



Published in final edited form as:

*Neuron*. 2019 December 04; 104(5): 987–999.e4. doi:10.1016/j.neuron.2019.08.040.

## Three rostromedial tegmental afferents drive triply dissociable aspects of punishment learning and aversive valence encoding

Hao Li<sup>1,2,\*</sup>, Peter J. Vento<sup>1,\*</sup>, Jeffrey Parrilla-Carrero<sup>1</sup>, Dominika Pullmann<sup>1</sup>, Ying S. Chao<sup>1</sup>, Maya Eid<sup>1</sup>, Thomas C. Jhou<sup>1</sup>

<sup>1</sup>Medical University of South Carolina, Department of Neuroscience, Charleston, SC 29425

<sup>2</sup>Current address: The Salk Institute for Biological Studies, La Jolla, CA, 92037

### Summary

Persistence of reward-seeking despite punishment or other negative consequences is a defining feature of mania and addiction, but brain regions implicated in these processes have overlapping yet disparate functions that are difficult to reconcile. We now show that the ability of an aversive punisher to inhibit reward-seeking depends on coordinated activity of three distinct afferents to the rostromedial tegmental nucleus (RMTg) arising from cortex, brainstem, and habenula that drive triply dissociable RMTg responses to aversive cues, outcomes, and prediction errors, respectively. The RMTg in turn drives negative, but not positive, valence encoding patterns in the ventral tegmental area (VTA). Hence, different aspects of aversive processing are both coordinated yet dissociated from each other and from reward processing.

### eTOC blurb:

Survival requires one to obtain rewards but also cease reward-seeking when it incurs high costs or punishment. Li, Vento et al. show that punishment learning requires highly coordinated activity patterns in distinct afferents and efferents of the rostromedial tegmental nucleus.

### Introduction

Seeking of rewards and avoidance of punishment are both essential for survival, and extensive research has investigated the neural bases of reward, most often examining

**Corresponding author:** Dr. Thomas C. Jhou, (843) 876-2290, jhou@musc.edu, 70 President St. Drug Discovery Bldg DD221, Charleston, SC 29425.

\*Both authors contributed equally to this work

#### Author Contributions

H.L., P.J.V, J.P.-C., D.P., M.E., and Y.S.C., conducted experiments; H.L. and P.J.V designed the experiments and wrote the paper; H.L., P.J.V and T.C.J. reviewed and edited the paper.

Conceptualization, H.L., P.J.V. and T.C.J.; Methodology, H.L., P.J.V. and T.C.J.; Investigation, H.L., P.J.V, J.P.C., D.P., M.E., and Y.S.C.; Writing – Original Draft, H.L. and P.J.V.; Writing – Review & Editing, H.L., P.J.V. and T.C.J.; Funding Acquisition, T.C.J. and P.J.V.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of Interests

No conflicts in interests

dopamine (DA) neurons and their various afferents and efferents (Hu, 2016; Watabe-Uchida et al., 2017). However, many severe and intractable neuropsychiatric disorders result when reward-seeking persists despite high costs or punishments, as in mania, substance abuse or gambling disorders (American Psychiatric Association, 2013). Despite the importance of these phenomena, relatively fewer studies have examined the neural circuitry of punishment or of failures to withhold reward-seeking in the face of punishment.

We recently showed that punishment learning greatly depends upon the rostromedial tegmental nucleus (RMTg), a major GABAergic afferent to midbrain DA neurons (Balcita-Pedicino et al., 2011; Jhou et al., 2013; Smith et al., 2019; Vento et al., 2017). In particular, lesions or optogenetic inactivation of the RMTg or of its projections to the ventral tegmental area (VTA) abolish the ability of cocaine's aversive effects to impede drug-seeking, and increase by 48–145% the intensity of footshock required to suppress ongoing reward-seeking. Furthermore, the RMTg exhibits anatomic connections to numerous structures involved in aversive learning or punishment, including a notable afferent from the prelimbic cortex (PL) and a strikingly dense input from the lateral habenula (LHb) (Jhou et al., 2009a; Jhou et al., 2009b; Kaufling et al., 2009), both of which are implicated in many of the same functions as the RMTg, including depression-like behavior, avoidance learning, punishment, and driving of encoding patterns in VTA DA neurons (Chang et al., 2018; Chen et al., 2013; Hong et al., 2011; Jhou et al., 2013; Lammel et al., 2012; Li et al., 2019b; Matsumoto and Hikosaka, 2007, 2009a; Stamatakis and Stuber, 2012; Trusel et al., 2019a). However, lesions or inactivations of the PL or LHb fail to abolish several aversive behaviors and VTA encoding patterns that are otherwise severely impaired by RMTg lesions or inactivations (Heldt and Ressler, 2006; Jean-Richard-Dit-Bressel and McNally, 2016; Jhou et al., 2009a; Li et al., 2019a; Tian and Uchida, 2015). Hence, despite their interconnectedness, and functional overlap, the roles of the PL, LHb, and RMTg are likely also to differ, but in ways that are not well understood. Additional studies have speculated that the RMTg and LHb interact with the brainstem parabrachial nucleus (PBN) to encode punishment signals in presumptive midbrain DA neurons (Campos et al., 2018; Coizet et al., 2010), but again the exact natures of these interactions are unknown. Finally, the RMTg role in punishment behavior may depend on its projections to VTA DA neurons (Balcita-Pedicino et al., 2011; Vento et al., 2017) whose inhibitions by punishment-related stimuli play major roles in aversive learning (Chang et al., 2016; Danjo et al., 2014). However, the RMTg contribution to these encoding patterns are also unknown, as midbrain DA responses to cues predicting punishment have been posited to arise not only from the RMTg, but also from at least three other sources (Tian et al., 2016).

Hence, prior studies have implicated many brain regions in punishment learning, but in disparate ways that are difficult to reconcile. Prior studies also used different behavioral and physiological paradigms that further complicate efforts to compare across studies. Thus, we sought an integrated understanding of punishment learning by using a uniform set of behavioral, electrophysiological, and optogenetic paradigms to examine five interconnected regions: the RMTg, VTA, LHb, PL, and PBN. We found that learning from aversive experiences depends upon a striking convergence onto the RMTg of distinct but complementary signals arising from the PL, PBN, and LHb that convey triply dissociable information about aversive cues, outcomes, and prediction errors, respectively. These

pathways in turn drive correspondingly distinct but coordinated phases of punishment learning behavior. We also found that the RMTg drives encoding in VTA neurons of punishment-related, but not reward-related, stimuli, while also substantially targeting non-VTA structures. Hence, punishment learning is driven by interacting circuits that exhibit distinct but highly coordinated roles.

## Results

### RMTg neurons encode valence or salience

Prior studies showed that RMTg neurons play multiple roles in acquisition and performance of punishment learning (Vento et al., 2017), possibly due to their activation by multiple types of negative motivational signals, including footshocks, shock-predictive cues, and unexpected reward omission (Jhou et al., 2009a). We hypothesized that these various roles would be driven by various afferents (Jhou et al., 2009b; Kaufling et al., 2009), but particularly the LHB due to its activation by highly similar stimuli as the RMTg (Hong et al., 2011; Matsumoto and Hikosaka, 2007, 2009a). Hence, we recorded from 273 neurons in the RMTg, as delineated by immunostaining for the transcription factor FOXP1 (Lahti et al., 2016; Smith et al., 2018). Recordings were performed both before and after inactivation of the LHB, and in response to rewards, shocks, and auditory cues (1kHz or 8kHz tones, counterbalanced) predicting them (Figures 1A–D, 2A–D, S1A, D–E). Prior to afferent inactivation, many RMTg neurons encoded motivational “valence”, as previously noted (Hong et al., 2011; Jhou et al., 2009a; Li et al., 2019a), with a majority of stimulus-responsive neurons showing inhibitions to reward-predictive cues, and activations by shocks, shock cues, or reward omissions (Figures 1E–H). Furthermore, many of these responses occurred in the same neurons, as a majority of reward cue-inhibited neurons were also activated by shocks or shock cues, or both (Figure S1B), while among reward cue-inhibited neurons, the magnitude of inhibitions to reward- and shock-predictive cues correlated inversely with each other (Figure S1C). In these neurons, neutral cues (predicting no outcome) produced much smaller responses, yielding a monotonic encoding of valence to cues predicting reward, no outcome, and shock (Figure S1I). Notably, a smaller proportion of neurons showed *activations* by reward-predictive cues, with responses to reward- and shock-predictive cues correlating *positively* with each other (Figure S1C) while neutral cues again produced minimal responses (Figure S1J), a pattern consistent with “salience” encoding. Notably, previous findings in mice indicate that RMTg neurons encoding valence- and salience-like patterns project preferentially to the VTA or dorsal raphe nucleus (DRN), respectively (Li et al., 2019a).

Neurons recorded just outside the FOXP1-delineated RMTg region (n=677) showed very different response patterns than those within the RMTg, and showed no overall discrimination between reward and shock-predictive cues as had been seen in the RMTg (Figure S1F).

### The LHB drives RMTg responses to unpredicted negative motivational events

Surprisingly, RMTg neurons, independent of their encoding patterns (Figure S5C), still showed strong excitations to shocks and shock cues after either pharmacological inactivation

of the LHb, or pathway-specific optogenetic inactivation of LHb terminals in the RMTg. In particular, during a time window 0–30ms after shocks or shock cues (when neural responses were most prominent), neither inactivation method reduced RMTg excitations to shocks (whether cued or uncued) nor to shock cues (Figures 2E–G). However, we also observed a slower RMTg response component 30–100ms after shock cues and uncued (but not cued) shocks, that was markedly attenuated by both pharmacological and optogenetic LHb inactivation (Figures 2E, J, G, L). Because this 30–100ms response was much larger after uncued than cued shocks (Figure 2M), we hypothesized it represented an aversive *prediction error* present only after surprising stimuli. This delayed response was also present after cues, but we could not determine whether it was surprise-related, as all cues were unpredictable due to randomization of trial types. Hence, we tested an additional group of animals using “double-cued” aversive trials in which two consecutive cues (4kHz and 8kHz) were presented 0.5 seconds apart and a shock delivered immediately after the second cue (Figure 2N). In these trials, we indeed observed that the 30–100ms response was greatly attenuated after the second (predictable) but not first (surprising) cue (Figure 2O), consistent with it representing a prediction error.

In addition to RMTg responses to surprising shocks and cues, we conducted a third test of the RMTg response to surprising stimuli, by omitting rewards after cue presentations in a randomly selected 20% of reward trials. These unexpected omissions activated RMTg neurons up to 500ms after the time of expected reward delivery, and this response was completely abolished by LHb inactivation (Figure 2I). Hence, the LHb contributes a slow RMTg response to at least three types of surprising stimuli (shock cues, uncued shocks, and reward omissions), while playing no role in the more rapid (<30ms) responses to shocks or shock cues.

Importantly, LHb inactivations did not alter RMTg inhibitions by reward cues (Figure 2H), even though reward cues were also unpredictable due to trial randomization. Hence, the LHb influence on the RMTg is unidirectional – driving increased responses to stimuli that are “worse” than expected but having no effect on RMTg responses to “better” than expected stimuli. LHb inactivation also did not alter RMTg baseline firing, nor animals’ behavioral performance as measured by the propensity or latency to approach the food receptacle (Figures S5A, B).

