



Published in final edited form as:

Neurobiol Learn Mem. 2023 April ; 200: 107734. doi:10.1016/j.nlm.2023.107734.

Optogenetic Disruption of the Prelimbic Cortex Alters Long-Term Decision Strategy but Not Valuation on a Spatial Delay Discounting Task

Amber E. McLaughlin^{1,2}, A. David Redish¹

¹Department of Neuroscience, University of Minnesota Twin Cities, Minneapolis, MN, USA

²Current affiliation: Icahn School of Medicine at Mount Sinai, New York, NY, USA

Abstract

Rats demonstrate a preference for smaller, immediate rewards over larger, delayed ones, a phenomenon known as delay-discounting (DD). Behavior arises from the interaction of multiple decision-making systems, and the medial prefrontal cortex (mPFC) has been identified as a central component in the mediation between these decision systems. To investigate the role of the prelimbic (PL) subregion of mPFC on decision strategy interaction, we compared two cohorts of rats (ChR2-opsin-expressing ‘Active’ and opsin-absent ‘Control’) on a spatial delay-discounting task while delivering in-vivo light stimulation into PL at the choice point of select trials. By analyzing the overall delay-adjustment along with deliberative and procedural behavioral strategy markers, our study revealed differences in the decision strategies used between the active and control animals despite both groups showing similar valuations. Control animals developed the expected shift from deliberative to procedural decision strategy on this task (indicated by reaching delay-stability, particularly during late-session laps); however, active-virus animals repeatedly over-adjusted around their preferred delay throughout the entire session, suggesting a significant deficit in procedural decision-making on this task. Active animals showed a significant decrease in proportion of vicarious trial and error events (VTE, a behavior correlated with deliberative processes) on delay adjustment laps relative to control animals. This points to a more nuanced role for VTE, not just in executing deliberation, but in shifting from deliberative to procedural processes. This opto-induced change in VTE was especially pronounced for late-session adjustment laps. We found no other session-by-session or lap-by-lap effects, leaving a particular role for PL in the long-term development of procedural strategies on this task.

1. Introduction

An animal’s success in foraging depends on its ability to process information from the environment and to use that information to best direct its actions toward reward. Current

AEM: Conceptualization, Methodology, Data collection, Data curation, Analysis, Writing. **ADR:** Conceptualization, Methodology, Writing, Supervision.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

theories suggest that behavior arises from the interaction of multiple decision-making systems, each with its own complex computational processes and neural components (O'Keefe & Nadel 1978; Daw, Niv, & Dayan 2005; Redish 2013; van der Meer et al. 2012). Two of these systems are the deliberative system (also known as model-based or goal-directed decision-making, Gilbert & Wilson 2007; Johnson & Redish 2007; Niv, Joel, & Dayan 2006; Redish 2016) and the procedural system (also known as model-free or habitual decision-making, Hull 1943; Daw, Niv, & Dayan 2005; Niv, Joel, & Dayan 2006; Graybiel 2008). Deliberative processes entail planning strategies that consider future outcomes (Redish 2016), while procedural processes use learned associations between situations and actions (Graybiel 1998; Dickinson 1994; Barnes et al. 2005). As such, deliberative/planning systems are sensitive to contingency and are better suited to drive behavior under conditions that are flexible and changing, whereas procedural/habitual systems are sensitive to action-contiguity and drive behavior in familiar, stable conditions (Balleine, Delgado, & Hikosaka 2007; McLaughlin, Diehl, & Redish 2021). However, the ways in which these distinct decision systems interact to produce dynamic behavior remains an ongoing area of research.

The combination of sophisticated task paradigms with neuromodulation techniques have spurred new directions for connecting complex behavior to neural mechanisms. An important strategy marker, the behavioral event of vicarious trial and error (VTE) has been found to correlate with neurophysiological deliberation during decision-making (Johnson & Redish 2007; Amemiya & Redish 2016; Papale et al. 2016; Kay et al. 2020) and to vanish with behavioral automation evidenced by habit learning (van der Meer et al. 2012; Gardner et al. 2013; Smith & Graybiel 2013). The medial prefrontal cortex (mPFC) has long been associated with both the deliberative (Fuster 1997; Killcross & Coutureau 2003; Rich & Shapiro 2007; Kesner & Churchwell 2011) and procedural (Jog et al. 1999; Coutureau & Killcross 2003; Barnes et al. 2005; Smith & Graybiel 2013b; Barker et al. 2017;) decision systems as well as in their interaction (Ragozzino et al. 1999; Miller & Cohen 2001; van Aerde, Heistek, & Mansvelder 2008; Heidbreder & Groenewegen 2003). Previous work suggests a broad role for mPFC in modulating complex, dynamic decision-making (Dalley, Cardinal, & Robbins 2004; Kesner & Churchwell 2011; Laubach et al. 2018; McLaughlin, Diehl, & Redish 2021).

The intricate anatomical connectivity mPFC has with areas such as hippocampus (Jones & Wilson 2005; Peyrache et al. 2009; Adhikari, Topiwala, & Gordon 2010; Bett et al. 2012; Hok et al. 2013; Ito et al. 2015; Guise & Shapiro 2017; Schmidt, Duin, & Redish 2019), ventral striatum (Floresco, Seamans, & Phillips 1997; Euston, Gruber, & McNaughton 2012) and orbitofrontal cortex (Chudasama & Robbins 2003; Sul et al. 2010) makes it a key region of interest for studying reward valuation, action selection, and task representation. Behavioral experiments have investigated its functional and anatomical complexity in rodents through economic tasks that use indicators, like VTE (Bett et al. 2012; Gardner et al. 2013; Schmidt et al. 2013; Redish 2016; Kidder et al. 2021) and habit formation (Coutureau & Killcross 2003; Smith & Graybiel 2013a) to distinguish deliberative from procedural strategies (Papale et al. 2012; Powell & Redish 2016; Sweis et al. 2018) while targeting specific subregions within the mPFC. Of its three known subregions (ACC, PL, and IL), prelimbic disruption has been associated with deficits in goal-directed behavior

(most notably in VTE reduction (Schmidt, Duin, & Redish 2019; Kidder et al. 2021), and recordings from the prelimbic region suggest a role in processing information relevant to environmental changes indicative of recognizing a need for a behavioral strategy shift (Peyrache et al. 2009; Durstewitz et al. 2010; Powell & Redish 2016; Barker et al. 2017). By selectively disrupting PL at the choice point of a spatial decision-making task that can measure valuation and delays to reward, we aimed to investigate how deliberative deficits contribute to strategy changes across different time scales.

Delay Discounting describes a reduction in the perceived value of a reward as the temporal distance (the delay) to that reward increases. Increases in delay discounting have been linked to impulsivity and found to be a risk factor for addiction and other disorders (Mischel, Ebbsen, & Zeiss 1972; Madden et al. 1997; Giordano et al. 2002; Odum, Madden, & Bickel 2002; Mitchell 2004; Madden & Bickel 2010; Lempert et al. 2019), making it an interesting metric to investigate maladaptive decision-making behaviors. Moreover, behavioral training designed to reduce discounting rates has been found to reduce addictive relapse (Stein, Daniel, Epstein, & Bickel 2015).

The Spatial Adjusting Delay Discounting task adapts the Mazur adjusting delay procedure (Mazur 1997) to a T-maze that builds on the naturalistic behavior of rats to alternate between foraging options (Papale et al. 2012). The Spatial Adjusting Delay Discounting task is a neuroeconomic task that requires subjects to repeatedly make left vs. right choices between a small, immediate reward on one side and a larger one that will only be delivered after a variable delay once the subject arrives at the reward location on the other (Fig. 1A). In this task, the delays change based on the rat's choices – choosing the smaller-sooner reward decreases the delay to the larger-later reward by 1s, while choosing the larger-later reward increases its delay by 1s (Fig. 1B). Thus, the subject can control the delay 'cost' of the larger reward by making choices on this task - titrating the delay up or down depending on its left/right choice proportion.

Well-trained rats exhibit three distinct phases over the course of a session in the Spatial Adjusting Delay Discounting Task (Papale et al. 2012; Bett et al. 2015; Kreher et al. 2019). Rats first show an investigation phase marked by alternation between the two sides, presumably to assess the parameters of the task for a given session (which side is delayed, how much reward is delivered on the delayed side, and what is the start delay). Rats then typically show a titration phase, in which they repeatedly run more laps to one side or the other, which drives the delay up (consecutive delay side laps) or down (consecutive non-delay side laps). Once the delay has reached the rat's individual willingness to wait for the larger reward, rats typically alternate between sides, holding the delay constant, which we identify as a maintenance phase (Fig. 1C–D). Importantly, the task does not enforce these phases on subjects; rather, these phases describe the patterns of behavior that well-trained rats typically exhibit.

Both deliberative and procedural decision-making systems can solve this task, but rats typically deliberate when titrating on the task, and then automate (proceduralize, automate into a habit) when maintaining, as evidenced by hippocampal involvement (Bett et al. 2012), changes in behavioral deliberation markers and the regularity of the path taken (Papale et

al. 2012; Bett et al. 2015; Papale et al. 2016; Kreher et al. 2019), and by hippocampal (Papale et al. 2016) and orbitofrontal and ventral striatal firing patterns on this task (Stott and Redish 2014). Analyzing the trajectory through the choice point on each lap provides a way to measure VTE and infer deliberation (Fig. 1E–F). The reward economy and strategy dynamics on this task are covert and internally driven by the subject (Powell & Redish 2016). This provides a useful way to investigate an animal's valuation algorithms (Fig. 1G) and self-guided shifts in strategy (Fig. 1H–I).

The three-phase structure from the Spatial Adjusting Delay Discounting task makes it a powerful tool for identifying changes in strategy between deliberation and procedural decision-making. Prelimbic (PL) firing patterns show strategy-related representations that align with the three phases (exploration, titration, maintenance), changing ensemble-alignments a few laps before a rat changes strategy (Powell & Redish 2016). On other tasks, previous research has shown that mPFC disruptions yield deficits in strategy changes that appear to follow subregional specificity —most notably, linking deficits in goal-directed (deliberative) behaviors with PL disruption (Ragozzino et al. 1999; Rich and Shapiro 2007; Tran-Tu-Yen et al. 2009; Dalton et al. 2016; Riaz et al. 2019; Schmidt, Duin, & Redish 2019; Kidder et al. 2021) while disruption of IL activity appears to inhibit habit formation (Coutureau and Killcross 2003; Killcross and Coutureau 2003; Smith and Graybiel 2013a). Given the known relationships between PL ensembles and within-task behavioral phases, we set out to examine the consequences of prelimbic mPFC disruption on the Spatial Adjusting Delay Discounting task. Given the known effects on choice behavior of PL disruption at a choice-point (Kidder et al. 2021), we targeted this disruption through optogenetic manipulation on a subset of laps specifically at the choice point.

