Title: Dorsomedial prefrontal neural ensembles reflect changes in task utility

Running title: dmPFC activity corresponds to task utility

Authors: Blake S. Porter and Kristin L. Hillman

Department of Psychology

Brain Health Research Centre

University of Otago

275 Leith Walk, Dunedin 9016, New Zealand

Corresponding Author: Blake S. Porter, blakeporterneuro@gmail.com

Conflict of Interest: The authors declare no competing financial interests.

Funding: This study was supported by Marsden Fund grant U001617 (K.L.H.) from the Royal Society of New Zealand Te Apārangi.

Abstract

When performing a physically demanding behavior, sometimes the optimal choice is to quit the behavior rather than persist and waste time and energy. The dorsomedial prefrontal cortex (dmPFC), consisting of the anterior cingulate cortex and secondary motor area, likely contributes towards such utility assessments. Here, we examined how rodent dmPFC single unit and ensemble level activity corresponded to changes in motivation and quitting in an effortful weight lifting task. Rats carried out two task paradigms: one that became progressively more physically demanding over time and a second fixed effort version. Rats could quit the task at any time. Dorsomedial PFC neurons were highly responsive to each behavioral stage of the task, consisting of rope pulling, reward retrieval, and reward area leaving. Activity was highest early in sessions, commensurate with the highest relative task utility, then decreased until the point of quitting. Neural ensembles showed stable task representations across the entirety of sessions. However, these representations drifted and became more distinct over the course of the session. These results suggest that dmPFC neurons represent behavioral states that are dynamically modified as behaviors lose their utility, culminating in task quitting.

Introduction

When engaged in a certain behavior, animals must continuously assess whether to persist in that behavior or change course. During foraging, for example, animals face recurring decisions as to whether to stay in their current patch of diminishing food or to incur the costs of exploration to search for a more abundant patch (Stephens and Krebs 1986; Hayden 2018). A third option also exists: rather than adjusting the current course of action, an animal could quit the behavior all together. The decision to quit may be driven by a multitude of factors including physiological states (e.g., satiation, fatigue), the cognitive assessment of diminishing utility, and inter-individual trait variability. However, the neurophysiological processes that underlie quitting behaviors are not well understood.

The prefrontal cortex is likely a critical region in quitting processes due to its role in encoding action-outcome contingencies and using these associations to drive goal-directed behaviors (Kennerley and Walton 2011; Euston et al. 2012; Zhao et al. 2016). In particular, the anterior cingulate cortex (ACC or Anterior Cg1 in rodents) plays a specialized role in evaluating the physical effort costs associated with certain behaviors; this enables net utility determinations and behavioral motivation towards difficult but worthwhile goals (Walton et al. 2003; Kennerley et al. 2006; Hosokawa et al. 2013; Klein-Flugge et al. 2016). Individual Cg1 neurons are known to respond to physical effort costs and reward amounts associated with different actions and their outcomes (Kennerley et al. 2009; Hayden and Platt, 2010; Cowen et al. 2012; Hillman and Bilkey, 2012; Hosokawa et al. 2013; Porter et al. 2019). As a whole, the ACC is thought to form a cognitive map of the “task space” that is used to drive decision-making towards actions that carry high relative utility (Kurniawan et al. 2013; Haber and Behrens, 2014; Heilbronner and Hayden, 2016; Azab and Hayden, 2017; Mashhoori et al. 2018; Hart et al. 2020).

The ACC is well connected to sensorimotor areas which may translate decisions into actions (Reep et al. 1987; Gu et al. 1999; Lapish et al. 2008; Rolls, 2019). The secondary motor area (M2) is laterally adjacent to Cg1 in rodents and is considered a higher order motor area that participates in action plan selection based on utility (Sul et al. 2011; Murakami et al. 2014; Klein-Flugge et al. 2016; Barthas and Kwan, 2017; Ebbesen et al. 2018). Together, these two regions in the rat dorsomedial prefrontal cortex (dmPFC) – the anterior Cg1 and M2 – demonstrate important roles in determining the utility and selection of behaviors that maximize benefit to the animal. However, it is not well-understood how neurons in these areas respond to progressive performance in a prolonged task, where quitting is ultimately the most beneficial action.

In order to investigate quitting behaviors we recently developed an effortful weight lifting task (WLT) for use with laboratory rats (Porter and Hillman, 2019). To date, most neurophysiological studies on cost-benefit decision making have not been structured to allow a subject to quit outright. Generally, subjects are provided with different cost-benefit actions (e.g., operant actions involving variable effort in rodent studies, Salamone et al. 1994; Cowen et al. 2012; Hillman and Bilkey, 2010) or cost-benefit contingencies (e.g., stimulus-outcome pictorials in primate studies, Kennerley and Wallis, 2009) to choose between, and subjects make recursive decisions over a fixed number of trials or a fixed duration. In our WLT, rats are challenged to pull a weight-adjustable rope in order to earn rewards. Weight-adjustment allows for dynamic control over the effort associated with the rope-pulling behavior. Furthermore, the animals were free to quit the task at any point.

Here we recorded single units from the dmPFC while rats carried out the WLT in order to characterize how neural activity patterns relate to quitting behaviors. Two different paradigms of the WLT were used: a progressive weight (PW) version where the weight increased over time, and a fixed weight (FW) version that used a constant, moderate weight.

We predicted that dmPFC neurons would be responsive to our effort-based task in that different neurons should represent the actions, outcomes, and salient behavior transitions associated with performing the task. Furthermore, these representations should be modulated by changes in utility, brought about by heavier weight-adjustments and/or fatigue. We predicted that Cg1 neurons would be more responsive during reward phases to monitor action-outcome contingencies while M2 neurons would be more responsive during rope pulling and the associated effort levels. Lastly, we predicted that neural populations in the dmPFC would undergo state transitions that reflect the rats' behavioral transitions towards the point of quitting.

Materials and Methods

Subjects

Male Sprague-Dawley rats (n=6, weighing between 400 - 480 g at the time of experiment; Hercus Taieri Resource Unit, New Zealand) were single housed in individually ventilated cages measuring 38 x 30 x 35 cm (Tecniplast, Italy). Animals were maintained on a 12 h reverse light cycle (07:00 to 19:00), with experimentation occurring during the animals' dark, active phase. Water was available *ad libitum* throughout training and experimentation. Daily chow (Teklad diet; Envigo, USA) was rationed to promote food motivation for training and during experimentation. Rats were weighed twice per week and daily food rations were adjusted to maintain each animal's body weight ≥ 85% of their free-feeding body weight. All procedures were approved by the University of Otago's Animal Ethics Committee, protocol 91/17.

Training

Prior to surgery, animals were trained in the WLT as previously described (Porter and Hillman, 2019). Briefly, the WLT is performed in a 120 x 90 x 60 cm arena (Figure 1a). An animal is progressively shaped to pull a rope 30 cm to trigger the dispensing of 0.25 mL of 20% sucrose solution. The rope can be weighted from 0 to 225 g to increase the effort demands of the task, however only 0 g and 45 g are used in the training phase. Once an animal can perform 20 successful pulls (10 pulls on 0 g followed immediately by 10 pulls on 45 g) in under 10 minutes, they are considered trained in the WLT and ready for surgery. Rats took an average of 26.3 days (Standard Deviation (S.D.) = 0.7 days) to reach the 20 successful trial criterion.

Microdrives

All rats were implanted with Halo-10 Microdrives (Neuralynx, USA) consisting of eight independently drivable tetrodes. Tetrodes were made from 18 micron Nichrome Formvar insulated wire (A-M Systems, Inc., USA). Tetrodes were gold plated over three to five days using a nanoZ (White Matter LLC, USA) such that electrode impedances were 200-300 kOhms prior to surgery. Tetrodes were secured to Neuralynx EIB-36's via gold pins to interface with our headstage (see *Data collection*, below). Neuralynx Standard A exit tips were used and tetrodes were positioned in the exit tip to minimize medial-lateral spread when implanted in the brain.

Surgery

Anesthesia was induced with 5% Isoflurane in oxygen and then maintained at 2-3% throughout the surgery. Rats were mounted in a Kopf (David Kopf Instruments, USA) stereotaxic frame with non-puncture ear bars. One mL of 2% Lidocaine diluted in 4 mL of 0.9% saline was injected subcutaneously above the skull. A midline incision was made and the skin over the skull retracted. A 3 mm diameter craniotomy was made centered at + 2.5 mm anterior-posterior and ± 0.7 mm medial-lateral from Bregma (Paxinos and Watson, 2007). Five rats were implanted on the right hemisphere and one on the left hemisphere. The dura was removed and the microdrive was lowered onto the surface of the brain via a stereotaxic arm such that the most medial tetrode was next to the central venous sinus. Two 4 mm stainless steel screws, bilaterally placed over the cerebellum, served as grounds. Six to eight additional 3 mm stainless steel screws were then put partially into the skull and the implant was secured to the skull and skull screws with ultraviolet cure self-etching dental cement (Maxcem Elite™, Kerr Corp., USA). Sutures were used to close the skin around the implants.

Post-operatively (post-op), rats were subcutaneously administered prophylactic antibiotic (30mg/kg of Amphoprim) and an anti-inflammatory (5 mg/kg of Carprofen) and recovered in a heated cage until they regained consciousness and were active. Rats were then returned to a clean home cage to recover for 10 days. The anti-inflammatory was administered for a minimum of two additional days post-op and as needed if rats showed any signs of pain. After the 10 days of post-op recovery, rats were re-introduced to the WLT and tested to criterion to ensure that they could still carry out the task and habituate to their implants. Rats took an average of 9.2 (S.D. = 4.9) days to return to criterion. Once rats reestablished criterion performance, experimentation started.

Data collection

Single unit data was collected using a Neuralynx Digital Lynx SX system with Cheetah software (version 6.3.2) via a Neuralynx HS-36-LED pre-amplifier headstage connected to a Neuralynx Saturn commutator. Data from the 32 electrodes were sampled at 32 kHz and bandpass filtered between 600 and 6,000 Hz. Every tetrode was referenced to an electrode on a different tetrode with no spiking activity. Single unit action potentials were detected online with a threshold of 65 µV and stored for offline spike sorting. Spikes were sorted offline using Plexon (USA) Offline Sorter (version 4) by primarily relying on peak-valley amplitude and principal component analysis. Tetrodes were not moved day-to-day if they had single units on them. If a tetrode did not have any single units it was advanced approximately 32 microns at the end of a recording session.

Each headstage contained two LEDs which enabled the animal's position to be tracked by Neuralynx's Cheetah software via an overhead camera (JAI Inc., USA). Video was sampled at 25 frames per second. An Arduino Uno (www.arduino.cc) was used to detect rope pulls and to dispense sucrose solution reward; TTL signals of all events were sent to the Digital Lynx SX system for timestamping and storage.

Progressive weight paradigm

Rats completed two WLT paradigms: a progressive weight (PW) paradigm and a fixed weight (FW) paradigm (Porter and Hillman, 2019). In the PW paradigm, a session started with a two minute pre-baseline period where the rat was placed in the WLT arena but the rope was not extended into the environment. After the pre-baseline, the rope was extended with 0 g of weight attached and the rat could begin earning rewards; each successful 30 cm pull dispensed 0.2 mL of 20% sucrose solution. After every 10 successful (rewarded) pulls of the rope, 45 g of weight was added to the rope until the rat quit or 225 g was reached. If a rat completed 10 trials of 225 g, a clip was placed on the rope in order to prevent the rat from pulling the weight high enough to trigger reward dispensing. However, no rats in this study completed 10 trials of 225 g.

