

## ***Pallidal circuits for aversive motivation and learning***

### **Highlights**

- *The GPh and VP play complementary roles in processing aversive information*
- *GPh encodes punishment prediction errors and updates action values*
- *VP encodes incentive and aversive state values and drives motivation*
- *GABAergic and glutamatergic VP neurons drive approach and avoidance respectively*

### **Abstract**

The basal ganglia are a group of subcortical nuclei that play a critical role in goal-directed behaviour by setting motivation, adjusting the vigour of actions, and driving reinforcement learning. Detailed knowledge exists of how these nuclei contribute to the different aspects of goal-directed behaviour in the context of reward-seeking, but far less is known about how they work in a framework of avoiding threats. This review will highlight recent work that has begun exploring how the non-motor output of the basal ganglia, from the ventral pallidum (VP) and the habenula-projecting globus pallidus (GPh), are essential in driving the motivation to avoid punishment and the learning from aversive outcomes. We consider the VP and GPh in turn and discuss the circuitry and neural encoding that support their functions, while underscoring the unanswered questions in the field.

### **Introduction**

A prominent trait of animals is their ability to predict rewards and punishments, and to change their behavior accordingly, such that they can obtain reward and avoid threats. In response to stereotyped rewarding and threatening cues animals display a range of hard-wired species-specific behavioral responses. Equally animals can learn to flexibly adapt their behavior to obtain rewards and avoid threats. The basal ganglia play a key role in flexibly adapting behavior to obtain reward [1-3]. Electrophysiological recordings have identified neurons in the basal ganglia that encode the value of actions [4, 5]. These representations develop with learning and are thought to provide the basis for selecting the most valuable action [6-8]. These values are learnt through experience, with dopaminergic input to the basal ganglia providing a reinforcement signal that strengthens the representation of actions that lead to reward [9, 10].

In the case of avoiding threats, it is equally important for animals to flexibly adapt their behavior but the mechanisms that support this are less well known. Evidence is emerging that the basal ganglia are vital for driving the motivation to avoid threats as well as for learning from aversive outcomes. In this review we discuss the evidence that two basal ganglia output nuclei the VP and GPh play complimentary roles in avoiding threats, with a subset of VP neurons driving avoidance behavior by encoding motivational state value and the GPh encoding punishment prediction errors to update action values.

### ***Ventral Pallidum***

The VP is the major output structure of the ventral basal ganglia [11]. It receives GABAergic projections from the nucleus accumbens (NAc) and transmits information to multiple brain regions involved in motor control and motivation such as the ventral tegmental area (VTA), lateral habenula (LHb), thalamic mediodorsal nucleus and pedunculo-pontine tegmental nucleus [12]. This connectivity places the VP in an ideal location to transform information about the expected value of stimuli into motivation [13] and has led to proposals that it forms the “final common path” for drug seeking [14] and for reward processing more

generally [15]. Indeed, a large body of work, comprehensively reviewed by others [15-17], has identified the VP as a crucial driver of reward-seeking behaviour. For example, the VP is important for the normal hedonic reactions to sucrose [18], and lesions to the VP decrease an animal's willingness to work for reward [18, 19]. Conversely, rats will work to electrically self-stimulate their VP [20, 21] and pharmacological activation and disinhibition can both trigger feeding in sated animals [22].

In addition to seeking reward an at least equally important motivational drive is the desire to avoid punishment. Experimental evidence suggests that the VP also plays a role in this motivational processes as pharmacological disruption of intra-VP mu-opioid activity is sufficient to drive conditioned place aversion [23] and activating mu-opioid receptors in the VP can impair conditioned taste avoidance [24]. In a similar manner, disinhibiting the VP through injections of the GABAergic antagonist bicuculline induce anxiety-related behaviours and increase avoidance in an approach/avoidance task in primates [25, 26]. Conversely, reducing the glutamatergic input to the VP can reduce anxiety and despair-like behaviours in rats [27]. D1 receptor agonists injected into the VP can also enhance inhibitory avoidance learning [28]. Together this suggests that the VP plays a role in the motivation to both seek reward and avoid punishment.

