

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 Louis[1], Corane Karoutchi[1], Bastien Berret[3,4,5], David Robbe[1†]

[1] Institut de Neurobiologie de la Mediterranée (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 Phenomènes 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*

Introduction

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

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

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

Results

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

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

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

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

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

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

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

Lesions of the dorsal striatum decreased the speed of reward-oriented locomotion while sparing basic motor control and reward consumption

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

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

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

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

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

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

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

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

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

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

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

Discussion

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

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

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

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

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

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

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

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

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

Acknowledgements

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

Author contributions

Maria-Teresa Jurado-Parras: Conceptualization, Investigation, Writing–review & editing; Mostafa Safaie: Conceptualization, Software, Visualization, Investigation, Formal analysis, Writing–original draft; Stefania Sarno: Software, Methodology, Formal analysis, Conceptualization, Visualization, Writing–review & editing; Jordane Louis: Investigation; Corane Karoutchi: Investigation; Bastien Berret: Conceptualization, Methodology, Formal analysis, Writing–review & editing; David Robbe: Funding acquisition, Project administration, Supervision, Visualization, Writing–original draft, Conceptualization, Software.

Declaration of Interests

The authors declare no competing interests.

Figure Legends

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

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

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

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

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

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

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

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

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact

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

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The Jupyter Notebooks and the raw data necessary for the full replication of the figures is publicly available via the Open Science Foundation in the following repository: https://osf.io/2vmus/

EXPERIMENTAL MODEL AND SUBJECT DETAILS

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

METHODS DETAILS

Apparatus

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

Habituation

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

Main Behavioral Task

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

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

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

Reverse Treadmill Task

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

Locomotor Activity Test

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

Lesion Surgery

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

Histology

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

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

Optimal Control Model Description

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

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

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

The term v/τ in the above equation is a resistive force per unit mass . The term τ is a friction coefficient (when τ increases resistance decreases) and $u(t)$ is the propulsive force per unit mass. In all the simulation we set $\tau = 1.8$ s. Qualitatively, the results we obtained are independent of τ. A key component of the behavioral task is that, to obtain a reward, animals must enter the reward area after GT. This reward area is delimited by the front wall of the treadmill equipped with a reward port and an infrared-beam located at 10 cm of the front wall. In agreement with the behavioral data, the initial and final positions of the animal in the model were both equal to the beam position $x_b = 10$ cm. The position of the rat is constrained by the treadmill length, which is $L_T = 90$ cm. The treadmill pushes the animal toward the rear wall of the treadmill (located at position 90 cm) with a constant positive velocity of $v_T = 10$ cm/s. In the coordinate adopted in the model, running toward the reward area is, therefore, occurring with negative velocity. The velocity of the rat was also constrained to be negative, i.e. $v_{max} = 0$ cm/s (the rat cannot run toward the rear wall of the) and its module could not be bigger than 70 cm/s, i.e. $v_{min} = -70$ cm/s (the rat cannot run toward the reward area at a speed faster than 70 cm/s). We also constrained the maximum speed of the rat at the moment of beam crossing to be less than 40 cm/s, i.e. v_C^{min} = - 40. Finally, the module of the acceleration was forced to be less than 180 cm/s^2. We relied on experimental data to constrain the dynamical variables (i.e. speed and acceleration) used in the model.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

QUANTIFICATION AND STATISTICAL ANALYSIS

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

Motor Routine Definition

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

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

Speed Calculation

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

Reverse Routine Definition

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

Definition of Frontal Trials

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

Speed Modulation Analysis

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

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

Max. Pos. Quantification

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

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

Normalizing Speed and Max. Pos.

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

Licking Analysis

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

Quantification of Lesion Size

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

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

Statistics

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

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

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

Finally, the relation between lesion size and magnitude of behavioral impairment was quantified 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., Pélégrini-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

