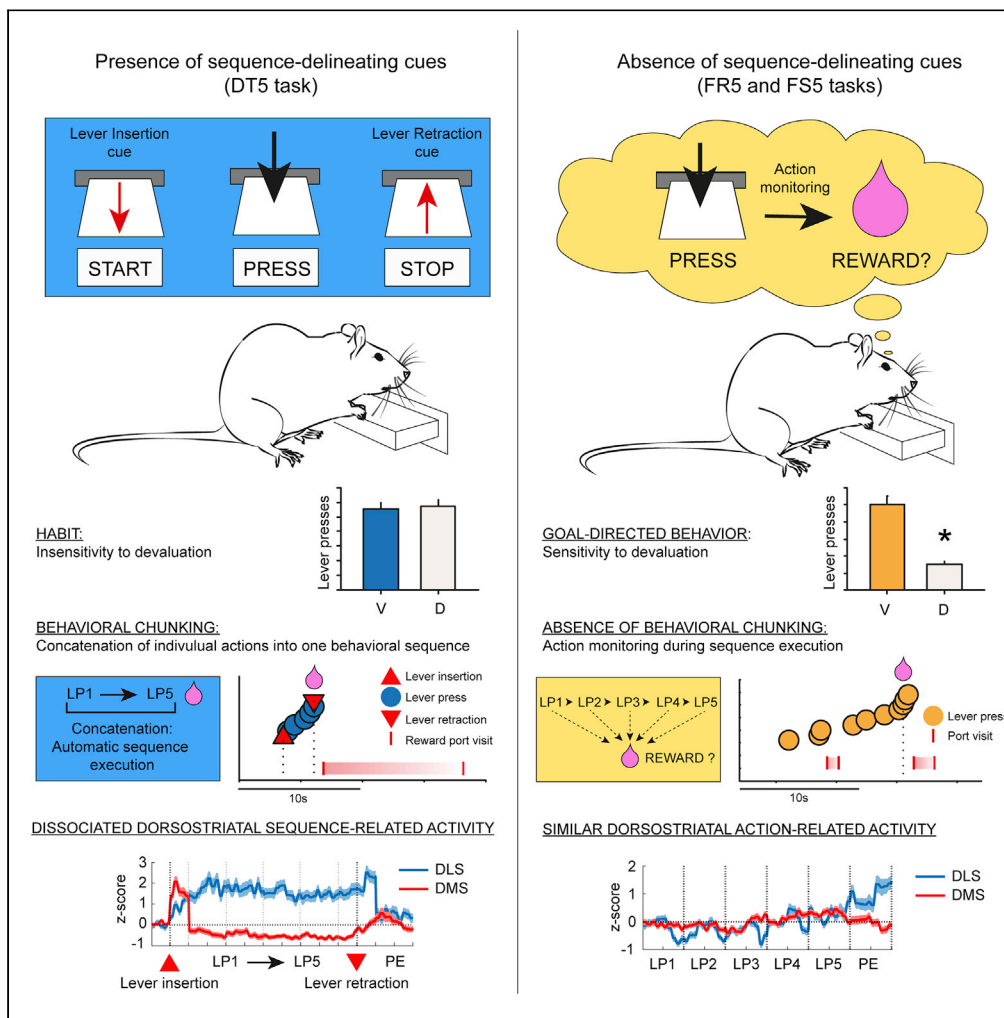**Article**

Lack of action monitoring as a prerequisite for habitual and chunked behavior: Behavioral and neural correlates

Youna Vandaele,
Patricia H. Janak

youna.vandaele@univ-poitiers.
fr

Highlights

Cues signaling action sequences boundaries promote habit and behavioral chunking

Requirements for action monitoring prevent habit and behavioral chunking

Skilled performance does not predict goal-directed vs. habitual behavior

Differential engagement of striatal subregions may reflect behavioral chunking

Article

Lack of action monitoring as a prerequisite for habitual and chunked behavior: Behavioral and neural correlates

Youna Vandaele[1,3,*] and Patricia H. Janak[1,2]

SUMMARY

We previously reported the rapid development of habitual behavior in a discrete-trials instrumental task in which lever insertion and retraction act as reward-predictive cues delineating sequence execution. Here we asked whether lever cues or performance variables reflective of skill and automaticity might account for habitual behavior in male rats. Behavior in the discrete-trials habit-promoting task was compared with two task variants lacking the sequence-delineating cues of lever extension and retraction. We find that behavior is under goal-directed control in absence of sequence-delineating cues but not in their presence, and that skilled performance does not predict goal-directed vs. habitual behavior. Neural activity recordings revealed an engagement of dorsolateral striatum and a disengagement of dorsomedial striatum during the sequence execution of the habit-promoting task, specifically. Together, these results indicate that sequence delineation cues promote habit and differential engagement of striatal subregions during instrumental responding, a pattern that may reflect cue-elicited behavioral chunking.

INTRODUCTION

Habits are relatively automatic or fixed behaviors that do not require planning or deliberation.[1] In contrast to goal-directed actions that rely on response-outcome associations, habitual actions are performed in absence of outcome representations, and are therefore insensitive to outcome devaluation or degradation of the instrumental contingency.[2,3] By definition, habits are also characterized by higher performance efficiency.

Under stable conditions, actions have been suggested to become "chunked" together into a behavioral unit, facilitating more efficient and consistent execution.[4–9] Although habit and behavioral chunking are conceptually related,[10,11] whether they constitute a differing expression of a single cognitive process or represent separate cognitive mechanisms remains unclear.[12–14] In addition, although behavioral chunking may lead to performance efficiency, some evidence suggest that performance efficiency or automaticity alone is not a clear indicator of whether or not a behavioral response is under habitual control, as defined by insensitivity to outcome devaluation.[12,15,16]

We have been interested in factors that might impact behavioral control when the same action sequence is required to earn rewards. Previously we found that responding under a free-running fixed-ratio-5 schedule (five lever presses required to earn each reward) is under goal-directed control, but by imposing a discrete trial structure on this response requirement, responding became habitual.[17] In this later procedure (termed discrete-trials fixed-ratio 5; DT5), each trial begins with lever insertion with lever retraction after five lever presses, followed by reward delivery. The reasons underlying these distinct behavioral mechanisms might include the utility of external cues signaling reward availability and delivery in the DT5 task as opposed to the need for internal monitoring and/or port checking in the FR5 task. Alternatively, habitual control might occur as behavior becomes highly regular and less variable, which we found was the case within days for the DT5 task, but not for the FR5 task. Here we sought to further investigate the behavioral mechanisms underlying the control of instrumental responding. To do so, we compared behavior within three different training procedures. The first was the DT5 task described above. The second two tasks also utilized a fixed-ratio 5 schedule of reinforcement but in absence of sequence initiation and cessation cues: 1) the standard fixed-ratio 5 (FR5) task, in which each sequence of five lever presses is followed by reward delivery;

[1]Department of Psychological and Brain Sciences, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD 21218, USA

[2]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, MD 21205, USA

[3]Lead contact

*Correspondence:
youna.vandaele@
univ-poitiers.fr

https://doi.org/10.1016/j.isci.2022.105818

and, 2) a version of the FR5 task in which mid-sequence reward port checks are penalized, termed the fixed-sequence 5 (FS5) task. We conducted tests of reward devaluation and analyzed indices relevant to automaticity and efficient performance.

In parallel, we monitored neural activity within the dorsomedial striatum (DMS) and the dorsolateral striatum (DLS). The dorsal striatum is linked to habit, skill learning, and behavioral chunking with distinct regional specificities.[6,7,13,18–24] Functional manipulations typically show that the dorsomedial striatum (DMS) is involved in goal-directed control early in training whereas the dorsolateral striatum (DLS) is engaged in the formation of habit and skill learning over more extended training.[18,19,22,25–29] How this shift in control from DMS to DLS occurs and regulates habit learning and/or behavioral chunking remains an important area of investigation.[30] Although habit formation and behavioral chunking are associated with the development of a task-bracketing activity in the DLS but not the DMS,[23,31–34] habit-like behavior correlates with dopamine release in the DMS but not the DLS in a new instrumental model of habit and dopamine release does not shift across striatal regions during extended training.[15] Further, it was recently shown that DLS and DMS compete, rather than cooperate, for the control of behavior, in that lesions of the DMS improved acquisition of skilled and automated action sequences.[16]

We have recently examined sequence-related activity during the execution of lever press sequences within the DT5 task that promotes habitual behavior and high-performance efficiency.[17,35] Notably, the activity pattern in DLS and DMS differed from previous studies and was not correlated with changes in performance during training despite continued improvements,[35] but instead appeared to reflect the cue-delineated behavioral sequence, congruent with a role in behavioral chunking. Here we sought to further explore this finding by comparing striatal activity within the DT5 task with that recorded within highly skilled action sequences not susceptible to behavioral chunking or habitual control.

Here, after analyzing sequence behavior across learning and probing for goal-directed control, we report in male rats the replication of our prior results showing that lever pressing is habitual in the DT5 task and goal-directed in the FR5 task.[17] Furthermore, we demonstrate that training rats in the FS5 task favor high-performance efficiency, but that behavior remains under goal-directed control. In both DT5 and FS5 procedures, rats learn to rapidly emit individual responses, a process congruent with skilled performance. Thus, measures of skilled performance fail to predict whether that behavior is under goal-directed or habitual control.[12]

Using these three well-characterized tasks, we observed in the DT5 task that average DLS spike activity was increased, and average DMS activity was decreased throughout the lever press sequence relative to pre-sequence activity, as we previously reported.[35] This pattern was specific to the DT5 task and was bracketed by neural responses to the lever insertion and retraction. In contrast, in the FR5 and FS5 tasks, mean dorsostriatal activity during goal-directed behavior was characterized by transient modulations around individual lever presses and relatively less excitation in the DLS relative to the DMS. This was especially notable when comparing the FS5 task to the DT5 task, two procedures with similar behavioral profiles. Further, activity tended to ramp upwards during sequence performance within the FS5, a task requiring ongoing behavioral monitoring, but not in the DT5, a task that does not require behavioral monitoring during sequence performance. Thus, relatively greater excitation in DLS and inhibition in DMS is correlated with the behavioral demonstration of habit and action chunking but not with skilled performance per se.

Together these results indicate that the rapid development of habitual performance is promoted by external cues signaling the start and/or end of action sequences, thereby eliminating requirements for action monitoring. Whether or not action monitoring is required determines the differential recruitment of striatal subregions for highly similar action sequence performance. We propose that lever cues in the DT5 task promote rapid action chunking and habit learning, associated with greater DLS excitation and DMS inhibition, whereas the presence of outcome expectations in the FR5 and FS5 tasks results in the absence of this clear population distinction in activity during performance of the five lever press response.

