



HAL
open science

The Dorsal Striatum Energizes Motor Routines

Maria-Teresa Jurado-Parras, Mostafa Safaie, Stefania Sarno, Jordane Louis,
Corane Karoutchi, Bastien Berret, David Robbe

► **To cite this version:**

Maria-Teresa Jurado-Parras, Mostafa Safaie, Stefania Sarno, Jordane Louis, Corane Karoutchi, et al.. The Dorsal Striatum Energizes Motor Routines. *Current Biology - CB*, 2020, pp.4362-4372. 10.1016/j.cub.2020.08.049 . hal-02964179

HAL Id: hal-02964179

<https://amu.hal.science/hal-02964179>

Submitted on 26 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

The dorsal striatum energizes motor routines

Maria-Teresa Jurado-Parras^{*1}, Mostafa Safaie^{*1}, Stefania Sarno^{*1,2},
Jordane Louis¹, Corane Karoutchi¹, Bastien Berret^{3,4,5}, David Robbe^{1†}

¹ Institut de Neurobiologie de la Méditerranée (INMED), Turing Center for Living Systems, INSERM, Aix-Marseille Université, Parc Scientifique et Technologique de Luminy, 13009 Marseille, France.

² Institut de Recherche sur les Phénomènes Hors Equilibres (IRPHE), Centrale Marseille, CNRS, Aix-Marseille Université, 49 Rue Frédéric Joliot Curie, 13013 Marseille, France.

³ Université Paris-Saclay CIAMS, 15 Rue Georges Clemenceau, 91400 Orsay, France.

⁴ CIAMS, Université d'Orléans, 2 allée du Château, 45062 Orléans, France

⁵ Institut Universitaire de France, 1 rue Descartes, 75231 Paris Cedex 05, France

* Co-first authors

†Lead Contact; E-mail: david.robbe@inserm.fr

Summary

The dorsal striatum (dS) has been implicated in storing procedural memories and controlling movement kinematics. Since procedural memories are expressed through movements, the exact nature of the dS function has proven difficult to delineate. Here we challenged rats in complementary locomotion-based tasks designed to alleviate this confound. Surprisingly, dS lesions did not impair the rats' ability to remember the procedure for the successful completion of motor routines. However, the speed and initiation of the reward-oriented phase of the routines were irreversibly altered by the dS lesion. Further behavioral analyses combined with modeling in the optimal control framework indicated that these kinematic alterations were well-explained by an increased sensitivity to effort. Our work provides evidence supporting a primary role of the dS in modulating the kinematics of reward-oriented actions, a function that may be related to the optimization of the energetic costs of moving.

Keywords: *dorsal striatum, vigor, procedural memory, effort, basal ganglia, motor control, rats, locomotion*

1 Introduction

2 Most of the daily life behaviors of humans and other animals require the ability to quickly re-
3 member what to do in a familiar context. This so-called procedural memory is required (among
4 many other sensory, motor, and cognitive processes) when, for instance, we commute between
5 our home and our workplace or when we routinely use a wide range of tools. It is generally be-
6 lieved that the basal ganglia, and most specifically the dorsal striatum (dS, or caudate-putamen),
7 store procedural memories and by doing so can drive the execution of well-learned actions [1, 2].
8 Evidence supporting this view is essentially derived from experiments in humans performing
9 associative learning [3] and motor sequences [4, 5] tasks and in rodents and non-human pri-
10 mates engaged in a variety of experimental settings involving navigation in mazes [6–8], arm
11 reaching sequences [9], lever-press instrumental conditioning [10–17] or accelerating rotarods
12 [18–20]. A role of the dorsal striatum in procedural memory (i.e., selecting and executing adap-
13 tive actions) is compatible with the fact that dS projection neurons forming the direct (indirect)
14 basal ganglia pathway are known to facilitate (prevent) movement production through disinhi-
15 bition (inhibition) of brainstem and forebrain motor regions [21–24]. However, this dichotomy
16 in the functional anatomy of the basal ganglia is also compatible with a bi-directional control
17 of behavioral invigoration [25, 26]. This alternative hypothesis is in agreement with bradyki-
18 nesia, one of the most common motor deficits seen in Parkinson’s disease (PD) [27] and the
19 repeated observation that dS activity correlates with the speed of movements and locomotion
20 [24, 28–32].

21 Delineating the dS function (i.e., procedural memory storage vs control of movement speed/
22 vigor) is challenging because movements are the readout of procedural memories [33]. Conse-
23 quently, in tasks used to probe the dS contribution to procedural memory through perturbation
24 of neuronal activity, it is nearly impossible to disentangle whether impaired performance arises
25 from an inability to implement a preserved procedural memory into actions (i.e., a motor control
26 deficit), or from a direct alteration of the stored procedural memory (i.e., not remembering what

27 to do) with preserved motor control. Such a performance confound is maybe most obvious in
28 the case of the accelerating rotarod task which requires both to control the speed of locomotion
29 and remembering the recipe for a successful trial (postural adjustments at trial onset and running
30 faster and faster on the small accelerating rod). Here, we attempted to address this conundrum
31 and limit the impact of the performance confound by examining the impact of dS lesion in rats
32 performing a series of locomotion-based tasks with varying requirements in terms of movement
33 speed control and procedural memory. On the one hand, we found that dS lesion did not prevent
34 the animals from remembering the procedural steps to follow to successfully perform a previ-
35 ously learned routine. On the other hand, dS lesion changed the kinematic parameters of the
36 routine execution in a way that is well accounted for by an increased sensitivity to effort. By
37 setting the sensitivity to effort, the dS may contribute to the modulation of the energy invested
38 into voluntary movements, in the general context of an optimal control of reward-oriented ac-
39 tions. Such an elementary function might explain the previous involvement of the dS in both
40 procedural memory and the control of the speed of movements.

41 **Results**

42 **Lesions of the dorsal striatum did not prevent the overall performance of** 43 **motor routine**

44 To understand how the dS contributes to the control of voluntary actions while limiting as much
45 as possible the aforementioned issue of performance confound, we challenged rats in a set of
46 motor tasks taking place on a motorized treadmill. In the first task, to obtain a drop of sucrose
47 solution, rats ($n = 67$) had to wait for a fixed goal time ($GT = 7$ s relative to treadmill
48 onset) before entering a reward area located at the front of the treadmill, while its belt was
49 slowly moving backward (Figure 1A) [34]. Across practice sessions composed of ~ 130 trials
50 (treadmill on) separated with intertrial intervals (treadmill off), animals learned to wait longer
51 and longer before entering the reward area, until reliably doing it very close to the GT (Figure

52 1B). Task proficiency was clearly associated with the acquisition and reliable performance of
53 the following routine (Figure 1C, compare left and middle, Video S1): 1) during the intertrial
54 interval, following the consumption of sweetened water, rats remained in the reward area; 2)
55 when the treadmill was turned on (trial onset), they did nothing and let the moving belt carry
56 them away from the reward area; 3) when they reached the rear wall of the treadmill, they outran
57 the opposing treadmill belt to re-enter the reward area just after 7 s (Entrance Time, $ET > GT$).
58 After 2-3 weeks of daily practice, most of the rats used this wait-and-run routine in about 75%
59 of the trials (Figure 1C, right). Performing this routine requires procedural memory (among
60 other sensory, motor and cognitive processes) insofar as animals must remember that a given
61 sensorimotor state must be associated with a given action (to run or to stay immobile) to obtain
62 a reward [34]. Learning this routine was paralleled by a robust invigoration of the running phase
63 toward the reward area (i.e., during the first 10 sessions, rats crossed the treadmill toward the
64 reward area with progressively faster speeds, Figure 1D). Rats could have used several other
65 strategies to perform proficiently in this task. For instance, they could have remained close to
66 the reward area by running at the same speed as the treadmill for several seconds, and then
67 perform a short acceleration to enter the reward area just after GT . Still, we recently reported
68 that the usage of the wait-and-run routine facilitated timing accuracy and that a majority of
69 animals relied on this strategy [34], despite the fact that it might not be the optimal solution in
70 terms of effort.

71 We then performed fiber-sparing lesions of the dS in 57 animals well-trained in this task.
72 The lesions targeted either the dorsolateral or the dorsomedial region of the striatum or both
73 territories (DLS, DMS, DS lesions, Figures 1E–1G, Figures S1A–S1D). Behavioral testing re-
74 sumed two weeks after the lesion surgery (Figures S1E–S1G). The average ET of animals with
75 the largest lesions (mostly DS lesions) dropped during the first post-lesion session (Figures
76 1F-1H and 1J) because these animals ran toward the reward area prematurely after trial onset
77 (Figures 1F and 1G, third column). Consequently, a reduction in the usage of the wait-and-run

78 routine was observed during this first post-lesion session (Figures 1I and 1K). But surprisingly,
79 most of these animals recovered from this initial impairment and after a few additional sessions,
80 task proficiency was similar to the pre-lesion level (Figures 1H–1K). Concretely, for the major-
81 ity of rats with DS lesion (14 out of 16), the ability to follow the previously learned procedural
82 steps required for successful routine performance was intact (e.g., remain in the reward area
83 following reward consumption, running straight into the reward area after reaching the back
84 region of the treadmill). In addition, after a lesion restricted to the DLS or DMS, rats success-
85 fully continued to rely on the wait-and-run routine to enter the reward area very close to the *GT*
86 (Figures 1E, 1H and 1I).

87 While these results seem to indicate that dS lesion spared procedural memory (and many
88 other basic sensory, motor and cognitive processes required to perform the wait-and-run se-
89 quence), it could be argued that the transient impairment in performance induced by large dS
90 lesions is caused by a deterioration of procedural memory and reflects a reversal to pre-learning
91 behavior (Figures 1B and 1C, left) [17], compensated in subsequent post-lesion sessions through
92 a dS-independent learning process. To test the validity of this interpretation, a different group of
93 animals was trained in a modified version of the task in which the treadmill belt moved slowly
94 toward the reward area (instead of away from it). Several animals ($n = 9$) became proficient
95 in this version of the task by learning to perform a *run-and-wait* routine: they moved to the
96 back of the treadmill during the intertrial interval (after licking the reward from the previous
97 trial and while the belt was immobile) and, following trial onset, they remained still while be-
98 ing passively transported toward the reward area (Figures 2A–2C, Video S2). This routine was
99 interesting for two reasons. First, regarding the general issue of performance confound, the
100 run-and-wait routine can be performed even if animals display motor control deficits. Indeed,
101 to be proficient, animals simply had to go to the back of the treadmill during the relatively long
102 (20 s) intertrial interval (i.e., they could do it very slowly if necessary) and to remain still after
103 trial onset until the treadmill carries them passively into the reward area. Second, if the dS was

104 contributing to the storage of procedural memories, well-trained animals with dS lesion would
105 be expected to forget running toward the back region of the treadmill during the intertrial in-
106 terval and, instead, should remain in the reward area, as they naturally did before learning. We
107 found that, following dS lesion, these 9 animals kept going toward the back of the treadmill af-
108 ter reward consumption and remained still while being carried toward the reward area (Figures
109 2B–2D). Importantly, the lack of effect of the dS lesion was not due to the fact that learning the
110 run-and-wait routine was easier than learning the wait-and-run routine (Figure S2). Altogether,
111 and in contrast to the general belief that the dS stores procedural memories, we observed that its
112 lesion spared the ability of rats to remember what to do during the performance of two simple
113 motor routines.

114 **Lesions of the dorsal striatum decreased the speed of reward-oriented lo-** 115 **comotion while sparing basic motor control and reward consumption**

116 Because of the potential role of the dS in controlling the vigor of reward-oriented movements
117 [25, 26], we examined the impact of dS lesions on the speed of the animals when they actively
118 crossed the treadmill toward the reward-area during performance of the wait-and-run routine
119 (Figure 3A, see also trajectory illustrations in Figure 1). For the three groups of animals (DLS,
120 DMS and DS), the lesion induced an immediate and irreversible decrease in running speed (Fig-
121 ures 3B and 3C), an effect that robustly correlated with lesion size (Figure 3D). The maintained
122 task proficiency following dS lesion suggested that this reduction in running speed occurred
123 while the motivation of the animals to obtain rewards (reward-seeking) was largely preserved.
124 Accordingly, animals with a dS lesion kept licking the sweetened water upon correct *ET*, al-
125 though licking initiation was delayed, which could be due to a slower approach of the reward
126 delivery device following routine completion (Figure S3).

127 To better understand the origin of this running speed reduction, we examined whether the dS
128 lesions affected the animals' elementary ability to run at different speeds. First, we compared
129 basic locomotor activity between non-lesioned (control) and lesioned rats, in a paradigm that

130 did not include reward-oriented runs, using a different treadmill. The locomotion test consisted
131 of several trials (30 s long) at fixed speeds (0 to 40 cm/s), interleaved with 30 s long intertrial
132 pauses. We found that both control and lesioned rats displayed similar exploratory locomotor
133 activity when the treadmill remained immobile (Figure 4A). In trials in which the treadmill was
134 turned on, both groups were similarly able to follow a reasonable range of imposed speeds even
135 though, as the speed increased, animals with a dS lesion ran with slightly slower speeds than
136 control animals (Figure 4B). It has been previously shown that the speed of reward-oriented
137 movements increases with movement distance, to minimize temporal discounting of rewards
138 [35, 36]. We compared running speed in trials during which running was initiated from the back
139 versus the middle of the treadmill (Figure 4C, i.e., long vs short run distance). As predicted,
140 animals ran faster when they initiated their runs from the back of the treadmill than from its
141 middle (Figure 4D). This modulation was maintained after dS lesion when the performance
142 of the animals was stable (modulation ratio, Before vs Stable: $p = 0.27$, permutation test),
143 although running speeds were generally slower following dS lesion (see Figure 3C).