### **Triply dissociable influence of LHb, PL, and PBN on RMTg firing**

Because RMTg early phase responses to shocks and cues were not dependent on the LHb (Figure 2), we surmised that they arose from other unknown RMTg afferents. Notably, in our RMTg recordings, RMTg responses to shock and shock-predictive cues were most often seen in distinct subpopulations. Among 152 neurons excited by either shock or shock cue, only a minority ( $36/152 = 23\%$ ) were activated by both stimuli, while the remaining neurons ( $66/152=44\%$  and  $50/152=33\%$ ) were activated by only shock cues or shocks, respectively (Figure S1G). This separation of shock versus shock cue responses led us to further hypothesize that responses to these stimuli arose from distinct afferents, possibly the PL and PBN which had been previously implicated in processing of cues and aversive outcomes, respectively (Campos et al., 2018; Chen et al., 2013; Coizet et al., 2010). Indeed, we found

that PL inactivation (either pharmacological or optogenetic) significantly reduced the RMTg's early (0–30ms) but not late (30–100ms) phase excitation by shock cues (Figures 3A–B, G, H). PL inactivations also had no effect on RMTg excitations to shocks (whether cued or not) nor to reward omission (Figures 3C–D, F, I–J), but they *did* reduce the inhibitory response of RMTg neurons to reward cues (Figure 3E), indicating a major PL contribution to RMTg responses to cues of both positive and negative valence. In contrast to PL inactivation, PBN inactivation (pharmacological or optogenetic) had no effect on RMTg responses to cues (of either positive or negative valence), but instead reduced the early (0–30ms) excitation to both uncued and cued shocks, without affecting late phase (30–100ms) excitation to uncued shocks (Figures 3K–N, Q–K). Responses to reward omissions were also unaffected by PBN inactivation (Figures 3O, P). Inactivations of the PL, PBN, and LHb afferents did not alter RMTg baseline firing rates, nor animals' behavioral performance as measured by reward receptacle approach (Figures S5A–B).

Taken together, our electrophysiological recordings show that the PL, PBN, and LHb influences on RMTg firing are *triple* dissociable, independently driving RMTg responses to three distinct negative motivational stimuli: shock-predictive cues, shocks, and prediction error signals, respectively. Notably, inactivations of the LHb, PL or PBN produced equal response impairments in valence- and salience-encoding neurons, as well as neurons non-responsive to reward cues (Figure S5C), suggesting that these afferents broadly target multiple subpopulations of RMTg neurons.

### **PL→RMTg, PBN→RMTg, and LHb→RMTg modulate triply dissociable phases of punished reward-seeking**

We next tested whether the strikingly separate PL, PBN, and LHb contributions to RMTg firing were reflected in correspondingly distinct behavioral influences. We employed a temporally-specific and pathway-specific optogenetic approach to inhibit PL→RMTg, PBN→RMTg, or LHb→RMTg projections during specific task phases in a previously established punishment learning paradigm (Vento et al., 2017) (Figures 4A–D). In this task, rats were trained to lever press for food, after which brief footshocks were introduced immediately following delivery of food pellets. Shock intensities were gradually escalated across trials within each test session until a “breakpoint” was reached, at which point subjects withheld responding (Figure 4B). Using this behavioral paradigm, we had previously shown increased reward-seeking when the RMTg projection to the VTA was optogenetically inhibited either: (1) at the beginning of trials when subjects are deciding whether to lever press for reward in spite of impending punishment (“decision phase”), or (2) at the precise moment when shock is delivered and presumably serves to guide future behavior (“shock phase”) (Figure 4E). We found that these two temporally specific roles for the RMTg were mirrored in the role of the PL and PBN in RMTg encoding of cues and shock, respectively. In particular, we found that optogenetic inhibition of the PL→RMTg projection increased shock breakpoint if light was delivered during the decision (cue) phase, but not the shock phase, while PBN→RMTg inhibition increased shock breakpoint when optical inhibition overlapped footshock presentation, but not the decision phase (Figures 4F, S6).

Interestingly, inhibition of the LHB→RMTg projection did not affect shock breakpoint during either the shock phase or the decision phase (Figure 4F). We hypothesized this was because the delivery of footshocks followed a predictable schedule. Hence, we next assessed the LHB→RMTg contribution when punishment was delivered at 50% probability, and hence unpredictable on *every* trial. Consistent with our hypothesis, inhibition of the LHB→RMTg pathway in this probabilistic shock task increased shock breakpoint, with inhibition being effective during both the decision and shock phases (Figure 4G).

### **RMTg drives VTA encoding of punishment-related, but not reward-related, stimuli**

We had previously shown that RMTg roles in punished reward seeking were mediated by its projections to the VTA (Vento et al., 2017), but the exact RMTg influence on VTA firing had not been broadly tested. Prior studies had posited a major influence of the RMTg on dopamine neurons, since a majority (67%) of RMTg neurons project to either the VTA or substantia nigra pars compacta (SNc), where they synapse overwhelmingly onto DA (versus non-DA) neurons and inhibit 94–95% of presumed midbrain DA neurons (Balcita-Pedicino et al., 2011; Bourdy et al., 2014; Hong et al., 2011; Zhou et al., 2009a; Li et al., 2019a; Smith et al., 2018; Vento et al., 2017). Most recently, we also found that VTA neuron inhibitions by punishers (e.g. footshocks) depend on the RMTg (Li et al., 2019a), but that study did not examine VTA responses to reward omissions, nor to cues *predicting* punishment, which prior studies had suggested could arise from numerous possible structures, including the RMTg but also the ventral pallidum, hypothalamus, and striatum (Tian et al., 2016). Finally, observations that some RMTg neurons are activated by reward cues raise the possibility that the RMTg might have mixed excitatory-inhibitory effects on VTA encoding patterns. To address these questions, we recorded VTA neurons in rats with or without selective virally-mediated apoptotic (taCasp3-driven) ablation of VTA-projecting RMTg neurons (Figures 5A, B, S7A, B) in rats trained to associate distinct auditory cues with sucrose reward, aversive footshock, or no outcome, respectively (Figure 5C). We found that in sham-lesioned rats, 32% of VTA neurons showed phasic activations restricted to a 200ms window after reward-predictive cues, while 36% of neurons showed sustained activations to reward cues extending 200–2000ms after stimulus onsets (Figure 5D). We denoted these neurons as type I and type II, respectively, analogous to previous work in mice demonstrating these response patterns in optogenetically identified DA or GABA neurons, respectively (Cohen et al., 2012; Eshel et al., 2015; Matsumoto et al., 2016; Mohebi et al., 2019). The remaining 13 neurons (28% of total) were by definition not activated by reward cues. Further analysis showed that these “unclassified” neurons were largely non-responsive to all stimuli, aside from a single neuron inhibited by shock-predictive cues, and two neurons inhibited by shocks.

Type I VTA neurons exhibited reward prediction error (RPE) encoding patterns. In particular, type I neurons were activated by reward cues in proportion to reward probability (60%, 90%, or 100%), and by rewards themselves in inverse proportion to reward probability, with the latter responses disappearing entirely when rewards were fully predicted (Figures 5E–F), as reported for DA and presumptive midbrain DA neurons in rats, mice, and non-human primates (Cohen et al., 2012; Mohebi et al., 2019; Schultz et al., 1997). Furthermore, type I neurons were predominantly inhibited (sometimes after a

transient <50ms excitation) by all three negative motivational stimuli tested: footshocks, shock cues, and reward omissions (Figures 5G–H) (7/15, 10/15, and 9/15 neurons, respectively), again consistent with previous reports in multiple species (Cohen et al., 2012; Matsumoto et al., 2016; Matsumoto and Hikosaka, 2009a). Type I neurons almost never showed activations by these three negative motivational stimuli (beyond the possible transient excitation), with the exception of a single neuron activated by shock cues.

In contrast to type I neurons, type II VTA neurons were mostly non-responsive to shock cues or shocks (13/19 and 10/19) (Figure S7D), or showed large transient (<50ms) excitations to shocks that immediately returned to baseline (6/19 neurons), a marked contrast to the mostly slower inhibitory responses of type I neurons.

Ablation of VTA-projecting RMTg neurons did not change the percentage of VTA neurons classified as type I or type II (Type I: 14/47 and 17/51 for sham and lesion groups,  $p = 0.79$ ; type II: 19/47 and 18/51 for sham and lesion groups,  $p=0.89$ , Chi square), nor did it alter magnitudes of responses to reward cues or unexpected rewards in either type I or type II neurons (Figures 5D–F, S7E). However, ablation of VTA-projecting RMTg neurons dramatically decreased the proportions of type I VTA neurons inhibited by footshocks, shock cues, and reward omissions (2/17, 2/17, and 1/17 neurons,  $p=0.02$ ,  $p=0.001$ , and  $p=0.0009$ , respectively, Chi-square versus sham condition) (Figures 5G, H), consistent with loss of a major inhibitory influence on these neurons. Furthermore, after ablation, some type I neurons (7/17) were significantly *excited* by shocks during a window 100–400ms post-stimulus, a pattern never seen in shams ( $p = 0.02$ , Chi square), again consistent with a loss of an inhibitory influence.

After RMTg ablation, type II neurons remained largely unresponsive to shock cues or shocks (13/18 and 8/18) ( $p=0.92$  and  $0.77$  versus sham, Chi square) (Figures S7C, D), and ablations also did not affect the proportion of neurons activated by footshock (10/18,  $p=0.35$  vs sham, Chi square). Unclassified neurons (neither type I nor II) remained largely unresponsive to stimuli after RMTg ablations (data not shown).

Notably, VTA neurons showing phasic inhibitions to reward cues were non-existent in shams and rare (1/51 neurons) in lesion groups, suggesting that reward cue-activated RMTg neurons (saliency-like) were not primary drivers of VTA encoding patterns. Taken together, the above data show that RMTg neurons have a highly selective effect on negative valence encoding in the VTA.

### **RMTg neurons broadcast punishment signals to multiple targets**

Our findings that the RMTg drives negative valence encoding in the VTA do not preclude RMTg influences on other targets. Indeed, prior studies had shown that largely separate RMTg subpopulations project to the VTA and SNC (Smith et al., 2019). Therefore, we examined shock- or shock cue-induced expression of the immediate-early gene cFos in rats with injections of the retrograde tracer cholera toxin B subunit (CTb) into either the VTA or SNC (Figures 6A–C). We found evidence that c-Fos activated neurons project to both targets, with 42% or 45% of shock- or shock cue-activated RMTg neurons expressing retrograde tracer after VTA injections (Figures 6D–E), while 21% and 23% of shock- or

shock cue-activated RMTg neurons expressed retrograde tracer after SNC injections (Figure 6E). Hence, although midbrain DA neurons are targeted overall by a majority of RMTg neurons, a substantial minority of RMTg shock- and shock cue-activated neurons likely project elsewhere.

## Discussion

The current study examined 5 interconnected brain regions that had all been previously implicated in punishment learning, albeit in separate studies using different testing paradigms that precluded direct comparisons across studies. Hence, we used a common set of physiological and behavioral tests to show that these regions play roles that are distinct but also highly coordinated, with dissociable punishment-related signals arising from the PL, PBN, and LHb and converging onto the RMTg to drive correspondingly distinct aspects of punishment learning and behavior (Figure 7). Furthermore, RMTg efferents in turn influence encoding of negative but not positive motivational stimuli in VTA neurons, while also likely conveying punishment signals to regions outside the VTA.

### Segregation of punishment signals into distinct brain pathways

One striking finding of this study is that distinct brain pathways carry highly segregated yet complementary neural signals about punishment, contributing to correspondingly distinct yet coordinated roles in behavior. For example, RMTg ablation severely reduced VTA responses to several types of punishment-related, but not reward-related stimuli, indicating a sharp separation between pathways carrying reward versus punishment information, consistent with earlier studies showing that inactivation of the RMTg→VTA pathway impairs punishment learning independently of reward learning (Vento et al., 2017).

In addition to segregation of punishment signals from reward signals, we also observed further segregation of different types of punishment signals from each other. In particular, we found that PL, PBN, and LHb projections to the RMTg drive triply dissociable components of responses to cues, shocks, and prediction errors, respectively. This separation of encoding patterns into distinct afferents was strikingly commensurate with their separate behavioral contributions, as the PL and PBN afferents contributed to punishment learning during the “decision” and “outcome” phases of our punished reward-seeking task, in contrast to the RMTg as a whole, whose projection to the VTA was critically involved in both phases (Vento et al., 2017).

