

# Serotonin, Inhibition and Negative Mood

Peter Dayan<sup>1</sup>      Quentin Huys<sup>1,2</sup>

<sup>1</sup>Gatsby Computational Neuroscience Unit, UCL  
17 Queen Square, London WC1N 3AR, UK

<sup>2</sup>Center for Theoretical Neuroscience,  
Columbia University, New York, 10032, NY, USA  
dayan@gatsby.ucl.ac.uk, qhuys@cantab.net

**PLoS Computational Biology**  
**In Press**

November 14, 2007

## Contents

<b>Abstract</b>	<b>2</b>
<b>Non-technical summary</b>	<b>2</b>
<b>1 Introduction</b>	<b>2</b>
<b>2 Methods</b>	<b>3</b>
2.1 The model: trains of thought	3
2.2 Serotonergic inhibition	4
2.3 Learning	5
2.4 Manipulations after learning	5
2.4.1 Tryptophan depletion	5
2.4.2 Recall bias and reward seeking	6
<b>3 Results</b>	<b>6</b>
3.1 Behavioural inhibition	6
3.2 Tryptophan depletion	6
3.3 Recall bias	8
3.4 Reward seeking	9
<b>4 Discussion</b>	<b>9</b>
4.1 Behavioural Inhibition System	10
4.2 Tryptophan depletion	10
4.3 Dopamine and Serotonin	11
4.4 Depression and anxiety	12
4.5 Conclusions	12
Acknowledgements	13
Competing Interests	13
<b>Supplement S1</b>	<b>13</b>
<b>References</b>	<b>15</b>

## Abstract

**Background:** Pavlovian predictions of future aversive outcomes lead to behavioral inhibition, suppression and withdrawal. There is considerable evidence for the involvement of serotonin in both the learning of these predictions and the inhibitory consequences that ensue, although less for a causal relationship between the two.

**Methodology:** In the context of a highly simplified model of chains of affectively-charged thoughts, we interpret the combined effects of serotonin in terms of pruning a tree of possible decisions, *ie* eliminating those choices that have low or negative expected outcomes.

**Conclusions / Significance:** We show how a drop in behavioural inhibition, putatively resulting from an experimentally- or psychiatrically-influenced drop in serotonin, could result in unexpectedly large negative prediction errors, and a significant aversive shift in reinforcement statistics. We suggest an interpretation of this finding that helps dissolve the apparent contradiction between the fact that inhibition of serotonin reuptake is the first-line treatment of depression although serotonin itself is most strongly linked with aversive rather than appetitive outcomes and predictions.

## Non-technical summary

Serotonin is an evolutionarily ancient neuromodulator probably best known for its role in psychiatric disorders. However, that role has long appeared contradictory to its role in normal function, and indeed its various roles in normal affective behaviours have been hard to reconcile. Here we model two predominant functions of normal serotonin function in a highly simplified reinforcement learning model and show how these may explain some of its complex roles in depression and anxiety.

## 1 Introduction

Serotonin (5-hydroxytryptophan, 5-HT) is a neuromodulator that appears to play a critical role in a wealth of psychiatric conditions, including depression, anxiety, panic, and obsessive compulsions. However, despite the importance of serotonergic pharmacotherapies, notably selective reuptake inhibitors (SSRIs), the roles that serotonin plays in normal and abnormal function are still mysterious. We start from three particular findings. First, 5-HT is involved in the prediction of aversive events, possibly as a form of opponent (1; 2; 3) to dopamine (4; 5; 6; 7; 8; 9; 10; 11). Second, 5-HT is involved in behavioural inhibition (12; 13; 14), preventing or curtailing ongoing actions in the light of predictions of aversive outcomes. The third finding is the collection of psychopharmacological data implicating 5-HT in animal models of depression and anxiety (15; 16; 17), together with the fact that depleting 5-HT (by dietary depletion of its precursor, tryptophan) in human subjects who have recovered from depression, can reinstate an acute, at times fulminant, re-experience of subjective symptoms of the disease, as assessed by various rating scales (18; 19; 20; 21). Furthermore, while SSRIs are used in the treatment of depression, genetically induced, constitutive decreases in the efficiency of 5-HT reuptake are a risk factor for depression (22; 23; 24). These findings are hard to connect: the second fact seems orthogonal to the first and third, which are themselves in apparent contradiction. If 5-HT is really involved in predicting aversive outcomes, then *depleting* it should surely have positive rather than negative affective consequences.

We suggest that the missing link comes from considering the interactions between Pavlovian predictions and ongoing action selection. The interaction is seen in conditioned suppression (25), a standard workhorse test for aversive predictions. Subjects are trained to emit appetitive instrumental actions (such as pressing a lever for reward), and to associate (by classical conditioning) a light with a shock. Presentation of the light during instrumental performance *reduces* the rate at which animals emit those responses. Neither the theoretical nor the neurobiological status of this interaction is completely resolved, though there is some evidence of the involvement of 5-HT in the nucleus accumbens in its realization (26; 27; 28).

Here, we treat a subset of the inhibitory processes associated with Gray's Behavioural Inhibition System (13; 7; 29; 30) in terms of what might be called a preparatory Pavlovian response. Consummatory Pavlovian responses are (evolutionarily) pre-programmed reactions to the presence of affectively significant outcomes such as food, water, or threats. Preparatory Pavlovian responses are similarly pre-programmed responses to predictions of those outcomes.

Even though the predictions are learned; the responses are not, and may therefore be behaviorally inappropriate in certain circumstances (31; 32). For our purposes, and as long noted by Deakin & Graeff (7), the most important preparatory Pavlovian response to a prediction of a (sufficiently distant) threat (30) is inhibition, in the form of withdrawal or disengagement. This explicitly links the first two findings discussed above, as the inhibition is directly associated with aversive predictions.

To explore the consequences of reflexive, direct inhibition of action for learning in affective settings, together with the repercussions when 5-HT is compromised, we built a highly simplified model that sought to isolate these effects from more general learning effects. More specifically, we built a model of trains of thoughts. In our treatment we considered thoughts as actions which lead from one belief state to the next. Trains of thought gained value through their connections with a group of terminal states that were preassigned either positive or negative affective values. 5-HT directly inhibited chains of thought predicted to lead towards negative terminal states. Our model can be seen in terms of 5-HT's pruning of a decision tree of outcome states and choices (33; 34).

We argue that the results on Tryptophan depletion (TrD) above now emerges when considering the consequences of this reflexive behavioral inhibition on ongoing learning about the world, and on subsequent action choice and predictions. The most notable effect in the model is a critical *bias* towards optimistic valuation. That is, states and actions with potentially negative consequences are under-explored and incorrectly (over)valued because of the reflexive inhibition. When inhibition fails, though, which is the last of the three issues mentioned above, there are two adverse consequences. First, the inhibition is no longer a crutch for instrumental action choice – so subjects have to learn to avoid potentially bad situations rather than being able to rely on this reflexive mechanism. Second, due to a mismatch between policy and value function, characteristic inconsistencies between the predicted and actual values arise, with the actual values encountered being more negative than predicted, though also actually more realistic. This mismatch between policy and value function also leads to an overall reduction in rewards obtained. Boosting 5-HT in the model again restores the *status quo*. Of course, this highly simplified model can not possibly, by itself, accommodate all the diverse and confusing roles of 5-HT. Nevertheless, it replicates some prominent behavioural and pharmacological facets of depression and anxiety in humans and animal models, which we discuss in section 4.

The next section defines the model of trains of thought more formally. Section 3 considers normal, biased, learning, and the consequences of impairments to 5-HT processing. We save for section 4 a broader discussion of data and theories pertaining to 5-HT.

## 2 Methods

### 2.1 The model: trains of thought

Figure 1 illustrates our underlying model of trains of thought. It is intended to emphasize a role for 5-HT in behavioural inhibition, and is therefore couched at an abstract level. Throughout, we will equate thoughts with actions, and revisit the more general action setting later. We initially focus on the effect of one inhibitory reflexive action in the context of otherwise fixed actions (a fixed policy).

A train of thoughts starts at one of a set of internal belief states ( $\mathcal{I}_+, \mathcal{I}_-$ ), may proceed through more such states, and ends in one of set of terminal outcome states ( $\mathcal{O}_+, \mathcal{O}_-$ ). The connectivity between belief states is sparse, with states  $\mathcal{I}_+$  leading preferentially to other states in  $\mathcal{I}_+$  and outcome states  $\mathcal{O}_+$  with positive values; and states  $\mathcal{I}_-$  leading preferentially to other states in  $\mathcal{I}_-$  and outcome states  $\mathcal{O}_-$  with negative values (red arrows), though each could also lead to states of opposite 'sign' (black arrows). In addition, trains of thought can be inhibited by 5-HT (see below). In this simple model, the value of an internal state is the average value of the terminal states to which it ultimately leads.

More formally, the model is a form of Markov decision process (see (35)), with four sets of sparsely interconnected states  $\{\mathcal{I}_\pm, \mathcal{O}_\pm\}$ . Two sets,  $\mathcal{O}_+$  and  $\mathcal{O}_-$  (each with 100 elements in the simulation) are associated respectively with positive ( $r(s) \geq 0, s \in \mathcal{O}_+$ ) and negative affective values ( $r(s) \leq 0, s \in \mathcal{O}_-$ ; both drawn from suitably truncated 0-mean, unit variance, Gaussian distributions, see inset histograms in figure 1) and are terminal states. The other sets,  $\mathcal{I}_+$  and  $\mathcal{I}_-$  (each with 400 elements) contain internal states and are not associated with immediate affective values ( $r(s) = 0, \forall s \in \mathcal{I}_- \cup \mathcal{I}_+$ ).