144 **Kinematic alterations induced by dS lesion are compatible with an increased** 145 **sensitivity to effort**

146 Given that the dS lesion spared the rats' ability to modulate their running speed, a parsimonious
147 explanation of our results is that lesioned animals "preferred" slower speeds. Interestingly, it
148 has been previously proposed that the slowness in reaching movements observed in PD patients
149 reflects an increased sensitivity to effort which leads them to perform slower (hence less ef-
150 fortful) movements [37]. Could the less vigorous runs toward the reward area observed after
151 dS lesion be also accounted for by an increased sensitivity to effort? To address this question,
152 we took advantage of the optimal control framework that relies on the assumption that animal
153 behavior is optimal with respect to a cost function [38]. We modeled the optimal trajectory of
154 a rat, taking into account costs related to locomotion control (effort) and those imposed by the
155 task rules (running in the front is costly as it can lead to premature *ET*s, which were penalized

156 in our task). The effort-related term $e(t)$ had a quadratic dependence on the instantaneous speed
157 produced by the animal at each time t during a trial. The “spatial” cost was dictated by the task
158 rules $p(x)$ and penalized positions x close to the reward area. Importantly, we computed optimal
159 trajectories in fixed time T ($= 7$ s) with known initial/final positions. This was done because,
160 both before and after lesion, rats initiated trials in the reward area and re-approached it close
161 to the GT . Thus, in a trial of duration T , we assumed that the rats minimized the total cost C ,
162 which was a linear combination of the effort and spatial penalty terms:

$$C = \int_0^T [\alpha e(t) + p(x)] dt \quad (1)$$

163 The parameter α governed the effort sensitivity. Increasing the effort sensitivity term (i.e., α) in
164 the model consistently resulted in optimal trajectories that laid closer and closer to the reward
165 area (Figure 5A). Similar results were obtained when different approximations of the effort-
166 related and spatial costs were combined to create alternative versions of the model (Figure
167 S4A). This result is explained by the fact that, in the context of the wait-and-run routine, running
168 for a relatively longer period of time at a slow speed is cumulatively less costly (i.e., requires
169 less effort) than running for a relatively shorter period of time at a faster speed (Figure 5B).
170 Moreover, an increased sensitivity to effort should also cause lesioned animals to start running
171 earlier in time (and thus in space) toward the reward area. This prediction is interesting at two
172 levels. First, it is reminiscent of the behavior observed during the first post-lesion session when
173 animals with a large dS lesion ran mainly close to the reward area (Figures 1F and 1G, see
174 “acute” column). Second, a careful visual inspection of the rats’ trajectories following DLS,
175 DMS and DS lesions when performance was stable revealed a tendency of the animals to start
176 running earlier toward the reward area, compared to their respective pre-lesion session (Figures
177 1E–1G, compare 2nd and 4th columns). To more directly test whether the alterations in the wait-
178 and-run routine performance is compatible with an increased sensitivity to effort, we used an
179 inverse optimal control approach and adjusted the effort sensitivity parameter α of the optimal
180 control model to fit the animals’ median trajectories (before and after dS lesion, Figure 5C). This

181 analysis was restricted to the animals in which the dS lesion induced a significant reduction
182 in running speed (i.e., 39 rats with relatively large lesions, Figure S4C). This allowed us to
183 estimate the change in effort sensitivity following dS lesion and to observe that its magnitude
184 correlated with lesion size (Figure 5D). Finally, we examined whether, following dS lesion,
185 rats started running earlier (i.e., on a more intermediate portion of the treadmill) toward the
186 reward area, as predicted by the optimal control model (Figure 5A). This was indeed the case
187 and this effect, while being maximal on the first post-lesion sessions, remained significant after
188 three weeks of daily testing and its magnitude correlated with lesion size (Figures 5E and 5F).
189 Finally, comparing the average maximum position in all the trials performed by all the control
190 and lesioned animals confirmed that those with dS lesions initiated their runs earlier (in space
191 and time) toward the reward area (Figure S4D).

192 The optimal control model suggested that remaining very close to the reward area minimizes
193 energy expenditure (effort) by limiting the usage of fast speeds (Figures 5A and 5B). In the first
194 post-lesion session, this type of behavior was observed in animals with the largest lesions (DS
195 group in Figure 1H) but led to premature entrances in the reward area and an abrupt drop in
196 correct trials. Noticeably, despite this initial impairment, most rats with a DS lesion (14 out of
197 16) progressively waited longer and longer and, after a few post-lesion sessions, they recovered
198 a behavior that resembled the one expressed before lesion (i.e., they relied on the wait-and-run
199 routine although its kinematics was altered, Figure 1H). This suggests the intriguing possibility
200 that the dS may not be critical to learn the wait-and-run routine in the first place. Accordingly,
201 we found that all the rats in which DLS or DMS lesions were performed before training learned
202 the waiting task similarly to non-lesioned rats (Figure 6A). A similar result was observed in
203 the majority (14 out 20) of rats in which DS lesions were performed before training (Figure
204 6A). Still, 6 rats with the largest lesions (DS lesion before training, mean lesion size of 75%,
205 Figure 6A, arrow) persisted in running in the reward area throughout the entire training period
206 (Figures 6B and 6C). In some occasions, these animals performed correct wait-and-run routines

207 (Figure 6B), which suggests that the lack of task improvement was not due to a fundamental
208 motor disability. Interestingly, a similar behavior (persistence to remain close to the reward
209 area) was also observed in 2 rats with DS lesion performed after training (lesion size >80%,
210 Figure 1J, note the two lines crossing the x-axis). Thus, the biggest alteration in learning and
211 performing the wait-and-run routine could be accounted for by a persistent and extreme attempt
212 to minimize locomotor costs (i.e., hypersensitivity to effort, Figure 5A). Finally, in animals in
213 which dS lesions were performed before training, a robust negative correlation between their
214 average running speed and their lesion size was observed (Figure 6D). A similar correlation
215 was found when the dS lesion was performed after learning (Figure 6D). Thus, dS lesions
216 performed before and after training affected the performance of wait-and-run routine in a way
217 that is compatible with an increased sensitivity to effort, as lesioned rats initiated earlier their
218 runs toward the reward area (Figure S4D) and used a slower speed.

219 **Discussion**

220 In this study, we report that lesioning the dorsal striatum (dS) irreversibly modified the kine-
221 matic parameters of execution of a motor routine in a way that is parsimoniously explained
222 by an increased sensitivity to effort. Specifically, after dS lesion, animals kept performing the
223 sequential steps of the wait-and-run routine, but they started to run toward the reward area ear-
224 lier (i.e., on a more intermediate portion of the treadmill) and used slower speeds. Despite
225 these kinematic alterations, dS-lesioned animals were able to increase their running speed in a
226 control task during which the treadmill belt moved progressively faster. Moreover, similarly to
227 non-lesioned rats, those with a dS lesion used faster speeds when initiating their reward-oriented
228 runs in the back region of the treadmill than when initiating runs from its middle portion. These
229 two observations suggest that the changes in kinematics caused by the dS lesion did not re-
230 sult from a general impairment in locomotion speed control per se. In addition, following dS
231 lesion, the total number of licks performed during the intertrial interval was stable, indicating

232 that reward-seeking was also globally preserved. Our modeling work using the optimal control
233 framework revealed that the kinematic changes displayed after dS lesion were compatible with
234 an increased sensitivity to effort. Specifically, by using an inverse optimal control procedure to
235 estimate the effort sensitivity of the animals from their trajectories before and after dS lesion,
236 we found that the magnitude of the changes in the effort sensitivity term of the fitted model
237 correlated with the size of the dS lesion. Indeed, in the context of the wait-and-run routine,
238 running for a longer period of time at a slower speed was less costly than remaining most of
239 the time immobile and running briefly at a fast speed. In other words, the same kinematic al-
240 terations were expected to occur had we forced non-lesioned rats to perform the wait-and-run
241 routine with extra weight on their back (Figure S4B). Altogether, our results indicate that the dS
242 is critical for the invigoration of reward-oriented motor routines and suggest that this function
243 is mediated by setting the animals' sensitivity to effort.

244 The original aim of our study was to investigate the contribution of the dS during the exe-
245 cution of a well-learned motor routine and, consequently, we did not systematically manipulate
246 the effort required to perform the task. Still, in agreement with a contribution of the dS in the
247 invigoration of motor sequences, rats were virtually unaffected by the dS lesion when perform-
248 ing another routine which required very little effort (Figure 2). Even if future investigations
249 should examine the contribution of the dS in motor tasks in which both effort and accuracy
250 requirements are parametrically and separately manipulated, it is striking that a similar increase
251 in sensitivity to movement cost with preserved motor function has been proposed to explain
252 slower arm reaching in early-stage Parkinson's disease (PD) patients, in which dopamine de-
253 pletion is primarily affecting dorsal regions of the striatum [37, 39]. A role of dS dopamine in
254 effort sensitivity is directly supported by the observation that striatal dopamine depletion in a
255 mice model of PD is associated with less vigorous reaching movements and decreased locomo-
256 tion [40, 41]. It is also compatible with the fact that striatum-wide manipulation of the type-2
257 dopaminergic receptor modifies the energy expenditure of rats engaged in a foraging task [42].

258 Pharmacological manipulation in behaving rats suggested that dopamine in the ventral striatum
259 is critical for the prolonged exertion of effort during reward-seeking behavior [43], a result that
260 raises the question of the potential difference between the function of dopamine in the ventral
261 and dorsal striatum. On the one hand, dS-lesioned rats kept performing the wait-and-run rou-
262 tine and licked similarly the sweetened solution following successful trials. On the other hand,
263 they ran more slowly toward the reward area and started licking the reward with an increased
264 delay. These results are in agreement with the possibility that dopamine in the ventral striatum
265 may determine the general motivation to work for a reward while, in the dorsal striatum, it may
266 control how much energy is put into the specific actions/motor sequences selected to obtain
267 rewards [44, 45].

268 It is generally believed that the DLS is essential for the long-term storage and automatic
269 retrieval of procedural skills while the DMS is more critical for their early development and
270 goal-directed usage [46, 47]. The wait-and-run routine required rats to remember to (i) remain
271 inside the reward area following reward consumption in the previous trial, (ii) remain immobile
272 following treadmill onset and (iii) run forward without pause once reaching the back of the
273 treadmill. If the DLS function was to store a procedural memory that rats retrieved when per-
274 forming the wait-and-run routine, its lesion after extensive practice (6 weeks of practice, 5 days
275 per week) should have caused major behavioral impairments. This is not what we observed
276 as DLS lesioned rats remained inside the reward area until trial onset and did not forget what
277 to do once they approached the rear portion of the treadmill. Still, we cannot rule out that the
278 preserved routine performance following dS lesion stems for the complex mixture of sensory
279 information (e.g., visual cues) and cognitive skills (e.g., impulse control) engaged during the 7 s
280 long wait-and-run routine, that are likely to depend on neuronal activity distributed across the
281 DMS, the DLS and outside the dS. In support of this possibility, a recent study reported that
282 DLS lesions impaired the performance of idiosyncratic short sequences of limb movements and
283 whole body postural adjustments that rats developed in order to respect a 700 ms long inter-

284 val between two lever presses [17]. Thus, the DLS may store some kind of motor programs
285 that drive the execution of brief stereotyped sequences of innate movements, especially when
286 these sequences are performed in a ritualistic manner. Alternatively, in [17], the fact that rats
287 following DLS lesion reversed to their natural tendency to press the lever twice in a row with a
288 much shorter interval could also be interpreted as an attempt to reduce energy expenditure with
289 preserved procedural memory (i.e., rats remembered to press the lever twice). Similarly, it is
290 interesting to note that the behavioral impairments following DLS perturbation in the acceler-
291 ating rotarod, which constitute one of the primary evidence for DLS contribution to procedural
292 memory [19, 20], can also be accounted for by a decrease in vigor (or increased sensitivity to
293 effort). Taking into account that we also observed that DMS lesions spared the acquisition of
294 the wait-and-run routine (Figure 6A), our results add to a growing body of studies whose con-
295 clusions are difficult to reconcile with a serial contribution of the DMS and then of the DLS to
296 procedural skill learning and performance [17, 48–51] and more specifically, with a role of the
297 sensorimotor regions of the basal ganglia in the storage of procedural memory [52, 53].

298 Our study indicates that dS lesion induced an irreversible reduction of the running speed
299 used to complete the wait-and-run routine, and that the magnitude of this reduction was robustly
300 correlated with lesion size. This result provides a direct support for the hypothesis that the basal
301 ganglia control the vigor of reward-oriented movements [25, 26]. It is particularly interesting
302 that pairing the execution of a reaching movement performed at a specific speed range (e.g.,
303 slow or fast) with a brief stimulation of dS projection neurons forming either the direct or
304 indirect basal ganglia pathway (dSPN or iSPN) increased or decreased the tendency of mice
305 to use this speed range during subsequent reaches, respectively [30]. This result is compatible
306 with a role of the dS related to movement cost optimization if one postulates that activation of
307 dSPNs and iSPNs, respectively, decreased and increased the sensitivity to the effort necessary
308 to execute these fast or slow reaching movements, which in turn affected the preference for one
309 speed range over another in subsequent executions (e.g., increasing the cost of slow movements

310 by pairing them with iSPN stimulation will result in avoiding slow movements and favor faster
311 ones). Importantly, it is possible that the role of the dS in setting effort sensitivity contributes
312 to a general cost-benefit analysis of action plans [54] that will not only affect the speed of
313 movements but also the choice of a particular action versus another (i.e., decision making),
314 hence the expression of procedural memories. For instance, it has been shown in mice well-
315 trained to perform a sequence of 4 presses on two levers (L-L-R-R) that pairing the pressing
316 of the first lever (L) with iSPN optogenetic stimulation caused the animals to avoid repeating
317 this action and switch to the second (R) lever [15]. This result is entirely compatible with an
318 increased cost (effort sensitivity) associated with the first action (press L lever) which will be
319 avoided in favor of the relatively less costly alternative (press R lever). Similarly, a role of the dS
320 in the cost /benefit analysis of action plans may explain the behavioral alterations that resulted
321 from d/iSPN activity manipulations in a variety of experimental settings [13, 20, 24, 55–57].

322 Expending effort to produce faster movements allows limiting the temporal discounting of
323 reward [36, 58]. In sensory guided decision-making tasks, the cost of time can also be reduced
324 by limiting the duration of deliberation [59]. Interestingly, recent evidence supports a role of
325 the basal ganglia in signaling the urgency to commit to a choice [60]. Future studies should
326 investigate whether signaling effort and urgency are the two sides of an underlying function
327 implemented in the basal ganglia to maximize the reward rate while minimizing costs.

