

Error-related Signaling in Nucleus Accumbens D2 Receptor-expressing Neurons Guides Avoidance-based Goal-directed Behavior

Tadaaki Nishioka

Osaka University <https://orcid.org/0000-0002-5080-0407>

Tom Macpherson

Osaka University

Kosuke Hamaguchi

Department of Biological Sciences, Kyoto University Graduate School of Medicine

<https://orcid.org/0000-0002-0540-320X>

Takatoshi Hikida (✉ hikida@protein.osaka-u.ac.jp)

Osaka University

Article

Keywords:

Posted Date: January 5th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1109714/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on April 21st, 2023. See the published version at <https://doi.org/10.1038/s41467-023-38025-3>.

Abstract

Learnt associations between environmental cues and the outcomes they predict (cue-outcome associations) play a major role in behavioral control, guiding not only which responses we should perform, but also which we should avoid, in order to achieve a specific goal. The encoding of such cue-outcome associations, as well as the performance of cue-guided goal-directed behavior, is thought to involve dopamine D1 and D2 receptor-expressing medium spiny neurons (D1-/D2-MSNs) of the nucleus accumbens (NAc). Here, using a visual discrimination task in mice, we assessed the role of NAc D1-/D2-MSNs in cue-guided goal-directed avoidance of inappropriate responding. Cell-type specific neuronal silencing and in-vivo imaging revealed NAc D2-MSNs to selectively contribute to cue-guided avoidance behavior, with activation of NAc D2-MSNs following response error playing an important role in optimizing future goal-directed behavior. Our findings indicate that error-signaling by NAc D2-MSNs underlies the ability to use environmental cues to avoid inappropriate behavior.

Introduction

Goal-directed behavior can be guided not only by strategies based upon the acquisition of a desirable outcomes, but also by those aimed at reducing undesirable outcomes. Indeed, we are often faced with situations in which the correct behavioral response to acquire a desired outcome is unknown or ambiguous, but prior experience of failure can be used to avoid inappropriate responses. The important role that negative outcomes play in guiding avoidance behavior has long been appreciated, dating back from Thorndike's (1927) *law of effect* to the more recent proposal of *loss aversion* in the field of behavioral economics¹⁻³. However, the neural mechanisms that underly the use of such avoidance-based strategies for goal-directed behavior are still unclear, and studies to date have primarily focused investigating the reinforcement of behaviors that directly result in rewarding outcomes⁴⁻⁶.

Within a limbic cortico-basal ganglia-thalamo-cortical signalling loop, the nucleus accumbens (NAc), in particular the NAc Core subregion, is suggested to play an important role in goal-directed decision making via its role in linking information processing concerning outcome values with that related to goal selection⁷⁻¹⁰. NAc medium spiny projection neurons (MSN), the major neuron type, are typically anatomically divided into two roughly equal subpopulations; dopamine D1 receptor-expressing MSNs (D1-MSN) that project predominantly to the ventral pallidum (VP) and substantia nigra pars reticulata (SNr), and dopamine D2 receptor-expressing MSNs (D2-MSN) that project predominantly to the VP^{11,12}. While previous studies have demonstrated NAc D1-MSNs to be implicated in reward-related learning and D2-MSNs to be involved in aversion-related learning and behavioral flexibility¹³⁻¹⁵, the specific role that NAc D1- and D2-MSNs may play in avoidance-based behavior are less clear.

To isolate and measure the ability for avoidance-based goal-directed behavior, we designed a novel visual discrimination-based cue-guided avoidance learning (VD-Avoid) task in which mice were required to avoid an instrumental response at a visual cue known to be associated with a reward omission, and instead respond at a random cue that had not previously been associated with any outcome in order to acquire a

liquid reward. In a series of experiments, miniature microscope *in vivo* calcium imaging was used to investigate the precise activity patterns of D1- and D2-MSNs during the VD-Avoid task, while chemogenetic and time-specific optogenetic silencing of NAc D1-MSN or D2-MSNs during the same task were used to establish whether inactivation of these two subpopulations impairs the utilization of avoidance-based behavioral strategies. Our findings indicate that while D1-MSNs are activated by rewarding outcomes and contribute to the ongoing performance of behaviors directly resulting in rewarding outcomes, D2-MSNs are activated by reward omission and are necessary for avoidance of responses resulting in non-rewarded outcomes.

Results

Mice Can Acquire Avoidance-based Goal-directed Behavior

To assess the ability for avoidance-based goal-directed behavior in mice, we created a novel avoidance-based visual discrimination task (VD-Avoid) in which mice used visual cues to determine which of two touchscreen response windows should be avoided in order to receive a liquid reward by responding at the alternate window (Fig. 1A). Following trial initiation, a visual cue was presented in each of the two response windows (Fig. 1B); one visual cue (correct cue) was randomly changed each trial (51 possible images) and resulted in delivery of a liquid reward (7 μ l condensed milk at the reward magazine) following a touchscreen response, while the other visual cue (incorrect cue) was kept consistent during all trials and resulted in no reward and a 5-sec time-out following a touchscreen response. This design meant that mice could not form a cue-outcome association for the correct cue due to its random nature, but must instead rely on the cue-outcome association of the incorrect cue to guide appropriate behavior (avoidance of the known visual cue and a touch response at the unknown visual cue).

Within 14 days of training, all C57BL/6J mice were able to reach the task criterion of $\geq 80\%$ correct responses in a 60-min session for two consecutive days, indicating that visual cues signalling incorrect responses are sufficient to guide avoidance-based goal-directed behavior (Fig. 1C, 1E and 1F). Indeed, as training progressed, the correct response latency gradually decreased, suggesting that mice were able to perform the task more efficiently following repeated training (Fig. 1D and 1G). On the other hand, there was no significant change in reward collection latency, suggesting that motivation to acquire the reward does not change across learning (Fig. 1D and 1H). Finally, as several studies have reported the phenomenon of post-error slowing, in which the subjects tend to show a slower response subsequent to an error trial^{16,17}, it was investigated whether the recent history of the previous trial affected the response latency on the subsequent trial. Mice showed a trend towards slower responses on trials that immediately followed a trial in which a response error was made (**Fig. S1**), suggesting that response latencies may also be modulated by recent history in the VD-Avoid task. Together, these data indicate the ability of mice to acquire a goal-directed avoidance of a visual cue signaling an incorrect response window, providing a framework for studying the neural mechanisms underlying avoidance-based goal-directed behavior.

NAc D1- and D2-MSNs Control Different Aspects of Goal-directed Behaviors

Given the proposed importance of the NAc in goal-directed behavior and inhibitory control^{10,18-20}, we investigated whether NAc D1- and D2-MSNs contribute to performance in the VD-Avoid task. We bilaterally injected a Cre-dependent adeno-associated virus (AAV) expressing an inhibitory designer receptor exclusively activated by a designer drug (iDREADD, hM4Di²¹) fused to the fluorescent marker mCherry (AAV5-hSyn-DIO-hM4Di-mCherry), or the mCherry marker by itself (AAV5-hSyn-DIO-mCherry) for control animals, into the NAc of D1-Cre or D2-Cre mice (Fig. 2A). hM4Di iDREADDs have been established to robustly suppress neuronal activity when the artificial ligand clozapine-N-oxide (CNO) is administered²², and have previously been demonstrated to be effective for inhibition of NAc activity²³. Injections sites were targeted at the core region of the NAc due to the proposed importance of this subregion in goal-directed behavior^{8,10,24,25} and were histologically confirmed by fluorescence microscopic observation of mCherry expression following the completion of experiments (Fig. 2B and S2). In D1-Cre mice, hM4Di-mCherry-positive axon terminals were observed in the ventral pallidum (VP) and substantia nigra pars reticulata (SNr), while in D2-Cre mice, hM4Di-mCherry-positive axon terminals were observed only in the VP but not SNr (**Fig. S3**), consistent with the canonical targets of D1-MSN or D2-MSN in the NAc^{11,12}.

