

Published in final edited form as:

Neuron. 2013 July 10; 79(1): . doi:10.1016/j.neuron.2013.04.039.

The thalamo-striatal pathway and cholinergic control of goal-directed action: Interlacing new and existing learning in the striatum

Laura A. Bradfield, Jesus Bertran-Gonzalez, Billy Chieng, and Bernard W. Balleine
Brain & Mind Research Institute, University of Sydney, Camperdown, NSW 2050, Australia

Summary

The capacity for goal-directed action depends on encoding specific action-outcome associations, a learning process mediated by the posterior dorsomedial striatum (pDMS). In a changing environment plasticity has to remain flexible requiring interference between new and existing learning to be minimized, yet it is not known how new and existing learning are interlaced in this way. Here we investigated the role of the thalamo-striatal pathway linking the parafascicular thalamus (Pf) with cholinergic interneurons (CINs) in the pDMS in this process. Removing the excitatory input from Pf to the CINs was found to reduce the firing rate and intrinsic activity of these neurons and produced an enduring deficit in goal-directed learning after changes in the action-outcome contingency. Disconnection of the Pf – pDMS pathway produced similar behavioral effects. These data suggest that CINs reduce interference between new and existing learning, consistent with claims that the thalamo-striatal pathway exerts state control over learning-related plasticity.

Introduction

For goal-directed actions to remain adaptive in a changing environment, animals have to exploit successful actions whilst continuing to explore new strategies to capitalize on the shifting environmental contingencies. Existing, well-learned solutions can, however, often proactively interfere with new learning (Dempster and Brainerd, 1995; Underwood, 1957), raising the issue of how new behavioral strategies resist interference during encoding (Rescorla, 1996). In brain areas such as the hippocampus and frontal cortex, it has been suggested that the flexibility that is required accurately to encode, for example, new routes for navigation, novel categories or paired associates, depends critically on the modulation of plasticity by the cholinergic innervation of these structures (De Rosa and Hasselmo, 2000; Hasselmo and Bower, 1993; Hasselmo and Sarter, 2011; Yu and Dayan, 2002). Thus, although acetylcholine and cholinergic agonists suppress transmission at intrinsic fibers linking pyramidal cells, they have little effect on the synaptic transmission at afferent fibers (Hasselmo et al., 1992; Linster et al., 1999) suggesting that acetylcholine plays a role in cortical neurotransmission through modulation of inhibitory plasticity in recurrent networks (Bonsi et al., 2008; Vogels et al., 2011).

© 2013 Elsevier Inc. All rights reserved.

Correspondence: Bernard Balleine PhD, Brain & Mind Research Institute, University of Sydney, bernard.balleine@sydney.edu.au.

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

Various models of acetylcholine function have proposed, therefore, that cholinergic activity reduces interference in associative plasticity by creating a cellular tag for synaptic events that occur in conjunction with acetylcholine release (Froemke et al., 2007; Hasselmo and Bower, 1993). Consistent with these views, changes in cholinergic activity do not affect initial learning or retrieval and often only affect new learning induced in the presence of that change (De Rosa and Hasselmo, 2000; Hasselmo and Bower, 1993; Newman et al., 2012; Ragozzino et al., 2009); as such changes in synaptic plasticity appear to depend on cholinergic tone and, in the absence of acetylcholine, new learning is likely to be subject to interference from existing learning, perhaps by increasing contextual uncertainty (Yu and Dayan, 2002).

With regard to goal-directed learning, it is now well documented that encoding the action-outcome associations necessary for goal-directed action depends on the posterior dorsomedial striatum (pDMS) (Shiflett et al., 2010; Yin et al., 2005a; Yin et al., 2005b). As revealed by various post training tests - most notably tests of sensitivity to outcome devaluation - lesions, reversible inactivation or pharmacological blockade of plasticity in the pDMS abolish the encoding of the action-outcome associations that constitute goal-directed learning and that mediate choice between distinct courses of action (Balleine and O'Doherty, 2010; Balleine et al., 2009). Previous research has found evidence from place learning that changes in *stimulus-outcome* associations can cause acetylcholine release in the anterior DMS (Brown et al., 2010). It is not known, however: what processes mediate new learning after changes in the *action-outcome* contingency; the role that striatal cholinergic activity plays in new goal-directed learning; nor, as goal-directed learning depends on the posterior DMS and not the anterior DMS (Yin et al., 2005b), whether new learning also depends specifically on the pDMS. Given the role of acetylcholine in other brain regions in reducing interference of this kind, however, one possibility is that, rather than influencing initial action-outcome encoding, cholinergic activity in the pDMS functions to integrate new with existing learning when instrumental contingencies change.

In the face of cholinergic depletion, therefore, this account predicts that, although initial learning should be intact, new learning induced by changes in the action-outcome contingency will result in interference between the initial and this new learning and so a loss of goal-directed control. Here we sought to assess this hypothesis by altering cholinergic activity in the pDMS both chronically, by disconnection of the thalamo-striatal pathway, and acutely, using local pharmacological manipulations, and examining the effects of these treatments on: (i) initial acquisition of specific action-outcome associations; (ii) sensitivity to the selective degradation of those action-outcome contingencies and (iii) the rats' ability to encode new action-outcome associations.

Results

Parafascicular nucleus of the thalamus regulates cholinergic activity in the posterior dorsomedial striatum

Cholinergic interneurons (CINs) provide the main source of acetylcholine in the striatum (Bolam et al., 1984; Contant et al., 1996). Although they constitute only ~3% of the neurons, they ramify extensively making cholinergic activity in the striatum amongst the highest in the brain (Sorimachi and Kataoka, 1975). Their activity can be influenced by a number of neuromodulators, most notably dopamine and acetylcholine itself (Calabresi et al., 1998; Threlfell and Cragg, 2011), although their activity is mostly determined by excitatory glutamatergic afferents arising in midline thalamic nuclei (Consolo et al., 1996a; Consolo et al., 1996b; Lapper and Bolam, 1992). Prior tracing studies suggest that the region of midline thalamus containing the parafascicular thalamic nucleus (Pf) projects massively and extensively throughout all portions of the striatum (Deschenes et al., 1996; Groenewegen

and Berendse, 1994). The specificity, however, of Pf afferents to the pDMS – the region we have previously shown to be critical for the acquisition of goal-directed learning in this species (Yin et al., 2005b) – has not been explicitly assessed. As a consequence, at the outset of these experiments we determined the specificity of this pathway by infusing the retrograde tracer fluorogold into the pDMS and examining the extent of retrograde labeling throughout the thalamus. We found only one area of labeling in the midline thalamic nuclei localized to the dorsal portion of the ipsilateral parafascicular thalamic nucleus (Pf – Figure 1A). Importantly, no labeling was observed in the thalamus in the hemisphere contralateral to the infusion (Figures 1B & 1C).

Next we examined the effect of an NMDA-induced unilateral cell body lesion of the Pf on the function of CINs in the pDMS ipsilateral and the pDMS contralateral to the lesion. For this experiment 5–6 week old rats (n=18) were first given a unilateral lesion of the Pf (Figure 2A). After 1 week we took 300 μm coronal sections through the DMS and, using a cell-attached configuration of patch-clamp electrophysiology with least perturbation of intracellular content, compared the spike frequency in CINs located in the pDMS either ipsilateral or contralateral to the Pf lesion (Figure 2B–E). As we have done previously (Bertran-Gonzalez et al., 2012), determination of CINs was based on their well-described morphological and electrophysiological characteristics (Bennett and Wilson, 1999), as well as post hoc biocytin-labeled histochemistry (Figures 2C & 2D). Importantly, we found that the frequency of action potentials was significantly reduced in CINs recorded in the hemisphere ipsilateral to the Pf lesion relative to those recorded in the contralateral hemisphere, $F(1,17) = 26.09$, $p < 0.001$, (Figure 2E, Table S1).

To confirm that the lesion-induced reduction in firing rate was specific to changes in the intrinsic activity of CINs, we used a recently described means of measuring functionally relevant changes in CIN activity based on fluctuations in phosphorylation levels of the ribosomal protein S6 (S6rp) assessed by immunofluorescence (Bertran-Gonzalez et al., 2012) (Figure 2F and 2G). We explored the state of phosphorylation of different C-terminal residues of the ribosomal protein S6 (S6rp), an integrant of the ribosomal complex modulated in striatal neurons (Bertran-Gonzalez et al., 2012; Santini et al., 2009; Valjent et al., 2011). In untreated rats, we recently described a persistent phosphorylation of the Ser^{240–244} phospho-pair of S6rp specifically in CINs of different striatal regions (Bertran-Gonzalez et al., 2012)(Figure 2F), likely reflecting the intrinsic translational tone of these neurons (Ruvinsky and Meyuh, 2006). Accordingly, in rats with unilateral PF lesions, we detected a reduction in the phospho-Ser^{240–244} signal in CINs in the pDMS ipsilateral to the lesion compared to those in the contralateral pDMS, $F(1,49) = 42.573$, $p < .001$, (Figures 2G-left panel). This effect was not observed in CINs in the adjacent dorso-lateral striatum (DLS) (Figure 2G-right panel) $F(1, 48) = 1.046$, $p = .312$. Together these results suggest that the functional activity of CINs in the pDMS is heavily regulated by the parafascicular thalamus via the thalamo-striatal projection.

Loss of parafascicular control of striatal cholinergic activity does not affect initial goal-directed learning

Having established that cytotoxic lesions of the Pf alter the function of CINs in the pDMS, we next assessed the effect of these lesions on goal-directed learning. We first examined whether Pf lesions affected initial encoding and retrieval of action-outcome associations, and then whether they affected the ability to encode changes in those action-outcome associations. Rats were first given bilateral NMDA-induced lesions of the Pf or sham lesions (cf. Figures 3B & 3J). They were then food deprived and trained to press two levers on random ratio schedules of reinforcement, one delivering food pellets and the other a sucrose solution (Figure 3A). Rats quickly learned to press the levers and increased their performance as the ratio requirement increased. Statistical analysis showed that Pf lesions

had no effect on acquisition (Figure 3C); there was no main effect of group, no group x acquisition interaction (all $F_s < 1$), and both the Sham, $F(1, 10) = 13.82$, $p = .004$, and the Pf group, $F(1, 10) = 14.34$, $p = .004$, linearly increased responding across training sessions. Pf lesions also had no effect on goal-directed behavior, as evaluated using sensory-specific satiety to devalue one or other instrumental outcome. Specifically, rats were given unrestricted access to either the pellets or sucrose for 1-hr followed by a choice extinction test in which both levers were available but no outcomes delivered. This test provides a direct assessment of the action-outcome associations encoded during training (cf. Balleine & Dickinson, 1998); if rats encoded specific lever press-outcome associations during training and integrated these with the current value of the two outcomes, they should have reduced their performance of the devalued action relative to the non-devalued action on test. We observed an outcome devaluation effect of similar magnitude in both Sham and Pf-lesioned rats (Figure 3D) suggesting that action-outcome encoding was intact in the Pf group. There was a main effect of devaluation, $F(1, 10) = 15.08$, $p = .003$, but no main effect of group, $F(1, 10) = .14$, $p = .71$, and no group x devaluation interaction, $F(1, 10) = .38$, $p > .376$.

