to controllable stress, and we pose that DA function is increased via 5HT receptors, then we should expect an increased release of DA in NAc in uncontrollable stress as well, and this is not observed(Bland *et al.*, 2003b) ; release is increased only in PFC.

The explanation we propose is that PFC inhibits DA release in NAc during uncontrollable stress. It has been observed that DA in PFC inhibits DA release in NAc during stress (King *et al.*, 1997) ; PFC excitatory connexions are also inhibiting NAc. So PFC is capable of inhibiting DA in NAc.
It can, but must it do so ?
In terms of adaptive behaviour, when a situation is deemed uncontrollable, there is no reason to boost values any more, both in term of facilitation and incentive value learning (see next paragraph and paragraph 3.1.4), so the only effect of maintaining a high DA release would be to increase the incentive value of irrelevant behaviours. Therefore, it makes sense to inhibit DA in NAc if the situation is uncontrollable.
In the model, the facilitatory term is $k.B_0$, and the action of opioids $(1-k)$ reduces the facilitation of DA release by 5HT.

Attribution of increased incentive value

As has been briefly mentioned in chapter 2, DA seems crucially involved in learning the rewarding value of a stimulus. According to an influential theory, DA signal reports an error in the prediction of future reward (Schultz *et al.*, 1997) ; the higher the signal, the higher the learnt rewarding value.

In the context of punishment we are examining, if a stimulus is sufficiently valuable to cross the upstream barrier, it might excite DA cells in an exceedingly high fashion, as we have seen in last paragraph, resulting in attribution of an elevated incentive value.

Thus, terminal facilitation of firing could account for excessive attribution of incentive value, as seen in (Leyton & Stewart, 1996)

State transitions

State transitions could be mediated by firing of NE in the PFC, in pretty much the same manner as has been described by Dayan and Yu (Dayan & Yu, 2005) More precisely :

- transition from basal state to punished state : the probability that a punishment is impending is computed by considering all relevant predicting cues (CSs, contextual cues, etc.). NE signal varies according to that probability ; if computed probability goes above a threshold probability, NE triggers the entry into punished state.
- abrupt transitions between states : similarly, NE signal follows the computation of the probability that one response is now better than the current one, and when needed, triggers the transition into the new state.

It should be noted that in the PFC, NE release is facilitated by 5HT2a receptors subtype (Millan *et al.*, 2000). There is also large evidence that 5HT2a activation increases impulsive and premature responding in attentional tasks (Koskinen *et al.*, 2003), and this effect is dependent on adrenergic alpha1 receptor activation(Koskinen *et al.*, 2003).
Thinking of impulsivity as an unjustified transition from a 'no-response' state to a 'response' state, 5HT2a receptors could increase impulsivity by facilitating NE-triggered state transitions.
In this view, activation of 5HT2a receptors in mPFC could result in facilitated transitioning from basal to punished responding, ie, increased anxiety.

3.2.2 Uncontrollable vs. controllable stress and learned helplessness

Learned helplessness

Learned helplessness would correspond to the selection of the passivity response in the model (response I). How in the model do we get to choose the passivity response ?
We have assumed that the passivity response has both intrinsic value and cost at 0. Therefore, the only remaining term in the computation of the preference index is the discomfort term : $P(I) = -B_0.k.S_0$. Thus, the highest preference index is obtained with maximal analgesia : $P(I) = 0$, for $k = 0$.

It means that passivity will be preferred if all other responses have preference indices *below* 0, that is, if all the values of coping responses have been sufficiently eroded by unsuccessful attempts.
However, if a non-coping response (a D response in the model) is available, it might be selected and prevent the I response. Indeed, actions like chewing a piece of wood are reported to attenuate the effects of uncontrollable stress.

If the I response is selected, $k = 1$, that is, according to our mapping, opioid function is maximal. This is consistent with the findings of a sensitization to morphine following uncontrollable stress (Will *et al.*, 1998, 2002, 2004; Bland *et al.*, 2004b).

