HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2020 August 07.

Published in final edited form as:

Neuron. 2019 August 07; 103(3): 445–458.e10. doi:10.1016/j.neuron.2019.05.026.

Prefrontal cortex regulates sensory filtering through a basal ganglia-to-thalamus pathway

Miho Nakajima[1], **L. Ian Schmitt**[1], **Michael M. Halassa**[1,2,3,*]

[1]McGovern Institute for Brain research and the Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA 02140, USA

[2]The Stanley Center for Psychiatric Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02140, USA

[3]Lead Contact

Summary

To make adaptive decisions, organisms must appropriately filter sensory inputs, augmenting relevant signals and suppressing noise. The prefrontal cortex (PFC) partly implements this process by regulating thalamic activity through modality-specific thalamic reticular nucleus (TRN) subnetworks. However, because the PFC does not directly project to sensory TRN subnetworks, the circuitry underlying this process had been unknown. Here, using anatomical tracing, functional manipulations and optical identification of PFC projection neurons, we find that the PFC regulates sensory thalamic activity through a basal ganglia (BG) pathway. Engagement of this PFC-BG-thalamus pathway enables selection between vision and audition by primarily suppressing the distracting modality. This pathway also enhances sensory discrimination, and is utilized for goal-directed background noise suppression. Overall, our results identify a new pathway for attentional filtering and reveal its multiple roles in sensory processing based on internal goals.

eTOC Blurb

Making sense of a noisy world depends on active filtering of behaviorally relevant sensory information. Nakajima et al. illuminates a pathway that implements this function, demonstrating how control circuits regulate early sensory processing to suppress distractors and improve behaviorally-relevant signals.

*Correspondence: mhalassa@mit.edu.

Author Contributions

M.N, L.I.S and M.M.H. conceived experiments and analyses, interpreted the data and wrote the manuscript. M.N. performed behavioral and tracing experiments. M.N. and L.I.S. performed optical and electrophysiological experiments and analyzed the data. M.M.H. supervised the experiment and directed the analysis. All authors read the final version of the manuscript.

Declaration of Interests

The authors declare no competing interests.

Introduction

An animal's survival depends on its ability to appropriately extract relevant signals from its environment (Heekeren et al., 2008; Romo and de Lafuente, 2013). Because such signals are often embedded in varying levels of sensory noise, the ability to selectively amplify relevant inputs and suppress distractors is of paramount importance (Cohen and Maunsell, 2009; Parr and Friston, 2018). Some of these 'filtering' operations are thought to be automatically implemented at multiple levels of sensory processing (Desimone and Duncan, 1995; Wark et al., 2007). Sensory filtering operations are also engaged by goal-directed (also known as 'top-down' or executive) control (Squire et al., 2013; Verhoef and Maunsell, 2016; Verhoef and Maunsell, 2017). For example, non-human primate experiments have shown that changes in sensory processing are influenced by executive control from prefrontal cortex (Moore and Fallah, 2004; Noudoost et al., 2010). The observation of this type of control has led to the notion that the PFC is a flexible network capable of linking behaviorally-relevant inputs to the regulation of ongoing sensorimotor transformations elsewhere (Miller and Cohen, 2001; Petersen and Posner, 2012).

In previous work, we showed that sensory thalamus is a locus for sensory filtering; when a mouse shifts between selecting and suppressing a broadband visual input with low feature complexity, changes in visual thalamic activity (lateral geniculate nucleus; LGN) predict successful behavior (Wimmer et al., 2015). These changes in LGN activity reflect inhibitory control from the visual subnetwork of the thalamic reticular nucleus (visTRN), which are in turn regulated by the prelimbic region of the PFC (Wimmer et al., 2015). This work in mice complemented previous findings in humans (O'Connor et al., 2002) and non-human primates (McAlonan et al., 2008) in establishing the geniculate as a locus for goal-directed attentional filtering (Phillips et al., 2016). However, a circuit-level understanding of this type of top-down control, starting from how the PFC transforms task-relevant cues to sensory control signals and ending with behaviorally-relevant changes within the sensory thalamus, is lacking. In addition, it is unclear whether such a pathway would only implement broad modality suppression relevant to cross-modal selection, or whether it could have more general roles in attentional filtering. Even if such top-down control had limited feature selectivity and contributed to modality-specific broad noise suppression for example, such a filtering operation would be a crucial prerequisite for subsequent cortical processes underlying selective attention (Hafter et al., 2008)

To address these challenges, we adopted a circuit-level approach in which we first sought to identify the pathways involved in executive control of thalamic sensory filtering. Our experiments revealed a novel pathway by which prefrontal control of visTRN activity was mediated through basal ganglia intermediaries. In addition to broad suppression of distracting sensory input with low feature complexity, we found that this same pathway was engaged to enhance auditory feature selectivity. More specifically, prefrontal circuits transformed learned cues to pathway-specific output signals that ultimately enhanced auditory TRN function, supporting the ability to maintain sparse activity in auditory thalamus despite broadband background noise. Our findings provide novel insights into the circuit mechanisms underlying the PFC's ability to exert goal-directed sensory filtering within and across modalities.

Results

The PFC implements goal-directed visual thalamic suppression through a basal ganglia pathway

We had previously found that neurons within the visual subnetwork of the TRN (visTRN) show bidirectional changes in spike rates predictive of appropriate crossmodal sensory selection task (Fig. S1A) (Rikhye et al., 2018; Schmitt et al., 2017; Wimmer et al., 2015). At the core of the task is sensory selection, where freely behaving mice selected between spatially conflicting visual and auditory targets to receive a reward (Fig. 1A). At the beginning of each trial, mice were presented with either 100 msec of low- or high-pass filtered white noise, which corresponded to two rules for sensory selection: *attend to vision* (ignore audition) and *attend to audition* (ignore vision), respectively.

By design, sensory selection in this task was asymmetric: the visual target was either a right- or left-sided flash that indicated the appropriate reward port location on that trial, while the binaural auditory target was either an up- or down-sweep indicating the right or left reward port locations, respectively (Fig. 1B). This asymmetry meant that visual selection only required the animal to identify simple contrast changes, while auditory selection required feature categorization. To distinguish these two operations, we refer to them as visual contrast ‘detection’ (Busse et al., 2011) and auditory feature ‘discrimination’ (Dragoi et al., 2002; Guo et al., 2017; Pinto et al., 2013).

In this task, visTRN neurons showed a decrease or increase in spike rate prior to stimulus onset depending on whether vision was selected or ignored, respectively (Wimmer et al., 2015). Because these changes were diminished when the PFC was suppressed, and because the PFC has been shown to encode and maintained cue information during target anticipation (Rikhye et al., 2018; Schmitt et al., 2017) we reasoned that the PFC exerts top-down control over thalamic sensory processing to regulate sensory selection (Fig. 1C). Interestingly, we found that enhanced spiking on *attend to audition* trials, which presumably suppresses an expected distracting visual input, was diminished by PFC suppression (Fig. 1D). This finding, along with the fact that visTRN neural suppression has been shown to selectively diminish performance on the same trial type (Wimmer et al., 2015), supports the notion that PFC recruitment of sensory TRN is utilized to broadly suppress a distracting modality.

How does the PFC exert its control on thalamic sensory processing? We found that the PFC does not directly project to visTRN (Fig. 1E, F; or audTRN, Fig. S1B), indicating that intermediate circuits are required. Since it was unclear how many links there were between PFC and visTRN, we took two complementary approaches; an anatomical one in which we identified putative inputs to visTRN and a functional one in which we systematically inactivated PFC output pathways, assessing which particular inactivation gave rise to modality-specific behavioral deficits like those seen when the visTRN itself is inactivated (Wimmer et al., 2015).

These two lines of investigation converged on the basal ganglia as key intermediate circuits coupling PFC to sensory thalamus. First, we performed monosynaptic rabies tracing to

identify inputs to visTRN neurons (Fig. 1G, Fig. S1C, D) modified from previous methods (Sun et al., 2014; Wall et al., 2010). Tracing showed a substantial source of presynaptic inputs from the caudal region of the globus pallidus (GP; Fig. 1H, Fig S1E, F). The GP region identified received GABAergic presynaptic input from the dorsal region of the tail of striatum (Fig. 1I) (Menegas et al., 2017), which is known to receive primary visual cortical inputs (Fig. S1G) (Jiang and Kim, 2018; Oh et al., 2014), as well as direct PFC inputs (Hunnicutt et al., 2016). We refer to this region as visual striatum (visStriatum). Second, we found that optogenetic suppression of PFC terminals in visStriatum disrupted performance only on trials that required animals to suppress the visual target (Fig. 1J, K, Fig. S1H). We found that a similar selective effect was produced by visTRN suppression (Fig. 1L). Such evidence for a PFC output carrying a modality-specific signal was not seen in all other output pathways tested, which either gave symmetric performance deficits across modalities (Fig. 1M, Fig. S1H) or no effects at all (Fig. S1I, J). These effects were unlikely to be artifacts of optogenetic intervention, as EGFP control experiments did not affect task performance (Fig. S1K).

PFC projection neurons to visStriatum selectively encode the *attend to audition* cue

To test the hypothesis that the PFC generates a modality specific signal propagated via the visStriatum and visTRN to suppress visual inputs, we recorded the activity of PFC neurons that project to visStriatum ($PFC_{visStriatum}$). If our hypothesis was correct, these neurons would display modality-specific neural signals during the 2AFC task delay period. As such, we performed optogenetic tagging of $PFC_{visStriatum}$ neurons using the retrograde non-toxic canine adenovirus (CAV2), allowing us to track their activity in the task (Fig. 2A-E; see methods).

Previous studies have shown that PFC neurons encode and maintain the task-relevant cues through sequential population activity in which individual neurons show transient increases in spike rate (Schmitt et al., 2017). Such ‘peaks’ in spike rate tile the delay period of the task, maintaining a temporally-sparse representation of the cue. Here, we found that a subset of $PFC_{visStriatum}$ neurons also showed peaks tiling the delay period. In these neurons, however, peaks were almost exclusively found for *attend to audition* trials (Fig. 2F-H). In agreement with these observations, cue information could be decoded from this population exclusively in the *attend to audition* trials (Fig. 2I, J). Combined, these results suggest that PFC outputs to visStriatum are generally involved in the suppression of thalamic visual transmission when vision is behaviorally-irrelevant.

Engagement of the PFC in the auditory selection component of the 2AFC is multifaceted

The cross-modal sensory selection 2AFC task is, by design, asymmetric. That is, the visual and auditory targets have different degrees of feature complexity. While visual targets are a right- or left-sided light flash whose presence requires simple ‘detection’, the auditory targets are binaurally-presented up- or down-sweeps requiring feature ‘discrimination’. A number of studies have shown that these types of operations are implemented in fundamentally different ways in the brain. More specifically, while detection only requires sensory cortical engagement when stimulus contrast is near threshold (Glickfeld et al., J. Neuro., 2013; Cone et al., eNeuro., 2019), discrimination engages sensory cortex more

generally (Letzkus et al., Nature, 2011; Lee et al., Nature, 2012). As such, these two processes can be differentially, and sometimes oppositely, impacted by a particular manipulation (Guo et al., Neuron, 2017; Olsen et al., Nature, 2012). In line with these studies, we previously had found that primary visual cortex (V1) inactivation causes a shift in visual detection threshold when the animal has to simply detect the location of a visual flash (right vs. left) in an uncued manner (Wimmer et al., 2015). In the cross-modal task, this translated to a relatively small detriment in performance on visual trials. Conversely, suppression of primary auditory cortex (A1) reduced performance down to chance level rather than producing a quantitative detriment in behavior.

The idea that auditory and visual stimuli might be processed differently due to the asymmetric design of the cross-modal sensory selection task led us to wonder if this difference is reflected in the engagement of TRN by the PFC-BG-thalamus pathway. Similar to visTRN, audTRN neurons receive inputs from the GP (Fig. S2A, B), suggesting that any observed differences could not be explained by fundamental differences in circuit architecture. Two-color rabies tracing also revealed that presynaptic inputs to the audTRN and visTRN are largely non-overlapping (3% overlap, 12/354, Fig. 3A, B) suggesting that these circuits are likely to be controlled separately, potentially by different striatal inputs. Because the auditory striatum (audStriatum, Fig. S2C) receives PFC inputs analogous to those of the visStriatum (Hunnicutt et al., 2016) we were able to record and manipulate their corresponding PFC control ($PFC_{audstriatum}$ neurons or their terminals in audStriatum) in the context of the 2AFC task.

Compared to $PFC_{visStriatum}$ neurons, $PFC_{audStriatum}$ neurons showed significant selectivity towards both task-relevant cues (Fig. 3C-F). More precisely, although these neurons were more likely to encode the *attend to vision* than the *attend to audition* cue (Fig. 3F), the relative selectivity between these cues was significantly lower compared to that observed for $PFC_{visStriatum}$ neurons (Fig. S2D). Moreover, in contrast to the selective behavioral effects seen with PFC terminal suppression in visStriatum (Fig. 1K), PFC terminal suppression in audStriatum led to impaired performance on both cueing conditions (Fig. 3G, H, Fig. S2E, F).

At this point, it was becoming clear that the asymmetry in the task design was indeed being translated to asymmetries in neural activity and causal engagement in the task along the visual and auditory PFC-BG-thalamus pathways. However, it remained unclear exactly why the auditory pathway was being engaged in both trial types. The auditory cues used to indicate trial types were unlikely to be the reason of this dual engagement, as previous studies showed that the cueing modality has little or no impact on either performance of this task or the associated patterns of prefrontal engagement (Rikhye et al., 2018).

To better understand how the PFC-audStriatum pathway engaged auditory processing in both trial types, we examined the activity in the auditory thalamus (ventral medial geniculate body; MGBv) during the task. Because our experiments in the visual pathway showed that top-down control is necessary for suppressing distracting visual thalamic inputs, we asked whether a similar effect was taking place at the level of MGBv neurons. Indeed, evoked MGBv neural responses were reduced in *attend to vision* compared to *attend to audition*

trials (Fig. 3I, J), an effect that was quantitatively comparable to that seen for LGN with respect to visual selection (Fig. 3K).

Although changes in MGBv spiking were observed across cueing conditions, they were non-uniform across the population; MGBv neurons that displayed a preference towards a specific auditory target (up- vs. down-sweep) emitted spikes that were more informative of these targets during the *attend to audition* compared to *attend to vision* trials (Fig. 3J, L). This finding, combined with the fact that *attend to audition* trials also engaged the PFC-to-audStriatum pathway suggested that auditory discrimination may require some degree of top-down auditory thalamic suppression. This suppression would enforce a sparse auditory thalamic output, which is consistent with our decoding values (Fig. 3L), and would render target stimulus features to be maximally separable at the level of cortical readout (Chechik et al., 2006; Willmore et al., 2014).

A close examination of our electrophysiological recordings supported the notion that auditory discrimination engages thalamic inhibition, and that this is an explanation for the neural asymmetries observed across the PFC-BG-thalamic pathway in this design of the 2AFC task. Specifically, we found that putative audTRN neurons (Fig S3A-G, see methods) showed an increase in spike rates during the delay period of the task for both *attend to audition* and *attend to vision* trials (Fig. 4A-C). Although this spike rate increase was quantitatively smaller for *attend to audition* trials compared to *attend to vision* trials, audTRN activity was qualitatively different than visTRN activity in this task, where the latter population showed reduced spiking only when the upcoming relevant target was visual (Fig. 4C).

The different behaviors of audTRN and visTRN support the notion that task-design asymmetries may have enforced asymmetric engagement of the identified PFC-BG-thalamus circuits, which was read out behaviorally when this pathway is suppressed (Fig. 1K, Fig. 3H, Fig. S2E), and neurally in the asymmetry seen across the different sensory TRN subnetworks (Fig. 4C). A potential caveat for this idea, however, is the inherent asymmetry between vision and audition being distinct sensory processing pathways. To investigate this issue, we trained mice on a symmetric 2AFC task, in which the auditory component was a simple detection of a left- or right-sided tone cloud (Fig. 4D). Remarkably, this task design resulted in audTRN activity profiles across the two cueing conditions that were qualitatively similar to visTRN activity profiles in the original 2AFC task, in which spike rates were suppressed prior to the preferred stimulus presentation (Fig. 4E, F). These findings confirmed that goal-directed control of sensory thalamic processing is distinct when the sensory operation is a simple detection versus when it is a feature discrimination, and suggested a possible general role for the PFC-BG-thalamus pathway in goal-directed sensory filtering relevant for feature extraction.