Parafascicular control of striatal cholinergic activity is necessary to encode changes in the action-outcome contingency

Next we examined the ability of Pf rats to adjust to a change in the action-outcome contingency. First, we assessed their sensitivity to a programmed reduction in the contingency on one lever from a positive contingency (around 0.05) to a zero contingency – see Figure 3A (Balleine and Dickinson, 1998). Sham-lesioned rats adjusted to this change in contingency; as is clear from Figures 3E and 3F they reduced their performance on the degraded action both during degradation training when the outcomes were delivered (Figure 3E) and in an extinction test (Figure 3F). In contrast, Pf-lesioned rats failed to adjust to the change in contingency and performed both actions equally during training and the test. Statistical analysis revealed no main effect of group, $F(1, 10) = .17$, $p = .69$, but a main effect of degradation, $F(1, 10) = 5.01$, $p = .049$, and a group x degradation interaction, $F(1, 10) = 6.62$, $p = .028$. Simple effects revealed that, whereas the Sham group reduced performance of the degraded action, $F(1, 10) = 9.92$, $p = .01$, the Pf group did not, $F(1, 10) = .07$, $p = .801$. Similar results emerged from the extinction test; i.e. no main effect of group, $F(1, 10) = .26$, $p = .621$, but an effect of degradation, $F(1, 10) = 10.78$, $p = .012$, and a group x degradation interaction, $F(1, 10) = 8.32$, $p = .016$. Simple effects found the Sham group differed on the degraded and non-degraded levers, $F(1, 10) = 16.3$, $p = .002$, whereas the Pf group did not, $F(1, 10) = .1$, $p = .763$.

To confirm that the Pf lesion affected the rats' ability to encode the *change* in contingency, and not simply the reduction in a positive contingency, we retrained the rats on the initial contingencies and then reversed the relationship between the actions and outcomes; i.e., the action that delivered sucrose now delivered pellets whereas the action that delivered pellets now delivered sucrose (Figure 3A). Both the Sham and the Pf groups performed similarly during the training phase on these new action-outcome contingencies (Figure 3G) and, statistically, although there was a main effect of linear acquisition, $F(1, 10) = 4.72$, $p = .041$, there was neither an effect of group, $F(1, 10) = .23$, $p = .638$, nor a group x acquisition interaction, $F(1, 10) = .03$, $p = .87$. Next, we assessed whether the rats encoded the new action-outcome contingencies using two tests: (i) an outcome devaluation test, as described above, and (ii) a test of outcome-selective reinstatement (Ostlund and Balline, 2007). We used these two tests because they allowed us to compare the ability of the rats to use action-outcome information both to decrease and to increase the selection of a specific action in the choice tests conducted in extinction. The results from the devaluation test are presented in Figure 3H and from the outcome-selective reinstatement test in Figure 3I. In marked contrast to the devaluation test conducted after the initial learning (cf. Figure 3D), after

reversal the rats in the Pf-lesioned group responded similarly on the two actions and were unable to choose appropriately when one of the two outcomes was devalued (Figure 3H). In contrast the sham rats responded appropriately, reducing performance of the action most recently associated with the now devalued outcome. Statistical analysis supported these observations revealing a main effect of group, $F(1, 10) = 11.56, p = .007$, and of devaluation, $F(1, 10) = 5.98, p = .035$, and, critically, a significant group \times devaluation interaction, $F(1, 10) = 12.91, p = .005$. Whereas the Sham group showed a reliable devaluation effect, $F(1, 10) = 15.63, p = .003$, the Pf group did not, $F(1, 10) = .79, p > .394$. These results imply that the PF rats were unable to perform in a manner consistent with either prior or new associations.

After the devaluation test the rats were retrained on the reversed contingencies and then both lever press actions were extinguished over 15 min for the outcome-selective reinstatement test: rats received four reinstatement trials before each of which a single outcome was delivered and responding during the next 2 min (i.e. the post-outcome period) recorded (Figure 3A). Performance in the final 2 min of extinction and prior to each of the reinstatement trials served as the pre-test period. As is clear from Figure 3-I, the Sham group reinstated performance of the action most recently associated with the delivered outcome (i.e. Post-reinst) relative to the other action (Post-other). In contrast, both actions were reinstated in the Pf-lesioned group in an undifferentiated manner. Accordingly, although there was no effect of group, $F(1, 10) = .8, p = .392$, responding on the reinstated action relative to the other action differed between groups; in the sham group the increase in responding was specific to the reinstated lever, $F(1, 10) = 11.68, p = .007$, whereas in the Pf group it was not and was similarly distributed between the two levers, $F(1, 10) = .99, p = .343$.

Together, the contingency degradation, devaluation, and outcome-specific reinstatement tests confirm that Pf-lesioned rats were unable to use action-outcome information to guide instrumental performance after the initial contingencies were altered.

Disconnection of the thalamo-striatal pathway confirms parafascicular involvement in the cholinergic control of action-outcome learning in the striatum

Although our initial results imply that the effects of Pf lesions on action-outcome encoding and retrieval were secondary to their effects on cholinergic activity in striatum, we sought to confirm this more directly in two further experiments in which we disconnected the Pf from the pDMS using (i) asymmetrical lesions; and (ii) pharmacological blockade of cholinergic activity. The logic behind disconnection experiments is straightforward (Everitt et al., 1991): if the Pf and pDMS are functionally connected, then, for example, combining a unilateral Pf lesion with a unilateral lesion of the pDMS in the contralateral hemisphere should disconnect these structures and disrupt this function. Hence, for this experiment we compared a group that received contralateral (Contra) Pf and pDMS lesions with a group that received ipsilateral (Ipsi) lesions of these structures, inducing the same degree of cell loss whilst preserving the Pf-DMS pathway in the one hemisphere. To evaluate any impairment in the Ipsi group, a Sham control was also included. Representative lesions are shown in Figures 4A and schematically in 4-I.

After recovery from surgery we conducted a direct replication of the behavioral procedures previously described the results of which are presented in Figure 4. As is clear from this figure, the effect of the disconnection was similar to that induced by bilateral Pf lesions. First note that no differences were found between the two control (Sham and Ipsi) groups throughout this experiment so we averaged across these two groups for statistical analyses. As with bilateral Pf lesions, no effect was observed on initial acquisition of instrumental performance (Figure 4B; there was an effect of linear acquisition, $F(1, 21) = 125.32, p =$

001, but not of group, $F(1, 21) = 3.79$, $p = .065$, and no interaction, $F(1, 21) = 2.32$, $p = .143$). Further, each group was similarly sensitive to outcome devaluation after initial training and adjusted their choice performance in the extinction test towards the non-devalued action (Figure 4C). Statistical analysis found an effect of devaluation, $F(1, 21) = 54.82$, $p = .001$, but no effect of group, $F(1, 21) = .05$, $p = .82$, and no interaction between these factors, $F(1, 21) = .84$, $p = .36$.

Rats were then retrained assessed for their sensitivity to contingency degradation. As is clear from Figures 4D and 4E, although the sham and ipsilateral control groups were both sensitive to degradation, rats given the disconnection, i.e. Group Contra, were not. Specifically, Groups Sham and Ipsi selectively reduced responding on the degraded lever during both training (Figure 4D) and test (Figure 4E), whereas Group Contra responded similarly on both levers clearly maintaining, or even mildly continuing to increase, performance on both actions across the sessions of degradation. In the analysis of degradation training data (Figure 4D) we found a main effect of group, $F(1, 21) = 13.73$, $p = .004$, and of degradation, $F(1, 21) = 18.27$, $p = .001$, and, importantly, a significant group x degradation interaction, $F(1, 21) = 7.86$, $p = .011$. Simple effects analyses reveal that this interaction consisted of a degradation effect (i.e. nondegraded > degraded) in both the Sham, $F(1, 21) = 6.92$, $p = .016$, and Ipsi, $F(1, 21) = 21.1$, $p = .001$, groups but no degradation effect (i.e. nondegraded = degraded) in Group Contra, $F(1, 21) = .03$, $p = .86$. Similar analysis of the test data (Figure 4E) again found no main effect of group, $F(1, 21) = 0.01$, $p = .934$, a main effect of degradation, $F(1, 21) = 16.43$, $p = .001$, and a group x degradation interaction, $F(1, 21) = 5.60$, $p = .028$, with simple effects showing a significant effect of degradation in the Sham, $F(1, 21) = 4.41$, $p = .048$, and Ipsi, $F(1, 21) = 20.37$, $p = .001$, groups but not in the Contra group, $F(1, 21) = .17$, $p = .684$.