The separation of punishment signals in distinct RMTg afferents contrasts with previous reports about DA neurons, whose activations by rewards and reward cues had been shown to arise from a confluence of inputs encoding mixed information about cues and rewards, rather than the segregated information we saw among RMTg inputs (Tian et al., 2016). The current findings also showed other asymmetries between reward and punishment processing. In particular, prediction errors were just one component of the RMTg response to punishment-related stimuli, in contrast to VTA neurons showing much purer prediction error responses to reward-related stimuli. For example, midbrain DA neuron responses to rewards largely disappear if these rewards are fully predicted (Figures 5E–F)(Montague et al., 1996; Schultz et al., 1997; Schultz and Dickinson, 2000), while RMTg neurons still show



responses to aversive stimuli (albeit reduced) even if they are fully predicted, a property also shared with LHB and VTA neuron responses to aversive stimuli (Matsumoto and Hikosaka, 2009a; Wang et al., 2017). Hence, even though reward and aversion have been posited to be mediated by a common neural “currency”, our data and others show several asymmetries between processing of rewards and punishers. The reasons for these differences are not known, but some possible hypotheses have been raised by behavioral economic studies that have long noted large asymmetries in human subjective reactions to economic gains versus losses (Kahneman and Tversky, 1979; Tom et al., 2007; Zhang et al., 2014).

One particular example of signal segregation was our initially surprising finding of a highly selective role for the LHB in augmenting RMTg responses to stimuli that were “worse” than expected. Many prior studies had posited a much broader role for the LHB, due to its activation by almost all stimuli known to activate the RMTg (Hong et al., 2011; Jhou et al., 2009a; Jhou et al., 2013). However, previous studies had also shown that individual LHB neurons exhibit heterogeneity in both projection targets and response patterns (Congiu et al., 2019; Matsumoto and Hikosaka, 2009a), raising the possibility of the LHB exhibiting selective influences on specific targets. The selective LHB effect in the current results also parallel previous findings that the LHB drives DA neuron inhibitions to reward omissions, but not airpuffs or its predictors (Tian and Uchida, 2015). Together, we posit that these results may help explain why prior studies have strongly implicated the LHB in some motivational tasks but not other seemingly similar tasks. In particular, several studies have noted that the LHB is more strongly involved during early acquisition than later maintenance of a punishment task (Trusel et al., 2019a), and is particularly involved in behavioral tasks requiring flexible adaptation to changing contingencies, all of which would be expected to generate high levels of prediction errors (Baker et al., 2017; Bromberg-Martin and Hikosaka, 2011; Heldt and Ressler, 2006; Montague et al., 1996; Stopper and Floresco, 2014; Stopper et al., 2014; Trusel et al., 2019a).

### **Reward- and punishment-related signals both encode retrospective and prospective information**

Despite the various asymmetries between midbrain encoding of punishment- versus reward-related stimuli, we also found evidence for parallel similarities between these processes, as both RMTg and DA neurons may play dual roles in learning. In particular, DA neurons have long been posited to provide both *retrospective* and *prospective* information about reward-seeking, with DA activation by rewards providing a “teaching” signal that reinforces actions completed in the recent *past*, while DA activation by reward-predictive cues energizes and potentiates pursuit of rewards to be obtained in the near *future* (Hamid et al., 2016; Kim et al., 2012; Phillips et al., 2003). Our current and previous findings suggest a similar but inverse role for the RMTg, whose activation by shocks provides an *aversive teaching signal* that punishes actions occurring in the recent past, while its responses to cues may inhibit the ongoing seeking of rewards to be obtained in the immediate future (Vento et al., 2017).

We further found evidence that the prospective (cue-related) functions of the RMTg are related to its PL afferents, as inhibiting this afferent to the RMTg during the “decision” phase increased lever pressing in the punished reward-seeking task. We posit that the

increased pressing reflected reduced signaling of upcoming punishment. However, we also found a PL contribution to RMTg responses to reward-predictive stimuli, raising the possibility that its inactivation impaired signaling of upcoming rewards instead. We believe the latter influence would be weaker, however, as impaired reward signaling would be expected to reduce lever pressing, the opposite of the observed effect.

### **Diversity of RMTg encoding patterns and output projections**

The RMTg appears relatively homogeneous neurochemically, with >90% of neurons expressing markers for GABA, nociceptin (also called orphanin FQ), and the transcription factor FOXP1 (Smith et al., 2019). However, the current study adds to a growing body of evidence that RMTg neurons exhibit heterogeneous encoding patterns and projection targets (Lavezzi et al., 2012; Li et al., 2019a; Segó et al., 2014). For example, we had previously shown that RMTg projections to the VTA and SNC in rats arise from topographically distinct subpopulations (Smith et al., 2019), while RMTg projections to the VTA versus dorsal raphe nucleus (DRN) also arise from distinct subpopulations that respectively show predominantly inhibitory versus excitatory responses to reward-predictive cues (Li et al., 2019a). The latter study notably shows a correlation between neural encoding patterns and projection targets, with most VTA-projecting neurons encoding valence, while projections to the DRN more often encoded salience-like patterns. The current study further confirmed the presence of a predominant valence-like encoding pattern in the RMTg accompanied by a smaller but substantial salience-encoding subpopulation. Interestingly, prior recordings of putative SNC DA neurons (Matsumoto and Hikosaka, 2009b) also found subsets of neurons encoding valence (opposite activation/inhibition to reward/aversion) and salience (activation by both rewards and punishers). The functional significance of these different encoding patterns is unknown, but influential theoretical work has suggested that valence and salience signals should play distinct, but complementary roles in learning, with valence signals indicating whether an outcome is “good” versus “bad”, and salience signals indicating whether a stimulus is motivationally relevant (regardless of valence) (Kahnt et al., 2014; Pearce and Hall, 1980). While our current and prior results strongly suggest that RMTg projections to the VTA drive valence encoding, the functions of other projections are not known, but could possibly convey salience signals. Notably, PL, PBN, and LHb projections to the RMTg seem to have equal influences on valence- and salience-encoding RMTg neurons, suggesting that they could influence multiple distinct subsets of RMTg neurons projecting to different targets and having different influences on behavior.

### **Methodological considerations: DA and GABA neurons in VTA versus RMTg**

In some of our recordings of VTA neurons, we ablated VTA-projecting cells in the RMTg using a dual virus approach. However, the adjacency of the RMTg to the VTA raises the possibility that we could have also ablated some neurons within the VTA itself, including possibly GABAergic neurons that may also influence reward prediction error encoding in the VTA (Cohen et al., 2012). However, our injections of AAV-FLEX-taCasp3 into the RMTg were placed 1.9mm caudal to the retroAAV-Cre injections into the VTA, and our cell counts of FOXP1 neurons confirmed that cell loss was most prominent at middle levels of the RMTg, with little or no loss of neurons at the rostral RMTg closest to the VTA. Because the AAV-FLEX-taCasp3 did not express its own fluorescent marker, its spread could not be

visualized directly, and other evidence of viral infection, e.g. the presence of apoptosis markers, would have likely disappeared by the time of sacrifice, several weeks after initial injections. However, pilot experiments in our lab involving the same volume and serotype of virus expressing GFP into the RMTg produced spread mostly within a  $\pm 1.2$ mm radius around the injection site, with minimal spread into the VTA.

In our recordings of VTA neurons, we did not attempt to classify VTA neurons as DA or non-DA, as almost all methods of electrophysiological classification have been challenged, including classification based on waveform shape, firing rate, and even genetic identity (Cohen et al., 2012; Grace and Bunney, 1983; Lammel et al., 2015; Poulin et al., 2018; Stuber et al., 2015; Ungless and Grace, 2012). However, the response patterns we saw in our type I or II neurons are strikingly similar to those previously shown in optogenetically tagged mouse DA and GABA neurons, respectively (Cohen et al., 2012; Eshel et al., 2015). In particular, we found that type I neurons predominantly exhibited classic reward prediction error (RPE) responses to rewards and reward cues across a range of reward probabilities, as widely reported in multiple species (Cohen et al., 2012; Mohebi et al., 2019; Schultz et al., 1997). Furthermore, light and electron microscopic studies have shown that RMTg neurons overwhelmingly target DA versus non-DA neurons in the VTA and SNC (Balcita-Pedicino et al., 2011; Bourdy et al., 2014; Zhou et al., 2009a), and the current study found that RMTg ablation selectively impaired responses in type I but not type II or unclassified VTA neurons. These results suggest a strong possibility that a large proportion of type I neurons are indeed DAergic. However, because not all DA neurons encode reward, and not all reward-activated neurons are DAergic (Brischoux et al., 2009; Tian and Uchida, 2015), the exact neurochemical identity of a VTA neuron appears to be neither a necessary nor sufficient indicator of the information it encodes. Hence, the current study analyzed VTA recordings with respect to encoding patterns rather than DAergic identity per se.

Finally, we used the proton pump archaerhodopsin to inhibit afferent terminals in the RMTg, as this opsin has been shown to effectively inhibit presynaptic release in response to brief pulses of light (200ms) (Mahn et al., 2016). However, it may produce paradoxical activation after prolonged (several minutes) light delivery, likely due to pH-dependent calcium influx (Mahn et al., 2016; Wiegert et al., 2017). However, the majority of the experiments described here used very brief (<1 sec) duration of light delivery and even the longest light duration used here (30 sec) was far less than the durations previously shown to increase calcium levels.

### Future directions

The current study shows that punishment learning relies on coordinated interactions of at least five brain regions, the VTA, RMTg, PL, PBN, and LHb. However, numerous other regions have also been implicated in punishment learning, including the pallidum, ventral pallidum, amygdala, paraventricular thalamus, lateral hypothalamus, and nucleus accumbens (Beyeler et al., 2016; Faget et al., 2018; Lammel et al., 2012; Lecca et al., 2017; Meye et al., 2016; Nieh et al., 2016; Penzo et al., 2015; Piantadosi et al., 2017; Stephenson-Jones et al., 2016; Trusel et al., 2019b). The relative contributions of each of these regions, and their interactions with each other, are not well understood, although it is notable that almost all of

them are interconnected with one or more of the structures examined here. Hence, the current study may provide a basis for future work that will further integrate our understanding of these phenomena.

## STAR Methods

### LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Thomas Jhou (jhou@musc.edu). This study did not generate new unique reagents.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Animals**—All procedures were conducted under the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and all protocols were approved by Medical University of South Carolina Institutional Animal Care and Use Committee. Adult male Sprague Dawley rats weighing 200–300 g from Charles River Laboratories were paired housed in standard shoebox cages with food and water provided *ad libitum*, unless otherwise stated. We used 87 rats in total: 21 rats for RMTg recordings with pharmacological inactivation (8 for LHb, 7 for PL, and 6 for PBN inactivation), 10 for RMTg recording with optogenetic inactivation (3 for LHb, 3 for PL and 4 for PBN), 5 for three-cued discrimination task, 11 for VTA recordings (5 lesioned and 6 sham-lesioned controls), 22 for optogenetic inhibition during punished food-seeking (6 for LHb→RMTg, 5 for PBN→RMTg, 6 for PL→RMTg, and 5 for no virus control). 18 for cFos induction with CTb injected in either VTA or SNC (11 for VTA and 7 for SNC). In VTA group, 7 tested for shocks and 4 tested for shock cues. In SNC group, 3 tested for shocks and 4 tested for shock cues.

### METHOD DETAILS

**Surgeries**—All surgeries were conducted under aseptic conditions using inhaled isoflurane anesthesia (1.5%-3% at 0.5–1.0 liter/min). Analgesic (ketoprofen, 5mg/kg) was administered subcutaneously immediately after surgery. Animals were given at least 5 days to recover from surgery, and at least 3 weeks for viruses to express.

**Pavlovian behavioral training**—Rats were food restricted to 85% of their *ad libitum* body weight and trained to associate distinct auditory cues (70dB 1kHz or white noise, 2 seconds duration, counterbalanced across animals) with either a food pellet (45mg, BioServ) or no outcome (Figure 1A). Training was conducted in Med Associates chambers (St. Albans, VT). Trial types were randomized with a 30s interval between successive trials. A “correct” response consisted of entering the food tray within 2 s after reward cue onset or withholding response for 2 s after neutral cue onsets. Rats were trained with 100 trials per session, one session per day, until they achieved 85% accuracy in any 20-trial block, after which animals were then trained with one extra session in which neutral cue trials were replaced by shock trials consisting of a 2 second 8kHz tone (75dB) followed by a 10ms 0.7mA footshock.

