The encoding of interoceptive-based predictions by the paraventricular nucleus of the thalamus D2R+ neurons

Briana Machen[1], Sierra N. Miller[1], Al Xin[2], Carine Lampert[1], Lauren Assaf[1], Julia Tucker[1], Sarah Herrell[1], Francisco Pereira[2], Gabriel Loewinger[2], Sofia Beas[1#]

1 – Department of Neurobiology, University of Alabama at Birmingham, Birmingham, AL, USA.

2 – Machine Learning Core, National Institute of Mental Health, Bethesda, MD, USA.

[#] Correspondence: Sofia Beas, Ph.D. Email: sbeas@uab.edu

ABSTRACT

Understanding how the brain integrates internal physiological states with external sensory cues to guide behavior is a fundamental question in neuroscience. This process relies on interoceptive predictions, which are internal models that anticipate changes in the body's physiological state based on sensory inputs and prior experiences. Despite recent advances in identifying the neural substrates of interoceptive predictions, the precise neuronal circuits involved remain elusive. In our study, we demonstrate that Dopamine 2 Receptor (D2R+) expressing neurons in the paraventricular nucleus of the thalamus (PVT) play key roles in interoception and interoceptive predictions. Specifically, these neurons are engaged in behaviors leading to physiologically relevant outcomes, with their activity highly dependent on the interoceptive state of the mice and the expected outcome. Furthermore, we show that chronic inhibition of PVT^{D2R+} neurons impairs the long-term performance of interoceptive-guided motivated behavior. Collectively, our findings provide insights into the role of PVT^{D2R+} neurons in learning and updating state-dependent predictions by integrating past experiences with current physiological conditions to optimize goal-directed behavior.

INTRODUCTION

All organisms perceive the world through a combination of internal and external senses. As such, understanding how the brain integrates internal physiological states with external sensory cues to guide behavior is a fundamental question in neuroscience. Although much is known about perception of external sensory cues, less is known about the mechanisms mediating the perception of internal bodily cues. Interoception is the ability to sense, track, and integrate signals originating from within the body, such as hunger, thirst, temperature, heart rate, and blood CO_2 levels, amongst others[1-5]. However, interoception plays a broader, more encompassing role within the systems governing emotion, motivation, decision making, and social cognition[4,6,7]. Moreover, impairments in the ability to monitor bodily signals exist across many psychiatric illnesses, such as panic disorder, anxiety, depression, and substance use disorders[2,3,8].

Prevailing models of interoception posit that physiological signals are sensed, tracked, integrated, and prioritized to guide behaviors[1-6,8,9]. Studies on the mechanisms mediating interoception have largely focused on the “sensing” and “tracking” aspects. Indeed, hypothalamic areas, such as the arcuate nucleus (ARC) and the lateral hypothalamus (LH), as well as hindbrain areas, such as the medulla and nucleus of the solitary tract (NTS), play a distinctive role in detecting hunger signals from the body and the regulation of feeding behaviors[10,11]. Additional regions, such as the median preoptic nucleus (MnPO) and the subfornical organs (SFO), mediate thirst and water drinking behaviors[12-14]. However, much less is known about how the brain integrates these signals and prioritizes them to create the perception of internal states.

Once sensed, these peripheral physiological signals are integrated with external sensory cues and information from prior experiences to create internal models[4,5,9]. These internal models enable individuals to anticipate changes in the body’s physiological state based on these factors (current physiological state, environmental cues, learned experiences) – a process known as

interoceptive prediction[4,5,9]. These predictive processes allow the brain to prepare for and respond proactively to physiological demands, ensuring efficient and adaptive regulation of bodily processes. Such mechanisms also represent a fundamental aspect of how the brain maintains homeostasis and guides behavior. By predicting physiological needs, the brain prioritizes actions that restore or sustain internal balance, ultimately shaping decision-making and motivation. While significant advances have been made in identifying the neural substrates underlying interoceptive predictions, the specific neuronal circuits responsible for encoding these processes remain poorly understood.

The paraventricular nucleus of the thalamus (PVT) has emerged as a critical node in the limbic system, yet its functional role remains unclear. Anatomical studies show that the PVT receives converging inputs from brain areas that process homeostatic and interoceptive signals[15-17]. Hypothalamic and brainstem areas that process physiological signals, such as hunger[11,18-20], thirst[12-14], thermoregulation[21,22], pain[23-25] , and stress[26-28] project to and recruit the PVT. Moreover, optogenetic activation of inputs to the PVT from hunger processing areas, such as the ARC and LH, result in feeding and food-seeking behaviors[18,20,29]; while stimulation from inputs originating in the MnPO result in water intake and seeking behavior[13,14]. These studies demonstrate that PVT neurons participate in interoception. In addition, the PVT receives signals about environmental cues from cortical areas, such as the medial prefrontal cortex (mPFC)[30-34]. The activity of PVT neurons is also altered by external cues when paired with physiologically relevant stimuli[18,27,30,32-36]. As such, the PVT is anatomically well-positioned to effectively bridge the interoceptive and exteroceptive domains to create internal states that guide behavior.

While commonly glutamatergic, PVT projection neurons are heterogeneous[26,37-39]. Particularly, our previous research found that a subtype of PVT neurons (PVT neurons expressing the dopamine receptor 2; PVT^{D2R+}) are tuned to changes in current physiological states[26,39] and, as such, play a more general role in integrating and encoding these states. In this study, we used fiber photometry to further investigate the *in vivo* dynamics of PVT^{D2R+}

neurons in mice performing a reward linear maze task. Specifically, we focused on thoroughly characterizing the precise role these neurons play in interoceptive perception that influence goal-directed behaviors. We discovered that PVT^{D2R+} neurons integrate factors, such as physiological needs, environmental cues, and previous experiences, to compute an interoceptive prediction. This prediction, in turn, guides motivated behavior. Moreover, we show that chronic chemogenetic inhibition of PVT^{D2R+} neurons results in long-term deficits in interoceptive guided motivated behavior. Collectively, our results highlight the reconceptualization of PVT^{D2R+} neurons as a neural mechanism that is critical for interoception and interoceptive predictions that are needed for the perception of internal states.

METHODS

Mice

All procedures were performed in accordance with the University of Alabama Birmingham *Institutional Animal Care and Use Committee (IACUC)* Guide for the Care and Use of Laboratory Animals. For this, female and male Drd2-Cre (GENSAT founder line ER44) or C57BL/6 mice were used. These mice were between 8 – 15 weeks of age and were group housed under a 12-h light-dark cycle (6 a.m. to 6 p.m. light), at a temperature of 70–74 °F and 40–65% humidity. After surgery, mice were single-housed and were provided with food and water *ad-libitum.* Two days before initiating behavioral training procedures, food was removed for food restriction and the amount of food provided was according to 85% of their free-feeding weight. Water was maintained *ad-libitum,* unless otherwise specified. Mice returned to *ad libitum* feeding after testing.

Viral vectors

For photometry experiments, we used AAV9-hSyn-FLEX-GCaMP6s-WPRE-SV40 which was produced by the Vector Core of the University of Pennsylvania. For our chemogenetic inhibition, we used AAV2-hSyn-DIO-hM4D(Gi)-mCherry which was produced by Addgene. For controls, we used AAV2-CAG-Flex-tdTomato which was produced by the vector core of the University of North Carolina. GABA sensors were AAV-Syn-iGABASnFR2 or AAV-Syn-Flex-iGABASnFR2[40] purchased from Addgene.

Stereotaxic surgery

Mice were given viral injections (approximately 1 μl) at the following stereotaxic coordinates (at a 6° angle): pPVT, −1.60 mm from Bregma, 0.25 mm lateral from midline, and −3.35 mm vertical from the cortical surface. Optical fibers with diameters of 400 μm (0.48 NA - Neurophotometrics) were used. These fibers were implanted over the pPVT immediately following viral injections (targeted 300-400 μm above the injection site). Following stereotaxic injections, animals were allowed to recover for 2–3 weeks to ensure maximal AAV expression.

Linear maze reward-seeking task

The linear maze task consisted of a linear maze (81.28 × 21.59 × 29.21, L ×W× H in cm). The Ethovision (Noldus) tracking system allowed us to track the movement of the mice through the maze. The maze consisted of three zones: trigger zone, corridor, and reward zone. The opposite ends of the track were designated as the trigger zone and reward zone and were connected by a long corridor. The task consisted of 60 min-long sessions of self-paced trials and mice received one test session per day. For each trial, food-restricted mice were trained to wait in the trigger zone for two seconds. An LED mounted on top of the reward magazine port was then illuminated, which signaled reward availability, allowing mice to run from the trigger zone down the corridor and into the reward zone to retrieve a food reward. The strawberry

milkshake (SMS) reward consisted of strawberry Ensure® mixed 1:1 with 20% sucrose water. The non-caloric reward consisted of 10% sucralose (Splenda®) solution. The sucrose reward consisted of 10% sucrose solution. Reward delivery and food consumption signaled trial termination; therefore, mice had to return to the trigger zone to initiate another trial. Experimental schedule and data acquisition were implemented through the Ethovision software for operant control. For training in this task and to avoid potential neophobia, mice were provided with 8 hours of SMS reward in their home cages for two consecutive days before training began. For the initial training session, mice performed the linear maze task without a wait requirement in the trigger zone, so no premature trials occurred. In subsequent sessions, a rule was implemented requiring mice to remain in the trigger zone for two seconds before approaching, allowing for baseline activity measurement prior to approach.

Blocked reward access sessions

Mice that had previously been trained on the linear maze reward-seeking task underwent testing in a modified session in which access to the reward port was physically obstructed. Each session comprised two distinct phases: a 15-minute *Block phase* followed by a 60-minute *Post-Block phase*. During the *Block phase*, a wire mesh tea strainer was used to block access to the reward port. Upon completion of this phase, the session was briefly paused to remove the wire mesh, after which the *Post-Block phase* commenced and continued uninterrupted for the remaining duration. These sessions were conducted as self-paced trials, with no rewards delivered at any point during the session.

Reward magnitude test sessions

Mice previously trained on the linear maze reward-seeking task were tested in a modified session where the amount of reward varied between sessions. Each 60-minute session consisted of three consecutive blocks of 33 trials, with each block delivering a different reward

amount: a small reward (1 µl), a medium reward (3 µl), and a large reward (6 µl). Mice were required to complete the small reward trial block before proceeding to the medium reward block, followed by the large reward block.

Bulk Ca^{2+} and fiber photometry

All photometry experiments were performed using an RZ5P acquisition system (Tucker-Davis Technologies; TDT) equipped with a real-time signal processor and controlled by a software interface (Synapse version 92). Specifically, the system is integrated with two continuous sinusoidally modulated LEDs at 473 nm (211 Hz) and 405 nm (531 Hz) that serve as the light source to excite GCaMP6s and an isosbestic autofluorescence signal, respectively. The LED intensity (ranging 10–15 µW) at the interface between the fiber tip and the animal was constant throughout the session. Fluorescence signals were collected by the same fiber implant that was coupled to a 400 µm optical patch-cord (0.48 NA) and focused on two separate photoreceivers (2151, Newport Corporation) connected to the emission ports of a custom-ordered fiber photometry Mini Cube (Doric Lenses). TTL pulses recorded by the same system were used to annotate the occurrence of behavioral manipulations. For the measurements of fluorescent calcium signals and ΔF/F analysis, a least-squares linear fit to the 405 nm signal to align it to the 470 nm signal was first applied. The resulting fitted 405 nm signal was then used to normalize the 473 nm as follows: ΔF/F = (473 nm signal − fitted 405 nm signal)/fitted 405 nm signal. All GCaMP signal data is presented as the z-score of the ΔF/F from a baseline (2 seconds) prior to the onset of events.

Four-bottle choice assessment

Mice were individually housed and tested in their home cages. After becoming accustomed to each solution, hungry mice were given free access to four bottles in their home cage. Each

bottle contained one of the following: water, a 10% sucrose solution, a non-caloric 10% sucralose solution, or a strawberry milkshake (strawberry Ensure® mixed with 20% sucrose water). After a 2-hour testing period, the total volume consumed (in mL) for each solution was recorded. Consumption was averaged across two sessions. To prevent place preference, the placement of the solutions in the cage was randomized across sessions.

Chemogenetic inhibition

For chemogenetic inhibition, deschloroclozapine (DCZ; Tocris Bioscience, catalog #7193) was dissolved in 1–2% dimethyl sulfoxide (DMSO) in saline and given intraperitoneally (I.P.) at a final dose of 0.1 mg/kg. An equivalent volume of 1-2% DMSO saline was used as vehicle control on SMS pre-exposure days and on post-inhibition testing days. Alternatively, CNO (Enzo Life Sciences, catalog# BML-NS105) was administered at a dose of 10 mg/kg. All injections were administered 30 minutes prior to behavioral testing.

***In vivo* data analysis**

After calculating the z-score of the ΔF/F for each event in every trial the mice performed, the latency to the reward zone was calculated. All trials that took longer than 11 seconds and the events corresponding to that trial (decision to approach, approach, reward zone arrival, reward delivery) were excluded. Less than 10% of the trials met this exclusion criterion. After that, the trials and events for each individual test session were sorted by their trial order within a session. After sorting the trials in each test session, these were then divided into 5 equivalent trial blocks. This binning strategy allowed for comparable representation per animal in both block trials and was used by prior studies to test the effect of trial order[39], which permits the comparison of the activity dynamics of PVT neurons and how they were modulated by the motivational states of the mice across the progression of each test session.

FLMM - Trial order effects were also confirmed using a Functional Linear Mixed Model (FLMM). This analysis, previously described in detail[39,41,42], combines linear mixed models (LMM) and functional regression. The FLMM framework produces plots for each covariate in the model (e.g., trial order effects) to show whether associations with the photometry signal are significant at each timepoint during the behavior. These FLMM output plots display the regression coefficient estimates for each covariate at each timepoint in the trial (solid lines). They also show estimated pointwise (dark-shaded area) and joint (light-shaded area) 95% confidence intervals (CI) for the coefficient estimates. The joint CIs allow for the interpretation of each covariate's effect on the photometry signals across multiple time intervals in the trial. Timepoints where the pointwise CI do not include zero indicate pointwise statistical significance, while timepoints where the joint CI do not include zero indicate joint statistical significance. For the trial order effects we used the FLMM association with continuous variables (https://rpubs.com/gloewinger/1159129). The random intercept model: $\mathbb{E}[Y_{i,j}(s)| X_{i,j} , \gamma_i(s)] = \beta_0(s) + X_{i,j}\ \beta_1(s) + \gamma_{0,i}(s)$ was chosen based on best fit determined by their AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) values.

Concurrent FLMM Model – We applied a concurrent FLMM model to analyze trials of different lengths and estimate how factors that vary within a trial (e.g., state) influence the photometric signals of PVT^{D2R+} neurons. This was necessary as animals took different amounts of time to travel between the trigger and reward zones and thus each “trial” could differ in length. Specifically, we included an “approach” variable (a binary indicator of whether an animal was approaching the reward zone at each trial timepoint) in the model, in addition to its interaction with the trial outcome type. For animal *i* on trial *j,* denote:

• $X_{1,I,j}$: the trial outcome, where $X_{1,I,j}$ = 1 for ‘strawberry milkshake’ (SMS) trials and $X_{1,I,j}$ = 0 for ‘other outcome’ trials.

• $X_{2,I,j}(s)$: approach, where $X_{2,I,j}(s)$ = 1 when mouse *i* is approaching the reward zone on trial *j* at trial timepoint *s* and $X_{2,I,j}$(s)= 0 for trial timepoints after the reward has been delivered.

• $Y_{i,j}(s)$ is the photometry response of mouse *i* on trial *j* at trial timepoint *s*.

$\beta(s)$ is the vector of coefficients at timepoint *s* and $\square_{i,j}(s)$ is Gaussian noise for mouse *i* on trial *j* at trial timepoint *s*.

We used the formula syntax: photometry ~ outcome * approach + (outcome |ID). The model formula was:

$$Y_{i,j}(s) = \beta_0(s) + \gamma_{0,i}(s) + X_{1,i,j}\,[\beta_1(s) + \gamma_{,i1}(s)] + X_{2,i,j}(s)\,\beta_2(s) + X_{1,i,j}\,X_{2,i,j}(s)\beta_3(s) + \square_{,i,j}(s),$$

where (*s*) indicates that a variable is functional (i.e., varies across trial timepoints). $X_{1,i,j}$ does not vary with respect to *s*, as it denotes a trial condition.

The coefficients are interpreted:

• $\beta_0(s)$: The mean signal when a mouse has received (i.e., after reward delivery) an “other” outcome (e.g. a non-caloric reward) at trial timepoint *s*.

• $\beta_1(s)$: The mean difference in signal between hungry strawberry milkshake reward (H-SMS) and the other outcome (e.g. a non-caloric reward) during reward approach (i.e., before non-caloric reward delivery) at trial timepoint *s*.

• $\beta_2(s)$: The mean difference between approach and outcome delivery for other outcome trials at trial timepoint *s*.

• $\beta_3(s)$: This interaction term is interpreted as a difference in differences. Denoting $A(s)$ as the mean signal difference between approach and reward in H-SMS (at trial timepoint *s*) trials, and

B(s) as the mean signal difference (at trial timepoint *s*) between approach and reward in the outcome of comparison (e.g. a non-caloric reward), $\beta_3(s) = A - B$. Thus, $\beta_3(s) = 0$ when the magnitudes of the approach vs. reward zone mean signal difference are equal (at timepoint *s*) on H-SMS and "other" outcome (e.g. a non-caloric reward) trials. Conversely, $\beta_3(s) \neq 0$ when approaching and H-SMS are together associated with an additional effect. Thus, $\beta_3(s)$ captures the degree to which the trial outcome type (e.g. H-SMS vs. non-caloric rewards) modulates the change in signal that occurs when an animal receives the reward (compared to approaching the reward zone).

Altogether, this concurrent FLMM model enables a detailed understanding of how context (state and outcome) and behavioral phase (approach vs. outcome delivery) dynamically influence neuronal activity, shedding light on the differences in anticipatory signaling during motivated behavior.

Linear Interpolation - Linear interpolation was used to estimate unknown values that fall within the range of known data points using the following mathematical formula:

$$y = y1 + (x2 - x1)(y2 - y1)(x - x1)$$

In this study, linear interpolation was employed to standardize the length of neuronal recording data across trials of varying durations. This method estimates intermediate values between known data points by assuming a linear relationship. Specifically, here, the recording of each trial was resampled to the same number of data points using linear interpolation, preserving the overall temporal structure while mitigating the effects of varying trial durations. This strategy allowed for consistent alignment and aggregation of neuronal activity patterns across trials. Linear interpolation was solely used to generate average traces for graphical purposes showing photometry signals during approach (Fig. 4A-F). These approach signals were analyzed by calculating the AUC of the signal during approach timepoints: for trial *j,* we summed the signal

values during approach timepoints and divided that sum by the number of approach timepoints). We then conducted standard hypothesis tests on these AUCs.

Histology

To verify GCaMP6s, GABASnFR2, hM4Di, and TdTomato expression and optical fiber placements, mice were injected with euthanasia solution and subsequently sacrificed via transcardial perfusion, first with PBS and then with paraformaldehyde (PFA; 4% in PBS). Brains were then post-fixed in 4% PFA at 4 °C overnight and cryoprotected using a 30% PBS-buffered sucrose solution for ~24–36 h. Coronal brain sections (50 µm) were generated using a freezing microtome. Images were taken using a Nikon microscope running NIS Elements AR 5.42.03 64-bit (Nikon).

Statistics and data presentation

All data were analyzed using GraphPad Prism (Domatics) and custom-made Python scripts (Anaconda Navigator). FLMM analyses were performed using R scripts (R Studio). For the FLMM Trial Order Effects, the associations with continuous variables and random intercept model was used[39,41,42]. For statistical analyses, two-tailed paired and unpaired t-tests (parametric; test statistic: t), repeated measures ANOVAs, and two-way ANOVAs (parametric; test statistic: F) were used. Post-hoc Tukey's, Dunnett's, and Šídák's multiple comparisons test were used when appropriate and if the corresponding omnibus test detected a significant difference. All data are presented as mean ± s.e.m.

RESULTS

PVT^{D2R+} neurons sense and track hunger states.

