Acquisition of auditory discrimination mediated by different processes through two distinct circuits linked to the lateral striatum

Susumu Setogawa[1,2,3,4,5,*], Takashi Okauchi[4], Di Hu[4], Yasuhiro Wada[6], Keigo Hikishima[7], Hirotaka Onoe[8], Kayo Nishizawa[3], Nobuyuki Sakayori[9], Hiroyuki Miyawaki[1,2], Takuma Kitanishi[2], Kenji Mizuseki[1,2], Yilong Cui[4], Kazuto Kobayashi[3*]

[1]Department of Physiology, Graduate School of Medicine, Osaka Metropolitan University, Osaka 545-8585, Japan

[2]Department of Physiology, Osaka City University Graduate School of Medicine, Osaka 545-8585, Japan

[3]Department of Molecular Genetics, Institute of Biomedical Sciences, Fukushima Medical University, Fukushima 960-1295, Japan

[4]Laboratory for Biofunction Dynamics Imaging, RIKEN Center for Biosystems Dynamics Research, Kobe 650-0047, Japan

[5]Japan Society for the Promotion of Science, Tokyo 102-0083, Japan

[6]Laboratory for Pathophysiological and Health Science, RIKEN Center for Biosystems Dynamics Research, Kobe 650-0047, Japan

[7]Medical Devices Research Group, Health and Medical Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8564, Japan

[8]Human Brain Research Center, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan

[9]Department of Physiology and Oral Physiology, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima 734-8553, Japan

28 *Address for correspondence:

29 Kazuto Kobayashi, Ph.D.

30 E-mail: kazuto@fmu.ac.jp

31 Department of Molecular Genetics, Institute of Biomedical Sciences, Fukushima Medical

32 University, Fukushima 960-1295, Japan.

33 Tel. +81-24-547-1667

34

35 Susumu Setogawa, Ph.D.

36 E-mail: setogawa@omu.ac.jp

37 Department of Physiology, Graduate School of Medicine, Osaka Metropolitan University, Osaka

38 545-8585, Japan

39 Tel. +81-6-6645-3717

40

Abstract

The striatum, the central hub of cortico-basal ganglia loops, contains functionally heterogeneous subregions distinguished by the topographic patterns of structural connectivity. These subregions mediate various processes of procedural learning. However, it remains unclear when and how striatal subregions engage in the acquisition of sensory stimulus-based decision-making. A neuroimaging of regional brain activity shows that the anterior dorsolateral striatum (aDLS) and posterior ventrolateral striatum (pVLS) in rats are activated in a different temporal pattern during the acquisition phase of auditory discrimination. Chronic and transient pharmacologic manipulations show that the aDLS promotes the behavioral strategy driven by the stimulus-response association while suppressing that by the response-outcome association, and that the pVLS contributes to forming and maintaining the stimulus-response strategy. Electrophysiological recording indicates that subpopulations of aDLS neurons predominantly represent the outcome of specific behaviors at the initial period of discrimination learning, and that pVLS subpopulations encode the beginning and ending of each behavior according to the progress of learning. In addition, other subpopulations of striatal neurons indicate sustained activation after obtaining reward with distinct patterns reflecting the stimulus-response associations. Our findings demonstrate that aDLS and pVLS neurons integrate the new learning of auditory discrimination in spatiotemporally and functionally different manners.

Introduction

The neural network connecting the cerebral cortex, basal ganglia, and thalamus forms functionally heterogeneous loops distinguished by the topographic patterns of connectivity[1,2], and is implicated in various learning processes by the integration of cognitive, sensorimotor, and reward information[3,4]. The striatum is a key node in the cortico-basal ganglia loops, and its segregated subregions are considered to mediate procedural learning, such as motor skills and instrumental behavior, through parallel processing via different loops[5-7]. Previous studies in humans and non-human primates have reported that the neural activity within the striatum shifts from the associative subregions (caudate and anterior putamen) to the sensorimotor subregions (posterior putamen) as motor learning progresses[8-11]. In rodents, an instrumental behavior initiates as goal-directed actions and then transits to stimulus-response habits[12,13], requiring the dorsomedial striatum (DMS; a homologue to the caudate in primates) and the dorsolateral striatum (DLS; a homologue to the putamen in primates), respectively[14,15]. Based on these results, a model of procedural learning has been proposed, in which the functional dominance changes from the associative to sensorimotor subregions in the striatum during the learning phases[16].

In contrast, deviations from the prior model are evident in other behavioral tasks on decision-making based on external sensory cues. For example, the neural activity in the caudate in humans is associated with learning of a visuomotor association task and the execution of learned behavior[17,18]. The tail of the caudate in non-human primates contributes to the automatic, habitual behavior in decision-making task using complex visual stimuli[19]. In rodents, the DLS is involved in the acquisiton and execution of the discrimination task with visual and auditory cues, wheras the DMS madiates the performance of learned behavior[20,21]. Moreover, other striatal subregions, such as the posterior striatum and ventrolateral striatum in rodents are also engaged in the auditory discrimination[22-24]. Based on these results, a hypothesis has arisen that the acquisition of sensory cue-based decision-making requires wide-range spatiotemporal processing within the striatum. However, the neural mechanism by which the functional circuits linked to striatal subregions

change throughout learning processes is not yet fully understood.

To investigate a large-scale striatal reorganization in accordance with the acquisition of external cue-dependent decisions, we conducted a two-alternative auditory discrimination task with rats[25]. Using a small-animal neuroimaging technique, we found that the anterior dorsolateral striatum (aDLS) and posterior ventrolateral striatum (pVLS) are activated to the highest level at the middle and late stages during the acquisition phase of auditory discrimination, respectively. Then, chronic pharmacologic manipulations confirmed that the aDLS and pVLS are necessary for discrimination learning, and the transient manipulations also indicated the function of these subregions in the learning at the corresponding stages. The resultant effects of the manipulations on the behavioral strategy showed that the aDLS promotes the strategy based on the stimulus-response association, while suppressing that based on response-outcome associations. Additionally, the pVLS engages in the formation and maintenance of the stimulus-response association strategy. Electrophysiological data of striatal neurons indicated the presence of aDLS subpopulations, mainly representing the outcome of specific behaviors at the initial period of auditory discrimination and pVLS subpopulations encoding the beginning and ending of behavior in association with the development of the discrimination. There were other subpopulations in the aDLS and pVLS showing long-term increased activity following reward with different patterns of the combination between the stimulus and response. These findings revealed that the two striatal subregions are involved in the acquisition of auditory discrimination through distinct spatiotemporal and functional fashions, challenging the prior model which proposed a transition from the associative to sensorimotor striatal subregions during learning.

Results

Leaning processes of auditory discrimination

We employed a two-alternative auditory discrimination task, in which rats are required to make new associations between a tone instruction cue (2 or 10 kHz) and a response (pressing the left or right

lever)[25], setting the interval between the stimulus presentation and lever insertion to 3 s (Figure 1A). The success rate gradually increased, reaching a plateau around Day 13, which persisted through Day 24 (Figure 1B). Based on changes observed in the success rate, we categorized the learning processes into two phases: the acquisition phase, which lasted until Day 13; and the learned phase, which began after Day 13 and then continued at least until Day 24. In addition, we divided the acquisition phase into three stages depending on the success rate that included the early (< 60%), middle (60–80%), and late (80–100%) stages. We chose Days 2, 6, and 10 at the early, middle, and late stages in the acquisition phase, respectively, as well as Days 13 and 24 in the learned phase for statistical comparison of the success rate. There were significant differences between Days 2 and 6, Days 6 and 10, as well as Days 10 and 13, showing no significant difference between Days 13 and 24 (Figure 1C). In contrast, the response time and omission rate did not vary among the training days (Figures 1D and 1E), and the variation of response bias decreased along with the progress of auditory discrimination (Figure 1F).

Distinct brain activity patterns in striatal subregions during learning processes

To investigate dynamic changes in regional brain activity in the entire striatum during auditory discrimination in the same animals, we conducted the positron emission tomography (PET) for small animals with 2-deoxy-2-[^{18}F]fluoro-D-glucose (^{18}F-FDG), which measures cerebral glucose metabolism correlated to brain activity[26-28] (Figure 2A). Before the auditory discrimination task, we performed the single lever press task, in which a lever pseudorandomly presented on either the left or right side, and the rats learned to press a lever associated with receiving a reward. This task was used as a control that does not require the decision process based on the auditory stimulus. ^{18}F-FDG-PET scanning was carried out on Day 4 of the single lever press task, and then during a series of training sessions including the acquisition (Days 2, 6, and 10) and learned (Day 24) phases (Figure 2B). The average number of lever presses in the single lever press task exceeded 80 times (81.6 ± 1.3, mean ± s.e.m.) on Day 2 and maintained that level until Day 4 (Figure 2C). In the

following auditory discrimination task, the success rate gradually increased along with the training days (Figure 2D), showing a significant difference among sessions in which we performed the ^{18}F-FDG-PET scanning (Figure 2E). The response time and omission rate were consistent among the days, and the variation of response bias became smaller as the task progresses (Figure 2E). These performances were similar to those in the intact rats (compared to Figures 1C–1F), suggesting that the procedure for ^{18}F-FDG-PET scans does not affect the acquisition of discrimination.

We investigated the brain regions, in which ^{18}F-FDG uptake significantly increased or decreased during the learning processes, applying a statistical significance threshold ($p < 0.001$, uncorrected) and an extent threshold, by using a voxel-based statistical parametric analysis[26]. We first analyzed task-related brain activity by comparing ^{18}F-FDG uptake on either day of the discrimination task with that on Day 4 of the single lever press task (summarized in Table S1). In the striatal subregions, there were significant increases of the activity in the unilateral aDLS and the bilateral pVLS on Days 6 and 10, respectively (Figures 2F and 2G). However, no significant change was observed in the DMS during the acquisition phase except a unilateral decrease on Day 24 (Figure 2F). Next, to evaluate the activity related to the progress of learning, ^{18}F-FDG uptake on Days 6, 10, and 24 was compared with that on Day 2. The learning-dependent brain activity was significantly increased in the bilateral aDLS and the unilateral pVLS on Days 6 and 10, respectively (Figure 2H), whereas the activity decreased in the bilateral DMS on Day 24 (Figure 2H). This decrease was also detected when ^{18}F-FDG uptakes were compared between Days 10 and 24 (Figure 2H), suggesting the down-regulation of the DMS activity in the learned phase.

To quantitatively validate changes in the brain activity in striatal subregions, we analyzed the amount of ^{18}F-FDG uptake in the voxel of interest of these subregions (Figure 2I). The ^{18}F-FDG uptake in the bilateral aDLS reached a peak on Day 6, and it decreased during the subsequent days (Figure 2J), whereas the uptake in the bilateral pVLS gradually increased along with the progress of learning, showing the maximal value on Day 10 (Figure 2K). The uptake in the DMS did not alter during the acquisition phase, and it decreased on Day 24 (Figure 2L). These results suggest that the

163 aDLS and pVLS contribute to learning processes in a different temporal pattern during the
164 acquisition phase, whereas the DMS may be engaged to the behavior in the learned phase.
165

166 **Excitotoxic lesion of the aDLS and pVLS, but not the DMS, disrupts the acquisition of**
167 **auditory discrimination**

168 To study whether the changed activity of striatal subregions is involved in learning processes, rats
169 received a bilateral injection of solution containing ibotenic acid (IBO; 8 mg/mL, 0.3 μL/site) or
170 phosphate buffered saline (PBS) into the subregions, and we conducted for the behavioral study.
171 The range of IBO lesion in the subregions was analyzed by immunostaining with a neuronal marker
172 NeuN after the behavioral tests (Figures 3A–3C). The number of single lever presses increased
173 similarly along with the training in both the PBS- and IBO-injected groups into the aDLS, pVLS, or
174 DMS along with the training (Figures 3D–3F). For the aDLS injections, increases in success rate
175 were impaired during the acquisition phase in the IBO group, compared to the PBS group, and the
176 impairments were mainly observed at the middle stage (Figures 3G). For the pVLS injection,
177 success rate impairments were observed in the IBO group showed compared to the PBS group,
178 although the period showing a significant difference was unclear (Figure 3H). The response time in
179 the aDLS-lesioned group did not alter during the discrimination training, whereas that in the pVLS-
180 lesioned group was continuously prolonged (Figure. S1A and S1B), and the omission rate and
181 response bias were similar between the lesioned and the corresponding control groups (Figures
182 S1D/S1G for aDLS lesion and Figures S1E/S1H for pVLS lesion). In contrast, for the DMS
183 injection the success rate in the discrimination task was comparable between the PBS and IBO
184 groups (Figure 3I). The DMS lesion lengthened the response time through the learning (Figure
185 S1C), whereas it did not alter the omission rate and response bias (Figures S1F and S1I). These
186 results indicate that the aDLS and pVLS are necessary for the acquisition of auditory
187 discrimination, and that the aDLS appears to be mainly involved in the learning process at the
188 middle stage. The results also reveal that the DMS contributes to the execution, but not learning, of

189 discriminative behavior.