The subsequent sensitization to morphine that has been observed is long-lasting (up to 7 days), and indicates a general sensitization to opioids ; one of its effects is that morphine-induced release of 5HT in mPFC is increased (Bland *et al.*, 2003a).
Furthermore, stress-induced analgesia is increased in uncontrollably stressed animals.
For our model, what could be expected of such an increase of analgesia, as we have seen previously, is a shift of the upper bound (B_0) towards the left of the plots. Thus, as we have seen, fight and flight responses are more likely to be overlooked, all the more as their intrinsic value is depressed by their lack of success. That could account for impaired coping behaviour in subsequent tests.

Relation with depression

Considering the roles we have assigned to neuromodulators during punishment, several features could bear some relation to depression :

- if dampening of rewards is exagerated, too few events are considered rewarding enough to elicit a response ; this would facilitate the onset of learned helplessness, and more generally lead to anhedonia. Indeed, there is evidence that hyperfunctionality of 5HT2c might account for depressive symptoms in an animal model of depression (Dremencov *et al.*, 2005).
- studies have found a 52% of 5HT2a receptors in PFC in another animal model of depression (Ossowska *et al.*, 2001). other studies have shown that the density of 5HT2 receptors is higher in the bran of sucicided people. In the perspective of the model, 5HT2a enhances NE release, that in turn triggers state transitions. Therefore, increased population of 5HT2a receptors could lead to frequent unjustified shift from basal to punished state, resulting in anxiety.

Chapter 4

Concluding remarks

4.1 Personal contribution

This work was aimed at integrating scattered and puzzling findings about 5HT and DA into a coherent theory.

4.1.1 Background from the literature

A review of literature on reward and punishment yields a contrasting picture : reward contexts seem quite straightforward, with approach and drive being unequivocally enhanced, without any harmful side effects. On the other hand, punishment contexts are more ambiguous, as coping response can involve either avoidance (flight) or approach (fight).

Alterations in monoamine functions reflect that contrast : DA response is fairly clearcut in reward context, correlating with increased drive (Roitman *et al.*, 2004), activity (Hamdani *et al.*, 2004), and reinforcement (Wise, 2004). Conversely, DA response to aversive context yields a mixed and complex picture of intertwined facilitation and inhibition (Cabib & Puglisi-Allegra, 1996).

Thus, the matched complexities of behavioural and DA response to punishment seem to arise from the necessity of optimally balancing two conflicting opponent urges : successful coping (avoidance of punishment) and gathering of available rewards.

Given the attested role of 5HT in inhibiting inborn responses, allowing reversal learning and withdrawal of acquired response, 5HT is likely to have a quite important role in that tuning. Thus, Deakin and Graeff (Graeff *et al.*, 1997) have postulated a role of 5HT in defense and anxiety to alter innate fear response in an adaptive way by inhibiting unnecessary innate flight and eliciting fight response to learnt predictors of punishment. 5HT is also a good candidate for the regulatory role assigned by Gray to the BIS (behavioural inhibitory system) over the BAS (behavioural approach system) and the FFFS (flight-fight-freezing system) (Gray & McNaughton, 2000).

Therefore, it is not surprising that 5HT should display an ambiguous influence over reward as well as fear processing, sometimes inhibitory, sometimes excitatory.

Given the high implication of DA in reward processing, this influence should be present at the level of regulation exerted by 5HT over DA. In fact, seemingly confusing results in the literature have extensively demonstrated that this is the case, with some of the numerous 5HT receptors subtypes enhancing DA function (especially 5HT2a receptors and 5HT3), and other inhibiting it (notably 5HT2c).

However, there have been no definite attempts to link those behavioural and physiological findings.

This has been the motivation of present work.

4.1.2 A model of behavioural regulation in a context of punishment

This work consists in a two-fold model that proposes a parallel account of coping behaviour and neuromodulatory interactions.
The proposed behavioural analysis of coping behaviour during punishment is centered on the necessity of a tradeoff between general facilitation of behaviour to enhance efficacy of fight or flight responses, and inhibition of behaviour irrelevant to coping.
The subsequent neuromodulatory mapping of that behavioural model allows an interpretation of the puzzling data about actions exerted by 5HT on DA in precisely this way, ie the necessity to achieve differential facilitation and inhibition of behaviour for different behavioural responses.
Such a fine tuning is bound to be fragile, and could fail in a number of ways, thereby explaining how mood disorders like depression might set in as a dysfunction of that regulatory system.

