Striatum expresses region-specific plasticity consistent with distinct memory abilities

Jonathan Touboul, L. Venance, Sylvie Perez, Yihui Cui, Gaëtan Vignoud, Elodie Perrin, Alexandre Mendes, Zhiwei Zheng

Striatum expresses region-specific plasticity consistent with distinct memory abilities

Sylvie PEREZ[1]†, Yihui CUI[1,2]†, Gaëtan VIGNOUD[1,3]†, Elodie PERRIN[1], Alexandre MENDES[1], Zhiwei ZHENG[2], Jonathan TOUBOUL[4]* and Laurent VENANCE[3,5]*

[1]Center for Interdisciplinary Research in Biology (CIRB), College de France, CNRS, INSERM, Université PSL, Paris, France.

[2] Department of Neurobiology, and Department of Neurology of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, 310058, Hangzhou, China.

[3]MAMBA-Modelling and Analysis for Medical and Biological Applications, Inria Paris, LJLL (UMR-7598) -Laboratory Jacques-Louis Lions, Paris, France.

[4]Department of Mathematics and Volen National Center for Complex Systems, Brandeis University, Waltham, MA, USA.

[5]Lead contact: Laurent VENANCE laurent.venance@college-de-france.fr

†: These authors contributed equally; *: co-senior authors

SUMMARY

The striatum mediates two learning modalities: goal-directed behavior in dorsomedial (DMS) and habits in dorsolateral (DLS) striatum. The synaptic bases of these learnings are still elusive. Indeed, while ample research has described DLS plasticity, little remains known about DMS plasticity and its involvement in procedural learning. Here, we find symmetric and asymmetric anti-Hebbian spike-timing-dependent plasticity (STDP) in DMS and DLS, respectively, with opposite plasticity dominance upon increasing corticostriatal activity. During motor skill learning, plasticity is engaged in DMS and striatonigral DLS neurons only during early learning stages, whereas striatopallidal DLS neurons are mobilized only during late phases. With a mathematical modelling approach, we find that symmetric anti-Hebbian STDP favored memory flexibility, while asymmetric anti-Hebbian STDP favored memory maintenance, consistent with memory processes at play in procedural learning.

INTRODUCTION

The dorsal striatum is critical for action selection and initiation (Yin and Knowlton, 2006; Graybiel and Grafton, 2015; Jin and Costa, 2015) and represents a major site for memory formation encoding for procedural learning (Perrin and Venance, 2019). The dorsal striatum is composed of two main anatomico-functional regions, the dorsolateral striatum (DLS) and dorsomedial striatum (DMS) based on topographic cortical glutamatergic afferents. Cortical inputs from the sensorimotor and premotor cortices project somatotopically to DLS, whereas associative and prefrontal cortices project, with a decreasing topography, to DMS (Hunnicutt et al., 2016; Hintiryan et al., 2016; Hooks et al., 2018) with fewer inputs from somatosensory cortical areas (Reig and Silberberg, 2014). Moreover, DLS and DMS appear to engage at different learning phases: the classical view posits that during reward-guided instrumental learning DMS supports goal-directed behavior, while DLS is gradually involved in later learning phases associated with habit formation and performance (Costa et al., 2004; Yin and Knowlton, 2006; Balleine and O'Doherty, 2010; Corbit and Janak, 2010; Thorn et al., 2010; Gremel and Costa, 2013; Burton et al., 2015; Vandaele et al., 2019). Similarly, during motor skill learning DMS appears to play a crucial role during initial phases of fast improvements, while DLS is determinant for slower learning phases as experience accumulates (Graybiel and Grafton, 2015; Jin and Costa, 2015; Costa et al., 2004; Yin et al., 2009; Xiong et al., 2015). Nevertheless, there is evidence that DLS does not only activate at late learning phases, but is engaged, together with DMS, from early training phases (Thorn et al., 2010; Gremel and Costa, 2013; Kimchi et al., 2009; Stalnaker et al., 2010; Kupferschmidt et al., 2017; Bergstrom et al., 2018). Acquisition and maintenance of motor skills and habits involve corticostriatal long-term synaptic efficacy changes (Perrin and Venance, 2019). Indeed, *in vivo* proxies for plasticity, such as changes in firing activity (Costa et al., 2004; Yin et al., 2009; Thorn and Graybiel, 2014; Barnes et al., 2011; Koralek et al., 2012; O'Hare et al., 2016; Athalye et al., 2018; Peters et al., 2021) or in evoked-LFP

(Xiong et al., 2015), were detected in the corticostriatal pathway throughout procedural learning. Conversely, triggering corticostriatal synaptic plasticity was shown to modify habitual behavior (Xiong et al., 2015; Ma et al., 2018). Although these findings clearly highlight a correlative and/or causal link between corticostriatal plasticity and procedural learning, the nature and contribution of DLS and DMS long-term plasticity remain to be fully determined.

To investigate plasticity implication in memory storage and retrieval in DMS and DLS, we characterized the spike-timing-dependent plasticity (STDP) and observed anti-Hebbian STDP in both DMS and DLS, with specific profiles: a symmetric long-term depression (LTD) arises in DMS, contrasting with an asymmetric anti-Hebbian STDP in DLS. When corticostriatal activity scaled up, we found opposite polarity plasticity in DMS and DLS, since LTD prevailed in DMS, while long-term potentiation (LTP) prevailed in DLS. Furthermore, striatal medium-sized spiny neurons (MSNs) from the DMS exhibited opposite plasticity, in a specific activity regime, depending on whether they belonged to the striatonigral (direct) or striatopallidal (indirect) pathway. During motor skill learning, we found that during early learning phases, plasticity was engaged for all DMS-MSNs, and striatonigral MSNs in DLS. In contrast, during late learning phases, we found that only striatopallidal MSNs in DLS were mobilized. We developed a mathematical model to quantify the capacity of these plasticity rules for memory formation and storage, and relearning, a capacity distinct from recall in that it is driven by rewards. Our model predicted that asymmetric anti-Hebbian STDP facilitated the maintenance of memory, whereas symmetric LTD allowed a swift turnover of memories, potentially enhancing memory flexibility. These findings reveal how distinct plasticity maps in DLS and DMS could endow striatum with complementary capacities for procedural learning allowing flexibility in memory acquisition and stabilization of memories potentially for habit formation.

RESULTS

Distinct anti-Hebbian STDP profiles in sensorimotor and associative striatum

We used two brain slice preparations that preserved connections between the sensorimotor cortex (S2) and DLS, or associative cortex (CG2) and DMS, and allowed to stimulate within cortical layer 5 while recording MSNs (Fig. 1) (Fino et al., 2018). At both synapses, we first applied the same STDP protocol consisting of 100 pairings at 1Hz with prescribed timing Δt_{STDP}~-15 or +15ms (Fig. 1a).

We investigated DLS-STDP (Fig. 1b and Table S1) and observed asymmetric (*i.e.* distinct plasticity polarity on both sides of $\Delta t_{STDP=}0$) anti-Hebbian STDP: LTP for post-pre pairings (Fig. 1b2) and LTD for pre-post pairings (Fig. 1b3). Anti-Hebbian qualifies STDP with pre-post LTD. Post-pre pairings induced LTP ($p<0.0001$, $n=16$; Fig. 1b2 and Fig. S1a), whereas pre-post pairings induced LTD ($p<0.0001$, $n=14$; Fig. 1b3 and Fig. S1b). This is in line with DLS-STDP displaying anti-Hebbian polarity in native conditions (Fino et al., 2005; Fino et al., 2010; Mendes et al., 2020). DLS-STDP with a Hebbian polarity has also been reported (Pawlak and Kerr, 2008; Shen et al., 2008), but caused by the use of $GABA_A$ receptor antagonists. Indeed, we previously showed that GABA acts as an Hebbian/anti-Hebbian switch (Paillé et al., 2013; Valtcheva et al., 2017), so polarity of the corticostriatal STDP depends on whether $GABA_A$ receptor antagonists are used (*in vitro* Hebbian STDP (Pawlak and Kerr, 2008; Shen et al., 2008)) or not (*in vitro* anti-Hebbian STDP (Fino et al., 2005; Fino et al., 2010; Mendes et al., 2020)). We thus recorded STDP in the absence of $GABA_AR$ antagonist to preserve the local striatal microcircuits and the anti-Hebbian polarity also observed *in vivo* (Schulz et al., 2010; Morera-Herreras et al., 2019).

In DMS (Fig. 1c and Table S1), post-pre and pre-post pairings induced a symmetric (*i.e.* similar plasticity polarity, here LTD, on both sides of $\Delta t_{STDP=}0$) anti-Hebbian STDP (Fig. 1c2 and 1c3).

Indeed, LTD was observed following post-pre pairings ($p<0.0001$, $n=11$; Fig. 1c2 and Fig. S1c) as well as following pre-post pairings ($p<0.0001$, $n=16$; Fig. 1c3 and Fig. S1d).

Corticostriatal STDP in DMS and DLS displayed both anti-Hebbian plasticity, with symmetric and asymmetric profiles, respectively.

Post-pre DLS-LTP and DMS-LTD are NMDAR-mediated, whereas pre-post DLS- and DMS-LTD are CB_1R-mediated.

Regarding DLS-STDP, LTP induced by post-pre pairings was NMDAR-mediated since prevented by D-AP5 (50μM) ($p=0.3775$, $n=9$), whereas LTD induced by pre-post pairings was CB_1R-mediated since precluded by AM251 (3μM) ($p=0.5688$, $n=5$) (Fig. 2a-b and Table S1), as previously reported (Shen et al., 2008; Fino et al., 2010).

We next investigated the signaling pathways associated with DMS-STDP (Fig. 2c-d and Table S1). We first tested whether post-pre and pre-post LTD were CB_1R-mediated, since CB_1R-mediated LTD is a dominant LTD form (Shen et al., 2008; Fino et al., 2010; Cui et al., 2016; Mathur et al., 2012) in DLS. We found that AM251 left unaffected LTD induced by post-pre pairings ($p<0.0001$, $n=8$) (Fig. 2c), while prevented LTD with pre-post pairings ($p=0.4614$, $n=5$) (Fig. 2d). We next investigated the involvement of NMDARs, and observed that D-AP5 prevented post-pre DMS-LTD from arising ($p=0.1059$, $n=7$) (Fig. 2c). We examined the involvement of L- and T-type voltage-sensitive calcium channels (VSCCs), which are activated by the back-propagating action potential (Feldman, 2012). Mibefradil (20μM), an antagonist of T-type VSCCs, prevented post-pre LTD ($p=0.0751$, $n=8$) (Fig. 2c). Therefore, DMS-LTD is NMDAR- and T-type VSCCs-mediated for post-pre pairings, and endocannabinoid-mediated for pre-post pairings.

Post-pre DLS-LTP and DMS-LTD are NMDAR-mediated, and pre-post DLS- and DMS-LTD are both CB_1R-mediated.

Dominance of opposite polarity of plasticity in DMS and DLS with increasing corticostriatal activity.

Expression map of STDP is not only shaped by spike timing (Δt_{STDP}) but also by the frequency at which pairings are presented ($F_{pairings}$) (Feldman, 2012). We further characterized the induction rules of corticostriatal STDP in DLS (n_{total}=130 DLS-MSNs) and DMS (n_{total}=125 DMS-MSNs) by varying together Δt_{STDP} ($-100 \leq \Delta t_{STDP} \leq +100$ms) and $F_{pairings}$ ($1 \leq F_{pairings} \leq 5$Hz). We varied $F_{pairings}$ up to 5Hz, (i) to stay in the STDP (and not rate-coding plasticity) domain since for $F_{pairings}>5$Hz Δt_{STDP} becomes shorter (*i.e.* 50ms) than the domain of STDP expression, and (ii) this frequency range is in line with *in vivo* corticostriatal cell assembly activities (Oberto et al., 2022).

We explored the domain in which STDP has been widely reported, *i.e.* $-30<\Delta t_{STDP}<+30$ms (Feldman, 2012), and then expanded investigation up to ±100ms. We observed two main facts: (i) the STDP expression domain (Δt_{STDP}) is narrower in DLS than in DMS for low $F_{pairings}$ and is widening with increasing firing activity, and (ii) the existence of an opposite dominance of LTP and LTD in DLS and DMS, respectively (Fig. 3 and Table S2).

LTP dominates in DLS when pairing frequency scales up. In DLS, we observed a dominance of LTP together with an enlargement of STDP expression domain with increasing $F_{pairings}$ (Fig. 3a-d). For $F_{pairings}$=1Hz (n=42 MSNs), plasticity was induced within a narrow Δt_{STDP} with LTP ($p<0.0001$, n=21) and LTD (p=0.0004, n=13) restricted to $-30<\Delta t_{STDP}<0$ms and $0<\Delta t_{STDP}<30$ms, respectively (Fig. 3a). Pairings for $\Delta t_{STDP} \approx -35$ or +35ms did not induce plasticity (p=0.6205, n=4 and p=0.4670, n=4). We have previously shown that no plasticity was induced for Δt_{STDP} beyond ±30ms even for wide Δt_{STDP}, *i.e.* ±250 and ±500ms (Valtcheva and Venance, 2016). For $F_{pairings}$=2.5Hz (n=36), plasticity domain broadened considerably up to $-100<\Delta t_{STDP}<+100$ms, with LTP being induced for-

100<Δt_{STDP}<+30ms and LTD for +20<Δt_{STDP}<+100ms (Fig. 3b). Indeed, most MSNs subjected to post-pre pairings (-100<Δt_{STDP}<0ms) displayed LTP (p=0.0003, n=22). LTP domain was extended also on the pre-post side (0<Δt_{STDP}<+30ms: p=0.0082, n=6) with 67% of LTP for only 33% of LTD (Fig. 3b); this should be compared to the 8% of LTP and 85% of LTD expressed for $F_{pairings}$=1 Hz and 0<Δt_{STDP}<+30ms. Interestingly, the expression domain of LTD was also enlarged and shifted to +30<Δt_{STDP}<+100ms where LTD was the exclusive form of plasticity (p=0.0006, n=8) (Fig. 3b). For $F_{pairings}$=5Hz (n=52), both post-pre and pre-post pairings spanning from -100 to +100ms induced LTP (-100<Δt_{STDP}<0ms: p<0.0001, n=33; 0<Δt_{STDP}<+100ms: p=0.0002, n=19) (Fig. 3c). Therefore, in DLS, increasing frequency in a time-coding paradigm, such as STDP, shapes plasticity map by favoring LTP (Fig. 3d). This feature was robust and conserved with a rate-coding paradigm at a high frequency. Indeed, a single round of high-frequency stimulation (HFS) induced LTP exclusively (p<0.0001, n=6) (Fig. 3e).

In DLS, the asymmetric anti-Hebbian STDP at 1Hz turned into a symmetric Hebbian STDP at 5Hz. Furthermore, the temporal window of STDP induction was enlarged with increasing frequency (from 60ms at 1Hz to 200ms at 5Hz) and LTP became the prominent form of plasticity (Fig. 3d and 3i).

LTD dominance in DMS when pairing frequency scales up. In DMS, for $F_{pairings}$ from 1 to 5 Hz, we observed a dominance of LTD, *i.e* a symmetric anti-Hebbian STDP, except at 2.5Hz for which asymmetric STDP appeared (Fig. 3f-i). For $F_{pairings}$=1Hz (n=39 MSNs), plasticity domain in DMS was wider than in DLS since comprised between -100 and +100ms (Fig. 3f). Indeed, LTD was induced for post-pre pairings with -100<Δt_{STDP}<0ms (p=0.0033, n=15) as well as for pre-post pairings with 0<Δt_{STDP}<+100ms (p<0.0001, n=24). For $F_{pairings}$=2.5Hz (n=59), a complex plasticity map was obtained since post-pre pairings with -30<Δt_{STDP}<0ms induced not only LTD (52% of the MSNs) but also LTP (41% of the MSNs) (Fig. 3g). Indeed, among 27 neurons subjected to post-pre

pairings (-30<Δt_{STDP}<0ms), 14 and 11 cells displayed LTD (72±3%, p<0.0001) and LTP (134±6%, p=0.0002), respectively; the 2 remaining ones did not show plasticity. There was no correlation (Pearson's correlation) between the Δt_{STDP} (p=0.7888), input resistance (Ri, p=0.5135), resting membrane potential (RMP, p=0.9740) or excitatory post-synaptic potential (PSP) amplitude (p=0.4476) of the recorded MSNs and STDP polarity for -30<Δt_{STDP}<0ms (Fig. S2). For pairings beyond -30 ms (-100<Δt_{STDP}<-30ms), LTD was observed (p=0.0085, n=10). For pre-post pairings (0<Δt_{STDP}<100ms), LTD was the main form of plasticity to be induced (p<0.0001, n=22), *i.e.* a similar picture than those obtained for pre-post pairings at 1Hz. For $F_{pairings}$=5Hz (n=27), both post-pre and pre-post pairings induced LTD (-100<Δt_{STDP}<0ms: p<0.0001, n=16; 0<Δt_{STDP}<+100ms: p=0.0200, n=11) (Fig. 3h). Notably, at $F_{pairings}$=5Hz, none of the recorded MSNs (n=11) at -30<Δt_{STDP}<0 ms exhibited anymore LTP, as it was the case for almost half of the MSNs for $F_{pairings}$=2.5Hz.

