NEURAL DYNAMICS OF COGNITIVE FLEXIBILITY: META-RPE SIGNALING WITHIN A PRELIMBIC CORTEX-VENTRAL TEGMENTAL AREA CIRCUIT EXPEDITES CONTINGENCY DEGRADATION DURING COGNITIVE FLEXIBILITY

Madelyn M. Hjort

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Garret D. Stuber, Chair

Sam A Golden

Paul E.M. Phillips

Nicholas A. Steinmetz

Program Authorized to Grant Degree:

Neuroscience

University of Washington

Abstract

Neural Dynamics of Cognitive Flexibility: Meta-RPE Signaling Within a Prelimbic Cortex-Ventral Tegmental Area Circuit Expedites Contingency Degradation During Cognitive Flexibility

Madelyn M. Hjort

Chair of the Supervisory Committee:

Garret D. Stuber

Departments of Anesthesiology and Pain Medicine & Pharmacology

Although tightly associated with prefrontal cortex (PFC), concrete cognitive flexibility signals have historically been ill defined. One common test of cognitive flexibility involves reversal learning, where the contingencies of discrete learned cues are enhanced or degraded, and an individual subsequently must flexibly remap their behavior. This work presents meta-RPE (mRPE), a cognitive flexibility signal that peaks in the middle of reversal behavior and represents the average of repeated, concentrated errors over many trials. Allowing mRPE to modulate canonical single-trial reward prediction errors (RPEs) expedites reversal learning and fits observed animal behavior better than models with static learning rates. Using novel statistical and imaging methods, this work identifies a subpopulation of neurons in prelimbic cortex (PL), a PFC subregion, that selectively encode a contingency degradation-related mRPE signal and can directly modify RPE via preferential representation in projections to VTA. Otherwise stable PL dynamics across reversal suggest the mRPE signal is unlikely attributable to representational drift. Dopaminergic innervation to PL does not predict the mRPE signal, instead representing a contingency elevation mRPE signal from elsewhere in the brain. Deriving mRPE and identifying its neural correlate in the PL-VTA circuit represents a quantitative advance in the field's understanding of cognitive flexibility signaling within the prefrontal cortex.

Dedication

To Sam, for encouraging me to focus on the reversal

To Paul, for recommending I look into the math

To Nick, for looking into the math with me

To Garret, for suggesting the math could parallel the neurobiology

To Jake, for making sure I was well fed

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION AND LITERATURE REVIEW

I. COGNITIVE FLEXIBILITY IN HEALTH AND DISEASE

We live in a volatile world where we must balance stable maintenance of value-guided choices with the ability to adapt choice behavior when the valuation of an outcome changes. Characterizing how the brain can flexibly degrade value of certain options while maintaining others has relevance for many cognitive disorders, including addiction, as a persistent high-value association between drug-related cues and drug taking is one of the main drivers of relapse (Sinha & Li, 2007). Understanding how the brain can selectively learn to reduce the value of cued outcomes is the first step towards future interventions that activate this circuitry to help individuals with substance use disorder (SUD) achieve long-term reduction of drug taking without affecting high value associations between other cue-outcome pairs.

When discussing contingency degradations of learned associations in response to changing environments, it is important to begin with shared terminology. *Contingency* is the predicted outcome or association between a cue or action and a reinforcing outcome. Outcomes become predicted through learning. *Learning* is a sustained change in value-guided behavior and can therefore be measured via changes in behavior. *Cognitive flexibility* is the ability to make changes to existing learning and adapt behavior to changing contingencies in dynamic environments. Contingency degradation is a specific subset of cognitive flexibility where behavior must adapt to decreases in the magnitude of a predicted outcome. Contingency degradation can be measured using *reversal learning* paradigms, which engage a specific type of cognitive flexibility where the contingencies of two or more learned cues or behavioral strategies are switched but all cues remain. Therefore, longitudinal characterization of how learned cue-reward associations are represented in the brain across reversal learning paradigms can characterize cognitive flexibility signaling. These definitions are summarized in the table below.

Table I-1: Definitions

Contingency	The predicted outcome between a cue or action and a reinforcer
Learning	A sustained change in value-guided behavior
Cognitive Flexibility	The ability to make changes to existing learning and adapt behavior to changing contingencies in dynamic environments
Reversal Learning	A specific type of cognitive flexibility where the contingencies of two or more learned cues or behavioral strategies must be switched but all cues remain

Deficits in cognitive flexibility have been reported in many human disorders, primarily via the Wisconsin Card Sorting Test (WCST). Used as a gold standard to test executive function in humans, the WCST asks a patient to match a given card to one of four (or more) other cards available on the screen. Cards differ in color, shape, or number, and which dimension must be employed for a successful match changes over the course of the session. Patients with cognitive flexibility deficits struggle to adapt to changing rules (Kopp et al., 2021). As a reward for accurate matching, patients are sometimes given small amounts of “money” per correct response, or at minimum the hope that their doctor concludes they do not need additional treatment. Deficits in WCST performance have been reported in many neuropsychiatric disorders including SUD (Faustino et al., 2021), schizoaffective disorders (Rady et al., 2012), bipolar disorder (Fleck et al., 2008), autism (Landry & Al-Taie, 2016), OCD (Henry, 2006), and, to a lesser extent, ADHD (Romine et al., 2004). Almost universally, affected patients struggle with perseverative errors—repeated responses to the previously rewarded dimension after a contingency reversal (Miles et al., 2021). This error profile suggests that the patients struggle to degrade the associated contingencies of outcomes predicted from the previous dimension. Thus, understanding the neural circuitry that underlies contingency degradation will act as an important first step towards future interventions to improve cognitive flexibility in individuals suffering from a wide scope of diseases, including SUD.

Human neuroimaging has provided some insight into the circuitry that may be responsible for contingency degradation. Studies imaging during the WCST repeatedly report high levels of

activation in the prefrontal cortex (PFC) during flexibility, most significantly in the anterior cingulate cortex (ACC, area 32) and orbitofrontal cortex (OFC) (for review see (Monchi et al., 2001)). Lesion studies likewise recapitulate the importance of the PFC for WCST performance and subsequently cognitive flexibility (for review see (Gläscher et al., 2019)). Beyond recording activations and lesion consequences, it is difficult to dissect the neural circuitry that may underlie these functions, necessitating animal models. Non-human primate analogs of the PFC share high projection and cytoarchitectural similarity to humans (for review see (Levy, 2024)) and generally display activation of the same regions (Premereur et al., 2018). Electrical recordings have given some insight into the relationship between cellular firing on individual days in well trained animals and task parameters (for review see (Klein-Flügge et al., 2022)), but large-scale behavioral interventions are difficult in primates. Rodents are much more tractable for large-scale longitudinal cellular and behavioral recordings during cognitive flexibility.

Although somewhat controversial, rodents do seem to have a prefrontal cortex (for review see (Laubach et al., 2018)). The primary point of dispute lies in a lack of rodent cytoarchitectural homology for the granular portions of the primate PFC (Laubach et al., 2018). Rodent PFC contains a number of subregions, medially including ACC, prelimbic cortex (PL), and infralimbic cortex (IL), and laterally subdivisions of the OFC (Paxinos & Franklin, 2019). Cytoarchitecturally, ACC, PL, and IL all lie within what primate researchers call ACC (Laubach et al., 2018), though they also share functional connectivity that is also similar with granular portions of primate medial PFC (Heilbronner et al., 2016). The strongest analog via both cytoarchitectural and projection-based classification is PL for primate area 32, which is reflected in newer versions of the mouse atlas (Laubach et al., 2018; Paxinos & Franklin, 2019). Fortunately, area 32 (hereafter referred to as PL) is also one of the most activated regions in humans and primates across cognitive flexibility studies (Monchi et al., 2001; Premereur et al., 2018), easily allowing more granular flexibility studies in rodent models.

Across species, PFC is a layered structure largely composed of pyramidal neurons intermixed with interneurons. Pyramidal neurons in the medial PFC (mPFC) receive input from many nodes in reward-learning circuitry including the ventral tegmental area (VTA), basolateral amygdala (BLA), and hippocampus, in addition to other cortical regions and the thalamus (Anastasiades & Carter, 2021). Excitatory in nature, pyramidal neurons in PL project back to many of these same structures, including VTA (~10%), BLA (~10%), hippocampus and thalamus (40%), as well as regions of the ventral striatum (~20%), in addition to other cortical areas (Anastasiades & Carter, 2021; Gabbott et al., 2005). These projections are generally divided into two large classes based on their projections within (intra-telencephalic, IT) or out of (pyramidal tract, PT) the cortex, and are spatially intermixed (Dembrow et al., 2010). Although human studies have implicated PL and OFC activation in cognitive flexibility behaviors, they cannot directly test where these signals are sent or how they are represented at the single-cell level. Adapting the WCST for animal models, especially rodents, allows for more granular dissection of cognitive flexibility signals within the PFC. Dissecting flexibility signals is an important first step towards future studies that may induce increased cognitive flexibility in individuals effected by many neuropsychiatric disorders, including SUD.

II. TESTING PREFRONTAL COGNITIVE FLEXIBILITY IN ANIMAL MODELS

In order to better understand the neural mechanisms that underlie cognitive flexibility, researchers have developed an array of cognitive flexibility tests that can be performed by animals. Cognitive flexibility tasks in animal models generally involve an initial association between sensory cues or spatial behavioral strategies and the probability of cue-associated reward delivery that is abruptly changed once the initial contingencies are well learned. The simplest forms involve two cues or spatial strategies, one 100% reward predictive, and one 0% reward predictive, that switch contingencies—or reverse—after an initial learning period. Since this task structure is relatively simple, there is a strong homology between rodent, primate, and human iterations of reversal learning tasks. In a more complicated form, researchers adapt the

WCST for animals (for review see (Izquierdo et al., 2017)). Primates perform a very similar version to the human task with virtual cards on a touch screen, while rodent adaptations are built more directly from simple paradigms. Rodent analogs of the WCST usually involve two bowls filled with differently scented and textured bedding, with one hiding a reward that must be dug out. After animals reliably dig in the correct odor cued bowl, rules can be changed on the following axes: *odor-odor,* where animals must dig in the other scented bowl for the treat, or *odor-texture,* where animals must ignore odor cues and pay attention to bedding texture, or texture-texture, where which bedding texture contains the treat is switched. Odor-odor or texture-texture transformations are deemed "intradimensional shifts", while odor-texture or texture-odor shifts are categorized as "extradimensional shifts" (for review see (Izquierdo et al., 2017)). The degree to which these differences in shifting are meaningful is not well characterized, and difficult to dissociate from other task parameters including discrimination difficulty. Nevertheless, a combination of simple reversal and modified WCST analogs have been used to probe the neural underpinnings of reversal learning in many animal models.

Studies of prefrontal contributions to cognitive flexibility have largely centered around the OFC. One of the more lateral regions of the PFC, OFC contains strong reciprocal connectivity with the hippocampus, and has been postulated to be involved in modifying and/or containing maps of task or value spaces (Padoa-Schioppa, 2007; Schoenbaum et al., 2009; Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014). OFC lesions impair reversal learning in simple task paradigms, including visual, auditory, spatial, and operant reversals in primates and auditory, visual, olfactory, tactile, and spatial reversals in rodents (for review see (Izquierdo et al., 2017)). In some cases, OFC silencing completely abolishes the ability of animals to adapt their behavior to new contingencies in a reversal task (Banerjee et al., 2020). Observations of orbitofrontal necessity for reversal learning have precipitated a number of more recent computationally driven studies where single neurons are tracked across learning to assess their roles across acquisition stages. Studies of OFC's role during acquisition of cue-reward associations suggest that OFC's

importance in reversals may lie in detecting that a change has occurred to reward contingencies within the landscape. In simple Pavlovian learning, longitudinally tracked orbitofrontal ensembles show high activity early in learning, but have a minimal, blunted response once the task is well learned (Namboodiri et al., 2021; P. Y. Wang et al., 2020). This observation has led to the theory that OFC may act as a gate for learning that expedites the learning rate when new cue-reward contingencies must be learned (Namboodiri et al., 2021). Analysis of learning patterns revealed that OFC appears to learn retrospectively from unpredicted reward deliveries (Namboodiri et al., 2021). The observation that OFC activity is related to the unfamiliarity in presentation of a reward following an associated cue suggests that one would expect OFC activity to be high once again immediately following a contingency reversal when animals first experience unpredicted reward deliveries from the new contingencies.

Longitudinal recordings of orbitofrontal neurons across a tactile reversal of go/no-go whisked texture cues suggest that OFC activity is indeed highest immediately following contingency reversal. Banerjee and colleagues established that the OFC is specifically necessary for the reversal phase of this task: receptor-mediated OFC inhibition does not impair initial learning but renders animals unable to remap their behavior to reversed contingencies. Longitudinally tracked characterization of OFC activity revealed that the largest proportion of the population is active immediately following the reversal, with significantly smaller ensembles active when the initial or reversed contingencies are well learned. Trial averages of OFC responses revealed the highest OFC activity occurred at the first false alarm trial--where animals respond to the previous go cue--which has been reversed to no-go. Within approximately three false alarm trials, OFC no longer responds, despite stable behavioral remapping of the new

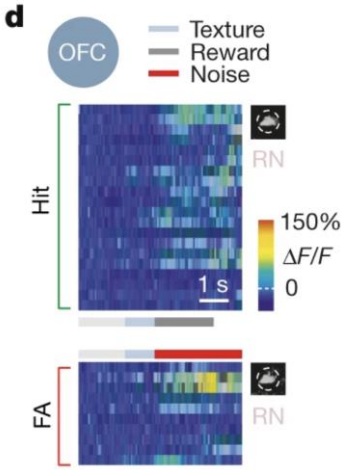

Figure I-1: **OFC activity following whisked texture contingency reversal, as represented in Banerjee et al. 2020** (4d reproduced without modification). For Hit (licks to new rewarded cue) and FA (false alarm trials, where animal licks in response to the old contingencies), trials from reversal is represented on the vertical axis in the time domain.

contingencies taking an additional several days (**Fig. I-1**). OFC also responds to new 'hit' trials as animals begin to respond to the new, reversed, go contingency but the amplitude is nearly half that of false alarm trials (Banerjee et al., 2020). This observation is somewhat at odds with the theory that OFC responding is primarily reward driven, since activity is highest at unexpected omissions. Regardless, it still supports the general idea that OFC activity is highest when cue-reward associations do not follow a familiar pattern, and that as a new pattern becomes more familiar, OFC activity reduces.

Blocking the OFC signal in this task prevented remodeling of S1 ensembles for preferential go-cue activation and the development of reversed behavioral contingencies (Banerjee et al., 2020) Somewhat in contrast, a different study using an odor reversal task revealed similar high OFC activation at the start of the contingency reversal, but did not find evidence of activity remodeling in piriform (odor sensory) cortex (P. Y. Wang et al., 2020). Taken together, these studies suggest a vital role for OFC signaling at the onset of unfamiliar cue-reward contingencies, including unexpected contingency reversal. Such a role is somewhat consistent with a trial-by-trial reward prediction error (RPE) framework: RPE is highest in magnitude for the unexpected rewards and omissions that occur immediately following the contingency reversal, as is OFC activity. However, OFC activity reduces too quickly following the reversal for its activity to be completely explained by RPE, as the animal's behavior takes days or weeks, not a handful of trials, to fully remap to the new contingencies.

This targeted role for OFC at the beginning of contingency reversal leaves open a fundamental question for reversal learning and cognitive flexibility: what neural circuitry is responsible for supporting behavioral remapping to reversed cue contingencies over the days it takes to remap to stable behavior? Such a role may be fulfilled by the nearby prelimbic cortex.

III. BEHAVIORAL EVIDENCE FOR THE INVOLVEMENT OF PRELIMBIC CORTEX IN COGNITIVE FLEXIBILITY

Unlike the nearby OFC, the role of the more medial PL in cognitive flexibility-driven behaviors like reversal learning remains unresolved. This hinges largely upon discrepancies (Izquierdo et al., 2017) in the findings of rodent behavioral lesion studies where transient optical, chemical or permanent surgical lesion of PL can impair, facilitate, or have no effect on how quickly animals can adapt to changes in their task structure, environment, or state. However, analysis of similarities between studies that report equivalent lesion effects reveals a discernable subset of reversal learning behaviors that require PL.

Studies that report impairments in reversal learning following PL lesion have a diversity of task conditions. In 1997, Bussey and colleagues repeatedly presented rats with a randomly ordered pair of visual stimuli that matched in size, shape, and luminance but differed in color. A nosepoke in the port adjacent to the rewarded stimulus (S+) achieved delivery of a sucrose pellet, while a poke towards the unrewarded stimulus (S-) achieved nothing. After animals poked for the correct stimulus 85% of the time in two consecutive sessions, the identities of the S+ and S- stimuli were reversed. Before the rats began the task, researchers removed the medial frontal cortex (lesion maps largely overlap with PL, with some inclusion of IL), ACC, or performed a surgical sham for control animals. All rats learned the initial contingency at comparable speeds. However, when S+ and S- were reversed, the PL lesioned rats took almost twice as long to reach the success criteria when compared to the ACC and sham rats. Researchers noticed that all animals performed equally poorly immediately following reversal, but PL lesioned rats needed thousands more trials than ACC or sham rats to prevent sliding back to a pattern of responding to the previous cue contingencies (Bussey et al., 1997).

Other research with more modern techniques, using optogenetics for transient and reversible lesions of PL, has reported similar results. Biró and colleagues trained mice to run in virtual reality towards a circle with a hashed pattern instead of a solid, filled circle in order to

achieve a sucrose reward until they chose the patterned stimulus 90% of the time in a given session. Next, the task rules were switched such that mice had to run towards the larger of the two circles, ignoring the pattern. Before training, researchers injected mice bilaterally with an excitatory opsin in PL interneurons, which subsequently inhibit PL pyramidal neurons. The inhibited mice learned the initial contingency as quickly as their non-inhibited counterparts even when they experienced PL inhibition during and following reward delivery, but when the rule switch occurred inhibited animals took over twice as long to reach criteria for the new contingencies. More granular analysis of their behavior revealed that inhibited animals did not take longer to respond to 'conflict' trials (where the new and old rules were at odds) and instead responded much faster, but often to the old contingencies (Biró et al., 2019). Other studies involving visual gratings (Bussey et al., 1997; Chudasama & Robbins, 2003; Nakayama et al., 2018), patterns (Brigman & Rothblat, 2008), colors (Baker & Ragozzino, 2014), or transitions from visual to auditory cues (Haddon & Killcross, 2006; Marton et al., 2018), are affected similarly by PL lesions: animals learn the initial contingencies without deficits, but struggle to adapt to changes in stimulus contingency within their environment.

PL lesion effects are not unique to visual stimuli. Animals struggle to adapt to changing stimulus contingencies between odors and textures (Birrell & Brown, 2000; Mukherjee & Caroni, 2018; Ragozzino et al., 2003), or extended levers (De Bruin et al., 2000; Kosaki & Watanabe, 2012) in the same way as visual stimuli. Transitions from cued to spatial task strategies, where the animal must transition from using external cues to an internal orientation system to always turn to a particular side (Oualian & Gisquet-Verrier, 2010; Ragozzino et al., 1999; Rich & Shapiro, 2007; Seamans et al., 1995; Young & Shapiro, 2009), cardinal direction (Ragozzino et al., 1999; Rich & Shapiro, 2007; Young & Shapiro, 2009), or reward dish (Birrell & Brown, 2000; Mukherjee & Caroni, 2018; Ragozzino et al., 2003) are likewise affected. Notably, transitions from a spatial to a cued strategy do not seem to be affected by PL lesion in the same manner. This phenomenon is well demonstrated by Oualian and colleagues. When they first trained rats to go to the brighter

illuminated arm of a two-arm y-maze to receive a reward and then reversed the task rules such that rats should always travel to the right (or left) arm to receive a reward, PL lesions impaired reversal. However, when rats were first trained to travel to the right (or left) arm, and then task rules reversed to travel to the brighter arm, PL lesions had no effect on reversal behavior. Similar reports have been reported by other groups (Floresco et al., 2008). Taken together, these studies suggest PL lesion effects preferentially effect transitions from previous task strategies that employ cue-reward associations, but not strategies that rely only on internal spatial orientation.

If a task must have discrete, external cue-reward associations to engage PL, it follows that other tasks that employ internally guided spatial tasks would not report impacts of PL lesions on reversal learning. This conclusion is supported by the literature. Lesions to PL do not have an effect on transitions between spatial only strategies, most often evaluated via switching between side and cardinal direction turning (Boulougouris et al., 2007; Ragozzino et al., 1999; Young & Shapiro, 2009). A notable exception to this is a report that PL inactivation impairs reversal learning in the Morris water maze when the platform is moved after its location is well learned(Latif-Hernandez et al., 2016). One interpretation of this discrepancy is that mice use sensory cues within the chamber or testing room, instead of an internal spatial model, to predict the platform location. The notion that the Morris water maze task uses cues for navigation is congruent with both the original description of the task, and other review literature (Morris, 1981; Othman et al., 2022; Vorhees & Williams, 2006). When the maze is enclosed by blackout curtains that prevent accessing distal cues from the experimental room, animal performance suffers (Morris, 1981; Williams et al., 2003), suggesting a reliance on distal cues for task performance. Although the water maze does not have the same discrete cues delivered purposefully in the task context, its link with reliance on distal cues within the room to predict platform location (and reward animals with a break from swimming at risk of drowning) evokes strong similarities to the aforementioned tasks reversing from more canonical cue-reward associations that are impaired with PL lesion.

A subset of studies involving cue-guided reversals track animal's behavior across multiple contingency switches. Notably, PL effects are strongest, and often only significant, the first time an animal experiences a reversal of a given contingency, whether from a visual cue (Bussey et al., 1997; Oualian & Gisquet-Verrier, 2010) or extended lever (De Bruin et al., 2000). A targeted importance for PL involvement in the first contingency reversal suggests studies that pre-train animals with multiple reversals before recording may overlook intitial PL involvement. In support of this, animals that are well trained on visual discrimination (Hervig et al., 2020; McAllister et al., 2015) or lever-guided (Dalton et al., 2016) tasks before lesion do not display reversal impairments. Why PL seems to have preferential importance for early experiences of contingency reversals remains to be discovered. Perhaps its importance lies in constructing or shaping downstream circuitry to be able to switch between task states, and once state sets are thereby generated, other brain areas toggle between constructed states following surprise signals, perhaps from OFC. Exactly where task state signals reside likewise remains unresolved, though several brain areas have been postulated including OFC, hippocampus, ACC, and amygdlar structures (see (Izquierdo et al., 2017)).

There appears to be a correlation between cue-discrimination difficulty and the number of reversals for which PL is needed. In addition to assessing reversals between size- and luminance-matched colored shapes, Bussey and colleagues also assessed reversal of shapes that were easier to discriminate. To aid discriminability, cues consisted of two consistent circles with different sizes, colors, patterns, and borders (smooth vs jagged). When rats with medial frontal cortex lesions were asked to reverse responding to the two shapes, PL lesions no longer had an effect on the speed of reversals (Bussey et al., 1997). Brigman and Rothblat observed similar results in mice. When asked to discriminate between shapes with different patterns, where lines in one direction predicted reward while lines in another direction did not, PL lesions significantly impair the speed by which the mice were able to adapt their behavior to reversal of the pattern-reward contingencies. Conversely, when asked to reverse predicted reward contingencies

between a black shape and a white shape, both presented on a gray background, PL lesions had no effect (Brigman & Rothblat, 2008).The notion that there is a threshold of cue or task complexity required to engage PL is supported by other experimental data, as switching from distinct, non-probabilistic odor cues is likewise not impacted by PL inactivation (Birrell & Brown, 2000; Mukherjee & Caroni, 2018) and accomplished in only a handful of trials, as would be expected from an easy task. These results suggest a possible alternate explanation of the spatial-only and pre-training findings: that the tasks are too simple to require PL. The majority of internally guided spatial tasks that have been used to assess the link between PL and reversal behavior utilize short behavioral motifs that largely consist of turning in a particular direction. As animals perfom simple turns that orient to different directions hundreds, if not thousands, of times per day, performing basic modifications to such a behavior may be so simple and innate that it is baked into neural circuitry that predates the emeregence of prefrontal cortex. Similarly, it is possible that when animals have experienced the same contingency reversal multiple times, the percieved complexity of the task decreases below the PL threshold since animals are familiar with dynamic changes in the environment, and have experienced identical changes previously. This interpretation suggests the existence of a hierarchy of neural circuitry for cognitive flexibility: something outside the prefrontal cortex for very simple tasks, and specialized prefrontal circuitry that activates once a complexity threshold is reached. Observations that the nearby OFC is still needed to reverse contingencies in tasks regardless of complexity (Izquierdo et al., 2017) is somewhat at odds with this interpretation.

Analysis of the error profile that results from PL lesions supports the idea that PL plays an active role in the construction of new ‘states’ or sets following contingency reversals. Whether animals switch from visual cues (Baker & Ragozzino, 2014; Biró et al., 2019; Floresco et al., 2008; Haddon & Killcross, 2006; Marton et al., 2018; Oualian & Gisquet-Verrier, 2010), texture cues (Spellman et al., 2021), levers (De Bruin et al., 2000; Kosaki & Watanabe, 2012), or navigating a Morris water maze (Latif-Hernandez et al., 2016), PL lesions extend responding to the previously

reinforced cue far past the contingency switch. This error profile, termed preservative or regressive error, in which animals consistently slide back into responding to the previously reinforced cue-reward association, suggests that PL may play a direct role in transitioning cued-responding away from the original and towards new contingencies. In support of this, studies of rats with particularly large prelimbic lesions report that animals can go days before they stop responding to the previous reinforcer (Rich & Shapiro, 2007; Young & Shapiro, 2009) even if they make progress towards the new contingency on a given day. This, when coupled with orbitofrontal data (see section II), suggests that animals are able to detect that a change in their environment has occurred (since OFC is intact) and make some progress in reacting to the change, but without PL that change is unable to persist to and be built upon in subsequent sessions until many days have passed. Such an impairment is consistent with the idea that PL circuitry, although perhaps not necessary to detect an initial change in the environment, is needed to expedite construction of a stable set or state of responding to the new cue contingencies. That this impairment seems specific to degradation of the previous contingency suggests that the mechanism of PL action to facilitate reversal behavior may rely more strongly upon modulation of negative reward prediction errors to facilitate contingency degradation, though this has yet to be tested directly.

IV. CUE REPRESENTATIONS IN PRELIMBIC CORTEX

Behavioral evidence suggests a fundamental role for PL in the degradation of cue-reward associations following contingency reversals. Understanding how cue representations are first constructed in PL is a vital first step to characterize how they are degraded during reversal learning. A small number of longitudinally tracked studies have quantified cue acquisition dynamics within PL. Otis and colleagues tracked PL pyramidal neurons across early and late stages of acquisition of a simple discrimination between tone cues with deterministic reward prediction. Before learning, only 3% of neurons displayed firing specificity for the future reward-paired cue (CS+) or non-reward-paired cue (CS-). Notably, once the contingencies were well learned, 34% of neurons fired preferentially for the CS+, while CS- coding remained around 3% (Otis et al., 2017). Wang and colleagues took this finding one step further in a paradigm with 2 CS+ and CS- cues. Looking only at cue-period, PL activity was able to decode CS+ from CS- trials over many days, and equally well between the two CS+ cues due to a similar shared activity profile (P. Y. Wang et al., 2020). Taken together, these results suggest PL displays stable, preferential encoding of reward-predictive cues within a single task context. This selective coding for the reward-paired cue could explain why PL lesion effects seem specific to degradation of reward paired cues: PL can only modify the information it contains, which is CS+ specific.

Electrical recordings of untracked PL populations provide preliminary evidence for whether this stable encoding extends across contexts following contingency reversal. In a study of serial reversals, Guise and Shaprio found they could reliably decode which of the two task states a rat was in from consistently shared elements of population activity and that decoding performance was strongest in rats that performed the task most accurately (Guise & Shapiro, 2017). This suggests that even after multiple contingency reversals, population activity reliably returns to a stable state space. Stable population coding before and after reversal has been observed by others, including Powell and Reddish who noted a distinct “disorganization” of PL dynamics during

the reversal period that contrasted with consistent dynamics before and after the reversal (Powell & Redish, 2016). Rich and Shapiro also observed stable pre- and post- reversal dynamics that could decode the task state, and more notably that the decoding shift from PL dynamics occurred before the rats began to reliably change their behavior to respond to the new contingencies and increased in strength over many trials (Rich & Shapiro, 2009). Comparable to others, Karlsson and colleagues observed a relatively stable period of population activity during stable days, and a “more volatile” (higher gamma) pattern of firing during behavioral change. What this volatility can be attributed to within the reversal remains unresolved. Like Rich and Shapiro, they observed that this change in firing preceded behavioral change, and added that it cannot be predicted or decoded from any characteristics of the current or previous trial (Karlsson et al., 2012). While most studies report a gradual shift between representations of task state, Durstewitz and colleagues suggest the transition is “rapid” from Bayesian modeling of task parameters (Durstewitz et al., 2010), though examination of their example data suggests that their model transitions much faster than the actual neural population dynamics. Taken together, this evidence suggests that PL encoding may be more stable within contexts than across them. Unfortunately, none of these studies looked at cue-encoding directly, but when paired with other evidence, they suggest that there is a stable PL population state before the reversal, which may be cue-related. Exactly what changes about this state, and whether reward-paired cue encoding is related to the changing states remains unresolved.

Studies of extradimensional shifting across cue modalities provide additional preliminary evidence that there may be a stable cue representation across task states and contexts in PL. In tasks where a cue indicated whether to lick left or right to receive reward, Bari and colleagues observed a cue-locked value code within PL neurons projecting to the DMS that consisted of a mix of relative value firing that persisted until the next trial, and absolute value firing that decayed during the inter-trial interval (Bari et al., 2019). Bissonette and Roesch also observed that cue

representations comprising about 30% of the population shifted from preferentially encoding light stimuli to encoding the new reward-predictive odor stimuli as animals adapted to the contingency reversal over days (Bissonette & Roesch, 2015). Marton and colleagues observed a similar shift in PL firing in high frequency bands to audio cues when mice switched from previously rewarded visual cues. (Marton et al., 2018). Whether this shift occurred within a stable subpopulation of neurons or is the source of the differential state decoding observed by others remained somewhat unresolved until recently. In a tracked single-cell imaging study, Reinert and colleagues found a subpopulation of PL neurons that preferentially activated toward go cues regardless of rule in a task that required shifting between visual grating rules. Notably, while both contexts had cells that represented the no-go cues, these cells did not generalize across contexts (Reinert et al., 2021). A second tracked single-cell imaging study where mice shifted between texture and odor go cues reports similar results in high task-locked activity correlations between cue cells to both texture and odor cues when they were reward predictive (Spellman et al., 2021). Taken together, these results support the presence of a neural population in PL that retains stable reward-paired cue encoding across state changes. Whether this population is universal across contexts not linked by a reversal or shift of cue contingencies remains unknown. The existence of such a finding could somewhat uniquely situate PL as the seat of a universal lookup table of cues associated with a given reward, which could explain why its lesion impairs degradation of high-value cues specifically. If the current state of the lookup table cannot be accessed, it is possible that animals default to the last accessed values, perseverating to the old cues for some time. However, it is also possible, and more likely, that behavior lesion effects do not result from an inability to access cue-reward associations (as cue locked behavior remains) but from an inability to *modify* the predictive power of cues previously paired with reward.

There is preliminary evidence of a PL population that specifically activates during the contingency reversal. When Bissonette and Roesch separated out “conflict” trials, where new

odor and old light rules were at odds, they found a unique firing increase in 11% of neurons (Bissonette & Roesch, 2015). Malagon-Vina and colleagues also observed a population with increased firing to “bad behavioral performance” during the reversal, that was distinct from the population that fired during “good behavioral performance” (i.e. stability) (Malagon-Vina et al., 2018). Rich and Shapiro identify one single unit with firing that parallels the “proportion [of] incorrect” choices during the reversal (Rich & Shapiro, 2009). The relationship of these “conflict”, “incorrect” or “poor performance” neurons to the “volatility” Karlsson et al. or “disorganization” that Powell et al. observed remains unknown, though their sudden appearance in PL following contingency reversal could explain poor decoding during behavioral transition. Although a reversal-specific PL population has been noticed by multiple groups, how it is activated, how its dynamics shift over time, how it interacts with other populations in PL, and its causal relationship to reversal behavior remains enigmatic

V. CATECHOLAMINERGIC INNERVATION OF PRELIMBIC CORTEX DURING COGNITIVE FLEXIBILITY

The source of PL activation during reversal has not been studied at length. However, the contribution of the catecholaminergic neurotransmitter dopamine, implicated in many forms of learning (for review see (Wise, 2004)), to reversal learning has been studied and may provide an explanatory mechanism. Of particular note, van der Meulen and colleagues performed microdialysis of mPFC dopamine before, during, and after a simple contingency reversal. Microdialysis is a technique that temporarily extracts the interstitial fluid between cells, and measures neurotransmitter levels in the sample via examination of unique molecular properties (for review see (Müller, 2002)). van der Meulen et al. detected minimal dopamine levels in mPFC during stable behavior, but noticed a sharp and significant rise during contingency reversal compared to controls that did not reverse. Of note, levels of norepinephrine, another catecholaminergic neurotransmitter associated with learning (for review see (Harley, 2004)) did not rise with reversal (van der Meulen et al., 2007). An important caveat is that microdialysis measurements are slow compared to learning trials (minutes vs. seconds), and therefore cannot be related to reversal behavior in a time-locked manner. A second-by-second profile of PL dopamine release during reversal learning would allow a more thorough dissection of how dopaminergic release in PL during reversal impacts behavior but has not yet been performed. Nevertheless, van der Meulen et al.'s dopamine increase during reversal could drive the increase in PL cells associated with reversal learning, though this has yet to be tested directly.

Local release of dopamine can impact neural activity in multiple ways. Dopamine primarily signals through two families of receptors, the D1 family of G_s- protein coupled receptors, which generate excitatory cAMP via increased adenylyl cyclase activation, and the D2 family of $G_{i/o}$-coupled receptors that inhibit adenylyl cyclase and subsequently cellular activity. Altogether, five dopamine receptors have been identified, with the D1 family consisting of D1 and D5 receptors, and the D2 family consisting of D2, D3, and D4 receptors. Located exclusively postsynaptically,

D1-family receptors (D1Rs) excite neurons in their target area following local dopamine release. Somewhat in contrast, D2-family receptors (D2Rs) are located both pre- and postsynaptically and inhibit both dopaminergic target and source neurons (for dopamine receptor review, see (Beaulieu & Gainetdinov, 2011)). Thus, van der Meulen et al.'s reported rising levels of dopamine in PL could potentially both excite and/or inhibit the activity of PL pyramidal neurons during contingency reversal depending on which receptors they express.

The distribution of D1Rs and D2Rs within mPFC has been characterized. Using fluorescently tagged agonists, Vincent and colleagues revealed that the majority of prefrontal dopamine receptors are in pyramidal neurons (91% for D1Rs and 78% of D2Rs). Within neurons, receptors were densest in mPFC output layers, where 22% of neurons expressed D1Rs and 47% D2Rs (Vincent et al., 1993). LeMoine and Gaspar report a similarly high density of dopamine receptors within mPFC layers V and VI (Le Moine & Gaspar, 1998). In contrast to Vincent et al., Santana and colleagues report that D2Rs localize to pyramidal neurons only, while D1Rs have mixed distribution across pyramidal and interneurons (Santana et al., 2009). With regards to PFC dopaminergic input, electron microscopy studies report the highest innervation density in PL (Van Eden et al., 1987). Like dopamine receptors, input fibers are strongest to deep projection layers V (Van Eden et al., 1987) or V/VI (Descarries et al., 1987). More recent studies using retrobeads in thalamus and corpus callosum have revealed that D2R-containing pyramidal neurons in PL exclusively project out of the cortex, while D1R neurons show mixed selectivity (Gee et al., 2012). Taken together, these histological findings suggest that dopamine may play a role in modulating output neurons within PL.

How dopamine receptor activation impacts pyramidal neuron activity in the mPFC has been extensively studied *ex-vivo*. Low concentrations of dopamine significantly increase spiking activity, while high dopamine levels inhibit mPFC neurons (Trantham-Davidson et al., 2004; Zheng et al., 1999). Applying D2R antagonists abolishes the inhibitory effect of dopamine at high levels (Seamans, Gorelova, et al., 2001; Tseng & O'Donnell, 2004; Zheng et al., 1999), while D1

agonists block any dopamine-mediated spiking increases (Seamans, Durstewitz, et al., 2001; J. Wang & O'Donnell, 2001). Experiments blocking D1Rs report that dopamine can still interact with D2Rs and decrease spiking at low concentrations (Otani et al., 1998; Seamans, Gorelova, et al., 2001; X. Wang et al., 2003). Taken together, these results suggest that dopaminergic interactions with D1Rs dominate signaling at low concentrations, but are overtaken by interactions with D2Rs at higher concentrations (Trantham-Davidson et al., 2004). As these studies occurred in cut brain slices, the consequences of changing dopamine receptor interactions on behavior could not be assessed directly.

The contribution of dopamine release and receptor signaling to cognitive flexibility *in vivo* has been well categorized at the behavioral level. The effects of D1R-modulation are inconsistent, with some studies reporting increased errors (Izquierdo et al., 2006; Ragozzino, 2002), while others report no effects (Alsiö et al., 2019). In contrast, when injected at the start of contingency reversal, D2R antagonists reliably impair reversals (Alsiö et al., 2019; Boulougouris et al., 2009), as do D2R knockouts (DeSteno & Schmauss, 2009; Kruzich et al., 2006; Kruzich & Grandy, 2004), with an error profile marked by perseverative errors (Alsiö et al., 2019; Kruzich et al., 2006). The similarity of this phenotype to that of PL lesion was not lost on researchers. In brainwide activation scans, D2-KO mice display reduced activation of the mPFC, including PL, following reversal (DeSteno & Schmauss, 2009), and in healthy mice, D2R expression in the mPFC is positively correlated with reversal performance (Laughlin et al., 2011). Local administration of D2R-antagonists in PL significantly increases perseverative errors and performance during early reversal (Floresco et al., 2006). This effect seems to be specific to D2 signaling in target regions, as D2 deletion from dopaminergic neurons does not significantly impair reversal (Linden et al., 2018). Taken together, these results suggest that the ability of dopamine to activate D2Rs in PL is necessary for reversal. When coupled with the *ex vivo* studies of dopamine-receptor dynamics, this implies that dopamine-mediated activity decreases may be more important than dopamine-mediated increases on PL activity during reversal. This

interpretation is somewhat at odds with reported firing increases in a subpopulation of PL neurons during reversal (see section IV). Given this, it is possible that falling dopamine concentrations from the negative RPEs that result from selecting previously, but no longer, rewarded cued outcomes facilitate PL activations via falling D2R- mediated inhibitions. Studies correlating trial-locked dopamine levels and PL neural activity are needed to test if this may be the case.

VI. SUMMARY AND HYPOTHESIS

Research across species has implicated a vital role for the prefrontal cortex (PFC) in cognitive flexibility behaviors. The orbitofrontal cortex (OFC) within the lateral PFC appears to play a vital role in early detection of changing environmental contingencies, but how animals subsequently remap their behavior to new contingencies has remained elusive. Rodent behavioral studies have also implicated the nearby prelimbic cortex (PL) in cognitive flexibility behaviors, perhaps specifically surrounding learning from errors following repeated reward omissions. Recordings of PL population activity during contingency reversals have detected elevations in population activity that persist until behavioral performance has stabilized. Thus, a major goal of this work is to longitudinally categorize PL activity across contingency reversal, with the *hypothesis* that there will be a subpopulation of PL neurons that activates specifically to repeated unexpected reward omissions. Since dopamine release in PL is also elevated following contingency reversal, we will characterize PL dopamine dynamics across reversal behavior and compare with PL neural dynamics to test the *hypothesis* that dopamine release in PL may provide an explanatory mechanism for elevations in PL activity during reversal. Understanding contingency reversal in PL mechanistically at the single-cell level is an important step towards future activations of this circuitry to promote cognitive flexibility in disorders with cognitive flexibility impairments, including substance use disorder.

References

Alsiö, J., Phillips, B. U., Sala-Bayo, J., Nilsson, S. R. O., Calafat-Pla, T. C., Rizwand, A., Plumbridge, J. M., López-Cruz, L., Dalley, J. W., Cardinal, R. N., Mar, A. C., & Robbins, T. W. (2019). Dopamine D2-like receptor stimulation blocks negative feedback in visual and spatial reversal learning in the rat: Behavioural and computational evidence. *Psychopharmacology*, *236*(8), 2307–2323. https://doi.org/10.1007/s00213-019-05296-y

Anastasiades, P. G., & Carter, A. G. (2021). Circuit organization of the rodent medial prefrontal cortex. *Trends in Neurosciences*, *44*(7), 550–563. https://doi.org/10.1016/j.tins.2021.03.006

Baker, P. M., & Ragozzino, M. E. (2014). The prelimbic cortex and subthalamic nucleus contribute to cue-guided behavioral switching. *Neurobiology of Learning and Memory*, *107*, 65–78. https://doi.org/10.1016/j.nlm.2013.11.006

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F. F., & Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature*, *585*(7824), 245–250. https://doi.org/10.1038/s41586-020-2704-z

Bari, B. A., Grossman, C. D., Lubin, E. E., Rajagopalan, A. E., Cressy, J. I., & Cohen, J. Y. (2019). Stable representations of decision variables for flexible behavior. *Neuron*, *103*(5), 922-933.e7. https://doi.org/10.1016/j.neuron.2019.06.001

Beaulieu, J.-M., & Gainetdinov, R. R. (2011). The Physiology, Signaling, and Pharmacology of Dopamine Receptors. *Pharmacological Reviews*, *63*(1), 182–217. https://doi.org/10.1124/pr.110.002642

Biró, S., Lasztóczi, B., & Klausberger, T. (2019). A Visual Two-Choice Rule-Switch Task for Head-Fixed Mice. *Frontiers in Behavioral Neuroscience*, *13*, 119. https://doi.org/10.3389/fnbeh.2019.00119

Birrell, J. M., & Brown, V. J. (2000). Medial Frontal Cortex Mediates Perceptual Attentional Set Shifting in the Rat. *The Journal of Neuroscience*, *20*(11), 4320–4324. https://doi.org/10.1523/JNEUROSCI.20-11-04320.2000

Bissonette, G. B., & Roesch, M. R. (2015). Neural correlates of rules and conflict in medial prefrontal cortex during decision and feedback epochs. *Frontiers in Behavioral Neuroscience*, *9*, 266. https://doi.org/10.3389/fnbeh.2015.00266

Boulougouris, V., Castañé, A., & Robbins, T. W. (2009). Dopamine D2/D3 receptor agonist quinpirole impairs spatial reversal learning in rats: Investigation of D3 receptor involvement in persistent behavior. *Psychopharmacology*, *202*(4), 611–620. https://doi.org/10.1007/s00213-008-1341-2

Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, *179*(2), 219–228. https://doi.org/10.1016/j.bbr.2007.02.005

Brigman, J. L., & Rothblat, L. A. (2008). Stimulus specific deficit on visual reversal learning after lesions of medial prefrontal cortex in the mouse. *Behavioural Brain Research*, *187*(2), 405–410. https://doi.org/10.1016/j.bbr.2007.10.004

Bussey, T. J., Muir, J. L., Everitt, B. J., & Robbins, T. W. (1997). Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behavioral Neuroscience*, *111*(5), 920–936. https://doi.org/10.1037/0735-7044.111.5.920

Chudasama, Y., & Robbins, T. W. (2003). Dissociable Contributions of the Orbitofrontal and Infralimbic Cortex to Pavlovian Autoshaping and Discrimination Reversal Learning: Further Evidence for the Functional Heterogeneity of the Rodent Frontal Cortex. *The Journal of Neuroscience*, *23*(25), 8771–8780. https://doi.org/10.1523/JNEUROSCI.23-25-08771.2003

Dalton, G. L., Wang, N. Y., Phillips, A. G., & Floresco, S. B. (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *Journal of Neuroscience*, *36*(6), 1996–2006. https://doi.org/10.1523/JNEUROSCI.3366-15.2016

De Bruin, J. P. C., Feenstra, M. G. P., Broersen, L. M., Van Leeuwen, M., Arens, C., De Vries, S., & Joosten, R. N. J. M. A. (2000). Role of the prefrontal cortex of the rat in learning and decision making: Effects of transient inactivation. In *Progress in Brain Research* (Vol. 126, pp. 103–113). Elsevier. https://doi.org/10.1016/S0079-6123(00)26010-X

Dembrow, N. C., Chitwood, R. A., & Johnston, D. (2010). Projection-Specific Neuromodulation of Medial Prefrontal Cortex Neurons. *The Journal of Neuroscience*, *30*(50), 16922–16937. https://doi.org/10.1523/JNEUROSCI.3644-10.2010

Descarries, L., Lemay, B., Doucet, G., & Berger, B. (1987). Regional and laminar density of the dopamine innervation in adult rat cerebral cortex. *Neuroscience*, *21*(3), 807–824. https://doi.org/10.1016/0306-4522(87)90038-8

DeSteno, D. A., & Schmauss, C. (2009). A ROLE FOR DOPAMINE D2 RECEPTORS IN REVERSAL LEARNING. *Neuroscience*, *162*(1), 118–127. https://doi.org/10.1016/j.neuroscience.2009.04.052

Durstewitz, D., Vittoz, N. M., Floresco, S. B., & Seamans, J. K. (2010). Abrupt Transitions between Prefrontal Neural Ensemble States Accompany Behavioral Transitions during Rule Learning. *Neuron*, *66*(3), 438–448. https://doi.org/10.1016/j.neuron.2010.03.029

Faustino, B., Oliveira, J., & Lopes, P. (2021). Diagnostic precision of the Wisconsin Card Sorting Test in assessing cognitive deficits in substance use disorders. *Applied Neuropsychology. Adult*, *28*(2), 165–172. https://doi.org/10.1080/23279095.2019.1607737

Fleck, D. E., Shear, P. K., Madore, M., & Strakowski, S. M. (2008). Wisconsin Card Sorting Test performance in bipolar disorder: Effects of mood state and early course. *Bipolar Disorders*, *10*(4), 539–545. https://doi.org/10.1111/j.1399-5618.2008.00582.x

Floresco, S. B., Block, A. E., & Tse, M. T. L. (2008). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behavioural Brain Research*, *190*(1), 85–96. https://doi.org/10.1016/j.bbr.2008.02.008

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. L. (2006). Multiple Dopamine Receptor Subtypes in the Medial Prefrontal Cortex of the Rat Regulate Set-Shifting. *Neuropsychopharmacology*, *31*(2), Article 2. https://doi.org/10.1038/sj.npp.1300825

Gabbott, P. L. A., Warner, T. A., Jays, P. R. L., Salway, P., & Busby, S. J. (2005). Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. *Journal of Comparative Neurology*, *492*(2), 145–177. https://doi.org/10.1002/cne.20738

Gee, S., Ellwood, I., Patel, T., Luongo, F., Deisseroth, K., & Sohal, V. S. (2012). Synaptic Activity Unmasks Dopamine D2 Receptor Modulation of a Specific Class of Layer V Pyramidal Neurons in Prefrontal Cortex. *The Journal of Neuroscience*, *32*(14), 4959–4971. https://doi.org/10.1523/JNEUROSCI.5835-11.2012

Gläscher, J., Adolphs, R., & Tranel, D. (2019). Model-based lesion mapping of cognitive control using the Wisconsin Card Sorting Test. *Nature Communications*, *10*(1), 20. https://doi.org/10.1038/s41467-018-07912-5

Guise, K. G., & Shapiro, M. L. (2017). Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron*, *94*(1), 183-192.e8. https://doi.org/10.1016/j.neuron.2017.03.011

Haddon, J. E., & Killcross, S. (2006). Prefrontal Cortex Lesions Disrupt the Contextual Control of Response Conflict. *The Journal of Neuroscience*, *26*(11), 2933–2940. https://doi.org/10.1523/JNEUROSCI.3243-05.2006

Harley, C. W. (2004). Norepinephrine and dopamine as learning signals. *Neural Plasticity*, *11*(3–4), 191–204. https://doi.org/10.1155/np.2004.191

Heilbronner, S. R., Rodriguez-Romaguera, J., Quirk, G. J., Groenewegen, H. J., & Haber, S. N. (2016). Circuit based cortico-striatal homologies between rat and primate. *Biological Psychiatry*, *80*(7), 509–521. https://doi.org/10.1016/j.biopsych.2016.05.012

Henry, J. D. (2006). A meta-analytic review of Wisconsin Card Sorting Test and verbal fluency performance in obsessive-compulsive disorder. *Cognitive Neuropsychiatry*, *11*(2), 156–176. https://doi.org/10.1080/13546800444000227

Hervig, M. E., Fiddian, L., Piilgaard, L., Božič, T., Blanco-Pozo, M., Knudsen, C., Olesen, S. F., Alsiö, J., & Robbins, T. W. (2020). Dissociable and Paradoxical Roles of Rat Medial and Lateral Orbitofrontal Cortex in Visual Serial Reversal Learning. *Cerebral Cortex*, *30*(3), 1016–1029. https://doi.org/10.1093/cercor/bhz144

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, *345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Izquierdo, A., Wiedholz, L. M., Millstein, R. A., Yang, R. J., Bussey, T. J., Saksida, L. M., & Holmes, A. (2006). Genetic and dopaminergic modulation of reversal learning in a touchscreen-based operant procedure for mice. *Behavioural Brain Research*, *171*(2), 181–188. https://doi.org/10.1016/j.bbr.2006.03.029

Karlsson, M. P., Tervo, D. G. R., & Karpova, A. Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science (New York, N.Y.)*, *338*(6103), 135–139. https://doi.org/10.1126/science.1226518

Klein-Flügge, M. C., Bongioanni, A., & Rushworth, M. F. S. (2022). Medial and orbital frontal cortex in decision-making and flexible behavior. *Neuron*, *110*(17), 2743–2770. https://doi.org/10.1016/j.neuron.2022.05.022

Kopp, B., Lange, F., & Steinke, A. (2021). The Reliability of the Wisconsin Card Sorting Test in Clinical Practice. *Assessment*, *28*(1), 248–263. https://doi.org/10.1177/1073191119866257

Kosaki, Y., & Watanabe, S. (2012). Dissociable roles of the medial prefrontal cortex, the anterior cingulate cortex, and the hippocampus in behavioural flexibility revealed by serial reversal of three-choice discrimination in rats. *Behavioural Brain Research*, *227*(1), 81–90. https://doi.org/10.1016/j.bbr.2011.10.039

Kruzich, P. J., & Grandy, D. K. (2004). Dopamine D2receptors mediate two-odor discrimination and reversal learning in C57BL/6 mice. *BMC Neuroscience*, *5*(1), 12. https://doi.org/10.1186/1471-2202-5-12

Kruzich, P. J., Mitchell, S. H., Younkin, A., & Grandy, D. K. (2006). Dopamine D2 receptors mediate reversal learning in male C57BL/6J mice. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(1), 86–90. https://doi.org/10.3758/CABN.6.1.86

Landry, O., & Al-Taie, S. (2016). A Meta-analysis of the Wisconsin Card Sort Task in Autism. *Journal of Autism and Developmental Disorders*, *46*(4), 1220–1235. https://doi.org/10.1007/s10803-015-2659-3

Latif-Hernandez, A., Shah, D., Ahmed, T., Lo, A. C., Callaerts-Vegh, Z., Van der Linden, A., Balschun, D., & D'Hooge, R. (2016). Quinolinic acid injection in mouse medial prefrontal cortex affects reversal learning abilities, cortical connectivity and hippocampal synaptic plasticity. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep36489

Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? *eNeuro*, *5*(5), ENEURO.0315-18.2018. https://doi.org/10.1523/ENEURO.0315-18.2018

Laughlin, R. E., Grant, T. L., Williams, R. W., & Jentsch, J. D. (2011). Genetic Dissection of Behavioral Flexibility: Reversal Learning in Mice. *Biological Psychiatry*, *69*(11), 1109–1116. https://doi.org/10.1016/j.biopsych.2011.01.014

Le Moine, C., & Gaspar, P. (1998). Subpopulations of cortical GABAergic interneurons differ by their expression of D1 and D2 dopamine receptor subtypes. *Molecular Brain Research*, *58*(1), 231–236. https://doi.org/10.1016/S0169-328X(98)00118-1

Levy, R. (2024). The prefrontal cortex: From monkey to man. *Brain*, *147*(3), 794–815. https://doi.org/10.1093/brain/awad389

Linden, J., James, A. S., McDaniel, C., & Jentsch, J. D. (2018). Dopamine D2 Receptors in Dopaminergic Neurons Modulate Performance in a Reversal Learning Task in Mice. *Eneuro*, *5*(1), ENEURO.0229-17.2018. https://doi.org/10.1523/ENEURO.0229-17.2018

Malagon-Vina, H., Ciocchi, S., Passecker, J., Dorffner, G., & Klausberger, T. (2018). Fluid network dynamics in the prefrontal cortex during multiple strategy switching. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-017-02764-x

Marton, T. F., Seifikar, H., Luongo, F. J., Lee, A. T., & Sohal, V. S. (2018). Roles of Prefrontal Cortex and Mediodorsal Thalamus in Task Engagement and Behavioral Flexibility. *Journal of Neuroscience*, *38*(10), 2569–2578. https://doi.org/10.1523/JNEUROSCI.1728-17.2018

McAllister, K. A. L., Mar, A. C., Theobald, D. E., Saksida, L. M., & Bussey, T. J. (2015). Comparing the effects of subchronic phencyclidine and medial prefrontal cortex dysfunction on cognitive tests relevant to schizophrenia. *Psychopharmacology*, *232*(21), 3883–3897. https://doi.org/10.1007/s00213-015-4018-7

Miles, S., Howlett, C. A., Berryman, C., Nedeljkovic, M., Moseley, G. L., & Phillipou, A. (2021). Considerations for using the Wisconsin Card Sorting Test to assess cognitive flexibility. *Behavior Research Methods*, *53*(5), 2083–2091. https://doi.org/10.3758/s13428-021-01551-3

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(19), 7733–7741. https://doi.org/10.1523/JNEUROSCI.21-19-07733.2001

Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, *12*(2), 239–260. https://doi.org/10.1016/0023-9690(81)90020-5

Mukherjee, A., & Caroni, P. (2018). Infralimbic cortex is required for learning alternatives to prelimbic promoted associations through reciprocal connectivity. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-05318-x

Müller, M. (2002). Microdialysis. *BMJ : British Medical Journal*, *324*(7337), 588–591.