**In vivo electrophysiological recordings**—RMTg recording sessions used the same Pavlovian configuration as the final training session, with the exception that reward delivery was omitted in a randomly selected 20% of reward trials, and an additional trial type was added consisting of uncued shocks (also 10ms 0.7mA) presented with no prior cue (Figure 1C). Each of the three trial types (reward, cued shock, and uncued shock) represented a random 33% of all trial types. We found that frequency of the tones did not affect RMTg responses to reward nor shock cues (Figure S1D, E).

In the three-cued discrimination task, reward trial, shock trial, and neutral trial which predicts no outcome were randomly selected at 33% probability (Figure S1H).

In VTA recording experiments, the probability of reward delivery in reward trials varied between sessions (60%, 90%, or 100%), but was fixed for any particular session (Figure 5C).

Drivable electrode arrays were implanted above the RMTg (AP: -7.4mm; ML: 2.1mm; DV: -7.4mm from dura, 10-degree angle) or VTA (AP: -5.5mm; ML: 2.5mm; DV: -7.8mm from dura, 10-degree angle towards midline). For inactivation experiments, guide cannulas (Plastics One Inc.) were implanted into the LHb (AP: -3.6mm; ML: 1.6mm; DV: -4.1mm from dura, 10-degree angle), PL (AP: 3mm; ML: 0.6mm; DV: 3mm from dura), or PBN (AP: -9.5mm; ML: 3mm; DV: -5.7mm. All guides were bilateral except for LHb inactivation experiments in which we implanted unilateral guides, and then electrolytically lesioned the contralateral fasciculus retroflexus (FR), the main axon bundle carrying LHb axons to the RMTg, using 2mA current delivered for 30s (AP: -4mm; ML:  $\pm$ 2mm; DV: -5.3mm from dura, 10-degree angle).

In some VTA recording experiments, we conducted pathway-specific lesions of VTA-projecting RMTg neurons using 200nl rAAV2-Retro-CAG-Cre (UNC, AV7703C) injected unilaterally into the VTA (DV: -8.2mm from dura) with the other side receiving saline. Also, rAAV5-Flex-taCasp3-TEVp (UNC) was injected bilaterally (200nl per side) into the RMTg (DV: -7.8mm from dura). This approach induces rapid Cre-dependent cell death specifically in VTA-projecting RMTg neurons (Yang et al., 2013). Electrodes were placed in the same region of the VTA where retroAAV-Cre was injected (Figure S7A).

Electrodes consisted of a bundle of sixteen 18  $\mu$ m Formvar-insulated nichrome wires (A-M Systems, Sequim, WA) attached to a customized printed circuit board (Seeed Studio, Shenzhen, China) onto which a 3D-printed microdrive was also mounted (Shapeways, Inc, New York, NY). Electrodes were grounded through a 37-gauge wire attached to a gold-plated pin (Mill-max, part # 1001-0-15-15-30-27-04-0) implanted into cortex. Recording electrodes were advanced 80–160  $\mu$ m at the end of each session. Recording signals were amplified by a unity gain headstage followed by further amplification at 250x gain with bandpass filtering of 300 Hz–6 kHz (Neurosys LLC). Analog signals were then digitized (18 bits, 15.625 kHz sampling rate) by a PCI card (National Instruments) controlled by customized acquisition software (Neurosys LLC). Spikes were initially detected via thresholding to remove signals less than twofold above background noise levels, and spikes were further discriminated and sorted using principal component analysis. Detected spikes on each recording wire were accepted only if a refractory period was seen, determined by

<0.2% of spikes occurring within 1ms of a previous spike, as well as by the presence of a large central notch in the auto-correlogram. Neurons that had large drifts in firing rates within a session were also excluded. Since the shock duration used in the present study was 10ms, the first 10ms of data after footshock were removed from analysis due to interference from shock artifacts.

In some RMTg recording sessions, shock trials were signaled by a double cue, in which a 0.5s duration 4kHz tone preceded a 2s 8kHz tone, and 10ms 0.7mA footshock was administered immediately thereafter (Figure 2N).

Inhibition of either the PL or PBN was accomplished via bilateral microinjection of a GABA<sub>A</sub>/GABA<sub>B</sub> agonist cocktail (0.05nmol muscimol and 0.5nmol baclofen) during the second half of each session. LHb inactivation used only a unilateral infusion, as we had lesioned the contralateral FR in these animals.

In RMTg recordings with optogenetic inhibitions, green laser light (532nm, 10mW, Dragon Lasers; Changchun, China) was delivered through an optical cable (50µm core, Precision Fiber Products, Milpitas, CA) mated to an implanted 2-mm-diameter stainless steel ferrule (Precision Fiber Products, Milpitas, CA) containing an optical fiber (200-µm core; Thorlabs, Newton, NJ) with its tip located at the dorsal edge of the RMTg. Laser light was delivered continuously beginning 100ms before the onset of the aversive cue or shock and persisting for 100ms after stimulus offset. Laser light was delivered during cued shocks, uncued shocks, or shock cues, on 33% of trials selected randomly within each test session.

**Punishment resistance experiments**—For LHb cannulation, 6mm length guide cannulae (Plastics One Inc.) were implanted bilaterally into the LHb (AP: -3.6mm; ML: 1.6mm; DV: -4.1mm from dura, 10-degree angle). For FR lesions, 2mA current was delivered bilaterally to the FR using a glass electrode (FHC Inc.) for 30s per side (AP: -4mm; ML: ±2mm; DV: -5.3mm from dura, 10-degree angle). For pathway-specific inhibition of PBN→RMTg or PL→RMTg, rAAV2-hSyn-eArch3.0-eYFP (UNC, AV5229C) or pAAV2-hSyn-HA-hM4D(Gi)-mCherry (AddGene) (500nL per side) was bilaterally injected into either the PBN or PL. Optical fibers (ThorLabs) or infusion cannulae (Plastic One Inc.) were implanted bilaterally into the RMTg.

Methods are similar to our previous publication (Vento et al., 2017). In brief, food-restricted rats (85% +/- 3% of *ad libitum* body weight) were trained to lever press for food rewards (45mg pellets, BioServ, San Diego, CA) on a fixed ratio 5 (FR5) schedule with a 15 second intertrial interval (ITI) in standard operant chambers (Med Associates). Cue lights were illuminated above both the active and inactive levers which turned off during the ITI, and responses on the inactive lever had no programmed consequences. Each test session began with 5 “no cost” trials where responding on the active lever was not punished. Beginning on the 6<sup>th</sup> trial, FR completion resulted in delivery of the food pellet and brief (33ms) footshock. The intensity of footshock was low in the first punished trials (0.25mA), but subsequently increased by approximately 30% after completion of every 3 trials thereafter. Failure to respond on the active lever for 30sec resulted in an additional 30 sec time-out period in which the active and inactive levers remained extended to record behavioral

responses, but cue lights turned off and lever presses yielded no consequence. After this timeout, cue lights were once again illuminated and a new trial initiated. After three consecutive timeouts, the session was terminated and the last shock intensity was recorded as the “shock breakpoint”. Data were averaged across multiple test sessions to reduce day-to-day variability (Vento et al., 2017).

In animals receiving temporally-specific inhibition of PL, PBN, or LHb projections to the RMTg, we injected inhibitory light-sensitive archaerhodopsin (rAAV2-hSyn-eArch3.0-eYFP) into the PL, PBN, or LHb two weeks before behavioral testing, and then (in the same surgery) implanted bilateral stainless steel ferrules containing indwelling optical targeted to the dorsal edge of the RMTg. During testing, green light (532 nm; Dragon Lasers; Changchun, China) was delivered at 15–30mW without pulsing, using optical splitters (Precision Fiber Products) mated to the chronically implanted stainless-steel ferrules. Laser light was delivered at two distinct time points in the food-seeking task: the “decision phase” (the beginning of each trial when levers are extended, and cue lights are illuminated indicating reward availability), or the “shock phase” (the brief time period (~200ms) overlapping footshock). We used 15mW during the decision phase, and 30mW during the (much shorter) shock phase. As a control for delivery of light overlapping the footshock (*synchronized* condition), control sessions were performed in which light was instead delivered immediately before or after footshock (*desynchronized* condition). A full discussion of this design has been presented in our prior publication (Vento et al., 2017), but in brief, this design controls for any possible shock-independent effects of light on behavior. For example, it controls for the possibility that RMTg inhibition is reinforcing in its own right, as all trials have the same amount of laser light, and the same amount of shock, with the same average delay from operant action (lever press) to food/light/shock outcomes, and the only difference being whether the laser overlaps the exact time of shock delivery or not.

**cFos induction**—For aversion-induced cFos, animals were habituated in the behavioral chamber 20 minutes at the first day (context A). On the second day, animals were trained with 15 pairs of a 2s 8kHz tone followed by a 5s 0.5mA footshock with ITI of 60s in a chamber with distinct beddings and an ethanol odor (context B). Then animals were again habituated in context A for a day. On the fourth day, animals received either 15 presentations of shock cues or 15 pairs of shock cues and shocks. For animals tested with shock-alone, the training session was omitted and they received 15 footshocks lasting 5–15 seconds (0.5mA).

**Perfusions and tissue sectioning**—After completion of experimental procedures, rats were anesthetized using inhaled isoflurane and transcardially perfused with 10% formalin in 0.1M phosphate buffered saline (PBS), pH 7.4. For electrophysiology experiments, immediately prior to perfusion rats were placed under isoflurane anesthesia and 100 $\mu$ A current was passed through the electrode to permit visualization of wire tips. Brains were removed from the skull, equilibrated in 20% sucrose, and cut into 40 $\mu$ m sections on a freezing microtome. Sections were stored in PBS with 0.05% sodium azide.

**Immunohistochemistry**—Free-floating sections were immunostained by overnight incubation in primary antibody (rabbit anti-FOXP1, Abcam, ab16645, 1:50,000 dilution; mouse anti-TH, Millipore, MAB-377, 1:50,000 dilution; rabbit anti-Cre, Novagen, 69050–3,

1:5,000 dilution; rabbit anti-cFos, Millipore, ABE457, 1:5,000 dilution; goat anti-CTb, List Biological Laboratories, 1:50,000 dilution), in PBS with 0.25% Triton-X and 0.05% sodium azide. Tissue was then washed three times in PBS (30s each) and incubated in biotinylated secondary (1:1000 dilution, Jackson ImmunoResearch, West Grove, PA) for 30 min, followed by three 30s rinses in PBS, and 1-hour incubation in avidin-biotin complex (Vector Laboratories, Burlingame, CA). Tissue was then rinsed in sodium acetate buffer (0.1M, pH 7.4), followed by incubation for 5 min in 0.05% diaminobenzidine (DAB, Vector Laboratories) with 0.1% nickel and 0.01% hydrogen peroxide, revealing a blue-black reaction product. For florescent staining, Cy3-conjugated anti-rabbit and 488-conjugated anti-goat (1:1000 dilution, Jackson ImmunoResearch, West Grove, PA) were used. Counterstaining was achieved using deep-red Nissl staining (Fisher, N21483).

**Quantification and statistical analysis**—Electrophysiological firing rates were analyzed using non-parametric paired tests, and were determined to have significant responses to stimuli if post-stimulus firing (averaged over time window of interest) was significantly different from the one-second average baseline firing calculated prior to each individual stimulus in a given session ( $p < 0.05$ , Wilcoxon signed-rank). RMTg responses to shock-predictive cues were examined in a 0–100ms time window after cue onsets, and 10–100ms after shock onsets (to avoid tabulating shock artifacts occurring in the first 10ms). Prior studies had shown that RMTg responses to reward-predictive cues are slower than responses to shock-related stimuli (Hong et al., 2011; Zhou et al., 2009a), and hence these responses were examined 100–500ms after cue onset.

In all graphs, normalized firing rates are calculated by dividing post-stimulus firing rates by the average firing during a one-second window prior to that stimulus (either cue or outcome). This ratio is calculated for each trial independently, then all trials are averaged for a given session.

