



Published in final edited form as:

*Curr Biol.* 2025 January 20; 35(2): 431–443.e4. doi:10.1016/j.cub.2024.11.069.

## Recurrent activity propagates through labile ensembles in macaque dorsolateral prefrontal microcircuits

Suzanne O. Nolan<sup>1</sup>, Patrick R. Melugin<sup>1</sup>, Kirsty R. Erickson<sup>1</sup>, Wilson R. Adams<sup>1</sup>, Zahra Z. Farahbakhsh<sup>1</sup>, Colleen E. Mcgonigle<sup>2</sup>, Michelle H. Kwon<sup>1</sup>, Vincent D. Costa<sup>3,7</sup>, Troy A. Hackett<sup>4</sup>, Verginia C Cuzon<sup>3</sup>, Christos Constantinidis<sup>5</sup>, Christopher C. Lapish<sup>6</sup>, Kathleen A. Grant<sup>3</sup>, Cody A. Siciliano<sup>1,6</sup>

<sup>1</sup>Department of Pharmacology, Vanderbilt Brain Institute, Vanderbilt Center for Addiction Research, Vanderbilt University, Nashville, TN 37232, USA

<sup>2</sup>Department of Psychology, Indiana University Indianapolis, Indianapolis, IN 46202, USA

<sup>3</sup>Division of Neuroscience, Oregon National Primate Research Center, Oregon Health and Science University, Beaverton, OR 97006, USA

<sup>4</sup>Department of Hearing and Speech Sciences, Department of Psychology, Vanderbilt University School of Medicine, Vanderbilt University, Nashville, TN 37232, USA

<sup>5</sup>Department of Biomedical Engineering, Department of Pharmacology, Vanderbilt University, Nashville, TN 37235, USA

<sup>6</sup>Department of Anatomy, Cell Biology, & Physiology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

<sup>7</sup>Division of Developmental and Cognitive Neuroscience, Emory National Primate Research Center, Atlanta, GA 30329, USA

### Summary

Human and non-human primate studies clearly implicate the dorsolateral prefrontal cortex (dlPFC) as critical for advanced cognitive functions<sup>1,2</sup>. It is thought that intracortical synaptic architectures

**Lead contact and co-corresponding author:** Cody A Siciliano, cody.siciliano@vanderbilt.edu, Twitter: @CodySiciliano. **Co-corresponding author:** Kathleen A Grant, grantka@ohsu.edu.

**Author Contributions:** Conceptualization: S.O.N., K.A.G., C.A.S.; Data curation: S.O.N., K.A.G., C.A.S.; Formal analysis: S.O.N., W.R.A., C.E.M., M.H.K., C.C.L., C.A.S.; Funding acquisition: C.C.L., K.A.G., C.A.S.; Investigation: S.O.N., P.R.M., K.R.E., W.R.A., Z.Z.F., T.A.H., V.C.C., C.A.S.; Methodology: S.O.N., P.R.M., K.R.E., W.R.A., Z.Z.F., T.A.H., C.C.L., K.A.G., C.A.S.; Project administration: C.C., K.A.G., C.A.S.; Resources: V.D.C., V.C.C., C.C., C.C.L., K.A.G., C.A.S.; Software: S.O.N., W.R.A., C.C.L., C.A.S.; Supervision: V.D.C., C.C., K.A.G., C.A.S.; Validation: S.O.N., K.R.E., C.C., C.A.S.; Visualization: S.O.N., C.C.L., C.A.S.; Writing – original draft: S.O.N., C.A.S.; Writing – review & editing: S.O.N., P.R.M., K.R.E., W.R.A., Z.Z.F., C.E.M., M.H.K., V.D.C., T.A.H., V.C.C., C.C., C.C.L., K.A.G., C.A.S.

As co-corresponding author, C.A.S. was responsible for overseeing all imaging data collection and analysis, as well as manuscript writing and revision. As co-corresponding author, K.A.G. was responsible for overseeing the cognitive behavioral testing protocols and the necropsy protocol. This included assignment, training and daily implementation of the procedures involving the research animals, as well as assuring the fidelity of the behavioral equipment and datasets provided.

**Declaration of Interests:** The authors declare no competing interests.

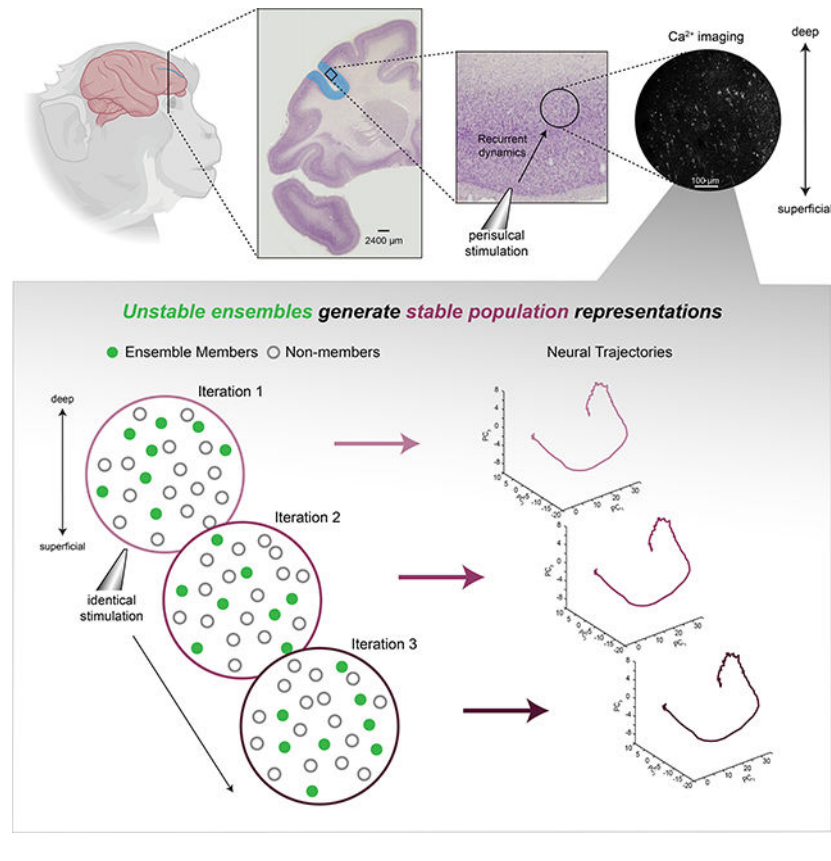
**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published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

within the dIPFC are the integral neurobiological substrate that gives rise to these processes<sup>3–7</sup>. In the prevailing model, each cortical column makes up one fundamental processing unit composed of dense intrinsic connectivity, conceptualized as the ‘canonical’ cortical microcircuit<sup>3,8</sup>. Each cortical microcircuit receives sensory and cognitive information from upstream sources which are represented by sustained activity within the microcircuit, referred to as persistent or recurrent activity<sup>4,9</sup>. Via recurrent connections within the microcircuit, activity propagates for a variable length of time, thereby allowing temporary storage and computations to occur locally before ultimately passing a transformed representation to a downstream output<sup>4,5,10</sup>. Competing theories regarding how microcircuit activity is coordinated have proven difficult to reconcile *in vivo* where intercortical and intracortical computations cannot be fully dissociated<sup>5,9,11,12</sup>. Here, we interrogated intrinsic features of isolated microcircuit networks using high-density calcium imaging of macaque dIPFC *ex vivo*. Using perisulcal stimulation to evoke recurrent activity in deep layers, we found that activity propagates through stochastically assembled intracortical networks whereby orderly, predictable low dimensional collective dynamics arise from ensembles with highly labile cellular memberships. Microcircuit excitability covaried with individual cognitive performance, thus anchoring heuristic models of abstract cortical functions within quantifiable constraints imposed by the underlying synaptic architecture. Together, our findings strongly suggest that the ability to produce high fidelity population-level signals from distributed, labile networks is an intrinsic feature of dIPFC microcircuitry.

### eTOC blurb:

Using *ex vivo* Ca<sup>2+</sup> imaging in macaque dorsolateral prefrontal cortex, Nolan et al. reveal that recurrent activity propagates through unstable, stochastically assembled ensembles to produce stable population-level events. Network excitability covaried with set-shifting performance, suggesting these features may underlie dIPFC’s role in cognition.

### Graphical Abstract



## Results & Discussion

Contemporary knowledge of cortical microcircuits relies heavily on studies in non-primate species, where *ex vivo* brain slice preparations allow the cytoarchitectural organization and synaptic weights dictating intracortical dynamics to be interrogated in the absence of intercortical and sensory inputs<sup>13–15</sup>. While studies in non-primate species have been critical to our understanding of microcircuitry in sensory cortices, the dlPFC is an evolutionary specialization distinct to primates without a direct homolog in other mammals<sup>16,17</sup>. As such, direct investigation of the intrinsic properties of intracortical dlPFC microcircuits represents a missing link in the chain of evidence connecting synaptic features at the single-cell level with the cognitive processes that the dlPFC subserves. This work establishes a paradigm for assessing intrinsic features of intracortical microcircuits in acute primate brain slices to explore three overarching questions: 1) to what degree does the intrinsic architecture of the dlPFC propagate activity in the absence of ongoing extrinsic inputs, 2) for a given discrete input, does network activity arise from dedicated, stable groups of cells or from a distributed and flexible network, and 3) do the static features of intracortical excitability observed *ex vivo* covary with individual differences in cognitive performance.

To interrogate microcircuit interactions that govern network-level properties, it is necessary to observe the activity of many cells simultaneously, ideally without sampling bias towards genetically- or morphologically-defined subpopulations. To achieve high-density, spatiotemporally resolved recordings of arbitrarily sampled cells, acute *ex vivo* brain slices

from adult rhesus macaques containing dlPFC area 46 were loaded with the cell-permeable fluorescent calcium dye fluo-4 acetoxymethyl ester (fluo-4 AM) (Figure S1A). Using necropsy and MRI-guided tissue preparation procedures optimized for obtaining viable slices with high anatomical repeatability<sup>18,19</sup>, slices displayed high loading efficiency with cells clearly visible with as little as a one-minute loading period in 5  $\mu\text{M}$  of fluo-4 AM (Figure S1B). Importantly, fluo-4 AM is cell-permeable and non-fluorescent until cleavage of the AM tag by intracellular esterases, which renders the dye cell-impermeable and activates its calcium-dependent fluorescence, thus trapping the active form within the cell. Because this process requires active esterases, loading efficiency also provides a within-sample, online verification of cell viability<sup>20,21</sup>. A 30-minute loading period was sufficient to visualize cellular processes and produced dense loading throughout the slice (Figure S1B–E). Co-staining with DAPI confirmed that cellular features aligned with nuclei as expected (Figure S1F). Using high-speed laser scanning confocal microscopy, we visualized temporally-resolved calcium dynamics in hundreds of cells simultaneously over a 762  $\mu\text{m}^2$  field of view (Video S1).