Quitting was defined as the rat making no attempts to pull the rope for two full minutes. When a rat quit the rope was retracted from the apparatus at the end of the two minute 'quit period', and a two minute post-baseline period started. Lastly, after the post-baseline period, two sucrose rewards were manually dispensed to check if the rats were satiated or not. Rats had one minute to consume the sucrose reward once dispensed. Every rat on every session consumed these manually dispensed rewards, suggesting that satiation was not occurring. After the sucrose satiation check the recording session was stopped and the animal was returned to its home cage. Rats ran one session per day.

Fixed weight paradigm

In the FW paradigm, a session again started with a two minute pre-baseline period, followed by the rope being extended into the arena with 0 g of weight attached. After 10 successful (rewarded) pulls on 0 g, the rope weight was immediately switched to a fixed weight, calibrated to each specific rat. The initial fixed weight was the maximum weight that the rat had completed in the PW paradigm (i.e., the weight on which they had successfully

performed 10 rewarded pulls). Rats were allowed to complete as many trials as they wanted on the FW until they quit or until 60 minutes elapsed, whichever came first. After 2-3 days (average 2.7 days, S.D. = 0.4 days), the weight was adjusted to be 45 g heavier or lighter depending on the rats' performance: if rats were working for fewer than 15 minutes their weight was reduced whereas if rats were working for more than 15 minutes their weight was increased. Rats then carried out the adjusted FW paradigm for 2-5 further days (average 3.2 days, S.D. = 1.0 days) depending on their performance and the number of single units being recorded.

Overall analyses

All analyses were carried out using Matlab (version R2018a; The MathWorks Inc., USA). Colorbrewer (https://github.com/axismaps/colorbrewer/) and mpl colormaps (https://bids.github.io/colormap/) were used for graphical color maps.

Behavioral analyses

Behavioral metrics were calculated as previously reported (Porter and Hillman, 2019). Briefly, for the PW paradigm, the quit weight was calculated as the highest weight in a given session on which the rat made at least one rope pulling attempt. For the FW paradigm, the quit duration was calculated as the time from when the FW was attached (exclusive of 0 g) until the last pull attempt.

For both the PW and FW paradigms, quartiles were determined based on the duration of each individual session. For the PW paradigm, the duration was calculated starting from when the 0 g weight was available until the rat quit. For the FW paradigm, quartiles were calculated starting from the time the FW was available (thus excluding the initial period of 10 trials on 0 g). For all FW paradigm analyses this initial 0 g period was excluded so as to avoid comparisons between 0 g-containing quartiles and FW-containing quartiles Thus, for behavioral and neural analyses of the FW paradigm, every quartile has the same weight

attached to the rope. Every quartile had to contain at least five completed trials; recording sessions that did not meet this minimum trial criterion were excluded from this manuscript. The four quartiles are abbreviated throughout the manuscript as Q1, Q2, Q3, and Q4 to denote the first, second, third, and fourth quartile.

Trials were defined as the sequence of events beginning with a rope pull attempt followed by reward area entry and concluding with reward area leaving. This provided three task phases: the attempt phase, the reward entry phase, and the reward leaving phase (Figure 1b). The attempt phase consisted of 1.5 seconds before and 2.0 seconds after a pull attempt. Pull attempts were timestamped by a TTL code from the Arduino that logged when the weight was lifted. Reward entry and reward leaving were based on the tracking data obtained from LEDs on the headstage. A 20 x 20 cm area was manually placed around the reward cap for each recording session based on the rat's tracking data and known cap location. Reward entry was centered on when the rat entered the reward zone and their instantaneous speed dropped below 15 cm/s. The reward entry phase consisted of 1.5 seconds before and 1.0 second after reward zone entry. Reward leaving consisted of 1.0 second before and 1.5 seconds after the point where the rat left the reward zone. Task-phase durations were chosen based on the distribution of durations across all recording sessions in order to maximize duration but minimize overlap between phases.

Trials were deemed "successful" when the rat pulled the weight sufficiently high enough (30 cm) to trigger a reward, or deemed "failures" when the rat did not pull the weight high enough but still went and sampled the reward cap. Events when the rat pulled the rope but did not check the reward cap were excluded. Trials were also parsed according to whether or not the rat "engaged" or "disengaged" from the task on the subsequent trial. Engagement was determined based on how long it took from the rat leaving the reward cap until the next pull attempt. If the rat made another attempt within 12 seconds the trial was considered an

engage trial; anything longer than 12 seconds was considered a disengage trial. Twelve seconds was chosen as a cut off as 75% of all next-attempts occurred within 12 seconds of leaving the reward cap.

The on-task ratio was calculated as the total time a rat spent working at the task divided by the total time of the session when the rope was available. Rats were considered working if they were both within the "work zone" and making attempts to pull the rope. The work zone consisted of a 60 x 64 cm area encompassing the rope pulling area and the reward cap. The reward interval was calculated as the time between earned rewards based on the reward TTL signals. The success-to-attempt ratio was calculated as the number of successful pull attempts divided by the total number of attempts made. For the on-task ratio and success-to-attempt ratio we used a repeated-measures one-way ANOVA to test for significant changes over the quartiles and Tukey-Kramer's test for post-hoc pairwise comparisons. Due to some rats having no rewards in Q4, and thus no reward intervals, we instead used a Kruskal-Wallis test to measure if reward interval times were affected by quartile and Tukey-Kramer's test for post-hoc comparisons.

Neural analyses – single neurons

For every recording session, sorted spikes from Plexon Offline sorter were imported into Matlab with the accompanying tracking and TTL data from that recording session. Neurons were assigned a brain area based on what tetrode they were recorded from. We tested whether or not neurons were task-responsive by randomly sampling every neuron's firing rate across 10,000 two second epochs drawn from the entire recording session (including the pre- and post-baselines). Neurons were considered to be task-responsive if, during at least one task-phase, their firing rate was significantly different (Wilcoxon rank sum test) to the random sample and they had a firing rate of at least 0.5 Hz. Furthermore, to avoid including neurons that may have been lost or gained over the course of a recording session, a

neuron also had to have a non-zero firing rate in at least one task-phase for all four quartiles. Neurons that did not meet this criteria were not included for analysis. Spike raster plots and peri-event time histograms were created for task-responsive neurons for each of the three task-phases across all trials. Spike times were aligned to the center of the task-phase event. Histograms were created with 100 ms bins across each task-phase. For each task-phase, a Chi-squared test against an equal distribution was used to test if there was an unequal distribution of task-responsive neurons across brain areas.

Neural analyses – population analyses

All task-responsive neurons from all sessions were pooled together for further analysis. A sequence plot was created with all neurons where each neuron's binned trial firing rates were z-scored then averaged across all trials in each 100 ms bin. All neurons were then sorted based on the trial bin that contained their maximum firing rate. For graphical purposes only, every neuron's activity was normalized to be between 0 and 1 where 1 corresponded to their maximum z-scored firing rate and 0 corresponded to their minimum z-scored firing rate. All analysis was done on the z-scored firing rates. A cross-correlation matrix was created using the sequence plot where the Pearson's correlation was calculated pairwise for every 100 ms bin of a trial.

Principal component analysis (PCA) was carried out on the z-score averaged firing rates of all task-responsive neuron in 100 ms bins across the trial using Matlab's pca function. PCA was also carried out on the successful versus failed trials and engage versus disengage trials. In these cases, each task-responsive neuron's firing rates were binned into 100 ms bins, z-scored, and then all neurons were pooled together. We computed the distance between the PCA score trajectories by calculating the distance between each 100 ms bin using the first three principal components. As a control, we also randomly split trials of a recording session into halves (for engage versus disengage comparisons, and success versus fail comparisons)

or quarters (for quartile analyses) and averaged a neuron's activity over the random trials. This was carried out 1,000 times per session. For every iteration, all neurons were pooled together and the distance between each 100 ms bin calculated. For the comparisons of engage versus disengage and success versus fail, a paired t-test was used to determine if the observed PCA distances were significantly different from the mean of the 1,000 random PCA score distances. A one-way ANOVA, with Tukey-Kramer's post-hoc, was used to test whether or not the observed PCA score distances across the three task-phases were significantly different.

For graphical purposes only, the PCA score plots were smoothed with a 500 ms moving average within each task-phase when comparing multiple trajectories. The PCA trajectories of all trials from the PW and FW paradigms (Figures 4c and 7c) are plotted with the raw scores. All distance calculations and statistics were done on the raw PCA scores.

Kruskal-Wallis tests were used on each task-responsive neuron to test firing rate variability across quartiles for each task-phase. To keep trial numbers consistent across sessions, we limited analysis to five trials in each quartile. For Q1, the first five trials were used. For Q2 and Q3, the middle five trials were used. For Q4, the last five trials were used. We then pooled the quartile averaged z-scored firing rates of all neurons with a significant ($p < 0.05$) effect for quartile for each task-phase. For each task-phase we determined what quartile each neuron was most active in based on their averaged fire rates in each quartile. We carried out a Chi-squared test on the observed distribution of peak activity across the four quartiles versus and equal distribution across all quartiles. Repeated-measures one-way ANOVA with a post-hoc Tukey-Kramer's test was used on the populations of neurons for each task-phase, with quartile as a factor.

Neuronal data was also broken up into two minute epochs across the entire recording session, including the pre- and post-baseline epochs where the rope was not available to the

rat. Four more epochs consisted of the four quartiles. For quartile one, the first two minutes of the 0 g (PW paradigm) or the fixed weight (FW paradigm) was used. For quartiles two and three, the middle two minutes of the quartile were used. For quartile four, the two minutes preceding the rats' final attempt were used. The quit epoch consisted of the two minutes proceeding the rats' final attempt where the rope was still available but the rat was not making attempts. For each task-responsive neuron a z-scored firing rate was calculated for each two minute epoch and all neurons were pooled. Pairwise cross-correlations were calculated using Pearson's correlation. Matlab's pdist function was used to calculate the pairwise correlation distances between epochs. Matlab's linkage function was then used to hierarchically cluster the correlation distances based on their average correlations. Matlab's dendrogram function was used to plot the clustering.

Histology

After completion of all behavioral paradigms, rats were deeply anesthetized with Isoflurane and transcardially perfused with 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). All tetrodes were wound up before the brain was removed to avoid damage. Brains were then cryoprotected in 30% sucrose in 4% PFA PBS until they sunk. Brains were sectioned at 40 µm using a Cryostat (CM1950, Lecia Biosystems, LLC., Germany). Brain sections were then mounted, stained with Thionine, and imaged with a light microscope (Lecia Biosystems, LLC.). Tetrodes were mapped based on their tracks in reference to Paxinos and Watson (2007) brain atlas, microdrive turn logs, and their relative locations in the microdrive (Figure 1c).

