Sst+ GPi output neurons provide direct feedback to key nodes of the basal ganglia and drive behavioral flexibility.

Weglage M.*[1], Ährlund-Richter S.*[1,2], Fuzik J.[1], Skara V.[1], Lazaridis I.[#1,3], Meletis K.[#1].

*equal contribution

[#]corresponding author

[1]Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[2]Current address; Picower Institute of Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA, USA.

[3]Current address; McGovern Institute, Massachusetts Institute of Technology, Cambridge, MA, USA.

Abstract

The GPi-LHb pathway is the main output of the basal ganglia suggested to shape motivated behaviors. We show here that Sst+ GPi-LHb neurons send direct feedback to key nodes of the basal ganglia: the GPe, the striatal striosomes, and dopamine neurons in the SNc. Chronic silencing of this pathway did not affect learning or execution of value-guided choices, but severely disrupted the ability to adapt choice-behavior and seek an alternative reward location after task reversal. Calcium imaging revealed that Sst+ GPi neurons did not signal outcome value or value updates during reversal learning. Instead, progressive suppression of the Sst+ GPi activity was linked to increased commitment to one choice, and activity increased during exploration of alternative choices. We propose that GPi Sst+ neurons drive behavioral flexibility through a direct feedback signal to balance the activity of key nodes in the basal ganglia.

Introduction

The basal ganglia are a network of nuclei that are important for the generation of voluntary movements, decision-making and motivated behaviors [1,2]. The selection of specific actions, and the process of updating the value of chosen actions, ultimately produces motivated behaviors that optimize gain from the environment. Environmental conditions can rapidly and unexpectedly change, such as availability of resources (e.g. food), and animals must be able to dynamically adapt their choices in order to survive. From this perspective, the choice to continue performing the same rewarded action (exploitation), versus the exploration of alternative options, is a critical computation. The basal ganglia and the habenula are organized in an evolutionary conserved circuit optimized to support dynamic decision-making and choice updating [3]. The inability to adapt and explore (i.e. behavioral inflexibility) is manifested in habitual or compulsive actions [4], a condition in humans associated with psychiatric disorders and considerable cost in terms of quality of life [5].

The internal segment of the globus pallidus (GPi) and the substantia nigra pars reticulata (SNr) are the output nuclei of the basal ganglia, which in the established basal ganglia model regulate behavior through direct projections to thalamus and other motor centers [6,7]. The mouse GPi (also called entopeduncular nucleus) can be subdivided into two main output pathways: the parvalbumin-expressing (Pv+) GPi neurons that project to thalamus and brainstem, and the somatostatin-expressing (Sst+) GPi neurons that project to lateral habenula (LHb) [8–10]. Sst+ GPi neurons co-transmit GABA and glutamate in LHb [11–13], and have been suggested to carry information on action value [14] although optogenetic activation of the Sst+ GPi-LHb neurons does not shift value [11]. Overall, the GPi-LHb pathway holds a unique anatomical position to integrate information from the basal ganglia in order to guide motivated behavior.

We here used genetic targeting of the GPi-LHb neurons based on their defining molecular profile: co-expression of Vglut2 and Vgat, or expression of Sst [11,13]. We found that visualization of the brain-wide projection pattern of Sst+ and Vglut2/Vgat+ GPi neurons revealed a striking and unexpected organization. In addition to the known axonal projection to LHb, GPi neurons also have dense projections to key nodes in the basal ganglia: the striatal striosomes, the external segment of the globus pallidus (GPe), and dopamine neurons in the substantia nigra pars compacta (SNc). This organization suggested that Sst+ GPi neurons shape motivated behaviors beyond signaling action value, by calibrating the activity of functional nodes in the basal ganglia circuitry. We therefore investigated the role of Sst+ GPi neurons in a simple maze-based choice task during the initial acquisition, and after overtraining and rule-reversal, using viral tetanus toxin-mediated silencing and calcium imaging of the population activity. We found that the Sst+ GPi neurons were necessary for adjusting behavior after reversal and that their activity reflected the transition to more flexible behavior. We propose that a neuromodulatory feedback signal from Sst+ GPi neurons onto key

nodes of the basal ganglia is central for flexible behavior by shifting the network away from ongoing and reinforced choices to drive exploration of alternative options.

Results

GPi-LHb neurons have feedback connections to key nodes of the basal ganglia. The GPi output neurons that project to the LHb co-transmit GABA and glutamate can be identified based on the specific expression of Sst [11,13]. We aimed to visualize the axons of GPi-LHb Sst+ neurons to define their brain-wide projection patterns. To specifically target the LHb-projecting GPi neurons, a Cre- and Flp-dependent retrograde adeno-associated viral construct expressing ChR2-EYFP (AAVrg-Con/Fon-ChR2-EYFP) was injected into the LHb of Vglut2-Cre/Vgat-Flp transgenic mice (experimental strategy referred to as Vglut2/Vgat$^{GPi\text{-}LHb}$) (N=7 mice) (**Figure 1a**). As expected, this strategy resulted in EYFP labeling of neurons in the GPi with a dense projection to the LHb (**Figure 1b**). In addition to the known GPi-LHb pathway, we found extensive and dense collateral projections of Vglut2/Vgat$^{GPi\text{-}LHb}$ neurons to other subcortical and cortical brain regions, most prominently targeting the GPe, the caudate putamen (CPu), the SNc, and the prefrontal cortex (PFC) (**Figure 1b-c and g, Supplementary Figure 1a**). The dense axonal terminals of Vglut2/Vgat$^{GPi\text{-}LHb}$ axons within the GPe were confirmed by immunostaining to detect parvalbumin-expressing (Pv+) GPe neurons; which clearly outlined borders of the GPe and the axonal innervation within (**Figure 1d**). The Vglut2/Vgat$^{GPi\text{-}LHb}$ collaterals projected further into the CPu, with distinct patterns of concentrated axonal arborizations most commonly found in the dorsal CPu (**Figure 1b-c**). Due to the specific labeling pattern, we asked whether the axonal branching preferentially targeted the mu-opioid receptor-expressing (i.e. Oprm1, MOR) striosomes in the CPu [15]. Indeed, we found that the axonal arborization of LHb-projecting GPi neurons was limited to MOR+ striosomes (**Figure 1f**). The Vglut2/Vgat$^{GPi\text{-}LHb}$ collateral projections further reached into posterior subcortical regions, preferentially targeting dopaminergic (TH-expressing) neurons of the SNc (**Figure 1e**). We also observed axonal labeling in the substantia innominata (SI), but these axons could not be confirmed to originate from Vglut2/Vgat$^{GPi\text{-}LHb}$ neurons since we observed a small number of EYFP-expressing cell bodies within the SI, suggesting that a population of Vglut2/Vgat-expressing SI neurons project to the LHb (**Supplementary Figure 1b**).

We validated the feedback projection pattern of GPi-LHb neurons to the GPe, the CPu striosomes, the SNc, and the PFC by using two additional experimental strategies. First, to directly label Vglut2/Vgat-expressing neurons inside the GPi, we injected a Cre- and Flp-dependent viral construct (AAV8-Con/Fon-ChR2-EYFP) into the GPi of Vglut2-Cre/Vgat-Flp mice (experimental strategy referred to as Vglut2/Vgat^{GPi}; N=6 mice; **Supplementary Figure 2a**). Second, we labeled Sst+ neurons that project to the LHb by injecting a retrograde Cre-dependent viral construct (AAVrg-DIO-jGCaMP7s) into the LHb of Sst-Cre mice (experimental strategy referred to as SST$^{GPi\text{-}LHb}$; N=9 mice; **Supplementary Figure 3a**). Both the Sst$^{GPi\text{-}LHb}$ and the Vglut2/Vgat^{GPi} experimental strategies confirmed the observed specific projection pattern of GPi-LHb neurons, targeting the GPe, the SNc, the PFC, as well as the highly specific targeting of CPu

striosomes, but neither strategy resulted in axonal labeling in the SI (**Supplementary Figure 2b-e and Supplementary Figure 3a-dd**).

The experimental strategy of a direct viral injection into the GPi (experimental strategy Vglut2/Vgat^{GPi}) resulted in additional labeling in the subthalamic nucleus (STN) and the SNr, which was inconsistent with the results of the other two strategies (**Supplementary Figure 4a-b**). A careful mapping of the Vglut2/Vgat EYFP-expressing neurons revealed that a majority of the cell bodies were as expected in the GPi, but also revealed a small number of cell bodies in the adjacent STN (**Supplementary Figure 2f and Supplementary Figure 4c**). To determine whether the observed Vglut2/Vgat-expressing neurons in the STN projected to the SN, we injected a retrograde Cre- and Flp-depended viral construct (AAVrg-Con/Fon-ChR2-EYFP) in the SN of Vglut2-Cre/Vgat-Flp mice. This experiment resulted in the retrograde labeling of EYFP-expressing cell bodies in the GPi as expected, but also in the STN, confirming that Vglut2/Vgat+ STN neurons form a direct STN-SN projection, which is distinct from the GPi-LHb pathway (**Supplementary Figure 4d-f**). In conclusion, we found that GPi-LHb output neurons form an extensive feedback network that targets key nodes of the basal ganglia, including the GPe, the CPu striosomes and the SNc.

GPi-LHb neurons synapse and co-transmit glutamate and GABA in the basal ganglia targets. To further determine the postsynaptic targets of LHb-projecting GPi neurons we performed a series of monosynaptic retrograde tracing experiments using modified rabies virus and patch-clamp recordings. To investigate whether GPi-LHb neurons preferentially form synapses on striosome (Oprm1-expressing) striatal projection neurons (SPNs) versus D2-expressings SPNs (iSPNs), we injected a helper virus (AAV-DIO-TVA-V5-RbG), followed by a modified rabies virus expressing EGFP into the CPu of Oprm1-Cre (N=6 mice) or A2A-Cre mice (N=4 mice; **Figure 2a and d**). The rabies tracing experiments confirmed that Sst+ GPi neurons synapse directly onto Oprm1-expressing but not A2A-expressing SPNs, consistent with our observation that they preferentially target striosomes (**Figure 2a-c, Supplementary Figure 5a-b and Supplementary Table 1**). To similarly determine whether GPi-LHb neurons directly target dopamine neurons in SNc, we used the modified rabies virus tracing strategy in DAT-Cre mice (N=3 mice). We found that DAT-expressing SNc neurons received monosynaptic inputs from Sst+ GPi neurons thereby confirming a direct pathway (**Figure 2d-f, Supplementary Figure 5c and Supplementary Table 1**).

GPi-LHb neurons have previously been reported to co-transmit glutamate and GABA in the LHb [12]. Therefore, we investigated whether Vglut2/Vgat$^{GPi-LHb}$ axons also co-transmit glutamate and GABA from their terminals in the GPe. We used optogenetic activation of Vglut2/Vgat^{GPi} axon terminals in the GPe together with electrophysiology recordings from GPe neurons. A Cre- and Flp-dependent viral construct

expressing ChR2 (AAV8-Con/Fon-ChR2-EYFP) was injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice, and GPe neurons were recorded using cell-attached and whole-cell patch clamp recordings (**Figure 2g**). The electrophysiological response to optogenetic Vglut2/Vgat^{GPi} axon activation, location, and molecular profile of each patched GPe neuron was characterized (**Figure 2h-j)**. In cell-attached recordings, the optogenetic activation of Vglut2/Vgat^{GPi} axons (3ms, blue light) resulted in an action potential doublet followed by firing pause, which interrupted fast tonic firing of GPe neurons (**Figure 2h**). The whole-cell recorded postsynaptic response of GPe neurons to optogenetic activation was monosynaptic (i.e. in TTX, 4-AP) with a delay of 2-3ms. The response consisted of both an excitatory and an inhibitory component and repeated optogenetic stimulations (10Hz) caused short term synaptic depression of both components (**Figure 2k-l**). Sequential pharmacological dissection (Gabazine 10uM, NBQX 20uM, AP-V 50uM) confirmed the synaptic co-transmission of glutamate and GABA from Vglut2/Vgat^{GPi} terminals in all GPe neurons with optogenetically-evoked postsynaptic response (N=9 neurons; **Figure 2m-n**). We further probed the identity of the responsive GPe neurons, and we found that all neurons were PV+ and lacked FoxP2 expression (**Figure 2j and Supplementary Figure 5d**) and displayed intrinsic electrophysiological properties characteristic of PV+ GPe neurons (**Supplementary Figure 5e-f**).

We next investigated the presence of vesicular transporters for glutamate (i.e. Vglut2) and GABA (i.e. Vgat) in the boutons of Vglut2/Vgat^{GPi} axons in target regions with confirmed co-transmission (i.e. the GPe and the LHb), as well as in the CPu and the PFC. To visualize the boutons of Vglut2/Vgat^{GPi} axons, we injected a Cre- and Flp-dependent viral construct (AAV8-Con/Fon-GCaMP6m) into the GPi of Vglut2-Cre/Vgat-Flp mice and mapped the presence of Vglut2 and Vgat by immunostaining (**Supplementary Figure 5g-h**). In all target regions, the vast majority of GFP-labeled boutons from Vglut2/Vgat^{GPi} axons contained both Vglut2 and Vgat (co-labeled boutons/total boutons: LHb 433/465; GPe 274/292; CPu 67/75; PFC 31/34; **Supplementary Figure 5i**), further supporting that Vglut2/Vgat^{GPi} axons can co-transmit glutamate and GABA in all basal ganglia targets.

Sst+ GPi-LHb neurons drive behavioral flexibility. To investigate the role of Sst+ LHb-projecting GPi neurons in motivated behaviors and specifically decision-making, we trained mice on a simple 2-choice maze reversal task. In this task, mice learn to initiate a trial in order to then run and turn into one arm of the maze to receive a reward. The correct, rewarded arm remained consistent for the first 14 training days (sessions) and was not cued by auditory or visual signals. After the 14th day, the reward location switched to the opposite arm, forcing the mice to adapt their choice behavior to this task reversal in order to maximize rewards.

We aimed to determine whether GPi neurons are critical for learning and executing the correct choice before and after the rule reversal. To do so, we functionally inactivated Sst+ GPi-LHb neurons by disabling their neurotransmitter release machinery by cell-type specific viral expression of the tetanus toxin light chain (TeTxLC) [16]. We injected a Cre-dependent retrograde virus expressing TeTxLC (AAVrg-DIO-TeTxLC-dTomato) bilaterally into the LHb of Sst-Cre mice (N=6 mice) and compared their behavior to a control group of fiber-implanted mice expressing jGCaMP7s (**Figure 3a-b and Supplementary Figure 6c**). The mice performed the reversal task in a maze similar to a T-maze, but with sharper angles for choice turns (i.e. arrow maze; **Figure 3c**). A trial consisted of the complete traversal of the center corridor and of one of the two "choice" corridors branching off at the maze's three-way junction. Mice received a water reward at the end of the corridor after choosing the correct side, and after correct initiation in the center corridor (see Methods for details and **Supplementary Table 2**).

We found that silencing of the Sst+ GPi neurons did not interfere with the learning or the execution of the correct choice turn in the first phase of the task, but it severely impaired the ability to reevaluate and/or suppress this well-trained response post-reversal (**Figure 3d-e**). During the pre-reversal phase, mice with TeTxLC expression in Sst+ GPi neurons reached accurate and asymptotic performance (**Figure 3d**). However, in the post-reversal phase, the choice-behavior of the same mice regressed to chance-level accuracy, from which it did not recover even after 14 days of training, whereas the control group achieved behavioral accuracy within 3 days of the reward location reversal (**Figure 3d**). Importantly, the performance variability increased substantially after the reversal, and poor performance was strongly associated with the number of Sst+ GPi neurons expressing TeTxLC (**Figure 3e**). We found that these behavioral deficits could not be explained by motor or motivational impairments, as the number of trials performed per session was independent of the extent of the TeTxLC expression, and did not differ between the impaired and the control group (**Figure 3f and Supplementary Figure 7a**).

A closer inspection of individual training sessions revealed a uniform behavioral pattern: well-trained mice in both the impaired and the control groups performed fewer trials and committed more errors as the session progressed. In the final week of pre-reversal sessions, mice completed half of the trials in each session at a rapid pace, but slowed down to about half, and finally a quarter, of that rate over the two subsequent quarters of trials (**Figure 3g-h**). This was also reflected in the decreased choice accuracy, from close to 100% correct in the first two quarters to approximately 75% in the fourth (**Figure 3h**). In summary, mice expressing TeTxLC in the Sst+ GPi population display a selective deficit in their ability to flexibly adapt their choice after reversal of the reward location, without affecting their normal motor pattern.

Sst+ GPi population activity does not reflect the choice value. We aimed to investigate how the Sst+ GPi neuron activity was correlated to the performance in the maze task, by focusing on the outcome evaluation signals or prediction errors, which could underlie the deficit in reversal learning. To that end, we used fiber photometry to record the population activity of Sst+ GPi neurons, as well as the activity of the Pv+ GPi neurons. Cell-type specific expression of the calcium indicator jGCaMP7s [17] was accomplished by an injection of a Cre-dependent retrograde viral construct (AAVrg-DIO-jGCaMP7s) in the LHb of Sst-Cre mice, or by a direct injection of a Cre-dependent virus (AAV-DIO-jGCaMP7s) in the GPi of Pv-Cre mice. We implanted optical fibers above the GPi to capture the neuronal activity-dependent fluorescence emitted by jGCaMP7s (**Figure 3b and Supplementary Figure 6a**).