Following training to criterion levels of performance ($\geq 80\%$ correct on two consecutive days) in the VD-Avoid task, mice expressing iDREADDs in NAc D1- or D2-MSNs and control mice were intraperitoneally administered Vehicle or CNO across 4 test sessions (Fig. 2C). Chemogenetic inhibition of either NAc D1- or D2-MSNs was sufficient to transiently decrease performance in the VD-Avoid task (Fig. 2D, **top**). Indeed, averaging performances across the two Vehicle and CNO test sessions revealed performance to be similarly reduced in mice expressing iDREADDs in NAc D1- or D2-MSNs, but not in control mice (D1- and D2-Cre mice data were merged for the control group as no significant effect of genotype was observed) (Fig. 2D, **bottom**). While both NAc D1- and D2-MSN inhibition impaired performance, trial-by-trial analysis of data revealed these two neuron populations to contribute to different aspects of the task. Chemogenetic inhibition of NAc D1-MSNs reduced the amount of correct responses in trials immediately following a correct response (Fig. 2E), whereas inhibition of NAc D2-MSNs reduced correct responses in trials immediately following either correct or incorrect responses (Fig. 2F). In control mice, correct responses in trials immediately following either correct or incorrect responses were unaffected by CNO delivery (Fig. 2G). These findings suggest that while NAc D1-MSNs may be important for the maintenance of correct goal-directed behavioral strategies, NAc D2-MSNs may play a role in error-based learning to avoid an incorrect strategy. Considering that the NAc also plays an important role in motivational control²⁶⁻²⁸, we additionally analyzed the effect of chemogenetic manipulation on motivation-related measures. iDREADD inhibition of NAc D1-MSNs, but not D2-MSNs, reduced the total number of earned rewards and increased the latency to collect the reward, indicating that activity in NAc D1-MSNs contributes to the motivation to perform the task (**Fig. S4**).

Finally, to further confirm the observed roles of NAc D1- and D2-MSNs in the VD-Avoid task, we additionally tested animals in a reversed visual discrimination task in which a consistent visual cue signaled the rewarded response window that should be attended to, while a randomly-assigned visual cue signaled an unrewarded response window (VD-Attend task) (**Fig. S5A**). Interestingly, although iDREADD inhibition of NAc D1-MSNs similarly decreased the performance of the VD-Attend task, inhibition of NAc D2-MSNs had no effects on performance (**Fig. S5B-S5F**). Additionally, similar to the VD-Avoid task, reduced performances observed in NAc D1-MSN inhibited mice were found to be the result of reduced performance in trials immediately following a correct response in the VD-Attend task (**Fig. S5D**). Collectively, our data suggest that NAc D1-MSNs contribute to attendance-based goal-directed behavior following correct responses, while NAc D2-MSNs are implicated in avoidance-based goal-directed behavior following response errors.

NAc D2-MSNs Are Transiently Activated Following Response Errors, While NAc D1-MSNs Are Transiently Activated Following Correct Responses

To investigate the precise temporal window in which these NAc D1- and D2-MSNs contribute to goal-directed behavior in the VD-Avoid task, we next performed *in-vivo* calcium imaging of each of these neuron types at the single-cell level using miniature microscope. An AAV expressing the fluorescent calcium indicator, jRCaMP7f²⁹, in a Cre dependent manner (AAV9-hSyn-FLEX-jRCaMP7f) was microinjected into the NAc Core of D1-Cre or D2-Cre mice, and a gradient-index (GRIN) lens implanted above the viral injection site (Fig. 3A, 3B, S6 and S7). Neural activity was recorded in freely moving mice on the first and criterion sessions of the VD-Avoid task, and both D1- and D2-Cre mice were able to acquire the task within 1-3 weeks of training (**Fig. S7**). A constrained non-negative matrix factorization method for microendoscopic images (CNMFfe;³⁰ was used to analyze the neural activities of individual NAc D1- or D2-MSNs during performance of the VD-Avoid task (Fig. 3C and 3D; **Movies S1 and S2**). In the criterion session, a total of 266 cells were identified in D1-Cre mice and 194 cells in D2-Cre mice. To determine which neurons' activities were modulated by correct or error responses, we performed hierarchical clustering and classified neurons into groups based upon their activity profiles (Fig. 3E, 3F and S8). The averaged activities of identified neurons during the ITI (-5-0 sec from trial initiation), Cue (0-2 sec from trial initiation), and Outcome period (0-5 sec from a response) were then compared between correct and error trials to identify the time window during which activity was altered for each neuron type (Fig. 3G-3J). Four groups of neurons were identified; Type I neurons, that demonstrated increased activity during cue presentation in error trials and inhibition during the outcome of correct trials (Fig. 3E and 3G), Type II neurons, that showed activation and inhibition during the outcome period of error trials and correct trials, respectively (Fig. 3E and 3H), Type III neurons, that showed inhibition during the outcome period of error trials (Fig. 3E and 3I), and Type IV neurons, that showed activation during the outcome period of correct trials (Fig. 3E and 3J). Of all identified NAc D1-MSNs, 24.4% (65 of 266 cells) were classified as Type I, 30.8% (82 of 266 cells) as Type II, 31.2% (83 of 266 cells) as Type III, and 13.5% (36 of 266 cells) as Type IV (Fig. 3F), whereas, in D2-MSNs, 29.9% were (58 of 194 cells) were classified as Type I, 59.8% (116 of 194 cells) as Type II, 6.7% (13 of 194 cells) as Type III, and 3.6% (7 of 194 cells) as Type IV

(Fig. 3F). Moreover, the proportion of neurons activated by error responses (Type II) was significantly higher in D2-MSN than in D1-MSN population, while the proportions of neurons with correct preference (Type III and Type IV; neurons inhibited by error responses and those activated by correct responses, respectively) were higher in D1-MSN than in D2-MSN population (Fig. 3F). Collectively, neural imaging of NAc D1- and D2-MSNs during performance in the criterion session of the VD-Avoid task indicates that the majority (59.8%) of D2-MSNs and a subset (30.8%) of D1-MSNs in the NAc were activated during the Outcome period of error trials, suggesting that error-related signaling by these neurons during this time window may play an important role in guiding avoidance-based goal-directed behavior following response errors. Our data also shows that D1-MSNs in the NAc have a larger proportion of the neurons that are inhibited following error responses (Type III) or activated following correct responses (Type IV) than that of D2-MSN, suggesting that correct-related signaling by D1-MSN also contributes to higher performance in the VD-Avoid task.

Error-induced Signaling in NAc D2-MSNs Was Increased by Learning

To investigate whether outcome-specific patterns of activation in NAc D1- and D2-MSNs are shaped by learning, we next analyzed changes in the responses of individual NAc MSNs pre- and post-acquisition of the VD-Avoid task (Fig. 4A and S7) using a previously-established cell registration method³¹. A total of 239 neurons (D1-Cre, 103 pairs; D2-Cre, 136 pairs) were able to be identified in both the first (novice) and criterion (expert) sessions. First, we examined whether neural dynamics in clusters with clear activation in the outcome period (Type II and IV) was shaped by learning and whether there were differences in the original activity patterns between D1-MSN and D2-MSN. Neurons activated by error response (Type II) in the criterion session were largely unresponsive in error trials in the first session, indicating that error activation of the NAc Type II neurons in criterion sessions emerged from non-responding population in both D1/D2-MSN types (Fig. 4B, 4C, 4F, and 4G). On the other hand, D1-MSNs activated by correct response (Type IV) were largely unresponsive in correct trials in the first session (Fig. 4D and 4H), while D2-MSNs activated by correct response (Type IV) were also activated by correct responses in session 1 (Fig. 4E and 4I). Next, comparing averaged neural activity in the Outcome period for all types D1-MSNs, we found that D1-MSNs with correct preference (Type III and IV in expert) were activated in more expert than novice mice after correct responses (**Fig. S9A**), while Type III became inhibited after error responses (**Fig. S9B**). In contrast, in D2-MSNs, neurons activated by error responses (Type II in expert) became activated in more expert than novice mice after only error responses (**Fig. S9C and S9D**). Interestingly, D1-MSNs activated by error responses (Type II in expert) also became activated in more expert than novice mice after error responses (Fig. 4F and S9B), though not as much as D2-MSNs (Fig. 4G and S9D). Taken together, D1-MSNs with correct preference (Type III and IV) became relatively more active after correct responses through learning, as responsiveness bidirectionally changed in both correct and error trials. In contrast, D2-MSNs with error preference (Type II) showed a selective change in responsiveness to error trials.