In DMS, LTD dominates regardless Δt_{STDP} and $F_{pairings}$ (Fig. 3i), with the noticeable exception of paired activity at $F_{pairings}$=2.5Hz with -30<Δt_{STDP}<0ms for which LTD as well as LTP were observed. Thus, in DMS with increasing $F_{pairings}$, the symmetric anti-Hebbian STDP observed at 1Hz was transitorily flipped at 2.5Hz to a mixture of symmetric and asymmetric anti-Hebbian STDP, and then at 5Hz only symmetric anti-Hebbian STDP was observed. We then tested the effect of HFS and only LTD was induced (p<0.0001, n=7) (Fig. 3j).

In conclusion, LTP dominates in DLS with increasing $F_{pairings}$ (and HFS) making the plasticity map evolving from asymmetric anti-Hebbian to symmetric Hebbian STDP, whereas in DMS LTD is the prominent form of plasticity showing mainly symmetric anti-Hebbian STDP (at the exception of $F_{pairings}$=2.5Hz).

Striatonigral and striatopallidal DMS-MSNs exhibit frequency-specific plasticity

We investigated whether the observed dichotomy regarding LTP/LTD expression in DMS-MSNs for $F_{pairings}$=2.5Hz overlapped that of striatonigral and striatopallidal MSNs. The two MSN subtypes express different dopaminergic receptors, D_1R- and D_2R-like for the striatonigral and striatopallidal pathways, respectively (Calabresi et al., 2014; Bonnavion et al., 2019). At this stage we used Drd1a-eGFP mice to investigate STDP in D_1^+ and D_1^- DMS-MSNs (Fig. 4 and Table S3). We performed double patch-clamp recordings of neighboring MSNs pairs (<50µm away, *n*=8 pairs) composed of one D_1^+ and one D_1^- MSNs. Both DMS-MSNs were subjected to the same STDP protocol: 100 post-pre pairings at $F_{pairings}$=2.5Hz and Δt_{STDP}=-15ms (Fig. 4a). Figure 4b shows an example of a D_1^+/D_1^- DMS-MSN pair which exhibits opposite plasticity, LTD (p<0.0001) and LTP (p<0.0001), respectively. In all 8 D_1^+/D_1^- MSN pairs, we found that D_1^+ MSNs displayed LTD (p<0.0001), whereas D_1^- MSNs displayed exclusively LTP (p<0.0001, n=8) (Fig. 4c).

We next investigated the signaling pathways associated with LTD and LTP expressed in D_1^+ and D_1^- MSNs. Both LTD in D_1^+ MSNs and LTP in D_1^- MSNs, induced by post-pre pairings at 2.5Hz, were NMDAR-mediated. Indeed, D-AP5 prevented plasticity in D_1^+ (p=0.1255, n=7) and D_1^- MSNs (p=0.5884, n=6) (Fig. 4d).

We observed in rats that at $F_{pairings}$=1Hz the majority (73%) of the DMS-MSNs randomly chosen exhibited LTD (Fig. 3f), suggesting that this LTD could be induced indistinguishably in $D1^+$ and $D1^-$ DMS-MSNs. Using Drd1a-eGFP mice, we confirmed this observation made in rats, since LTD was induced in both $D1^+$ and $D1^-$ DMS-MSNs when subjected to post-pre pairings at 1Hz with Δt_{STDP}=-15ms (D_1^+ MSNs: p=0.0005, n=5; D_1^- MSNs: p<0.0001, n=6) (Fig. 4e).

In DMS the LTP/LTD dichotomy observed for $F_{pairings}$=2.5Hz relies on the belonging of MSNs to the striatonigral or striatopallidal pathway, whereas for other $F_{pairings}$ (or HFS), similar plasticity (LTD) is induced in most of DMS-MSNs.

Region-specific involvement of STDP during procedural learning.

We next investigated the engagement of evoked-LTP and -LTD in DLS and DMS during a motor skill learning (Fig. 5 and Table S4). For this purpose, we performed *ex vivo* occlusion experiments and tested STDP expression (post-pre pairings for $F_{pairings}$=2.5Hz) 24 hours after habituation (control group, n=9 mice), the first day (early-trained group, n=8) or the seventh day (late-trained group, n=10) of Rotarod training (Fig. 5a and Methods). We chose $F_{pairings}$=2.5Hz since it allows to capture most of the DMS- and DLS-STDP features, and more particularly the DMS-plasticity dichotomy between striatonigral and striatopallidal DMS-MSNs. Here, we used Drd2-EGFP mice to investigate STDP in D_2^+ (striatopallidal) and D_2^- (striatonigral) DMS- and DLS-MSNs; it allows also to ensure that the observation with Drd1a-eGFP mice (Fig. 4) still stand in another mouse line. Most of mice reached a plateau in term of motor learning performance by the second or third day of training (Fig. 5a). We exploited the fact that the Rotarod is a non-lateralized motor task to test DLS- and DMS-STDP induction in the same animals: for each mouse, one brain hemisphere was used to assess DLS-STDP and the other for DMS-STDP.

In DMS, habituated mice displayed LTD in D_2^- MSNs (p=0.0002, n=7) (Fig. 5b1) and LTP in D_2^+ MSNs (p<0.0001, n=6) (Fig. 5b2), in line with results obtained using Drd1a-eGFP mice (Fig. 4a-c). Plasticity magnitudes observed in habituated (Fig. 5b1 and 5b2) and in naïve mice (Fig. 4c) were not significantly different (LTD: p=0.9174; LTP: p=0.4770). After one day of training on the accelerating Rotarod, the early-trained group showed an occlusion of plasticity in both D_2^-

(p=0.1566, n=6) and D_2^+ (p=0.5850, n=7) DMS-MSNs (Fig. 5c1 and c2). After seven days of training, the late-trained group showed again expression of LTD in D_2^- DMS-MSNs (p<0.0001, n=10) (Fig. 5d1) as well as LTP in D_2^+ DMS-MSNs (p<0.0001, n=8) (Fig. 5d2). These later plasticity magnitudes were similar to those in habituated mice (LTD: p=0.9365; LTP: p=0.9523). There was a significant interaction between synaptic plasticity and stages of motor skill learning (habituated, early- and late-stages) as well as the belonging of DMS-MSNs to the striatonigral or striatopallidal pathways (two-way ANOVA, $F_{(2,38)}$=5.721, p=0.007). These results indicate that DMS-STDP (LTP and LTD) are engaged during early phases of training and disengaged at later stages of motor skill learning.

In DLS, habituated mice displayed LTP in both D_2^- (p<0.0001, n=7) and D_2^+ (p<0.0001, n=7) MSNs (Fig. 5e1 and e2), in line with evoked-STDP in Drd1a-eGFP mice (Fig. 4a-c). The early-trained group exhibited distinct synaptic efficacy changes depending on D_2^- and D_2^+ DLS-MSNs. Indeed, D_2^- MSNs showed an occlusion of plasticity (p=0.2795, n=9) (Fig. 5f1), while LTP was induced in D_2^+ MSNs (p<0.0001, n=7) (Fig. 5f2). This later LTP was of similar magnitude to the evoked-LTP in habituated mice (p=0.4933). Conversely, the late-trained group showed a reverse plasticity picture to the one observed for the early-trained group: LTP was successfully evoked in D_2^- DLS-MSNs (p<0.0001, n=10) (Fig. 5g1), whereas plasticity was occluded this time in D_2^+ DLS-MSNs (p=0.4223, n=7) (Fig. 5g2). Evoked-LTP in D_2^- DLS-MSNs displayed similar amplitude in late-trained and habituated groups (p=0.4923). There was an interaction between synaptic plasticity and habituated/early-/late-stages as well as striatonigral and striatopallidal DLS-MSNs phenotype (two-way ANOVA, $F_{(2,41)}$=9.531, p<0.001).

Striatonigral and striatopallidal DLS-MSNs are selectively engaged, in terms of STDP, depending on the stages of motor skill learning: only striatonigral DLS-MSNs are involved at early stages, and

during the late stages only striatopallidal DLS-MSNs are engaged. DMS-MSN plasticity is mobilized in both MSN populations during early stages and then showed a disengagement during late stages. Interestingly, striatonigral DLS-MSNs show the same plasticity profiles, and occlusion, as striatopallidal DMS-MSNs.

Reduced mathematical model of the striatal network

To investigate whether the distinct forms of STDP in DMS and DLS (Fig. 1) subtend different learning properties, we considered a simplified model of the corticostriatal system, composed of a fixed number P of cortical neurons projecting to one MSN. We quantified the capacity of this system to retain memory as a function of the form of corticostriatal STDP, all parameters equal otherwise. The MSN, modeled as an integrate-and-fire neuron, received inputs from P cortical neurons, as well as activity from other cells represented by a Poisson process (Fig. 6). The synaptic weight W_i between cortical neuron $i \in \{1, \cdots, P\}$ and MSN is subject to plasticity, through an all-to-all pair-based learning rule with instantaneous updates given by

$$\Delta W_i = \begin{cases} A_{pre-post} \exp\left(-\frac{\Delta t}{T_{STDP}}\right) & if\ the\ MSN\ spikes\ at\ time\ t_{post}, \\ A_{post-pre} \exp\left(\frac{\Delta t}{T_{STDP}}\right) & if\ cortical\ neuron\ i\ spikes\ at\ time\ t_{pre,i}, \end{cases}$$

with $\Delta t = t_{post} - t_{pre,i}$.

By convention, we fixed $A_{pre-post} = -1$, and varied the parameter $A_{post-pre}$; $A_{post-pre} < 0$ corresponds to symmetric anti-Hebbian STDP, while $A_{post-pre} > 0$ corresponds to asymmetric anti-Hebbian learning rules (Fig. 6).

The MSN was presented with various patterns of cortical activity, that included two types of input: (i) *Cortical patterns:* coincident spikes from a prescribed sub-group of N_{stim} cortical input neurons,

with a Gaussian jitter of standard deviation τ_p and (ii) *Random cortical firing:* Poisson spikes from all cortical neurons with rate λ_r/P. Examples of such patterns are presented in Figure 6: pattern A, with cortical neurons 1, 3 and 4 fired and pattern B cortical neurons 2, 3 and 4 (spikes in green), with superimposed random spikes (grey). N_p patterns were built according to these principles and split randomly with 50% chance into *rewarded* and *non-rewarded* patterns (A is a non-rewarded pattern $(-)$ and B is a rewarded pattern $(+)$).

Rewarded patterns are associated to abstract *reward signals* representing neuromodulation (*e.g.* dopaminergic signaling) (Foncelle et al., 2018; Gerstner et al., 2018; Brzosko et al., 2019), and resulting in a potentiation of the synapses associated with cortical neurons that fired in a window around the presentation of rewarded patterns (including both the neurons that were associated with the pattern and those associated with noise; red bands in Figure 6). Because of the prominent role of depression in anti-Hebbian learning, particularly in the symmetric case, this potentiation mechanism (referred to as reward-LTP) is essential to maintain some spiking activity (Thorn and Graybiel, 2014).

To quantify the occurrence of the presentation of patterns, we opted for simplicity to split time into windows of 100 ms, and, in each of this time bins, to present a randomly chosen pattern with a probability η (and no pattern presented, with probability $1-\eta$, bins labeled $\emptyset$ in Figure 6). A rewarded cortical pattern is learnt when the MSN spiked in response to the synchronous cortical activity. Conversely, a non-rewarded pattern is learnt if the MSN did not fire during pattern presentation. The accuracy of striatal learning was estimated during test protocols conducted throughout the task on a network devoid of any plasticity and noise, with a metric combining the fraction of rewarded patterns correctly eliciting spikes from the MSN and of non-rewarded ones that did not trigger any spike.

To avoid transient effects associated with the initialization of synaptic variables, we simulated an initial phase of *spontaneous activity* of the cortical network, defined by the presentation of patterns with probability $\eta = \eta_m$ in the absence of reward-LTP. During the *learning* phase, patterns are presented at each iteration and rewards were provided for rewarded patterns. This phase emulates a learning stage, with the MSN eventually gaining an ability to discriminate patterns by spiking in response to rewarded patterns and not spiking in response to non-rewarded patterns. The capacity of the network to keep patterns in memory was then estimated during a *maintenance* phase, with stimulus presented with probability $\eta = \eta_m$ in the absence of reward-LTP. Finally, the capacity to relearn previously learned patterns was tested in a protocol identical to the learning phase (Table S5 and S6).

Random input activity was fixed at $\lambda_r = \lambda_{MSN} = 5Hz$ during (re-)learning phases to mimic *in vivo* MSN firing rate (Mahon et al., 2006). It was set to $\lambda_r = 4\lambda_{MSN} = 20Hz$ in the initial and maintenance phases to amplify forgetting within comparable timescales as learning (Methods; choosing λ_r =5Hz in these phases yields similar relative results, with higher final accuracies for similar maintenance phase durations).

Asymmetric anti-Hebbian STDP favors memory maintenance, whereas symmetric LTD allows accrued flexibility

Numerical experiments showed significant differences between the learning capability with symmetric or asymmetric anti-Hebbian STDP, particularly during the maintenance and relearning phases. Figure 7a reports changes in learning accuracy as a function of time throughout the four different phases and for various non-zero values of $A_{post-pre}$ (solid lines) or controls with $A_{pre-post} = 0$ (dotted lines). The final accuracies at the end of each phase were reported in Figure 7b

as a function of $A_{post-pre}$. Accuracy essentially remains at chance levels during the initial phase, as expected in the absence of reward. In the learning phase, rewards drove a rapid increase and stabilization of accuracy in all conditions. When rewards were no more provided in the maintenance phase, the accuracy gradually plummeted, with different dynamics that depended on $A_{post-pre}$. Finally, during the relearning phase, the system learns again previously memorized patterns when reward-LTP is applied anew, with again distinct kinetics depending on $A_{post-pre}$.

Learning phase. The rapid rise of accuracy in the learning phase highlights the ability of all tested networks to store patterns in the presence of pattern-specific reward-LTP (Fig. 7a). Heuristically, the cumulated increase in the synaptic weights induces on neurons associated with rewarded patterns increases the probability of the MSN to spike in response to the rewarded patterns compared to non-rewarded signals. This effect is counterbalanced by the presence of pre-post LTD, which leads to depression of the synaptic weights, and therefore favors an equilibrium where the synaptic weights are high enough to trigger the spike, but still remain bounded, and prevents firing after non-rewarded patterns. Indeed, if the MSN happened to spike after presentation of a non-rewarded pattern, the pre-post LTD induced a decrease in synaptic weights for neurons associated with the pattern, that will contribute to a decay of MSN firing probability. To highlight the crucial importance of the anti-Hebbian pre-post LTD in learning of patterns, we computed learning accuracies in the absence thereof ($A_{pre-post} = 0$; Fig. 7a). In this case, the absence of the pre-post LTD counterbalancing reward-LTP led to a continual growth of the synaptic weights, which resulted in the MSN spiking for non-rewarded patterns and therefore reducing learning accuracy.