2. Methods

2.1 Subjects

For this investigation, all subjects were first-generation fisher-brown Norway (FBNF-1) hybrid rats ($n = 16$; 8 females, 8 males, ages 6–13 mos. at time of experiment) bred in-house. One female control rat and one male active rat lost their implants and only completed half of the experimental sessions and they are not included in the analyses reported. One female control rat was found post-hoc to have poor transfection and is not included in the analyses reported. This left us with 7 active and 6 control rats for analysis. All procedures were approved by the University of Minnesota Internal Animal Care and Use Committee and were done in accordance with NIH guidelines.

Before training, each rat was acclimated to the experimenter through five days of 30-minute handling sessions. During this time, rats were individually housed with access to freely available food and water in their home cages and on a 12-hour light/dark cycle. Fully handled rats were then food-restricted (placed in home cages with access to water but no food) and offered 12g of [45mg full-nutrition Bioserve plain] pellets for 45 minutes once/day for 5 days. For each rat, this pellet training occurred within the same 2-hour window sometime during the light cycle, conditioning them to expect their daily food ration at a consistent time. Body weight was recorded for all subjects prior to feeding each day and maintained at above 80% (baseline weight calculated as the average across handling

sessions). Once accustomed to eating the reward pellets, rats were then given access to freely available food in their home cage so their body weight could return to baseline in preparation for surgery.

2.2 Surgery

All subjects underwent a single surgical procedure targeting the prelimbic mPFC subregion for bilateral viral infusion and optic fiber implantation. Subjects were randomly assigned to either the active (n=8; 4 female, 4 male) or control (n=8; 4 female, 4 male) group. These two cohorts are distinguished only by the type of virus used during the procedure. Active animals were transduced with AAV5-ChR2-CaMKIIa-mCherry, while control animals were transduced with AAV5-CaMKIIa-mCherry (Fig. 2B). Rats were anesthetized for the entirety of the stereotaxic surgery using from 1% to 2% isoflurane and medical-grade oxygen. The virus was sourced from UNC Vector Core via AddGene (Active: AAV5-CaMKIIa-hChR2(H134R)-mCherry with titer of 6.8×10^{12} GC/mL; Control AAV5- CaMKIIa-mCherry with titer of 3.3×10^{12} GC/mL).

To access mPFC, craniotomies were drilled [coordinates from bregma: A/P +2.80mm(female) +3.00mm(male), M/L +/-0.70mm], and bilateral cannulae were lowered into the craniotomies to [from skull surface: D/V -3.8mm(female) -4.0mm(male)]. One microliter of virus was delivered into each hemisphere at a rate of 200 nanoliters/minute. The cannulae were slowly raised following a 15-minute wait period after injection. Optic fibers, attached through an in-house procedure to an LED capable of delivering opsin-activating light were then bilaterally implanted through the craniotomies, and directed just dorsal to the injection site [from skull surface: D/V -3.6mm(female) -3.8mm(male)]. The implant was secured to the skull with Metabond and dental acrylic.

Rats spent 5 days recovering from surgery with access to freely available food (chow and pellets) in their home cages. To further prevent infection, rats were given an antibiotic sequence of Baytril (2.27%) for 5 recovery days.

2.3 Task Training

Following surgery and recovery, rats were trained to perform the Spatial Adjusting Delay Discounting task where they repeatedly faced a choice to decide between two reward offers: a fixed, non-delayed offer that delivered a smaller reward (one 45mg food pellet) after 1s, and a delayed offer that delivered a larger reward (3 pellets for females, 4 pellets for males) after some variable delay (D). Each lap taken to the non-delayed side decreased the subsequent delay (D) by one second, and each lap taken to the delayed side increased the delay by one second. In this way, rats were able to use their choices to adjust the larger reward delay time D and titrate it up or down to a preferred duration (Fig. 1C–D). 1 second was the smallest delay achievable on this task, and there was no cap for how long the delay could be titrated up to.

This task required rats to learn the “correct” direction to progress within a figure-eight maze (Fig. 1A). From the start point, rats must proceed up the center track and enter the choice point from that direction. From the choice point, animals could either turn left or right. Entry into the respective wait zone would trigger the countdown. So long as the rat remained in the

wait zone, the countdown continued, and the reward was delivered after the delay time was up. Movement in either direction out of the wait zone terminated the offer. Rats had to enter into the start zone to initiate a lap sequence. Any entry into a zone that went counter to the task flow would terminate the lap offer and another lap would not begin until the rat entered the start zone (from any direction). In practice, all rats ran consistently in a figure-eight pattern through the central track returning to the start zone through the left or right return rails, and they consistently waited out the delays in the wait zone before continuing.

To learn this task, rats progressed through blocks of training stages across days with each stage introducing an additional task complexity (Fig. 2A). The first set introduced rats only to one side (left or right, changing each day) of the task for the entirety of the daily session. During this time, rats were given 45 minutes to earn all their food for the day. Each lap in the correct direction cued an auditory tone followed by 2 food pellets 1s later. Left and right sides were blocked on alternate days for 4–6 days until the rat ran more than 50 laps to each side. The second training set exposed rats to both the left and right task sides, following the standard task protocol. The location of the delayed side was randomized for each 45-minute session and always began with a 1s start delay that the rats learned got progressively longer the more they chose it over the non-delay side. For these first two training stages, rats were prevented from running backwards by the experimenter manually blocking them with a long stick (a thin PVC pipe). Rats then learned to maneuver through the task while fixed with a tethered head-stage during the third training set, with the starting delay was set to 5s. At this stage, rats were not prevented from running backwards, but the experimenter would still block as needed to encourage learning. The final training set exposed tethered rats to randomized starting delays of either 1s, 5s, or 10s. Once rats ran 5 consecutive days at sufficient performance with no experimenter blocking, the experimental sequence began. These last 5–6 days of “training” were used as the “baseline” data for each rat. Rats were not manually blocked during this final training set or through the experimental sequence.

2.4 Experimental Sequence

Once rats completed training, they entered an 18 session (1 session/day) experimental sequence whereby the start delay was randomized, and the value range was broadened to include sessions that started with greater initial economic scarcity than those experienced during training (Sessions started at 1s, 5s, 10s, 16s, 22s, 30s, counterbalancing the delayed side). To test the effects of optogenetic stimulation on mPFC during this task, light stimulation was delivered on $\frac{1}{3}$ of the experimental sessions as a 7 Hz sine wave and $\frac{1}{3}$ of the experimental sessions as white noise. Because no differences were found between these two experimental conditions (Fig. S1A–E), they have been combined for analysis into a single ‘Opto’ condition ($\frac{2}{3}$ of experimental sessions) and compared to the remaining $\frac{1}{3}$ of the sessions in which no-opto stimulation (N.O.) was delivered.

Each experimental session had two lap types (‘stim’ and ‘no stim’), whereby each entry into the choice point (each lap) had a random 50% chance of receiving stim. During control sessions, no stimulation was delivered at the choice point, though a lap type was still recorded for proper comparison. For opto sessions, 3 seconds of light (Stim) was triggered upon entry into the choice point for a random 50% of laps and controlled with no

stimulation (Sham) on the remaining 50%. “No- opto” sessions did not deliver any light. The stimulation protocol that dictated the pattern of light delivered at the choice point (Opto or no-opto see stimulation parameters) and the delay side location (left, right) were randomized and remained unchanged for each session. These task parameters were counterbalanced across the 18 experimental sessions. No distinguishable differences were detected between sham and stim laps (Fig. S2A–C).

2.5 Stimulation Parameters

The LED attached to the ends of the fibers on each rat’s implant was controlled by an in-house stimulation generator/amplifier. Each LED was tested prior to surgical implantation and emitted light with approximately 7mW of power at the tip of the fiber prior to surgical implantation. Three patterns of light were used in experimental sessions: Noise [a white noise random walk], theta [7hz sine wave], and no-opto [no light delivered]. Again, no discernable differences were seen between the noise and theta stimulations (Fig. S1A–E). Therefore, data from noise and theta stimulation days were grouped together as ‘opto’ days and compared against the control ‘no-opto’ days.

2.6 Analysis

2.6.1 Final Delay—To assess each rat’s valuation for the larger-later and smaller-sooner sides, the preferred delay was determined by averaging the delay to the larger-later side for the last 20 laps of a given session. See Fig. 1G.

2.6.2 Behavioral tracking and Vicarious Trial and Error (VTE)—Rat movement through each session of the task was recorded by an overhead camera and positional tracking was accomplished by the detection of headstage LEDs. Pixels exceeding a user-defined luminance were digitized and time-stamped by an analog Cheetah data acquisition system (Neuralynx). The choice point was defined as beginning halfway up the center task arm and ending midway through the left and right choice arms. VTE for each lap was quantified by the z-scored integrated absolute change in angular velocity of the head within the choice point ($zIdPhi$) following the methods of Papale et al. 2012. This integrated angular velocity incorporates a measure of time through the choice point such that VTE reflects pausing and re-orienting behaviors. Events with a $zIdPhi$ score > 1 were considered VTE events. See Fig. 1E,F. Analysis was performed using in house programs written in MATLAB. The tracking for one male control rat was insufficient for VTE analysis but was sufficient for task procedures and thus all the other analyses. Thus, this rat is not included in the VTE analyses, but is included in the other analyses.

2.6.3 Final Delay Slope—The rate of change in delay to the larger-later side over the final 20 laps provided a measure for the degree of stability in the rat’s behavior by the end of the session. See Fig. 1H. We used the absolute value of this final delay slope because detecting changes in titration and alternation strategy patterns were independent of the direction of change for the delay.

2.6.4 Proportion of Adjustment Laps—Each individual lap was identified as either adjustment or alternation depending on if it was to the same (adjustment) or opposite

(alternation) side as the lap preceding it. Each session phase was determined by the proportion of these types of laps. To quantify the dynamics of this proportion on a single session, each lap was labeled with either a 1 (adjustment) or 0 (alternation). The proportion of adjustments was determined by averaging within each bin (encompassing 10% of the session) and visualized with a shaded error line plot. See Fig. 1I.

2.6.5 Statistics—Multiway analysis of variance (anovan, Matlab, Mathworks, Natick MA) was used to detect significance for the main effect of virus and session type as well as their interaction. Each metric was analyzed with two different session grouping methods: 1) separating session type by opto delivery (comparing baseline sessions to opto and no-opto experimental sessions) and 2) combining these two experimental groups (post opto) for comparison against the baseline (pre opto) data. For each analysis, the mean value for each rat was calculated such that $n = \# \text{ rats}$. Because individual rats were assigned to only one of the two virus groups, the rat variable was nested within the virus variable. P-values less than 0.05 are reported as significant.

