Archival Report

Activity in the Dorsomedial Striatum Underlies Serial Reversal Learning Performance Under Probabilistic Uncertainty

Madison K. Young, Kyna-Anne Conn, Joyosmita Das, Simin Zou, Suzy Alexander, Thomas H.J. Burne, and James P. Kesby

ABSTRACT

BACKGROUND: Corticostriatal circuits, particularly the dorsomedial striatum (DMS) and lateral orbitofrontal cortex, are critical for navigating reversal learning under probabilistic uncertainty. These same areas are implicated in the reversal learning impairments observed in individuals with psychosis as well as their psychotic symptoms, suggesting that they may share a common neurobiological substrate. To address this question, we used psychostimulant exposure and specific activation of the DMS during reversal learning in mice to assess corticostriatal activity.

METHODS: We used amphetamine treatment to induce psychosis-relevant neurobiology in male mice during reversal learning and to examine pathway-specific corticostriatal activation. To determine the causal role of DMS activity, we used chemogenetics to drive midbrain inputs during a range of probabilistic contingencies.

RESULTS: Mice treated with amphetamine showed altered punishment learning, which was associated with decreased shifting after losses and increased perseverative errors after reversals. Reversal learning performance and strategies were dependent on increased activity in lateral orbitofrontal cortex to DMS circuits as well as in the DMS itself. Specific activation of midbrain to DMS circuits also decreased shifting after losses and reversal learning performance. However, these alterations were dependent on the probabilistic contingency.

CONCLUSIONS: Our work suggests that the DMS plays a multifaceted role in reversal learning. Increasing DMS activity impairs multiple reversal learning processes dependent on the level of uncertainty, confirming its role in the maintenance and selection of incoming cortical inputs. Together, these outcomes suggest that elevated dopamine levels in the DMS could contribute to decision-making impairments in individuals with psychosis.

https://doi.org/10.1016/j.bpsgos.2022.08.005

Cognitive flexibility is an essential executive function that allows an organism to appropriately adapt behaviors in response to changes in the external environment. Impairments in cognitive flexibility have been observed in a range of neuropsychiatric disorders including schizophrenia, Parkinson’s disease, obsessive-compulsive disorder, and substance abuse (1–4). The pathophysiology of these deficits is commonly investigated via reversal learning paradigms that have been developed for use across species, including humans, nonhuman primates, and rodents (5).

Cognitive dysfunction is a core deficit in individuals with schizophrenia (6). These cognitive impairments often include disturbances in decision making, including difficulties learning from positive and negative feedback that can manifest as maladaptive reward learning (7). This has been confirmed in multiple studies that demonstrate that people with schizophrenia achieve significantly fewer reversals than healthy control subjects in reversal learning tasks (3,5–10). Behavioral impairments are often paired with a reduction in the use of optimal trial-by-trial strategies, such as win-stay (selecting the same stimulus after a win) and lose-shift (selecting the opposing stimulus after a loss), implemented by individuals to aid in task navigation (8). Impairments in reversal learning have been observed in both the initial discrimination phase and the subsequent reversal phases (3, 8,11–13).

The etiology of schizophrenia and its accompanying cognitive deficits is still unknown, but evidence highlights a role for subcortical dopamine systems in reward learning and reinforcement (7). An area of particular interest is the associative striatum (dorsomedial striatum [DMS] in rodents) (1–4), which is the site of increased dopamine function in psychosis and an important area in goal-directed action and decision making (15). Functional neuroimaging studies in healthy subjects support a role for subcortical dopamine in reversal learning. For example, increased dopamine receptor availability in the caudate nucleus (associative striatum) has been shown to correlate with more reversal learning errors (16).

Typically, probabilistic reversal learning (PRL) paradigms are used to study cognitive flexibility, with stimuli presented as both high (80%) and low (20%) probability reward outcomes (80:20 contingency). This requires subjects to make decisions in the presence of misleading feedback, allowing accumulated
evidence across previous trials to guide choices. Once a set performance criterion has been achieved, the high- and low-rewarded stimuli are reversed, prompting the subject to adjust their behavior based on reward-outcome feedback. Numerous studies, both clinical and preclinical, have identified the importance of, and the complex relationship between, the striatum and the orbitofrontal cortex during the serial reversal learning phases of PRL tasks (7,17–22). However, despite the clinical relevance of reversal learning, low performance levels in rodents (23,24) and the complex role of the DMS have constrained prior studies.