For both behavioral and electrophysiological experiments, repeated measures two-way ANOVA with Holm-Sidak correction for multiple comparisons, e.g. across multiple time windows, and two-tailed paired t-test were used to compare across experimental conditions, respectively, if not otherwise specified. Differences in shock breakpoint were assessed using two-way ANOVA with Sidak's correction for multiple comparisons, or paired t-tests. \* $p < 0.05$ , \*\* $p < 0.01$  and \*\*\* $p < 0.001$ . Calculations were performed using MATLAB and Prism 7 software (Graph Pad).

## DATA AND CODE AVAILABILITY

The datasets and code have not been deposited in a public repository due to their volume and complexity, but are available from the corresponding author upon request.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## Acknowledgements:

This work was supported by National Institutes of Health grants R21 DA037327, R01 DA037327, P50 DA015369 (all to TCJ), and F32 DA040379 (to PJV).

## References

- American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders, 5th edn (Washington, DC).
- Baker PM, Raynor SA, Francis NT, and Mizumori SJ (2017). Lateral habenula integration of proactive and retroactive information mediates behavioral flexibility. *Neuroscience* 345, 89–98. [PubMed: 26876779]
- Balcita-Pedicino JJ, Omelchenko N, Bell R, and Sesack SR (2011). The inhibitory influence of the lateral habenula on midbrain dopamine cells: ultrastructural evidence for indirect mediation via the rostromedial mesopontine tegmental nucleus. *J Comp Neurol* 519, 1143–1164. [PubMed: 21344406]
- Beyeler A, Namburi P, Glober GF, Simonnet C, Calhoon GG, Conyers GF, Luck R, Wildes CP, and Tye KM (2016). Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. *Neuron* 90, 348–361. [PubMed: 27041499]
- Bourdy R, Sanchez-Catalan MJ, Kaufling J, Balcita-Pedicino JJ, Freund-Mercier MJ, Veinante P, Sesack SR, Georges F, and Barrot M (2014). Control of the nigrostriatal dopamine neuron activity and motor function by the tail of the ventral tegmental area. *Neuropsychopharmacology* 39, 2788–2798. [PubMed: 24896615]
- Brischoux F, Chakraborty S, Brierley DI, and Ungless MA (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci U S A* 106, 4894–4899. [PubMed: 19261850]
- Bromberg-Martin ES, and Hikosaka O (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nat Neurosci* 14, 1209–1216. [PubMed: 21857659]
- Campos CA, Bowen AJ, Roman CW, and Palmiter RD (2018). Encoding of danger by parabrachial CGRP neurons. *Nature* 555, 617–622. [PubMed: 29562230]
- Chang CY, Esber GR, Marrero-Garcia Y, Yau HJ, Bonci A, and Schoenbaum G (2016). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nat Neurosci* 19, 111–116. [PubMed: 26642092]
- Chang CY, Gardner MPH, Conroy JC, Whitaker LR, and Schoenbaum G (2018). Brief, But Not Prolonged, Pauses in the Firing of Midbrain Dopamine Neurons Are Sufficient to Produce a Conditioned Inhibitor. *J Neurosci* 38, 8822–8830. [PubMed: 30181136]
- Chen BT, Yau HJ, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW, and Bonci A (2013). Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature* 496, 359–362. [PubMed: 23552889]
- Cohen JY, Haesler S, Vong L, Lowell BB, and Uchida N (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88. [PubMed: 22258508]
- Coizet V, Dommert EJ, Klop EM, Redgrave P, and Overton PG (2010). The parabrachial nucleus is a critical link in the transmission of short latency nociceptive information to midbrain dopaminergic neurons. *Neuroscience* 168, 263–272. [PubMed: 20363297]
- Congiu M, Trusel M, Pistis M, Mameli M, and Lecca S (2019). Opposite responses to aversive stimuli in lateral habenula neurons. *Eur J Neurosci*.
- Danjo T, Yoshimi K, Funabiki K, Yawata S, and Nakanishi S (2014). 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* 111, 6455–6460. [PubMed: 24737889]
- Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, and Uchida N (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* 525, 243–246. [PubMed: 26322583]

- Faget L, Zell V, Souter E, McPherson A, Ressler R, Gutierrez-Reed N, Yoo JH, Dulcis D, and Hnasko TS (2018). Opponent control of behavioral reinforcement by inhibitory and excitatory projections from the ventral pallidum. *Nat Commun* 9, 849. [PubMed: 29487284]
- Grace AA, and Bunney BS (1983). Intracellular and extracellular electrophysiology of nigral dopaminergic neurons--1. Identification and characterization. *Neuroscience* 10, 301–315. [PubMed: 6633863]
- Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, and Berke JD (2016). Mesolimbic dopamine signals the value of work. *Nat Neurosci* 19, 117–126. [PubMed: 26595651]
- Heldt SA, and Ressler KJ (2006). Lesions of the habenula produce stress- and dopamine-dependent alterations in prepulse inhibition and locomotion. *Brain Res* 1073–1074, 229–239.
- Hong S, Zhou TC, Smith M, Saleem KS, and Hikosaka O (2011). Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J Neurosci* 31, 11457–11471. [PubMed: 21832176]
- Hu H (2016). Reward and Aversion. *Annual review of neuroscience* 39, 297–324.
- Jean-Richard-Dit-Bressel P, and McNally GP (2016). Lateral, not medial, prefrontal cortex contributes to punishment and aversive instrumental learning. *Learn Mem* 23, 607–617. [PubMed: 27918280]
- Zhou TC, Fields HL, Baxter MG, Saper CB, and Holland PC (2009a). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* 61, 786–800. [PubMed: 19285474]
- Zhou TC, Geisler S, Marinelli M, Degarmo BA, and Zahm DS (2009b). The mesopontine rostromedial tegmental nucleus: A structure targeted by the lateral habenula that projects to the ventral tegmental area of Tsai and substantia nigra compacta. *J Comp Neurol* 513, 566–596. [PubMed: 19235216]
- Zhou TC, Good CH, Rowley CS, Xu SP, Wang H, Burnham NW, Hoffman AF, Lupica CR, and Ikemoto S (2013). Cocaine drives aversive conditioning via delayed activation of dopamine-responsive habenular and midbrain pathways. *J Neurosci* 33, 7501–7512. [PubMed: 23616555]
- Kahneman D, and Tversky A (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47, 28.
- Kahnt T, Park SQ, Haynes JD, and Tobler PN (2014). Disentangling neural representations of value and salience in the human brain. *Proc Natl Acad Sci U S A* 111, 5000–5005. [PubMed: 24639493]
- Kaufling J, Veinante P, Pawlowski SA, Freund-Mercier MJ, and Barrot M (2009). Afferents to the GABAergic tail of the ventral tegmental area in the rat. *J Comp Neurol* 513, 597–621. [PubMed: 19235223]
- Kim KM, Baratta MV, Yang A, Lee D, Boyden ES, and Fiorillo CD (2012). Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS One* 7, e33612. [PubMed: 22506004]
- Lahti L, Haugas M, Tikker L, Airavaara M, Voutilainen MH, Anttila J, Kumar S, Inkinen C, Salminen M, and Partanen J (2016). Differentiation and molecular heterogeneity of inhibitory and excitatory neurons associated with midbrain dopaminergic nuclei. *Development* 143, 516–529. [PubMed: 26718003]
- Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, and Malenka RC (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature* 491, 212–217. [PubMed: 23064228]
- Lammel S, Steinberg EE, Foldy C, Wall NR, Beier K, Luo L, and Malenka RC (2015). Diversity of transgenic mouse models for selective targeting of midbrain dopamine neurons. *Neuron* 85, 429–438. [PubMed: 25611513]
- Lavezzi HN, Parsley KP, and Zahm DS (2012). Mesopontine rostromedial tegmental nucleus neurons projecting to the dorsal raphe and pedunculopontine tegmental nucleus: psychostimulant-elicited Fos expression and collateralization. *Brain structure & function* 217, 719–734. [PubMed: 22179106]
- Lecca S, Meye FJ, Trusel M, Tchenio A, Harris J, Schwarz MK, Burdakov D, Georges F, and Mameli M (2017). Aversive stimuli drive hypothalamus-to-habenula excitation to promote escape behavior. *Elife* 6.

- Li H, Pullmann D, Cho JY, Eid M, and Zhou TC (2019a). Generality and opponency of rostromedial tegmental (RMTg) roles in valence processing. *Elife* 8.
- Li H, Pullmann D, and Zhou TC (2019b). Valence-encoding in the lateral habenula arises from the entopeduncular region. *Elife* 8.
- Mahn M, Prigge M, Ron S, Levy R, and Yizhar O (2016). Biophysical constraints of optogenetic inhibition at presynaptic terminals. *Nat Neurosci* 19, 554–556. [PubMed: 26950004]
- Matsumoto H, Tian J, Uchida N, and Watabe-Uchida M (2016). Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife* 5.
- Matsumoto M, and Hikosaka O (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111–1115. [PubMed: 17522629]
- Matsumoto M, and Hikosaka O (2009a). Representation of negative motivational value in the primate lateral habenula. *Nat Neurosci* 12, 77–84. [PubMed: 19043410]
- Matsumoto M, and Hikosaka O (2009b). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841. [PubMed: 19448610]
- Meye FJ, Soiza-Reilly M, Smit T, Diana MA, Schwarz MK, and Mameli M (2016). Shifted pallidal co-release of GABA and glutamate in habenula drives cocaine withdrawal and relapse. *Nat Neurosci* 19, 1019–1024. [PubMed: 27348214]
- Mohebi A, Pettibone JR, Hamid AA, Wong JT, Vinson LT, Patriarchi T, Tian L, Kennedy RT, and Berke JD (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*.
- Montague PR, Dayan P, and Sejnowski TJ (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16, 1936–1947. [PubMed: 8774460]
- Nieh EH, Vander Weele CM, Matthews GA, Presbrey KN, Wichmann R, Leppla CA, Izadmehr EM, and Tye KM (2016). Inhibitory Input from the Lateral Hypothalamus to the Ventral Tegmental Area Disinhibits Dopamine Neurons and Promotes Behavioral Activation. *Neuron* 90, 1286–1298. [PubMed: 27238864]
- Pearce JM, and Hall G (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* 87, 532–552. [PubMed: 7443916]
- Penzo MA, Robert V, Tucciarone J, De Bundel D, Wang M, Van Aelst L, Darvas M, Parada LF, Palmiter RD, He M, et al. (2015). The paraventricular thalamus controls a central amygdala fear circuit. *Nature* 519, 455–459. [PubMed: 25600269]
- Phillips PE, Stuber GD, Heien ML, Wightman RM, and Carelli RM (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* 422, 614–618. [PubMed: 12687000]
- Piantadosi PT, Yeates DCM, Wilkins M, and Floresco SB (2017). Contributions of basolateral amygdala and nucleus accumbens subregions to mediating motivational conflict during punished reward-seeking. *Neurobiology of learning and memory* 140, 92–105. [PubMed: 28242266]
- Poulin JF, Caronia G, Hofer C, Cui Q, Helm B, Ramakrishnan C, Chan CS, Dombeck DA, Deisseroth K, and Awatramani R (2018). Mapping projections of molecularly defined dopamine neuron subtypes using intersectional genetic approaches. *Nat Neurosci* 21, 1260–1271. [PubMed: 30104732]
- Schultz W, Dayan P, and Montague PR (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. [PubMed: 9054347]
- Schultz W, and Dickinson A (2000). Neuronal coding of prediction errors. *Annual review of neuroscience* 23, 473–500.
- Sego C, Goncalves L, Lima L, Furigo IC, Donato J Jr., and Metzger M (2014). Lateral habenula and the rostromedial tegmental nucleus innervate neurochemically distinct subdivisions of the dorsal raphe nucleus in the rat. *J Comp Neurol* 522, 1454–1484. [PubMed: 24374795]
- Smith RJ, Vento PJ, Chao YS, Good CH, and Zhou TC (2018). Gene expression and neurochemical characterization of the rostromedial tegmental nucleus (RMTg) in rats and mice. *Brain Struct Funct*.
- Smith RJ, Vento PJ, Chao YS, Good CH, and Zhou TC (2019). Gene expression and neurochemical characterization of the rostromedial tegmental nucleus (RMTg) in rats and mice. *Brain structure & function* 224, 219–238. [PubMed: 30302539]
- Stamatakis AM, and Stuber GD (2012). Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nat Neurosci* 15, 1105–1107. [PubMed: 22729176]