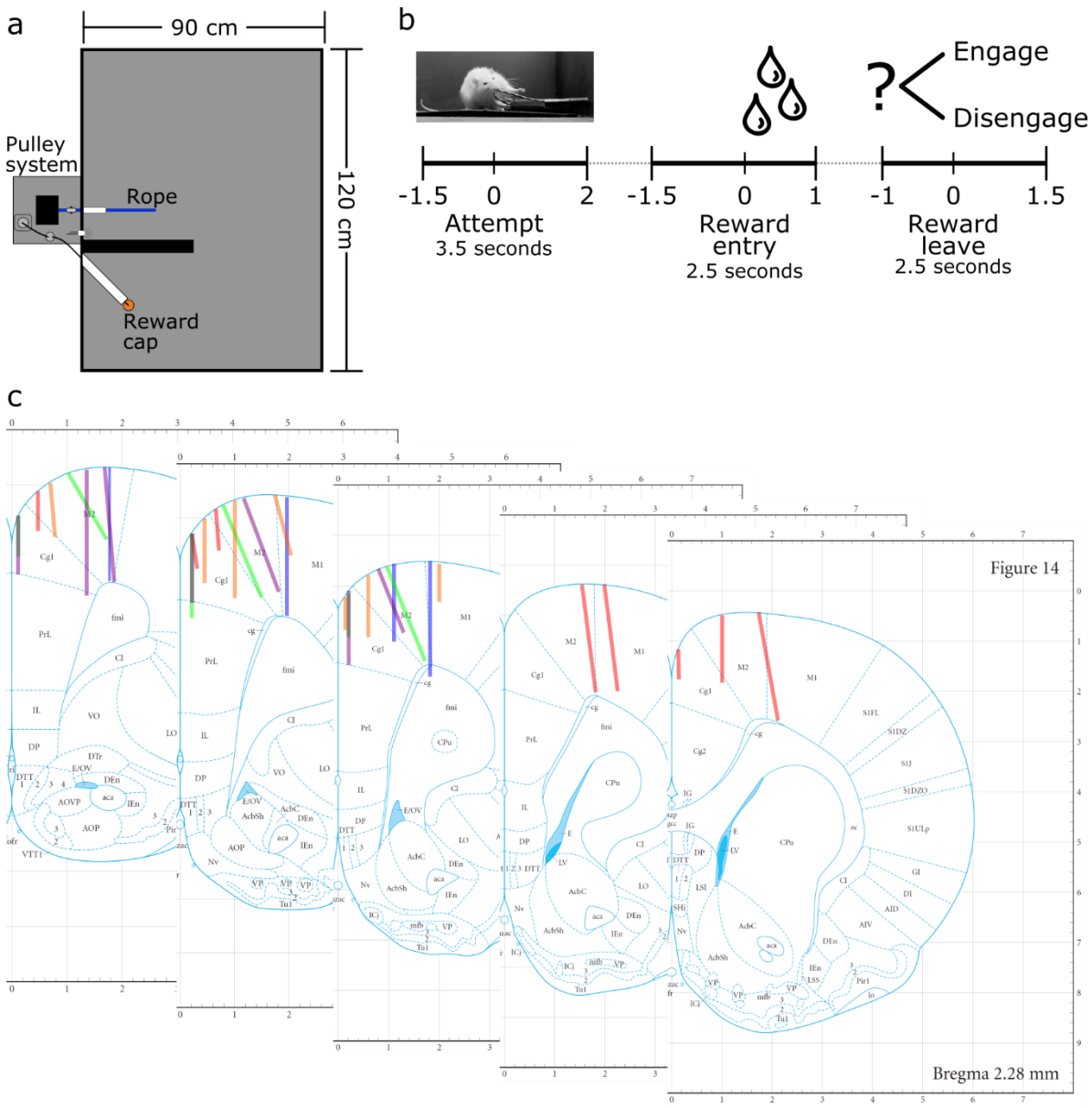

Figure 1: Task overview and histology. a) Weight lifting task schematic as reported in Porter and Hillman (2020). b) Trial breakdown illustrating the three phases of interest. In the Attempt phase, rats are physically pulling the rope. In Reward entry, rats enter the reward cap area and consume sucrose reward if present. In the Reward leave, rats leave the reward cap area and choose to re-engage in a rope pull or disengage. c) Histological reconstruction of tetrode positions. Each color represents the tetrodes from one rat. Tetrodes spanned +3.72 mm to +2.28 mm relative to Bregma. Schematics modified from Paxinos and Watson (2007).

Results

Progressive weight behavior

Rats first carried out the PW paradigm; here the rope weight started out at 0 g and was increased by 45 g every 10 successful pulls, to a maximum weight of 225 g. Rats were free to quit the task at any point. Six rats performed a total of 30 sessions of the PW task (3-7 sessions per rat; average 5.0 ± 1.7 S.D.). Consistent with our previous studies (Porter and Hillman, 2019), rats were more likely to quit on heavier weights (Figure 2a). No rat was able to complete 10 successful pulls of 225 g. Due to the inter- and intra-individual variance of quit weights across sessions, each session varied in time duration and number of trials. Sessions varied in length from 874 to 2,056 seconds (median = 1,589 ± 373). Across all sessions rats earned from 25 to 55 rewards (median = 47 ± 8). Because of this variability, we broke each session into quartiles based on time and used those quartiles to analyze behavioral performance.

We computed a time-on-task ratio that reflected the proportion of time rats spent engaged in the task. A ratio of 1 indicated that the rats were regularly making attempts to pull the rope while a ratio of 0 reflected rats making no attempts. Across quartiles, rats spent significantly less time on-task as the session progressed into the third and fourth quartiles ($F(2.6, 76.4) = 177.0$, $p < 0.0001$; Figure 2b). When rats were on-task and working, it took longer to earn rewards as the session progressed ($H(3, 107) = 56.0$, $p < 0.0001$; Figure 2c). Increases in reward interval, notably in Q4, were likely in part due to an increasing number of failed rope pulls as weights got heavier. Across quartiles, there was a significant decrease in the success-to-attempt ratio ($F(2.4, 68.4) = 88.1$, $p < 0.0001$; Figure 2d).

Fixed weight behavior

On the FW paradigm, the rope weight started out at 0 g and after 10 successful pulls the rope weight was set to a moderate FW, calibrated to each rat (see Methods). Rats could

perform as many weighted rope pulls as desired, up to a maximum duration of 60 minutes. Rats were free to quit the task at any point. Because we set a minimum trial number criterion for data analysis inclusion (see Methods), only nine sessions of the FW task were available for analysis (three rats performed these nine sessions, 2-4 sessions per rat; average 3.0 ± 1). Quit durations ranged from 1,803 to 3,548 seconds (median = 2,711 ± 571 seconds Figure 2e) and rats earned from 24 to 106 rewards (median = 81 ± 27 rewards) on the FW.

Similar to the PW paradigm, rats spent significantly less time on-task as the FW sessions progressed ($F(1.7, 13.3) = 47.5$, $p < 0.0001$; all pairwise comparisons to Q4, p's < 0.0001; Figure 2f). There was a main effect for quartile on reward interval ($H(3, 31) = 9.8$, $p = 0.02$; Figure 2g), with Q4 having a higher reward interval than Q3 ($p = 0.030$) and Q2 ($p = 0.035$). There was a significant main effect of quartile for the success-to-attempt ratio ($F(2.1, 17.0) = 5.9$, $p = 0.01$; Figure 2h). However, post-hoc pairwise analyses only showed modest differences between quartiles: success-to-attempt ratio was significantly higher on Q2 compared to Q3 ($p = 0.041$) and Q4 ($p = 0.043$). Compared to behavior in the PW task, where weights incrementally increased by 45 g every 10 trials, rats may have struggled with the FW initially in Q1 where the weight suddenly increased from 0 g to the FW value greater than 135 g. This likely accounted for the lower success rate and greater reward interval observed in Q1 in the FW task.

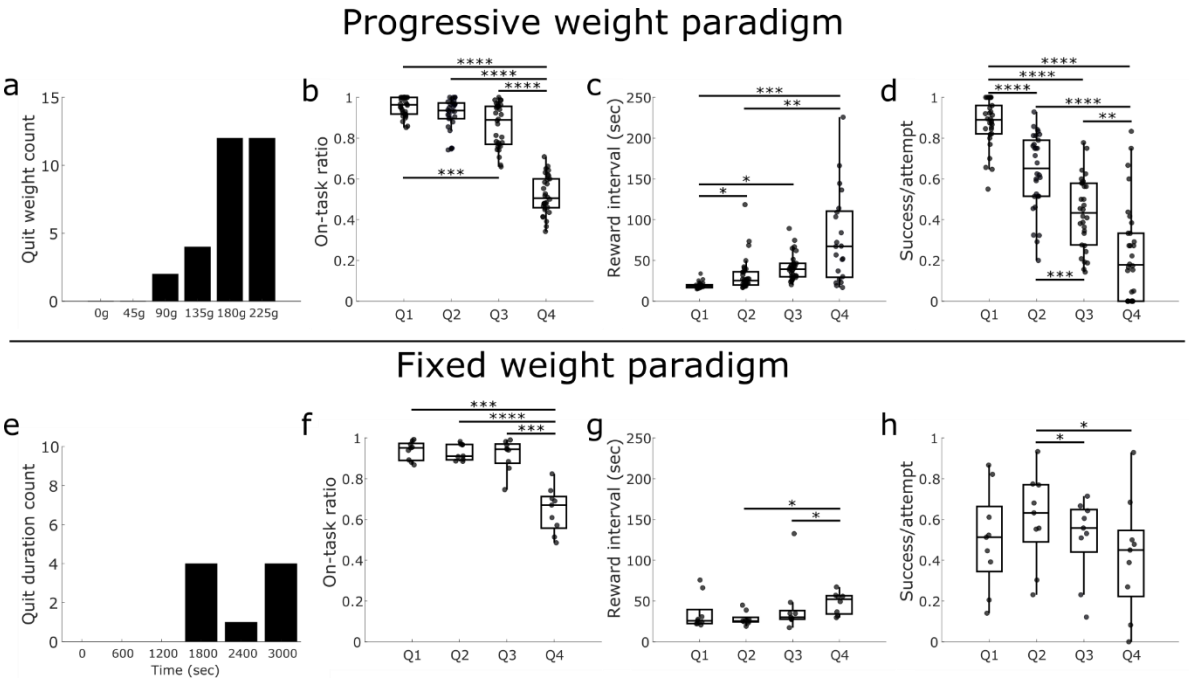

Figure 2: Behavioral measures on the PW and FW paradigms. a-d) Behavioral metrics for the PW task. e-h) Behavioral metrics for the FW task. For both tasks, each recording session was broken up into four equal duration quartiles, Q1-Q4. The time on-task ratio was calculated as the time rats spent working (making attempts) divided by the total time in each quartile. The success-to-attempt ratio was calculated as the number of successful attempts divided by the total number of attempts for each quartile. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Neuronal task selectivity – Progressive weight paradigm

For neuronal analysis we focused on three task-phases of interest: the attempt phase, the reward entry phase, and the reward leaving phase (see Methods and Figure 1b). In the PW paradigm, we recorded a total of 360 neurons across the 30 recording sessions. Based on histological reconstruction (see Figure 1c), 123 of these neurons were from the Cg1, 93 were between Cg1-M2, and 144 were from Motor cortices 2/1 (M2/1).

A variety of activity patterns were observed within and across task-phases as well as over time (Figure 3a-h). Some neurons were especially active during the attempt phase (3a,

b). The Cg1 neuron shown in Figure 3a appeared to have increased its firing rate prior to the attempt and then fired rhythmically at approximately 2 Hz. The motor neuron in Figure 3b was highly selective to the attempt phase with high firing right before the attempt followed by a rhythmic firing pattern during rope pulling. Other neurons had phasic firing to very specific time points of a trial. For example, the Cg1-M2 neuron shown in Figure 3c fired just prior to reward entry while the M2/1 neuron in Figure 3d fired as the rat left the reward zone. Furthermore, some neurons showed multiple phasic peaks such as the motor neuron in Figure 3e that was active both right before reward entry and at reward leaving. More complex activity profiles were also observed such as the Cg1 neuron shown in Figure 3f that fired tonically during the all three phases but appeared to be briefly inhibited prior to reward entry and after reward leaving. Some neurons showed a clear change in firing rate over the course of the recoding. For example, the neuron in Figure 3g was highly active before reward entry on early trials (top lines of firing rate raster plot) but decreased its firing rate at reward entry as the weight became heavier over the session (moving vertically down the raster plot). The motor neuron shown in Figure 3h doesn’t have a clear preference for a trial-phase but does increase in firing over time.