RESULTS

Differences in expression of habit and performance in the DT5, FR5, and FS5 tasks

Sensitivity to devaluation in the FS5 and FR5 tasks, but not the DT5 task

Male rats were trained in three different lever pressing tasks under fixed-ratio 5 schedules of reinforcement (Figures 1A–1C) and recordings were obtained from the DMS and DLS (Figures 1D, 1H, and 1L). The first
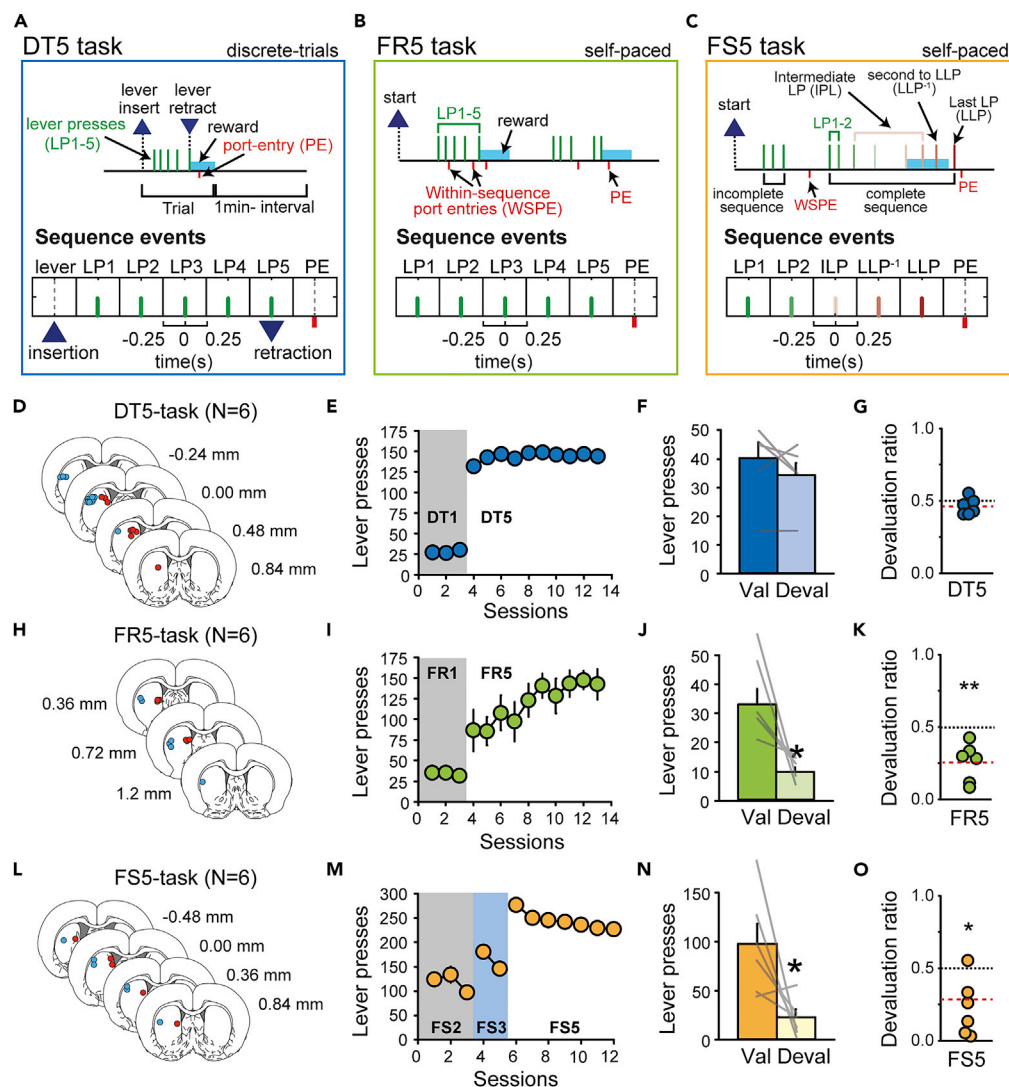
Figure 1. Differences in habitual learning in the DT5, FR5, and FS5 tasks

(A–C) Diagrams of task -structure and events in the DT5 (A), FR5 (B), and FS5 tasks (C).

(D) Electrode placements in the DT5 group.

(E) Mean number of lever presses across training sessions in the DT5 task.

(F) Mean number of lever presses in the valued and devalued conditions. Gray lines represent individual rats.

(G) Devaluation ratio of individual rats in the DT5 group. Devaluation ratios at or above the black dotted line indicate habit. The red dotted line indicates the group average.

(H–K) Same as D–G for the FR5 group.

(L–O) Same as D-G for the FS5 group. Error bars indicate SEM *p < 0.05; **p < 0.01.

group was trained in the discrete trial fixed-ratio 5 task (DT5 group N = 6; Figures 1A and 1D–1G). In this task, every trial began with the insertion of a lever after which rats had to complete a sequence of 5 lever presses to obtain access to a reward, signaled by lever retraction occurring at the fifth lever press (Figure 1A). Rats quickly learned to complete every ratio of the session as shown by the number of lever presses rapidly reaching the maximum allowed per session (i.e. 150 lever presses for 30 rewards; Figure 1E). After 10 DT5 sessions, we confirmed that, on average, responding was insensitive to outcome devaluation induced by sensory-specific satiety (Figures 1F and 1G; $F_{(1,5)} = 2.39$, $p > 0.1$; devaluation ratio t-test against 0.5: $t = 1.51$, $p > 0.1$).

The second group was trained in the FR5 task (N = 6; Figures 1B and 1H−1K). In this free-operant task, rats completed 5 lever presses on a lever that was continuously available, and in absence of any explicit reward

predictive cue (Figure 1B). The total number of lever presses progressively increased to reach the maximum possible around the last three sessions (Figure 1I). We confirmed at the end of this training that behavior was goal-directed as indicated by sensitivity to satiety-induced devaluation (Figures 1J and 1K; $F_{(1,5)} = 11.14$, $p < 0.05$; devaluation ratio t-test against 0.5: t = 4.54; $p < 0.01$).

The third group of rats was trained in the fixed sequence length 5 task (FS5, N = 6; Figures 1C and 1L−1O). In this free-operant task, rats had to complete a sequence of 5 consecutive lever presses without checking the reward port to obtain a reward (Figure 1C). A port entry before the completion of the ratio reset the ratio and that lever press sequence was considered incomplete. However, additional presses after completion of the ratio were without consequence and considered as part of the completed sequence (Figure 1C). Importantly, no external cues indicated the correct completion of the sequence. After initial training with FS2 and FS3 requirements, rats were tested in the final FS5 schedule for 7 sessions (Figure 1M and STAR methods). Sensitivity to satiety-induced devaluation was assessed at the end of FS5 training and revealed that behavior was under goal-directed control (Figures 1N and 1O; $F_{(1,5)} = 7.48$, $p < 0.05$; devaluation ratio t-test against 0.5: t = 3.37, $p < 0.05$). Collectively, these results extend previous findings[17,35] showing that lever pressing is habitual in presence of lever cues signaling the opportunity to respond and the termination of the sequence (DT5 task) but is under goal-directed control in their absence (FR5 and FS5 tasks).

Performance in the DT5 and FS5 tasks is associated with greater response rates and reduced variability in responding than performance in the FR5 task

Behavior across the three tasks also differed with respect to sequence learning and performance. Rats rapidly learned to complete the required ratio in a sequence of consecutive lever presses in the DT5 task by pressing the lever until its retraction on every trial (Figure 2A). This resulted in the suppression of within-sequence port entries by the second session (Figure 2B; session Friedman ANOVA $\chi^2 = 41.76$, $p < 0.0001$), and an increase in within-sequence response rate across the training period (Figure 2B; $F_{(9,45)} = 9.37$, $p < 0.0001$), accompanied by a decrease in trial-by-trial variability (Figure 2B; $F_{(9,45)} = 2.60$, $p < 0.05$). In contrast, in absence of lever cues in the FR5 task, rats regularly checked the port after a few lever presses and before completing the ratio (Figure 2C). This resulted in a high number of within-sequence port entries, a low within-sequence response rate, and higher trial-by-trial variability (Figure 2D; last three sessions DT5 vs FR5: within-sequence port entries Z = 2.88, $p < 0.01$; within-sequence response rate $F_{(1,10)} = 45.36$, $p < 0.0001$; coefficient of variation Z = 2.72, $p < 0.01$). Although the within-sequence response rate increased across sessions (session $F_{(9,36)} = 7.97$, $p < 0.0001$), it never reached the level observed in the DT5 task (Figures 2B and 2D). These results suggest that lever insertion and retraction in the DT5 task constitute salient cues that shape behavior and promote the expression of highly efficient and habitual responding. The rapid suppression of within-sequence port entries in the DT5 task suggests that rats learned to concatenate individual lever presses into the execution of a single behavioral sequence, a process referred to as behavioral chunking. Performance of the action sequence requires little ongoing monitoring since sequence termination and reward availability is signaled by lever retraction.

In line with our predictions, the sequence requirement of the FS5 task greatly improved performance as rats learned to suppress port checking during sequence execution (Figures 2E and 2F). The number of within-sequence port entries declined over training and was significantly lower than in the FR5 task (Figure 2F; main effect of session $F_{(6,30)} = 12.12$, $p < 0.0001$; last three sessions FR5 vs FS5: Z = 2.88; $p < 0.01$). During complete sequences, the within-sequence response rate was higher than in the FR5 task (main effect of task $F_{(1,9)} = 43.10$, $p < 0.0001$) and did not significantly differ from the DT5 task (Figure 2F; FS5 vs DT5: main effect of task $F_{(1,10)} = 0.26$, $p > 0.5$). Trial-by-trial variability within the FS5 task was similar to the DT5 task (Figure 2F; FS5 vs DT5: main effect of task $F_{(1,10)} = 2.87$, $p > 0.1$). In contrast to the DT5 task, the FS5 task required rats to abstain from checking the port prematurely, thereby requiring active monitoring of behavioral output and/or the passage of time.