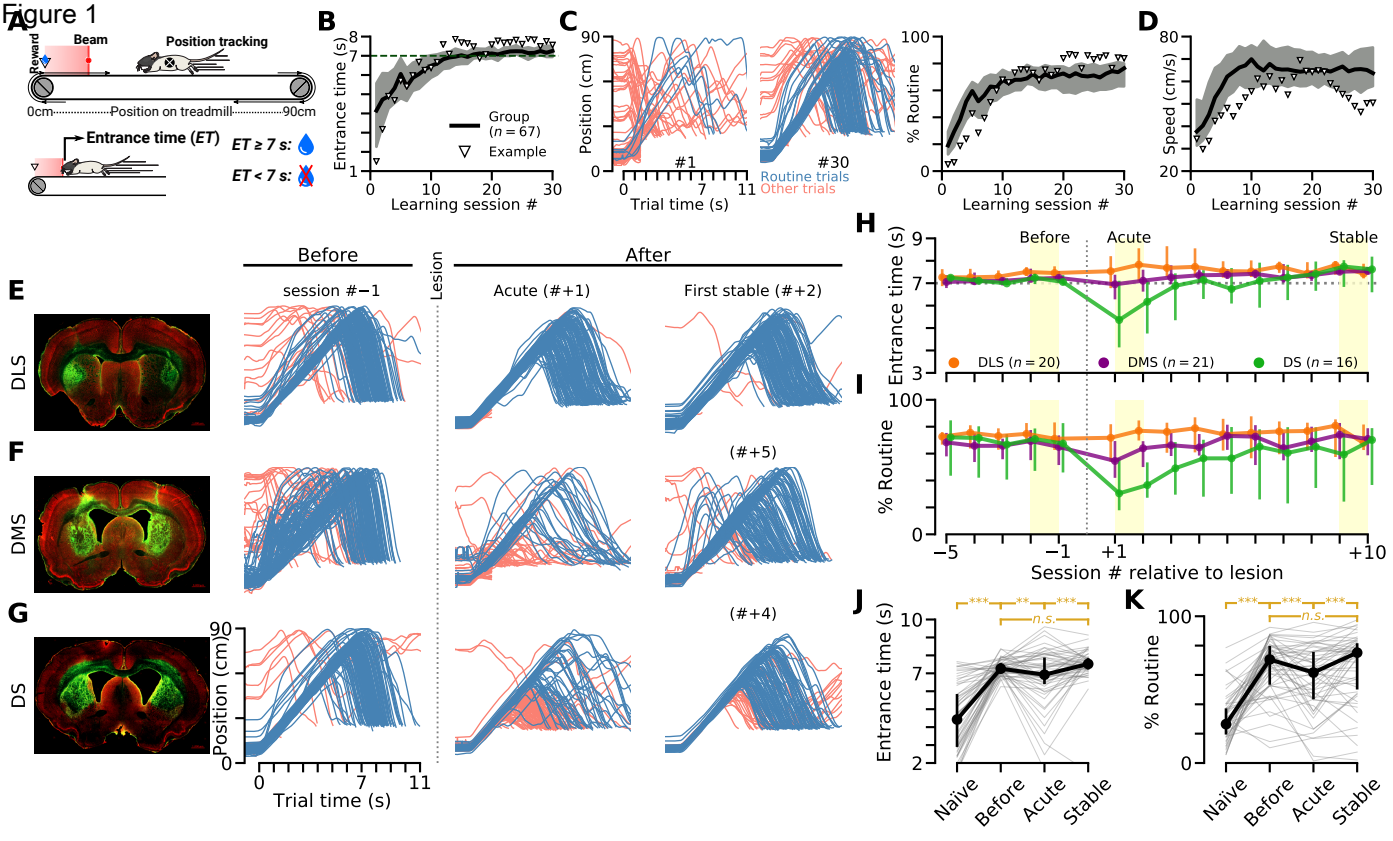

A
Reward
Beam
Position tracking
0cm
Position on treadmill
90cm
Entrance time (ET)
ET ≥ 7 s:
ET < 7 s:
B
Entrance time (s)
Learning session #
Group (n = 67)
Example
C
Position (cm)
#1
#30
Trial time (s)
Routine trials
Other trials
% Routine
Learning session #
D
Speed (cm/s)
Learning session #
E
DLS
F
DMS
G
DS
Before
After
Lesion
session #−1
Acute (#+1)
First stable (#+2)
(#+5)
(#+4)
Position (cm)
Trial time (s)
H
Before
Acute
Stable
Entrance time (s)
DLS (n = 20)
DMS (n = 21)
DS (n = 16)
I
% Routine
Session # relative to lesion
J
Entrance time (s)
n.s.
Naive
Before
Acute
Stable
K
% Routine
n.s.
Naive
Before
Acute
Stable