Lastly, to investigate whether the difference of the calcium activity between error and correct choices is maintained over the course of learning, we calculated the area under the curve receiver operating characteristic (auROC) for the z-scored activities of individual neurons (**Fig. S10**). By tracking the same neurons pre- and post-acquisition of the VD-Avoid task, we found that the activity of NAc D2-MSNs, but not D1-MSNs, in expert mice was positively correlated with that in novice mice (**Fig. S10**), indicating that D2-MSNs signaling the difference of correct and error outcomes were relatively maintained over learning. Nevertheless, we detected a significant number of cells that acquired discriminability between correct and error outcomes (**Fig. S10**) in both cell types. These results indicate that the neuronal activation of D1-MSN and D2-MSN after making correct choices and error choices, respectively depends on the learning. The acquisition of responsiveness by learning of these neurons is likely to contribute to the high performance in the VD-Avoid task. Given that inhibition of D1-MSN in the NAc decreased the performance on the trial following a correct trial (Fig. 2E and S5D), it could be that neuronal activation of D1-MSN in the NAc contributes to keep same strategy with confidence.

Optogenetic Inhibition of NAc D2-MSNs During the Outcome Period of Error Trials Impairs Avoidance-based Goal-directed Behavior

Our iDREADD experiment showed that the chance of making a correct response following an error trial was reduced in D2-MSN inhibited mice but not in D1-MSN inhibited mice (compare Fig. 2E and 2F), suggesting the selective role of D2-MSN in error correction. In addition, our *in-vivo* calcium imaging experiments showed that larger proportion of D2-MSNs were activated during the outcome period of error trials (Fig. 3 and 4). To confirm the functional importance of the outcome period for NAc D2-MSN control of avoidance-based goal-directed behavior, we optogenetically inhibited the activity of NAc D2-MSNs in a time-locked manner. We first bilaterally injected a Cre-dependent AAV expressing the light-driven outward proton pump, archaerhodopsin (ArchT)³², fused to the fluorescent marker tdTomato (AAV5-FLEX-ArchT3.0-tdTomato), or the eYFP marker (AAV5-DIO-eYFP) for control animals, into the NAc of D2-Cre mice, then implanted optic fibers directly above the NAc (Fig. 5A, 5B, and S11). This technique has previously been established to suppress the activity of ArchT-expressing NAc D2-MSNs³³. We separated the task into 3 time periods (ITI period; the last 5 sec of the inter-trial interval (ITI), Cue period; the time from trial initiation to the response, Outcome period; 5 sec after the response, Fig. 5C) to test whether NAc D2-MSN activities in different time periods have different effects on the performance of the VD-Avoid task. The tests were performed on animals that had reached the criterion ($\geq 80\%$ correct on two consecutive days or $\geq 75\%$ correct on three consecutive days), and light stimulation was performed in a random order in 50% of trials. The timing of stimulation (ITI, Cue, or Outcome) was changed for each session, and the order of stimulation timing was randomized among mice (**Fig. S12**). We found that optogenetic inhibition of NAc D2-MSNs during the Outcome period, but not the ITI or Cue periods, impaired performance in the VD-Avoid task (Fig. 5D-5F). Trial-by-trial analysis of Outcome period D2-MSN inhibition sessions revealed that reduced performance was primarily caused by response errors in trials immediately following a response error, rather than response errors following a correct response, indicating a reduction in error signal-induced avoidance behavior (Fig. 5G and 5H). No changes in the

percentage of correct responses immediately following correct or error trials were observed when NAc D2-MSNs were inhibited during ITI or Cue periods (**Fig. S13A-S13D**). Interestingly, optogenetic inhibition of NAc D2-MSNs during the Outcome period, but not ITI or Cue periods, speed up responses in trials immediately following a response error (**Fig. S14A-S14F**), suggesting that activation of NAc D2-MSNs during the Outcome period of a response error contributes to the phenomenon of post-error slowing. Taken together, these findings indicate that activation of NAc D2-MSNs following the Outcome period of response errors causally contributes to the ability for cue-guided avoidance of inappropriate behavior.

Discussion

The NAc is a key component of basal ganglia and is thought to contribute to reward evaluation and motivation control by integrating glutamatergic and dopaminergic inputs from the cerebral cortex and ventral tegmental area (VTA), respectively^{5,10,26,34,35}. However, while the importance of the NAc has been discussed in several model of behavioral control^{7,8,10,36,37}, a precise understanding of the role of the NAc, and its constituent cell-types, in goal-directed behavior has remained elusive. Here we established a novel visual discrimination task in mice to assess the neural mechanisms underlying cue-guided avoidance-based goal-directed behavior without the influence of reward-associated cues and revealed that NAc D1- and D2-MSN contribute to goal-directed behavior through dissociable mechanisms. D1-MSNs in the NAc have a larger proportion of the neurons that are inhibited following error responses (Type III) or activated following correct responses (Type IV) than that of D2-MSN and D1-MSN activity was revealed to be important for consecutively correct responding in our visual discrimination task. In contrast, the majority of NAc D2-MSNs were found to be activated by error responses and error-related signaling in NAc D2-MSNs during the response outcome period was demonstrated to play an important role in avoidance of inappropriate behavior in the immediate future.

The NAc receives dense projections from dopamine (DA) neurons of the VTA, which are known to encode reward prediction error signals^{38,39}. Upon encounter with a reward greater than that predicted from previous experience, VTA DA neuron activity, and local DA release in the NAc Core, are reported to increase, while the opposite pattern is observed when a reward smaller than that predicted is encountered^{5,38,40}. Within the NAc, the excitability of D1- and D2-MSNs are likely to be bidirectionally modulated by local DA release^{41,42}. Indeed, while DA binding at Gs-protein coupled D1 receptors stimulates cAMP signaling, increasing cellular excitability, binding at Gi-protein coupled D2 receptors inhibits cAMP signaling, reducing the cell's excitability⁴³. Thus, the activity of NAc D1- and D2-MSNs observed in the present study corresponds with that expected according to the reward prediction error theory and the molecular properties of DA receptors; with a rewarding outcome activating NAc D1-MSNs, likely as a result of augmented local DA release, and a negative outcome activating NAc D2-MSNs, likely as the disinhibitory effect of a reduction in local DA release. These findings also support previous evidence indicating that NAc D1- and D2-MSNs play important roles in signaling reward and aversion, respectively^{13,44-47}. Interestingly, studies in humans have shown that increased DA concentration in the brain following L-dopa treatment, reduces the ability of participants to avoid choices that lead to negative

outcomes, but does not alter the ability to learn from positive outcomes in goal-directed learning tasks^{48,49}. Given that chemogenetic and optogenetic suppression of NAc D2-MSNs in our study similarly disrupted the ability of mice to avoid a behavioral response leading to a negative outcome when error-signals were blocked, but not when reward signals were blocked, it could be speculated that the ability of L-dopa treatment to impair avoidance-based goal-directed behavior in humans may have been the result of DA-induced hypoactivity of NAc D2-MSNs.

The temporal localization of NAc activity to the Outcome period of behavioral responses in our study suggests that NAc D1- and D2-MSNs may be important for monitoring and updating of goal-directed behavior, rather than for action selection itself. Interestingly, previous evidence has suggested that optogenetic stimulation of the dorsomedial striatum (DMS) during Cue presentation biases action selection⁵⁰, supporting models proposing the DMS to act as an actor and NAc to act as a critic in action selection and action evaluation, respectively^{51,52}. The NAc forms part of a limbic processing loop that has been reported to converge with associative/cognitive and motor processing loops, involving the DMS and DLS, respectively, at the level of the SNr^{53,54}. This circuitry provides a mechanism through which information about action values, provided by the limbic loop, can be integrated with information about current goals, mediated by the associative/cognitive loop, to dynamically control goal-directed behavior, as has been suggested by recent computational models^{10,36}.

An interesting finding of our *in-vivo* imaging data was while ensembles of NAc D2-MSNs activated by the negative outcome (reward omission) of a response error remained consistent throughout task learning, ensembles of NAc D1-MSNs activated by the rewarding outcome (reward delivery) of a correct response changed across task learning. A potential explanation for these activity patterns may be that NAc signalling does not simply encode the value of a specific outcome, but rather encodes more complex information about the value of outcomes associated with specific cues. Thus, it could be speculated that the random nature of the reward-associated cue, but not the non-reward-associated cue, in our avoidance-based visual discrimination (VD-Avoid) task resulted in the changing patterns of activated NAc D1-MSNs, but not NAc D2-MSNs. Alternatively, it is possible that rewarding outcomes are signalled by more general, summated patterns of activity than negative outcomes. Indeed, recent evidence suggests that NAc D1-MSN activity controls generalized Pavlovian learning of cue-outcome associations, while NAc D2-MSNs contribute to the ability to discriminate between Pavlovian cues⁵. A further surprising finding of our *in-vivo* imaging study was that, contrary to canonical role of NAc D1-MSNs in reward signaling^{13,44,45,55,56}, a small subset of NAc D1-MSNs were found to be activated by the reward omission. These data hint at heterogeneous functionality of NAc D1-MSN subpopulations and suggest the potential existence of NAc D1-MSNs responsive to negative outcomes, as has recently been reported in dorsal striatal D1-MSNs⁵⁷. Future studies investigating the precise activity patterns of ensembles of NAc D1- and D2-MSNs in response to a variety of rewarding and negative stimuli will likely help to fully elucidate the roles of these neural populations in value signaling.

Finally, impairments in the ability for response control are associated with risk-taking behaviors in drug addiction and attention-deficit/ hyperactivity disorder^{49,58-60}. A previous study showed D2-MSN activation in the NAc decreased risky choices in the risk-seeking rats, suggesting that D2-MSN activity in the NAc is important for avoiding risk-taking behavior⁶¹. Our results largely support their data and extend them to show that NAc D2-MSN activity is involved in suppressing the action associated with the negative outcome as well as a risky choice. Additionally, their data showing that the probability of making a risky choice decreases on the next trial after failing to obtain a reward and that optogenetic activation of NAc D2-MSNs reduces risky behavior are in agreement with our findings. Repeated cocaine treatment has also been shown to reduce the frequency of miniature excitatory postsynaptic currents in D2-MSNs⁶². This study fits with a model in which increased excitability of D2-MSNs leads to the strategy to avoid a bad option, while decrease in excitability of D2-MSN causes disability to avoid a bad option^{49,56,63}. These bidirectional effects on the strategy support our hypothesis that activation of D2-MSN plays an important role in avoiding a bad option.

In conclusion, we provide evidence that NAc D2-MSNs contribute to the ability to use environmental cues to guide avoidance behavior. Moreover, activation of D2-MSNs in the NAc by response errors plays an important role in execution of a strategy to avoid an undesirable response. These findings indicate that modulating the neural activity of D2-MSN in the striatum including the NAc by D2 receptor-selective drugs may be beneficial for the treatment of disorders associated with impaired ability for inhibitory control, such as drug addiction and attention-deficit/ hyperactivity disorder^{49,58-60}. In addition, our findings suggest that the VD-Avoid task is a useful paradigm for investigating the neural mechanisms that underly avoidance-based goal-directed behavior.

Methods

Animals

Wild-type C57BL/6J mice (male, 8-10 weeks old) were used for validation of behavioral experiments. For DREADD, optogenetic, and calcium imaging experiments, male heterozygous D1-Cre (FK150Gsat) and D2-Cre (ER44Gsat) mice were used (DREADD, D1-Cre, n = 12, one mouse was excluded due to no viral expression; D2-Cre, n = 12, one mouse was excluded due to unstable behavior; optogenetics, D2-Cre, n = 19, 2 mice were excluded due to insufficient conditioning; calcium imaging, D1-Cre, n = 3; D2-Cre, n = 4, one mouse was excluded because of incorrect GRIN lens placement). D1-Cre and D2-Cre were maintained in a C57BL/6J background. Animals were housed on a 12-hour light/dark cycle. Behavioral studies were conducted during the light cycle. Mice were kept on water restriction during behavioral testing. For all behavioral experiments except for calcium imaging experiment, mice were grouped housed throughout the experiments. For calcium imaging experiments, mice were singly housed after GRIN lens implantation. All experiments conformed to the guidelines of the National Institutes of Health experimental procedures, and were approved by the Animal Experimental Committee of Institute for Protein Research at Osaka University (approval ID 29-02-1).

Stereotaxic Surgery

All mice used in this study were anesthetized with ketamine (100 mg/kg) and xylazine (20 mg/kg) for surgical procedures and placed in a stereotaxic frame (Kopf Instruments).

For DREADD experiments, heterozygous D1-Cre or D2-Cre mice were bilaterally injected with 400 nl of AAV5-hSyn-DIO-hM4Di(Gi)-mCherry (5.1×10^{12} GC/ml, UNC) or AAV5-hSyn-DIO-mCherry (5.2×10^{12} GC/ml, UNC) using a Nanoject II instrument (Drummond) at a rate of 100 nl/min (coordinates in mm: AP +1.20, ML ± 1.25 from bregma, and DV -3.50 from brain surface). The injection pipette remained in place for 5–10 min to reduce backflow.

For optogenetics experiments, heterozygous D2-Cre mice were bilaterally injected with 400 nl of AAV5-CAG-FLEX-ArchT3.0-tdTomato (1.3×10^{13} GC/ml, Addgene) or AAV5-EF1a-DIO-eYFP (1.3×10^{13} GC/ml, Addgene) were using a Nanoject III instrument (Drummond) at a rate of 100 nl/min (coordinates in mm: AP +1.20, ML ± 1.25 from bregma, and DV -3.50 from brain surface). The injection pipette remained in place for 5–10 min to reduce backflow. After retraction, 200 μ m diameter (NA 0.37) optic fibers (Thorlabs) were bilaterally implanted and fixed in place with the dental cement (Superbond) at AP +1.20, ML ± 1.30 from bregma, and DV -3.20 from brain surface.

For calcium imaging experiments, heterozygous D1-Cre or D2-Cre mice were unilaterally injected with 1200 nl of AAV9-FLEX-jGCaMP7f (9.6×10^{12} GC/ml, Addgene) were stereotaxically injected using a Nanoject III instrument (Drummond) at a rate of 100 nl/min (coordinates in mm: AP +1.20, ML ± 1.25 from bregma, and DV -3.60 and -3.10 from brain surface). The injection pipette remained in place for 5–10 min to reduce backflow. After virus injection, a sterile 21-gauge needle was slowly lowered into the brain to a depth of -2.0 mm from the brain surface to aspirate brain tissue above the NAc. A GRIN lens (600 μ m diameter, Inscopix) was slowly lowered into the brain to a depth of -3.20 mm from the brain surface by using a GRIN lens holder (Inscopix). We secured the GRIN lens to the skull with dental cement (Superbond). A silicone elastomer (Kwik-Cast; World Precision Instruments) was applied to the top of the lens to prevent external damage. Four-to-six weeks after lens implantation, a baseplate (Inscopix) attached to the miniature microscope (nVista; Inscopix) was positioned above the GRIN lens. The focal plane was adjusted until blood vessels could be clearly observed. After adjustment, the baseplate was secured in place with the dental cement.

Behavioral Experiments

Apparatus. Training and testing were conducted in a Bussey-Saksida touchscreen chamber (Lafayette Instrument). A black plastic mask with 2 windows (70 \times 75 mm² spaced, 5 mm apart, 16 mm

above the floor) was placed in front of the touchscreen. ABET II and WhiskerServer software (Lafayette) were used to control operant system and data collection.

Pretraining. As the first phase (3 days), mice were habituated to the chamber in 40-min sessions. Diluted condensed milk (7 μ l, Morinaga Milk) was dispensed in the reward magazine every 10 sec. In the following phase (1 day), a stimulus was randomly displayed in 1 of the 2 windows. After a 30-sec stimulus presentation, the milk reward (20 μ l) was delivered with a tone (3 kHz) and the inside of the magazine was illuminated. When mice collected the reward, the magazine light went out, and the next trial commenced (60 trials, or up to 60 min) with a new stimulus after a 20-sec intertrial interval (ITI). In the next phase, stimuli were randomly displayed in one of 2 windows, and mice were obligated to touch the stimulus to receive a reward. In the final phase of the pretraining, when a blank window was touched, mice were punished with a 5-sec time-out. After reaching criterion (77% correct for 2 consecutive days), mice moved on to basic training.

Basic training. Mice were tested 5–6 days per week (60 trials per day, or up to 60 min). Each trial was initiated after mice nose-poked in the magazine. Visual cues were presented until mice responded at either window.

For the VD-Attend task, two visual cues (marble and a random image) were presented in the touchscreen. The random image was pseudorandomly chosen from 51 images. If the mouse responded to the correct (marble) visual cue, a milk reward (7 μ l) was delivered with a tone (3 kHz) and the magazine was illuminated. When mice collected the reward, the magazine light went out, and the next trial commenced (60 trials, or up to 60 min) with a new stimulus after a 20-sec intertrial interval (ITI). If the mouse responded to the incorrect (random) visual cue, the mouse was punished with a 5-sec time-out (house light on).

For the VD-Avoid task, two visual cues (flag and a random image) were presented in the touchscreen. If the mouse responded to the correct (random) visual cue, a milk reward (7 μ l) was delivered with a tone (3 kHz) and the magazine was illuminated. If the mouse responded to the incorrect (flag) visual cue, the mouse was punished with a 5-sec time-out (house light on).

A response at a random image was recorded as a correct response, while a response to visual cue “Flag” was recorded as an incorrect response.

After reaching criterion (>80% correct for 2 consecutive days), mice moved on to the test phase for DREADD experiments or cable habituation for optogenetic experiments, respectively.

For DREADD experiments, vehicle on days 1 and 3, or CNO (1.0 mg/kg diluted with vehicle, Sigma Aldrich) on days 2 and 4, was intraperitoneally administered 30 min before the session.

For optogenetic inhibition experiments, once the performance stabilized (>80% correct for 2 consecutive days) with the fiber optic cables attached, optogenetic stimulation experiments were commenced. LED power was set to 1-3 mW. Stimulation schedule was counterbalanced.

For calcium imaging experiments, data was acquired at 20 Hz with 0.6 mW LED at the first session of the basic training (Novice) and the session after reaching criterion (the criterion session; Expert). After acquisition, calcium recording files were temporally (factor of 2) and spatially (factor of 4) downsampled and motion-corrected using Inscopix Data Processing software ver 1.3.0. The fluorescent traces of individual neurons were extracted from these images by CNMFe³⁰. Z-scores were calculated from all recording data. The first three principal components (PCs) of the Z-scores of all neurons in the correct and error trials were calculated using principal component analysis (PCA), with the singular value decomposition algorithm. Hierarchical clustering of the first three PCs was then performed using a Euclidean distance metric and a complete agglomeration method.

Percentage correct (correct trials divided by correct plus incorrect trials, recorded as percent), and latencies to correct response, incorrect response, and reward collection were monitored in all behavioral experiments.

Histology

Animals were deeply anesthetized and transcardially perfused with 0.01 M PBS followed by 4% paraformaldehyde (PFA) in 0.1 M PB (pH 7.4). Brains were removed and post-fixed with 4% PFA at 4 °C for 2 days. After cryoprotection, brains were embedded in OCT compound and cryosectioned (40 µm). Sections were mounted with antifade mounting medium with DAPI (Vectashield). Stitched images were acquired using a Keyence BZ-X800 microscope.

Statistical analyses

Prism (Graphpad) software was used for statistical analyses. The behavioral performances in wild-type were analyzed using unpaired t-test. DREADD data were analyzed using two-way RM ANOVA with Group (hM4Di, mCherry) and Drug Treatment (vehicle, CNO) or one-way ANOVA (D1-hM4Di, D2-hM4Di, D1/D2-mCherry). Optogenetic data were analyzed using two-way RM ANOVA with Group (ArchT, eYFP) and Light stimulation (OFF, ON) or History (After Correct, After Error) and Light stimulation (OFF, ON). Post hoc Sidak's multiple comparisons test was performed when F-ratios were significant ($p < 0.05$). Comparisons of proportion of cells were made using chi-squared test. Simple linear regression was made to calculate the correlation of auROC between novice and expert. All data are expressed as means \pm SEM.

Declarations

The authors declare no competing financial interests.

References

1. Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263 (1979).
2. Rasmussen, E. B. & Newland, M. C. Asymmetry of reinforcement and punishment in human choice. *J Exp Anal Behav* 89, 157–67 (2008).
3. Kubanek, J., Snyder, L. H. & Abrams, R. A. Reward and punishment act as distinct factors in guiding behavior. *Cognition* 139, 154–67 (2015).
4. Horner, A. E. *et al.* The touchscreen operant platform for testing learning and memory in rats and mice. *Nat Protoc* 8, 1961–1984 (2013).
5. Iino, Y. *et al.* Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature* 1–6 (2020).
6. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* 521, (2015).
7. Balleine, B. W. The Meaning of Behavior: Discriminating Reflex and Volition in the Brain. *Neuron* 104, 47–62 (2019).
8. Floresco, S. B. The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annu Rev Psychol* 66, 25–52 (2015).
9. Macpherson, T. *et al.* Parallel and hierarchical neural mechanisms for adaptive and predictive behavioral control. *Neural Networks* 144, 507–521 (2021).
10. Mannella, F., Gurney, K. & Baldassarre, G. The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Front Behav Neurosci* 7, 135 (2013).
11. Kupchik, Y. M. *et al.* Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nat Neurosci* 18, 1230–1232 (2015).
12. Lu, X. Y., Ghasemzadeh, M. B. & Kalivas, P. W. Expression of D1 receptor, D2 receptor, substance P and enkephalin messenger RNAs in the neurons projecting from the nucleus accumbens. *Neuroscience* 82, 767–780 (1997).
13. Hikida, T., Kimura, K., Wada, N., Funabiki, K. & Nakanishi, S. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* 66, 896–907 (2010).
14. Yawata, S., Yamaguchi, T., Danjo, T., Hikida, T. & Nakanishi, S. Pathway-specific control of reward learning and its flexibility via selective dopamine receptors in the nucleus accumbens. *Proc National Acad Sci* 109, 12764–12769 (2012).
15. Soares-Cunha, C., Coimbra, B., Sousa, N. & Rodrigues, A. J. Reappraising striatal D1- and D2-neurons in reward and aversion. *Neurosci Biobehav Rev* 68, 370–386 (2016).
16. Rabbitt, P. & Rodgers, B. What does a Man do after he Makes an Error? An Analysis of Response Programming. *Q J Exp Psychol* 29, 727–743 (1977).

17. Notebaert, W. *et al.* Post-error slowing: An orienting account. *Cognition* 111, 275–279 (2009).
18. Roitman, J. D. & Loriaux, A. L. Nucleus accumbens responses differentiate execution and restraint in reward-directed behavior. *J Neurophysiol* 111, 350–360 (2014).
19. Ambroggi, F., Ghazizadeh, A., Nicola, S. M. & Fields, H. L. Roles of Nucleus Accumbens Core and Shell in Incentive-Cue Responding and Behavioral Inhibition. *J Neurosci* 31, 6820–6830 (2011).
20. Taha, S. A. & Fields, H. L. Inhibitions of Nucleus Accumbens Neurons Encode a Gating Signal for Reward-Directed Behavior. *J Neurosci* 26, 217–222 (2006).
21. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc National Acad Sci* 104, 5163–5168 (2007).
22. Wulff, P. & Arenkiel, B. R. Chemical genetics: receptor–ligand pairs for rapid manipulation of neuronal activity. *Curr Opin Neurobiol* 22, 54–60 (2012).
23. Luo, Y. J. *et al.* Nucleus accumbens controls wakefulness by a subpopulation of neurons expressing dopamine D1 receptors. *Nat Commun* 9, 1576 (2018).
24. Parkes, S. L., Bradfield, L. A. & Balleine, B. W. Interaction of Insular Cortex and Ventral Striatum Mediates the Effect of Incentive Memory on Choice Between Goal-Directed Actions. *J Neurosci* 35, 6464–6471 (2015).
25. Cardinal, R. N. & Cheung, T. H. Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat. *Bmc Neurosci* 6, 9 (2005).
26. Gallo, E. F. *et al.* Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nat Commun* 9, 1086 (2018).
27. Schultz, W., Apicella, P., Scarnati, " Eugenio & Ljungbergc, T. Neuronal Activity in Monkey Ventral Striatum Related to the Expectation of Reward. *The Journal of Neuroscience* 12, 4595–4610 (1992).
28. Roitman, M. F., Wheeler, R. A. & Carelli, R. M. Nucleus Accumbens Neurons Are Innately Tuned for Rewarding and Aversive Taste Stimuli, Encode Their Predictors, and Are Linked to Motor Output. *Neuron* 45, 587–597 (2005).
29. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat Methods* 16, 649–657 (2019).
30. Zhou, P. *et al.* Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *Elife* 7, e28728 (2018).
31. Sheintuch, L. *et al.* Tracking the Same Neurons across Multiple Days in Ca²⁺ Imaging Data. *Cell Reports* 21, 1102–1115 (2017).
32. Han, X. *et al.* A High-Light Sensitivity Optical Neural Silencer: Development and Application to Optogenetic Control of Non-Human Primate Cortex. *Frontiers Syst Neurosci* 5, 18 (2011).
33. Kang, B. J. *et al.* Effect of optogenetic manipulation of accumbal medium spiny neurons expressing dopamine D2 receptors in cocaine-induced behavioral sensitization. *Eur J Neurosci* 46, 2056–2066 (2017).