We first analyzed the GPi population responses to trial outcomes that were rendered surprising to mice by the reversal of the reward location (**Figure 4a-f**). Based on the hypothesis that GPi projections to LHb would encode negative value, we predicted that their activity would increase in response to unexpected negative value (reward omissions), and decrease in response to unexpected positive value (reward delivery), effectively signaling the inverse of the dopamine reward prediction error. In contrast, we predicted that the Pv+ GPi population activity would predominantly correlate with motor behavior (i.e. locomotion).

We first observed that the mice only briefly slowed down at the water spout when the expected water drops were not delivered (**Figure 4a**), whereas they stopped for approximately one second to consume unexpected rewards (**Figure 4d**). The average population traces recorded from Pv+ GPi neurons reflected these movements, slowly decreasing when mice were mostly immobile and consumed the reward, and increasing as they accelerated away from the unrewarded spout (**Figure 4b and e**). In contrast to our prediction, the LHb-projecting Sst+ GPi population did not show any clear activity pattern reflecting value or value update (positive or negative), and instead decreased regardless of the outcome after mice reached the spout (**Figure 4b and e, Supplementary Figure 7c**). The mean fluorescence changes recorded from Sst+ neurons following arrival at the spouts did not evolve over the course of learning as expected of a prediction error signal (i.e. the absolute magnitude did not decrease as the initially unexpected outcomes became gradually more expected). In fact, the activity changes of Sst+ neurons tended to grow more negative across the unrewarded trials, and did not vary considerably across the rewarded trials (**Figure 4c and f**).

The absence of a clear, population-level outcome evaluation or prediction error signal in the activity of LHb-projecting Sst+ GPi neurons was further strengthened by a second experiment, in which we trained mice on a two-choice probabilistic reversal task (**Figure 4g**). After initiating at a central nosepoke port, mice had to choose between two adjacent choice ports, one of which yielded a water reward with 75% probability. The location of the reward reversed unpredictably every 7-23 rewarded trials, forcing mice to constantly evaluate the recent reward history in order to infer the correct choice. The mice performed this task

successfully, allowing us to fit a Q-learning model that accurately captured the behavior and generated prediction error estimates for each trial (**Supplementary Figure 7d-e**). For both the Pv+ and Sst+ GPi populations, we found that the mean fluorescence changes during reward consumption were best explained by a negative correlation with the duration of rewarded nosepokes rather than with the estimated value updates (**Figure 4h**). To investigate this in more detail, we split both rewarded and unrewarded choice trials (contraversive) into low value update trials, in which the current outcome matched the previous one, and into high value update trials in which it did not (**Figure 4i**). The mean fluorescence traces of the Sst+ population showed a slow decrease subsequent to both rewarded and unrewarded choice outcomes during the period when mice remained in the port (**Figure 4j**). Importantly, the reward and the no-reward (omission) responses did not differ between low value and high value update trials (**Figure 4i-j**, **Supplementary Figure 7f**). To further investigate the possible encoding of negative value in the GPi-LHb output, we recorded the activity of Sst+ and Pv+ GPi neurons during classical fear conditioning where an auditory cue predicted the delivery of mild foot shocks. We found that the activity of both the Sst+ and Pv+ GPi populations clearly lacked responses to the negative value, and the shock-predicting auditory cue, and both populations instead showed correlation with mobility (**Supplementary Figure 7g**). Based on the three different behavioral paradigms, we conclude that the Sst+ LHb-projecting GPi neurons are not broadcasting a prediction or evaluation signal.

Sst+ GPi-LHb population activity is suppressed during the choice phase. Since the Sst+ GPi population lacked value or value update signals we considered that impaired performance in the reversal task after tetanus toxin silencing could instead be explained by an impaired ability to shift the persistent choice to an alternative option. To first determine the GPi activity during spontaneous movements, we decided to characterize the activity of Sst+ GPi neurons during exploration in an open field. We found that the Sst+ GPi population activity increased monotonically with movement velocity, similar to the Pv+ GPi population activity, although the Sst+ population did not show a preference for contraversive turns in contrast to the Pv+ population (**Figure 5a-b**). Surprisingly, we found that the activity during running and turning in rewarded trails in the maze was very different compared to the open field for both Sst+ and Pv+ GPi neurons. In pre-reversal sessions, Sst+ GPi neuron activity was suppressed during early contraversive choice runs, and activity dropped as mice approached the turn (i.e. the choice phase). This choice-related suppression was less pronounced as mice continued to perform rewarded trials, and the activity increased during the choice phase during later trials within the same session. Strikingly, the increased activity during the choice phase in late trials was associated with a decreased velocity (**Figure 5c-d**). The activity of Sst+ neurons during ipsiversive choice runs was not distinctly modulated in early post-reversal trials and

increased towards the turn in every session quartile (**Figure 5d**). In contrast, Pv+ GPi neurons displayed a constant elevated activity level in both contra- and ipsiversive runs and largely independent of turning (**Figure 5d**).

We then asked whether the transient suppression of the Sst+ GPi pathway during the choice phase, which was evident in the first quarter of trials in pre-reversal sessions, could be related to task performance. To investigate this, we plotted the neuronal activity at the onset of the turn versus the fraction of trials completed for each session separately, and fitted linear regression lines for each mouse (**Figure 5e**). Interestingly, the suppression of Sst+ GPi neurons emerged over the first week of training, as mice developed a strong preference for the rewarded, contraversive choice corridor. After the task reversal, once mice began to consistently redirect their turns towards the opposite, ipsiversive choice corridor, the suppression of Sst+ GPi neurons vanished. This pattern is concisely summarized by the regression coefficients, where the intercept shows the suppression of Sst+ GPi neurons at the start, and the slope the magnitude of the change by the end of the session. We found that the coefficients fitted for the Sst-Cre mice changed significantly over the 20 days of the reversal task, whereas those for the Pv-Cre mice did not (**Figure 5f**). The behavioral acquisition of the prepotent contraversive choice response in the task thus coincided with a substantial suppression of the activity of LHb-projecting Sst+ GPi neurons, which resolved as the mice returned to a more flexible behavioral state and successfully reversed their behavior. Overall, our findings on the dynamics of the choice-related activity over the learning process and adaptation to rule reversal support that the Sst+ GPi pathway is a critical node in regulating behavioral flexibility.

Discussion

The basal ganglia play a central role in guiding choices and evaluating actions, processes that are thought to depend on the GPi output nucleus that projects to the LHb [3,18]. In addition to the known projections to LHb, we found that Sst+ GPi neurons make monosynaptic connections with the GPe, the striatal striosomes, and dopamine neurons in the SNc. This neuroanatomical organization suggests a complex functional role for Sst+ GPi neurons, beyond transmitting negative value to LHb [14,19–21].

To investigate the role of GPi-LHb neurons in motivated behavior, we silenced Sst+ GPi neurons in a 2-choice reversal task to determine their role in the learning process, expert-level performance, and adaptation to task reversal. We found that cell-type specific silencing of the Sst+ GPi population did not affect the ability to learn or to perform at expert level in the task. Instead, we found that the activity of Sst+ GPi neurons was necessary for the ability to flexibly adapt the behavior upon task reversal, and to select the alternative choice when the reward location changed. This impairment in behavioral flexibility could be the result of deficits in encoding the action, or in outcome evaluation, based on the role of LHb in signaling negative value [18,22,23]. To directly map the activity of Sst+ GPi neurons during encoding of negative value, we recorded their calcium activity in two different choice tasks using fiber photometry. Surprisingly, the activity of Sst+ GPi neurons did not track reward omissions in the 2-choice maze reversal task, and their activity did not reflect the action value in the 2-choice probabilistic reversal task. We therefore analyzed the Sst+ GPi activity in relation to other task-related actions. We found that Sst+ GPi neuron activity was highly modulated when mice decided which reward spout to approach, i.e. in the choice phase of the task. As mice learned to perform the task, the Sst+ GPi activity became increasingly suppressed during the choice phase, showing the largest decrease of activity in expert-level mice that persistently and rapidly executed the task. The suppression of Sst+ GPi population activity developed over days, and followed the behavioral performance of the mice, indicating a role in shaping the learning. Furthermore, Sst+ GPi neurons showed a ramping of activity within single sessions that corresponded to reduced accuracy in the choice task and increased exploration of the maze. After task reversal, Sst+ GPi activity was elevated in the choice phase as the mice returned to a more flexible behavioral state and successfully reversed their behavior.

The GPi-LHb pathway has been proposed to encode negative value, most likely through control over dopamine neuron activity [14,19,22,24,25]. However, more recent work has suggested that the negative value encoding found in the LHb originates from excitatory inputs from the LHA [11,26], and other regions [27–29]. Overall, our findings suggest that the Sst+ GPi neurons regulate the level of commitment to one action, and are not encoding the action value. The task-specific activity pattern of Sst+ GPi neurons suggests that the inability to reverse actions upon silencing of the GPi-LHb pathway is unlikely to be due to an evaluation

deficit. Instead, the inflexible behavior is the result of an impairment in the exploration of alternative choices, manifesting due to the lack of activity of Sst+ GPi neurons.

Our data support a model where GPi Sst+ activity is suppressed during repetitive, habit-like actions, and elevated during the exploration of alternative options in the environment. Based on these findings, we propose that the Sst+ GPi neurons drive behavioral flexibility in order to counteract habitual or compulsive choices, and that this function emerges from the feedback of a GABA/Glutamate signal from Sst+ GPi neurons to specific nodes of the basal ganglia. Our results raise an interesting function for GABA/glutamate co-transmission to modulate the activity of each target region during learning, habit formation, and in the reinstatement of behavioral flexibility. Future work will have to address the possible synaptic plasticity or dynamics of GABA/glutamate co-transmission in the GPi feedback circuitry and the potential role of the feedback system in compulsive disorders or Parkinson's disease.

Within the basal ganglia, the activity in striosomes and of dopamine neurons are thought to track state-dependent reward expectations and prediction errors. This "critic" circuitry influences the "actor" circuitry (e.g. in dorsal striatum) to perform the most situation-appropriate motor action [3,30–32]. In dorsal striatum, the dorsomedial striatum is primarily promoting goal-directed behavior and the dorsolateral striatum is engaged in more stereotyped and habitual motor actions [33–36]. Our findings challenge the notion that GPi-LHb neurons are a part of the "critic" circuitry of the basal ganglia, for example by encoding the action value and prediction errors. Instead, our data suggests that Sst+ GPi neurons influence the balance between the goal-directed versus habitual domain of the "actor" circuitry of the basal ganglia, for example by regulating dopamine release in the dorsal striatum and via striosome-dependent mechanisms [33,36–38].

We have here shown how the GPi feedback pathway expands the complexity of connectivity diagrams that depict the basal ganglia circuit as a set of one-directional parallel pathways. Our results point to an unexpected connectivity in the basal ganglia, revealing a major feedback system from a GPi output population that can act as a neuromodulatory signal to suppress the persistent commitment to repeat the same action, thereby promoting the exploration of alternative options in the environment.

Acknowledgements

We thank the Animal Behavior Core Facility (ABCF) at Karolinska Institutet for access to behavioral equipment. We thank Karl Deisseroth and Christina Kim for generous support with establishing the fiber photometry system (FIP). Funding for this study was provided by the Swedish Society for Medical Research (Svenska Sällskapet för Medicinsk Forskning) and the Wenner Gren Foundation to S.Ä-R.; the Swedish Research Council (Vetenskapsrådet), the Swedish Brain Foundation (Hjärnfonden), Karolinska Institutet (KID doctoral funding for M.W.), the Swedish Foundation for Strategic Research, and the Knut och Alice Wallenberg Foundation to K.M..

Author Contributions

S.Ä-R. and I.L. designed, interpreted and performed the anatomical characterization. S.Ä-R. analyzed and visualized the anatomical data. M.W. and I.L. designed and performed the behavioral experiments. M.W. analyzed, interpreted and visualized all behavioral data. V.S. performed anatomical and behavioral experiments. J.F. performed all electrophysiology experiments and analysis. M.W., S.Ä-R. and K.M. wrote the manuscript. I.L. and K.M. supervised and conceptualized the work. All of the authors contributed, discussed, and commented on the manuscript.

Competing Interests statement

The authors declare no competing interests.

Data availability

All data presented in this paper is available upon request to the corresponding authors.

Code availability

Custom code is available upon request.

References

1. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
2. Graybiel, A. M., Aosaki, T., Flaherty, A. W. & Kimura, M. The Basal Ganglia and Adaptive Motor Control. *Science* **265**, 1826–1831 (1994).
3. Grillner, S. & Robertson, B. The Basal Ganglia Over 500 Million Years. *Curr. Biol.* **26**, R1088–R1100 (2016).
4. Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V. & Graybiel, A. M. Building Neural Representations of Habits. *Science* **286**, 1745–1749 (1999).
5. Robbins, T. W., Vaghi, M. M. & Banca, P. Obsessive-Compulsive Disorder: Puzzles and Prospects. *Neuron* **102**, 27–47 (2019).
6. Alexander, G. E. & Crutcher, M. D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* **13**, 266–271 (1990).
7. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
8. Rajakumar, N., Elisevich, K. & Flumerfelt, B. A. Parvalbumin-containing GABAergic neurons in the basal ganglia output system of the rat. *J. Comp. Neurol.* **350**, 324–336 (1994).
9. Van Der Kooy, D. & Carter, D. A. The organization of the efferent projections and striatal afferents of the entopeduncular nucleus and adjacent areas in the rat. *Brain Res.* **211**, 15–36 (1981).
10. Vincent, S. R. & Brown, J. C. Somatostatin immunoreactivity in the entopeduncular projection to the lateral habenula in the rat. *Neurosci. Lett.* **68**, 160–164 (1986).
11. Lazaridis, I. *et al.* A hypothalamus-habenula circuit controls aversion. *Mol. Psychiatry* 1 (2019) doi:10.1038/s41380-019-0369-5.
12. Shabel, S. J., Proulx, C. D., Piriz, J. & Malinow, R. GABA/glutamate co-release controls habenula

output and is modified by antidepressant treatment. *Science* **345**, 1494–1498 (2014).

13. Wallace, M. L. *et al.* Genetically Distinct Parallel Pathways in the Entopeduncular Nucleus for Limbic and Sensorimotor Output of the Basal Ganglia. *Neuron* **94**, 138-152.e5 (2017).
14. Stephenson-Jones, M. *et al.* A basal ganglia circuit for evaluating action outcomes. *Nature* **539**, 289–293 (2016).
15. Märtin, A. *et al.* A Spatiomolecular Map of the Striatum. *Cell Rep.* **29**, 4320-4333.e5 (2019).
16. Pashkovski, S. L. *et al.* Structure and flexibility in cortical representations of odour space. *Nature* **583**, 253–258 (2020).
17. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).
18. Hikosaka, O. The habenula: from stress evasion to value-based decision-making. *Nat. Rev. Neurosci.* **11**, 503–513 (2010).
19. Hong, S. & Hikosaka, O. The globus pallidus sends reward-related signals to the lateral habenula. *Neuron* **60**, 720–729 (2008).
20. Meye, F. J. *et al.* Shifted pallidal co-release of GABA and glutamate in habenula drives cocaine withdrawal and relapse. *Nat. Neurosci.* **19**, 1019–1024 (2016).
21. Shabel, S. J., Proulx, C. D., Trias, A., Murphy, R. T. & Malinow, R. Input to the Lateral Habenula from the Basal Ganglia Is Excitatory, Aversive, and Suppressed by Serotonin. *Neuron* **74**, 475–481 (2012).
22. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
23. Matsumoto, M. & Hikosaka, O. Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci.* **12**, 77–84 (2009).
24. Christoph, G. R., Leonzio, R. J. & Wilcox, K. S. 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 (1986).

25. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212–217 (2012).

26. Trusel, M. *et al.* Punishment-Predictive Cues Guide Avoidance through Potentiation of Hypothalamus-to-Habenula Synapses. *Neuron* **102**, 120-127.e4 (2019).

27. Ottenheimer, D. J. *et al.* A quantitative reward prediction error signal in the ventral pallidum. *Nat. Neurosci.* **23**, 1267–1276 (2020).

28. Root, D. H. *et al.* Distinct Signaling by Ventral Tegmental Area Glutamate, GABA, and Combinatorial Glutamate-GABA Neurons in Motivated Behavior. *Cell Rep.* **32**, 108094 (2020).

29. Tooley, J. *et al.* Glutamatergic Ventral Pallidal Neurons Modulate Activity of the Habenula–Tegmental Circuitry and Constrain Reward Seeking. *Biol. Psychiatry* **83**, 1012–1023 (2018).

30. Barto, A. G., Sutton, R. S. & Anderson, C. W. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Trans. Syst. Man Cybern.* **SMC-13**, 834–846 (1983).

31. Doya, K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* **12**, 961–974 (1999).

32. Houk, J. C., Adams, J. L. & Barto, A. G. A Model of How the Basal Ganglia Generate and Use Neural Signals That Predict Reinforcement. in *Models of Information Processing in the Basal Ganglia* (The MIT Press, 1994). doi:10.7551/mitpress/4708.003.0020.

33. Graybiel, A. M. Habits, Rituals, and the Evaluative Brain. *Annu. Rev. Neurosci.* **31**, 359–387 (2008).

34. Ragozzino, M. E. The Contribution of the Medial Prefrontal Cortex, Orbitofrontal Cortex, and Dorsomedial Striatum to Behavioral Flexibility. *Ann. N. Y. Acad. Sci.* **1121**, 355–375 (2007).

35. Thorn, C. A., Atallah, H., Howe, M. & Graybiel, A. M. Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron* **66**, 781–795 (2010).

36. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476 (2006).

37. Jenrette, T. A., Logue, J. B. & Horner, K. A. Lesions of the Patch Compartment of Dorsolateral

Striatum Disrupt Stimulus–Response Learning. *Neuroscience* **415**, 161–172 (2019).

38. Nadel, J. A. *et al.* Optogenetic stimulation of striatal patches modifies habit formation and inhibits dopamine release. *Sci. Rep.* 1–16 (2021) doi:10.1038/s41598-021-99350-5.
39. Ährlund-Richter, S. *et al.* A whole-brain atlas of monosynaptic input targeting four different cell types in the medial prefrontal cortex of the mouse. *Nat. Neurosci.* **22**, 657–668 (2019).
40. Kim, C. K. *et al.* Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat. Methods* **13**, 325–328 (2016).
41. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
42. Nath, T. *et al.* Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat. Protoc.* **14**, 2152–2176 (2019).
43. Lopes, G. *et al.* Bonsai: an event-based framework for processing and controlling data streams. *Front. Neuroinformatics* **9**, (2015).
44. Steiner, H.-C. Firmata : Towards Making Microcontrollers Act Like Extensions of the Computer. *Proceedings of the International Conference on New Interfaces for Musical Expression* 125–130 (2009) doi:10.5281/zenodo.1177689.
45. Weglage, M. *et al.* Complete representation of action space and value in all dorsal striatal pathways. *Cell Rep.* **36**, 109437 (2021).
46. Tai, L.-H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* **15**, 1281–1289 (2012).

METHODS

Viral Constructs. *Anatomy and histology.*

AAVrg-DIO-ChR2-mCherry; retroAAV-pAAV-Ef1a-double floxed-hChR2(H134)-mCherry, Addgene, cat. no. 20297 (7x10e12 viral genomes/ml).

AAVrg-Con/Fon-ChR2-EYFP; pAAV-hSyn-Con/Fon-hChR2(H134)-EYFP, Addgene, cat. no. 55645 (7x10e12 viral genomes/ml).

AAVrg-DIO-jGCaMP7s; retroAAV-pGP-AAV-syn-FLEX-jGCaMP7s-WPRE, Addgene, cat. no. 104491 (7x10e12 viral genomes/ml).

AAV-Con/Fon-ChR2-EYFP; AAV8-pAAV-hSyn-Con/Fon-hChR2(H134)-EYFP, Addgene, cat. no. 55645 (2x10e13 viral genomes/ml).

AAV-Con/Fon-GCaMP6m; AAV8-EF1a-Con/Fon-GCaMP6m, produced in the laboratory of Dr. Karl Deisseroth (Stanford University), AAV#1746 (3,5x10e12 viral genomes/ml).

Rabies tracing.

The rabies helper virus AAV-DIO-TVA-V5-Rabies G; AAV5-EF1a-DIO-TVA-V5-t2A-Rabies G, was created as described in Ährlund-Richter et al. (2019) (Addgene cat. plasmid no. 119743). AAV5 viral particles were produced by the Virus Vector Core facility at the University of North Carolina (1.2x10e13 viral particles/ml). The EnvA pseudotyped Rabies Glycoprotein-deleted Rabies Virus (Rb) (1x10e10 infection unit/ml) was produced in-house as described in [39].

Behavior.

Tetanus Toxin.

AAVrg-DIO-TeTxLC-dTomato; pAAV-hSyn-FLEX-TeLC-P2A-dTomato was a gift from Sandeep Datta, Addgene, cat. plasmid no.159102. retroAAV viral particles were produced by the Virus Vector Facility at the University of Zurich (7x10e12 viral particles/ml).

Fiber Photometry.

AAVrg-DIO-jGCaMP7s; retroAAV-pGP-AAV-syn-FLEX-jGCaMP7s-WPRE, Addgene, cat.no. 104491 (7x10e12 viral genomes/ml).

AAV1-DIO-jGCaMP7s; AAV1-pGP-AAV-syn-FLEX-jGCaMP7s-WPRE, Addgene, cat. no. 104491 (1x10e13 viral genomes/ml).

Animal subjects. Experiments were conducted using adult male and female mice, wild type C57BL/6J (Charles River Laboratories) or transgenic mouse lines; Vglut2-Cre: Slc17$^{a6tm2(cre)Lowl}$, Jackson stock no. 028863; Sst-Cre: B6N.Cg-Sst$^{tm2.1(cre)Zjh}$, Jackson stock no. 018973, Pv-Cre: Pvalb$^{tm1(cre)Arbr}$, Jackson stock no. 008069. Vgat-Flpo transgenic mice were produced by Hongkui Zeng (Allen Institute for Brain Science, USA). All transgenic mice used in experiments were heterozygous for the transgenes. Mice were maintained under standard housing conditions with a 12-hour light cycle and with *ad libitum* access to food and water unless placed on a food or water restriction schedule (see behavioral paradigms for details). All procedures were approved and performed in accordance and compliance with the guidelines of the Stockholm Municipal Committee (approval no. N166/15, 7362-2019 and 15440-2020).

Animal cohorts;

Anatomy and histology.: N=7 retro LHb injection Vgat-Flpo/Vglut2-Cre, N=3 retro SNr/SNc injection Vgat-Flpo/Vglut2-Cre, N=6 direct injection GPi Vgat-Flpo/Vglut2-Cre, N=9 retro LHb injection SST-Cre mice.

Electrophysiology: N=8 Vgat-Flpo/Vglut2-Cre mice.

Rabies: N=3 Rabies injection in the CPu of A2A-Cre, N=6 Rabies injection in the CPu of Oprm1-Cre, N=3 Rabies injection in the SNc of DAT-Cre mice.

Fiber photometry: N=8 SST-Cre mice and N=6 PV-Cre mice.

Tetanus Toxin: N=6 SST-Cre mice.

See also **Supplementary Table 2**.

Viral injections and implants. *General procedure***.** Mice were anesthetized with isoflurane (2%) and placed into a stereotaxic frame (Harvard Apparatus, Holliston, MA). Before the first incision the analgesic Buprenorphine (0.1 mg/kg) and local analgesic Xylocain/Lidocain (4 mg/kg) was administered subcutaneously . The temperature of the mice was maintained at 36 °C with a feedback-controlled heating pad. For viral injections a micropipette attached on a Quintessential Stereotaxic Injector (Stoelting, Wood Dale, IL) was used. The pipette was held in place for 5 min after the injection before being slowly retracted from the brain. The analgesics Carpofen (5mg/kg) was given at the end of the surgery, followed by a second dose 18-24h after the surgery.

Anatomy. For anatomical characterization of the GPi-LHb pathway in SST-Cre mice, 0.3µl of either AAVrg-DIO-ChR2-mCherry or AAVrg-DIO-jGCaMP7s was unilaterally injected into the LHb (coordinates: AP -1.65mm, ML 0.95mm, DV -2.45mm). Targeting of the GPi-LHb pathway in Vgat-Flp/Vglut2-Cre mice was achieved by either an unilateral injection of 0.3µl of AAVrg-Con/Fon-ChR2-EYFP into the LHb (coordinates: AP -1.65mm, ML 0.95mm, V -2.45mm) or a direct injection of 0.2µl of AAV8-Con/Fon-ChR2-EYFP or AAV8-Con/Fon-GCaMP6m in the GPi (coordinates: AP -1.3mm, ML 1.75mm, DV -4.1mm).

Rabies tracing. For cell-type specific monosynaptic retrograde tracing mice were unilaterally injected with 0.3µl of the helper virus AAV-DIO-TVA-V5 in the SNc (DAT-Cre; coordinates: AP -3.08mm, L 1.2mm, V- 4.25mm) or 0.4µl in the CPu (A2A-Cre or Oprm1-Cre; coordinates: AP +0.5mm, L 2.1mm, V- 2.0mm). After 21 days, 0.3µl of Rabies-EGFP virus was injected at the same coordinates. Mice were perfused and brains were collected for histology 7 days post rabies injection.

Slice electrophysiology. Targeting and labeling of neuronal inputs was achieved by a unilateral injection of 0.2µl AAV8-Con/Fon-ChR2-mCherry into the GPi (coordinates: AP -1.22 mm, ML 1.75 mm, DV -4.1 mm) of Vgat-Flp/Vglut2-Cre mice.

Fiber Photometry. Adult male and female SST-Cre mice were injected with 0.3µl of AAVrg-DIO-jGCaMP7s [17] into the LHb (coordinates: AP -1.65 mm, ML ±0.95 mm, DV -2.45 mm). PV-Cre mice were injected with 0.2µl of AAV1-DIO-jGCaMP7s directly in the GPi (coordinates: AP -1.22 mm, ML ±1.75mm, DV -4.1mm). Injected mice were implanted one week later with optical fibers (200µm diameter) aimed directly above the GPi (coordinates: AP -1.22mm, ML ±1.75mm, DV -3.9mm).

Tetanus toxin. Male and female SST-Cre mice were bilaterally injected with 0.3µl of AAVrg-DIO-TeTxLC-dTomato [16] in the LHb (coordinates: AP -1.65mm, ML ±0.95mm, V -2.45mm). Training of the mice in the arrow maze reversal task started 21 days post injection.

Histology. *General procedure***.** Mice were deeply anaesthetized with pentobarbital and then transcardially perfused with 0.1M PBS followed by 4% paraformaldehyde in PBS 0.1M. Brains were removed and post-

fixed in 4% paraformaldehyde in PBS 0.1M overnight at 4°C and then washed and stored in 0.1M PBS. Coronal (60μm thickness) or sagittal sections (80μm thickness) were cut using a vibratome (Leica VT1000, Leica Microsystems, Nussloch GmbH, Germany).

*Immunostaining***.** Immunostaining was performed on free-floating sections. Briefly, sections were incubated for one hour in 0.3% TritonX-100 in Tris-buffered saline (0.3% TBST). For difficult stainings, the brain sections were treated with a preheated (40°C) antigen retrieval solution (10mM sodium citrate, 0,05% Tween20, pH:6) for 1-2 minutes. In order to block non-specific antibody binding, sections were incubated in 5% Normal Donkey Serum in 0.3% TBST, for one hour at room temperature (RT). Sections were subsequently incubated overnight at RT with primary antibodies followed by a 4h incubation with the secondary antibodies. The sections where thereafter washed consecutively in 0.3% TBST, 1x TBS and 1x PBS (15min each) and mounted on glass slides (Superfrost Plus, Thermo Scientific) and coverslipped (Thermo Scientific) using glycerol: 1x PBS (50:50).

Primary antibodies used (*abbreviation; dilution; company; catalogue number*): Guinea pig anti-parvalbumin (PV; 1:1000; Synaptic Systems; 195 004), Rabbit anti-FoxP2 (FoxP2, 1:1000; Abcam; ab16046), Rabbit anti-Mu Opioid Receptor (MOR; 1:200; Abcam; ab134054), Chicken 1), Sheep anti-Tyrosine Hydroxylase (TH, 1:1000; Millipore; AB1542), Rabbit anti-Green Fluorescent Protein (GFP; 1:1000; Invitrogen; A6455), Chicken anti-Green Fluorescent Protein (GFP; 1:1000; Invitrogen; A10262), Goat anti-Green Fluorescent Protein (GFP; 1:1000 Abcam; ab5450), Rabbit anti-Red fluorescent Protein (RFP; 1:1000; Rockland; 600401379).

Secondary antibodies used (*fluorophore; dilution; company; catalogue number*): Donkey anti-rabbit (Cy3; 1:1000; Jackson; 711-165-152), Donkey anti-rabbit (Cy5; 1:1000; Jackson; 711-175-152), Donkey anti-chicken (Cy3; 1:1000; Jackson; 703-165-155), Donkey anti-chicken (Cy5; 1:1000; Jackson; 703-175-155), Donkey anti-guinea pig (Alexa Flour 405; 1:500; Jackson; 705-475-147), Donkey anti-guinea pig (Cy3; 1:1000; Jackson; 706-165-148), Donkey anti-guinea pig (Cy5; 1:1000; Jackson; 706-175-148), Donkey anti-goat (Alexa Fluor 488; 1:1000, Jackson; 705-545-147).

Anatomical and Histological analysis. *Anatomical mapping of GPi-LHb axons and cell bodies.* For viral constructs and strategies see *Viral Constructs*. Tiled whole brain images of sagittal or coronal cut sections

were acquired at 10x magnification and close-up z-stacked images were acquired at 20-63x magnification using a Zeiss (LSM800) confocal microscope.

For the whole-brain axonal patterning plots, each sagittal cut brain section was manually mapped to a ML coordinate and x-y coordinate for each fluorescent image pixel (axonal fluorescent labeling) was segmented out in ImageJ. Cell bodies were manually mapped using the CellCounter in ImageJ. A 2D reconstruction of the axonal arborization/cell bodies was thereafter plotted within the Allen Reference Atlas framework in R.

For axon location confirmation within a target region, immunohistological labeling was performed on coronally sectioned brains.

*Neurotransmitter labeling in axonal terminals***.** Immunostaining was performed to quantify the co-labeling of GPi boutons with Vgat and Vglut2. GPi boutons were identified by injecting AAV-Con/Fon-GCaMP6m directly into the GPi of Vglut2-Cre/Vgat-Flp mice. 1-3 Z-stack images were captured of each brain region of interest on a Zeiss (LSM800) confocal microscope. Images were acquired using identical pinhole, gain, and laser settings for all brain regions and analyzed using ImageJ and/or Imaris software. Boutons were identified by GCaMP6m expression and the fluorescence intensity of the Vglut2 and Vgat stainings within the area of each identified bouton was measured.

Fiber placement and virus expression in behavioral cohorts. The brains of mice in the fiber photometry and tetanus toxin cohorts were collected as specified in *General procedure*, cut (60μm thickness) coronally and all consecutive sections covering the GPi and adjacent brain regions were collected and mounted. Assessment of fiber placement was based on the lesion from the fiber in the tissue. Virus expression level and specificity was manually evaluated and animals with misplaced fibers or lack of virus expression/specificity were excluded from the study.

Behavior. *Fiber photometry and locomotor tracking.* The fiber photometry calcium imaging setup was as previously reported [40]. We acquired calcium-dependent fluorescence, elicited at jGCaMP7's [17] optimal excitation wavelength (470 nm) and indicative of neuronal activity, as well as interleaved calcium-independent control signals at the calcium indicator's isosbestic wavelength (405 nm). The light power of the 470 nm LED was dialed to about 20-30 μW, with the power of the 405 nm LED ranging slightly lower, usually between 10-25 μW.

In conjunction with the photometry signals, we captured videos of the behaving mice and, using the DeepLabCut software [41,42] post-hoc, extracted the xy-coordinates of three body parts: the base of the tail, the center of the body, and the snout/head. For alignment purposes, video and photometry frames, as well as important task events, were timestamped, usually using the Bonsai software [43], which communicated with external hardware via an Arduino Micro microcontroller (Arduino) running the firmata protocol [44]. The photometry traces and the behavior videos were recorded at varying frame rates (see below). Before further processing, we therefore uniformly downsampled all photometry and locomotor tracking data recorded at higher frame rates to the 10 Hz rate of the majority of the photometry recordings, by binning and averaging.

Each photometry recording was further preprocessed as follows: The 405 nm control signal was first scaled to the 470 nm neuronal activity signal, using linear least-squares regression, and then subtracted from it. This removed motion and autofluorescence-related artifacts. Next, we estimated the baseline fluorescence of the raw activity signal over time, by fitting a least-squares regression line to the scaled control signal, and then divided the corrected activity signal at each time point by the corresponding raw baseline estimate. This resulted in a normalized and photobleaching-adjusted, motion and autofluorescence-corrected neuronal activity trace ($\Delta F/F$). Finally, we computed the rolling z-score (5 min moving window) of this trace to remove very slow activity transients, as well as to standardize the measurements across animals (F in our figures).

The downsampled spatial trajectories of all three tracked body parts were further preprocessed using a Kalman smoother, assuming constant movement velocity frame-to-frame. Based on the smoothed trajectories, we computed the velocity of the body mid-point (cm/s) as well as its angular velocity (°/s) relative to the base of the tail. To segment out left and right turns, forward motion and pauses, we used our published optimization algorithm [45].

Arrow maze reversal task. The arrow maze consisted of a 40x40 cm box with a clear plastic floor and a number of removable inserts. To fashion 5 cm-wide corridors along the 4 walls of the box and a diagonal corridor crossing the center, we inserted two "half pyramid"-shaped (triangular base, 2 sides slanted) inserts of clear plastic into the box. The slanting of the inserts' walls towards the center of the box allowed the photometry patch cable to move freely without catching on corners or edges. Inserts equipped with water spouts were placed in three corners of the box, blocking off a corner at one end of the center-crossing corridor and the two corners next to it. The open fourth corner and the three adjacent corridors formed the

"arrow-maze" (a pointed T-maze). Various visual cues, visible through the clear plastic throughout the arena, aided navigation. Facing the 3-way corner, the box's walls left of the center corridor were white, the walls on the right black. Moreover, each wall featured additional visual cues of different shapes and colors and a DUPLO panda figure (Lego) was placed at the center of the left half pyramid insert. The whole arena was illuminated by a white LED light strip glued along the rim of the clear plastic floor.

To motivate foraging for water rewards, mice were water-restricted, receiving 1 ml of water per day in total (in-task and researcher-supplied). Every 12 days of water restriction were followed by 2 days of rest with ad-libitum water access.

