Title: Concomitant processing of choice and outcome in frontal corticostriatal ensembles correlates with performance of rats

Authors: Takashi Handa^{1,3*, 5}, Rie Harukuni³, Tomoki Fukai^{2,3 *, 4}

Affiliations:

1, Department of Behavior and Brain Organization, Center Advanced European Study and Research (caesar), Bonn, 53175 Germany.

2, Neural Coding and Brain Computing Unit, Okinawa Institute of Science and Technology, Tancha 1919-1, Onna-son, Okinawa 904-0495, Japan.

3, Laboratory for Neural Coding and Brain Computing, RIKEN Center for Brain Science, Hirosawa 2-1, Wako, Saitama 351-0198, Japan.

4, Lead contact

5, present address: Department of Neurobiology, Graduate School of Biomedical and Health Sciences (Medicine), Hiroshima University, Hiroshima, 734-8553, Japan.

***Correspondence :**

email: handanb@hiroshima-u.ac.jp (T.H.), tomoki.fukai@oist.jp (T.F.)

19 **Summary**

20 The frontal cortex-basal ganglia network plays a pivotal role in adaptive goal-directed behaviors.

21 Medial frontal cortex (MFC) encodes information about choices and outcomes into sequential

22 activation of neural population, or neural trajectory. While MFC projects to the dorsal striatum (DS),

23 whether DS also displays temporally coordinated activity remains unknown. We studied this

24 question by simultaneously recording neural ensembles in the MFC and DS of rodents performing

25 an outcome-based alternative choice task. We found that the two regions exhibited highly parallel

26 evolution of neural trajectories, transforming choice information into outcome-related information.

27 When the two trajectories were highly correlated, spike synchrony was task-dependently modulated

28 in some MFC-DS neuron pairs. Our results suggest that neural trajectories concomitantly process

29 decision-relevant information in MFC and DS with increased spike synchrony between these regions.

30

31 **Keywords:** win-stay lose-shift; frontal corticostriatal ensemble; multi-neuron recordings;

32 population coding; Linear Fisher's discriminant

33

34 **Introduction**

35 Deciding actions based on the outcome of past actions is crucial for the survival of animals. The
36 frontal cortex-basal ganglia network is engaged in various forms of decision making involved in
37 adaptive goal-directed behaviors [1-4]. These studies have shown how choice- and/or outcome-
38 related information is processed in the frontal cortex and the basal ganglia. However, how the
39 cortico-basal ganglia network encodes behavioral information at the neural ensemble level remains
40 largely unknown. How the related brain regions communicate such information during outcome-
41 based action selection is yet to be clarified [5]. In this study, we explore whether and how task-
42 related neural activity is coordinated between the medial frontal cortex (MFC) and dorsal striatum
43 (DS), which is a major input structure of the basal ganglia [6-8], of rats performing an outcome-based
44 alternative choice task.

45

46 The MFC is engaged in encoding reward and error signals [9,10] and hence is considered to
47 play a role in monitoring positive and negative outcomes from an action. In particular, the rostral
48 agranular medial cortex or the secondary motor cortex of rodent has been implicated in sensory-
49 cued [11-13] and outcome-based action selection [14,15]. Along the frontal cortico-basal ganglia
50 axis, excitatory MFC outputs project monosynaptically to DS [16,17], which also receives
51 glutamatergic inputs from some thalamic nuclei [7,18] and are modulated by midbrain dopaminergic
52 inputs conveying outcome information [19] and reward prediction error [20]. Accordingly, DS is

thought to associate a specific action with the resultant outcome [21-24]. Consistent with this, lesions in MFC [14,15] and DS [25,26] impair outcome-based choice behaviors.

Evidence from MFC, or more specifically posterior parietal cortex (PPC) and orbitofrontal cortex (OFC), shows that sequential activation of cortical neuron ensembles, which is termed neural trajectory, underlies decision making [12,27,28]. We explore how this trajectory information is conveyed to striatal neurons during decision making. If the role of the striatum is selection of cortical inputs for action generation, as conventionally thought, such selection may occur sparsely in time, making continuously evolving trajectories less relevant to striatal functions. We therefore clarify to what extent cortical and striatal neuron ensembles are coherently activated during decision making. We hypothesize that neural ensembles in the MFC and DS utilize synchronized spikes to communicate with one another during adaptive behavior. Such spikes are ubiquitous in cortical circuits [29,30] and were suggested to enhance information transmission and association between directly or indirectly interconnected brain regions [31-32].

We trained head-restrained rats to lick either left or right spouts for earning reward (positive outcome) and simultaneously recorded multi-neuron activity from the MFC and DS of these rats. Because a rewarded spout was changed intermittently, the rats had to switch their choice responses to maximize reward, exhibiting behavioral responses similar to a typical win-stay lose-away behavior.

Unexpectedly, we found that neural trajectories emerge and evolve highly coherently in both MFC and DS during the entire period of the alternative choice task. In some MFC-DS neuron pairs, the number of coincident spikes was modulated by task events when the trajectory evolution displayed an enhanced coherence. Our results suggest that the adaptive control of outcome-based decision making relies on the coevolution of neural trajectories in MFC and DS and that spike synchrony plays a role in this temporal coordination.

Results

Adaptive outcome-based behavior under head-restrained condition

Twenty head-restrained rats performed an outcome-based two choice task, in which the rats were required to lick one of two spouts (Left and Right) and in each block of trials one spout delivered reward, with a high reward acquisition probability (>75%, see Methods) (Fig.1a). The choice-reward contingency was systematically reversed without any sensory feedback after the cumulative number of rewarded trials reached 10 in each block. The rats figured out the reversal of choice-reward contingency through the monitoring of no-reward events, time-out, and reward acquisition within subsequent several trials (Fig. 1b). Rats tended to select the same spout as that chosen in the last rewarded trial, but they switched the choice following one to several unrewarded trials (Fig. 1b, c). This choice pattern is close to the so-called ‘win-stay’ and ‘lose-shift’ which is the optimal strategy

Figure 1

Figure 1. Adaptive choice patterns in response to change in choice-reward contingency.

(a) *Left and middle*: Snapshots of a head-restrained rat making lick responses toward Left (orange arrowhead) and Right (blue arrowhead) spouts. A circle indicates the tongue position. *Right*: A schematic illustration of an outcome-based two alternative choice task. Individual trial started with 'Start' tone (3 kHz) presentation. Head-fixed rats were required to wait for 'Go' tone (10 kHz) without a lick during '1st Delay' period, and then to make a choice response by licking either one of spouts within 5 s after 'Go' tone onset. Reward was given after '2nd Delay' of various durations (0.3-0.7 s) if the chosen position corresponded to current reward location. Otherwise no reward, no sensory feedback, and timeout of 5 s were given. (b) Representative flexible choice behavior by a rat across 9 block-reversals of reward position. Vertical lines and colored areas denote the block-reversal and reward position in each block, respectively. Circle and cross indicate the outcome, reward and no-reward respectively, after choice response (Left or Right). Square shows choice pattern (Switch or Repeat). (c) Probability of repeating choice after rewarded and unrewarded outcomes (30 sessions, 20 rats). Circle indicates individual sessions. Horizontal lines and error bars represent mean and SD, respectively. The statistical significance was confirmed by Wilcoxon signed-rank test. Inset: Probability of repeating choice following one or two consecutive unrewarded trials. (d) Effects of past choice (purple) and outcome (green) on (*left*) current choice position and (*right*) current choice type (repeat and switch). The ordinate shows averaged coefficients of a logistic regression model (N = 30 sessions). Error bars show SD. (e) Licking patterns during rewarded and unrewarded trials in the session shown in b. Rastergrams denote licking times and locations (Left: orange, Right: blue). Histograms show averaged Left and Right licking rates in rewarded (solid line) and unrewarded (line with cross) trials. (f) Median licking rate in an epoch between Go tone onset and reward delivery (Go-Rwd), during outcome period of 4-s after reward delivery (PostRwd), and for 4-s after response in rewarded (PostResp in black) and unrewarded (PostResp in red) trials. Asterisks indicate statistical significance (Wilcoxon signed-rank test, $p < 0.05$). (g) Comparison of reaction times among different choice patterns. Circles denote median reaction times in individual trials. Horizontal and vertical lines indicate median and 1st-3rd quartiles across trials, respectively. Trials were classified by last and current outcomes at each session (e.g., U-R denotes unrewarded last and rewarded current choices, and R+1, R+2, ... denote the number of continuous rewarded outcomes).

91 to maximize reward in the present task. To examine if the choice pattern (repeating post rewarded
92 trial and switching post no-reward trial) was biased by past outcome (reward or no-reward) or by
93 past choice position (Left and Right), we estimated the effect of past outcomes and past choices by
94 a logistic regression analysis (Methods). The last outcome provided the greatest effect on the current
95 choice type (repeat or switch) regardless of past choice position (t-test with Bonferroni correction,
96 $p = 3.73 \times 10^{-15}$). In contrast, the current choice position was the most affected by the last choice
97 position regardless of outcome (t-test with Bonferroni correction, $p = 5.17 \times 10^{-6}$) (Fig. 1d).

98 Trial-by-trial licking behavior was strongly influenced by outcome events. Rats licked either
99 one of the two spouts immediately after the appearance of Go tone, and then showed fast and
100 rhythmic licking at around 6 Hz just after reward delivery (Fig. 1e). In contrast, rats showed sparse
101 licking at around 1 Hz during outcome period in unrewarded trials. The lick rate of population data
102 was significantly different before and after reward delivery in rewarded trials (Fig. 1f, Wilcoxon
103 signed-rank test, $p = 0.00295$) and between rewarded and unrewarded outcome periods (Wilcoxon
104 signed-rank test, $p = 1.50 \times 10^{-6}$). In behavioral studies on primates and rodents, prolonged response
105 times, called post-error slowing, are observed after no reward outcome in sensory-cued choice tasks
106 [33,34]. At 13 sessions, we observed significantly differential reaction times among different choice
107 patterns classified by recent outcome history (Kruskal-Wallis test, $p < 0.05$). Although the reaction
108 times were indeed longer after no or first reward outcome than after consecutive reward outcomes

in some sessions, we did not find any tendency toward post-error slowing at the population level ($p = 0.989$, Kruskal-Wallis test, $n = 30$ sessions) (Fig. 1g).