3. Results

3.1 Training and Histology

Histology obtained post-experimentation confirmed viral targeting of the mPFC centered within the prelimbic subregion (Fig. 2C). Of the original 16 rat subjects, two (one active-virus and one control-virus) had incomplete experimental session data and were excluded from the study, leaving us with 14 subjects. 5/7 active-virus rats and 5/7 of the control-virus rats showed good bilateral transfection. 2/7 of the active and 1/7 of the control rats showed unilateral transfection. One control rat had poor transfection and was not included in the analyses. These data provided us with $n=7$ active rats and $n=6$ control rats.

At the end of training, both control and active cohorts demonstrated proficiency across all three primary metrics on the task: both cohorts titrated the delay to similar ending values (final delay) measured as the average delay for the last 20 laps (Fig. 2D, ANOVA, no effect of virus: $p=.75$ [$df=1$; $F=0.11$]) and they had similar slopes for this final delay average (Fig. 2E, ANOVA, no effect of virus: $p=0.50$ [$df=1$; $F = 0.48$]). Additionally, the baseline proportion of adjustment laps for each cohort were indistinguishable from one another. The two groups are not distinguishably different in any of these metrics, suggesting that their learned strategies were comparable prior to light delivery from the experimental phase.

3.2 Valuation Metric: Final Delay

To assess the effects of optogenetic stimulation on task performance, we first analyzed the number of laps and amount of reward earned between active and control cohorts over the 18 experimental sessions. No session-level opto effects were detected in either cohort when comparing opto to no-opto experimental conditions, but this analysis did suggest a difference in total lap number on experimental sessions between active and control cohorts (Fig. 3A, ANOVA, effect of rat: $p<10^{-10}$ [$df = 11$; $F = 29.4$]; a significant main effect of virus: $p=0.001$ [$df=1$; $F=17.9$]; no main effect of stimulation day (opto vs no-opto): $p=0.27$ [$df=1$; $F=1.4$]; and no interaction: $p=0.12$ [$df=1$; $F=2.9$]). When the opto and no-opto

sessions were combined into a single (post) group for each cohort, this virus effect was no longer significant, but a marked increase in the average number of laps (from pre to post) was detectable with no interaction effect (Fig. 3D, ANOVA, effect of rat: $p=0.001$ [df=11; $F=7.7$]; no main effect of virus: $p=0.14$ [df=1; $F=2.6$]; a significant effect of pre-vs-post: $p=0.01$ [df=1; $F=9.7$]; and no interaction: $p=0.35$ [df=1; $F=1.0$]). This suggests that active and control groups both showed improvement in task performance with extended exposure to the experimental conditions.

We then assessed how the increase in lap number influenced the total food earnings. Again, no effect of opto was detected at the session-level in either cohort, but this analysis suggested greater overall earnings in the control cohort compared to the active animals (Fig. 3B, ANOVA, effect of rat: $p<10^{-10}$ [df = 11; $F = 29.4$]; a significant main effect of virus: $p=0.0001$ [df=1; $F=11.3$]; no main effect of stimulation session (opto vs no-opto): $p=0.23$ [df=1; $F=1.6$]; and no interaction: $p=0.15$ [df=1; $F=2.4$]). In the combined pre vs. post analysis, no significant effects were detected. However, there was a trend toward an increase in overall food intake from pre to post and it appeared that this trend was driven primarily by the control rats (Fig. 3E, ANOVA, effect of rat: $p=0.001$ [df=11; $F=8.2$]; no main effect of virus: $p=0.25$ [df=1; $F=1.5$]; trend effect of pre-vs-post: $p=0.06$ [df=1; $F=4.5$]; no interaction: $p=0.15$ [df=1; $F=2.4$]).

We then asked if there were any effects of opto stimulation on the larger-later reward valuation. Both control and active cohorts titrated to similar final delays across opto and no-opto experimental sessions (Fig. 3C, ANOVA, effect of rat: $p=0.013$ [df = 11; $F = 9.4$]; no main effect of virus: $p=0.69$ [df=1; $F=0.16$]; no main effect of stimulation session (opto vs no-opto): $p=0.29$ [df=1; $F=1.2$]; and no interaction: $p=0.27$ [df=1; $F=1.3$]), and this value was not distinguishably different in experimental conditions when compared to baseline (pre) (Fig. 3F, ANOVA, no effect of rat: $p=0.07$ [df=11; $F=2.5$]; no main effect of virus: $p=0.60$ [df=1; $F=0.3$]; no main effect of pre-vs-post: $p=0.51$ [df=1; $F=0.5$]; and no interaction: $p=0.85$ [df=1; $F=0.04$]). This finding suggests that rats were consistently trading off wait-time for the larger reward such that the final delay was a preserved goal: driven up when the initial delay was below the goal and driven down when the initial delay was above it. Indeed, the probability of choosing the larger-later reward was strongly correlated with distance from the final delay equally for both cohorts and across both opto conditions (Fig. S4E). Additionally, there was an inverse correlation between final delay and initial delay for all animals with no effect of virus or session opto (Fig. S4B) – animals tended to have higher final delays the lower the initial delay condition was for a given session. These results were consistent with those seen in earlier experiments on this task (Papale et al. 2012; Bett et al. 2015; Kreher et al. 2019).

3.3 Strategy Metrics: Adjustment Laps and Final Delay Slope

The three-phase structure for this task can be described as a shift in strategy from exploration to titration to exploitation or maintenance behavior (Fig 1C–D). To analyze potential differences in the development and implementation of these strategies between the two groups, we compared the proportion of adjustment laps as rats progressed through each session. Consistent with previous studies (Papale et al. 2012; Bett et al. 2015; Kreher

et al. 2019), control rats showed a marked increase in adjustments during the titration phase followed by an increase in alternation late within each session during the exploitation or maintenance phase (Fig. 4C). This strategy employed by control animals during the experimental phase differed from their pre-opto baseline, with no apparent difference between opto and no-opto sessions (Fig. 4D). However, active-virus rats did not show this strategy development (Fig. 4E). Instead, they appeared to maintain a high late-session adjustment proportion, suggesting that they were continuing to titrate around their target delay, even to the end of the session. An example of this “unstable” strategy is shown compared to the “stable” phase structure (Fig. 4A–B) to illustrate how the stability or “plateau” in adjustment proportion is inversely related to that of the delay.

The key to maintaining a delay on this task is alternating between sides, which holds that delay constant over a pair of laps. If late-session alternation is occurring, we would expect to see a flat slope for the final delay indicative of its low rate of change brought on by alternation laps (Fig. 5A). In contrast, if late-session titration is occurring, we would expect to see the final delay continuing to change (either up or down) over those final laps (Fig. 5B). To quantify this final delay stability, we analyzed the absolute value of its slope.

Control animals showed significantly flatter final delay slopes compared to active-virus animals with no differences detected in either group between opto session type (Fig. 5D, ANOVA, effect of rat: $p=0.003$ [df=11; $F=5.9$]; effect of virus: $p=0.0001$ [df=1; $F=33.6$]; no main effect of stimulation session (opto-vs-no opto): $p=0.91$ [df=1; $F=0.01$]; no interaction: $p=0.46$ [df=1; $F=0.58$]). This significant effect between cohorts was upheld when experimental conditions were combined. There was a significant interaction detected between virus and pre vs post when compared to baseline (Fig. 5D, ANOVA, effect of rat: $p=0.04$ [df=11; $F=3.1$]; effect of virus: $p=0.007$ [df=1; $F=11.1$]; no main effect pre-vs-post: $p=0.21$ [df=1; $F=1.8$]; significant interaction between virus and pre-vs-post: $p=0.05$ [df=1; $F=4.85$]). Importantly, both groups showed comparable final delay slopes at the end of the training prior to the experimental days, but in the experimental days, the control animals decreased their final delay slopes while the active animals did not. This flatter final delay slope is indicative of a robust maintenance phase that developed over the course of experimentation in the control, but not the active, cohort.

3.4 Vicarious Trial and Error (VTE)

Lastly, we looked at the proportion of vicarious trial and error (VTE) events occurring at the choice point to probe how the two groups may differ in their deliberation. VTE is a behavioral event (Fig. 6A) in which rats pause at a choice point and re-orient back and forth between options (Muenzinger & Gentry, 1931; Muenzinger, 1938; Tolman, 1938, Redish, 2016). We quantified the overall proportion of VTE events and found no significant differences between cohorts and no effects of opto at the session level (Fig. S3A). The correlation between VTE and neurophysiological deliberation (Johnson & Redish 2007) prompted us to analyze its proportion in relation to the adjustment laps, which require flexibility and thus likely deliberation systems (Redish 2016). Note that previous studies of the Spatial Adjusting Delay Discounting task have found VTE to occur preferentially on adjustment laps during titration (Papale et al. 2012; Bett et al. 2015; Kreher et al. 2019).

Consistent with these previous studies, all animals demonstrated a greater proportion of VTE on adjustment laps compared to alternation laps (Fig. S3B). Focusing on adjustment lap VTE, no significant differences were detected between opto and no-opto sessions in either group, but this analysis suggested that active animals may have reduced VTE compared to the control cohort (Fig. 6D, ANOVA, no effect of rat: $p=0.14$ [df=10; $F=2.0$]; a significant effect of virus: $p=0.02$ [df=1; $F=7.5$]; no effect of stimulation session (opto-vs-no opto): $p=0.40$ [df=1; $F=0.77$]; and no interaction: $p=0.68$ [df=1; $F=0.18$]). When combined for pre-post analysis, we found no detectable differences in VTE between the two groups or from their baseline data (Fig. 6F, ANOVA, no effect of rat: $p=0.67$ [df=10; $F=0.75$]; no effect of virus $p=0.25$ [df=1; $F=1.5$]; no effect of pre-vs-post: $p=0.58$ [df=1; $F=0.33$]; and no interaction: $p=0.97$ [df=1; $F=0.0$]).

However, by examining VTE on adjustment laps occurring only late in the session (last 50%) (Fig. 6C), we found that the active cohort appeared to have decreased VTE compared to the control group. Though this effect appeared to be primarily driven by the active group's opto sessions, no detectable differences were seen between session-type (Fig. 6E, ANOVA, no effect of rat: $p=0.69$ [df=10; $F=0.72$]; significant effect of virus: $p=0.02$ [df=1; $F=7.7$]; no significant main effect of stimulation session: $p=0.28$ [df=1; $F=1.3$]; and no interaction: $p=0.95$ [df=1; $F=0$]). This effect of virus was upheld in the pre-post comparison (Fig. 6G, ANOVA, significant effect of rat: $p=0.009$ [df=10; $F=5.0$], significant effect of virus: $p=0.009$ [df=1; $F=10.4$]; no significant main effect of pre-vs-post: $p=0.31$ [df=1; $F=1.2$], but a significant interaction: $p=0.02$ [df=1; $F=7.4$]). Importantly, the significant interaction effect showed that, while control animals had no change in VTE from their baseline, active animals had a reduction in their VTE. This provides some evidence of deliberative impairment resulting from optogenetic manipulation of mPFC at the choice point. Note that late-session adjustment laps in control animals are rare events that likely entail a correction to procedural alternation between right/left choice behavior during alternation. In contrast, active-virus rats continued to show a plethora of adjustment laps, even during late phases of each session (Fig. 4C).