Nakayama, H., Ibañez-Tallon, I., & Heintz, N. (2018). Cell-Type-Specific Contributions of Medial Prefrontal Neurons to Flexible Behaviors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(19), 4490–4504. https://doi.org/10.1523/JNEUROSCI.3537-17.2018

Namboodiri, V. M. K., Hobbs, T., Trujillo-Pisanty, I., Simon, R. C., Gray, M. M., & Stuber, G. D. (2021). Relative salience signaling within a thalamo-orbitofrontal circuit governs learning rate. *Current Biology*, *31*(23), 5176-5191.e5. https://doi.org/10.1016/j.cub.2021.09.037

Otani, S., Blond, O., Desce, J.-M., & Crépel, F. (1998). Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex. *Neuroscience*, *85*(3), 669–676. https://doi.org/10.1016/S0306-4522(97)00677-5

Othman, M. Z., Hassan, Z., & Che Has, A. T. (2022). Morris water maze: A versatile and pertinent tool for assessing spatial learning and memory. *Experimental Animals*, *71*(3), 264–280. https://doi.org/10.1538/expanim.21-0120

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017).

Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), 103–107. https://doi.org/10.1038/nature21376

Oualian, C., & Gisquet-Verrier, P. (2010). The differential involvement of the prelimbic and infralimbic cortices in response conflict affects behavioral flexibility in rats trained in a new automated strategy-switching task. *Learning & Memory*, *17*(12), 654–668. https://doi.org/10.1101/lm.1858010

Padoa-Schioppa, C. (2007). Orbitofrontal Cortex and the Computation of Economic Value. *Annals of the New York Academy of Sciences*, *1121*(1), 232–253. https://doi.org/10.1196/annals.1401.011

Paxinos, G., & Franklin, K. B. J. (2019). *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates*. Academic Press.

Powell, N. J., & Redish, A. D. (2016). Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms12830

Premereur, E., Janssen, P., & Vanduffel, W. (2018). Functional MRI in Macaque Monkeys during Task Switching. *The Journal of Neuroscience*, *38*(50), 10619–10630. https://doi.org/10.1523/JNEUROSCI.1539-18.2018

Rady, A., Elsheshai, A., Abou el Wafa, H., & Elkholy, O. (2012). WCST Performance in Schizophrenia and Severe Depression with Psychotic Features. *ISRN Psychiatry*, *2012*, 373748. https://doi.org/10.5402/2012/373748

Ragozzino, M. E. (2002). The Effects of Dopamine D1 Receptor Blockade in the Prelimbic–Infralimbic Areas on Behavioral Flexibility. *Learning & Memory*, *9*(1), 18–28. https://doi.org/10.1101/lm.45802

Ragozzino, M. E., Detrick, S., & Kesner, R. P. (1999). Involvement of the Prelimbic–Infralimbic Areas of the Rodent Prefrontal Cortex in Behavioral Flexibility for Place and Response

Learning. *Journal of Neuroscience*, *19*(11), 4585–4594. https://doi.org/10.1523/JNEUROSCI.19-11-04585.1999

Ragozzino, M. E., Kim, J., Hassert, D., Minniti, N., & Kiang, C. (2003). The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. *Behavioral Neuroscience*, *117*(5), 1054–1065. https://doi.org/10.1037/0735-7044.117.5.1054

Reinert, S., Hübener, M., Bonhoeffer, T., & Goltstein, P. M. (2021). Mouse prefrontal cortex represents learned rules for categorization. *Nature*, *593*(7859), Article 7859. https://doi.org/10.1038/s41586-021-03452-z

Rich, E. L., & Shapiro, M. (2009). Rat Prefrontal Cortical Neurons Selectively Code Strategy Switches. *Journal of Neuroscience*, *29*(22), 7208–7219. https://doi.org/10.1523/JNEUROSCI.6068-08.2009

Rich, E. L., & Shapiro, M. L. (2007). Prelimbic/Infralimbic Inactivation Impairs Memory for Multiple Task Switches, But Not Flexible Selection of Familiar Tasks. *Journal of Neuroscience*, *27*(17), 4747–4755. https://doi.org/10.1523/JNEUROSCI.0369-07.2007

Romine, C. B., Lee, D., Wolfe, M. E., Homack, S., George, C., & Riccio, C. A. (2004). Wisconsin Card Sorting Test with children: A meta-analytic study of sensitivity and specificity. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *19*(8), 1027–1041. https://doi.org/10.1016/j.acn.2003.12.009

Santana, N., Mengod, G., & Artigas, F. (2009). Quantitative Analysis of the Expression of Dopamine D1 and D2 Receptors in Pyramidal and GABAergic Neurons of the Rat Prefrontal Cortex. *Cerebral Cortex*, *19*(4), 849–860. https://doi.org/10.1093/cercor/bhn134

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, *10*(12), 885–892. https://doi.org/10.1038/nrn2753

Seamans, J. K., Durstewitz, D., Christie, B. R., Stevens, C. F., & Sejnowski, T. J. (2001). Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. *Proceedings of the National Academy of Sciences*, *98*(1), 301–306. https://doi.org/10.1073/pnas.98.1.301

Seamans, J. K., Floresco, S. B., & Phillips, A. G. (1995). Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. *Behavioral Neuroscience*, *109*(6), 1063–1073. https://doi.org/10.1037//0735-7044.109.6.1063

Seamans, J. K., Gorelova, N., Durstewitz, D., & Yang, C. R. (2001). Bidirectional Dopamine Modulation of GABAergic Inhibition in Prefrontal Cortical Pyramidal Neurons. *Journal of Neuroscience*, *21*(10), 3628–3638. https://doi.org/10.1523/JNEUROSCI.21-10-03628.2001

Sinha, R., & Li, C. S. R. (2007). Imaging stress- and cue-induced drug and alcohol craving: Association with relapse and clinical implications. *Drug and Alcohol Review*, *26*(1), 25–31. https://doi.org/10.1080/09595230601036960

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., & Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

Trantham-Davidson, H., Neely, L. C., Lavin, A., & Seamans, J. K. (2004). Mechanisms Underlying Differential D1 versus D2 Dopamine Receptor Regulation of Inhibition in Prefrontal Cortex. *Journal of Neuroscience*, *24*(47), 10652–10659. https://doi.org/10.1523/JNEUROSCI.3179-04.2004

Tseng, K. Y., & O'Donnell, P. (2004). Dopamine–Glutamate Interactions Controlling Prefrontal Cortical Pyramidal Cell Excitability Involve Multiple Signaling Mechanisms. *Journal of Neuroscience*, *24*(22), 5131–5139. https://doi.org/10.1523/JNEUROSCI.1021-04.2004

van der Meulen, J. A. J., Joosten, R. N. J. M. A., de Bruin, J. P. C., & Feenstra, M. G. P. (2007). Dopamine and Noradrenaline Efflux in the Medial Prefrontal Cortex During Serial

Reversals and Extinction of Instrumental Goal-Directed Behavior. *Cerebral Cortex*, *17*(6), 1444–1453. https://doi.org/10.1093/cercor/bhl057

Van Eden, C. G., Hoorneman, E. M. D., Buijs, R. M., Matthijssen, M. A. H., Geffard, M., & Uylings, H. B. M. (1987). Immunocytochemical localization of dopamine in the prefrontal cortex of the rat at the light and electron microscopical level. *Neuroscience*, *22*(3), 849–862. https://doi.org/10.1016/0306-4522(87)92964-2

Vincent, S., Khan, Y., & Benes, F. (1993). Cellular distribution of dopamine D1 and D2 receptors in rat medial prefrontal cortex. *The Journal of Neuroscience*, *13*(6), 2551–2564. https://doi.org/10.1523/JNEUROSCI.13-06-02551.1993

Vorhees, C. V., & Williams, M. T. (2006). Morris water maze: Procedures for assessing spatial and related forms of learning and memory. *Nature Protocols*, *1*(2), 848–858. https://doi.org/10.1038/nprot.2006.116

Wang, J., & O'Donnell, P. (2001). D1 Dopamine Receptors Potentiate NMDA-mediated Excitability Increase in Layer V Prefrontal Cortical Pyramidal Neurons. *Cerebral Cortex*, *11*(5), 452–462. https://doi.org/10.1093/cercor/11.5.452

Wang, P. Y., Boboila, C., Chin, M., Higashi-Howard, A., Shamash, P., Wu, Z., Stein, N. P., Abbott, L. F., & Axel, R. (2020). Transient and Persistent Representations of Odor Value in Prefrontal Cortex. *Neuron*, *108*(1), 209-224.e6. https://doi.org/10.1016/j.neuron.2020.07.033

Wang, X., Zhong, P., Gu, Z., & Yan, Z. (2003). Regulation of NMDA Receptors by Dopamine D4 Signaling in Prefrontal Cortex. *The Journal of Neuroscience*, *23*(30), 9852–9861. https://doi.org/10.1523/JNEUROSCI.23-30-09852.2003

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, *17*(8), 513–523. https://doi.org/10.1038/nrn.2016.56

Williams, M. T., Morford, L. L., Wood, S. L., Wallace, T. L., Fukumura, M., Broening, H. W., & Vorhees, C. V. (2003). Developmental D-methamphetamine treatment selectively induces spatial navigation impairments in reference memory in the Morris water maze while sparing working memory. *Synapse (New York, N.Y.)*, *48*(3), 138–148. https://doi.org/10.1002/syn.10159

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, *81*(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483–494. https://doi.org/10.1038/nrn1406

Young, J. J., & Shapiro, M. L. (2009). Double Dissociation and Hierarchical Organization of Strategy Switches and Reversals in the Rat PFC. *Behavioral Neuroscience*, *123*(5), 1028. https://doi.org/10.1037/a0016822

Zheng, P., Zhang, X.-X., Bunney, B. S., & Shi, W.-X. (1999). Opposite modulation of cortical N-methyl-d-aspartate receptor-mediated responses by low and high concentrations of dopamine. *Neuroscience*, *91*(2), 527–535. https://doi.org/10.1016/S0306-4522(98)00604-6

CHAPTER 1: NEW METHODS AND GENERAL EXPERIMENTAL PROTOCOLS

Successful investigation of neural correlates of contingency reversal in prelimbic cortex (PL) required methodological advancements. First, large neural populations needed to be longitudinally trackable across learning. Since it is extremely difficult to reliably track neurons across days using electrophysiological methods (Yuan et al., 2024), optical-based approaches are needed where cells are directly visualized across days and matched using similar morphology and position in the field of view (FOV). The most reliable way to optically track cells across days is via high-quality two-photon imaging (Benninger & Piston, 2013).

PL, although relatively superficial, is not superficial enough to visualize using two-photon microscopy (Theer & Denk, 2006) without an intermediate implant to grant access deeper into the brain. Traditional methods involve gradient-index (GRIN) lenses, which warp images from one surface of the lens to the other, resulting in reliable relay that is marked by notable aberrations in the periphery (Jung et al., 2004; Levene et al., 2004; Meng et al., 2019; Otis et al., 2017; Ziv et al., 2013). These aberrations make it much more difficult to reliably register cells across days. Newer approaches utilizing microprisms also relay two-photon excitation deeper into the brain with reduced aberrations and signal loss (Andermann et al., 2013; Low et al., 2014; Redman et al., 2022; Spellman et al., 2021), but were poorly characterized, not accessible to the deep brain, and never directly compared to GRIN lens counterparts. Thus, the first methodological advancement of this work involved developing a reproducible microprism implant protocol, establishing reliable cross-day tracking, characterizing microprisms applied to PL, and extending the utility of microprism imaging to the deeper brain for applications in future work.

After imaging and tracking cells, it is time to evaluate their coding properties. Generally, coding is evaluated in one of two ways: *encoding*, which assesses whether task-related predictors including behavior can predict cellular activity, or *decoding*, which assesses whether cellular activity can predict task parameters. Since task predictors can covary, for example licking when sucrose delivery occurs, they can contaminate the results of decoding analysis where the neural

activity could explain either or both of the behavioral parameters. In contrast, encoding analysis can test the individual contribution of each task parameter to the observed neural signal. Generalized Linear Models (GLMs) are one way to perform encoding analysis. Building a predictive model with all predictors, and comparing how much worse the model becomes when any individual predictor is removed evaluates the importance of each task parameter to recorded neural activity (Steinmetz et al., 2019). The importance of a given parameter is assumed to be proportional to how much worse the predictive model is able to reproduce the neural data without its inclusion. However, thresholds of exactly how much worse model performance following a predictor omission must reach to be meaningful can bias analysis of coding proportion within brain areas (Ottenheimer et al., 2023). Adapting GLM analysis for standard F-statistic based significance testing can ground encoding results to traditional methods in neuroscience and increase reproducibility across labs, brain areas, and experiments. Thus, the goal of the second methodological advancement in this work was to derive a significance testing algorithm for GLM analysis.

Finally, correlates of reversal learning in the brain cannot be assessed without a reversal learning task, nor can the work in this dissertation be reproduced without a description of general experimental methods. Consequently, the purpose of the final section of this chapter is detail the general experimental methods that produced the data in the following chapters.

Thus, this chapter is subdivided into three sections:

I. Microprism Development for Improved Deep Brain Recordings (*a manuscript reprint*)

II. Adapting Generalized Linear Models for Significance Testing

III. General Experimental Methods

I. MICROPRISM DEVELOPMENT FOR IMPROVED DEEP BRAIN RECORDINGS

The following is a reprint of (Hjort et al., 2024) with minimal formatting adjustments

INTRODUCTION

Understanding how neural ensemble dynamics evolve over time is crucial for grasping the mechanisms of neural and behavioral processes. Longitudinal identification of single cells profiled during these processes using activity alone is challenging when neural activity is anticipated to change over time. To address this, optical imaging approaches, such as those capturing calcium-induced fluorescence changes, align cell morphology and field position instead of activity to consistently track cell activity over days to weeks. A prominent optical method that both visualizes and records cellular activity is fluorescence microscopy of genetically-encoded voltage, neuromodulator, or calcium indicators like GCaMP (Chen et al., 2013). Since somatic calcium influx occurs coincident with neural action potentials (Hodgkin & Huxley, 1952), the rise in GCaMP fluorescence can serve as an indicator for neural activity. Additionally, this fluorescence can be averaged throughout a session to capture a spatial 'footprint' of the cells within the field of view (FOV) (Chen et al., 2013).

Voltage, neuromodulator, and calcium indicators can be excited by either single or multiphoton wavelengths of light (Chen et al., 2013). Notably, multiphoton imaging results in sharper cell FOVs, typically facilitating more accurate cell tracking. However, a limitation of two-photon microscopy is its shallow penetration depth; it can only reach a few hundred microns from the brain surface, restricting visualization to the upper cortical layers without additional tradeoffs including reduced resolution and yield (Theer & Denk, 2006), or use of hyper-specialized dyes (Tischbirek et al., 2017) or equipment (Lu et al., 2017; Wang et al., 2015; Yildirim et al., 2019).

Several approaches have been developed to circumvent this limitation, most popularly using a gradient-index (GRIN) lens chronically implanted above the population of interest to relay multiphoton excitation deep into the brain (Jung et al., 2004; Levene et al., 2004; Meng et al., 2019; Otis et al., 2017; Ziv et al., 2013). As GRIN lenses employ continuous changes in refractive

index across the lens to focus samples from one end of the lens to the other, relatively minor changes in microscope working distance can translate to accurate focus of cell populations even if they are deep within the brain (Meng et al., 2019). These benefits are not without their tradeoffs, namely non-uniform excitation and fluorescence collection within the FOV, and notable warping in visualized cellular morphology around the lens periphery (Xu et al., 2020). Recently, the use of microprisms has emerged as an alternative to GRIN lenses for two-photon imaging in the cortex and superficial hippocampus (Andermann et al., 2013; Low et al., 2014; Redman et al., 2022; Spellman et al., 2021). Since microprisms utilize linear optics, excitation and morphology visualization are uniform across the field of view, but microscopes must be able to travel the microprism length to visualize the population of interest without colliding with the microprism or head-fixation device. Given the limited working distance (WD) of objectives for standard microscopy, microprisms have been somewhat incompatible for imaging below the superficial cortical surface, leaving many brain structures out of reach. However, recent advances in microscope objective development have produced objectives with compatible working distances (20 mm) for subcortical microprism imaging (Yu et al., 2023). Here we detail and characterize a protocol for imaging the deep brain using microprisms, affording stable, longitudinal tracking of neuronal activity across multiple days over a large FOV. Altogether, this microprism two-photon imaging approach represents a substantial increase in resolution and throughput of trackable cells over days to weeks.

METHODS

2.1 Mice

All cortical GRIN and microprism imaging took place in C57BL6/J (7) or *CaMKIIa-tta x Teto-GCamp6s* (1) mice. DS microprism imaging occurred in D1-cre x Ai14 mice (4), and VS microprism imaging in *Vgat*-Cre mice. All procedures outlined were approved by UW IACUC (Protocol # 4450-01). See Legaria et al.(Legaria et al., 2022) for additional details for DS GRIN

lens mice, though in brief *Drd1*-cre x Ai14 mice (4) were injected with AAV2/9-CaMKII-GCaMP6s and implanted with a 1x4mm GRIN lens (Legaria et al., 2022).

2.2 Microprisms

All microprisms were manufactured by and purchased from OptoSigma. The 1.5x1.5x3 mm microprisms used for cortical imaging are the same as reported by Spellman et al. (Spellman et al., 2021) (OS PN 160712BK01), while the 1.5x1.5x8 mm microprisms were custom designed (OS PN 22031003567). To facilitate easier implant of the cortical prisms and to provide a larger surface area for glue, we attached the 1.5x1.5x3 mm prisms to an 8mm #1 coverslip (Harvard Apparatus 64-0701) using optical glue (NOA68) applied via a 50cc syringe. As NOA68 is viscous and difficult to load into the syringe, we removed the plunger, squeezed about 0.1mL into the syringe in the space vacated by the plunger, and then reinserted the plunger to apply one drop of glue to the coverslip. We also attached the coverslip to a metal head-fixation ring (Gordon-Fennell et al., 2023). Since our 1.5x1.5x8 mm microprisms had sufficient surface area above the skull to apply glue, we omitted the coverslip in preparations involving that implant. We suggest leaving at least 2 mm of additional microprism length above the surface of the brain to allow for adequate fixation to the skull for any custom designs.

2.3 Surgical preparation

We performed surgery following our previous protocol (Resendez et al., 2016) with the following modifications (**Fig. 1-1**). Please refer to Resendez et al. for more information on specific surgical equipment, viral dilution testing, and simple troubleshooting steps. A more detailed protocol for this specific surgery, including images, is available upon reasonable request.

We induced anesthesia with 5% isoflurane, headfixed mice on a stereotax, and maintained anesthesia with <1.5% isoflurane. Anesthesia was verified by lack of a toe pinch response (Resendez et al., 2016). We performed an incision with surgical scissors or a #10 scalpel blade to expose the skull. Once the surface of the skull was exposed (**Fig. 1-1A**), we marked the corners of our square craniotomy by gently touching the stereotax-mounted drill bit to the surface

of the skull. Refer to **Table 1** for “craniotomy corners.” We also drilled pilot holes for #00-90 skull screws (B010MMSJJO, Amazon) located over the cerebellum (2x) and olfactory bulb (1x).

We hand-drilled our craniotomy until the drilled bone easily flexed when we gently tapped on the center piece of bone. Following application of sterile saline, we removed the center bone piece using #5 forceps. (**Fig. 1-1B**) After extracting any bone shards, we applied hemostatic sponge (Goodwill AGS111) pre-wetted with sterile saline to control any remaining bleeding. We then inserted skull screws into the drilled pilot holes (**Fig. 1-1C**). If the dura did not release when the craniotomy was exposed, we made a superficial incision using a microscalpel (Graham-Field 2979#30) and carefully removed it using a micro curette (FST 10080-05) (**Fig. 1-1D**). We applied additional saline-wetted hemostatic sponge to control bleeding if necessary.

For prefrontal cortex surgeries, we injected 4x400nL of AAVDJ-CaMKIIa-GCaMP6s (UNC Vector Core lot av6364) at the locations indicated in **Table 1**. Mice for experiments shown in **Figure 1-4** received a 3:1 mix of GCaMP6s and AAV8-CAMKIIa-ChRmine(Marshel et al., 2019). For striatal surgeries, we injected 4x300nL of AAVDJ-hSyn-GCaMP6s (Stanford Vector Core) at the locations indicated in **Table 1-1**. All injections were performed using a Nanoject II (Drummond) at a rate of 1 nL/s. For injections performed at the same ML coordinates, we did not pause between the ventral and dorsal injections, but waited at least 15 min between the dorsal injection and removal of the glass pipette from the brain (**Fig. 1-1E**).

Following all virus injections, brain surface was cleaned until it was free of active bleeds and/or blood clots near the site of the microprism face implant. Active bleeding or blood clots near the implant site can travel with the microprism to the imaging FOV and cause obstructions. Aspiration or pre-lowering of a needle into the craniotomy was not necessary for any of our applications, and we did not observe obvious indicators of tissue compaction in post-recording histology (**Fig. 1-2A**).

We held the 1.5x1.5x8 mm microprism with a surgical bulldog clamp (WPI 14119) covered with heat shrink tubing (Qualtec Qkit 1) on a stereotaxic electrode holder (Kopf, Model 1773). We

held the 1.5x1.5x3 mm microprism apparatus using a miniscope clamp (Inscopix), though other methods are also likely viable.

We measured our microprism implant coordinates at Bregma from the medial caudal corner of the microprism, as we found this easiest to visualize and reliably locate. After positioning the microprism over the desired implant site, we removed any remaining hemostatic sponge, applied additional saline, and lowered the microprism at a maximum of 500 µm/min (**Fig. 1-1F**). If any profuse bleeding occurred, we removed the microprism, stopped the bleeding, and reinserted the microprism.

Once the microprism was at the desired coordinates, we glued (Loctite 234790) it in place and cured the glue using a dental cement activator (Jet 1406). We then removed the lowering clamp and added a layer of C&B Metabond (Parkell S380) to further affix the implant on the skull. Successful implants allow for visualization of brain tissue through the implant.

For 1.5x1.5x8mm implants, we lowered a metal headring (see Gordon-Fennell et al. 2023 for narrow sidewall .stl file) around the implant, ensuring the microprism was slightly recessed by creating a superglue platform to rest the headring as needed. We then secured the headring to the skull using dental cement (Jet 1230-P). To better identify animals, we wrote a number into the dental cement when mostly dry using a permanent marker (Sharpie 1735792) (**Fig. 1-1G**). We protected the microprism by covering it with Silicone Sealant (WPI KWIK-CAST) (**Fig. 1-1H**). Animals recovered for at least two weeks before initial FOV visualization, and at least four weeks before experiments began.

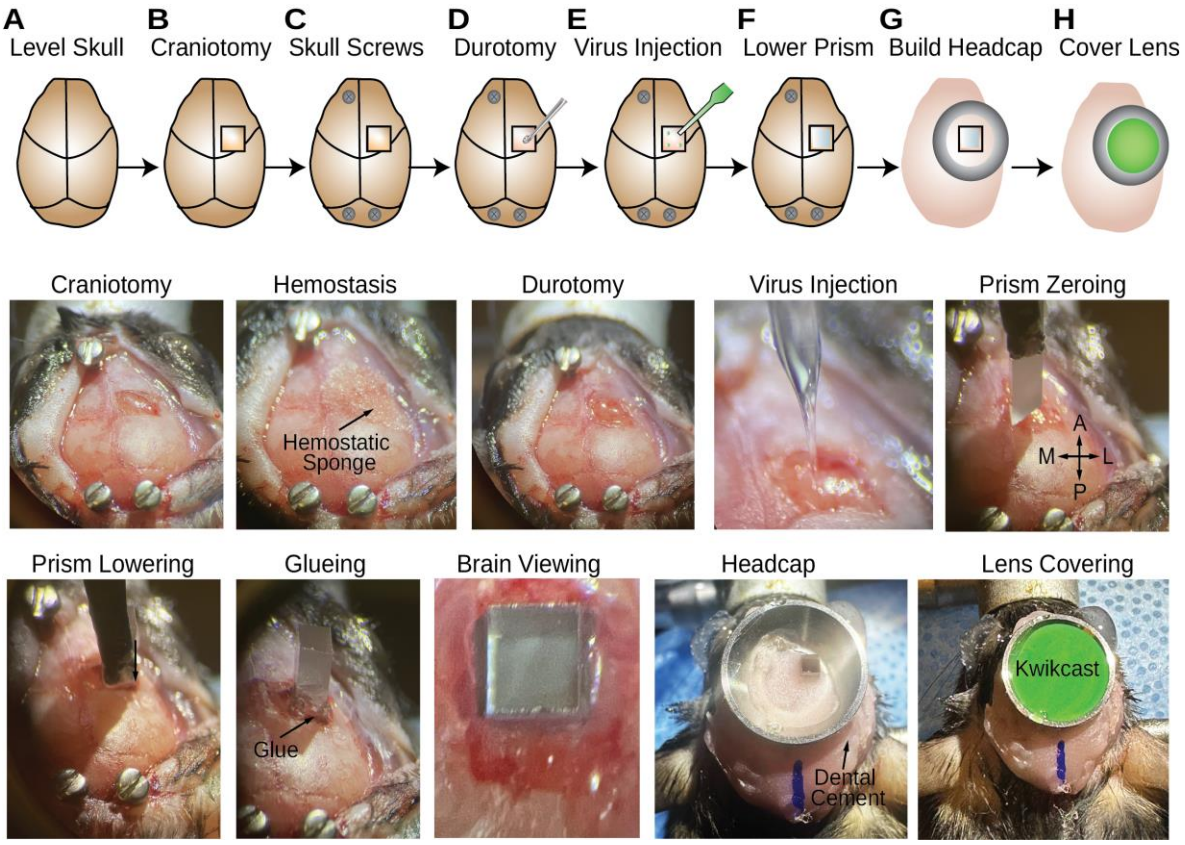

Figure 1-1: Surgical Implantation of Microprism. Schematic of major surgical steps **A**. Begin with cleaned, leveled exposed skull. **B.** Craniotomy with central skull piece removed. **C.** Insertion of skull screws for headcap stability **D.** Durotomy with microcurette **E.** Injection of virus into the craniotomy. **F.** Microprism lowered through craniotomy into target brain area. **G.** Microprism secured to the skull with superglue and dental cement. **H.** Final preparation with microprism, headring, and headcap, and protective covering A detailed protocol with images is available upon request.

Surgical images omitted from publication at Neurophotonics' request are included in this reprint.

Table 1-1: Surgical Coordinates

Microprism	Location	Craniotomy Corners	Virus Injections	Microprism Corner
1.5x1.5x3	Prefrontal Cortex	AP 1.85 and 3.60 ML 0.25 and 2.00	AP 1.94 ML 0.75 and 1.25 DV -2.25 and -1.85	AP 2.05 ML 0.3 DV -2.5
1.5x1.5x8	Dorsal Striatum	AP 0.0 and 1.75 ML 2.0 and 0.25	AP 0.5 and 1.5 ML 1.65 DV -3.25 and -3.0	AP 0.25 ML 1.75 DV -3.5
1.5x1.5x8	Ventral Striatum	AP -0.3 and 2.3 ML 1.65 and 3.3	AP 0.85 and 1.50 ML 1.65 DV -4.0 and -3.25	AP 0.65 ML 1.75 DV -4.25

2.4 *In-vivo* imaging and tracking of neurons visualized through microprisms

We performed all 1.5x1.5x8mm microprism imaging on a Bruker Investigator two-photon microscope through a Cousa objective (20 mm WD) (Yu et al., 2023). All 1.5x1.5x3 mm microprism were imaged on an Olympus FVMPE-RS with an XLPLN10xSVMP (NA0.6, 8 mm WD) immersion objective except for point spread function and stimulation data which took place on a Bruker2p+ using the Cousa and TL10X-2P objectives, respectively, and all cortical GRIN lens imaging on a Bruker 2p+ though a TL10X-2P (NA 0.5, 8 mm WD) objective. On Bruker microscopes we imaged with an InSight X3 Laser at 820 nm for plane registration, and 920 nm for functional GCaMP imaging. Olympus microscopes utilized a MaiTai DeepSee Laser at the same wavelengths. All data was acquired at 7.5 Hz on resonant galvos (4 frame averaging). All GRIN and PFC microprism data was acquired at 512x512 pixel resolution, while DS and VS microprism data was collected at 1024x1024, which allowed higher quality signal in extracted transients. The DS GRIN dataset in **Figure 1-3** was used with permission from the authors of Legaria et al. (Legaria et al., 2022), and was collected using a N16XLWD-PF (0.8 NA, 3mm WD) objective at 30Hz. The PFC GRIN lens dataset used in **Figure 1-3** was previously published by our group (Ottenheimer et al., 2023).

We used an Open-Source Head-fixed Rodent Behavioral Experimental Training System (OHRBETS) (Gordon-Fennell et al., 2023) for all 1.5x1.5x8 mm microprism imaging. To facilitate better leveling of the imaging plane, we mounted the OHRBETS stage on a Thorlabs goniometer platform (TTR001/M). We verified that the microprism was level in three ways: First, after cleaning the microprism of debris on a dissecting microscope, we adjusted the goniometer until the light of the microscope illuminated the microprism evenly (white). Second, when visualizing the microprism using the eyepieces of the two-photon microscope, we ensured that all 4 corners were even and sharply in focus. Third, we ensured a uniform laser flare on the microprism face before lowering the objective to the imaging plane. Imaging with the microprisms, as with GRIN lenses, is possible without the goniometer stage, but in our experience often has a dimmer, lower yield FOV and impaired tracking across days. To register a consistent imaging plane across days, we first imaged at 820 nm (**Fig. 1-2B)**, the approximate isosbestic wavelength at which GCaMP6 is equally fluorescent calcium free and calcium bound (Barnett et al., 2017). Since this image was independent of the activity state of the system, we used 820 nm acquired images to compare resolvable cells across multiple daily imaging sessions during data acquisition.

Following recording, we tracked imaged cells over days by concatenating all imaging sessions and running them through Suite2p (Pachitariu et al., 2017). This both corrected motion within days and small misalignments across days with some modifications from the default. In brief, enabling two_step_registration=1, increasing the snr_threshold (to 2), and increasing maxregshiftNR (to 50) improved automated cross-day registration. We manually verified cells as tracked (i.e. sorted to 'cells') if they had at least one visible transient each recording day.

2.5 Data Analysis

Peri-event time histogram creation: following motion correction, we standardized fluorescence between cells by subtracting and subsequently dividing by the mean of each trace over the entire session. We then generated peri-event time histograms of our data for **Fig. 1-2H**.

Field of view quality: To quantify variations in field fluorescence between GRIN lens and microprisms, we averaged pixel fluorescence in concentric circles from the FOV center. We fit the resulting measurement with a linear model to quantify the relationship between field fluorescence and the distance from the center of the imaging field and used the slope (β) to compare this relationship between recording methods.

Cell mask variance and counts: we isolated the cells within our FOV using Suite2p for PFC GRIN and DS microprism imaging, and manually for PFC microprism and DS GRIN Imaging. We then calculated the area of each mask using custom python scripts. To compare variation in mask size (with the goal of highlighting irregular and stretched cells around the GRIN lens periphery), we calculated the variance of these area measurements, and divided by the mean mask size to standardize across cells of different sizes in PFC and DS GRIN and microprism preps. We also compared the total number of cells visualized in each preparation per unit implant area, and quantified the relationship between cell density and distance from implant center in three concentric circles. All statistical testing utilized balanced Two-Factor ANOVA with replication to compare brain areas (PFC and DS) and implant types (GRIN or prism) or One-way RM ANOVA within a single area, both at $\alpha<0.05$ (Supplementary Table 1). All analysis scripts available on request.

2.6 Microprisms for spatial light modulation applications: We assessed the suitability of microprisms for single cell stimulation and characterized the point spread function (PSF) of the microprisms following previous methods (Piantadosi et al., 2023). All statistical testing utilized Repeated Measures One-Way ANOVA with Tukey Honest Significant Difference test comparing all columns to 0 offset at $\alpha<0.05$

2.7 Code, data, and materials availability: All data and analysis scripts are available from the corresponding author upon reasonable request. Microprisms can be purchased from OptoSigma using OS PN 160712BK01 (3mm) OS PN 22031003567 (8mm).

RESULTS: To determine the capabilities of 1.5x1.5x8 mm microprisms, we implanted the microprism into the dorsal striatum (DS), a subcortical region important for reward processing and movement (**Fig. 1-2A**). After mice recovered from surgery, we leveled the microprisms for imaging (see methods), and used the long working distance Cousa 10x objective (Yu et al., 2023) to descend to the focal plane above the neurons we aimed to visualize (**Fig. 1-2B-E**). Manual alignment under the microscope and automated registration allowed us to track cells in the DS across multiple days (**Fig. 1-2B-D**) even though mice were returned to their home cages at the end of each recording session. We visualized a total of 10,727 stable and trackable cells with this preparation, with 2,681 +/- 516 (Mean +/- SEM) trackable cells per FOV in each of four mice (**Fig. 1-2B-D, F, I**). Extracted activity demonstrates that neurons were active and dynamic on all recording days (**Fig. 1-2G,H**). When aligned to spout extension and sucrose delivery, many cells show a robust increase in activity that is consistent across days (**Fig. 1-2H-J**). Notably, most of these neurons are also trackable over multiple weeks (**Fig. 1-2K-M**). We also implanted the Ventral Striatum (VS) using the same method and we visualized trackable cells within an imaging plane (**Fig. 1-2N-O)**. Taken together, these results illustrate the utility of our preparation for functional, longitudinal imaging of subcortical neural populations.

To characterize the optical quality of our microprism imaging method (**Fig. 1-3A-D**), we compared to a published preparation(Spellman et al., 2021) utilizing a microprism for functional characterization of prefrontal cortical (PFC) ensembles (**Fig. 1-3F-I**). Both preparations have high-quality lateral resolution (~2 μm) that is much lower than the average diameter of neurons in our preparations (~20 μm) (**Fig. 1-3D, I; Fig. 1-4D**), and suggests they may be useful for studying some subcellular processes, though we did not test this directly. On average, we visualized 730 +/- 201 PFC and 2,781 +/- 505 DS neurons in each microprism preparation on a single day (Mean+/-SEM) (**Fig. 1-3M**). Given the high level of spatial specificity, this preparation will be suitable to image most cell types in the brain regions accessible with the microprisms.

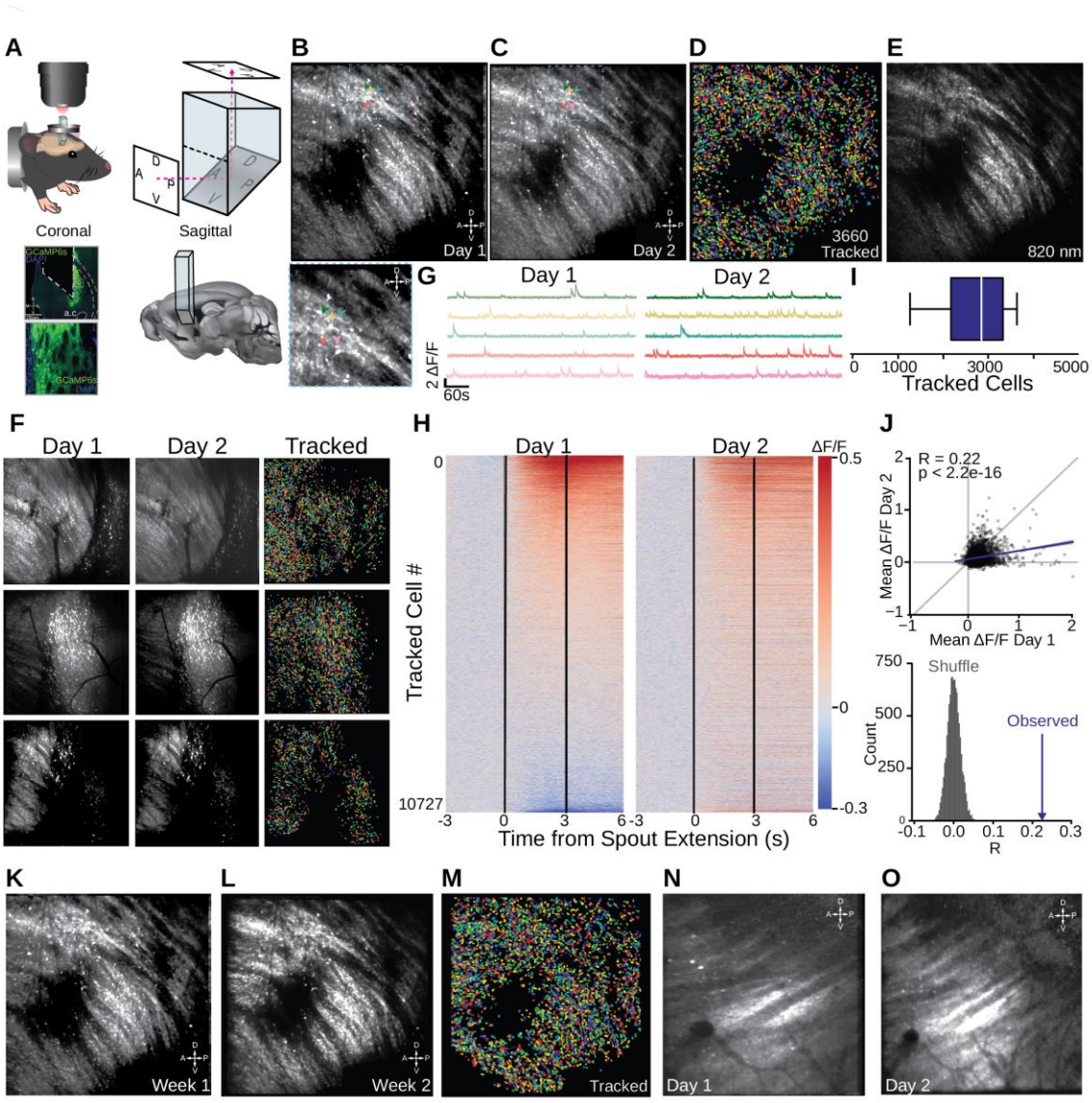

Figure 1-2: Visualization of the Dorsal (DS) and Ventral Striatum (VS) with microprism approach. A. 3D rendering of prism in mouse brain, imaging plan transformation through prism, and track schematics on sagittal sections adapted from Allen CCF, and coronal histological section (10x upper, 40x lower, showing implant track, GCaMP6s and DAPI). **B**. (Upper) Sample Field of View for Representative Mouse on Day 1 (Lower) higher-resolution zoom with cells indicated. **C.** Sample Field of View for Representative Mouse on Day 2 **D.** Tracked ROIs from B and C. **E.** Registration plane at 820nm for activity-independent tracking. **F.** Tracked FOVs from additional mice. **G.** Sample extracted transients for DS cells on Days 1 (left) and 2 (right). Cells are marked in B-C. **H.** Peri-stimulus Time histograms of spout retraction/sucrose-delivery aligned activity from all mice on Day 1 (left) and 2 (right). Vertical lines indicate spout extension and retraction. **I.** Quantification of cells tracked across days in the DS in one FOV for four mice (B-D, F). **J.** Correlation analysis of mean fluorescence in H during access period (0-3 s) across days. **K**. B reproduced. **L.** FOV from B/K tracked over 2 weeks. **M.** Extracted masks from tracking of K and L (3726 cells). **N.** Sample VS FOV on Day 1. **O.** Sample VS FOV on Day 2.

We also compared our microprism recordings to those using industry-standard 1x4mm gradient-index (GRIN) lenses (**Fig. 1-3E, J**) with comparable objectives and zoom (see Methods). Both GRIN lens datasets were sourced from recent literature from our (PFC) and other (DS) groups(Legaria et al., 2022; Ottenheimer et al., 2023). The microprisms facilitate a statistically significant decrease in the relationship between radial distance from lens center and fluorescence, indicating a more uniform brightness across the FOV (**Fig. 1-3K**). Microprisms also produce more consistently sized cell masks (**Fig. 1-3L**). On average, we were able to visualize hundreds of cells through GRIN lenses over a 0.785 mm^2 surface, and thousands of cells across a single plane in our microprism preparation over a 2.25 mm^2 surface (**Fig. 1-3M**). We also evaluated the relationship between cell density and distance from the center of the imaging field in GRIN and microprism preparations. Notably, there is not a significant difference in cell density at the center of GRIN and microprism preparations, though cell density significantly decreases with distance from the center in GRIN recordings (**Fig. 1-3N**). These results indicate that this approach produces a marked improvement in FOV quality and neuron yield in microprism preparations compared to GRIN lens imaging.

Finally, we characterized the suitability of microprism preparations for single cell optogenetics using a spatial light modulator (SLM) to activate the two-photon compatible opsin, ChRmine (Marshel et al., 2019) (**Fig. 1-4A**). The lateral resolution within the visualized plane was within the diameter of a single cell (**Fig. 1-4 D, E-F**), while the Axial (Z) resolution only slightly larger than the diameter of the imaged cells (**Fig. 1-4 C-D, G**). This physiological resolution will further depend on opsin and other biological factors in additional preparations. Notably, we did not observe any cells with a vertical alignment (as would appear by additional peaks in **Fig. 1-4C**) where this z-resolution would stimulate off target cells. Taken together, these results suggest that microprism preparations can be used to both read neural activity and perform holographic spatial light modulation for single cell optogenetics with higher throughput.

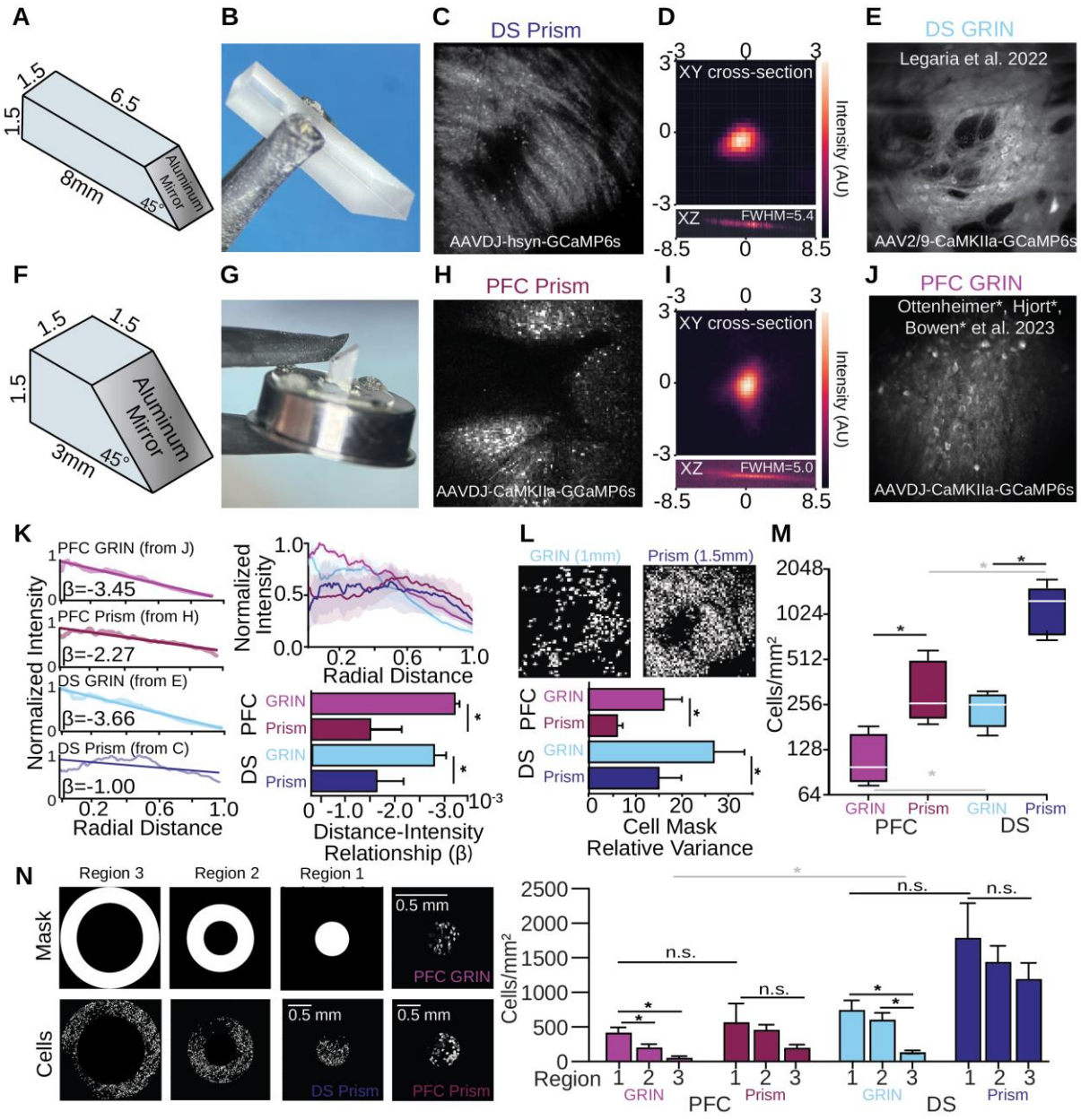

Figure 1-3: Optical properties of microprism preparations and comparison with field standard GRIN lens preparations. *(caption continued on subsequent page)*

A. Schematic of 1.5x1.5x8mm microprism. **B.** Photo of 1.5x1.5x8mm microprism. **C.** Sample FOV from GRIN microprism PFC prep. **D.** PSF for 1.5x1.5x8mm microprism of 100nm bead **E.** Sample FOV from GRIN lens DS prep, reproduced from Legaria et al. 2022[12] with permission **F.** Schematic of 1.5x1.5x3mm microprism (from Spellman et al. 2021[10]). **G.** Photo of 1.5x1.5x3mm microprism. **H.** Sample FOV from microprism PFC prep. **I.** PSF for 1.5x1.5x3mm microprism of 100nm bead. **J.** Sample FOV for GRIN PFC prep. **K.** (**Left**) Intensity-Distance Relationships for example FOVs in C-D and H-I. (**Right, Upper**) Measurement of normalized intensity (% of maximum) from center of lens FOV for PFC GRIN (magenta), PFC prism (maroon), DS GRIN (cyan) and DS prism (indigo) preparations. Mean +/- SEM. (**Right, Lower**) The relationship between distance and intensity (slope of linear fits from K, left) is significantly stronger for GRIN lenses compared to prisms in both the PFC and DS (Two-Way ANOVA, Brain Area: F(1,12)=0.0708, **Implant Type: F(1,12)=9.2387, p=0.01028**, Interaction: F(1,12)=0.4768). **L.** (**Uppe**r) Visualization of cell masks for DS GRIN lens (left, from C) and prism (right, from D). (**Lower**) Quantification of Relative Variance (variance divided by mean to account for different mask size in imaging fields) for GRIN and prism FOVs in DS and PFC. Masks vary more significantly in size in GRIN compared to prism lens preparations. (Two-Way ANOVA, Brain Area: F (1, 12) =4.7125, **Implant Type: F (1, 12) =5.8072, p=0.03292,** Interaction: F (1, 12) =0.0491. **M.** Cell density of GRIN and prism preparations. There is a significant difference between GRIN and prism cell density in PFC and DS, as well as a difference in the number of cells visualized between regions. (Two-Way ANOVA, **Brain Area: F (1, 12) =18.13, p=0.0011 Implant Type: F (1, 12) =24.07, p=0.0004, Interaction: F (1, 12) =10.10, p=0.0080**). **N**. Quantification of cell density across distance from center of GRIN or microprism. Mean +/- SEM There is not a significant difference between cell densities at FOV center across GRIN and microprism (Two-Way ANOVA, **Brain Area: F (1, 12)=6.973, p=0.02155,** Implant Type: F (1, 12) =4.24, p=0.06, Interaction: F (1, 12) =2.4096, p=0.1466), though intensity decreases with distance for GRIN lenses in the PFC (One-Way RM ANOVA, **F(2,3)= 14.6647, p=0.001473**) and DS (**F(2,3)= 11.7578, p=0.003088**). For all comparisons, asterisk = significance at $\alpha<0.05$.