Within-sequence action timing distinguishes DT5, FS5, and FR5

During the last training session, we observed a significant task difference in the mean port entry latency following sequence completion (Figure 2G; $F_{(2,15)} = 15.74$, $p < 0.001$) and a trend for trial-by-trial variability in this measure (Figure 2G; $F_{(2,15)} = 3.50$, $p = 0.057$). Post-hoc analysis revealed that reward retrieval was faster in the DT5 task and slower in the FS5 task (DT5 vs FR5: $p < 0.05$, DT5 vs FS5: $p < 0.001$, FR5 vs FS5: $p < 0.05$). These results are consistent with behavioral chunking presumably occurring in the DT5 task and the maintenance of behavioral monitoring and deliberation in the FS5 task (Dezfouli and Balleine, 2013).
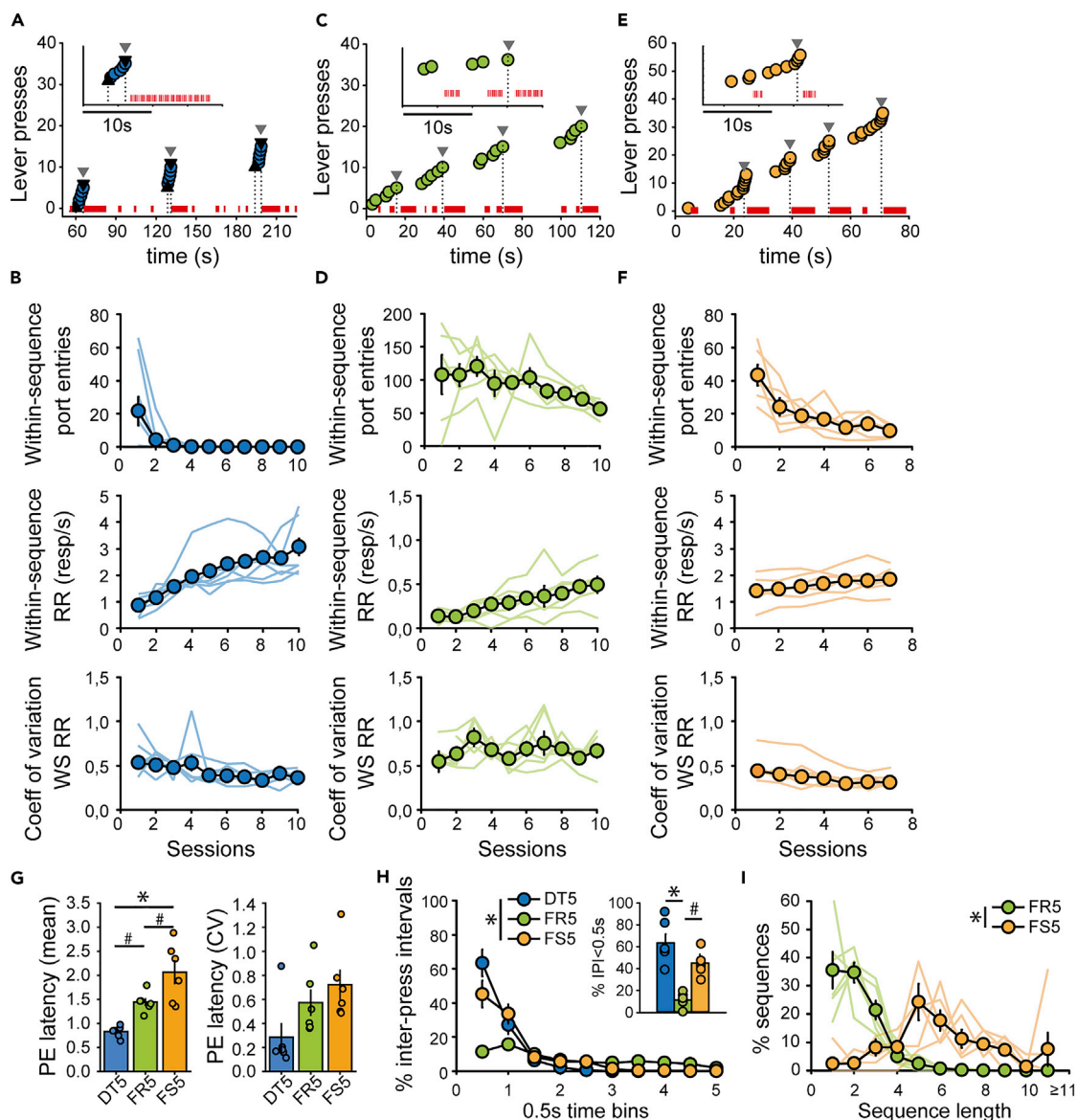
Figure 2. Task differences in performance

(A) Microstructure of instrumental performance in a representative rat during DT5 training. Circles represent lever presses; red ticks indicate time in the reward port; upward and downward triangles represent lever insertion and retraction, respectively. Inset: enlargement of the second trial.

(B) Mean number of within-sequence port entries (top), within-sequence response rate (WS RR; middle), and coefficient of variation of within-sequence response rate (bottom) across DT5 sessions.

(C and D) Same as A-B for the FR5 task.

(E and F) Same as A-B for the FS5 task. In C and E, downward triangles indicate reward delivery.

(G) Mean (left panel) and coefficient of variation (right panel) of port entry latency during the last training session in the DT5, FR5, and FS5 tasks. #p < 0.05; *p < 0.001.

(H) Mean distribution of inter-press intervals (IPI) during the last training session in the DT5, FR5, and FS5 tasks. *p < 0.0001. Inset: percentage of IPI shorter than 0.5s in the three tasks. *p < 0.001; #p < 0.01.

(I) Mean distribution of sequence lengths during the last training session in the FR5 and FS5 tasks. *p < 0.0001. Error bars represent the SEM Transparent lines in B, D, F and I represent individual rats.

Analysis of the distributions of inter-press intervals (IPI) during the last training session in each task revealed that while FR5 performance is characterized by a wide range of IPI, performance in the DT5 and FS5 tasks was characterized by a majority of IPI shorter than 0.5s (Figure 2H; bins by task interaction $F_{(18,135)} = 10.97$, p < 0.0001; %short IPI: main effect of task $F_{(2,15)} = 15.95$, p < 0.001; post-hoc FR5 vs DT5: p < 0.001; FR5 vs

FS5: p < 0.01; DT5 vs FS5: p = 0.16). Overall, the performance in the FS5 task was more comparable to the DT5 task but strongly differed from the FR5 task. Although the FR5 and FS5 tasks shared the same free-operant structure with higher uncertainty about reward delivery, rats trained in these 2 tasks adopted very different strategies. More specifically, rats made a majority of short sequences in the FR5 task (1-3 lever presses) whereas performance in the FS5 task was characterized by a wider range of sequence lengths, with a peak at 5 lever presses (Figure 2I; task*length $F_{(10,90)}$ = 21.38, p < 0.0001). In fact, rats trained in the FS5 task learned to avoid making incomplete sequences by pressing more persistently on the lever, which resulted in a high proportion of sequences longer than 5 lever presses (Figure 2I; 54.3 ± 10.7%). The lack of a clear sequence termination signal and the requirement for behavioral monitoring is apparent in this relatively broad distribution of sequence length in the FS5 task suggesting that an invariant action unit, or "behavioral chunk," does not form under this training regimen of a few weeks.

Although lever pressing performance did not differ on average between the DT5 and FS5 tasks (Figure 2H), closer inspection of the microstructure of operant behavior seems to reveal more subtle task differences (Figures 2A and 2E). Notably, the response rate was not homogeneous during the execution of complete sequences in the FS5 task, as rats tended to press faster on the lever toward the end of the sequence (Figure 2E). To address more subtle within-sequence differences in responding across tasks, we analyzed peri-event time histograms of the lever presses aligned to the beginning (LP1), middle (LP3, ILP), and end (LP5, LLP) of the sequence during the first and last training sessions (Figure 3). For both the first and last sessions, we found significant task by time bins interactions after the first lever press (Figures 3A and 3D; first session post-LP1: $F_{38,285}$ = 2.27, p < 0.0001; last session post-LP1: $F_{38,285}$ = 5.08, p < 0.0001), around the intermediate lever press (Figures 3B and 3E; first session peri-ILP: $F_{36,270}$ = 1.96, p < 0.01; last session peri-ILP: $F_{36,270}$ = 2.28, p < 0.001) and before the last lever press (Figures 3C and 3F; first session pre-LLP: $F_{38,285}$ = 2.29, p < 0.0001; last session pre-LLP: $F_{38,285}$ = 7.33, p < 0.0001). The main effects of the task were significant for each event and session (F-values>10.7, p values<0.01). However, post-hoc analyses revealed that, while the FR5 task differed from the other tasks for every event and on both sessions (p values<0.05), performance on the DT5 and FS5 tasks only significantly differed on the last session, after the first lever press (Figure 3D, p < 0.05) and before the last lever press (Figure 3F, p < 0.01). The frequency of inter-press intervals shorter than 1-s was nonetheless comparable between the two tasks.

Distinct sequence-related striatal neural activity in the tasks promoting habitual vs goal-directed behavior

Training in the three tasks favored distinct behavioral phenotypes, differing in terms of sensitivity to devaluation and performance. We next compared DMS and DLS activity in these 3 tasks to determine how specific activity patterns might relate to instrumental performance and strategy. Medium spiny neurons (MSN) were distinguished from interneuron populations using firing rates and waveform properties (Figures S1A–S1E; STAR Methods).[34,36,37] Putative MSNs represented 94%, 87%, and 91% of the recorded units in the DT5, FR5, and FS5 groups, respectively. Due to the low number of interneurons, we restricted analyses in this study to the population of putative MSN. Electrode placements were similar across groups (Figure S1).

As previously reported,[35] in the DT5 task, task-related neurons represented on average 86.3 ± 1.8% of the MSN population (DLS, range N = 16-34/session; DMS, range N = 50-64/session). The relative proportions of task-related increases and decreases differed with a majority of increases in DLS (Figure S2; 75.7 ± 2.2% of task-related units) and a majority of decreases in DMS (Figure S2; 73.4 ± 1.4% of task-related units). These differences led to differences in the average normalized activity of task-related neurons as a function of brain region and across events (Figures 4A–4D; main effect of region, $F_{(1,749)}$ = 68.33, p < 0.0001; region*-event interaction, $F_{(11,8239)}$ = 25.3, p < 0.0001). Specifically, relative to baseline, there was on average a sustained excitation throughout the behavioral sequence, from the insertion of the lever to the port entry following reward delivery, in the DLS (Figure 4A). In contrast, DMS activity was characterized by a sustained depression in firing during lever pressing and excitation at the boundaries of the sequence, at the time of lever insertion and retraction (Figure 4B).