- Stephenson-Jones M, Yu K, Ahrens S, Tucciarone JM, van Huijstee AN, Mejia LA, Penzo MA, Tai LH, Wilbrecht L, and Li B (2016). A basal ganglia circuit for evaluating action outcomes. *Nature* 539, 289–293. [PubMed: 27652894]
- Stopper CM, and Floresco SB (2014). What’s better for me? Fundamental role for lateral habenula in promoting subjective decision biases. *Nat Neurosci* 17, 33–35. [PubMed: 24270185]
- Stopper CM, Tse MTL, Montes DR, Wiedman CR, and Floresco SB (2014). Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron* 84, 177–189. [PubMed: 25220811]
- Stuber GD, Stamatakis AM, and Kantak PA (2015). Considerations when using cre-driver rodent lines for studying ventral tegmental area circuitry. *Neuron* 85, 439–445. [PubMed: 25611514]
- Tian J, Huang R, Cohen JY, Osakada F, Kobak D, Machens CK, Callaway EM, Uchida N, and Watabe-Uchida M (2016). Distributed and mixed information in monosynaptic inputs to dopamine neurons. *Neuron* 91, 1374–1389. [PubMed: 27618675]
- Tian J, and Uchida N (2015). Habenula lesions reveal that multiple mechanisms underlie dopamine prediction errors. *Neuron* 87, 1304–1316. [PubMed: 26365765]
- Tom SM, Fox CR, Trepel C, and Poldrack RA (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518. [PubMed: 17255512]
- Trusel M, Nuno-Perez A, Lecca S, Harada H, Lalive AL, Congiu M, Takemoto K, Takahashi T, Ferraguti F, and Mameli M (2019a). Punishment-Predictive Cues Guide Avoidance through Potentiation of Hypothalamus-to-Habenula Synapses. *Neuron* 102, 120–127 e124. [PubMed: 30765165]
- Trusel M, Nuno-Perez A, Lecca S, Harada H, Lalive AL, Congiu M, Takemoto K, Takahashi T, Ferraguti F, and Mameli M (2019b). Punishment-Predictive Cues Guide Avoidance through Potentiation of Hypothalamus-to-Habenula Synapses. *Neuron*.
- Ungless MA, and Grace AA (2012). Are you or aren’t you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci* 35, 422–430. [PubMed: 22459161]
- Vento PJ, Burnham NW, Rowley CS, and Zhou TC (2017). Learning from one’s mistakes: a dual role for the rostromedial tegmental nucleus in the encoding and expression of punished reward seeking. *Biol Psychiatry* 81, 1041–1049. [PubMed: 27931744]
- Wang D, Li Y, Feng Q, Guo Q, Zhou J, and Luo M (2017). Learning shapes the aversion and reward responses of lateral habenula neurons. *Elife* 6.
- Watabe-Uchida M, Eshel N, and Uchida N (2017). Neural Circuitry of Reward Prediction Error. *Annual review of neuroscience* 40, 373–394.
- Wiegert JS, Mahn M, Prigge M, Printz Y, and Yizhar O (2017). Silencing Neurons: Tools, Applications, and Experimental Constraints. *Neuron* 95, 504–529. [PubMed: 28772120]
- Yang CF, Chiang MC, Gray DC, Prabhakaran M, Alvarado M, Juntti SA, Unger EK, Wells JA, and Shah NM (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* 153, 896–909. [PubMed: 23663785]
- Zhang R, Brennan TJ, and Lo AW (2014). The origin of risk aversion. *Proc Natl Acad Sci U S A* 111, 17777–17782. [PubMed: 25453072]

**Highlights.**

Rostromedial tegmental (RMTg) neurons are activated by punishment-related stimuli

Distinct RMTg afferents communicate distinct punishment-related signals

These RMTg afferents drive correspondingly distinct aspects of punishment learning

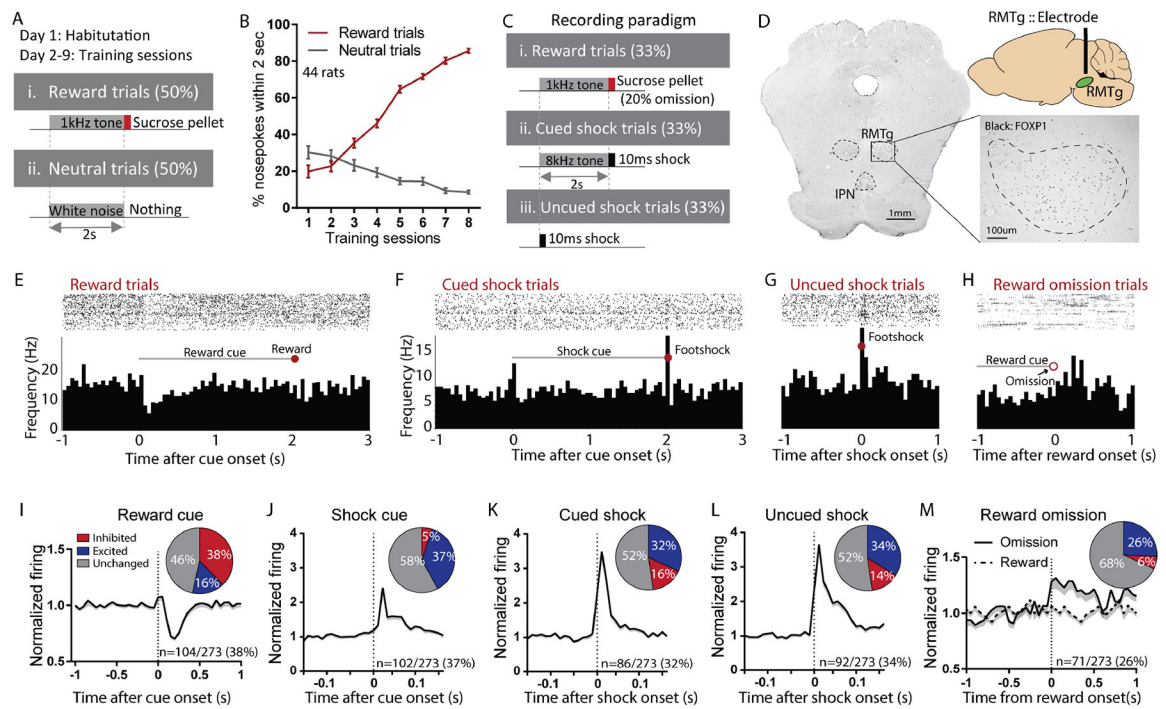
Negative valence encoding in the ventral tegmental area depends on the RMTg

Author Manuscript

Author Manuscript

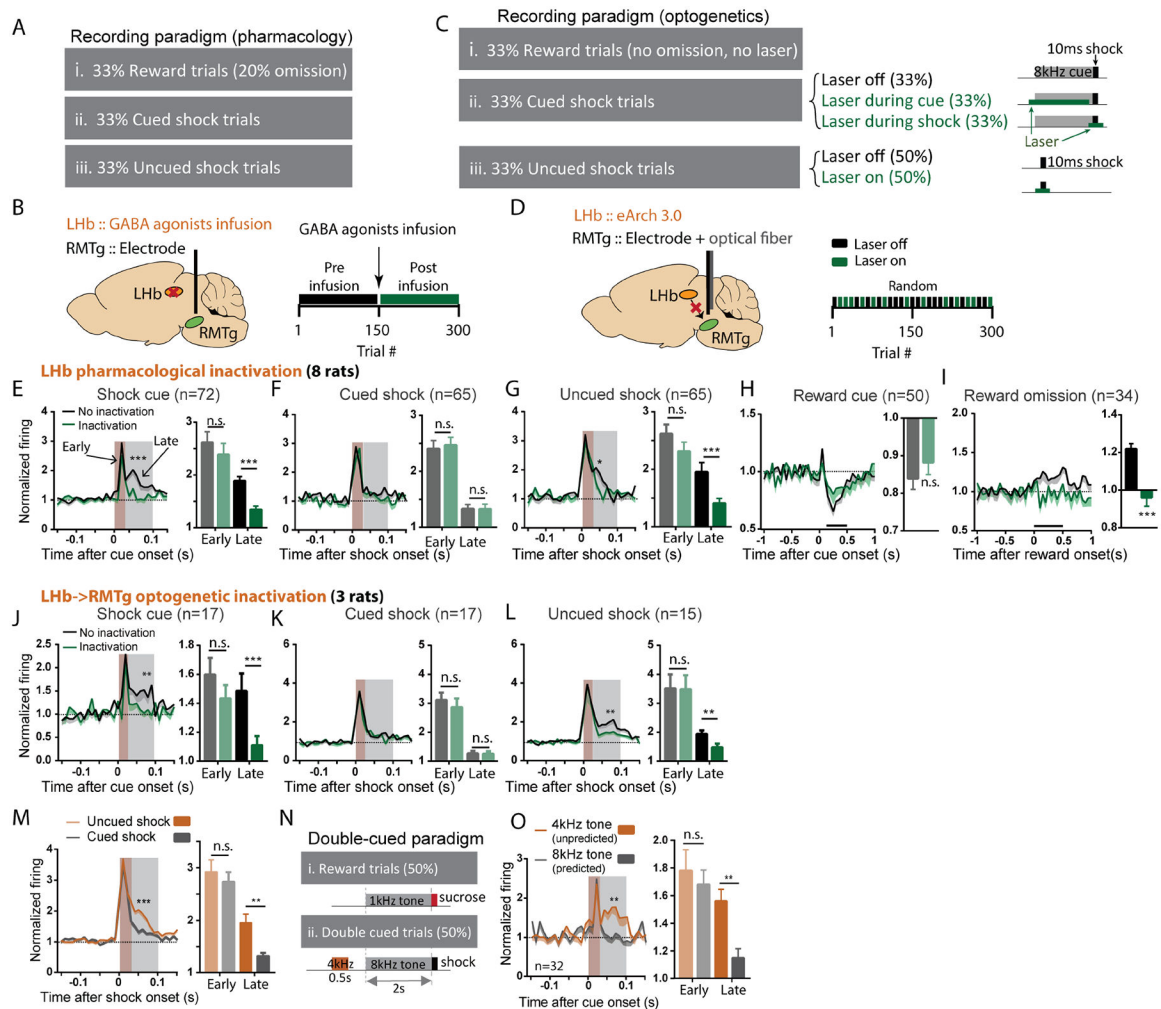
Author Manuscript

Author Manuscript



**Figure 1. Rostromedial tegmental (RMTg) neurons respond to affective stimuli.**

(A-B) Animals used for RMTg recordings initially underwent Pavlovian training until they achieved >85% accuracy discriminating reward versus neutral cues. (C) During recordings, cued and uncued shock trials were also added, along with occasional reward omissions. (D) Electrode placements were determined with respect to RMTg boundaries delineated by FOXP1 immunostaining. (E-H) Raster plots and histograms of representative RMTg neuron responses to reward, cued shock, reward omission, or uncued shock trials. (I-M) Averaged responses and proportions of RMTg neurons that were inhibited by reward cues and excited by shock predictive cues, shocks (cued or uncued) and reward omission.