190 Previous studies have reported that the dorsal striatum is involved in both behavioral
191 strategies based on the stimulus-response association and the response-outcome association and
192 during instrumental learning[12,13]. To assess the impact of striatal lesions on the behavioral strategies,
193 we analyzed the proportions of responses attributed to two strategies in all responses of each
194 session. One is the "win-shift-win" strategy, which is considered to reflect the behavioral strategy
195 based on the stimulus-response association[29,30]. In this strategy, after a correct response in the
196 previous trial, the rats press the opposite lever in the current trial in response to a shift of the
197 instruction cue, resulting in the correct response. Another strategy is the "lose-shift-lose" strategy,
198 which is considered to appear as a consequence of the behavioral strategy based on the response-
199 outcome association[29]. In this strategy, after an error response in the previous trial, the rats press the
200 opposite lever in the current trial, despite a shift of the instruction cue, leading to another error
201 response. In the control group into either of three striatal subregions, the proportions of win-shift-
202 win and lose-shift-lose strategies increased and decreased, respectively, during the acquisition phase
203 (Figures 3J–3L). The changes in the proportion of the two strategies were similar to those in the
204 intact rats (Figure S2). For the aDLS injection, the increases in the win-shift-win proportion were
205 markedly impaired in the IBO group compared to the PBS group with significant differences
206 mainly at the middle stage, and the decrease in the lose-shift-lose proportion was delayed compared
207 to the same group (Figure 3J). For the pVLS-injection, the increasing win-shift-win proportion was
208 also impaired in the IBO group compared to the PBS group, although the change in the lose-shift-
209 lose proportion was comparable between the two groups (Figure 3K). In contrast, for the DMS
210 lesion both win-shift-win and lose-shift-lose proportions in the IBO group did not show significant
211 differences from the corresponding values in their control group (Figure 3L). These results indicate
212 that the aDLS promotes the win-shift-win strategy through stimulus-response association and
213 suppresses the lose-shift-lose strategy rooted in response-outcome association, and that the pVLS
214 selectively facilitates the win-shift-win strategy acquisition. Thus, these two subregions not only

215 play roles in the acquisition of auditory discrimination but also control distinct behavioral strategies
216 during the learning processes.
217 After the completion of discrimination training, we tested locomotion and feeding behavior
218 of rats with aDLS and pVLS lesions. There was no difference in locomotor activity in the open field
219 and the amount of food intake in both lesioned groups (Figures S1J-S1M), suggesting that learning
220 deficits following excitotoxic lesions of the striatal subregions cannot be attributable to
221 abnormalities in these general motor behaviors.
222
223 **Inhibiting aDLS and pVLS activity impairs discriminative behavior at different stages of the**
224 **acquisition phase**
225 Our ^{18}F-FDG-PET imaging study indicated that the brain activity in the aDLS and pVLS was
226 increased in a different temporal pattern during the acquisition phase of auditory discrimination. To
227 examine whether neuronal activity in striatal subregions is linked with the processes at different
228 stages, we temporarily inhibited the activity in each subregion by the treatment with a $GABA_A$
229 receptor agonist muscimol (MUS). In a pilot study, we tested the effect of the bilateral injection of
230 several doses of MUS (0.1, 0.2, and 0.5 μg/μL) into the aDLS or pVLS on the single lever press
231 behavior and confirmed that the administration of the lowest dose of MUS did not affect the
232 frequency of single lever presses and the quantity of food intake (Figures S3A–S3D). In addition, to
233 achieve the stage-specific inactivation of the aDLS or pVLS neurons, rats were given bilateral
234 injections of MUS or saline (SAL) at three different timings during a series of the training. The first
235 injection was conducted on Day 2 to target the early stage. The second and third injections were
236 conducted on the day after the success rate had reached 60% and 80% for the first time through the
237 training, respectively, to target the middle or late stage. After the injections, the rats were used for
238 the behavioral test. Placement locations of the injection cannula into the striatal subregions were
239 assessed by cresyl violet staining after the behavioral tests (Figures 4A and 4B). The success rates
240 on the day before the injection (day N-1) at the early, middle, and late stages were similar between

241 the two groups that received the bilateral injections of MUS or SAL (Figures S3E and S3F). Rats
242 were then administered injections into the aDLS on the next day (Day N), and the success rate in
243 the MUS-injected groups significantly decreased at the middle stage compared to the corresponding
244 SAL-injected groups, but not at the early and late stages (Figure 4C, left). In contrast, the success
245 rate in the MUS groups into the pVLS significantly decreased at the late stage, but not at the early
246 and middle stages, compared with the respective SAL groups (Figure 4D, left). To normalize the
247 variation of learning speed in the individual animals, we calculated the value by subtracting the
248 success rate on Day N-1 from that on Day N in the individual animals. There were significant
249 reductions in the subtracted rate at the middle stage in the MUS-injected group into the aDLS
250 (Figure 4C, right) and at the late stage in the MUS-injected group into the pVLS (Figure 4D, right).
251 These results indicate that the aDLS and pVLS mainly function at the middle and late stages,
252 respectively, in the acquisition phase of auditory discrimination, supporting the results obtained
253 from our imaging study.

254 Next, we analyzed the influence of temporal striatal inactivation on the behavioral strategy
255 at the middle and late stages of the acquisition phase. We calculated the proportions of the win-
256 shift-win and lose-shift-lose strategies, and subtracted the values of these proportions on Day N-1
257 from those on Day N. The proportion of both strategies on Day N-1 were comparable between the
258 MUS- and SAL-injected groups (Figures S3G and S3H). In the aDLS, the subtracted proportion of
259 the lose-shift-lose strategy at the middle stage significantly increased in the MUS group as
260 compared to the SAL group, although the subtracted proportion of the win-shift-win strategy tended
261 to decrease in the MUS group as compared to the control group (Figure 4E), and the subtracted
262 proportions of the win-shift-win or lose-shift-lose strategy at the late stage was similar between
263 both groups (Figure 4F). In the pVLS, the subtracted proportions of the win-shift-win and lose-
264 shift-lose strategies at the middle stage were similar between the MUS and SAL groups (Figure
265 4G), and the subtracted win-shift-win proportion at the late stage significantly decreased in the
266 MUS group as compared to the SAL group, although the subtracted lose-shift-lose proportion was

comparable between the two groups (Figure 4H). Our results show that transient inhibition of aDLS activity at the middle stage affects the lose-shift-lose strategy and has a subtle effect on the win-shift-win strategy, supporting that the aDLS mediates the acquisition process of auditory discrimination at the middle stage at least partly through the suppression of the behavior based on the response-outcome association. Although the impact of inhibition on the win-shift-win strategy was subtle, excitotoxic lesion of the aDLS influenced the win-shift-win strategy in addition to the lose-shift-lose strategy (see Figure 3J), suggesting that the aDLS acts to enhance the auditory discrimination by promoting the stimulus-response association, together with the suppression of the response-outcome association. The discrepancy in the extent of changes in the win-shift-win strategy between the two treatments may be generated from the difference in the timing or sustainability of drug inhibition of the striatal subregion. In addition, our results show that pVLS inactivation at the late stage selectively impairs the win-shift-win strategy, supporting contribution of the pVLS to the process via progression of the stimulus-response association.

Neuronal activity in the aDLS and pVLS during the acquisition phase

To examine the firing activity in striatal subregions during the acquisition phase of auditory discrimination, we simultaneously recorded neuronal activity in the aDLS and pVLS at the unilateral (left) side by using a multi-unit recording procedure (Figure 5A). Using immunostaining for tyrosine hydroxylase (TH) after the recordings, we verified the location of electrodes within the striatal subregions (Figure 5B). After the single lever press task, we conducted the auditory discrimination task for the recordings. In correct trials, the delay period (0.5 ± 0.2 s) was inserted between a lever press and a sound indicating the reward to distinguish neuronal activity related to the two events (Figure 5C). The behavioral data in individual rats were divided into three grouped sessions corresponding to the early, middle, and late stages, based on a sigmoid function analysis[31] of the success rate in the task. The success rates in the grouped sessions were 52.7 ± 0.8 %, 61.6 ± 0.7 %, and 85.3 ± 1.7 % (mean ± s.e.m.) for the early, middle, and late stages, respectively, and the

293 rate was significantly increased throughout the three stages (Figure S4A). However, the response
294 time, omission rate, and response bias were similar among the stages (Figures S4B-S4D). Thus,
295 these stages corresponded to the timing window defined in the acquisition phase of the transient
296 MUS inhibition experiment of neuronal activity.

297 We identified 1,062 and 423 well-isolated neurons from the aDLS and pVLS, respectively,
298 across the three stages (Figure S4E) and focused on the major population of striatal neurons,
299 putative medium spiny neurons (MSNs; Figures S4F and S4G). The total numbers of MSNs at the
300 early, middle, and late stages were 295, 364, and 360 neurons for the aDLS, and 87, 110, and 190
301 neurons for pVLS, respectively. Based on the combinatorial pattern of the tone instruction cue and
302 lever press in our discrimination task, we categorized the electrophysiological data into four trial
303 types: (1) high-frequency tone/right lever press (HR) and (2) low-frequency tone/left lever press
304 (LL) as correct responses; and (3) low-frequency tone/right lever press (LR) and (4) high-frequency
305 tone/left lever press (HL) as error responses (see Figure 1A). We identified HR or LL type neurons
306 showing significant changes in firing rate related to the cue onset, choice response, reward sound,
307 or first licking as compared to the baseline firing rate in each of two trials (Mann-Whitney *U* test, p
308 < 0.01). These neurons were further divided into two groups based on increased or decreased
309 activity (Figures S5A and S5B). In the following analyses, we focused on the HR and LL type
310 neurons with increased event-related activity to explore the firing patterns of neurons underlying
311 behavioral processes of auditory discrimination.

312

313 **Subpopulations of aDLS neurons, but not of pVLS neurons, show firing activity related to the**
314 **behavioral outcome**

315 Previous studies have reported that neurons in the dorsal striatum show the outcome-related activity
316 in the instrumental conditioning tasks[24,32]. We thus investigated whether aDLS and pVLS neurons
317 exhibit a response related to the behavioral outcome during the acquisition of auditory
318 discrimination. In the aDLS, the firing activity of reward sound-related HR type neurons showed a

higher extent in the increase during a period, when the reward sound is presented in the correct responses, in the HR trial as compared with the LR, LL, or HL trial throughout the three stages (Figure 5D, top rows). The rate of these neurons also moderately increased during the same period in the LR trial compared to the LL or HL trial across the three stages, although the extent of the increase was lower than that in the HR trial (Figure 5D, top rows). The area under the receiver operating characteristic (auROC) values revealed a significant difference of the HR trial from the LR trial during the reward sound period, whereas the values of HL and LL trials indicated a significant difference in the opposite direction from the LR trial during the same period (Figure 5D, bottom rows). The averaged firing rate during the reward sound period in the HR or LR trial appeared to modestly increase at the middle stage, but no significant difference existed among the stages (Figure 5E). To directly compare the number of neurons showing event-related activity among different stages, we calculated the number of reward sound-HR neurons by matching the trial number at each stage (Figure 5F). The proportion decreased progressively from the early to middle stages, and to the late stage. In addition, the firing rate of reward sound-related LL type neurons exhibited a higher increase in the LL trial and a modest increase in the HL trial while the averaged firing rate during the reward sound period was similar among the three stages. Furthermore, the proportion of reward sound-LL neurons reduced progressively through the stages (Figures S6A–S6C). These results indicate that the aDLS contained neurons related to the outcome of choice, showing a greater increase in the firing activity for the correct responses (HR and LL trials) than the error responses (LR and HL trials), and that the proportion of these subpopulations was gradually reduced through the progress of learning. In contrast, during the reward sound period at all stages, the firing rate and auROC values of reward sound-HR and reward sound-LL neurons in the pVLS revealed no significant differences across the four kinds of trial (Figures S6D and S6E), indicating that the firing activity of pVLS neurons did not appear to be associated with the behavioral outcome.

345 **Subpopulations of aDLS and pVLS neurons show sustained firing activity after the reward in**
346 **different manners**
347 The dorsal striatum is known to demonstrate a response of neuronal activity after the positive
348 outcome in the instrumental behaviors[33,34]. We thus examined whether aDLS and pVLS neurons
349 show the response after obtaining the reward during the acquisition phase of auditory
350 discrimination. In the aDLS, the firing rate of the first licking-related HR type neurons indicated a
351 sustained elevation after the first licking in the HR trial as compared with any of the other three
352 trials across the early, middle, and late stages (Figure 6A, top rows). The auROC values revealed a
353 significant difference of the LR, LL, or HL trial from the HR trial across the three stages (Figure
354 6A, middle rows). In addition, we calculated the averaged firing rate and increased firing period,
355 defined as the total number of time bins above 3 s.d. of the baseline firing rate for 5 s after the first
356 licking in the HR trial. The averaged firing rate was comparable between the early and middle
357 stages, but the rate was significantly lower at the late stage compared to either of the other two
358 stages (Figure 6B). The increased firing period was also comparable between the early and middle
359 stages, showing a significant reduction at the late stage (Figure 6C). The proportions of first licking-
360 HR neurons were similar between the early and late stages, although they showed a slight decrease
361 at the middle stage (Figure 6D). There were few neurons with significant correlations between the
362 firing rate and licking performance (Figure 6E). In addition, the firing rate of the first licking-
363 related LL type neurons exhibited the largest increase in the LL trial at the three stages, showing a
364 tendency to reduce the average firing rate and increased firing period after the first licking across
365 the stages (Figures S7A-S7C). The proportion of first licking-LL neurons during the learning
366 processes displayed similar changes as observed in the first licking-HR neurons (Figure S7D).
367 These results indicate that the aDLS contained neurons showing long-lasting, increased activity
368 after the reward in the specific correct combination of the stimulus and response (HR or LL trial),
369 and that the extent of firing activity and length of firing period in the response appeared to be
370 gradually reduced with no correlation between the proportion of neurons and the progress of the

371 learning stages.

372 In the pVLS, the first licking-HR neurons showed a sustained elevation in the firing rate
373 after receiving a reward in the HR trial compared to that in the LR or HL trial throughout the three
374 stages (Figure 6F, top rows). A similar elevation in the firing rate was also observed in the LL trial
375 throughout the stages (Figure 6F, top rows). The auROC values revealed a significant difference of
376 the LR or HL trial, but not the LL trial, from the HR trial throughout the stages (Figure 6F, middle
377 rows). The average firing rate did not show significant changes among the three stages (Figure 6G),
378 and the increased firing period was also similar across the stages (Figure 6H). The proportion of
379 first licking-HR neurons increased at the late stage compared to the early or middle stage (Figure
380 6I). The firing rate showed almost no significant correlation to the licking performance (Figure 6J).
381 In addition, the first licking-LL neurons exhibited similar firing patterns and changes in the
382 proportion of neurons across the stages (Figures S7F–S7J). Therefore, the pVLS included neurons
383 representing the long-lasting increased activity after the reward in the two kinds of correct
384 responses (HR and LL trials) with the increased number of these neurons at the late stage of
385 learning.

386

387 **Neurons in the pVLS, but not in the aDLS, exhibit firing patterns related to the beginning and**
388 **ending of the behavior**

389 The dorsal striatum plays a key role in the acquisition and execution of behaviors based on the
390 stimulus-response association, and the association is accompanied by a transient activation of
391 neuronal firing immediately after the stimulus presentation or action in trials, which is considered to
392 represent the beginning and ending of a learned behavior[35-37]. We first examined whether the firing
393 activity of cue onset-related HR type or cue onset-related LL type neurons corresponds to the
394 beginning of the behavior in the auditory discrimination task. In the aDLS, the firing rate of the cue
395 onset-HR and cue onset-LL neurons showed a significant difference at some time points in each
396 trial compared to those in any of the other three trials, with a significant difference of auROC

397 values between the trials (Figures S8A and S8C). However, the transiently increased activity did not
398 emerge at the time points corresponding to the beginning of the behavior. In contrast, the cue onset-
399 HR neurons in the pVLS showed a transient increase in the firing rate immediately after the cue
400 onset, and the activity appeared at the middle and late stages (Figure 7A, top). The transient
401 activation was observed not only in the HR trial but also in the other three trials, being consistent
402 with the results of a previous report[38]. The increased firing rate showed a slight but notable
403 difference between the HR and LL trials with a significant difference in the auROC values between
404 the two trials (Figure 7A, low), suggesting that different combinations between the stimulus and
405 response may affect the level of firing activity. Differences in the averaged firing rate between the
406 two kinds of trial was observed at the middle and late stages (Figure 7B). The proportion of cue
407 onset-HR neurons increased at the middle and late stages compared to the early stage (Figure 7C).
408 In addition, the firing rate of the cue onset-LL neurons exhibited patterns similar to those observed
409 in the cue onset-HR neurons (Figure S8E). The data indicate that subpopulations of pVLS neurons
410 show firing activity related to the beginning of the behavior, with a rising number of neurons
411 throughout discrimination learning.

