Mesolimbic Dopamine Dynamically Tracks, and Is Causally Linked to, Discrete Aspects of Value-Based Decision Making

Supplemental Information

Supplemental Methods

Animals

Singly-housed rats were approximately 90 to 120 days old weighing 300 to 350 grams at the start of experiments. Animals were maintained at no less than 85% of pre-experimental bodyweights by food restriction (~10-15 g of Purina laboratory chow each day in addition to approximately 1 g of sucrose consumed during behavioral sessions) except during the post-operative recovery period when food was given *ad libitum*. All experiments were done during the light cycle. Animal procedures were conducted in accordance with the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals, and were approved by the University of North Carolina at Chapel Hill Institutional Animal Care and Use Committee (IACUC).

Surgery

For all surgical procedures, rats were anesthetized with a ketamine hydrochloride (100 mg/kg) and xylazine hydrochloride (10 mg/kg) mixture (intramuscular) and placed in a stereotaxic frame.

Fast-Scan Cyclic Voltammetry (FSCV): Rats were surgically prepared for voltammetric recordings as previously described (1). A guide cannula (Bioanalytical Systems, West Lafayette, IN) was positioned dorsally in the NAc core (1.3 mm anterior, 1.3 mm lateral from bregma). A guide cannula for the Ag/AgCl reference electrode was placed in the contralateral hemisphere.

The bipolar stimulating electrode (Plastics 1 Inc., Roanoake, VA) was placed dorsally to the VTA (5.2 mm posterior, 1.0 mm lateral from bregma and 7 mm ventral from brain surface). Stainless steel skull screws and dental cement were used to secure all items. The bipolar stimulating electrode was lowered in 0.2 mm increments until physical responses to electrical stimulation diminished indicative of proper electrode placement. The stimulating electrode was then fixed with dental cement.

Optogenetics: Rats were infused with a Cre-dependent adeno-associated viral construct encoding ChR2 with EYFP (AAV5-DIO-ChR2-EYFP) at a titer of 4 x 10^{12} virus molecules/ml, into the VTA (4 injections: AP -5.4 mm, ML ±0.7 mm, DV -8.4 mm; and AP -6.2 mm, ML ±0.7 mm, DV -7.4 mm from skull at bregma, 1 µl per site infused at a rate of ~0.5 µl/min with a 2 µl Hamilton syringe). The syringe was held in place for an additional 15 min before removal. To stimulate fibers arising from these cells, optical fibers (200 µm diameter core) coupled to ferrules (2.25 mm diameter, 250 µm bore) were chronically implanted over the NAc bilaterally at AP +1.8 mm, ML ±2.5 mm and DV -6.9 mm from skull at bregma, and angled in the ML plane away from the midline 10˚. Optical fibers were held in place with dental cement. For a subset of subjects destined for anaesthetized recordings (n = 2), rats were adminisered virus infusions as above, but no optical fibers were implanted above the NAc.

Behavioral Procedures

All behavioral experiments were conducted at least 1 week post-surgery. For all operant tasks involving sucrose pellet reinforcement, rats underwent similar pretraining before beginning each behavioral task. First, rats were trained to consume 10 psuedorandomly delivered (avg: 2 min) sucrose pellets (45 mg, Purina) from the foodcup in a single session. Next, rats were trained

to press two distinct levers in which each response was reinforced on a continuous schedule of reinforcement. Reinforced responses resulted in the delivery of a sucrose pellet to a centrally located food cup. Animals were trained to criterion of 50 presses on each response lever over 5 sessions before moving to the behavioral tasks outlined below.

Delay Discounting Task (for voltammetry studies): Each delay discounting session consisted of three blocks of 30 discrete trials (Fig. 1A). The first 20 trials of each block were either forced choice delay (Fig. 1A, *left*) or forced choice immediate (Fig. 1A, *middle*) trials (10 trials of each). For forced choice delay trials, a single cue light was illuminated over one lever for 5 s, followed by extension of both levers. Responses (FR1) on the lever below the illuminated cue light were rewarded with three sucrose pellets delivered after a delay (either no delay [block 1], short delay [10 s; block 2] or long delay [20 s; block 3]). For forced immediate trials, the cue light above the second lever was illuminated for 5 s then both levers were extended into the chamber. Lever presses (FR1) under the illuminated cue light within 10 s were rewarded with a single sucrose pellet across the 3 blocks. For both trial types, presses on the non-illuminated lever were not rewarded and counted as an error. Similarly, if the animal did not respond on either lever within 10 s, both levers retracted and the trial was counted as an omission. The next 10 trials within each block were free choice trials (Fig. 1A, *right*) in which both cue lights simultaneously illuminated for 5 s, and both levers were extended. Once either lever was pressed, both levers retracted and the animal was rewarded based on the contingency of reinforcement for the chosen lever within that block. Failure to choose a response within 10 s resulted in the levers retracting and the trial counted as an omission. Importantly, each trial was of fixed duration (60 s) so that reward choice was not influenced by how quickly the rat could complete the task, i.e., choosing the smaller immediate reward did not lead to the next trial more

quickly. Following 25 training sessions, all rats were be prepared for electrochemical recording in the NAc core. After recovery, animals underwent additional training sessions until behavior reached the pre-surgery baseline (at least 5 sessions).

Delay Discounting Task (for optogenetic studies): A modified version of the delay discounting task was used for optogenetic studies in another set of rats. Following shaping of lever press behavior, rats were trained in a discrete trial task to learn to associate each cue light with a specific response. During initial shaping, rats received either forced choice left or forced choice right trials in a random order. For forced left trials, the left cue light was illuminated for 5 s prior to both levers extending. Presses on the left lever resulted in the delivery of a single sucrose pellet. On forced choice right trials, presses on the right lever following cue light illumination resulted in a single sucrose pellet. Presses on the incorrect lever were counted as errors. Shaping sessions consisted of 35 trials of each trial type, and animals were trained on this task for at least 15 sessions until reaching behavioral criterion (85% accuracy).

Following initial shaping, rats began training on the decision making task. In this task one lever was designated the small-immediate lever and the other the large-delay lever. These lever assignments remained constant for each animal across training days but were counterbalanced across rats. Within a given session there were three trial types: forced choice immediate, forced choice delay, and free choice. During forced choice immediate trials (20 trials), one cue light was illuminated for 5 s followed by both levers extending for 10 s. Presses on the correct lever resulted in a single sucrose pellet delivered with 0 s delay. Presses on the non-signaled lever were counted as errors. On forced choice delay trials (20 trials), the other cue light was illuminated for 5 s followed by both levers extending. If the correct lever was chosen, the large reward was delivered after a specified delay. Importantly, the magnitude and delay of the larger

reinforcer remained constant throughout the entire training session, but varied across sessions. The magnitude of reinforcement was set in a given session to either 1, 2, or 3 pellets while the delay to reinforcer was set to either 0, 10, 20, 40, or 60 s following press. As on forced choice immediate trials, presses on the non-signaled lever were counted as errors and went unrewarded. Finally, during free choice trials (30 trials), both cue lights illuminated for 5 s followed by both levers extending. Responses were rewarded based on the contingency of the lever chosen. Free choice trials provided a behavioral readout of response preference for each rat on each day. During the first 20 trials of each session, only forced choice immediate and forced choice delay trials were presented (10 of each) to ensure that the animals learned the reward contingency for each session prior to any free choice trials. Animals were trained such that they experienced each delay and magnitude pairing.

Following training, animals underwent the optical stimulation tests. During stimulation sessions, rats (*TH::Cre*$^{(+/-)}$, n = 11, and littermate controls, n = 12) were connected to patch cables with optical fiber (200 μm core, 0.22 NA, Doric Lenses), encased in a durable plastic covering. These cables terminated with a ferrule connector (Precision Fiber Products) that were secured to the rat's optical fiber implant with a fitted ceramic sleeve (Precision Fiber Products), and were attached at the other end to an optical commutator (Doric Lenses). This commutator allowed for bilateral stimulation of NAc terminals and provided unrestrained movement for the animal. The commutator was connected via a second optical patch cable to a 150 mW DPSS 473 nm laser (OEM Laser Systems). Optical stimulation was controlled by a computer running Med PC IV (Med Associates) software which also recorded behavioral events. Each stimulation test occurred during separate training sessions. To test how dopamine signaling is causally linked with reward value (either intrinsic, delay to reward or subjective, reward magnitude), terminal

dopamine release was optically stimulated during cues that predicted lower value options. For the delay-based manipulation during test, the delay to reinforcement was set at 10 seconds and the value of the large reinforcer was 1 pellet, such that animals were choosing between one pellet immediately and one pellet after a 10 second delay. Rats followed the same training as before, however during forced choice cues that predicted the delay option, optical stimulation of dopamine terminals was administered during the 5 s cue period. Animals then made behavioral choices without stimulation.