Water-restricted mice were first habituated to the arrow maze for 3 days. During habituation we removed the spout-equipped inserts, thus making all corridors and corners accessible. In the following 14 experimental days (i.e. excluding rest days), mice were trained to shuttle back and forth between the spouts at the ends of the center "initiation" corridor and one of the "choice" corridors, where water rewards (13 μl) were delivered upon approach in alternating fashion. Approaching the spout in the "incorrect" choice corridor reset the reward location to the initiation spout. For the fiber-implanted mice, the correct choice corridor was on the side contralateral to their implant, whereas for the tetanus toxin-expressing mice, the correct choice was the right corridor, at this stage. The two choice corridors swapped roles, the rewards now being obtainable in the formerly incorrect arm, in the subsequent reversal phase of 6 (photometry group) or 14 (tetanus group) days. Other than the rewards, no cues signaled the correct choice or the reversal. All sessions lasted 20 minutes.

The task was implemented using the Bonsai software [43], used to detect entries into the corridor end zones from a live video feed, and two Arduino Micro microcontrollers (Arduino), controlling the water spouts and communicating with Bonsai via the firmata protocol [44], respectively. Video frames were acquired at 60 Hz using an Oryx 10GigE camera (Flir) placed below the arena. Fiber photometry fluorescence data was acquired at 10 Hz.

Nose-poke probabilistic reversal task. Food-restricted mice, kept above 85% of their free-feeding body weight, were trained on an established probabilistic reversal task [11,14,46]. In brief, mice initiated trials at a central nose poke port, then chose between two identical ports located to its left and right. A nose poke into the "correct" port resulted in a liquid sucrose reward (15%, 3.75 μl) with 75% probability to be delivered,

whereas the "incorrect" port yielded nothing. The roles assigned to the choice ports reversed every 7-23 rewarded trials, at random and without indication. LED lights mounted inside the ports signaled only whether the initiation or choice ports were actionable. Nose pokes were recorded as infrared beam breaks by a pyboard microcontroller (pyboard.org), which operated the 15x15 cm custom-built operant chamber. The pyboard also registered every recorded photometry and video frame. Training and recording sessions lasted 30 to 60 minutes. A BlackFly USB3 camera (Flir) captured video from above at 20 or 40 Hz, whilst fiber photometry fluorescence data was acquired at half these rates.

We fitted relative action values using Q-learning, as reported previously [45]. The relative action value was computed as $\beta \times (Qr - Ql) + b$, where Qr and Ql, β, and b denote the Q-values of the left and right ports, the explore-exploit trade-off, and the static port bias, respectively. The relative action values are the log-odds of the animal choosing the right port, such that negative values favor left port and positive values favor right port choices. Reward prediction errors were calculated as the relative action value shift in favor of the action taken on each trial. Importantly, a relative action value decrease was treated as a positive error, and a relative action value increase was treated as a negative error, on leftward but not rightward trials.

Auditory fear conditioning. The conditioned stimulus (CS) was a 10 s, 4kHz, pulsed tone (250 ms on/off) measuring 55 dBA inside the chamber; the unconditioned stimulus (US) was a 1 s, 0.3 mA foot shock. Each day, mice were exposed to 20 presentations of the CS, separated by random inter-trial-intervals (10-70 s). On days 1 and 2 (habituation), no US was delivered. On day 3 (conditioning), every CS was followed by the US. On day 4 (extinction), a plastic mat covered the grid floor. In this novel environment, the CS was once again presented without the US.

The task was controlled by an Arduino Micro microcontroller (Arduino) equipped with a piezo buzzer. The shocks were generated by a TTL-triggered aversive stimulator (ENV-414S, Med Associates) connected to the grid floor of the conditioning chamber (ENV-307W-CT, Med Associates). The chamber was located inside an unilluminated, sound-attenuating cubicle (ENV-016MD, Med Associates). Videos were recorded from above at 30 Hz, using an infrared webcam. Fiber photometry fluorescence data was acquired at 10 or 20 Hz.

Open field. Overnight food or water-deprived mice were left to freely explore a 49x49 cm box with a clear plastic floor for 15 to 30 minutes. Except for red and infrared light sources, all lights in the room were

dimmed. A BlackFly USB3 or an Oryx 10GigE camera (both Flir) recorded video from below at 20, 40, or 60 Hz. Fiber photometry fluorescence data was acquired at 10 or 20 Hz.

Electrophysiology. Vglut2-Cre/Vgat-Flpo mice were injected with AAV-Con/Fon-ChR2-EYFP at 8-9 weeks and recorded 10-12 weeks of age (for details on the virus injection see "Viral injections and implants" above. 250 µm thick coronal slices were cut with a vibratome (VT1200S, Leica, Germany) in ice-cold cutting solution, containing (in mM): 40 NaCl, 2.5 KCl, 1.25 NaH2PO4, 26 NaHCO3, 20 glucose, 37.5 sucrose, 20 HEPES, 46.5 NMDG, 46.5 HCl, 1 L-ascorbic acid, 0.5 CaCl2, 5 MgCl2. The slices were incubated in cutting solution at 34°C for 13 minutes, and then maintained at room temperature in extracellular solution until recording, containing (in mM): 124 NaCl, 2.5 KCl, 1.25 NaH2PO4, 26 NaHCO3, 20 glucose, 2 CaCl2, 1 MgCl2.

For recording, the slices were superfused with extracellular solution at 33-35°C. Neurons were visualized using a 60x objective (Olympus, Tokyo, Japan) and a DIC microscope (Scientifica, Uckfield, UK). Patch pipettes (resistance 7-10 MΩ), pulled using a horizontal puller (P-87 Sutter Instruments, Novato, CA, USA) were filled with Cs-internal solution containing (in mM): 130 CsMeSO3, 4 QX-314, 5 Na2-phosphocreatine, 1.5 MgCl2, 10 HEPES, 5 Mg-ATP, 0.35 Na-GTP, 1 EGTA, 8 biocytin for voltage clamp and with 130mM K-gluconate, 5mM KCl, 10mM HEPES, 10mM Na2-phosphocreatine, 4mM ATP-Mg, 0.3mM GTP-Na internal solution for current-clamp recordings. Signals were recorded with an Axon MultiClamp 700B amplifier and digitized at 20 kHz with an Axon Digidata 1550B digitizer (Molecular Devices, San Jose, CA, USA). Access resistance and pipette capacitance were compensated for. In voltage clamp experiments we bath applied TTX (1uM; Tocris) and 4-AP (5 mM; Sigma-Aldrich) to isolate monosynaptic responses. During optogenetic stimulation protocols in all experiments the membrane potential was first depolarized to -40 mV to detect both excitatory and inhibitory currents, then to 0 mV to isolate inhibitory currents which got pharmacologically blocked by bath application GABAA antagonist Gabazine (10uM; Sigma-Aldrich). Then cells were hyperpolarized to -80 mV to isolate AMPA currents which were then were blocked by NBQX (20uM; Tocris) and finally the cells got depolarized to +40 mV to measure NMDA-mediated currents which were blocked by the selective NMDA antagonist AP-V (50 uM; Tocris). The synaptic properties of GPi Vglut2+/Vgat+ projections onto GPe neurons were probed using multiple light pulse train protocols with 3 ms blue light pulses with ~2.5 mW light power from a pE-4000 (CoolLED, Andover, UK). light source. To assess passive and active membrane properties, neurons recorded in current-clamp mode were held at a membrane potential of -70 mV. The rheobase current was

determined by applying near-threshold current steps, followed by steps proportional to the rheobase current with a duration of 1 s. Intrinsic electrical properties, EPSC and IPSC amplitudes were extracted using a custom-written Matlab (MathWorks, Natick, MA, USA) script.

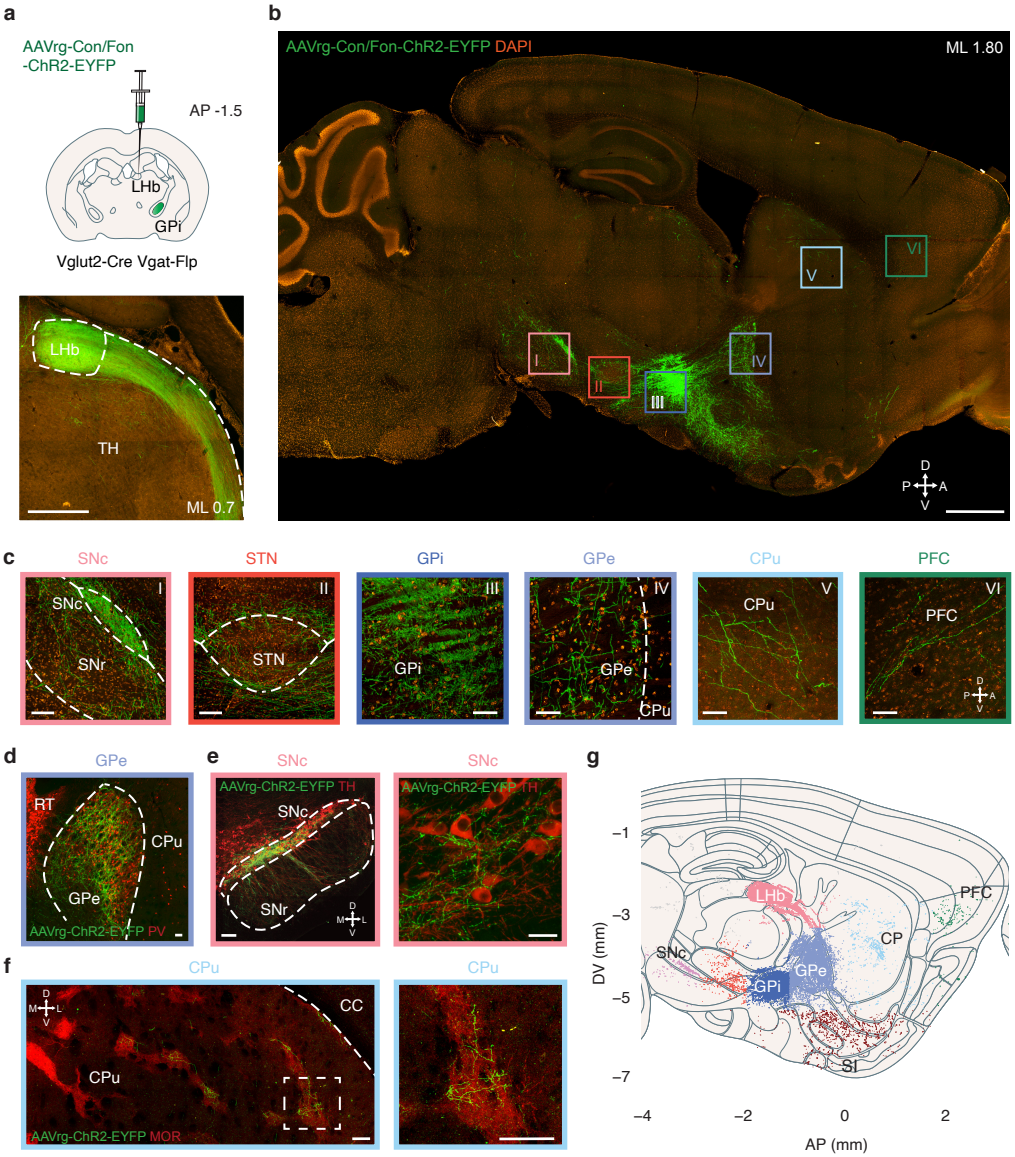

Figure 1 | Collateral projection targets of LHb-projecting Vglut2+/Vgat+ GPi neurons.

a, Experimental strategy to achieve selective ChR2-EYFP expression in LHb-projecting Vglut2+/Vgat+ GPi neurons (Vglut2+/Vgat+$^{GPi\text{-}LHb}$).

b, Major collateral projections of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ neurons. Left: axons in the LHb. Green: AAVrg-ChR2-EYFP; orange: DAPI. Roman numerals indicate regions of interest in **c**. Scale bars: 1mm [overview], 500µm [left].

c, Close-up images of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ cell bodies in the GPi (III) and axonal projections in the SNc (I), STN (II), GPe (IV), CPu (V) and PFC (VI). Roman numerals indicate brain regions marked in **b**. Green: AAVrg-ChR2-EYFP; orange: DAPI. Scale bars: 100µm [SNc, STN]; 50µm [GPi, GPe, CPu, PFC].

d, Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons in the GPe confirmed by Parvalbumin (PV) immunostaining. Green: AAVrg-ChR2-EYFP; red: PV. Scale bar: 100µm.

e, Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons in the SNc confirmed by tryptophan hydroxylase (TH) staining. Green: AAVrg-ChR2-EYFP; red: TH. Scale bars: 100µm [left]; 50µm [right].

f, Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axonal collaterals target striosomes in the CPu. Striosomes identified by µ-Opioid Receptor 1 (MOR) staining. Green: AAVrg-ChR2-EYFP; red: MOR. Scale bars: 100µm.

g, Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons throughout the brain, colored based on their target brain region (as in ARA). Intrahemispheric brain sections (n = 15) superimposed onto the ARA plate #100884129 (~1.95 ML).

Representative images in **b**, **c** and **d**, **e** and **f** replicated in N = 3 mice and N = 4 mice, respectively, see also **Supplementary Figure 3**.

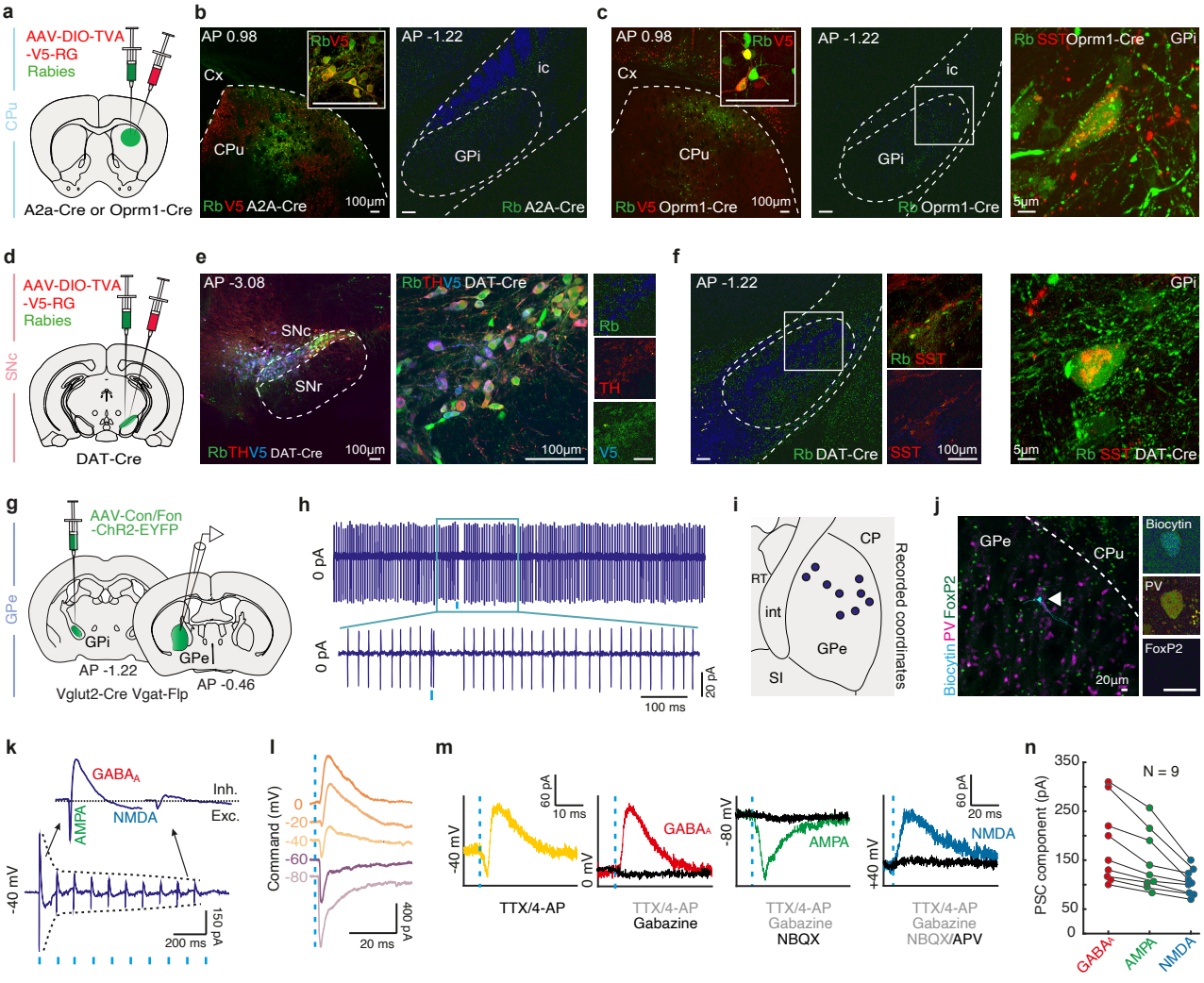
a
AAV-DIO-TVA
-V5-RG
Rabies
CPu
A2a-Cre or Oprm1-Cre
b
AP 0.98
Cx
CPu
Rb V5 A2A-Cre
100µm
AP -1.22
ic
GPi
Rb A2A-Cre
c
AP 0.98
Cx
CPu
Rb V5 Oprm1-Cre
100µm
AP -1.22
ic
GPi
Rb Oprm1-Cre
Rb SST Oprm1-Cre
GPi
5µm
d
AAV-DIO-TVA
-V5-RG
Rabies
SNc
DAT-Cre
e
AP -3.08
SNc
SNr
Rb TH V5 DAT-Cre
100µm
Rb
TH
V5
f
AP -1.22
Rb DAT-Cre
Rb SST
SST
100µm
GPi
5µm
Rb SST DAT-Cre
g
AAV-Con/Fon
-ChR2-EYFP
GPe
GPi
AP -1.22
GPe
AP -0.46
Vglut2-Cre Vgat-Flp
h
0 pA
0 pA
20 pA
100 ms
i
CP
RT
int
GPe
SI
Recorded coordinates
j
GPe
CPu
Biocytin PV FoxP2
20µm
Biocytin
PV
FoxP2
k
GABA$_A$
Inh.
Exc.
AMPA
NMDA
-40 mV
150 pA
200 ms
l
Command (mV)
0
-20
-40
-60
-80
400 pA
20 ms
m
60 pA
10 ms
-40 mV
0 mV
TTX/4-AP
GABA$_A$
TTX/4-AP
Gabazine
-80 mV
AMPA
TTX/4-AP
Gabazine
NBQX
60 pA
20 ms
+40 mV
NMDA
TTX/4-AP
Gabazine
NBQX/APV
n
N = 9
PSC component (pA)
350
250
150
50
0
GABA$_A$
AMPA
NMDA