Simultaneous multi-neuron recordings from MFC and DS

To test whether and how neuronal population activities of the MFC-DS circuit are collectively coordinated during the outcome-based action selection, we recorded multi-neuron activity in both MFC and DS using two silicon probes during task performance. In addition, we simultaneously performed juxtacellular recordings of Neurobiotin-loaded neurons near the silicon probe for the MFC recording (Fig. 2a). We confirmed the projection from the recording site in the MFC to that in the DS by using a retrograde tracer Fluoro-Gold (FG) injected into the DS (Methods). FG-labelled corticostriatal neurons were primarily observed in certain restricted regions of the MFC (Fig. 2b and c), particularly in the rostral agranular medial cortex as previously reported [6,7]. Observing the track of the silicon probe insertion, Neurobiotin-loaded neurons (in some cases), FG-labelled neurons in layers 3 and 5 of MFC (Fig. 2b-c and Fig. S1) and the tracks of another probe for the DS recording at or near FG injection site in DS (Fig. 2b), we confirmed that multi-neuron activity was presumably recorded from a subset of MFC-DS circuit.

Single MFC and DS cells encode both positive and negative outcomes

Figure 2

Figure 2. Simultaneous recording from MFC and DS. (a) Left: A schematic illustration of recordings. One silicon probe (black) and a glass pipette electrode (gray) were inserted into MFC and another silicon probe into DS (red). Right: Multiunit activity recorded with tetrodes in MFC and DS and juxtacellularly recorded activity (gray). Spikes of isolated units are depicted below multiunit signals. (b) Left: Fluorescent images show corticostriatal projection neurons labelled with Fluoro-Gold (FG) in MFC (top) and the injection site (white arrowhead) of FG in DS (bottom). Right: Nissl stained brain sections specified by the AP coordinates from Bregma. One shown at bottom left is identical to the lower fluorescent image and one at bottom right indicates silicon probe tracks (scars, green arrowhead) in DS. The area surrounded by a dash-lined box was magnified (top). Scale bars show 1 mm. (c) Left: Juxtacellularly recorded activity of a neuron during task performance (top) and current injection (bottom). Horizontal bars indicate epochs of periodic current injection whereas long vertical lines are artifacts by the on/off cycles of injection. Middle: A Nissl stained brain section including probe track (scar, green arrowhead) in MFC (agranular medial cortex). Right: Fluorescent images reveal FG labelled corticostriatal neurons and a Neurobiotin (Nb)-labeled neuron in MFC. The rightmost image enlarges the inside of the box area.

Figure S1

Figure S1. Juxtacellular recording during task performance and labelling the recorded neuron. Top left: Juxtacellular activity (gray) during task performance. Black, orange, and blue lines indicate event signal, left licking signal, and right licking signal, respectively. Top right: Post hoc visualization of the recorded neuron and corticostriatal projection neurons. Bottom left: Auto-correlogram and spike waveform of the identified neuron. Bottom right: Firing patterns of the labelled neuron aligned to response execution.

127 We analyzed event-related firing patterns of well-isolated single neurons recorded simultaneously
128 from MFC ($n = 468$, mean $\pm$ SD = 39 ± 18 cells per session) and DS ($n = 489$, mean $\pm$ SD = 40 ± 14 cells
129 per session) in 10 rats. The recordings were performed in the left hemisphere while these rats were
130 performing the learned task through 12 recording sessions (Methods). The firing rates of both MFC
131 and DS cells were correlated with choices (Left vs. Right) and outcomes (Reward vs. No-reward) (two-
132 way ANOVA, $p < 0.05$, Methods).

133

134 Many outcome-modulated neurons decreased firing rate during the period of positive
135 outcome but not during that of negative outcome (Fig. 3a, b). However, some neurons increased
136 firing rate most strongly for negative outcomes (Fig. 3c), and others raised firing rate for positive
137 outcomes (Fig. 3d). We then calculated peri-event time histograms (PETHs), which revealed similar
138 sequential dynamics of MFC and DS cells (Fig. 3e and Fig. S2a). The PETHs also indicated a prominent
139 decrease in firing rate in the majority of MFC and DS cells after reward delivery. The fractions of
140 choice- and outcome-modulated neurons evolved similarly in MFC and DS, but these fractions
141 differed significantly between MFC and DS in large portions of task period (Fig. 3f). We quantified
142 the degree and preference of selectivity across different sessions and neurons by performing
143 receiver operating characteristic (ROC) analysis of neurons' firing rates and calculating the area
144 under ROC curve (AROC) (Methods). In both MFC and DS, choice-selective neurons exhibited a bias
145 towards right (i.e., contralateral) choice responses whereas outcome-selective neurons towards

Figure 3

Figure 3. Choice and outcome coding in single neurons of MFC and DS. Below, neuronal activities are aligned to choice response (vertical line). (a-d) Representative firing patterns modulated by choices and outcomes on MFC and DS single neurons. Simultaneously recorded Right-choice selective responses of (a) MFC and (b) DS cells. These cells decreased firing rate immediately after reward delivery (cyan circle). Another DS cell increased firing rate during post-outcome periods in (c) unrewarded and (d) rewarded trials. Top row shows spike waveforms (mean $\pm$ SD) in four channels of a tetrode and auto-correlogram. Peri-event time histogram (PETH) shows mean $\pm$ SEM of firing rate (bin = 200-ms, sliding 20 ms), and horizontal bars indicate the times at which firing rate was significantly different (two-way ANOVA, $p < 0.05$) with respect to choice (purple) or outcome (green) factor. (e) Normalized PETHs of MFC ($n = 468$) and DS ($n = 489$) neurons are shown in rewarded and unrewarded trials. In each condition, units are sorted by the latency of maximum absolute peaks. (f) *Top*: Proportion of MFC (black) and DS (red) cells showing (*left*) choice or (*right*) outcome-selective activity (two-way ANOVA, $p < 0.05$). Blue bars represent the times at which the cell proportions were significantly different between MFC and DS (χ^2 test, $p < 0.05$). *Middle and bottom*: The number of choice- and outcome-selective neurons are shown together with AROC values revealing preference between contralateral/ipsilateral choices or unrewarded/rewarded outcomes, respectively. (g) Comparison of selectivity index for choice (*left*) and outcome (*right*) between MFC (black) and DS (red). Thick solid lines and shaded area represent median and 1st/3rd quartiles, respectively. Blue bars represent the times at which the selectivity indices were significantly different between MFC and DS (Mann-Whitney U test, $p < 0.01$).

Figure S2

unrewarded outcomes (Fig. 3f). At the group level, selectivity indices were significantly larger in MFC than in DS for both choice- and outcome-selective neurons in large portions of task period (Fig. 3g, Mann-Whitney U test, $p < 0.01$). A small number of MFC and DS cells dynamically changed their firing patterns according to the histories of outcome and choice (Supplementary information and Fig. S2b).

Thus, choice- and outcome-modulated neurons show similar dynamical evolution in MFC and DS, but their fractions and degrees of selectivity were significantly different between the two brain regions. These results suggest parallelism between MFC and DS in the neural population coding of task-relevant information in this outcome-based choice task.

Functional correlation of neural trajectories between MFC and DS

The above similarity in neural dynamics raises a question about the way neural population activities are coordinated in MFC and DS to encode information about decision making. To clarify this, we explored how the neural trajectories encoding Left and Right decisions were separated in each session (12 sessions in total) by using Fisher's linear discriminant (Methods). This method yields a hyperplane that optimally separates datasets belonging to two categories, such as Left/Right choices or positive/negative outcomes, in the high dimensional space of the instantaneous population rate vectors [12,13,35]. We calculated discriminability index d' , which is the distance between the categorized rate vectors measured orthogonally to such a hyperplane. To keep a sufficient number

165 of cells ($n \geq 25$ in both regions per session, Supplementary information), below we omitted three
166 data sets (R986-r1, R1000-r1, R1004-r1 in Fig. S3a) and only used 9 datasets (R982-r1, R983-r1, R983-
167 r2, R985-r1, R985-r2, R991-r1, R1005-r1, R1009-r1, R1012-r1).

168

169 For choice-selective or outcome-selective trajectories, the index d' peaked around the time of
170 choice response or outcome, respectively (Figs. 4a and S3c). However, the peak times were not
171 significantly different between MFC and DS (Fig. 4a). Then, we studied how neural trajectories
172 evolved in each session by projecting the instantaneous rate vectors onto the axes orthogonal to
173 choice (Ch) and outcome (Ot) hyperplanes obtained by the linear discriminant analysis (Fig. 4b).
174 Hereafter, these trajectories are termed “Ch-projected” and “Ot-projected” trajectories. The
175 temporal patterns of these trajectories varied from session to session in both MFC and DS (Fig. S4).
176 However, their trial-averages exhibited transiently or gradually correlated increases between MFC
177 and DS around choice ($r = 0.62$) and outcome ($r = 0.50$) events, respectively (Fig. 4b and 4e). The
178 inter-trajectory correlations were particularly strong (Pearson correlation coefficient was about 0.5
179 or greater) in 4 sessions (R983-r2 in Fig. 4, R983-r1, R985-r1, and R985-r2 in Fig. S4). The neural
180 trajectories diverged along with choice responses (Left and Right) around the response time (t_1 : 0.2
181 s before choice response), and afterwards diverged along with outcomes (reward and no-reward)
182 during outcome period (t_2 : 1 s after choice response), as exemplified in Fig. 4c. Figure 4d shows the
183 trial-by-trial values of Ch-projected and Ot-projected trajectories in MFC and DS at times t_1 and t_2 ,

Figure S3

a

b

c

Figure S3. Degree of discrimination by population activity at individual sessions. (a) Time evolution of discrimination degree (d') on choice (purple) and outcome (green) axes around choice responses. Vertical line indicates the peak time of d' and cyan bar the interval of reward delivery. (b) Correlations between peak d' values and the number of neurons in population data. (c) Group data of d' on choice and outcome axes in MFC (black) and DS (red). Data were aligned at start cue end, go cue end, and response execution.

Figure 4

Figure 4. Temporal correlations of neural trajectories between MFC and DS. Unless otherwise stated, all results are shown for a typical session (R983-r2). (a) *Left*: Time series of discrimination degree (d') for neural populations simultaneously recorded in MFC ($n = 58$, *Top*) and DS ($n = 48$, *Bottom*). Vertical lines indicate the peak times of d' measured between different choices (purple) and outcomes (green). *Right*: Comparison of the peak times of d' between MFC and DS ($N = 9$ sessions) at Ch- and Ot- hyperplanes. Squares and error bars denote mean and SD. P-value was assessed by paired t-test. (b) Trial-averaged (*left*) choice-projected and (*right*) outcome-projected trajectories of population vectors in MFC (*top*) and DS (*bottom*). Gray lines show standard errors. Vertical lines indicate times t_1 (-0.2 s) and t_2 (+1.0 s) at which the trajectories were near maximally separated on choice and outcome axes, respectively. (c) Evolution of trial-averaged population vectors from -3 s to t_1 (*left*) or t_2 (*right*) in the space spanned by choice and outcome axes. Circles indicate population activities at t_1 (*left*) and t_2 (*right*). (d) The distances of choice- and outcome-projected trajectories from the separating hyperplanes at t_1 (*top*) and t_2 (*bottom*) are shown in MFC (*1st row*) and DS (*2nd row*). (e) Time evolution of Pearson's correlation coefficients was calculated between MFC and DS projected trajectories. Purple and green lines represent the mean correlation coefficients, and gray lines show individual correlation coefficient calculated from randomly sampled 431 trials. Circles indicate the times of the maximum correlation. (f) *Top*: Mutual information about choice (*left*) and outcome (*right*) obtained from the trajectories in MFC (black) and DS (red) shown in b. *Bottom*: Mutual information from 9 sessions (thin) and the average (thick). Asterisk indicates statistical significance (Mann-Whitney U test, $p < 0.05$). (g) Reward acquisition probability was significantly different between sessions with lower (< 0.5) and higher (> 0.5) trajectory correlations (t test, $P < 0.0092$).