Theoretical work has suggested that the motivation to obtain reward or the desire to escape a threat are driven by the same process as in both cases the incentive is a positive outcome, i.e obtaining reward or safety [29-31]. However, in the case of the VP recent work has shown that approach and avoidance are controlled by genetically distinct populations of neurons. Optogenetic activation of GABAergic neurons in the VP, as with non-specific VP stimulation, can drive self-stimulation and real-time place preference [32, 33]. In contrast, optogenetic activation of glutamatergic VP neurons has an opposing effect and drives real-time place avoidance [32, 33]. The glutamatergic VP neurons also adaptively constrain reward seeking as they are necessary for limiting the effort expended to obtain a reward and for learning to avoid sucrose rewards that are paired with nausea inducing lithium chloride [33]. As the glutamatergic VP neurons can drive avoidance behavior it is possible that the GABAergic and glutamatergic VP neurons may also work in concert in an aversive setting with the GABAergic neurons adaptively constraining avoidance behavior when there is potential for reward. This suggests that glutamatergic and GABAergic VP populations may encode the drive for approach and avoidance respectively with the balance between their activity determining the overall behavioral strategy.

Interestingly, the projection pattern of these two VP populations is qualitatively similar, suggesting that they exert their motivational effects through opposing roles on common downstream structures (figure 1). Both the GABAergic and glutamatergic neurons project to the lateral habenula, VTA and rostromedial tegmental nucleus (RMTg) [33]. Indeed, stimulating the GABAergic or glutamatergic axons in the VTA drives reward or aversive responses respectively [32]. In contrast, while optogenetic stimulation of the glutamatergic VP terminals in the LHb is aversive [32, 34] optogenetic stimulation of the GABAergic input to the LHb has not been reported to be rewarding [32]. In both the LHb and the VTA the GABAergic and glutamatergic neurons make synaptic connections with greater than fifty percent of the neurons recorded suggesting that they may synapse on the same population of neurons [32]. If this is indeed the case it would provide a simple integrative mechanism for downstream targets to determine the balance between the GABAergic and glutamatergic VP activity.

In the VTA the VP glutamatergic projection neurons synapse directly onto both dopaminergic and GABAergic neurons, while the net effect of activating the glutamatergic VP neurons in vivo is to excite putative VTA GABA neurons and inhibit putative DA neurons [33]. This inhibition of dopamine neurons could contribute to the aversive responses driven by the glutamatergic VP input, since optogenetic inhibition of the VTA DA neurons alone induces real-time place avoidance and drives negative prediction error [35, 36]. However, the dopaminergic response to glutamatergic VP input is likely more complex. For one, dopamine release increases in a number of brain areas in response to aversive stimuli, including in the

dorsal lateral striatum [37], medial prefrontal cortex [38, 39], the medial shell of the NAc [40] and the tail of the striatum [41]. In addition, dopaminergic activity is needed for active avoidance, and blocking dopamine receptor activity in the medial shell of the NAc can impair avoidance behaviour and promote passive defensive strategies [29, 42]. Therefore, if glutamatergic VP neurons are needed for motivating active avoidance this population cannot simply suppress dopamine. One possibility is that as only a subset of midbrain dopaminergic neurons are activated by aversive stimuli glutamatergic VP input to the VTA could differentially affect subpopulations of dopamine neurons [39, 43-45]. In this case, the VP could inhibit the classic value-coding dopamine neurons, that are inhibited by aversive stimuli, and excite dopamine neurons that are activated by aversive stimuli. This suggests that aversive and appetitive motivation could be implemented in part by the different target regions that are modulated by dopamine in an appetitive or aversive context.

The aversive nature of the glutamatergic VP input to the VTA may also be mediated by non-dopaminergic neurons in the VTA. Recent experiments have shown that a population of glutamatergic neurons in the VTA project to the medial shell of the NAc and that photoactivation of this pathway drives aversion [46]. This effect is due to excitation of parvalbumin interneurons that in turn inhibit spiny projection neurons in the NAc [46].

Consistent with a role in both appetitive and aversive motivation, *in vivo* recordings in the VP have identified two main types of neurons that are activated by the prediction of either reward or punishment [25, 47]. A subset these of VP neurons are active in a sustained manner from the time of a cue predicting a rewarding outcome until the outcome occurs, while a second population are inhibited in a sustained manner in response to appetitive outcomes [19, 25, 47, 48]. While it has not been demonstrated it is tempting to speculate that these functional classes correspond to the GABAergic and glutamatergic VP neurons. These activity patterns led Hikosaka and colleagues to propose that the VP neurons encode the current value of a given state rather than the difference between the actual and expected values, which is calculated by reward prediction error-coding dopamine neurons [43, 47]. Interestingly, such sustained activity in response to reward- and punishment-predictive cues has also been observed in the dorsal raphe [49]. Furthermore, this sustained activity has been observed in optogenetically-identified serotonergic neurons and was shown to change dynamically with the motivational state [50]. This could be driven in part by the VP as the VP projects to the medial portions of the LHb, a region that in turn project strongly to the dorsal raphe [16]. The sustained activity in the VP and DR would be useful for maintaining motivation in either a rewarding or aversive setting. Indeed, during reward-seeking the activity of the VP neurons is predictive of response latencies and the vigor of animal behaviour during instrumental tasks [19, 47, 48].