Likewise, magnitude-based decision testing was conducted such that rats received two pellets delivered immediately for pressing the large (delay) lever. Here, rats were evaluating a choice between two pellets immediately or one pellet immediately. Optical stimulation occurred during forced choice immediate cues that predicted the smaller (one pellet) option. Again, stimulation only occurred during forced choice cue presentations and not during choice cue presentations, behavioral responses, or reward delivery. Therefore, any alterations in behavior resulted from changes in the predicted reward values related to magnitude as a result of the stimulation.

Intracranial Self Stimulation: Rats (*TH::Cre*$^{(+/-)}$, n = 11, and littermate controls, n = 12) were trained to nosepoke for optical stimulation of the NAc, though due to fiber loss during this portion, only a subset completed all portions (self-administration, extinction, reinstatement and signal shift) of the task (*TH::Cre*$^{(+/-)}$, n = 10, and littermate controls, n = 9). During each session, a houselight illuminated the chamber and a single white LED lamp recessed in the rear of the nosepoke receptacle indicated that entries would be rewarded. During the first five acquisition sessions, a single nosepoke resulted in a 5 s bilateral optical stimulation (20 Hz, 5 ms pulsewidth, 20 mW). During this 5 s stimulation period, the light in the nosepoke hole extinguished

signifying that any further nosepokes during this period would not be reinforced. Both rewarded nosepokes and total nosepokes were recorded, over 30 minute behavioral sessions. Following the 5 acquisition sessions, the laser was switched off, and animals were able to respond during one 75 min extinction session. During this session, all behavioral events were identical to acquisition, however nosepoke responses were not reinforced, and recorded every 15 minutes. Behavior was considered extinguished following either two 15 minute periods of no responses or following 75 minutes. Directly following extinction, the laser was turned back on and one 5 s "priming" stimulation was delivered. Animals were then allowed to nosepoke for optical stimulation during a 45 min reinstatement session. Following reinstatement, the intensity of the laser was decreased to a 1 s stimulation (20 Hz, 5 ms pulsewidth, 20 mW) while all other parameters of the task remained identical. Rats were then allowed to nosepoke for this lower level 1s stimulation over a 30 min behavioral session.

Fast-Scan Cyclic Voltammetry

For awake/behaving voltammetry recordings, a carbon-fiber electrode, housed in the micromanipulator, was lowered into the NAc core on the recording day and was used to measure dopamine changes during task performance. The potential of the carbon-fiber electrode was held at -0.4V versus the Ag/AgCl reference electrode. The carbon fiber and Ag/AgCl electrodes were connected to a head-mounted voltammetric amplifier attached to a commutator (Crist Instrument Company, Hagerstown, MD). Prior to recording, the carbon fiber electrode was allowed to equilibrate for 20-30 minutes in the brain to minimize current drift. Voltammetric recordings were made every 100 ms by applying a triangular waveform that drives the potential to +1.3V and back to -0.4V at a scan rate of 400V/s. Application of the triangular waveform results in the

oxidation and reduction of chemical species that are electroactive within this potential range (including dopamine), producing a measurable change in current at the carbon fiber. Dopamine was verified by plotting this change in current against the applied potential to produce a cyclic voltammogram (2, 3). The stable contribution of current produced by oxidation and reduction of surface molecules on the carbon fiber was removed by using a differential measurement (i.e., background subtraction) between a time when such signals were present but dopamine was not. For data collected during the behavioral session, this background period was obtained during the baseline window (10 s prior to cue onset). Following equilibration, dopamine release was electrically evoked by stimulating the VTA using a range of stimulation parameters (2-24 biphasic pulses, 20-60 Hz, 120 µA, 2 ms per phase) to make sure that the carbon fiber electrode was placed close to dopamine release sites and to create a training set for principal component analysis (1-4). Animals then underwent task performance and electrochemical recordings were made continuously with 100 ms temporal resolution. A second computer and software system (Med Associates Inc.) controlled behavioral events and sent digital outputs for each event to the voltammetry recording computer to be time stamped along with the electrochemical data. Following termination of the behavioral session, VTA stimulation was repeated to verify the stability of the electrode and ensure that the location of the electrode still supported dopamine release.

To verify that optical stimulation was sufficient to promote phasic dopamine release in the NAc, *TH::Cre*$^{(+/-)}$ rats expressing ChR2 in the NAc were anesthetized with urethane (i.p.) (Sigma-Aldrich, St. Louis, MO) and prepared for electrochemical recordings. First, we confirmed that cell body stimulation was sufficient to promote dopamine release *in vivo* by coupling an optical fiber to the electrical stimulation probe and lowered them into the VTA. A

carbon fiber electrode was lowered into the NAc. We electrically stimulated (60 Hz 24 biphasic pulses, 120 µA, 2 ms per phase) the VTA to verify that the recording electrode was in a location that supported dopamine release. We then optically stimulated (20 mW, 20 Hz, 40 pulses) cell bodies while recording dopamine release in the NAc. Next, a carbon fiber electrode coupled to an optical fiber was lowered into the NAc to verify that dopamine terminal stimulation was sufficient to promote release *in vivo*, as this had previously only been shown *in vitro* (5). Dopamine release was recorded while terminals were optically stimulated at varying intensities (20 mW, 20 Hz, 5-100 pulses) to verify that there was an intensity dependent relationship between optical stimulation and dopamine release.

Signal Identification and Separation

Following recording sessions, dopamine release evoked by electrical stimulation of the VTA was used to identify dopamine transients. Stimulation of the VTA leads to two well-characterized electrochemical events: an immediate but transient increase in dopamine and a delayed but longer lasting basic pH shift. To separate these signals a training set was constructed from representative, background-subtracted cyclic voltammograms for dopamine and pH, as previously described (2, 3). The background period was obtained at the minima for the dopamine signal 5 s before event onset. This training set was used to perform principal component regression on data collected during the recording session (1-4). Principal components were selected such that at least 99% of the variance in the training set was accounted for by the model. All data presented here fit the resulting model at the 95% confidence level.

Histology

Voltammetry Electrode Placements: Rats were deeply anesthetized with a ketamine/xylazine mixture (100 mg/kg and 10 mg/kg, respectively). In order to mark the placement of the electrode tip, a 150 to 250 µA current was be passed through a stainless steel electrode for 5 s using established procedures (4, 6). Animals were decapitated and brains removed and post-fixed in 10% formalin. After post-fixing and freezing, brains were sliced at the level of the NAc at 30 µm coronal sections and mounted on microscope slides. The specific position of electrodes was assessed by visual examination of successive coronal sections in comparison to visual landmarks and anatomical organization of the NAc represented in a stereotaxic atlas (7), Fig. S2.

