Text S1 Impulsivity in a deep environment.

Supplementary Text S1 for Dayan P, Huys QJM (2008) Serotonin, inhibition and negative mood. PLoS Comput Biol 4(1):e4 doi:10.1371/journal.pcbi.0040004.sd001

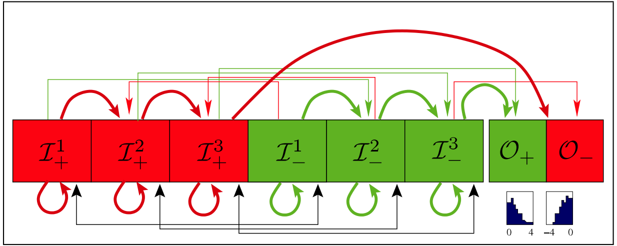
The present implementation of inhibition could lead to short-sighted actions in deeper environments. To explore this, we consider an environment with explicit deep structure, where all large rewards are hidden behind small punishments, and vice-versa. Large α_{5HT} may prevent visits to the states associated 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 positively valued states are primarily accessible via negatively valued states.

Supplementary Figure 1 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 (Supplementary Figure 1 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 $\mathcal{O}$.

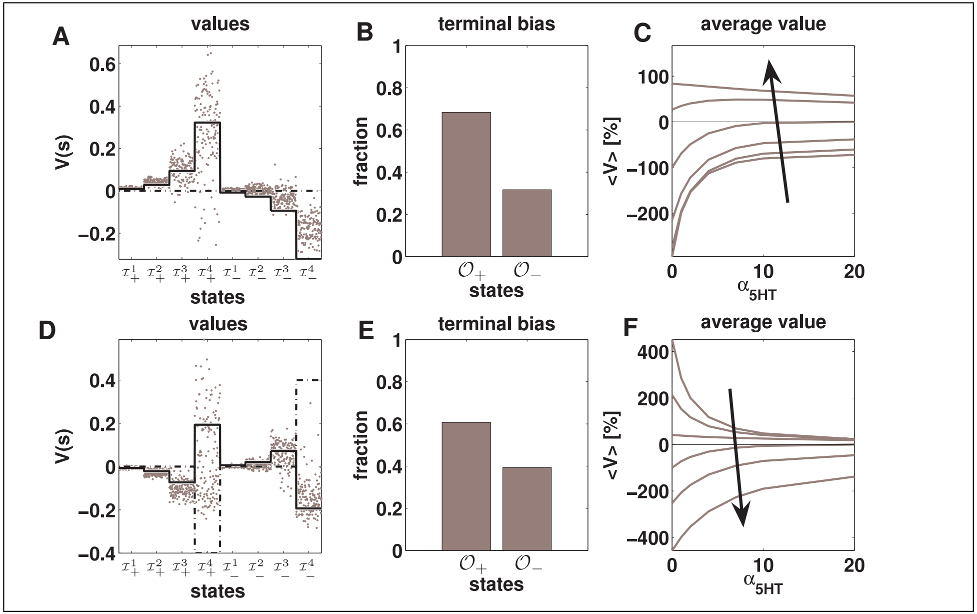
Supplementary Figure 2A shows the true values of states without inhibition and their estimated values with inhibition. There is a clear positive bias for all negatively valued states. Supplementary Figure 2B shows that the outcomes are still more frequently positive than negative, and Supplementary Figure 2C shows the effect of altering θ. As in Figure 5B, negative θ lead to preferential selection of actions leading to states with negative values, and the overall average value is increased by increasing θ.

However, the situation changes if the states with large positive outcomes are primarily accessible through punished states. The dash-dotted line in Supplementary Figure 2D 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 (Supplementary Figure 2E), and diminishing α_{5HT} with $\theta = 0$ is still unfavourable (Supplementary Figure 2F, black line). However, the effect of θ is now reversed (see arrow in Supplementary Figure 2F): Negative θ (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 α_{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.



Supplementary Figure 1: 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 $\mathcal{O}_-$.

Supplementary Figure 1 for Dayan P, Huys QJM (2008) Serotonin, inhibition and negative mood. PLoS Comput Biol 4(1):e4 doi:10.1371/journal.pcbi.0040004.sg001



Supplementary Figure 2: 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 θ. In (C), larger θ are advantageous, in (F), smaller θ are better.

Supplementary Figure 2 for Dayan P, Huys QJM (2008) Serotonin, inhibition and negative mood. PLoS Comput Biol 4(1):e4 doi:10.1371/journal.pcbi.0040004.sg002