We first examined how physiological states impacted PVT^{D2R+} neuronal activity while mice performed a linear maze task that assesses motivation. For this, we trained hungry mice to perform a linear maze task while investigating the *in vivo* dynamics of PVT^{D2R+} neurons (Fig. 1A,B). In this task, mice were placed in a linear maze equipped with a reward delivery port and a bright cue LED light at the end wall of the maze (reward zone). The other far end of the maze was set as the “trigger zone” (Supp. Fig. 1A). Mice had to learn by trial and error to perform trials requiring them to go and remain in the trigger zone for 2 sec to initiate a rewarded trial, which was signaled by a cue light above the reward port. Once the cue light signaled reward availability, mice could approach the reward zone to obtain a reward (strawberry milkshake, SMS). We found that when mice approached the reward zone, PVT^{D2R+} neuronal activity significantly increased (Supp. Fig. 1B). In accordance with their previously suggested role in sensing interoceptive stimuli, such as hunger, we found that the hunger states modulated PVT^{D2R+} neuronal activity. To show this, we first sorted trials performed by trial order, labeling the earliest trials of the session as ‘Trial Group - G1’ and the latest trials as ‘G5’[39] (Supp. Fig. 1C). We implemented this strategy to model the assumption that well-trained food-restricted mice are hungrier at the beginning of each 60 min test session than at the end (when they have eaten several rewards). Importantly, here, the average latency to approach did not differ across the trial groups throughout test sessions (Supp. Fig. 1D). When comparing between early (G1) and late (G5) trials, the mean PVT^{D2R+} neuronal responses were significantly higher in the early trials (Supp. Fig. 1E,F), suggesting that stronger hunger levels lead to higher PVT^{D2R+} neuronal activity. We also found a strong negative correlation between the area under the curve (AUC) during approach and trial order (Supp. Fig. 1G), such that higher PVT^{D2R+} neuronal activity was associated with early trials and diminished activity with later trials. The association between trial order and decreases in signal of PVT^{D2R+} neurons was confirmed using a functional linear mixed modeling (FLMM) statistical framework[39,41,42] (Supp. Fig. 1H). Notably, no correlation was found between approach AUC and latency to approach (Supp. Fig. 1I).

To further probe the ability of PVT^{D2R+} neurons to sense and track hunger levels, we recorded the responses of these neurons during two separate additional test sessions (Fig. 1C) in which: 1) sated mice received a calorie-rich SMS reward (Sated- SMS; S-SMS), and 2) hungry mice received a sweet non-caloric reward (Hungry- Non-Cal). Importantly, we used a within-animal comparison strategy to circumvent potential confounds due to differences in viral expression or fiber placement. For these test sessions, we hypothesized that sated mice would display an overall decrease in PVT^{D2R+} neuronal activity levels across all trials in the session. Moreover, we also hypothesized that if PVT^{D2R+} neurons tracked the hunger levels of the mice when receiving a non-caloric reward (H-NCal), this would result in hungry mice remaining hungry, and thus, PVT^{D2R+} neuronal activity would not be modulated across the entire test session. Indeed,

we observed that when mice were tested in a sated state, there was a significant decrease in the overall responses of PVT^{D2R+} neurons during the reward approach compared to when they were tested while hungry (Fig. 1D). Additionally, trials conducted across the different 'trial groups' in the sated sessions showed significantly higher latencies to approach (Fig. 1E) and more pronounced decreases in PVT^{D2R+} neuronal activity (Fig. 1F-H). PVT^{D2R+} neuronal activity during sated states was also negatively associated with trial order (Fig. 1I,J), such that lower signals were associated with later trials, but not with latency to approach (Fig. 1K). Conversely, when the same mice were tested while hungry and given a sweet non-caloric reward, PVT^{D2R+} neuronal activity significantly increased during the reward approach (Fig. 1D). However, comparing early (G1) and late (G5) group trials in the H-NCal sessions showed similar response magnitudes (Fig. 1L,M). Additionally, the approach AUC was no longer significantly correlated with trial order (Fig. 1N,O) or latency (Fig. 1P). The lack of modulation by the physiological state during H-NCal sessions persisted even after mice were exposed to a non-caloric reward for 4 hours daily over two weeks, teaching them that non-caloric rewards do not satisfy them (Supp. Fig. 2).

Interestingly, in a new cohort of mice tested first while hungry using a caloric reward (H-SMS sessions) and a non-caloric reward (H-NCal sessions), followed by a sated non-caloric reward (S-NCal sessions; Fig. 1R), we found no differences in the magnitude of the PVT^{D2R+} neuronal photometric responses between these three distinct sessions (Fig. 1S-V). These findings show that even when mice are sated, a non-caloric reward still resulted in robust increases in PVT^{D2R+} neuronal activity during reward approach. Importantly, sated mice performed equivalent number of trials as hungry mice (Fig. 1Q,V), suggesting that the decreases in PVT^{D2R+} neuronal photometric responses during sated states are not solely due to an overall reduction in motivation. Altogether, these results strongly suggest that PVT^{D2R+} neurons can track the internal state of the animal. Additionally, these results also suggest that one way they track hunger levels is by sensing the caloric intake.

PVT^{D2R+} neuronal activity is influenced by the interoceptive value of the outcome.

Our findings showed that a non-caloric reward elicited the same response magnitude regardless of state, indicating PVT^{D2R+} neurons might signal the potential interoceptive value of the reward. Indeed, interoceptive value refers to the subjective value assigned to the internal physiological state or bodily sensations used to influence decision-making and behavior signals[2,4,9]. To explore this, we tested whether the interoceptive value of the outcome affected PVT^{D2R+} neuronal activity. We compared the responses of PVT^{D2R+} neurons during three separate sessions where we varied the relative and the interoceptive values of the outcomes (Fig. 2A). First, we examined the responses of PVT^{D2R+} neurons when hungry mice received SMS reward which has a ‘higher relative value’, 10% sucrose solution (Suc) which has a ‘medium relative value’, and plain water (H_2O) which has a ‘low relative value’. We hypothesized that if PVT^{D2R+} neurons were sensitive to the relative value of the reward, switching from SMS to a lower value outcome would decrease the photometric signal during reward approach. However, we found that the response magnitudes during reward approach were similar between

H-SMS and H-Suc (Fig. 2B). Trials within the H-Suc sessions also showed state-dependent effects (Fig. 2C-F), though these were less pronounced than with the calorie-dense SMS reward. Additionally, there was a negative association between PVT^{D2R+} neuronal signal and trial order in the H-Suc sessions (Fig. 2G,H). When comparing reward approach during sessions where hungry mice received water (H- H_2O) as the outcome, we observed a significant decrease in PVT^{D2R+} neuronal responses (Fig. 2B). This decrease occurred quickly upon encountering the new outcome, with no significant differences in signal magnitude between early (G1) and late (G5) trial groups (Fig. 2I-K). In addition, there was no longer an association between signal magnitude and trial order during the approach (Fig. 2L), but the FLMM analysis revealed a negative association after reward zone entry (Fig. 2M). These decreases were likely not due to changes in overall motivation, as mice performed an equivalent number of trials across the test sessions (Fig. 2N). These findings suggest that PVT^{D2R+} neurons are activated by behaviors leading to physiologically relevant outcomes, likely encoding interoceptively-valuable stimuli rather than relative value. Indeed, in a 4-bottle choice test, mice strongly preferred the calorie-rich SMS when given the choice between SMS, 10% sucrose, a non-caloric reward, and water (Fig. 2O).

To further demonstrate that PVT^{D2R+} neurons encode physiologically relevant outcomes, we compared the mean photometric response magnitude of PVT^{D2R+} neurons across three conditions (Fig. 2P): 1) sessions where mice were tested while hungry and given the SMS reward (H-SMS), 2) sessions where mice were tested while hungry and given water (H-H_2O), and 3) sessions where mice had *ad libitum* access to food, water was restricted, and water was given as a reward (T-H_2O). We found that when mice were water-restricted and given water as a reward, there were robust increases in the photometric responses of PVT^{D2R+} neurons during reward approach. The magnitude of these responses was similar to those observed when mice were tested while hungry and given the SMS reward (Fig. 2Q). Importantly, we again observed state-dependent effects, with higher activity during early trials as compared to late trials (Fig.

2R,S), and a negative association between signal magnitude and trial order (Fig. 2T). Overall, these data strongly suggest that PVT^{D2R+} neurons are encoding the behaviors that lead to physiologically relevant outcomes.

PVT^{D2R+} neuronal activity is state and outcome-dependent.

Our findings showed that motivated behavior engages PVT^{D2R+} neurons only when the outcome is physiologically relevant, such as when hungry mice receive food or thirsty mice receive water. However, we previously observed that PVT^{D2R+} neurons remain active even when no reward is given during the linear task[39]. We hypothesized that this activity persists during reward omission due to the expectation of a reward. To test this, we trained mice for 8 weeks in rewarded sessions. This ensured that hungry mice maintained high response rates despite reward omission (H-OM), allowing us to test the effects of reward expectation on PVT^{D2R+} neuronal activity. We also included a session where sated mice experienced reward omission (S-OM) (Fig. 3A). We found that reward approach still elicited robust PVT^{D2R+} neuronal responses after extended training (H-SMS; Fig. 3B). However, complete reward omission (H-OM) significantly decreased PVT^{D2R+} neuronal responses during reward approach compared to H-SMS sessions (Fig. 3B) and increased the latency to approach (Fig. 3C). Additionally, state-dependent modulation of PVT^{D2R+} neurons was no longer observed (Fig. 3E,F), and there was no significant correlation between approach AUC and trial order (Fig. 3G). Surprisingly, FLMM analysis revealed a positive association between trial order and PVT^{D2R+} neuronal response magnitude (Fig. 3H). During the sated omission session (S-OM), PVT^{D2R+} neuronal responses were not significantly modulated by reward approach (Fig. 3B,I-M). Interestingly, mice in the H-OM session performed more trials than in the H-SMS sessions (Fig. 3N). These findings suggest that the decreases in signals were likely not solely due to a loss of motivation.

To further investigate how reward expectation affects PVT^{D2R+} neuronal activity during reward omission, we examined the neuronal responses when access to a reward tray was

physically blocked (Fig. 3O-T), and whether removing this block (without delivering the reward) would increase reward expectation and thus enhance PVT^{D2R+} photometric responses. We hypothesized that a physical barrier serves as a cue signaling inaccessibility, thereby lowering reward expectation. Conversely, removing the barrier could raise expectation by altering the animal's predictive model of reward availability. To test this, we first trained hungry mice in a standard SMS pre-test to establish baseline neuronal responses (Fig. 3P,Q). Next, we conducted an omission (OM) pre-test, confirming that reward omission reduced PVT^{D2R+} activity (Fig. 3P,Q). Finally, the same mice underwent a blocked reward test session (Fig. 3P), which had two phases: 1) *Block Phase* (15 minutes)— a wire mesh physically blocked access to the reward port, 2) *Post-Block Phase*— The mesh was removed, but the reward remained omitted. This design allowed us to assess whether the removal of the physical block alone could elevate reward expectation and modulate PVT^{D2R+} activity. We found that physically blocking the reward port significantly reduced PVT^{D2R+} photometric responses (Fig. 3Q,R). Notably, once the physical block was removed (despite the continued absence of the reward) the photometric responses significantly increased, reaching levels comparable to those observed during the H-SMS pre-test (Fig. 3Q,R).

Next, to further test the influence of outcome expectation on PVT^{D2R+} neuronal activity, we also analyzed the photometric responses of PVT^{D2R+} neurons during 'premature' trials performed by hungry mice. In the linear maze task, mice had to wait for 2 seconds in the trigger zone to receive the cue signaling reward availability. However, in some trials, mice did not wait properly in the trigger zone before approaching, resulting in the absence of the cue during the reward approach. These trials were categorized as 'premature approaches'. We found that during premature approaches, mean PVT^{D2R+} neuronal activity was significantly lower during the approach to the reward zone (Supp. Fig. 3A). Importantly, the latencies to approach were equal between the "correct" rewarded approaches and unrewarded premature approaches (Supp. Fig. 3B). Interestingly, we observed state-dependent effects, with higher activity during early trials

(G1) compared to late trials (G5) in the premature trials (Supp. Fig. 3C,D). Together, these findings suggest that PVT^{D2R+} neuronal activity is modulated within a session depending on outcome expectation.

Lastly, we aimed to test the hypothesis that PVT^{D2R+} neurons are responding to two dissociable but related events: 1) the expectation of an outcome during the approach phase, and 2) the outcome itself. However, because our behavioral task is entirely self-paced, the duration of each trial varies. This variability in trial latencies made it difficult to cleanly separate the neural responses associated with reward expectation during approach from those associated with reward delivery when averaging across grouped trials. To address this limitation and better dissociate PVT^{D2R+} neuronal responses to 'approach only' versus 'reward only' events, we conducted two complementary analyses. In the first analysis, we restricted each trial to either the approach phase or the reward delivery phase. We then calculated the AUC for each approach-only trial within each outcome group (see methods). This approach allowed us to independently examine neural activity during approach and during reward delivery. During the approach phase, we observed that in most test sessions, PVT^{D2R+} neuronal responses were significantly higher in correct trials compared to premature trials (Fig. 4A–F, M). Furthermore, the magnitude of these responses was modulated by the expected value of the outcome: sessions involving higher-value outcomes (e.g., H-SMS) elicited stronger PVT^{D2R+} activity than those with lower-value outcomes (Fig. 4A–F). Interestingly, in H-OM sessions, we found no significant difference in PVT^{D2R+} responses between correct and premature trials (Fig. 4F, M). This lack of differentiation is likely due to the ambiguity introduced by the absence of any outcome in both trial types during H-OM sessions. Unlike blocked access, which directly cued the inaccessibility of reward (Fig. 3O-T), the animal may be unable to determine whether the task contingencies have changed or whether the outcome was simply omitted, leading to a disruption in expectation. In the second analysis, we isolated and compared PVT^{D2R+} neuronal responses specifically during reward delivery and feeding bouts (the first 10 sec after reward

delivery). We found that these photometric responses were also influenced by both the nature of the outcome and the internal state of the mice (Fig. 4G–L). Notably, we observed significant state-dependent changes in PVT^{D2R+} activity within sessions that included caloric rewards (Fig. 4G, H, J, N). These state-dependent modulations were absent in sessions that did not involve caloric rewards, such as H-NCal (Fig. 4I, N), H-H_2O (Fig. 4E, N), and H-OM (Fig. 4F, N).

In our second analysis, we applied a concurrent functional linear mixed model (FLMM), which allows for comparisons across sessions while accounting for variability in trial durations. This model included a binary functional covariate that marked timepoints when the animal was approaching the reward versus when it was receiving the reward. This approach enabled us to examine how behavioral phase (approach vs. outcome delivery), internal state (e.g., hunger), and outcome type (e.g., caloric vs. non-caloric) dynamically influence PVT^{D2R+} photometric responses. The concurrent FLMM analysis confirmed our previous findings and revealed additional differences that were not detectable using trial-averaged methods (Supp. Fig. 4). Specifically, during reward delivery, the mean photometry signal of PVT^{D2R+} neurons was significantly higher in the H-SMS condition compared to S-SMS (Supp. Fig. 4A $\beta_1(s)$), H-H_2O (Supp. Fig. 4D $\beta_1(s)$), and H-OM (Supp. Fig. 4E $\beta_1(s)$). However, in the H-NCal condition, the mean photometry signal was significantly higher than H-SMS during the later stages of reward delivery (Supp. Fig. 4B $\beta_1(s)$), suggesting that PVT^{D2R+} neurons respond similarly to caloric and non-caloric rewards, but only caloric rewards induce state-dependent decreases during prolonged feeding. No significant differences were observed between H-SMS and H-Suc (Supp. Fig. 4C $\beta_1(s)$), indicating that PVT^{D2R+} neurons do not encode relative reward value. Across all conditions, the model showed that PVT^{D2R+} mean signal was consistently lower during approach than during reward delivery (Supp. Fig. 4A–E $\beta_2(s)$). Lastly, this FLMM model enabled us to evaluate the interaction effects between behavioral states (e.g., approaching vs. receiving), internal states (e.g., hungry vs. sated), and outcome types (e.g., SMS vs. NCal), and how these factors collectively influence the neural response of PVT^{D2R+} neurons (Supp. Fig. 4A-F $\beta3(s)$).

Specifically, we observed that the interaction for the H-SMS vs. S-SMS comparison is negative, indicating that the mean differences in photometric responses between these two outcome types are larger during the outcome delivery than during outcome approach. The results were similar for the interaction in the H-SMS vs. H-H_2O comparisons (Supp. Fig. 4D $\beta_3(s)$) and the H-SMS vs. H-OM comparisons (Supp. Fig. 4E $\beta_3(s)$). For the H-SMS vs. NCal comparison, we found that the interaction is first negative and then is positive, showing that the differences in mean neural activity between approach and reward delivery timepoints are lower in H-NCal (than H-SMS) during early portions of the trial. Then, during later portions of the trial, the difference in mean signals between approach and reward timepoints is higher in the H-SMS condition (Supp. Fig. 4B $\beta_3(s)$). Finally, we found no differences between H-SMS and H-Suc (Supp. Fig. 4C $\beta_3(s)$). Overall, the concurrent FLMM analysis demonstrates that PVT^{D2R+} photometric signals are shaped by a combination of the behavioral phase, expected outcome, and internal state.

PVT^{D2R+} neuronal activity during return is modulated by state, outcome, and outcome expectation.

Next, we sought to test the *in vivo* dynamics of these PVT^{D2R+} neurons during the return from the reward zone to the trigger zone. We reasoned that by examining the *in vivo* dynamics during this time, where mice are engaging in the same motivated behavior (running) but not expecting a reward, would allow us to further dissociate the PVT^{D2R+} neuronal responses between expectation of the outcome and the outcome delivery. Interestingly, we found returning from the reward zone to the trigger zone resulted in a robust decrease in the photometric responses of PVT^{D2R+} neurons in all outcome sessions (Fig. 5). Moreover, the responses were also robustly modulated by the state of the mice, the outcome delivered, and the outcome expectation (Fig. 5A-H). Surprisingly, our analysis revealed state-dependent differences when comparing early trials ('G1') to late trials ('G5') in sessions where outcomes that led to greater

satiation were provided (Fig. 5A,B,D,F). Specifically, stronger decreases in the signal of PVT^{D2R+} neurons were observed in early trials as compared to late trials. These effects again were absent in those sessions where non-caloric outcomes were provided (H-NCal, H-H_2O, and H-OM sessions; Fig. 5C,E,G). Importantly, the decreases observed during return and the differences in modulation across various conditions (Fig. 5H) were not due to activity levels returning to baseline after reward delivery. First, mice spent ~20 seconds in the reward zone before returning (Supp. Fig. 5), allowing reward-evoked signals to return to baseline in most groups (Fig. 4G–L). Second, trials with reward delivery and return responses were aligned and analyzed based on reward zone arrival and sorted by the time spent in the reward zone. These trials showed that the initiation of return closely aligned with decreases in the signal of PVT^{D2R+} neurons (Fig. 5I-L). Additionally, injections of the GABA sensor (GABASnFR2) into the PVT revealed increased GABA signaling during return (Supp. Fig. 6A-G), indicating that the decreases during the return phase are due to increases in inhibitory inputs. Interestingly, GABAergic inhibitory signaling remained consistent across trial groups (Supp. Fig. 6H-O), indicating that the observed modulations during return are unlikely due to changes in inhibitory signaling onto these neurons. Further analysis showed that PVT^{D2R+} neuronal responses increased immediately prior to return initiation (Supp. Fig. 7A; Fig. 5I-L). Moreover, the magnitude of these peaks in the signal during the decision to return drove the observed modulation (Supp. Fig. 7B-J). These findings suggest that PVT^{D2R+} neuronal activity increases during the decision to return, which is modulated by internal state, outcome, and expectation, and is then suppressed by inhibitory signaling during return onset.

Motivational states are associated with PVT^{D2R+} neuronal responses across different conditions.

Our findings showed that PVT^{D2R+} neurons mediate interoceptive predictions by integrating information about the animal's state and the expected outcome. Previously, we

demonstrated that PVT^{D2R+} neuronal activity was closely linked to the latency to approach the reward zone in a linear maze task[39]. Based on these findings, we hypothesized that PVT^{D2R+} neuronal activity is associated with latency to approach because latency reflects the motivational strength of the interoceptive prediction. Specifically, lower latencies (fast trials) would be associated with stronger predictions and thus higher PVT^{D2R+} neuronal responses, while higher latencies (slow trials) would be associated with weaker predictions and thus attenuated PVT^{D2R+} neuronal responses. To test this, we sorted trials by approach latency (labeling fast trials as 'L1' and slow trials as 'L5'; Supp. Fig. 8) and calculated the approach slope under various conditions. Consistent with our prior work, we found significant correlations between latency and slope in most conditions (Supp. Fig. 8A–D, F), with fast trials showing steeper slopes. However, this relationship was absent in the H-H_2O (Supp. Fig. 8E) and S-OM (Supp. Fig. 8G) conditions. These results suggest that PVT^{D2R+} activity tracks the motivational strength of interoceptive predictions, modulated by state, outcome, and expectation.