Immunohistochemistry: Animals were perfused transcardially with physiological saline and a 4% paraformaldehyde solution. Brains were then removed, post-fixed, and frozen. Brains were sectioned coronally at 30 µm with half of the sections mounted on slides to verify optical fiber placement with a light microscope and the other half placed in 0.1M PB for immunohistochemistry to verify ChR2 expression. Free floating sections were washed in Triton-X (0.5% solution in phosphate-buffered saline, (PBST)) and 0.1M PBS. Sections then were incubated with 10% normal donkey serum (Jackson Immunoresearch Laboratories) in 0.1% PBST for 60 minutes. Sections were then incubated for 72 hours at 4°C in TH polyclonal antibody raised in sheep (1:500, Abcam) and then washed and incubated for 2 hours at room temperature in donkey anti-sheep secondary antibodies conjugated to Alexa Fluor 647 (1:800, Jackson Immunoresearch Laboratories). Next, sections were washed and incubated for 60 minutes in 2% NeuroTrace (435/455 nm, Invitrogen LifeTech), then washed and mounted onto microscope slides in phosphate-buffered water and coverslipped with Vectashield mounting

medium (Vector Laboratories). Sections were visualized on a confocal microscope to evaluate virus expression as well as optical fiber placement (Figs. S6A, S7 and S8). Rats with misplaced or broken optical fibers or *TH::Cre*$^{(+/-)}$ rats that failed to show evidence of expression of ChR2 in the VTA and NAc were excluded from analysis. Imaged neurons from *TH::Cre*$^{(+/-)}$ rats displayed a high degree of co-expression of TH and CHr2-EYFP as in Fig. S6A.

Data Analysis

In the FSCV experiments, we calcuated a variety of behavioral factors to determine the rats' performance. On the forced trials, we calculated the percent accuracy and error types. Specifically, errors could result from either pressing the incorrect lever (error of commission) or failing to respond within the appropriate trial duration (error of omission). For the free choice trials, we calculated the proportion of time each rat chose each of the levers during a given block. Whichever option the rat chose most within a block, the rat was considered to "prefer" that option for both the free choice and forced choice trials for that block. Notably, a rat's preference could shift across blocks as delay to reinforcement increased, and for purposes of the study, the preference measure was computed for each rat based on performance on the day of recording.

Analysis of behavior during training and test sessions included examination of overall response rates and allocation, number of errors committed, number of omissions committed, and lever preference. To determine if rats reliably acquired the decision making tasks, we evaluated the number of errors and correct responses during training and test sessions. Further, we evaluated the ability of subjects to adjust behaviors as the value of reinforcers changed (delay and magnitude manipulations) by using a repeated measures ANOVA to compare responses during choice trials for the large delayed reward across the different reward contingencies. For

the optogenetics experiments it was also important to confirm that there were no differences in value-based decision making behaviors across the two groups of animals. As such, we compared responding between the two groups to confirm that there were no differences in baseline responding during task performance. To evaluate if stimulation of dopamine terminals during lower value options was sufficient to shift behavioral responding we compared response allocation on choice trials during stimulation versus nonstimulation sessions and across both groups of rats. Further, the percentage of correct responses on forced trials was compared between stimulated versus non-stimulated sessions to confirm that laser stimulation did not alter the ability to perform the task. For optical self-stimulation tests, we evaluated behavioral responding across days using a 2-way repeated measures ANOVA to determine if there was a significant increase in nosepoke behavior in the *TH::Cre*$^{(+/-)}$ animals compared to controls as they learned the contingency of the task. We then made specific planned comparisons of responding on the final session of training to the extinction session, reinstatement session, and 1 s stimulation using *t*-tests.

Changes in extracellular dopamine concentration were assessed by aligning dopamine concentration traces to each behavioral event. Increases or decreases in NAc dopamine concentration from baseline for the cue presentations and were evaluated separately for each cue type (across both block and type of cue) using a one-way repeated measures ANOVA with Dunnett's correction for multiple comparisons. This analysis compared the baseline mean dopamine concentration (10 s prior to cue onset) to each data point (100 msec bin) obtained within 5 s following the task event. The effect of cue type (large versus small) and reward delay (no delay block, short delay block, long delay block) were evaluated with a 2-way repeated

measures ANOVA that compared peak changes in dopamine levels following each event, with Tukey's correction for multiple post-hoc comparisons.

For choice trials, rats' data were grouped by which option they selected most out of the total number of responses for that block. For example, in the No Delay block, if a rat chose the Large/Delay option 8 times and the Small/Immediate option 2 times, that rat for that block would be said to have a preference for the Large/Delay option. If, in the next block (Short Delay), the rat chose the Small/Immediate option 6 times and the Large/Delay option 4 times, his preference has now shifted to the Small/Immediate option for that block. Unlike the forced-choice trials where rats were required to perform a fixed number of trials at each lever (10 each per block), the free choice trials required rats to allocate choice responses based on preference. As such, there were unequal numbers of responses for each option within a block for the free choice trials; indeed, on the blocks where the rats had the strongest preference (10 out of 10 responses on one of the options), there would be no within-subjects comparison with which to compare the alterative option within that block. Further, due to differences in variance, it would similarly be inappropriate to compare DA release for the average of 9 preferred choices compared to the one non-preferred choice in a block. To address this, we used the DA release for each trial and averaged across all subjects within a given block based on whether it was the rat's preferred option. Data for this were analyzed using a 2-way between-subjects ANOVA using block (No Delay, Short Delay, Long Delay) and preference (Preferred Option, Non-Preferred Option) as factors. Because we could not assume normality or equal variance, post-hoc tests were done using nonparametric Kolmogorov-Smirnov tests between the Preferred and Non-Preferred options within each block.

All analyses were considered significant at $\alpha = 0.05$. Statistical and graphical analysis was performed using Graphpad Prism (Graphpad Software, Inc) and STATISTICA (Statsoft, Tulsa, OK).

Supplementary Results

Delay Discounting Behavior

During performance of the delay discounting task in FSCV recorded sessions, we analyzed various factors of the behavioral responses for the different options. While in general rats showed accurate overall performance on the Forced Choice trials (84%), this was modulated by the delay length in current block of the task. A two-way ANOVA that examined Block x Forced Choice type (Large/Delay vs Small/Immediate) found a significant interaction between Block and Type, $F_{(2,14)} = 15.62$, $p < 0.001$ (Fig. S1A). Posthoc comparisons found that accuracy for the Small/Immediate option was the same across all blocks (No Delay: 94%, Short Delay: 92%, Long Delay: 89%; all $p > 0.95$), while the accuracy for the Large/Delay option was less stable. Rats showed good accuracy for the No Delay (96%) and Short Delay option (86%) that was statistically similar ($p = 0.60$), but significantly worse accuracy on the Long Delay block (44%; Short Delay vs Long Delay, $p < 0.001$). As such, there were no differences in accuracy for the different trial types on the No Delay block ($p = 0.99$) and the Short Delay block ($p = 0.91$), but rats were less accurate for the Large/Delay trials than the Small/Immediate trials in the Long Delay block ($p < 0.001$).

To understand these differences, we then looked at the two kinds of errors that rats could make: omissions (no response) and commissions (press the wrong lever). A two-way ANOVA for Block and Trial Type found that rats made an increasing number of omissions across blocks,

(main effect Block, $F_{(2,14)} = 8.69$, $p < 0.01$), but there was no effect of Trial Type or Block X Trial Type (both $F < 1$; Fig. S1B). Posthoc analysis showed that rats in both groups made more omissions in the Long Delay block (21%) than the No Delay (1%) block (Large/Delay Trials, $p = 0.04$; Small/Immediate Trials, $p = 0.06$), but there were no differences between Trial Type within any block (all $p > 0.96$).

In contrast, commission errors were differentially distributed across the Trial Types (Fig. S1C), as indicated by a significant two-way ANOVA, $F_{(2,14)} = 9.97$, $p < 0.01$. Similar to the accuracy measure, we found that commission errors were stable across blocks for the Small/Immediate trials (No Delay: 8%, Short Delay: 3%, Long Delay: 1%; all $p > 0.90$), but increased significantly for the Large/Delay trials (No Delay: 0%, Short Delay: 8%, Long Delay: 36%). There were no differences in commission errors between Trial Type in the No Delay block ($p = 0.88$) or in the Short Delay block ($p = 0.97$), but there were significantly more commission errors on the Large/Delay option than the Small/Immediate option in the Long Delay block ($p < 0.01$).