Figure S4

Figure S4. MFC and DS trajectories and their session-by-session correlations. Each gray box shows results from one session. Within a box, neural trajectory in MFC (*top*) and DS (*middle*) on choice (*left*) and outcome (*right*) axes are shown. Color codes indicate Left rewarded trials (orange), Left unrewarded trials (brown), Right rewarded trials (light blue) and Right unrewarded trials (dark blue). Thin lines represent S.E.M. (*Bottom*) Time evolution of correlation coefficient between MFC and DS trajectories. Gray lines show the correlation coefficients calculated with randomly sampled 431 trials in session R985-r1, which contains the minimum trial numbers among the 9 sessions. The averaged correlation coefficients between MFC and DS were displayed for Ch-projected (purple) or Ot-projected (green) trajectories. Circles indicate the times of maximum correlation.

184 respectively. These trajectories were highly correlated between the two areas around the
185 corresponding time points (Fig. 4e). In the other 5 sessions, while the Ot-projected trajectories were
186 similar, the Ch-projected ones behaved differently in MFC and DS and the inter-trajectory
187 correlations were also weaker between the two areas (Fig. S4).

188

189 The enhanced correlations of Ch-projected and Ot-projected trajectories likely reflect choice
190 responses and the monitoring of reward, respectively. As the degree of correlations varied during
191 the task in all sessions, we addressed the behavioral variables correlated to the high MFC-DS
192 correlation of neural trajectories. We investigated the temporal relationships among licking behavior,
193 choice responses, reward delivery and the times of peak inter-trajectory correlations. As shown in
194 Fig. 1e, the licking rate quickly reached a peak when the choice response was made and then
195 decayed slowly during the outcome period ('reward licking period') in rewarded trials. In
196 unrewarded trials, the decay was much faster. The peak correlation times of Ch-projected
197 trajectories tended to cluster around choice response (t-test, $p = 0.366$) and reward delivery (t-test,
198 $p = 0.620$), while those of Ot-projected trajectories were significantly later than choice response (t-
199 test, $p = 2.65 \times 10^{-3}$) and reward delivery (t-test, $p = 0.0297$) (Fig. S5a), and were retained high during
200 the following licking behavior (Fig. S5b). These results suggest that the epochs of highly correlated
201 Ch- and Ot-projected trajectories are related to licking behavior or the intake of reward, respectively.
202 The trajectory correlations exhibited different temporal profiles in rewarded and unrewarded trials

Figure S5

Figure S5. Behavioral variable when the correlation of trajectories reached peak.

(a) Session-by-session licking rate, choice response (black) and reward delivery (light blue) are aligned at the time when Ch-projected (purple) and Ot-projected (green) trajectories showed peak correlations between MFC and DS. The origin of time refers to the time of peak correlations. Means (filled symbols) and SDs (horizontal lines) of choice response and reward delivery time are shown (N = 9 sessions).

(b) Time evolution of trajectory correlations are shown for a rat in rewarded (middle), unrewarded (bottom), and mixed (top) trials. Arrows indicate the times of peak correlation in the mixed trial condition. We randomly selected 63 trials in each condition and calculated correlation coefficient (gray line). We repeated this procedure by 10 times. Averaged correlation is represented by colored thick line. Circles indicate the peak values and times.

(c) Comparison of peak correlation coefficients between rewarded and unrewarded trials.

(Fig. S5b), but the peak correlations of Ch- (Wilcoxon signed-rank test, $p = 0.218$) and Ot-projected ($p = 0.546$) trajectories were not significantly different between the two trial types, suggesting that the enhanced correlations reflect the monitoring of reward rather than actual reward delivery.

It is intriguing to examine whether neural trajectory switched in either or both of MFC and DS when the rats switched their choices after unrewarded trials. In the present task, rats switched their choices after a reversal of reward block within a couple of unrewarded trials (Fig. 1b and c). We hypothesized that during the unrewarded trials neural trajectory should cross the hyperplane (Unless otherwise stated, hyperplane refers to Ch-hyperplane) from one side to the opposite side (i.e. 'Right to Left' or 'Left to Right') before a choice response in the next trial. Out of 9 sessions, in 3 sessions we observed a gradual crossing of MFC trajectory during ITIs, while a similar crossing of DS trajectory was observed only in one session (Fig. S6). Although the crossing was somewhat more prominent in MFC than in DS, it was difficult to determine which region dominated switching choice responses.

We further verified the analogous encoding of choice and outcome information in MFC and DS by calculating the time evolution of the information carried by the neural trajectories (Fig. 4f). The amount of information varied from session to session and was slightly greater in MFC than in DS. However, the mutual information averaged over nine sessions was not statistically different between

Figure S6

Figure S6. Crossing trajectories between two switching choice trials during inter-trial interval (ITI) period. Neural trajectory during ITI period post unrewarded episode in 3 sessions (left: R985-r2, Ch-projected trajectory; middle: R982-r1, Ch-projected trajectory; right: R1005-r1, Ot-projected trajectory) in MFC (1st row) and DS (3rd row). Dark blue and brown denote the neural trajectories when rats switched choices from Right to Left and from Left to Right, respectively. Black thin lines denote standard errors. Differences in the two trajectories were quantified by AROC values in MFC (2nd row, black) and DS (4th row, red). A gradual shift from high values (> 0.5) to low values (< 0.5), suggests that the trajectories evolved to the opposite side of the corresponding hyperplane. Statistical significance of AROC values is estimated by Mann-Whitney U test ($p < 0.05$) and represented by thick lines. Two vertical lines show the duration of start cue presentation.

222 MFC and DS at most of the times around choice response (Mann-Whitney U test, $p > 0.05$). These
 223 results indicate that neural trajectories evolve in parallel in MFC and DS and transiently increase the
 224 mutual correlation to transfer choice and outcome information without a large loss. Reward
 225 acquisition probability was significantly larger ($p = 0.0092$, two-sample t test) in sessions with high
 226 inter-trajectory correlations (> 0.5) during choice and outcome periods than in sessions with low
 227 correlations (Fig. 4g). In this analysis, the session R1005-r1 was included in the low-correlation group
 228 because the correlation was low during choice and outcome periods. Omission of this session did
 229 not change the essential result of comparison ($p = 0.0193$, two-sample t test).

230

231 ***Synchronous spiking when neural trajectories were highly correlated between MFC and DS***

232 As shown above, neural trajectories in MFC and DS exhibit task event-dependent modulations of
 233 coherence. This result raises questions about whether neuronal firing is also temporally correlated
 234 between these regions and whether correlated spikes, if any, are related to neural trajectories. To
 235 study these questions, we calculated cross-correlograms (CCGs) of MFC-DS cell pairs in the time
 236 window in which MFC-DS correlations of trajectories were maximized (see Fig.4e and Fig. S4)
 237 (Methods). Then, we compared task-event relevant and irrelevant CCGs in three trial categories (i.e.,
 238 Left-choice, Right-choice and rewarded trials).

239

Our analysis revealed MFC-DS cell pairs displaying statistically significant spike synchrony in a task-event-dependent manner (Fig. 5a). Shuffling spike trains across trials within the individual pairs eliminated the peaks in CCGs (Fig. S7), implying that these cell pairs fired with a close temporal relationship in each trial. The peak values of CCGs showed significant correlations between MFC and DS (> 4 SD of the base line) in task-event related windows, but these values were greatly decreased if CCGs were calculated in task-event-irrelevant windows (i.e., at random time points) (Fig. 5a and b). These results suggest that the emergence of spike synchrony between MFC and DS cells is correlated with task events.

Spike synchrony tended to occur when neural trajectories were highly correlated between MFC and DS, which further supports the role of spike synchrony in coordinating neural population activities in the two regions. The spike count of MFC cells exhibited statistically significant peaks at near zero time lags from spikes of DS cells in 4 sessions (R983-r1, R983-r2, R985-r1 and R985-r2 in Fig. 5c). At this epoch, the maximum correlation coefficients between MFC and DS trajectories were near or over 0.5 (Fig. 4e and Fig. S4). By contrast, the CCG distributions showed no significant peaks in the other 5 sessions (R982-r1, R991-r1, R1005-r1, R1009-r1, and R1012-r1). However, in these 5 sessions the correlations of neural trajectories were also weaker compared to the former 4 sessions (Fig. S4). Interestingly, in rewarded trials the correlation coefficient between Ot-projected trajectories was correlated with the fraction of cell pairs showing CCG peaks within ± 2 ms (Pearson

Figure S7

Figure S7. Significant cross correlations between MFC and DS cell were dependent on the spike timing in the ongoing trial but not on the number of spikes. Left: CCGs of the cell pair #1 in Fig. 5a were calculated in the original (left) and shuffled (right) trial orders in rewarded trials. Right: Normalized CCG peak amplitudes were calculated for all cell pairs showing significant CCG peaks in windows Ch (purple) and Ot (green). The peak amplitude are compared with those obtained in shuffled trial orders. Horizontal lines and error bars show the median and 1st/3rd quartiles, respectively. P-values were calculated by Wilcoxon signed-rank test.