412 Next, we examined the firing activity of choice response-related HR type or choice
413 response-related LL type neurons, corresponding to the ending of a behavior in auditory
414 discrimination. In the aDLS, the firing rate of these neurons did not exhibit the transient increase
415 related to the ending of the behavior, instead, there was a modest increase before the lever press
416 (Figures S8B and S8D). In contrast, the choice response-HR neurons in the pVLS were transiently
417 activated immediately after the lever press in the four kinds of trial at the middle and late stages,
418 and the rate appeared to gradually decline toward the reward sound presentation (Figure 7D). The
419 averaged firing rate was similar between the HR and LL trials though the stages (Figure 7E). The
420 proportion of the choice response-HR neurons increased at the middle and late stages compared to
421 the early stage (Figure 7F). In addition, the rate of choice response-LL neurons displayed similar
422 firing patterns to those observed in the choice response-HR neurons (Figure S8F). Thus,

423 subpopulations of pVLS neurons also represent the activity related to the ending of the behavior in
424 auditory discrimination, increasing the neuron number throughout the learning stages in a similar
425 pattern to neurons relevant to the beginning of the behavior in the task.
426

427 ## Discussion

428 Our findings demonstrate that the aDLS and pVLS are required for the acquisition of auditory
429 discrimination at different stages of learning, and that the DMS is not necessary for the acquisition
430 of discrimination, although it appears to contribute to the execution of actions, in particular to the
431 regulation of response time. These results not only support some recent studies reporting no
432 requirement of the DMS in the procedural learning[20,39,40], but also reveal important roles of two
433 different subregions in the lateral striatum in the external cue-dependent decision-making. We
434 propose that aDLS and pVLS neurons act to integrate the new learning of auditory discrimination in
435 spatiotemporally and functionally distinct manners, showing a remarkable difference from the prior
436 learning model that proposes the functional dominance changes from the DMS to DLS subregions
437 during procedural learning[12-16].

438 Our ^{18}F-FDG-PET imaging study demonstrated that the brain activity in the aDLS reaches
439 a peak at the middle stage of the acquisition phase of auditory discrimination (Figure 2I). The
440 results of excitotoxic lesion experiment of striatal subregions suggest that the aDLS functions to
441 enhance the acquisition of discrimination through the promotion of behavior based on the stimulus-
442 response association and the suppression of behavior based on the response-outcome association
443 (Figure 3J). Transient inhibition experiment of neuronal activity confirmed this aDLS function at
444 the middle stage of the learning, although the drug inhibition tended to reduce the stimulus-response
445 behavior (Figures 4E and 4F). In addition, we found subpopulations of aDLS neurons (reward
446 sound-related neurons) which represent the behavioral outcome and differentiate the correct (HR or
447 LL) and incorrect (LR or HL) choices (Figures 5D and S6A). The distinct responses of these
448 neurons may explain the mechanism by which the aDLS regulates the stimulus-response and

449 response-outcome associations through distinct neural circuits leading to a differential control of the
450 two associations. Alternatively, the responses mainly promote the behavior based on the stimulus-
451 response association, thereby secondarily inhibiting the response-outcome association. Indeed,
452 some previous studies have suggested that an action based on either the stimulus-response
453 association or response-outcome association is generated as a result of competition between
454 different striatal subregions[41]. We also found other subpopulations of aDLS neurons (first licking-
455 related neurons) showing long-lasting increased activity after obtaining reward in response to a
456 specific combination between the stimulus and response (HR or LL trial) (Figures 6A and S7A).
457 This activity may function as the feedback signal to promote the stimulus and response association,
458 resulting in the enhancement of discrimination learning. The proportion of neurons related to the
459 behavioral outcome showed a gradual reduction through the progress of auditory discrimination
460 (Figures 5E and S6C), and the extent of the firing activity and the length of the firing period of
461 neurons showing sustained activation after reward were also decreased at the late stage of learning
462 (Figures 6B and 6C). The learning-dependent changes in the properties of these subpopulations in
463 the aDLS may explain the mechanism underlying the earlier acquisition process of the auditory
464 discrimination.

465 The ^{18}F-FDG-PET imaging indicated that the pVLS activity was gradually elevated
466 through the acquisition phase of auditory discrimination, showing the maximal value at the late
467 stage (Figure 2J). The lesion experiment suggested that the pVLS drives the formation of
468 discrimination learning via the selective progression of the stimulus-response association (Figure
469 3K), and the transient inhibition experiment ascertained the pVLS function at the late stage of
470 learning (Figures 4G and 4H). Previous studies have shown that the stimulus-response association
471 is encoded in the DLS or VLS as the transient neuronal activity representing the beginning and
472 ending of a learned behavior[35-37,42]. A recent study reported that the VLS is causally related to
473 learned performance based on the auditory stimulus, while the DLS shows a weaker relationship to
474 the behavior[22]. In the present study, we found subpopulations of pVLS neurons that represent the

475 beginning and ending of a behavior (cue onset- and choice response-related neurons, respectively)
476 (Figures 7A, 7D, S8E, and S8F), showing an increased number of neurons along with the progress
477 of discrimination learning (Figures 7C and 7F). These data suggest that pVLS neurons contribute to
478 the formation of the stimulus-response association through the learning processes. In addition, we
479 found another subpopulation of pVLS neurons (first licking-related neurons) that are continuously
480 activated after the reward outcome in response to two kinds of correct choices (HR and LL trials)
481 (Figures 6F and S7F), and the proportion of neurons was increased at the late stage (Figures 6I and
482 S7I). This activity may act as a feedback signal to consolidate the stimulus and response
483 association, leading to the maintenance of associative memory at the late stage during the
484 acquisition phase of learning.

485 In the present study, we identified neurons with four kinds of event-related activity in the
486 aDLS and pVLS, and classified them into the HR and LL type neurons in which their firing activity
487 showed a significant increase in the HR and LL trials, respectively, as compared to the baseline
488 (Figures 5–7 and S6-S8). Parts of these two types of neurons with the respective event-related
489 activity overlapped, and the numbers of overlapping neurons varied between striatal subregions
490 (Figures S9A and S9B). In the pVLS, the overlapping neurons for the same event-related activities
491 were 35.1%, 45.2%, and 67.0% in the cue onset, choice response, and first licking, respectively. In
492 contrast, the aDLS had lower overlaps of 12.6% and 22.7% in the reward sound and first licking,
493 respectively. Thus, distinct neurons in the aDLS tend to have properties of a single trial type while
494 the same neurons in the pVLS seem to possess properties of both types. One possible explanation
495 for the diversity in the proportion of overlapping neurons between the two subregions may be the
496 difference of input patterns into striatal neurons or plastic changes in functional connectivity during
497 learning processes. The detailed mechanism by which aDLS and pVLS neurons react to the events
498 in a different manner should be investigated in the future.

499 Anatomically, striatal subregions are known to form heterogeneous cortico-basal ganglia
500 loops in mice[2,43,44]. We examined the neural circuits linked to the aDLS and pVLS in rats by using

bidirectional tracers. The results showed that two striatal subregions receive innervations from different subregions in the cerebral cortex, intralaminar thalamus, and substantia nigra pars compacta, and send projections to different subregions in the basal ganglia nuclei, including the globus pallidus, entopeduncular nucleus, and substantia nigra pars reticulata (Figure S10), confirming that the aDLS and pVLS form the heterogenous cortico-basal ganglia loops in the rat brain. These data suggest that the aDLS and pVLS mediate the acquisition of auditory discrimination via the parallel loops, and cooperate to achieve neural processing related to the dual regulation of the stimulus-response and response-outcome associations. Further studies are needed to elucidate the mechanism by which the two kinds of loops are processed separately and integrate the new learning of discrimination.

Limitations

In the ^{18}F-FDG PET neuroimaging study, the brain activity in the aDLS reached a peak at the middle stage of the acquisition phase of auditory discrimination (Figure 2I). In the multi-unit electrophysiological recording experiment, the firing activity of aDLS neuron subpopulations related to the behavioral outcome showed no significant differences among the three stages (Figure 5E), while the proportions of these subpopulations gradually reduced through the progress of learning stages (Figure 5F). The extent of firing activity and length of firing period of other subpopulations showing sustained activation after the reward appeared to show learning-dependent decrease (Figure 6B and 6C), although the proportions of these subpopulations indicated no correlation to the progress of the learning (Figure 6D). The patterns of the temporal changes in the brain activity of the striatal subregions during the learning stages did not match completely the time variation in the property and proportion of specific event-related neurons. In our electrophysiological analysis, we identified well-isolated neurons from striatal subregions across the behavioral task, focusing on putative MSNs (Figures S4E–S4G). Based on the combinatorial pattern of the tone instruction cue and lever press, we categorized the electrophysiological data into

527 the four trials, including the HR, LL. LR, and HL. We identified HR or LL type neurons showing
528 significant changes in the firing rate related to specific events compared to the baseline firing rate.
529 These neurons were further divided into two groups with increased or decreased activity relative to
530 the baseline firing (Figures S5A and S5B). In the present study, we focused on event-related
531 neurons with the increased activity. Because the analysis was limited to neuronal subpopulations
532 related to specific events with the increased activity, we consider that it is difficult to explain fully
533 dynamic shifts in the brain activity of the striatal subregions dependent on the progress of learning
534 by the time variation of firing activity of individual event-related neurons. The activity of other
535 subpopulations in the striatum may be involved in shifts in brain activity during the learning
536 processes. In addition, recent studies have reported that the activity of glial cells influences the
537 uptake of ^{18}F-FDG[28] and that these cells regulate spike timing-dependent plasticity[45]. Changes in
538 glial cellular activity, through the control of synaptic plasticity, may partly contribute to the pattern
539 formation of learning-dependent shift of brain activity.
540

541 ## Materials and Methods

542 **Animals**

543 Animal care and handling procedures were carried out in accordance with the guidelines established
544 by RIKEN Center for Biosystems Dynamics Research, Fukushima Medical University, and Osaka
545 City University. All procedures were approved by their Institutional Animal Care and Use
546 Committees. All efforts were made to minimize the number of animals used and their suffering
547 throughout the course of the experiments. Male Long-Evans rats (8–13 weeks old, Institute for
548 Animal Reproduction (Ibaraki, Japan)) were used for the present study. The rats were maintained at
549 22 ± 2 °C and 60% humidity in a 12-h light/12-h dark cycle, and food and water were continuously
550 available unless stated otherwise.
551

552 **Surgery**

553 The rats underwent a stereotaxic surgery under isoflurane anesthesia (1–5%). For excitotoxic lesion
554 of striatal neurons, IBO was dissolved in PBS (0.8 mg/mL, Wako Pure Chemical), and IBO solution
555 (0.3 μL per site) was injected bilaterally with the coordinates (mm) from bregma and dura for the
556 aDLS (AP +1.6, ML ±3.3, DV -3.9), pVLS (AP -1.6, ML ±4.6, and DV -5.6), or DMS (AP -1.6, ML
557 ±1.4, and DV -4.8) according to the atlas of rat brain[46]. Injection was performed through a glass
558 pipette (diameter: 60 μm) at a constant velocity of 0.1 mL/min with the microinfusion pump
559 (EICOM). For transient inactivation of striatal neurons, two guide cannulae (length: 10 mm, outer
560 diameter: 26 gauge) were bilaterally placed on the skull and fixed using four screws and dental
561 cement. The chip of the cannula was placed at 1 mm above the target coordinates. A solution
562 containing MUS (Sigma-Aldrich) in saline (0.1 μg/μL, 0.25 μL per site) was injected into the target
563 coordinates through the internal cannula (outer diameter: 30 gauge) at a constant velocity of 0.25
564 μL/min. For anterograde and retrograde tracing, a solution containing cholera toxin subunit B (CTb;
565 1.0 mg/mL) conjugated to Alexa 488 or Alexa 555 (Invitrogen) as a bidirectional tracer was injected
566 into the aDLS and pVLS (0.3 μL/site), respectively. The coordinates for the injection into striatal
567 subregions were the same as described above. For the electrophysiological experiments, surgery
568 was performed as described previously[47]. A 64-channel silicon probe (Buzsaki64spl; NeuroNexus)
569 attaching to a 3D-printed microdrive with a movable screw was chronically implanted with the
570 coordinates (mm) from the bregma and dura into the left aDLS (AP +1.6, ML -3.3, and DV -2.5)
571 and pVLS (AP -1.6, ML -4.6, and DV -3.6). A faraday cage composed of copper mesh was secured
572 on the skull with dental cement to reduce electrical noise and avoid damage to the implants. The
573 rats were allowed to recover for 1 week after the surgeries.
574

575 **Behavioral analysis**

576 Behavioral tasks were conducted in operant conditioning chambers (30.5 × 24.1 × 29.2 cm; ENV-
577 007CT; MED Associates) equipped with two retractable levers on both sides of a reward port in the
578 front panel (ENV-203M-45, ENV-652AM, and ENV-201A; MED Associates). A multi-tone

generator (ENV-223; MED Associates) and room light were mounted on the top of the center and corner panels on the rear side, respectively. For the electrophysical experiment, a 3D printed reward port was used to detect the licking and to prevent contact with the head attachment and front panel. The behavioral experiment in intact rats was performed using grain pellets (F0021, Bioserv) as a food reward. To prevent spillover of radioactivity into the temporal cortex as a consequence of excessive ^{18}F-FDG accumulation into the muscles by their activity to eat the pellets[48,49], the PET imaging experiments were carried out using 0.02% saccharine solution (B0131, Tokyo Chemical Industry) as a liquid reward. The electrophysiological experiment was also carried out using the saccharine solution. The pharmacological experiment was performed using the pellets.