PVT^{D2R+} neurons detect changes in reward magnitude.

Interoception models have posited that brain areas encoding interoceptive predictions have the capability to detect changes in reward magnitude[4-6,9,43]. Since interoceptive predictions help anticipate and regulate bodily conditions, they inherently detect changes in reward magnitude which allow the brain to efficiently allocate effort and attention to optimizing resource acquisition[4,43]. As such, we next tested whether PVT^{D2R+} neurons could detect changes in reward magnitude. To investigate this, we first trained hungry mice to perform the linear maze task to obtain a non-caloric reward (Fig. 6A). We used a non-caloric reward to avoid state-dependent changes that could confound our results. Once we confirmed this during a pre-test (Fig. 6B, E), the mice underwent a 'reward magnitude test session' consisting of three within-session trial blocks, each offering different reward sizes (small: 1µl, medium: 3µl, and large: 6µl; Fig. 6A). We found that PVT^{D2R+} photometric responses were modulated by the different reward

magnitudes, such that bigger rewards resulted in higher response signaling of these neurons, while smaller reward magnitudes resulted in lower response magnitudes (Fig. 6C). No significant changes in approach latency were found between reward magnitude trial blocks (Fig. 6D). Notably, variations in reward magnitude did not alter photometric responses during the return phase (Fig. 6F), although return latencies were reduced (Fig. 6G). Altogether, these findings strongly suggest that PVT^{D2R+} neurons can detect changes in reward magnitude.

PVT^{D2R+} neuronal responses emerge from learned cue-outcome associations.

Brain areas involved in interoceptive predictions also have the distinct ability to predict internal states based on the current state, expected outcomes, and learned environmental cues[4-6,9]. Thus, we next tested whether PVT^{D2R+} neuronal responses also depend on the environmental cues associated with the outcome and state. Indeed, we showed that when mice returned from the reward zone to the trigger zone in the maze task, the photometric responses of PVT^{D2R+} neurons displayed robust decreases that were state- and outcome- dependent (Fig. 5). Thus, we hypothesized that the differences in PVT^{D2R+} neuronal responses between the reward approach and the return to the trigger zone could be due to the learned association between a side of the maze and the outcome associated with it (reward zone = reward vs. trigger zone = no reward). If the increases in PVT^{D2R+} neuronal activity during the approach, and decreases during the return, are due to the learned cue-outcome associations, then these would appear gradually as mice learn to associate each side of the maze with their respective outcome. To test this, we assessed the photometric responses of PVT^{D2R+} neurons as mice learned to perform our linear maze task. Here, we first exposed mice to the SMS and ensured mice were consuming it reliably (Fig. 7A). We found a small modulation of PVT^{D2R+} neuronal responses during the reward approach (Fig. 7B; Supp. Fig. 9A) and during the return to the trigger zone (Fig. 7E; Supp. Fig. 9B) throughout the first training session (S1). However, as mice received more training sessions and learned to associate the reward zone with the delivery of

the SMS and the return with the lack of reward, the PVT^{D2R+} neuronal responses during approach (Fig. 7B; Supp. Fig. 9A) and during return emerged (Fig. 7E; Supp. Fig. 9B) in a training- and session-dependent manner. To identify precisely when responses emerged, we compared the dynamics between the early trial group (G1) versus the late trial group (G5) in training session 1. Surprisingly, no differences were found between them (Fig. 7H,J). We then analyzed training session 2 (Fig. 7I,K) and compared the dynamics between the early (G1) versus late trial group (G5). Again, we found small increases in the overall activity of the PVT^{D2R+} neurons that were present even in the early trial group (G1). These increases were slightly bigger in late trials of both sessions (Fig. 7H, I), however these were not statistically significant. Notably, the lack of within-session ('G1' vs. 'G5') statistical differences for S1 and S2 sessions is likely due to responses being occluded by the satiety effects present, particularly in later trials. Nonetheless, these findings show that these neuronal responses emerge slowly despite successful task acquisition, as evidenced by the increases in trials performed during later sessions (Fig. 7D). For the return responses, we found similar dynamics in the emergence of the responses (Fig. 7J,K). However, by session 2, we found a significant difference between early and late trials (Fig. 7K), presumably due to the certainty of not receiving a reward after return. Altogether, these findings show that PVT^{D2R+} neuronal responses develop as mice learn cue-outcome associations. These findings are also in accordance with the new proposed role of PVT^{D2R+} neurons in the signaling of interoceptive predictions. Indeed, it has been suggested that predictive interoceptive responses develop more slowly than exteroceptive predictions[4] because interoceptive signals tend to be subtler and more diffused than exteroceptive signals[4].

Chronic PVT^{D2R+} neuronal inhibition impairs performance in the linear maze tasks.

Since PVT^{D2R+} neurons encode interoceptive predictions and their responses emerge slowly, across several sessions, we then aimed to investigate whether chronic chemogenetic inhibition of PVT^{D2R+} neurons would impair performance in our linear maze task. To do this, D2-

Cre mice were injected with either an AAV-hSyn-DIO-hM4D(Gi)-mCherry virus or a control virus (AAV-CAG-Flex-TdTomato) in the PVT. Once the virus was expressed, hM4Di- PVT^{D2R+} mice and control- PVT^{D2R+} mice were injected with deschloroclozapine (DCZ; Fig. 8A,B) or CNO (Supp. Fig. 10) 30 mins prior to training in the linear maze task. Given that PVT^{D2R+} neuronal responses were slow to emerge (Supp. Fig. 9), we repeated our chemogenetic inhibition for three consecutive test sessions, followed by three sessions in which the mice were injected with saline prior to testing (Fig. 8B). We found that the chemogenetic inhibition of PVT^{D2R+} neurons did not result in acute changes in performance for both correct (Fig. 8C) and premature trials (Fig. 8E). We found, however, significant differences in the latencies to perform correct trials during the S2 and S3 (Fig. 8D), but no changes in latency to perform premature trials (Fig. 8F). Furthermore, we found that when mice were tested for post-inhibition effects, control mice performed a significantly higher number of correct trials compared to hM4Di-PVT^{D2R+} mice (Fig. 8C). No changes in the performance of premature trials were observed between controls and hM4Di-PVT^{D2R+} mice during post-inhibition testing (Fig. 8E). Lastly, we found that chemogenetic inhibition of PVT^{D2R+} neurons after task acquisition did not affect task performance (Supplementary Fig. 10 I,J). This suggests that these neurons are not essential for the execution or maintenance of task performance after the learning phase. Collectively, these results indicate that PVT^{D2R+} neurons play a critical role in early learning. Furthermore, prolonged inhibition of these neurons during training impairs the sustained performance of behaviors guided by interoceptive motivation.

DISCUSSION

In this study, we characterized a specific role for PVT^{D2R+} neurons in using interoceptive cues to guide goal-directed behavior. We showed that PVT^{D2R+} neurons compute an interoceptive prediction by integrating factors, such as current physiological needs, previous experiences, and environmental cues, to guide motivated behavior. Using a linear maze task

that assesses motivated behavior in combination with *in vivo* fiber photometry of PVT^{D2R+} neurons, we found a significant increase in the response of these neurons during reward approach and reward delivery. Furthermore, these increases were directly related to the current state of the animal but also depend on the outcome expectation. We only observed increases in PVT^{D2R+} neuronal activity during motivated behavior of hungry animals expecting a food reward, but not during the motivated behavior of hungry animals expecting a water outcome, or during the return to the trigger zone. In addition, when mice were thirsty and a water outcome was provided, these neurons again showed increases during reward approach.

Sensing of current physiological states by PVT neurons

Previous anatomical and functional studies have already established the PVT as a critical hub for sensing and responding to changes in homeostasis[15-17]. These studies highlighted the role of PVT neurons as key detectors of homeostatic imbalances, including changes in energy levels[18,29,44-49], hydration[13,14,50], stress[26-28,51], pain[23-25], thermoregulation[21,22], itch[52,53], social behavior[54-56], and arousal states[57-59]. Anatomically, PVT neurons are richly connected to brain regions that monitor these homeostatic signals, such as the hypothalamus, which relays information about hunger, thirst, and stress through hormonal and neuropeptide signals, including orexin[18,60,61], neuropeptide Y[19], and others[62]. The PVT also shares rich connections with brainstem areas, such as rostral ventrolateral medulla[29] and the nucleus of the solitary tract[47], which receive information about the state of the body through somatic afferents and chemosensory receptors[11].

Furthermore, PVT neurons express some of these chemosensory receptors, which enable them to directly sense and respond to homeostatic challenges. For example, PVT neurons express the Glut2 glucose transporter and are capable of directly detecting glucose levels[63]. Electrophysiological recordings show that glucose directly modulates the firing activity of PVT neurons, with some neurons exhibiting increased excitability in response to low glucose

levels[63]. PVT neurons also sense changes in extracellular glucose concentrations via expression of enzymes like glucokinase[48]. Moreover, changes in hydration status, such as dehydration, activate PVT neurons, which in turn promote thirst-related behaviors[13,14]. Beyond basic survival needs, the PVT contributes to maintaining arousal and circadian rhythms, responding to dynamic changes in sleep-wake states and coordinating transitions between them[37,57,64]. Similarly, the PVT regulates stress responses via inputs from the hypothalamus and brainstem and projections to the amygdala and prefrontal cortex, influencing both emotional and physiological adaptations[26,27,33-35,65]. Experimental manipulation of PVT activity using optogenetics and chemogenetics further underscores its role in homeostasis. For instance, activating or inhibiting specific PVT subpopulations can alter feeding[18,31], drinking[13,14], stress reactivity[26,33,35], and arousal levels[37,58]. Collectively, our findings are consistent with previous research showing that the magnitude of the activity of PVT neurons encodes the salience of current internal states, such as the degree of hunger or the level of stress. We also expand on this knowledge by showing that PVT^{D2R+} neurons are likely involved in these state-dependent processes.

Notably, in our study, we also found that a non-caloric reward showed increased PVT^{D2R+} neuronal responses during reward approach and reward delivery for both hunger and sated states. This was somewhat surprising given that a non-caloric reward presumably does not relay information about nutritional value or physiological significance. However, when these non-caloric rewards were provided, the within-session state-dependent modulation was no longer observed, and this effect persisted even with prolonged exposure to the non-caloric solution. Other studies have also found differences in neurocircuit involvement following non-caloric sweetener consumption. The intake of the non-caloric sweetener, sucralose, fails to deactivate hypothalamic areas compared to the ingestion of glucose and fructose[66]. Additionally, compared to caloric options, sucralose consumption shows reduced activity in areas related to reward, including the lateral hypothalamus (LHA), ventral tegmental area (VTA), and nucleus

accumbens (NAc)[67]. However, despite their wide use, relatively little is known about the effects of non-caloric sweeteners in the brain. Indeed, one of the suggested potential risks of non-caloric food intake is the decoupling of sweet taste from the satiety effects needed for maintaining the finely tuned regulation of energy homeostasis[68]. Future research should investigate the underlying mechanisms of how sweet tastants activate PVT^{D2R+} neurons and the observed decoupling of reward and satiety we observed.

Modulation of PVT^{D2R+} neurons by environmental cues and context

Previous studies established that PVT neurons are particularly sensitive to external signals that predict outcomes critical for survival, such as food, safety, or danger[18,19,27,32,33,35,47,59]. When an environmental cue is associated with a physiologically relevant outcome, PVT neurons become activated by the predictive cues[18,33]. This activation is proposed to facilitate the integration of relevant sensory information, to enable the prediction of future internal physiological states to drive adaptive responses accordingly. For instance, in a hunger state, the PVT exhibits heightened activity in response to food-related cues[18]. This heightened response to physiologically relevant cues has been documented in other states like thirst[13,14], fear[26,33,35], safety[27], and social[55] behaviors. These responses engage with downstream brain regions like the NAc[18,26,29,39], amygdala[24,33,35] and prefrontal cortex[27,30,33,34] to prioritize behaviors that will reestablish homeostasis[13,15,29,69]. We confirmed and expanded this previous knowledge by showing that cues and context modulated the responses of PVT^{D2R+} neurons, as shown by the differential responses to the reward zone approach vs. trigger zone return (Fig. 1 and Fig. 6) and correct vs. premature trials (Fig. 4O). These context and cue-driven activations reflect the ability of PVT^{D2R+} neurons to dynamically adjust their activity based on the motivational significance of the stimulus, ensuring that resources are directed toward the most urgent physiological needs. Additionally, a major finding of this study is the slow emergence of these cue-induced responses, which are in contrast to the fast emergence of cue-induced responses

in other brain regions[33,34,70]. Indeed, the slow emergence of cue related responses in PVT neurons has been previously documented in the literature[33,34] and are in accordance with our new proposed role of PVT^{D2R+} neurons in driving interoceptive predictions. This is consistent with modern interoceptive models, which posit that predictive interoceptive responses to environmental cues tend to emerge at a much slower pace than exteroceptive predictions[4,5,9]. This slow emergence has been attributed to interoceptive signals being more subtle and diffused than exteroceptive ones[4,5]. Importantly, dysregulation of cue-driven activation has been implicated in conditions like substance use disorder, where maladaptive cue-responses lead to compulsive behaviors. This highlights the critical balance the PVT^{D2R+} neurons might maintain in the processing and prioritizing of environmental stimuli.

Modulation of PVT^{D2R+} neurons by learned experiences and their role in interoceptive predictions

Models of interoception also posit that beyond sensing of current internal signals from the body, the brain also makes predictions of future internal states[3-6,9]. Indeed, these same models postulate that predictive coding of internal states are not represented in a single brain area but in the connectivity of various hierarchical populations of neurons. This hierarchy enables representation at various levels of abstraction, with higher levels integrating information across multiple sensory domains to create more abstract predictions, while lower levels integrate more exact sensory inputs[9]. Previous studies in humans and rodents show that one key higher-level area involved in these predictions is the insula cortex[1,50,71]. Importantly, the PVT shares strong anatomical connections with the insula cortex that participate in the predictive coding of future internal states[50,72]. Here, we also expand on this knowledge to show that PVT^{D2R+} neurons are part of the hierarchy of populations making predictions of internal states. We now show that beyond their role in detecting current internal states, PVT^{D2R+} neurons also dynamically adjust their activity to reflect both the internal state of the organism and the

expected value of the outcome. This was most obviously demonstrated by the increased PVT^{D2R+} neuronal activity observed when mice were approaching the reward zone. Conversely, the same motivated behavior, 'running' towards the trigger zone, resulted in decreases in the responses of these neurons. Furthermore, these reductions in photometric responses of PVT^{D2R+} neurons during the return phase were also associated with enhanced GABAergic inhibitory signaling. Indeed, the PVT receives diverse and wide-ranging direct inhibitory inputs from brain areas such as the zona incerta, anterior hypothalamic nucleus, and periaqueductal gray, among others[26]. Additionally, it receives indirect inhibitory projections from the prelimbic cortex through the anterior ventral thalamic reticular nucleus[73,74]. Recent previous studies showed that this indirect pathway is critical for shaping goal-directed decisions through a disinhibitory mechanism[73,74]. Importantly, and consistent with their ability to make interoceptive predictions, the responses during the decision to return were influenced by the state of the animal and the expected outcome. In addition, analyzing premature approaches allowed us to assess the responses of PVT^{D2R+} neurons while animals ran with the same vigor and in the same context (toward the reward zone) but without a cue light signaling the reward. During these premature approaches, the activity of PVT^{D2R+} neurons was attenuated. Moreover, these differences between correct and premature responses were not observed when the outcome was no longer relevant (Hungry- H_2O sessions and Hungry- OM sessions). These findings strongly suggest that when the outcome is not relevant, PVT^{D2R+} neurons adjust their activity to update future interoceptive predictions.

Interoception deficits as a major underlying mechanism for prominent psychiatric conditions

Interoception is one of the core drivers of motivation and emotion[2-4,6,8]. However, when interoceptive systems become dysregulated, it can lead to motivational and emotional deficits seen in many prominent psychiatric conditions, including chronic stress, major depressive

disorder (MDD), autism spectrum disorders (ASDs), and others[2,3,6,8]. Both human and rodent studies show that the paraventricular thalamus (PVT) contributes to these disorders in different ways. For example, chronic stress manifests physiologically with increases in cortisol levels and/or impaired hypothalamus-pituitary-adrenal (HPA) axis activation[75,76]. PVT activity is altered after chronic stress[26,77], and lesions or inactivation of the PVT impair animals' ability to adapt to novel stressors[28]. MDD, characterized by low mood and low motivation, is highly linked to autonomic dysfunction and inflammation[78]. Studies in animals show that inflammation results in PVT dysfunction and motivational impairments[79-82], indicating that dysregulation of PVT neurons could be a key driver of low motivation in MDD. Additionally, anxiety disorders, eating disorders, substance use disorders, and pain disorders are other well-known conditions resulting from dysregulation of interoceptive processes[2,3,6,8]. The PVT has been implicated in these conditions in both humans and rodents[83].

Limitations

Previous studies have demonstrated that PVT neurons exhibit heterogeneity, which has important functional implications[37-39]. For example, our previous research has shown that other types of neurons, such as PVT neurons that do not express the D2 receptor, are functionally distinct from PVT^{D2R+} neurons. During motivated behavior, these PVT^{D2R-} neurons show decreases in photometric responses and are not as attuned to the physiological state of the animal[39]. Therefore, one main limitation of the present study is the use of fiber photometry to investigate the activity of PVT^{D2R+} neurons and their role in interoceptive signaling and predictions. It is possible that the observed PVT^{D2R+} neuronal activity shown here includes responses from other PVT subpopulations within the PVT^{D2R+} neurons. Indeed, using single nucleus RNA, a previous study showed that PVT^{D2R+} neurons can be further subdivided into two subtypes: Esr1+ and Col12a1+[38]. Thus, it is possible that the distinct effects observed are mediated by these distinct subpopulations of PVT^{D2R+} neurons. Due to these limitations, future

studies should use single-cell calcium imaging of PVT neurons to fully characterize the activity of PVT^{D2R+} and PVT^{D2R-} neurons during interoceptive prediction. This approach will help understand how factors such as physiological states, environmental cues, and previous experiences impact these neurons and their role in interoceptive predictions.

In conclusion, pinpointing the functional role of PVT neurons and their contribution to motivated behavior has been challenging. This difficulty arises from the recently discovered heterogeneity in PVT neurons[37-39,84] and their opposing functional responses[39]. Additionally, the PVT has been implicated in various conditions and roles, adding to the complexity. The findings presented here illuminate the role of PVT^{D2R+} neurons in mediating interoceptive predictions, proposing a more unified role that contextualizes previous findings. Collectively, and as highlighted here, interoception signaling and predictions are central to creating internal states that guide adaptive behavior. Dysregulation of interoceptive processes results in a range of psychiatric and neurological conditions. Thus, a deeper understanding of the neural basis of interoception may offer new insights into potential therapeutics for these debilitating conditions. By elucidating key areas and neuronal populations involved and how they respond to different factors, this research may help identify new therapeutic targets aimed at restoring healthy interoceptive signaling in individuals.

ACKNOWLEDGEMENTS

We thank Dr. Zachary Knight, Dr. Kirstie Cummings, Dr. Jamie Peters, Dr. Barry Setlow, Dr. Chase Francis and Jonathan Ross Folmar for their comments on the manuscript. This work was supported by Department of Neurobiology at the University of Alabama Birmingham, an R00 MH126429 to S.B., a NARSAD Young Investigator Award by the Brain and Behavior Research Foundation to S.B., a T32 by the National Cancer Institute of the National Institutes of Health under award number T32CA047888 to B.M., a NIGMS K12 Institutional Research and Career

Development Award to S.N.M. and by the NIMH Intramural Research Program (project ZIC-MH002968, authors A.X., F.P., and G.L.).

AUTHOR CONTRIBUTIONS

B.M., C.L., and S.N.M. performed the behavioral experiments and surgeries. L.A., J.T., and S.H. assisted with behavioral experiments and histology. B.M. and S.B. performed all the data analysis. A.X., G.L., and F.P. assisted with the FLMM concurrent model analyses. S.B. designed the study and interpreted the results. S.B. wrote the manuscript. B.M., S.N.M., and G.L., provided extensive edits and comments to the manuscript.

COMPETING INTERESTS

The authors declare no competing interests. This content is solely the responsibility of the authors and does not necessarily represent the views of the NIH.