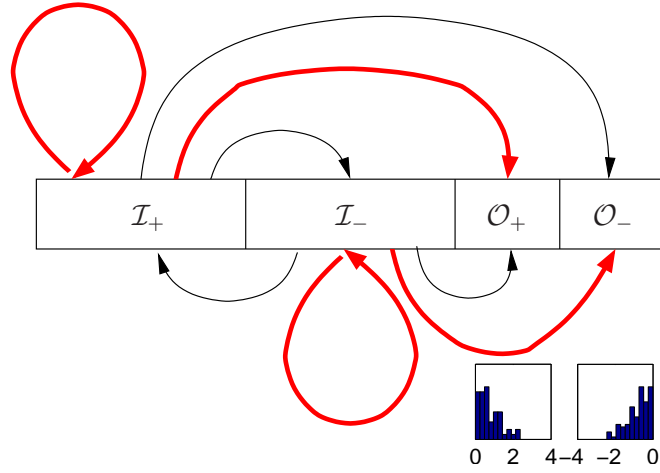


Figure 1: Markov models of thought. The abstract state space is divided into the 4 blocks shown. The right two,  $\mathcal{O}_+$  and  $\mathcal{O}_-$ , are associated with direct affective values  $r(s)$  (inset histograms); the left two,  $\mathcal{I}_-$  and  $\mathcal{I}_+$ , are internal. Transitions between (belief) states are determined by actions (thoughts). We initially focus on a fixed policy, leading to the transition between states shown in the figure: states in each internal block  $\mathcal{I}_+$  and  $\mathcal{I}_-$  preferentially connect with each other and their respective outcome states  $\mathcal{O}_+$  and  $\mathcal{O}_-$ . However, each state has links to states in the other block. The model is approximately balanced as a whole, with an equal number of positive and negative states.

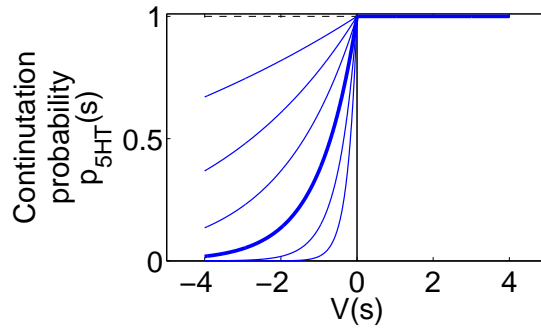


Figure 2: Probability of continuing a train of thoughts. For values  $V(s) > 0$ , thoughts are continued with probability 1. Conversely, when the state  $s$  has negative value, the probability of continuation drops off as an exponential function of the value. The rate of the exponential is set by  $\alpha_{5HT}$ .

## 2.2 Serotonergic inhibition

A policy is a (probabilistic) mapping from states to actions  $a \leftarrow \pi(s)$  and defines the transition matrix between the states in the model. For simplicity, we consider a fixed, basic, policy  $\pi^0$ . In this, each element of  $\mathcal{I}_+$  effectively has 8 outgoing connections, 3 to other (randomly chosen) elements in  $\mathcal{I}_+$ ; 3 to randomly chosen elements in  $\mathcal{O}_+$ ; and 1 each to randomly chosen elements in  $\mathcal{I}_-$  and  $\mathcal{O}_-$ . Similarly, each element of  $\mathcal{I}_-$  has 8 outgoing connections, 3 to other (randomly chosen) elements in  $\mathcal{I}_-$ ; 3 to randomly chosen elements in  $\mathcal{O}_-$ ; and 1 each to randomly chosen elements in  $\mathcal{I}_+$  and  $\mathcal{O}_+$ . Thoughts are modelled as actions  $a$  following these connections, labelled by the identities of the states to which they lead. Supplement S1 gives details of a more complex environment in which we explicitly explore effects of impulsivity.

To isolate the effect of 5-HT in inhibiting actions in aversive situations, we consider the highly simplified proposal that serotonin stochastically terminates trains of thoughts when these reach aversive states. More specifically, under serotonergic influence the transition probabilities are modified in a manner that depends on states' values. We let the probability of continuing a train of thought (of continuing along the fixed policy  $\pi^0$ ) be dependent (and inversely related to) the value  $V(s)$  of a state:

$$p_{5HT}(s) = \min(1, \exp(\alpha_{5HT} V(s))) \quad (1)$$

where  $\alpha_{5HT}$  is a multiplicative factor that scales the impact of 5-HT (see figure 2). When thoughts are not continued (inhibited), they stop and restart in a randomly chosen state  $\mathcal{I}$  (though see below for relaxations of this). The more disastrous the potential sequelæ of state  $s$ , the more negative  $V^\pi(s)$ , and so the less likely the chain was to be continued. On the other hand, even slightly positive values would essentially veto any termination. This introduces an asymmetry into the model defined by the simple base policy. Other possibilities for the information reported by 5-HT and for the dynamical interaction between 5-HT and dopamine are considered in the discussion, and the fixed base policy  $\pi^0$  is relaxed below.

## 2.3 Learning

The value of each state represents the expected reward obtainable from that state when following a particular policy. Under the fixed policy  $\pi^0$ , dynamic programming techniques (35) allow the value function  $V^\pi(s)$  over states  $s$  to be written, and solved for, concisely as:  $V^\pi(s) = r(s)$ ,  $s \in \mathcal{O}_\pm$ , and

$$V^\pi(s) = \gamma \sum_a \pi_s(a) V^\pi(a) = \gamma \sum_{s'} \pi_s(s') V^\pi(s') \quad (2)$$

where  $\gamma$  is a discount factor ( $\gamma = 0.9$  in our simulations). Dynamic programming also employs a  $Q^\pi(s, a)$  function (36) over states and thoughts defined for those actions that exist by

$$Q^\pi(s, a) = \gamma V^\pi(a) \quad (3)$$

Optimal values  $V^*(s)$  and  $Q^*(s, a)$  are those value functions which are associated with any policy  $\pi_s^*(a)$  that maximizes the long-run affective outcomes of the train.

While it is not possible to use these techniques directly to evaluate the value function under serotonergic influence (the inhibition depends on the value function itself and thus represents a nonlinear interaction), the temporal difference learning rule (TD; 35) can be used to acquire estimates  $V_{\text{est}}^{\pi(\alpha_{5HT})}(s)$  of the values  $V^{\pi(\alpha_{5HT})}(s)$  of states under serotonergically modified policies  $\pi(\alpha_{5HT})$ . TD specifies an online learning rule for which the change in the estimated value  $V_{\text{est}}^{\pi(\alpha_{5HT})}(s)$  based on taking action  $a$  at state  $s$  and therefore arriving deterministically at  $s' = a(s)$  is:

$$\Delta V_{\text{est}}^{\pi(\alpha_{5HT})}(s) = \epsilon \begin{cases} -V_{\text{est}}^{\pi(\alpha_{5HT})}(s) & \text{if the train is inhibited} \\ r(a) + \gamma V_{\text{est}}^{\pi(\alpha_{5HT})}(s') - V_{\text{est}}^{\pi(\alpha_{5HT})}(s) & \text{otherwise} \end{cases} \quad (4)$$

where the learning rate  $\epsilon = 0.05$ . A slightly simpler alternative rule suggests that learning of  $V_{\text{est}}^{\pi(\alpha_{5HT})}(s)$  is itself prevented by termination:

$$\Delta V_{\text{est}}^{\pi(\alpha_{5HT})}(s) = \epsilon \begin{cases} 0 & \text{if the train is inhibited} \\ r(a) + \gamma V_{\text{est}}^{\pi(\alpha_{5HT})}(a) - V_{\text{est}}^{\pi(\alpha_{5HT})}(s) & \text{otherwise} \end{cases} \quad (5)$$

That  $V_{\text{est}}^{\pi(\alpha_{5HT})}(s)$  does not change under this rule given termination implies that learning is only slowed for these states, rather than being biased towards 0. We generally report results from this variant.

In the sequel, we show values after substantial learning (20000 trains), plus the consequences of manipulating serotonin (by manipulating  $\alpha_{5HT}$ ) once the values are already acquired.

## 2.4 Manipulations after learning

### 2.4.1 Tryptophan depletion

Given the values  $V_{\text{est}}^{\pi(\alpha_{5HT})}(s)$  learned under a policy  $\pi(\alpha_{5HT})$  determined by  $\alpha_{5HT} = 20$ , the steady-state transitions probabilities can be calculated for any new  $\alpha_{5HT} \neq 20$  simply by working out the probability of inhibition for each state. In particular, this allows  $\alpha_{5HT}$  to be reduced to model a pharmacological or psychiatric reduction in serotonin function. To separate the effect of this reduction from that of learning, we only learn up to the reduction and then look at the behaviour after the reduction in the absence of further learning.

### 2.4.2 Recall bias and reward seeking

To account for the effect of recall biases often seen in depression, we will additionally consider the effect of biased resampling after behavioural inhibition. A simple way of achieving this in a manner that relates to the affective value of states is to let

$$p_{\text{start}}(s) \propto \exp(\beta V(s))$$

whereby values of  $\beta < 0$  will bias resampling towards states with *lower* values  $V(s)$ , i.e. states in  $\mathcal{I}_-$ .

So far, only serotonergically-determined inhibitory responses have been considered. Mirror to these are dopaminergically-controlled approach responses (32) which actually favour actions  $a$  with positive state-action values  $Q(s, a)$  (under a policy). The combined effect can be incorporated in a straightforward manner by choosing action  $a$  in state  $s$  according to a softmax

$$\pi_s(a) \propto \exp(\theta Q(s, a)), \quad (6)$$

where  $\theta$  controls the degree of influence of the  $Q$  value. Note that, in this simple model, instrumental and Pavlovian control are essentially indistinguishable.

## 3 Results

### 3.1 Behavioural inhibition

By construction, the environment in figure 1 is symmetric with respect to rewards and punishments, and so the overall statistics of the values of states are balanced about 0. Indeed, figure 3A shows that for the base policy, 20000 learning steps are ample to acquire a reasonable values  $V_{\text{est}}(s)$  for the states (the remaining discrepancies from  $V_{\text{true}}(s)$ , here defined for  $\alpha_{5\text{HT}} = 0$ ) arise from the stochasticity in the choice of action together with the fixed learning rate). Critically, there is no bias in either  $V_{\text{est}}(s)$  or  $V_{\text{true}}(s)$ .

By contrast, figure 3D shows the substantial bias in  $V_{\text{est}}^{\pi(\alpha_{5\text{HT}})}(s)$  consequent on setting a large value of  $\alpha_{5\text{HT}} = 20$ . In this case, low valued states are much less well visited and explored. The bias comes despite the use of learning rule 5, which only slows down learning for low value states rather than also distorting it. Of course, in this case, the extent of the bias depends on the initial values for the states (all of which are set to 0 in the simulation).