In the FR5 task, the relative proportion of increases and decreases in spike activity related to task events was 35.7 ± 3.5% (increases) and 64.3 ± 3.5 (decreases) for DLS and 44.8 ± 2.7 (increases) and 55.2 ± 2.7% (decreases) for DMS (Figure S2). The mean activity of task-related neurons in DLS and DMS, representing on average 86.9 ± 2.8% of recorded MSNs (DLS, range N = 21-43/session; DMS, range N = 41-67/session),
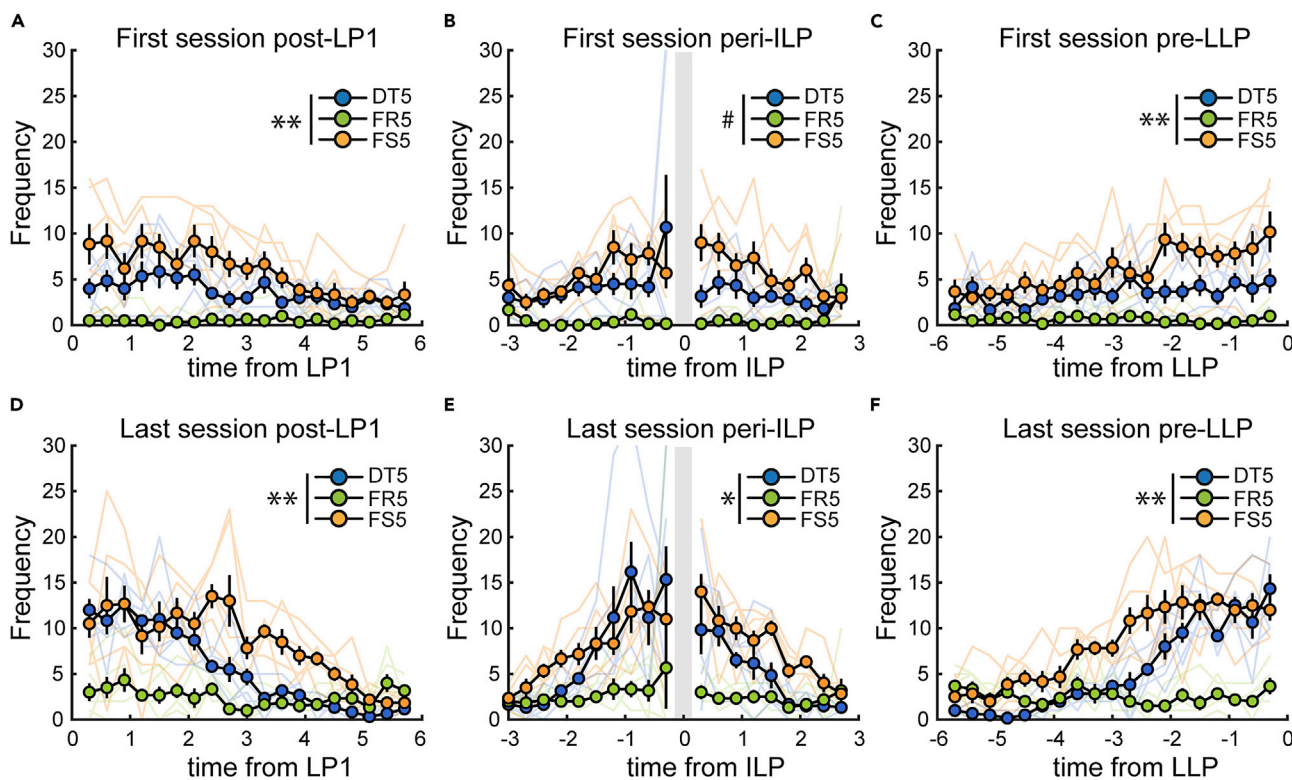
Figure 3. Local task differences in response rates at the beginning, middle, and end of lever press sequences

(A–C) Peri-event time histograms of the lever presses aligned to the first lever press (A), intermediate lever press (B), and last lever press (C) during the first training session in the DT5, FR5, and FS5 tasks.

(D–F) Same as A–C for the last training session. **p < 0.0001; *p < 0.001; #p < 0.01. Error bars represent the SEM Transparent lines represent individual rats.

differed from the DT5 group, with transient increases around each individual lever press (Figures 4D–4F). There was no main effect of region ($F_{(1,757)}$ = 2.31, p > 0.1). DLS and DMS activity significantly differed across events (Figures 4D and 4E; region*event interaction, $F_{(11,8327)}$ = 10.64, p < 0.0001) but not across sessions (Figure 4F; main effect of session: $F_{(9,757)}$ = 0.93, p > 0.4; region*session interaction, $F_{(9,757)}$ = 1.72, p = 0.08). While DLS mean population activity tended to peak at the time of lever presses (Figures 4D and 4F), DMS activity peaked after each press with greater increases following the last lever press in the sequence (Figures 4E and 4F).

In the FS5 task, the average proportion of task-related neurons across sessions was 90.8 ± 6.1% (DLS, range N = 7-27/session; DMS, range N = 42-60/session), and the relative proportion of increases and decreases in spike activity related to task events was 58.1 ± 3.8% (increases) and 41.9 ± 3.8 (decreases) for DLS and 46.3 ± 1.9 (increases) and 53.6 ± 1.9% (decreases) for DMS (Figure S2). Again, there was no significant main effect of region ($F_{(1,466)}$ = 0.13, p > 0.5), but an interaction of region x events in this task (Figures 4G and 4H; region*event interaction $F_{(11,5126)}$ = 12.02, p < 0.0001). Specifically, mean task-related neural activity in the DLS transiently decreased after each lever press and increased around the port entry whereas activity in the DMS slowly ramped up toward the end of the lever press sequence (Figures 4G–4I). We did not observe any change in activity across training sessions (Figure 4I; $F_{(6,466)}$ = 0.93, p > 0.4).

The above analysis demonstrates distinct patterns of activity as a function of region in each task. In the two tasks with goal-directed behavior, there was a distinct lack of overall mean population differences in firing between DLS and DMS. We next asked if neural activity patterns differed among the three tasks using a statistical comparison. A 3-way ANOVA with task, region, and session as between-factors (analysis restricted to first 7 sessions) and task event as a repeated factor, revealed significant effects of task ($F_{(21,425)}$ = 21.55, p < 0.0001), and a task by region ($F_{(21,492)}$ = 6.14, p < 0.01) and task by region by event ($F_{(22,16,412)}$ = 11.25, p < 0.0001) interaction. Post hoc analysis revealed significant differences in DLS and DMS activity between the DT5 and the other two tasks (p values<0.0001) but not between the FR5 and
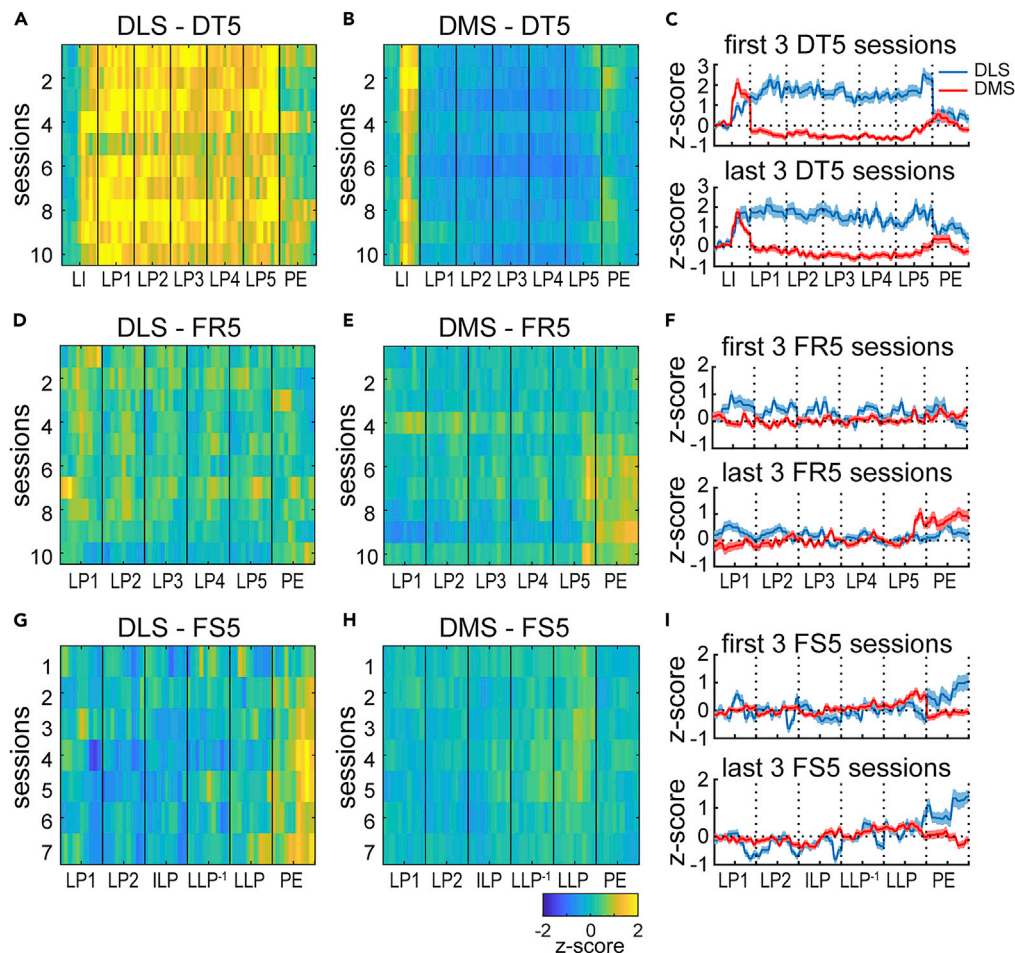
Figure 4. Distinctions among sequence-related activity in the DT5, FR5, and FS5 tasks

(A and B) Heatmaps of mean task-related neural activity along the behavioral sequence and across training sessions in the DLS (A) and DMS (B).

(C) Mean ± SEM PSTH of DLS and DMS task-related neurons over the first three (top) and last three (bottom) DT5 training sessions.

(D–F) Same as A–C for the FR5 task.

(G–I) Same as A–C for the FS5 task. See also Figures S1 and S2.

FS5 tasks (p values>0.3). Together, these results reveal differences in mean activity in the DLS and DMS that distinguish among the three tasks, with a relatively greater excitation and inhibition in DLS and DMS, respectively, in the DT5 task in which behavior is habitual.

Persistent task differences in activity pattern after controlling for differences in performance

The above analyses revealed important task differences in activity patterns that could be attributed to differences in habitual vs goal-directed control and behavioral chunking. However, the sensorimotor dynamics during sequence execution in the DT5 task robustly differed from the FR5 task and slightly differed from the FS5 task (Figure 3). Since the dorsostriatal activity is strongly influenced by sensorimotor dynamics,[38–40] we were especially interested in whether these differences in performance could account for the distinct activity patterns observed in the DT5 task and FS5 task.