Final accuracy at the end of the learning phase depended on $A_{post-pre}$ (Fig. 7b). Indeed, the system did not reach a fixed equilibrium during the learning phase, but stabilized at a stationary regime alternating multiple correct spiking responses interspersed by erroneous silences. This phenomenon

344 is due to the fact that the synaptic weights W associated with the pattern decreased progressively for
345 each accurate answer according to the superposition of reward-LTP and pre-post LTD. This decrease
346 persisted until the MSN stops firing, leading to a jump in W (since in the absence of a spike, only
347 reward-LTP is applied), after what the process starts afresh. As a consequence of this mechanism, we
348 expect that larger reward-LTP will allow for a larger fraction of experiments with spiking in response
349 to a rewarded signal, in turn increasing accuracy. However, when reward-LTP exceeds pre-post LTD,
350 an instability arises with a divergence of synaptic weights that would reduce accuracy, since it will
351 lead to non-specific spiking in response to non-reward signals. Accordingly, there exists a value of
352 $A_{post-pre}$ where the accuracy is maximal ($A_{post-pre} \approx 0.5$; Fig. 7b).
353 A combination of pre-post LTD and reward-LTP enabled discrimination of rewarded and non-
354 rewarded patterns, for all values of $A_{post-pre}$, with significant differences compared to the initial
355 phase.
356 *Maintenance phase.* The drop in accuracy during the maintenance phase was found to be faster for
357 symmetric anti-Hebbian learning than for asymmetric anti-Hebbian learning, with a significant
358 impact of the value of $A_{post-pre}$ on the accuracy at the end of the maintenance (Fig. 7b), as visible in
359 the distinct characteristic decay time $T_{maintenance}$ of accuracy (Fig. 7c1). Asymmetric STDP
360 allowed maintenance of higher accuracies for longer durations in the absence of rewards (Fig. 7c2).
361 Phenomenologically, symmetric LTD tends to induce a global depression in response to random
362 stimuli, which therefore can lead the MSN to stop firing to any patterns. The presence of LTP in
363 asymmetric anti-Hebbian STDP limits this phenomenon, since only pre-post pairings will lead to
364 LTD, which can be compensated by post-pre pairings. This more balanced response allows for a
365 more durable conservation of the relative W magnitudes in the absence of rewards. To show the
366 correlations with W dynamics, we computed the deviation from W values at the end of the learning
367 phase in terms of amplitude and change in orientation (Fig. 7c3). We observed that asymmetric

STDP ($A_{post-pre}$>0) led to a smaller deviation of the synaptic weights than for what was observed for symmetric STDP.

Relearning phase. After the maintenance phase, relearning started with resuming the delivery of a reward-LTP during presentation of rewarded patterns. While the system reaches similar accuracies as those obtained at the end of the learning phase, the kinetics of relearning appeared distinct for the different types of plasticity studied. Compared to symmetric anti-Hebbian STDP, relearning is faster for asymmetric anti-Hebbian STDP, as long as $A_{post-pre}$ is small enough to avoid runaway potentiation and instability (Fig. 7c4: quantification for the characteristic time of relearning $T_{relearning}$).

Influence of the pattern presentation rate. The maintenance phase can be seen as an unlearning of rewarded patterns: an MSN that will have maintained memory will fire in response to a rewarded pattern, and associated weights will decrease, leading the MSN to stop firing after enough presentations. This will lead to a drop in accuracy (Fig. 7d1), faster when patterns are presented more frequently (*i.e.* larger η_m), and found to be more rapid for symmetric anti-Hebbian STDP (Fig. 7d2).

The pattern presentation rate η_m chosen in the maintenance phase also played a significant role in the relearning phase, with higher rates of presentations leading to a slower relearning, potentially indicating a more dramatic deviation of W from their after-learning values (Fig. 7d3). However, the difference between symmetric and asymmetric STDP persisted for varying presentation rates η_m: even if the network loses its capacity to recall correctly the patterns, W may have conserved a signature of learnt patterns.

Influence of noise. While the network is able to learn in the presence of noise (Fig. S3), we expect the uncertainty and variability it creates on patterns to impact final accuracies and learning efficiency. In absence of noise ($\lambda_{MSN} = 0$), the system reached high accuracies in the learning phase, and the maximal value reached was identical for both symmetric and asymmetric STDP, that were conserved

throughout the maintenance phase in the absence of noise (Fig. S3). While the network was still learning in the presence of noise, increasing noise levels progressively impaired learning ability, and for Poisson noise with frequency $\lambda_{MSN} > 10Hz$, the network showed poor learning and maintenance abilities. Realistic noise levels on the same order as typical MSN firing rate is between these two regimes and allow both learning and maintenance, with different capacities for symmetric and asymmetric anti-Hebbian STDP as discussed above.

Robustness to $(P, N_p,\ N_{stim})$ *parameters of the model.* Beyond noise levels, all observations reported were robust to variations in the number of cortical neurons, patterns presented and number of stimulations with $(P, N_p,\ N_{stim})$ equals to $(10, 10, 3)$, $(10, 15, 5)$ and $(20, 30, 3)$ (Fig. S4).

Interpretation of the model on the impact of different STDP in DMS and DLS. For low stimulus frequency, the asymmetric anti-Hebbian STDP observed in DLS at 1Hz could support maintenance of stimulus associations for longer durations and relearning almost immediately previously learned associations. In contrast, the symmetric DMS-LTD leads to a faster erasure of associations, making the system available to learn new patterns. STDP elicited at striatopallidal DMS-MSN at 2.5Hz switched from symmetric LTD to asymmetric anti-Hebbian STDP. We postulate that with more frequent stimulations, striatopallidal MSNs adapt their behavior so as to store patterns for longer times than striatonigral MSNs.

Overall, when presenting patterns at a slow rate, DMS with symmetric LTD, is able to forget quickly, whereas asymmetric anti-Hebbian STDP maintains memory in DLS. This observation on a simplified model suggests a possible cellular basis for the various observations on the differential involvement of DMS and DLS in motor skill learning (Yin and Knowlton, 2006; Jin and Costa, 2015): both learn the task during the first trials, and then DMS disengages when habit learning mediated by the DLS takes over initial phases of motor training or goal-directed learning.

DISCUSSION

To explore how the striatum is able to achieve distinct learning modalities, from goal-directed behavior to maintaining habits, we explored long-term plasticity in DMS and DLS. We found distinct anti-Hebbian STDP: symmetric in DMS and asymmetric in DLS. Hebbian and anti-Hebbian STDP have been reported in the dorsal striatum depending whether GABAergic transmission inhibitors (Paillé et al., 2013; Valtcheva et al., 2017) were applied (Hebbian STDP (Pawlak and Kerr, 2008; Shen et al., 2008)) or not (anti-Hebbian STDP (Fino et al., 2005; Fino et al., 2010; Mendes et al., 2020)). These studies targeted DLS-MSNs, except Shen et al. (2008) where MSNs were recorded indifferently in DLS and DMS. *In vivo* recordings confirmed the anti-Hebbian polarity of striatal STDP (Schulz et al., 2010; Moera-Herreras et al., 2019; but see Fisher et al., 2017). With increasing cortical activity, plasticity followed opposite polarity in DMS and DLS, with LTD and LTP dominance, respectively. Another difference between DMS- and DLS-STDP upon increasing $F_{pairings}$, is that plasticity expression domain remained wide in DMS, whereas it was broadened in DLS. Thus, DLS appears highly sensitive to increasing $F_{pairings}$ with a profound remodeling, namely a double invasion of the plasticity domain by LTP: (i) LTP which was restricted to narrow Δt_{STDP} ($-30<\Delta t_{STDP}<0$ms) progressively invades the pre-post side, but also (ii) gains ground for larger Δt_{STDP} values ($-100<\Delta t_{STDP}<+100$ms). Regarding DMS, a similar plasticity map is observed for $F_{pairings}$ of 1-5Hz, with LTD expressed in the whole Δt_{STDP} range ($-100<\Delta t_{STDP}<100$ms), with the notable exception of striatopallidal MSNs exhibiting an opposite transient polarity, LTP, for a specific $F_{pairings}$ (2.5Hz). Although there is less overall sensitivity to $F_{pairings}$ in DMS than in DLS, there is a specific activity regime that allows for LTP expression in striatopallidal DMS-MSNs. The dominance of LTD *vs* LTP in DMS and DLS, respectively, is not restricted to time-coding paradigm but is also observed under rate coding activity (HFS), and therefore characterize each striatal sub-compartment. This is in line with studies, using other cell conditioning paradigms, reported distinct plasticity in DMS and

DLS (Partridge et al., 2000; Shan et al., 2014; Hawes et al., 2015; Munoz et al., 2020). In DMS, anti-Hebbian LTD occurred in a larger temporal window than in DLS, suggesting that noisy signals in terms of jitter and correlation can induce LTD in DMS, while no plasticity would be triggered in DLS. This feature could participate to behavioral flexibility of DMS (Ragozzino et al., 2002a; Ragozzino et al., 2007) and its involvement at early stages of learning where correlated inputs can be more jittered than subsequently upon repetitions.

Various experimental conditions including the angle of brain slicing, location of the stimulation electrode (cortex, corpus callosum or striatum), pharmacological agents, and/or protocols (HFS, LFS, theta-burst, STDP) can account for the great variety of observed plasticity (Lovinger, 2010; Cerovic et al., 2013) and subsequent controversies. A key element governing the polarity of STDP can also be the recruited neuromodulator(s) (Foncelle et al., 2018; Brzosko et al., 2019). Indeed, stimulation within the cortex, corpus callosum or striatum is not recruiting similarly neuromodulators. Specific recruitment at $F_{pairings}$=2.5Hz of dopaminergic fibers or cholinergic neurons could account for distinct plasticity at striatonigral and striatopallidal DMS-MSNs. Supporting this hypothesis, it was shown that dopamine and adenosine orientate *in vivo* STDP polarity in the DLS (Fisher et al., 2017), and that pharmacological inhibition of muscarinic receptors or opto-activation of adenosine A_2A receptors in DMS suppresses behavioral flexibility (Ragozzino et al., 2002b; Li et al., 2016). A $F_{pairings}$ of 2.5Hz could be reminiscent of the upper range of thalamocortical delta waves occurring during sleep (Steriade et al., 1993) and associated with weakening or forgetting of memories (Kim et al., 2019).

Anti-Hebbian STDP, mainly reported in Purkinje-like cells, has been proposed to promote storage and retrieval of a temporally structured negative image of prior sensory stimuli (Roberts and Bell, 2000; Roberts and Leen, 2010; Requarth and Sawtell, 2011). Our model shows that asymmetric anti-Hebbian STDP leads to the maintenance of learned patterns, whereas symmetric LTD causes a rapid

decrease in memory performance in the absence of reward. Similar to Roberts and Bell (2000), the LTP/LTD alternation in asymmetric anti-Hebbian STDP, if correctly tuned, forces the synaptic weights to retain some information on previously learned patterns. On the contrary, with symmetric LTD, the synaptic weights converge to zero because they are only subject to depression, leaving the system fresh to construct new associations and identify novel stimuli. Our study shows that DLS-MSNs exhibit asymmetric anti-Hebbian STDP, consistent with DLS role in habit behavior, where rewards are no longer presented. Conversely, we hypothesize that thanks to symmetric anti-Hebbian STDP in DMS-MSNs, DMS should be able to adapt quickly between different action-outcome associations and therefore forget rapidly previous information. Our model shows that this role is consistent with anti-Hebbian symmetric STDP. This is in line with the role of DMS which is essential for behavioral flexibility such as strategy-shifting or reversal learning (Bonnavion et al., 2019; Ragozzino et al., 2002a; Ragozzino, 2007).

It is assumed that DMS is involved in early phases of learning and goal-directed behavior, whereas DLS is involved in late phases of learning and habit acquisition (Graybiel and Grafton, 2015; Jin and Costa, 2015). In a motor skill learning, we found that DMS (striatonigral and striatopallidal MSNs), via DMS plasticity, is involved in early phases of learning and is disengaging upon reiteration and acquisition of the motor skill. In line with previous studies (Balleine and O'Doherty, 2010; Thorn et al., 2010; Kimchi et al., 2009; Stalnaker et al., 2010; Kupferschmidt et al., 2017; Bergstrom et al., 2018), DLS plasticity can also be engaged from the early phases of motor skill training. However, we found that this concerns the striatonigral DLS-MSNs and not striatopallidal DLS-MSNs, these later being mobilized only during the late stages of learning. Moreover, striatonigral DLS-MSN plasticity follows DMS-MSNs engagement since LTP can be induced again during the late stages of learning. The fact that part of the DLS-MSNs (striatonigral DLS-MSNs) are mobilized during early training while striatopallidal DLS-MSNs take over at later stages agrees with the notion of sequential learning

implying part of the DLS at early learning stages (Bergstrom et al., 2018) and with the antagonistic control operated by striatonigral and striatopallidal DLS-MSNs during consolidation phase (Smith et al., 2021). Interestingly, when striatopallidal DMS-MSNs express asymmetric anti-Hebbian STDP, similar to DLS-STDP, this plasticity exhibits the same sequence of engagement/disengagement than in DLS. On the contrary, when symmetric anti-Hebbian STDP is expressed in DMS, this later evolves oppositely to that of the DLS. In DMS, we observed different plasticity depending on striatonigral and striatopallidal MSNs but with similar timing of engagement, whereas the reverse situation was observed in DLS with similar plasticity in MSNs but different timing of engagement of striatonigral and striatopallidal MSNs. The engagement of striatopallidal DLS-MSN, but not direct DLS-MSNs, in term of plasticity during extensive training agrees with Yin et al. (2009) showing an increased *ex vivo* corticostriatal transmission specifically in the striatopallidal MSNs and not in striatonigral MSNs after Rotarod extensive training. For 2.5Hz STDP, striatopallidal DMS-MSNs display asymmetric anti-Hebbian STDP, *i.e.* similar STDP than in DLS-MSNs, and exhibit similar engagement/disengagement than striatonigral DLS-MSNs. It indicates, according to model predictions, that striatonigral and striatopallidal DLS-MSNs favor the maintenance of learned patterns during early and late stages of learning, respectively. It also suggests that in DMS half of the MSNs, *i.e.* striatonigral DMS-MSNs showing a symmetric LTD, are more flexible than the other half, *i.e.* striatopallidal DMS-MSNs exhibiting asymmetric anti-Hebbian STDP. Our findings support the ideas that DMS and DLS display distinct plasticities, are both engaged from the early training phases, with a progressive disengagement of DMS, theoretically consistent with the plasticity they express.

Limitations of the study

The observations, about the LTD *vs* LTP dominance in DMS *vs* DLS or the specific activity regime for opposite plasticity in striatopallidal DMS-MSNs, remain to be examined in other corticostriatal domains and also *in vivo* with the tracking of plasticity across days in behaving animals (Xiong et al., 2015). Because *ex vivo* (post-Rotarod) results and conclusions are limited to the form of STDP tested here, post-pre pairings at 2.5Hz, they need to be extended to other frequencies and pre-post pairings. The specific properties of striatonigral and striatatopallidal MSNs in DMS and DLS (Alegre-Cortés et al., 2021) may also contribute to the variety of learning abilities. Therefore, future work will be needed to characterize and integrate such differences in models and study how the network combines striatonigral and striatopallidal pathways endowed with distinct plasticity rules for procedural learning.

Acknowledgments

This work was supported by LabEx MemoLife (Paris, France); Ecole Normale Supérieure (Paris, France), ANR Mopla, Fondation Bettencourt Schueller, Grant 2022-A0100612385 (HPC resources from GENCI-IDRIS), Zhejiang Provincial Natural Science Foundation of China (LR19C090001) and the National Natural Science Foundation of China to Y.C. (31922031).

We thank Sonia Garel (Ecole Normale Supérieure, Paris, France) for kindly providing Drd2-eGFP mice.

Author contributions

Conceptualization: LV, JT, GV

Investigation: SP, YC, GV, EP, AM, ZZ

Supervision: LV, JT

Writing—original draft: LV, JT, GV

Writing—review & editing: LV, JT, GV, YC

Declaration of interests

The authors declare no competing interests.

Figures

Figure 1: Distinct anti-Hebbian STDP profiles in DLS and DMS.

(**a**) STDP pairings at 1 Hz with $\Delta t_{STDP}<0$ and $\Delta t_{STDP}>0$ referring to post-pre (**a1**) and pre-post (**a2**) pairings. Pre- and postsynaptic stimulations were applied either in the sensorimotor (**b**) or the associative (**c**) cortical and striatal areas. (**b**) DLS-STDP displays an asymmetric anti-Hebbian polarity in rats. (**b1**) Experimental setup. (**b2** and **b3**) Averaged time-courses of (**b2**) LTP induced by 100 post-pre pairings (n=16) and (**b3**) LTD induced by 100 pre-post pairings (n=14). (**c**) DMS-STDP displays symmetric anti-Hebbian polarity. (**c1**) Experimental setup. (**c2** and **c3**) Averaged time-courses of LTD induced by (**c2**) 100 post-pre pairings (n=11) and (**c3**) 100 pre-post pairings (n=16).