Group Contra were similarly impaired following reversal of the instrumental contingencies. Again, the lesions did not significantly affect performance on the levers during retraining on the reversed contingencies (Figure 4F); there was a main effect of linear acquisition, $F(1, 21) = 22.71$, $p = .001$, but no effect of group, $F(1, 21) = 1.43$, $p = .245$, and no group x acquisition interaction ($F < 1$). Nevertheless, the ability of the rats in Group Contra to retrieve the new contingencies on test was significantly impaired: Group Sham and Group Ipsi both showed a significant outcome devaluation effect (i.e. nondevalued > devalued), whereas Group Contra responded similarly on both levers (Figure 4G). Statistical analysis revealed no effect of group, $F(1, 21) = .7$, $p = .412$, but an effect of devaluation, $F(1, 21) = 4.71$, $p = .042$, and a group x devaluation interaction, $F(1, 21) = 4.40$, $p = .048$; whereas the Sham, $F(1, 21) = 5.10$, $p = .035$ and Ipsi groups, $F(1, 21) = 3.84$, showed reliable devaluation effects the Contra group did not, $F(1, 21) = .211$, $p = .651$. In the outcome selective reinstatement test (Figure 4H) Group Sham and Group Ipsi, both showed selective reinstatement but Group Contra did not. There was no effect of group, $F(1, 21) = .38$, $p = .545$, a main effect of responding in the pre vs. post periods, $F(1, 21) = 12.61$, $p = .002$, however the post-outcome reinstatement was specific to the lever associated with that outcome only in Group Sham, $F(1, 21) = 6.81$, $p = .016$, and Group Ipsi, $F(1, 21) = 6.1$, $p = .022$, but was divided equally between levers in Group Contra, $F(1, 21) = .17$, $p = .898$. The impairments observed in Group Contra here echo those previously observed as a result of bilateral PF lesions.

Disconnection of the thalamo-striatal pathway reduces CIN but not MSN activity

To confirm the effect of the Pf lesions on CIN function in the pDMS, we examined p-Ser²⁴⁰⁻²⁴⁴-S6rp intensity in ChAT immunoreactive neurons in the intact pDMS in rats drawn from the Sham, Ipsi and Contra groups perfused immediately after the reinstatement test. To assess specificity we also compared p-S6rp intensity in ChAT immunoreactive neurons in the dorsolateral striatum (DLS) in these groups. The results of these analyses are

presented in Figures 5A, 5B and 5C. As is clear from these figures, different levels of p-S6rp intensity in the pDMS were observed amongst groups: p-S6rp signal was significantly reduced in CINs from Group Contra (the disconnection group) compared to CINs from both Group Ipsi and Group Sham (the controls), based on the quantification presented in Figures 5B and 5C ($F(1,9) = 17.54, p < .001$). These differences were specific to the pDMS and, as observed previously (cf. Figure 2G), were not observed in the DLS; $F(1,9) = .32, p = .587$.

Using brain sections from the same experiment, we further examined whether the Pf lesion principally affected CINs in the pDMS, or whether the medium spiny neurons (MSNs) in this region were affected as well, based on the proportion of Pf glutamatergic inputs to MSNs and the complex regulation of MSNs by the Pf (Ellender et al., 2013). We took advantage of the phospho-Thr²⁰²-Tyr²⁰⁴-extracellular regulated kinase 1/2 (phospho-ERK1/2) detection in the striatum, a method shown to reliably reflect neuronal activation in MSN populations (Bertran-Gonzalez et al., 2008; Shiflett and Balleine, 2011a, b). Confocal analysis of phospho-ERK1/2 revealed considerable levels of activation in the pDMS of all groups (Figure 5D, top panels), and this activation principally occurred in MSNs, as shown by co-localization with DARPP-32 immunostaining (Figure 5D, bottom panels) (Matamales et al., 2009). Importantly, whereas the levels of ERK1/2 activation in the pDMS did not differ between Sham and Ipsi groups, $F(1,21) = .414, p = .527$, a significant increase of activated MSNs was observed in the group Contra, $F(1, 21) = 4.565, p = .045$ (Figure 5E). Moreover, we detected very few phospho-ERK1/2 neurons in the DLS (Figure 5F), in line with the more critical role of the pDMS relative to the DLS in the context of goal-directed action (Shiflett et al., 2010).

These data suggest that the expected decrease of Pf glutamatergic input to the pDMS had a direct effect on the activity of CINs but did not produce a similar effect on MSNs. Rather it resulted in an increase in MSN activity most likely due to the loss of the general inhibitory effect of CINs on striatal MSNs. The effect of Pf lesions on MSN activation reported here supports the recently described neuromodulatory nature of these specific projections (Ellender et al., 2013), and points to the importance of the Pf-CIN synapses in controlling striatal processes (Ding et al., 2010; Threlfell et al., 2012).

Disconnection of parafascicular thalamus and anterior DMS does not affect new goal-directed learning

In a separate group of rats, we investigated whether the impairments we observed following Pf-pDMS disconnection were specific to the posterior DMS, or whether disconnection of the Pf from anterior DMS (aDMS) would produce a similar effect. It is well known that the Pf projects to both the aDMS and pDMS (Deschenes et al., 1996), and a previous study observed an increase in acetylcholine in aDMS as rats learned new stimulus-outcome associations in a place task (Brown et al., 2010). The Pf-aDMS pathway, however, appears not to be required to learn new action-outcome contingencies; we found that rats with contralateral Pf and aDMS lesions showed intact initial learning (Figure S1) and, unlike the pDMS disconnection, also showed intact outcome devaluation (Figure 4G) and outcome-specific reinstatement (Figure S1) after the reversal of the action-outcome contingencies (Figure 4G; Figure S1). Statistical analysis showed that the lesion had no effect on reversal training ($F < 1$) and, on test, that there was an effect of devaluation (nondevalued > devalued), $F(1, 13) = 8.69, p = .011$, but no group x devaluation interaction, $F < 1$.

The results of this experiment suggest, therefore, that the thalamo-striatal pathway connecting the Pf and aDMS does not play a role in either initial learning or the acquisition of new goal-directed actions and confirm, therefore, that the findings following disconnection of the Pf-pDMS pathway on new learning are specific to that pathway. This is

consistent with the argument that the Pf alters the functional role of cholinergic interneurons specifically in the pDMS to enable the encoding of new action-outcome associations.

Reduction in cholinergic-activity in the pDMS using the M2/M4 agonist Oxotremorine-S replicates the effect of Pf – pDMS disconnection on goal-directed learning

The observed effects of Pf lesion on CIN function in pDMS suggests that the observed behavioral effects of bilateral Pf and contralateral Pf-pDMS lesions are most likely regulated by alterations in CIN function in pDMS. This experiment sought to determine this directly. We capitalized on previous findings showing that activation of the muscarinic M2 autoreceptor, highly expressed with high specificity on the membrane of CINs (Hersch et al., 1994), inhibits the function of these cells (Calabresi et al., 1998). In a variety of studies, the M2/M4 agonist, Oxotremorine-S (Oxo-S), has been shown to increase the trafficking and expression of M2 receptors on the membrane of CIN and to inhibit the function of these neurons (Bernard et al., 1998; Ragozzino et al., 2009). In this study, therefore, we coupled a unilateral Pf lesion with the infusion of Oxo-S or vehicle into the contralateral pDMS during the training of the reversed contingencies as described in previous studies.

Prior to the experiment, we first confirmed the previously reported expression of M2 receptors (M2Rs) on the membrane of CINs using immunohistochemistry and, secondly, the influence of Oxo-S on the firing of isolated CINs *in vitro*. As shown in Figure 6A, we found clear evidence for the localization of M2Rs on the membrane of ChAT positive neurons in the pDMS as previously reported. For the electrophysiological studies, we took 300 μm coronal sections through the pDMS and used cell-attached recordings to assess the effect of Oxo-S on activity of CINs identified as described previously. We confirmed that the pharmacological effects were likely due to post-synaptically expressed muscarinic receptors by synaptically isolating recorded neurons through application of a cocktail of glutamatergic and GABAergic synaptic blockers (CNQX, AP5 and picrotoxin), a treatment that does not affect CIN's intrinsic firing (Bennett and Wilson, 1999; Bertran-Gonzalez et al., 2012). As shown in Figure 6B, we found that Oxo-S produced a clear silencing of action potentials recorded from these neurons in a manner comparable to voltage-gated sodium channel blocker tetrodotoxin (TTX), and that the effect was reversed by the muscarinic antagonist scopolamine. Finally, to confirm the effect of Oxo-S on the activity of CINs we assessed the Ser²⁴⁰⁻²⁴⁴ phosphorylation signal of S6rp in CINs in viable brain slices that had been incubated with Oxo-S for 1 hour compared to the contralateral hemispheres taken from the same animal and that were incubated for 1 hour without Oxo-S (Control). Again, a clear reduction in activity of the CINs exposed to Oxo-S was observed; the phosphorylation signal of S6rp was significantly reduced in Oxo-S-incubated hemisections as compared to control hemisections (Figure 6C; $F(1, 109) = 17.27, p < .001$).

We next gave two groups of rats unilateral lesions of the Pf and implanted guide cannulae aimed at the contralateral pDMS (see Figures 6D, 6J and 6K and gave them instrumental training on the initial action-outcome contingencies as described previously. Acquisition of lever pressing was similar to that previously observed (Figure 6E) and we confirmed that both groups encoded the initial action-outcome associations using an outcome devaluation test (Figure 6F) finding a main effect of devaluation, $F(1, 12) = 59.05, p < .001$, but neither an effect of group, $F(1, 12) = 1.61, p = .229$, nor a group x devaluation interaction, $F(1, 12) = .01, p = .918$. Subsequently, we retrained the rats for four sessions on the new, reversed contingencies. Prior to each session of training on the new contingencies rats were given an infusion of either Oxo-S or vehicle into the pDMS (Figure 6D). Although there was a clear trend for Oxo-S to mildly reduce lever pressing during these sessions (Figure 6G), statistically, the groups did not differ, $F(1, 12) = 4.08, p = .066$. Furthermore, lever press rates during these sessions were robust and the linear increase in performance was similar to

vehicle infused rats suggesting that acquisition was otherwise normal. After training, we again gave outcome devaluation and outcome-selective reinstatement tests, conducted drug free. In these tests, intra-pDMS infusions of Oxo-S during training produced a clear deficit in new action-outcome encoding: rats that received these infusions pressed both levers at similar rates on test whereas rats given intra-pDMS infusions of vehicle showed a reliable outcome devaluation effect (nondevalued > devalued; Figure 6H). Statistical analysis found no main effect of group, $F(1, 12) = .25$, $p = .623$, but a main effect of devaluation, $F(1, 12) = 11.46$, $p = .005$, and a group x devaluation interaction, $F(1, 12) = 6.18$, $p = .029$. Simple effects showed that the vehicle-infused group responded more on the nondevalued than the devalued lever, $F(1, 12) = 17.23$, $p = .001$, whereas the Oxo-S infused group did not: $F(1, 12) = .41$, $p = .536$. In the outcome-selective reinstatement test, rats that received intra-pDMS infusions of vehicle showed selective reinstatement (reinstated > nonreinstated, post-outcome delivery) whereas rats given the Oxo-S during training showed non-selective reinstatement (reinstated = nonreinstated). Statistical analysis of the test performance revealed no effect of group, $F(1, 12) = 1.32$, $p = .404$, an effect of pre vs. post reinstatement, $F(1, 12) = 37.27$, $p = .001$, but this post-reinstatement increase in responding was specific to the reinstated lever only for the Vehicle group, $F(1, 12) = 5.35$, $p = .039$, and was similarly distributed on the two levers in the Oxo-S group, $F(1, 12) = 1.81$, $p = .203$. The results from the devaluation and reinstatement tests were, therefore, similar to those observed after bilateral Pf lesions or disconnection of the Pf from the DMS, suggesting that the behavioral effects were induced by changes in CIN function in the pDMS.