Figure 5

Figure 5. Relationships among MFC-DS spike synchrony, trajectory correlation and task performance. (a) Spike CCGs of two MFC-DS cell pairs. For each cell pair, time lags of MFC cell spike firing relative to DS cell spike firing were calculated in task-event relevant (*middle row*) and irrelevant (*bottom row*) time windows. Average and SD (gray shade) of spike waveforms recorded with tetrodes in MFC (black) and DS (red) are also shown (*top row*). *Left and middle columns*: Left or Right choice trials with arbitrary outcomes, respectively. *Right column*: rewarded trials with arbitrary choice positions. Windows Ch (purple) and Ot (green) refer to the analyzed time windows centered around the peak MFC-DS correlations of Ch-projected and Ot-projected trajectories, respectively. Task-event irrelevant CCG gives a control and was calculated at a randomly selected time point. Arrowhead indicates the peak of CCGs above the mean + 4SD of the baseline defined by spike counts from -100 to -80 ms and from +80 to +100 ms (gray shaded ranges). (b) CCG peak values of cell pairs in windows Ch (purple) and Ot (green) are compared with those in task-event irrelevant windows (black). Each dot represents a cell pair with a significant CCG peak (peak > 4SD). Horizontal lines and error bars show the median and 1st/3rd quartiles, respectively. P-values were calculated by Wilcoxon signed-rank test. (c) Proportion of MFC-DS cell pairs that displayed an excess amount of spike coincidences in rewarded trials was plotted for the 9 sessions against CCG peak time lags. The numbers of such cell pairs are given in parentheses. (d) (*Top*) Correlation coefficients between MFC and DS Ot-projected trajectories are plotted against the fraction of MFC-DS cell pairs showing a CCG peak within ± 2 ms for different time windows. (*Bottom*) Relationship between such MFC-DS cell pairs and animal's reward acquisition probability. Symbols represent individual sessions, and Pearson correlation coefficient and P-value are denoted for each diagram.

correlation test, window O_t , $r = 0.689$, $p = 0.0398$) (Fig. 5d, top). Further, the fraction of such cell pairs was correlated with the probability of animal's reward acquisition (Fig. 5d, bottom). These results suggest that precisely correlated firing of MFC and DS cells underlies the coherent evolution of neural trajectories between these areas and boost animal's task performance.

Inactivation of MFC attenuated neural representation of choice and outcome in DS cells

To confirm the contributions of inter-trajectory correlations between MFC and DS to adaptive control, we examined the effect of bilateral MFC inhibition on the choice behavior and DS-cell activity in 5 rats (muscimol group) (Fig. 6a, Methods). It is known that MFC is one of the main regions projecting excitatory inputs to DS [6-8]. Another 10 rats were examined the choice behavior without bilateral MFC inhibition as a control group. The muscimol injection induced variable choice patterns in five muscimol-treated rats (F test, $F = 16.95$, $p = 0.0179$). After the injection (Day 2), the reward acquisition probabilities of four rats were significantly different from those prior to the injection (Day 1) (χ^2 test, $p < 0.05$). Among these four, the reward acquisition behavior was impaired in three while another rat rather improved the behavioral performance (Fig. 6b). Figure 6c shows an example of impaired task performance. This rat reduced "stay" behavior after preceding rewarded outcomes but increased such behavior after preceding unrewarded outcomes (Fig. 6c-d). Indeed, the fractions of "stay" and "switch" choice patterns were significantly different between Day1 and Day2 in the muscimol group (two-way ANOVA, $F_{1,4} = 5.774$, $p = 0.021$) but not in the control group ($F_{1,9} = 0.248$,

Figure 6

Figure 6. Inactivation of MFC attenuated neuronal representation of choice and outcome in DS. In statistical comparisons, horizontal and vertical lines indicate the median and 1st/3rd quartiles, respectively. (a) Experimental schedules for control (N = 10 rats) and muscimol (N = 5 rats) groups. At the first session (Day1), all 15 rats underwent the recordings of multi-neuron activity from MFC and DS without any pharmacological treatments. At the second session (Day2), we injected muscimol into the MFC of five rats (muscimol group) and recorded multi-neuron activity from DS. For the remaining 10 rats (control group), neural activities were recorded from MFC and DS without any injection. (b) Reward acquisition probabilities in Day1 and Day2 are shown in control (black) and muscimol (red) groups. Each symbol shows data from a rat. Colored (black or red) symbols indicate the individual rats that showed significantly different probabilities (χ^2 test, $p < 0.05$) between Day1 and Day2. Arrowhead indicates the rat shown in c. The differences in reward acquisition probabilities between Day1 and Day2 were compared between groups by two-sample t-test. (c) An example of impaired adaptive choice behavior of a muscimol-injected rat is shown on Day1 (top, no injection) and Day2 (bottom, muscimol injection). (d) Mean number of choice patterns per block on Day1 (left) and Day2 (right). Left: control group, Right: muscimol group. R: rewarded in last trial, U: unrewarded in last trial, Rp: repeat in current trial, Sw: switch in current trial. (e) Firing patterns of DS cells were compared (Mann-Whitney U test) in the muscimol group in terms of firing rate (Left) and inter-spike interval (Right) between Day1 (n = 130, orange) and Day2 (n = 77, red). Circles indicate the firing rates of individual DS neurons. (f) Time evolution of the fraction of (top) choice-modulated and (bottom) outcome-modulated DS cells (two-way ANOVA, 200-ms-long sliding-by-20-ms bins, $p < 0.05$) on Day1 (orange) and Day2 (red). Blue bars show the epochs of significantly different fractions between Day1 and Day2 (χ^2 test, $p < 0.05$).

278 $p = 0.620$). Accordingly, the control and muscimol groups showed significantly different reward
279 acquisition probabilities between Day 1 and Day2 (two-sample t-test, $p = 0.0472$). The behavioral
280 variability between the two groups was unlikely attributed to the attenuation of their motivations
281 for the task because the number of trials was not significantly different between Day1 (mean $\pm$ SD =
282 731 ± 86 trials) and Day2 (740 ± 96 trials) in the muscimol group (unequal variance t-test, $p = 0.877$)
283 and these numbers were not significantly different from those of the control group. Those results
284 suggest that MFC is not a sole contributor to the present adaptive choice behavior.

285

286 The inhibition of MFC diminished neuronal activities and their selective modulations in DS.
287 The firing rates of DS cells ($n = 77$) were significantly lower (Mann-Whitney U-test, $p = 0.0157$) and
288 the median of inter-spike intervals was significantly longer (Mann-Whitney U-test, $p = 0.0051$) on
289 Day2 than on Day1 ($n = 130$) (Fig. 6e), suggesting an attenuated excitatory drive or an enhanced
290 inhibition on DS cells. The MFC inhibition also altered the task-related activities of DS cells, especially
291 those encoding choices and outcomes (two-way ANOVA, $p < 0.05$). The proportion of choice-
292 modulated cells around and after choice responses was significantly smaller on Day2 than on Day1
293 (χ^2 test, $p < 0.05$) (Fig. 6f, top), and so was the proportion of outcome-modulated cells during the
294 latter part of outcome period (more than 2 s after a choice response). However, this proportion was
295 not significantly different during the former part of outcome period (Fig. 6f, bottom). Taken together,
296 our results suggest that neuronal activation in MFC is engaged in DS neural coding both choice-

297 related information around response time and outcome-related information during reward licking
298 behavior.

299

300

301 **Discussion**

302 This study revealed a highly parallel evolution of MFC and DS neural trajectories regulating outcome-
303 based choice behavior. These trajectories initially encoded a current choice and later reflected the
304 outcome of the choice response. Despite variation in temporal patterns of the trajectories across
305 sessions and individuals, the amount of information carried by the trajectories was not significantly
306 different between MFC and DS. Furthermore, cross-area spike coincidences in the millisecond range
307 were enhanced in rewarded trials when the trajectories were highly correlated between the two
308 regions, suggesting that the parallel trajectory evolution in MFC and DS actively participates in
309 coordinating the behavioral task.

310

311 Along the frontal cortex-basal ganglia axis, the MFC-DS channel is thought to be the stage
312 to process action selection based on past outcomes [1,3]. A widely hypothesized mechanism for this
313 process is that neural ensemble in DS selectively gates the necessary information (or suppresses the
314 unnecessary information) received from MFC. However, our results provide no support for this
315 hypothesis because neural populations in MFC and DS preserve similar amounts of information

316 about choices as well as outcomes and their trajectories evolve quite similarly in the two regions
317 (Fig. 4).

318

319 Our data suggest in contrast to the gating hypothesis that the neural trajectory of DS
320 ensembles reflects that of MFC ensembles, which is consistent with previous work demonstrating a
321 simultaneous and correlated activation of time-dependent ramping activity in MFC and DS during a
322 temporal judgment task [36]. An alternative interpretation, given that MFC innervates DS
323 unidirectionally, is that MFC cells exhibit a choice-related signal prior to DS cells and then convey the
324 signal to DS in a sequence, as previously suggested [14,37]. However, our data does not support this
325 interpretation because choice-related signals concomitantly emerge in MFC and DS. This discrepancy
326 may be attributed to several differences in experimental conditions. For instance, unlike other
327 studies, we recorded neuronal activities in a head-restrained condition, did not change the
328 probability of reward delivery during the task, and directly compared simultaneous neuronal
329 activations between MFC and DS.

330

331 A possible role of the corticostriatal circuit suggested by our findings is that construction of
332 temporal correlation of neural activation between MFC and DS could lead to consolidating efficient
333 transmission of information from frontal cortex to downstream of basal-ganglia or to Inhibiting
334 inappropriate action. A previous study demonstrated the synchronization of oscillatory local field

potentials at 5-13 Hz across motor cortex, DS and substantia nigra pars reticulata, which is the output terminal of the basal ganglia, and the propagation of spindle-like spike-and-wave oscillations along this cortico-basal ganglia pathway in freely moving rats. Spikes of individual neurons were phase-locked to the oscillatory local field potential and synchronized within the cortical-basal ganglia network [38]. Since synchronized neuronal discharges may enable a reliable information transmission [29,30], the temporal correlation between neural ensembles in the frontal cortex-basal ganglia network could be a consequence of learning the decision behavior. In accordance with this view, during skill learning, neural ensembles in both motor cortex and DS are known to develop spike correlations [39,40] and precise temporal spiking patterns [31,41]. In the present study, synchronous spiking was coordinated between MFC and DS cells when neural trajectories in these areas were strongly correlated in the well learned rats, supporting the hypothesis (Fig. 5).