Figure 2

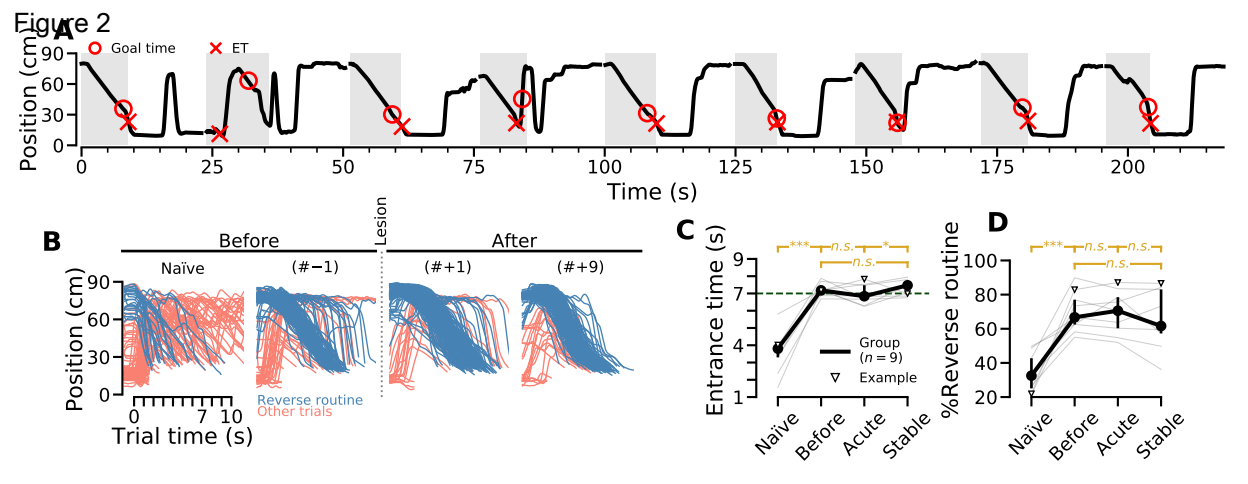

A
Goal time
ET
Position (cm)
90
60
30
0
0
25
50
75
100
125
150
175
200
Time (s)
B
Before
Lesion
After
Naïve
(#−1)
(#+1)
(#+9)
Position (cm)
90
60
30
0
0
7
10
Trial time (s)
Reverse routine
Other trials
C
Entrance time (s)
9
7
4
1
***
n.s.
*
n.s.
Group
(n = 9)
Example
Naïve
Before
Acute
Stable
D
%Reverse routine
100
80
60
40
20
***
n.s.
n.s.
n.s.
Naïve
Before
Acute
Stable

Figure 3

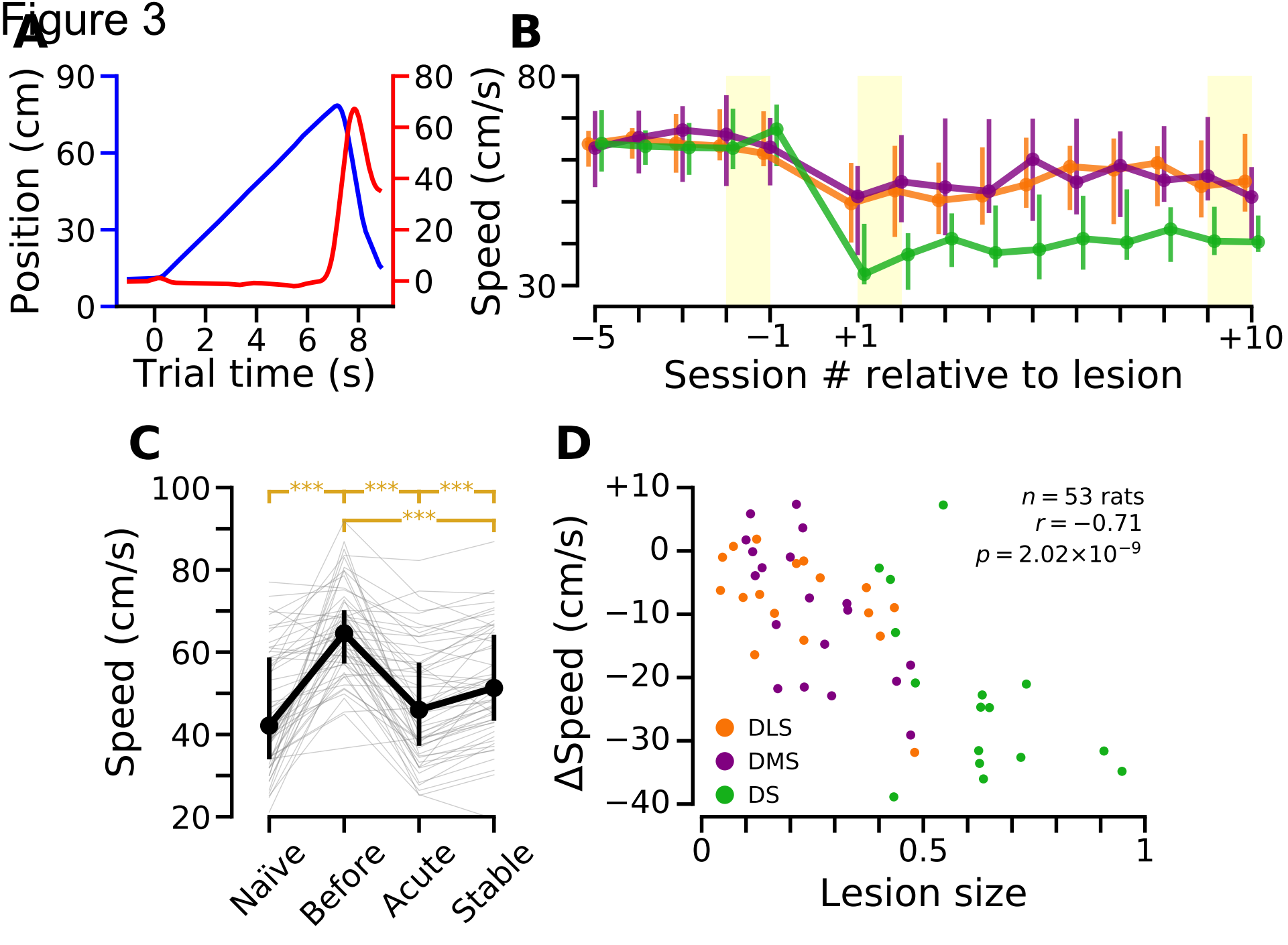

A
Position (cm)
90
60
30
0
80
60
40
20
0
0 2 4 6 8
Trial time (s)
B
Speed (cm/s)
80
30
−5
−1
+1
+10
Session # relative to lesion
C
Speed (cm/s)
100
80
60
40
20
***
***
***
***
Naïve
Before
Acute
Stable
D
ΔSpeed (cm/s)
+10
0
−10
−20
−30
−40
n = 53 rats
r = −0.71
p = 2.02×10−9
DLS
DMS
DS
0
0.5
1
Lesion size

