

# Tonic activity in lateral habenula neurons acts as a neutral valence brake on reward-seeking behavior

5 Ryan J. Post<sup>1,2,3</sup>, David A. Bulkin<sup>1,2,3</sup>, R. Becket Ebitz<sup>4</sup>, Vladlena Lee<sup>1</sup>, Kasey Han<sup>1</sup>, Melissa R. Warden<sup>1,2,5,\*</sup>

<sup>1</sup>Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853, USA

<sup>2</sup>Cornell Neurotech, Cornell University, Ithaca, NY 14853 USA

10 <sup>3</sup>These authors contributed equally

<sup>4</sup>Department of Neuroscience, Université de Montréal, Montréal, QC H3T 1J4, Canada

<sup>5</sup>Lead Contact

\*Correspondence: mrwarden@gmail.com

15

## SUMMARY

Survival requires both the ability to persistently pursue goals and the ability to determine when it is time to stop, an adaptive balance of perseverance and disengagement. Neural activity in the lateral habenula (LHb) has been linked to negative valence, but its role in regulating the balance between engaged reward-seeking and disengaged behavioral states remains unclear. Here, we show that LHb neural activity is tonically elevated during minutes-long periods of disengagement from reward-seeking behavior, both when due to repeated reward omission (negative valence) and when sufficient reward has been consumed (positive valence). Further, we show that LHb inhibition extends ongoing reward-seeking behavioral states but does not prompt task re-engagement. We find no evidence for similar tonic activity changes in ventral tegmental area dopamine neurons. Our findings support a framework in which tonic activity in LHb neurons suppresses engagement in reward-seeking behavior in response to both negatively- and positively-valenced factors.

Keywords: lateral habenula, ventral tegmental area, dopamine, reward, tonic activity, calcium imaging, motivated behavior, behavioral state, disengagement, valence

## INTRODUCTION

Animals transition between directed pursuit of rewards and exploratory or quiescent behavioral states on a timescale of minutes to hours (Ferster and Skinner, 1957; Cohen et al., 2007; Flavell et al., 2013; Hills et al., 2015; Stern et al., 2017; Ebitz et al., 2018; Marques et al., 2020). Factors that influence the persistence of reward-seeking behavioral states include current and predicted homeostatic need (Aponte et al., 2011; Chen et al., 2015), the history of action successes and failures (Vroom, 1964; Charnov, 1976; Ullsperger and von Cramon, 2003; Ebitz et al., 2019), reward proximity (Howe et al., 2013; McGinty et al., 2013; Westbrook and Frank, 2018; Guru et al., 2020), opportunity costs (Niv et al., 2007; Kurzban et al., 2013; Boureau et al., 2015), and environmental threats (Lecca et al., 2017; Alhadeff et al., 2018). These factors differ in modality and valence, but all can contribute to suppressing reward-seeking behavior.

The past decade has seen a surge of interest in the role of the lateral habenula (LHb) in regulating behavior. The LHb, part of the epithalamus, is a major conduit of information from the forebrain to brainstem neuromodulatory centers, and LHb neural activity reflects events and behavioral states at a range of timescales (Bianco and Wilson, 2009; Hikosaka, 2010; Proulx et al., 2014; Hu et al., 2020). At the sub-second timescale, LHb neurons fire phasically when predicted rewards are omitted, when cues that predict reward omission or punishment appear, and when shocks or air-puffs are delivered, and phasic LHb activity scales with negative value (Matsumoto and Hikosaka, 2007, 2009; Lecca et al., 2017). At longer timescales, neural and metabolic activity in the LHb is elevated during helplessness and passive coping in aversive environments (Caldecott-Hazard et al., 1988; Morris et al., 1999; Shumake et al., 2003; Mirrione et al., 2014; Proulx et al., 2018; Yang et al., 2018; Andalman et al., 2019), and excitatory synaptic transmission onto LHb neurons is potentiated in depression-like behavioral states (Li et al., 2011,

2013; Shabel et al., 2014; Lecca et al., 2016). These findings have suggested a connection between LHb neural activity and negative valence (Friedman et al., 2011; Stamatakis and Stuber, 2012; Lammel et al., 2012; Proulx et al., 2014).

Recent evidence has also implicated the LHb in the suppression of reward-seeking behavior. LHb stimulation reduces the number of actions that animals are willing to perform for rewards (Proulx et al., 2018), and blocking GABAergic signaling within the LHb disrupts anticipatory licking in reward tasks (Lalivert et al., 2021). Additionally, lesioning the LHb or silencing excitatory inputs to the LHb elevates consumption of palatable food but not contaminated food (Paul et al., 2011; Stamatakis et al., 2016), suggesting that the LHb is necessary for satiety- but not disgust-based feeding suppression. Ventral tegmental area (VTA) dopamine (DA) neurons play an essential role in reward and sustained goal pursuit (Schultz et al., 1997; Salamone and Correa, 2012; Dolan and Dayan, 2013; Howe et al., 2013; Guru et al., 2020), and the LHb suppresses VTA dopamine DA neural activity via the GABAergic rostromedial tegmental nucleus (RMTg) (Christoph et al., 1986; Ji and Shepard, 2007; Matsumoto and Hikosaka, 2007; Jhou et al., 2009; Hong et al., 2011). The LHb-RMTg-VTA pathway has been implicated in the transmission of negative reward information to VTA DA neurons (Matsumoto and Hikosaka, 2007; Hong et al., 2011; Stamatakis and Stuber, 2012), and the RMTg has been hypothesized to function as a 'brake' on VTA DA neural activity (Barrot et al., 2012).

As reward-seeking can be suppressed by both negative factors (reward omission, threats) and positive factors (satiety, exploration), it is conceivable that LHb neural activity may play a role in both negatively- and positively-motivated disengagement from reward-seeking. Here, we combine fiber photometry, multi-unit electrophysiology, and optogenetics in freely behaving mice to investigate the role of LHb neural activity in regulating the persistence of reward-seeking behavioral states in different contexts. Our findings reveal that large, sustained increases in tonic LHb neural activity accompany minutes-long states of disengagement from reward-seeking behavior, and that these tonically excited LHb states can be driven either by task disengagement due to repeated reward omission (negative valence) or by task disengagement following the consumption of sufficient reward (positive valence). Further, we demonstrate a causal role for LHb neural activity in regulating reward-seeking behavior by showing that LHb inhibition extends reward-seeking behavioral states. Our findings support the hypothesis that tonically elevated LHb neural activity reduces engagement in reward-seeking behavior in response to both negatively- and positively-valenced factors.

## RESULTS

### Calcium Dynamics in Genetically Targeted LHb Neurons

The LHb and surrounding brain tissue are composed primarily of glutamatergic neurons, and achieving targeted expression of genetically encoded tools in the LHb while confidently excluding expression in adjacent brain regions is challenging. In the *klf8-Cre* (NP171) mouse line (GENSAT; Proulx et al., 2014) Cre recombinase is expressed in the habenula but not in adjacent brain regions, a feature that allowed us to restrict the expression of genetically-encoded optical indicators and actuators to the LHb. Although Cre is also expressed in the medial habenula (MHb)

in this line, we were able to consistently minimize MHb expression by refinement of viral vector serotype, injection volume, and coordinates (see Methods). To characterize Cre-dependent gene expression in *Klk8-Cre* (NP171) mice, we injected AAV9-CAG-Flex-GFP in the LHb and quantified co-localization of GFP and NeuN, a marker of neuronal identity (Mullen et al., 1992). We found that GFP expression was highly specific to neurons (97.8%  $\pm$  0.49% specificity), and that the majority of NeuN-labeled neurons at the injection site were GFP-positive (71.6%  $\pm$  2.97% penetrance,  $n$  = 365 cells, **Figure 1A**).

To assess the utility of the *Klk8-Cre* (NP171) mouse line for monitoring LHb activity, we used fiber photometry (Cui et al., 2013; Gunaydin et al., 2014) to record LHb population calcium activity during reward consumption and omission. We injected a genetically-encoded calcium indicator, AAV9-CAG-Flex-GCaMP6s (Chen et al., 2013), into the LHb of *Klk8-Cre* (NP171) mice, and implanted an optical fiber over the LHb to monitor calcium-dependent fluorescence (**Figures 1C and 1D**). Because LHb neurons are inhibited by reward consumption and excited by the omission of predicted rewards (Matsumoto and Hikosaka, 2007, 2009; Bromberg-Martin and Hikosaka, 2011), we examined LHb activity during performance of a simple self-paced operant task. In this 'poke-reward' task, mice poked their nose into a port on one side of an operant chamber in order to trigger the delivery of a water reward on the other side (**Figure 1B**). We recorded LHb activity while mice performed a probabilistic version of this task in which each trial had a 20% chance to yield no reward, a 60% chance to yield a medium reward (10  $\mu$ l), and a 20% chance to yield a large reward (20  $\mu$ l). LHb neural activity during the reward consumption epoch was negatively correlated with reward value (GCaMP6s ( $n$  = 34 sessions, 7 mice), GFP ( $n$  = 15 sessions, 3 mice),  $p$  < 0.0001, two-sample t-test, **Figures 1E and 1F**). We also recorded activity from VTA DA neurons in *DAT-Cre* mice (Bäckman et al., 2006) using fiber photometry during performance of the same task. Consistent with previous findings (Schultz et al., 1997; Matsumoto and Hikosaka, 2007; Cohen et al., 2012; Guru et al., 2020; Kim et al., 2020), VTA DA neural activity was positively correlated with reward value (GCaMP6 ( $n$  = 28 sessions, 6 mice), GFP ( $n$  = 12 sessions, 3 mice),  $p$  < 0.0001, two-sample t-test, **Figures 1G-1J**). Thus, LHb neural activity during reward consumption in the *Klk8-Cre* (NP171) line is qualitatively similar to previous findings.

To compare these results to other methods used for recording LHb neural activity, we used fiber photometry to record the activity of genetically-targeted glutamatergic LHb neurons in *Vglut2-ires-Cre* mice (Vong et al., 2011) during performance of the same task (**Figures S1A and S1B**). Most neurons in the LHb are glutamatergic and express *Vglut2* (Brinschwitz et al., 2010), and *Vglut2-ires-cre* mice have previously been used to target the LHb (Wang et al., 2017; Flanigan et al., 2020). We also recorded LHb neural activity in C57BL6/J mice using multi-unit electrophysiology (**Figures S1E and S1F**). Both LHb *Vglut2-ires-Cre* neural activity (GCaMP6 ( $n$  = 15 sessions, 3 mice), GFP ( $n$  = 15 sessions, 3 mice),  $p$  < 0.0001, two-sample t-test, **Figures S1C and S1D**) and LHb multi-unit activity ( $n$  = 448 recordings across 40 electrodes in 6 mice,  $p$  < 0.0001, two-sample t-test compared to shuffled sample, **Figures S1G and S1H**) were negatively correlated with reward value, consistent with our findings in *Klk8-Cre* (NP171) mice. Thus, phasic reward-related responses were consistent between *Klk8-Cre* (NP171), *Vglut2-ires-Cre*, and multi-unit electrophysiology recordings. Neural activity during reward approach was similar between *Klk8-Cre* (NP171) and multi-unit electrophysiology recordings but distinct from that observed in *Vglut2-ires-Cre* recordings (**Figures 1E, S1C, and S1G**). This may reflect the challenges in restricting vector expression to the LHb via glutamatergic targeting using the *Vglut2-*

ires-Cre line, and highlights the utility of the Klk8-Cre (NP171) line in driving regionally-restricted expression.

### Tonically Elevated LHb Activity During Spontaneous Task Disengagement

The largest LHb activity changes we observed were periods of sustained excitation that occurred when mice spontaneously disengaged from task performance, typically toward the end of recording sessions after the consumption of numerous rewards. LHb neural activity in Klk8-Cre (NP171) mice was tonically elevated throughout these disengaged periods and decreased when mice spontaneously re-engaged in the task (**Figures 2A and 2B**). Here, we use tonic to refer to population activity that is elevated across all task epochs, as opposed to phasic activity restricted to reward omission. Although the single neuron spiking activity patterns that underlie this calcium signal are not known and may reflect changes in either tonic firing or burst rate, there is some evidence that tonic spiking activity is preferentially represented in calcium signals due to the longer integration window (Wei et al., 2020).

To quantify behavior, we used a two-state hidden Markov model (HMM, see Methods) to identify periods of time during which mice were less engaged in task performance (low engagement states) and periods during which mice were more strongly engaged (high engagement states) (Murphy, 2012). We have previously used this approach to differentiate periods of exploration and exploitation (Ebitz et al., 2018, 2019). Note that although the HMM assigns a binary state to each time point, this does not imply that the true underlying engagement state or LHb neural activity dynamics are similarly abrupt.

We regressed task engagement state onto  $\Delta F/F$  for each session, and included running speed as a predictor to assess whether locomotor activity may also contribute to tonic LHb activity. LHb neural activity was negatively correlated with task engagement state, but was not correlated with running speed or the interaction between task engagement state and running speed (GCaMP6 (n = 34 sessions, 7 mice), GFP (n = 15 sessions, 3 mice), engagement state:  $p < 0.0001$ , running speed:  $p = 0.8826$ , engagement state X running speed:  $p = 0.9387$ , two-sample t-test, **Figure 2C**). Likewise, we found that LHb activity in Vglut2-ires-Cre mice was negatively correlated with task engagement state, but was not correlated with running speed or the interaction between task engagement state and running speed (GCaMP6 (n = 15 sessions, 3 mice), GFP (n = 13 sessions, 3 mice), engagement state:  $p = 0.0021$ , running speed:  $p = 0.2577$ , engagement state X running speed:  $p = 0.4584$ , two-sample t-test, **Figures S2A and S2B**). LHb neural activity recorded via multi-unit electrophysiology was negatively correlated with task engagement state, but was positively correlated with running speed (n = 448 recordings across 40 electrodes in 6 mice, engagement state:  $p < 0.0001$ , running speed:  $p < 0.0001$ , t-test compared to shuffled sample, **Figures S2C and S2D**). Thus, tonically elevated LHb neural activity reflects periods of spontaneous disengagement from reward-seeking task performance.

When we examined VTA DA neural activity in DAT-Cre mice during this task we found no correlation with task engagement state or with interaction between task engagement state and running speed, but we observed a small negative correlation with running speed (GCaMP6 (n = 28 sessions, 6 mice), GFP (n = 12 sessions, 3 mice), engagement state:  $p = 0.6700$ , running speed:  $p = 0.0039$ , engagement state X running speed:  $p = 0.1227$ , two-sample t-test, **Figures 2D-2F**). The absence of tonic VTA DA neural activity changes between high and low task

engagement states may be related to previous reports of stable baseline VTA DA neural activity through changing schedules of rewards and punishments (Cohen et al., 2015; Mohebi et al., 2019).

## Baseline Tonic LHb Activity Reflects Task Engagement State

To confirm that tonic disengagement-related LHb excitation was not simply a consequence of state-dependent differences in phasic reward-related activity, we examined LHb activity during two task epochs: the baseline epoch (1 second prior to each trial-initiating poke) and the reward epoch (1 second following the first lick). The baseline epoch occurred during the run to the nose-poke, the best option for analysis during this self-paced task without an inter-trial interval. We found greater mean LHb neural activity in low compared to high engagement states during both baseline and reward epochs in *Klk8-Cre* (NP171) mice ( $n = 34$  sessions, 7 mice, baseline:  $p = 0.0003$ , reward:  $p = 0.0001$ , paired samples t-test, **Figures 3A and 3B**). We found the same effect in *Vglut2-ires-Cre* mice ( $n = 15$  sessions, 3 mice, baseline:  $p = 0.0115$ , reward:  $p = 0.0104$ , paired samples t-test, **Figures S3A and S3B**) and via multi-unit electrophysiology ( $n = 448$  recordings across 40 electrodes in 6 mice, baseline:  $p < 0.0001$ , reward:  $p < 0.0001$ , paired samples t-test, **Figures S3C and S3D**).