These results raise the question about how such a precise spike synchrony, which is indicated by the short latency of CCG peak between MFC and DS cells (Fig. 5), was generated between the cortical and subcortical regions. There are at least two possible explanations. One possibility is that corticostriatal neurons in MFC make monosynaptic connections to the postsynaptic DS cells. In this case, presynaptic MFC cells should discharge prior to the firing of postsynaptic DS cells, but this was not the case at all in the CCGs between these cells (Fig. 5). Another possibility is that the firing of MFC and DS cells was driven by a common input from other corticostriatal neurons

354 in the MFC [42,43] or neurons in other cortical areas projecting to both MFC and DS. Medium spiny
355 DS cells receive excitatory inputs from ipsilateral and contralateral MFC [16], while MFC cells' output
356 projects to both ipsilateral and contralateral cortical areas and DS through corpus callosum, the so-
357 called crossed-corticostratial (or intratelencephalic) cells [6,44]. Since excitatory connections are
358 abundant between MFC corticostratial pyramidal cells [45], synchronous discharges of these
359 neurons generate large EPSPs in their postsynaptic target cells in MFC and DS, which increases the
360 probability of synchronous firing between MFC and DS cells. Alternatively, both MFC [46] and DS
361 [7,8] receive afferent connections from higher-order cortical areas such as the ventrolateral OFC and
362 PPC. Since the OFC and PPC are implicated in goal-directed decision-making [47,48], excitatory
363 inputs from a common subpopulation of OFC and/or PPC neurons may evoke synchronous spiking
364 of MFC and DS cells during decision-making. In the present study, bilateral MFC inactivation
365 attenuated, but did not completely eliminate, choice and outcome representations in DS cells (Fig.
366 6), suggesting that brain regions other than MFC also drive DS cells in outcome-based action
367 selection. For instance, optogenetic inhibition of PPC corticostratial projection neurons altered the
368 history dependent choice bias during decision making [49]. To distinguish between the different
369 causes of spike synchrony between MFC and DS cells further recordings from multiple regions
370 involving MFC, OFC, PPC, and DS are required.

371

372 It was recently shown in the three-layered cortex of the reptile that just a few spikes of

373 single pyramidal neurons can trigger a cascade of firing sequences of neuron ensembles [50], as was
374 suggested by theoretical studies [29,51]. Our results indicate the possibility that a similar cascade of
375 firing sequences is triggered in the cortico-basal ganglia pathway during the parallel evolution of
376 frontal cortical and striatal neural trajectories. Such spike sequences may not only enhance reliable
377 information transmission (i.e., spikes) from the cortex to the basal ganglia, but also may contribute
378 to regulating the gain of synaptic plasticity together with modulations by dopaminergic afferents
379 [52,53]. Further experiments are required to clarify these points.

380

381 In sum, we have shown that neural trajectories evolve in DS during goal-directed decision
382 making in parallel with a similar trajectory evolution in MFC. When the parallel trajectory evolution
383 in DS and MFC exhibits a strong coherence, spike synchrony is also enhanced between the two
384 regions and behavioral performance tends to be improved. Our results suggest that the parallel
385 evolution of cortical and striatal neural trajectories is a hallmark of successful decision making.

386

387

388 **Acknowledgments**

389 We thank T. Sharp and T. Takekawa for their advice and supports for initial analysis on our data. We
390 are grateful to M. Gilson, P. Goncalves, J. Igarashi, Y. Isomura, T. Kurikawa, and M. Tatsuno for their
391 valuable comments on our results. We thank J. Wickens for his critical reading of our manuscript and

392 Y. Goda for her advices. This work was partly supported by the Grants-in-Aid for Scientific Research
393 (KAKENHI) from MEXT (no. 24700345 to TH, nos. 18H05213 and 19H04994 to TF).

394

395 **Author contributions**

396 TH and TF designed the project. TH and RH conducted experiments. TH analyzed data. TH and TF
397 wrote the manuscript.

398

399 **Declaration of Interests**

400 The authors declare no competing interests.

401

References

1. Nambu A. (2008) Seven problems on the basal ganglia. *Curr Opin Neurobiol*, vol. 18: 595-604.
<https://doi.org/10.1016/j.conb.2008.11.001>
2. Balleine BW, O'Doherty JP. (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, vol. 35: 48-69. <https://doi.org/10.1038/npp.2009.131>
3. Hikosaka O, Isoda M. (2010) Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn Sci*, vol. 14: 154-161.
<https://doi.org/10.1016/j.tics.2010.01.006>
4. Friedman A, Homma D, Gibb LG, Amemori KI, Rubin SJ, Hood AS, Riad MH, Graybiel AM. (2015) A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell*, vol. 161: 1320-33. <https://doi.org/10.1016/j.cell.2015.04.049>
5. Brown EN, Kass RE, Mitra PP. (2004) Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat Neurosci*, vol. 7: 456-461. <https://doi.org/10.1038/nn1228>
6. Wilson CJ. (1987) Morphology and synaptic connections of crossed corticostriatal neurons in the rat. *J Comp Neurol*, vol. 263: 567-580. <https://doi.org/10.1002/cne.902630408>
7. Cheatwood JL, Reep RL, and Corwin JV. (2003) The associative striatum: cortical and thalamic projections to the dorsocentral striatum in rats. *Brain Res*, vol. 968: 1-14.
[https://doi.org/10.1016/S0006-8993\(02\)04212-9](https://doi.org/10.1016/S0006-8993(02)04212-9)

- 421 8. Reep RL, Cheatwood JL, Corwin JV. (2003) The associative striatum: organization of cortical
422 projections to the dorsocentral striatum in rats. *J Comp Neurol*, vol. 467: 271-292.
423 <https://doi.org/10.1002/cne.10868>
- 424 9. Narayanan NS, Cavanagh JF, Frank MJ, Laubach M. (2013) Common medial frontal mechanisms
425 of adaptive control in humans and rodents. *Nat Neurosci*, vol. 16: 1888-1895.
426 <https://doi.org/10.1038/nn.3549>
- 427 10. Hyman JM, Holroyd CB, Seamans JK. (2017) A novel neural prediction error found in anterior
428 cingulate cortex ensembles. *Neuron*, vol. 95: 447-456.
429 <https://doi.org/10.1016/j.neuron.2017.06.021>
- 430 11. Erlich JC, Bialek M, Brody CD. (2011) A cortical substrate for memory-guided orienting in the
431 rat. *Neuron*, vol. 72: 330-343. <https://doi.org/10.1016/j.neuron.2011.07.010>
- 432 12. Handa T, Takekawa T, Harukuni R, Isomura Y, Fukai T. (2017) Medial frontal circuit dynamics
433 represents probabilistic choices for unfamiliar sensory experience. *Cereb Cortex*, vol. 27: 3818-
434 3831. <https://doi.org/10.1093/cercor/bhx031>
- 435 13. Kurikawa T, Haga T, Handa T, Harukuni R, Fukai T. (2018) Neuronal stability in medial frontal
436 cortex sets individual variability in decision-making. *Nat Neurosci*, vol. 21: 1764-1773.
437 <https://doi.org/10.1038/s41593-018-0263-5>
- 438 14. Sul JH, Jo S, Lee D, Jung MW. (2011) Role of rodent secondary motor cortex in value-based
439 action selection. *Nat Neurosci*, vol. 14: 1202-1208. <https://doi.org/10.1038/nn.2881>

- 440 15. Gremel C, Costa R. (2013a) Premotor cortex is critical for goal-directed actions. Front in
441 Comput Neurosci, vol.7: 110. <https://doi.org/10.3389/fncom.2013.00110>
- 442 16. Wilson CJ. (1986) Postsynaptic potentials evoked in spiny neostriatal projection neurons by
443 stimulation of ipsilateral and contralateral neocortex. Brain Res, vol. 367: 201-213.
444 [https://doi.org/10.1016/0006-8993\(86\)91593-3](https://doi.org/10.1016/0006-8993(86)91593-3)
- 445 17. McGeorge AJ, Faull RL. (1989) The organization of the projection from the cerebral cortex to
446 the striatum in the rat. Neuroscience, vol. 29: 503–537. [https://doi.org/10.1016/0306-](https://doi.org/10.1016/0306-4522(89)90128-0)
447 [4522\(89\)90128-0](https://doi.org/10.1016/0306-4522(89)90128-0)
- 448 18. Smith Y, Raju DV, Pare JF, Sidibe M. (2004) The thalamostriatal system: a highly specific
449 network of the basal ganglia circuitry. Trends in Neurosci, vol. 27: 520-527.
450 <https://doi.org/10.1016/j.tins.2004.07.004>
- 451 19. Wickens JR, Reynolds JNJ, Hyland BI. (2003) Neural mechanisms of reward-related motor
452 learning. Curr Opin Neurobiol, vol. 13: 685-690. <https://doi.org/10.1016/j.conb.2003.10.013>
- 453 20. Schultz W. (2006) Behavioral theories and the neurophysiology of reward. Annu Rev Psychol,
454 vol. 57: 87-115. <https://doi.org/10.1146/annurev.psych.56.091103.070229>
- 455 21. Reynolds, JNJ, Hyland BI, Wickens JR. (2001) A cellular mechanism of reward-related learning.
456 Nature, vol. 413: 67-70. <https://doi.org/10.1038/35092560>
- 457 22. Lauwereyns J, Watanabe K, Coe B, Hikosaka O. (2002) A neural correlate of response bias in
458 monkey caudate nucleus. Nature, vol. 418: 413-417. <https://doi.org/10.1038/nature00892>

- 459 23. Isomura Y, Takekawa T, Harukuni R, Handa T, Aizawa H, Takada M, Fukai T. (2013) Reward-
460 modulated motor information in identified striatum neurons. J Neurosci, vol. 33: 10209 -
461 10220. <https://doi.org/10.1523/JNEUROSCI.0381-13.2013>
- 462 24. Nonomura S, *et al.* (2018) Monitoring and updating of action selection for goal-directed
463 behavior through the striatal direct and indirect pathways. Neuron, vol. 99: 1302-1314.
464 <https://doi.org/10.1016/j.neuron.2018.08.002>
- 465 25. Packard MG, Hirsh R, White NM. (1989) Differential effects of fornix and caudate nucleus
466 lesions on two radial maze tasks: evidence for multiple memory systems. J Neurosci, vol. 9:
467 1465-1472. <https://doi.org/10.1523/JNEUROSCI.09-05-01465.1989>
- 468 26. Skelin I, Hakstol R, VanOyen J, Mudiayi D, Molina LA, Holec V, Hong NS, Euston DR, McDonald
469 RJ, Gruber AJ. (2014). Lesions of dorsal striatum eliminate lose-switch responding but not
470 mixed-response strategies in rats. Eur J Neurosci, vol. 39: 1655-1663.
471 <https://doi.org/10.1111/ejn.12518>
- 472 27. Harvey CD, Coen P, Tank DW. (2012) Choice-specific sequences in parietal cortex during a
473 virtual-navigation decision task. Nature, vol. 484: 62–68. <https://doi.org/10.1038/nature10918>
- 474 28. Bakhurin KI, Goudar V, Shobe JL, Claar LD, Buonomano DV, Masmanidis SC. (2017) Differential
475 encoding of time by prefrontal and striatal network dynamics. J Neurosci, vol. 37: 854-870.
476 <https://doi.org/10.1523/JNEUROSCI.1789-16.2016>
- 477 29. Salinas E, Sejnowski TJ. (2001) Correlated neuronal activity and the flow of neural information.

- 478 Nat Rev Neurosci, vol. 2: 539-550. <https://doi.org/10.1038/35086012>
- 479 30. Buzsáki G, Schomburg EW. (2015) What does gamma coherence tell us about inter-regional
480 neural communication? Nat Neurosci, vol. 18: 484-489. <https://doi.org/10.1038/nn.3952>
- 481 31. Koralek AC, Costa RM, Carmena JM. (2013) Temporally precise cell-specific coherence
482 develops in corticostriatal networks during learning. Neuron, vol. 79: 865-872.
483 <https://doi.org/10.1016/j.neuron.2013.06.047>
- 484 32. Girardeau G, Inema I, Buzsáki G. (2017) Reactivations of emotional memory in the
485 hippocampus–amygdala system during sleep. Nat Neurosci, vol. 20: 1634-1642.
486 <https://doi.org/10.1038/nn.4637>
- 487 33. Isoda M, Hikosaka O. (2007) Switching from automatic to controlled action by monkey medial
488 frontal cortex. Nat Neurosci, vol. 10: 240-248. <https://doi.org/10.1038/nn1830>
- 489 34. Narayanan NS, Laubach M. (2008) Neuronal correlates of post-error slowing in the rat
490 dorsomedial prefrontal cortex. J Neurophysiol, vol. 100: 520-525.
491 <https://doi.org/10.1152/jn.00035.2008>
- 492 35. Bishop C. (2006) Pattern Recognition and Machine Learning. New York: Springer.
- 493 36. Emmons EB, De Corte BJ, Kim Y, Prker KL, Matell MS, Narayanan NS. (2017) Rodent medial
494 frontal control of temporal processing in the dorsomedial striatum. J Neurosci, vol. 37: 8718-
495 8733. <https://doi.org/10.1523/JNEUROSCI.1376-17.2017>
- 496 37. Ma L, Hyman JM, Lindsay AJ, Phillips AG, Seamans JK. (2014) Differences in the emergent

- 497 coding properties of cortical and striatal ensembles. *Nat Neurosci*, vol. 17: 1100-1106.
- 498 <https://doi.org/10.1038/nn.3753>
- 499 38. Dejean C, Gross CE, Bioulac B, Boraud T. (2007) Synchronous high-voltage spindles in the
- 500 cortex-basal ganglia network of awake and unrestrained rats. *Eur J Neurosci*, vol. 25: 772-784.
- 501 <https://doi.org/10.1111/j.1460-9568.2007.05305.x>
- 502 39. Costa RM, Cohen D, Nicolelis MA. (2004) Differential corticostriatal plasticity during fast and
- 503 slow motor skill learning in mice. *Current Biology*, vol. 14: 1124-1134.
- 504 <https://doi.org/10.1016/j.cub.2004.06.053>
- 505 40. Santos FJ, Oliveira RF, Jin X, Costa RM. (2015) Corticostriatal dynamics encode the refinement
- 506 of specific behavioral variability during skill learning. *Elife*, vol. 4: e09423. [10.7554/eLife.09423](https://doi.org/10.7554/eLife.09423)
- 507 41. Lemke SM, Ramanathan DS, Guo L, Won SJ, Ganguly K. (2019) Emergent modular neural
- 508 control drives coordinated motor actions. *Nat Neurosci* vol. 22: 1122–1131.
- 509 <https://doi.org/10.1038/s41593-019-0407-2>
- 510 42. Shepherd GM. (2013) Corticostriatal connectivity and its role in disease. *Nat Rev Neurosci*, vol.
- 511 14: 278-291. <https://doi.org/10.1038/nrn3469>
- 512 43. Kawaguchi Y. (2017) Pyramidal cell subtypes and their synaptic connections in layer 5 of rat
- 513 frontal cortex. *Cereb Cortex*, vol. 27: 5755-5771. <https://doi.org/10.1093/cercor/bhx252>
- 514 44. Reiner A, Jiao Y, Del Mar N, Laverghetta AV, Lei WL. (2003) Differential morphology of
- 515 pyramidal tract-type and intratelencephalically projecting-type corticostriatal neurons and

- 516 their intrastriatal terminals in rats. *J Comp Neurol*, vol. 457: 420-440.
- 517 <https://doi.org/10.1002/cne.10541>
- 518 45. Morishima M, Kawaguchi Y. (2006) Recurrent connection patterns of corticostriatal pyramidal
- 519 cells in frontal cortex. *J Neurosci*, vol. 26: 4394–4405.
- 520 <https://doi.org/10.1523/JNEUROSCI.0252-06.2006>
- 521 46. Reep RL, Goodwin GS, Corwin JV. (1990) Topographic organization in the corticocortical
- 522 connections of medial agranular cortex in rats. *J Comp Neurol*, vol. 294: 262-280.
- 523 <https://doi.org/10.1002/cne.902940210>
- 524 47. Gremel C, Costa R. (2013b) Orbitofrontal and striatal circuits dynamically encode the shift
- 525 between goal-directed and habitual actions. *Nat Commun*, vol. 4: 2264.
- 526 <https://doi.org/10.1038/ncomms3264>
- 527 48. Erlich JC, Brunton BW, Duan CA, Hanks TD, Brody CD. (2015) Distinct effects of prefrontal and
- 528 parietal cortex inactivations on an accumulation of evidence task in the rat. *Elife*, vol. 4:
- 529 e05457. [10.7554/eLife.05457](https://doi.org/10.7554/eLife.05457)
- 530 49. Hwang EJ, *et al.* (2019) Corticostriatal flow of action selection bias. *Neuron*, vol. 104: 1126-
- 531 1140. <https://doi.org/10.1016/j.neuron.2019.09.028>
- 532 50. Hemberger M, Shein-Idelson M, Pammer L, Laurent G. (2019) Reliable sequential activation of
- 533 neural assemblies by single pyramidal cells in a three-layered cortex. *Neuron*, vol. 104: 353-
- 534 369. <https://doi.org/10.1016/j.neuron.2019.07.017>

- 535 51. Diesmann M, Gewaltig MO, Aertsen A. (1999) Stable propagation of synchronous spiking in
536 cortical neural networks. *Nature*, vol. 402: 529-533. <https://doi.org/10.1038/990101>
- 537 52. Pawlak V, Kerr JN. (2008) Dopamine receptor activation is required for corticostriatal spike-
538 timing-dependent plasticity. *J Neurosci*, vol. 28: 2435-2446.
539 <https://doi.org/10.1523/JNEUROSCI.4402-07.2008>
- 540 53. Shen W, Flajolet M, Greengard P, Surmeier DJ. (2008) Dichotomous dopaminergic control of
541 striatal synaptic plasticity. *Science*, vol. 321: 848-851. [10.1126/science.1160575](https://doi.org/10.1126/science.1160575)
- 542 54. Isomura Y, Harukuni R, Takekawa T, Aizawa H, Fukai T. (2009) Microcircuitry coordination of
543 cortical motor information in self-initiation of voluntary movements. *Nat Neuroscience*, vol.
544 12: 1586-1593. <https://doi.org/10.1038/nn.2431>
- 545 55. Narayanan NS, Horst NK, Laubach M. (2006) Reversible inactivations of rat medial prefrontal
546 cortex impair the ability to wait for a stimulus. *Neuroscience*, vol. 139: 865-876.
547 <https://doi.org/10.1016/j.neuroscience.2005.11.072>
- 548 56. Paxinos G, Watson C. (2009) *The rat brain in stereotaxic coordinates*. New York: Elsevier.
- 549 57. Takekawa T, Isomura Y, Fukai T. (2010) Accurate spike sorting for multi-unit recordings. *Eur J*
550 *Neurosci*, vol. 31: 263-272. <https://doi.org/10.1111/j.1460-9568.2009.07068.x>
- 551 58. Takekawa T, Isomura Y, Fukai T. (2012) Spike sorting of heterogeneous neuron types by
552 multimodality-weighted PCA and explicit robust variational Bayes. *Front Neuroinform*, vol. 6:
553 5. <https://doi.org/10.3389/fninf.2012.00005>

554 59. Hazan L, Zugaro M, Buzsáki G. (2006) Klusters, NeuroScope, NDManager: a free software suite
555 for neurophysiological data processing and visualization. J Neurosci Methods, vol. 155: 207–
556 216. <https://doi.org/10.1016/j.jneumeth.2006.01.017>

557 60. Britten KH, Shadlen MN, Newsome WT, Movshon JA. (1992) The analysis of visual motion: a
558 comparison of neuronal and psychophysical performance. J Neurosci, vol. 12: 4745-4765.
559 <https://doi.org/10.1523/JNEUROSCI.12-12-04745.1992>

560 61. Shannon CE. (1948) A mathematical theory of communication. Bell Syst Tech J, vol. 27 :379–
561 423.
562

563 **Methods**

564 ***Animal preparation.*** All experiments were approved and carried out in accordance with the
565 Animal Experiment Plan by the Animal Experiment Committee of RIKEN. Male Long-Evans rats (N =
566 34, 6 weeks, 200-220 g, Japan SLC, Inc.) were used. Home cages were set in a temperature and
567 humidity controlled environment with lights maintained on a 12-h light/dark cycle. Prior to a primary
568 surgery, rats were handled briefly and habituated to a stainless-steel cylinder in the cages. All surgical
569 procedures were operated under sterile circumstances. Animals underwent three surgical
570 procedures. Rats were anaesthetized with 2% isoflurane. Their body temperature was monitored
571 with a rectal probe and maintained at ~37°C on a heating pad during the surgery.

572 At a primary surgery, a sliding head-attachment (Narishige) was implanted on the skull with
573 dual-curing resin cement (Panavia, Kuraray Noritake Dental) and dental resin (Unifast II, GC) as done
574 before in our laboratory [12,13,23,54]. Reference and grounding electrodes (teflon coated silver
575 wire) were put on dura mater above the cerebellum. After recovery from the surgery, the rats were
576 deprived water intake in the home cages in order to utilize water as a reward for the execution of a
577 task, although food was available *ad libitum*. Water (10 ml) was supplemented at every weekend. At
578 a second surgery 3 days prior to a first recording session, we injected a retrograde tracer Fluoro-Gold
579 (FG) into the DS in order to confirm if our recording sites corresponded to the MFC sending
580 corticostriatal projection to the DS. A glass micropipette filled with 2% FG (Fluorochrome) dissolved
581 in 0.1 M cacodylic acid was installed on a micromanipulator angled medially by 27 deg. The pipette

was inserted through a tiny burr hole drilled in the skull over left hemisphere (+1.5 mm to Bregma, 1.0 mm lateral to midline, 4.3 mm traveling distance) so that the pipette tip reached the dorsocentral part of striatum (+1.5 mm anterior to Bregma, approximately 3.0 mm lateral to midline, approximately 3.8 mm ventral to pia mater), which rostral agranular medial cortex (AGm) pyramidal neurons directly innervate (Figs. 2 and S1). FG was iontophoretically loaded through 7 s pulse of +5.0 μ A by 7 s interval for 30 or 60 minutes with an iontophoresis pump (BAB-501, Kation Scientific). Then, the tiny hole was covered with sterilized spongel and dental silicone sealants (Matsukaze). At a third surgery after the training sessions were completed, two cranial windows (1.2 mm diameter) were made above the DS and MFC of left hemisphere (+1.0 and +3.0 mm to Bregma, 3.0 and 1.0 mm to midline for DS and rostral MFC, respectively) and then its dura maters were removed. For bilateral muscimol injection session, another tiny burr hole was additionally drilled in the skull above MFC of right hemisphere (+3.0 mm to Bregma, 1.0 mm to midline). The cranial windows were covered with the silicone sealants.