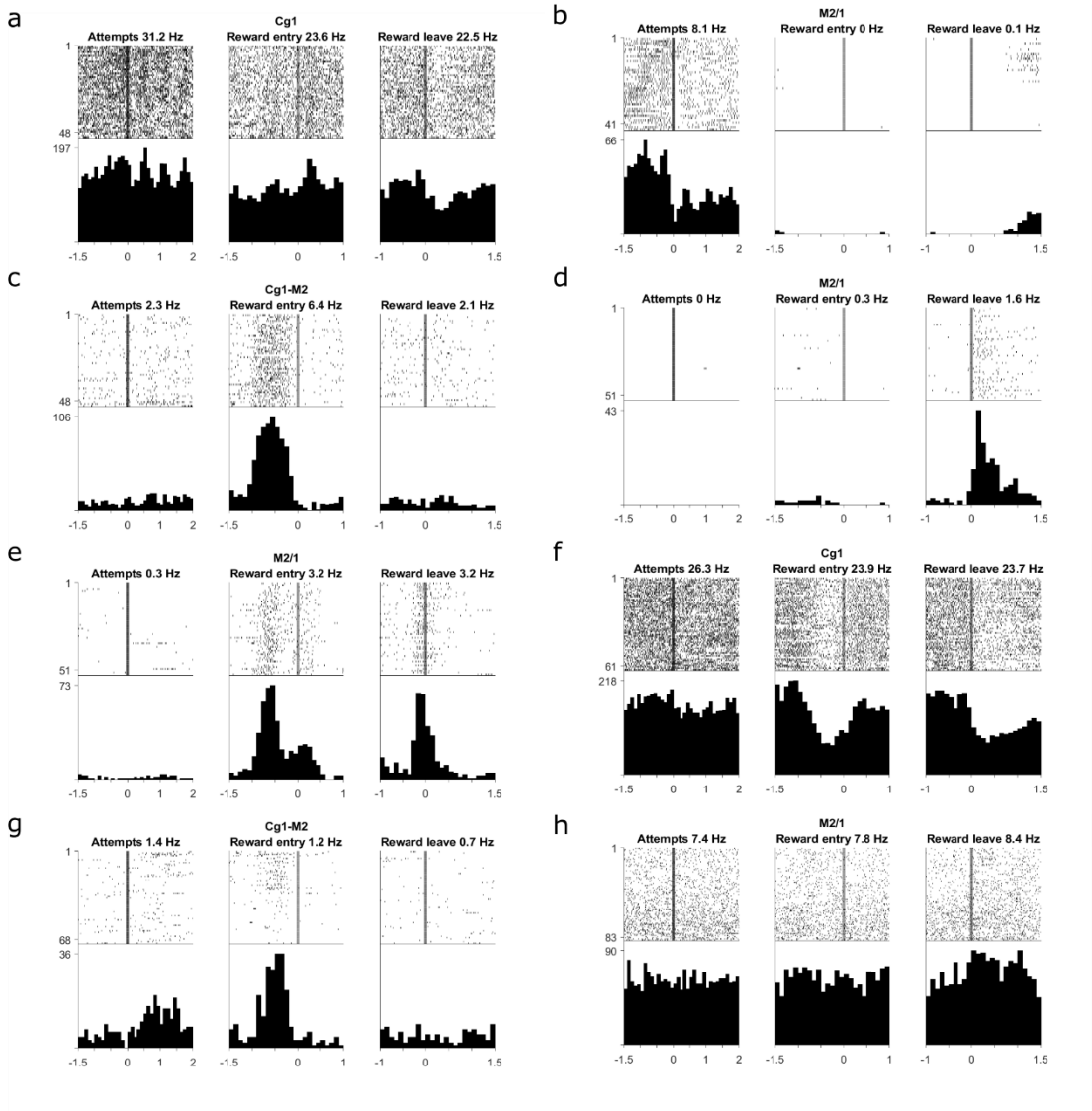

Figure 3: Examples of single unit activity across task-phases during PW sessions. a-h) Each panel shows the activity of one neuron across the three task-phases with a raster plot and peri-event time histogram. Neuron location (Cg1, Cg1-M2, M2/1) is indicated in the title line for each panel. For the raster plots, each row is a trial where the top-most row corresponds to the first trial and descends chronologically. Vertical lines indicate the time of event: attempt, reward entry, reward leaving. Histograms were calculated with 100 ms bins. The y-tick number indicates the highest number of spikes across all three phases. The y-axis is consistent across all three phases.

There were a total of 257 (71%) unique task-responsive neurons; 162 of these neurons were responsive to the attempt phase, 155 to reward entry, and 148 to reward leave. Ninety-seven neurons were responsive in only one epoch while 112 were responsive in two epochs and 48 were responsive in all three. There was no significant difference in the proportion of responsive neurons across the three brain locations (X^2 all p's > 0.56). Thus, unless otherwise stated, all neurons responsive to at least one task-phase were pooled together across brain locations for further analysis.

The normalized, trial averaged activity from these task-responsive neurons is shown in Figure 4a. Trials were defined as the sequence of events beginning at a rope pull attempt followed by reward zone entry and concluding with reward zone leaving (see Methods). Neural activity during the attempt phase was highly correlated across task-responsive neurons which may reflect tonic activity patterns during the rope pulling action (Figure 4b). In contrast, neural activity during reward entry and reward leaving was more phasic with lower correlation values from early compared to later time points within these phases. Activity in the attempt phase was highly dissimilar compared to the two reward phases. To further visualize neural activity across trials, we carried out PCA on the trial averaged activity of the task-responsive neurons (Figure 4c). Supporting the cross-correlation analysis, principal component 1 (PC1; 31% of variance) differentiated trial-bins by attempt phase versus the two reward phases (reward entry and reward leave). PC2 and PC3 explained a further 16% and 10% of the observed variance, respectively. Trial-bins in the attempt phase clustered together before transitioning to the reward phases. From the start of reward entry, the trajectory traces out a path from reward entry to the end of reward leaving; there is some clustering around the transition point from reward entry to reward leaving that may reflect reward consumption (see Figure 4c).

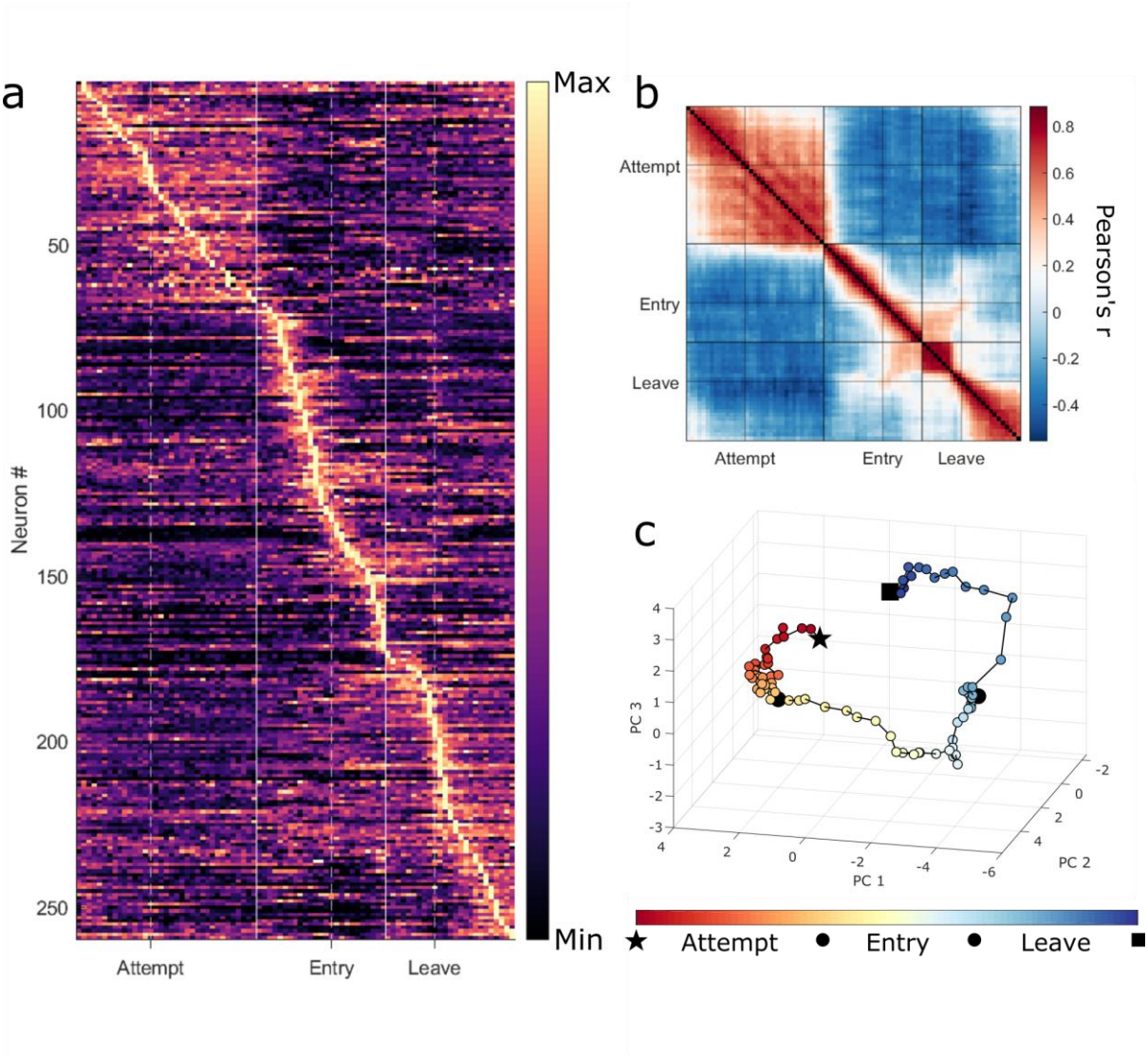

Figure 4: Population activity of task-responsive neurons during the PW paradigm. a) Trial averaged activity of all 257 task-responsive neurons, sorted by trial-bin in which neurons had their highest activity. Each trial was segmented into 100ms time bins. For graphical purposes, each neuron's trial averaged activity was scaled between 0 and 1. b) Cross-correlation matrix of the average, normalized activity of all active neurons across trial phases. c) PCA scores from the average neural activity across trials in the first three principal components. Every point represents the top three PC scores from one 100ms bin. Red represents the start of the trial and progresses to blue at the end of the trial. The first time bin is highlighted with a star and last bin with a square. Black circles represent trial-phase transitions from Attempt to Reward Entry, and Reward Entry to Reward Leave.

To test whether neural activity reflected differences in future behavior, we parsed trials based on whether or not the rat went on to engage in another rope pull attempt after reward consumption. A trial was deemed an “engage” trial if – after the reward leave phase – the rat went on to perform another rope pull within 12 seconds. Twelve seconds was chosen as a cut-off since, on approximately 75% of all trials, rats made another attempt within 12 seconds (Figure 5a). The PCA trial-phase trajectories for engage versus disengage trials were highly similar (PC1, 23%; PC2, 12%; PC3, 7%; Figure 5b) and distances between bins tended to be within 2 S.D. of the mean distance between randomly sampled trajectories (Figure 5c). However, towards the end of the Reward leave phase, the engage and disengage trajectories started to diverge. There was a main effect for task-phase ($F(2, 82) = 20.3$, $p < 0.0001$). Post-hoc pairwise testing showed no significant difference between the distances of engage-disengage trajectories for attempt versus reward entry ($p = 0.99$) but reward leaving distances were significantly greater compared to both the attempt phase ($p < 0.0001$) and reward entry ($p < 0.0001$).