4.2 Future directions

4.2.1 Hypothesis testing about neural substrates

This work has proposed a number of hypotheses about neural substrates of coping behaviour, which results in testable predictions. For example, punishment with 5HT2c receptors inactivated by a specific antagonist should yield a generalized indiscriminate activation of behaviour, and possibly impaired coping response learning in the presence of many motivational distractors. On the contrary, inactivation of 5HT2a and 5HT3 by specific antagonists should lead to enhanced passivity.

The role of opioids in triggering the 'uncontrollable-stress response' has already been established by previous studies.

As for the proposed role of NE, the effect of injections of specific agonists or antagonists of adrenergic receptors in responding to fear conditioned CS should be examined ; the prediction here is that facilitating alpha1 /inhibiting alpha2 function would result in increased anxiety and freezing to the CS.

Finally, a role for DA D3 receptor should be examined. D3 receptors function is still poorly understood; however, remission from depression correlates with higher D3 density, be it through pharmacological treatment or ECT. Moreover, agonism at D3 receptors is associated with enhanced morphine-conditioned place preference (mCPP) acquisition (Frances *et al.*, 2004b), but inhibited mCPP expression(Frances *et al.*, 2004a). Given that enhanced mCPP acquisition is a hallmark of the uncontrollable-stress response, D3 receptors might have some part here.

4.2.2 Improving the model

The model in its current form could be adequately completed by a somewhat more 'economics-inspired' account of the processing of rewards during punishment, with suitable treatment of expected sums of future rewards associated with each behavioural response.

As we have mentioned before, DA cells can fire in two different modes (tonic and phasic). We have remained quite vague about the type of firing that should be involved in the response to punishment. Given the correlations between drive and tonic DA on the one hand, reward prediction error on the other hand, it could be expected that elevation of tonic DA as captured by microdialysis corresponds to an overall energization of behaviour, to help performing the selected responses more efficiently, while an elevation of phasic DA would allow for efficient and fast learning of the winning coping strategy once it has led to release from punishment. But these predictions obviously need validation.

It would also be important to determine how NE and 5HT fire during punishment, especially given that NE firing plays a particularly critical role in triggering response and regulating the level of focus of attention (Clayton *et al.*, 2004).

Finally, it has been proposed that in punishment, DA in mPFC inhibits DA in NAc, but the functions of DA and 5HT in mPFC remain unclear. More generally speaking, the different roles of mPFC, NAc, amygdala, striatum in behavioural facilitation and inhibition need considerable clarification.

4.2.3 Morphine and serotonin analgesiae

Fear-induced analgesia seems to be 5HT-, and not opioid-mediated in early punishment(Coimbra & Brandao, 1997). Given that the 'uncontrollable-stress response' pattern depends so crucially on opioids, it would be interesting to assess the respective role of opioids and 5HT in late uncontrollable stress analgesia, especially in the view that spinal cord opioid analgesia is inhibited if 5HT function is blocked.

4.2.4 Alternating punishment and reward

Due to the particularities of interaction of reward and punishment on Dopamine, several interesting effects are to be expected on the conditioning of cues alternatively in an appetitive and aversive contexts, eg. for a sufficiently high punsihment :

- for a S-R link that has been only weakly reinforced prior to punishment, presentation of the stimulus during punishment fails to elicit the response ; after punishment release, S-R link is unaltered
- if the S-R link has been strongly reinforced before punishment onset, presentation of S results in R, and due to increased DA function while the threshold has been crossed, the S-R link gets considerably strengthened and remains so after release from punishment

4.2.5 Parallels with conflicting choices in punishment-free states

The importance of dampening coping-irrelevant behaviours during punishment to favour coping responses even if their value is virtual (that is, the reward constituted by punishment offset is temporally distant) amounts to putting the emphasis on absolute expected value rather than immediate availability (eg., when having to choose between one lever that delivers one food pellet per press and another that ends punishment after 40 lever presses, choosing the punishment-ending lever).

Therefore, coping behaviour and preference for higher delayed rewards over smaller immediately available ones, are likely to share a common neural mechanisms. Better understanding of coping behaviour under punishment should then give insight into punishment-free reward choices.