The audTRN improves signal to noise ratio of MGBv sensory responses

To directly investigate the role of audTRN in auditory feature discrimination, we performed targeted electrophysical and behavioral studies. First, we recorded sensory responses of MGBv neurons to complex auditory stimuli (dynamic random chords, see methods) while performing randomized, interleaved, optogenetic inactivation of audTRN (Fig. 5A). We

found a marked decrease in feature selectivity in MGBv neurons when audTRN was suppressed, an effect which was worsened by noise (example Fig. 5B). We quantified the impact of this reduced selectivity at the population level through two metrics (population decoding: Fig. 5C; Fano Factor: Fig. S4A), both of which supported the necessity of intact audTRN function to maintain discriminability of neural signals in MGBv. Second, we employed a Go/No Go task that we had previously developed to interrogate auditory discrimination (Schmitt et al., 2017). Briefly, after initiation (500 msec), mice were presented with one of the three different tones (100 msec duration) with varying Signal-to-Noise ratio (SNR). We parametrically varied the SNR of target stimuli by adding background white noise of different intensity levels, which we reasoned would also change the overall feature complexity of the stimuli. Mice were trained to perform a nosepoke in a response port following the presentation of a 20kHz tone, but to withhold following the presentation of either 16kHz or 24kHz tones with both appropriate response types (hits and correct rejections) rewarded at a reward port (Fig. 5D). We found that audTRN suppression (Fig. S4B) during the stimulus presentation resulted in marked behavioral deficits on this task (Fig. 5E). Behavioral impairment varied in relation to the SNR of the target stimuli (Fig. S4C). These results directly support to the idea that enhanced audTRN activity improves MGBv neurons SNR to help maintain sparse stimulus representation necessary for sound discrimination.

Suppressing GP improves stimulus decodability of MGBv through audTRN

Having established a role for audTRN in controlling auditory SNR, we next sought to examine how executive circuits might engage this function. Our data thus far allowed us to construct a circuit diagram linking PFC to audTRN via striatum and GP (Fig. 6A). How might this pathway affect audTRN activity? We found that optogenetic GP suppression, a proxy for how we hypothesized top-down inputs would exert their control, resulted in increased audTRN activity (Fig. 6B-D).

We next examined how the GP-audTRN control pathway impacted responses in MGBv (Fig. 6A) and found that GP suppression diminished MGBv neural sensory responses (Fig. 6E-G). Despite this reduction in sensory-evoked responses, MGBv neural spikes more readily reflected their preferred stimulus, an effect that was particularly evident under high background noise conditions (Fig. 6H). Consistent with this neural SNR enhancement being behaviorally relevant, we found that GP suppression enhanced behavioral performance on the auditory discrimination task described above (Fig. 6I).

The PFC transforms noise-predicting cues to enhanced thalamic SNR through the basal ganglia and audTRN

To isolate the role of the PFC-BG-thalamus pathway in goal-directed background noise suppression relevant for auditory feature discrimination, we modified the auditory Go/No Go task to include a cueing component (Fig. 7A). In each session, 40% of trials did not include added broadband noise (stimulus SNR = 120), while 60% of trials included noise that varied between 0.5 and 10 stimulus SNR. We provided a consistently predictive visual cue (100msec pulse of either U.V. or green light, varied on an animal-by-animal basis) at the beginning of 50% of noise trials. The other 50% of the noise trials were either uncued or

were preceded by the other cue (green or U.V. light, depending on the animal) which was also presented on a random subset of baseline trials and so did not predict whether noise was going to be presented on an upcoming trial. In agreement with the intended design, animals learned to use the 'noise-predictive cue' to improve their auditory discrimination behavior (Fig. 7B).

Consistent with the notion that this type of behavioral design would engage top-down prefrontal control (Fig. 7C), we found that PFC suppression selectively eliminated the advantage of the predictive cue particularly in higher noise conditions (e.g. SNR 1.8 vs. 5.5; Fig. 7D; Fig. S5A). Suppressing PFC terminals in audStriatum also eliminated the cue effect, suggesting that the same PFC-BG pathway is involved (Fig. 7E). The idea that this frontal engagement ultimately enhances audTRN activity was further supported by the observation that audTRN suppression gave broadly similar effects while also impacting uncued trials (Fig. 7F). This latter result is not surprising given that, in contrast to PFC, the audTRN would not be expected to be selectively involved in the cueing component of auditory filtering. Overall, these results support the idea that filtering of background noise in the thalamus within a modality is engaged by the same prefrontal control pathway involved in cross-modal sensory selection. To directly examine the interactions between the PFC and auditory thalamic filtering, we turned to a high-throughput optical method to measure thalamic inhibition in behaving animals (Wells et al., 2016; Wimmer et al., 2015). This method, termed chloride photometry, uses bulk ratiometric measurements of the fluorescence resonance energy transfer (FRET) indicator SuperClomeleon to report changes in intracellular chloride, providing a proxy for GABAergic inhibition (Grimley et al., 2013; Gunaydin et al., 2014; Wells et al., 2016; Wimmer et al., 2015). By using multipoint sampling, we were able to further improve the sensitivity of this measurement (Fig. 7G, see methods), allowing us to more easily attribute changes in MGBv inhibition to the two distinct task phases associated with stimulus anticipation and delivery (Fig. 7H). Presenting a predictive cue augmented evoked MGBv inhibition solely due to its effect on stimulus anticipation, not delivery (Fig. 7I). Importantly, PFC suppression eliminated this anticipatory change in MGBv inhibition. This loss of FRET signal with PFC suppression is unlikely to be due to the laser itself in frontal regions, because we previously observed that laser in task-irrelevant frontal regions didn't affect the FRET signals in the sensory thalamus (Wimmer et al., 2015). This confirmed that the PFC-BG-thalamus pathway is ultimately controlling auditory discrimination through changes in thalamic inhibition.

Our data to this point allowed us to infer that the PFC is engaged during goal-directed auditory noise suppression, but the neural coding principles underlying this engagement remained unclear. For example, does the PFC encode and maintain the noise-predicting cues prior to target stimulus presentation? To address this question, we examined prefrontal physiology in the cued auditory discrimination task using multi-electrode recordings (Fig. 7J). We found that PFC neurons indeed encoded the noise-predictive cues, and that the population exhibited a temporally-sparse sequential response similar to what we had observed in the cross-modal 2AFC task following task-relevant cue presentation (Fig. 7K, L). PFC neural population decoding revealed encoding and maintenance of the noise-predicting but not the unpredictive cue (Fig. 7M, Fig. S5B). Critically, the representation of the predictive cue was much easier to decode from optogenetically-tagged $PFC_{audStriatum}$

neurons than from the PFC population more broadly (Fig. 7N, Fig. S5C), consistent with the notion that these neurons may be the main cognitive control output population for this task.

Discussion

To cope with the challenge introduced by distracting inputs and sensory noise, animals use a variety of filtering mechanisms (Wark et al., 2007; Warrant, 2016) capable of modulating neuronal response to sensory inputs based on the characteristics of those inputs (Angelucci and Bressloff, 2006; Carandini and Heeger, 2011). However, the need to engage sensory filtering often varies depending the degree of sensory noise and particular behavioral goals (Gilbert and Li, 2013). To accommodate these changing requirements, flexible networks in the frontal cortex can selectively modulate filtering of behaviorally-relevant inputs through attentional mechanisms (Miller and Cohen, 2001).

In this study, we identify and characterize a clear example of this kind of coordinated top-down control: PFC dependent regulation of thalamic sensory filtering. Our results demonstrate that the PFC engages inhibitory circuits in the thalamus via a novel PFC-BG-thalamus pathway to implement control of sensory processing. This newly identified circuit can be utilized either to broadly suppress a distracting modality or improve signal-to-noise ratio (SNR) for noisy input by maintaining a sparse response profile within a single modality.

A “triangulation” approach to functional circuit mapping

Our identification of a new PF-BG-thalamus sensory filtering pathway relied on the use of multiple, independent lines of investigation. This “triangulation” approach provides a means to understand the role of a given set of circuits in behavior while also helping to rule out alternative interpretations. In the current study, we used two distinct tasks which independently showed that PFC-BG-thalamus pathway engagement was general to sensory filtering regardless of whether the filtering operation itself was embedded within a more complex hierarchical design or whether its top-down control was initiated by visual or auditory cues. In addition, the task-general findings that silencing of PFC terminals in a sensory striatal territory selectively disrupted behavior on trials requiring suppression of that same sensory modality support the idea of modality-specific sensory filtering and rules out optogenetic artifacts.

Although the efficacy of presynaptic terminal suppression using eArch3.0 has been questioned, with some studies pointing to paradoxical excitation following prolonged activation of this opsin (Wiegert et al., 2017), our relatively brief stimulation (500msec) is unlikely to be subject to these types of effects (El-Gaby et al., 2016). Although the cannot entirely rule out all potential artifacts, our triangulation approach makes it unlikely that such issues would lead to incorrect interpretations for multiple reasons. First, the clearly distinguishable effects produced by suppressing terminals in audStriatum and visStriatum strongly supports the notion that the two outputs are functionally distinct. Second, these pathway-specific optogenetic effect are supported by the independent observations that $PFC_{visStriatum}$ and $PFC_{audStriatum}$ neurons encoded task-relevant cues with different degrees

of selectivity. Lastly, electrophysiological and optical recording from the thalamus demonstrate that sensory TRN activity mirrored pathway-specific cortical engagements.

PFC flexibly transforms learned cues to fixed outputs in order to control sensory processing

PFC neurons are known to exhibit a high degree of task-selectivity rather than fixed sensory responses, a feature thought to underlie the ability of this brain region to flexibly control sensorimotor transformations in other brain regions (Donahue and Lee, 2015; Miller and Cohen, 2001; Squire et al., 2013; Szczepanski and Knight, 2014). In this study, we showed that although internal PFC circuitry can be flexible, PFC outputs are anatomically fixed. As a key example of this fixed output circuitry, we investigated PFC projections to different territories of the basal ganglia involved in the regulation of sensory processing in different modalities. We find that outputs to these different parts of the basal ganglia must be linked to behaviorally relevant inputs to control sensory processing based on learned rules. This result is broadly relevant to understanding the diverse functions of the PFC as it places a fundamental constraint on how this circuit can remodel its internal network to coordinate activity in outside circuits that it regulates. Moreover, the approaches used in this study can be applied to identify functional engagement of PFC output neurons in different behaviors potentially facilitating the investigation of the role of the many behaviors in which PFC is implicated (Miller and Cohen, 2001).

Star*Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and be fulfilled by the Lead Contact, Michael Halassa (mhalassa@mit.edu)

Experimental Model and Subject Details

Vgat-Cre mice (016962) and C57BL/6J mice (000664) were obtained from Jackson Laboratory. All KO and Vgat-Cre mice used in this study were backcrossed to C57BL/6J mice for at least 6 generations to obtain a homogeneous genetic background. All mice tested were between 2–12 months of age and housed on a 12-h light/dark cycle. Male mice were used for behavioral testing to reduce potential confounds from placing mice both genders sequentially in the same behavioral testing environment, while mice of both sexes were used for all other experiments. Throughout training and testing, all procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee at the New York University Langone Medical Center as well as the Committee on Animal Care at the Massachusetts Institute of Technology. All procedures are also in accordance with guidelines issued by the US National Institutes of Health.

Method Details

Behavioral Training and Testing

Testing/Training Setups for Behavioral Task: Behavioral training and testing for all tasks (cross-modal sensory selection task, basic auditory discrimination task, and cued noisy

auditory discrimination task), took place in grid-floor mounted, custom-built enclosures made of acrylic plastic (maximum dimensions in cm: length: 15.2; width: 12.7; height: 24). All enclosures contained custom-designed operant ports, each of which was equipped with an IR LED/IR phaototransistor pair (Digikey, Thief River Falls, MN) for nose-poke detection. For all tasks, trial initiation was achieved through an initiation port mounted on the grid floor ~6 cm away from the 'response ports' and 'reward ports' located at the front of the chamber. A pair of electrostatic speakers (Tucker Davis Technologies) producing the auditory stimuli were placed outside of the training apparatus and sound stimuli were conveyed via cylindrical tubes to apertures located at either side of the initiation port, allowing consistent delivery of stereotypical stimuli across trials. All stimuli and auditory cues across tasks were generated by a TDT Rx8 sound system (Tucker-Davis Technologies, Alachua, FL). Sound stimuli and auditory cues were recorded and assessed for intensity using a prepolarized icp array microphone (PCB Piezotronics, Depew, NY) after which frequency production was equalized using software-based calibration via SigCalRP (Tucker-Davis Technologies, Alachua FL). Visual stimuli for cross-modal sensory selection task were produced by two dimmable, white-light-emitting diodes (Mouser, El Cajon, CA) mounted on each side of the initiation port and controlled by an Arduino Mega microcontroller (Ivrea, Ital). Noise cues for noisy auditory discrimination task (see below) were produced by UV (320–380nm) or Green (495510nm) light emitting diodes (Mouser, El Cajon, CA) mounted on the top of the enclosure and controlled by Arduino Mega microcontroller (Ivrea, Italy). For cross-modal sensory selection tasks, two response ports were mounted at the angled front wall 7.5 or 5 cm apart, respectively. Response ports were separated by 1 cm divider walls and each was capable of delivering a milk reward (10μl evaporated milk delivered via a single-syringe pump (New Era Pump Systems, Farmingdale, NY) when a correct response was made. In the case of the auditory GO/NO GO task environment used for the basic and cued noisy auditory discrimination tasks, response and reward ports were stacked with the response port on top. The stacked ports were located centrally behind the initiation poke. Access to all response and reward ports was restricted by vertical sliding gates which were moved via a custom 3D printed rack and pinion gear system (MakerBot replicator, Brooklyn, NY) powered by a servo motor (Tower Hobbies, Champaign, IL). The TDT Rx8 sound production system (Tucker Davis Technologies, Alachua, FL) was controlled through MATLAB (MathWorks, Natick, MA), interfacing with a custom written software running on an Arduino Mega (Ivrea, Italy) for trial logic control. For all tasks, mice were food restricted to 85–90% of their *ad libitum* body weight before training.

Cross-modal Sensory Selection Task Training and testing: A total of 18 control and 6 Vgat Cre mice were trained on this task. No differences were observed in learning between these groups. Training was performed as previously described (Schmitt et al., 2017; Wimmer et al., 2015) and required ~ 2 months of daily training for each mouse. Particular steps were taken throughout the training and testing periods to ensure that mice used the rules for sensory selection.

Training was carried out in multiple stages. First, 10 μl of evaporated milk (reward) was delivered randomly to each reward port for shaping and reward habituation. Second, mice

were trained to initiate individual trials and make association of side of reward poke with target stimuli (no conflict trials). Initially, mice had to briefly (50msec) break the infrared beam in the initiation port to trigger target stimulus presentation and render reward ports accessible. Mice were trained to hold their snouts for up to 700msec. Broadband white noise indicated trial availability, which prompted a mouse to initiate a trial. Upon successful initiation, the white noise was replaced by either brown (10-kHz low-pass-filtered white noise) or blue noise (11-kHz high-pass-filtered white noise) for 0.1sec to indicate whether to *attend to vision* or *attend to audition*. This was followed by a delay period (0.4–0.6 sec) prior to target stimuli presentation. Target stimuli were presented in blocks of six trials consisting of single-modality stimulus presentation (no conflict). In *attend to vision* trials (brown noise), the location of the rewarded port was signaled by a white LED (visual target) mounted underneath it in order to establish an association between the location of the visual target and the location of the reward port. In *attend audition* trials (blue noise), mice learned the association between the auditory target, up-sweep (10–15kHz) or a down-sweep (16–12kHz) with right or left ports respectively. For symmetric task, a 0.1s tone cloud consisting of a random set of overlapping pure tones (12.5 ms tone length with an overlap of between 2–8 tones at any given time) spanning a range of 4–46 kHz was played from the speaker on one side only. The same tone clode was used for right and left sides, and mice were trained to associate the location of sound and the location of the reward port. Rewards were available for 15sec following correct response, followed by a 5sec inter-trial interval (ITI). An incorrect response immediately renders the response ports inaccessible and was punished with a time-out, which consisted of a 30sec ITI. Third, conflict trials were introduced, in which auditory and visual targets were co-presented indicating reward at opposing locations and mice need to use trial cues (brown or blue noise) to match the target modality. Four different trial types were presented in repeating blocks: (1) three auditory-only trials; (2) three visual-only trials; (3) six conflict trials with auditory target; and (4) six conflict trials with visual target. Once mice performed successfully on conflict trials, single-modality trials were removed and block length was reduced to three trials. Fourth, during the final stage of training, all block structure was removed and trial type was randomized. For each mouse, training was continued until the animal's performance level reached at least 65 % for both trial cues, after which they were injected with viral vectors and implanted with optic fibers or microdrive (see relevant sections below). Following recovery, each animal was re-trained to the original performance criteria. For electrophysiological recordings and experiments with optical manipulation, testing conditions were equivalent to the final stage of training. To make sure mice did not have any biases towards one modality and that they were able to use trial cues to choose modalities, only sessions with balanced performance (both auditory and visual trial types above 65%) across target modalities in baseline trials (those without laser manipulations) were used for the further analysis.