To test whether neural activity reflected differences in rope pulling outcome, we also parsed trials based on whether or not the trial was a “success” (i.e., the weight was lifted 30 cm to trigger reward dispensing) or a “failure” (i.e., rope pulling was initiated but the weight was not lifted to the full required height of 30 cm). For 75% of recording sessions the average success rate was above 50% (Figure 5d). The PCA trial phase trajectories of successful versus failed trials were distinct but had similar overall shapes (PC1, 15%; PC2, 8%; PC3, 6%; Figure 5e). The distances between successful and failed time bins in PC space were greater than 2 S.D. of randomly sampled trajectories (Figure 5f). No significant effect of task-phase was observed ($F(2, 82) = 1.02$, $p = 0.36$).

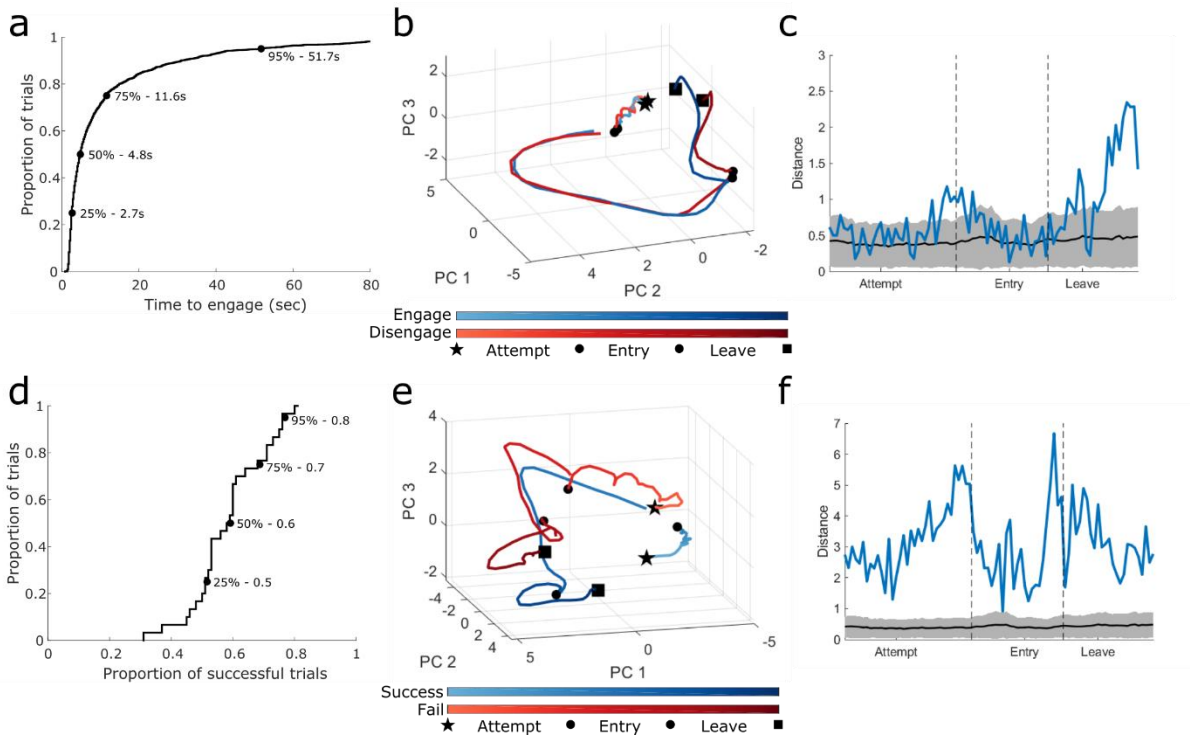

Figure 5: dmPFC neural activity changes during task engagement and failure in the PW paradigm. a) Cumulative distribution function of the time it took rats to engage in another trial following reward zone leaving. Quartile boundaries indicated by text. b) Neural trajectories in PC space during trials where rats subsequently engaged in a rope pulling trial (blue) or disengaged (red). c) Distances between engage vs disengage trials in PC space (blue line) for each 100ms time bin. Solid black line is the mean distance between random trajectories and grey area represents ± 2 S.D. d) Cumulative distribution function for the proportion of successful trials across sessions. Quartile boundaries indicated by text. e) Neural trajectories in PC space during trials where rats succeeded in lifting the weight 30 cm (blue) or failed (red). f) Distances between successful vs failed trials in PC space (blue line) for each 100ms time bin. Solid black line is the mean distance between random trajectories and grey area represents ± 2 S.D. b & e) Lighter color indicates earlier within the trial. The first time bin is highlighted with a star and last bin with a square. Black circles represent trial-phase transitions from Attempt to Entry and Entry to Leave.

To test whether neural activity in each task-phase was changing over time as rats carried out the PW paradigm, we parsed trials based on their temporal quartile of occurrence. Thirty-five percent (57/162) of attempt phase neurons, 34% (53/155) of reward entry neurons, and 30% (45/148) of reward leave neurons had a significant effect for quartile. Of the neurons with a quartile effect, we determined in which quartile they had the highest firing rate. For all three task-phases, the observed frequencies of peak activity across the quartiles were significantly different than an equal distribution across the quartiles (Attempts, $X^2(3) = 18.1$, $p < 0.0004$; Entry, $X^2(3) = 17.4$, $p = 0.0006$; Leaving, $X^2(3) = 20.5$, $p < 0.0002$; Figure 6a). The majority of neurons, across task-phases, had their highest firing rates in Q1. To test if neural activity was changing over the quartiles, we carried out one-way repeated measures ANOVAs on the population of task-active quartile-effect neurons in each task-phase (Figure 6b). Across the four quartiles, there was a significant effect for quartile for the attempt phase ($F(1.8, 101.7) = 9.5$, $p < 0.0003$), reward entry ($F(2.4, 126.3) = 3.6$, $p = 0.023$), and reward leave phase ($F(1.9, 85.2) = 9.5$, $p < 0.0003$). Specifically, pairwise post-hoc testing showed Q1 had significantly higher neural activity compared to Q3 and Q4 during attempts and reward leaving (all p's < 0.04), but not during reward entry (no pairwise comparisons were significant). For all three task-phases, we did not see a difference in the observed numbers of quartile selective neurons across brain regions (Chi-square test, all p's > 0.06; Figure 6c). These data may indicate that a subpopulation of dmPFC neurons have their peak activity early on in the session and then decrease firing over time.

To visualize any changes in neural activity over the quartiles, we carried out PCA on the trial averaged activity of all 257 task-responsive neurons across the quartiles (PC1, 16%; PC2, 9%; PC3, 7%). The PCA trial-phase trajectories were highly conserved across quartiles (Figure 6d). Similar to the PCA trial phase trajectory that was observed when all trials across a session were averaged together (see Figure 4c), each quartile's trajectory had activity

clustered tightly during attempts, an abrupt shifts to reward entry, then a smooth transition through reward entry to reward leaving (Figure 6d). PC3, however, appears to differentiate the trajectories by quartile in that distinct trajectories for each of the four quartiles can be seen (Figure 6e, f). Pairwise distances for every trial-bin in PC space across the four quartiles showed a significant effect for quartile pairs ($F(5, 504) = 500$, $p < 0.0001$; Figure 6f). The distances between Q1-Q2 and Q2-Q3 were equivalent ($p = 0.99$) whereas all other pairwise comparisons were significantly different from one another (all p's < 0.003).Q3-Q4 had the closest trajectories, whereas Q1-Q4 were the most distant. All distances from real trajectories were significantly greater than those from randomly generated quartiles (paired t-tests, all p's < 0.0001). Thus, across sessions the dmPFC may have a stable task representation that is modified over time.

To determine if neural activity changes could be seen more broadly across a recording session – as rats went from entering the apparatus, to starting the task, to quitting the task – we averaged the activity of the 257 task-active neurons in 2 minute epochs. These epochs consisted of the pre-baseline period, the four quartiles of task performance, the quit period where the rat could continue to do the task (rope available) but chose not to, and the final post-baseline period (see Methods). We predicted that neural activity would be the most similar when comparing the pre-baseline, quit period, and post-baseline epochs, as behaviorally the rat is not engaged in the task during any of these periods. Cross-correlation matrices revealed two distinct patterns which were contrary to our predictions: neural activity was highly similar across the pre-baseline, Q1, and Q2 epochs, and highly similar across the Q3, Q4, quit period, and post-baseline epochs (Figure 6g). Specifically, neural activity was highly correlated between pre-baseline and Q1 ($r = 0.50$), Q4 and the quit period ($r = 0.56$), as well as Q4 and the post-baseline period ($r = 0.66$). Clustering analyses confirmed these patterns showing two distinct clusters relating to the first half of the session and the latter half

of the session (Figure 6h). Thus, neural activity in the dmPFC appears to shift as the session progresses, irrespective of the specific behaviors being performed.

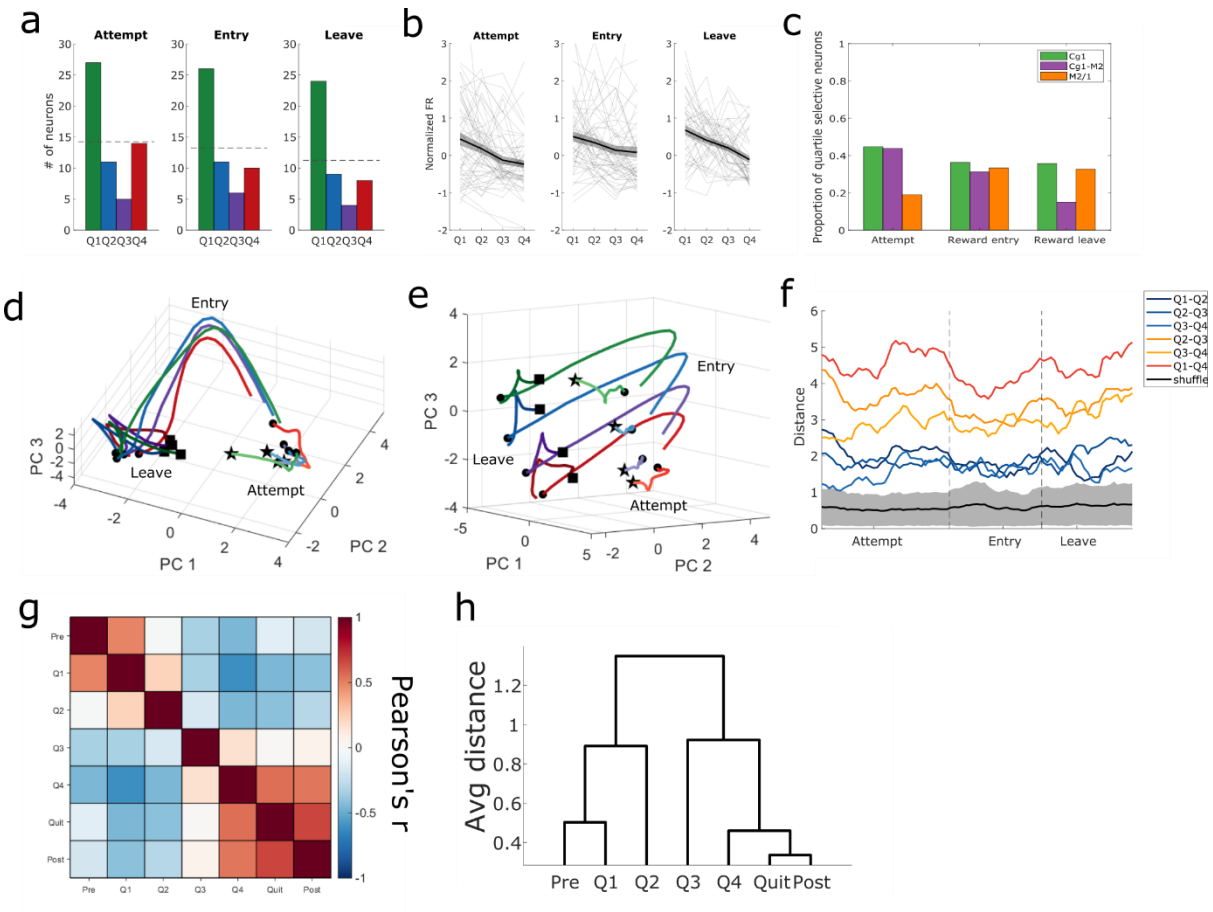

Figure 6: dmPFC neural activity across quartiles during the PW paradigm. a) Distributions of peak neural activity across the four quartiles for each task-phase. Dotted line indicates equal distribution across quartiles. b) Average population activity of task-active neurons with a quartile effect. Activity averaged over all neurons shown as solid black line with grey shading representing S.E.M. Individual neuron activity plotted as light grey lines. c) Proportion of task-responsive neurons with a significant quartile effect by brain region. d & e) PCA of trial activity across the four quartiles, viewed from two different angles. Q1, green; Q2, blue; Q3, purple; Q4, red. f) Pairwise distances between PC trajectories (colored) and average distance between shuffled quartiles (black) with ± 2 S.D. in grey. Cooler colors represent temporally closer pairs while warmer colors indicate more temporally distant pairs. g) Cross-correlation matrix of 2 min epoch activity across sessions. Pre, pre-baseline; post, post-baseline. h) Dendrogram of hierarchically clustered correlation distances.