The rats were subjected to dietary restrictions of 12 g/day (CE-2, CLEA Japan) and water access was limited to 6 to 9 h in accordance with the reward used in the experiments for 3 days before the beginning of the training schedule, and the body weights of the rats were maintained at ~85% of their normal weights. The training schedule consisted of four continuous steps. In the first step (magazine training), the rats were habituated to the operant conditioning chambers, and the reward was presented with the click every 20 s. Each daily session was started by illumination of the room light, and continued for 30 min. The sessions were conducted for 3 consecutive days. In the second step (shaping), only one lever on the left or right side was inserted. When the rats pressed the lever, they earned a reward accompanied by a click. Each daily session was started by illumination of the light, and lasted 30 min or until the criteria, in which the rats received 90 rewards. When the number of lever presses on either side reached the criteria, the lever was inserted in the opposite side on the next day. When the rats achieved the criteria on both sides for 2 or 3 days, they moved to the following step. In the third step (single lever press task), each trial was started by illumination of the light and insertion of one lever on either the left or right side in a pseudorandom manner with equal frequencies on both sides for every set of four trials. When the rats pressed the lever, the reward was presented with the sound, and the lever was immediately retracted. The room light turned off 4 s after the lever press. When they did not press the lever

605 within 10 s after the start, the trial was terminated when the light turned off. The trial was repeated
606 every 20 s. Each daily session continued for 30 min, and the sessions were conducted for 4
607 consecutive days. The number of lever presses in each session was counted. In the final step (two-
608 alternative auditory discrimination task), each trial was started by the presentation of a high (10
609 kHz) or low (2 kHz) frequency tone as an instruction cue in a pseudorandom manner. After three
610 seconds, the room light was illuminated, and two retractable levers were inserted at the same time.
611 The rats were required to press the left or right levers in response to the low or high tone cues,
612 respectively. When the rats pressed the correct lever, the reward was presented with the sound, and
613 the levers were immediately retracted. Correct response allowed the rats to access the reward: the
614 retractable spout delivering water was inserted for 1 s for the PET imaging experiment; a pellet was
615 delivered by a food dispenser for the pharmacological experiment; and three drops of water were
616 delivered through a solenoid valve for the electrophysiological experiment. The room light turned
617 off 4 s after the correct response. On the other hand, error response resulted in the retraction of
618 levers and the turning off of the room light. No response within 10 s after the lever insertion was
619 recorded as an omission response. The success rate in each session was calculated by dividing the
620 number of correct response trials with the number of lever press trials minus the number of
621 omission trials. The trial was repeated every 20 s. Each daily session continued for 60 min, and the
622 sessions were conducted for 14 or 24 days according to the experiments.

623 To assess the behavioral strategies in the auditory discrimination task, a trial-by-trial
624 analysis was conducted; the win-shift-win strategy was defined as a correct response, in which,
625 after a correct response during the previous trial, the rats press the opposite lever during the current
626 trial in response to the shift of the tone instruction cue; the lose-shift-lose strategy was defined as an
627 error response, in which, after an error response during the previous trial, the rats press the opposite
628 lever during the current trial despite the shift of the instruction cue. The proportion of both
629 strategies was calculated by dividing the number of each strategy with the number of lever press
630 trials minus the number of omission trials.

631 For the transient inactivation experiment, MUS injections were bilaterally performed 20
632 min before start of the session. For habituation, all rats were injected with saline bilaterally into the
633 corresponding striatal subregions on the last day (Day 4) of the single lever press task. During the
634 auditory discrimination task, the first injection was conducted on Day 2 of the task. The second and
635 third injections were conducted the next day, after the success rate had exceeded 60% and 80%,
636 respectively. Each session continued for 15 min.
637 For the electrophysiological experiments, a delay period (0.5 ± 0.2 s) was added in the
638 auditory discrimination task between the lever press and reward sound to distinguish the motor- and
639 reward-related responses of neurons. When the rats pressed the lever, the lever was immediately
640 retracted. After the delay period, the reward was presented with a sound. The room light turned off
641 4 s after the lever press. The trial was repeated every 20 s, and each daily session continued for 60
642 min.
643
644 **PET imaging analysis**
645 An ^{18}F-FDG-PET scan was performed with a microPET Focus220 (Siemens Medical Solutions,
646 Knoxville) designed for high-resolution imaging of small animals (spatial resolution of 1.4 mm in
647 full width at half maximum at the center of the field of view) as described previously[26]. Serial ^{18}F-
648 FDG-PET scans were performed on the last day of the single lever press task and on Days 2, 6, 10,
649 and 24 of auditory discrimination task. To administer the ^{18}F-FDG under the awake rats, an
650 indwelling catheter was attached to the tail under anesthesia with 5% induction and 2.0–2.5%
651 maintenance of isoflurane (median [interquartile range]: 6 [5–8] min). The recovery period was set
652 to at least 2 hours until the beginning of the behavioral session, with the aim of minimizing the
653 impact of anesthesia as much as possible. A solution of ^{18}F-FDG (ca. 74 MBq/0.4 mL) was
654 intravenously injected just before the beginning of each behavioral session through the indwelling
655 catheter. After the 30-min session, the rats were returned to their home cage, and then a 30-min
656 static PET scan was performed under anesthesia with 2.0–2.5% isoflurane 55-min after the ^{18}F-FDG

injection. During the PET scan, the body temperature of the anesthetized rats was maintained at 37°C using a small animal warmer (BWT-100A, Bio Research Center). The acquired emission data were sorted into a single sinogram and reconstructed by a standard 2D filtered back projection (FBP) with a ramp filter and cutoff frequency at 0.5 cycles per pixel or by a statistical maximum a *posteriori* probability algorithm (MAP), 12 iterations with point spread function effect.

Voxel-based analysis of PET images was conducted according to the method described previously[26]. Briefly, individual MAP-reconstructed FDG (MAP-FDG) images were aligned to a FDG template image of Long Evans rats using PMOD software package (version 3.2, PMOD Technologies), and then the aligned images were transformed into the space of a T2 weighted MRI templates for the Long-Evans rats (https://www.nitrc.org/projects/tpm_rat/). The transformation parameters obtained from individual MAP-FDG images were used to match the individual FBP-reconstructed FDG (FBP-FDG) images with the MRI templates. The voxel size of the template was magnified by 10 to a size similar to that of a human brain. All FBP-FDG images were resampled with a voxel size of 1.2 × 1.2 × 1.2 mm and spatially smoothed through an isotropic Gaussian kernel (6-mm FWHM). Voxel-based statistical analysis was performed by using the Statistical Parametric Mapping (SPM) software (version 8; https://www.fil.ion.ucl.ac.uk/spm/). For visualization of the results, SPM t-statistic values were overlayed on MRI templates to define the voxels showing significantly increased or decreased brain activity, using a high threshold (p = 0.001, uncorrected) and an extent threshold (with the output values from the voxel-based analysis). The voxel of interest (VOI) was determined ass voxels with more than 85% of the maximum T-value within activated areas obtained by voxel-based statistical comparison of FBP-FDG images. The average regional ^{18}F-FDG uptake was calculated on the basis of the values of the VOI in FBP-FDG images.

Electrophysiological recording

Extracellular multi-unit recording was conducted at a sampling rate of 20 kHz from the striatum of

683 freely moving rats by using a 512-ch acquisition board (Open Ephys) via 64-ch recording
684 headstages (C3325; Intan Technologies). The tips of the probe were lowered by turning the screw (>
685 140 μm/day) until it reached the target region (the aDLS and pVLS) during the single lever press
686 task. The tips were sometimes lowered by the same procedure (140 to 280 μm/day) after the daily
687 recording session to improve cell yield.

688 Spike sorting was performed automatically by the Kilosort software
689 (https://github.com/cortex-lab/KiloSort). Clear noise clusters were removed in the first manual
690 sorting using the phy GUI (https://github.com/cortex-lab/phy). The remaining clusters were then re-
691 clustered by using the Klusta Kwik (https://klusta.readthedocs.io/en/latest/), and the clusters were
692 adjusted in the second manual sorting by using the phy GUI again. The sorted clusters satisfying all
693 the following criteria were used for further analysis: L-ratio < 0.05, isolation distance > 15, and ISI
694 index < 0.2[50,51]. Well-isolated single neurons (n = 1,485; aDLS, n = 1,062; pVLS, n = 423) were
695 classified as putative medium spiny neurons (MSNs), fast spiking interneurons (FSIs), tonically
696 active neurons (TANs), and unclassified interneurons (UIs). Putative MSNs and TANs were
697 separated from FSIs and UIs by trough-to-peak duration of mean spike waveform (> 0.6 ms)
698 according to previous studies, with slight modifications[52,53]. Putative FSIs were then separated from
699 UIs by utilizing the proportion of time associated with long inter-spike intervals. This was achieved
700 by summing the inter-spike intervals longer than 2 s and subsequently dividing the resulting sum by
701 the total recording time ($Prop_{ISIs} > 2$ s)[54]. Neurons with values of $Prop_{ISIs} > 2$ s less than 0.4 were
702 classified as putative FSIs, while neurons with $Prop_{ISIs} > 2$ s greater than 0.4 were classified as
703 putative UIs. The putative MSNs and TANs were separated by measuring the post-spike
704 suppression[54]. We measured the length of suppressed time during which the firing rate of a neuron
705 was suppressed following a spike by counting the number of 1-ms bins in its autocorrelation
706 function until the rate of the neuron was equal to or greater than its averaged rate over the 600-ms to
707 900-ms autocorrelation bins. Putative TANs were separated by post-spike suppression of less than
708 40 ms, and the remaining neurons were classified as putative MSNs. Because the numbers of FSIs,

TANs, and UIs were low (79 out of 1,485, 5.3%), they were excluded from further analysis.

To divide the three stages during the acquisition phase, the sigmoidal function was fitted using the following four-parameter logistic equation:

$$f_{(x)} = \frac{1 - \gamma - (1 - \lambda)}{1 + \exp\left(-\beta(x - \alpha)\right)} + \gamma$$

Where α is the inflection point, β is a variable for slope, and λ and γ are parameters related to the lower asymptote (γ) and upper asymptote (λ) of the learning curve. The function was used to evaluate the day when the success rate reached 20% ($t_{20\%}$) or 90% ($t_{90\%}$) of the maximum rate change defined by the difference between the γ and λ. Three stages were defined according to the following criteria: early stage (up to 3 days before the middle stage); middle stage (3 days including before and after the day of $t_{20\%}$); and late stage (3 days including the day of $t_{90\%}$ and before). One rat learned the discrimination faster, and the success rate reached $t_{90\%}$ 2 days after the day of $t_{20\%}$. One day before the day of $t_{90\%}$ was assigned to the late stage. For the following analysis, spike data in each striatal subregion were separated into the early, middle, and late stages (n = 295, 364 and 360, respectively, for the aDLS; and n = 87, 110 and 190, respectively, for the pVLS).

To determine the neurons showing significant changes in firing rate related to the task events, electrophysiological data were grouped into four trial types based on the instruction cue (high or low tone) and the choice (right or left lever); HR, high tone and right lever press; HL, high tone and left lever press; LR, low tone and right lever press; and LL, low tone and left lever press. Neurons with three or more trial data were analyzed. We identified the HR or LL type neurons, detecting significant increases or decreases of the firing rate by applying the Wilcoxon rank-sum test ($p < 0.01$) to trial-by-trial spike number between the baseline period (BL, -1,500 to -500 ms or -1,100 to -500 ms before presentation of the instruction cue) and any of four peri-event periods (cue onset, 0 to +1,000 ms after the onset of the instruction cue; choice response, -700 to +300 ms around the time of lever press; reward sound, -300 to 300 ms around the time of sound onset instructing the reward delivery; and first licking, 0 to +1,000 ms after the time of first licking after

the choice response). Any spikes around the events were averaged in 50-ms bins. The event-related activity was shown as averaged firing rate aligned based on the timing of an event for each trial. The auROC values were calculated by using the spike counts between the HR or LR trials and any of the other three trials. Bins showing significant differences in the auROC data were determined by applying the Wilcoxon rank-sum test ($p < 0.05$) to the distribution of 200 ms bin-by-bin auROC values and 0.5.

To compare the number of neurons showing event-related activity among different stages, the number of HR or LL type neurons were calculated by matching the trial number at each stage. Twenty trials were randomly selected for a neuron, and neurons with less than 20 trials were removed from the analysis. The neurons with event-related activity were identified as described above. The proportion (percentage) of neurons relative to the total number of neurons was calculated at each individual stage. The same analysis was repeated 30 times and the distribution of the proportion of neurons with event-related activity at the stage was obtained. The cumulative probability against the proportion was calculated using cumulative distribution function (cdfplot, MATLAB code from Mathworks).

Licking responses were detected with an infrared photobeam sensor placed in the 3D printed reward port, and generated timestamps for the onset of each spout contact according to the analog voltage signal. The licking responses corresponding to four kinds of trial in the first licking-related neurons were shown as averaged rates (50-ms bins) aligned based on the timing of the choice response. To calculate the correlation coefficient between neuronal activity and licking responses, the spikes and licking for 5 s after the first licking were counted every trial for each first licking-related neuron. Trials in which neurons never responded were removed from both the spike and licking data. The correlation coefficients between the two remaining data were calculated for each neuron (corr, MATLAB code from Mathworks), and neurons showing significant differences were determined by applying the Pearson correlation coefficient ($p < 0.01$).

760 **Histology**

761 After the behavioral experiments, the rats were anesthetized with pentobarbital (50 mg/kg body
762 weight) and isoflurane (3%), then transcardially perfused 4% paraformaldehyde in phosphate-buffer
763 (PB). The brains were post-fixed overnight at 4 °C and cryoprotect with 10, 20, and 30% sucrose in
764 PB. Brain tissues were cut into 30-μm sections by using a cryostat (Leica). For the lesion
765 experiment, cresyl violet staining or immunostaining were performed to define the lesioned areas in
766 the striatum. The sections were stained with a mouse monoclonal anti-NeuN IgG (1:1000 dilution;
767 Millipore, Cat# MAB377, RRID:AB 2298772), and then with a goat anti-mouse IgG conjugated
768 with Alexa Fluor 488 (1:400 dilution; Thermo Fisher Scientific, Cat# A-11029, RRID:AB
769 2534088). The sections were counterstained with 0.3 μM 4',6-diamidino-2-phenylindole (DAPI;
770 Sigma-Aldrich). For transient inactivation experiment, sections were stained by cresyl violet to
771 visualize the positions of guide cannulae. For the electrophysiological experiments, small lesions
772 were made by applying an electrolytic current (3 μA, 10 s) through top and bottom channels of each
773 shank by using the stimulus isolator (A365, World Precision Instrument). Sections through the
774 striatum were stained with a mouse monoclonal anti-TH IgG (1:1,000 dilution; Millipore, Cat#
775 MAB318, RRID:AB 2201528), and then with a goat anti-mouse IgG conjugated with Alexa Fluor
776 488 (1:400 dilution; Thermo Fisher Scientific, Cat# A-11029, RRID:AB 2534088). The recording
777 sites were verified by the staining with DAPI and red fluorescent Nissl stain solution (1:200
778 dilution; Thermo Fisher Scientific, Cat# N21482, RRID:AB 2620170). Fluorescence and
779 immunostained signals were visualized under fluorescence microscopes (BZ-X800; Keyence) and
780 confocal microscope (A1R; Nikon).