Finally, we examined whether there was a relationship between Forced Choice errors of commission and preference within a block. We hypothesized that if rats were using preference to guide responding, then they should make more errors with their non-preferred lever during Forced Choice trials. In confirmation, regression analysis indicated a significant negative linear trend between the number of errors on the Large/Delay lever and the proportion of times it was chosen in the Free Choice trials, $r = -0.70$, $p < 0.0001$. In contrast, there was no relationship between the number of errors on the Small/Immediate lever and the number of times chosen in the Free Choice trials, $r = -0.28$, $p = 0.19$. Further, this effect was predominantly due to variations in preference within the Long Delay block, where preference for the Small/Immediate

lever was correlated with increased errors on the Large/Delay Forced Choice trials, r = -0.89, p = 0.003, while this preference had no effect on the error rate on the Small/Immediate lever, r = -0.36, p = 0.38. Collectively, these findings suggest that rats were not making errors because they did not understand the task contingencies. If this was the case, rats would have been expected to make errors randomly throughout the blocks and regardless of the forced choice trial type or preference. Instead, rats only increased errors for the Large/Delay option in the Long Delay block, while accuracy in all other blocks was high and stable. These behavioral findings are consistent with the idea that the reward associated with the Large/Delay option is devalued by the delay.

Histology

Placement of electrodes for recordings (n = 8) were confirmed to be in the nucleus accumbens core (Fig. S2).

Phasic Dopamine Signaling During Free Choice Trials

During free choice trials, phasic DA release following cue onset was related to the value of the best available option. Consistent with previous work (1, 8), peak DA at the time of the free choice cue decreased across trial blocks as the subjective value of expected rewards was decreased by the temporal delay (Fig. S3A, $F_{(2,14)}$ = 9.75, p = 0.002). Specifically, free choice cue DA was significantly greater during the No Delay block than in the Short Delay ($p < 0.05$) and the Long Delay ($p < 0.05$) block, though there was no difference between DA at the Long and Short Delays (Fig. S3B). Indeed, further analysis indicated that the rats' preferences modulated DA signaling during both free choice and forced choice trials in similar ways.

Subject-by-subject dopamine during the preferred forced choice cues was significantly correlated with DA during the free choice cues only when the rat subsequently chose its preferred option, $r = 0.64$, $p < 0.001$, but not when the rat chose its non-preferred option, $p > 0.1$ (Fig. S4, red symbols). In contrast, there was no significant relationship between DA release for the non-preferred forced choice cues compared to the free choice cues when the rat chose either its preferred or non-preferred option (both $p > 0.1$; Fig S4, gray symbols). Thus, phasic DA release following predictive cues was similar in both forced choice and free choice trials in a manner that strongly predicted the rats' subjective preference within a block.

Previous work has suggested that DA in the free choice trials may differ not at the time of cue, but instead at the time of action (8). To assess if this were the case for the present experiment, we evaluated DA release at the time of press and in the period immediately post-press, which corresponds to the action and evaluation of the action/reward, respectively. Data were analyzed using all trials grouped by preference and aligned to the time of press. Unlike for cues, we found no evidence of a difference in DA signal at the time of the chosen action, regardless of preference based on a two-way between-subjects ANOVA, $F_{(2, 210)} = 0.07$, $p = 0.92$, Fig. S5A. Likewise, there were no main effects or interactions of Block or Choice when the DA concentrations were grouped by expected reward size rather than preference (all main effects and interaction, $F < 1$).

In contrast, DA concentrations were modestly modulated by the post-reward evaluation of the response (Fig. S5B). Looking at peak DA in the 2 s following press relative to a pre-press baseline, there was a trend towards a significant interaction of Block by Reward Size (Large vs Small), $F_{(2,210)} = 2.35$, $p = 0.09$. Posthoc tests indicated that the peak DA in the period following presses on the Small/Immediate choice slightly increased (Kolmogorov-Smirnov D = 0.25, $p =$

0.10), while DA following presses on the Large/Delay choice slightly decreased (Kolmogorov-Smirnov D = 0.38, $p = 0.07$). While there were no differences in post-press DA for the different reward options in either the No Delay block (Kolmogorov-Smirnov D = 0.16, $p = 0.95$) or the Short Delay block (Kolmogorov-Smirnov D = 0.25, $p = 0.23$), there was a significant decrease in DA for the Large/Delay choice following press relative to the Small/Immediate press (Kolmogorov-Smirnov D = 0.49, $p < 0.01$). Thus, in the Free Choice trials, we see evidence that post-response DA was linked not to the value of the reward (as DA was similar whether the animal received 3 pellets or 1), but instead possibly to evaluation of less-desired choices that was explicitly linked to delay costs rather than reward size.

Optical Stimulation of CHR2-Expressing Dopamine Cells *TH::Cre*$^{(+/-)}$ Rats Is Sufficient to Promote Phasic Dopamine Release in the NAc

Optical stimulation of dopamine cell bodies in the VTA of anaesthetized rats elicited significant dopamine release in the NAc in *TH::Cre*$^{(+/-)}$ rats expressing ChR2, that was time-locked to laser onset, and peaked at laser offset (Fig. S6B). Next, optical stimulation limited to dopaminergic terminals in the NAc likewise resulted in a significant time-locked increase in dopamine release (Fig. S6C). There was a significant linear relationship between stimulation intensity and dopamine release ($r^2 = 0.83$, $p < 0.0001$, Fig. S6D) with the greatest stimulation intensity eliciting the highest dopamine release.

Stimulation of Dopamine Terminals in the NAc Is Sufficient to Promote Goal-Directed Responding

Similar to previous findings using cell body stimulation (5, 9), dopamine release in the NAc via optical stimulation of arising VTA fibers was sufficient to promote goal-directed actions. *TH::Cre*$^{(+/-)}$ (n = 10) and littermate controls (n = 9) received injections of Cre-dependent ChR2 virus in the VTA, along with chronic implantation of an optical fiber with tips in the NAc (Figs. S8 and S9). At least eight weeks following surgery, rats were able to nosepoke for a 5 s train of light pulses in the NAc (FR1; 20 Hz, 100 p 5 msec pulse duration; Fig. S7). We found a main effect of group ($F_{(1,17)}$ = 8.26, p = 0.01), session ($F_{(11,17)}$ = 7.94, p < 0.0001) and a significant interaction ($F_{(11,17)}$ = 5.258, p < 0.0001). *TH::Cre*$^{(+/-)}$ rats showed a significant increase ($t_{(9)}$ = 3.56, p = 0.0062), and controls a significant decrease ($t_{(8)}$ = 3.61, p = 0.0069), in nosepokes from session 1 to 5, with *TH::Cre*$^{(+/-)}$ animals responding significantly more than controls during session 5 ($t_{(17)}$ = 3.004, p = 0.008).

Following acquisition, animals underwent extinction (nosepokes no longer elicited laser stimulation). *TH::Cre*$^{(+/-)}$ rats rapidly extinguished responding, with significantly fewer nosepokes at the end of extinction compared to the final training session, $t_{(9)}$ = 3.676, p = 0.005. However, when laser stimulation was resumed, *TH::Cre*$^{(+/-)}$ rats rapidly reinstated nosepoking to pre-extinction levels, $t_{(9)}$ = 1.591, p = 0.146. During the final session, nosepoking was reinforced with a lower-intensity 1 s (20 Hz 20 p 20 mW) optical stimulation. The rate of 1 s self-stimulation was significantly blunted compared to the 5 s stimulation period during the final day of training $t_{(9)}$ = 3.308, p = 0.0091.