Figure 2 | GPi-LHb neurons provide synaptic feedback in the basal ganglia and co-transmit glutamate and GABA.
a, AAV-DIO-TVA-V5-RbG (V5) and Rabies-EGFP virus (Rb) were injected into the dorsolateral CPu of A2A-Cre and Oprm1-Cre mice, 21 days apart.
b, Representative image of starter neurons co-expressing V5 (red) and Rb (green) in the CPU of an A2A-Cre mouse. Insert: close-up of starter neurons. Right: No Rb neurons in the GPi. Scale bars: 100µm. Replicated in N = 4 A2A-Cre mice; for cell counts see **Supplementary Table 1**. See also **Supplementary Figure 5a**.
c, Representative image of starter neurons co-expressing V5 (red) and Rb (green) in the CPU of an Oprm1-Cre mouse. Insert: closeup of starter neurons. Middle: Rb presynaptic neurons in the GPi. Right: Somatostatin+ (SST; red) Rb (green) neuron in the GPi. Scale bars: 100µm [left, middle]; 5µm [right]. Replicated in N = 6 Oprm1-Cre mice; for cell counts see **Supplementary Table 1**. See also **Supplementary Figure 5b**.
d, AAV-DIO-TVA-V5-RbG (V5) and Rabies-EGFP virus (Rb) were injected into the SNc of DAT-Cre mice, 21 days apart.
e, Representative image of starter neurons co-expressing the V5 (blue), Rb (green), and TH (red) in the SNc of a DAT-Cre mouse. Right: closeup of starter neurons. Scale bars: 100µm.
f, Representative Rb GPi neurons (green) targeting DAT+ SNc neurons. Close-up images to the right show SST+ (red) Rb (green) GPi neurons. Right: Close-up of one SST+ Rb GPi neuron. Scale bars: 100µm [left], 5µm [right]. Replicated in N = 3 DAT-Cre mice; for cell counts see **Supplementary Table 1**.
g, AAV-Con/Fon-ChR2-EYFP was unilaterally injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice. ChR2-expressing GPi axon terminals were activated by blue light and postsynaptic currents in GPe neurons were recorded in whole-cell voltage-clamp mode.
h, Optogenetic activation (3ms, blue tick mark) of GPi axons results in an action potential doublet followed by a silent phase in a postsynaptic GPe neuron (replicated in n = 3 neurons). Scale bars: 100 ms, 20 pA.
i, Location of each recorded GPe neuron responsive to GPi axonal activation (n = 9).
j, Responsive GPe neuron (white arrowhead) filled with biocytin (cyan) co-expresses PV (pink) but not FoxP2 (green), as evident from the close-up images shown to the right. Scale bars: 20µm. Replicated in n = 9 recorded responsive GPe neurons. See also **Supplementary Figure 5c**.
k, Response of a GPe neuron to optogenetic stimulation (10 Hz, 3 ms, blue tick marks) of GPi axon terminals. The envelope (dashed line) indicates a symetrical short-term depression of the GPi-GPe synapse. Upper traces: Excitatory (AMPA, NMDA) and inhibitory ($GABA_A$) postsynaptic current components of the first and tenth pulse (replicated in n = 3 neurons). Holding potential was -40 mV.
l, Optogenetically evoked postsynaptic currents at different holding potentials. Averages of 5 traces for n = 9 neurons.
m, Pharmacological dissection of postsynaptic currents evoked by opto-stimulation (blue dashed line). Gabazine (10 µM), NBQX (20 µM) and APV (50 µM) were applied sequentially. TTX (1 µM) and 4-AP (5 mM) were applied throughout the entire recording. Averages of 5 traces for n = 9 neurons.
n, Maximal current amplitude of $GABA_A$, AMPA- and NMDA-currents measured at the holding potentials shown in **m**.

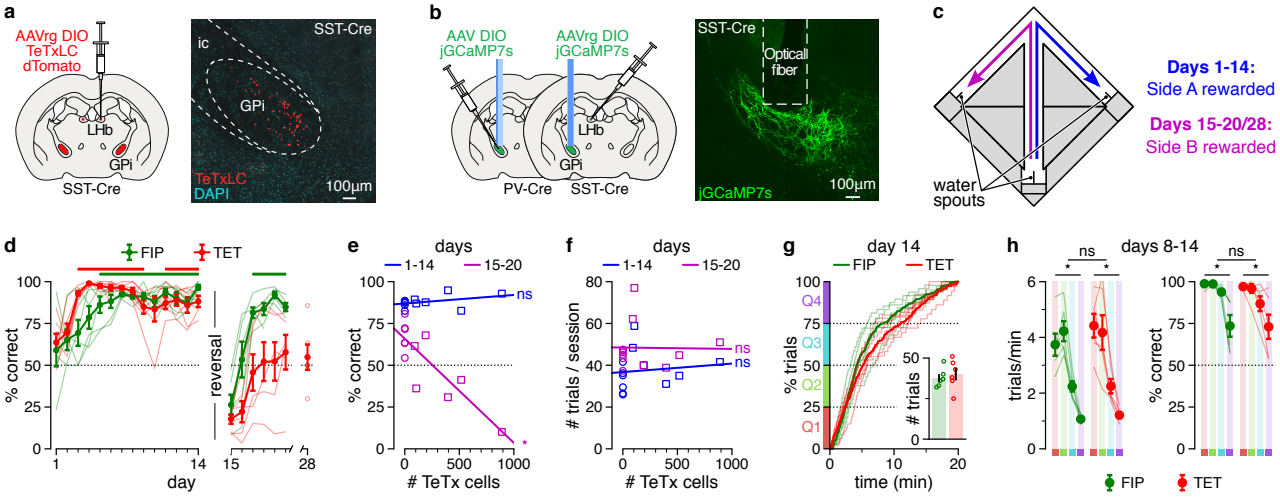

Figure 3 | SST+ GPi neurons drive behavioral flexibility.

a, AAVrg-DIO-TeTxLC-dTomato was injected bilaterally in the LHb of SST-Cre mice. Right: TeTxLC (red) and DAPI (cyan) in the GPi of SST-Cre mice. Scale bar: 100µm.

b, AAVrg-DIO-jGCaMP7s was injected uni- or bilaterally in the LHb of SST-Cre mice. AAV-DIO-jGCaMP7s was injected unilaterally in the GPi of PV-Cre mice. 200µm fibers were implanted above the GPi. Right: jGCaMP7s (green) in the GPi of a SST-cre mouse. Scale bar: 100µm. See also **Supplementary Figure 6** and **Supplementary Table 2**.

c, The arrow maze reversal task. For 14 days, the choice to turn right (TeTxLC mice), or contraversive relative to the imaged hemisphere (jGCaMP7s mice), was reinforced with a water drop (13 µl) at the end of the corridor (blue arrow). On day 15 and thereafter, the rewarded choice was reversed (magenta arrow). See **Methods** for details.

d, Accuracy of the fiber photometry (FIP, green) and TeTxLC-expressing (TET, red) groups in the maze task. Mean ± SEM. *: $p<0.05$, Bonferroni-corrected, 1-sided t-test (greater than 50% correct). N = 6 mice per group.

e, Mice's average accuracy pre- (blue) and post-reversal (magenta) regressed on the number of TeTxLC-expressing cells (lines). Circles: individual FIP mice; squares: individual TET mice. *: $p<0.001$, t-test of the slope of the regression line. N = 12 mice.

f, Average number of trials completed per session pre- (blue) and post-reversal (magenta) regressed on the number of TeTxLC-expressing cells. Circles: individual FIP mice; squares: individual TET mice. ns: $p>0.05$, t-test of the slope of the regression line. N = 12 mice.

g, The fraction of trials completed plotted against time for training day 14. Colorbar on y-axis: quartile bins (Q1-4), defined by the quartiles of the trial number distributions. Transparent lines: individual FIP (green) and TET (red) mice, opaque lines: group averages. Inset: average number of trials completed in the session, by group. Circles: individual mice. Mean ± SEM. N = 6 mice per group.

h, Left: Average trial rates in each trial quartile bin (background shading, colors as in g). Right: Average accuracy in each trial quartile bin, by group. Transparent lines: individual FIP (green) and TET (red) mice, points: group averages. Mean ± SEM. *: $p<0.001$, repeated-measures ANOVA, within groups. ns: $p>0.05$, t-test for each quartile, between groups. N = 6 mice per group.

See **Supplementary Table 3** for detailed statistical values for **d–f** and **h**.

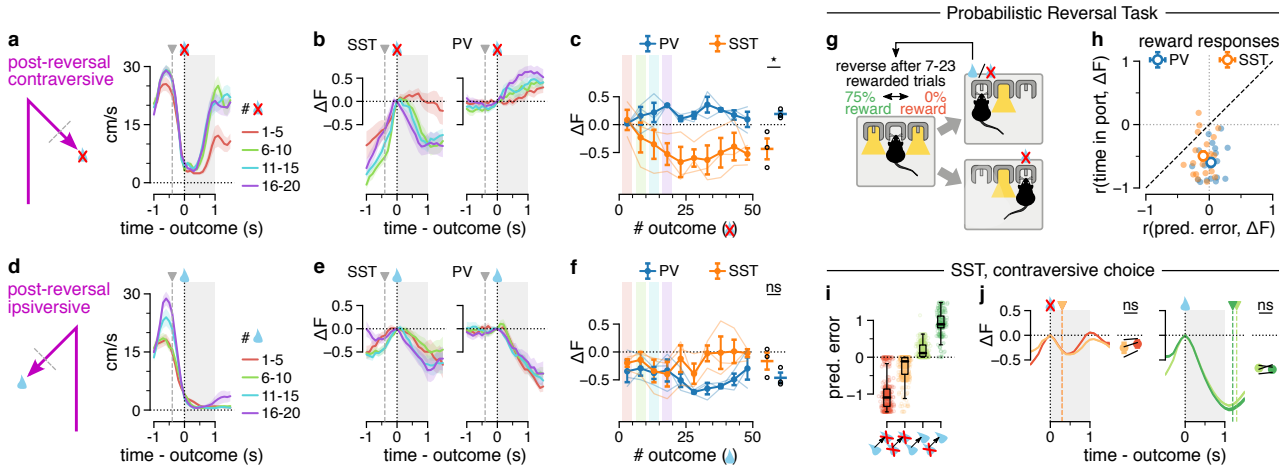

Figure 4 | SST+ GPi neuronal activity does not reflect the updating of value.

a, Left: post-reversal, entering the end zone of the contraversive choice corridor (grey dotted line) did not produce the expected water rewards. Right: movement velocities during the first 20 unrewarded contraversive end zone entries. Trace color-coding: bins of 5 trials. Grey shading: 1s time window in which mice consume the reward, if available. Mean ± SEM. n = 30 trials (pooled, N = 6 mice).

b, Outcome-aligned SST (left) and PV (right) population activity traces during the first 20 unrewarded error trials post-reversal. ΔF = F − F0 where F0 is the fluorescence at time 0 (reaching the water spout). Trace color-coding and grey shading: see **a**. Mean ± SEM. n = 15 trials (pooled, N = 3 mice per group).

c, Mean PV (blue) and SST (orange) population responses to the first 50 unrewarded outcomes post-reversal. Mean ΔF computed over the 1s time window shaded grey in **a** and **b**. Bins of 5 trials. Shading: color indicates bins used in **a** and **b**. Mean ± SEM. Right plot: comparison of the group means over all 50 trials. Circles: individual mice. *: p<0.05, t-test. N = 3 mice per group.

d–f, Same as **a–c**, but for the first rewarded outcomes obtained post-reversal by entering the ipsiversive choice corridor end zone.

g, Probabilistic reversal task. After trial initiation at the center port, a reward becomes available (with 75% probability) at one of the two peripheral ports. The rewarded choice reverses unpredictably every 7-23 rewarded trials.

h, Mean fluorescence change during rewarded nose pokes, correlated with (x-axis) the reward prediction error versus (y-axis) the time spent in the port. ΔF = F − F0 where F0 is the fluorescence at port entry. Filled circles: correlation coefficients for individual PV (blue) and SST (orange) sessions. Open circles: group averages. PV/SST n = 18/20 sessions (pooled, N = 3 mice per group).

i, Reward prediction errors plotted for all two-trial outcome histories. Open circles: single trials. Boxplot whisker range: maximum/minimum value within quartiles ± interquartile range. n = 405/589/365/210 trials (pooled, N = 3 mice).

j, SST ΔF (see **h**) responses to omissions (left) and rewards (right), split by outcome on the previous trial. Color coding as in **i**. Triangles and dashed lines: median port exit time points. Grey shading: time window used to compute means for the paired t-tests shown to the right. Black lines: individual mice. Mean ± SEM. Traces: n = 405/589/365/210 trials (pooled, N = 3 mice). ns: p>0.05, paired t-tests. N = 3 mice.

See **Supplementary Table 3** for detailed statistical values for **c–f** and **j**.

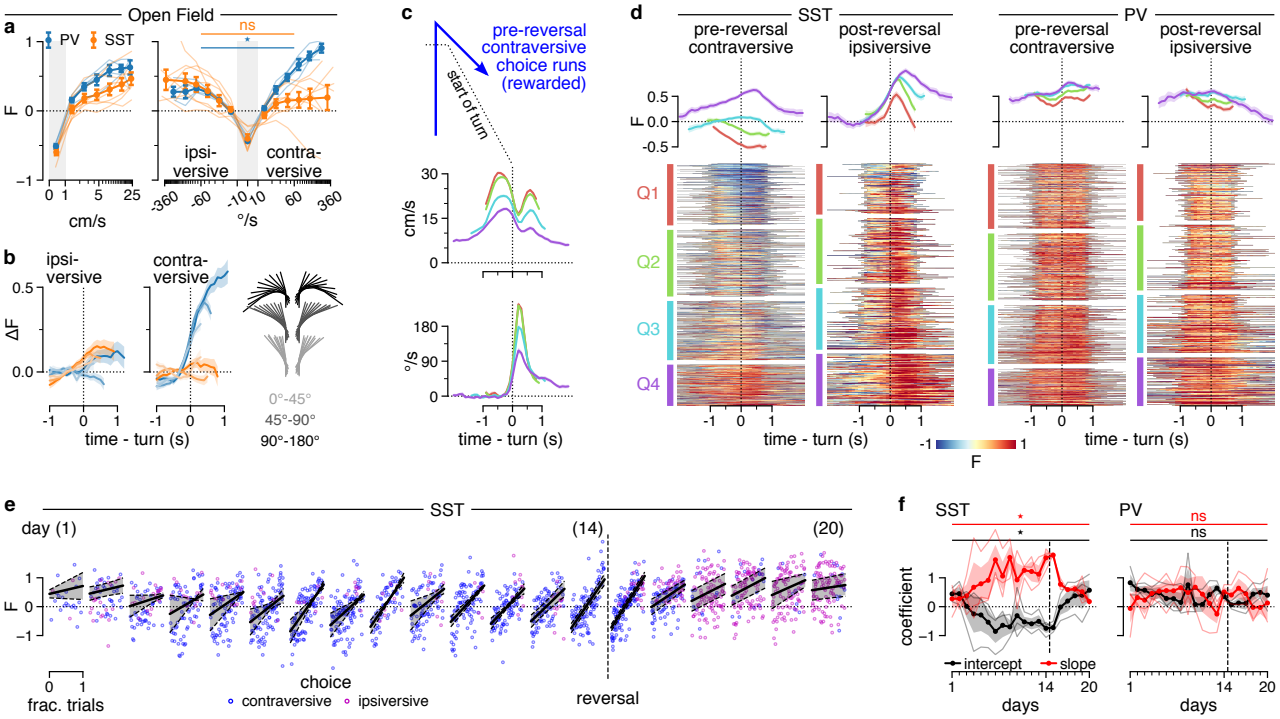

Figure 5 | Flexible choice coincides with the disinhibition of the SST+ GPi population.

a, Mean PV (blue) and SST (orange) activity at different velocities (left) and angular velocities (right) in the open field. The bins' edges were spaced evenly along the log-scaled x-axes. Grey shading: sections of x-axes scaled linearly. Mean ± SEM. *: p<0.05, paired t-test of the log(ipsiversive °/s) X F versus log(contraversive °/s) X F correlation coefficients. PV/SST N = 4/7 mice.

b, PV (blue) and SST (orange) activity during ipsi- (left) and contraversive (middle) turns, grouped by the total angle of the turn. ΔF: F - mean F in the second prior to turn onset. The mean activity traces end at the grouped turns' median end time points. Mean turn trajectories (right): Lines delineate the mice's average posture (tail base to body midpoint to snout) in one fraction of the turn (10 bins). Mean ± SEM. PV/SST n = 811/2480 turns total (pooled, PV/SST N = 4/7 mice).

c, Average velocities (middle) and angular velocities (bottom) during pre-reversal contraversive choice runs (top) in each trial quartile bin (see **Figure 1f**), aligned to turn entry. The traces are cut at the median initiation zone exit (run start) and reward zone entry (run end) time points. Mean ± SEM. n = 586/622/549/366 trials (pooled, N = 6 mice).

d, SST and PV population activity during all individual correct choice runs, recorded pre-reversal (contraversive) and post-reversal (ipsiversive). The trials were sorted by fraction of prior trials completed that session and grouped by trial quartile (Q1-4). The mean traces for each quartile, shown above the rasters, were cut as in **c**. Quartile groups are unevenly-sized due to the exclusion of trials in which mice entered the unrewarded corridor. Mean ± SEM. SST/PV n = 1513/1562 trials total (pooled, N = 3 mice per group).

e, SST population activity at the choice turn onset (dashed line in **c** and **d**) plotted against the fraction of trials completed for days 1-20. Dots: single contraversive (blue) and ipsiversive (magenta) trials. Black lines and shading: mean ± SEM of the regression fits for individual mice (N = 3).

f, For each day, the activity at the choice turn onset was regressed on the fraction of trials completed (cf. **e**). The plot shows the coefficients (intercept in black and slope in red) for SST (left) and PV (right) mice across days. Mean ± SEM. *: p<0.05; repeated-measures ANOVA. N = 3 mice per group.