Plasticity values and statistics: **Table S1**.

Bar graphs represent the average of all STDP experiments and each point represents the percentage of change in EPSC amplitude at 50-60 min after STDP pairings in a single STDP experiment. Insets correspond to the average EPSC amplitude at baseline (black) and at 50-60 min after STDP pairings (grey). Error bars represent the SEM. ****: $p<0.0001$ by one sample t test.

Figure 2: Signaling pathways involved in DMS- and DLS-STDP.

(**a**) Post-pre DLS-LTP was NMDAR-mediated since prevented by D-AP5 (50µM) (n=9) and (**b**) pre-post LTD was CB_1R-mediated since prevented by AM251 (3µM) (n=5). (**c**) Post-pre DMS-LTD was not CB_1R-mediated, because left unaffected by AM251 (n=8), but was NMDAR- and T-type VSCC-mediated because prevented by D-AP5 (n=7) or mibefradil (20µM) (n=8). (**d**) Pre-post DMS-LTD was CB_1R-mediated since prevented with AM251 (n=5).

Plasticity values and statistics: **Table S1**.

Error bars represent the SEM. ****: $p<0.0001$; ns: not significant by one sample t test.

Figure 3: LTD and LTP dominates in DMS and DLS, respectively, when pairing frequency scales up.

(**a**-**d**) DLS-STDP by varying Δt_{STDP} ($-100<\Delta t_{STDP}<100$ms) and $F_{pairings}$ (1, 2.5 and 5Hz) (n=130 DLS-MSNs). Summary graphs of STDP in relation with Δt_{STDP} showed at 1Hz (**a**) an asymmetric anti-Hebbian STDP in a restricted time window ($-30<\Delta t_{STDP}<30$ms) (n=42), at 2.5Hz (**b**) a widening of the temporal window of STDP expression and LTP being also induced for short pre-post pairings (n=36), and at 5Hz (**c**) dominance of LTP (n=52). (**d**) STDP profiles illustrates the progressive

dominance of LTP and the evolution of the asymmetric anti-Hebbian STDP into symmetric Hebbian STDP with increasing $F_{pairings}$. (**e**) Averaged time-courses of HFS-LTP (p<0.0001, n=6). (**f**-**i**) DMS-STDP by varying Δt_{STDP} and $F_{pairings}$ (n=125 DMS-MSNs). Summary graphs of STDP showed at 1Hz (**f**) symmetric anti-Hebbian STDP in a broad time window (-100<Δt_{STDP}<100ms) (n=39), at 2.5Hz (**g**) mainly LTD except for narrow (-30<Δt_{STDP}<0ms) post-pre pairings for which half of the MSNs exhibited LTD while the other half displayed LTP (n=59), and at 5Hz (**h**) dominance of LTD (n=27). (**i**) STDP profiles show the dominance of LTD and the unique plasticity map at 2.5Hz for which LTP emerged for post-pre pairings. (**j**) Averaged time-courses of HFS-LTD (p<0.0001, n=7).

STDP values and statistics: **Table S2**.

Panels **d** and **i**, linear interpolation with a Savitzky-Golay filter. Each point: % of change in EPSC amplitude 50-60 min after STDP pairings. Error bars represent the SEM. *: p<0.05, **: p<0.01, ***: p<0.001, ****: p<0.0001 by one sample t test, Mann-Whitney test or Kolmogorov-Smirnov test.

Figure 4: D_1^+ and D_1^- DMS-MSNs express opposite plasticity at $F_{pairings}$=2.5 Hz.

(**a**-**d**) DMS-STDP for post-pre pairings at $F_{pairings}$=2.5Hz. (**a**-**c**) Whole-cell recordings in $D1^+$ and $D1^-$ MSN pairs (n=8 cell pairs) in Drd1a-eGFP mice. (**a**) Experimental set-up (**b**) Example of a D_1^+ and D_1^- MSN pair recording: LTD (Δt_{STDP}=-19 ms) in a D_1^+ and LTP (Δt_{STDP}=16ms) in a D_1^- DMS-MSN. (**c**) Averaged time-courses of MSN pair recordings showing LTD in D_1^+ MSNs (7/8 LTD), and LTP in D_1^- MSNs (8/8 LTP); Right: summary graph with synaptic weights determined 40-50 minutes after pairings. (**d**) LTD and LTP in, respectively, D_1^+ (n=7, p=0.1255) and D_1^- (n=6, p=0.5884) MSN are NMDAR-mediated since prevented by D-AP5. (**e**) At $F_{pairings}$=1Hz, LTD was the unique form of plasticity at in D_1^+ and D_1^- DMS-MSNs (p=0.0005, n=5, and p<0.0001, n=6).

Plasticity values and statistics: **Table S3**.

Error bars represent the SEM. ***: p<0.001; ****: p<0.0001; ns: not significant by one sample t test.

Figure 5: Region-specific involvement of STDP during motor skill learning

(**a**) Experimental set-up: Rotarod and *ex vivo* STDP. Drd2-EGFP mice were divided in habituated (n=9 mice), early-trained (n=8) and late-trained (n=10) groups. Motor skill learning of mice were evaluated with the latency to fall from the Rotarod. (**b**-**d**) Occlusion experiments in DMS and averaged time-courses of DMS-MSNs recorded in habituated (**b**), early-trained (**c**) and late-trained (**d**) mice after post-pre pairings at 2.5Hz (Δt_{STDP}~-15ms). (**b**) Habituated mice: LTD (p=0.0002, n=7)

and LTP (p<0.0001, n=6) in D_2^- (**b1**) and D_2^+ (**b2**) DMS-MSNs, respectively. (**c**) Early-trained mice: no synaptic plasticity detected in D_2^- (p=0.1566, n=6) (**c1**) and D_2^+ (p=0.5850, n=7) (**c2**) MSNs. (**d**) Late-trained mice: LTD (p<0.0001, n=10) and LTP (p<0.0001, n=8) in D_2^- (**d1**) and D_2^+ (**d2**) DMS-MSNs, respectively, similarly to the habituated group. (**e**-**g**) Occlusion experiments in DLS-MSNs. (**e**) Habituated mice: LTP in D_2^- (p<0.0001, n=7) (**e1**) and D_2^+ (p<0.0001, n=7) (**e2**) DLS-MSNs. (**f**) Early-trained mice: no synaptic plasticity detected in D_2^- (p=0.2795, n=9) (**f1**) while LTP was induced in D_2^+ (p<0.0001, n=7) (**f2**) DLS-MSNs. (**g**) Late-trained mice: a reverse plasticity map was observed when compared to (**f**): LTP (p<0.0001, n=10) in D_2^- (**g1**) and no plasticity (p=0.4223, n=7) in D_2^+ (**f2**) DLS-MSNs. D_2^- DLS-MSNs are engaged at early stages, while D_2^+ DLS-MSNs remain available for plasticity, and conversely at late stages. D_2^- and D_2^+ DMS-MSNs show the same behavior: they are engaged in the early-trained stages and show a disengagement at late-trained stages.

Plasticity values and statistics: **Table S4**.

Error bars represent the SEM. ***: p<0.001; ***: p<0.0001; ns: not significant by one sample t test.

Figure 6: A pattern recognition task to test the learning, maintenance and relearning in a computational model of the striatal network.

Striatal network with $P = 4$ cortical neurons (green), a random input neuron with rate λ_r (yellow) and one MSN, represented by its membrane potential V (brown). Two mechanisms of synaptic plasticity are considered in the dynamics of the synaptic weight W: anti-Hebbian STDP and LTP related to the reward signal (reward-LTP). Anti-Hebbian STDP is modeled using exponential kernels, with different values for $A_{post-pre}$ and $A_{pre-post} = -1$. Example of the learning task (left), separated in four phases (**Table S5**) with four iterations in each phase. $Np = 2$ patterns A and B are presented to the network, with A being non-rewarded $(-)$ and B rewarded $(+)$. $\emptyset$: iteration with no pattern presentation. Spiking activity of the cortical neurons (green: pattern spikes; grey: random spikes) and the random input neuron are represented along with the MSN membrane potential V.

Figure 7: Influence of symmetric and asymmetric anti-Hebbian STDP on learning, maintenance and relearning of patterns in a striatal network.

(**a**) Learning dynamics for $P = 10$ input neurons, $N_{stim} = 3$ stimulations by pattern, $N_p = 15$ patterns and the proportion of pattern presentation in the initial/maintenance phases $\eta_m = 0$. Time-evolution of the learning accuracy through an initialization, learning, maintenance and relearning phases, for distinct anti-Hebbian plasticities (blue: symmetric, brown: no post-pre learning, orange: asymmetric). Averaged simulations (plain lines) and controls with $A_{pre-post} = 0$ (dotted lines). (**b**) Accuracy at the end of each phase as a function of the type of plasticity (where $A_{post-pre}$ parametrizes the plasticity). (**c1**) Examples of fits obtained from a set of 20 simulations for the accuracy during the maintenance phase. Averaged simulations (for 20 simulations, open circles), associated fit (plain lines), tangent at origin (dashed) and fitted final accuracy (dotted). Below: sets of values for $T_{maintenance}$ obtained with this method for $A_{post-pre} = 0$. (**c2**) $T_{maintenance}$ as a function of $A_{post-pre}$. (**c3**) Weight similarity measures $d_2(W)$ and $sp(W)$ as a function of $A_{post-pre}$. (**c4**) $T_{relearning}$ as a function of $A_{post-pre}$. (**d**) Dependency of $T_{maintenance}$ (**d1**), accuracy at the end of the maintenance phase (**d2**) and $T_{relearning}$(**d3**) on the type of plasticity $A_{post-pre}$ and stimulus presentation frequency η_m.

Mean results computed over 200 simulations (**a**). Mean of results over 10 sets of 20 simulations with errors bars $\pm SD/2$ (**b-d**). t-test *: $p < 0.05$, **: $p < 0.005$, ***: $p < 0.0005$.

STAR METHODS TEXT

RESSOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Laurent Venance (laurent.venance@college-de-france.fr)

Material availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the Lead Contact upon request.

All original code has been deposited at https://github.com/gvignoud/striatalLearning and is publicly available as the date of publication. DOIs are listed in the key resources table.

- Any additional information required to reanalyze the data reported in this work paper is available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All experiments were performed in accordance with the guidelines of the local animal welfare committee and the EU (directive 2010/63/EU). Every precaution was taken to minimize stress and the number of animals used in each series of experiments. Sprague Dawley rats $P_{25\text{-}35}$ (Charles River, L'Arbresle, France) and C57BL/6 Drd1a-GFP and Drd2-EGFP (heterozygous) mice $P_{25\text{-}50}$ were used for brain slice electrophysiology, and behavior and *ex vivo* electrophysiology, respectively. Animals of both sexes were used. Animals were housed in standard 12-hour light/dark cycles and food and water were available *ad libitum*.

METHOD DETAILS

Cortico-DLS and -DMS brain slices

We used horizontal and para-sagittal (33° angle) brain slice preparations to stimulate specifically cortical inputs (layer 5) originating from the sensorimotor S2 or associative cortices CG2, and projecting to the DLS or DMS, respectively (Fino et al., 2018). Brain slices (300 and 270 µm-thick for rats and mice, respectively) were prepared with vibrating blade microtome (VT1200S, Leica Microsystems, Nussloch, Germany). The brains were sliced in an ice-cold cutting solution (in mM: 125 NaCl, 2.5 KCl, 25 glucose 25 $NaHCO_3$, 1.25 NaH_2PO_4, 2 $CaCl_2$, 1 $MgCl_2$, 1 pyruvic acid) through which 95% O_2/5% CO_2 was bubbled. The slices were transferred to the same solution at 34°C for 45-60 minutes and then to room temperature, before electrophysiology.

Single and dual patch-clamp recordings

For whole-cell recordings, borosilicate glass pipettes of 6-8MΩ resistance were filled with (in mM): 122 K-gluconate, 13 KCl, 10 HEPES, 10 phosphocreatine, 4 Mg-ATP, 0.3 Na-GTP, 0.3 EGTA (adjusted to pH 7.35 with KOH). The composition of the extracellular solution was (mM): 125 NaCl, 2.5 KCl, 25 glucose, 25 $NaHCO_3$, 1.25 NaH_2PO_4, 2 $CaCl_2$, 1 $MgCl_2$, 10 µM pyruvic acid through which 95% O_2 and 5% CO_2 was bubbled. Signals were amplified using EPC10-2 amplifiers (HEKA Elektronik, Lambrecht, Germany). All recordings were performed at 34°C, using a temperature control system (Bath-controller V, Luigs&Neumann, Ratingen, Germany). Recordings were sampled at 10kHz, using the Patchmaster v2x32 program (HEKA Elektronik).

689 **Spike-timing-dependent plasticity protocols**

690 Electrical stimulations were performed with concentric bipolar electrodes (Phymep, Paris, France)
691 placed in layer 5 of S2 or CG2 cortex. Electrical stimulations were monophasic, at constant current
692 (ISO-Flex stimulator and Master 9, A.M.P.I., Jerusalem, Israel) and adjusted to evoke 100-300pA
693 EPSCs. Repetitive control stimuli were applied at 0.1Hz. STDP protocols consisted of 100 pairings
694 of pre- and postsynaptic stimulations (at 1, 2.5 or 5 Hz) separated by a specific Δt_{STDP} kept constant
695 during a given STDP pairing; $\Delta t_{STDP}<0$ indicating that postsynaptic stimulation preceded presynaptic
696 stimulation, post-pre pairings, and $\Delta t_{STDP}>0$ indicating that presynaptic stimulation preceded
697 postsynaptic activation, pre-post pairings. Presynaptic stimulations corresponded to cortical
698 stimulations and the postsynaptic stimulation of an action potential evoked by a depolarizing current
699 step (30ms duration) in the recorded MSN. Recordings were made over a period of 10 minutes at
700 baseline, and for at least 50 minutes after the STDP protocols; long-term changes in synaptic weight
701 were measured in the last 10 minutes. Experiments were excluded if the mean input resistance (Ri)
702 varied by more than 20% through the experiment.

703

704 **High frequency stimulation protocols**

705 High-frequency stimulation (HFS) consisted in a single train of cortical stimuli at 100 Hz for 1
706 second and the depolarization of the postsynaptic element from its holding membrane potential (-
707 75mV) to 0 mV was coincident with the presynaptic stimulation.

708

709 **Chemicals**

710 DL-2-amino-5-phosphono-pentanoic acid (D-AP5, 50µM) (Tocris, Ellisville, MO, USA) and
711 (1S,2S)-2-[2-[[3-(1H-benzimidazol-2yl)propyl]methylamino]ethyl]-6-fluoro-1,2,3,4-tetrahydro-1-(1-

712 methylethyl)-2-naphthalenyl methoxyacetoacetate dihydrochloride (Mibefradil; 20μM) were
713 dissolved directly in the extracellular solution. N-(piperidin-1-yl)-5-(4-iodophenyl)-1-(2,4-
714 dichlorophenyl)-4-methyl-1H-pyrazole-3-carboxamide (AM251; 3μM) was dissolved in ethanol and
715 added to the external solution, such that the final concentration of ethanol was 0.01-0.1%.
716
717 **Behavioral task: accelerating Rotarod**
718 Mice (P58-62) were trained on the accelerating Rotarod (Ugo Basile, Gemonio, Italy) in a single
719 session of 10 trials per day for 1 (early-trained group) or 7 consecutive (late-trained group) days.
720 Each trial consisted of a continuous acceleration from 4 to 40 rpm over 300 seconds. Before the first
721 trial of the first day, mice were placed on the rod for 30 seconds without acceleration (4 rpm). The
722 control animal group consisted in mice subjected to the habituation phase, consisting of a single
723 session with non-accelerating Rotarod (4 rpm) for 300 seconds. The latency to fall off the rod was
724 measured. Between each trial, mice were placed in their homecage to rest for 15 minutes.
725
726 ***Ex vivo* occlusion experiments**
727 Mice subjected to Rotarod were divided into three groups for occlusion experiments: (1) the
728 habituated mice, subjected to non-accelerating (4 rpm) Rotarod for 300 seconds, (2) the early-trained
729 mice, subjected to accelerating Rotarod for one day (the first one) with 10 trials, and (3) the late-
730 trained mice, subjected to accelerating Rotarod for seven days (10 trials/day). After Rotarod task
731 completion, mice were left in their homecages for 24 hours and then sacrificed for *ex vivo* patch-
732 clamp recordings. Plasticity occlusion was tested in the following STDP paradigm: 100 pairings at
733 2.5Hz with $\Delta t_{STDP=}$-15ms. For each mouse, one brain hemisphere was sliced to obtain cortico-DLS
734 slices and the other hemisphere was used for cortico-DMS slices. Then, cortico-DLS and -DMS
735 STDP were assessed in the same mice in most of the cases.