To systematically evaluate the degree to which a temporally and spatially discrete stimulation produces persistent activity, a bipolar stimulating electrode was placed superficially along the bank of the principal sulcus while calcium recordings were obtained distally in the deep pyramidal layers (Figure 1A–D). This configuration allows for input-output relationships to be directly assessed in the absence of extrinsic connections, which necessitates that any polysynaptic transient changes in activity must be propagated through feedback and/or feedforward connections within the slice (referred to throughout as recurrent activity). By recording distal to the stimulating electrode, recurrent activity can be assessed downstream of the direct depolarizations induced by the electrical field potential generated between the poles of the electrode. Though stimulation cannot be equated with endogenous inputs occurring in an intact system, any propagation of activity downstream of the initial event must conform to the synaptic architecture and intrinsic constraints imposed by the intracortical microcircuitry. Using this approach, we first examined how recurrent activity patterns arise from a single pulse stimulation across a range of intensities ( $n=8$  slices from 8 subjects, 2–10 mA single pulse stimulations). We found that single pulse stimulation was sufficient to produce robust transients in the bulk fluorescence signal (mean intensity over time across the field of view) which scaled with increasing stimulation intensities (Figure 1E). Intensity-dependent increases were observed in both peak amplitude as well as the area under the curve (AUC) of the decay phase, measured in a sliding 2s window beginning one frame after the peak fluorescence intensity for each trace (Figure 1G–H). Importantly, the stimulation duration (4 ms) and the off-kinetics of the sensor (typically 50–200 ms decay times<sup>22,23</sup>) are shorter than the duration of the observed signals (tau decay time  $\times$  stimulation intensity curve reached an asymptote of  $480 \pm 38$  ms (mean  $\pm$  SEM)) (Figure S1G–H). This indicates propagation of activity via multi-synaptic events and/or persistent changes in basal activity patterns. Moreover, the loading efficiency, as assessed by the number of cells segmented in each field of view, did not covary with the amplitude of the whole field response (Figure S1J), confirming that these signals represent network dynamics rather than the density of loaded cells. The input-output curves for both signal amplitude and AUC of the decay phase displayed a sigmoidal shape over scaling intensities. Indeed, over

a relatively restricted range of input intensity (5-fold), both curves scaled to asymptote and displayed similar half-maximal intensities of roughly 5 mA (Figure 1G–H). These results indicate that propagation of intracortical activity can be readily evoked in this preparation and that, despite the exogenous nature of the input stimulation, disparate input parameters produce distinct and predictable population-level responses.

In a parallel experiment ( $n=5$  slices from 5 subjects), using the same recording configuration, we next evoked activity using a single pulse stimulation matched to the half-maximal parameters observed in the stimulation intensity curve (5 mA, all parameters identical). We then examined the input-output relationship across an orthogonal parameter axis by delivering stimulations of fixed intensity with increasing train lengths (1 – 20 pulses, 20 Hz, 5 mA throughout) (Figure 1F). The amplitudes and area of the decay phase evoked by the single pulse, 5 mA stimulation did not differ between experiments (unpaired  $t$ -test<sub>(1,11)</sub>=0.66,  $p = 0.52$  for peak amplitude; unpaired  $t$ -test<sub>(1,11)</sub>=0.86,  $p = 0.41$  for AUC), confirming that this recording strategy allows repeatable input-output relationships to be observed and compared across slices. Across varying pulse number, the tau decay constants again far outlasted the decay kinetics of the sensor (average across all stimulation conditions  $770 \pm 34$  ms (mean $\pm$ SEM)) (Figure S1I). In contrast to the intensity curve, both amplitude and AUC of the decay phase increased linearly and reached much greater magnitudes with increasing train lengths (Figure 1I–J). These results indicate inputs received in close temporal succession, 50 ms apart in this case, produce summation in downstream activity. This temporal summation is cumulative, scaling linearly even when inputs arrive over a relatively extended total timeframe (1s between the first and last pulse in the 20 pulse condition). Temporal summation of cortical inputs is thought to be key, perhaps requisite, for recurrent network activity, as it provides a means for a recurrent network to modulate the conditions under which persistent activity is initiated or terminated<sup>24</sup>. Several biophysical mechanisms have been implicated in temporal summation, such as membrane capacitance, resistance, and surface area, which together determine the membrane time constant. In addition, ionic currents, such as  $I_h$ , and through ligand-gated receptors as well as calcium clearance/handling, have been shown to be critical in governing the integrating properties of neurons<sup>25–28</sup>. Together, these mechanisms can dynamically and rapidly regulate the integration of incoming stimulations, in both linear and nonlinear fashions. The present works expand on these findings to further characterize how input-output relationships reveal aspects of this process in a fixed (i.e. *ex vivo*) system.

To determine how calcium dynamics in individual cells (Video S1) give rise to the population-level signals described above, we next extracted activity traces corresponding to single cells from the recordings. We visualized the activity aligned to stimulation onset for the 2,189 cells recorded across the intensity curve, revealing clear excitatory and inhibitory responses (Figure 1K). Normalized single-cell activity traces were then thresholded to classify the cells according to responsivity (excited, inhibited, no response), based on the AUC during the 2 seconds following the stimulation onset (further details in the Methods). Separation into excited or inhibited subpopulations illustrated stimulation-evoked divergent activity between the subpopulations and unclassified (no response) cells as expected, and demonstrated that this divergent activity persisted well beyond the window used for classification (Figure 1L). Comparing the distribution of excited, inhibited, and

non-responsive cells across stimulations showed moderate increases in the number of excited and inhibited cells as the stimulation intensity increased (Figure 1M). Indeed, increasing stimulation intensity led to a linear rise in the proportion of the population recruited (Figure 1N). Further, analysis of individual amplitudes for each cell across the intensity curve showed that the cells classified as having no response showed minimal-to-no input-dependent scaling, corroborating that the thresholds set for classification effectively captured separate subpopulations (Figure S1K–R). To further validate the integrity of these measures, we also confirmed that both excitatory and inhibitory responses were sensitive to the calcium chelator BAPTA-AM (Figure S1S–T).

Examination of single-cell responses over the input-output curve for increasing train length ( $n = 549$  cells) again demonstrated that persistent excitatory and inhibitory changes were evident in single cells well beyond the initial stimulation-evoked response (Figure 1O–P). It is interesting to note that the bulk fluorescence signal differs from the means of the individual cell populations (c.f. Figure 1E–F, 1K–P). This disparity is likely due to the contribution of neuropil signal to the bulk fluorescence traces, which may reflect activity arising from presynaptic terminals from long-range inputs, horizontal intralaminar axons, and dendrites of the recorded cells. The demixing method used to extract the single-cell traces removes the contribution of global and local neuropil signals to ensure that single-cell dynamics are faithfully represented<sup>29–32</sup>. The apparent responsiveness of the neuropil in the deep layers, distal to the stimulator, provides further evidence that the stimulations/recording configuration evokes signals which are mediated by largely polysynaptic events. We also see that the proportion of the population recruited by the stimulation scaled with increasing train lengths (Figure 1Q). In contrast to the intensity curve, where proportions of the population recruited scaled linearly over a limited range, increasing the train length led to robust and distinctly nonlinear increases in the proportion of cells responsive to the stimulation (Figure 1R). Thus, though there was clear scaling in the amplitude of the bulk responses over both curves (Figure 1E–F), the input-output relationship was disparate. Together these results suggest that temporal summation of inputs is capable of recruiting a larger downstream network than is possible with any single input, regardless of intensity.

We next sought to determine if the input-output relationships described above are mediated by specific networks of cells. The networks within a cortical microcircuit and their contribution to neural processes are the subject of long-standing debate. Though the terminology has taken many forms over the decades, a repeated theme is whether networks of co-active cells, which we will refer to here as ensembles, are composed of dedicated versus labile members. Theories have spanned the spectrum of this dichotomy<sup>33–36</sup>. On one extreme, units of cognitive information are stored in a physical location, conceptualized as an engram or localist architecture, within an ensemble of cells that each have a dedicated contribution to the ensemble dynamics when activated<sup>33</sup>. On the other extreme, ensembles are functionalized through a population-level code where any given output can be achieved by distinct combinations of cellular members such that contributions of each are interchangeable, and no single component is indispensable<sup>35,36</sup>. *In vivo* recordings have demonstrated the existence and importance of ensembles in sensory and cognitive representations<sup>37,38</sup>; however, at any given time, an unknown number of ensembles are active, due to encoding of sensory stimuli and internal cognitive processes which are

constantly co-occurring *in vivo*. This complexity, along with a multitude of technical limitations in recording from identified cells over time, has prevented progress towards reconciling competing theories of ensemble membership stability in the dIPFC.

To explore the range of possibilities permitted by the intrinsic connectivity of the dIPFC, we analyzed the ensembles evoked across the input-output curves of stimulation intensity and train length with respect to cellular identities. Activity traces, detailed in Figure 1, were extracted from a single concatenated, co-registered video of all stimulation replicates per slice/subject. Thus, any possible methodological weights on detection thresholds were applied equally across all stimulations and each cell in the dataset received an unambiguous, binary classification indicating the presence or absence of activity at each stimulation replicate.

Throughout both experiments, the stimulating electrode remained in a static location, allowing for observations of ensembles evoked by a stimulus ( $S$ ) with known features in the absence of any non-recurrent inputs to the network. Stimulations were delivered at least 3 minutes apart, thus ensembles evoked by each iteration ( $x$ ) of  $S$  were considered as the representation of  $S^x$  (Figure 2A). The population of cells recorded throughout the experiment are referred to as  $S$ , and probabilities were expressed as a proportion of the number of cells in the  $S$  population. We first determined the shared versus unique membership between the ensembles evoked by the first, lowest intensity stimulation ( $S^A$ , corresponding to the 2 mA stimulation shown in Figure 1) and the stimulation which produced a half-maximal population-level response ( $S^A \rightarrow S^D$ ), and then again between the half-maximal stimulation and the final, highest intensity stimulation which was in the asymptotic portion of the input-output curve ( $S^D \rightarrow S^I$ ). A Sankey diagram depicting the percentage of  $S$  recruited to each ensemble and the proportion of crossover between each, revealed that although ensemble size was essentially identical for  $S^A \rightarrow S^D$  (26% and 25% of  $S$ , respectively), greater than 50% of the  $S^D$  ensemble represented unique members which were not recruited in  $S^A$  (Figure 2B). At the highest intensity stimulation, the ensemble size scaled to 35%, and again more than 50% of the membership was unique from  $S^D \rightarrow S^I$ . Visualizing membership crossover between one, ten, and twenty pulse stimulations again showed a striking degree of membership instability (Figure 2C). Performing the same analysis within only the cells which were members of the  $S^A$  ensemble yielded values strikingly similar to the full population with 25% and 36% of the  $S^A$  ensemble being recruited to  $S^D$  and  $S^I$ , respectively (Figure 2D). Dividing cells by whether they displayed excitatory or inhibitory responses revealed that a similar degree of turnover occurred within each response class (Figure 2D–E). These results indicate that in the absence of ongoing inputs, regardless of whether a broad or directionality-specific membership criterion is applied, recursive activity originating from a fixed spatial location within the dIPFC microcircuit propagates through predominantly labile ensembles.

We next considered whether these dynamics were reflective of 1) a flexible but predictable ensemble, with the same sets of cells switching in and out at each iteration of  $S$ , or 2) probabilistic recruitment of unique members at each iteration such that a high degree of turnover would be observed at any intersection of  $S$  ensembles. Therefore, we determined the set similarity expressed as Jaccard's similarity index, where values of 1 and 0 represent

identical vs non-overlapping memberships for a given comparison, between each of the ensemble iterations (Figure 2F). The resulting set similarity maps clearly support the second scenario: across the 46 total comparisons resulting from pairwise analyses within each dataset similarity coefficients ranged from 0.12 to 0.47, demonstrating a high degree of turnover between any intersection of  $S$  ensembles (Figure 2G, I). These findings strongly suggest that the synaptic architecture of dIPFC microcircuitry is organized such that population-level recurrent signals propagate through probabilistically recruited cellular networks.

Given the surprising degree of apparent instability, we next sought to determine whether ensemble recruitment was entirely probabilistic, with all cells having an equal likelihood of recruitment to the ensemble or, alternatively, if recruitment was weighted towards or against specific cells. To test these possibilities, a matrix was formed with columns equal to the number of cells in the dataset and rows equal to the number of iterations/stimulations. Each value in the matrix was a 0 or 1 indicating whether the cell was a member of the ensemble at each iteration. To determine the distribution of memberships that would be expected if recruitment probability was equally weighted across all cells, a shuffled dataset was created by randomizing the location of values in the matrix. Thus, the shuffled dataset had the same number of cells and stimulations, as well as the same total number of ensemble memberships, but any influence of cell identity on probability of ensemble membership across iterations was due to chance. Recruitment probability per cell [number of times included in the ensemble/number of times tested] was calculated for the observed and shuffled datasets, and the distributions of recruitment probability were compared across 10,000 rounds of reshuffling (Figure S2A). We found that the distribution of recruitment probability to the  $S$  ensembles was consistently different than the shuffled control distributions, indicating that recruitment probability was unequally weighted across the population (Figure 2H, Figure S2).

To provide a more intuitive readout of how recruitment probability was weighted across the population, we categorized each cell based on whether it was activated at each iteration of  $S$  into three groups: those that were members at all iterations of  $S$  (termed “Always”), those that were not members of any of the  $S$  ensembles (termed “Never”), and those that were members and non-members of at least one  $S$  ensemble (termed “Sometimes”) (Figure 2H). The mean recruitment probability across iterations of  $S$  was 0.2731, thus the probability  $P$  of a cell falling into the ‘Always’ category if each of the 9 iterations of  $S$  is assumed to be a fully independent event can be determined as  $P = (0.2731)^9$  and would be expected to occur in 1 out of every 188,333 cells (0.000845%). The probability of falling into the ‘Never’ category assuming that recruitment at each iteration is independent is  $P = (1 - 0.2731)^9$  or 1 out of ~18 cells (5.66%). In contrast, out of 2,189 cells observed across the 9 iterations of ensemble  $S$ , Always cells represented 0.18% of the population, a 213 times greater proportion than would be expected if ensemble recruitment was unweighted across the population. Never cells represented 8.82% of the population, 1.6 times more than expected without weighting. Though the total number of cells categorized as Always and Never were few, just 4 and 124 respectively, the observed distributions differed markedly from shuffled controls (Figure S2A–E). These results suggest that although the large majority of ensemble

members are stochastically recruited to participate in a given iteration, there is a small minority of core members which contribute to all iterations of the ensemble.

For the ensembles evoked by stimulations with increasing train length, where the mean recruitment probability over the 5 iterations was 0.4393 (Figure 2I), Always cells represented 6.01% of the population, 3.7 times greater than expected if no weights were applied, while Never cells represented 11.11% of the population, 2 times greater than expected without weighting. Statistical comparisons of the number of cells in each category with each of the 10,000 reshuffled controls revealed that the observed categorizations were indeed consistently different than would be expected by chance (Figure S2). To determine whether weighting of recruitment probability was related to cells' contribution to the ensemble, we performed the same analysis when only excited cells were considered members of the ensemble (termed excited subassembly) or when only inhibited cells were considered members of the ensemble (termed inhibited subassembly) (Figure 2K). We found that, across both datasets, within the excited and inhibited subassemblies the proportion of the population classified as Always was reduced or eliminated entirely, demonstrating that the majority of Always cells in the full ensemble were recruited through both excitatory and inhibitory activity over iterations (Figure 2L–O). In contrast, the proportion of Never cells was greatly increased in both subassemblies, such that the overall distributions of Always, Sometimes, and Never cells within the subassemblies were consistently different than their respective shuffle controls (Figure S2F–M). Together, these results further support that although the large majority of ensemble members are labile, there are core members that are consistently recruited and participate in all iterations of a given ensemble. Even within the core members however, the contribution to the ensemble remains labile with cells displaying excited and inhibited responses across iterations. Further, it appears that while the microcircuitry is organized such that most cells are recruitable to any given network, a minority are excluded or have a very low probability of being recruited by a given input.

To test whether instability of ensemble membership is reflective of the differential stimulation parameters used above, rather than inherent stochasticity, we performed an additional experiment ( $n = 677$  cells) where identical stimulations (single pulse, 5 mA) were delivered at the same interval as the stimulation curve experiment (i.e. 3 minutes apart). The results bear striking similarity to those observed in the increasing stimulation intensity experiment, wherein a high degree of turnover between subsequent stimulations (both  $S^A \rightarrow S^B$  and  $S^B \rightarrow S^C$ ) was observed (Figure 3A–B). A direct comparison of the three identical stimulations (5, 5, 5 mA) and the ascending (4,5,6 mA) stimulations shows that the fraction of the population that remains unrecruited across stimulations (i.e. had not yet participated in any ensemble) does not differ across the two paradigms (Figure 3C). Further, the average percent of the population that is either newly recruited (i.e. not participating in the previous stimulation) or retained (i.e. participating in the previous stimulation) is also unchanged (Figure 3D). In separate slices ( $n=1,043$  cells), to determine if stimulation replicates were sufficiently spaced to avoid cross-trial effects, we repeated the stimulation intensity curve experiment (2 to 10 mA) but randomized the order in which stimulations were delivered, which were then reordered during analysis to facilitate comparison to the prior experiments. Again, inherent stochasticity prevailed throughout the iterations, as over 50% of the cells

recruited to the half-maximal  $S^D$  stimulation were not recruited to the lowest stimulation intensity  $S^A$  (Figure S3A–B) and overall the percent of cells recruited with each successive stimulation did not differ whether stimulation intensities increased in ascending order or in a randomized fashion (Figure S3C). The results demonstrate that the outcome of each iteration was not directly influenced by the intensity of the prior stimulation, and that the degree of turnover is high across any of the conditions tested.

Notably, on the single-cell level, there was a striking degree of spontaneous activity in the absence of exogenous stimulation (Video S1). This is in stark contrast to previous reports that rodent PFC displays high rates of spontaneous activity *in vivo* but is largely silent when observed in the absence of extrinsic inputs<sup>13,39,40</sup>. To explore this observation, we first used event detection methods to quantify event rates throughout the recorded single cell traces. Descriptive statistics summarizing all cells in the intensity curve experiment revealed that indeed, high event rates were commonly seen throughout the population, which had a mean of 16.37 events/minute (SEM:  $\pm 0.16$ ) (Figure S3D). To explore whether probability of recruitment to ensembles might be weighted based on physiological features, we next quantified event rate only during each of the pre-stimulation baseline periods and correlated this value with each cells' recruitment probability over the ensemble iterations. This analysis revealed a negative relationship between spontaneous event rate and recruitment probability (Figure S3E), an effect which was particularly pronounced in the few Always cells which displayed event rates in the lower quartile of the population prior to stimulation. Intrinsic excitability is thought to be a critical determinant of ensemble formation<sup>41</sup>; however, in mouse somatosensory cortices cells, which are spontaneously active are more likely to participate in evoked ensembles<sup>43,44</sup>, opposite of the relationship observed here. It is intriguing to speculate that preferential recruitment of low activity cells in dlPFC may support greater diversity of ensemble membership as compared to the sensory cortices. It is also conceivable that it may be advantageous for ensembles generated from internal cognitive processes to compete (i.e. be generated in parallel) but not directly interrupt an ensemble event (i.e. override ongoing activity). Thus, recruitment of low activity cells may protect members of spontaneous ensembles from being commandeered by a cooccurring intracortical ensemble activation.

Next, to determine if recruitment probability might also be related to relative spatial location, we rank ordered cells by their y pixel coordinate relative to the co-recorded cells within each field of view. Because each slice was recorded with the cortical surface, where the stimulator was placed, in a consistent orientation with respect to the objective, the rank location serves as a relative measure of distance from the surface/stimulator. We found a modest relationship between location and recruitment probability, with greater distance from the sulcus being associated with higher likelihood of recruitment (Figure S3F). This effect was driven entirely by inhibitory responses however, as the same relationship was observed when considering only the inhibitory subassembly and was not present in the excited subassembly (Figure S3G–H). Together, these results lend further credence to the idea that there are underlying cellular features that dictate ensemble recruitment for certain members despite the apparent stochasticity at the population level, and indicate that there is a spatial organization underlying the recruitment of inhibitory networks during intracortical

signal propagation. While future work will be needed to elucidate the cell types involved, the spatial organization of inhibitory networks may be more readily apparent due to the relative sparsity and laminar expression patterns of inhibitory neurons<sup>44</sup>; we posit that the increased likelihood of inhibitory recruitment as a function of distance from the cortical surface may reflect feed-forward inhibition through GABAergic interneurons.

Our results thus far demonstrate that the intrinsic architecture of dlPFC microcircuitry imposes significant lability in the paths through which recurrent signals are propagated. This leads to strikingly high rates of ensemble turnover, even over short timescales and in the absence of co-occurring inputs. Such variability raises important questions as to how this architecture influences computations within the microcircuit: does such a seemingly random wiring schema necessitate degraded representational features of the network, or can it coexist with stable, high-fidelity encoding? To address this, we assessed the collective dynamics of the microcircuit ensembles by plotting neural trajectories derived from principal component analysis (PCA) of the single-cell calcium traces (Figure 3E–J, Figure S3I–J). Remarkably, despite the diverse cellular composition, conserved dynamics were observed, as indicated by the similar paths taken in PCA space across both stimulation intensity ensembles (Figure 3E) and train length ensembles (Figure 3H).

To further explore the relationship, or lack thereof, between ensemble membership fluidity and collective dynamics, we performed the same analysis on subpopulations of cells that display consistent membership across all ensemble iterations (i.e. Always and Never cells) compared to cells which change membership status for at least one iteration (Sometimes cells). For both the varying stimulation intensity (Figure 3F–G) and varying train length experiments (Figure 3I–J), the trajectories exhibited qualitatively similar properties. These results demonstrate that despite continuous turnover in ensemble membership, the representation of the stimulation conveyed by the collective ensemble dynamics retains high fidelity across iterations.

Finally, we sought to determine whether individual differences in behavioral metrics of cognitive performance were related to microcircuit excitability observed in isolated dlPFC *ex vivo* and, if so, whether these relationships would be reflected in readouts of the bulk fluorescence signals, ensemble membership stability, or low dimensional features of the single cell dynamics. All subjects from the recordings presented had gone through behavioral testing which included attentional set-shifting (Figure 4A). Based on extensive prior characterization of this task, we selected three measures of individual performance *a priori*, which have been shown to reliably capture meaningful individual differences in performance: the median latency to touch the screen over sessions (Figure 4B), change in performance over sessions as measured by the slope of the performance index ( $P_1$ ) across sessions (Figure 4C), and the mean perseverative error (PE) rate across sessions (Figure 4D)<sup>42,46–49</sup>. To probe for relationships between *ex vivo* activity and *in vivo* cognitive flexibility, we correlated the three summary metrics of behavioral performance with three facets of microcircuit dynamics assayed in the stimulation intensity curve experiment: overall excitability, ensemble membership stability, and fidelity of low dimensional dynamics across stimulations. As a metric of overall excitability, we calculated the linear slope of the peak amplitudes of the bulk fluorescence signal across stimulations (Figure

S4A). To create a summary metric for ensemble stability, the rate at which unique members were recruited across ensemble iterations was calculated for each animal (Figure S4E). As a measure of encoding similarity across ensemble iterations, we summed the absolute residuals of pairwise subtractions of the neural trajectories and rank ordered animals based on the outcome (Figure S4I). Of these three summary metrics, only overall excitability was found to covary with individual behavioral performance whereby high excitability was associated with high performance index (Figure 4E). Thus, our data directly link the sum total of intrinsic features of dIPFC microcircuits, observable in an isolated preparation, with higher-order cognitive functions and suggest that the ability to produce similar population-level signals from predominantly distributed, labile networks is a key feature of dIPFC function.

In summary, we have established a viable approach for the observation of large-scale population dynamics in *ex vivo* non-human primate brain tissue, circumventing long-standing technical limitations, and thereby revealing inherent properties of the dIPFC. We demonstrated that the intrinsic architecture of the dIPFC is capable of propagating signals through recurrent connections that can be effectively shaped into distinct representations of a variety of stimulus features. Our findings further indicated that activity in response to discrete inputs propagates through largely disparate and flexible networks of cells, supporting the wealth of evidence for distributed population representations across a cortical region, and also suggested that a small minority of core members in the network may play specialized or essential roles. The degree of ensemble turnover, even when extrinsic factors are removed, argues strongly against an engram or localist encoding schema. Finally, we revealed covariation between the static features of intracortical excitability observed *ex vivo* and individual differences in cognitive performance, thus linking cognitive abilities with microcircuit properties in an evolutionarily specialized region of PFC specific to primates. These results contribute to the understanding of intracortical microcircuitry and its relevance to cognitive processing, providing a foundation for further investigations into the synaptic mechanisms underlying cognition.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Cody Siciliano (cody.siciliano@vanderbilt.edu).

### Materials availability

This study did not generate new unique reagent.

### Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## STAR METHODS

### EXPERIMENTAL MODEL DETAILS

Fifteen rhesus macaques (*Macaca mulatta*, nine female and six male) averaging 9 years of age were used for the current studies. Animals were individually housed in quadrant cages (0.8 × 0.8 × 0.9m) with constant temperature (20–22°C) and humidity (65%) and an 11-hour-light cycle (lights on at 8 am). Animals were housed in the same room with visual and auditory access to all subjects in the room. In addition, monkeys were socially housed with a same-sex “partner” monkey for at least 2 hours per day or provided grooming contact if full pairing was contraindicated.

Each housing cage had a computer-operated panel embedded in the side wall that provided all food, fluids, and cognitive task presentations (additional details in<sup>45–50</sup>). Body weights were taken weekly. All animal procedures and experiments were approved by the Oregon National Primate Research Center (ONPRC) Institutional Animal Care and Use Committee and followed recommendations of the *Guide for the Care and Use of Laboratory Animals* (National Research Council 2011). Additionally, ONPRC is fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International.

### METHOD DETAILS

**Tissue extraction and slice preparation**—Tissue was prepared as described previously and delivered to experimenters blinded to individual animal identities. For the duration of the experiment and subsequent analyses, experimenters were kept blinded to all aspects of the subjects identity and assigned the animals a novel identification number which was decoded following analysis video processing, trace extraction, and initial analyses. Briefly, subjects were deeply anesthetized with ketamine (10 mg/kg) and maintained on isoflurane as previously described<sup>18,19</sup>. Monkeys were then perfused with ice-cold oxygenated perfusion solution [containing (in mM) 124 NaCl, 23 NaHCO<sub>3</sub>, 3 NaH<sub>2</sub>PO<sub>4</sub>, 5 KCl, 2 MgSO<sub>4</sub>, 10 D-glucose, 2 CaCl<sub>2</sub>]. Next, brains were removed quickly and sectioned along the coronal plane using a brain matrix (Electron Microscopy Sciences, Hartfield, PA, USA). The brain knife position was preplanned according to each subject’s prior MRI results. An isolated tissue block containing prefrontal area 46 was placed in ice-cold oxygenated perfusion solution on ice for subsequent transport between where monkeys were housed and the building where slicing was conducted.

A ceramic blade was attached to a vibrating tissue slicer to prepare 250 μm thick coronal brain sections containing area 46 of the dorsolateral prefrontal cortex as previously described<sup>19,51,52</sup>. Tissue was sliced in cold, oxygenated cutting solution containing the following in mM: 194 sucrose, 30 NaCl, 4.5 KCl, 1 MgCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 1.2 NaH<sub>2</sub>PO<sub>4</sub>, and 10 glucose.

**Slice Incubation and Dye Loading**—Once sliced, tissue was transferred to an incubation chamber containing oxygenated artificial CSF (aCSF). aCSF was used for all experimental procedures post-slicing, and contained the following in mM: 126 NaCl, 2.5 KCl, 1.2 NaH<sub>2</sub>PO<sub>4</sub>, 2.4 CaCl<sub>2</sub>, 1.2 MgCl<sub>2</sub>, 25 NaHCO<sub>3</sub>, 11 D-glucose, and 0.4 L-ascorbic

acid (pH adjusted to 7.4). Slices were divided into floating mesh-bottom wells which held the slices just below the surface of the aCSF reservoir, which was held inside a lighttight incubation chamber (Braincubator, PAYO Scientific, Kingswood, AUS)<sup>53</sup>. The incubation chamber consisted of a Peltier plate which held the reservoir, temperature and pH probes, aerator delivering 95%/5% carbogen, and a recirculating pump. Temperature was maintained at 15 degrees C and pH was adjusted back to 7.4 if any drift occurred. To maximize slice viability, aCSF was continuously removed from one side of the reservoir and returned on the other after passing through a UVC irradiation filter in a secondary chamber which was optically isolated from the slices. This ensures a constant supply of oxygenated, clean aCSF and leads to extended slice viability<sup>54,55</sup>. Slices were incubated in this chamber for at least 1 hour before dye loading.

After a minimum of 1 hour in the incubation chamber post-slicing, slices were removed and loaded with 5  $\mu$ M fluo-4 AM. Dye solution was made by dissolving lipophilized fluo-4 AM in pure DMSO containing 1% w/v pluronic acid (made fresh within the hour) to a concentration of 5 mM. Based on extensive parameterization of dye loading protocols from Yuste and colleagues<sup>56,57</sup>, the dye solution was pipetted directly onto the surface of the slice such that a high concentration was briefly achieved before rapidly diffusing into the chamber and equilibrating to the final concentration. 4  $\mu$ L of dye solution was pipetted on top of the slice which was held in a low-volume oxygenated loading chamber containing aCSF (Brain Slice Keeper, Scientific Systems Designs Inc, Ontario, CAN). The total volume in the loading chamber was 4 mL, thus final concentrations were: 0.1% v/v DMSO, 0.001% w/v pluronic acid, and 5  $\mu$ M fluo-4 AM in aCSF. Slices were protected from light throughout the process. Following loading, slices were rinsed briefly in dye-free, pre-oxygenated aCSF before transferring them back to the incubation chamber. Slices were allowed to rest for at least 30 minutes before transferring to the imaging system.

**Overview of Imaging System**—The system used for recording *ex vivo* calcium dynamics consisted of a customized upright laser scanning confocal microscope (Thorlabs Imaging Systems, Sterling, VA) integrated with a slice perfusion chamber (Warner Instruments, Hamden, CT). While in the imaging chamber, the flow rate was monitored online using pass-through flow sensors (Sensirion Inc, Chicago, IL) to ensure a stable supply of dissolved oxygen was delivered to the slice and that pharmacological agents were delivered in a reproducible manner. To ensure microsecond repeatability of synchronization between image capture and stimulation, the microscope and stimulus isolator (NL800A Current Stimulus Isolator, Digitimer, Hertfordshire, UK) were triggered using one function generator with a delay line such that both systems were controlled by a single master clock (PulsePal, Open Ephys, Rochester, NY). Fluorescence excitation was achieved with a 488 nm laser line fiber-coupled to the scan head. Emission was collected through a pinhole confocal with the imaging plane, and passed to a PMT detection module through a multimode fiber. Emission was then split by a 550nm longpass dichroic mirror and the transmitted beam passed through a 525nm emission filter at the face of a cooled GaAsP PMT sampled at 80 MHz.

**Image Acquisition**—Following a minimum 30-minutes rest period after dye loading, slices were transferred to the imaging chamber and immersed in aCSF which was flowing at a rate of 2 mL/min. For evoked recordings, a 4x magnification objective (0.13 NA, 17.2 mm WD, Nikon) was first utilized to assess gross anatomical features and place the stimulating electrode before switching to a 16x water dipping objective (0.80 NA, 3.0 mm WD, Nikon) for the duration of the recordings.

During experimental recordings, images were acquired at  $1024 \times 1024$  pixel density across a  $762\mu\text{m}^2$  field of view ( $744 \text{ nm}^2$  pixel size) at 15.14 frames per second (fps). For evoked recordings, stimulation onset was 5 seconds from the start of the frame capture and activity was recorded for 14.82 seconds following stimulation onset (19.82 seconds per recording, corresponding to 300 frames). The resonance scanner was kept constantly running during the inter-stimulation interval to reduce thermal drift.

**Electrical and Pharmacological Manipulations**—Localized electrical stimulation was delivered via a twisted bipolar electrode (0.127mm diameter stainless steel 2-channel electrode, PlasticsOne, Roanoke, VA) held by a micromanipulator (MM-3 Micromanipulator, Narishige, Amityville, NY) and positioned superficially with contacts parallel to the principal sulcus. Calcium dynamics were evoked using a variety of stimulation parameters delivered to the slice. To determine the input-output relationship with varying stimulation amperage, we delivered a single pulse ranging from 2–10 mA, 4 ms wide, monophasic, constant current. The stimulation intensities selected here are within the range used in clinical application of intracortical stimulations<sup>58–60</sup>. For the multi-pulse data, pulse intensities were held constant at 5 mA, and consisted of 4 ms wide monophasic pulses with train lengths increasing from 1 to 20 pulses. To determine the stability of ensembles and the replicability of our findings, two additional stimulation paradigms were tested in separate slices. For the first one, a single pulse at 5 mA, 4ms wide, monophasic, constant current, was repeated every three minutes for three collections to determine if ensemble membership lability was similar to or different from the increasing intensity paradigm. In a second experiment, the same stimulation parameters (a single pulse ranging from 2–10 mA, 4 ms wide, monophasic, constant current) were randomly ordered rather than delivered in an ascending series. For the calcium chelation experiment, aCSF containing  $10 \mu\text{M}$  BAPTA-AM (CAS: 126150–97-8, Tocris Biosciences, Minneapolis, MN, USA) was perfused onto the slice at a rate of 2 mL per minute for at least 30 minutes after the baseline recordings were taken.

### Calcium Imaging Analysis

**Image processing and bulk fluorescence signal extraction:** Following the acquisition, tiff frames were collated into one file for each sweep/recording using customized FIJI macros. Videos were then visually inspected and motion corrected, if needed, in Inscopix Data Processing software (Inscopix, Mountain View, CA). The bulk fluorescence signal, or whole field activity, was determined by averaging pixel intensities over the entire field of view to create a single brightness over time trace for each video, expressed in units of  $\Delta F/F$  calculated as  $(F-F_0)/F_0$  for each sample in the trace  $F$  and  $F_0$  equal to the mean intensity of the entire trace. To quantify the stimulation-evoked activity, traces were z-score normalized

to the first 70 frames (the pre-stimulation period spanned from frame 1:75), and the traces were then transformed such that the value of the last frame before the stimulation was set to zero and this differential was applied to the entire trace. Parameters of interest, peak amplitude following stimulation and AUC of the decay phase were then calculated using custom MATLAB scripts.

**Single cell demixing and quantification:** After the steps listed above, videos that were processed for whole-field analysis were then co-registered across sweeps using PatchWarp<sup>61</sup> to ensure that cells could be tracked effectively across the time series. Following co-registration, videos were then frame averaged by a factor of 2 bringing the final framerate to 7.58 fps and binned by a factor of 2 bringing the final pixel size to 1.49  $\mu\text{m}^2$ . We then used a constrained non-negative matrix factorization (CNMF) algorithm to extract fluorescence traces from cells within the field of view<sup>62</sup>. All spatial footprints were then visually inspected following CNMF and any overlapping cells or cells that contained multiple cells in one footprint were removed.

For stimulus-evoked activity, representative heatmaps were created using traces that had been z-scored to a pre-stimulus window (-32 through -7 frames prior to the stimulation onset, roughly corresponding to 1 after the start of the recording through 1 second prior to the stimulation). Cells were classified as excited, inhibited, or no response based on the area of the curve during the 2 seconds following the stimulation onset. For each stimulation, cells with an AUC of above 0.85 were classified as excited while cells with an AUC of below -0.85 were classified as inhibited, and the remainder were considered not responsive. These thresholds were empirically verified as effectively separating subpopulations and control experiments confirmed that both excited and inhibited responses reflected stimulation-evoked calcium dynamics (Figure S1).

### **Attentional Set-Shifting Task (ASST)**

**Apparatus:** For the set-shifting experiments, a computer-operated touchscreen panel was embedded in a side wall of each housing cage as previously described<sup>42,48</sup> and individual computers were controlled and inputs time-stamped using a centrally located computer. In addition to the touchscreen, there was a dowel below the screen, two drinking spouts on the screen's left and right sides, and an infrared finger-poke on the righthand side (Med Associates Inc, US). These operanda were connected to a food pellet dispenser that delivered a 1g banana-flavored food pellet as a reward. All panels were linked to a network and the main computer via a National Instruments interface and Labview software (LabView 2011, SPI, National Instruments, Texas, United States).