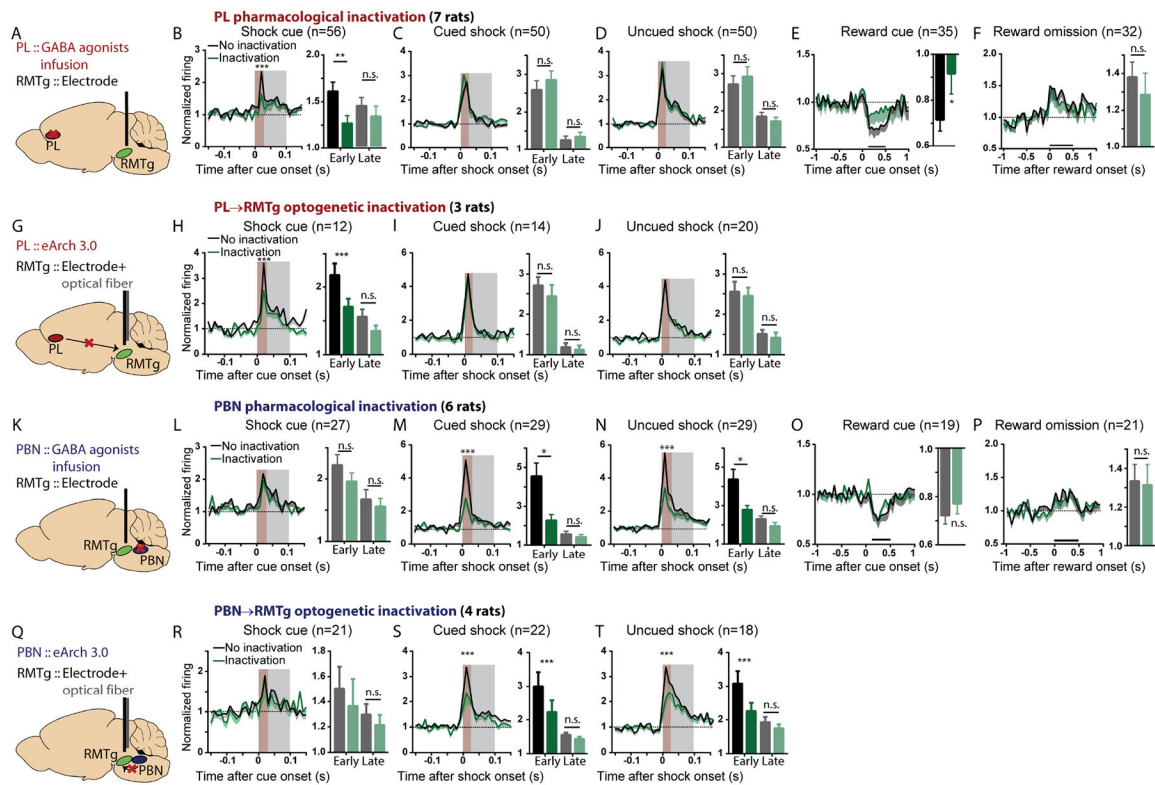


**Figure 2. The lateral habenula (LHb) contributes to RMTg activations by aversive prediction errors.**

(A-B) We recorded RMTg neuron responses to rewards, shocks, and their predictive cues before and after pharmacological inhibition of the LHb (via baclofen/muscimol cocktail infusion) during the latter half of recording sessions. (C-D) For optogenetic inhibition, virus expressing inhibitory eArch 3.0 was bilaterally injected into the LHb and an optical fiber implanted in the RMTg along with the recording electrode. Laser was turned on during randomly selected trials and remained off in others. (E-I) Time course graphs aligned to cues or outcomes (as denoted above each graph), along with adjacent bar graphs of average firing during intervals indicated by shaded windows or black bars, show that LHb inactivation selectively attenuated RMTg responses to unexpected cues and shocks (grey shaded windows in E, G), and unexpected reward omission (black bar in I), all of which represent “worse than expected” prediction errors ( $p=0.1069$  and  $p<0.0001$  for reward cue and reward omission;  $p=0.0871$  and  $p=0.001$ ,  $p=0.8615$  and  $p=0.9794$ ,  $p=0.074$  and  $p=0.0025$  for early and late of shock cue, cued shock, and free shock, respectively). (J-L) Optogenetic inhibition of LHb inputs to the RMTg selectively decreased the same responses. (M) Uncued and cued shocks elicited similar excitations during an “early” time window

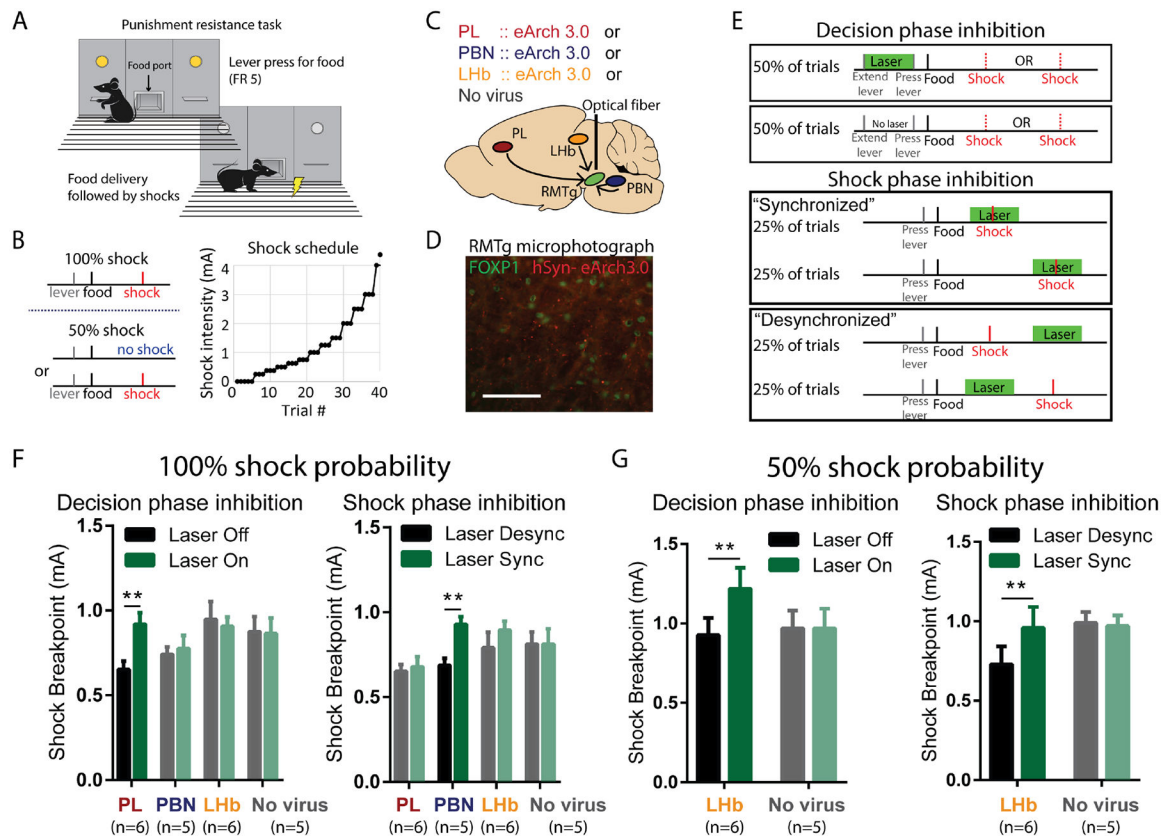
after stimulus onset (0–30ms,  $p > 0.05$ ), while uncued shocks produced a much larger response during a “late” time window (30–100ms,  $p < 0.0001$ ), consistent with an aversive prediction error. (**N**) Some trials used a double cue consisting of two consecutive auditory tones (4kHz and 8kHz) predicting a subsequent shock, making the second cue fully predicted by the first cue. (**O**) Both cues elicited similar early phase firing, but the first (unpredicted) cue elicited much greater late phase activation ( $p = 0.889$  and  $p = 0.005$ ), again consistent with an aversive prediction error. All p-values are from paired t-tests and repeated measures two-way ANOVA, with Holm-Sidak multiple comparisons post-hoc tests. Brown and grey shaded areas: early and late phase response windows. Data for each individual neuron is shown in Figures S2A–E, and individual animal histology in **S2F**.





**Figure 3. The prelimbic cortex (PL) and parabrachial nucleus (PBN) drive RMTg activations by predictive cues and shocks, respectively.**

(A, G, K, Q) Illustrations of pharmacological and optogenetic inhibition. Trial designs were the same as in Figures 2A, C. (B-F, H-J) PL inactivation selectively attenuated RMTg early phase (0–30ms, brown shaded stripe) responses to shock cues and reward cues, but not other stimuli ( $p=0.0038$  and  $p=0.2777$ ,  $p=0.1331$  and  $p=0.5463$ ,  $p=0.4690$  and  $p=0.5075$  for early and late components of shock cue, cued shock, and free shock, respectively;  $p=0.0264$  and  $p=0.3924$  for reward cue and reward omission). (L-P, R-T) PBN inactivation selectively attenuated RMTg early phase responses to shocks, without affecting other responses ( $p=0.1484$  and  $p=0.4233$ ,  $p=0.0009$  and  $p=0.8003$ ,  $p<0.0001$  and  $p=0.2680$ , for early and late responses to shock cue, cued shock, and uncued shock, respectively;  $p=0.2011$  and  $p=0.8521$  for reward cue and reward omission). Neurons that were not inhibited by reward cues or not activated by shock cues, shocks, or reward omission were excluded from analyses in Figures 2 and 3, but data for all neurons, along with histology, are shown in Figures S2, S3, and S4. All p values are from paired t-tests and repeated measures two-way ANOVA, with Holm-Sidak multiple comparisons post-hoc tests. Brown and grey shaded areas: early and late phases.



**Figure 4. PL, PBN, and LHb inputs to the RMTg modulate reward seeking under punishment during distinct phases of decision-making.**

(A, B) Behavioral paradigm in which lever pressing for food reward was punished by progressively increasing footshock intensity until a behavioral breakpoint is reached. (C) Schematic of optogenetic procedures in which AAV expressing eArch3.0 is injected into PL, PBN, or LHb, and optical fiber is targeted to RMTg. (D) Photomicrograph of eArch 3.0-expressing terminals (red) intermingled among RMTg neurons identified by FOXP1 immunostaining (green). (E) Laser light (green horizontal bar) was delivered during either the “decision phase” or “shock phase” of the task. Laser delivery during footshock (“synchronized” condition) is compared with control trials in which laser is delivered just before/after footshock (“desynchronized”). (F) Optogenetic inhibition of PL or PBN projections to RMTg increased shock breakpoint when light was delivered during the decision phase or shock phase, respectively, but not vice versa. Inhibition of LHb projection to RMTg was without effect (for decision phase, interaction of pathway  $\times$  inactivation:  $p=0.04$ , with *post-hoc* tests showing  $p=0.002$  for PL and  $p=0.923$  for PBN and LHb; for shock phase, interaction of pathway  $\times$  inactivation:  $p=0.014$ , with *post-hoc* tests showing  $p=0.0019$  for PBN,  $p=0.888$  for PL and LHb). (G) Optogenetic inhibition of LHb projection to RMTg increased shock breakpoints during both the decision phase and shock phase when shocks were delivered at 50% probability, even though no effect had been seen earlier with shocks delivered at 100% probability (for decision phase, interaction:  $p=0.038$ , *post-hoc*:  $p=0.014$  and  $p>0.999$  for LHb and no virus groups; for shock phase, interaction:  $p=0.058$ ,