In contrast, VTA DA neural activity during the baseline epoch did not depend on task engagement state ( $n = 28$  sessions, 6 mice,  $p = 0.4183$ , paired samples t-test, **Figures 3C and 3D**). However, VTA DA neural activity during the reward epoch was significantly greater during high engagement states ( $n = 28$  sessions, 6 mice,  $p < 0.0001$ , paired samples t-test, **Figures 3C and 3D**), which may be related to the diminished phasic VTA DA reward response during states of satiety (Bassareo and Chiara, 1999; Branch et al., 2013; Papageorgiou et al., 2016). Although task engagement state is distinct from satiety and evolves on a faster timescale, satiety contributes to the probability of being in a disengaged state in reward-based tasks (Benelam, 2009; Allen et al., 2019). Indeed, in the poke-reward task disengaged states tended to occur toward the end of each recording session following the consumption of numerous rewards. An alternative hypothesis is that task disengagement may have been induced by fatigue. Fatigue is associated with a reduction in spontaneous locomotor activity in mice (Harrington, 2012), thus if fatigue contributed to task disengagement mice would be expected to move less during low engagement states. To investigate this possibility we compared locomotor activity during periods of high and low task engagement. We found that average speed was not significantly different between states (high engagement speed:  $5.6326 \pm 0.1348$  cm/s, low engagement speed:  $5.0406 \pm 0.1062$  cm/s,  $n = 105$  sessions,  $p = 0.6031$ , two-sample t-test), inconsistent with this hypothesis.

As rewards during this task were omitted on 20% of trials, reward omission is another factor that may have contributed to task disengagement. If so, entries into low engagement states would be expected to be preceded by non-rewarded trials more often than expected by chance. Contrary to this hypothesis, we found that entries into low engagement states were actually less likely to be preceded by reward omissions and more likely to be preceded by medium or large rewards than expected by chance (no reward:  $p < 0.0001$ , medium reward:  $p = 0.0037$ , large reward:  $p = 0.0001$ , binomial test, **Figure S3E**). We also found that entries into high engagement states were more likely to be preceded by reward omissions and less likely to be preceded by medium rewards than expected by chance (no reward:  $p < 0.0001$ , medium reward:  $p < 0.0001$ , large reward:

0.8418, binomial test, **Figure S3E**). Thus, mice were less likely to disengage if a reward was omitted on the previous trial and more likely to disengage when the previous trial was rewarded, a finding perhaps related to the burst of reward-seeking behavior that occurs at the onset of extinction training (Skinner, 1938; Cooper et al., 1987). Together, these findings argue against the hypothesis that fatigue or reward omission drove entry into low engagement states in the probabilistic poke-reward task, and further support the hypothesis that satiety contributed to task disengagement.

### **LHb Activity is Tonically Elevated During Task Disengagement Due to Repeated Reward Omission**

We next asked whether tonic LHb excitation is specific to satiety-driven task disengagement, or if LHb neural activity might also be tonically elevated when mice disengage from task performance in response to repeated reward omission. To address this question, we recorded LHb neural activity in *Klk8-Cre* (NP171) mice while mice performed a blocked version of the poke-reward task. In this task, 5-minute blocks of medium-reward trials alternated with 5-minute blocks of no-reward trials (**Figure 1B**). As in the probabilistic poke-reward task, LHb neural activity was high during disengaged states and low during engaged task performance (**Figures 4A and 4B**). We regressed task engagement state onto  $\Delta F/F$  for each session and included block reward contingencies and running speed as predictors. We found that LHb neural activity was negatively correlated with task engagement state and block reward, but was not correlated with running speed (GCaMP6 ( $n = 35$  sessions, 6 mice), GFP ( $n = 15$  sessions, 3 mice), engagement state:  $p = 0.0259$ , block reward:  $p = 0.0090$ , running speed:  $p = 0.4529$ , two-sample t-test, **Figure 4C**). Regression coefficients for two- and three- way interactions were not different between GCaMP and GFP groups ( $p > 0.05$  for all comparisons, paired samples t-tests).

We found similar results in *Vglut2-ires-Cre* mice (GCaMP6 ( $n = 15$  sessions, 3 mice), GFP ( $n = 15$  sessions, 3 mice), engagement state:  $p = 0.0052$ , block reward:  $p = 0.0151$ , running speed:  $p = 0.4283$ , two-sample t-test, **Figures S4A and S4B**). As in the probabilistic reward task, LHb neural activity recorded via multi-unit electrophysiology was negatively correlated with task engagement state and block reward, but was positively correlated with running speed ( $n = 96$  recordings across 40 electrodes in 6 mice, engagement state:  $p < 0.0001$ , block reward:  $p < 0.0001$ , running speed:  $p < 0.0001$ , two-sample t-test, **Figures S4C and S4D**). VTA DA neural activity in *DAT-Cre* mice positively correlated with block reward, but we found no correlation with engagement state or running speed (GCaMP6 ( $n = 30$  sessions, 6 mice), GFP ( $n = 13$  sessions, 3 mice), engagement state:  $p = 0.5563$ , block reward:  $p = 0.0447$ , running speed:  $p = 0.4294$ , two-sample t-test, **Figures 4D-4F**). Together, these data reveal that tonically elevated activity in LHb neurons accompanies task disengagements prompted by factors with either negative (repeated reward omission) or positive (satiety) valence.

### **Rising LHb Activity Precedes Task Disengagement**

In order to characterize the time course of the emergence of elevated tonic LHb activity during task disengagement, we examined the trials leading up to and following transitions into low engagement states. In the probabilistic poke-reward task, we observed that LHb activity in *Klk8-*

Cre (NP171) mice increased across the trials preceding disengagement, while VTA DA activity did not (**Figure 5A**, data aligned to first lick for each trial). To quantify this effect, we used baseline LHb neural activity (1 s prior to the trial-initiating poke) from each of the 10 trials preceding disengagement to calculate the slope of the change in activity from the first to the last of these trials (Klk8-Cre (NP171): GCaMP6 (n = 34 sessions, 7 mice), GFP (n = 15 sessions, 3 mice), p = 0.0195; DAT-Cre: GCaMP6 (n = 28 sessions, 6 mice), GFP (n = 12 sessions, 3 mice), p = 0.4850, two-sample t-test, **Figure 5B**). We found similar effects during the blocked poke-reward task when we examined the 10 trials preceding disengagement (Klk8-Cre (NP171): GCaMP6 (n = 35 sessions, 6 mice), GFP (n = 15 sessions, 3 mice), p = 0.0060; DAT-Cre: GCaMP6 (n = 30 sessions, 6 mice), GFP (n = 13 sessions, 3 mice), p = 0.7910, two-sample t-test, **Figures 5C and 5D**). We calculated the slope of the change in baseline activity across a symmetrical 10 trial window slid in 1 trial steps across the state transition, and found that in both tasks LHb activity began increasing prior to entry into low engagement states (**Figures 5E and 5F**).

### LHb Inhibition Extends Reward-Seeking Behavioral States

If tonically elevated LHb neural activity promotes task disengagement, inhibiting the LHb should disrupt the ability to disengage from task performance. To test this hypothesis, we optogenetically inhibited LHb neurons while mice performed the poke-reward task. We bilaterally injected AAV-EF1 $\alpha$ -DIO-eNpHR3.0-eYFP into the LHb of Klk8-Cre (NP171) mice and implanted optical fibers above the LHb for light delivery (**Figures 6A and 6B**). We asked two questions: 1) does LHb inhibition during disengaged states prompt re-engagement in task performance? 2) does LHb inhibition during task-engaged states disrupt the ability to disengage from ongoing task performance?

To examine whether LHb inhibition during disengaged states prompts task re-engagement, we allowed mice to freely perform poke-reward trials for medium rewards and waited until they spontaneously refrained from task performance for two minutes. At this point light delivery was initiated, and inhibition was maintained until mice spontaneously performed another trial (**Figure 6C**). We found that LHb inhibition during disengaged states had no effect on the latency to re-engage (NpHR (n = 5), GFP (n = 5), p = 0.6669, paired samples t-test, **Figure 6D**). To examine whether LHb inhibition during engaged states disrupts the ability to disengage from ongoing task performance we began as above, but instead of starting inhibition after two minutes without trials we continued to wait until mice spontaneously completed another trial. At this point light delivery was initiated, and inhibition was maintained until mice again refrained from task performance for two minutes. Here, we found that LHb inhibition increased the mean number of trials completed before the next disengagement (NpHR (n = 5), GFP (n = 5), p = 0.0316, paired samples t-test, **Figure 6E**). Thus, LHb inhibition extends ongoing task-engaged states but does not prompt re-engagement in disengaged mice.

Finally, we asked whether LHb inhibition prolonged task-engaged states only when rewards were available, or whether LHb inhibition promoted persistent task performance even in the absence of reward. To address this question, mice were allowed to perform the poke-reward task as above, but no rewards were delivered at any time during task performance. When rewards were not available we found no effect of LHb inhibition on either the latency to re-engage (NpHR (n = 5), GFP (n = 5), p = 0.8208, paired samples t-test, **Figure S6A**) or the number of trials



completed before the next disengagement (NpHR (n = 5), GFP (n = 5),  $p = 0.3367$ , paired samples t-test, **Figure S6B**). Thus, LHb inhibition only prolongs task-engaged states when rewards are available. This finding is consistent with the hypothesis that LHb neural activity serves as a brake on the neural systems that promote reward-seeking behavior – releasing the brake would not prompt task re-engagement if a signal that promotes reward-seeking behavior is not active.

## DISCUSSION

Here, we sought to investigate how LHb neural activity reflects and regulates the persistence of reward-seeking behavioral states. LHb recordings during a self-paced reward-seeking task revealed that neural activity was tonically low during engaged task performance, but tonically high during minutes-long task disengagements. We observed these tonic activity increases both during disengagements that followed repeated reward omission (negative valence) and during spontaneous disengagements after sufficient reward consumption (positive valence). Importantly, LHb neural activity did not rise continuously with the number of rewards consumed (and thus did not directly reflect satiety), but rather rose when animals disengaged from task performance and fell when they re-engaged. Likewise, LHb neural activity rose during disengagements prompted by repeated reward omission and fell during re-engagement. Phasic LHb neural activity upon individual reward omissions was also observed, as previously demonstrated (Matsumoto and Hikosaka, 2007), but was moderate in comparison to state-dependent tonic changes. In contrast, we observed large phasic reward signals in VTA DA neural activity but did not detect engagement state-dependent tonic changes. Finally, we found that inhibiting LHb neurons prolonged ongoing reward-seeking behavioral states but did not prompt disengaged mice to resume task performance. Together, these findings are consistent with the hypothesis that tonically elevated LHb neural activity acts as a valence-neutral brake on reward-seeking behavior.

Disengagement from reward-seeking behavior can be triggered by a variety of factors with positive or negative valence. For example, an animal might stop attempting to obtain water either because it has already consumed enough water and its homeostatic needs have been met, or because its actions to obtain water have proven ineffective. Satiety is positively valenced (Betley et al., 2015; Garfield et al., 2015; Schéle et al., 2017) and action failure is negatively valenced (Skinner, 1953; Leitenberg, 1965; Amsel, 1992; Papini and Dudley, 1997), but both factors can prompt task disengagement. In support of this notion, the LHb receives its strongest afferents from basal ganglia circuits involved in action selection and evaluation and hypothalamic circuits involved in homeostatic regulation (Herkenham and Nauta, 1977; Hong and Hikosaka, 2008; Shabel et al., 2012; Proulx et al., 2014; Yetnikoff et al., 2015; Stamatakis et al., 2016; Stephenson-Jones et al., 2016; Zahm and Root, 2017; Barker et al., 2017).

Other factors that can prompt task disengagement include imminent threats and the emergence of other profitable behavioral options. In threatening situations, LHb neural activity may play a role in terminating ongoing reward-seeking behavioral states in order to permit defensive circuits to assume control of behavior. There is anatomical support for this idea, as the LHb receives inputs from regions implicated in defensive behavior, including the periaqueductal gray, superior colliculus, and raphe nuclei (Zhao et al., 2014; Shang et al., 2015; Yetnikoff et al., 2015; Tovote et al., 2016; Zahm and Root, 2017; Huang et al., 2017; Evans et al., 2018; Seo et

al., 2019). Finally, the LHb is essential for flexibility in behavioral choice. LHb inactivation impairs context-appropriate strategy switching (Thornton and Evans, 1982; Stopper and Floresco, 2014; Baker et al., 2015; Mizumori and Baker, 2017), and the LHb receives inputs from frontal cortical regions that support behavioral flexibility (Greatrex and Phillipson, 1982; Miller and Cohen, 2001; Monsell, 2003; Kim and Lee, 2012). Thus, a diverse array of LHb afferents provides potential mechanistic substrates for disengaging from ongoing reward-seeking task performance in response to a variety of negatively and positively valenced factors.

Importantly, this model does not exclude a role for phasic LHb activity in signaling negatively valenced information, for which there is abundant evidence (Matsumoto and Hikosaka, 2007, 2009; Lecca et al., 2017). Because of the methodological constraints of photometry and multiunit electrophysiology we cannot rule out the possibility that separate populations of LHb neurons contribute to phasic and tonic LHb signals, although some overlap is likely given the high penetrance of expression in *Klk8-Cre* (NP171) mice. Likewise, our findings are consistent with prior reports that LHb activity is elevated in depression-like behavioral states (Caldecott-Hazard et al., 1988; Morris et al., 1999; Shumake et al., 2003; Li et al., 2011, 2013; Mirrione et al., 2014; Shabel et al., 2014; Lecca et al., 2016; Proulx et al., 2018; Yang et al., 2018; Andalman et al., 2019), which are associated with a reduction in reward-seeking behavior (Willner et al., 1992). Our results complement these observations and suggest a framework in which tonically elevated LHb neural activity not only supports disengagement from reward-seeking behavior in response to negatively-valenced factors such as chronic stress, but also in response to positively-valenced factors such as the satisfaction of homeostatic needs.

An interesting observation was the absence of tonic engagement-related signals in VTA DA neurons, particularly given the strength of the LHb projection to the RMTg and VTA (Herkenham and Nauta, 1979; Araki et al., 1988; Jhou et al., 2009; Gonçalves et al., 2011). Previous reports have shown that baseline VTA DA neural activity remains stable through changing reward and punishment schedules (Cohen et al., 2015; Mohebi et al., 2019), findings that may be related to this observation. The LHb is known to contribute to phasic suppression of VTA DA neural activity – single-pulse electrical stimulation of the LHb transiently inhibits VTA DA neurons (Christoph et al., 1986; Matsumoto and Hikosaka, 2007; Ji and Shepard, 2007), and LHb lesions reduce VTA DA neuron inhibition upon reward omission (Tian and Uchida, 2015) – but the role of tonic LHb activity in regulating VTA DA neural activity is less well understood. One hypothesis is that tonic LHb signals related to sustained disengagement are not, in fact, conveyed to VTA DA neurons, but are instead transmitted from the LHb to the raphe nuclei (Proulx et al., 2014). Indeed, tonic activity in dorsal raphe serotonin neurons reflects recently experienced rewards and punishments (Nakamura et al., 2008; Bromberg-Martin et al., 2010; Cohen et al., 2015). However, chemogenetic inhibition of the LHb-dorsal raphe projection actually reduces perseverative reward seeking (Coffey et al., 2020). The LHb also sends a major projection to the median raphe which may have a distinct function (Proulx et al., 2014).