To control for local differences in response rate between the DT5 and FS5 tasks at the beginning, middle, and end of the sequence, we sub-selected trials in each recording session based on the inter-press intervals following the first lever press, around the intermediate lever press or before the last lever press (Figure 5A). Specifically, only trials with an inter press interval less than 1-s after the first lever press, less than 2-s before
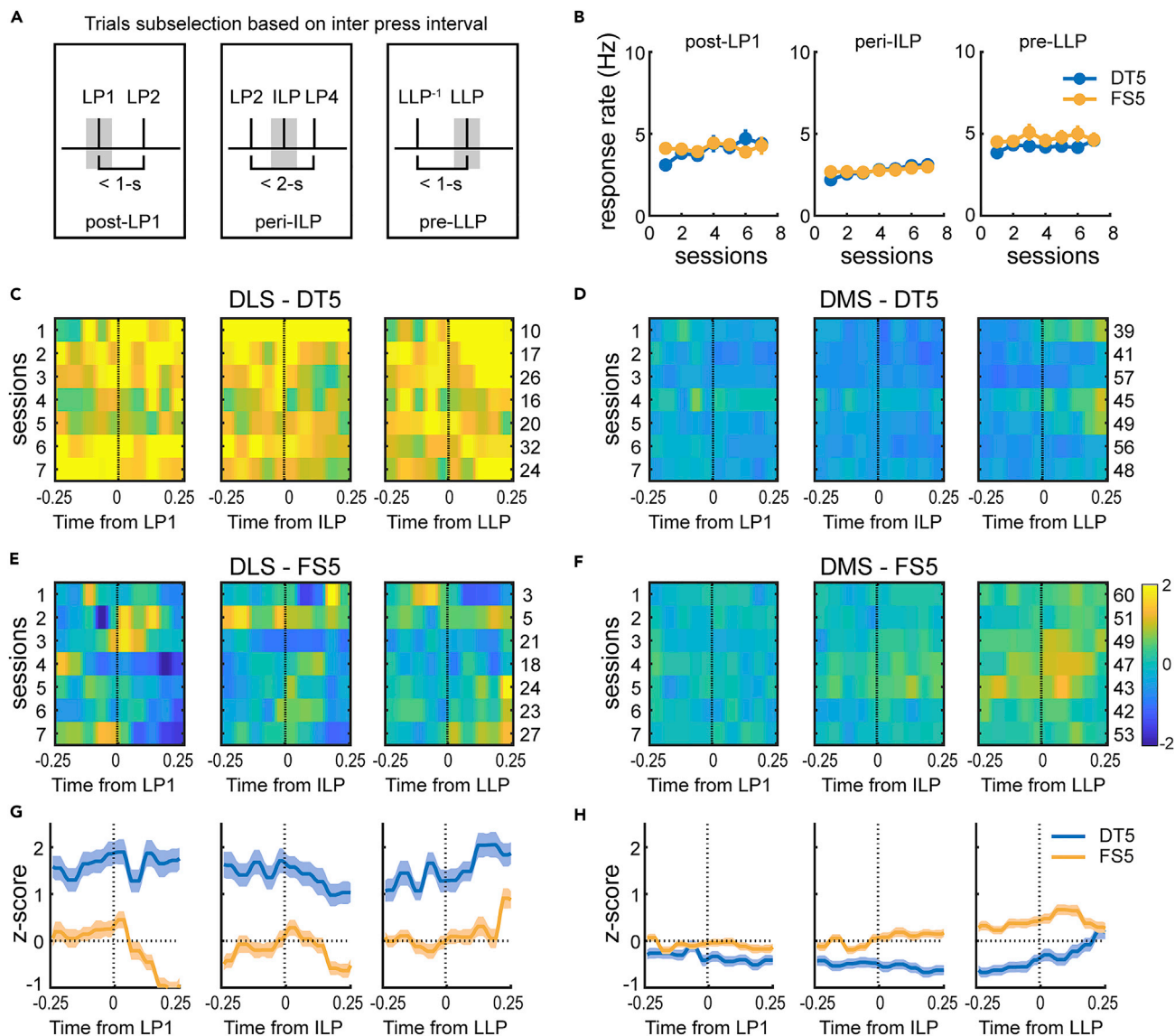
Figure 5. Persistent task differences in striatal activity after controlling for local differences in the response rate

(A) Diagram representing inter-press interval criteria for trials sub-selection at the beginning, middle, and end of the sequence.

(B) Mean local response rate ($\pm$SEM) after the first lever press (LP1; left panel), around the intermediate lever press (ILP; middle panel), and before the last lever press (LLP; right panel) across training sessions in the DT5 and FS5 tasks, after trials sub-selection.

(C and D) Heatmaps of the mean neural activity in DLS (C) and DMS (D) around LP1 (left panel), ILP (middle panel), and LLP (right panel) across sessions in the DT5 task after trials sub-selection.

(E and F) Same as C and D for the FS5 task. Numbers on the right of the heatmaps indicate the number of neurons.

(G and H) Mean PSTH ($\pm$SEM) of DLS (G) and DMS (H) neural activity in the FS5 task and in the DT5 task after trials sub-selection.

and after the middle press, and less than 1-s before the final lever press were considered for analysis (Figure 5A). Using this approach, local response rates at the beginning, middle, and end of lever press sequences were comparable between tasks and across sessions (Figure 5B; effect of task: F-values<2.32, p values>0.1, task by session interaction: F-values<0.9, p values>0.4). However, the average task-related activity in DLS and DMS remained significantly different between tasks following trial sub-selection (Figures 5C–5H). 2-way ANOVAs with task and session as between-factors and time bin (0.25 ms each) as a repeated factor revealed a main effect of the task in the DLS around the first lever press ($F_{(1,252)}$ = 33.99, p < 0.0001), the intermediate lever press ($F_{(1,252)}$ = 29.08, p < 0.0001), and the last lever press ($F_{(1,266)}$ = 40.98, p < 0.0001; Figure 5G). Compared to the FS5 task, the mean activity in the DLS showed

a greater excitation compared to the baseline in the DT5 task (Figures 5C, 5E, and 5G). In the DMS, we also found the main effect of task around the first lever press ($F_{(1,666)} = 6.75$, $p < 0.01$), the intermediate lever press ($F_{(1,666)} = 29.29$, $p < 0.0001$), and the last lever press ($F_{(1,666)} = 38.05$, $p < 0.0001$; Figure 5H). Mean activity in the DMS showed an inhibition in the DT5 task (Figures 5D and 5H), but remained near baseline in the FS5 task (Figures 5F and 5H). Thus, neuronal activity patterns in the DT5 task significantly differed from the FS5 task despite similar response rates at the start, middle, and termination of the sequence.

The analysis above was restricted to local neuronal activity at the start, middle, and termination of the sequence. We next sought to compare population activity in the DT5 and FS5 tasks during the full behavioral sequence. For each recording session, we matched trials from the DT5 dataset to the middle tertile of sequence response rates in the FS5 dataset (Figure 6A). Using this approach, the within-sequence response rate was overall higher in the FS5 task than in subsampled trials from the DT5 task (Figure 6B; main effect of task: $F_{(11,332)} = 5.56$, $p < 0.05$). However, session-by-session analysis revealed no task difference (F-values<2.19, p values>0.1). The mean activity in DLS and DMS during training in the DT5 task remained significantly different from the FS5 task (Figures 6C–6E). A 3-way ANOVA with task, region, and session as between-factors and event as a repeated factor revealed a main effect of task ($F_{(1,691)} = 42.81$, $p < 0.0001$), a task by region interaction ($F_{(1,691)} = 132.18$, $p < 0.0001$), a task by event interaction ($F_{(11,7601)} = 4.72$, $p < 0.001$) and a task by region by event interaction ($F_{(11,7601)} = 11.20$, $p < 0.0001$). Compared to the FS5 task, the average task-related activity in the DT5 task showed greater excitation in the DLS and inhibition in the DMS (Figures 6C and 6E), while the mean activity in both regions during the FS5 task tended to ramp and to return to baseline after any deviation during the five lever-press responses (Figures 6D and 6E). Together, these results suggest that higher performance alone cannot explain the sequence-related activity pattern of sustained DMS inhibition or DLS excitation observed in the DT5 task.

DISCUSSION

In this study, we report distinctions in behavioral strategy for three different instrumental procedures sharing the same response requirement. The DT5 task was the only task promoting habitual responding, although both the DT5 and FS5 tasks resulted in equivalent efficient, skilled performance. We conclude that the discrete lever cues predicting reward availability and delivery that bracket and constrain the performance of the lever-press sequence account for the susceptibility of this behavior to habitual control, and that automatic low-variability instrumental performance per se is not sufficient to generate habit. The similarity of lever-pressing performance within the DT5 and FS5 tasks allowed for the comparison of striatal activity under habitual versus goal-directed control. We found that habitual control was associated with relatively greater activity in the DLS and relatively suppressed activity in the DMS and that goal-directed control in the more cognitively demanding FS5 task was associated with ramping activity during the course of the sequence. Together, these findings suggest that sensorimotor variables alone may not be sufficient to account for population-level striatal neural activity patterns, but that the presence or absence of action monitoring/cognitive control is also relevant.

We replicated previous results showing that the presence of lever cues in the DT5 task promoted habitual learning with high-performance efficiency whereas their absence in the FR5 task resulted in a goal-directed behavior with low-performance efficiency.[17] Here, we extend these findings, by showing that training in the FS5 task allowed for sequence learning but did not promote habitual behavior. In both the FR5 and FS5 tasks, the lever was continuously extended and did not constitute a dynamic cue predictive of reward availability or delivery. The high number of within-sequence port entries in the FR5 task and the rats' tendency to keep pressing on the lever after the reward delivery for varying sequence lengths in the FS5 task, together with longer port entry latency upon sequence completion, reveal the uncertainty associated with reward delivery in these conditions and highlight the ongoing behavioral monitoring during sequence execution. For these reasons, sequence performance in the FR5 and FS5 tasks is unlikely to reflect a "chunked" behavior, since these by definition are invariant concatenated actions. It was recently suggested that high reinforcer predictability could favor the development of habit by reducing the attentional demand on behavioral monitoring.[41–46] This could explain the rapid development of habit in the DT5 task, in which the lever retraction directly signals the completion of the sequence and can ameliorate any requirement for monitoring the number of lever press responses. This is congruent with the notion that the lever retraction promoted behavioral chunking such that individual actions are executed as a behavioral unit without outcome representation until the completion of the sequence. The rapid decrease in within-sequence port entries across sessions (Figure 2), and even within the first session,[35] directly supports this hypothesis.
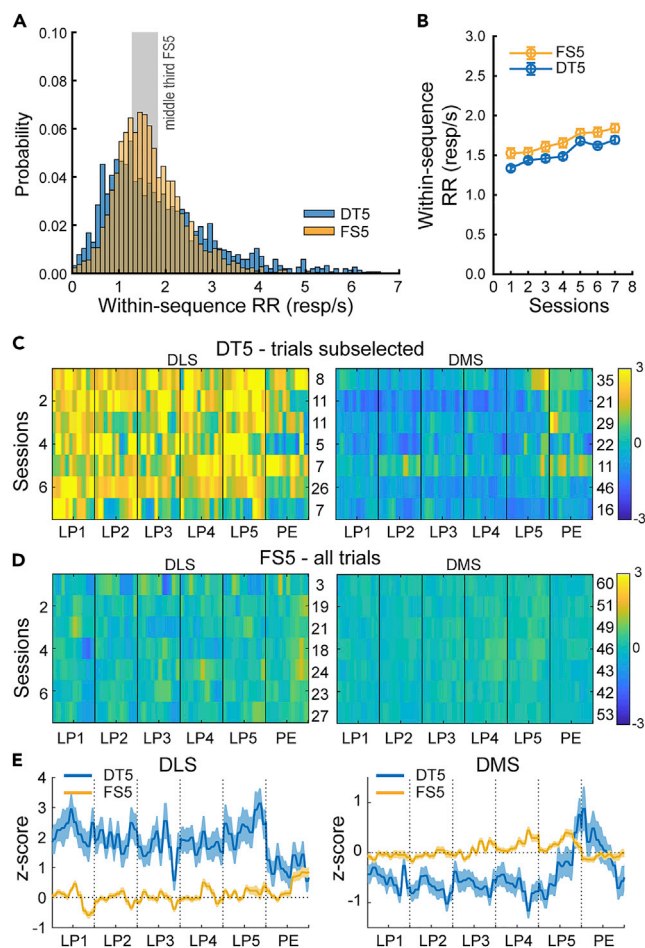
Figure 6. Persistent task differences in striatal activity after controlling for average difference in within-sequence response rate

(A) Distribution of within-sequence response rates across recording sessions in the DT5 and FS5 tasks. The gray area represents the middle tertile of within-sequence response rates in the FS5 task, in which we selected trials from the DT5 task.