Figure 4

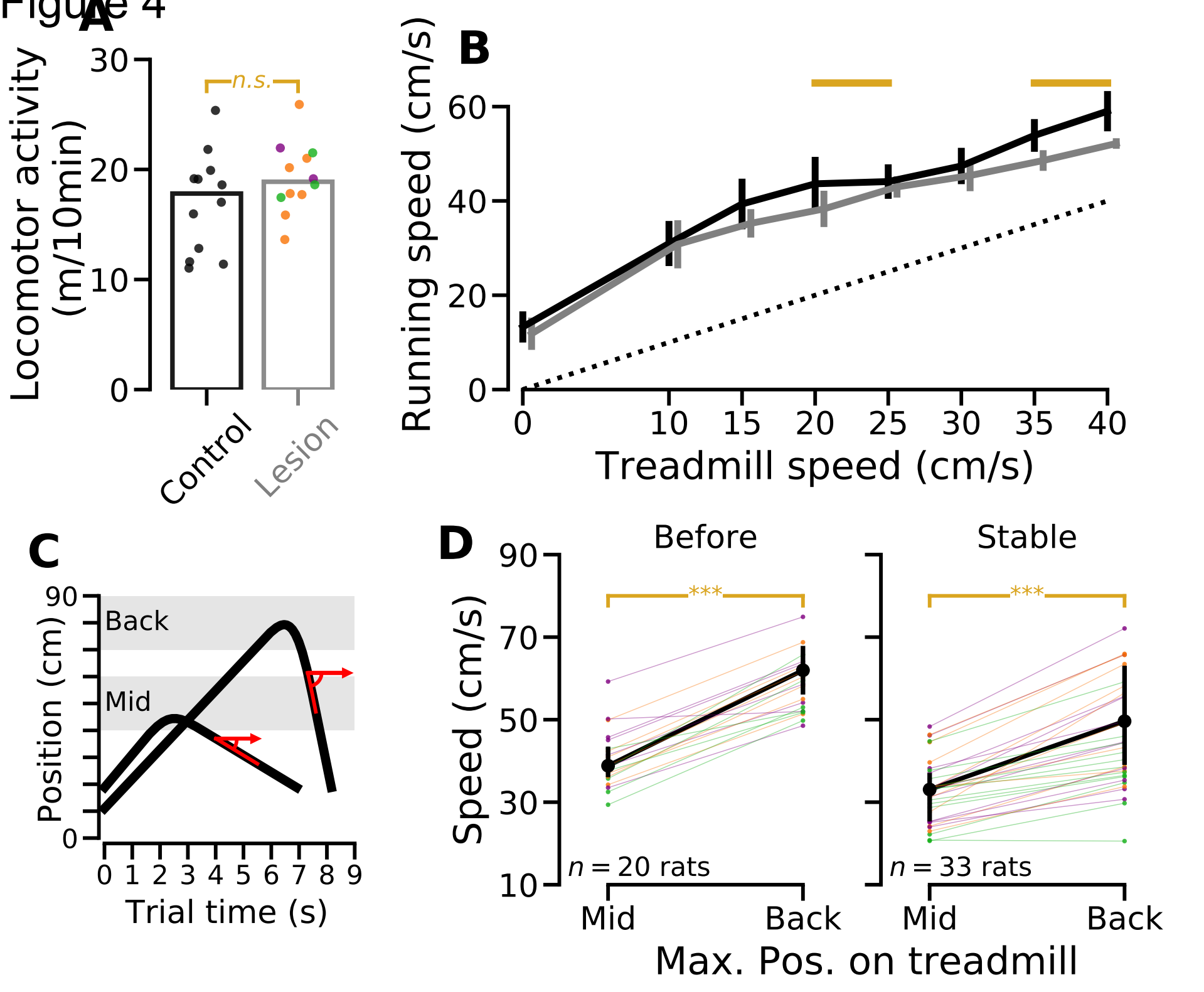

A
Locomotor activity (m/10min)
30
20
10
0
n.s.
Control
Lesion
B
Running speed (cm/s)
60
40
20
0
0
10
15
20
25
30
35
40
Treadmill speed (cm/s)
C
Position (cm)
90
0
Back
Mid
0 1 2 3 4 5 6 7 8 9
Trial time (s)
D
Before
Stable
Speed (cm/s)
90
70
50
30
10
***
***
n = 20 rats
n = 33 rats
Mid
Back
Mid
Back
Max. Pos. on treadmill