4. Discussion

The medial prefrontal cortex is known to be involved in multiple decision-making strategies and the ability to shift between them (McLaughlin, Diehl, & Redish 2021). We used in-vivo optogenetic manipulation to selectively disrupt the mPFC (targeting the prelimbic cortex) at the choice point on a spatial adjusting delay discounting task. With controls on the lap, session, and whole experimental scales, we sought to further understand the timing of these effects on short- and long-term strategy behavior. We found that the signatures of strategy shifting from deliberative titration to procedural alternation that rats typically show on the spatial adjusting delay-discounting task developed with experience in the control group but not in the active group, yet delay-valuation remained intact for both groups. Importantly, this differentiation in performance was revealed on the scale of the whole experiment; strategy signatures of the control group changed over sessions while the active animals' strategic behavior showed little variance from their pre-opto performance. This suggests an important role for prelimbic in strategy-shift learning, long-term procedural strategy development, or a combination of both.

Using the strategy-phase language developed in previous spatial adjusting delay-discounting experiments (Papale et al. 2012), control rats demonstrated strategic behavior consistent with the expected three-phase structure, identifiable as investigation, titration, and exploitation or maintenance. The important shift from titration to maintenance is evident by the emergence of prolonged alternation and creates a delay-stable structure which was robust in the control group. However, the strategy structure of the active cohort was different: they continued to show steeper delay slopes and a higher proportion of adjustment laps compared to the control group, which revealed a strategy profile with more titration and less alternation. This delay-unstable strategy profile appeared to reflect a state of prolonged titration, using adjustment laps to ‘over-correct’ and oscillate broadly around a preferred delay.

On the spatial DD task, titration has been linked to deliberation, and alternation has been linked to the procedural decision system due to rats’ natural tendency to alternate between reward sources on foraging tasks (Richman et al. 1986; Papale et al. 2012; Powell & Redish 2016). Vicarious trial and error is a more fine-scale behavior with robust evidence to support a correlation of VTE events with deliberative decision making (Johnson & Redish 2007; Bett et al 2012, 2015; Redish 2016, Kidder et al. 2021). Analyzing the choice point behavior during late-session adjustment laps showed an overall reduction in VTE for the active animals compared to the control group. We found that the proportion of these VTE events were inversely correlated with initial start delay (Fig. S4D). No correlation was detected between initial delay and proportion of late session adjustment laps (Fig. S4C). Proportion of adjustment laps were correlated with current lap delay, but no differences were detected between groups (Fig. S4F) and no differences were detected for correlation between lap delay and proportion of VTE (Fig. S4G).

Our data revealed no immediate (lap-by-lap) effects of optogenetic mPFC disruption on behavior at the choice point (Fig. S2A–C). This suggests that disruption to mPFC on this task contributed to impairments in deliberative behavioral flexibility that were more-likely critical for long-term strategy learning and development than short-term performance.

These results not only provide further support for mPFC’s role in deliberation, but they also offer further insight into the role that PL plays in the development of procedural strategies. The long-term deficits in active animal’s habit formation combined with short-term deliberative (but not procedural) impairment suggests an interesting interaction between deliberation and the development of habit strategies. Moreover, these data suggest that PL and its role in deliberation may also play a role in habit learning and development beyond strategy switching at a given moment.

Our VTE results are consistent with another similar optogenetic study (Kidder et al. 2021) on a slightly different Spatial Delayed Alternation task. Kidder et al. found deficits in task performance that were correlated with a reduction in VTE events following mPFC perturbation. They also found that these impairments were specific to mPFC disruption during points of deliberation (choice points) and the effects were seen on a session-by-session level. Kidder et al. concluded that deliberation relies on information from mPFC as it relates to working memory and that their data provided further evidence for the deliberative role of VTE. Taken together with the results from our study, one would expect

that a reduction in VTE is likely evidence for deliberative deficits. However, other studies have suggested that titration on the spatial adjusting delay-discounting task is a deliberative strategy (Papale et al. 2012; Powell & Redish 2016). How do those results align with our observation that PL-disrupted rats showed a decrease in alternation laps (maintenance) and an increase in adjustment laps (titration) while also showing decreased VTE?

Unlike the other decision-making tasks on which VTE has been studied (Hu & Amsel 1995; Johnson & Redish 2007; Amemiya et al. 2014; Papale et al. 2016; Bett et al. 2012; Kay et al. 2020; Kidder et al. 2021), the spatial adjusting delay discounting task does not have clear right/wrong choices. Rather, the choice patterns reveal the subjects' internally driven strategy shifts. Because strategic cues are not experimentally controlled, deliberation becomes more ambiguous to measure and VTE may be playing a more nuanced role in this task. For example, Kreher et al. (2019) found that disrupting the perirhinal cortex (PRC) with the GABA(A) agonist muscimol on the spatial delay discounting task led to similar strategy performance deficits (increased adjustment laps in late-stage behavior) but opposite VTE results (increased VTE at the choice point in late-stage behavior). Like our observations, this group found no significant differences in the indifference point (valuation) between the PRC-disrupted and undisrupted cohorts. Kreher et al. interpreted the strategy changes as an increase in deliberation which prevented procedural task stabilization.

An interesting possibility is that our PL disruption reduced the recognition of stability which impaired the development of a procedural strategy. Embedding the mPFC in the context of learning and memory as it relates to strategy shift signaling could provide a mechanistic framework that better fits all these data. Neural ensemble recordings have consistently found strategy-related representational transitions in PL (Durstewitz et al. 2010; Hyman et al. 2012; Ma et al. 2016), including on the spatial adjusting delay discounting task (Powell & Redish 2016). Powell & Redish found that PL ensembles changed their task-related characteristics with strategy changes, and that these representational transitions preceded behavioral transitions on the spatial adjusting delay discounting task. This suggests that mPFC is important for processing contextual information related to a need for strategy shift, and it follows that disruption to this region could inhibit the strategy change from titration to alternation on this task.

Other physiology experiments have sought to understand mechanisms of decision making by investigating the coherence between mPFC and hippocampus (HPC) due to its implication in spatial representation and memory consolidation. Benchenane et al. 2010 found that theta oscillations between these two regions peaked at the choice point on a Y maze during strategy transitions and suggested that this coherence supports memory consolidation for long-term reward prediction learning. Hasz and Redish (2020) found that mPFC and hippocampal ensembles dynamically represented task contingencies that updated with changing contexts. They found that transitions between representations occurred in mPFC before hippocampus. This suggests that disrupting mPFC might inhibit the information flow that communicates task context, which could result in strategy-change deficits. Behaviorally, this could appear as an inhibition in the formation of procedural action chains in the long-term.

Differentiating between a strategy-development and a strategy-switching mechanism for mPFC is difficult to do with only behavioral measurements. However, large-scale comparison between the experimental phase and pre-opto (baseline) data seemed to reveal that the control-virus cohort developed robust alternation over time that drove their strategy profile to differ from the active cohort, whose strategy appeared to remain unchanged through exposure to the task. This learning deficit would also explain our VTE data, as reduced VTE has been associated with learning and memory deficits particularly related to hippocampal disruption on spatial tasks (Hu & Amsel 1995, Bett et al. 2012). A hippocampal lesion study using this spatial DD task (Bett et al. 2015) also found reduced adjustment lap VTE and maintenance deficits while showing the same indifference in valuation between sham and lesion groups seen in our data (above) and other studies. Interestingly, the alternation deficits in Bett et al. were more apparent early in a session and recovered in later laps, suggesting a role for hippocampus in the early investigation phases and the settling down of the titration phase, but no hippocampal deficit in alternation, in contrast to the maintenance consequences of our PL disruption.

Broadly put, these data suggest that PL may be particularly important for communicating and building context associations that are necessary for connecting a situation with a decision strategy that is appropriate for the degree of behavioral flexibility perceived to be required. The long-standing hypothesis aligning deliberative decision processes and the prelimbic cortex points to its importance in cognitive flexibility and planning (Fuster 1997; Ragozzino et al. 1999; Killcross & Coutureau 2003; Dalley, Cardinal, & Robbins 2004; Tran-Tu-Yen et al. 2009). However, complexities in navigating dynamic environments point to the need for a more nuanced role for the mPFC to account for behavioral tasks with multiple decision-system interactions. Our results highlight the importance of PL in cognitive flexibility but suggest broader effects than on deliberative function alone. With no clear right/wrong choice metric, the spatial adjusting delay discounting task measures subjective valuation using the animal's preferred delay (final delay). This value was preserved through PL disruption, as were short-scale (lap-by-lap) deliberative behaviors (titration, VTE). It was only at a large-scale (across sessions) that strategy shifts appeared between the two cohorts, whereby active animals showed reduced VTE on late-session adjustment laps and showed deficits in the development of procedural strategy markers. Our data thus suggests that the PL region of medial prefrontal cortex contributes to flexible decision-making processes essential for strategy learning and shifting from deliberative into procedural strategies while keeping valuation algorithms intact.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Kelsey Seeland for providing technical support in optic-fiber implant design, construction, and surgical procedures, Ayaka Sheehan for performing histology, Elizabeth Dean and Anneke Duin for piloting this experiment that yielded promising preliminary data, and Grant Noble and Kevin Singh for their help training and running rats. This work was funded by NIH R01 MH112688.