34. Ikemoto, S. & Panksepp, J. The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res Rev* 31, 6–41 (1999).
35. Reynolds, S. M. & Berridge, K. C. Positive and Negative Motivation in Nucleus Accumbens Shell: Bivalent Rostrocaudal Gradients for GABA-Elicited Eating, Taste “Liking”/“Disliking” Reactions, Place Preference/Avoidance, and Fear. *J Neurosci* 22, 7308–7320 (2002).
36. Mannella, F., Mirolli, M. & Baldassarre, G. Goal-Directed Behavior and Instrumental Devaluation: A Neural System-Level Computational Model. *Front Behav Neurosci* 10, 181 (2016).
37. Nicola, S. M. The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology* 191, 521–550 (2007).
38. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science (New York, N.Y.)* 275, 1593–1599 (1997).
39. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88 (2012).
40. Jong, J. W. de *et al.* A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. *Neuron* 101, 133-151.e7 (2019).
41. Lee, S. J. *et al.* Cell-type-specific asynchronous modulation of PKA by dopamine in learning. *Nature* 1–6 (2020).
42. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity. *Science* 321, 848–851 (2008).
43. Surmeier, D. J., Ding, J., Day, M., Wang, Z. & Shen, W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci* 30, 228–235 (2007).
44. Lobo, M. K. *et al.* Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science (New York, N.Y.)* 330, 385–390 (2010).
45. Macpherson, T., Morita, M. & Hikida, T. Striatal direct and indirect pathways control decision-making behavior. *Frontiers in psychology* 5, 1301 (2014).
46. Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S. & Nakanishi, S. Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens. *Proceedings of the National Academy of Sciences of the United States of America* 111, 6455–6460 (2014).
47. Klawonn, A. M. & Malenka, R. C. Nucleus Accumbens Modulation in Reward and Aversion. *Cold Spring Harb Sym* 83, 119–129 (2018).
48. Frank, M. J., Samanta, J., Moustafa, A. A. & Sherman, S. J. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* 318, 1309–1312 (2007).
49. Frank, M. J., Seeberger, L. C. & O’Reilly, R. C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943 (2004).

50. Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* 15, 1281–1289 (2012).
51. Bornstein, A. M. & Daw, N. D. Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr Opin Neurobiol* 21, 374–380 (2011).
52. Penner, M. R. & Mizumori, S. J. Y. Neural systems analysis of decision making during goal-directed navigation. *Prog Neurobiol* 96, 96–135 (2012).
53. Aoki, S. *et al.* An open cortico-basal ganglia loop allows limbic control over motor output via the nigrothalamic pathway. *Elife* 8, e49995 (2019).
54. Foster, N. N. *et al.* The mouse cortico-basal ganglia-thalamic network. *Biorxiv* 2020.10.06.326876 (2020).
55. Cole, S. L., Robinson, M. J. F. & Berridge, K. C. Optogenetic self-stimulation in the nucleus accumbens: D1 reward versus D2 ambivalence. *Plos One* 13, e0207694 (2018).
56. Nakanishi, S., Hikida, T. & Yawata, S. Distinct dopaminergic control of the direct and indirect pathways in reward-based and avoidance learning behaviors. *Neuroscience* 282, 49–59 (2014).
57. Xiao, X. *et al.* A Genetically Defined Compartmentalized Striatal Direct Pathway for Negative Reinforcement. *Cell* 183, 211-227.e20 (2020).
58. Strickland, J. C., Bolin, B. L., Lile, J. A., Rush, C. R. & Stoops, W. W. Differential sensitivity to learning from positive and negative outcomes in cocaine users. *Drug Alcohol Depen* 166, 61–68 (2016).
59. Robbins, T. W., Gillan, C. M., Smith, D. G., Wit, S. de & Ersche, K. D. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci* 16, 81–91 (2012).
60. Constantinople, C. M., Piet, A. T. & Brody, C. D. An Analysis of Decision under Risk in Rats. *Curr Biol* 29, 2066-2074.e5 (2019).
61. Zalocusky, K. A. *et al.* Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* 531, 642 646 (2016).
62. Kim, J., Park, B. H., Lee, J. H., Park, S. K. & Kim, J. H. Cell Type-Specific Alterations in the Nucleus Accumbens by Repeated Exposures to Cocaine. *Biol Psychiat* 69, 1026–1034 (2011).
63. Frank, M. J. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks* 19, 1120–1136 (2006).

Figures

Figure 1

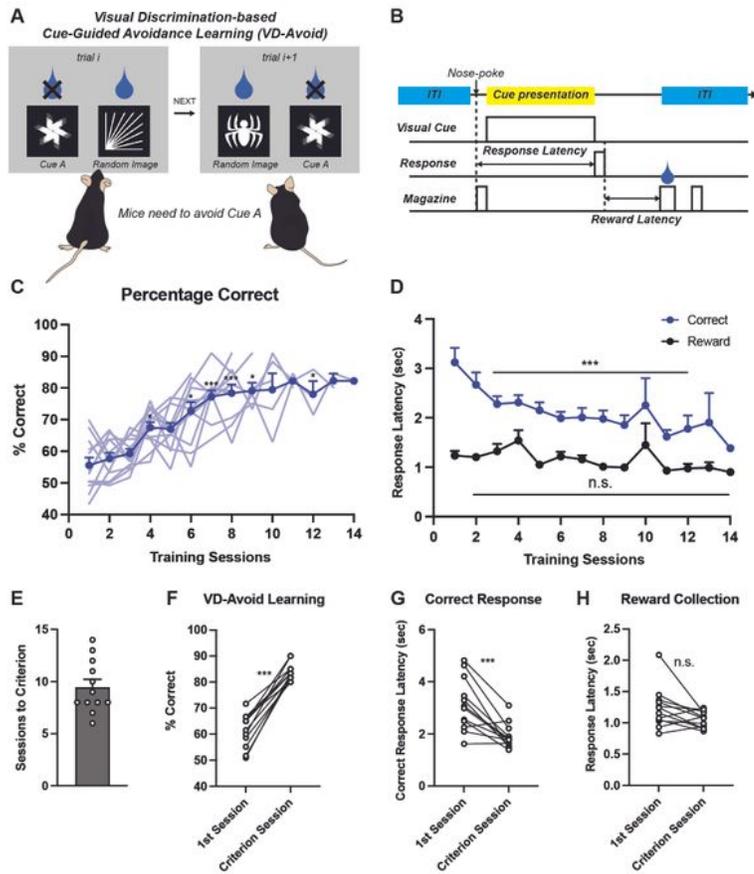


Figure 1

Experimental Design and Behavioral Performance

(A) Experimental design.

(B) Timeline of the task events and the definition of the behavioral parameters.

(C) The percentage of correct in each session (n = 12, One-way RM-ANOVA with Sidak correction, *p < 0.05 vs Session 1; **p < 0.01 vs Session 1, ***p < 0.001 vs Session 1).

(D) Correct response latency and reward collection latency in each session (Two-way RM-ANOVA with Sidak correction, *p < 0.05 vs Session 1; **p < 0.01 vs Session 1, ***p < 0.001 vs Session 1).

(E) Sessions to criterion.

(F) The percentage of correct in the 1st and the last session (paired t-test, $t_{11} = 9.799$, ***p < 0.0001).

(G) Correct response latency in the 1st and the last session (paired t-test, $t_{11} = 4.616$, ***p = 0.0007).

(H) Reward collection latency in the 1st and the last session (paired t-test, $t_{11} = 2.077$, p = 0.0620).

Data are presented as mean \pm SEM.

Figure 2

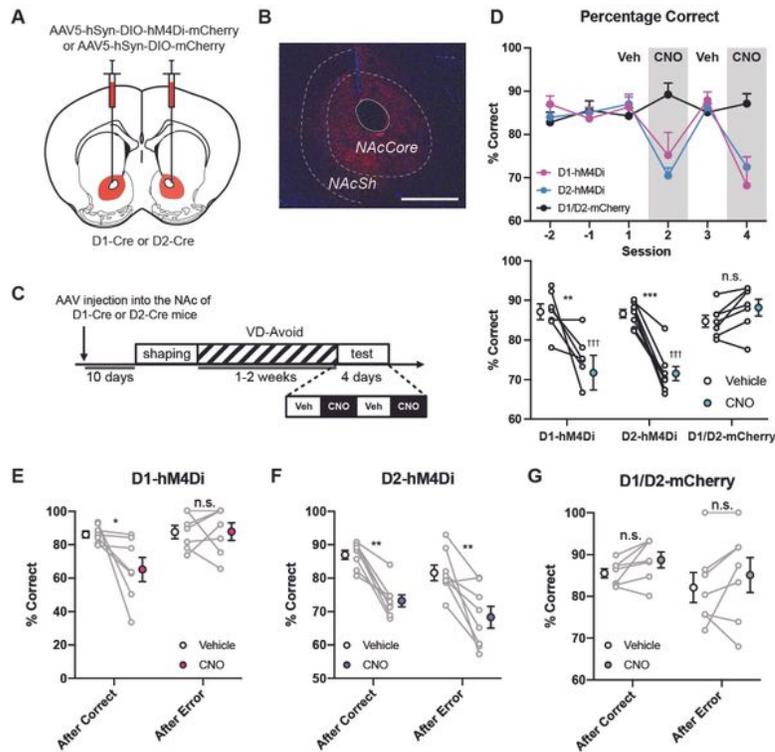


Figure 2

DREADD Suppression of D1-MSNs and D2-MSNs in the NAc Decreases the Performance of the VD-Avoid Task through Different Mechanisms

(A) A schematic of viral injection.

(B) Representative injection site of AAV5-hSyn-DIO-hM4Di-mCherry. Scale bar, 500 μ m. NAcCore, nucleus accumbens core; NAcSh, nucleus accumbens shell.

(C) Experimental timeline.

(D) The effects of DREADD suppression on the behavioral performance of the VD-Avoid task in D1-Cre mice (n = 7), D2-Cre mice (n = 8), and D1/D2-mCherry mice (n = 7) in each session (Top). DREADD suppression of D1-MSN and D2-MSN in the NAc decreased the behavioral performance in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,19} = 20.06$, $***p = 0.0003$; D1-hM4Di, $**p = 0.0011$; D2-hM4Di, $***p = 0.0007$; D1/D2-mCherry, $p = 0.7209$; D1-hM4Di vs D1/D2-mCherry in CNO, $†††p < 0.0001$; D2-hM4Di vs D1/D2-mCherry in CNO, $†††p < 0.0001$).

(E) DREADD suppression of D1-MSNs in the NAc decreased the behavioral performance after correct trials in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type \times Treatment interaction, $F_{1,12} = 5.603$, $*p = 0.0356$; After Correct, $*p = 0.0133$; After Error, $p = 0.9990$).

(F) DREADD suppression of D2-MSNs in the NAc decreased the behavioral performance after correct and error trials in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,14} = 29.79$, $***p < 0.0001$; After Correct, $**p = 0.0012$; After Error, $**p = 0.0014$).

(G) CNO treatment in D1/D2-mCherry mice did not affect the behavioral performance after correct or error trial in the VD-Avoid task (Two-way RM-ANOVA with Sidak correction, Trial type \times Treatment interaction, $F_{1,12} = 0.001573$, $p = 0.9690$; After Correct, $p = 0.7057$; After Error, $p = 0.7319$).

Data are presented as mean \pm SEM.

Figure 3

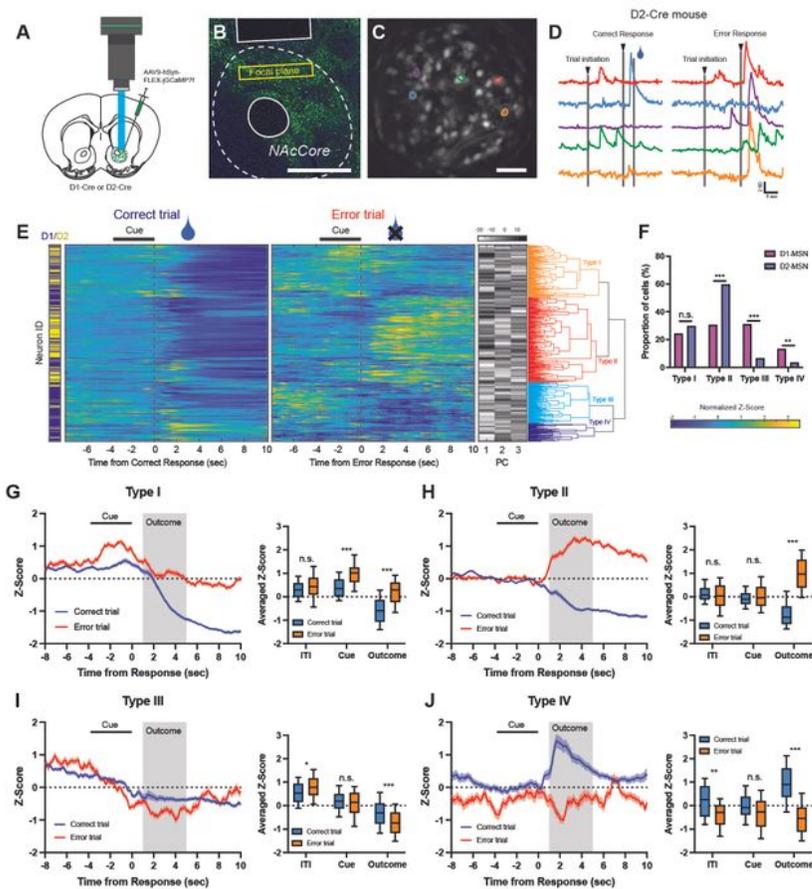


Figure 3

Individual D2-MSNs Are Predominantly Activated by Response Errors

(A) A schematic of viral injection and GRIN lens implantation.

(B) Representative coronal image of jRCaMP7f expression in D2-MSNs. Scale bar, 500 μ m.

(C) Maximum projection image of a representative imaging plane. Scale bar, 50 μm .

(D) Example traces of individual neurons from a representative mouse performing the VD-Avoid task.

(E) Heatmap of neuronal activity of all neurons recorded in the correct and error trials (Left). Each row represents one neuron. First three principal components (PC) and hierarchical clustering dendrogram showing the relationship of each neuron within the four clusters (Right).

(F) Proportion of each type of cells in D1-MSN and D2-MSN (Chi-square test, Type I, $p = 0.3393$; Type II, $***p < 0.001$, Type III, $***p < 0.001$; Type IV, $**p = 0.0053$).

(G) Averaged traces of Type I neuron in correct and error trials (Left). Averaged Z-score during ITI (-5-0 sec from trial initiation), Cue (0-2 sec from trial initiation), and Outcome (1-5 sec from response) period (Right; Two-way RM ANOVA with Sidak correction, ITI, $p = 0.1641$; Cue, $***p < 0.001$; Outcome, $***p < 0.001$).

(H) Averaged traces of Type II neuron in correct and error trials (Left). Averaged Z-score during ITI, Cue, Outcome period (Right; Two-way RM ANOVA with Sidak correction, ITI, $p = 0.1820$; Cue, $p = 0.5652$; Outcome, $***p < 0.001$).

(I) Averaged traces of Type III neuron in correct and error trials (Left). Averaged Z-score during ITI, Cue, Outcome period (Right; Two-way RM ANOVA with Sidak correction, ITI, $*p = 0.0106$; Cue, $p = 0.1826$; Outcome, $***p < 0.001$).

(J) Averaged traces of Type IV neuron in correct and error trials (Left). Averaged Z-score during ITI, Cue, Outcome period (Right; Two-way RM ANOVA with Sidak correction, ITI, $**p = 0.0025$; Cue, $p = 0.2613$; Outcome, $***p < 0.001$).

Data are presented as mean \pm SEM. Boxes and whiskers show the respective median and the 25th to 75th and 10th to 90th percentiles.