A second possibility is that tonic LHb activity may regulate the gain of phasic VTA DA reward-related responses, rather than tonic activity in VTA DA neurons. The LHb inhibits VTA DA neurons by exciting GABA neurons in the RMTg and VTA, but it also sends a direct glutamatergic projection to VTA DA neurons (Omelchenko et al., 2009; Brinschwitz et al., 2010), and there is ultrastructural evidence for an approximately equal fraction of LHb synapses on VTA DA and

GABA neurons (Omelchenko et al., 2009). In cortical circuits, balanced changes in background excitation and inhibition have been proposed as a mechanism for regulating the gain of stimulus-driven neural responses independently of spontaneous firing rate (Chance et al., 2002; Ferguson and Cardin, 2020), and we speculate that a similar mechanism may be at work in the VTA. Consistent with this idea, we observed diminished phasic reward-related responses in VTA DA neurons during disengaged states, when LHB activity is tonically elevated, and prior work has shown a similar regulation of VTA DA phasic neural activity by satiety (Bassareo and Chiara, 1999; Branch et al., 2013; Papageorgiou et al., 2016) which may be related.

Taken together, our results support the hypothesis that a key role of tonic LHB neural activity is to regulate engagement in reward-seeking behavior in response to both negative and positive factors, and add to mounting evidence that the functional role of the LHB extends beyond the processing of negative information. This broadening of our understanding of LHB function will be helpful for assessing the potential and limitations of the LHB as a therapeutic entry point for mood, anxiety, and motivational disorders, and provides a conceptual framework of potential utility for deepening our understanding of the functional roles of downstream neuromodulatory circuits.

## ACKNOWLEDGMENTS

We thank B.J. Sleezer, I.T. Ellwood, J.R. Fetcho, J.H. Goldberg, H.K. Reeve, A. Guru, C. Seo, E.L. Troconis, Y.Y. Ho, W. Gu, and Y. Baumel for helpful discussions; B.J. Sleezer and A.K. Recknagel for expert technical and analytical assistance; and the Warden laboratory and Cornell Neurobiology and Behavior for training and support. This work was supported by funding from the National Institutes of Health (R01 MH127510 and DP2 MH109982 to M.R.W.), the New York Stem Cell Foundation (M.R.W.), the Alfred P. Sloan Foundation (M.R.W.), the Whitehall Foundation (M.R.W.), the Brain and Behavior Research Foundation (D.A.B., M.R.W.), the Mong Family Foundation (R.J.P., D.A.B.), and Cornell University.

## AUTHOR CONTRIBUTIONS

Conceptualization, R.J.P., D.A.B., and M.R.W.; Methodology, R.J.P., D.A.B., R.B.E., and M.R.W.; Software, R.J.P., D.A.B., and R.B.E.; Investigation, R.J.P., D.A.B., V.L., K.H., and M.R.W.; Formal Analysis, R.J.P.; Writing – Original Draft, R.J.P. and M.R.W.; Writing – Review & Editing, all authors; Supervision, M.R.W.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

- 5 Alhadeff, A.L., Su, Z., Hernandez, E., Klima, M.L., Phillips, S.Z., Holland, R.A., Guo, C., Hantman, A.W., De Jonghe, B.C., and Betley, J.N. (2018). A Neural Circuit for the Suppression of Pain by a Competing Need State. *Cell* 173, 140–152. <https://doi.org/10.1016/j.cell.2018.02.057>.
- 10 Allen, W.E., Chen, M.Z., Pichamoorthy, N., Tien, R.H., Pachitariu, M., Luo, L., and Deisseroth, K. (2019). Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science* 364, 253. <https://doi.org/10.1126/science.aav3932>.
- 15 Andelman, A.S., Burns, V.M., Lovett-Barron, M., Broxton, M., Poole, B., Yang, S.J., Grosenick, L., Lerner, T.N., Chen, R., Benster, T., et al. (2019). Neuronal Dynamics Regulating Brain and Behavioral State Transitions. *Cell* 177, 970–985. <https://doi.org/10.1016/j.cell.2019.02.037>.
- 20 Aponte, Y., Atasoy, D., and Sternson, S.M. (2011). AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat. Neurosci.* 14, 351–355. <https://doi.org/10.1038/nn.2739>.
- Araki, M., McGeer, P.L., and Kimura, H. (1988). The efferent projections of the rat lateral habenular nucleus revealed by the PHA-L anterograde tracing method. *Brain Res.* 441, 319–330. [https://doi.org/10.1016/0006-8993\(88\)91410-2](https://doi.org/10.1016/0006-8993(88)91410-2).
- Bäckman, C.M., Malik, N., Zhang, Y., Shan, L., Grinberg, A., Hoffer, B.J., Westphal, H., and Tomac, A.C. (2006). Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. *Genesis* 44, 383–390. <https://doi.org/10.1002/dvg.20228>.
- 25 Baker, P.M., Oh, S.E., Kidder, K.S., and Mizumori, S.J.Y. (2015). Ongoing behavioral state information signaled in the lateral habenula guides choice flexibility in freely moving rats. *Front. Behav. Neurosci.* 9, 295. <https://doi.org/10.3389/fnbeh.2015.00295>.
- 30 Barker, D.J., Miranda-Barrientos, J., Zhang, S., Root, D.H., Wang, H.L., Liu, B., Calipari, E.S., and Morales, M. (2017). Lateral Preoptic Control of the Lateral Habenula through Convergent Glutamate and GABA Transmission. *Cell Rep.* 21, 1757–1769. <https://doi.org/10.1016/j.celrep.2017.10.066>.
- Barrot, M., Sesack, S.R., Georges, F., Pistis, M., Hong, S., and Jhou, T.C. (2012). Braking Dopamine Systems: A New GABA Master Structure for Mesolimbic and Nigrostriatal Functions. *J. Neurosci.* 32, 14094–14101. <https://doi.org/10.1523/JNEUROSCI.3370-12.2012>.
- 35 Bassareo, V., and Chiara, G.D. (1999). Modulation of feeding-induced activation of mesolimbic dopamine transmission by appetitive stimuli and its relation to motivational state. *Eur. J. Neurosci.* 11, 4389–4397. <https://doi.org/10.1046/j.1460-9568.1999.00843.x>.
- Benelam, B. (2009). Satiation, satiety and their effects on eating behaviour. *Nutr. Bull.* 34, 126–173. <https://doi.org/10.1111/j.1467-3010.2009.01753.x>.

- Betley, J.N., Xu, S., Cao, Z.F.H., Gong, R., Magnus, C.J., Yu, Y., and Sternson, S.M. (2015). Neurons for hunger and thirst transmit a negative-valence teaching signal. *Nature* 521, 180–185. <https://doi.org/10.1038/nature14416>.
- 5 Bianco, I.H., and Wilson, S.W. (2009). The habenular nuclei: a conserved asymmetric relay station in the vertebrate brain. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1005–1020. <https://doi.org/10.1098/rstb.2008.0213>.
- Boureau, Y.-L., Sokol-Hessner, P., and Daw, N.D. (2015). Deciding How To Decide: Self-Control and Meta-Decision Making. *Trends Cogn. Sci.* 19, 700–710. <https://doi.org/10.1016/j.tics.2015.08.013>.
- 10 Branch, S.Y., Goertz, R.B., Sharpe, A.L., Pierce, J., Roy, S., Ko, D., Paladini, C.A., and Beckstead, M.J. (2013). Food Restriction Increases Glutamate Receptor-Mediated Burst Firing of Dopamine Neurons. *J. Neurosci.* 33, 13861–13872. <https://doi.org/10.1523/JNEUROSCI.5099-12.2013>.
- 15 Brinschwitz, K., Dittgen, A., Madai, V.I., Lommel, R., Geisler, S., and Veh, R.W. (2010). Glutamatergic axons from the lateral habenula mainly terminate on GABAergic neurons of the ventral midbrain. *Neuroscience* 168, 463–476. <https://doi.org/10.1016/j.neuroscience.2010.03.050>.
- Bromberg-Martin, E.S., and Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* 14, 1209–1216. <https://doi.org/10.1038/nn.2902>.
- 20 Bromberg-Martin, E.S., Hikosaka, O., and Nakamura, K. (2010). Coding of Task Reward Value in the Dorsal Raphe Nucleus. *J. Neurosci.* 30, 6262–6272. <https://doi.org/10.1523/JNEUROSCI.0015-10.2010>.
- 25 Caldecott-Hazard, S., Mazziotta, J., and Phelps, M. (1988). Cerebral correlates of depressed behavior in rats, visualized using <sup>14</sup>C-2-deoxyglucose autoradiography. *J. Neurosci.* 8, 1951–1961. .
- Chance, F.S., Abbott, L.F., and Reyes, A.D. (2002). Gain Modulation from Background Synaptic Input. *Neuron* 35, 773–782. [https://doi.org/10.1016/S0896-6273\(02\)00820-6](https://doi.org/10.1016/S0896-6273(02)00820-6).
- Charnov, E.L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* 9, 129–136. [https://doi.org/10.1016/0040-5809\(76\)90040-X](https://doi.org/10.1016/0040-5809(76)90040-X).
- 30 Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R. a., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300. <https://doi.org/10.1038/nature12354>.
- 35 Chen, Y., Lin, Y.C., Kuo, T.W., and Knight, Z.A. (2015). Sensory Detection of Food Rapidly Modulates Arcuate Feeding Circuits. *Cell* 160, 829–841. <https://doi.org/10.1016/j.cell.2015.01.033>.
- Christoph, G.R., Leonzio, R.J., and Wilcox, K.S. (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J. Neurosci.* 6, 613–619. <https://doi.org/10.1523/JNEUROSCI.06-03-00613.1986>.

- Coffey, K.R., Marx, R.G., Vo, E.K., Nair, S.G., and Neumaier, J.F. (2020). Chemogenetic inhibition of lateral habenula projections to the dorsal raphe nucleus reduces passive coping and perseverative reward seeking in rats. *Neuropsychopharmacology* 45, 1115–1124. <https://doi.org/10.1038/s41386-020-0616-0>.
- 5 Cohen, J.D., McClure, S.M., and Yu, A.J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 933–942. <https://doi.org/10.1098/rstb.2007.2098>.
- Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88. <https://doi.org/10.1038/nature10754>.
- 10 Cohen, J.Y., Amoroso, M.W., and Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *ELife* 4, e06346. <https://doi.org/10.7554/eLife.06346>.
- Cooper, J.O., Heron, T.E., and Heward, W.L. (1987). *Applied Behavior Analysis* (Columbus, OH: Merrill).
- 15 Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* 494, 238–242. <https://doi.org/10.1038/nature11846>.
- Dolan, R.J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. <https://doi.org/10.1016/j.neuron.2013.09.007>.
- 20 Ebitz, R.B., Albarran, E., and Moore, T. (2018). Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex. *Neuron* 97, 450–461. <https://doi.org/10.1016/j.neuron.2017.12.007>.
- Ebitz, R.B., Sleezer, B.J., Jedema, H.P., Bradberry, C.W., and Hayden, B.Y. (2019). Tonic exploration governs both flexibility and lapses. *PLOS Comput. Biol.* 15, e1007475. <https://doi.org/10.1371/journal.pcbi.1007475>.
- 25 Evans, D.A., Stempel, A.V., Vale, R., Ruehle, S., Lefler, Y., and Branco, T. (2018). A synaptic threshold mechanism for computing escape decisions. *Nature* 558, 590–594. <https://doi.org/10.1038/s41586-018-0244-6>.
- Ferguson, K.A., and Cardin, J.A. (2020). Mechanisms underlying gain modulation in the cortex. *Nat. Rev. Neurosci.* 21, 80–92. <https://doi.org/10.1038/s41583-019-0253-y>.
- 30 Ferster, C.B., and Skinner, B.F. (1957). *Schedules of Reinforcement* (New York: Appleton-Century-Crofts).
- Flanigan, M.E., Aleyasin, H., Li, L., Burnett, C.J., Chan, K.L., LeClair, K.B., Lucas, E.K., Matikainen-Ankney, B., Durand-de Cuttoli, R., Takahashi, A., et al. (2020). Orexin signaling in GABAergic lateral habenula neurons modulates aggressive behavior in male mice. *Nat. Neurosci.* 23, 638–650. <https://doi.org/10.1038/s41593-020-0617-7>.
- 35

- Flavell, S.W., Pokala, N., Macosko, E.Z., Albrecht, D.R., Larsch, J., and Bargmann, C.I. (2013). Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in *C. elegans*. *Cell* 154, 1023–1035. <https://doi.org/10.1016/j.cell.2013.08.001>.
- 5 Friedman, A., Lax, E., Dikshtein, Y., Abraham, L., Flaumenhaft, Y., Sudai, E., Ben-Tzion, M., and Yadid, G. (2011). Electrical stimulation of the lateral habenula produces an inhibitory effect on sucrose self-administration. *Neuropharmacology* 60, 381–387. <https://doi.org/10.1016/j.neuropharm.2010.10.006>.
- 10 Garfield, A.S., Li, C., Madara, J.C., Shah, B.P., Webber, E., Steger, J.S., Campbell, J.N., Gavrilova, O., Lee, C.E., Olson, D.P., et al. (2015). A neural basis for melanocortin-4 receptor–regulated appetite. *Nat. Neurosci.* 18, 863–871. <https://doi.org/10.1038/nn.4011>.
- GENSAT The Gene Expression Nervous System Atlas (GENSAT) Project, NINDS Contracts N01NS02331 & HHSN271200723701C to The Rockefeller University (New York, NY).
- 15 Gonçalves, L., Sego, C., and Metzger, M. (2011). Differential projections from the lateral habenula to the rostromedial tegmental nucleus and ventral tegmental area in the rat. *J. Comp. Neurol.* 1300, 1278–1300. <https://doi.org/10.1002/cne.22787>.
- Greatrex, R.M., and Phillipson, O.T. (1982). Demonstration of synaptic input from prefrontal cortex to the habenula in the rat. *Brain Res.* 238, 192–197. [https://doi.org/10.1016/0006-8993\(82\)90782-X](https://doi.org/10.1016/0006-8993(82)90782-X).
- 20 Gunaydin, L.A., Grosenick, L., Finkelstein, J.C., Kauvar, I.V., Fenno, L.E., Adhikari, A., Lammel, S., Mirzabekov, J.J., Airan, R.D., Zalocusky, K.A., et al. (2014). Natural neural projection dynamics underlying social behavior. *Cell* 157, 1535–1551. <https://doi.org/10.1016/j.cell.2014.05.017>.
- 25 Guru, A., Seo, C., Post, R.J., Kullakanda, D.S., Schaffer, J.A., and Warden, M.R. (2020). Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. *BioRxiv* 2020.05.21.108886. <https://doi.org/10.1101/2020.05.21.108886>.
- Harrington, M.E. (2012). Neurobiological studies of fatigue. *Prog. Neurobiol.* 99, 93–105. <https://doi.org/10.1016/j.pneurobio.2012.07.004>.
- 30 Herkenham, M., and Nauta, W.J. (1977). Afferent connections of the habenular nuclei in the rat. A horseradish peroxidase study, with a note on the fiber-of-passage problem. *J. Comp. Neurol.* 173, 123–145. <https://doi.org/10.1002/cne.901730107>.
- Herkenham, M., and Nauta, W.J. (1979). Efferent connections of the habenular nuclei in the rat. *J. Comp. Neurol.* 187, 19–47. <https://doi.org/10.1002/cne.901870103>.
- Hikosaka, O. (2010). The habenula: from stress evasion to value-based decision-making. *Nat. Rev. Neurosci.* 11, 503–513. <https://doi.org/10.1038/nrn2866>.
- 35 Hills, T.T., Todd, P.M., Lazer, D., Redish, A.D., Couzin, I.D., Bateson, M., Cools, R., Dukas, R., Giraldeau, L.A., Macy, M.W., et al. (2015). Exploration versus exploitation in space, mind, and society. *Trends Cogn. Sci.* 19, 46–54. <https://doi.org/10.1016/j.tics.2014.10.004>.