Figure 3E shows how frequently each of the outcome states was reached in a run (as a function of its outcome  $r(s)$ ). Since behavioural inhibition terminates trains on their way to potential disaster, aversive terminal states are sampled less (shown by the red regression line), which is consistent with the bias of the estimated value. Figures 3C;F show these effects as a function of  $\alpha_{5\text{HT}}$ . The greater the inhibition, the worse estimated are the values (C), particularly for aversive states; but the more benign is the exploration (F). Learning with greater inhibition leads to a more optimistic set of values; however, this is coupled with a more aggressive rejection of all actions even mildly associated with negative outcomes.

### 3.2 Tryptophan depletion

Reducing the value of  $\alpha_{5\text{HT}}$  after learning a value function  $V_{\text{est}}^{\pi(\alpha_{5\text{HT}})}(s)$  under its influence can be expected to have various consequences as it introduces a mismatch between policy and value function. The most obvious one is a more negative average affective outcome (the average value of trains of thought) in the model. This is because choices are less biased against actions that are predicted to have aversive consequences, and so the latter occur more frequently. A second consequence is that there will be substantial adverse surprises associated with transitions that previously were inhibited. The surprise at reaching an actual outcome can be measured using the prediction error

$$\delta = r(a) - V_{\text{est}}^{\pi(\alpha_{5\text{HT}})}(s) \quad (7)$$

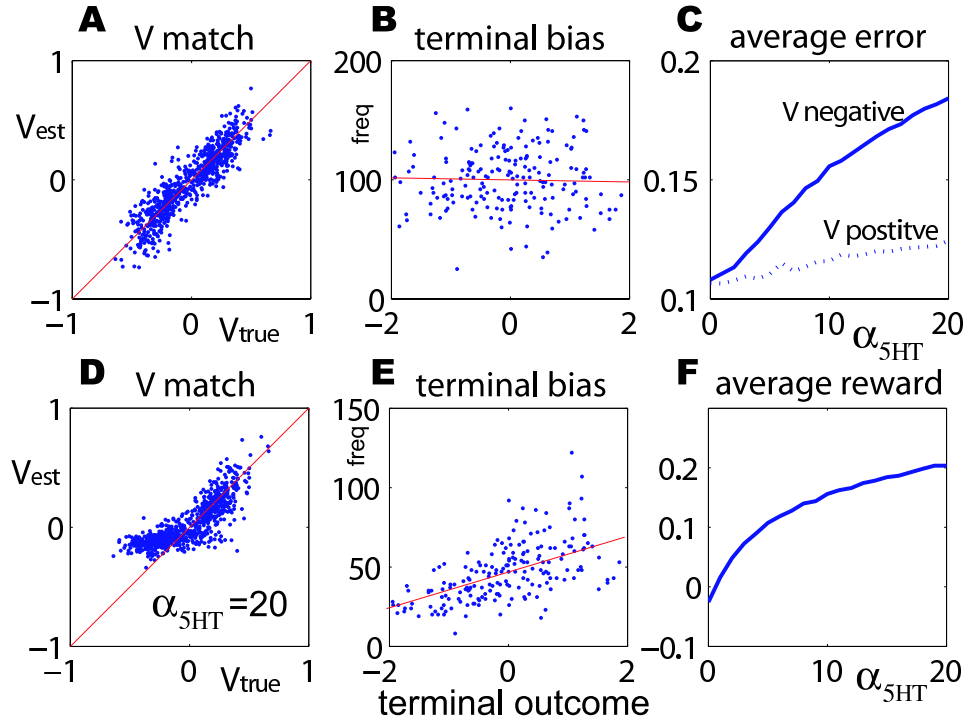


Figure 3: Learning with behavioural inhibition. (A;B) with  $\alpha_{5HT} = 0$ , for one particular learning run, the values  $V_{est}$  match their true values  $V_{true}$  (inferred through dynamic programming) under an equal-sampling exploration policy (A); and trains of thought end in terminal states  $\mathcal{O}_-, \mathcal{O}_+$  equally often as a function of their actual outcomes (B; the red line is the regression line). (D;E) with  $\alpha_{5HT} = 20$ , negative  $V$  values are poorly estimated (since exploration is progressively inhibited for larger  $\alpha_{5HT}$ ), and the more negative the value of the outcome, the less frequently that outcome gets visited over learning (E). Importantly, there is an optimistic *underestimate* of the negative value of state. (C) shows the root mean square error (averaging over 20 runs) for states with positive (dotted) and negative  $V_{true}$  values as a function of  $\alpha_{5HT}$ . The effect of the sampling bias is strikingly apparent, preventing accurate estimates mainly of the negatively valued states. (F) shows the average reward received during learning as a function of  $\alpha_{5HT}$  – the benefits of behavioural inhibition are apparent.

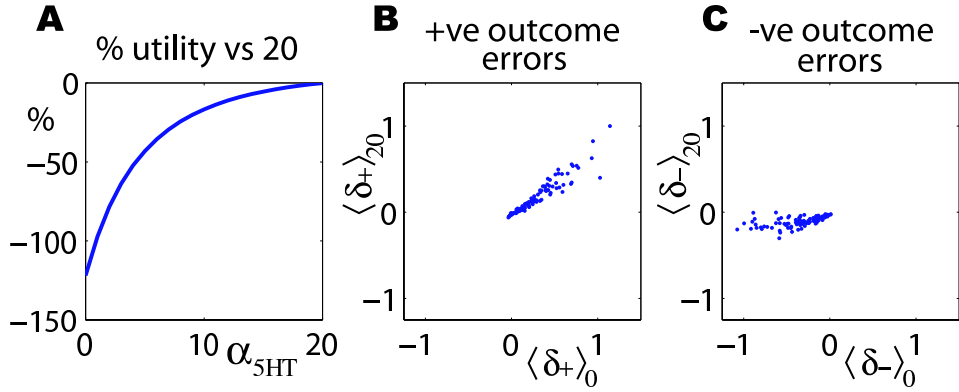


Figure 4: Reduced inhibition. These graphs show statistics of the effect of learning  $V$  values with  $\alpha_{5HT} = 20$ , and then suffering from reduced serotonin  $\alpha_{5HT} < 20$  during sampling of thoughts. For a given thought environment, these are calculated in closed form, without estimation error. A) as is also evident in figure 3F, the average affective return is greatly reduced from the value with  $\alpha_{5HT} = 20$ , in fact for the extreme value of  $\alpha_{5HT} = 0$ , it becomes slightly negative (reflecting a small sample bias in the particular collection of outcomes). B;C) normalized *outcome* prediction errors at the time of transition to  $\mathcal{O}_+$  (B) or  $\mathcal{O}_-$  (C) for  $\alpha_{5HT} = 20$  against  $\alpha_{5HT} = 0$ . These reflect the individual probability that each terminal transition goes to  $r(s)$  from  $V(s')$  for  $s \in \mathcal{O}$  and  $s' \in \mathcal{I}$ , including all the probabilistic contingencies of termination, *etc.* They are normalized for the two values of  $\alpha_{5HT}$ . Terminations in  $\mathcal{O}_+$  are largely unaffected by the change in inhibition; terminations in  $\mathcal{O}_-$  with negative consequences, have greatly increased negative prediction error.

for the last transition of a chain from state  $s \in \mathcal{I}_\pm$  to a state  $a \in \mathcal{O}_\pm$ . We may expect negative prediction errors

$$\delta_- = \begin{cases} \delta & \text{if } \delta < 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

to be of special importance, because of substantial evidence that aversive outcomes whose magnitudes and timing are expected so they can be prepared for, have substantially less disutility than outcomes that are more aversive than expected (at least for physiological pains, see 37).

Figure 4 shows the consequences of learning under full inhibition and then wandering through state space with reduced inhibition. The change in the average terminal affective value as a percentage of the case during learning that  $\alpha_{5HT} = 20$  is shown in figure 4A. As was already apparent in figure 3F (which averages over the whole course of learning), large costs are incurred for large reductions in inhibition. For  $\alpha_{5HT} = 0$ , the average reward is actually negative, which is why the curve dips below  $-100\%$ . This value is relevant, since the internal environment is approximately symmetric in terms of the appetitive and aversive outcomes it affords. Subjects normally experience an *optimistic* or rosy view of it, by terminating any unfortunate trains of thought (indeed 55% of their state occupancy is in  $\mathcal{I}_+$  compared with  $\mathcal{I}_-$ ). Under reduced 5-HT, subjects see it more the way it really is (the ratio becomes 50%).

Figure 4B;C show comparative scatter plots of the terminal prediction errors. Here, we consider just the last transition from an internal state to an outcome state. Prediction errors here that are large and negative, with substantially more aversive outcomes than expected may be particularly damaging. Figure 4C compares the average terminal prediction errors for all transitions into states in  $\mathcal{O}_-$  with no serotonergic inhibition  $\alpha_{5HT} = 0$ , to those for the value  $\alpha_{5HT} = 20$  that was used during learning. For the case that  $\alpha_{5HT} = 20$ , the negative prediction errors are on average very small (partly since the probability of receiving one is very low). With reduced inhibition, the errors become dramatically larger, potentially leading to enhanced global aversion. By comparison, as one might expect, the positive prediction errors resulting from transitions into  $\mathcal{O}_+$  are not greatly affected by the inhibition (figure 4B).

