UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Modeling the prefrontal cortex as a lifelong learning system

Permalink

https://escholarship.org/uc/item/77b0056h

Author

Batra, Kanha

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Modeling the prefrontal cortex as a lifelong learning system

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Electrical Engineering (Machine Learning and Data Science)

by

Kanha Batra

Committee in charge:

Professor Kay M. Tye, Chair
Professor Vikash Gilja, Co-chair
Professor Terrence J. Sejnowski, Co-chair
Professor Kanaka Rajan
Professor Edward J. Wang

2024

Copyright

Kanha Batra, 2024

All rights reserved.

The Dissertation of Kanha Batra is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

To my family, friends and all those who endeavor to change the world.

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE iii

DEDICATION iv

TABLE OF CONTENTS v

LIST OF FIGURES viii

ACKNOWLEDGEMENTS x

VITA xiv

ABSTRACT OF THE DISSERTATION xvi

Chapter 1 Introduction 1

1.1 The pursuit of artificial intelligence 1

1.2 The need for lifelong learning 4

1.3 The brain with all the answers 7

1.4 Goals of this dissertation 11

1.5 Bibliography 12

Chapter 2 Cortical ensembles orchestrate social competition through hypothalamic outputs 19

2.1 Abstract 19

2.2 Introduction 20

2.3 mPFC neurons encode competition behavior 24

2.4 mPFC reflects rank and winning 25

2.5 mPFC predicts future wins 28

2.6 Rank-dependent mPFC representations 29

2.7 mPFC-LH neurons modulate social dominance 33

2.8 Conclusion 36

2.9 Appendix 38

2.10 Supplemental figures 73

2.11 Bibliography 94

2.12 Acknowledgements 98

Chapter 3 Visualizing the Longitudinal Development of Stress-Induced Anhedonia From Representations of Valence in the PFC 100

3.1 Abstract 100

3.2 Introduction 101

3.3 Chronic mild stress induces anhedonia and is predictive during Pavlovian Conditioning task 103

3.4 mPFC population activity predicts anhedonia phenotypes prior to stress 107

3.5 mPFC population dynamics exhibit dramatic changes during both early and late stress in hedonic groups 111

3.6 Appendix 115

3.7 Supplemental figures 125

3.8 Bibliography 129

3.9 Acknowledgements 131

Chapter 4 Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning 132

4.1 Introduction 132

4.2 Model architecture 134

4.3 Behavioral task and simulation design 138

4.4 Foraging reveals context specific behavioral strategies 140

4.5 DA in mPFC selectively encodes contextual information 143

4.6 mPFC ensembles follow expert-like behavior 147

4.7 Simulations with NeMoE generate diverse strategies resembling behavior 151

4.8 Discussion and future work 153

4.9 Bibliography 155

4.10 Acknowledgements 156

Chapter 5 Conclusion..... 158

LIST OF FIGURES

Figure 1-1 Schematic describing PFC function adapted from Miller and Cohen 2001. ... 9
Figure 1-2 Dopamine function in mediating signal-to-noise ratio of mPFC neurons adapted from Vander Weele et al. 2018). ... 10
Figure 2-1 Novel social dominance assay and deep learning tool for tracking multiple animals. ... 22
Figure 2-2 Social rank and competitive success is decoded by mPFC population dynamics. ... 26
Figure 2-3 Relative dominants have more reward-seeking-behavior cells whereas subordinates have larger responses to competitor behavior. ... 31
Figure 2-4 mPFC–LH pathway encodes social competition and modulates social dominance behavior. a, Left, viral strategy to stimulate projectors. ... 34
Figure 3-1 Chronic mild stress induces anhedonia and is predictive during Pavlovian Conditioning task. ... 105
Figure 3-2 mPFC population activity predicts anhedonia phenotypes prior to stress. . 108
Figure 3-3 mPFC population dynamics exhibit dramatic changes during both early and late stress in hedonic groups. ... 113
Figure 4-1 Components of the architecture and their biological counterparts. ... 136
Figure 4-2 Schematics of the behavioral task along with the simulation strategy. ... 139
Figure 4-3 Experimental validation of the Seasonal Foraging Task. ... 141
Figure 4-4 Context representation in dopamine response. ... 145
Figure 4-5 Context representation in mPFC neural response. ... 149
Figure 4-6 Deep reinforcement learning simulations on the Go/No-Go SFT. ... 152

Supplemental Figure 2-1 Venn Diagram indicating overlap of mice recorded across experiments. ... 44
Supplemental Figure 2-2 Flow chart describing the interaction between variables in HMM-GLM. ... 62
Supplemental Figure 2-3 Additional behavioral metrics during reward competition in unimplanted mice. ... 74
Supplemental Figure 2-4 AlphaTracker tracking metrics. ... 76
Supplemental Figure 2-5 AlphaTracker unsupervised clustering results. ... 77
Supplemental Figure 2-6 Histological validation of electrode placements. ... 79
Supplemental Figure 2-7 Behavior in competition with logger and HMM-GLM model controls. ... 80

Supplemental Figure 2-8 Additional data for mPFC population dynamics during social competition. 82

Supplemental Figure 2-9 mPFC population dynamics during social competition are not driven by location or mouse identity. 84

Supplemental Figure 2-10 Decoding performance for relative and absolute social rank, and competitive success with different datasets. 87

Supplemental Figure 2-11 Additional data for mPFC single unit responses to task-relevant events during social competition. 89

Supplemental Figure 2-12 mPFC-LH photostimulation does not affect other motivated behaviors. 92

Supplemental Figure 3-1 Ketamine treatment after chronic mild stress decreases variance in social index in anhedonic mice. 126

Supplemental Figure 3-2 Anhedonia groups decoded via mPFC population activity. . 127

Supplemental Figure 3-3 mPFC neuronal population dynamics exhibit dramatic changes during select weeks. 128

ACKNOWLEDGEMENTS

This dissertation would not have happened without the collective effort of several individuals whom I would like to thank.

I would like to start with my parents. Neither of them is a STEM major and yet they encouraged me to keep pursuing my dreams. I thank them for their support throughout the past ten years of school that I have been through from bachelor's to master's to finally my PhD. I also thank them for listening to me explain my research every single time though I know they have stopped following what I study since high school. I learned the value of hard work from them and can never thank them enough.

I would like to thank my elder brother, Tanmay. We rarely see eye to eye, and I have spent a good part of the last decade trying to compete with him on academic merit, but I learned my most important lesson with him: you need to find your own passions to follow for your achievements to feel satisfying.

I would like to thank my partner, Akanksha. We met shortly before the preliminary exam that admitted me to the PhD program at UC San Diego. She has seen me through every single lab meeting presentation, conference, and examination throughout my PhD and has funneled an infinite amount of support for me through it all. I have also seen her grow through her career and become someone I admire more each day. I learned the value of sheer perseverance in face of anxiety from her.

I would like to thank all the friends I have made through life. They say friends are the family that you choose, and I have been really fortunate in that. They have not only made going through the hard moments in academia easier, but even enjoyable. It's a

comradery I cherish deeply and hope to preserve for the rest of my life. I learned the value of finding amusement in life even in the hardest of times from them.

I would like to thank my advisor, Kay Tye, for taking me on as a student despite my scarce prior knowledge of neuroscience. She has shown me the path to transform my vision into a hypothesis to an experimental approach and finally to a thesis. In all my meetings with her, not once have I left without a whole host of new ideas brimming with potential. She has shown me the ropes of academia, the importance of transparency and communication in science. I learned the value of believing in your ideas (no matter how ambitious they sound) from her.

I would like to thank my co-advisor, Terry Sejnowski, for providing his sage wisdom and ability to focus on the bigger picture like an optimizer that finds its way out of a local minimum. I learned from him the value of inculcating knowledge, deep and wide, and to not take everything so seriously.

I would like to thank Nancy, my postdoc mentor in the lab before she moved on to start her own lab like the baller that she is. She taught me my first action potential, that was really the window to the wonderful and complex world of neuroscience. She was unbelievably patient with all my mathematical rambling and always managed to shed light on the focal points of the research. She gave me advice, both personal and professional, with so much sincerity it might have been her own life she's leading. I learned from her the value of how every single scientific endeavor, no matter how big or small, whether it's annotating images or crafting complex algorithms, brings you one step closer to your goals.

I would like to thank members of the Tye lab. I have learned something from every single person that I have interacted with, but the most important was the impact of having a collaborative and supportive scientific community. I have made a number of meaningful connections that have been instrumental in my social and academic development. I have also had the pleasure of working with a team of wonderful undergraduates, research technicians and fellow graduate students, without whom my thesis project would have been merely a shell of the dream that it became.

I would like to thank my committee members, Vikash Gilja, Kanaka Rajan and Edward Wang, for their invaluable advice and feedback that they provided time and again throughout my PhD.

I would also like to thank the broader scientific communities at the Salk Institute for Biological Studies and the University of California San Diego. They provided a collective platform for me to thrive as a graduate student and scientist.

Chapter 2, in full, is a reprint of the material as it appears in: Nancy Padilla-Coreano*, Kanha Batra*, Makenzie Patarino, Zexin Chen, Rachel R. Rock, Ruihan Zhang, Sébastien B. Hausmann, Javier C. Weddington, Reesha Patel, Yu E. Zhang, Hao-Shu Fang, Srishti Mishra, Deryn O. LeDuke, Jasmin Revanna, Hao Li, Matilde Borio, Rachelle Pamintuan, Aneesh Bal, Laurel R. Keyes, Avraham Libster, Romy Wichmann, Fergil Mills, Felix H. Taschbach, Gillian A. Matthews, James P. Curley, Ila R. Fiete, Cewu Lu, and Kay M. Tye. Cortical ensembles orchestrate social competition through hypothalamic outputs. *Nature*, 603 (7902), 667-671, 2022. The dissertation author was a co-primary investigator (*) and author of this paper.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Austin A. Coley, Jeremy M. Delahanty, Kanha Batra, Assaf Ramot, Rachelle Pamintuan, Laurel Keyes, Jim Hagemann, Christopher R. Lee, Vivian Liu, Harini Advikoluna, Jianna Cressy, Caroline Jia, Francesca Massa, Deryn O. LeDuke, Moumen Gabir, Bra'a Durubeh, Lexe Linderhof, Reesha Patel, Analetta Ehler, Romy Wichmann, Hao Li, Kyle B. Fischer, Talmo Pereria, Takaki Komiyama, Kay M. Tye. *In preparation*, 2024. The dissertation author was a co-author of this paper.

Chapter 4, in full, is currently being prepared for submission for publication of the material. Kanha Batra, Anousheh Bakhti-Suroosh, Clara N. Yi, May G. Chan, Ruby Tseng, Kelbi Banducci, Jiaqi Zhang, Sophie E. Poelmann, Christopher R. Lee, Tanisha Roy, Kelly Chang, Tristan B. Tuazon, Romy Wichmann, Caroline Jia, Ben Tsuda, Bryan Nielsen, Laurel R. Keyes, Yulong Li, Talmo Pereira, Terrence J. Sejnowski, Kay M. Tye. Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning. *In preparation*, 2024. The dissertation author was the primary investigator and author of this paper.

VITA

2018 Bachelor of Technology in Chemical Engineering, Indian Institute of Technology Delhi

2020 Master of Science in Electrical Engineering (Machine Learning and Data Science), University of California San Diego

2024 Doctor of Philosophy in Electrical Engineering (Machine Learning and Data Science), University of California San Diego

PUBLICATIONS

Kanha Batra, Anousheh Bakhti-Suroosh, Clara N. Yi, May G. Chan, Ruby Tseng, Kelbi Banducci, Jiaqi Zhang, Sophie E. Poelmann, Christopher R. Lee, Tanisha Roy, Kelly Chang, Tristan B. Tuazon, Romy Wichmann, Caroline Jia, Ben Tsuda, Bryan Nielsen, Laurel R. Keyes, Yulong Li, Talmo Pereira, Terrence J. Sejnowski, Kay M. Tye. Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning. *In preparation*, 2024.

Reesha R Patel, Makenzie Patarino, Kelly Kim, Rachelle Pamintuan, Felix H Taschbach, Hao Li, Christopher R Lee, Aniek van Hoek, Rogelio Castro, Christian Cazares, Raymundo L Miranda, Caroline Jia, Jeremy Delahanty, Kanha Batra, Laurel R Keyes, Avraham Libster, Romy Wichmann, Talmo D Pereira, Marcus K Benna, Kay M Tye. Social isolation recruits amygdala-cortical circuitry to escalate alcohol drinking. *In submission*, 2024.

Fergil Mills,* Christopher R Lee*, James R Howe, Hao Li, Shan Shao, Maria N Keisler, Mackenzie E Lemieux, Felix H Taschbach, Laurel R Keyes, Matilde R Borio, Hannah S Chen, Reesha R Patel, Alexa L Gross, Jeremy Delahanty, Christian Cazares, Liezl Maree, Romy Wichmann, Talmo D Pereira, Marcus K Benna, Cory M Root, Kay M Tye. Amygdalostriatal transition zone neurons encode sustained valence to direct conditioned behaviors. *In submission*, 2024.

Chris A. Leppla, Laurel R. Keyes, Gordon Glober, Gillian A. Matthews, Kanha Batra, Maya Jay, Yu Feng, Hannah S. Chen, Fergil Mills, Jeremy Delahanty, Jacob M. Olsen, Edward H. Nieh, Praneeth Namburi, Craig Wildes, Romy Wichmann, Anna Beyelar, Eyal Y. Kimchi, Kay M. Tye. Thalamus sends information about arousal but not valence to the amygdala. Psychopharmacology, 2023.

Hao Li*, Praneeth Namburi*, Jacob M. Olson*, Matilde Borio, Mackenzie E. Lemieux, Anna Beyeler, Gwendolyn G. Calhoon, Natsuko Hitora-Imamura, Austin A. Coley, Avraham Libster, Aneesh Bal, Xin Jin, Huan Wang, Caroline Jia, Sourav R. Choudhury, Xi Shi, Ada C. Felix-Ortiz, Verónica de la Fuente, Vanessa P. Barth, Hunter O. King, Ehsan M. Izadmehr, Jasmin S. Revanna, Kanha Batra, Kyle B. Fischer, Laurel R. Keyes, Nancy Padilla-Coreano, Cody A. Siciliano, Kenneth M. McCullogh, Romy Wichmann, Kerry J. Ressler, Ila R. Fiete, Feng Zhang, Yulong Li, Kay M. Tye. Neurotensin orchestrates valence assignment in the amygdala. Nature, 2022.

Nancy Padilla-Coreano*, Kanha Batra*, Makenzie Patarino, Zexin Chen, Rachel R. Rock, Ruihan Zhang, Javier Weddington,
Sebastien Hausmann , Yu Eva Zhang, Reesha Patel, Laurel R. Keyes, Gillian Matthews, James P. Curley, Ila R. Fiete, Cewu Lu,
Kay M. Tye. Cortical ensembles orchestrate social competition through hypothalamic outputs. Nature, 2022.

Leonardino A. Digma, Christine H. Feng, Christopher C. Conlin, Ana E. Rodríguez-Soto, Allison Y. Zhong, Troy S. Hussain, Asona J. Lui, Kanha Batra, Aaron B. Simon, Roshan Karunamuni, Joshua Kuperman, Rebecca Rakow-Penner, Michael E. Hahn,
Anders M. Dale, Tyler M. Seibert. Correcting B0 inhomogeneity-induced distortions in whole-body diffusion MRI of bone metastases. Scientific Reports, 2022.

Christine H Feng, Christopher Charles Conlin, Kanha Batra, Ana Rodriguez-Soto, Roshan Karunamuni, Aaron Simon, Joshua Kuperman, Rebecca Rakow-Penner, Michael E Hahn, Anders M Dale, Tyler M Seibert. Voxel-level Classification of Prostate Cancer on Magnetic Resonance Imaging: Improving Accuracy Using Four-Compartment Restriction Spectrum Imaging. Journal of Magnetic Resonance Imaging, 2021.

*Authors have contributed equally

ABSTRACT OF THE DISSERTATION

Modeling the prefrontal cortex as a lifelong learning system

by

Kanha Batra

Doctor of Philosophy in Electrical Engineering (Machine Learning and Data Science)

University of California San Diego, 2024

Professor Kay M. Tye, Chair
Professor Vikash Gilja, Co-chair
Professor Terrence J. Sejnowski, Co-chair

The quest to understand and reverse engineer natural intelligence has captivated humanity for centuries. While we have made breakthroughs in neuroscience and machine learning individually, the intersection of the two remains largely unexplored and this dissertation aims to tap into its unknown potential. Using a dual modeling-experimental

methodology, I study the cognitive capabilities of the prefrontal cortex, a brain-region essential for higher-order cognitive functions.

I first define the postulates of lifelong learning, the idea of a continually learning algorithm capable of mastering innumerable tasks sequentially with efficient transfer of knowledge (Chapter 1). I discuss the state-of-the-art solutions that have been proposed in engineering and their shortcomings. I subsequently describe the systematic approach that I undertake over the course of my PhD.

I first look at the role of the prefrontal cortex in the modulation of behavior across social contexts. I use state space analysis to identify distinct neural representations within the PFC linked to social rank and competition outcomes. I then devise a novel adaptation of the Hidden Markov Model to identify latent states associated with flexible social behavior, revealing potential mechanisms for context-specific neural activation within the PFC.

Function and disease are two sides of the same coin. In order to broaden my understanding of prefrontal mechanisms, I next pursue a collaborative study focused on the detrimental impact of chronic stress on PFC function and behavior (Chapter 3). I identify a method to categorize animals based on stress-induced anhedonia. I use state space analysis to identify distinctions in the anhedonic state from the hedonic in PFC activity and study its evolution over a chronic stress paradigm.

Given the findings from previous chapters, I develop a hierarchical modular neural network architecture based on the PFC, integrating internal state dynamics and a novel neuromodulatory filtering mechanism (Chapter 4). This model successfully reproduces animal behavior and PFC function while being simulated on a contextual foraging task.

This underscores its validity and potential for understanding the neural underpinnings of lifelong learning as well as related neuropsychiatric conditions.

Through this dissertation, I aim to advance our understanding of the neural basis of lifelong learning. The novel computational models and experimental findings presented here have far-reaching implications for the development of artificial intelligence systems with enhanced cognitive capabilities, as well as for potential therapeutic interventions targeting PFC dysfunction in various neuropsychiatric conditions.

Chapter 1 Introduction

In this chapter, I describe the central question behind my dissertation. I lay out the historical developments that inspired me to pursue this question as a part of my doctoral research. I discuss the various existing approaches geared towards solving the problem along with their limitations. Lastly, I summarize the various research projects that I pursued over the course of my thesis which is elaborated upon in the following chapters.

1.1 The pursuit of artificial intelligence

The origins of Artificial Intelligence (AI) lie in a rich history spanning centuries of philosophical, mathematical, physical, and literary inquiry. The Turing era introduced computer systems based on symbolic, rule-based computations, capable of pre-programmed tasks. While excelling in well-defined domains, these systems struggled with the complexity and ambiguity of real-world problems. Developments in statistics and modeling gave rise to the sub-field of machine learning (Samuel 1959). Machine learning shifted the focus towards systems that could learn patterns directly from data, enabling significant progress in areas like perception and pattern recognition.

Over the decades, the field transformed due to multi-disciplinary advancements. One was the development of complex computational architectures such as the perceptron (Rosenblatt 1958), support vector machine (SVM) (Boser, Guyon, and Vapnik 1992), artificial neural networks (ANNs) (Lecun et al. 1998) and ensemble methods (Dietterich 2000). These architectures are capable of capturing both linear and non-linear relationships between distributions of data (LeCun, Bengio, and Hinton 2015), modeling

sequential data (Bai, Kolter, and Koltun 2018), and learning new tasks with minimal data using few-shot and meta-learning (Finn, Abbeel, and Levine 2017). A second advancement was in computation: the massive parallelization of training using graphic processing units (GPUs) (Pandey et al. 2022) as well as the development of distributed optimization algorithms (Dean et al. 2012; Mnih et al. n.d.; Niu et al. 2011). Lastly, the explosion of open-source datasets across academia and industry has fueled the ability of models to generalize learning to near-human levels, especially in vision and speech (Banko and Brill 2001; Deng et al. 2009; Halevy, Norvig, and Pereira 2009; Krizhevsky, Sutskever, and Hinton 2012). These are but a few significant milestones in the ongoing evolution of AI.

Modern AI has rapidly advanced, tackling some of humanity's most complex challenges, especially in the realms of biology. A delayed cancer diagnosis can cause a significant increase in the mortality rate of patients(Hanna et al. 2020). A major cause of this is a shortage of medical personnel as well as resources. Since diagnostics are inherently data-driven, they present a compelling opportunity for machine learning. Deep convolutional neural networks now detect malignant cancer anomalies from medical images with specificity and sensitivity rivaling, and sometimes exceeding, those of expert physicians (Esteva et al. 2017; McKinney et al. 2020). Another critical problem, the inability to computationally predict protein structure, has hindered our understanding of genetic diseases and the design of targeted drugs and enzymes (Dill and MacCallum 2012). Recent breakthroughs in deep reinforcement learning have not only enabled highly accurate predictions of known protein structures but also offer a path toward designing entirely novel proteins (Jumper et al. 2021). A third arena where AI is transforming

healthcare is surgery. Global surgeon shortages lead to limited access to care and contribute to surgeon burnout (Elmore et al. 2016; Holmer et al. 2015). Advancements in robotics and machine learning are beginning to address these challenges, with studies demonstrating successful cases of both autonomous and semi-assisted surgeries (Saeidi et al. 2022).

While the applications to biology and healthcare are highly relevant to this dissertation, AI's potential naturally extends far beyond medicine. From self-driving cars to surprisingly creative text generation, AI is reshaping industries. A key driver behind these advancements is the development of the transformer, a novel neural network architecture that has revolutionized the field of AI (Vaswani et al. 2017). Transformers excel due to their self-attention mechanism, which allows the model to dynamically focus on the most relevant parts of the input sequence, whether it's words in a sentence or image patches in a picture. Unlike traditional architectures, transformers capture long-range dependencies within data efficiently, leading to a deeper understanding of complex relationships. Additionally, transformers are highly parallelizable, meaning they can be efficiently trained on massive datasets using multiple processing units. Consequently, transformers have become a cornerstone of various machine learning domains.

In natural language processing, transformers have powered the creation of Large Language Models (LLMs) that have demonstrated impressive abilities in tasks like writing different kinds of creative content, translating languages, and composing realistic dialogue (Brown et al. 2020; Devlin et al. 2019). Similarly, transformers have made significant inroads into computer vision. Large Vision Models (LViMs), based on transformer architectures, are achieving excellent results in tasks traditionally

dominated by convolutional neural networks (CNNs). For example, the Vision Transformer (ViT) utilizes transformers to classify images with high accuracy, demonstrating the model's ability to learn complex visual patterns (Dosovitskiy et al. 2021). Similarly, Detection Transformer (DETR) employs transformers for object detection, pinpointing and identifying objects within an image (Carion et al. 2020) Understanding the strengths of the transformer architecture illuminates the capabilities essential for the next generation of AI. It’s also important to mention that despite these successes, AI systems still face limitations, can inherit the biases present in their training data, and raise complex ethical questions about fairness and decision-making transparency. The journey of AI is far from over, and navigating both its potential and potential pitfalls will be crucial in the years to come.

1.2 The need for lifelong learning

While modern AI has achieved remarkable feats in solving complex tasks, a fundamental limitation persists in its ability to continuously learn, adapt, and evolve as seamlessly as biological systems. This enduring gap underscores the crucial need for lifelong learning, a key step to bridging the divide between artificial and natural intelligence (Kudithipudi et al. 2022).

Lifelong learning in AI aims to enable machine learning models to continuously learn, adapt, and build upon their existing knowledge over time, in contrast to traditional models that often struggle to retain information when presented with new tasks or changing environments. Inspired by biological learning systems, key elements of lifelong learning include overcoming catastrophic forgetting by developing strategies to prevent new learning from erasing previously acquired knowledge. Additionally, lifelong learning

promotes adaptability, allowing models to seamlessly adjust their behavior and knowledge to handle unexpected scenarios, new tasks, or shifts in data patterns. It also encourages knowledge transfer, facilitating the use of existing knowledge to learn new tasks faster and more efficiently, echoing how humans often build on existing skills when acquiring new ones. Lifelong learning also strives for models that can learn effectively from smaller amounts of data with less computational overhead. Finally, it involves enabling AI systems to “learn how to learn” through meta-learning, allowing them to improve their own learning strategies and become more adaptable throughout their lifespan.

There have been several attempts over the years to tackle the problem of lifelong learning. One influential solution is Elastic Weight Consolidation (EWC), a regularization-based method that protects parameters in a neural network deemed important for previously learned tasks (Kirkpatrick et al. 2017). Inspired by the concept of synaptic memory consolidation in the brain, EWC aims to mitigate catastrophic forgetting (Benna and Fusi 2016; Squire et al. 2015), setting a benchmark for evaluating newer methods. While EWC helps a network use its existing capacity more efficiently, it has limitations. Its focus on preserving knowledge within a fixed-size network means it can reach a saturation point while learning multiple tasks. Additionally, its reliance on a fixed metric to determine parameter importance can lead to the overwriting of old learning or hinder the network's ability to learn significantly different tasks.

Gradient Episodic Memory (GEM) offers another approach to lifelong learning that mitigates catastrophic forgetting (Lopez-Paz and Ranzato 2022). It does this by storing a small set of representative examples from each task that the model has encountered.

These examples guide the learning process on new tasks, helping the network retain old knowledge while minimizing disruption. However, this technique has limitations. GEM's memory footprint grows as more tasks are learned, and it introduces computational overhead by requiring gradient calculations with respect to both the current task and the stored past examples.

A third powerful example is that of Incremental Classifier and Representation Learning (iCaRL) (Rebuffi et al. 2017). This strategy focuses on compressing knowledge gained from previous tasks into a smaller representation through knowledge distillation, reducing the tendency for catastrophic forgetting. iCaRL also maintains a carefully selected set of examples, aiding in both the knowledge distillation process and classifying new data. Compared to GEM, this approach offers better scalability by limiting the growth of its memory footprint. However, limitations persist. Memory storage still scales with the number of tasks, and the distillation process is computationally intensive, introducing a trade-off. Furthermore, the distillation process, driven by a single network, carries the risk of negative transfer or saturation when learning diverse tasks.

While these examples represent significant strides in the pursuit of lifelong learning, the fundamental challenge of seamlessly retaining and building upon knowledge remains unsolved. These solutions, while effective in certain scenarios, often succumb to issues like parameter redundancy, memory overhead, and computational inefficiency. A key contributor to these limitations is the tendency towards over-engineering – developing increasingly complex mechanisms to combat forgetting rather than seeking more elegant solutions that perhaps mirror the efficiency of biological learning processes. The quest for true lifelong learning continues, demanding a shift in

focus and potential inspiration from the natural systems that so effortlessly learn and adapt.

1.3 The brain with all the answers

Diverse natural organisms, from simple insects to complex mammals, demonstrate extraordinary lifelong learning capabilities. For instance, *Formica fusca* ants rapidly form lasting olfactory associations, exhibiting resistance to forgetting (Piqueret, Sandoz, and d'Ettorre 2019). This allows them to navigate complex environments and efficiently locate food sources. Mice (particularly *Mus musculus*), with their more complex brains, possess specialized regions supporting cognitive and behavioral flexibility across various tasks (Anacker and Hen 2017; Murray et al. 2015). Their well-developed hippocampus, for example, plays a crucial role in spatial learning and memory. Humans (*Homo sapiens*) with the most intricate neural architecture, exemplify lifelong learning. We can acquire and adapt knowledge throughout our lifespan, developing expertise in diverse domains – enough to pursue a dissertation on lifelong learning. Crucially, the neural architectures underlying these abilities are products of evolution, shaped by selective pressures to optimize learning within remarkably constrained memory and energy budgets compared to state-of-the-art machine learning models (Strubell, Ganesh, and McCallum 2019).

There are several brain regions implicated in lifelong learning such as the hippocampus (Moser, Kropff, and Moser 2008; Toda et al. 2019), anterior cingulate cortex (Stevens et al. 2011) and the prefrontal cortex (Blumenfeld and Ranganath 2007; Friedman and Robbins 2022; Miller and Cohen 2001; Wilson et al. 2014), to name a few. Given the depth and breadth of these regions and associated neural circuits, I would like to focus on the prefrontal cortex (PFC) in this dissertation.

The PFC is a neocortical brain region with connections to most other cortical regions such as sensory and motor areas as well as subcortical regions such as the thalamus and basal ganglia (Miller and Cohen 2001; Simons and Spiers 2003; Yeterian et al. 2012). These connections are generally reciprocal, enabling the PFC to integrate multi-dimensional inputs and exert top-down control in complex behaviors. Thus, the PFC plays a key role in cognition and associated executive functions such as reasoning, planning, language and social behavior (Jobson et al. 2021). Many theories have been classically put forward to explain the function of the PFC in these intricate behaviors (Desimone and Duncan 1995; Goldman-Rakic et al. 1997; Miller and Cohen 2001). While they differ in semantics and intellectual lenses, the central idea remains the same. Essentially, given a system of complex mappings between inputs and outputs, a given scenario can present the agent with a vast array of actions to choose from. The different mappings compete on grounds of saliency to facilitate the selection of different actions. The PFC does not represent the mappings themselves but produces biases towards certain patterns of neural activations that lead the brain to the corresponding action or behavior. One of the most popular analogies describes the PFC as a switch operator with the rest of the brain being a railroad network where the PFC guides the flow of information through specific sub-pathways or sub-circuits in order to facilitate an appropriate response (Figure 1.1) (Miller and Cohen 2001). This property of the PFC makes it an important center for context dependent decision-making – the precursor to lifelong learning.

The PFC is also a center for a diverse and heterogenous distribution of neuromodulators that together create a sensitive neurochemical environment for optimal

function (Cools and Arnsten 2022; Zhong et al. 2022). These neuromodulators have complex response profiles and play a myriad of roles in the PFC. One of the prime examples is dopamine – a neuromodulator popular for its role in learning and reward prediction error (Schultz 1998). Recently, several studies have highlighted dopamine's non-canonical functions in neocortical areas such as gating of behaviorally relevant sensory information, activation of PFC neurons to maintain working memory and relaying choice response signals to downstream regions (Floresco 2013; Ott and Nieder 2019). A

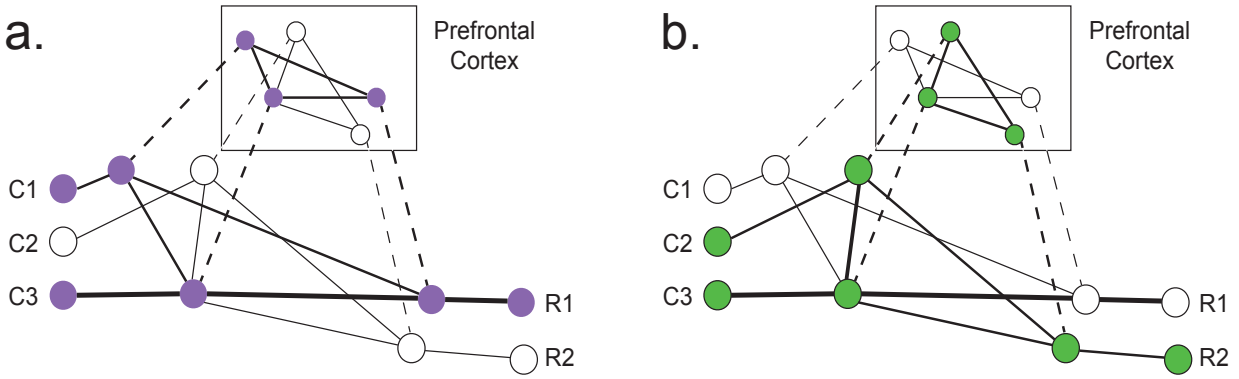


Figure 1-1 Schematic describing PFC function adapted from Miller and Cohen 2001.

a, Hypothesized neural network architecture with contextual input units (C1, C2 and C3), hidden units (part of the PFC) and behavioral output units (R1 and R2); C1 and C3 activations are directed via the PFC to output R1. b, C2 and C3 activations are directed to output R2.

specific study of high relevance to this dissertation shows that optogenetic stimulation of mesocortical dopamine terminals in the mPFC of mice increases the signal to noise ratio of neurons projecting to the dorsal periaqueductal gray (dPAG) in response to aversive stimuli but not reward stimuli (Figure 1.2) (Vander Weele et al. 2018).

Two important aspects of this experiment are that the stimuli were presented without any paired cue and that the authors did not observe this in mPFC neurons projecting to the nucleus accumbens (NAc). Thus, dopamine seems to be influencing the

firing of mPFC neurons in a context and anatomy specific manner. This idea can be extended to other neuromodulators of note such as serotonin (Andrade 2011; Athilingam et al. 2017; Celada, Puig, and Artigas 2013; Morgan et al. 2023), acetylcholine (Bloem, Poorthuis, and Mansvelder 2014; Cools and Arnsten 2022; Howe et al. 2017) and norepinephrine (Basu et al. 2022; Ventura et al. 2003; Zhang et al. 2013), each of which

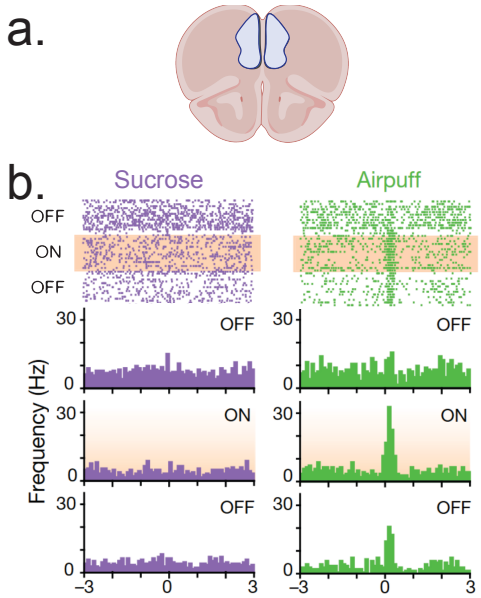


Figure 1-2 Dopamine function in mediating signal-to-noise ratio of mPFC neurons adapted from Vander Weele et al. 2018).

A, Schematic of a brain slice with the mPFC highlighted. B, (top) Response of example mPFC-dPAG neuron to presentations of sucrose (purple) and airpuff (green); (bottom) peristumulus time histogram over the trials.

has been shown to play a role in cognitive flexibility of the PFC.

Finally, dysfunction of the PFC is also associated with a variety of neuropsychiatric disorders and diseases. This ranges from social deficits found in patients of Autism Spectrum Disorder and schizophrenia (Bicks et al. 2015) to anxiety and Major Depressive Disorder (Pizzagalli and Roberts 2022). Thus, modeling the executive functioning of the PFC would lead to an architecture that is not only able to perform lifelong learning but can also be used as a computational testbed for studying diseased states.

1.4 Goals of this dissertation

This dissertation aims to study the cognitive capabilities of the PFC systematically with tight integration between computational modeling and experimental results.

Chapter 2 considers the PFC as a latent state system that constructs multiple mappings between neural population activations and behavior in a social competition behavioral paradigm. Using state space analysis, the PFC neural manifold is first segregated into subspaces based on social rank and competition outcome to study the distinct neural representations of the competition task. This largely involves using principal component analysis, a popular dimensionality reduction technique, to align the neural activity of multiple animals to the same reduced subspace for qualitative and quantitative comparisons between populations of dominant and subordinate animals. Then, a novel adaptation of a computational architecture called the Hidden Markov Model Generalized Linear Model (HMMGLM) is used to identify latent states in the PFC, each of which is paired with a separate linear mapping from neural activity to behavior. This model is used to predict behavioral dynamics in conjunction with evolving neural dynamics. The findings of this study highlight a possible mechanism that the PFC utilizes to facilitate context specific sub-circuit activation.

The line between adaptive and maladaptive behaviors is a complex one and must be considered when building models of the brain. Chapter 3 dives into this by studying the effect of a chronic mild stress paradigm on the PFC and associated behavioral dysfunction. Animals are first clustered into anhedonic phenotypes based on their behavior in a sucrose preference task. State space analysis is then used to study how

population activity in the PFC evolves over the course of the stress paradigm for the different phenotypes. Lastly, effectiveness of ketamine is tested as a therapeutic drug by studying the reversal in neural and behavioral changes to pre-stress levels.

Chapter 4 builds on the learning from Chapters 2 and 3 and considers the PFC as a modular hierarchical neural network architecture. It leverages the internal state dynamics from the HMM GLM with a novel neuromodulatory filtering mechanism driving these changes in the model. The model is validated on a contextual foraging task geared towards lifelong learning and is capable of re-producing PFC dysfunction via neuromodulatory manipulations.

1.5 Bibliography

Anacker, Christoph, and René Hen. 2017. “Adult Hippocampal Neurogenesis and Cognitive Flexibility — Linking Memory and Mood.” *Nature Reviews Neuroscience* 18(6): 335–46. doi:10.1038/nrn.2017.45.

Andrade, Rodrigo. 2011. “Serotonergic Regulation of Neuronal Excitability in the Prefrontal Cortex.” *Neuropharmacology* 61(3): 382–86. doi:10.1016/j.neuropharm.2011.01.015.

Athilingam, Jegath C, Roy Ben-Shalom, Caroline M Keeshen, Vikaas S Sohal, and Kevin J Bender. 2017. “Serotonin Enhances Excitability and Gamma Frequency Temporal Integration in Mouse Prefrontal Fast-Spiking Interneurons” ed. Marlene Bartos. *eLife* 6: e31991. doi:10.7554/eLife.31991.

Bai, Shaojie, J. Zico Kolter, and Vladlen Koltun. 2018. “An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling.” doi:10.48550/arXiv.1803.01271.

Banko, Michele, and Eric Brill. 2001. “Scaling to Very Very Large Corpora for Natural Language Disambiguation.” In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, Toulouse, France: Association for Computational Linguistics, 26–33. doi:10.3115/1073012.1073017.

Basu, Aakash, Jen-Hau Yang, Abigail Yu, Samira Glaeser-Khan, Jiesi Feng, John H. Krystal, Yulong Li, and Alfred P. Kaye. 2022. “Prefrontal Norepinephrine Represents a Threat Prediction Error under Uncertainty.” : 2022.10.13.511463. doi:10.1101/2022.10.13.511463.

Benna, Marcus K., and Stefano Fusi. 2016. “Computational Principles of Synaptic Memory Consolidation.” *Nature Neuroscience* 19(12): 1697–1706. doi:10.1038/nn.4401.
Bicks, Lucy K., Hiroyuki Koike, Schahram Akbarian, and Hirofumi Morishita. 2015. “Prefrontal Cortex and Social Cognition in Mouse and Man.” *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.01805.

Bloem, Bernard, Rogier Bernard Poorthuis, and Huibert Daniel Mansvelder. 2014. “Cholinergic Modulation of the Medial Prefrontal Cortex: The Role of Nicotinic Receptors in Attention and Regulation of Neuronal Activity.” *Frontiers in Neural Circuits* 8. doi:10.3389/fncir.2014.00017.
Blumenfeld, Robert S., and Charan Ranganath. 2007. “Prefrontal Cortex and Long-Term Memory Encoding: An Integrative Review of Findings from Neuropsychology and Neuroimaging.” *The Neuroscientist* 13(3): 280–91. doi:10.1177/1073858407299290.

Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik. 1992. “A Training Algorithm for Optimal Margin Classifiers.” In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT ’92, New York, NY, USA: Association for Computing Machinery, 144–52. doi:10.1145/130385.130401.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. “Language Models Are Few-Shot Learners.” doi:10.48550/arXiv.2005.14165.

Carion, Nicolas, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. 2020. “End-to-End Object Detection with Transformers.” In *Computer Vision – ECCV 2020*, eds. Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm. Cham: Springer International Publishing, 213–29. doi:10.1007/978-3-030-58452-8_13.

Celada, Pau, M. Victoria Puig, and Francesc Artigas. 2013. “Serotonin Modulation of Cortical Neurons and Networks.” *Frontiers in Integrative Neuroscience* 7. https://www.frontiersin.org/articles/10.3389/fnint.2013.00025 (February 10, 2024).

Cools, Roshan, and Amy F. T. Arnsten. 2022. “Neuromodulation of Prefrontal Cortex Cognitive Function in Primates: The Powerful Roles of Monoamines and Acetylcholine.” *Neuropsychopharmacology* 47(1): 309–28. doi:10.1038/s41386-021-01100-8.

Dean, Jeffrey, Greg Corrado, Rajat Monga, Kai Chen, Matthieu Devin, Mark Mao, Marc’ aurelio Ranzato, et al. 2012. “Large Scale Distributed Deep Networks.” In *Advances in Neural Information Processing Systems*, Curran Associates, Inc. https://papers.nips.cc/paper_files/paper/2012/hash/6aca97005c68f1206823815f66102863-Abstract.html (March 11, 2024).

Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. “ImageNet: A Large-Scale Hierarchical Image Database.” In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, , 248–55. doi:10.1109/CVPR.2009.5206848.

Desimone, Robert, and John Duncan. 1995. “Neural Mechanisms of Selective Visual Attention.” *Annual Review of Neuroscience* 18(Volume 18, 1995): 193–222. doi:10.1146/annurev.ne.18.030195.001205.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. “BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.” doi:10.48550/arXiv.1810.04805.

Dietterich, Thomas G. 2000. “Ensemble Methods in Machine Learning.” In *Multiple Classifier Systems*, Lecture Notes in Computer Science, Berlin, Heidelberg: Springer Berlin Heidelberg, 1–15. doi:10.1007/3-540-45014-9_1.

Dill, Ken A., and Justin L. MacCallum. 2012. “The Protein-Folding Problem, 50 Years On.” *Science* 338(6110): 1042–46. doi:10.1126/science.1219021.

Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, et al. 2021. “An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale.” doi:10.48550/arXiv.2010.11929.

Elmore, Leisha C., Donna B. Jeffe, Linda Jin, Michael M. Awad, and Isaiah R. Turnbull. 2016. “National Survey of Burnout among US General Surgery Residents.” *Journal of the American College of Surgeons* 223(3): 440–51. doi:10.1016/j.jamcollsurg.2016.05.014.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. “Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks.” *Nature* 542(7639): 115–18. doi:10.1038/nature21056.

Finn, Chelsea, Pieter Abbeel, and Sergey Levine. 2017. “Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks.” doi:10.48550/arXiv.1703.03400.

Floresco, Stan B. 2013. “Prefrontal Dopamine and Behavioral Flexibility: Shifting from an ‘Inverted-U’ toward a Family of Functions.” *Frontiers in Neuroscience* 7. doi:10.3389/fnins.2013.00062.

Friedman, Naomi P., and Trevor W. Robbins. 2022. “The Role of Prefrontal Cortex in Cognitive Control and Executive Function.” *Neuropsychopharmacology* 47(1): 72–89. doi:10.1038/s41386-021-01132-0.

Goldman-Rakic, P. S., A. R. Cools, K. Srivastava, A. C. Roberts, Trevor William Robbins, and Lawrence Weiskrantz. 1997. “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(1346): 1445–53. doi:10.1098/rstb.1996.0129.

Halevy, Alon, Peter Norvig, and Fernando Pereira. 2009. “The Unreasonable Effectiveness of Data.” *IEEE Intelligent Systems* 24(2): 8–12. doi:10.1109/MIS.2009.36.

Hanna, Timothy P., Will D. King, Stephane Thibodeau, Matthew Jalink, Gregory A. Paulin, Elizabeth Harvey-Jones, Dylan E. O’Sullivan, et al. 2020. “Mortality Due to Cancer Treatment Delay: Systematic Review and Meta-Analysis.” *BMJ* 371: m4087. doi:10.1136/bmj.m4087.

Holmer, Hampus, Adam Lantz, Teena Kunjumen, Samuel Finlayson, Marguerite Hoyler, Amani Siyam, Hernan Montenegro, et al. 2015. “Global Distribution of Surgeons, Anaesthesiologists, and Obstetricians.” *The Lancet Global Health* 3: S9–11. doi:10.1016/S2214-109X(14)70349-3.

Howe, William M., Howard J. Gritton, Nicholas A. Lusk, Erik A. Roberts, Vaughn L. Hetrick, Joshua D. Berke, and Martin Sarter. 2017. “Acetylcholine Release in Prefrontal Cortex Promotes Gamma Oscillations and Theta–Gamma Coupling during Cue Detection.” *The Journal of Neuroscience* 37(12): 3215–30. doi:10.1523/JNEUROSCI.2737-16.2017.

Jobson, Dan D, Yoshiki Hase, Andrew N Clarkson, and Rajesh N Kalaria. 2021. “The Role of the Medial Prefrontal Cortex in Cognition, Ageing and Dementia.” *Brain Communications* 3(3): fcab125. doi:10.1093/braincomms/fcab125.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, et al. 2021. “Highly Accurate Protein Structure Prediction with AlphaFold.” *Nature* 596(7873): 583–89. doi:10.1038/s41586-021-03819-2.

Kirkpatrick, James, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, et al. 2017. “Overcoming Catastrophic Forgetting in Neural Networks.” *Proceedings of the National Academy of Sciences* 114(13): 3521–26. doi:10.1073/pnas.1611835114.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton. 2012. “ImageNet Classification with Deep Convolutional Neural Networks.” In *Advances in Neural Information Processing Systems*, Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html (March 11, 2024).

Kudithipudi, Dhireesha, Mario Aguilar-Simon, Jonathan Babb, Maxim Bazhenov, Douglas Blackiston, Josh Bongard, Andrew P. Brna, et al. 2022. “Biological Underpinnings for Lifelong Learning Machines.” *Nature Machine Intelligence* 4(3): 196–210. doi:10.1038/s42256-022-00452-0.

Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. “Gradient-Based Learning Applied to Document Recognition.” *Proceedings of the IEEE* 86(11): 2278–2324. doi:10.1109/5.726791.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” *Nature* 521(7553): 436–44. doi:10.1038/nature14539.

Lopez-Paz, David, and Marc’Aurelio Ranzato. 2022. “Gradient Episodic Memory for Continual Learning.” doi:10.48550/arXiv.1706.08840.

McKinney, Scott Mayer, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, et al. 2020. “International Evaluation of an AI System for Breast Cancer Screening.” *Nature* 577(7788): 89–94. doi:10.1038/s41586-019-1799-6.

Miller, E. K., and J. D. Cohen. 2001. “An Integrative Theory of Prefrontal Cortex Function.” *Annual Review of Neuroscience* 24: 167–202. doi:10.1146/annurev.neuro.24.1.167.

Mnih, Volodymyr, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Tim Harley, Timothy P Lillicrap, David Silver, and Koray Kavukcuoglu. “Asynchronous Methods for Deep Reinforcement Learning.”

Morgan, Ashlea A., Nuno D. Alves, Gregory S. Stevens, Tamanna T. Yeasmin, Alexandra Mackay, Saige Power, Derya Sargin, et al. 2023. “Medial Prefrontal Cortex Serotonin Input Regulates Cognitive Flexibility in Mice.” : 2023.03.30.534775. doi:10.1101/2023.03.30.534775.

Moser, Edvard I., Emilio Kropff, and May-Britt Moser. 2008. “Place Cells, Grid Cells, and the Brain’s Spatial Representation System.” *Annual Review of Neuroscience* 31(1): 69–89. doi:10.1146/annurev.neuro.31.061307.090723.

Murray, Andrew J., Marta U. Woloszynowska-Fraser, Laura Ansel-Bollepalli, Katy L. H. Cole, Angelica Foggetti, Barry Crouch, Gernot Riedel, and Peer Wulff. 2015. “Parvalbumin-Positive Interneurons of the Prefrontal Cortex Support Working Memory and Cognitive Flexibility.” *Scientific Reports* 5(1): 16778. doi:10.1038/srep16778.

Niu, Feng, Benjamin Recht, Christopher Re, and Stephen J. Wright. 2011. “HOGWILD!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent.” doi:10.48550/arXiv.1106.5730.

Ott, Torben, and Andreas Nieder. 2019. “Dopamine and Cognitive Control in Prefrontal Cortex.” *Trends in Cognitive Sciences* 23(3): 213–34. doi:10.1016/j.tics.2018.12.006.

Pandey, Mohit, Michael Fernandez, Francesco Gentile, Olexandr Isayev, Alexander Tropsha, Abraham C. Stern, and Artem Cherkasov. 2022. “The Transformational Role

of GPU Computing and Deep Learning in Drug Discovery." *Nature Machine Intelligence* 4(3): 211–21. doi:10.1038/s42256-022-00463-x.

Piqueret, Baptiste, Jean-Christophe Sandoz, and Patrizia d'Ettorre. 2019. "Ants Learn Fast and Do Not Forget: Associative Olfactory Learning, Memory and Extinction in Formica Fusca." *Royal Society Open Science* 6(6): 190778. doi:10.1098/rsos.190778.

Pizzagalli, Diego A., and Angela C. Roberts. 2022. "Prefrontal Cortex and Depression." *Neuropsychopharmacology* 47(1): 225–46. doi:10.1038/s41386-021-01101-7.

Rebuffi, Sylvestre-Alvise, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert. 2017. "iCaRL: Incremental Classifier and Representation Learning." In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, 5533–42. doi:10.1109/CVPR.2017.587.

Rosenblatt, F. 1958. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review* 65(6): 386–408. doi:10.1037/h0042519.

Saeidi, H., J. D. Opfermann, M. Kam, S. Wei, S. Leonard, M. H. Hsieh, J. U. Kang, and A. Krieger. 2022. "Autonomous Robotic Laparoscopic Surgery for Intestinal Anastomosis." *Science Robotics* 7(62): eabj2908. doi:10.1126/scirobotics.abj2908.

Samuel, A. L. 1959. "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development* 3(3): 210–29. doi:10.1147/rd.33.0210.

Schultz, Wolfram. 1998. "Predictive Reward Signal of Dopamine Neurons." *Journal of Neurophysiology* 80(1): 1–27. doi:10.1152/jn.1998.80.1.1.

Simons, Jon S., and Hugo J. Spiers. 2003. "Prefrontal and Medial Temporal Lobe Interactions in Long-Term Memory." *Nature Reviews Neuroscience* 4(8): 637–48. doi:10.1038/nrn1178.

Squire, Larry R., Lisa Genzel, John T. Wixted, and Richard G. Morris. 2015. "Memory Consolidation." *Cold Spring Harbor Perspectives in Biology* 7(8): a021766. doi:10.1101/cshperspect.a021766.

Stevens, Francis L., Robin A. Hurley, Katherine H. Taber, Robin A. Hurley, L. Anne Hayman, and Katherine H. Taber. 2011. "Anterior Cingulate Cortex: Unique Role in Cognition and Emotion." *The Journal of Neuropsychiatry and Clinical Neurosciences* 23(2): 121–25. doi:10.1176/jnp.23.2.jnp121.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. 2019. "Energy and Policy Considerations for Deep Learning in NLP." doi:10.48550/arXiv.1906.02243.

Toda, Tomohisa, Sarah Parylak, Sara B. Linker, and Fred H. Gage. 2019. “The Role of Adult Hippocampal Neurogenesis in Brain Health and Disease.” *Molecular psychiatry* 24(1): 67–87. doi:10.1038/s41380-018-0036-2.

Vander Weele, Caitlin M., Cody A. Siciliano, Gillian A. Matthews, Praneeth Namburi, Ehsan M. Izadmehr, Isabella C. Espinel, Edward H. Nieh, et al. 2018. “Dopamine Enhances Signal-to-Noise Ratio in Cortical-Brainstem Encoding of Aversive Stimuli.” *Nature* 563(7731): 397–401. doi:10.1038/s41586-018-0682-1.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. “Attention Is All You Need.” In *Advances in Neural Information Processing Systems*, Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (December 8, 2023).

Ventura, Rossella, Simona Cabib, Antonio Alcaro, Cristina Orsini, and Stefano Puglisi-Allegra. 2003. “Norepinephrine in the Prefrontal Cortex Is Critical for Amphetamine-Induced Reward and Mesoaccumbens Dopamine Release.” *The Journal of Neuroscience* 23(5): 1879–85. doi:10.1523/JNEUROSCI.23-05-01879.2003.

Wilson, Robert C., Yuji K. Takahashi, Geoffrey Schoenbaum, and Yael Niv. 2014. “Orbitofrontal Cortex as a Cognitive Map of Task Space.” *Neuron* 81(2): 267–79. doi:10.1016/j.neuron.2013.11.005.

Yeterian, Edward H., Deepak N. Pandya, Francesco Tomaiuolo, and Michael Petrides. 2012. “The Cortical Connectivity of the Prefrontal Cortex in the Monkey Brain.” *Cortex; a journal devoted to the study of the nervous system and behavior* 48(1): 58–81. doi:10.1016/j.cortex.2011.03.004.

Zhang, Zizhen, Steven Cordeiro Matos, Sonia Jego, Antoine Adamantidis, and Philippe Séguéla. 2013. “Norepinephrine Drives Persistent Activity in Prefrontal Cortex via Synergistic A1 and A2 Adrenoceptors.” *PLOS ONE* 8(6): e66122. doi:10.1371/journal.pone.0066122.

Zhong, Wen, Swapnali Barde, Nicholas Mitsios, Csaba Adori, Per Oksvold, Kalle von Feilitzen, Liam O’Leary, et al. 2022. “The Neuropeptide Landscape of Human Prefrontal Cortex.” *Proceedings of the National Academy of Sciences* 119(33): e2123146119. doi:10.1073/pnas.2123146119.

Chapter 2 Cortical ensembles orchestrate social competition through hypothalamic outputs

In this chapter, I study the role of the medial prefrontal cortex in a social competition assay and hypothesize a latent state system as a possible underlying mathematical model. I use data from an experimental paradigm designed and executed by my co-author Nancy Padilla-Coreano. I apply state space analysis along with dimensionality reduction to identify neural manifolds where the mPFC operates when an agent is acting as a dominant vs subordinate in a social competition task. I calculate the separation of the manifold subspaces under different conditions such as social rank, competition outcome and performing the task alone. I adapt a Hidden Markov Model Generalized Linear Model to explain the data observed in the experiment. I use decoding theory to identify potential downstream brain regions from the mPFC where social hierarchy information might be directed.

2.1 Abstract

Most social species self-organize into dominance hierarchies (Chiao 2010; Wang, Kessels, and Hu 2014), which decreases aggression and conserves energy (Bernstein 1981; Dewsbury 1982), but it is not clear how individuals know their social rank. We have only begun to learn how the brain represents social rank (Hou et al. 2016; Karamihalev et al. 2020; So et al. 2015; Wang et al. 2011; Zhou et al. 2017) and guides behavior on the basis of this representation. The medial prefrontal cortex (mPFC) is involved in social dominance in rodents (Wang et al. 2011; Zhou et al. 2017) and humans (Ligneul et al.

2016; Zink et al. 2008). Yet, precisely how the mPFC encodes relative social rank and which circuits mediate this computation is not known. We developed a social competition assay in which mice compete for rewards, as well as a computer vision tool (AlphaTracker) to track multiple, unmarked animals. A hidden Markov model combined with generalized linear models was able to decode social competition behavior from mPFC ensemble activity. Population dynamics in the mPFC predicted social rank and competitive success. Finally, we demonstrate that mPFC cells that project to the lateral hypothalamus promote dominance behavior during reward competition. Thus, we reveal a cortico-hypothalamic circuit by which the mPFC exerts top-down modulation of social dominance.

2.2 Introduction

The mPFC is best known for its role in higher cognitive functions, with theoretical emphasis on mPFC integration of sensory and limbic information to flexibly guide behavior on the basis of task rules (Miller and Cohen 2001). Notably, mPFC circuitry has also been implicated in social cognition, social memory and dominance (Levy et al. 2019; Ligneul et al. 2016; Murugan et al. 2017; Wang et al. 2011; Zhou et al. 2017). We hypothesized that mPFC neurons encode social rank and are part of top-down circuits to guide behavior on the basis of social rank (Lee, Chen, and Tye 2021).

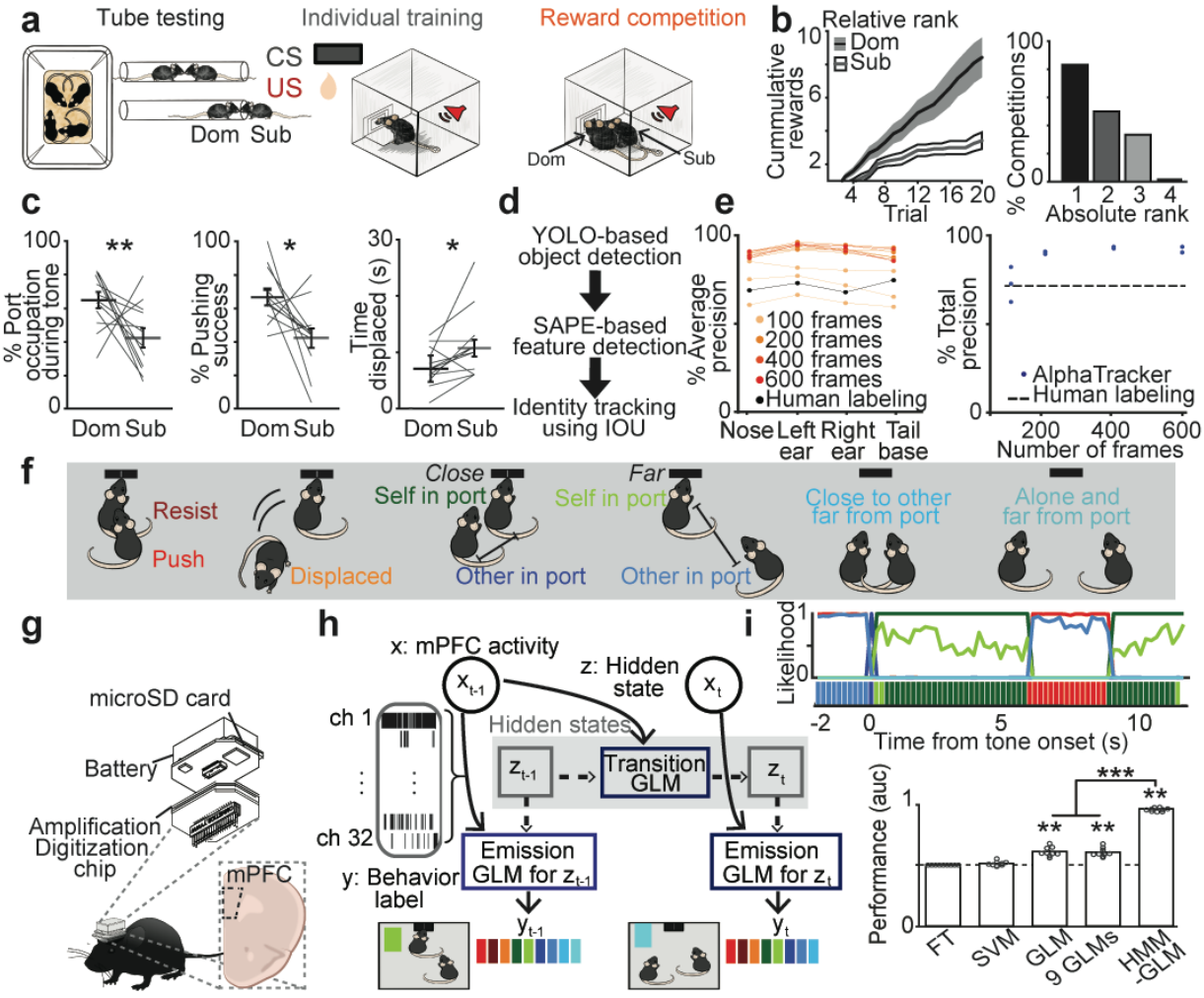
We designed a reward competition assay wherein mice that were linearly ranked among their cage mates competed for a liquid reward delivered during a tone. This task design optimized rigorous statistical examination of ethologically relevant behaviors in a trial structure (Fig. 2.1a). We considered relative social ranks within each competing pair, enabling a within-subject comparison for intermediate-ranked animals. After individually

learning that the tone predicted reward delivery (Supp. Fig. 2.3a), mice competed for rewards with a cage mate. Dominant mice, as defined by the tube test (Wang et al. 2011), obtained more rewards, spent more time at the reward port and were more successful at displacing competitors (Fig. 2.1b, c and Supp. Fig. 2.3).

To automatically track the behavior of multiple, unmarked mice, we developed AlphaTracker, a deep-learning computer vision tool that combines two neural networks, one to create a bounding box for each subject, and another for pose estimation to detect multiple, unmarked animals (Fig. 2.1d, e). AlphaTracker also applies another algorithm to track animal identity across frames considering animal positions from the previous frame (Fig. 2.1d; see Methods). The performance of AlphaTracker surpasses human accuracy when tracking two or four mice (Supp. Fig. 2.4) and includes unsupervised clustering of high-dimensional tracking output data to aid in the identification of novel behavioral motifs (Supp. Fig. 2.5).

Figure 2-1 Novel social dominance assay and deep learning tool for tracking multiple animals.

a, Reward competition behavioral paradigm. b, Mice with higher relative ranks (dominant (Dom)) collected more rewards than relative subordinates (Sub) when competing in dyads (n = 12 dyads; sign rank test on total rewards, P = 0.008). Left, cumulative rewards across trials. Right, percentage of competitions won by absolute rank (n = 6 competitions per rank). c, Port occupation, pushing success (pushing that resulted in displacement of competitor) and time displaced from port differed by relative rank (n = 12 dyads; sign rank test, occupation P = 0.04; pushing success P = 0.02; displaced P = 0.016). d, Architecture of AlphaTracker, which combines two convolutional neural networks (YOLO, you only look once; SAPE, single-animal pose estimation, which is a squeeze-and-excitation network) and intersection over union (IOU) for identity tracking. e, AlphaTracker precision for tracking two unmarked, near-identical mice is higher than human precision separated by body parts (left) or total (right; average precision across body parts; n = 3 subsampled replicates). f, Social competition behavioral labels used for decoder models. g, Wireless device to record neural activity from the mPFC (n = 965 trials from 32 sessions from 13 mice). Image modified from SpikeGadgets. h, Architecture of the HMM–GLM model to describe the relationship between neural activity and behavioral states. Hidden states are shown on the grey background. Ch, channel; x, mPFC activity; y, behavior label; z, hidden state. i, Top, example trial with real behavioral labels and prediction for HMM–GLM six-state model (colors from f). Bottom, performance across models based on area under the receiver operating curve (AUC; n = 9 behaviors; Kruskal–Wallis test P = 6.7 × 10−8; model versus chance sign test P = 0.004 for six-state HMM–GLM, 9 GLMs and GLM; Wilcoxon rank sum HMMGLM versus GLM P = 4 × 10−5, HMM–GLM versus 9 GLMs P = 4 × 10−5, GLM versus 9 GLMs P = 0.54). FT, frequency of behaviors table. Data are presented as mean values ± s.e.m. *P < 0.05, **P < 0.01, ***P < 0.001.


a
Tube testing
Individual training
Reward competition
CS
US
Dom Sub
Dom
Sub
b
Relative rank
Dom
Sub
Cummulative rewards
Trial
% Competitions
Absolute rank
c
% Port occupation during tone
% Pushing success
Time displaced (s)
Dom Sub
d
YOLO-based object detection
SAPE-based feature detection
Identity tracking using IOU
e
% Average precision
100 frames
200 frames
400 frames
600 frames
Human labeling
Nose
Left ear
Right ear
Tail base
% Total precision
AlphaTracker
Human labeling
Number of frames
f
Resist
Push
Displaced
Close
Self in port
Other in port
Far
Self in port
Other in port
Close to other far from port
Alone and far from port
g
microSD card
Battery
Amplification Digitization chip
mPFC
h
x: mPFC activity
z: Hidden state
ch 1
ch 32
y: Behavior label
Hidden states
Transition GLM
Emission GLM for z_{t-1}
Emission GLM for z_t
i
Likelihood
Time from tone onset (s)
Performance (auc)
FT
SVM
GLM
9 GLMs
HMM -GLM

2.3 mPFC neurons encode competition behavior

To investigate whether mPFC neurons encode competition behaviors, we used wireless head-mounted devices to record cellular-resolution activity during the social competition (Fig. 2.1f, g, Supp. Fig. 2.6 and Supp. Fig. 8.1a–g). After AlphaTracker facilitated identification of nine different behavioral labels (Fig. 2.1f and Supp. Fig. 8.1h, i), we investigated whether the mPFC predicted specific behavioral outputs. Given the ability of mPFC neurons to be selective for different stimulus features under different contexts (Rigotti et al. 2013), we posited that mPFC neural activity could be dynamic, and that representations may be hierarchical, and influenced by internal hidden states. We turned to an unsupervised method to identify hidden states by combining a hidden Markov model (HMM) with generalized linear models (GLMs) (Calhoun, Pillow, and Murthy 2019; Escola et al. 2011) and adapted it to use mPFC neural activity to predict behavior. In our HMM–GLM model, one set of multinomial GLMs predicts the transition probabilities between hidden states, and each hidden state is paired with another multinomial GLM that describes the relationship between neural activity and behavior (Fig. 2.1h).

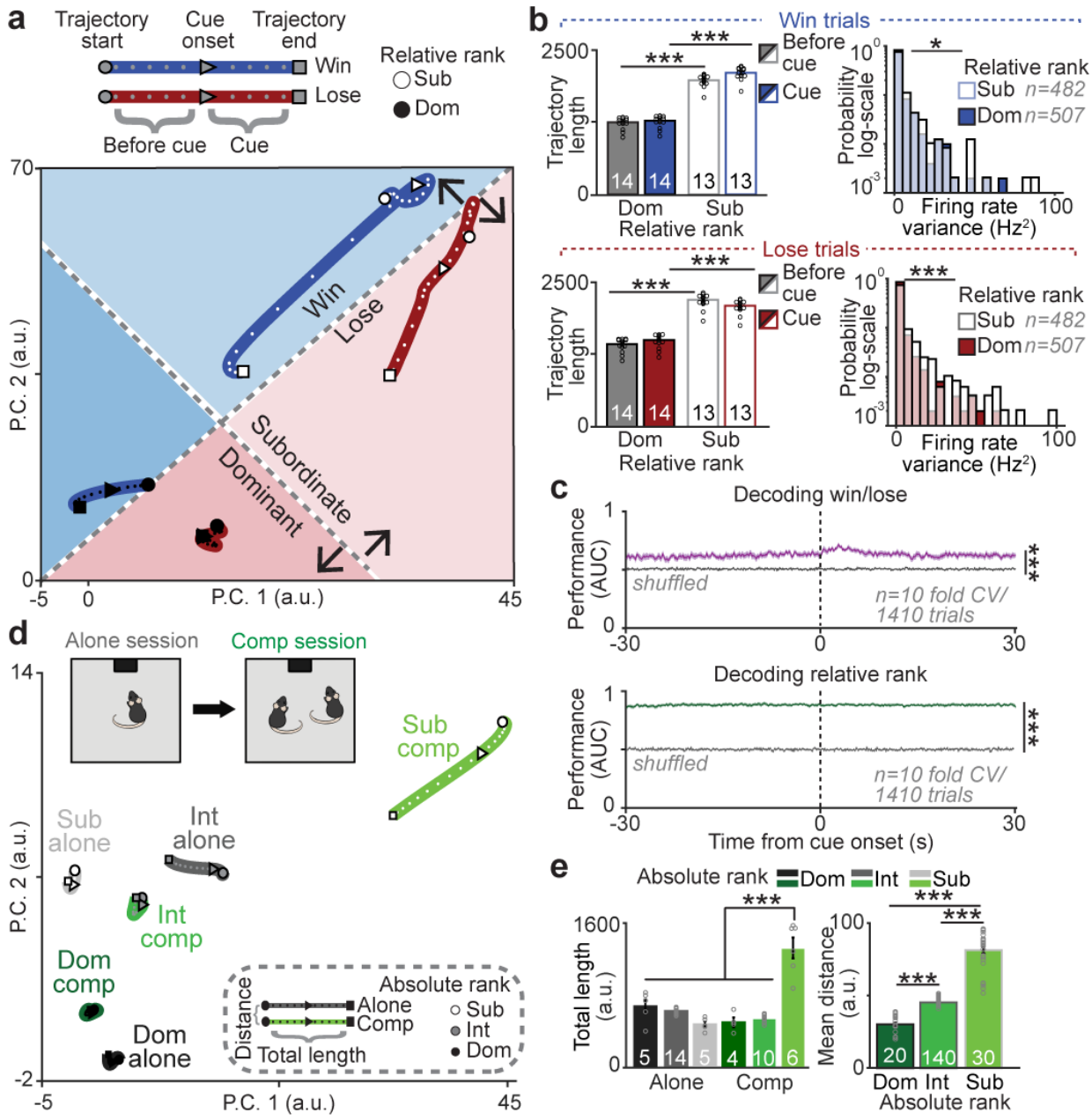
An HMM–GLM model with six hidden states decoded behavioral labels from neural activity with superior cross-validated performance compared with that of static models (Fig. 2.1i and Supp. Fig. 2.7j). The model performed equally well when training for one relative rank and testing on the other (Supp. Fig. 2.7k, l), suggesting that mPFC encoding of social competition behavior is generalizable across relative ranks. Given this finding, we then considered whether there is a stable and simple encoding of rank in mPFC neural representations, and whether these variables could themselves predict behavior.

2.4 mPFC reflects rank and winning

We next investigated whether mPFC neural activity could be used to decode relative social rank, and whether the neural representation of relative social rank is triggered by discrete task-relevant events (cued competition trials or port entries) or stably represented throughout the task. To visualize population activity, we plotted the population activity vector for task-relevant events (Supp. Fig. 2.8a) in a lower-dimensional firing-rate space using principal component analysis. Neural trajectories during the cue and port entries of the self or other (competitor) for win or lose trials occupied segregated neural activity subspaces even before the cue onset, suggesting separable brain states preceding each trial (Fig. 2.2a and Supp. Fig. 2.8), consistent with primate studies (Piva et al. 2019). Relative subordinates had longer neural trajectories compared with those of relative dominants (Fig. 2.2b and Supp. Fig. 2.8b–f). Indeed, our analyses revealed larger firing rate variance, but not faster firing rate changes, for relative subordinate mice (Fig. 2.2b and Supp. Fig. 2.8c). We ruled out the possible contributions of potential confounds (for example, subject location, distance to reward port and identity) to the differences in neural trajectories across ranks (Supp. Fig. 2.9).

Figure 2-2 Social rank and competitive success is decoded by mPFC population dynamics.

a, Neural trajectories of mPFC firing rate differ by relative rank during the tone presentation for win and lose trials in a lower-dimensional common principal component (PC) subspace (trajectories include a total of 20 mice; 27 sessions; dominants, n = 507 cells; subordinates, n = 482 cells). a.u., arbitrary units. b, Neural trajectory lengths for win (top; two-way repeated measures analysis of variance (ANOVA) effects of relative rank $F_{1,25} = 1{,}090$, $P = 2 \times 10^{-16}$ and interaction $F_{1,25} = 1{,}660$, $P = 3 \times 10^{-14}$) and lose (bottom; two-way repeated measures ANOVA effect of relative rank $F_{1,25} = 883$, $P = 9 \times 10^{-16}$ and interaction $F_{1,25} = 2{,}995$, $P = 9 \times 10^{-16}$) trials. Firing rate variance is higher for relative subordinates (number of neurons indicated in plots; top, win trials: Kolmogorov–Smirnov test $P = 0.01$, Wilcoxon rank sum $P = 0.19$; bottom, lose trials: Kolmogorov–Smirnov test $P = 5 \times 10^{-7}$, Wilcoxon rank sum $P = 2 \times 10^{-9}$). Light blue and pink indicate overlapping bars. c, SVM performance is higher than chance for decoding competitive success (top) and relative rank (bottom) (grey: shuffled data performance; Wilcoxon rank sum: competitive success $P = 2 \times 10^{-4}$, relative rank $P = 2 \times 10^{-4}$; competitive success versus relative rank $P = 2 \times 10^{-4}$). CV, cross-validation. d, Neural trajectories of mPFC population firing rate by absolute rank (dominant = rank 1; intermediates (Int) = ranks 2 and 3; subordinate = rank 4) when performing the reward task alone versus in competition (Comp) in a lower-dimensional common PC subspace (neurons in alone session: dominant = 111, intermediate = 259, subordinate = 140; competition: dominant = 309, intermediate = 359, subordinate = 330). e, Left, trajectory during the tone is higher for subordinates during competition (two-way ANOVA main effect of rank $F_{2,38} = 30.4$, $P = 1 \times 10^{-8}$, task $F_{1,38} = 26.1$, $P = 9 \times 10^{-6}$ and interaction $F_{2,38} = 70.1$, $P = 1 \times 10^{-13}$). Right, distance between alone and competition tone trajectories increases with rank (n reflects all possible combinations of alone versus competition trajectories; one-way ANOVA main effect of rank $F_{2,187} = 536$, $P = 3 \times 10^{-78}$). Post-hoc comparisons are Bonferroni-corrected t-tests, $^{*}P < 0.05$, $^{***}P < 0.001$. Two-way ANOVAs were for rank and event (baseline versus event) or rank and trial type and sample size indicated in plots. Data are presented as mean ± s.e.m.


a
Trajectory start
Cue onset
Trajectory end
Win
Lose
Before cue
Cue
Relative rank
Sub
Dom
Win
Lose
Subordinate
Dominant
P.C. 2 (a.u.)
P.C. 1 (a.u.)
b
Win trials
Lose trials
Before cue
Cue
Trajectory length
Dom
Sub
Relative rank
Probability log-scale
Firing rate variance (Hz²)
Relative rank
Sub n=482
Dom n=507
c
Decoding win/lose
Decoding relative rank
Performance (AUC)
shuffled
n=10 fold CV/ 1410 trials
Time from cue onset (s)
d
Alone session
Comp session
Sub alone
Int alone
Int comp
Sub comp
Dom comp
Dom alone
Distance
Total length
Absolute rank
Alone
Comp
Sub
Int
Dom
e
Absolute rank Dom Int Sub
Total length (a.u.)
Mean distance (a.u.)
Alone
Comp
Dom Int Sub
Absolute rank

2.5 mPFC predicts future wins

To directly test the hypothesis that the mPFC encodes relative rank and competitive success at the population level, we trained a support vector machine (SVM) classifier to decode these binary states from single-trial data (Supp. Fig. 2.10a). An SVM was able to decode both competitive success and relative rank—even before cue onset (Fig. 2.2c and Supp. Fig. 2.10b–e), consistent with the notion that state differences in the mPFC correlate with future winning (Zhou et al. 2017). Social rank was more accurately decoded than competition outcome from mPFC neural activity (Fig. 2.2c), perhaps reflecting the relative stability of rank versus competitive success. Although our data are consistent with the idea of a 'winning effect' or a 'losing streak' (Dugatkin 1997; Zhou et al. 2017), the decoding accuracy across the trial was consistently above chance. PFC neural activity could predict whether the next trial would be a win or a loss ~30 s before the competition trial began, providing cellular evidence supporting the psychological concept of 'a winning mindset'.

Notably, we can decode the absolute social rank of individuals from mPFC activity, even when they are alone (Supp. Fig. 2.10f–h). To visualize differences between tone responses to the reward while alone versus in competition, we plotted the neural trajectories across tasks in the same subspace of the principal component analysis (Fig. 2.2d). Subordinate mice (rank 4) had larger changes induced by competition with longer tone trajectory lengths during competition (Fig. 2.2e). By contrast, dominants (rank 1) showed the smallest differences between the alone and competition state. To confirm that population dynamics differed between receiving the reward alone versus winning, we recorded the same neurons while animals performed the reward task alone versus in

competition and found that an SVM could decode trial type from mPFC population dynamics (Supp. Fig. 2.10i, j). Notably, relative rank could be predicted in intermediate-ranking animals (Supp. Fig 2.10e). However, we cannot rule out the possibility that the representation may reflect social identity and the associated social history with that individual rather than relative rank alone; indeed, it is yet unclear whether the brain is capable of separably representing rank and identity. Together, these data demonstrate that the mPFC has a dynamic, but consistent, representation of social rank and competitive success despite having multiple, rank-independent hidden states for encoding behavior during social competition.

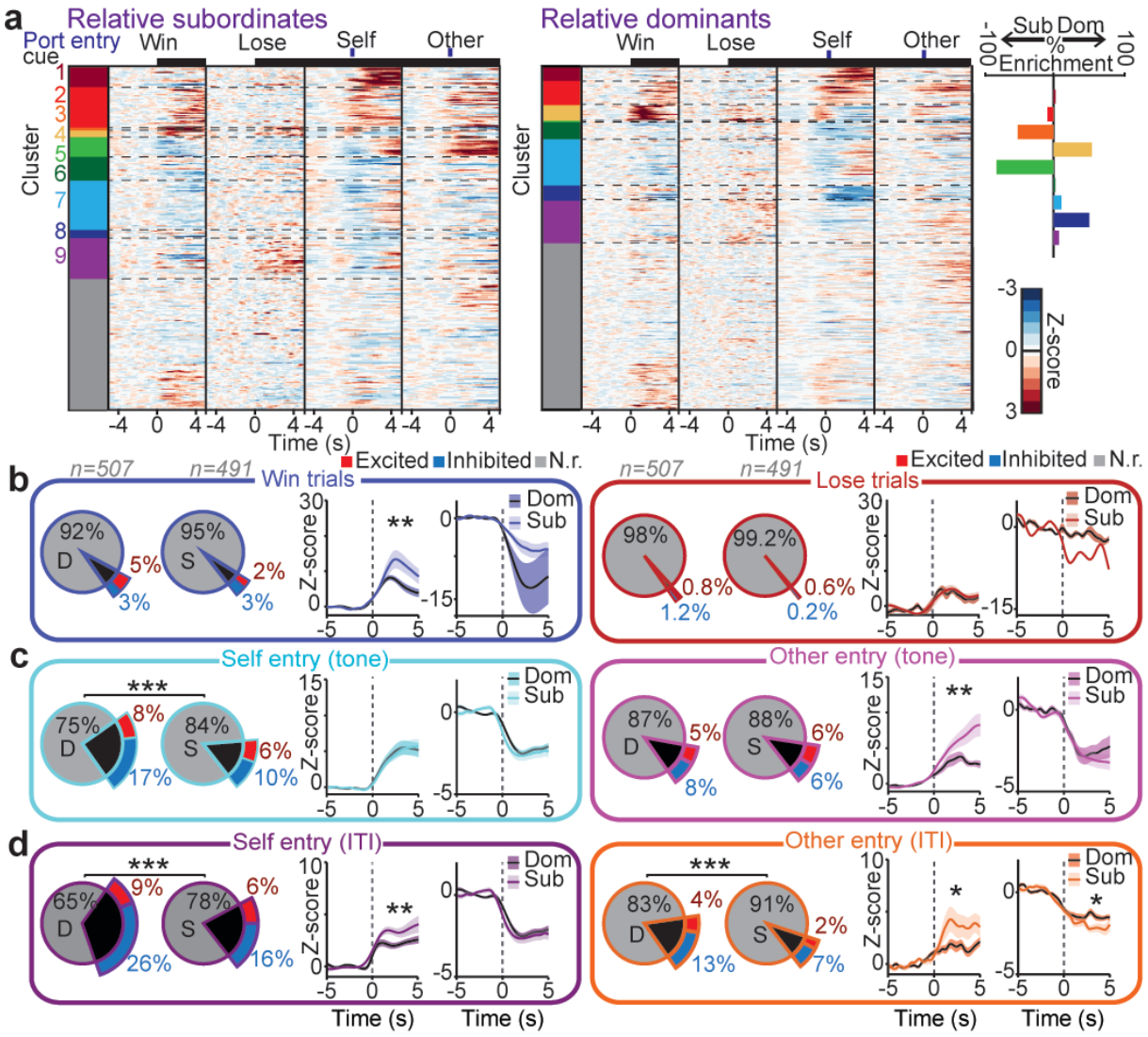
2.6 Rank-dependent mPFC representations

Given that the mPFC encodes social rank and competitive success, we posited that specific ensembles of cells might encode distinct task-relevant events in a rank-dependent manner to provide a distributed representation of social rank and competitive success. To investigate whether social rank is represented within the mPFC at the single-cell level, we analyzed the firing rate of individual mPFC neurons during discrete reward competition events. mPFC single units showed diverse responses to the tone for win-or-lose trials and to port entries performed by self or the other (that is, the competitor) that differed by social rank (Fig. 2.3a and Supp. Fig. 2.11a). We quantified the ensemble sizes and magnitude of responses to the task-relevant events while animals were alone versus in social competition (Fig. 2.3b–d and Supp. Fig. 2.11b–d). During competition, but not while alone, relative dominants had more cells that were responsive to self port entries whereas subordinates had larger responses to win trials and port entries of the other (Fig. 2.3b–d and Supp. Fig. 2.5b–d). Furthermore, the mPFC neurons of relative

subordinates exhibited larger phasic responses in response to task events, consistent with the longer neural trajectories observed (Fig 2.2).

Figure 2-3 Relative dominants have more reward-seeking-behavior cells whereas subordinates have larger responses to competitor behavior.

a, Left, heatmaps for mPFC firing rate responses to task-relevant events during the reward competition. Colors indicate clusters obtained by hierarchical clustering. Cells included in heatmap have a Z score greater than 2 or less than −1 (dominant n = 326; subordinate n = 305); clusters with a non-responsive population average are labelled grey. Right, difference between relative dominant and subordinate cells (percentage enrichment) across functional clusters. Black horizontal lines represent the cue. b, Left, response magnitude for mPFC cells during win trials differed across relative ranks (D, dominant; S, subordinate; Fisher's exact test, total responsive cells per group P = 0.30; Wilcoxon rank sum across groups: excited P = 0.01; inhibited P = 0.06). Right, number of responsive cells and response magnitude for lose trials did not differ by relative rank (Fisher's exact test, total responsive per group P = 0.17; Wilcoxon rank sum: excited P = 0.62; inhibited P = 0.28). c, Left, number of mPFC cells responsive to self tone port entry was higher for relative dominants (Fisher's exact test, total responsive cells per group P = 0.0003; Wilcoxon rank sum: excitation P = 0.28 and inhibition P = 0.99). Right, there was no difference in number of cells responsive to other tone port entry whereas the excitation magnitude was higher for relative subordinates (Fisher's exact test, total responsive cells per group P = 0.84; Wilcoxon rank sum: excitation P = 0.006, inhibition P = 0.11). d, Left, relative dominants had more mPFC cells responsive to self port entries during the inter-trial interval (ITI) whereas relative subordinates had larger excitation magnitude (Fisher's exact test, P = 9.8 × 10−6; Wilcoxon rank sum: excitation P = 0.006 and inhibition P = 0.28). Right, relative dominants had more mPFC cells responsive to other port entries during the inter-trial interval whereas subordinates had larger magnitudes (Fisher's exact test, P = 0.00019; Wilcoxon rank sum: excitation P = 0.015 and inhibition P = 0.04). Recordings were collected from 20 mice; sample indicated in plots is cells. Data are presented as mean ± s.e.m. *P < 0.05, **P < 0.01, ***P < 0.001.


a
Relative subordinates
Relative dominants
Port entry cue
Win
Lose
Self
Other
Cluster
Time (s)
Sub Dom
% Enrichment
-100
100
Z-score
b
Win trials
Lose trials
n=507
n=491
Excited
Inhibited
N.r.
92%
D
5%
3%
95%
S
2%
3%
98%
0.8%
1.2%
99.2%
0.6%
0.2%
Dom
Sub
c
Self entry (tone)
Other entry (tone)
75%
8%
17%
84%
6%
10%
87%
5%
8%
88%
6%
6%
d
Self entry (ITI)
Other entry (ITI)
65%
9%
26%
78%
6%
16%
83%
4%
13%
91%
2%
7%

2.7 mPFC-LH neurons modulate social dominance

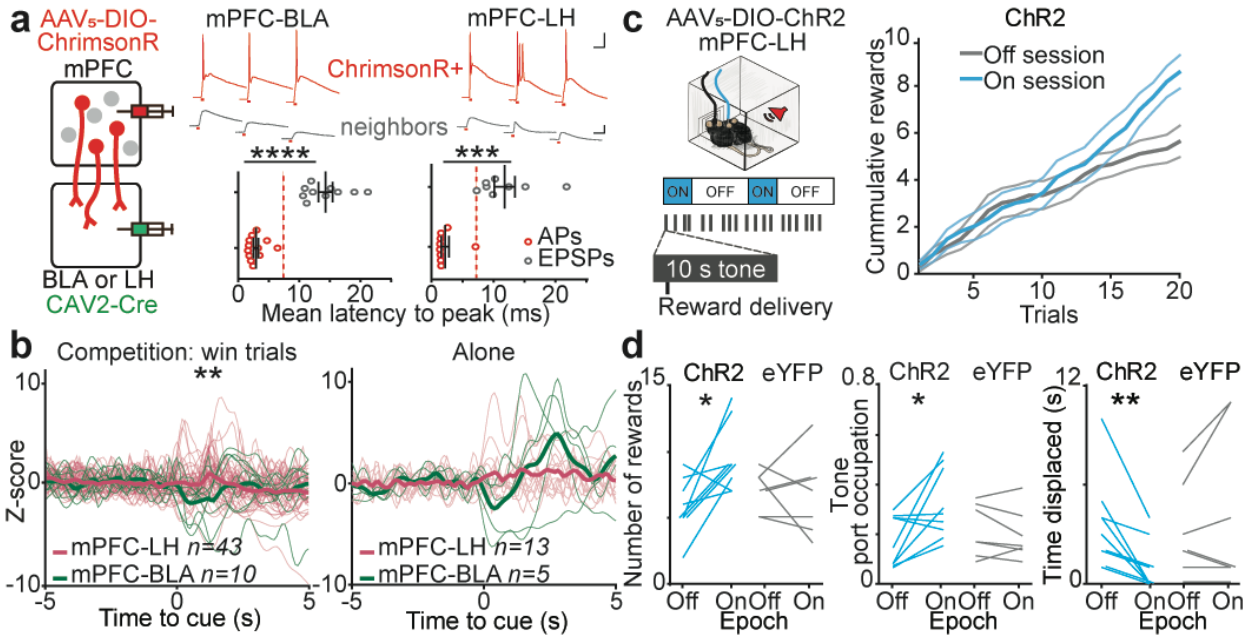
Given the functional diversity of neural responses from individual mPFC neurons, we next investigated how information was routed out of the mPFC during social competition to downstream subcortical targets.

The lateral hypothalamus (LH) comprises a diversity of cell types and has been shown to drive hypersocial behavior and social investigation (Nieh et al. 2016), and to modulate social defensive behaviors (Li et al. 2018; Rangel et al. 2016). Further, the LHplays a critical role in energy balance homeostasis (Burton, Rolls, and Mora 1976)—demonstrating the capacity to serve as a homeostatic control center (Cannon 1929). On the basis of the conceptual framework for social homeostasis, after social information is detected and evaluated in a rank-dependent manner, it would be sent to a control center for comparison to a social homeostatic set point (Lee, Chen, and Tye 2021; Matthews and Tye 2019).

We also investigated the mPFC projection to the basolateral amygdala (BLA) because recent evidence suggests that BLA firing rates correlate with the social rank of conspecific faces in non-human primates (Munuera, Rigotti, and Salzman 2018) and the BLA is an important point of convergence for socially derived information (Allsop et al. 2018) to be associated with emotional valence (Allsop et al. 2018; Felix-Ortiz et al. 2016; Janak and Tye 2015).

Figure 2-4 mPFC–LH pathway encodes social competition and modulates social dominance behavior. a, Left, viral strategy to stimulate projectors.

Top middle and right, mPFC projector cells (ChrimsonR+; red) and neighbor cells (grey) responding to pulses of red light. Top scale, 40 ms and 20 mV; bottom scale, 40 ms and 4 mV. Bottom middle and right, average photo-latency of action potentials (APs) or excitatory postsynaptic potentials (EPSPs) for mPFC–BLA (left plot) and mPFC–LH (right plot) ChrimsonR+ cells compared to neighboring cells (mPFC–BLA ChrimsonR+ n = 12 versus neighbours n = 10; t-test, ****P < 0.0001; mPFC–LH ChrimsonR+ n = 9 versus neighbours n = 8; t-test, ***P< 0.001). Dashed lines indicate photoresponse latency threshold used for phototagging projector. AAV5, adeno-associated virus serotype 5; DIO, double-floxed inverted open reading frame. b, Left, firing rate of mPFC–LH is higher than that of mPFC–BLA during the reward delivery (0–2 s) in win trials (mPFC–BLA n = 10 neurons, mPFC–LH n = 42 neurons, Wilcoxon rank sum P = 0.015). Right, firing rate for projector populations during tones for the reward task alone (alone data: mPFC–BLA n = 5 neurons, mPFC–LH n = 13 neurons, Wilcoxon rank sum P = 0.50; alone versus competition mean Z score during tone; mPFC–LH P = 0.0189, mPFC–BLA P = 0.43). Thicker lines represent the mean. c, Left, during reward competition light-off or light-on sessions in which light was delivered in epochs (5-min light epoch of four 5-ms light pulses at 100 Hz every 200 ms). Right, cumulative rewards obtained by ChR2 mice in the light-off versus light-on session (n = 9 mice). Data are presented as mean values ± s.e.m. d, mPFC–LH cell stimulation increased the number of trials won (left; ChR2 n = 9, eYFP n = 7; two-way repeated measures ANOVA interaction of virus and light $F_{1,14} = 5.22$, P = 0.03; Bonferroni-corrected t-test ChR2 P = 0.01), time spent occupying the reward port (middle; ChR2 n = 9, eYFP n = 7; two-way repeated measures ANOVA interaction of virus and light $F_{1,14} = 6.73$, P = 0.02; Bonferroni-corrected t-test, ChR2 P = 0.02) and decreased time spent being displaced (right; ChR2 n = 9 paired t-test, P = 0.005; eYFP n = 7 paired t-test, P = 0.28). *P < 0.05, **P < 0.01.


a
AAV5-DIO-ChrimsonR
mPFC
BLA or LH
CAV2-Cre
mPFC-BLA
mPFC-LH
ChrimsonR+
neighbors
****
***
APs
EPSPs
Mean latency to peak (ms)
b
Competition: win trials
Alone
**
Z-score
mPFC-LH n=43
mPFC-BLA n=10
mPFC-LH n=13
mPFC-BLA n=5
Time to cue (s)
c
AAV5-DIO-ChR2
mPFC-LH
ON
OFF
ON
OFF
10 s tone
Reward delivery
ChR2
Off session
On session
Cummulative rewards
Trials
d
ChR2
eYFP
Number of rewards
Tone port occupation
Time displaced (s)
Off
On
Off
On
Epoch

To identify mPFC cells that project monosynaptically to the LH or BLA, we used an intersectional viral strategy to express ChrimsonR in each projection, validated with ex vivo recordings (Fig. 2.4a and Sup. Fig. 2.12a, b). We then wirelessly recorded mPFC neural activity while animals were alone or competing and delivered red light pulses at the end of the competition session to photoidentify mPFC–LH or mPFC–BLA neurons. We found that mPFC–LH neurons had stronger excitation to reward delivery than mPFC–BLA neurons during reward competition, but not when performing the task alone (Fig. 2.4b and Supp. Fig. 2.12c).

Given the selective unmasking of a robust mPFC–LH response to the reward-predictive tone only in the context of social competition (Fig. 2.4b), we speculated that mPFC–LH neurons could modulate reward-related social competition. To directly test the hypothesis that mPFC–LH neurons have a causal relationship with social-dominance-related behavior, we expressed either channelrhodopsin 2 (ChR2) or enhanced yellow fluorescent protein (eYFP) in mPFC–LH neurons and implanted an optic fibre in the mPFC (Fig. 2.4c and Supp. Fig. 2.12d). ChR2-expressing mice won more rewards during the entire competition, had greater reward port occupation, and spent less time being displaced when they received optical stimulation (Fig. 2.4d). Stimulating mPFC–LH neurons did not affect reward-seeking behavior while performing the reward competition assay alone, feeding in the home cage, anxiety, sociability, place preference or general effort (Supp. Fig. 2.12e–k).

2.8 Conclusion

Together, these data demonstrate that the mPFC neural activity predicts future competitive success, can be decoded to predict both relative and absolute social rank,

and uses cortico-hypothalamic circuits to guide social competition behavior. Development of an ethologically relevant social competition task that incorporates a trial structure allowed us to reveal how related variables updated on different timescales might be parsed and represented separately. Indeed, social rank and competitive success representations occupied orthogonal activity spaces (Fig. 2.2a), which we speculate is an adaptive strategy that the PFC can use to parse related variables updated on different timescales.

Importantly, the way that mPFC ensembles encode behavior is dynamic, which suggests a model in which internal states influence how the mPFC modulates behavior, consistent with a role in flexibly guiding behavior. Our data demonstrate that cortico-hypothalamic circuits carry social rank information that could potentially modulate the many different neuropeptide- and hormone-expressing subpopulations in the hypothalamus to achieve behavioral modulation. Indeed, we speculate that the mPFC serves as a rank identification node that works in concert with the anterior cingulate cortex to function as a ‘detector’ to extract signals from social agents and that downstream projections to the hypothalamus may function as the detector node output to a social homeostatic ‘control center’, within a purported social homeostatic circuit (Lee, Chen, and Tye 2021).

This study not only unveils a number of technological advances that together provide a platform for the investigation of social hierarchies, but also begins to integrate pieces of evidence that together support the notion that there is a neural circuit for social homeostasis.

2.9 Appendix

2.9.1 Subjects

Adult (8-10 week old at the time of first experimental procedure), wild-type male C57/BL6J mice from the Jackson Laboratory (RRID: IMSR_JAX:000664) were used for all experiments described. Mice were housed with randomly assigned littermates in cages of 3-4, maintained under a 12-hour reverse light/dark cycle, and given access to food and water ad libitum except where indicated in experimental proceedings. Experiments were conducted during the light-off light cycle phase. Experiments conducted at MIT were approved by the MIT IACUC while experiments conducted at the Salk Institute were approved by the Salk Institute IACUC. Sample size was not predetermined and instead estimated based on similar studies (Wang et al. 2011; Zhou et al. 2017).

2.9.2 Surgery

Surgeries were performed aseptically using a digital small animal stereotaxic instrument (David Kopf Instruments). Anesthesia was induced with 5% isoflurane and mice maintained at 1%-3% isoflurane in the stereotactic frame, with a heating pad stabilizing body temperature. Immediately post-surgery, all animals were injected subcutaneously with 1.0mL saline for rehydration and moved to a clean cage. Animals were allowed a recovery period of ≥7 days prior to experimental procedures and resumption of tube test ranking.

A beveled 33-gauge microinjection needle was used for all injections, and virus was delivered at a rate of 100nL/min using a 1mL microsyringe (nanofil; WPI) with both a microsyringe pump (UMP3; WPI) and controller (Micro4; WPI). For all injections,

coordinates were measured relative to bregma. For the optogenetic stimulation experiment mPFC->LH experimental mice received two 200nL injections of AAV5-ef1-DIO-hCHR2-eYFP into the mPFC (AP: +1.8 ML: +0.3 DV: -2 and -2.5), an optical fiber in mPFC (AP +1.8 ML +0.3 DV: -2.25) and CAV2-Cre mixed AAV5-CAMKIIa-mCherry for visualization (3:1) into the LH (AP: -0.7 & -1.0 ML: 2.86 DV: -5.0) 20° vertical angle. Identical parameters were used for non-opsin control animals except the virus used for mPFC injections was AAV5-ef1-DIO-eYFP. For the phototagging experiments, we used a red-shifted excitatory opsin ChrimsonR for better light penetrance with the wireless devices3. Mice received two injections of 200 nL of AAV8-hSyn-flex-ChrimsonR-tdtomato into the mPFC (AP: +1.8 ML: +0.3 DV: -2 and -2.5) and CAV2-Cre into the LH (AP: -0.7 & -1.0 ML: 2.86 DV: -5.0 at 20° vertical angle) or into the BLA (AP:-1.4 ML: 3.28 DV:-5.0). The optical fiber for phototagging experiments was placed at an angle to allow tethering the laser while the logger was connected. The mPFC fiber was implanted in AP: 1.8, ML-1.5 DV -1.3 with a 45° vertical angle. In addition, a ground wire was implanted in the left hemisphere touching the brain surface. Electrode implants and fiber implantation for the phototagging experiment were done in a second surgery 4 weeks after viral injection. The electrode consisted of a bundle of 32 single nichrome wires of 25 um diameter that was implanted into mPFC (AP: 1.8, ML: +0.5, DV: -2.25). All fibers and electrodes were secured using C&B-Metabond® Quick Adhesive Luting Cement (Parkell) and several additional layers of dental cement (Ortho-Jet powder, Lang Dental) were also applied.

2.9.3 Food restriction

Subjects were fed ad libitum until initiation of experiments requiring food restriction (reward training). Prior to these experiments, subjects were weighed to the nearest tenth

of a gram to attain a baseline for free-feeding body weight. During experiments requiring restriction, subjects were weighed daily and food-restricted to a limit of 85% of this baseline weight. Restriction was attained through limiting daily access to food pellets, providing approximately 2.5-4g/animal per day or 3 hrs of unlimited access to food. Animals had free access to water throughout.

2.9.4 Tube testing

After at least 1-week post-surgery recovery, mice were handled for 15-min/day for two days to gain familiarity with experimenters and reduce stress during experiments. Tube training began after handling. The mouse tube test assay was adapted from Lindzey and colleagues (Lindzey, Winston, and Manosevitz 1961) and validated by Wang et al. (Wang et al. 2011) for social dominance in mice. Mice were individually trained to walk through the tube for at least two days, until they were comfortably walking across without resistance. For testing to determine rank, all mice in each cage were tested in a round robin design in a randomized order. For each pair, mice were released at opposite ends of the tube simultaneously. The mouse that either itself backed out or was pushed out from the end where it was released was designated as “loser/subordinate” whereas the other mouse was designated as “winner/dominant.” Social ranks obtained with the tube test were considered stable when obtaining the same results for 3 or more days in a row.

2.9.5 Reward training and competition

The reward competition assay was validated in unimplanted male mice (data presented in Fig. 1; n=8). Food restricted mice were trained individually in preparation for the reward competition. During training sessions, single mice were placed in a MED-PC chamber with an adapted 3D-printed reward port. Trials began with presentation of a red

LED located above the reward port and a 60dB, 5kHz tone, which both lasted for 10 seconds. A 10uL reward of vanilla Ensure® was delivered 2 seconds after tone onset. Inter-trial intervals (ITIs) lasted 60-100 seconds. A house light and white noise were on throughout the session. Training sessions consisted of at least 20 trials. Training consisted of 5-10 daily sessions for unimplanted mice, and 12-16 days for mice with recording devices. Training continued until all mice within a cage reached learning criteria of collecting the rewards with a mean latency of 5 seconds or less from tone onset. Reward port entries were detected through an electrical lickometer or infrared beam. After all mice in a cage attained learning criteria, we initiated reward competition. During all competition sessions, two mice were placed in one chamber. Mice competed against one cagemate per session. Competition sessions lasted 30-minutes (20 rewards) or 45minutes (30 rewards). During the reward competition behavior, was recorded via webcams (at 25 or 30 fps). A custom-made MATLAB GUI was used to score pushing, resisting and displacement by a trained observer. The mouse that occupied the port for 3 seconds surrounding the reward delivery, or the first mouse to the port after reward delivery was considered the winner of the trial. Trials where these criteria were not met and thus who got the reward was unclear were excluded from analysis. AlphaTracker was used to track the nose, ears and tail base of both mice during the competition. For AlphaTracker analysis only videos with the same camera angle and zoom were pooled.

For optogenetic reward competition experiments, 11 cages of 4 mice each were injected and implanted with fibers in two separate experiments (data presented in Fig. 4 and Supp. Fig 12). Cages were randomly assigned to experimental or control group. Scoring and analysis of behavior was done blinded to experimental group and rank when

possible. All mice were reward trained but only intermediate or subordinate mice were stimulated during competition. After all animals in each cage reached criteria, competition sessions began. After 3 days of robin round competition, pairs of mice were selected for optogenetic reward competition experiments that met the following criteria: stable tube ranks and stable winning/losing relationship in the reward competition for two consecutive days or more, and winner being the higher ranking animal. These select pairs performed the reward competition two more additional consecutive days: one day for a light OFF competition and the other for a light ON competition for the relative subordinate animal only. For the light ON competition, the subordinate animal received laser stimulation in alternating epochs of 5 minutes of bursting stimulation (four 5 ms pulses at 100 Hz every 200 ms) and 10 minutes of light off. Both animals were tethered to the patch cords during training prior to the competition. To prevent the two patch cords from tangling we designed a commutator to split the laser into two with a dual commutator and then into single commutators for each individual mouse.

2.9.6 Reward competition and “alone” recording experiments

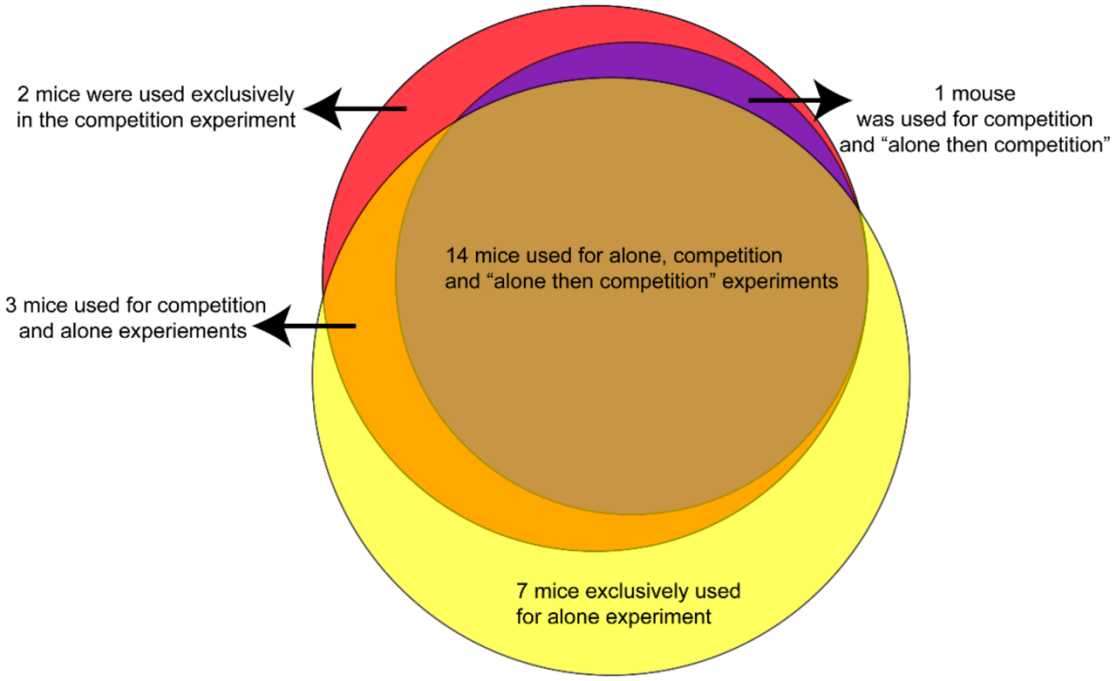
Wireless devices consisted in MiniLogger (SpikeGadgets) that recorded locally and digitized mPFC neural activity. During the reward competitions we recorded continuous signals from the 32 channels filtered for spikes (600-6,000 Hz) and a common average reference was applied using Trodes Software (SpikeGadgets). Tone times, port entries and laser pulses were detected with a main control unit that also triggered the miniloggers to record (Main Control Unit, SpikeGadgets). After recording sessions, logger data was extracted and merged with the Main Control Unit event related data. Spike

sorting was done with Klusta (Rossant et al. 2016) and subsequently verified by a trained observer for merging, deleting or splitting spike clusters.

For the mPFC recordings during the reward competition 24 new mice were implanted with electrodes in three batches of recording experiments. Out of those 24 mice only 20 mice had cells but the remaining 4 were still used for behavioral purposes as competitors. Additional animals were implanted (7 mice) and together with a subset of the mice used for competition (17 mice), had recording sessions after reaching training criteria while performing the task alone before ever having a social competition session (24 mice in total). Finally, in 15 out of the 20 mice used for competition recording we performed an additional experiment consisting in alone trials followed by competition trials (Supp. Fig. 2.10i-j) in the same recording session. For those recording sessions, the mouse started the reward task alone and the experimenter opened the chamber after 10 to 15 trials and added a competitor who had simultaneously received the same number of trials alone in a different chamber. For all experiments, several times recording sessions were excluded because of battery failures. Competition sessions during recording experiments consistent of 30 trials to maximize the amount of data obtained, both mice always wore loggers to match conditions across competitors. However, in 6 mice we utilized first generation loggers that had shorter battery life and therefore sessions for those animals consisted in 20 trials. Electrode-implanted mice were trained to wear a weight matched dummy headstage (mimicking the wireless logger weight) during their individual reward training. Individual reward training continued until mice reached the mean 5 sec latency criteria.

2.9.7 Reward competition behavioral analysis

If the port was occupied during reward delivery, winning was assigned based on who was at the reward port for seconds 1.5-3 post tone onset. If there was complete switch of who was at the reward port during the period of reward delivery the trial was considered complicated and not assigned as a win nor lose. If the port was unoccupied during the reward delivery winning and losing was assigned based on latency to enter



Supplemental Figure 2-1 Venn Diagram indicating overlap of mice recorded across experiments.

Most of the mice were used in all experiments, however, to consider absolute rank differences we added additional animals to the alone experiments.

port postdelivery of reward. Reward competition data was scored via two methods: semi-automated based on distance to the port tracked using AlphaTracker and confirmed by trained observer or using a custom-made Matlab GUI to score videos by a trained observer. Using the GUI we also scored pushing, resisting and displacing behavior, while relative location and distance to port were scored automatically via post processing of AlphaTracker tracking data. Pushing success was calculated as percent of pushing that

resulted in displacing competitor. For the latency to port logger vs no logger comparison (Supp. Fig. 2.7a), we used 12 animals that met the criteria of having reward alone sessions without logger within one week of reward alone sessions wearing loggers. Behavioral data across groups was compared using ANOVAs and t-test except when data was shown to be not normally distributed. For non-normal data we used Wilcoxon tests. Analyses in Supp. Fig. 2.3c-d and Supp. Fig. 2.9 only include animals with the same weight recording device (logger), excluding a subset of mice that were recorded with an older version logger, while those in Supp. Fig. 3e-g are paired analyses that include all mice.

2.9.8 Effort T-maze

The effort T-maze task was adapted from Walton et al. (2002) (Walton, Bannerman, and Rushworth 2002). Mice were trained on a T-maze (main arm: 40x8x15cm, side arms 30x8x15cm) to associate one side with a high reward (HR; 20µL droplet of 100% vanilla Ensure®) and the other side with a low reward (LR; 20µL droplet of 25% vanilla Ensure® mixed with water). The side with the high arm and low arm were counterbalanced across subjects. Training lasted nine days. Mice were first habituated to the T-maze apparatus with the cagemates where droplets of Ensure® were placed through the arms. The next four days, mice were individually given forced trials to sample the LR and HR arms and free choice trials to determine arm preference. Each training day always started with a forced sampling of both arms. After mice consistently showed preference for the high arm side, on day 5, a 7 cm wall was introduced in the HR arm. Mice were habituated to performing the task (including climbing the walls) while tethered to the patch cord. They were trained with forced and free choice trials with the 7 cm wall

and on day 8 the HR arm wall was replaced with a 15cm wall. After two days of training with the 15 cm wall, on day 10 mice received stimulation on 4 out of 8 trials, in blocks of two trials.

2.9.9 Three-chamber Sociability Test

This assay was adapted from Brodkin et al. (2014) (Brodkin et al. 2004). For each trial, the subject was placed in a test chamber (57.15 x 22.5 x 30.5 cm), with the empty central compartment connected to one additional compartment of equal size on both sides. The compartments on both sides contained a single perforated cup. One cup was empty, whereas the other held a conspecific (novel juvenile) underneath. The subject was allowed to explore the chamber for a period of 10-min. Each subject underwent two social interaction tests separated by 24-hrs, with one test paired with optical stimulation and one without. All groups were counterbalanced for both order of light stimulation and a different novel juvenile was used on both days. EthoVision XT video tracking system (Noldus, Netherlands) was used to record and quantify time spent within each chamber.

2.9.10 Open Field Test

Subjects were tested individually, and placed in the center of a transparent, plastic chamber (53 x 53 cm) divided into a central and a peripheral field. The test consisted of a 10-min period divided into five 2-min epochs (OFF-ON-OFF-ON-OFF) in which the mouse freely investigated the chamber. Animals were recorded with a video camera and tracked for the duration of the test using EthoVision XT video tracking system.

2.9.11 Conditioned Place Preference (CPP)

A test chamber (57.15 x 22.5 x 30.5 cm) with a monochromatic wallpaper of different patterns on each side was used. On day 1, each mouse was allowed 15-min free exploration of the apparatus. Mice with strong side preference (≥60%) were excluded from days 2-3. On day 2, mice received two (≥2-hr separation) 30-min conditioning sessions while confined to one side of the chamber. Mice received optogenetic stimulation in one side only, counterbalancing for order and side across animals. On day 3, mice were allowed to explore freely for 45-min, the first 10-min used for analysis. All sessions were EthoVision-tracked and video-recorded.

2.9.12 Food Assay

Each animal was individually placed in a clean cage along with a single food pellet for 30 min and stimulated in an epoch manner OFF-ON-OFF with 10 min epochs. Both animal and pellet were weighed before and after assay initiation. Sessions were EthoVision-tracked and recorded via side-webcam. All food assay experiments were done in the afternoon under dim lighting conditions (40 lux).

2.9.13 Optogenetic stimulation

Mice were stimulated with bursting 473nm light delivered in pulses of four 5 ms pulses at 100Hz in 5Hz intervals using a DPSS laser. Intensity was set to 10-11 mW at the end of each 200 nm patch cord (Thorlabs). For all experiments involving stimulation, all animals were tethered during training or habituation as well as testing.

For reward competition or alone sessions with phototagging, after the reward session, we tethered the patch cord to the fiber in mPFC and delivered 5 ms pulses at 10 or 20 Hz of red light (635 nm) at 8 mW (measured from patch cord). Neurons were identified as phototagged if they significantly increased firing rate with the photolatency

threshold established during the ex-vivo experiment. For the in vivo data included in the manuscript, 3 out of 8 mice had mPFC-BLA phototagged cells and 3 out of 12 mice had mPFC-LH phototagged cells.

To calculate if cells were phototagged, mean firing rate in the 10 ms pre to vs fist 10 ms post light pulse presentation was calculated. Next, the latency of excitation was considered the first post light onset millisecond bin in which the mean firing rate was 4 standard deviations higher than the mean firing rate of the 10 ms baseline period. A given cell was considered phototagged if the mean firing rate in the 10ms pre vs post was significantly excited based on a Wilcoxon sign rank test with alpha=0.001 and latency of excitation was 7 ms or shorter.

2.9.14 Histology

After experiments, mice were anesthetized deeply with sodium pentobarbital (200 mg/kg; intraperitoneal injection) and transcardially perfused with 10 mL of Ringer's solution followed by 10 mL of cold 4% PFA in 1xPBS. Mice were decapitated and brains were extracted, fixed in 4% PFA for ≥24-hr at 4°C, then transferred to 30% sucrose in 1xPBS until they sank to the bottom of the solution. Coronal sections were then cut at 50 um and stored in 1xPBS at 4°C until processing. Sections were added to 25mL of 1xPBS and 2uL of DAPI and placed on a rocker at room temperature for 10-min. Sections were then mounted on glass microscope slides, cover-slipped with PVA-DABCO, and stored at room temperature.

2.9.15 Electrophysiology data analysis

2.9.15.1 Neural trajectories

Principal component analysis (PCA) was conducted to evaluate the population-level firing rate dynamics (Cunningham and Yu 2014; Gao and Ganguli 2015; Linsenbardt, Timme, and Lapish 2019; Williamson et al. 2019). PCA is a commonly used dimensionality reduction tool that preserves the direction of largest variance in the data (Cunningham and Yu 2014). A single global PCA was done on a matrix containing all the data (concatenated mean firing rate per event type: win, lose, self port entry during tone, other port entry during tone, self port entry during ITI and other port entry during ITI) for all animals such that we could compare neural trajectories across groups (Linsenbardt, Timme, and Lapish 2019). This matrix had neurons in rows, and in the columns had mean firing rates during -5 to 5 seconds post task-relevant event using 100 ms bins. The neural trajectories for each task-relevant event were created per group by multiplying the coefficients obtained in the PCA by the mean firing rates across trials or port entries. For each trajectory, the geodesic length was calculated as the sum of Euclidean distances between adjacent 100 ms bins (Shifrin n.d.). Distance between trajectories was calculated as the Euclidean distance between the two trajectories bin-by-bin. To allow for statistical comparisons, the neural trajectory metrics were calculated using the leave one out (LOO) method, leaving out all the neurons from a single animal per group, thus the number of iterations is the number of mice in that group. The n reported in trajectory quantifications corresponds to the number of mice utilized for the LOO. Importantly, in every iteration the same PCA coefficients per cell were used for the neural trajectory, since the PCA was done once prior to this step, but the neurons included varied. For visualization purposes we plotted the first 2 or 3 PC subspaces, but for quantification of trajectory lengths and distance between trajectories the first 148 PCs were used to capture 90% of the variance.

For all trajectory visualizations and quantifications, we matched the number of neurons for each group to control for number of neurons being different across groups. For just the PC space visualizations, we smoothed the trajectories using MATLAB's smoothdata function. To compute the alignment, we used the following formulation as described by Miri et al. (Miri et al. 2017): say D_1 and D_2 are data matrices for the neural activity of neurons (rows being neurons and columns being time bins) under two different conditions (such as win vs lose trials), and P_1 and P_2 are the corresponding principal component vectors:

Definition 2.1: Alignment Index

$$Alignment\ Index\ := \frac{trace(P_1^T cov(D_2) P_1)}{trace(P_2^T cov(D_2) P_2)}$$

2.9.15.2 SVM classifier

To test if rank or competitive success could be decoded from single trial mPFC population activity, we used a support vector machine (SVM) with a Gaussian kernel. To obtain single trial mPFC population activity we used the coefficients obtained for each neuron in the global PCA and created a single trial neural trajectory using the firing rate for that trial. We trained the SVM using the first 8 PCs per trial and timepoints (-30 to 30 sec post cue or -5 to 5 sec post cue) as features and either rank or the competitive outcome of the trial as the label. We did a 10-fold cross-validation (CV), briefly, the data was split into 10 subsets and in each iteration the training consisted of a different 90% subset of the data and the testing was done with the remaining 10% of the data. We utilized the default hyperparameters as defined in MATLAB 2019b. Since the standard SVM is not probabilistic, we used an appropriate score-to-posterior-probability transformation function as proposed by Platt (Advances in Large-Margin Classifiers n.d.).

For the 10-fold CV, we computed the area under the receiver operating characteristic curve (AUC score) for the test data since it's sensitive to class imbalance. Further, we tried to maintain a class balance wherever possible (e.g. win vs lose trials used). We used this model for a variety of decoding tests to decode competitive success, relative rank and absolute rank. For the decoding of absolute social rank in Supp. Fig. 2.10, we used a one-vs-all strategy, where we generate one performance curve for each absolute rank by converting the multi-class problem into four binary problems (to test whether we can decode absolute rank 1 vs the rest, absolute rank 2 vs the rest and so on). Since we're using neural activity recorded from two mice competing in the task at a time to decode absolute social rank defined on a cage of four mice, we cannot be certain whether the decoding algorithms are predicting absolute social rank directly or are instead predicting a variable correlated to the absolute social rank, such as competition specific patterns in neural activity.

2.9.15.3 Firing rate analysis and hierarchical clustering for neural activity

To determine if a cell was significantly responding to an event, we compared the firing rate in a baseline period vs the event onset (5 second window for baseline and event) using a Wilcoxon sign rank test. If firing rate change was significant, excitation or inhibition was determined based on the average Z-score during the 5 second response window (if it was positive the cell was considered excited, while if it was negative the cell was considered inhibited). Note that the mean firing rate per trial type was smoothed using MATLAB's smoothdata function prior to z-score calculations for this analysis. Mean firing rate across events for the cue winning and losing trials, and port entries for self and

other (tone and ITI) were concatenated per cell and arranged across animals in a matrix where column was time and row was neuron number. Using custom made MATLAB code, this matrix was clustered using agglomerative hierarchical clustering based on Euclidean distance and a threshold of 0.4 was used to determined functional clusters. For the hierarchical clustering cells that had a firing rate lower than 0.05 Hz during the baseline period were excluded from the analysis given that cells with zero spikes in baseline could not be z-scored such that we only used 913/998 cells. In Fig. 3 we only display neurons with strong responses to events (firing rate during event larger or equal to 2 Z-scores or smaller or equal to -1 Z-scores).

2.9.16 AlphaTracker

2.9.16.1 Architecture of AlphaTracker

Inspired by similar tools like DLC and SLEAP we turned to deep learning for multi-animal tracking (Lauer et al. 2022; Pereira et al. 2022; Segalin et al. 2021), however, we took a top-down approach based on AlphaPose (Fang et al. 2017). First, a convolutional neural network is trained to predict the whole mouse location as a bounding box. For this, YOLOv349 was adapted to achieve high detection speed. Then the bounding box becomes the input of an additional convolutional neural network Squeeze-andExcitation Networks (SENet) (Hu, Shen, and Sun 2018) that tracks the body points of interest. Finally, to track animals across frames we use intersection over union (IOU). First, each animal's position and body parts are represented using a set of hierarchical bounding boxes. Then for two adjacent frames, we calculate the similarity of all descriptor pairs between these two frames according to equation 2.2 while the IOU is calculated using equation 2.3.

Definition 2.2: Descriptor Similarity

$$Sim(D_i^t, D_i^{t+1}) := IOU(Box_i^t, Box_i^{t+1}) + \frac{1}{n}\sum_{k=1}^{n} IOU(P_{i,k}^t, P_{j,k}^{t+1})$$

where D_i^t is the descriptor of animal i at frame t, Box_i^t is the bounding box of animal i at frame t predicted by the convolutional neural network, $P_{i,k}^t$ is the box that wraps the k^{th} body point of animal i at frame t.

Definition 2.3: Intersection Over Union

$$IOU(Box_1, Box_2) := \frac{Area\ Of\ Intersection\ (Box_1, Box_2)}{Area\ Of\ Union(Box_1, Box_2)}$$

The descriptor similarities are then sorted in descending order. The animal in the current frame and the animal in the previous frame that has the highest similarity are first matched and assigned with the same tracking ID. Then, for the remaining mice the ones with the second highest similarity are matched and so on until all animals in the current frame are assigned with a tracking ID or until no animals are left unmatched in the previous frame. Finally, a Kalman filter (Kalman 1960) is used to model the velocity and acceleration to predict the expected location of body points to try to overcome errors in cases with occlusion or tracking errors.

2.9.16.2 Evaluation of AlphaTracker

We randomly extract and labeled 800 frames from 4 different videos of two unmarked mice interacting in a clean cage, 200 of those frames were annotated by two people to determine human level accuracy. We split the data multiple times into non-overlapping sequences using different training-testing ratios (1:7, 1:3, 1:1 and 3:1) and trained 12 different networks with either 100,200, 400 or 600 frames (3 different splits for each ratio) and tested the tracking quality using the remaining frames. Furthermore, we

extracted 600 frames from a single video of four unmarked mice interacting in a clean cage, 200 of which were human-annotated to determine ground-truth values, and a subset of 100 were annotated by two different humans to calculate error across humans. We then trained 2 different networks with either 200 or 400 frames.

We quantified average precision (Andriluka et al. 2018; Pishchulin et al. 2016), the root mean square error (RSME) of the different networks, and how they compared to the precision and error across two humans. RMSE, in this case, was defined as the Euclidean distance of pixels between a ground truth point and a network-predicted point. Importantly, RMSE was done after identity error correction in a subset of the frames coming from the same video to quantify accuracy of keypoints and percent identity error in the same dataset. We excluded 3 frames in which AlphaTracker failed to find one of the mice. Specifically, for each instance (animal) in a frame of video, the Euclidean distance (pixels) was calculated for each target point (head, left ear, right ear, and tail). Then, the distances for each body part were averaged across all animals in the current frame (either two or four mice). Lastly, this was performed across all evaluation frames coming from the same video and a final average was calculated as the average RMSE per frame across all body parts.

2.9.16.3 Training and data processing

Reward competition videos were annotated using the open source annotation tool Sloth (https://github.com/cvhciKIT/sloth). AlphaTracker was trained with at least 200 annotated images for a given camera angle and chamber. After training, AlphaTracker, tracked competition videos and generated x,y coordinates for both mice that were imported into MATLAB for further analysis. A human annotated the x,y location of the

reward port for each video and that location was used to quantify distance to port and to normalize mouse locations for the 2D histogram of locations of the reward chamber.

2.9.16.4 Unsupervised clustering for behavioral motifs

AlphaTracker output was used to calculate features for behavioral motif clustering inspired by other tools (Wiltschko et al. 2015). We implemented two popular techniques; agglomerative hierarchical clustering to generate a dendrogram, and Uniform Manifold Approximation & Projection (UMAP) (McInnes, Healy, and Melville 2020) to generate a manifold. The hierarchical clustering is a bottom-up technique of building clusters based on similarity between data points without having to specify the number of clusters and for which results are presented in a dendogram. In our case the clustering was done based on the Euclidean distance of a combination of behavioral features (see below) for every 500 ms clip. UMAP is a dimensionality reduction technique which preserves local as well as global structure in the data and defines an embedding that can be easily visualized in two dimensions. The data processing for these techniques was as follows: First, the data was split into 500 ms clips. For each video clip, we extract the features of the individual mouse and for frames where mice were close to each other we also extract features of the interaction between mice from the pose estimation and tracking result as well as the original image. For each clip we normalized the data by setting the origin of coordinates of all the poses (left ear, right ear, nose and tailbase) to the nose position of the target mouse in the middle frame of the clip. For the features shown in Supp. Fig. 2.5 z-score normalization was applied using the mean across the whole clip. In addition, the direction of x axis of coordinates of all the poses was set to be the orientation of the target mouse's body vector in the middle frame of the clip. Fourier Transform was applied to the contour

images of the mice. For each mouse the following features were used: length of head, length of body, angle between head and body, ear to nose length for each ear, ear-body angle for each ear, displacement of the mouse, angle of the body across frames, Fourier transform of the displacement and Fourier transform of the contour of the mouse. For social clustering, the following additional features were used: distance between noses, angles between mice, distance between tail of one mouse and nose of the other. All angles, distances and displacement are reported using polar coordinates. All these features were then used to cluster the AlphaTracker output using hierarchical clustering or UMAP.

2.9.17 Ex vivo electrophysiological recordings and analysis

A separate cohort of mice (n= 6) was injected using the same viral strategy described above for unilateral ChrimsonR expression in mPFC-LH or mPFC-BLA neurons. Following 6 weeks of postoperative recovery (time matched for behavioral recordings), mice were anesthetized with sodium pentobarbital (200 mg/kg-1; intraperitoneal injection), decapitated and brains were quickly extracted in oxygenated (95% O2 and 5% CO2) high sucrose cutting solution containing (in mM): 87 NaCl, 2.5 KCl, 1.3 NaH2PO4, 7 MgCl2, 25 NHCO3, 75 sucrose, 5 ascorbate, and 0.5 CaCl2 (~320-300 mOsm; pH = 7.3). Coronal sections (300 μm) containing the mPFC and LH or BLA were collected using a vibrating blade microtome (VT120; Leica, Buffalo Grove, IL). Slices were then incubated in oxygenated, modified artificial cerebrospinal fluid (ACSF) containing (in mM): 75 sucrose, 87 NaCL, 2.5 KCl, 1.3 NaH2PO4, 7 MgCl2, 0.5 CaCl2, 25 NaHCO3 and 5 ascorbic acid (~300 mOsm; pH =7.3) in a 31°C water bath for 1 hour before recordings were obtained.

Viral injections sites in the mPFC and LH or BLA for mPFC-LH and mPFC-BLA neurons, respectively, were confirmed. Whole-cell patch-clamp recordings were obtained from visually identified neurons expressing ChrimsonR (ChrimsonR+) and adjacent non-expressing neighboring neurons. Recordings of mPFC-LH and mPFC-BLA neurons were obtained using glass pipettes with a pipette resistance of 3-5 MΩ fabricated with a horizontal puller (P-1000, Sutter, CA) and filled with internal solution containing (in mM): 125 potassium gluconate, 20 HEPES, 10 NaCl, 3 MgATP, and 8 biocytin (~290mOsm; pH = 7.3). Recorded signals were acquired, amplified, and digitized at 10kHz using a Multiclamp 700B amplifier, Digidata144, and pClamp10 software (Molecular Devices, Sunnyvale, CA). Neurons were voltage-clamped at -70mV and given a 1 second constant red (589nm, 10mW light power at tip of fiber) light pulse to confirm ChrimsonR-mediated inward photocurrents. Neurons were then given a 1 Hz train of 5 ms red light pulses for 10 pulses in current-clamp under resting membrane conditions to assess photo-response latencies. Data were analyzed using Clampfit software (Molecular Devices, Sunnyvale, CA. The average photo-response latency of action potentials (APs) and excitatory postsynaptic potentials (EPSP) were measured as the time from light onset to the peak AP or EPSP for the 10 pulses for each cell.

2.9.18 Decoding of behavior

2.9.18.1 Dataset

The dataset used for modeling consisted of mPFC multiunit firing rate (X) and behavioral states (Y) determined from AlphaTracker output along with annotation from trained experts. Multiunit firing rate activity was determined by applying a common reference average and determining spikes using 4 standard deviations from the mean.

Since behavioral conditions needed to be consistent across animals for the AlphaTracker tracking, we selected 13 animals with equivalent video conditions. Data from 13 mice was pooled and arranged as:

X = #trials x #channels x #time_stamps

Y = #trials x #time_stamps

The dataset consists of 965 trials, with spiking information from 32 channels across 70 timestamps of 200 ms intervals. For each time stamp, we have assigned one of 9 behavior labels, which served as the 'classes' for the classification problem (see Fig 1f). These classes were determined based on distance thresholds applied using AlphaTracker tracking output, and pushing, resisting and displaced were annotated by humans. The classification problem was to input the multi-unit activity sequence and obtain the corresponding behavior label sequence for the same trial.

We utilized five decoding models with increasing complexity, with the HMM-GLM being the most complex. The goal was to build a decoder by iteratively adding a richer feature space and complex relationships with the output behaviors. In each case, we have used a 10-fold cross-validation method, training with 90% of the trials and testing on the held-out 10% of trials and computed the area under the receiver operating characteristic curve (AUC scores).

For the HMM-GLM, in addition to the 10-fold cross-validation, we also performed leave-one-animal-out (LOO) cross-validation to test the model's performance in a more stringent situation, the results from which are summarized in Supp. Fig. 2.7j, k. In the LOO the model was trained with all but one held out mouse, and the model is trained on the held-out mouse, this was done 13 times, one per mouse.

2.9.18.2 Frequency Table (FT)

We constructed a table with the frequencies of the 9 behavioral labels normalized by the total occurrence of all of them. This helps generate fixed probability values for each of the behaviors, which is used as the likelihood for all time points, independent of the neural activity. This is the simplest possible classifier, solely based on the distribution of the behavior labels in the training set. The AUC scores for the FT model were close to chance (Fig. 1i).

2.9.18.3 Generalized Linear Model (GLM)

We next wanted to incorporate neural activity using a simple decoder. We used a Multinomial GLM with the logit as link function and with a 9-output softmax. The GLM takes the neural activity at each time point as input and produces a likelihood for the 9 behaviors as output. The AUC scores for the GLM model were significantly higher than chance, describing a relationship between the behavior labels and the neural activity (Fig 1i).

2.9.18.4 Support Vector Machine (SVM)

Our next strategy was to employ an SVM. Because the SVM with a Gaussian kernel was successful in decoding rank type as well as trial outcome, we utilized that. Since SVMs are traditionally binary classifiers, we utilized a one-vs-all strategy wherein for each of the 9 behavioral labels, we re-labeled the whole dataset as zero (not-that-label) or one (that-label) and then ran the model on the dataset with 10-fold cross-validation. The SVM takes the neural activity at each time point as input and produces a likelihood for the 9 behaviors as output. The AUC scores for the SVM model were close to chance (Fig 1i).

2.9.18.5 Ensemble of GLMs (9 GLMs)

We then utilized an ensemble of 9 Multinomial GLMs, each with the logit as the link function, for the 9 behaviors. Since this meant that each GLM is trained independently to solve a binary class problem instead of one GLM to solve a multi-class problem, it seemed like a natural extension of the standard GLM. The AUC scores for this model were comparable to that of the single multi-class GLM, and the additional complexity was not useful (Fig 1i).

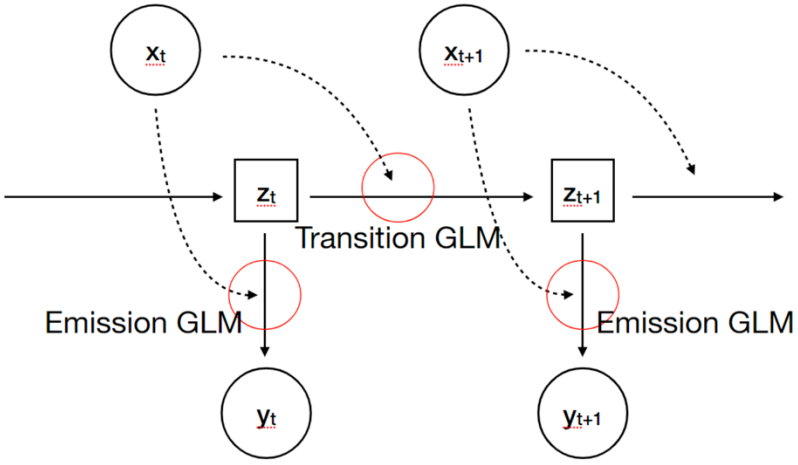
2.9.18.6 Hidden Markov Model – Generalized Linear Model (HMM-GLM)

We hypothesized that while the Multinomial GLM is able to partially capture mPFC dynamics, it still had two shortcomings. The first is the 'static' nature of the model and its inability to store past information when making predictions, which is key to common inference problems involving time series. The second is the simplistic nature of the model —which does not resemble the neural computation of the PFC —and hence is unable to utilize the statistical power of the novel trial structure used in this study. Moreover, the 9-class problem is a far more complex problem than a 2-class problem (rank or win/lose) with a much noisier dataset. This also explains the lack of improvement with 9-Multinomial-GLMs since it's still a linear and static model.

To incorporate temporal context into the model, we looked at the models proposed by Calhoun et al. (Calhoun, Pillow, and Murthy 2019) as well as Escola et al. (Escola et al. 2011) Calhoun et al. first introduced the idea of using a Hidden Markov Model (HMM) for modeling the change in the internal state of the organism as transitions of 'hidden states' and the neural spikes as emissions obtained from the hidden states. Escola and colleagues used Poisson GLMs for computing probabilities of each transition between

hidden states as well as for determining emissions of neural spikes. The input x_t for the transition and emission steps were the time-varying stimuli and the emitted observations y_t was the firing rate of the neurons. This approach provided a mechanism where both the state transition probabilities as well as neural spike emission probabilities varied with time. Calhoun et al. proposed a modification of the HMM-GLM in which both the input and output were behaviors, with a logit GLM instead of Poisson. While trying to decode the behavior of a male Drosophila during courtship rituals, they fed certain physical feedback cues that the male observed in the female into the GLM layers. These feedback cues from the past served as the input x_t and the current behavior was predicted via the emissions y_t. For our study, we modified the model further by having mPFC neural activity be the input to the GLM layers. The model formulation and training algorithm are similar to Escola et al. Note that the GLMs used in this model are binomial GLMs instead of multinomial. This is simply for the purpose of a simple interpretation for them. The primary differences are: 1) the input to the transition and state GLMs for the current time point is neural activity from the immediately previous time point only and not further back, 2) the GLM link function being a sigmoid/softmax has different gradients and dynamics as compared to the Poisson, 3) The outputs are categorical behaviors in our case instead of the binned spike counts and 4) we utilize an iterative optimization technique since the second order derivatives in our case are highly complex and computationally inefficient to use.

2.9.18.7 HMM-GLM Model Architecture & Assumptions



Supplemental Figure 2-2 Flow chart describing the interaction between variables in HMM-GLM.

Both the state transitions and output emissions follow the one-step Markov assumption:

Equation 2.1

$$P(z_{t+1}|z_1, z_2, \dots, z_t, y_1, y_2, \dots, y_t;\ x_1, x_2, \dots, x_t, \theta) = P(z_{t+1}|z_t; x_t, F_{tr})$$

Equation 2.2

$$P(y_t|z_1, z_2, \dots, z_t, y_1, y_2, \dots, y_{t-1}\ ;\ x_1, x_2, \dots, x_t, \theta) = P(y_t|z_t; x_t, F_{em})$$

Where $z_t \epsilon \{1, 2, \dots, K\}$ is the set of discrete hidden states, $y_t \epsilon \{1, 2, \dots, J\}$ is the set of categorical behavioral labels, θ is the set of all parameters $x_t \epsilon \mathbb{R}^D$ is the input at any time (T is the number of time stamps in a trial and D is the number of dimensions/channels of the input), F_{tr} is the set of transition probability parameters and F_{em} is the set of emission probability parameters.

The first assumption states that given the current hidden state, the future state is independent of both the past states as well as behavior outputs. The second assumption states that given the current hidden state, the current emission is independent of both the

past states as well as behavior outputs. Furthermore, the neural activity is treated as a deterministic variable since it is fixed for any given trial.

2.9.18.8 HMM-GLM Model Training

Let the data pertaining to one trial be concatenated into large matrices as:

$X = [x_1, x_2, x_3, \dots, x_T] \, \epsilon \, \mathbb{R}^{D \times T}$

$Y = [y_1, y_2, y_3, \dots, y_T] \epsilon [1, 2, \dots, J]^{1 \times T}$

$Z = [z_1, z_2, z_3, \dots, z_T] \epsilon [1, 2, \dots, K]^{1 \times T}$

where each array is either a single dimensional or multidimensional discrete time series or sequence.

Now, let us look at the parametrization of the model:

1) For state initialization:

Definition 2.4

$P(z_1 = k; \Pi) := \pi_k$

where: $\Pi = [\pi_1, \pi_2, \pi_3, \dots, \pi_K]$ represents the set of prior probabilities for all possible hidden states for the first time point in a sequence.

2) For state transitions:

Definition 2.5

$$\alpha^t_{z_{t-1}, z_t} := P(z_t = k | z_{t-1} = k'; F_{tr}) = \frac{\exp\left(F^T_{tr_{k',k}} \begin{bmatrix} x_t \\ 1 \end{bmatrix}\right)}{\sum_{k''=1}^{K} \exp\left(F^T_{tr_{k',k''}} \begin{bmatrix} x_t \\ 1 \end{bmatrix}\right)}$$

where: $F_{tr} \epsilon \mathbb{R}^{(D+1) \times K \times K}$ represents the array of filters (or weights) used for state transitions. For a transition from state m to state n, we use the weight vector $F_{tr_{m,n}}$. The data vector x_t is augmented with a '1' at the end to incorporate a bias.

3) For emissions:

Definition 2.6

$$\eta^{t}_{z_t,y_t} \coloneqq P(y_t = j | z_t = k;\ F_{em}) = \frac{\exp\left(F^{T}_{em_{k,j}}\left[\frac{x_t}{1}\right]\right)}{\sum_{j'=1}^{J}\exp\left(F^{T}_{em_{k,j'}}\left[\frac{x_t}{1}\right]\right)}$$

where: $F_{em} \epsilon \mathbb{R}^{(D+1)\times K\times J}$ represents the array of filters (or weights) used for output emissions. For an emission of output p from state q, we use the weight vector $F_{em_{q,p}}$

Thus, $\theta = \{\Pi, F_{tr}, F_{em}\}$ is the set of all parameters such that $|\theta| =$ total number of parameters $= K + (D+1)K^2 + (D+1)K$ scales linearly with the number of inputs and quadratically with the number of hidden states.

The introduction of the hidden states makes the estimation of the parameters an unsupervised learning process. We use an adapted version of the Baum-Welch implementation of the Expectation Maximization (E.M.) algorithm for this.

For one sample sequence, the complete data likelihood is described by:

Equation 2.3

$$P(Y, Z; \theta) = P(z_1; \theta)\prod_{t=2}^{T} P(z_t | z_{t-1}; \theta)\prod_{t=1}^{T} P(y_t | z_t; \theta)$$

$$= \pi_{z_1}\prod_{t=2}^{T}\alpha^{t}_{z_{t-1},z_t}\prod_{t=1}^{T}\eta^{t}_{z_t,y_t}$$

While $P(Y, Z; \theta)$ represents the "complete" likelihood, it is a misnomer since the hidden state sequence is not known. This then becomes the *marginal* likelihood instead. The E.M. algorithm helps us maximize the true likelihood, i.e., using the law of total probability, $P(Y; \theta) = \sum_{Z\epsilon Z'} P(Y, Z; \theta)$ via iteratively maximizing the marginal likelihood in a roundabout manner. Here, $Z' = [1, 2, \dots, K]^{1\times T}$ is the set of all possible hidden state sequences.

E-Step:

Definition 2.7

$Q(\theta; \theta^{old}) \coloneqq \mathbb{E}_{Z|Y;\theta^{old}}[\log L(\theta; Y, Z)]$

where $L(\theta; Y, Z) \coloneqq P(Y, Z; \theta)$ is the likelihood function.

In order to compute the objective function Q, known as the Baum Auxiliary Function, we utilize a pair of algorithms called the forward and backward algorithms. Note that we are referring to Q as the 'objective function' since it will be optimized in the M-Step of the E.M. algorithm.

Forward Algorithm:

Definition 2.8

$a_k(t) \coloneqq P(y_1, y_2, \dots, y_t, z_t = k; \theta)$

Thus:

$$a_k(1) = P(y_1, z_1 = k; \theta)$$
$$= P(z_1 = k; \theta) P(y_1 | z_1 = k; \theta)$$
$$= \pi_k \eta^t_{k,y_1}$$

And:

$$a_k(t) = \eta^t_{k,y_t} \sum_{k'=1}^{K} a_{k'}(t-1) \alpha^t_{k',k}$$

Backward Algorithm:

Definition 2.9

$b_k(t) \coloneqq P(y_{t+1}, y_{t+2}, \dots, y_T | z_t = k; \theta)$

Thus: $b_k(T) = 1$, since it is deterministic, and:

$$b_k(t) = \sum_{k'=1}^{K} \alpha_{k,k'}^{t+1} \eta_{k',y_{t+1}}^{t+1} b_{k'}(t+1)$$

Using these relationships, we can compute the single and consecutive pairwise marginal probabilities of the posterior distribution of the state sequence given the emission sequence as:

Definition 2.10

$$\gamma(z_t = k) \coloneqq P(z_t = k|Y;\theta)$$

$$= \frac{a_k(t)b_k(t)}{P(Y;\theta)}$$

$$= \frac{a_k(t)b_k(t)}{\sum_{k'=1}^{K} a_{k'}(t)b_{k'}(t)}$$

Definition 2.11

$$\zeta(z_t = k, z_{t+1} = k') \coloneqq P(z_t = k, z_{t+1} = k'|Y;\theta)$$

$$= \frac{a_k(t)\alpha_{k,k'}^{t}\eta_{k',y_{t+1}}^{t+1} b_{k'}(t+1)}{P(Y;\theta)}$$

$$= \frac{a_k(t)\alpha_{k,k'}^{t+1}\eta_{k',y_{t+1}}^{t+1} b_{k'}(t+1)}{\sum_{k''=1}^{K}\sum_{k'''}^{K} a_{k''}(t)\alpha_{k'',k'''}^{t+1}\eta_{k''',y_{t+1}}^{t+1} b_{k'}(t+1)}$$

We can now describe the whole posterior, $P(Z|Y;\theta)$, using these new definitions.

Returning to the likelihood function $L(\theta;Y,Z)$:

Equation 2.4

$$\log L(\theta;Y,Z) = \log\big(P(Y,Z;\theta)\big)$$

$$= \log \pi_{z1} + \sum_{t=2}^{T} \log \alpha_{z_{t-1},z_t}^{t} + \sum_{t=1}^{T} \log \eta_{z_t,y_t}^{t}$$

Plugging all the expressions into the objective function $Q(\theta;\theta^{old})$:

Equation 2.5

$$Q(\theta;\theta^{old}) = \mathbb{E}_{Z|Y;\theta^{old}}[\log L(\theta;Y,Z)]$$

$$= \sum_{Z\epsilon Z'} P(Z|Y;\ \theta^{old}) \log L(\theta;Y,Z)$$

$$= \sum_{k=1}^{K} \gamma^{old}(z_1 = k) \log \pi_k + \sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{k'=1}^{K} \zeta^{old}(z_{t-1} = k, z_t = k') \log \alpha_{k,k'}^{t}$$

$$+ \sum_{t=1}^{T}\sum_{k=1}^{K} \gamma^{old}(z_t = k) \log \eta_{k,y_t}^{t}$$

Up until this point, only one sample trial sequence has been considered. Say we have M number of trials. Then, under the i.i.d. (independent & identically distributed) assumption, the overall objective function becomes:

Equation 2.6

$$Q(\theta;\ \theta^{old}) = \sum_{m=1}^{M}\left(\sum_{k=1}^{K} \gamma^{m^{old}}(z_1^m = k) \log \pi_k \right.$$

$$+ \sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{k'=1}^{K} \zeta^{m^{old}}(z_{t-1}^m = k, z_t^m = k') \log \alpha_{k,k'}^{m^t}$$

$$\left. + \sum_{t=1}^{T}\sum_{k=1}^{K} \gamma^{m^{old}}(z_t^m = k) \log \eta_{k,y_t}^{m^t}\right)$$

M-Step:

This step involves the optimization of the objective function using an iterative scheme. While several methods can be used to achieve this, we resorted to gradient ascent. We initialize Π (the prior) to chance and other elements of θ with a Gaussian distribution of zero mean and small standard deviation ($\mu = 0, \sigma = 0.001$), and then iteratively update θ as:

Equation 2.7

$$\theta^{new} = \theta^{old} + \rho \nabla_{\theta} Q|_{\theta=\theta^{old}}$$

where ρ is the learning rate.

In order to set a dynamic learning rate, we utilize a technique called Learning Rate Annealing, which is expressed as $\rho = \frac{\rho_0}{1+\frac{ep}{T_0}}$, where ρ_0 is the initial learning rate, ep is the number of epochs that have passed and T_0 is the annealing temperature. Thus, as more training epochs pass, the learning rate becomes smaller.

Before we can compute the gradients, we need to incorporate the Gaussian prior into our objective function. This would help regularize the expression for the assumption of the Gaussian distribution for our parameters.

Equation 2.8

$$\begin{aligned} Q(\theta;\, \theta^{old}) = & \sum_{m=1}^{M} \left(\sum_{k=1}^{K} \gamma^{m^{old}}(z_1^m = k) \log \pi_k \right. \\ & + \sum_{t=2}^{T} \sum_{k=1}^{K} \sum_{k'=1}^{K} \zeta^{m^{old}}(z_{t-1}^m = k, z_t^m = k') \log \alpha_{k,k'}^{m^t} \\ & \left. + \sum_{t=1}^{T} \sum_{k=1}^{K} \gamma^{m^{old}}(z_t^m = k) \log \eta_{k,y_t}^{m^t} \right) \\ & - \frac{1}{2\sigma^2} \left(\sum_{k=1}^{K} \pi_k^2 + \sum_{k=1}^{K} \sum_{k'=1}^{K} {F_{tr}}_{c,c'}^{T} F_{tr_{c,c'}} + \sum_{j=1}^{J} \sum_{k=1}^{K} {F_{em}}_{c,j}^{T} F_{em_{c,j}} \right) \end{aligned}$$

Finally, let us look at the gradients:

$\forall r \epsilon \{1, 2, \dots, K\}$, $\forall s \epsilon \{1, 2, \dots, K\}$ and $\forall l \epsilon \{1, 2, \dots, J\}$

Equation 2.9

$$\nabla_{\pi_r} Q(\theta;\, \theta^{old})$$

$$= \sum_{m=1}^{M}\sum_{k=1}^{K}\left(\frac{\delta_{c,r}}{\pi_k}\gamma^{m^{old}}(z_1^m = k)\right) - \frac{1}{\sigma^2}\sum_{k=1}^{K}\delta_{r,k}\pi_k$$

$$= \sum_{m=1}^{M}\frac{1}{\pi_r}\gamma^{m^{old}}(z_1^m = r) - \frac{1}{\sigma^2}\pi_r$$

Equation 2.10

$\nabla_{F_{tr_{r,s}}} Q(\theta;\ \theta^{old})$

$$= \sum_{m=1}^{M}\sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\delta_{k,r}\left(\delta_{k',s} - \frac{\sum_{k''=1}^{K}\exp\left(F^T_{tr_{k,k''}}\left[\frac{x_t^m}{1}\right]\right)\delta_{k'',s}}{\sum_{k''=1}^{K}\exp\left(F^T_{tr_{k,k''}}\left[\frac{x_t^m}{1}\right]\right)}\right)\zeta^{m^{old}}(z_{t-1}^m k, z_t^m\right.$$

$$\left.= k')\left[\frac{x_t^m}{1}\right]\right) - \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\delta_{k,r}\delta_{k',s}F_{tr_{k,k'}}$$

$$= \sum_{m=1}^{M}\sum_{t=2}^{T}\left(\zeta^{m^{old}}(z_{t-1}^m = r, z_t^m = s) - \alpha_{r,s}^{m^t}\gamma^{m^{old}}(z_{t-1}^m = r)\right)\left[\frac{x_t^m}{1}\right] - \frac{1}{\sigma^2}F_{tr_{r,s}}$$

Equation 2.11

$\nabla_{F_{em_{r,l}}} Q(\theta;\theta^{old})$

$$= \sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{K}\left(\delta_{k,r}\left(\delta_{y_t^m,l} - \frac{\sum_{j=1}^{J}\exp\left(F^T_{em_{k,j}}\left[\frac{x_t^m}{1}\right]\right)\delta_{j,l}}{\sum_{j=1}^{J}\exp\left(F^T_{em_{k,j}}\left[\frac{x_t^m}{1}\right]\right)}\right)\gamma^{m^{old}}(z_t^m = k)\left[\frac{x_t^m}{1}\right]\right)$$

$$- \frac{1}{\sigma^2}\sum_{j=1}^{J}\sum_{k=1}^{K}\delta_{k,r}\delta_{j,l}F_{em_{k,j}}$$

$$= \sum_{m=1}^{M}\sum_{t=1}^{T}\left(\left(\delta_{y_t^m,l} - \eta_{r,l}^{m^t}\right)\gamma^{m^{old}}(z_t^m = r)\left[\frac{x_t^m}{1}\right]\right) - \frac{1}{\sigma^2}F_{em_{c,j}}$$

where $\delta_{i,j} = 1\ if\ i = j\ and\ 0\ otherwise$ is the Kronecker delta.

Since a closed-form solution for these updates would be very complicated to obtain, we relied on numerical computing to optimize the objective function and train the model.

We followed the tips laid out by Rabiner (Rabiner 1989) to deal with problems of numerical overflow/underflow while computing the forward and backward steps of the training algorithm. We also binned all probability distributions to at least 10^{-300}, to prevent numerical errors that MATLAB's precision levels could not handle.

2.9.18.9 Cross-validation & Model Selection

Since we had to manually select the number of hidden states to be used, we resorted to a 10-fold cross validation mechanism to remove any biases in the process. In each fold, we trained the model on 9/10ths of the data, and then computed the likelihood on the 10th held-out part. We then plotted the log-likelihoods (trial average) of the held-out data from all the folds as a function of number of hidden states with standard error of the mean as error bars (Supp. Fig. 7j). We found that as you increase the number of hidden states, the log likelihood increases, reaching a plateau at the 6-state model.

2.9.18.10 Testing

Identifying the most probable hidden state sequence

Since $P(z_t = k|Y;\theta) = \gamma(z_t = k)$, then for a given trial, $z_t^* = \arg\max_k \gamma(z_t = k)$ is the most probable hidden state for time t.

Predicting the behavior of mice

We utilized the one-step-prediction method to decode the behavior of mice from neural activity. In short, the information from prior time point was indirectly used together with neural activity of the current time point to predict the behavioral label. For the time

points of a given trial the most probable hidden state of the previous time point was stored. This helped compute the hidden state distribution for each time point.

For $t = 1$, there is no hidden state prior to this time. However, the distribution of hidden state is given by Π.

$\forall\, t \,\epsilon\, \{2, 3, \dots, T\}$, say the most probable hidden state for the previous time point is given by z_{t-1}^{*}. Now:

Equation 2.12

$$P(y_t|y_1, y_2, \dots, y_{t-1};\ x_1, x_2, \dots, x_t, \theta\ )$$

$$= \sum_{k=1}^{K} P(y_t, z_t = k|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

$$= \sum_{k=1}^{K} P(y_t|y_1, y_2, \dots, y_{t-1}, z_t = k; x_1, x_2, \dots, x_t, \theta)\, P(z_t = k|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

$$= \sum_{k=1}^{K} P(y_t|z_t = k; x_t, \theta) \sum_{k'=1}^{K} P(z_t = k, z_{t-1} = k'|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

$$= \sum_{k=1}^{K} P(y_t|z_t = k; x_t, \theta) \sum_{k'=1}^{K} P(z_t = k|y_1, y_2, \dots, y_{t-1}, z_{t-1} = k'; x_1, x_2, \dots, x_t, \theta) P(z_{t-1}$$

$$= k'|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

$$= \sum_{k=1}^{K} P(y_t|z_t = k; x_t, \theta) \sum_{k'=1}^{K} P(z_t = k|z_{t-1}$$

$$= k'; x_{t-1}, \theta) P(z_{t-1} = k'|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

Since z_{t-1}^{*} is the most probable value for z_{t-1}, the probabilities $P(z_{t-1} \neq z_{t-1}^{*}|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta\ ) = 0$ while that for $z_{t-1} = z_{t-1}^{*}$ would be unity.

Thus:

Equation 2.13

$$P(y_t = j|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta)$$

$$= \sum_{k=1}^{K} P(y_t = j|z_t = k; x_t, \theta) P(z_t = k|z_{t-1} = z_{t-1}^*; x_{t-1}, \theta)$$

This is how we generated the likelihoods for all the behavior labels for time t. The behavior label with the highest likelihood was selected as the most probable behavior for time t, i.e.,

Equation 2.14

$$y_t^* = \arg\max_j P(y_t = j|y_1, y_2, \dots, y_{t-1}; x_1, x_2, \dots, x_t, \theta).$$

Further we selected the state that leads to the highest likelihood for the correct behavior as compared with the ground truth as the most probable hidden state for time t, i.e.,

Equation 2.15

$$z_t^* = \arg\max_k P(y_t = y_t^{true}|z_t = k; x_t, \theta) P(z_t = k|z_{t-1} = z_{t-1}^*; x_{t-1}, \theta)$$

In this way we obtain the likelihoods for the behavior labelfor all the time points of a trial in an inductive manner.

Lastly, we observed a time lag of one time bin (200 ms) in the model output predictions. Since this was present uniformly across the entire dataset, we simply shifted the model output backwards in time by one bin.

2.9.19 Data and code availability

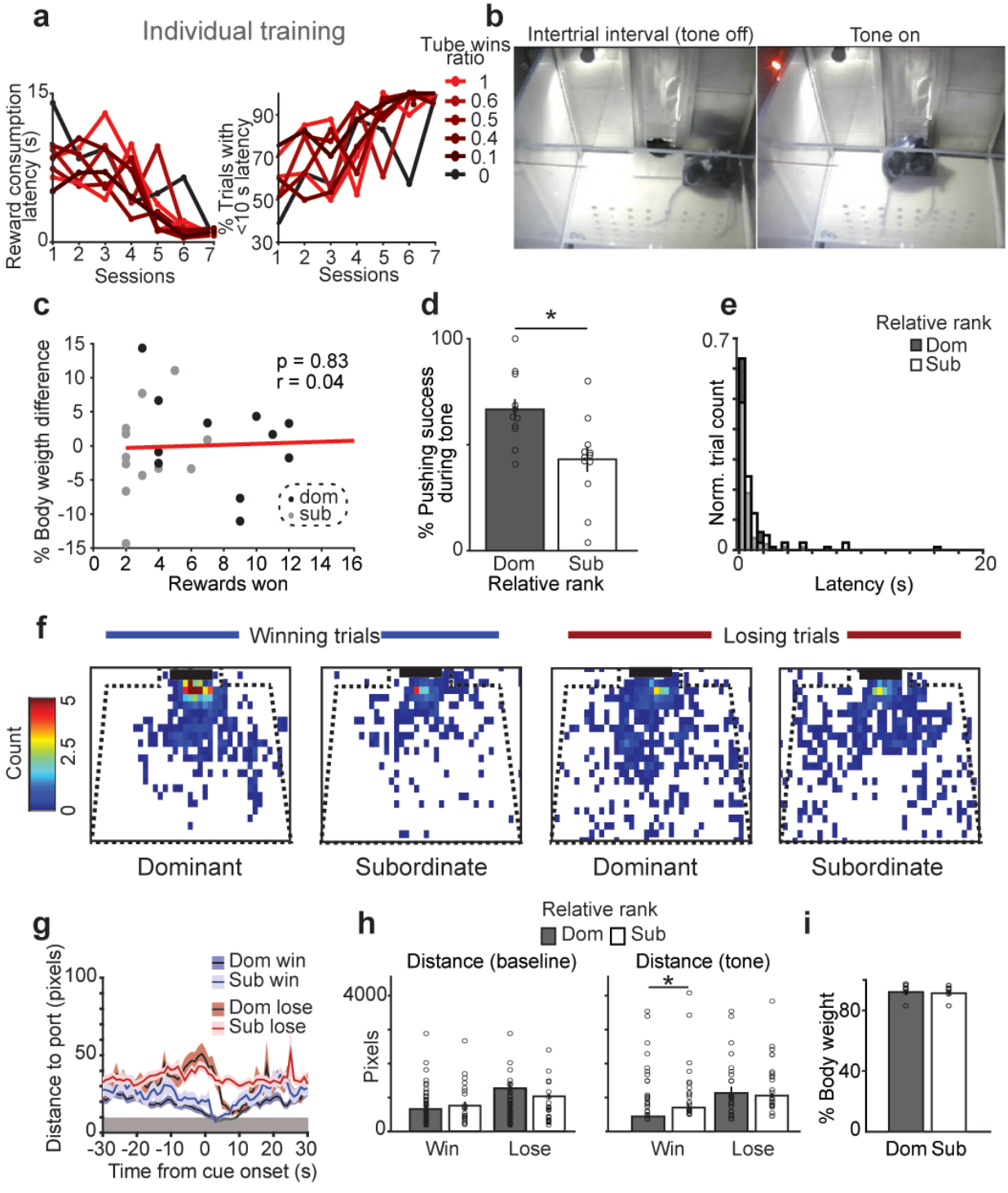
Source data are provided with this paper. Data will be made available upon reasonable request to the corresponding authors. Source data are provided with this

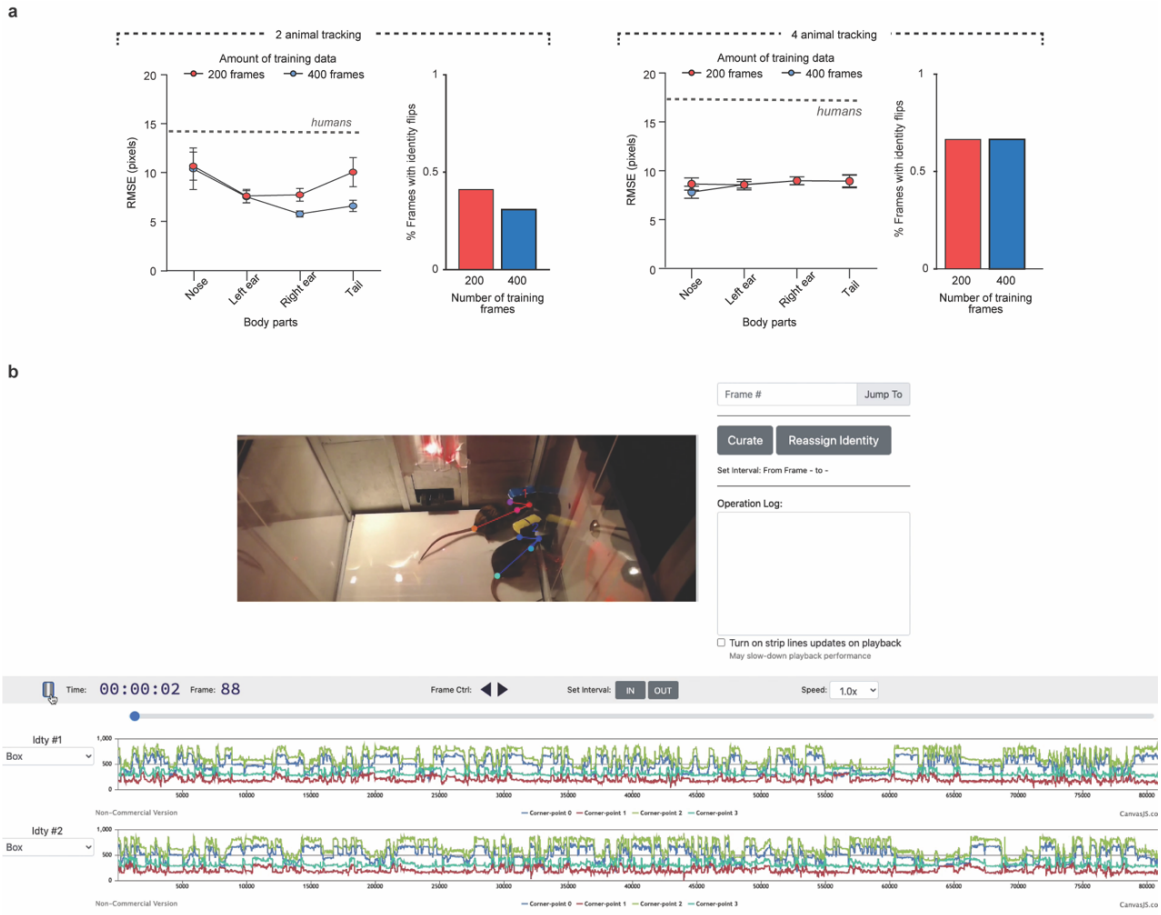
paper. Code for AlphaTracker can be found at https://tyelab.org/tools/ and code for the HMM–GLM model can be found at https://github.com/ Tyelab/HMMGLM.

2.10 Supplemental figures

Supplemental Figure 2-3 Additional behavioral metrics during reward competition in unimplanted mice.

a, Mice learned the tone reward association at the same rate across social ranks. Left, latency to reward consumption following tone onset decreased over sessions. Right, the percent of trials with a reward consumption latency of less than 10 s increased over sessions (n = 8 mice). Data are plotted as a function of social rank as measured by wins in the tube test. b, Example frames from reward competition assay showing intertrial interval time and during the tone period. c, Body weight difference between competitors does not correlate with rewards won (n = 12 dyads from 8 mice, Pearson's correlation, p = 0.83). d, Relative dominant mice have higher pushing success during the tone (n = 12 per group, paired t-test p = 0.025). e, Latency to pick up the reward across trials for relative dominants vs subordinates (dom=101, sub = 41, Two-sample Kolmogorov-Smirnov test p = 0.29). f, Area occupied by dominants or subordinates in the 10 s prior to the tone onset for win vs lose trials (n = 12 dyads). g, Distance to reward port across time by trial type and relative rank (trials n: dom win=68, dom lose = 24, sub win = 24, sub lose = 68 from 12 dyads; early baseline −30 to −20 s prior to cue there is no effect of trial nor relative rank; 2-way ANOVA using the mean distance from −5 s to cue onset: main effect of trial type $F(1,180) = 44.4$, $p = 3x10{-}10$, rank p = 0.94, interaction p=0.09; 2-way ANOVA using the mean distance from 5 s prior to tone until 10 s post tone: main effect of trial type $F(1,180) = 68$, $p = 2.5x10{-}14$, rank p = 0.071, interaction p = 0.79). Gray rectangle indicates contact range for the reward port. h, Total distance traveled immediately before the tone and during the tone period (baseline:10 s prior to tone; tone: 10 s of the tone) across trial types for relative dominant and subordinate mice (dom win = 68, dom lose = 24, sub win = 24, sub lose = 68 from 12 dyads; Wilcoxon rank-sum, baseline win p = 0.79, baseline lose p = 0.59, tone win p = 0.028, tone lose p = 0.86). Gray zone indicates contact with port. i, Percent body weight during food restriction did not differ across relative dominant and subordinate mice (n = 12 dyads, paired t-test, p = 0.23).


a
Individual training
Reward consumption latency (s)
Sessions
% Trials with <10 s latency
Tube wins ratio
1
0.6
0.5
0.4
0.1
0
b
Intertrial interval (tone off)
Tone on
c
p = 0.83
r = 0.04
% Body weigth difference
Rewards won
dom
sub
d
% Pushing success during tone
Dom
Sub
Relative rank
e
Relative rank
Dom
Sub
Norm. trial count
Latency (s)
f
Winning trials
Losing trials
Count
Dominant
Subordinate
Dominant
Subordinate
g
Dom win
Sub win
Dom lose
Sub lose
Distance to port (pixels)
Time from cue onset (s)
h
Relative rank
Dom
Sub
Distance (baseline)
Distance (tone)
Pixels
Win
Lose
Win
Lose
i
% Body weight
Dom
Sub

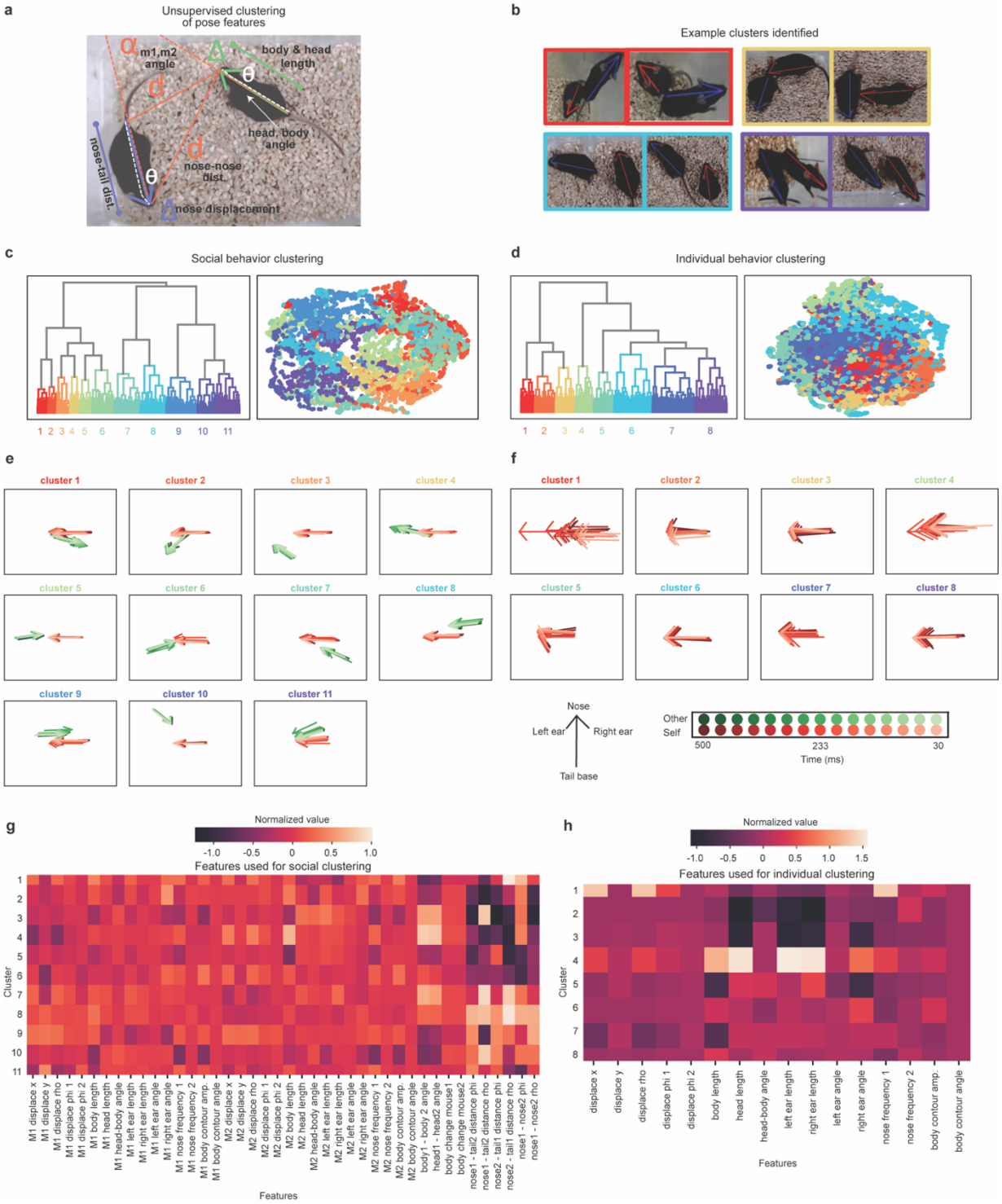


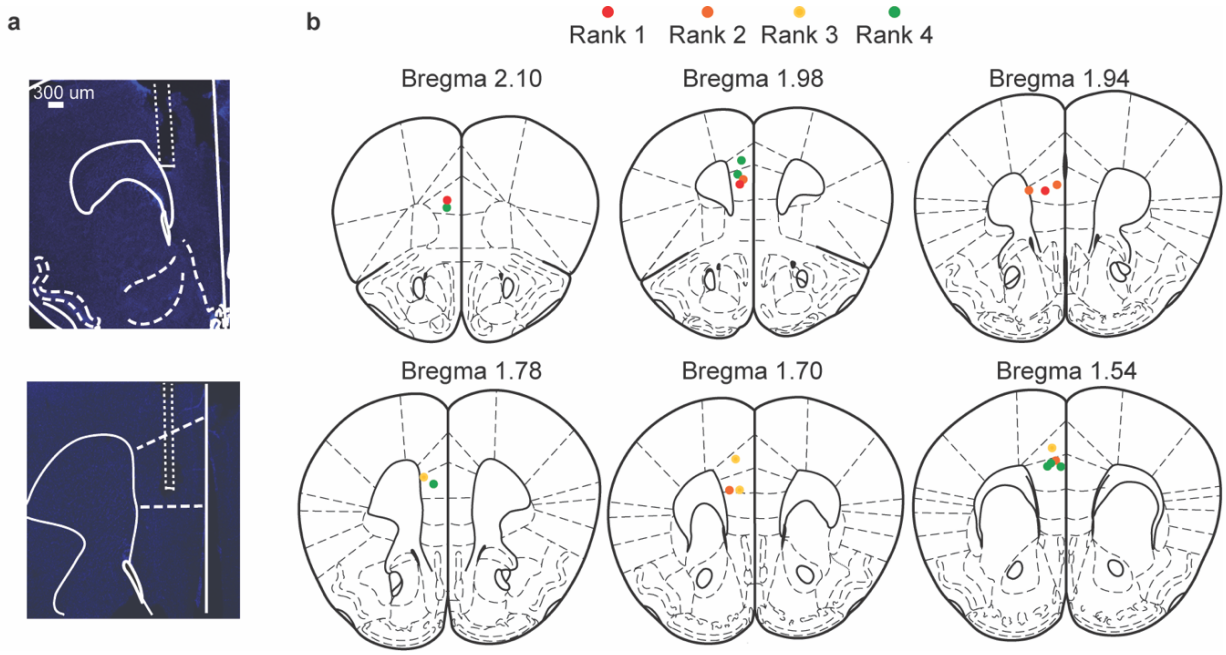
Supplemental Figure 2-4 AlphaTracker tracking metrics.

a, Root mean square error (RMSE) and identity error rate of AlphaTracker when tracking different body parts in videos with high resolution (1920x1080 pixels). Left plots have training and tracking done on 2 unmarked mice videos and right plots have training and tracking done on 4 unmarked mice videos. For both datasets two humans annotated the data and the RMSE between humans is indicated with the dashed line. For identity error rate 2 mouse tracking done with 9737 frame video and 4 mouse tracking done with 6020 frame video. b, Screenshot of user interface (UI) to fix errors made by AlphaTracker tracking. In addition, this UI can be used for exploring the clustering data.

Supplemental Figure 2-5 AlphaTracker unsupervised clustering results.

a, Diagram depicting features used for AlphaTracker's unsupervised clustering of the tracking datapoints. a, The features include head length, body length, body-head angle, displacement of the nose, distance between mice, angle between mice. b, Example frames from clips belonging to a specific cluster (cluster ID indicated with the color outline in c). c, Dendrogram and UMAP plot showing all video clips color coded by cluster ID for social behavior clustering. The mean cluster outputs are shown in (e) and features used are shown in (g). d, Dendrogram and UMAP plot showing all video clips color coded by cluster ID for individual behavior clustering. The mean cluster outputs for this clustering are shown in (f) and features used are shown in (h). e, Average normalized skeleton for nose, ears and tail base across clusters for the social behavior clustering across 500 ms of video clip time. Red arrow indicates self skeleton and green indicates the other skeleton. Each arrow represents 33.3 ms of data (1 frame). f, Average normalized skeleton for nose, ears and tail base across clusters for the individual behavior clustering across 500 ms of video clip time. Each arrow represents 33.3 ms of data (1 frame). Legend in bottom applies to panels e-f. g, Heatmap of normalized values for the self and other features used for social behavior clustering. h, Heatmap of normalized values for the self features used for individual behavior clustering.


a
Unsupervised clustering of pose features
m1,m2 angle
body & head length
head, body angle
nose-nose dist.
nose-tail dist.
nose displacement
b
Example clusters identified
c
Social behavior clustering
1 2 3 4 5 6 7 8 9 10 11
d
Individual behavior clustering
1 2 3 4 5 6 7 8
e
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
cluster 9
cluster 10
cluster 11
f
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
Nose
Left ear
Right ear
Tail base
Other
Self
500
233
30
Time (ms)
g
Normalized value
-1.0 -0.5 0 0.5 1.0
Features used for social clustering
Cluster
Features
M1 displace x
M1 displace y
M1 displace rho
M1 displace phi 1
M1 displace phi 2
M1 body length
M1 head length
M1 head-body angle
M1 left ear length
M1 right ear length
M1 left ear angle
M1 right ear angle
M1 nose frequency 1
M1 nose frequency 2
M1 body contour amp.
M1 body contour angle
M2 displace x
M2 displace y
M2 displace rho
M2 displace phi 1
M2 displace phi 2
M2 body length
M2 head length
M2 head-body angle
M2 left ear length
M2 right ear length
M2 left ear angle
M2 right ear angle
M2 nose frequency 1
M2 nose frequency 2
M2 body contour amp.
M2 body contour angle
body1 - body 2 angle
head1 - head2 angle
body change mouse1
body change mouse2
nose1 - tail2 distance phi
nose1 - tail2 distance rho
nose2 - tail1 distance phi
nose2 - tail1 distance rho
nose1 - nose2 phi
nose1 - nose2 rho
h
Normalized value
-1.0 -0.5 0 0.5 1.0 1.5
Features used for individual clustering
Cluster
Features
displace x
displace y
displace rho
displace phi 1
displace phi 2
body length
head length
head-body angle
left ear length
right ear length
left ear angle
right ear angle
nose frequency 1
nose frequency 2
body contour amp.
body contour angle

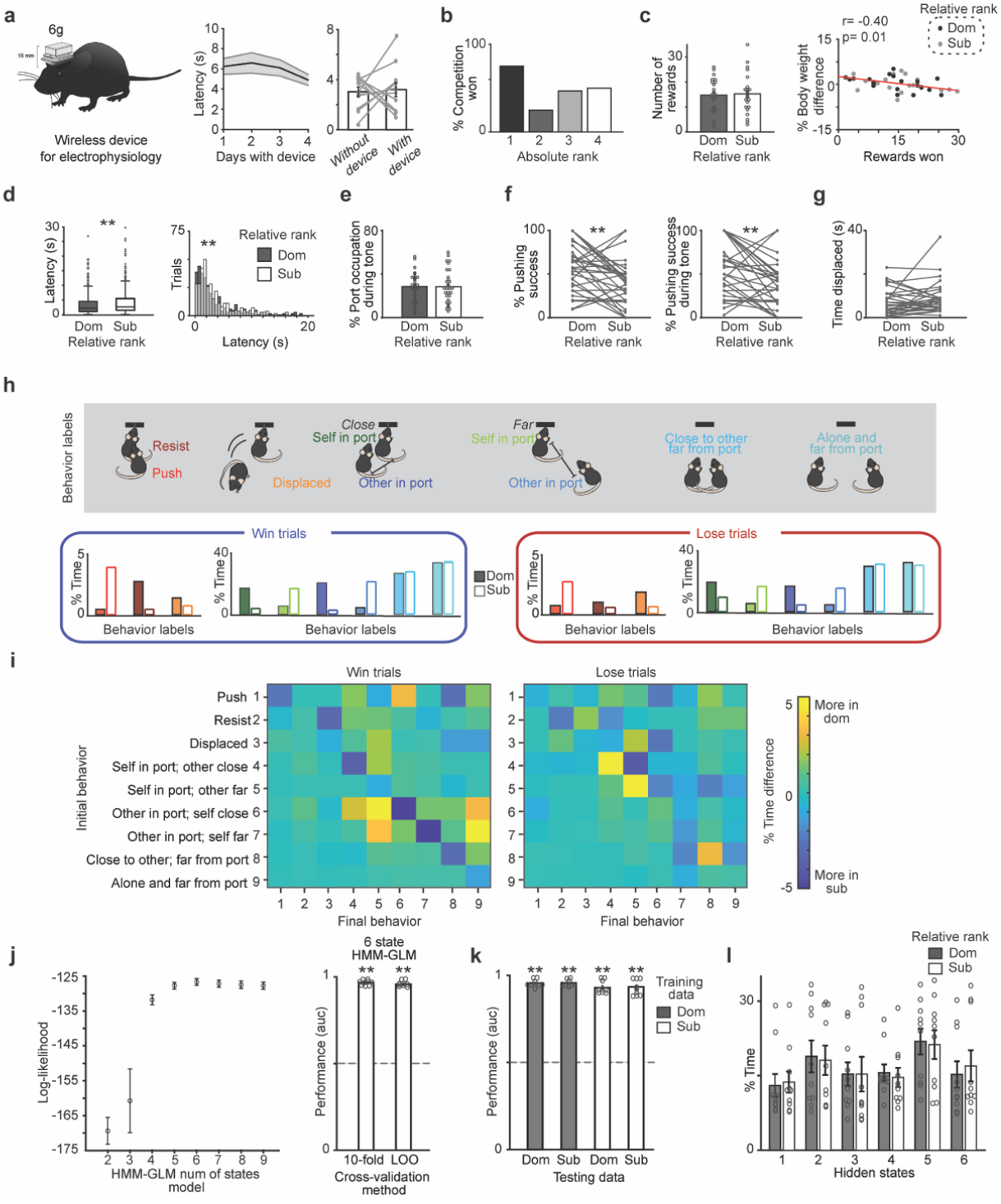


Supplemental Figure 2-6 Histological validation of electrode placements.

a, Representative images showing electrode track and lesions of mPFC electrode wires. b, Location of center for electrode lesions for all mice color coded by absolute rank across animals.

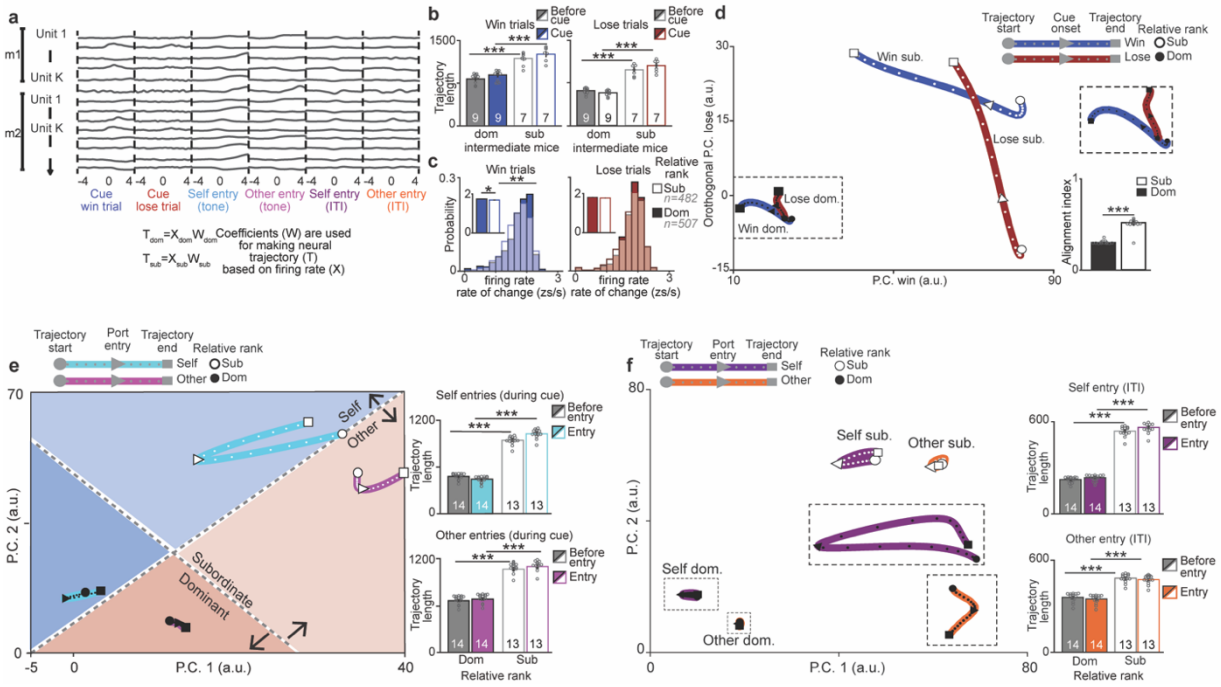
Supplemental Figure 2-7 Behavior in competition with logger and HMM-GLM model controls.

a, Left, diagram of wireless electrophysiology recording device used for mPFC recordings. Image modified from SpikeGadgets' MiniLogger product illustration. Middle, latency to collect reward over four days of training (n = 16 mice). Right, latency to collect reward while performing reward task alone was not affected by wearing the logger (n = 12 mice; paired t-test, p = 0.83). b, Percent competitions won by absolute rank is highest for rank 1 mice in dataset used for mPFC recordings (number of competitions per rank 1 n = 12; rank 2 n = 12; rank 3 n = 15; rank 4 n = 14). c, Left, number of rewards obtained by relative dominants (dom) and subordinates (sub) during the reward competitions between animals wearing loggers (n = 22 mice per group; paired t-test, p = 0.86). Right, % body weight difference between competitors significantly correlates with rewards won (sub n = 19 dom n = 20, Pearson's correlation, *p = 0.01). For correlation only mice with same day weight measurements were used. d, Subordinates had longer latencies to pick up the reward during win trials (center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers). Left, latency per group. Right, histogram of the distribution of latencies across all trials (dom trials n = 326, sub trials n = 358, Wilcoxon rank-sum, p = 0.012; Two-sample Kolmogorov-Smirnov test, dom vs sub trials p = 0.015; One way RM-ANOVA F(1,24) = 2.06, p = 0.002). e, Percent port occupancy during tone across relative rank (n = 31 sessions, paired t-test p = .90). f, Relative dominants were more successful displacing subordinates from the reward port throughout the competition (left; n = 32 sessions, paired t-test, p = 0.002) and during the tone time (right; n = 31 sessions, paired t-test, p = 0.005). g, Total time being displaced from reward port by relative rank in dataset used for mPFC recordings (n = 31 sessions; paired t-test p = 0.15). h, Percent time (normalized by total time per behavior) for 9 behaviors analyzed for win and lose trials separated by relative social rank. i, Percent time difference between relative dominant and subordinates for behavioral transitions during win trials (left) vs lose trials (right). j, Left, model selection for HMM-GLM state number using 10-fold cross-validation method results in a 6 state model being optimal. Error bars indicate standard error across the 10 cross-validations. Right, HMM-GLM 6 state model performance predicts behavioral label regardless of training method utilized (AUC n = 9, one per each behavior label; Sign test of model performance vs chance p = 0.004 for both methods). k, HMM-GLM 6 state model predicted behavioral label regardless of which dataset was used for training or testing (n = 9 behavior labels using 482 trials for dom vs 478 trials for sub; Sign test performance vs 0.5 (chance) p = 0.004 for all tests). l, Distribution of percent time spent in each hidden state by relative rank group (n = 10 cross-validations using 482 trials for dom vs 478 trials for sub from 14 mice).


a
6g
Wireless device
for electrophysiology
Latency (s)
Days with device
Without device
With device
b
% Competition won
Absolute rank
c
Number of rewards
Relative rank
r= -0.40
p= 0.01
Relative rank
Dom
Sub
% Body weight difference
Rewards won
d
Latency (s)
Relative rank
Trials
Latency (s)
e
% Port occupation during tone
f
% Pushing success
% Pushing success during tone
g
Time displaced (s)
h
Behavior labels
Resist
Push
Displaced
Close
Self in port
Other in port
Far
Self in port
Other in port
Close to other far from port
Alone and far from port
Win trials
Lose trials
% Time
Behavior labels
i
Win trials
Lose trials
Initial behavior
Push 1
Resist 2
Displaced 3
Self in port; other close 4
Self in port; other far 5
Other in port; self close 6
Other in port; self far 7
Close to other; far from port 8
Alone and far from port 9
Final behavior
% Time difference
More in dom
More in sub
j
Log-likelihood
HMM-GLM num of states model
6 state HMM-GLM
Performance (auc)
10-fold
LOO
Cross-validation method
k
Performance (auc)
Training data
Dom
Sub
Testing data
l
Relative rank
% Time
Hidden states

Supplemental Figure 2-8 Additional data for mPFC population dynamics during social competition.

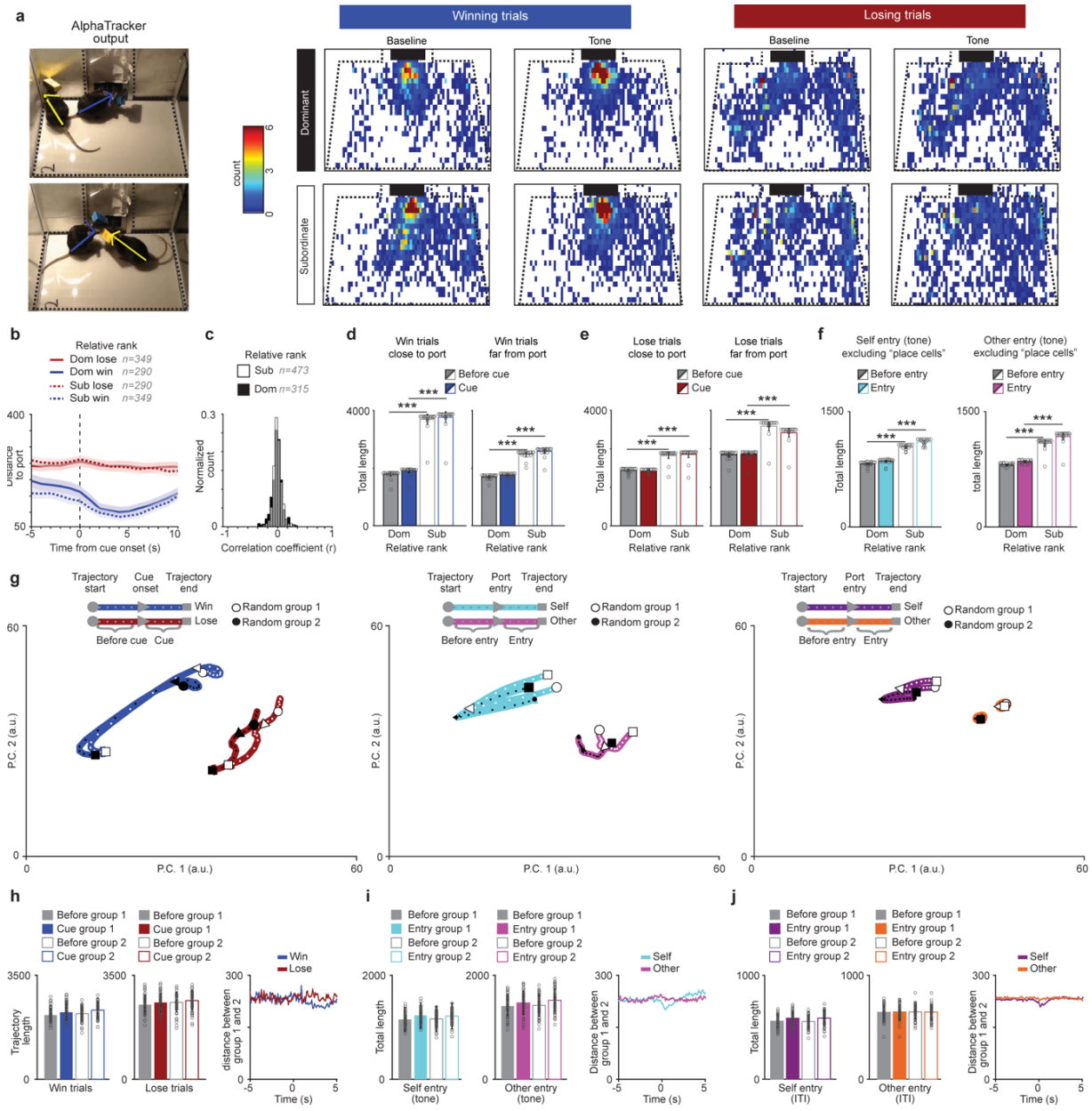
a, Data arrangement across all animals (m1 = mouse 1, m2 = mouse 2) for the dimensionality reduction to a common subspace for the six task-relevant events. Neural trajectories were created for dominant and subordinate data using mean firing rate per event and the principal component analysis coefficients. b, Neural trajectory lengths (using principal components that captured 90% of variance) for win and lose trials are longer for relative subordinates in intermediate (ranks 2 or 3) mice (n indicated on plots; win 2-way RM-ANOVA main effects of relative rank $F(1,14) = 165$, $p = 2x10{-}6$; lose 2-way RM-ANOVA effect of relative rank $F(1,14) = 262$, $p = 6x10{-}7$). c, Firing rate rate of change is higher for relative dominants only in win trials (number of neurons indicated in plots, inset plot has average across groups; win trials rate of change: Kolmogorov-Smirnov (KS) test $p = 0.009$, Wilcoxon rank sum $p = 0.01$; lose trials rate of change: KS test $p = 0.40$, Wilcoxon rank sum $p = 0.19$). d, Neural trajectories for win and lose trials plotted in the first Principal Component (PC) for win and the orthogonal lose subspace show little overlap. Top right, inset of dominant neural trajectories. Bottom right, alignment of win and lose trajectories was significantly lower for dominant mice ($n = 13$ per group; Wilcoxon rank-sum, $p = 1.5x10{-}5$). e, Left, neural trajectories of mPFC population firing rate differ by relative rank for port entries that occur during the tone period in a lower dimensional common principal component (PC) sub-space (trajectories are the average across leave one out iterations leaving out one mouse at a time, total neurons recorded from dominants: $n = 507$ and subordinates: $n = 490$ units from 20 mice). Self entry events are aligned to port entries of the subject mouse while other entry events are aligned to the competitor's port entries. Right, trajectory lengths (using PCs that captured 90% of variance) for self entry (top) and other entry (bottom) during the tone are longer for relative subordinates (self entry 2-way RM-ANOVA effect of relative rank $F(1,25) = 452$, $p = 5x10{-}14$ and interaction of relative rank and event $F(1,25) = 5,950$, $p = 1x10{-}17$; other entry 2-way ANOVA effect of relative rank $F(1,25) = 728$, $p = 3x10{-}15$ and interaction of relative rank and event $F(1,25) = 90$, $p = 5x10{-}7$). f, Left, Neural trajectories of mPFC population firing rate for port entries that occur during inter-trial interval (ITI) projected into the first two principal components of the common behavioral subspace. Insets show closer look to the dominant trajectories. Right, neural trajectory lengths for self entry (top) and other entry (bottom) during the ITI ($n = 14$ relative dom mice, $n = 13$ relative sub mice; self entry: 2-way RM-ANOVA main effect of rank $F(1,25) = 77.7$, $p = 1x10{-}9$; other entry: 2-way RM-ANOVA main effect of rank $F(1,25) = 110$, $p = 2x10{-}10$. Self entry events are aligned to port entries of the subject mouse while other entry events are aligned to the competitor's port entries. ITI port entries refer to port entries that occurred outside of the tone period.


a
Unit 1
m1
Unit K
Unit 1
m2
Unit K
Cue win trial
Cue lose trial
Self entry (tone)
Other entry (tone)
Self entry (ITI)
Other entry (ITI)
b
Win trials
Lose trials
Before cue
Cue
Trajectory length
dom
sub
intermediate mice
c
Probability
firing rate
rate of change (zs/s)
Relative rank
Sub
Dom
d
Trajectory start
Cue onset
Trajectory end
Relative rank
Win
Lose
Win sub.
Lose sub.
Lose dom.
Win dom.
Orthogonal P.C. lose (a.u.)
P.C. win (a.u.)
Alignment index
e
Trajectory start
Port entry
Trajectory end
Relative rank
Self
Other
Dom
Sub
P.C. 2 (a.u.)
P.C. 1 (a.u.)
Subordinate
Dominant
Self entries (during cue)
Other entries (during cue)
Before entry
Entry
f
Self sub.
Other sub.
Self dom.
Other dom.
Self entry (ITI)
Other entry (ITI)

Supplemental Figure 2-9 mPFC population dynamics during social competition are not driven by location or mouse identity.

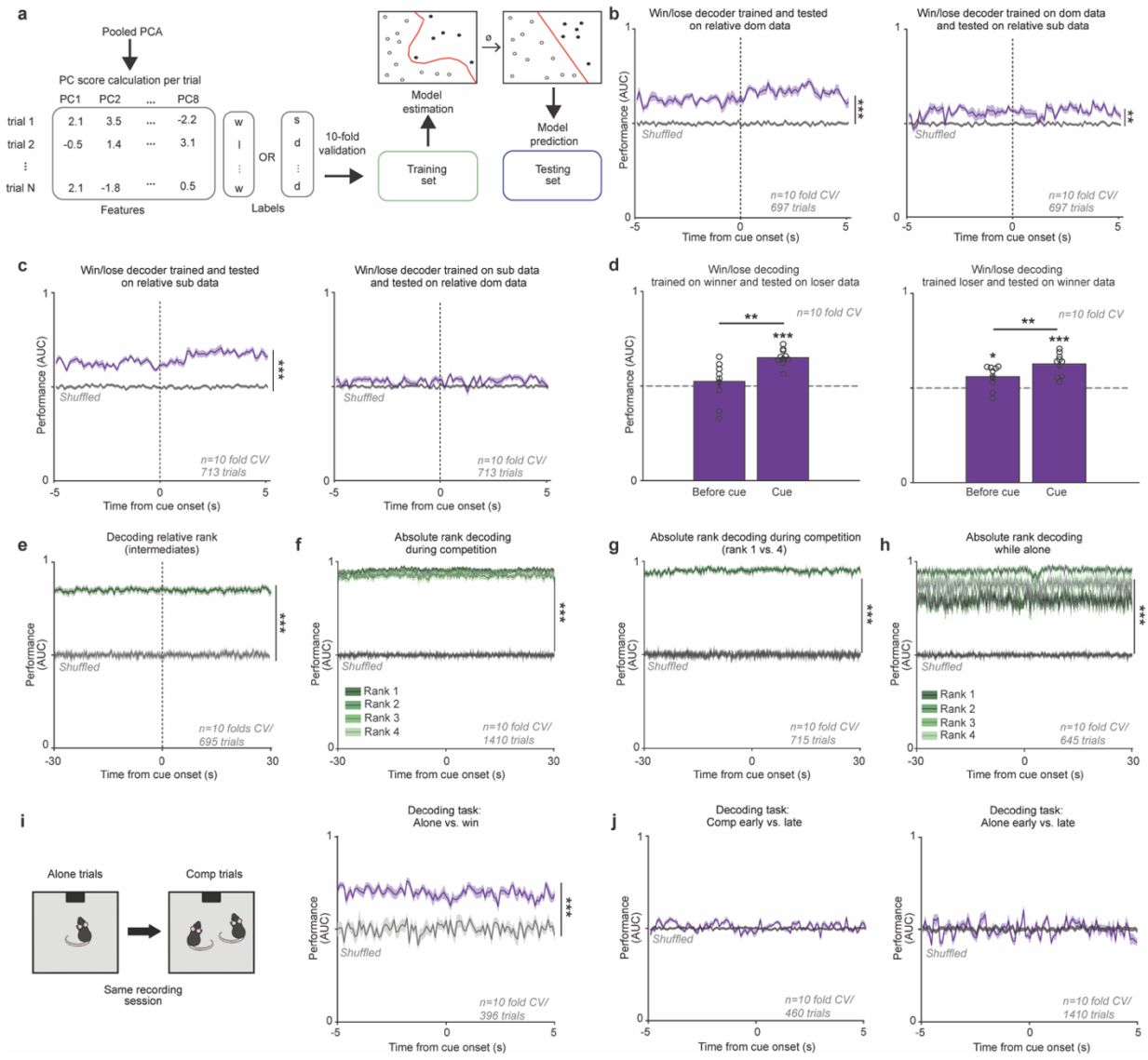
a, Average occupation in different parts of the chamber for win vs lose trials for the five seconds prior to tone vs first five seconds of tone. Black squares represent the reward port location. b, Distance to reward port differed by trial-type but not by rank (trials: dom win = 290, dom lose = 349, sub win = 349, sub lose = 290; 2-way ANOVA, main effect of trial-type $F(1,1274) = 353$, $p = 8.8\text{x}10{-}70$, rank $p = 0.098$ and interaction $p = 0.066$). c, Distribution of the correlation coefficients for firing rate and distance to port for the population of mPFC single units did not differ by rank (dom = 321, sub = 479; KS test, $p = 0.48$). d, To determine if distance to reward port affected the population dynamics during win and lose trials a subset of data with matched video conditions was split by distance to reward port. Neural trajectory lengths were higher for relative subordinates during win trials in which mice were close or far to the reward port during tone onset (dom n = 19 sessions, sub n = 18 sessions; win close to port: 2-way RM-ANOVA main effect of rank $F(1,35)=738$, $p = 5\text{x}10{-}21$; win far from port: 2-way RM-ANOVA main effect of rank $F(1,35) = 588$, $p=3\text{x}10{-}20$). e, Neural trajectory lengths were higher for relative subordinates during lose trials in which mice were close or far from reward port during tone onset (dom n = 19 sessions, sub n = 18 sessions; lose close to port: 2-way RM-ANOVA main effect of rank $F(1,35) = 588$, $p = 3\text{x}10{-}20$; lose far from port: 2-way RM-ANOVA main effect of rank $F(1,35) = 46.7$, $p = 5\text{x}10{-}11$). f, To determine if reward port “place cells” contributed to neural trajectory rank differences we calculated the neural trajectory lengths without cells that were correlated to distance to port in a subset of data with equivalent video settings (video resolution and camera angle). Left, neural trajectories for self entry during the tone are highest for relative subordinates without the distance correlated cells (dom n = 18 sessions, sub n = 18 sessions; 2-way RM-ANOVA main effect of rank $F(1,34) = 94.4$, $p = 1\text{x}10{-}13$). Right, neural trajectories are highest for relative subordinates without the distance correlated cells (dom n = 18 sessions, sub n = 18 sessions; excluding correlated cells: 2-way RM-ANOVA main effect of rank $F(1,34) = 100$, $p = 1\text{x}10{-}13$). g, Neural trajectories of mPFC population activity for two randomly selected halves of the data for (left) win and lose trials, (middle) port entries during the tone and (right) ITI port entries (data from 49 recording sessions from 20 mice). All trajectories reflect the mean trajectories across 50 bootstrapping iterations. h, Left, trajectory lengths for win and lose trials when data is divided randomly show no effect of group indicating that the effect of rank is not due to chance (n = 50; win: 2-way ANOVA, event $F(1,196) = 8.41$, $p = 0.004$, group $p = 0.62$; lose: event $p = 0.13$, group $p = 0.65$). Right, mean trajectory distances between groups for win and lose trials. i, Left, trajectory lengths for port entries during the tone when data is divided randomly show no effect of group (n = 50; self entry: 2-way ANOVA, event $F(1,196) = 14.2$, $p = 0.0002$, group $p = 0.97$; other entry: $F(1,196) = 6.76$, $p = 0.01$, group $p = 0.31$). Right, mean trajectory distances between groups for self-entry

and other entry during the tone. j, Left, trajectory lengths for ITI port entries when data is divided randomly show no effect of group (n = 50; self entry: 2-way ANOVA, event $F(1,196) = 10.3$, $p = 0.001$, group $p = 0.93$; other entry: event $p = 0.96$, group $p = 0.87$). Right, mean trajectory distances between groups for self entry and other entry during the ITI.


a
AlphaTracker output
Winning trials
Losing trials
Baseline
Tone
Dominant
Subordinate
count
b
Relative rank
Dom lose n=349
Dom win n=290
Sub lose n=290
Sub win n=349
Distance to port
Time from cue onset (s)
c
Relative rank
Sub n=473
Dom n=315
Normalized count
Correlation coefficient (r)
d
Win trials close to port
Win trials far from port
Before cue
Cue
Total length
Dom
Sub
Relative rank
e
Lose trials close to port
Lose trials far from port
f
Self entry (tone) excluding "place cells"
Other entry (tone) excluding "place cells"
Before entry
Entry
g
Trajectory start
Cue onset
Trajectory end
Win
Lose
Random group 1
Random group 2
Before cue
Cue
Port entry
Self
Other
Before entry
Entry
P.C. 1 (a.u.)
P.C. 2 (a.u.)
h
Before group 1
Cue group 1
Before group 2
Cue group 2
Trajectory length
Win trials
Lose trials
distance between group 1 and 2
Time (s)
i
Entry group 1
Entry group 2
Self entry (tone)
Other entry (tone)
Distance between group 1 and 2
j
Self entry (ITI)
Other entry (ITI)

Supplemental Figure 2-10 Decoding performance for relative and absolute social rank, and competitive success with different datasets.

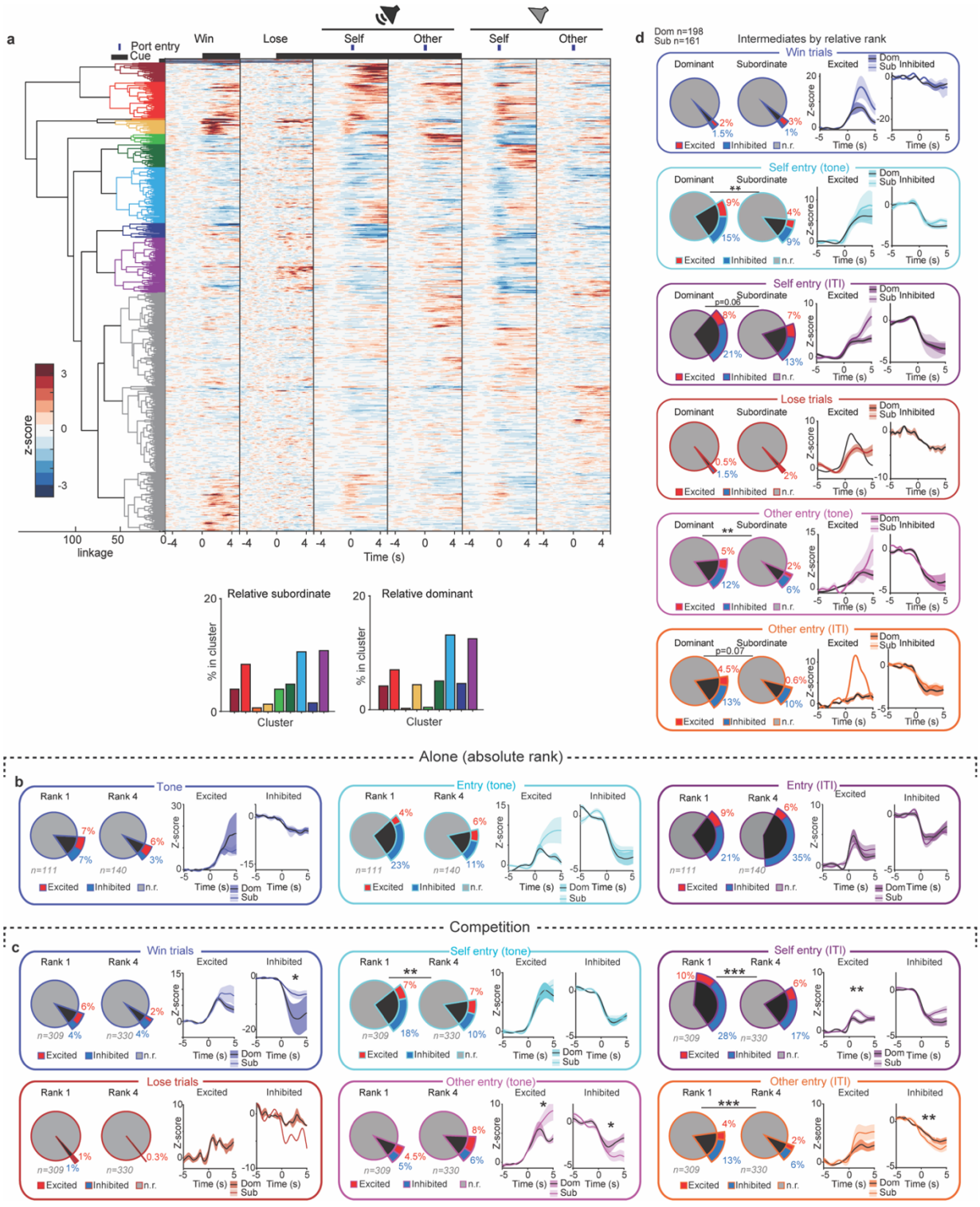
a, Support Vector Machine (SVM) data pipeline to decode rank or competition outcome based on single trial population mPFC data in the common behavioral subspace. b, mPFC population encoding of win/lose in relative dominants generalizes to relative subordinates. Decoding performance (area under the receiving operating curve; AUC) when (left) training and testing on relative dominant data or (right) training on dominant and tested on relative subordinate data was higher than chance (shuffled performance indicated in gray). (Wilcoxon rank sum, dom/dom $p = 0.0002$, dom/sub $p = 0.003$). c, mPFC population encoding of win/lose in relative subordinates does not generalize to relative dominants. Decoding performance (area under the receiving operating curve; AUC) when (left) training and testing on relative subordinate data was higher than chance but not when (right) testing on relative dominant data (shuffled performance indicated in gray). (Wilcoxon rank sum, sub/sub $p = 0.0002$, sub/dom $p = 0.14$). d, Decoder performance for classifying competition outcome using training data from winner data (e.g. mouse won majority of trials) and testing data from loser data (e.g. mouse lost majority of trials) and using training data from loser data and testing data from winner data (Wilcoxon rank sum: left, baseline vs shuffle $p = 0.10$, left, cue vs shuffle $p = 0.0002$, right, baseline vs shuffle $p = 0.02$, right, cue vs shuffle $p = 0.0002$; Wilcoxon sign rank: loser base vs cue $p = 0.002$, winner base vs cue $p = 0.004$). All error bars indicate standard error from 10-fold cross-validation. e, SVM performance for decoding relative rank specifically for intermediate (ranks 2 or 3) mice; mean AUC vs shuffled AUC Wilcoxon rank sum: $p = 0.0002$). f, Absolute rank can be decoded from mPFC population activity during social competition. One model was trained per absolute rank (mean performance across ranks vs shuffled data; Wilcoxon rank sum $p = 0.0002$). g, Absolute rank can be decoded for rank 1 and 4 animals from mPFC population activity during social competition. One model was trained to discriminate rank 1 trials from rank 4 (mean performance across ranks vs shuffled data: Wilcoxon rank sum $p = 0.0002$). h, Absolute rank can be decoded from mPFC population activity in mice performing reward task alone. One model was trained per absolute rank (mean performance across ranks vs shuffled data; Wilcoxon rank sum $p = 0.0002$). i, Left, experimental design. In 15 mice the same neurons were recorded during alone trials and followed by competition trials. Right, mPFC population activity can decode between alone tone presentations and win trials during the competition trials (shuffle performance indicated by gray line; mean AUC vs shuffled AUC Wilcoxon rank sum $p = 0.0002$). j, mPFC population activity is not sufficient to decode early vs late trials within task (alone mean AUC vs shuffle AUC Wilcoxon sum rank $p = 0.47$; comp mean AUC vs shuffle AUC Wilcoxon sum rank $p = 0.47$).


a
Pooled PCA
PC score calculation per trial
PC1 PC2 ... PC8
trial 1 2.1 3.5 ... -2.2
trial 2 -0.5 1.4 ... 3.1
trial N 2.1 -1.8 ... 0.5
Features
w l w
OR
s d d
Labels
10-fold validation
Model estimation
Training set
Model prediction
Testing set
b
Win/lose decoder trained and tested on relative dom data
Win/lose decoder trained on dom data and tested on relative sub data
Performance (AUC)
Shuffled
n=10 fold CV/ 697 trials
Time from cue onset (s)
c
Win/lose decoder trained and tested on relative sub data
Win/lose decoder trained on sub data and tested on relative dom data
n=10 fold CV/ 713 trials
d
Win/lose decoding trained on winner and tested on loser data
Win/lose decoding trained loser and tested on winner data
n=10 fold CV
Before cue
Cue
e
Decoding relative rank (intermediates)
n=10 folds CV/ 695 trials
f
Absolute rank decoding during competition
Rank 1
Rank 2
Rank 3
Rank 4
n=10 fold CV/ 1410 trials
g
Absolute rank decoding during competition (rank 1 vs. 4)
n=10 fold CV/ 715 trials
h
Absolute rank decoding while alone
n=10 fold CV/ 645 trials
i
Alone trials
Comp trials
Same recording session
Decoding task: Alone vs. win
n=10 fold CV/ 396 trials
j
Decoding task: Comp early vs. late
n=10 fold CV/ 460 trials
Decoding task: Alone early vs. late
n=10 fold CV/ 1410 trials

Supplemental Figure 2-11 Additional data for mPFC single unit responses to task-relevant events during social competition.

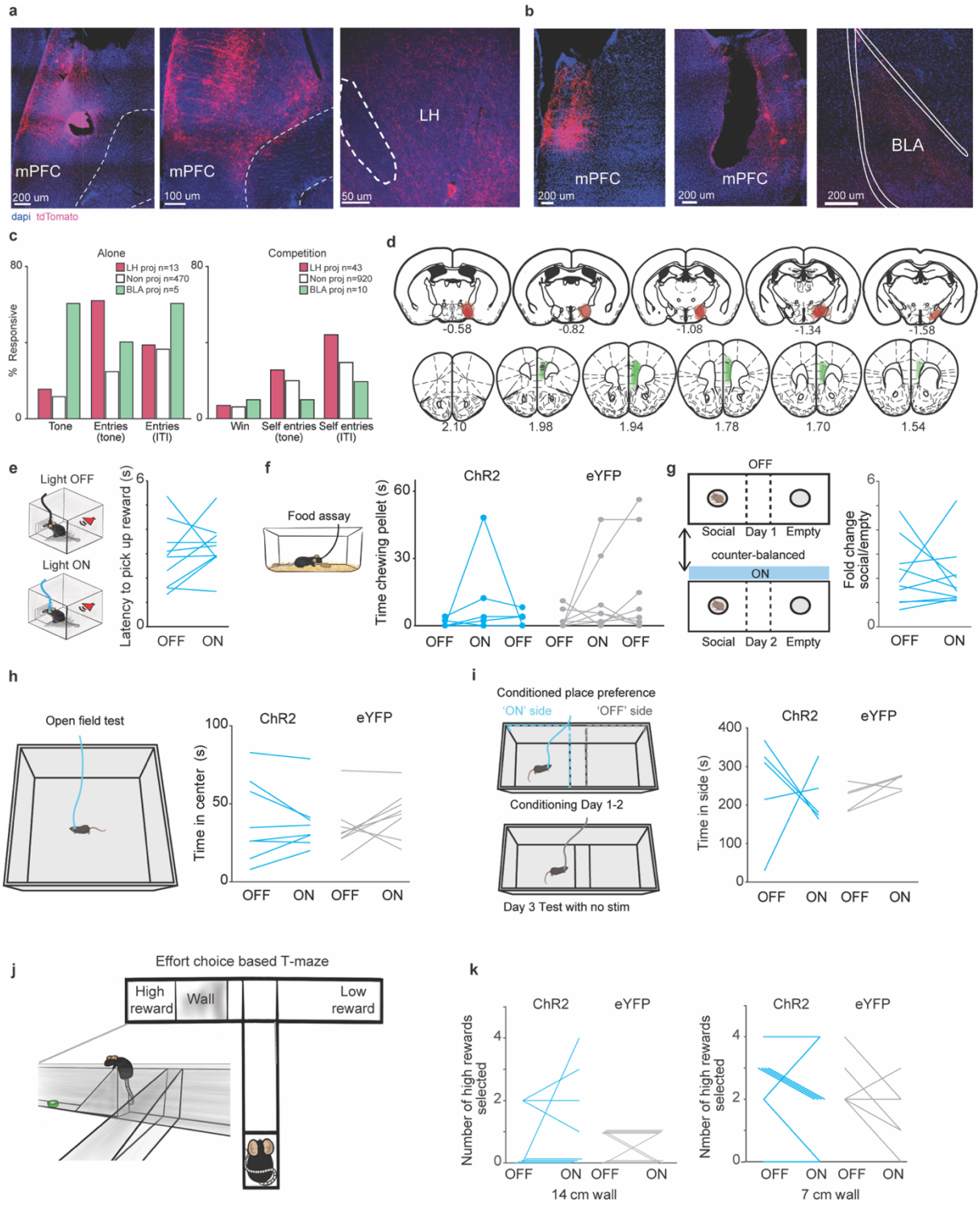
a, Top, Dendrogram for functional clusters and heatmap of mean firing rate for all the neurons included in the hierarchical clustering (n = 913 cells). Gray cells in the dendrogram indicate cells in functional clusters that did not meet criteria of mean z-score being higher than 2 or lower than −1 for at least one event. Bottom, distribution of mPFC cells across functional clusters in relative subordinates and relative dominants. b, Left, mPFC tone responsive cells when mice perform the reward task alone. Number of responsive cells and response magnitude to the tone does not differ between rank 1 and rank 4 mice (rank 1 exc = 8 rank 1 inh = 8 rank 4 exc = 8 rank 4 inh = 4; Fisher's exact test, total responsive per group p = 0.16; Wilcoxon rank sum across groups: exc p = 0.87, inh p = 1.0). Middle, mPFC tone port entries responsive cells when mice perform the reward task alone. Number of responsive cells and response magnitude to port entries during tone does not differ across dom (rank 1) vs sub (rank 4) mice (dom exc = 5 dom inh = 25 sub exc = 9 sub inh = 16; Fisher's exact test, total responsive per group p = 0.09; Wilcoxon rank sum across groups: exc p = 0.23, inh p = 0.62). Right, mPFC inter trial interval (ITI) port entries responsive cells when mice perform the reward task alone. Number of responsive cells and response magnitude to port entries during ITI does not differ between rank 1 and rank 4 mice (rank 1 exc = 10, rank 1 inh = 23 rank 4 exc = 9 rank 4 inh = 49; Fisher's exact test, total responsive per group p = 0.06; Wilcoxon rank sum across groups: exc p = 0.84, inh p = 0.17). c, Total responsive cells and response magnitude to task-relevant event during social competition for absolute rank 1 vs rank 4 (win trials: dom exc = 20, dom inh = 11, sub exc = 7, sub inh = 14, Fisher's exact test p = 0.11, Wilcoxon rank sum exc p = 0.23, inh p = 0.03; lose trials: dom exc = 3 dom inh = 3, sub exc = 0, sub inh = 1, Fisher's exact test p=0.12, Wilcoxon rank sum inh p = 0.50; self entries tone: dom exc = 23, dom inh = 57, sub exc = 24, sub inh=32, Fisher's exact test p = 0.006, Wilcoxon rank sum exc p = 0.42, inh p = 0.77; other entries tone: dom exc=14, dom inh 16, sub exc = 27, sub inh = 19, Fisher's exact test p = 0.11, Wilcoxon rank sum exc p = 0.049, inh p = 0.04; self entries ITI dom exc=31, dom inh = 89, sub exc = 21, sub inh=56, Fisher's exact test p = 2x10-5, Wilcoxon rank sum exc p = 0.01, inh p = 0.41; other entries ITI dom exc=13, dom inh = 41, sub exc = 8, sub inh = 21, Fisher's exact test p = 0.001, Wilcoxon rank sum exc p = 0.11, inh p = 0.008). d, Total responsive cells and response magnitude to task-relevant event during social competition for intermediate rank mice (ranks 2 and 3) by relative rank (win trials: dom exc = 4, dom inh = 3, sub exc = 5, sub inh=2, Fisher's exact test p = 0.76, Wilcoxon rank sum exc p = 0.11, inh p = 0.80; lose trials: dom exc = 1 dom inh = 3, sub exc = 3, sub inh = 0, Fisher's exact test p = 1, Wilcoxon rank sum exc p = 1; self entries tone: dom exc = 17, dom inh = 30, sub exc = 7, sub inh = 14, Fisher's exact test p = 0.01, Wilcoxon rank sum exc p = 0.89, inh p = 0.57; other entries tone: dom exc = 10, dom inh 23, sub exc = 3, sub inh = 10, Fisher's exact test p = 0.01, Wilcoxon

rank sum exc p = 0.46, inh p = 0.79; self entries ITI dom exc=15, dom inh = 42, sub exc = 11, sub inh = 21, Fisher's exact test p = 0.06, Wilcoxon rank sum exc p = 0.11, inh p = 0.44; other entries ITI dom exc = 9, dom inh = 26, sub exc = 1, sub inh = 16, Fisher's exact test p = 0.07, Wilcoxon rank sum exc p = 0.20, inh p = 0.90).


a
Port entry
Cue
Win
Lose
Self
Other
Self
Other
z-score
3
0
-3
100
50
0
linkage
-4 0 4 -4 0 4 -4 0 4 -4 0 4 -4 0 4
Time (s)
Relative subordinate
Relative dominant
% in cluster
Cluster
d
Dom n=198
Sub n=161
Intermediates by relative rank
Win trials
Self entry (tone)
Self entry (ITI)
Lose trials
Other entry (tone)
Other entry (ITI)
Dominant
Subordinate
Excited
Inhibited
n.r.
Z-score
Time (s)
Dom
Sub
Alone (absolute rank)
b
Tone
Entry (tone)
Entry (ITI)
Rank 1
Rank 4
Competition
c
Win trials
Lose trials
Self entry (tone)
Other entry (tone)
Self entry (ITI)
Other entry (ITI)

Supplemental Figure 2-12 mPFC-LH photostimulation does not affect other motivated behaviors.

a, Representative images showing electrode lesions and mPFC-LH cells and LH axon terminals (tdTomato). b, Representative images showing electrode lesions and mPFC-BLA cells and BLA axon terminals (tdTomato). c, Responsive cells to tones and port entries while performing the reward task alone vs in social competition (alone: tone mPFC-LH n = 2/13, mPFC-BLA n = 3/5, non-phototagged n = 54/470; entries during tone: mPFC-LH n = 8/13, mPFC-BLA n = 2/5, non-phototagged n = 115/470; entries during ITI: mPFC-LH n = 5/13, mPFC-BLA n = 3/5, non-phototagged n = 170/470; competition win trials: mPFC-LH n = 3/43, mPFC-BLA n = 1/10, non-phototagged n = 62/920; self entries during tone mPFC-LH: n = 11/43, mPFC-BLA n = 1/10, non-phototagged n = 193/920; self entries during ITI: mPFC-LH n = 19/43, mPFC-BLA n = 2/10, non-phototagged n = 271/920, Fisher's exact test nonphoto vs LH p = 0.011). d, Summary of mPFC optical fiber location (indicated with horizontal gray lines), mPFC viral expression and LH CAV2-Cre injection sites across mice for experiments shown below and in Figure 4. Distance to bregma is indicated under each brain slice. Top row shows LH injection and bottom row shows mPFC injection and fiber. e, mPFC-LH photostimulation in ChR2 mice did not change latency to pick reward while performing reward task alone (n = 10; paired t-test, p = 0.42). f, mPFC-LH photostimulation did not increase chow eating in the homecage (eYFP n = 8, ChR2 n = 7; 2-way RM ANOVA no significant effect of light, virus or interaction). g, mPFC-LH photostimulation in ChR2 mice did not change time spent in social chamber in the 3-chamber social interaction assay (n = 10; paired t-test, p = 0.79). h, mPFCLH photostimulation did not change anxiety-like behavior in the open field (ChR2 n = 8, eYFP n = 8; 2-way repeated measures (RM) ANOVA no significant effect of light, virus or interaction). i, mPFC-LH photostimulation did not evoke conditioned placed preference or aversion (ChR2 n = 5, eYFP n = 5; 2-way RM ANOVA no significant effect of light, virus or interaction). j, Effort based T-maze allows mice to choose between a low reward low effort arm or a high reward high effort arm in which they must climb a wall to obtain the reward. k, mPFC-LH photostimulation did not increase high effort choice in the effort T-maze (ChR2 n = 8, eYFP n = 9; 2-way RM ANOVA no significant effect of light, virus or interaction for both 14 and 7 cm walls).


a
mPFC
200 um
mPFC
100 um
LH
50 um
dapi tdTomato
b
mPFC
200 um
mPFC
200 um
BLA
200 um
c
Alone
LH proj n=13
Non proj n=470
BLA proj n=5
% Responsive
Tone
Entries (tone)
Entries (ITI)
Competition
LH proj n=43
Non proj n=920
BLA proj n=10
Win
Self entries (tone)
Self entries (ITI)
d
-0.58
-0.82
-1.08
-1.34
-1.58
2.10
1.98
1.94
1.78
1.70
1.54
e
Light OFF
Light ON
Latency to pick up reward (s)
OFF ON
f
Food assay
ChR2
eYFP
Time chewing pellet (s)
OFF ON OFF OFF ON OFF
g
OFF
Social Day 1 Empty
counter-balanced
ON
Social Day 2 Empty
Fold change social/empty
OFF ON
h
Open field test
ChR2
eYFP
Time in center (s)
OFF ON OFF ON
i
Conditioned place preference
'ON' side
'OFF' side
Conditioning Day 1-2
Day 3 Test with no stim
ChR2
eYFP
Time in side (s)
OFF ON OFF ON
j
Effort choice based T-maze
High reward
Wall
Low reward
k
ChR2
eYFP
Number of high rewards selected
OFF ON OFF ON
14 cm wall
ChR2
eYFP
Nmber of high rewards selected
OFF ON OFF ON
7 cm wall

2.11 Bibliography

"Advances in Large-Margin Classifiers." *MIT Press*. https://mitpress.mit.edu/9780262194488/advances-in-large-margin-classifiers/ (May 15, 2024).

Allsop, Stephen A., Romy Wichmann, Fergil Mills, Anthony Burgos-Robles, Chia-Jung Chang, Ada C. Felix-Ortiz, Alienor Vienne, et al. 2018. "Corticoamygdala Transfer of Socially Derived Information Gates Observational Learning." *Cell* 173(6): 1329-1342.e18. doi:10.1016/j.cell.2018.04.004.

Andriluka, Mykhaylo, Umar Iqbal, Eldar Insafutdinov, Leonid Pishchulin, Anton Milan, Juergen Gall, and Bernt Schiele. 2018. "PoseTrack: A Benchmark for Human Pose Estimation and Tracking." In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, , 5167–76. doi:10.1109/CVPR.2018.00542.

Bernstein, Irwin S. 1981. "Dominance: The Baby and the Bathwater." *Behavioral and Brain Sciences* 4(3): 419–29. doi:10.1017/S0140525X00009614.

Brodkin, Edward S., Andrea Hagemann, Sondra Maureen Nemetski, and Lee M. Silver. 2004. "Social Approach–Avoidance Behavior of Inbred Mouse Strains towards DBA/2 Mice." *Brain Research* 1002(1): 151–57. doi:10.1016/j.brainres.2003.12.013.

Burton, M. J., E. T. Rolls, and F. Mora. 1976. "Effects of Hunger on the Responses of Neurons in the Lateral Hypothalamus to the Sight and Taste of Food." *Experimental Neurology* 51(3): 668–77. doi:10.1016/0014-4886(76)90189-8.

Calhoun, Adam J., Jonathan W. Pillow, and Mala Murthy. 2019. "Unsupervised Identification of the Internal States That Shape Natural Behavior." *Nature Neuroscience* 22(12): 2040–49. doi:10.1038/s41593-019-0533-x.

Cannon, Walter B. 1929. "Organization for Physiological Homeostasis." *Physiological Reviews* 9(3): 399–431. doi:10.1152/physrev.1929.9.3.399.

Chiao, Joan Y. 2010. "Neural Basis of Social Status Hierarchy across Species." *Current Opinion in Neurobiology* 20(6): 803–9. doi:10.1016/j.conb.2010.08.006.

Cunningham, John P., and Byron M. Yu. 2014. "Dimensionality Reduction for Large-Scale Neural Recordings." *Nature Neuroscience* 17(11): 1500–1509. doi:10.1038/nn.3776.

Dewsbury, Donald A. 1982. "Dominance Rank, Copulatory Behavior, and Differential Reproduction." *The Quarterly Review of Biology* 57(2): 135–59. doi:10.1086/412672.

Dugatkin, Lee Alan. 1997. "Winner and Loser Effects and the Structure of Dominance Hierarchies." *Behavioral Ecology* 8(6): 583–87. doi:10.1093/beheco/8.6.583.

Escola, Sean, Alfredo Fontanini, Don Katz, and Liam Paninski. 2011. “Hidden Markov Models for the Stimulus-Response Relationships of Multistate Neural Systems.” *Neural Computation* 23(5): 1071–1132. doi:10.1162/NECO_a_00118.

Fang, Hao-Shu, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. 2017. “RMPE: Regional Multi-Person Pose Estimation.” In *2017 IEEE International Conference on Computer Vision (ICCV)*, , 2353–62. doi:10.1109/ICCV.2017.256.

Felix-Ortiz, A. C., A. Burgos-Robles, N. D. Bhagat, C. A. Leppla, and K. M. Tye. 2016. “Bidirectional Modulation of Anxiety-Related and Social Behaviors by Amygdala Projections to the Medial Prefrontal Cortex.” *Neuroscience* 321: 197–209. doi:10.1016/j.neuroscience.2015.07.041.

Gao, Peiran, and Surya Ganguli. 2015. “On Simplicity and Complexity in the Brave New World of Large-Scale Neuroscience.” *Current Opinion in Neurobiology* 32: 148–55. doi:10.1016/j.conb.2015.04.003.

Hou, Xun Helen, Minsuk Hyun, Julian Taranda, Kee Wui Huang, Emmalee Todd, Danielle Feng, Emily Atwater, et al. 2016. “Central Control Circuit for Context-Dependent Micturition.” *Cell* 167(1): 73-86.e12. doi:10.1016/j.cell.2016.08.073.

Hu, Jie, Li Shen, and Gang Sun. 2018. “Squeeze-and-Excitation Networks.” In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, , 7132–41. doi:10.1109/CVPR.2018.00745.

Janak, Patricia H., and Kay M. Tye. 2015. “From Circuits to Behaviour in the Amygdala.” *Nature* 517(7534): 284–92. doi:10.1038/nature14188.

Kalman, R. E. 1960. “A New Approach to Linear Filtering and Prediction Problems.” *Journal of Basic Engineering* 82(1): 35–45. doi:10.1115/1.3662552.

Karamihalev, Stoyo, Elena Brivio, Cornelia Flachskamm, Rainer Stoffel, Mathias V Schmidt, and Alon Chen. 2020. “Social Dominance Mediates Behavioral Adaptation to Chronic Stress in a Sex-Specific Manner” eds. Matthew N Hill, Kate M Wassum, and Jaideep Singh Bains. *eLife* 9: e58723. doi:10.7554/eLife.58723.

Lauer, Jessy, Mu Zhou, Shaokai Ye, William Menegas, Steffen Schneider, Tanmay Nath, Mohammed Mostafizur Rahman, et al. 2022. “Multi-Animal Pose Estimation, Identification and Tracking with DeepLabCut.” *Nature Methods* 19(4): 496–504. doi:10.1038/s41592-022-01443-0.

Lee, Christopher R., Alon Chen, and Kay M. Tye. 2021. “The Neural Circuitry of Social Homeostasis: Consequences of Acute versus Chronic Social Isolation.” *Cell* 184(6): 1500–1516. doi:10.1016/j.cell.2021.02.028.

Levy, Dana Rubi, Tal Tamir, Maya Kaufman, Ana Parabucki, Aharon Weissbrod, Elad Schneidman, and Ofer Yizhar. 2019. “Dynamics of Social Representation in the Mouse Prefrontal Cortex.” *Nature Neuroscience* 22(12): 2013–22. doi:10.1038/s41593-019-0531-z.
Li, Yi, Jiawei Zeng, Juen Zhang, Chenyu Yue, Weixin Zhong, Zhixiang Liu, Qiru Feng, and Minmin Luo. 2018. “Hypothalamic Circuits for Predation and Evasion.” *Neuron* 97(4): 911-924.e5. doi:10.1016/j.neuron.2018.01.005.

Ligneul, Romain, Ignacio Obeso, Christian C. Ruff, and Jean-Claude Dreher. 2016. “Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex.” *Current Biology* 26(23): 3107–15. doi:10.1016/j.cub.2016.09.015.

Lindzey, G., H. Winston, and M. Manosevitz. 1961. “Social Dominance in Inbred Mouse Strains.” *Nature* 191(4787): 474–76. doi:10.1038/191474a0.

Linsenbardt, David N., Nicholas M. Timme, and Christopher C. Lapish. 2019. “Encoding of the Intent to Drink Alcohol by the Prefrontal Cortex Is Blunted in Rats with a Family History of Excessive Drinking.” *eNeuro* 6(4). doi:10.1523/ENEURO.0489-18.2019.
Matthews, Gillian A., and Kay M. Tye. 2019. “Neural Mechanisms of Social Homeostasis.” *Annals of the New York Academy of Sciences* 1457(1): 5–25. doi:10.1111/nyas.14016.

McInnes, Leland, John Healy, and James Melville. 2020. “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.” doi:10.48550/arXiv.1802.03426.
Miller, E. K., and J. D. Cohen. 2001. “An Integrative Theory of Prefrontal Cortex Function.” *Annual Review of Neuroscience* 24: 167–202. doi:10.1146/annurev.neuro.24.1.167.

Miri, Andrew, Claire L. Warriner, Jeffrey S. Seely, Gamaleldin F. Elsayed, John P. Cunningham, Mark M. Churchland, and Thomas M. Jessell. 2017. “Behaviorally Selective Engagement of Short-Latency Effector Pathways by Motor Cortex.” *Neuron* 95(3): 683-696.e11. doi:10.1016/j.neuron.2017.06.042.

Munuera, Jérôme, Mattia Rigotti, and C. Daniel Salzman. 2018. “Shared Neural Coding for Social Hierarchy and Reward Value in Primate Amygdala.” *Nature Neuroscience* 21(3): 415–23. doi:10.1038/s41593-018-0082-8.

Murugan, Malavika, Hee Jae Jang, Michelle Park, Ellia M. Miller, Julia Cox, Joshua P. Taliaferro, Nathan F. Parker, et al. 2017. “Combined Social and Spatial Coding in a Descending Projection from the Prefrontal Cortex.” *Cell* 171(7): 1663-1677.e16. doi:10.1016/j.cell.2017.11.002.

Nieh, Edward H., Caitlin M. Vander Weele, Gillian A. Matthews, Kara N. Presbrey, Romy Wichmann, Christopher A. Leppla, Ehsan M. Izadmehr, and Kay M. Tye. 2016. “Inhibitory Input from the Lateral Hypothalamus to the Ventral Tegmental Area

Disinhibits Dopamine Neurons and Promotes Behavioral Activation." *Neuron* 90(6): 1286–98. doi:10.1016/j.neuron.2016.04.035.

Pereira, Talmo D., Nathaniel Tabris, Arie Matsliah, David M. Turner, Junyu Li, Shruthi Ravindranath, Eleni S. Papadoyannis, et al. 2022. "SLEAP: A Deep Learning System for Multi-Animal Pose Tracking." *Nature Methods* 19(4): 486–95. doi:10.1038/s41592-022-01426-1.

Pishchulin, Leonid, Eldar Insafutdinov, Siyu Tang, Bjoern Andres, Mykhaylo Andriluka, Peter Gehler, and Bernt Schiele. 2016. "DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation." doi:10.48550/arXiv.1511.06645.

Piva, Matthew, Kayla Velnoskey, Ruonan Jia, Amrita Nair, Ifat Levy, and Steve WC Chang. 2019. "The Dorsomedial Prefrontal Cortex Computes Task-Invariant Relative Subjective Value for Self and Other" eds. Thorsten Kahnt and Michael J Frank. *eLife* 8: e44939. doi:10.7554/eLife.44939.

Rabiner, Lawrence R. 1989. "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition." *Proceedings of the IEEE* 77(2): 257–86.

Rangel, Miguel J. Jr, Marcus V. C. Baldo, Newton S. Canteras, and Joel D. Hahn. 2016. "Evidence of a Role for the Lateral Hypothalamic Area Juxtadorsomedial Region (LHAjd) in Defensive Behaviors Associated with Social Defeat." *Frontiers in Systems Neuroscience* 10. doi:10.3389/fnsys.2016.00092.

Rigotti, Mattia, Omri Barak, Melissa R. Warden, Xiao-Jing Wang, Nathaniel D. Daw, Earl K. Miller, and Stefano Fusi. 2013. "The Importance of Mixed Selectivity in Complex Cognitive Tasks." *Nature* 497(7451): 585–90. doi:10.1038/nature12160.

Rossant, Cyrille, Shabnam N. Kadir, Dan F. M. Goodman, John Schulman, Maximilian L. D. Hunter, Aman B. Saleem, Andres Grosmark, et al. 2016. "Spike Sorting for Large, Dense Electrode Arrays." *Nature Neuroscience* 19(4): 634–41. doi:10.1038/nn.4268.

Segalin, Cristina, Jalani Williams, Tomomi Karigo, May Hui, Moriel Zelikowsky, Jennifer J Sun, Pietro Perona, David J Anderson, and Ann Kennedy. 2021. "The Mouse Action Recognition System (MARS) Software Pipeline for Automated Analysis of Social Behaviors in Mice" eds. Gordon J Berman, Kate M Wassum, and Asaf Gal. *eLife* 10: e63720. doi:10.7554/eLife.63720.

Shifrin, Theodore. "DIFFERENTIAL GEOMETRY: A First Course in Curves and Surfaces."

So, Nina, Becca Franks, Sean Lim, and James P. Curley. 2015. "A Social Network Approach Reveals Associations between Mouse Social Dominance and Brain Gene Expression." *PLOS ONE* 10(7): e0134509. doi:10.1371/journal.pone.0134509.

Walton, Mark E., David M. Bannerman, and Matthew F. S. Rushworth. 2002. "The Role of Rat Medial Frontal Cortex in Effort-Based Decision Making." *The Journal of*

Neuroscience: The Official Journal of the Society for Neuroscience 22(24): 10996–3. doi:10.1523/JNEUROSCI.22-24-10996.2002.

Wang, Fei, Helmut W. Kessels, and Hailan Hu. 2014. “The Mouse That Roared: Neural Mechanisms of Social Hierarchy.” *Trends in Neurosciences* 37(11): 674–82. doi:10.1016/j.tins.2014.07.005.

Wang, Fei, Jun Zhu, Hong Zhu, Qi Zhang, Zhanmin Lin, and Hailan Hu. 2011. “Bidirectional Control of Social Hierarchy by Synaptic Efficacy in Medial Prefrontal Cortex.” *Science* 334(6056): 693–97. doi:10.1126/science.1209951.

Williamson, Ryan C., Brent Doiron, Matthew A. Smith, and Byron M. Yu. 2019. “Bridging Large-Scale Neuronal Recordings and Large-Scale Network Models Using Dimensionality Reduction.” *Current Opinion in Neurobiology* 55: 40–47. doi:10.1016/j.conb.2018.12.009.

Wiltschko, Alexander B., Matthew J. Johnson, Giuliano Iurilli, Ralph E. Peterson, Jesse M. Katon, Stan L. Pashkovski, Victoria E. Abraira, Ryan P. Adams, and Sandeep Robert Datta. 2015. “Mapping Sub-Second Structure in Mouse Behavior.” *Neuron* 88(6): 1121–35. doi:10.1016/j.neuron.2015.11.031.

Zhou, Tingting, Hong Zhu, Zhengxiao Fan, Fei Wang, Yang Chen, Hexing Liang, Zhongfei Yang, et al. 2017. “History of Winning Remodels Thalamo-PFC Circuit to Reinforce Social Dominance.” *Science* 357(6347): 162–68. doi:10.1126/science.aak9726.

Zink, Caroline F., Yunxia Tong, Qiang Chen, Danielle S. Bassett, Jason L. Stein, and Andreas Meyer-Lindenberg. 2008. “Know Your Place: Neural Processing of Social Hierarchy in Humans.” *Neuron* 58(2): 273. doi:10.1016/j.neuron.2008.01.025.

2.12 Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears in: Nancy Padilla-Coreano*, Kanha Batra*, Makenzie Patarino, Zexin Chen, Rachel R. Rock, Ruihan Zhang, Sébastien B. Hausmann, Javier C. Weddington, Reesha Patel, Yu E. Zhang, Hao-Shu Fang, Srishti Mishra, Deryn O. LeDuke, Jasmin Revanna, Hao Li, Matilde Borio, Rachelle Pamintuan, Aneesh Bal, Laurel R. Keyes, Avraham Libster, Romy Wichmann, Fergil Mills, Felix H. Taschbach, Gillian A. Matthews, James P. Curley, Ila R. Fiete, Cewu

Lu, and Kay M. Tye. Cortical ensembles orchestrate social competition through hypothalamic outputs. *Nature*, 603 (7902), 667-671, 2022. The dissertation author was a co-primary investigator (*) and author of this paper.

Chapter 3 Visualizing the Longitudinal Development of Stress-Induced Anhedonia From Representations of Valence in the PFC

To successfully build a holistic model of the prefrontal cortex for decision-making, it is imperative to understand how it would perform under diseased states. In this chapter, I study precisely that by looking at the effect of stress induced anhedonia on the representation of valence in the PFC.

Using unsupervised machine learning, I categorize animals based on their anhedonic behavior post stress. Further, I use state space analysis to study the evolution of the neural representation of valence in the PFC over the course of a chronic mild stress paradigm. Lastly, I use a GLM based decoder to study how the separation between the neural manifolds of hedonic vs anhedonic group changes over time.

3.1 Abstract

A critical issue within the mental health field is the lack of granularity in diagnostic practices. We speculate that many distinct pathologies are currently being diagnosed as the same disease, and prescribed the same treatments – and that the key to developing effective antidepressants that work for everyone, we need to first identify a strategy to differentiate between heterogeneous conditions. Anhedonia, a condition identified in multiple neuropsychiatric disorders, is described as the inability to experience pleasure and is linked to anomalous medial prefrontal cortex (mPFC) activity. The mPFC is responsible for higher order functions, such as valence processing; however, it remains

unknown how mPFC valence-specific neuronal population activity is affected during anhedonic conditions. To test this, we implemented the unpredictable chronic mild stress (CMS) protocol in rodents and examined hedonic behaviors following stress and ketamine treatment. We used unsupervised clustering as a spectrum-based approach to delineate individual variability in hedonic values in response to stress. We then performed in vivo 2-photon calcium imaging to longitudinally track mPFC valence-specific neuronal population activity and dynamics during Pavlovian conditioning tasks. We observed reductions in both reward and aversive processing mPFC neurons within anhedonic mice, in contrast to hedonic groups. Using dimensionality reduction analysis, we identified a critical time point that shows an imbalance in mPFC dynamics that predicts resiliency to chronic mild stress, and rebounds following ketamine treatment. Also, we utilized behavioral pose-estimation tracking systems to detect a reduction in specific facial features during Pavlovian conditioning tasks during anhedonia. Lastly, we applied a linear classifier to decode resiliency to chronic mild stress based on mPFC valence-encoding properties. These results indicate that mPFC valence encoding properties are predictive of anhedonic states. Altogether, these experiments point to the need for increased granularity in the measurement of both behavior and neural activity, as these factors can decode the induction conditions of stress-induced anhedonia.

3.2 Introduction

Diagnosing neuropsychiatric disorders is challenging due to the overlap in symptoms and conditions, making it difficult for effective treatment interventions. As a result, mental health practices are transitioning towards spectrum-based approaches to provide granularity to prevent misdiagnosis, allowing more precise treatment (Diagnostic

and Statistical Manual of Mental Disorders n.d.). For example, a patient experiencing excessive sleeping and eating may receive the same antidepressant as one with insomnia and appetite loss. This may explain the low rates of efficacy for first line antidepressants and begs further dissection.

Anhedonia—described as the inability to experience pleasure and hedonic feeling (Destoop et al. 2019; Pizzagalli 2014)—is an underlying condition and core feature observed in both schizophrenia (SCZ), major depressive disorder (MDD)(Whitton and Pizzagalli 2022), and bipolar disorder (BD) (Dimick et al. 2021; Lee et al. 2015), and is suggested to be linked to anomalous medial prefrontal cortex (mPFC) activity. The mPFC, a higher order cortical region primarily responsible for cognition (Bechara, Damasio, and Damasio 2000; Miller and Buschman 2013), working memory (Euston, Gruber, and McNaughton 2012; Sawaguchi and Goldman-Rakic 1991), sociability(Forbes and Grafman 2010), and emotional control (Etkin, Egner, and Kalisch 2011), is also involved in valence processing (Coley et al. 2021), essential for discerning positive and negative hedonic values (Tye 2018). For instance, prefrontal cortex pathways participate in reward processing functions such as anticipatory pleasure, reward consumption (Burgos-Robles, Bravo-Rivera, and Quirk 2013), reward prediction (Schultz 2016), aversion (Anstrom and Woodward 2005; Vander Weele et al. 2018), and anhedonia (Ferenczi et al. 2016; Warden et al. 2012). Stress plays a major role in disrupting mPFC processes leading to depressive-phenotypes and is highly responsive to treatment. Indeed, imaging studies conducted in MDD patients reveal biomarkers within the mPFC predictive of response to therapy (Drysdale et al. 2017). This led us to hypothesize that mPFC valence-encoding processes and behavioral features are predictive during stress-induced anhedonia.

3.3 Chronic mild stress induces anhedonia and is predictive during Pavlovian Conditioning task

To test this, we implemented the unpredictable chronic mild stress (CMS) protocol (Fig. 1a) to induce anhedonia and assessed consummatory pleasure, despair, motivation, and sociability across weeks. We used sucrose preference test (SPT) as a measure of anhedonia and utilized unsupervised k-means clustering to classify subjects into resilient, neutral, and susceptible clusters (Fig. 1b, c). We then evaluated SPT scores in non-stressed (control) groups, hedonic (resilient and neutral) and anhedonic (susceptible) groups. Our results showed anhedonic groups display a significant reduction in sucrose preference following post-stress and is restored following ketamine administration and sustained post-ketamine time point (Fig. 1d-e). To examine whether anhedonia groups displayed similar scores prior to stress we analyzed SPT scores during baseline. We observed no change in anhedonia clusters at baseline.

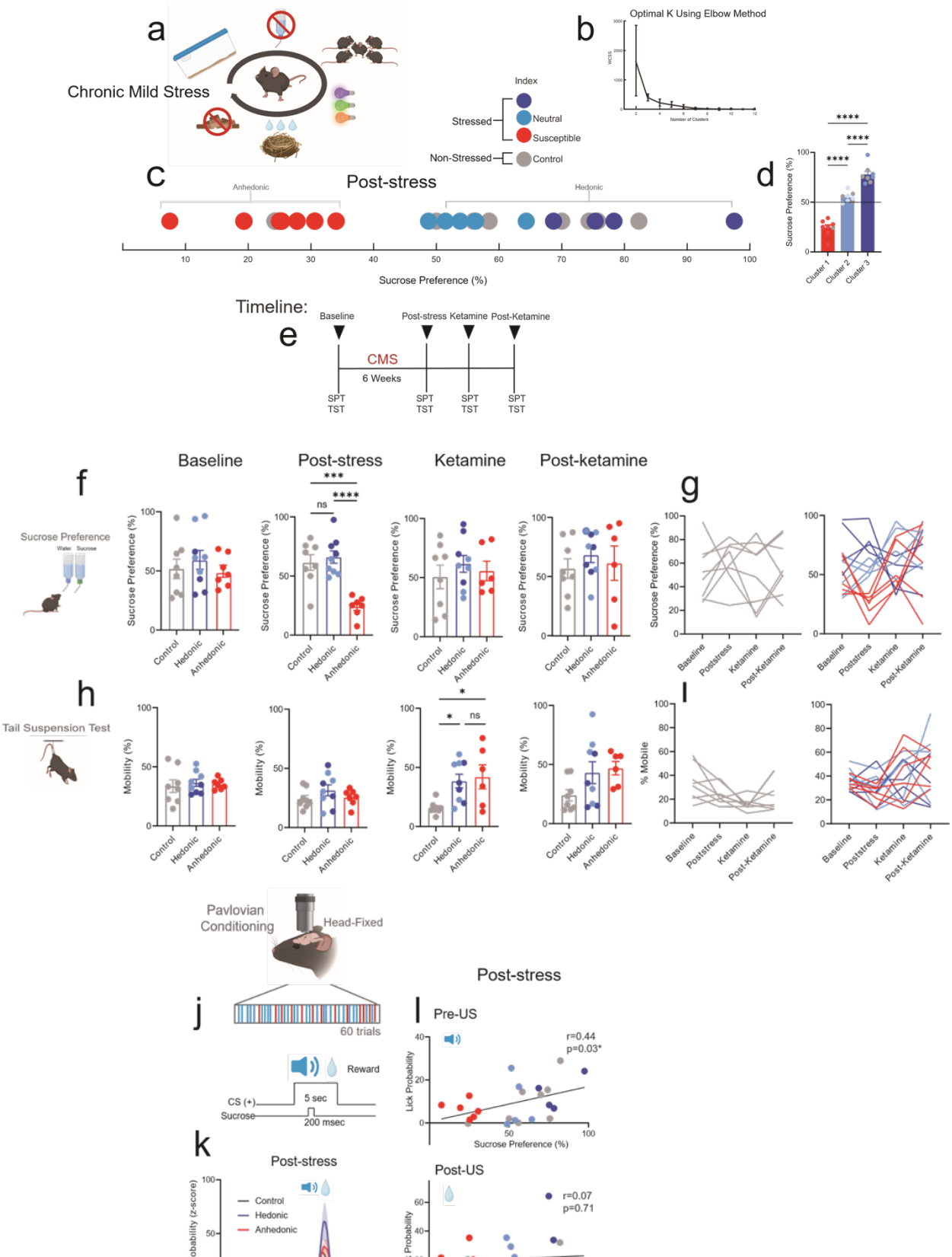
Additionally, CMS mice revealed no difference during tail suspension test (TST) at baseline or post-stress time points, indicating no change in behavioral despair or motivation; but showed an increase following ketamine treatment (Fig. 1f, g) that was diminished at post-ketamine time point. Indeed, we noticed a decline in control groups during ketamine administration (Fig. 1g), that was restored post-ketamine. This data suggests ketamine selectively increased behavioral despair in non-stressed groups. Interestingly, we observed no difference in social preference in anhedonia groups in response to chronic mild stress but noticed an increase in social variability during post-stress (Supp. Fig. 3.1c). Social variability diverges in response to ketamine, showing an

increase in hedonic groups but a reduction in anhedonic group, and is sustained post-ketamine.

To understand the association between anhedonia phenotypes and learning, we trained mice in a “head-fixed” Pavlovian conditioning task used to discriminate reward and aversive stimuli (Fig. 1h). During the task, one condition stimulus (CS) is paired (tone) with sucrose delivery (US-unconditioned stimulus) as reward (30% sucrose), and a different conditioned stimulus is paired with an air puff for aversion. Our results showed no difference in lick probability activity in reward trials during the anticipatory phase (CS) and appetitive phase (US) at post-stress (Fig. 1i). Further, we noticed a significant correlation in lick probability and sucrose preference during the conditioned stimulus at post-stress timepoint; revealing anhedonia subjects display both a reduction in lick probability and sucrose preference (Fig. 1j). We observed no correlation during the appetitive phase. These findings suggest that anhedonia classification is predictive of conditioned stimulus performance during post-stress timepoint.

Figure 3-1 Chronic mild stress induces anhedonia and is predictive during Pavlovian Conditioning task.

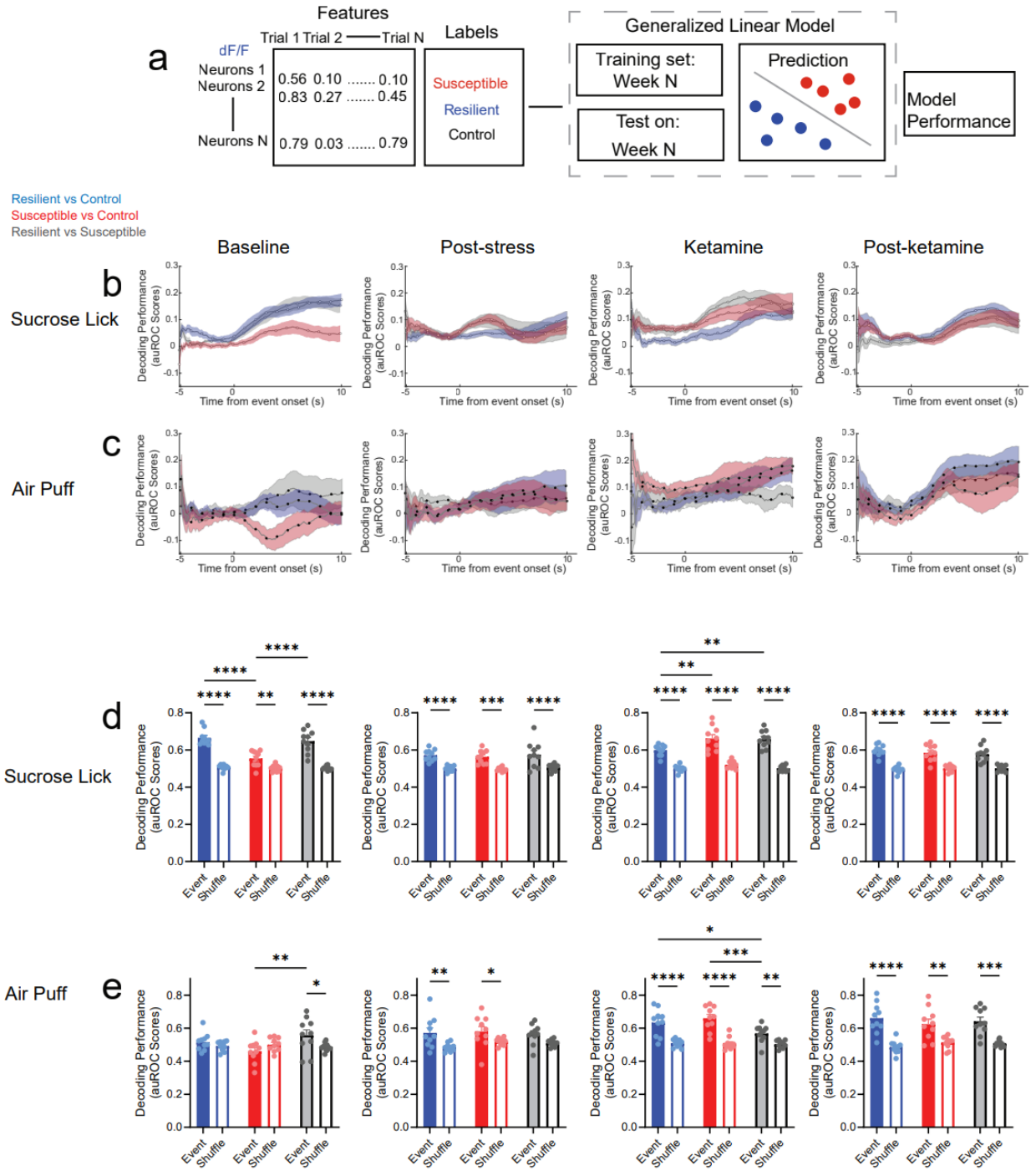
a, Schematic of unpredictable chronic mild stress (CMS) protocol. CMS mice were exposed to 2-3 stressors per day for 6 weeks that consisted of cage tilting, strobe light illumination, white noise, crowded housing, light/dark cycle manipulations, food deprivation, water deprivation and damp bedding. b, The optimal k elbow method using the within-cluster-sum-of-square (WCSS) was applied to determine the appropriate number of clusters derived from sucrose preference scores of mice post-stress time point. c, k-means clustering (k=3) of sucrose preference scores was applied to determine resilient (dark blue), neutral (light blue), and susceptible (red) subjects in both stressed (n=8) and non-stressed control (gray) groups (n=14). d, Cluster analysis of sucrose preference scores for resilient (cluster 3), neutral (cluster 2), and susceptible (cluster 1) groups. One-way ANOVA, $F(2,21)=100.3$, $p<0.0001$. Tukey Post hoc, Cluster 1\Cluster 2, $p<0.0001$, Cluster 1\Cluster 3, $p<0.0001$, Cluster 2\Cluster 3, $p<0.0001$. e, Timeline of measurements for sucrose preference test (SPT) and tail suspension test (TST) during chronic mild stress and ketamine treatment. f, Sucrose preference test scores at multiple time points in control, hedonic, and anhedonic mice. Baseline: One-way ANOVA, $F(2,21)=0.4606$, $p=0.6371$; Post-stress: One-way ANOVA, $F(2,21)=16.95$, $p<0.0001$, Tukey Post hoc, control\hedonic, $p=0.8051$; control\anhedonic, $p=0.0003$; control\hedonic, $p<0.0001$. Ketamine: One-way ANOVA, $F(2,20)=0.4637$, $p=0.6356$, Post-ketamine: One-way ANOVA, $F(2,20)=0.4364$, $p=0.6524$. g, Longitudinal description showing non-stressed control mice (left) and stressed (resilient, neutral, and susceptible) mice during sucrose preference test. h, Tail suspension test scores at multiple time points in control, hedonic, and anhedonic mice. Baseline: One-way ANOVA, $F(2,21)=0.3632$, $p=0.6997$; Post-stress: One-way ANOVA, $F(2,21)=1.185$, $p=0.3253$; Ketamine: One-way ANOVA, $F(2,20)=5.376$, $p=0.0135$, Tukey Post hoc, control\hedonic, $p=0.0309$; control\anhedonic, $p=0.0246$; hedonic\anhedonic, $p=0.9187$. Post-ketamine: One-way ANOVA, $F(2,20)=2.702$, $p=0.0915$. i, Longitudinal description showing non-stressed control mice (left) and stressed (resilient, neutral, and susceptible) mice during tail suspension test. j, Pavlovian conditioning paradigm in a head-fixed mouse showing US paired with a 5-second pure tone as the conditioned stimulus (CS), with the tone frequency set at 9 kHz for the rewarding CS. k, Peri-stimulus time histogram (PSTH) of lick probability during reward trials in control, hedonic, and anhedonic mice. l, Pearson’s correlation of lick probability and sucrose preference test in control, hedonic, and anhedonic mice. Top, Pre-US, $r=0.44$, $p=0.03$; Bottom, Post-US, $r=0.07$, $p=0.71$.


a
Chronic Mild Stress
b
Optimal K Using Elbow Method
WCSS
Number of Clusters
Index
Stressed
Neutral
Susceptible
Non-Stressed
Control
c
Post-stress
Anhedonic
Hedonic
Sucrose Preference (%)
d
Sucrose Preference (%)
Cluster 1
Cluster 2
Cluster 3
Timeline:
e
Baseline
Post-stress
Ketamine
Post-Ketamine
CMS
6 Weeks
SPT
TST
f
Baseline
Post-stress
Ketamine
Post-ketamine
Sucrose Preference
Water
Sucrose
Sucrose Preference (%)
Control
Hedonic
Anhedonic
ns
g
Baseline
Poststress
Ketamine
Post-Ketamine
h
Tail Suspension Test
Mobility (%)
i
% Mobile
Pavlovian Conditioning
Head-Fixed
j
60 trials
Reward
CS (+)
Sucrose
5 sec
200 msec
k
Post-stress
Lick Probability (z-score)
Time (s)
l
Post-stress
Pre-US
r=0.44
p=0.03*
Lick Probability
Sucrose Preference (%)
Post-US
r=0.07
p=0.71

3.4 mPFC population activity predicts anhedonia phenotypes prior to stress

To determine if mPFC population activity encodes anhedonia behavioral phenotype classification, we utilized longitudinal in vivo 2-Photon calcium imaging to track neuronal population activity performing a Pavlovian conditioning task during chronic mild stress and ketamine treatment. Using this data, we utilized a generalized linear model (GLM) to predict if mPFC neuronal population activity could decode control, hedonic and anhedonia groups (Fig. 2a). We trained and tested neural data acquired from the first sucrose lick during reward trials and air puff stimulus during aversive trials across weeks, and analyzed decoding performance for hedonic vs control groups, anhedonic vs control groups, and hedonic vs. anhedonic groups. Our results showed there is a high decoding accuracy for resilient vs control groups compared to shuffled data during reward trials for each week, with the exception of week 4 (Supp. Fig. 3.2). Similarly, we observed a high decoding performance during reward trials at all time points in anhedonic vs control groups, and hedonic vs. anhedonic groups (Supp. Fig. 3.2). This data suggest that we can use mPFC population activity to discern anhedonia phenotypes in response to reward stimuli.

Figure 3-2 mPFC population activity predicts anhedonia phenotypes prior to stress. a, Schematic depicts feature and label inputs for the generalized linear model classifier used for decoding performance. b, Time series traces depicting decoding performance during the first sucrose lick (reward trials) in hedonic vs control groups (blue), anhedonic vs control groups (red), and hedonic vs. anhedonic groups (dark gray) at baseline, post-stress, ketamine, and post-ketamine time points. c, Time series traces depicting decoding performance during air puff (aversive trials) in hedonic vs control groups (blue), anhedonic vs control groups (red), and hedonic vs. anhedonic groups (dark gray) at baseline, post-stress, ketamine, and post-ketamine time points. d, Decoding accuracy during Sucrose lick, Baseline: Two-way ANOVA, event $F(1,27)=187.8$, $p<0.0001$, groups $F(2,27)=13.10$, $p<0.0001$, interaction, $F(2,27)=14.32$, $p<0.0001$. Post-stress: Two-way ANOVA, event $F(1,27)=64.98$, $p<0.0001$, groups $F(2,27)=0.3770$, $p=0.6895$, interaction, $F(2,27)=0.0303$, $p=0.9731$. Ketamine: Two-way ANOVA, event $F(1,27)=203.4$, $p<0.0001$, groups $F(2,27)=6.495$, $p=0.0050$, interaction, $F(2,27)=3.472$, $p=0.0455$. Post-ketamine: Two-way ANOVA, event $F(1,27)=141.9$, $p<0.0001$, groups $F(2,27)=0.334$, $p=0.7190$, interaction, $F(2,27)=1.257$, $p=0.3007$. e, Decoding accuracy during Air puff, Baseline: Two-way ANOVA, event $F(1,27)=1.412$, $p=0.2450$, groups $F(2,27)=2.372$, $p=0.1125$, interaction, $F(2,27)=4.357$, $p=0.0229$. Post-stress: Two-way ANOVA, event $F(1,27)=18.34$, $p=0.0002$, groups $F(2,27)=0.5826$, $p=0.5653$, interaction, $F(2,27)=0.4530$, $p=0.6405$. Ketamine: Two-way ANOVA, event $F(1,27)=74.79$, $p<0.0001$, groups $F(2,27)=4.066$, $p=0.0286$, interaction, $F(2,27)=3.786$, $p=0.0355$. Post-ketamine: Two-way ANOVA, event $F(1,27)=65.09$, $p<0.0001$, groups $F(2,27)=0.03431$, $p=0.9663$, interaction, $F(2,27)=1.192$, $p=0.3191$. All post-hoc comparisons are sidak t-tests, $*p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$ All 2-way ANOVAs were for event (event vs shuffle) and groups (hedonic vs control, anhedonic vs control, and hedonic vs. anhedonic).


a
Features
Trial 1 Trial 2 —— Trial N
dF/F
Neurons 1
Neurons 2
Neurons N
0.56 0.10 ..... 0.10
0.83 0.27 ..... 0.45
0.79 0.03 ..... 0.79
Labels
Susceptible
Resilient
Control
Generalized Linear Model
Training set: Week N
Test on: Week N
Prediction
Model Performance
Resilient vs Control
Susceptible vs Control
Resilient vs Susceptible
Baseline
Post-stress
Ketamine
Post-ketamine
b
Sucrose Lick
c
Air Puff
Decoding Performance (auROC Scores)
Time from event onset (s)
d
Sucrose Lick
e
Air Puff
Event
Shuffle

We then compared decoding accuracy between hedonic vs control groups, anhedonic vs control groups, and hedonic vs. anhedonic groups at baseline, post-stress, ketamine, and post-ketamine time points. Interestingly, at baseline, we observed a significant increase in decoding performance in hedonic vs control groups compared to anhedonic vs control groups (Fig. 2d). Consistently, we noticed a significant increase in hedonic vs. anhedonic groups compared to anhedonic vs control groups. This data suggest that mPFC neural population activity response to reward stimuli predicts resilient behavior prior to stress. We observed no differences in anhedonia groups at post-stress. During the ketamine time point, we observed that anhedonic vs control groups displayed a significantly higher decoding accuracy compared to hedonic vs control groups, and is corroborated with an increase in hedonic vs. anhedonic groups compared to hedonic vs control groups. This data indicates mPFC population activity decoding accuracy in anhedonia mice is mostly affected in response to ketamine during reward trials.

We then examined decoding performance in response to air puff between hedonic vs control groups, anhedonic vs control groups, and hedonic vs. anhedonic groups at baseline, post-stress, ketamine, and post-ketamine time points (Supp. Fig. 3.2). We observed no difference in decoding accuracy for resilient vs control groups compared to shuffled data in response to air puff at baseline but observed a significant difference for all other weeks. Similarly, we noticed no difference in anhedonic vs control groups compared to shuffled data at baseline. This data suggests that mPFC decoded aversive stimuli only following the first week of stress.

Although, we observed no difference between hedonic vs control groups and anhedonic vs control groups, we detected an increase in hedonic vs. anhedonic groups

compared to anhedonic vs control groups at baseline during aversive trials (Fig. 2e), also showing a significant increase in hedonic vs. anhedonic groups compared to shuffled data. At post-stress timepoint, the hedonic vs. anhedonic groups displayed no difference compared to shuffled data. However, hedonic vs. anhedonic groups showed a significant increase in compared to hedonic vs control groups and anhedonic vs control groups during ketamine time-point. There was no difference between the groups at post-ketamine time point.

3.5 mPFC population dynamics exhibit dramatic changes during both early and late stress in hedonic groups

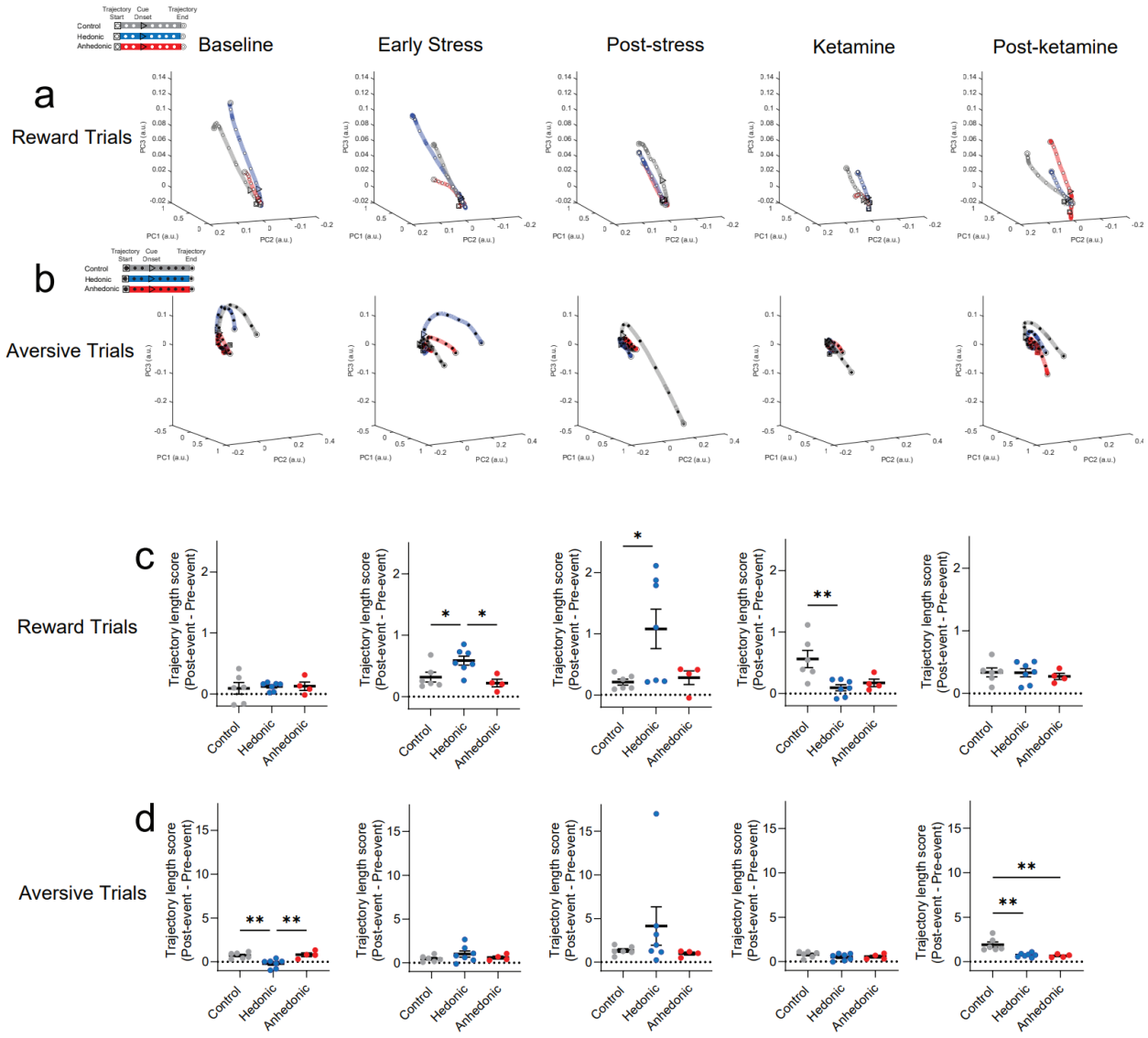
To examine mPFC population activity, we utilized principal component analysis (PCA) to plot activity in a lower dimensional space during reward and aversive trials during chronic mild stress and treatment (Fig. 3a, b). We examined population dynamics across weeks in control, hedonic and anhedonic groups by measuring trajectory length post-event during reward trials, aversive trials, and ratio (reward/aversive) (Supp. Fig. 3.3). Our results showed that control groups exhibit a dramatic increase in trajectory length following ketamine administration during reward trials, that was not observed in hedonic or anhedonic groups, suggesting diametrically opposed responses to ketamine due to stress. Also, interestingly, hedonic groups displayed a large increase in trajectory length during early stress (Week 2), that was not observed in control or anhedonic groups, indicating a potential biomarker for resiliency. No surprise, we noticed an increase in

length during late stress in hedonic groups during reward trials, consistent with resilient behavior to chronic mild stress.

We then analyzed the difference scores for trajectory lengths (post-event – pre-event) for reward trials at specific time points and compared across control, hedonic and anhedonic groups (Fig. 3a). We observed no change during baseline but noticed a significant increase in trajectory length difference score in hedonic groups during early stress in response to reward (Fig. 3c). Similarly, we revealed a significant increase in hedonic groups compared to control groups during reward at post-stress, but no difference compared to anhedonic groups. Additionally, we showed hedonic groups displayed a dramatic reduction in difference score compared to controls during ketamine time point, however no difference was seen in anhedonia groups. We observed no difference at post-ketamine time point, suggesting a restoration in a reward response.

Next, we measured trajectory length post-event during aversive trials, and noticed an increase in control mice during ketamine time point (Supp. Fig. 3.3), suggesting a similar sensitivity to valence stimuli. Also, during aversive trials we observed an increase in trajectory length in hedonic groups during early stress (Week 2), a similar trend to the response during reward trials. Also, we observed a variable response during week 6 in hedonic groups. Both hedonic and anhedonic groups display a reduction at ketamine time point and is restored following post-ketamine.

Figure 3-3 mPFC population dynamics exhibit dramatic changes during both early and late stress in hedonic groups.
a, Neural trajectories of mPFC populations during reward trials in control, hedonic, and anhedonic groups in a lower dimensional common principal component (PC) sub-space at time points: baseline, early stress, post-stress, ketamine, and post-ketamine. b, Neural trajectories of mPFC populations during aversive trials in control, hedonic, and anhedonic groups in a lower dimensional common principal component (PC) sub-space at time points: baseline, early stress, post-stress, ketamine, and post-ketamine. c, Difference score of trajectory lengths (post-event – pre-event) during reward trials (using principal components that captured 90% of variance) for control, hedonic, and anhedonic groups, Baseline: One-way ANOVA, $F(2,14)=0.08102$, $p=0.9226$, Early stress: $F(2,14)=6.207$, $p=0.0118$, Post-stress: $F(2,14)=4.444$, $p=0.032$, Ketamine: $F(2,14)=7.168$, $p=0.0072$, Post-ketamine: $F(2,14)=0.2385$, $p=0.7909$. d, Difference score of trajectory lengths (post-event – pre-event) during aversive trials (using principal components that captured 90% of variance) for control, hedonic, and anhedonic groups, Baseline: One-way ANOVA, $F(2,14)=10.04$, $p=0.0020$, Early stress: $F(2,14)=1.141$, $p=0.3475$, Post-stress: $F(2,14)=1.231$, $p=0.3218$, Ketamine: $F(2,14)=1.118$, $p=0.3543$, Post-ketamine: $F(2,14)=12.26$, $p=0.0008$. All post-hoc comparisons are Tukey t-tests, $*p<0.05$, $**p<0.01$.


a
Reward Trials
b
Aversive Trials
Baseline
Early Stress
Post-stress
Ketamine
Post-ketamine
Trajectory Start
Cue Onset
Trajectory End
Control
Hedonic
Anhedonic
PC1 (a.u.)
PC2 (a.u.)
PC3 (a.u.)
c
Reward Trials
d
Aversive Trials
Trajectory length score (Post-event - Pre-event)
Control
Hedonic
Anhedonic

We then analyzed the difference scores for trajectory lengths for aversive trials at specific time points and compared across groups (Fig, 3b). Our results showed a dramatic reduction in difference score in hedonic groups at baseline, however no changes at early stress, post-stress, or ketamine. Interestingly, we observed a dramatic reduction in both stressed groups at post-ketamine time point. Lastly, we measured the trajectory length valence ratio (reward/aversion) to understand the strength of valence properties across different time points. In control groups we noticed an increase in the trajectory length ratio at ketamine time points, indicating a preference towards reward processing. However, we observe no difference of trajectory length valence ratio in hedonic or anhedonic groups during stress or treatment.

3.6 Appendix

3.6.1 Animals and housing

Adult, male C57BL/6J mice with the HET DAT-Cre genotype (at the minimum age of weeks) and arrival from Jackson Laboratory (RRID: IMSRJAX:000,664), were utilized for this study. The mice were housed in a reverse light cycle, with ad libitum access to food and water, until the commencement of behavioral tests or imaging sessions. The animals were accommodated in cages with up to three nest mates. All animal handling procedures adhered to the guidelines stipulated by the National Institute of Health (NIH) and were approved by the UCSD Institutional Animal Care and Use Committee (IACUC).

3.6.2 Stereotaxic surgery

Under aseptic conditions, surgery was conducted on all subjects using a small animal stereotax (David Kopf Instruments, Tujunga, CA, USA), with body temperature

maintenance achieved through the use of a heating pad. Anesthesia was induced using a 5% mixture of isoflurane and oxygen, which was subsequently reduced to 2–2.5% and maintained throughout the procedure (0.5 L/min oxygen flow rate). Once the subjects reached an adequate level of anesthesia, a 1mg/kg Buprenex injection was administered intraperitoneally (IP), the ophthalmic ointment was applied to protect the eyes, hair was clipped from the incision site, the area was scrubbed alternatively three times with betadine and 70% ethanol, and lidocaine was subcutaneously (SQ) injected at the incision site. All measurements for viral injections were referenced from Bregma as the origin. Following the surgery, the subjects were IP injected with 1mL Ringer's Lactate and placed in clean cages containing water-softened mouse chow to facilitate recovery. The cages were positioned on a heating pad to aid in the recovery process.

3.6.3 Viral injection and GRIN lens placement surgery

To enable recording from medial prefrontal cortex (mPFC) neurons and optogenetic stimulation of axons originating from the ventral tegmental area (VTA), a combination viral approach was implemented. Following the aforementioned general surgical procedures, an incision was made to expose the skull. After skull leveling, craniotomies were performed above the mPFC and the VTA regions. For selective expression of ChrimsonR in VTA neurons projecting to the mPFC, 700 nL and 700 nL of AAV9-hSyn-FLEX-ChrimsonR-tdTomato were respectively injected into the VTA at stereotaxic coordinates of -4.2 mm and -4.1 mm dorsoventral from Bregma (stereotaxic coordinates from Bregma, in mm: -3.3 anteroposterior, 0.35 mediolateral). Subsequently, 300 nL of AAV8-hSyn-jGCaMP7f was injected into the mPFC at stereotaxic coordinates of 1.9 mm anteroposterior, 0.40 mm mediolateral, and -2.2 mm dorsoventral from

Bregma. The injections were carried out using a 10 µL Nanofil syringe (WPI, Sarasota, FL, USA) driven at a rate of 0.1 µL/min with a microsyringe pump and controller (Micro4; WPI, Sarasota, FL, USA). Following each viral injection for each virus, the needle was allowed to stay in place for 5-10 minutes to allow viral material penetration before extraction. To prevent contamination, the needle was thoroughly flushed with 70% ethanol and sterile water. Viral aliquots were sourced from the University of North Carolina Vector Core (Chapel Hill, NC). Subsequent to viral injections, a 1 x 4 mm gradient refractive index (GRIN) lens (Proview, Inscopix Inc, Mountain View, CA, USA) was inserted into the mPFC at stereotaxic coordinates of 1.9 mm anteroposterior, 0.4 mm mediolateral, and -2.18 mm dorsoventral from Bregma. The GRIN lens was then secured to the skull and headplate using C&B Metabond and cement (Parkell), respectively.

3.6.4 Behavioral testing

All behavioral testing occurred after a minimum of three weeks post-surgery recovery. Mice were individually handled for 15 minutes each day for five days to gain familiarity with experimenters and reduce stress during experiments.

3.6.5 Sucrose preference test

The sucrose preference test (SPT) was used to measure anhedonia and was conducted in operant chambers (Med Associates, Inc) placed within sound-attenuated cubicles. Each SPT session lasted for 60 minutes and involved the use of two electrical lickometers and a house light set at an intensity of 40 lux. The lickometers were connected to bottles containing either tap water or a 1% sucrose solution in tap water. The MedPC IV software (Med Associates, Inc) was utilized to detect and record each lick event. No additional food sources were available within the operant chambers. To ensure variability,

the bottle configuration was different in each of the six operant chambers used. This allowed for repeated measures experiments, enabling animals to be re-tested and re-establish learning during each session.

3.6.6 3-Chamber Sociability test

The 3-chamber sociability test was used to measure sociability and was performed in a clear rectangular plexiglass arena. Prior to each session, the subject mouse is habituated in the empty arena for 3 minutes. Subsequently, the mouse is taken out of the arena, and a novel male mouse is placed inside a barred cup on one side of the arena together with an empty barred cup on the opposite side. The subject mouse is placed in the arena for 7 minutes during which footage is taken with a digital video camera above the arena. Ethovision XT software (Noldus, Wageningen, Netherlands) was used to analyze the mouse's time spent in each zone (left, center, right) of the arena.

3.6.7 Tail suspension test

The tail suspension test was used to measure behavioral despair. The tail of each mouse was placed between two strips of autoclave labeling tape. The end of one strip of tape was then secured to a horizontal bar 40 cm from the ground, ensuring that the animal could not make other contact or climb during the assay. Video recording was started 90 s from the time that the animal was inverted and taped. Time spent struggling was measured by blind scoring each minute of video material after the testing was completed and was reported in seconds for each minute of the assay.

3.6.8 Unpredictable Chronic Mild Stress protocol

To induce anhedonic symptoms, the chronic mild stress (CMS) protocol was implemented within a mouse model(Tye et al. 2013). Mice in the CMS group were exposed to 2-3 stressors per day for 6 weeks that consisted of cage tilting, strobe light illumination, white noise, crowded housing, light/dark cycle manipulations, food deprivation, water deprivation, and damp bedding. Stressors were imposed over all cages and randomized across all the days.

3.6.9 Ketamine administration

After the 6-week chronic mild stress protocol, all mice were IP injected with saline (0.01-0.04 ml). The following week all mice were IP injected with ketamine (1 mg/kg, 0.01-0.04 ml) to alleviate anhedonia.

3.6.10 Anhedonia Classification

Mice were classified following chronic mild stress using unsupervised k-means clustering method (k=3). Number of clusters were determined by using the optimal k elbow method within-clusters sum of squares (WCSS). 3 groups were identified (Resilient, neutral, and susceptible). Groups were then reclassified into control (non-stressed), hedonic (stressed), and anhedonic (stressed). Resilient and neutral groups were pooled into hedonic groups.

3.6.11 Pavlovian conditioning paradigm and trial structure

In this Pavlovian paradigm, a highly palatable 30% sucrose solution (200 ms) served as the rewarding unconditioned stimulus (US), while a mildly aversive air puff to the subject's face (~ 10 psi, 100 ms) acted as the aversive US. Both the rewarding and

aversive US were paired with a 5-second pure tone as the conditioned stimulus (CS), with the tone frequency set at 9 kHz for the rewarding CS and 2 kHz for the aversive CS. The reward trial started with the CS followed by a lick contingent reward US with a 2-second delay. After the CS ended, the US was vacuumed away from the spout. The aversive trials started with the CS followed by the aversive US with a 2-second delay. The reward and aversive catch trials both consisted of the respective CS with no US. The trials were separated by a 25-30 second inter-trial interval (ITI).

Subjects were first head-fix trained in a closed box for 20 reward trails with no lick-contingency and no US delay. Each box was equipped with a replica of the acquisition setup, without the microscope. This consisted of a head-fix clamp fixed above the tube with the subject. A spout connected to a voltage recorder was fixed in front of the subject. The air-puff spout and camera were fixed to opposite sides of the subject. Training sessions were ramped up to 60 trials over 3 sessions, after which lick contingency was turned on with a 2-second US delay for 2 sessions. Subsequently, Discrimination training sessions started, where 20% of trials changed to aversive trials. Before acquisition trials started subjects were trained under the 2-photon microscope for another 3 sessions. If subjects did not perform correct anticipatory lick responses to > 50% of reward trials, learning was deemed unsuccessful.

The acquisition sessions consisted of 60 trials. Each session was divided into three consecutive blocks of trials: pre-stimulation, stimulation, and post-stimulation trials. Within the stimulation block, there were 5 aversive trials, 5 rewarding trials, and 2 empty trials (no CS and US), all preceded by photostimulation. In both the pre- and post-stimulation blocks, there were 8 aversive trials, 2 aversive catch trials, 36 reward trials without lick

contingency, and 2 reward catch trials. These trials were pseudorandomized across the two blocks, with the requirements that the first 3 trials of the pre-stimulation block were reward trials, there were no consecutive sequences of 3 aversive trials, and the catch trials occurred in the last 15% of the trials. During each trial, facial footage, in vivo calcium imaging, and lick behavior was recorded.

3.6.12 In vivo 2-photon calcium imaging

We used a two-photon microscope (Bruker Ultima Investigator, Bruker Nano) with a 20 × objective (0.45 NA, Olympus) and 920 nm excitation wavelength (Ti-Sapphire laser, Newport) for in vivo calcium imaging. Images were acquired using Prairieview (Bruker Nano). Each field-of-view (FOV) (512 × 512 pixels covering 524 × 524 μm) was scanned at ~29.8 Hz.

3.6.13 Optogenetic stimulation

Optogenetic stimulation for stimulation trials was carried out using a 593 nm LED Driver (LEDD1B T-Cube, Thorlabs; Newton, NJ). The light stimulation paradigm consisted of 100 light pulses at a frequency of 20 Hz, with each pulse lasting for 5 seconds. The stimulation occurred 7 seconds before the initiation of each stimulation trial.

3.6.14 Signal processing

Images from 2-photon calcium imaging were processed using Suite2P. We used Suite2P to correct motion artifacts, define regions of interest (ROIs) corresponding to individual neurons, and extract their GCaMP fluorescence(Pachitariu et al. 2017). We selected only cellular ROIs by manual curation. Sessions and trials that contained motion artifacts and technical issues were taken out for further analysis.

3.6.15 Perfusion

Following the conclusion of recording experiments subjects were deeply anesthetized with an injection of sodium pentobarbital (200 mg/kg, intraperitoneal injection) and perfused transcardially with 20 mL of ice-cold lactated Ringer's solution, followed by 20 mL ice-cold paraformaldehyde (4%; PFA) in phosphate-buffered saline (PBS). Brains were extracted and placed in 4% PFA for 24 h. The tissue was then equilibrated in a cryo-protectant solution (30% sucrose in PBS, w/v). Coronal slices measuring 60um were taken from the tissue using a sliding microtome (HM430; Thermo Fisher Scientific, Waltham, MA), and stored in PBS at 4 °C.

3.6.16 Epifluorescence imaging

Stained tissue slices were imaged using an epifluorescence microscope (Keyence BZ-X). Images were taken using a 2x objective lens. Following imaging, the images were evaluated to determine the location of viral expression as seen via fluorescence from the ChrimsonR-tdTomato and GCaMP7f. Recording sites were located using GRIN lens lesion locations.

3.6.17 Unsupervised hierarchical clustering

The neuronal response to GCaMP7f fluorescence signal for each neuron in the FOV was calculated. Each neuron's mean Z-score was calculated as (F(t) − Fm)/SD where F (t) is the $\Delta F/F0$ at time t and Fm is the mean of $\Delta F/F0$ in a baseline period. The Z-scored normalized trace was then averaged across session for reward and aversive trials per subject. The population mean response was then calculated by averaging the mean Z-scored normalized trace of all neurons for subject groups. The mean Z-score

response for each neuron was used for unsupervised hierarchical clustering with a Ward linkage method and Euclidean distance metric. Z-scored traces from all neurons from each cluster were then averaged to create a peri-event time histogram. The population averages each cluster represented activity during reward trials and aversive trials.

3.6.18 Principal component analysis

Principal component analysis (PCA) was used to measure population firing rate dynamics in the mPFC(Padilla-Coreano et al. 2022). A local PCA was done on a matrix containing all Z-scored normalized data (Reward CS tone, Aversive CS tone, Reward first lick, Reward US, Aversive US) for all animals such that we could compare neural trajectories across groups (Control, Hedonic, and Anhedonic). The matrix had neurons in rows, and in the columns had mean Z-score response during -5 to 10 seconds post CS event using 100 ms bins. The neural trajectories for each task-relevant event were created per group by multiplying the coefficients obtained in the PCA by the mean Z-score response across trials per week. For each neural trajectory, the length was calculated as the sum of Euclidean distances between adjacent 100 ms bins. Also, neural trajectories distances were calculated as the Euclidean distance between the two trajectories bin-by-bin. For statistical comparison analysis, the neural trajectory metrics were calculated using the leave-one-out (LOO) method, leaving out all the neurons from a single animal per group, therefore the number of iterations is the number of mice in that group. Thus, in every iteration the same PCA coefficients per cell were used for neural trajectory analysis. For quantification of trajectory lengths and distance between trajectories the first 40 PCs were used to capture 90% of the variance. For all trajectory visualizations and

trajectory quantifications, we matched the number of neurons for each group (Control, Hedonic, and Anhedonic) for comparison analysis across weeks.

3.6.19 Generalized linear model classifier

To test if anhedonia phenotype groups (Control, Hedonic, and Anhedonic) could be decoded during reward and aversive trials from mPFC population activity, we used a generalized linear model (GLM) classifier. To obtain anhedonia group mPFC population activity we used the coefficients obtained for each neuron in the local PCA and created a neural trajectory using the mean Z-score responses for the Reward and Aversive trials (Reward CS tone, Aversive CS tone, Reward first lick, Reward US, Aversive US). We trained the GLM using the first 8 PCs per session per week (-5 to 10 seconds post CS event) as features. We did a 10-fold cross-validation (CV), where the data was split into 10 subsets and in each iteration the training consisted of a different 90% subset of the data, then the testing was done with the remaining 10% of the data. For the 10-fold CV, we computed the area under the receiver operating characteristic curve (AUC score) for the test data. We used this model decode control versus resilient, control versus susceptible, and resilient versus susceptible. We then compared decoding performance (auROC scores) against shuffled data across weeks.

3.6.20 Statistical methods

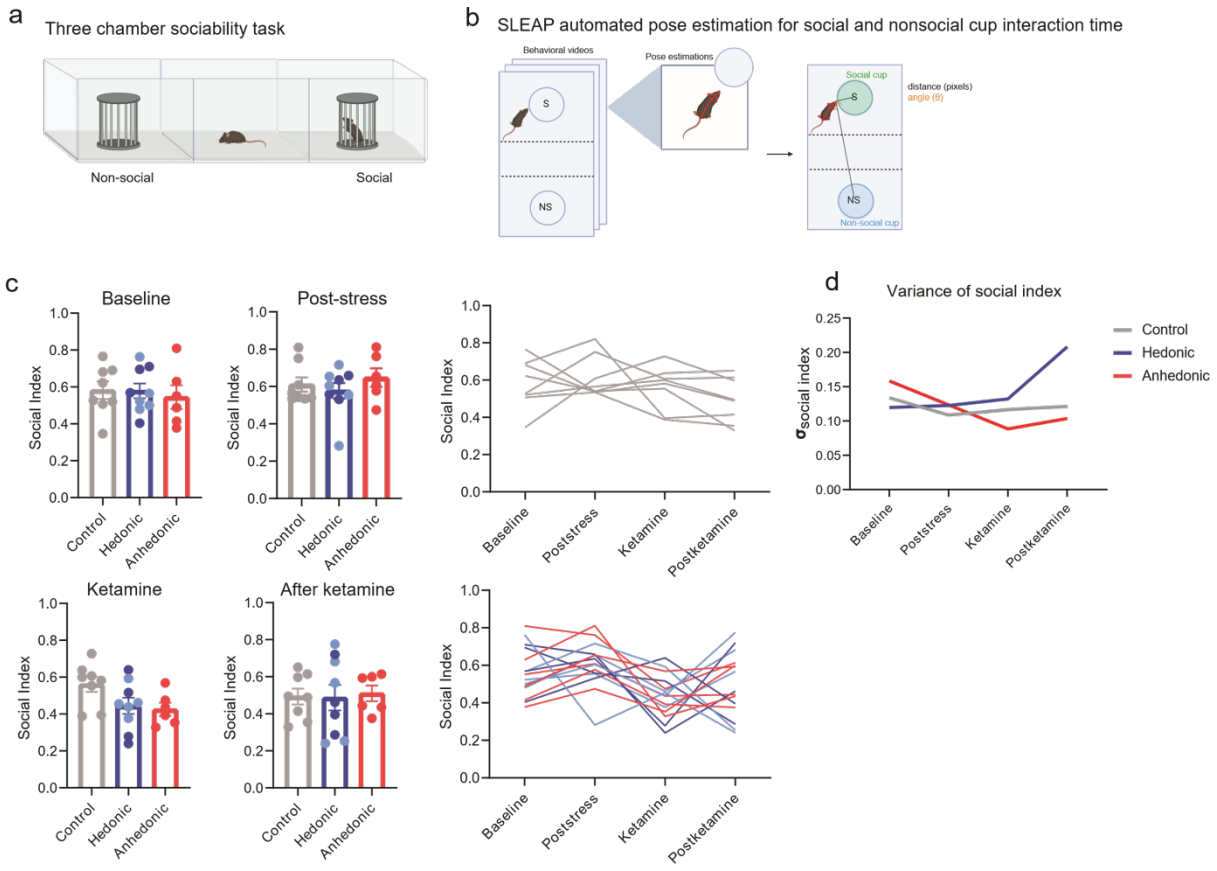
The thresholds for significance were placed at *$p<0.05$, **$p<0.01$, ***$p<0.001$, and ****$p<0.0001$ unless stated otherwise. All data are shown as mean and SEM. Wilcoxon signed rank-sum test, Pearson correlation, one-way ANOVA, Repeated-measure ANOVA, and mixed-effects model followed by a Tukey's posthoc test were performed using GraphPad Prism 6 or MATLAB. The p values were corrected for multiple

comparisons. Ward's linkage hierarchical clustering utilizing Euclidean distance was performed using MATLAB.

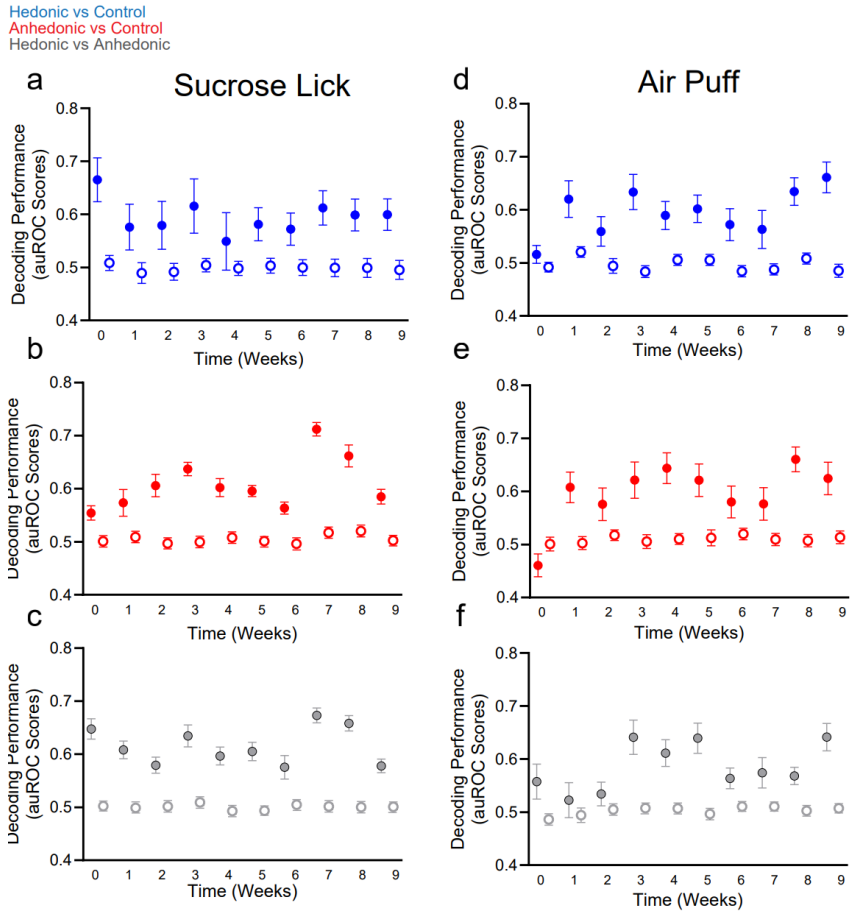
3.6.21 Social analysis

Social LEAP Estimates Animal Poses (SLEAP)(Pereira et al. 2022) was used to automatically detect mouse social interaction behavior. The behavioral videos for 3-chamber social interaction were initially attained using Noldus EthoVision XT and a Basler GenICam at 25 frames/second.

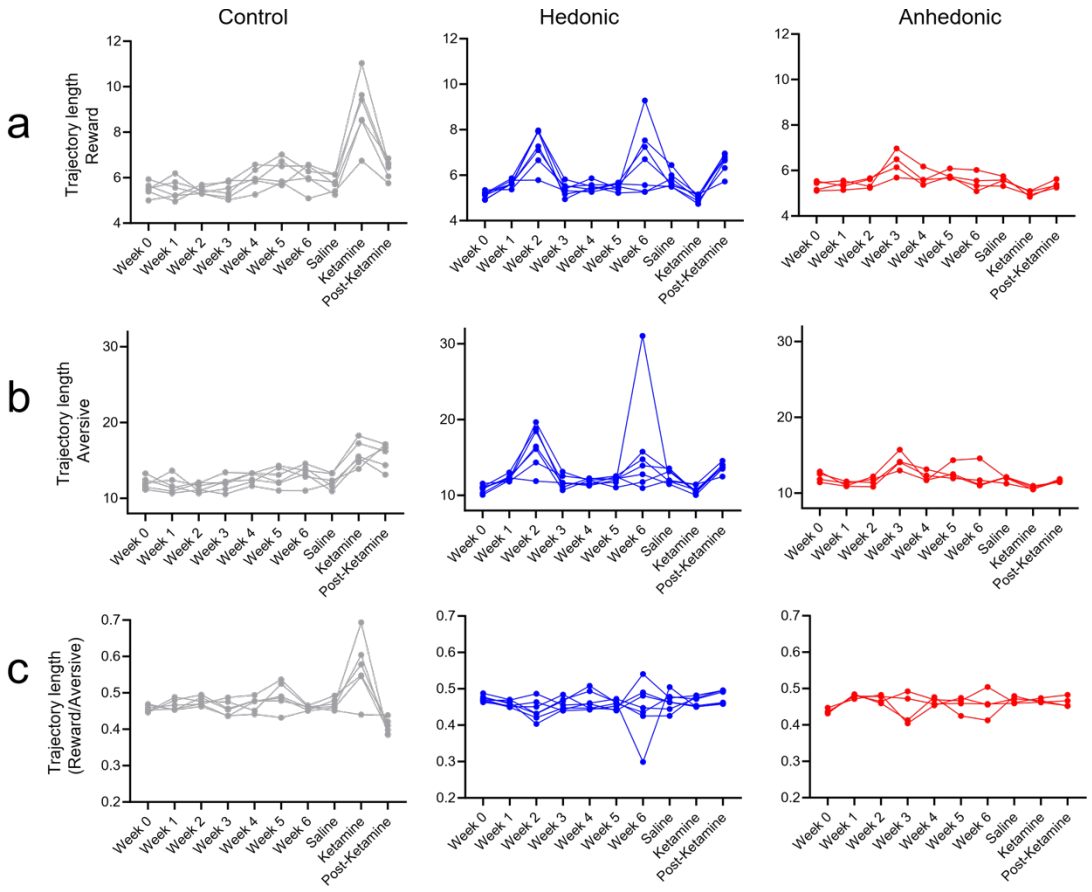
3.7 Supplemental figures



Supplemental Figure 3-1 Ketamine treatment after chronic mild stress decreases variance in social index in anhedonic mice.
a, Schematic of three-chamber sociability task assessing social preference. b, Workflow for SLEAP automated pose tracking, used to precisely quantify interaction time based on the subject's distance in pixels and angle to both the social and non-social cups. c, Social index, calculated as a ratio of time spent interacting with the social cup over combined social cup and non-social cup interaction times, measured at baseline (one-way ANOVA, Tukey's post hoc, interaction effect: $F(2, 20) = 0.1539$, $p = 0.8583$), post-stress (one-way ANOVA, Tukey's post hoc, interaction effect: $F(2, 20) = 0.09649$, $p = 0.5403$), ketamine (one-way ANOVA, Tukey's post hoc, interaction effect: $F(2, 20) = 0.2762$, $p = 0.0726$), and one week after ketamine treatment (one-way ANOVA, Tukey's post hoc, interaction effect: $F(2, 20) = 3.173$, $p = 0.9614$) time points. Error bars represent mean +/- SEM. d, Standard deviation plot of social index across baseline (Control: n=8, SD = 0.1338; Hedonic: n = 9, SD =0.1196; Anhedonic: n = 5, SD = 0.1585), post-stress (Control: n=8, SD = 0.1087; Hedonic: n = 9, SD =0.1255; Anhedonic: n = 5, SD = 0.1234), ketamine (Control: n=8, SD = 0.1166; Hedonic: n = 9, SD =0.1322; Anhedonic: n = 5, SD = 0.08842), and after ketamine (Control: n=8, SD = 0.1213; Hedonic: n = 9, SD =0.2082; Anhedonic: n = 5, SD = 0.1036) timepoints.



Supplemental Figure 3-2 Anhedonia groups decoded via mPFC population activity. a, Decoding accuracy in response to sucrose lick in hedonic vs control groups across weeks. Two-way ANOVA, event $F(1,90)=527$, $p<0.0001$, weeks $F(9,90)=6.583$, $p<0.0001$, interaction $F(9,90)=4.668$, $p<0.0001$. b, Decoding accuracy in response to sucrose lick in anhedonic vs control groups across weeks. Two-way ANOVA, event $F(1,90)=388.3$, $p<0.0001$, weeks $F(9,90)=8.851$, $p<0.0001$, interaction $F(9,90)=6.783$, $p<0.0001$. c, Decoding accuracy in response to sucrose lick in hedonic vs anhedonic groups across weeks. Two-way ANOVA, event $F(1,90)=406.1$, $p<0.0001$, weeks $F(9,90)=4.133$, $p=0.0002$, interaction $F(9,90)=3.771$, $p=0.0005$. d, Decoding accuracy in response to Air puff in hedonic vs control groups across weeks. Two-way ANOVA, event $F(1,90)=106.1$, $p<0.0001$, weeks $F(9,90)=2.428$, $p=0.0161$, interaction $F(9,90)=2.056$, $p=0.0419$. e, Decoding accuracy in response to Air puff in anhedonic vs control groups across weeks. Two-way ANOVA, event $F(1,90)=81.65$, $p<0.0001$, weeks $F(9,90)=3.172$, $p=0.0023$, interaction $F(9,90)=3.190$, $p=0.0022$. f, Decoding accuracy in response to Air puff in hedonic vs anhedonic groups across weeks. Two-way ANOVA, event $F(1,90)=93.74$, $p<0.0001$, weeks $F(9,90)=2.619$, $p=0.0098$, interaction $F(9,90)=2.570$, $p=0.0111$. All post-hoc comparisons are sidak t-tests, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$ All 2-way ANOVAs were for event (event vs shuffle) and weeks (0-9).



Supplemental Figure 3-3 mPFC neuronal population dynamics exhibit dramatic changes during select weeks.

a, Trajectory lengths post-event (0-10 sec at on-set of CS) during reward trials. Repeated measures One-way ANOVA control groups: subjects, $F(5,45)=5.089$, $p=0.0009$, weeks, $F(2.215, 11.07)=26.28$, $p<0.0001$; hedonic groups: subjects, $F(6,54)=1.835$, $p=0.1094$, weeks, $F(1.923, 11.54)=12.44$, $p=0.0014$; anhedonic groups: subjects, $F(3,27)=5.739$, $p=0.0036$, weeks, $F(2.558, 7.673)=9.185$, $p=0.0073$.

b, Trajectory lengths post-event (0-10 sec at on-set of CS) during aversive trials. Repeated measures One-way ANOVA control groups: subjects, $F(5,45)=5.878$, $p=0.0003$, weeks, $F(3.416, 17.08)=17.21$, $p<0.0001$; hedonic groups: subjects, $F(6,54)=1.056$, $p=0.4005$, weeks, $F(1.401, 8.405)=4.758$, $p=0.0505$; anhedonic groups: subjects, $F(3,27)=3.762$, $p=0.0223$, weeks, $F(1.501, 4.504)=6.963$, $p=0.0456$.

c, Trajectory lengths post-event (0-10 sec at on-set of CS) ratio of reward/aversive trials. Repeated measures One-way ANOVA control groups: subjects, $F(5,45)=1.197$, $p=0.3258$, weeks, $F(1.443, 7.216)=9.732$, $p=0.0119$; hedonic groups: subjects, $F(6,54)=0.2708$, $p=0.9482$, weeks, $F(1.745, 10.47)=0.9845$, $p=0.3943$; anhedonic groups: subjects, $F(3,27)=2.504$, $p=0.0804$, weeks, $F(1.801, 5.403)=1.573$, $p=0.2861$.

All post-hoc comparisons are Tukey t-tests, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

3.8 Bibliography

Anstrom, Kristin K., and Donald J. Woodward. 2005. “Restraint Increases Dopaminergic Burst Firing in Awake Rats.” *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 30(10): 1832–40. doi:10.1038/sj.npp.1300730.

Bechara, A., H. Damasio, and A. R. Damasio. 2000. “Emotion, Decision Making and the Orbitofrontal Cortex.” *Cerebral Cortex (New York, N.Y.: 1991)* 10(3): 295–307. doi:10.1093/cercor/10.3.295.

Burgos-Robles, Anthony, Hector Bravo-Rivera, and Gregory J. Quirk. 2013. “Prelimbic and Infralimbic Neurons Signal Distinct Aspects of Appetitive Instrumental Behavior.” *PLOS ONE* 8(2): e57575. doi:10.1371/journal.pone.0057575.

Coley, Austin A., Nancy Padilla-Coreano, Reesha Patel, and Kay M. Tye. 2021. “Valence Processing in the PFC: Reconciling Circuit-Level and Systems-Level Views.” *International Review of Neurobiology* 158: 171–212. doi:10.1016/bs.irn.2020.12.002.
Destoop, Marianne, Manuel Morrens, Violette Coppens, and Geert Dom. 2019.

“Addiction, Anhedonia, and Comorbid Mood Disorder. A Narrative Review.” *Frontiers in Psychiatry* 10: 311. doi:10.3389/fpsyt.2019.00311.

“Diagnostic and Statistical Manual of Mental Disorders.” *DSM Library*. https://dsm.psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596 (June 9, 2024).

Dimick, Mikaela K., Megan A. Hird, Lisa M. Fiksenbaum, Rachel H. B. Mitchell, and Benjamin I. Goldstein. 2021. “Severe Anhedonia among Adolescents with Bipolar Disorder Is Common and Associated with Increased Psychiatric Symptom Burden.” *Journal of Psychiatric Research* 134: 200–207. doi:10.1016/j.jpsychires.2020.12.031.

Drysdale, Andrew T, Logan Grosenick, Jonathan Downar, Katharine Dunlop, Farrokh Mansouri, Yue Meng, Robert N Fetcho, et al. 2017. “Resting-State Connectivity Biomarkers Define Neurophysiological Subtypes of Depression.” *Nature medicine* 23(1): 28–38. doi:10.1038/nm.4246.

Etkin, Amit, Tobias Egner, and Raffael Kalisch. 2011. “Emotional Processing in Anterior Cingulate and Medial Prefrontal Cortex.” *Trends in Cognitive Sciences* 15(2): 85–93. doi:10.1016/j.tics.2010.11.004.

Euston, David R., Aaron J. Gruber, and Bruce L. McNaughton. 2012. “The Role of Medial Prefrontal Cortex in Memory and Decision Making.” *Neuron* 76(6): 1057–70. doi:10.1016/j.neuron.2012.12.002.

Ferenczi, Emily A., Kelly A. Zalocusky, Conor Liston, Logan Grosenick, Melissa R. Warden, Debha Amatya, Kiefer Katovich, et al. 2016. “Prefrontal Cortical Regulation of Brainwide Circuit Dynamics and Reward-Related Behavior.” *Science (New York, N.Y.)* 351(6268): aac9698. doi:10.1126/science.aac9698.

Forbes, Chad E., and Jordan Grafman. 2010. “The Role of the Human Prefrontal Cortex in Social Cognition and Moral Judgment.” *Annual Review of Neuroscience* 33: 299–324. doi:10.1146/annurev-neuro-060909-153230.

Lee, Jung Suk, Suwon Jung, Il Ho Park, and Jae-Jin Kim. 2015. “Neural Basis of Anhedonia and Amotivation in Patients with Schizophrenia: The Role of Reward System.” *Current Neuropharmacology* 13(6): 750–59. doi:10.2174/1570159X13666150612230333.

Miller, Earl K., and Timothy J. Buschman. 2013. “Cortical Circuits for the Control of Attention.” *Current Opinion in Neurobiology* 23(2): 216–22. doi:10.1016/j.conb.2012.11.011.

Pachitariu, Marius, Carsen Stringer, Mario Dipoppa, Sylvia Schröder, L. Federico Rossi, Henry Dalgleish, Matteo Carandini, and Kenneth D. Harris. 2017. “Suite2p: Beyond 10,000 Neurons with Standard Two-Photon Microscopy.” : 061507. doi:10.1101/061507.

Padilla-Coreano, Nancy, Kanha Batra, Makenzie Patarino, Zexin Chen, Rachel R. Rock, Ruihan Zhang, Sébastien B. Hausmann, et al. 2022. “Cortical Ensembles Orchestrate Social Competition through Hypothalamic Outputs.” *Nature* 603(7902): 667–71. doi:10.1038/s41586-022-04507-5.

Pereira, Talmo D., Nathaniel Tabris, Arie Matsliah, David M. Turner, Junyu Li, Shruthi Ravindranath, Eleni S. Papadoyannis, et al. 2022. “SLEAP: A Deep Learning System for Multi-Animal Pose Tracking.” *Nature Methods* 19(4): 486–95. doi:10.1038/s41592-022-01426-1.

Pizzagalli, Diego A. 2014. “Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model.” *Annual review of clinical psychology* 10: 393–423. doi:10.1146/annurev-clinpsy-050212-185606.

Sawaguchi, T., and P. S. Goldman-Rakic. 1991. “D1 Dopamine Receptors in Prefrontal Cortex: Involvement in Working Memory.” *Science (New York, N.Y.)* 251(4996): 947–50. doi:10.1126/science.1825731.

Schultz, Wolfram. 2016. “Dopamine Reward Prediction Error Coding.” *Dialogues in Clinical Neuroscience* 18(1): 23–32.

Tye, Kay M. 2018. “Neural Circuit Motifs in Valence Processing.” *Neuron* 100(2): 436–52. doi:10.1016/j.neuron.2018.10.001.

Tye, Kay M., Julie J. Mirzabekov, Melissa R. Warden, Emily A. Ferenczi, Hsing-Chen Tsai, Joel Finkelstein, Sung-Yon Kim, et al. 2013. “Dopamine Neurons Modulate Neural Encoding and Expression of Depression-Related Behaviour.” *Nature* 493(7433): 537–41. doi:10.1038/nature11740.

Vander Weele, Caitlin M., Cody A. Siciliano, Gillian A. Matthews, Praneeth Namburi, Ehsan M. Izadmehr, Isabella C. Espinel, Edward H. Nieh, et al. 2018. “Dopamine Enhances Signal-to-Noise Ratio in Cortical-Brainstem Encoding of Aversive Stimuli.” *Nature* 563(7731): 397–401. doi:10.1038/s41586-018-0682-1.

Warden, Melissa R., Aslihan Selimbeyoglu, Julie J. Mirzabekov, Maisie Lo, Kimberly R. Thompson, Sung-Yon Kim, Avishek Adhikari, et al. 2012. “A Prefrontal Cortex-Brainstem Neuronal Projection That Controls Response to Behavioural Challenge.” *Nature* 492(7429): 428–32. doi:10.1038/nature11617.

Whitton, Alexis E., and Diego A. Pizzagalli. 2022. “Anhedonia in Depression and Bipolar Disorder.” *Current Topics in Behavioral Neurosciences* 58: 111–27. doi:10.1007/7854_2022_323.

3.9 Acknowledgements

Chapter 3, in full, is currently being prepared for submission for publication of the material. Austin A. Coley, Jeremy M. Delahanty, Kanha Batra, Assaf Ramot, Rachelle Pamintuan, Laurel Keyes, Jim Hagemann, Christopher R. Lee, Vivian Liu, Harini Advikoluna, Jianna Cressy, Caroline Jia, Francesca Massa, Deryn O. LeDuke, Moumen Gabir, Bra’a Durubeh, Lexe Linderhof, Reesha Patel, Analetta Ehler, Romy Wichmann, Hao Li, Kyle B. Fischer, Talmo Pereria, Takaki Komiyama, Kay M. Tye. *In preparation*, 2024. The dissertation author was a co-author of this paper. The data described in this chapter is preliminary and shall be updated in the final version of the manuscript.

Chapter 4 Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning

In this chapter I consider the prefrontal cortex as a hierarchical modular neural network architecture. The modules of the architecture are based on known mechanisms of computations in the PFC at the individual neuron, population and circuit levels. I design a multi-context foraging task to validate the architecture for lifelong learning. To develop the architecture, I follow two parallel branches: experimental and computational. For the experimental branch, I utilize behavioral and systems neuroscience techniques with a rodent model to systematically collect biological data to validate the assumptions behind the architecture as well as compare results in the task. For the computational branch, I utilize deep reinforcement learning techniques to simulate the task with the architecture. The overall goal of this chapter is to bridge the gap between the two branches through a bidirectional comparison of internal mechanisms as well as behavioral output, to produce a biologically plausible architecture.

4.1 Introduction

The ability to adapt behavior flexibly to changing environments is necessary for survival. This entails learning appropriate behaviors for a plethora of contexts without forgetting old ones, switching between contexts with minimal latency and using knowledge of previously learned contexts to adapt to new contexts rapidly and efficiently (Kudithipudi et al. 2022). These features together enable organisms to learn continually.

Modern artificial intelligence (AI) has developed enough to be able to learn these individual contexts and navigate through them optimally given enough training data. However, the problem of continuous or lifelong learning is still an unsolved problem. Given that nature has already solved this problem within biology of natural agents, it is prudent to study the brain to solve this problem.

The prefrontal cortex (PFC) is a higher order cognitive region of the brain implicated in a variety of functions such as executive planning, working memory, attention control, decision making and social behavior (Miller and Cohen 2001). Thus, it plays an important role in multi-context learning and flexible behavior. Many experimental studies have pointed to evidence that the PFC routes information through different sub-circuits downstream to direct the most appropriate behavior for a given context. The PFC is also a center for a diverse set of neuromodulatory receptors which have aided in its flexibility. Dopamine for example has been hypothesized to modulate the signal to noise ratio of different ensembles depending upon the valence of the incoming stimuli (Vander Weele et al. 2018). There have also been studies trying to model this phenomenon as a meta-learning system (Tsuda et al. 2020; Wang et al. 2018). However, there is disconnect between the two either because the models are being tested on non-ethological tasks or because the neural signature for context representation and transitions in the PFC has not been established with enough granularity to adequately inform the models.

In this exploratory study, we aim to build a novel computational architecture modeling the neuromodulatory filtering of signals in PFC neural ensembles as a lifelong learning algorithm. We also propose validating and informing the model using systems neuroscience methodologies to probe how neuromodulatory signals as well as population

activity in the PFC encode behavioral motifs during a multi-context foraging behavioral task.

4.2 Model architecture

We have adapted the classical mixture of experts (MoE) architecture to incorporate biological mechanisms inspired by the PFC. In the classical version, MoE is described as an ensemble algorithm where a “gating network” is trained to select the most appropriate “expert network” out of a given set in order to produce the most appropriate output given the input (R. A. Jacobs et al. 1991). In this design, the “gating” mechanism can only select one expert at a time and the “experts” are feedforward networks pre-trained individually on the tasks that the architecture is likely to encounter. While this was an elegant way to utilize the expertise of multiple networks while sparing the computational requirements of activating them simultaneously, it suffers from limitations: a) the learning capacity of the model scales with the number of experts leading to an exponential increase in the parameter space, b) the experts don’t share knowledge leading to repetitive learning of overlapping task elements and c) energy savings from single expert activation come at the cost of inability to solve tasks beyond the ability of any individual expert.

In related work, the PFC has been modeled as a Dynamic Mixture of Experts (DynaMoE) network with the following advantages: the experts are trained in conjunction with the gating network, the gating network can recruit new experts dynamically in order to expand learning capacity and the gating mechanism is explained using dopaminergic mediation of signal to noise ratio of mPFC neurons (Tsuda et al. 2020; Vander Weele et al. 2018). This study enhances the original architecture with the ability to grow along with knowledge transfer of experts – instilling some of the core requirements of lifelong

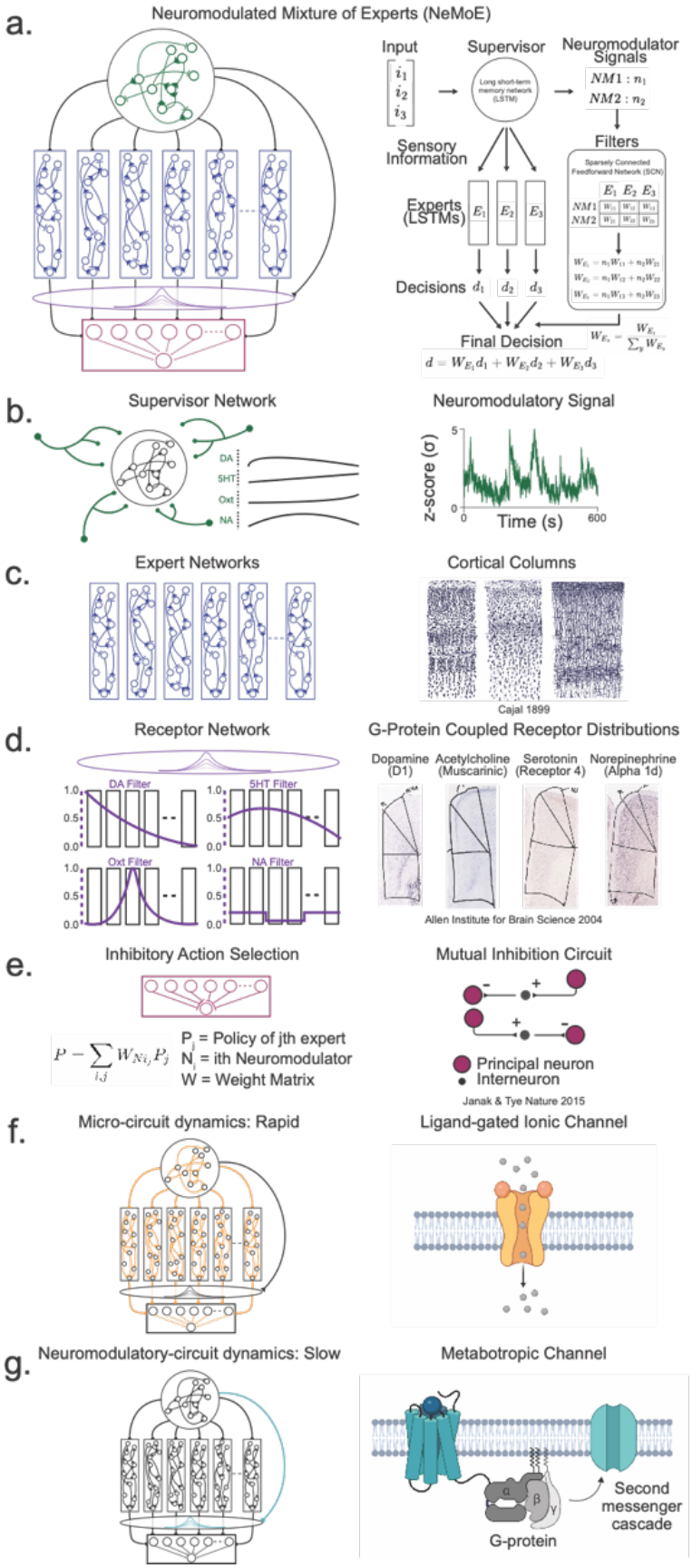
learning. However, the reliance on network expansion to increase learning capacity as well as single expert inference similar to classical MoE leads to parameter and energy inefficiency.

In this study, we have adapted the MoE architecture to include neuromodulation as a mechanism of context-specific neuronal signal filtering, creating the Neuromodulated Mixture of Experts (NeMoE) (Fig. 4.1a). The distinction from the traditional “gating” enables NeMoE to activate multiple experts at once and weigh their expertise with relevance to the input. This provides the network with the ability to navigate the trade-off between granular single-expert decisions vs all-expert decisions both of which lie at two extremes.

The “supervisor network” is a recurrent neural network that models the neuromodulatory signals transmitted to the mPFC from the rest of the brain (Fig. 4.1b). The relative “strengths” of the signals is proportional to the extent to which the corresponding receptors are activated. The “expert networks” are recurrent neural networks modeling cortical columns in the PFC (Oberländer 2012) (Fig. 4.1c). Each expert, mirroring a neuron ensemble with uniform neuromodulatory receptor functions, can generate its own policy over the action space, with a weighted contribution to the agent’s final decisions. The “receptor” network is a sparse feedforward network that simulates the PFC’s receptor sensitivity and distribution, applying “filters” to expert outputs based on neuromodulator levels (Allen Brain Atlas 2004) (Fig. 4.1d). An inhibitory action selection module, implemented as a softmax, integrates these policies into a final decision (Janak and Tye 2015) (Fig. 4.1e).

Figure 4-1 Components of the architecture and their biological counterparts.

a, (left) NeMoE architecture; (right) Flow chart describing the weighted decision-making process. b, (left) Supervisor network as a Long Short-Term Memory (LSTM) unit, schematic of modeled neuromodulatory signal transients (DA = Dopamine, 5HT = Serotonin, Oxt = Oxytocin, NA = Noradrenaline); (right) Schematic of an example trace z-score obtained using fluorescent biosensors. c, (left) Expert networks as LSTM units; (right) Cortical columnar structure adapted from Cajal 1899. d, (left) Receptor network as a sparsely connected feedforward network, schematic of neuromodulatory filter weights for the expert networks; (right) Section of mPFC showing receptor expression patterns adapted from the Allen Brain Atlas for dopamine, acetylcholine, serotonin, and norepinephrine annotated with sub-regions of the mPFC. e, (left) Inhibitory action-selection module as a softmax layer; (right) schematic of mutual inhibition adapted from Janak and Tye 2015. f, Micro-circuit connections and g, Neuromodulatory-circuit connections highlighted in the NeMoE architecture (left) along withs schematics of ligand-gated ionic channel and metabotropic channel (right) created using Biorender®.


a.
Neuromodulated Mixture of Experts (NeMoE)
Input
Supervisor
Neuromodulator Signals
Long short-term memory network (LSTM)
NM1 : n1
NM2 : n2
Sensory Information
Filters
Sparsely Connected Feedforward Network (SCN)
Experts (LSTMs)
Decisions
Final Decision
b.
Supervisor Network
Neuromodulatory Signal
DA
5HT
Oxt
NA
z-score (σ)
Time (s)
c.
Expert Networks
Cortical Columns
Cajal 1899
d.
Receptor Network
G-Protein Coupled Receptor Distributions
DA Filter
5HT Filter
Oxt Filter
NA Filter
Dopamine (D1)
Acetylcholine (Muscarinic)
Serotonin (Receptor 4)
Norepinephrine (Alpha 1d)
Allen Institute for Brain Science 2004
e.
Inhibitory Action Selection
Mutual Inhibition Circuit
Pj = Policy of jth expert
Ni = ith Neuromodulator
W = Weight Matrix
Principal neuron
Interneuron
Janak & Tye Nature 2015
f.
Micro-circuit dynamics: Rapid
Ligand-gated Ionic Channel
g.
Neuromodulatory-circuit dynamics: Slow
Metabotropic Channel
Second messenger cascade
G-protein

Lastly, most artificial neural networks (ANN) operate on one time scale for learning and inference, inspired by neurotransmission and dopaminergic reward prediction error theory (Schultz 1998). The brain however operates on multiple time scales (Gao et al. 2020) each of which adds an additional layer of flexibility to cognitive tasks. The advantages of incorporating multiple time scales into ANNs has been theorized (Botvinick et al. 2019) but not been taken up as a popular notion so far. To build towards similar ideas, NeMoE supports dual timescales for neurotransmitter and neuromodulator dynamics, promoting meta-learning and generalization (Fig. 4.1f). The supervisor and expert networks operate at the “fast” neurotransmitter time scale while the receptor network operates at the “slow” neuromodulator time scale.

4.3 Behavioral task and simulation design

Traditionally, lifelong learning studies use tasks like the Wisconsin Card Sorting Task or the Single/Multi Armed Bandit Task. While these tasks are capable of testing basic cognitive skills, they lack an ethological basis, causing the objective functions that the artificial agents learn to be distinct from their natural counterparts.

To overcome these limitations, we developed a novel multi-context Seasonal Foraging Task (SFT) to study naturalistic decision-making. In this setup, mice navigate an arena (Med PC rat chamber) with an insulated home area (opaque plexiglass plate) and a foraging area with a shock-delivering grid (Fig. 4.2a). The mice are first trained to retrieve a liquid reward (Ensure®) from the reward port whenever a tone (5kHz) plays (reward context). The reward is delivered 3s after the tone onset and remains in the port for the duration of the tone (total 20s) after which a vacuum system sucks it out of the port. They are then presented with sequences of different “seasonal contexts” where we

introduce randomly delivered foot shocks in the foraging area as a proxy for cost (Fig. 4.2b). These contexts alter reward and cost dynamics, with "summer" offering high rewards at low cost, "winter" providing low rewards at high cost, and "fall" combining high rewards with high costs (Fig. 4.2c). The value of the reward is modulated by changing the volume of Ensure® while the value of the cost is modulated by varying the average inter-

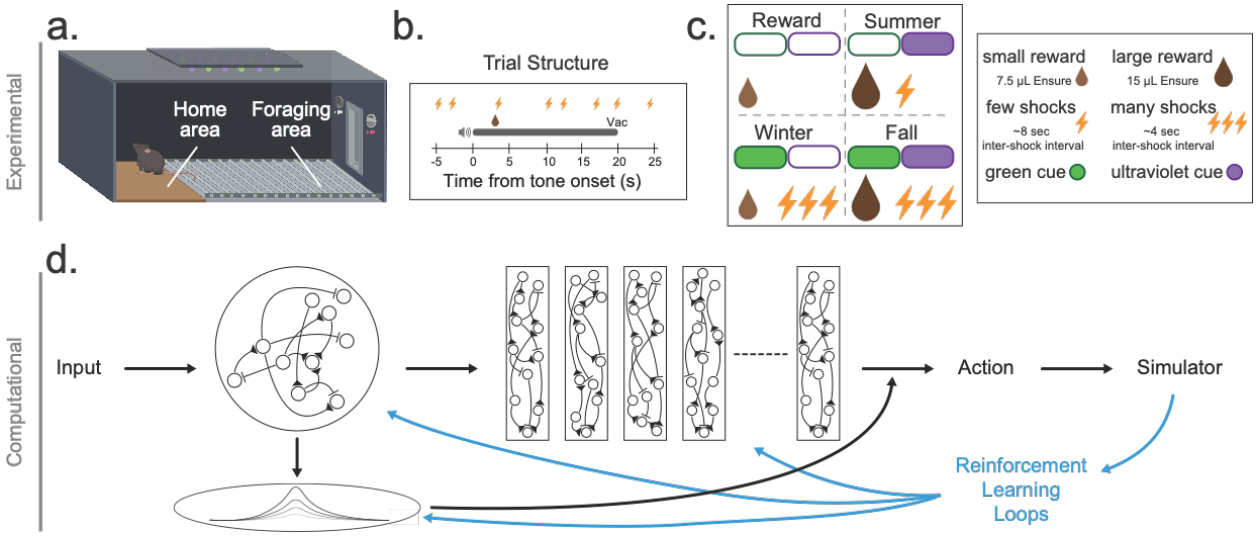


Figure 4-2 Schematics of the behavioral task along with the simulation strategy.

a, Arena for behavioral experiments. b, Trial structure. c, Seasonal contexts with varying reward and cost contingencies. d, Reinforcement learning pipeline for the simulations. Schematics created using Biorender®.

shock-interval with each shock lasting half a second. Contexts are signaled to the mice visually by a system of green and ultraviolet ambient lights (that are counterbalanced for summer and winter), chosen for their visibility to mice, to simulate distinct environments (G. H. Jacobs, Williams, and Fenwick 2004; Szatko et al. 2020).

To simulate the NeMoE architecture on the same task, we developed a custom deep reinforcement learning pipeline based on the Advantage Actor Critic (A2C) algorithm (Mnih et al. n.d.). This setup integrates the sensory and historical inputs received by the

mice into a vector input that is fed into the architecture, which then produces appropriate actions. The actions are then passed onto the simulator which provides real time feedback into the form of a reinforcement which is the scalar sum of the positive reward and negative cost. The reinforcement accumulated over multiple consecutive trials is used to update the parameters of the various modules of the architecture using backpropagation (or backpropagation through time).

4.4 Foraging reveals context specific behavioral strategies

Unimplanted adult (8-12 weeks old at the time of the first experimental procedure) wild-type male C57/BL6J mice under food restriction (to a limit of 85% baseline weight) were used to validate the behavioral task. The mice were subjected to the task over the course of a four-week behavioral paradigm, with complexity of the experiment increasing incrementally over time (Fig. 4.3a). The mice were first trained on the Reward Context for four days or until they could successfully retrieve at least 80% of the rewards within the first 10s of the tone duration. We titrate the shock magnitudes for each subject by presenting them with inescapable shocks with increasing values until we visually confirm a reaction on three successive deliveries. Following this, the subjects are presented with the foraging task with different sequences of contexts, each sequence being repeated at least twice. In each experimental session, the first context is always the Reward Context to serve as a baseline for comparability over the days.

Figure 4-3 Experimental validation of the Seasonal Foraging Task.

a, Behavioral paradigm table, each experiment has blocks of trials per context presented sequentially, R = Reward, ShkTitr = Shock Titration, S = Summer, W = Winter, F = Fall. b, (left) Port entry probability vs time relative to tone onset plotted per seasonal context for data from the RSW (11th day), RSWS (13th day), RWSW (15th day), RSFW(17th day) and RFSW (19th day) experimental sessions; (right) port entry probability mean during the baseline (time -20s to 0s) and tone (0s to 20s) periods (Two-way ANOVA (n=8 mice) Interaction: S-W: $F(1,14)=6.026$, *p=0.0278; S-W-S: $F(2,21)=5.964$, **p=0.0089, W-S-W: $F(2,21)=7.652$, **p=0.0032; S-F-W: $F(2,21)=11.64$, ***p=0.0004; Bonferroni's multiple comparison test (S-W): S-W: **p=0.0026; (S-W-S): S-W: ***p=0.0002, W-S: **p=0.0026, (W-S-W): W-S: **p=0.0048, S-W: ****p<0.0001; (S-F-W): S-F: ***p=0.0001, S-W: ****p<0.0001, F-W: ***p=0.0009). c, Behavior feature extraction pipeline using SLEAP. d, Decoding performance (measured as area under the receiver operator characteristic curve) using GLM decoder vs time relative to tone onset normalized to the shuffled control for the four contexts. e, Mean decoding performance during the tone (0s to 20s) period (Two-way ANOVA (n=10 folds) Interaction: S-W: $F(3,72)=84.57$, ****p<0.0001 Bonferroni's multiple comparison test: (R-A) ****p<0.0001; (S-A) ****p<0.0001; (F-A) ****p<0.0001, (W-A) p<0.0001).

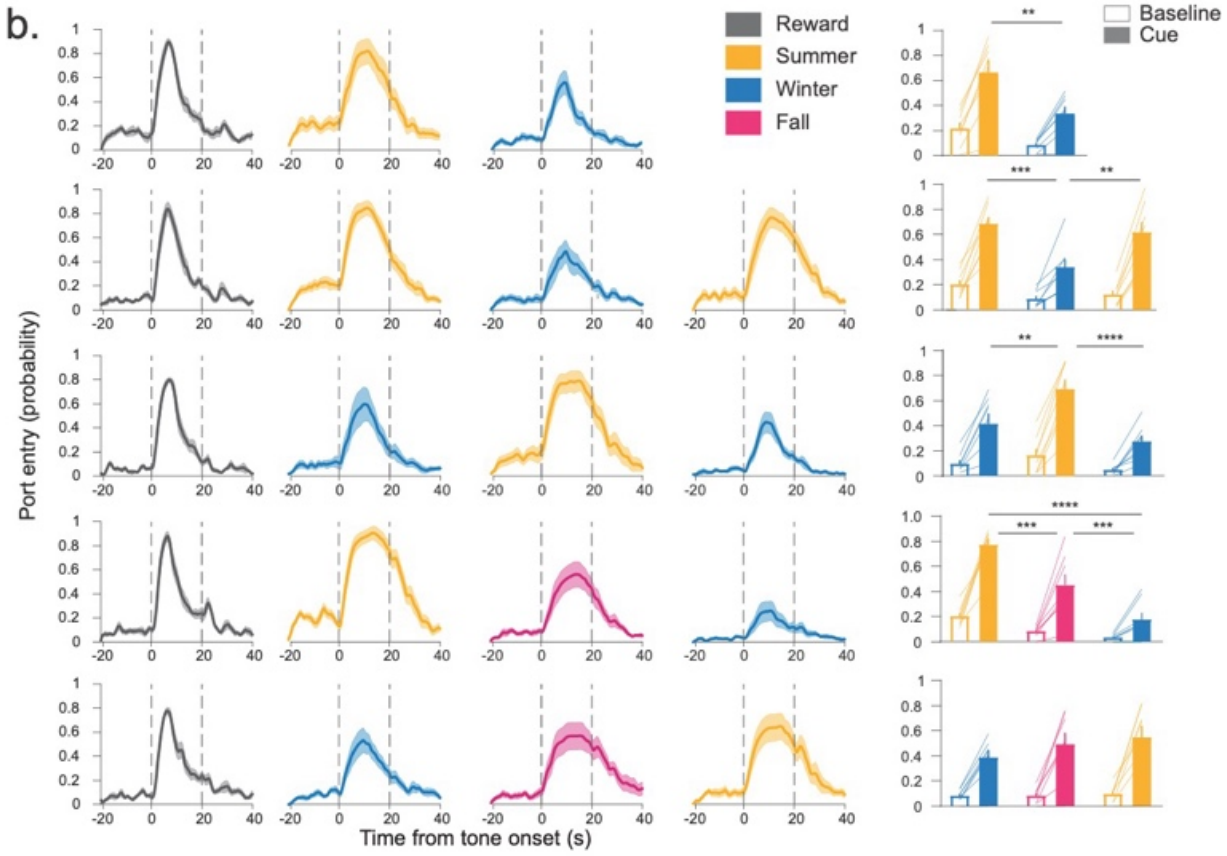
a.

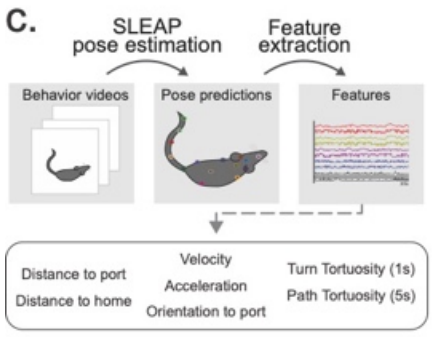
Day	1	2	3	4	5
Exp	R	R	R	R	ShkTitr

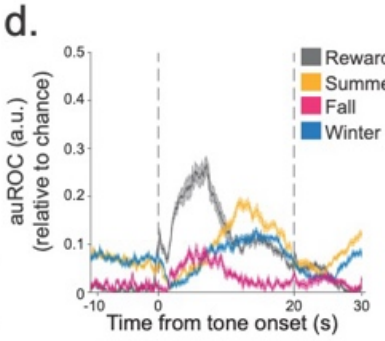
Day	6	7	8	9	10
Exp	RS	RS	RS	RSW	RSW

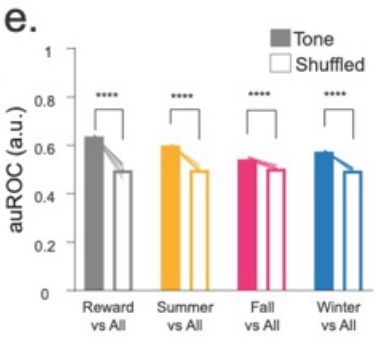
Day	11	12	13	14	15
Exp	RSW	RSWS	RSWS	RWSW	RWSW

Day	16	17	18	19	20
Exp	RSFW	RSFW	RWFS	RWFS	









We computed the trial-averaged port entry probability over the different experimental sessions (Fig. 4.3b). The four rows in Fig. 4.3b represent data from the last repetition of the corresponding context sequences (Reward-Summer-Winter, Reward-Summer-Winter-Summer, Reward-Winter-Summer-Winter, Reward-Summer-Fall-Winter and Reward-Winter-Fall-Summer). The animals exhibited context-specific port entry behavior independent of the order of the sequences with the values being proportional to the reward to cost ratio. Further, while the overall trends remain independent of the order of the contexts, there is an effect of marginal utility as observed the RWSW and RWFS experiments where the early Winter Context has relatively higher port entries while the late Summer Context has relatively lower port entries.

Using the pose estimation algorithm SLEAP, we tracked the movements of the animals as they performed the behavioral task (Pereira et al. 2022) (Fig. 4.3c). We used this data to extract a repertoire of behavioral features relevant to foraging (Fig. 4.3c). We used a Generalized Linear Model (GLM) to predict the type of context from these behavioral features and found that all four contexts are represented distinctly from each other, demonstrating that the mice have learned stereotypical behavioral strategies for each context.

4.5 DA in mPFC selectively encodes contextual information

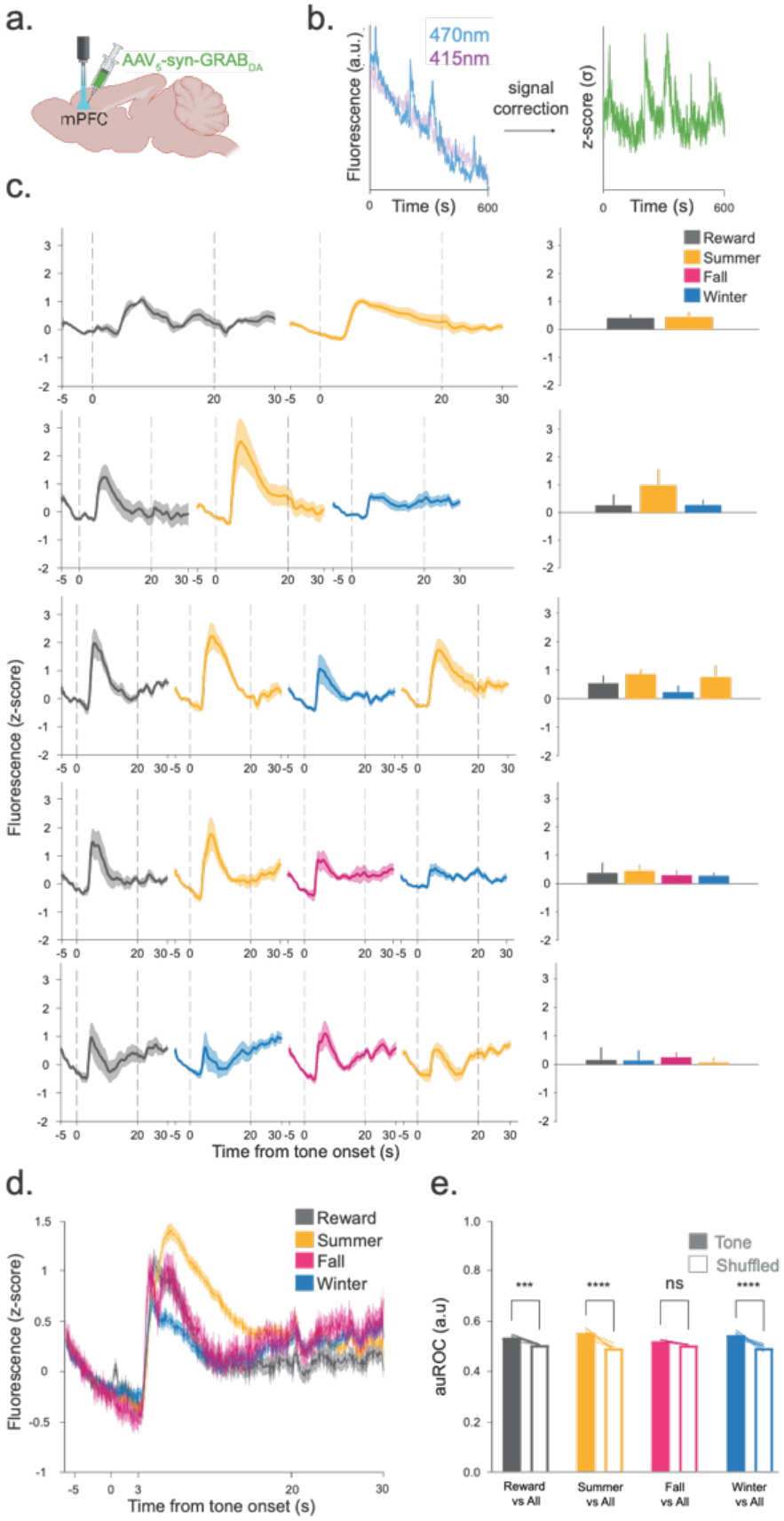
To investigate the role of dopamine input in contextual modulation of neural activity in the mPFC, we administered AAV5-syn-GRAB-DA into the mPFC of wild type male mice. We implanted an optical fiber into the mPFC after 2-4 weeks of viral expression

(Fig. 4.4a). We ran behavioral experiments using the Seasonal Foraging Task while measuring fluorescence in the mPFC at 470nm and 415nm to adjust for background illumination and other discrepancies using an isosbestic point. To extract dopamine transients in response to behavioral events, we utilized a linear signal correction method to regress the signal against the isosbestic point (Fig. 4.4b).

We computed the z-scored fluorescence over the different experimental sessions in the paradigm (Fig. 4.4c). Like in the previous section, the five rows in Fig. 4.4c represent data from the last repetition of the context sequences (Reward-Summer, Reward-Summer-Winter, Reward-Summer-Winter-Summer, Reward-Summer-Fall-Winter and Reward-Winter-Fall-Summer). The dopamine response to tone exhibited specificity to the Summer and Winter contexts with the value being proportional to the reward to cost ratio, similar to the behavioral response. The response during either the Fall context or the experiments starting with Winter was more variable, the latter being influenced by marginal utility as in behavior. To further probe this effect, we plotted the distribution of the response across all the trials/experiments (Fig. 4.4d). We found that while the response to Reward, Summer and Winter contexts followed prototypic profiles, the response to Fall overlapped significantly with the others. We used a GLM to predict the type of context from dopamine response and found that all the contexts apart from Fall were decodable.

Figure 4-4 Context representation in dopamine response.

a, Schematic showing viral strategy to express GRABDA in the mPFC and measure dopamine signal using fiber photometry. b, Signal correction pipeline using isosbestic control. c, (left) Dopamine signal (z-scored) vs time relative to tone onset plotted per seasonal context for data from the RS (8th day), RSWS (13th day), RSFW(17th day) and RFSW (19th day) experimental sessions; (right) port entry probability mean during the tone (0s to 20s) period (n = 4 mice). d, Dopamine signal (z-scored) pooled over all trials/experiments vs time relative to tone onset plotted per seasonal context. e, Mean decoding performance (measured as area under the receiver operator characteristic curve) using GLM decoder vs time relative during the tone (0s to 20s) period for the four contexts (Two-way ANOVA (n=5 folds) Main Effect of Decoding: $F(1,32)=140.8$, ****$p<0.0001$ Bonferroni's multiple comparison test: R: ***$p<0.001$; S: ****$p<0.0001$; W: ****$p<0.0001$).


a.
AAV5-syn-GRABDA
mPFC
b.
470nm
415nm
Fluorescence (a.u.)
Time (s)
signal correction
z-score (σ)
c.
Reward
Summer
Fall
Winter
Fluorescence (z-score)
Time from tone onset (s)
d.
Reward
Summer
Fall
Winter
Fluorescence (z-score)
Time from tone onset (s)
e.
Tone
Shuffled
auROC (a.u)
***
****
ns
****
Reward vs All
Summer vs All
Fall vs All
Winter vs All

4.6 mPFC ensembles follow expert-like behavior

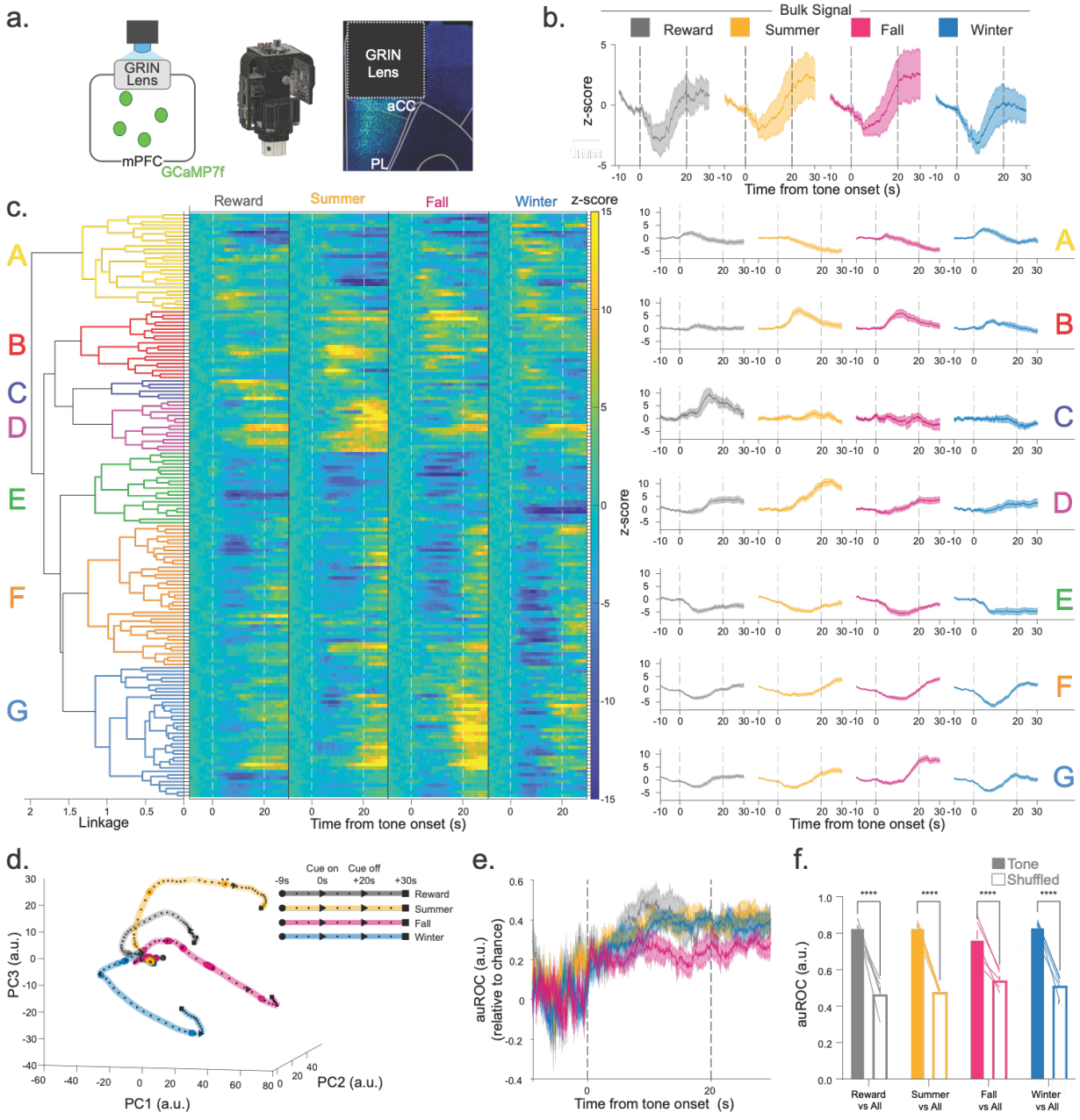
Next, we sought to investigate the representation of the SFT contexts in the neuronal activity in the mPFC. We injected AAV1-syn-GCaMP7f into the mPFC of mice and subsequently implanted a gradient refractive index (GRIN) lens (Fig. 4.5a). Around 4 weeks post-implantation, a baseplate was fixed at an optimal focus height for a miniature microscope. We ran behavioral experiments 6-8 weeks after viral injection, using the paradigm depicted in Fig. 4.3a while recording calcium indicator fluorescence from the mPFC. We used NoRMCorre (Pnevmatikakis and Giovannucci 2017) to correct motion artifacts in the imaging data. We computed the z-scored bulk neural activity in response to the tone onset using the entire field-of-view in the imaging data for the four contexts from the final Reward-Summer-Fall-Winter (17th day) and found that the signal was largely similar (Fig. 4.4b).

Further, we used CNMF-E (Zhou et al. 2018) to identify single units from the same dataset (day 17th) and extract their activity. We applied a hierarchical agglomerative clustering (HAC) method on the z-scored activity of the single units developed in house to identify ensembles of units with similar patterns of activations (Fig. 4.4c). Unlike the bulk signal, we were able to identify ensembles with unique responses to the tone onset across the four contexts. In response to the reward delivery (3s after tone onset), Ensemble B shows an excitatory effect during Summer & Fall while ensemble C does the same during Reward. Ensembles D and G show an excitatory response towards the middle to end of the tone period, presumably when the reward retrieval occurs. Lastly, ensemble F shows an inhibitory response to the tone onset in all four contexts with a higher affinity for Winter.

While HAC helps identify groups of individual units with similar activation patterns, it does not directly provide information about the population level coding. To uncover that, we used principal component analysis to reduce the dimensionality of the neural activity to a 2D space and found that the population activity trajectories for the four contexts have a similar shape but occupy distinct neural subspaces (Fig. 4.4d). To quantify this separation, we used a GLM decoder to predict the type of context from neural activity and were able to do so better than chance (Fig. 4.4e, f).

Figure 4-5 Context representation in mPFC neural response.

a, Schematic showing the viral strategy to express GCaMP7f in the mPFC and miniscope setup for one photon calcium imaging of neurons. b, Bulk neuronal signal (z-scored) vs time relative to tone onset plotted per seasonal context for data from the RSFW (17th day) experimental sessions. c, Left, Heatmaps for mPFC single unit activity (z-scored) vs time relative to tone onset along with dendrogram from the hierarchical clustering output. Right, Mean response (z-scored) of the single units from each ensemble vs time relative to tone onset plotted per seasonal context (n = 2 mice, 150 units). e, Decoding performance (measured as area under the receiver operator characteristic curve) using GLM decoder vs time relative to tone onset normalized to the shuffled control for the four contexts. f, Mean decoding performance during the tone (0s to 20s) period (Two-way ANOVA (n=5 folds) Main Effect of Decoding: $F(1,32)=239.7$, ****$p<0.0001$ Bonferroni's multiple comparison test: R: ****$p<0.0001$; S: ****$p<0.0001$; F: ****$p<0.0001$; W: ****$p<0.0001$).


a.
GRIN Lens
mPFC
GCaMP7f
GRIN Lens
aCC
PL
b.
Bulk Signal
Reward
Summer
Fall
Winter
z-score
Time from tone onset (s)
c.
Reward
Summer
Fall
Winter
z-score
A
B
C
D
E
F
G
Linkage
Time from tone onset (s)
d.
Cue on
Cue off
-9s
0s
+20s
+30s
Reward
Summer
Fall
Winter
PC1 (a.u.)
PC2 (a.u.)
PC3 (a.u.)
e.
auROC (a.u.)
(relative to chance)
Time from tone onset (s)
f.
Tone
Shuffled
auROC (a.u.)
Reward vs All
Summer vs All
Fall vs All
Winter vs All

4.7 Simulations with NeMoE generate diverse strategies resembling behavior

In order to simulate the experimental paradigm with the NeMoE architecture, we first created a simulator for a simplified version of the SFT. In this implementation, each trial is reduced to a single Go/No-go decision based on a set of discrete inputs for the tone, context cues, reward history and action history per trial (Fig. 4.6a). We adapted the loss function detailed in Mnih et al (Mnih et al. n.d.) with an additional term for the regularization of the receptor layer of the network, in order to impose sparsity in the connections (Fig. 4.6b) – inspired directly by the sparse non-uniform distribution of neuromodulatory receptors in the mPFC as shown in Fig. 4.1d. In any individual simulation, an agent with the NeMoE architecture is initialized with two neuromodulators and three experts with randomly sampled parameters. The agent was presented with the same experimental paradigm as in Fig. 4.3a using deep reinforcement learning simulations.

Due to the stochastic nature of the simulations, we found that the agent learned different neural strategies to solve the task. Fig. 4.6c shows the neuromodulatory signals evolving over the course of an example simulation simulation along with the mean expert weights in the last four contexts. Due to the random initialization, the two neuromodulators seemed to follow the same response profile for all contexts but after repeated exposure to the different context sequences of the task the signals converge to context specific profiles similar to what was observed with dopamine in Fig. 4.4c. Further, the weights

assigned to the experts also converge to fixed values and describe the relative importance of the experts across the contexts.

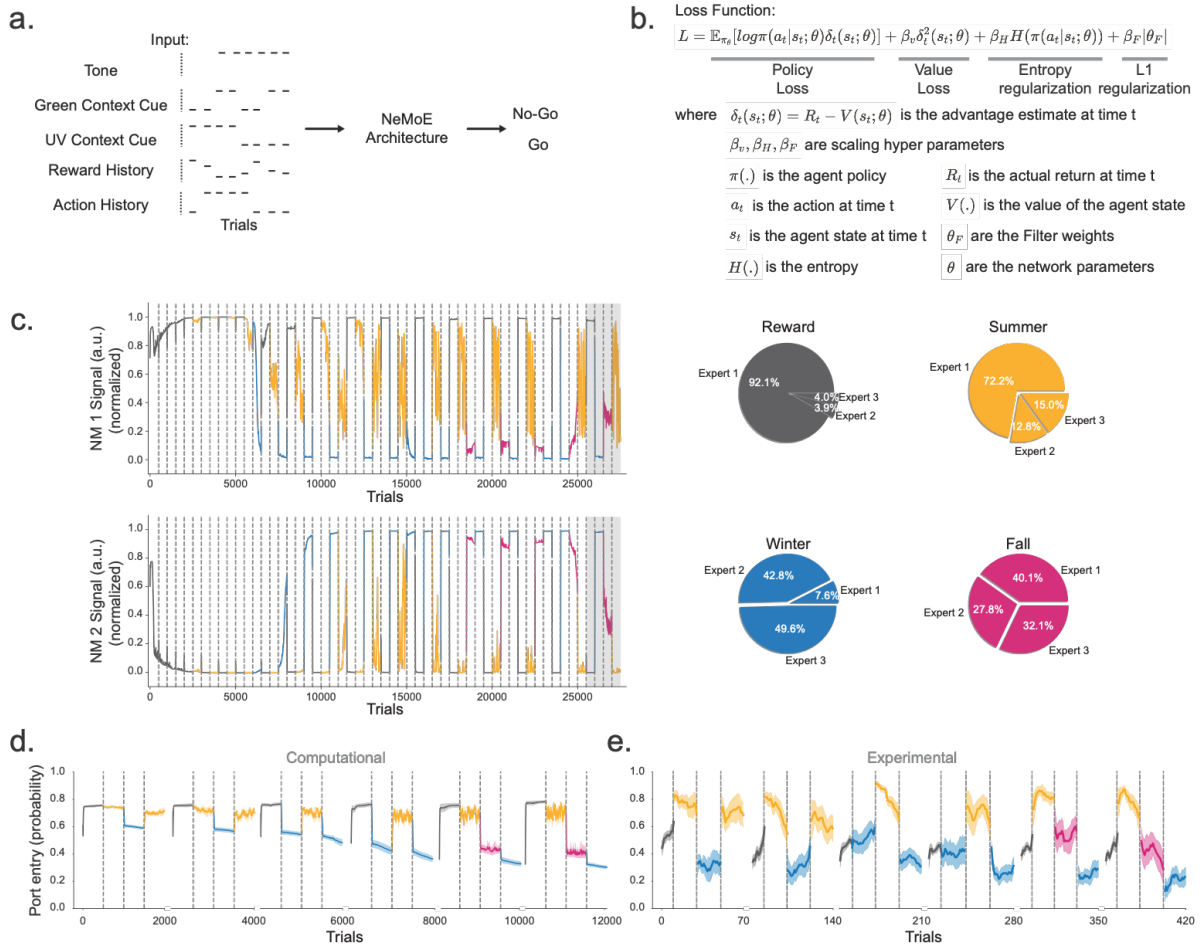


Figure 4-6 Deep reinforcement learning simulations on the Go/No-Go SFT.

a, Schematic showing details of the simulator. b, Loss function used to optimize the network. c, Left, neuromodulatory signals vs trial number. Right, mean weights assigned by the supervisor to the experts during the last four contexts in the paradigm. d, Port entry probability vs trial number for the NeMoE architecture (n = 5 simulations). e, Port entry probability vs trial number for the behavioral data (n = 8 animals).

Lastly, we ran five independent simulations where the neural network learned similar patterns of activations for the SFT. However, despite this diverse set of neural responses, the agent converges to a stable behavioral response (Fig. 4.6d). When

compared to the real data, the PFC inspired NeMoE architecture mimics biological behavior well (Fig. 4.6e).

4.8 Discussion and future work

In this study, we have attempted a multi-modal approach to develop a novel neural network architecture to solve the behemoth of a problem that is lifelong learning. The Seasonal Foraging Task offers a high dimensionality and ethological grounding that has largely been absent from typical tests found in literature. The behavioral readout from animal experiments provides robust evidence that there exist non-overlapping behavioral strategies that a natural agent would employ for the different contexts of the task.

The fiber photometry experiments show that dopaminergic inputs to the mPFC attain context-specific response profiles. The decodability of context from a one-dimensional bulk signal is indicative that despite low sensitivity, the signal is quite specific to the type of context. While Reward, Summer and Winter are decodable, the lack of Fall decoding is also an intriguing finding. One hypothesis could be that dopamine might be encoding “unique” contexts as Fall can be considered a combination of Summer and Winter. Another hypothesis is that one of the other neuromodulators active in the mPFC might be encoding for such conflicted contexts, encouraging future experiments with sensors for them. In alignment with this evidence, the optogenetic stimulation results from Vander Weele et al. (Vander Weele et al. 2018) published previously in our lab show that specific mPFC subpopulations are impacted by dopamine in a context-specific manner. Together, this forms the foundation for our central hypothesis that a neuromodulation driven subpopulation signal filtering is one of the mechanisms that enable the PFC to perform lifelong learning.

Further, the neural activity recorded from the one photon calcium imaging experiments show single units that can be clustered into ensembles that have context specific response profile as well. This data serves as evidence for the existence of “expert”-like ensembles in the PFC playing context-specific roles. The decodability of context from the population neural activity shows that the signals of the neurons interact within and across ensembles to encode information relevant to each context - explaining the need to use recurrent neural networks to model the experts. Lastly, the simulations with the Go/No-go version of the task show that the architecture is capable of mimicking natural agents both at the neural and behavioral level while learning the task.

This “binary” simplification of the task makes the simulations comparable to studies conducted on conventional tasks like the Wisconsin Card Sorting or Multi Armed Bandit. However, it also under-utilizes the high dimensionality of the task. To leverage it fully, a multi-step simulation where each trial is a timeseries and the actions space is a one- or two-dimensional arena would serve as an appropriate and highly relevant next step. It would show case the ability of the model to scale with the dimensionality of the task with limited increase in the number of parameters. It would also enable more granular comparisons with the neural and behavioral data, both of which are timeseries’ with high temporal resolution. Another avenue for future work is the scaling up of experiments to include more animals to be able to show significance of the trends that were observed along with probing the function of other neuromodulators such as serotonin, norepinephrine and acetylcholine.

This study lays the groundwork for this extension as well as future research where multiple experimental modalities are integrated to create bio-realistic models of the brain.

It also creates an architecture that serves as a computational testbed for various experimental paradigms associated with the PFC including studies focusing on neuropsychiatric illnesses such as depression and schizophrenia which are highly relevant to dysfunction in this brain region.

4.9 Bibliography

"Allen Brain Atlas." 2004.

Botvinick, Matthew, Sam Ritter, Jane X. Wang, Zeb Kurth-Nelson, Charles Blundell, and Demis Hassabis. 2019. "Reinforcement Learning, Fast and Slow." *Trends in Cognitive Sciences* 23(5): 408–22. doi:10.1016/j.tics.2019.02.006.

Gao, Richard, Ruud L. van den Brink, Thomas Pfeffer, and Bradley Voytek. 2020. "Neuronal Timescales Are Functionally Dynamic and Shaped by Cortical Microarchitecture." *eLife* 9: e61277. doi:10.7554/eLife.61277.

Jacobs, Gerald H., Gary A. Williams, and John A. Fenwick. 2004. "Influence of Cone Pigment Coexpression on Spectral Sensitivity and Color Vision in the Mouse." *Vision Research* 44(14): 1615–22. doi:10.1016/j.visres.2004.01.016.

Jacobs, Robert A., Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. 1991. "Adaptive Mixtures of Local Experts." *Neural Computation* 3(1): 79–87. doi:10.1162/neco.1991.3.1.79.

Janak, Patricia H., and Kay M. Tye. 2015. "From Circuits to Behaviour in the Amygdala." *Nature* 517(7534): 284–92. doi:10.1038/nature14188.

Kudithipudi, Dhireesha, Mario Aguilar-Simon, Jonathan Babb, Maxim Bazhenov, Douglas Blackiston, Josh Bongard, Andrew P. Brna, et al. 2022. "Biological Underpinnings for Lifelong Learning Machines." *Nature Machine Intelligence* 4(3): 196–210. doi:10.1038/s42256-022-00452-0.

Miller, E. K., and J. D. Cohen. 2001. "An Integrative Theory of Prefrontal Cortex Function." *Annual Review of Neuroscience* 24: 167–202. doi:10.1146/annurev.neuro.24.1.167.

Mnih, Volodymyr, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Tim Harley, Timothy P Lillicrap, David Silver, and Koray Kavukcuoglu. "Asynchronous Methods for Deep Reinforcement Learning."

Oberländer, Marcel. 2012. "Beyond the Cortical Column."

Pereira, Talmo D., Nathaniel Tabris, Arie Matsliah, David M. Turner, Junyu Li, Shruthi Ravindranath, Eleni S. Papadoyannis, et al. 2022. “SLEAP: A Deep Learning System for Multi-Animal Pose Tracking.” *Nature Methods* 19(4): 486–95. doi:10.1038/s41592-022-01426-1.

Pnevmatikakis, Eftychios A., and Andrea Giovannucci. 2017. “NoRMCorre: An Online Algorithm for Piecewise Rigid Motion Correction of Calcium Imaging Data.” *Journal of Neuroscience Methods* 291: 83–94. doi:10.1016/j.jneumeth.2017.07.031.

Schultz, Wolfram. 1998. “Predictive Reward Signal of Dopamine Neurons.” *Journal of Neurophysiology* 80(1): 1–27. doi:10.1152/jn.1998.80.1.1.

Szatko, Klaudia P., Maria M. Korympidou, Yanli Ran, Philipp Berens, Deniz Dalkara, Timm Schubert, Thomas Euler, and Katrin Franke. 2020. “Neural Circuits in the Mouse Retina Support Color Vision in the Upper Visual Field.” *Nature Communications* 11(1): 3481. doi:10.1038/s41467-020-17113-8.

Tsuda, Ben, Kay M. Tye, Hava T. Siegelmann, and Terrence J. Sejnowski. 2020. “A Modeling Framework for Adaptive Lifelong Learning with Transfer and Savings through Gating in the Prefrontal Cortex.” *Proceedings of the National Academy of Sciences* 117(47): 29872–82. doi:10.1073/pnas.2009591117.

Vander Weele, Caitlin M., Cody A. Siciliano, Gillian A. Matthews, Praneeth Namburi, Ehsan M. Izadmehr, Isabella C. Espinel, Edward H. Nieh, et al. 2018. “Dopamine Enhances Signal-to-Noise Ratio in Cortical-Brainstem Encoding of Aversive Stimuli.” *Nature* 563(7731): 397–401. doi:10.1038/s41586-018-0682-1.

Wang, Jane X., Zeb Kurth-Nelson, Dharshan Kumaran, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Demis Hassabis, and Matthew Botvinick. 2018. “Prefrontal Cortex as a Meta-Reinforcement Learning System.” *Nature Neuroscience* 21(6): 860–68. doi:10.1038/s41593-018-0147-8.

Zhou, Pengcheng, Shanna L Resendez, Jose Rodriguez-Romaguera, Jessica C Jimenez, Shay Q Neufeld, Andrea Giovannucci, Johannes Friedrich, et al. 2018. “Efficient and Accurate Extraction of in Vivo Calcium Signals from Microendoscopic Video Data” ed. David C Van Essen. *eLife* 7: e28728. doi:10.7554/eLife.28728.

4.10 Acknowledgements

Chapter 4, in full, is currently being prepared for submission for publication of the material. Kanha Batra, Anousheh Bakhti-Suroosh, Clara N. Yi, May G. Chan, Ruby Tseng, Kelbi Banducci, Jiaqi Zhang, Sophie E. Poelmann, Christopher R. Lee, Tanisha

Roy, Kelly Chang, Tristan B. Tuazon, Romy Wichmann, Caroline Jia, Ben Tsuda, Bryan Nielsen, Laurel R. Keyes, Yulong Li, Talmo Pereira, Terrence J. Sejnowski, Kay M. Tye. Neuromodulated mixture of experts: A prefrontal cortex inspired architecture for lifelong learning. *In preparation*, 2024. The dissertation author was the primary investigator and author of this paper.

Chapter 5 Conclusion

Nature has perfected intelligence over millennia, each evolutionary iteration propelling organisms to survive and thrive in an ever-changing environment. The epitome of this process is the emergence of lifelong learning – the concept of learning and retaining knowledge of multiple contexts, switching between them seamlessly and use past knowledge to facilitate new learning. To truly capture this in a model, it is imperative to develop the computational architectures while studying biological systems at the micro and macro scales. Unfortunately, there is a wide gap between the latest and greatest machine learning algorithms and our understanding of behavioral and systems neuroscience. This causes our machines to learn to optimize objective functions that aren't ethological, which I believe is the main reason why we still haven't achieved lifelong learning in our systems.

Via the medium of this dissertation, I have attempted to make a small contribution to bridge this gap by using machine learning methods to identify the neural basis of contextual decision-making and in turn use this basis as an inductive bias to develop better computational architectures. While I mainly focused on the prefrontal cortex as the central piece of my modeling venture, it is but a piece of a bigger, more complex puzzle that is the brain.

I first look at how the mPFC assesses social context and modulates behavior as animals participate in their naturally established social hierarchies using state space analysis (Chapter 2). It reveals an internal state like representation of key variables such as (relative) social rank and outcome of a reward competition assay. This indicates that a latent state system might be working at the heart of mPFC neural computations. Inspired

by this, I adapt a classical architecture to model these latent states as well as be able to dynamically predict behavior from evolving neural population activity in the mPFC: the Hidden Markov Model Generalized Linear Model (HMM GLM). By letting the data decide the appropriate number of latent states along with influencing the probabilities of state transition and behavior prediction at the millisecond time resolution, the model was able to explain the experimental data with high precision. This suggests a need for specialized secondary mechanisms to modulate neural activity while mapping it to behavior.

I then look at how computations in the mPFC are impacted due to disruptions caused by a chronic mild stress paradigm (Chapter 3). I first use unsupervised clustering to identify naturally forming categories for animals based on their anhedonic behavior – suggesting that there are several gradations of anhedonia instead of it just being binary which is unfortunately how current treatments are planned for. I further use a combination of state space analysis and neural decoding techniques to study how the neural population representation of valence is distinct in hedonic and anhedonic animals, and the evolution of this distinction over the course of the paradigm. This suggests that cognitive dysfunction can cause drastic changes in population level representation of learned environmental variables – which might be a compensatory or affected mechanism of the brain.

Based on these findings, I develop a novel modular neural network architecture to model the PFC – Neuromodulated Mixture of Experts (NeMoE). The primary work horse of the architecture is an adapted Mixture of Experts (MoE) module where a distribution of experts can be selected at any time to weigh in on a decision. This “filtering” effect is enabled by a secondary neuromodulatory network which mediates the signal to noise

ratio of different experts to produce the most relevant combination of subpopulations to handle varied tasks and contexts. The modularity also enables the architecture to simulate various diseased states (such as anhedonia) and study them with greater granularity than before. Unlike the HMM GLM which is capable of decoding behavior from given neural activity, NeMoE takes it a step further to simulate the neural activity dynamics themselves. I develop a task with elements of lifelong learning based on contextual foraging. I first run behavioral and systems experiments with mice on the task to validate the model behaviorally as well as provide evidence for why this particular architecture is suitable for modeling the prefrontal cortex. Lastly, I compare the model with the data to find stark similarities between the two in terms of behavioral and neural dynamics, representation of task variables and even inter-animal differences. This demonstrates that developing artificial intelligence models in conjunction with biological experiments is an immensely powerful and underrated methodology to mimic ethological objective functions with biological plausibility.

While the scope of this dissertation is limited to these chapters, it certainly opens up a large number of avenues to pursue in order to further our understanding of natural and artificial intelligence. A general next step would be to incorporate several time scales of neural processing, multiple brain regions with inter-region interactions and using inverse reinforcement learning methods to further probe ethological objective functions. Developing models with this methodology would also create computational testbeds to simulate future experimental paradigms to help narrow down the parameter spaces of interest. Lastly, by simulating perturbations in a “healthy” or “normal” system, we can

create models of diseased states in human brains to be able to provide precise and individualized therapies.