Noisy Auditory Discrimination Task Training and Testing: A total of 18 control and 5 Vgat Cre mice were trained on this task. No differences were observed in learning between these groups. For both the basic and cued auditory discrimination tasks, animals were first trained to initiate, then to perform the basic discrimination task (~3 weeks daily training). In training and testing, mice initiated each trial by holding their snout in an initiation port for at least 50msec to initiate a delay period. Following this 500msec delay, a pure tone stimulus

was played for 100ms from speakers on both sides of the initiation port at an intensity of 60 dB. Animals were trained to hold their head in the initiation port throughout the delay and stimulus delivery and trials were interrupted if they failed to do so. The pure tone stimuli were pseudo-randomly varied on a trial by trial basis, with trials divided between the “GO” stimulus (20 kHz, ~40% of trials) and two “NO GO” stimuli (16 and 24 kHz, ~ 30% of trials per frequency). The order of “NO GO” stimuli also followed a pre-determined pseudorandomized sequence. After stimulus presentation, the response port was made accessible for a 2.5sec trial period. In “GO” trials, the mouse was required to poke in this response port within the trial period (a “Hit”) in which case a reward port directly underneath the response port became accessible, and reward was delivered. For a “Miss” in which the mouse failed to poke within the trial period, the reward port was not made accessible. For a “correct rejection”, withholding for the full 2.5sec when the “NO GO” stimulus was played, the reward port was made accessible. For a “False Alarm” response on a “NO GO” trial, the reward port was not made accessible. For both types of incorrect response (False Alarm and Miss) animals were punished with delay in the availability of the next trial by an additional 15sec time-out period.

To parametrically vary stimulus signal to noise ratios (SNR) white noise was added to pure tone stimuli at fixed stimulus SNR ratios as labelled in each relevant figure. In the basic discrimination task, 60% of total trials were randomly masked by noise (average session length = 212 trials). Following initial introduction of noise, performance on noise trials dropped precipitously even for low noise levels. For the lowest noise levels, however, performance recovered close to baseline levels and partial recovery was observed for high noise conditions as well. Because of this, mice were trained for an additional 1–2 weeks following noise introduction with a mixture of noise intensities prior to testing. For the cued noisy auditory discrimination task a Green (495–510 nm) or UV (320–380 nm) light was activated for 100msec following the start of initiation. Through multiple sessions prior to testing (~ 2 months daily training following initial task acquisition) mice were trained to expect a noise trial following one of the two light colors on 100 % of trials (predictive cue) and a 50% chance of noise following the other color (unpredictive cue). Mice were divided into two equal groups with the predictive cue assigned to UV in one group and green in the other. Groups were of equal size for each experiment with a minimum of 2 mice in each. After the light cue, or a 100 msec cue free period for uncued trials, animals were required to hold through an additional 400 msec delay period prior to the sound stimulus playing. During testing in the cued noisy auditory discrimination task, trials were divided into subsets for different noise/cue combinations. In each session, ~40% of total trials were pure tone without noise (SNR =120) and were either uncued or cued with the unpredictive cue (uncued 25% total, unpredictive cue 15% total). The remaining 60% of total trials were noise trials (SNR 0.5–10) divided among all three cueing conditions (predictive cue 30% total, unpredictive cue 15% total, uncued 15% total). Each session included only one SNR level for the noise trials. After this second training stage, mice were injected with viral vectors and implanted with optic fibers or microdrive (see relevant sections below). Following recovery, each animal was re-trained to a performance level of > 70% in pure tone trials (SNR = 120). Only sessions with the performance above 70% in pure tone trials (SNR = 120) were used for analysis.

Behavioral Analysis

Cross-Modal Sensory Selection Task: Analysis of behavioral experiments was performed as described previously(Rikhye et al., 2018; Schmitt et al., 2017; Wimmer et al., 2015). Briefly, for all experiments with optogenetic manipulations, only sessions where baseline performance was at ≥ 65% correct for both trial types were included in the analysis. To account for variation associated with baseline performance across sessions and across mice used as well as multiple comparisons using the same data, statistical analysis was performed for all manipulations in tandem. Details of this approach are given in the Quantification and Statistical Analysis section, below.

Basic and Cued Auditory Discrimination Tasks: Performance on the discrimination task was initially assessed using the d' statistic (d' = Z(hit rate) – Z(false alarm rate)). For all experiments including psychometric function estimation across stimulus SNR levels, optogenetic, and pharmacological manipulations sessions were only included if baseline performance was ≥ 65% correct (SNR = 120). Noise-masking behavior with multiple SNR levels (Fig. 5, 7) was averaged across sessions (d' was calculated for each session and noise-level, then averaged) and fit with a logistic function. For optogenetic behavioral (audTRN suppression, Fig. 5; GP suppression, Fig. 6) and predictive cue behaviors (Fig. 7) these trial types were initially grouped within sessions and d' values were calculated on a session by session basis. Only sessions in which at least 25 trials of a given type occurred were included. For comparisons at multiple SNR levels, performance on each trial type was pooled within sessions for each SNR level (one SNR level was included per session) and d' was calculated. To estimate parameters of psychometric functions across noise levels, d' averages for all SNR levels were fit with the logistic function:

$$F(x;\alpha,\beta,\lambda) = \frac{(\lambda)}{1 + exp(-\beta(x-\alpha))}$$

where x corresponds to the inverse of the ratio between the intensity of added noise and the maximum sound intensity of the stimulus (i.e. stimulus SNR) in log-scale, α corresponds to the detection threshold and λ corresponds to the maximum performance associated with behavioral "lapse rate" (Wimmer et al., 2015). Fitting was made using the Palamedes psychophysical toolbox (http://www.palamedestoolbox.org/) via maximum likelihood estimation. Confidence interval estimates were then made using a bootstrapping procedure in which subsets of sessions were selected at random across mice (60% selection per subset per SNR level) with parameters estimated by fitting the resulting data for each subset.

To estimate the performance benefit by a noise predictive (or unpredictive) cue, the performance (d') for individual SNR levels in uncued trials within each session was subtracted from the corresponding performance on cued trials. The resulting function, being a change between logistic cumulative distribution functions, was then fit using the logistic probability density function (Treisman and Faulkner, 1985):

$$F(x;\alpha,\beta,\lambda) = \frac{exp^{-\beta(x-\alpha)}}{(1 + exp(-\beta(x-\alpha)))^{2}/\beta}$$

Where the parameters are the same as in the fitting of the original function above.

Optogenetics in Behavior—For experiments with optical stimulation (Fig. 1, 3, 5, 6, 7, Fig. S1, 4, 5), testing conditions were equivalent to the final stage of training. Laser trains of yellow (560 nm for eNpHR3.0 or eArch3.0 activation) light consisting of 50Hz, 18-ms pulses (90% duty cycle) at an intensity of 8–12 mW (measured at tip of the optic fibers) were delivered via TTL triggered laser system (Omicron-Laserage, Dudenhofen, Germany) on a random subset of trials. Laser light intensity was calibrated prior to each testing session using a laser power meter (Thorlabs, NJ, USA). Because behavior and recording systems were automated and stimulus sequence and optogenetic manipulations varied on a trial by trial basis, researchers were not blinded to trial type. Laser was turned on either during delay and the tone presentation (Fig. 5, 6, 7, Fig. S4, 5) or during the delay period (Fig. 1,3, Fig. S1, 2) of pseudo-randomly selected subsets of trials.

Optical Chloride Measurements in Behavior—For combined TRN optical recordings with/without optogenetic PFC disruption (Fig. 7G-I), laser trains of yellow light were delivered during the initiation period on a random subset of trials as described above. FRET-based measurement of sound-evoked $[Cl^-]_I$ responses was performed as previously described (Wells et al., 2016; Wimmer et al., 2015) with some methodological improvements to enhance signal level, as described below. For these recordings, excitation of SuperClomeleon CFP and YFP along with their emitted light were carried through chronically implanted optical fibers using a specialized, custom constructed triple fiber (total inner diameter 660 μm, Doric lenses, see Fig. 4K) which was connected with a 400μm, 0.48NA optic patch cord to the recording system. This triple fiber consisted of three angled mirror fibers (60 degrees, NA 0.66) which surrounded the sampled structure. These fibers were oriented towards the sampled structure allowing them to both provided CFP excitation (430 nm light) and collect emissions within the optimally excited zone. The three fibers were collimated through a custom lens system to connect to a common patch cord through which excitation light was also delivered. Recordings were made using the Assisted Rotating Fluorescence Mini Cube (ARFMC) for FRET system (Doric lenses, Quebec, Canada). Although minimal artifacts were observed, to reduce laser-light induced artifacts optogenetic manipulations were delivered via angled optical fibers oriented away from the recording fibers, as described below.

Virus Set and Injection Coordinates—All AAVs were produced by either the vector core at the University of North Carolina Chapel Hill or Massachusetts Institute of Technology viral core with titers above 10^{12} VG/ml. All FuGB2LV were produced in our laboratory with titers above 10^8 VG/ml. The vector for TVA was modified from the original lentivector pFCGW. FuGB2LV was produced as described previously (Halassa et al., 2014). The expression plasmid and two helper plasmids, delta8.9 and FuGB2 (Kato et al., 2011), were transfected into human embryonic kidney 293FT cells with Polyethylenimine “Max” (PEI; Polyscience, Inc; 24765). Viral particles were collected from the cell culture medium, pelleted by ultracentrifugation at 50,000 × g(m/s^2) for 2 hours. Pseudotyped, RG-deleted rabies virus, RVΔG was purchased through Salk Institute with titers above 10^8 VG/ml. Canine adenovirus 2 encoding Cre (CAV2-Cre) was produced in Institute de Genetique

Moleculaire de Montpellier. Coordinates for each injection were as follows (in mm, A/P, M/L from Bregma, D/V from brain surface): MGBv: A/P: −3.2, M/L: ±2.0, D/V: −3.0; PFC: A/P: 2.4, M/L: ±0.6, D/V: −1.4; LGN: A/P: −2.1, M/L: ±2.0, D/V: −2.5; visTRN: A/P: −1.6, M/L: ±2.0, D/V: −2.4; audTRN: A/P: −1.8, M/L: ±2.3, D/V: −3.0; visStriatum: A/P: −1.46, M/L: ±3.0, D/V: −1.9; audStriatum: A/P: −1.8, M/L: ±2.8, DV: −3.0; GP: A/P: −1.4, M/L: ±2.6, D/V: −3.0. Mice were anesthetized using 1% isoflurane and mounted on a stereotactic frame for virus injections. For behavioral experiments, mice were allowed to recover for 2–4 weeks following virus injection to allow expression prior to testing.

Tracing: For monosynaptic retrograde tracing experiments, we used a two-step injection protocol. First, 2 helper viruses were injected: 0.4μl FuGB2LV-EF1α-DIO-EGFP2aTVA was injected in sensory thalamus (either LGN or MGBv), and 0.1μl of AAV2-EF1α-DIO-HistonGFP2aB19G was injected in sensoryTRN. Second, after 1–2 months, 0.2μl of RVΔ*G*, (EnvA-SAD ΔG-mCherry, or EYFP) was injected to the sensory thalamus (either LGN or MGBv).

Optogenetic Experiments: For TRN manipulation, visTRN and audTRN neurons were labeled through injections (0.4–0.6μl) of FuGB2LV-EF1α-DIO-ChR2-EYFP (for activation) or FuGB2LV-EF1α-DIOeNpHR3.0-eYFP (for inactivation) into LGN or MGBv of Vgat-Cre mice. To tag specific PFC projection neurons, CAV2-Cre virus was injected into visStriatum or audStriatum and an AAV virus harboring Cre-dependent ChR2-EYFP was injected into PFC, resulting in expression of ChR2 in each population. For optogenetic PFC soma or terminal suppression during behavior, 0.4μl of AAV2-hsyn-eArch3.0-EYFP was injected into PFC. To examine the projection of neurons in the tail of striatum to GP, 0.1μl of 1:1 mixture of AAV8 EF1α-DIO-synaptophysin-mCherry and AAV1-hsyn-ChR2-EYFP was injected in visStriatum.

Optic fiber Implantation—As with viral injections, mice were anesthetized using 1% isoflurane and mounted on a stereotactic frame. For behavioral optogenetic experiments, up to four pairs of 200 μm optic fibers (Doric Lenses, Quebec, Canada) were stereotactically inserted at the following coordinates (in mm A/P, M/L from Bregma, D/V from brain surface): PFC: A/P: 2.6, M/L: ±0.6, D/V: −1.0; visStriatum: A/P: −1.7, M/L: ±2.8 D/V: −1.3, visTRN: A/P: −1.6, M/L ±2.2, D/V: −-2.2; audTRN: A/P: −1.8, M/L: ±2.3, D/V: −2.5; MD: A/P: −1.46, M/L: ±0.5 D/V: −2.5, Amygdala: A/P: −2.8, M/L: ±3.35 D/V: −3.0; HDB /P: 0.2, M/L: ±1.1 D/V: −4.7, audStriatum: A/P: −1.8, M/L: ±2.8, DV: −2.5, GP: A/P: −1.4, M/L: ±2.6, D/V: −2.5. For fiber photometry experiments, custom constructed triple fibers were implanted bilaterally in the MGBv (A/P: −3.2, M/L: ±1.8). To optimize signal quality in these experiments, fibers were implanted into pre-injected mice two weeks after virus injection. During implantation, fluorescence measurements were made continuously while fibers were slowly advanced towards the target. During this targeting, broadband sounds were delivered to the contralateral ear at regular intervals. Once sound-related events were observed in the optical signal, the depth was recorded and fibers were fixed in place using dental cement.

In combined optogenetic and fiber-photometry experiments, 45° angled optical fibers were implanted posterior to the PFC (A/P: 2.6, M/L: ±0.6, D/V: −1.0) and oriented towards the

anterior to minimize light contamination in the optical recordings. Up to 3 stainless-steel screws were used to anchor the implant to the skull and everything was bonded together with dental cement. Mice were allowed to recover with *ad libitum* access to food and water for one week after which they were brought back to food regulation and behavioral training resumed.

Microdrive Array Construction and Implantation—Custom drive housings were designed using 3D CAD software (SolidWorks, Concord, MA) and printed in Accura 55 plastic (American Precision Prototyping, Tulsa, OK) as described previously (Liang et al., 2017). Prior to implantation, each drive was loaded with 12–24 independently movable microdrives carrying 12.5μm Stablohm 650 (California Fine Wire Company, Grove Beach, CA) tetrodes. Electrodes were pinned to custom-designed 96- or 128-channel electrode interface boards (EIB, Sunstone Circuits, Mulino, OR) along with a common reference wire (A-M systems, Carlsborg, WA). For combined optogenetic tagging and electrophysiological recordings of audTRN or visTRN, mirror-tipped optical fibers delivering the light beam at right angles (MFC_200/245–0.37_34mm_MF1.25_MA45, Doric Lenses Inc., Quebec, Canada) were embedded in our implants anterior to the electrode arrays and oriented posteriorly towards the audTRN or visTRN. For high-density bilateral recordings of MGBv, we constructed drives with static, non-movable electrodes (implantation targeting is described below for these drives).

For combined optogenetic manipulations and electrophysiological recordings of the PFC, optic fibers delivering the light beam lateral (MFC_200/245–0.37_34mm_MF1.25_MA45, Doric Lenses Inc., Quebec, Canada) were embedded adjacent to the electrodes.