- Hong, S., and Hikosaka, O. (2008). The Globus Pallidus Sends Reward-Related Signals to the Lateral Habenula. *Neuron* 60, 720–729. <https://doi.org/10.1016/j.neuron.2008.09.035>.
- 5 Hong, S., Jhou, T.C., Smith, M., Saleem, K.S., 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. <https://doi.org/10.1523/JNEUROSCI.1384-11.2011>.
- Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500, 575–579. <https://doi.org/10.1038/nature12475>.
- 10 Hu, H., Cui, Y., and Yang, Y. (2020). Circuits and functions of the lateral habenula in health and in disease. *Nat. Rev. Neurosci.* 21, 277–295. <https://doi.org/10.1038/s41583-020-0292-4>.
- Huang, L., Yuan, T., Tan, M., Xi, Y., Hu, Y., Tao, Q., Zhao, Z., Zheng, J., Han, Y., Xu, F., et al. (2017). A retinorecipient projection regulates serotonergic activity and looming-evoked defensive behaviour. *Nat. Commun.* 8, 14908. <https://doi.org/10.1038/ncomms14908>.
- 15 Jhou, T.C., Geisler, S., Marinelli, M., Degarmo, B.A., and Zahm, D.S. (2009). 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. <https://doi.org/10.1002/cne.21891>.
- 20 Ji, H., and Shepard, P.D. (2007). Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *J. Neurosci.* 27, 6923–6930. <https://doi.org/10.1523/JNEUROSCI.0958-07.2007>.
- Kim, U., and Lee, T. (2012). Topography of descending projections from anterior insular and medial prefrontal regions to the lateral habenula of the epithalamus in the rat. *Eur. J. Neurosci.* 35, 1253–1269. <https://doi.org/10.1111/j.1460-9568.2012.08030.x>.
- 25 Kim, H.R., Malik, A.N., Mikhael, J.G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S.J., et al. (2020). A Unified Framework for Dopamine Signals across Timescales. *Cell* 183, 1600–1616.e25. <https://doi.org/10.1016/j.cell.2020.11.013>.
- Kurzban, R., Duckworth, A., Kable, J.W., and Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behav. Brain Sci.* 36, 661–679. <https://doi.org/10.1017/S0140525X12003196>.
- 30 Lalive, A.L., Congiu, M., Clerke, J.A., Tchenio, A., Ge, Y., and Mameli, M. (2021). Synaptic inhibition in the lateral habenula shapes reward anticipation. *BioRxiv* <https://doi.org/10.1101/2021.01.21.427572>.
- 35 Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K., and Malenka, R.C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature* 491, 212–217. <https://doi.org/10.1038/nature11527>.
- Lecca, S., Pelosi, A., Tchenio, A., Moutkine, I., Lujan, R., Hervé, D., and Mameli, M. (2016). Rescue of GABAB and GIRK function in the lateral habenula by protein phosphatase 2A



inhibition ameliorates depression-like phenotypes in mice. *Nat. Med.* 22, 254–261. <https://doi.org/10.1038/nm.4037>.

5 Lecca, S., Meye, F.J., Truscel, M., Tchenio, A., Harris, J., Schwarz, M.K., Burdakov, D., Georges, F., and Mameli, M. (2017). Aversive stimuli drive hypothalamus-to-habenula excitation to promote escape behavior. *ELife* 6, e30697. <https://doi.org/10.7554/eLife.30697>.

Leitenberg, H. (1965). Is time-out from positive reinforcement an aversive event? A review of the experimental evidence. *Psychol. Bull.* 64, 428–441. <http://dx.doi.org/10.1037/h0022657>.

10 Li, B., Piriz, J., Mirrione, M., Chung, C., Proulx, C.D., Schulz, D., Henn, F., and Malinow, R. (2011). Synaptic potentiation onto habenula neurons in the learned helplessness model of depression. *Nature* 470, 535–539. <https://doi.org/10.1038/nature09742>.

Li, K., Zhou, T., Liao, L., Yang, Z., Wong, C., Henn, F., Malinow, R., Yates, J.R., and Hu, H. (2013).  $\beta$ CaMKII in lateral habenula mediates core symptoms of depression. *Science* 341, 1016–1020. <https://doi.org/10.1126/science.1240729>.

15 Marques, J.C., Li, M., Schaak, D., Robson, D.N., and Li, J.M. (2020). Internal state dynamics shape brainwide activity and foraging behaviour. *Nature* 577, 239–243. <https://doi.org/10.1038/s41586-019-1858-z>.

Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111–1115. <https://doi.org/10.1038/nature05860>.

20 Matsumoto, M., and Hikosaka, O. (2009). Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci.* 12, 77–84. <https://doi.org/10.1038/nn.2233>.

McGinty, V.B., Lardeux, S., Taha, S.A., Kim, J.J., and Nicola, S.M. (2013). Invigoration of Reward Seeking by Cue and Proximity Encoding in the Nucleus Accumbens. *Neuron* 78, 910–922. <https://doi.org/10.1016/j.neuron.2013.04.010>.

25 Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>.

Mirrione, M., Schulz, D., Lapidus, K., Zhang, S., Goodman, W., and Henn, F. (2014). Increased metabolic activity in the septum and habenula during stress is linked to subsequent expression of learned helplessness behavior. *Front. Hum. Neurosci.* 8, 29. <https://doi.org/10.3389/fnhum.2014.00029>.

30 Mizumori, S.J.Y., and Baker, P.M. (2017). The Lateral Habenula and Adaptive Behaviors. *Trends Neurosci.* 40, 481–493. <https://doi.org/10.1016/j.tins.2017.06.001>.

Mohebi, A., Pettibone, J.R., Hamid, A.A., Wong, J.M.T., Vinson, L.T., Patriarchi, T., Tian, L., Kennedy, R.T., and Berke, J.D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature* 570, 65–70. <https://doi.org/10.1038/s41586-019-1235-y>.

35 Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7).

- Morris, J.S., Smith, K.A., Cowen, P.J., Friston, K.J., and Dolan, R.J. (1999). Covariation of activity in habenula and dorsal raphe nuclei following tryptophan depletion. *NeuroImage* 10, 163–172. <https://doi.org/10.1006/nimg.1999.0455>.
- 5 Mullen, R.J., Buck, C.R., and Smith, A.M. (1992). NeuN, a neuronal specific nuclear protein in vertebrates. *Development* 116, 201–211. .
- Murphy, K.P. (2012). *Machine learning: a probabilistic perspective* (Cambridge, MA: MIT Press).
- Nakamura, K., Matsumoto, M., and Hikosaka, O. (2008). Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J. Neurosci.* 28, 5331–5343. <https://doi.org/10.1523/JNEUROSCI.0021-08.2008>.
- 10 Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* 191, 507–520. <https://doi.org/10.1007/s00213-006-0502-4>.
- Omelchenko, N., Bell, R., and Sesack, S.R. (2009). Lateral habenula projections to dopamine and GABA neurons in the rat ventral tegmental area. *Eur. J. Neurosci.* 30, 1239–1250. <https://doi.org/10.1111/j.1460-9568.2009.06924.x>.
- 15 Papageorgiou, G.K., Baudonnat, M., Cucca, F., and Walton, M.E. (2016). Mesolimbic Dopamine Encodes Prediction Errors in a State-Dependent Manner. *Cell Rep.* 15, 221–228. <https://doi.org/10.1016/j.celrep.2016.03.031>.
- Papini, M.R., and Dudley, R.T. (1997). Consequences of Surprising Reward Omissions. *Rev. Gen. Psychol.* 1, 175–197. <https://doi.org/10.1037/1089-2680.1.2.175>.
- 20 Paul, M.J., Indic, P., and Schwartz, W.J. (2011). A role for the habenula in the regulation of locomotor activity cycles. *Eur. J. Neurosci.* 34, 478–488. <https://doi.org/10.1111/j.1460-9568.2011.07762.x>.
- Proulx, C.D., Hikosaka, O., and Malinow, R. (2014). Reward processing by the lateral habenula in normal and depressive behaviors. *Nat. Neurosci.* 17, 1146–1152. <https://doi.org/10.1038/nn.3779>.
- 25 Proulx, C.D., Aronson, S., Milivojevic, D., Molina, C., Loi, A., Monk, B., Shabel, S.J., and Malinow, R. (2018). A neural pathway controlling motivation to exert effort. *Proc. Natl. Acad. Sci.* 115, 5792–5797. <https://doi.org/10.1073/pnas.1801837115>.
- Salamone, J.D., and Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. *Neuron* 76, 470–485. <https://doi.org/10.1016/j.neuron.2012.10.021>.
- 30 Schéle, E., Cook, C., Le May, M., Bake, T., Luckman, S.M., and Dickson, S.L. (2017). Central administration of ghrelin induces conditioned avoidance in rodents. *Eur. Neuropsychopharmacol.* 27, 809–815. <https://doi.org/10.1016/j.euroneuro.2017.05.001>.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A Neural Substrate of Prediction and Reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>.
- 35

- Seo, C., Guru, A., Jin, M., Ito, B., Sleezer, B.J., Ho, Y., Wang, E., Boada, C., Krupa, N.A., Kullakanda, D.S., et al. (2019). Intense threat switches dorsal raphe serotonin neurons to a paradoxical operational mode. *Science* 342, aau8722. <https://doi.org/10.1126/science.aau8722>.
- 5 Shabel, S.J., Proulx, C.D., Trias, A., Murphy, R.T., and Malinow, R. (2012). Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin. *Neuron* 74, 475–481. <https://doi.org/10.1016/j.neuron.2012.02.037>.
- Shabel, S.J., Proulx, C.D., Piriz, J., and Malinow, R. (2014). GABA/glutamate co-release controls habenula output and is modified by antidepressant treatment. *Science* 345, 1494–1498. <https://doi.org/10.1126/science.1250469>.
- 10 Shang, C., Liu, Z., Chen, Z., Shi, Y., Wang, Q., Liu, S., Li, D., and Cao, P. (2015). A parvalbumin-positive excitatory visual pathway to trigger fear responses in mice. *Science* 348, 1472–1477. <https://doi.org/10.1126/science.aaa8694>.
- Shumake, J., Edwards, E., and Gonzalez-Lima, F. (2003). Opposite metabolic changes in the habenula and ventral tegmental area of a genetic model of helpless behavior. *Brain Res.* 963, 274–281. [https://doi.org/10.1016/s0006-8993\(02\)04048-9](https://doi.org/10.1016/s0006-8993(02)04048-9).
- 15 Skinner, B.F. (1938). *The behavior of organisms: an experimental analysis* (New York: Appleton-Century).
- Skinner, B.F. (1953). *Science and Human Behavior* (New York: Macmillan).
- Stamatakis, A.M., and Stuber, G.D. (2012). Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nat. Neurosci.* 15, 1105–1107. <https://doi.org/10.1038/nn.3145>.
- 20 Stamatakis, A.M., Van Swieten, M., Basiri, M.L., Blair, G.A., Katak, P., and Stuber, G.D. (2016). Lateral Hypothalamic Area Glutamatergic Neurons and Their Projections to the Lateral Habenula Regulate Feeding and Reward. *J. Neurosci.* 36, 302–311. <https://doi.org/10.1523/JNEUROSCI.1202-15.2016>.
- 25 Stephenson-Jones, M., Yu, K., Ahrens, S., Tucciarone, J.M., Van Huijstee, A.N., Mejia, L.A., Penzo, M.A., Tai, L.H., Wilbrecht, L., and Li, B. (2016). A basal ganglia circuit for evaluating action outcomes. *Nature* 539, 289–293. <https://doi.org/10.1038/nature19845>.
- Stern, S., Kirst, C., and Bargmann, C.I. (2017). Neuromodulatory Control of Long-Term Behavioral Patterns and Individuality across Development. *Cell* 171, 1649–1662. <https://doi.org/10.1016/j.cell.2017.10.041>.
- 30 Stopper, C.M., and Floresco, S.B. (2014). What's better for me? Fundamental role for lateral habenula in promoting subjective decision biases. *Nat. Neurosci.* 17, 33–35. <https://doi.org/10.1038/nn.3587>.
- 35 Thornton, E.W., and Evans, J.C. (1982). The role of habenular nuclei in the selection of behavioral strategies. *Physiol. Psychol.* 10, 361–367. <https://doi.org/10.3758/BF03332965>.

- Tian, J., and Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. *Neuron* 87, 1304–1316. <https://doi.org/10.1016/j.neuron.2015.08.028>.
- 5 Tovote, P., Esposito, M.S., Botta, P., Chaudun, F., Fadok, J.P., Markovic, M., Wolff, S.B.E., Ramakrishnan, C., Fenno, L., Deisseroth, K., et al. (2016). Midbrain circuits for defensive behaviour. *Nature* 534, 206–212. <https://doi.org/10.1038/nature17996>.
- Ullsperger, M., and von Cramon, D.Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J. Neurosci.* 23, 4308–4314. .
- 10 Vong, L., Ye, C., Yang, Z., Choi, B., Chua, S., and Lowell, B.B. (2011). Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* 71, 142–154. <https://doi.org/10.1016/j.neuron.2011.05.028>.
- Vroom, V.H. (1964). *Work and motivation* (New York: Wiley).
- 15 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, e23045. <https://doi.org/10.7554/eLife.23045>.
- Wei, Z., Lin, B.-J., Chen, T.-W., Daie, K., Svoboda, K., and Druckmann, S. (2020). A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology. *PLOS Comput. Biol.* 16, e1008198. <https://doi.org/10.1371/journal.pcbi.1008198>.
- 20 Westbrook, A., and Frank, M. (2018). *Dopamine and proximity in motivation and cognitive control* (Elsevier Ltd).
- Willner, P., Muscat, R., and Papp, M. (1992). Chronic mild stress-induced anhedonia: A realistic animal model of depression. *Neurosci. Biobehav. Rev.* 16, 525–534. [https://doi.org/10.1016/S0149-7634\(05\)80194-0](https://doi.org/10.1016/S0149-7634(05)80194-0).
- 25 Yang, Y., Cui, Y., Sang, K., Dong, Y., Ni, Z., Ma, S., and Hu, H. (2018). Ketamine blocks bursting in the lateral habenula to rapidly relieve depression. *Nature* 554, 317–322. <https://doi.org/10.1038/nature25509>.
- Yetnikoff, L., Cheng, A.Y., Lavezzi, H.N., Parsley, K.P., and Zahm, D.S. (2015). Sources of input to the rostromedial tegmental nucleus, ventral tegmental area, and lateral habenula compared: A study in rat. *J. Comp. Neurol.* 523, 2426–2456. <https://doi.org/10.1002/cne.23797>.
- Zahm, D.S., and Root, D.H. (2017). Review of the cytology and connections of the lateral habenula, an avatar of adaptive behaving. *Pharmacol. Biochem. Behav.* 162, 3–21. <https://doi.org/10.1016/j.pbb.2017.06.004>.
- 35 Zhao, X., Liu, M., and Cang, J. (2014). Visual cortex modulates the magnitude but not the selectivity of looming-evoked responses in the superior colliculus of awake mice. *Neuron* 84, 202–213. <https://doi.org/10.1016/j.neuron.2014.08.037>.

# STAR★Methods

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and Virus Strains</b>		
AAV9-CAG-Flex-GCaMP6s	Penn Vector Core	AV-9-PV2818
AAVDJ-hSyn-DIO-GCaMP6m	Stanford Gene Vector and Virus Core	GVVC-AAV-95
AAV5-CAG-Flex-GCaMP6f	Penn Vector Core	AV-5-PV2816
AAV9-CAG-Flex-GFP	UNC Vector Core	N/A
AAV5-CAG-Flex-GFP	UNC Vector Core	N/A
AAV9-EF1 $\alpha$ -DIO-eNpHR3.0-eYFP	Vector Biolabs	VB4584
<b>Chemicals</b>		
DAPI	Sigma-Aldrich	D9542-5mg
DABCO	Sigma-Aldrich	290734-100ML
<b>Experimental Models: Organisms/Strains</b>		
Mouse: Tg(Klk8-cre)NP171Gsat/Mmucd	MMRRC	RRID: MMRRC_036080-UCD
Mouse: B6.SJL- <i>Slc6a3</i> <sup>tm1.1(cre)Bkmn</sup> /J	The Jackson Laboratory	RRID:IMSR_JAX:006660
Mouse: <i>Slc17a6</i> <sup>tm2(cre)Low</sup> /J	The Jackson Laboratory	RRID:IMSR_JAX:016963
Mouse: C57BL/6J	The Jackson Laboratory	RRID:IMSR_JAX:000664
<b>Software and Algorithms</b>		
MATLAB script for analysis	Brianna J. Sleezer	N/A
Med Associates script for operant conditioning	Dave A. Bulkin, Ryan J. Post	N/A
MATLAB toolbox for multi-unit electrophysiology processing	Daniel N. Hill, Samar B. Mehta, David Kleinfeld	N/A

## 5 RESOURCE AVAILABILITY

### Lead Contact

Further information and requests for resources and reagents should be directed to the Lead Contact, Melissa R. Warden (mrwarden@gmail.com).

### Materials Availability

This study did not generate new reagents.

## Data and Code Availability

Data and source code supporting the current study will be made available upon request.

## 5 EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Animals

10 All procedures conformed to guidelines established by the National Institutes of Health and have been approved by the Cornell University Institutional Animal Care and Use Committee. Male and female mice (postnatal 3-6 months) were used in this study. Mice were group housed in a vivarium on a reversed 12h light/dark cycle. All experiments were conducted during the dark portion of the cycle. Klk8-Cre (NP171) (MMRRC, University of California, Davis, CA), DAT-Cre (The Jackson Laboratory, Bar Harbor, ME), and Vglut2-ires-Cre (The Jackson Laboratory, Bar Harbor, ME) mice were used for photometry experiments. Klk8-Cre (NP171) mice were used for optogenetic inhibition experiments. C57BL/6J mice were used for electrophysiology experiments. All Cre driver lines were fully backcrossed to C57BL/6J mice. Mice were provided with *ad libitum* access to food and water prior to training on the poke-reward task.

## METHOD DETAILS

### 20 Viral Vectors

For photometry experiments, Klk8-Cre (NP171) and Vglut2-ires-Cre mice were injected with AAV9-CAG-Flex-GCaMP6s (Penn Vector Core, Philadelphia, PA). DAT-Cre mice were injected with AAV5-Syn-Flex-GCaMP6f (Penn Vector Core, Philadelphia, PA) or AAV-DJ-EF1 $\alpha$ -DIO-GCaMP6m (Stanford Vector Core, Stanford, CA). For optogenetic experiments, Klk8-Cre (NP171) mice were injected with AAV9-EF1 $\alpha$ -DIO-eNpHR3.0-eYFP (Vector Biolabs, Malvern, PA). For all experiments, control animals were injected with AAV9-CAG-Flex-GFP (UNC Vector Core, Chapel Hill, NC).

### 30 Surgical Procedures

Mice were deeply anesthetized with isoflurane (5%). Fur was trimmed, and mice were placed in a stereotaxic frame (Kopf Instruments, Tujunga, CA) on a heating pad to prevent hypothermia. Isoflurane was delivered at 1-3% throughout surgery; this level was adjusted to maintain a constant surgical plane. Ophthalmic ointment was used to protect the eyes. Lactated ringers (500 ml, subcutaneous) was administered before the start of surgery. A mixture of 0.5% lidocaine and 0.25% bupivacaine (100  $\mu$ l) was injected subdermally along the incision line. The scalp was disinfected with betadine and alcohol. A midline incision exposed the skull, which was thoroughly cleaned, and a craniotomy was made above the LHb or VTA. Virus was targeted to the LHb (-1.80 AP, 0.40 ML, -2.80 and -2.60 DV) or VTA (-3.10 AP, 0.35 ML, -4.60 and -4.30 DV), and slowly pressure injected (100 nl min<sup>-1</sup>) using a 10  $\mu$ l Hamilton syringe (nanofil; WPI, Sarasota,

FL), a 33 gauge beveled needle, and a micro-syringe pump controller (Micro 4; WPI, Sarasota, FL). After each injection, the needle was left in place for 10 minutes and then slowly withdrawn.

For photometry experiments, a total of 600 nl (300 nl at each DV site) of vector was injected and an optical fiber (400  $\mu$ m diameter, 0.48 NA, Doric Lenses, Québec, Canada) was implanted in the right hemisphere. Implants were targeted to LHb (-1.80 AP, 0.40 ML, -2.40 DV) or VTA (-3.10 AP, 0.35 ML, -4.45 DV). For LHb optogenetic inhibition experiments, a total of 1200 nl (300 nl at each DV site in each hemisphere) of vector was injected and an optical fiber (200  $\mu$ m diameter, 0.22 NA, Thorlabs Inc., Newton, NJ) was implanted at a 20 degree angle in each hemisphere (-1.80 AP, +1.23 ML, -2.21 DV). A layer of metabond (Parkell, Inc., Edgewood, NY) and dental acrylic (Lang Dental Manufacturing, Wheeling, IL) was applied to firmly hold the fiber in place, and the surrounding skin was sutured closed. Post-operative buprenorphine (0.05 mg/kg) and carprofen (5 mg/kg) were administered subcutaneously. Virus was allowed to express for a minimum of 6 weeks before behavioral testing.

## Immunohistochemistry

To determine the specificity and penetrance of the Klk8-Cre (NP171) line for LHb neurons, male mice were injected with AAV9-CAG-Flex-GFP in the LHb, as described above. After 4 weeks of expression, mice were perfused with 4% paraformaldehyde in PBS. Brains were extracted, post-fixed in 4% PFA for 24 h, and equilibrated in 30% sucrose in PBS for at least 3 days. 50  $\mu$ m coronal sections were cut on a freezing microtome and stored in cryoprotectant at 4°C. Sections containing the LHb were blocked (10% normal goat serum; ThermoFisher Scientific, Waltham, MA) and incubated with 1:400 anti-NeuN primary antibody (ABN78, EMD Millipore, Darmstadt, DE) and 1:500 Cy3-expressing secondary antibody (AB\_2307443, Jackson ImmunoResearch, West Grove, PA) (Zhong et al., 2017). Sections were incubated in 4',6-diamidino-2-phenylindole (DAPI, 1:50000) and mounted on slides with PVA-DABCO. Images were acquired using a Zeiss LSM800 confocal scanning laser microscope (Zeiss) with a 20X air objective.

## Fiber Photometry

Fiber photometry was performed using a Doric photometry system (Doric Lenses, Québec, Canada). A 490 nm LED was sinusoidally modulated at 211 Hz and passed through a GFP excitation filter. A 405 nm LED was modulated at 531 Hz and passed through a 405 nm bandpass filter. Both light streams were coupled to an optical patch cord (0.48 NA; 400  $\mu$ m core) (Doric Lenses, Quebec, Canada), which was connected to an optical fiber brain implant in each mouse. Emitted fluorescence was collected by the same fiber, passed through a GFP emission filter, and focused onto a photoreceiver.

## Multi-Unit Electrophysiology

To record multi-unit activity (MUA), we implanted male C57BL/6 mice with 16-channel electrode arrays (35  $\mu$ m tungsten electrodes with 200  $\mu$ m electrode spacing, 200  $\mu$ m row spacing, 6 mm in length; Innovative Neurophysiology, Inc., Durham, NC) centered over the right LHb. Target coordinates, relative to bregma, were: AP: -1.20 mm to -2.80 mm (anterior and posterior edges

of the array); ML: 0.30 mm and 0.50 mm (medial and lateral electrode rows, respectively); DV: - 2.55 mm. A ground wire, affixed to the array, was attached to two stainless steel screws placed in the cerebellum. We used a Tucker-Davis Technologies acquisition system and Synapse software to record spike data (Tucker-Davis Technologies, Alachua, FL). Voltage measurements were collected and saved at 24.414 kHz. MUA was extracted from the raw voltage trace through a series of offline processing steps. First, signals were filtered and large artifactual voltage fluctuations were removed from each channel using stationary wavelet decomposition/transform. Next, a common average reference for each array was calculated by taking the sample by sample average of all channels; this global average was then subtracted from the signal on each channel. This method of referencing has been found to outperform alternative referencing methods, such as single best electrode referencing (Ludwig et al., 2009). MUA spiking activity was then extracted using the toolbox UltraMegaSort2000 and a voltage threshold of 2.5 standard deviations above the mean of the voltage trace.

To determine the location of electrode tips, we used an optical clearing technique that allowed us to visualize the entire electrode tract (including electrode tips) for all 16 electrodes in each array. Following collection of MUA data, mice were perfused using saline and 4% paraformaldehyde and decapitated. Skin and other tissues were removed from the head (keeping electrode arrays intact) and skulls (with intact arrays) were drop fixed in 4% paraformaldehyde for 48 to 72 hours at 4°C. Brains were then carefully dissected from the skull and arrays were gently removed. A 1.5 mm thick sagittal slice of brain tissue centered around the electrode array was taken from each brain. Slices were then drop fixed in glutaraldehyde for 24 hours at 4°C, washed in PBS-T for 24 hours, and then placed in 6% SDS at 37°C and checked daily to monitor clearing progress. Slices were typically sufficiently cleared in 6 to 10 days. Slices were then washed in 0.3% Triton-X in PBS at 37°C for 48 hours and subsequently placed in an iodixanol solution composed of 50 g diatrizoic acid, 40 g N-methyl-d-glucamine, 55 g iodixanol, and 0.02% sodium azide per 100 ml water (Murray et al., 2015). Slices were gently swirled daily and monitored for transparency and refraction changes over the course of 2 to 5 days. Samples were then transferred to a fresh iodixanol solution for 24 hours and, finally, mounted in the iodixanol solution between two cover glasses which were separated by 1.5 mm rubber gaskets. Slices were then imaged using a confocal microscope (LSM800, Zeiss). A 647 nm wavelength excitation light was used for imaging. To determine the location of the medial and lateral electrode rows, z-stacks were constructed from optical sections taken in 50 µm increments. Given that electrode arrays extended posterior to LHb, the posterior most 3 to 4 electrodes in both the medial and lateral rows were typically excluded from further analysis based on inspection of the electrode tract locations.

### Optogenetic Inhibition

During behavioral testing, external patch cords (200 µm diameter, 0.22 NA, Doric Lenses, Québec, Canada) were coupled to implanted fiber optic cannulae (CFM22U-20, Thor Labs, NJ, US) with zirconia sleeves. Cannulae were placed above LHb bilaterally as described above. An optical commutator allowed for unrestricted rotation (Doric Lenses, Québec, Canada). Optical inhibition was provided with a 594 nm diode pumped solid state laser (Mambo 100, Colbalt, Solna, SE). Inhibition experiments used 4 mW light (127 mW/mm<sup>2</sup> at the fiber tip). While it has been demonstrated that continuous illumination can cause non-specific effects on physiology and



behavior, 4 mW is within a range that minimizes heating (Yizhar et al., 2011) and has been shown to have no effect on behavior in recent work (Owen et al., 2019). A non-opsin GFP group was included in all optogenetic experiments to control for heating and other potential artifacts (Yizhar et al., 2011), and experiments were designed to avoid prolonged continuous light. Instead of starting illumination at the beginning of the behavioral testing session, we waited until animals spontaneously disengaged and then turned the laser on.

## Behavioral Testing

In all versions of the poke-reward task, mice were first water restricted over the course of three to five days until they reached 80% of their pre-restriction body weight. After restriction, mice were trained to poke their nose into a hole (ENV-313W; Med Associates, VT) on one wall of a 20 x 22 cm operant chamber (ENV-307W-CT) housed in a sound-attenuating box (ENV-022MD) and containing a lickometer (ENV-250B) on the end of the chamber opposite the nose-poke. At the start of each session, a light in the nose-poke hole was illuminated. Upon completion of a successful nose-poke, the light in the nose-poke hole was turned off, a masking soft white noise was turned on, and a water reward was delivered via syringe pump (PHM100) to a fluid port on the opposite wall of the chamber. A nose-poke entry followed by a lick was considered a single trial. Mice were free to run back and forth between the nose-poke hole and the reward spout and complete trials at their own pace. Training continued daily until mice were able to perform 90 or more trials within a 30-minute session across two consecutive daily sessions. Mice were run on one 30-minute session each day, 7 days a week.

*Probabilistic task.* In the probabilistic version of the poke-reward task, each successful nose-poke had a 20% chance of yielding a large reward (20  $\mu$ l of water), a 60% chance of yielding a medium reward (10  $\mu$ l of water) and a 20% chance of yielding no reward (0  $\mu$ l of water). No cues were provided to indicate the reward size or probabilities, and care was taken to avoid sound from the syringe pump upon reward delivery. Throughout the manuscript we refer to these three conditions as 'reward value', rather than reward size, as reward value is defined as the product of reward size and probability (Glimcher, 2005; Matsumoto and Hikosaka, 2009).

*Blocked task.* In the blocked version of the poke-reward task, no reward trials (0  $\mu$ l of water) and reward trials (10  $\mu$ l of water) were grouped into 5-minute long blocks of trials that alternated between reward available and no reward available. Because we wanted to be sure that mice experienced each block, block reward contingencies were only changed once animals completed a successful nose-poke following 5 minutes within a block - this design minimized instances in which mice may not complete any trials during a 5-minute long period and subsequently would not be aware of the block transition. As such, the absolute duration of each block was a minimum of 5 minutes, but varied depending on mouse behavior. All photometry experiments for the poke-reward task were run using the same mice and occurred in the following order: training (5-10 days), probabilistic task (5 days), blocked task (5 days).

*Optogenetic inhibition: Re-engagement task.* Mice performed a task designed to probe the capacity for LHB inhibition to promote task re-engagement after entering a disengaged state (**Figure 6C**). In this task, we provided mice with a medium reward (10 µl water) following each nose poke, consistent with the training paradigm above. We then waited for mice to disengage from the task, defined as 2 minutes after the last nose poke. We then turned the laser on (or began a sham laser off period), and measured the latency for mice to re-engage in the task, defined as a nose poke. We collected behavioral data over the course of four consecutive daily sessions, alternating between sessions with and without laser. The order of sessions with and without laser was counterbalanced across mice. Following the completion of these four sessions, mice performed an additional four sessions of the re-engagement task. In these additional sessions, all task and laser conditions were the same, but no rewards were delivered.