5. References

- Adhikari A, Topiwala MA, Gordon JA (2010). Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety. *Neuron*, 65(2), 257–269. 10.1016/j.neuron.2009.12.002 [PubMed: 20152131]
- Amemiya S, Noji T, Kubota N, Nishijima T, Kita I. (2014). Noradrenergic modulation of vicarious trial-and-error behavior during a spatial decision-making task in rats. *Neuroscience*, 265, 291–301. 10.1016/j.neuroscience.2014.01.031 [PubMed: 24480363]
- Amemiya S, Redish AD (2016). Manipulating Decisiveness in Decision Making: Effects of Clonidine on Hippocampal Search Strategies. *Journal of Neuroscience*, 36(3), 814–827. 10.1523/JNEUROSCI.2595-15.2016 [PubMed: 26791212]
- Balleine BW, Delgado MR, Hikosaka O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, 27(31), 8161–8165. 10.1523/JNEUROSCI.1554-07.2007 [PubMed: 17670959]
- Barker JM, Glen WB, Linsenbardt DN, Lapish CC, Chandler LJ (2017). Habitual Behavior Is Mediated by a Shift in Response-Outcome Encoding by Infralimbic Cortex. *eNeuro*, 4(6). 10.1523/ENEURO.0337-17.2017
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, 437(7062), 1158–1161. 10.1038/nature04053 [PubMed: 16237445]
- Benchenane K, Peyrache A, Khamassi M, Tierney PL, Gioanni Y, Battaglia FP, Wiener SI (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron*, 66(6), 921–936. 10.1016/j.neuron.2010.05.013 [PubMed: 20620877]
- Bett D, Allison E, Murdoch LH, Kaefer K, Wood ER, Dudchenko PA (2012). The neural substrates of deliberative decision making: contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. *Frontiers in Behavioral Neuroscience*, 6, 70. 10.3389/fnbeh.2012.00070 [PubMed: 23115549]
- Bett D, Murdoch LH, Wood ER, Dudchenko PA (2015). Hippocampus, delay discounting, and vicarious trial-and-error. *Hippocampus* 25(5), 643–654. 10.1002/hipo.22400 [PubMed: 25483408]
- Chudasama Y, Robbins TW (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional *Journal of Neuroscience*. <https://www.jneurosci.org/content/23/25/8771.short>
- Coutureau E, Killcross S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behavioural Brain Research*, 146(1–2), 167–174. 10.1016/j.bbr.2003.09.025 [PubMed: 14643469]
- Dalley JW, Cardinal RN, Robbins TW (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience and Biobehavioral Reviews*, 28(7), 771–784. 10.1016/j.neubiorev.2004.09.006 [PubMed: 15555683]
- Dalton GL, Wang NY, Phillips AG, Floresco SB (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *Journal of Neuroscience*, 36(6), 1996–2006. 10.1523/JNEUROSCI.3366-15.2016 [PubMed: 26865622]
- Redish AD (2013). *The Mind Within the Brain: How We Make Decisions and How Those Decisions Go Wrong*. Oxford University Press.
- Daw ND, Niv Y, Dayan P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience* 8(12), 1704–1711. 10.1038/nn1560 [PubMed: 16286932]
- Dickinson A. (1994). Instrumental conditioning. *Animal Learning and Cognition*, 45–79. https://www.researchgate.net/profile/Anthony-Dickinson-3/publication/303655719_Instrumental_Conditioning/links/5ecba21d92851c11a888502c/Instrumental-Conditioning.pdf

- Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, 66(3), 438–448. 10.1016/j.neuron.2010.03.029 [PubMed: 20471356]
- Euston DR, Gruber AJ, McNaughton BL (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76(6), 1057–1070. 10.1016/j.neuron.2012.12.002 [PubMed: 23259943]
- Evenden JL, Ryan CN (1996). The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* 128(2), 161–170. 10.1007/s002130050121 [PubMed: 8956377]
- Floresco SB, Seamans JK, Phillips AG (1997). Selective roles for hippocampal, prefrontal cortical, and ventral striatal circuits in radial-arm maze tasks with or without a delay. *Journal of Neuroscience*, 17(5), 1880–1890. <https://www.ncbi.nlm.nih.gov/pubmed/9030646> [PubMed: 9030646]
- Fuster JM (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Lippincott-Raven. <https://play.google.com/store/books/details?id=YupqAAAAMAAJ>
- Gardner RS, Uttaro MR, Fleming SE, Suarez DF, Ascoli GA, Dumas TC (2013). A secondary working memory challenge preserves primary place strategies despite overtraining. *Learning and memory*, 20(11), 648–656. 10.1101/lm.031336.113 [PubMed: 24136182]
- Gilbert DT, Wilson TD (2007). Propection: experiencing the future. *Science*, 317(5843), 1351–1354. 10.1126/science.1144161 [PubMed: 17823345]
- Giordano LA, Bickel WK, Loewenstein G, Jacobs EA, Marsch L, Badger GJ (2002). Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money. *Psychopharmacology*, 163(2), 174–182. 10.1007/s00213-002-1159-2 [PubMed: 12202964]
- Graybiel AM (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, 70(1–2), 119–136. 10.1006/nlme.1998.3843 [PubMed: 9753592]
- Graybiel AM (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31, 359–387. 10.1146/annurev.neuro.29.051605.112851
- Guise KG, Shapiro ML (2017). Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron*, 94(1), 183–192.e8. 10.1016/j.neuron.2017.03.011 [PubMed: 28343868]
- Hasz BM, Redish AD (2020). Dorsomedial prefrontal cortex and hippocampus represent strategic context even while simultaneously changing representation throughout a task session. *Neurobiology of Learning and Memory*, 171, 107215. 10.1016/j.nlm.2020.107215 [PubMed: 32276121]
- Heidbreder CA, Groenewegen HJ (2003). The medial prefrontal cortex in the rat: evidence for a dorso-ventral distinction based upon functional and anatomical characteristics. *Neuroscience and Biobehavioral Reviews*, 27(6), 555–579. 10.1016/j.neubiorev.2003.09.003 [PubMed: 14599436]
- Hok V, Chah E, Save E, Poucet B. (2013). Prefrontal cortex focally modulates hippocampal place cell firing patterns. *Journal of Neuroscience*, 33(8), 3443–3451. 10.1523/JNEUROSCI.3427-12.2013 [PubMed: 23426672]
- Hull CL (1943). *Principles of behavior: an introduction to behavior theory*. 422. <https://psycnet.apa.org/fulltext/1944-00022-000.pdf>
- Hyman JM, Ma L, Balaguer-Ballester E, Durstewitz D, Seamans JK (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13), 5086–5091. 10.1073/pnas.1114415109 [PubMed: 22421138]
- Ito HT, Zhang S-J, Witter MP, Moser EI, Moser M-B (2015). A prefrontal–thalamo–hippocampal circuit for goal-directed spatial navigation. *Nature*, 522(7554), 50–55. 10.1038/nature14396 [PubMed: 26017312]
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999). Building neural representations of habits. *Science*, 286(5445), 1745–1749. 10.1126/science.286.5445.1745 [PubMed: 10576743]
- Johnson A, Redish AD (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45), 12176–12189. 10.1523/JNEUROSCI.3761-07.2007 [PubMed: 17989284]

- Jones MW, Wilson MA (2005). Theta Rhythms Coordinate Hippocampal–Prefrontal Interactions in a Spatial Memory Task. *PLoS Biology*, 3(12), e402. 10.1371/journal.pbio.0030402 [PubMed: 16279838]
- Kay K, Chung JE, Sosa M, Schor JS, Karlsson MP, Larkin MC, Liu DF, Frank LM (2020). Constant Sub-second Cycling between Representations of Possible Futures in the Hippocampus. *Cell*, 180(3), 552–567. 10.1016/j.cell.2020.01.014 [PubMed: 32004462]
- Kesner RP, Churchwell JC (2011). An analysis of rat prefrontal cortex in mediating executive function. *Neurobiology of Learning and Memory*, 96(3), 417–431. 10.1016/j.nlm.2011.07.002 [PubMed: 21855643]
- Kidder KS, Miles JT, Baker PM, Hones VI, Gire DH, Mizumori SJY (2021). A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays. *Hippocampus*, 31(7), 690–700. 10.1002/hipo.23306 [PubMed: 33507595]
- Killcross S, Coutureau E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, 13(4), 400–408. 10.1093/cercor/13.4.400 [PubMed: 12631569]
- Kreher MA, Johnson SA, Mizell J-M, Chetram DK, Guenther DT, Lovett SD, Setlow B, Bizon JL, Burke SN, Maurer AP (2019). The perirhinal cortex supports spatial intertemporal choice stability. *Neurobiology of Learning and Memory*, 162, 36–46. 10.1016/j.nlm.2019.05.002 [PubMed: 31125611]
- Laubach M, Amarante LM, Swanson K, White SR (2018). What, If Anything, Is Rodent Prefrontal Cortex? *eNeuro*, 5(5). 10.1523/ENEURO.0315-18.2018
- Lempert KM, Steinglass JE, Pinto A, Kable JW, Simpson HB (2019). Can delay discounting deliver on the promise of RDoC?. *Psychological medicine*, 49(2), 190–199. 10.1017/S0033291718001770 [PubMed: 30070191]
- Ma L, Hyman JM, Durstewitz D, Phillips AG, Seamans JK (2016). A Quantitative Analysis of Context-Dependent Remapping of Medial Frontal Cortex Neurons and Ensembles. *Journal of Neuroscience*, 36(31), 8258–8272. 10.1523/JNEUROSCI.3176-15.2016 [PubMed: 27488644]
- Madden GJ, Bickel WK (2010). *Impulsivity: The Behavioral and Neurological Science of Discounting*. American Psychological Association. https://play.google.com/store/books/details?id=_A9sPgAACAAJ
- Madden GJ, Petry NM, Badger GJ, Bickel WK (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology* 5(3), 256–262. 10.1037/1064-1297.5.3.256 [PubMed: 9260073]
- Mazur JE (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning and Behavior*, 25(2), 131–147. 10.3758/BF03199051
- McLaughlin AE, Diehl GW, Redish AD (2021). Potential roles of the rodent medial prefrontal cortex in conflict resolution between multiple decision-making systems. *International Review of Neurobiology*, 158, 249–281. 10.1016/bs.irn.2020.11.009 [PubMed: 33785147]
- Miller EK, Cohen JD (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202. 10.1146/annurev.neuro.24.1.167
- Mischel W, Ebbsen EB, Zeiss AR (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, 21(2), 204–218. 10.1037/h0032198 [PubMed: 5010404]
- Mitchell SH (2004). Measuring Impulsivity and Modeling Its Association With Cigarette Smoking. *Behavioral and Cognitive Neuroscience Reviews* 3(4), 261–275. 10.1177/1534582305276838 [PubMed: 15812110]
- Muenzinger KF (1938). Vicarious Trial and Error at a Point of Choice: I. A General Survey of its Relation to Learning Efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology*, 53(1), 75–86. 10.1080/08856559.1938.10533799
- Muenzinger KF, Gentry E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology*, 12(2), 195–206. 10.1037/h0072238
- Niv Y, Joel D, Dayan P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, 10(8), 375–381. 10.1016/j.tics.2006.06.010 [PubMed: 16843041]
- O’keefe J, Nadel L. (1978). *The hippocampus as a cognitive map*. Oxford university press. <https://discovery.ucl.ac.uk/id/eprint/10103569/1/HCMComplete.pdf>

