Intelligent behavior relies on recognizing structure in the environment. Rather than treating each situation as unique, the brain extracts rules that link diverse experiences – a process called generalization. This ability depends on cognitive maps: internal representations encoding abstract relationships between environmental features and behavioral outcomes, enabling flexible transfer to novel situations[1]. Cognitive map formation involves a division of labor across brain regions: the hippocampus builds maps of specific task environments and behavioral consequences[2], while the medial prefrontal cortex (mPFC) encodes abstract structures that generalize across contexts[3]. **Yet, how mPFC neural population activity reorganizes during learning to extract these abstract structures remains uncharacterized at the cellular level.**

I propose to combine large-scale, cellular-resolution imaging with machine learning to **characterize how mPFC neural population dynamics evolve during learning (Aim 1) and identify the computational principles underlying biological generalization (Aim 2).** Drawing on my experience in physics, computational modeling, optical engineering, and systems neuroscience, I am uniquely equipped to integrate advanced imaging with quantitative methods to address how prefrontal circuits achieve structured generalization.

Intellectual Merit

Currently, researchers are attempting to understand PFC function during generalization by studying well-trained animals using electrophysiological recordings or limited imaging approaches. Progress has been hindered as electrophysiology lacks the ability to track stable cell identity across weeks of learning, while conventional imaging has typically captured only well-trained neural activity, missing critical dynamics. Using the state-of-the-art two-photon random access mesoscope (2P-RAM)[4] in the Sun Lab at Cornell, I will **longitudinally record from thousands of stably identified mPFC neurons** (Fig. 1a) **to track population dynamics over several weeks as animals learn to generalize task structure across sensory modalities** (Fig. 2; Aim 1). This cross-modal design tests whether mPFC encodes abstract task rules versus sensory-specific information. Through an existing collaboration with Prof. Kevin Ellis (Cornell Computer Science)[5], I will train a machine learning model to identify the computational principles underlying mPFC encoding strategies (Aim 2). **Hypothesis:** the mPFC will encode both abstract structure and modality-specific information through distinct subpopulations, producing a spectrum of tuning properties with some neurons showing modality-invariant responses while others retain sensory specificity (Aim 1). Computational models with intermediate network sparsity will best replicate these mPFC representations across learning stages, providing insight into the mechanisms underlying biological generalization (Aim 2).

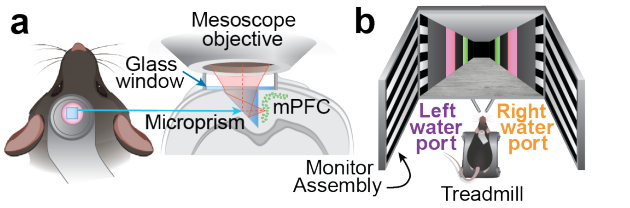

Fig. 1. Experimental setup for imaging mPFC during virtual reality behavior. (a) Chronic microprism implantation enables two-photon imaging of mPFC through a glass cranial window. (b) Head-fixed mouse navigates a linear virtual reality corridor on a treadmill with left and right water ports for reward delivery.

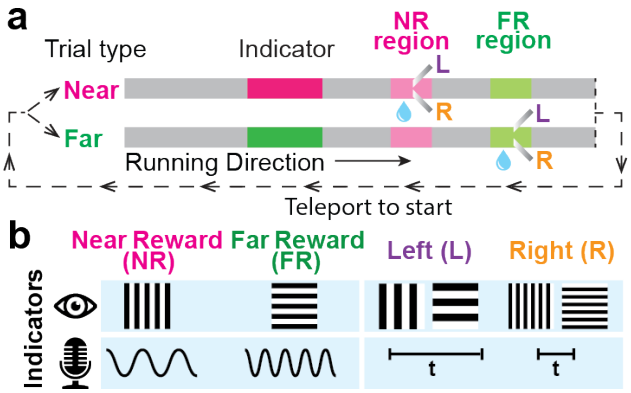

Fig. 2. Cross-modal task design. (a) Abstract task structure with two decision variables: reward location (Near vs. Far) and lick direction (Left vs. Right). (b) Sensory indicators implementing identical task structure across visual and auditory modalities.

Aim 1. Characterize population dynamics in mPFC during learning and generalization. Head-fixed transgenic mice expressing calcium indicator GCaMP8s will navigate a linear virtual reality corridor while I perform 2P-RAM imaging of mPFC activity through a chronic microprism implantation (Fig. 1)[6]. I will train mice on tasks with identical abstract structure (Fig. 2a) but different sensory modalities: vision and audition (Fig. 2b). In the visual task, stripe orientation (vertical vs. horizontal) determines reward location (near vs. far) while pattern width (thick vs. thin) determines lick direction (left vs. right). In the auditory task, tone pitch (low vs. high) determines reward location while tone duration (long vs. short) determines lick

direction, with visual displays disabled. After learning the initial task, mice will learn the task in the opposing sensory modality. I will use dimensionality reduction techniques (UMAP, Γ-VAE[7]) and representational similarity analysis[8] to track population reorganization across learning stages. **Hypothesis: mPFC will encode both abstract structure and modality-specific information through distinct subpopulations.** This predicts a spectrum of tuning properties which evolve over the course of the learning process, with some neurons showing modality-invariant responses while others retain sensory specificity. Using generalized linear models[3], I will decode abstract vs. sensory contributions for each cell, revealing which task dimensions merge first during learning. If cross-modal transfer is incomplete, I will examine intermediate training stages to identify when and how modality-specific representations emerge, testing whether additional training promotes abstraction. **Guaranteed outcome:** this approach will provide the first large-scale, cellular-resolution characterization of prefrontal population dynamics during cross-modal generalization, establishing foundational principles for how neural circuits implement flexible learning.

Aim 2. Decode task structure from mPFC dynamics using machine learning models. To identify the computational principles underlying cross-modal generalization in mPFC, I will leverage existing collaborations between Cornell's Neurobiology and Computer Science departments[5]. I will train a multimodal recurrent neural network (RNN) on the same task structure presented to the mice (Fig. 2) using visual or auditory stimulus input to predict output behavioral choices (reward location, lick direction). I will analyze the representations learned in the mPFC and track alignment of recorded neural activity (Aim 1) with learned model representations (Aim 2). RNNs are particularly suited for this investigation because their recurrent dynamics flexibly implement context-dependent transformations[9] - the core computation required for this task. I will systematically vary network hyperparameters (learning rate, connectivity strength, input representations) to identify regimes replicating mPFC learning stages. **Hypothesis: network sparsity will be the critical parameter controlling cross-modal generalization capabilities.** Drawing on recent findings that soft winner-takes-all (sWTA) mechanisms promote decorrelated cognitive maps[2], I will modulate sparsity by systematically varying sWTA inhibition strength. Critically, mPFC and hippocampus employ distinct representational strategies. While the hippocampus exhibits strong global inhibition driving completely separated representations for similar contexts[2], the mPFC exhibits weaker inhibition and maintains overlapping, distributed representations even for distinct task states[10]. **This predicts that optimal mPFC-like representations will occupy an intermediate sparsity regime** - enabling both abstract structure extraction and modality-specific encoding. **Guaranteed outcome:** I will produce a publicly available computational framework and associated dataset to serve as a benchmark enabling the comparison of various artificial neural network models of mPFC abstraction.

Broader Impacts

Understanding prefrontal mechanisms for data-efficient generalization directly addresses a critical challenge in AI development: despite massive datasets and computational resources[11,12], current AI systems exhibit "jagged intelligence" - excelling at complex tasks while unpredictably failing at simpler ones that lie just outside the training distribution. This project—a collaboration between the Neuroscience and Computer Science departments at Cornell - exemplifies interdisciplinary approaches crucial for American leadership in AI. I will advance open science by making all code and data available on GitHub, publishing in open-access journals, and uploading preprints to bioRxiv, facilitating rapid knowledge transfer across academia and AI industries. To broaden STEM participation, I am developing a four-session course through Cornell's Graduate Student School Outreach Program (GRASSHOPR) introducing neuroscience and microscopy concepts to local middle and high school students (*Personal Statement*). To support this project in the Sun lab, I will recruit two to three undergraduate researchers annually through programs supporting underrepresented students (*Personal Statement*), providing comprehensive training in two-photon microscopy, computational analysis, and behavioral neuroscience - mentoring them through all research stages from experimental design to manuscript preparation.

References: [1] E. C. Tolman, Psychol. Rev. (1948). [2] W. Sun, et al., Nature (2025). [3] M. El-Gaby, et al., Nature (2024). [4] N. J. Sofroniew, et al., eLife (2016). [5] A. Noviello, et al., arXiv preprint (2025). [6] R.J. Low, et al., Proc. Natl. Acad. Sci. U.S.A. (2014). [7] J. Z. Kim, et al., arXiv preprint (2024). [8] H. Nili, et al., PLOS Comput. Biol. (2014). [9] V. Mante et al., Nature (2013). [10] M. Rigotti, et al., Nature (2013). [11] J. Kaplan, et al., arXiv (2020). [12] OpenAI, "Announcing The Stargate Project" (2025).