*Optogenetic inhibition: Persistence task.* Mice were run on a task that was designed to probe the capacity for LHB inhibition to promote persistent task engagement (**Figure 6E**). In this task, we provided mice with a medium reward (10 µl water) following each nose poke, and waited for mice to disengage from the task, defined as 2 minutes after the last nose poke. We then waited for mice to re-engage in the task, defined as a nose poke, and then turned the laser on (or began a sham laser period). We terminated the session once mice disengaged from the task once again. We measured persistence as the number of trials (nose pokes) performed during the laser on (or sham laser) period. We collected behavioral data over the course of four consecutive daily sessions, alternating between sessions with and without laser. The order of sessions with and without laser was counterbalanced across mice. Following the completion of these four sessions, mice performed an additional four sessions of the persistence task. In these additional sessions, all task and laser conditions were the same, but no rewards were delivered.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Fiber Photometry

The 405 nm reference channel was fit to the 490 nm channel using linear least squares. Relative fluorescence changes, reported as  $\Delta F/F$ , were calculated using the following equation:

$$\frac{\Delta F}{F_0} = \frac{490 \text{ nm signal-fitted } 405 \text{ nm signal}}{\text{fitted } 490 \text{ nm signal}} \times 100$$

### Identification of State Transitions (Hidden Markov Model)

To identify periods of high and low engagement, we modeled high and low engagement as two latent states that could underlie behavior. Because the level of engagement could be measured in multiple ways (the frequency of licks and/or the frequency of nose pokes), we used a multivariate HMM to simultaneously account for both types of observations. Behavior was coded as binary vectors indicating the presence or absence of each behavior during each second of the task. These emissions were modelled as Poisson random variables whose probability of occurring

was free to differ across latent states. The model was fit via expectation-maximization using the Baum Welch algorithm (Bilmes, 1998; Murphy, 2012), which finds a (possibly local) maxima of the complete-data likelihood. The behavioral data was oversampled relative to the slow changes in engagement of interest here, so we used two methods to highlight the slow changes in engagement states. First, observations were smoothed across neighboring time bins (10 bins) to disrupt the irrelevant local structure that occurred because licks and nose pokes happened at opposite ends of the chamber. Second, we added a regularization term that penalized frequent transitions between states (Montanez et al., 2015). The algorithm was initialized with a random seed once, and the model that maximized the observed (incomplete) data log likelihood was ultimately taken as the best for each session. Finally, we used the Viterbi algorithm to discover the most probable a posteriori sequence of latent states, given the model and behavioral observations (Murphy, 2012).

Although we use a two-state model here, true underlying engagement state likely varies continuously. Our use of a two-state model is a coarse estimate of this state, which we can't observe directly. Here, we are trying to capture the distinction between highly engaged states, where trials are performed rapidly one after another, and low engagement states, where trials are performed at a low rate. We could also have chosen to model 'high', 'medium', and 'low' engagement states, or an even greater number of states, but all these models require assumptions about an underlying state that is not directly observable.

## Statistics

All statistical analyses were performed using custom-written scripts in MATLAB (MathWorks, Natick, MA).

*Epochs.* In all figures, 'first lick-aligned' indicates neural activity aligned to the first spout contact after a poke, whether or not reward was received. The reward epoch is defined as the 1 second following the first spout contact (0 seconds to 1 seconds). The baseline epoch is defined as the 1 second prior to nose poke entry (-1 seconds to 0 seconds). This epoch was used as baseline because there is no inter-trial interval in the poke-reward task.

*Regressions.* Linear regressions in Figures 1 and S1 were performed using baseline-subtracted reward epoch neural data. Regression coefficients were partially standardized by z-scoring reward value. All other linear regressions were performed using full session  $\Delta F/F$  data and binary values from the two-state HMM (high [1] or low [0] engagement) at each time point.

*Significance.* In all figures, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , and \*\*\*\*  $p < 0.0001$ . Effects with  $p < 0.05$  were considered significant. Error bars represent standard error of the mean (SEM).

## METHODS REFERENCES

Bilmes, J.A. (1998). A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models (University of California, Berkeley).

Glimcher, P.W. (2005). Indeterminacy in Brain and Behavior. *Annu. Rev. Psychol.* 56, 25–56.

Ludwig, K.A., Miriani, R.M., Langhals, N.B., Joseph, M.D., Anderson, D.J., and Kipke, D.R. (2009). Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays. *J. Neurophysiol.* 101, 1679–1689.

5 Matsumoto, M., and Hikosaka, O. (2009). Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci.* 12, 77–84.

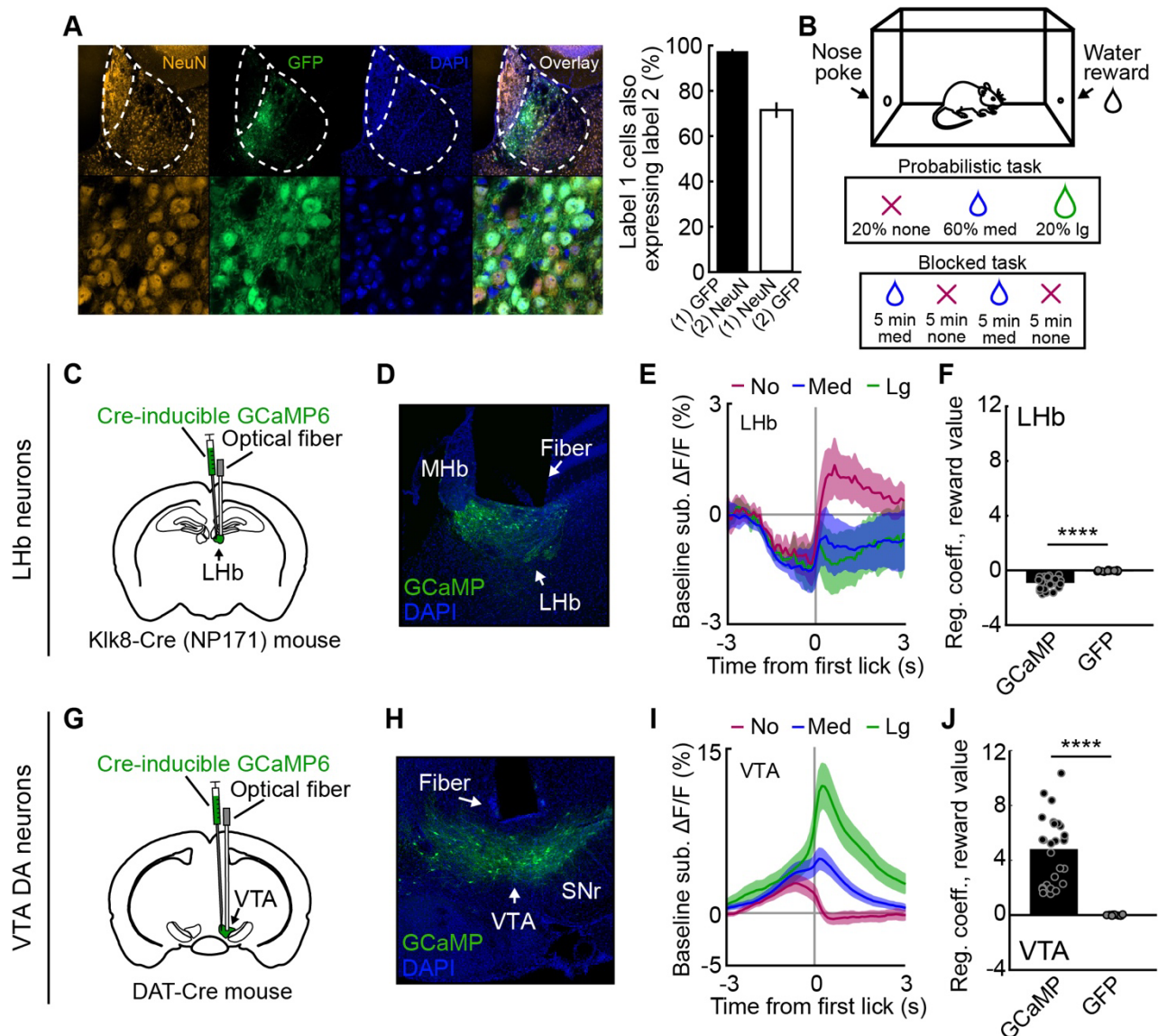
Montanez, G.D., Amizadeh, S., and Laptev, N. (2015). Inertial Hidden Markov Models: Modeling Change in Multivariate Time Series. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Murphy, K.P. (2012). *Machine learning: a probabilistic perspective* (Cambridge, MA: MIT Press).

10 Owen, S.F., Liu, M.H., and Kreitzer, A.C. (2019). Thermal constraints on in vivo optogenetic manipulations. *Nat. Neurosci.* 22, 1061–1065.

Yizhar, O., Fenno, L.E., Davidson, T.J., Mogri, M., and Deisseroth, K. (2011). Optogenetics in Neural Systems. *Neuron* 71, 9–34.

Zhong, P., Vickstrom, C.R., Liu, X., Hu, Y., Yu, L., Yu, H.-G., and Liu, Q. (2017). HCN2 channels in the ventral tegmental area regulate behavioral responses to chronic stress. *ELife* 7, e32420.



**Figure 1. Calcium Dynamics in Genetically Targeted LHB Neurons**

(A) Left: GFP expression and NeuN and DAPI staining in the LHb of a Klk8-Cre (NP171) mouse. Right: Quantification of NeuN and GFP colocalization in Klk8-Cre (NP171) mice ( $n = 365$  cells).

(B) Schematic of probabilistic and blocked poke-reward task designs.

(C) Schematic of LHb viral vector injection and optical fiber placement.

(D) LHb GCaMP6 expression, Klk8-Cre (NP171) mouse.

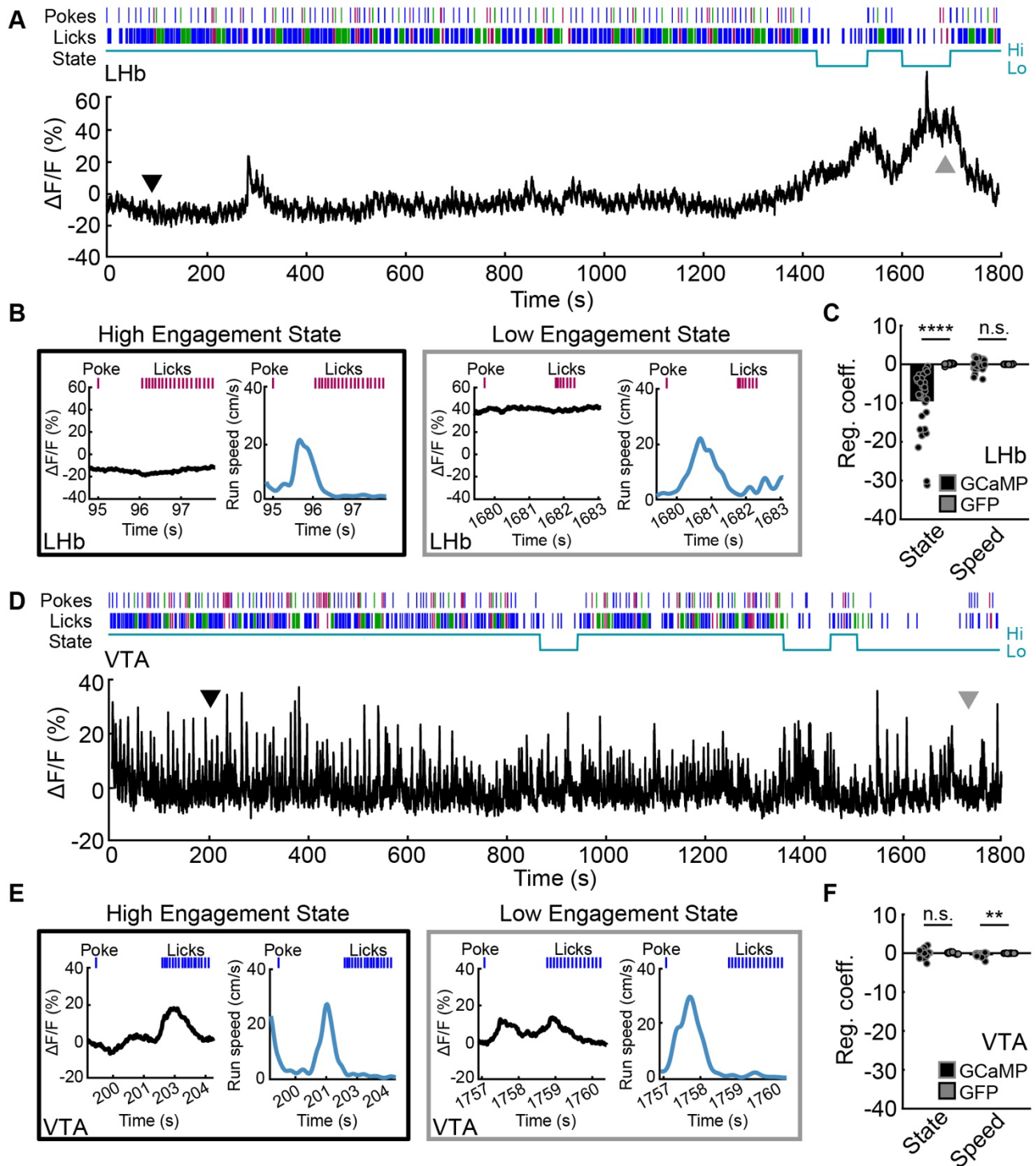
(E) Average baseline-subtracted  $\Delta F/F$  in GCaMP Klk8-Cre (NP171) mice ( $n = 7$ ) during probabilistic task sessions, aligned to first lick and separated by reward value (no reward, red; medium reward, blue; large reward, green).

(F) Average regression coefficients.  $\Delta F/F$  0-1 s following first lick was regressed onto reward value in GCaMP ( $n = 34$  sessions, 7 mice) and GFP ( $n = 15$  sessions, 3 mice) Klk8-Cre (NP171) mice.

(G-J) Same as C-F, but for VTA DA neurons in GCaMP ( $n = 30$  sessions, 6 mice) and GFP ( $n = 13$  sessions, 3 mice) DAT-Cre mice.

\*\*\*\* $p < 0.0001$ , two-sample t-test. Shaded regions indicate SEM.

See also Figure S1.



**Figure 2. Tonically Elevated LHB Activity During Spontaneous Task Disengagement**

(A) Example LHB Kik8-Cre (NP171) photometry during a probabilistic task session. Ticks indicate poke and lick times (no reward, red; medium reward, blue; large reward, green). Engagement state, cyan line. Example trials in B indicated by black and gray arrows.

(B) Example trials from A. Left (black box): trial during high engagement state. Right (gray box): trial during low engagement state.

(C) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state and running speed in GCaMP (n = 34 sessions, 7 mice) and GFP (n = 15 sessions, 3 mice) Klk8-Cre (NP171) mice.

(D) Example VTA DAT-Cre photometry during a probabilistic task session, same conventions as A. Example trials in E indicated by black and gray arrows.

5 (E) Example trials from D, same conventions as B.