Behavioral training apparatus. After recovery from the first surgery, the rats were trained to perform a behavioral task controlled under a customized multiple-rats training system (O'hara), which enabled training several rats to learn a task paradigm in parallel [23,54]. The behavioral task was controlled by a custom-written software with LabVIEW (National Instruments). In an isolated training chamber, each rat was placed at a body-supporting cylinder, and its head was fixed rigidly

601 and painlessly by screwing a sliding head-holder on a stereotaxic frame. Auditory stimuli were
 602 presented at 60-dB SPL via a speaker placed in front of the rat. The timing and direction of licking
 603 movement were detected when its tongue interrupted an infrared beam placed below the left and
 604 right spouts. The detection sensitivity was much enough to detect tongue movements during licking
 605 spout, but not sensitive enough to detect other potential movements such as whisking and
 606 mastication. White plastic plates were located at both sides of the face in order to prohibit rats
 607 from approaching the beam detectable spaces by forelimb reaching (Fig. 1a). Spouts were connected
 608 to a syringe set on a single-syringe pump (AL-1000, World Precision Instruments) via silicon tubing.
 609 Water delivery from each spout was regulated by an audible pinch valve triggered by TTL signal
 610 which also triggered the syringe pumping.

611

612 **Behavioral task.** A trial began with a pure tone presentation (3 kHz, 1 s) (Start in Fig. 1a).
 613 Rats were required not to lick any spouts from 0.3 s after Start cue offset until an appearance of
 614 another auditory cue (10 kHz, 0.2 s) (Go). If the rats licked any spouts during the delay period, the
 615 trial was immediately aborted. This pseudorandom delay period ranged between 0.7-2.3 s. After the
 616 Go signal onset, the rats were allowed to lick one of either left or right spout within a response
 617 window (5 s). A first lick was judged as a choice response (Choice). When the chosen spout location
 618 corresponded to ongoing reward location, 0.1% saccharin water (15 μ l) was delivered as a reward
 619 after a pseudorandom delay period ranging between 0.3 and 0.7 s (2nd Delay). After an outcome

620 period (4 s, Outcome), a next trial began. On the other hand, when the rats chose no-reward spout,
621 they did not receive any sensory feedback but had an additional time-out of 5 s after the outcome
622 period. After the time-out, a next trial began.

623 At a first training session, the rats learned to associate licking a single spout located at left
624 (or right) side with a reward delivery. Once the rats voluntarily licked the spout to acquire a reward,
625 they learned to associate licking another spout set at the opposite side with a reward acquisition
626 through another spout. At this stage, the second delay period after choice was fixed as zero (no
627 delay). At a second training session, the reward-associated spout location was switched after a
628 bunch of cumulative rewarded trials were observed at each side. As rats reliably showed repetitive
629 rewarded choices rather than random choices, the number of reversal of reward-location increased.
630 Eventually, once the cumulative rewarded trial number reached 10 within each block, the reward-
631 associated spout position was reversed without any feedback like sensory or physical differences in
632 the task. Therefore, rats could not know the block reversal *a priori* without experiencing forthcoming
633 trials. From 13th and 17th sessions, the first and second delay periods were prolonged by extending
634 to 0.7-2.3 s and to 0.3-0.7 s, respectively. After each rat was trained at a training chamber for 19
635 days, they were afterwards habituated to another task chamber, where electrophysiological
636 recording experiments were conducted, through two more training sessions.

637

638 ***Electrophysiological recordings.*** Two daily recording experiments were conducted. Multi-neuron

activity was simultaneously recorded from MFC and DS of left hemisphere with two 32-channels silicon probes consisting of four shanks (0.4 mm shank separation), on which tetrode-like electrode sites were spaced vertically by 0.5 mm (A4x2-tet-7/5mm-500-400-312, NeuroNexus Technologies). Each probe was connected to a custom-made headstage installed on either one of two fine micromanipulators (1760-61; David Kopf Instruments) mounted on a stereotaxic frame (SR-8N, Narishige). A silicon probe was penetrated vertically (depth from pia mater: 1.2 mm) into MFC (+2.4-4.2 mm to Bregma, 1.0-1.4 mm to midline), and the shanks were aligned along midline. Another silicon probe angled posteriorly by 6 deg was inserted into DS through a cranial window (+0.6-1.0 mm to Bregma, 2.1-3.7 mm to midline, 4.0 mm traveling distance), and the shanks were aligned along coronal suture. Multiunit signals were amplified by the headstages before being fed into main amplifiers (Nihon Kodan, 2000 gain) with a band-pass filter (0.5 Hz to 10 kHz). At the second recording session, juxtacellular activity was recorded from the MFC of left hemisphere. A glass microelectrode was prepared by a laser puller (P-2000; Sutter Instrument) and filled with 2-3% Neurobiotin (Vector Laboratories) dissolved in 0.5 M potassium chloride (9-19 MΩ). The electrode was inserted into MFC through the cranial hole for MFC silicon probe with a stereotaxic hydraulic micromanipulator (SM-25C, Narishige). Juxtacellular activity was amplified (final gain 1000) with two amplifiers (IR-283, Cygnus Technologies; EX4-400, Dagan) and filtered (0.3-10 kHz). All neural data was sampled at 20 kHz with two hard-disc recorders (LX-120, TEAC), together with time of task events and licking each spout (left and right). After juxtacellular recording, we tried to electroporate

658 Neurobiotin into the recorded cell with positive current injection (2-14 nA, duration of 0.5s at 1 Hz
659 interval, for 5-15 mins) in order to visualize the recorded cell *post hoc* and to verify the recording
660 location.

661

662 ***Inactivation of the MFC.*** Of 34 rats, 15 reached a performance criterion (>75%) at the first recording
663 session. Of the 15 rats, 10 and 5 were tested without and with injection of GABAa agonist muscimol
664 into bilateral MFC at the second recording session, respectively. For the muscimol group, the rats
665 were placed in the body restrainer and their heads were fixed on the stereotaxic frame. A
666 microinjection syringe (Hamilton) filled with 0.1% muscimol (muscimol hydrobromide, Sigma-
667 Aldrich) dissolved in 0.1M phosphate-buffered saline [55] was set on a microsyringe pump (Legato
668 130, KD Scientific) which was installed on the micromanipulator (1760-61; David Kopf Instruments).
669 An injection needle (25 or 31 gauge) was vertically inserted into the MFC (+3.0 mm to Bregma, 1.0
670 mm to midline, -1.5 mm from pia mater). Muscimol solution was injected via the microsyringe pump
671 at 0.2 µl/min by totally 0.5 ul in each hemisphere. After injections were completed, the needle was
672 left for 5 minutes to allow for diffusion and then slowly retracted. Then, a silicon probe was inserted
673 into DS of left hemisphere as described above. Then, the rats were tested the behavioral task 1 hour
674 after end of muscimol infusion. For control group, multi-neuron and juxtacellular recordings were
675 conducted.

676

Histology. Animals were deeply anaesthetized with Urethane (2-3 g/kg, i.p.) and then perfused intracardially with 0.9% chilled saline followed by 4% paraformaldehyde (PFA) dissolved in 0.1 M phosphate buffer (PB). The fixed brain was stored in 4% PFA overnight and then stored in 30% sucrose solution in 0.1 M PB over 2 weeks. Postfixed brains were frozen and coronally sliced into 50- μ m thick serial sections with a microtome Cryostat (HM5000M, Microm). The brain sections were stored in 0.1 M PB at 4°C overnight. The brain sections were subject for immunostaining to detect FG and Neurobiotin (Nb). For fluorescent visualizations of FG labelled neurons and Nb-loaded neuron, the brain sections were incubated with a rabbit antibody of FG (AB153, 1:3000 dilution, Millipore) at 4°C overnight, followed by incubation with a goat anti-rabbit antibody of FG conjugated with Alexa-594 (A11012, 1:500 dilution, Invitrogen) together with antibody Strept Avidine Alexa-488 (A11012, 1:250 dilution, Invitrogen) of Nb over 2 hours. The slices were washed with PB and mount on a slide glass and dried in a shaded box overnight. Fluorescent images were imaged with a fluorescence microscopy (Olympus, AX70) in order to check if FG labeled neurons were observed around the silicon probe recording locations in MFC and near the FG injection site in DS as well as if the Nb-loaded neurons were observed in MFC. After the fluorescent imaging, the brain sections were re-stained the Nb-loaded neuron using the avidin-biotin-horseradish peroxidase complex (Vectastain Elite ABC; Vector Laboratories, 1:200 dilution) with diaminobenzidine and nickel as demonstrated previously [23,54]. Finally, the slices were counterstained with Neural Red Nissl, and observed with a microscopy to check the tracks (scars) due to probe insertion. We judged the

696 recording locations in MFC and DS and the AP coordinate of brain sections according to the rat brain
697 atlas [56].

698
699 **Data analysis.** All of behavioral and neuronal data were analyzed by custom-written MATLAB
700 scripts (The MathWorks, Inc.). We analyzed behavioral data on 20 rats performing the task at high
701 reward acquisition probability (>0.75) through 30 recording sessions.

702
703 *Logistic regression analysis.* We estimated effects of past outcome and past choice on current choice
704 (repeat or switch) by the following regression model [14].

$$705 \quad \log\left(\frac{p_r}{p_s}\right) = \sum_{k=1}^{10} \beta_{Rk} R(i-k) + \sum_{k=1}^{10} \beta_{Ck} (C_L(i-k) - C_R(i-k)) + \beta_0$$

706 where p_r (or p_s) is the probability of repeat choice (or switch choice) in the i th trial. The variables $R(i)$
707 and $C_L(i)$ (or $C_R(i)$) indicate presence of reward delivery (0 or 1) and left (or right) choice (0 or 1) in
708 the i th trial, respectively. The coefficient β_{Rk} and β_{Ck} represent the effect of past rewards and choices,
709 respectively. β_0 is an intercept. These coefficients were calculated at each session. We checked if the
710 coefficient was significantly different from zero by t-test with Bonferroni correction (Fig. 1d).