Another hallmark of VP neuronal activity is that neurons are only excited by reward predictive cues when the available reward is “wanted” [15, 26, 51]. For example, high concentration salt solutions are normally aversive for rats but when a rat is salt depleted, high salt concentration solutions become appetitive. Correspondingly VP neurons are not activated by cues predicting unpalatable salt solutions but after salt depletion cues gain an incentive value and VP neurons are robustly activated by the prediction of the salt solution [51]. This coding of incentive salience or “wanting” may be encoded by GABAergic VP neurons and raises the possibility that glutamatergic VP neurons in contrast could encode aversive salience. Together the recording data suggests that ventral pallidal neurons dynamically track the incentive and aversive value on both long and short timescales to set the motivational state.

Together there is emerging evidence that the VP plays a fundamental role in driving the motivation to not only seek reward, but to also avoid punishment. These motivational states are driven by GABAergic and glutamatergic neurons, which may encode incentive and aversive state values respectively. It is tempting to speculate that these two populations work together to balance the behavioural requirements, constraining

reward seeking when there are risks/costs involved and limiting avoidance behavior when there is potential for reward.

### ***Habenula-projecting globus pallidus (GPh)***

Another key aspect of evading threats, in addition to motivating avoidance, is the ability to learn what actions to choose in order to avoid a threat [30, 31]. LHb is an important area of the brain involved in learning from aversive outcomes [52]. Neurons in this nucleus encode negative prediction errors, such they are phasically excited by the prediction of punishment or when an outcome is worse than expected [53, 54]. Activation of the LHb is aversive and can promote negative reinforcement [55] and leads to inhibition of reward prediction error coding dopaminergic neurons through a di-synaptic connection including the RMTg [56, 57]. Tian and colleagues showed with lesions of the LHb and in vivo recording of dopamine neurons that inhibition of dopamine when an outcome is worse than expected is dependent on the LHb [58]. These dips in dopamine when an outcome is worse than expected are sufficient to drive negative reinforcement [36].

This role in learning from negative outcomes may at least in part be driven by the basal ganglia. The LHb receives a large excitatory projection from basal ganglia output neurons located in the globus pallidus interna (GPi) (also called the entopeduncular nucleus in mice) [59-62]. These GPh neurons are located in the rostral two thirds of the GPi in mice and in the border region between the GPi and globus pallidus externa (GPe) in primates [59, 63, 64]. A subpopulation of GPh neurons encode negative reward prediction error signals and this activity precedes the activity in the LHb [60, 63, 64]. Optogenetically stimulating the GPh can drive activity in the lateral portion of the LHb and is aversive [61]. In a probabilistic switching task optogenetic activation of the GPh can devalue choices and preventing the excitation of the GPh impairs the ability of mice to use negative feedback to guide their decisions in the task [60]. Together this suggests that the GPh is crucial for aversive learning as it drives the activity in the LHb when an outcome is aversive or when it is worse than expected.

Beyond learning to avoid punishment, the GPh and LHb neurons play a broader role in reinforcement learning. Neurons in both nuclei are also inhibited by the prediction or unexpected delivery of reward [60, 63] and optogenetic inhibition of the GPh or inhibition of the LHb is rewarding and can reinforce actions [60, 65]. In line with this inactivation of the LHb does not merely impair the processing of aversive information but rather prevents mice from forming a preference between choices with different subjective value [66].

Recent single cell sequencing data has revealed that the GPh is actually composed of two genetically distinct populations [67]. One type of GPh neurons expresses somatostatin and co-releases glutamate and GABA, a second population expresses parvalbumin and are exclusively glutamatergic [67]. Interestingly in primates two functional classes of GPh neurons have also been identified, a reward negative type that bi-directionally encodes prediction errors and a reward positive type that display sustained activation to reward predictive information [63]. How these genetic and functional populations relate to each other still needs to be determined as does their individual roles in behaviour.