In this study, we aimed to 1) establish a PRL task in mice that featured more reversals per session and above-chance performance at a range of probabilistic reward contingencies; 2) identify how amphetamine treatment alters reversal learning and corticostriatal network activation; and 3) chemogenetically activate midbrain inputs to the DMS to determine its specific role in reversal learning under a variety of probabilistic contingencies. Our work highlights a multifaceted role for the DMS in reversal learning.

METHODS AND MATERIALS

Animals

For detailed protocols see Supplemental Methods. Ten-week-old male C57BL/6JArc mice (Animal Resources Centre) were used (31 for experiment 1 and 23 for experiment 2). All procedures were performed with approval from the University of Queensland Animal Ethics Committee (QBI/079/17).

PRL Protocol

PRL testing was conducted in 8 Plexiglas operand chambers (model ENV-307A; Med Associates Inc.). The chambers featured 2 nosepoke holes and a reward magazine located on the opposite wall, which could deliver strawberry-flavored milk. The target nosepoke was defined as that with the highest reward rate. No punishment followed an unrewarded response; rather, punishment/loss reflected a lack of expected reward. After 6 consecutive target nosepokes, the target and nontarget reward contingencies were reversed. Behavioral measures included the total number of trials, latency to respond, number of reversals per 100 trials, proportion of win-stay and lose-shift strategies, and number of perseverative errors (consecutive errors after a reversal).

Experiment 1: Amphetamine Treatment and Corticostriatal Networks

See Figure 1 (top panel) for a detailed experimental timeline.

Surgery and Tract-Tracing Procedures. Cholera toxin subunit B (CTb) conjugated with Alexa Fluor 555 (C-34776; Thermo Fisher Scientific) was injected into the right posterior DMS (0.5 μL; anteroposterior [AP], +0.01; medial-lateral [ML], −0.18; dorsoventral [DV], −0.29 [from the skull surface], in mm relative to bregma) (29) and a CTb conjugated with Alexa Fluor 647 (C-34778; Thermo Fisher Scientific) was injected into the left anterior DMS (0.5 μL; AP, +0.11; ML, +0.145; DV, −0.29).

Simulation and Computational Modeling. Simulations (5000 simulations of 500 trials) with random choices were coded in R (version 3.6) to quantify the likelihood of chance reversals. We modeled latent task variables using the hBayesDM package for R (26). A reward/punishment learning model with parameters for reward learning rate, punishment learning rate, and inverse temperature was selected (27).

Histology, Immunofluorescence, and cFos Quantification. Brain sections were stained for cFos and DAPI (Table S1 for details). Sections were imaged on a spinning-disk confocal system and 40× objective. Image acquisition was performed using SlideBook 6.0 (Si Inc.), Image preparation and regional masks (25) were completed using ImageJ (28). Nuclei were segmented with Cellpose (29) and fed into CellProfiler version 4.1 for quantification (30). cFos+ cells were determined as those with >1.1-fold mean cFos intensity when compared with the mean intensity of all cells per image.

Figure 1. Timeline of experimental manipulations. Top panel: Experiment 1. Mice were trained until PRL performance was stable and then underwent tract-tracing surgery. After recovery and rebaselining, mice were tested for each contingency (80:20, 70:30, 80:40, and 90:10), separated by 2 days. Following 5 additional days of baseline PRL testing, mice were administered amphetamine (1 mg/kg, intraperitoneally) or saline (15 minutes prior to testing) at an 80:20 contingency. Mice were perfused 60 minutes after completion of testing (i.e., 90 minutes after the middle of the reversal learning session) for subsequent cFos analysis. Bottom panel: Experiment 2. After DREADDs surgery and PRL training, mice underwent acute CNO administration (30 minutes prior to testing). This was conducted using a within-subject Latin-square design at the 80:20 contingency with CNO (0, 0.5, 1, and 2 mg/kg; in 0.5% DMSO and 9.9% saline). This was followed by a crossover design with CNO (0 and 2 mg/kg) at the 80:40 contingency. Subsequently, mice were given fresh water daily (0.25% DMSO) for 5 days at 80:20 followed by a water solution containing CNO (~8 mg/kg/day). CNO exposure was continued for 23 days as follows: 14 days at 80:20, 4 days at 80:40, and 5 days at 70:30. After testing was complete, a small cohort of mice were administered 2 mg/kg CNO intraperitoneally and perfused 2 hours later for cFos analysis. CNO, clozapine N-oxide; DMSO, dimethyl sulfoxide; DREADD, designer receptors exclusively activated by designer drug; PRL, probabilistic reversal learning.
The Dorsomedial Striatum Drives Reversal Learning in Mice

Experiment 2: Acute and Chronic Activation of Midbrain to DMS Cells With DREADDs

See Figure 1 (bottom panel) for a detailed experimental timeline.