REFERENCES

1. Chen, W.G., Schloesser, D., Arensdorf, A.M., Simmons, J.M., Cui, C., Valentino, R., Gnadt, J.W., Nielsen, L., Hillaire-Clarke, C.S., Spruance, V., et al. (2021). The Emerging Science of Interoception: Sensing, Integrating, Interpreting, and Regulating Signals within the Self. Trends Neurosci *44*, 3-16. 10.1016/j.tins.2020.10.007.
2. Khalsa, S.S., Adolphs, R., Cameron, O.G., Critchley, H.D., Davenport, P.W., Feinstein, J.S., Feusner, J.D., Garfinkel, S.N., Lane, R.D., Mehling, W.E., et al. (2018). Interoception and Mental Health: A Roadmap. Biol Psychiatry Cogn Neurosci Neuroimaging *3*, 501-513. 10.1016/j.bpsc.2017.12.004.
3. Quadt, L., Critchley, H.D., and Garfinkel, S.N. (2018). The neurobiology of interoception in health and disease. Ann N Y Acad Sci *1428*, 112-128. 10.1111/nyas.13915.
4. Feldman, M.J., Bliss-Moreau, E., and Lindquist, K.A. (2024). The neurobiology of interoception and affect. Trends Cogn Sci *28*, 643-661. 10.1016/j.tics.2024.01.009.
5. Berntson, G.G., and Khalsa, S.S. (2021). Neural Circuits of Interoception. Trends Neurosci *44*, 17-28. 10.1016/j.tins.2020.09.011.
6. Tsakiris, M., and Critchley, H. (2016). Interoception beyond homeostasis: affect, cognition and mental health. Philos Trans R Soc Lond B Biol Sci *371*. 10.1098/rstb.2016.0002.
7. Adolfi, F., Couto, B., Richter, F., Decety, J., Lopez, J., Sigman, M., Manes, F., and Ibanez, A. (2017). Convergence of interoception, emotion, and social cognition: A twofold fMRI meta-analysis and lesion approach. Cortex *88*, 124-142. 10.1016/j.cortex.2016.12.019.
8. Bonaz, B., Lane, R.D., Oshinsky, M.L., Kenny, P.J., Sinha, R., Mayer, E.A., and Critchley, H.D. (2021). Diseases, Disorders, and Comorbidities of Interoception. Trends Neurosci *44*, 39-51. 10.1016/j.tins.2020.09.009.
9. Petzschner, F.H., Garfinkel, S.N., Paulus, M.P., Koch, C., and Khalsa, S.S. (2021). Computational Models of Interoception and Body Regulation. Trends Neurosci *44*, 63-76. 10.1016/j.tins.2020.09.012.
10. Saper, C.B., Chou, T.C., and Elmquist, J.K. (2002). The need to feed: homeostatic and hedonic control of eating. Neuron *36*, 199-211. 10.1016/s0896-6273(02)00969-8.
11. Zimmerman, C.A., and Knight, Z.A. (2020). Layers of signals that regulate appetite. Curr Opin Neurobiol *64*, 79-88. 10.1016/j.conb.2020.03.007.
12. Grove, J.C.R., and Knight, Z.A. (2024). The neurobiology of thirst and salt appetite. Neuron *112*, 3999-4016. 10.1016/j.neuron.2024.10.028.
13. Leib, D.E., Zimmerman, C.A., Poormoghaddam, A., Huey, E.L., Ahn, J.S., Lin, Y.C., Tan, C.L., Chen, Y., and Knight, Z.A. (2017). The Forebrain Thirst Circuit Drives Drinking through Negative Reinforcement. Neuron *96*, 1272-1281 e1274. 10.1016/j.neuron.2017.11.041.
14. Allen, W.E., DeNardo, L.A., Chen, M.Z., Liu, C.D., Loh, K.M., Fenno, L.E., Ramakrishnan, C., Deisseroth, K., and Luo, L. (2017). Thirst-associated preoptic neurons encode an aversive motivational drive. Science *357*, 1149-1155. 10.1126/science.aan6747.
15. Penzo, M.A., and Gao, C. (2021). The paraventricular nucleus of the thalamus: an integrative node underlying homeostatic behavior. Trends Neurosci *44*, 538-549. 10.1016/j.tins.2021.03.001.
16. McGinty, J.F., and Otis, J.M. (2020). Heterogeneity in the Paraventricular Thalamus: The Traffic Light of Motivated Behaviors. Front Behav Neurosci *14*, 590528. 10.3389/fnbeh.2020.590528.

17. Kirouac, G.J. (2015). Placing the paraventricular nucleus of the thalamus within the brain circuits that control behavior. Neurosci Biobehav Rev *56*, 315-329. 10.1016/j.neubiorev.2015.08.005.
18. Meffre, J., Sicre, M., Diarra, M., Marchessaux, F., Paleressompoulle, D., and Ambroggi, F. (2019). Orexin in the Posterior Paraventricular Thalamus Mediates Hunger-Related Signals in the Nucleus Accumbens Core. Curr Biol *29*, 3298-3306 e3294. 10.1016/j.cub.2019.07.069.
19. Horio, N., and Liberles, S.D. (2021). Hunger enhances food-odour attraction through a neuropeptide Y spotlight. Nature *592*, 262-266. 10.1038/s41586-021-03299-4.
20. Betley, J.N., Cao, Z.F., Ritola, K.D., and Sternson, S.M. (2013). Parallel, redundant circuit organization for homeostatic control of feeding behavior. Cell *155*, 1337-1350. 10.1016/j.cell.2013.11.002.
21. Kanai, M., Kamiizawa, R., Hitora-Imamura, N., and Minami, M. (2022). Exposure to hot and cold environments activates neurons projecting from the paraventricular thalamic nucleus to brain regions related to approach and avoidance behaviors. J Therm Biol *103*, 103157. 10.1016/j.jtherbio.2021.103157.
22. Bhatnagar, S., and Dallman, M.F. (1999). The paraventricular nucleus of the thalamus alters rhythms in core temperature and energy balance in a state-dependent manner. Brain Res *851*, 66-75. 10.1016/s0006-8993(99)02108-3.
23. Mindaye, S.A., Chen, W.H., Lin, S.C., Chen, Y.C., Abdelaziz, M.A., Tzeng, Y.S., Shih, A.C., Chen, S.Y., Yang, S.B., and Chen, C.C. (2024). Separate anterior paraventricular thalamus projections differentially regulate sensory and affective aspects of pain. Cell Rep *43*, 114946. 10.1016/j.celrep.2024.114946.
24. Liang, S.H., Zhao, W.J., Yin, J.B., Chen, Y.B., Li, J.N., Feng, B., Lu, Y.C., Wang, J., Dong, Y.L., and Li, Y.Q. (2020). A Neural Circuit from Thalamic Paraventricular Nucleus to Central Amygdala for the Facilitation of Neuropathic Pain. J Neurosci *40*, 7837-7854. 10.1523/JNEUROSCI.2487-19.2020.
25. Chang, Y.T., Chen, W.H., Shih, H.C., Min, M.Y., Shyu, B.C., and Chen, C.C. (2019). Anterior nucleus of paraventricular thalamus mediates chronic mechanical hyperalgesia. Pain *160*, 1208-1223. 10.1097/j.pain.0000000000001497.
26. Beas, B.S., Wright, B.J., Skirzewski, M., Leng, Y., Hyun, J.H., Koita, O., Ringelberg, N., Kwon, H.B., Buonanno, A., and Penzo, M.A. (2018). The locus coeruleus drives disinhibition in the midline thalamus via a dopaminergic mechanism. Nat Neurosci *21*, 963-973. 10.1038/s41593-018-0167-4.
27. Wang, Z., Wang, Z., and Zhou, Q. (2024). Modulation of learning safety signals by acute stress: paraventricular thalamus and prefrontal inhibition. Neuropsychopharmacology *49*, 961-973. 10.1038/s41386-023-01790-2.
28. Bhatnagar, S., and Corbett, B. (2017). Paraventricular Thalamic Regulation of Habituation to Repeated Stress: Molecular and Network Mechanisms. (NATURE PUBLISHING GROUP MACMILLAN BUILDING, 4 CRINAN ST, LONDON N1 9XW, ENGLAND), pp. S557-S558.
29. Beas, S., Gu, X., Leng, Y., Koita, O., Rodriguez-Gonzalez, S., Kindel, M., Matikainen-Ankney, B.A., Larsen, R.S., Kravitz, A.V., Hoon, M.A., and Penzo, M.A. (2020). A ventrolateral medulla-midline thalamic circuit for hypoglycemic feeding. Nat Commun *11*, 6218. 10.1038/s41467-020-19980-7.
30. Otis, J.M., Zhu, M., Namboodiri, V.M.K., Cook, C.A., Kosyk, O., Matan, A.M., Ying, R., Hashikawa, Y., Hashikawa, K., Trujillo-Pisanty, I., et al. (2019). Paraventricular Thalamus Projection Neurons Integrate Cortical and Hypothalamic Signals for Cue-Reward Processing. Neuron *103*, 423-431 e424. 10.1016/j.neuron.2019.05.018.

31. Do-Monte, F.H., Minier-Toribio, A., Quinones-Laracuente, K., Medina-Colon, E.M., and Quirk, G.J. (2017). Thalamic Regulation of Sucrose Seeking during Unexpected Reward Omission. Neuron *94*, 388-400 e384. 10.1016/j.neuron.2017.03.036.
32. Campus, P., Covelo, I.R., Kim, Y., Parsegian, A., Kuhn, B.N., Lopez, S.A., Neumaier, J.F., Ferguson, S.M., Solberg Woods, L.C., Sarter, M., and Flagel, S.B. (2019). The paraventricular thalamus is a critical mediator of top-down control of cue-motivated behavior in rats. Elife *8*. 10.7554/eLife.49041.
33. Do-Monte, F.H., Quinones-Laracuente, K., and Quirk, G.J. (2015). A temporal shift in the circuits mediating retrieval of fear memory. Nature *519*, 460-463. 10.1038/nature14030.
34. Quinones-Laracuente, K., Vega-Medina, A., and Quirk, G.J. (2021). Time-Dependent Recruitment of Prelimbic Prefrontal Circuits for Retrieval of Fear Memory. Front Behav Neurosci *15*, 665116. 10.3389/fnbeh.2021.665116.
35. Penzo, M.A., Robert, V., Tucciarone, J., De Bundel, D., Wang, M., Van Aelst, L., Darvas, M., Parada, L.F., Palmiter, R.D., He, M., et al. (2015). The paraventricular thalamus controls a central amygdala fear circuit. Nature *519*, 455-459. 10.1038/nature13978.
36. Paniccia, J.E., Vollmer, K.M., Green, L.M., Grant, R.I., Winston, K.T., Buchmaier, S., Westphal, A.M., Clarke, R.E., Doncheck, E.M., Bordieanu, B., et al. (2023). Restoration of a paraventricular thalamo-accumbal behavioral suppression circuit prevents reinstatement of heroin seeking. Neuron. 10.1016/j.neuron.2023.11.024.
37. Gao, C., Leng, Y., Ma, J., Rooke, V., Rodriguez-Gonzalez, S., Ramakrishnan, C., Deisseroth, K., and Penzo, M.A. (2020). Two genetically, anatomically and functionally distinct cell types segregate across anteroposterior axis of paraventricular thalamus. Nat Neurosci *23*, 217-228. 10.1038/s41593-019-0572-3.
38. Gao, C., Gohel, C.A., Leng, Y., Ma, J., Goldman, D., Levine, A.J., and Penzo, M.A. (2023). Molecular and spatial profiling of the paraventricular nucleus of the thalamus. Elife *12*. 10.7554/eLife.81818.
39. Beas, S., Khan, I., Gao, C., Loewinger, G., Macdonald, E., Bashford, A., Rodriguez-Gonzalez, S., Pereira, F., and Penzo, M.A. (2024). Dissociable encoding of motivated behavior by parallel thalamo-striatal projections. Curr Biol *34*, 1549-1560 e1543. 10.1016/j.cub.2024.02.037.
40. Ilya Kolb, J.P.H., Akihiro Matsumoto, Benjamin J. Arthur, Yan Zhang, Arthur Tsang, Daniel Reep, Getahun Tsegaye, Jihong Zheng, Ronak Patel, Loren L. Looger, Jonathan S. Marvin, Wyatt L. Korff, Keisuke Yonehara, Glenn C. Turner (2025). iGABASnFR2: Improved genetically encoded protein sensors of GABA. bioRxiv. https://doi.org/10.1101/2025.03.25.644953.
41. Loewinger, G., Cui, E., Lovinger, D., and Pereira, F. (2023). A Statistical Framework for Analysis of Trial-Level Temporal Dynamics in Fiber Photometry Experiments. bioRxiv. 10.1101/2023.11.06.565896.
42. Cui, E., Leroux, A., Smirnova, E., and Crainiceanu, C.M. (2022). Fast Univariate Inference for Longitudinal Functional Models. J Comput Graph Stat *31*, 219-230. 10.1080/10618600.2021.1950006.
43. Fujimoto, A., Murray, E.A., and Rudebeck, P.H. (2021). Interaction between decision-making and interoceptive representations of bodily arousal in frontal cortex. Proc Natl Acad Sci U S A *118*. 10.1073/pnas.2014781118.
44. Zhang, X., and van den Pol, A.N. (2017). Rapid binge-like eating and body weight gain driven by zona incerta GABA neuron activation. Science *356*, 853-859. 10.1126/science.aam7100.
45. Zhang, J., Chen, D., Sweeney, P., and Yang, Y. (2020). An excitatory ventromedial hypothalamus to paraventricular thalamus circuit that suppresses food intake. Nat Commun *11*, 6326. 10.1038/s41467-020-20093-4.

46. Ye, Q., Nunez, J., and Zhang, X. (2022). Oxytocin Receptor-Expressing Neurons in the Paraventricular Thalamus Regulate Feeding Motivation through Excitatory Projections to the Nucleus Accumbens Core. J Neurosci *42*, 3949-3964. 10.1523/JNEUROSCI.2042-21.2022.
47. Ong, Z.Y., Liu, J.J., Pang, Z.P., and Grill, H.J. (2017). Paraventricular Thalamic Control of Food Intake and Reward: Role of Glucagon-Like Peptide-1 Receptor Signaling. Neuropsychopharmacology *42*, 2387-2397. 10.1038/npp.2017.150.
48. Kessler, S., Labouebe, G., Croizier, S., Gaspari, S., Tarussio, D., and Thorens, B. (2021). Glucokinase neurons of the paraventricular nucleus of the thalamus sense glucose and decrease food consumption. iScience *24*, 103122. 10.1016/j.isci.2021.103122.
49. Cheng, J., Wang, J., Ma, X., Ullah, R., Shen, Y., and Zhou, Y.D. (2018). Anterior Paraventricular Thalamus to Nucleus Accumbens Projection Is Involved in Feeding Behavior in a Novel Environment. Front Mol Neurosci *11*, 202. 10.3389/fnmol.2018.00202.
50. Livneh, Y., Sugden, A.U., Madara, J.C., Essner, R.A., Flores, V.I., Sugden, L.A., Resch, J.M., Lowell, B.B., and Andermann, M.L. (2020). Estimation of Current and Future Physiological States in Insular Cortex. Neuron *105*, 1094-1111 e1010. 10.1016/j.neuron.2019.12.027.
51. R, G.A., Campbell, E.J., Walker, L.C., S, S.C.n., Muthmainah, M., F, S.K., A, M.G., O'Shea, M.J., He, S., Dayas, C.V., et al. (2023). A paraventricular thalamus to insular cortex glutamatergic projection gates "emotional" stress-induced binge eating in females. Neuropsychopharmacology *48*, 1931-1940. 10.1038/s41386-023-01665-6.
52. Yun-Qing Li, J.-N.L., Liu-Jie Zhao et al. (2025). Central Recruitment mechanism of Itch and Pain Through the Paraventricular Thalamic Nucleus in a Pathway- and Intensity-Dependent Manner. Research Square PREPRINT https://doi.org/10.21203/rs.3.rs-5615052/v1.
53. Li, J.N., Wu, X.M., Zhao, L.J., Sun, H.X., Hong, J., Wu, F.L., Chen, S.H., Chen, T., Li, H., Dong, Y.L., and Li, Y.Q. (2023). Central medial thalamic nucleus dynamically participates in acute itch sensation and chronic itch-induced anxiety-like behavior in male mice. Nat Commun *14*, 2539. 10.1038/s41467-023-38264-4.
54. Dong, X., Li, Y., and Kirouac, G.J. (2015). Blocking of orexin receptors in the paraventricular nucleus of the thalamus has no effect on the expression of conditioned fear in rats. Front Behav Neurosci *9*, 161. 10.3389/fnbeh.2015.00161.
55. Dong, X., Li, S., and Kirouac, G.J. (2020). A projection from the paraventricular nucleus of the thalamus to the shell of the nucleus accumbens contributes to footshock stress-induced social avoidance. Neurobiol Stress *13*, 100266. 10.1016/j.ynstr.2020.100266.
56. Kooiker, C.L., Birnie, M.T., and Baram, T.Z. (2021). The Paraventricular Thalamus: A Potential Sensor and Integrator of Emotionally Salient Early-Life Experiences. Front Behav Neurosci *15*, 673162. 10.3389/fnbeh.2021.673162.
57. Ren, S., Wang, Y., Yue, F., Cheng, X., Dang, R., Qiao, Q., Sun, X., Li, X., Jiang, Q., Yao, J., et al. (2018). The paraventricular thalamus is a critical thalamic area for wakefulness. Science *362*, 429-434. 10.1126/science.aat2512.
58. Hua, R., Wang, X., Chen, X., Wang, X., Huang, P., Li, P., Mei, W., and Li, H. (2018). Calretinin Neurons in the Midline Thalamus Modulate Starvation-Induced Arousal. Curr Biol *28*, 3948-3959 e3944. 10.1016/j.cub.2018.11.020.
59. Zhu, Y., Nachtrab, G., Keyes, P.C., Allen, W.E., Luo, L., and Chen, X. (2018). Dynamic salience processing in paraventricular thalamus gates associative learning. Science *362*, 423-429. 10.1126/science.aat0481.

60. Kelley, A.E., Baldo, B.A., and Pratt, W.E. (2005). A proposed hypothalamic-thalamic-striatal axis for the integration of energy balance, arousal, and food reward. J Comp Neurol *493*, 72-85. 10.1002/cne.20769.
61. Iglesias, A.G., and Flagel, S.B. (2021). The Paraventricular Thalamus as a Critical Node of Motivated Behavior via the Hypothalamic-Thalamic-Striatal Circuit. Front Integr Neurosci *15*, 706713. 10.3389/fnint.2021.706713.
62. Curtis, G.R., Oakes, K., and Barson, J.R. (2020). Expression and Distribution of Neuropeptide-Expressing Cells Throughout the Rodent Paraventricular Nucleus of the Thalamus. Front Behav Neurosci *14*, 634163. 10.3389/fnbeh.2020.634163.
63. Labouebe, G., Boutrel, B., Tarussio, D., and Thorens, B. (2016). Glucose-responsive neurons of the paraventricular thalamus control sucrose-seeking behavior. Nat Neurosci *19*, 999-1002. 10.1038/nn.4331.
64. Kolaj, M., Zhang, L., Ronnekleiv, O.K., and Renaud, L.P. (2012). Midline thalamic paraventricular nucleus neurons display diurnal variation in resting membrane potentials, conductances, and firing patterns in vitro. J Neurophysiol *107*, 1835-1844. 10.1152/jn.00974.2011.
65. Engelke, D.S., Zhang, X.O., O'Malley, J.J., Fernandez-Leon, J.A., Li, S., Kirouac, G.J., Beierlein, M., and Do-Monte, F.H. (2021). A hypothalamic-thalamostriatal circuit that controls approach-avoidance conflict in rats. Nat Commun *12*, 2517. 10.1038/s41467-021-22730-y.
66. van Opstal, A.M., Kaal, I., van den Berg-Huysmans, A.A., Hoeksma, M., Blonk, C., Pijl, H., Rombouts, S., and van der Grond, J. (2019). Dietary sugars and non-caloric sweeteners elicit different homeostatic and hedonic responses in the brain. Nutrition *60*, 80-86. 10.1016/j.nut.2018.09.004.
67. Frank, G.K., Oberndorfer, T.A., Simmons, A.N., Paulus, M.P., Fudge, J.L., Yang, T.T., and Kaye, W.H. (2008). Sucrose activates human taste pathways differently from artificial sweetener. Neuroimage *39*, 1559-1569. 10.1016/j.neuroimage.2007.10.061.
68. Low, Y.Q., Lacy, K., and Keast, R. (2014). The role of sweet taste in satiation and satiety. Nutrients *6*, 3431-3450. 10.3390/nu6093431.
69. Zhang, L., Koller, J., Ip, C.K., Gopalasingam, G., Bajaj, N., Lee, N.J., Enriquez, R.F., and Herzog, H. (2021). Lack of neuropeptide FF signalling in mice leads to reduced repetitive behavior, altered drinking behavior, and fuel type selection. FASEB J *35*, e21980. 10.1096/fj.202100703R.
70. Akash, G., Changwoo, S., Ryan, J.P., Durga, S.K., Julia, A.S., and Melissa, R.W. (2020). Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. bioRxiv, 2020.2005.2021.108886. 10.1101/2020.05.21.108886.
71. Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., Yang, Y., Luo, J., Wu, Y., Gao, Z., and Fan, J. (2019). Anterior insular cortex plays a critical role in interoceptive attention. Elife *8*. 10.7554/eLife.42265.
72. Takemoto, M., Kato, S., Kobayashi, K., and Song, W.J. (2023). Dissection of insular cortex layer 5 reveals two sublayers with opposing modulatory roles in appetitive drinking behavior. iScience *26*, 106985. 10.1016/j.isci.2023.106985.
73. Aquino-Miranda, G., Jalloul, D., Zhang, X.O., Li, S., Kirouac, G.J., Beierlein, M., and Do Monte, F.H. (2024). Functional properties of corticothalamic circuits targeting paraventricular thalamic neurons. Neuron *112*, 4060-4080 e4067. 10.1016/j.neuron.2024.10.010.
74. Ma, J., O'Malley, J.J., Kreiker, M., Leng, Y., Khan, I., Kindel, M., and Penzo, M.A. (2024). Convergent direct and indirect cortical streams shape avoidance decisions in mice via the midline thalamus. Nat Commun *15*, 6598. 10.1038/s41467-024-50941-6.