736

737 **Mathematical models**

738 **Neuronal network model.** To simulate the impact of plasticity on learning, we built a simple
739 neuronal network model that includes P cortical neurons serving as input neurons to one MSN. The
740 MSN integrated cortical and external input (section below) and fired when hitting a threshold,
741 according to the classical leaky integrate-and-fire model (Burkitt, 2006; Gerstner et al., 2014).
742 Between two spikes, the membrane potential V of the neuron satisfies a linear differential equation:

743 $$\tau \frac{dV}{dt} = -(V(t) - V_{eq}) + R\, I(t) \qquad (1)$$

744 Spikes were emitted when the voltage exceeded a threshold V_{th}, at which time the voltage of the
745 neuron was instantaneously reset to V_{eq} and resumed input integration. We set $\tau = 16ms$, $V_{eq} =$
746 $-80mV$ and $R = 80M\Omega$, $V_{th} = -45mV$, and the reset potential was the resting potential V_{eq} (Yim et
747 al., 2011).

748 **Connectivity and input to the MSN.** The input $I(t)$ received by the MSN is the superposition of the
749 input received from P cortical neurons, noted $I_{stim}(t)$, and an external (to the network) input $I_{ext}(t)$
750 modeled as a Poisson process with rate λ_r:

751 $$I(t) = I_{stim}(t) + I_{ext}(t). \qquad (2)$$

752 Spikes from cortical neurons and the external source induce instantaneous jumps in the MSN
753 membrane potential. Jumps associated with cortical sources have amplitudes that vary through
754 plasticity mechanisms (next section). These amplitudes are modeled through the collection of
755 synaptic weights $W(t) = \big(W_i(t)\big)_{1\le i\le P}$. Denoting t_i^k the k-th spike time of input neuron i and δ the
756 Dirac mass, we have

757 $$I_{stim}(t) = \tau \sum_{1\le i\le P} \sum_{t_i^k \le t} W_i\big(t_i^k -\big)\delta(t - t_i^k) \qquad (3)$$

where we noted, for a function f being potentially discontinuous at time t, $f(t-)$ the value reached immediately before the jump.

Contrasting with the network input described above whose synaptic weights are allowed to vary in time according to plasticity rules (next section), the external input is assumed to induce jumps of fixed amplitude $W_{ext} = 1nA$ (high enough to evoke spiking activity in MSN):

$$I_{ext}(t) = \tau W_{ext} \sum_{t_{ext}^k \leq t} \delta(t - t_{ext}^k) \quad (4)$$

where $\left(t_{ext}^k\right)_{k\geq 0}$ denotes the sequence of external spike times, which have exponentially distributed inter-spike intervals.

The factor τ needs to be added in both currents expressions because we chose to use a simple model of synaptic inputs, where spikes induce a dirac of activity. This scaling was needed to relate synaptic weight W to EPSC amplitudes measured in experiments. The membrane potential has the following expression, between spikes of the postsynaptic neuron,

$$V(t) = V_{eq} + R \sum_{1\leq i\leq P} \sum_{t_i^k \leq t} W_i\left(t_i^k -\right) e^{-(t-t_i^k)/\tau} + RW_{ext} \sum_{t_{ext}^k \leq t} e^{-(t-t_{ext}^k)/\tau} \quad (5)$$

Cortico-striatal plasticity. We implemented a pair-based model of STDP, where synaptic weights W were updated after each spike (all-to-all implementation (Morrison et al., 2008)), according to the spike timing relative to all previous spikes of the other neuron. In detail:

— if the MSN spikes at time t_{post} (postsynaptic spike), then all weights are updated. Noting $t_{pre,i}$ the previous spikes of cortical neuron i, W_i is updated according to:

$$W_i\left(t_{post}\right) = W_i\left(t_{post} -\right) + \varepsilon \sum_{t_{pre,i}\leq t_{post}} \Phi(t_{post} - t_{pre,i}) \quad (6)$$

where ε denotes the plasticity rate, chosen in our simulations as $\varepsilon = 0.02$.

— if presynaptic cortical neuron $i \in \{1,\cdots,P\}$ spikes at time $t_{pre,i}$, noting t_{post} the times of the MSN spikes, then W_i is updated as:

$$W_i(t_{pre,i}) = W_i(t_{pre,i}-) + \varepsilon \sum_{t_{post} \le t_{pre,i}} \Phi(t_{post} - t_{pre,i}) \qquad (7)$$

Denoting $\Delta t = t_{post} - t_{pre}$ the timing between the presynaptic (cortical) spike and the postsynaptic (MSN) spike, we use an exponential STDP kernel (Kusmierz et al., 2017):

$$\Phi(\Delta t) = \begin{cases} A_{post-pre} \exp\left(\frac{\Delta t}{\tau_s}\right) & if\ \Delta t < 0, \\ A_{pre-post} \exp\left(-\frac{\Delta t}{\tau_s}\right) & if\ \Delta t > 0, \end{cases} \qquad (8)$$

with $\tau_s = 20\ ms$.

Consistent with the anti-Hebbian form of the corticostriatal STDP (Paillé et al., 2013; Valtcheva et al., 2017), we consider $A_{pre-post} = -1$, corresponding to synaptic depression subsequent to a pre-post paired stimulation. The sign of $A_{post-pre}$ allows distinguishing between symmetric anti-Hebbian STDP ($A_{post-pre} \le 0$) reported at DMS corticostriatal synapses from asymmetric anti-Hebbian STDP ($A_{post-pre} \ge 0$) reported at DLS corticostriatal synapses. Here, we focused on the influence of $A_{post-pre}$ on learning and relearning.

During a learning task (next section), the system is presented with a succession of cortical patterns. Each pattern corresponds to a temporal window of a fixed duration ($100ms$), where a subset of N_{stim} cortical neurons spike at time $t_{offset} = 50ms$. Two types of noise are modeled at the level of a single cortical neuron. First, each neuron involved in the pattern spikes at a time normally distributed with mean t_{offset} and standard deviation $\tau_p = 0.2ms$, modeling variability of the spike times. Second, cortical spikes unrelated to the pattern are added through Poisson spikes with rate λ_r/P, representing the random firing of the cortical neuron. Moreover, the influence of external inputs is

modeled at the level of the postsynaptic neuron, directly with the spikes of the random input presented above.

A pattern can either be rewarded or not through a simple additive mechanism. If a pattern is rewarded, then each time a presynaptic neuron i fires during the pattern (even if it is noise), its associated synaptic weight gets potentiated, following,

$$\Delta W_i = dt\varepsilon A_{reward} > 0. \quad (9)$$

If the pattern is not rewarded, the synaptic weight is not modified.

Detailed models, in particular three-factor learning rules (Foncelle et al., 2018; Gerstner et al., 2018; Kusmierz et al., 2017), are thus approximated here by the presence of a simple reward signal consisting in the potentiation of the synaptic weight of all presynaptic neurons that spiked during the pattern (both those involved in the pattern and those associated with noisy inputs).

A framework for corticostriatal plasticity was developed along with the use of dopamine-dependent STDP curves (Gurney et al., 2015), but only focused on Hebbian STDP. Following the same principles about the role of dopamine in the reward system, we chose to fix the STDP curves and modeled the reward influence through an additive potentiation as used in most existing models of anti-Hebbian STDP (Roberts and Bell, 2000; Williams et al., 2003; Rumsey and Abbott, 2004).

Eventually, synaptic weights are clipped within a realist range $[w_{min}, w_{max}] = [0.,2.]nA$. The initial synaptic weights are drawn from a uniform distribution on $[0.,0.05]nA$.

Learning with anti-Hebbian STDP rules. To characterize the capacity of learning associated with each STDP forms, we defined a fixed set of N_p cortical patterns. The system was presented either with a randomly chosen pattern of correlated cortical activity (from the set N_p patterns) with probability η, or with probability $1-\eta$ only with noise. Among the set of N_p patterns, a fixed subset was chosen to be rewarded (rewarded patterns were randomly chosen among all patterns, each

pattern having a probability $1/2$ to be rewarded). A rewarded pattern was deemed learnt if the MSN fired in response to the presentation of the pattern. Moreover, non-rewarded patterns should not elicit any spike.

The accuracy of the learning process was estimated through the averaged numbers of correct responses:

$$\text{Accuracy} = \frac{1}{N_p}\sum_{1\leq k\leq N_p}[r_k\sigma_k + (1-r_k)(1-\sigma_k)] \quad (10)$$

where $r_k = 1$ if k is a rewarded pattern and 0 otherwise, $\sigma_k = 1$ if the MSN spiked and 0 otherwise.

Each simulation emulated learning throughout four phases (Table S5), all of which including STDP and differing in the frequency of pattern presentation and presence of rewards:

(a) The initial phase of spontaneous activity, where patterns are presented randomly ($\eta = \eta_m$) and with noise. This phase is useful to avoid transient effects due to the initialization by reaching a realistic synaptic weight regime based on the plasticity rule. Noise was set to $\lambda_r = 4\lambda_{MSN} = 20Hz$.

(b) The learning phase during which neurons display spontaneous random activity with pattern presented at each iteration ($\eta = 1$), and Poisson noise with intensity $\lambda_r = \lambda_{MSN} = 5Hz$ consistent with *in vivo* MSN firing (Mahon et al., 2006). The reward signal was present and potentiated all synapses of presynaptic neurons active during a rewarded pattern. This phase emulates learning, as MSN learns to discriminate patterns by spiking in response to rewarded patterns and not spiking in response to non-rewarded patterns.

(c) The maintenance phase models spontaneous activity with $\lambda_r = 4\lambda_{MSN} = 20Hz$ and random presentations of patterns ($\eta = \eta_m$) in the absence of rewards, allowing to evaluate the system's ability to sustain a discrimination between learnt patterns. We chose to take $\lambda_r = 4\lambda_{MSN}$ to shorten our simulations and speed up the decrease of memory. All results are still true for $\lambda_r = \lambda_{MSN}$, except that memory is maintained for longer times than our simulations permitted.

(d) The relearning phase, with the same parameters as the learning phase (a), is used to measure the system ability to learn again patterns, after a period of spontaneous activity.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data analysis for patch-clamp recordings

Off-line analysis was performed with Fitmaster (Heka Elektronik), Igor-Pro 6.0.3 (Wavemetrics, Lake Oswego, OR, USA) and Matlab R2012b (Mathworks). Statistical analyses were performed with Prism 5.02 software (San Diego, CA, USA). "n" refers to the number of experiments (performed on a single neuron from a single slice). All results are expressed as mean±SEM. Statistical significance was assessed in one-sample *t* tests, unpaired *t* tests, paired *t* tests or ANOVA as appropriate, using the indicated significance threshold (*p*). For the fits in Figures 4d and 4i, we interpolated 100 points linearly using scipy.interpolate.interp1d for each STDP curve and then applied a Savitzky-Golay filter (from scipy.signal.savgo_filter) with parameters window_length=25 and polyorder=2. We also added the same interpolated data for significant LTP and LTD, in the range [-30, 0ms] at $F_{pairings}$=2.5Hz for DMS-MSNs (Fig. 3g and 3i).

Mathematical model simulations and statistical analysis.

Simulations were performed on Python 3.X, using the Anaconda suite (Anaconda Software Distribution, Computer software Version 2-2.4.0. Anaconda, Nov. 2016. Web. https ://anaconda.com.). The Python libraries of numeric calculus *numpy* and plotting *matplotlib* were used. Custom code is freely accessible on https://github.com/gvignoud/dms_dls. Simulations were run on the INRIA CLEPS cluster and HPC resources from GENCI-IDRIS, using GNU parallel (Tange, O. (2020, May 22). GNU Parallel 20200522 ('Kraftwerk'). Zenodo. https

867 ://doi.org/10.5281/zenodo.3841377). We used a Euler scheme to simulate our network and Poisson
868 processes, with $dt = 0.2ms$.
869 To study the network evolution during the different phases and compute learning accuracy, we
870 evaluated some properties of the network every 50 pattern iterations (except at the beginning of each
871 phase, where we recorded every 5 pattern iterations). During these test sessions, we froze the network
872 structure by considering that
873 — Both types of plasticity are turned off;
874 — The three noise components described above were suppressed ($\lambda_r = 0, \tau_p = 0$);
875 — All patterns were successively presented to the network, and the accuracy was computed using the
876 MSN responses, as described in Equation (1);
877 — Between each pattern, the MSN potential was reset to its equilibrium value, in order to avoid
878 influence of one pattern to another.
879 For each set of parameters, we ran 200 different simulations. Statistics were computed on a random
880 split of the simulations into 10 sets of 20 simulations, to compare statistics of performance of the
881 network across multiple conditions. We use statistical t-test from scipy.stats Python library (* $p <$
882 0.05, ** $p < 0.005$, *** $p < 0.0005$).
883 We started by collecting the mean accuracy at the end of each phase A_j for $j = 1,2,3,4$.
884 To characterize the kinetics of learning in the maintenance phase, we fitted a sigmoidal function to
885 the curve of accuracy as a function of time. This generic sigmoid with 5 parameters, was given by:

886 $$S[A_{init}, t_{init}, A_{end}, d, g](t) = A_{end} + (A_{init} - A_{end}) \times \left(1 + \left(\frac{t - t_{init}}{g}\right)\right)^{-d} \quad (11)$$

887 where A_{init} is the initial value of the sigmoid, t_{init} the starting time of the second phase (in terms of
888 pattern iterations), A_{end} is the final value, g is a timescale parameter, while d is a shape parameter of
889 the sigmoid. Fitting the sigmoid allows comparing the dynamics in the maintenance phase for various

conditions, and in particular the performance of symmetric and asymmetric anti-Hebbian STDP. The initial value was set at $A_{init} = A_2$ and the ending value at $A_{end} = A_1$.The fits were realized by estimating the best values of d and g to reproduce the accuracy dynamics and were performed using the scipy.optimize.curvefit function of the scipy Python library.

We compared maintenance of the learning task through the characteristic time of decay to represent the speed of decrease,

$$T_{maintenance} = \min\left(\max\left(0, \frac{g}{d}\right), 1000\right). \quad (12)$$

Finally, to measure the relearning characteristic time, i.e. the time necessary for the system to relearn patterns after the maintenance phase, we define $T_{relearning}$ as follows. Remembering that A_1 (respectively, A_2) is the accuracy after the initial phase (respectively, learning phase), we define $T_{relearning}$ as the time needed in the relearning phase to learn again at least 60% of the previously learned accuracy,

$$T_{relearning} = \inf\{t > 0 | Accuracy(t) - A_1 > 0.6 \times (A_2 - A_1)\}. \quad (13)$$

We investigated how synaptic weight during the maintenance phase deviate from those at the end of the learning phase. To this end, we defined W_{ref} the synaptic weights at the end of the learning phase, and used various metrics to analyze the divergence of the weights from this value during maintenance:

- The divergence of the L^2 norm,

$$d_2(W) = \frac{1}{1 + \frac{\sqrt{\sum_i \left(W_i - W_{ref,i}\right)^2}}{\sqrt{\sum_i W_{ref,i}{}^2}}} \quad (14)$$

which is equal to 1 when $W = W_{ref}$ and decays to 0 as the Euclidean distance between the two weight vectors increases. A large $d_2(W)$ (*i.e.* close to 1) means that weights remained similar to the

reference, and a decay of that quantity estimates how quickly the weight vector deviates from reference.

- One may consider that relative values of weights, rather than their absolute amplitude, contain a particularly important information in learning. W_{ref} provides a direction in the space of weights, and we estimated the alignment of the weight vector at a given time with W_{ref} through the *cosine similarity* of the centered synaptic weight:

$$sp(W) = \frac{\sum_i (W_i - \overline{W_i}) \times (W_{ref,i} - \overline{W_{ref,i}})}{\sqrt{\sum_i (W_i - \overline{W_i})^2} \times \sqrt{\sum_i (W_{ref,i} - \overline{W_{ref,i}})^2}} \quad (15)$$

where $\bar{x}$ for $x \in \mathbb{R}^P$ denotes the average of the vector component.