Neuronal task selectivity – Fixed weight paradigm

The PW paradigm affords an animal continuously decreasing utility, due primarily to the increasing effort costs of the increasing weight. Rats must decide when to quit the task when the weight becomes too physically demanding. However, quitting in the PW task may be spurred by other primary factors such as fatigue or satiation. In order to try and understand the neural activity behind quitting behaviors more broadly, we analyzed dmPFC neural activity as rats carried out the FW paradigm, which lacks progressive increases in weight.

We recorded a total of 62 dmPFC neurons from the three rats that contributed the nine FW sessions shown in Figure 3 (2-4 sessions per rat). Of these 62 neurons, a total of 53 (86%) met our task-responsive criteria for further analysis; trial averaged activity is shown in Figure 7a. Thirty-eight of the task-responsive neurons were active in the attempt phase, 37 in the reward entry phase, and 30 in the reward leave phase. Thirteen of the task-responsive neurons were active in one phase, 28 in two phases, and 12 in all three. A cross-correlation matrix across all bins (Figure 7b) shows a similar pattern to that observed for the PW paradigm (see Figure 4b). Despite the relatively small number of neurons, PCA of the trial trajectory resulted in a similar low-dimensional trajectory (PC1, 37%; PC2, 16%; PC3, 10%; Figure 7c) to that seen for the PW trajectory (see Figure 4c).Activity clustered tightly during the attempt phase before transitioning through the reward entry phase with a relatively wide trajectory. Activity then clustered again from the reward entry to reward leave transition.

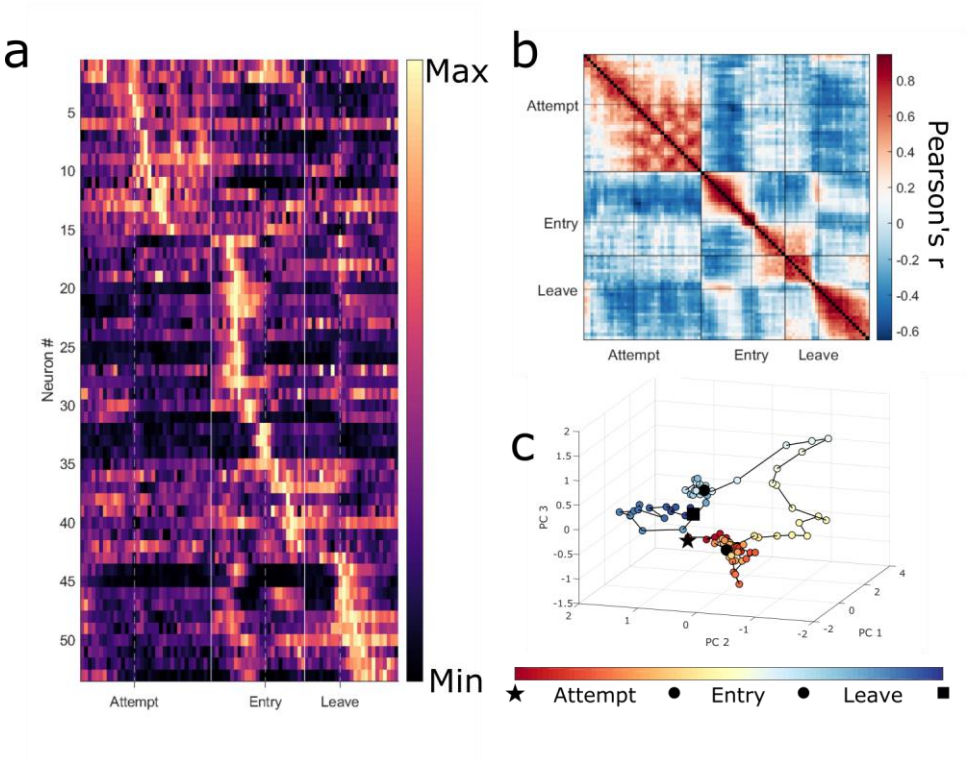

Figure 7: Population activity of task-responsive neurons during the FW paradigm. a) Trial averaged activity of all 53 task-responsive neurons, sorted by trial-bin in which neurons had their highest activity. Each trial was segmented into 100ms time bins. For graphical purposes, each neuron's trial averaged activity was scaled between 0 and 1. b) Cross-correlation matrix of the average, normalized activity of all active neurons across trial phases. c) PCA scores from the average neural activity across trials in the first three principal components. Every point represents the top three PC scores from one 100ms bin. Red represents the start of the trial and progresses to blue at the end of the trial. The first time bin is highlighted with a star and last bin with a square. Black circles represent trial-phase transitions from Attempt to Reward Entry, and Reward Entry to Reward Leave.

To test whether neural activity was changing over the course of the FW paradigm, we analyzed the activity of neurons across session quartiles. Kruskal-Wallis tests were used on each task-responsive neuron to test firing rate variability across quartiles for each task-phase. Of the 38 neurons active in the attempt phase, 15 (40%) showed a significant change in firing rate across the quartiles. Fourteen of the 36 (39%) reward entry and 11 of the 30 (37%) reward leave active neurons showed a significant effect for quartile. As was seen in the PW paradigm, the majority of task-active neurons with a significant effect of quartile had their highest firing rate in Q1 of the FW task (Attempts, $X^2(3) = 14.1$, $p = 0.003$; Entry, $X^2(3) = 17.4$, $p < 0.0006$; Leaving, $X^2(3) = 19.9$, $p < 0.00001$; Figure 8a). Across quartiles, the normalized firing rates from the population of task-active neurons showed a significant effect for quartile for all three task-phases (Attempts, $F(2.7, 40.5) = 17.1$ $p < 0.0001$; Entry, $F(2.4, 33.3) = 13.1$, $p < 0.0004$; Leaving, $F(2.3, 25.0) = 21.4$, $p < 0.00001$; Figure 8b). Post-hoc pairwise comparisons revealed that the averaged population activity was significantly higher during Q1 compared to Q4 (all p's < 0.0003) and during Q2 compared to Q4 (all p's < 0.012) in all three task-phases. Thus, similar to the PW paradigm, a subset of dmPFC neurons exhibit a decrease in activity across quartiles in the FW paradigm. Quartile selective neurons by brain region are shown in Figure 8c, however statistics are not provided as there were too few neurons to conduct Chi-Square tests.

PCA was applied to the averaged normalized activity of the 53 task-active neurons across the four quartiles during the FW task (PC1, 22%; PC2, 10%; PC3, 8%; Figure 8d, e). Generally, the same trajectory shape across the whole trial can be seen across all four quartiles (Figure 8d). However, as was seen in the PW task, the trial trajectories for each quartile can be distinguished (Figure 8e). The attempt phase remains relatively similar across quartiles, however, the wide neural trajectories during the reward entry phase show differing paths. During reward leaving, trajectories become smaller over the quartiles reflective of

more stable, stationary neural states. Pairwise distances were calculated between every quartile across trial bins (Figure 8f). As was seen in the PW, temporally distant trajectories tended to be farther apart than temporally proximal trajectories ($F(5, 504) = 113$, $p < 0.0001$). Differing from the PW paradigm, distances between Q2-Q3 and Q3-Q4 were equal ($p = 0.85$) in the FW paradigm, otherwise, all other pairwise comparisons were significantly different (all p's < 0.008). In order of mean distance, Q1-Q2 was more distant than Q2-Q3 and Q3-Q4. Unique to the FW, distances between trajectories tended to be farthest apart prior to the reward entry point and after reward leaving point. The average distances of every real trajectory were significantly greater than those from randomly sampled trajectories (paired t-tests, all p's < 0.0001).

Neural activity during the FW paradigm was also analyzed in broader two minute epochs across the session, which included the pre- and post-baseline periods (Figure 8g, h). Neural activity was highly correlated between pre-baseline and Q1 ($r = 0.60$), before becoming less correlated quartile to quartile. Neural activity across theQ4, Quit, and post-baseline periods was also highly correlated, with significant Pearson's r values observed between Q4 and Quit ($r = 0.63$), Q4 and post-baseline ($r = 0.71$), and Quit versus post-baseline ($r = 0.80$). Overall the FW paradigm results are similar to the PW results in that a stable task representation appears to be modified across quartiles; in the FW paradigm, however, this cannot be directly attributed to increases in absolute physical effort across time (i.e., increments in rope weight).