(B) Mean within-sequence response rate (±SEM) across training sessions in the DT5 and FS5 tasks after the subsampling of DT5 trials.

(C and D) Heatmaps of the mean task-related neural activity in DLS (left) and DMS (right) across sessions in the DT5 task after trials subsampling (C) and in the FS5 task (D). Numbers on the right of the heatmaps indicate the number of neurons.

(E) Average PSTH (±SEM) of DLS (left) and DMS (right) neurons in the FS5 task and in the DT5 task after trial subsampling.

In the DT5 task, it is possible that behavioral chunking and habit form because lever press sequences are physically constrained by lever presentation and retraction. Thus, it is not clear whether the presentation of external cues is sufficient to explain behavior in this task or whether constraining action sequences is also required. It was recently shown that replacing lever cues with an auditory cue in a discrete-trial procedure can also support habitual learning, which suggests that constraining action sequences by inserting and retracting the lever is not a requirement for habit formation.[45] Though, to directly address this question in our conditions, it would be interesting to show that habitual learning can form under a free-operant FR5 schedule in presence of external cues that do not interfere with ongoing actions. For instance, insertion and retraction of a second inactive lever could provide discriminative cues signaling reward availability and delivery, without interfering with lever-pressing behavior. Another particularity of the DT5 task is the 1-min time limit for sequence completion. It could be argued that this time constraint contributes to faster sequence execution in this task compared to the FR5 and FS5 tasks. However, during DT5 training several rats completed the ratio is less than 20s from the first session despite experiencing very rare omission (1 trial or less). Thus, it is unlikely that the observed increase in response rate resulted from the time constraint imposed on trial duration.

Overall, these findings are not consistent with prior literature suggesting that random interval schedules of reinforcement,[47,48] and the associated temporal uncertainty[49,50] promote habit. Our findings are however consistent with other studies employing discrete trials and/or discriminated operant procedures,[41,43–45] such as in the DT5 task, where the presentation of the lever can act as a discriminative cue. Furthermore, several neuro-computational models suggest that arbitration between habitual and goal-directed decision-making could depend on the relative uncertainty of each system's predictions.[51–53] Notably, the low uncertainty about reinforcement learning predictions derived from lever cues could explain the rapid development of habit in the DT5 task.

In contrast to the DT5 task, the sequence requirement and the uncertainty about reward delivery in the FS5 task may require rats to maintain a representation of the outcome by keeping track of the time elapsed or the number of lever presses made before checking the port.[54] This cognitive demand may prevent the development of habit and explain rats' sensitivity to outcome devaluation.[55] Interestingly, it was recently shown that introducing uncertainty about task contingencies before testing can be sufficient to re-engage goal-directed control on an otherwise, habitual behavior.[46,56,57] These results show that performing an action sequence in a highly skilled manner and with low variability does not imply that responding is habitual and, by extension, that goal-directed control should not be equated with a lack of automaticity, in agreement with recent findings.[12] Relatedly, sequence learning in the FS5 task, even when executed quite expertly, does not necessitate behavioral chunking. Since rats had to monitor progress during sequence execution, we consider that behavioral chunking was prevented in that individual actions cannot be executed as a single behavioral sequence without ongoing behavioral monitoring and mid-sequence outcome representation. It is worth noting that behavioral chunking cannot be inferred from rats' behavior in the FS5 task, the low number of within-sequence port entries being a task requirement. Future studies are thus needed to directly address this hypothesis.

The average activity of task-related neurons in DMS and DLS differed across the three tasks. In the DT5 task, we observed dissociated sequence-related activity in DMS and DLS characterized by greater mean increased firing throughout the behavioral sequence in the DLS, and greater mean decreased firing during lever pressing with increases at the boundaries of the sequences in the DMS (Vandaele et al. 2019a), in agreement with the relative proportions of task-related excited and inhibited units. The distinct activity patterns of DLS excitation and DMS inhibition in the only task promoting habit might seem to indicate that this pattern of sequence-related activity encodes habit. However, the absence of change across sessions suggests otherwise, since habit formation is expected to occur from the fifth to the 10th session, based on prior experiments.[17] Instead, we interpret the activity patterns observed in the striatum of rats performing the DT5 task as reflecting the cue-delineated behavioral chunking, and surmise that this chunking may facilitate the development of or induce sensitivity to habitual control. In marked contrast, the mean modulation of spiking activity in the FR5 and FS5 tasks was much less distinct comparing DMS and DLS reflecting the more balanced proportions of units with task-related increases and decreases. Of note, in the FS5 task, mean neural activity tended to ramp up as rats progressed in the execution of the sequence. In fact, we have recently shown that DMS neurons track progress across sequence execution within the FS5 task,[54] a neural pattern congruent with ongoing behavioral monitoring rather than the production of a chunked set of actions.

The role of dorsal striatum in motor control and skill learning raises the possibility that task differences in the dorsostriatal activity reflect differences in sensorimotor dynamics.[24,39,58–60] Although instrumental performance in the FR5 task strongly differed from the DT5 task, the analysis of complete sequences in the FS5 task indicates a performance more comparable to DT5 performance with similar within-sequence response rates during several sessions (i.e. sessions 3-5) and low trial-by-trial variability in this measure (Figure 2). Of note, the activity patterns in DLS and DMS significantly differed between the DT5 and FS5 tasks after controlling for local differences in response rate (Figures 5 and 6), suggesting that differences in sensorimotor dynamics are not sufficient to explain the results reported here. Further, if sequence-related activity in DLS and DMS neurons merely reflected differences in sensorimotor dynamics, one would expect the dorsostriatal activity to change in step with performance improvement across sessions in the DT5 task. However, at the level of population activity, this is not what we observed. We note differences in activity patterns cannot be attributed to differences in electrode placements across tasks, or to differences in response requirement, reward exposure, or training duration. However, different reward contingencies may have been experienced in the FS5 task, which, by its nature, resulted in a larger proportion of unrewarded lever presses due to incomplete

sequences and extra-lever presses in complete sequences. Further research is needed to determine whether and to what extent the perceived reward contingency may have affected striatal activity.

Here, we hypothesize that the mean activity pattern of DMS neurons in the DT5 task, characterized by decreases in firing during lever presses followed by excitation at the termination of the sequence, could be understood as a neural correlate of action chunking. According to this hypothesis, DMS disengagement during lever pressing would gate behavioral chunking by allowing the DLS to take control. This hypothesis is supported by previous findings showing that DMS disengagement predicts skill learning[24] while DLS activity promotes habit and behavioral vigor.[61] Accordingly, it was suggested that DMS activity can modulate DLS access to the control of action during navigation in a T-maze.[32] Recently, DLS and DMS were shown to exert opposing role during the acquisition of skilled action sequences, with accelerated task acquisition following DMS loss of function.[16] Thus, distinct dorsostriatal activity patterns in the DT5 task could encode behavioral chunking through the inhibition of DMS neurons, reflecting the suspension of outcome expectation during lever presses, and the gated excitation of DLS neurons, reflecting behavioral vigor during sequence execution.

To conclude, it is tempting to speculate from the comparison of behavior and dorsostriatal activity across the three tasks that although rats are constrained to execute sequences of consecutive lever presses in the FS5 task, they only concatenate individual actions into unitary sequences in the DT5 task. Rapid behavioral chunking in the DT5 task (occurring within the first session)[35] and its presumed absence in the FR5 and FS5 tasks, could explain why striatal activity largely differs across tasks and does not appear to track habitual learning or improved task performance across training sessions in our conditions. This suggests that, although related, habitual learning and behavioral chunking develop along distinct timescales, with chunking preceding habit development, a hypothesis deserving further research.

Limitations of the study

All experiments in this study were conducted with male rats. Moving forward, it will be important to consider sex as a biological variable to determine whether the current findings generalize to female rats. In the current study, electrophysiological recordings allow investigating correlational patterns of DLS and DMS activity across three different tasks but do not demonstrate whether these patterns are causal for habit formation. Future experiments could for instance use optogenetic tools to directly address this question. Finally, due to their low proportion, we did not investigate the activity of interneurons, which could play a role in habitual learning and behavioral chunking (Martiros et al., 2018). Furthermore, the techniques used here do not differentiate between the two main populations of medium spiny neurons (D1 versus D2).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Experimental groups
  - Behavioral training
  - Outcome devaluation by sensory-specific satiety
  - Electrophysiological recordings
  - Analysis of electrophysiological recordings
  - Histology
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105818.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.V. and P.H.J.; methodology, Y.V., P.H.J.; formal analysis, Y.V.; investigation, Y.V.; writing – original draft, Y.V. and P.H.J.; writing – review & editing, Y.V. and P.H.J.; visualization, Y.V.; funding acquisition, P.H.J.; resources, P.H.J.; supervision, P.H.J.

DECLARATION OF INTERESTS

The authors declare no competing interests. The current institutional affiliation of YV is: Université de Poitiers, INSERM, U-1084, Laboratoire des Neurosciences Expérimentales et Cliniques, Poitiers, France.

INCLUSION AND DIVERSITY

We support the inclusive, diverse, and equitable conduct of research.

REFERENCES

1. Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., and Boakes, R.A. (1995). Motivational control after extended instrumental training. Anim. Learn. Behav. 23, 197–206. https://doi.org/10.3758/BF03199935.

2. Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. Phil. Trans. Biol. Sci. 308, 67–78. https://doi.org/10.1098/rstb.1985.0010.

3. Balleine, B.W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37, 407–419. https://doi.org/10.1016/S0028-3908(98)00033-1.

4. Smith, K.S., and Graybiel, A.M. (2016). Habit formation. Dialogues Clin. Neurosci. 18, 33–43.

5. Jin, X., and Costa, R.M. (2015). Shaping action sequences in basal ganglia circuits. Curr. Opin. Neurobiol. 33, 188–196.

6. Graybiel, A.M. (1998). The basal ganglia and chunking of action repertoires. Neurobiol. Learn. Mem. 70, 119–136.

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. https://doi.org/10.1126/science.286.5445.1745.

8. Dezfouli, A., and Balleine, B.W. (2013). Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. PLoS Comput. Biol. 9, e1003364.

9. Dezfouli, A., Lingawi, N.W., and Balleine, B.W. (2014). Habits as action sequences: hierarchical action control and changes in outcome value. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369, 20130482.