**Task design:** As depicted in Figure 4, all animals underwent set-shift testing, which has been examined extensively and described previously for other cohorts<sup>42,48</sup>. Briefly, animals were given the set shift sessions inside their housing cage each morning, alongside all other monkeys in the cohort. To establish touching the screen, highly preferred photographs (e.g., fruit, other monkeys) were projected and, when touched, expanded to the full screen for 3 sec<sup>42</sup>. Following the acquisition of a reliable operant response (i.e., touching the screen), set-shifting training began. During training, the monkeys were presented with a

random presentation of two geometrical shapes, and the trial ended following the selection of one of the shapes or after 30 seconds elapsed. Discrimination between which shape was predetermined as “correct” or “incorrect” was achieved by the contingent presentation of a 1g banana flavored food pellet and a presentation of the preferred photograph for a correct response and the contingent presentation of a blank screen for 10 seconds following an incorrect response. Subjects were gradually moved to a second order of FR1 (FR3) where every correct response earned a presentation of the preferred photo and every 3<sup>rd</sup> correct response resulted in a pellet, to prevent satiety over the session. The criterion for the training phase was that responding was maintained throughout the 45-minute session. All animals received 7 training days on the second-order schedule.

Following training, animals were given 30 consecutive sessions to assess set-shifting performance followed by an additional 10 days at a later point. All behavioral variables were calculated by taking an average of the performance at the two timepoints. Within each session, there was a maximum of 4 original types of discrimination sets – simple, compound, and intradimensional and extradimensional shifts (shown in Figure 4A) each followed by a reversal set. A set consisted of a pair of shapes and each set retained the same two shapes and colors for all trials. The color of each shape and the side (right or left of center screen) of the screen it was presented was randomly chosen in each trial. Reversal trials included the same two shapes and colors with only the correct contingency being reversed. A set was considered acquired when a running twelve out of 15 consecutive trials was deemed correct. In simple discrimination, both shapes were either black or white, therefore only shape could be used as the basis of the discrimination. In the compound discriminations, the shapes were new as were the colors of the shapes, with the discrimination still based on the shape of the object. In the intradimensional shift set, there was a different combination of shapes and colors, and the discrimination was based on shape again. The fourth and final set represented an extradimensional shift, presenting a different combination of unique colors and shapes but with the correct discrimination now based on color. Progression through the sets was self-paced and the overall session ended after 45 minutes elapsed or criteria being met for all eight sets in the allotted time.

**Behavioral data analysis and correlations:** Behavioral variables of interest were selected *a priori* based on extensive characterizations in previous work<sup>42,46–49</sup>. From these works, we chose the performance index, the median latency to touch the screen, and the average ratio of the perseverative errors to total errors (non-reinforced stimuli selection) as our three axes of behavioral performance. Perseverative errors were defined as errors during the reversal phase, where choices were based on the prior contingency<sup>63</sup>. All variables were examined as a composite measure across the first 30 days and the second 10-day re-test. These behavioral variables were correlated with three facets of *ex vivo* activity: 1) overall excitability, as measured by the linear slope of the peak amplitudes of bulk fluorescence signal across stimulations (Figure 1G, shown individually in Figure S4A), 2) ensemble stability, as measured by the Hill slope of the cumulative recruitment curve (shown in Figure S4E), and 3) encoding fidelity, measured as the absolute residuals for each pairwise comparison of the PCA trajectory paths across stimulations (shown in Figure S4J). For trajectory comparisons, the absolute residuals from all comparisons were summed to create

an overall index of dissimilarity between trajectories. Each subject was assigned a rank based on this value, with a rank of 1 indicating the highest degree of similar between trajectories.

## QUANTIFICATION AND STATISTICAL ANALYSIS

All details regarding individual statistical comparisons can be found within the relevant figure legends. Statistical comparisons were conducted using Graphpad Prism 10 Software (Boston, MA, USA) except for the permutation/reshuffling analyses in Figure 2 and Figure S2, which was performed using custom MATLAB scripts.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments:

This work was supported by NIH grants R00 DA04510 (C.A.S.), R01 AA030115 (C.A.S.), U01 AA029971 (C.A.S.), U01 AA013510 (K.A.G.), R24 AA019431 (K.A.G.), U24 AA013641 (K.A.G.), P51 OD0119092 (K.A.G.) as well as support from Alkermes Pathways Research Award (C.A.S.), the Brain Research Foundation (C.A.S.), and the Whitehall Foundation (C.A.S.). S.O.N. (F32 DA051136), P.R.M. (F31 AA029626), K.R.E. (F31 DA056202), and Z.Z.F. (F31 DA056196) were supported by NIH fellowships.

## REFERENCES

1. Miller EK, and Cohen JD (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. [PubMed: 11283309]
2. Goldman-Rakic PS (1995). Architecture of the prefrontal cortex and the central executive. *Ann. N. Y. Acad. Sci.* 769, 71–83. [PubMed: 8595045]
3. Douglas RJ, and Martin KA (1991). A functional microcircuit for cat visual cortex. *The Journal of physiology* 440, 735–769. [PubMed: 1666655]
4. Douglas RJ, Koch C, Mahowald M, Martin KA, and Suarez HH (1995). Recurrent excitation in neocortical circuits. *Science* 269, 981–985. [PubMed: 7638624]
5. Christos C, Shintaro F, Daeyeol L, John DM, Xue-Lian Q, Min W, and Amy FTA (2018). Persistent Spiking Activity Underlies Working Memory. *The Journal of Neuroscience* 38, 7020. [PubMed: 30089641]
6. Goldman-Rakic PS, Cools AR, Srivastava K, Roberts AC, Robbins TW, and Weiskrantz L (1996). The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351, 1445–1453. [PubMed: 8941956]
7. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, and Friston KJ (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. [PubMed: 23177956]
8. Douglas RJ, and Martin KAC (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451. [PubMed: 15217339]
9. Zylberberg J, and Strowbridge BW (2017). Mechanisms of Persistent Activity in Cortical Circuits: Possible Neural Substrates for Working Memory. *Annu. Rev. Neurosci.* 40, 603–627. [PubMed: 28772102]
10. Wimmer K, Nykamp DQ, Constantinidis C, and Compte A (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* 17, 431–439. [PubMed: 24487232]
11. Stokes MG (2015). ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences* 19, 394–405. [PubMed: 26051384]

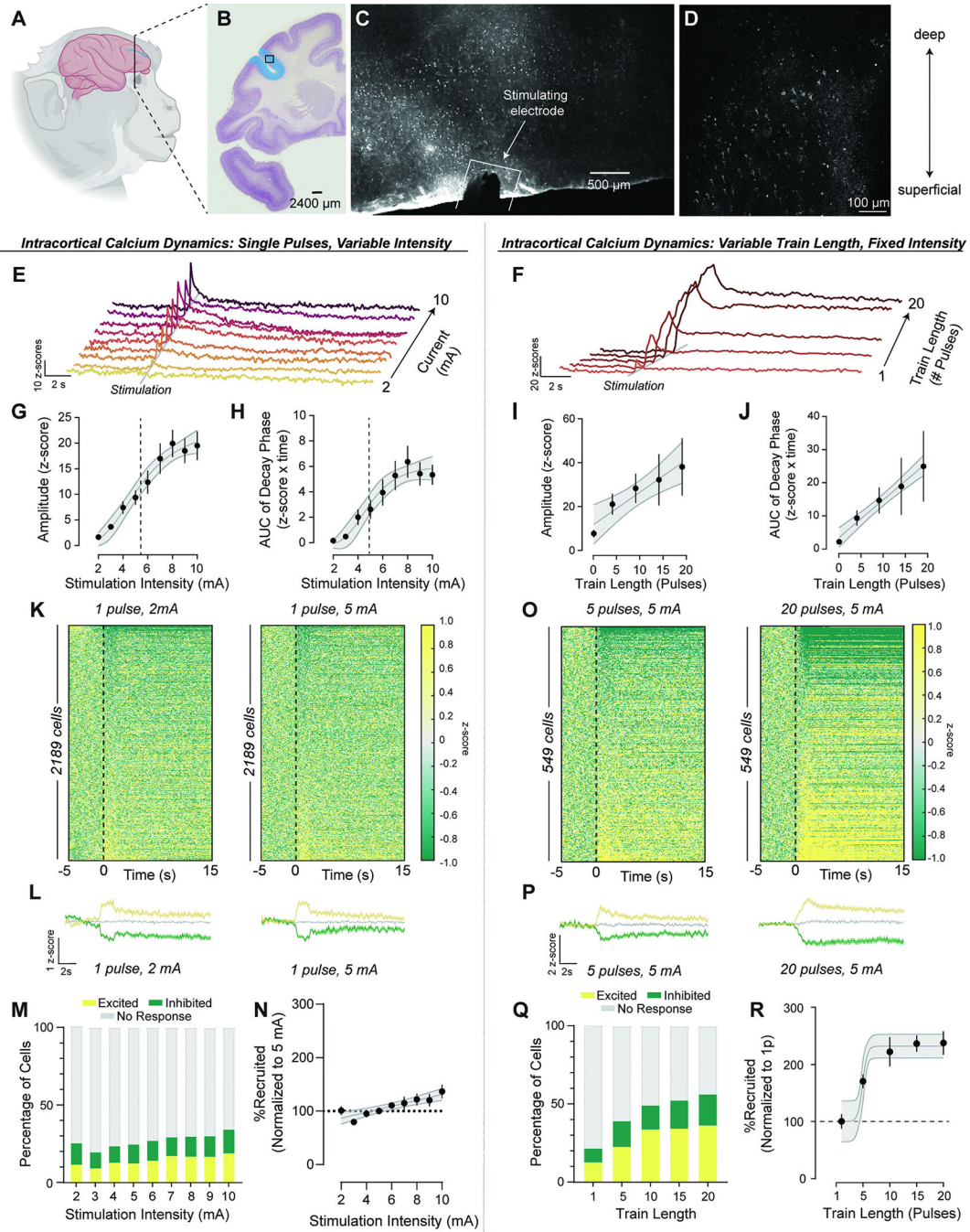
12. Lundqvist M, Herman P, and Miller EK (2018). Working Memory: Delay Activity, Yes! Persistent Activity? Maybe Not. *J Neurosci* 38, 7013–7019. [PubMed: 30089640]
13. Blaeser AS, Connors BW, and Nurmikko AV (2017). Spontaneous dynamics of neural networks in deep layers of prefrontal cortex. *Journal of Neurophysiology* 117, 1581–1594. [PubMed: 28123005]
14. Kodama NX, Feng T, Ullett JJ, Chiel HJ, Sivakumar SS, and Galán RF (2018). Anti-correlated cortical networks arise from spontaneous neuronal dynamics at slow timescales. *Scientific Reports* 8, 666. [PubMed: 29330480]
15. Luczak A, and Maclean JN (2012). Default activity patterns at the neocortical microcircuit level. *Front. Integr. Neurosci.* 6, 30. [PubMed: 22701405]
16. Preuss TM (1995). Do rats have prefrontal cortex? The rose-woolsey-akert program reconsidered. *J. Cogn. Neurosci.* 7, 1–24. [PubMed: 23961750]
17. Carlén M (2017). What constitutes the prefrontal cortex? *Science* 358, 478–482. [PubMed: 29074767]
18. Cuzon Carlson VC, Seabold GK, Helms CM, Garg N, Odagiri M, Rau AR, Daunais J, Alvarez VA, Lovinger DM, and Grant KA (2011). Synaptic and morphological neuroadaptations in the putamen associated with long-term, relapsing alcohol drinking in primates. *Neuropsychopharmacology* 36, 2513–2528. [PubMed: 21796110]
19. Daunais JB, Kraft RA, Davenport AT, Burnett EJ, Maxey VM, Szeliga KT, Rau AR, Flory GS, Hemby SE, Kroenke CD, et al. (2010). MRI-guided dissection of the nonhuman primate brain: A case study. *Methods* 50, 199–204. [PubMed: 19364532]
20. Gee KR, Brown KA, Chen WN, Bishop-Stewart J, Gray D, and Johnson I (2000). Chemical and physiological characterization of fluo-4 Ca(2+)-indicator dyes. *Cell Calcium* 27, 97–106. [PubMed: 10756976]
21. Monette R, Small DL, Mealing G, and Morley P (1998). A fluorescence confocal assay to assess neuronal viability in brain slices. *Brain Research Protocols* 2, 99–108. [PubMed: 9473610]
22. Lock JT, Parker I, and Smith IF (2015). A comparison of fluorescent Ca<sup>2+</sup> indicators for imaging local Ca<sup>2+</sup> signals in cultured cells. *Cell Calcium* 58, 638–648. [PubMed: 26572560]
23. Reese AL, and Kavalali ET (2015). Spontaneous neurotransmission signals through store-driven Ca<sup>2+</sup> transients to maintain synaptic homeostasis. *eLife* 4, e09262. [PubMed: 26208337]
24. Izhikevich EM (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting* (The MIT Press).
25. Day M, Carr DB, Ulrich S, Ilijic E, Tkatch T, and Surmeier DJ (2005). Dendritic excitability of mouse frontal cortex pyramidal neurons is shaped by the interaction among HCN, Kir2, and K<sub>leak</sub> channels. *J. Neurosci.* 25, 8776–8787. [PubMed: 16177047]
26. Kwon S-K, Sando R 3rd, Lewis TL, Hirabayashi Y, Maximov A, and Polleux F (2016). LKB1 Regulates Mitochondria-Dependent Presynaptic Calcium Clearance and Neurotransmitter Release Properties at Excitatory Synapses along Cortical Axons. *PLoS Biol.* 14, e1002516. [PubMed: 27429220]
27. Ferguson KA, and Cardin JA (2020). Mechanisms underlying gain modulation in the cortex. *Nat. Rev. Neurosci.* 21, 80–92. [PubMed: 31911627]
28. Cardin JA, Kumbhani RD, Contreras D, and Palmer LA (2010). Cellular mechanisms of temporal sensitivity in visual cortex neurons. *J. Neurosci.* 30, 3652–3662. [PubMed: 20219999]
29. Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R, and Paninski L (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* 104, 3691–3704. [PubMed: 20554834]
30. Friedrich J, Yang W, Soudry D, Mu Y, and Paninski L (2016). Multi-scale approaches for high-speed imaging and analysis of large neural populations. 10.1101/091132.
31. Vander Weele CM, Siciliano CA, Matthews GA, Namburi P, Izadmehr EM, Espinel IC, Nieh EH, Schut EHS, Padilla-Coreano N, Burgos-Robles A, et al. (2018). Dopamine enhances signal-to-noise ratio in cortical-brainstem encoding of aversive stimuli. *Nature* 563, 397–401. [PubMed: 30405240]
32. Siciliano CA, and Tye KM (2019). Leveraging calcium imaging to illuminate circuit dysfunction in addiction. *Alcohol* 74, 47–63. [PubMed: 30470589]

33. Hubel DH, and Wiesel TN (1959). Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148, 574–591. [PubMed: 14403679]
34. Buschman TJ, Denovellis EL, Diogo C, Bullock D, and Miller EK (2012). Synchronous Oscillatory Neural Ensembles for Rules in the Prefrontal Cortex. *Neuron* 76, 838–846. [PubMed: 23177967]
35. Chiang F-K, Wallis JD, and Rich EL (2022). Cognitive strategies shift information from single neurons to populations in prefrontal cortex. *Neuron* 110, 709–721.e4. [PubMed: 34932940]
36. Duncan J (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience* 2, 820–829. [PubMed: 11715058]
37. Balaguer-Ballester E, Lapiush CC, Seamans JK, and Durstewitz D (2011). Attracting Dynamics of Frontal Cortex Ensembles during Memory-Guided Decision-Making. *PLOS Computational Biology* 7, e1002057. [PubMed: 21625577]
38. Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, Adam KCS, Valls-Solé J, Constantinidis C, and Compte A (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature Neuroscience* 23, 1016–1024. [PubMed: 32572236]
39. Pirot S, Jay TM, Glowinski J, and Thierry AM (1994). Anatomical and electrophysiological evidence for an excitatory amino acid pathway from the thalamic mediodorsal nucleus to the prefrontal cortex in the rat. *Eur. J. Neurosci.* 6, 1225–1234. [PubMed: 7524967]
40. Pires J, Nelissen R, Mansvelter HD, and Meredith RM (2021). Spontaneous synchronous network activity in the neonatal development of mPFC in mice. *Developmental Neurobiology* 81, 207–225. [PubMed: 33453138]
41. Hansel C, and Yuste R (2024). Neural ensembles: role of intrinsic excitability and its plasticity. *Front. Cell. Neurosci.* 18, 1440588. [PubMed: 39144154]
42. MacLean JN, Watson BO, Aaron GB, and Yuste R (2005). Internal dynamics determine the cortical response to thalamic stimulation. *Neuron* 48, 811–823. [PubMed: 16337918]
43. Carrillo-Reid L, Yang W, Kang Miller J-E, Peterka DS, and Yuste R (2017). Imaging and optically manipulating neuronal ensembles. *Annu. Rev. Biophys.* 46, 271–293. [PubMed: 28301770]
44. Ma S, Skarica M, Li Q, Xu C, Risgaard RD, Tebbenkamp ATN, Mato-Blanco X, Kovner R, Krsnik Ž, de Martin X, et al. (2022). Molecular and cellular evolution of the primate dorsolateral prefrontal cortex. *Science* 377, eabo7257. [PubMed: 36007006]
45. Shnitko TA, Allen DC, Gonzales SW, Walter NAR, and Grant KA (2017). Ranking Cognitive Flexibility in a Group Setting of Rhesus Monkeys with a Set-Shifting Procedure. *Front. Behav. Neurosci.* 11, 55. [PubMed: 28386222]
46. Shnitko TA, Gonzales SW, and Grant KA (2019). Low cognitive flexibility as a risk for heavy alcohol drinking in non-human primates. *Alcohol* 74, 95–104. [PubMed: 30097387]
47. Shnitko TA, Gonzales SW, Newman N, and Grant KA (2020). Behavioral Flexibility in Alcohol-Drinking Monkeys: The Morning After. *Alcohol. Clin. Exp. Res.* 44, 729–737. [PubMed: 31984521]
48. Grant KA, Newman N, Gonzales S, and Shnitko TA (2021). Replicability in measures of attentional set-shifting task performance predicting chronic heavy drinking in rhesus monkeys. *Alcohol* 96, 93–98. [PubMed: 34509594]
49. Grant KA, Newman N, Lynn C, Davenport C, Gonzales S, Carlson VCC, and Kroenke CD (2022). Brain Functional Connectivity Mapping of Behavioral Flexibility in Rhesus Monkeys. *The Journal of Neuroscience* 42, 4867–4878. [PubMed: 35552233]
50. Baker EJ, Farro J, Gonzales S, Helms C, and Grant KA (2014). Chronic alcohol self-administration in monkeys shows long-term quantity/frequency categorical stability. *Alcoholism: Clinical and Experimental Research* 38, 2835–2843. [PubMed: 25421519]
51. Siciliano CA, Calipari ES, Cuzon Carlson VC, Helms CM, Lovinger DM, Grant KA, and Jones SR (2015). Voluntary ethanol intake predicts  $\kappa$ -opioid receptor supersensitivity and regionally distinct dopaminergic adaptations in macaques. *J. Neurosci.* 35, 5959–5968. [PubMed: 25878269]
52. Siciliano CA, Calipari ES, Yorgason JT, Lovinger DM, Mateo Y, Jimenez VA, Helms CM, Grant KA, and Jones SR (2016). Increased presynaptic regulation of dopamine neurotransmission in

- the nucleus accumbens core following chronic ethanol self-administration in female macaques. *Psychopharmacology* 233, 1435–1443. [PubMed: 26892380]
53. Krimer LS, and Goldman-Rakic PS (1997). An interface holding chamber for anatomical and physiological studies of living brain slices. *Journal of Neuroscience Methods* 75, 55–58. [PubMed: 9262144]
  54. Breen PP, and Buskila Y (2014). Braincubator: An incubation system to extend brain slice lifespan for use in neurophysiology (IEEE).
  55. Buskila Y, Breen PP, Tapson J, van Schaik A, Barton M, and Morley JW (2014). Extending the viability of acute brain slices. *Scientific reports* 4, 5309. [PubMed: 24930889]
  56. Ikegaya Y, Le Bon-Jego M, and Yuste R (2005). Large-scale imaging of cortical network activity with calcium indicators. *Neurosci. Res.* 52, 132–138. [PubMed: 15893573]
  57. Yuste R, MacLean J, Vogelstein J, and Paninski L (2011). Imaging Action Potentials with Calcium Indicators. *Cold Spring Harbor Protocols* 2011, db.prot5650.
  58. Johnson LA, Wander JD, Sarma D, Su DK, Fetzi EE, and Ojemann JG (2013). Direct electrical stimulation of the somatosensory cortex in humans using electrocorticography electrodes: a qualitative and quantitative report. *Journal of Neural Engineering* 10, 036021. [PubMed: 23665776]
  59. Signorelli F, Guyotat J, Mottolese C, Schneider F, D’Acunzi G, and Isnard J (2004). Intraoperative electrical stimulation mapping as an aid for surgery of intracranial lesions involving motor areas in children. *Child’s Nervous System* 20, 420–426.
  60. Kramer DR, Kellis S, Barbaro M, Salas MA, Nune G, Liu CY, Andersen RA, and Lee B (2019). Technical considerations for generating somatosensation via cortical stimulation in a closed-loop sensory/motor brain-computer interface system in humans. *J. Clin. Neurosci.* 63, 116–121. [PubMed: 30711286]
  61. Hattori R, and Komiyama T (2022). PatchWarp: Corrections of non-uniform image distortions in two-photon calcium imaging data by patchwork affine transformations. *Cell Reports Methods*, 100205. [PubMed: 35637910]
  62. Giovannucci A, Friedrich J, Gunn P, Kalfon J, Brown BL, Koay SA, Taxidis J, Najafi F, Gauthier JL, Zhou P, et al. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. *Elife* 8. 10.7554/eLife.38173.
  63. Grant KA, Newman N, Lynn C, Davenport C, Gonzales S, Cuzon Carlson VC, and Kroenke CD (2022). Brain Functional Connectivity Mapping of Behavioral Flexibility in Rhesus Monkeys. *J. Neurosci.* 42, 4867–4878. [PubMed: 35552233]

**Highlights:**

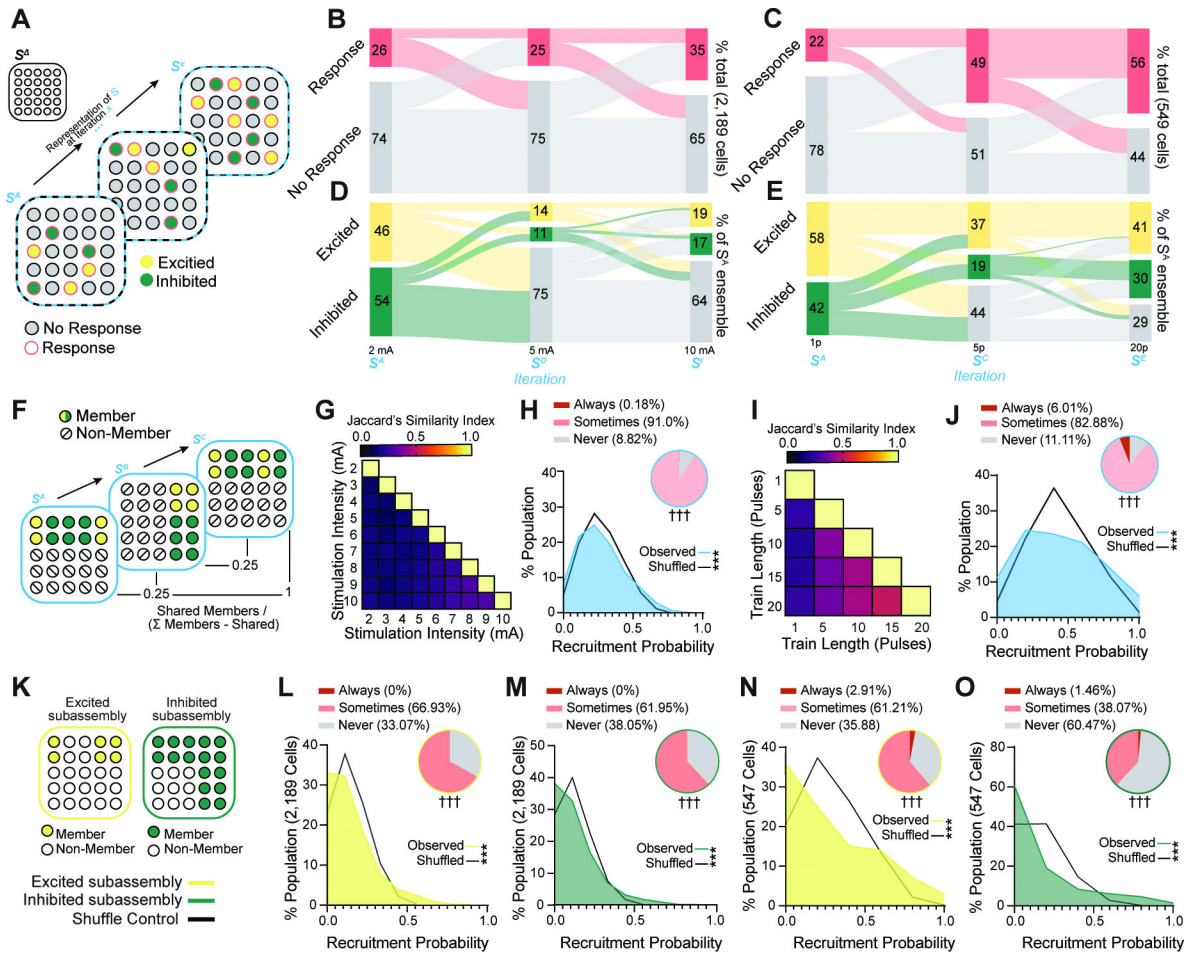
- dlPFC intrinsically maintains recurrent activity in the absence of extrinsic inputs
- Identical inputs are routed through labile, stochastically assembled ensembles
- Stable population-level representations arise from unstable ensembles
- Microcircuit excitability covaried with individual cognitive flexibility



**Figure 1. Localized peri-sulcal stimulation elicits far-field representations in population-level signals via recursion within scalable ensembles.**

(A) *Ex vivo* brain slices from adult rhesus macaques were prepared immediately post-mortem from tissue blocks sectioned in the coronal plane. (B) All recordings were performed in right Brodmann’s area 46 of the dorsolateral prefrontal cortex, indicated in blue (atlas image was reproduced from [BrainMaps.Org](https://www.brainmaps.org/)). (C) To deliver a localized input stimulus to the superficial cortex without inducing an electrical potential across layers, a twisted bipolar electrode was placed peri-sulcal with contacts parallel to the bank of the

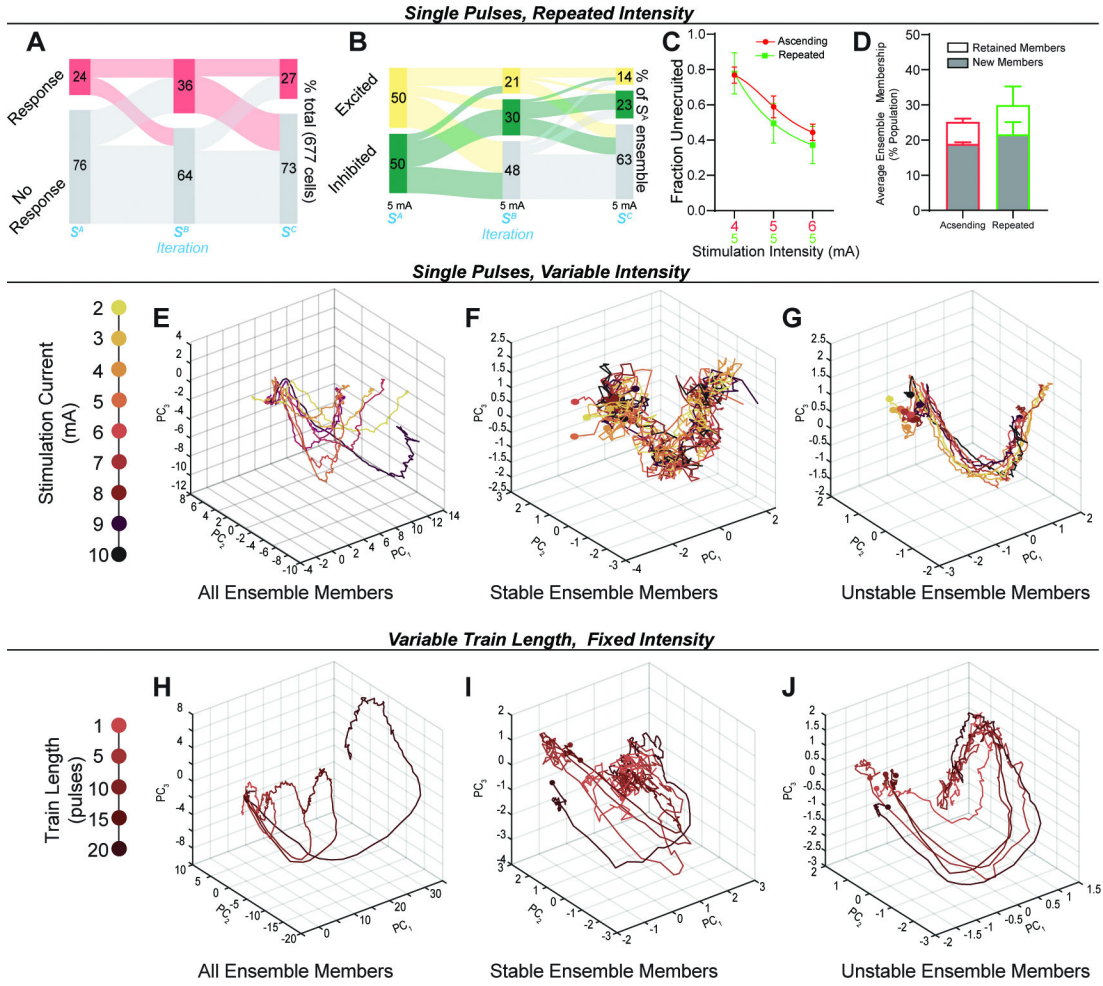
principal sulcus. The image depicts stimulation placement and gross anatomical features, which were assessed using a low-magnification air objective (a water dipping objective was used for all experimental recordings). **(D)** Cells were bulk loaded with the cell-permeable fluorescent calcium indicator fluo-4 AM and *ex vivo* calcium dynamics were observed across a 762  $\mu\text{m}^2$  field of view at 15.14 fps. The image shows a standard deviation T-projection from a representative field of view. **(E)** Single pulses with varying amplitude were delivered to test the input-output relationship across input intensities while all other parameters were held constant (monophasic, 4 ms pulse width). Representative brightness-over-time traces of bulk fluorescence demonstrate far-field responses evoked by peri-sulcal single pulse stimulations (2–10 mA constant current, ascending order, stimulation onset at 5 seconds). **(F)** In a parallel experiment, input-output relationships were examined for stimulations of varying train lengths while all other parameters were held constant (1–20 pulses, ascending order, 20 Hz). **(G)** The amplitude of evoked bulk fluorescence activity scales nonlinearly with stimulation intensity and was better described by a four-parameter symmetrical logistic function than by a linear fit (extra sum-of-squares F test,  $F_{(1,6)} = 9.28$ ,  $p < 0.05$  for fit comparison;  $R^2 = 0.98$  for nonlinear fit). Vertical line denotes the EA<sub>50</sub> value (5.42 mA, [95% CI: 4.51, 10.11]). **(H)** Area under the curve (AUC) of the decay phase, measured in a sliding window starting one frame after the peak of each trace, also scales nonlinearly as a function of input intensity (extra sum-of-squares F test,  $F_{(1,6)} = 12.05$ ,  $p < 0.05$  vs linear fit,  $R^2 = 0.96$  for nonlinear fit). Vertical line denotes the EA<sub>50</sub> value (4.91 mA [95% CI: 4.14, 7.13]). **(I)** Across stimulations of increasing train length, there was a linear increase in the amplitude of the bulk fluorescence response (extra sum-of-squares F test,  $F_{(1,2)} = 17.98$ ,  $p > 0.05$  vs linear fit,  $R^2 = 0.92$  for linear fit). **(J)** Similarly, the AUC of the decay phase scaled linearly with train length (extra sum-of-squares F test,  $F_{(1,2)} = 1.92$ ,  $p > 0.05$  vs linear,  $R^2 = 0.99$  for linear fit). **(K)** Heatmaps of single-cell dynamics for single pulse stimulations (2 and 5 mA) with stimulation onset indicated by the dotted black line. **(L)** Activity over time traces for population averages among cells classified as excited, inhibited, or not responsive to 2 mA and 5 mA single pulse stimulations. **(M)** Distributions of response types (Excited, Inhibited, No response) as a percent of the population differs across stimulation intensities (chi-squared test,  $\chi^2 = 197.8$ ,  $p < 0.0001$ ). **(N)** The percentage of cells recruited by the stimulation scales linearly with increasing stimulation intensity (extra sum-of-squares F test,  $F_{(2,68)} = 0.62$ ,  $p > 0.05$  vs linear fit,  $R^2 = 0.24$  for linear fit;  $F_{(1,70)} = 22.54$ ,  $p < 0.0001$  vs zero). **(O)** Representative heatmaps of single-cell dynamics evoked by multiple pulses (5 pulses and 20 pulses at 5 mA). **(P)** Activity over time traces for population averages among cells classified as excited, inhibited, or no response to 5 pulse and 20 pulse stimulations. **(Q)** Proportion of the population recruited increases across stimulations with increasing train length (chi-squared test,  $\chi^2 = 179.0$ ,  $p < 0.0001$ ). **(R)** Increasing train length leads to nonlinear increases in ensemble size (extra sum-of-squares F test,  $F_{(1,22)} = 8.22$ ,  $p < 0.05$  vs linear fit,  $R^2 = 0.68$  for nonlinear fit). Error bars indicate SEM. Continuous error bands indicate 95% confidence intervals of best-fit values derived through curve fitting. **See Figure S1 and Video S1 for additional data.**



**Figure 2. dIPFC microcircuit ensembles are stochastically assembled around a minority of stable core members.**

(A) Diagram of possible ensemble dynamics underlying the population level signals reported in Figure 1.  $S^x$  refers to the population observed throughout the recordings regardless of activity, and  $S^x$  refers to the ensemble activated by a stimulus  $S$  at a given iteration  $x$ . In this example, there is a gradual increase in the proportions of the population participating in the ensemble over iterations, but each iteration is represented by mostly unique members. (B) Sankey diagram showing retention/crossover of cells categorized as responsive vs not response to stimulation for the lowest intensity/first stimulation to the half-maximal/fifth stimulation ( $S^A \rightarrow S^D$ ) and from the half-maximal intensity stimulation to the final stimulation ( $S^D \rightarrow S^I$ ). (C) Sankey diagram showing retention/crossover of cells categorized as responsive vs no response to stimulation for the single pulse stimulation to the ten pulse/third stimulation in the train length curve dataset and from the ten-pulse stimulation to the twenty-pulse stimulation. (D) Sankey diagram of the same ensemble iterations as depicted in panel B, from the stimulation intensity curve experiment, but only including cells which were recruited to  $S^A$  and with response directionality denoted. (E) Sankey diagram of the same ensemble iterations as depicted in panel C, from the train length curve experiment, but only including cells which were recruited by the first stimulation in the experiment and with response directionality denoted. (F) Diagram depicting potential

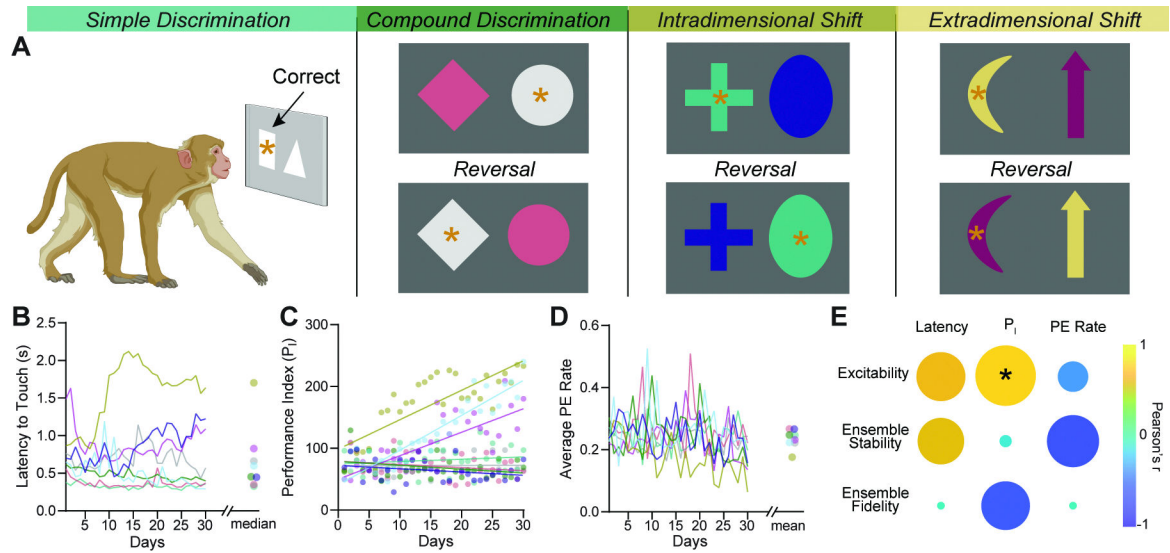
ensemble dynamics that could underlie the high degree of turnover observed in both datasets (panels B-E). In the example, one set of cells switches in and out of the ensemble, resulting in a low similarity index between some of the iterations while others are identical (Jaccard's index of 1). **(G)** Heatmaps of Jaccard's similarity index for the 36 comparisons resulting from pairwise analysis of the ensemble memberships across the 9 iterations/stimulations intensities for the intensity curve dataset. An index of 1 indicates that members are identical, while an index of 0 indicates that members are non-overlapping. **(H)** Distribution of recruitment probability over the 9 iterations of ensemble *S* (number of times recruited to the ensemble / number of iterations). Shuffled distribution shown is the first of 10,000 reshuffles, for illustrative purposes (**see Figure S2 for reshuffle statistics and distributions**). Inset: cells were classified based on whether they were recruited to all of the ensemble iterations ('Always'), at least one but not all ('Sometimes'), or were not recruited to any of the ensemble iterations ('Never'). **(I)** Heatmaps of Jaccard's similarity index for the 10 comparisons resulting from pairwise analysis of the ensemble memberships across the 5 iterations/stimulations from the variable train length dataset. **(J)** Distribution of recruitment probability over the 5 iterations of the ensemble evoked by stimulations of increasing train length (number of times recruited to the ensemble / number of iterations). The shuffle distribution shown is the first of 10,000 reshuffles, for illustrative purposes. Inset: cells were classified as 'Always', 'Sometimes', or 'Never' based on whether they were recruited to all of the ensemble iterations, at least one but not all, or were not recruited to any of the ensemble iterations. **(K-O)** To determine if weighting of recruitment probability differed for recruitment via excitation vs inhibition, datasets were reanalyzed twice using modified definitions of ensemble membership; once in which only stimulation-evoked increases in activity were considered for membership (excited subassembly) and once in which only stimulation-evoked decreases in activity were considered for membership (inhibited subassembly). **(K)** Diagram illustrating how differential weighting of recruitment probability across excited and inhibited cells can affect subassembly membership. The memberships depicted represent the result of applying the subassembly definitions to the hypothetical scenario depicted in panel F. **(L)** Recruitment probability over the 9 iterations in the stimulation intensity experiment when only excitatory responses are considered. Inset: associated classifications into Always excited, Sometimes excited, and Never excited. **(M)** Recruitment probability for the stimulation intensity experiment when only inhibitory responses are considered. Inset: associated classifications into Always inhibited, Sometimes inhibited, and Never inhibited. **(N)** Recruitment probability for the train length experiment when only the excitatory subassembly of the ensemble is considered. Inset: associated classifications. **(O)** Recruitment probability for the train length experiment when only the inhibitory subassembly of the ensemble is considered. Inset: associated classifications. \*\*\*median Kolmogorov-Smirnov D for observed vs shuffled from 10,000 reshuffles > critical value when  $\alpha=0.001$ ; †††median  $X^2$  value for observed vs shuffled from 10,000 reshuffles > critical value when  $\alpha=0.001$ . **See Figure S2 for additional data.**



**Figure 3. Ensembles display high-fidelity population-level dynamics despite membership instability.**

(A-C) To determine the degree to which instability in ensemble recruitment reflects inherent stochasticity in microcircuit dynamics versus differential stimulation parameters, we repeated identical perisulcal stimulations in triplicate (single pulse, 5 mA) and compared these to our previous results of ascending stimulation intensity (4,5,6 mA). Sankey diagram showing retention/crossover of cells categorized as responsive vs not responsive for the first to second identical stimulation ( $S^A \rightarrow S^B$ ) and from the second to the third identical stimulation ( $S^B \rightarrow S^C$ ). (B) Sankey diagram of the same ensemble iterations as depicted in panel A, from the repeated stimulation experiment, but only including cells recruited to  $S^A$  and with response directionality denoted. (C) To compare the repeated vs ascending stimulation conditions, we first examined the percent of the population not yet recruited to at least one ensemble (excited/inhibited) across the three stimulation iterations. The proportion of cells recruited did not differ compared to three stimulations from the ascending stimulation intensity experiment (two-way ANOVA [condition x stimulation replicate], condition:  $F_{(1,9)} = 2.11$ ,  $p = 0.26$ ; stimulation replicate:  $F_{(1,181), 10.63} = 71.6$ ,  $p < 0.001$ ; Šidák’s test, 4 mA vs 5 mA:  $t_{(2,23)} = 0.14$ ,  $p = 0.99$ ; 5 mA vs 5 mA:  $t_{(2,46)} = 1.36$ ,  $p = 0.64$ ; 6 mA vs 5 mA:  $t_{(2,30)} = 1.13$ ,  $p = 0.74$ ). (D) We next determined the

average proportion of each ensemble that had been recruited to the iteration immediately prior (retained members) and the proportion that had not yet been recruited to the iteration one prior (new members). Comparing the repeated versus ascending stimulation conditions, we found no differences in the average proportion of new ( $t_{(9)} = 1.33$ ,  $p = 0.22$ ) or retained members ( $t_{(9)} = 0.61$ ,  $p = 0.56$ ). **(E-J)** To examine the low dimensional dynamics across varying stimulation parameters, we plotted neural trajectories from the first three principal components of each stimulation, normalized by z-score to remove the magnitude changes evoked by varying the stimulation parameters, allowing direct comparison of the shape of the trajectories. We found that conserved dynamics are clearly observable across different stimulations through the qualitatively similar paths taken in PCA space. Panel E shows all cells together across the increasing intensity, single pulse curve. Pairwise comparisons between the nine trajectories yielded  $r$  values ranging from 0.73 to 0.99. **(F)** To further characterize this point, we also examined the dynamics observed in cells which do not change designations across stimulations, i.e. Always and Never cells, compared to **(G)** those recruited at least once (i.e. Sometimes), finding that the low dimensional dynamics are conserved among each of these populations and do not differ appreciably from the entire population (panel E). **(H)** The same PCA analysis was repeated for all cells observed in the variable train length experiment. Pairwise comparisons between the five trajectories yielded  $r$  values ranging from 0.56 to 0.95. We also compared between the **(I)** Always and Never cells and the **(J)** Sometimes cells, showing the same pattern of conserved low dimensional dynamics regardless of the membership stability. Error bars indicate SEM. **See Figure S3 for additional data.**



**Figure 4. Cognitive flexibility is associated with intrinsic excitability of dIPFC microcircuits.**

(A) Diagram of set-shifting behavioral paradigm. Monkeys were trained to select one of two geometrical shapes, which would result in a food pellet reward, and as the animals reached the performance criterion they would advance to the next, more challenging phase (e.g., Simple Discrimination → Complex Discrimination). Each session was 45 minutes, and began with Simple Discrimination, regardless of the previous session's performance. Three behavioral metrics of interest were selected *a priori* for analyses based on prior characterization of the task: median latency to touch the screen, slope of the performance index ( $P_1$ ), and mean perseverative error rate. (B) Latency to touch the screen from trial onset for each of the 8 subjects over 30 sessions. While there is session to session variance a clear pattern of individual differences can be seen, which are summarized by the median latency over sessions. (C) Overall performance, as measured by the performance index ( $P_1$ ) over sessions. There are wide individual differences in rate of improved performance over sessions, which can be summarized by the slope of the linear regression. (D) Average perseverative error rate fluctuates over sessions. A perseverative error refers to continued utilization of the previous strategy in a new phase of the task. Perseverative error rate (defined as perseverative errors as a proportion of the total number of errors) from each session was averaged together per subject and this value was used for correlation analysis. (E) The three behavioral summary metrics were correlated with three summary metrics of microcircuit activity: overall excitability (see figure S4A), ensemble stability (see Figure S4E), and encoding fidelity (see Figure S4I). Greater overall excitability was positively correlated with slope of the performance index (Pearson's  $r_{(8)} = 0.78$ ,  $p < 0.05$ ). See Figure S4 for additional data.

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Chemicals, Peptides, and Recombinant Proteins</i>		
fluo-4 acetoxymethyl ester	Thermo Fisher Scientific	Cat#: F14201; CAS: 273221-67-3
Dimethyl sulfoxide (DMSO)	Tocris	Cat#: 3176; CAS: 67-68-5
DAPI	Thermo Fisher Scientific	Cat#: 62248
<i>Experimental Models: Organisms/Strains</i>		
Rhesus Macaque	Oregon National Primate Research Center	N/A
<i>Software and Algorithms</i>		
GraphPad Prism	GraphPad Software	<a href="https://www.graphpad.com/">https://www.graphpad.com/</a>
MATLAB	Mathworks	<a href="https://www.mathworks.com/">https://www.mathworks.com/</a>
ThorImage	Thorlabs Imaging Systems	<a href="https://www.thorlabs.com/">https://www.thorlabs.com/</a>
Patchwarp	Hattori et al. <sup>61</sup>	<a href="https://github.com/ryhattori/PatchWarp/">https://github.com/ryhattori/PatchWarp/</a>
Inscopix Data Processing Software	Inscopix	<a href="https://inscopix.com/">https://inscopix.com/</a>
CaImAn	Giovannucci et al. <sup>62</sup>	<a href="https://github.com/flatironinstitute/CaImAn">https://github.com/flatironinstitute/CaImAn</a>