328 **Acknowledgements**

329 We thank Drs. I. Bureau, R. Cossart, J. Epsztein, E. Fino and T. Verstynen for critical read-
330 ing of earlier versions of this manuscript, and Dr A. de Chevigny for advice with histological
331 verification of the lesions. This work was supported by the European Research Council (ERC-
332 2013-CoG – 615699 – NeuroKinematics; D.R.), a Centuri postdoctoral fellowship (S.S.), and a
333 nEUro*AMU, ANR-17-EURE-0029 grant (M.S.).

334 **Author contributions**

335 Maria-Teresa Jurado-Parras: Conceptualization, Investigation, Writing–review & editing; Mostafa Safaie:
336 Conceptualization, Software, Visualization, Investigation, Formal analysis, Writing–original
337 draft; Stefania Sarno: Software, Methodology, Formal analysis, Conceptualization, Visual-
338 ization, Writing–review & editing; Jordane Louis: Investigation; Corane Karoutchi: Investi-
339 gation; Bastien Berret: Conceptualization, Methodology, Formal analysis, Writing–review &
340 editing; David Robbe: Funding acquisition, Project administration, Supervision, Visualization,
341 Writing–original draft, Conceptualization, Software.

342 **Declaration of Interests**

343 The authors declare no competing interests.

344 **Figure Legends**

345 **Figure 1. Preserved performance a wait-and-run motor routine following dS lesion. (A)**
346 Experimental apparatus and task rules. **(B)** Entrance time across training sessions for all the
347 rats trained in this task. Shaded area represents the interquartile range. **(C)** Trajectories of
348 an example animal on the treadmill, for all the trials performed during sessions #1 and #30
349 (left). Percentage of trials during which animals performed the wait-and-run routine, across
350 sessions (right). **(D)** Running speed when animals ran toward the reward area, across sessions.
351 Triangles in B to D indicate the performance for the example animal shown in C (left). **(E-G)**
352 Histology (1st column, GFAP in green shows gliosis, red is NeuN) and trajectories of single
353 animals with bilateral lesions of the dorsolateral, dorsomedial and entire dorsal striatum (E:
354 DLS, F: DMS, G: DS). # indicates session number relative to lesion break. **(H-I)** Time course
355 of the lesion effect on *ET* (H) and percentage of routine usage (I). **(J-K)** Similar to H-I, group
356 statistical comparison between different times relative to lesion and/or learning stages (10,000
357 resamples, *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$ and *n.s.* indicates non-significant $p \geq 0.05$
358). Trajectories in C, E, F and G are cut after *ET*. Error bars in H-K (and throughout the figures)
359 indicate the median range (25th to 75th percentiles) across animals. See also Figure S1 and
360 Video S1.

361 **Figure 2. Preserved performance of a run-and-wait routine following dS lesion. (A)** Tra-
362 jectory of a proficient animal trained in a version of the waiting task in which the belt moved
363 toward the reward area (rather than away from it). 9 consecutive trials (shaded areas) and in-
364 tertrials (white areas) are shown. **(B)** Trajectories from a single representative animal in two
365 sessions before and two sessions after lesion. **(C-D)** Comparison of *ET* (C) and percentage
366 of run-and-wait (reverse) routine usage (D), in different stages of training. Same significance
367 symbols as in Figure 1. See also Figure S2 and Video S2.

368 **Figure 3. dS lesions induced a robust decrease in the speed at which rats ran toward**
369 **the reward.** (A) Position on the treadmill (blue) and speed relative to the treadmill belt (red)
370 for a typical example of wait-and-run routine performance. (B) Time course of the dS lesion
371 effect on running speed. Same color code for lesion types and same animals as in Figure 1.
372 (C) Similar to B, group statistical comparison between different times relative to lesion and/or
373 learning stages. Same significance symbols as in Figure 1. (D) Average change in running
374 speed (speed after lesion – speed before lesion) versus lesion size, for 53 rats that received a
375 dS lesion after training. See also Figure S3.

376 **Figure 4. Preserved spontaneous locomotor activity and modulation of running speed**
377 **following dS lesion.** (A) Distance ran while exploring a new (and immobile) treadmill for
378 non-lesioned (control) and lesioned rats ($n = 12$ in both groups, same color code for lesion
379 types as in Figure 1). (B) Average running speed in a free running task (no reward) in which
380 control and lesioned rats were submitted to trials with incremental treadmill speed (same color
381 code and same animals as in A). Golden lines indicated significant differences between groups
382 (corrected for multiple comparisons). (C) Trials were split into 2 categories depending on
383 whether rats initiated their run from the middle or back portion of the treadmill. Speed was
384 computed and averaged across trial types. (D) Speed of the runs initiated from either the middle
385 or back portion of the treadmill, and calculated for each animal over the last 5 sessions before
386 lesion (left) and the last 5 stable sessions after lesion (right).

387 **Figure 5. Kinematic changes after dS lesion are compatible with an increased effort sen-**
388 **sitivity.** (A) Optimal trajectories predicted by an optimal control model (see main text for de-
389 tails) with increasing effort sensitivity. (B) Predicted total cost of 5 different trajectories when
390 effort sensitivity is fixed ($\alpha = 10$). The positions, speeds and cumulative costs during trial
391 are shown from left to right. (C) Top, Trajectories of a single animal during a single session
392 before and after dS lesion. Only trials in which the routine was executed (thin blue lines) were

393 taken into account to find the trajectory of the trial with the median maximum position (bold
394 blue line). Bottom, median trajectories (same as upper row) and best model fit (dashed brown).
395 Effort sensitivity (α) values of the best model are shown. **(D)** Effect of dS lesion on α vs size
396 of the lesion. **(E)** Effect of dS lesion on the maximum position of routine trials. **(F)** Change in
397 maximum position (difference in average Max. Pos. of the last 5 sessions before lesion and the
398 first 5 stable sessions after lesion) versus lesion size. Analyses shown in D, E, F were restricted
399 to animals in which the dS lesion induced a reduction in running speed (see Figure S4C). In D
400 and F, compared to E, 2 animals were removed due to incomplete histological quantification of
401 the lesion (see Methods). Same color code for lesion types as in Figure 1. Green triangles in
402 D, E and F are data points from the example animal whose trajectories and model fit, before
403 and after lesion, are shown in C. See also Figure S4.

404 **Figure 6. Effect of DLS, DMS and DS lesions performed before training.** **(A)** Session-by-
405 session change in performance (*ET*, upper panels; Percentage of trials in which the routine was
406 performed, lower panels) for animals without lesion (Control, left) and for animals that received
407 a lesion before training (DLS, DMS, DS from left to right). Black lines indicate Control group
408 median. Thin colored lines indicate single animals. Thick colored lines (same color code as in
409 Figure 1) in 3 rightmost columns indicate group performance for comparison (8 lesion animals
410 with fewer than 30 training sessions are not shown, which explains the difference in the number
411 of animals with Figure S1C). Horizontal golden lines indicate significant differences between
412 control and lesion groups (corrected for multiple comparisons). **(B)** Trajectories before and after
413 extensive training (sessions #1 and #30) for two animals with large DS lesions. Note that, after
414 extensive practice, Rat238 was capable of performing the wait-and-run routine. **(C)** Percentage
415 of trials in which animals remained in the front region of the treadmill (computed for sessions
416 #25 to #30) versus lesion size. Dots with black circles represent the 6 animals (arrow in the
417 upper right panel in A that remained in the reward area for the majority of the trials across
418 their first 30 training sessions). **(D)** Correlation between speed of individual rats when they ran

419 forward toward reward area and lesion size. Animals in which lesion was performed before
420 (early group) and after (late group) learning the task were pooled together. Separate correlation
421 results for each group: early group: $r = -0.58$, $p = 1.35 \times 10^{-6}$; late group: $r = -0.76$,
422 $p = 2.82 \times 10^{-11}$. Same color code as in Figure 1. Speed was computed for all the routine trials
423 performed during the last 5 sessions in which performance was stable.

424 **STAR METHODS**

425 **RESOURCE AVAILABILITY**

426 **Lead Contact**

427 Further information and requests for resources and reagents may be directed to and will be
428 fulfilled by the Lead Contact, David Robbe (david.robbe@inserm.fr).

429 **Materials Availability**

430 This study did not generate new unique reagents.

431 **Data and Code Availability**

432 The Jupyter Notebooks and the raw data necessary for the full replication of the figures is pub-
433 licly available via the Open Science Foundation in the following repository: <https://osf.io/2vmus/>

434 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

435 A total of 166 male Long-Evans rats were used in this study (number for each experimental
436 condition is directly reported in the figures). They were 12 weeks old at the beginning of the ex-
437 periments, housed in groups of 4 rats in temperature-controlled ventilated racks and kept under
438 a 12 h–12 h light/dark cycle. All the experiments were performed during the light cycle. Food
439 was available *ad libitum* in their homecage. Rats had restricted access to water while their body
440 weights were regularly measured. No animal was manually excluded from the analysis. All
441 experimental procedures were conducted in accordance with standard ethical guidelines (Euro-
442 pean Communities Directive 86/60 - EEC) and were approved by the relevant national ethics
443 committee (Ministère de l'enseignement supérieur et de la recherche, France, Authorizations
444 #00172.01 and #16195).

445 **METHODS DETAILS**

446 **Apparatus**

447 Four identical treadmills were used for the experiments. Each treadmill was placed inside a
448 sound-attenuating box. Treadmills were 90 cm long and 14 cm wide, surrounded by plexiglass
449 walls such that the animals were completely confined on top of the treadmill conveyor belt.
450 The treadmill belt covered the entire floor surface and was driven by a brushless digital motor
451 (BGB 44 SI, Dunkermotoren). The front wall (relative to the turning direction of the belt) was
452 equipped with a device delivering drops of 10% sucrose water solution (maximal drop size
453 $\sim 80 \mu\text{L}$). An infrared beam, located at 10 cm of this device, defined the limit of the reward
454 area. A loudspeaker placed outside the treadmill was used to play an auditory noise (1.5 kHz,
455 65 db) to signal incorrect behavior (see below). Two strips of LED lights were installed on the
456 ceiling along the treadmill to provide visible and infrared lighting during trials and intertrials,
457 respectively (see below). The animals' position was tracked via a ceiling-mounted camera
458 (Imaging Source, 25 fps). A custom-made algorithm detected the animal's body and recorded
459 its centroid to approximate its position on the treadmill. After trial onset, the first interruption
460 of the beam was registered as entrance time in the reward area (*ET*). The entire setup was fully
461 automated by a custom-made program (LabVIEW, National Instruments). The experimenter
462 was never present in the behavioral laboratory during the experiments.

463 **Habituation**

464 Animals were handled 30 min per day for 3 days, then habituated to the treadmill for 3 to 5
465 daily sessions of 30 min, while the treadmill's motor remained turned off and a drop of sucrose
466 solution was delivered every minute. Habituation sessions resulted in systematic consumption
467 of the reward upon delivery and a tendency of the animals to spend more time in the reward
468 area.

469 **Main Behavioral Task**

470 Training started after handling and habituation. Each animal was trained once a day, 5 times
471 a week (no training on weekends). Each of the daily sessions lasted for 55 min and contained
472 ~ 130 trials. Trials were separated by intertrial periods of 15 s. During intertrials, the treadmill
473 remained dark and infrared ceiling-mounted LEDs were turned on to enable video tracking
474 of the animals. Position was not recorded during the last second of the intertrials to avoid
475 buffer overflow of our tracking routine and allow for writing to the disk (see the gaps in the
476 position trace in Figure 2A). The beginning of each trial was cued by turning on the ambient
477 light, 1 s before motor onset. Since animals developed a preference for the reward area during
478 habituation, the infrared beam was turned on 1.5 s after trial start. This delay was sufficient
479 to let the animals be carried out of the reward area by the treadmill, provided they did not
480 move forward. After the first 1.5 s, the first interruption of the beam was considered as ET .
481 The outcome of the trial depended solely on the value of the ET , compared to the goal time
482 ($GT = 7$ s). In a correct trial ($ET \geq GT$), an infrared beam crossing stopped the motor, turned
483 off the ambient light, and triggered the delivery of reward. In an error trial ($ET < GT$), there
484 was an extended running penalty. During the penalty, the motor kept running, the ambient light
485 stayed on and an audio noise indicated an error trial. The duration of the penalty period was
486 anticorrelated with the magnitude of error, between 1 s and 10 s (see [34] for more details). In
487 trials wherein animals didn't cross the beam in 15 s since trial onset, trial stopped and reward
488 was not delivered.

489 The magnitude of the reward was a function of the ET and animal's performance in previous
490 sessions (only in early training). Reward was maximal at $ET = GT$ and dropped linearly
491 to a minimum (= 38% of the maximum) for ET 's approaching 15 s (i.e., the maximum trial
492 duration). Moreover, in the beginning of the training, partial reward was also delivered for
493 error trials with $ET > ET_0$, where ET_0 denotes the minimum threshold for getting a reward.
494 The magnitude of this additional reward increased linearly from zero for $ET = ET_0$, to its

495 maximum volume for $ET = GT$. In the first session of training, $ET_0 = 1.5$ s and for every
496 following session, it was updated to the maximum value of median ET s of the past sessions.
497 Once ET_0 reached the GT , it was not updated anymore.

498 **Reverse Treadmill Task**

499 This task differed from the normal task in three critical properties: 1) the treadmill moving
500 direction was reversed, i.e., the conveyor belt moved toward the reward port; 2) the treadmill
501 speed was set at 8 cm/s (instead of 10 cm/s) to ensure that starting the trial in the back of the
502 treadmill and remaining still after trial onset would be rewarded, i.e., $ET \geq 7$. 3) the intertrial
503 duration was 20 s, instead of 15 s, to allow sufficient time for the animals to move to the back
504 of the treadmill while the motor was still off.