Figure 5

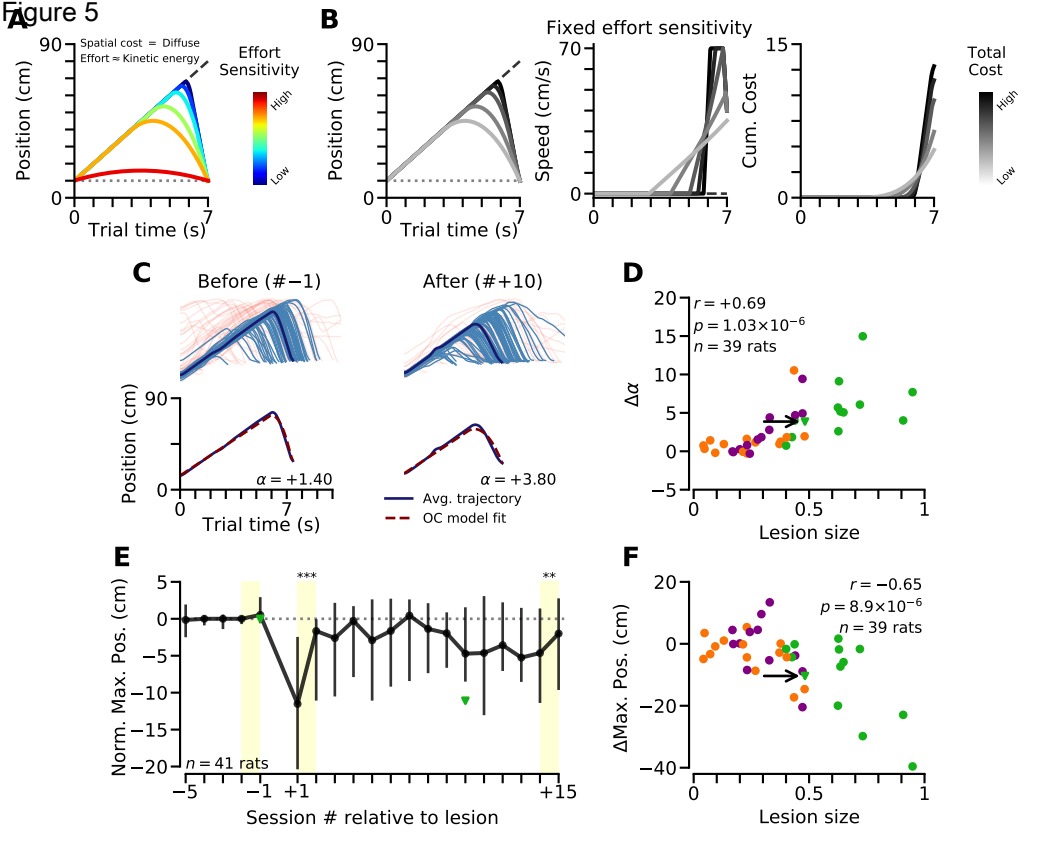
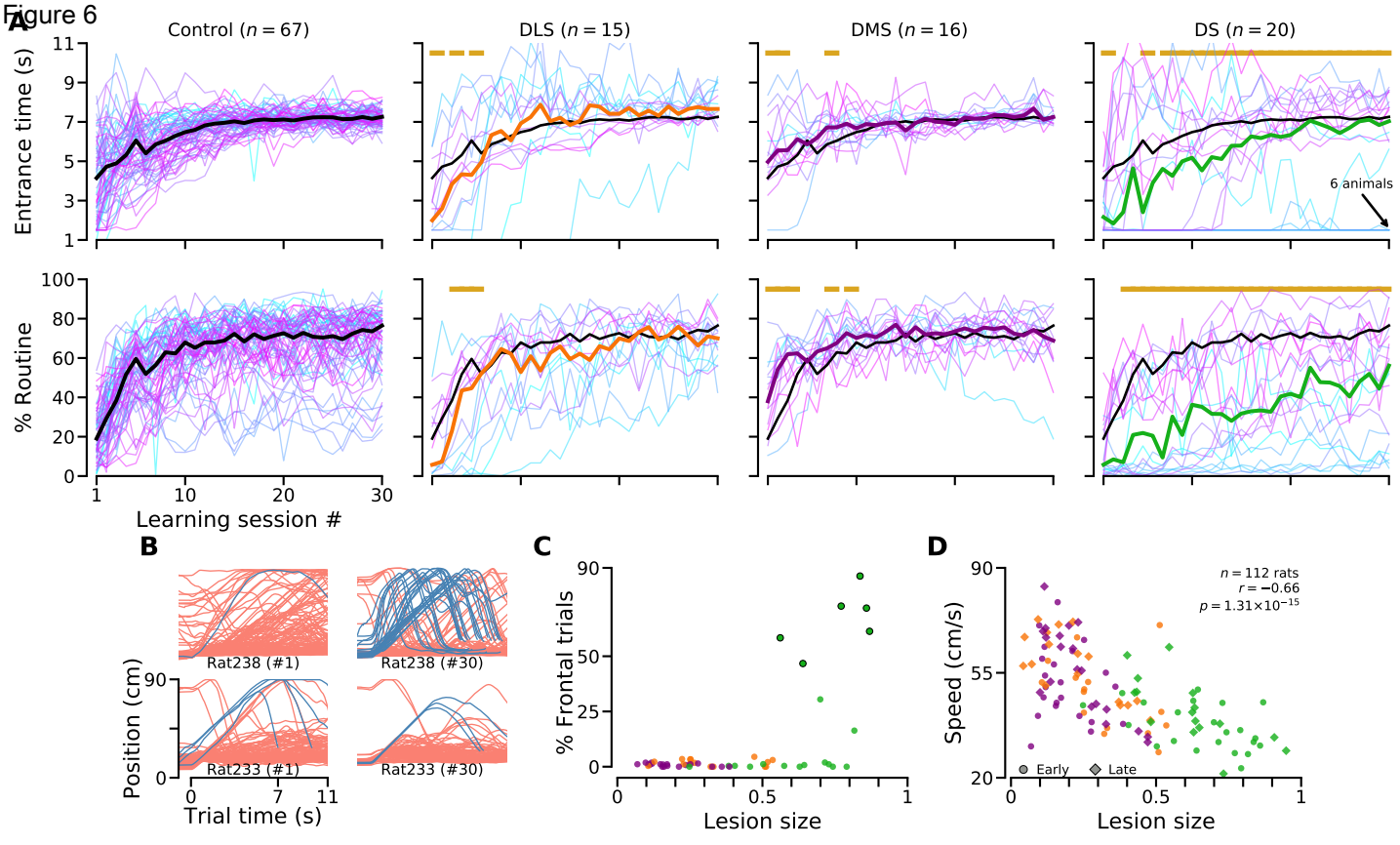