Discussion

The current series of experiments found that manipulating the thalamo-striatal pathway to influence cholinergic function in the posterior dorsomedial striatum did not affect the initial acquisition of goal-directed actions but strongly attenuated the rats' ability to encode new action-outcome contingencies involving those actions. Rather than affecting new learning specifically, we found that the deficit in cholinergic function had a more profound effect, inducing interference between both new and existing action-outcome encoding in the pDMS. As noted above, adapting to changes, temporary or otherwise, in existing action-outcome contingencies requires animals not just to exploit successful solutions to decision problems but also to explore alternative solutions. In order to do so, however, it is necessary that existing memories be interlaced with new learning in a manner that reduces interference between them, otherwise the new, the existing, or indeed both new and existing learning could be lost. The current experiments suggest this latter outcome is induced by a decrement in striatal cholinergic function. Thus, our results suggest that cholinergic activity, mediated by the CINs in the pDMS, serves the function of interlacing new goal-directed learning with existing plasticity to reduce interference between them.

The primary evidence for these claims comes from the pattern of behavioral effects induced by treatments affecting cholinergic function; i.e. the effects of lesioning the inputs to the CINs, and the disconnection of these inputs from their target in the pDMS, either by asymmetrical lesion- or oxotremorine infusion. These treatments induced robust interference in the encoding of action-outcome contingencies, but only *after* changes in these contingencies were made. Thus, bilateral lesions of the Pf or disconnection of the Pf from the pDMS, rendered the rats insensitive to contingency degradation, an effect that was not due simply to a loss in general activity; performance was maintained throughout degradation training and, indeed, appeared, if anything, to increase across sessions after the disconnection treatment. Nor were these effects produced by a failure to attend to the change in contingency, as might be proposed on an attentional theory of cholinergic function (Matsumoto et al., 2001). If this were true, although the new learning might have been lost, initial learning, which was demonstrably intact prior to the change in contingency, should

have been unaffected. However, when a positive contingency was maintained but the identity of the action-outcome associations was reversed, impaired cholinergic function did not simply result in the failure to encode the new learning but resulted in the inability to express either the old or the new learning, leaving the rats unable to choose based on either contingency.

Finally, this interference was produced both in tests involving outcome devaluation, which necessitate a selective reduction in the performance of an action based on the change in value of its associated outcome, and in tests assessing outcome-selective reinstatement – which generates a selective elevation in the reinstated action based on the delivery of its associated outcome during extinction. In both tests the ability of rats in the control groups to choose between the two actions was substantial and respected the most recently trained contingencies, whereas choice was indifferent in the groups for whom the thalamo-striatal pathway and attendant striatal cholinergic function were compromised. However, in neither the devaluation nor the reinstatement tests was this indifference due merely to a failure to choose; the overall rate of choice performance summed across both actions during these tests was generally similar to the controls, particularly in the reinstatement tests.

It is also important to note that the effect of the change in contingency was not due to an interaction with the pre-training treatment; post-training inactivation of the CINs using oxotremorine had a similar effect when introduced only during new learning after the initial training phase was complete. Indeed, this similarity in the effects of the lesion- and oxotremorine-induced disconnection suggests that the source of the effects of both treatments was likely similar. In addition to their expression on CINs, however, M2 receptors are also expressed on cortical terminals (Ding et al., 2010; Goldberg et al., 2012) and so, in addition to inhibiting acetylcholine release at the CINs, oxotremorine can also suppress glutamate release and ongoing motor behavior (Hersch et al., 1994). Nevertheless, although oxotremorine differed from the Pf lesion by mildly suppressing instrumental performance during training, the overall similarity in the effects of these treatments both behaviorally and on CIN function, suggests that it was the latter influence of the drug, rather than its effect on cortical terminals, that was functionally the more critical in the current study.

The thalamo-striatal pathway and instrumental conditioning

The current results suggest that the thalamo-striatal pathway contributes to new goal-directed learning through its projections specifically to the posterior, and not the anterior, DMS during instrumental conditioning. We found that this pathway largely governs CIN activity, as demonstrated by clear changes in activity in, and the pharmacological correlates of, the disconnection procedure. Nevertheless, it is important to recognize: (1) that the effects of Pf manipulation could be mediated by indirect thalamo-striatal connections and, more critically, (2) that any effects of altered CIN function can only be manifest through changes in projection MSN activity; in this case changes in the segregation of plasticity at the MSNs after new learning. Indeed, in animals perfused right after expressing goal-directed behaviors, we found evidence of enhanced neuronal responses in MSNs when the Pf projections had been interrupted. These results cannot be explained by a loss in the direct drive of canonical glutamatergic inputs onto MSNs, as Pf denervation would reduce, rather than increase, activity on these neurons. Instead, our observations support more recent views of how the Pf inputs modulate striatal function. In a recent study, Ellender and colleagues (Ellender et al., 2013) demonstrated heterogeneity in the regulation exerted by inputs from different types of thalamostriatal afferents on MSNs. In particular, they highlighted the modulatory nature of the inputs provided by specific parafascicular afferents for long-term plasticity, which contrasted with the excitatory influence of adjacent centrolateral afferents. Generally, therefore, although requiring further study, growing evidence supports the major

involvement of parafascicular-cholinergic synapses in the regulation of striatal function (Ding et al., 2010; Threlfell et al., 2012).

From this perspective, during goal-directed learning striatal CINs in the pDMS do not serve a simple attentional or arousal function as has been proposed in other task situations (Dalley et al., 2008; Robbins and Roberts, 2007), although the thalamo-striatal pathway as a whole could be described as serving a related function by regulating the ‘bottom-up’ activation of CINs within the striatal network (Ding et al., 2010; Kimura et al., 2004). Certainly the connectivity of the Pf is consistent with this kind of role, with many of its afferent inputs coming from reticular and sensory thalamic areas (Groenewegen and Berendse, 1994). This suggestion ignores, however, the substantial inputs from motor areas including motor cortex and pedunculopontine tegmentum, and motivational areas such as the amygdala central nucleus, and parabrachial nucleus (Cornwall and Phillipson, 1988). Indeed, together with a number of recent behavioral findings, these inputs to the Pf have suggested to some researchers the view that, together with other modulators of CINs in striatum, the thalamo-striatal pathway may generate an internal context, producing, broadly, a ‘context for action’ based on temporal, sensory and motivational factors (Apicella, 2007; Kimura et al., 2004). On this account, the Pf-pDMS pathway functions to provide a distinct context on which specific action-outcome associations become conditional.

This contextual control hypothesis of thalamo-striatal function is attractive not only because it is consistent with the modulatory function of acetylcholine but also because ‘contextual’ or ‘state’ cues of this kind have long been advanced as the simplest solution to the computation problems presented by the need to encode changes in contingency (French, 1991, 1999). Indeed, conditional control of this kind, although adding computational complexity, may be what allows new and existing learning to be spatially and temporally segregated (French, 1999), something that should be expected to become far more important after contingencies change. Furthermore, within the broader corticobasal ganglia network, the Pf, as part of the thalamo-striatal pathway, has long been considered a ‘loop nucleus’ because it projects to the striatum and receives inputs from the substantia nigra pars reticulata and so forms part of an internal modulatory loop running parallel to the larger corticobasal ganglia loops (Parent and Hazrati, 1995). Structurally, therefore, this pathway would appear well situated to monitor and to switch between cortical inputs to the striatum based on changes in well-predicted external contingencies (Kimura et al., 2004). Indeed, the recent suggestion that striatal CINs may form a recurrent inhibitory network anticipates context- or state-specific plasticity of this kind, with each CIN potentially modulating a distinct region of cortico-striatal plasticity under the control of the thalamo-striatal pathway (Sullivan et al., 2008).

State prediction errors and cholinergic function

At a formal level, contextual or state cues of this kind have emerged as a critical component of computational models of goal-directed action derived from model-based reinforcement learning (Daw et al., 2005). Such cues are argued to exert conditional control over actions and to produce a *state prediction error* when changes in such control occur. Model-based reinforcement learning uses experienced state-action-state transitions to build a model of the environment by generating state prediction errors produced by any discrepancy induced by a state transition based on the current estimates of state-action-state transition probabilities (Glascher et al., 2010). The notion of state prediction errors is in contrast with that of reward prediction errors derived from temporal difference models of learning (Sutton and Barto, 1998) that have been shown to reliably correlate with the phasic action of midbrain dopamine neurons (Schultz and Dickinson, 2000). However, reward prediction error is negligible particularly in the reversal experiments in the current series; the animal is expecting one outcome and receives another which creates an error signal but one that is

unrelated to rewarding prediction per se (the amount of reward earned is unchanged). This kind of signal is consistent with recent suggestions that CINs may participate in a form of prediction error signal in the DMS during reversal of previously learned contingencies (Apicella et al., 2011). Indeed, similar studies assessing prediction errors in, what are at least nominally, instrumental conditioning tasks have found that TANs (nominally CINs) preferentially encode prediction errors to situational events rather than reward (Apicella et al., 2011; Stalnaker et al., 2012). Taken together, the effect of impaired pDMS CIN function on contingency degradation and the learning of new, but not initial, action-outcome contingencies is consistent with a deficit in computing reductions in state prediction errors that lead to reductions in contingency knowledge (see Supplemental Text).

Whatever the role of CINs in conditional control, the current data suggest that the thalamo-striatal pathway and its influence on CINs is critical for encoding changes in the instrumental contingency. Although this pathway does not appear to play any direct role in encoding action-outcome associations (this paper) or in striatal LTP (Bonsi et al., 2008), it appears to be essential for ensuring the integration of learning about changes in the instrumental contingency with existing learning in a manner that preserves both forms of plasticity. This is in line with the general behavioral finding that prior instrumental learning is preserved in the face of changes in contingency (Rescorla, 1991, 1996) and provides a novel mechanism for this preservation.