- Odum AL, Madden GJ, Bickel WK (2002). Discounting of delayed health gains and losses by current, never- and ex-smokers of cigarettes. *Nicotine and Tobacco Research*, 4(3), 295–303. 10.1080/14622200210141257 [PubMed: 12215238]
- Papale AE, Stott JJ, Powell NJ, Regier PS, Redish AD (2012). Interactions between deliberation and delay-discounting in rats. *Cognitive, Affective and Behavioral Neuroscience*, 12(3), 513–526. 10.3758/s13415-012-0097-7
- Papale AE, Zielinski MC, Frank LM, Jadhav SP, Redish AD (2016). Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. *Neuron*, 92(5), 975–982. 10.1016/j.neuron.2016.10.028 [PubMed: 27866796]
- Peyrache A, Khamassi M, Benchenane K, Wiener SI, Battaglia FP (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature neuroscience*, 12(7), 919–926. 10.1038/nn.2337 [PubMed: 19483687]
- Powell NJ, Redish AD (2014). Complex neural codes in rat prelimbic cortex are stable across days on a spatial decision task. *Frontiers in Behavioral Neuroscience*, 8, 120. 10.3389/fnbeh.2014.00120 [PubMed: 24795579]
- Powell NJ, Redish AD (2016). Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behavior. *Nature Communications*, 7, 12830. 10.1038/ncomms12830
- Ragozzino ME, Wilcox C, Raso M, Kesner RP (1999). Involvement of rodent prefrontal cortex subregions in strategy switching. *Behavioral Neuroscience*, 113(1), 32–41. 10.1037//0735-7044.113.1.32 [PubMed: 10197904]
- Redish AD (2016). Vicarious trial and error. *Nature Reviews. Neuroscience*, 17(3), 147–159. 10.1038/nrn.2015.30 [PubMed: 26891625]
- Riaz S, Puveendrakumaran P, Khan D, Yoon S, Hamel L, Ito R. (2019). Prelimbic and infralimbic cortical inactivations attenuate contextually driven discriminative responding for reward. *Scientific Reports*, 9(1), 3982. 10.1038/s41598-019-40532-7 [PubMed: 30850668]
- Rich EL, Shapiro ML (2007). Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *Journal of Neuroscience*, 27(17), 4747–4755. 10.1523/JNEUROSCI.0369-07.2007 [PubMed: 17460087]
- Richman CL, Dember WN, Kim P. (1986). Spontaneous alternation behavior in animals: A review. *Current Psychology*, 5(4), 358–391. 10.1007/bf02686603
- Schmidt B, Papale A, Redish AD, Markus EJ (2013). Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. *Learning and memory (Cold Spring Harbor, N.Y.)*, 20(3), 130–138. 10.1101/lm.028753.112 [PubMed: 23418392]
- Schmidt B, Redish AD (2021). Disrupting the medial prefrontal cortex with designer receptors exclusively activated by designer drug alters hippocampal sharp-wave ripples and their associated cognitive processes. *Hippocampus* 31(10), 1051–1067. 10.1002/hipo.23367 [PubMed: 34107138]
- Schmidt B, Duin AA, Redish AD (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *Journal of Neurophysiology*, 121(6), 1981–2000. 10.1152/jn.00793.2018 [PubMed: 30892976]
- Smith KS, Graybiel AM (2013a). A Dual Operator View of Habitual Behavior Reflecting Cortical and Striatal Dynamics. *Neuron* 79(3), 608. 10.1016/j.neuron.2013.07.032
- Smith KS, Graybiel AM (2013b). Using optogenetics to study habits. *Brain Research* 1511, 102–114. 10.1016/j.brainres.2013.01.008 [PubMed: 23313580]
- Stein JS, Daniel TO, Epstein LH, Bickel WK (2015). Episodic future thinking reduces delay discounting in cigarette smokers. *Drug and Alcohol Dependence* 156, e212. 10.1016/j.drugalcdep.2015.07.571
- Sul JH, Kim H, Huh N, Lee D, Jung MW (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66(3), 449–460. 10.1016/j.neuron.2010.03.033 [PubMed: 20471357]
- Sweis BM, Larson EB, Redish AD, Thomas MJ (2018). Altering gain of the infralimbic-to-accumbens shell circuit alters economically dissociable decision-making algorithms. *Proceedings of the National Academy of Sciences of the United States of America*, 115(27), E6347–E6355. 10.1073/pnas.1803084115 [PubMed: 29915034]

- Tolman EC (1939). Prediction of vicarious trial and error by means of the schematic sowbug. *Psychological Review*, 46(4), 318–336. 10.1037/h0057054
- Tran-Tu-Yen DAS, Marchand AR, Pape J-R, Di Scala G, Coutureau E. (2009). Transient role of the rat prelimbic cortex in goal-directed behaviour. *European Journal of Neuroscience* 30(3), 464–471. 10.1111/j.1460-9568.2009.06834.x [PubMed: 19614748]
- van Aerde KI, Heistek TS, Mansvelder HD (2008). Prelimbic and infralimbic prefrontal cortex interact during fast network oscillations. *PloS One*, 3(7), e2725. 10.1371/journal.pone.0002725 [PubMed: 18628964]
- van der Meer M, Kurth-Nelson Z, Redish AD (2012). Information Processing in Decision-Making Systems. *The Neuroscientist* 18(4), 342–359. 10.1177/1073858411435128 [PubMed: 22492194]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

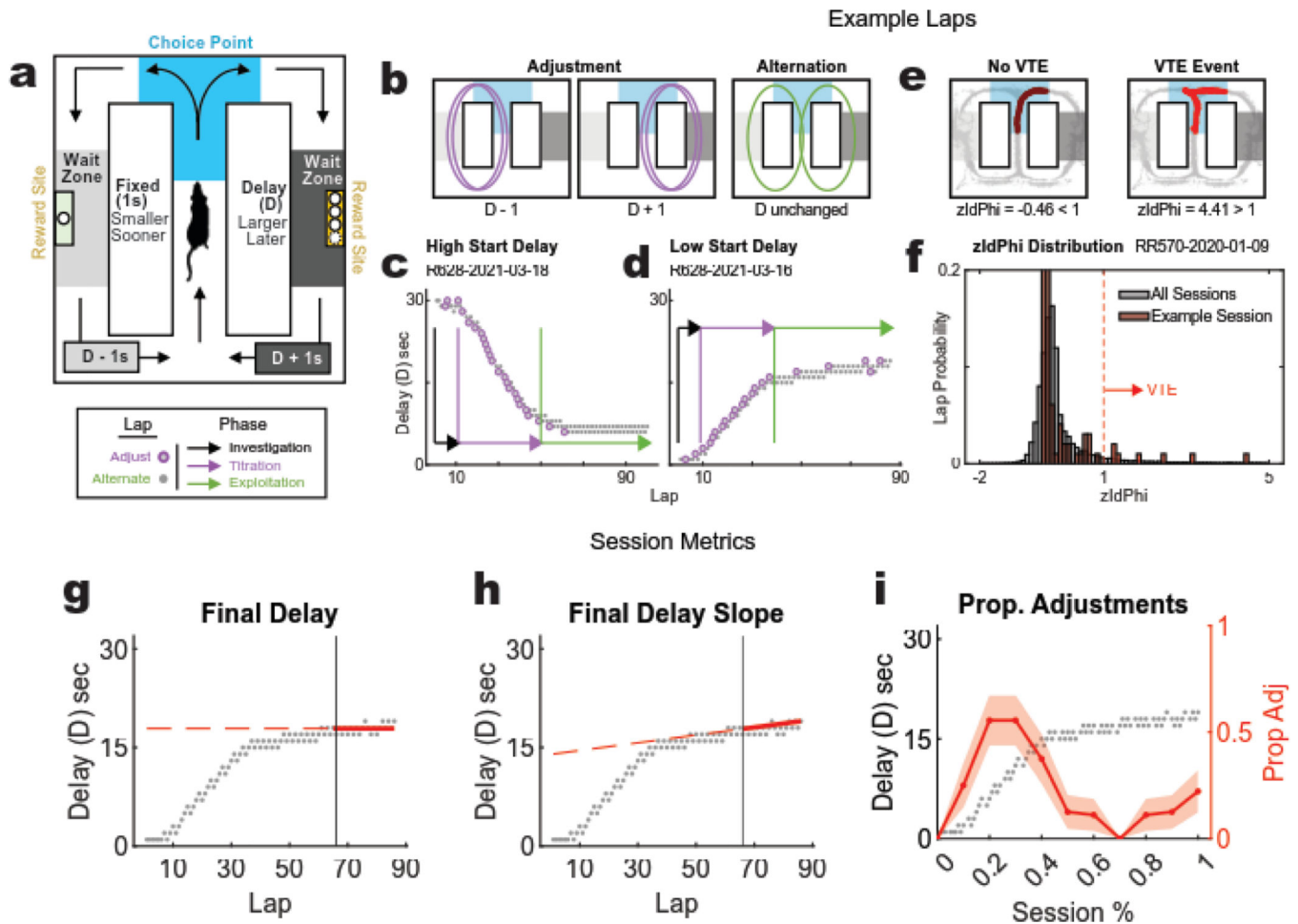


Figure 1: Spatial Delay Discounting task, behavior phases, and metrics

(a) DD Task Schematic: rats repeatedly faced left/right decisions between a small (1 pellet) reward delivered after a fixed 1s delay, and a large (3 pellets for females, 4 for males) reward dispensed after a variable delay (D). Rats manipulated the variable delay throughout each 45-minute session by increasing it by 1s each time the larger reward was earned and decreasing it by 1s each time the smaller one was earned. The delay was indicated by audible tones that were triggered upon entry into the reward zones (delay tone decreased in 250 Hz increments, the higher the delay tone pitch, the longer the delay). (b) Adjustment vs. Alternation Laps: on each lap, the rat can either adjust the delay down (left) or up (middle) by turning to the same side as the previous lap, or the rat can alternate (right) to maintain the delay. Well-trained rats typically exhibit three distinct behavioral phases. (c, d) Example sessions: the change in delay to the delayed side across laps as the rat first investigates, then titrates the delay down (c) and up (d) until reaching a preferred duration that is held steady through exploitation. (e) Example VTE laps: the rat's trajectory through the choice point was assessed for each lap. Paths with angular shifts (right) indicate a VTE event, whereas smooth paths (left) do not. (f) Distribution of zIdPhi: zIdPhi was calculated for each trajectory through the choice point. The distribution of these values for all experimental sessions (gray) is shown with that of one example session (red). Laps with zIdPhi > 1 were defined as VTE events. (g, h, i) Session-level metrics: (g) The average delay across the final

20 laps of a session (Final delay) measures this preferred delay that the rat is willing to take to balance out the larger reward. (h) The absolute value of the slope of the average final delay, and (i) the proportion of adjustment laps (consecutive laps to the same side) are markers to distinguish titration from exploitation.