A
Spatial cost = Diffuse
Effort ≈ Kinetic energy
Effort Sensitivity
High
Low
Position (cm)
90
0
0
7
Trial time (s)
B
Fixed effort sensitivity
Position (cm)
Speed (cm/s)
70
0
Cum. Cost
15
0
Total Cost
High
Low
C
Before (#−1)
After (#+10)
Position (cm)
α = +1.40
α = +3.80
Trial time (s)
Avg. trajectory
OC model fit
D
r = +0.69
p = 1.03×10−6
n = 39 rats
Δα
Lesion size
E
Norm. Max. Pos. (cm)
***
**
n = 41 rats
−5
−1
+1
+15
Session # relative to lesion
F
r = −0.65
p = 8.9×10−6
n = 39 rats
ΔMax. Pos. (cm)
Lesion size

Figure 6


A
Control (n = 67)
DLS (n = 15)
DMS (n = 16)
DS (n = 20)
Entrance time (s)
% Routine
Learning session #
6 animals
B
Rat238 (#1)
Rat238 (#30)
Rat233 (#1)
Rat233 (#30)
Position (cm)
Trial time (s)
C
% Frontal trials
Lesion size
D
Speed (cm/s)
Lesion size
n = 112 rats
r = −0.66
p = 1.31×10−15
Early
Late