In line with this idea, prefrontal 5HT has been involved in inhibiting impulsive responding (that is, preference for the small immediate reward), whereas 5HT2a receptor activation enhances impulsive responding. Furthermore, depression is associated with decreased anterior cingulate function (Anand *et al.*, 2005), and this area is critically involved with conflict resolution.

Another related suggestion is that uncertainty of a reward could have the same effect as temporal distance, in that an immediately available reward is more certain than a distant one. Thus, when exposed to punishment for the first time and having to choose between testing behavioural responses with unknown escape-value, and responding on a familiar food-delivering lever, a rat should choose to explore behavioural responses to get the highest coping chance.
A shift of emphasis from certainty to expected value in punishment with respect to pure reward schedules could account for the well-documented risk-seeking in punishment, and risk-aversion in reward.

To study the validity of this idea, it would be interesting to design experiments where the rat has to choose between temporally distant (and possibly uncertain) escape and immediate reward.

Another potentially interesting experiment would be to test the alteration of temporal discounting in subjects that are submitted to punishment, eg. through intertwined tasks, with one assessing temporal discounting, and the other involving a competition between reward and punishment (eg. one quickly learnt rewarded response vs. hard to guess responses that end punishment). If promoting coping behaviour means the same as choosing rather delayed higher rewards than small immediate ones, a decreased temporal discounting might be observed.

Chapter 5

Appendix - Methods for evaluating neuromodulatory function

5.1 Electrophysiology

Principle : electrodes measure extracellularly or intracellularly the current in neurons.

Temporal resolution : very high. Order of a ms.

Restrictions :
- intracellular recordings cannot be used with freely moving animals.
- limited information on population activity
- gives no information on the quantity of neurotransmitter actually released for one spike, and this can vary widely

Advantages :
- high temporal resolution.
- only method that differentiates bursts from single spikes

5.2 Microdialysis

Principle : a microdialysis probe is implanted in the area of interest in the brain ; analysis of the collected dialysate informs on the concentration of different chemicals extracellularly. Also allows local infusion of a substance.

Temporal resolution : quite poor. Order of a minute.

Restrictions : poor temporal resolution.

Advantages :
- gives an idea of global population activity
- recordings can be maintained over an extended time

5.3 Fast-cyclic voltammetry

Principle : electrochemical measurements that can track the signal of monoamines.

Temporal resolution : high. Order of 100 ms.

Restrictions : limited selectivity : signal of DA and NE are undistinguishable.

Advantages :
- gives an idea of global population activity
- high temporal resolution

Bibliography

Amat, J., Baratta, M.V., Paul, E., Bland, S.T., Watkins, L.R. & Maier, S.F. (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat Neurosci*, **8**, 365–71.

Anand, A., Li, Y., Wang, Y., Wu, J., Gao, S., Bukhari, L., Mathews, V.P., Kalnin, A. & Lowe, M.J. (2005). Antidepressant effect on connectivity of the mood-regulating circuit: an FMRI study. *Neuropsychopharmacology*, **30**, 1334–44.

Barnes, N.M. & Sharp, T. (1999). A review of central 5-ht receptors and their functions. *Neuropharmacology*, **38**, 1083–1152.

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

Bland, S.T., Hargrave, D., Pepin, J.L., Amat, J., Watkins, L.R. & Maier, S.F. (2003a). Stressor controllability modulates stress-induced dopamine and serotonin efflux and morphine-induced serotonin efflux in the medial prefrontal cortex. *Neuropsychopharmacology*, **28**, 1589–96.

Bland, S.T., Twining, C., Watkins, L.R. & Maier, S.F. (2003b). Stressor controllability modulates stress-induced serotonin but not dopamine efflux in the nucleus accumbens shell. *Synapse*, **49**, 206–8.

Bland, S.T., Schmid, M.J., Watkins, L.R. & Maier, S.F. (2004a). Prefrontal cortex serotonin, stress, and morphine-induced nucleus accumbens dopamine. *Neuroreport*, **15**, 2637–41.

Bland, S.T., Twining, C., Schmid, M.J., Der-Avakian, A., Watkins, L.R. & Maier, S.F. (2004b). Stress potentiation of morphine-induced dopamine efflux in the nucleus accumbens shell is dependent upon stressor uncontrollability and is mediated by the dorsal raphe nucleus. *Neuroscience*, **126**, 705–15.