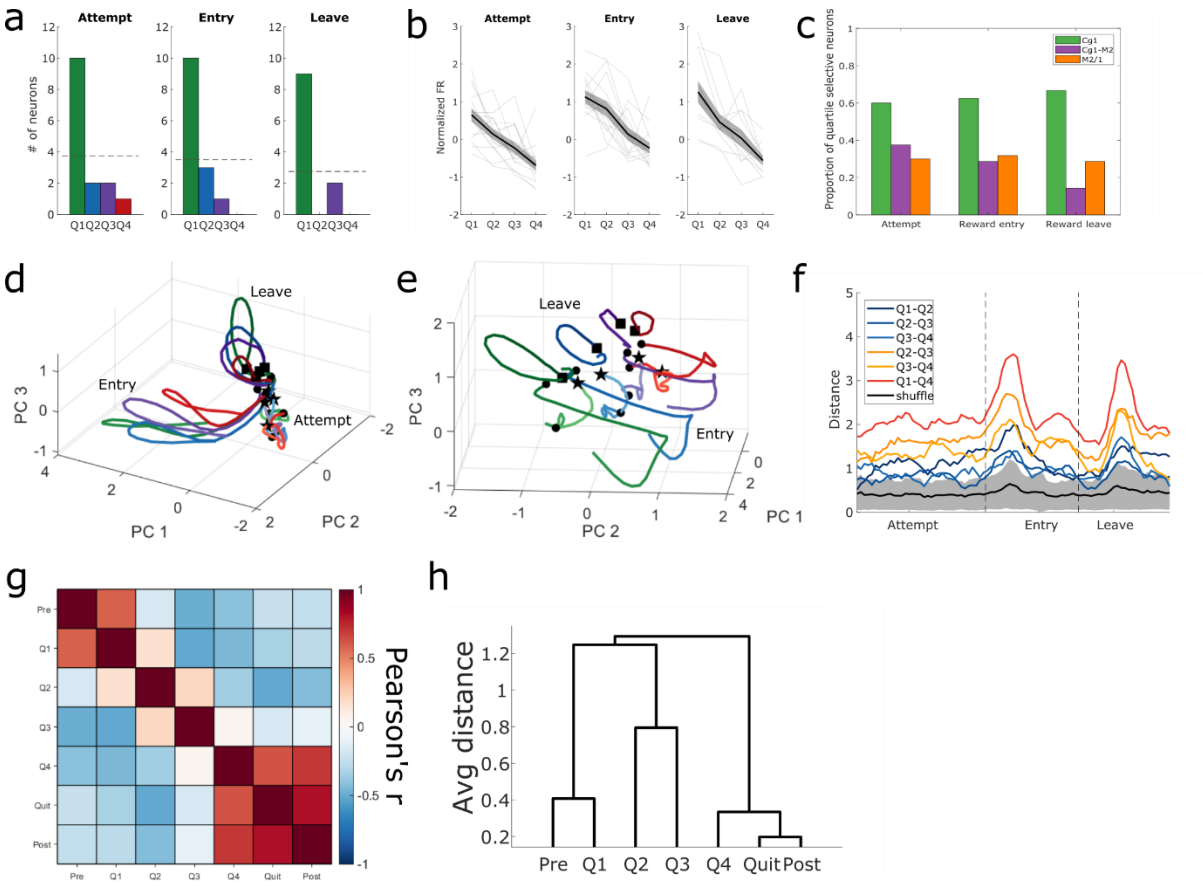

Figure 8: dmPFC neural activity across quartiles during the FW paradigm. a) Distributions of peak neural activity across the four quartiles for each task-phase. Dotted line indicates equal distribution across quartiles. b) Average population activity of task-active neurons with a quartile effect. Activity averaged over all neurons shown as solid black line with grey shading representing S.E.M. Individual neuron activity plotted as light grey lines. c) Proportion of task-responsive neurons with a significant quartile effect by brain region. d & e) PCA of trial activity across the four quartiles, viewed from two different angles. Q1, green; Q2, blue; Q3, purple; Q4, red. f) Pairwise distances between PC trajectories (colored) and average distance between shuffled quartiles (black) with ± 2 S.D. in grey. Cooler colors represent temporally closer pairs while warmer colors indicate more temporally distant pairs. g) Cross-correlation matrix of 2 minute epoch activity across sessions. Pre, pre-baseline; post, post-baseline. h) Dendrogram of hierarchically clustered correlation distances.

Discussion

We utilized two complimentary effortful behavioral paradigms to investigate how dmPFC neural activity responded to changes in task utility, and to test whether dmPFC ensemble activity reflected a rat's progression towards task quitting. In the PW paradigm rats were challenged with progressively heavier weights across the duration of a single session. The FW paradigm offered a stable cost-benefit ratio across the duration of a single session. In both paradigms, behavioral metrics indicate that rats were highly engaged in the task until the final quartile (Q4) where engagement waned. In the PW paradigm, rats took longer to earn rewards and failed rope-pulling more often quartile-by-quartile as they approached their quit weight in Q4. In the FW paradigm, reward intervals and failed attempts only increased in Q4 relative to Q2 and Q3. Overall, these data show that the PW paradigm ramps up in difficulty with commensurate changes in behavior across quartiles, leading towards the quit point. On the other hand, the FW paradigm presents an endurance-like task where behavior is relatively stable in middle quartiles until a change in Q4, which immediately precedes the quit point.

Despite the behavioral differences between the two paradigms, we observed similar neural activity patterns across quartiles in both paradigms. Neural activity was highest during the Q1 and decreased across quartiles, leading to the quit point. Principal component analyses showed conserved neural trajectories for the trial phases (attempt, reward entry, reward leave), which remained relatively stable across the session. However, each quartile had a distinct trajectory. The more temporally close quartiles had closer PCA trajectories, with Q1 and Q4 being the most distinct in both task paradigms. Broader temporal analysis of neural activity encompassing the entire recording session – which included the baseline periods where the rope was unavailable – revealed distinct shifts in activity from the start of the session to the end. Overall, our results suggest that the dmPFC maintains a stable task

representation that is responsive to the changing utility of an environment and reflective of the animal's motivational state for task engagement.

Similar response patterns in Cg1 and M2 neurons

We observed similar activity patterns in neurons from Cg1, Cg1-M2, and M2/1 during WLT performance. The proportion of neurons from each brain area that were active in each task-phase were not significantly different. Furthermore, an equal proportion of task-responsive neurons were quartile-responsive in the PW paradigm. These results were somewhat surprising as we expected M2 neurons to be more responsive to the rope pulling actions of the attempt phase, and Cg1 neurons to be more responsive to the reward phases. However, our results do align with the overlapping roles ascribed to the two brain areas. Generally, M2 has been proposed to select and adapt action plans based on sensory feedback (Reep et al. 1987; Barthas and Kwan, 2017; Olson et al. 2020) while the Cg1 monitors behaviors to determine their utility and bias behavior to maximize value (Shenhav et al. 2013; Holroyd and McClure, 2015). Both brain regions may play complimentary roles in physical effort-based cost-benefit decision making: M2 selects actions and evaluates sensorimotor feedback of the costs and outcomes of those actions, while anterior Cg1 evaluates the overall utility of the selected action plan in comparison to available alternatives. For example, Sul et al. (2010, 2011) recorded from M2 and Cg1 neurons from rats carrying out a probability-based cost-benefit decision making task on a T-maze. M2 neurons encoded choice selection earlier than other brain regions, while the Cg1 was the first region to encode the value of the selected action (Sul et al. 2011). While our task was not designed for the specific timing of choice and value information, we demonstrate that neurons in both regions encode task relevant behaviors and that these are modified by current utility. Our results add further support that anterior Cg1 and M2 likely play similar but complimentary roles in cost-benefit

decision making, specifically at the intersection of action selection, action-outcome monitoring, and adjusting future behavior.

Value encoding in the dmPFC

Here we show that dmPFC neurons on average have high firing rates early on when a task is initiated, but that these firing rates progressively decrease as the task is repetitively performed. Since these patterns were observed in both paradigms (PW and FW) they cannot be solely attributed to progressive increases in rope weight. Furthermore, in both paradigms we observed a general shift in activity patterns from the start of a recording session to the end of the session, which included baseline periods where the rope was not available. We propose that the observed decrease in activity across a session reflects the decrease in task utility across time, and that this corresponds to waning motivation and eventual quitting of the task.

Numerous studies have shown that the Cg1/ACC and M2 (or supplementary motor area in primate) respond to the relative value of behaviors, with firing rates or BOLD responses corresponding to high value options (Croxson et al. 2009; Kennerley et al. 2009; Kennerley and Wallis, 2009; Sul et al. 2011; Hillman and Bilkey, 2010, 2012; Cowen et al. 2012; Porter et al. 2019; Hart et al. 2020). However, in most previous studies subjects had to choose between, or are at least were presented with, multiple stimuli or options in a given task; quitting not being one of the options (though see Zhao et al. 2016). Furthermore, there has generally been an emphasis on the evaluation of the *relative* value of behaviors in the context of all possible behaviors (Hillman and Bilkey, 2010, 2012; Wallis and Kennerley, 2011; Hart et al. 2020) even if there is no overt choice between multiple options (Cowen et al. 2012; Porter et al. 2019). For example, Hart et al. (2020) demonstrated that Cg1 neurons have higher activity when rats are carrying out a sucrose-rewarded progressive lever pressing task with no other option available, compared to when standard rat-chow was concurrently

available. Thus, the presence of an alternative option (chow) discounted the value of the progressive lever pressing action.

In our task rats could only choose between two options – carrying out the task or not; there were no other extrinsically rewarded courses of action available in the environment. In accordance with the dmPFC encoding the relative value of a behavior, we observed a value signal that decreased over the course of the session, culminating in the quit point (quitting being the only alternative course of action). Overall our results further support the idea that dmPFC encodes the relative value of behaviors in order to choose actions that maximize utility (Rushworth et al. 2004). At a certain point, quitting can become the action with optimal utility in response to changing environmental demands and internal state (Noakes, 2012; van der Linden, 2010; Hockey, 2013).

Behavioral state representation in dmPFC activity

When we compared trials where the animal went on to engage in another rope pull, versus trials where the animal disengaged from the task, dmPFC activity in PC space showed general overlap in the neural trajectories. However, after the rat left the reward cap the neural trajectories began to differ significantly between engage and disengage trials. The timing of this divergence was somewhat surprising given the dmPFC's role in evaluating outcome and encoding value, as discussed above. We had predicted that neural activity would diverge earlier in the trial, such as during reward entry/consumption, where the value of the task would presumably be updated. If the value was high, rats would engage in another trial whereas if the value was low, rats would disengage. Indeed, activity in the human ACC during the feedback phase of a task has been correlated to quitting behavior (Zhao et al. 2016). Our results suggest that value updating may be occurring later than we predicted (i.e., after the rat leaves the reward zone), and/or that the dmPFC is more involved in action

monitoring after the choice to re-engage has been made (Botvinick, 2007; Shenhav et al. 2016; but see: Heilbronner and Hayden, 2016).

In this study we also compared neural activity during successful trials versus activity during failed, non-rewarded trials. In PCA space the neural trajectories were highly distinct between these trial types. This distinction could be due to multiple reasons but primarily the behavior on these trial types is different. During successful trials, a rat successfully lifts the weight 30 cm and this triggers an LED and Arduino-controlled sucrose dispensing; the rat goes to the reward cap and consumes the sucrose, and then leaves the reward zone. On failed trials, a rat may very briefly pull the rope or pull it for a long duration but fail to pull the weight to the required height of 30 cm. The rat may not check the reward cap at all on failed trials, or check the reward cap briefly (< 2 seconds as sucrose was not present); we constrained our analysis of failed trials to trials where the animal did visit the reward cap. In addition to these behavioral differences between successful and failed trials, level of engagement could also account for some of the neural activity pattern differences. It is possible that on certain trials the animal attempted the rope pull but was less engaged or motivated at that particular moment, resulting in a failed attempt and reflected in the neural activity. Regardless of the potential causes, taken together these results demonstrate that dmPFC ensemble activity reflects the behavioral states of the rat as they carry out the WLT.

Previous studies that have examined cost-benefit decision making and quitting have tended to use progressive ratio lever pressing paradigms (Hodos, 1961; Schweimer and Hauber, 2005; Rickard et al. 2009; Swanson et al. 2019). In these studies, rodents are required to press a lever progressively more times in order to earn the same reward until they quit at their “breakpoint” ratio. Lesions to the Cg1 (Schweimer and Hauber, 2005) or more ventral prelimbic cortex (Swanson et al. 2019) do not alter the breakpoint of rats in these paradigms. However, lesions to the Cg1 of rats has consistently resulted in avoidance of high-

effort options (climbable-barriers, high-ratio lever presses) when relatively lower effort options are available (Rudebeck et al. 2006; Hauber and Sommer, 2009; Holec et al. 2014; Hart et al. 2017). These results seem to be at odds with one another: lesioning the Cg1 results in rats shifting their choice preference towards lower effort options, but does not affect high effort performance on progressive ratio tasks. Wang et al. (2017) provided some clarity on this discrepancy by showing that muscimol-inactivation of the rat Cg1 impairs performance in a novel "do more get more" task, suggesting that the Cg1 is critical for self-paced effort expenditure. Furthermore, Hart et al. (2020) demonstrated that chemogenetic manipulation of the ACC resulted in rats putting less effort into a progressive ratio task only when alternative rat-chow was available. Thus, in the context of quitting behaviors, the ACC/Cg1 may provide a value signal of the current effortful behavior relative to the other available behaviors in a task environment and this value signal influences the decision to persist at the current behavior or quit (Hillman and Bilkey, 2013).

dmPFC's cognitive map of task space

Two proposed, complimentary functions of the dmPFC are reinforcement-learning and action selection (Paus, 2001; Hadland et al. 2003; Holroyd and McClure, 2015). These two functions work synergistically where the dmPFC learns about and updates a mental model of the world forming an abstract task space in order to select actions that maximize the value of an environment (Heilbronner and Hayden, 2016). Using PCA we demonstrated that the dmPFC contains stable task-space representations of important behaviors during our task: effortful rope pulling (attempt phase), reward approach and consumption (reward entry phase), and prospective task engagement at the end of a trial (reward leave phase). Similar results have been reported showing the Cg1/ACC representing task-space (Lapish et al. 2008; Cowen et al. 2012) including the ACC being able to recall the locations of unchosen rewards when mistakes were made, providing evidence of task-space representations (Mashhoori et

al. 2018). In addition, Wang et al. (2017) reported ACC representations of the sequential steps of a delay-based cost-benefit task where rats were free to nose-poke for longer durations to earn proportionally larger rewards (Wang, et al. 2017). We found similar results to Wang et al. (2017) in the proportion of recorded neurons representing task states (Wang et al. 60%, our results PW: 71%; FW: 86%) and subpopulation responding to cost-benefit contingencies or quartiles (Wang et al. 34%; our results PW average: 33%; FW average: 39%). Thus, the dmPFC may have flexible, sequential representations of the entirety of task-relevant behaviors that can be modified based on cost-benefit contingencies for optimal action selection.