75. Herman, J.P., McKlveen, J.M., Ghosal, S., Kopp, B., Wulsin, A., Makinson, R., Scheimann, J., and Myers, B. (2016). Regulation of the Hypothalamic-Pituitary-Adrenocortical Stress Response. Compr Physiol *6*, 603-621. 10.1002/cphy.c150015.
76. Herman, J.P. (2013). Neural control of chronic stress adaptation. Front Behav Neurosci *7*, 61. 10.3389/fnbeh.2013.00061.
77. He, Y., Ren, Y., Chen, X., Wang, Y., Yu, H., Cai, J., Wang, P., Ren, Y., and Xie, P. (2024). Neural and molecular investigation into the paraventricular thalamus for chronic restraint stress induced depressive-like behaviors. J Adv Res. 10.1016/j.jare.2024.10.025.
78. Cui, L., Li, S., Wang, S., Wu, X., Liu, Y., Yu, W., Wang, Y., Tang, Y., Xia, M., and Li, B. (2024). Major depressive disorder: hypothesis, mechanism, prevention and treatment. Signal Transduct Target Ther *9*, 30. 10.1038/s41392-024-01738-y.
79. Cao, Y., Wu, Z., Zhang, M., Ji, R., Zhang, H., and Song, L. (2024). Microglial adenosine A(2A) receptor in the paraventricular thalamic nucleus regulates pain sensation and analgesic effects independent of opioid and cannabinoid receptors. Front Pharmacol *15*, 1467305. 10.3389/fphar.2024.1467305.
80. Chang, J., Li, Z., Yuan, H., Wang, X., Xu, J., Yang, P., and Qin, L. (2024). Protective role of aconitate decarboxylase 1 in neuroinflammation-induced dysfunctions of the paraventricular thalamus and sleepiness. Commun Biol *7*, 1484. 10.1038/s42003-024-07215-0.
81. Cheng, J., Ma, X., Li, C., Ullah, R., Wang, X., Long, J., Yuan, Z., Liu, S., Fu, J., Chen, Z., et al. (2022). Diet-induced inflammation in the anterior paraventricular thalamus induces compulsive sucrose-seeking. Nat Neurosci *25*, 1009-1013. 10.1038/s41593-022-01129-y.
82. Kasahara, T., Takata, A., Kato, T.M., Kubota-Sakashita, M., Sawada, T., Kakita, A., Mizukami, H., Kaneda, D., Ozawa, K., and Kato, T. (2016). Depression-like episodes in mice harboring mtDNA deletions in paraventricular thalamus. Mol Psychiatry *21*, 39-48. 10.1038/mp.2015.156.
83. Flagel, S.B. (2022). The paraventricular nucleus of the thalamus and its potential role in psychopathology. Neuropsychopharmacology *47*, 385-386. 10.1038/s41386-021-01112-4.
84. Shima, Y., Skibbe, H., Sasagawa, Y., Fujimori, N., Iwayama, Y., Isomura-Matoba, A., Yano, M., Ichikawa, T., Nikaido, I., Hattori, N., and Kato, T. (2023). Distinctiveness and continuity in transcriptome and connectivity in the anterior-posterior axis of the paraventricular nucleus of the thalamus. Cell Rep *42*, 113309. 10.1016/j.celrep.2023.113309.

FIGURE LEGENDS

Figure 1. PVT^{D2R+} neuronal activity state-dependent modulation is contingent on calorie intake.

(A) Schematic of the stereotaxic injections and optical fiber implantation for targeting D2R+ neurons in the PVT. **(B)** Representative image from a Drd2-Cre mouse expressing GCaMP6s in the PVT. **(C)** Timeline schematic for the test sessions. **(D)** Average GCaMP6s responses (n=5) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach in test sessions: H-SMS: n = 705 trials, S-SMS: n = 368 trials, and H-NCal: n = 928 trials. Repeated measures ANOVA, *p<0.05. H-SMS vs. S-SMS Dunnett's multiple comparisons test **p<0.05*. H-SMS vs. H-NCal Dunnett's multiple comparisons test *p=0.54, ns*. **(E)** Latencies to reach the reward zone across 'trial group blocks' (G1 – G5) for H-SMS and S-SMS sessions. Two-way ANOVA, Effect of trial group block ****p<0.0001*; Effect of test session (satiety) ****p<0.0001;* Interaction *p=0.46*, *ns*. H-SMS and H-NCal sessions. Two-way ANOVA, Effect of trial group block ****p<0.0001; Effect of test session *p=0.44*, *ns*; Interaction *p=0.99, ns*. **(F)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the S-SMS and the H-NCal test sessions. **(G)** Heatmap of reward approach responses from PVT^{D2R+} neurons during sated-SMS test sessions, time-locked to approach onset, sorted by trial order, and binned into 5 'trial group blocks' (G1 – G5). **(H)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the sated-SMS test sessions. Repeated measures ANOVA, **p<0.001*. G1 vs. G2 Dunnett's multiple comparisons test **p<0.05*. G1 vs. G5 Dunnett's multiple comparisons test **p<0.05*. **(I)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the sated-SMS test session. **(J)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point results for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot

shows a negative association between PVT^{D2R+} GCaMP6s responses and trial order before and after reward zone entry for the S-SMS test session. **(K)** No correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and approach latency within the S-SMS test session. **(L)** Heatmap of reward approach responses from PVT^{D2R+} neurons during H-NCal test sessions. **(M)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the H-NCal test sessions. Repeated measures ANOVA, *p=0.078, ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.43, ns*. **(N)** No correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the H-NCal test sessions. **(O)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point results for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot shows no significant association between PVT^{D2R+} GCaMP6s responses and trial order for the H-NCal test sessions. **(P)** No correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and approach latency within the H-NCal test sessions. **(Q)** Quantification of trials completed within a session comparing trials during the H-SMS, the S-SMS and the H-NCal test sessions. Repeated measures ANOVA **p<0.05.* H-SMS vs. S-SMS Dunnett's multiple comparisons test *p=0.26, ns.* H-SMS vs. H-NCal Dunnett's multiple comparisons test *p=0.06, ns.* **(R)** Schematic of the timeline for the test sessions for the second cohort of mice. **(S)** Average GCaMP6s responses (n=4) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach in test sessions: H-SMS, H-NCal, and S-NCal. H-SMS: n = 540 trials, H-NCal: n = 414 trials and S-NCal: n = 515 trials. Repeated measures ANOVA, *p=0.86, ns*. **(T)** Latencies to reach the reward zone across 'trial group blocks' (G1 – G5) for H-SMS, H-NCal, and S-NCal sessions. Two-way ANOVA, Effect of trial group block ****p<0.0001*; Effect of test session **p<0.05;* Interaction *p=0.82, ns*. **(U)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the S-NCal test sessions. Repeated measures ANOVA, *p=0.83, ns*. **(V)** Quantification of trials completed during H-SMS, H-NCal and

S-NCal test sessions. Repeated measures ANOVA, *p=0.64, ns*.

Figure 2. PVT^{D2R+} neuronal activity is modulated by the interoceptive value of the outcome.

(A) Timeline schematic for the test sessions. **(B)** Average GCaMP6s responses (n=5) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach in test sessions: H-SMS: n = 460 trials, H-Suc: n = 885 trials and H-H_2O: n = 549 trials. Repeated measures ANOVA, ***p<0.01*. H-SMS vs. H-Suc Dunnett's multiple comparisons test *p=0.12, ns.* H-SMS vs. H-H_2O Dunnett's multiple comparisons test ***p<0.01*. **(C)** Latencies to reach the reward zone across 'trial group blocks' (G1 – G5) for H-SMS, H-Suc, and H-H_2O sessions. Two-way ANOVA, Effect of trial group block **p<0.05*; Effect of test session *****p<0.0001;* Interaction **p<0.05*. **(D)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the H-Suc test sessions. **(E)** Heatmap of reward approach responses from PVT^{D2R+} neurons during H-Suc test sessions, time-locked to approach onset, sorted by trial order, and binned into 5 'trial group blocks' (G1 – G5). **(F)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the H-Suc test sessions. Repeated measures ANOVA, **p<0.05*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.05, ns.* **(G)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the H-Suc test sessions. **(H)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point results for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot shows a negative association between PVT^{D2R+} GCaMP6s responses and trial order before and after reward zone entry for the H-Suc test sessions. **(I)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the H-H_2O test sessions. **(J)** Heatmap of reward approach responses from PVT^{D2R+} neurons during H- H_2O test sessions. **(K)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the H-H_2O test sessions.

Repeated measures ANOVA, *p=0.50, ns*. **(L)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the H-H_2O test sessions. **(M)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point results for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot shows a negative association between PVT^{D2R+} GCaMP6s responses and trial order only after reward zone entry for the H-H_2O test sessions. **(N)** Quantification of trials completed during H-SMS, H-Suc, and H-H_2O test sessions. Repeated measures ANOVA, *$p<0.05$. H-SMS vs. H-Suc Dunnett's multiple comparisons test *p=0.63, ns*. H-SMS vs. H-H_2O Dunnett's multiple comparisons test *p=0.06, ns*. **(O)** Quantification of average consumption (mLs) in a 4-bottle test of SMS, 10% Suc, Non-Cal and H_2O *ad libitum* consumption. Repeated measures ANOVA, ****$p<0.0001$. **(P)** Schematic of the timeline for the test sessions. **(Q)** Average GCaMP6s responses (n=7) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach in test sessions: H-SMS, T-H_2O, and H-H_2O. H-SMS: n = 1111 trials, T-H_2O: n = 864 trials and H-H_2O: n = 762 trials. Repeated measures ANOVA, **$p<0.01$. H-SMS vs. T-H_2O Dunnett's multiple comparisons test *p=0.97, ns.* H-SMS vs. H-H_2O Dunnett's multiple comparisons test **$p<0.01$. **(R)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early ‘G1’ and late ‘G5’ trials within the T-H_2O test sessions. **(S)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the T-H_2O test sessions. Repeated measures ANOVA, **$p<0.01$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$. **(T)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the T-H_2O test sessions.

Figure 3. PVT^{D2R+} neuronal activity is state and outcome dependent.

(A) Timeline schematic for the test sessions. **(B)** Average GCaMP6s responses (n=4) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach in test sessions: H-SMS, H-OM, and S-OM. H-SMS: n = 705 trials, H-OM: n= 797 trials and S-OM: n = 253 trials from 5 mice. Repeated measures ANOVA, ***p<0.01*. H-SMS vs. H-OM Tukey's multiple comparisons test ***p<0.01*. H-SMS vs. S-OM Tukey's multiple comparisons test ***p<0.01.* H-OM vs. S-OM Tukey's multiple comparisons test *p=0.58, ns.* **(C)** Latencies to reach the reward zone across 'trial group blocks' (G1 – G5) for H-SMS, H-OM, and S-OM sessions. Two-way ANOVA, Effect of trial group block *p=0.54*, *ns.* Effect of test session ****p<0.0001;* Interaction *p=0.44, ns*. **(D)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the H-OM test sessions. **(E)** Heatmap of reward approach responses from PVT^{D2R+} neurons during H-OM test sessions, time-locked to approach onset, sorted by trial order, and binned into 5 'trial group blocks' (G1 – G5). **(F)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the H-OM test sessions. Repeated measures ANOVA, *p=0.45, ns*. **(G)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within the H-OM test sessions. **(H)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point results for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot shows a positive association between PVT^{D2R+} GCaMP6s responses and trial order before and after reward zone entry for the H-OM test sessions. **(I)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the S-OM test sessions. **(J)** Heatmap showing reward approach responses from PVT^{D2R+} neurons during S-OM test sessions. **(K)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks within the S-OM test sessions. Repeated measures ANOVA, *p=0.45*, *ns.* **(L)** No correlation between the AUC of the reward approach-evoked changes in

GCaMP6s activity and trial order within the S-OM test sessions. **(M)** AUC quantification of the reward approach across 'trial group blocks' (G1 – G5) for H-SMS, H-OM, and S-OM sessions. Two-way ANOVA, Effect of trial group block *p=0.26*; *ns.* Effect of test session ***p<0.01;* Interaction *p=0.27, ns*. **(N)** Quantification of trials completed within a session comparing trials performed during the H-SMS, the H-OM and the S-OM test sessions. Repeated measures ANOVA, ***p<0.01.* H-SMS vs. H-OM Tukey's multiple comparisons test **p<0.05*. H-SMS vs. S-OM Tukey's multiple comparisons test *p=0.60, ns*. H-OM vs. S-OM Tukey's multiple comparisons test ***p<0.01*. **(O)** Schematic illustrating the blocked reward test in the linear maze task. **(P)** *Left*: Timeline schematic for the blocked reward test sessions. *Right*: Schematic of the blocked reward test session showing that each session was divided into 2 parts in which the reward tray was blocked for the first 15 min of the session and was then unblocked for the remainder of the session. **(Q)** Average GCaMP6s responses from PVT^{D2R+} neurons (n=4) and **(R)** AUC quantification comparing ramping activity during reward approach for the H-SMS pre-test, H-OM pre-test, Blocked OM, and H-OM post-block tests sessions. Repeated measures ANOVA, ***p<0.01.* H-SMS pre-test vs. H-OM pre-test Tukey's multiple comparisons test **p<0.05*. H-SMS pre vs. Blocked OM Tukey's multiple comparisons test ***p<0.01*. H-SMS pre vs. H-OM post-block Tukey's multiple comparisons test *p=0.69, ns*. **(S)** Quantification of total trials completed for the H-SMS pre-test, H-OM pre-test, Blocked OM, and H-OM post-block tests sessions. Repeated measures ANOVA, ****p<0.0001.* H-SMS pre-test vs. H-OM pre-test Tukey's multiple comparisons test **p<0.05*. H-SMS pre-test vs. Blocked OM Tukey's multiple comparisons test ****p<0.0001*. H-SMS pre-test vs. H-OM post-block Tukey's multiple comparisons test ***p<0.001*. H-OM pre-test vs. H-OM post-block Tukey's multiple comparisons test *p=0.92, ns*. **(T)** Quantification of latency to reward zone for the H-SMS

pre-test, H-OM pre-test, Blocked OM, and H-OM post-block tests sessions. Repeated measures ANOVA, ****p<0.001.* H-SMS pre-test vs. H-OM pre-test Tukey's multiple comparisons test ***p<0.01*. H-SMS pre-test vs. Blocked OM Tukey's multiple comparisons test ****p<0.001*. H-SMS pre-test vs. H-OM post-block Tukey's multiple comparisons test *p=0.52, ns*.

Figure 4. PVT^{D2R+} neuronal activity during approach only is modulated by outcome and outcome expectation.

(A-F) Average GCaMP6s responses from PVT^{D2R+} neurons during approach only, normalized by linear interpolation comparing correct and premature trials in the test sessions: **(A)** H-SMS, **(B)** S-SMS, **(C)** H-NCal, **(D)** H-Suc**, (E)** H-H_2O**,** and **(F)** H-OM**. (G-L)** Average GCaMP6s responses from PVT^{D2R+} neurons (top), comparisons between early 'G1' and late 'G5' trials (middle) and AUC quantifications across trial groups (bottom) during reward delivery for the **(G)** H-SMS test sessions; Repeated measures ANOVA, ****$p<0.0001$. G1 vs. G5 Dunnett's multiple comparisons test **$p<0.01$, **(H)** for S-SMS test sessions; Repeated measures ANOVA, $p=0.10$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.08$, *ns*, **(I)** for H-NCal test sessions; Repeated measures ANOVA, $p=0.17$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.65$, *ns*, **(J)** for H-Suc test sessions; Repeated measures ANOVA, ***$p<0.001$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$, **(K)** for H-H_2O test sessions; Repeated measures ANOVA, $p=0.48$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.87$, *ns*, **(L)** and for H-OM test sessions; Repeated measures ANOVA, $p=0.32$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.98$, *ns*. **(M)** AUC quantifications of the reward approach only evoked changes in GCaMP6s activity across the different test sessions for correct 'C' and premature 'P' trials. Two-tailed paired t-tests. H-SMS-C vs. H-SMS-P: ***$p<0.001$. S-SMS-C vs. S-SMS-P: *$p<0.05$. H-NCal-C vs. H-NCal-P: $p=0.07$, *ns.* H-Suc-C vs. H-Suc-P: **$p<0.01$. H-H_2O-C vs. H-H_2O-P $p=0.69$, *ns.* H-OM-C vs. H-OM-P $p=0.75$, *ns.* Two-tailed unpaired t-tests. H-SMS-C vs. H-OM-C *$p<0.05$. H-SMS-P vs. H-OM-P **$p<0.01$. **(N)** AUC quantifications of the reward delivery and feeding bouts-evoked changes in GCaMP6s activity across the different test sessions. Two-tailed paired t-tests. H-SMS baseline (base) vs. H-SMS: ***$p<0.001$. Base vs. S-SMS: **$p<0.01$. Base vs. H-NCal: ****$p<0.0001$. Base vs. H-Suc: ***$p<0.001$. Base vs. H-H_2O: $p=0.50$, *ns.* Base vs. H-OM: $p=0.60$, *ns.*

Figure 5. PVT^{D2R+} neuronal activity during return is modulated by outcome and outcome expectation.

(A-F) Average GCaMP6s responses from PVT^{D2R+} neurons (top), comparisons between early 'G1' and late 'G5' trials (middle), and AUC quantifications across trial groups (bottom). These responses are time-locked to return onset, sorted by trial order, and binned into 5 trial group blocks (G1 – G5) during return to the trigger zone for **(A)** H-SMS test sessions: Repeated measures ANOVA, ****$p<0.0001$. G1 vs. G5 Dunnett's multiple comparisons test **$p<0.01$; **(B)** for S-SMS test sessions: Repeated measures ANOVA, **$p<0.01$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$; **(C)** for H-NCal test sessions: Repeated measures ANOVA, $p=0.06$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.06$, *ns*; **(D)** for H-Suc test sessions: Repeated measures ANOVA, *$p<0.05$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$; **(E)** for H-H_2O test sessions: Repeated measures ANOVA, $p=0.09$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.07$, *ns*; **(F)** for T-H_2O test sessions: Repeated measures ANOVA, **$p<0.01$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$, *ns*; **(G)** and for H-OM test sessions: Repeated measures ANOVA, $p=0.41$, *ns.* G1 vs. G5 Dunnett's multiple comparisons test $p=0.63$, *ns*. **(H)** AUC quantifications of the return-evoked changes in GCaMP6s activity across the different test sessions. Two-tailed paired t-tests. H-SMS baseline (base) vs. H-SMS: ****$p<0.0001$. Base vs. S-SMS: *$p<0.05$. Base vs. H-NCal: ****$p<0.0001$. Base vs. H-Suc: ***$p<0.001$. Base vs. H-H_2O: $p=0.08$, *ns.* Base vs. H-OM: ****$p<0.0001$. Two-tailed unpaired t-tests. H-SMS vs H-H_2O: *$p<0.05$. H-H_2O vs. H-OM: *$p<0.05$. **(I-L)** Heatmap of trials containing reward delivery and return responses from PVT^{D2R+} neurons, time-locked to reward zone arrival, sorted by duration in reward zone. Yellow dots indicate return onset for test sessions: **(I)** H-SMS (strong responses modulated by state), **(J)** S-SMS (weak responses modulated by state), **(K)** H-NCal (strong responses not modulated by state) and **(L)** H-OM (weak responses not modulated by state).

Figure 6. PVT^{D2R+} neurons detect changes in reward magnitude.

(A) *Left:* Schematic timeline for test sessions measuring changes in reward magnitude. *Right:* Schematic of the reward magnitude test session showing that each session was divided into 3 trials blocks in which mice received NCal reward of different magnitudes: small (1µl), medium (3µl), and big (6µl). **(B)** *Left:* Average GCaMP6s responses (n=4) from PVT^{D2R+} neurons during the pre-test with a non-caloric reward. *Right:* AUC quantification comparing ramping activity during reward approach for early (G1) and late (G5) trials. Two-tailed paired t-test. G1 vs. G5, *p=0.48, ns.* **(C)** *Left:* Average GCaMP6s responses from PVT^{D2R+} neurons to the different reward amounts during approach. *Right:* AUC quantification comparing ramping activity during reward approach for the small, medium, and big reward. Repeated measures ANOVA, *$p<0.05$. Small vs. Medium: Tukey's multiple comparisons test *p=0.27, ns*. Small vs. Big: Tukey's multiple comparisons test *$p<0.05$. Medium vs. Big: Tukey's multiple comparisons test *p=0.16, ns*. **(D)** Latencies to reach the reward zone for the small, medium and big rewards. Repeated measures ANOVA, *p=0.06, ns.* Small vs. Medium: Tukey's multiple comparisons test *p=0.06, ns*. Small vs. Big: Tukey's multiple comparisons test *p=0.18, ns*. Medium vs. Big: Tukey's multiple comparisons test *p=0.38, ns*. **(E)** *Left:* Average GCaMP6s responses from PVT^{D2R+} neurons during return phase of the pre-test session. *Right:* AUC quantification comparing ramping activity during reward approach for early (G1) and late (G5) trials. Two-tailed paired t-test. G1 vs. G5, *p=0.21, ns.* **(F)** *Left:* Average GCaMP6s responses from PVT^{D2R+} neurons during return. *Right:* AUC quantification showing comparisons between return activity for the small, medium, and big reward. Repeated measures ANOVA, *p=0.47, ns.* Small vs. Medium: Tukey's multiple comparisons test *p=0.59, ns*. Small vs. Big: Tukey's multiple comparisons test *p=0.96, ns*. Medium vs. Big: Tukey's multiple comparisons test *p=0.48, ns*. **(G)** Latencies for the return for the small, medium and big rewards. Repeated measures ANOVA, *$p<0.05$. Small vs. Medium:

Tukey's multiple comparisons test *$p<0.05$. Small vs. Big: Tukey's multiple comparisons test *$p<0.05$. Medium vs. Big: Tukey's multiple comparisons test $p=0.33$, *ns*.

Figure 7. PVT^{D2R+} neuronal responses during approach and return emerge from the learning of cue–outcome associations.

(A) Timeline schematic for the test sessions. **(B)** Average GCaMP6s responses (n=8) from PVT^{D2R+} neurons and AUC quantification comparing ramping activity during reward approach for the first 'S1', second 'S2', third 'S3' and fifth 'S5' training session in our linear maze task. Repeated measures ANOVA, ****p<0.001.* S1 vs. S3 Tukey's multiple comparisons test **p<0.05*. S1 vs. S5 Tukey's multiple comparisons test ****p<0.001*. S3 vs. S5 Tukey's multiple comparisons test *p=0.60, ns*. **(C)** Latencies to reach the reward zone for training sessions 'S1', 'S3' and 'S5'. Repeated measures ANOVA, ***p<0.01.* S1 vs. S3 Tukey's multiple comparisons test *p=0.06, ns*. S1 vs. S5 Tukey's multiple comparisons test **p<0.05*. S3 vs. S5 Tukey's multiple comparisons test *p=0.08, ns*. **(D)** Correct approach trials performed during training sessions 'S1', 'S3' and 'S5'. Repeated measures ANOVA, **p<0.05.* S1 vs. S3 Tukey's multiple comparisons test *p=0.25, ns*. S1 vs. S5 Tukey's multiple comparisons test **p<0.05*. S3 vs. S5 Tukey's multiple comparisons test **p<0.05.* **(E)** Average GCaMP6s responses from PVT^{D2R+} neurons and AUC quantification during return to trigger zone for the first 'S1', third 'S3', and fifth 'S5' training session. Repeated measures ANOVA, **p<0.05.* S1 vs. S3 Tukey's multiple comparisons test *p=0.54, ns.* S1 vs. S5 Tukey's multiple comparisons test **p<0.01*. S3 vs. S5 Tukey's multiple comparisons test *p=0.14, ns*. **(F)** Latencies to return to the trigger zone for training sessions 'S1', 'S3', and 'S5'. Repeated measures ANOVA, **p<0.05.* S1 vs. S3 Tukey's multiple comparisons test *p=0.15, ns*. S1 vs. S5 Tukey's multiple comparisons test **p<0.05*. S3 vs. S5 Tukey's multiple comparisons test **p<0.05*. **(G)** Total return trials performed during training sessions 'S1', 'S3', and 'S5'. Repeated measures ANOVA, **p<0.05.* S1 vs. S3 Tukey's multiple comparisons test *p=0.25, ns*. S1 vs. S5 Tukey's multiple comparisons test **p<0.05*. S3 vs. S5 Tukey's multiple comparisons test **p<0.05.* **(H)** Average GCaMP6s responses and AUC quantification from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late

‘G5’ trials for the first training session. Two-tailed paired t-test. G1 vs. G5, *p=0.31, ns.* **(I)** Average GCaMP6s responses and AUC quantification from PVT^{D2R+} neurons comparing reward approach between early ‘G1’ and late ‘G5’ trials for the second training session. Two-tailed paired t-test. G1 vs. G5, *p=0.91, ns.* **(J)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing return to trigger zone between early ‘G1’ and late ‘G5’ trials for the first training session. Two-tailed paired t-test. G1 vs. G5, *p=0.05, ns.* **(K)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing return to trigger zone between early ‘G1’ and late ‘G5’ trials for the second training session. Two-tailed paired t-tests. G1 vs. G5, ***p<0.01.*

Figure 8. Chronic PVT^{D2R+} neuronal inhibition impairs performance in the linear maze task.

(A) *Left*: Schematic of the stereotaxic injections for targeting D2R+ neurons in the PVT. *Right*: Representative image from a Drd2-Cre mouse expressing hM4Di-mcherry in the PVT. **(B)** Diagram showing the timeline for the chemogenetic inhibition of PVT^{D2R+} neurons. **(C)** Quantification of correct trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group ****$p<0.0001$. Interaction $p=0.47$, *ns*. Šídák's multiple comparisons test: (S4) control vs. hM4Di, *$p<0.05$; (S5) control vs. hM4Di, *$p<0.05$; (S6) control vs. hM4Di, *$p<0.05$. **(D)** Quantification of latency to approach during correct trials for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group ****$p<0.0001$; Interaction $p=0.28$, *ns*. Šídák's multiple comparisons test: (S2) control vs. hM4Di, *$p<0.05$; (S3) control vs. hM4Di, **$p<0.01$. **(E)** Quantification of premature trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group *$p<0.05$. Interaction $p=0.99$, *ns*. **(F)** Quantification of latency to approach during premature trials for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group *$p<0.05$. Interaction $p=0.29$, *ns*.

Supplementary Figure 1. PVT^{D2R+} neurons sense and track hunger states.

(A) Schematic depicting the linear maze reward-seeking task. **(B)** *Left:* Average GCaMP6s responses from PVT^{D2R+} neurons showing ramping activity during reward approach. *Right:* Quantification of the approach-evoked changes in GCaMP6s fluorescence in PVT^{D2R+} neurons. Area under the curve (AUC), n = 705 trials from 5 mice. Comparisons between baseline (base) and approach (H-SMS); two-tailed paired t-test, ****$p<0.0001$. **(C)** Heatmap showing excitatory reward approach responses from PVT^{D2R+} neurons, time-locked to approach onset, sorted by trial order, and binned into 5 'trial group blocks' (G1 – G5). **(D)** Latencies to reach the reward zone across 'trial group blocks' (G1 – G5). Repeated measures ANOVA, **$p<0.01$; G1 vs. G5 Dunnett's multiple comparisons test $p=0.98$, *ns.* **(E)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the test sessions. **(F)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks. Repeated measures ANOVA, ***$p<0.001$. G1 vs. G5 Dunnett's multiple comparisons test *$p<0.05$. **(G)** Correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and trial order within a test session. **(H)** FLMM coefficient estimates plot of the approach trial order effect and statistical significance at each trial time-point for the photometric responses of PVT^{D2R+} neurons for reward zone entry. The plot shows a negative association between PVT^{D2R+} GCaMP6s responses and trial order before and after reward zone entry. **(I)** No correlation between the AUC of the reward approach-evoked changes in GCaMP6s activity and approach latency within a test session.

Supplementary Figure 2. Long-term exposure to the non-caloric reward did not affect the magnitude of PVT^{D2R+} neurons' photometric responses.

(A) Timeline schematic for non-caloric reward exposure and test sessions. **(B)** Average GCaMP6s responses (n=5) from PVT^{D2R+} neurons for the non-caloric reward after two weeks of exposure. **(C)** Average GCaMP6s responses from PVT^{D2R+} neurons comparing reward approach between early 'G1' and late 'G5' trials within the test sessions. **(D)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks. Repeated measures ANOVA, *p=0.39*, *ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.50, ns*.

Supplementary Figure 3. Premature trials are associated with reduced photometric responses of PVT^{D2R+} neurons during reward approach.

(A) Average GCaMP6s responses (n=5) from PVT^{D2R+} neurons and AUC quantification showing comparisons between ramping activity during reward approach in correct and premature trials in the H-SMS test sessions. Two-tailed paired t-test ***p<0.01*. **(B)** Latencies to reach the reward zone for correct and premature trials performed in the H-SMS test sessions. Two-tailed paired t-test *p=0.60, ns*. **(C)** AUC quantification of the reward approach-evoked changes in GCaMP6s activity across trial group blocks for the H-SMS premature trials. Repeated measures ANOVA, ****p<0.001*. G1 vs. G5 Dunnett's multiple comparisons test ***p<0.01.* **(D)** AUC quantification of the reward approach across 'trial group blocks' (G1 – G5) for H-SMS-C (correct) and H-SMS-P (premature) trials. Two-way ANOVA, Effect of trial group block *****p<0.0001*. Effect of trial type *****p<0.0001;* Interaction *p=0.61, ns*.

Supplementary Figure 4. Concurrent FLMM model of between conditions comparisons.

(A-E) FLMM coefficient estimates plots and statistical significance at each trial time-point for the mean changes in photometric responses of PVT^{D2R+} neurons during approach and reward delivery for the between outcomes comparisons (i.e., H-SMS vs. 'other outcome'). Each plot shows the four beta coefficients that result from these comparisons. *Top left* - $\beta_0(s)$: shows the mean signal when the mouse received the 'other outcome'. *Top right* - $\beta_1(s)$: shows the mean difference in signal between H-SMS and the 'other outcome' during reward delivery. *Bottom left* - $\beta_2(s)$: shows the mean difference between approach and outcome delivery for the 'other outcome'. *Bottom right* - $\beta_3(s)$: shows the difference of differences, capturing the additional effect on signal when the mouse is approaching the 'other outcome' that is not already captured by $\beta_0(s)+\beta_1(s)+\beta_2(s)$. As such, $\beta_3(s)$ captures the interaction effect between behavioral state (approaching vs. reward delivery) and outcome condition (H-SMS vs. 'other outcome').

(A) Beta coefficients for the comparisons of the H-SMS and S-SMS conditions. $\beta_0(s)$: shows that during reward delivery in the S-SMS condition, the mean signal is not significantly modulated. $\beta_1(s)$: shows that during reward delivery, the mean signal is significantly higher for H-SMS compared to S-SMS conditions. $\beta_2(s)$: shows that the mean signal during approach is significantly lower than the mean signal during reward delivery for the S-SMS condition. $\beta_3(s)$: shows a significant negative interaction effect between behavioral states (approaching vs. receiving) and conditions (H-SMS vs. S-SMS), showing that the mean changes in photometric responses are larger during the outcome delivery than during approach. Specifically, there are larger mean changes during H-SMS than in the S-SMS conditions.

(B) Beta coefficients for the comparisons of the H-SMS and H-NCal conditions. $\beta_0(s)$: shows that during reward delivery in the H-NCal condition, the mean signal is significantly positively modulated by the NCal rewards. $\beta_1(s)$: shows that during reward delivery, the mean signal is

significantly lower for H-SMS compared to H-NCal conditions. $\beta_2(s)$: shows that the mean signal during approach is significantly lower than the signal during reward delivery for the H-NCal condition. $\beta_3(s)$: shows a significant negative interaction early in the trial, followed by positive interaction effects later in the trial and during reward delivery. This indicates that the differences in neural activity during the early approach are significantly lower, but later in the trial, and during reward consumption, they are significantly higher in the NCal than in the H-SMS conditions.

(C) Beta coefficients for the comparisons of the H-SMS and H-Suc conditions. $\beta_0(s)$: shows that during reward delivery in the H-Suc condition, the mean signal is positively modulated by the Suc rewards, however, this was not significant. $\beta_1(s)$: shows that during reward delivery, the mean signal for H-SMS is not significantly different than the signal during the H-Suc conditions. $\beta_2(s)$: shows that the mean signal during approach is significantly lower than the signal during reward delivery for the H-Suc condition. $\beta_3(s)$: shows no significant interaction effects between behavioral states (approaching vs. receiving) and conditions (H-SMS vs. H-Suc).

(D) Beta coefficients for the comparisons of the H-SMS and H-H_2O conditions. $\beta_0(s)$: shows that during reward delivery in the H-H_2O condition, the mean signal is not significantly modulated. $\beta_1(s)$: shows that during reward delivery, the mean signal is significantly higher for H-SMS compared to H-H_2O conditions. $\beta_2(s)$: shows that the mean signal during approach is significantly lower than the signal during outcome delivery for the H-H_2O condition. $\beta_3(s)$: shows a significant negative interaction effect between behavioral states (approaching vs. receiving) and conditions (H-SMS vs. H-H_2O), showing that the mean changes in photometric responses are larger during the outcome delivery than during approach. Specifically, there are larger mean changes during H-SMS than in the H-H_2O conditions.

(E) Beta coefficients for the comparisons of the H-SMS and H-OM conditions. $\beta_0(s)$*:* shows that during reward delivery in the H-OM condition, the mean signal is positively modulated initially by the omission of the rewards, but the signal quickly goes down over time. $\beta_1(s)$*:* shows that during reward delivery, the mean signal is significantly higher for H-SMS compared to H-OM conditions. $\beta_2(s)$*:* shows that the mean signal during approach is significantly lower than the signal during outcome delivery for the H-OM condition. $\beta_3(s)$*:* shows a significant negative interaction effects in the trial and during reward delivery.

Supplementary Figure 5. Time spent in the reward zone after reward delivery does not differ across outcome groups or trial groups.

(A-F) Quantifications of time spent in the reward zone after reward delivery and prior to return across trial groups for the **(A)** H-SMS test sessions; Repeated measures ANOVA, *p=0.07, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.52, ns*. **(B)** S-SMS test sessions; Repeated measures ANOVA, *p=0.11, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.07, ns*. **(C)** H-NCal test sessions; Repeated measures ANOVA, *p=0.16, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.77, ns*. **(D)** H-Suc test sessions; Repeated measures ANOVA, *p=0.44, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.95, ns*. **(E)** H-H_2O test sessions; Repeated measures ANOVA, *p=0.11, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.06, ns*. **(F)** H-OM test sessions; Repeated measures ANOVA, *p=0.51, ns.* G1 vs. G5 Dunnett's multiple comparisons test *p=0.99, ns*. **(G)** Quantifications of time spent in the reward zone after reward delivery and prior to return across outcome groups. Repeated measures ANOVA, *p=0.26, ns.* H-SMS vs. S-SMS Dunnett's multiple comparisons test *p=0.99, ns*. H-SMS vs. H-NCal Dunnett's multiple comparisons test *p=0.63, ns*. H-SMS vs. H-Suc Dunnett's multiple comparisons test *p=0.99, ns*. H-SMS vs. H-H_2O Dunnett's multiple comparisons test *p=0.32, ns*. H-SMS vs. H-OM Dunnett's multiple comparisons test *p=0.49, ns*.

Supplementary Figure 6. GABAergic signaling dynamics in PVT neurons during performance in the linear maze task.

(A) Schematic of the stereotaxic injections and optical fiber implantation for targeting PVT and PVT^{D2R+} neurons. **(B)** Representative image from a C57BL/6 mouse expressing the GABA sensor iGABASnFR2 in the PVT. **(C)** Timeline schematic for the test sessions. **(D-G)** Average iGABASnFR2 responses from PVT neurons and PVT^{D2R+} neurons during reward approach and during return in H-SMS test sessions. PVT neuronal imaging: 1474 trials from 5 mice. PVT^{D2R+} neuronal imaging: 1283 trials from 5 mice. **(H-K)** Average iGABASnFR2 responses from PVT and PVT^{D2R+} neurons comparing reward approach and return between early 'G1' and late 'G5' trials within the H-SMS test sessions. **(L-O)** AUC quantification of the reward approach-evoked changes and return-evoked changes in iGABASnFR2 responses from PVT neurons and PVT^{D2R+} neurons activity across trial group blocks. **(L)** Repeated measures ANOVA, *p=0.30*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.22, ns*. **(M)** Repeated measures ANOVA, *p=0.66*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.63, ns*. **(N)** Repeated measures ANOVA, *p=0.22, ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.23, ns*. **(O)** Repeated measures ANOVA, *p=0.18, ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.57, ns*.

Supplementary Figure 7. PVT^{D2R+} photometric responses increased during the decision to return, and these increases were modulated by state and outcome expectation.

(A) Representative single trace of the photometric responses of PVT^{D2R+} neurons at the different stages of the trial. **(B)** Quantification of the maximum peak responses during decision to return (1 second before return) -evoked PVT^{D2R+} GCaMP6s activity for sessions H-SMS (strong responses modulated by state), S-SMS (weak responses modulated by state), H-NCal (strong responses not modulated by state), and H-OM (weak responses not modulated by state). Repeated measures ANOVA, ***p<0.01*. H-SMS vs. S-SMS Dunnett's multiple comparisons test *p=0.31, ns*. H-SMS vs. H-NCal Dunnett's multiple comparisons test *p=0.99, ns*. H-SMS vs. H-OM Dunnett's multiple comparisons test ***p<0.01*. **(C-F)** Quantification of the maximum peak responses during decision to return-evoked PVT^{D2R+} GCaMP6s activity across trial group block for the H-SMS, S-SMS, H-NCal, and H-OM test sessions. **(C)** H-SMS: Repeated measures ANOVA, ****p<0.001*. G1 vs. G5 Dunnett's multiple comparisons test **p<0.05*. **(D)** S-SMS: Repeated measures ANOVA, ****p<0.001*. G1 vs. G5 Dunnett's multiple comparisons test ***p<0.01*. **(E)** H-NCal: Repeated measures ANOVA, *p=0.11, ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.86, ns*. **(F)** H-OM: Repeated measures ANOVA, *p=0.69, ns*. G1 vs. G5 Dunnett's multiple comparisons test *p=0.99, ns*. **(G-J)** Correlations between the maximum peak in signal 1 sec before return and the trial order for H-SMS, S-SMS, H-NCal, and H-OM test sessions.

Supplementary Figure 8. Motivational states are associated with PVT^{D2R+} neuronal responses across different conditions.