10. Du, Y., Krakauer, J.W., and Haith, A.M. (2022). The relationship between habits and motor skills in humans. Trends Cogn. Sci. 26, 371–387. https://doi.org/10.1016/j.tics.2022.02.002.

11. Haith, A.M., and Krakauer, J.W. (2018). The multiple effects of practice: skill, habit and reduced cognitive load. Curr. Opin. Behav. Sci. 20, 196–201. https://doi.org/10.1016/j.cobeha.2018.01.015.

12. Garr, E., and Delamater, A.R. (2019). Exploring the relationship between actions, habits, and automaticity in an action sequence task. Learn. Mem. 26, 128–132. https://doi.org/10.1101/lm.048645.118.

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

14. Robbins, T.W., and Costa, R.M. (2017). Habits. Curr. Biol. 27, R1200–R1206. https://doi.org/10.1016/j.cub.2017.09.060.

15. van Elzelingen, W., Warnaar, P., Matos, J., Bastet, W., Jonkman, R., Smulders, D., Goedhoop, J., Denys, D., Arbab, T., and Willuhn, I. (2022). Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. Curr. Biol. 32, 1163–1174.e6. https://doi.org/10.1016/j.cub.2021.12.027.

16. Turner, K.M., Svegborn, A., Langguth, M., McKenzie, C., and Robbins, T.W. (2022). Opposing roles of the dorsolateral and dorsomedial striatum in the acquisition of skilled action sequencing in rats. J. Neurosci. 42, 2039–2051. https://doi.org/10.1523/jneurosci.1907-21.2022.

17. Vandaele, Y., Pribut, H.J., and Janak, P.H. (2017). Lever insertion as a salient stimulus promoting insensitivity to outcome devaluation. Front. Integr. Neurosci. 11, 23. https://doi.org/10.3389/fnint.2017.00023.

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

19. Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. 166, 189–196. https://doi.org/10.1016/j.bbr.2005.07.012.

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

21. Balleine, B.W., Liljeholm, M., and Ostlund, S.B. (2009). The integrative function of the basal ganglia in instrumental conditioning. Behav. Brain Res. 199, 43–52. https://doi.org/10.1016/j.bbr.2008.10.034.

22. Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. 19, 181–189. https://doi.org/10.1111/j.1460-9568.2004.03095.x.

23. Smith, K., and Graybiel, A. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron 79,

608. https://doi.org/10.1016/j.neuron.2013.07.032.

24. 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.e5. https://doi.org/10.1016/j.neuron.2017.09.040.

25. Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci. 22, 513–523. https://doi.org/10.1111/j.1460-9568.2005.04218.x.

26. Corbit, L.H., and Janak, P.H. (2016). Habitual alcohol seeking: neural bases and possible relations to alcohol use disorders. Alcohol Clin. Exp. Res. 40, 1380–1389.

27. Corbit, L.H., and Janak, P.H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. Eur. J. Neurosci. 31, 1312–1321. https://doi.org/10.1111/j.1460-9568.2010.07153.x.

28. Corbit, L.H., Nie, H., and Janak, P.H. (2014). Habitual responding for alcohol depends upon both AMPA and D2 receptor signaling in the dorsolateral striatum. Front. Behav. Neurosci. 8, 301.

29. Corbit, L.H., Nie, H., and Janak, P.H. (2012). Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol. Psychiatry 72, 389–395.

30. Lerner, T.N. (2020). Interfacing behavioral and neural circuit models for habit formation. J. Neurosci. Res. 98, 1031–1045. https://doi.org/10.1002/jnr.24581.

31. Regier, P.S., Amemiya, S., and Redish, A.D. (2015). Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. J. Neurophysiol. 114, 1399–1416. https://doi.org/10.1152/jn.00189.2015.

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

33. 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, 1158–1161. https://doi.org/10.1038/nature04053.

34. Martiros, N., Burgess, A.A., and Graybiel, A.M. (2018). Inversely active striatal projection neurons and interneurons selectively delimit useful behavioral sequences. Curr. Biol. 28, 560–573.e5. https://doi.org/10.1016/j.cub.2018.01.031.

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

36. Schmitzer-Torbert, N.C., and Redish, A.D. (2008). Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. Neuroscience 153, 349–360. https://doi.org/10.1016/j.neuroscience.2008.01.081.

37. Stalnaker, T.A., Berg, B., Aujla, N., and Schoenbaum, G. (2016). Cholinergic interneurons use orbitofrontal input to track beliefs about current state. J. Neurosci. 36, 6242–6257. https://doi.org/10.1523/JNEUROSCI.0157-16.2016.

38. Coffey, K.R., Nader, M., and West, M.O. (2016). Single body parts are processed by individual neurons in the mouse dorsolateral striatum. Brain Res. 1636, 200–207. https://doi.org/10.1016/j.brainres.2016.01.031.

39. Robbe, D. (2018). To move or to sense? Incorporating somatosensory representation into striatal functions. Curr. Opin. Neurobiol. 52, 123–130. https://doi.org/10.1016/j.conb.2018.04.009.

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

41. Thrailkill, E.A., Trask, S., Vidal, P., Alcalá, J.A., and Bouton, M.E. (2018). Stimulus control of actions and habits: a role for reinforcer predictability and attention in the development of habitual behavior. J. Exp. Psychol. Anim. Learn. Cogn. 44, 370–384. https://doi.org/10.1037/xan0000188.

42. Vandaele, Y., and Ahmed, S.H. (2021). Habit, choice, and addiction. Neuropsychopharmacology 46, 689–698. https://doi.org/10.1038/s41386-020-00899-y.

43. Vandaele, Y., Guillem, K., and Ahmed, S.H. (2020). Habitual preference for the nondrug reward in a drug choice setting. Front. Behav. Neurosci. 14, 78. https://doi.org/10.3389/fnbeh.2020.00078.

44. Vandaele, Y., Vouillac-Mendoza, C., and Ahmed, S.H. (2019). Inflexible habitual decision-making during choice between cocaine and a nondrug alternative. Transl. Psychiatry 9, 109. https://doi.org/10.1038/s41398-019-0445-2.

45. Thrailkill, E.A., Michaud, N.L., and Bouton, M.E. (2021). Reinforcer predictability and stimulus salience promote discriminated habit learning. J. Exp. Psychol. Anim. Learn. Cogn. 47, 183–199. https://doi.org/10.1037/xan0000285.

46. Bouton, M.E. (2021). Context, attention, and the switch between habit and goal-direction in behavior. Learn. Behav. 49, 349–362. https://doi.org/10.3758/s13420-021-00488-z.

47. Dickinson, A., Nicholas, D.J., and Adams, C.D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. Q. J. Exp. Psychol. B 35, 35–51. https://doi.org/10.1080/14640748308400912.

48. Perez, O.D., and Dickinson, A. (2020). A theory of actions and habits: the interaction of rate correlation and contiguity systems in free-operant behavior. Psychol. Rev. 127, 945–971. https://doi.org/10.1037/rev0000201.

49. de Russo, A.L., Fan, D., Gupta, J., Shelest, O., Costa, R.M., and Yin, H.H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. Front. Integr. Neurosci. 4, 17. https://doi.org/10.3389/fnint.2010.00017.

50. Garr, E., Bushra, B., Tu, N., and Delamater, A.R. (2020). Goal-directed control on interval schedules does not depend on the action-outcome correlation. J. Exp. Psychol. Anim. Learn. Cogn. 46, 47–64. https://doi.org/10.1037/xan0000229.

51. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8, 1704–1711. https://doi.org/10.1038/nn1560.

52. Lee, S.W., Shimojo, S., and O'Doherty, J.P. (2014). Neural computations underlying arbitration between model-based and model-free learning. Neuron 81, 687–699. https://doi.org/10.1016/j.neuron.2013.11.028.

53. Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. PLoS Comput. Biol. 7, e1002055. https://doi.org/10.1371/journal.pcbi.1002055.

54. Vandaele, Y., Ottenheimer, D.J., and Janak, P.H. (2021). Dorsomedial striatal activity tracks completion of behavioral sequences. eNeuro 8. ENEURO.0279,21.2021. https://doi.org/10.1101/2021.04.01.437899.

55. Kosaki, Y., and Dickinson, A. (2010). Choice and contingency in the development of behavioral autonomy during instrumental conditioning. J. Exp. Psychol. Anim. Behav. Process. 36, 334–342. https://doi.org/10.1037/a0016887.

56. Trask, S., Shipman, M.L., Green, J.T., and Bouton, M.E. (2020). Some factors that restore goal-direction to a habitual behavior. Neurobiol. Learn. Mem. 169, 107161. https://doi.org/10.1016/j.nlm.2020.107161.

57. Bouton, M.E., Broomer, M.C., Rey, C.N., and Thrailkill, E.A. (2020). Unexpected food outcomes can return a habit to goal-directed action. Neurobiol. Learn. Mem. 169, 107163. https://doi.org/10.1016/j.nlm.2020.107163.

58. 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 article No discrete start/stop signals in the dorsal striatum of mice performing a learned action. Curr. Biol. 28, 3044–3055.e5. https://doi.org/10.1016/j.cub.2018.07.038.

59. Rueda-Orozco, P.E., and Robbe, D. (2015). The striatum multiplexes contextual and kinematic information to constrain motor habits execution. Nat. Neurosci. 18, 453–460. https://doi.org/10.1038/nn.3924.

60. Dudman, J.T., and Krakauer, J.W. (2016). The basal ganglia: from motor commands to the control of vigor. Curr. Opin. Neurobiol. *37*, 158–166. https://doi.org/10.1016/j.conb.2016.02.005.

61. Crego, A.C.G., Štoček, F., Marchuk, A.G., Carmichael, J.E., van der Meer, M.A.A., and Smith, K.S. (2020). Complementary control over habits and behavioral vigor by phasic activity in the dorsolateral striatum. J. Neurosci. *40*, 2139–2153. https://doi.org/10.1523/JNEUROSCI.1313-19.2019.

62. Ottenheimer, D., Richard, J.M., and Janak, P.H. (2018). Ventral pallidum encodes relative reward value earlier and more robustly than nucleus accumbens. Nat. Commun. 9, 4350. https://doi.org/10.1038/s41467-018-06849-z.