505 **Locomotor Activity Test**

506 A group of animals with a striatal lesion (7 DLS, 2 DMS, and 3 DS), and another group of non-
507 lesioned animals ($n = 12$) were used in this test to assess their general locomotor activity. Prior
508 to this task, animals had full access to water and food for at least 3 days. Then, they were placed
509 on an unfamiliar treadmill, with a different structure (slanted walls and covered reward port)
510 compared to the treadmill in which they were trained, while their position was being recorded
511 using a side-mounted high-speed camera (200 fps). During the first 10 min, the ambient light
512 was turned off and the treadmill remained immobile. Their exploratory locomotor activity, i.e.,
513 how much they moved along the treadmill, during this period, is presented in Figure 4A. Then,
514 in a free-running task, they ran in trials of 30 s while the treadmill speed progressively increased
515 across trials (5 trials at 0 cm/s, 2 trials at 10 cm/s, 3 trials at 15 cm/s and 5 trials at 20, 25, 30,
516 35, and 40 cm/s, data shown in Figure 4B). Each trial was followed by an intertrial (30 s-long),
517 with the ambient light and the treadmill motor turned off. The running speed reported in Figure
518 4B is the average running speed of animals during the trials of any given treadmill speed.

519 **Lesion Surgery**

520 Anesthesia was induced with an intraperitoneal (IP) injection of a mixture of 100 mg/kg ke-
521 tamine and 10 mg/kg xylazine and was maintained during the surgery with inhalant isoflurane
522 gas (less than 3%). After shaving and cleaning the scalp, the animal was placed in the stereo-
523 taxic frame (Kopf Instruments) and a local anesthetic (lidocaine) was injected under the scalp.
524 Then, an incision along the midline of the skull was made, allowing for cleaning the exposed
525 skull and drilling the craniotomies above the targeted areas. To perform fiber-sparing lesion of
526 the striatum, ibotenic acid (1% in 0.1M NaOH, Fisher Scientific) was infused (Pump 11 Elite
527 Nanomite, Harvard Apparatus, using a 10 μ L WPI Nanofil syringe) in 6 sites symmetrically ar-
528 ranged relative to the midline (i.e. 3 sites in each hemisphere), at a rate of 90 nl/min. Injection
529 coordinates (in mm, with reference to Bregma, according to Paxinos) are shown in Figure S1A
530 (each injection at -5.6 dorsoventral). The infused volume in each site was 200 nL for DLS and
531 DMS lesions, and 400 nL for DS lesions. The needle remained in place for 10 min following
532 the injection to allow for the diffusion of the excitotoxin. Then the needle was retracted slowly
533 to avoid backflow of the drug. Once all the injections were performed, craniotomies were filled
534 with bone wax, the skull was disinfected and the skin was sutured. Animals were allowed to
535 recover for two weeks before resuming behavioral training. After surgery, animals were housed
536 alone for 3 days in a warmed cage, to avoid getting hurt by their cagemates, and were force-
537 fed if needed. Post-surgery pain was reduced as much as possible using an opiate painkiller
538 (Buprenorphine) and if necessary a non-steroidal anti-inflammatory drug (Carprofen).

539 **Histology**

540 Animals were euthanized with an overdose IP injection of 2 mL pentobarbital or with an in-
541 jection of Zoletil (40mg/kg) and Domitor (0.6 mg/kg). Then, they were perfused with 4%
542 paraformaldehyde and their brains were harvested for histological analysis of the lesion size
543 and location. Brains were coronally sliced on a vibratome at 60 μ m thickness. For each animal,

544 six sections spanning the dorsal striatum along the rostrocaudal axis were selected (usually the
545 following slice numbers: 5, 15, 25, 35, 45, and 55 for consistency) and submerged in a 0.1 M
546 PBS solution. Then, PBS was replaced with citrate buffer (10 mM) for 10 min at room tem-
547 perature. Next, slices were submerged with a blocking solution, consisting of PBS with 0.3%
548 Triton and 15% normal goat serum (NGS), for 2 hrs at room temperature. Then the solution was
549 replaced with a solution consisting of 2 uL anti-NeuN antibody (Merks Millipore, MAB377)
550 and 0.5 uL of anti-GFAP antibody (Agilent, Z033429-2) diluted in 200 uL of the blocking
551 solution, kept overnight at 4°C. Sections were then rinsed twice in PBS for 10 min at room
552 temperature, before being resubmerged in 1 uL of donkey-anti-mouse antibody (AL555, red),
553 1 uL of donkey-anti-rabbit antibody (AL488, green) diluted in 400 uL of PBS for 2 hr at room
554 temperature. Finally, they were washed twice in PBS for 10 min and mounted for microscopy.

555 **Optimal Control Model Description**

556 The optimal control (OC) theory relies upon the assumption that behavior is governed by op-
557 timality principles with respect to some cost function [38]. This means that, when making a
558 decision and generating movements, the default tendency of the brain is to maximize reward
559 while minimizing effort. Here, we aimed at simulating the optimal trajectories of rats while
560 they performed the wait-and-run motor sequence and when their sensitivity to effort was ma-
561 nipulated. Critically, because striatal lesions changed the kinematics of the motor sequence
562 execution, we wanted to test the hypothesis that these changes in kinematics were compatible
563 with an increased sensitivity to effort. After lesion, rats kept initiating the trial in the reward
564 area and arrived in the reward area close to the Goal Time ($GT = 7s$, Figure 1). We thus
565 computed the optimal trajectory in fixed time T ($= 7$ s) with known initial/final states, given a
566 system dynamics and cost.

567 **Equation of Motion and Constraints.** We assumed that the speed of a rat v satisfies the
 568 following equation of motion[61]:

$$\frac{dv}{dt} = u(t) - \frac{v}{\tau} \quad (2)$$

569 The term v/τ in the above equation is a resistive force per unit mass . The term τ is a friction co-
 570 efficient (when τ increases resistance decreases) and $u(t)$ is the propulsive force per unit mass.
 571 In all the simulation we set $\tau = 1.8$ s. Qualitatively, the results we obtained are independent of
 572 τ . A key component of the behavioral task is that, to obtain a reward, animals must enter the
 573 reward area after GT . This reward area is delimited by the front wall of the treadmill equipped
 574 with a reward port and an infrared-beam located at 10 cm of the front wall. In agreement with
 575 the behavioral data, the initial and final positions of the animal in the model were both equal to
 576 the beam position $x_b = 10$ cm. The position of the rat is constrained by the treadmill length,
 577 which is $L_T = 90$ cm. The treadmill pushes the animal toward the rear wall of the treadmill
 578 (located at position 90 cm) with a constant positive velocity of $v_T = 10$ cm/s. In the coordinate
 579 adopted in the model, running toward the reward area is, therefore, occurring with negative ve-
 580 locity. The velocity of the rat was also constrained to be negative, i.e. $v_{max} = 0$ cm/s (the rat
 581 cannot run toward the rear wall of the) and its module could not be bigger than 70 cm/s, i.e.
 582 $v_{min} = -70$ cm/s (the rat cannot run toward the reward area at a speed faster than 70 cm/s).
 583 We also constrained the maximum speed of the rat at the moment of beam crossing to be less
 584 than 40 cm/s, i.e. $v_C^{min} = -40$. Finally, the module of the acceleration was forced to be less than
 585 180 cm/s². We relied on experimental data to constrain the dynamical variables (i.e. speed and
 586 acceleration) used in the model.

587 We defined the components of the state vector $\mathbf{x} = [x^0, x^1]$, respectively as the position
 588 and speed of the rat in the laboratory frame of reference. Their dynamics is governed by the
 589 equations:

$$\begin{cases} \dot{x}^0 = x^1 \\ \dot{x}^1 = u - (x^1 - v_T)/\tau \end{cases} \quad (3)$$

590 The state variables are constrained by the following initial conditions and inequalities:

$$\begin{cases} x^0(0) = x^0(T) = x_b \\ x^1(0) = v_T \\ v_C^{min} + v_T \leq x^1(T) \leq 0 \\ 0 \leq x^0(t) \leq L_T \quad \forall t \in [0, T] \\ v_{min} + v_T \leq x^1(t) \leq v_T \quad \forall t \in [0, T] \end{cases} \quad (4)$$

591 **Control Variable and Cost Function.** In the OC framework, we assumed that a rat modulates
 592 its propulsive force on a moment-to-moment basis, so $u(t)$ is the control. The infinitesimal
 593 energetic cost $c(t)$, for a rat of mass m , is assumed to be the linear combination of two terms.
 594 The first term is an effort-related term $e(t)$ that is either proportional to the kinetic energy or
 595 has a quadratic dependence on force (i.e., on the control):

$$e(t) = \begin{cases} m(x^1)^2 & (\text{Effort} \approx \text{Kinetic energy}) \\ (mu)^2 & (\text{Effort} \approx \text{Force}^2) \end{cases} \quad (5)$$

596 The second term is a cost related to the task rules, namely that running in the reward area
 597 before GT is associated with a penalty. The modeled trajectory must respect the following rule:
 598 $x^0(t) > x_b \quad \forall t < T$ (x_b being the position of the infrared beam). We modelled this "spatial"
 599 cost, using a differentiable approximation of the Heaviside function of height $A = 10$:

$$p(x^0) = \frac{A}{1 + e^{1+k(x^0-x_b)}} \quad (6)$$

600 The parameter k governs the steepness of this spatial cost . We used $k = 100$ and $k = 1$ to
 601 model the spatial cost in a way that is either localized or diffuse, respectively. The rationale for
 602 doing so is that it is difficult to precisely know how the animals perceived the risk of running
 603 close to the reward area.

604 Therefore, in a trial of duration T , the rat wants to minimize the total cost C :

$$C = \int_0^T [\alpha e(t) + p(x)] dt \quad (7)$$

605 The parameter α governs the effort sensitivity. In the simulations showing the effect of effort
 606 sensitivity on the optimal trajectory (Figure 5A and Figure S4A), we fixed the mass of the

607 animal to $m = 1$ and used six values for the parameter α , namely $\alpha = \{0.1, 1, 2, 5, 10, 100\}$.
608 High values of α correspond to high effort sensitivity. In the simulations showing the effect
609 of the rat's weight on the optimal trajectory (Figure S4B) we fixed $\alpha = 1$ and varied the mass
610 according to $m = \{0.1, 1, 2, 5, 10\}$. Finally, to compute the cumulative cost for the 5 example
611 trajectories in Figure 5B, we assumed that effort was proportional to the kinetic energy, that the
612 spatial cost was diffuse and we set $m = 1$ and $\alpha = 10$.

613 It is worth to note that, combining Equation 5 and 7, the coefficient multiplying the effort
614 term in the cost function is proportional either to αm or to αm^2 (respectively when the effort
615 is proportional to the kinetic energy or to the force). Therefore, considering that the dynamics
616 in Equation 2 is independent of the mass m , in the optimal control model, varying the mass is
617 formally equivalent to varying the effort sensitivity. This equivalence explains the similarity of
618 the results obtained in Figures S4A and S4B.

619 **Numerical Implementation.** We used the *CasAdi* [62] software and direct collocation method
620 to numerically find the optimal trajectories. In all simulations, to obtain the optimal trajectory,
621 we used 200 collocation points.

622 **Inverse Optimal Control Procedure.** We use numerical inverse optimal control techniques
623 to estimate the effect of lesion on the effort sensitivity parameter. This approach allows to au-
624 tomatically recover the cost function (i.e. the effort parameter in this case) from experimental
625 trajectories, which are assumed to be optimal [63]. The inverse optimal control method was
626 implemented as follows: first, for each animal, we computed the trajectory corresponding to
627 the median max position over the last 5 sessions before lesion and the last 5 stable sessions
628 after lesion (same trial inclusion as in Max. Pos. analysis). Then, for each animal and for
629 each session, we adjusted the effort parameter α such as to minimize the mean squared error
630 between this trajectory and the optimal trajectory generated with the OC model. We assumed
631 that effort was proportional to the kinetic energy, that the spatial cost was diffuse and we set

632 $m = 1$, as in Figure 5A. Simulated trajectories were constrained to have the same initial po-
633 sition, final position and entrance time of the corresponding real trajectory. The minimization
634 process was achieved via the Trust-Region Constrained Algorithm implemented in the SciPy
635 optimize module. The effort parameter was constrained to values between 0 and 100. Note that
636 the optimal trajectories for $\alpha = 0.1$ and for $\alpha = 100$ correspond respectively to the blue line
637 and the red line in Figure 5A. For each animal, once the best-fit α parameter for each session
638 is obtained, we computed the difference $\Delta\alpha$ between its median value after and before lesion
639 (denoted respectively $\langle \alpha_A \rangle$ and $\langle \alpha_B \rangle$). The relation between lesion size and change
640 in effort sensitivity was quantified using the Pearson’s correlation coefficient. In the correla-
641 tion analysis presented above, fixing $m = 1$ implicitly corresponds to assume that the weight
642 differences between rats are negligible. To verify if the correlation obtained is independent of
643 this assumption, we repeated the correlation analysis using a normalized version of $\Delta\alpha$ defined
644 as $\Delta\alpha_N = \Delta\alpha / (\langle \alpha_B \rangle + \langle \alpha_A \rangle)$. We obtained again a positive significant correlation
645 between $\Delta\alpha_N$ and the lesion size ($r = +0.60$ and $p = 5.08 \times 10^{-5}$).

646 **QUANTIFICATION AND STATISTICAL ANALYSIS**

647 Data from each behavioral session was stored in separate text files, containing position infor-
648 mation, entrance times, treadmill speeds, and all the task parameters. Position information was
649 then smoothed (Gaussian kernel, $\sigma = 0.3$ s). The entire data processing pipeline was imple-
650 mented in Python, using open-source libraries and custom-made scripts. We used a series of
651 Jupyter Notebooks to process, quantify, and visualize every aspect of behavior, to develop and
652 run the optimal control simulations, and to generate all the figures in this manuscript.

653 **Motor Routine Definition**

654 We quantified the percentage of trials in which animals performed the wait-and-run motor rou-
655 tine in each session (Figure 1I). A trial was considered *routine* if all the following three condi-
656 tions were met: 1) the animal started the trial in the front (initial position $< 30cm$); 2) the ani-

657 mal reached the rear portion of the treadmill during the trial (maximum trial position $> 50\text{cm}$);
658 3) the animal completed the trial (i.e., it crossed the infrared beam).

659 **Speed Calculation**

660 Unless otherwise stated, speed in this manuscript refers to the velocity with which animals
661 crossed the treadmill toward the reward port. For every trial, it is calculated based on the time
662 the animal takes to run from 60 cm to 40 cm along the treadmill (Figure 3A). Speed for each
663 training session is the average speed across its trials. Similar effects of dS lesion were obtained
664 when we used alternative ways of computing their speed when crossing the treadmill toward
665 the reward area.