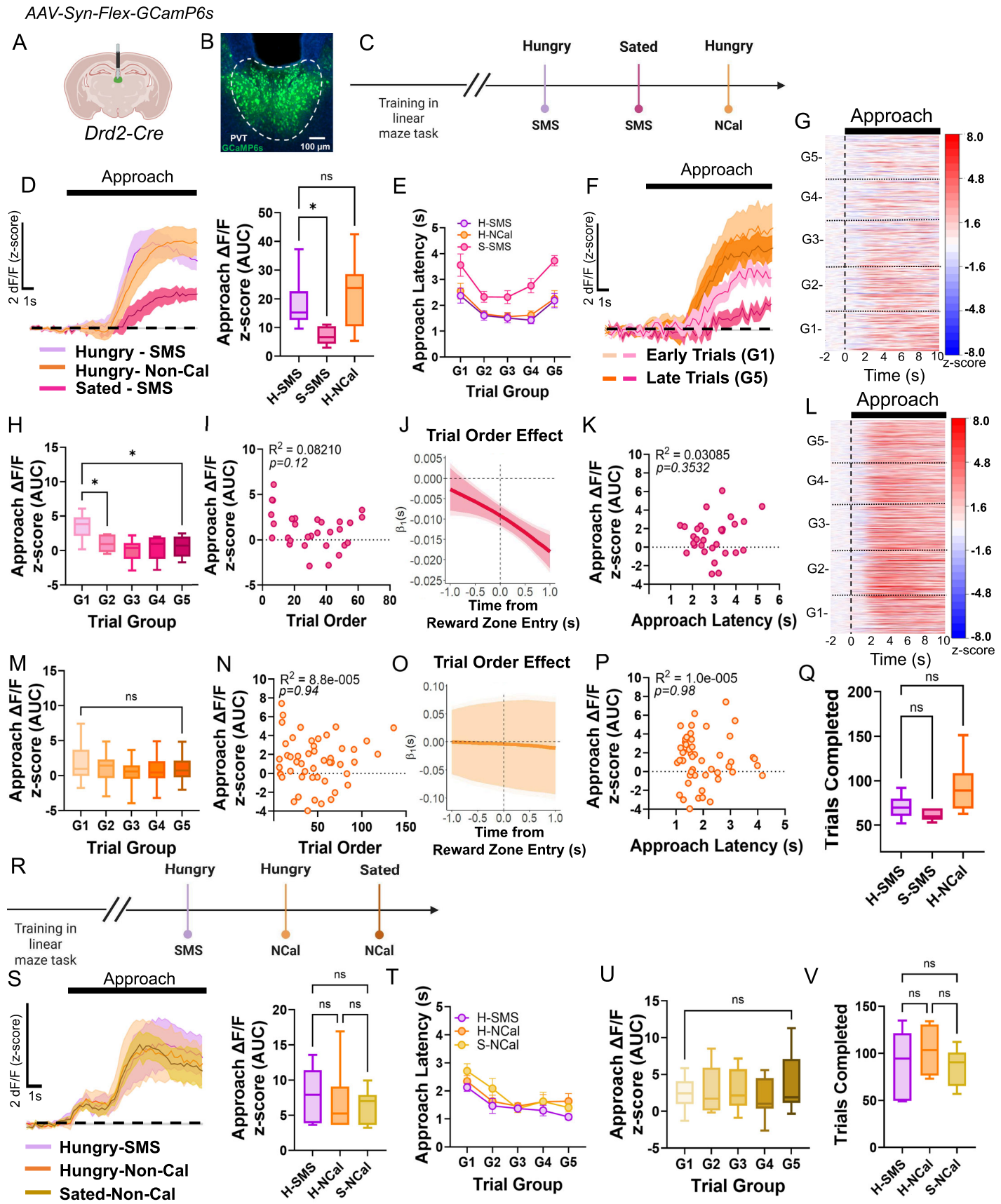
(A-G) Average GCaMP6s responses from PVT^{D2R+} neurons comparing between fast 'L1' and slow 'L5' trials (top). AUC quantifications across latency groups during approach (middle). Correlations between the approach slope and the latency to approach (bottom) for the **(A)** H-SMS test sessions; Repeated measures ANOVA, ***p<0.01.* L1 vs. L5 Dunnett's multiple comparisons test ***p<0.01*. **(B)** S-SMS test sessions; Repeated measures ANOVA, *p=0.06, ns.* L1 vs. L5 Dunnett's multiple comparisons test ***p<0.01*. **(C)** H-NCal test sessions; Repeated measures ANOVA, **p<0.05.* L1 vs. L5 Dunnett's multiple comparisons test **p<0.05*. **(D)** H-Suc test sessions; Repeated measures ANOVA, ***p<0.01.* L1 vs. L5 Dunnett's multiple comparisons test **p<0.05*. **(E)** H-H_2O test sessions; Repeated measures ANOVA, *p=0.62, ns.* L1 vs. L5 Dunnett's multiple comparisons test *p=0.95, ns*. **(F)** H-OM test sessions; Repeated measures ANOVA, ***p<0.01.* L1 vs. L5 Dunnett's multiple comparisons test ****p<0.001*. **(G)** S-OM test sessions; Repeated measures ANOVA, *p=0.57, ns.* L1 vs. L5 Dunnett's multiple comparisons test *p=0.99, ns*.

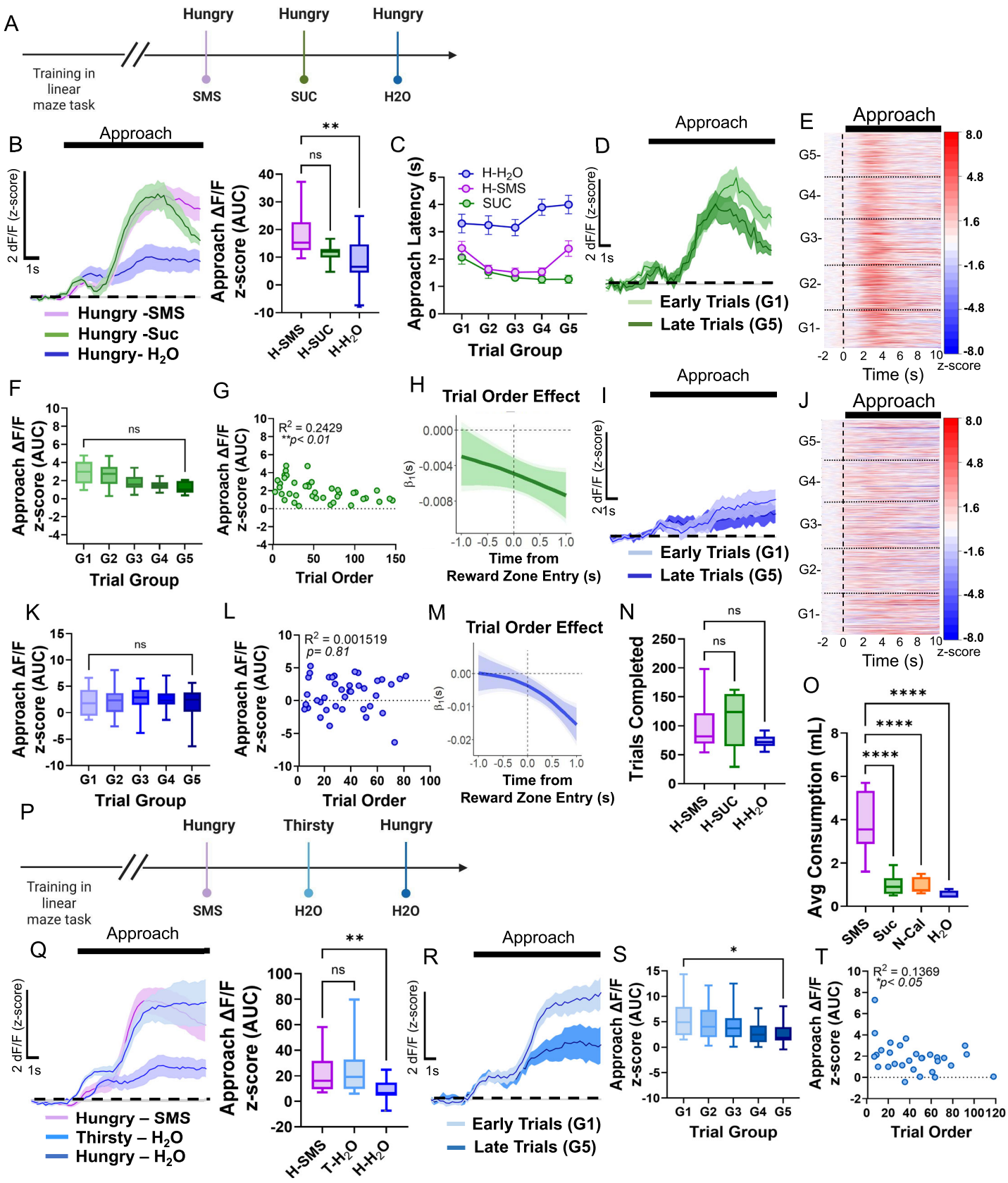
Supplementary Figure 9. PVT^{D2R+} responses in individual animals during approach and return over learning.

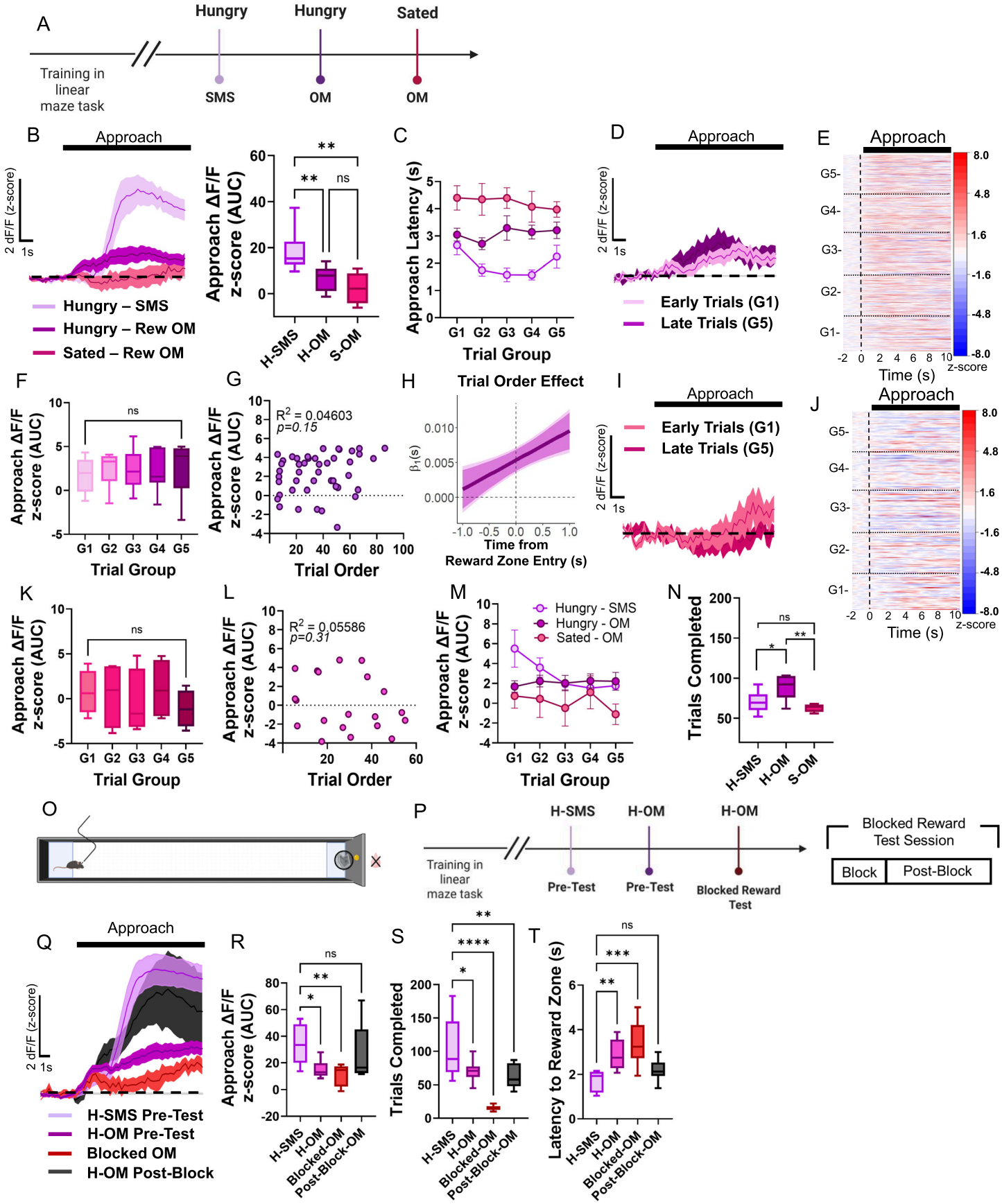
(A) Mean GCaMP6s responses of PVT^{D2R+} neurons during approach, per animal, across learning. **(B)** Mean GCaMP6s responses of PVT^{D2R+} neurons during return, per animal, across learning.

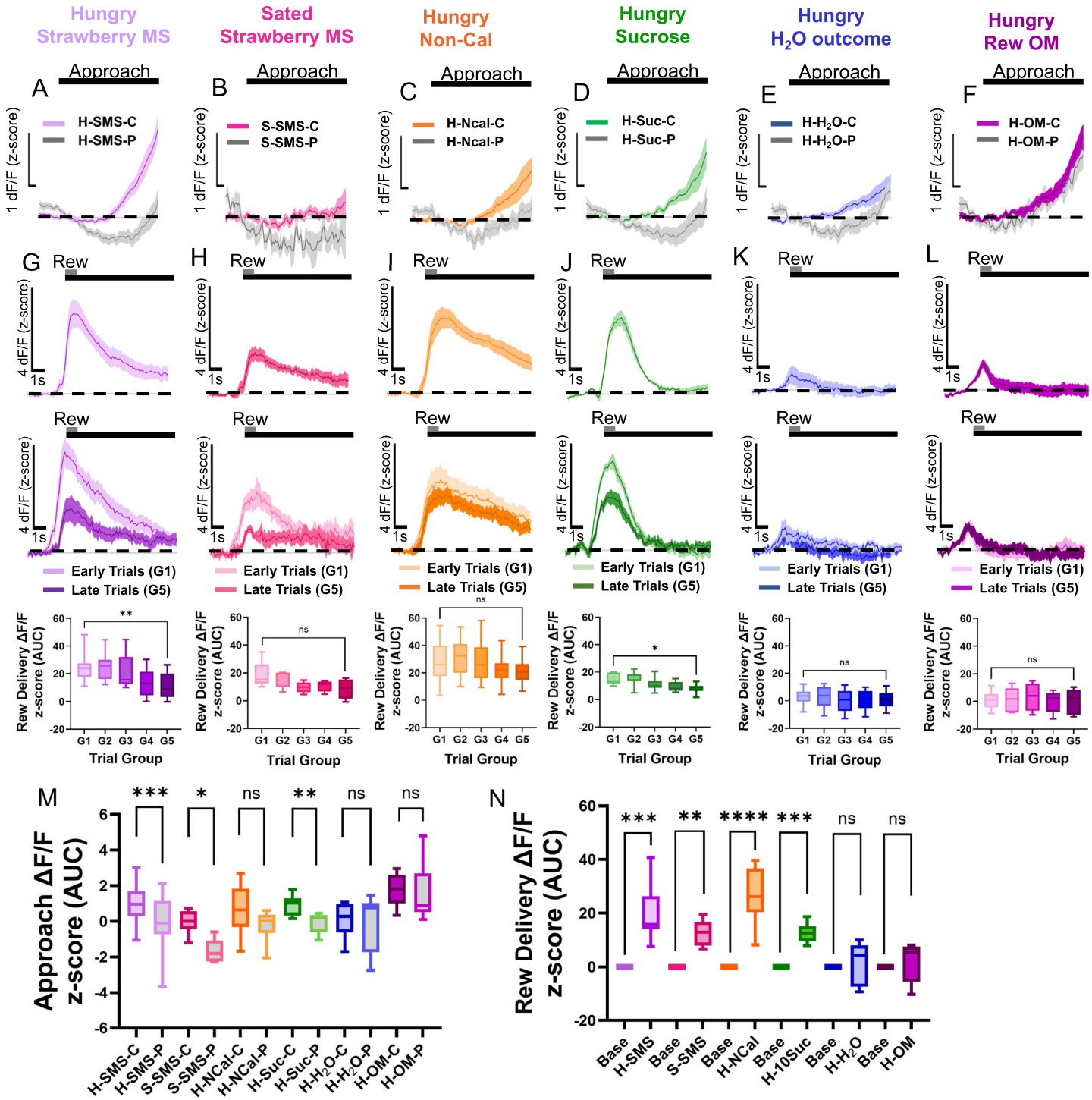
Supplementary Figure 10. Chronic, but not acute, chemogenetic inhibition of PVT^{D2R+} neurons via CNO administration results in impaired performance on the linear maze task.

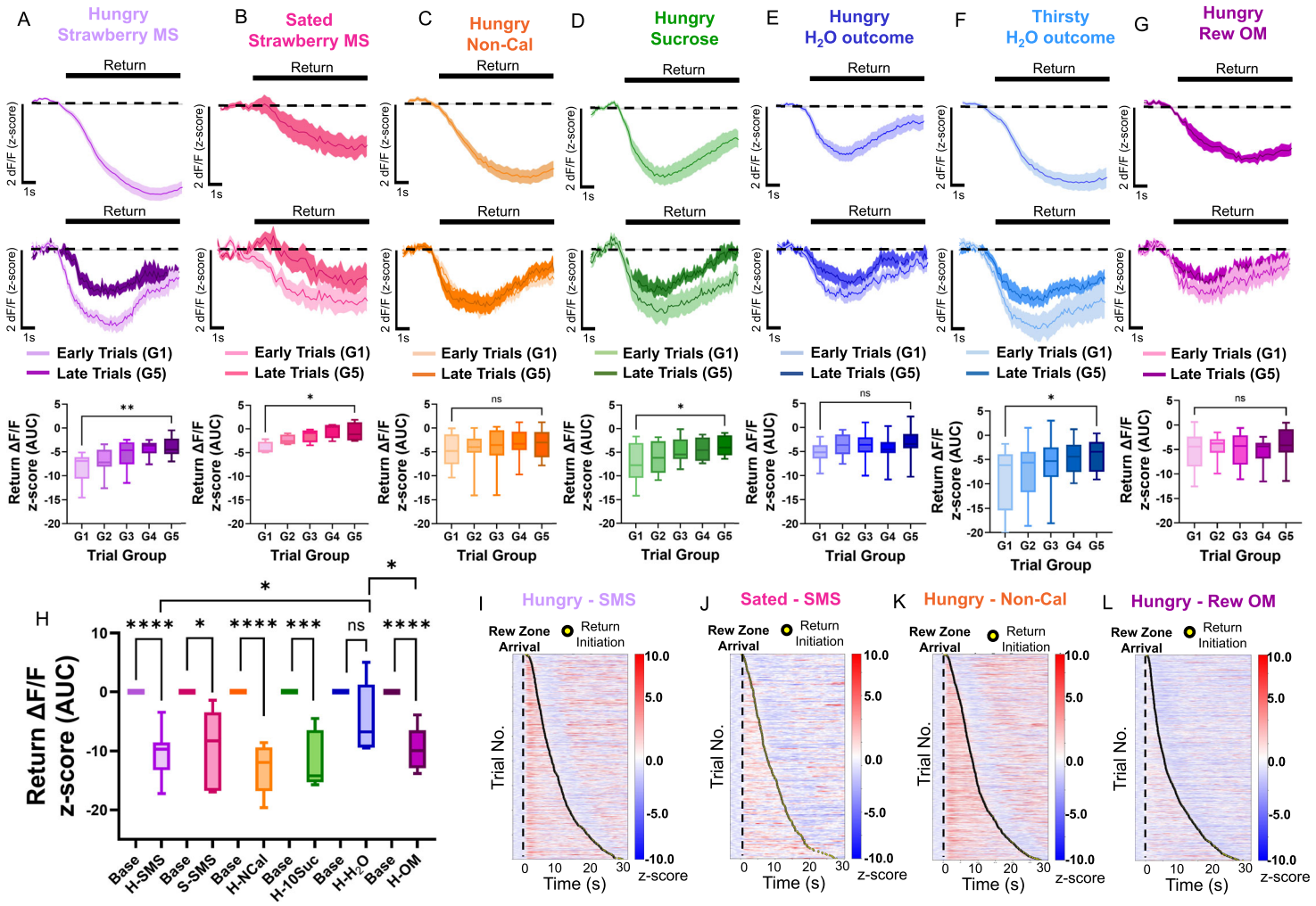
(A) *Left:* Schematic of the stereotaxic injections for targeting D2R+ neurons in the PVT. *Right:* Representative image from a Drd2-Cre mouse expressing hM4Di-mcherry in the PVT. **(B)** Schematic of the timeline for the chemogenetic chronic inhibition of PVT^{D2R+} neurons with CNO**.** **(C)** Quantification of correct trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group **$p<0.01$. Interaction ***$p<0.001$. *Šídák's multiple comparisons test: (S5) control vs. hM4Di, *$p<0.05$; (S6) control vs. hM4Di, ***$p<0.001$.* **(D)** Quantification of the latency to approach during correct trials for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group ***$p<0.001$; Interaction *p=0.50, ns.* Šídák's multiple comparisons test: (S3) control vs. hM4Di, *$p<0.05$. **(E)** Quantification of premature trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session ****$p<0.0001$. Effect of viral group *p=0.91,ns*. Interaction *p=0.86, ns*. **(F)** Quantification of latency to approach during premature trials for hM4Di and Controls groups. Two-way ANOVA, Effect of test session *$p<0.05$. Effect of viral group **$p<0.01$. Interaction *p=0.24*, ns. Šídák's multiple comparisons test: (S2) control vs. hM4Di, **$p<0.01$. **(G)** Schematic of the stereotaxic injections for targeting D2R+ neurons in the PVT. **(H)** Schematic of the timeline for the chemogenetic acute inhibition of PVT^{D2R+} neurons with CNO**. (I)** Quantification of correct trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session *p=0.07, ns*. Effect of viral group *p=0.73, ns*. Interaction *p=0.88, ns*. **(J)** Quantification of premature trials completed for hM4Di and Controls groups. Two-way ANOVA, Effect of test session *p=0.05, ns.* Effect of viral group *p=0.34, ns.* Interaction *p=0.98, ns.*