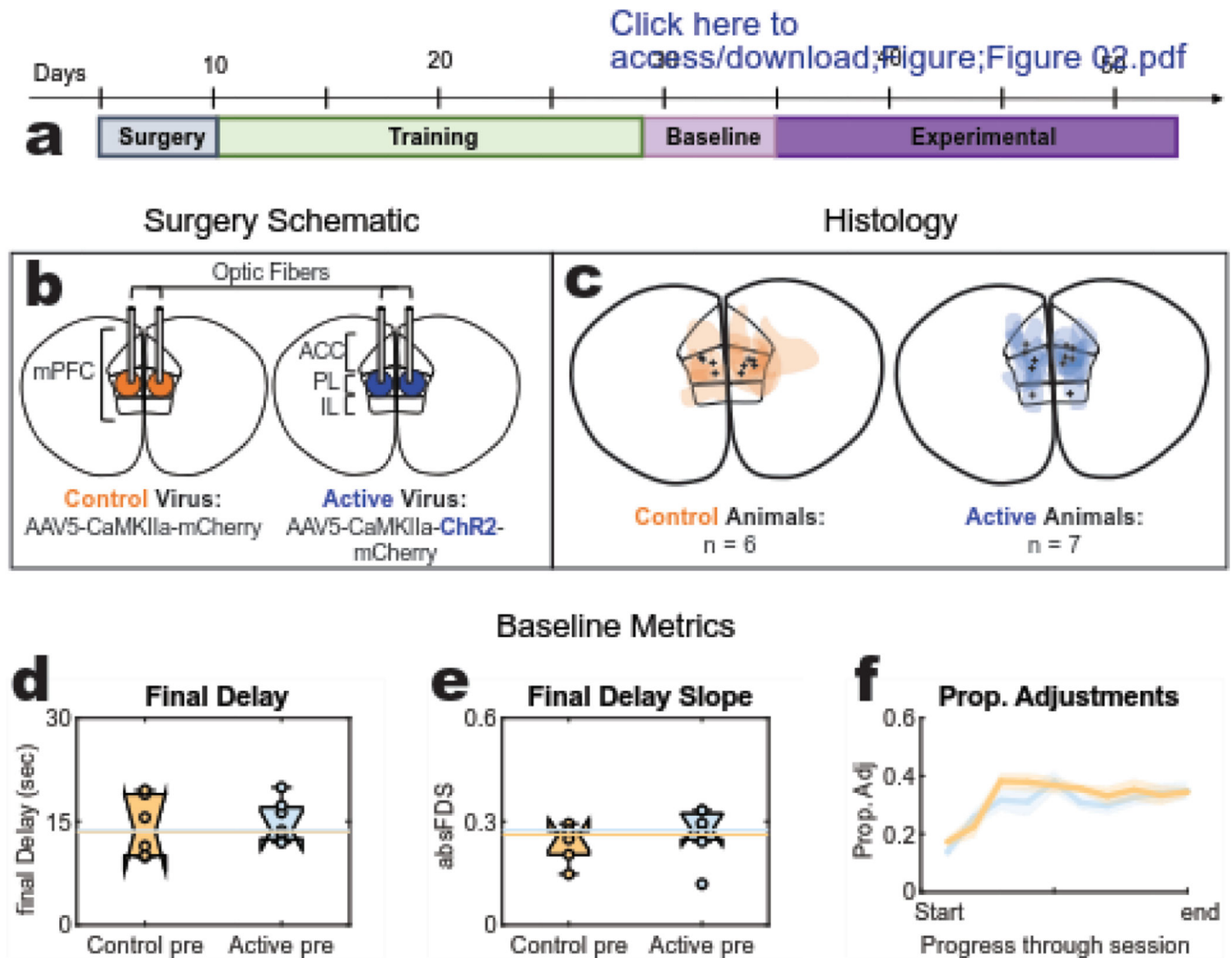


Figure 2: Task stages, Histology, and baseline behaviour

(a) Task Stages: Following surgery recovery, rats entered step-wise training blocks that introduced them to the DD task until they were running reliably with a head tether. Baseline behavior was collected for 6 days prior to the 18 day opto-delivery experimental sequence

(b) Surgery schematic: the prelimbic (PL) region was targeted for bilateral injection of either an active (opsin expressing) or control (opsin absent) virus and optic fiber implantation.

(c) Histology: target variation for viral expression (color splotches) and fiber tips (black marks) of each subject in either cohort (three rats presented with unilateral expression and will be identified in the results with dashed lines).

(d-f) Baseline session-level metrics: “pre” data was collected during the baseline stage that confirmed both cohorts were performing sufficiently on the task by achieving comparable (d) final delay values, (e) final delay slopes, and (f) proportion of adjustment laps.

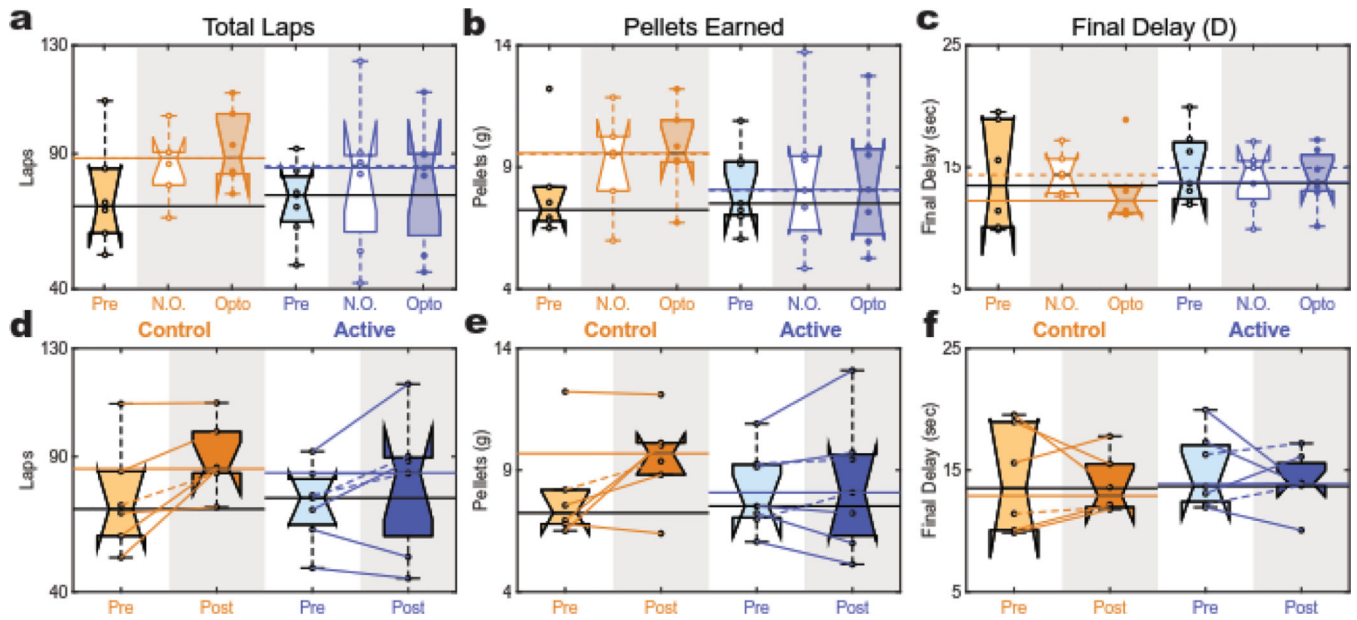


Figure 3: Task Performance and Final Delay

(a,d) Total Laps: Both cohorts had significant increases in total lap average from (d) baseline to experimental (post) sessions. (a) No difference was detected between opto and no opto sessions in either cohort. (b,e) Total Pellets: No change was detected in total pellet earning in the active animals. However, the control group showed a trend to increase pellet intake from pre to post (e) that was not detected in the active group. (b) No difference was detected between opto and no opto sessions in either cohort. (c,f) Final Delay: Both cohorts showed no change in their average final delay when compared to their baseline (f). (c) No difference was detected between opto and no opto sessions in either cohort. These data suggest the subjective valuation of the larger-later reward (measured by final delay) was stable across time and comparable between groups.