See **Supplementary Table 3** for detailed statistical values for **a** and **f**.

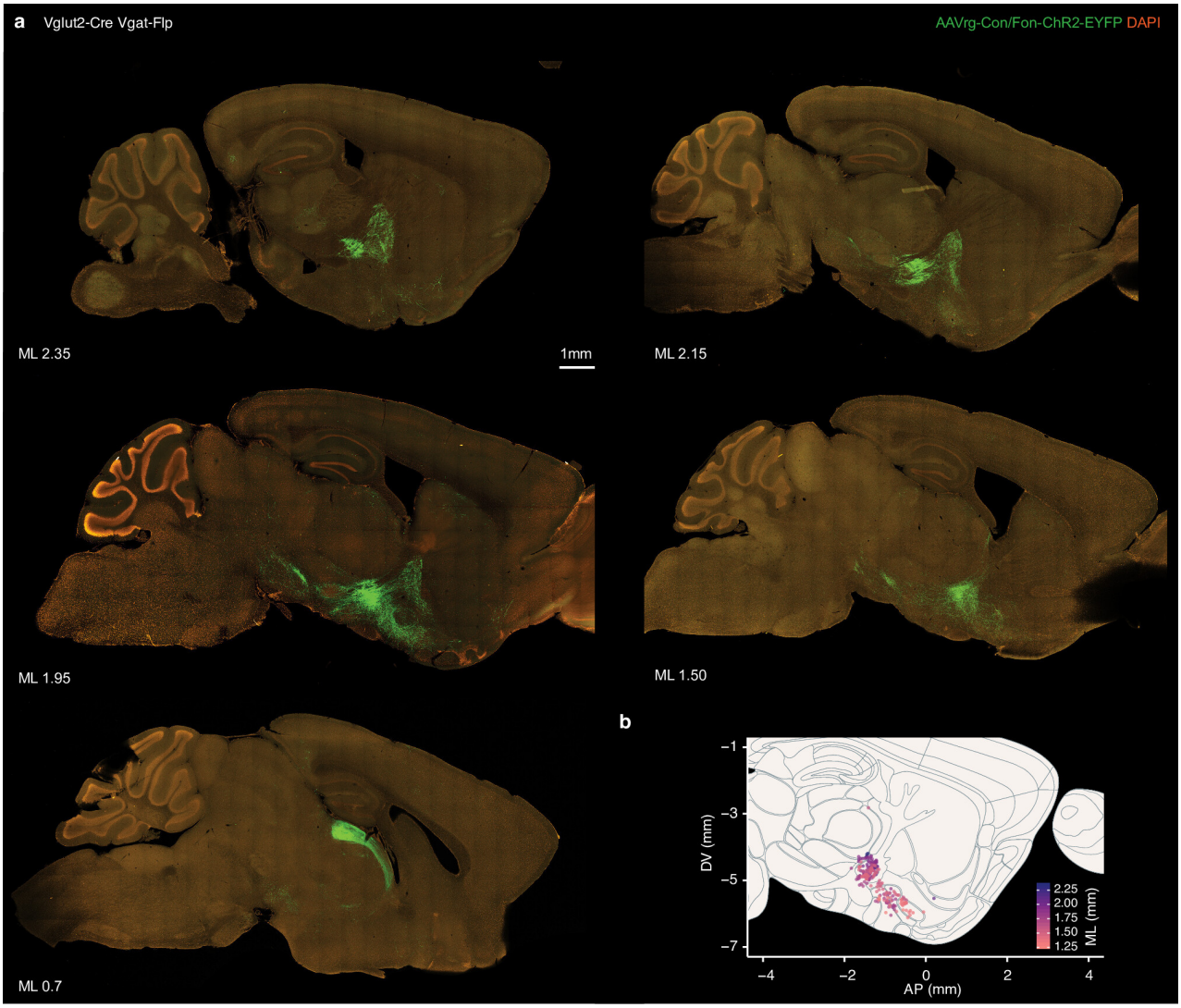

Figure S1 | Vglut2+/Vgat+ LHb-projection neurons can be identified in the GPi and the SI.

a, AAVrg-Con/Fon-ChR2-EYFP (AAVrg-ChR2-EYFP) was unilaterally injected in the LHb of Vglut2-Cre/Vgat-Flp mice. Sagittal sections from the injected hemisphere of a representative mouse show Vglut2+/Vgat+$^{GPi\text{-}LHb}$ axons throughout the hemisphere. Green: AAVrg-ChR2-EYFP; Orange: DAPI. Scale bar: 1mm.

b, LHb-projecting Vglut2+/Vgat+ AAVrg-ChR2-EYFP expressing cell bodies were identified in the GPi and the SI. Each dot represents one cell body identified within the unilateral hemisphere. All neurons are superimposed onto the ARA plate #100884129 (~1.95 ML), each neuron is color-coded by its Medial-Lateral (ML) stereotaxic coordinate (blue: lateral; light pink: medial). n = 21 sections counted.

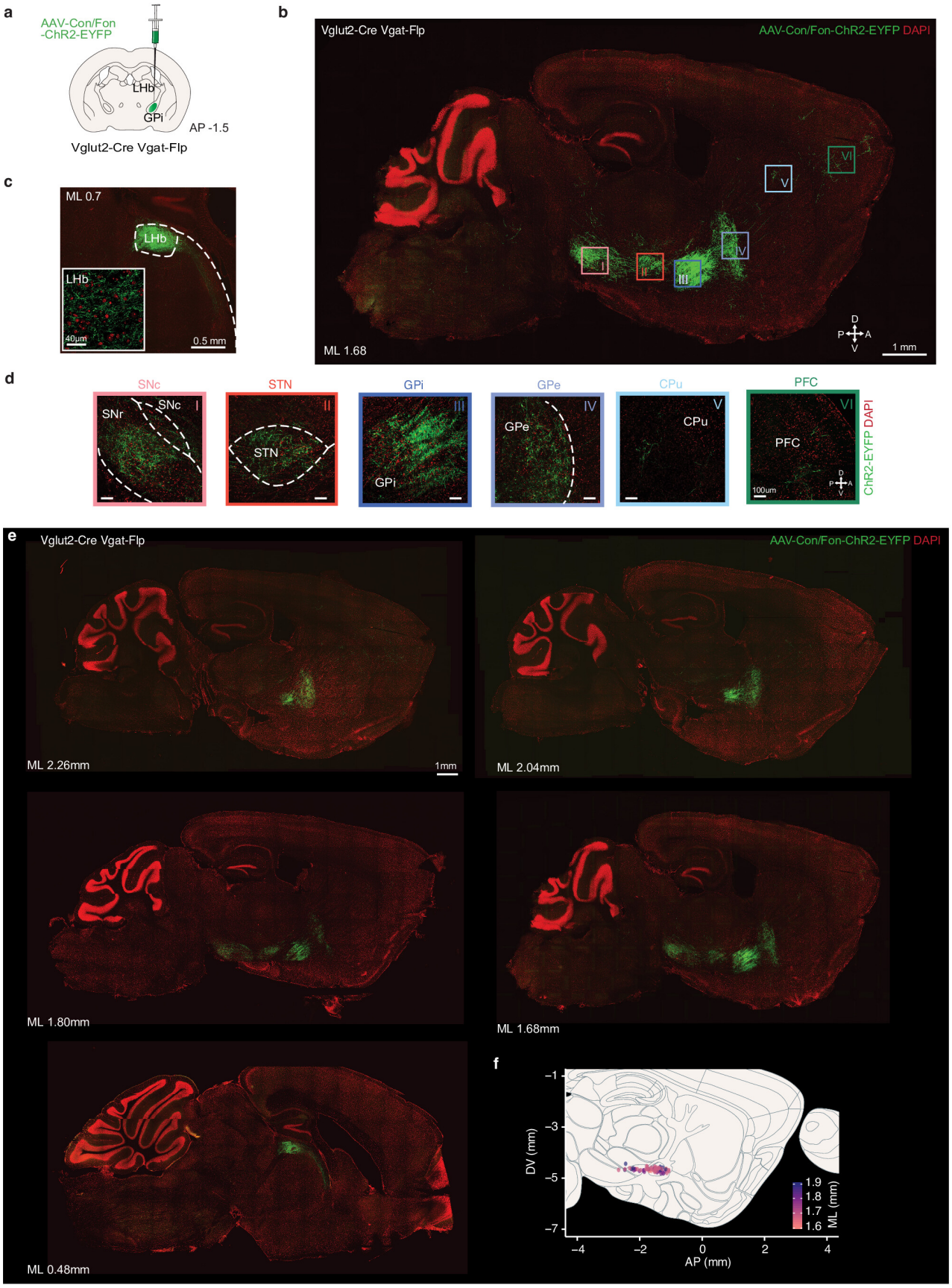
a
AAV-Con/Fon
-ChR2-EYFP
LHb
GPi
AP -1.5
Vglut2-Cre Vgat-Flp
b
Vglut2-Cre Vgat-Flp
AAV-Con/Fon-ChR2-EYFP DAPI
ML 1.68
1 mm
c
ML 0.7
LHb
40μm
0.5 mm
d
SNc
STN
GPi
GPe
CPu
PFC
SNr
ChR2-EYFP DAPI
100um
e
Vglut2-Cre Vgat-Flp
AAV-Con/Fon-ChR2-EYFP DAPI
ML 2.26mm
1mm
ML 2.04mm
ML 1.80mm
ML 1.68mm
ML 0.48mm
f
DV (mm)
AP (mm)
ML (mm)

Figure S2 | Collateral projection targets of GPi Vglut2+/Vgat+ neurons confirmed with a direct viral injection into the GPi.

a, AAV-Con/Fon-ChR2-EYFP (AAV-ChR2-EYFP) was unilaterally injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice.

b, Sagittal section from a representative mouse show GPi Vglut2+/Vgat+ neurons efferent projections throughout the ipsilateral hemisphere. Green: AAV-ChR2-EYFP; Red: DAPI. Roman numerals and colored squares indicate regions of interest displayed in **d**. Scale bar: 1mm.

c, GPi Vglut2+/Vgat+ axons were confirmed in the LHb as described previously (Lazaridis et al. 2019). Green: AAV-ChR2-EYFP; Red: DAPI. Scale bar: 0.5mm, insert: 40µm

d, Close up images of cell bodies in the GPi and afferent projections in the SNc, STN, GPe, CPu and PFC, roman numerals and color code indicate regions marked in **b**. Green: AAV-ChR2-EYFP; red: DAPI. Scale bar: 100µm.

e, Sagittal brain sections from the ipsilateral hemisphere of a representative mouse demonstrate GPi Vglut2+/Vgat+ afferent projections throughout the ipsilateral hemisphere. Green: AAV-ChR2-EYFP, red: DAPI. Scale bar: 1mm. Representative images in **b**, **c**, **d** and **e** replicated in N = 3 mice.

f, Vglut2+/Vgat+ AAV-ChR2-EYFP expressing neurons were identified in the GPi and the STN. Each dot represents one cell body identified within the unilateral hemisphere. All neurons are superimposed onto the ARA plate #100884129 (~1.95 ML), each neuron is color coded by its Medial-Lateral (ML) stereotaxic coordinate (blue: lateral, light pink: medial). n = 9 sections counted.

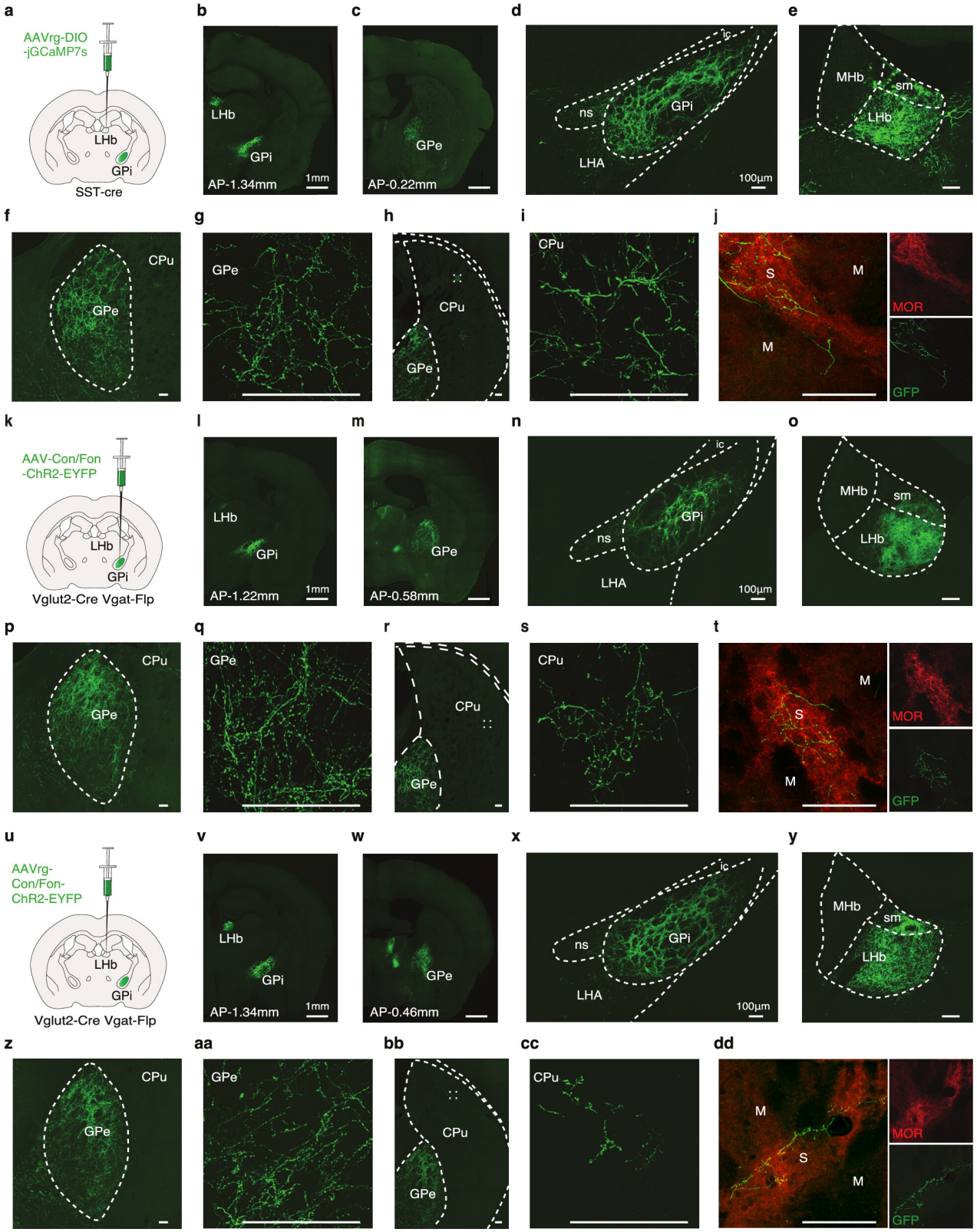
a
AAVrg-DIO
-jGCaMP7s
LHb
GPi
SST-cre
b
LHb
GPi
AP-1.34mm
1mm
c
GPe
AP-0.22mm
d
ic
ns
GPi
LHA
100μm
e
MHb
sm
LHb
f
CPu
GPe
g
GPe
h
CPu
GPe
i
CPu
j
S
M
M
MOR
GFP
k
AAV-Con/Fon
-ChR2-EYFP
LHb
GPi
Vglut2-Cre Vgat-Flp
l
LHb
GPi
AP-1.22mm
1mm
m
GPe
AP-0.58mm
n
ic
GPi
ns
LHA
100μm
o
MHb
sm
LHb
p
CPu
GPe
q
GPe
r
CPu
GPe
s
CPu
t
M
S
M
MOR
GFP
u
AAVrg-
Con/Fon-
ChR2-EYFP
LHb
GPi
Vglut2-Cre Vgat-Flp
v
LHb
GPi
AP-1.34mm
1mm
w
GPe
AP-0.46mm
x
ic
ns
GPi
LHA
100μm
y
MHb
sm
LHb
z
CPu
GPe
aa
GPe
bb
CPu
GPe
cc
CPu
dd
M
S
M
MOR
GFP