711
712 *Spike sorting, clustering, and refining.* Spike event from multiunit (or juxtacellular) activity was
713 isolated by a custom-made semi-automatic spike-sorting program EToS [12 (or 5) feature dimensions
714 for 4 (or 1) channels; high-pass filter, 300 Hz; time-resolution, 20 kHz; spike-detection interval, >0.5

ms) [57,58]. The sorted spike clusters were combined, divided, and discarded manually to refine single-neuron clusters by Klusters [59]. To avoid overlapping of detection of same units recorded among spatially distinct tetrodes, we checked cross-correlations of spike times among isolated units across all of tetrodes. If there was a high correlation peak only at zero time between a pair of units, one of the units was excluded from further analyses because those spikes which originated from the same neuron were presumably recorded through different tetrodes.

For comparison of neural activity between MFC and DS, we analyzed behavioral and electrophysiological data (isolated cells from multi-neuron activity) in 12 sessions from 10 rats: experiment-id (reward acquisition probability, the number of well isolated cells), R982-r1 ($p = 0.828$, MFC/DS = 45/43), R983-r1 ($p = 0.837$, MFC/DS = 35/55), R983-r2 ($p = 0.834$, MFC/DS = 58/48), R985-r1 ($p = 0.853$, MFC/DS = 30/60), R985-r2 ($p = 0.867$, MFC/DS = 66/26), R986-r1 ($p = 0.813$, MFC/DS = 16/20), R991-r1 ($p = 0.806$, MFC/DS = 46/31), R1000-r1 ($p = 0.784$, MFC/DS = 12/32), R1004-r1 ($p = 0.792$, MFC/DS = 20/51), R1005-r1 ($p = 0.794$, MFC/DS = 65/64), R1009-r1 ($p = 0.774$, MFC/DS = 35/33), R1012-r1 ($p = 0.755$, MFC/DS = 40/26). To examine effects of muscimol injection in MFC on neuronal activation of DS cells, we analyzed data from 130 and 77 DS cells of the 3 rats in muscimol group at the first and second recording sessions, respectively.

Event-related neuronal activity. Task-event-related neuronal activity was examined on well isolated units. To make peri-event time histograms (PETHs), we calculated mean and SEM of instantaneous

firing rate (FR) in a 20-ms bin around task events (Start cue onset, Go tone onset, Choice response, and Reward delivery) and the PETHs were smoothed with a Gauss filter (SD = 40 ms). To compare firing pattern across neurons and recording sessions, the PETHs were transformed the FR into z-score by using the mean and SD of control activity, which were FRs in 200-ms at pseudorandom times (1000 time points) in each recording session. Then, the z-scores were normalized by the absolute peak value (Fig. 3e).

Neuronal selectivity for choice and outcome. Statistical significance of neuronal modulation by choice and outcome was analyzed by a two-way ANOVA with choice (left and right) and outcome (reward and no-reward) factors ($p < 0.05$) for each FR in 0.2-s wide sliding window by 20 ms increment around choice response (± 3 s) (Fig. 3f). To quantify the preference of selectivity for choice and outcome, we used a receiver-operating characteristic (ROC) analysis for neurons showing the statistical significance by two-way ANOVA at each window [60]. The proportion of trials for a condition (left choice or rewarded trials) was plotted against that of trials for another condition, (right choice or unrewarded) while changing a criterion for the FR. If the area under the plotted ROC curve (AROC) was 0.5 (0 or 1), the two conditions are indistinguishable (or completely different, respectively). If AROC was larger (smaller) than 0.5, FR is larger in right choice (left choice) or unrewarded outcome (rewarded outcome) than under the opposite condition. To compare the selectivity for choice and outcome modulated activity between MFC and DS, the AROC was

transformed as a selectivity index. Selectivity index = $| \text{AROC} - 0.5 | + 0.5$ (Fig. 3g).

Analysis of neural trajectory separation. All analyses for population activity were separately performed on data sampled at individual recording session. Fisher's linear discriminant [35] was used to find the degree of discrimination between neural trajectories classified into two conditions (e.g., Left and Right choices). To define a Ch-hyperplane, we constructed the distributions of the corresponding $\mathbf{r}(t)$, which is N -dimensional population FR vector (N corresponds to the number of population of neurons), separately for Left and Right choice trials in each 0.2-s sliding window by 50 ms increment time around choice response (± 3 s). Our task is to find a hyperplane, or equivalently the normal vector $\mathbf{w}$ of this hyperplane, that best divides the two distributions projected onto the direction of $\mathbf{w}$. If the two distributions have the means and covariance matrices $\mathbf{m}_1$, Σ_1 and $\mathbf{m}_2$, Σ_2 , respectively, we can obtain $\mathbf{w}$ by maximizing the ratio of the between-class variance to the within-class variance,

$$S = (\mathbf{w}^T \mathbf{m}_1 - \mathbf{w}^T \mathbf{m}_2)^2 / \mathbf{w}^T (\Sigma_1 + \Sigma_2) \mathbf{w},$$

where $\mathbf{w}^T \mathbf{m}_1$ and $\mathbf{w}^T \mathbf{m}_2$ are the means of the two projected distributions and $\mathbf{w}^T \Sigma_1 \mathbf{w}$ and $\mathbf{w}^T \Sigma_2 \mathbf{w}$ are their variances. We can show that S is maximized if $\mathbf{w} \propto (\Sigma_1 + \Sigma_2)^{-1} (\mathbf{m}_2 - \mathbf{m}_1)$. The coefficient of proportion can be determined by the normalization condition: $|\mathbf{w}| = 1$.

Thus, we can calculate the discrimination function between Left and Right choice trials as

$$S = (r_L - r_R)^2 / (\sigma_L^2 + \sigma_R^2) \text{ from the means } (r_L, r_R) \text{ and standard deviations } (\sigma_L, \sigma_R) \text{ of the trial-by-}$$

trial firing rates in these trials. The discrimination degree between the two clusters is defined as d'
 $= S$.

The $\mathbf{r}(t)$ in each 0.2-s wide window (0.2 s increment time) around choice response (± 3 s) was projected onto the axis orthogonal to the hyperplane by calculating inner product of $\mathbf{r}(t)$ and the normal vector $\mathbf{w}_{\text{choice}}$, which yielded maximal degree of discrimination (d'). For instance, population rate at time t vectors was projected onto the axis orthogonal to “Ch-hyperplane”, $\mathbf{v}_t = \langle \mathbf{r}(t), \mathbf{w}_{\text{choice}} \rangle$. In this case, we termed the projected rate vector “Ch-projected trajectory”. We performed the same procedure to find “Ot-hyperplane” ($\mathbf{w}_{\text{outcome}}$) by classifying trials into rewarded and unrewarded conditions, and to obtain the projected rate vector “Ot-projected trajectory”.

Mutual information. To quantify the difference in Ch-projected trajectory (or Ot-projected trajectory), we calculated the mutual information [61] between choices (or outcomes) and the projected population rate vectors $\mathbf{v}_t$:

$$I(C; V_t) = H(C) - H(C | V_t) = -\sum_c p(c) \log p(c) + \left\langle \sum_c p(c | \mathbf{v}_t) \log_2 p(c | \mathbf{v}_t) \right\rangle_{\mathbf{v}},$$

where C and V_t denote the set of choices (or outcome) c and the set of projected population activity $\mathbf{v}_t$, respectively, and $p(c)$ is the choice probability (or outcome probability), $p(c | \mathbf{v}_t)$ is the conditional probability of choice (or outcome) c when $\mathbf{v}_t$ is observed at time t , and the parenthesis means an averaging over the values of $\mathbf{v}_t$. The mutual information was calculated at each time window (bin = 0.2 s) around choice response (± 3 s), and was normalized by the total entropy of choice $H(C)$ (or total

791 entropy of outcome) to provide the percentage of information as demonstrated before [12].

792

793 *Pairwise correlation analysis of population activity.* A Pearson correlation was performed
 794 for Ch-projected trajectory and Ot-projected trajectory between MFC and DS at each time window
 795 (bin = 0.2 s) around choice response (± 3 s). Because the correlations are sensitive to the number of
 796 trials, we randomly selected 431 trials, which was the minimum total number of trials among 9
 797 sessions, and calculated correlation coefficient so that the correlation coefficient was calculated with
 798 the same total number of trials sampled. We repeated this procedure 10 times to obtain mean
 799 correlation coefficients (Fig. 4e and Fig. S4). To compare the correlations of trajectories between
 800 rewarded and unrewarded trials, we calculated correlation coefficient by randomly selected 63 trials,
 801 which was the minimum total number of unrewarded trials among 9 sessions, for each outcome
 802 condition (only rewarded trials, only unrewarded trials, trials without distinction). We repeated this
 803 procedure 10 times to obtain mean correlation coefficients (Fig. S5b).

804

805 *Cross-correlation analysis of spike times.* We calculated cross-correlograms (CCGs) under 3
 806 conditions (left choice trials, right choice trials, and rewarded trials) for all MFC-DS cell-pairs. We
 807 used spike times in a 5-s wide window centered on a specific time at which a maximum absolute
 808 correlation of neural trajectories was yielded at Ch-hyperplane (or Ot-hyperplane) referred as
 809 window Ch (or window Ot) (Fig. 4e and Fig. S4). As a control, we calculated task-event irrelevant

810 CCGs by randomly sampling X time points within the same recording session for the same cell-pair.

811 X corresponds to the number of trials used for calculation of task-event relevant CCGs under each

812 condition. For example, the CCG between MFC and DS cells was calculated by counting spikes of

813 MFC cell in a window ranging from -0.1 to +0.1 s relative to each spike time of DS cell (bin = 1 ms).

814 Spike synchrony was detected by comparing the peak value in CCGs within ± 20 ms with the

815 mean+4SD of values in a baseline window (-100 to -80 ms and +80 to +100 ms) under each condition.

816 If the peak value was larger than this criterion, we judged that the neuron pair showed spike

817 synchrony. We did not assess the spike synchrony if the mean of baseline values was less than 1 or

818 the SD was 0 due to scarce spiking. We did not look into negative spike correlation. For each cell-pair

819 showing significant CCG peak values, we also calculated CCGs by shuffling trials between MFC and

820 DS cells (Fig. S6). To compare peak value of CCGs, each CCG was z-scored by the mean and SD of the

821 baseline values (Fig. S6 and 5b).