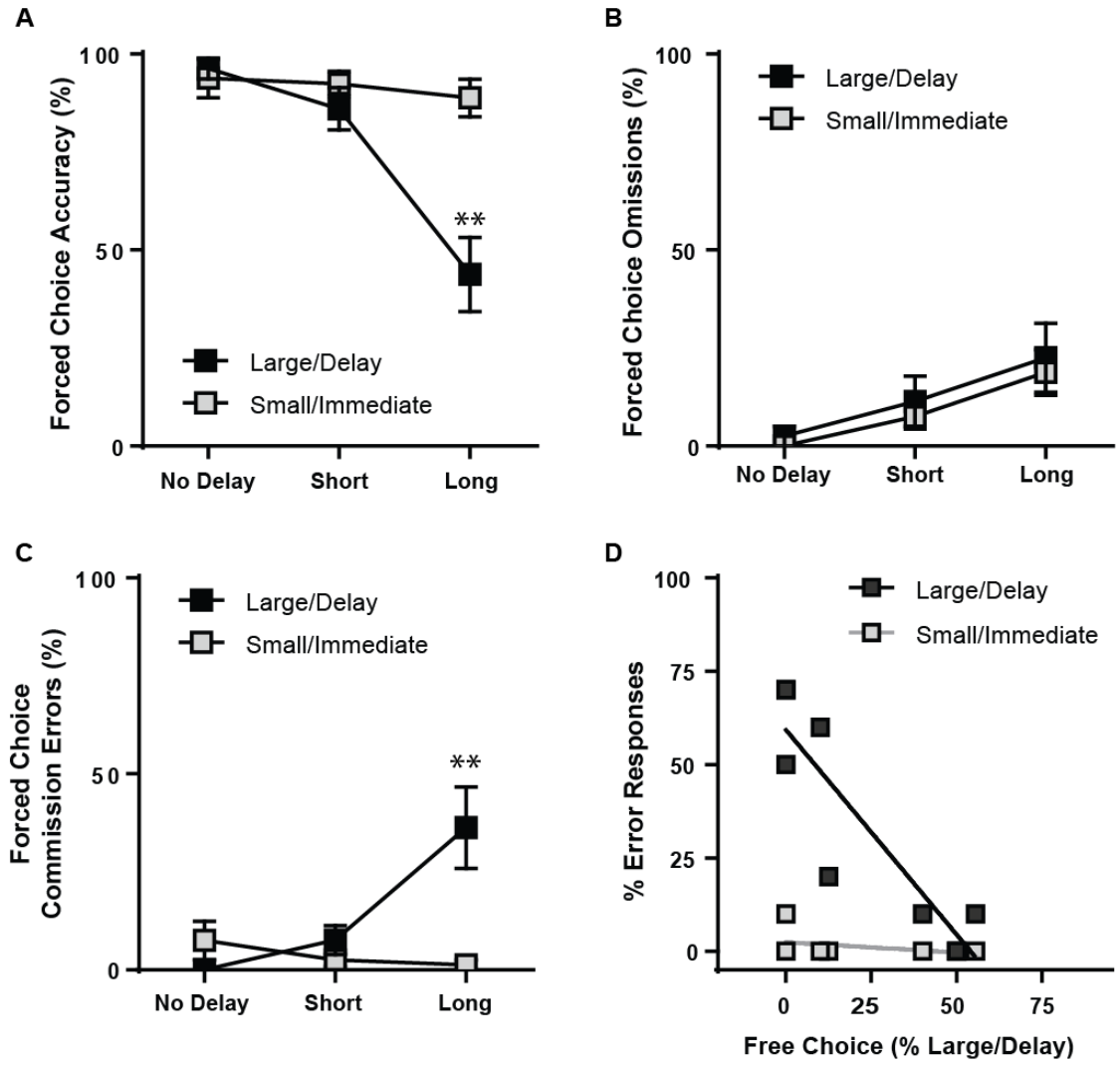

Figure S1. During FSCV recordings, rats showed accurate performance on the Forced Choice trials, which was modulated by the length of delay for the Large/Delay outcome on those trials. **(A)** Accuracy was stable and high for the Forced Small/Immediate trials across all blocks, but declined for the Large/Delay trials as the delay to reinforcement increased across blocks. **(B)** Omitted responses increased across blocks, but did not differ between trial type. **(C)** Errors of commission (i.e., presses on the wrong lever) were selectively increased for the Large/Delay trials during the Long Delay block. Specifically, when the Large/Delay lever was presented, there was a significant increase in presses on the (non-reinforced) Small-Immediate lever. **(D)** Errors of commission were not random during the Long Delay block. Preference for the Small/Immediate option during the Free Choice trials significantly predicted ($p < 0.005$) the error of commission rate on the Forced Large/Delay trials, while there was no correlation between choice preference and error rate on the Forced Small/Immediate trials. $^{**}p < 0.01$, Lg/Del vs Sm/Imm. A-C display mean ± SEM.

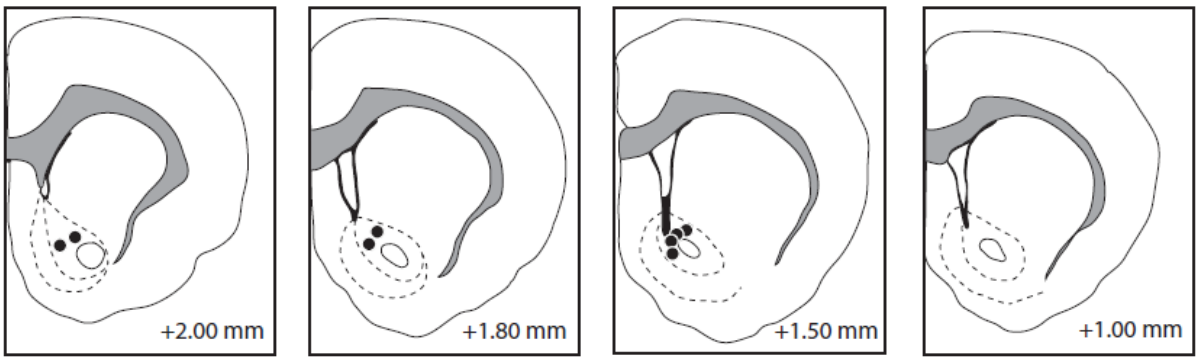

Figure S2. Histological verification of electrode tips for the delay discounting task for electrochemical measurement sites in the NAc core. Coronal sections show carbon fiber electrode tip locations (black dots) for 8 recording locations in 7 rats in the delay task. Numbers indicate section distance rostral to bregma.

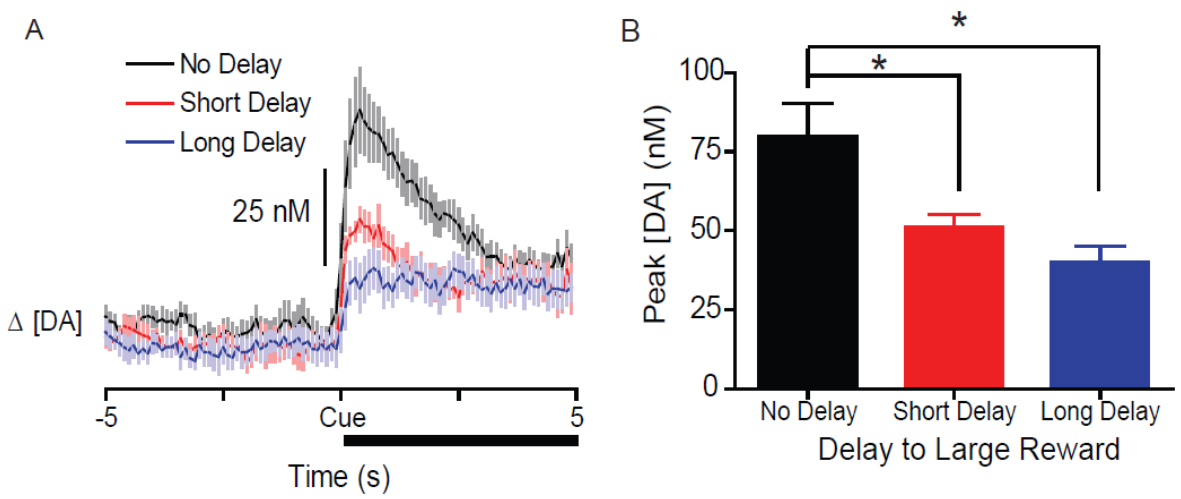

Figure S3. Dopamine release during free choice trials on the delay discounting task. **(A)** Dopamine concentration aligned to cue onset (black bar, time 0 s) on free choice trials during each of the three blocks of the task. Free choice trials were associated with significant increases in dopamine during cue presentation indicated by repeated measures ANOVA $p < 0.05$ for all comparisons of cue period versus baseline. **(B)** Peak dopamine concentration during free choice cue presentation. Choice cues during the no delay block evoked significantly greater dopamine than choice trials during the short or long delay blocks, though there was no difference in dopamine signaling during the short and long delay blocks. *$p < 0.05$. All data are mean ± SEM.