Figure S3 | LHb-projecting GPi neurons targeted with three different viral strategies show projection targets.
a, Experimental strategy: AAVrg-DIO-jGCaMP7s was unilaterally injected into the LHb of Sst-Cre mice.
b–j, Representative coronal sections depicting the main projection targets of retrogradely targeted SST+ GPi neurons after injection of AAVrg-DIO-jGCaMP7s in LHb in SST-cre mice. Images from one representative mouse, replicated in N = 9 mice.
k, Experimental strategy: AAV-Con/Fon-ChR2-EYFP was unilaterally injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice.
l–t, Representative coronal sections depicting the main projection targets of directly targeted Vglut2+/Vgat+ GPi neurons after injection of AAV-Con/Fon-ChR2-EYFP in the GPi in Vglut2-cre/Vgat-Flp mice. Images from one representative mouse, replicated in N = 3 mice.
u, Experimental strategy: AAVrg-Con/Fon-ChR2-EYFP was unilaterally injected in the LHb of Vglut2-Cre/Vgat-Flp mice.
v–dd, Representative coronal sections depicting the main projection targets of retrogradely targeted Vglut2+/Vgat+ GPi neurons after injection of AAVrg-Con/Fon-ChR2-EYFP in LHb in Vglut2-cre/Vgat-Flp mice. Images from one representative mouse, replicated in N = 4 mice.
b, d, e, l, n, o, v, x, y, Coronal sections of the injected hemisphere displaying cell bodies in the GPi and axons in the LHb (green) for the three viral strategies, respectively. Scale bar: 1mm (**b, l, v**), close-ups: 100µm (**d, e, n, o, x, y**).
c, f, g, m, p, q, w, z, aa, Axonal arborizations in the GPe of GPi LHb-projecting neurons (green) were confirmed for the three viral strategies implemented. Scale bar: 1mm (**c, m, w**), close-ups: 100µm (**f, g, p, q, z, aa**).
h, i, r, s, bb, cc, GPi LHb-projecting neurons' afferent projections show discrete patches of arborization within the CPu. Close-up images of axonal arborization are indicated with a white square in the overview image. Scale bar: 100µm.
j, t, dd, The three viral strategies all display the afferent projections of GPi LHb-projecting neurons (green) discretely targeting the striosome compartments of the CPu identified with MOR immunostaining (red). Scale bar: 100µm.
Abbreviations; LHA: Lateral Hypothalamus, ns: nigrostriatal bundle, MHb: Medial Habenula, sm: stria medullaris, S: striosome compartment, M: matrix compartment.

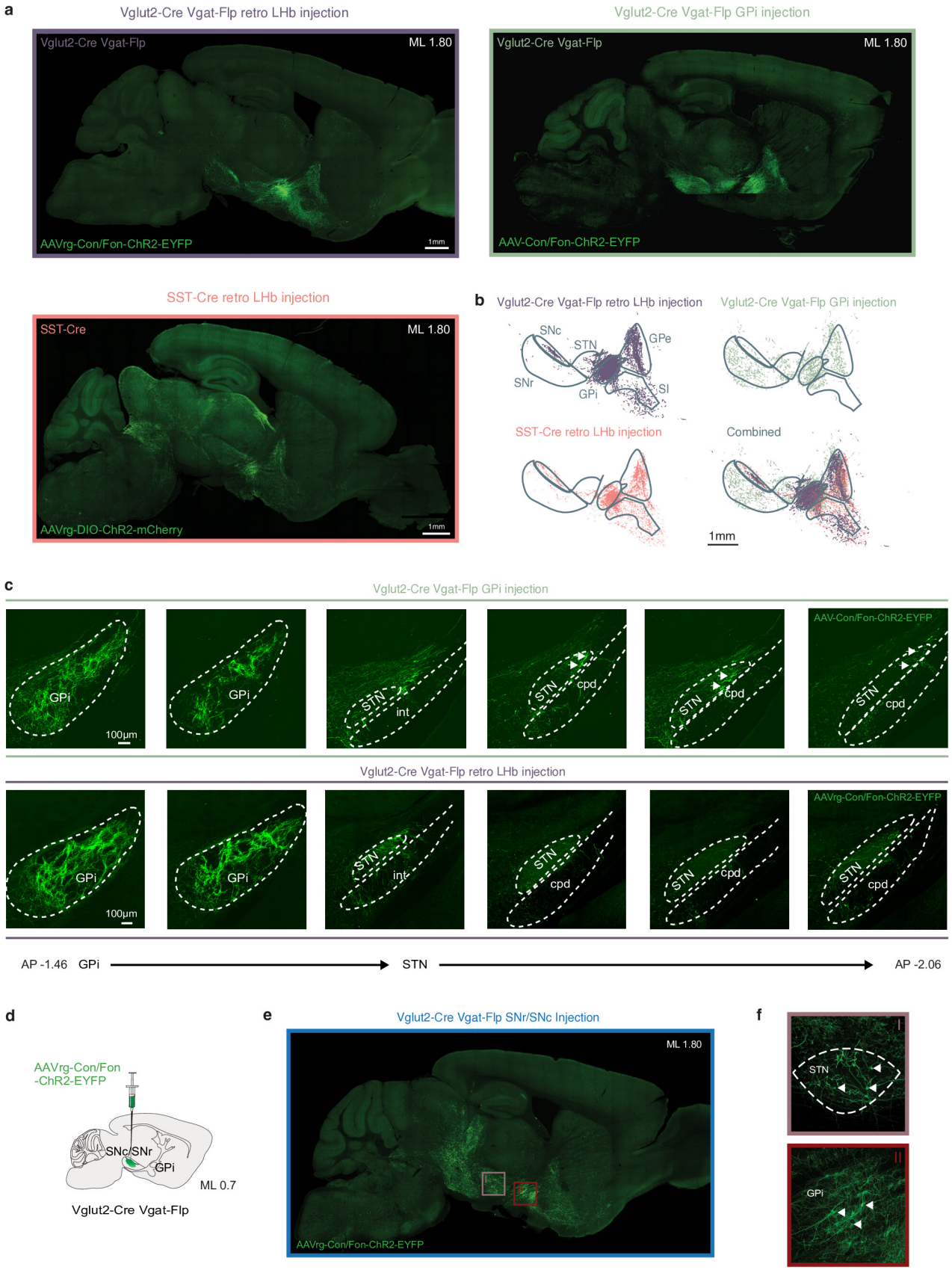
a
Vglut2-Cre Vgat-Flp retro LHb injection
Vglut2-Cre Vgat-Flp
ML 1.80
AAVrg-Con/Fon-ChR2-EYFP
1mm
Vglut2-Cre Vgat-Flp GPi injection
Vglut2-Cre Vgat-Flp
ML 1.80
AAV-Con/Fon-ChR2-EYFP
SST-Cre retro LHb injection
SST-Cre
ML 1.80
AAVrg-DIO-ChR2-mCherry
1mm
b
Vglut2-Cre Vgat-Flp retro LHb injection
Vglut2-Cre Vgat-Flp GPi injection
SNc
STN
GPe
SNr
GPi
SI
SST-Cre retro LHb injection
Combined
1mm
c
Vglut2-Cre Vgat-Flp GPi injection
GPi
100μm
STN
int
cpd
AAV-Con/Fon-ChR2-EYFP
Vglut2-Cre Vgat-Flp retro LHb injection
AAVrg-Con/Fon-ChR2-EYFP
AP -1.46 GPi
STN
AP -2.06
d
AAVrg-Con/Fon
-ChR2-EYFP
SNc/SNr
GPi
ML 0.7
Vglut2-Cre Vgat-Flp
e
Vglut2-Cre Vgat-Flp SNr/SNc Injection
ML 1.80
AAVrg-Con/Fon-ChR2-EYFP
f
STN
GPi

Figure S4 | Direct comparison of three viral strategies targeting LHb-projecting GPi neurons confirm axonal collaterals in SNc but not STN.

a, Representative sagittal section from three viral strategies targeting the Vglut2+/Vgat+ and SST+ GPi LHb-projecting neurons. AAVrg-Con/Fon-ChR2-EYFP unilaterally injected in the LHb of Vglut2-Cre/Vgat-Flp mice (purple), AAV-Con/Fon-ChR2-EYFP unilaterally injected into the GPi of Vglut2-Cre/Vgat-Flp mice (light green) and AAVrg-DIO-jGCaMP7s unilaterally injected into the LHb of Sst-Cre mice (pink), respectively. Scale bar: 1mm.

b, The Vglut2+/Vgat+ or Sst+ GPi-LHb neurons afferent projections from the sagittal brain sections displayed in **a** displayed separately and overlayed for a direct comparison of projection targets. The three viral strategies for targeting Vglut2+/Vgat+ or Sst+ GPi-LHb neurons give rise to efferent projections with differential targeting of the SNr, STN and SI. Viral strategy color coded as in **a**. Both retrograde injections result in axons in the SI (purple and pink), not confirmed with a direct injection in the GPi (light green), while a direct viral injection into the GPi gives rise to axons in the STN not present with a retrograde viral strategy.

c, A direct injection of AAV8-Con/Fon-ChR2-EYFP into the GPi of Vglut2-Cre/Vgat-Flp mice (light green) results in Vglut2+/Vgat+ ChR2-expressing cell bodies within the STN (indicated by white arrows), alas suspected to result in axons in the SNr (as seen in **a**, **b**: light green). AAVrg-Con/Fon-ChR2-EYFP unilaterally injected in the LHb of Vglut2-Cre/Vgat-Flp mice (purple) display no cell bodies in the STN. Scale bar: 100µm.

d, Experimental strategy: AAVrg-Con/Fon-ChR2-EYFP was unilaterally injected into the SNr and SNc of Vglut2+/Vgat+ mice.

e, Representative sagittal section from a Vglut2+/Vgat+ mouse with a unilateral injection of AAVrg-Con/Fon-ChR2-EYFP into the SNr/SNc (blue). Roman numerals and colored boxes indicate regions of interest displayed in **f**.

f, A unilateral injection of AAVrg-Con/Fon-ChR2-EYFP into the SNr/SNc of Vglut2+/Vgat+ mice results in the detection of cell bodies (indicated by white arrows) in the STN (I; brown) and the GPi (II; red). Representative images in **d**, **e** and **f** replicated in N = 3 mice.

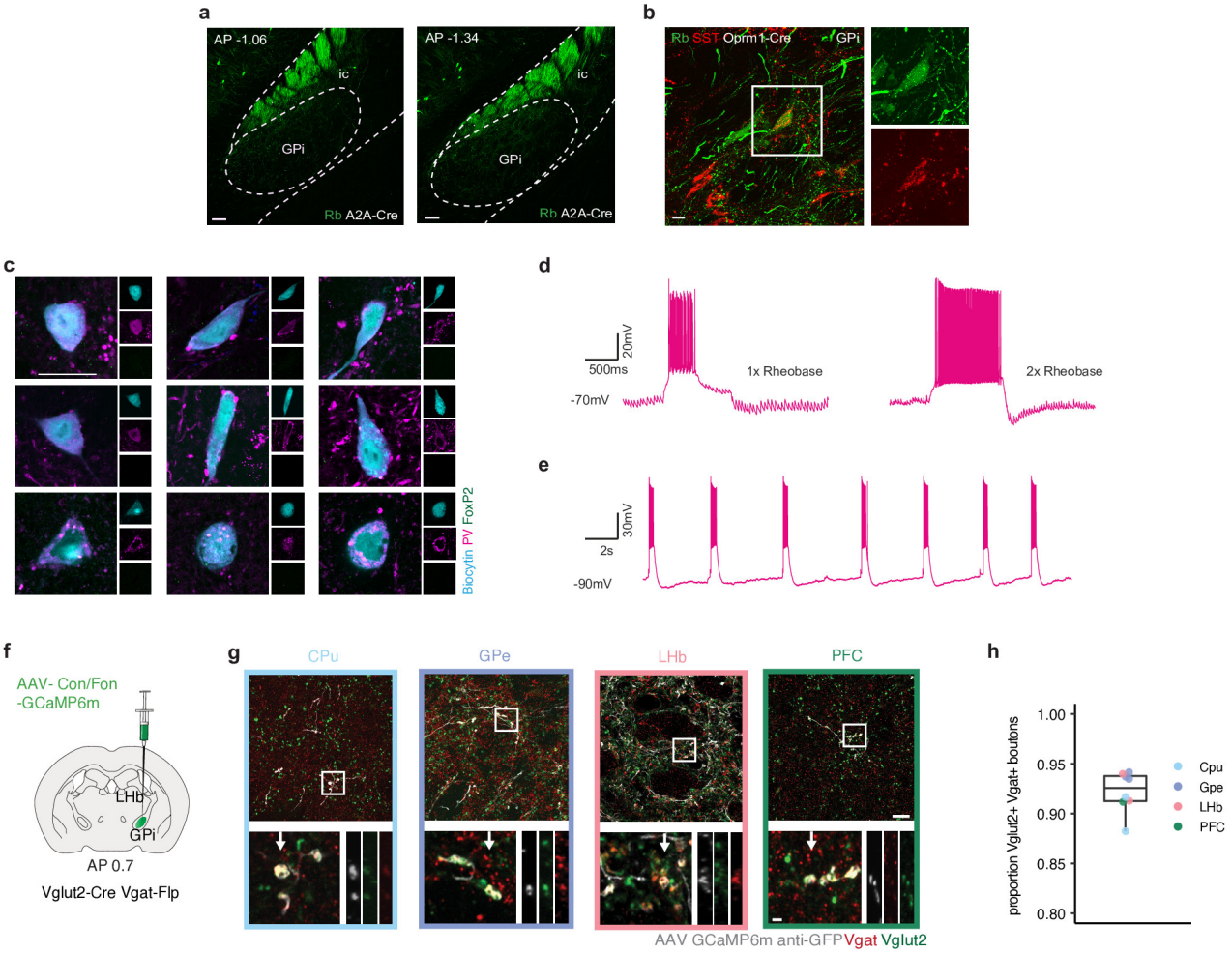
a
AP -1.06
ic
GPi
Rb A2A-Cre
AP -1.34
ic
GPi
Rb A2A-Cre
b
Rb SST Oprm1-Cre
GPi
c
Biocytin PV FoxP2
d
20mV
500ms
1x Rheobase
2x Rheobase
-70mV
e
30mV
2s
-90mV
f
AAV- Con/Fon
-GCaMP6m
LHb
GPi
AP 0.7
Vglut2-Cre Vgat-Flp
g
CPu
GPe
LHb
PFC
AAV GCaMP6m anti-GFP Vgat Vglut2
h
proportion Vglut2+ Vgat+ boutons
1.00
0.95
0.90
0.85
0.80
Cpu
Gpe
LHb
PFC

Figure S5 | Electrophysiological and immunohistological characterization of Vglut2/Vgat GPi neurons' input to the Cpu and the GPe.

a, Rabies tracing experiment of A2A-expressing CPu neurons. No Rabies-expressing monosynaptic input neurons could be detected in the GPi. Related to **Figure 2b**. Scale bar: 100μm.

b, Rabies tracing experiment of Oprm1-expressing CPu neurons. Left: Representative Rabies-expressing neuron (green) in the GPi targeting Ormp1-expressing CPu neurons co-labeled with SST (red) by immunohistological staining. Right: closeup of neuron indicated to the left. Related to **Figure 2c**. Scale bar: 5μm.

c, Immunohistological identification of biocytin-filled (turquoise; n = 9), electrophysiologically-responsive, patched neurons in the GPe show expression of PV (pink) and not FoxP2 (green). Scale bar 20μm.

d, Rheobasic bursting response (left) of a PV-expressing GPe neuron and two-times rheobasic response (right) of another PV-expressing GPe fast spiking neuron. Interestingly there is a high rate of spontaneous activity recorded as postsynaptic potentials in most of the PV-positive GPe neurons innervated by the GPi.

e, Spontaneous burst oscillation of PV-positive GPe neuron upon hyperpolarization to -90 mV.

f, Experimental strategy: AAV-Con/Fon-GCaMP6m was unilaterally injected directly into the GPi of Vglut2-Cre/Vgat-Flp mice.

g, Up: Representative images of immunostaining of Vglut2 (green) and Vgat (red) on GCaMP6m-expressing GPi neuron terminals (white) in the main projection targets regions of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ neurons (from left to right: CPu (light blue), GPe (blue), LHb (pink), and PFC (green)). Scale bar: 10μm. Down: close-ups of area indicated in the top images, displaying the overlay of Vgat, Vglut2 and GCaMP6m-expressing GPi efferents in the Z-plane. Left, down; Identified bouton in the coronal plane, the white arrow indicates the area of cross-sectioning in the z-plane displayed to the right. Right, down; The imaged z-plane of the bouton indicated to the left, for each imaged channel (white: GCaMP6m anti-GFP; green: Vglut2; red: Vgat). Scale bar: 1μm.

h, Proportion of GCaMP6m-expressing GPi boutons co-expressing Vglut2 and Vgat within the four main projection targets of Vglut2+/Vgat+$^{GPi\text{-}LHb}$ neurons (proportion of total boutons identified by the expression of the viral fluorophore: light blue; CPu (n = 65/75), blue; GPe (n= 274/292), pink; LHb (n = 433/465), green; PFC (n = 31/34)). Boxplot: center lines, median; edges, upper and lower quartiles; whiskers, largest and smallest value no more than 1.5×IQR of the edges. Each dot represents the proportion of GPi boutons co-expressing Vglut2 and Vgat for each analyzed brain section, color-coded by the brain area analyzed.