Experimental Procedures

Full details of the experiment procedures are provided in the Supplemental Methods.

Subjects—For the behavioral studies, male Long-Evans rats, weighing between 300–380g at the beginning of the experiment, were used as subjects. For electrophysiology experiments male Long-Evans rats between 5 and 6 weeks old were used, weighing between 120–150 g. Rats that experienced behavioral training and testing were maintained at ~ 85% of their free-feeding body weight by restricting their food intake to between 8 and 12g of their maintenance diet per day. All procedures were approved by the University of Sydney Ethics Committee.

Behavioral procedures

Training & Devaluation

Magazine training: On days 1 and 2 all rats were placed in operant chambers for ~20 min. In each session of each experiment the house light was illuminated at the start of the session and turned off when the session was terminated. No levers were extended during magazine training. 20 pellet and 20 sucrose outcomes were delivered to the magazine on an independent random time (RT) 60 s schedule.

Lever training: The animals were next trained to lever press on random ratio schedules of reinforcement. Each lever was trained separately each day and the specific lever-outcome assignments were fully counterbalanced. The session was terminated after 20 outcomes were earned, or after 30 min. For the first 2 days lever pressing was continuously reinforced. Rats were shifted to a random ratio (RR)-5 schedule for the next 3 days (i.e. each action delivered an outcome with a probability of .2), then to an RR-10 schedule (or a probability of .1) for 3 days, then to an RR-20 schedule (or a probability of .05) for the final 3 days.

Devaluation extinction tests: After the final day of RR-20 training, rats were given free access to either the pellets (25g place in a bowl) or the sucrose solution (100mL in a drinking bottle) for 1 hr in the devaluation cage. The aim of this prefeeding procedure was to

satiates the animal specifically on the preferred outcome, thereby reducing its value relative to the non-preferred outcome (cf. Balleine & Dickinson, 1998). Rats were then placed in the operant chamber for a 10 min choice extinction test. During this test both levers were extended and lever presses recorded, but no outcomes were delivered. The next day a second devaluation test was administered with the opposite outcome. Rats were then placed back into the operant chambers for a second 10 min choice extinction test.

Contingency Degradation Procedure

Contingency degradation training: During the 6 days of contingency degradation training rats continued to receive these same action-outcome pairings on an RR-20 schedule. In addition, one of the two outcomes (either pellets or sucrose) was delivered outside of the lever press-outcome contingency; i.e. in each second that no lever pressing occurred, either sucrose or pellets were delivered with the same probability [$p(\text{outcome/no action}) = .05$] that a lever press earned that outcome. As a result, the probability of earning one of the two outcomes was the same whether the animal pressed the lever or not. The other action-outcome contingency was non-degraded because the rat was still required to press the lever to receive that outcome. For half of the animals the lever press – pellet contingency was degraded, and the lever press-sucrose contingency remained intact. The remaining animals received the opposite arrangement. Rats were given two 20 min training sessions each day, one on each lever.

Contingency degradation extinction test: Following the final day of contingency training, rats in both groups received a 10 min choice extinction test. During this test both levers were extended and lever presses recorded, but no outcomes were delivered.

Contingency reversal procedure

Contingency reversal training: Subsequent to the contingency degradation extinction test rats were trained to lever press on an RR-20 schedule with the previously- trained contingencies reversed. That is, the lever that previously earned pellets now earned sucrose, and the lever that previously earned sucrose now earned pellets. Contingency reversal training continued for 4 days.

Devaluation Extinction Tests: Devaluation extinction tests took place as described for outcome devaluation.

Reinstatement Testing: Rats were retrained on the reversed contingencies on an RR-20 schedule for 1 day. The next day an outcome-selective instrumental reinstatement test was conducted. The test session began with a 15 min period of extinction to lower the rats' rate of responding on both levers. They then received 4 reinstatement trials separated by 7 min each. Each reinstatement trial consisted of a single delivery of either the sucrose solution or the grain pellet. All rats received the same trial order: sucrose, pellet, pellet, sucrose. Responding was measured during the 2 min periods immediately before (Pre) and after (Post) each delivery.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research reported in the manuscript was supported by grants from the National Institute of Mental Health #MH56446, the National Health & Medical Research Council of Australia #633267, and a Laureate Fellowship

from the Australian Research Council, #FL0992409. The authors thank Amir Dezfouli for his comments on the manuscript.

References

- Apicella P. Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci.* 2007; 30:299–306. [PubMed: 17420057]
- Apicella P, Ravel S, Deffains M, Legallet E. The role of striatal tonically active neurons in reward prediction error signaling during instrumental task performance. *J Neurosci.* 2011; 31:1507–1515. [PubMed: 21273435]
- Balleine B, O'Doherty J. Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology.* 2010; 35:48–69. [PubMed: 19776734]
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology.* 1998; 37:407–419. [PubMed: 9704982]
- Balleine BW, Liljeholm M, Ostlund SB. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res.* 2009; 199:43–52. [PubMed: 19027797]
- Bennett BD, Wilson CJ. Spontaneous activity of neostriatal cholinergic interneurons in vitro. *J Neurosci.* 1999; 19:5586–5596. [PubMed: 10377365]
- Bernard V, Laribi O, Levey AI, Bloch B. Subcellular redistribution of m2 muscarinic acetylcholine receptors in striatal interneurons in vivo after acute cholinergic stimulation. *J Neurosci.* 1998; 18:10207–10218. [PubMed: 9822774]
- Bertran-Gonzalez J, Bosch C, Maroteaux M, Matamalas M, Herve D, Valjent E, Girault JA. Opposing patterns of signaling activation in dopamine D1 and D2 receptor-expressing striatal neurons in response to cocaine and haloperidol. *J Neurosci.* 2008; 28:5671–5685. [PubMed: 18509028]
- Bertran-Gonzalez J, Chieng BC, Laurent V, Valjent E, Balleine BW. Striatal cholinergic interneurons display activity-related phosphorylation of ribosomal protein S6. *PLoS ONE.* 2012; 7:e53195. [PubMed: 23285266]
- Bolam JP, Wainer BH, Smith AD. Characterization of cholinergic neurons in the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry, Golgi-impregnation and electron microscopy. *Neuroscience.* 1984; 12:711–718. [PubMed: 6382048]
- Bonsi P, Martella G, Cuomo D, Platania P, Sciamanna G, Bernardi G, Wess J, Pisani A. Loss of muscarinic autoreceptor function impairs long-term depression but not long-term potentiation in the striatum. *J Neurosci.* 2008; 28:6258–6263. [PubMed: 18550768]
- Brown HD, Baker PM, Ragozzino ME. The parafascicular thalamic nucleus concomitantly influences behavioral flexibility and dorsomedial striatal acetylcholine output in rats. *J Neurosci.* 2010; 30:14390–14398. [PubMed: 20980596]
- Calabresi P, Centonze D, Pisani A, Sancesario G, North RA, Bernardi G. Muscarinic IPSPs in rat striatal cholinergic interneurons. *J Physiol.* 1998; 510(Pt 2):421–427. [PubMed: 9705993]
- Consolo S, Baldi G, Giorgi S, Nannini L. The cerebral cortex and parafascicular thalamic nucleus facilitate in vivo acetylcholine release in the rat striatum through distinct glutamate receptor subtypes. *Eur J Neurosci.* 1996a; 8:2702–2710. [PubMed: 8996820]
- Consolo S, Baronio P, Guidi G, Di Chiara G. Role of the parafascicular thalamic nucleus and N-methyl-D-aspartate transmission in the D1-dependent control of in vivo acetylcholine release in rat striatum. *Neuroscience.* 1996b; 71:157–165. [PubMed: 8834399]
- Contant C, Umbriaco D, Garcia S, Watkins KC, Descarries L. Ultrastructural characterization of the acetylcholine innervation in adult rat neostriatum. *Neuroscience.* 1996; 71:937–947. [PubMed: 8684624]
- Cornwall J, Phillipson OT. Afferent projections to the parafascicular thalamic nucleus of the rat, as shown by the retrograde transport of wheat germ agglutinin. *Brain Research Bulletin.* 1988; 20:139–150. [PubMed: 2836036]
- Dalley JW, Mar AC, Economidou D, Robbins TW. Neurobehavioral mechanisms of impulsivity: fronto-striatal systems and functional neurochemistry. *Pharmacol Biochem Behav.* 2008; 90:250–260. [PubMed: 18272211]

- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci.* 2005; 8:1704–1711. [PubMed: 16286932]
- De Rosa E, Hasselmo ME. Muscarinic cholinergic neuromodulation reduces proactive interference between stored odor memories during associative learning in rats. *Behav Neurosci.* 2000; 114:32–41. [PubMed: 10718260]
- Dempster, FN.; Brainerd, CJ. *Interference and Inhibition in Cognition.* San Diego: Academic Press; 1995.
- Deschenes M, Bourassa J, Doan VD, Parent A. A single-cell study of the axonal projections arising from the posterior intralaminar thalamic nuclei in the rat. *Eur J Neurosci.* 1996; 8:329–343. [PubMed: 8714704]
- Ding JB, Guzman JN, Peterson JD, Goldberg JA, Surmeier DJ. Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron.* 2010; 67:294–307. [PubMed: 20670836]
- Ellender TJ, Harwood J, Kosillo P, Capogna M, Bolam JP. Heterogeneous properties of central lateral and parafascicular thalamic synapses in the striatum. *J Physiol.* 2013; 591:257–272. [PubMed: 23109111]
- Everitt BJ, Morris KA, O'Brien A, Robbins TW. The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. *Neuroscience.* 1991; 42:1–18. [PubMed: 1830641]
- French, RM. Using semi-distributed representations to overcome catastrophic forgetting in connectionist networks. *Proceedings of the Thirteenth Annual Cognitive Science Society Conference;* 1991. p. 173-178.
- French RM. Catastrophic forgetting in connectionist networks: Causes, consequences and solutions. *trends in cognitive. Sciences.* 1999; 3:128–135.
- Froemke RC, Merzenich MM, Schreiner CE. A synaptic memory trace for cortical receptive field plasticity. *Nature.* 2007; 450:425–429. [PubMed: 18004384]
- Glascher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron.* 2010; 66:585–595. [PubMed: 20510862]
- Goldberg JA, Ding JB, Surmeier DJ. Muscarinic modulation of striatal function and circuitry. *Handbook of experimental pharmacology.* 2012:223–241. [PubMed: 2222701]
- Groenewegen HJ, Berendse HW. The specificity of the 'nonspecific' midline and intralaminar thalamic nuclei. *Trends Neurosci.* 1994; 17:52–57. [PubMed: 7512768]
- Hasselmo ME, Anderson BP, Bower JM. Cholinergic modulation of cortical associative memory function. *J Neurophysiol.* 1992; 67:1230–1246. [PubMed: 1597709]
- Hasselmo ME, Bower JM. Acetylcholine and memory. *Trends Neurosci.* 1993; 16:218–222. [PubMed: 7688162]
- Hasselmo ME, Sarter M. Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology.* 2011; 36:52–73. [PubMed: 20668433]
- Hersch SM, Gutekunst CA, Rees HD, Heilman CJ, Levey AI. Distribution of m1-m4 muscarinic receptor proteins in the rat striatum: light and electron microscopic immunocytochemistry using subtype-specific antibodies. *J Neurosci.* 1994; 14:3351–3363. [PubMed: 8182478]
- Kimura M, Minamimoto T, Matsumoto N, Hori Y. Monitoring and switching of cortico-basal ganglia loop functions by the thalamo-striatal system. *Neuroscience research.* 2004; 48:355–360. [PubMed: 15041188]
- Lapper SR, Bolam JP. Input from the frontal cortex and the parafascicular nucleus to cholinergic interneurons in the dorsal striatum of the rat. *Neuroscience.* 1992; 51:533–545. [PubMed: 1488113]
- Linster C, Wyble BP, Hasselmo ME. Electrical stimulation of the horizontal limb of the diagonal band of Broca modulates population EPSPs in piriform cortex. *J Neurophysiol.* 1999; 81:2737–2742. [PubMed: 10368393]
- Matamales M, Bertran-Gonzalez J, Salomon L, Degos B, Deniau JM, Valjent E, Herve D, Girault JA. Striatal medium-sized spiny neurons: identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. *PLoS ONE.* 2009; 4:e4770. [PubMed: 19274089]

- Matsumoto N, Minamimoto T, Graybiel AM, Kimura M. Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J Neurophysiol.* 2001; 85:960–976. [PubMed: 11160526]
- Newman EL, Gupta K, Climer JR, Monaghan CK, Hasselmo ME. Cholinergic modulation of cognitive processing: insights drawn from computational models. *Frontiers in behavioral neuroscience.* 2012; 6:24. [PubMed: 22707936]
- Ostlund SB, Balline BW. Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Learn Behav.* 2007; 35:43–52. [PubMed: 17557390]
- Parent A, Hazrati LN. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res Brain Res Rev.* 1995; 20:91–127. [PubMed: 7711769]
- Ragozzino ME, Mohler EG, Prior M, Palencia CA, Rozman S. Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol Learn Mem.* 2009; 91:13–22. [PubMed: 18845266]
- Rescorla RA. Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes.* 1991; 17:465–474.
- Rescorla RA. Response-outcome associations remain functional through interference treatments. *Animal Learning & Behavior.* 1996; 24:450–458.
- Robbins TW, Roberts AC. Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cereb Cortex.* 2007; 17(Suppl 1):i151–160. [PubMed: 17725997]
- Ruvinsky I, Meyuhos O. Ribosomal protein S6 phosphorylation: from protein synthesis to cell size. *Trends in Biochemical Sciences.* 2006; 31:342–348. [PubMed: 16679021]
- Santini E, Alcacer C, Cacciatore S, Heiman M, Herve D, Greengard P, Girault JA, Valjent E, Fisone G. L-DOPA activates ERK signaling and phosphorylates histone H3 in the striatonigral medium spiny neurons of hemiparkinsonian mice. *J Neurochem.* 2009; 108:621–633. [PubMed: 19187092]
- Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci.* 2000; 23:473–500. [PubMed: 10845072]
- Shiflett MW, Balleine BW. Contributions of ERK signaling in the striatum to instrumental learning and performance. *Behav Brain Res.* 2011a; 218:240–247. [PubMed: 21147168]
- Shiflett MW, Balleine BW. Molecular substrates of action control in cortico-striatal circuits. *Prog Neurobiol.* 2011b; 95:1–13. [PubMed: 21704115]
- Shiflett MW, Brown RA, Balleine BW. Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *The Journal of neuroscience: the official journal of the Society for Neuroscience.* 2010; 30:2951–2959. [PubMed: 20181592]
- Sorimachi M, Kataoka K. High affinity choline uptake: an early index of cholinergic innervation in rat brain. *Brain Res.* 1975; 94:325–336. [PubMed: 1148873]
- Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G. Reward prediction error signaling in posterior dorsomedial striatum is action specific. *J Neurosci.* 2012; 32:10296–10305. [PubMed: 22836263]
- Sullivan MA, Chen H, Morikawa H. Recurrent inhibitory network among striatal cholinergic interneurons. *J Neurosci.* 2008; 28:8682–8690. [PubMed: 18753369]
- Sutton, RS.; Barto, AG. *Reinforcement Learning.* Cambridge, Mass: MIT Press; 1998.
- Threlfell S, Cragg SJ. Dopamine signaling in dorsal versus ventral striatum: the dynamic role of cholinergic interneurons. *Frontiers in systems neuroscience.* 2011; 5:11. [PubMed: 21427783]
- Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron.* 2012; 75:58–64. [PubMed: 22794260]
- Underwood BJ. Interference and forgetting. *Psychological Review.* 1957; 64:49–60. [PubMed: 13408394]
- Valjent E, Bertran-Gonzalez J, Bowling H, Lopez S, Santini E, Matamalas M, Bonito-Oliva A, Herve D, Hoeffler C, Klann E, et al. Haloperidol regulates the state of phosphorylation of ribosomal protein S6 via activation of PKA and phosphorylation of DARPP-32. *Neuropsychopharmacology.* 2011; 36:2561–2570. [PubMed: 21814187]

- Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*. 2011; 334:1569–1573. [PubMed: 22075724]
- Yin HH, Knowlton BJ, Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci*. 2005a; 22:505–512. [PubMed: 16045503]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci*. 2005b; 22:513–523. [PubMed: 16045504]
- Yu AJ, Dayan P. Acetylcholine in cortical inference. *Neural Netw*. 2002; 15:719–730. [PubMed: 12371522]

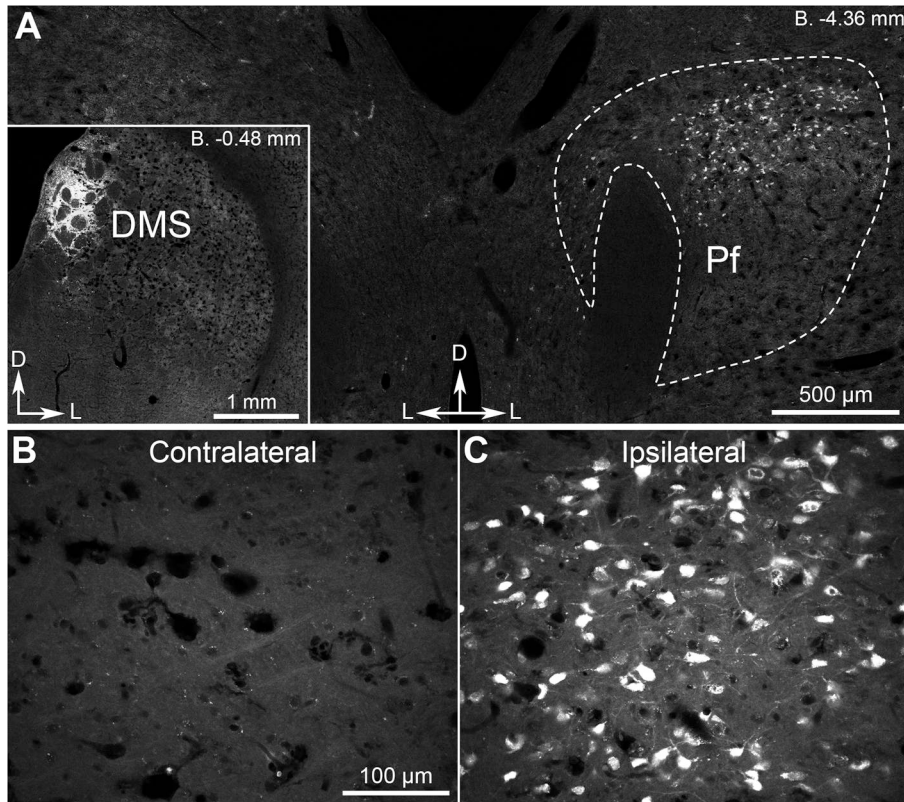


Figure 1. Parafascicular thalamic neurons directly and unilaterally project to dorsomedial striatal territories
 (A) Fluorescent signal recorded in the parafascicular thalamic nucleus (Pf) 4 d after injection of retrograde tracer fluorogold in the DMS (inset). (B, C) Confocal higher-magnification images showing fluorogold fluorescence in Pf ipsilateral or contralateral to the DMS injection. D, dorsal; L, lateral.