781

782 **Statistical analysis**

783 Statistical analysis was performed using MATLAB (MathWorks) and SPSS Statistics (Ver. 27,
784 IBM). Comparisons of data between two groups were performed using the Student's *t*-test (paired),
785 the Wilcoxon signed rank test (paired), and the Wilcoxon rank-sum test (unpaired). Comparisons of

data among three groups were constructed using a one-way repeated measures ANOVA, two-way repeated measures ANOVA, or the Kruskal-Wallis test (without assuming normal distribution and homogeneity of variances). All statistical tests were two-sided. The Bonferroni post-hoc test and Mann–Whitney U-test with Bonferroni adjustment were carried out as needed to compare means. For boxplots, the data points below the lower (25th percentile) -1.5 interquartile range (IQR) and above the upper (75th percentile) quartiles +1.5 IQR were regarded as outliers, and they are not included in the boxplots.

Acknowledgements

We thank Drs. Keiji Ota, Hikaru Yokoyama, and Sho Aoki for instructing the behavioral analysis and Dr. Hideyuki Matsumoto for technical advice and discussion. We thank the Research Support Platform, Osaka Metropolitan University Graduate School of Medicine for microscopic imaging. This work was supported by grants-in-aid for JSPS Fellows (19J01997) and Scientific Research (C) (21K11556) from the Japan Society for the Promotion of Science (S.S.); a grant-in-aid for Scientific Research on Transformative Research Areas (A) Adaptive Circuit Census (21H05244) from the Ministry of Education, Science, Sports, and Culture of Japan (K.K.); the Grant for Casio Science Promotion Foundation, Takeda Science Foundation; and the Nakatomi Foundation (S.S.).

Author contributions

S.S., Y.C., and K.K. designed the study. S.S. performed the small-animal neuroimaging, multi-unit recordings, tracing studies, and pharmacological experiments with technical support from all authors; T.O., D.H., Y.W., H.O., K.H., and Y.C. assisted with the neuroimaging studies. K.N. performed the tracing study, and N.S. assisted with the immunostaining. H.M., T.K., and K.M. assisted with the multi-unit extracellular recordings. S.S. and K.K. wrote the manuscript with assistance from all coauthors.

Declaration of interest

The authors declare no competing interest.

Data availability

All source data underlying the graphs and charts are available in Supplementary Data. All other data are available upon request. The datasets generated and analyzed in this study are deposited on Mendeley data (DOI: 10.17632/ghzr838bdk.1).

References

1. Alexander, G.E. & Crutcher, M.D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* **13**, 266-271 (1990).
2. Foster, N.N., *et al.* The mouse cortico-basal ganglia-thalamic network. *Nature* **598**, 188-194 (2021).
3. Kim, H.F. & Hikosaka, O. Parallel basal ganglia circuits for voluntary and automatic behaviour to reach rewards. *Brain* **138**, 1776-1800 (2015).
4. Yin, H.H. & Knowlton, B.J. The role of the basal ganglia in habit formation. *Nat Rev Neurosci* **7**, 464-476 (2006).
5. Yin, H.H., *et al.* Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* **12**, 333-341 (2009).
6. Kupferschmidt, D.A., Juczewski, K., Cui, G., Johnson, K.A. & Lovinger, D.M. Parallel, but dissociable, processing in discrete corticostriatal inputs encodes skill learning. *Neuron* **96**, 476-489 e5 (2017).
7. Thorn, C.A., Atallah, H., Howe, M. & Graybiel, A.M. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* **66**, 781-795 (2010).
8. Doyon, J., *et al.* Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav Brain Res* **199**, 61-75 (2009).
9. Lehéricy, S., *et al.* Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proc Natl Acad Sci U S A* **102**, 12566-12571 (2005).
10. Miyachi, S., Hikosaka, O. & Lu, X. Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp Brain Res* **146**, 122-126 (2002).
11. Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z. & Rand, M.K. Differential roles of monkey striatum in learning of sequential hand movement. *Exp Brain Res* **115**, 1-5 (1997).
12. Dickinson, A. Actions and habits: the development of behavioural autonomy. *Phil Trans R Soc Lond B* **308**, 67-78 (1985).

13. O'Hare, J.K., *et al.* Pathway-specific striatal substrates for habitual behavior. *Neuron* **89**, 472-479 (2016).
14. Yin, H.H., Knowlton, B.J. & Balleine, B.W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* **19**, 181-189 (2004).
15. Yin, H.H., Ostlund, S.B., Knowlton, B.J. & Balleine, B.W. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22**, 513-523 (2005).
16. Redgrave, P., *et al.* Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat Rev Neurosci* **11**, 760-772 (2010).
17. Toni, I. & Passingham, R.E. Prefrontal-basal ganglia pathways are involved in the learning of arbitrary visuomotor associations: a PET study. *Exp Brain Res* **127**, 19-32 (1999).
18. Seger, C.A. & Cincotta, C.M. The roles of the caudate nucleus in human classification learning. *J Neurosci* **25**, 2941-2951 (2005).
19. Kim, H.F. & Hikosaka, O. Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron* **79**, 1001-1010 (2013).
20. Featherstone, R.E. & McDonald, R.J. Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination task. *Behav Brain Res* **150**, 15-23 (2004).
21. Featherstone, R.E. & McDonald, R.J. Lesions of the dorsolateral striatum impair the acquisition of a simplified stimulus-response dependent conditional discrimination task. *Neuroscience* **136**, 387-395 (2005).
22. Lee, J., Wang, W. & Sabatini, B.L. Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. *Nat Neurosci* **23**, 1388-1398 (2020).
23. Xiong, Q., Znamenskiy, P. & Zador, A.M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348-351 (2015).
24. Guo, L., Weems, J.T., Walker, W.I., Levichev, A. & Jaramillo, S. Choice-selective neurons in the

auditory cortex and in its striatal target encode reward expectation. *J Neurosci* **39**, 3687-3697 (2019).

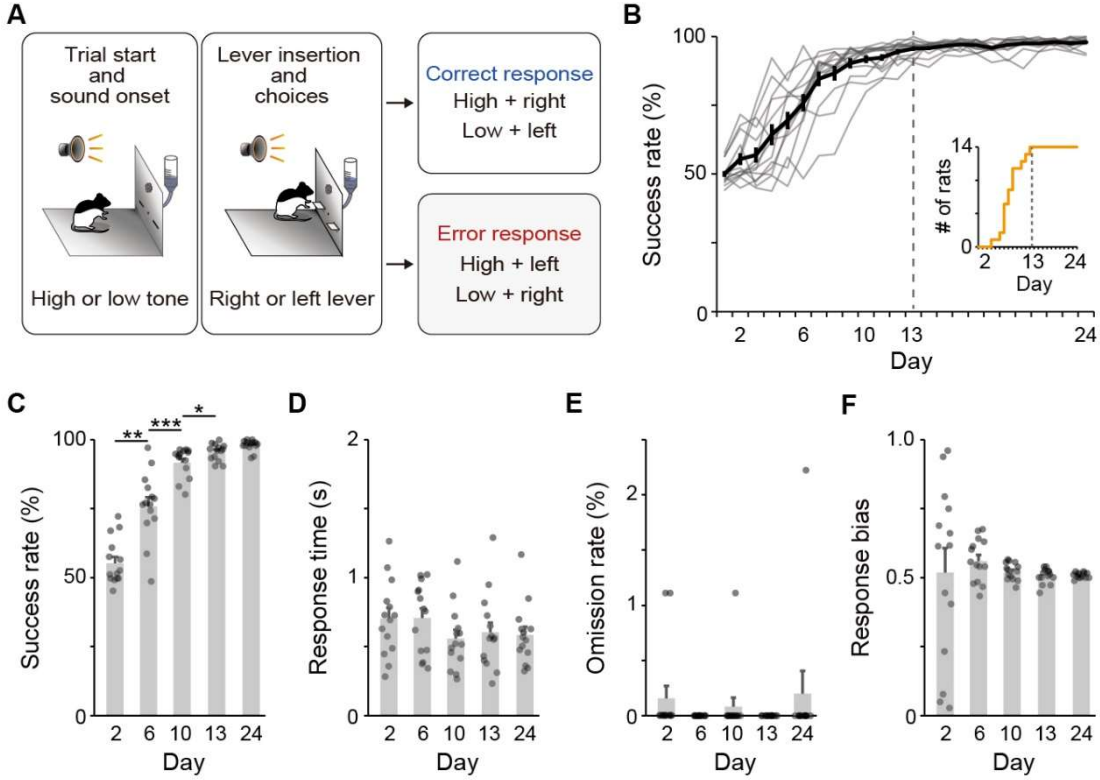
25. Nishizawa, K.*, et al.* Striatal indirect pathway contributes to selection accuracy of learned motor actions. *J Neurosci* **32**, 13421-13432 (2012).
26. Cui, Y.*, et al.* A voxel-based analysis of brain activity in high-order trigeminal pathway in the rat induced by cortical spreading depression. *Neuroimage* **108**, 17-22 (2015).
27. Kornblum, H.I.*, et al.* In vivo imaging of neuronal activation and plasticity in the rat brain by high resolution positron emission tomography (microPET). *Nat Biotechnol* **18**, 655-660 (2000).
28. Zimmer, E.R.*, et al.* [18 F] FDG PET signal is driven by astroglial glutamate transport. *Nat Neurosci* **20**, 393-395 (2017).
29. Bergstrom, H.C.*, et al.* Dorsolateral striatum engagement interferes with early discrimination learning. *Cell Rep* **23**, 2264-2272 (2018).
30. Packard, M.G., Hirsh, R. & White, N.M. Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci* **9**, 1465-1472 (1989).
31. Bathellier, B., Tee, S.P., Hrovat, C. & Rumpel, S. A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. *Proc Natl Acad Sci U S A* **110**, 19950-19955 (2013).
32. Nonomura, S.*, et al.* Monitoring and updating of action selection for goal-directed behavior through the striatal direct and indirect pathways. *Neuron* **99**, 1302-1314. e5 (2018).
33. Williams, Z.M. & Eskandar, E.N. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* **9**, 562-568 (2006).
34. Smith, K.S. & Graybiel, A.M. Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. *J Neurophysiol* **115**, 1487-1498 (2016).
35. Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V. & Graybiel, A.M. Building neural representations of habits. *Science* **286**, 1745-1749 (1999).
36. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z. & Graybiel, A.M. Activity of striatal neurons reflects

898 dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158-1161 (2005).
899 37. Jin, X. & Costa, R.M. Start/stop signals emerge in nigrostriatal circuits during sequence learning.
900 *Nature* **466**, 457-462 (2010).
901 38. Barnes, T.D.*, et al.* Advance cueing produces enhanced action-boundary patterns of spike activity
902 in the sensorimotor striatum. *J Neurophysiol* **105**, 1861-1878 (2011).
903 39. Smith, A.C.W.*, et al.* Opposing roles for striatonigral and striatopallidal neurons in dorsolateral
904 striatum in consolidating new instrumental actions. *Nat Commun* **12**, 5121 (2021).
905 40. Wolff, S.B.E., Ko, R. & Ölveczky, B.P. Distinct roles for motor cortical and thalamic inputs to
906 striatum during motor skill learning and execution. *Sci Adv* **8**, eabk0231 (2022).
907 41. Gremel, C.M.*, et al.* Endocannabinoid modulation of orbitostriatal circuits gates habit formation.
908 *Neuron* **90**, 1312-1324 (2016).
909 42. Chen, Z.*, et al.* Direct and indirect pathway neurons in ventrolateral striatum differentially
910 regulate licking movement and nigral responses. *Cell Rep* **37**, 109847 (2021).
911 43. Hunnicutt, B.J.*, et al.* A comprehensive excitatory input map of the striatum reveals novel
912 functional organization. *eLife* **5**, e19103 (2016).
913 44. Hintiryan, H.*, et al.* The mouse cortico-striatal projectome. *Nat Neurosci* **19**, 1100-1114 (2016).
914 45. Valtcheva, S. & Venance, L. Astrocytes gate Hebbian synaptic plasticity in the striatum. *Nature*
915 *communications* **7**, 13845 (2016).
916 46. Paxinos, G. & Watson, C. The rat brain in stereotaxic coordinates IV Ed (Academic Press, 1998).
917 47. Kitanishi, T., Umaba, R. & Mizuseki, K. Robust information routing by dorsal subiculum
918 neurons. *Sci Adv* **7**, eabf1913 (2021).
919 48. Endepols, H.*, et al.* Effort-based decision making in the rat: an [18F]fluorodeoxyglucose micro
920 positron emission tomography study. *J Neurosci* **30**, 9708-9714 (2010).
921 49. Marx, C.*, et al.* Conflict processing in the rat brain: behavioral analysis and functional μPET
922 imaging using [F]fluorodeoxyglucose. *Front Behav Neurosci* **6**, 4 (2012).
923 50. Fee, M.S., Mitra, P.P. & Kleinfeld, D. Automatic sorting of multiple unit neuronal signals in the

presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* **69**, 175-188 (1996).

51. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A.D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**, 1-11 (2005).
52. Bakhurin, K.I., Mac, V., Golshani, P. & Masmanidis, S.C. Temporal correlations among functionally specialized striatal neural ensembles in reward-conditioned mice. *J Neurophysiol* **115**, 1521-1532 (2016).
53. Kim, N.*, et al.* A striatal interneuron circuit for continuous target pursuit. *Nat Commun* **10**, 2715 (2019).
54. Peters, A.J., Fabre, J.M.J., Steinmetz, N.A., Harris, K.D. & Carandini, M. Striatal activity topographically reflects cortical activity. *Nature* **591**, 420-425 (2021).