During drive implantation, mice were deeply anesthetized with 1% isofluorane and mounted on a stereotaxic frame. Burr holes were drilled for optical fibers when necessary. A craniotomy was drilled centered at A/P 2 mm, M/L 0.6 mm for PFC recordings (~1 × 2.5 mm), at A/P −1.8 mm, M/L 2.0 mm for audTRN recordings (~2 × 2 mm), at A/P −2mm, M/L 2.5mm for visTRN recordings(~3 × 2 mm), or at A/P −3.2 mm, M/L 2.0 mm for MGBv recordings (~2 × 2 mm). The dura was carefully removed and the drive implant was lowered into the craniotomy using a stereotaxic arm until electrode tips touched the cortical surface. Surgilube (Savage Laboratories, Melville, NY) was applied around electrodes to guard against immobilization by the dental cement used to affix the drive (see below). Stainless steel screws were implanted into the skull to provide electrical ground and mechanical stability for drives and the whole construct was bonded to the skull using C&B-Metabond luting cement (Parkell, Edgewood NY). For head-fixation experiments, a custom-designed 3D-printed hexagonal plastic crown (MakerBot Replicator, Brooklyn NY) was implanted encircling the drive at its base.

In the subset of surgeries that used static implants for MGBv recordings, online targeting was necessary to ensure accurate electrode placement. In these cases, the drive was connected to our data-acquisition system for electrophysiological recordings (see below, Electrophysiological Recordings) when being lowered into the brain. Once the drive was lowered to within 500 μm of the target depth, we presented bilateral auditory stimuli (dynamic random chords presented with EC1 electrostatic speakers with an ED1 speaker

driver, Tucker Davis Technologies, Alachua FL). The drive was then advanced in steps of 100 μm until we observed auditory-responsive units. The drive was then advanced an additional 250 μm to target the ventral MGB. If auditory responsive units were still observed, the drive was and then bonded to the skull as described above. Otherwise, the drive was raised with 50 μm steps until sound responsive units were observed by for bonding.

Head Fixed Recording System—Recordings of MGBv and audTRN sensory responses were conducted in a custom head-fixation/sound-delivery system enclosed in a ventilated sound-proof chamber (IAC Acoustics, North Aurora, IL). The head-fixation system consisted of a pair of custom 3D printed plastic fixation clamps (MakerBot Replicator, Brooklyn NY) used to lock the implanted plastic crown at the base of the implant into place during recordings. These were fixed to an acrylic plastic frame which also supported a platform on which the animal stood. The platform was composed of low-friction acrylic and was adjusted based on the height of the animal and spring-loaded to minimize torque on the implant. For head fixed recordings, stimuli were delivered from a pair of electrostatic speakers on either side of the animal via straight, plastic tubes 2 cm long which terminated 2.5 mm from the left and right ears.

Electrophysiological Recordings—Signals from tetrodes were acquired using a Neuralynx multiplexing digital recording system (Neuralynx, Bozeman MT) via a combination of 32- and 64-channel digital multiplexing headstages plugged to the 96- or 128-channel EIB of each implant. Signals from each electrode were amplified, filtered between 0.1 Hz and 9 kHz and digitized at 30 kHz. For audTRN recordings, tetrodes were lowered over the course of 1–2 weeks from the cortex into the target structure. For PFC recordings, adjustments were more targeted, consistent with the more superficial position of the region of interest. The system used for recordings (head fixed and in-behavior) was entirely automated so no investigator blinding of genotype or drug conditions was required for electrophysiological experiments. Following acquisition, spike sorting was performed offline based on relative spike amplitude and energy within electrode pairs automatically using MountainSort as previously described (Chung et al., 2017; Rikhye et al., 2018). After initial clustering, units were divided into fast spiking (FS) and regular spiking (RS) based on waveform characteristics as previously described (Halassa et al., 2014). Briefly, peak to trough time was measured in all spike waveforms, and showed a distinct bimodal distribution (Hartigan's dip test, $p < 10^{-5}$). These distributions separated at 187 μs, and cells with peak to trough times above this threshold were considered RS while those with peak to trough times below were considered FS cells. This initial identification was subsequently further validated in two feature dimensions (Half Trough time vs Peak to Trough time) using k-means clustering which showed good agreement with the single dimension separation (97% overlap in cell selection).

MGBv and audTRN Specific Methods: For recordings made within the MGB, many units were observed in which the spike waveform showed characteristics of an FS neuron, despite the absence of interneurons in this structure (Winer and Larue, 1996). Based on previous results, we considered that these spikes reflected terminal responses from audTRN

projections (Barthó et al., 2014). This idea was supported by three lines of evidence. First, the waveform features of FS neurons in MGBv and opto-tagged audTRN neurons were similar (Fig. S3C, D). Second, responses of FS neurons were reduced when optogenetic suppression of audTRN neurons was delivered via an optic fiber position above the portion of the TRN where there soma are generally located (Fig. S3E, F). Third, in contrast to RS units, FS units in the MGB tended to show complex, multi-peaked sound responses (Fig.S3G). Based on these observations, we considered single units showing FS-like waveforms to correspond to audTRN neuron activity.

In MGB, audTRN and audTRN terminal recordings, neurons were considered sound responsive if their firing rate was significantly elevated across at least 20 percent of the stimulus period (8×25 ms bins). For somatic recordings, MGB projecting TRN neurons (audTRN) were identified using retrograde optogenetic tagging resulting in expression of either ChR2 or NpHR 3.0. Neurons were considered tagged if their firing rate showed a significant increase (ChR2) or decrease (NpHR 3.0) in firing rate within 25 ms of laser pulse onset. Significance testing was based on estimation of the 95% confidence intervals for peri-stimulus time histograms (PSTHs) generated from 100 sound stimuli or laser pulses.

For firing rate estimates, as well as Fano factor quantification and some of the included population decoding (Fig. 5, 6), stimuli consisted of 20 dynamic random chord (DRC) sound stimuli repeated 20 times each. To record responses to noisy sound stimuli, same number of DRC stimuli were repeated 20 times each with added broadband noise (SNR 3.2). Firing rates for the “sound evoked responses” were obtained during stimulus (DRC or DRC + noise) presentation and were normalized to “spontaneous” firing rates estimated over the 250 msec period prior to the onset of sound or laser on optogenetic manipulations. For optogenetic manipulations, laser onset preceded the onset of sound by 250 msec with a ramping offset to prevent rebound spiking. To estimate the effect of adding broadband noise, we calculated the change in firing rate produced by noise by subtracting the average firing rate for each DRC in DRC only trials from the average firing rate in the corresponding DRC +Noise trials.

PFC Specific Methods: For PFC recordings made in animals performing either the cross-modal or cued noisy auditory discrimination tasks, we identified $PFC_{visStriatum}$ and $PFC_{audStriatum}$ based on their projections using a retrograde optogenetic tagging approach. For this identification, CAV2-Cre virus was injected into visStriatum or audStriatum and an AAV virus harboring Cre-dependent ChR2-EYFP was injected into PFC, resulting in expression of ChR2 in each population. This allowed us to identify neurons based on their response to optogentic stimulation, which was performed after each behavioral session. To assess how consistently neurons responded to multiple pulse of stimuli, 50 trains (at 10 s intervals) of 5 light pulses (10ms pulses 473 nm laser light delivered at 50 Hz via angled optic fibers within the implanted drives, see Microdrive array construction and implantation, above) were delivered after each behavioral session. For this optogenetic tagging, mice were removed from the behavioral enclosure and placed in an adjacent enclosure. Neurons were considered to be potentially tagged if their firing rate showed a significant increase in firing rate within 20 ms of the first laser pulse. To ensure that we considered neurons expressing ChR2 (directly tagged), as opposed to those which responded due to indirect activation due

to input from directly tagged neurons (post-synaptic), additional analytical steps were employed to separate these populations (see Fig. 2). We used two independent measures to designed to assess 1. Response latency (evoked response onset, measured as the time to half maximal response) and 2. Response timing consistency (peak jitter, measured as the median separation of spike times from the time of peak evoked firing rate) of potentially tagged neurons. Using these quantifications of optogenetic response properties we were able to separate the different populations using a clustering based approach (via the k-means clustering function implemented in MATLAB). Based on previous approaches (Lima et al., 2009) we assumed that directly tagged neurons would have smaller response latencies (smaller evoked response onset) and more consistent responses (smaller peak jitter).

To assess response latency, we first constructed a high resolution PSTH (2 msec bin size, interpolated via linear interpolation to 1 msec temporal resolution) for each potentially tagged neuron. The first peak after each laser pulse was then identified using the findpeaks function in MATLAB (MathWorks, Natick, MA). The onset of the response (evoked response onset) then determined as the time after laser pulse onset at which the PSTH reached half the firing rate at the identified peak. Next, we assessed the consistency of the response by measuring the difference between the peak response time across all trials (identified with the findpeaks function) and the nearest spike in each trial (peak jitter). We found that these measures clustered into two easily distinguished groups (see Fig. 2E) which were separated via k-means clustering. A neuron was considered tagged if its peak jitter and evoked response onset appeared in the cluster with the smaller centroid values for both axis.

Fano Factor—Fano factor values were computed for each MGBv neuron based on their responses across 20 repeated deliveries of each DRC stimulus. This computation was made using MATLAB code which is included in the variance toolbox (available online at https://churchland.zuckermaninstitute.columbia.edu/content/code) as described previously (Churchland et al., 2010). Briefly, spike counts were computed in 25 ms windows aligned to the chords of the DRC for each trial. Spike count means and variance were then computed across trials. The mean and variance across DRC stimulus chords were compiled and fitted with a regression line. The slope of this line was taken as the Fano factor for this cell. This “raw” Fano factor was used across neurons and conditions (audTRN suppression or no laser) for comparisons.

Identification of peaks in task-modulated neurons—In assessing delay period responses in both the cross-modal task (as per our previous results) and the cued noisy discrimination task we did not observe individual PFC neurons that exhibited sustained increases in spiking relative to baseline. For both tasks, however, a subset of cells showed a brief elevation (peak) of spiking activity at a defined moment in the delay period (Fig. 2, 3, 7). Across both tasks, these neurons were identified based on consistency in their spike timing across correct trials, as well as cross-trial elevation in spike rate as follows:

First, periods of increased consistency in spike-timing across trials were identified using a Matching-Minimization algorithm (Wu and Srivastava, 2011). This was used to determine the best moments of spike time alignment across trials (putative peak times). These putative peak times were obtained as the solution of the equation:

$$S = \arg_{C \in S} \min \sum_{k=1}^{N} d_2(S_k, C)^2$$

In which the putative peak times across trains (S) was obtained by minimizing the sum of the distance function (d_2) of the observed spike trains (S_k) and the current peak time estimate (C) given a set of penalty coefficients associated with spike time translation. Initially, spikes in C were placed arbitrarily within the sample window. The number of spikes (n) included in the initial estimate was obtained by minimizing the equation:

$$\sum_{k=1}^{N} |n_k - n|$$

Where n_k is the number of spikes in a trial and n is the number of putative peak locations. Thus the initial number of putative peaks is equal to the median number of spikes observed within the sample period across spike trains. From this starting condition, putative peak times were iteratively adjusted to minimize the distance function (d_2) between the observed spikes and the putative peak-time estimate. This adjustment was based on the relative ISI values of the peak time estimate (f) and the spike train for each trial (g) based on the equation:

$$d_p(f,g) = \left(\lambda \sum_{k=1}^{M+1} |(\Delta gs_k)^{\frac{1}{p}} - (\Delta fs_k)^{\frac{1}{p}}|^p \right)^p$$

Where Δgs_k and Δfs_k are vectors of the interspike intervals associated with the peak time estimate and the spike train of a given trial, respectively, M is equal to the total number of spikes, λ is the cost penalty weight and p is the comparison parameter (in this case 2 for pairwise comparisons). If the distance for a given putative peak in the estimate was optimal (e.g. the distance is at the local minimum) then it was left in place, otherwise it was moved via linear interpolation between its current location and the measured spike times across trials, placing it in the center of the interpolation line. Finally, the overall distance was minimized by adjusting the interspike interval using the spike ISI average metric, a solution to the minimization equation (above). The putative peak times were updated using the newly calculated ISIs after which the variance was computed. These steps were iterated until the variance converged, that is no observed decrease in variance was produced by repeating the algorithm. The resulting spike times were taken as putative peak locations.

To determine whether a peak occurs at any of these putative locations, we applied two further criteria. First, for 75% of the trials, at least one spike must fall within +/− 25 ms of the putative peak time. This conservative threshold was based on the median firing rates observed during the task period which in the majority of cells is less than 10 Hz predicting that the most spike intervals between trials will be greater than 50 ms (½ peak ISI for this firing rate). Second, we incorporated a z-score criterion which is sensitive to changes in the

number of spikes occurring in a particular time-bin across trials. The z-score was computed relative to the pre-delay baseline (10 ms binning, convolved with a 25 ms half-width gaussian kernel). If the z-score within the 50 ms window surrounding the potential peak exceeded 1.5 for a cell meeting the first criterion, then this time point was considered a true peak and the cell identified as a task-modulated unit.

Decoding Analysis—To assess tone representation in the MGBv and predictive/unpredictive cue representations in the PFC we applied a population decoding approaches, the Poisson Naïve Bayes (PNB) classifier, as implemented in the neural decoding toolbox (Meyers, 2013). This analysis was applied to sound responsive MGBv neurons as well as to all neurons in the PFC recorded during behavior. In each case, neurons recorded from either structure each of which were pooled into a pseudo-population for each subset of cells within a condition (e.g. MGBv with and without audTRN suppression or GP suppression or PFC recordings for each trial types in behavior). For classification, neuron spiking activity was modeled as a Poisson random variable with each neuron's activity assumed to be independent. The model was based on spike counts of these pseudo-populations (Meyers, 2013). Analysis was performed based either on spiking associated with stimulus (MGBv) or cue type (PFC) as described below.

PFC Decoding for cross-modal sensory-selection task: Analysis of data recorded in these sessions was limited to optogenetically tagged neurons which we identified as putative visStriatum projecting ($PFC_{visStriatum}$) or audStriatum projecting ($PFC_{audStriatum}$). Spike trains from these populations, which included spiking 1 second before and 1.5 seconds after initiation, from correct trials were first divided into those in which animals were cued to *attend to audition* and those in which they were cued to *attend to vision*. To assess the degree to which each rule was encoded separately, a third 'jittered-control' spike train was generated for each neuron in which average firing rate was preserved but task-related changes in spiking were eliminated. This was produced by first jittering all recorded spikes for a given neuron in each session by a time step drawn for each spike from a uniform distribution (range = ± 250 ms). Spike trains were then sampled from this jittered spiking data over the same time ranges used for the real data. Although the number of trials included in the jitter-control trains was the same as the number of correct trials for a given cue type, these trials were randomly sampled irrespective of trial type or performance. Decoding was performed separately for visual and auditory cue types against their respective jitter-control data. The effectiveness of this approach to capture different types of cue related information is supported by the differential encoding of the two rules by $PFC_{visStriatum}$ neurons (see Fig. 2I).

To train the classifier, the included spike trains were repeatedly and randomly subsampled (60 resampling runs) and divided into training and test subsets (10 trials sampled 6 used for training and 4 for testing). For each subsampling, the classifier was trained using the training subset to produce a predictive mean response template ($\bar{x}$) for each trial type (i). Templates were constructed separately for 50 ms overlapping windows across the trace (step size = 25 ms) and with the classifier trained for each template allowing a temporal profile to be estimated. In the cross-validation step, these templates were used to predict the class for

each test trial in the test set (x^*) by maximizing the log likelihood decision function ($\left[i^* = arg\ \max_i \mathrm{LL}(x^*, \bar{x})\right]$) via well-established methods (Duda et al., 2000) assuming a Poisson distribution for firing rates. The overall likelihood value can then be calculated by multiplying the probabilities for each neuron together (under the assumption that each feature is independent). The prediction accuracy (decision values) were quantified as normalized rank in the posterior probability list (Meyers, 2013; Mitchell et al., 2004). To determine the variability of this estimate, a bootstrapping procedure was applied in which 60% of neurons were subsampled from the overall population and the same procedure was repeated (50 resampling runs). The resulting traces were used to estimate the 95% confidence intervals of the initial estimate from the full population. For group comparisons (such as those in Fig. 1R), prediction accuracy was estimated based on the activity of these subsamples at each time point during the delay period and the decoding accuracy was grouped across time points/subsamples. The relative specificity of encoding visual versus auditory cues for $PFC_{visStriatum}$ and $PFC_{audStriatum}$ populations was quantified using an "Encoding Selectivity Index" (Fig. S2D). This index was derived from the population estimates by comparing encoding in each subset of neurons (60% subsampling as described above) for the two cues. To adjust this metric to a 100% scale, we first subtracted the chance level of the decoder (50%) from either type of decoding. Next we took the absolute value of the difference in decoding accuracy (Δdecoding) between visual and auditory cues (within subsets). Finally, to compare the two PFC populations, we divided Δdecoding values for both $PFC_{visStriatum}$ and $PFC_{audStriatum}$ populations by the average of Δdecoding values for the $PFC_{visStriatum}$ (as this distribution had the larger mean Δdecoding).