Surgery for Pathway-Specific DREADDs. Surgery was performed as for experiment 1. Transfection of striatal pathways was achieved by bilaterally microinjecting Cre-dependent hM3D(Gq) flex-switch DREADDs (designer receptors exclusively activated by designer drugs) virus (pAAV5-hSyn-DIO-hM3D(Gq)-mCherry; Addgene; 0.5 µL) into the substantia nigra (AP, −3.10; ML, ±1.25; DV, −4.20) and a retrograde Cre-recombinase virus (pENN.AAV.hSyn.HI.eGFP-Cre.WPRE.SV40; Addgene; 0.7 µL) into the DMS (AP, +0.11; ML, ±1.25; DV, −2.90). A control virus not expressing DREADDs (pAAV5-hSyn-DIO-mCherry; Addgene) was used for half of the animals. From 23 mice, strong viral expression in both hemispheres was confirmed in 12 control and 6 DREADDs mice.

Histology and cFos Quantification. Sectioning and histology were performed as for experiment 1 (Table S1 for antibody details). Striatal sections were stained for mCherry and tyrosine hydroxylase. Midbrain sections were stained for mCherry and cFos. Images were taken at 20× magnification.

General Statistics

Analyses were performed using SPSS Statistics (version 26; IBM Corp.) and R (version 4.1.2). Data were analyzed using analysis of variance and repeated-measures analysis of variance when within-subject factors were present. Post hoc comparisons were performed using Sidák corrections. Graphical results are expressed as mean ± standard error of the mean (SEM).

RESULTS

A Minimum Level of Probabilistic Uncertainty Is Required to Impair Performance

By optimizing task parameters, we were able to demonstrate that mice can perform PRL across a range of contingencies. Mice performed the task well above chance, even at more difficult contingencies (Figure 2A). The total number of trials completed was not affected by differing contingencies (Figure 2B), but a main effect of contingency on the number of reversals per 100 trials was observed ($F_{3,28} = 19.5, p < .001$) (Figure 2C). Mice achieved fewer reversals at all contingencies compared with the 90:10 contingency ($p < .001$). This suggests that 90:10 contingencies do not provide much probabilistic uncertainty, and the 80:20 contingency is an effective probabilistic contingency for use in mice [and the most common in humans (5)]. This is further supported by the lack of differences in reversal performance when comparing deterministic (100:0, $6.45 \pm 0.26$ reversals/100 trials) with 90:10 (6.15 ± 0.30 reversals/100 trials) contingencies. Alterations in contingencies also affected choice strategy, with a significant difference in win-stay probability being observed ($F_{3,28} = 3.5, p < .05$) (Figure 2D). Mice had a decrease in win-stay probability during the 80:40 contingency compared with the 90:10 contingency ($p < .05$), commensurate with the increased misleading feedback for 80:40. In contrast, lose-shift probability was not significantly different between contingencies (Figure 2E). This confirms that mice are capable of navigating PRL well above chance when task parameters are optimized. We also assessed stability over 5 days at 80:20 (Figure S1), with mice demonstrating consistent performance and low variance in win-stay and lose-shift probability. These data parallel studies in humans, highlighting that PRL tasks have high retest stability (31). Taken together, this task protocol allows for within-session manipulations and the use of more uncertain/difficult contingencies in preclinical rodent studies.

Amphetamine Treatment Alters Punishment Learning

Psychomimetic drugs such as amphetamine are often used to model the increased dopaminergic function associated with psychosis (15,32–38). However, increasing brain dopaminergic function did not alter the number of reversals per 100 trials (Figure 3A), the response latency (Figure 3B), or the win-stay probability (Figure 3C) during reversal learning. However, amphetamine treatment led to a decrease in lose-shift probability ($F_{1,19} = 10.4, p < .01$) (Figure 3D) and an increase in the number of reversals per 100 trials (Figure 3E). The probability of being rewarded on the high/low lever was tested at 90:10, 80:20, 80:40, and 70:30 ($n = 23$). Mice performed well above chance when given the 90:10 and 80:20 contingencies (note: x-axis is not linear). Even at the more difficult 80:40 and 70:30 contingencies, more than 90% of mice performed above chance levels. Mice maintained a high number of trials per session under all contingencies (B). As the difference between high vs. low decreased (80:40 and 70:30), mice were less efficient at completing reversals (C). Win-stay probability (D) was significantly greater at 90:10 compared with 80:40, whereas lose-shift probability (E) was not significantly affected. Data are expressed as mean ± standard error. *$p < .05$, **$p < .001$. 

Figure 2. Mouse performance when compared with chance at various contingencies. The number of reversals (as a proportion of all mice/simulations) achieved during the contingency modifications (A). The probability of being rewarded on the high/low lever was tested at 90:10, 80:20, 80:40, and 70:30 ($n = 23$). Mice performed well above chance when given the 90:10 and 80:20 contingencies (note: x-axis is not linear). Even at the more difficult 80:40 and 70:30 contingencies, more than 90% of mice performed above chance levels. Mice maintained a high number of trials per session under all contingencies (B). As the difference between high vs. low decreased (80:40 and 70:30), mice were less efficient at completing reversals (C). Win-stay probability (D) was significantly greater at 90:10 compared with 80:40, whereas lose-shift probability (E) was not significantly affected. Data are expressed as mean ± standard error. *$p < .05$, **$p < .001$. 

Figure 3. Amphetamine treatment alters punishment learning. Amphetamine treatment led to a decrease in lose-shift probability ($F_{1,19} = 10.4, p < .01$) (Figure 3D) and an increase in the number of reversals per 100 trials (Figure 3E). The probability of being rewarded on the high/low lever was tested at 90:10, 80:20, 80:40, and 70:30 ($n = 23$). Mice performed well above chance when given the 90:10 and 80:20 contingencies (note: x-axis is not linear). Even at the more difficult 80:40 and 70:30 contingencies, more than 90% of mice performed above chance levels. Mice maintained a high number of trials per session under all contingencies (B). As the difference between high vs. low decreased (80:40 and 70:30), mice were less efficient at completing reversals (C). Win-stay probability (D) was significantly greater at 90:10 compared with 80:40, whereas lose-shift probability (E) was not significantly affected. Data are expressed as mean ± standard error. *$p < .05$, **$p < .001$. 

Figure S1. Stability over 5 days at 80:20. Mice perform well above chance when given the 90:10 and 80:20 contingencies (note: x-axis is not linear). Even at the more difficult 80:40 and 70:30 contingencies, more than 90% of mice performed above chance levels. Mice maintained a high number of trials per session under all contingencies (B). As the difference between high vs. low decreased (80:40 and 70:30), mice were less efficient at completing reversals (C). Win-stay probability (D) was significantly greater at 90:10 compared with 80:40, whereas lose-shift probability (E) was not significantly affected. Data are expressed as mean ± standard error. *$p < .05$, **$p < .001$. 

Table S1. Antibody details.
Amphetamine Treatment Increases Activity in the DMS and Lateral Orbitofrontal Cortex Inputs to the DMS During Reversal Learning

We were interested in the role of the DMS and its corticostriatal inputs during reversal learning and after amphetamine. To identify how activity in these corticostriatal circuits modulates PRL, we next quantified cFos expression in combination with CTb retrograde labeling of DMS inputs (Figure 4A). We determined the percentage of cFos+ cells and their intensity distribution in a range of corticostriatal regions (Figure 4B–D). As others have observed for these corticostriatal circuits (39–41), innervation of the DMS was greatest from the anterior cingulate cortex (ACC), ventral orbitofrontal cortex, and lateral orbitofrontal cortex (LO), with more than 75% of CTb-labeled cells originating in these areas (Figure 4E). Of the regions assessed, only the DMS had a significantly greater number of cFos+ cells after amphetamine treatment (Figure 4H) (\(r_{19} = 2.5, p < .05\)) (see Table S2). Other studies looking at cFos+ cell number after amphetamine have also observed specific increases in the DMS (42), suggesting that the DMS is more susceptible to activity-induced changes after amphetamine treatment.

In contrast to cFos+ cell percentage, amphetamine treatment increased the intensity of cFos expression in multiple areas (Tables S2 and S3). For example, all cortical areas showed a significant shift to the right of the intensity histograms (significant bin \(\times\) drug interaction), although the large number of cells quantified suggests that this reflects a small real-world increase in intensity. For the corticostriatal inputs (CTb-labeled cells), the ACC showed a significant bin \(\times\) drug interaction (Figure 4J) (\(F_{17,289} = 3.6, p < .01\)), and the LO had a significantly higher mean cFos intensity (Figure 4K) ( inset; \(f_{1.3} = 3.4, p < .01\)). In the striatum, a significant interaction of bin \(\times\) drug (right-shifted curve) was observed in the DMS (Figure 4L) (\(F_{17,323} = 3.9, p < .05\)) but not in the dorsolateral striatum (DLS) (Figure 4M) or the nucleus accumbens subregions (Table S3). Significant Kolmogorov-Smirnov tests of the cumulative distribution frequencies supported increased cFos intensities in the ACC and LO inputs to the DMS (Figure 4N, Q). Taken together, these results demonstrate that amphetamine alters activity in the DMS and corticostriatal projections from the ACC and LO to the DMS. Furthermore, these results show that the level of cFos expression used to classify cFos+ cells may have dramatic effects on study outcomes.

Reversal Learning Performance Is Associated With LO and DMS Corticostriatal Networks

Having identified specific differences in cFos expression within the ACC and in LO inputs to the DMS, we were interested in determining how these relate to behavioral performance. We considered saline- and amphetamine-treated mice as a single group because we did not know their individual dopamine levels and because amphetamine-induced cortical dopamine release is modest (~2-fold of baseline levels) (43,44), especially considering that studies in humans have indicated that manipulating striatal dopamine levels can differentially alter...
Figure 4. Corticostriatal cFos activation during reversal learning and amphetamine treatment. Fluorescently labeled CTb was injected into the DMS to retrogradely label cortical inputs to the DMS (A). cFos expression in CTb-labeled and CTb-unlabeled cells was quantified in cortical and striatal areas (B–D). The ACC and VO/LO areas had the largest proportion of CTb-labeled cells (E), highlighting their dense connectivity to the DMS. Percentage of cFos+ cells in corticostriatal inputs (ACC and LO) to the DMS and the DLS (F–I). Amphetamine treatment only increased the percentage of cFos+ cells in the DMS. There was significant right-shift in cFos + cell intensity levels (fold-increase from average of all cells) after amphetamine in ACC cells projecting to the DMS and in the DMS itself (J–M). Amphetamine increased the average cFos intensity in LO cells projecting to the DMS (inset) and in the DMS (inset), suggesting greater cFos activation. Cumulative distribution plots (N–Q) of cells in the ACC and LO projecting to the DMS show significantly greater increases in cFos intensity after amphetamine treatment. Similarly, amphetamine increased cFos intensity in the DMS, but not in the DLS. Data are expressed as mean ± standard error.
reward and punishment learning depending on baseline dopamine turnover (45). Moreover, we were interested in how corticostriatal activity relates to behavioral performance more broadly. Therefore, we used a principal component analysis to identify which cFos measures (ACC and LO inputs, as well as striatal outcomes) were associated with key performance outcomes (i.e., sharing strong loadings). This resulted in a 4-factor solution that accounted for 78% of the overall variance (Figure 5A). Factor 1 indicated strong relationships between LO inputs to the DMS, the DMS itself, and behavioral outcomes. cFos intensity in these areas was positively associated with reversal performance and win-stay probability, but negatively associated with lose-shift probability. Factor 2 grouped ACC and LO function, suggesting a positive relationship between cortical activity in the ACC and LO independent of behavioral outcomes. Factor 3 indicated that activity in the nucleus accumbens is positively associated with win-stay probability, whereas factor 4 indicated a negative relationship exists between DLS activity and lose-shift probability. The outcomes in factor 1 were supported by significant correlations between the mean cFos intensity of DMS cells and inputs from the LO (Figure 5B), reversal performance (Figure 5C), win-stay probability (Figure 5D), and lose-shift probability (Figure 5E). Taken together, these findings demonstrate that LO inputs to the DMS and activity within the DMS itself are associated with multiple aspects of PRL. Systemic amphetamine not only compromises learning after losses but also drives greater corticostriatal recruitment and activity, which may help negate broad impairments in reversal performance.

**Activation of Midbrain to DMS Cells With DREADDs**

There is a strong relationship between the orbitofrontal cortex and DMS function in reinforcement learning, particularly for dopaminergic systems (20,21). We were interested in whether increased DMS activity, independent of altered LO inputs, could lead to the observed reversal learning phenotypes. Therefore, we used pathway-specific DREADDs to activate midbrain to DMS projections (Figure 6A), hypothesizing that if increased DMS activity is a critical factor (downstream of LO activity or otherwise), then we should observe a behavioral phenotype akin to that after amphetamine treatment. DREADDs can be activated chronically via CNO (clozapine N-oxide) (AK Scientific) in drinking water (46), allowing us to

---

*(n = 8–11). *p < .05, **p < .01. ACC, anterior cingulate cortex; AcbC, nucleus accumbens core; AcbSh, nucleus accumbens shell; CTb, cholera toxin subunit B; DLS, dorsolateral striatum; DMS, dorsomedial striatum; eCDF, empirical cumulative distribution function; IL, infralimbic cortex; KS, Kolmogorov-Smirnov test; LO, lateral orbitofrontal cortex; MO, medial orbitofrontal cortex; PL, prelimbic cortex; VO, ventral orbitofrontal cortex.*
manipulate these circuits acutely and chronically during reversal learning. We confirmed that DREADDs receptors remained effective following all testing. The number of cFos+ cells was significantly greater in the mCherry+ cells of hM3Dq mice compared with control mice ($p < .05$). The number of cFos+ cells was also significantly greater in hM3Dq mCherry+ cells compared with control mCherry+ cells ($p < .01$). These results confirm results of prior studies demonstrating the longevity of DREADDs activity after chronic activation (46).

**Acute Activation Decreases Lose-Shift Probability**

As for amphetamine treatment, acute DREADDs activation did not significantly alter the number of completed trials or reversal performance. Activating cells projecting from the midbrain to the DMS significantly decreased lose-shift probability at the 80:40 contingency. Data are expressed as mean ± standard error. *$p < .05$, **$p < .01$. CNO, clozapine N-oxide; DMS, dorsomedial striatum; DREADDs, designer receptors exclusively activated by designer drugs; TH, tyrosine hydroxylase.
(Figure 6D), reversal performance (Figure 6E), or win-stay probability (Figure 6F) at 80:20 or 80:40 contingencies. Lose-shift probability was not significantly altered by DREADDs activation at 80:20 (Figure 6G), but at 80:40, there was a significant interaction of dose × group ($F_{1,45} = 5.8, p < .05$), with CNO administration in hM3Dq mice significantly decreasing lose-shift probability compared with control mice ($p < .05$) and vehicle treatment ($p < .01$). This result mimicked our observations after amphetamine, but at a more difficult contingency (i.e., 80:40 vs. 80:20). It is also consistent with our recent work on reversal learning in individuals with early psychosis (47) showing specific decreases in lose-shift probability at more uncertain contingencies (i.e., 80:40).

**Chronic Activation Independently Impairs Reversal Performance and Lose-Shift Probability**

The DMS is important for action selection and maintaining optimal strategy use in reversal learning (14,15,19). Therefore, well-trained animals may rely on learned strategies to maintain performance after acute DMS manipulations. Chronic activation may induce performance alterations by gradually impairing the navigation of reversal learning. We used CNO in drinking water to chronically activate midbrain to DMS pathways during the navigation of reversal learning. We used CNO in drinking water to chronically activate midbrain to DMS pathways during daily PRL testing. This protocol did not affect basic operant outcomes such as the number of trials completed (Figure 7B) or response latency (data not shown). In contrast, for reversal performance (Figure 7C), there was a significant effect of stage ($F_{3,45} = 16.2, p < .001$) and a significant interaction of stage × group ($F_{3,45} = 3.8, p < .05$). hM3Dq mice completed fewer reversals per 100 trials compared with control mice at the 80:40 ($p < .05$) and 70:30 ($p < .01$) contingencies. There were no significant effects of DREADDs activation on win-stay probability (Figure 7D), but we did observe a significant effect of stage ($F_{3,45} = 4.3, p < .01$) and a significant interaction of stage × group ($F_{3,45} = 3.1, p < .05$) for lose-shift probability (Figure 7E). Control mice maintained baseline levels of lose-shift probability at all contingencies, whereas hM3Dq mice had a significantly lower lose-shift probability at 80:20 ($p < .05$) and 80:40 ($p < .05$) than baseline. Together, these results indicate that lose-shift probability is sensitive to DMS activation, and DMS activation produces a phenotype akin to acute amphetamine treatment. However, DMS activation can also impair reversal performance, which tracks with increasing probabilistic uncertainty.

**DISCUSSION**

In this study, we investigated the relationship between corticostratal activity and reversal learning after dopaminergic manipulations in the DMS of male mice. Our results demonstrate that the DMS regulates multiple processes in reversal learning performance. Experimental timeline featuring 5 days of baseline testing at 80:20 with new water bottles containing DMSO for acclimatization (A). Mice were then given water bottles containing CNO and assessed for an additional 14 days at 80:20, followed by 4 days at 80:40 and 5 days at 70:30. All mice averages are represented as a percentage of their group’s baseline (e.g., controls normalized to baseline [n = 11] and hM3Dq normalized to hM3Dq baseline [n = 6]). Chronic CNO treatment to stimulate midbrain to DMS pathways did not alter the average number of trials completed at any contingency (B). DREADDs activation during both the 80:40 and 70:30 contingencies significantly decreased reversal performance (C). No differences were observed for win-stay probability (D), but DREADDs activation significantly decreased lose-shift probability during both the 80:20 and 80:40 contingencies (E). Data are expressed as mean ± standard error. *p < .05, **p < .01, ***p < .005 compared with hM3Dq baseline. CNO, clozapine N-oxide; DMS, dorsomedial striatum; DREADDs, designer receptors exclusively activated by designer drugs; PRL, probabilistic reversal learning.
The Dorsomedial Striatum Drives Reversal Learning in Mice

learning. Amphetamine treatment altered punishment learning and increased activity in LO inputs to the DMS and in the DMS itself. Next, we confirmed that amphetamine-induced phenotypes can be replicated via specific activation of midbrain to DMS pathways. Moreover, chronic activation of midbrain to DMS pathways resulted in a decline in reversal performance when probabilistic uncertainty is increased. These outcomes suggest that DMS dopamine dysfunction may contribute to reversal learning deficits in disorders such as psychosis and highlight a complex interplay between probabilistic uncertainty and DMS function.

Serial PRL in Mice

Previous studies have suggested that mice perform PRL tasks at levels at or slightly above chance (24), highlighting the need for improved task protocols. In our optimized PRL paradigm, more than 90% of mice performed above chance at complex contingencies (i.e., 80:40 and 70:30). Mice maintained levels of win-stay and lose-shift probability comparable to those reported in rats; however, win-shift probability was lower than in humans (17,48–50). We observed decreased win-shift probability at 80:40, concordant with data from human studies using more complex contingency settings (51). These data demonstrate that complex PRL tasks can be translated for use in rodents.

Amphetamine Administration Impairs Punishment Learning

Striatal dopamine function is considered an important moderator of reversal learning (5,7,14,16,45,52). Increasing brain dopamine with amphetamine significantly reduced lose-shift probability and increased perseverative errors after a reversal. Acute amphetamine treatment has been shown to alter lose-shift performance in other protocols with misleading feedback (53), and dopamine depletion in the DMS increases lose-shift probability in rats (54). Moreover, increased perseverations and alterations in corticostriatal activity have been observed in stimulant abusers (55). These outcomes complement the computational modeling parameters that suggest that amphetamine impairs punishment learning and the response to loss (defined as a lack of reward in this protocol). Thus, amphetamine-treated mice place less emphasis on current unrewarded trials (losses) after a reversal, leading to a longer period selecting the now-less-optimal outcome. These findings align with results of previous studies highlighting the potential role of dopamine in learning, whereby dopamine depletion improves punishment-based reversal learning (56,57).

Dopamine function often acts in an inverted-U-shaped response, with too little or too much impairing cognitive function (58), meaning that the same dose of amphetamine can improve performance in some subjects while impairing it in others (16,59,60). Therefore, outcomes often differ between studies. For example, cocaine and D-amphetamine have been shown to decrease total reversals and win-shift probability; however, lose-shift probability and perseverative errors remained consistent with those in saline-treated rats (61). Similarly, PRL performance and perseverative behavior in humans are not affected by administration of methylphenidate, a dopamine transporter blocker (62). These disparities may be due to differences in baseline dopamine function or competing regional effects after systemic administration.

LO and DMS Networks Underlie Reversal Performance in Mice

Our cFos data demonstrate that amphetamine increases the activation of DMS cells and LO inputs to the DMS. Selective effects of amphetamine on DMS activity have been observed previously (42), and the dorsal striatum may be more susceptible to stimulant-induced alterations in dopamine (63). The LO is also a core area associated with serial reversal learning, lose-shift probability, and identifying changes in outcome contingencies (22,64). The LO and DMS have overlapping roles, with studies in marmosets demonstrating that lesions of the caudate produce reversal learning phenotypes similar to those of the orbitofrontal cortex (20). This suggests that LO inputs to the DMS may be responsible for downstream changes in activity and reversal learning outcomes. For example, our PCA suggests that increased activity in both the LO inputs to the DMS and the DMS itself negatively associated with lose-shift probability. This aligns with the decreased lose-shift probability observed after amphetamine. However, these same measures were also positively associated with improved reversal performance and win-shift probability, indicating that LO to DMS signaling may drive multiple aspects of serial reversal learning. The balance of these competing roles, dose of amphetamine, and behavioral context may underlie the disparity seen between studies (65).