REFERENCES

- Alegre-Cortés, J., Sàez, M., Montanari, R., and Reig, R. (2021). Medium spiny neurons activity reveals the discrete segregation of mouse dorsal striatum. eLife 10:e60580. 10.7554/eLife.60580.

- Athalye, V. R., Santos, F. J., Carmena, J. M., and Costa, R. M. (2018). Evidence for a neural law of effect. Science 359(6379):1024-1029. 10.1126/science.aao6058.

- Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35(1):48-69. 10.1038/npp.2009.131.

- Barnes, T. D., Mao, J-B., Hu, D., Kubota, Y., Dreyer, A. A., Stamoulis, C., Brown, E. N., and Graybiel, A. M. (2011). Advance cueing produces enhanced action-boundary patterns of spike activity in the sensorimotor striatum. J. Neurophysiol. 105:1861–1878. 10.1152/jn.00871.2010.

- Bergstrom, H. C., Lipkin, A. M., Lieberman, A. G., Pinard, C. R., Gunduz-Cinar, O., Brockway, E. T., Taylor, W. W., Nonaka, M., Bukalo, O., Wills, T. A., et al. (2018). Dorsolateral Striatum Engagement Interferes with Early Discrimination Learning. Cell Rep. 23(8):2264-2272. 10.1016/j.celrep.2018.04.081.

- Bonnavion, P., Fernández, E. P., Varin, C., and de Kerchove d'Exaerde, A. (2019). It takes two to tango: Dorsal direct and indirect pathways orchestration of motor learning and behavioral flexibility. Neurochem. Int. 124:200-214. 10.1016/j.neuint.2019.01.009.

- Brzosko, Z., Mierau, S. B., and Paulsen, O. (2019). Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. Neuron 103(4):563-581. 10.1016/j.neuron.2019.05.041.

Burkitt, A. N. (2006). A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. Biological Cybernetics 95(1) :1–19. 10.1007/s00422-006-0068-6.

- Burton, A. C., Nakamura, K., and Roesch, M. R. (2015). From ventral-medial to dorsal-lateral striatum: neural correlates of reward-guided decision-making. Neurobiol. Learn. Mem. 117: 51-59. 10.1016/j.nlm.2014.05.003

- Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., and Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: a critical reappraisal. Nat. Neurosci. 17: 1022–1030. 10.1038/nn.3743.

Cerovic, M., d'Isa, R., Tonini, R., and Brambilla, R. (2013). Molecular and cellular mechanisms of dopamine-mediated behavioral plasticity in the striatum. Neurobiol. Learn. Mem. 105, 63-80. 10.1016/j.nlm.2013.06.013.

- Costa, R. M., Cohen, D., and Nicolelis, M. A. L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. 14: 1124-1134. 10.1016/j.cub.2004.06.053

- Corbit, L. H., and Janak, P. H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. Eur J Neurosci. 31(7):1312-21. 10.1111/j.1460-9568.2010.07153.x

- Cui, Y., Prokin, I., Xu, H., Delord, B., Genet, S., Venance, L., and Berry, H. (2016). Endocannabinoid dynamics gate spike-timing dependent depression and potentiation. Elife 5: 1–32. 10.7554/eLife.13185

Feldman, D. E. (2012). The spike-timing dependence of plasticity. Neuron 75: 556-571. 10.1016/j.neuron.2012.08.001.

- Fino, E., Vandecasteele, M., Perez, S., Saudou, F., and Venance, L. (2018). Region-specific and state-dependent action of striatal GABAergic interneurons. Nat Commun. 9(1):3339. 10.1038/s41467-018-05847-5.

- Fino, E., Glowinski, J., and Venance, L. (2005). Bidirectional activity-dependent plasticity at corticostriatal synapses. J. Neurosci. 25: 11279-11287. 10.1523/JNEUROSCI.4476-05.2005.

- Fino, E., Paille, V., Cui, Y., Morera-Herreras, T., Deniau, J., and Venance, L. (2010). Distinct coincidence detectors govern the corticostriatal sipke timing dependent plasticity. J. Physiol. 588: 3045-3062. 10.1113/jphysiol.2010.188466.

- Fisher, S. D., Robertson, P. B., Black, M. J., Redgrave, P., Sagar, M. A., Abraham, W. C., and Reynolds, J. N. J. (2017). Reinforcement determines the timing dependence of corticostriatal synaptic plasticity in vivo. Nat. Commun. 8(1):334. 10.1038/s41467-017-00394-x.

- Foncelle, A., Mendes, A., Jędrzejewska-Szmek, J., Valtcheva, S., Berry, H., Blackwell, K. T., and Venance, L. (2018). Modulation of Spike-Timing Dependent Plasticity: Towards the Inclusion of a Third Factor in Computational Models. Front. Comput. Neurosci. 12:49. 10.3389/fncom.2018.00049.

Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). Neuronal dynamics : from single neurons to networks and models of cognition. Cambridge University Press, New York, NY, USA.

- Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., and Brea, J. (2018). Eligibility traces and plasticity on behavioral time scales : experimental support of neoHebbian three-factor learning rules. Front. Neural Circuits 12:53. 10.3389/fncir.2018.00053.

- Graybiel, A. M., and Grafton, S. T. (2015). The striatum: where skills and habits meet. Cold Spring Harb Perspect. Biol. 7: a021691. 10.1101/cshperspect.a021691.

- Gremel, C. M., and Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. 4:2264. 10.1038/ncomms3264.

- Gurney, K. N., Humphries, M. D., and Redgrave, P. (2015). A new framework for cortico-striatal plasticity: behavioural theory meets in vitro data at the reinforcement-action interface. PLoS Biol. 13, e1002034. 10.1371/journal.pbio.1002034.

- Hawes, S. L., Evans, R. C., Unruh, B. A., Benkert, E. E., Gillani, F., Dumas, T. C., and Blackwell, K. T. (2015). Multimodal Plasticity in Dorsal Striatum While Learning a Lateralized Navigation Task. J. Neurosci. 35:10535–10549. 10.1523/JNEUROSCI.4415-14.2015.

- Hintiryan, H., Foster, N. N., Bowman, I., Bay, M., Song, M. Y., Gou, L., Yamashita, S., Bienkowski, M. S., Zingg, B., Zhu, M., et al.. (2016). The mouse cortico-striatal projectome. Nat. Neurosci. 19(8):1100-14. 10.1038/nn.4332.

- Hooks, B. M., Papale, A. E., Paletzki, R. F., Feroze, M. W., Eastwood, B. S., Couey, J. J., Winnubst, J., Chandrashekar, J., and Gerfen, C. R. (2018). Topographic precision in sensory and motor corticostriatal projections varies across cell types and cortical area. Nat. Commun. 9(1):3549. 10.1038/s41467-018-05780-7.

- Hunnicutt, B. J., Jongbloets, B. C., Birdsong, W. T., Gertz, K. J., Zhong, H., and Mao, T. (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. ELife 5: e19103. 10.7554/eLife.19103.

- Jin, X., and Costa, R. M. (2015). Shaping action sequences in basal ganglia circuits. Curr. Op. Neurobiol. 33, 188-196. 10.1016/j.conb.2015.06.011.

- Kim, J., Gulati, T., and Ganguly, K. (2019). Competing Roles of Slow Oscillations and Delta Waves in Memory Consolidation versus Forgetting. Cell 179(2):514-526.e13. 10.1016/j.cell.2019.08.040.

- Kimchi, E. Y., Torregrossa, M. M., Taylor, J. R., and Laubach, M. (2009). Neuronal correlates of instrumental learning in the dorsal striatum. J. Neurophysiol. 102: 475-489. 10.1152/jn.00262.2009.

- Koralek, A. C., Jin, X., Long, J. D., Costa, R. M., and Carmena, J. M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483: 331–335. 10.1038/nature10845.

- Kupferschmidt, D. A., Juczewski, K., Cui, G., Johnson, K. A., and Lovinger, D. M. (2017). Parallel, but Dissociable, Processing in Discrete Corticostriatal Inputs Encodes Skill Learning. Neuron 96(2):476-489.e5. 10.1016/j.neuron.2017.09.040.

- Kuśmierz, Ł., Isomura, T., and Toyoizumi, T. (2017). Learning with three factors: modulating Hebbian plasticity with errors. Curr. Opin. Neurobiol. 46:170-177. 10.1016/j.conb.2017.08.020.

- Li, Y., He, Y., Chen, M., Pu, Z., Chen, L., Li, P., Li, B., Li, H., Huang, Z. L., Li, Z., and Chen, J. F. (2016). Optogenetic Activation of Adenosine A2A Receptor Signaling in the Dorsomedial Striatopallidal Neurons Suppresses Goal-Directed Behavior. Neuropsychopharmacology 41(4):1003-13. 10.1038/npp.2015.227.

- Lovinger, D. M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. Neuropharmacology 58, 951-961. 10.1016/j.neuropharm.2010.01.008.

- Ma, T., Cheng, Y., Roltsch Hellard, E., Wang, X., Lu, J., Gao, X., Huang, C. C. Y., Wei, X-Y., Ji, J-Y., and Wang, J. (2018). Bidirectional and long-lasting control of alcohol-seeking behavior by corticostriatal LTP and LTD. Nat. Neurosci. 21: 373–383. 10.1038/s41593-018-0081-9.

- Mahon, S., Vautrelle, N., Pezard, L., Slaght, S. J., Deniau, J. M., Chouvet, G., and Charpier, S. (2006). Distinct patterns of striatal medium spiny neuron activity during the natural sleep-wake cycle. J. Neurosci. 26(48):12587-95. 10.1523/JNEUROSCI.3987-06.2006.

- Mathur, B. N., and Lovinger, D. M. (2012). Endocannabinoid–Dopamine Interactions in Striatal Synaptic Plasticity. Front. Pharmacol. 3, :66. 10.3389/fphar.2012.00066.

- Mendes, A., Vignoud, G., Perez, S., Perrin, E., Touboul., J., and Venance, L. (2020). Concurrent Thalamostriatal and Corticostriatal Spike-Timing-Dependent Plasticity and Heterosynaptic Interactions Shape Striatal Plasticity Map. Cereb. Cortex. 30(8):4381-4401. 10.1093/cercor/bhaa024.

- Morera-Herreras, T., Gioanni, Y., Perez, S., Vignoud, G., and Venance, L. (2019). Environmental enrichment shapes striatal spike-timing-dependent plasticity in vivo. Sci. Rep. 9(1):19451. 10.1038/s41598-019-55842-z.

- Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. Biological Cybernetics 98(6) :459–478. 10.1007/s00422-008-0233-1.

- Muñoz, B., Haggerty, D. L., and Atwood, B. K. (2020). Synapse-specific expression of mu opioid receptor long-term depression in the dorsomedial striatum. Sci. Rep. 10(1):7234. 10.1038/s41598-020-64203-0.

- O'Hare, J. K., Ade, K. K., Sukharnikova, T., Van Hooser, S. D., Palmeri, M. L., Yin, H. H., and Calakos, N. (2016). Pathway-Specific Striatal Substrates for Habitual Behavior. Neuron 89(3):472-9. 10.1016/j.neuron.2015.12.032.

- Oberto, V. J., Boucly, C. J., Gao, H. Y., Todorova, R., Zugaro, M. B., and Wiener, S. I. (2022). Distributed cell assemblies spanning prefrontal cortex and striatum. Curr. Biol. 32, 1-13. 10.1016/j.cub.2021.10.007.

- Paillé, V., Fino, E., Du, K., Morera Herreras, T., Perez, S., Hellgren Kotaleski, J., and Venance, L. (2013). GABAergic circuits control spike-timing-dependent plasticity. J Neurosci 33: 9353-9363. 10.1523/JNEUROSCI.5796-12.2013.

- Partridge, J. G., Tang, K. C., and Lovinger, D. M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. J. Neurophysiol. 84(3):1422-9. 10.1152/jn.2000.84.3.1422.

- Pawlak, V., and Kerr, J. N. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. J Neurosci 28: 2435-2446. 10.1523/JNEUROSCI.4402-07.2008.

- Perrin, E., and Venance, L. (2019). Bridging the gap between striatal plasticity and learning. Curr Opin Neurobiol. 54:104–112. 10.1016/j.conb.2018.09.007.

- Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D., and Carandini, M. (2021). Striatal activity topographically reflects cortical activity. Nature 591(7850):420-425. 10.1038/s41586-020-03166-8.

- Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J., and Kesner, R. P. (2002a). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. Behav. Neurosci. 116(1):105-15. 10.1037//0735-7044.116.1.105.

- Ragozzino, M. E., Jih, J., and Tzavos, A. (2002b). Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. Brain Res. 953(1-2):205-14. 10.1016/s0006-8993(02)03287-0.

- Ragozzino, M. E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. Ann. N. Y. Acad. Sci. 1121:355-75. 10.1196/annals.1401.013.

- Reig, R., and Silberberg, G. (2014). Multisensory integration in the mouse striatum. Neuron 83(5):1200-12. 10.1016/j.neuron.2014.07.033.

- Requarth, T., and Sawtell, N. B. (2011). Neural mechanisms for filtering self-generated sensory signals in cerebellum-like circuits. Curr. Opin. Neurobiol. 21(4):602-8. 10.1016/j.conb.2011.05.031.

- Roberts, P. D., and Leen, T. K. (2010). Anti-hebbian spike-timing-dependent plasticity and adaptive sensory processing. Front. Comput. Neurosci. 4:156. 10.3389/fncom.2010.00156.

- Roberts, P. D., and Bell, C. C. (2000). Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. J. Comput. Neurosci. 9, 67–83. 10.1023/a:1008938428112.

- Rumsey, C. C., and Abbott, L. F. (2004). Equalization of synaptic efficacy by activity- and timing-dependent synaptic plasticity. J. Neurophysiol. 91(5):2273-80. 10.1152/jn.00900.2003.

- Schulz, J. M., Redgrave, P., and Reynolds, J. N. (2010). Cortico-striatal spike-timing dependent plasticity after activation of subcortical pathways. Front. Synaptic Neurosci 2: 23. 10.3389/fnsyn.2010.00023.

- Shan, Q., Ge, M., Christie, M. J., and Balleine, B. W. (2014). The acquisition of goal-directed actions generates opposing plasticity in direct and indirect pathways in dorsomedial striatum. J. Neurosci. 34: 9196–9201. 10.1523/JNEUROSCI.0313-14.2014.

- Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. Science 321: 848-851. 10.1126/science.1160575.

- Smith, A. C. W., Jonkman, S., Difeliceantonio, A. G., O'Connor, R. M., Ghoshal, S., Romano, M. F., Everitt, B. J., and Kenny, P. J. (2021). Opposing roles for striatonigral and striatopallidal neurons in dorsolateral striatum in consolidating new instrumental actions. Nat. Commun. 12(1):5121. 10.1038/s41467-021-25460-3.

- Stalnaker, T. A., Calhoon, G. G., Ogawa, M., Roesch, M. R., and Schoenbaum, G. (2010). Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front. Integr. Neurosci. 4:12. 10.3389/fnint.2010.00012.

- Steriade, M., McCormick, D. A., and Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. Science 262, 679-685. 10.1126/science.8235588.

- Thorn, C. A., Atallah, H., Howe, M., and Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron 66, 781-795. 10.1016/j.neuron.2010.04.036.

- Thorn, C. A., and Graybiel, A. M. (2014). Differential entrainment and learning-related dynamics of spike and local field potential activity in the sensorimotor and associative striatum. J. Neurosci. 34(8):2845-59. 10.1523/JNEUROSCI.1782-13.2014.

- Valtcheva, S., and Venance, L. (2016). Astrocytes gate Hebbian synaptic plasticity in the striatum. Nat. Commun. 7: 13845. 10.1038/ncomms13845.

- Valtcheva, S., Paillé, V., Dembitskaya, Y., Perez, S., Gangarossa, G., Fino, E., and Venance, L. (2017). Developmental control of spike-timing-dependent plasticity by tonic GABAergic signaling in striatum. Neuropharmacology 121: 261–277. 10.1016/j.neuropharm.2017.04.012.

- Vandaele, Y., Mahajan, N. R., Ottenheimer, D. J., Richard, J. M., Mysore, S. P., and Janak, P. H. (2019). Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. Elife 8:e49536. 10.7554/eLife.49536.