Cabib, S. & Puglisi-Allegra, S. (1994). Opposite responses of mesolimbic dopamine system to controllable and uncontrollable aversive experiences. *J Neurosci*, **14**, 3333–40.

Cabib, S. & Puglisi-Allegra, S. (1996). Stress, depression and the mesolimbic dopamine system. *Psychopharmacology (Berl)*, **128**, 331–42.

Cabib, S., Oliverio, A. & Puglisi-Allegra, S. (1989). Stress-induced decrease of 3-methoxytyramine in the nucleus accumbens of the mouse is prevented by naltrexone pretreatment. *Life Sci*, **45**, 1031–7.

Caggiula, A.R. (1972). Shock-elicited copulation and aggression in male rats. *J Comp Physiol Psychol*, **80**, 393–7.

Clayton, E.C., Rajkowski, J., Cohen, J.D. & Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. *J Neurosci*, **24**, 9914–20.

Coimbra, N.C. & Brandao, M.L. (1997). Effects of 5-HT2 receptors blockade on fear-induced analgesia elicited by electrical stimulation of the deep layers of the superior colliculus and dorsal periaqueductal gray. *Behav Brain Res*, **87**, 97–103.

Corr, P.J., Pickering, A.D. & Gray, J.A. (1997). Personality, punishment, and procedural learning: a test of J.A. Gray's anxiety theory. *J Pers Soc Psychol*, **73**, 337–44.

Daw, N., Kakade, S. & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Netw*, **15**, 603–616.

Dayan, P. & Yu, A.J. (2005). Norepinephrin and Neural Interrupts, unpublished.

De Deurwaerdere, P., Stinus, L. & Spampinato, U. (1998). Opposite change of in vivo dopamine release in the rat nucleus accumbens and striatum that follows electrical stimulation of dorsal raphe nucleus: role of 5-HT3 receptors. *J Neurosci*, **18**, 6528–38.

De Deurwaerdere, P., Cervo, L., Stinus, L. & Spampinato, U. (2002). Central 5-HT(4) receptors and dopamine-dependent motor behaviors: searching for a functional role. *Pharmacol Biochem Behav*, **71**, 627–33.

De Deurwaerdere, P., Navailles, S., Berg, K.A., Clarke, W.P. & Spampinato, U. (2004). Constitutive activity of the serotonin2C receptor inhibits in vivo dopamine release in the rat striatum and nucleus accumbens. *J Neurosci*, **24**, 3235–41.

Deakin, J. & Graeff, F. (1991). 5-HT and mechanisms of defence. *J Psychopharmacology*, **5**, 305–316.

Di Chiara, G. & Tanda, G. (1997). Blunting of reactivity of dopamine transmission to palatable food: a biochemical marker of anhedonia in the CMS model? *Psychopharmacology (Berl)*, **134**, 351–3; discussion 371–7.

Di Chiara, G., Loddo, P. & Tanda, G. (1999). Reciprocal changes in prefrontal and limbic dopamine responsiveness to aversive and rewarding stimuli after chronic mild stress: implications for the psychobiology of depression. *Biol Psychiatry*, **46**, 1624–33.

Di Matteo, V., De Blasi, A., Di Giulio, C. & Esposito, E. (2001). Role of 5-HT(2C) receptors in the control of central dopamine function. *Trends Pharmacol Sci*, **22**, 229–32.

Dickinson, A. & Balleine, B. (2002). The role of learning in motivation. In C.R. Gallistel, ed., *Stevens' handbook of experimental psychology*, vol. 3, New York:Wiley.

Dremencov, E., Newman, M.E., Kinor, N., Blatman-Jan, G., Schindler, C.J., Overstreet, D.H. & Yadid, G. (2005). Hyperfunctionality of serotonin-2C receptor-mediated inhibition of accumbal dopamine release in an animal model of depression is reversed by antidepressant treatment. *Neuropharmacology*, **48**, 34–42.

Ferre, S., Cortes, R. & Artigas, F. (1994). Dopaminergic regulation of the serotonergic raphe-striatal pathway: microdialysis studies in freely moving rats. *J Neurosci*, **14**, 4839–46.