The DMS Regulates Multiple Aspects of Reversal Learning

To clarify the role of the DMS in reversal learning, independent of alterations in activity from LO inputs, we used chemogenetic activation of the DMS. Acute activation did not impair reversal performance, but there was a decrease in lose-shift probability at the 80:40 contingency. Specific decreases in lose-shift probability have been found at 80:40, but not at 80:20, in those with early psychosis (47), which features increased dopamine function in the associative striatum (DMS equivalent in humans) (7,15). In addition, imaging studies in those at risk of developing psychosis (those with high levels of subthreshold psychotic symptoms) have observed altered caudate activation in response to unexpected feedback during PRL (66). The specific effects of amphetamine and DMS activation on lose-shift, but not win-stay, probability was unexpected. However, studies using cortical lesions in rats have suggested that value-related information (both reward and loss) is distributed across cortical areas, allowing for redundancy and parallel processing (67). The DMS is potentially more critical in navigating the response to loss than reward.

Then we used chronic activation to see whether sustained impairments in DMS function could impair performance. Sustained DMS activation decreased lose-shift probability at 80:20 contingencies akin to amphetamine. Furthermore, reversal performance was heavily affected at more difficult contingencies (i.e., 80:40 and 70:30). This demonstrates that DMS dysfunction can replicate amphetamine-induced reversal learning phenotypes and that DMS-induced alterations in
lose-shift probability and reversal performance are distinct. For example, lose-shift probability was decreased at 80:20, but reversal performance was maintained. In contrast, lose-shift probability was maintained at 70:30, but reversal performance was significantly impaired. Although altered activation of the ventral striatum is often observed during reversal learning in individuals with schizophrenia (7), there is evidence of altered caudate function. For example, decreased caudate activation in individuals with schizophrenia has been associated with deficits in probabilistic learning and after positive feedback (68,69). Together, these data extend former work on the DMS that highlights its complex role in the maintenance and reliable execution of a selected strategy (14,19). Moreover, the DMS is important for evidence accumulation during learning [perhaps uniquely so in the brain (70,71)], which may become more important in more uncertain, complex environments. These multifaceted computations and roles support the possibility of multiple behavioral phenotypes due to DMS dysfunction, dependent on probabilistic uncertainty and the level of training or chronicity. Imaging studies focused on how corticostriatal activation changes in response to probabilistic uncertainty may provide more clarity on the role of the DMS (or the caudate) during PRL. Based on our data and the other data discussed here, we hypothesize that individuals with psychosis may show larger activation differences compared with healthy control subjects as uncertainty is increased.

Conclusions

Here, we demonstrated that mice can perform many reversals in a single session of PRL and that LO inputs to the DMS (and DMS function) are critical for navigating reversal learning. Furthermore, altering dopaminergic function globally and in the DMS can induce multiple reversal learning phenotypes, highlighting a complex interplay between contingency and DMS dopamine systems. These data suggest that the deficits observed in people with psychosis could be driven by subcortical dopaminergic dysfunction in the DMS, an effect that could occur independent of, or in addition to, cortical dysfunction in these individuals.

ACKNOWLEDGMENTS AND DISCLOSURES

Imaging and analyses were performed at the Queensland Brain Institute’s Advanced Microscopy Facility. This work was supported by an Advanced Queensland Research Fellowship (Queensland Government; No. AQRF04115-16RD1 [to JPK]), a University of Queensland Early Career Researcher Grant (Grant No. UQECR1718021 [to JPK]), the Brain & Behavior Research Foundation (Maltz Prize [to JPK]), and a National Health and Medical Research Council Project Grant (Grant No. GNT1139960 [to JPK]).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia (MKY, K-AC, JD, SZ, SA, THJB, JPK); Queensland Centre for Mental Health Research, Brisbane, Queensland, Australia (SA, THJB, JPK); and QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia (JPK). MKY and K-AC contributed equally to this work.

Address correspondence to James P. Kesby, Ph.D., at jkesby@uq.edu.au.

Received May 31, 2022; revised Aug 16, 2022; accepted Aug 17, 2022.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2022.08.005.

REFERENCES

The Dorsomedial Striatum Drives Reversal Learning in Mice