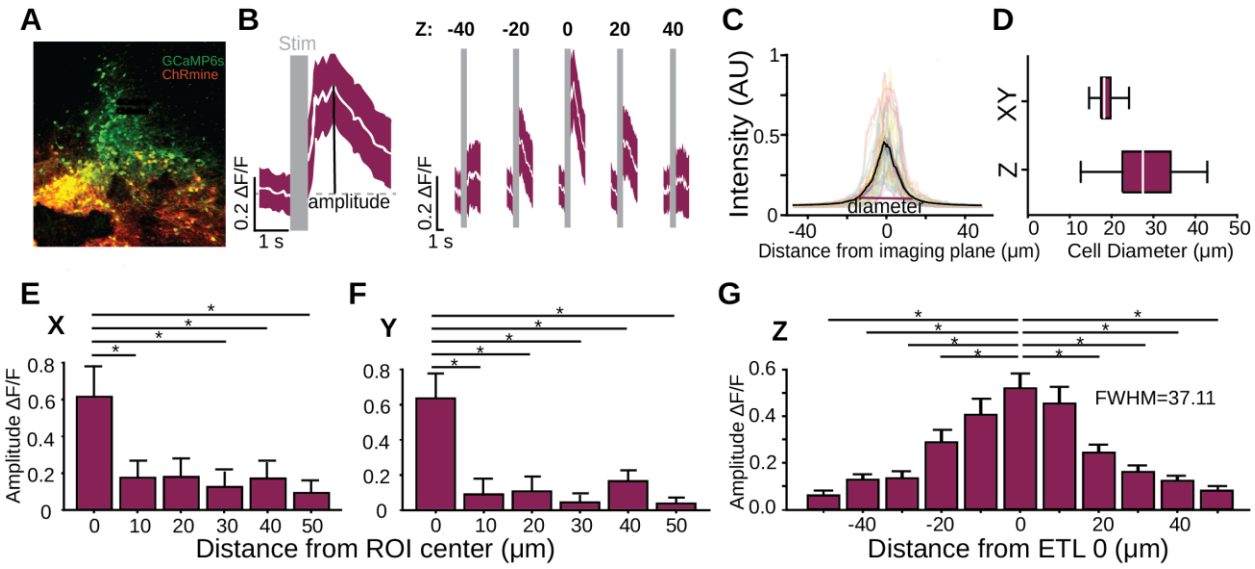

Figure 1-4: Characterizing microprism suitability for single cell stimulation. A. Sample FOV through PFC prism using TL10x. **B. (Left)** Sample evoked transient with stim period and amplitude measurement used for E-G indicated **(Right)** sample stim response at +/- 0, 20, and 40 μm. White mean +/- magenta SEM. **C**. Z-stack through stimulation ROIs (from G) to quantify cell diameter, 95% of peak width indicated and was used for (D). **D.** Quantification of cell dimensions for lateral (XY) and axial (Z) resolutions. **E.** Average evoked signal from stimulation ROI offsets in X plane in increments of 10 μm for 30 cells from 3 animals. Asterisks indicate significant differences from approximate cell centroid at α<0.05 in One-Way repeated measures ANOVA with Tukey HSD test. **X: F (5, 95) = 2.9758, p=0.01538, 0-10 p=0.03371,** 0-20 p=0.1075**, 0-30 p= 0.01084, 0-40 p= 0.03373, 0-50 p= 0.0076534. F.** Like E but in Y plane. **Y: F (5, 95) =10.4351, p=0.000448, 0-10 p= 0.003537, 0-20 p <0.0001, 0-30 p= 0.001098, 0-40 p= 0.005559, 0-50 p= 0.0004289. G.** Average evoked signal from stimulation ROI offsets in Z plane in increments of 10 μm for 30 cells from 3 animals (see B). Asterisks indicate significant differences from ETL 0 at α<0.05 in One-Way repeated measures ANOVA with Tukey HSD test. **Z: F (29, 290) = 17.81, p<0.0001, 0 - -50 p <0.0001, 0 - -40 <0.0001, 0 - -30 p <0.0001, 0 - -20 p= 0.0263**, 0 - -10 p= 0.4793, 0 - 10 p= 0.6582**, 0 - 20 p= 0.0011, 0 - 30 p= 0.0001, 0 - 40 p <0.0001, 0 - 50 p <0.0001.**

DISCUSSION

Here we present a high cell yield imaging approach to visualize deep regions of the brain via two microprisms (3 mm and 8 mm in length). Compared to commonly used GRIN lens approaches, the microprism preparation (**Fig. 1-1**) provides a ~10x increase in cell yield, a more uniform FOV visualization, and a much larger FOV (**Fig. 1-3**). Using these microprisms we were able to track thousands of dynamic neurons across days within a single FOV in the DS, which can be aligned to task events and further analyzed (**Fig. 1-2**). This microprism-mediated preparation achieves the necessary optical resolution for cellular—and perhaps subcellular—imaging (**Fig. 1-3**). Stable FOVs can be tracked over multiple sessions (**Fig. 1-2**) and stimulated using optogenetics (**Fig. 1-4**). Altogether, this technique provides a marked improvement upon canonical preparations to visualize and track single cells in deeper brain structures.

Although the Cousa objective has a larger FOV that could account for differences in cell density yield between GRIN and microprism recordings, it is unlikely it is the main driver of this effect. Moreover, the periphery of the GRIN lens recordings (**Fig. 1-3 E,J**) already display substantial aberrations and limited resolvable ROIs. Increasing the imaging FOV is unlikely to resolve these aberrations that are inherent to the GRIN lens preparation. Furthermore, the increased yield is not attributable only to differences in microprism and GRIN size, as both PFC and DS microprism preparations have a higher density of cells per squared millimeter (**Fig. 1-3M**). Likewise, it is unlikely that there are substantial axial resolution differences that account for the differences we observed. Cell density yield between GRIN and microprism preparations did not significantly differ at the center of the imaging field (**Fig. 1-3N**) and the physiological axial resolution for microprisms (**Fig. 1-4G**) is comparable to published literature measurements for GRIN lenses on the same microscope (Piantadosi et al., 2023). Our results (**Fig. 1-3N**) suggest that this differential yield is attributable to a significant reduction in the density of resolvable cell ROIs near the edges of GRIN lens preparations. While differences in implant size, microscopes objectives used, or axial resolution could impact total cell number, our results suggest that the

improved optics of the microprism, including its more uniform FOV drive the improvements in cell density we report.

The length of our implant is sufficient to image at any depth in the mouse brain, though its large footprint may certainly limit its suitability for some areas. We chose to fabricate a single size that was feasible for recording in any area to allow for ordering in large batches, which in our experience significantly reduces microprism cost. Other users may consider reducing the implant length according to their needs, though an important consideration is to allow sufficient surface area above the brain for gluing or cementing the microprism to the skull if attaching the microprism to a coverslip is undesirable.

This approach can likely be adapted for one-photon miniscope recordings with some modifications to the miniscope apparatus. Since the optics must move a much longer distance to focus down the length of microprisms, compared to GRIN lenses, miniscopes will likely need to adjust the mounting distance between the imaging optics and microprism or the range and position of the internal focus. Ordering microprisms of the correct length for each application, as opposed to a more universal 8 mm length, may help mitigate these risks.

In conclusion, we developed an approach which improves upon canonical approaches to allow high-quality tracking of thousands of neurons across days in the deep brain.

Code, data, and materials availability statement

All data and analysis scripts are available from the corresponding author upon reasonable request. Microprisms can be purchased from OptoSigma using OS PN 160712BK01 (3mm) OS PN 22031003567 (8mm).

Acknowledgements

We would like to thank Alex Legaria, Ben Yang, Alexxai Kravitz and Jones Parker for sharing their DS GRIN lens dataset for our use in **Figure 1-3**. We would also like to thank Kyle Haddock and colleagues at OptoSigma for microprism design assistance, as well as Charles Zhou for assistance with protocols for optical characterization of the microprisms. We would like to thank Sean Piantadosi for comments on this manuscript. This work was supported by F31DA053706 (MH), K99DA059709 (RG), NSF 1934288 & R37DA032750 (GDS), P30DA048736 (GDS & MRB) and R37DA033396 (MRB), and the Weill Neurohub (MRB).

Author Contributions

Conceptualization: MH, GDS, RG, MRB. *Methodology*: microprism and surgical design- MH and GDS. MH, RG, VMKN, and AGF performed surgeries; MH, LT, and AGF took surgery photos. *Investigation*: MH (**Fig. 1-2 N-O, Fig. 1-3 H, J, Fig 4**), RG (**Fig. 1-2 B-F, K-M**), LT (**Fig. 1-3 D, I**). *Data Curation*: MH, RG. *Formal Analysis and Visualization*: MH, AGF (**Fig. 1-2J**), LT (**Fig. 1-3 D, I**). *Writing- Original Draft*: MH. *Writing- Review and Editing*: MH, RG, AGF, VMKN, GDS, MRB. *Funding acquisition*: GDS, MH, RG, MRB. Supervision: GDS, MRB.

Competing interests: The authors have no competing interests to declare.

II. ADAPTING GENERALIZED LINEAR MODELS FOR SIGNIFICANCE TESTING

Neurons often display multivariate coding, so it can be difficult to dissociate the contributions of individual predictors from traces or peri-stimulus maps alone. For example, in a task where an odor cue evokes anticipatory licking towards a future sucrose delivery, activity increases during the cue period could be related to the odor, an abstract cue-reward association, the licking, or any combination thereof. Running independent regressions for any of the variables would suggest a significant contribution to the neural activity that could actually have arisen from one of the other covariates. Especially given widespread representation of motor correlates in cortical regions compared to task-specific representations (Engel & Steinmetz, 2019), it is important to account for all potential variables in encoding models concurrently.

Generalized linear models (GLMs) simultaneously test the relationship between many potential predictors and a given neural signal. Standard packages in many programming languages perform this function and return a single significance measure regarding whether all variables together can meaningfully predict the neural data. While useful as a first pass, the goal of encoding analysis is to determine which (if any) of the individual predictors can uniquely contribute to the modeled signal. One technique to determine the individual contribution of a predictor to the signal is to train the full GLM model (**Fig. 1-5.1**), then train individual models with one predictor removed (**Fig. 1-5.2**) and assess how much worse the leave-one-out model performed compared to the original model in explaining the variance in the neural data (Steinmetz et al., 2019). In this strategy, the explained variance only decreases in the leave-one-out model if another variable cannot also predict the neural signal in place of the left-out variable. This adequately accesses the unique involvement of each individual predictor in the neural signal. However, setting a variance explained threshold to determine whether a predictor's contribution is meaningful can skew the proportion of neurons that display coding of a particular variable (Ottenheimer et al., 2023). Instead, adapting the leave-one-out GLM for standard significance

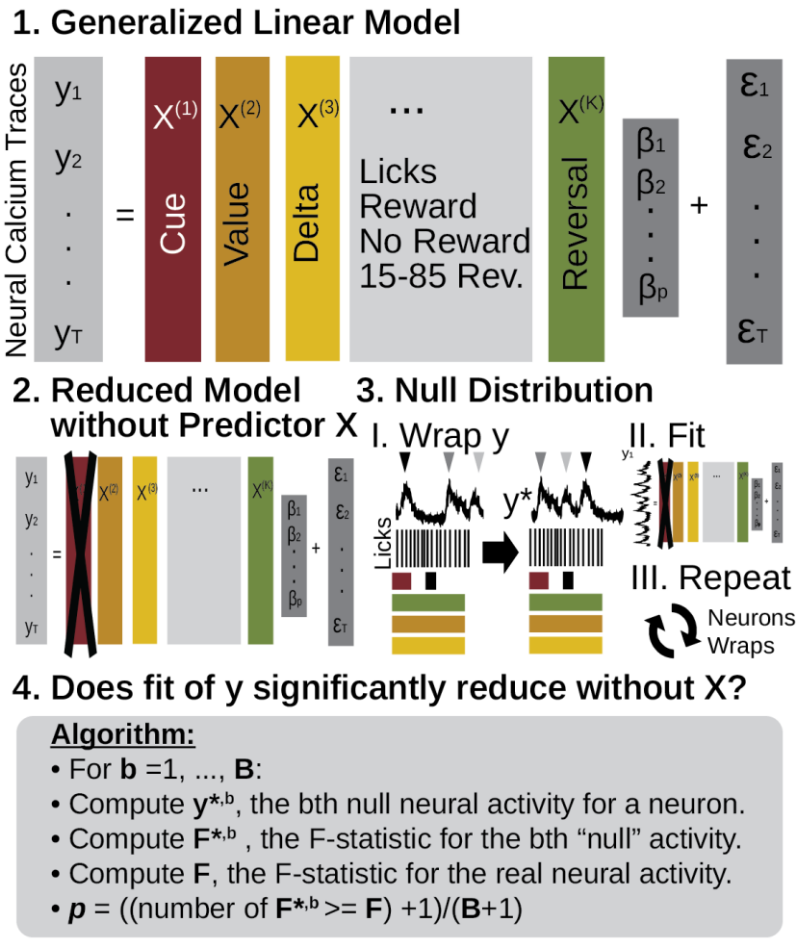

Figure 1-5: Visualization of Generalized Linear Model Significance Testing Method.

testing can generate a more reproducible encoding threshold in alignment with traditional statistics. Thus, statisticians were required. Fortunately, Daniela Witten and Ethan Ancell were willing to collaborate.

The mathematics, not being a particular strong suit of the author of this work, will likely be more adequately described in the upcoming collaborative publication on the topic. Nevertheless, two quantities are needed for significant testing: the observed data, and a null distribution of data that could have been observed but was not. Hypothesis testing asks how likely the observed data would have occurred in the null distribution. Generating a null distribution from timeseries data comes with a major caveat due to kinetics of calcium indicators like GCaMP as points observed at a given time are not independent from those observed in close temporal proximity. Random shuffles destroy the dependent time structure of the neural data and therefore do not represent a distribution of data that could have been observed. Instead, time-wrapping such that the entire timeseries experiences a uniform time shift preserves the time structure of the data minus one non-linearity from filling the gap at the front of the series with data from the very end of the trace. Wrapping each neuron on the dataset once with a random time shift generates a large null distribution of data that could have happened (and indeed did happen) but is randomly misaligned from the task parameters that could predict it (**Fig. 1-5.3**). For fiber photometry data, the null distribution was constructed in largely the same way but instead of representing many neurons on the same day, it was built using all of the signals from an indicator in a given brain area on all recording days to provide a sufficient number of observations.

Significance can then be determined by comparing the performance of the leave-one-out model on a given neuron (*b*) properly aligned to the task predictors (F) to the performance of the shuffled dataset over many signals (F^{*b}). The p-value signifies the number of times the relationship between the left-out predictor and the neural data (F) is also represented in the null distribution of randomly time shifted data (F*) (**Fig. 1-5.4**). Some linear algebra shortcuts to expedite the model fitting also took place but are outside the scope of this work. Finally, after

setting a significant threshold ($p<0.01$ for neural data, or $p<0.05$ for photometry data owing to lower signal to noise), neurons could be assigned significant encoding for any task predictor(s).

For the reversal learning experiments, the following parameters predicted the neural signal: cue (85-15, 85-85, 15-85, 15-15 on stable days, or a single combined kernel on reversal days since representations could be unstable as contingencies shifted), licks (shifted forward and backward in time 300ms to account for preparatory feedback-related encoding), sucrose delivery (1s) or omission (1s), Rescorla-Wagner modeled value (see **Chapter 1-III**), reward prediction error (delta), and modeled reversal signals (**Fig. 3-1**). For the combined GCaMP and dopamine imaging experiments, the dopamine kernel represented the $rGRAB_{DA}$ signal shifted in time from 0-1s to account for potential delays in GPCR-related signaling cascades.

Reanalyzing data first characterized using the change in variance explained method, yielded similar stability results (**Fig. 2-2E&F**). This agreement between previous methods and the novel p-value GLM suggests mathematical validity for the new method, since it largely reproduces previous results. Another feature of the p-value GLM compared to setting a single variance explained threshold is that the F statistic for each predictor is computed individually and the threshold can scale with the likelihood of false positive encoding within the time-shifted dataset. This is particularly relevant as some predictors occur frequently (like licking) while others are relatively sparse (like reversal signals to a single cue), and the probability of a random alignment with peaks in neural activity can be expected to increase when a predictor happens more often. Altogether, the p-value GLM method brings strong statistical grounding to set meaningful thresholds for leave-one-out GLMs, improving their utility for encoding models.

III. GENERAL METHODS

Reversal Task: Rodent behavioral tasks that evaluate cue-based reversal learning generally involve two cues, one with high value and one with low value, that are switched at a pre-determined timepoint. While these tasks include both contingency degradation (high-low) and enhancement (low-high) components, they can be difficult to disentangle since information about the contingency switch of one cue can be applied to the other. A task that includes cues that both do and do not reverse can circumvent this issue because identifying that one cue has changed contingency does not necessarily provide information about the contingency status of another cue.

Animals are also likely to quickly notice contingencies that are more salient to them. Of the five traditional senses, mice rely primarily on olfactory information (for review see Lankford et al., 2020) to navigate their environment in the wild, so they may find olfactory information most salient. Thus, the task in the majority of this work contained four counterbalanced odor cues at 10% in mineral oil with neutral valence (Lee et al., 2013; Saraiva et al., 2016), so that no cue started with any inherent meaning to the mice. Two cues began as high value (85% probability of 2.5uL 10% sucrose reward), and two as low value (15% probability of the same reward). After animals displayed stable task performance for two days, the contingencies of one high value (85-15) and one low value (15-85) cue were reversed, while the other two remained constant (85-85, 15-15). Throughout this work, cues are denoted A-B, where A indicates the reward probability before the reversal, and B after. **Figure 1-6** depicts this task.

Value Modeling: Rescorla-Wagner (RW) value models estimated how an animal's internal valuation of stable and reversed cues changed throughout the task. Following a published protocol (Rhoads & Gan, 2022), an animal's "choices" (presence or absence of anticipatory licking above baseline in attempt to collect a predicted reward before its delivery)

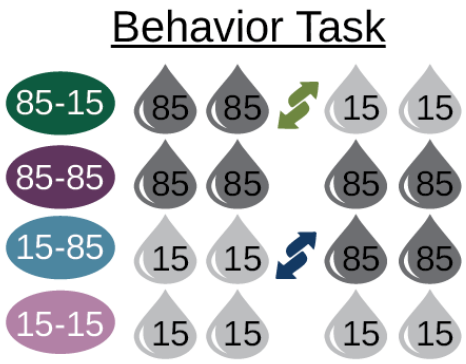

Figure 1-6: Pavlovian Reversal Task

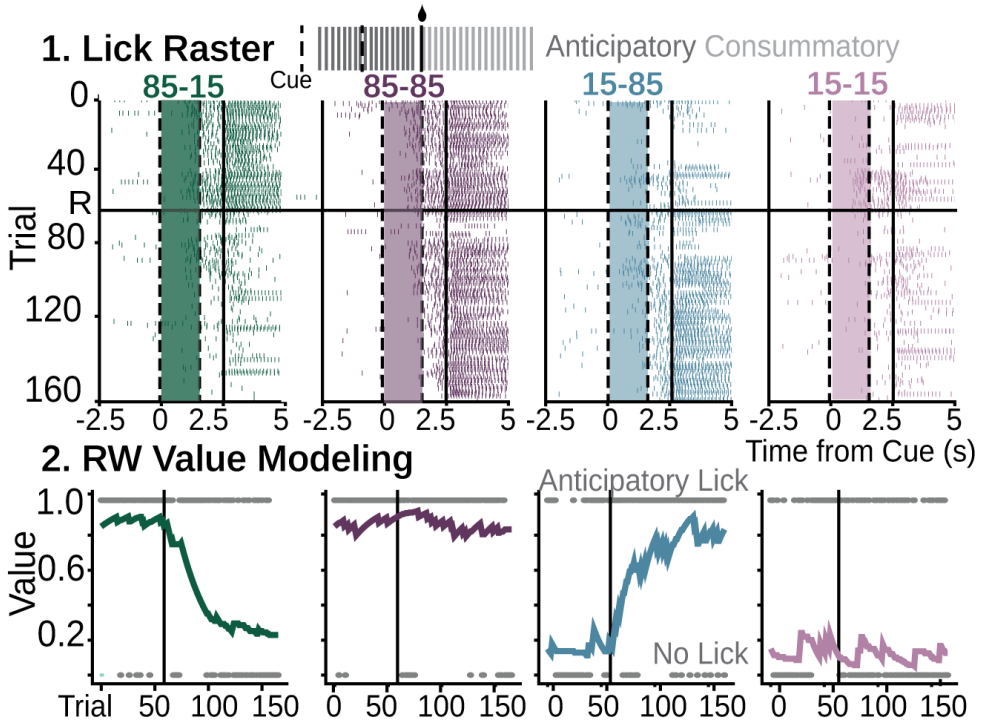

Figure 1-7: Rescorla-Wagner Modeling of Reversal Behavior. Sample task performance in an example animal and estimated value curves using anticipatory licking as a choice proxy.

were individually fit to each of the four reversal cues (**Fig. 1-7**). The RW modeling returned an estimated value (Q) for each trial x cue-type, in addition to how the outcome differed from the prediction (δ), also known as the reward prediction error (RPE). Both the value and RPE subsequently acted as predictors for generalized-linear modeling of the neural signal. See **Chapter 3** for more details about the application of RW models to neural data.

Neural Recordings: This work represents four major imaging cohorts: a cue set stability cohort detailed in **Appendix A**, and three new two-photon cohorts. The primary PL imaging cohort contains six mice, for a total of 2158 cells tracked across reversal (3544 untracked), acquired on an Olympus FVMPE-RS with an XLPLN10xSVMP. All mice received implants of 1.5x3mm microprisms (**Fig. 1-3G**) following the protocol in **Chapter 1-I**. The other two-photon imaging cohorts utilized 1.5x8mm lenses due to availability, but otherwise followed the same implant/injection protocol. Both were acquired on Bruker Investigator microscopes using Cousa objectives. They are underpowered due to untimely laser failure, and therefore only represent single animals, but were included in this work due to complimentary results from adequately powered associated fiber photometry cohorts. Thus, the dual GRABDA3h/GCaMP6s results (**Fig. 4-2**) were acquired using dual 920/1080 Spectra Physics Insight laser excitation and represent 546 cells from one animal, and the PFC-VTA results represent 16 cells from one animal. VTA injections were targeted towards rostro-lateral VTA at -2.8, +1.22 , -4.4 angled 10 degrees. The VTA cohort was imaged at 920nm using a Coherent Chameleon laser. Planes were registered across days and cells tracked following the method in **Chapter 1-I**.

Bulk Fluorescence Recordings: In total, eight mice injected with AAV2/9-hSyn-GRAB-DA3h along with four mice injected with AAVDJ-CaMKIIa-GCAMP6s at +1.94, 0.4, -2.4 and implanted with photometry fibers at +1.94, 0.4, -2.3 also performed the reversal task under research technician Zoe Garrett's supervision. A 465 nm LED excited both indicators, and although a 405nm LED was also present, it's signal was discarded given a lack of motion artifacts in headfixed mice. Following collection, each signal underwent photobleaching correction via a

5th degree polynomial and denoising using the OASIS algorithm (Friedrich et al., 2017), adapted to fit published indicator kinetics (Zhuo et al., 2023) where necessary.

Data Analysis: All analysis was performed in Python 3.7/Jupyter Notebook using custom scripts. In addition to the protocols outlined in **Chapter 1-I**, GLM details in **Chapter 1-II**, and value modeling detailed earlier in this chapter, we utilized the following packages: *pandas.rollling(x).mean* to calculate meta-RPE, where x was the retrospective window that best matched the smoothed RW derivative (see **Chapter 3**), *numpy.intersect1d()* on the indices of GLM-identified ensembles to quantify their overlap within and across days. *statsmodels.regression.linear_model.OLS* to quantify the correlation coefficient (r^2) between datasets and significance statistics for single linear regressions, and *scipy.signal.gaussian(1)* to smooth lick traces where appropriate. *Matplotlib* generated most graphs, with *seaborn* utilized for heatmaps. GraphPad Prism 9 calculated population statistics on bar graphs using either Repeated-Measures One-way ANOVA or Two-way ANOVA, where appropriate, both with Tukey HSD for multiple comparisons. The BIC scores in Figure 3-4E are a feature of the utilized RW modeling package (Rhoads & Gan, 2022). The color schemes represented in this work utilize MetBrewer's Redon Pandora 1914 palette (Mills, 2023).

References

Andermann, M. L., Gilfoy, N. B., Goldey, G. J., Sachdev, R. N. S., Wölfel, M., McCormick, D. A., Reid, R. C., & Levene, M. J. (2013). Chronic cellular imaging of entire cortical columns in awake mice using microprisms. *Neuron*, *80*(4), 900–913. https://doi.org/10.1016/j.neuron.2013.07.052

Barnett, L. M., Hughes, T. E., & Drobizhev, M. (2017). Deciphering the molecular mechanism responsible for GCaMP6m's Ca2+-dependent change in fluorescence. *PLOS ONE*, *12*(2), e0170934. https://doi.org/10.1371/journal.pone.0170934

Benninger, R. K. P., & Piston, D. W. (2013). Two-Photon Excitation Microscopy for the Study of Living Cells and Tissues. *Current Protocols in Cell Biology / Editorial Board, Juan S. Bonifacino ... [et Al.]*, *0 4*, Unit-4.1124. https://doi.org/10.1002/0471143030.cb0411s59

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultra-sensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. https://doi.org/10.1038/nature12354

Engel, T. A., & Steinmetz, N. A. (2019). New perspectives on dimensionality and variability from large-scale cortical dynamics. *Current Opinion in Neurobiology*, *58*, 181–190. https://doi.org/10.1016/j.conb.2019.09.003

Friedrich, J., Zhou, P., & Paninski, L. (2017). Fast online deconvolution of calcium imaging data. *PLoS Computational Biology*, *13*(3), e1005423. https://doi.org/10.1371/journal.pcbi.1005423

Gordon-Fennell, A., Barbakh, J. M., Utley, M. T., Singh, S., Bazzino, P., Gowrishankar, R., Bruchas, M. R., Roitman, M. F., & Stuber, G. D. (2023). An open-source platform for head-fixed operant and consummatory behavior. *eLife*, *12*, e86183. https://doi.org/10.7554/eLife.86183

Hjort, M., Raajaram Gowrishankar, Lucy Tian, Adam Gordon-Fennell, Vijay M. K. Namboodiri, Michael R. Bruchas, & Garret D. Stuber. (2024). Microprisms enable enhanced throughput and resolution for longitudinal tracking of neuronal ensembles in deep brain structures. *Neurophotonics*, *11*(3), 033407. https://doi.org/10.1117/1.NPh.11.3.033407

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, *117*(4), 500–544.

Jung, J. C., Mehta, A. D., Aksay, E., Stepnoski, R., & Schnitzer, M. J. (2004). In Vivo Mammalian Brain Imaging Using One- and Two-Photon Fluorescence Microendoscopy. *Journal of Neurophysiology*, *92*(5), 3121–3133. https://doi.org/10.1152/jn.00234.2004

Lankford, C. K., Laird, J. G., Inamdar, S. M., & Baker, S. A. (2020). A Comparison of the Primary Sensory Neurons Used in Olfaction and Vision. *Frontiers in Cellular Neuroscience*, *14*, 595523. https://doi.org/10.3389/fncel.2020.595523

Lee, J., Linster, C., & Devore, S. (2013). Odor preferences shape discrimination learning in rats. *Behavioral Neuroscience*, *127*(4), 498–504. https://doi.org/10.1037/a0033329

Legaria, A. A., Matikainen-Ankney, B. A., Yang, B., Ahanonu, B., Licholai, J. A., Parker, J. G., & Kravitz, A. V. (2022). Fiber photometry in striatum reflects primarily nonsomatic changes in calcium. *Nature Neuroscience*, *25*(9), Article 9. https://doi.org/10.1038/s41593-022-01152-z

Levene, M. J., Dombeck, D. A., Kasischke, K. A., Molloy, R. P., & Webb, W. W. (2004). In vivo multiphoton microscopy of deep brain tissue. *Journal of Neurophysiology*, *91*(4), 1908–1912. https://doi.org/10.1152/jn.01007.2003

Low, R. J., Gu, Y., & Tank, D. W. (2014). Cellular resolution optical access to brain regions in fissures: Imaging medial prefrontal cortex and grid cells in entorhinal cortex. *Proceedings of the National Academy of Sciences*, *111*(52), 18739–18744. https://doi.org/10.1073/pnas.1421753111

Lu, R., Sun, W., Liang, Y., Kerlin, A., Bierfeld, J., Seelig, J., Wilson, D. E., Scholl, B., Mohar, B., Tanimoto, M., Koyama, M., Fitzpatrick, D., Orger, M. B., & Ji, N. (2017). Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nature Neuroscience*, *20*(4), 620–628. https://doi.org/10.1038/nn.4516

Marshel, J. H., Kim, Y. S., Machado, T. A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., Shane, J. C., McKnight, D. J., Yoshizawa, S., Kato, H. E., Ganguli, S., & Deisseroth, K. (2019). Cortical layer–specific critical dynamics triggering perception. *Science*, *365*(6453), eaaw5202. https://doi.org/10.1126/science.aaw5202

Meng, G., Liang, Y., Sarsfield, S., Jiang, W., Lu, R., Dudman, J. T., Aponte, Y., & Ji, N. (2019). High-throughput synapse-resolving two-photon fluorescence microendoscopy for deep-brain volumetric imaging in vivo. *eLife*, *8*, e40805. https://doi.org/10.7554/eLife.40805

Mills, B. (2023). *GitHub—BlakeRMills/MetBrewer: Color palette package in R inspired by works at the Metropolitan Museum of Art in New York*. https://github.com/BlakeRMills/MetBrewer?tab=readme-ov-file

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), 103–107. https://doi.org/10.1038/nature21376

Ottenheimer, D. J., Hjort, M. M., Bowen, A. J., Steinmetz, N. A., & Stuber, G. D. (2023). A stable, distributed code for cue value in mouse cortex during reward learning. *eLife*, *12*, RP84604. https://doi.org/10.7554/eLife.84604

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. D. (2017). *Suite2p: Beyond 10,000 neurons with standard two-photon microscopy* (p. 061507). bioRxiv. https://doi.org/10.1101/061507

Piantadosi, S. C., Zhou, Z. C., Pizzano, C., Pedersen, C. E., Nguyen, T. K., Thai, S., Stuber, G. D., & Bruchas, M. R. (2023). Holographic stimulation of opposing amygdala ensembles bidirectionally modulates valence-specific behavior via mutual inhibition. *Neuron*. https://doi.org/10.1016/j.neuron.2023.11.007

Redman, W. T., Wolcott, N. S., Montelisciani, L., Luna, G., Marks, T. D., Sit, K. K., Yu, C.-H., Smith, S., & Goard, M. J. (2022). Long-term transverse imaging of the hippocampus with glass microperiscopes. *eLife*, *11*, e75391. https://doi.org/10.7554/eLife.75391

Resendez, S. L., Jennings, J. H., Ung, R. L., Namboodiri, V. M. K., Zhou, Z. C., Otis, J. M., Nomura, H., McHenry, J. A., Kosyk, O., & Stuber, G. D. (2016). Visualization of cortical, subcortical, and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nature Protocols*, *11*(3), 566–597. https://doi.org/10.1038/nprot.2016.021

Rhoads, S. A., & Gan, L. (2022). Computational models of human social behavior and neuroscience: An open educational course and Jupyter Book to advance computational training. *Journal of Open Source Education*, *5*(47), 146. https://doi.org/10.21105/jose.00146

Saraiva, L. R., Kondoh, K., Ye, X., Yoon, K., Hernandez, M., & Buck, L. B. (2016). Combinatorial effects of odorants on mouse behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), E3300–E3306. https://doi.org/10.1073/pnas.1605973113

Spellman, T., Malka, S., Jesse, K., Gabriela, M.-N., & Conor, L. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, *576*(7786), 266–273. https://doi.org/10.1038/s41586-019-1787-x

Theer, P., & Denk, W. (2006). On the fundamental imaging-depth limit in two-photon microscopy. *JOSA A*, *23*(12), 3139–3149. https://doi.org/10.1364/JOSAA.23.003139

Tischbirek, C. H., Birkner, A., & Konnerth, A. (2017). In vivo deep two-photon imaging of neural circuits with the fluorescent Ca2+ indicator Cal-590. *The Journal of Physiology*, *595*(10), 3097–3105. https://doi.org/10.1113/JP272790

Wang, K., Sun, W., Richie, C. T., Harvey, B. K., Betzig, E., & Ji, N. (2015). Direct wavefront sensing for high-resolution in vivo imaging in scattering tissue. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms8276

Xu, S., Yang, H., Menon, V., Lemire, A. L., Wang, L., Henry, F. E., Turaga, S. C., & Sternson, S. M. (2020). Behavioral state coding by molecularly defined paraventricular hypothalamic cell type ensembles. *Science*, *370*(6514), eabb2494. https://doi.org/10.1126/science.abb2494

Yildirim, M., Sugihara, H., So, P. T. C., & Sur, M. (2019). Functional imaging of visual cortical layers and subplate in awake mice with optimized three-photon microscopy. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-018-08179-6

Yu, C.-H., Yu, Y., Adsit, L. M., Chang, J. T., Barchini, J., Moberly, A. H., Benisty, H., Kim, J., Young, B. K., Heng, K., Farinella, D. M., Leikvoll, A., Pavan, R., Vistein, R., Nanfito, B. R., Hildebrand, D. G. C., Otero-Coronel, S., Vaziri, A., Goldberg, J. L., … Smith, S. L. (2023). *The Cousa objective: A long working distance air objective for multiphoton imaging in vivo* (p. 2022.11.06.515343). bioRxiv. https://doi.org/10.1101/2022.11.06.515343

Yuan, A. X., Colonell, J., Lebedeva, A., Okun, M., Charles, A. S., & Harris, T. D. (2024). Multi-day Neuron Tracking in High Density Electrophysiology Recordings using EMD. *bioRxiv: The Preprint Server for Biology*, 2023.08.03.551724. https://doi.org/10.1101/2023.08.03.551724

Zhuo, Y., Luo, B., Yi, X., Dong, H., Wan, J., Cai, R., Williams, J. T., Qian, T., Campbell, M. G., Miao, X., Li, B., Wei, Y., Li, G., Wang, H., Zheng, Y., Watabe-Uchida, M., & Li, Y. (2023). Improved dual-color GRAB sensors for monitoring dopaminergic activity in vivo. *bioRxiv*, 2023.08.24.554559. https://doi.org/10.1101/2023.08.24.554559

Ziv, Y., Burns, L. D., Cocker, E. D., Hamel, E. O., Ghosh, K. K., Kitch, L. J., El Gamal, A., & Schnitzer, M. J. (2013). Long-term dynamics of CA1 hippocampal place codes. *Nature Neuroscience*, *16*(3), 264–266. https://doi.org/10.1038/nn.3329

CHAPTER 2: TESTS OF CUE STABILITY IN PRELIMBIC CORTEX

Before exploring how reversal learning is reflected in the prelimbic cortex (PL), it is important to establish that the isolated dynamics reflect changes to existing learning, and not stochastic fluctuations within the brain. While the previous chapter discussed methodological approaches to this end, this chapter will present an experimental foundation to test this necessity. Existing literature provides preliminary support for a stable cue code within PL because two distinct odor cues that predict the same reward contingency within the same task cue set are represented similarly enough within PL ensembles that decoders perform equally across the two cues (Wang et al., 2020). That cue responses emerge in PL with learning (Otis et al., 2017) provides additional support that PL cue representation is likely a level of abstraction above a direct sensory representation as although the sensory properties of a cue do not change over learning, the representation in PL does change. Wang et al.'s findings are amenable to a similar interpretation, as based on decoder performance it appears that discrete odor cues with the same contingencies share a high level of similarity in how they are represented in PL. Three major open questions remain regarding cue stability in PL:

I. Does the stability of PL cue representations extend across cue sets?
II. Are cue-reward contingencies represented stably across cognitive flexibility, including reversals?
III. What drives the stability of PL cue-reward representations?

Answering these questions is very important before attempting to characterize a flexibility signal in PL for the following reasons, broken down by question.

I/II: Some have argued that behavioral responses across cognitive flexibility are represented as different "sets" (Monsell, 2003). If cue-reward contingencies are represented differently in these sets, it would be difficult to isolate a cognitive flexibility signal from the intersection of the disappearance of the old set and emergence of the new set representation, neither

of which would be truly specific to the reversal period or distinctly “flexibility.” However, if cue-reward contingencies are represented stably across information sets, then changes unique to the transition could be isolated as flexibility signals. Characterizing whether cues are stable across sets not linked by cognitive flexibility (I) is an important precursor to studies that attempt to do so across cognitive flexibility events, like reversals (II).

III. Many regions of the brain reflect motor signals, including licking (Zagha et al., 2022). The similarity in how well Wang et al.’s cues with the same reward contingency could be explained by documented similarities in licking behavior across cues with the same reward contingencies. Such a finding would not be evidence for a stable cue code, but instead a stable movement correlate reflected in PL. Including licking information in encoding models that test for stability is vital to test whether any observed stability is actually related to cue properties, instead of cue-evoked motor responses

Therefore, the purpose of this chapter is to test the hypothesis that PL cue encoding is not a motor correlate, and instead represents a stable cue-reward association across odor sets, days, and cognitive flexibility events.

I. DOES THE STABILITY OF PRELIMBIC CUE REPRESENTATIONS EXTEND ACROSS CUE SETS?

To examine the representational similarity of distinct cue sets in PL, mice trained on two odor sets (A and B) in sequence on different days (**Fig. 2-1A**). Each set continued three neutral odors diluted 10% in mineral oil that predicted a 2.5uL 10% sucrose reward with 100%(CS+), 50%(CS50), or 0%(CS-) probability (**Fig. 2-1B**). On a given day, headfixed animals (**Figure 2-1C**) received 50 randomly interleaved trials of CS+, CS50, or CS- cues with 8-12s between presentations, for a total of 150 cues per day. On all days, animals behaved concurrently with two-photon calcium imaging through a 1x4mm GRIN lens of excitatory neural ensembles in PL injected with AAVDJ-CaMKIIa-GCaMP6s (**Fig. 2-1C**). Neurons were tracked across days using a

combination of concatenated Suite2p recordings and manual curation using custom python scripts. For more methodological information, see **Appendix A**. Mice took three days to achieve mastery of the task, measured using anticipatory licking. Anticipatory licking counts the licks an animal performs in between cue presentation and reward delivery, indicating that an animal is attempting to collect a future predicted reward. In both sets A and B, animals displayed the highest amount of anticipatory licking for CS+ cues, followed by CS50 cues, and almost no licking for CS- cues (**Fig. 2-1D**). Licking following cues with the same contingency was not significantly different between the two sets, indicating that mice likely treated the cues from sets A and B equivalently. With equal behavior for sets A and B in hand, it was then possible to ask whether sets A and B were represented similarly within PL neurons.

Similarity maps for cues across sets A3 and B3 comparing the difference in neural activity 0-2.5s before and after cue delivery can easily be qualitatively compared for potential coding similarities. Astonishingly, these maps share an incredible amount of visual similarity even though they were constructed on different days and with different CS+ odors (**Fig. 2-2A**). Across all cue types, there is a striking similarity in trial averaged activity in cells tracked across days and within cue types (**Fig. 2-2B**). These findings are congruent to those from Wang et al. However, a number of trial variables could drive this similarity independent of a stable neural representation of cue-reward associations, most notably the motor components of licking behavior. Especially since mice display statistically insignificant differences in cue-evoked licking behavior across odor sets A and B (**Fig. 2-1D**), neural reflections of this licking behavior could easily explain the observed similarity in cue representations.

Training an encoding-based generalized liner model (GLM) using previously published methods (Steinmetz et al., 2019) allowed for an examination of the specific contributions of different task parameters including cues, licks, and reward delivery to the observed stability. Cells were attributed “significant” coding of a parameter when the variance explained (VE) of the overall

predictive model dropped at least 2% when the parameter was omitted from the model. Under this criterion, approximately 50% of recorded cells were significant for at least one task parameter. As expected, about 30% of cells were lick responsive, but about 10% of those cells also had significant cue encoding (“both”) in addition to another approximately 10% that coded for cues independently of licks (**Fig. 2-2C**). The trained encoding models provided meaningful predictions of neural activity, as they outperformed control models with task elements in shuffled in time (**Fig. 2-2D**). Isolating the unique variance attributable to cue, lick, or other variable coding within cue, lick, both, and other (mostly reward) cells on days A3 and B3 involved an encoding model with beta-weights calculated on A3 and applied to both days. Despite a statistically significant amount of coding, there was not a significant difference in the amount (VE) of cue coding across A3 and B3 in cue or both cells, even with licking was accounted for in the GLM (**Fig. 2-2E**). Taken together, these GLM results suggest that motor correlates of learning cannot fully account for the similarity in cue coding between sets A and B.

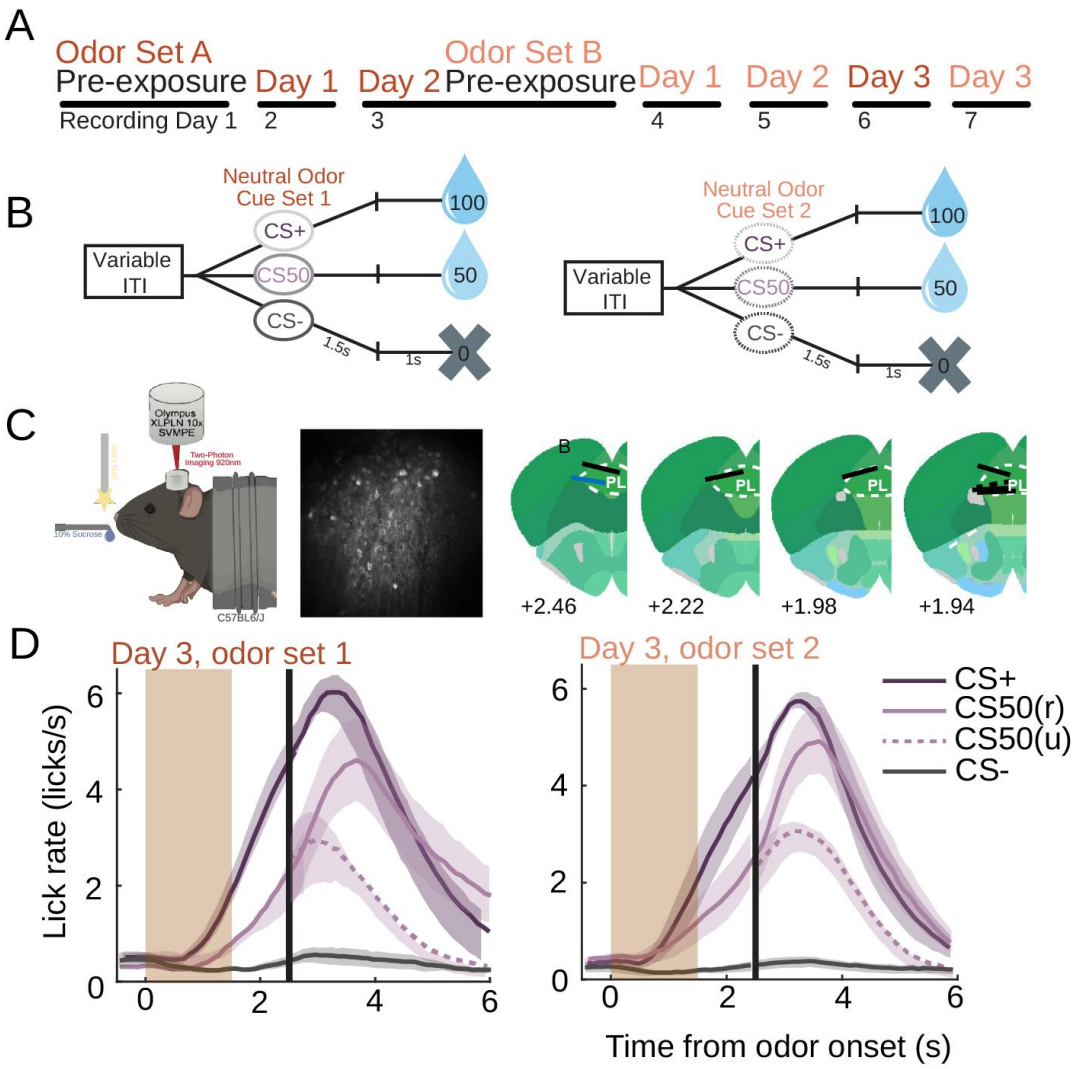

Figure 2-1: Cue Set Task and Behavior. *Adapted from Ottenheimer*, Hjort*, Bowen* et al. 2023* **A.** Schematic of set task and recording timeline. **B**. Representation of cue contingencies for CS+ (100%), CS50 (50%), and CS-(0%) for sets A and B. **C.** Mouse recording schematic, example PL FOV, and location of GRIN Lenses in PL. **D.** Animal performance in sets 1 and 2 (also referred to as A and B)

The results thus far established that cue representations in PL are stable independent of licking representations, but a fundamental question remained: how stable are they actually? Understanding the relative stability of PL cue coding within and across sets can help set thresholds for assessing stability across cognitive flexibility events, like reversals. Answering this question, required re-analysis of the stability data with the p-value GLM (**Chapter 1**). When comparing the proportion of cells that pass the $\alpha<0.01$ threshold within a single odor set (A) between days A2 and A3, the vast majority of cue cells encode the CS+ on both days (**Fig. 2-2F**). Surprisingly, most cue cells also significantly code for the CS50 on both days, with little proportional difference from the CS+ despite a 50% reduction in reward likelihood (**Fig. 2-2F**). Few cells displayed significant coding for the CS-, consistent with behavioral (**Fig. 2-1D**) and trial-averaged neural (**Fig. 2-2B**) data representations. Coding stability was slightly reduced across sets A3 and B3, but still approximately 50% of cue cells were significantly detected across odor sets for both the CS+ and the CS50 (**Fig. 2-2F**). Taken together, these data suggest that cue coding is largely stable across odor sets.

That there seems to be little difference in coding ensemble size for 100% and 50% likely rewards suggests that value representations are unlikely to drive differences in the size of cue-reward associations within PL. This is perhaps unexpected, as generally the strength of cue-reward associations is thought to scale with the strength (or value) of the association, and at the level of the number of neurons in an ensemble this does not appear to be true. However, the magnitude of the activity within ensembles active during the cue period (**Fig. 2-2B**) does appear larger for the CS+ compared to the CS50. Perhaps the structure of cue-reward association representations in PL is hierarchical- first whether the reward is ‘wanted’ at the ensemble level, and then how much the reward is wanted within that ensemble via activation magnitude. Future studies may be able to clarify if this is indeed the case.

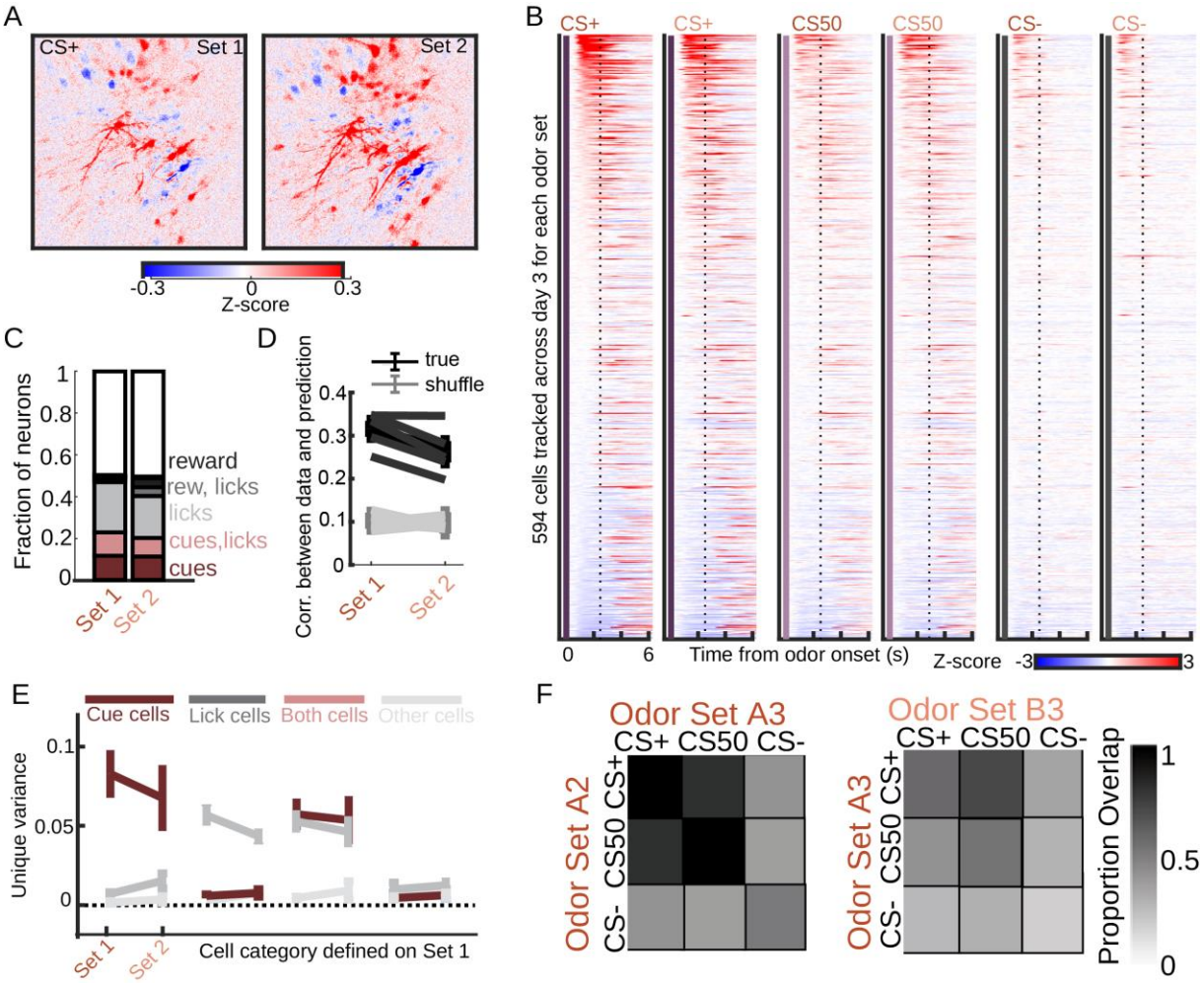

Figure 2-2: Cue Stability across Odor Sets. *Adapted from Ottenheimer*, Hjort*, Bowen* et al. 2023.* **A.** Averaged z-scored activity for A3 and B3 CS+ cues. **B**. Trial-average extracted activity for all tracked cells across A3 and B3 for CS+, CS50, CS- cues. **C**. Proportion of neurons that passed GLM VE>0.02 threshold. **D**. There is not a significant difference between A3-trained GLM performance on B3. **E**. Unique variance explained by cues, licks, and other parameters for cue, lick and both cells. **F**. Proportion overlap for cells that significantly code for CS+, CS50, and CS- in p-value GLM at α<0.01. Proportions on each map are represented via comparison to the largest number of cells sorted into a group on any day. Approximately 50% of cue cells are stable across A3 and B3 for CS+ and CS50 conditions.