666 **Reverse Routine Definition**

667 Percentage of reverse (run-and-wait instead of wait-and-run) routine trials is analogous to the
668 percentage of routine trials, except that it is performed in the version of the task in which the
669 treadmill belt is moving toward the reward. A trial was considered *reverse routine* if the follow-
670 ing conditions were met: 1) the animal started the trial in the back region of the treadmill(initial
671 position $> 60\text{cm}$); 2) the animal completed the trial (i.e., it crossed the infrared beam).

672 **Definition of Frontal Trials**

673 Frontal trials are defined as trials in which the animal remained in the frontal portion of the
674 treadmill (i.e., position $< 30\text{ cm}$) for the first 5 s after trial onset.

675 **Speed Modulation Analysis**

676 In Figure 4D, we split the trials in which rats performed the wait-and-run routine according to
677 the position at which they started to run forward (Figure 4C). Runs were either initiated in the
678 middle of the treadmill (Mid, maximum position between 40 and 60 cm) or in the back region
679 of the treadmill (Back, maximum position between 70 and 90 cm). The data were pooled from
680 the last 5 sessions before lesion (Figure 4D, left) and the last 5 stable sessions after the lesion

681 (Figure 4D, right). To improve the reliability, animals were discarded if they did not have at
682 least 10 trials in the Mid and 10 trials in the Back condition (trials that strictly followed the
683 wait-and-run routine, their maximum position was within the range, and for which the speed
684 could have been defined). The fewer number of animals in the left panel was due to the fact that
685 most animals before lesion performed the wait-and-run routine by going all the way to the rear
686 portion of the treadmill, thus they did not have enough Mid trials.

687 **Max. Pos. Quantification**

688 The maximum position an animal reached along the treadmill before initiating the run epoch
689 toward the reward in the wait-and-run routine was quantified as Max. Pos. in Figure 5E. There-
690 fore, Max. Pos. was only calculated for trials that strictly followed a wait-and-run routine, i.e.,
691 total immobility followed by continuous running until reaching the infrared beam. A trial was
692 included if the following conditions were met: 1) the animal started the trial in the front (initial
693 position $< 30\text{cm}$); 2) the animal remained still while being pushed backward by the treadmill
694 (any movement shorter than 0.1 s or slower than 5 cm/s was ignored to avoid trial rejection due
695 to jitter in position tracking); 3) the animal performed an uninterrupted running epoch (staying
696 immobile/moving backward for a duration shorter than 0.1 s was ignored to avoid trial rejection
697 due to jitter in position tracking); 4) the animal completed the trial (i.e., it crossed the infrared
698 beam). Notice that compared to the definition of the routine trials, the threshold for maximum
699 position is relaxed to allow the detection of trials with a reduced maximum position. To increase
700 the reliability, sessions with fewer than 5 trials for which Max. Pos. could be defined were ex-
701 cluded from further analysis. The reported value of Max. Pos. for each session is the average
702 value across its trials (Figure 5E). Furthermore, in Figure 5E, we only analyzed the animals
703 that had an effect on their running speed after dS lesion (shown in Figure S4C). Animals were
704 assigned to the $\Delta\text{Speed} < 0$ group if the average speed of 5 consecutive stable sessions after
705 the lesion (i.e., session +8 to +13) were lower than that of 5 consecutive sessions before the
706 surgery (i.e., sessions -5 to -1). Similar effects of dS lesion on Max. Pos. (Figures 5E and

707 5F) were obtained if the trial inclusion criteria were more restrictive.

708 **Normalizing Speed and Max. Pos.**

709 In Figures 5E and S5C, to normalize each animal's performance according to its own behavior
710 prior to the lesion, behavioral measures of individual animals (Max. Pos. in Figure 5E and
711 speed in Figure S4C) were subtracted from the median value of the respective measure during
712 the 5 last sessions before lesion. Animals were included in this analysis only if these behavioral
713 measures (Max. Pos and speed) could be defined in at least half of the sessions analyzed. A few
714 animals were discarded: 2 rats from the speed analysis (57 rats in Figure 1 compared to 44+11
715 rats in Figure S5C) and 3 rats from the Max. Pos. analysis (41 rats in Figure 5E compared to 44
716 rats with reduced speed in Figure S5C). Those rats either ran toward the back of the treadmill
717 before lesion (Max. Pos. could not be computed) or remained in the reward area (speed could
718 not be computed).

719 **Licking Analysis**

720 Licking was recorded for 12 animals (Figure S3). For rewarded trials, to calculate the peak lick
721 frequency (Figure S3E), lick rate was measured with a sliding 0.5 s window (90% overlap).
722 Then the average lick rate per session was calculated and the maximum value was reported
723 as peak lick frequency. Moreover, lick duration was calculated as the longest interval during
724 which lick rate was at least 10 percent of peak lick frequency (similar results were obtained
725 with thresholds between 5% and 15%).

726 **Quantification of Lesion Size**

727 Whole slices were imaged using an Apotome microscope (Zeiss), and stitched together in the
728 processing software (Zeiss Zen). Then, for each slice, ventricle, striatum, and lesioned area
729 were manually outlined (Figure S1B) bilaterally in the image processing software (ImageJ,
730 Fiji). This procedure was performed blindly to the behavioral results. The size and the centroid

731 coordinates were automatically computed for all of the above-mentioned areas. The anteropos-
732 terior location of each slice was also approximated according to the rat brain atlas (Paxinos).
733 The lesion size reported in this paper is the ratio of the lesion volume over the volume of the
734 striatum. The region of interest was approximated as a truncated cone between any two con-
735 secutive sections, and the volume was accordingly calculated. The type of lesion (DLS, DMS
736 or DS) was also determined visually and confirmed by comparing the centroid location of the
737 lesion to that of the entire striatum (see Figure S1). Animals with a DLS lesion in one hemi-
738 sphere and a DMS in another ($n = 7$) were excluded from this manuscript. In four animals
739 with a dS lesion performed after learning the task, the lesion size quantification could not be
740 properly performed (e.g., slices with poor histological labelling). These animals were classi-
741 fied according to their injection coordinates in the surgery (3 DLS and 1 DMS), however, they
742 were excluded from any analyses that required the lesion size (hence the difference between the
743 number of animals in Figure 3D and the total number in Figure 1).

744 **Statistics**

745 All statistical comparisons were performed using resampling methods (permutation test and
746 bootstrapping, in every case, $n = 10000$ iterations). The permutation test used in this manuscript
747 has already been described and implemented [34]. In brief, to test if two groups of data points
748 (e.g., speed before vs speed after lesion) are different, random reassignment of data points to
749 surrogate groups should not generate group differences similar to that of the original groups.
750 By repeating this process over and over and building a distribution of surrogate group differ-
751 ences, we estimated the likelihood of the original group difference to belong to this distribution.
752 This test was used to compare two independent groups, controlling for multiple comparisons,
753 in Figure 4B and Figure 6A. A similar permutation test was also used to compare two sets of
754 unpaired data points in Figure 4A and Figure S4D.

755 For paired comparisons (Figure 1J and 1K, Figure 2C and 2D, Figure 3C, Figure 4D, Figure
756 5E, Figure S1E to S1G, Figure S3B to S3E), we generated the bootstrap distribution of mean

757 differences ($n = 10000$ with replacement). Significance was reported if 95% Confidence Inter-
758 val (CI) of the pairwise differences differed from zero (i.e., zero was not within the CI). Then,
759 the p-value was reported as the fraction of samples in the bootstrap distribution with their sign
760 opposite to that of the CI. For example, in Figure 4D, right, the 95% CI of pairwise differ-
761 ences is $(-9.45, -15.25)$. Since this interval does not contain zero, it is reported significant,
762 and since none of the resamples were positive (as opposed to the CI), so $p < 0.0001$, i.e., less
763 than one chance in 10000 resamples. On the other hand, in Figure 1K (Before vs Stable), the
764 CI is $(-0.04, 0.03)$ which includes zero, and hence the difference between these two groups of
765 data is reported to be non-significant. Statistical differences were reported using the following
766 notation: *** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$, and *n.s.* for not significant.

767 Finally, the relation between lesion size and magnitude of behavioral impairment was quan-
768 tified using the Pearson's correlation coefficient.

Supplemental Videos

Video S1. Proficient performance of the wait-and-run routine. Related to Figure 1.

Video S2. Proficient performance of the run-and-wait routine. Related to Figure 2.

References

1. Squire, L.R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory* 82, 171–177.
2. Graybiel, A.M. (2008). Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* 31, 359–387.
3. Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.
4. Doyon, J., Laforce Jr, R., Bouchard, G., Gaudreau, D., Roy, J., Poirier, M., Bédard, P.J., Bédard, F., and Bouchard, J.P. (1998). Role of the striatum, cerebellum and frontal lobes in the automatization of a repeated visuomotor sequence of movements. *Neuropsychologia* 36, 625–641.
5. Lehéricy, S., Benali, H., Van de Moortele, P.F., Pélégriani-Issac, M., Waechter, T., Ugurbil, K., and Doyon, J. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proceedings of the National Academy of Sciences* 102, 12566–12571.
6. Packard, M.G. and McGaugh, J.L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of learning and memory* 65, 65–72.
7. Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., and Graybiel, A.M. (1999). Building neural representations of habits. *Science* 286, 1745–1749.
8. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437, 1158.

9. Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., and Rand, M.K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Experimental brain research* *115*, 1–5.
10. Jin, X. and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* *466*, 457.
11. Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* *494*, 238.
12. Jin, X., Tecuapetla, F., and Costa, R.M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature neuroscience* *17*, 423–430.
13. Tecuapetla, F., Jin, X., Lima, S.Q., and Costa, R.M. (2016). Complementary contributions of striatal projection pathways to action initiation and execution. *Cell* *166*, 703–715.
14. 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*, 472–479.
15. Geddes, C.E., Li, H., and Jin, X. (2018). Optogenetic editing reveals the hierarchical organization of learned action sequences. *Cell* *174*, 32–43.
16. Sheng, M.j., Lu, D., Shen, Z.m., and Poo, M.m. (2019). Emergence of stable striatal D1R and D2R neuronal ensembles with distinct firing sequence during motor learning. *Proceedings of the National Academy of Sciences* *116*, 11038–11047.
17. Dhawale, A.K., Wolff, S.B., Ko, R., and Ölveczky, B.P. (2019). The basal ganglia can control learned motor sequences independently of motor cortex. *bioRxiv* , 827261.

18. Dang, M.T., Yokoi, F., Yin, H.H., Lovinger, D.M., Wang, Y., and Li, Y. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proceedings of the National Academy of Sciences* *103*, 15254–15259.
19. Yin, H.H., Mulcare, S.P., Hilário, M.R., 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. *Nature neuroscience* *12*, 333.
20. Durieux, P.F., Schiffmann, S.N., and de Kerchove d’Exaerde, A. (2012). Differential regulation of motor control and response to dopaminergic drugs by D1R and D2R neurons in distinct dorsal striatum subregions. *The EMBO journal* *31*, 640–653.
21. Alexander, G.E. and Crutcher, M.D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in neurosciences* *13*, 266–271.
22. Chevalier, G. and Deniau, J. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in neurosciences* *13*, 277–280.
23. Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological reviews* *80*, 953–978.
24. Kravitz, A.V., Freeze, B.S., Parker, P.R., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* *466*, 622.
25. Turner, R.S. and Desmurget, M. (2010). Basal ganglia contributions to motor control: a vigorous tutor. *Current opinion in neurobiology* *20*, 704–716.
26. Dudman, J.T. and Krakauer, J.W. (2016). The basal ganglia: from motor commands to the control of vigor. *Current opinion in neurobiology* *37*, 158–166.

27. Robbe, D. and Dudman, J.T. (2020). The Basal Ganglia invigorate actions and decisions. In *The Cognitive Neurosciences* (MIT Press), pp. 527–540.
28. Kim, N., Barter, J.W., Sukharnikova, T., and Yin, H.H. (2014). Striatal firing rate reflects head movement velocity. *European Journal of Neuroscience* *40*, 3481–3490.
29. Rueda-Orozco, P.E. and Robbe, D. (2015). The striatum multiplexes contextual and kinematic information to constrain motor habits execution. *Nature neuroscience* *18*, 453.
30. Yttri, E.A. and Dudman, J.T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* *533*, 402.
31. Sales-Carbonell, C., Taouali, W., Khalki, L., Pasquet, M.O., Petit, L.F., Moreau, T., Rueda-Orozco, P.E., and Robbe, D. (2018). No discrete start/stop signals in the dorsal striatum of mice performing a learned action. *Current Biology* *28*, 3044–3055.
32. Fobbs, W.C., Bariselli, S., Licholai, J.A., Miyazaki, N.L., Matikainen-Ankney, B.A., Creed, M.C., and Kravitz, A.V. (2020). Continuous representations of speed by striatal medium spiny neurons. *Journal of Neuroscience* *40*, 1679–1688.
33. Shmuelof, L. and Krakauer, J.W. (2011). Are we ready for a natural history of motor learning? *Neuron* *72*, 469–476.
34. Safaie, M., Jurado-Parras, M.T., Sarno, S., Louis, J., Karoutchi, C., Petit, L.F., Pasquet, M.O., Eloy, C., and Robbe, D. (2020). Turning the body into a clock: Accurate timing is facilitated by simple stereotyped interactions with the environment. *Proceedings of the National Academy of Sciences* *117*, 13084–13093.
35. Shadmehr, R., De Xivry, J.J.O., Xu-Wilson, M., and Shih, T.Y. (2010). Temporal discounting of reward and the cost of time in motor control. *Journal of Neuroscience* *30*, 10507–10516.

36. Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement vigor as a reflection of subjective economic utility. *Trends in neurosciences* 42, 323–336.
37. Mazzoni, P., Hristova, A., and Krakauer, J.W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *Journal of neuroscience* 27, 7105–7116.
38. Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience* 7, 907.
39. Baraduc, P., Thobois, S., Gan, J., Broussolle, E., and Desmurget, M. (2013). A common optimization principle for motor execution in healthy subjects and parkinsonian patients. *Journal of Neuroscience* 33, 665–677.
40. Galter, D., Pernold, K., Yoshitake, T., Lindqvist, E., Hoffer, B., Kehr, J., Larsson, N.G., and Olson, L. (2010). MitoPark mice mirror the slow progression of key symptoms and L-DOPA response in Parkinson's disease. *Genes, Brain and Behavior* 9, 173–181.
41. Panigrahi, B., Martin, K.A., Li, Y., Graves, A.R., Vollmer, A., Olson, L., Mensh, B.D., Karpova, A.Y., and Dudman, J.T. (2015). Dopamine is required for the neural representation and control of movement vigor. *Cell* 162, 1418–1430.
42. Mourra, D., Gnazzo, F., Cobos, S., and Beeler, J.A. (2020). Striatal Dopamine D2 Receptors Regulate Cost Sensitivity and Behavioral Thrift. *Neuroscience* 425, 134–145.
43. Salamone, J.D. and Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. *Neuron* 76, 470–485.
44. Shadmehr, R. and Krakauer, J.W. (2008). A computational neuroanatomy for motor control. *Experimental brain research* 185, 359–381.

45. Berke, J.D. (2018). What does dopamine mean? *Nature neuroscience* *21*, 787–793.
46. Yin, H.H. and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience* *7*, 464–476.
47. Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M.C., Lehericy, S., Bergman, H., Agid, Y., DeLong, M.R., and Obeso, J.A. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson’s disease. *Nature Reviews Neuroscience* *11*, 760–772.
48. Moussa, R., Poucet, B., Amalric, M., and Sargolini, F. (2011). Contributions of dorsal striatal subregions to spatial alternation behavior. *Learning & memory* *18*, 444–451.
49. Sippy, T., Lapray, D., Crochet, S., and Petersen, C.C. (2015). Cell-type-specific sensorimotor processing in striatal projection neurons during goal-directed behavior. *Neuron* *88*, 298–305.
50. 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*, 476–489.
51. 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*.
52. Horak, F.B. and Anderson, M.E. (1984). Influence of globus pallidus on arm movements in monkeys. I. Effects of kainic acid-induced lesions. *Journal of Neurophysiology* *52*, 290–304.
53. Desmurget, M. and Turner, R.S. (2010). Motor sequences and the basal ganglia: kinematics, not habits. *Journal of Neuroscience* *30*, 7685–7690.

54. Hwang, E.J. (2013). The basal ganglia, the ideal machinery for the cost-benefit analysis of action plans. *Frontiers in neural circuits* 7, 121.
55. Ferguson, S.M., Eskenazi, D., Ishikawa, M., Wanat, M.J., Phillips, P.E., Dong, Y., Roth, B.L., and Neumaier, J.F. (2011). Transient neuronal inhibition reveals opposing roles of indirect and direct pathways in sensitization. *Nature neuroscience* 14, 22–24.
56. Kravitz, A.V., Tye, L.D., and Kreitzer, A.C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature neuroscience* 15, 816.
57. Tai, L.H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nature neuroscience* 15, 1281.
58. Shadmehr, R., Huang, H.J., and Ahmed, A.A. (2016). A representation of effort in decision-making and motor control. *Current biology* 26, 1929–1934.
59. Carland, M.A., Thura, D., and Cisek, P. (2019). The Urge to Decide and Act: Implications for Brain Function and Dysfunction. *The Neuroscientist* 25, 491–511.
60. Thura, D. and Cisek, P. (2017). The basal ganglia do not select reach targets but control the urgency of commitment. *Neuron* 95, 1160–1170.
61. Keller, J.B. (1974). Optimal velocity in a race. *The American Mathematical Monthly* 81, 474–480.
62. Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., and Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11, 1–36.
63. Berret, B. and Jean, F. (2016). Why don't we move slower? the value of time in the neural control of action. *Journal of neuroscience* 36, 1056–1070.

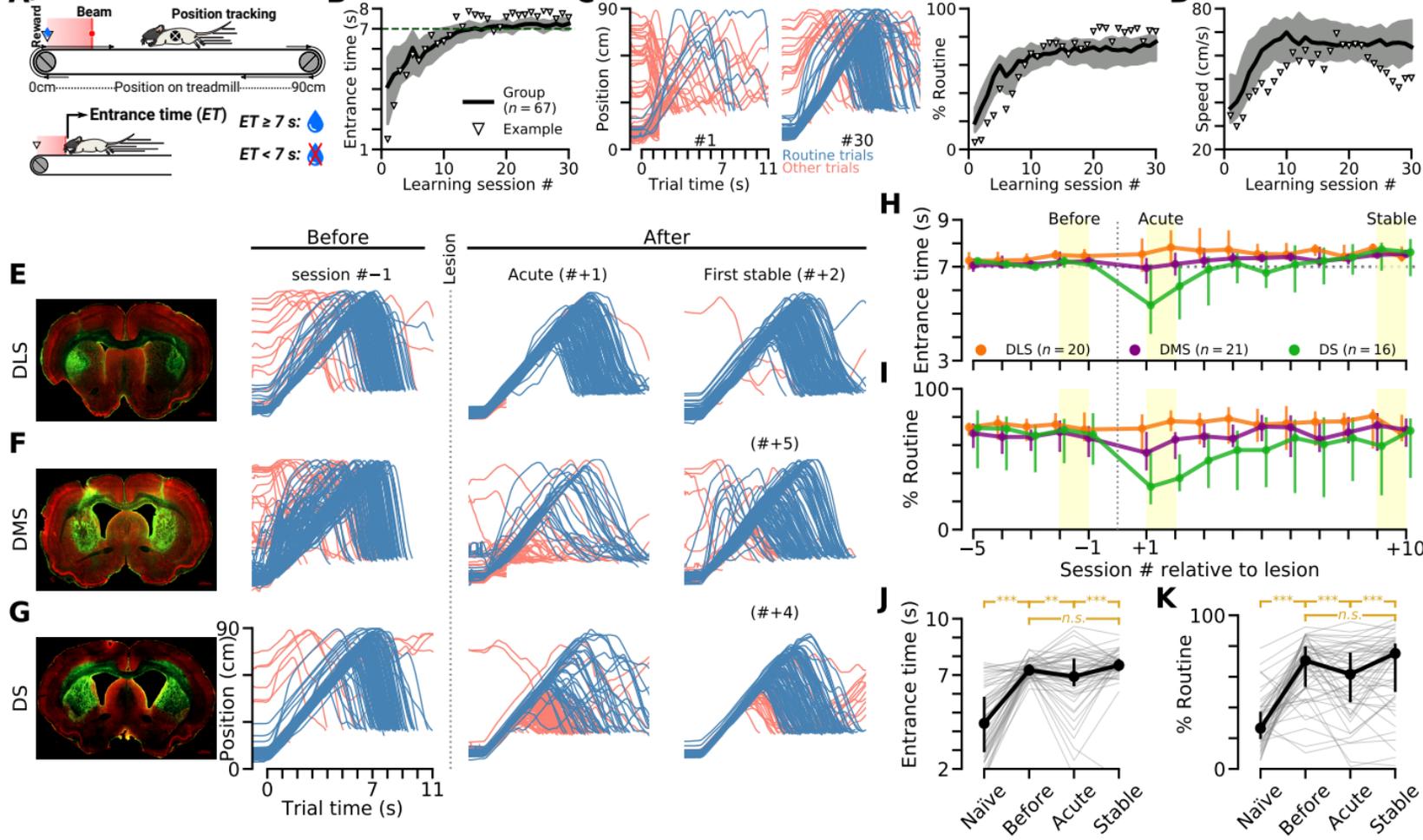
Figure 1

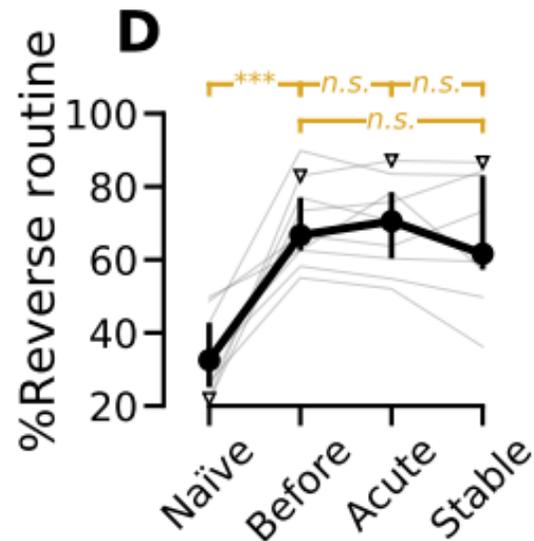
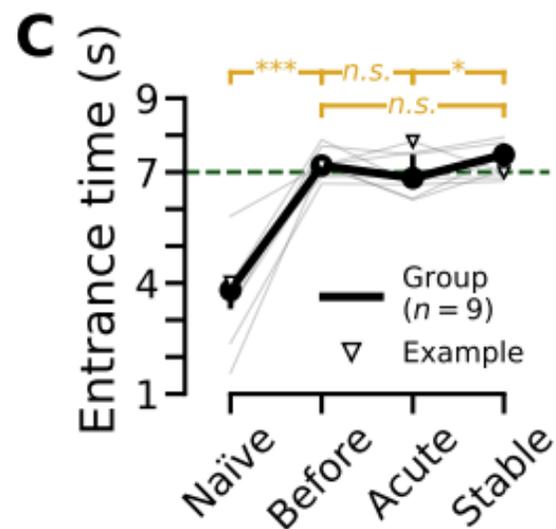
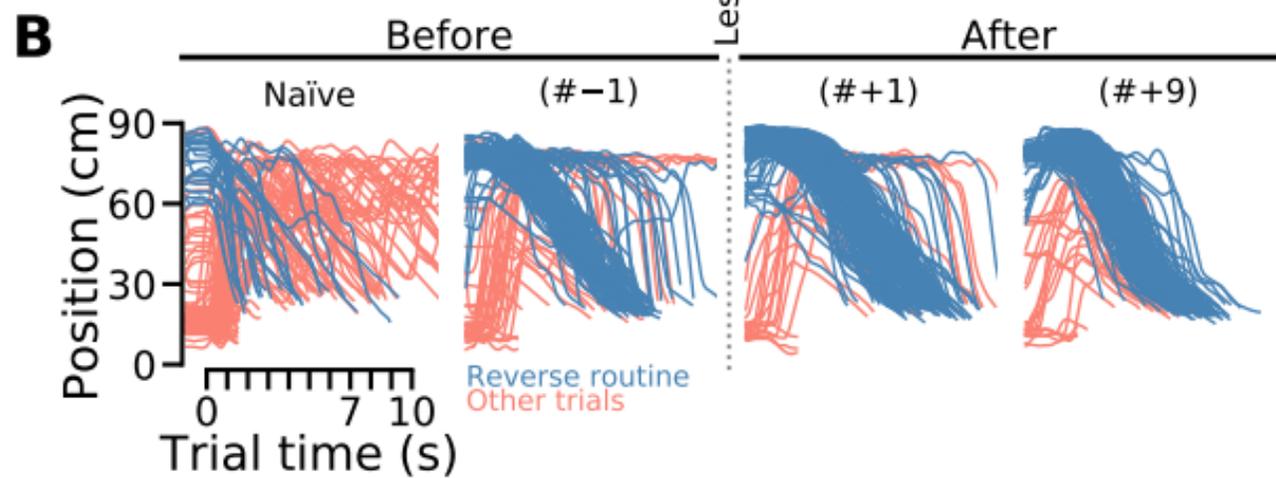
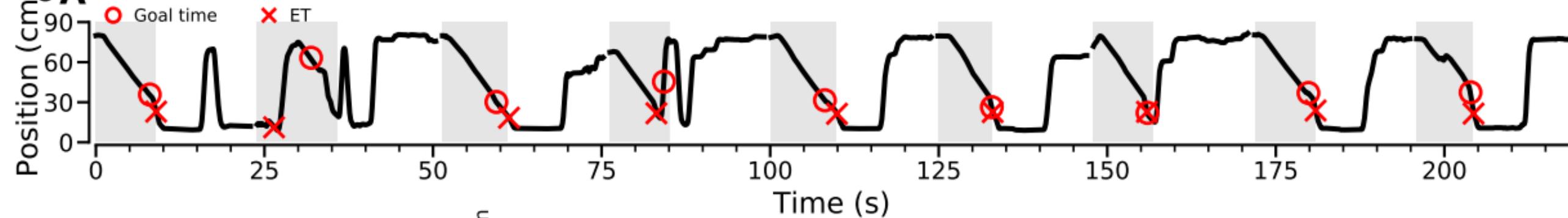
Figure 2