934 **Figures and legends**



935 **Figure 1. Behavioral performance of two-alternative auditory discrimination learning.**

936 (A) Schematic diagram of the auditory discrimination task. Each trial started by presentation of a
937 tone instruction cue with the high (10 kHz) or low (2 kHz) frequency. Three seconds later, the room
938 light was illuminated, and two retractable levers were inserted at the same time. The rats were
939 required to press the right and left levers in response to the high and low tones, respectively. (B)
940 Learning curve of the auditory discrimination in intact rats (n = 14). Inset indicates the cumulative
941 curve of number achieved the success rate of more than 90%. (C) Success rate at Days 2, 6, 10, 13,
942 and 24 (one-way repeated ANOVA, $F[1.897,24.659] = 106.706$, $p = 1.1 \times 10^{-12}$, *post hoc*
943 Bonferroni test, Day 2 vs. Day 6, $p = 0.001$, Day 6 vs. Day 10, $p = 3.6 \times 10^{-4}$, Day 10 vs. Day 13, p
944 $= 0.040$, and Day 13 vs. Day 24, $p = 0.080$). (D) Response time (one-way repeated ANOVA,
945 $F[2.203,28.642] = 2.386$, $p = 0.105$). (E) Omission rate ($F[4,52] = 0.699$, $p = 0.596$). (F) Response
946 bias (one-way repeated ANOVA, $F[1.129,14.674] = 0.250$, $p = 0.653$). Data are indicated as the
947 mean ± s.e.m., and individual data are overlaid. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

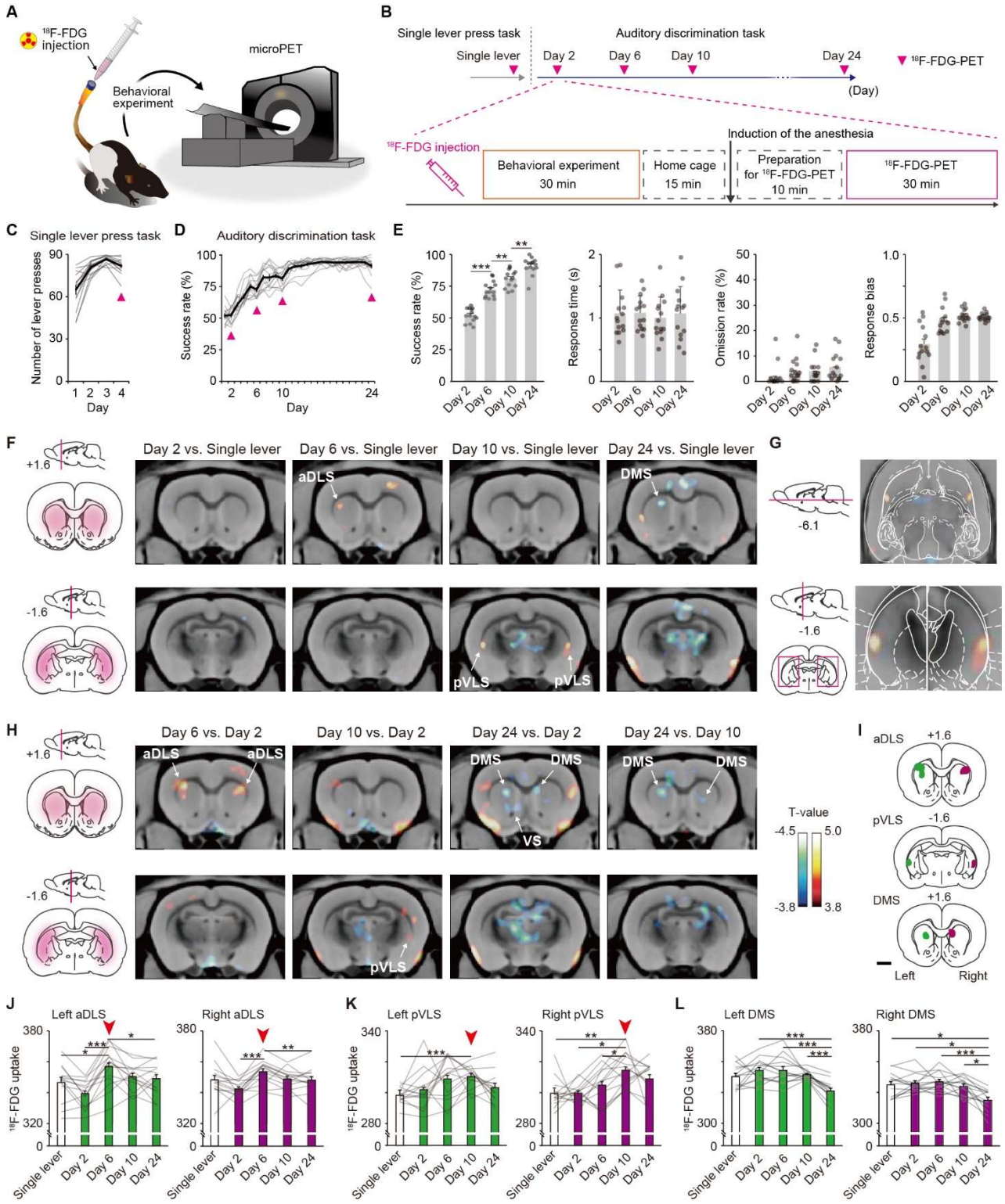


Figure 2. Dissociable brain activity patterns during auditory discrimination learning among three striatal subregions. (A) Schematic illustration of an awake rat received intravenously ^{18}F-FDG injection through an indwelling catheter attached to the tail. The rat was conducted for the behavioral experiment and then used for microPET imaging. (B) Schedule for the behavioral training and ^{18}F-FDG-PET scan. After ^{18}F-FDG injection, rats (n = 14 rats) were subjected to the

953 behavioral experiment (30 min) and returned to the home cage (15 min). Ten min before the PET
954 scan started, the rats were anesthetized and then the scan was started (30 min). (C and D) Learning
955 curves of the single lever press task (C) and auditory discrimination task (D). Arrowheads indicate
956 the PET scan days. (E) Behavioral performance on the scan days during the auditory discrimination.
957 Success rate (one-way repeated ANOVA, $F[3,39] = 125.012$, $p = 4.8 \times 10^{-20}$, *post hoc* Bonferroni
958 test; Day 2 vs. Day 6, $p = 5.0 \times 10^{-7}$; Day 6 vs. Day 10, $p = 0.001$; Day 10 vs. Day 24, $p = 0.003$),
959 response time (one-way repeated ANOVA, $F[3,39] = 0.156$, $p = 0.926$), omission rate (one-way
960 repeated ANOVA, $F[3,39] = 1.559$, $p = 0.215$), and response bias (one-way repeated ANOVA,
961 $F[1.779,23.130] = 17.734$, $p = 3.6 \times 10^{-5}$) are shown. (F) Representative images of coronal sections
962 that compared the brain activity in the single lever press task with the activity on Day 2, 6, 10, or 24
963 in the discrimination task. Left panels show schematic illustrations of striatal subregions. (G)
964 Horizontal (top) and coronal (bottom) images of striatal activation areas shown on Day 10 vs. single
965 lever in (F). (H) Representative images of coronal section that compared the brain activity on Day 2
966 with that on Day 6, 10, or 24 in the discrimination task. Color bars indicate the T-values, and a
967 value of 3.8 was used as the threshold corresponding to the uncorrected threshold ($p < 0.001$). (I)
968 Schematic pictures showing voxel of interests for the aDLS, pVLS, and DMS (green, left
969 hemisphere; and purple, right hemisphere). (J-L) Regional ^{18}F-FDG uptakes in the aDLS (J, one-
970 way repeated ANOVA, left aDLS, $F[4,52] = 10.322$, $p = 3.0 \times 10^{-6}$, *post hoc* Bonferroni test; single
971 lever vs. Day 6, $p = 0.016$; Day 2 vs. Day 6, $p = 6.5 \times 10^{-5}$; Day 2 vs. Day 10, $p = 0.019$; Day 2 vs.
972 Day 24, $p = 0.017$; Day 6 vs. Day 24, $p = 0.041$; right aDLS, $F[4,52] = 7.462$, $p = 7.9 \times 10^{-5}$, *post*
973 *hoc* Bonferroni test, Day 2 vs. Day 6, $p = 2.9 \times 10^{-4}$; Day 2 vs. Day 10, $p = 0.036$; Day 6 vs. Day
974 24, $p = 0.005$), pVLS (K, one-way repeated ANOVA, left pVLS, $F[2.368,30.784] = 4.152$, $p =$
975 0.020, *post hoc* Bonferroni test; single lever vs. Day 10, $p = 1.8 \times 10^{-5}$; right pVLS, $F[4,52] =$
976 5.995, $p = 4.8 \times 10^{-4}$, *post hoc* Bonferroni test, single lever vs. Day 10, $p = 0.001$; Day 2 vs. Day 10,
977 $p = 0.011$; Day 6 vs. Day 10, $p = 0.032$), and DMS (L, one-way repeated ANOVA, left DMS,
978 $F[4,52] = 12.836$, $p = 2.4 \times 10^{-7}$, *post hoc* Bonferroni test; Day 2 vs. Day 24, $p = 1.8 \times 10^{-5}$; Day 6

979 vs. Day 24, $p = 4.8 \times 10^{-4}$; Day 10 vs. Day 24, $p = 5.5 \times 10^{-5}$; right DMS, $F[4,52] = 10.717$, $p = 2.0$

980 $\times 10^{-6}$, *post hoc* Bonferroni test; single lever vs. Day 24, $p = 0.036$; Day 2 vs. Day 24, $p = 0.011$;

981 Day 6 vs. Day 24, $p = 2.6 \times 10^{-4}$; Day 10 vs. Day 24, $p = 0.011$). Arrowheads indicate the day with

982 the most activations throughout the learning process. Data are indicated as the mean ± s.e.m., and

983 individual data are overlaid. The anteroposterior coordinates from bregma (mm) are shown (F–I).

984 Scale bar; 2 mm (I). *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

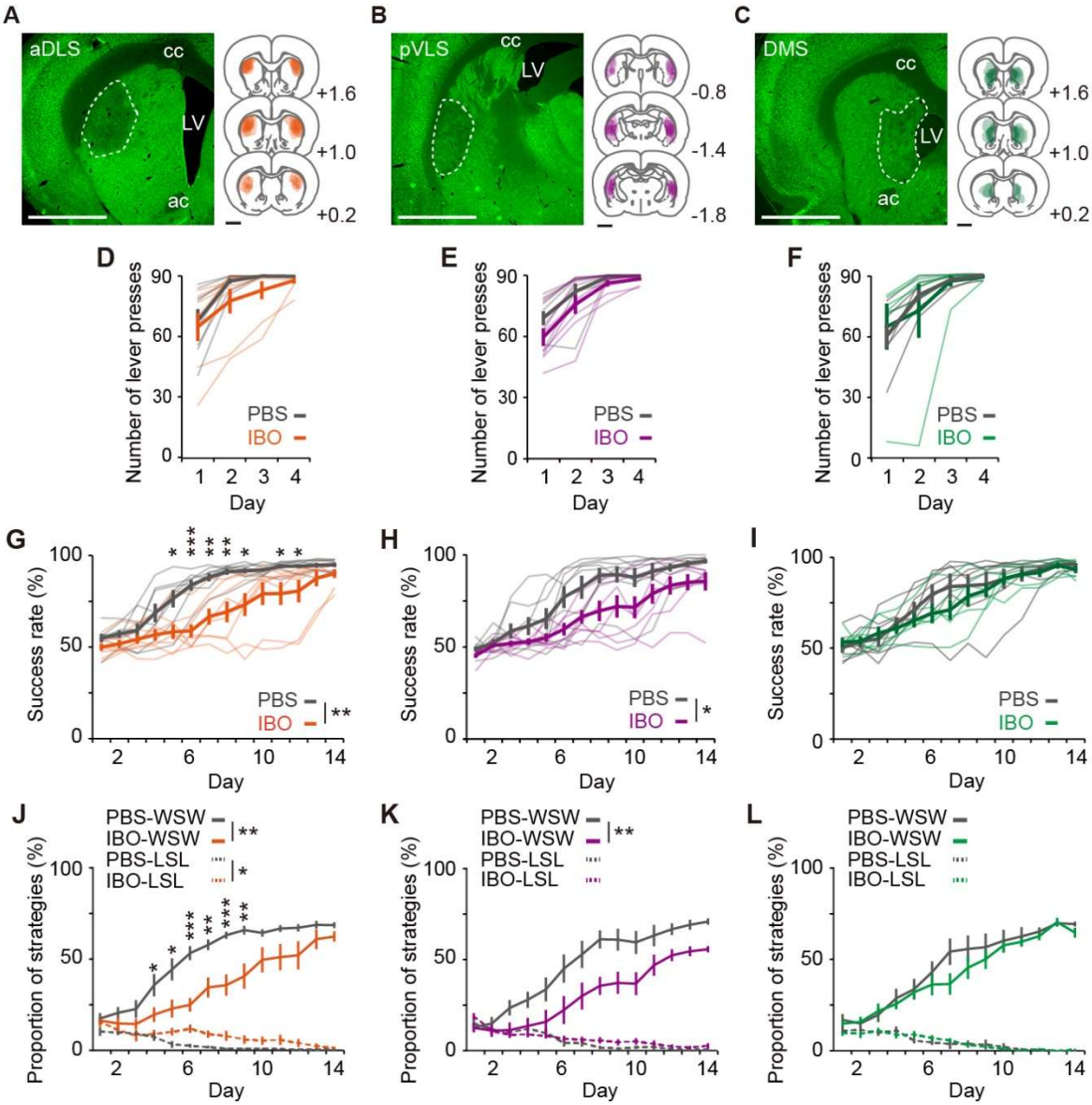


Figure 3. **Impacts of excitotoxic lesion of striatal subregions on the acquisition of auditory discrimination.** Rats were given intracranial injection of PBS or IBO solution into the aDLS (n = 8 for each injection), pVLS (n = 8 for each injection), or DMS (n = 7 for PBS injection, and n = 6 for IBO injection) before start of the single lever press task. (A-C) Representative images of NeuN immunostaining and individual schematic illustrations showing lesioned area in the aDLS (A), pVLS (B), or DMS (C). Dotted lines in the images indicate the range of the lesioned area. ac, anterior commissure; cc, corpus callosum; and LV, lateral ventricle. (D-F) Learning curves during the single lever press task in the groups injected into the aDLS (D, two-way repeated ANOVA, group, $F[1,14] = 1.275$, $p = 0.278$, day, $F[1.512,21.171] = 26.281$, $p = 7.0 \times 10^{-6}$, group × day, $F[1.512,21.171] = 0.834$, $p = 0.418$), pVLS (E, two-way repeated ANOVA, group, $F[1,14] = 2.542$,