### 3.3 Recall bias

Two additional effects enrich this partial picture. One, which plays a particularly important role in the cognitive behavioural therapy literature, is that depressed patients have a tendency to prefer to *recall* aversive states or memories (38; 39). Figure 5A shows the consequence of doing this according to a simple softmax (see methods). These curves,

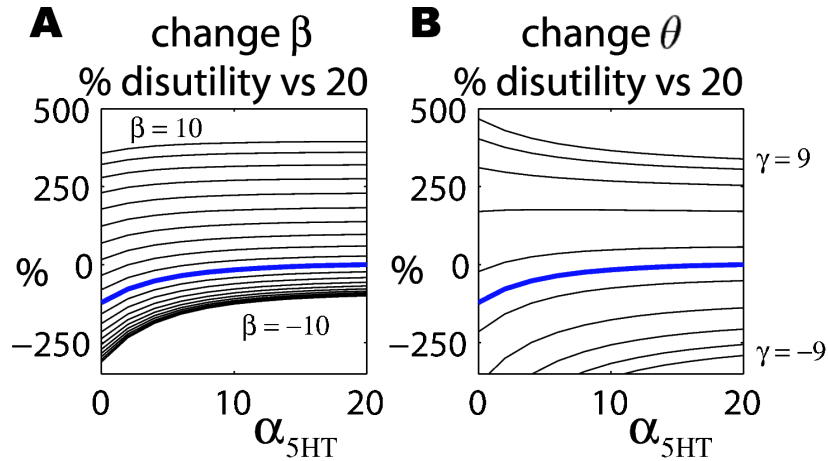


Figure 5: Both plots are in the same form as figure 4A, showing the percentage utilities compared with the standard learning case  $\alpha_{5HT} = 20$ , as a function of  $\alpha_{5HT}$  (the emboldened blue curve is exactly that in figure 4A). A) Given a mood-dependent bias on the starting state, with  $p_{\text{start}}(s) \propto \exp(\beta V(s))$ , the plots show the consequences of various values of  $\beta$ . Negative  $\beta$ , favoring low value states, leads to substantially negative average outcomes. B) Instrumental control of action choice, a putative model of dopaminergic effects, can also either exacerbate or improve the outcomes, depending on the value of the parameter  $\theta$  governing a softmax choice of actions.

as in figure 4A show the percentage average utility compared with  $\alpha_{5HT} = 20, \beta = 0$  across values of  $\alpha_{5HT}$ , and for  $\beta = -10, -9, \dots, 10$ . As might be expected, biasing the starting point to  $\mathcal{I}_-$ , and, even worse, to those particular states in  $\mathcal{I}_-$  that are most deleterious, has a big negative impact on average utility. For  $\alpha_{5HT} = 0; \beta = -10$ , occupancy of  $\mathcal{I}_+$  relative to  $\mathcal{I}_-$  became a paltry 27% as subjects ruminate (40; 41) negatively.

### 3.4 Reward seeking

The second factor is our restriction to just inhibition of trains of thought rather than a more fine-scale manipulation of the relative probabilities of different thoughts. We now relax this and explore the effect of additionally allowing preferential transitions towards certain states. In equation 6, for positive values, the parameter  $\theta$  biases action choice towards actions leading to positively valued states, whereas for negative values it does the opposite (i.e. subjects prefer to transition to negatively valued states). Figure 5B shows the effects of  $\theta$ . It is apparent that rather extreme values of  $\theta$  can both significantly aggravate or suppress the effect of  $\alpha_{5HT}$ . For the highest positive values of  $\theta$  the curves reverse shape, showing that it can be beneficial *not* to inhibit trains of thought. This arises since the model of figure 1 was chosen to have the extreme property that there is always the possibility of avoidance (in that all the states in  $\mathcal{I}_-$  admit at least one action that leads to  $\mathcal{I}_+$ ), and inhibiting trains of thought removes this outcome. A different, and rather counterintuitive, interaction between inhibition and reward seeking obtains in environments where rewards are hidden behind punishments (see supplementary material S1).

## 4 Discussion

We studied a very simple Markov decision process model of affectively-charged thoughts, and showed various aspects of the influence of behavioural inhibition on the experience of appetitive and aversive outcomes, predictions and prediction errors. The model formalises behavioural inhibition as a Pavlovian control process that arrests internally-directed thoughts (and likewise externally-directed actions) that are predicted to lead to aversive consequences. Overall this is favourable, leads to enhanced average rewards, and is related to adaptive pruning (33; 34). However, the consequences can also be deleterious (31; 32). Compromising inhibition in the model has two related consequences. First, the values of states are revealed to be overly optimistic. Second, control is disturbed, with aversive chains being insufficiently deselected.

While this work shows how several prominent aspects of serotonin's manifold putative functions and effects can be

reconciled within a unifying framework, we acknowledge that we have neglected a wide range of other issues, and certainly do not claim that this is an exhaustive account of the data. There are also interesting alternative view of 5-HT, such as that due to (42) who suggested that it is involved in controlling the appropriate timescale of behaviour by determining the discount factor for future affective outcomes (parameter  $\gamma$  in equation 2). In this theory, 5-HT depletion reduces the effective value of  $\gamma$ , making subjects appear more impulsive (43; 44; 45). Our model captures impulsivity through reduced 5-HT more directly, suggesting that actions that are comparatively worse lose direct inhibition that was previously restraining them, and are therefore more likely to be executed.

## 4.1 Behavioural Inhibition System

We suggested that this form of behavioural inhibition arises through predictions of aversive outcomes, tied to serotonin's putative role in reporting aversive prediction errors as an opponent to dopamine. This comes directly from the original notions of behavioural inhibition and serotonergic effects from Gray, Deakin, Graeff and their colleagues (13; 6; 7; 29; 30); however, it is perhaps best seen as a subset of the current version of Gray's Behavioural Inhibition System (BIS; 29). One salient difference is that BIS is suggested as being primarily engaged by *conflict*, rather than ongoing predictions of future aversive outcomes. Of course, a main source of conflict is that between approach and avoidance, with the latter coming from these aversive predictions. An interesting consequence of dividing the prediction of the value of future outcomes between two separate opponent systems is that it is indeed possible to have simultaneous appetitive and aversive expectations, as opposed to just one combined, net, prediction. Although we used the net prediction to control inhibition, it would be interesting to explore other possibilities associated with the BIS view, such as that any aversive prediction could arrest ongoing action, even if outweighed by appetitive predictions.

Further, rather than have the aversive predictive *values* of states lead to termination of trains of thought, it is possible that the negative prediction error ( $\delta_-$  from equation 8), which (10) suggested is being reported by phasic serotonin, could be responsible instead. Alternatively, in the mirror reflection of the proposal that a tonic dopaminergic signal reports average reward (and controllable/avoidable punishment) and energises behaviour (46; 47), it could be that a more tonic serotonergic signal, averaging aversion over longer time horizons and favoring quiescence, could be responsible.

Another difference between our account and the full BIS is that, in the latter, although actions are indeed inhibited in the face of conflict, the BIS is then suggested as initiating a set of behaviours (such as exploration or risk assessment) to resolve that conflict. The set of preparatory Pavlovian actions associated with aversive predictions appears to be more refined than that associated with appetitive predictions (mostly just approach), with a wide range of different defensive possibilities being selected between according to the nature and proximity of the threat (48; 30). One class of these is even laid out along columns of the dorsal periaqueductal gray (PAG; 49). Nevertheless, any of these defensive manoeuvres would interrupt the ongoing chain of actions, and this is what we modelled. Risk assessment and exploration are of most obvious use in the face of uncertainty and ignorance, whereas conditioned suppression, and thus the sort of inhibition that we consider, remains even after substantial learning. It would certainly be worth going one stage further, modelling the interruption in terms of a switch between different Markov decision problems, with new information changing the transition and payoff structures.

## 4.2 Tryptophan depletion

One of our central results is the effect of an acute reduction in  $\alpha_{5HT}$  *after* learning with elevated  $\alpha_{5HT}$  has taken place. In our model, this leads to a *decrease* in behavioural inhibition of actions leading to negative states. Although specific effects might arise from local manipulations of 5-HT concentrations or receptor responsivity, key data comes from the systemic manipulation associated with acute tryptophan depletion (TrD; 50), in which plasma levels of tryptophan and, at least in animals, CNS levels of serotonin, are drastically reduced (by up to 90%). Although the particular chains of thoughts analysed here have not been the subject of experimental scrutiny, there is by now a considerable body of literature on the effects of TrD on normal human functioning. In broad agreement with our results, various effects have been related to decreased reward processing (51; 39; 52), decreased behavioral inhibition (53; 54; 55; 56; 44; 57; 45), rumination (21), facial fear recognition (58) and, more indirectly, increased aggressiveness (54; 59; 60).

Perhaps of most direct relevance to our implementation are the results of a recent study which decoupled rewards from correct performance of an action from the outcomes of the actions (61). This study actually involved a sophisticated

assessment of the effects of TrD on reversal learning. However, one way of viewing a portion of the results stems from an abstract representation of the task. Subjects had to press one of two buttons (A or B) in response to one of two stimuli (also called A and B), with presses associated with A leading to a symbolic reward; and presses associated with B leading to a symbolic punishment. Critically, these outcomes were independent of the rectitude of the subjects' responses; so they couldn't avoid the punishment by making errors. In this case, subjects more often failed to press button B correctly than button A, and this difference disappeared after tryptophan depletion. This is directly consistent with the present interpretation of serotonergic inhibition of actions that lead to aversive outcomes.

Famously, tryptophan depletion does not have a uniform effect on all subjects. There is an important genetic polymorphism in the 5-HT reuptake mechanism, with subjects having the less efficient version generally showing greater effects (62; 63; 57; 52; 64; 65; 66). For this to be consistent with our formulation, the difference in functional 5-HT levels before and after TrD has to be greater in the subjects with less efficient reuptake. This in turn might most simply be due to increased levels of 5-HT (and behavioural inhibition) throughout development in carriers of the short 5HTTLPR allele. Perhaps related to this is the finding that TrD produces a dose-dependent relapse of depressive symptomatology in formerly depressed patients (18; 19; 41; 20), or in patients with risk factors such as a family history of depression (63) (although the three-way interaction between TrD, 5HTTLPR and past depression is hard to fit into this framework (67)).

There is a significant body of work on the effects of serotonergic manipulations on affective processing, particularly on processing of facial expressions (58; 68; 69; 70). It is difficult to interpret this work in our context for several reasons: first, there have often been effects on recognition of specific aversive facial expressions (e.g. fear) but not others (e.g. disgust). Our model does not speak to these distinctions. Secondly, in these tasks, subjects identify stimuli by pressing a button. Thus, there is a Pavlovian association between certain buttons and the aversive stimuli, and, interpreting these tasks in the same framework as we interpreted the work of Cools *et al.* (61), one might predict that TrD would increase, rather than decrease accuracy. The precise effect, however, would depend on the relative strength of the instructed and the reflexive Pavlovian response, and on the antagonism between the responses. Indeed, both aspects have been found: acute manipulation of serotonin increased recognition accuracy of fearful faces with increasing serotonin (58; 68; 70), whereas a more chronic increase in 5-HT (via SSRIs) yields a decrease in recollection of negative memories (69). Furthermore, while the exact relationship between behavioral inhibition and amygdala activation still needs clarification, it is additionally possible that increased amygdala activation may relate to perceptual mood congruency effects (38): after disinhibition, thoughts often visit negative states, and it is possible that this may affect prior expectations about stimulus which in turn could speed up processing of negatively valenced information.

Tryptophan depletion (or indeed selective serotonin reuptake inhibitors) have not previously been used in tasks like the Markov decision problem of the type we discussed. A direct prediction of the model is that subjects trained under TrD would explore states less when tested in a normal regime, while those trained under SSRIs would do so more (assuming that SSRIs indeed elevate 5-HT levels). Similar predictions hold for subjects with short or long alleles of the 5-HT reuptake mechanism on these tasks. This would essentially represent a generalisation of the findings by Cools *et al.* (61) to the domain of sequential decisions. The tasks could use external, observable, actions; more directly, it would also be useful to monitor the execution of affective trains of thought, and study the perturbation of this under serotonergic manipulations. In designing such studies, it is important to bear in mind the potentially opponent instrumental and Pavlovian effects, in just the same way that boosting dopamine and monitoring the effects on negative automaintenance may be confusing. Note that although there are various important datasets as to the effects of TrD on simple probabilistic and delay-discounting tasks (56; 71; 72; 73; 51; 74; 75; 52), these studies do not encompass the sorts of behavioral chains that we propose 5-HT to be able to halt.

### 4.3 Dopamine and Serotonin

One of the backdrops for the present theory was the extensive modeling of phasic dopamine as a prediction error for future reward, and the results that (a) the baseline firing rates of dopamine cells are insufficient to report prediction errors for negative rewards (i.e. punishments); (b) the ample psychological evidence for the existence of a pair of systems, one associated with appetitive outcomes and the other with aversive outcomes; and (c) the evidence that at least some aspects of 5-HT and dopamine are in mutual opposition. Indeed, based on these data and the theories of Deakin & Graeff (6; 7), Daw *et al.* (10) suggested that serotonin rather than dopamine reports negative prediction errors based on an antagonism between serotonin and dopamine at both a behavioral and pharmacological level. For

example, in rodents, 5-HT antagonises the general excitatory effects of DA (4), the self-administration of amphetamine and ICSS (76; 77), the effects of dopamine on appetitive learning (8) and the potentiation of appetitive learning by amphetamine (78).

However, pure opponency is far too simple. For instance, there is by now extensive evidence that 5-HT modulates dopaminergic activity both through receptors in the VTA and by modulation of distal release sites, and that this modulation can occur in both inhibitory and excitatory directions (4; 79; 80; 81; 82; 77; 83; 84; 85; 86; 87; 88; 89). Even the rise in 5-HT due to SSRIs has overall pro-dopaminergic effects, both at behavioural and physiological level (90; 91; 81; 92), and there is one report that DA antagonists reverse the antidepressant effect of SSRIs (93). Further, there is evidence that DA itself is released in many aversive circumstances (94; 95), and is involved in aversively-motivated behaviours like avoidance (96; 97; 98).

In our terms, apart from the aspects of the interaction of dopamine and 5-HT that were explored in figures 5b and 7, there are a couple of other effects. First, inhibition in our model has the consequence of increasing the average expected reward. As such, tonic dopamine, which has been suggested to report such a quantity (46; 99; 47; 100; 101), would be increased when 5-HT is boosted, and potentially vice-versa (92; 88). This would compete with the more direct effect that 5-HT inhibits actions, and particularly inhibits actions supported by dopaminergic predictions or rewards (8; 78; 102) and thus high levels of 5-HT might also depress levels of tonic dopamine, more in line with accounts that stress the opponent role of dopamine and serotonin (4; 10; 103).

The second complexity (Boureau, personal communication) is that active defense (such as active avoidance) requires energizing, and indeed appears to be controlled by the (presumably dopaminergically-reported) appetitive outcome of reaching a state of safety rather than the (presumably serotonergically-mediated) outcome of leaving a state of fear. That is, it appears that dopamine reports the rewards reaped from avoiding or controlling aversive outcomes (15; 94; 104).

We mentioned the mirror notion that the relationship between 5-HT and inhibition arising through aversive predictions is parallel to the obverse relationship for dopamine and engagement/approach through appetitive predictions (32). In this case, appetitively-directed chains of thoughts would be favored. Indeed, Smith *et al* (105; 106), in their work on the conditioned avoidance model of schizophrenia, suggested something rather like this. In their account, dopamine controls the extent of search through a forward model, although they did not couple this to dopamine's involvement in appetitive prediction.

In all, disentangling and elucidating these varied relations between dopamine and serotonin is a pressing task.

#### 4.4 Depression and anxiety

It would be reasonable to argue that the present model is more relevant to anxiety than depression. There is at best a somewhat fuzzy distinction between the two in terms of risk factors (107) and pharmacology (108), and they are extraordinarily comorbid (109). There is also no complete definition of either disease in terms of the sort of reinforcement mechanisms that we have been considering.

While depressive (but not anxious) symptoms can be reinduced by TrD in a subpopulation of patients, TrD it is not the only such manipulation; and it is not effective in all patients. Patients who are responsive to SNRIs are more sensitive to catecholamine depletion by  $\alpha$ -methyl-tyrosine (110; 111) than TrD and a recent report with a DA antagonist successfully re-induced depressive symptoms in formerly depressed people (93). The latter authors suggest that DA may be a "final common path" for depression, and may relate more to the depressive state than serotonin, which in turn may be more important in defining a trait (15; 112; 113). In addition, only 50% of formerly depressed subjects do respond to TrD (41; 114), and a pooled analysis of 71 formerly depressed subjects found that previous response to SSRIs had less predictive power for TrD response than chronicity of the depressive disorder and gender (114). As mentioned, resolving the actual relative contribution of serotonergic inhibition will be tricky.

#### 4.5 Conclusions

In sum, the findings in this study argue for an involvement of the serotonin reuptake mechanism in mood disorders such as anxiety and depression in the following manner: due to a decreased efficiency of the transporter, increased

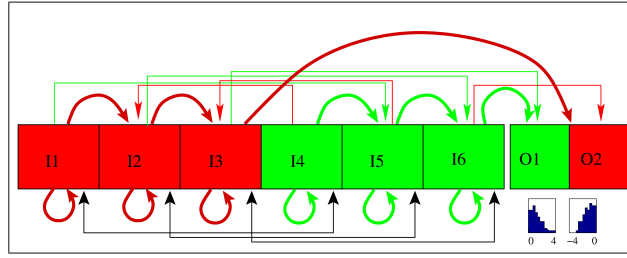


Figure 6: Similar state space to figure 1, but with a more explicitly deep structure. State in  $\mathcal{I}_+^1$  mainly lead to  $\mathcal{I}_+^2$ , or back to themselves. The last states in each of the two chains (here  $\mathcal{I}_+^3$  and  $\mathcal{I}_-^3$ ) always preferentially lead to the outcome state  $\mathcal{O}_+$  and  $\mathcal{O}_-$ .

behavioural inhibition results in acquisition of overly optimistic values. Such value functions are adaptive, but only in conjunction with strong behavioural inhibition. On the other hand, they do render the individual highly sensitive to large decreases in average experienced rewards when serotonin function is reduced. This might underlie a (controversially) larger sensitivity to TrD and SSRIs of persons with the short 5HTTLPR allele (see (115)). Returning to the sequential decision-making tasks suggested above, this study would predict that the short 5HTTLPR allele would be associated with more reflexive avoidance of states predictive of punishment, and it may be possible to assess this with differential effects of TrD on carriers with the short and long 5HTTLPR allele.

A further, more involved, conjecture, which returns to the fact that serotonin is not the sole causative agent in depression, is that it is the effects of reduced 5-HT on affective experience that leads to the various symptoms of depression, acting via the otherwise normative operation of the multiple systems involved in behavioural control. For instance, we have argued that the consequences of 5-HT reduction include unexpected punishments, large negative prediction errors, and a drop in average reward. These changes in the statistics of reward demand explanation, for example in terms of a shift in the characteristics of the environment, and should cause *normative* behavioural responses. In particular, the unsignalled aversion that comes independent of the subject’s actions can be seen as a form of uncontrollable punishment. Uncontrollability lies at the heart of an important characterization of depression centred around learned helplessness (116; 104).

We concentrated on the effects of reduced 5-HT rather than the reasons for this reduction. The obvious option is that it is a pathological result from processes operating at a purely cellular level. However, it could also arise as a normative meta-adaptation to the statistics of experienced punishments and rewards. Formalizing this fully would require a more general theory of inhibition – what level of inhibition is optimal? Tools for the characterisation of the trade-off between accurate knowledge about a state’s value and the cost incurred in learning about it are already in existence (34; 117; 118) and might be applicable to aspects of the present case.

## Acknowledgements

Funded by the Gatsby Charitable Foundation and a Bogue Research Fellowship. We are grateful to Y-Lan Boureau, Roshan Cools, Nathaniel Daw, Hanneke Den Ouden, Karl Friston, Michael Moutoussis, Jon Roiser, Barbara Sahakian, Douglas Steele, Jonathan Williams and Paul Willner for helpful discussions. We would also like to thank anonymous reviewers for helpful comments.

## Competing Interests

The authors declare no competing interests.

## S1 Supplement 1

The present implementation of inhibition could lead to short-sighted actions in deeper environments. To explore this, we consider an environment with explicit deep structure, where all large rewards are hidden behind small

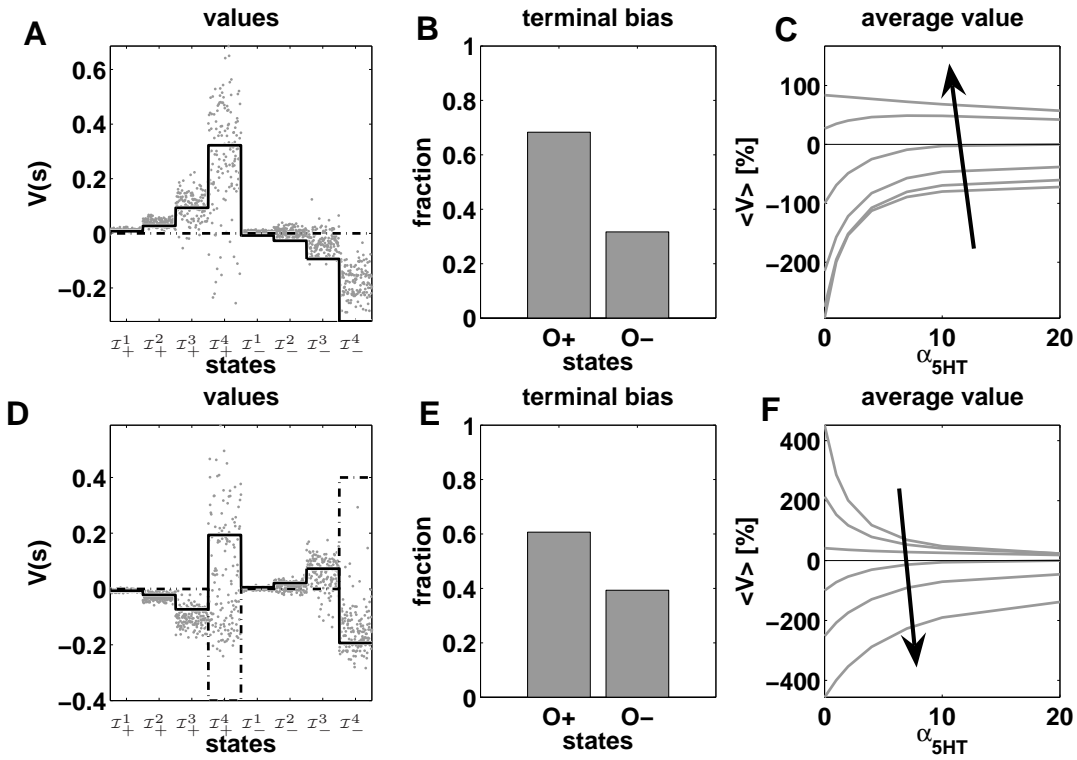


Figure 7: Inhibition in a deep environment. The outcomes  $\mathcal{O}$  are approached by sequentially walking through  $K = 4$  levels. Only  $\mathcal{I}^4$  states lead to outcomes. (A,D): True values without inhibition are shown by black line. It is constant for each level and valence as, or illustration, all outcomes were assigned the same positive value (+1 or -1). The reward of the states  $\mathcal{I}$  is zero and shown by the dash-dotted line. The grey point display the estimated values of the states under inhibition  $\alpha_{5HT} = 20$ . There is a positive bias in all states, but it is more pronounced in the states with true negative values. In (D), the dash-dotted line indicates that states  $\mathcal{I}_+^4$  now carry reward  $-0.4$ , while states  $\mathcal{I}_-^4$  carry reward  $+0.4$ . States  $\mathcal{I}_+^k$  for  $k = \{1, 2, 3\}$  now have true negative values and  $\mathcal{I}_-^k$  for  $k = \{1, 2, 3\}$  have true positive values. (B,E): Probabilities of ending thought sequence in  $\mathcal{O}_+$  or  $\mathcal{O}_-$ . (C,F): Effect of preferentially choosing actions according to their valence on the average value of states. The arrow indicates increasing  $\gamma$ . In (C), larger  $\gamma$  are advantageous, in (F), smaller  $\gamma$  are better.

punishments, and vice-versa. Large  $\alpha_{5HT}$  may prevent visits to the states associated with small punishments and thus prevent visits to the states behind them that are associated with large rewards. This effect is closely related to impulsivity, which itself has been associated with an impaired behavioural inhibitory system (55). We first show that all the conclusions up to now also apply to an environment of depth  $K$  in which the reward structure is still bipartite, i.e. in which no good states are primarily accessible via negative states.

Figure 6 shows the slightly augmented environment that we used to explore the effects of impulsivity. In this, the sets of non-terminal states are subdivided into  $K$  groups (figure 6 shows this for  $K = 3$ ). The last of the groups ( $\mathcal{I}_+^K$  is connected to the terminal states  $\mathcal{O}_-$ , as before; for the other groups with  $k < K$ , the role of  $\mathcal{O}_+$  is replaced by  $\mathcal{I}_+^{k+1}$ . The same applies to the states with negative valence. Identical probabilities as for figure 1A are used.

Specifically, we let states  $\mathcal{I}_+^k$  of positive valence and at level  $1 \leq k \leq K - 1$  preferentially lead to states in the same level and valence  $\mathcal{I}_+^k$  (with probability 3/8) or to states  $\mathcal{I}_+^{k+1}$  (with probability 3/8) of the same valence, but one level closer to the outcomes  $\mathcal{O}$ . With smaller probability (1/8) they can also lead to states of the opposite valence at the same level  $\mathcal{I}_-^k$  or one level up  $\mathcal{I}_-^{k+1}$ . States  $\mathcal{I}_-^{K-1}$  have connections as shown in figure 1B and are the only states that can lead to outcomes  $\mathcal{O}$ .

Figure 7A shows the true values of states without inhibition and their estimated values with inhibition. There is a clear positive bias for all negatively valued states. Figure 7B shows that the outcomes are still more frequently positive than negative, and figure 7C shows the effect of altering  $\gamma$ . As in figure 5B, negative  $\gamma$  lead to preferential selection of actions leading to states with negative values, and the overall average value is increased by increasing  $\gamma$ .

However, the situation changes if the states with large positive outcomes are primarily accessible through punished states. The dash-dotted line in figure 7D shows that the states  $\mathcal{I}_+^K$  at the final level  $K$  before the outcomes are now punished, while the states  $\mathcal{I}_-^K$  are rewarded. The values of states  $\mathcal{I}_+^3$  is now more negative than their counterpart without inhibition. This is because thoughts are often interrupted before they can proceed through (the punished)  $\mathcal{I}_+^4$  to (the rewarded)  $\mathcal{O}_+$ . Overall, the states in  $\mathcal{I}_+$  are now predominantly negative, and those in  $\mathcal{I}_-$  predominantly positive. However, inhibition does still lead to an overall positive outcome bias (figure 7E), and diminishing  $\alpha_{5HT}$  with  $\gamma = 0$  is still unfavourable (figure 7F, black line). However, the effect of  $\gamma$  is now reversed (see arrow in figure 7F): Negative  $\gamma$  (less reward seeking) will now predominantly lead to choices in  $\mathcal{I}_+$ , and thus to choices with longer-term positive outcomes (as there is no more inhibition). Thus, when  $\alpha_{5HT}$  is lowered in an environment in which rewards lurk behind punishments, a reduction of reward-seeking behaviour may play a role in adaptively compensating for the lack of inhibition.

## References

- [1] Solomon RL, Corbit JD (1974) An opponent-process theory of motivation. i. temporal dynamics of affect. *Psychol Rev* 81:119–145. [2](#)
- [2] Dickinson A, Dearing MF (1979) Appetitive-aversive interactions and inhibitory processes. In: Dickinson A, Boakes RA, editors, *Mechanisms of learning and motivation*, Hillsdale, NJ: Erlbaum. pp. 203–231. [2](#)
- [3] Dickinson A, Balleine B (2002) The role of learning in the operation of motivational systems. In: Gallistel R, editor, *Stevens' handbook of experimental psychology*, New York: Wiley, volume 3. pp. 497–534. [2](#)
- [4] Carter CJ, Pycocck CJ (1978) Differential effects of central serotonin manipulation on hyperactive and stereotyped behaviour. *Life Sci* 23:953–960. [2](#), [12](#)
- [5] Costall B, Hui SC, Naylor RJ (1979) The importance of serotonergic mechanisms for the induction of hyperactivity by amphetamine and its antagonism by intra-accumbens (3,4-dihydroxy-phenylamino)-2-imidazoline (dpi). *Neuropharmacology* 18:605–609. [2](#)
- [6] Deakin JFW (1983) Roles of brain serotonergic neurons in escape, avoidance and other behaviors. *J Psychopharmacol* 43:563–77. [2](#), [10](#), [11](#)
- [7] Deakin JFW, Graeff FG (1991) 5-HT and mechanisms of defence. *Journal of Psychopharmacology* 5:305–16. [2](#), [3](#), [10](#), [11](#)
- [8] Fletcher PJ (1996) Injection of 5-HT into the nucleus accumbens reduces the effects of d-amphetamine on responding for conditioned reward. *Psychopharmacology (Berl)* 126:62–9. [2](#), [12](#)
- [9] Kapur S, Remington G (1996) Serotonin-dopamine interaction and its relevance to schizophrenia. *Am J Psychiatry* 153:466–76. [2](#)

- [10] Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Networks* 15:603–16. 2, 10, 11, 12
- [11] Esposito E (2006) Serotonin-dopamine interaction as a focus of novel antidepressant drugs. *Curr Drug Targets* 7:177–185. 2
- [12] Soubrié P (1986) Reconciling the role of central serotonin neurons in human and animal behaviour. *Behav Brain Sci* 9:319–364. 2
- [13] Gray JA (1991) The psychology of fear and stress, volume 5 of *Problems in the behavioural sciences*. Cambridge, UK: Cambridge University Press, 2 edition. 2, 10
- [14] Schmajuk NA, Gray JA, Lam YW (1996) Latent inhibition: a neural network approach. *J Exp Psychol Anim Behav Process* 22:321–349. 2
- [15] Willner P (1985) *Depression: A psychobiological synthesis*. New York: John Wiley & Sons. 2, 12
- [16] Graeff FG, Guimaraes FS, De Andrade TGCS, Deakin JFW (1998) Role of 5HT in stress, anxiety and depression. *Pharm Biochem Behav* 54:129–41. 2
- [17] Maier SF, Watkins LR (2005) Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neurosci Biobehav Rev* 29:829–41. 2
- [18] Young SN, Smith SE, Pihl RO, Ervin FR (1985) Tryptophan depletion causes a rapid lowering of mood in normal males. *Psychopharmacology (Berl)* 87:173–177. doi:10.1007/BF00431803. 2, 11
- [19] Delgado PL, Charney DS, Price LH, Aghajanian GK, Landis H, et al. (1990) Serotonin function and the mechanism of antidepressant action. reversal of antidepressant-induced remission by rapid depletion of plasma tryptophan. *Arch Gen Psychiatry* 47:411–418. 2, 11
- [20] Moreno FA, Gelenberg AJ, Heninger GR, Potter RL, McKnight KM, et al. (1999) Tryptophan depletion and depressive vulnerability. *Biol Psychiatry* 46:498–505. 2, 11
- [21] Smith KA, Fairburn CG, Cowen PJ (1999) Symptomatic relapse in bulimia nervosa following acute tryptophan depletion. *Arch Gen Psych* 56:171–6. 2, 10
- [22] Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, et al. (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301:386–89. 2
- [23] Millan MJ (2006) Multi-target strategies for the improved treatment of depressive states: Conceptual foundations and neuronal substrates, drug discovery and therapeutic application. *Pharmacol Ther* 110:135–370. doi:10.1016/j.pharmthera.2005.11.006. 2
- [24] Sibille E, Lewis DA (2006) Sert-ainly involved in depression – but when? *Am J Psychiatry* 163:8–11. 2
- [25] Estes W, Skinner B (1941) Some quantitative aspects of anxiety. *J Exp Psychol* 29:390–400. 2
- [26] Fletcher PJ (1995) Effects of combined or separate 5,7-dihydroxytryptamine lesions of the dorsal and median raphe nuclei on responding maintained by a DRL 20s schedule of food reinforcement. *Brain Res* 675:45–54. 2
- [27] Fletcher PJ, Korth KM (1999) Activation of 5-HT<sub>1B</sub> receptors in the nucleus accumbens reduces amphetamine-induced enhancement of responding for conditioned reward. *Psychopharmacology* 142:165–74. 2
- [28] Graeff FG (2002) On serotonin and experimental anxiety. *Psychopharmacology (Berl)* 163:467–476. doi:10.1007/s00213-002-1112-4. 2
- [29] Gray JA, McNaughton N (2003) *The neuropsychology of anxiety*. OUP, 2nd edition. 2, 10
- [30] McNaughton N, Corr PJ (2004) A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neurosci Biobehav Rev* 28:285–305. doi:10.1016/j.neubiorev.2004.03.005. 2, 3, 10
- [31] Breland K, Breland M (1961) The misbehavior of organisms. *American Psychologist* 16:681–4. 3, 9
- [32] Dayan P, Niv Y, Seymour B, Daw ND (2006) The misbehavior of value and the discipline of the will. *Neural Netw* 19:1153–1160. doi:10.1016/j.neunet.2006.03.002. 3, 6, 9, 12
- [33] Knuth D, Moore R (1975) An Analysis of Alpha-Beta Pruning. *Artificial Intelligence* 6:293–326. 3, 9
- [34] Baum EB, Smith WD (1997) A bayesian approach to relevance in game playing. *Artificial Intelligence* 97:195–242. 3, 9, 13
- [35] Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press. 3, 5
- [36] Watkins CJCH (1989) *Learning from delayed rewards*. Ph.D. thesis, King’s College, Cambridge University. 5
- [37] Rachman S, Arntz A (1991) The overprediction and underprediction of pain. *Clinical Psychology Review* 11:339–55. 8
- [38] Blaney PH (1986) Affect and memory: a review. *Psychol Bull* 99:229–46. 8, 11
- [39] Klaassen T, Riedel WJ, Deutz NE, Van Praag HM (2002) Mood congruent memory bias induced by tryptophan depletion. *Psychol Med* 32:167–72. 8, 10
- [40] Nolen-Hoeksema S (1991) Responses to depression and their effects on the duration of depressive episodes. *J Abnorm Psychol* 100:569–82. 9
- [41] Smith KA, Fairburn CG, Cowen PJ (1997) Relapse of depression after rapid depletion of tryptophan. *Lancet* 249:915–9. 9, 11, 12

- [42] Doya K (2000) Metalearning, neuromodulation and emotion. In: Hatano G, Okada N, Ta H, editors, *Affective minds*, Amsterdam: Elsevier Science. pp. 101–4. [10](#)
- [43] Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, et al. (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* 7:887–93. [10](#)
- [44] Schweighofer N, Shishida K, Han CE, Okamoto Y, Tanaka SC, et al. (2006) Humans can adopt optimal discounting strategy under real-time constraints. *PLoS Comput Biol* 2:e152. doi:10.1371/journal.pcbi.0020152. [10](#)
- [45] Schweighofer N, Tanaka SC, Doya K (2007) Serotonin and the evaluation of future rewards: Theory, experiments, and possible neural mechanisms. *Ann N Y Acad Sci* doi:10.1196/annals.1390.011. [10](#)
- [46] Niv Y, Daw N, Dayan P (2005) How fast to work: Response vigor, motivation and tonic dopamine. In: *Advances in Neural Information Processing*, MIT Press. pp. 1019–26. [10](#), [12](#)
- [47] Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520. doi:10.1007/s00213-006-0502-4. [10](#), [12](#)
- [48] Blanchard DC, Blanchard RJ (1988) Ethoexperimental approaches to the biology of emotion. *Annu Rev Psychol* 39:43–68. doi:10.1146/annurev.ps.39.020188.000355. [10](#)
- [49] Bandler R, Shipley MT (1994) Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends Neurosci* 17:379–389. [10](#)
- [50] Bell C, Abrams J, Nutt D (2001) Tryptophan depletion and its implications for psychiatry. *Br J Psychiatry* 178:399–405. [10](#)
- [51] Murphy FC, Smith KA, Cowen PJ, Robbins TW, Sahakian BJ (2002) The effects of tryptophan depletion on cognitive and affective processing in healthy volunteers. *Psychopharm* 163:42–53. [10](#), [11](#)
- [52] Roiser JP, Blackwell AD, Cools R, Clark L, Rubinsztein DC, et al. (2006) Serotonin transporter polymorphism mediates vulnerability to loss of incentive motivation following acute tryptophan depletion. *Neuropsychopharmacology* 31:2264–2272. doi:10.1038/sj.npp.1301055. [10](#), [11](#)
- [53] LeMarquand DG, Benkelfat C, Pihl RO, Palmour RM, Young SN (1999) Behavioral disinhibition induced by tryptophan depletion in nonalcoholic young men with multigenerational family histories of paternal alcoholism. *Am J Psychiatry* 156:1771–1779. [10](#)
- [54] Bjork JM, Dougherty DM, Moeller FG, Swann AC (2000) Differential behavioral effects of plasma tryptophan depletion and loading in aggressive and nonaggressive men. *Neuropsychopharmacology* 22:357–369. doi:10.1016/S0893-133X(99)00136-0. [10](#)
- [55] Deakin JF (2003) Depression and antisocial personality disorder: two contrasting disorders of 5ht function. *J Neural Transm Suppl* 64:79–93. [10](#), [15](#)
- [56] Anderson IM, Richell RA, Bradshaw CM (2003) The effect of acute tryptophan depletion on probabilistic choice. *J Psychopharmacol* 17:3–7. doi:10.1177/0269881103017001687. [10](#), [11](#)
- [57] Hayward G, Goodwin GM, Cowen PJ, Harmer CJ (2005) Low-dose tryptophan depletion in recovered depressed patients induces changes in cognitive processing without depressive symptoms. *Biol Psychiatry* 57:517–524. doi:10.1016/j.biopsych.2004.11.016. [10](#), [11](#)
- [58] Harmer CJ, Rogers RD, Tunbridge E, Cowen PJ, Goodwin GM (2003) Tryptophan depletion decreases the recognition of fear in female volunteers. *Psychopharmacology (Berl)* 167:411–417. doi:10.1007/s00213-003-1401-6. [10](#), [11](#)
- [59] Walsh MT, Dinan TG (2001) Selective serotonin reuptake inhibitors and violence: a review of the available evidence. *Acta Psychiatr Scand* 104:84–91. [10](#)
- [60] Lesch KP, Merschdorf U (2000) Impulsivity, aggression, and serotonin: a molecular psychobiological perspective. *Behav Sci Law* 18:581–604. [10](#)
- [61] Cools R, Robinson OJ, Sahakian B (2007) Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharm In Press*:000. [10](#), [11](#)
- [62] Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, et al. (1996) Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 274:1527–31. [11](#)
- [63] Neumeister A, Konstantinidis A, Stastny J, Schwarz MJ, Vitouch O, et al. (2002) Association between serotonin transporter gene promoter polymorphism (5HTTLPR) and behavioral responses to tryptophan depletion in healthy women with and without family history of depression. *Arch Gen Psychiatry* 59:613–20. [11](#)
- [64] Roiser JP, Rogers RD, Cook LJ, Sahakian BJ (2006) The effect of polymorphism at the serotonin transporter gene on decision-making, memory and executive function in ecstasy users and controls. *Psychopharmacology (Berl)* 188:213–227. doi:10.1007/s00213-006-0495-z. [11](#)
- [65] Hariri AR, Holmes A (2006) Genetics of emotional regulation: the role of the serotonin transporter in neural function. *Trends Cog Sci* 10:182–91. [11](#)
- [66] Roiser JP, Mller U, Clark L, Sahakian BJ (2007) The effects of acute tryptophan depletion and serotonin trans-