Fletcher, P.J., Korth, K.M., Robinson, S.R. & Baker, G.B. (2002). Multiple 5-HT receptors are involved in the effects of acute MDMA treatment: studies on locomotor activity and responding for conditioned reinforcement. *Psychopharmacology (Berl)*, **162**, 282–91.

Floresco, S.B., West, A.R., Ash, B., Moore, H. & Grace, A.A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci*, **6**, 968–73.

Frances, H., Foll, B.L., Diaz, J., Smirnova, M. & Sokoloff, P. (2004a). Role of DRD3 in morphine-induced conditioned place preference using drd3-knockout mice. *Neuroreport*, **15**, 2245–9.

Frances, H., Smirnova, M., Leriche, L. & Sokoloff, P. (2004b). Dopamine D3 receptor ligands modulate the acquisition of morphine-conditioned place preference. *Psychopharmacology (Berl)*, **175**, 127–33.

Grace, A.A. & Bunney, B.S. (1984a). The control of firing pattern in nigral dopamine neurons: burst firing. *J Neurosci*, **4**, 2877–90.

Grace, A.A. & Bunney, B.S. (1984b). The control of firing pattern in nigral dopamine neurons: single spike firing. *J Neurosci*, **4**, 2866–76.

Graeff, F.G., Viana, M.B. & Mora, P.O. (1997). Dual role of 5-HT in defense and anxiety. *Neurosci Biobehav Rev*, **21**, 791–9.

Gray, J.A. & McNaughton, N. (2000). *The neuropsychology of anxiety : an enquiry into the functions of the septo-hippocampal system (second edition)*. Oxford University Press.

Hamdani, S., Rajabi, H., Stewart, J., Arvanitogiannis, A. & Shizgal, P. (2004). Sustained dopamine efflux in the rat nucleus accumbens during prolonged electrical stimulation of the medial forebrain bundle. Program No. 438.1. 2004 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience.

Higgins, G.A. & Fletcher, P.J. (2003). Serotonin and drug reward: focus on 5-HT2C receptors. *Eur J Pharmacol*, **480**, 151–62.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, **96**, 651–6.

King, D., Zigmond, M.J. & Finlay, J.M. (1997). Effects of dopamine depletion in the medial prefrontal cortex on the stress-induced increase in extracellular dopamine in the nucleus accumbens core and shell. *Neuroscience*, **77**, 141–53.

Koskinen, T., Haapalinna, A. & Sirvio, J. (2003). Alpha-adrenoceptor-mediated modulation of 5-HT2 receptor agonist induced impulsive responding in a 5-choice serial reaction time task. *Pharmacol Toxicol*, **92**, 214–25.

Leyton, M. & Stewart, J. (1996). Acute and repeated activation of male sexual behavior by tail pinch: opioid and dopaminergic mechanisms. *Physiol Behav*, **60**, 77–85.

Lucas, G. & Spampinato, U. (2000). Role of striatal serotonin2A and serotonin2C receptor subtypes in the control of in vivo dopamine outflow in the rat striatum. *J Neurochem*, **74**, 693–701.

Lucki, I. (1998). The spectrum of behaviors influenced by serotonin. *Biol Psychiatry*, **44**, 151–62.

Millan, M.J., Lejeune, F. & Gobert, A. (2000). Reciprocal autoreceptor and heteroreceptor control of serotonergic, dopaminergic and noradrenergic transmission in the frontal cortex: relevance to the actions of antidepressant agents. *J Psychopharmacol*, **14**, 114–38.

Missale, C., Nash, S.R., Robinson, S.W., Jaber, M. & Caron, M.G. (1998). Dopamine receptors: from structure to function. *Physiol Rev*, **78**, 189–225.

Morley, J.E. & Levine, A.S. (1980). Stress-induced eating is mediated through endogenous opiates. *Science*, **209**, 1259–61.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. & Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**, 452–4.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H. & Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, **38**, 329–37.

Ossowska, G., Nowa, G., Kata, R., Klenk-Majewska, B., Danilczuk, Z. & Zebrowska-Lupina, I. (2001). Brain monoamine receptors in a chronic unpredictable stress model in rats. *J Neural Transm*, **108**, 311–9.