In summary, we used two different effort-based WLT paradigms to demonstrate that neurons in the rat dmPFC represent sequential phases of the task. Neurons in both anterior Cg1 and M2 responded similarly, showing selective firing activity to sustained motor plans (rope pulling), reward processing, and the dynamic utility of the task. Our results add further support that the dmPFC forms a representation of the behaviors that can be carried out in a given environment, and that these task space representations can be modified based on their utility to the animal. This utility signal may be conveyed to other brain regions to decide whether to continue to persist at putting effort into a task, or to quit.

.

References:

Azab H and Hayden BY. 2017. Correlates of decisional dynamics in the dorsal anterior cingulate cortex. *PLoS Biology.* 15(11):1–25.

Barthas F and Kwan AC. 2017. Secondary Motor Cortex: Where 'Sensory' Meets 'Motor' in the Rodent Frontal Cortex. *Trends in Neurosciences*. 40(3):181–193.

Botvinick MM. 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive Affective and Behavioral Neuroscience.* 7(4):356–366.

Cowen SL, Davis GA, and Nitz DA. 2012. Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. *Journal of Neurophysiology.* 107:2393–2407.

Croxson PL, Walton ME, O'Reilly JX, Behrens TEJ, and Rushworth MFS. 2009. Effort-based cost-benefit valuation and the human brain. *The Journal of Neuroscience.* 29(14):4531–4541.

Ebbesen CL, Insanally MN, Kopec CD, Murakami M, Saiki A, and Erlich JC. 2018. More than just a "motor": Recent surprises from the frontal cortex. *Journal of Neuroscience.* 38(44): 9402–9413.

Euston DR, Gruber AJ, and McNaughton BL. 2012. The Role of Medial Prefrontal Cortex in Memory and Decision Making. *Neuron.* 76(6):1057–1070.

Gu X, Staines WA, and Fortier PA. 1999. Quantitative analyses of neurons projecting to primary motor cortex zones controlling limb movements in the rat. *Brain Research.* 835(2):175–187.

Haber SN and Behrens TEJ. 2014. The Neural Network Underlying Incentive-Based Learning: Implications for Interpreting Circuit Disruptions in Psychiatric Disorders. *Neuron.* 83(5):1019–1039.

Hadland KA, Rushworth MFS, Gaffan D, and Passingham RE. 2003. The anterior cingulate and reward-guided selection of actions. *Journal of Neurophysiology.* 89(2):1161–1164.

Hart EE, Blair GJ, O'Dell TJ, Blair HT, and Izquierdo A. 2020. Chemogenetic Modulation and Single-Photon Calcium Imaging in Anterior Cingulate Cortex Reveal a Mechanism for Effort-Based Decisions. *The Journal of Neuroscience.* 40(29):5628–5643.

Hart EE, Gerson JO, Zoken Y, Garcia M, and Izquierdo A. 2017. Anterior cingulate cortex supports effort allocation towards a qualitatively preferred option. *European Journal of Neuroscience.* 46(1):1682–1688.

Hauber W and Sommer S. 2009. Prefrontostriatal circuitry regulates effort-related decision making. *Cerebral Cortex.* 19(10):2240–2247.

Hayden BY. 2018. Economic choice: the foraging perspective. *Current Opinion in Behavioral Sciences.* 24:1–6.

Hayden BY and Platt ML. 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *The Journal of Neuroscience.* 30(9):3339–3346.

Heilbronner SR and Hayden BY. 2016. Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annual Review of Neuroscience.* 39:149–170.

Hillman KL and Bilkey DK. 2010. Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. *The Journal of Neuroscience.* 30(22):7705–7713.

Hillman KL and Bilkey DK. 2012. Neural encoding of competitive effort in the anterior cingulate cortex. *Nature Neuroscience.* 15(9):1290–1297.

Hillman KL and Bilkey DK. 2013. Persisting through subjective effort: A key role for the anterior cingulate cortex? *Behavioral and Brain Sciences.* 36:691–692.

Hockey R. 2013. *The Psychology of Fatigue.* Cambridge (UK): Cambridge University Press.

Hodos W. 1961. Progressive Ratio as a Measure of Reward Strength. *Science.* 134(3483):943–944.

Holec V, Pirot HL, and Euston DR. 2014. Not all effort is equal: the role of the anterior cingulate cortex in different forms of effort-reward decisions. *Frontiers in Behavioral Neuroscience.* 8, 12.

Holroyd CB and McClure SM. 2015. Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological Review.* 122(1):54–83.

Hosokawa T, Kennerley SW, Sloan J, and Wallis JD. 2013. Single-Neuron Mechanisms Underlying Cost-Benefit Analysis in Frontal Cortex. *Journal of Neuroscience.* 33(44):17385–17397.

Kennerley SW, Dahmubed AF, Lara AH, and Wallis JD. 2009. Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience.* 21(6):1162–1178.

Kennerley SW and Wallis JD. 2009. Evaluating choices by single neurons in the frontal lobe: Outcome value encoded across multiple decision variables. *European Journal of Neuroscience.* 29:2061–2073.

Kennerley SW and Walton ME. 2011. Decision making and reward in frontal cortex: complementary evidence from neurophysiological and neuropsychological studies. *Behavioral Neuroscience*. 125(3):297–317.

Kennerley SW, Walton ME, Behrens TEJ, Buckley MJ, and Rushworth MFS. 2006. Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience.* 9(7):940–947.

Klein-Flugge MC, Kennerley SW, Friston K, and Bestmann S. 2016. Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. *Journal of Neuroscience.* 36(39):10002–10015.

Kurniawan IT, Guitart-Masip M, Dayan P, and Dolan RJ. 2013. Effort and valuation in the brain: the effects of anticipation and execution. *The Journal of Neuroscience.* 33(14):6160–6169.

Lapish CC, Durstewitz D, Chandler LJ, and Seamans JK. 2008. Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America.* 105(33):11963–11968.

Mashhoori A, Hashemniayetorshizi S, McNaughton BL, Euston D, and Gruber AJ. 2018. Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. *eLife.* 7:e29793.

Murakami M, Vicente MI, Costa GM, and Mainen ZF. 2014. Neural antecedents of self-initiated actions in secondary motor cortex. *Nature Neuroscience*. 17(11):1574–1582.

Noakes TD. 2012. Fatigue is a brain-derived emotion that regulates the exercise behavior to ensure the protection of whole body homeostasis. *Frontiers in Physiology.* 3, 82:1–13.

Olson JM, Li JK, Montgomery SE, and Nitz DA. 2020. Secondary Motor Cortex Transforms Spatial Information into Planned Action during Navigation. *Current Biology.* 30(10):1845-1854.

Paus T. 2001. Primate anterior cingulate cortex: where motor control drive and cognition interface. *Nature Reviews Neuroscience.* 2:417–424.

Paxinos G and Watson C. 2007. *The Rat Brain in Stereotaxic Coordinates* (Sixth). Elsevier.

Porter BS and Hillman KL. 2020. Regional activity in the rat anterior cingulate cortex and insula during persistence and quitting in a physical-effort task. *eNeuro.* 7(5)ENEURO0243-202020.

Porter BS and Hillman KL. 2019. A Novel Weight Lifting Task for Investigating Effort and Persistence in Rats. *Frontiers in Behavioral Neuroscience.* 13:1–13.

Porter BS, Hillman KL, and Bilkey DK. 2019. Anterior cingulate cortex encoding of effortful behavior. *Journal of Neurophysiology.* 121(2):701–714.

Reep RL, Corwin JV, Hashimoto A, and Watson RT. 1987. Efferent Connections of the Rostral Portion of Medial Agranular Cortex in Rats. *Brain Research Bulletin.* 19(2):203–221.

Rickard JF, Body S, Zhang Z, Bradshaw CM, and Szabadi E. 2009. Effect of Reinforcer Magnitude on Performance Maintained By Progressive-Ratio Schedules. *Journal of the Experimental Analysis of Behavior.* 91(1):75–87.

Rolls ET. 2019. The cingulate cortex and limbic systems for emotion action and memory. *Brain Structure and Function.* 224(9):3001–3018.

Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, and Rushworth MFS. 2006. Separate neural pathways process different decision costs. *Nature Neuroscience.* 9(9):1161–1168.

Rushworth MFS, Walton ME, Kennerley SW, and Bannerman DM. 2004. Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences.* 8(9):410–417.

Salamone JD, Cousins MS, and Bucher S. 1994. Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural Brain Research.* 65:221–229.

Schweimer J and Hauber W. 2005. Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. *Learning and Memory.* 12:334–342.

Shenhav A, Botvinick M, and Cohen J. 2013. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron.* 79(2):217–240.

Shenhav A, Cohen JD, and Botvinick MM. 2016. Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience.* 19(10):1286–1291.

Stephens DW and Krebs JR. 1986. *Foraging Theory.* Princeton (USA). Princeton University Press.

Sul JH, Jo S, Lee D, and Jung MW. 2011. Role of rodent secondary motor cortex in value-based action selection. *Nature Neuroscience.* 14(9):1202–1210.

Sul JH, Kim H, Huh N, Lee D, and Jung MW. 2010. Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron.* 66(3):449–460.

Swanson K, Goldbach HC, and Laubach M. 2019. The rat medial frontal cortex controls pace but not breakpoint in a progressive ratio licking task. *Behavioral Neuroscience.* 133(4):385–397.

van der Linden D. 2010. The urge to stop: The cognitive and biological nature of acute mental fatigue. In: P. L. Ackerman, editor. *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications.* Washington DC (USA). American Psychological Association. p 149–164.

Wallis JD and Kennerley SW. 2011. Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Annals of the New York Academy of Sciences.* 1239:33–42.

Walton ME, Bannerman DM, Alterescu K, and Rushworth MFS. 2003. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *The Journal of Neuroscience.* 23(16):6475–6479.

Wang S. Hu SH. Shi Y. and Li BM. 2017. The roles of the anterior cingulate cortex and its dopamine receptors in self-paced cost–benefit decision making in rats. *Learning and Behavior.* 45(1):89–99.

Wang S, Shi Y, and Li BM. 2017. Neural representation of cost–benefit selections in rat anterior cingulate cortex in self-paced decision making. *Neurobiology of Learning and Memory.* 139:1–10.

Zhao W, Kendrick KM, Chen F, Li H, and Feng T. 2016. Neural circuitry involved in quitting after repeated failures: Role of the cingulate and temporal parietal junction. *Scientific Reports.* 6, 24713:1–12.