As the GPh bi-directionally responds to cues of opposing valence it suggests the GPh is able to integrate reward and punishment information to evaluate whether an action was better or worse than expected. Indeed cell-type-specific synaptic manipulations in mice have demonstrated that inhibitory and excitatory inputs to the GPh are necessary for mice to appropriately learn from positive and negative feedback, respectively [60]. This suggests that reward and punishment information is encoded in separate circuits upstream of the GPh, with the GPh representing the key point of convergence for information of opposing valence. Rabies tracing from the form the GPh and GPi has demonstrated that the circuits that control the motor output of

the basal ganglia are distinct from the circuits that regulate the GPh [60, 67] (figure 2). Both types of GPh neurons receive direct inhibitory input from both striosome and matrix compartments of the striatum, while the GPi receives inhibitory input exclusively from the matrix [60, 67]. As GABAergic inputs to the GPh are necessary for processing of positive feedback [60], the direct inhibitory input from the striatum to the GPh may provide the information about predicted rewards. One candidate for this inhibition are striosomes as striatal neurons provide direct inhibitory input to the GPh [60, 67] and striosomes have been proposed to play a role in reinforcement learning [68, 69]. Indeed, striosomal neurons are crucial for the rewarding effect of opioid drugs [70] and loss of the mu opioid receptors on striosomes prevents both the rewarding effect of morphine and abolishes the induced dopamine release in the NAc [70].

Until recently, directly testing the function of striosomes has been hampered by the lack of methods to specifically target these neurons. The development of non-inducible and embryonically inducible transgenic mouse lines that specifically express GFP or Cre-recombinase in striosomal neurons now permits an investigation of their function [71, 72]. The role of striosomes in encoding predicted rewards was directly tested for the first time in two recent studies, each using in vivo calcium imaging but different genetic strategies to specifically record from striosomal neurons. Both studies showed that striosomal neurons encode information about the expected reward that develops with learning [72, 73]. Striosome neurons are therefore excellent candidates for inhibiting the GPh when an outcome is predicted to lead to reward.

The best candidates for exciting the GPh when an outcome is aversive or when it is worse than expected are distinct circuits associated with the 'indirect pathway'. The GPi, which receives excitatory input from the core of the subthalamic nucleus (STN) and input from parvalbumin-positive prototypical GPe neurons [67]. The GPh receives glutamatergic input from the subthalamic cells located in the 'limbic' region of the STN, on the medial border of this nucleus and in the surrounding parasubthalamic nucleus (pSTN) and input from mainly parvalbumin-negative GPe neurons [60, 67] (figure 2). The pSTN is a good candidate for providing information about negative valence to the GPh as glutamatergic input to GPh neurons is needed for mice to process negative feedback [60]. Furthermore, a subpopulation of subthalamic cells are activated by aversive outcomes or when an outcome is worse than expected [74]. This suggests information of about predicted rewards and punishments may be carried by the direct and indirect pathway inputs to the GPh respectively. This hypothesis is supported by optogenetic experiments, where stimulating the direct and indirect pathways not only influences movement parameters but also drives reward and punishment respectively [75]. Interestingly predicted punishment activity is also over represented in striosomes as compared to matrix and may be encoded by indirect pathway striosomal neurons [73]. These indirect striosomal neurons could provide the input to the indirect limbic pathway consisting of the PV-ve GPe neurons and the pSTN to convey information about predicted punishment to the GPh.

Together this model suggests that direct and indirect striosomal neurons encode the predicted benefits and costs associated with certain stimuli. This information of opposing valence would then be integrated in the GPh to give the total predicted value of how good or how bad an outcome is expected to be. In line with this hypothesis, individual striosomal neurons independently encode information about predicted rewards or punishments [73]. Furthermore, Graybiel and colleagues showed that a prefrontal-striosomal pathway in rats plays a specific role in decisions making when there are both costs and benefits associated with particular choices [76, 77]. The GPh is a likely candidate for where information about costs and benefits is combined to predict the relative value of choices or determine whether the outcome was better or worse than expected.

Conclusion

Theories of avoidance have noted that there are two separate types of learning that occur in active avoidance [29, 78-80]. First animals need to learn the value of states so that threatening situations can be predicted and avoided as well as motivating escape if threats are encountered. Second, in case an animal does encounter a threatening situation it needs to learn which actions to choose to escape the threat. Interestingly the VP and the GPh could fulfill these different roles. VP glutamatergic neurons can drive avoidance by encoding negative motivational states. If these neurons not only encode the current state value but also serve as a teaching signal to learn the predicted value of an aversive state, then animals could learn to avoid negative states before encountering a threat. In contrast, GPh neurons encode reward and punishment prediction errors. In classic reinforcement learning models RPE's are used to update the value of actions such that in any given context animals are able to choose the most valuable action [9, 81, 82]. Indeed, manipulations have shown that RPE signaling in the GPh can bias action selection by updating action values [60]. In this way if an animal were to avoid a threat, the outcome would be better than expected and the resulting GPh RPE signal could positively reinforce the action that led to the threat being avoided. Taken together, the VP and GPh could have complementary roles in learning, with the VP learning motivational state values and the GPh updating action values. Such a dichotomy may also occur within subpopulations of dopamine neurons as a recent study has shown that VTA and SNc dopaminergic neurons are able to respectively reinforce general motivation or invigorate specific actions [83]. This suggests that the VP and GPh could be nodes of two extended reinforcement systems: one for learning state value and another for learning action value.

**Figure 1.** Schematic of the major GABAergic and glutamatergic projections from the ventral pallidum and the GPh as well as some selected additional circuitry. This schematic highlights the circuitry discussed in the review and is by no means exhaustive.

**Figure 2.** Schematic showing the proposed selection and evaluation circuits within the basal ganglia (adapted from ref 60). Here the direct and indirect pathway striosomal neurons are proposed to independently encode reward and punishment information. The information of opposing valence is then integrated in the GPh. For brevity the circuit linking the GPh and the dopaminergic neurons is left out. This circuit contains a projection to the LHB and then direct projection from the LHB to dopaminergic neurons as well as di-synaptic connections to dopamine neurons via the RMTg.

#### Paper summaries

Ref 32 \*\* This is the first paper to highlight the opposing roles of GABAergic and glutamatergic VP neurons in reward and aversion. Viral tracing and path clamp recording are used to demonstrate the similarity of the projections of both populations of VP neurons indicating that downstream circuits may be sensitive to the balanced activity of the two pathways.

Ref 33 \*\* First paper to show that VP glutamatergic neurons can drive aversive behaviour. Viral ablation of the VP glutamatergic neurons also demonstrates that these neurons are necessary for constraining the

energy expended in the pursuit of reward suggesting the GABAergic and glutamatergic VP populations adaptively work together to control reward-seeking.

Ref 34 \* A detailed analysis of a subpopulation of VP neurons that express parvalbumin. This population is predominantly glutamatergic but also contains some GABAergic neurons. They map the projections and valence of the VP projections but also add some excellent analysis (not mentioned in the review) of the alterations of the VP in depression.

Ref 60 \*\* First recording of GPh neurons in mice and first demonstration of an alternative basal ganglia circuit that regulates the GPh, with striosomes and specific populations of indirect pathway neurons preferentially innervating the GPh. Cell-type specific manipulations also demonstrate the influence the GPh has on updating action values.

Ref 67 \*\* Tour de force using single cell sequencing and viral tracing to show that there are two genetically distinct populations of GPh neurons. One that expresses somatostatin and co-releases glutamate and GABA and another that expresses parvalbumin and is glutamatergic.

Ref 72 \*\* A landmark paper that uses an inducible cre-line to label striosomes and demonstrate with two-photon calcium imaging that they preferentially encode information related to the prediction of reward. The paper also demonstrates a new method for targeting striosomes by combining an inducible cre line with tamoxifen inductions at different embryonic time points to birthdate label different populations of striatal neurons.

Ref 73 \* This paper demonstrates that different populations of striosome neurons encode information related to the prediction of reward and punishment, with the responses proportional to the reward or punishment intensity. The activity develops during learning with different populations of striosomes neurons active in either the early or late phase of reward learning.

Ref 76 \* Paper describing the information encoded by striosomes during a conflict task. Manipulations of the afferent input to striosomes is used to demonstrate these neurons play a specific role in cost-benefit decision making.

1. Hikosaka, O., et al., *Basal ganglia circuits for reward value-guided behavior*. Annu Rev Neurosci, 2014. **37**: p. 289-306.
2. Turner, R.S. and M. Desmurget, *Basal ganglia contributions to motor control: a vigorous tutor*. Curr Opin Neurobiol, 2010. **20**(6): p. 704-16.
3. Graybiel, A.M., *The basal ganglia: learning new tricks and loving it*. Curr Opin Neurobiol, 2005. **15**(6): p. 638-44.
4. Lau, B. and P.W. Glimcher, *Value representations in the primate striatum during matching behavior*. Neuron, 2008. **58**(3): p. 451-63.

5. Samejima, K., et al., *Representation of action-specific reward values in the striatum*. Science, 2005. **310**(5752): p. 1337-40.
6. Znamenskiy, P. and A.M. Zador, *Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination*. Nature, 2013. **497**(7450): p. 482-5.
7. Barnes, T.D., et al., *Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories*. Nature, 2005. **437**(7062): p. 1158-61.
8. Pasupathy, A. and E.K. Miller, *Different time courses of learning-related activity in the prefrontal cortex and striatum*. Nature, 2005. **433**(7028): p. 873-6.
9. Lee, D., H. Seo, and M.W. Jung, *Neural basis of reinforcement learning and decision making*. Annu Rev Neurosci, 2012. **35**: p. 287-308.
10. Lau, B., T. Monteiro, and J.J. Paton, *The many worlds hypothesis of dopamine prediction error: implications of a parallel circuit architecture in the basal ganglia*. Curr Opin Neurobiol, 2017. **46**: p. 241-247.
11. Heimer, L., et al., *Substantia innominata: a notion which impedes clinical-anatomical correlations in neuropsychiatric disorders*. Neuroscience, 1997. **76**(4): p. 957-1006.
12. Haber, S.N. and B. Knutson, *The reward circuit: linking primate anatomy and human imaging*. Neuropsychopharmacology, 2010. **35**(1): p. 4-26.
13. Mogenson, G.J., D.L. Jones, and C.Y. Yim, *From motivation to action: functional interface between the limbic system and the motor system*. Prog Neurobiol, 1980. **14**(2-3): p. 69-97.
14. Kalivas, P.W. and N.D. Volkow, *The neural basis of addiction: a pathology of motivation and choice*. Am J Psychiatry, 2005. **162**(8): p. 1403-13.
15. Smith, K.S., et al., *Ventral pallidum roles in reward and motivation*. Behav Brain Res, 2009. **196**(2): p. 155-67.
16. Root, D.H., et al., *The ventral pallidum: Subregion-specific functional anatomy and roles in motivated behaviors*. Prog Neurobiol, 2015. **130**: p. 29-70.
17. Humphries, M.D. and T.J. Prescott, *The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward*. Prog Neurobiol, 2010. **90**(4): p. 385-417.
18. Farrar, A.M., et al., *Forebrain circuitry involved in effort-related choice: Injections of the GABAA agonist muscimol into ventral pallidum alter response allocation in food-seeking behavior*. Neuroscience, 2008. **152**(2): p. 321-30.
19. Richard, J.M., et al., *Ventral Pallidum Neurons Encode Incentive Value and Promote Cue-Elicited Instrumental Actions*. Neuron, 2016. **90**(6): p. 1165-1173.
20. Panagis, G., et al., *Ventral pallidum self-stimulation: a moveable electrode mapping study*. Behav Brain Res, 1995. **68**(2): p. 165-72.
21. Panagis, G., et al., *Ventral pallidum self-stimulation induces stimulus dependent increase in c-fos expression in reward-related brain regions*. Neuroscience, 1997. **77**(1): p. 175-86.
22. Stratford, T.R., A.E. Kelley, and K.J. Simansky, *Blockade of GABAA receptors in the medial ventral pallidum elicits feeding in satiated rats*. Brain Res, 1999. **825**(1-2): p. 199-203.
23. Skoubis, P.D. and N.T. Maidment, *Blockade of ventral pallidal opioid receptors induces a conditioned place aversion and attenuates acquisition of cocaine place preference in the rat*. Neuroscience, 2003. **119**(1): p. 241-9.