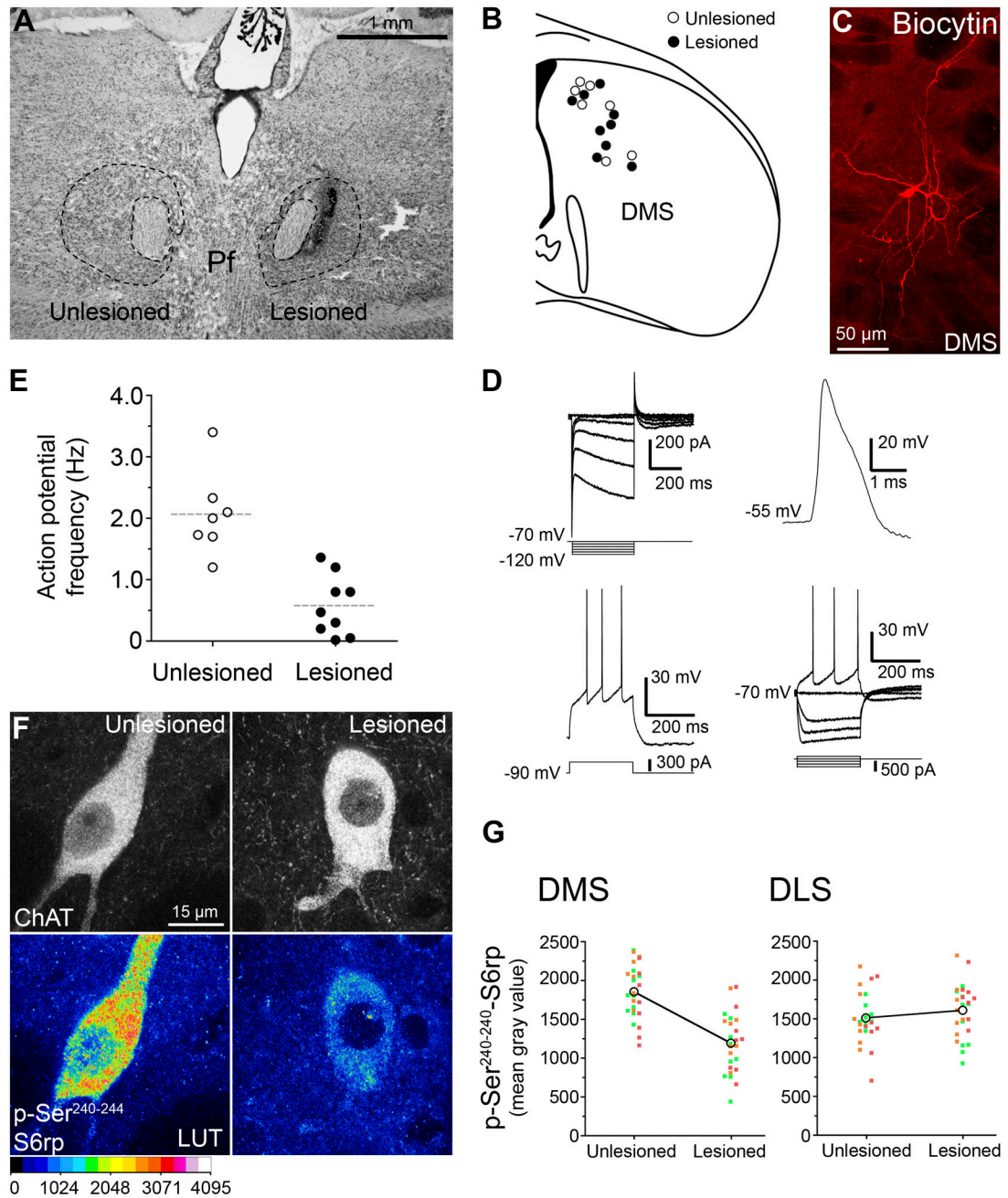


Figure 2. Effect of parafascicular thalamic lesions on cholinergic interneurons of the dorsomedial striatum

(A) Rat Nissl-stained section showing unilateral NMDA excitotoxic lesions of the parafascicular nucleus (Pf). (B) Anatomical localization of randomly selected cholinergic interneurons from unlesioned and lesioned cerebral hemispheres. Each circle represents a single neuron sampled from electrophysiological recording. (C) Example of cholinergic interneuron labeled with biocytin in the DMS. (D) Cellular physiological characteristics of a recorded neuron under whole-cell patch-clamp. Current-voltage relationship recorded by stepping the cell to various hyperpolarizing membrane potentials (top-left panel). Under current-clamp configuration, whole-cell action potential (top-right panel) and

depolarization-triggered action potential firing (bottom panels) were routinely sampled for comparisons with known CIN cellular characteristics. (E) Frequency distribution plot showing basal action potential firing in cholinergic interneurons of lesioned and unlesioned hemisections. (F) High-magnification confocal images showing p-Ser²⁴⁰⁻²⁴⁴-S6rp intensity in ChAT immunoreactive neurons of the DMS ipsilateral (lesioned) or contralateral (unlesioned) to the Pf lesion. A 16 pseudo-color palette (Lookup Table, LUT) highlights the intensity of p-S6rp fluorescence (display range: 0 – 4096). (G) Within-individuals quantification of p-Ser²⁴⁰⁻²⁴⁴-S6rp signal in ChAT immunoreactive neurons of dorsomedial (DMS) and dorsolateral (DLS) striatal territories ipsilateral or contralateral to the Pf lesion. In scatterplots each dot corresponds to one neuron; each color corresponds to a different animal.

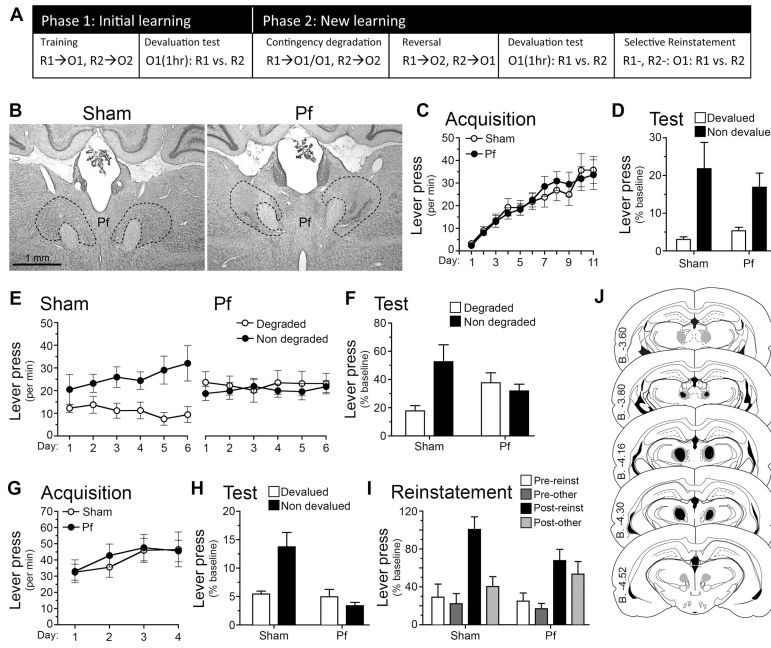


Figure 3. Parafascicular thalamus encodes changes in action-outcome contingencies
 (A) Rats were trained on two lever press responses, R1 and R2, to earn pellets and sucrose outcomes (O1 and O2, counterbalanced). For test rats were satiated for 1 hr on one outcome, i.e. O1 (1hr), prior to a choice test, R1 vs. R2. To induce new learning we used contingency degradation then reversal, which was tested using outcome devaluation and selective reinstatement tests. See supplemental methods for a full description of these procedures. (B) Rat Nissl-stained sections showing bilateral Sham or NMDA excitotoxic lesions of the parafascicular nucleus (Pf). (B–H) Mean or % baseline rate of lever pressing (± 1 SEM) during: (C) acquisition of initial action-outcome contingencies averaged over levers; (D) outcome devaluation testing; (E) acquisition of contingency degradation; (F) contingency degradation testing in extinction; (G) acquisition of the reversed contingencies; (H) outcome devaluation testing of the reversed contingencies; (I) reinstatement of reversed contingencies. (J) Minimal (black) and maximal (grey) extent of NMDA-induced excitotoxic lesions of the Pf.

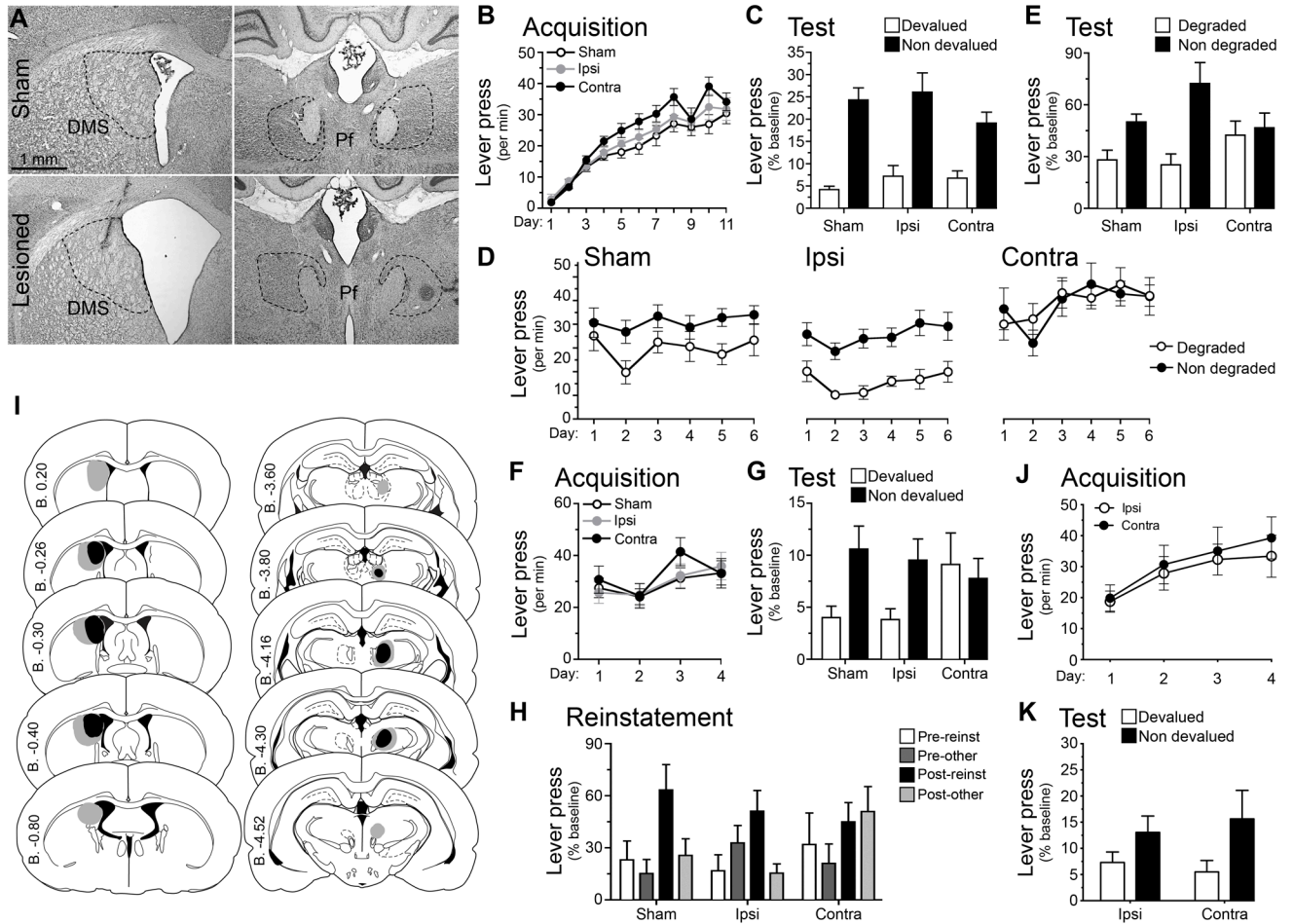


Figure 4. Disconnection of the thalamo-striatal pathway confirms parafascicular involvement in action-outcome learning in the posterior dorsomedial striatum