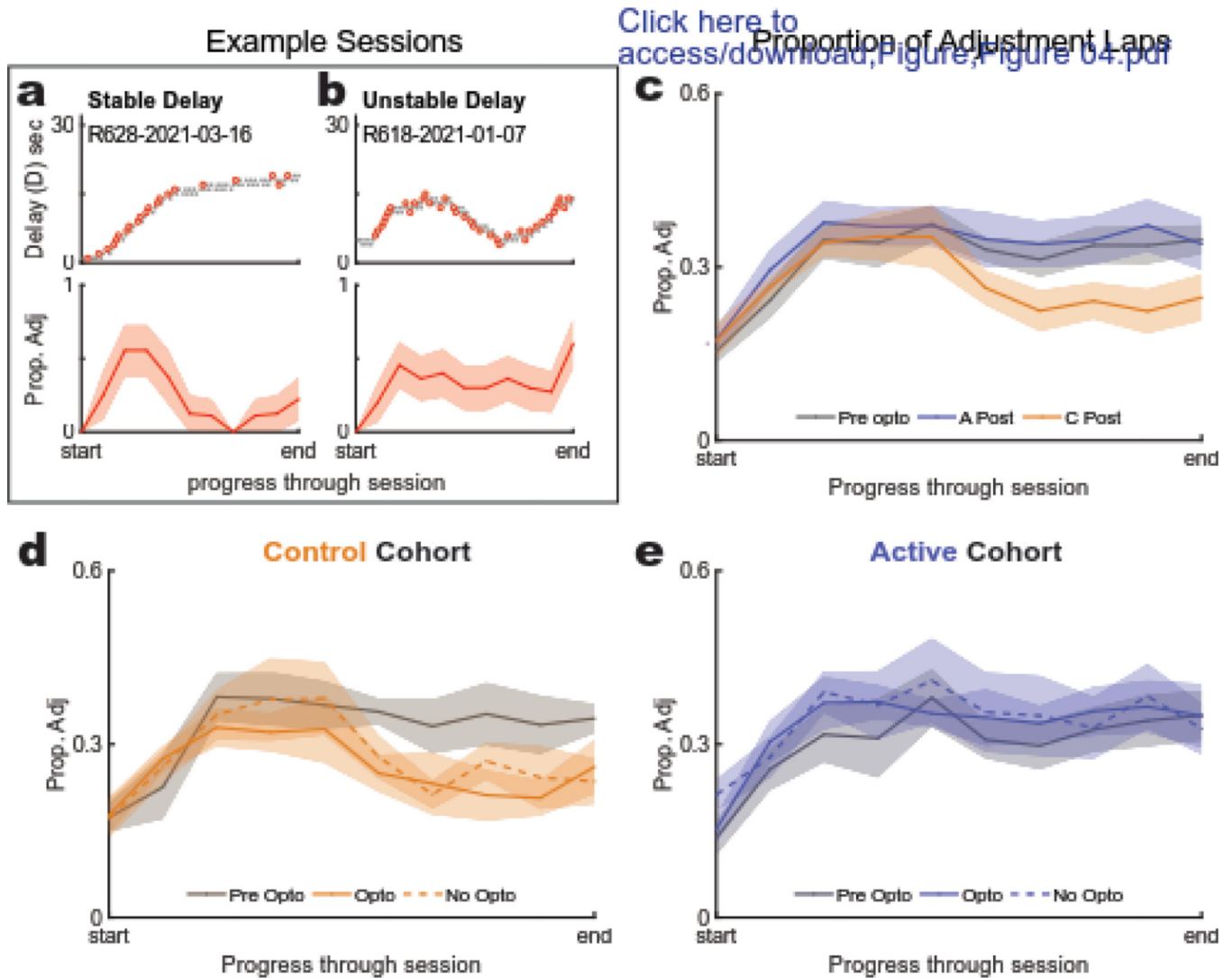


Figure 4: Proportion of Adjustment Laps

(a,b) Example sessions: the adjustment laps (top) and their running proportion to total laps (bottom) for a typical three phase structure for a session in which the rat maintained a stable delay in the maintenance/exploitation phase (a) and a session in which the rat did not (b). (c) Baseline to experimental (pre-post) comparison: Post stage proportion of adjustment laps for the active group, control group, and their combined baseline data. Each session was normalized on the same start-end scale (from the first to the last lap). (d,e) Effect of opto between cohorts: Proportion of adjustment laps (Prop Adj) during baseline and experimental sessions for control (d) and active (e) cohorts. (d) Control animals showed a significant decrease in proportion of adjustment laps in their late halves of their experimental (post) sessions compared to their baseline sessions. (e) This developing decrease was not seen in the active cohort. No detectable effects were noted between opto and no-opto sessions in either cohort.

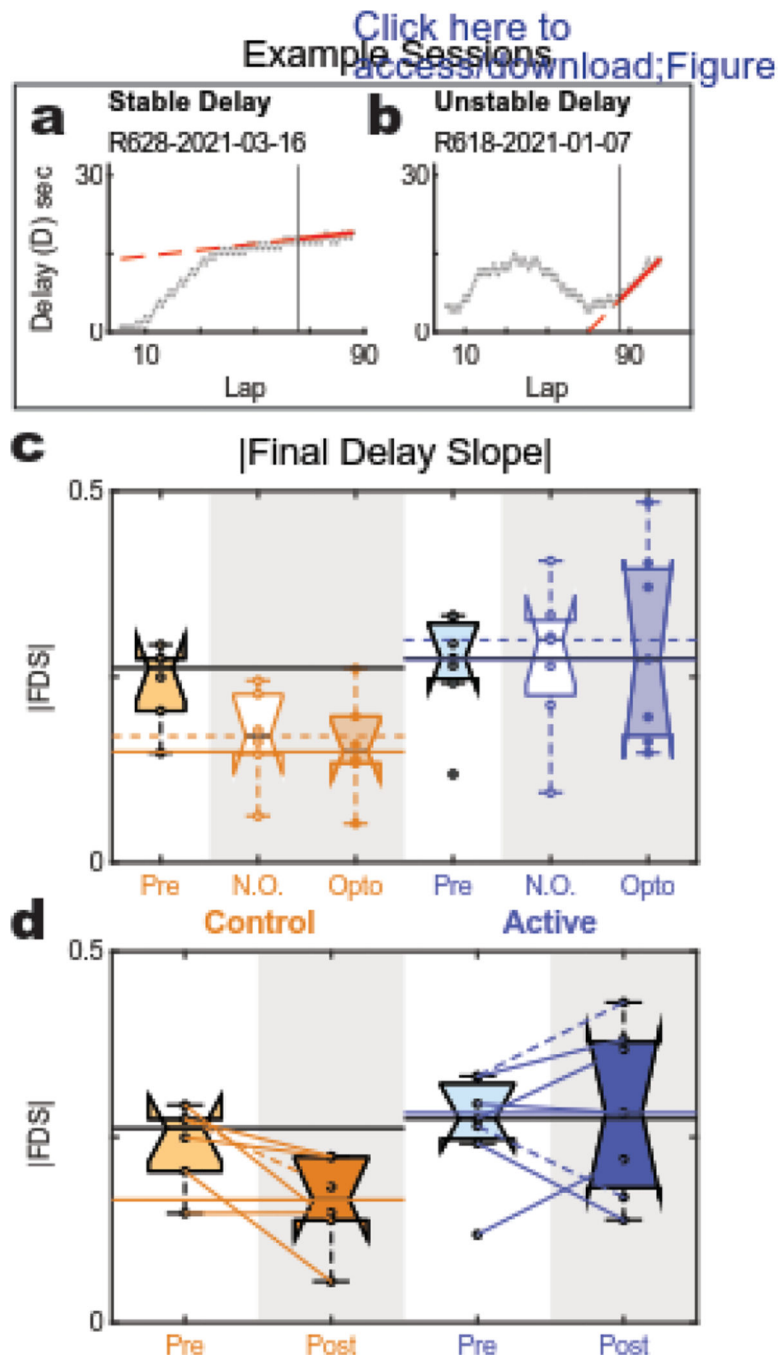


Figure 5: Final Delay Slope

(a,b) Example sessions: the final delay slope for a typical three phase structure on a session in which the rat maintained a stable delay in the last half of the session (a) and a session in which the rat did not (b). (c) Effect of opto between cohorts: No difference was detected between opto and no opto sessions in either cohort. (d) Overall effect of opto on Final Delay Slope: Control animals showed a significant flattening in their post sessions compared to their baseline sessions. Active-cohort animals showed no change in Final Delay Slope between their baseline (pre) and experimental (post) sessions.

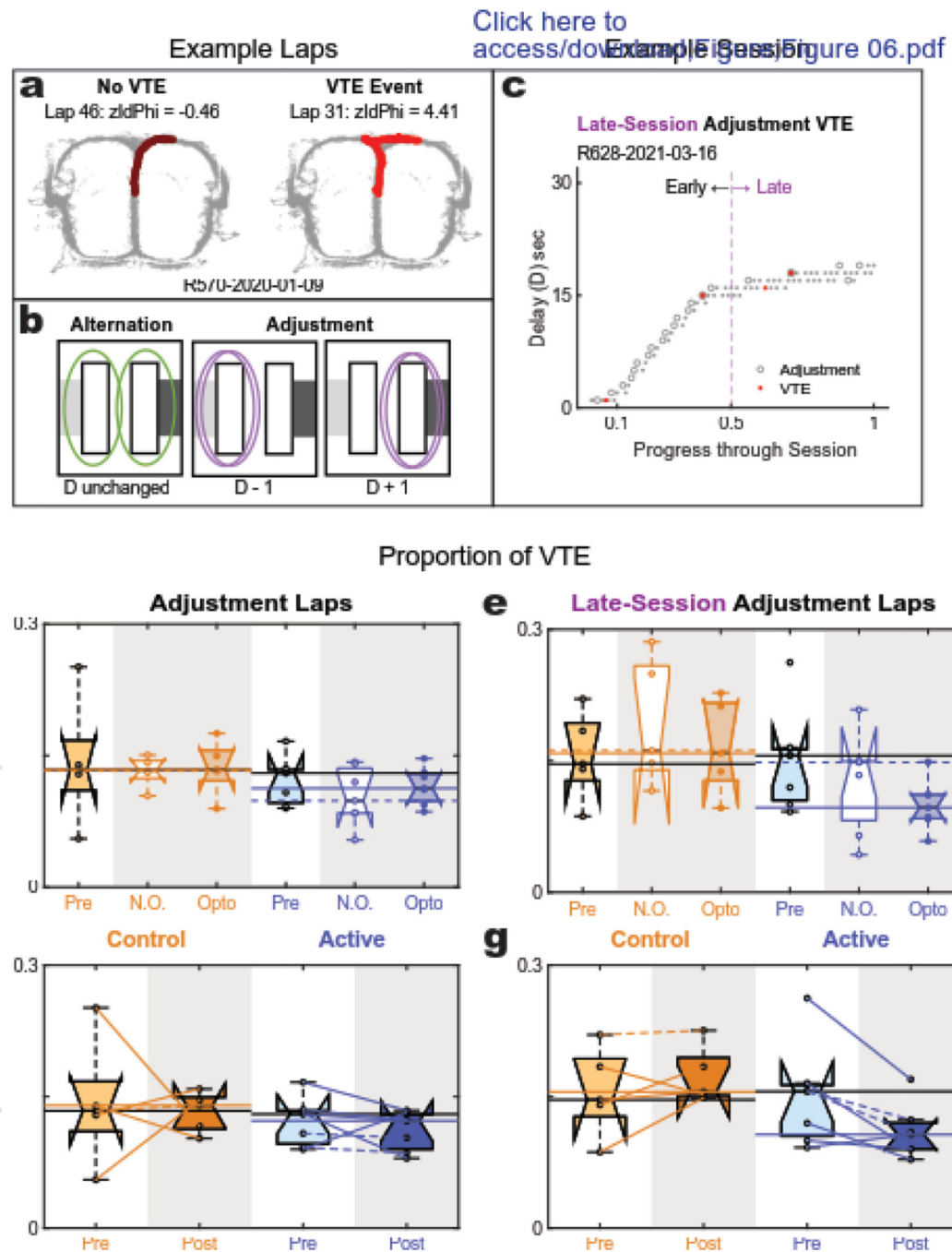


Figure 6: Vicarious trial and error (VTE) behaviors

(a) Example VTE laps: spatial tracking of a rat through the choice point during a non-VTE (left) and VTE (right) lap. (b) Schematic reminder of alternation vs adjustment laps. (c) Example Session: Adjustment Laps are marked with open circles and VTE is marked in red. Laps in the last 50% of the session were identified as “late” laps. (d,f) Proportion of adjustment laps showing VTE: No difference was detected between either group and no effect of opto was detected at the session level. However, when restricting the analysis to adjustment laps occurring only in the later half of the session (e,g), the proportion of VTE

for the post-opto active group was found to be significantly less than the baseline and control group sessions. Interestingly, this reduction in VTE appeared to be driven primarily by the opto sessions.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript