Title

Differential processing of decision information in subregions of rodent medial prefrontal cortex

Authors/Affiliations

Geoffrey W Diehl, A David Redish*
Department of Neuroscience, University of Minnesota, Minneapolis, MN, United States
* Correspondence: redish@umn.edu

Abstract

Decision-making involves multiple cognitive processes requiring different aspects of information about the situation at hand. The rodent medial prefrontal cortex (mPFC) has been hypothesized to be central to these abilities. Functional studies have sought to link specific processes to specific anatomical subregions, but past studies of mPFC have yielded controversial results, leaving the precise nature of mPFC function unclear. To settle this debate, we recorded from the full dorso-ventral extent of mPFC in each of 8 rats, as they performed a complex economic decision task. These data revealed four distinct functional domains within mPFC that closely mirrored anatomically-identified subregions, including novel evidence to divide prelimbic cortex into dorsal and ventral components. We found that dorsal aspects of mPFC (ACC, dPL) were more involved in processing information about active decisions, while ventral aspects (vPL, IL) were more engaged in motivational factors.

Acknowledgments

We thank K.Seeland, C. Boldt, A. Sheehan, and C. Bogner for technical assistance as well as members of the Redish lab for useful discussion. We also thank S. Heilbronner for discussion and guidance regarding prefrontal anatomy, and the G. Buszáki lab for guidance on silicon probe recordings. This work was supported by NIH grants R01-MH112688 and T32-MH115886.

Author Contributions

G.W.D and A.D.R designed the study. G.W.D. collected and analyzed the data. G.W.D and A.D.R wrote the manuscript.

Declaration of Interests

The authors declare no competing interests.

Impact Statement

Diehl and Redish describe how the rodent medial prefrontal cortex is composed of functionally distinct subregions that process economic decision information differentially along a dorso-ventral gradient.

Introduction

The medial prefrontal cortex (mPFC) is considered central to the ability of mammals to succeed in the complex decision situations common to natural behavior (Dalley et al., 2004; Euston et al., 2012; Miller and Cohen, 2001). We engaged broad mPFC functioning using our unique Restaurant Row task: a foraging-based task that integrates sensory discrimination, value assessment, deliberative planning, re-evaluation of choice, and motivational factors (Steiner and Redish, 2014; Sweis et al., 2018a). Critically, Restaurant Row has been shown to engage mPFC at multiple stages of the task, with different causal manipulations of this brain area eliciting measurable changes in decision behavior (Schmidt et al., 2019; Sweis et al., 2018b; Sweis et al., 2018c). We coupled this strategy of engaging diverse mPFC functioning with neural ensemble recordings across the full dorso-ventral extent of mPFC. Using linear silicon probes that spanned nearly 3mm of the medial wall, we simultaneously measured neural activity across the mPFC. This novel behavioral and physiologic approach to mPFC allowed us to identify the relationships between prefrontal activity and a wide array of variables relevant for task performance, to explicitly determine how mPFC firing patterns related to decision processing at multiple well-controlled spatiotemporal moments in the task, and to directly compare these functional properties across the dorso-ventral anatomy of the medial prefrontal cortex.

The mPFC is typically described as consisting of the anterior cingulate cortex (ACC), the prelimbic cortex (PL), and the infralimbic cortex (IL), each of which have been identified with different functional roles (Groenewegen and Uylings, 2000; Heidbreder and Groenewegen, 2003; Hoover and Vertes, 2007; Kolb, 1990; Sesack et al., 1989; Uylings et al., 2003).

Previous studies have linked the anterior cingulate cortex (ACC) to three prominent functions, each associated with selection or evaluation of potential actions. One set of work has highlighted the role of ACC in choosing between a pair of options or general strategies, particularly within the context of shifting to a new strategy or diverting from a “default” option (Brockett et al., 2020; Ito et al., 2015; Karlsson et al., 2012; Lapish et al., 2008; Ma et al., 2016; Tervo et al., 2021). Physiologic recordings of ACC have also been linked to evaluation of the outcome of a recent choice or of potential alternatives that were not taken (Hyman et al., 2017; Hyman et al., 2013; Mashhoori et al., 2018; Sul et al., 2010). Behavioral manipulations of ACC have similarly supported a role in the post-hoc assessment of decision making and use of that feedback to guide future choices (Akam et al., 2021; Dalton et al., 2016; Hart et al., 2020). Finally, a set of studies have identified a role for ACC in cost-benefit analyses, particularly when the cost entails exertion of physical or mental effort (Hillman and Bilkey, 2010; 2012; Hosking et al., 2014; Walton et al., 2003).

In contrast, studies of the prelimbic cortex (PL) have pointed to a role in strategy representations, goal directed responses, and the expression of learned associations. This expression has been observed both for avoidance learning and fear conditioning (Bravo-Rivera et al., 2014; Corcoran and Quirk, 2007; Diehl et al., 2018; Sierra-Mercado et al., 2011; Vidal-Gonzalez et al., 2006). These learned associations have also been implicated within the context of goal directed decision making in which performance relies on acquired knowledge of a task structure (Killcross and Coutureau, 2003; Tran-Tu-Yen et al., 2009). Consistent with this idea, neural recordings from PL have revealed firing patterns associated with reward delivery, stimulus presentation, and specific behavioral response patterns (Burgos-Robles et al., 2013; Hok et al., 2005; Passecker et al., 2019; Pratt and Mizumori, 2001; Takehara-Nishiuchi et al., 2020). Prelimbic cortex has also been implicated in the initial learning of these associations and in

holding information on-line for a short period (Burgos-Robles et al., 2009; da Silva et al., 2020; Gilmartin and McEchron, 2005; Gilmartin et al., 2013; Mukherjee and Caroni, 2018; Ragozzino et al., 1998; Seamans et al., 1995), highlighting an alternative role in working memory. Finally, PL has been found to be important in shifting between strategies and the development of goal directed response patterns (Birrell and Brown, 2000; Killcross and Coutureau, 2003; Powell and Redish, 2016; Ragozzino et al., 1999; Rich and Shapiro, 2007; Rivalan et al., 2011; Seamans *et al.*, 1995; Tran-Tu-Yen *et al.*, 2009).

Infralimbic cortex (IL), has been most studied within the contexts of fear conditioning and addiction. In particular, IL appears involved in extinction behavior and the shift towards new, learned associations that it entails (Do-Monte et al., 2015; Milad and Quirk, 2002; Mukherjee and Caroni, 2018; Vidal-Gonzalez *et al.*, 2006). Interestingly, IL has been linked to extinction of both appetitive and aversive associations (Bravo-Rivera *et al.*, 2014; Capuzzo and Floresco, 2020; Laque et al., 2019; Milad and Quirk, 2002; Peters et al., 2008; Rhodes and Killcross, 2004; Riaz et al., 2019; Sierra-Mercado *et al.*, 2011), suggesting a more generalized role in suppression of past associations or replacement with newer ones. As it relates to decision making, IL has been implicated in the development and execution of habitual actions. Manipulations of the IL subregion prevent the development of goal-independent responding (Coutureau and Killcross, 2003; Killcross and Coutureau, 2003; Smith and Graybiel, 2013; Smith et al., 2012), and neural activity within IL is reminiscent of that in dorso-lateral striatum, a region associated with habitual action (Smith and Graybiel, 2013).

As will be described below, we present direct evidence that there are actually four functionally separable components of the mPFC (ACC, dorsal PL, ventral PL, IL), corresponding roughly to predictions from anatomical connectivity (Berendse et al., 1992; Heidbreder and Groenewegen, 2003; Voorn et al., 2004). Furthermore, these four components process information about different components of the Restaurant Row task, pointing to a coordinated engagement across mPFC subregions during complex decision making. We find the dorsal components more engaged in action selection components of the task, while the ventral components to be more engaged in motivational aspects of the task. Critically, these new understandings of mPFC function were possible because of our integrated approach to the problem --- engaging a wide array of functions via Restaurant Row, combined with simultaneously measuring neural activity across the full medial prefrontal cortex.

Results

Restaurant Row as an economic decision-making task

In order to effectively engage the mPFC during naturalistic decision making, we employed a complex foraging task that requires interaction across multiple different cognitive domains: Restaurant Row (Steiner and Redish, 2014; Sweis *et al.*, 2018a). In the Restaurant Row task (RRow, Fig 1A,B), rats had one hour to forage for their day's food by making a series of decisions about whether or not to wait for delayed rewards at each of four food reward sites (restaurants), each providing a different flavor of food reward. At each restaurant, rats first entered the Offer Zone (OZ) where the frequency of an auditory cue signaled the temporal delay that the rat must wait to earn the food reward (1-30s delays, drawn at random each visit). While the rat was in the OZ, the tone was played at the same frequency, once per second, until rats made a decision to either accept the offer by entering into the Wait Zone (WZ) where they would wait for their reward, or skip the opportunity, move to the Transition Zone (TZ), and advance to the next restaurant where they would receive a new offer. Upon accepting an offer and entering the WZ, the delay began to count down with tones played each second at progressively lower frequencies. The delay countdown/tones continued until either the delay reached zero and the respective reward was delivered (earned), or the rat elected to quit the offer and leave the WZ for the next site. Upon completing the delay and earning a reward, rats were denoted as being in the Reward Zone (RZ) during which time they could consume their reward and linger at the site until they elected to leave for the next restaurant. Note that while the OZ and TZ were spatially defined, the WZ and RZ were spatially coincident and were defined as before (WZ) and after (RZ) reward delivery respectively. In the event that rats rejected an offer, either by leaving the OZ in a skip or by leaving the WZ in a quit, the offer was rescinded and the tones ceased. Effective performance on RRow thus required assessment of the presented delay and the prospective reward that could be earned.

Consistent with previous studies (Schmidt *et al.*, 2019; Steiner and Redish, 2014; Sweis *et al.*, 2018b; Sweis et al., 2018d), we found rats' behavior during RRow to largely reflect the use of economically well-founded behavioral strategies (Fig 1C-G, Fig 1-Supp1). Decisions in the task could be described by a threshold, with rats tending to accept and ultimately earn delay offers below threshold and skip or quit those above (Fig 1C). Importantly, these thresholds were derived from the subject's behavior and not imposed externally. Thresholds varied across rats and restaurant (Fig 1D) but were generally consistent from one day to the next for a given rat at a given restaurant (Fig 1-Supp1A,B; Wilcoxon Signed-rank test of variance in thresholds across restaurants vs across sessions; rank = 2, $p < 0.05$), indicating that they reflected subjective assessments of the worth of different reward flavors. In the OZ, accepting vs skipping an offer was closely tied to its distance from the threshold with progressively better offers accepted at higher rates and worse offers accepted at lower rates (Fig 1E). Likewise, in the WZ quit decisions were progressively more likely for poorer offers of longer delays that were farther above the threshold value (Fig 1F). These results suggest that we can consider the signed distance from threshold as a measure of subjective value (Value = threshold – offered cost). Note that in this equation, the threshold is a function of both rat and restaurant.

Decisions in the OZ were generally made quickly (1-3 sec), with rats making accept decisions faster than skip decisions (Fig 1-Supp1C,D; Wilcoxon Signed-Rank; $z = -6.6$, $p < 0.001$). Among all accept choices, reaction time was inversely related to the subjective value of the presented offer with better offers leading to faster decisions ($r = -0.30$, $p < 0.001$; n = 8 rats and 20 value bins). This relationship however only held for accepted offers of positive value, or subjectively advantageous decisions (Positive Values: r

= -0.31, p <0.01; Negative Values: r = 0.08, p = 0.46; n = 8 rats and 10 value bins). The shorter reaction time for accept choices is consistent with earlier studies on the Restaurant Row task (Sweis *et al.*, 2018c; Sweis *et al.*, 2018d), and suggests that accepting an offer is the faster default response and skipping reflects a slower override process.

Quit decisions in the WZ were more common earlier in the delay period and average time to quit an offer was inversely related to its subjective value (Fig 1-Supp1E,F; r = -0.52, p < 0.001). In the RZ, rats generally lingered at least five seconds after earning their rewards with a skewed distribution including long, relatively infrequent linger bouts (Fig 1-Supp1G). Interestingly, we did not find a significant relationship between linger time and offer value (Fig 1-Supp1H; r = 0.05, p = 0.51) in contrast to previous studies (Sweis *et al.*, 2018a).

Firing of prefrontal cells is related to ongoing behavior

To record large ensembles of precisely localized mPFC units we implanted rats with linear, 64 channel silicon probes directed along the full extent of the medial wall of the prefrontal cortex (n = 13 total probes; Fig1-Supp2A, Table 1). All recordings were taken after rats were well trained on Restaurant Row, highly accustomed to performing the task under recording procedures, and exhibiting stable behavioral performance. Post-hoc histological verification of probe locations confirmed that 11 of the 13 probes were completely contained within the mPFC (ACC, PL, or IL regions). The remaining two probes were in the VO and MO regions and these data were excluded from all analyses. In total, we recorded 3017 single units from the mPFC of 8 rats (104 total sessions; Table 2) and precisely localized each one within a standardized 3-dimensional space (Fig1-Animiation1; Fig1-Supp2D) using a combination of waveform profiles across recording channels (Csicsvari et al., 2003), histological information, and reconstruction of probe movement through the brain (Fig1-Supp2; see Methods). Of these 3017 recorded cells, we identified putative principal cells (n = 2441) and interneurons (n = 576) on the basis of cells' average firing rates and the width of their spike waveform (Fig S3A, B).

As a first assessment of neural activity in mPFC, we computed a variety of basic spiking and waveform characteristics (Fig 1-Supp3C). To then directly evaluate if these basic physiologic properties varied systematically across the mPFC, we arranged cells according to their location along the DV axis. Overall, we found a broad distribution of basic spiking properties and, outside of the division between principal cells and interneurons, no clear groupings emerged. However, a few basic properties varied systematically along the dorso-ventral axis. Overall firing rate was higher in dorsal aspects of mPFC, whereas spiking was more variable in ventral aspects. Though, in all cases there remained large amounts of overlap in the full distribution of mPFC spiking characteristics. Thus, overall basic spiking characteristics throughout the mPFC were highly diverse, with some properties exhibiting significant, though limited, gradation along the dorso-ventral axis.

To evaluate how mPFC activity was related to the behavioral execution of the RRow task, we directly compared the firing of mPFC cells with a series of task relevant behavioral variables. We observed many individual cells with clear relationships between their firing rate and relevant task variables including: the accept/skip decision made in the Offer Zone, the temporal delay of the presented offer, the overall time that had elapsed through the behavioral session, and the subjective rank of the prospective reward (Fig 2A). Importantly, all of the task relevant variables that we analyzed were significantly represented within the mPFC population (Fig 2B,C; Wilcoxon Signed-Rank vs Shuffled Spikes; all z > 23.5, all p < 0.001). Notably, the proportion of cells with significant relation to behavior was not equivalent across all

behavioral variables. Furthermore, representations of task variables were highly distributed, with cells often exhibiting a significant correlation to many different variables (Fig 2D,F). Importantly, while some behavioral variables in RRow were strongly correlated (e.g. OZ choice depends on offer delay and the rat's subjective threshold; Fig 2E), multimodal coding by mPFC cells was clear and obvious even across completely unrelated behavioral dimensions. As an additional control, we performed a pair of stepwise regression procedures to account for correlations between behavioral variables (Fig 2-Supp1). Even when progressively removing the influence of each behavioral variable on mPFC firing, we still observed multimodal coding of behavioral variables across the population. These findings highlight that mPFC is strongly engaged during RRow with neural activity patterns that are relevant to the decision processes of this task, and are consistent with other studies of mPFC function that find mixed selectivity tuning in prefrontal cells (Duncan, 2001; Powell and Redish, 2014; Rigotti et al., 2013; Seamans et al., 2008; Stokes et al., 2013; Wallis et al., 2001).

Four distinct functional subdivisions exist within the mPFC

To directly examine how information is processed across the mPFC, we computed the functional coupling between pairs of simultaneously recorded cells using transfer entropy. Transfer entropy is a measure between a pair of cells that quantifies the degree to which knowing about one cell's activity allows one to describe that of another cell's activity, above and beyond what can be known from the second cell's own past activity (Timme and Lapish, 2018). It therefore provides a metric of the functional communication between pairs of cells and can provide insight into information flow through a network. Furthermore, we leveraged the ability provided by silicon probes to precisely localize recorded cells in the brain in order to specifically examine the relay of information across the prefrontal dorso-ventral axis.

While the firing of some mPFC cell pairs was unrelated (23%), overall, the distribution of transfer entropy was significantly above chance level (Wilcoxon Signed-Rank, $z = 230.1$, $p < 0.001$) with a strong rightward skew, indicating substantive communication between select pairs of cells (Fig 3A). Organizing cell pairs according to their respective anatomical locations provided direct insight into the nature of these high-coupling relationships (Fig 3B). Transfer entropy was elevated along the identity diagonal, indicating increased communication between cells that were located at similar positions. Yet importantly, this increase was not uniform across the full dorso-ventral extent, but rather was characterized by distinct anatomical positions in which transfer entropy was elevated. In other words, functional communication was enriched at a handful of specific locations within the mPFC, indicative of localized processing subdivisions. We quantitatively identified these subdivisions, and the boundaries between them, by averaging the transfer entropy data from bins proximal to the identity diagonal (within 200µm) at each dorso-ventral depth along the medial wall (Fig 3C). This analysis revealed four peaks separated by identifiable valleys between them (data were not unimodal; Hartigan's Dip Test, $D = 0.0291$, $p < 0\ .001$). Thus, examination of communication between simultaneously recorded cells revealed that during this decision-making task, the mPFC was characterized by four distinct functional subunits.

While transfer entropy calculated throughout the full behavioral period revealed distinct subunits within mPFC, RRow is an inherently dynamic task with distinct phases of choice, re-evaluation, reward, and transition (Steiner and Redish, 2014; Sweis *et al.*, 2018a; Sweis *et al.*, 2018b; Sweis *et al.*, 2018c; Sweis *et al.*, 2018d). Thus, we repeated our transfer entropy analysis, taking spiking data from each of the four task phases independently (Fig 3-Supp1). Notably, data from each task phase produced comparable

transfer entropy profiles with comparable subregional boundaries. These similarities indicate that the functional subregions are consistent throughout the RRow task structure and strongly suggest that they reflect a generalized processing architecture within mPFC.

Importantly, these subdivisions that were identified purely functionally largely matched the stereotaxically-defined subdivisions identified in previous anatomical studies, as outlined in the Paxinos and Watson atlas (Fig 3D) (Paxinos and Watson, 2007). Functionally defined boundaries at the dorsal and ventral end of the mPFC closely aligned with the anatomically defined boundaries between ACC and PL dorsally and between PL and IL ventrally. Accordingly, we will designate our most dorsal and ventral functional subunits as ACC and IL respectively. However, our functional analysis also revealed a transition in the intermediate range of the mPFC, roughly bisecting the anatomical PL subregion. This finding implies that prelimbic cortex is not a functionally homogenous area, but rather should be further split into dorsal (dPL) and ventral (vPL) subregions, a premise that has some historical anatomical support (Berendse *et al.*, 1992; Heidbreder and Groenewegen, 2003; Vogt and Paxinos, 2014; Voorn *et al.*, 2004). While this anatomical separation has not yet been widely adopted, particularly in functional studies, our data strongly suggest that these regions need to be considered separately.

Having identified the existence of four distinct functional subregions within mPFC (ACC, dPL, vPL, IL), we examined the degree of information transfer between these four areas during the full behavioral session (Fig 3F). In line with the fact that subregions were originally identified as areas with elevated communication, transfer entropy between cell pairs was largest when both cells were located within the same subregion (intra-regional communication). Functional communication between cell pairs that spanned different subregions (inter-regional communication) was significantly weaker (Mann-Whitney U test, Intra vs Inter region; $z = 19.87$, $p < 0.001$). Interestingly, functional connectivity varied as a function of distance between subregions, with adjacent areas exhibiting stronger interactions than more distally located areas (ANOVA grouped by Subregion Step; $F(3) = 179.5$, $p < 0.001$; HSD post-hoc: Step 0 vs 1, $p < 0.001$, Step 1 vs 2, $p < 0.001$, Step 2 vs 3, $p = 0.76$; correlation between pairwise transfer entropy and subregion steps: $r = -0.07$, $p < 0.001$). We also examined the neural representation of our array of task relevant variables in each subregion independently (Fig 2-Supp2). Importantly, we found each behavioral variable to be represented across all four areas, though the overall degree of this coding varied between subregions, suggesting differential processing of task relevant information between them. Finally, because our data set contained comparatively fewer ACC cells relative to the other subregions, we performed a bootstrapping procedure in which we repeated our analyses after down-sampling the data from the dPL, vPL, and IL subregions to match that of ACC. Importantly, even with a subsampled data set, we identified subregional boundaries and strength of inter-regional communication levels that matched our main findings (Fig 3-Supp2).

Generalized task variables are differentially represented across mPFC subregions

Although transfer entropy measures revealed the existence of four distinct functional subregions, it only measures the extent of information transfer, and does not provide any insight into what information is represented within these brain areas as decisions are made. Mirroring natural conditions, decision making in the RRow task is not performed as isolated events but rather in a broader task context: The specific nature of decision that must be made varies according to the current behavioral phase in the task (task phase). Rats are forced to operate under a global time constraint both achieving satiety and reducing their overall opportunity over the course of the session (session time). Economic decisions in the Offer Zone are made according to the subjective quality of the currently available reward (reward

rank) and the offer presented to the rat (offer delay). Importantly, each of these broader task components were encoded across all four mPFC subdivisions, though to varying degrees (Fig 4A; Wilcoxon Signed-Rank: Task Phase, $z = 47.5$, $p < 0.001$; Session Time, $z = 45.4$, $p < 0.001$; Reward Rank, $z = 41.9$, $p < 0.001$; Offer Delay, $z = 35.5$, $p < 0.001$; OZ Choice, $z = 44.0$, $p < 0.001$; ANOVA across variables; $F(4) = 925.1$, $p < 0.001$). For each variable, we evaluated the degree of neural coding by calculating the mutual information between the firing rates of mPFC cells and the behavior of interest. Briefly, mutual information quantifies the degree of information about a particular behavior of interest that can be extracted from neural spike rates and describes how well neural spiking can differentiate between behaviors (Timme and Lapish, 2018).

Success in the RRow task requires an understanding of the current phase of the task and its objectives, and an ability to employ an appropriate behavioral strategy given that understanding. In line with a role of the mPFC in executive functioning, we found strong coding for task phase (OZ, WZ, RZ, or TZ) in each of the four prefrontal subregions (Wilcoxon Signed-Rank: all $z > 7.4$, all $p < 0.001$). However, mutual information between cell firing rate and task phase was not equivalent across subregions (Fig 4B, C; ANOVA across subregions; $F(3) = 39.3$, $p < 0.001$). Instead, this information was most strongly represented in ACC and dPL, weaker in vPL, and then elevated again in the ventral-most IL subregion (ANOVA HSD post-hoc: all pairwise comparisons $p < 0.001$). In general, average neuronal firing rates were highest while rats were in the OZ making their decisions, and then tended to decrease during the WZ and the RZ. These variations in firing rate across task phase were strongest in the ACC subregion, as would be expected based on mutual information calculations

At the broader task level, RRow also requires an understanding of the progression through the behavioral session. Because the task design limits rats to only one hour to earn rewards, progression through the session is inversely related to the overall opportunity remaining. In parallel, progression through the session is intertwined with other variables likely to influence value-based decisions such as satiety, fatigue, and general motivation. Again, we found significant coding for elapsed time into the session across all four mPFC subregion (Wilcoxon Signed-Rank: all $z > 5.9$, all $p < 0.001$). However, in contrast to task phase, neural coding for session time was strongest in the ventral components of mPFC, vPL and IL (Fig 4D, E; ANOVA across subregions; $F(3) = 21.4$, $p < 0.001$; HSD post-hoc: ACC vs dPL, $p = 0.24$; vPL vs IL, $p = 0.66$; all other comparisons $p < 0.001$). Interestingly, on average the firing rates of cells in vPL and IL were strongly elevated in the first third of the session and then decreased and remained stable for the remaining duration. In contrast, firing rates in ACC were initially low at the beginning of the session, and then increased over the initial 10 minutes. In light of these changes in mPFC activity over the course of the session, we revisited our main analyses of mPFC functioning across subregions to determine the degree to which they varied as motivational factors evolved over the session (Fig 4-Supp1).

Lastly, RRow decision making involves evaluating a presented offer within the context of the subjective value of a prospective reward option. Because the location of each reward flavor remained static throughout the entire behavioral sequence, rats could easily predict the potential reward at each restaurant visit and appropriately evaluate its subjective worth. Neural coding for this information about prospective rewards was present in each mPFC subregion (Fig 4F,G; Wilcoxon Signed-Rank: all $z > 7.0$, $p < 0.001$). Yet across the four subregions mutual information about reward ranking was most robust in the dorsal areas (ACC and dPL) and weaker in the ventral areas (vPL and IL) (ANOVA across subregions; $F(3) = 45.4$, $p < 0.001$; HSD post-hoc: ACC vs dPL, $p = 0.29$; vPL vs IL, $p < 0.05$; all other comparisons $p <$

0.001). Interestingly, examining raw firing rate differences across reward ranks revealed that in ACC and dPL, firing rates tended to be lower for the more preferred reward options.

In conjunction with assessing the subjective quality of their potential reward, rats must evaluate the delay of the offer presented to them in making their choice. We observed significant coding of offer delay throughout the mPFC (Fig 4H; Wilcoxon Signed-Rank: all $z > 6.9$, all $p < 0.001$). Yet by nature, accept/skip decisions on RRow should be inexorably linked to the offer delay, with rats that behave economically largely accepting short delays and skipping long delays (Fig 1B, 2E). Accordingly, we also found strong coding of offer zone choice across the mPFC (Fig 4H; Wilcoxon Signed-Rank: all $z > 7.3$, all $p < 0.001$) which was strongest in the dorsal subregions (ANOVA across subregions; $F(3) = 134.0$, $p < 0.001$; HSD post-hoc: vPL vs IL, $p < 0.05$; all other comparisons $p < 0.001$). To distinguish which of these related variables was more closely tied to mPFC firing we split our data according to OZ choice (Accept or Skip) and recalculated the MI between firing rate and offer delay (Fig 4I). Interestingly, coding for offer delay was largely eliminated when accounting for OZ choice (Accept: ACC $z = -0.2$, $p = 0.79$; dPL, vPL, IL all $z > 9.0$, all $p < 0.001$; Skip: ACC, dPL, IL all $Z < 0.5$, all $p > 0.05$; vPL $z = 3.4$, $p < 0.001$), suggesting that neural responses were more tightly associated with rats' decision making in the Offer Zone than the offer delay that had been presented to them.

Integrating our findings across these behavioral variables implies a biasing of functional information across the mPFC. Representations of task phase, reward ranking, and OZ choice were stronger in dorsal portions (ACC and dPL) of mPFC, while session time was more strongly represented in ventral portions of mPFC (vPL and IL). Both task phase and reward ranking are important components of successful decision-making processing in Restaurant Row. Decisions in the OZ, WZ, RZ, and TZ all encompass different considerations that must be accounted for when making decisions, but reward preferences play a role in each. Our data suggest a preferential involvement of dorsal mPFC in these processes. In contrast, the increased representations of progress through the session in ventral portions of mPFC (vPL and IL) suggest that more generalized and prolonged motivational aspects preferentially engage the ventral components of mPFC.

Activity in mPFC strongly encodes the identity of the upcoming choice

In the RRow task, value-based decisions occur predominantly within the Offer Zone (Fig 1G,H). Therefore, examination of the firing rate profiles of mPFC cells while rats are in the OZ can provide insight into how neural processing differs when rats accept a presented offer vs when they elect to skip the offer and transition to the next restaurant. We directly measured these differences in neural activity by computing the mutual information between cell firing rates over the decision period (time normalized across visits to equate durations) and the choice that was ultimately made. Across all four subregions, coding for the eventual choice was initially at chance level as rats approached the OZ, prior to receiving information about the delay to earn reward (Fig 5A). However, shortly after entering the OZ, coding for the eventual choice increased above chance and remained elevated through the end of the decision period (Repeated measures ANOVA for subregions over time bins; Interaction: $F(75) = 87.9$, $p < 0.001$. Post-hoc on each time bin via one sided Wilcoxon Signed-rank vs 0 corrected for multiple comparisons). Importantly, this choice coding emerged strongest within the ACC population where mutual information for choice peaked just prior to rats exiting the OZ and registering their choice, although it was also quite prominent in dPL and was detectable in vPL and IL (ANOVA grouped by subregion; $F(3) = 144.4$, $p < 0.001$; HSD Post-hoc: ACC vs others, all $p < 0.001$). Directly examining the net firing rate of mPFC cells between accept decisions and skip decisions pointed to important differences

across the subregions (Fig 5B). Net changes in FR were dramatically larger in ACC and dPL as compared to ventral mPFC (vPL and IL), a result consistent with the differences in magnitude of mutual information. Interestingly, net firing rates also revealed that ACC cells tended to be more responsive during skip decisions (68% net skip; X^2 test for equal proportions vs 50%, $X^2 = 4.9$, $p < 0.05$) whereas dPL cell activity was more biased towards accept decisions (55% net accept; $X^2 = 4.8$, $p < 0.05$). In the ventral areas, vPL was again biased towards skip decisions (55% net skip; $X^2 = 6.2$, $p < 0.05$), but the preferred responses across the IL population was balanced between accept and skip choices (50% net skip; $X^2 = 0.0$, $p = 0.93$). These features of net firing rate highlight important differences between prefrontal subregions, and in the case of ACC fits with a body of literature pointing towards a role in enabling the agent to abandon an available reward in favor of other options (skips in this task) (Hayden et al., 2011; MacDonald et al., 2000; Shenhav et al., 2013; Tervo *et al.*, 2021).

In the RRow task, rats register their decision, accept or skip, via a motor action either into the Wait Zone or the Transition Zone. Thus, cognitive decisions and motor actions are inherently intertwined. In order to dissociate between decision making and motor execution, we leveraged the fact that accept decisions in RRow always involve a left turn whereas skip decisions are always a right turn. Accordingly, neurons in the left hemisphere will be ipsilateral to the motor output of an accept action and contralateral to that of a skip. Past work in the cingulate cortex and related premotor structures suggests that directed motor responses preferentially engage contralateral neural systems (Erlich et al., 2011). Therefore, if our neural responses were tied to motor actions, we would expect that left hemisphere cells should be preferentially engaged during skip decisions (contralateral actions) whereas right hemisphere cells should be engaged during accept decisions. To test this, we segregated our recordings according to hemisphere (Fig 5-Supp1). Critically, we found comparable neural responses from both left and right hemispheres with no strong bias towards preferential firing towards a contralateral motor action. This then suggests that our observed neural changes throughout mPFC are more likely related to decision making than to motor outputs.

Optimal decision making on Restaurant Row requires that rats determine the subjective value of a presented offer by integrating information about identity of the current restaurant and the offer delay. Thus, we examined the degree to which current restaurant and offer delay were individually represented within the mPFC. Consistent with our prior analysis across the full task period, coding for the identity of the current restaurant was abundant across all four mPFC subregions while rats were in the OZ, although they differed in overall magnitude (Fig 5C; Rm ANOVA ; Accept: $F(75) = 7.0$, $p < 0.001$; Skip: $F(75) = 2.6$, $p < 0.001$; Each time bin via one sided Wilcoxon Signed-rank vs 0 corrected for multiple comparisons; ANOVA grouped by subregion; Accept: $F(3) = 45.6$, $p < 0.001$; Skip: $F(3) = 49.5$, $p < 0.001$). This was the case for both accept and skip decisions and throughout the full decision period, indicating that rats maintained an understanding of their current position within the task regardless of what offer they received or how they responded to it. In stark contrast, neural coding for the delay offered (cost) was absent across all four mPFC subregions (Fig 5D; Rm ANOVA; Accept: $F(75) = 2.1$, $p < 0.001$; Skip: $F(75) = 1.3$, $p < 0.05$; Each time bin via one sided Wilcoxon Signed-rank vs 0 corrected for multiple comparisons, no time bins reached significance, all $p > 0.05$; ANOVA by subregion; Accept: $F(3) = 2.2$, $p = 0.09$; Skip: $F(3) = 1.2$, $p = 0.31$). To verify that any conjunctive representation of reward rank was not obscuring our ability to detect delay coding in mPFC we repeated this analysis independently for each restaurant rank (Fig 5-Supp2). Even accounting for potential multiplexed coding, we did not detect significant representations of delay during the OZ. Fascinatingly, although mPFC (particularly ACC and

dPL) encoded the decision made, the firing rates of mPFC cells never consistently represented the quality of the offer that was being considered. This finding suggests that effective decision making in the Offer Zone cannot involve mPFC operating in isolation, but rather must require active communication with other brain area(s) that do contain the necessary information about what delay has been presented.

One potential consideration in our reported results is the lower sampling of ACC cells relative to other subregions. However, we note that our findings of ACC involvement in decision making is consistent with the broader literature (Hyman et al., 2012; Hyman *et al.*, 2013; Ito *et al.*, 2015; Lapish *et al.*, 2008; Sul *et al.*, 2010) and that the bias that we observe for ACC engagement during skip decisions is in line with a proposed role of this structure in bypassing prospective rewards and searching out alternatives (Brockett *et al.*, 2020; Hayden *et al.*, 2011; MacDonald *et al.*, 2000; Mashhoori *et al.*, 2018; Shenhav *et al.*, 2013; Tervo *et al.*, 2021). To directly test the potential consequences of limited sampling on our reported results we repeated all analyses after subsampling our recordings of dPL, vPL, and IL to match those of ACC (Fig 5-Supp3). Importantly, the qualitative nature of all results held, though predictably the statistical power was reduced. Therefore, while sampling bias from ACC and insufficient statistical power remain possibilities, our findings of mPFC responding in the OZ are largely robust.

Re-Evaluative decisions are represented in the mPFC

While these rats made most of their decisions in the Offer Zone, after entering the Wait Zone rats are free to re-evaluate their decision and quit the offer countdown. Furthermore, multiple manipulation studies have found differential effects on decision making in the OZ vs in the WZ (Sweis *et al.*, 2018b; Sweis *et al.*, 2018c), presenting a unique opportunity in our task to examine a distinct secondary decision process. We directly measured the neural coding for the decision in the WZ (either quit or earn) by computing the mutual information between neural firing rates and the eventual choice. At the beginning of the WZ, there was no coding for choice across any of the four subregions (Fig 6A; Repeated measures ANOVA for subregions over time bins; Interaction: $F(207) = 1.3$, $p < 0.01$; Post-hoc on each time bin via one sided Wilcoxon Signed-rank vs 0 corrected for multiple comparisons, no time bins reached significance, all $p > 0.05$). However, examining neural activity at the end of the WZ, just before either the quit decision was executed and the rat left the WZ or the delay ended and food was delivered, we found a significant increase in mutual information for the re-evaluative choice made (Fig 6B; Rm ANOVA; Interaction: $F(207) = 4.4$, $p < 0.001$; Each time bin via one sided Wilcoxon Signed-rank vs 0 corrected for multiple comparisons). While a significant increase occurred across all four subregions, the magnitude of this increase was stronger in dorsal mPFC (ACC and dPL) than in ventral mPFC (vPL and IL) (ANOVA grouped by subregion; $F(3) = 17.3$, $p < 0.001$; Mann-Whitney U test, dmPFC (ACC & dPL) vs vmPFC (vPL & IL); $z = 4.2$, $p < 0.001$). Furthermore, the increase occurred earliest and persisted for longest in the dPL subregion, suggesting particularly strong coding for re-evaluative decision making in dPL. Interestingly, whereas ACC activity in the OZ coded the primary decision more strongly than was detected in other prefrontal areas, this was not the case for re-evaluation decisions in the WZ where the magnitude of MI change was comparable between ACC and dPL (ANOVA HSD post-hoc ACC vs dPL, $p = 0.93$). Although we sampled fewer ACC cells, down-sampling of data from the other subregions yielded qualitatively similar results to those from our full data set (Fig 6-Supp1). It is therefore likely that while we were unable to detect significant responses in ACC early in the decision process, the qualitative response pattern we observed is an accurate representation of the subregion's activity pattern.

Collectively, these results indicate that while coding for re-evaluative decisions was apparent across the full mPFC, it exhibited a gradient behavior with more prominent responses observed in the dorsal areas.

In the RRow Offer Zone, choices are registered based on a motor output either to the Wait Zone (Accept) or the Transition Zone (Skip). In the Wait Zone, the decision entails either passively remaining at the location (Earn) or actively leaving the site (Quit). To determine if our neural response in the WZ may be more indicative of a motor response, we computed the net firing difference between earn and quit events for each cell (Fig 6-Supp2). Importantly, we found cells with firing rate biases for both options, as would be expected if the decision process were driving the neural activity. Additionally, we found that ACC cells trended toward being more likely to increase their firing rate for a quit decision, or one in which rats bypassed available food for an alternative (64% net Quit; $X^2 = 2.9$, $p = 0.08$). This mirrors both our findings in the OZ (Fig 5B), and past work on ACC (Hayden *et al.*, 2011; Mashhoori *et al.*, 2018; Tervo *et al.*, 2021).

As in the OZ, decision making in the WZ relies on the evaluation of the delay to be waited and a determination of whether or not that time expenditure is advantageous. In the OZ this information was surprisingly absent from representations in mPFC (Fig 5D), pointing to the need for inter-regional communication. However, after rats entered the WZ, firing rates in dPL, vPL, and IL all varied systematically according to the offer delay that had been accepted, with higher firing rates associated with longer delays (Fig 6C, D). Notably, no relationship existed for the ACC subregion, the area with the strongest coding for the initial choice (ACC: $p > 0.62$ at all time points). While it remains possible that our limited sampling of ACC prevented detection of a relationship between firing rate and delay, or that sampling biases might have skewed our findings, down-sampling our recordings from dPL, vPL, and IL to a comparable size did not eliminate the effects seen elsewhere (Fig 6-Supp1). Consistent with our mutual information results, no significant correlation existed between average firing rate and delay in the 1s period prior to WZ entry while rats were still in the OZ (Pre WZ: all $p > 0.19$). However, at both the beginning and the end of the WZ there was a significant correlation between firing rate and delay, indicative of coding for the value of the accepted offer (Early WZ: dPL, vPL, IL all $p < 0.01$; Late WZ: dPL $p = 0.06$, vPL and IL $p < 0.01$). Notably, this coding for offer value persisted past the end of the delay period (Post WZ: dPL and vPL $p < 0.05$, IL $p = 0.06$) indicating that although value coding was not present within mPFC while the initial OZ decision was being made, it strongly developed across multiple subregions during the re-evaluation period.

Responses to reward delivery differ across mPFC subregions

After accepting the offer and waiting out the full delay, rats in RRow earn a food reward. This event provides both a highly salient natural reward, and also serves as a learning signal for the rats to determine if their past decisions were a worthwhile use of the time expended. We examined mPFC activity patterns at the beginning of the Reward Zone as reward was delivered and observed qualitatively different neural responding across the four subregions (Fig 7A,B, Fig 7-Supp1; ANOVA grouped by subregion; $F(3) = 13.1$, $p < 0.001$). ACC activity was characterized by a strong decrease in firing rate that persisted for several seconds. Across the remaining three subregions the responses were heterogenous across the populations in each area, but the average responding varied from dorsal to ventral with dPL exhibiting a net increase, IL a net decrease, and vPL falling into a middle ground with balanced increases and decreases in firing rate. Interestingly, responding to reward delivery tended to occur sooner in the ventral portions of mPFC, though across the population there was considerable

variation (Fig 7C; ANOVA grouped by subregion; $F(3) = 10.5$, $p < 0.001$; Mann-Whitney U test, dmPFC (ACC & dPL) vs vmPFC (vPL & IL); $z = 4.8$, $p < 0.001$).

After earning a reward, rats then must advance to the next restaurant in order to receive a new offer and earn more food. However, this advancement is un-cued and entirely at the volition of the rat. Interestingly, rats tended to spend considerable portions of their total time lingering in the RZ, consuming their earned food and waiting to proceed to the next site (Fig 1-Supp1G), consistent with prior observations of rats and mice on this task (Schmidt and Redish, 2021; Steiner and Redish, 2014; Sweis *et al.*, 2018a; Sweis *et al.*, 2018c; Sweis *et al.*, 2018d). Because this lingering time is self-directed and presents a limiting factor to earning more food, we hypothesized that neural activity during this time may provide additional insight into subjective decision making. Indeed, we found that firing rates within mPFC subregions systematically varied according to the lingering behavior of the rat (Fig 7D, E).

Interestingly, when examining the first half of the lingering period, we found significant relationships in the ventral portions of mPFC, vPL and IL, between firing rate and eventual linger time (Pre RZ: all $p > 0.66$; Early RZ: vPL $p < 0.05$, IL $p = 0.07$). However, in the second half of the lingering period this relationship between firing rate and lingering time disappeared in the ventral mPFC and an opposing relationship developed in the dPL subregion (Late RZ: dPL $p < 0.01$; Post RZ: all $p > 0.39$). While we did not observe any systematic effect in the ACC subregion, this may be attributable to the limited sampling or potential biases within our data set. Notably, relationships between firing rate and linger time in the other three subregions were obscured when down-sampling the total population (Fig 7-Supp1). Together, these findings suggest the potential for multiple cognitive processes during the lingering period and a gradual shift of neural engagement along the dorso-ventral gradient of the medial wall. During the early portion of the lingering time, when rats were presumably processing the fact that they had received food of a specific flavor, ventral portions of mPFC were more strongly engaged, potentially related to motivational or learning factors. During the later portion of the lingering time, when rats were considering moving on to the next foraging site, dorsal portions were more strongly engaged, potentially related to the more acute decision of specifically taking the action to transition to the next restaurant.

Discussion

Clear functional distinctions exist along the medial wall of the prefrontal cortex. Our results identified four subregional units that closely match well-established anatomical designations, thus presenting a strong correspondence between the functional and anatomical understanding of the mPFC. Examination of the activity of each of these subregions during complex decision-making further revealed subtle but significant differences in their information processing and pointed to a gradient of function across the prefrontal cortex. While each prefrontal subregion was engaged across a wide array of functional components of decision making, a biasing emerged between the dorsal and ventral ends of prefrontal cortex. Neural activity in dorsal portions of the mPFC was more strongly engaged during acute decision making, both in the Offer Zone where initial value-based choices were made, and in the Wait Zone where choices were re-evaluated. In conjunction, ventral portions of mPFC were more strongly associated with variables that might be linked to more motivational factors such as time into the behavioral session, and how long to linger after receiving food before electing to pursue the next reward.

Prefrontal functional divisions mirror anatomical divisions, with an added segregation of the prelimbic cortex

The medial prefrontal cortex has long been attributed a role in a diverse array of functions (Dalley *et al.*, 2004; Euston *et al.*, 2012; Miller and Cohen, 2001) and from an anatomical standpoint it has long been known that different portions of the medial wall have differing connectivity profiles (Groenewegen and Uylings, 2000; Kolb, 1990; Uylings *et al.*, 2003). While these anatomical differences imply that the ACC, PL, and IL regions likely also differ functionally, no single study had yet fully addressed this question, meaning that to draw any conclusions about functional differences between mPFC subregions requires integration across different experimental approaches (Carlén, 2017; Kesner and Churchwell, 2011; Laubach et al., 2018; McLaughlin et al., 2021; Miller and Cohen, 2001). Furthermore, any direct correspondence between functional and anatomical domains has thus far remained elusive. Using the precise anatomical localization afforded by silicon probes and the diverse engagement of mPFC provided by Restaurant Row, we discovered that functional communication within the mPFC occurs in a localized pattern consisting of four distinct processing subregions. Furthermore, the anatomical boundaries between these processing hubs closely match anatomically defined subregions, highlighting a tight coupling between the functional and anatomical domains.

However, the emergent story is more complex than the standard anatomical understanding of mPFC as composed of anterior cingulate, prelimbic, and infralimbic cortices. In addition to the boundaries between these three subregions, we also report evidence for functional division of the prelimbic cortex into dorsal and ventral subregions. While there is anatomical data supporting a division of PL into dorsal and ventral parts (Berendse *et al.*, 1992; Heidbreder and Groenewegen, 2003; Vogt and Paxinos, 2014; Voorn *et al.*, 2004), our work shows that they process information differently as well.

Distributed vs Discrete vs Gradient: Functional activity in the prefrontal cortex

What is the best way to characterize functional activity throughout the mPFC? Some reports of prefrontal activity have described distributed representations where many cells across the cortex are all engaged in a given functional process, or in turn where the activity of each cell multiplexes information across a large set of multiple processes (Duncan, 2001; Rigotti *et al.*, 2013; Seamans *et al.*, 2008; Stokes *et al.*, 2013; Wallis *et al.*, 2001). In some aspects of mPFC function we find a similar picture (Fig 8). For each behavioral variable we examined, we found it to be represented by a distributed population of cells spread across all four prefrontal subregion (Fig 2). Each individual cell also responded to multiple different variables as would be expected in a distributed coding scheme.

Alternatively, one can think of discrete units of functional processing within mPFC with each component primarily involved in a particular function. The idea of distinct component of mPFC is strongly rooted in anatomical work that points to the prefrontal areas of anterior cingulate, prelimbic (dorsal and ventral), and infralimbic cortices (Groenewegen and Uylings, 2000; Heidbreder and Groenewegen, 2003; Hoover and Vertes, 2007; Kolb, 1990; Sesack *et al.*, 1989; Uylings *et al.*, 2003). In parallel, many functional studies, particularly those that seek to identify functional dissociations between subregions, have worked to attribute distinct cognitive processes to each prefrontal subregion (Burgos-Robles *et al.*, 2013; Capuzzo and Floresco, 2020; Dalton *et al.*, 2016; Killcross and Coutureau, 2003; Mukherjee and Caroni, 2018; Ragozzino *et al.*, 1998; Sierra-Mercado *et al.*, 2011; Vidal-Gonzalez *et al.*, 2006). Notably, our measurement of functional communication within the mPFC using transfer entropy points to the

existence of four distinct processing units (Fig 3). Thus, in some respects mPFC appears to be composed of distinct functional subunits (Fig 8).

Finally, activity within the prefrontal cortex could be thought of as a gradient of functional processing. Indeed, many aspects of our data point to this perspective. Dorsal aspects of mPFC (ACC & dPL) were strongly engaged while rats were making their decisions (Fig 5, 6); however ventral aspects (vPL & IL) also represented these same decision components, but to a lesser degree. In parallel, neural activity in ventral aspects was strongly associated with motivational factors (Fig 4, 7), but dorsal cells also represented this information. While all aspects of mPFC were engaged in each of the functional processes we examined, the degree of engagement was non-uniform. Biases existed and a functional gradient emerged in moving from dorsal to ventral along the medial wall (Fig 8).

The combination of these three perspectives points to a nuanced picture of mPFC function: Discrete subregions exist, which preferentially processes different aspects of the cognitive needs of a task (here RRow). But, these subregions, like the cells within them, multiplex the necessary cognitive information along a dorso-ventral gradient. By combining anatomical information, measures of network communication (such as through transfer entropy), and analyses of the relationships between neural activity and behavior (such as through mutual information), it becomes possible to determine how complex brain regions, such as the mPFC, process information. Future studies should build on this work and examine the consequences of manipulating these functional relationships and the causal roles of the distinct mPFC subregions in learning and action selection. Complex tasks such as RRow provide a particularly useful avenue for such study as they provide multiple access points for teasing apart the complex functions of prefrontal cortex (Schmidt and Redish, 2021; Sweis *et al.*, 2018b).

Restaurant Row as a window into decision making

Decision making is an inherently complex and interconnected topic. One method for addressing this complexity entails isolating a particular component of the decision process in a specifically designed behavioral task. Here, we took an alternative approach that built on the complexity of the naturalistic behavior available in the Restaurant Row task (Steiner and Redish, 2014; Sweis *et al.*, 2018a). Successful performance in RRow requires monitoring and integrating a wide swath of relevant variables, all of which we found to be represented within the mPFC. But the task also presents a series of independent and interacting decision processes. The Offer Zone and Wait Zone have been linked to distinct decision processes, with experimental manipulations affecting decisions in each independently (Sweis *et al.*, 2018b; Sweis *et al.*, 2018c). Our data reinforce this understanding with functional differences between ACC and dPL between these task phases. Manipulations specifically of the dorsal mPFC have been found to alter interactions with hippocampus and signatures of deliberative strategies in multiple planning stages of this task (Schmidt *et al.*, 2019; Schmidt and Redish, 2021). Here again our data inform this past work, finding that choice coding preferentially emerged within these same dorsal areas of mPFC that are interconnected with areas such as hippocampus that are implicated in deliberation. Thus, Restaurant Row provides an important middle ground in the study of decision making. It provides a scenario that mirrors decision making in the natural world and engages a wide array of neural processes operating in concert, while at the same time providing instances of reliable and well-controlled windows into specific decision-making processes for study with causal manipulations or neural recordings.

The foraging economics of RRow parallels similar results seen in other rodent economic tasks. Another study that directly sought to answer how prefrontal firing is involved in economic decision making is the

work of Sul et al. (Sul *et al.*, 2010). Using a bandit task, this study found that firing across the prefrontal cortex, both medial prefrontal (ACC and PL/IL) and orbitofrontal (OFC), represented the identify of an upcoming decision just prior to rats making their choice. Furthermore, this coding for choice was stronger in ACC as compared to PL/IL, a finding that matches our results. When examining neural signals of value, Sul et al. found limited coding within the ACC and PL/IL subregions with stronger representations in OFC. Here again our results match this past work as we find no evidence for coding of value (here being offer delay) while rats make their initial value decision in the OZ.

One limitation that exists in many naturalistic decision-making tasks is the ability of an experimenter to distinguish between neural activity associated with a decision process as compared to a motor output. Externally, it is only possible to gauge that a decision has been made based on an animal's motor outputs (or lack of outputs), intrinsically linking the two processes. While naturalistic decision tasks cannot escape this reality, RRow provides two cases that can begin to dissociate between decisions and motor actions. First, decisions in the Offer Zone entail prescribed directions of movement: accept decisions are always left turns and skip decisions are always right turns. Accordingly, examining neural activity either ipsilateral or contralateral to the direction of movement can provide clues of what may be related to motor output vs decision (Fig 5-Supp1) (Erlich *et al.*, 2011). Second, decisions in the Wait Zone require rats either to actively leave the site (quits) or to passively remain (earn). Thus, if neural activity is associated with motor responding one should observe firing increases only during active quit decisions. In our data, we found firing rate increases during both responses, pointing towards engagement of mPFC in the decision component (Fig 6-Supp2). Additionally, we can compare neural response when rats quit a decision in the Wait Zone and when they eventually leave the Reward Zone after consuming their food (Fig 6-Supp3), as both cases entail similar motor actions, but reflect markedly different cognitive states. While there were correlations between the neural responses between these two periods, especially in the more dorsal aspects of mPFC, firing of the majority of cells differed markedly between quitting an offer and leaving the reward zone after consuming a reward. Thus, our data suggest that in RRow, decision processing is a stronger driver of mPFC activity than motor output.

Prefrontal cortex and its involvement in decision making

Current theories suggest that a central purpose of the mPFC is to support effective decision making as organisms interact with their world (Dalley *et al.*, 2004; Euston *et al.*, 2012; McLaughlin *et al.*, 2021; Miller and Cohen, 2001). Building off of this framework, our findings can provide valuable insight into how the different components of mPFC may operate in concert to support decision-making, particularly in naturalistic foraging scenarios such as RRow (Fig 8). We found that each task-relevant variable that we examined was represented within the neural firing patterns of mPFC to a significant degree, presumably to be employed in furthering the rat's goals of understanding their world and securing rewards. Yet information was not represented uniformly across the full mPFC. We found that two distinct cognitive domains emerged, and that these were anatomically biased between the dorsal and ventral aspects of mPFC. In the dorsal mPFC (ACC and dPL), variables related to active decision making were most prominent, and strong changes to firing patterns occurred at critical decision points in the task. In contrast, neural activity in the ventral mPFC (vPL and IL) was more related to task variables that changed slowly across the session, such as motivation-related factors. Interestingly, this functional biasing between dorsal and ventral mPFC has a strong anatomical basis (Heidbreder and Groenewegen, 2003; Voorn *et al.*, 2004): dorsal components of mPFC more strongly project to structures such as dorsal hippocampus, dorsal striatum, and premotor cortex which are collectively involved in the selection and

execution of specific actions (Heidbreder and Groenewegen, 2003; Heilbronner et al., 2016; Hoover and Vertes, 2007; Sesack *et al.*, 1989), while ventral mPFC has more connectivity with ventral hippocampus, ventral striatum, amygdala, and hypothalamus which are more tied to motivation components and generalized status monitoring (Heidbreder and Groenewegen, 2003; Heilbronner *et al.*, 2016; Hurley et al., 1991; Sesack *et al.*, 1989).

Moving beyond this initial perspective of a functional gradient from dorsal to ventral mPFC, our data highlight differences across the four subregions. ACC was more strongly engaged during Offer Zone decisions with notable weaker choice coding and transfer entropy signals in the Wait Zone. Comparatively, dPL was strongly engaged in the Wait Zone period where firing rate differences reliably differentiated earn vs quit decisions and encoded information about offer delay required in making these decisions. Therefore, our data suggests a bias towards ACC engagement during primary decision making (Accept vs Skip in the Offer Zone) with dPL then engaged during re-evaluation of these decisions (Earn vs Quit in the Wait Zone). In the ventral mPFC, we found that neural activity in both vPL and IL were strongly correlated to motivation-related components of the task (task progression and the first part of the linger time). While analysis of real-time coupling between cells clearly separated the two ventral subregions, further analysis in the task found only limited differences between them. One possibility is that this lack of dissociation between vPL and IL is a product of RRow as a task where overall motivational factors are less critical for success than the active decision making that drives dorsal mPFC. Future studies should specifically aim to parse out any functional differences in ventral mPFC through directed examination of motivational components on decision making.

Methods

Subjects

Eight adult Fisher-Brown Norway F1 hybrid rats (4M, 4F) were used in the study. Sample size (n = 8 rats) was based on past studies and did not employ a formal power test. All rats were bred in house and at age 9-15 months during experiments. Rats were single housed in a temperature-controlled colony room with a 12-hour light/dark cycle (lights on at 8:00 AM). Throughout the experiment, rats were food restricted to at or above their 80% free feeding weight, except for approximately one week prior to implantation of silicon probes during with rats were given full access to food. All training and testing sessions were conducted during the light phase and each rat performed the behavioral task at approximately the same time each day. All procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC) and were performed following NIH guidelines.

Each subject was treated independently to yield a sample of 8 biologic subjects contained within the study. No formal replication was performed, though the behavioral data reported here largely replicate previously published works.

Restaurant Row task

Prior to all experiments rats were handled daily to acclimate them to human contact. They were then placed on food deprivation and for four successive days were offered one hour of access to flavored food pellets (45mg/pellet, TestDiet, Richmond IN). After familiarization with the available food rewards rats were trained on the economic decision task Restaurant Row (Steiner and Redish, 2014; Sweis *et al.*, 2018a). All behavioral events were operated through custom built Matlab software (Matlab v2015b, Mathworks) to track rat position, control task state, play tones, and trigger reward delivery. Video was collected via an overhead camera sampling at 30Hz and tracking an LED that was attached to a backpack strapped to the rats (pre-op) or mounted to the implant (post-op). Tones were played on a set of external speakers and food delivery was via a Matlab-Arduino interface (Arduino Uno) sending TTL pulses to Med-Associate feeders (Fairfax, VT).

In Restaurant Row, rats had one hour each day to run clockwise around a raised octagonal track (130cm x 130cm overall footprint) and collect flavored food pellets at four reward sites (restaurants) located at the end of radial spokes. At each reward site, automated feeders delivered two reward pellets (90mg total) of a unique flavor: cherry, banana, plain (unflavored), or chocolate. All food pellets were nutritionally equivalent and varied only in their flavor. Throughout the entire experiment flavors were fixed to their respective reward sites.

For each restaurant visit, rats first entered the Offer Zone (OZ) where the frequency of an auditory cue signaled the temporal delay that the rat was required to wait to earn a reward. Delays ranged from 1 sec to 30 sec, drawn pseudorandomly from a uniform distribution, with higher frequency tones corresponding to longer delays. A 1kHz tone corresponded to the 1 sec delay with 175Hz steps up in frequency for each additional second. Tones were played for 100ms and were repeated at one second intervals. While rats were in the OZ, tones continued to play at the same frequency, and time did not count down, until rats elected to either <u>Accept</u> the offer and move into the Wait Zone (WZ) or <u>Skip</u> the opportunity and advance to the next restaurant. Upon the rat accepting an offer and moving down the reward spoke into the WZ, the delay began to count down with tones of progressively lower frequency played each second. If the rat remained in the WZ until the delay reached zero, they <u>Earned</u> a reward

and flavored reward pellets were delivered. Upon earning the reward rats were designated as now being in the Reward Zone (RZ) and could eat their reward and linger in the zone until ready to advance to the next restaurant. Note that the WZ and RZ were spatially coincident and were defined according to pre vs post reward delivery. However, at any time the rat could instead elect to Quit the offer, leave the WZ, and transition to the next restaurant. Upon exiting the WZ the offer was rescinded, tones stopped, and the reward could no longer be earned, but otherwise quit actions carried no additional punishment. After leaving a given restaurant (Skip, Quit, or Earn), rats proceeded through the Transition Zone (TZ) to the next site where a new delay offer would be presented, and a different flavor reward could be earned. Rats were permitted one hour of time on Restaurant Row each day, which represented their sole opportunity to receive food each day, presenting an inherent economic pressure.

Restaurant Row initial training

Training on the Restaurant Row task occurred over the course of 25 days. Progressively longer potential delays were introduced over the course of training, but the task was otherwise identical to the full version. For training days 1-5 all offers were 1 sec delays. For days 6-10 offer delays ranged from 1-5 sec. Additionally, on days 1-7 a combined OZ/WZ was used in which all offers were automatically accepted, and rats simply had to remain within either zone to earn the reward. For days 11-15 offer delays ranged from 1-15 sec. For days 16-25 delays spanned the full range of 1-30 sec and was the full version of the Restaurant Row task. Two rats (R536/R543) received an additional 5 days of training in the final 1-30 sec range. At the completion of this training sequence, rats were given free access to food for at least one week prior to surgery for implantation of silicon probes.

Surgery

At the time of surgery rats were anesthetized with isoflurane (2% in O_2) and were placed into a stereotaxic frame. Carprofen (5mg/kg, SC), dual-cillin (0.2ml, SC), and saline (3ml, SC) were administered for analgesia, prophylaxis, and hydration respectively (All drugs provided by UMN Research Animal Resources). A pair of ground screws were implanted over cerebellum and a 3D printed ring was attached to the skull via an anchor screw and C&B Metabond adhesive. Craniotomies were drilled over the mPFC (±0.6-0.7 ML, +2.5-3.0 AP) and either one or two (bilateral) linear 64-channel silicon probes (Cambridge Neurotech H3 model) were implanted targeting either the dorsal (ACC/PL) or ventral (PL/IL) extent of the PFC medial wall. Each probe had recording contacts spaced at 20µm intervals spanning a total of 1260µm along the dorso-ventral axis. For dual implants, one probe targeted dorsal while the other targeted ventral mPFC to span the full extent (~3mm) of the medial wall. Targeting of dorsal and ventral mPFC on the left and right hemisphere was counterbalanced across rats, accounting for sex, but not considering any behavioral metrics.

Probes were housed in custom built 3D printed microdrives that allowed for independent adjustment of each probe along the dorso-ventral axis. After lowering the probes into brain (~1.5mm; All recording sites into brain), craniotomies were filled with a sterile mixture of bone wax and mineral oil (3/1 by weight) and microdrives were adhered to the skull with Metabond. A 3D printed shell and head cap were attached to the skull mounted ring to protect the implant. Finally, an LED was attached to the rear of the shell to facilitate video tracking during recordings. At the completion of surgery rats were given an injection of Baytril (10mg/kg, SC) and additional saline (20ml/kg, SC) as well as oral administration of Children's Tylenol (0.8ml).

Restaurant Row post-op training

Rats were given three days of recovery following surgery and were then reintroduced to the Restaurant Row task for retraining. Retraining followed a similar sequence to original training but in an expeditated fashion. Retraining lasted 10 days total: 1 day of 1 sec, 2 days of 1-5 sec, 2 days of 1-15 sec, 5 days of 1-30 sec. Additionally, for the final three days of retraining (1-30 sec delays), rats ran the task with the recording tether attached, familiarizing them with the process of performing the task under full task and recording conditions. After this 10 day retraining sequence rats continued to perform Restaurant Row during mPFC recordings for an additional 14 days (7 days for R506). Over the course of this retraining sequence, probes were gradually lowered to the appropriate depths targeting the dorsal (ACC/PL) or ventral (PL/IL) mPFC.

Recordings

Recordings were performed over 14 days (7 days for R506) immediately following post-op retraining and positioning of the probes. A technical glitch during one session for R535 (day 5) precluded synchronization of neural data to behavior and was thus excluded. A total of 104 sessions of data were included across all analysis. The Restaurant Row task during recording sessions was identical to the full task used in training; rats were given one hour on the track and offer delays ranged from 1 sec to 30 sec sampled pseudorandomly from a uniform distribution. Before and after the 1 hour task period rats sat on a flower pot for 5 minutes to record baseline neural activity during quiet resting. These pre/post data were used in the process of single unit isolation but were otherwise not analyzed here.

All recordings were collected at a 30kHz sampling rate using a 256 channel Intan RHD recording system connected to 64-channel Intan RHD headstages. An Arduino interface between Matlab and the Intan system facilitated time syncing of neural data to ongoing behavior and task events. Periodically throughout the 14 day recording sequence probes were lowered (35-150µm) to increase the likelihood of recording a diverse array of units and to increase the overall sampling across the dorso-ventral axis of the mPFC. All movement of probes occurred after completion of a day's recordings leaving 23 hours for probes to restabilize before the next recording session. While we took efforts to record from different cells across the experiment, we are unable to claim that all recorded cells are unique.

Single unit identification was achieved using Kilosort v2.0 (https://github.com/MouseLand/Kilosort) with post clustering curation performed manually using Phy (https://github.com/cortex-lab/phy). Each silicon probe was analyzed for units independently and included data from the 60 min behavioral session as well as the two 5 min pre/post periods. Prior to Kilosort processing, data underwent a median reference subtraction across the 64-channels of the given probe, and subsequent unit identification used a 600Hz high-pass filter. Collectively we identified 4030 single units across all recording sessions. All isolation of single units was performed blind to the session's corresponding behavioral data.

Histology

At the completion of all recordings, 3 seconds of 4µA of current was passed through 20 channels on each probe to induce gliosis at the final recording location. Two days later rats were overdosed with pentobarbital and were perfused transcardially with phosphate-buffered saline and paraformaldehyde. Probes were left in place for 3-5 hours and were then carefully removed from the brain, leaving a visible track of their final position. Brains were then extracted and post-fixed in paraformaldehyde for 24 hours before being transferred to a mixture of paraformaldehyde and sucrose for cryoprotection. Coronal sections (40µm) through the mPFC were cut on a cryostat, mounted to slides, and stained with cresyl

violet. For each probe, the entry point into the brain, final tip location, and trajectory were determined from serial sections. Together with records of experimenter movement of each probe through the brain, this information was used in localizing units in 3D space within mPFC (see below). Across all histological sections we compared the expected depth of probe tips based on movement records to the depth measured in processed tissue. We found that minor tissue shrinkage occurred during histological processing with depth in sections to be on average 92 ± 4% smaller than the in vivo placements. For two out of the 13 total probes, histology revealed the probe trajectory to be outside of mPFC (one in VO, and one in MO). Data from these two probes (1013 units) were excluded from all analyses leaving a total of 3017 units verified to be located in the mPFC. While a large proportion of ACC and dPL cells came from one rat each, for all analyses we verified that results were qualitatively similar across the remaining animals (data not shown).

Data Analysis

Behavior

Each visit to a Restaurant Row restaurant was identified as a Skip, a Quit, or an Earn dependent on whether or not rats entered the WZ (Skip) and subsequently if they remined in the WZ throughout the full delay (Quit/Earn). For each visit, decision reaction time was taken as the elapsed time between entering and exiting the OZ either into the WZ in the case of accepting the offer or moving to the next restaurant for Skips. Quit time was the time spent in the WZ prior to a quit decision being made and rats exiting the WZ. Note that, by definition, for a given restaurant visit the quit time must be less than the accepted delay. Linger time was defined as the time spent in the RZ after delivery of the reward up to the point at which rats exited the zone. Offer decision thresholds were calculated for each restaurant (flavor) independently by the least-squares fit of a Heaviside function of receiving food at each visit as a function of offer delay. Note that both Quit and Skip instances were treated identically (no food earned). To calculate the consistency of thresholds across days we correlated the four-element vector of restaurant thresholds between each pair of recording sessions, either both taken from the same rat (day-to-day consistency) or from different rats (null distribution). Threshold consistency was also computed as the variance for a given restaurant threshold across sessions as compared to the variance across restaurant thresholds for a given session. Reward rank was determined in each session by sorting restaurants by decision threshold with the largest threshold (the reward the rat was willing to wait longest for) ranked 1st. Ties between restaurant thresholds were broken by the total number of earned rewards with the greater number of earns identified as the more preferred restaurant. For some behavioral analyses we computed the “subjective value” of each offer delay as its signed distance from threshold. Note that thresholds were a function of both rat and restaurant. Good offers were those with a positive value and bad offers had a negative value. Progressive distance from zero reflected the magnitude of this offer quality.

$$Subjective\ Value\ = Threshold_{RR} - Delay$$

Unit localization in standardized 3D space

To precisely evaluate functional activity along prefrontal medial wall every unit recorded from the mPFC was localized to its 3-dimensional location in a standardized atlas space. This was achieved by integrating waveform profiles across probe recording channels, physical locations of these channels on the probe shank, daily records of turning depths of each probe through the brain, histological

information about probe trajectories through the brain, and stereotaxic coordinates in a standardized atlas framework. First, for each cell we evaluate the 64 waveform profiles across all recording channels of a given probe and identified the channel with the maximal amplitude as the putative location of the cell body (Csicsvari *et al.*, 2003). We then computed the linear distance of this channel from the ventral tip of the probe based on 20µm spacing between recording channels. Using daily records of turning depth of the probe tip into the brain we were able to identify of the depth along the probe trajectory of the identified recording channel and putative cell body location. Next, for the probe of interest we identified in histological sections the 3D location of the entry point of the probe into the brain and the final location of the probe tip. Note that because we only ever turned probes ventral, this histological position marks the maximal distance that the probe was turned into the brain. The relative location of these entry and end points were then identified in a standard Paxinos and Watson atlas (Paxinos and Watson, 2007). While our histological sections provided an AP resolution of 40µm, identification of corresponding points in the atlas provided a much coarser assessment along this axis. Finally, we computed the location of our recorded cell in standardized 3D space by converting the relative distance of our putative cell body location along the probe trajectory in real space (histological and probe turn distance), to the relative position along the trajectory within atlas normalized space.

Basic spiking and waveform properties

Average firing rate for each cell was defined as the total number of spikes detected over the full behavioral session divided by the one-hour duration. This metric did not include spikes during the 5 min pre/post recording periods. Waveforms for each cell were taken from the recording channel with the largest amplitude and thus inferred to be most proximally located to the cell body. The time between the peak and valley of a cell's average waveform was taken as the spike width. Coefficient of variation for each cell was computed by binning spiking over the full one-hour session into 50ms time bins, calculating the average firing rate in each bin, and dividing the standard deviation across these firing rates by their mean. For each cell we calculated both the median inter-spike-interval (ISI) and standard deviation across the ISI during over the course of the one-hour behavioral session. As a measure of bursting we computed for each cell the ratio between short ISIs (<10ms) and long ISIs (>125ms; ~1 theta cycle). To distinguish between principal cells and interneurons we compared average firing rate and spike width for our recorded cells. Data clustered into two largely separable groups and boundaries between them were drawn by hand by a trained experimenter.

Normalization of Offer and Reward Zone time

Progression through both the Offer Zone and the Reward Zone were self-paced and varied across each restaurant visit. Therefore, we evaluated neural activity during these two periods with respect to the relative time spent in the zone. For each zone visit we divided the total time spent in the zone into 20 bins such that each bin composed 5% of the total time. For the period before and after we took the time spent in the preceding zone and subsequent zone and divided it into 10 bins (10% relative time). Analyses included 3 bins pre and post the decision period for a total of 26 normalized time bins. This normalization method was used in Figures 5 and 7D.

Peri-Event Time Histograms (PETHs)

Peri-event time histograms (PETHs) were computed using standard methods over either normalized time bins (OZ and RZ) or absolute time bins (WZ). Analysis of the WZ used bin widths of 100ms. For each visit to a restaurant, we totaled the spike count in each time bin and divided by the bin width to

compute average firing rate in each bin. We then computed the mean firing rate across all restaurant visits for each time point. To specifically examine firing profile differences across various task conditions additional PETHs were constructed in which only a subset of restaurant visits was included (e.g. only visits that led to a Skip decision, or only visits with an offer delay between 1 and 5 sec).

To account for chance behavior of mPFC spiking, not directly associated with ongoing behavior or task variables we computed a series of PETHs based on permutated spike times of each cell. For each cell, the full spike train was circular shifted in time ("shuffled") by a random amount ranging from 1 minute to 59 minutes to dissociate neural spiking from ongoing task events. PETHs were then calculated using the behavioral event times and the time shifted spike train ("shuffled PETH"). This process was repeated 30 times and the mean PETH across the distribution of shuffles was taken as the by chance, expected activity. For all analyses, PETHs of individual cells were normalized by subtracting this shuffled average.

Mutual Information

To directly relate the spiking of mPFC cells to behavior during decision making we computed the mutual information (MI) between firing rate and specific task conditions (Timme and Lapish, 2018). Intuitively, MI quantifies how much information neural firing can provide in predicting a concurrent behavioral condition. For each behavioral variable data were binned into 5 bins of uniform width. For MI calculations over time, we followed a parallel approach to that of our PETH calculations in computing MI at each time bin treated independently. After calculating firing rates in each time bin (PETHs), firing rates were discretized into 5 uniformly spaced bins. For each time point, the binned spiking rates across all restaurant visits were collected along with the corresponding behavioral/task condition data. The mutual information between the two matched vectors was then computed according to:

$$MI = \sum p(B,S) log_2(\frac{p(B,S)}{p(B),p(S)})$$ where ***B*** and ***S*** reflect binned behavior and spiking rate

As for PETHs, we normalized our MI calculations by repeating all analyses after time shifting spike trains of our mPFC cells. Shuffling and MI computation was repeated 30 times for each cell and the mean across this shuffled distribution was taken as the chance level. All analyses subtracted this shuffled average from the MI computed on raw, unshuffled spike times.

Correlation between Firing and Behavior

To relate firing rate of prefrontal cells to behavioral variables we computed the average firing rate in each task zone for each restaurant visit. For each behavioral variable, the corresponding value at each time point was taken and the Pearson's correlation was computed between the vectors of firing rate and behavioral values. To relate behavioral variables to each other we used this same process but correlated a pair of vectors across different behaviors.

Stepwise Regressions between Firing and Behavior

To account for linear relationships between behavioral variables in the RRow task we used a stepwise linear regression procedure to relate firing of mPFC cells to independent behavioral variables. We used two complementary methods for performing the stepwise regression but in both cases, we performed a sequential, stepwise linear regression to progressively remove the influence of a given behavioral variable on the neural firing before relating the next behavioral variables. That is, subsequent behaviors were fit to the residual firing rates after accounting for preceding behaviors.

In the first method we fit each behavioral variable to mPFC spiking in a predetermined order of behaviors. Thus, each cell underwent an identical analysis and behavioral variables were regressed out in the same order across the full population. Order of behaviors was determined from the overall prevalence within the data set based on a standard correlation analysis (Fig 2C). In the second method we used the Matlab function "`stepwisefit`" which fits the stepwise regression for each cell independently. Regression order is determined by the strength of the regression for each individual, meaning that the order that each behavioral variable is regressed out will vary across the population. Note that in both cases, the "Regression Rank" identifies the overall strength of the relationship to neural firing, not necessarily the original order in which it was applied in the analysis.

Transfer Entropy

To measure the degree to which the activity of mPFC cells may be informative about and/or influence each other we computed the transfer entropy (TE) between the firing rates of pairs of simultaneously recorded mPFC cells (Timme and Lapish, 2018). Related to granger causality, TE measures the information provided by the past activity of an input (cell X) about the future of a response (cell Y) above and beyond the information derived from the response's own past. Note that TE is directional and that TE from X to Y does not necessarily equal the TE from Y to X. First, activity of all cells over the full Restaurant Row session was binned using 10ms time steps. For each pair of simultaneously recorded cells, we computed the transfer entropy between an input cell X and a response cell Y according to:

$$TE_{X\rightarrow Y} = \sum p(Y_t, Y_{t-1}, X_{t-1}) log_2 \frac{p(Y_t|Y_{t-1}, X_{t-1})}{p(Y_t|Y_{t-1})}$$

To normalize the TE values, we repeated the above analysis after time shifting spiking data for only the predictor X cell. The full spike train of the predictor X cell was circular shifted in time and then TE between the shifted predictor cell X and the unshifted response cell Y was calculated to yield a single by-chance TE value between the pair. This shuffling procedure was repeated 30 times to yield a distribution of chance TE values (30 chance values). A mean was then taken across these 30 shuffles and taken as the by-chance expected relationship between the pair of cells. This mean was subtracted from the raw calculation taken from unshifted spike times, and this net TE is reported.

To examine the relationship between pairwise TE and anatomical locations we binned resultant TE values into a 2D histogram (spatial resolution of 100µm) according to the DV position of the two cells included in the pair, paying mind to which was analyzed as the input cell X and which was the response cell Y. The mean TE was taken for each bin. Transfer entropy was quantified at depths along the medial wall by averaging the normalized TE in bins along the identity diagonal. At each DV step we calculated the mean across all bins within 200µm; that is to say, bins in which both pairs of cells were within 200µm of the dorso-ventral point of interest. Valleys in the quantification (subregion boundaries) were computed as the local minima of the quantification along the DV axis.

To calculate the TE values for activity in each task phase independently we repeated the above procedures but only included time bins from when rats were located in the task phase of interest.

Subsampling of Data in dPL, vPL, and IL

To account for the lower sampling of ACC cells relative to other subregions we repeated main analyses after subsampling the populations of the dPL, vPL, and IL subregions to match that of ACC.

For TE calculations, we subsampled populations at random to match the cell count of ACC (72 cells) and repeated our quantification along the dorso-ventral depth and identification of subregional bounds as for the full data set. This subsampling procedure was repeated 500 times with subregional bounds computed for each iteration. Because TE is a pairwise metric, we also subsampled our data to match the average number of cell pairs formed by our ACC data set. For each other subregion we randomly sampled a number of cell pairs to match the number of within region pairs of our ACC data set (334 pairs) and the average number of across region cell pairs for our ACC data set (397 pairs). We then computed the subregional bounds on this subsampled data set and repeated the process 500 times. To compute distributions of TE across subregions we used these same subsampling methods and report one iteration in the corresponding supplemental figure, though repeated sampling yielded comparable results.

For analyses of OZ, WZ, and RZ responding we subsampled our populations of dPL, vPL, and IL cells to match that of ACC (72 ells). We then repeated our analyses for each subregion using only this subsampled data set. One such iteration is shown in supplemental figures, though repeated iterations yielded comparable results.

Changes in Functional Properties over the Session

To test for any changes in functional properties over the session we divided our data into equal size blocks over the course of the session and repeated our transfer entropy analyses, and the main analyses in each of the OZ, WZ, and RZ. For TE calculations, the session was divided into 5 equal durations (12 min each, Blocks 1 – 5 from start to end of the session). Quantification of TE across depth, and average TE for within subregion cell pairs were calculated independently for each time block. For OZ, WZ, and RZ analyses the session was divided into 3 equal durations (20 min bins, Blocks 1-3 from start to end). For OZ and WZ the cumulative MI over the respective decisions periods was calculated in each block. For the RZ the average firing rate was calculated in in the 2 seconds following reward delivery.

Correlation between Firing Rates and Offer Delay/Linger Time

To relate firing of mPFC cells to either offer delay (WZ, Fig 6) or linger time (RZ, Fig 7) we computed the average firing rate profile for each subregion for each rat independently. For linger time, behavioral bin windows were taken as percentages from the minimum to the maximum linger time computed independently in each behavioral session. Note that because of the long tail in the distribution of linger times, 0-30 percent linger time (range in Fig 7D) accounts for 87% of all RZ visits. Average PETHs were computed for each behavioral bin (offer delay or linger time) yielding up to 48 total PETHs (8 rats x 6 bins) for each mPFC subregion. For offer delay behavioral bins correlations between these 48 PETHs and the respective behavioral bins were computed during four time windows. For the Wait Zone and offer delay analyses these were: Pre WZ (0-1 sec before WZ entry), Early WZ (0-3 sec after WZ entry), Late WZ (0-3 sec before WZ exit), Post WZ (0-1 sec after WZ exit). For the Reward Zone and linger time analyses these were: Pre RZ (3 bins before RZ enter), Early RZ (first half of RZ; 10 bins), Late RZ (second half of RZ; 10 bins), Post RZ (3 bins after RZ exit). For each time window the average firing rate over the window was computed for each PETH and the Pearson's correlation was computed between firing rate and center of the behavioral bin.

Statistics

Comparison of threshold variance across restaurants vs across sessions was performed with a paired Wilcoxon Signed-Rank test with the rats as the samples. Behavioral analysis of time in the offer zone (reaction time) between accept and skip decisions was through a paired Wilcoxon Signed-Rank test between average reaction time at each offer value (data points for each rat). This method accounts for any variation in reaction time that may be driven by the offer value itself and directly compares between the two decision processes. Correlations between behavioral times (reaction time, quit time, linger time) and offer value were similarly performed after averaging metrics at each offer value (data points for each rat).

All analyses of single cell firing patterns were normalized by subtracting the mean of “shuffled” data generated by circularly time-shifting spiking with respect to behavior. 30 iterations of each shuffle were performed and the mean was taken across the distribution. An identical procedure was used for pairwise TE analysis but only spiking from the input X cell was shuffled.

Comparison of cell spiking to behavior was made on the distribution of correlations for each behavior of interest. In each case the raw correlations were compared against a distribution generated from the same behavioral data but using shuffled spike times using a Wilcoxon Signed-Rank test. Spike shuffling was repeated 30 times and the mean correlation across the shuffles was taken for each cell.

Comparison of overall TE metrics against chance were performed with a Wilcoxon Signed-Rank test of shuffle normalized TE values against 0. Note this is equivalent to a matched test between raw data and shuffled data. TE quantified along the DV axis was evaluated as multimodal via Hartigan’s Dip Test (Hartigan and Hartigan, 1985). Comparison between TE within a subregion vs across subregions was performed with a Mann-Whitney U test. To evaluate TE as a function of distance between subregions, pairwise TE was grouped according to the number of steps between a pair of mPFC subregions (0 Steps for within subregion, 1 Step for adjacent regions, etc.). Comparison between step sizes was via an ANOVA and individual pairs of step sizes evaluated with Tukey-Kramer HSD post-hoc criterion.

For generalized task variables (Task Phase, Session Time, Reward Rank), mutual information values were computed for each cell and normalized as described above against shuffled spike times. Normalized mutual information values were identified as significant using a Wilcoxon Signed Rank test. Comparisons across mPFC subregions was via an ANOVA and HSD post-hoc testing.

Comparisons of mutual information metrics (OZ and WZ choice) over time were performed using a repeated measures ANOVA with time bins as the repeated variables. OZ comparisons used 26 time bins and WZ used 70 bins. Post-hoc comparisons for each subregion were computed using a series of Wilcoxon Signed-Rank tests comparing mutual information at each time point against 0. Tests were one-tailed in the positive direction and significance levels for these tests were evaluated after Bonferroni correction for multiple comparisons; n = 26 or 70 time bins. To compare responses across subregions we computed the average response for each metric over a specific period of interest to yield a single value for each cell. In the OZ this period was from the OZ entry through the end of the analysis window (end of post-OZ). For the beginning and end of the WZ this period was the first 2 second after WZ entry or the last 2 seconds before WZ exit. A comparable method was used for summarizing average responses of mPFC firing rates in response to reward delivery. In the RZ the period of interest was from 0.5 seconds after RZ entry to 3 seconds after RZ entry. Comparisons across mPFC subregions was via an ANOVA and

HSD post-hoc testing. In cases of comparing dorsal mPFC (ACC and dPL) vs ventral mPFC (vPL and IL) we used a Mann-Whitney U test.

To compare proportions of mPFC cells that were biased toward accept vs skip decisions in the OZ we used a chi-square (X^2) test for equal proportions. The proportion of skip preferring cells was determined as those with an average net firing rate greater than 0 and was compared against a null hypothesis of 50% of cell in the respective mPFC subregion. Note that because accept vs skip is a binary choice the proportion of accept preferring cells is the direct inverse with identical statistical results.

Correlation between average firing rates and offer delay (Fig 6) and linger time (Fig 7) were performed after first averaging firing rate PETHs for each rat independently. For each rat, PETHs were generated for each behavioral bin yielding up to 48 firing rate profiles (8 rats x 6 bins). For comparing offer delay firing in the WZ behavior was grouped into evenly spaced bins according to the offer presented in the OZ. For comparisons to linger time we used bins composed of percent of the total linger time. In each session the 1st and 99th percentiles of linger time were taken as the session minimum and maximum. For each visit to the RZ the total linger time was computed as the relative percentage between this minimum and maximum. Using this method, included linger times between 0 and 30 percent accounted for 87% of all RZ visits. From these PETH profiles average firing rates were computed in four time windows for each analysis. For the WZ, time bins were 0-1 sec before WZ entry (Pre WZ), 0-3 sec after WZ entry (Early WZ), 0-3 sec before WZ exit (Late WZ), and 0-1 sec after WZ exit (Post WZ). For the RZ, time bins were the 3 bins prior to RZ entry (Pre RZ), the first 10 bins of the RZ (Early RZ), the second 10 bins of the RZ (Late RZ), and the 3 bins after RZ exit (Post RZ).

All correlations between sets of values used Pearson correlations and significance assessed at $\alpha = 0.05$. Unless otherwise indicated, significance was assessed using two-sided tests and an $\alpha = 0.05$.

Data and Code

All data and code necessary to replicate the presented results can be found at: https://osf.io/s5xqm/

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
strain, strain background (include species and sex here)	Fisher-Brown Norway F1 Hybrid rats	Bred in House		4M, 4F; Aged 9-15 months
chemical compound, drug	Isoflurane	UMN Research Animal Resources (RAR)		
chemical compound, drug	Carprofin	UMN RAR		
chemical compound, drug	Dual-Cillin	UMN RAR		
chemical compound, drug	Baytril	UMN RAR		
chemical compound, drug	C&B Metabond	Patterson Dental	Cat #: 553-3484	
chemical compound, drug	Pentobarbitol	UMN RAR		

chemical compound, drug	Paraformal dehyde	Sigma-Aldrich	Cat #: 158127	
software, algorithm	Matlab v2017b	Mathworks	RRID: SCR_001622	For Formal Analysis
software, algorithm	Matlab v2015b	Mathworks	RRID: SCR_001622	For Running Task
software, algorithm	Kilosort v2.0	Marius Pachitariu	https://github.com/MouseLand/Kilosort	Downloaded Jan 2020
software, algorithm	Phy v2.0 Beta 1	Cyrille Rossant	https://github.com/cortex-lab/phy	Downloaded Feb 2020
software, algorithm	Analysis Code	This Paper	https://osf.io/s5xqm/	
other	Physiology and Behavioral Data	This Paper	https://osf.io/s5xqm/	
other	Food Pellets (45mg, 5TUL)	TestDiet	Cat #: 1811155; 1811443; 1812298; 1811645	Flavors: Plain; Chocolate; Fruit “Cherry”; Banana
other	64 Ch Silicon Probe	Cambridge Neurotech	ASSY-156-H3	H3 Model
other	256 Ch Intan RHD Recording System	Intan	Cat #: C3100	
other	64 Ch Intan RHD Headstage	Intan	Cat #: C3325	w/ Accelerometer

References

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R.F., Dayan, P., and Costa, R.M. (2021). The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. Neuron *109*, 149-163 e147. 10.1016/j.neuron.2020.10.013.

Berendse, H.W., Galis-de Graaf, Y., and Groenewegen, H.J. (1992). Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. J Comp Neurol *316*, 314-347. 10.1002/cne.903160305.

Birrell, J.M., and Brown, V.J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. J Neurosci *20*, 4320-4324.

Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F., and Quirk, G.J. (2014). Neural structures mediating expression and extinction of platform-mediated avoidance. J Neurosci *34*, 9736-9742. 10.1523/JNEUROSCI.0191-14.2014.

Brockett, A.T., Tennyson, S.S., deBettencourt, C.A., Gaye, F., and Roesch, M.R. (2020). Anterior cingulate cortex is necessary for adaptation of action plans. Proceedings of the National Academy of Sciences *117*, 6196-6204. 10.1073/pnas.1919303117.

Burgos-Robles, A., Bravo-Rivera, H., and Quirk, G.J. (2013). Prelimbic and infralimbic neurons signal distinct aspects of appetitive instrumental behavior. PLoS One *8*, e57575. 10.1371/journal.pone.0057575.

Burgos-Robles, A., Vidal-Gonzalez, I., and Quirk, G.J. (2009). Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. J Neurosci *29*, 8474-8482. 10.1523/JNEUROSCI.0378-09.2009.

Capuzzo, G., and Floresco, S.B. (2020). Prelimbic and Infralimbic Prefrontal Regulation of Active and Inhibitory Avoidance and Reward-Seeking. J Neurosci *40*, 4773-4787. 10.1523/JNEUROSCI.0414-20.2020.

Carlén, M. (2017). What constitutes the prefrontal cortex? Science *358*, 478-482. 10.1126/science.aan8868.

Corcoran, K.A., and Quirk, G.J. (2007). Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. J Neurosci *27*, 840-844. 10.1523/JNEUROSCI.5327-06.2007.

Coutureau, E., and Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behav Brain Res *146*, 167-174. 10.1016/j.bbr.2003.09.025.

Csicsvari, J., Henze, D.A., Jamieson, B., Harris, K.D., Sirota, A., Bartho, P., Wise, K.D., and Buzsaki, G. (2003). Massively parallel recording of unit and local field potentials with silicon-based electrodes. J Neurophysiol *90*, 1314-1323. 10.1152/jn.00116.2003.

da Silva, T.R., Sohn, J.M.B., Andreatini, R., and Stern, C.A. (2020). The role of prelimbic and anterior cingulate cortices in fear memory reconsolidation and persistence depends on the memory age. Learn Mem *27*, 292-300. 10.1101/lm.051615.120.

Dalley, J.W., Cardinal, R.N., and Robbins, T.W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev *28*, 771-784. 10.1016/j.neubiorev.2004.09.006.

Dalton, G.L., Wang, N.Y., Phillips, A.G., and Floresco, S.B. (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. J Neurosci *36*, 1996-2006. 10.1523/JNEUROSCI.3366-15.2016.

Diehl, M.M., Bravo-Rivera, C., Rodriguez-Romaguera, J., Pagan-Rivera, P.A., Burgos-Robles, A., Roman-Ortiz, C., and Quirk, G.J. (2018). Active avoidance requires inhibitory signaling in the rodent prelimbic prefrontal cortex. Elife *7*. 10.7554/eLife.34657.

Do-Monte, F.H., Manzano-Nieves, G., Quinones-Laracuente, K., Ramos-Medina, L., and Quirk, G.J. (2015). Revisiting the role of infralimbic cortex in fear extinction with optogenetics. J Neurosci *35*, 3607-3615. 10.1523/JNEUROSCI.3137-14.2015.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. Nat Rev Neurosci *2*, 820-829. 10.1038/35097575.

Erlich, J.C., Bialek, M., and Brody, C.D. (2011). A cortical substrate for memory-guided orienting in the rat. Neuron *72*, 330-343. 10.1016/j.neuron.2011.07.010.

Euston, D.R., Gruber, A.J., and McNaughton, B.L. (2012). The role of medial prefrontal cortex in memory and decision making. Neuron *76*, 1057-1070. 10.1016/j.neuron.2012.12.002.

Gilmartin, M.R., and McEchron, M.D. (2005). Single neurons in the medial prefrontal cortex of the rat exhibit tonic and phasic coding during trace fear conditioning. Behav Neurosci *119*, 1496-1510. 10.1037/0735-7044.119.6.1496.

Gilmartin, M.R., Miyawaki, H., Helmstetter, F.J., and Diba, K. (2013). Prefrontal activity links nonoverlapping events in memory. J Neurosci *33*, 10910-10914. 10.1523/JNEUROSCI.0144-13.2013.

Groenewegen, H.J., and Uylings, H.B.M. (2000). The prefrontal cortex and the integration of sensory, limbic and autonomic information. In Progress in Brain Research, (Elsevier), pp. 3-28. https://doi.org/10.1016/S0079-6123(00)26003-2.

Hart, E.E., Blair, G.J., O'Dell, T.J., Blair, H.T., and Izquierdo, A. (2020). Chemogenetic Modulation and Single-Photon Calcium Imaging in Anterior Cingulate Cortex Reveal a Mechanism for Effort-Based Decisions. J Neurosci *40*, 5628-5643. 10.1523/JNEUROSCI.2548-19.2020.

Hartigan, J.A., and Hartigan, P.M. (1985). The Dip Test of Unimodality. The Annals of Statistics *13*, 70-84, 15.

Hayden, B.Y., Pearson, J.M., and Platt, M.L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. Nat Neurosci *14*, 933-939. 10.1038/nn.2856.

Heidbreder, C.A., and Groenewegen, H.J. (2003). The medial prefrontal cortex in the rat: evidence for a dorso-ventral distinction based upon functional and anatomical characteristics. Neurosci Biobehav Rev *27*, 555-579. 10.1016/j.neubiorev.2003.09.003.

Heilbronner, S.R., Rodriguez-Romaguera, J., Quirk, G.J., Groenewegen, H.J., and Haber, S.N. (2016). Circuit-Based Corticostriatal Homologies Between Rat and Primate. Biol Psychiatry *80*, 509-521. 10.1016/j.biopsych.2016.05.012.

Hillman, K.L., and Bilkey, D.K. (2010). Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. J Neurosci *30*, 7705-7713. 10.1523/JNEUROSCI.1273-10.2010.

Hillman, K.L., and Bilkey, D.K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. Nat Neurosci *15*, 1290-1297. 10.1038/nn.3187.

Hok, V., Save, E., Lenck-Santini, P.P., and Poucet, B. (2005). Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc Natl Acad Sci U S A *102*, 4602-4607. 10.1073/pnas.0407332102.

Hoover, W.B., and Vertes, R.P. (2007). Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. Brain Structure and Function *212*, 149-179. 10.1007/s00429-007-0150-4.

Hosking, J.G., Cocker, P.J., and Winstanley, C.A. (2014). Dissociable contributions of anterior cingulate cortex and basolateral amygdala on a rodent cost/benefit decision-making task of cognitive effort. Neuropsychopharmacology *39*, 1558-1567. 10.1038/npp.2014.27.

Hurley, K.M., Herbert, H., Moga, M.M., and Saper, C.B. (1991). Efferent projections of the infralimbic cortex of the rat. J Comp Neurol *308*, 249-276. 10.1002/cne.903080210.

Hyman, J.M., Holroyd, C.B., and Seamans, J.K. (2017). A Novel Neural Prediction Error Found in Anterior Cingulate Cortex Ensembles. Neuron *95*, 447-456 e443. 10.1016/j.neuron.2017.06.021.

Hyman, J.M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., and Seamans, J.K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. Proc Natl Acad Sci U S A *109*, 5086-5091. 10.1073/pnas.1114415109.

Hyman, J.M., Whitman, J., Emberly, E., Woodward, T.S., and Seamans, J.K. (2013). Action and outcome activity state patterns in the anterior cingulate cortex. Cereb Cortex *23*, 1257-1268. 10.1093/cercor/bhs104.

Ito, H.T., Zhang, S.J., Witter, M.P., Moser, E.I., and Moser, M.B. (2015). A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. Nature *522*, 50-55. 10.1038/nature14396.

Karlsson, M.P., Tervo, D.G., and Karpova, A.Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. Science *338*, 135-139. 10.1126/science.1226518.

Kesner, R.P., and Churchwell, J.C. (2011). An analysis of rat prefrontal cortex in mediating executive function. Neurobiol Learn Mem *96*, 417-431. 10.1016/j.nlm.2011.07.002.

Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex *13*, 400-408. 10.1093/cercor/13.4.400.

Kolb, B. (1990). Prefrontal cortex. In The cerebral cortex of the rat., (The MIT Press), pp. 437-458.

Lapish, C.C., Durstewitz, D., Chandler, L.J., and Seamans, J.K. (2008). Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. Proc Natl Acad Sci U S A *105*, 11963-11968. 10.1073/pnas.0804045105.

Laque, A., G, L.D.N., Wagner, G.E., Nedelescu, H., Carroll, A., Watry, D., T, M.K., Koya, E., Hope, B.T., Weiss, F., et al. (2019). Anti-relapse neurons in the infralimbic cortex of rats drive relapse-suppression by drug omission cues. Nat Commun *10*, 3934. 10.1038/s41467-019-11799-1.

Laubach, M., Amarante, L.M., Swanson, K., and White, S.R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? eNeuro *5*, ENEURO.0315-0318.2018. 10.1523/ENEURO.0315-18.2018.

Ma, L., Hyman, J.M., Durstewitz, D., Phillips, A.G., and Seamans, J.K. (2016). A Quantitative Analysis of Context-Dependent Remapping of Medial Frontal Cortex Neurons and Ensembles. J Neurosci *36*, 8258-8272. 10.1523/JNEUROSCI.3176-15.2016.

MacDonald, A.W., 3rd, Cohen, J.D., Stenger, V.A., and Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science *288*, 1835-1838. 10.1126/science.288.5472.1835.

Mashhoori, A., Hashemnia, S., McNaughton, B.L., Euston, D.R., and Gruber, A.J. (2018). Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. Elife *7*. 10.7554/eLife.29793.

McLaughlin, A.E., Diehl, G.W., and Redish, A.D. (2021). Potential roles of the rodent medial prefrontal cortex in conflict resolution between multiple decision-making systems. Int Rev Neurobiol *158*, 249-281. 10.1016/bs.irn.2020.11.009.

Milad, M.R., and Quirk, G.J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. Nature *420*, 70-74. 10.1038/nature01138.

Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu Rev Neurosci *24*, 167-202. 10.1146/annurev.neuro.24.1.167.

Mukherjee, A., and Caroni, P. (2018). Infralimbic cortex is required for learning alternatives to prelimbic promoted associations through reciprocal connectivity. Nat Commun *9*, 2727. 10.1038/s41467-018-05318-x.

Passecker, J., Mikus, N., Malagon-Vina, H., Anner, P., Dimidschstein, J., Fishell, G., Dorffner, G., and Klausberger, T. (2019). Activity of Prefrontal Neurons Predict Future Choices during Gambling. Neuron *101*, 152-164 e157. 10.1016/j.neuron.2018.10.050.

Paxinos, G., and Watson, C. (2007). The rat brain in stereotaxic coordinates, 6th Edition (Academic Press/Elsevier).

Peters, J., LaLumiere, R.T., and Kalivas, P.W. (2008). Infralimbic prefrontal cortex is responsible for inhibiting cocaine seeking in extinguished rats. J Neurosci *28*, 6046-6053. 10.1523/JNEUROSCI.1045-08.2008.

Powell, N.J., and Redish, A.D. (2014). Complex neural codes in rat prelimbic cortex are stable across days on a spatial decision task. Front Behav Neurosci *8*, 120. 10.3389/fnbeh.2014.00120.

Powell, N.J., and Redish, A.D. (2016). Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. Nat Commun *7*, 12830. 10.1038/ncomms12830.

Pratt, W.E., and Mizumori, S.J. (2001). Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. Behav Brain Res *123*, 165-183. 10.1016/s0166-4328(01)00204-2.

Ragozzino, M.E., Adams, S., and Kesner, R.P. (1998). Differential involvement of the dorsal anterior cingulate and prelimbic-infralimbic areas of the rodent prefrontal cortex in spatial working memory. Behav Neurosci *112*, 293-303. 10.1037//0735-7044.112.2.293.

Ragozzino, M.E., Detrick, S., and Kesner, R.P. (1999). Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. J Neurosci *19*, 4585-4594.

Rhodes, S.E., and Killcross, S. (2004). Lesions of rat infralimbic cortex enhance recovery and reinstatement of an appetitive Pavlovian response. Learn Mem *11*, 611-616. 10.1101/lm.79704.

Riaz, S., Puveendrakumaran, P., Khan, D., Yoon, S., Hamel, L., and Ito, R. (2019). Prelimbic and infralimbic cortical inactivations attenuate contextually driven discriminative responding for reward. Sci Rep *9*, 3982. 10.1038/s41598-019-40532-7.

Rich, E.L., and Shapiro, M.L. (2007). Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. J Neurosci *27*, 4747-4755. 10.1523/JNEUROSCI.0369-07.2007.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585-590. 10.1038/nature12160.

Rivalan, M., Coutureau, E., Fitoussi, A., and Dellu-Hagedorn, F. (2011). Inter-individual decision-making differences in the effects of cingulate, orbitofrontal, and prelimbic cortex lesions in a rat gambling task. Front Behav Neurosci *5*, 22. 10.3389/fnbeh.2011.00022.

Schmidt, B., Duin, A.A., and Redish, A.D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. J Neurophysiol *121*, 1981-2000. 10.1152/jn.00793.2018.

Schmidt, B., and Redish, A.D. (2021). Disrupting the medial prefrontal cortex with designer receptors exclusively activated by designer drug alters hippocampal sharp-wave ripples and their associated cognitive processes. Hippocampus *31*, 1051-1067. 10.1002/hipo.23367.

Seamans, J.K., Floresco, S.B., and Phillips, A.G. (1995). Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. Behav Neurosci *109*, 1063-1073. 10.1037//0735-7044.109.6.1063.

Seamans, J.K., Lapish, C.C., and Durstewitz, D. (2008). Comparing the prefrontal cortex of rats and primates: insights from electrophysiology. Neurotox Res *14*, 249-262. 10.1007/BF03033814.

Sesack, S.R., Deutch, A.Y., Roth, R.H., and Bunney, B.S. (1989). Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: An anterograde tract-tracing study with Phaseolus vulgaris leucoagglutinin. Journal of Comparative Neurology *290*, 213-242. https://doi.org/10.1002/cne.902900205.

Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron *79*, 217-240. 10.1016/j.neuron.2013.07.007.

Sierra-Mercado, D., Padilla-Coreano, N., and Quirk, G.J. (2011). Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. Neuropsychopharmacology *36*, 529-538. 10.1038/npp.2010.184.

Smith, K.S., and Graybiel, A.M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron *79*, 361-374. 10.1016/j.neuron.2013.05.038.

Smith, K.S., Virkud, A., Deisseroth, K., and Graybiel, A.M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. Proc Natl Acad Sci U S A *109*, 18932-18937. 10.1073/pnas.1216264109.

Steiner, A.P., and Redish, A.D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. Nat Neurosci *17*, 995-1002. 10.1038/nn.3740.

Stokes, M.G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. Neuron *78*, 364-375. 10.1016/j.neuron.2013.01.039.

Sul, J.H., Kim, H., Huh, N., Lee, D., and Jung, M.W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron *66*, 449-460. 10.1016/j.neuron.2010.03.033.

Sweis, B.M., Abram, S.V., Schmidt, B.J., Seeland, K.D., MacDonald, A.W., 3rd, Thomas, M.J., and Redish, A.D. (2018a). Sensitivity to "sunk costs" in mice, rats, and humans. Science *361*, 178-181. 10.1126/science.aar8644.

Sweis, B.M., Larson, E.B., Redish, A.D., and Thomas, M.J. (2018b). Altering gain of the infralimbic-to-accumbens shell circuit alters economically dissociable decision-making algorithms. Proc Natl Acad Sci U S A *115*, E6347-E6355. 10.1073/pnas.1803084115.

Sweis, B.M., Redish, A.D., and Thomas, M.J. (2018c). Prolonged abstinence from cocaine or morphine disrupts separable valuations during decision conflict. Nat Commun *9*, 2521. 10.1038/s41467-018-04967-2.

Sweis, B.M., Thomas, M.J., and Redish, A.D. (2018d). Mice learn to avoid regret. PLoS Biol *16*, e2005853. 10.1371/journal.pbio.2005853.

Takehara-Nishiuchi, K., Morrissey, M.D., and Pilkiw, M. (2020). Prefrontal Neural Ensembles Develop Selective Code for Stimulus Associations within Minutes of Novel Experiences. J Neurosci *40*, 8355-8366. 10.1523/JNEUROSCI.1503-20.2020.

Tervo, D.G.R., Kuleshova, E., Manakov, M., Proskurin, M., Karlsson, M., Lustig, A., Behnam, R., and Karpova, A.Y. (2021). The anterior cingulate cortex directs exploration of alternative strategies. Neuron *109*, 1876-1887 e1876. 10.1016/j.neuron.2021.03.028.

Timme, N.M., and Lapish, C. (2018). A Tutorial for Information Theory in Neuroscience. eNeuro *5*. 10.1523/ENEURO.0052-18.2018.

Tran-Tu-Yen, D.A., Marchand, A.R., Pape, J.R., Di Scala, G., and Coutureau, E. (2009). Transient role of the rat prelimbic cortex in goal-directed behaviour. Eur J Neurosci *30*, 464-471. 10.1111/j.1460-9568.2009.06834.x.

Uylings, H.B., Groenewegen, H.J., and Kolb, B. (2003). Do rats have a prefrontal cortex? Behav Brain Res *146*, 3-17. 10.1016/j.bbr.2003.09.028.

Vidal-Gonzalez, I., Vidal-Gonzalez, B., Rauch, S.L., and Quirk, G.J. (2006). Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. Learn Mem *13*, 728-733. 10.1101/lm.306106.

Vogt, B.A., and Paxinos, G. (2014). Cytoarchitecture of mouse and rat cingulate cortex with human homologies. Brain Struct Funct *219*, 185-192. 10.1007/s00429-012-0493-3.

Voorn, P., Vanderschuren, L.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci *27*, 468-474. 10.1016/j.tins.2004.06.006.

Wallis, J.D., Anderson, K.C., and Miller, E.K. (2001). Single neurons in prefrontal cortex encode abstract rules. Nature *411*, 953-956. 10.1038/35082081.

Walton, M.E., Bannerman, D.M., Alterescu, K., and Rushworth, M.F. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. J Neurosci *23*, 6475-6479.

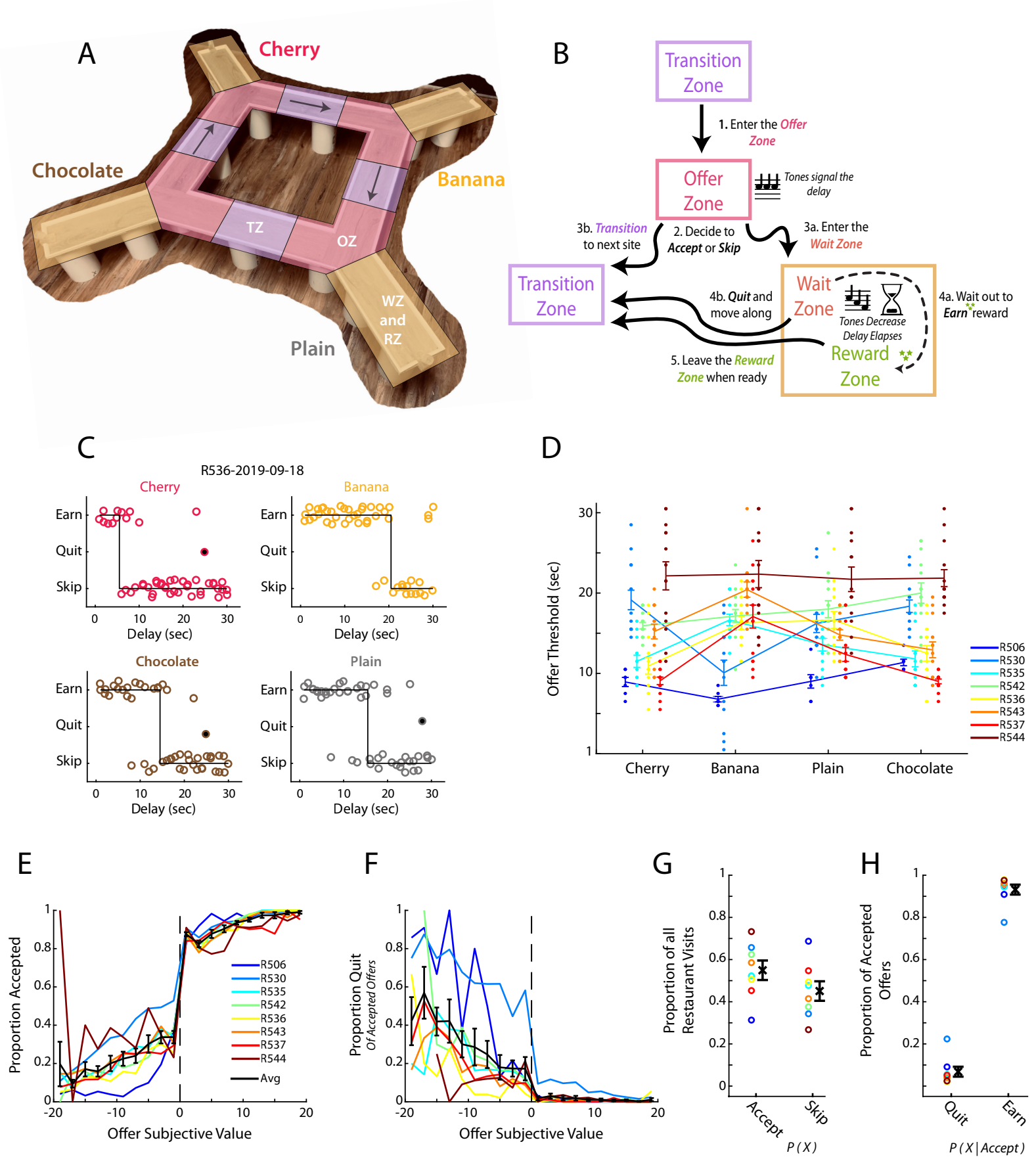
A
Cherry
Chocolate
Banana
TZ
OZ
WZ and RZ
Plain
B
Transition Zone
1. Enter the Offer Zone
Offer Zone
Tones signal the delay
3b. Transition to next site
2. Decide to Accept or Skip
3a. Enter the Wait Zone
Transition Zone
Wait Zone
4b. Quit and move along
Tones Decrease Delay Elapses
4a. Wait out to Earn reward
5. Leave the Reward Zone when ready
Reward Zone
C
R536-2019-09-18
Cherry
Banana
Chocolate
Plain
Earn
Quit
Skip
Delay (sec)
D
Offer Threshold (sec)
Cherry
Banana
Plain
Chocolate
R506
R530
R535
R542
R536
R543
R537
R544
E
Proportion Accepted
Offer Subjective Value
Avg
F
Proportion Quit
Of Accepted Offers
Offer Subjective Value
G
Proportion of all Restaurant Visits
Accept
Skip
P (X)
H
Proportion of Accepted Offers
Quit
Earn
P (X | Accept)

Figure 1: Restaurant Row as a decision-making task

(A) Illustration of the Restaurant Row (RRow) maze. (B) Schematic of the task structure and the decision progression through an individual reward site (restaurant). (C) Example choice responding and thresholds across one behavioral session on RRow. Each circle represents one restaurant visit and is plotted according to the presented delay and the behavioral response. Offers that were quit are shown as filled circles. (D) Thresholds across the four restaurants for each rat in each session (n=8 Rats; 104 Total session). Threshold Means ± SEM for each rat are shown as error bars. (E, F) Proportion of offers that were accepted (E) or Quit (F) as a function of the subjective value of the offer (Value = Threshold – Offer Delay). Data are presented for each rat with the Mean ± SEM across rats shown in black. Note that Quits (F) are calculated as the proportion of offers that were initially accepted. (G) Proportion of all restaurant visits that were Accepted or Skipped. (H) Proportion of Accepted offers that were subsequently Quit or Earned.

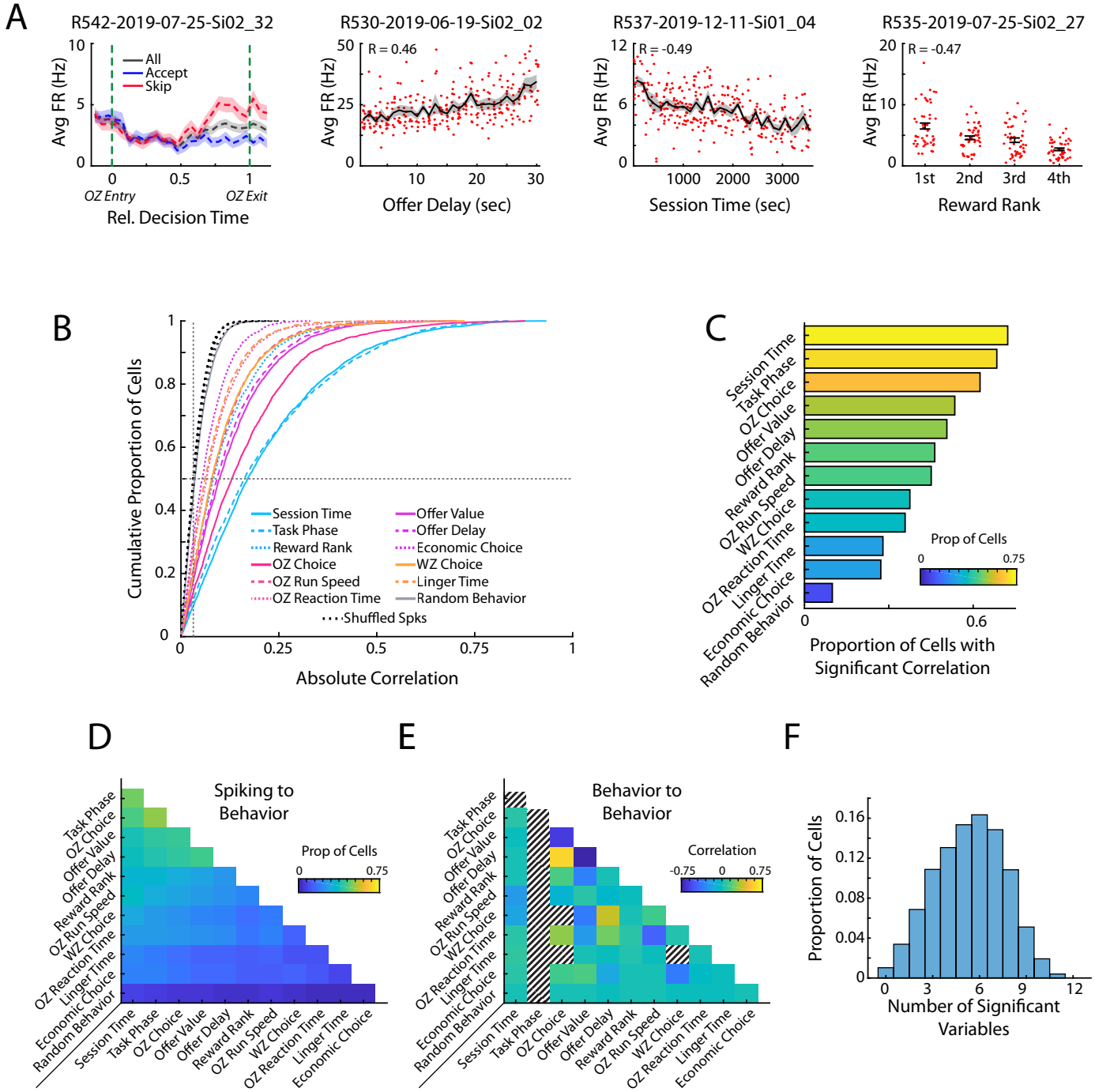
A
R542-2019-07-25-Si02_32
R530-2019-06-19-Si02_02
R537-2019-12-11-Si01_04
R535-2019-07-25-Si02_27
Avg FR (Hz)
All
Accept
Skip
OZ Entry
OZ Exit
Rel. Decision Time
R = 0.46
Offer Delay (sec)
R = -0.49
Session Time (sec)
R = -0.47
1st
2nd
3rd
4th
Reward Rank
B
Cumulative Proportion of Cells
Absolute Correlation
Session Time
Task Phase
Reward Rank
OZ Choice
OZ Run Speed
OZ Reaction Time
Offer Value
Offer Delay
Economic Choice
WZ Choice
Linger Time
Random Behavior
Shuffled Spks
C
Prop of Cells
Proportion of Cells with Significant Correlation
D
Spiking to Behavior
E
Behavior to Behavior
Correlation
F
Proportion of Cells
Number of Significant Variables

Figure 2: Prefrontal cells respond to task relevant variables

(A) Four example mPFC cells that respond to offer zone choice, offer delay, session time, and reward rank. (B) Cumulative density functions of the unsigned correlation between firing rate of mPFC cells (n=3017) and various task relevant behaviors. Chance values are computed by correlating firing rates to randomly drawn values (Random Behavior) and by time shifting spiking of each cell and correlating the dissociated spiking data to the behaviors of interest (Shuffled Spks). Firing rates in mPFC were significantly more correlated to each behavioral variable than to the respective chance distributions. (C) Proportion of mPFC cells that were significantly correlated to each behavioral variable. (D) Proportion of the mPFC cell population that was significantly correlated to both of a pair of behavioral variables. (E) Different pairs of behavioral variables were correlated to each other. The average correlation across the 104 behavioral sessions is shown for each pair of variables. Note that Task Phase varied within each restaurant visit and thus could not be compared to other variables. Additionally, some combinations of variables were conditionally impossible to co-occur and thus could not be correlated (e.g. OZ Choice and WZ Choice). (F) Individual mPFC cells were often significantly correlated to multiple behaviors. Distribution of multimodal coding across the mPFC cell population. Each bar indicates the proportion of mPFC cells with firing rates that were significantly correlated to the indicated number of behavioral variables.

Figure 3: Transfer entropy reveals subregions within mPFC

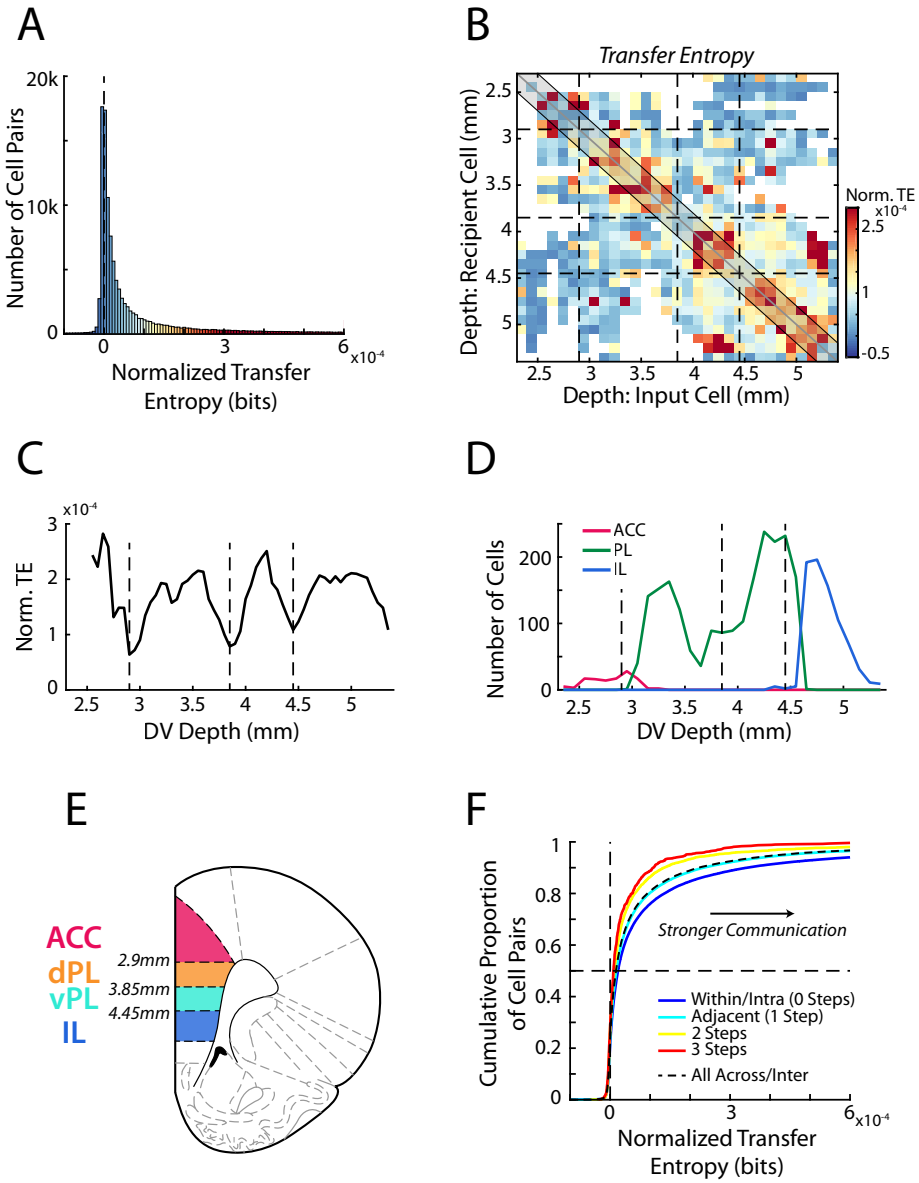

Figure 3: Transfer entropy reveals subregions within mPFC

(A) Transfer entropy (TE) between simultaneously recorded pairs of cells (n=107,418 pairs) while rats performed RRow. Data were normalized based on shuffled spike times of the input cell meaning that TE values above zero (dashed line) reflect cell interactions that are stronger than expected by chance. Across the population, TE was significantly higher than chance. Data along the x-axis are color coded according to the scale in panel B. (B) Mean normalized TE (norm. TE) between pairs of cells as a function of the DV location of the input and recipient cells. Black dashed lines represent the functional boundaries between mPFC subregions computed in panel C. (C) Quantification of TE between cell pairs within 200 µm of the identity diagonal (shaded region in B). Data are the mean TE of bins within 200µm of each DV depth. Dashed lines denote the local minima of the curve and represent derived transition bounds between subregions. (D) Number of recorded ACC, PL, and IL cells at each DV position, as classified based purely on anatomical information as defined in the Paxinos and Watson atlas. TE derived bounds are repeated from panel C. (E) A schematic of four subdivisions of mPFC based on the boundaries derived from analysis of functional communication (TE). (F) Cumulative distributions of normalized TE values between pairs of cells. Data are grouped according to the relative proximity (number of steps) between the subregions of the two cells. Black dashed line groups all inter-regional pairs regardless of distance. Note that a more rightward distribution corresponds to overall stronger TE.

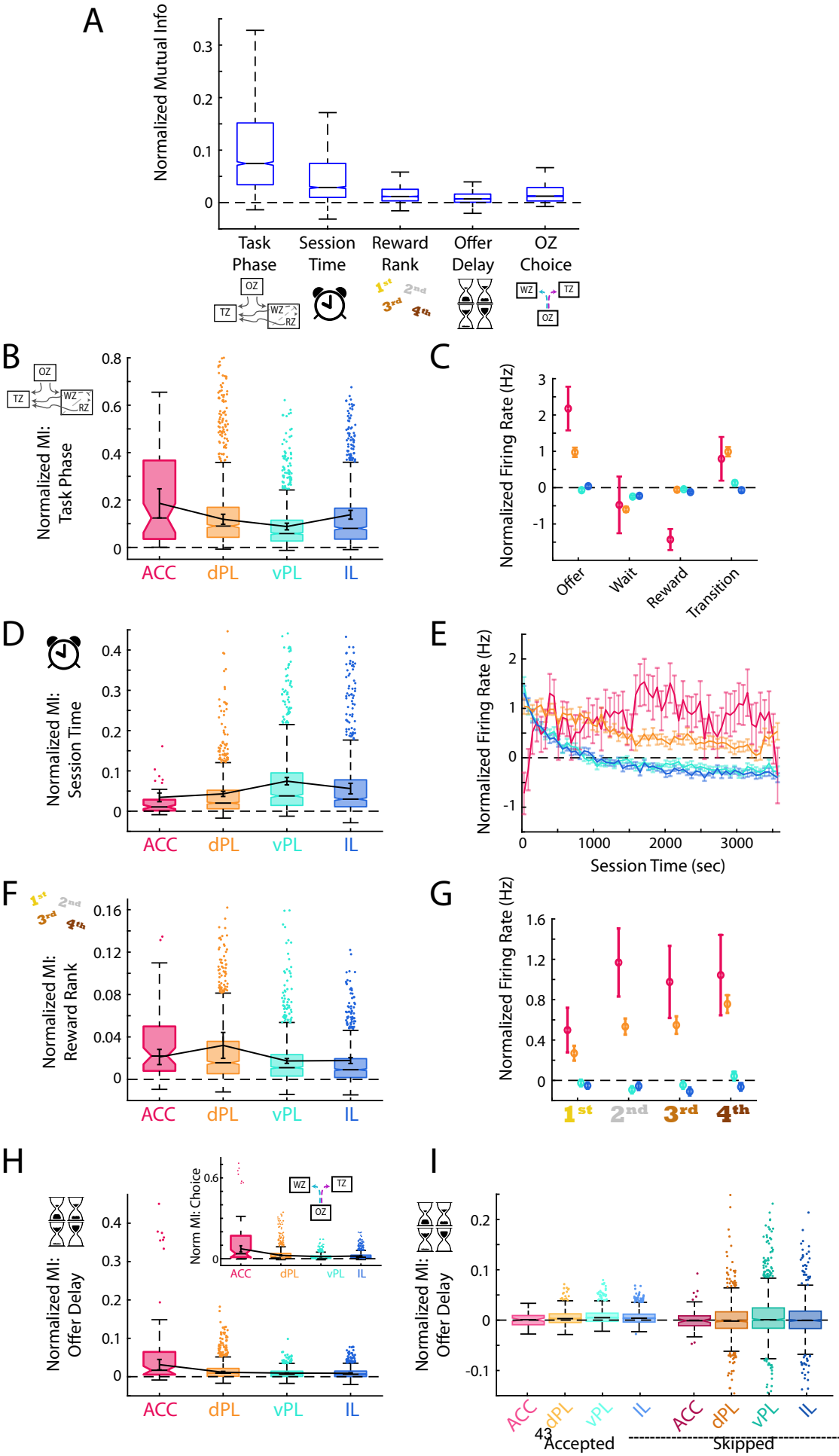
A
Normalized Mutual Info
0.3
0.2
0.1
0
Task Phase
Session Time
Reward Rank
Offer Delay
OZ Choice
OZ
TZ
WZ
RZ
1st
2nd
3rd
4th
B
Normalized MI: Task Phase
0.8
0.6
0.4
0.2
0
ACC
dPL
vPL
IL
C
Normalized Firing Rate (Hz)
3
2
1
0
-1
Offer
Wait
Reward
Transition
D
Normalized MI: Session Time
0.4
0.3
0.2
0.1
0
ACC
dPL
vPL
IL
E
Normalized Firing Rate (Hz)
2
1
0
-1
0
1000
2000
3000
Session Time (sec)
F
Normalized MI: Reward Rank
0.16
0.12
0.08
0.04
0
ACC
dPL
vPL
IL
G
Normalized Firing Rate (Hz)
1.6
1.2
0.8
0.4
0
1st
2nd
3rd
4th
H
Normalized MI: Offer Delay
0.4
0.3
0.2
0.1
0
Norm MI: Choice
0.6
0
ACC
dPL
vPL
IL
I
Normalized MI: Offer Delay
0.2
0.1
0
-0.1
ACC
dPL
vPL
IL
ACC
dPL
vPL
IL
Accepted
Skipped

Figure 4: Generalized task variables are differentially represented

(A) Mutual information (MI) calculations for mPFC cells between the task phase, elapsed session time, reward site ranking, offer delay, and choice made in the offer zone. For each cell, MI values were normalized based on data generated with shuffled spike times. For clarity, distribution outliers (beyond 1.5x IQR) are not displayed. (B, D, F, H) Normalized MI to task phase (B), session time (D), reward rank (F), and offer delay (H) for each mPFC subregion. Boxplots with outliers show the data for all recorded cells in each subregion. Data were also averaged within each rat and the Mean ± SEM across rats is shown overlayed in black. The inset in panel (H) shows the corresponding MI data for offer zone choice. Note that because choice is based largely on the presented delay these two data are highly correlated. (C, E, G) Mean ± SEM firing rate of cells in each mPFC subregion across the four task phases (C), over the session duration (E), and for each restaurant rank (G). Note that firing rates were normalized based on shuffled spike times. Reward rank was determined in each session by sorting restaurants by decision threshold with the largest threshold (willing to wait longest to earn a reward) ranked 1st. (I) MI for offer delay split by whether rats accepted or skipped in the OZ. Note that after accounting for the choice made, MI for delay is largely eliminated suggesting that the neural responses are more associated with choice than delay.

Figure 5: Activity in mPFC strongly encodes the identity of the upcoming choice

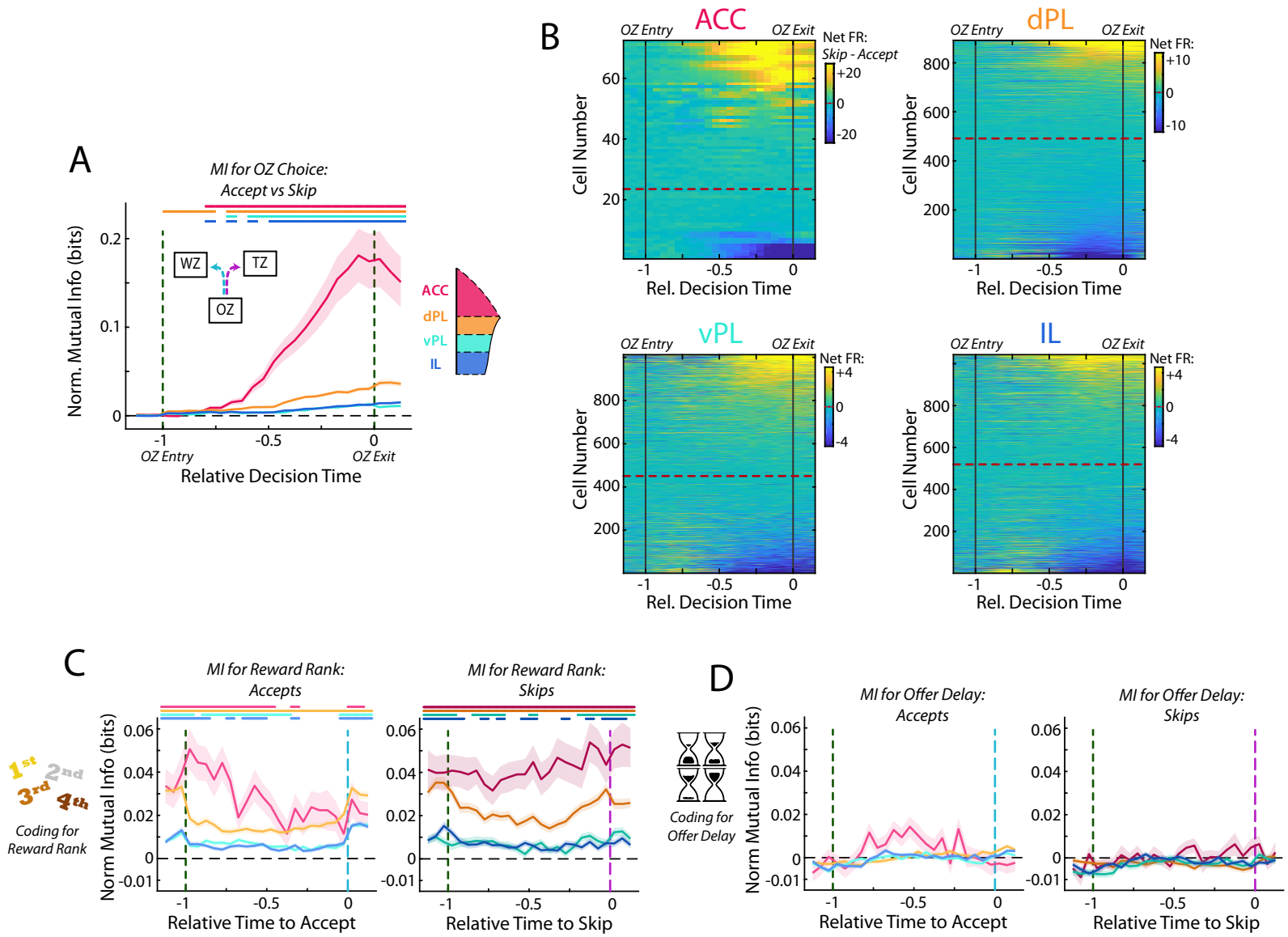

Figure 5: Activity in mPFC strongly encodes the identity of the upcoming choice

(A) Mutual information (MI) between firing rate and OZ choice (accept vs skip) for cells in each mPFC subregion. Note that durations in the OZ are normalized for each lap to yield relative duration for analysis. (B) Net firing rate of mPFC cells between accept and skip decisions (Skip FR – Accept FR). For each subregion, cells are sorted according to the net firing rate after entering the OZ. Color scales are identified for each subregion but note the change in scale between subregions. (C) Mutual information between firing rate of mPFC cells and restaurant ranking and (D) between firing rate and the presented offer delay. Data are segregated according to the decision that was made. For (A,C,D), data are shown as the Mean ± SEM across cells in each mPFC subregion. All data were normalized against analyses performed using shuffled spike times. Time bins with a significant MI response (one-tailed Wilcoxon Signed-rank, $\alpha = 0.05$ after multiple comparison correction; $n = 26$ time bins) are identified by colored bars above each plot.

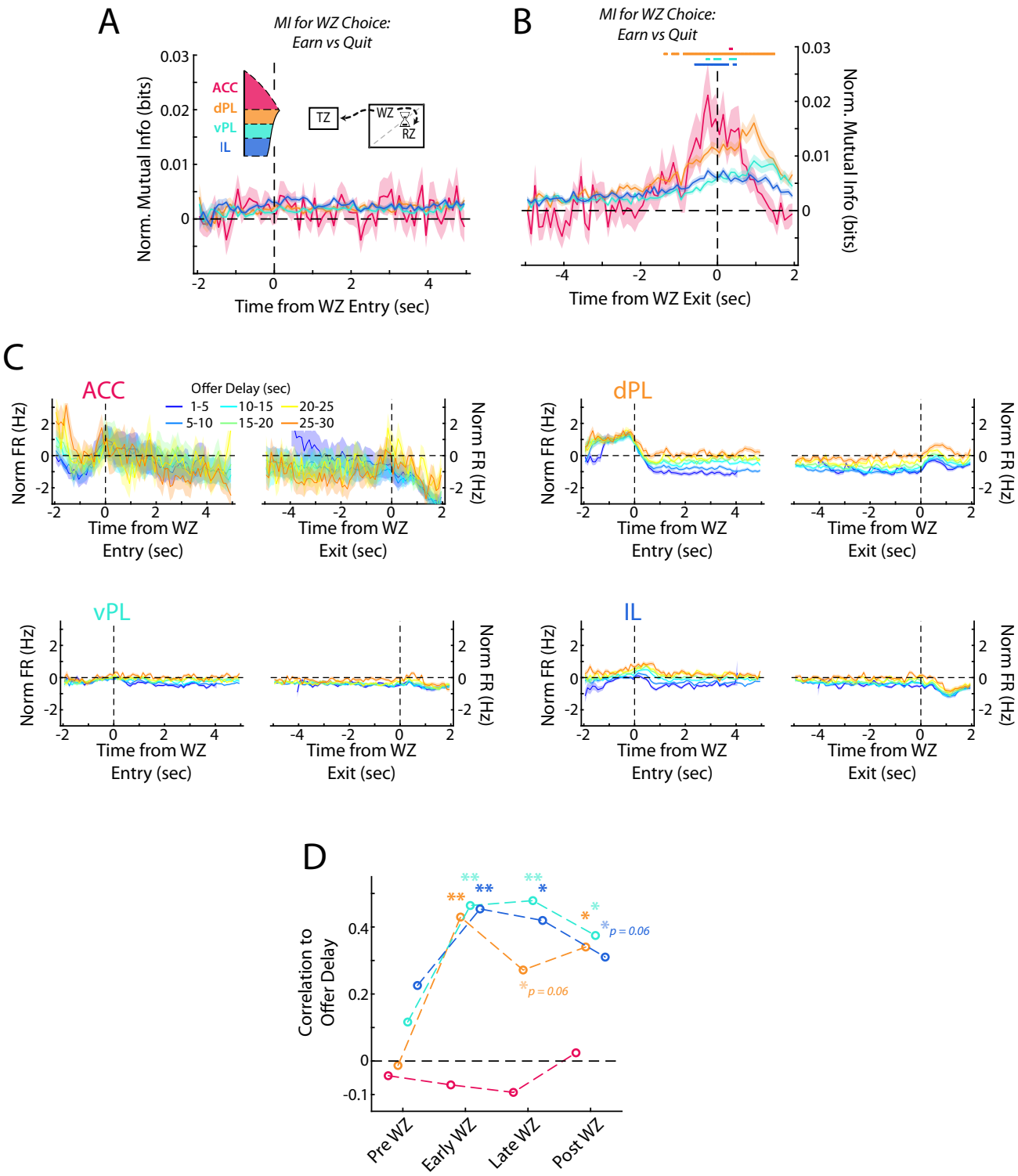
A
MI for WZ Choice:
Earn vs Quit
Norm. Mutual Info (bits)
ACC
dPL
vPL
IL
TZ
WZ
RZ
Time from WZ Entry (sec)
B
MI for WZ Choice:
Earn vs Quit
Norm. Mutual Info (bits)
Time from WZ Exit (sec)
C
ACC
Offer Delay (sec)
1-5
5-10
10-15
15-20
20-25
25-30
Norm FR (Hz)
Time from WZ Entry (sec)
Time from WZ Exit (sec)
dPL
vPL
IL
D
Correlation to Offer Delay
p = 0.06
Pre WZ
Early WZ
Late WZ
Post WZ

Figure 6: Re-Evaluative decisions are represented in mPFC

(A, B) Mutual information (MI) between firing rate and WZ choice (earn vs quit) for cells in each mPFC subregion at the beginning and end of delay period. Data are show in (A) aligned to WZ entry and in (B) aligned to WZ exit. Time bins with a significant MI response (one-tailed Wilcoxon Signed-rank, $\alpha = 0.05$ after multiple comparison correction; n = 70 time bins) are identified by colored bars above each plot. (C) Average firing rate of mPFC cells for each subregion separated by the offer delay that had been accepted. Data are aligned to the beginning of the WZ and the end of the WZ. (D) Correlation between average firing rate across mPFC cells and the offer delay that had been accepted. Average firing rate profiles were computed across all cells for each rat independently in each offer delay bin (up to 48 values: 8 rats x 6 delay bins). These firing rates were then correlated against offer delay during four time windows: Pre WZ, Early WZ, Late WZ, and Post WZ. See methods for additional details. For (A-C), data are shown as the Mean ± SEM across cells in each mPFC subregion. All data were normalized against analyses performed using shuffled spike times.

Figure 7: Responses to reward delivery differ across mPFC subregions

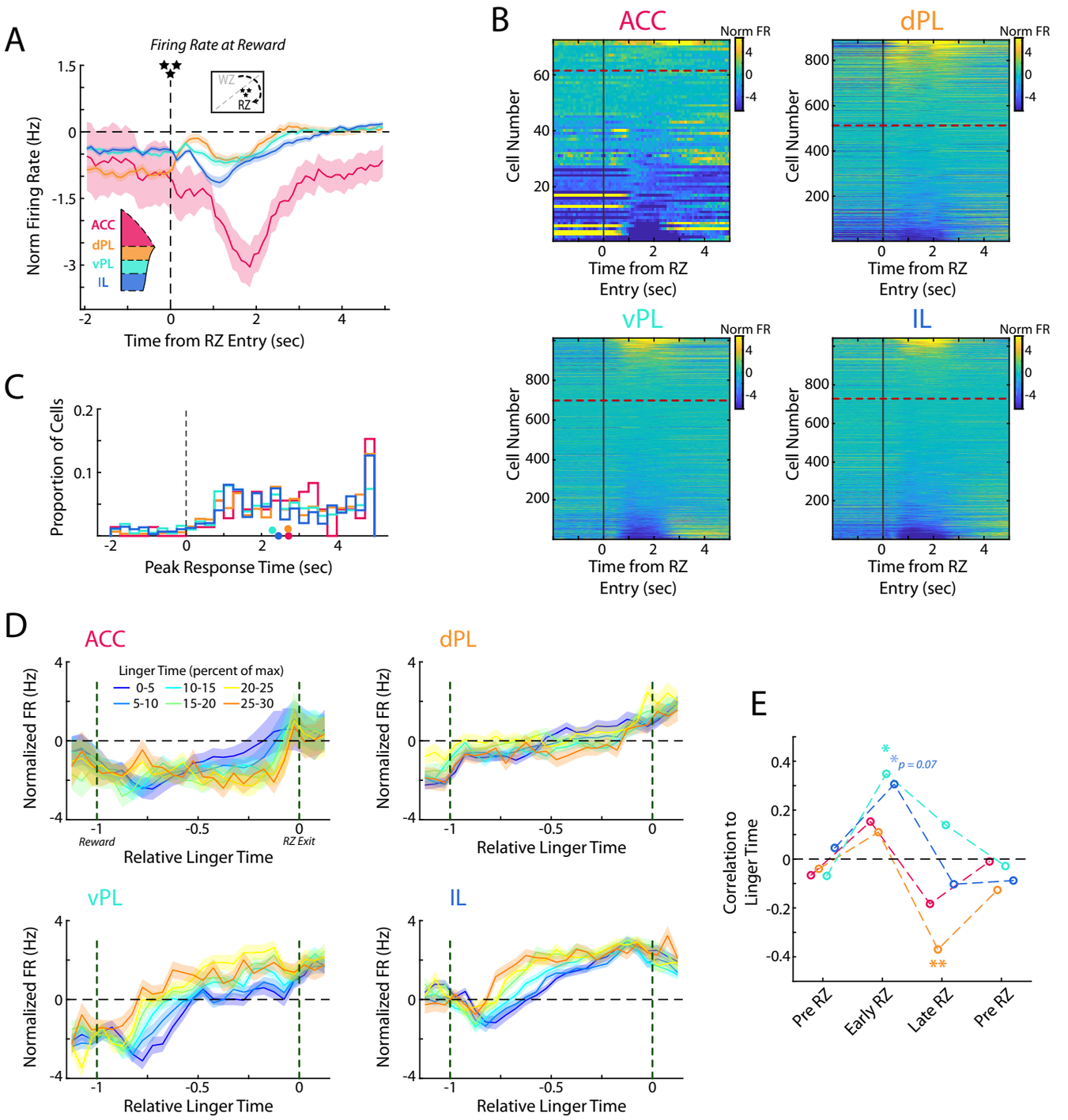

Figure 7: Responses to reward delivery differ across mPFC subregions

(A) Average firing rate for cells in each mPFC subregion as rats entered the Reward Zone and received their reward. (B) Distribution of all cell responses for each subregion at RZ entry. Cells are sorted according to the average activity from 0.5-3 sec after RZ entry. (C) Time of the peak cellular response in the first portion of the RZ period (-2 – 5 sec from RZ entry). Distributions of all cells are shown as histograms and mean response times for each subregion are indicated with a colored dot. (D) Average firing rate of mPFC cells for each subregion separated by the amount of time spent in the RZ (linger time). Data are grouped according to the percentage of maximal linger time computed independently for each session. Note that duration in the RZ is normalized across laps to produce a relative duration for analysis. (E) As for Fig 6D, correlations between average firing rate across mPFC cells and the linger time percentage. Average rate profiles were computed for each rat in each linger time bin (up to 48 values) and correlated to percentage of linger time. Average firings rates and correlations were computed in four time windows: Pre RZ, Early RZ, Late RZ, Post RZ. See methods for additional details. For (A, D, E), data are shown as the Mean ± SEM across cells in each mPFC subregion. All data were normalized against analyses performed using shuffled spike times.

Figure 8: Functional activity in the medial prefrontal cortex

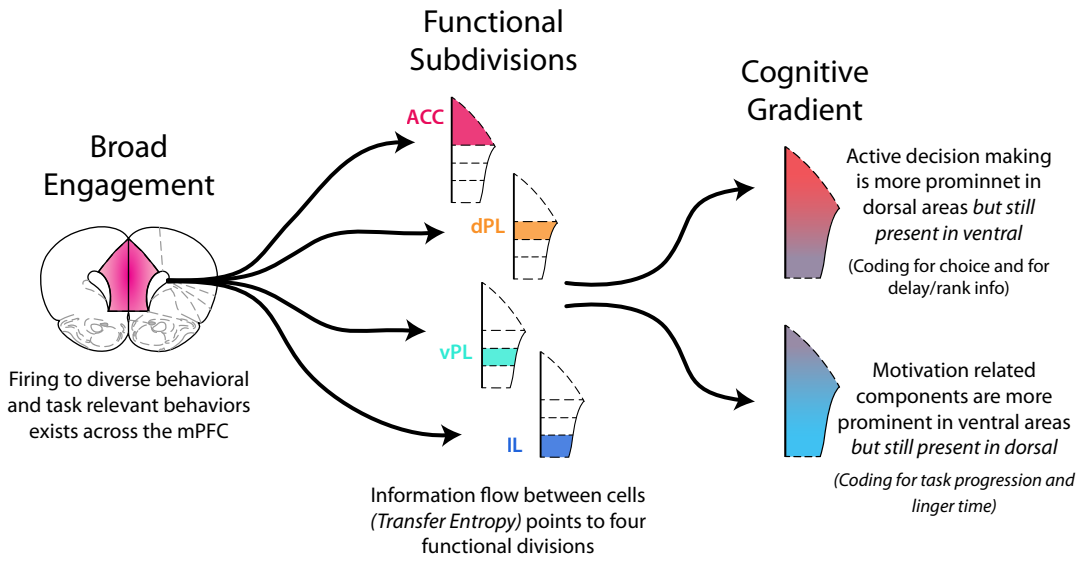

Figure 8: Functional activity in the medial prefrontal cortex

Functional activity in mPFC can be described in multiple ways. There was broad, distributed engagement across subregions for many behavioral variables. Analysis of functional coupling between cells using TE revealed distinct subregions. In many cases functional activity exhibited a gradient along the dorso-ventral axis. Neural engagement during active decision processing was more prominent in dorsal regions while motivation related components were more strongly represented in ventral areas.

Figure 1 – Animation 1: 3D Anatomical location of recorded mPFC cells

3-dimensional reconstruction of the mPFC and the anatomical location of all cells included in the study. Each dot corresponds to a single unit (n=3017) and are colored according to rat identity (n=8). Shaded regions indicate subregions of the mPFC as defined in the Paxinos and Watson atlas. See Fig 1-Supplament2,D for legend keys.

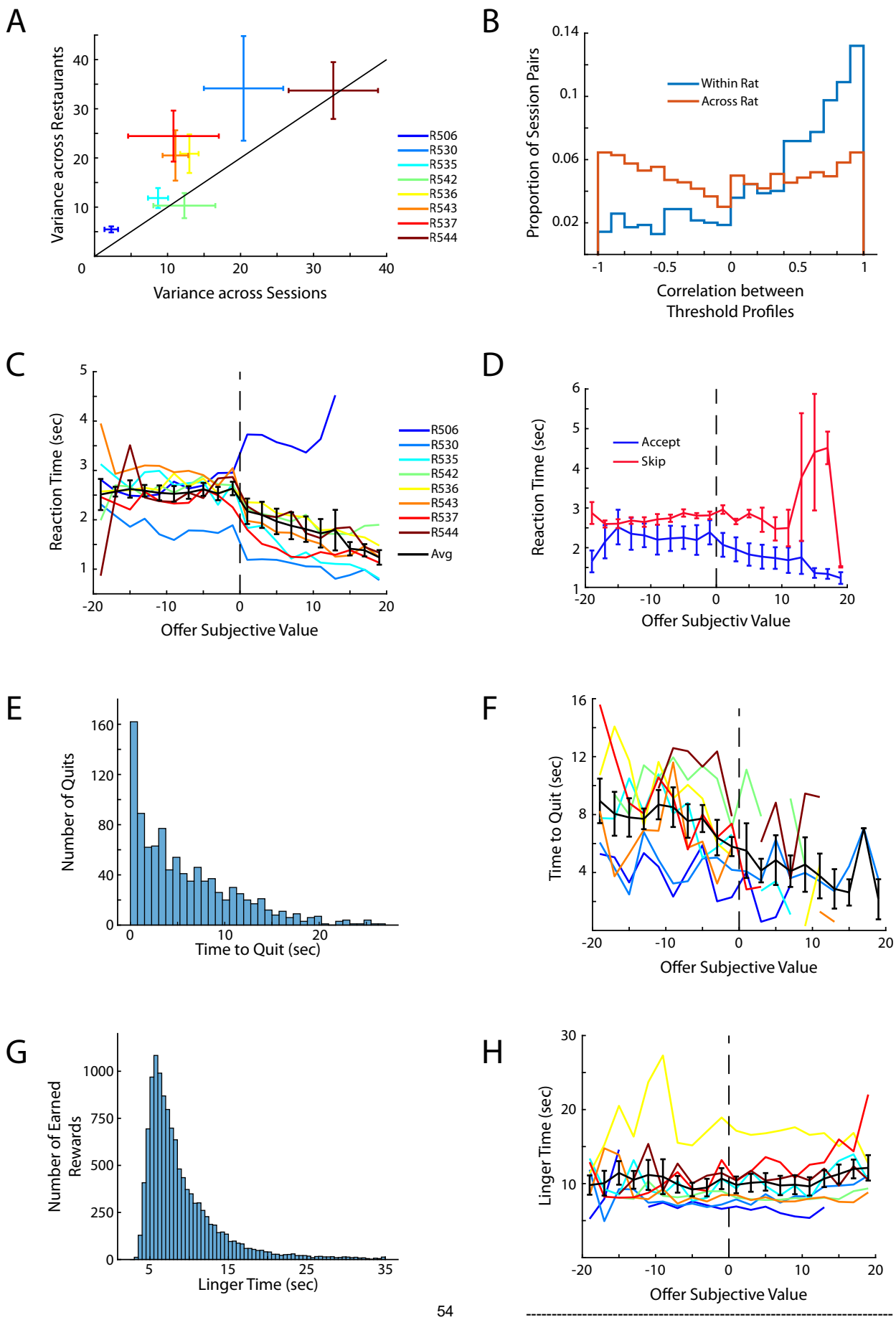
A
Variance across Restaurants
Variance across Sessions
R506
R530
R535
R542
R536
R543
R537
R544
B
Proportion of Session Pairs
Within Rat
Across Rat
Correlation between Threshold Profiles
C
Reaction Time (sec)
Offer Subjective Value
Avg
D
Reaction Time (sec)
Accept
Skip
Offer Subjectiv Value
E
Number of Quits
Time to Quit (sec)
F
Time to Quit (sec)
Offer Subjective Value
G
Number of Earned Rewards
Linger Time (sec)
H
Linger Time (sec)
Offer Subjective Value

Figure 1 – Supplement 1: Additional behavior on RRow

(A) Average variance in threshold across sessions plotted against the average variance in threshold across restaurants. Data are plotted as Mean ± SEM of the threshold variance for each rat. Rats showed more variance in thresholds across restaurants than across sessions indicating that rats exhibited preferences across reward flavors but that the subjective assessment of each flavor was comparably stable across successive sessions. (B) Correlation of restaurant thresholds across behavioral sessions. Threshold profiles across sessions for a given rat tended to be strongly correlated whereas thresholds between rats were unrelated. (C) Average time in the Offer Zone (reaction time) before making a decision and exiting the OZ. Data are presented for each rat with the Mean ± SEM across rats shown in black. (D) Reaction time in the OZ split between accept and skip decisions. (E) Histogram of the time to quit an accepted offer in the WZ. Note that, by definition, time to quit must be less than accepted delay. (F) Average time to quit as a function of the subjective value of the accepted offer. (G) Histogram of the total time spent in the RZ (linger time) after earning a reward before rats left the restaurant and transitioned to the next site. (H) Linger time as a function of subjective value.

Figure 1 - Supplement 2: Localization of single unit recordings from medial prefrontal cortex

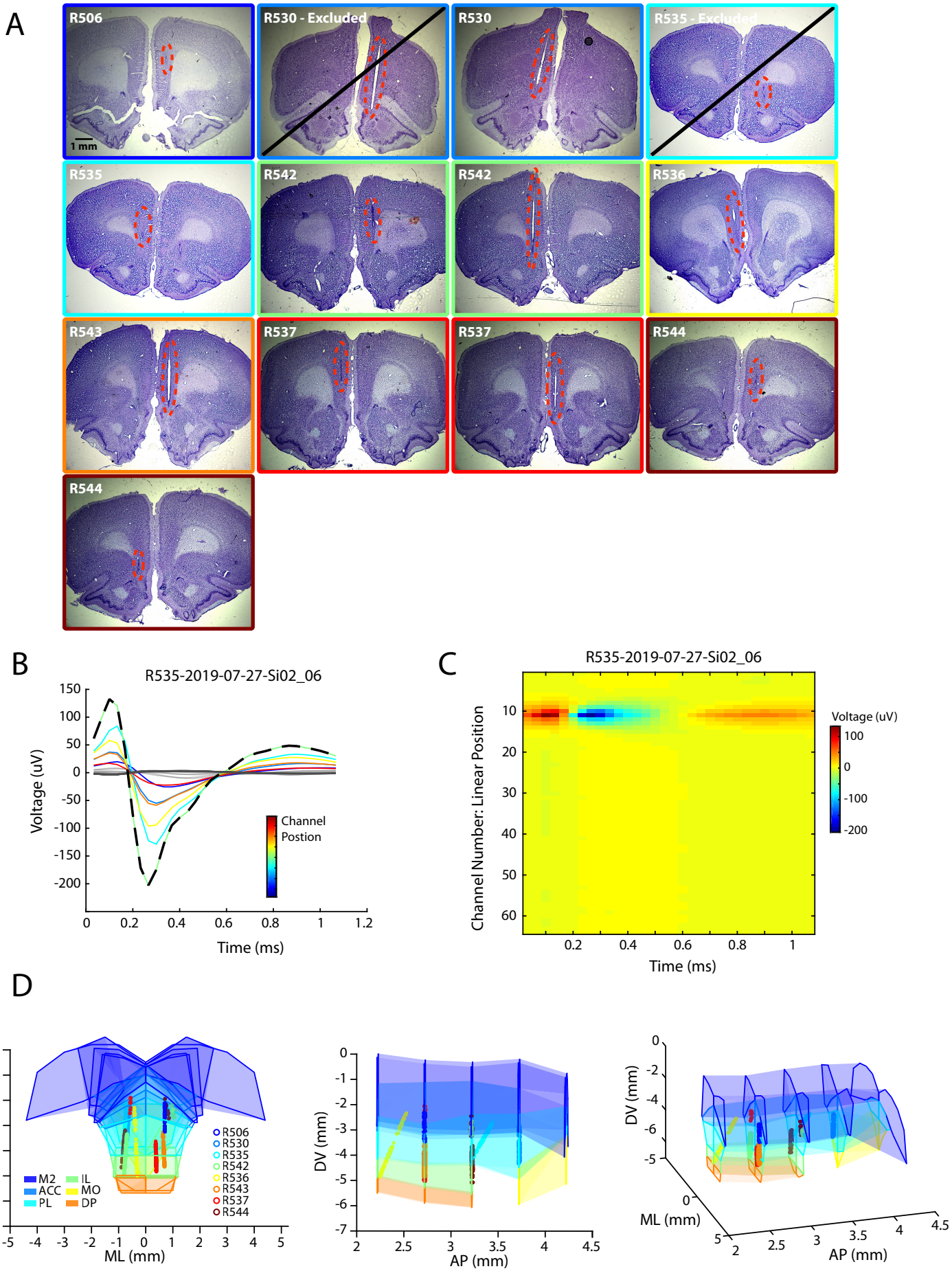

Figure 1 – Supplement 2: Localization of single unit recordings from medial prefrontal cortex

(A) Histological sections of silicon probe trajectories through the mPFC. Note that 2 probes were located outside of the mPFC and were excluded from subsequent analysis. (B,C) Line and image plots of the waveform of an example cell recorded across the 64 linear channels of a silicon probe. For the line plot the 8 largest amplitude channels are colored according to their sequential positions on the probe. The channel with the largest amplitude waveform (black dashed line) can be identified and treated as the location most proximal to the cell body of the recorded unit. (D) Static views of the 3-dimensional reconstruction of the mPFC and the anatomical location of all cells included in the study. Each dot corresponds to a single unit (n=3017) and are colored according to rat identity (n=8). Shaded regions indicate subregions of the mPFC as defined in the Paxinos and Watson atlas.

Figure 1 - Supplement 3: Basic spiking characteristics of medial prefrontal cells

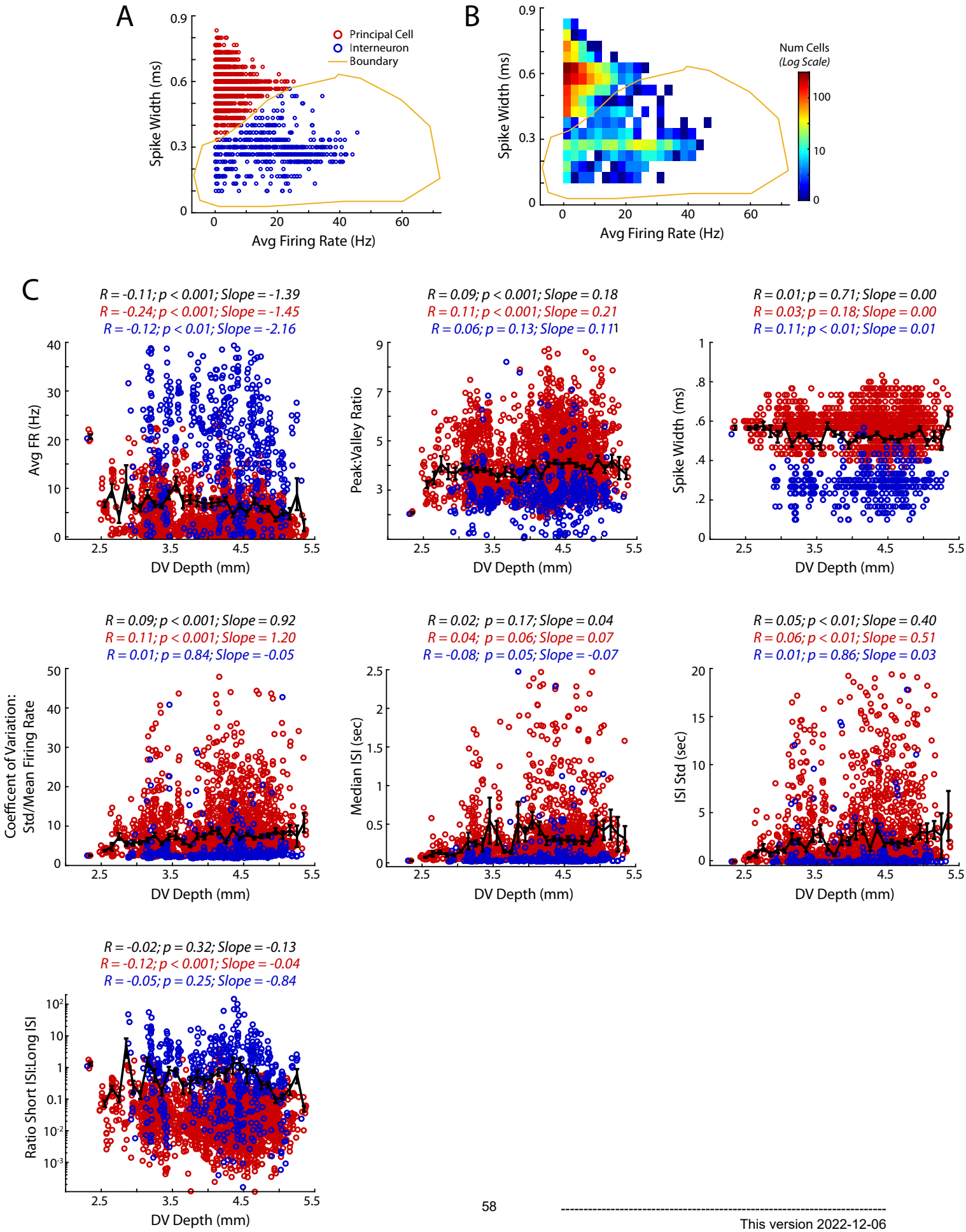

Figure 1 – Supplement 3: Basic spiking characteristics of medial prefrontal cells

(A,B) Segregation of Principal cells and Interneurons in mPFC based on firing rate and waveform shape. Data are shown as points for individual cells (A), or as a density heat map (B). (C) Basic spiking parameters of mPFC cells as a function of DV location of each cell (see Methods for details about each metric). Each cell is colored according to the principal/interneuron designation derived in panel A. For each spiking parameter the correlation coefficient between the metric and DV location, its p-value, and linear slope are reported above the respective plot. Statistics for each plot are reported for all cells (black), and for principal cells (red) and interneurons (blue) separately.

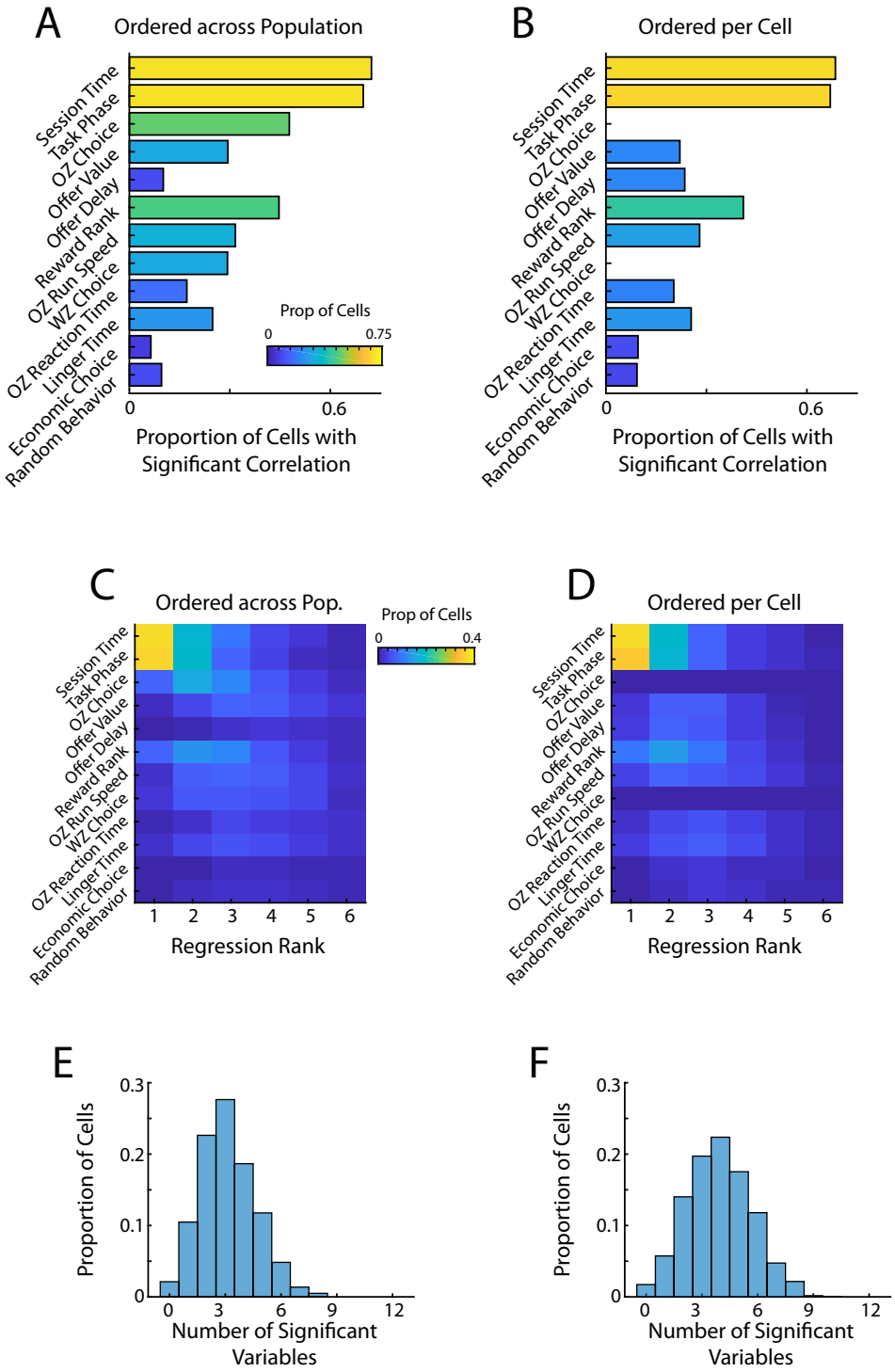
A
Ordered across Population
Session Time
Task Phase
OZ Choice
Offer Value
Offer Delay
Reward Rank
OZ Run Speed
WZ Choice
OZ Reaction Time
Linger Time
Economic Choice
Random Behavior
Prop of Cells
0
0.75
0
0.6
Proportion of Cells with Significant Correlation
B
Ordered per Cell
Session Time
Task Phase
OZ Choice
Offer Value
Offer Delay
Reward Rank
OZ Run Speed
WZ Choice
OZ Reaction Time
Linger Time
Economic Choice
Random Behavior
0
0.6
Proportion of Cells with Significant Correlation
C
Ordered across Pop.
Prop of Cells
0
0.4
Session Time
Task Phase
OZ Choice
Offer Value
Offer Delay
Reward Rank
OZ Run Speed
WZ Choice
OZ Reaction Time
Linger Time
Economic Choice
Random Behavior
1 2 3 4 5 6
Regression Rank
D
Ordered per Cell
Session Time
Task Phase
OZ Choice
Offer Value
Offer Delay
Reward Rank
OZ Run Speed
WZ Choice
OZ Reaction Time
Linger Time
Economic Choice
Random Behavior
1 2 3 4 5 6
Regression Rank
E
Proportion of Cells
0 0.1 0.2 0.3
0 3 6 9 12
Number of Significant Variables
F
Proportion of Cells
0 0.1 0.2 0.3
0 3 6 9 12
Number of Significant Variables

Figure 2 – Supplement 1: Stepwise regression of spiking to behavior

Detection of behavioral coding using two stepwise regression approaches (A, C, E) The stepwise order of behavioral regressors was dictated externally based on the overall prevalence in the baseline correlation analysis (Fig 2C). The same ordering was used for all cells in the population. (B, D, F) The stepwise order of behavioral regressors was determined independently for each cell in order of explanatory power. (A, B) Proportion of cells with a significant regression weight for each behavioral variable when using a stepwise regression based approach. (C, D) Breakdown of significant behavioral regressors split by the explanatory rank in the stepwise regressor. A variable of rank 1 would be the most explanatory and a rank 2 would be the next most explanatory. Note that the regressor rank is not always equivalent to the step of the regression at which the variable was added. (E, F) Number of significant explanatory variables (multimodal coding) for mPFC cells when using a stepwise regression. Note that by using a regression based approach this represents mathematically orthogonal variables.

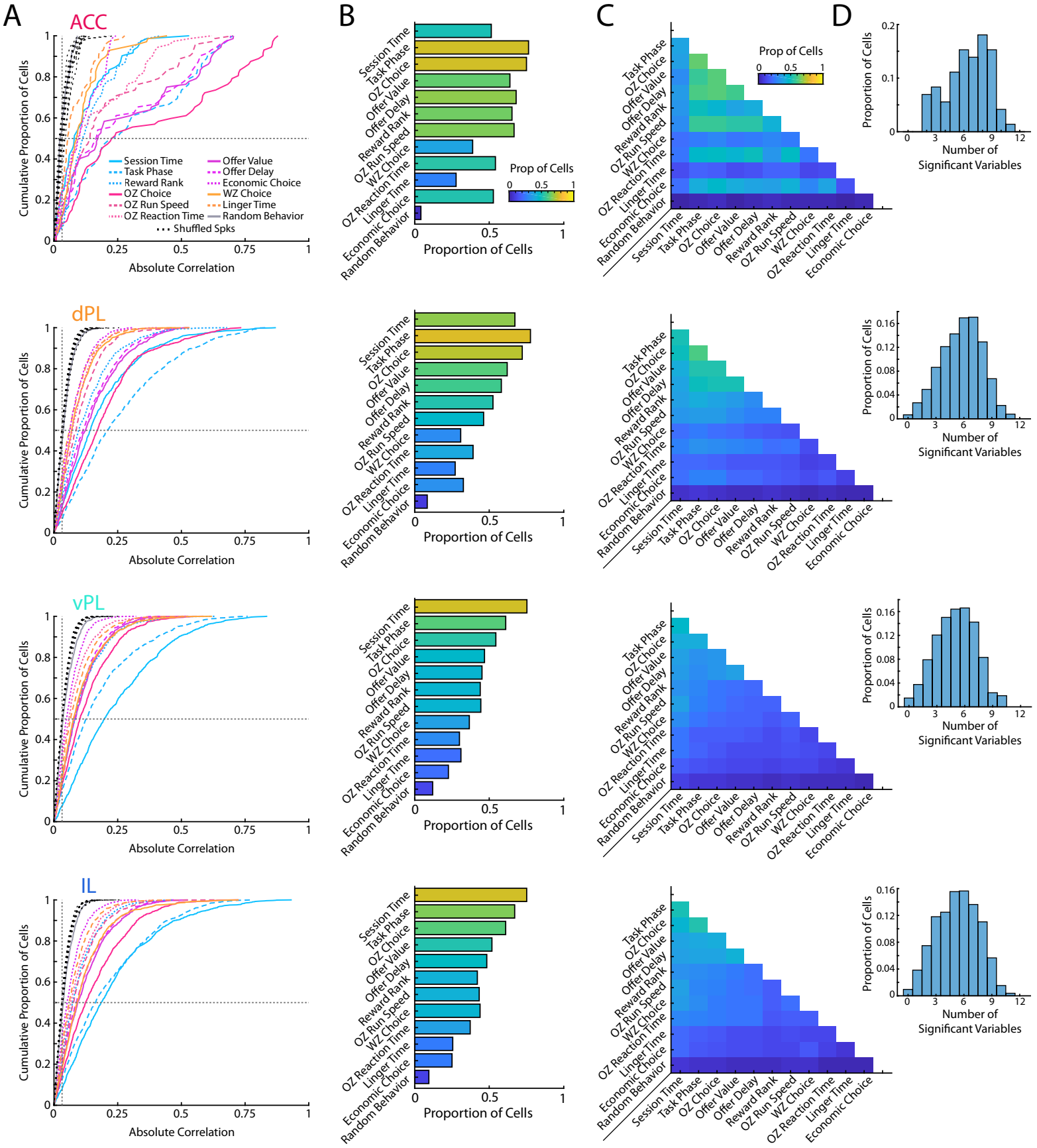
A
ACC
dPL
vPL
IL
Cumulative Proportion of Cells
Absolute Correlation
Session Time
Task Phase
Reward Rank
OZ Choice
OZ Run Speed
OZ Reaction Time
Offer Value
Offer Delay
Economic Choice
WZ Choice
Linger Time
Random Behavior
Shuffled Spks
B
Proportion of Cells
Prop of Cells
C
D
Number of Significant Variables

Figure 2 – Supplement 2: Segregation of prefrontal responses by subregion

Data are presented as in Fig 2B, C, D, F but here segregated by mPFC subregion (ACC, dPL, vPL, IL) defined using functional communication and TE. The first column (A) shows cumulative density functions of the unsigned correlation between firing rate of mPFC cells and various task relevant behaviors. The second column (B) shows the proportion of cells that were significantly correlated to each behavioral variables. The third column (C) shows the proportion of cells from each subregion that were significantly correlated to pairs of behavioral variables. The fourth column (D) shows the distribution of multimodal coding for cells in each subregion.

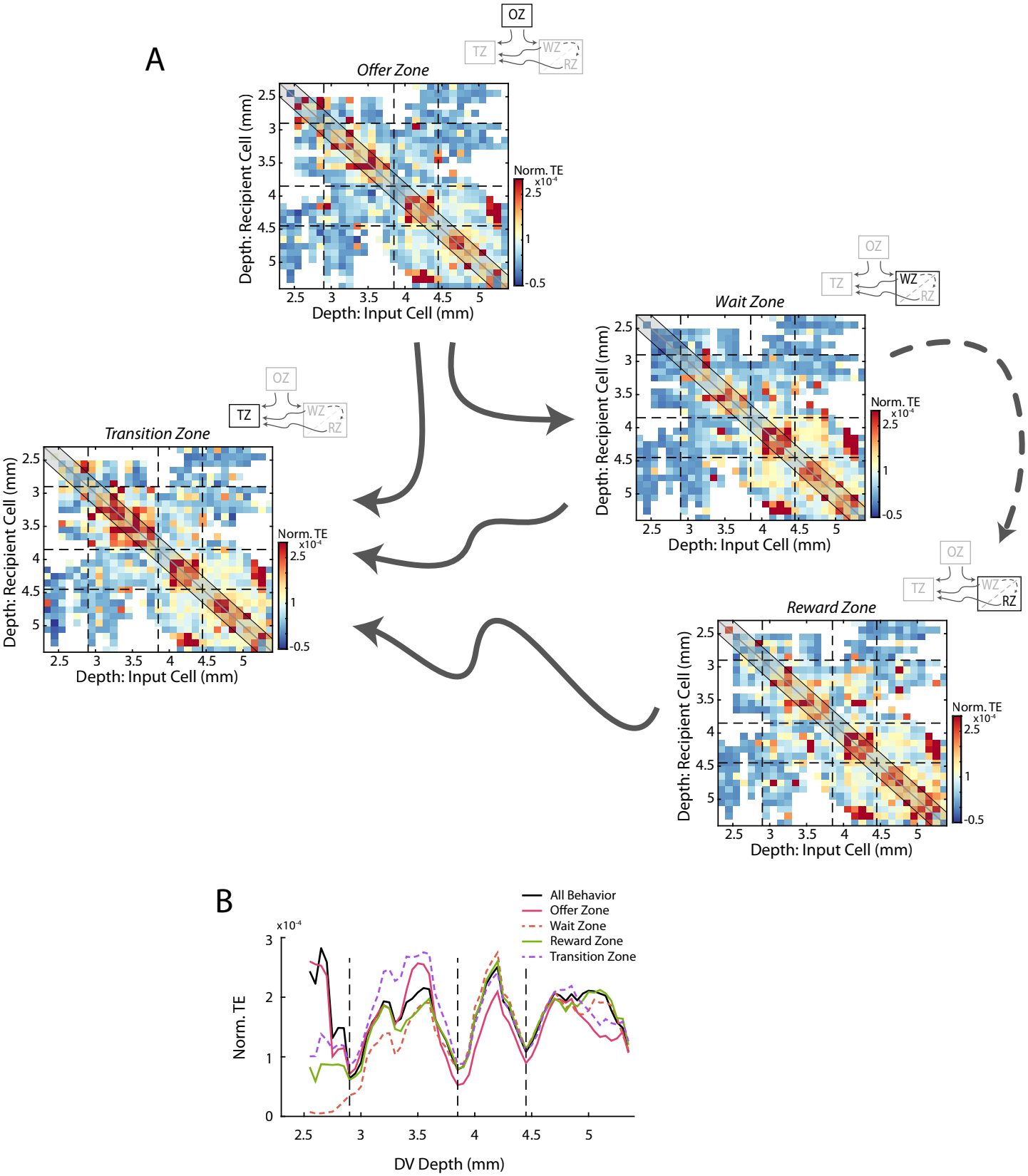
A
Offer Zone
Wait Zone
Transition Zone
Reward Zone
OZ
TZ
WZ
RZ
Depth: Recipient Cell (mm)
Depth: Input Cell (mm)
Norm. TE
x10⁻⁴
2.5
1
-0.5
B
All Behavior
Offer Zone
Wait Zone
Reward Zone
Transition Zone
Norm. TE
DV Depth (mm)

Figure 3 – Supplement 1: Transfer entropy analyses during different task phases

(A) As in Fig 3B: Mean normalized transfer entropy (norm TE) between pairs of cells as a function of the DV location of the input and recipient cells. Data are segregated according to task phase (Offer Zone, Wait Zone, Reward Zone, Transition Zone) including only time points during which rats were in the respective task phase. Black dashed lines represent the functional boundaries between mPFC subregions and are reproduced here from Fig 3 and calculations from the full data set. TE profiles were largely comparable across task phases (all pairwise correlations $r > 0.66$). (B) Quantification of TE between cells pairs around the identity diagonal as in Fig 3C. Data from the full task (Fig 3C) is reproduced here in black.

Figure 3 - Supplement 2: Subsampling of mPFC recordings in TE calcuations and subregion bound identificaiton

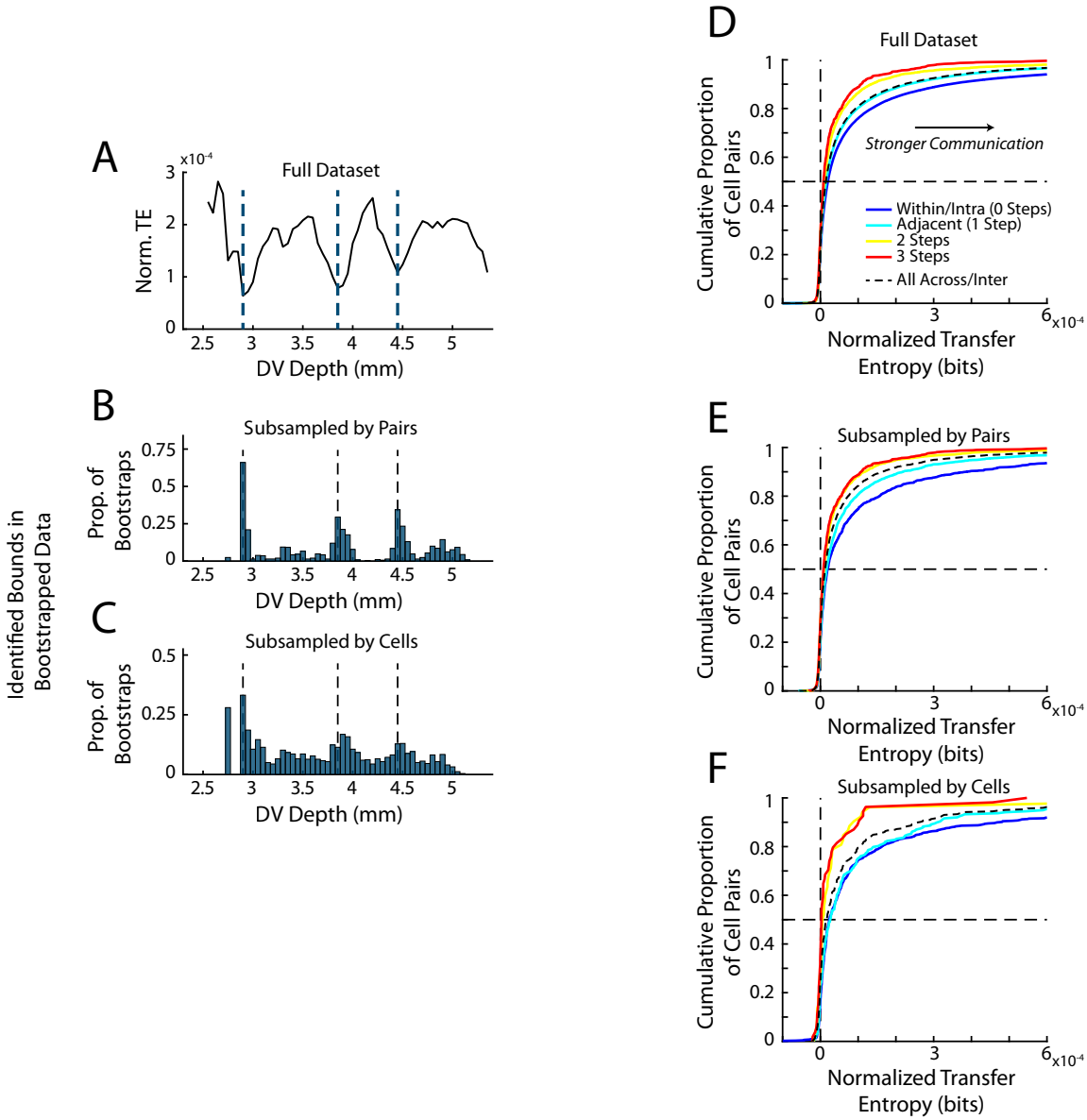

Figure 3 – Supplement 2: Subsampling of mPFC recordings for TE calculations

(A) Plot of the TE quantification by depth derived from the full data set, repeated from Fig 3C. (B) The number of cell pairs included in the analysis was down sampled to match those of the ACC subregion. For each subregion the total number of pairs within the subregion (334 pairs) and the average number of pairs across subregions (397 pairs) was equated. Quantifications of TE were then computed and subregion bounds were identified from this quantification as for the full data. This process was repeated 500 times. The histogram shows the proportion of these 500 iterations in which a subregion boundary was detected at each respective depth. Note that the majority of subsamples yielded detected bounds at the same anatomical position as those of the full data set. (C) Data are presented as in B but here the total number of cells in each subregion was subsampled to equal that of ACC (72 cells). Note that because TE requires simultaneously recorded cells and our subsampling did not consider this, subsampling by cells yields a more stringent control. However, again the same anatomical bounds are largely detected in the subsampled data. (D, E, F) Quantification of the normalized transfer entropy within and between subregions. (D) The full data is reproduced from Fig 3F. (E) Data are subsampled according to number of cell pairs. (F) Data are subsampled according to total number of included cells.

Figure 4 - Supplement 1: Analysis of main functional effects over the session

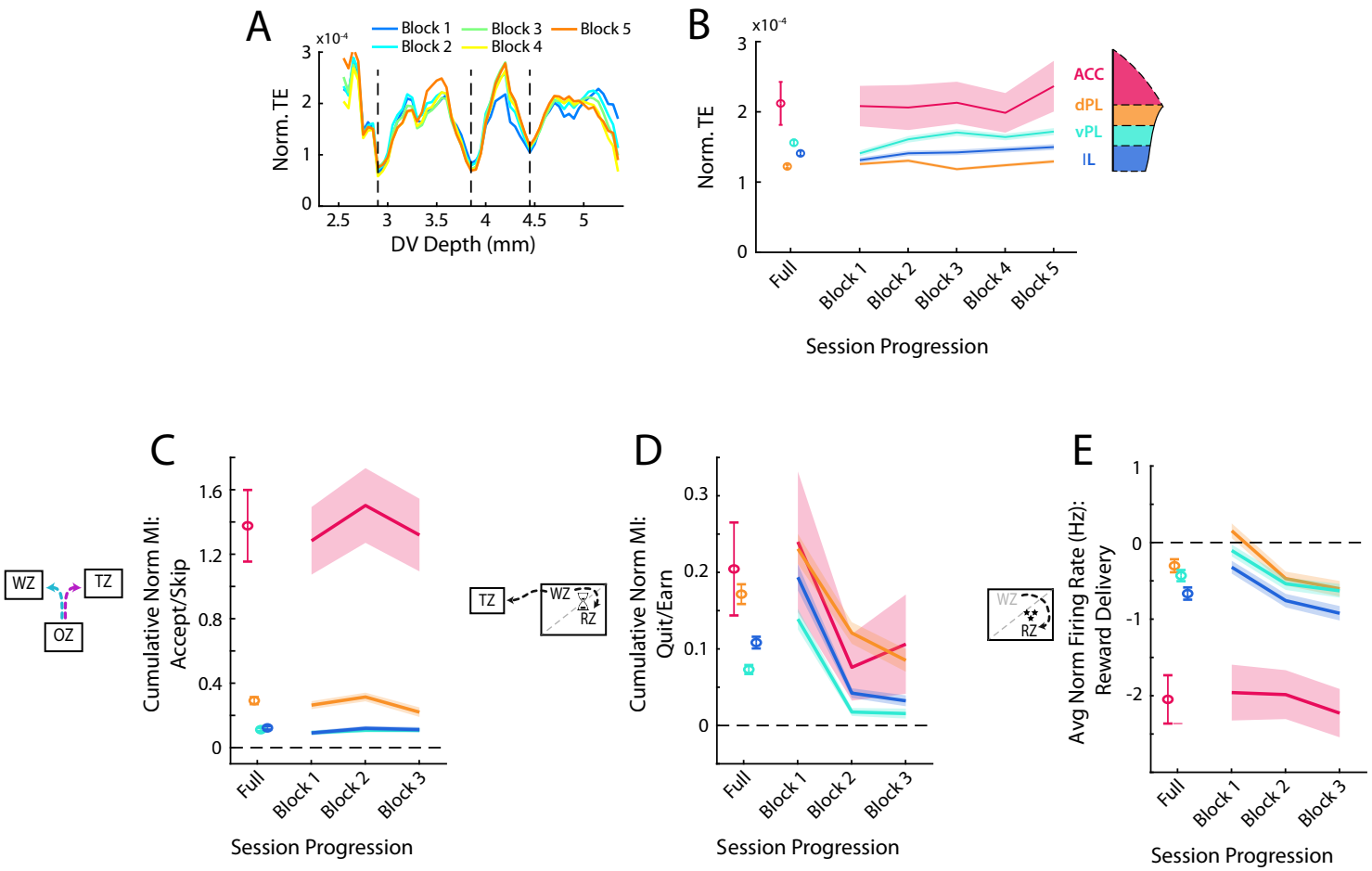

Figure 4 – Supplement 1: Analysis of main functional effects over the session

(A) Transfer entropy calculations for deriving subregion bounds subdivided by portion of the behavioral session. Data were divided into 5 equal portions of 12 min each (Blocks) and the TE analysis was repeated independently for each time block. Block 1 is the beginning of the session and Block 5 is the end. (B) Average within subregion TE coupling over the course of the session. Data are shown as Mean ± SEM of normalized TE taken only from those cell pairs within the respective subregion. Data from the full session are shown first and from each block of 12 minutes in sequence. (C, D, E) Quantification of the main response patterns in the OZ (C), WZ (D), and RZ (E) split by progression through the session. Note that here the session is divided into 3 equal portions of 20 minutes each. In the OZ and WZ these correspond to the cumulative MI during the decision period. In the RZ this corresponds to the average firing during reward delivery.

Figure 5 - Supplement 1: Offer Zone respones by recording hemisphere

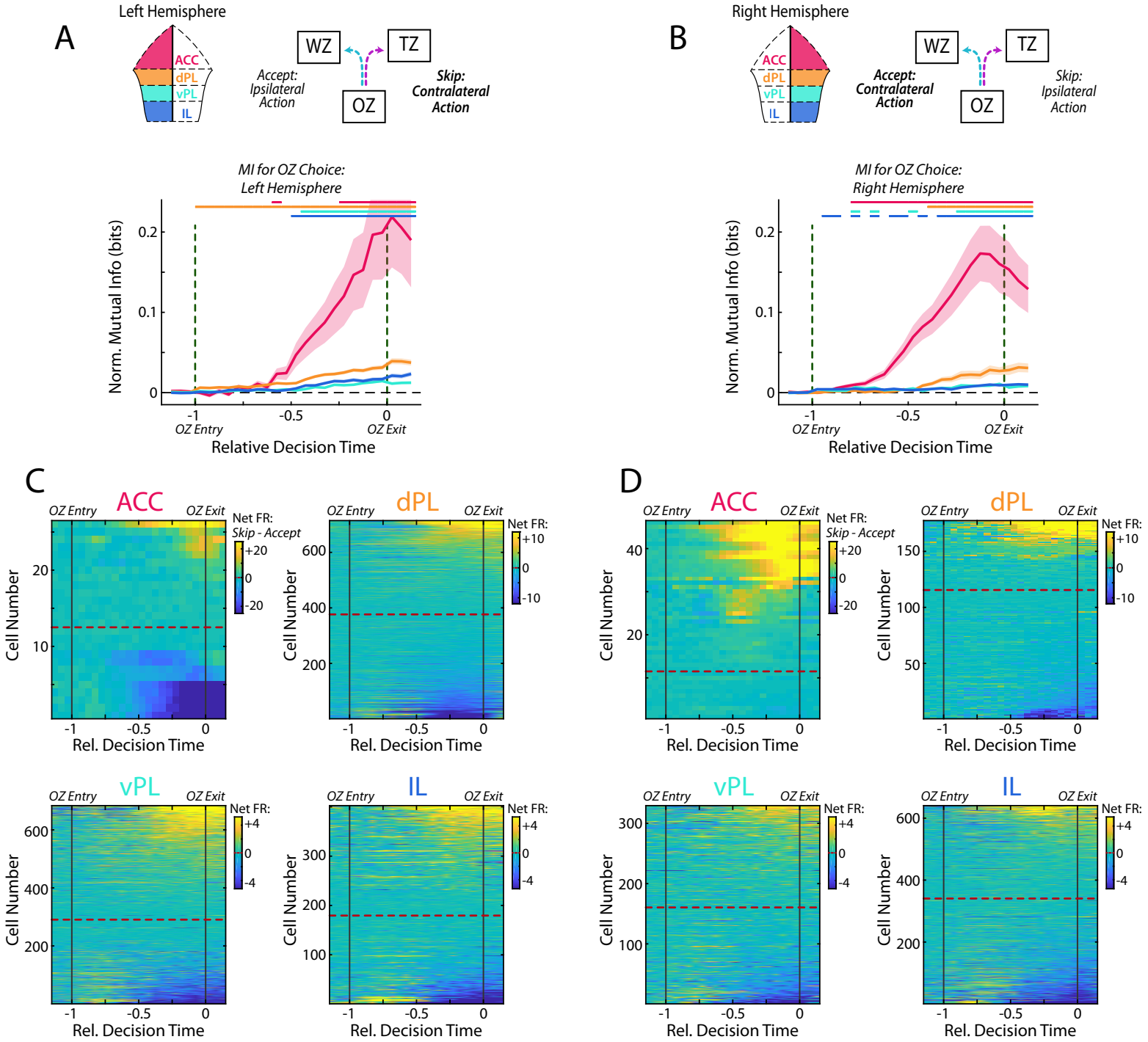

Figure 5 – Supplement 1: Offer Zone responses by recording hemisphere

Cell firing in the Offer Zone was segregated according to which hemisphere each cell was recorded on to distinguish between ipsilateral and contralateral direction of movement. The left column (A, C) denotes cells that were recorded from the left hemisphere and the right column (B, D) is cells recorded from the right. (A, B) Normalized MI during decision making in the Offer Zone. Data are presented as in Fig 5A. (C, D) Net firing rates for each cell between Skip decisions and Accept decisions. Data are presented as in Fig 5B. Note that if neural responses were strongly related to motor action, we would expect a bias toward increased firing during contralateral movement. We did not observe such a bias in our OZ data.