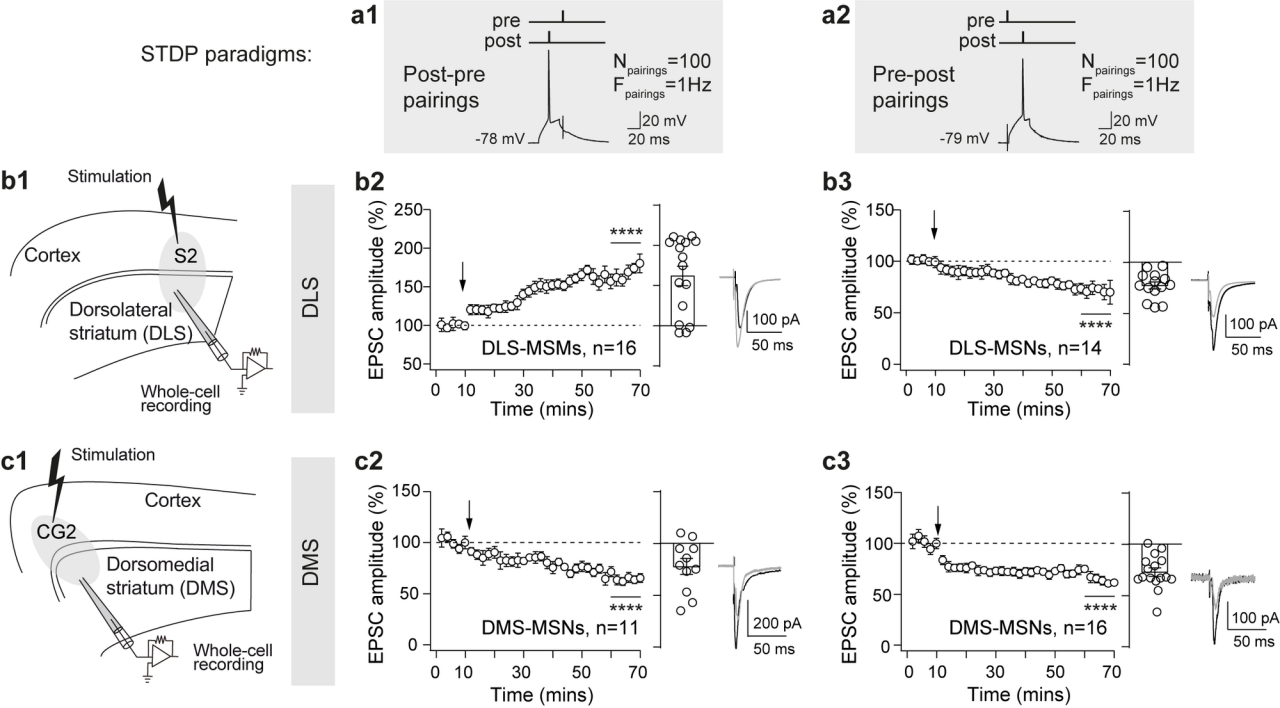
- Williams, A., Roberts, P. D., and Leen, T. K. (2003). Stability of negative-image equilibria in spike-timing-dependent plasticity. Phys Rev E 68(2 Pt 1):021923. 10.1103/PhysRevE.68.021923.

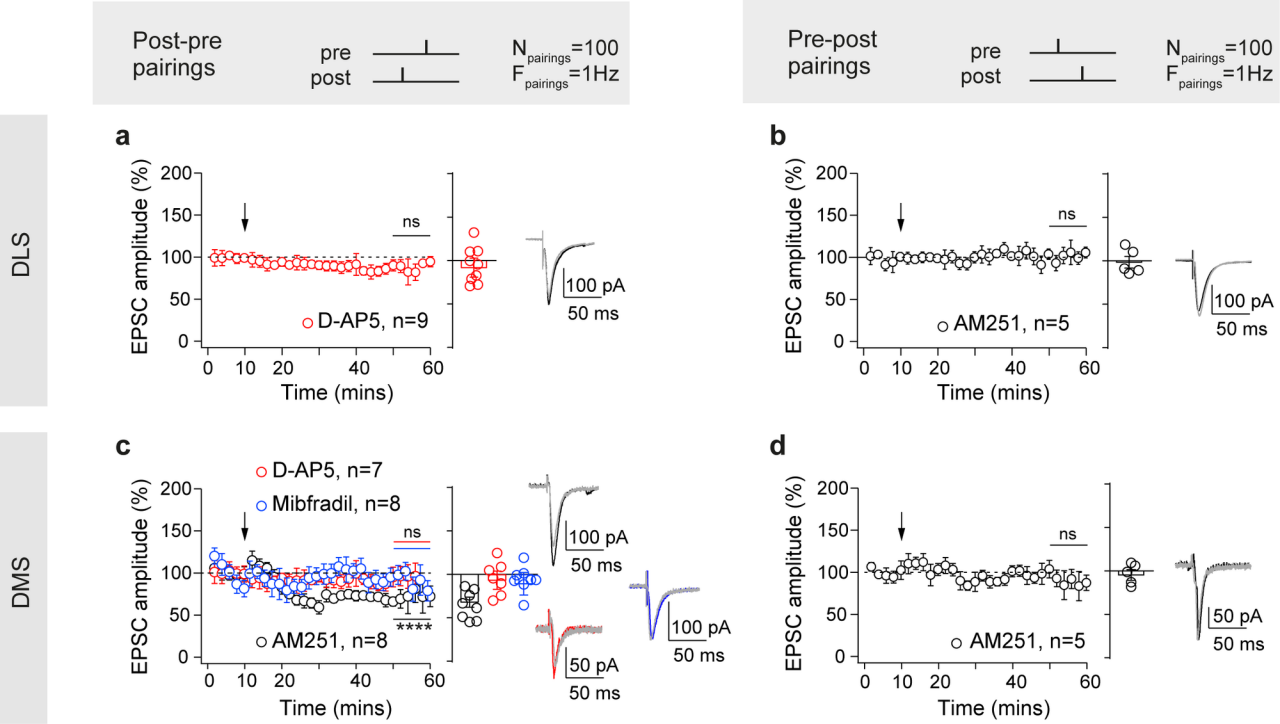
- Xiong, Q., Znamenskiy, P., and Zador, A. M. (2015). Selective corticostriatal plasticity during acquisition of an auditory discrimination task. Nature 521:348–351. 10.1038/nature14225.

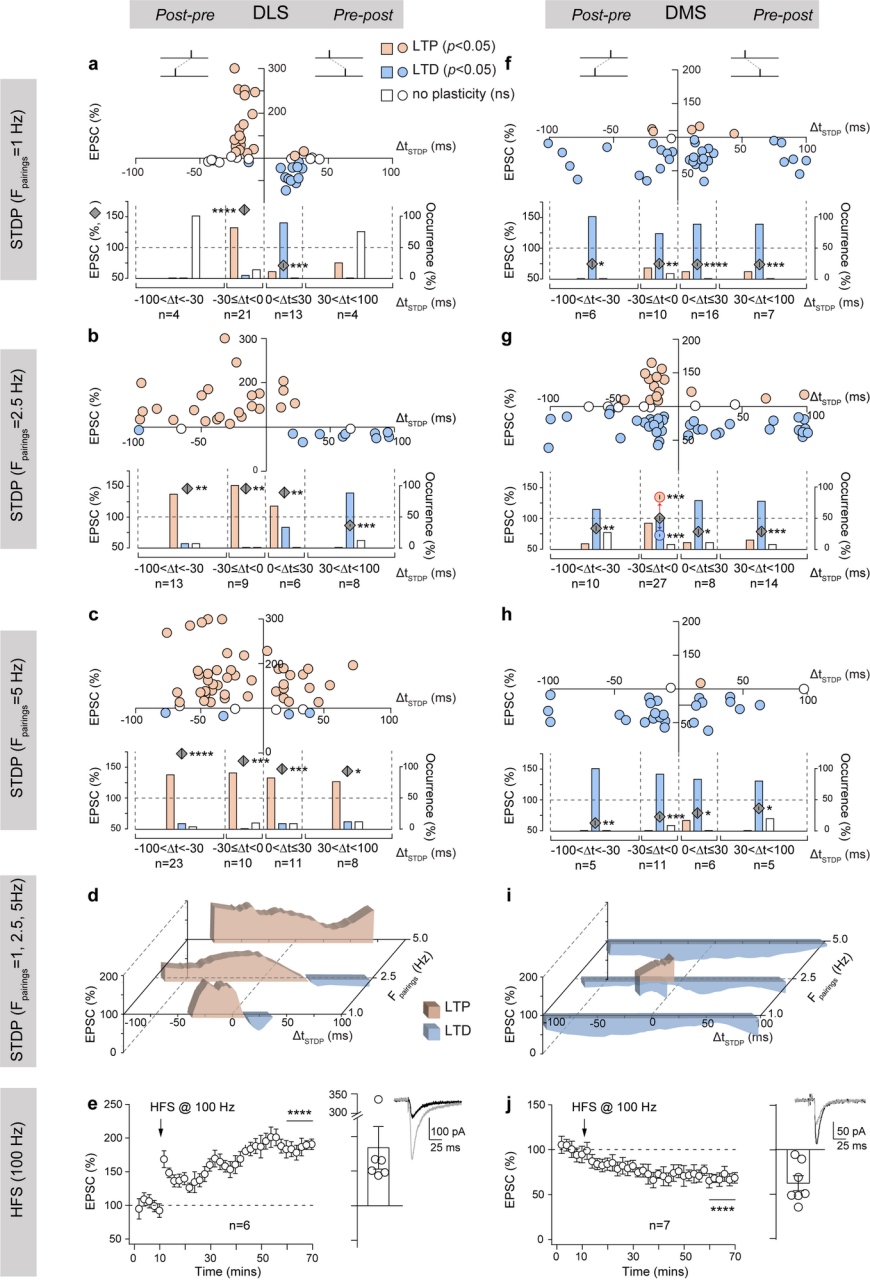
- Yim, M. Y., Aertsen, A., and Kumar, A. (2011). Significance of input correlations in striatal function. PLoS Computational Biology 7(11): e1002254. 10.1371/journal.pcbi.1002254.

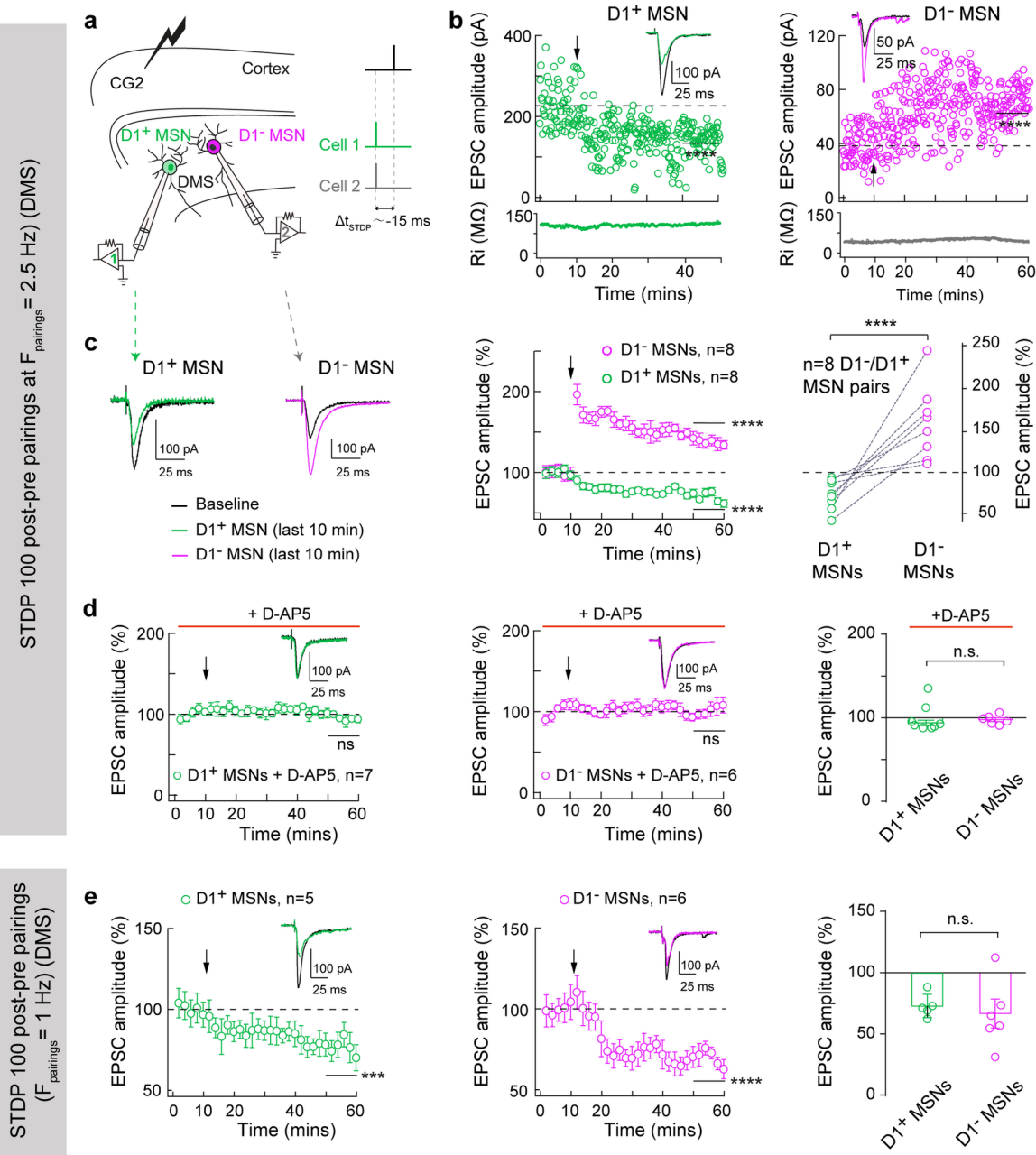
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. 7: 464–476. 10.1038/nrn1919.

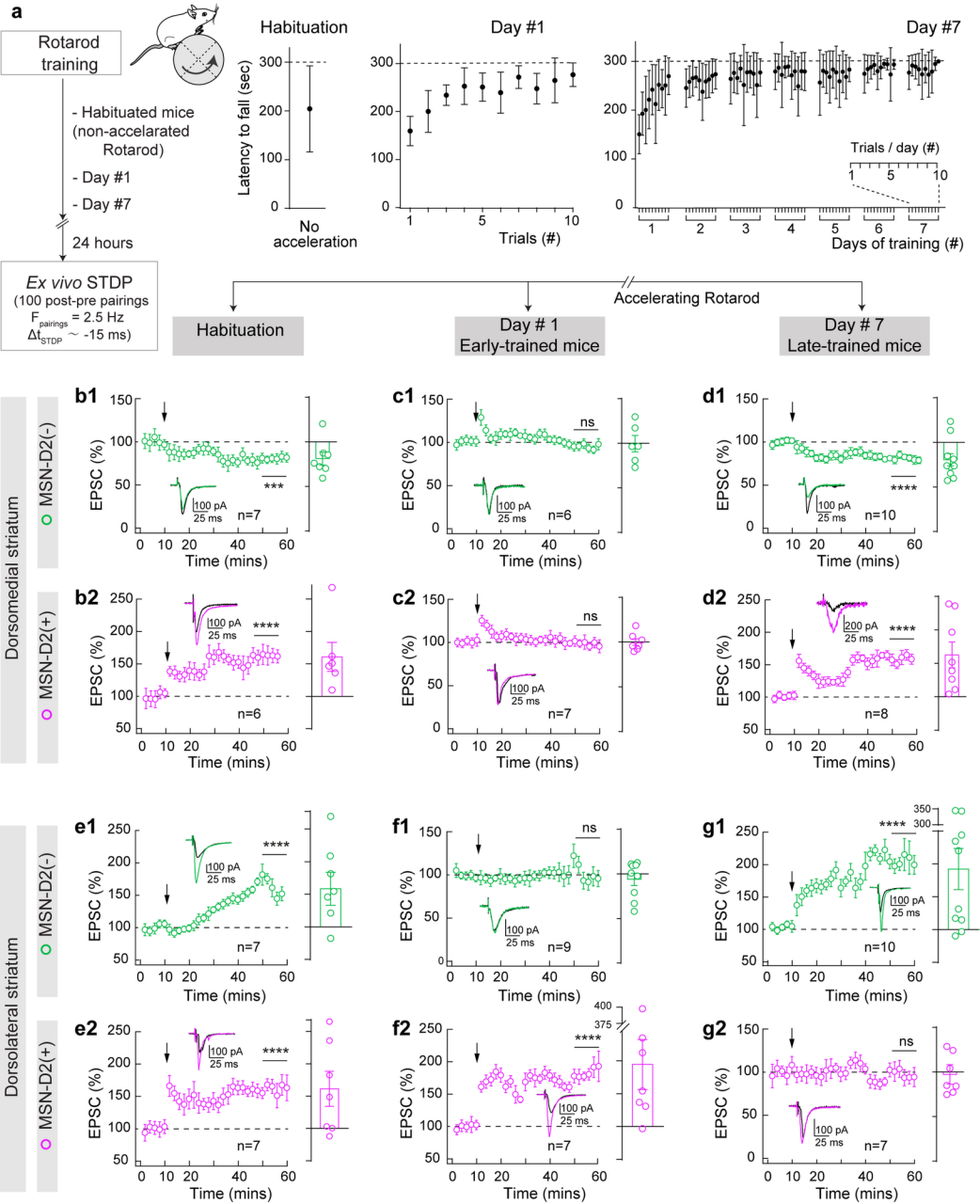
- Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., and Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. 12: 333–341. 10.1038/nn.2261.