MGBv Decoding: To effectively assess the encoding of sounds relevant to behavior, we decoded the activity of neural populations recorded in the MGBv that represents sounds used in behavior. We focused on decoding pure tones with frequencies of 20 kHz and 24 kHz, the more difficult discrimination. To quantify the effect of noise, we assessed the response to these pure tones with added noise. Spike trains were taken from MGBv responses elicited by three types of stimuli: 1. tones embedded in DRCs, 2. tones presented alone or 3. Frequency "sweeps" (up-sweep (10–15kHz) or a down-sweep (16–12kHz)) used in the 2 AFC behavior. In the first case, a subset of "chords" within the DRC set were used in which only the tone of interest was presented. A total of 12 such "chords", 6 per behaviorally related tone, were present in the DRCs each of which was presented 20 times (tone length = 25 ms, 120 trials per tone) either with or without noise (SNR 3.2, used in Fig. 5B or SNR 1.8, used if Fig. 6H). To more efficiently evaluate the effect of noise, 25 ms tones were presented alone in some recordings (used in Fig. 6H) flanked by 100 ms of silence or added noise alone (100 repetitions per noise condition). In these two cases, spike trains of MGBv responses used for decoding began 25 ms before the tone onset and terminated 50 ms after (75 ms total). Finally, MGBv population encoding of sound sweeps (used in Fig. 3L) was performed using responses during behavior beginning 25 ms before the sweep onset and terminating 25 ms after (150 ms total). Only correct trials were used and sessions were only included if greater than 20 correct trials per condition were available.

To train the classifier, spike trains were repeatedly and randomly subsampled (60 resampling runs) and divided into training and test subsets (10 trials sampled 6 used for training and 4

for testing). For each subsampling, the classifier was trained using the training subset to produce a predictive mean response template ($\bar{x}$) for each stimulus tone (i). Templates were constructed separately for 25 ms time bins across the trace (step size = 25 ms) and with the classifier trained for each template. In the cross-validation step, these templates were used to predict the class for each test trial in the test set (x^*) by maximizing the log likelihood decision function ($\left[i^* = arg\ \max_i \mathrm{LL}(x^*, \bar{x})\right]$) as described previously(Duda et al., 2000). The overall likelihood value can then be calculated by multiplying the probabilities for each neuron together (under the assumption that each neuron is independent). The prediction accuracy (decision values) were quantified as normalized rank in the posterior probability list (Meyers, 2013; Mitchell et al., 2004). We used the maximum prediction accuracy within the second and third time bins corresponding to the first 50 ms after tone onset. This allows us to look at activity elicited by the tone at various offsets during which it is physiologically plausible that the tone is encoded in the MGBv (Anderson et al., 2009). To determine the variability of this estimate, a bootstrapping procedure was applied in which 60% of neurons were subsampled from the overall population and the same procedure was repeated (50 resampling runs) and the maximum prediction accuracy in the latter two time bins was estimated based on the activity of these subsamples. These decoding accuracy values were grouped across time points/subsamples for each condition and used for comparisons (Fig. 5C, 6H, 3L).

PFC Decoding for cued noisy auditory discrimination: Spike trains, which included spiking 1 second before and 1.5 seconds after initiation, from correct trials were first divided into those in which animals were cued with a predictive cue, those in which an unpredictive cue was presented and those in which no cue was presented. Decoding was performed comparing trials in which the predictive cue was delivered with those in which the unpredictive cue was delivered. This approach was chosen to reduce the chance that stimulus related activity, rather than the rule meaning, was primarily responsible for the information decoded from the population. The effectiveness of this approach was supported by the observation that the unpredictive cue could not be decoded from uncued trials based on population activity (see Fig. 7M).

To train the classifier, spike trains were repeatedly and randomly subsampled (60 resampling runs) and divided into training and test subsets (10 trials sampled 6 used for training and 4 for testing). For each subsampling, the classifier was trained using the training subset to produce a predictive mean response template ($\bar{x}$) for each trial type (i). Templates were constructed separately for 50 ms overlapping windows across the trace (step size = 25 ms) and with the classifier trained for each template allowing a temporal profile to be estimated. In the cross-validation step, these templates were used to predict the class for each test trial in the test set (x^*) by maximizing the log likelihood decision function ($\left[i^* = arg\ \max_i \mathrm{LL}(x^*, \bar{x})\right]$) as described previously (Duda et al., 2000). The overall likelihood value can then be calculated by multiplying the probabilities for each neuron together (under the assumption that each feature is independent). The prediction accuracy (decision values) were quantified as normalized rank in the posterior probability list (Meyers, 2013; Mitchell et al., 2004). To determine the variability of this estimate, a bootstrapping procedure was applied in which 60% of neurons were subsampled from the overall population and the same

procedure was repeated (50 resampling runs). The resulting traces were used to estimate the 95% confidence intervals of the initial estimate from the full population.

To determine whether optogenetically tagged $PFC_{audStriatum}$ neurons were more likely to be engaged in encoding the predictive cue than the average PFC neuron we assessed task variable decoding from different sized subsets of these neurons from either $PFC_{audStriatum}$ or non-tagged neurons. More specifically, we assessed how the encoding of the predictive cue (against the unpredictive cue) varied with the number of neurons included. To do this, we randomly drew increasing numbers of neurons from each group (i.e. subsets 2,3,4 … 100 $PFC_{audStriatum}$ or non-tagged neurons) and assessed the maximum decoding accuracy from the resulting subsets using the approach described above. For each size of subset, this approach was repeated 25 times, allowing us to construct confidence intervals across subsets.

To assess the encoding of task relevant information among $PFC_{audStriatum}$ or non-tagged neurons, we used an approach intended to match the number of neurons included while also ensuring that neurons included showed task relevant activity (Fig. S5C). More specifically, neurons recorded during task performance were sorted based on their maximum change in spiking rate during delay periods in which either a predictive or unpredictive cue was delivered, normalized relative to baseline rates. Only the top 100 neurons from each group was used for subsequent analysis.

Histology—To examine tracing results, mice were deeply anesthetized and transcardially perfused with phosphate buffered saline (PBS) followed by 4% paraformaldehyde. Brains were dissected, postfixed overnight at 4°C and sectioned in 50μm thickness using a vibratome (LEICA, Buffalo Grove, IL). All sections were imaged on a Zeiss LSM710 confocal microscope (Zeiss, Oberkochen, Germany).

Quantification and Statistical Analysis

Behavior: For behavioral studies, preliminary studies provided sufficient information on effect size so that power analyses could be performed to determine the number of mice and sessions needed. The sample number needed was estimated using power analysis in MATLAB (sampsizepwr) with a β of 0.7 (70%). Using this strategy, the required number of animals was determined to be between 3 and 6 mice per cohort across testing conditions, with >/= 4 sessions per animal. For multiple comparisons, non-parametric ANOVA (Kruskal-Wallis H-test), repeated measures 2-way ANOVA (Friedman Test) was performed followed by pairwise post-hoc analysis. For behavioral data obtained for multiple optogenetic manipulations in the cross-modal sensory selection task, all conditions were initially grouped and analyzed via a single multiple ANOVA (MANOVA) prior to pairwise testing. This included baseline variance for individual groups estimated as follows: Baseline variance was estimated by sub-sampling alternating trials with no optogenetic manipulation for each cue type (visual and auditory) in individual sessions for each type of optogenetic manipulation. The performance for each cue type in one sub-sample was then subtracted from that of the other and these artificial “performance deltas” were used to estimate change from baseline due to chance. All post-hoc pairwise comparisons, including comparisons

made between optogenetic manipulations and variation in control conditions, used non-parametric rank-sum (unpaired samples) or sign-rank (paired samples) tests.

Spiking Data: As for behavioral data, with the exception of visual TRN neural recordings that were conducted previously (Wimmer et al., 2015)), preliminary studies provided sufficient information on effect size so that power analyses could be performed to determine the number of mice and sessions needed. The sample number needed was estimated using power analysis in MATLAB (sampsizepwr) with a β of 0.7 (70%). For all included electrophysiological experiments, three sessions were recorded from a single mouse and used for power analysis based on the effect size observed in this preliminary cohort. For each statement of statistical difference included in the manuscript, an appropriate statistical comparison was performed. As in behavioral experiments, for multiple comparisons, non-parametric ANOVA (Kruskal-Wallis H-test), repeated measures 2-way ANOVA (Friedman Test) was performed followed by pairwise post-hoc analysis. For large sample sets in electrophysiological recordings, the Kolmogorov–Smirnov normality test was first performed on the data to determine whether parametric or non-parametric tests were required. However, non-parametric tests (rank-sum or sign-rank tests) were used by default for pairwise comparisons.

Chloride Photometry: For statistical analysis, individual parts of the traces associated with behavioral trials were separated based on their relationship to the stimulus and initiation. Prior to extracting individual components of the signal, the overall trace was smoothed with a 100 ms Gaussian filter to reduce noise. A diagram illustrating distinct response components of an ideal single trial response in the cued noisy discrimination behavior is shown in Figure 7H.

Because the chloride signal is negative going, components were multiplied by negative one to place the signal in the positive range. The components extracted from trial signals include: 1. Anticipation, 2. Stimulus and 3. Evoked. The first component, anticipation, was taken as the difference from baseline (equal to the average of the 500 ms prior to initiation) for the average of the over the 100 ms prior to stimulus presentation. The second component, associated with the stimulus, was taken as the difference between the anticipation signal and the minimum value in the 100 ms that follow stimulus presentation. This was done to separate chloride signal engaged by the sound stimulus from changes that might be due to anticipation of noise (see Fig. 7H for illustration). Finally, the evoked signal, corresponding to the peak chloride signal during stimulus, was taken as difference between the minimum value in the 100 ms following stimulus presentation from the original baseline (again equal to the average of the 500 ms prior to initiation).

Once extracted, these signal components were grouped by genotype and condition and used for subsequent analysis. For all statistical analysis, N values used corresponded to the number of mice. For average traces shown, 95% confidence intervals were calculated based on session averages within mice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank S. N. Oline and R.D. Wimmer for technical assistance. M.N. is supported by a JSPS fellowship. L.I.S. is supported by a NARSAD Young Investigator award. M.M.H. is supported by grants from NIMH, NINDS, Brain and Behavior, Simons, Sloan and Klingenstein Foundations as well as the Human Frontiers Science Program.

References

Anderson LA, Christianson GB, and Linden JF (2009). Stimulus-specific adaptation occurs in the auditory thalamus. J Neurosci 29, 7359–7363. [PubMed: 19494157]

Angelucci A, and Bressloff PC (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog Brain Res 154, 93–120. [PubMed: 17010705]

Barthó P, Slézia A, Mátyás F, Faradzs-Zade L, Ulbert I, Harris KD, and Acsády L (2014). Ongoing network state controls the length of sleep spindles via inhibitory activity. Neuron 82, 1367–1379. [PubMed: 24945776]

Brouwer GJ, and Heeger DJ (2011). Cross-orientation suppression in human visual cortex. J Neurophysiol 106, 2108–2119. [PubMed: 21775720]

Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Schölvinck ML, Zaharia AD, and Carandini M (2011). The detection of visual contrast in the behaving mouse. J Neurosci 31, 11351–11361. [PubMed: 21813694]

Carandini M, and Heeger DJ (2011). Normalization as a canonical neural computation. Nat Rev Neurosci 13, 51–62. [PubMed: 22108672]

Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, and Nelken I (2006). Reduction of information redundancy in the ascending auditory pathway. Neuron 51, 359–368. [PubMed: 16880130]

Chung JE, Magland JF, Barnett AH, Tolosa VM, Tooker AC, Lee KY, Shah KG, Felix SH, Frank LM, and Greengard LF (2017). A Fully Automated Approach to Spike Sorting. Neuron 95, 1381–1394.e1386. [PubMed: 28910621]

Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13, 369–378. [PubMed: 20173745]

Cohen MR, and Maunsell JH (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12, 1594–1600. [PubMed: 19915566]

Desimone R, and Duncan J (1995). Neural mechanisms of selective visual attention. Annu Rev Neurosci 18, 193–222. [PubMed: 7605061]

Donahue CH, and Lee D (2015). Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. Nat Neurosci 18, 295–301. [PubMed: 25581364]

Dragoi V, Sharma J, Miller EK, and Sur M (2002). Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. Nat Neurosci 5, 883–891. [PubMed: 12161755]

Duda RO, Hart PE, and Stork DG (2000). Pattern classification, 2nd edn (New York; Chichester: Wiley-Interscience).

Dunn FA, Doan T, Sampath AP, and Rieke F (2006). Controlling the gain of rodmediated signals in the Mammalian retina. J Neurosci 26, 3959–3970. [PubMed: 16611812]

El-Gaby M, Zhang Y, Wolf K, Schwiening CJ, Paulsen O, and Shipton OA (2016). Archaerhodopsin Selectively and Reversibly Silences Synaptic Transmission through Altered pH. Cell Rep 16, 2259–2268. [PubMed: 27524609]

Gilbert CD, and Li W (2013). Top-down influences on visual processing. Nat Rev Neurosci 14, 350–363. [PubMed: 23595013]

Gregoriou GG, Gotts SJ, and Desimone R (2012). Cell-type-specific synchronization of neural activity in FEF with V4 during attention. Neuron 73, 581–594. [PubMed: 22325208]

Grimley JS, Li L, Wang W, Wen L, Beese LS, Hellinga HW, and Augustine GJ (2013). Visualization of synaptic inhibition with an optogenetic sensor developed by cell-free protein engineering automation. J Neurosci 33, 16297–16309. [PubMed: 24107961]

Gunaydin LA, Grosenick L, Finkelstein JC, Kauvar IV, Fenno LE, Adhikari A, Lammel S, Mirzabekov JJ, Airan RD, Zalocusky KA, et al. (2014). Natural neural projection dynamics underlying social behavior. Cell 157, 1535–1551. [PubMed: 24949967]

Guo W, Clause AR, Barth-Maron A, and Polley DB (2017). A Corticothalamic Circuit for Dynamic Switching between Feature Detection and Discrimination. Neuron 95, 180–194.e185. [PubMed: 28625486]

Hafter, Sarampalis, and Loui (2008). Auditory Attention and Filters, Vol 29 (Springer).

Halassa MM, Chen Z, Wimmer RD, Brunetti PM, Zhao S, Zikopoulos B, Wang F, Brown EN, and Wilson MA (2014). State-dependent architecture of thalamic reticular subnetworks. Cell 158, 808–821. [PubMed: 25126786]

Heekeren HR, Marrett S, and Ungerleider LG (2008). The neural systems that mediate human perceptual decision making. Nat Rev Neurosci 9, 467–479. [PubMed: 18464792]

Hunnicutt BJ, Jongbloets BC, Birdsong WT, Gertz KJ, Zhong H, and Mao T (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. Elife 5.

Jiang H, and Kim HF (2018). Anatomical Inputs From the Sensory and Value Structures to the Tail of the Rat Striatum. Front Neuroanat 12, 30. [PubMed: 29773980]

Kato S, Kobayashi K, Inoue K, Kuramochi M, Okada T, Yaginuma H, Morimoto K, Shimada T, and Takada M (2011). A lentiviral strategy for highly efficient retrograde gene transfer by pseudotyping with fusion envelope glycoprotein. Hum Gene Ther 22, 197–206. [PubMed: 20954846]

Liang L, Oline SN, Kirk JC, Schmitt LI, Komorowski RW, Remondes M, and Halassa MM (2017). Scalable, Lightweight, Integrated and Quick-to-Assemble (SLIQ) Hyperdrives for Functional Circuit Dissection. Front Neural Circuits 11, 8. [PubMed: 28243194]

Lima SQ, Hromádka T, Znamenskiy P, and Zador AM (2009). PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. PLoS One 4, e6099. [PubMed: 19584920]

McAlonan K, Cavanaugh J, and Wurtz RH (2008). Guarding the gateway to cortex with attention in visual thalamus. Nature 456, 391–394. [PubMed: 18849967]

Menegas W, Babayan BM, Uchida N, and Watabe-Uchida M (2017). Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. Elife 6.