Parsons, L.H. & Justice, J.B.J. (1993). Perfusate serotonin increases extracellular dopamine in the nucleus accumbens as measured by in vivo microdialysis. *Brain Res*, **606**, 195–9.

Pavlov, I. (1927). *Conditioned reflexes*. Oxford University Press.

Phillips, P.E.M., Stuber, G.D., Heien, M.L.A.V., Wightman, R.M. & Carelli, R.M. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature*, **422**, 614–8.

Roitman, M., Stuber, G., Phillips, P., Wightman, R. & Carelli, R. (2004). Dopamine operates as a subsecond modulator of food seeking. *J Neurosci*, **24**, 1265–1271.

Salamone, J.D., Correa, M., Mingote, S.M. & Weber, S.M. (2005). Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Curr Opin Pharmacol*, **5**, 34–41.

Santiago, M., Matarredona, E.R., Machado, A. & Cano, J. (1998). Influence of serotoninergic drugs on in vivo dopamine extracellular output in rat striatum. *J Neurosci Res*, **52**, 591–8.

Sasaki-Adams, D.M. & Kelley, A.E. (2001). Serotonin-dopamine interactions in the control of conditioned reinforcement and motor behavior. *Neuropsychopharmacology*, **25**, 440–52.

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

Solomon, R. & Corbit, J. (1974). An opponent-process theory of motivation i.temporal dynamics of affect. *Psychol Rev*, **81**, 119–145.

Stuber, G.D., Wightman, R.M. & Carelli, R.M. (2005). Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. *Neuron*, **46**, 661–9.

Takase, L.F., Nogueira, M.I., Baratta, M., Bland, S.T., Watkins, L.R., Maier, S.F., Fornal, C.A. & Jacobs, B.L. (2004). Inescapable shock activates serotonergic neurons in all raphe nuclei of rat. *Behav Brain Res*, **153**, 233–9.

Tao, R. & Auerbach, S.B. (2002). Opioid receptor subtypes differentially modulate serotonin efflux in the rat central nervous system. *J Pharmacol Exp Ther*, **303**, 549–56.

Ungless, M.A., Magill, P.J. & Bolam, J.P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science*, **303**, 2040–2.

Vianna, D.M., Landeira-Fernandez, J. & Brandao, M.L. (2001). Dorsolateral and ventral regions of the periaqueductal gray matter are involved in distinct types of fear. *Neurosci Biobehav Rev*, **25**, 711–9.

Will, M.J., Watkins, L.R. & Maier, S.F. (1998). Uncontrollable stress potentiates morphine's rewarding properties. *Pharmacol Biochem Behav*, **60**, 655–64.

Will, M.J., Der-Avakian, A., Pepin, J.L., Durkan, B.T., Watkins, L.R. & Maier, S.F. (2002). Modulation of the locomotor properties of morphine and amphetamine by uncontrollable stress. *Pharmacol Biochem Behav*, **71**, 345–51.

Will, M.J., Der-Avakian, A., Bland, S.T., Grahn, R.E., Hammack, S.E., Sparks, P.D., Pepin, J.L., Watkins, L.R. & Maier, S.F. (2004). Electrolytic lesions and pharmacological inhibition of the dorsal raphe nucleus prevent stressor potentiation of morphine conditioned place preference in rats. *Psychopharmacology (Berl)*, **171**, 191–8.

Winstanley, C.A., Theobald, D.E., Dalley, J.W., Cardinal, R.N. & Robbins, T.W. (2005). Double Dissociation between Serotonergic and Dopaminergic Modulation of Medial Prefrontal and Orbitofrontal Cortex during a Test of Impulsive Choice. *Cereb Cortex*, epub ahead of print.

Wise, R. (2004). Dopamine, learning and motivation. *Nat Rev Neurosci*, **5**, 483–494.

Zangrossi, H.J., Viana, M.B., Zanoveli, J., Bueno, C., Nogueira, R.L. & Graeff, F.G. (2001). Serotonergic regulation of inhibitory avoidance and one-way escape in the rat elevated T-maze. *Neurosci Biobehav Rev*, **25**, 637–45.