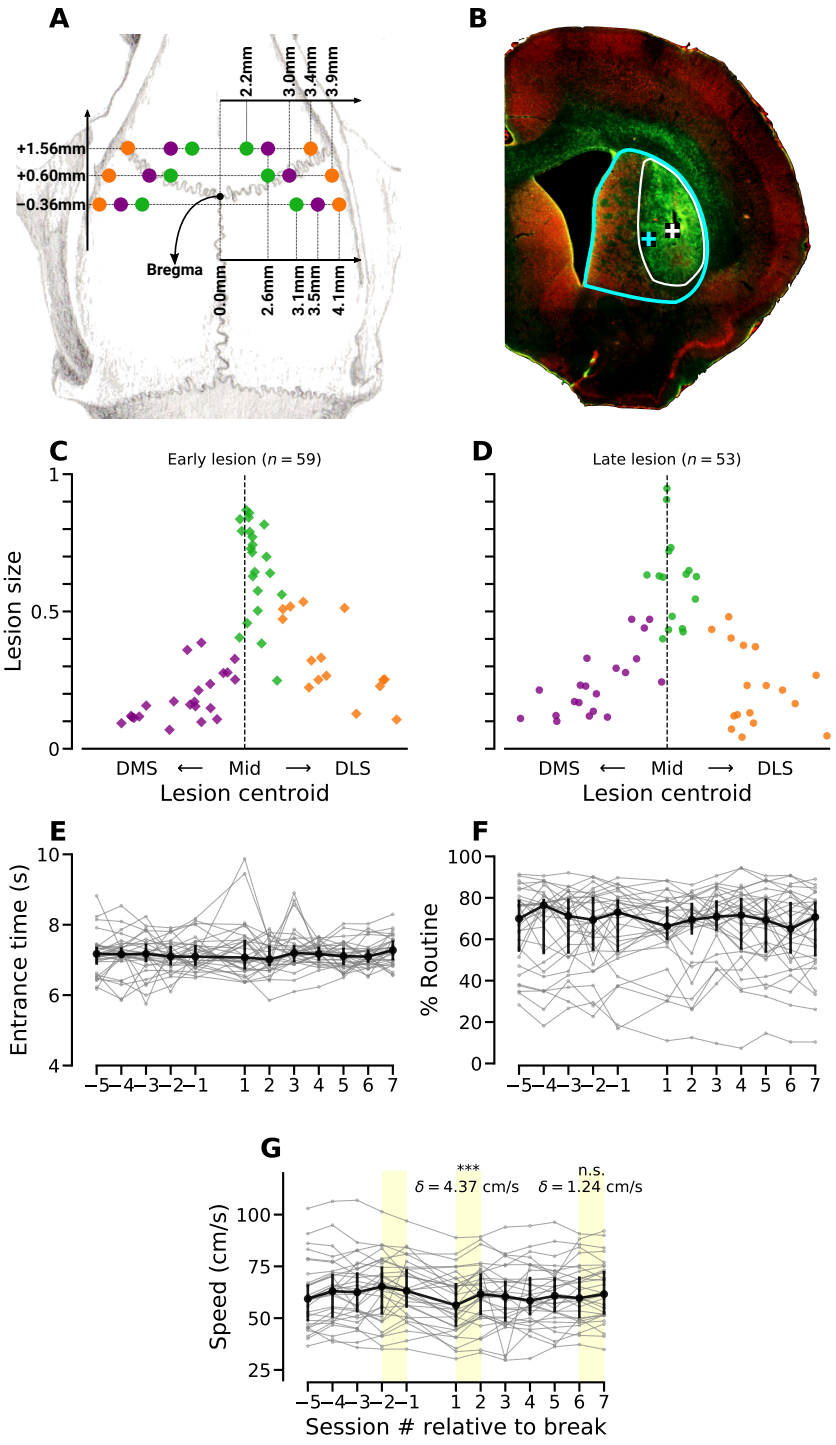


Figure S1: Dorsal striatum lesion quantification and effect of the two-weeks long break on task performance. Related to Figure 1.

(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).

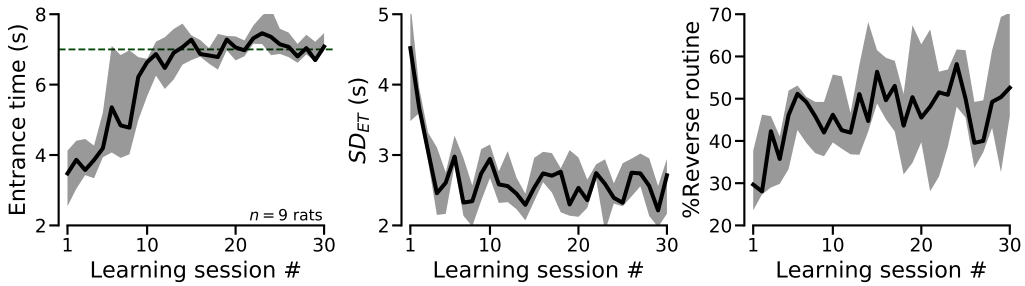


Figure S2: Performance improvement in the reverse treadmill task. Related to Figure 2. Left: Entrance time across learning sessions. Middle: Session-by-session standard deviation of ET. Right: Percentage of trials during which animals used the run-and-wait (reverse) routine.