Meyers EM (2013). The neural decoding toolbox. Front Neuroinform 7, 8. [PubMed: 23734125]

Miller EK, and Cohen JD (2001). An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24, 167–202. [PubMed: 11283309]

Mitchell TM, Hutchinson R, Niculescu RS, Pereira F, Wang X, Just M, and Newman S (2004). Learning to Decode Cognitive States from Brain Images. Machine Learning 57, 145–175.

Moore T, and Fallah M (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. J Neurophysiol 91, 152–162. [PubMed: 13679398]

Noudoost B, Chang MH, Steinmetz NA, and Moore T (2010). Top-down control of visual attention. Curr Opin Neurobiol 20, 183–190. [PubMed: 20303256]

O'Connor DH, Fukui MM, Pinsk MA, and Kastner S (2002). Attention modulates responses in the human lateral geniculate nucleus. Nat Neurosci 5, 1203–1209. [PubMed: 12379861]

Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, et al. (2014). A mesoscale connectome of the mouse brain. Nature 508, 207–214. [PubMed: 24695228]

Olsen SR, Bhandawat V, and Wilson RI (2010). Divisive normalization in olfactory population codes. Neuron 66, 287–299. [PubMed: 20435004]

Parr T, and Friston KJ (2018). Attention or salience? Curr Opin Psychol 29, 1–5. [PubMed: 30359960]

Pazo JH, Barceló AC, Bellantonio E, Pazo VC, and Almarares N (2013). Electrophysiologic study of globus pallidus projections to the thalamic reticular nucleus. Brain Res Bull 94, 82–89. [PubMed: 23500178]

Petersen SE, and Posner MI (2012). The attention system of the human brain: 20 years after. Annu Rev Neurosci 35, 73–89. [PubMed: 22524787]

Phillips JM, Kambi NA, and Saalmann YB (2016). A Subcortical Pathway for Rapid, Goal-Driven, Attentional Filtering. Trends Neurosci 39, 49–51. [PubMed: 26743499]

Pinto L, Goard MJ, Estandian D, Xu M, Kwan AC, Lee SH, Harrison TC, Feng G, and Dan Y (2013). Fast modulation of visual perception by basal forebrain cholinergic neurons. Nat Neurosci 16, 1857–1863. [PubMed: 24162654]

Reynolds JH, and Heeger DJ (2009). The normalization model of attention. Neuron 61, 168–185. [PubMed: 19186161]

Rikhye RV, Gilra A, and Halassa MM (2018). Thalamic regulation of switching between cortical representations enables cognitive flexibility. Nat Neurosci 21, 17531763.

Romo R, and de Lafuente V (2013). Conversion of sensory signals into perceptual decisions. Prog Neurobiol 103, 41–75. [PubMed: 22472964]

Schmitt LI, Wimmer RD, Nakajima M, Happ M, Mofakham S, and Halassa MM (2017). Thalamic amplification of cortical connectivity sustains attentional control. Nature 545, 219–223. [PubMed: 28467827]

Simoncelli EP, and Heeger DJ (1998). A model of neuronal responses in visual area MT. Vision Res 38, 743–761. [PubMed: 9604103]

Squire RF, Noudoost B, Schafer RJ, and Moore T (2013). Prefrontal contributions to visual selective attention. Annu Rev Neurosci 36, 451–466. [PubMed: 23841841]

Sun Y, Nguyen AQ, Nguyen JP, Le L, Saur D, Choi J, Callaway EM, and Xu X (2014). Cell-type-specific circuit connectivity of hippocampal CA1 revealed through Cre-dependent rabies tracing. Cell Rep 7, 269–280. [PubMed: 24656815]

Szczepanski SM, and Knight RT (2014). Insights into human behavior from lesions to the prefrontal cortex. Neuron 83, 1002–1018. [PubMed: 25175878]

Treisman M, and Faulkner A (1985). ON THE CHOICE BETWEEN CHOICE THEORY AND SIGNAL-DETECTION THEORY. Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology 37, 387–405.

Verhoef BE, and Maunsell JH (2016). Attention operates uniformly throughout the classical receptive field and the surround. Elife 5.

Verhoef BE, and Maunsell JHR (2017). Attention-related changes in correlated neuronal activity arise from normalization mechanisms. Nat Neurosci 20, 969–977. [PubMed: 28553943]

Villalobos N, Oviedo-Chávez A, Alatorre A, Ríos A, Barrientos R, Delgado A, and Querejeta E (2016). Striatum and globus pallidus control the electrical activity of reticular thalamic nuclei. Brain Res 1644, 258–266. [PubMed: 27208494]

Wall NR, Wickersham IR, Cetin A, De La Parra M, and Callaway EM (2010). Monosynaptic circuit tracing in vivo through Cre-dependent targeting and complementation of modified rabies virus. Proc Natl Acad Sci U S A 107, 21848–21853. [PubMed: 21115815]

Wark B, Lundstrom BN, and Fairhall A (2007). Sensory adaptation. Curr Opin Neurobiol 17, 423–429. [PubMed: 17714934]

Warrant EJ (2016). Sensory matched filters. Curr Biol 26, R976–R980. [PubMed: 27780072]

Wells MF, Wimmer RD, Schmitt LI, Feng G, and Halassa MM (2016). Thalamic reticular impairment underlies attention deficit in Ptchd1(Y/-) mice. Nature 532, 58–63. [PubMed: 27007844]

Wiegert JS, Mahn M, Prigge M, Printz Y, and Yizhar O (2017). Silencing Neurons: Tools, Applications, and Experimental Constraints. Neuron 95, 504–529. [PubMed: 28772120]

Willmore BD, Cooke JE, and King AJ (2014). Hearing in noisy environments: noise invariance and contrast gain control. J Physiol 592, 3371–3381. [PubMed: 24907308]

Wimmer RD, Schmitt LI, Davidson TJ, Nakajima M, Deisseroth K, and Halassa MM (2015). Thalamic control of sensory selection in divided attention. Nature 526, 705–709. [PubMed: 26503050]

Winer JA, and Larue DT (1996). Evolution of GABAergic circuitry in the mammalian medial geniculate body. Proc Natl Acad Sci U S A 93, 3083–3087. [PubMed: 8610172]

Wu W, and Srivastava A (2011). An information-geometric framework for statistical inferences in the neural spike train space. J Comput Neurosci 31, 725–748. [PubMed: 21584775]

Highlights

- A Prefrontal-Basal Ganglia-thalamus pathway controls inhibition in sensory thalamus
- This pathway enables sensory selection by suppressing distracting inputs
- Goal-directed noise filtering via thalamic inhibition enhances discrimination

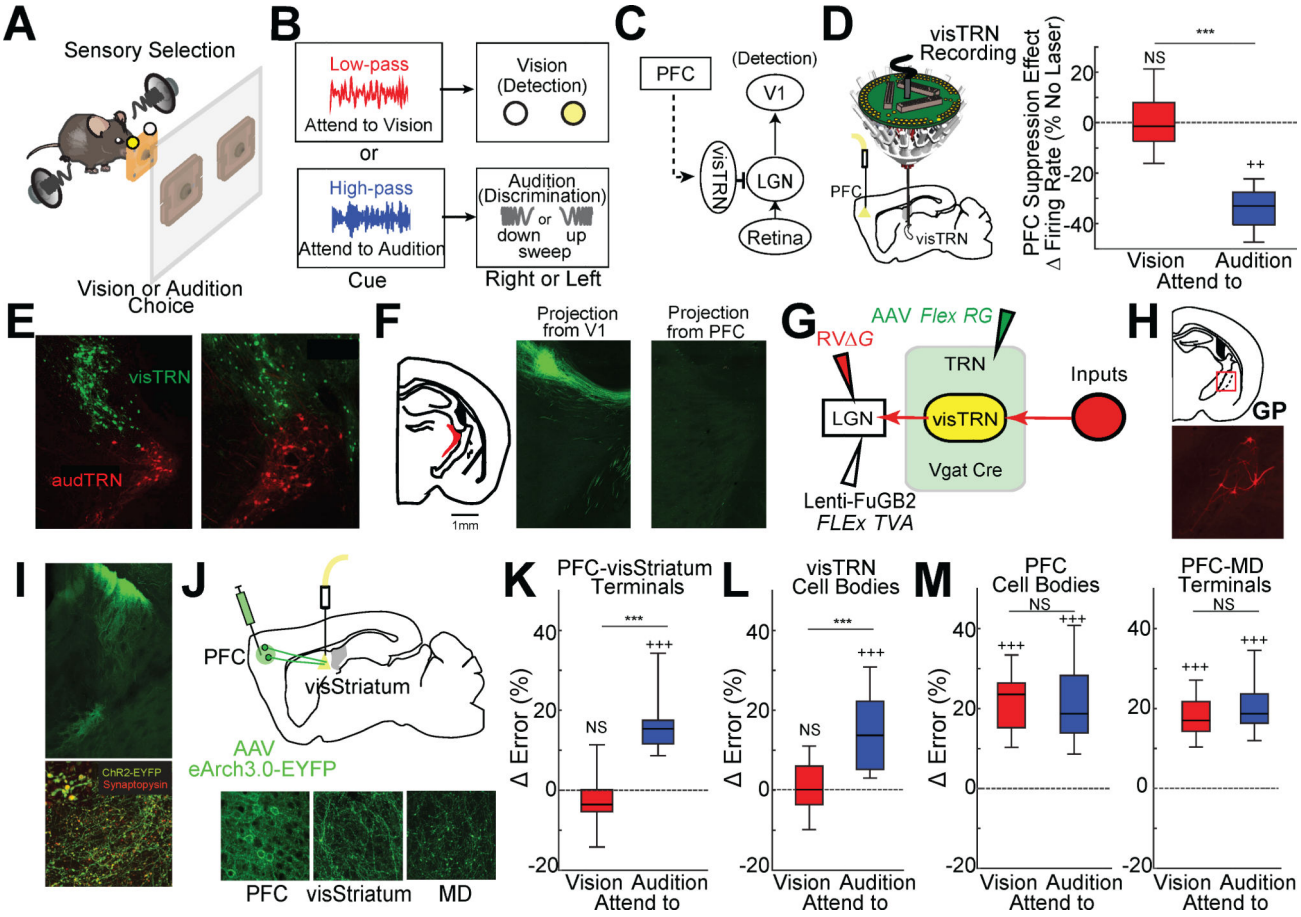

Figure 1: PFC suppresses behaviorally-irrelevant thalamic visual inputs via a basal ganglia pathway

A. Cartoon of the 2AFC crossmodal task. Freely behaving mice are simultaneously presented with two sensory targets (100msec of a visual flash and auditory sweep) to indicate reward port location. The relevant target is cued on a trial-by-trial basis using filtered high- or low-pass (HP or LP) noise.

B. Schematic of cue-to-target mapping. Each trial, a mouse is presented with either LP or HP, each of which cueing the relevant target (flash vs. sweep). Target processing is asymmetric; vision requires simple detection while audition requires discrimination.

C. Cartoon of the hypothesized pathway for PFC regulation of thalamic visual transmission through visTRN. Solid lines indicate direct pathways, while the dotted line indicates an unknown functional connection.

D. (Left) Schematic of multi-electrode drive targeting visTRN along with optogenetic suppression of PFC. (Right) Normalized firing rate changes for visTRN neurons with PFC suppression during the delay period of the task for the two different trial types (*attend to vision*, red, vs *attend to audition*, blue). PFC suppression preferentially impacts *attend to audition* trials (N= 4 mice, n = 220 neurons; ***p<0.001, pairwise signrank test; ++ p<0.005, vs baseline). Data replotted (Wimmer et al., 2015).

E. Confocal images showing retrograde labelling of visTRN (LGN injection) and audTRN (MGB injection) using different colors of RVΔ*G*.

F. Putative visTRN neurons, which are known to be localized to the dorsal sector of the TRN, receive projections from sensory cortices but not from the PFC.

G. Strategy for identifying inputs into visTRN using monosynaptic rabies. We first expressed the TVA receptor in visTRN neurons via injection of the retrograde lentivirus (FuGB2LV) harboring Cre-dependent TVA-mCherry into LGN of Vgat Cre mice. A second injection of AAV containing the Cre-dependent rabies glycoprotein broadly delivered this gene to TRN neurons. The overlap of these injection was specific to visTRN. Lastly, mutated G-deleted rabies virus (RVΔ*G*) expressing mCherry was injected in LGN to enable the assembly of rabies viruses capable of monosynaptic labeling of inputs.

H. Top: Cartoon of a brain section indicating the location of the caudal globus pallidus (GP), where the largest source of monosynaptic inputs to visTRN were identified. Bottom: example confocal image of these caudal GP neurons (red).

I. Neurons in visual striatum (visStriatum) send GABAergic projections to visTRN-projecting GP. Top: Injection of ChR2-EYFP (green) into visStriatum shows putative terminals in GP. Bottom: combining this injection with synaptophysin-mCherry (red) definitively identifies terminals in that region of GP (yellow dots). (Inset) Higher resolution magnification.

J. Top: Cartoon of experimental setup for projection-specific PFC suppression. AAV-hSyn-eArch3.0-EYFP was injected into PFC and optical fibers implanted in multiple putative targets. Bottom: Robust labelling of both PFC cell bodies (left) and their terminals in two projection targets: visStriatum (center) and MD (right).

K. Optogenetic suppression of PFC terminals in visStriatum led to a greater increase in errors for *attend to audition* trials (blue) compared to *attend to vision* trials (red) (N = 7 mice, n = 24 sessions; ***p<0.001, pairwise sign-rank test; +++ p<0.001, vs baseline).

L. visTRN cell body suppression led to qualitatively similar results (N = 6 mice, n = 24 sessions; ***p<0.001, pairwise sign-rank test; +++ p<0.001, vs baseline).

M. Optogenetic silencing of PFC cell bodies or terminals in MD diminished performance across both trial types equivalently (N = 6 mice, n = 24 sessions; pairwise sign-rank test; +++ p<0.001, vs baseline)

Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

See also Fig. S1.

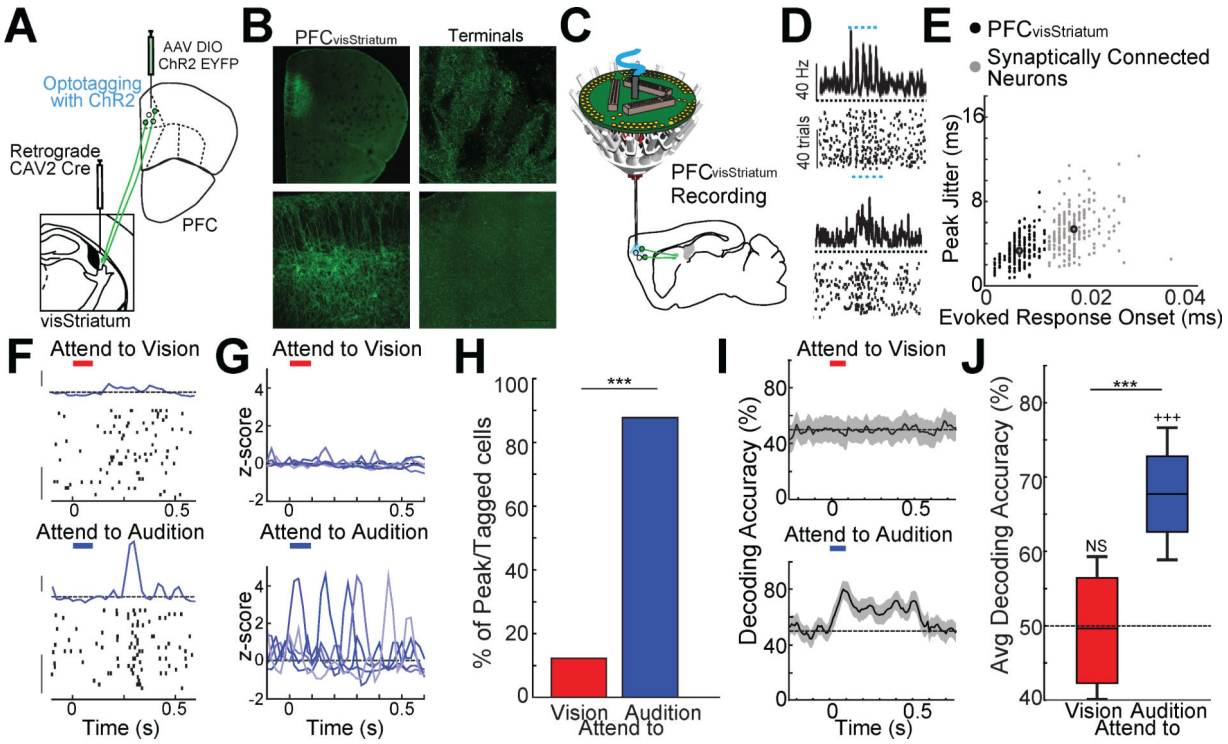

Figure 2: PFC projection neurons to visStriatum encode specific rules

A. Strategy of intersectional CAV2-Cre based retrograde labeling of PFC neurons projecting to visStriatum ($PFC_{visStriatum}$) to optogenetically tag this population. The CAV2Cre virus was injected into visStriatum while an AAV virus harboring Cre-dependent ChR2-EYFP was injected into PFC.