63. Richard, J.M., Stout, N., Acs, D., and Janak, P.H. (2018). Ventral pallidal encoding of reward-seeking behavior depends on the underlying associative structure. Elife *7*, e33107. https://doi.org/10.7554/eLife.33107.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Electrophysiological datasets and Matlab code used for analysis and figures	GIN repository	https://doi.org/10.12751/g-node.p53gw9
Experimental models: Organisms/strains		
Long Evans rats	Envigo	HsdBlu:LE
Software and algorithms		
MATLAB	Mathworks	N/A
Statistica	StatSoft 7.0	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources, analysis and methodology should be directed to and will be fulfilled by the lead contact, Youna Vandaele (youna.vandaele@univ-poitiers.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Original electrophysiological data and Matlab code used for analysis and figures have been deposited at G-node and are publicly available as of the date of publication. DOI is listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

18 male Long Evans rats (Envigo) were individually housed in a temperature (21°C) and light-controlled (12-h light-dark cycle, lights ON at 7 a.m.) vivarium, with partial enrichment. All experiments were performed during the light cycle. Rats were given free access to water throughout the experiment and were maintained at 90% of their free-feeding weight. This study was carried out in accordance with the recommendations of the Guide for the Care and Use of Laboratory Animals (Institute of Instrumental Training Laboratory Animal Resources, Commission of Life Sciences, National Research Council, 1996). The protocol was approved by the institutional animal care and use committee of Johns Hopkins University.

METHOD DETAILS

Experimental groups

Following instrumental training under continuous reinforcement schedules, rats in the 3 different groups received surgery and neurons in dorsal striatum were recorded throughout acquisition of the DT5 task (N = 6), the FR5 task (N = 6) or the FS5 task (N = 6). Sensitivity to satiety-induced devaluation was assessed at the end of recording in all the rats, tethered in the recording chamber.

Behavioral training

Initial instrumental training

Rats were trained with a small aliquot of a solution of 20% sucrose (0.1 mL delivered over 3s). In each task, the house-light, located on the ceiling of the operant chamber remained illuminated during sessions. After a single 30-min magazine training session under a variable interval 60s schedule, rats were trained to press the left lever to earn the reward, delivered in the adjacent magazine. Sessions were limited to 1 h or 30

reward deliveries. Rats were trained for 3 to 5 sessions, until they earned the 30 rewards in less than 1 h. Conditioning chambers housed within sound-attenuating boxes (Med Associates, St Albans, VT) and designed for *in vivo* neural recording were used for all training and testing.

DT5 task training

Each trial (max: 30) of the session begin with the insertion of the left lever and completion of the lever press requirement results in the retraction of the lever, the reward delivery, and the initiation of a new inter-trial interval. The response requirement was 1 for 3 sessions (discrete-trial FR1; DT1) and increased to 5 for 10 sessions (discrete-trial FR5; DT5). Failure to complete the ratio within 1 min was considered as an omission and resulted in lever retraction and initiation of a new inter-trial interval. Data from this group of rats were presented in a previous study aimed at comparing DMS and DLS activity across early and extended training in the DT5 task.[35] Here we only selected rats from the early training group trained with liquid 20% sucrose.

FR5 task training

In this free-operant task, the lever was continuously presented and 5 presses on the lever resulted in a 3-s delivery of the liquid sucrose reward (Fixed ratio 5, FR5). Rats were trained for ten FR5 sessions, limited to 30 min or 30 reward deliveries.

FS5 task training

In this free-operant task, the lever was continuously presented. Rats had to complete a sequence of at least 5 consecutive lever presses without checking the port to obtain a reward, whose delivery was not signaled. A port entry before completion of the ratio resulted in resetting the ratio and the lever press sequence was considered as incomplete. However, additional presses after completion of the ratio were without consequences and considered as part of the completed sequence. Rats were first trained with a fixed sequence length of 2 lever presses for a minimum of 3 sessions or until they earned 30 rewards. The response requirement was then increased to 3 lever presses for a minimum of 2 sessions (or until earning 30 rewards) before training in the final fixed sequence length 5 schedule (FS5) for 7 sessions. Sessions were limited to 30 min or 30 reward deliveries.

Outcome devaluation by sensory-specific satiety

Each rat received 2 days of testing, separated by one reinforced training session. Rats were given 1 h free access to their training reward (sucrose 20%; devalued condition) or to a control reward, which never served as a reinforcer (grain-based pellet; valued condition). Pre-feeding occurred in feeding cages in the experimental room. Immediately after pre-feeding, rats were placed in the recording chambers for a test session conducted under extinction. Test sessions were limited to 10 trials in the DT5 task and 10 min in the FR5 and FS5 tasks. On the second test session, animals were pre-fed with the alternative reward prior to the extinction test.

Electrophysiological recordings

Surgeries and recording

After acquisition of instrumental responding, rats underwent surgeries. In the DT5 group, rats were implanted with 2 unilateral arrays of 8 wires aimed at DMS and DLS (0.004′ steel wires arranged in a 2 × 4 configuration, each array spaced 2 mm apart, Microprobes). In the FR5 and FS5 groups, 16 tungsten wires were soldered on two connectors and arranged in 2 bundles spaced 2 mm apart. For each group, target coordinates were +0.25 mm AP, ±2.3 mm ML, −4.6 mm DV for DMS and +0.25 mm AP, ±4.3 mm ML, −4.6 mm DV for DLS. Surgeries were performed under isoflurane anesthesia (0.5-5%) with pre-operative injections of cefazolin (75 mg/kg, antibiotic) and carprofen (5 mg/kg, analgesic). Topical lidocaine was applied for local analgesia.

After a minimum of 5 days post-operative recovery, rats were accustomed to tethering with the recording cable for a few CRF sessions before the recording began. Cables were connected at one end to rats' headsets and at the other end to a commutator allowing free movement throughout acquisition of single-unit activity during the recording sessions. The multichannel acquisition processor (MAP) neural recording system (Plexon Inc, TX) was used to store and process amplified signals and timestamps of behavioral events.

Analysis of electrophysiological recordings

Spike sorting

As previously described,[62,63] individual units were isolated offline using Offline Sorter (Plexon Inc, TX). Analysis of interspike intervals distribution, auto-correlograms and cross-correlograms (Offline Sorter v3 and Neuro-Explorer 3.0, Plexon Inc, TX) was conducted to control for correct isolation of single units. Average waveforms and timestamps of units and event were exported from Neuro-Explorer 3.0 to MATLAB (MathWorks, MA) for further analysis. Analyses were restricted to units with well-defined waveforms and consistent characteristics throughout the entire recording session.

Waveform analysis

Neurons were classified as putative Medium Spiny Neurons (MSN), Fast Spiking Interneurons (FSI) or Tonically Active Neurons (TAN), based on the half-valley width of the average waveform and the overall firing rate (Figures S1A–S1E). Putative-FSI were defined by high firing rate (>20 Hz) and narrow waveforms (<0.15 ms) whereas putative-TAN were defined by low firing rate (<5 Hz) and wide waveform (>0.45 ms) as previously reported.[34,36,37] Neurons not classified as interneurons but showing features intermediate to MSNs and interneurons were unclassified (firing rate between 12.5 and 20 Hz and half-width between 0.4 and 0.45 ms). All the analyses in this study were conducted on the population of putative-MSNs.

Characterization of task-responsive neurons and z-scores

In each task, we analyzed spiking activity across twelve 0.25s periods before and after each lever press and port entry events. Since the performance in the FS5 task was characterized by a large proportion of sequences longer than 5 lever presses, the lever press events considered for the analysis were defined as follow: the first and second lever presses, one randomly selected intermediate lever press, the second to last lever press and the last lever press.

For each neuron, the presence of significant excitation or inhibition in response to an event was detected by running a t-test on the firing rate during the 0.25s periods pre- and post-event in comparison to a 1-s baseline period beginning 1-s before trial onset (defined by the lever insertion in the DT5 task and by the first lever press in the FR5 and FS5 tasks). Neurons expressing a significant response (p < 0.01) to at least one of the 12 behavioral events were considered as task responsive neurons.

Neural activity for each individual neuron was z-scored as follow: $(F_i - F_{mean})/F_{sd}$, where F_{sd} and F_{mean} represent the SD and mean of the firing rate during the 1-s baseline period, and F_i is the firing rate at the ith bin of the peristimulus time histogram (PSTH). Heatmaps and average PSTH presented in this study represent the z-scores from −0.25s to 0.25s around each event of the behavioral sequence.

In Figure 5, to control for local differences in response rate between the DT5 and FS5 tasks, we sub-selected trials based on inter-press intervals. Specifically, trials at the beginning and end of the sequence were selected if the inter-press interval following the first lever press or preceding the last lever press was shorter than 1s. Trials for mid-sequence activity were selected if the inter-press interval between the presses preceding and following the intermediate lever press was shorter than 2-s. Only recording sessions comprising at least 5 trials for a given time in the sequence were included in this analysis.

In Figure 6, we matched trials from the DT5 dataset to the middle tertile of response rates in the FS5 dataset. Specifically, for each FS5 sessions, we determined the 33rd and 66th percentile of within-sequence response rates across subjects and trials. On each DT5 recording session and until the seventh session, we restricted the analysis of neuronal spiking activity to trials with response rates between the 33rd and 66th percentile of the FS5 response rates distribution. We only considered DT5 sessions comprising at least 5 trials in this response rate range.

Histology

Electrode sites were labeled by passing a DC current through each electrode, under deep anesthesia with pentobarbital. All rats were perfused intracardially with 0.9% saline followed by 4% paraformaldehyde (FR5 and FS5 groups with tungsten wires), or 4% paraformaldehyde with 3% potassium ferricyanide (DT5 group with steel wires). Brains were extracted, post-fixed in 4% paraformaldehyde for 4–24 h, and transferred in

20% sucrose for >48 h for cryo-protection. To verify the electrodes placement, the brains were sectioned at 50 μm on a cryostat and slices were stained with cresyl violet and analyzed using light microscopy.

QUANTIFICATION AND STATISTICAL ANALYSIS

Within-sequence response rate (in resp/s) was computed by dividing the number of lever presses per sequence by the time from the first to the last lever press. Local response rates at the beginning and end of lever press sequences were computed by dividing number of lever presses included in the analysis (i.e. 2) by the inter-press interval following the first lever press or preceding the last lever press, respectively. Mid-sequence response rates were estimated by dividing the number of lever presses included in the analysis (i.e. 3) by the interval between the lever presses preceding and following the intermediate lever press. Port entry latency was defined by the time separating completion of the 5 lever press requirement and the port entry following reward delivery. The devaluation ratio was defined by the ratio of the number of lever presses in the devalued condition divided by the total number of lever presses in the valued and devalued conditions. Thus, a devaluation ratio of 0.5 indicates habitual responding. In the FS5 task, only complete sequences were considered in behavioral and electrophysiological analyses.

Data following a normal distribution were subjected to repeated measures ANOVA. The Mann Whitney test was used when normality assumptions were violated. Mean z-scores of units from DMS and DLS were compared across 12 consecutives events using repeated-measures ANOVA(with Geisser-Greenhouse correction for violation of sphericity), with events as within-design factor and regions, sessions and, when relevant, task, as between-design factors. Significant interactions were analyzed using the HSD Tukey post-hoc test. Events consisted of the average Z score during 0.25s periods before and after each lever press and port entry events. All analyses were conducted on MATLAB (MathWorks) and Statistica (StatSoft 7.0). Statistical details (tests used, value and nature of n,..etc.) are described in the results section. The definition of center and dispersion measures is described in figure legends. Significance was defined as a p value<0.05.