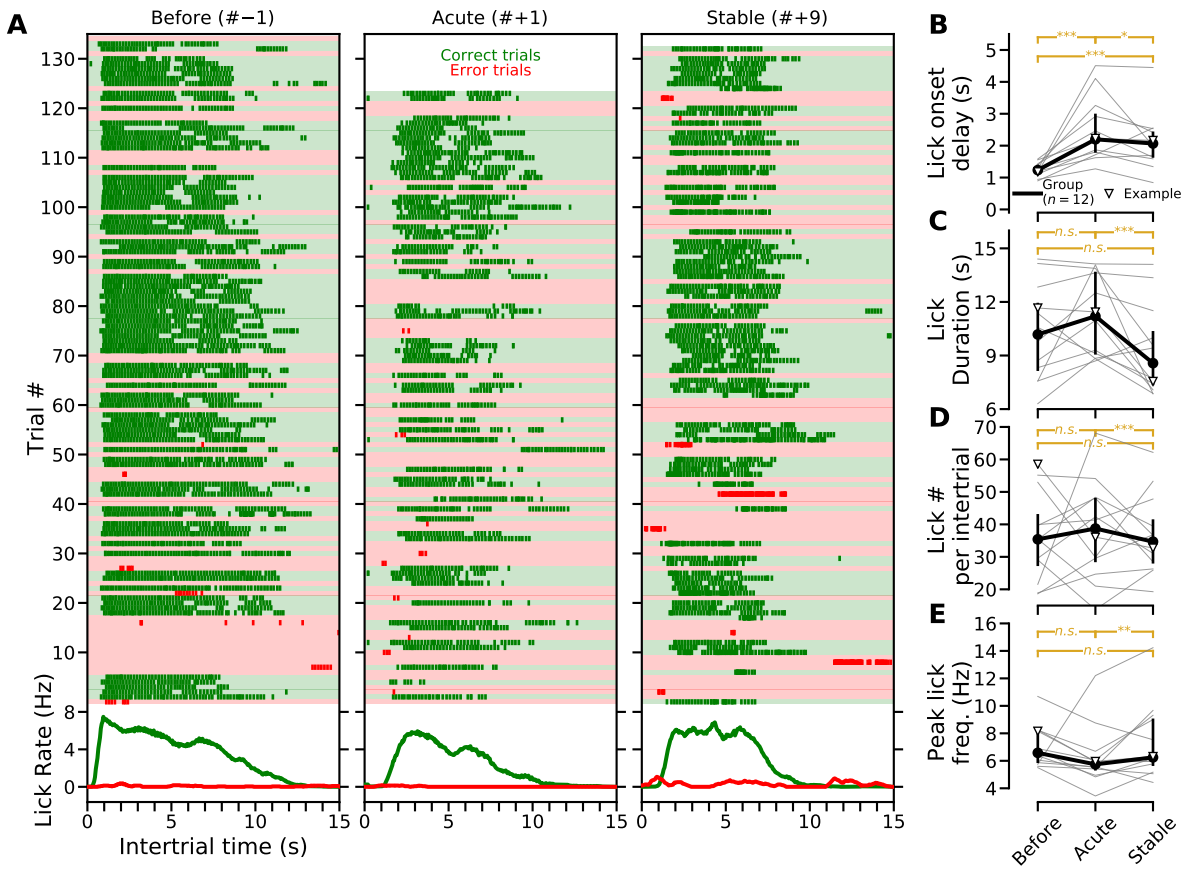


Figure S3: Licking behavior after dS lesion. Related to Figure 3. **(A)** Trial-by-trial licking pattern (top) and average lick rate aligned to intertrial onset for a single animal in 3 sessions (1 session just before and 2 after lesion). **(B-E)** Effect of dS lesion on lick onset delay (B), lick duration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

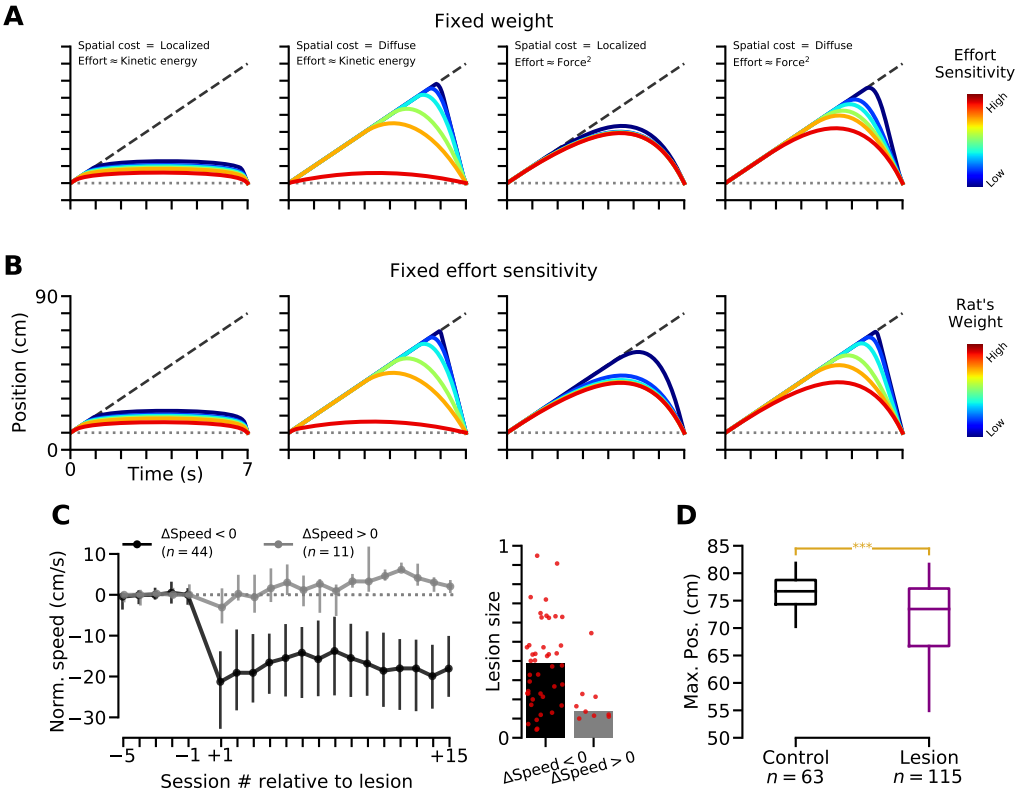


Figure S4: Optimal control models and effect of dS lesion on maximal position. Related to Figure 5. (A) Effect of effort sensitivity on optimal trajectories. Trajectories were obtained using 4 models that combine two approximation of effort and spatial costs. The cost of premature entrance in the reward 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, 25^{th} and 75^{th} percentiles; whiskers, 5^{th} and 95^{th} 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.