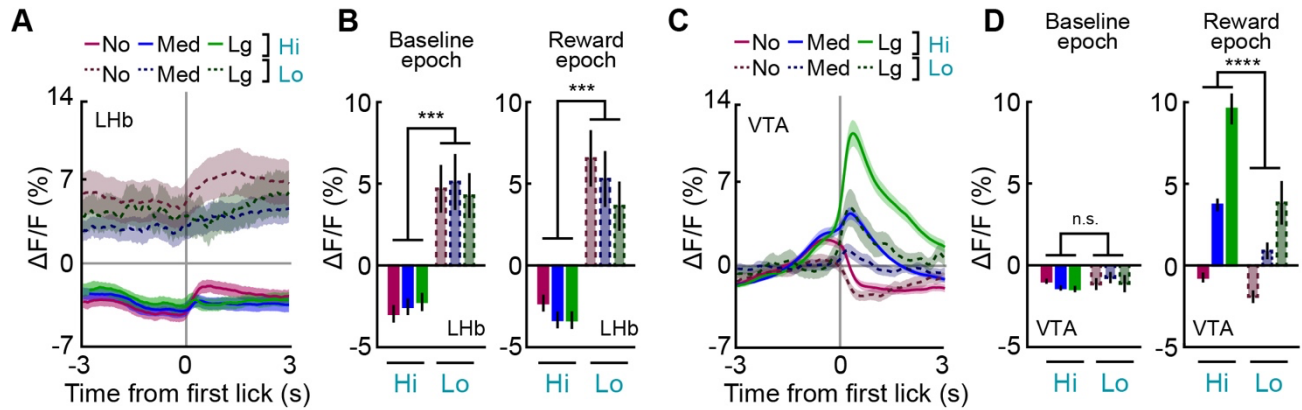
(F) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state and running speed in GCaMP (n = 28 sessions, 6 mice) and GFP (n = 12 sessions, 3 mice) DAT-Cre mice.

\*\*p < 0.01, \*\*\*\*p < 0.0001, two-sample t-test.

See also Figure S2.

10





**Figure 3. Baseline LHB Activity Reflects Task Engagement State**

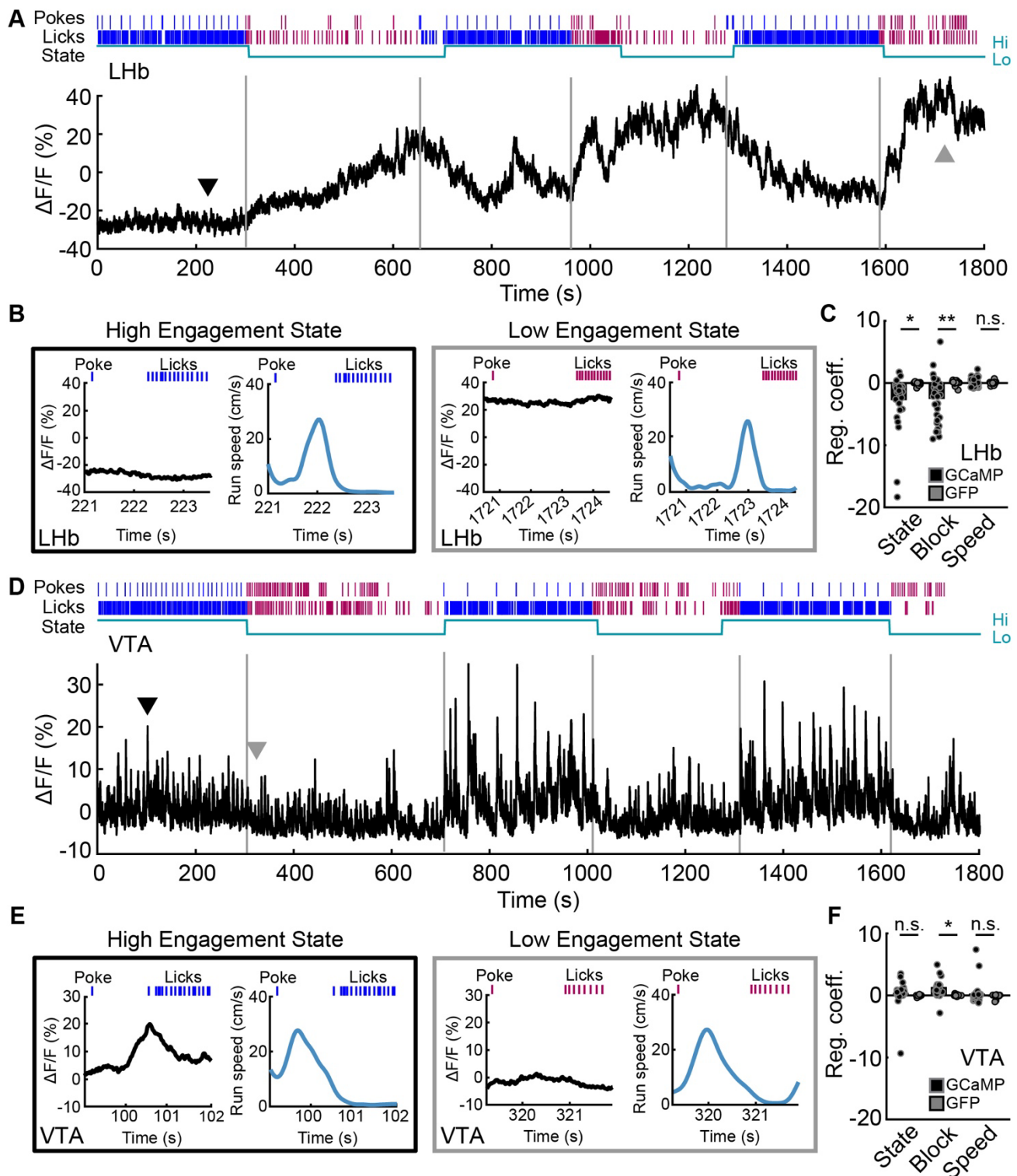
(A) Average  $\Delta F/F$  in GCaMP Kik8-Cre (NP171) mice during probabilistic task sessions ( $n = 34$  sessions, 7 mice), aligned to first lick and separated by engagement state (solid versus dotted lines) and reward value (no reward, red line; medium reward, blue; large reward, green).

(B) Average  $\Delta F/F$  during baseline and reward epochs in high and low engagement states.

(C-D) Same as A-B but for VTA DA neurons in DAT-Cre mice ( $n = 28$  sessions, 6 mice).

\*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , paired samples t-test. Shaded regions indicate SEM.

See also Figure S3.



**Figure 4. Tonic Activity in LHb During Task Disengagement Due to Repeated Reward Omission**

(A) Example LHb Kik8-Cre (NP171) photometry during a blocked task session. Ticks indicate poke and lick times (no reward, red; medium reward, blue). Engagement state, cyan line. Example trials in B indicated by black and gray arrows.

(B) Example trials from A. Left (black box): trial during high engagement state. Right (gray box): trial during low engagement state.

(C) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state, block reward, and running speed in GCaMP (n = 35 sessions, 6 mice) and GFP (n = 15 sessions, 3 mice) Kik8-Cre (NP171) mice.

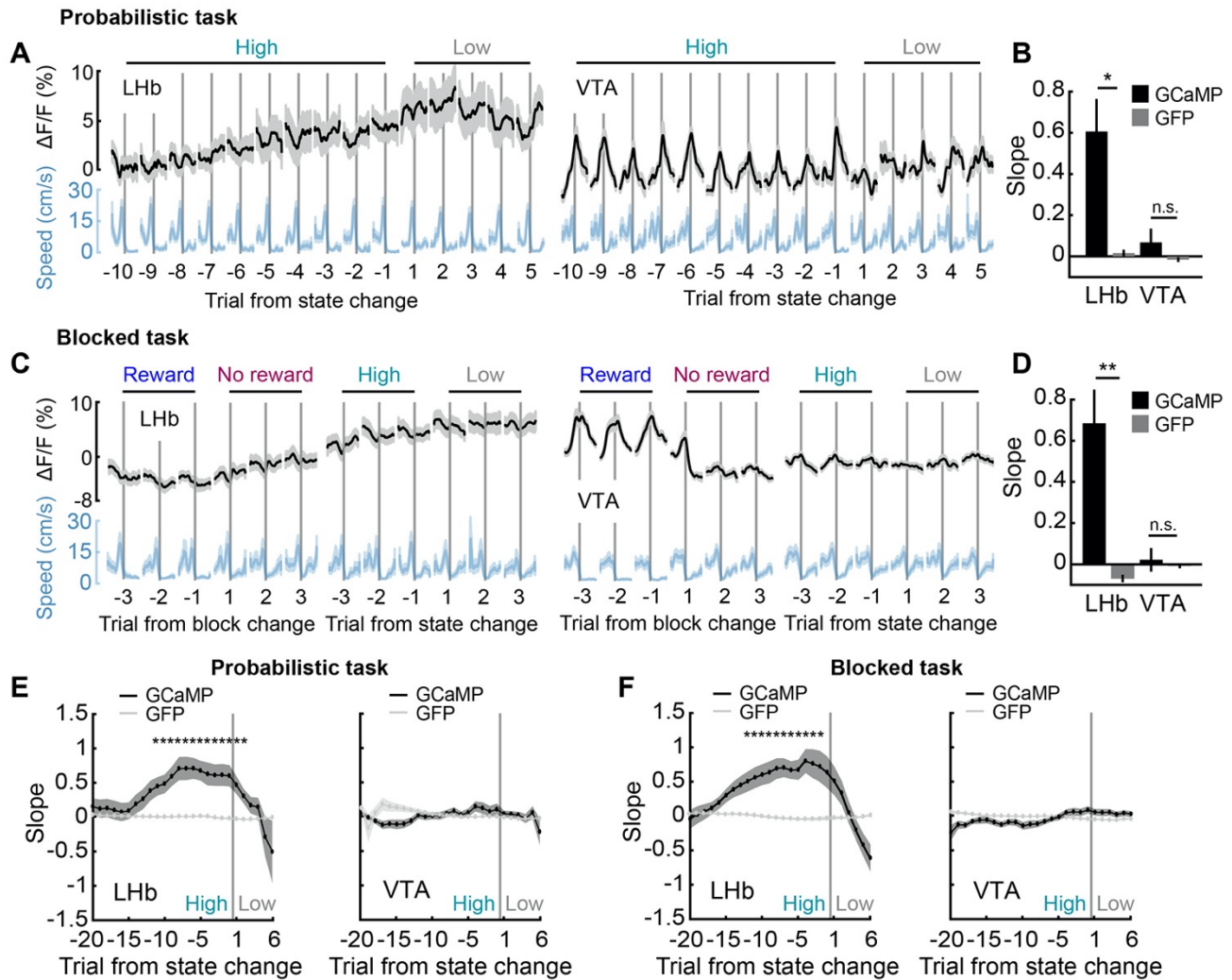
(D) Example VTA DAT-Cre photometry during a blocked task session, same conventions as A. Example trials in E indicated by black and gray arrows.

(E) Example trials from D, same conventions as B.

(F) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state, block reward, and running speed in GCaMP (n = 30 sessions, 6 mice) and GFP (n = 13 sessions, 3 mice) DAT-Cre mice.

\*p < 0.05, \*\*p < 0.01, two-sample t-test.

See also Figure S4.



**Figure 5. Rising LHb Activity Precedes Task Disengagement**

(A) Average first lick-aligned  $\Delta F/F$  for the trials immediately before and after entry into low engagement states, probabilistic task sessions. Left, LHb  $\Delta F/F$  in Kik8-Cre (NP171) mice. Right: VTA DA  $\Delta F/F$ .

(B) Slope of the increase in baseline  $\Delta F/F$  over the 10 trials before entry into low engagement states in LHb Kik8-Cre (NP171) (GCaMP (n = 34 sessions, 7 mice), GFP (n = 15 sessions, 3 mice)) and VTA DAT-Cre (GCaMP (n = 28 sessions, 6 mice), GFP (n = 12 sessions, 3 mice)), probabilistic task sessions.

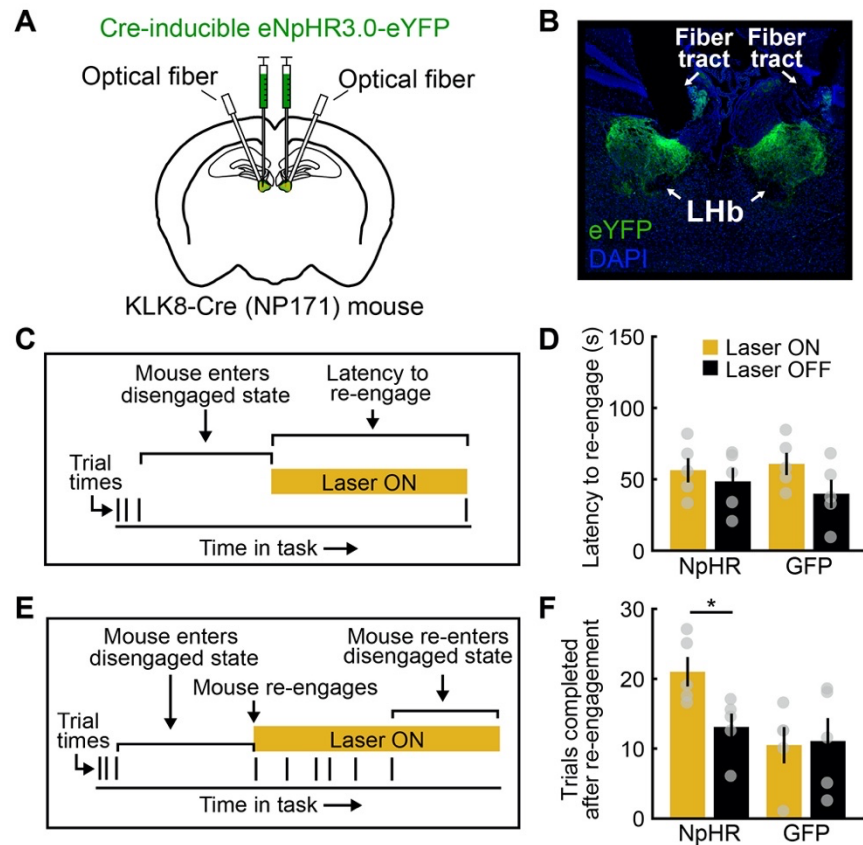
(C) Average first lick-aligned  $\Delta F/F$  for the trials immediately before and after the start of no-reward blocks and entry into low engagement states, blocked task sessions. Left, LHb  $\Delta F/F$  in Kik8-Cre (NP171) mice. Right: VTA DAT-Cre  $\Delta F/F$ .

(D) Slope of the increase in baseline  $\Delta F/F$  over the 10 trials before entry into low engagement states in LHb Kik8-Cre (NP171) (GCaMP (n = 35 sessions, 6 mice), GFP (n = 15 sessions, 3 mice)) and VTA DAT-Cre (GCaMP (n = 30 sessions, 6 mice), GFP (n = 13 sessions, 3 mice)), blocked task sessions.

(E) Slope of the increase in baseline  $\Delta F/F$  over a 10 trial sliding window, probabilistic task.

(F) Slope of the increase in baseline  $\Delta F/F$  over a 10 trial sliding window, blocked task.

\*p < 0.05, \*\*p < 0.01, two-sample t-test. Shaded regions indicate SEM.



**Figure 6. LHb Inhibition Extends Reward-Seeking Behavioral States**

(A) Schematic of LHb viral vector injection and optical fiber placement.

(B) eNpHR3.0-eYFP expression in LHb neurons in a Klk8-Cre (NP171) mouse.