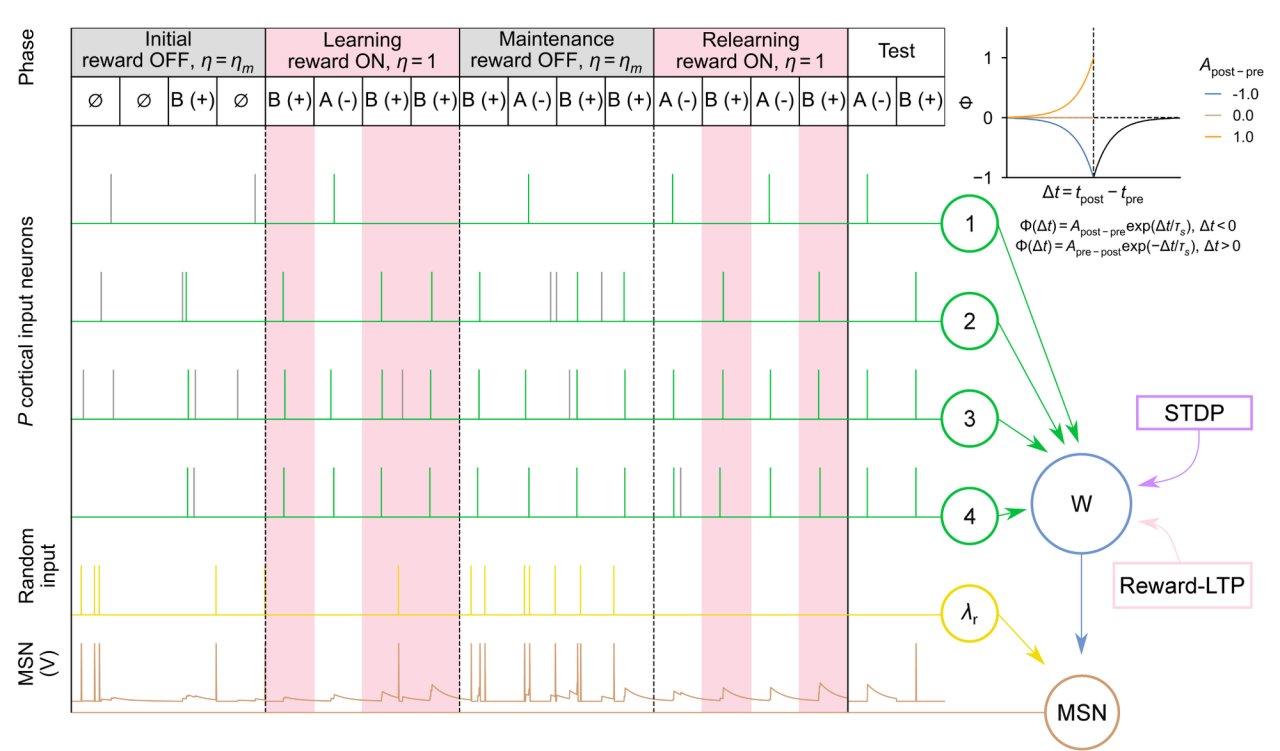

STDP paradigms:
a1
pre
post
Post-pre pairings
N pairings=100
F pairings=1Hz
20 mV
20 ms
-78 mV
a2
pre
post
Pre-post pairings
N pairings=100
F pairings=1Hz
20 mV
20 ms
-79 mV
b1
Stimulation
Cortex
S2
Dorsolateral striatum (DLS)
Whole-cell recording
DLS
b2
EPSC amplitude (%)
DLS-MSMs, n=16
****
Time (mins)
100 pA
50 ms
b3
EPSC amplitude (%)
DLS-MSNs, n=14
****
Time (mins)
100 pA
50 ms
c1
Stimulation
Cortex
CG2
Dorsomedial striatum (DMS)
Whole-cell recording
DMS
c2
EPSC amplitude (%)
DMS-MSNs, n=11
****
Time (mins)
200 pA
50 ms
c3
EPSC amplitude (%)
DMS-MSNs, n=16
****
Time (mins)
100 pA
50 ms

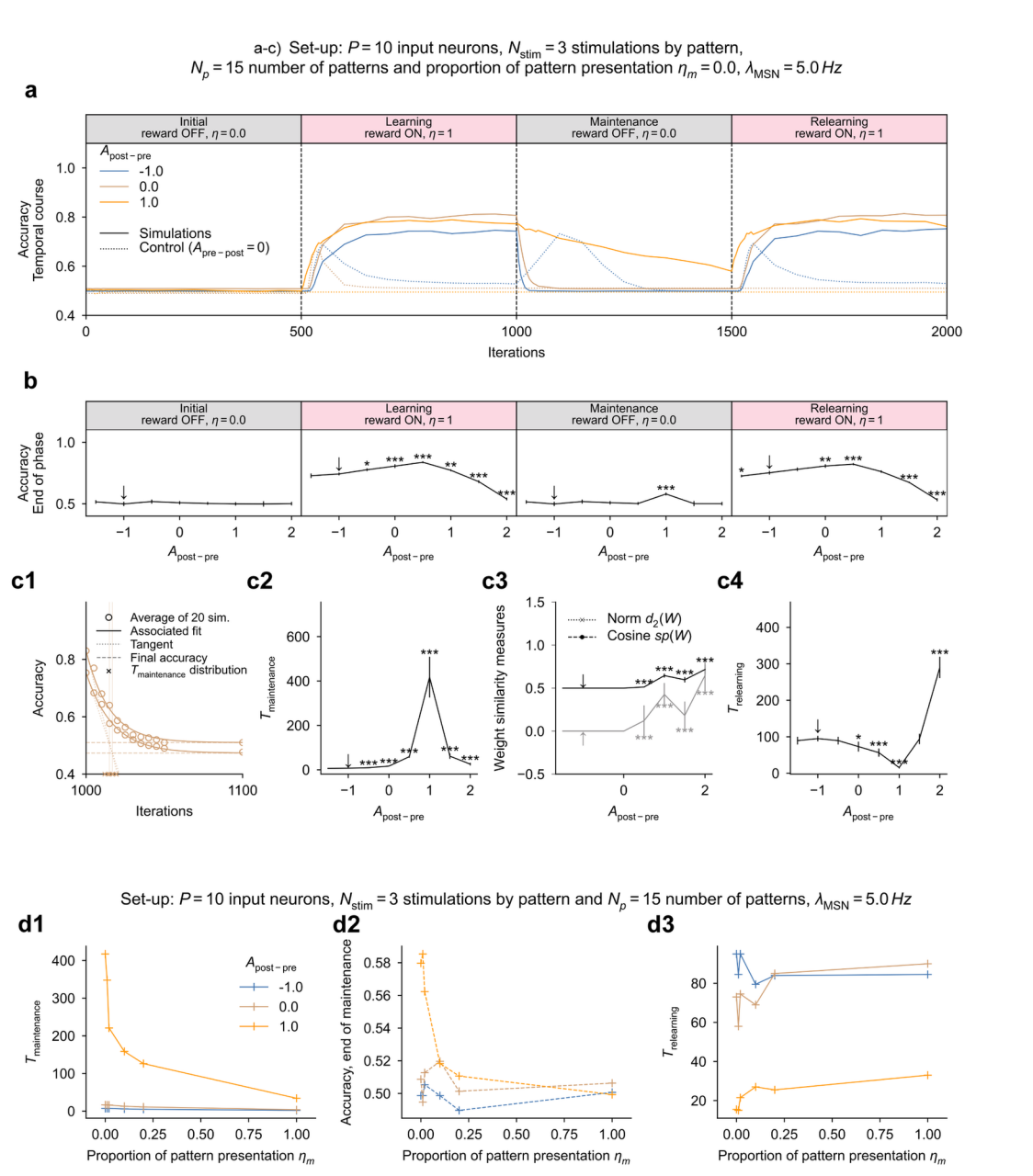

Post-pre pairings
pre
post
N pairings=100
F pairings=1Hz
Pre-post pairings
pre
post
N pairings=100
F pairings=1Hz
DLS
DMS
a
b
c
d
EPSC amplitude (%)
Time (mins)
ns
D-AP5, n=9
AM251, n=5
D-AP5, n=7
Mibfradil, n=8
AM251, n=8
****
100 pA
50 ms
50 pA
50 ms


Post-pre
DLS
Pre-post
Post-pre
DMS
Pre-post
LTP (p<0.05)
LTD (p<0.05)
no plasticity (ns)
STDP (F pairings=1 Hz)
STDP (F pairings=2.5 Hz)
STDP (F pairings=5 Hz)
STDP (F pairings=1, 2.5, 5Hz)
HFS (100 Hz)
a
b
c
d
e
f
g
h
i
j
EPSC (%)
Occurrence (%)
Δt STDP (ms)
-100<Δt<-30
-30≤Δt<0
0<Δt≤30
30<Δt<100
n=4
n=21
n=13
n=4
n=13
n=9
n=6
n=8
n=23
n=10
n=11
n=8
n=6
n=10
n=16
n=7
n=10
n=27
n=8
n=14
n=5
n=11
n=6
n=5
F pairings (Hz)
LTP
LTD
HFS @ 100 Hz
n=6
n=7
Time (mins)
100 pA
25 ms
50 pA
25 ms


STDP 100 post-pre pairings at F pairings = 2.5 Hz) (DMS)
STDP 100 post-pre pairings (F pairings = 1 Hz) (DMS)
a
CG2
Cortex
D1+ MSN
D1- MSN
DMS
Cell 1
Cell 2
Δt STDP ~-15 ms
b
D1+ MSN
D1- MSN
EPSC amplitude (pA)
Ri (MΩ)
100 pA
25 ms
50 pA
****
Time (mins)
c
D1+ MSN
D1- MSN
100 pA
25 ms
Baseline
D1+ MSN (last 10 min)
D1- MSN (last 10 min)
D1- MSNs, n=8
D1+ MSNs, n=8
EPSC amplitude (%)
n=8 D1-/D1+ MSN pairs
D1+ MSNs
D1- MSNs
d
+ D-AP5
ns
D1+ MSNs + D-AP5, n=7
D1- MSNs + D-AP5, n=6
n.s.
e
D1+ MSNs, n=5
D1- MSNs, n=6
***


a
Rotarod training
- Habituated mice (non-accelarated Rotarod)
- Day #1
- Day #7
24 hours
Ex vivo STDP
(100 post-pre pairings
F pairings = 2.5 Hz
Δt STDP ~ -15 ms)
Habituation
Day #1
Day #7
Latency to fall (sec)
No acceleration
Trials (#)
Trials / day (#)
Days of training (#)
Accelerating Rotarod
Habituation
Day # 1
Early-trained mice
Day # 7
Late-trained mice
Dorsomedial striatum
Dorsolateral striatum
MSN-D2(-)
MSN-D2(+)
EPSC (%)
Time (mins)
100 pA
25 ms
200 pA
25 ms
n=7
n=6
n=10
n=8
n=9
ns
***
****

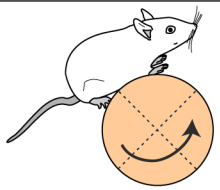

Phase
Initial
reward OFF, $\eta = \eta_m$
Learning
reward ON, $\eta = 1$
Maintenance
reward OFF, $\eta = \eta_m$
Relearning
reward ON, $\eta = 1$
Test
∅ ∅ B (+) ∅
B (+) A (-) B (+) B (+)
B (+) A (-) B (+) B (+)
A (-) B (+) A (-) B (+)
A (-) B (+)
P cortical input neurons
Random input
MSN (V)
1
2
3
4
λ_r
W
STDP
Reward-LTP
MSN
Φ
$A_{post-pre}$
-1.0
0.0
1.0
$\Delta t = t_{post} - t_{pre}$
$\Phi(\Delta t) = A_{post-pre}\exp(\Delta t/\tau_s)$, $\Delta t < 0$
$\Phi(\Delta t) = A_{pre-post}\exp(-\Delta t/\tau_s)$, $\Delta t > 0$


a-c) Set-up: $P = 10$ input neurons, $N_{stim} = 3$ stimulations by pattern,
$N_p = 15$ number of patterns and proportion of pattern presentation $\eta_m = 0.0$, $\lambda_{MSN} = 5.0\,Hz$
a
Initial
reward OFF, $\eta = 0.0$
Learning
reward ON, $\eta = 1$
Maintenance
reward OFF, $\eta = 0.0$
Relearning
reward ON, $\eta = 1$
$A_{post-pre}$
-1.0
0.0
1.0
Simulations
Control ($A_{pre-post} = 0$)
Accuracy
Temporal course
Iterations
b
Accuracy
End of phase
$A_{post-pre}$
c1
Average of 20 sim.
Associated fit
Tangent
Final accuracy
$T_{maintenance}$ distribution
Accuracy
Iterations
c2
$T_{maintenance}$
$A_{post-pre}$
c3
Norm $d_2(W)$
Cosine $sp(W)$
Weight similarity measures
$A_{post-pre}$
c4
$T_{relearning}$
$A_{post-pre}$
Set-up: $P = 10$ input neurons, $N_{stim} = 3$ stimulations by pattern and $N_p = 15$ number of patterns, $\lambda_{MSN} = 5.0\,Hz$
d1
$A_{post-pre}$
-1.0
0.0
1.0
$T_{maintenance}$
Proportion of pattern presentation η_m
d2
Accuracy, end of maintenance
Proportion of pattern presentation η_m
d3
$T_{relearning}$
Proportion of pattern presentation η_m

Corticostriatal plasticity

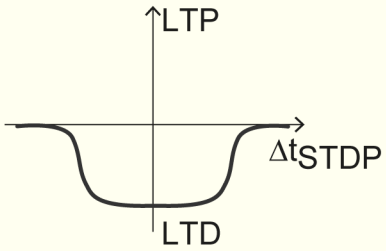
Anti-Hebbian STDP

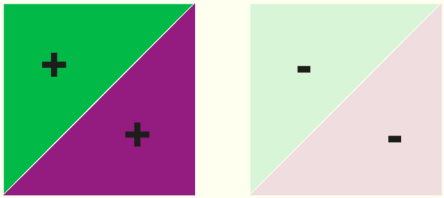
Motor skill learning



Early- *vs* late-trained mice

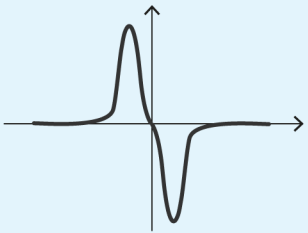
Dorsomedial striatum

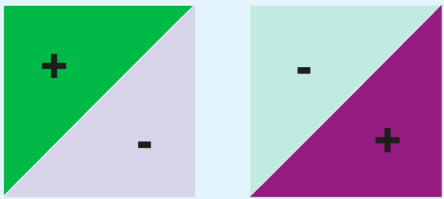




Symmetric ⟶ Memory flexibility

Dorsolateral striatum





Asymmetric ⟶ Memory maintenance

+/- STDP occlusion

Striatonigral MSNs

Striatopallidal MSNs