AAV-Syn-Flex-GCamP6s
A
Drd2-Cre
B
PVT
GCaMP6s
100 µm
C
Hungry
Sated
Hungry
Training in linear maze task
SMS
SMS
NCal
D
Approach
2 dF/F (z-score)
1s
Hungry - SMS
Hungry- Non-Cal
Sated - SMS
Approach ΔF/F z-score (AUC)
ns
*
H-SMS
S-SMS
H-NCal
E
Approach Latency (s)
H-SMS
H-NCal
S-SMS
G1 G2 G3 G4 G5
Trial Group
F
Approach
2 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
G
Approach
G5-
G4-
G3-
G2-
G1-
8.0
4.8
1.6
-1.6
-4.8
-8.0
z-score
Time (s)
H
Approach ΔF/F z-score (AUC)
G1 G2 G3 G4 G5
Trial Group
I
R2 = 0.08210
p=0.12
Approach ΔF/F z-score (AUC)
Trial Order
J
Trial Order Effect
β1(s)
Time from Reward Zone Entry (s)
K
R2 = 0.03085
p=0.3532
Approach ΔF/F z-score (AUC)
Approach Latency (s)
L
Approach
G5-
G4-
G3-
G2-
G1-
8.0
4.8
1.6
-1.6
-4.8
-8.0
z-score
Time (s)
M
Approach ΔF/F z-score (AUC)
ns
G1 G2 G3 G4 G5
Trial Group
N
R2 = 8.8e-005
p=0.94
Approach ΔF/F z-score (AUC)
Trial Order
O
Trial Order Effect
β1(s)
Time from Reward Zone Entry (s)
P
R2 = 1.0e-005
p=0.98
Approach ΔF/F z-score (AUC)
Approach Latency (s)
Q
ns
ns
Trials Completed
H-SMS
S-SMS
H-NCal
R
Hungry
Hungry
Sated
Training in linear maze task
SMS
NCal
NCal
S
Approach
2 dF/F (z-score)
1s
Hungry-SMS
Hungry-Non-Cal
Sated-Non-Cal
Approach ΔF/F z-score (AUC)
ns
ns
ns
H-SMS
H-NCal
S-NCal
T
Approach Latency (s)
H-SMS
H-NCal
S-NCal
G1 G2 G3 G4 G5
Trial Group
U
Approach ΔF/F z-score (AUC)
ns
G1 G2 G3 G4 G5
Trial Group
V
ns
ns
ns
Trials Completed
H-SMS
H-NCal
S-NCal


A
Hungry
Hungry
Hungry
Training in linear maze task
SMS
SUC
H2O
B
Approach
2 dF/F (z-score)
1s
Hungry -SMS
Hungry -Suc
Hungry- H2O
Approach ΔF/F z-score (AUC)
ns
**
H-SMS
H-SUC
H-H2O
C
Approach Latency (s)
H-H2O
H-SMS
SUC
G1 G2 G3 G4 G5
Trial Group
D
Approach
2 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
E
Approach
G5-
G4-
G3-
G2-
G1-
Time (s)
z-score
F
Approach ΔF/F z-score (AUC)
ns
G1 G2 G3 G4 G5
Trial Group
G
Approach ΔF/F z-score (AUC)
R2 = 0.2429
**p< 0.01
Trial Order
H
Trial Order Effect
β1(s)
Time from Reward Zone Entry (s)
I
Approach
2 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
J
Approach
G5-
G4-
G3-
G2-
G1-
Time (s)
z-score
K
Approach ΔF/F z-score (AUC)
ns
G1 G2 G3 G4 G5
Trial Group
L
Approach ΔF/F z-score (AUC)
R2 = 0.001519
p= 0.81
Trial Order
M
Trial Order Effect
β1(s)
Time from Reward Zone Entry (s)
N
Trials Completed
ns
ns
H-SMS
H-SUC
H-H2O
O
Avg Consumption (mL)
****
****
****
SMS
Suc
N-Cal
H2O
P
Hungry
Thirsty
Hungry
Training in linear maze task
SMS
H2O
H2O
Q
Approach
2 dF/F (z-score)
1s
Hungry – SMS
Thirsty – H2O
Hungry – H2O
Approach ΔF/F z-score (AUC)
ns
**
H-SMS
T-H2O
H-H2O
R
Approach
2 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
S
Approach ΔF/F z-score (AUC)
*
G1 G2 G3 G4 G5
Trial Group
T
Approach ΔF/F z-score (AUC)
R2 = 0.1369
*p< 0.05
Trial Order

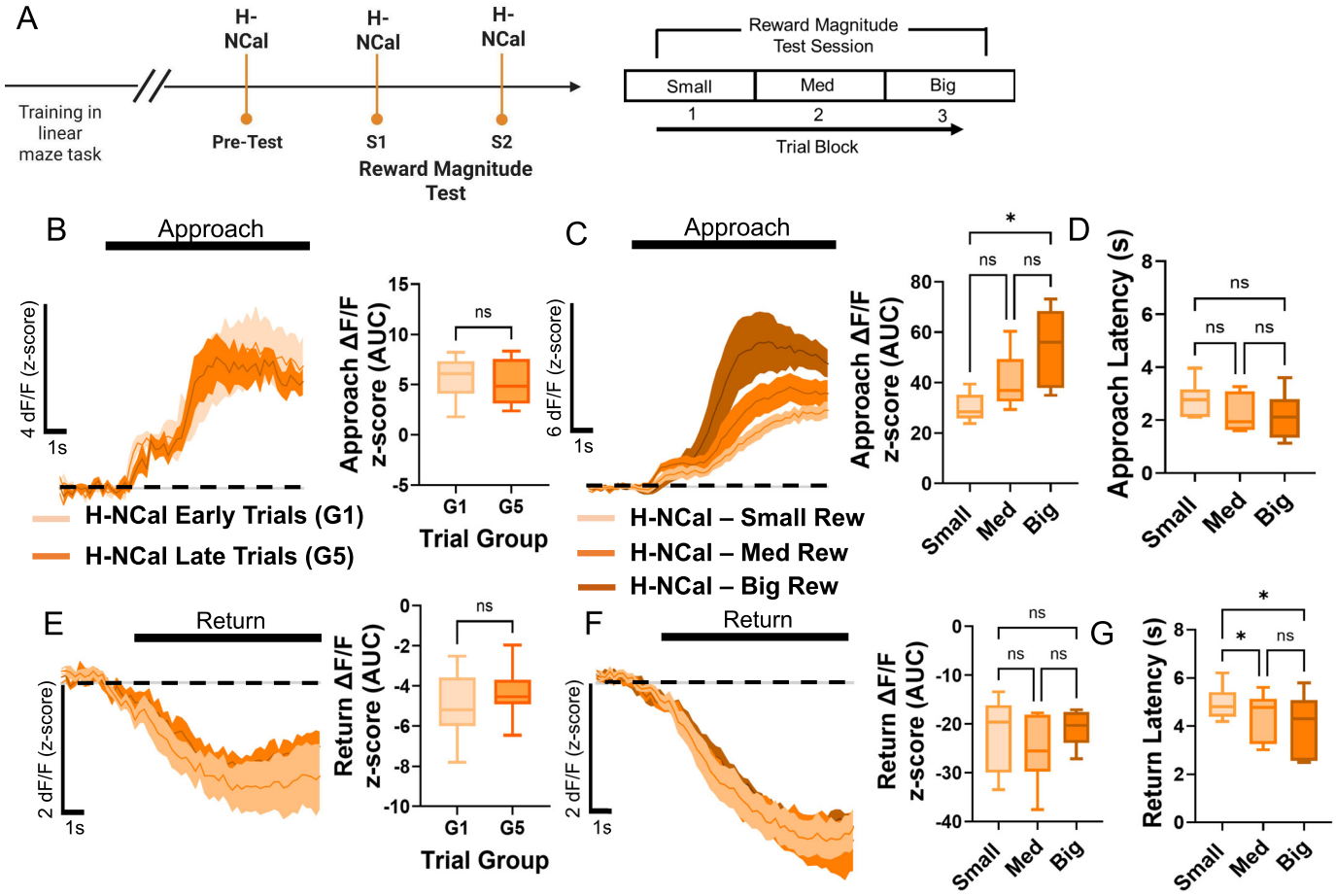

A
Hungry
Hungry
Sated
Training in linear maze task
SMS
OM
OM
B
Approach
2 dF/F (z-score)
1s
Hungry – SMS
Hungry – Rew OM
Sated – Rew OM
Approach ΔF/F z-score (AUC)
H-SMS
H-OM
S-OM
C
Approach Latency (s)
Trial Group
D
Approach
Early Trials (G1)
Late Trials (G5)
E
Approach
z-score
Time (s)
F
Approach ΔF/F z-score (AUC)
Trial Group
G
R2 = 0.04603
p=0.15
Trial Order
H
Trial Order Effect
Time from Reward Zone Entry (s)
I
Approach
Early Trials (G1)
Late Trials (G5)
J
Approach
K
Trial Group
L
R2 = 0.05586
p=0.31
Trial Order
M
Hungry - SMS
Hungry - OM
Sated - OM
Trial Group
N
Trials Completed
O
P
H-SMS
H-OM
H-OM
Training in linear maze task
Pre-Test
Pre-Test
Blocked Reward Test
Blocked Reward Test Session
Block
Post-Block
Q
Approach
H-SMS Pre-Test
H-OM Pre-Test
Blocked OM
H-OM Post-Block
R
Approach ΔF/F z-score (AUC)
Blocked-OM
Post-Block-OM
S
Trials Completed
T
Latency to Reward Zone (s)

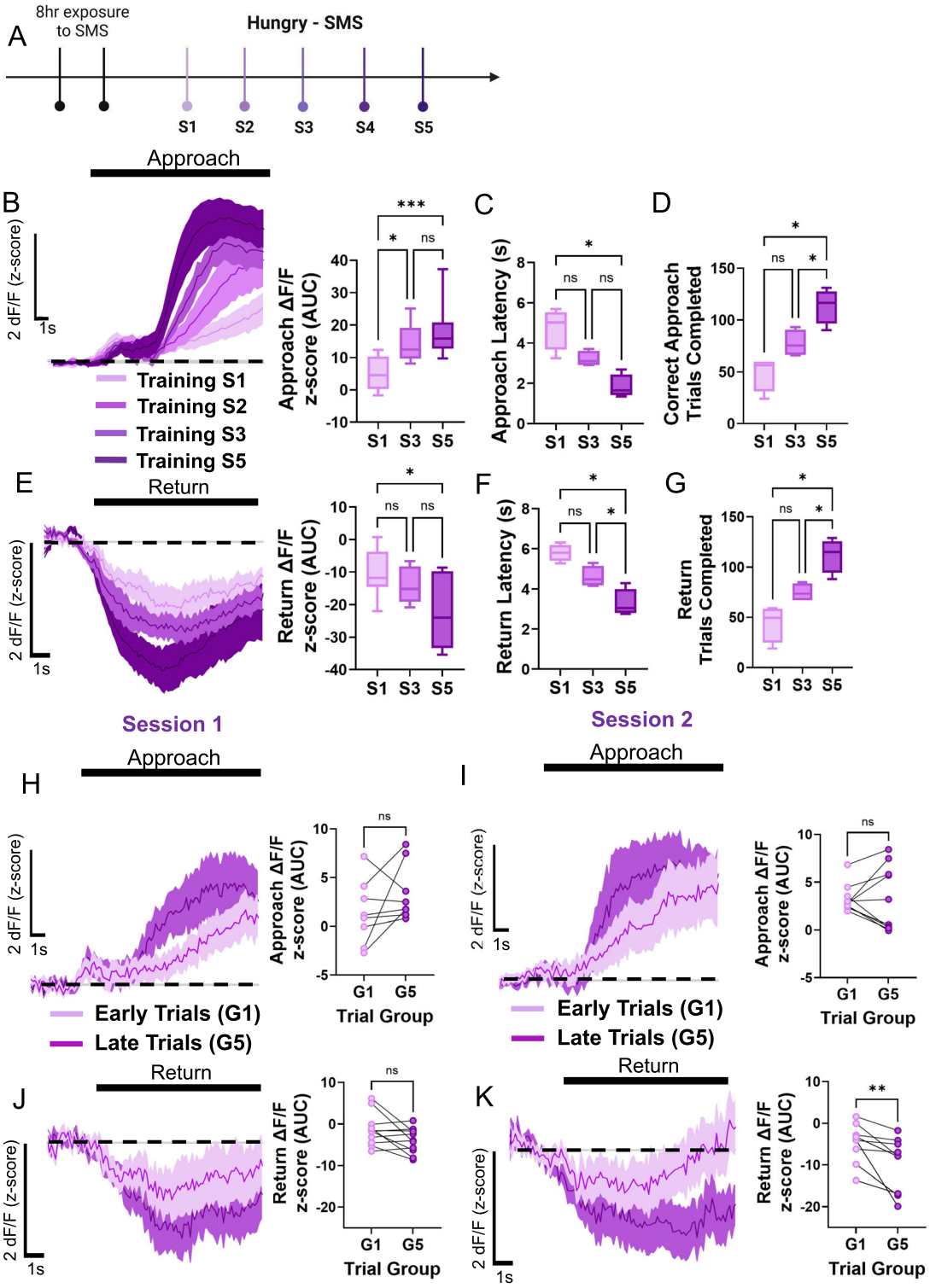

Hungry Strawberry MS
Sated Strawberry MS
Hungry Non-Cal
Hungry Sucrose
Hungry H2O outcome
Hungry Rew OM
Approach
Rew
A
B
C
D
E
F
G
H
I
J
K
L
H-SMS-C
H-SMS-P
S-SMS-C
S-SMS-P
H-Ncal-C
H-Ncal-P
H-Suc-C
H-Suc-P
H-H2O-C
H-H2O-P
H-OM-C
H-OM-P
1 dF/F (z-score)
4 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
Rew Delivery ΔF/F z-score (AUC)
Trial Group
G1 G2 G3 G4 G5
**
ns
*
M
Approach ΔF/F z-score (AUC)
***
*
ns
**
ns
ns
H-SMS-C
H-SMS-P
S-SMS-C
S-SMS-P
H-NCal-C
H-NCal-P
H-Suc-C
H-Suc-P
H-H2O-C
H-H2O-P
H-OM-C
H-OM-P
N
Rew Delivery ΔF/F z-score (AUC)
***
**
****
***
ns
ns
Base
H-SMS
S-SMS
H-NCal
H-10Suc
H-H2O
H-OM

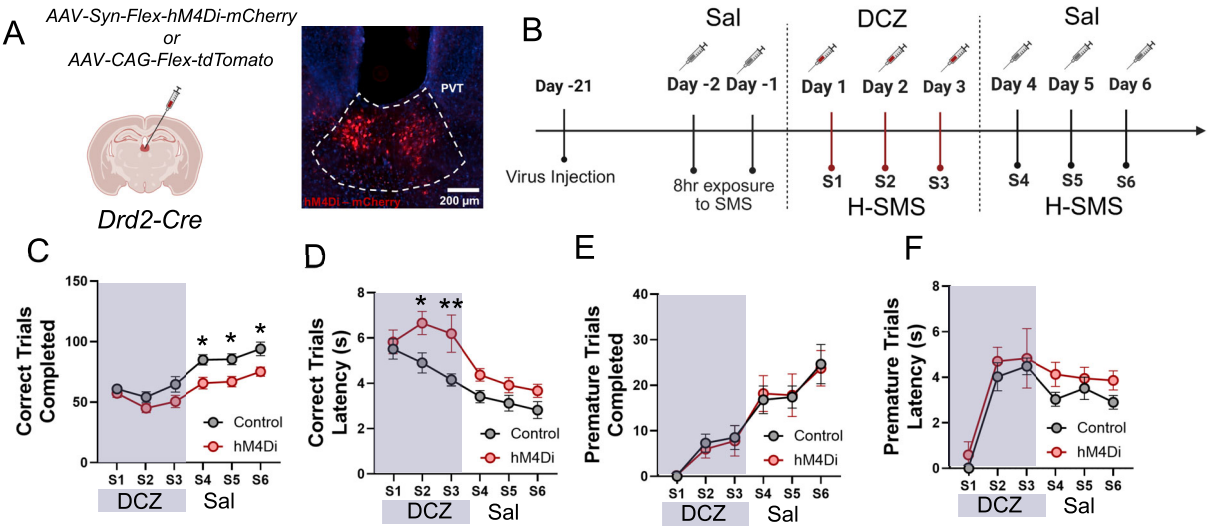

A
Hungry
Strawberry MS
B
Sated
Strawberry MS
C
Hungry
Non-Cal
D
Hungry
Sucrose
E
Hungry
H2O outcome
F
Thirsty
H2O outcome
G
Hungry
Rew OM
Return
2 dF/F (z-score)
1s
Early Trials (G1)
Late Trials (G5)
Return ΔF/F
z-score (AUC)
Trial Group
G1 G2 G3 G4 G5
**
*
ns
*
ns
*
ns

H
Return ΔF/F
z-score (AUC)
****
*
****
***
ns
****
*
*
Base
H-SMS
Base
S-SMS
Base
H-NCal
Base
H-10Suc
Base
H-H2O
Base
H-OM

I
Hungry - SMS
J
Sated - SMS
K
Hungry - Non-Cal
L
Hungry - Rew OM
Rew Zone
Arrival
Return
Initiation
Trial No.
Time (s)
z-score
10.0
5.0
0.0
-5.0
-10.0


A
H-NCal
H-NCal
H-NCal
Training in linear maze task
Pre-Test
S1
S2
Reward Magnitude Test
Reward Magnitude Test Session
Small
Med
Big
1
2
3
Trial Block
B
Approach
4 dF/F (z-score)
1s
H-NCal Early Trials (G1)
H-NCal Late Trials (G5)
Approach ΔF/F z-score (AUC)
ns
G1
G5
Trial Group
C
Approach
6 dF/F (z-score)
1s
H-NCal – Small Rew
H-NCal – Med Rew
H-NCal – Big Rew
Approach ΔF/F z-score (AUC)
*
ns
ns
Small
Med
Big
D
Approach Latency (s)
ns
ns
ns
Small
Med
Big
E
Return
2 dF/F (z-score)
1s
Return ΔF/F z-score (AUC)
ns
G1
G5
Trial Group
F
Return
2 dF/F (z-score)
1s
Return ΔF/F z-score (AUC)
ns
ns
ns
Small
Med
Big
G
Return Latency (s)
*
*
ns
Small
Med
Big


A
8hr exposure to SMS
Hungry - SMS
S1
S2
S3
S4
S5
Approach
B
2 dF/F (z-score)
1s
Training S1
Training S2
Training S3
Training S5
Approach ΔF/F z-score (AUC)
C
Approach Latency (s)
D
Correct Approach Trials Completed
E
Return
Return ΔF/F z-score (AUC)
F
Return Latency (s)
G
Return Trials Completed
Session 1
Session 2
H
I
Early Trials (G1)
Late Trials (G5)
Trial Group
G1
G5
J
K
ns
*
**
***


A
AAV-Syn-Flex-hM4Di-mCherry
or
AAV-CAG-Flex-tdTomato
Drd2-Cre
PVT
hM4Di - mCherry
200 µm
B
Sal
DCZ
Sal
Day -21
Day -2
Day -1
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Virus Injection
8hr exposure to SMS
S1
S2
S3
S4
S5
S6
H-SMS
H-SMS
C
Correct Trials Completed
Control
hM4Di
DCZ
Sal
D
Correct Trials Latency (s)
Control
hM4Di
DCZ
Sal
E
Premature Trials Completed
Control
hM4Di
DCZ
Sal
F
Premature Trials Latency (s)
Control
hM4Di
DCZ
Sal