II. DOES PRELIMBIC CUE STABILITY EXTEND ACROSS REVERSAL?

To determine whether PL cue representational stability extended across reversal, headfixed mice (**Fig. 2-3A**) implanted with microprisms (**Chapter 1**) performed a reversal learning task adapted from the Pavlovian odor learning task (**Chapter 2-I**) with some modifications (**Fig. 2-3B**). First, all cues were probabilistic to better capture potential value encoding within PL cue or other cell types. Value models are much more dynamic when outcomes cannot be predicted with certainty. Mice experienced a “high” value outcome of 85% reward probability and a “low” value outcome of 15% reward probability spread across four literature-validated neutral odor cues (Lee et al., 2013; Saraiva et al., 2016) selected randomly for each mouse from a set of eight (butanol: banana, limonene: lime, carvone: mint, benzaldehyde: almond, geraniol: floral, ethyl butyrate: pineapple, 3-hexenal: grass, pinene: pine). Having four total cues, two that start at 85% reward probability, and two at 15% reward probability allows for internal controls (85-85 and 15-15) when the contingency of the other two odors (85-15, 15-85) are reversed (**Fig. 2-3B**). This design allows dissociation of contingency degradation (85-15) and elevation (15-85) within a larger context with some stability (85-85, 15-15). Trials were separated by a longer 16-20s ITI to provide additional data without task predictors to improve p-value GLM fit (**Chapter 1**). Animals learn this task well: anticipatory licking is not significantly different within 85% or 15% cues on stable behavior sessions before, after, or across the reversal (**Fig. 2-3C**). This snapshot perspective implies an astonishing behavioral stability within animals throughout the task: they will lick about 2 times per second over the 2.5s between cue presentation and reward delivery for 85% cues, and 1 time per second for 15% cues. Whether this stability extended to PL neural representations remained to be tested.

Training p-value GLMs on four stable behavior days (pre1 and pre2 before reversal, and post1 and post2 after reversal) revealed a notable amount of stability in cue encoding across all comparisons (**Fig. 2-4A**). This degree of stability is likely not attributable to midway snapshots of

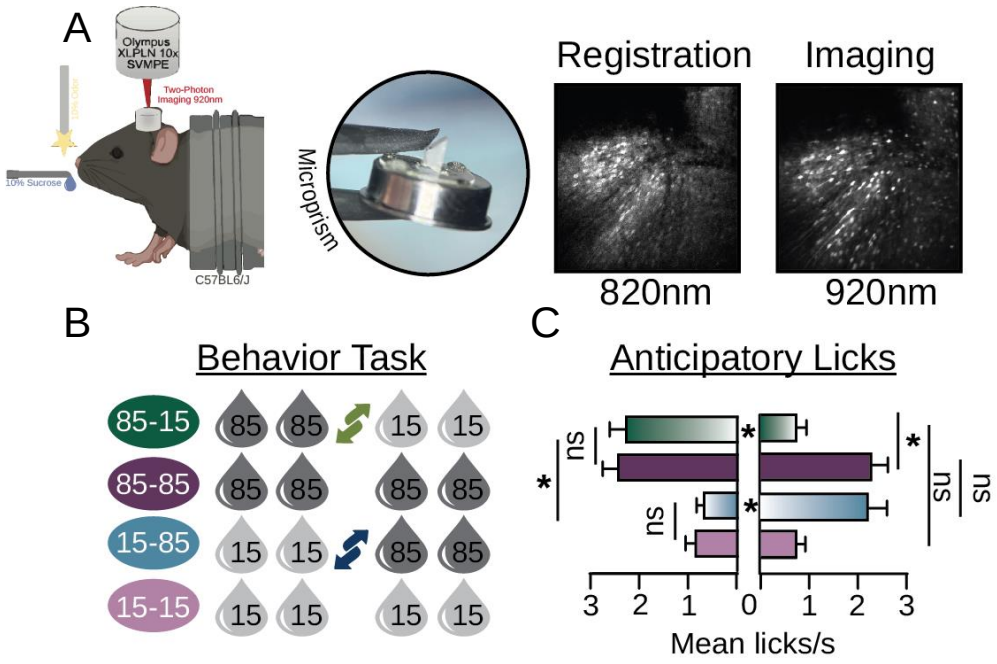

Figure 2-3: Pavlovian Reversal Task. A. Headfixed mouse receives odor cues and 10% sucrose rewards while under a two-photon microscope. Mice were imaged through microprisms (**see Chapter 1**) with a registration scan at 820nm for tracking and imaging scans at 920 nm. **B.** Schematic of 4-odor reversal task. After back to back days with consistent performance, the contingency of two cues (85-15 and 15-85) were switched halfway through a third session. Animals adapted to the new contingencies until they reached stable back to back performance. **C.** Behavioral quantification of learning. There is not a significant difference in anticipatory licking to 85% cues before or after reversal, nor for 15% cues across reversal.

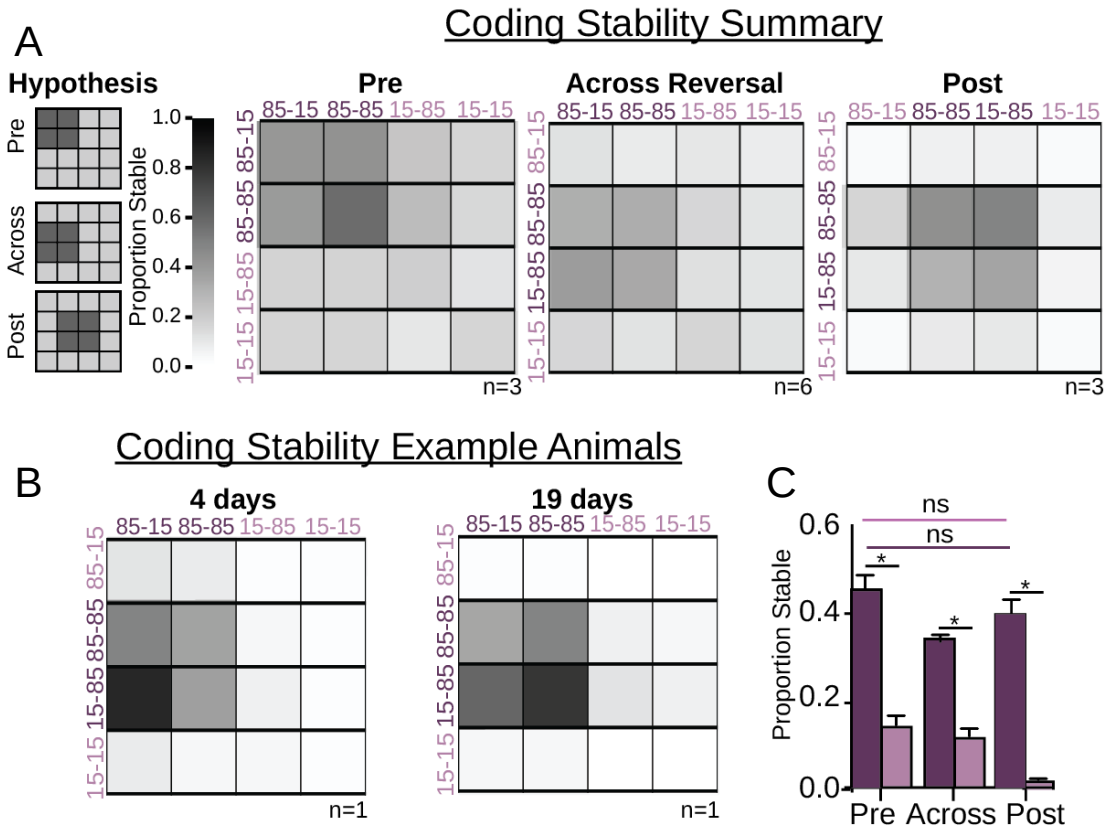

Figure 2-4: Cue Stability Across Reversal. A. Hypothesis for cue stability cell proportions, from result in **Fig. 2-2F** compared to neural results before (pre), after (post), and across reversal. **B.** Example animals with different spacing between stable pre and post day compared across reversal, in **A**. **C.** Statistics on the proportion of stable cue cells for 85% (dark purple) and 15% (light purple) cues before, after, and across reversal. There is not a significant difference in the proportion of coding within a cue type in any of the conditions.

representational drift, as animals with pre2 and post1 ("across reversal") sessions spaced days and weeks apart displayed similar levels of stability (**Fig. 2-4B**). In fact, stability did not significantly decline anywhere in the learning process (**Fig 2-4C**), with about 40% of cue cells stably coding high value relationships before, after, and across reversal.

Although 40% of cue cells appear to stably encode for high value cues across reversal, their activity profile could different before and after reversal, conveying set information. Peri-stimulus time histograms (PSTH) from tracked cells pre2 and post1 display striking similarity for high value cues before and after the reversal (**Fig. 2-5A**), as do their average traces (**Fig 2-5B**). Quantitatively, there is not a significant difference in the mean cue-evoked activity of stable cue cells within high or low value cue identities (**Fig. 2-5C**). Thus, cue information is not only stably represented in a subset of PL neurons, but also via their activity profiles.

This seeming immunity from the effects of representational drift quantified extensively in other cortical areas (for review see (Driscoll et al., 2022)) is notable for a number reasons. First, these results suggest that drift is not an inherent property of the cortex, as some have suggested. This opens an interesting future avenue of inquiry regarding why some areas of cortex drift and others do not, though this is outside the scope of the current work. Speculating, it is possibly that regions of the brain involved in, but not responsible for, learning like V1 receive information about learned associations but do not stably store them since they reside elsewhere in the brain. Secondly, this profound stability in PL suggests that PL may be a long-term substrate for storage of cue-reward associations in memory since it does not seem to remap frequently like other canonical memory substrates like hippocampus (Colgin et al., 2008). Finally, that PL displays largely stable cue-evoked dynamics before and after the contingency reversal permits for the careful isolation of potential cognitive flexibility signals during the behavioral remapping that takes place during the reversal (**Chapter 3**).

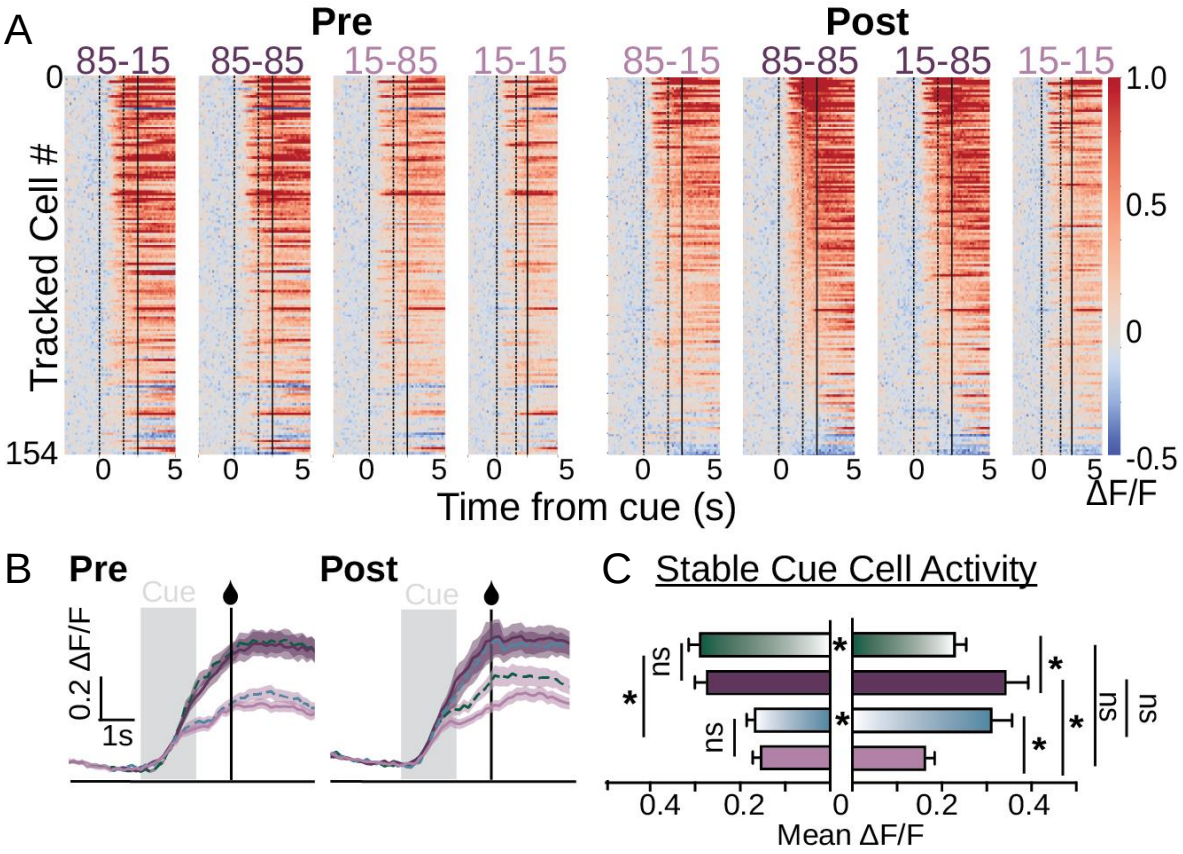

Figure 2-5: Trial-Averaged Dynamics in Stable Cue Cells. A. Trial-averaged time histogram of stable cue cell activity before (pre) and after (post) contingency reversal, sorted by post 85-85 activity. **B.** Mean +/-SEM traces for stable cue cells on 85-15 (green dash), 85-85 (dark purple solid), 15-85 (light blue dash) and 15-15 (light purple solid). Error bars indicate 85% (dark purple) or 15% (light purple) contingencies. **C**. Quantification of average stable cue cell activity during cue and trace interval (0-2.5s) split by cue. There is not a significant difference within coding of 85% or 15% cues before or after reversal.

III. WHAT DRIVES THE STABILITY OF PRELIMIBIC CUE-REWARD REPRESENTATIONS?

Cue representations in PL are stable within (**Fig. 2-4**) and across cue sets (**Fig. 2-2**), as well as across contingency reversals within a single cue set (**Fig. 2-4&2-5**). The stable cue representation does not derive solely from lick encoding, as although behavior is extremely similar for cues with the same reward contingencies (**Fig. 2-1D, Fig. 2-3C**), cells with significant cue coding arise (**Fig. 2-2C&E, Fig. 2-4A**) even when licking is accounted for as a modeling predictor. Thus, this stable representation is at least partially independent from motor signals, and extends across sensorially distinct cues with the same contingencies (**Fig. 2-4A**) even when these cues belong to distinct sets (**Fig. 2-2F**). That the stability of PL cue encoding does not depend on similarities in motor response or specific sensory properties, instead reflecting the future reward probability, suggests that PL represents a more abstract quantity about the cue. What this abstract quantity is can only be speculated upon, but it could be a representation of the cue-reward association itself. That PL cue representations are so stable across contexts, including odor sets and different locations in the cognitive flexibility landscape lends credence to this possibility.

If the stable cue contingency coding in PL is a direct representation of a cue-reward association, PL cannot be the only place in the brain such information is stored, as PL is unequivocally not needed for acquisition or expression of simple Pavlovian contingencies (**Introduction III**). Once again, what PL uses this relationship for during Pavlovian learning can only be speculated upon. Since PL appears to be necessary for degrading learned cue-reward contingencies (**Introduction III**), it is possible that the stable cue representation is utilized in some way, perhaps to detect changes in what was previously a stable reward landscape. Whether a reversal-related ensemble exists, and whether cue cells also perform to this function is the focus of **Chapter 3.**

References

Colgin, L. L., Moser, E. I., & Moser, M.-B. (2008). Understanding memory through hippocampal remapping. *Trends in Neurosciences*, *31*(9), 469–477. https://doi.org/10.1016/j.tins.2008.06.008

Driscoll, L. N., Duncker, L., & Harvey, C. D. (2022). Representational drift: Emerging theories for continual learning and experimental future directions. *Current Opinion in Neurobiology*, *76*, 102609. https://doi.org/10.1016/j.conb.2022.102609

Lee, J., Linster, C., & Devore, S. (2013). Odor preferences shape discrimination learning in rats. *Behavioral Neuroscience*, *127*(4), 498–504. https://doi.org/10.1037/a0033329

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., & Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), 103–107. https://doi.org/10.1038/nature21376

Saraiva, L. R., Kondoh, K., Ye, X., Yoon, K., Hernandez, M., & Buck, L. B. (2016). Combinatorial effects of odorants on mouse behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), E3300–E3306. https://doi.org/10.1073/pnas.1605973113

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, *576*(7786), 266–273. https://doi.org/10.1038/s41586-019-1787-x

Wang, P. Y., Boboila, C., Chin, M., Higashi-Howard, A., Shamash, P., Wu, Z., Stein, N. P., Abbott, L. F., & Axel, R. (2020). Transient and Persistent Representations of Odor Value in Prefrontal Cortex. *Neuron*, *108*(1), 209-224.e6. https://doi.org/10.1016/j.neuron.2020.07.033

Zagha, E., Erlich, J. C., Lee, S., Lur, G., O'Connor, D. H., Steinmetz, N. A., Stringer, C., & Yang, H. (2022). The Importance of Accounting for Movement When Relating Neuronal Activity to Sensory and Cognitive Processes. *Journal of Neuroscience*, *42*(8), 1375–1382. https://doi.org/10.1523/JNEUROSCI.1919-21.2021

CHAPTER 3: REVERSAL LEARNING, PRELIMBIC CORTEX, AND META-RPE

Since cue representations are stable across reversal (**Chapter 2**), changes in neural activity during the reversal period can be attributed to a flexibility signal and not a simple byproduct of representational drift. Once possible, the next step to isolate a cognitive flexibility signal is to generate a hypothesis of its activity profile. After constructing a hypothesis signal, it can be added to the p-value GLM as a predictor of the neural activity, removed, and tested for significance (**Chapter 1**). Since the reversal paradigm tests a relatively simple cognitive flexibility process with contingency degradation of the 85-15 cue and contingency elevation of the 15-85 cue, simple models can represent a potential reversal learning signal.

One definition for learning is as a *sustained change in value-guided behavior*. Thereupon, a learning signal can be represented mathematically as the sustained (smoothed) change (derivative) of a behaviorally grounded value function. Rescorla-Wagner (RW) is a simple value model that fits Pavlovian reversal behavior well (Siegel & Allan, 1996), updating a cached value signal for the next trial (Q[t+1]) using a learning rate (α) and the difference (δ) between the actual reward (r[t]) and predicted (Q[t]) value on the current trial (t). Since rewards are a set volume that is present (1) or omitted (0) in this task, r[t] is a binary variable. This model is represented in **Equations 3-1 and 3-2**.

Equation 3-1: $Q[t+1] = Q[t] + \alpha\delta$

Equation 3-2: $\delta = r[t] - Q[t]$

When calculated across the last stable pre-reversal day, the reversal day(s), and the first stable post reversal day, Q[t] strongly resembles a sigmoid function for the 85-15 and 15-85 cues, with some minor fluctuations from probabilistic reward history (**Fig. 3-1A**). The presence (1) or absence (0) of anticipatory licking as a choice-driven proxy for an animal's internal estimate of each cue's value can be used to calculate the learning rate (α) separately for each cue. **Figure 3-1A** represents the simulated RW value function for an example animal using the 85-15 cue's fit, $\alpha_{85\text{-}15}$. Since learning represents the *sustained* change in value guided behavior, fitting the RW

value curve with a sigmoid function removes minor changes from probabilistic reward history that are not sustained over many trials. This smoothing follows the general shape of the reversal curve well (**Fig. 3-1A**). Such a smoothed curve representing sustained behavior is not a representation of learning on its own, however.

If learning is the sustained *change* in value-guided behavior, then it can be represented as the *derivative* of the smoothed RW value curve (**Fig. 3-1A**). This representation suggests "negative" learning for 85-15 contingency degradation and "positive" learning for the 15-85 contingency enhancement. Since it is more of interest whether a flexibility signal is larger or smaller, instead of positive or negative, the absolute value of the smoothed value derivative represents learning equally across the two conditions in a directionally independent manner, as displayed in **Figure 3-1A**. This signal suggests that reversal learning, and by proxy cognitive flexibility, is highest in the middle of the behavioral transition, and lowest when behavior has stabilized (**Fig. 3-1A**).

Canonically, the reward prediction error (RPE) signal (**Eq. 3-2**) has been used to represent a trial-by-trial form of learning used to iteratively update value estimates and is proportional to the derivative of the unsmoothed RW value curve (**Eq. 3-1**). The learning signal put forth in this work shares a strong similarity to the traditional RPE signal in that it is also a derivative of an RW value curve (albeit smoothed), but is distinct because instead of representing small, iterative updates to value *trials*, it characterizes a much larger scale change to value *states* of a given cue. Verily, there is a striking dissimilarity between RPE (**Fig. 3-1B, left**) and this reversal learning signal (**Fig. 3-1A**) as RPE has the largest amplitude immediately following reversal, while the learning signal has the largest amplitude in the middle of the reversal. In contrast, smoothing RPEs over many trials generates a signal that is much more similar to the derived reversal learning signal (**Fig. 3-1B, right**) as both peak sometime after the initial contingency reversal. This interpretation suggests such a reversal learning signal may also reflect a state with higher error frequency than

expected, providing error information regarding cumulative trial errors, perhaps suggesting that the state of the cue-reward landscape has changed and contingencies must be re-updated. This learning signal will therefore hereby be referred to as *meta-RPE* since it represents error about errors. Fundamentally, the meta-RPE signal reflects the idea that learning takes place over many trials, so learning signals should also span multiple trials.

If meta-RPE represents a state error signal, then it should be highest when an animal's internal state model generates the highest frequency of incorrect choices. Representing incorrect choices (1) during 85% cues as an absence of anticipatory licking, and incorrect choices (1) during 15% cues as the presence of anticipatory licking, with the converses as correct choices (0), and an abrupt transition at the task reversal point, the calculated proportion incorrect signal over a short retrospective window also peaks in the middle of the reversal like the meta-RPE signal (**Fig. 3-1B**). Whether it originates from the derivative of a smoothed RW curve (**Fig. 3-1A**), a retrospective rolling mean of RPE over many trials (**Fig. 3-1B**), or the incorrect choice proportion (**Fig. 3-1C**), meta-RPE represents a fundamental error signal that suggests the state of the previously stable cue-reward landscape has changed, and behavior should subsequently also be changed. Such a signal could produce a sustained change in value-guided behavior, and therefore meta-RPE reflects reversal learning, and by proxy cognitive flexibility.

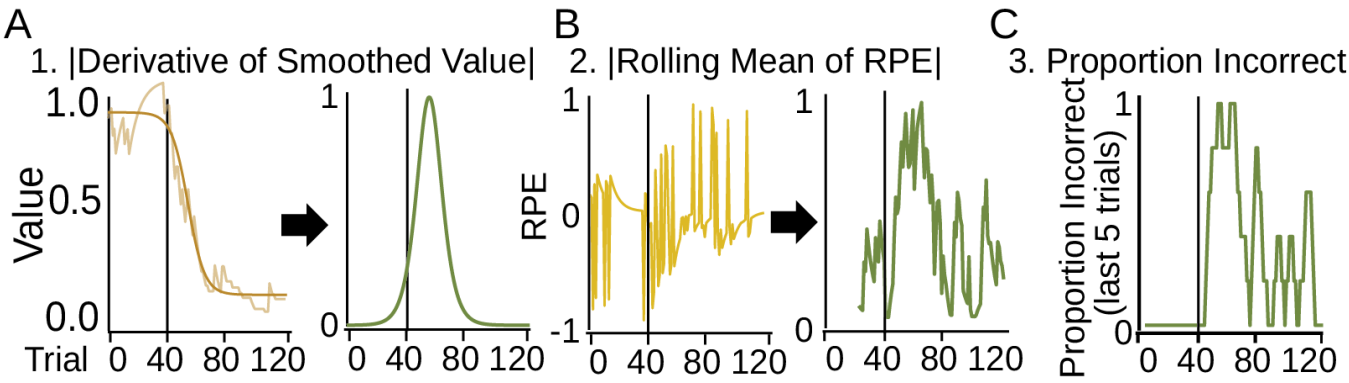

Figure 3-1: Deriving a Cognitive Flexibility Signal. **A**. (**Left**) RW value curve for 85-15 cue from example animal (shaded) smoothed with a sigmoid function. (**Right**) The reversal signal as the absolute value of the smoothed value curve. **B**. (**Left**) Corresponding RPE signals for the shaded curve in A. (**Right**) The reversal signal as RPE smoothed with a large retrospective rolling mean window. **C**. The reversal signal as the proportion incorrect trials in a retrospective window, from same sample animal.

Although the presence of a meta-RPE signal can derive multiple ways from animal choice behavior, whether such a signal was present within the mouse prelimbic cortex (PL) remained untested. In addition to cue stability across the four-odor reversal task (**Fig. 2-3B**), the p-value GLM also evaluated tracked cells for meta-RPE signal encoding of the 85-15 ($Rev_{85\text{-}15}$) and 15-85 ($Rev_{15\text{-}85}$) cues (**Fig. 3-2A**). Surprisingly, hundreds of the thousands of cells in the dataset displayed significant encoding specifically for the $Rev_{85\text{-}15}$ signal, while significantly fewer cells in a separate ensemble coded the $Rev_{15\text{-}85}$ signal (**Fig. 3-2B**). When the reversal signal is artificially moved from the day with the largest coding to stable pre or post reversal days (session permutation analysis), significantly fewer cells encode the $Rev_{85\text{-}15}$ signal, while the same proportion of cells pass the significance threshold for the $Rev_{15\text{-}85}$ signal (**Fig. 3-2C**). This suggests that the $Rev_{85\text{-}15}$ cells' activity pattern is not reflected in their regular everyday firing and is therefore specific to the reversal day. In contrast, the $Rev_{15\text{-}85}$ signal encoding does not appear unique or time locked to the reversal behavior and may therefore not be meaningful. Especially since $Rev_{85\text{-}15}$ encoding occurs in a much higher proportion of cells, it appears that PL preferentially encodes the contingency degradation ($Rev_{85\text{-}15}$) portion of the reversal behavior. This conclusion is consistent with the behavioral findings of others (**Introduction-III**) where permanent or temporary lesions of PL specifically impair contingency degradation of cues following a change in the cue-reward landscape. The strong presence of a PL $Rev_{85\text{-}15}$ ensemble, and relative absence of a corresponding $Rev_{15\text{-}85}$ ensemble, could mechanistically explain why PL lesions specifically impact behavioral changes towards degraded contingency. Consequently, the remainder of this chapter will focus explicitly on the $Rev_{85\text{-}15}$ contingency degradation signal.

There are three potential ways to derive a meta-RPE signal that follows the activity profile of $Rev_{85\text{-}15}$ cells (**Fig. 3-2D**), and although the neurons were isolated using model 1 (**Fig. 3-1A**) another model could better reflect their activity. Qualitatively, trial-averaged $Rev_{85\text{-}15}$ cell activity shares strong similarities to all three models (**Fig. 3-2D**) in that all have low activity early in the

task, that rises shortly following the contingency reversal, peaks partway through the reversal, and decreases to lower activity levels (**Fig. 3-2D**) once the behavior and value function largely stabilize (**Fig. 3-1A**). Notably, when correlating the modeled meta-RPE signals with the observed $Rev_{85\text{-}15}$ cell activity, there was not a significant difference in how well (r^2) each model fit the neural firing (**Fig. 3-2E**). This indicates that all three models are likely representative of the same underlying phenomenon: a continuous aggregation of repeated seeking errors towards omitted rewards following a cue with degraded contingency. Likewise, the $Rev_{85\text{-}15}$ ensemble parallels large-scale changes in the internal value function (**Fig. 3-1A&3-2D**), a high frequency of reward prediction errors over a series of trials (**Fig. 3-1B&3-2D**), and frequent incorrect choices (**Fig. 3-1C&3-2D**) equally well (**Fig. 3-2E**). Altogether, the $Rev_{85\text{-}15}$ ensemble activity reflects the summation of repeated behavioral errors over many trials following contingency reversal.

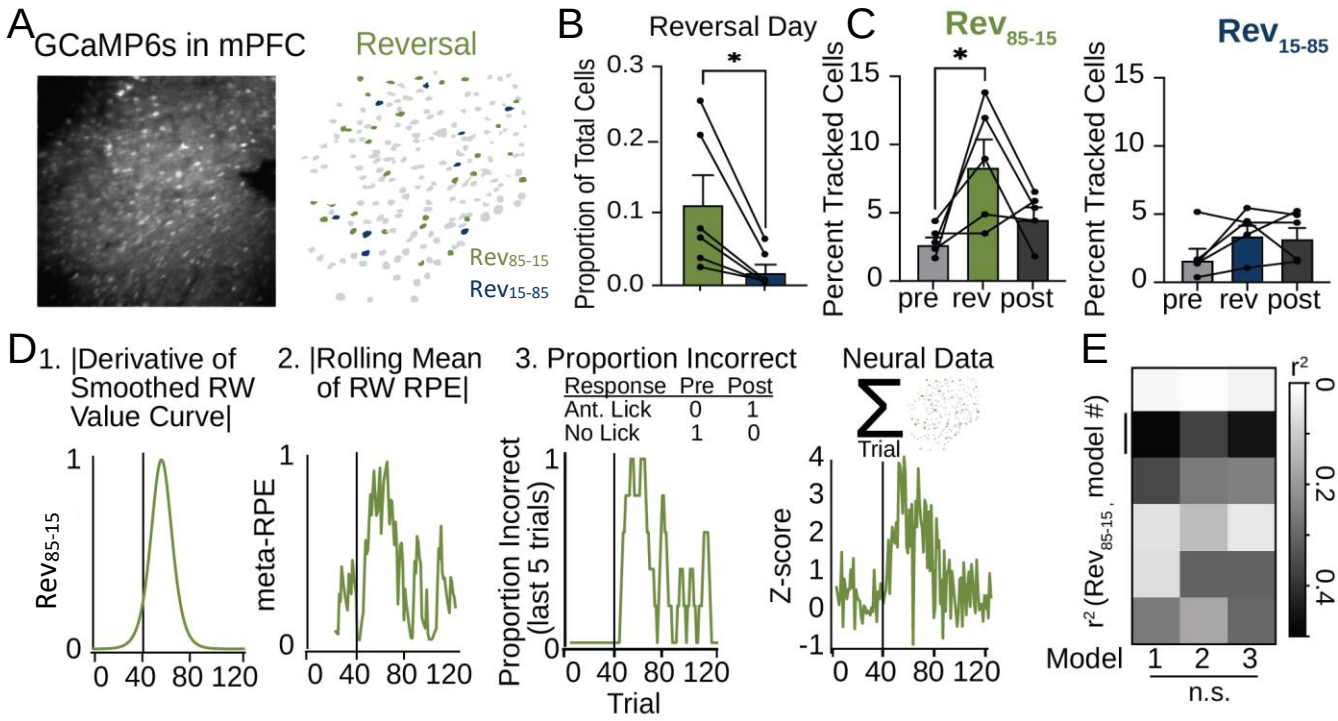

Figure 3-2: Detecting a Cognitive Flexibility Signal in Prelimbic Cortex. **A**. Sample field of view (FOV) in PL, with extracted cell masks (gray) overlaid with significant $Rev_{85\text{-}15}$ (green) or $Rev_{15\text{-}85}$ (blue) cells from the p-value GLM results. **B.** Proportion of cells in reversal day FOV significant for $Rev_{85\text{-}15}$ or $Rev_{15\text{-}85}$ signals. There are significantly more $Rev_{85\text{-}15}$ cells than $Rev_{15\text{-}85}$ cells in all FOVs for all animals (n=6). **C**. Results of session permutation analysis: When the reversal day $Rev_{85\text{-}15}$ signal is moved to stable pre- or post-reversal days, there are significantly fewer tracked cells that display this pattern of behavior. There is not a significant difference for cells displaying the $Rev_{15\text{-}85}$ activity pattern across days. **D**. Comparison of derived $Rev_{85\text{-}15}$ signals from choice behavior to extracted activity from $Rev_{85\text{-}15}$ cells in example animal. Vertical line indicates reversal point. **E**. Model signal correlations to neural data. There is not a significant difference in how well models 1-3 fit the neural data across animals.

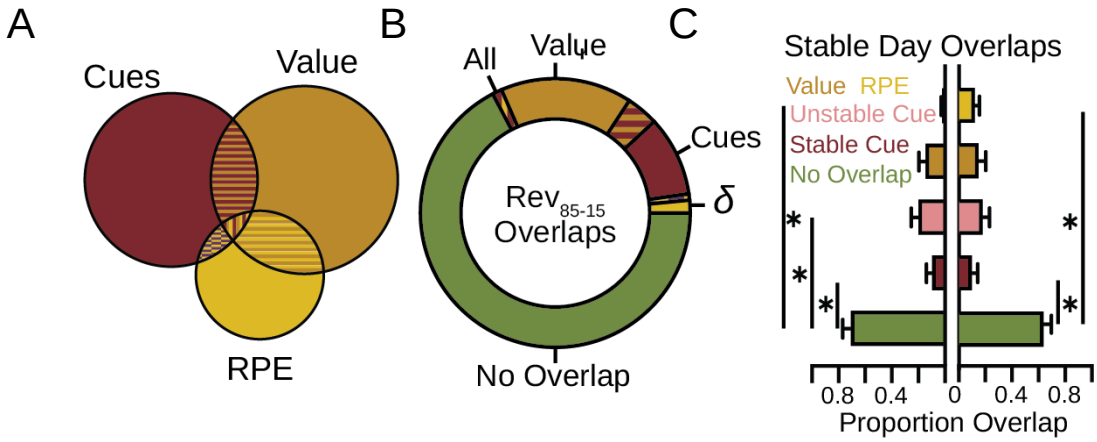

Figure 3-3: The $Rev_{85\text{-}15}$ Ensemble is Distinct from Other Task-related Ensembles. **A**. Extracted cue (red), value (orange), and RPE (yellow) ensembles from p-value GLM are largely non-overlapping. **B**. On the reversal day, the majority of $Rev_{85\text{-}15}$ cells are not significant for other task functions, including cue, value, and RPE, displaying little overlap (green). **C**. When tracked across days, $Rev_{85\text{-}15}$ cells largely do not overlap with stable or unstable cue, value, or RPE cells.

Although the $Rev_{85\text{-}15}$ ensemble does not display a comparable activity pattern on stable behavioral days (**Fig. 3-2C**), it could perform another function related to the task, perhaps tracking cues, value, or single-trial RPEs that extends additionally to reversal days. Interestingly, cues, value and RPE are represented in distinct ensembles on the reversal day (**Fig. 3-3A**). In some ways this is unsurprising as the three quantities represent distinct questions: what is that? (cue), do I want that? (value), and should I have wanted that? (RPE). The $Rev_{85\text{-}15}$ ensemble activity reflects computations derived from all three signals, in that the activity is specific for the 85-15 *cue* (**Fig. 3-2B**) and tracks repeated *RPEs* (**Fig. 3-1B**) in choice behavior (**Fig. 3-1C**) using *value* predictions (**Fig. 3-1A**), so the signals could be partially overlapping. However, when comparing additional GLM categories for which $Rev_{85\text{-}15}$ cells display significant encoding, the vast majority of $Rev_{85\text{-}15}$ cells are non-overlapping with cue, value, or RPE cells on reversal (**Fig. 3-3B**) or stable (**Fig. 3-3C**) behavior days. This suggests that the $Rev_{85\text{-}15}$ ensemble, and by proxy, cognitive flexibility more generally, may ask a distinct question— Has the state of the cue reward landscape changed? —using a unique ensemble from those used to ask iterative trial-by-trial questions about the behavior task.

The presence of a specific contingency degradation signal in PL within $Rev_{85\text{-}15}$ cells is interesting on its own (**Fig. 3-4A**). However, perhaps more interesting still is what such a signal could be used for beyond passively tracking changes in task structure. The PL lesion literature (**Introduction-III**) suggests that animals with temporary or permanent lesioning of PL function take *much longer* to degrade the contingency of previously rewarded cues, but that they do eventually adapt their behavior to the new contingencies. Since animals without a functioning PL learn reversals slower, and functioning PL contains a $Rev_{85\text{-}15}$ ensemble, it is possible that the $Rev_{85\text{-}15}$ signal expedites the reversal learning process. Fitting RW models to pre-reversal and post-reversal periods in addition to the full reversal paradigm (**Fig. 3-4B**) can test whether animals show an increased learning rate specifically for reversed cues during the reversal period. Notably,

animals displayed a significantly elevated learning rate (α) in RW models that included the reversal period compared to those trained only on pre- or post-reversal days for the 85-15 cue (**Fig. 3-4C**). The 85-15 reversal learning rate was also significantly elevated compared to stable 85-85 or 15-15 cues but similarly elevate to the learning rate for the other 15-85 reversed cue (**Fig. 3-4C**). What neural circuitry could account for the learning rate increase for contingency elevation of the 15-85 cue is addressed in **Chapter 4** of this work.

Although both the $Rev_{85\text{-}15}$ cell signal and 85-15 learning rate were higher during the reversal period, the two had yet to be connected directly. An enhanced form of the R-W equation (**Eq. 3-3**) can test whether allowing the averaged $Rev_{85\text{-}15}$ signal (Rev_N) to modify learning rate can account for the observed elevation in learning rate (**Fig. 3-4C**) during the reversal period and compare the fit to actual animal choice behavior of a value model that includes a dynamic meta-RPE signal to one that does not.

Equation 3-3: $Q[t+1] = Q[t] + \alpha\delta * Rev_N$

Excitingly, the model from **Equation 3-3** generates learning rates for the 85-15 cue across the reversal period (**Fig. 3-4D**) that are not significantly different from those observed in stable behavior pre and post reversal (**Fig. 3-4E**). This suggests that the z-scored activity of the $Rev_{85\text{-}15}$ ensemble is able to completely mathematically account for the elevated learning rate during the reversal, and correlationally suggests that the $Rev_{85\text{-}15}$ ensemble may be responsible for the observed increase in learning rate.

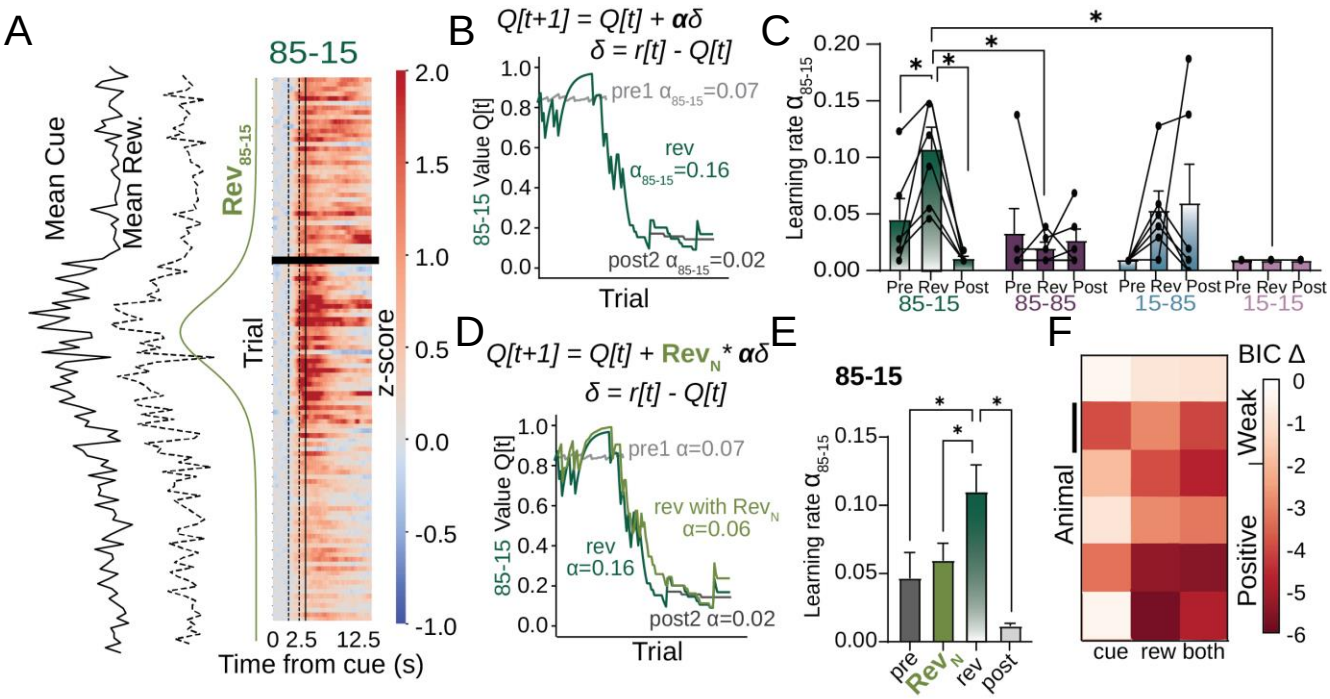

Figure 3-4: $Rev_{85\text{-}15}$ Signal Improves Value Model fit to Choice Behavior. A. Trial-average cue and reward signals plus PSTH for $Rev_{85\text{-}15}$ ensemble across the reversal task. **B**. RW models fit before (light gray), across (dark green), and after (dark gray) the reversal in sample animal. **C**. Summary statistics for learning rate (α) across cues and reversal stages. There is a significant elevation in the learning rate for 85-15 and 15-85 cues during reversal. **D**. Enhanced value model (**Eq. 3-3**) that allows $Rev_{85\text{-}15}$ signal (Rev_N) to scale αδ term, in same representative animal. **E**. The enhanced model abolishes learning rate increases observed in **B**. **F**. The enhanced model better fits animal choice behavior in all cases, using a Bayes Information Criterion (BIC) fit score. The lowest scores involve the $Rev_{85\text{-}15}$ signal averaged across cue and reward.

Moreover, adding the $Rev_{85\text{-}15}$ activity to a standard RW model performs dynamic scaling of the αδ term since the signal rises and falls over the course of many trials. This suggests that the learning rate is actually faster still than the standard value model (**Fig. 3-4B&C**) estimates in the middle of reversal learning. But does allowing αδ to dynamically scale in an enhanced value model actually improve the fit to animal choice behavior? The Bayes Information Criterion (BIC), a score that penalizes models for adding extra terms, asks how well its predictions fit ground truth data and is supposed to be minimized like golf, is lower (**Fig. 3-4F**) for the enhanced model (**Eq. 3-3**) compared to the standard RW model (**Eq. 3-1**). The best fit to behavior (i.e. lowest BIC) derives from the model that allows αδ to scale by the mean $Rev_{85\text{-}15}$ ensemble activity from "both" the cue and reward periods (**Fig. 3-4F**). That the inclusion of the signal during bother periods best fits behavior suggests that the $Rev_{85\text{-}15}$ may modulate choice behavior during both the choice to anticipatory lick during the cue period and modulate the updating of future choice behavior.

These results provoke an additional interesting question: why does allowing the $Rev_{85\text{-}15}$ signal to dynamically modulate trial-by-trial RPE (**Fig. 3-4D**) better fit the animal's choice behavior (**Fig. 3-4F**) than a statically elevated learning rate (**Fig. 3-4B**)? Answering this question requires a more careful analysis of the relationship between RPE and behavioral change during the reversal task. Somewhat unique to this task, in addition to quantifying whether an animal "chose" to seek a reward after presentation of a given cue (anticipatory licking above background levels), it is also possible to analyze *how much* the animal chose to seek that reward via *how many* anticipatory licks the animal performed above background. This enables additional analysis comparing how well simple value models predict the amount of choice behavior through direct comparison to the actual changes in the numbers of anticipatory licks to a given cue. Looking closely at the 85-15 cue, RPE significantly predicts the change in the number of normalized anticipatory licks on the subsequent trial in the epochs before and after reversal, as the RW model predicts (**Fig. 3-5A**). Surprisingly, the significant relationship between RPE and behavioral change is not maintained during the actual reversal in any animals studied (**Fig. 3-5A&B**), nor at the

population level (**Fig. 3-5B**). That RPEs are still occurring, represented in a subset of PL neurons (**Fig. 3-3A**), and could on their own generate a reversal curve (**Fig. 3-4B**), but do not significant predict changes to the actual behavior of the animal (**Fig. 3-5B**), has not been well characterized in previous literature.

Spatial frequency maps further visualize the disconnect between RPEs and licking behavior during the reversal period. Anticipatory licking across the task largely follows a bimodal distribution, with distinct peaks for pre- and post-reversal licking, and an intermediate area mostly populated with reversal-associated licking decreases (**Fig. 3-5C top**). When compared against RPE, it is immediately apparent that despite large RPEs, licking remains aligned with the pre-reversal distribution (white dashed arrow in **Fig. 3-5C center**), whereas intermediate RPEs dominate the transition zone. In contrast, when comparing anticipatory licking against the Rev_{85-15} signal (mRPE), the largest mRPEs are associated most strongly with licking away from pre and post reversal stable points (**Fig. 3-5C, lower**). Altogether, the mRPE signal visually parallels the reversal licking behavior more reliably than trial-by-trial RPE.

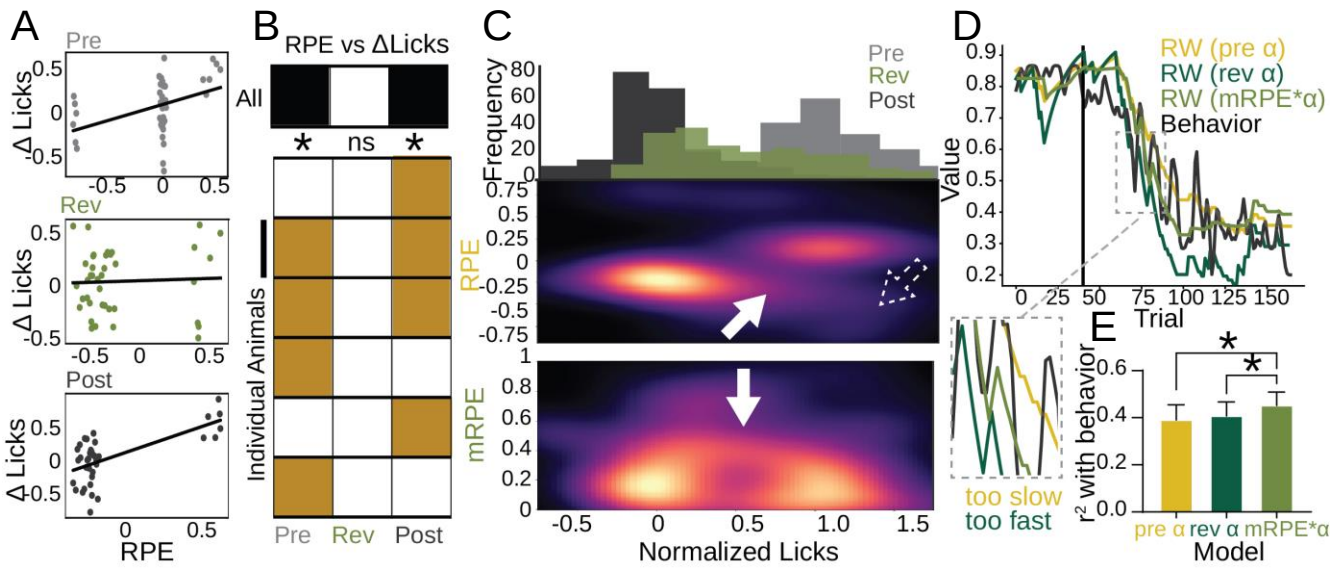

Figure 3-5: RPE Alone is Insufficient to Explain Reversal Licking Behavior. **A**. Scatterplots of relationship between RPE on one trial, and the change in normalized licking in the subsequent trial before (light gray) during (green) and after (dark gray) reversal for an example animal (indicated by black line on *B*). There is a significant relationship between RPE and behavioral change before and after, but not during reversal. **B**. Summary statistics from regressions between RPE and behavioral change before (left column) during (center) and after (right) reversal. Significant relationships are colored, non-significant are white. RPE vs. lick change is significant before and after reversal in the pooled population (black) and in majority of individual animals (gold), but not during the reversal. **C**. (**Top**) Frequency histogram of normalized licks pre (light gray), during (light green) and post (dark gray) reversal. Note the distinct separation of the pre and post-reversal lick distributions. (**Center**) Spatial heatmap of the relationship between the lick distributions and trial-associated RPE signals. When RPE is largest (i.e. immediately following reversal), licking has not shifted from the pre-reversal distribution (dashed arrow). The reversal portion of the distribution (in between the pre and post distributions) is associated with only moderate RPEs (white solid arrow). (**Lower**) Spatial heatmap comparing normalized licks and the reversal signal (mRPE). Note that when mRPE is highest, licking away from the pre- and post-hotspots in the map (white arrow). **D**. Value plots of example animal reversal behavior using the pre-reversal learning rate (yellow), learning rate fit across reversal (dark green), and learning rate fit across reversal when mRPE is also allowed to modulate RPE (light green), as well as licking behavior (black). Zoomed inset highlights period the relationship between models and reversal behavior. **E**. Correlations between value model predictions (as in *C*) and reversal behavior during reversal period across animals. There is a significant increase in the correlation between value predictions and behavior when mRPE signal is added to the model.

Testing whether the mRPE*α model better fits licking behavior than alternatives can increase confidence in the results from fitting the mRPE*α model to choice behavior. First, a pre-α model tested whether any increase from the pre-reversal learning rate was necessary to explain the reversal licking behavior. This model reduces the value function too slowly compared to the licking behavior (**Fig. 3-5D**). Thus, an increased learning rate during reversal is indeed needed to explain the behavior. A rev-α model tested whether a dynamically increased learning rate is necessary to explain reversal, or if an increased static α will do. This model reduces the value function too quickly compared to the licking behavior (**Fig. 3-5D**). In contrast to the pre-α and rev-α models, the mRPE*α value function visually decreases at a similar rate to the licking behavior in the example animal (**Fig. 3-5D**). At the population level, the mRPE*α model has a significantly stronger correlation coefficient with anticipatory licking behavior compared to the pre-α and rev-α models (**Fig. 3-5E**). Therefore, via two behavioral metrics—choice (**Fig. 3-4**) and number of anticipatory licks (**Fig. 3-5**)—allowing mRPE to scale RPEs fits reversal behavior the best, providing support that the observed $Rev_{85\text{-}15}$ signal is biologically meaningful.

Taken together, these results suggest that although animals could perform contingency degradation based solely upon trial-by-trial RPE, they do not. One can only speculate on why animals adopt the strategy where mRPEs drive the largest changes in reversal behavior during intermediate RPEs. The general need for an increased impact of RPE on behavioral change (i.e. not pre-α model) during contingency degradation is clear: the more behavior decreases following a reward omission, the faster animals stop performing work towards a cue with little utility. The R-W model fits the generalities of such a strategy across the reversal (the rev-α model). Truly, in near deterministic tasks, this is the most efficient strategy. However, the world is largely probabilistic, so not abandoning a strategy that has previously been successful (perseveration) and changing behavior in the face of new information (contingency degradation) must be more carefully weighed.

Perhaps the role of mRPE is to perform as a signal that balances abandoning a previously successful task strategy against increasing evidence that the cue-reward landscape has changed. That it can derive from either a sum of RPEs over a long time period or the number of incorrect choices in a smaller window provides mathematical support for such a notion in that both involve collecting errors over time. As the number of errors increases, so can an animal's confidence that their previously strategy is no longer useful. Each animal weights accumulated errors differently before they decide to act, and a longer retrospective window can be directly associated with an mRPE signal that takes longer to rise. Exactly how animals determine their mRPE window size is outside the scope of this work, but an interesting future question. Nevertheless, the modeling in **Figure 3-4&5** suggests that only once the mRPE signal has risen significantly does the animal truly begin to change his seeking behavior in earnest.

That a reversal learning signal (mRPE) should peak in the middle of contingency reversal is somewhat at odds with the meta-learning models of other groups. In addition to primate researchers (ex: Soltani & Izquierdo, 2019), Grossman and Cohen also note an elevated learning rate during mouse reversal but suggest this is due to an increased impact of the first large magnitude RPEs (Grossman et al., 2022). Although it agrees in dynamic modulation of the relationship between RPE and behavioral change, their "unexpected uncertainty" model of reversal behavior is at odds with the behavioral conclusions of this work because that it predicts the large RPEs following contingency reversal drive larger behavioral changes still, in direct contrast to the lack of behavioral change to the largest RPEs in **Figure 3-5**. In support of the mRPE model, their model appears to reflect too quick of a decrease in the early stages of reversal compared to the mouse behavior (see Grossman et al. 2022 Fig. 2F) that mRPE does not.

The existence of a dynamic mRPE signal that is highest when behavior is changing the fastest, computed by accumulating errors over many trials, and effectively models contingency degradation, has not been characterized in depth behaviorally or in the brain prior to this work.

Whether this signal can be explained by input from other systems implicated in PL reversal is the focus of **Chapter 4** of this work.

References

Grossman, C. D., Bari, B. A., & Cohen, J. Y. (2022). Serotonin neurons modulate learning rate through uncertainty. *Current Biology*, *32*(3), 586-599.e7. https://doi.org/10.1016/j.cub.2021.12.006

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*(3), 314–321. https://doi.org/10.3758/BF03210755

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, *20*(10), 635–644. https://doi.org/10.1038/s41583-019-0180-y

CHAPTER 4: PRELIMBIC CORTEX, DOPAMINE, AND THE VENTRAL TEGMENTAL AREA

A significant subset of prelimbic cortex (PL) cells significantly encode a meta-RPE (mRPE) signal that reflects the error profile during the contingency degradation component of reversal learning over many trials (**Chapter 3**). Mathematically, this signal expedites learning through a dynamically increased impact (α) of intermediate reward prediction errors (RPEs) in the middle of contingency degradation, which fits choice and licking behavior better than static α models. This signal could either arise via direct computation within PL or reflect a copy of input from other brain regions. Given the tight linkage between RPE and the activity of the ventral tegmental area (VTA) dopamine neurons (Schultz, 2016), along with reciprocal connectivity between PL and VTA (Gongwer et al., 2023; Hui & Beier, 2022), it is possible that the meta-RPE signal in PL reflects computations that have already taken place alongside trial-by-trial RPE within the VTA. Conversely, it is also possible that PL computes the longer timescale mRPE signal and sends it to VTA to physiologically modulate RPE encoding. Thus, the purpose of this chapter is twofold:

I. To test whether the PL 85-15 mRPE signal reflects VTA dopaminergic input.

II. To determine if the neurons that project to the VTA encode the 85-15 mRPE signal.

I. DOES DOPAMINE RELEASED INTO PRELIMBIC CORTEX REFLECT META-RPE?

It is possible that although a subset of PL neurons significant encode 85-15 mRPE signal, their activity reflects input from other regions. Since the mRPE signal is inherently a learning signal, and learning signals are traditionally associated with neuromodulator input (Bazzari & Parri, 2019), including meta-learning (Doya, 2002), it is likely that if the PL mRPE signal reflects input from other brain regions, it would arrive via neuromodulatory input. Adding credence to the longer-timescale mRPE signal reflecting neuromodulatory input, neuromodulators are molecules released from neural cells that affect cellular signaling via complicated cascades with longer timescales than traditional fast neurotransmitters (Weis & Kobilka, 2018). Since the mRPE signal

takes time to develop, and neuromodulator-mediated signaling likewise take time to affect cellular activity, the timescales of the two signals are theoretically compatible.

Behavior experiments provide valuable information about the contribution of neuromodulatory signaling in PL to reversal learning behavior. The neuromodulators most commonly associated with learning are dopamine (DA), norepinephrine (NE), and serotonin (5-HT) (Bazzari & Parri, 2019). Prelimbic cortex has receptors for all three neuromodulators (Arnsten, 1998; Puig & Gulledge, 2011), so $Rev_{85\text{-}15}$ cell activity could reflect input from any of the them. However, since $Rev_{85\text{-}15}$ cells are tightly linked with reversal behavior, if neuromodulatory input explained their firing, behavioral manipulations of that signal would likely impair the reversal behavior. Although serotonin signaling has been linked to meta-learning signals (Grossman et al., 2022), selective blockade of serotonin signaling in PL does not impair reversal behavior (Alsiö et al., 2020). In contrast, serotonergic blockade of the nearby orbitofrontal cortex (OFC) does impair reversal learning (Alsiö et al., 2020). Given that Grossman and colleagues link serotonergic signals to meta-learning *immediately following* contingency reversal, and OFC activity is highest in the same period (Banerjee et al., 2020), it is possible that serotonin modulates OFC activity to expedite early-stage detection of a change in the task environment. However, since the mRPE signal is highest in the *middle* of the reversal, and serotonin blockade in PL does not impair reversal behavior, such a serotonergic signal is unlikely to explain the $Rev_{85\text{-}15}$ signal in PL.

Thus, the catecholamines norepinephrine and dopamine remain as candidates to explain the $Rev_{85\text{-}15}$ signal in PL. Microdialysis provides a direct measurement of signaling molecule concentration in a brain area (Müller, 2002), and has been used to measure DA and NE levels in PL during contingency reversal. Notably, dopamine levels significantly rise during first contingency reversal compared to control animals with the same training history, but that do not experience a reversal (van der Meulen et al., 2007). In contrast, NE levels remain constant during reversal and do not differ from controls (van der Meulen et al., 2007). Supporting a lack of NE contribution to PL-mediated effects on contingency reversal, blocking PL NE signaling in normal

rats does not impact reversal behavior (Jett & Morilak, 2013). However, like serotonin, NE blockade in nearby OFC does impact reversal behavior (Ogg et al., 2023). Dopamine thus remains as the most likely candidate to influence reversal activity in PL. That selective DA antagonism in PL leaves initial learning unchanged but impairs reversal (Floresco et al., 2006) further supports this claim. Therefore, the purpose of the first portion of this chapter is to test the hypothesis that $Rev_{85\text{-}15}$ cell signaling reflects the profile of dopaminergic input into PL.

Before assessing the direct relationship between DA and $Rev_{85\text{-}15}$ cell activity in PL, it is important to characterize DA release in PL on its own. Towards this end, eight C57BL6/J mice injected with AAV2/9-hSyn-$GRAB_{DA}$3h and implanted with 6x400um photometry fibers performed the reversal learning task (**Fig. 4-1A&B**). $GRAB_{DA}$ is an indicator much like GCaMP, but built off a dopamine D1-receptor backbone instead of calmodulin to imbue dopamine specificity (Zhuo et al., 2023). Since cortical dopamine levels are generally low compared to those in the deeper brain, especially striatum, the 3h variant was specifically constructed for cortical regions (Zhuo et al., 2023), and perfect for this application. After collection on a TDT RZ10x with 470nm single-photon excitation, an OASIS-based algorithm (Friedrich et al., 2017) denoised the $GRAB_{DA}$ signal using published kinetics (Zhuo et al., 2023) to improve signal-to-noise and remove fluctuations in the signal too fast to be attributable to $GRAB_{DA}$ activity. All animals learned the task well, displaying stable licking behavior before and after reversal (**Fig. 4-1C&D**)

Like the PL neurons (**Fig. 2-2F&2-4A**) and the animal's behavior (**Fig. 4-1 C&D**), PL DA cue-evoked activity was stable before and after reversal in an example animal (**Fig. 4-1E**) and across the population (**Fig. 4-1F**). DA cue signaling for the 85-15 significantly decreased across reversal, and significantly increased for the 15-85. However, again like the PL neurons (**Fig. 2-5C**), there was not a significant difference after reversal between 85-15 and 15-15 mean cue amplitude, nor between the 85-85 and 15-85. Encoding of licking-related motor signals and not anything meaningful about cue information could explain this result, since the behavior follows the same significance pattern (**Fig. 4-1D**). Yet, when analyzed using a generalized linear model

(GLM) that included lick dynamics as a predictor (**Chapter 1**), cue-encoding still emerged as a significant variable in the majority of animals (**Fig. 4-1G**). Thus, although the PL DA cue dynamics parallel the behavior, they are not simply a motor correlate and are likely meaningful in a more abstract associative context. Since cues with the same associated rewards share comparable evoked activity before and after reversal , this suggests that dopamine release in PL does not carry set information about cues and appears to focus on contingency with stronger representation of higher value cues that predict more likely rewards. Furthermore, the stability in cue-evoked DA dynamics across reversal also allows for an isolation of potential signatures during the reversal period are not only biproducts of representational drift.

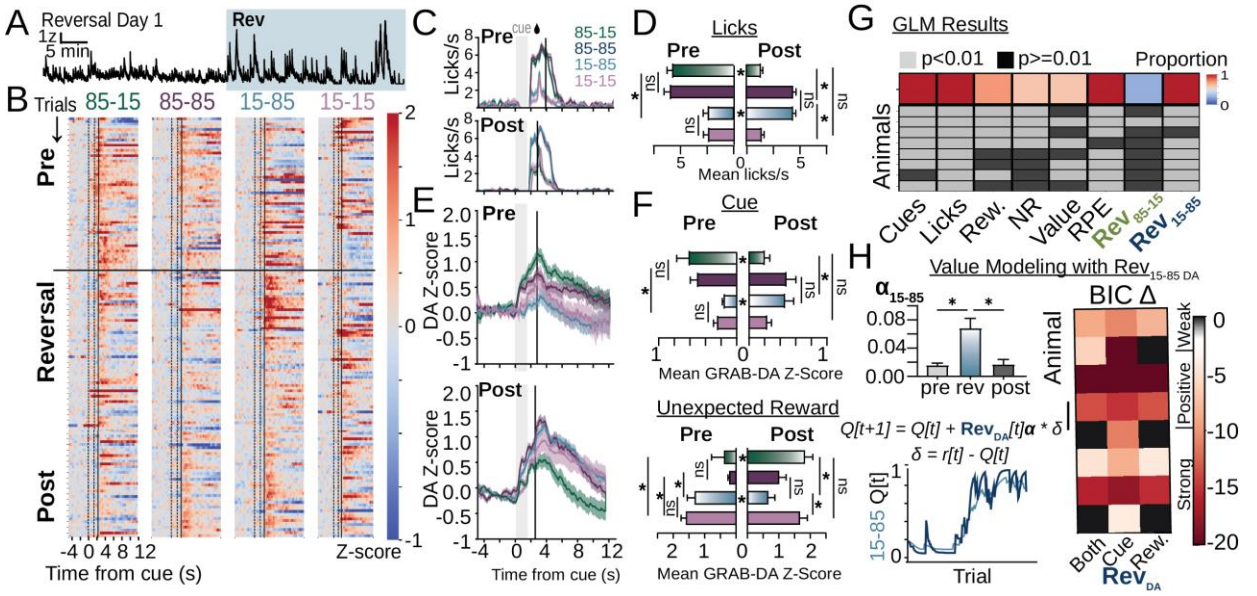

Figure 4-1: Fiber Photometry Recordings of Dopamine Release in Prelimbic Cortex. **A.** Representative $GRAB_{DA}$-3h signal during reversal day 1. **B**. Cue-locked time heat plots across trials for recordings on stable pre, post, and reversal days for example animal. **C**. Lick traces before and after reversal for sample animal. **D**. Mean +/- SEM $GRAB_{DA}$-3h traces for pre and post reversal stable days. **E**. Population statistics for photometry animal performance in reversal task. There is not a significant difference between licking to 85% or 15% cues before or after reversal. **F**. (**Top**) Mean cue activity before and after reversal on stable days. There is not a significant difference in DA release between 85% or 15% cues before or after reversal. (**Lower**) Dopamine release is elevated to rewards released following 15% cues without preceding anticipatory licking both before and after reversal. **G.** GLM results for $GRAB_{DA}$-3h signal across animals, where light gray indicates significant encoding. Population summary proportions indicate DA signal is significant in the most animals for cue, lick, RPE, and $Rev_{15\text{-}85}$ encoding. **H.** RW value models also show elevated learning rate for 15-85 cues in photometry animals. Adding the $GRAB_{DA}$-3h signal during the cue period to scale learning rate improves model fit.

Dopaminergic signals are most strongly associated with reward prediction error (RPE), the difference between the actual and predicted reward (Schultz, 2016). If PL DA signals likewise reflect RPE, then the amplitude of the $GRAB_{DA}$ signal would be highest following reward delivery when there is the largest difference between the predicted and actual reward. Using the lack of anticipatory licking as a surrogate for a lack of reward expectation, the $GRAB_{DA}$ signal amplitude is indeed high on trials with behaviorally unexpected rewards (**Fig. 4-1F**), especially following cues with low reward probability (15%). The magnitude of $GRAB_{DA}$ signal towards unexpected rewards on 15% trials remains elevated following reversal, only shifting to reflect the changed cue probabilities (**Fig. 4-1F**). That the DA amplitude is significantly lower for rewards following 85% cues (**Fig. 4-1F**) even when animals do not anticipatory lick on a given trial suggests that the rewards may still be neurally predicted. Especially since there is a significant increase in amplitude to the behaviorally unexpected rewards following the 85-15 cue after its contingency has been degraded across reversal, this is likely to be the case. Given differences in signaling towards delivery of rewards of equal volume, the observed RPE coding is unlikely to be attributable solely to motor correlates, but the GLM tested this directly. The vast majority of animals displayed significant encoding of the Rescorla-Wagner (RW) modelled RPE signal within PL DA dynamics (**Fig. 4-1G**).

The PL DA signal is noticeably higher during reversal (**Fig. 4-1A&B**), consistent with previous reports (van der Meulen et al., 2007). Since the PL DA signal significantly encodes RPE information, and RPE-associated amplitudes are generally larger than cue-associated amplitudes (**Fig. 4-1 E&F**), it follows that observed increases in PL DA could simply reflect an increased frequency of RPEs following contingency elevation of the 15-85 cue. Looking at cue-locked trial activity of an example animal (**Fig. 4-1B**) there is a strong increase in the DA signal following 15-85 reversal, consistent with the observed significant RPE coding (**Fig. 4-1G**). Since the DA signal encodes trial-by-trial RPE during reversal of the 15-85 cue, it could also encode a mRPE signal for the 15-85 cue similar to that towards the 85-15 cue (**Fig. 3-2**) in the PL neurons. Modeling

such a signal via inclusion in the GLM revealed significant encoding of a mRPE contingency elevation signal in PL DA in the vast majority of animals (**Fig. 4-1G**). Not only did the PL DA signal reflect standard RPE coding, but it also reflected longer timescale mRPE encoding only for the 15-85 cue. Removing the 85-15 mRPE signal from the GLM did not impact encoding predictions for the majority of animals (**Fig. 4-1G**). This suggests that dopaminergic input is unlikely to explain the 85-15 mRPE encoding observed in PL neurons. Still, since PL DA reflected 15-85 mRPE signaling, its information could be useful for the reversal behavior in a different way.

Like the 85-15 (**Fig. 3-4C**), RW modeling revealed an elevated learning rate towards the reversed 15-85 cue in the $GRAB_{DA}$ cohort (**Fig. 4-1H**). Training a similar mRPE*α model using the cue, reward, or combined $GRAB_{DA}$ signal expedites the learning and fits the choice behavior better across animals only during the cue period (**Fig. 4-1H**). This is somewhat in contrast to the neural results where while the cue model did fit the data better than a static α term, the average of $Rev_{85\text{-}15}$ cell activity across both the cue and reward fit the choice behavior better still (**Fig. 3-4F**). One possible explanation for this discrepancy is that the dopamine release, clearance from the synapse, and downstream GPCR signaling takes time, so releasing a mRPE signal computed from previous trials ahead of a given trial's RPE primes dopamine sensitive cells towards an enhanced response to the trial's RPE. If both the RPE and RPE enhancement signal were released concurrently, this could have an additive instead of multiplicative effect on RPE and change behavior more slowly.

The analysis of DA signaling dynamics in PL revealed that DA significantly codes for cue, RPE, and 15-85 mRPE. This suggests that it is unlikely that the $Rev_{85\text{-}15}$ cell signal we observed towards the 85-15 mRPE could be a copy of dopaminergic input. Performing concurrent, dual-color two-photon calcium imaging of GCaMP6s (920nm excitation) and the red rGRABDA-3h variant (1040nm excitation) more directly tested the relationship between PL DA and PL neural activity (**Fig. 4-2A**). In addition to standard GLM variables, including the DA signal (shifted in time

0-1s) as a predictor allowed a test of the hypothesis that despite both displaying elevations during the reversal, PL DA activity does not significantly predict the activity of PL $Rev_{85\text{-}15}$ neurons (**Fig. 4-2B**). Of 546 total cells tested, only the majority of $Rev_{15\text{-}85}$ (of which there were very few) and RPE coding cells also displayed significant encoding of the PL DA signal (**Fig. 4-2C**). Given that the DA signal alone also had very strong coding of both RPE and a $Rev_{15\text{-}85}$ signal (**Fig. 4-1G**), it is encouraging that the neural populations that significantly code these signals also significantly reflect the dopamine environment in their activity. Moreover, in alignment with the findings that the DA elevation surrounding reversals (**Fig. 4-1A**) is centered around the 15-85 contingency enhancement (**Fig. 4-1B**), and therefore does not encode the 85-15 meta-RPE signal (**Fig. 4-1G**), most $Rev_{85\text{-}15}$ cells did not significantly encode the dopamine state of their surrounding environment (**Fig. 4-2C**). This suggests that although both reversal cells and prefrontal dopamine increase during the reversal period, they play distinct and dissociable roles.

Surprisingly, although PL DA significantly encodes cue information (**Fig. 4-1G**), this information does not dramatically drive cue representations in the majority of PL cue cells (**Fig. 4-2C**). One explanation of this discrepancy is that while both PL neurons and PL DA release are active during the cue period, their signaling dynamics are incongruent. The example traces in **Fig. 4-1E** and summary data in **Fig. 2-5B** do appear to have different rise times, with dopamine rising early during the odor presentation and cue cells largely increasing their activity later during the delivery and through the trace interval. Furthermore, the modeling in **Fig. 4-1H** suggests that the $Rev_{15\text{-}85}$ signal modulates choice behavior towards the reversed 15-85 cue most strongly during the cue period and given the paucity of $Rev_{15\text{-}85}$ neurons (**Fig. 3-2C**) compared to cue neurons (**Fig. 3-3A**), it follows that the neural cue signal and the $Rev_{15\text{-}85}$ signal are separable quantities, encoded separately.

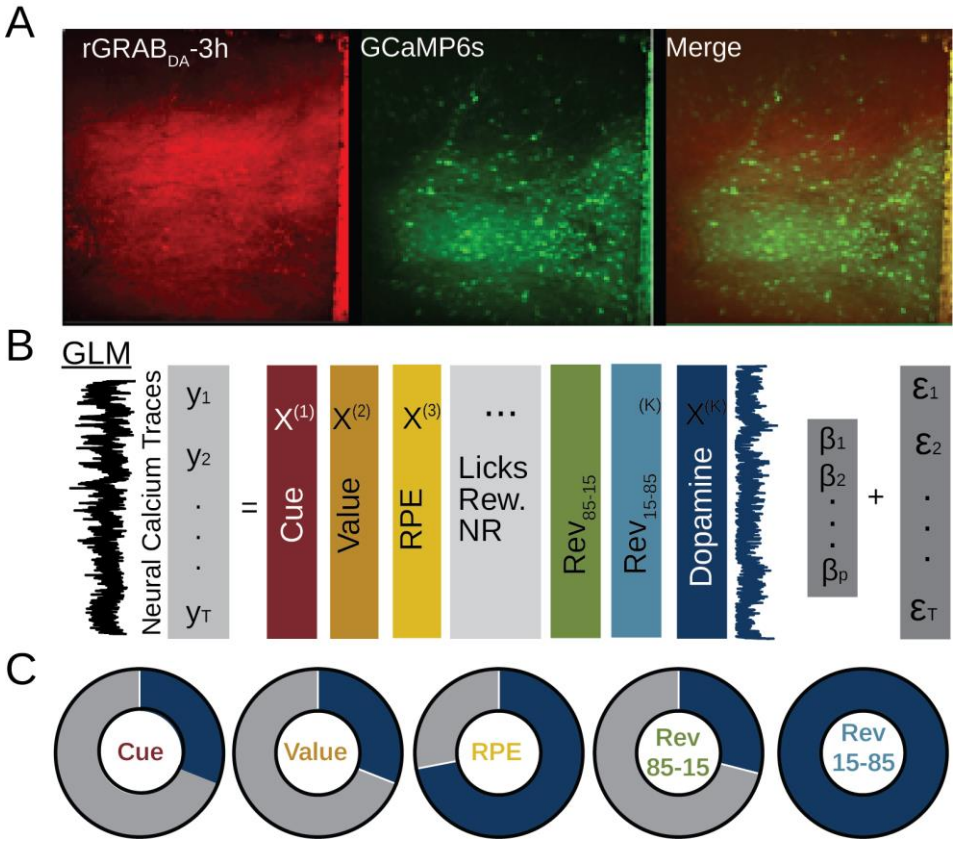

Figure 4-2: Dual Recordings of Neural Activity and Dopamine Release in Prelimbic Cortex. **A.** Representative field of view (FOV) with rGRAB$_{DA}$-3h signal (left), GCaMP6s signal (center) and merged signal. Extracted GCaMP6s ROIs were excluded from rGRAB$_{DA}$-3h channel before analysis. **B**. Structure of GLM, revised to include DA signal as a predictor of neural activity. DA kernel included 0-1s time offset activity from rGRAB$_{DA}$-3h channel. **C**. Proportion of cells that significantly encode both dopamine kernel and other task variables of interest (dark blue) or not (gray).

The results from the fiber photometry (**Fig. 4-1**) and simultaneous neural (**Fig. 4-2**) $GRAB_{DA}$ recordings strongly suggest the 85-15 contingency degradation mRPE signal within $Rev_{85\text{-}15}$ neurons in PL is not of dopaminergic origin. However, the dopamine that reaches PL seems to have a mRPE signal of its own, but for the 15-85 contingency enhancement portion of the reversal learning. Given the minimal $Rev_{15\text{-}85}$ coding in PL (**Fig. 3-2C**), and sensitivity of $Rev_{15\text{-}85}$ cells to dopaminergic input (**Fig. 4-2C**), it is unlikely that the contingency elevation mRPE signal originates in PL. This indicates that computation of mRPE signals is unlikely to be restricted to PL. However, since the $Rev_{85\text{-}15}$ signal was not represented in the DA input to PL, and VTA DA neurons are a convergence site for learning information across the brain (Stuber, 2023), it is possible that the $Rev_{85\text{-}15}$ signal is first computed in PL. Thus, the bulk photometry and simultaneous singe cell two-photon calcium and $GRAB_{DA}$ imaging agree that dopamine released into PL does not contain the $Rev_{85\text{-}15}$ signal nor explain signaling in the majority of $Rev_{85\text{-}15}$ cells. It is therefore extremely unlikely that the $Rev_{85\text{-}15}$ signal originates in VTA.

II. IS THE PRELIMBIC META-RPE SIGNAL PREFERENTIALLY SENT TO VTA?

Mathematically, the $Rev_{85\text{-}15}$ signal improves the fit of RPE-based learning models to behavior when allowed to dynamically modulate RPE signals. Neurobiologically, RPE signals are thought to arise from dopaminergic neurons in VTA (Schultz, 2016), and the $Rev_{85\text{-}15}$ signal is strongly represented in PL. It is therefore possible that $Rev_{85\text{-}15}$ neurons in PL could directly project to VTA dopamine neurons and thereby modulate RPE. PL represents the largest PFC input to VTA, with a distinct projection (Babiczky & Matyas, 2022; Gongwer et al., 2023), equally targeting VTA dopamine neurons and the GABAergic neurons that inhibit them (Carr & Sesack, 2000). Since the simplest mechanism is that PL $Rev_{85\text{-}15}$ neurons directly interface with VTA DA neurons, it is logical to center investigation towards the more rostral and lateral portion of the VTA, where the density of DA neurons and PFC terminals are highest (Simon et al., 2023). The purpose of this section is to test the hypothesis that the PL- VTA projection conveys the $Rev_{85\text{-}15}$ signal.

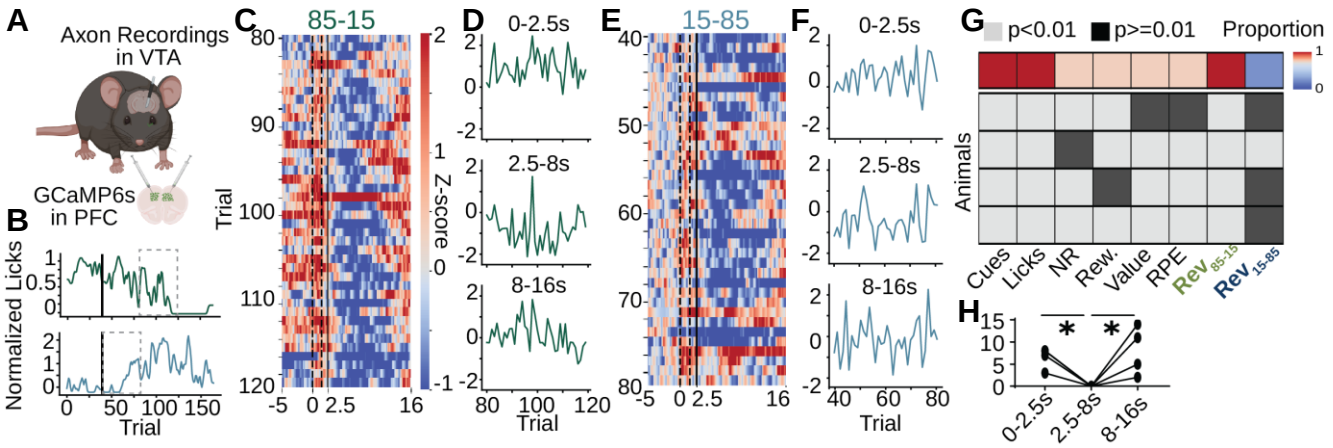

Figure 4-3: PFC Terminal Recordings in VTA Convey Reversal Signal. **A.** Schematic of experimental preparation. **B**. Behavioral acquisition curves for example animal, normalized as in Fig. 3-5, with reversal days marked via dotted lines for additional analysis. **C**. Trial-centered PFC-VTA terminal activity for day where the example animal behaviorally reversed the 85-15 cue. **D**. Trial-averaged PFC-VTA teriminal activity surrounding cue (top), reward (center) and post-reward (lower) periods. **E**. Same as C but for 15-85 cue. **F**. Same as D but for 15-85 cue. **G**. GLM results across animals. Every animal had significant encoding of cue, lick, and Rev85-15 signals. **H**. Although the cue and post-reward periods also display characteristic reversal shape, this activity peaks significantly later than the reward centered activity.

To test whether PL sends the $Rev_{85\text{-}15}$ signal to VTA, C57BL6J mice injected with GCaMP6s in PL and implanted with photometry fibers targeting rostro-lateral VTA performed the reversal task (**Fig. 4-3A**). Acqusition curves (**Fig. 4-3B**) enabled separate identification of the days where the majority of the 85-15 or 15-85 reversal occurred for additional analysis. Trial-centered activity maps surrounding the 85-15 cue (**Fig. 4-3C**) revealed a strong, transient inhbition that develops during the reward delivery and consumption (8-16s) phase of the task (**Fig. 4-3D**), and significantly encoded the $Rev_{85\text{-}15}$ signal in every animal (**Fig. 4-3G**). Conversely, there were increases both during the cue (0-2s) and post-ingestion (8-16s) period that peaked partway through the reversal period (**Fig. 4-3D**), though significantly later than the consumption inihbition (**Fig. 4-3H**). As trials were spaced 16-20s apart, the post-ingestion signal is also observable in the baseline before cue-delivery (**Fig. 4-3C**). In contrast, although 15-85 was also inhibited compared to baseline during consumption (**Fig. 4-3E**), it did not change in magnitude across the session (**Fig. 4-3F**) or reflect significant $Rev_{15\text{-}85}$ coding (**Fig. 4-3G**).

That every animal displayed significant encoding of the $Rev_{85\text{-}15}$ signal in PFC-VTA terminal activity strongly suggests that VTA receives the $Rev_{85\text{-}15}$ signal from PFC, perhaps from reversal cells. Performing projection-specific single cell two-photon calcium imaging of PFC-VTA cells directly tested whether they preferentially displayed significant encoding of the $Rev_{85\text{-}15}$ signal. In contrast to a pan-neuronal approach (**Fig. 3-2A**), injection of rAAV5-pgk-cre into rostro-lateral VTA, AAVDJ-EF1a-DIO-GCaMP6s in PL, and subsequent microprism implant sparsely restricted recording to only PFC-VTA projecting cells (**Fig. 4-4A**). Trial-averaged activity of PFC-VTA single units on the session with the largest behavioral change to the 85-15 cue (**Fig. 4-4B**) revealed a transient inhibition that deepened and then returned to baseline over a handful of trials only during the reward delivery and consumption phase of the task (**Fig. 4-4C**). When pooled together, most (65%) of the PFC-VTA neurons displayed significant encoding of the $Rev_{85\text{-}15}$ signal and none encoded the $Rev_{15\text{-}85}$ signal (**Fig. 4-4D**). This is a sizably larger proportion than the bulk

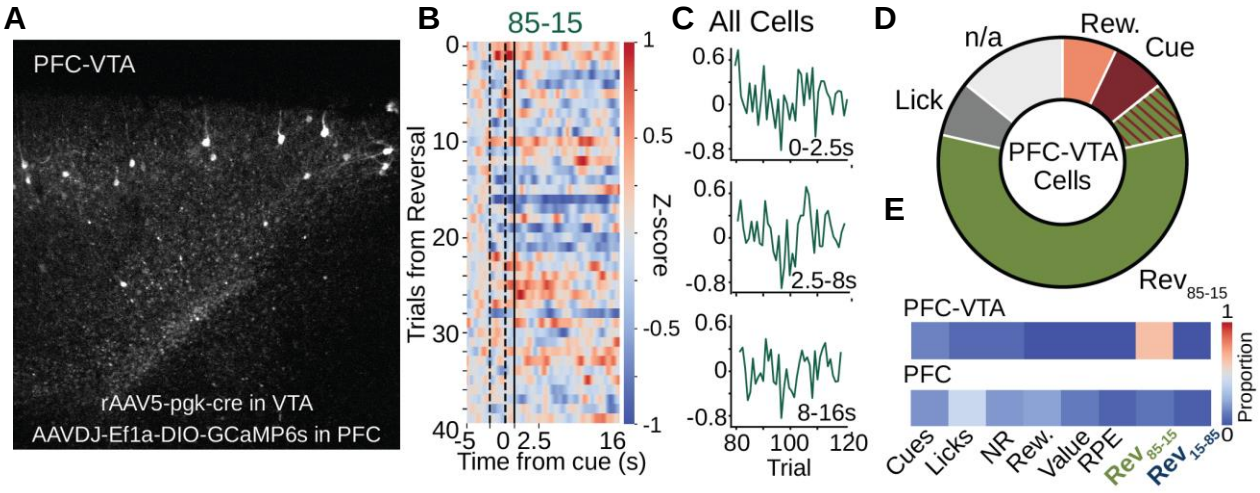

Figure 4-4: PFC Neurons Projecting to VTA Convey Reversal Signal. **A.** Representative field of view with viral strategy **B**. Trial-averaged activity for all cells in an example animal following contingency reversal. **C**. Trial-averaged PFC-VTA neuronal activity surrounding cue (top), reward (center) and post-reward (lower) periods. **D**. Distribution of GLM-derived significant encoding parameters for all evaluated cells. **E.** Comparison of the proportion of cells that encode each parameter in the PFC-VTA cohort compared to the main PFC cohort (See **Chapter 3**).

population (**Fig. 4-4E**). When coupled with fiber photometry results (**Fig. 4-3**), these single cell neuron results strongly support the conclusion that PL sends the mRPE $Rev_{85\text{-}15}$ signal to VTA via $Rev_{85\text{-}15}$ cells.

It is extremely encouraging that the $Rev_{85\text{-}15}$ cells observed generally within PL are greatly overrepresented in the population that projects to VTA. However, the directionality of the activity observed in the PFC-VTA cells (**Fig. 4-4B**) and many of the $Rev_{85\text{-}15}$ cells (**Fig. 3-4A**) is different. Many $Rev_{15\text{-}85}$ cells displayed strong increases paralleling the $Rev_{85\text{-}15}$ signal, while PFC-VTA cells displayed strong decreases. There are two potential explanations for this discrepancy. First, the non-restricted recordings in PL were performed with GCaMP6s under the CaMKIIα promoter. Although CaMKIIα is generally thought to have expression restricted to excitatory pyramidal neurons (Wang et al., 2013), there is a substantial amount of CaMKIIα virus expression leakage into inhibitory interneurons, up to 88% in the case of parvalbumin-containing interneurons in PFC (Veres et al., 2023). Since only pyramidal neurons project outside of PFC to VTA (Vázquez-Borsetti et al., 2009), it is possible that $Rev_{85\text{-}15}$ cells with excitations were interneurons, and therefore not represented in the PFC-VTA projection experiment. Since interneurons generally inhibit pyramidal neurons, interneuron increases would be expected to couple with pyramidal neuron decreases, as observed. Future studies will be needed to dissect the exact contribution of pyramidal cells and interneurons to the $Rev_{85\text{-}15}$ signal, but regardless of the exact microcircuitry that generates the signal, it reaches VTA from PFC, and decreases in activity during reversal.

Another possibility is that both excited and inhibited $Rev_{85\text{-}15}$ cells are pyramidal neurons, but with different projection targets. The rostro-lateral rAAV and fiber photometry implant coordinates are biased towards dopaminergic cells in VTA (Simon et al., 2023), and therefore likely skew towards representing PL cellular signals sent to VTA dopamine neurons. PL neurons also target VTA GABA neurons, which inhibit VTA dopamine neurons (Carr & Sesack, 2000), but are generally located more medially (Simon et al., 2023). It is possible the implant and injection

coordinates did not capture medial VTA-projecting PL neurons. Regardless, If the PL-VTA_{DA} projection decreases its activity during reversal, and the PL-VTA_{GABA} projection increases its activity, the net effect would be the same: decreased activity of VTA dopamine neurons.

The notion that PL-VTA activity reduces during reversal, possibly decreasing activity in dopaminergic neurons, is compatible with the results of the mathematical modeling in **Figure 3-4&5**. Putting the modeling results into words, allowing the $Rev_{85\text{-}15}$ signal to increase the impact of trial-by-trial RPE on choice behavior during the 85-15 contingency degradation improves predictions of the animal's choice behavior (**Fig. 3-4F**) and changes in anticipatory licking (**Fig. 3-5E**). Were this to be implemented physiologically, a $Rev_{85\text{-}15}$ signal would need to reach and modulate RPE signaling. This work has identified a $Rev_{85\text{-}15}$ signal in PL , and extensive previous work has identified VTA as the site of RPE signaling (for review see (Schultz, 2016)). This chapter has characterized the presence of the $Rev_{85\text{-}15}$ signal in the PL axons that reach VTA (**Fig. 4-3**), as well as in the PL pyramidal neurons that project to VTA (**Fig. 4-4**). Thus, the mathematically defined relationship between $Rev_{85\text{-}15}$ cells and RPE appears to have a direct physiological representation in the PL-VTA projection. Although mathematically multiplying a negative RPE by positive reversal signal produces an even more negative RPE, physiologically increasing excitatory input (a "positive" neural signal) onto inhibited neurons (a "negative" RPE) would likely negate the inhibition, instead of enhancing it. Enhancing an inhibition would require *decreases* in excitatory drive which could be physiologically multiplicative instead of summatory as VTA dopamine neurons display non-linear responses to excitatory input (Bonci & Malenka, 1999). Consequently, for excitatory pyramidal $Rev_{85\text{-}15}$ neurons to increase the magnitude of negative RPEs, they would need to *decrease* their activity, as observed in **Fig. 4-3G&4-4D**. Thus, the activity of the $Rev_{85\text{-}15}$ neurons within the PFC-VTA circuit represents the physiological implementation of the mRPEα model generated in **Chapter 3** of this work.

References

Alsiö, J., Lehmann, O., McKenzie, C., Theobald, D. E., Searle, L., Xia, J., Dalley, J. W., & Robbins, T. W. (2020). Serotonergic Innervations of the Orbitofrontal and Medial-prefrontal Cortices are Differentially Involved in Visual Discrimination and Reversal Learning in Rats. *Cerebral Cortex (New York, NY)*, *31*(2), 1090–1105. https://doi.org/10.1093/cercor/bhaa277

Arnsten, A. F. T. (1998). Catecholamine modulation of prefrontal cortical cognitive function. *Trends in Cognitive Sciences*, *2*(11), 436–447. https://doi.org/10.1016/S1364-6613(98)01240-6

Babiczky, Á., & Matyas, F. (2022). Molecular characteristics and laminar distribution of prefrontal neurons projecting to the mesolimbic system. *eLife*, *11*, e78813. https://doi.org/10.7554/eLife.78813

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F. F., & Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature*, *585*(7824), 245–250. https://doi.org/10.1038/s41586-020-2704-z

Bazzari, A. H., & Parri, H. R. (2019). Neuromodulators and Long-Term Synaptic Plasticity in Learning and Memory: A Steered-Glutamatergic Perspective. *Brain Sciences*, *9*(11), 300. https://doi.org/10.3390/brainsci9110300

Bonci, A., & Malenka, R. C. (1999). Properties and Plasticity of Excitatory Synapses on Dopaminergic and GABAergic Cells in the Ventral Tegmental Area. *The Journal of Neuroscience*, *19*(10), 3723–3730. https://doi.org/10.1523/JNEUROSCI.19-10-03723.1999

Carr, D. B., & Sesack, S. R. (2000). Projections from the Rat Prefrontal Cortex to the Ventral Tegmental Area: Target Specificity in the Synaptic Associations with Mesoaccumbens and Mesocortical Neurons. *Journal of Neuroscience*, *20*(10), 3864–3873. https://doi.org/10.1523/JNEUROSCI.20-10-03864.2000

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, *15*(4), 495–506. https://doi.org/10.1016/S0893-6080(02)00044-8

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. L. (2006). Multiple Dopamine Receptor Subtypes in the Medial Prefrontal Cortex of the Rat Regulate Set-Shifting. *Neuropsychopharmacology*, *31*(2), Article 2. https://doi.org/10.1038/sj.npp.1300825

Friedrich, J., Zhou, P., & Paninski, L. (2017). Fast online deconvolution of calcium imaging data. *PLoS Computational Biology*, *13*(3), e1005423. https://doi.org/10.1371/journal.pcbi.1005423

Gongwer, M. W., Klune, C. B., Couto, J., Jin, B., Enos, A. S., Chen, R., Friedmann, D., & DeNardo, L. A. (2023). Brain-Wide Projections and Differential Encoding of Prefrontal Neuronal Classes Underlying Learned and Innate Threat Avoidance. *The Journal of Neuroscience*, *43*(32), 5810–5830. https://doi.org/10.1523/JNEUROSCI.0697-23.2023

Grossman, C. D., Bari, B. A., & Cohen, J. Y. (2022). Serotonin neurons modulate learning rate through uncertainty. *Current Biology*, *32*(3), 586-599.e7. https://doi.org/10.1016/j.cub.2021.12.006

Hui, M., & Beier, K. T. (2022). Defining the interconnectivity of the medial prefrontal cortex and ventral midbrain. *Frontiers in Molecular Neuroscience*, *15*. https://doi.org/10.3389/fnmol.2022.971349

Jett, J. D., & Morilak, D. A. (2013). Too Much of a Good Thing: Blocking Noradrenergic Facilitation in Medial Prefrontal Cortex Prevents the Detrimental Effects of Chronic Stress on Cognition. *Neuropsychopharmacology*, *38*(4), 585–595. https://doi.org/10.1038/npp.2012.216

Müller, M. (2002). Microdialysis. *BMJ : British Medical Journal*, *324*(7337), 588–591.

Ogg, M. C., Franks, H. T., Lansdell, B. J., Hughes, A. C., Lee, J., Nolen, H. G., Shirinifard, A., & Schwarz, L. A. (2023). *Locus Coeruleus Norepinephrine Neurons Facilitate Orbitofrontal*

Cortex Remapping and Behavioral Flexibility (p. 2023.12.15.571858). bioRxiv. https://doi.org/10.1101/2023.12.15.571858

Puig, M. V., & Gulledge, A. T. (2011). Serotonin and Prefrontal Cortex Function: Neurons, Networks, and Circuits. *Molecular Neurobiology*, *44*(3), 449–464. https://doi.org/10.1007/s12035-011-8214-0

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

Simon, R. C., Loveless, M. C., Yee, J. X., Hashikawa, K., Stuber, G. D., Zweifel, L. S., & Soden, M. E. (2023). *Opto-seq reveals input-specific immediate early gene induction in ventral tegmental area cell types*. https://doi.org/10.1101/2023.06.22.546124

Stuber, G. D. (2023). Neurocircuits for motivation. *Science*, *382*(6669), 394–398. https://doi.org/10.1126/science.adh8287

van der Meulen, J. A. J., Joosten, R. N. J. M. A., de Bruin, J. P. C., & Feenstra, M. G. P. (2007). Dopamine and Noradrenaline Efflux in the Medial Prefrontal Cortex During Serial Reversals and Extinction of Instrumental Goal-Directed Behavior. *Cerebral Cortex*, *17*(6), 1444–1453. https://doi.org/10.1093/cercor/bhl057

Vázquez-Borsetti, P., Cortés, R., & Artigas, F. (2009). Pyramidal Neurons in Rat Prefrontal Cortex Projecting to Ventral Tegmental Area and Dorsal Raphe Nucleus Express 5-HT2A Receptors. *Cerebral Cortex*, *19*(7), 1678–1686. https://doi.org/10.1093/cercor/bhn204

Veres, J. M., Andrasi, T., Nagy-Pal, P., & Hajos, N. (2023). CaMKIIα Promoter-Controlled Circuit Manipulations Target Both Pyramidal Cells and Inhibitory Interneurons in Cortical Networks. *eNeuro*, *10*(4). https://doi.org/10.1523/ENEURO.0070-23.2023

Wang, X., Zhang, C., Szábo, G., & Sun, Q.-Q. (2013). Distribution of CaMKIIα expression in the brain in vivo, studied by CaMKIIα-GFP mice. *Brain Research*, *1518*, 9–25. https://doi.org/10.1016/j.brainres.2013.04.042

Weis, W. I., & Kobilka, B. K. (2018). The Molecular Basis of G Protein–Coupled Receptor Activation. *Annual Review of Biochemistry*, *87*, 897–919. https://doi.org/10.1146/annurev-biochem-060614-033910

Zhuo, Y., Luo, B., Yi, X., Dong, H., Wan, J., Cai, R., Williams, J. T., Qian, T., Campbell, M. G., Miao, X., Li, B., Wei, Y., Li, G., Wang, H., Zheng, Y., Watabe-Uchida, M., & Li, Y. (2023). Improved dual-color GRAB sensors for monitoring dopaminergic activity in vivo. *bioRxiv*, 2023.08.24.554559. https://doi.org/10.1101/2023.08.24.554559

SUMMARY AND BROADER IMPLICATIONS

Cognitive flexibility is important to make changes to existing learning when environmental contingencies change (**Introduction I**). This work has explored a simple form of cognitive flexibility, reversal learning, where the contingencies of two or more learned cues or behavioral strategies must be switched but all cues remain. Studies of prefrontal contributions to cognitive flexibility, including reversal learning have traditionally focused on orbitofrontal cortex (OFC). Other research has implicated OFC in detecting the initial changes in the environment following contingency reversal, as OFC activity is highest at reversal onset and lesions abolish the ability to perform cued reversals (**Introduction II**). However, OFC activity patterns decline too quickly to underlie behavioral remapping once an environmental change has been detected.

In contrast to OFC, lesions to the nearby prelimbic cortex (PL) also affect reversal behavior by prolonging the behavioral reversal time course, with impaired ability to learn from repeated errors (**Introduction III**). Neurally, PL population dynamics during reversal are more variable (**Introduction IV**) and also experience increased dopaminergic innervation (**Introduction V**). What precipitates changes in PL dynamics, their relationship to repeated errors, and the time-locked role of dopamine during contingency reversal was not well characterized in previous work. This work identified a distinct ensemble of PL neurons that activate specifically towards repeated negative reward prediction errors (RPEs) in close succession (**Chapter 3**), are not explained by dopaminergic input (**Chapter 4**), and project to the ventral tegmental area (VTA) (**Chapter 4**) using novel imaging and statistical methods (**Chapter 1**). Since PL population dynamics are relatively stable across reversal (**Chapter 2**), the activity of the reversal ensemble is unlikely to be attributable to representational drift. Allowing reversal cell activity to scale RPE improves the fit of trial-by-trial error models to behavior, because they account for non-linearities between RPE and behavioral change (**Chapter 3**). Reversal cells dominate the PL-VTA projection and represent

the physiological representation of the mathematical model (**Chapter 4**). Including the additional motivations for study, this work has addressed four major open questions in the field:

I. Why is learning during contingency reversal faster than stable learning rates predict?

II. Why might lesions of prelimbic cortex impact reversal learning and does this relate to previously detected firing changes during adaptation to contingency reversal?

III. How are dopamine increases in prelimbic cortex related to contingency reversal?

IV. What information does the PFC-VTA projection convey during cognitive flexibility?

The answers and implications for each of these questions are considered in more detail below.

I. WHY IS LEARNING DURING CONTINGENCY REVERSAL FASTER THAN STABLE LEARNING RATES PREDICT?

The idea of meta-learning emerged at the nexus of computer and neural science, and is often defined as "*the ability to learn how to learn, such as learning to adjust hyperparameters of existing learning algorithms*" (Langdon et al., 2022). Generally thought of as a process that expedites learning, meta-learning ideas have been applied to simple contingency degradation rodent tasks, like those in this work (Grossman et al., 2022; Namboodiri et al., 2021). This work and previous models unanimously agree that learning is faster following a change in environmental contingencies (**Fig. 3-4&3-5D**). However, previous models assume that learning happens fastest immediately following unexpected new information (i.e. at the contingency reversal), when RPEs are *largest* (Grossman et al., 2022; Namboodiri et al., 2021). As behavioral change is used as a correlate for learning in rodent models, if learning happens fastest immediately following contingency reversal, there should be an enhanced relationship between RPE and behavioral change during early reversal. Surprisingly, this does not appear to be the case as there is not a significant correlation between RPE and behavioral change during reversal (**Fig. 3-5B**) owing to the presence of the largest behavioral changes at *intermediate* RPEs (**Fig. 3-5C**). Thus, one would expect a meta-learning signal to peak at an *intermediate* point in the reversal, and not at its onset as previously argued.

Since many environments are inherently stochastic, it is logical to wait to commit to a large-scale behavioral change until evidence accumulates that the state of the environment has changed, instead of immediately changing behavior following a single unexpected datapoint. The meta-RPE (mRPE) signal serves such a function: increasing only when many negative errors have occurred in close succession, and almost all choices were incorrect (**Fig. 3-1**). That mRPE accounts for detected increases in learning rate (**Fig. 3-4E**) and improves fit to binary (**Fig. 3-4F**) and continuous (**Fig. 3-5E**) choice behavior strongly suggests it is meaningful. One way to conceptualize mRPE is as a state prediction error signal rising proportional to the quantity of previous trials where a behavioral strategy that was successful in a previous task state no longer yields the predicted outcome. Once a new strategy is adopted that better fits the new contingencies of the changed task state landscape, mRPE decreases. Such a signal qualifies as a meta-learning signal because it adjusts α (the Rescorla-Wagner learning rate parameter) but is better explained as a meta-RPE signal since it is both calculated from RPEs over a longer timescale and directly modifies the impact of trial-by-trial RPEs on behavioral change.

Altogether, the mRPE signal represents an error profile a level of abstraction above standard trial-by-trial RPE, better explains learning behavior than an RPE-only model, is highest when the largest behavioral changes occur, and produces behavioral estimates that strongly parallel those of real animals. Thus, mRPE explains the faster learning observed during contingency reversal.

II. WHY MIGHT LESIONS OF PRELIMBIC CORTEX IMPACT REVERSAL LEARNING AND DOES THIS RELATE TO PREVIOUSLY DETECTED FIRING CHANGES DURING ADAPTATION TO CONTINGENCY REVERSAL?

Behavioral data strongly suggests that PL lesions prolong the time course of reversal learning due to an impaired ability to learn from negative errors (**Introduction III**). This suggests that learning from negative errors may be an inherent property of a cellular ensemble within PL. Previous studies have observed changes in PL firing rates during reversal, but have not systematically explored their connection to learning, negative errors, or other components of reversal learning tasks (**Introduction IV**). A major goal of this study was to characterize the time-

locked activity of single cells in PL to reversal learning behavior. Identifying the meta-RPE (mRPE) signal that averages single-trial RPEs over a longer timescale, modulates RPE, and fits reversal behavior better than trial-by-trial RPE alone (**Fig. 3-4F&5E**) provided a candidate signal that could explain PL involvement in reversal learning. Without a mRPE signal, if behavior is predicted using a static learning rate comparable to pre-reversal behavior, the time course to stable post-reversal behavior is extended (**Fig. 3-5D**) as is observed in PL lesion animals (**Introduction III**). Microprisms allowed simultaneous recordings from thousands of PL cells across reversal (**Chapter 1-I**), and p-value leave-one-out generalized linear models (GLMs) allowed significance testing of the relationship between individual cells and the mRPE signal in addition to other task parameters (**Chapter 1-II**). Thus, advances in statistical and calcium imaging methodology allowed direct testing of whether PL units encoded mRPE.

Approximately 10% of the thousands of recorded neurons significantly encoded the mRPE signal only on the reversal day and only for the contingency degraded 85-15 cue (**Fig. 3-2B&C**). The activity of the reversal cells is strikingly elevated during the reversal period (**Fig. 3-4A**) and largely does not overlap with significant encoding of cue, value, or single-trial RPE signals (**Fig. 3-3B**). The emergence of the mRPE signal only during the reversal following repeated, concentrated errors and incorrect choices in a population of cells that had previously been fairly quiescent likely explains the firing of Bissonnette's "conflict" cells, Malagon-Vina's "bad behavioral performance" cells, Karlsson's increased "signaling volatility", Powell's "disorganization", and Rich's unit tracking "proportion incorrect" (Bissonette & Roesch, 2015; Karlsson et al., 2012; Malagon-Vina et al., 2018; Powell & Redish, 2016; Rich & Shapiro, 2009). Considered individually, increased reversal cell activity only during reversal would increase the *volatility* and *disorganization* of the overall PL population statistics that are otherwise fairly stable before and after the reversal (**Fig. 2-4A**). The reversal cells fire during the *conflict* between old and new contingency expectations because as animals choose to lick in anticipation to the 85-15 cue, subsequent reward delivery is unlikely (15%), in *conflict* with the pre-reversal 85% probability. As

animals make continued seeking decisions towards the devalued reward, they are making *incorrect* choices that yield *bad performance* in the task. Therefore, it is likely that the activity of the $Rev_{85\text{-}15}$ ensemble is responsible for previously observed PL firing changes during reversal.

Since the mRPE signal expedites learning following contingency degradation (**Summary I**), and the PL $Rev_{85\text{-}15}$ cells significantly encode the mRPE signal, it follows that the $Rev_{85\text{-}15}$ ensemble likely expedites learning following contingency degradation. Because the mRPE signal derives from repeated negative reward prediction errors, and PL lesion results in an impaired ability to learn from repeated negative errors, it is likely removal of the $Rev_{85\text{-}15}$ ensemble is responsible for this effect. The distinct activity profile between $Rev_{85\text{-}15}$ and OFC neurons (Banerjee et al., 2020), where $Rev_{85\text{-}15}$ cell activity peaks in the *middle* of reversal, while OFC activity peaks at the *start* of reversal, may likely explain the differential lesion effects where OFC animals never reverse while PL animals reverse more slowly. If OFC is important for detecting an initial change in the environment, and the change is never detected, behavior will never change. In contrast, if $Rev_{85\text{-}15}$ cells within PL are responsible for expediting behavioral change following reversal, their lesion will only slow progress.

Thus, PL lesions likely slow reversal learning from negative errors because they remove the $Rev_{85\text{-}15}$ ensemble, which expedites behavioral remapping to reversed cues following repeated negative errors. Increases in $Rev_{85\text{-}15}$ cell activity during contingency reversal can explain previously detected change in PL population firing statistics during contingency reversal.

III. HOW ARE PRELIMBIC DOPAMINE INCREASES RELATED REVERSAL?

A discussion of learning is incomplete without a discussion of dopamine. Tightly linked to learning via iterative errors, dopaminergic cells within the VTA are responsible for reward prediction error (RPE) signals across the brain (Schultz, 2016). Since VTA shares reciprocal connectivity with PL (Gongwer et al., 2023; Hui & Beier, 2022), RPE signals that reach PL likely originate from VTA. Since mRPE signals are computed from RPE, it is possible that they also arrive in PL via dopaminergic innervation. Adding support for this hypothesis, dopamine levels are selectively elevated in PL during contingency reversal (van der Meulen et al., 2007), the same period in which $Rev_{85\text{-}15}$ cells are active and the mRPE signal is highest. Before this work, the exact sub-second timescale of dopamine release in PL during reversal, and its impact on neurons and behavior was unexplored.

Surprisingly, the PL dopamine increase was not directly related to the contingency degradation or $Rev_{85\text{-}15}$ mRPE signal. Instead, the increases in activity were attributable to the contingency elevation (**Fig. 4-1B**) in both single-trial RPEs and multi-trial mRPEs (**Fig. 4-1G**). This strongly suggested dopamine release in PL did not drive $Rev_{85\text{-}15}$ cell activity. Simultaneous recordings of PL neurons and dopamine released into the environment confirmed this, as GLM analysis revealed only significant dopamine modulations in the majority of RPE and $Rev_{15\text{-}85}$ mRPE cells—not $Rev_{85\text{-}15}$ mRPE cells (**Fig. 4-2C**). Thus, dopamine release in PL is not directly related to the contingency degradation portion of reversal behavior. However, since the $Rev_{85\text{-}15}$ cellular signaling is compatible with mRPE, a signal that is computed by adding RPEs over many trials, may still be less directly linked to dopaminergic signaling because single-trial RPEs are significantly encoded in dopamine released into PL. That dopamine receptor antagonists affect reversal similarly to prelimbic lesions (**Introduction V**) suggests this may be the case. The exact mechanism by which $Rev_{85\text{-}15}$ cells compute the mRPE signal and whether dopaminergic innervation is required for this function remains an open question for future work.

That the dopamine signal is strongly linked with the contingency elevation phase of the reversal is compatible with its documented role conveying single-trial RPE. Given the pre-reversal task strategy, where animals do not expect reward deliveries (only 15% likelihood) following the 15-85 cue, when the reward contingency shifts to reward deliveries most of the time (85%), the resulting RPEs from reward deliveries could increase the amount of dopamine independent of any special role in reversal learning. However, dopamine does appear to have a special role in contingency elevation as the signal itself significantly encoded the 15-85 mRPE signal (**Fig. 4-1G**) and predicted the activity of the handful of $Rev_{15\text{-}85}$ cells (**Fig. 4-2C**). Given the paucity of 15-85 compared to 85-15 reversal cells (**Fig. 3-2B**), it is unlikely that this signal originates in PL.

Since PL selectively encodes the 85-15 mRPE signal, and dopamine selectively encodes the 15-85 mRPE signal, these signals are dissociable and likely distinct. This suggests that contingency degradation (85-15) and elevation (15-85) are separate processes that co-occur in tasks with a shared reversal point. This is understandable, as contingency degradation involves learning that a cue that used to meaningfully predict reward does not anymore, while contingency elevation involves learning that something that used to have a low association with reward, now meaningfully predicts it. The definition of contingency elevation (15-85) is very similar to normal acquisition (0-85) as both go from a low cue-reward predictive relationship to a high one and may therefore share a neural mechanism. This interpretation suggests that the cognitive flexibility portion of this reversal learning task is specifically the 85-15 contingency degradation signal (making changes to existing learning), which does not come from VTA dopamine neurons but is encoded in PL ensembles. Since contingency elevation and standard acquisition likely share a common mechanism, and PL DA encodes mRPE for contingency elevation, this suggests that mRPE signals may not be restricted to reversal learning.

The source of the 15-85 mRPE signal in VTA can only be speculated upon. Like PL, the basolateral amygdala (BLA) is reciprocally connected to VTA (An et al., 2021), contains representations of cue-reward associations that develop across learning, and BLA lesions only

effect reversal learning strategies involving transitions from cue-reward associations (Izquierdo et al., 2017). Prior interpretations of the BLA literature suggest its role involves high-level representation of enhanced cue-reward relationships (Wassum & Izquierdo, 2015), which is consistent with a contingency elevation function. That BLA signals reflect reward history over time (Morrison & Salzman, 2010) and activity perturbations during learning persist over many trials (Wassum & Izquierdo, 2015) suggests a time-scale compatibility with mRPE. Future work can identify whether BLA is indeed the source of the 15-85 mRPE signal reflected in the dopamine.

In conclusion, dopamine release in PL significantly encodes the contingency elevation during reversal, and do not have a direct significant relationship with contingency degradation behavior or PL ensembles. In addition to reflecting an increase in positive RPEs during contingency elevation, dopamine also significantly reflects a contingency elevation mRPE signal.

IV. WHAT INFORMATION DOES THE PFC-VTA PROJECTION CONVEY DURING COGNITIVE FLEXIBILITY?

The task-related function of the relatively weak prefrontal projection to VTA has been an open question for some time (Carr & Sesack, 2000; Gao et al., 2007; Gongwer et al., 2023). Substance use disorder (SUD) researchers have long hypothesized that decreased prefrontal innervation of subcortical circuitry is related to an impaired ability to devalue drug rewards during recovery but the mechanistic understanding of how PFC relates to this cognitive inflexibility has been understudied and the PFC-VTA projection largely ignored (Coley et al., 2021; Limpens et al., 2015; Volkow et al., 2019; Volkow & Morales, 2015). When mentioned at all, the PFC-VTA projection is postulated to directly perform a high-order flexibility-related modulation of the dopamine system, but this had yet to be tested directly (Koob & Volkow, 2016; Menon & D'Esposito, 2022). Although this study did not directly test the relationship of PFC-VTA activity during drug cue contingency degradation, it still characterized the activity of PFC-VTA neurons and axons during contingency degradation between salient odor cues and a natural sucrose

reinforcer. Understanding the basic biology that may go awry during SUD is an important first step to support future interventions of neural circuitry to promote SUD recovery.

As many hypothesized, this study established that the prelimbic subregion of PFC contains a discrete cognitive flexibility signal that represents repeated errors following contingency degradation (**Chapter 3**), termed meta-RPE (mRPE). Mathematical modeling suggests mRPE increases the impact of intermediate RPEs partway through behavioral remapping to the reversed cues, expediting contingency degradation (**Chapter 3**). Physiologically, a subset of PL neurons significantly encode the mRPE signal, RPE is tightly linked to the activity of VTA (Schultz, 2016), and PL represents the largest prefrontal connectivity to VTA (Carr & Sesack, 2000; Gongwer et al., 2023). Therefore, it is possible that mRPE's modeled-impact on RPE is physiologically implemented in the activity of the PL-VTA circuit.

Recordings of PL-VTA circuitry revealed that the mRPE signal reaches VTA in PL axons (**Fig. 4-3G**) and significantly predicts activity of a much larger proportion (65%) of PL-VTA neurons than the PL population (10%) overall (**Fig. 4-4E**). Thus, the mathematical modeling and neural connectivity are in alignment: the mRPE signal in PL is able to reach the RPE signal within VTA, likely modulating the magnitude of negative RPEs via an inhibition of excitatory tone (**Chapter 4**). Altogether, this suggests that the mRPE signal in PL that is likely responsible for behavioral lesion impairments during contingency reversal (**Introduction III**) due to an inability to modulate RPE signaling in VTA via the PL-VTA projection primarily composed of mRPE cells.

It remains an open question whether the activity of $Rev_{85\text{-}15}$ cells is constrained entirely to the PL-VTA projection population. Another recent work exploring PL contributions to cognitive flexibility (but that did not derive any such signal) reported increased activity during and immediately following expected reward delivery only during intra- or extra-dimensional shifts (Spellman et al., 2021). It is possible that mRPE cells in PL drove this change in activity, since they show similar activity both within PL (**Chapter 3**) and the PL-VTA population (**Chapter 4**). Spellman's work did not characterize the timescale of the changed PL activity but did explore

whether it was preferentially represented in PL projections to the thalamus (MDT) or striatum (VMS). The flexibility-related activity change was not preferentially restricted within PL-MDT or PL-VMS populations, but did show a concentration in deep layer V/VI neurons (Spellman et al., 2021). Notably, this is also the area that contains the largest concentration of PL-VTA neurons, over projections to thalamus or striatum (Gongwer et al., 2023). Thus, there is literature support for the notion that mRPE signaling is restricted to the PL-VTA projection, but future studies can directly test whether this is the case.

Altogether, PL $Rev_{85\text{-}15}$ neurons strongly convey the mRPE cognitive flexibility signal to VTA and represents the physiological implementation of modeled mRPE effects on RPE.

FINAL SUMMARY

This work presents meta-RPE for the first time, a cognitive flexibility signal that peaks in the middle of contingency degradation behavior and represents the average of repeated negative errors concentrated over many trials. Allowing meta-RPE to modulate canonical single-trial RPEs expedites reversal learning compared to models incorporating static pre-reversal learning rates, and fits observed animal behavior better than models that immediately increase RPE gain concurrent with contingency reversal. Physiologically, meta-RPE is represented within a subpopulation of neurons in PL and can directly modify RPE via preferential representation in projections to VTA. The mRPE signal within PL is unlikely attributable to representation drift, since population dynamics are consistent on stable behavior days across reversal. Likewise, dopaminergic innervation to PL does not predict the contingency degradation mRPE, instead representing a contingency elevation mRPE signal from elsewhere in the brain. Altogether, PL contributions to reversal behavior are likely explained via contingency degradation mRPE encoding, and conveyance of this signal to modulate downstream RPEs in VTA.

References

An, S., Li, X., Deng, L., Zhao, P., Ding, Z., Han, Y., Luo, Y., Liu, X., Li, A., Luo, Q., Feng, Z., & Gong, H. (2021). A Whole-Brain Connectivity Map of VTA and SNc Glutamatergic and GABAergic Neurons in Mice. *Frontiers in Neuroanatomy*, *15*. https://doi.org/10.3389/fnana.2021.818242

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F. F., & Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature*, *585*(7824), 245–250. https://doi.org/10.1038/s41586-020-2704-z

Bissonette, G. B., & Roesch, M. R. (2015). Neural correlates of rules and conflict in medial prefrontal cortex during decision and feedback epochs. *Frontiers in Behavioral Neuroscience*, *9*, 266. https://doi.org/10.3389/fnbeh.2015.00266

Carr, D. B., & Sesack, S. R. (2000). Projections from the Rat Prefrontal Cortex to the Ventral Tegmental Area: Target Specificity in the Synaptic Associations with Mesoaccumbens and Mesocortical Neurons. *Journal of Neuroscience*, *20*(10), 3864–3873. https://doi.org/10.1523/JNEUROSCI.20-10-03864.2000

Coley, A. A., Padilla-Coreano, N., Patel, R., & Tye, K. M. (2021). Valence processing in the PFC: Reconciling circuit-level and systems-level views. In *International Review of Neurobiology* (Vol. 158, pp. 171–212). Elsevier. https://doi.org/10.1016/bs.irn.2020.12.002

Gao, M., Liu, C.-L., Yang, S., Jin, G.-Z., Bunney, B. S., & Shi, W.-X. (2007). Functional Coupling between the Prefrontal Cortex and Dopamine Neurons in the Ventral Tegmental Area. *The Journal of Neuroscience*, *27*(20), 5414–5421. https://doi.org/10.1523/JNEUROSCI.5347-06.2007

Gongwer, M. W., Klune, C. B., Couto, J., Jin, B., Enos, A. S., Chen, R., Friedmann, D., & DeNardo, L. A. (2023). Brain-Wide Projections and Differential Encoding of Prefrontal

Neuronal Classes Underlying Learned and Innate Threat Avoidance. *The Journal of Neuroscience*, *43*(32), 5810–5830. https://doi.org/10.1523/JNEUROSCI.0697-23.2023

Grossman, C. D., Bari, B. A., & Cohen, J. Y. (2022). Serotonin neurons modulate learning rate through uncertainty. *Current Biology*, *32*(3), 586-599.e7. https://doi.org/10.1016/j.cub.2021.12.006

Hui, M., & Beier, K. T. (2022). Defining the interconnectivity of the medial prefrontal cortex and ventral midbrain. *Frontiers in Molecular Neuroscience*, *15*. https://doi.org/10.3389/fnmol.2022.971349

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, *345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Karlsson, M. P., Tervo, D. G. R., & Karpova, A. Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science (New York, N.Y.)*, *338*(6103), 135–139. https://doi.org/10.1126/science.1226518

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: A neurocircuitry analysis. *The Lancet. Psychiatry*, *3*(8), 760–773. https://doi.org/10.1016/S2215-0366(16)00104-8

Langdon, A., Botvinick, M., Nakahara, H., Tanaka, K., Matsumoto, M., & Kanai, R. (2022). Meta-learning, social cognition and consciousness in brains and machines. *Neural Networks*, *145*, 80–89. https://doi.org/10.1016/j.neunet.2021.10.004

Limpens, J. H. W., Damsteegt, R., Broekhoven, M. H., Voorn, P., & Vanderschuren, L. J. M. J. (2015). Pharmacological inactivation of the prelimbic cortex emulates compulsive reward seeking in rats. *Brain Research*, *1628*, 210–218. https://doi.org/10.1016/j.brainres.2014.10.045

Malagon-Vina, H., Ciocchi, S., Passecker, J., Dorffner, G., & Klausberger, T. (2018). Fluid network dynamics in the prefrontal cortex during multiple strategy switching. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-017-02764-x

Menon, V., & D'Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 90–103. https://doi.org/10.1038/s41386-021-01152-w

Morrison, S. E., & Salzman, C. D. (2010). Re-valuing the amygdala. *Current Opinion in Neurobiology*, *20*(2), 221–230. https://doi.org/10.1016/j.conb.2010.02.007

Namboodiri, V. M. K., Hobbs, T., Trujillo-Pisanty, I., Simon, R. C., Gray, M. M., & Stuber, G. D. (2021). Relative salience signaling within a thalamo-orbitofrontal circuit governs learning rate. *Current Biology*, *31*(23), 5176-5191.e5. https://doi.org/10.1016/j.cub.2021.09.037

Powell, N. J., & Redish, A. D. (2016). Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms12830

Rich, E. L., & Shapiro, M. (2009). Rat Prefrontal Cortical Neurons Selectively Code Strategy Switches. *Journal of Neuroscience*, *29*(22), 7208–7219. https://doi.org/10.1523/JNEUROSCI.6068-08.2009

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*(1), 23–32.

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., & Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell*, *184*(10), 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047

van der Meulen, J. A. J., Joosten, R. N. J. M. A., de Bruin, J. P. C., & Feenstra, M. G. P. (2007). Dopamine and Noradrenaline Efflux in the Medial Prefrontal Cortex During Serial Reversals and Extinction of Instrumental Goal-Directed Behavior. *Cerebral Cortex*, *17*(6), 1444–1453. https://doi.org/10.1093/cercor/bhl057

Volkow, N. D., Michaelides, M., & Baler, R. (2019). The Neuroscience of Drug Reward and Addiction. *Physiological Reviews*, *99*(4), 2115–2140. https://doi.org/10.1152/physrev.00014.2018

Volkow, N. D., & Morales, M. (2015). The Brain on Drugs: From Reward to Addiction. *Cell*, *162*(4), 712–725. https://doi.org/10.1016/j.cell.2015.07.046

Wassum, K. M., & Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neuroscience and Biobehavioral Reviews*, *57*, 271–283. https://doi.org/10.1016/j.neubiorev.2015.08.017

APPENDIX A: A reprint of Ottenheimer*, Hjort*, Bowen* et al. 2023
Available at: https://elifesciences.org/articles/84604

RESEARCH ARTICLE

A stable, distributed code for cue value in mouse cortex during reward learning

David J Ottenheimer[1,2,3†], Madelyn M Hjort[1,2†], Anna J Bowen[3†], Nicholas A Steinmetz[1,3*‡], Garret D Stuber[1,2,4*‡]

[1]Center for the Neurobiology of Addiction, Pain and Emotion, University of Washington, Seattle, United States; [2]Anesthesiology and Pain Medicine, University of Washington, Seattle, United States; [3]Department of Biological Structure, University of Washington, Seattle, United States; [4]Department of Pharmacology, University of Washington, Seattle, United States

Abstract The ability to associate reward-predicting stimuli with adaptive behavior is frequently attributed to the prefrontal cortex, but the stimulus-specificity, spatial distribution, and stability of prefrontal cue-reward associations are unresolved. We trained head-fixed mice on an olfactory Pavlovian conditioning task and measured the coding properties of individual neurons across space (prefrontal, olfactory, and motor cortices) and time (multiple days). Neurons encoding cues or licks were most common in the olfactory and motor cortex, respectively. By quantifying the responses of cue-encoding neurons to six cues with varying probabilities of reward, we unexpectedly found value coding in all regions we sampled, with some enrichment in the prefrontal cortex. We further found that prefrontal cue and lick codes were preserved across days. Our results demonstrate that individual prefrontal neurons stably encode components of cue-reward learning within a larger spatial gradient of coding properties.

eLife assessment

This study makes **valuable** observations about the representation of "value" in the mouse brain, by using a nice task design and recording from an impressive number of brain regions. The combination of state-of-the-art imaging and electrophysiology data offer **solid** support for the authors' conclusions. The paper will be of interest to a broad audience of neuroscientists interested in reward processing in the brain.

***For correspondence:** nick.steinmetz@gmail.com (NAS); gstuber@uw.edu (GDS)

[†]These authors contributed equally to this work
[‡]These authors also contributed equally to this work

Competing interest: The authors declare that no competing interests exist.

Funding: See page 25

Sent for Review 10 November 2022
Preprint posted 11 November 2022
Reviewed preprint posted 24 February 2023
Reviewed preprint revised 07 June 2023
Version of Record published 30 June 2023

Reviewing Editor: Michael B Eisen, University of California, Berkeley, United States

Introduction

Association of environmental stimuli with rewards and the subsequent orchestration of value-guided reward-seeking behavior are crucial functions of the nervous system linked to the prefrontal cortex (PFC) (***Miller and Cohen, 2001***; ***Klein-Flügge et al., 2022***). PFC is heterogeneous, with many studies noting subregional differences in both neural coding ***Kennerley et al., 2009***; ***Sul et al., 2010***; ***Hunt et al., 2018***; ***Wang et al., 2020a*** and functional impact on ***Dalley et al., 2004***; ***Rudebeck et al., 2008***; ***Buckley et al., 2009***; ***Kesner and Churchwell, 2011*** value-based reward seeking in primates and rodents. Furthermore, functional manipulations of PFC subregions exhibiting robust value signals do not always cause a discernible impact on reward-guided behavior (***Chudasama and Robbins, 2003***; ***St Onge and Floresco, 2010***; ***Dalton et al., 2016***; ***Verharen et al., 2020***; ***Wang et al., 2020a***), encouraging investigation of differences between value signals across PFC. Within individual PFC subregions, multiple studies have observed evolving neural representations across time, calling into question the stability of PFC signaling (***Hyman et al., 2012***; ***Malagon-Vina et al., 2018***). A systematic

comparison of coding properties across rodent PFC and related motor and sensory regions, as well as across days and stimulus sets, is necessary to provide a full context for the contributions of PFC subregions to reward processing.

Identifying neural signals for value requires a number of considerations. One issue is that other task features can vary either meaningfully or spuriously with value. In particular, action coding is difficult to parse from value signaling, given the high correlations between behavior and task events (***Musall et al., 2019***; ***Zagha et al., 2022***) and widespread neural coding of reward-seeking actions (***Steinmetz et al., 2019***). Additionally, without a sufficiently rich value axis, it is possible to misidentify neurons as 'value' coding even though they do not generalize to valuations in other contexts (***Stalnaker et al., 2015***; ***Hayden and Niv, 2021***; ***Zhou et al., 2021***). Because reports of the value have come from different experiments across different species, it is difficult to compare the presence of value signaling even across regions within the prefrontal cortex (***Kennerley et al., 2009***; ***Sul et al., 2010***; ***Stalnaker et al., 2015***; ***Otis et al., 2017***; ***Hunt et al., 2018***; ***Namboodiri et al., 2019***; ***Wang et al., 2020a***; ***Hayden and Niv, 2021***; ***Zhou et al., 2021***).

In this work, we sought to address the existing ambiguity in the distribution and stability of value signaling. We implemented an olfactory Pavlovian conditioning task that permitted the identification of value correlates within the domain of reward probability across two separate stimulus sets. With acute in vivo electrophysiology recordings, we were able to assess the coding of this task across 11 brain regions, including five PFC subregions, as well as olfactory and motor cortex, in a single group of mice, permitting a well-controlled comparison of coding patterns across a large group of the task-relevant regions in the same subjects. Unexpectedly, in contrast to the graded cue and lick coding across these regions, the proportion of neurons encoding cue value was more consistent across regions, with a slight enrichment in PFC but with similar value decoding performance across all regions. To assess coding stability, we performed 2-photon calcium imaging of neurons in the PFC for multiple days and determined that the cue and lick codes we identified were stable over time. Our data demonstrate the universality and stability of cue-reward coding in the mouse cortex.

Results

Distributed neural activity during an olfactory Pavlovian conditioning task

We trained mice on an olfactory Pavlovian conditioning task with three cue (conditioned stimulus) types that predicted reward on 100% ('CS+'), 50% ('CS50'), or 0% ('CS−') of trials (***Figure 1A***). Each mouse learned two odor sets (odor sets A and B), trained and imaged on separate days and then, for electrophysiology experiments, presented in six alternating blocks of 51 trials during the recording sessions (***Figure 1B***). Mice developed anticipatory licking (***Figure 1C–D***), and the rate of this licking correlated with reward probability (***Figure 1—figure supplement 1***), indicating that subjects successfully learned the meaning of all six odors.

Using Neuropixels 1.0 and 2.0 probes (***Jun et al., 2017***; ***Steinmetz et al., 2021***), we recorded the activity of individual neurons in PFC, including anterior cingulate area (ACA), frontal pole (FRP), prelimbic area (PL), infralimbic area (ILA), and orbital area (ORB) (***Wang et al., 2020b***; ***Laubach et al., 2018***). We also recorded from: secondary motor cortex (MOs), including anterolateral motor cotex (ALM), which has a well-characterized role in licking ***Chen et al., 2017***; olfactory cortex (OLF), including dorsal peduncular area (DP), dorsal taenia tecta (TTd), and anterior olfactory nucleus (AON), which receive input from the olfactory bulb (***Igarashi et al., 2012***; ***Mori and Sakano, 2021***); and striatum, including caudoputamen (CP) and nucleus accumbens (ACB), which are major outputs of PFC (***Heilbronner et al., 2016***, ***Figure 1E–F***). In a separate group of mice, we performed longitudinal 2-photon calcium imaging through a Gradient Refractive Index (GRIN) lens to track the activity of individual neurons in PL across several days of behavioral training (***Figure 1G–H***). Both techniques permitted robust measurement of the activity of neurons of interest and generated complementary results (***Figure 1I–J***, ***Figure 1—figure supplement 2***).

Graded cue and lick coding across the recorded regions

In the electrophysiology experiment, we isolated the spiking activity of 5332 individual neurons in regions of interest across 5 mice (449-1550 neurons per mouse, ***Figure 2A***, ***Figure 2—figure***

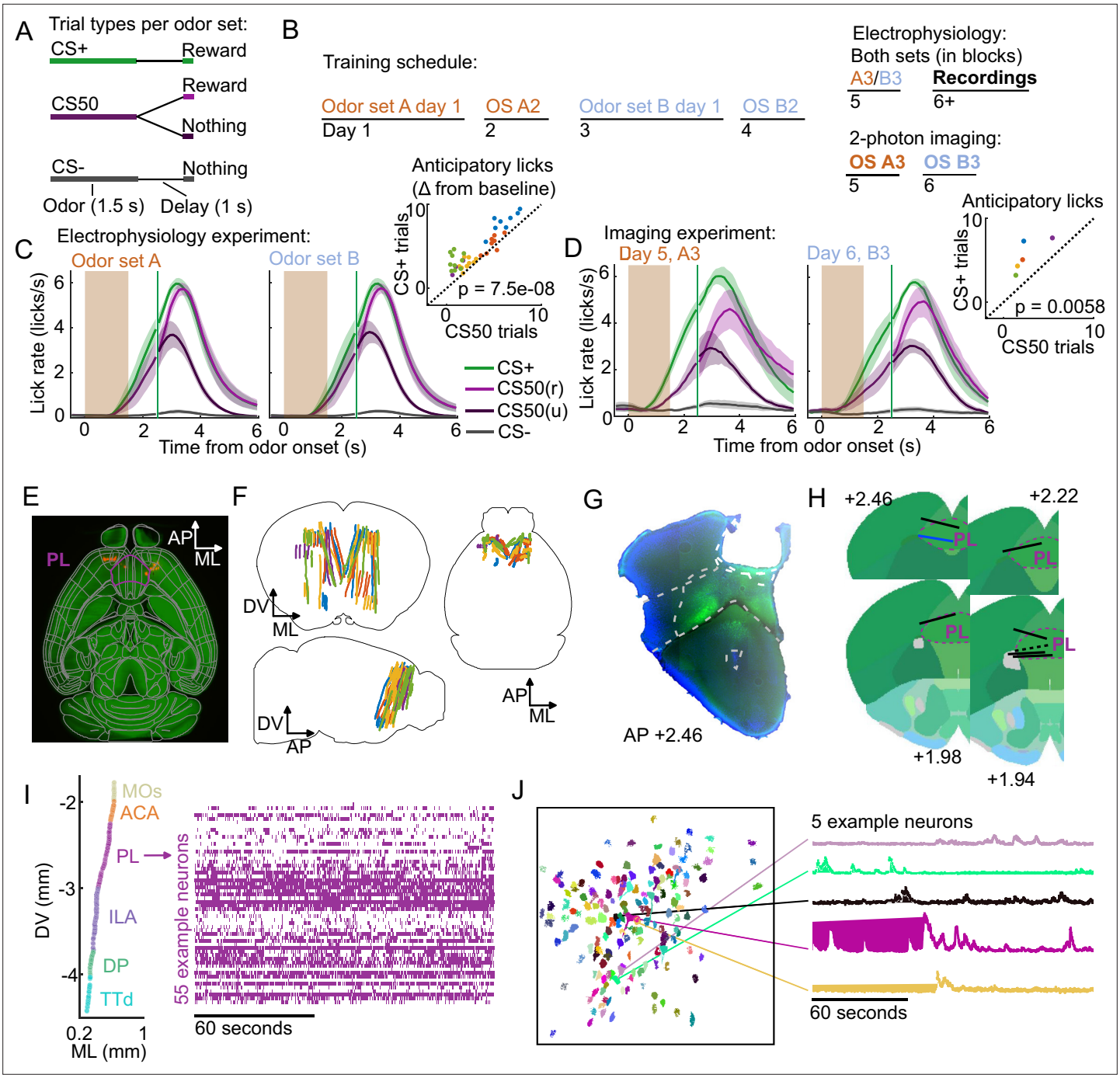

Figure 1. Electrophysiology and calcium imaging during olfactory Pavlovian conditioning. (**A**) Trial structure in Pavlovian conditioning task. (**B**) Timeline for mouse training. (**C**) Mean (+/− standard error of the mean (SEM)) lick rate across mice ($n = 5$) on each trial type for each odor set during electrophysiology sessions. CS50(r) and CS50(u) are rewarded and unrewarded trials, respectively. Inset: mean anticipatory licks (change from baseline) for the CS+ and CS50 cues for every session, color-coded by mouse. $F(1, 66) = 36.6$ for a main effect of cue in a two-way ANOVA including an effect of subject. (**D**) Same as (**C**), for the third session of each odor set ($n = 5$ mice). $t(4) = -5.4$ for a t-test comparing anticipatory licks on CS+ and CS50 trials. (**E**) Neuropixels probe tracks labeled with fluorescent dye (red) in cleared brain (autofluorescence, green). AP, anterior/posterior; ML, medial/lateral; DV, dorsal/ventral. Allen common-coordinate framework (CCF) regions delineated in gray. Outline of prelimbic area in purple (**F**) Reconstructed recording sites from all tracked probe insertions ($n = 44$ insertions, $n = 5$ mice), colored by mouse. (**G**) Sample histology image of lens placement. Visualization includes DAPI (blue) and GCaMP (green) signal with lines indicating cortical regions from Allen Mouse Brain Common Coordinate Framework. (**H**) Location of all lenses from experimental animals registered to Allen Mouse Brain Common Coordinate Framework. Blue line indicates location of lens in (**A**). The dotted black line represents approximate location of tissue that was too damaged to reconstruct an accurate lens track. The white

Figure 1 continued on next page

Figure 1 continued

dotted line indicates prelimbic area (PL) borders.(**I**) ML and DV coordinates of all neurons recorded in one example session, colored by region, and spike raster from example PL neurons. (**J**) ROI masks for identified neurons and fluorescence traces from five example neurons.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Anticipatory licking during the electrophysiology sessions.

Figure supplement 2. Similar neural activity in prelimbic area using electrophysiology and calcium imaging.

supplement 1A). The activity of neurons in all regions exhibited varying degrees of modulation in response to the six cues (***Figure 2B***). Broadly, there was strong modulation on CS+ and CS50 trials that appeared to be common to both odor sets (***Figure 2—figure supplement 1B***). Across regions, there was heterogeneity in both the magnitude and the timing of the neural modulation relative to odor onset (***Figure 2—figure supplement 1C***).

To quantify the relative contribution of cues and conditioned responding (licking) to the activity of neurons in each region, we implemented reduced rank kernel regression (***Steinmetz et al., 2019***), using cues, licks, and rewards to predict neurons' activity on held-out trials (***Figure 2C***, ***Figure 2—figure supplement 2A***). To determine the contribution of cues, licks, and rewards to each neuron's activity, we calculated unique variance explained by individually removing each predictor from the model and calculating the reduction in model performance (***Figure 2D***).

We identified individual neurons encoding cues, licks, or rewards as those for which that predictor uniquely contributed to 2% or more of their variance (a cutoff permitting no false positives and identifying neurons with robust task modulation, see Methods and ***Figure 2—figure supplement 3***). Neurons encoding cues (24% of all neurons), licks (11%), or both (16%) were most common. Neurons with any response to reward (independent of licking) were rare (5%) (***Horst and Laubach, 2013***). Cue neurons were characterized by sharp responses aligned to odor onset; in contrast, lick neurons' responses were delayed and peaked around reward delivery (***Figure 2—figure supplement 2B–C***), consistent with the timing of licks (***Figure 1C***). The activity of cue neurons on rewarded and unrewarded CS50 trials validated our successful isolation of neurons with cue but not lick responses (***Figure 2—figure supplement 2D***). The spatial distributions of cue and lick cells were noticeably different (***Figure 2E***). The differences could be described as graded across regions, with the most lick neurons in ALM, and the most cue neurons in olfactory cortex and ORB, though each type of neuron was observed in every region (***Figure 2F–G***, ***Figure 2—figure supplement 4***). Thus, our quantification of task encoding revealed varying proportions of cue and lick signaling across all regions.

Cue value coding is present in all regions

To expand upon our analysis identifying cue-responsive neurons, we next assessed the presence of cue value coding in this population. The three cue types (CS+, CS50, or CS−) in our behavioral tasks varied in relative value according to the predicted probability of reward (***Fiorillo et al., 2003***; ***Eshel et al., 2016***; ***Winkelmeier et al., 2022***). We reasoned that a neuron encoding cue value should have activities that scaled with the relative value of the cues (***Figure 3A***). We modeled this relationship on a per-neuron basis by scaling a single cue kernel by its reward probability (0, 0.5, or 1, see Methods, ***Figure 3B***). This model describes cue activity as similar across odors of the same value, and scaling in magnitude according to each odor's value. To consider alternative cue coding patterns, we also fit each neuron with 152 additional models containing all possible permutations of these values across the six cues, as well as models with selective responses for 1, 2, 3, 4, 5, or 6 cues, and determined which model best fit each neuron (***Figure 3—figure supplement 1***). If cue responses were exclusively sensory and followed known olfactory coding properties (***Stettler and Axel, 2009***; ***Pashkovski et al., 2020***), there would be no bias toward the ranked value model (CS+>CS50>CS−). We found, however, that this model was the most frequent best model, accounting for 14% of cue neurons (***Figure 3C***). We refer to these neurons as value cells. There were two additional patterns that emerged across the population of cue neurons. First, there was a large fraction best explained by the model with equivalent responses to all 6 cues, which we term untuned cells (14% of cue neurons). Second, many of the alternative models had coding patterns that were similar to the ranked value model, and these appeared to be overrepresented among cue neurons, as well. We quantified the similarity to ranked value by correlating the values assigned to each cue in each model with those assigned to the cues

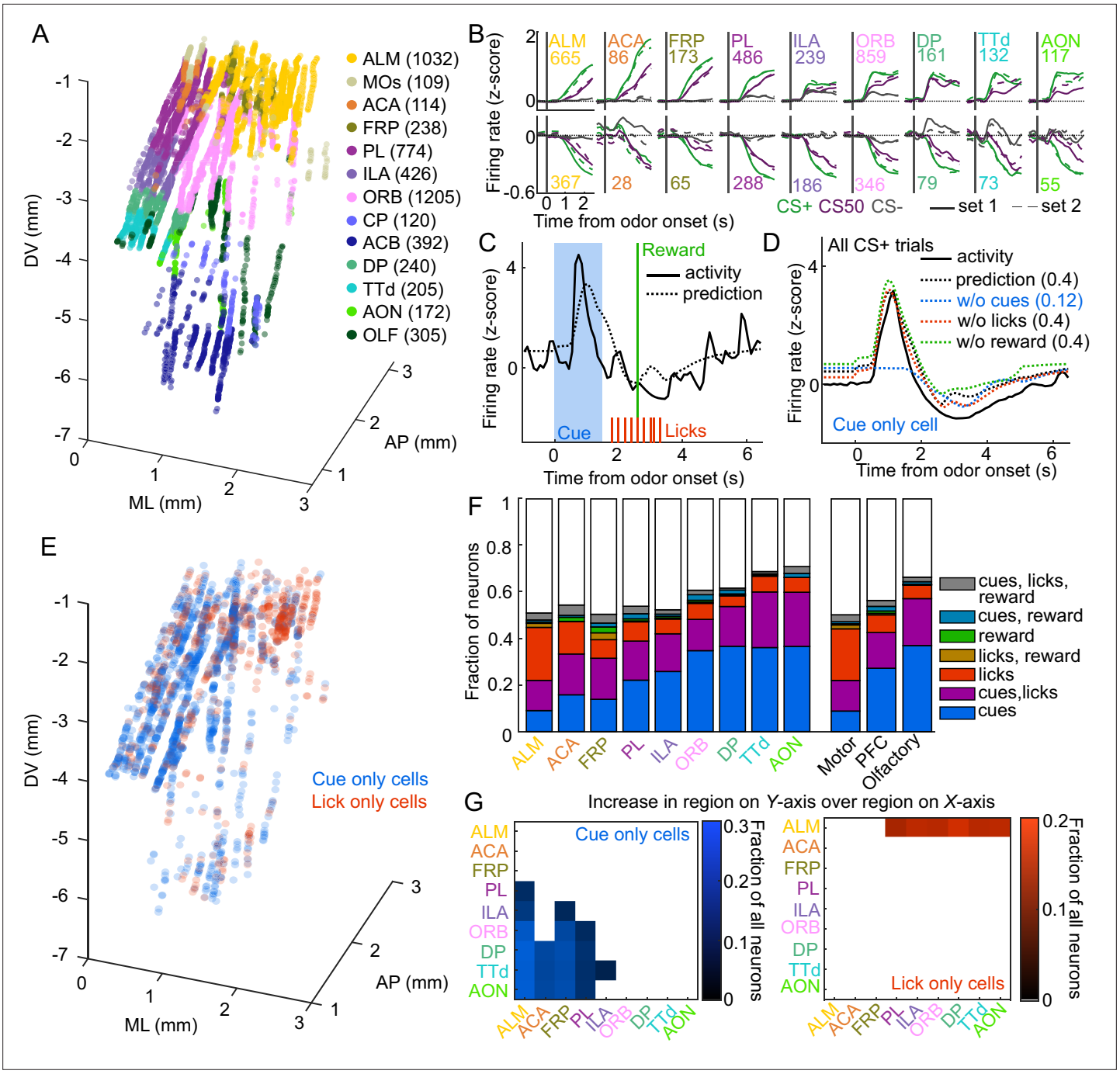

Figure 2. Graded cue and lick coding across the recorded regions. (**A**) Location of each recorded neuron relative to bregma, projected onto one hemisphere. Each neuron is colored by common-coordinate framework (CCF) region. Numbers indicate total neurons passing quality control from each region. (**B**) Mean normalized activity of all neurons from each region, aligned to odor onset, grouped by whether peak cue activity (0–2.5 s) was above (top) or below (bottom) baseline in held out trials. Number of neurons noted for each plot. (**C**) Example kernel regression prediction of an individual neuron's normalized activity on an example trial. (**D**) CS+ trial activity from an example neuron and predictions with full model and with cues, licks, and reward removed. Numbers in parentheses are model performance (fraction of variance explained). (**E**) Coordinates relative to bregma of every neuron encoding only cues or only licks, projected onto one hemisphere. (**F**) Fraction of neurons in each region and region group classified as coding cues, licks, reward, or all combinations of the three. (**G**) Additional cue (left) or lick (right) neurons in region on Y-axis compared to region on x-axis as a fraction of all neurons, for regions with statistically different proportions (see Methods).

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Task-related neural activity across brain regions.

Figure 2 continued on next page

Figure 2 continued

Figure supplement 2. Identification of cue and lick cells with GLM.

Figure supplement 3. Validation of variance cutoff for variable coding.

Figure supplement 4. Comparing proportions of cue and lick neurons across regions.

in the ranked value model; this approach revealed an enrichment in neurons best fit by models most similar to ranked value (35% of cue neurons, ***Figure 3C–D***). We refer to neurons best fit by models most similar to the value model as value-like cells.

Value characteristics were particularly strong among the neurons we identified as value cells. In particular, there was strong modulation for the CS+ odors, moderate modulation for CS50 odors, and the least modulation for CS− odors (***Figure 3E***). These characteristics were present to varying degrees in value-like cells, as well (***Figure 3F***). A key characteristic of value cells, however, was the singular value axis on which the cues were encoded. This was evident when projecting population activity onto the dimensions separating CS+ trials from CS− trials and CS50 trials from CS− trials (***Figure 3—figure supplement 2A***); the trajectory of value neurons traveled the same angle in this space for CS+ and CS50 trials, but differed for value-like (***Figure 3—figure supplement 2B***). We additionally characterized the coding properties of these populations with single-unit and pseudo ensemble decoding. For individual neurons decoding the six cue identities, performance was better using value cells than value-like or untuned cells (***Figure 3G***). At the population level, however, all groups of neurons performed similarly (***Figure 3H***). A key feature of a value signal beyond decoding cue identity, though, is the ability to represent many distinct cues along a shared value axis. Therefore, the value cells should be able to decode the value of a cue never presented during the training of the model. With this approach, models trained on value cells had better predictions of held-out cue value, leading to higher decoding accuracy (CS+, CS50, or CS−), compared to value-like and untuned cells (***Figure 3I***). Therefore, we successfully identified a population of neurons strongly encoding key features of value.

Interestingly, the frequency of value cells was similar across the recorded regions (***Figure 4A***). Despite the regional variability in the number of cue cells broadly (***Figure 2F–G***), there were very few regions that statistically differed in their proportions of value cells (***Figure 4A***, ***Figure 4—figure supplement 1***). Overall, there were slightly more value cells across all of PFC than in motor and olfactory cortex (***Figure 4A***, ***Figure 4—figure supplement 1***). Although the olfactory cortex had the most cue cells, these were less likely to encode value than cue cells in other regions (***Figure 4—figure supplement 2***). Value-like cells were also widespread; they were less frequent in the motor cortex as a fraction of all neurons, but they were equivalently distributed in all regions as a fraction of cue neurons (***Figure 4B***, ***Figure 4—figure supplement 1***, ***Figure 4—figure supplement 2***).

We next investigated the robustness of the value representation in each of our recorded regions. Principal component analysis on value and value-like cells from each region revealed similarly strong value-related dynamics across motor, prefrontal, and olfactory regions (***Figure 4C–D***). We quantified the robustness of value coding in each region by decoding cue value using selections of value cells from each region and found similar performance across all regions (***Figure 1E***). Taken together, these data illustrate that, in contrast to cue and lick coding broadly, value coding is similarly represented across the regions we sampled. In fact, this observation extended to the striatal regions we sampled as well, indicating that such value coding is widespread even beyond cortex (***Figure 4—figure supplement 3***).

Because cue valuations can be influenced by preceding reward outcomes, we next considered whether the cue value signaling we detected was sensitive to the history of reinforcement (***Nakahara et al., 2004***; ***Ottenheimer et al., 2020***; ***Winkelmeier et al., 2022***). To estimate the subjects' trial-by-trial cue valuation, we fit a linear model predicting the number of anticipatory licks on each trial using cue type, overall reward history, and cue type-specific reward history as predictors. We found a strong influence of cue type-specific reward history and a more modest influence of overall reward history (***Figure 5A***). We used the model prediction of licks per trial as our estimate of trial value; the effects of reward history on lick rate were apparent when grouping trials by the value estimates from the trial value model (***Figure 5B***).

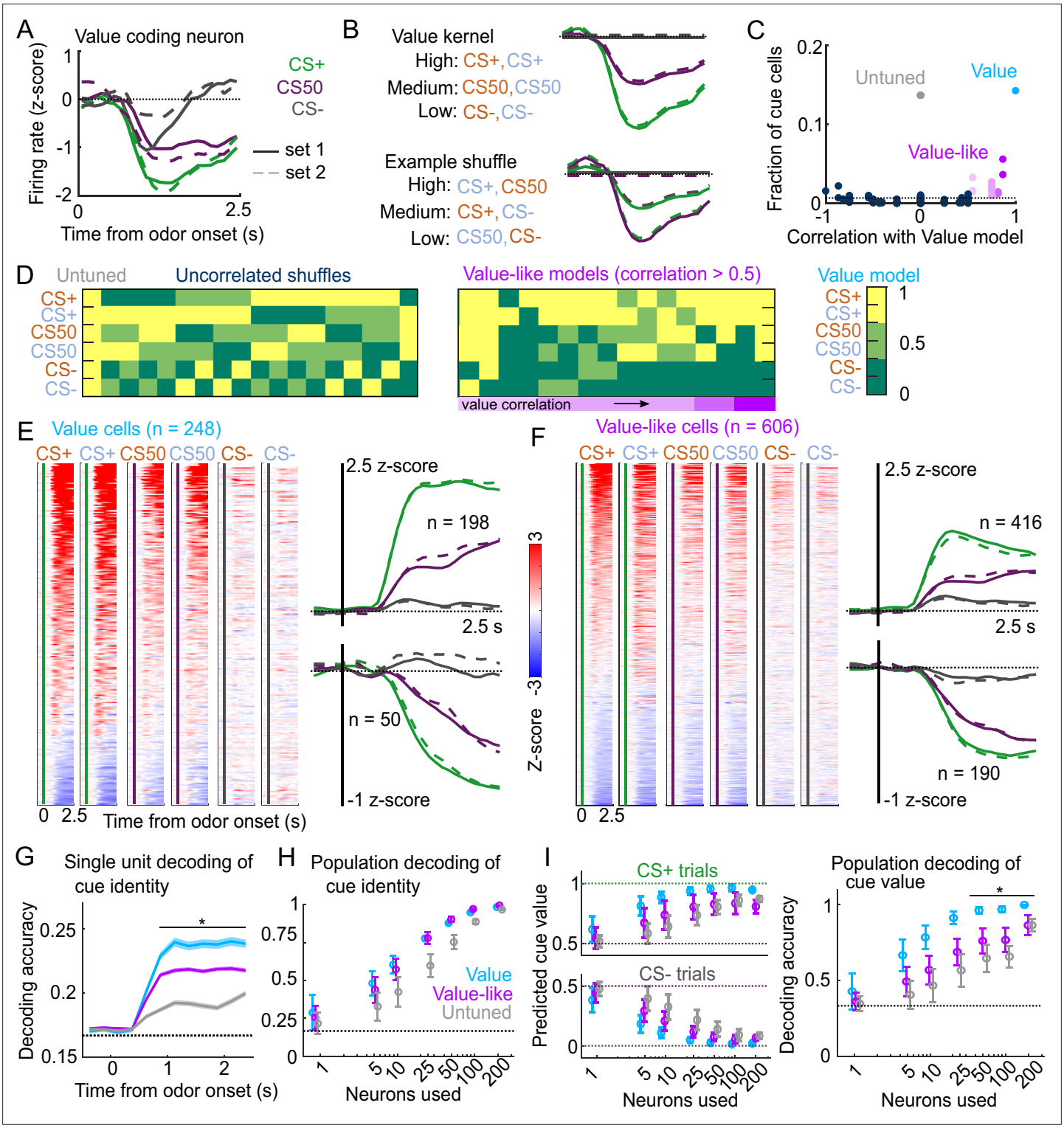

Figure 3. Robust value encoding and decoding among cue cells. (**A**) Normalized activity of an example value cell with increasing modulation for cues with higher reward probability.(**B**) For the same neuron, model-fit cue kernel for the original value model and with one of the 152 alternatively-permuted cue coding models. (**C**) Distribution of best model fits across all cue neurons. Light blue is value model, purple is value-like models, gray is untuned model, and the remaining models are dark blue. Value-like models are shaded according to their correlation with ranked value, as illustrated in (**D**). Dashed line is chance proportion when assuming even distribution. (**D**) Schematic of value assigned to each of the six cues for many of the cue coding models (full schematic in ***Figure 3—figure supplement 1***). Value-like models are sorted by their correlation with the ranked value model. (**E**) Left: normalized activity of every value cell, sorted by mean firing 0–1.5s following odor set A CS+ onset. Right: mean normalized activity of all value

Figure 3 continued on next page

Figure 3 continued

cells, grouped by whether peak cue activity (0–2.5s) was above (top) or below (bottom) baseline in held out trials. Number of neurons noted for each plot. (**F**) As in (**E**), for value-like cells. (**G**) Accuracy (mean ± SEM across neurons) of decoded cue identity for single neurons of value ($n = 248$), value-like ($n = 606$), and untuned ($n = 238$) neurons. * indicates where value, value-like, and untuned neurons significantly differed from each other and baseline (all $p < 0.001$, Bonferroni corrected). All pairwise comparisons in ***Supplementary file 2***. (**H**) Accuracy (mean ± SD across bootstrapped iterations) of decoded cue identity using different numbers of neurons. (**I**) Left: estimated value (mean ± SD across 1000 bootstrapped iterations) of held out CS+ (top) and CS− (bottom) trials using linear models trained on the activity of value, value-like, or untuned neurons. Right: accuracy (mean ± SD across bootstrapped iterations) of decoded cue value using these value estimates. * indicates where the accuracy of value neurons exceeded value-like and untuned neurons (all $p < 0.016$, bootstrapped). All pairwise comparisons in ***Supplementary file 2***.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Schematic of value model shuffles.

Figure supplement 2. Population analysis of value coding schemes.

We, therefore, investigated whether value cells showed similar trial-by-trial differences in their cue-evoked firing rates (***Figure 5C***). To test this, we compared the fit of our original cue coding models (***Figure 3B–D***) with an alternative model in which the kernel scaled with the per-trial value estimates from our trial value model (***Figure 5D***). Overall, 5% of cue cells, including 15% of the value cells, were best fit by the history model. Although the number of anticipatory licks per trial was used to generate the trial value estimates, the precise licking pattern on those trials was a poorer predictor of neural responses than the trial value-scaled cue kernel model (***Figure 5E***). To further evaluate the history component of these neurons, we calculated these neurons' activity on CS50 trials of varying value estimates from the trial value model and projected it onto the population dimension maximizing the separation between CS+ and CS−. We hypothesized that high value CS50 trials would be closer to CS+ activity while low value CS50 trials would be closer to CS− activity. Indeed, history cells (and lick cells) demonstrated graded activity along this dimension, in contrast to non-history value, value-like, and untuned cells (***Figure 5F–H***). Finally, we examined the regional distribution of history cells and found low numbers across all regions, but with a higher prevalence overall in PFC than in motor and olfactory cortex (***Figure 5I***), lending additional support for slightly enhanced value coding in PFC.

Cue coding emerges along with behavioral learning

To determine the timescales over which these coding schemes emerged and persisted, we performed longitudinal 2-photon calcium imaging and tracked the activity of individual neurons across several days of behavioral training (***Figure 6A***). We targeted a GRIN lens to PL, a location with robust cue and lick coding (***Figure 2F***) and where cue responses were predominantly value or value-like (***Figure 4A–B***, ***Figure 4—figure supplement 2***). Mice ($n = 8$) developed anticipatory licking during the first sessions of odor set A (A1) that differentiated CS+ trials from CS50 ($t(7) = 3.2$, $p = 0.015$) and CS− ($t(7) = 7.0$, $p = 0.0002$) trials and CS50 trials from CS− ($t(7) = 3.7$, $p = 0.008$) trials (***Figure 6B–C***). Visualizing the normalized activity across the imaging plane following CS+ presentation early and late in session A1 revealed a pronounced increase in modulation across this first session (***Figure 6D–E***). Individual neurons ($n = 705$, 41-165 per mouse) also displayed a notable increase in modulation in response to the CS+ after task learning (***Figure 6F***).

To determine whether this increase in activity was best explained by a cue-evoked response, licking, or both, we again used kernel regression to fit and predict the activity of each neuron for early, middle, and late trials in session A1. The number of individual neurons encoding cues more than doubled from early to late A1 trials (***Figure 6G***). The unique variance cues increased across this first session, in contrast to licks and reward (***Figure 6H***). This stark change in cue coding was also noticeable when plotting neurons encoding cues, licks, or both, as defined at the end of the sessions, on both early and late trials (***Figure 6I***). These data indicated that PFC neural activity related to cues (but not licks) rapidly emerge during initial learning of the behavioral task.

Cue and lick coding is stable across days

We next assessed whether cue and lick coding were stable across days. By revisiting the same imaging plane on each day of training, we were able to identify neurons that were present on all three days of odor set A training ($n = 371$, 20-65 per mouse) (***Figure 7A–B***). There was remarkable conservation of task responding across days, both on an individual neuron level (***Figure 7C***) and across all

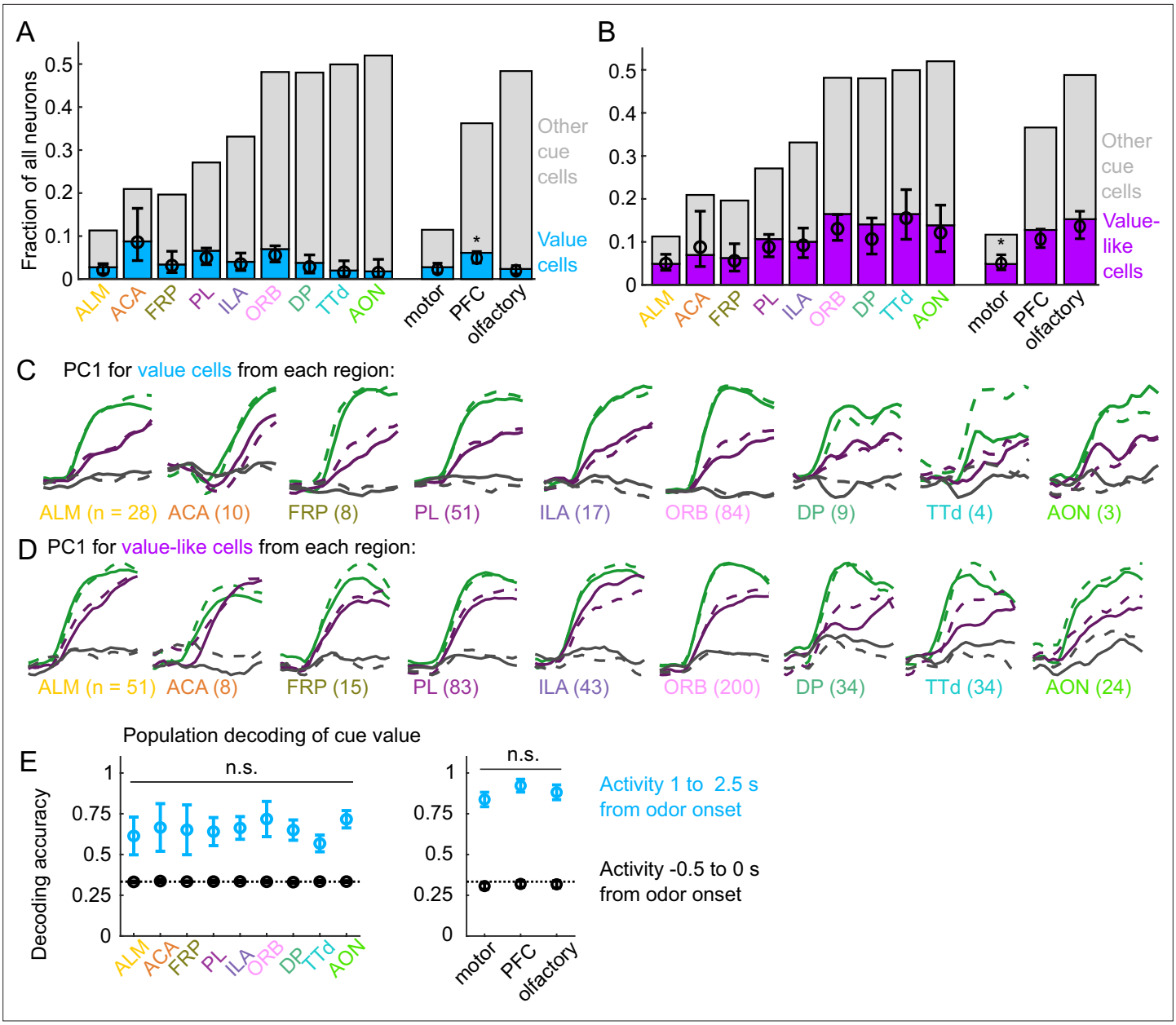

Figure 4. Widespread cue value coding. (**A**) Fraction of neurons in each region and region group classified as value cells (blue) and other cue neurons (gray), as well as fraction (± 95% CI) estimated from a linear mixed effects model with random effect of session (see Methods). Prefrontal cortex (PFC) has more value cells than motor ($p = 0.002$) and olfactory ($p = 0.00005$) cortex. All pairwise comparisons in ***Supplementary file 3***. (**B**) As in (**A**), for value-like cells. Motor cortex has fewer value-like cells than PFC ($p = 8 * 10^{-6}$) and olfactory cortex ($p = 4 * 10^{-8}$). All pairwise comparisons in ***Supplementary file 3***. (**C**) First principal component value cells from all regions. (**D**) As in (**C**), for value-like cells. (**E**) Accuracy of decoded cue value (mean ± SD across 1000 bootstrapped iterations) as in ***Figure 3I***, using five (with replacement) value cells from each region (left) and 25 value cells from each region group (right) using cue-evoked (blue) and baseline (black) activity. No regions or region groups significantly differed from each other ($p > 0.46$, Bonferroni corrected). All pairwise comparisons in ***Supplementary file 3***.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Relative proportions of value and value-like cells across regions.

Figure supplement 2. Value coding as a proportion of cue cells.

Figure supplement 3. Comparing prefrontal cortex (PFC) and striatum.

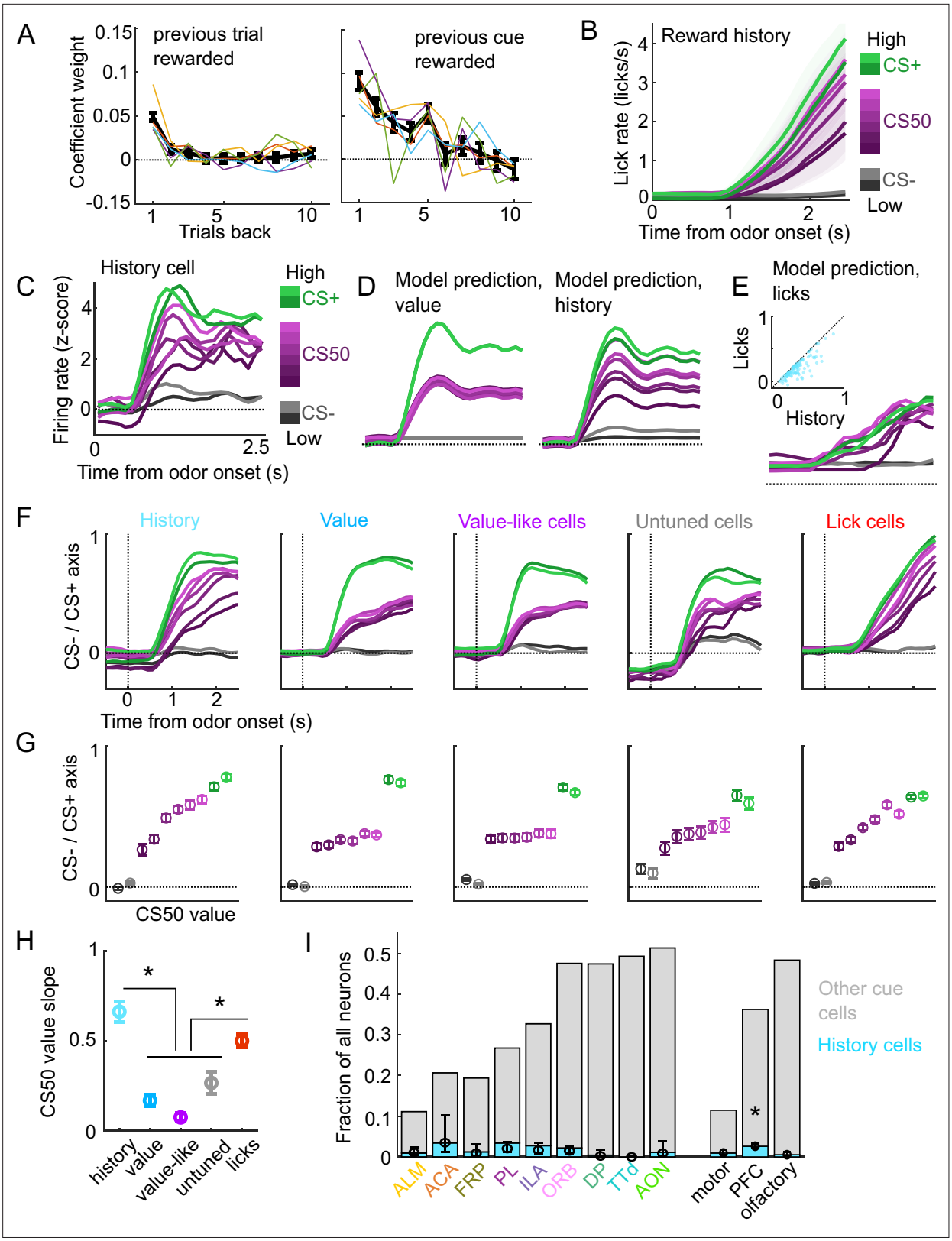

Figure 5. A subset of cue cells incorporate reward history. (**A**) Coefficient weight (± standard error from model fit) for reward outcome on the previous 10 trials of any type (left) and on the previous 10 trials of the same cue type (right) for the 'trial value' model: a linear model predicting the number of anticipatory licks on every trial of every session. Lick rates were normalized so that the maximum lick rate for each session was equal to 1. Colored lines are models fit to each individual mouse. (**B**) Mean (± SEM) lick rate across mice ($n = 5$ mice) on trials binned according to value estimated from

Figure 5 continued on next page

Figure 5 continued

the trial value model. (**C**) Normalized activity of an example history value cell with increasing modulation for cues of higher value. (**D**) For the same neuron, model-predicted activity with the original value model (left) and with the history model, which uses trial-by-trial value estimates from the trial value model (right). (**E**) For the same neuron, model-predicted activity using licks. Inset: variance explained using licks versus history for history neurons. (**F**) The activity of all cells in each category projected onto the coding dimension maximally separating CS− and CS+ for trials binned by value estimated from the trial value model. (**G**) The mean (± SD across 5000 bootstrapped selections of neurons) activity (1–2.5s from odor onset) along the coding dimension maximally separating CS− and CS+ for trials binned by value estimated from the lick model. (**H**) The mean (± SD across 5000 bootstrapped selections of neurons) slope of the activity on CS50 trials regressed onto the trial value model estimate for those trials. History and lick cells had greater slopes than the other groups ($p < 0.0003$, see ***Supplementary file 4***). (**I**) Fraction of neurons in each region and region group classified as history cells (light blue) and other cue neurons (gray), as well as estimated fraction (± 95% CI) with random effect of session (see Methods). Prefrontal cortex (PFC) had more history cells than motor ($p = 0.0016$) and olfactory ($p = 0.00053$) cortex. All pairwise comparisons in ***Supplementary file 4***.

imaged neurons (***Figure 7D***). In fact, neurons were much more correlated with their own activity on the subsequent day than would be expected by chance (***Figure 7E***, ***Figure 7—figure supplement 1A***). To further quantify coding stability, we fit our kernel regression to the activity of each neurons on session A3 (***Figure 7F***) and then used these models to predict activity in early, middle, and late trials on sessions A1-3. Session A3 model predictions were most highly correlated with true activity during A3, but they outperformed shuffle controls at all time points, demonstrating preservation of a learned coding scheme (***Figure 7G***, ***Figure 7—figure supplement 1B***). We then asked more specifically whether cells coding cues, licks, and both maintained their coding preferences across days. For each group of cells, we calculated their unique cue, lick, and reward variance at each time point. The preferred coding of each group, as defined in session A3, was preserved in earlier days (***Figure 7H***). Thus, cue and lick coding are stable properties of PFC neurons across multiple days of behavioral training.

A subset of mice ($n = 5$) also learned a second odor set (odor set B), presented on separate days. Activity was very similar for both odor sets, evident across the entire imaging plane (***Figure 8A***), for individual tracked neurons ($n = 594$, 81-153 per mouse) (***Figure 1—figure supplement 2B***), and for kernel regression classification of these neurons (***Figure 8B***). Notably, odor set A models performed similarly well at predicting both odors set A and odor set B activity (***Figure 8C***). Moreover, cue, lick, and both neurons maintained their unique variance preference across odor sets (***Figure 8D***). Finally, to investigate the presence of value coding across odor sets over separate days, we fit tracked cue neurons with the value model and its shuffles. Even with odor sets imaged on separate days (days 5 and 6 of training, A3 and B3), we again found that the value and value-like models were the best models for sizable fractions (9% and 47%, respectively) of cue neurons, demonstrating that value coding is conserved across stimulus sets on consecutive days (***Figure 8E–G***). Given the prominence of value-like signals in this imaged population, we then assessed the stability of cue cells with preferential CS+ responses across the tracked A1-3 sessions and found conservation of a value-like coding pattern (***Figure 8H***) and, as with the whole population (***Figure 7G***), greater correlation in activity across days than expected by chance (***Figure 8I***).

Discussion

Our experiments assessed how coding for reward-predicting cues and reward-seeking actions differed across brain regions and across multiple days of training. We found coding for cues and licks in all regions we sampled, but their proportions varied in a graded way across those regions. In contrast to regional differences in the proportion of cue-responsive neurons, cue-value cells were present in all regions and value could be decoded from them with similar accuracy regardless of the region. Coding for cue value was greatly overrepresented compared to alternative cue coding schemes and, in a subset of neurons, incorporated the recent reward history. Cue coding was established within the first day of training and neurons encoding cues or licks maintained their coding preference across multiple days of the task; the value characteristics of cue cells were also maintained across days. These results demonstrate widespread value coding and stability of cue and lick codes in PFC.

Graded cue and lick coding across regions

We found robust and separable coding for licks and cues (and combined coding of both) in all regions using electrophysiology and in PL using calcium imaging. The widespread presence of lick coding

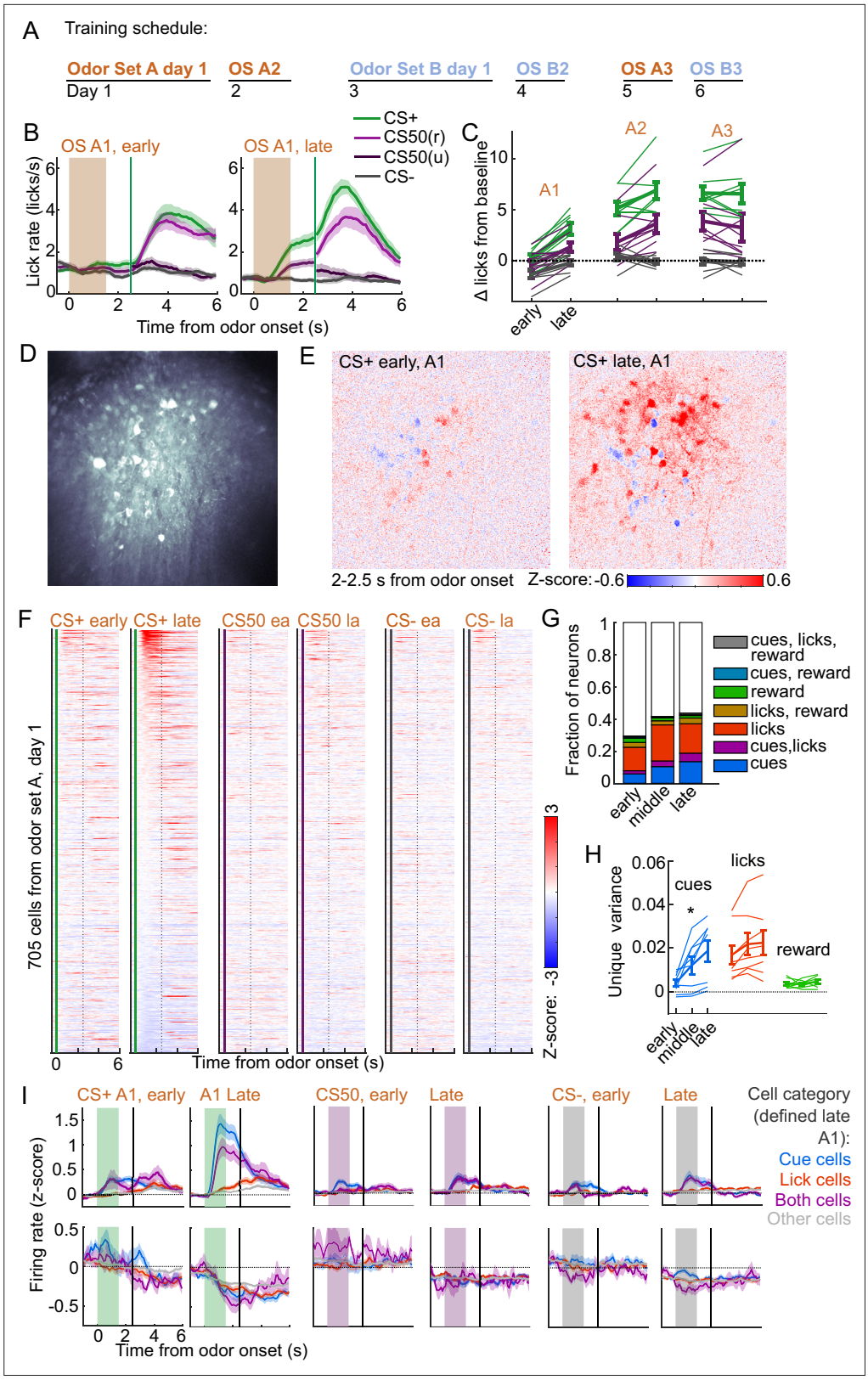

Figure 6. Acquisition of conditioned behavior and cue encoding in prefrontal cortex (PFC). (**A**) Training schedule for five of the mice in the calcium imaging experiment. An additional three were trained only on odor set A. (**B**) Mean (± SEM) licking on early (first 60) and late (last 60) trials from day 1 of odor set A ($n = 8$ mice). (**C**) Mean (± SEM) baseline-subtracted anticipatory licks for early and late trials from each day of odor set A. Thin lines are

Figure 6 continued on next page

Figure 6 continued

individual mice ($n = 8$ mice). (**D**) Standard deviation of fluorescence from example imaging plane. (**E**) Normalized activity of each pixel following CS+ presentation on early and late trials of session A1. (**F**) Normalized deconvolved spike rate of all individual neurons on early and late trials of session A1. (**G**) Proportion of neurons classified as coding cues, licks, rewards, and all combinations for each third of session A1. (**H**) Mean(± SEM across mice) unique variance explained by cues, licks, and rewards for neurons from each mouse. Thin lines are individual mice. Unique variance was significantly different across session thirds for cues ($F(2, 21) = 3.71,\ p = 0.04$) but not licks ($F(2, 21) = 0.37,\ p = 0.69$) or reward ($F(2, 21) = 0.65,\ p = 0.53$, $n = 8$ mice, one-way ANOVA). (**I**) Mean (± SEM) normalized deconvolved spike rate for cells coding cues ($n = 84$ above, $n = 28$ below), licks ($n = 91$ above, $n = 40$ below), both ($n = 31$ above, $n = 9$ below), or neither ($n = 307$ above, $n = 153$ below) on early and late trials, sorted by whether peak cue activity (0–2.5 s) was above (top) or below (bottom) baseline for late trials.

is consistent with recent reports of distributed movement and action coding (***Stringer et al., 2019***; ***Musall et al., 2019***; ***Steinmetz et al., 2019***); however, we saw sizable differences in the amount of lick coding across recorded regions. Notably, ALM had the greatest number of lick neurons, as well as the fewest cue neurons, perhaps reflecting its specialized role in the preparation and execution of licking behavior (***Chen et al., 2017***). Conversely, the olfactory cortical regions DP, TTd, and AON had the most cue neurons (especially non-value coding cue neurons), suggesting a role in early odor identification and processing (***Mori and Sakano, 2021***). PFC subregions balanced lick and cue coding, consistent with their proposed roles as association areas (***Miller and Cohen, 2001***; ***Klein-Flügge et al., 2022***), but there was variability within PFC as well. In particular, ORB had a greater fraction of cue cells than any other subregions, consistent with its known dense inputs from the olfactory system (***Price, 1985***; ***Price et al., 1991***; ***Ekstrand et al., 2001***). Thus, our results establish that the neural correlates of this Pavlovian conditioned behavior consist of a gradient of cue and response coding rather than segmentation of sensory and motor responses.

Widespread value signaling

Value signals can take on many forms and occur throughout task epochs. In our experiments, we focused on the predicted value associated with each conditioned stimulus, which is crucial for understanding how predictive stimuli produce motivated behavior (***Berridge, 2004***). Surveys of value coding in primate PFC have found individual neurons correlated with stimulus-predicted value in many subregions, with the strongest representations typically in ORB (***Roesch and Olson, 2004***; ***Sallet et al., 2007***; ***Kennerley et al., 2009***; ***Hunt et al., 2018***). In rodents, there is also a rich literature on value signaling in ORB (***Schoenbaum et al., 2003***; ***van Duuren et al., 2009***; ***Sul et al., 2010***; ***Stalnaker et al., 2014***; ***Namboodiri et al., 2019***; ***Kuwabara et al., 2020***; ***Wang et al., 2020a***), but there have also been many reports of value-like signals in frontal cortical regions beyond ORB (***Otis et al., 2017***; ***Allen et al., 2019***; ***Wang et al., 2020a***; ***Kondo and Matsuzaki, 2021***). In our present experiment, we sought to expand upon these rodent results by separating cue activity from licking, which tracks the value and may confound interpretation, by including more than two cue types, which provided a rich space to assess value coding, and by sampling from many frontal regions in the same experiment.

When considering the number of neurons responsive to cues rather than licks, our data confirmed the importance of ORB, which has more cue-responsive neurons than the motor and other prefrontal regions, but, beyond cue responsiveness, we were interested in identifying specific cue coding patterns pertaining to value. By analyzing the activity of cue-responsive neurons across all six odors predicting varying probabilities of reward, we were able to isolate neurons coding value, as well as those with value-like signals that could easily be misconstrued as value-coding in a task with fewer cues and value levels. Included in the value-like models are coding patterns that bias their activity for higher value odors without fitting our strict linear ranked value criteria; for instance, selective firing for one or two of the CS+ odors. The enrichment of these models among cue responsive neurons, even in the olfactory cortex, indicates the prevalence of value-biased coding schemes for odor-responsive neurons across brain regions. The question remains of where odor information is first shaped according to value. There have been multiple reports of some association-related modification of odor representations as early as the olfactory bulb (***Doucette et al., 2011***; ***Li et al., 2015***; ***Chu et al., 2016***; ***Koldaeva et al., 2019***). Considering we detected value and especially value-like coding in AON, DP, and TTd, perhaps these regions are a crucial first step in processing and amplifying task-related input from the

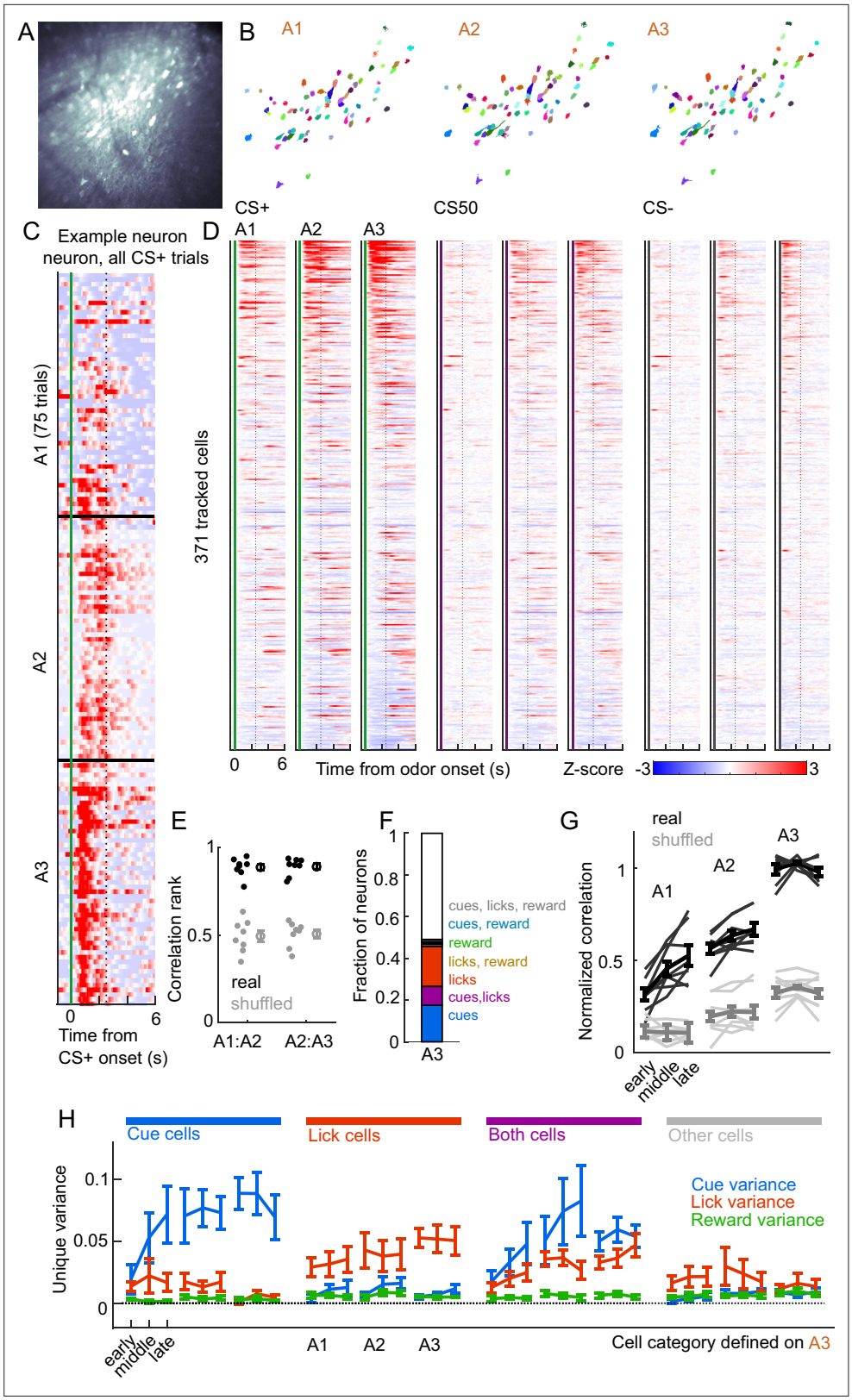

Figure 7. Cue and lick coding is stable across days. (**A**) Standard deviation fluorescence from example imaging plane. (**B**) Masks (randomly colored) for all tracked neurons from this imaging plane. (**C**) Deconvolved spike rate on every CS+ trial from all three sessions of odor set A for an example neuron. Vertical dashed line is reward delivery. Color axis as in (**D**). (**D**) Normalized deconvolved spike rate for all tracked neurons on all three sessions of odor

Figure 7 continued on next page

Figure 7 continued

set A. (**E**) Correlation between the activity of a given neuron in one session and its own activity in the subsequent session, quantified as a percentile out of correlations with the activity of all other neurons on the subsequent day. Plotted as the median for each subject and the mean (± SEM) across these values. Real data was more correlated than shuffled data ($p = 0.0078$ for both comparisons, Wilcoxon signed-rank test). (**F**) Fraction of tracked neurons coding cues, licks, rewards, and their combinations on day 3. (**G**) Model performance when using models from session A3 to predict the activity of individual neurons across session thirds of odor set A training, plotted as mean (± SEM) correlation between true and predicted activity across mice, normalized to the correlation between model and training data. Thin lines are individual mice. Performance was greater than shuffled data at all time points ($p < 0.002$, Bonferroni-corrected, $n = 8$ mice). Non-normalized data in ***Figure 7—figure supplement 1***. (**H**) Mean (± SEM across mice) unique cue, lick, and reward variance for cells classified as coding cues, licks, both, or neither on session A3. A3 cue cells had increased cue variance in A2 ($p < 10^{-7}$, see Methods) and A1 ($p < 0.03$) relative to lick and reward variance. Same pattern for A3 lick cells in A2 ($p < 0.0001$) and A1 ($p < 0.01$).

The online version of this article includes the following figure supplement(s) for figure 7:

Figure supplement 1. Correlation across days in prelimbic area (PL).

olfactory bulb. Because they provide input to PFC (***Igarashi et al., 2012***; ***Bhattarai et al., 2022***), they may be an important source of the cue coding we observed there.

The distribution of cue cells with linear coding of value was mostly even across regions, with slight enrichment overall in PFC compared to the motor and olfactory cortex, but no subregional differences in PFC. Importantly, cue value could be decoded from value cells in each region with similar accuracy. One consequence of a widely distributed value signal is that manipulating only one subregion would be less likely to fully disrupt value representations, which is consistent with the results of studies comparing functional manipulations across PFC (***Chudasama and Robbins, 2003***; ***St Onge and Floresco, 2010***; ***Dalton et al., 2016***; ***Verharen et al., 2020***; ***Wang et al., 2020a***). Different subregional impacts on behavior may reveal biases in how the value signal in each region contributes to reward-related behaviors, for instance during learning or expression of a reward associations (***Otis et al., 2017***; ***Namboodiri et al., 2019***; ***Wang et al., 2020a***). A related interpretation is that, in this task, there may be other properties that correlate with cue value, and the homogeneous value representation we observed across regions masks regional differences in tuning to these other correlated features, such as motivation (***Roesch and Olson, 2004***) and a host of related concepts, including salience, uncertainty, vigor, and arousal (***Stalnaker et al., 2015***; ***Hayden and Niv, 2021***; ***Zhou et al., 2021***), which can have different contributions to behavior. This interpretation is consistent with broader views that observations of 'value' signals are often misconstrued (***Zhou et al., 2021***) and that pure abstract value may not be encoded in the brain at all (***Hayden and Niv, 2021***). Although the identification of value in our task was robust to three levels of reward probability across two stimulus sets, the fact that this signal was widespread contributes to the case for revisiting the definition and interpretation of value to better understand regional specialization.

In our analysis, we uncovered a distinction between neurons encoding the overall value of cues and those with value representations that incorporated the recent reward history. Neurons with history effects were rare and most frequent in PFC. These neurons may have a more direct impact on behavioral output in this task, because the lick rate also incorporated recent reward history. Notably, the impact of reward history on these neurons was noticeable even prior to cue onset, consistent with a previously proposed mechanism for persistent value representations encoded in the baseline firing rates of PFC neurons (***Bari et al., 2019***).

Stability of PFC codes

Previous reports have observed drifting representations in PFC across time (***Hyman et al., 2012***; ***Malagon-Vina et al., 2018***), and there is compelling evidence that odor representations in piriform drift over weeks when odors are experienced infrequently (***Schoonover et al., 2021***). On the other hand, it has been shown that coding for odor association is stable in ORB and PL, and that coding for odor identity is stable in piriform (***Wang et al., 2020a***), with similar findings for auditory Pavlovian cue encoding in PL (***Otis et al., 2017***; ***Grant et al., 2021***) and ORB (***Namboodiri et al., 2019***). We were able to expand upon these data in PL by identifying both cue and lick coding and showing separable, stable coding of cues and licks across days and across sets of odors trained on separate days. We

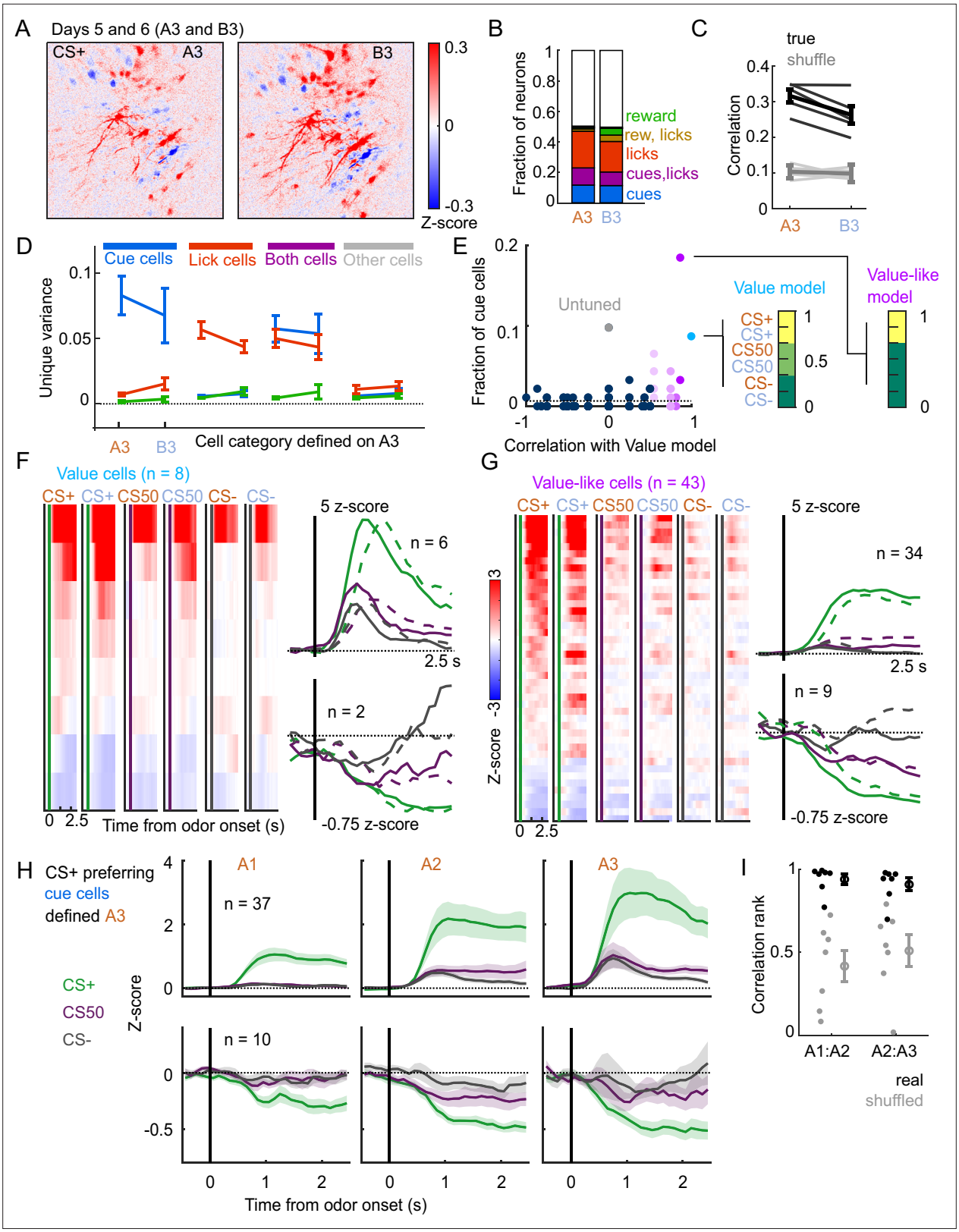

Figure 8. Stable cue coding across separately trained odor sets. (**A**) Normalized activity of all pixels in the imaging plane following CS+ presentation on the third day of each odor set (A3 and B3, days 5 and 6 of training). (**B**) Fraction of neurons coding for cues, licks, rewards, and their combinations in A3 and B3 (days 5 and 6). (**C**) Mean (± SEM, across mice) correlation between activity predicted by odor set A3 models and its training data (A3, cross-validated) or activity in B3, for true (black) and trial shuffled (gray) activity. Thin lines are individual mice. $F(1, 16) = 3.2,\ p = 0.09$ for main effect

Figure 8 continued on next page

Figure 8 continued

of odor set, $F(1,16) = 135$, $p < 10^{-8}$ for main effect of shuffle, $F(1,16) = 2.2$, $p = 0.16$ for interaction, $n = 5$ mice, two-way ANOVA. (**D**) Mean (± SEM, across mice) unique cue, lick, and reward variance for cells classified as coding cues, licks, both, or neither for odor set A. For each category, odor set A unique variance preference was maintained for odor set B ($p < 0.04$) except for both cells, for which lick and reward variance were not different in odor set B ($p = 0.22$, Bonferroni-corrected, $n = 5$ mice). (**E**) Distribution of best model fits across all cue cells, with colors from ***Figure 3C***. Dashed line is chance proportion when assuming even distribution. (**F**) Left: normalized activity of every value cell, sorted by mean firing 0–1.5s following odor set A CS+ onset. Right: mean normalized activity of all value cells, grouped by whether peak cue activity (0–2.5 s) was above (top) or below (bottom) baseline in held out trials. Number of neurons noted for each plot. (**G**) As in (**E**), for value-like cells. (**H**) Mean (± SEM, across neurons) activity of cue cells tracked across A1, A2, and A3 with preferential CS+ firing, defined on half of A3 trials and plotted for the other half of A3 trials and all of A1 and A2 trials. (**I**) For neurons in (**H**), correlation between a neuron's activity in one session and its own activity in the subsequent session, quantified as a percentile out of correlations with the activity of all other neurons on the subsequent day. Plotted as the median for each subject ($n = 7$ with CS+ preferring cue cells) and the mean (± SEM) across these values. Real data was more correlated than shuffled data ($p = 0.016$ A1:A2, $p = 0.031$ A2:A3, Wilcoxon signed-rank test).

were also able to detect value coding common to two stimulus sets presented on separate days, and conserved value features across the three training sessions. Notably, the model with responses only to CS+ cues best fit a larger fraction of imaged PL neurons than the ranked value model, a departure from the electrophysiology results. It would be interesting to know if this is due to a bias introduced by the calcium imaging approach, the slightly reduced CS50 licking relative to CS+ licking in the imaging cohort, or the shorter imaging experimental timeline.

The consistency in cue and lick representations we observed indicates that PL serves as a reliable source of information about cue associations and licking during reward-seeking tasks, perhaps contrasting with other representations in PFC (***Hyman et al., 2012***; ***Malagon-Vina et al., 2018***). Interestingly, the presence of lick, but not cue coding at the very beginning of the first session of training suggests that lick cells in PL are not specific to the task but that cue cells are specific to the learned cue-reward associations. Future work could expand upon these findings by examining stimulus-independent within session value coding across many consecutive days.

Overall, our work emphasizes the importance of evaluating the regional specialization of neural encoding with systematic recordings in many regions using the same task. Future work will clarify whether cue value is similarly widely represented in other reward-seeking settings and whether there are regional differences in the function of the value signal.

Materials and methods

Subjects

Subjects ($n = 5$ for electrophysiology, $n = 8$ for calcium imaging) were male and female C57BL/6 mice single-housed on a 12 hr light/dark cycle and aged 12–28 weeks at the time of recordings. Imaging experiments were performed during the dark cycle, electrophysiology during the light cycle. Mice were given free access to food in their home cages for the duration of the experiment. Mice were water restricted for the duration of the experiments and maintained at around 85% of their baseline weight (***Guo et al., 2014a***). All experimental procedures were performed in strict accordance with protocols 4450–01 and 4461–01 approved by the Animal Care and Use Committee at the University of Washington.

Surgical procedures

Mice were anesthetized with isoflurane (5%) and maintained under anesthesia for the duration of the surgery (1–2%). Mice received injections of carprofen (5 mg/kg) prior to incision.

Electrophysiology

A brief (1 hr) initial surgery was performed, as previously described (***Guo et al., 2014b***; ***Steinmetz et al., 2017***; ***Steinmetz et al., 2019***), to implant a steel headbar (approximately 15 × 3 × 0.5 mm, 1 g) for head fixation and a 3D-printed recording chamber exposing the skull for subsequent craniotomies. Briefly, an oval incision was made extending from the interparietal bone to the frontonasal suture, skirting the ocular area. The skin and periosteum were removed to expose the entire dorsal surface of the skull. Skull yaw, pitch, and roll were leveled, and exposed bone was texturized with a

brief application of green activator (Super-Bond C&B, Sun Medical). The incision was secured to the skull with the application of cyanoacrylate (VetBond; World Precision Instruments), and the 3D-printed recording chamber was attached to the skull with L-type radiopaque polymer (Super-Bond C&B). A thin layer of cyanoacrylate was applied to the skull inside the chamber and allowed to dry. Multiple (2-4) thin layers of UV-curing optical glue (Norland Optical Adhesives #81, Norland Products) were applied to the skull inside the chamber and cured with UV light to protect the exposed bone. The headbar was attached to the skull over the interparietal bone posterior to the chamber with Super-Bond polymer, and more polymer was applied around the headbar and chamber. Following recovery, a second brief (15–30 min) surgery was conducted to perform craniotomies for Neuropixels probe insertion. Briefly, following induction of anesthesia a small (2 × 1.5 mm (w × h)) craniotomy was made over the frontal cortex (+2.5–1 mm AP, ± 2.5–0.3 mm ML) with a handheld dental drill. The craniotomy was covered with a soft silicone gel (DOWSIL 3–4680) and the recording chamber was covered with a 3D-printed lid sealed with Kwik-Cast elastomer to protects craniotomy from dust.

Calcium imaging

A Gradient-Refractive Index (GRIN) lens and metal headcap were implanted following previously described procedures (***Namboodiri et al., 2019***) with the following modifications. In most mice, once the dura was removed from the craniotomy, we performed two injections of 0.5 μL of virus (1 μL total) containing the GCaMP gene construct (AAVDJ-CamKIIa-GCaMP6s, $5.3 * 10^{12}$ viral particles/mL from UNC Vector core lot AV6364) using a glass pipette microinjector (Nanoject II) at Bregma +1.94 mm AP, 0.3, and 1.2 mm ML, –2 mm DV. Ten minutes elapsed before the microinjector withdrawal to allow the virus to diffuse away from each infusion site. Then, mice were implanted with a 1 × 4 mm GRIN lens (Inscopix) aimed at +1.94 mm AP, 0.6 mm ML, and –1.8 mm DV. A subset of mice did not receive viral injections; instead, a lens with the imaging face coated 1 μL of the GCaMP6s virus mixed with 5% aqueous silk fibroin solution (***Jackman et al., 2018***) was implanted at the same coordinate. GCaMP expression and transients were similar in both preparations. Mice were allowed to recover for at least 5 weeks before experiments began.

Behavioral training

Mice were headfixed during training and recording sessions using either a headring (imaging experiments) or headbar (electrophysiology experiments). After initial habituation to head fixation, mice were first trained to lick for $2.5 \mu L$ rewards of 10% sucrose solution, delivered every 8–12 s through a miniature inert liquid valve (Parker 003-0257-900). After 4–5 days of lick training, mice experienced their first odor exposure (without reward delivery). Odors were delivered for a total of 1.5 s using a 4-channel olfactometer (Aurora 206 A) with 10% odor flow rate and 800 SCCM overall flow rate of medical air. Odors were randomly assigned to sets and cue identities, counterbalanced across mice. Odors were -carvone, -limonene, alpha-pinene, butanol, benzaldehyde, and geranyl acetate (Sigma Aldrich 124931, 218367, 147524, 281549, 418099, 173495, respectively), selected because of they are of neutral valence to naive mice (***Devore et al., 2013***; ***Saraiva et al., 2016***). Odors were diluted 1:10 in mineral oil and 10 µL was pipetted onto filter paper within the odor delivery vials (Thermo Fisher SS246-0040) prior to each session. Airflow was constant onto the mouse's nose throughout the session and switched from clean air to scented air for the 1.5 s duration odor delivery on each trial.

On days 1–2 of Pavlovian conditioning, mice received 50–75 trials each of three odor cues (odor set A), followed by reward on 100% (CS+), 50% (CS50), or 0% (CS−) of trials, 2.5 s following the odor onset, with 8–12 s between odor presentations. On days 3–4 mice then received training for 2 days with a second odor set (odor set B) with three new odors. For electrophysiology experiments, the odors were subsequently presented in the same sessions in six blocks of 51 trials. Odor set order alternated and was counterbalanced across days. For imaging experiments, mice received the third day of odor set A on day 5 and the third day of odor set B on day 6 of conditioning. An additional three imaging mice were only trained on one odor set.

Electrophysiological recording and spike sorting

During recording sessions, mice were headfixed. Recordings were made using either Neuropixels 1.0 or Neuropixels 2.0 electrode arrays (***Jun et al., 2017***; ***Steinmetz et al., 2021***), which have 384 selectable recording sites. Recordings were made with either 1.0 (1 shank, 960 sites), 2.1 (1 shank,

1280 sites),, or 2.4 (4 shanks, 5120 sites) probes, depending on the regions of interest. Probes were mounted to a dovetail and affixed to a steel rod held by a micromanipulator (uMP-4, Sensapex Inc). For later electrode track localization within the brain, probes were coated with a fluorescent dye (DiI, ThermoFisher Vybrant V22888) by holding 2 μl in a droplet on the end of a micropipette and painting the probe shank. In each session, one or two probes were advanced through the silicone gel covering the craniotomy over the frontal cortex, then advanced to their final position at approximately 3 μm s^{-1}. Electrodes were allowed to settle for around 15 min before starting recording. Recordings were made in internal reference mode using the 'tip' reference site, with a 30 kHz sampling rate. Recordings were repeated at different locations on each of multiple subsequent days, performing an additional craniotomy over the contralateral frontal cortex. The resulting data were automatically spike sorted with Kilosort2.5 and Kilosort3 (https://github.com/MouseLand/Kilosort; RRID:SCR_016422; ***Pachitariu et al., 2023***), v2.5 and 3.0. Extracellular voltage traces were preprocessed with common-average referencing by subtracting each channel's median to remove baseline offsets, then subtracting the median across all channels at each time point to remove common electrical artifacts. Sorted units were curated using automated quality control (***Banga, et al., 2022***): exclusions were based on spike floor violations (the estimated proportion of spikes that were missed because they fell below the noise level of the recording, estimated false negative rate), and refractory period violations (the estimated proportion of spikes arising from the non-primary neuron, the estimated false positive rate due to contamination, with a 10% cutoff). Quality control accuracy was assessed by manually reviewing a subset of the data using the phy GUI (https://github.com/kwikteam/phy; ***Rossant et al., 2021***). Because Kilosort2.5 and Kilosort3 use different clustering algorithms that can be advantageous for different types of recordings (stability, region, number of channels), for each session, we used units sorted with either Kilosort2.5 or Kilosort3 depending on which yielded the greatest number of high-quality units for that session. Brain regions were only included for subsequent analysis if there were recordings from at least three subjects and a total of over 100 neurons in the region. When we analyzed all of the motor cortex together, we included ALM and MOs neurons. When we analyzed all of the olfactory cortex, we included DP, TTd, AON, and other neurons in PIR, EPd, and OLF. We relabeled PIR and EPd as OLF because there were not enough neurons to analyze them as separate regions.

Imaging and ROI extraction

During imaging sessions, mice were headfixed and positioned under the 2-photon microscope (Bruker Ultima2P Plus) using a 20 x air objective (Olympus LCPLN20XIR). A Spectra-Physics InSight X3 tuned to 920 nm was used to excite GCaMP6s through the GRIN lens. Synchronization of odor and 10% sucrose delivery, lick behavior recordings, and 2-photon recordings were achieved with custom Arduino code. After recording, raw TIF files were imported into suite2p (https://github.com/MouseLand/suite2p; RRID:SCR_016434; ***Stringer et al., 2023***), v0.13.0. We used their registration, region-of-interest (ROI) extraction, and spike deconvolution algorithms, inputting a decay factor of $\tau = 1.3$ to reflect the dynamics of GCaMP6s, and manually reviewed putative neuron ROIs for appropriate morphology and dynamics. To find changes in activity across the entire imaging plane, found the mean pixel intensity for frames in the time of interest (2–2.5 s from CS+), subtracted the mean intensity of each pixel prior to cue onset (−2–0 s from all cues), and divided by the standard deviation for each pixel across those frames prior to cue onset.

Histology

Animals were anesthetized with pentobarbital or isoflurane. Mice were perfused intracardially with 0.9% saline followed by 4% paraformaldehyde (PFA).

Electrophysiology

Brains were extracted immediately following perfusion and post-fixed in 4% paraformaldehyde for 24 h. In preparation for light sheet imaging brains were cleared using organic solvents following the 3DISCO protocol (***Ertürk et al., 2012***) (https://idisco.info/), with some modification. Briefly, on day 1 brains were washed 3 X in PBS and dehydrated in a series of increasing MeOH concentrations (20%, 40%, 60%, 80%, 100%, 100%; 1 hr each) then incubated overnight for lipid extraction in 66% dichloromethane (DCM) in MeOH. On day 2 brains were washed 2 X twice in 100% MeOH for 1 hr each, then bleached overnight in 5% H_2O_2 in MeOH at 4 °C. On day 3 brains were washed 2 X in 100% MeOH,

then final lipid extraction was accomplished in a series of DCM incubations (3 hr in 66% DCM in MeOH, 2X 100% DCM for 15 min each) before immersion in dibenzyl ether (DBE) for refractive index matching. Brains were imaged on a light sheet microscope (LaVIsion Biotec UltraScope II) 2–7 days after clearing. Brains were immersed in DBE in the imaging well secured in the horizontal position, and illuminated by a single light sheet (100% width, 4 μm thick) from the right. Images were collected through the 2 X objective at 1 X magnification, from the dorsal surface of the brain to the ventral surface in 10 μm steps in 488 nm (autofluorescence, 30% power) and 594 nm (DiI, 2–10% power) excitation channels. The 1000 raw TIF images were compiled into a single multi-image file with 10 μm voxels, then spatially downsampled to 25 μm voxels for transformation to the Allen common-coordinate framework (CCF) volume (***Wang et al., 2020b***) using the Elastix algorithm (***Shamonin et al., 2013***). CCF-transformed volumes were used to generate CCF fluorescent probe tract locations (pixel coordinates along the probe tract) using Lasagna (https://github.com/SainsburyWellcomeCentre/lasagna; ***Campbell et al., 2020***). Probe tract CCF pixel coordinates (origin front, top, left) were transformed to bregma coordinates (origin bregma, x==ML, y==AP, and z==DV) in preparation for final integration with electrophysiology recordings using the International Brain Lab electrophysiology GUI (Faulkner M, Ephys Atlas GUI; 2020. https://github.com/int-brain-lab/iblapps/tree/master/atlaselectrophysiology; ***Faulkner, 2020***). For recording alignment, sorted spikes and RMS voltage on each channel were displayed spatially in relation to the estimated channel locations in Atlas space from the tracked probe. The recording sites were then aligned to the Atlas by manually identifying a warping such that recording sites were best fit to the electrophysiological characteristics of the brain regions (e.g. matching location of ventricles or white matter tracts with low firing activity bands). This procedure has been estimated to have a 70 μm error (***Steinmetz et al., 2019***; ***Liu et al., 2021***). Individual neuron locations were determined using the recording channel brain coordinates of each unit's maximum-amplitude waveform. We additionally assigned MOs neurons to the anterolateral motor cortex (ALM) if they were within a 0.75 mm radius of 2.5 mm AP, and 1.5 mm ML (***Chen et al., 2017***).

Calcium imaging

Following perfusion, intact heads were left in PFA for an additional week before brain extraction. Brains were then sliced on a Leica Vibratome (VT1000S) at 70 μm before mounting and nuclear staining via Fluoroshield with DAPI (Sigma-Aldrich F6057-20ML). Slices with GRIN lens tracks were then imaged on a Zeiss Axio Imager M2 Upright Trinocular Phase Contrast Fluorescence Microscope with ApoTome. The resulting images were manually aligned to the Allen Brain Atlas to reconstruct the location of each GRIN lens.

Neuron tracking

To identify the same neurons across imaging sessions, we used two approaches. To track neurons across the two odor sets on days 5 and 6, we concatenated the TIF files from each session and extracted ROIs simultaneously. To track neurons across training days 1–3 for a single odor set, we manually identified ROIs from the ROI masks outputted by suite2p. We linked the ROIs using a custom Python script that permitted the selection of the same ROI across the three imaging planes using OpenCV and saved the coordinates on each day. The tracking results across days 1–3 from one subject is displayed in ***Figure 7B***.

Behavioral analysis

For electrophysiology experiments, the subject was illuminated with infrared light (850 nm, CMVision IR30) and eye and face movements were monitored. The right eye was monitored with a camera (FLIR CM3-U3-13Y3M-CS) fitted with a zoom lens (Thorlabs MVL7000) and long-pass filter (Thorlabs FEL0750), recording at 70 fps. Face movements were monitored with another camera (FLIR CM3-U3-13Y3M-CS, zoom lens Thorlabs MVL16M23, long-pass filter Thorlabs FEL0750) directed at a 2 × 3 cm mirror reflecting the left side of the face, recording at 70 fps. Licks were detected from the face video by thresholding the average intensity of an ROI centered between the lips and the lick spout, calculated for every frame. Interlick intervals were thresholded at 0.083 s for a maximum lick rate of 12 licks s^{-1}. For calcium imaging experiments, eye and face movements were not monitored, and licks were detected with a capacitance sensor (MPR121, Adafruit Industries) connected to an Arduino board. To determine the impact of cues and previous outcomes on anticipatory licking, we fit a linear model

on all electrophysiology sessions simultaneously (and for each mouse). We predicted the number of licks 0–2.5 s from odor onset using cue identity, outcomes on the previous 10 trials, outcomes on the previous 10 of that cue type, and the total number of presentations of that cue type so far (to account for cue-specific satiety) using 'fitlm' in MATLAB. When dividing sessions into 'early' and 'late,' we used the first 60 and last 60 trials of the session. When dividing sessions into thirds for the GLM ('early,' 'middle,' 'late'), we used even splits of trials into thirds.

PSTH creation

Peri-stimulus time histograms (PSTHs) were constructed using 0.1 s bins surrounding cue onset.

Electrophysiology

Neuron spike times were first binned into 0.02 s bins and smoothed with a half-normal causal filter ($\sigma = 300$ ms) across 50 bins. PSTHs were then constructed in 0.1 s bins surrounding each cue onset. Each bin of the PSTH was z-scored by subtracting the mean firing rate and dividing the standard deviation across the 0.1 s bins in the 2 s before all trials. When splitting responses by polarity (above/below baseline, ***Figures 2B, 3E–F and 8H***, ***Figure 2—figure supplement 2B***), we used even trials to determine polarity and plotted the mean across odd trials for cross-validation.

Calcium imaging

Frames were collected at 30 Hz with 2-frame averaging, so the fluorescence for each neuron and the estimated deconvolved spiking was collected at 15 Hz. We interpolated the smoothing filter from the electrophysiology analysis (which was calculated at 50 Hz) and applied it to the deconvolved spiking traces. We then constructed PSTHs in 0.1 s bins surrounding each cue onset and z-scored (same as electrophysiology).

Licks

Licking PSTHS were constructed in 0.1 s bins surrounding cue onset. Each trial was then smoothed with a half-normal causal filter ($\sigma = 800$ ms). For the GLM, the lick rate was calculated across the whole session by first counting licks in either the 0.02 s (electrophysiology) or 15 Hz (imaging) bins, smoothed with a half-normal causal filter over 25 bins, and then converted to 0.1 s bins relative to each cue.

Kernel regression

To identify coding for cues, licks, and rewards in individual neurons, we fit reduced rank kernel-based linear model (***Steinmetz et al., 2019***).

Data preparation

The discretized firing rates $f_n(t)$ for each neuron n were calculated as described above for PSTH creation. We used the activity –1–6.5 s from each cue onset on every trial for our GLM analysis.

Predictor matrix

The model included predictor kernels for cues (CS+, CS50, and CS− for each odor set, as relevant), licks (individual licks, lick bout start, and lick rate), and reward (initiation of consummatory bout). The cue kernels were supported over the window 0–5 s relative to the stimulus onset. The lick predictor kernels were supported from –0.3–0.3 s relative to each lick, from –0.3–2 s relative to lick bout start, and lick rate was shifted from –0.4–0.6 s in 0.2 s increments from original rate. The reward kernel was supported 0–4 s relative to first lick following reward delivery. For electrophysiology experiments, the model also included six constants that identified the block number, accounting for tonic changes in firing rate across blocks. Because not all cues were present in every block, this strategy prevented the cue kernels from being used to explain the baseline changes across blocks. For each kernel to be fit we constructed a Toeplitz predictor matrix of size $T \times l$, in which T is the total number of time bins and l is the number of lags required for the kernel. The predictor matrix contains diagonal stripes starting each time an event occurs and 0 otherwise. The predictor matrices were horizontally concatenated to yield a global prediction matrix $\mathbf{P}$ of size $T \times L$ containing all predictor kernels. Rate vectors of all N neurons were horizontally concatenated to form $\mathbf{F}$, a $T \times N$ matrix.

Reduced-rank regression

To prevent noisy and overfit kernels we implemented reduced-rank regression (***Steinmetz et al., 2019***), which allows regularized estimation by factorizing the kernel matrix $\mathbf{K}$ into the product of a $L \times r$ matrix $\mathbf{B}$ and a $r \times N$ matrix $\mathbf{W}$, minimizing the total error: $E = ||\mathbf{F} - \mathbf{PBW}||^2$. The $T \times r$ matrix $\mathbf{PB}$ consists of a set of ordered temporal basis functions that can be linearly combined to estimate the neuron's firing rate over the whole training set and which results in the best possible prediction from any rank r matrix. To estimate each neuron's kernel functions we generated the reduced rank predictor matrix $\mathbf{PB}$ for $r = 20$, and estimated the weights $\mathbf{w_n}$ to minimize the squared error $E_n = |\mathbf{f_n} - \mathbf{PBw_n}|^2$ with elastic net regularization (using the MATLAB function 'lassoglm') with parameters $\alpha = 0.5$ and $\lambda = [0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5]$, using fourfold cross-validations to determine the optimal value for λ for each neuron. The kernel functions for neuron n was then unpacked from the L-length vector obtained by multiplying the first $r = 20$ columns of $\mathbf{B}$ by $\mathbf{w_n}$ (i.e. $\mathbf{k_n} = \mathbf{B}_{1:L,1:20}\mathbf{w_n}$). Predictor unique contributions. To assess the importance of each group of kernels for predicting a neuron's activity we first fit the activity of each neuron using the full reduced-rank regression procedure, then fit a reduced model (with fourfold cross-validation), holding out the kernels of the predictor to be tested (cues, licks, or rewards). If the difference in variance explained between the full and held-out model was $> 2\%$, and the total variance explained by the full model was $> 2\%$, the neuron was deemed selective for those predictors (***Steinmetz et al., 2019***). We validated this cutoff by comparing our results when adjusting the cutoff from 0.5–0.5% (***Figure 2—figure supplement 3***). The pattern of results was similar regardless of the cutoff. When we refit the reduced ranks to neural activity with the onset time of each trial shuffled, the 2% cutoff was the smallest that allowed no false positive identification of any neurons uniquely coding any variable (***Figure 2—figure supplement 3B***). Using a higher cutoff led to mislabeling neurons with clear cue responses as non-coding (***Figure 2—figure supplement 3E***).

Cue coding models

To assess cue coding schemes, we fit a new set of models focusing on a more restricted time window (−1–2.5 s from cue onset) using only cues and licks as predictors. Cue and lick neurons were identified as before, and subsequent cue characterization was performed on neurons with only a unique contribution of cues. To identify value coding among cue neurons, we fit a new kernel models with a single cue kernel that scaled according to the cue as well as six block constants (as above) with full rank. We inputted cue values as 1, 0.5, and 0 for each CS+, CS50, and CS−, respectively, ranked according to their reward probability. We fit 152 additional models with alternative configurations of cue value: all permutations of 1, 1, 0.5, 0.5, 0, 0 across the six cues, as well as all permutations of high responses (1) for 6 (we call this the 'untuned' model), 5, 4, 3, 2, or 1 cues, with other cues set to 0. Among the 153 total models, some were more similar to the ranked value model, which we quantified by correlating the six cue values of each alternative model with the ranked model. We termed all models with a correlation greater than 0.5 as 'value-like.' For each neuron, we found the model that best explained its activity. The models, their correlation with value, and the proportion of cue neurons best fit by each model are illustrated in ***Figure 3—figure supplement 1***. To verify the robustness of value coding in the neurons best fit by the ranked value model, we fit each of those neurons with 1000 iterations of the cue value model with shuffled cue order to create a null distribution. The fits of the original value model exceeded the 98th percentile of the null for all value neurons.

History model

For a more nuanced estimation of the value of the cue on each trial, we generated per trial value predictions using the lick linear model (described in section 'Behavioral analysis') with cue type, 10 previous outcomes, and 10 previous cue outcomes as predictors. These values were used to scale the height of cue kernel on each trial and were, on average, 0.05, 0.35, and 0.5 for CS−, CS50, and CS+, respectively, but varied on each trial according to the specific reward history. We compared the performance of this model to all the other cue coding models for value and value-like neurons to find neurons better explained by the history model. For neurons better fit by the history model, we also fit 1000 additional models with shuffled trial values within each cue (disrupting the trial history effect). Neurons exceeding the 95th percentile of this null distribution were deemed history neurons.

Coding stability

In the calcium imaging experiments, we used a number of approaches to assess the stability of neural codings. First, for neurons tracked across the first three sessions of odor set A, we took the trial-averaged activity of a given neuron for CS+, CS50, and CS− trials in one session and correlated it with the same neuron's trial-averaged activity from the subsequent session. We quantified the strength of the correlation as its percentile among correlations between that neuron in the first sessions and every other neuron on the subsequent session and compared this value to shuffle control (neuron identity shuffled) (***Figure 7E*** and ***Figure 7—figure supplement 1A***). To assess coding stability of these neurons, we used the kernels resulting from fitting the full model on day 3 and the predictors from each session third to predict neural activity at those time points. We assessed the accuracy of the prediction by correlating it with the true activity versus the correlation with the trial-shuffled data and present this data in original form (***Figure 7—figure supplement 1B***) and normalize to model performance when predicting the (cross-validated) data from the entire training session (***Figure 7G***). This shuffle maintains the temporal dynamics of each trial but shuffles the link between predictors on a given trial and the neural activity for that trial; correlation of predictors (like licks) across trials preserves some prediction accuracy even with this shuffle. We also trained models with data from the third day of odor set A training (A3, day 5) and tested on training days A3 and B3 (days 5 and 6). To determine how unique variable contributions (cues, licks, rewards) evolved over times, we fit our kernel regression model independently to each session third (early, middle, late) of sessions 1–3 for neurons tracked across the three odor set A sessions (***Figure 7H***). To assess value coding across the third sessions of odor set A and B (A3 and B3, days 5 and 6) we fit the 153 cues coding models described in Cue coding models to the neurons imaged on separate days (***Figure 8E***), concatenating the data from each odor set and adding a constant for each day to account for day differences found the model with the best fit for each neuron. We also looked at the stability of value-like signals across the three days of odor set A training by identifying CS+− preferring cue cells using half of the trials in session A3 and plotting the activity of those neurons for the remaining A3 trials and all trials from A1 and A2 (***Figure 8H***).

Principal component analysis

To visualize the dominant firing pattern of PL neurons (***Figure 1—figure supplement 2***), and of value and value-like cells (***Figure 4***), irrespective of direction (excitation or inhibition), we performed principal component analysis ('PCA' in MATLAB) on the concatenated PSTHs across all six cues for the neurons of interest, with each neuron's activity normalized by peak modulation so that each neuron's concatenated PSTH peaked at –1 or 1. We then plotted the score of the top components.

Decoding cue identity

We adapted the approach in ***Ottenheimer et al., 2018*** to use single units as well as random selections of pseudo ensembles to decode cue identity (out of six options) (***Figure 3G***). First, we binned the activity of each neuron into 0.25 s bins spanning –0.5–2.5 s from the onset of every cue in the sessions. These bins were labeled as 1–6 corresponding to the 6 different cues. For all decoding, we randomly selected 50 trials of each cue to use (most sessions had 51 of each cue, but a few had only 50). For single unit decoding of cue identity, we used fivefold cross-validation to train a linear discriminant model ('fitcdiscr' in MATLAB) on 80% of the data and tested correct classification of the six cues on the remaining 20%. We plotted the mean ± SEM performance over time for value, value-like, and untuned neurons, and compared their performance using an ANOVA with fixed effects of neuron type and time point and random effect of the session, making pairwise comparisons with Bonferroni correction. For population decoding, we pooled the activity between 1 and 2.5 s from cue onset (a period with stable decoding in the single unit analysis) and randomly selected groups of 1, 5, 10, 25, 50, 75, 100, or 200 value, value-like, or untuned neurons from all regions. We used the same linear discriminant model (with regularization $\gamma = 1$ in 'fitcdiscr') and fivefold cross-validations. We performed 1000 selections of neurons at each level, plotted the mean and standard deviations of classification accuracy across these iterations, and made pairwise comparisons across groups by calculating the number of iterations where the second group was greater or equal to the first; we repeated this one-way test for both directions of all pairs of groups

and used a Bonferroni corrected α. Pattern of results was unchanged when population activity was standardized with PCA. Pattern of results was also unchanged when we trained on one odor set and tested on the other.

Decoding cue value

Data were prepared for population decoding of cue identity, but with cues labeled as 0, 0.5, or 1 for CS−, CS50, and CS+ trials, respectively. Instead of a linear discriminant model, we used a linear model (elastic net, $\alpha = 0.5$) to regress cue value onto the activity of pseudo ensembles of neurons. To balance our model, we used 50 of each cue type for training and tested on 50 held out trials for a cue never seen by the model; this setup thoroughly tested the idea that value is encoded on a linear scale and thus should be able to generalize to a new cue in same value domain. For example to predict the value of 50 CS+ in odor set B trials, we used for training 50 trials of CS+ A, 0 trials of CS+ B, 25 trials of CS50A, 25 trials of CS50B, 25 trials of CS-A, and 25 trials of CS-B, maximizing coverage of the data while maintaining a balanced design. These models produced predicted values for each cue. We plot the predicted value for CS+ and CS− cues on the left in ***Figure 3H***. To convert these predictions to an accuracy score, we labeled values from –0.25–0.25 as CS−, 0.25–0.75 as CS50, and 0.75–1.25 as CS+ (values outside this range were automatically labeled incorrect). We performed this analysis on random groups of 1, 5, 10, 25, 50, 75, 100, or 200 value, value-like, or untuned neurons (***Figure 3H***), as well as random groups of five neurons (with replacement) from each region and 25 neurons (with replacement) from each region group (***Figure 4E***). We compared region decoding to decoding using a baseline window of –0.5–0 s from odor onset using neurons from each region. We performed 1000 selections of neurons at each level, plotted the mean and standard deviation of classification accuracy across these iterations, and made pairwise comparisons for cue identity.

Cue coding dimension

To project population activity onto the coding dimensions separating CS− activity from CS+ and CS50 activity, respectively, we adapted an approach from ***Li et al., 2016***. We first max normalized the odor set A PSTH activity of each neuron to prevent neurons with particularly large z-score values from dominating the dimension. We then defined coding dimensions from the even trials of odor set A. To find the 'consensus' cue-difference coding dimension across the group defined by each neuron's maximal difference across cue responses, we found the 0.5 s bin in the range 0–2.5 s from cue onset with the peak difference between CS− and CS+ activity or CS- and CS50 activity, for each neuron. This comprised a difference vector of length N defining the maximum cue difference coding across the neuron group. This difference vector was then multiplied by the original z-score values of each neuron's peak difference bin to find the values of peak CS+ vs CS− coding; these values were used to transform the data onto a 0 (CS−) to 1 (CS+) relative cue coding scale. To transform population activity onto the CS− to CS+ dimension at each moment in CS+, CS50, and CS− trials, we multiplied the activity of all neurons in each 0.1 s bin of remaining odd odor set A trials (z-score) by the difference vector and used the same 0–1 scale conversion ('same odor set'). We also multiplied the activity of neurons for cues in odor set B by the difference vector ('other odor set'). We repeated the same process for CS− and CS50 activity. To find the angle between the CS+ and CS50 projections, we bootstrapped the vectors that connected baseline activity to peak activity of CS50 and CS+ along the CS-/CS+ and CS−/CS50 axes and found the angle between these vectors. To find population activity along the CS+/CS− dimension at each moment for CS50 trials of various values, we multiplied the activity (z-score) of all neurons in each 0.1 s bin of the CS50 PSTHs (grouped by value estimated from the lick linear model) by the difference vector and used the same conversion to 0–1 scale. To estimate the distribution of values along the CS+/CS− dimension for each CS50 value condition, we bootstrapped (5000 iterations, with replacement) the population projection and took the mean 1–2.5 s from odor onset. We calculated the slope of the activity on CS50 trials by linearly regressing the estimated position of the population on the CS+/CS− dimension against the value from the lick linear model used to group those trials (5000 iterations, with replacement). To compare slopes across cell groups, we generated a p-value by calculating the number of iterations where the second group was greater or equal to the first; we repeated in this one-way test for both directions of all pairs of groups and used a Bonferroni corrected α.

Statistics

All statistical tests were performed in MATLAB (MathWorks). To compare the fraction of neurons of a specific coding type across regions, we fit a generalized linear mixed-effects model ('fitglme' in MATLAB) with logit link function and with fixed effects of intercept and region and a random effect of the session and then found the estimated mean and 95% confidence interval for each region. For pairwise comparisons across regions, we used a specific contrast for each pair of regions ('coefTest' in MATLAB) to find the p-value that these regions differed from each other and used a Bonferroni-corrected *alpha* for significance. To compare the number of anticipatory licks on different trial types, we found the mean number of anticipatory licks for each cue in each session, and then performed a two-way ANOVA with effects of cue and subject and session as our n (***Figure 1C***). To compare the variance explained during each third of the first session, we found the mean value across neurons from each mouse and then performed a one-way ANOVA on those means with mouse as our n (***Figure 6H***). To compare day 3 model performance on true and shuffled data across each time point (***Figure 7F***), we found the mean value across neurons from each mouse at each time point and then performed a two-way ANOVA with main effects of shuffle and time point, with mouse as our n. We then calculated pairwise statistics using 'multcompare' in MATLAB with Bonferroni correction. To compare cue, lick, and reward unique variance at each time point for each cell category (determined on day 3, ***Figure 7G***), we found the mean from the cells in that category in each mouse at each time point and performed a two-way ANOVA with main effects of variable and day, with mouse as our n. We then calculated pairwise statistics using 'multcompare' in MATLAB with Bonferroni correction.

Acknowledgements

Thank you to Vijay Namboodiri and Charles Zhou for assistance with the imaging. Thank you to Noam Roth for the spike sorting quality control metrics and feedback on the decoding analysis. This work was supported by National Institutes of Health grants F32DA053714 (DJO), F31DA053706 (MMH), T32DK007247 (AJB), R37DA032750 (GDS), and P30DA048736 (GDS), a UW Center for the Neurobiology of Addiction, Pain, and Emotion 2-photon pilot project grant (DJO), a Klingenstein-Simons Fellowship in Neuroscience (NAS), and the Pew Biomedical Scholars Program (NAS).

Additional information

Funding

Funder	Grant reference number	Author
National Institute on Drug Abuse	DA053714	David J Ottenheimer
National Institute on Drug Abuse	DA053706	Madelyn M Hjort
National Institute of Diabetes and Digestive and Kidney Diseases	DK007247	Anna J Bowen
National Institute on Drug Abuse	DA032750	Garret D Stuber
National Institute on Drug Abuse	DA048736	Garret D Stuber

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

David J Ottenheimer, Conceptualization, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Madelyn M Hjort, Conceptualization, Data curation, Validation, Investigation, Writing – original draft, Writing – review and editing; Anna J Bowen, Data curation, Software, Validation,

Investigation, Visualization, Writing – original draft, Writing – review and editing; Nicholas A Steinmetz, Conceptualization, Resources, Software, Supervision, Funding acquisition, Methodology, Writing – original draft, Project administration, Writing – review and editing; Garret D Stuber, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs
David J Ottenheimer http://orcid.org/0000-0003-4882-1898
Madelyn M Hjort http://orcid.org/0000-0001-9932-2349
Anna J Bowen http://orcid.org/0000-0002-8911-2572
Nicholas A Steinmetz http://orcid.org/0000-0001-7029-2908
Garret D Stuber http://orcid.org/0000-0003-1730-4855

Ethics
All experimental procedures were performed in strict accordance with protocols 4450-01 and 4461-01 approved by the Animal Care and Use Committee at the University of Washington.

Peer review material
Consensus Public Review: https://doi.org/10.7554/eLife.84604.3.sa1
Author Response: https://doi.org/10.7554/eLife.84604.3.sa2

Additional files

Supplementary files
- Supplementary file 1. Statistics related to ***Figure 2G*** and ***Figure 2—figure supplement 4***. Bonferroni-corrected p-values from region contrast in generalized linear mixed-effects model.
- Supplementary file 2. Statistics related to ***Figure 3***. Top: Bonferroni-corrected p-values from pairwise comparisons between the decoding accuracy of each group of neurons at each time point with their performance at baseline and with the other neuron groups at that time point. Middle, Bottom: Bonferroni-corrected p-values for pairwise comparisons of bootstrapped distributions (1000 samples) of decoding performance using increasing numbers of neurons in each group.
- Supplementary file 3. Statistics related to ***Figure 4***. Top, Middle: Bonferroni-corrected p-values from region contrasts in generalized linear mixed-effects model. Bottom: Bonferroni-corrected p-values for pairwise comparisons of bootstrapped distributions (1000 samples) of decoding performance using value cells from each region.
- Supplementary file 4. Statistics related to ***Figure 5***. Top: Bonferroni-corrected p-values for pairwise comparisons of bootstrapped distributions (5000 samples) of the slope of population activity of each group of neurons across CS50 trials of increasing value. Bottom: Bonferroni-corrected p-values from region contrasts in generalized linear mixed-effects model.
- Supplementary file 5. Statistics related to ***Figure 2G*** and ***Figure 2—figure supplement 4***. Bonferroni-corrected p-values from region contrast in generalized linear mixed-effects model.
- MDAR checklist

Data availability
The data and code for this manuscript are publicly available at https://doi.org/10.5281/zenodo.6686927 (***Ottenheimer et al., 2022***).

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Ottenheimer D, Hjort M, Bowen A, Stuber G, Steinmetz N	2022	Electrophysiology and two-photon imaging data from an olfactory Pavlovian conditioning task	https://doi.org/10.6084/m9.figshare.21365598	figshare, 10.6084/m9.figshare.21365598
Ottenheimer DJ, Steinmetz N	2023	SteinmetzLab/ottenheimer-et-al-2022: Revised manuscript	https://doi.org/10.5281/zenodo.7718542	Zenodo, 10.5281/zenodo.7718542

References

Allen WE, Chen MZ, Pichamoorthy N, Tien RH, Pachitariu M, Luo L, Deisseroth K. 2019. Thirst regulates motivated behavior through modulation of Brainwide neural population Dynamics. *Science* **364**:253. DOI: https://doi.org/10.1126/science.aav3932, PMID: 30948440

Banga,K, Benson,J, Bonacchi N, Bruijns SA, Campbell R, Chapuis GA, Churchland AK, Davatolhagh MF, Lee HD, Faulkner M, International-Brain-Laboratory. 2022. Reproducibility of In-Vivo Electrophysiological Measurements in Mice. [bioRxiv]. DOI: https://doi.org/10.1101/2022.05.09.491042

Bari BA, Grossman CD, Lubin EE, Rajagopalan AE, Cressy JI, Cohen JY. 2019. Stable representations of decision variables for flexible behavior. *Neuron* **103**:922–933.. DOI: https://doi.org/10.1016/j.neuron.2019.06.001, PMID: 31280924

Berridge KC. 2004. Motivation concepts in behavioral Neuroscience. *Physiology & Behavior* **81**:179–209. DOI: https://doi.org/10.1016/j.physbeh.2004.02.004

Bhattarai JP, Etyemez S, Jaaro-Peled H, Janke E, Leon Tolosa UD, Kamiya A, Gottfried JA, Sawa A, Ma M. 2022. Olfactory modulation of the medial Prefrontal cortex circuitry: implications for social cognition. *Seminars in Cell & Developmental Biology* **129**:31–39. DOI: https://doi.org/10.1016/j.semcdb.2021.03.022, PMID: 33975755

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PGF, Kwok SC, Phillips A, Tanaka K. 2009. Dissociable components of rule-guided behavior depend on distinct medial and Prefrontal regions. *Science* **325**:52–58. DOI: https://doi.org/10.1126/science.1172377, PMID: 19574382

Campbell R, Blot A, Rousseau C, Fabbri M, West SJ. 2020. Lasagna. v.1.0. Github. https://github.com/SainsburyWellcomeCentre/lasagna

Chen TW, Li N, Daie K, Svoboda K. 2017. A map of anticipatory activity in mouse motor cortex. *Neuron* **94**:866–879. DOI: https://doi.org/10.1016/j.neuron.2017.05.005, PMID: 28521137

Chu MW, Li WL, Komiyama T. 2016. Balancing the robustness and efficiency of odor representations during learning. *Neuron* **92**:174–186. DOI: https://doi.org/10.1016/j.neuron.2016.09.004, PMID: 27667005

Chudasama Y, Robbins TW. 2003. Dissociable contributions of the Orbitofrontal and Infralimbic cortex to Pavlovian Autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *The Journal of Neuroscience* **23**:8771–8780. DOI: https://doi.org/10.1523/JNEUROSCI.23-25-08771.2003, PMID: 14507977

Dalley JW, Cardinal RN, Robbins TW. 2004. Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews* **28**:771–784. DOI: https://doi.org/10.1016/j.neubiorev.2004.09.006

Dalton GL, Wang NY, Phillips AG, Floresco SB. 2016. Multifaceted contributions by different regions of the Orbitofrontal and medial Prefrontal cortex to probabilistic reversal learning. *The Journal of Neuroscience* **36**:1996–2006. DOI: https://doi.org/10.1523/JNEUROSCI.3366-15.2016, PMID: 26865622

Devore S, Lee J, Linster C. 2013. Odor preferences shape discrimination learning in rats. *Behavioral Neuroscience* **127**:498–504. DOI: https://doi.org/10.1037/a0033329, PMID: 23895061

Doucette W, Gire DH, Whitesell J, Carmean V, Lucero MT, Restrepo D. 2011. Associative cortex features in the first olfactory brain relay station. *Neuron* **69**:1176–1187. DOI: https://doi.org/10.1016/j.neuron.2011.02.024, PMID: 21435561

Ekstrand JJ, Domroese ME, Johnson DM, Feig SL, Knodel SM, Behan M, Haberly LB. 2001. A new subdivision of anterior Piriform cortex and associated deep nucleus with novel features of interest for Olfaction and epilepsy. *The Journal of Comparative Neurology* **434**:289–307. DOI: https://doi.org/10.1002/cne.1178, PMID: 11331530

Ertürk A, Becker K, Jährling N, Mauch CP, Hojer CD, Egen JG, Hellal F, Bradke F, Sheng M, Dodt H-U. 2012. Three-dimensional imaging of solvent-cleared organs using 3Disco. *Nature Protocols* **7**:1983–1995. DOI: https://doi.org/10.1038/nprot.2012.119, PMID: 23060243

Eshel N, Tian J, Bukwich M, Uchida N. 2016. Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience* **19**:479–486. DOI: https://doi.org/10.1038/nn.4239, PMID: 26854803

Faulkner M. 2020. Iblapps. bde9a6e. Github. https://github.com/int-brain-lab/iblapps/tree/master/atlaselectrophysiology

Fiorillo CD, Tobler PN, Schultz W. 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**:1898–1902. DOI: https://doi.org/10.1126/science.1077349, PMID: 12649484

Grant RI, Doncheck EM, Vollmer KM, Winston KT, Romanova EV, Siegler PN, Holman H, Bowen CW, Otis JM. 2021. Specialized coding patterns among Dorsomedial Prefrontal neuronal ensembles predict conditioned reward seeking. *eLife* **10**:e65764. DOI: https://doi.org/10.7554/eLife.65764, PMID: 34184635

Guo ZV, Hires SA, Li N, O'Connor DH, Komiyama T, Ophir E, Huber D, Bonardi C, Morandell K, Gutnisky D, Peron S, Xu N, Cox J, Svoboda K. 2014a. Procedures for behavioral experiments in head-fixed mice. *PLOS ONE* **9**:e88678. DOI: https://doi.org/10.1371/journal.pone.0088678, PMID: 24520413

Guo ZV, Li N, Huber D, Ophir E, Gutnisky D, Ting JT, Feng G, Svoboda K. 2014b. Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**:179–194. DOI: https://doi.org/10.1016/j.neuron.2013.10.020, PMID: 24361077

Hayden BY, Niv Y. 2021. The case against economic values in the Orbitofrontal cortex (or anywhere else in the brain). *Behavioral Neuroscience* **135**:192–201. DOI: https://doi.org/10.1037/bne0000448, PMID: 34060875

Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN. 2016. Circuit-based corticostriatal homologies between rat and primate. *Biological Psychiatry* **80**:509–521. DOI: https://doi.org/10.1016/j.biopsych.2016.05.012, PMID: 27450032

Horst NK, Laubach M. 2013. Reward-related activity in the medial Prefrontal cortex is driven by consumption. *Frontiers in Neuroscience* **7**:56. DOI: https://doi.org/10.3389/fnins.2013.00056, PMID: 23596384

Hunt LT, Malalasekera WMN, de Berker AO, Miranda B, Farmer SF, Behrens TEJ, Kennerley SW. 2018. Triple dissociation of attention and decision computations across Prefrontal cortex. *Nature Neuroscience* **21**:1471–1481. DOI: https://doi.org/10.1038/s41593-018-0239-5, PMID: 30258238

Hyman JM, Ma L, Balaguer-Ballester E, Durstewitz D, Seamans JK. 2012. Contextual Encoding by ensembles of medial Prefrontal cortex neurons. *PNAS* **109**:5086–5091. DOI: https://doi.org/10.1073/pnas.1114415109

Igarashi KM, Ieki N, An M, Yamaguchi Y, Nagayama S, Kobayakawa K, Kobayakawa R, Tanifuji M, Sakano H, Chen WR, Mori K. 2012. Parallel mitral and Tufted cell pathways route distinct odor information to different targets in the olfactory cortex. *The Journal of Neuroscience* **32**:7970–7985. DOI: https://doi.org/10.1523/JNEUROSCI.0154-12.2012, PMID: 22674272

Jackman SL, Chen CH, Chettih SN, Neufeld SQ, Drew IR, Agba CK, Flaquer I, Stefano AN, Kennedy TJ, Belinsky JE, Roberston K, Beron CC, Sabatini BL, Harvey CD, Regehr WG. 2018. Silk Fibroin films facilitate single-step targeted expression of Optogenetic proteins. *Cell Reports* **22**:3351–3361. DOI: https://doi.org/10.1016/j.celrep.2018.02.081, PMID: 29562189

Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, Lee AK, Anastassiou CA, Andrei A, Aydın Ç, Barbic M, Blanche TJ, Bonin V, Couto J, Dutta B, Gratiy SL, Gutnisky DA, Häusser M, Karsh B, Ledochowitsch P, et al. 2017. Fully integrated Silicon probes for high-density recording of neural activity. *Nature* **551**:232–236. DOI: https://doi.org/10.1038/nature24636, PMID: 29120427

Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. 2009. Neurons in the frontal lobe Encode the value of multiple decision variables. *Journal of Cognitive Neuroscience* **21**:1162–1178. DOI: https://doi.org/10.1162/jocn.2009.21100, PMID: 18752411

Kesner RP, Churchwell JC. 2011. An analysis of rat Prefrontal cortex in mediating executive function. *Neurobiology of Learning and Memory* **96**:417–431. DOI: https://doi.org/10.1016/j.nlm.2011.07.002, PMID: 21855643

Klein-Flügge MC, Bongioanni A, Rushworth MFS. 2022. Medial and orbital frontal cortex in decision-making and flexible behavior. *Neuron* **110**:2743–2770. DOI: https://doi.org/10.1016/j.neuron.2022.05.022, PMID: 35705077

Koldaeva A, Schaefer AT, Fukunaga I. 2019. Rapid task-dependent tuning of the mouse olfactory bulb. *eLife* **8**:e43558. DOI: https://doi.org/10.7554/eLife.43558, PMID: 30724732

Kondo M, Matsuzaki M. 2021. Neuronal representations of reward-predicting cues and outcome history with movement in the frontal cortex. *Cell Reports* **34**:108704. DOI: https://doi.org/10.1016/j.celrep.2021.108704, PMID: 33535051

Kuwabara M, Kang N, Holy TE, Padoa-Schioppa C. 2020. Neural mechanisms of economic choices in mice. *eLife* **9**:e49669. DOI: https://doi.org/10.7554/eLife.49669, PMID: 32096761

Laubach M, Amarante LM, Swanson K, White SR. 2018. What, if anything, is rodent Prefrontal cortex *ENeuro* **5**:ENEURO.0315-18.2018. DOI: https://doi.org/10.1523/ENEURO.0315-18.2018, PMID: 30406193

Li A, Gire DH, Restrepo D. 2015. Spike-field coherence in a population of olfactory bulb neurons Differentiates between odors irrespective of associated outcome. *The Journal of Neuroscience* **35**:5808–5822. DOI: https://doi.org/10.1523/JNEUROSCI.4003-14.2015, PMID: 25855190

Li N, Daie K, Svoboda K, Druckmann S. 2016. Robust neuronal Dynamics in Premotor cortex during motor planning. *Nature* **532**:459–464. DOI: https://doi.org/10.1038/nature17643

Liu LD, Chen S, Hou H, West SJ, Faulkner M, Laboratory TIB, Economo MN, Li N, Svoboda K. 2021. Accurate localization of linear probe electrode Arrays across multiple brains. *ENeuro* **8**:ENEURO.0241-21.2021. DOI: https://doi.org/10.1523/ENEURO.0241-21.2021, PMID: 34697075

Malagon-Vina H, Ciocchi S, Passecker J, Dorffner G, Klausberger T. 2018. Fluid network Dynamics in the Prefrontal cortex during multiple strategy switching. *Nature Communications* **9**:309. DOI: https://doi.org/10.1038/s41467-017-02764-x, PMID: 29358717

Miller EK, Cohen JD. 2001. An integrative theory of Prefrontal cortex function. *Annual Review of Neuroscience* **24**:167–202. DOI: https://doi.org/10.1146/annurev.neuro.24.1.167, PMID: 11283309

Mori K, Sakano H. 2021. Olfactory circuitry and behavioral decisions. *Annual Review of Physiology* **83**:231–256. DOI: https://doi.org/10.1146/annurev-physiol-031820-092824, PMID: 33228453

Musall S, Kaufman MT, Juavinett AL, Gluf S, Churchland AK. 2019. Single-trial neural Dynamics are dominated by richly varied movements. *Nature Neuroscience* **22**:1677–1686. DOI: https://doi.org/10.1038/s41593-019-0502-4, PMID: 31551604

Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O. 2004. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**:269–280. DOI: https://doi.org/10.1016/s0896-6273(03)00869-9, PMID: 14741107

Namboodiri VMK, Otis JM, van Heeswijk K, Voets ES, Alghorazi RA, Rodriguez-Romaguera J, Mihalas S, Stuber GD. 2019. Single-cell activity tracking reveals that Orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nature Neuroscience* **22**:1110–1121. DOI: https://doi.org/10.1038/s41593-019-0408-1, PMID: 31160741

Otis JM, Namboodiri VMK, Matan AM, Voets ES, Mohorn EP, Kosyk O, McHenry JA, Robinson JE, Resendez SL, Rossi MA, Stuber GD. 2017. Prefrontal cortex output circuits guide reward seeking through divergent cue Encoding. *Nature* **543**:103–107. DOI: https://doi.org/10.1038/nature21376, PMID: 28225752

Ottenheimer D, Richard JM, Janak PH. 2018. Ventral Pallidum Encodes relative reward value earlier and more Robustly than nucleus accumbens. *Nature Communications* **9**:4350. DOI: https://doi.org/10.1038/s41467-018-06849-z, PMID: 30341305

Ottenheimer DJ, Bari BA, Sutlief E, Fraser KM, Kim TH, Richard JM, Cohen JY, Janak PH. 2020. A quantitative reward prediction error signal in the ventral Pallidum. *Nature Neuroscience* **23**:1267–1276. DOI: https://doi.org/10.1038/s41593-020-0688-5, PMID: 32778791

Ottenheimer DJ, Hjort MM, Bowen AJ, Steinmetz NA, Stuber GD. 2022. Data and code for Ottenheimer et al 2022. Zenodo. https://doi.org/10.5281/zenodo.6686927DOI: https://doi.org/10.5281/zenodo.6686927

Pachitariu M, Steinmetz N, Colonell J, Rossant C, Bondy AG, Winter O, Guzman J, kushbanga, Janke A, Czuba T, Bhagat J. 2023. Kilosort. d55179f. GitHub. https://github.com/MouseLand/Kilosort

Pashkovski SL, Iurilli G, Brann D, Chicharro D, Drummey K, Franks KM, Panzeri S, Datta SR. 2020. Structure and flexibility in cortical representations of odour space. *Nature* **583**:253–258. DOI: https://doi.org/10.1038/s41586-020-2451-1

Price JL. 1985. Beyond the primary olfactory cortex: olfactory-related areas in the neocortex, thalamus and hypothalamus. *Chemical Senses* **10**:239–258. DOI: https://doi.org/10.1093/chemse/10.2.239

Price JL, Carmichael ST, Carnes K, Clugnet M, Kuroda M, Ray J. 1991. *Olfactory Input to the Prefrontal Cortex. Olfaction: A Model System for Computational Neuroscience* Cambridge, MA: MIT Press.

Roesch MR, Olson CR. 2004. Neuronal activity related to reward value and motivation in Primate frontal cortex. *Science* **304**:307–310. DOI: https://doi.org/10.1126/science.1093223, PMID: 15073380

Rossant C, Harris K, Carandini M, International Brain Laboratory. 2021. Phy. v. 2.0. Github. https://github.com/kwikteam/phy

Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MFS. 2008. Frontal cortex Subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience* **28**:13775–13785. DOI: https://doi.org/10.1523/JNEUROSCI.3541-08.2008, PMID: 19091968

Sallet J, Quilodran R, Rothe M, Vezoli J, Joseph J-P, Procyk E. 2007. Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience* **7**:327–336. DOI: https://doi.org/10.3758/CABN.7.4.327

Saraiva LR, Kondoh K, Ye X, Yoon K, Hernandez M, Buck LB. 2016. Combinatorial effects of Odorants on mouse behavior. *PNAS* **113**:E3300–E3306. DOI: https://doi.org/10.1073/pnas.1605973113

Schoenbaum G, Setlow B, Saddoris MP, Gallagher M. 2003. Encoding predicted outcome and acquired value in Orbitofrontal cortex during cue sampling depends upon input from Basolateral amygdala. *Neuron* **39**:855–867. DOI: https://doi.org/10.1016/s0896-6273(03)00474-4, PMID: 12948451

Schoonover CE, Ohashi SN, Axel R, Fink AJP. 2021. Representational drift in primary olfactory cortex. *Nature* **594**:541–546. DOI: https://doi.org/10.1038/s41586-021-03628-7, PMID: 34108681

Shamonin DP, Bron EE, Lelieveldt BPF, Smits M, Klein S, Staring M. 2013. Fast parallel image registration on Cpu and GPU for diagnostic classification of Alzheimer's disease. *Frontiers in Neuroinformatics* **7**:50. DOI: https://doi.org/10.3389/fninf.2013.00050, PMID: 24474917

St Onge JR, Floresco SB. 2010. Prefrontal cortical contribution to risk-based decision making. *Cerebral Cortex* **20**:1816–1828. DOI: https://doi.org/10.1093/cercor/bhp250, PMID: 19892787

Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, Schoenbaum G. 2014. Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nature Communications* **5**:1–13. DOI: https://doi.org/10.1038/ncomms4926, PMID: 24894805

Stalnaker TA, Cooch NK, Schoenbaum G. 2015. What the Orbitofrontal cortex does not do. *Nature Neuroscience* **18**:620–627. DOI: https://doi.org/10.1038/nn.3982, PMID: 25919962

Steinmetz NA, Buetfering C, Lecoq J, Lee CR, Peters AJ, Jacobs EAK, Coen P, Ollerenshaw DR, Valley MT, de Vries SEJ, Garrett M, Zhuang J, Groblewski PA, Manavi S, Miles J, White C, Lee E, Griffin F, Larkin JD, Roll K, et al. 2017. Aberrant cortical activity in multiple Gcamp6-expressing transgenic mouse lines. *ENeuro* **4**:ENEURO.0207-17.2017. DOI: https://doi.org/10.1523/ENEURO.0207-17.2017, PMID: 28932809

Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD. 2019. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**:266–273. DOI: https://doi.org/10.1038/s41586-019-1787-x, PMID: 31776518

Steinmetz NA, Aydin C, Lebedeva A, Okun M, Pachitariu M, Bauza M, Beau M, Bhagat J, Böhm C, Broux M, Chen S, Colonell J, Gardner RJ, Karsh B, Kloosterman F, Kostadinov D, Mora-Lopez C, O'Callaghan J, Park J, Putzeys J, et al. 2021. Neuropixels 2.0: A Miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**:eabf4588. DOI: https://doi.org/10.1126/science.abf4588, PMID: 33859006

Stettler DD, Axel R. 2009. Representations of odor in the Piriform cortex. *Neuron* **63**:854–864. DOI: https://doi.org/10.1016/j.neuron.2009.09.005, PMID: 19778513

Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. 2019. Spontaneous behaviors drive multidimensional, Brainwide activity. *Science* **364**:255. DOI: https://doi.org/10.1126/science.aav7893, PMID: 31000656

Stringer C, Del Grosso NA, Pachitariu M, Ki C, Pierre A, Dincer T, Couto J, Willem de Gee J, Amsalem O, winnubstj, bjeon90. 2023. Suite2P. 9808eca. GitHub. https://github.com/MouseLand/suite2p

Sul JH, Kim H, Huh N, Lee D, Jung MW. 2010. Distinct roles of rodent Orbitofrontal and medial Prefrontal cortex in decision making. *Neuron* **66**:449–460. DOI: https://doi.org/10.1016/j.neuron.2010.03.033, PMID: 20471357

van Duuren E, van der Plasse G, Lankelma J, Joosten RN, Feenstra MGP, Pennartz CMA. 2009. Single-cell and population coding of expected reward probability in the Orbitofrontal cortex of the rat. *The Journal of Neuroscience* **29**:8965–8976. DOI: https://doi.org/10.1523/JNEUROSCI.0005-09.2009, PMID: 19605634

Verharen JPH, den Ouden HEM, Adan RAH, Vanderschuren LJMJ. 2020. Modulation of value-based decision making behavior by Subregions of the rat Prefrontal cortex. *Psychopharmacology* **237**:1267–1280. DOI: https://doi.org/10.1007/s00213-020-05454-7, PMID: 32025777

Wang PY, Boboila C, Chin M, Higashi-Howard A, Shamash P, Wu Z, Stein NP, Abbott LF, Axel R. 2020a. Transient and persistent representations of odor value in Prefrontal cortex. *Neuron* **108**:209–224. DOI: https://doi.org/10.1016/j.neuron.2020.07.033, PMID: 32827456

Wang Q, Ding SL, Li Y, Royall J, Feng D, Lesnar P, Graddis N, Naeemi M, Facer B, Ho A, Dolbeare T, Blanchard B, Dee N, Wakeman W, Hirokawa KE, Szafer A, Sunkin SM, Oh SW, Bernard A, Phillips JW, et al. 2020b. The Allen mouse brain common coordinate framework: a 3D reference Atlas. *Cell* **181**:936–953. DOI: https://doi.org/10.1016/j.cell.2020.04.007, PMID: 32386544

Winkelmeier L, Filosa C, Hartig R, Scheller M, Sack M, Reinwald JR, Becker R, Wolf D, Gerchen MF, Sartorius A, Meyer-Lindenberg A, Weber-Fahr W, Clemm von Hohenberg C, Russo E, Kelsch W. 2022. Striatal Hub of dynamic and stabilized prediction coding in forebrain networks for olfactory reinforcement learning. *Nature Communications* **13**:3305. DOI: https://doi.org/10.1038/s41467-022-30978-1, PMID: 35676281

Zagha E, Erlich JC, Lee S, Lur G, O'Connor DH, Steinmetz NA, Stringer C, Yang H. 2022. The importance of accounting for movement when relating neuronal activity to sensory and cognitive processes. *The Journal of Neuroscience* **42**:1375–1382. DOI: https://doi.org/10.1523/JNEUROSCI.1919-21.2021, PMID: 35027407

Zhou J, Gardner MPH, Schoenbaum G. 2021. Is the core function of Orbitofrontal cortex to signal values or make predictions *Current Opinion in Behavioral Sciences* **41**:1–9. DOI: https://doi.org/10.1016/j.cobeha.2021.02.011, PMID: 33869678