B. Low magnification (top, left) or high magnification (bottom, left) confocal images of $PFC_{visStriatum}$ neurons. Labeled terminals were also found in visStriatum (top, right) but not in MD (bottom, right).

C. Schematic of multi-electrode drive targeting the PFC along with optogenetic tagging of $PFC_{visStriatum}$ neurons.

D. Example spike rasters and peri-stimulus time histograms (PSTH) for a ChR2 tagged $PFC_{visStriatum}$ neuron (top) and a synaptically connected neuron (bottom) showing more temporally precise and consistent laser-evoked responses in the directly labeled neuron.

E. K-means clustering identification of directly-tagged (black) and synaptically-connected population (grey) using response onset and spike jitter. Larger black outlined circles indicate cluster centroids (N = 4 mice, n= 216 tagged and 227 post-synaptic neurons).

F. Response of an optogenetically-tagged $PFC_{visStriatum}$ neuron recording in the cross-modal 2AFC task. Only correct trials are shown, which have been separated according to the cueing condition. Zero time indicates cue presentation (100msec duration, LP–Red bar, HP–Blue bar). Note that the neuron shows a brief increase in spike rate at ~300msec following cue presentation selectively in the *attend to audition* trials. This can be seen as a 'peak' in the PSTH of that condition (PSTH y-axis scale bar: 1 Zscore, Raster y-axis scale bar: 10 trials).

G. Example PSTHs of multiple $PFC_{visStriatum}$ neurons showing task-related peaks. The PSTH peaks tile the delay period selectively for the *attend to audition* trials.

H. Majority of $PFC_{visStriatum}$ neurons showed peaks only in *attend to audition* trial (blue) but not during *attend to vision* trials (red) (N = 2 mice per condition, n = 112 neurons; *** $p < 0.001$ pairwise binomial test).

I. Poisson naïve bayes classifier for $PFC_{visStriatum}$ neuronal spiking over short time bins (50msec) shows selective encoding of the *attend to audition* cue, consistent with the notion that this population is selectively engaged when vision needs to be suppressed. Responses to each cue type were compared against jittered spike times (±0–250msec jitter), shaded regions indicate 95% confidence intervals (N = 2 mice, n = 112 neurons).

J. Boxplots of the bootstrapped distribution of average decoding accuracy (60% of cells included per decoding run). Only decoding of *attend to audition* cues (blue) was consistently above chance (N = 2 mice, n = 112; ***$p<0.001$, pairwise sign-rank test; +++ $p<0.001$, vs baseline).

Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

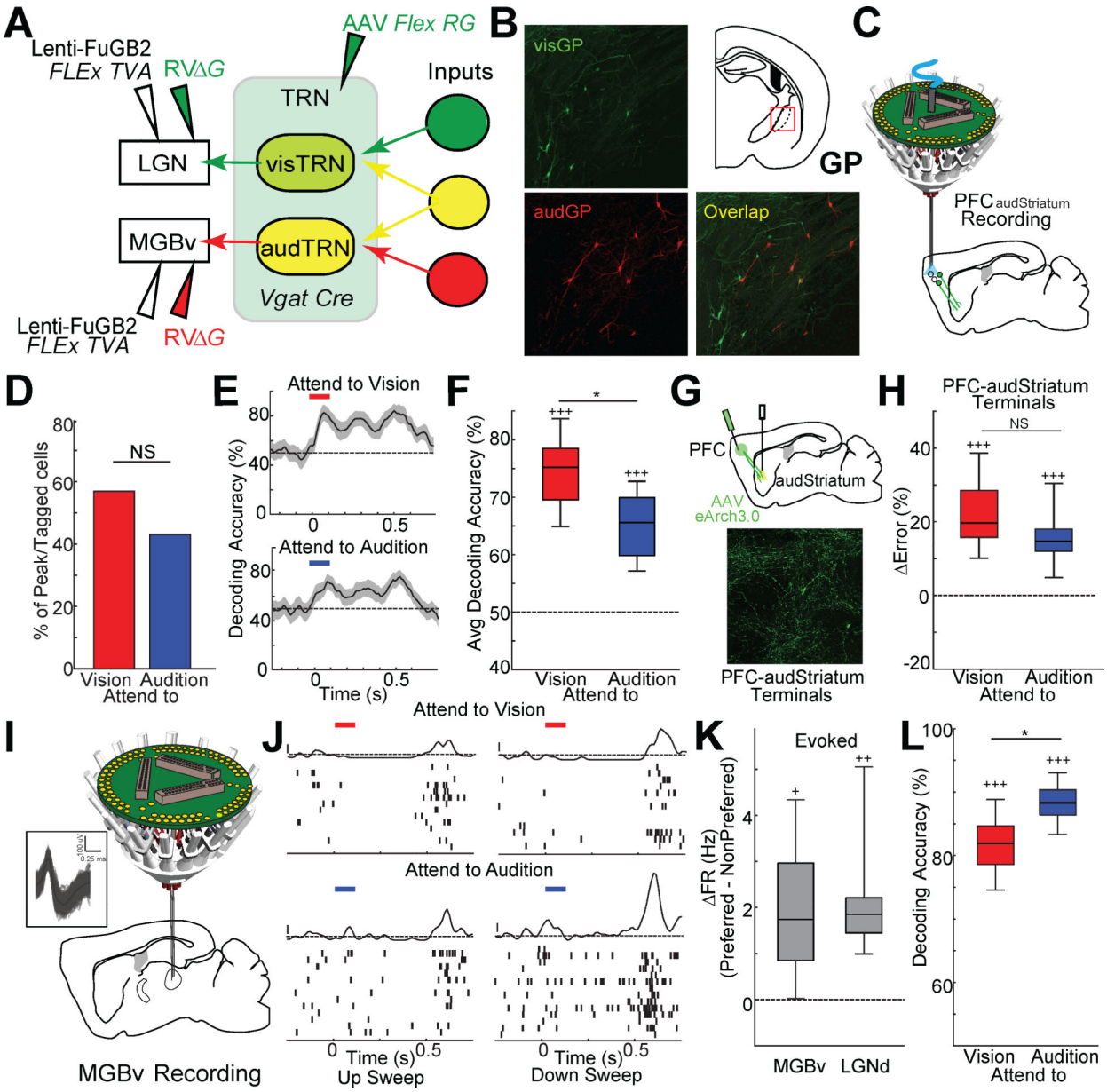

Figure 3: Multifaceted engagement of $PFC_{audStriatum}$ outputs in the cross-modal sensory selection task

A. Strategy for identifying inputs into both visTRN and audTRN using monosynaptic rabies, similar to Figure 1 G. Common inputs to visTRN and audTRN would be labeled as yellow.

B. (Left) Examples of confocal images showing EGFP labelled inputs to visTRN (visGP, top) and mCherry labelled inputs to audTRN (audGP, bottom) in caudal GP. (Right) Cartoon of a brain section indicating the location of GP (top), showing minimal overlap.

C. Schematic of multi-electrode drive targeting the PFC along with optogenetic tagging of $PFC_{audStriatum}$ neurons.

D. Percentage of retrograde optogenetically-tagged neurons ($PFC_{audStriatum}$) showing peaks for each trial type. Unlike $PFC_{visStriatum}$ neurons, $PFC_{audStriatum}$ neurons showed peaks in both trial types (N = 2 mice per condition, n = 104 neurons; pairwise binomial test).

E. Poisson naïve bayes classifier based on $PFC_{audStriatum}$ neuronal spike rates (50msec bins) reflects encoding of both cues, albeit to different degrees. Responses to each cue type were compared against jittered spike times (±0–250msec jitter), shaded regions indicate 95% confidence intervals (N = 2 mice, n = 104 neurons).

F. Boxplots of the bootstrapped distribution of average decoding accuracy (60% cells included per decoding run). The *attend to vision* cue is more strongly encoded across this population than the *attend to audition* cue, although both are consistently encoded above chance (N = 2 mice, n = 104 neurons; *p<0.05, pairwise sign-rank test; +++ p<0.001, vs baseline).

G. Cartoon of experimental preparation used to suppress PFC projections to audStriatum (top). Confocal image showing robust labelling of PFC terminals in audStriatum (bottom).

H. Suppression of PFC terminals in the audStriatum significantly disrupted performance on both trial types (N = 5 mice, n = 20 sessions; +++ p<0.001, vs baseline)

I. Schematic of multi-electrode drive targeting the MGBv. (Inset) Example waveforms of RS neurons from MGBv recording.

J. Response of a MGBv neuron recording in the task. Auditory stimuli (up or down sweep; 100msec duration) were presented at 0.5s. Note that the neuron shows stronger and more selective response to down sweep sound in the *attend to audition* cueing condition (bottom) compared with *attend to vision* condition (top). (PSTH y-axis scale bar: 1 Zscore).

K. Change in firing rate of evoked responses to target stimuli between trial types (Preferred-NonPreferred). MGBv neurons showed a reduction in evoked responses to auditory stimuli in *attend to vision* trials that was similar in magnitude to the one in LGNd responses for the opposite trial type.

L. Poisson naïve bayes classifier based on MGBv neuronal spiking (50msec bins) evoked by up- or down-sweep auditory target stimuli. Target sounds were more easily classified during *attend to audition* trials compared to *attend to vision* trials. Boxplots of the bootstrapped distribution of average classification accuracy (60% cells included per decoding run). (N = 2 mice, n = 693 neurons; *p<0.001, pairwise sign-rank test;).

Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

See also Fig. S2.

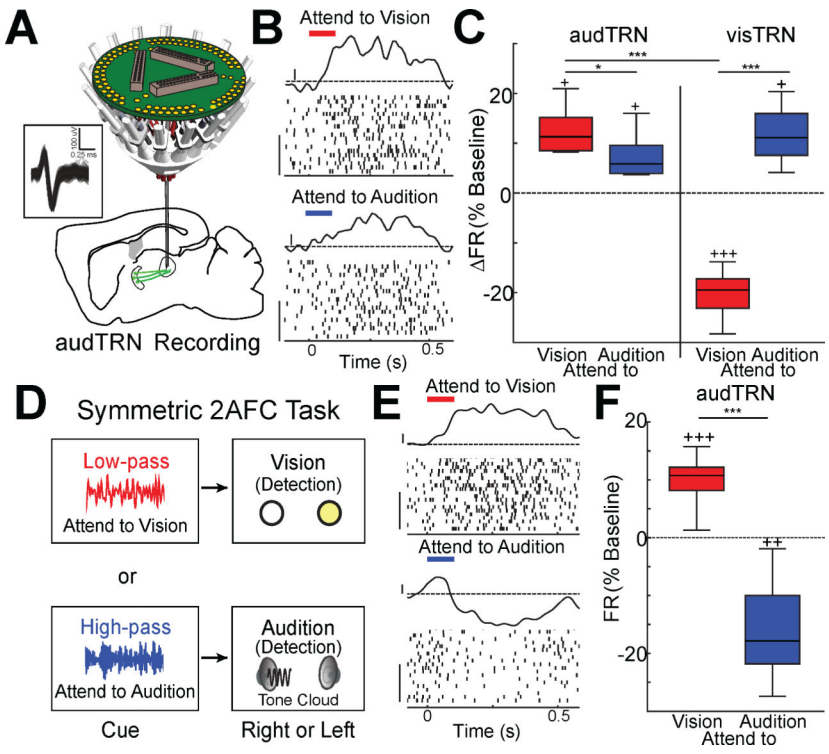

Figure 4: Increased audTRN neuron activation is associated with sound discrimination and suppression auditory distractors

A. Schematic of multi-electrode drive targeting the MGBv for terminal recordings of audTRN neurons. (Inset) Example waveforms of a FS neuron terminal recorded in MGBv.

B. Responses of an audTRN neuron recording in the crossmodal 2AFC task. While the neuron show delay responses for both trial types, the stronger response is observed in *attend to vision* trials. (PSTH y-axis scale bar: 1 Zscore; Raster y-axis scale bar: 10 trials)

C. Boxplots of normalized changes in firing rates for visTRN or audTRN neurons during the delay period of the crossmodal 2AFC task. (N > 2 mice per recording type, n > 200 neurons per condition; *p<0.05, *** p<0.001, pairwise sign-rank test; + p<0.05, +++ p<0.001, vs baseline).

D. Schematic of the symmetric 2AFC task; the auditory target is a tone cloud presented from either a right or left speaker.

E. Responses of an audTRN neuron recording in the symmetric task. The neuron shows increased response during *attend to vision* but decreased responses for *attend to audition* trials (PSTH y-axis scale bar: 1 Zscore; Raster y-axis scale bar: 10 trials).

F. Normalized changes of audTRN neuron firing rates during the delay period of the symmetric task for the two different trial types. (N = 2 mice, n = 139 FS neurons; *** p<0.001, pairwise rank-sum test; ++ p<0.01, +++ p<0.001 vs baseline).

Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

See also Fig. S3.

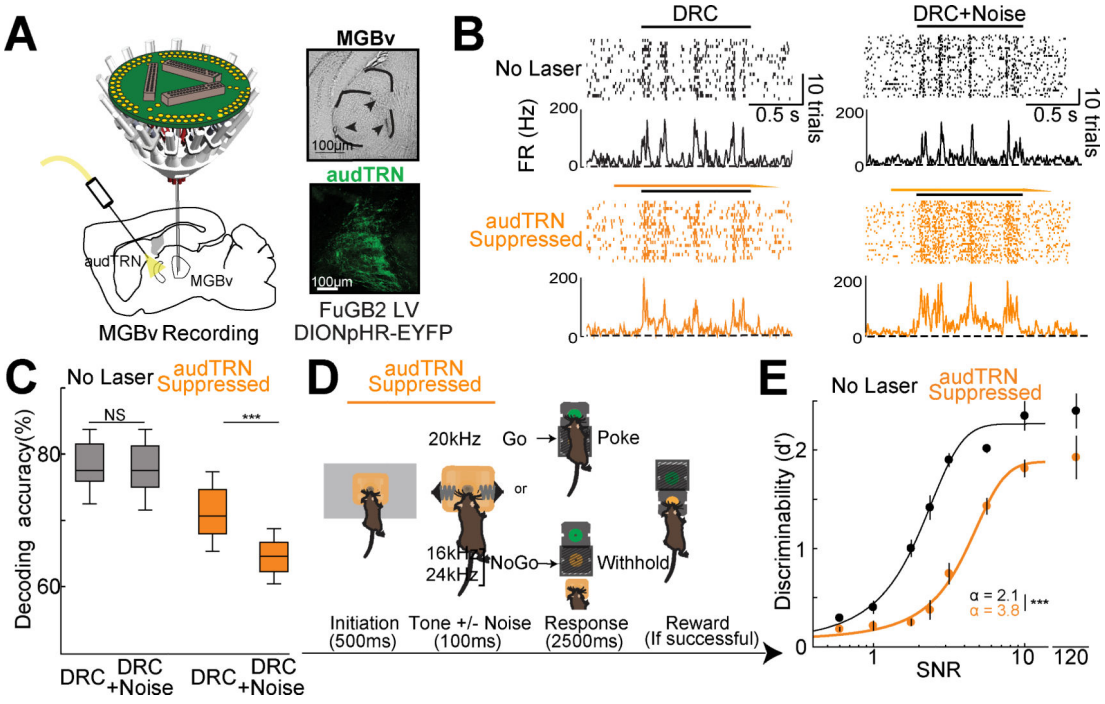

Figure 5: Activity of the audTRN supports discrimination by enforcing sparse sound representations in the MGBv

A. (Left) Schematic of multielectrode targeting of MGBv with optogenetic audTRN suppression. (Right) Example brain sections showing electrolytic MGBv electrode tip lesions (top). Confocal images of audTRN showing eNpHR3.0-eYFP expression (bottom).