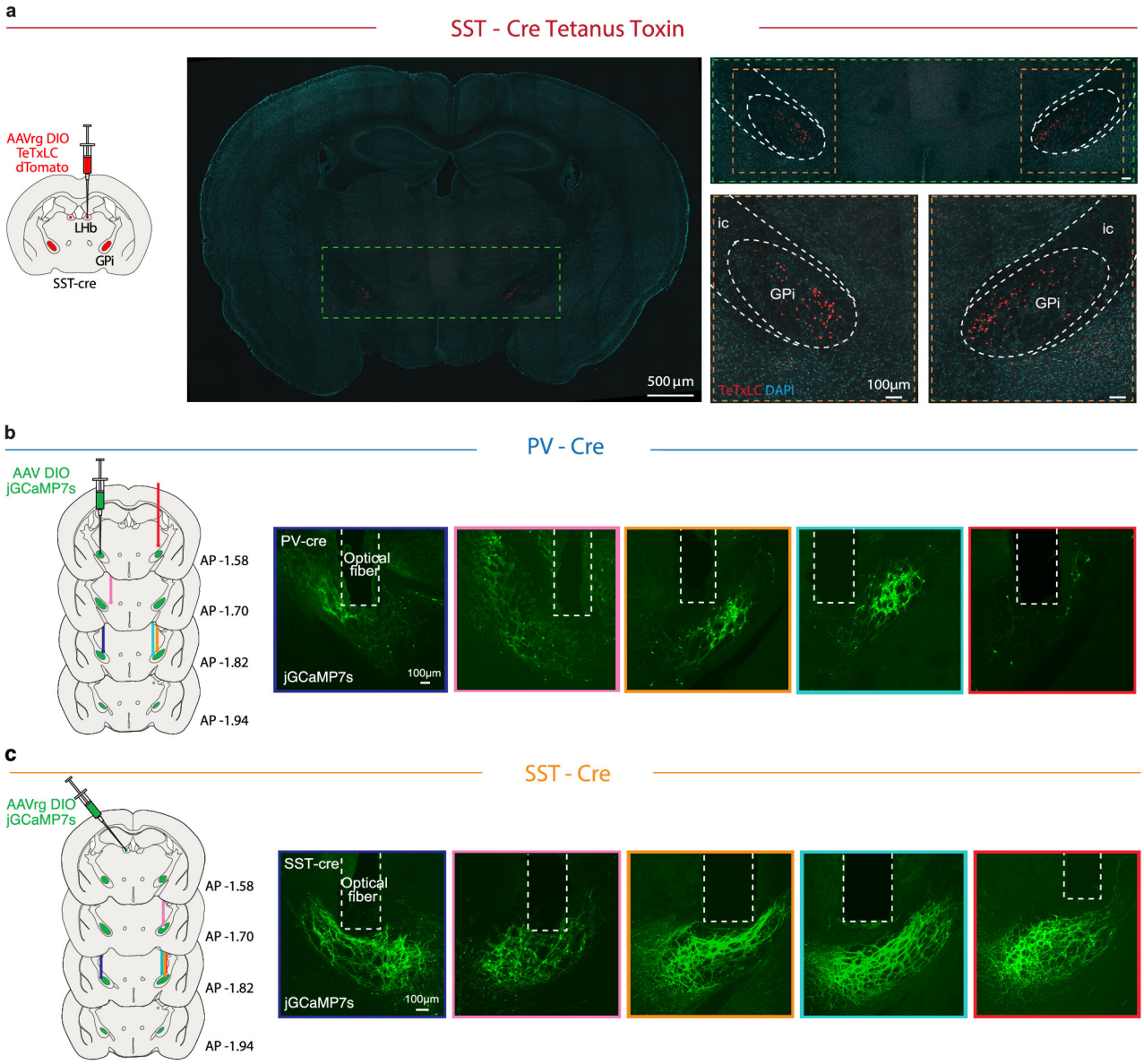

Figure S6 | Fiber placement and virus expression in behavioral cohorts.

a, Left: Experimental strategy for silencing SST+ LHb-projecting GPi neurons. SST-Cre mice were injected with AAVrg-DIO-TeTxLC-dTomato in the LHb. Right: Images from a representative mouse depicting TeTxLC-dTomato-expressing neurons (red) in the GPi. Scale bars; overview: 500µm, close-up: 100µm.

b, Left: Experimental strategy and illustration of viral expression and fiber placement for fiber photometry of PV+ GPi neurons. PV-Cre mice were injected with AAV-DIO-jGCaMP7s and implanted with optical fibers in GPi. Right: Images depicting fiber placement and jGCaMP7s expression (green) in individual mice. Scale bar: 100µm.

c, Left: Experimental strategy and illustration of viral expression and fiber placement for fiber photometry of LHb-projecting SST+ GPi neurons. SST-Cre mice were injected with AAVrg-DIO-jGCaMP7s in LHb and implanted with optical fibers in GPi. Right: Images depicting fiber placement and jGCaMP7s expression (green) in individual mice. Scale bar: 100µm.

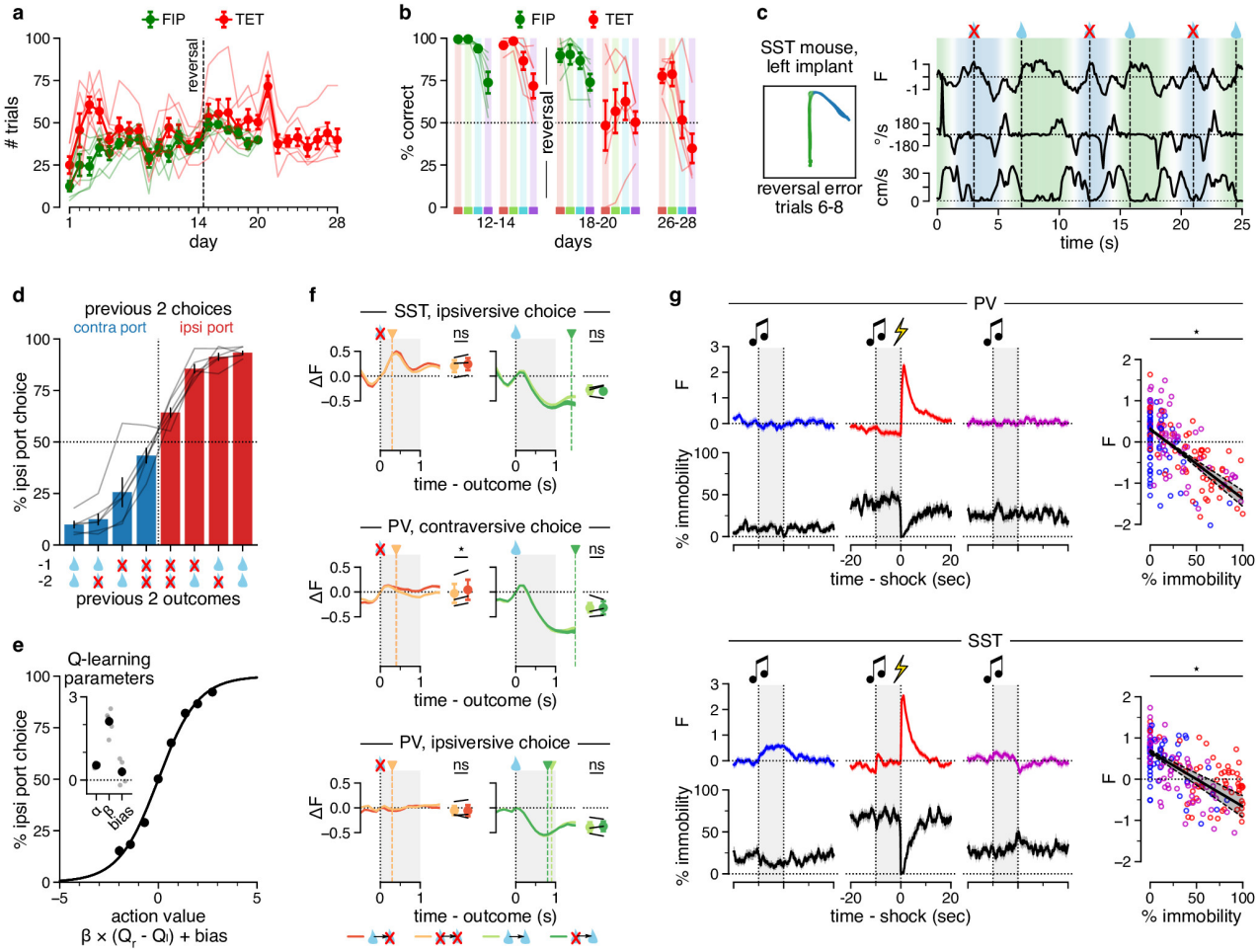
a
FIP
TET
trials
reversal
day
b
% correct
reversal
days
12-14
18-20
26-28
c
SST mouse, left implant
reversal error trials 6-8
F
°/s
cm/s
time (s)
d
previous 2 choices
contra port
ipsi port
% ipsi port choice
previous 2 outcomes
e
Q-learning parameters
α
β
bias
% ipsi port choice
action value
β × (Qr - Ql) + bias
f
SST, ipsiversive choice
PV, contraversive choice
PV, ipsiversive choice
ΔF
ns
time - outcome (s)
g
PV
SST
F
% immobility
time - shock (sec)
% immobility

Figure S7 | Supplementary data for behavioral experiments.

a, Number of trials completed by fiber photometry (FIP, green) and TeTxLC-expressing (TET, red) mice in each session of the maze reversal task. Mean ± SEM. N = 6 mice per group.

b, Average accuracy in each trial quartile bin, by group, pre- and post-reversal. Transparent lines: individual FIP (green) and TET (red) mice, points: group averages. Mean ± SEM. N = 6 mice per group. See also **Figure 3h**.

c, Example data from the first reversal session performed by an SST-Cre mouse (unrewarded error trials 6-8 shown). Left: movement trajectory of the mouse in the maze, color-coded according to the occupied corridor. Right: extracted fluorescence and movement kinematics. Background shading indicates the mouse's position in the maze. Blue: unrewarded choice corridor; green: initiation corridor.

d, Probability of ipsiversive port choice following various outcome histories (two trials back) in the probabilistic reversal task. For simplicity, only trials in which the previous two choices matched the current choice were included in this analysis. Blue bars: contraversive port choices; red bars: ipsiversive port choices. Mean ± SEM. N = 6 mice.

e, Probability of ipsiversive port choice as a function of the outcome history-based relative action value. Action values were fitted per mouse using Q-learning (see Methods). Circles: averages of mice for 8 evenly-spaced action value bins. Line: logistic regression fit, trials pooled over all mice. Inset: Q-learning parameters. Grey circles: individual mice; black circles: averages of mice. Mean ± SEM. N = 6 mice.

f, The remaining SST and PV ΔF responses to omissions (left) and rewards (right), split by outcome on the previous trial (see **Figure 4i** & **j**). Triangles and dashed lines: median port exit time points. Grey shading: time window used to compute means for the paired t-tests shown to the right. Black lines: individual mice. Mean ± SEM. Traces pooled over N = 3 mice per group. *: $p<0.05$, paired t-test. N = 3 mice per group.

g, Auditory fear conditioning of PV (top) and SST (bottom) mice. Traces (left) show the fluorescence response (blue, red, and magenta) and the probability of mice being stationary (black) for trials of the second habituation, the conditioning, and the extinction sessions (left to right, see event pictograms). Mice were considered stationary when median body and head velocity in a 300 ms window averaged lower than 1 cm/s. Grey shading: auditory conditioned stimulus (CS) presentation. Mean ± SEM, traces pooled over animals (N = 3 per group). Regressions (right) assess the relation of fluorescence and immobility during the CS period. Circles: single-trial CS responses, trial type color-coded following the traces. Line and shading: mean ± SEM of the regression fits. *: $p<0.001$, Wald test of the main effect of the slope; linear mixed-effects model. N = 3 mice per group.

See **Supplementary Table 3** for detailed statistical values for **f–g**.

Supplementary Table 1 | Rabies experiments cell counts.

Mouse line	Injection Site	No. Starter Cells	Rabies-expressing neurons GPi	Animal No.
A2A-cre	CPu	1240	0	1
A2A-cre	CPu	1055	0	2
A2A-cre	CPu	1048	0	3
A2A-cre	CPu	388	0	4
OPRm1-cre	CPu	125	8	1
OPRm1-cre	CPu	80	5	2
OPRm1-cre	CPu	133	6	3
OPRm1-cre	CPu	31	8	4
OPRm1-cre	CPu	31	4	5
OPRm1-cre	CPu	48	3	6
DAT-cre	SNc	388	76	1
DAT-cre	SNc	692	92	2
DAT-cre	SNc	232	23	3

Supplementary Table 2 | Mice used in behavioral experiments.

Mouse ID	Experimental group	Recorded hemisphere	Arrow maze	Probabilistic 2-choice	Fear conditioning	Open field
4178	PV-Cre FIP	right	x session #11 missing FIP			
5036	PV-Cre FIP	right	x			x
6145	PV-Cre FIP	right	x session #11 missing FIP, #19 & #20 missing			
6881	PV-Cre FIP	right		x	x	x
6892	PV-Cre FIP	right		x	x	x
6914	PV-Cre FIP	right		x	x	x
0520	SST-Cre FIP	left	x			
6376	SST-Cre FIP	right	x			x
7061	SST-Cre FIP	left	x			x
1944	SST-Cre FIP	right		x	x	x
2133	SST-Cre FIP	right		x	x	x
2144	SST-Cre FIP	right		x	x	x
0419	SST-Cre FIP	left				x
6996	SST-Cre FIP	left				x
380	SST-Cre TET		x			
382	SST-Cre TET		x			
383	SST-Cre TET		x			
385	SST-Cre TET		x			
386	SST-Cre TET		x			
388	SST-Cre TET		x			

Supplementary Table 3 | Statistical tests.

Figure	Test	n	Test statistic	P
3d	1-sample t-test, 1-sided (greater than 50% correct)	6 mice	t(5)	*: P<0.05, Bonferroni-corrected
3e	1-sample t-test of the regression slope, 2-sided	12 mice	pre-reversal: t(10)=1.4016; post-reversal: t(10)=-5.0279	pre-reversal: P=0.1913; post-reversal: P=0.0005
3f	1-sample t-test of the regression slope, 2-sided	12 mice	pre-reversal: t(10)=0.3850; post-reversal: t(10)=-0.0535	pre-reversal: P=0.7083; post-reversal: P=0.9584
3h	2-sample t-test, unpaired, 2-sided	12 mice	trial rate Q1/Q2/Q3/Q4: t(10)=-1.1723/0.0701/-0.0267-1.1740 accuracy Q1/Q2/Q3/Q4: t(10)=1.4800/1.0463/1.5162/0.0520	n.s.: no P<0.05; trial rate Q1/Q2/Q3/Q4: P=0.2683/0.9455/0.9792/0.2676; accuracy Q1/Q2/Q3/Q4: P=0.1697/0.3201/0.1604/0.9596
3h	repeated-measures ANOVA	6 mice	FIP trial rate: F(3,15)=43.2231; FIP accuracy: F(3,15)=12.7165; TET trial rate: F(3,15)=31.5422; TET accuracy: F(3,15)=10.7703	FIP trial rate: P=1.3e-7; FIP accuracy: P=0.0002; TET trial rate: P=1.0e-6; TET accuracy: P=0.0005
4c	2-sample t-test, unpaired, 2-sided	6 mice	t(4)=-3.2541	P=0.0313
4f	2-sample t-test, unpaired, 2-sided	6 mice	t(4)=1.6756	P=0.1691
4j	2-sample t-test, paired, 2-sided	3 mice	non-rewarded: t(2)=-2.5660; rewarded: t(2)=-0.7610	non-rewarded: P=0.1242; rewarded: P=0.5261
5a	2-sample t-test of the log10(angular velocity) X fluorescence correlation coefficients, paired, 2-sided	PV: 4 mice; SST: 7 mice	PV: t(3)=-4.7436; SST: t(6)=1.2165	PV: P=0.0178; SST: P=0.2695

5f	repeated-measures ANOVA of regression intercepts and slopes (PV: 4 missing sessions filled with coefficients from previous session)	3 mice	PV intercept: F(19,38)=1.3034; PV slope: F(19,38)=0.9100; SST intercept: F(19,38)=4.8681; SST slope: F(19,38)=2.0611	PV intercept: P=0.2376; PV slope: P=0.5752; SST intercept: P=1.7e-5; SST slope: P=0.0285
S7f	2-sample t-test, paired, 2-sided	3 mice	SST ipsiversive non-rewarded/rewarded: t(2)=-2.8314/-0.8399; PV contraversive non-rewarded/rewarded: t(2)=-6.7051/0.9973; PV ipsiversive non-rewarded/rewarded: t(2)=0.0970/0.6538	SST ipsiversive non-rewarded/rewarded: P=0.1054/0.4893; PV contraversive non-rewarded/rewarded: P=0.0215/0.9973; PV ipsiversive non-rewarded/rewarded: P=0.9315/0.5804
S7g	Wald test of the main effect of the slope; linear mixed-effects model with random slope and random intercept for each animal	3 mice	PV: z=-8.5650; SST: z=-5.0575	PV: P=1.1e-17; SST: P=4.2e-7