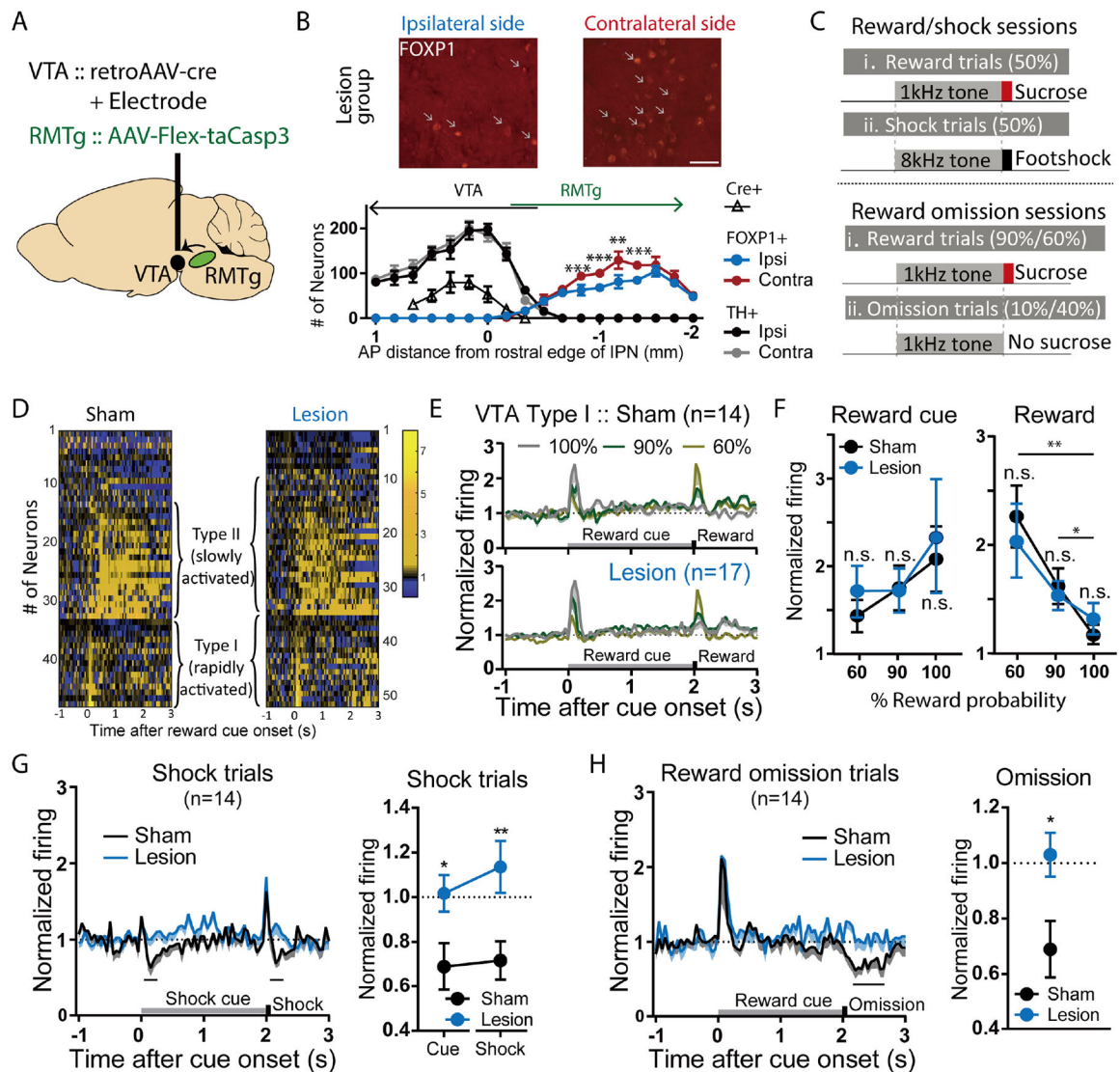
*post-hoc*:  $p=0.04$  and  $p=0.7575$  for Lhb and no virus groups). All status use repeated measures two-way ANOVA, followed by Holm-Sidak multiple comparison *post-hoc* test.

Author Manuscript

Author Manuscript

Author Manuscript

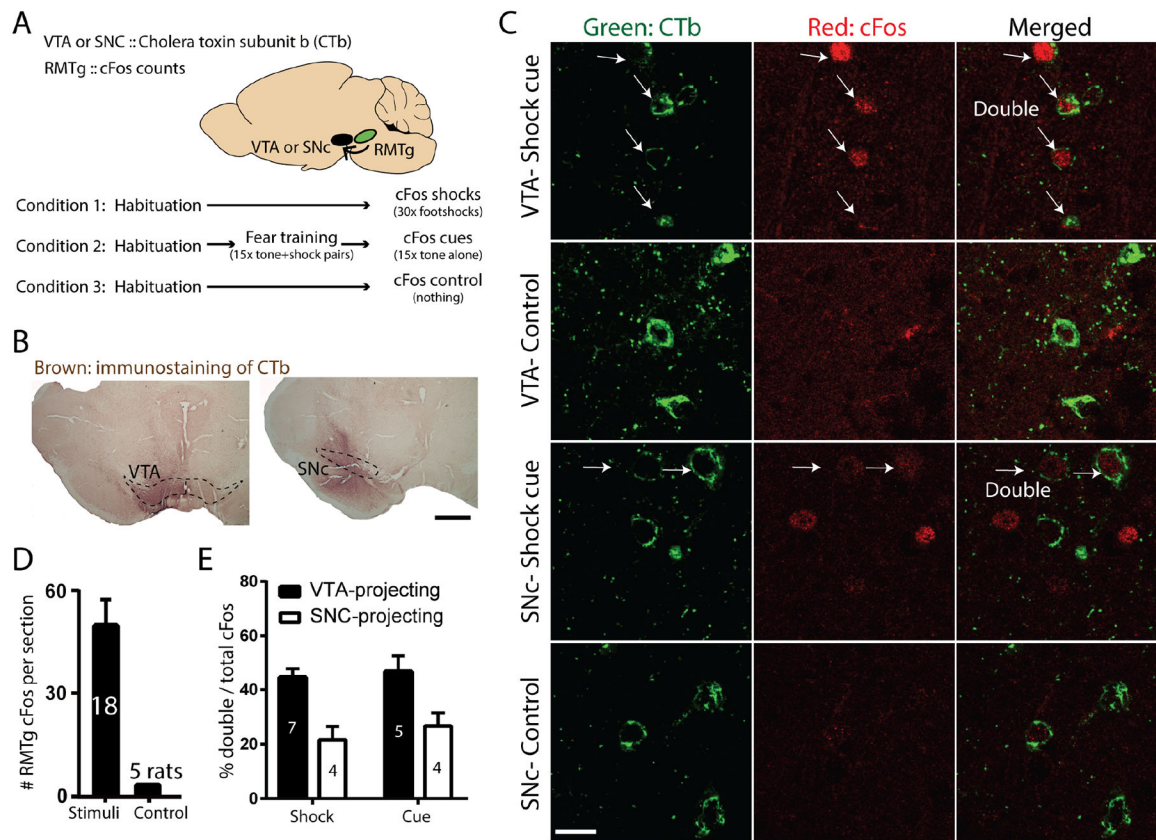
Author Manuscript



**Figure 5. Selective ablation of VTA-projecting RMTg neurons abolishes VTA neuron inhibition by aversion-related signals, but not excitation to reward-related signals.**

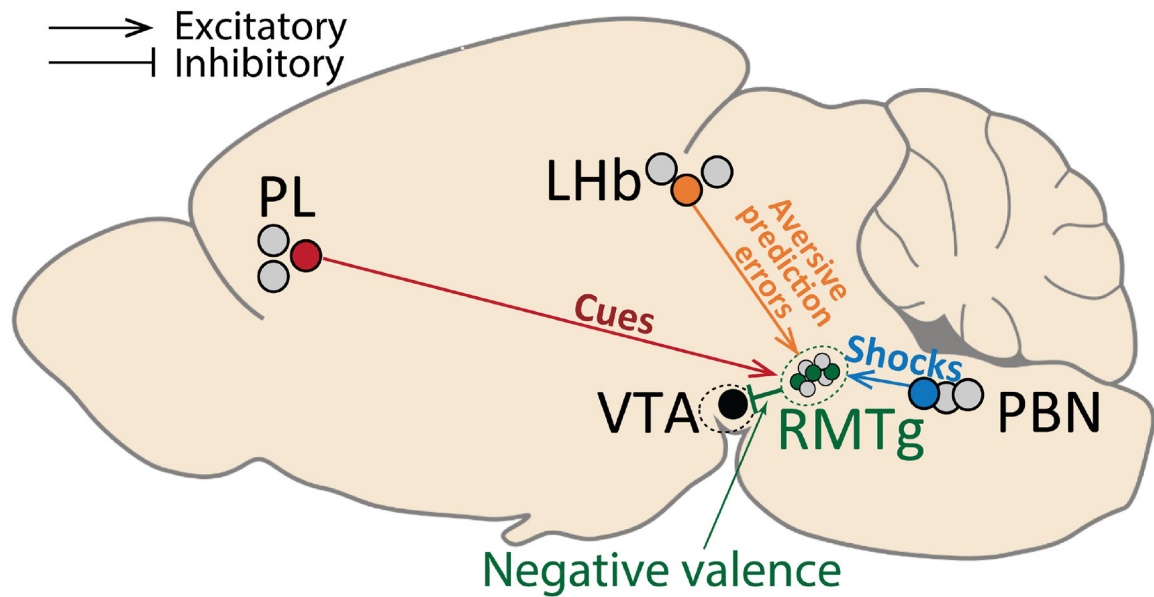
(A) Schematic of viral injections of AAV-Cre into the VTA, and AAV-FLEX-taCasp3 into the RMTg, selectively ablating VTA-projecting RMTg neurons. (B) Quantitation of RMTg FOXP1-positive cells and VTA TH-positive cells shows reduction of FOXP1 ipsilateral to AAV-Cre injection, compared with the contralateral side. Injections of FLEX-taCasp3 into RMTg were placed 1.9mm caudal to AAV-Cre injections, minimizing spread to the rostral RMTg and VTA. Scale bar: 50 $\mu$ m. (C) Schematic of recording sessions in which distinct auditory tones are followed by sucrose at 60%, 90%, or 100% probability, or shock at 100% probability. (D) Heatmaps of individual VTA neuron responses to reward trials. (E, F) Both sham and lesion groups show neurons that are rapidly (type I) or slowly (type II) activated by reward cues. Responses to reward cues or rewards were not affected by ablation at any of the three reward probabilities (For reward cue response,  $p=0.99$  for 60%, 90%, and 100%, sham and lesion group;  $p=0.99$  and  $p=0.92$ , and  $p=0.99$  and  $p=0.77$  for 60% vs. 90% and 60% vs. 100% in sham and lesion group, respectively. For reward response,  $p=0.92$  for 60%,

90%, and 100% compared between sham and lesion group;  $p=0.01$  and  $p=0.001$ , and  $p=0.03$  and  $p=0.12$  for 60% vs. 90% and 60% vs. 100% in sham and lesion group, respectively, two-way ANOVA, with Holm-Sidak multiple comparisons test). (**G, H**) Type I VTA neurons were inhibited by footshocks, shock-predictive cues, and 10% reward omission. Ablation of VTA-projecting RMTg neurons eliminated all three of these inhibitions (blue traces/symbols) (shock:  $p=0.008$ , cue:  $p=0.044$ ; two-way ANOVA, with Holm-Sidak multiple comparisons test; 10% omission:  $p=0.022$ , paired t-test for firing rates during analysis windows shown with black bars).



**Figure 6. RMTg neurons broadcast punishment signals to multiple targets.**

(A) Schematic of experimental design in which we injected retrograde tracer CTb into either the VTA or SNC, and induced cFos with either footshocks or shock-predictive auditory cues. (B) Immunostaining of CTb injection sites in VTA and SNC. Scale bar: 500 $\mu$ m. (C) Representative photomicrographs of CTb and cFos in the RMTg. (D) Shocks and shock cues greatly increased cFos in RMTg versus unstimulated animals, which showed extremely low RMTg cFos levels. (E) Shock- and shock-cue activated RMTg neurons were enriched in both VTA-projecting, and to a lesser extent SNC-projecting, subpopulations.



**Figure 7. Summary of functions subserved by discrete RMTg afferents.**

RMTg activations by reward-and punishment-predictive cues, footshocks, and “worse than expected” prediction errors are dependent on triply dissociable inputs from PL (red symbols/lines), PBN (blue symbols/lines), and LHb (orange symbols/lines), respectively, that drive corresponding aspects of punishment learning. Information about all negative motivational stimuli is further transmitted from RMTg to the VTA (green symbols/lines), where it is combined with information about reward-related stimuli. Grey symbols represent neurons that do not project to the same target as the colored symbols, reflecting heterogeneity of projection targets (and likely function) in these regions.