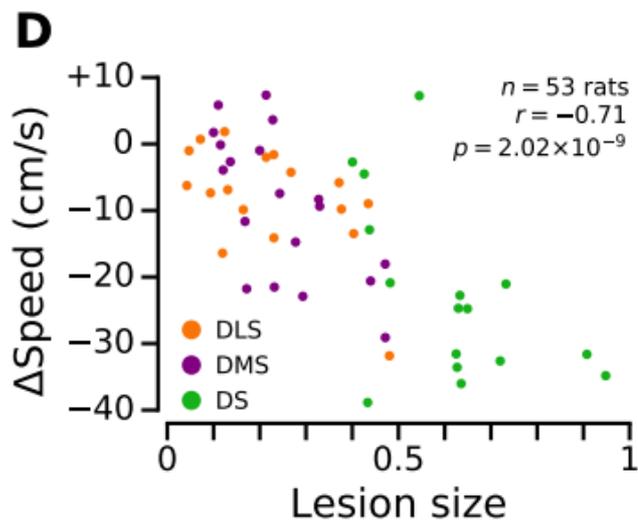
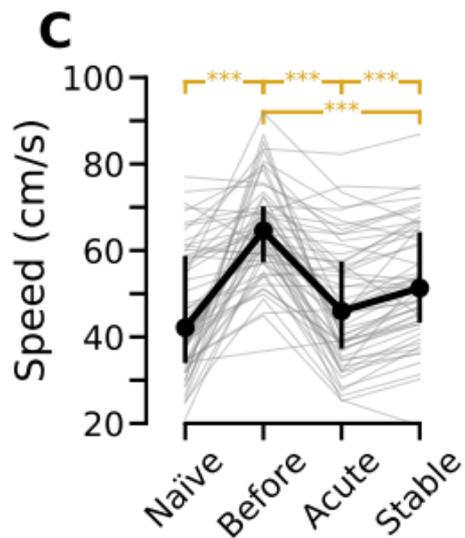
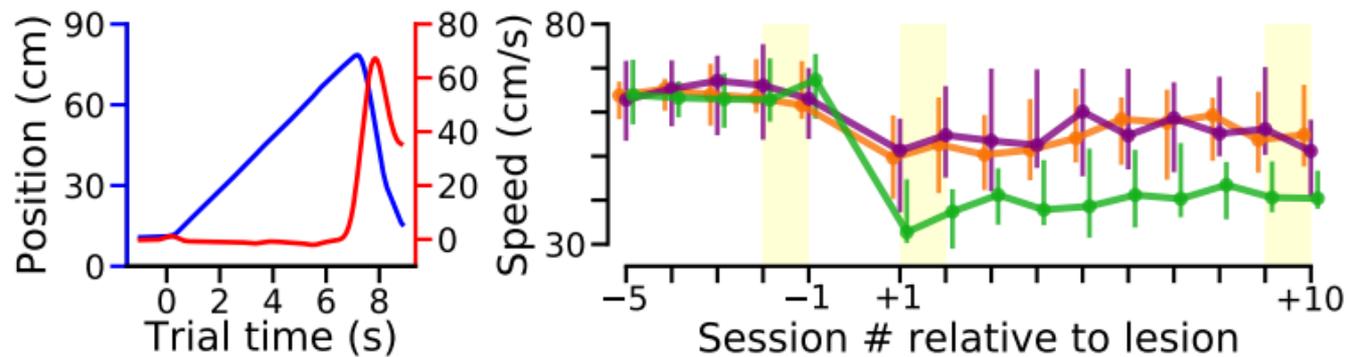
Figure 3

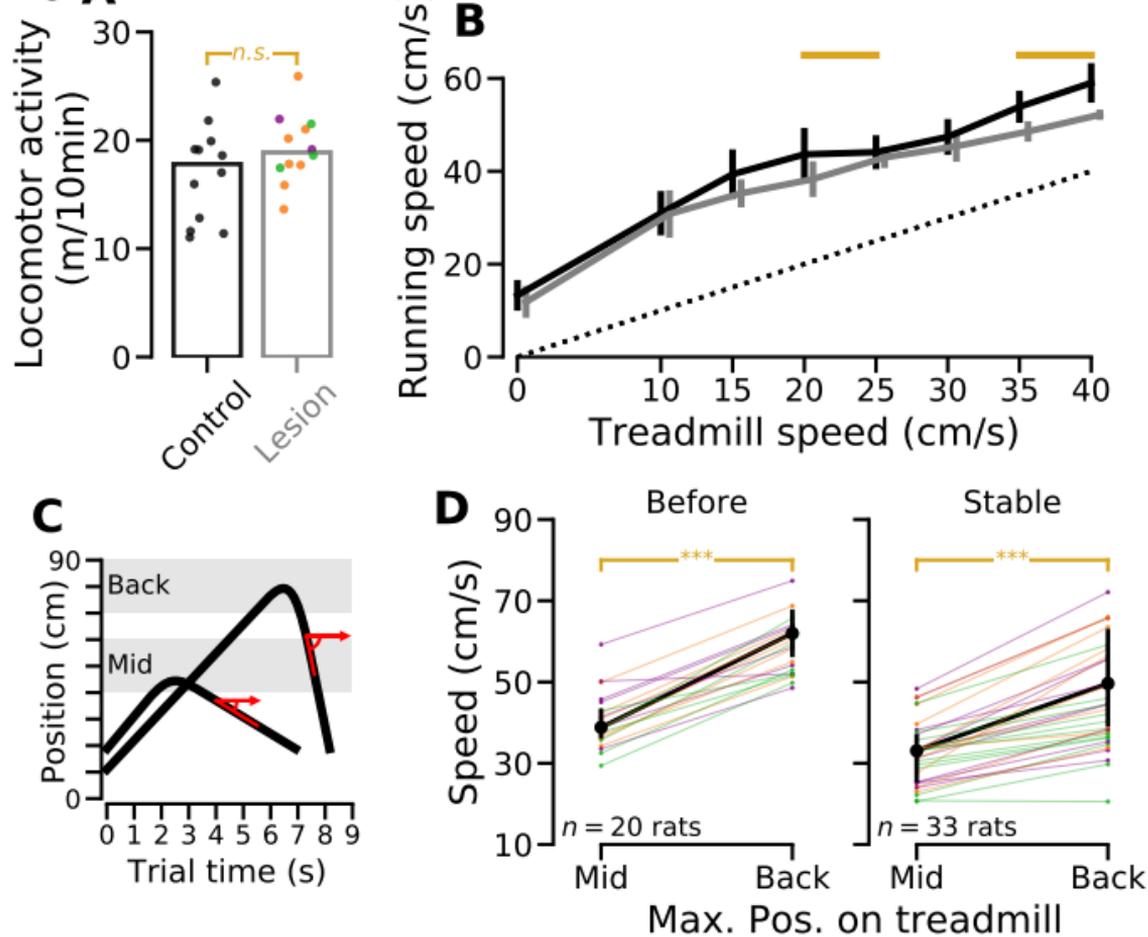
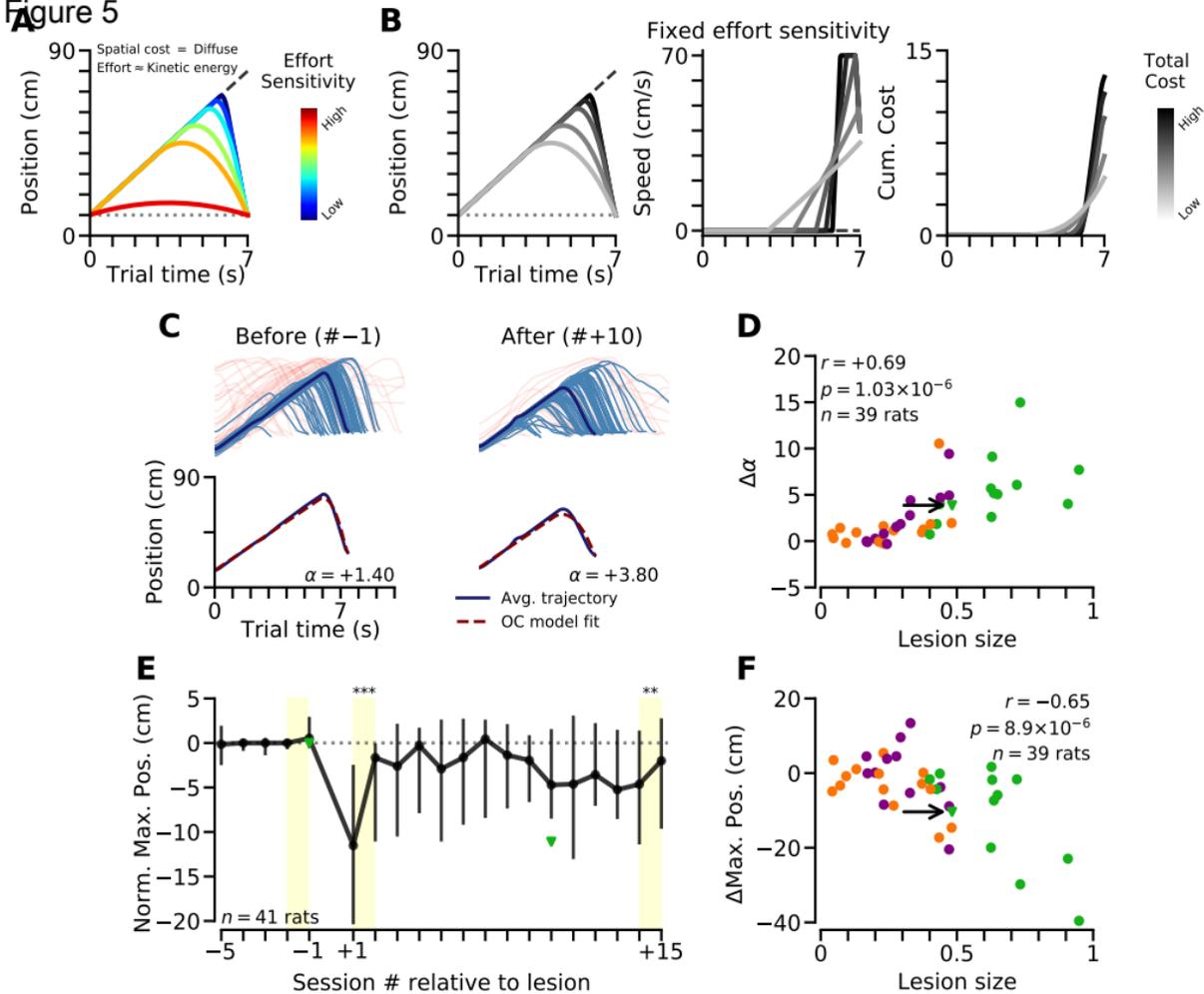
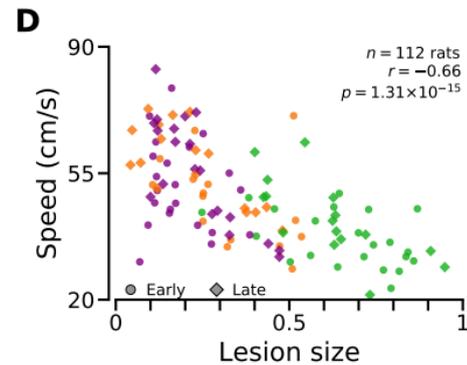
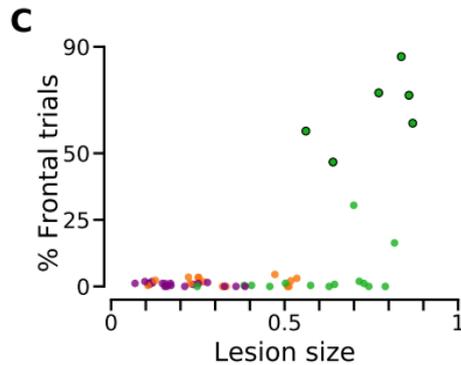
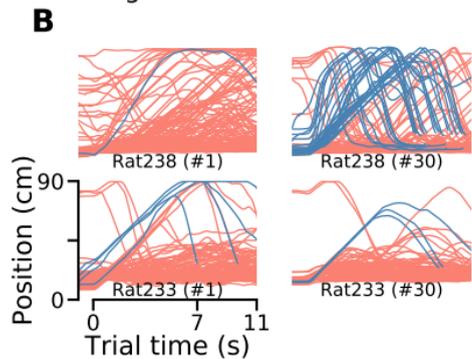
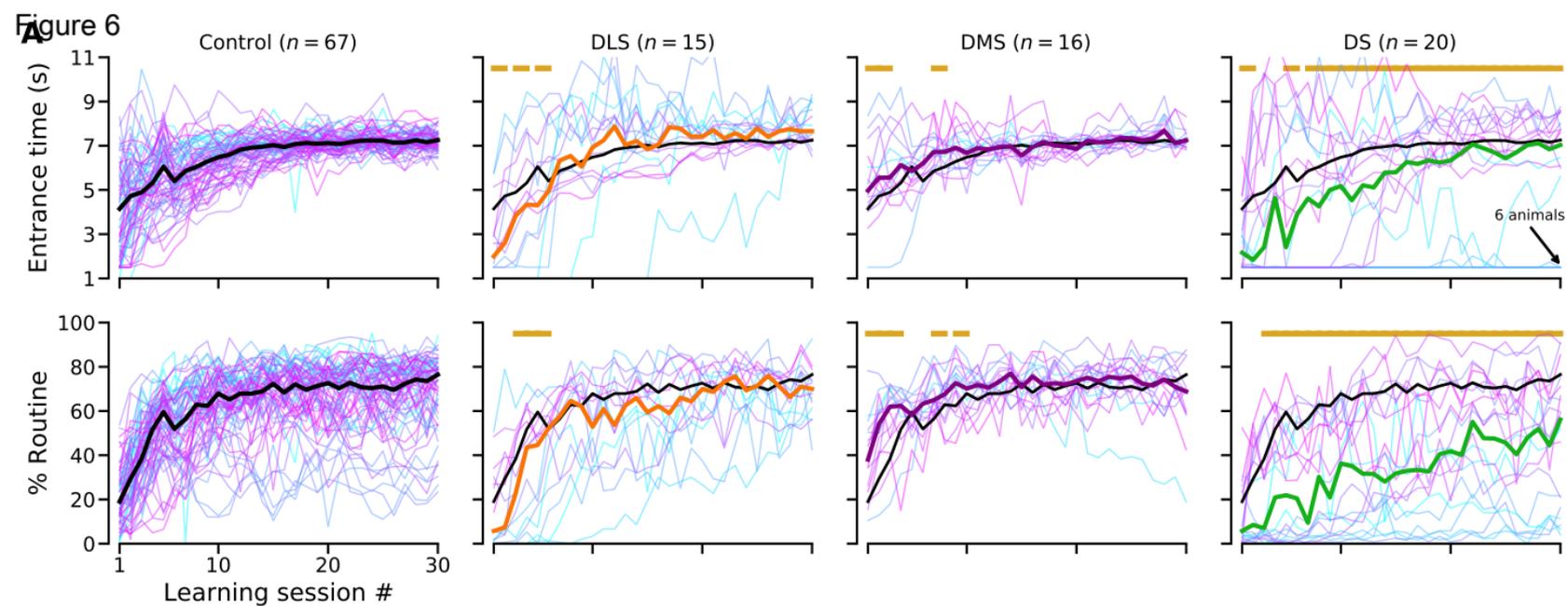
Figure 4