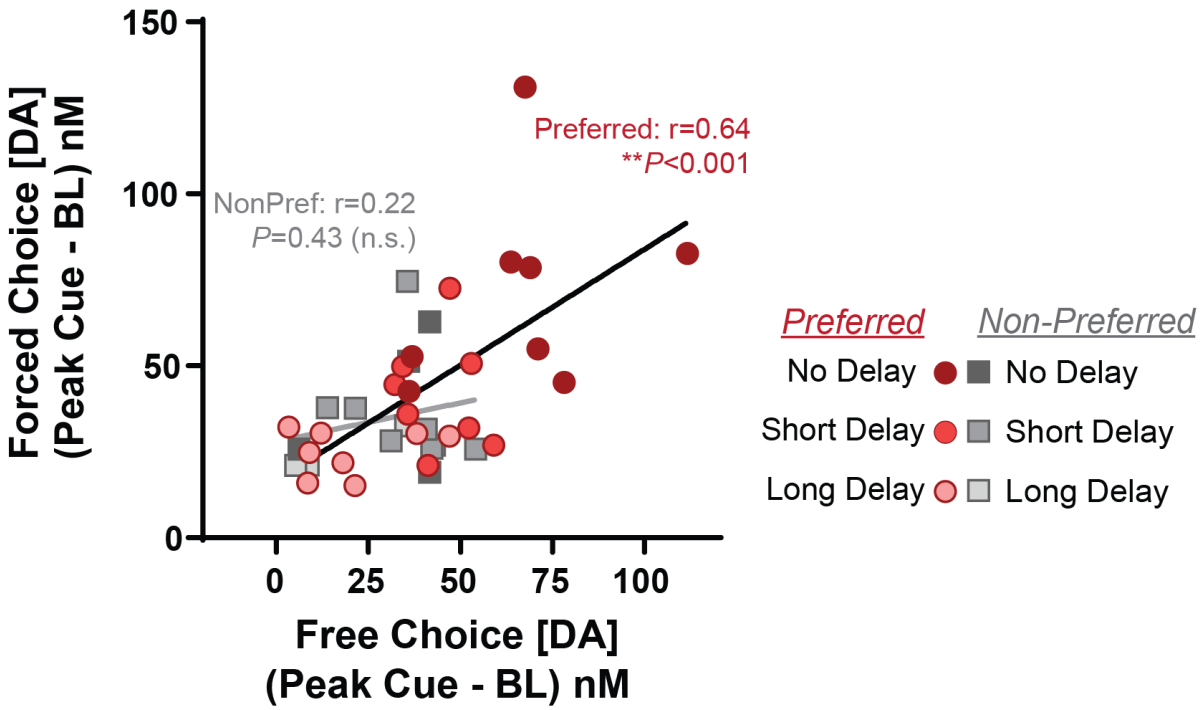

Figure S4. Dopamine release during cues in the preferred forced choice and free choice trials were correlated when rats made presses on the preferred option, but not when they made presses on the non-preferred option. Data show preferred forced choice trials versus free choice trials when the preferred option was later chosen (red symbols, black regression line), and non-preferred forced choice trials versus free choice trials when the non-preferred option was later chosen (gray symbols, gray regression line).

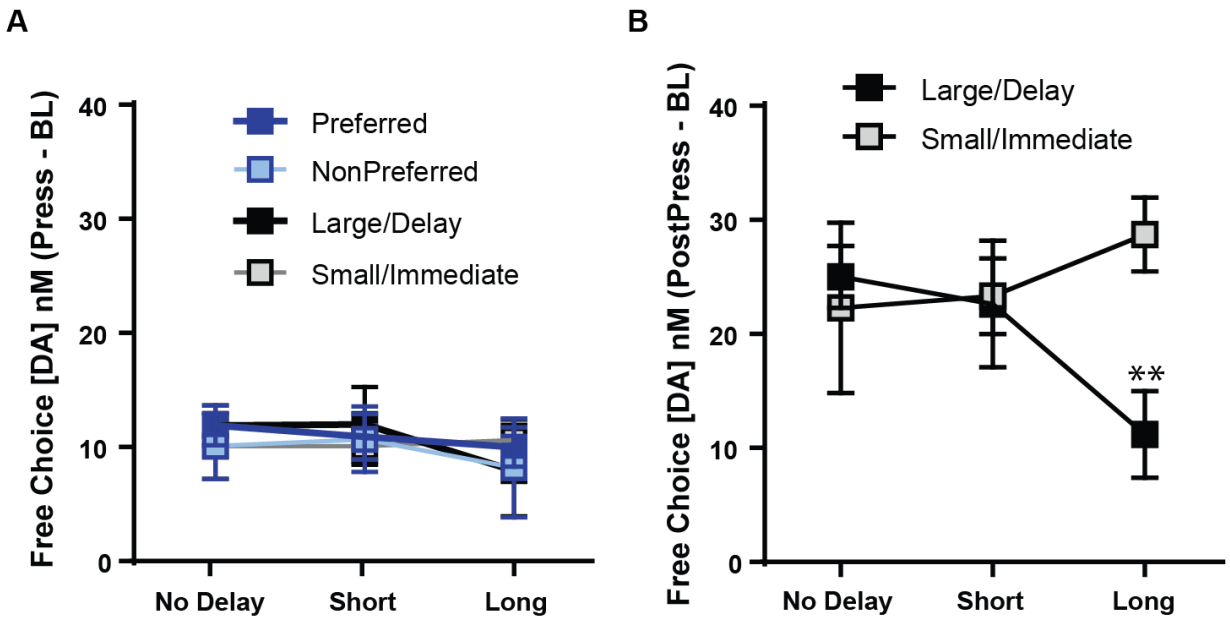

Figure S5. DA release during free choice trials at the time of lever press **(A)** and in the 2 s period following press **(B)**. DA release at the time of press was not modulated by either the rat's preference or size of anticipated reward **(A)**. In contrast, post-response DA following presses on the Large/Delay option was associated with decreased levels at the longest delays. $^{**}p < 0.01$. All data are mean ± SEM.