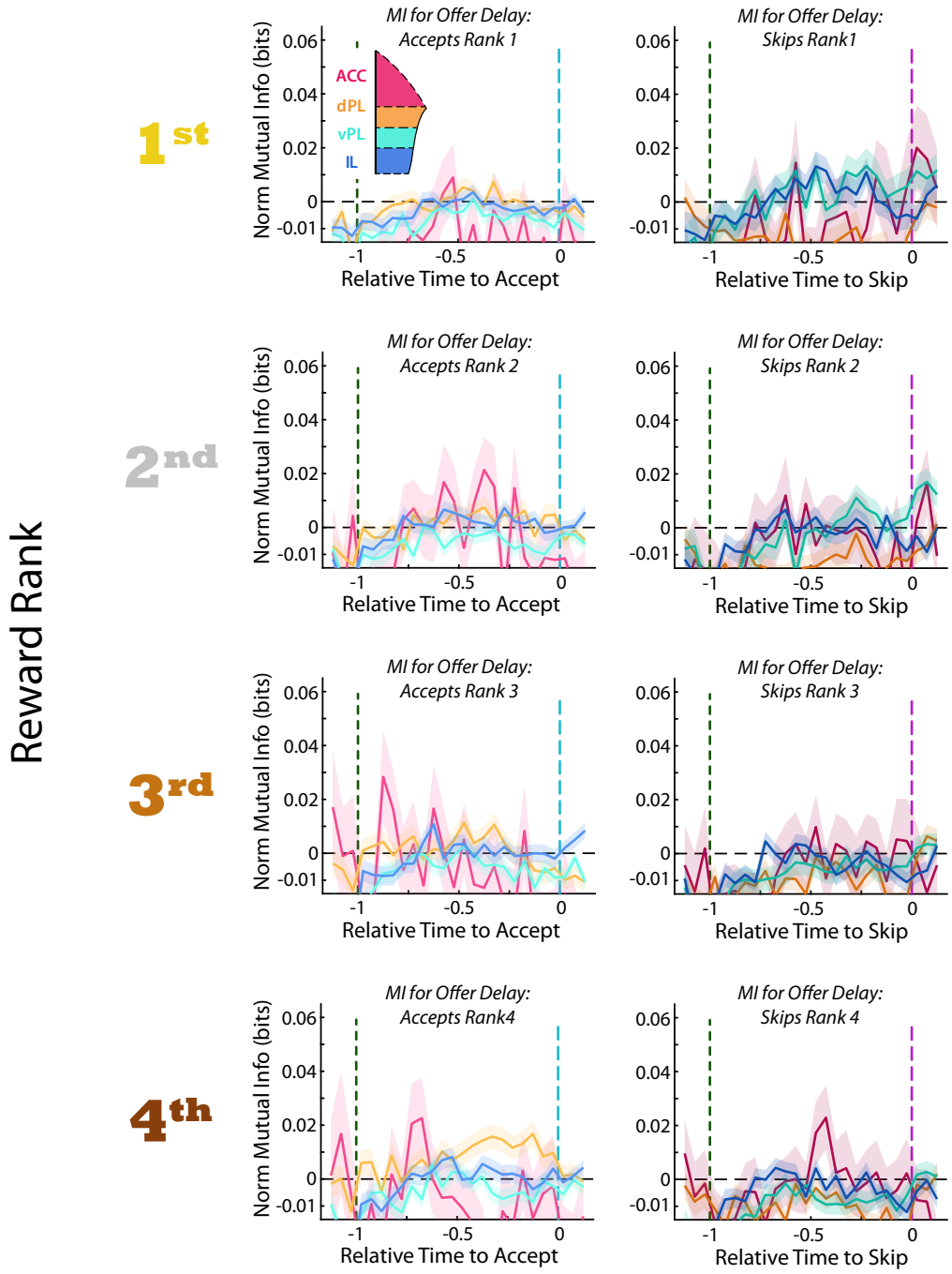
Reward Rank
1st
2nd
3rd
4th
MI for Offer Delay: Accepts Rank 1
MI for Offer Delay: Skips Rank1
MI for Offer Delay: Accepts Rank 2
MI for Offer Delay: Skips Rank 2
MI for Offer Delay: Accepts Rank 3
MI for Offer Delay: Skips Rank 3
MI for Offer Delay: Accepts Rank4
MI for Offer Delay: Skips Rank 4
ACC
dPL
vPL
IL
Norm Mutual Info (bits)
0.06
0.04
0.02
0
-0.01
-1
-0.5
0
Relative Time to Accept
Relative Time to Skip

Figure 5 – Supplement 2: Multiplexed coding of Rank and Delay

Mutual information between firing rate and offer delay. Data are segregated according to choice made in the Offer Zone (Accept or Skip) and by restaurant at which that decision was made. Under no conditions did we observe evidence for coding of offer delay in any mPFC subregions.

Figure 5 - Supplement 3: Responses in the Offer Zone from subsampled data

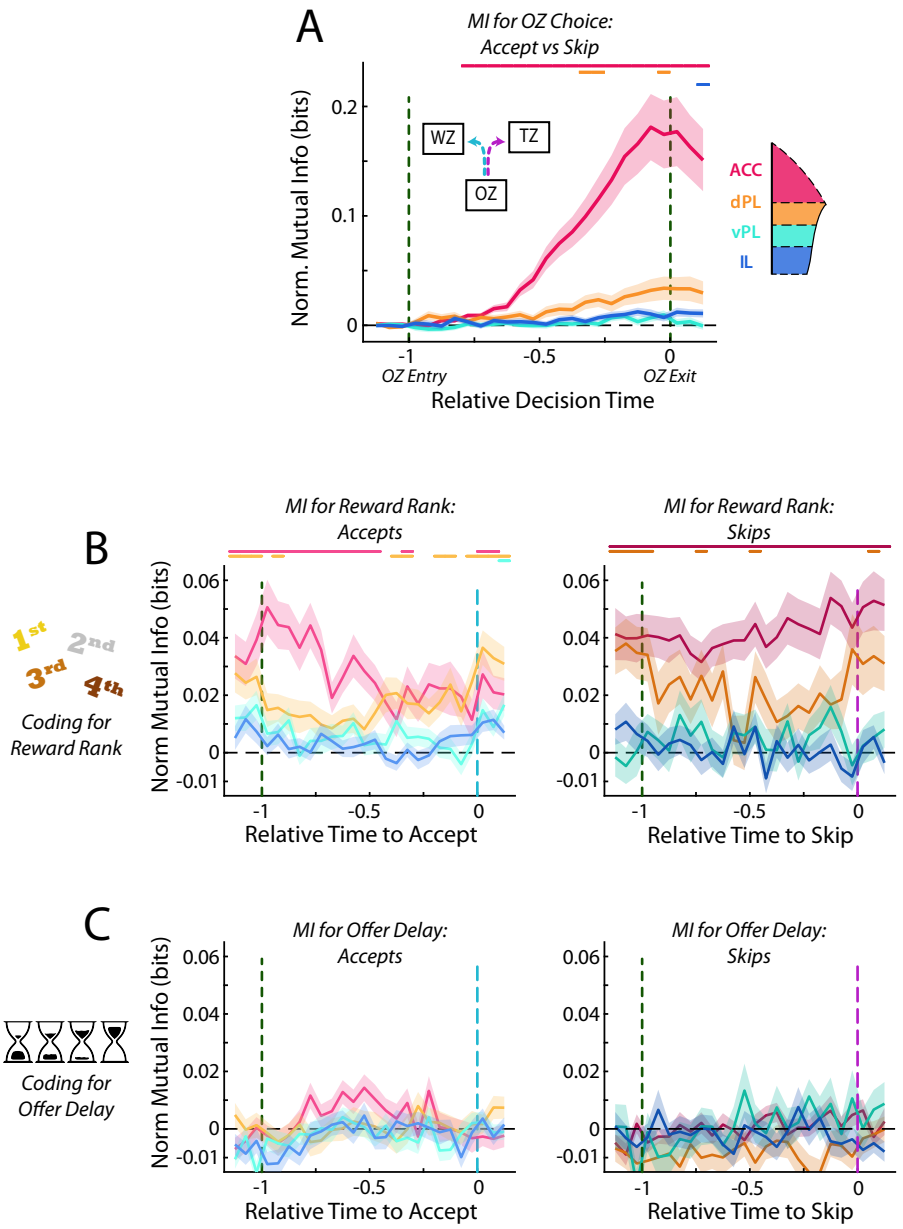

Figure 5 – Supplement 3: Responses in the Offer Zone from subsampled data

Populations of dPL, vPL, and IL cells were subsampled to match that of the ACC data set (72 Cells). Analyses from Fig 5 were repeated using the subsampled data and results are presented as in the main figure. (A) Mutual information between firing rate and choice in the OZ. (B) MI between firing rate and reward ranking. (C) MI between firing rate and offer delay.

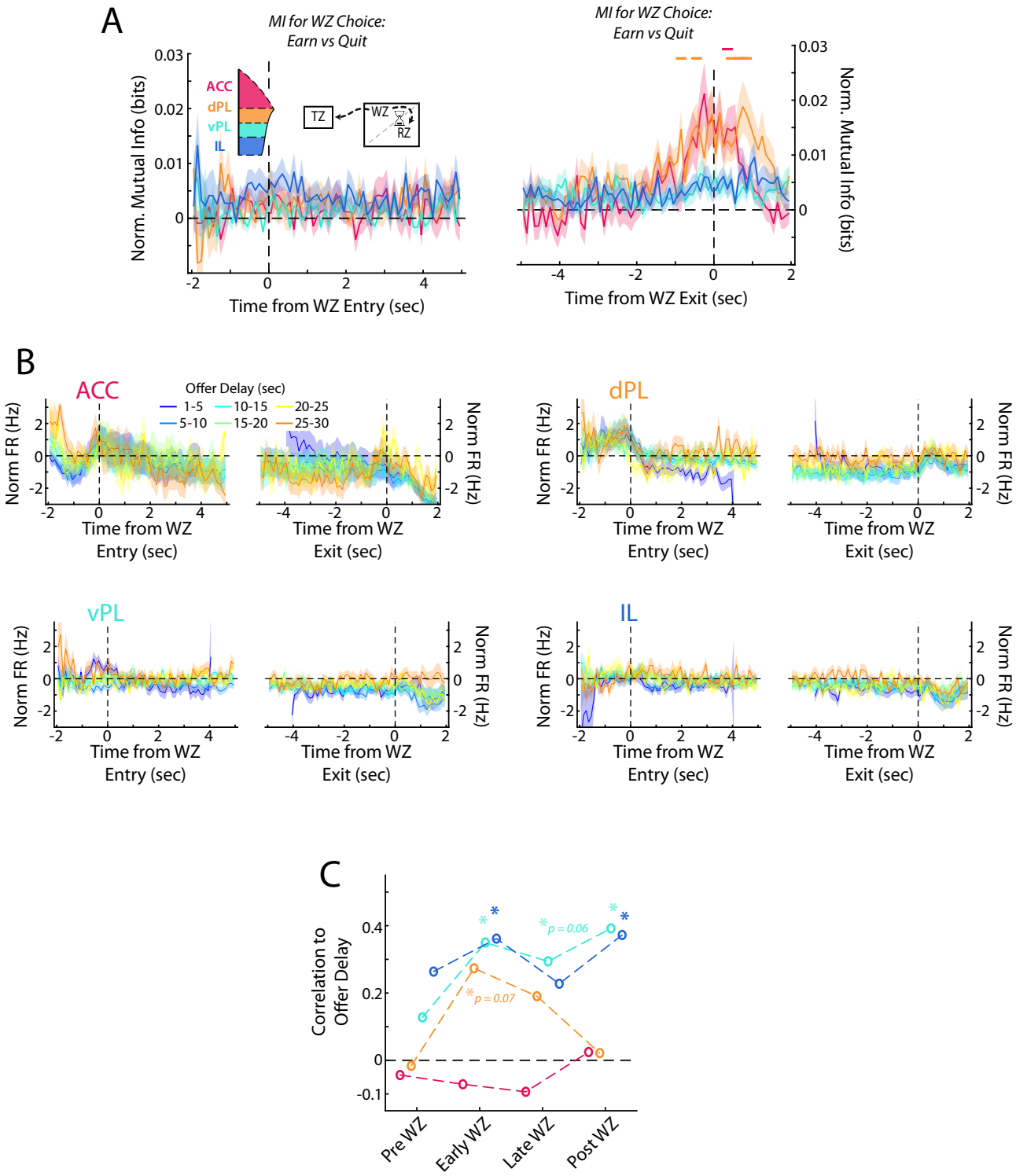
A
MI for WZ Choice: Earn vs Quit
Norm. Mutual Info (bits)
ACC
dPL
vPL
IL
TZ
WZ
RZ
Time from WZ Entry (sec)
Time from WZ Exit (sec)
B
Offer Delay (sec)
1-5
5-10
10-15
15-20
20-25
25-30
Norm FR (Hz)
Time from WZ Entry (sec)
Time from WZ Exit (sec)
C
Correlation to Offer Delay
p = 0.06
p = 0.07
Pre WZ
Early WZ
Late WZ
Post WZ

Figure 6 – Supplement 1: Responses in the Wait Zone from subsampled data

Populations of dPL, vPL, and IL cells were subsampled to match that of the ACC data set (72 Cells). Analyses from Fig 6 were repeated using the subsampled data and results are presented as in the main figure. (A) Mutual information between firing rate and choice at the beginning and end of the WZ. (B) Average firing rate separated by the offer delay that had been accepted (C) Correlation between average firing rate across mPFC cells and offer delay that was accepted.

Figure 6 - Supplement 2: Net firing between Quit and Earn responses

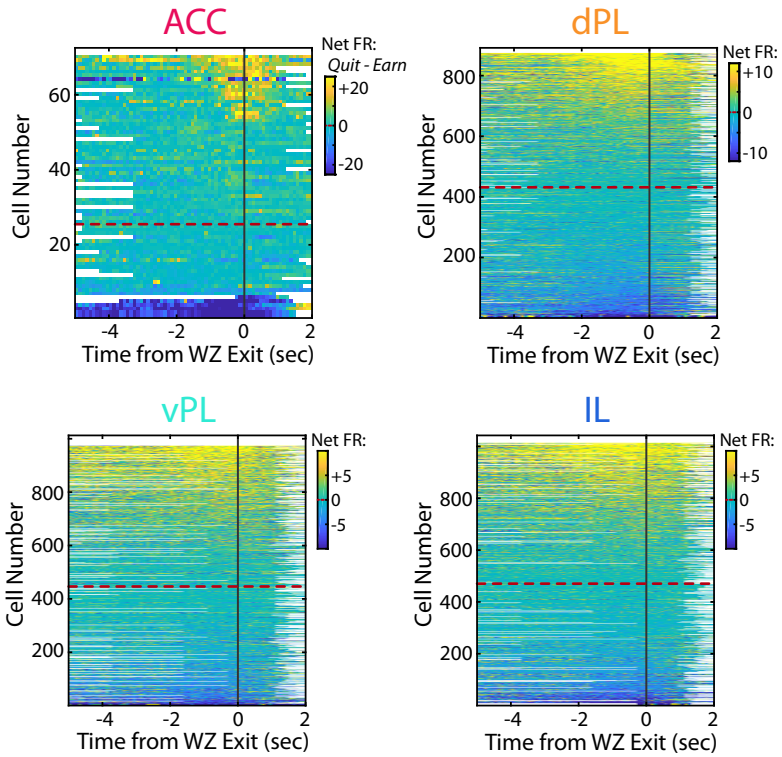

Figure 6 – Supplement 2: Net firing rate between Quit and Earn responses

Net firing rate of mPFC cells between quit and earn decisions in the Wait Zone (Quit FR – Earn FR). For each subregion cells are sorted according to net firing rate in the last 2 seconds in the WZ. Color scales are shown for each subregion but note the differences in scale between subregions.

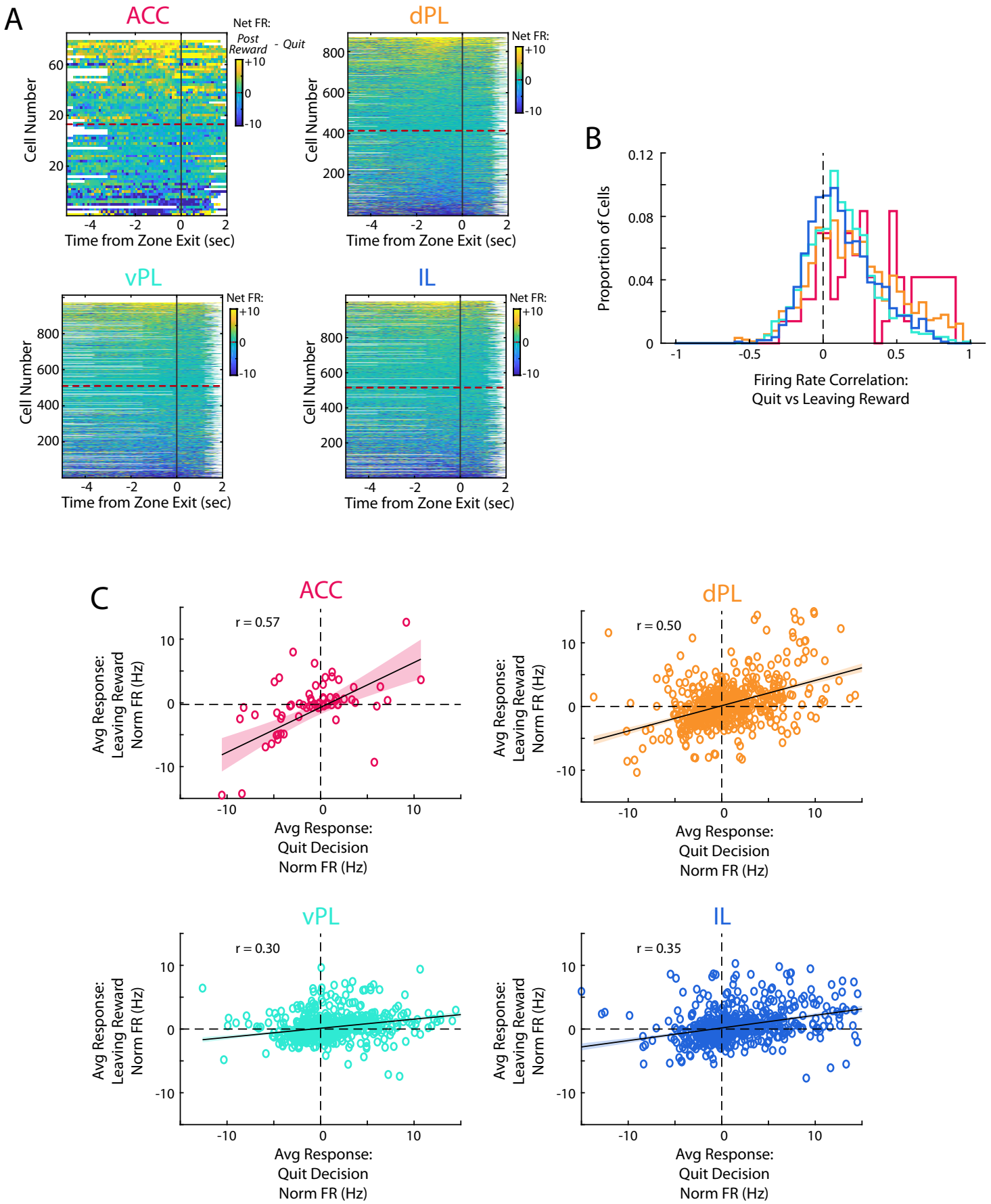
A
ACC
dPL
vPL
IL
Net FR:
Post Reward - Quit
+10
0
-10
Cell Number
Time from Zone Exit (sec)
B
Proportion of Cells
Firing Rate Correlation:
Quit vs Leaving Reward
C
r = 0.57
r = 0.50
r = 0.30
r = 0.35
Avg Response:
Leaving Reward
Norm FR (Hz)
Avg Response:
Quit Decision
Norm FR (Hz)

Figure 6 – Supplement 3: Similarity of quit decisions and leaving after reward

(A) Net firing rate of mPFC cells between the end of the time in the WZ when rats quit and offer and the time when they leave the RZ after consuming a reward (Post Reward FR - Quit FR). Note that these reflect comparable motor actions but occur during different cognitive situations. Cells are sorted according to net firing in the last 2 seconds in the zone. (B) Correlation of the PETHs between quit events and rats leaving the RZ after consuming a reward. (C) Scatter plot of the average firing rate response as rats make a quit decision and when they leave the RZ. Data are fit with a linear regression, shown with 95% confidence bounds, and the correlation between the two periods is noted for each plot.

Figure 7 - Supplement 1: Responses in the Reward Zone from subsampled data

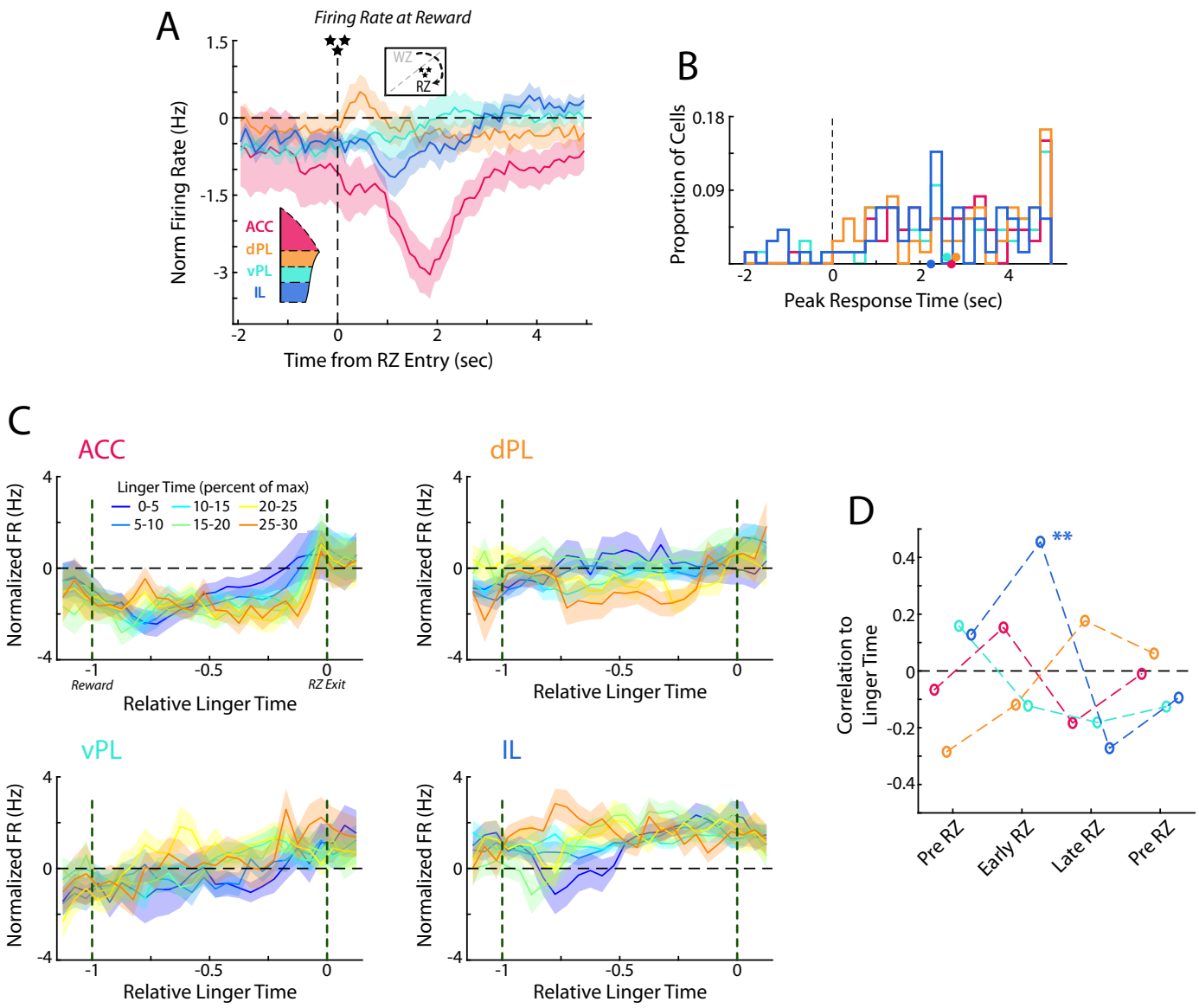

Figure 7 – Supplement 1: Responses in the Reward Zone from subsampled data

Populations of dPL, vPL, and IL cells were subsampled to match that of the ACC data set (72 Cells). Analyses from Fig 7 were repeated using the subsampled data and results are presented as in the main figure. (A) Average normalized firing rate as rats enter the RZ and earn their reward. (B) Histogram of peak responses of mPFC cells. (C) Average firing rates of mPFC cells segregated by the time spend lingering in the RZ. (D) Correlation between average firing rate across mPFC cells and the time that was spent lingering in the RZ.

Table 1: Probe Counts per Rat

Rat ID	Sex	Number of Probes	Dorsal Probe (ACC/PL) Hemisphere	Ventral Probe (PL/IL) Hemisphere
R506	Male	1	Right	***N/A***
R530	Female	2	Left	***Right (In MO; Excluded)***
R535	Female	2	Left	***Right (In VO; Excluded)***
R542	Male	2	Right	Left
R536	Female	1	Left* *Single probe moved from dorsal to ventral*	
R543	Male	1	***N/A***	Right
R537	Female	2	Right	Left
R544	Male	2	Right	Left

Table 1: Probe counts by rat

Breakdown of the sex and allocation of probe targeting across rats.

Table 2: Cell Counts per Rat

			Per Session Counts				Totals Counts				Per Session Averages			
Rat ID	Total Sessions	Total Cells	Avg Cells	Min Cells	Max Cells	Std Cells	ACC Cells	dPL Cells	vPL Cells	IL Cells	ACC Cells	dPL Cells	vPL Cells	IL Cells
R506	7	109	15.6	7	20	4.6	41	68	0	0	5.9	9.7	0.0	0.0
R530	14	206	14.7	3	34	10.6	0	92	102	12	0.0	6.6	7.3	0.9
R535	13	611	47.0	35	63	9.9	0	563	48	0	0.0	43.3	3.7	0.0
R542	14	562	40.1	29	50	5.7	1	40	310	211	0.1	2.9	22.1	15.1
R536	14	145	10.4	3	24	5.8	12	44	46	43	0.9	3.1	3.3	3.1
R543	14	464	33.1	19	47	8.0	0	9	172	283	0.0	0.6	12.3	20.2
R537	14	530	37.9	20	54	10.8	14	2	157	357	1.0	0.1	11.2	25.5
R544	14	390	27.9	18	41	7.1	4	73	177	136	0.3	5.2	12.6	9.7
TOTAL	**104**	**3017**	**28.3**	**16.8**	**41.6**	**7.8**	**72**	**891**	**1012**	**1042**	**1.0**	**8.9**	**9.1**	**9.3**

Table 2: Cell counts by rat

Breakdown of the number of cells recorded from each rat. Blue shaded cells denote the sum across a respective column and green shaded cells denote the mean. Note that while a large proportion of ACC and dPL cells come from one rat each, for all analyses we confirmed that results were qualitatively similar across the remaining animals (data not shown).