24. Inui, T. and T. Shimura, *Activation of mu-opioid receptors in the ventral pallidum decreases the negative hedonic evaluation of a conditioned aversive taste in rats*. Behav Brain Res, 2017. **320**: p. 391-399.
25. Saga, Y., et al., *Ventral Pallidum Encodes Contextual Information and Controls Aversive Behaviors*. Cereb Cortex, 2017. **27**(4): p. 2528-2543.
26. Smith, K.S. and K.C. Berridge, *The ventral pallidum and hedonic reward: neurochemical maps of sucrose "liking" and food intake*. J Neurosci, 2005. **25**(38): p. 8637-49.
27. Chang, C.H. and A.A. Grace, *Amygdala-ventral pallidum pathway decreases dopamine activity after chronic mild stress in rats*. Biol Psychiatry, 2014. **76**(3): p. 223-30.
28. Péczely, L., et al., *Role of D1 dopamine receptors of the ventral pallidum in inhibitory avoidance learning*. Behav Brain Res, 2014. **270**: p. 131-6.
29. Lloyd, K. and P. Dayan, *Safety out of control: dopamine and defence*. Behav Brain Funct, 2016. **12**(1): p. 15.
30. Boureau, Y.L. and P. Dayan, *Opponency revisited: competition and cooperation between dopamine and serotonin*. Neuropsychopharmacology, 2011. **36**(1): p. 74-97.
31. Dayan, P., *Instrumental vigour in punishment and reward*. Eur J Neurosci, 2012. **35**(7): p. 1152-68.
32. Faget, L., et al., *Opponent control of behavioral reinforcement by inhibitory and excitatory projections from the ventral pallidum*. Nat Commun, 2018. **9**(1): p. 849.
33. Tooley, J., et al., *Glutamatergic Ventral Pallidal Neurons Modulate Activity of the Habenula-Tegmental Circuitry and Constrain Reward Seeking*. Biol Psychiatry, 2018.
34. Knowland, D., et al., *Distinct Ventral Pallidal Neural Populations Mediate Separate Symptoms of Depression*. Cell, 2017. **170**(2): p. 284-297.e18.
35. Danjo, T., et al., *Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens*. Proc Natl Acad Sci U S A, 2014. **111**(17): p. 6455-60.
36. Chang, C.Y., et al., *Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors*. Nat Neurosci, 2016. **19**(1): p. 111-6.
37. Lerner, T.N., et al., *Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits*. Cell, 2015. **162**(3): p. 635-47.
38. Lammel, S., et al., *Input-specific control of reward and aversion in the ventral tegmental area*. Nature, 2012. **491**(7423): p. 212-7.
39. Lammel, S., et al., *Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli*. Neuron, 2011. **70**(5): p. 855-62.
40. Badrinarayan, A., et al., *Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell*. J Neurosci, 2012. **32**(45): p. 15779-90.
41. Menegas, W., et al., *Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice*. Elife, 2017. **6**.
42. Faure, A., et al., *Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens*. J Neurosci, 2008. **28**(28): p. 7184-92.
43. Matsumoto, M. and O. Hikosaka, *Two types of dopamine neuron distinctly convey positive and negative motivational signals*. Nature, 2009. **459**(7248): p. 837-41.

44. Lammel, S., B.K. Lim, and R.C. Malenka, *Reward and aversion in a heterogeneous midbrain dopamine system*. *Neuropharmacology*, 2014. **76 Pt B**: p. 351-9.
45. Brischoux, F., et al., *Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli*. *Proc Natl Acad Sci U S A*, 2009. **106**(12): p. 4894-9.
46. Qi, J., et al., *VTA glutamatergic inputs to nucleus accumbens drive aversion by acting on GABAergic interneurons*. *Nat Neurosci*, 2016. **19**(5): p. 725-733.
47. Tachibana, Y. and O. Hikosaka, *The primate ventral pallidum encodes expected reward value and regulates motor action*. *Neuron*, 2012. **76**(4): p. 826-37.
48. Richard, J.M., et al., *Ventral pallidal encoding of reward-seeking behavior depends on the underlying associative structure*. *Elife*, 2018. **7**.
49. Bromberg-Martin, E.S., O. Hikosaka, and K. Nakamura, *Coding of task reward value in the dorsal raphe nucleus*. *J Neurosci*, 2010. **30**(18): p. 6262-72.
50. Cohen, J.Y., M.W. Amoroso, and N. Uchida, *Serotonergic neurons signal reward and punishment on multiple timescales*. *Elife*, 2015. **4**.
51. Tindell, A.J., et al., *Ventral pallidum firing codes hedonic reward: when a bad taste turns good*. *J Neurophysiol*, 2006. **96**(5): p. 2399-409.
52. Hikosaka, O., *The habenula: from stress evasion to value-based decision-making*. *Nat Rev Neurosci*, 2010. **11**(7): p. 503-13.
53. Matsumoto, M. and O. Hikosaka, *Lateral habenula as a source of negative reward signals in dopamine neurons*. *Nature*, 2007. **447**(7148): p. 1111-5.
54. Matsumoto, M. and O. Hikosaka, *Negative motivational control of saccadic eye movement by the lateral habenula*. *Prog Brain Res*, 2008. **171**: p. 399-402.
55. Stamatakis, A.M. and G.D. Stuber, *Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance*. *Nat Neurosci*, 2012. **15**(8): p. 1105-7.
56. Hong, S., et al., *Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates*. *J Neurosci*, 2011. **31**(32): p. 11457-71.
57. Jhou, T.C., et al., *The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses*. *Neuron*, 2009. **61**(5): p. 786-800.
58. Tian, J. and N. Uchida, *Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors*. *Neuron*, 2015. **87**(6): p. 1304-1316.
59. Stephenson-Jones, M., et al., *Independent circuits in the basal ganglia for the evaluation and selection of actions*. *Proc Natl Acad Sci U S A*, 2013. **110**(38): p. E3670-9.
60. Stephenson-Jones, M., et al., *A basal ganglia circuit for evaluating action outcomes*. *Nature*, 2016. **539**(7628): p. 289-293.
61. Shabel, S.J., et al., *Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin*. *Neuron*, 2012. **74**(3): p. 475-81.
62. Shabel, S.J., et al., *Mood regulation. GABA/glutamate co-release controls habenula output and is modified by antidepressant treatment*. *Science*, 2014. **345**(6203): p. 1494-8.
63. Hong, S. and O. Hikosaka, *The globus pallidus sends reward-related signals to the lateral habenula*. *Neuron*, 2008. **60**(4): p. 720-9.
64. Rajakumar, N., K. Elisevich, and B.A. Flumerfelt, *Compartmental origin of the striato-entopeduncular projection in the rat*. *J Comp Neurol*, 1993. **331**(2): p. 286-96.