(A) Rat Nissl-stained sections showing anatomical disconnection through unilateral DMS and Pf NMDA excitotoxic lesions (sham and contralateral groups shown). (B–H) Mean or % baseline rate of lever pressing (± 1 SEM) during: (B) acquisition of initial action-outcome contingencies averaged over levers; (C) outcome devaluation testing; (D) acquisition of contingency degradation; (E) contingency degradation testing; (F) acquisition of the reversed contingencies; (G) outcome devaluation testing of the reversed contingencies; (H) reinstatement of reversed contingencies; (I) Minimal (black) and maximal (grey) lesions in the Pf and posterior DMS. Lesion side was counterbalanced; and (J) Reversal acquisition and (K) outcome devaluation after Pf-aDMS disconnection. See also Figure S1.

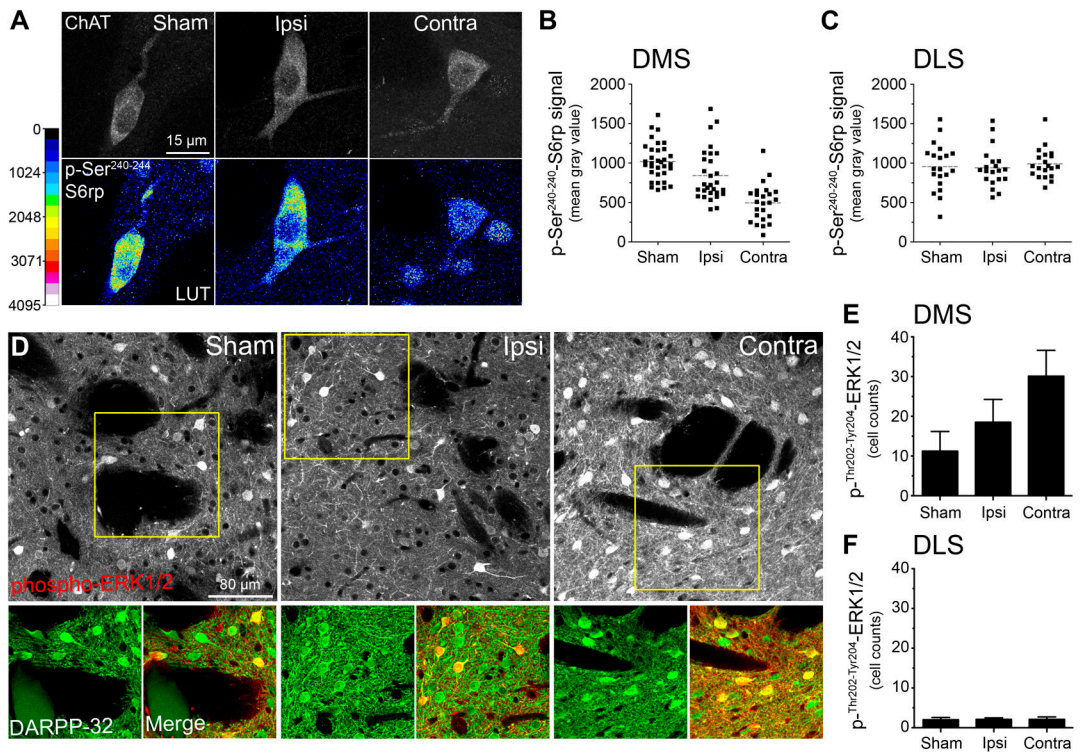


Figure 5. Parafascicular thalamic lesions predict CIN but not MSN activity in posterior dorsomedial striatum of goal-directed rats

(A) High-magnification confocal images showing p-Ser²⁴⁰⁻²⁴⁴-S6rp intensity in ChAT immunoreactive neurons of the intact DMS of sham, ipsilaterally or contralaterally lesioned rats immediately after reinstatement test. A 16 pseudo-color palette (Lookup Table, LUT) highlights the intensity of p-S6rp fluorescence (display range: 0 – 4096) (B–C) Quantification of p-Ser²⁴⁰⁻²⁴⁴-S6rp signal in all ChAT immunoreactive neurons of posterior dorsomedial (DMS; B) and dorsolateral (DLS; C) striatal territories of the different groups. In scatterplots, each dot corresponds to one neuron. (D) Confocal images showing double immunofluorescence of phospho-Thr²⁰²-Tyr²⁰⁴-ERK1/2 (red) and DARPP-32 (green) in the pDMS of rats treated as in A. (E–F) Quantification of phospho-Thr²⁰²-Tyr²⁰⁴-ERK1/2 in pDMS and adjacent DLS in the same rats.

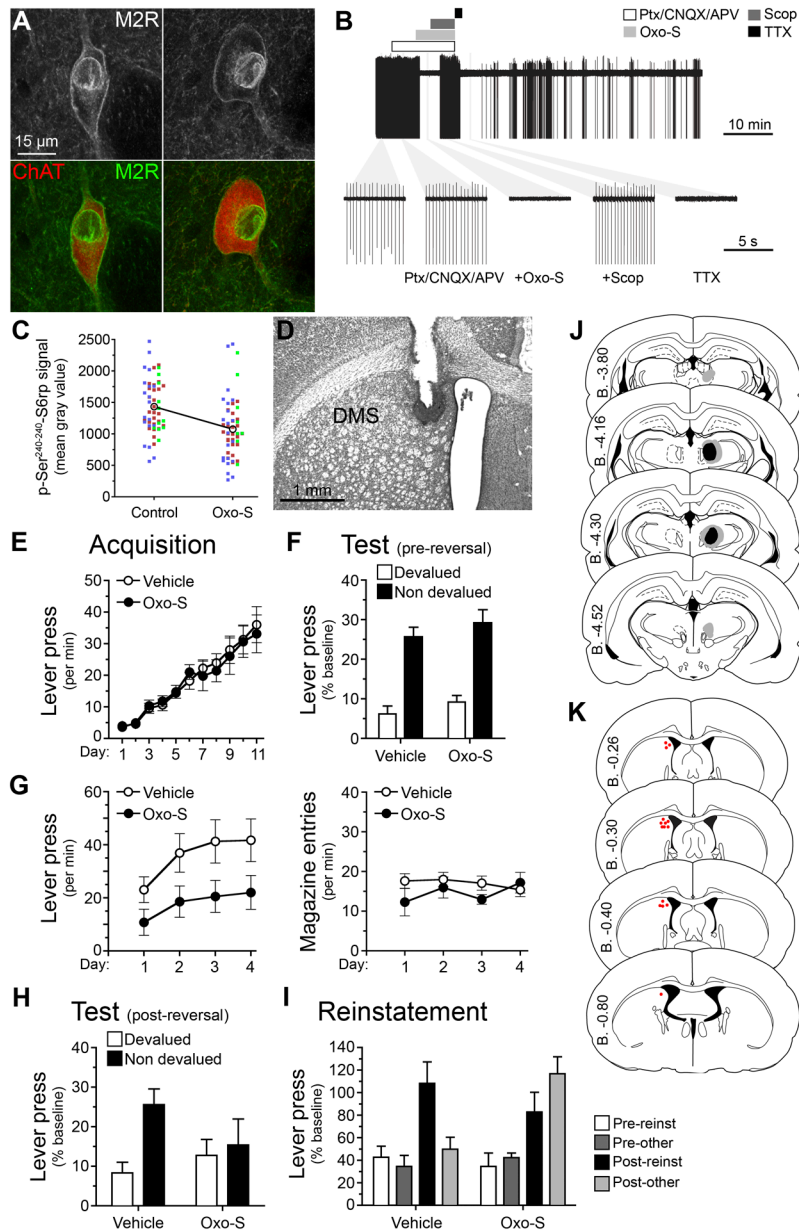


Figure 6. Reduction in cholinergic-activity in the DMS using the M2/M4 agonist Oxotremorine-S replicates the effect of Pf – DMS disconnection on goal-directed learning

(A) High-magnification confocal images showing ChAT and M2-muscarinic receptor (M2R) immunoreactivities. (B) Top, a representative raw trace showing inhibition of spontaneous action potential firing in a cholinergic interneuron by oxotremorine (Oxo-S, 1 μ M) and tetrodotoxin (TTX, 100 nM) in the presence of synaptic blockers – picrotoxin (Ptx, 100 μ M), CNQX (10 μ M) and APV (100 μ M). The bars above indicate time periods of drug applications. Concentration of scopolamine (Scop) was 3 μ M. Bottom, expanded time periods of the top trace during various drug applications, with the first left being at basal condition. All recordings were done in brain slices. (C) Within-individuals quantification of p-Ser²⁴⁰⁻²⁴⁴-S6rp signal in striatal ChAT immunoreactive neurons after exposing counterbalanced hemisections to 1-hour control/Oxo-S (1 μ M) incubation. Each dot corresponds to one neuron; each color corresponds to a different animal. (D) Rat Nissl-

stained section showing unilateral cannulation for Oxo-S or vehicle infusion into the DMS. (E–I) Mean or % baseline rate of lever pressing (\pm M1 SEM) during: (E) acquisition of initial action-outcome contingencies averaged over levers; (F) outcome devaluation testing; (G) acquisition of the reversed contingencies, lever press responding (left) and magazine entries (right); (H) outcome devaluation testing of the reversed contingencies; (I) reinstatement of reversed contingencies; (J) Minimal (black) and maximal (grey) lesions in the Pf. Lesion side was counterbalanced. (K) Cannula placements in the posterior DMS. Placement side was counterbalanced.