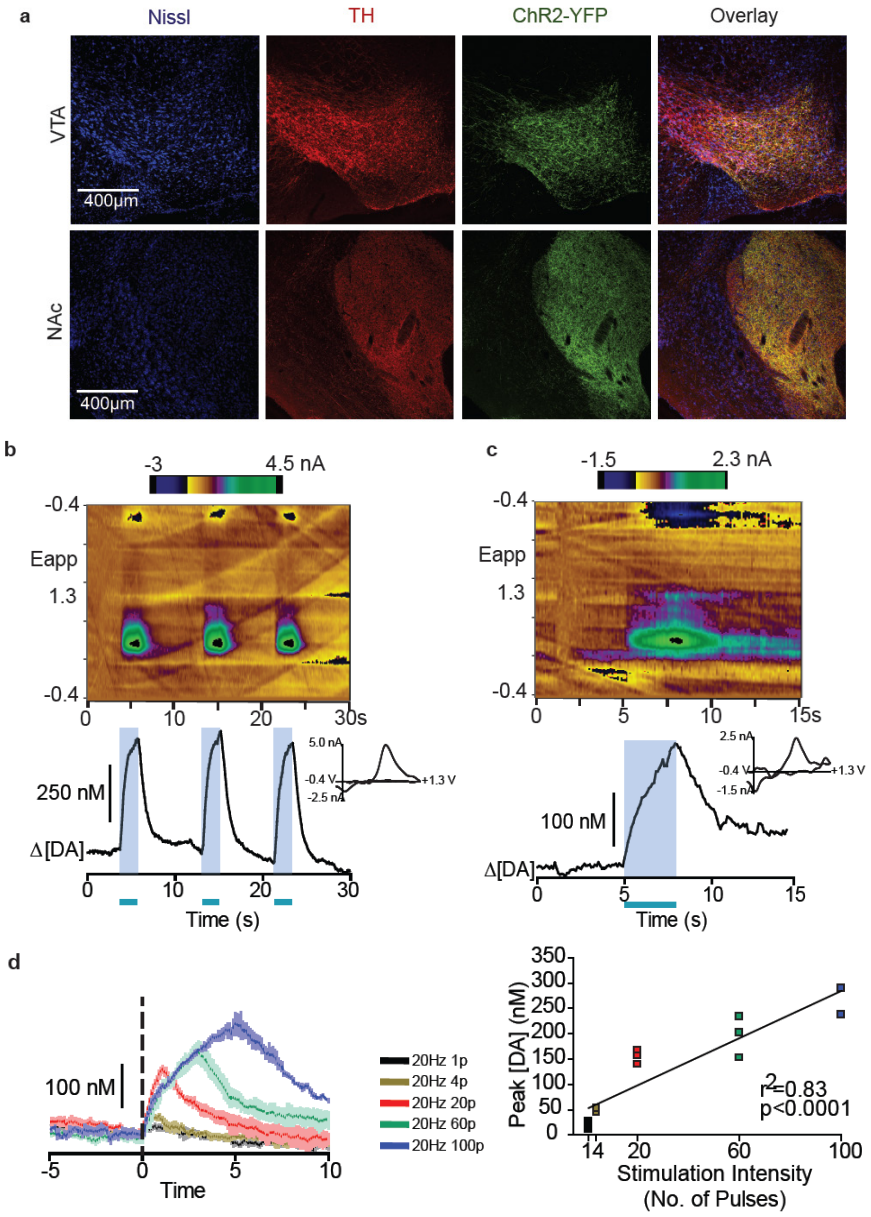

Figure S6. Optical stimulation of mesolimbic dopamine neurons promotes phasic dopamine release. **(A)** *Top*: TH staining and ChR2-EYFP expression in coronal slices display co-localization in cell bodies and projection neurons of the VTA. *Bottom*: TH staining and ChR2-EYFP expression in coronal slices display co-localization in dopamine varicosities in the terminal region of the NAc. Expression is not seen in the cell bodies in the terminal region of the NAc. **(B)** Optical stimulation of dopamine cell bodies in the VTA is sufficient to promote phasic dopamine release in the NAc. *Top*: Three-dimensional representation of electrochemical data during 20 Hz 40 p 20 mW 5 ms pulsewidth optical stimulation and corresponding dopamine concentration trace *(bottom)*. Cyclic voltammogram confirming that the signal measured is

dopamine shown to the right of the dopamine trace. Optical stimulation denoted by blue shaded bars. **(C)** Optical stimulation of dopamine terminals in the NAc was sufficient to promote phasic dopamine release; conventions follow from **(B)**. **(D)** *Left*: Amount of phasic dopamine released was dependent on the stimulation intensity. *Right*: Correlation of peak dopamine release and stimulation intensity.

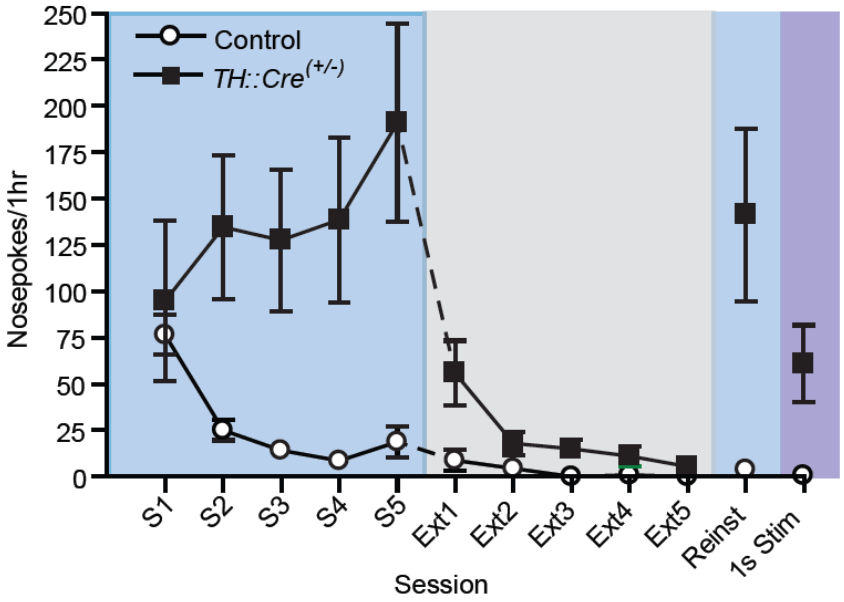

Figure S7. Nosepoke responses (nosepokes/hr) for optical stimulation for all animals in both groups. Blue area signifies when the laser was active while gray area signifies when the laser was off. Purple area signifies when responding was rewarded with a 1 s optical stimulation rather than 5 s.

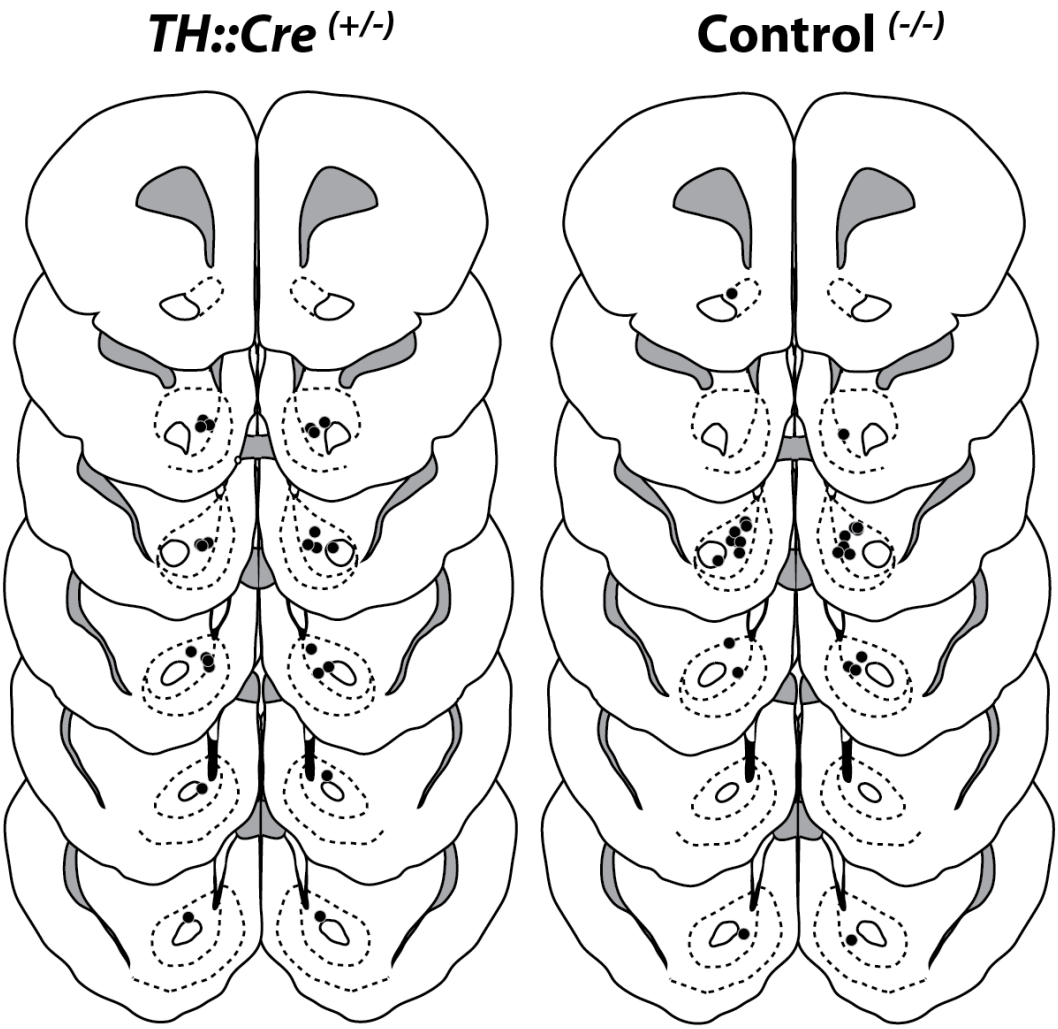

Figure S8. Placement of tips of optical fibers for subjects who received optical stimulation of DA fibers in the delay manipulation.

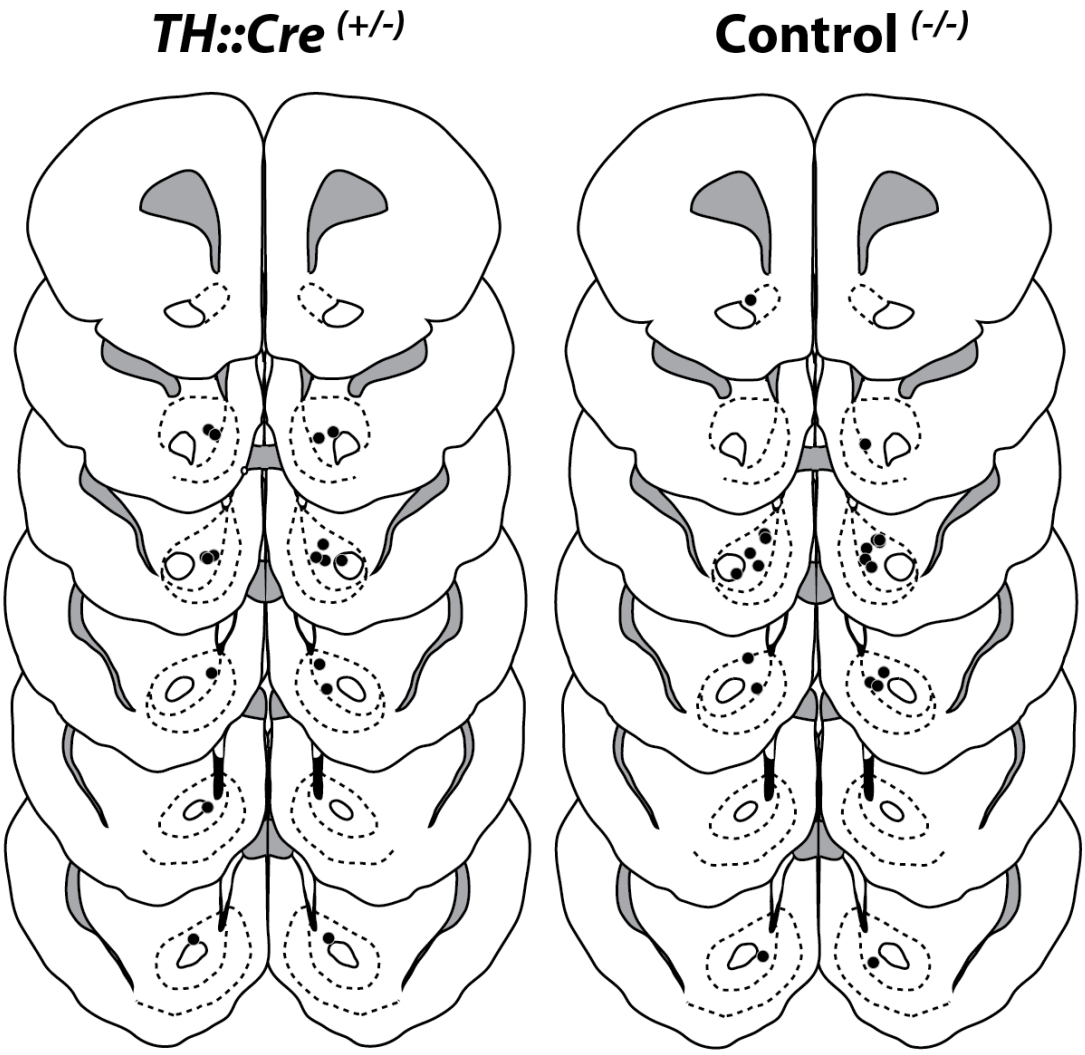

Figure S9. Placement of tips of optical fibers for subjects who received optical stimulation of DA fibers in the magnitude manipulation. Note that all subjects in the magnitude test received stimulation in the delay manipulation as well in the same location.

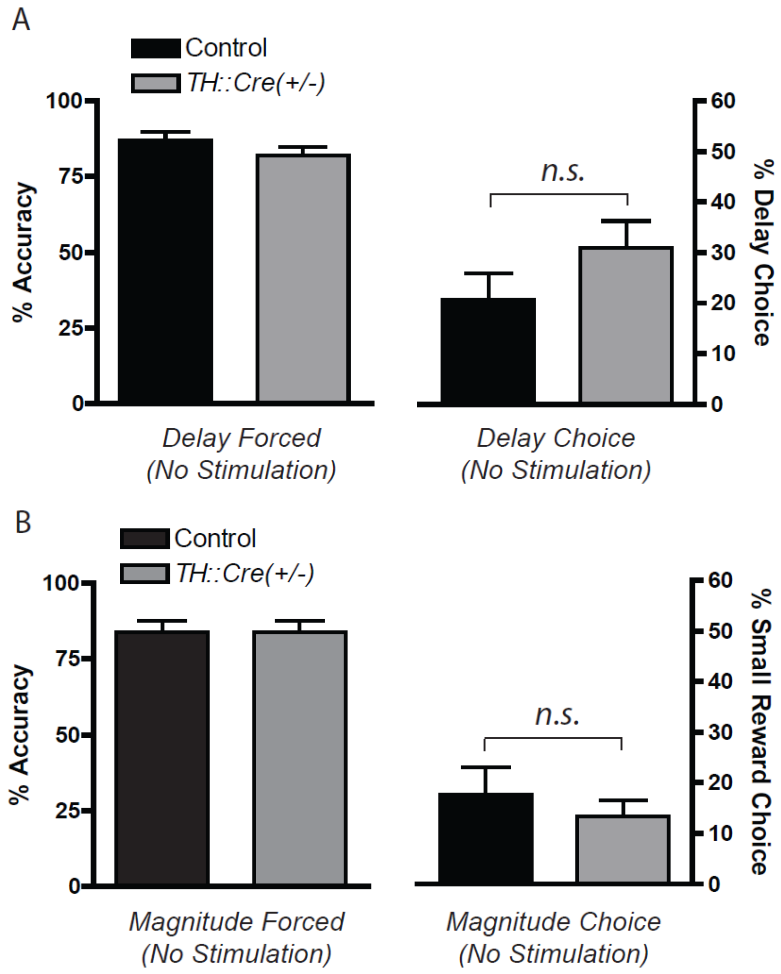

Figure S10. Rats were run in an identical behavioral decision making task as during the laser stimulation sessions, but the laser was turned off for the session. In both the Delay test **(A)**, and the Magnitude test **(B)**, rats performed similarly regardless of genetic background. In Forced Choice trials (**A**, *left* and **B**, *left*), both *TH::Cre*$^{(+/-)}$ and controls responded with similar levels of accuracy, and were not different than Forced Choice accuracy when the laser was active in the laser stimulation sessions. On Free Choice trials in the Delay test (**A**, *right*), rats in both groups similarly preferred the Immediate option over the Delay option. Likewise, in Free Choice trials in the Magnitude test (**B**, *right*), rats in both groups preferred the Large reward option over the Small reward option. All data mean ± SEM.

Supplemental References

1. Sugam JA, Day JJ, Wightman RM, Carelli RM (2012): Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. *Biol Psychiatry*. 71:199-205.

2. Heien ML, Khan AS, Ariansen JL, Cheer JF, Phillips PE, Wassum KM, *et al.* (2005): Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. *Proc Natl Acad Sci U S A*. 102:10023-10028.

3. Heien ML, Johnson MA, Wightman RM (2004): Resolving neurotransmitters detected by fast-scan cyclic voltammetry. *Anal Chem*. 76:5697-5704.

4. Day JJ, Roitman MF, Wightman RM, Carelli RM (2007): Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci*. 10:1020-1028.

5. Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, *et al.* (2011): Recombinase-driver rat lines: Tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 72:721-733.

6. Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM (2004): Dopamine operates as a subsecond modulator of food seeking. *J Neurosci*. 24:1265-1271.

7. Paxinos G, Watson C (1997): *The rat brain in stereotaxic coordinates*. New York: Academic Press.

8. Roesch MR, Calu DJ, Schoenbaum G (2007): Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*. 10:1615-1624.

9. Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013): A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*. 16:966-973.