B. (Left) Responses of an MGBv neuron to repeated presentation of a single dynamic random chord (black bar). On some DRC presentations audTRN was suppressed using eNpHR3.0 activation with ramped offset (yellow bar). (Right) Similar to the left panels, but with the addition of broad-band noise to the DRC (equivalent to stimulus SNR level of 3.2).

C. Poisson naïve bayes based population decoding (average) of 20kHz and 24kHz tone representations based on MGBv responses to DRCs. audTRN suppression reduced population encoding, an effect that was worsened by the addition of noise (N = 3 mice, n = 430 Neurons, *** p<0.001 pairwise sign-rank test).

D. Schematic of an auditory discrimination Go/NoGo task in which auditory stimulus signal-to-noise ratio (SNR) is parametrically controlled. After initiation, one of three different tones with varying SNRs. Following the Go tone, the mouse is required to nose-poke in the response port, which would subsequently open a reward port (hit). Following either of the NoGo tones, the mouse is required to withhold until the reward port naturally opens (correct rejection). NpHR activation (yellow bar) was included throughout initiation and tone representation on a pseudorandom subset of trials.

E. Psychometric function of control trials (no laser) in black shows the expected behavioral changes as a function of stimulus SNR. audTRN suppression results in diminished behavioral performance across stimulus SNR conditions (No noise condition shown off-scale for clarity, N = 5 mice n = 6 sessions/mouse/condition). In addition to showing a point-wise reduction in performance, audTRN suppression increased discrimination threshold (values as shown; *** p<0.001, pairwise rank-sum test) suggesting a greater requirement for

audTRN engagement when SNR is low. Discrimination threshold (α) values for baseline and audTRN suppression were also significantly different (inset).
Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)
See also Fig. S4.

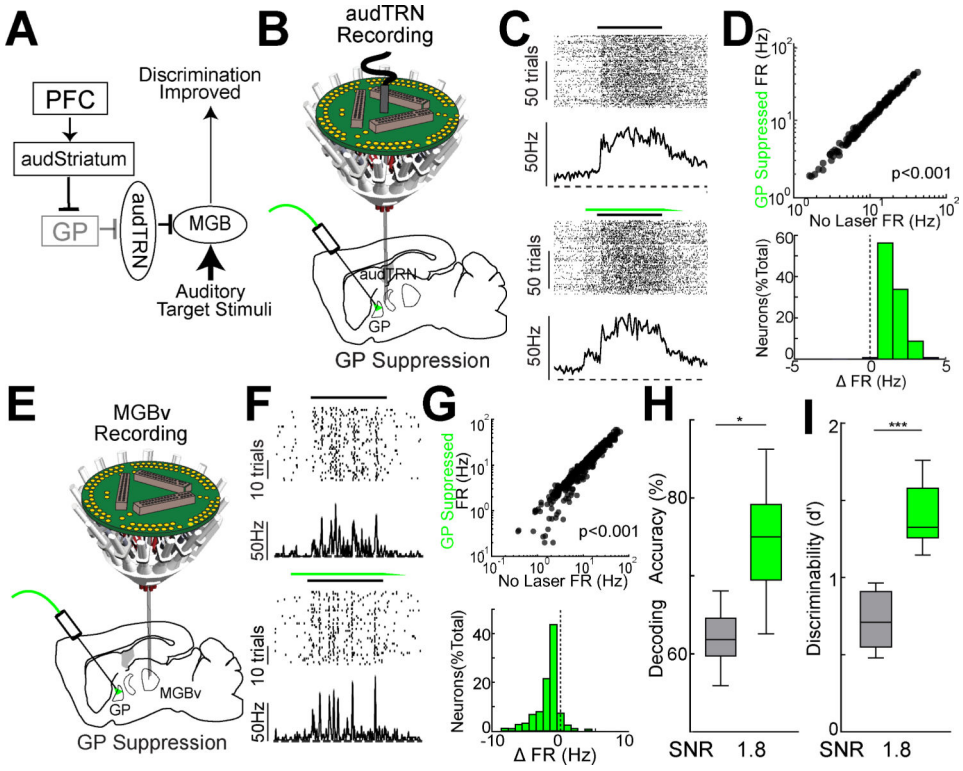

Figure 6: Activity of the audTRN supports discrimination by enforcing sparse sound representations in the MGBv

A. Cartoon summary of the PFC-basal ganglia-sensory thalamus circuit and its hypothesized engagement in noisy discrimination. This diagram suggests that PFC suppression of the GP would disinhibit audTRN neurons to improve auditory discrimination, which is tested in subsequent panels.

B. Schematic of multi-electrode drive targeting the audTRN along with optogenetic suppression of GP.

C. (Left) Responses of an audTRN neuron to repeated presentation of a single dynamic random chord (black bar). On some DRC presentations, GP was suppressed using eNpHR3.0 activation with ramping offset (green bar).

D. (Top) Comparison of sound-evoked responses of audTRN neurons with or without GP suppression. Evoked responses were significantly increased by suppressing GP (N = 3 mice, 151 neurons, rank-sum test). (Bottom) Distribution showing the changes in sound evoked responses of audTRN neurons produced by GP suppression.

E. Schematic of multi-electrode drive targeting the MGBv along with optogenetic suppression of GP.

F. (Left) Responses of a MGBv neuron to repeated presentation of a single dynamic random chord (black bar). On some DRC presentations, GP was suppressed via eNpHR3.0 activation. Although the spike rates of MGBv neurons decreases with GP suppression, the temporal precision of sound responses are increased (bottom).

G. (Top) Comparison of sound-evoked responses of MGBv neurons with or without GP suppression. Evoked responses were significantly decreased by suppressing GP (N = 3 mice,

531 neurons, rank-sum test). (Bottom) Distribution showing the changes in sound evoked responses of MGBv neurons produced by GP suppression.

H. Poisson naïve bayes based population decoding (average) based on MGBv responses to pure tones (20 kHz vs 24 kHz) with addition of high (SNR 1.8) levels of broadband noise (see methods). GP suppression (green) significantly improved encoding (N = 3 mice, 531 neurons, sign-rank test), an observation consistent with the notion that top-down control circuit improved sound discriminability in MGBv through audTRN.

I. Effect of GP suppression on behavioral performance of auditory discrimination Go/NoGo task (SNR 1.8). Interleaved, trial-by-trial optogenetic GP suppression (green) during the interval surrounding auditory tone delivery results in improved behavioral performance (N = 4 mice, n > 4 sessions/mouse, *** $p < 0.001$ rank-sum test).

Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

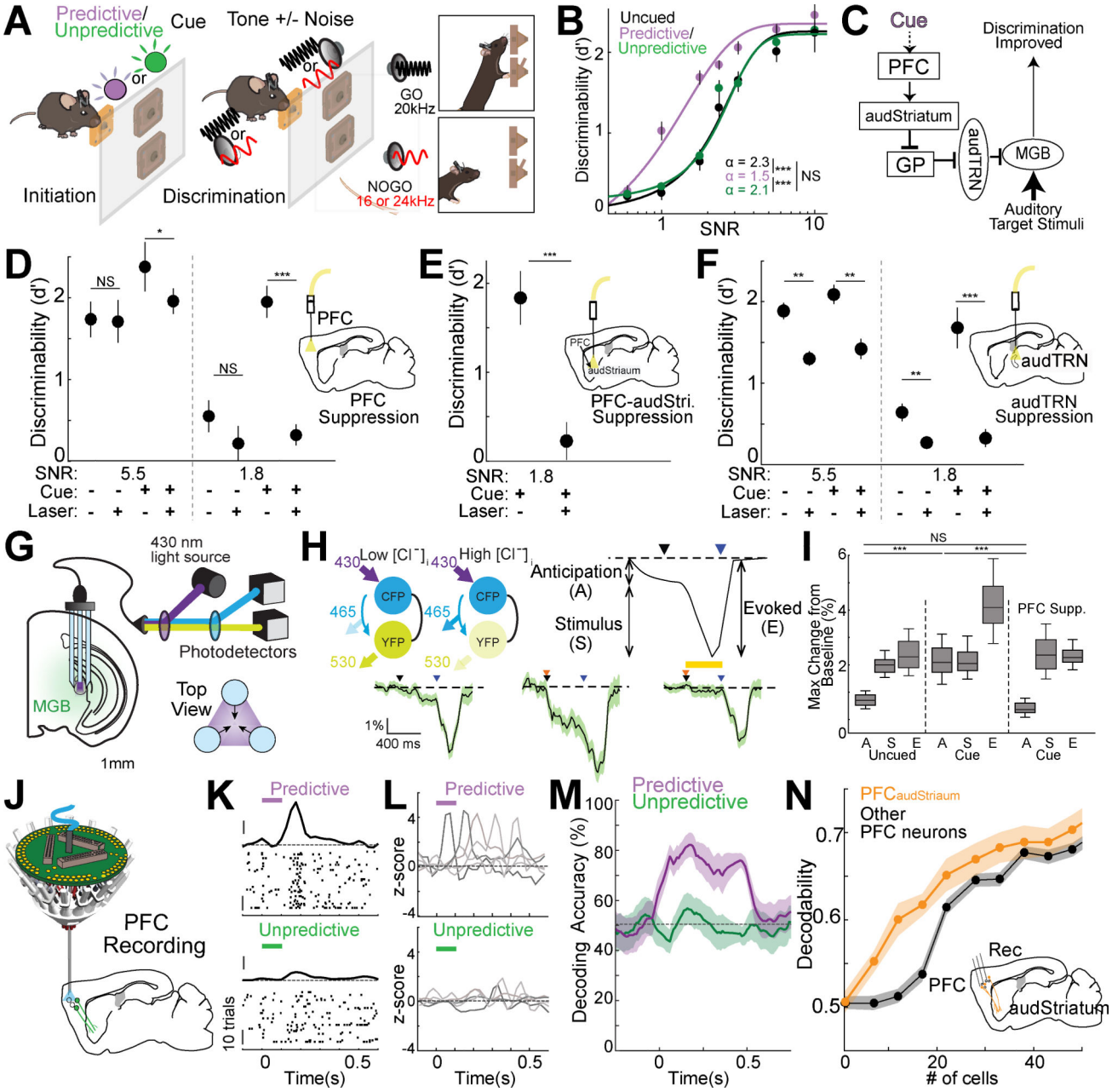

Figure 7: PFC improves performance on the cued auditory noisy discrimination task through the same PFC-basal ganglia-sensory thalamus pathway used for sensory selection

A. Schematic of a cued noisy auditory discrimination (Go/NoGo) task. On an interleaved subset of trials, mice were cued with a 100 msec pulse of either U.V. or green light followed by a 400 msec delay period before target stimulus presentation. During training each mouse learned that one of two light colors is always followed by a noise trial (predictive cue) while the other color (unpredictive cue) is not.

B. Task performance on the cued noisy auditory discrimination task across SNR levels for trials in which no cue (uncued, black), a predictive cue (predictive, violet) or an unpredictive cue (unpredictive, green) was delivered. Performance was significantly enhanced by the

predictive cue for lower SNR levels (N = 6 mice, n = 6 sessions/condition/mouse) suggesting mice were able to use the predictive cue to more effectively filter noise. Discrimination threshold was significantly reduced for predictive cue trials compared with uncued (value as shown, see methods; *** p<0.001 pairwise rank-sum test).

C. Cartoon summary of the PFC-basal ganglia-sensory thalamus circuit and its hypothesized engagement in cued noisy discrimination showing the proposed effect of PFC engagement by predictive cue through the identified control pathway.

D. Effect of PFC suppression (Laser) on behavioral performance for either predictive cue (Cue +) or uncued (Cue -) trials under two SNR conditions (5.5 and 1.8). PFC suppression (inset cartoon), selectively diminished the cue-dependent effect on performance for both SNR levels. (N = 6 mice, n = 10 sessions/condition/mouse; * p<0.05, *** p<0.001 pairwise rank-sum test).

E. The noise-predicting cue effect was diminished when PFC projections to audStriatum were optogenetically silenced. (Inset, N = 4 mice, 18 sessions; *** p < 0.001, pairwise rank-sum test)

F. Experiment similar to D, but with audTRN suppression (inset). Suppressing audTRN reduced performance for both cue and uncued trials consistent with the notion that this circuit is necessary for discrimination with and without top-down control (N = 5 mice, n = 10 sessions/condition/mouse; ** p<0.01, *** p<0.001 pairwise rank-sum test).

G. Schematic showing modified chloride photometry setup employed to measure thalamic inhibition in the MGBv using the fluorescent resonance energy transfer (FRET) indicator SuperClomeleon. A custom three terminal fiber was used to improve signal acquisition (see methods).

H. (Top, Left) Schematic showing the chloride related fluorescence signal from the SuperClomeleon indicator. This FRET indicator contains cyan fluorescent protein (CFP) as FRET donor and yellow fluorescent protein (YFP) as FRET acceptor. Under condition of elevated $[Cl^-]_i$, YFP is quenched and FRET is reduced. (Top, Right) Diagram showing distinct response components of an ideal single trial response in the cued noisy discrimination behavior. Average traces of FRET responses in MGB showing distinct response profiles following initiation (black arrow) and cue presentation (orange arrow) as well as when sound stimuli was presented (blue arrow). In uncued trials (bottom, left) a small increase in inhibitory signal was observed for anticipation. This signal was increased by cue (bottom center) and eliminated by PFC suppression (bottom, right; yellow bar indicates suppression period). (SNR = 1.8, N = 6 mice, > 24 session per condition, shaded area indicates the 95% confidence interval)

I. Quantification of the effect of cue and PFC suppression on behavior-related inhibitory chloride signal response components (shown in M: A-Anticipation, S-Stimulus, E-Evoked). The cue related increase during anticipation and overall evoked response increase on cued trials was eliminated by PFC suppression while sound stimulus responses were unaffected suggesting that only the anticipatory increase in inhibition was PFC dependent (SNR = 1.8; N = 6 mice, > 24 sessions/condition; *** p<0.001 pairwise rank-sum test).

J. Schematic of multi-electrode drive targeting the PFC along with optogenetic tagging of $PFC_{audStriatum}$ neurons.

K. Responses of a putative excitatory PFC neuron recorded during performance of the cued noisy discrimination task. Only correct trials are shown, which have been separated

according to the cueing condition. Zero time indicates cue presentation (100msec duration, the Predictive cue – purple, the Unpredictive cue – green bar). Note that the neuron shows a brief increase in spike rate at ~200msec following the noise-predictive cue presentation (top) but not following the unpredictive cue (bottom). (PSTH y-axis scale bar: 1 Zscore, Raster y-axis scale bar: 10 trials).

L. Example PSTHs from 5 simultaneously recorded PFC neurons in a mouse performing the cued noisy auditory discrimination task. These example neurons showed predictive cue selective response peaks tiling the delay period.

M. Poisson naïve bayes decoding of predictive or unpredictive cues against uncued trials showing that the PFC selectively encodes the predictive cue. Shaded regions indicate 95% confidence intervals.

N. Mean decoding accuracy for the predictive cue obtained by drawing increasing numbers of neurons either from all non-tagged PFC neurons recorded or for tagged $PFC_{audStriatum}$ populations (see methods). Inset shows cartoon summarizing the approach used to record retrogradely-labelled, optogenetically tagged $PFC_{audStriatum}$ neurons. Classification accuracy scaled more rapidly for $PFC_{audStriatum}$ compared with non-tagged neurons ($p = 8.9\times10^{-5}$ KS test, N = 4 mice, n = 112 tagged, 751 nontagged neurons) suggesting this population was more likely to encode the cue. Shaded regions indicate standard error of the mean. Boxplots: median (line), quartiles (box), 95% confidence interval (whiskers)

See also Fig. S5