Figure 4

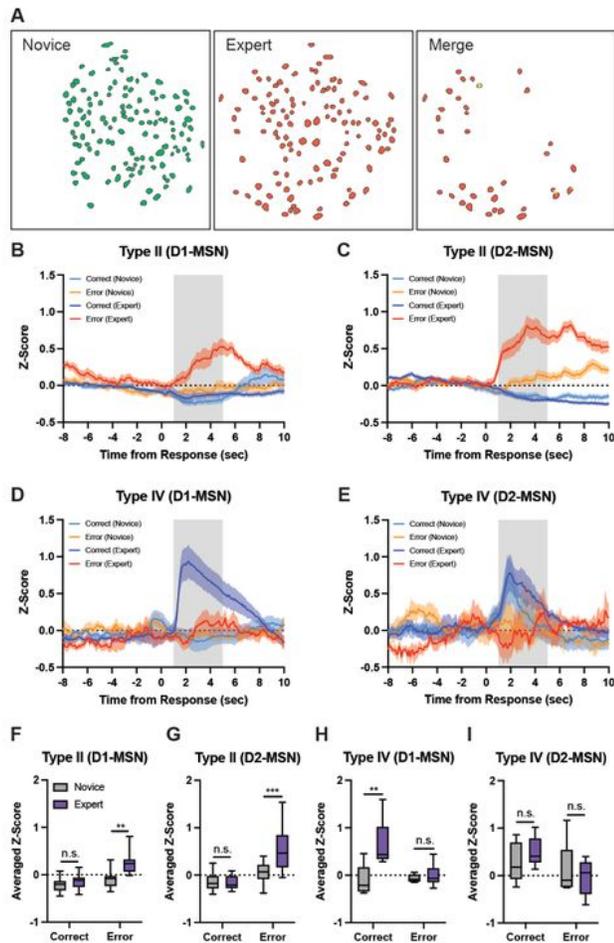


Figure 4

More Non-Responsive D2-MSNs Acquire Error Responses than D1-MSNs through Learning

(A) Representative image of contour map from novice (green) and expert (red).

(B) Averaged traces of Type II in D1-MSNs acquiring error responses through learning. Averaged calcium traces of novices and experts in correct and error trials.

(C) Averaged traces of Type II in D2-MSNs acquiring error responses through learning. Averaged calcium traces of novices and experts in correct and error trials.

(D) Averaged traces of Type IV in D1-MSNs acquiring correct response through learning. Averaged calcium traces of novices and experts in correct and error trials.

(E) Averaged traces of Type IV in D2-MSNs through learning. Averaged calcium traces of novices and experts in correct and error trials.

(F) Averaged Z-score of Type II in D1-MSN during Outcome period (Two-way RM ANOVA with Sidak correction, Correct, $p = 0.6853$; Error, $**p = 0.0011$).

(G) Averaged Z-score of Type II in D2-MSN during Outcome period (Two-way RM ANOVA with Sidak correction, Correct, $p = 0.9609$; Error, $***p < 0.0001$).

(H) Averaged Z-score of Type IV in D1-MSN during Outcome period (Two-way RM ANOVA with Sidak correction, Correct, $**p = 0.0038$; Error, $p = 0.8951$).

(I) Averaged Z-score of Type IV in D2-MSN during Outcome period (Two-way RM ANOVA with Sidak correction, Correct, $p = 0.3851$; Error, $p = 0.6701$).

Data are presented as mean \pm SEM. Boxes and whiskers show the respective median and the 25th to 75th and 10th to 90th percentiles. Clustering types were defined by criterion session data.

Figure 5

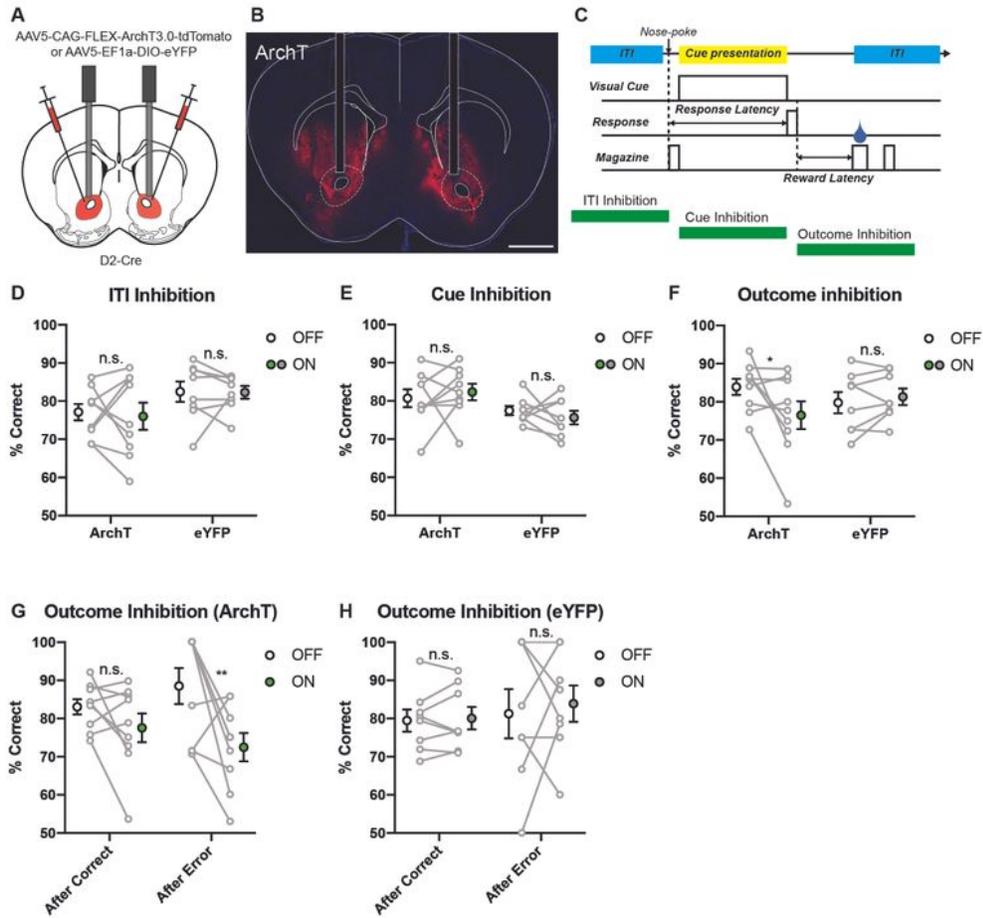


Figure 5

Post-Error Activation of D2-MSNs is Necessary for Avoidance-Based Goal-Directed Behavior

(A) Schematic of viral injection and optic fiber implantation.

(B) Representative coronal section with optic fibers. Scale bar, 1 mm.

(C) Schematic of optical stimulation protocol.

(D) Optogenetic suppression of D2-MSNs in the NAc during ITI period did not affect the behavioral performance (Two-way RM-ANOVA with Sidak correction, Group \times Treatment interaction, $F_{1,15} = 0.05089$, $p = 0.8246$; ArchT, $n = 9$, $p = 0.9121$; eYFP, $n = 8$, $p = 0.9981$).

(E) Optogenetic suppression of D2-MSNs in the NAc during Cue period did not affect the behavioral performance (Two-way RM-ANOVA with Sidak correction, Group \times Treatment interaction, $F_{1,15} = 0.8700$, $p = 0.3657$; ArchT, $p = 0.7815$; eYFP, $p = 0.7571$).

(F) Optogenetic suppression of D2-MSNs in the NAc during Outcome period decreased the behavioral performance in the next trial (Two-way RM-ANOVA with Sidak correction, Group \times Treatment interaction, $F_{1,15} = 5.340$, $*p = 0.0355$; ArchT, $p = 0.0271$; eYFP, $p = 0.8368$).

(G and H) Optical stimulation of D2-MSNs in the NAc during Outcome period after error response decreased the behavioral performance of ArchT mice (G) in the next trial (Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,8} = 7.134$, $*p = 0.0283$; After Correct, $p = 0.2634$; After Error, $**p = 0.0029$) but not eYFP mice (H, Two-way RM-ANOVA with Sidak correction, Treatment effects, $F_{1,7} = 0.2682$, $p = 0.6205$; After Correct, $p = 0.9909$; After Error, $p = 0.8477$).

Data are presented as mean \pm SEM.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS11210907.jpg](#)
- [FigureS12optogeneticsschedule.jpg](#)
- [FigureS13210907.jpg](#)
- [FigureS14210907.jpg](#)
- [FigureS1210821.jpg](#)
- [FigureS2210907.jpg](#)
- [SupplementaryMovieS1Correct.mp4](#)
- [FigureS3210821.jpg](#)
- [SupplementaryMovieS2Error.mp4](#)
- [FigureS4210907.jpg](#)
- [FigureS5210907.jpg](#)
- [FigureS6210907.jpg](#)
- [FigureS8211024.jpg](#)

- [FigureS9211117.jpg](#)
- [FigureS12optogeneticsschedule.jpg](#)
- [FigureS3210821.jpg](#)
- [FigureS7imagingschedule211013.jpg](#)