$p = 0.133$, day, $F[1.547,21.658] = 40.183$, $p = 2.2 \times 10^{-7}$, group × day, $F[1.547,21.658] = 0.953$, $p = 0.380$), or DMS (F, two-way repeated ANOVA, group, $F[1,11] = 0.025$, $p = 0.876$, day, $F[1.205,13.252] = 14.818$, $p = 0.001$, group × day, $F[1.205,13.252] = 0.610$, $p = 0.478$). (G-I) Learning curves during the auditory discrimination task in the groups injected into the aDLS (G, two-way repeated ANOVA, group, $F[1,14] = 11.578$, $p = 0.004$, day, $F[2.805,39.274] = 60.733$, $p = 1.8 \times 10^{-14}$, group × day, $F[2.805,39.274] = 3.863$, $p = 0.018$, simple main effect, PBS vs. IBO, Day 1, $p = 0.140$, Day 2, $p = 0.085$, Day 3, $p = 0.284$, Day 4, $p = 0.068$, Day 5, $p = 0.011$, Day 6, $p = 9.5 \times 10^{-5}$, Day 7, $p = 0.001$, Day 8, $p = 0.002$, Day 9, $p = 0.012$, Day 10, $p = 0.059$, Day 11, $p = 0.034$, Day 12, $p = 0.046$, Day 13, $p = 0.136$, Day 14, $p = 0.096$), pVLS (H, two-way repeated ANOVA, group, $F[1,14] = 7.221$, $p = 0.018$, day, $F[3.646,51.041] = 51.794$, $p = 9.2 \times 10^{-17}$, group × day, $F[3.646,51.041] = 1.612$, $p = 0.190$), or DMS (I, two-way repeated ANOVA, group, $F[1,11] = 0.469$, $p = 0.507$, day, $F[2.677,29.452] = 48.860$, $p = 3.9 \times 10^{-11}$, group × day, $F[2.677,29.452] = 1.094$, $p = 0.362$). (J-L) Changes in the proportions of the win-shift-win or lose-shift-lose strategies through the discrimination learning in the groups treated into the aDLS (J, two-way repeated ANOVA, group, $F[1,14] = 13.451$, $p = 0.003$, day, $F[3.048,42.673] = 54.777$, $p = 9.1 \times 10^{-15}$, group × day, $F[3.048,42.673] = 3.288$, $p = 0.029$, simple main effect, PBS vs. IBO, Day 1, $p = 0.728$, Day 2, $p = 0.225$, Day 3, $p = 0.165$, Day 4, $p = 0.044$, Day 5, $p = 0.015$, Day 6, $p = 1.1 \times 10^{-4}$, Day 7, $p = 0.002$, Day 8, $p = 5.1 \times 10^{-4}$, Day 9, $p = 0.004$, Day 10, $p = 0.060$, Day 11, $p = 0.053$, Day 12, $p = 0.081$, Day 13, $p = 0.198$, Day 14, $p = 0.075$ for win-shift-win; group, $F[1,14] = 5.039$, $p = 0.041$, day, $F[3.866,54.119] = 14.578$, $p = 5.2 \times 10^{-8}$, group × day, $F[3.866,54.119] = 2.471$, $p = 0.057$ for lose-shift-lose), pVLS (K, two-way repeated ANOVA, group, $F[1,14] = 9.251$, $p = 0.009$, day, $F[3.206,44.881] = 47.168$, $p = 2.7 \times 10^{-14}$, group × day, $F[3.206,44.881] = 1.862$, $p = 0.146$ for win-shift-win; group, $F[1,14] = 0.762$, $p = 0.397$, day, $F[4.773,66.815] = 24.018$, $p = 1.9 \times 10^{-13}$, group × day, $F[4.773,66.815] = 1.338$, $p = 0.260$ for lose-shift-lose), or DMS (L, two-way repeated ANOVA, group, $F[1,11] = 0.703$, $p = 0.420$, day, $F[3.052,33.568] = 48.424$, $p = 2.3 \times 10^{-12}$, group × day, $F[3.052,33.568] = 0.846$, $p = 0.480$ for win-shift-win; group, $F[1,11] = 0.002$, $p = 0.965$, day,

$F[3.442,37.866] = 13.787$, $p = 1.0 \times 10^{-6}$, group × day, $F[3.442, 37.866] = 0.581$, $p = 0.654$ for lose-shift-lose). Data are indicated as the mean ± s.e.m., and individual data are overlaid (except for panels J–L). The anteroposterior coordinates from bregma (mm) are shown (A–C). Scale bars; 2 mm (A–C). *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. Abbreviations: WSW. win-shift-win; and LSL, lose-shift-lose.

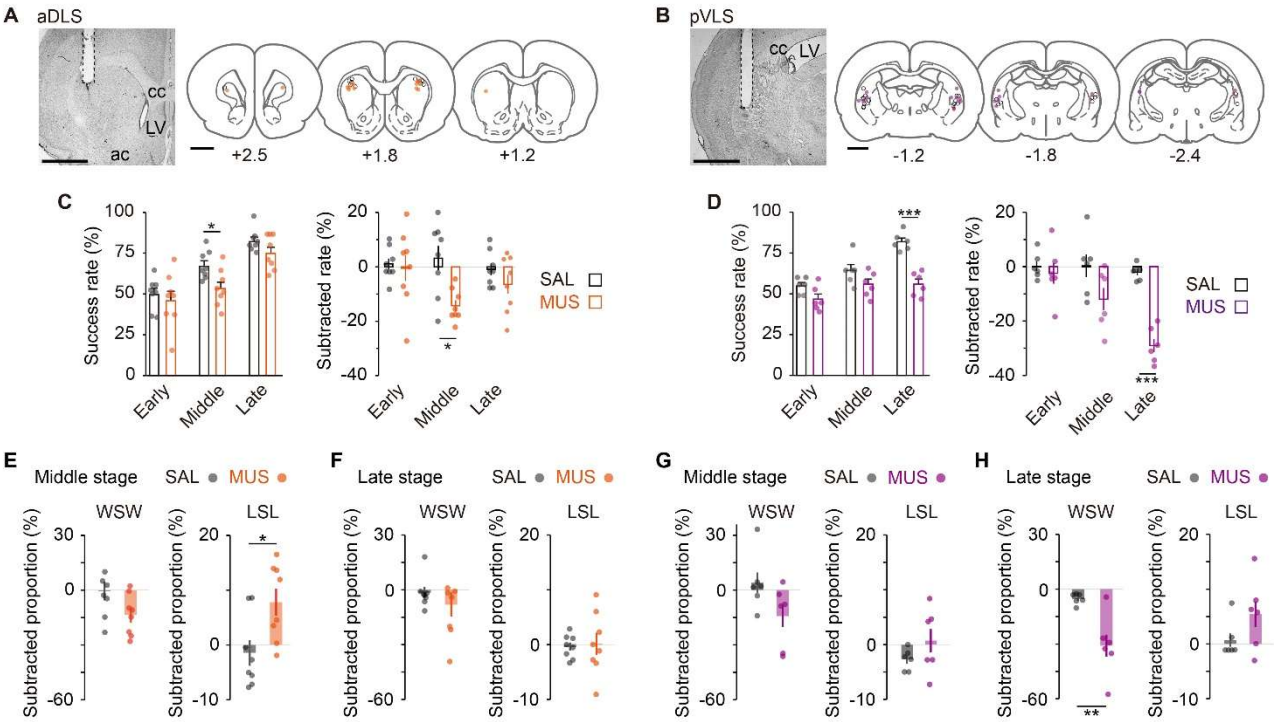


Figure 4. Influence of transient inhibition of striatal subregions at different timings on the performance of auditory discrimination. Rats received intracranial injection of SAL or MUS solution into the aDLS (n = 8 for each injection) or pVLS (n = 6 for each injection). (A and B) Representative images of cresyl violet staining and schematic illustrations showing placement sites of the tip of guide cannula in the aDLS (A) or pVLS (B). Dotted lines in the images indicate the position of the cannula placement. Ac, anterior commissure; cc, corpus callosum; and LV, lateral ventricle. (C and B) Effects of transient striatal inhibition on the performance. For aDLS inhibition (C), success rate (left, two-way repeated ANOVA, stage, $F[2,28] = 28.708$, $p = 1.7 \times 10^{-7}$, group, $F[1,14] = 8.840$, $p = 0.010$, stage × group, $F[2,28] = 0.757$, $p = 0.478$; unpaired Student's t-test, early, $t[14] = 0.558$, $p = 0.586$; middle, $t[14] = 2.764$, $p = 0.015$; Welch's t-test, late, $t[12.010] = 1.617$, $p = 0.132$, $p = 0.016$ after a Bonferroni correction method) and subtracted success rate on day N-1 from that on day N (right, two-way repeated ANOVA, stage, $F[2,28] = 1.231$, $p = 0.307$, group, $F[1,14] = 10.797$, $p = 0.005$, stage × group, $F[2,28] = 2.360$, $p = 0.113$; unpaired Student's t-test, early, $t[14] = 0.285$, $p = 0.780$, late, $t[14] = 1.271$, $p = 0.225$, Welch's t-test, middle, $t[9.603] = 3.331$, $p = 0.008$, $p = 0.016$ after a Bonferroni correction method). For pVLS inhibition (D), success rate (left, two-way repeated ANOVA, stage, $F[2,20] = 17.642$, $p = 3.8 \times 10^{-5}$, group, $F[1,10] =$

43.942, $p = 5.9 \times 10^{-5}$, stage × group, $F[2,20] = 4.729$, $p = 0.012$; simple main effect, early, $p = 0.055$, middle, $p = 0.109$, late, $p = 4.2 \times 10^{-5}$) and subtracted success rate on day N-1 from that on day N (right, two-way repeated ANOVA, stage, $F[2,20] = 12.105$, $p = 3.6 \times 10^{-4}$, group, $F[1,10] = 16.310$, $p = 0.002$, stage × group, $F[2,20] = 8.500$, $p = 0.002$; simple main effect, early, $p = 0.595$, middle, $p = 0.085$, late, $p = 3.0 \times 10^{-6}$). (E and F) Subtracted proportion of the win-shift-win or lose-shift-lose strategies in the groups injected into the aDLS at the middle stage (E, unpaired Student's t-test, $t[14] = 2.038$, $p = 0.061$ for win-shift-win, and $t[14] = 2.714$, $p = 0.017$ for lose-shift-lose) and the late stage (F, unpaired Student's t-test, $t[14] = 0.898$, $p = 0.384$ for win-shift-win, and $t[14] = 0.226$, $p = 0.824$ for lose-shift-lose). (G and H) Subtracted proportion of the win-shift-win or lose-shift-lose strategies in the groups injected into the pVLS at the middle stage (G, unpaired Student's t-test, $t(10) = 1.924$, $p = 0.083$ for win-shift-win, and Welch's t-test, $t[6.095] = 1.364$, $p = 0.221$ for lose-shift-lose) and the late stage (H, unpaired Student's t-test, $t[10] = 3.629$, $p = 0.005$ for win-shift-win, and $t[10] = 1.577$, $p = 0.146$ for lose-shift-lose). Data are indicated as the mean ± s.e.m., and individual data are overlaid. The anteroposterior coordinates from bregma (mm) are shown (A and B). Scale bars; 2 mm (A and B). *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. Abbreviations: WSW, win-shift-win; and LSL, lose-shift-lose.

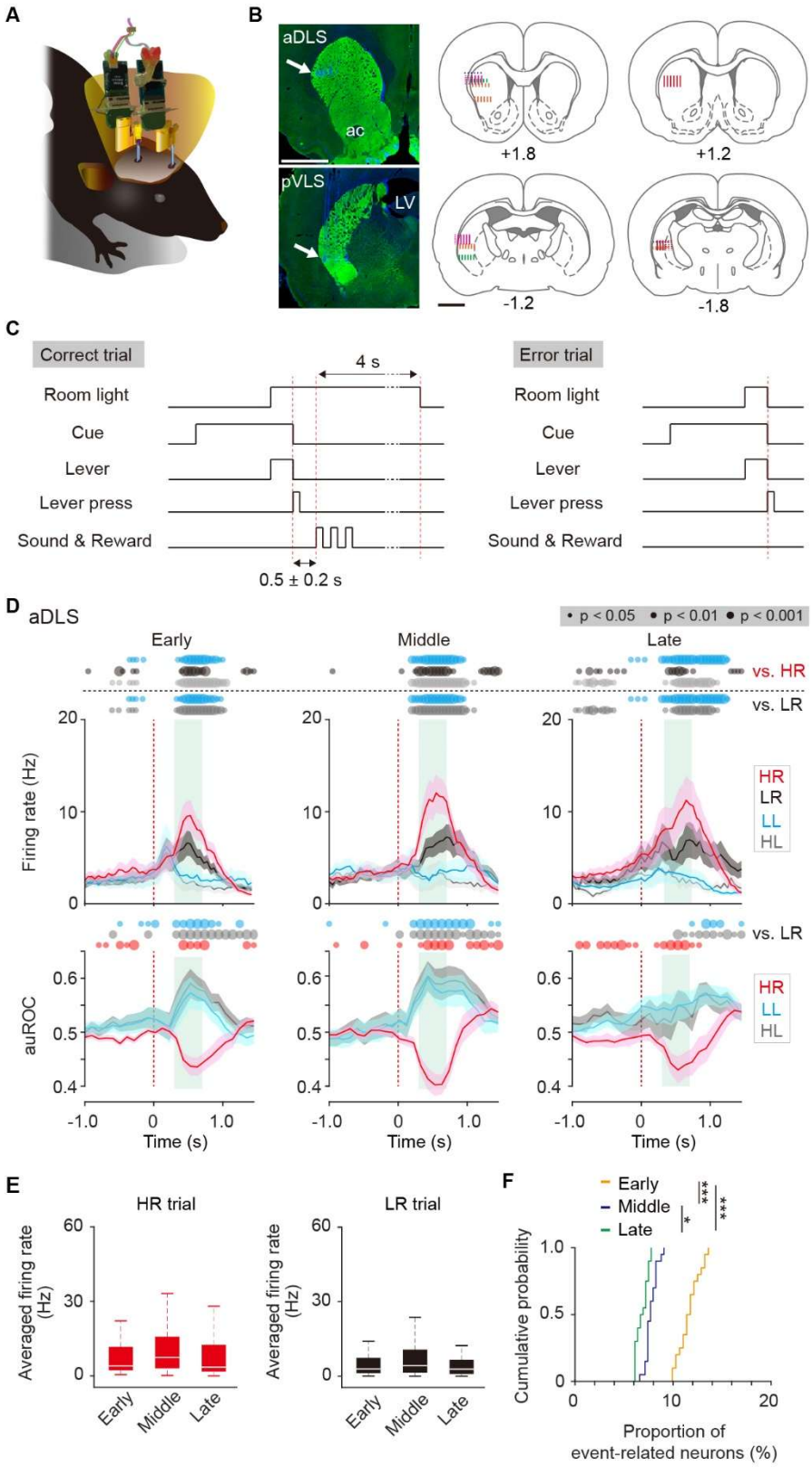


Figure 5. Multi-unit recording of neurons in striatal subregions and firing activity related to the behavioral outcome of reward sound-related HR type neurons in the aDLS. (A) Schematic illustration of a freely moving rat used for simultaneous multi-unit recordings in the aDLS and pVLS. Six rats were used for the following analysis. (B) Representative images showing positions