(C) Experimental schematic, task re-engagement.

(D) Average latency to re-engage in task performance. LHb inhibition initiated during disengagement. NpHR (n = 5) and GFP (n = 5) Klk8-Cre (NP171) mice.

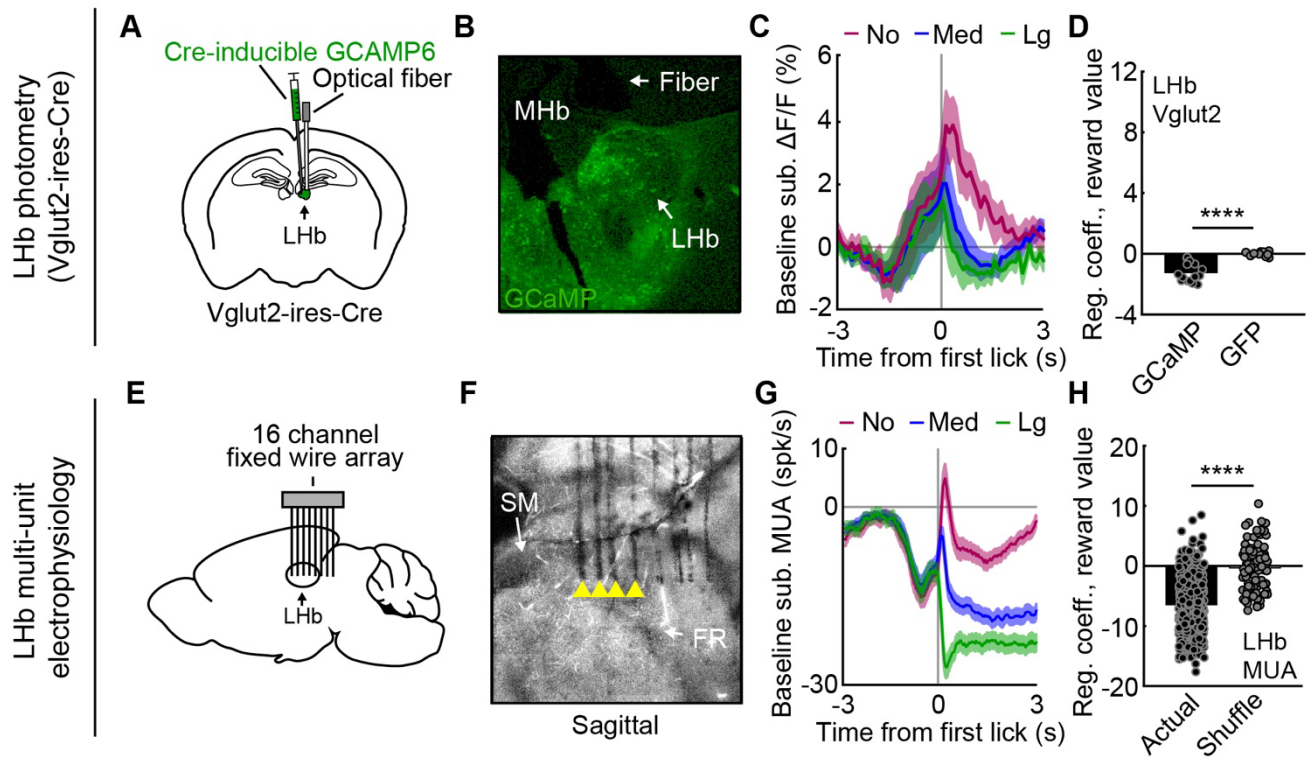
(E) Experimental schematic, task persistence.

(F) Average number of trials completed after task re-engagement. LHb inhibition initiated upon re-engagement. NpHR (n = 5) and GFP (n = 5) Klk8-Cre (NP171) mice.

\*p < 0.05, paired samples t-test. Error bars indicate SEM.

See also Figure S5.

## Supplemental Information



**Figure S1 (Related to Figure 1). Phasic LHB Activity Upon Reward Omission, Vglut2-ires-Cre Photometry and Multi-unit Electrophysiology**

(A) Schematic of LHB viral vector injection and optical fiber placement, Vglut2-ires-Cre mice.

(B) LHB GCaMP6 expression, Vglut2-ires-Cre mouse.

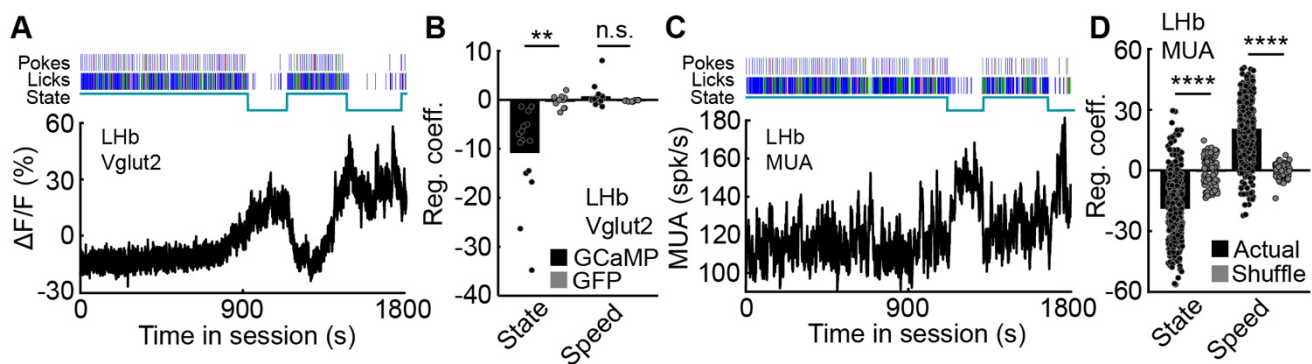
(C) Average baseline-subtracted  $\Delta F/F$  in GCaMP Vglut2-ires-Cre mice ( $n = 15$  sessions, 3 mice) during probabilistic task sessions, aligned to first lick and separated by reward value (no reward, red; medium reward, blue; large reward, green).

(D) Average regression coefficients.  $\Delta F/F$  0-1 s following first lick was regressed onto reward value in GCaMP ( $n = 15$  sessions, 3 mice) and GFP ( $n = 15$  sessions, 3 mice) Vglut2-ires-Cre mice. \*\*\*\* $p < 0.0001$ , two-sample t-test.

(E-H) Same as A-D, but for LHB multi-unit electrophysiology ( $n = 40$  electrodes in 6 mice). SM = stria medullaris, FR = fasciculus retroflexus. \*\*\*\* $p < 0.0001$ , t-test compared to shuffled sample.

Shaded regions indicate SEM.





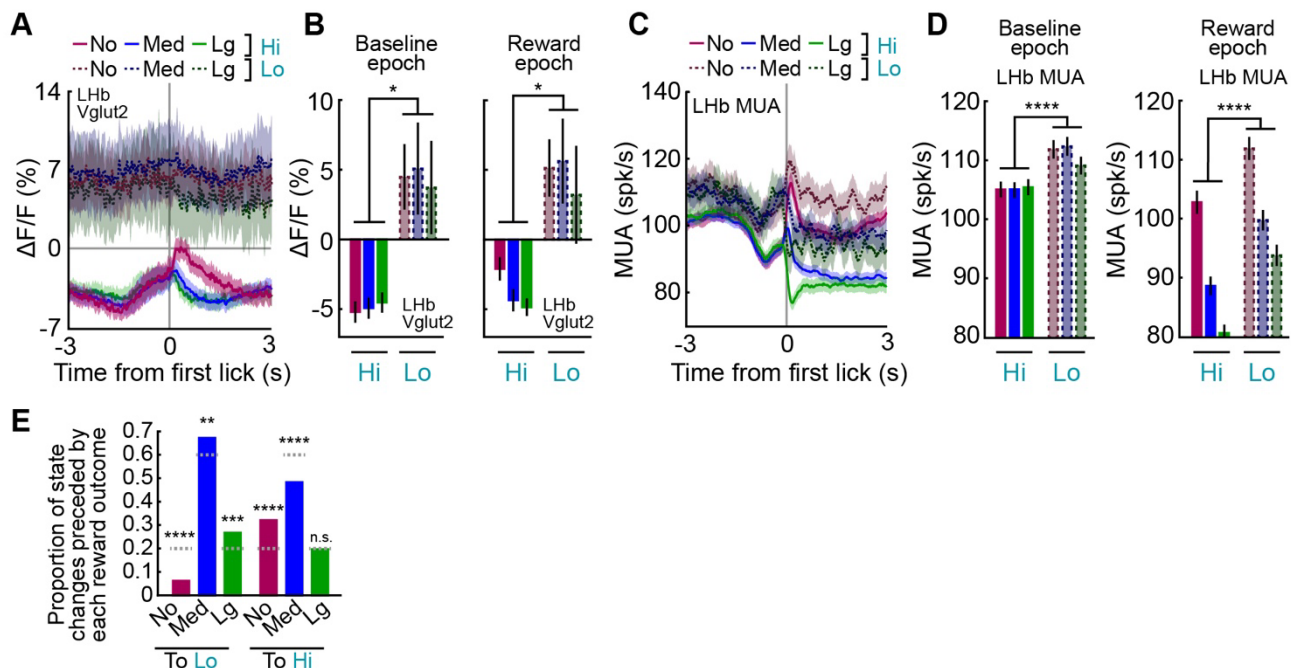
**Figure S2 (Related to Figure 2). Tonicity Elevated LHb Activity During Spontaneous Task Disengagement, Vglut2-ires-Cre Photometry and Multi-unit Electrophysiology**

(A) Example LHb Vglut2-ires-Cre photometry during a probabilistic task session. Ticks indicate poke and lick times (no reward, red; medium reward, blue; large reward, green). Engagement state, cyan line.

(B) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state and running speed in GCaMP ( $n = 15$  sessions, 3 mice) and GFP ( $n = 13$  sessions, 3 mice) Vglut2-ires-Cre mice. \*\* $p < 0.01$ , two-sample t-test.

(C) Example LHb multi-unit activity during a probabilistic task session, same conventions as A.

(D) Average regression coefficients. Multi-unit activity was regressed onto engagement state and running speed ( $n = 448$  sessions across 40 electrodes in 6 mice). \*\*\*\* $p < 0.0001$ , t-test compared to shuffled sample.



**Figure S3 (Related to Figure 3). Baseline LHB Activity Reflects Task Engagement State, Vglut2-ires-Cre Photometry and Multi-unit Electrophysiology**

(A) Average  $\Delta F/F$  in GCaMP Vglut2-ires-Cre mice (n = 15 sessions, 3 mice) during probabilistic task sessions, aligned to first lick and separated by engagement state (solid versus dotted lines) and reward value (no reward, red line; medium reward, blue; large reward, green).

(B) Average  $\Delta F/F$  during baseline and reward epochs in high and low engagement states. \*p < 0.05, paired samples t-test.

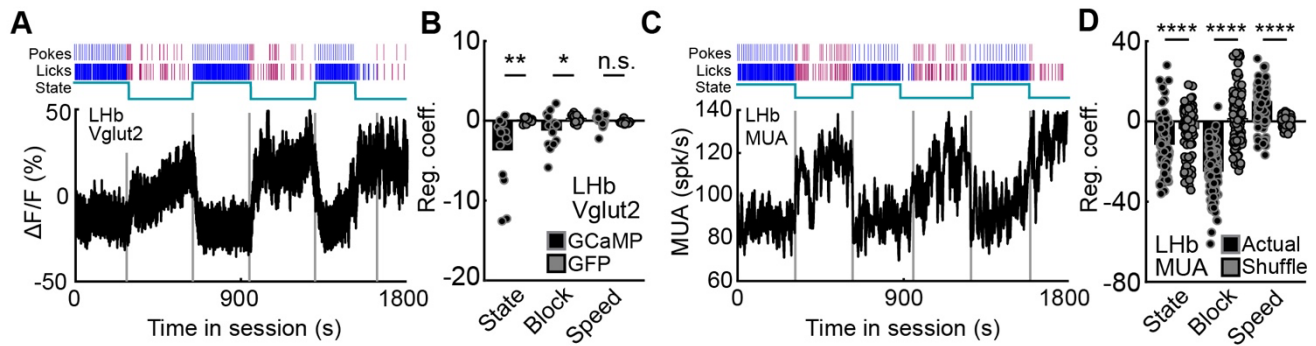
(C) Average multi-unit activity (n = 448 sessions across 40 electrodes in 6 mice) during probabilistic task sessions, aligned to first lick and separated by reward value. Same conventions as A.

(D) Average multi-unit activity during baseline and reward epochs in high and low engagement states. \*\*\*\*p < 0.0001, paired samples t-test.

(E) Proportion of transitions to low and high engagement states preceded by trials that yielded no reward, medium reward, or large reward. Chance proportions indicated by dashed gray lines. \*\*p < 0.01, \*\*\*p < 0.001, \*\*\*\*p < 0.0001, binomial test.

Shaded regions and error bars indicate SEM.





**Figure S4 (Related to Figure 4). Tonicity Elevated LHb Activity During Task Disengagement Due to Repeated Reward Omission, Vglut2-ires-Cre Photometry and Multi-unit Electrophysiology**

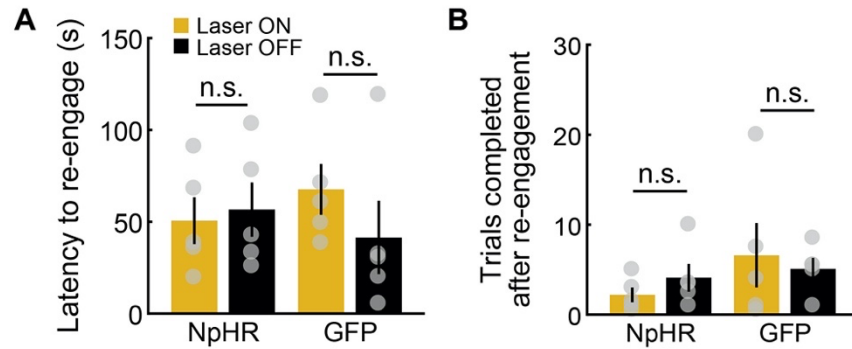
(A) Example LHb Vglut2-ires-Cre photometry during a blocked task session. Ticks indicate poke and lick times (no reward, red; medium reward, blue). Engagement state, cyan line.

(B) Average regression coefficients.  $\Delta F/F$  was regressed onto engagement state, block reward, and running speed in GCaMP ( $n = 15$  sessions, 3 mice) and GFP ( $n = 15$  sessions, 3 mice) Vglut2-ires-Cre mice.  $*p < 0.05$ ,  $**p < 0.01$ , two-sample t-test.

(C) Example LHb multi-unit activity during a blocked task session, same conventions as A.

(D) Average regression coefficients. Multi-unit activity was regressed onto engagement state, block reward, and running speed ( $n = 96$  sessions across 40 electrodes in 6 mice).  $****p < 0.0001$ , t-test compared to shuffled sample.

Error bars indicate SEM.



**Figure S5 (Related to Figure 6). LHb Inhibition Does Not Extend Task-Engaged States When Rewards Are Not Available**

(A) Average latency to re-engage in task performance without rewards. LHb inhibition initiated during disengagement. NpHR (n = 5) and GFP (n = 5) KIk8-Cre (NP171) mice.

(B) Average number of trials completed after task re-engagement without rewards. LHb inhibition initiated upon re-engagement. NpHR (n = 5) and GFP (n = 5) KIk8-Cre (NP171) mice.

Paired samples t-test. Error bars indicate SEM.

5