- porter polymorphism on emotional processing in memory and attention. *Int J Neuropsychopharmacol* 10:449–461. doi:10.1017/S146114570600705X. 11
- [67] Neumeister A, Hu XZ, Luckenbaugh DA, Schwarz M, Nugent AC, et al. (2006) Differential effects of 5-HTTLPR genotypes on the behavioral and neural responses to tryptophan depletion in patients with major depression and controls. *Arch Gen Psych* 63:978–86. 11
- [68] Harmer CJ, Bhagwagar Z, Perrett DI, Vllm BA, Cowen PJ, et al. (2003) Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology* 28:148–152. doi:10.1038/sj.npp.1300004. 11
- [69] Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM (2004) Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *Am J Psychiatry* 161:1256–1263. 11
- [70] Browning M, Reid C, Cowen PJ, Goodwin GM, Harmer CJ (2007) A single dose of citalopram increases fear recognition in healthy subjects. *J Psychopharmacol* 21:684–690. doi:10.1177/0269881106074062. 11
- [71] Rogers RD, Blackshaw AJ, Middleton HC, Matthews K, Hawtin K, et al. (1999) Tryptophan depletion impairs stimulus-reward learning while methylphenidate disrupts attentional control in healthy young adults: implication for the monoaminergic basis of impulsive behaviour. *Psychopharm* 146:428–91. 11
- [72] Mobini S, Chiang TJ, Al-Ruwaitea AS, Ho MY, Bradshaw CM, et al. (2000) Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology* 149:313–8. 11
- [73] Mobini S, Chiang TJ, Ho MY, Bradshaw CM, Szabadi E (2000) Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* 152:390–7. 11
- [74] Rogers RD, Tunbridge EM, Bhagwagar Z, Drevets WC, Sahakian BJ, et al. (2003) Tryptophan depletion alters the decision-making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology* 28:153–62. 11
- [75] Cools R, Blackwell A, Clark L, Menzies L, Cox S, et al. (2005) Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* 30:1362. 11
- [76] Redgrave P (1978) Modulation of intracranial self-stimulation behaviour by local perfusions of dopamine, noradrenaline and serotonin within the caudate nucleus and nucleus accumbens. *Brain Res* 155:277–95. 12
- [77] Higgins GA, Fletcher PJ (2003) Serotonin and drug reward: focus on 5-HT<sub>2c</sub> receptors. *Eur J Pharmacol* 480:151–162. 12
- [78] Fletcher PJ, Korth KM, Chambers JW (1999) Selective destruction of brain serotonin neurons by 5,7-dihydroxytryptamine increases responding for a conditioned reward. *Psychopharmacology (Berl)* 147:291–9. 12
- [79] Parsons LH, Justice Jr JB (1993) Perfusate serotonin increases extracellular dopamine in the nucleus accumbens as measured by in vivo microdialysis. *Brain Res* 606:195–99. 12
- [80] Galloway MP, Suchowski CS, Keegan MJ, Hjorth S (1993) Local infusion of the selective 5-HT<sub>1b</sub> agonist cp-93,129 facilitates striatal dopamine release in vivo. *Synapse* 15:90–2. 12
- [81] D’Aquila PS, Collu M, Gessa GL, Serra G (2000) The role of dopamine in the mechanism of action of antidepressant drugs. *Eur J Pharmacology* 405:365–73. 12
- [82] Di Matteo V, Cacchio M, Di Giulio C, Esposito E (2002) Role of serotonin(2C) receptors in the control of brain dopaminergic function. *Pharmacol Biochem Behav* 71:727–734. 12
- [83] Di Matteo V, Di Giovanni G, Di Mascio M, Esposito E (1999) SB 242084, a selective serotonin<sub>2C</sub> receptor antagonist, increases dopaminergic transmission in the mesolimbic system. *Neuropharmacology* 38:1195–205. 12
- [84] Gobert A, Rivet J, Lejeune F, Newman-Tancredi A, Adhumeau-Auclair A, et al. (2000) Serotonin (2C) receptors tonically suppress the activity of mesocortical dopaminergic and adrenergic, but not serotonergic, pathways: a combined dialysis and electrophysiological analysis in the rat. *Synapse* 36:205–21. 12
- [85] Grottick A, Fletcher P, Higgins G (2000) Studies to Investigate the Role of 5-HT<sub>2C</sub> Receptors on Cocaine- and Food-Maintained Behavior. *Journal of Pharmacology and Experimental Therapeutics* 295:1183–1191. 12
- [86] Bortolozzi A, Díaz-Mataix L, Scorza M, Celada P, Artigas F (2005) The activation of 5-HT<sub>2A</sub> receptors in prefrontal cortex enhances dopaminergic activity. *J Neurochem* . 12
- [87] Diaz-Mataix L, Scorza M, Bortolozzi A, Toth M, Celada P, et al. (2005) Involvement of 5-HT<sub>1A</sub> Receptors in Prefrontal Cortex in the Modulation of Dopaminergic Activity: Role in Atypical Antipsychotic Action. *Journal of Neuroscience* 25:10831–10843. 12
- [88] Benaliouad F, Kapur S, Rompr PP (2006) Blockade of 5-HT<sub>2a</sub> receptors reduces haloperidol-induced attenuation of reward. *Neuropsychopharmacology* doi:10.1038/sj.npp.1301136. 12
- [89] Fletcher P, Sinyard J, Higgins G (2006) The effects of the 5-HT<sub>2C</sub> receptor antagonist SB242084 on locomotor activity induced by selective, or mixed, indirect serotonergic and dopaminergic agonists. *Psychopharmacology*

- 187:515–525. [12](#)
- [90] Serra G, Argiolas A, Klimek V, Fadda F, Gessa GL (1979) Chronic treatment with antidepressants prevents the inhibitory effect of small doses of apomorphine on dopamine synthesis and motor activity. *Life Sci* 25:415–423. [12](#)
- [91] Besson A, Privat AM, Eschali er A, Fialip J (1999) Dopaminergic and opioidergic mediations of tricyclic antidepressants in the learned helplessness paradigm. *Pharmacol Biochem Behav* 64:541–8. [12](#)
- [92] Sasaki-Adams DM, Kelley AE (2001) Serotonin-dopamine interactions in the control of conditioned reinforcement and motor behaviour. *Neuropsychopharm* 25:440–52. [12](#)
- [93] Willner P, Hale AS, Argyropoulos S (2005) Dopaminergic mechanism of antidepressant action in depressed patients. *J Affect Disord* 86:37–45. [12](#)
- [94] Cabib S, Puglisi-Allegra S (1996) Stress, depression and the mesolimbic dopamine system. *Psychopharmacology* 128:331–42. [12](#)
- [95] Horvitz JC (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96:651–6. [12](#)
- [96] Beninger R, Mason S, Phillips A, Fibiger H (1980) The use of extinction to investigate the nature of neuroleptic-induced avoidance deficits. *Psychopharmacology* 69:11–18. [12](#)
- [97] Prinssen EPM, Kleven MS, Koek W (1996) Effects of dopamine antagonists in a two-way active avoidance procedure in rats: interactions with 8-OH-DPAT, ritanserin, and prazosin. *Psychopharmacology* 128:191–7. [12](#)
- [98] Smith A, Li M, Becker S, Kapur S (2006) Dopamine, prediction error and associative learning: a model-based account. *Network* 17:61–84. doi:10.1080/09548980500361624. [12](#)
- [99] Niv Y, Daw ND, Dayan P (2006) Choice values. *Nat Neurosci* 9:987–988. doi:10.1038/nn0806-987. [12](#)
- [100] Floresco SB, West AR, Ash B, Moore H, Grace AA (2003) Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci* 6:968–973. doi:10.1038/nn1103. [12](#)
- [101] Goto Y, Grace AA (2005) Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* 8:805–812. doi:10.1038/nn1471. [12](#)
- [102] Fletcher PJ, Azampanah A, Korth KM (2002) Activation of 5-HT(1B) receptors in the nucleus accumbens reduces self-administration of amphetamine on a progressive ratio schedule. *Pharmacol Biochem Behav* 71:717–21. [12](#)
- [103] Ungless MA (2004) Dopamine: the salient issue. *Trends Neurosci* 27:702–6. [12](#)
- [104] Huys QJM (2007) Reinforcers and control. Towards a computational  etiology of depression. Ph.D. thesis, Gatsby Computational Neuroscience Unit, UCL, University of London. [12](#), [13](#)
- [105] Smith A, Li M, Becker S, Kapur S (2004) A model of antipsychotic action in conditioned avoidance: a computational approach. *Neuropsychopharm* 29:1040–9. [12](#)
- [106] Smith AJ, Becker S, Kapur S (2005) A computational model of the functional role of the ventral-striatal d2 receptor in the expression of previously acquired behaviors. *Neural Comput* 17:361–95. [12](#)
- [107] Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS (2006) A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry* 163:857–864. doi:10.1176/appi.ajp.163.5.857. [12](#)
- [108] Ressler KJ, Nemeroff CB (2000) Role of serotonergic and noradrenergic systems in the pathophysiology of depression and anxiety disorders. *Depression Anxiety* 12:2–19. [12](#)
- [109] Kaufman J, Charney D (2000) Comorbidity of mood and anxiety disorders. *Depression Anxiety* 12:69–76. [12](#)
- [110] Delgado PL (2000) Depression: the case for a monoamine deficiency. *J Clin Psychiatry* 61 Suppl 6:7–11. [12](#)
- [111] Iversen L (2005) The monoamine hypothesis of depression. In: Licinio J, Wong ML, editors, *Biology of depression*, Weinheim, Germany: Wiley, volume 1. pp. 71–86. [12](#)
- [112] Heinz A (1999) Anhedonie – nosologie ubergreifendes Korrelat einer Dysfunktion des dopaminergen Verst arkungssystem? *Nervenarzt* 70:391–398. [12](#)
- [113] Willner P (2002) Dopamine and depression. In: Chiara GD, editor, *Handbook of Physiology: Dopamine in the CNS.*, Berlin: Springer. pp. 387–416. [12](#)
- [114] Booij L, der Does WV, Benkelfat C, Bremner JD, Cowen PJ, et al. (2002) Predictors of mood response to acute tryptophan depletion. a reanalysis. *Neuropsychopharmacology* 27:852–861. doi:10.1016/S0893-133X(02)00361-5. [12](#)
- [115] Kraft J, Peters E, Slager S, Jenkins G, Reinalda M, et al. (2007) Analysis of association between the serotonin transporter and antidepressant response in a large clinical sample. *Biol Psychiatry* 61:734–42. [13](#)
- [116] Abramson LY, Alloy LB, Hogan ME, Whitehouse WG, Cornette M, et al. (1998) Suicidality and cognitive vulnerability to depression among college students: a prospective study. *J Adolesc* 21:473–487. doi:10.1006/jado.1998.0167. [13](#)
- [117] Dearden R, Friedman N, Russell S (1998) Bayesian Q-learning. In: *Proceedings of the fifteenth National*

- Conference on Artificial Intelligence. pp. 761–8. 13
- [118] Dearden R, Friedman N, Andre D (1999) Model-based Bayesian exploration. In: Proceedings of the fifteenth Conference on Uncertainty in Artificial Intelligence. Stockholm, pp. 150–9. 13