65. Stamatakis, A.M., et al., *A unique population of ventral tegmental area neurons inhibits the lateral habenula to promote reward*. *Neuron*, 2013. **80**(4): p. 1039-53.
66. Stopper, C.M. and S.B. Floresco, *What's better for me? Fundamental role for lateral habenula in promoting subjective decision biases*. *Nat Neurosci*, 2014. **17**(1): p. 33-5.
67. Wallace, M.L., et al., *Genetically Distinct Parallel Pathways in the Entopeduncular Nucleus for Limbic and Sensorimotor Output of the Basal Ganglia*. *Neuron*, 2017. **94**(1): p. 138-152.e5.
68. Morita, K., et al., *Reinforcement learning: computing the temporal difference of values via distinct corticostriatal pathways*. *Trends Neurosci*, 2012. **35**(8): p. 457-67.
69. Keiflin, R. and P.H. Janak, *Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry*. *Neuron*, 2015. **88**(2): p. 247-63.
70. Cui, Y., et al., *Targeted expression of  $\mu$ -opioid receptors in a subset of striatal direct-pathway neurons restores opiate reward*. *Nat Neurosci*, 2014. **17**(2): p. 254-61.
71. Gerfen, C.R., R. Paletzki, and N. Heintz, *GENSAT BAC cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits*. *Neuron*, 2013. **80**(6): p. 1368-83.
72. Bloem, B., et al., *Two-photon imaging in mice shows striosomes and matrix have overlapping but differential reinforcement-related responses*. *Elife*, 2017. **6**.
73. Yoshizawa, T., M. Ito, and K. Doya, *Reward-Predictive Neural Activities in Striatal Striosome Compartments*. *eNeuro*, 2018. **5**(1).
74. Breyse, E., Y. Pelloux, and C. Baunez, *The Good and Bad Differentially Encoded within the Subthalamic Nucleus in Rats(1,2,3)*. *eNeuro*, 2015. **2**(5).
75. Kravitz, A.V., L.D. Tye, and A.C. Kreitzer, *Distinct roles for direct and indirect pathway striatal neurons in reinforcement*. *Nat Neurosci*, 2012. **15**(6): p. 816-8.
76. Friedman, A., et al., *A Corticostriatal Path Targeting Striosomes Controls Decision-Making under Conflict*. *Cell*, 2015. **161**(6): p. 1320-33.
77. Friedman, A., et al., *Chronic Stress Alters Striosome-Circuit Dynamics, Leading to Aberrant Decision-Making*. *Cell*, 2017. **171**(5): p. 1191-1205.e28.
78. Mower, O., *On the dual nature of learning: a reinterpretation of "conditioning" and "problem-solving"*. . 1947: *Harv Educ Rev*. p. 102-50.
79. Dinsmoor, J., *Punishment I. the avoidance hypothesis*. 1954: *Psychol Rev*. p. 34-46.
80. Miller, N., *Studies of fear as an acquirable drive: I. fear as motivation and fear-reduction as reinforcement in the learning of new responses*. 1948: *J Exp Psychol*. p. 89-101.
81. Sutton, R., *Reinforcement Learning: An Introduction*, A. Barto, Editor. 1998, MIT Press, Cambridge MA.
82. Daw, N.D. and K. Doya, *The computational neurobiology of learning and reward*. *Curr Opin Neurobiol*, 2006. **16**(2): p. 199-204.
83. Saunders, B.T., et al., *Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties*. *Nat Neurosci*, 2018. **21**(8): p. 1072-1083.