Figure 5



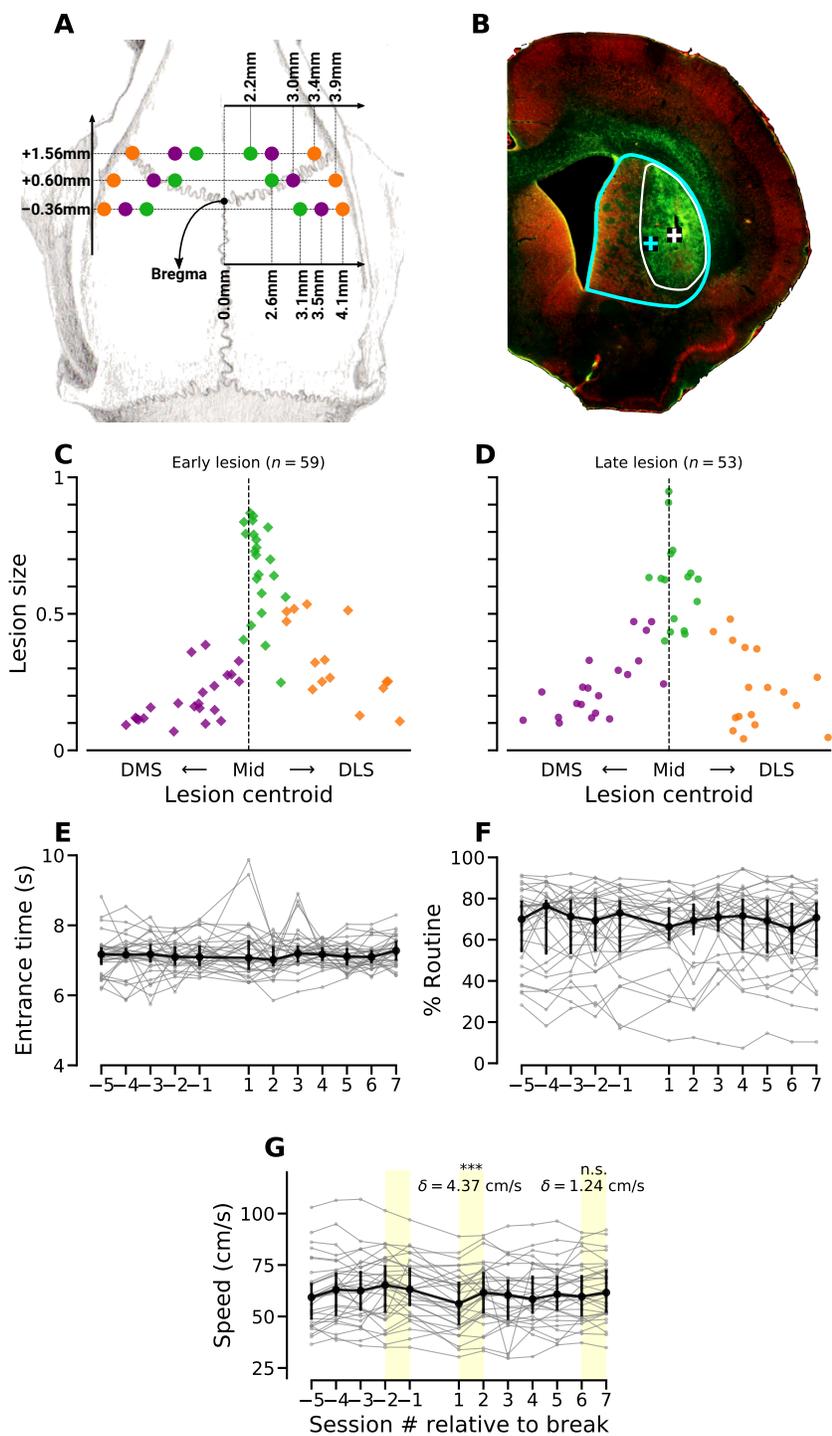


Figure S1: Dorsal striatum lesion quantification and effect of the two-weeks long break on task performance. Related to Figure 1.

(A) Schematic of the injection sites for the lesion. **(B)** Illustration of the quantification of the lesion size. For each coronal slide and hemi-striatum, the contour of the lesion was manually outlined using GFAP staining. The relative size of the lesion (compared to the full dS, manually outlined on the NeuN staining) and the coordinates of the lesion/striatum centroid was calculated. For each animal, the size and laterality were obtained by averaging data along the anteroposterior axis, for both left and right hemispheres. **(C, D)** Lesion size versus laterality for animals that underwent lesion before (Early, C) and after (Late, D) training. Lesion quantification was performed blindly relative to the behavioral analysis. In four animals with a dS lesion performed after learning the task (late lesion), the lesion size quantification could not be properly performed. These animals were classified according to their injection coordinates in the surgery (3 DLS and 1 DMS), however they were excluded from any analysis that required the lesion size (hence the difference between the number of “late lesion” animals in this figure, $n = 53$, and the total number of animals in Figure 1, $n = 57$). **(E-G)** Task performance before and after a two weeks long break in practice. Non-lesioned animals had stable performance before the two weeks-long break (same duration as lesion recovery period) in practice. (E) Entrance time. (F) Percentage of trials during which animals used the wait-and-run routine. (G) Speed of the animals when they ran toward the reward area. A small but significant reduction in running speed was observed just after the break (δ denotes the effect size).

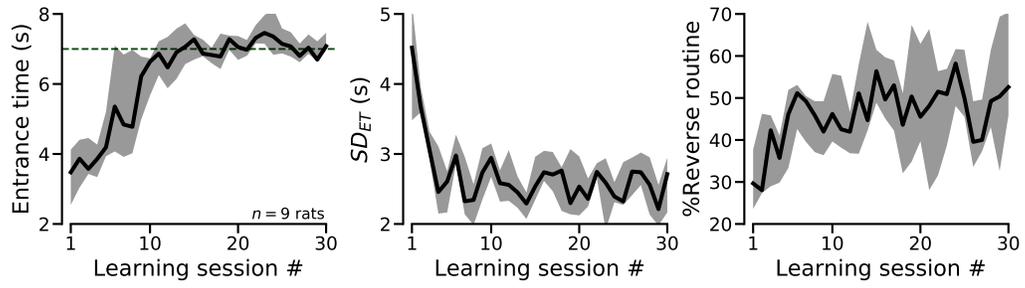


Figure S2: Performance improvement in the reverse treadmill task. Related to Figure 2. Left: Entrance time across learning sessions. Middle: Session-by-session standard deviation of ET . Right: Percentage of trials during which animals used the run-and-wait (reverse) routine.

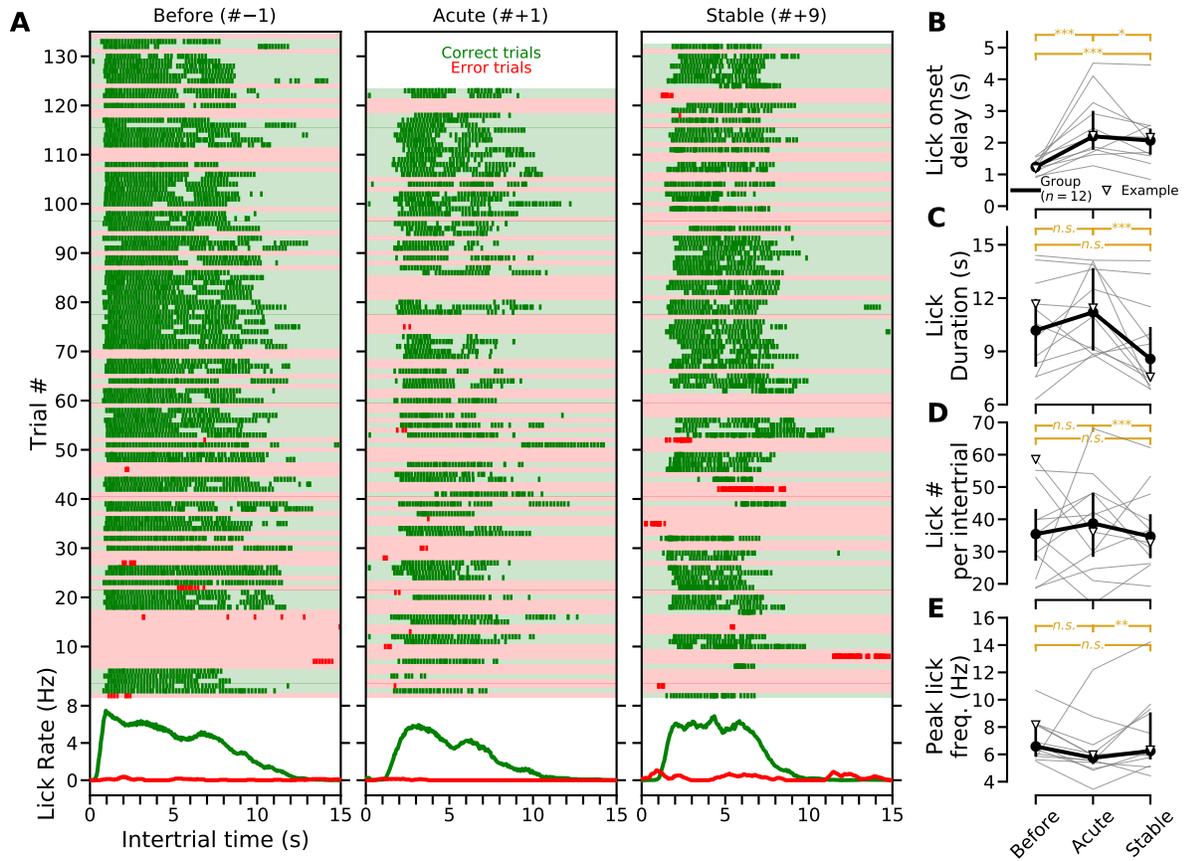


Figure S3: Licking behavior after dS lesion. Related to Figure 3. (A) Trial-by-trial licking pattern (top) and average lick rate aligned to intertrial onset for a single animal in 3 sessions (1 session just before and 2 after lesion). (B-E) Effect of dS lesion on lick onset delay (B), lick duration (C), number of licks per intertrial (D) and peak lick frequency (E). Same color code for dS lesion types as in Figure 1.

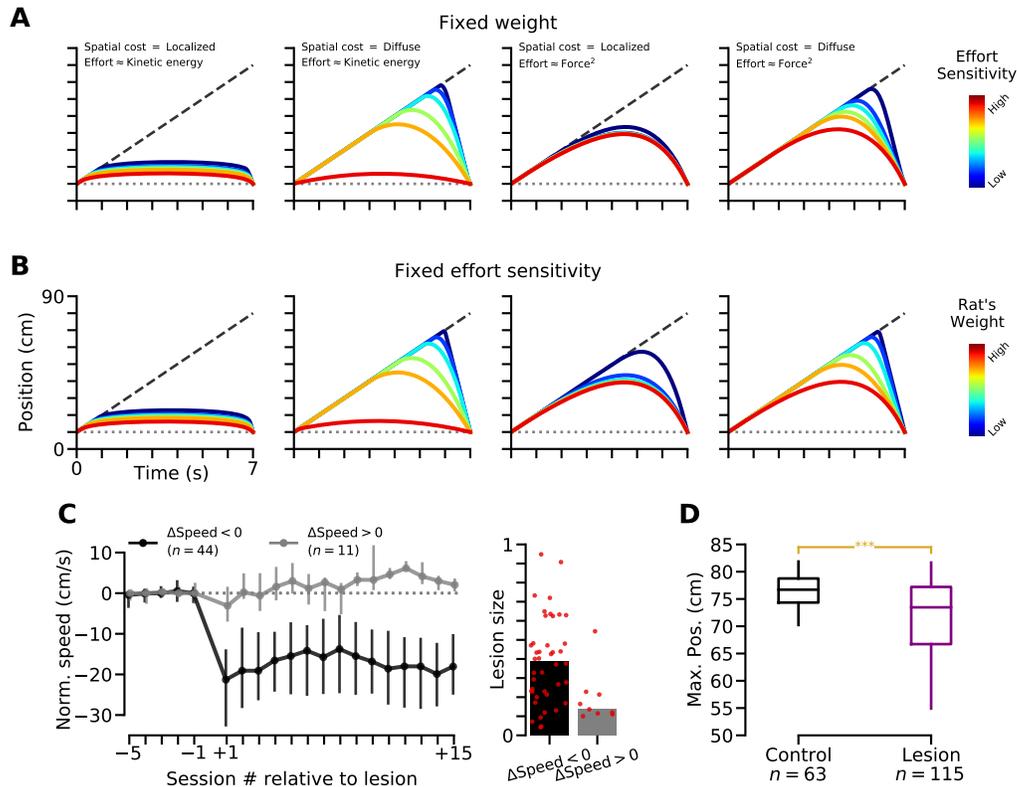


Figure S4: Optimal control models and effect of dS lesion on maximal position. Related to Figure 5. (A) Effect of effort sensitivity on optimal trajectories. Trajectories were obtained using 4 models that combine two approximation of effort and spatial costs. The cost of premature entrance in the reward area (spatial cost) was simulated using a Heaviside function that was either localized (\sim step function with non-zero value within the reward area) or diffuse (\sim a sigmoid function whose value gradually decreases toward zero away from the reward area). Effort was approximated either as the kinetic energy or as the square value of the muscular force produced by the animals. The weight of the animal was fixed to the value $m = 1$. (B) Similar to A, but for fixed value of effort sensitivity ($\alpha = 0.1$) and variable rat's weight. (C) Left, animals were divided into two groups based on the dS lesion effect on running speed. Right, lesion size for animals in those two groups. 3 out of the 44 animals with decreased running speed after lesion could not be included in the Max. Pos. analysis in Figure 5E (not enough routine trials in a few sessions analyzed, see STAR Methods). 2 out of the 41 remaining animals could not be included in the correlation plots in Figures 5D and 5F due to incomplete histological quantification. (D) Maximum position of control and lesioned rats. Each boxplot represents the range of the Max. Pos. (center line, median; box, 25th and 75th percentiles; whiskers, 5th and 95th percentiles) for control and lesioned (both before and after training) animals. For each animal, the median value of their average trajectory was computed over the last 5 sessions performed without lesion (Control group, at least 30 training sessions), and/or the last 5 sessions performed after dS lesion (Lesion group). The difference in number of control animals between this ($n = 63$) and Figure 6 ($n = 67$) is due to 4 animals that did not perform the wait-and-run routine reliably, precluding the Max. Pos. computation. Statistical comparison using permutation test (10000 permutations). Same significance symbols as in Figure 1.