of electrode tips (arrows) in the aDLS and pVLS (left rows). The recording sites estimated by electrode tracks and electrical marks in individual rats are shown (center and right rows). ac, anterior commissure; LV, lateral ventricle. (C) Sequence of some events in correct and error trials. The delay periods were pseudorandomly added between a correct lever press and the reward sound (0.5 ± 0.2 s). The room light turned off the extended time (4 s) after the correct response or immediately after the error response. (D) Mean firing rate (top rows) and auROC values (bottom rows) of reward sound-related HR type neurons in the aDLS at the early (n = 43 neurons), middle (n = 37 neurons), and late (n = 44 neurons) stages. These data are aligned to the choice response. Green shadows show the period when the reward sound is presented (500 ± 200 ms after the lever press). Time bins with significant differences between the mean firing rate in the HR or LR trial and any of the rates in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. (E) Averaged firing rate during the reward sound period in the HR and LR trials at the early, middle, and late stages (Kruskal-Wallis test: HR trial, $\chi^2 = 3.028$, $p = 0.220$; and LR trial, $\chi^2 = 1.249$, $p = 0.536$). (F) Cumulative probability of the proportion of reward sound-related HR type neuron number at the three stages (Kruskal-Wallis test, $\chi^2 = 45.378$, $p = 1.4 \times 10^{-10}$; *post hoc* Tukey-Kramer test, $p = 7.1 \times 10^{-5}$ for early vs. middle, $p = 1.0 \times 10^{-9}$ for early vs. late, and $p = 0.040$ for middle vs. late). Data are indicated as the mean ± s.e.m. (D) or the median and quartiles with the maximal and minimal values (E). The anteroposterior coordinates from bregma (mm) are shown (B). Scale bars; 2 mm (B). *$p < 0.05$ and ***$p < 0.001$.

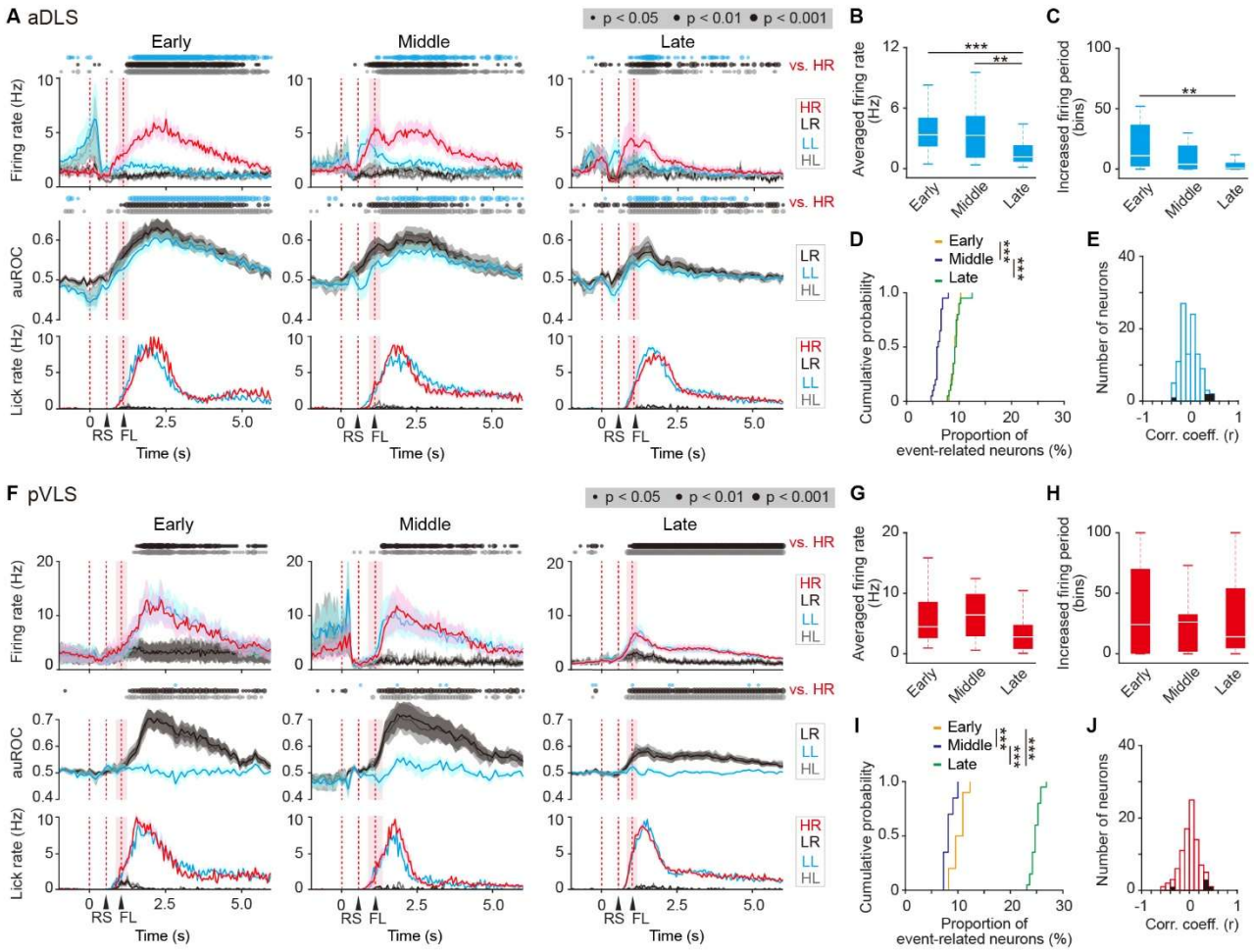


Figure 6. Sustained activity after the reward of first licking-related HR type neurons in the aDLS and pVLS. (A) Firing rate (top rows), auROC (middle rows), and licking rate (bottom rows) after the reward of first licking-related HR type neurons in the aDLS at the early (n = 27 neurons), middle (n = 31 neurons), and late (n = 49 neurons) stages. These data are aligned to the choice response. Time bins with significant differences between the mean firing rate in the HR trial and any of the rate in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. Timings of the reward sound and first licking are shown. (B) Averaged firing rate for 5 s after the first licking in the HR trial (Kruskal-Wallis test, $\chi^2 = 16.396$, $p = 2.8 \times 10^{-4}$, *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.726$, early vs. late, $p = 7.2 \times 10^{-4}$, middle vs. late, $p = 0.009$). (C) Averaged total number of time bins above 3 S.D. of the baseline firing of aDLS neurons for 5 s after the first licking in the HR trial (Kruskal-Wallis test, $\chi^2 = 12.803$, $p = 0.002$, *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.173$, early vs. late, $p = 0.001$, middle vs. late, $p = 0.219$). (D) Cumulative probability of the

proportion of first licking-related HR type neuron number (Kruskal-Wallis test, $\chi^2 = 39.280$, $p = 3.0 \times 10^{-9}$, *post hoc* Tukey-Kramer test, early vs. middle, $p = 1.3 \times 10^{-7}$, early vs. late, $p = 0.995$, middle vs. late, $p = 2.3 \times 10^{-7}$). (E) Distribution of the correlation coefficient between the numbers of spikes of first licking-related HR type neurons and licking. Closed column indicates the number of neurons showing significant correlations between the two parameters (5 out of 107 neurons). (F) Firing rate (top rows), auROC (middle rows), and licking rate (bottom rows) after the reward of FL-HR neurons in the pVLS at the early ($n = 17$ neurons), middle ($n = 11$ neurons), and late ($n = 64$ neurons) stages. These data are aligned to the choice response. Time bins with significant differences between the rate in the HR trial and any of the rate in other trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. The timings of the reward sound and first licking are shown. (G) Averaged firing rate of pVLS neurons after the first licking in the HR trial (Kruskal-Wallis test, $\chi^2 = 7.652$, $p = 0.022$; *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.975$, early vs. late, $p = 0.076$, and middle vs. late, $p = 0.096$). (H) Averaged total number of time bins above 3 S.D. of the baseline firing after the first licking in the HR trial (Kruskal-Wallis test, $\chi^2 = 0.023$, $p = 0.989$). (I) Cumulative probability of the proportion of first licking-related HR type neuron number (Kruskal-Wallis test, $\chi^2 = 47.922$, $p = 3.9 \times 10^{-11}$; *post hoc* Tukey-Kramer test, early vs. middle, $p = 0.011$, early vs. late, $p = 1.7 \times 10^{-4}$, and middle vs. late, $p = 9.7 \times 10^{-10}$). (J) Distribution of the correlation coefficient between the number of spikes in the pVLS and the licking. Closed column indicates the number of neurons showing significant correlations between the two parameters (5 out of 92 neurons). Data are indicated as the mean ± s.e.m. (A and F) or the median and quartiles with the maximal and minimal values (B, C, G, and H). $^{**}p < 0.01$ and $^{***}p < 0.001$. Abbreviations: RS, reward sound; and FL, first licking.

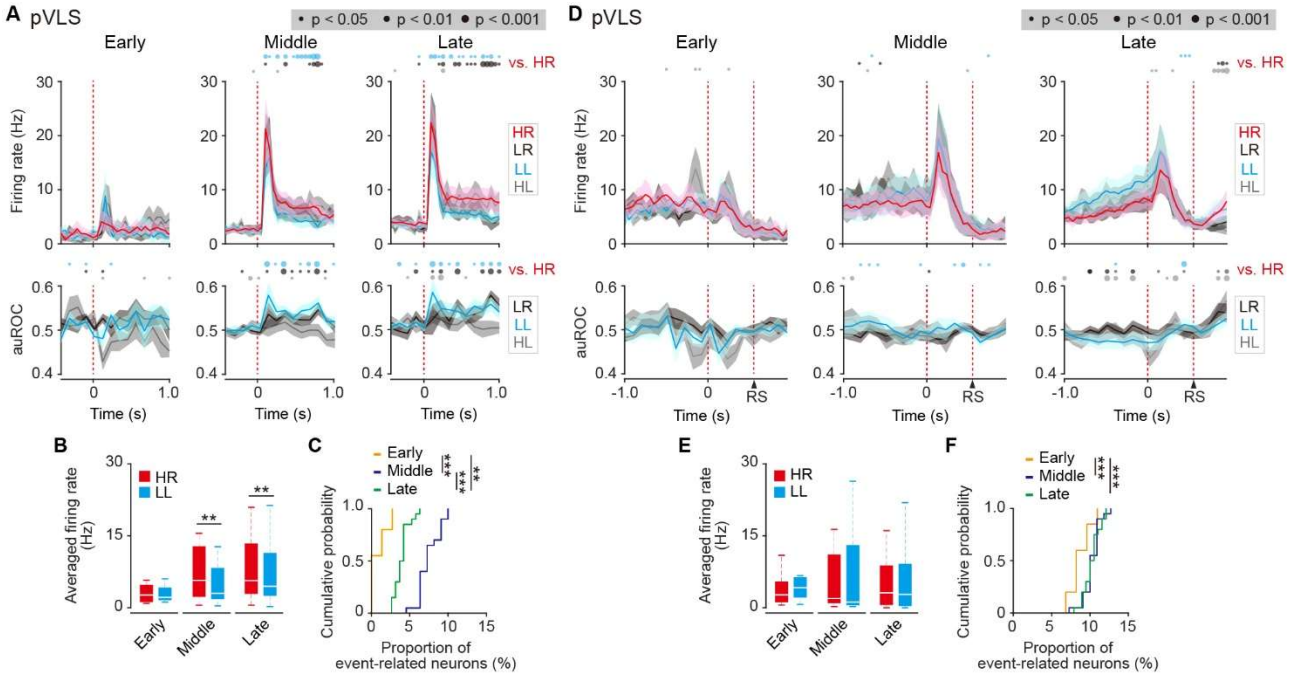


Figure 7. Transient activity related to the beginning and ending of a behavior of cue onset-HR and choice response-related HR type neurons in the pVLS. (A) Firing rate (top rows) and auROC values (bottom rows) of cue onset-related HR type neurons in the pVLS at the early (n = 4 neurons), middle (n = 21 neurons), and late (n = 24 neurons) stages. These data are aligned to the cue onset. Time bins with significant differences between the mean firing rate in the HR trial and either of the rate in other three trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. (B) Averaged firing rate of the cue onset-related HR type neurons in HR and LL trials (0-600 ms after the cue onset) (Wilcoxon signed rank test; early, p = 0.875, middle, p = 0.006, and late, p = 0.004). (C) Cumulative probability of the proportion of the cue onset-related HR type neuron number (Kruskal-Wallis test, χ^2 = 51.114, p = 8.0 × 10^{-12}; *post hoc* Tukey-Kramer test, early vs. middle, p = 9.6 × 10^{-10}, early vs. late, p = 0.002, and middle vs. late, p = 6.4 × 10^{-4}). (D) Firing rate (top rows) and auROC values (bottom rows) of the choice response-related HR type neurons in the pVLS at the early (n = 11 neurons), middle (n = 22 neurons), and late (n = 41 neurons) stages. These data are aligned to the choice response. Time bins with significant differences between the mean rate in the HR trial and any of the rate in other trials (Wilcoxon signed rank test) or the distribution of auROC values and 0.5 (Wilcoxon rank-sum test) are represented by the circles at the top. The timing of the

reward sound is presented. (E) Averaged firing rate of the choice response-related HR type neurons in HR and LL trials (0-600 ms after the choice response) (Wilcoxon signed rank test; early, p = 0.278, middle, p = 0.758, and late, p = 0.279). (F) Cumulative probability of the proportion of the choice response-related HR type neuron number (Kruskal-Wallis test, χ^2 = 13.786, p = 0.001; *post hoc* Tukey-Kramer test, early vs. middle, p = 0.003, early vs. late, p = 0.005, and middle vs. late, p = 0.976). Data are indicated as the mean ± s.e.m. (A and B) or the median and quartiles with the maximal and minimal values (B and E). **p < 0.01 and ***p < 0.001. Abbreviation: RS, reward sound.