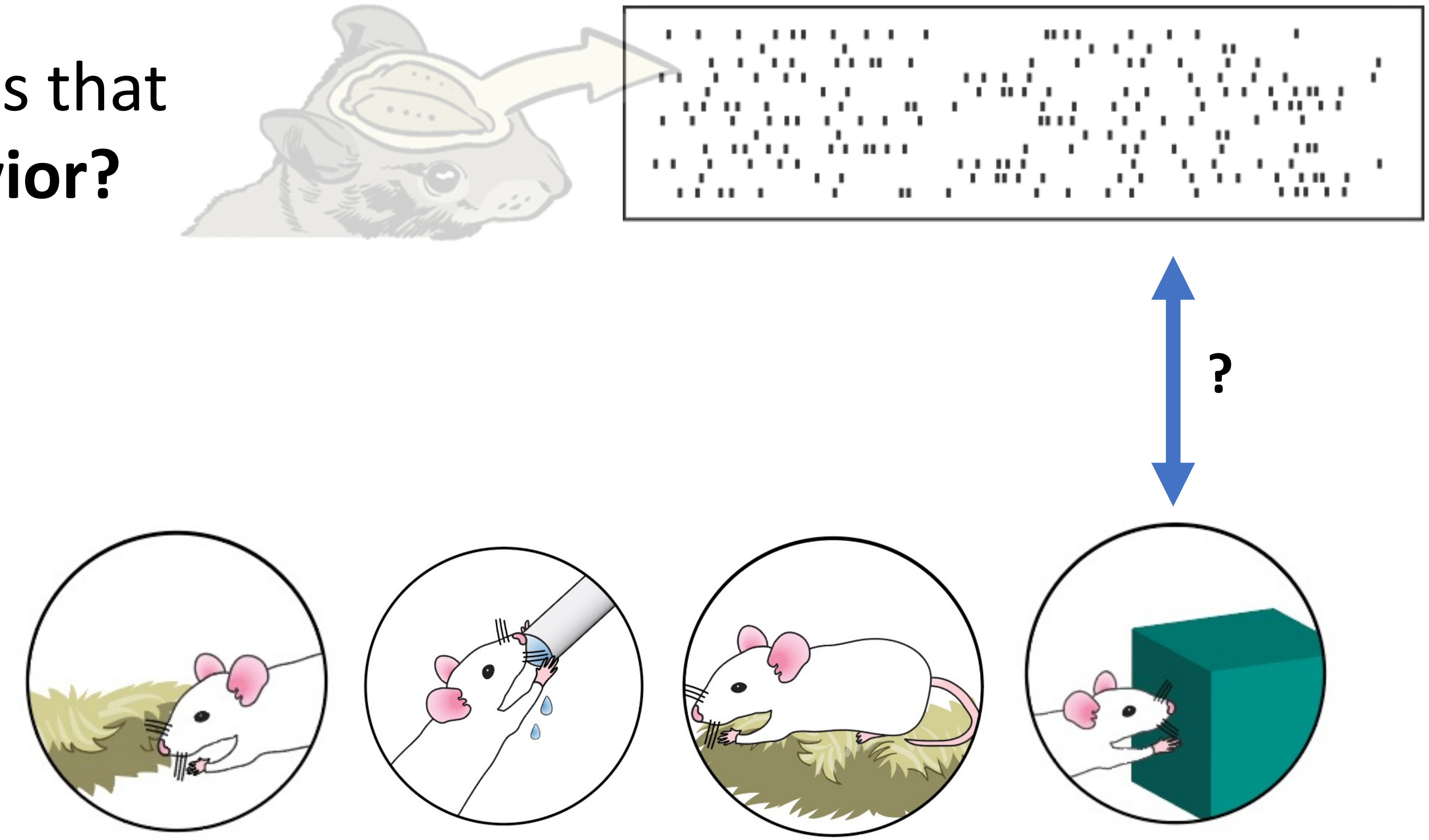
Learning representations of neural states through self-supervised learning

Eva Dyer, Georgia Tech

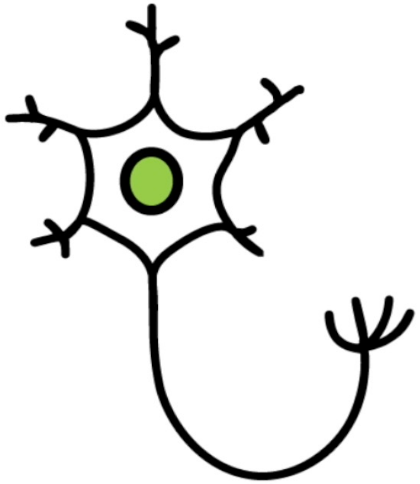
Caltech Data Science and AI for Neuroscience Summer School: July 14, 2022

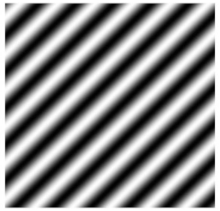
Aims

How can we build models that **link the brain and behavior?**



Challenges



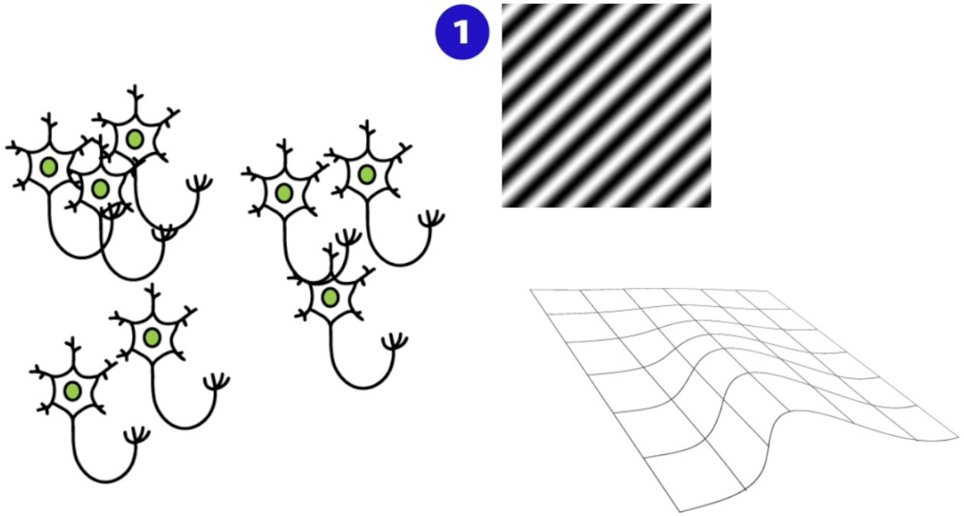


Neural responses are highly variable!

Even when behavior is stable, the responses of **individual neurons** appears stochastic and irregular.

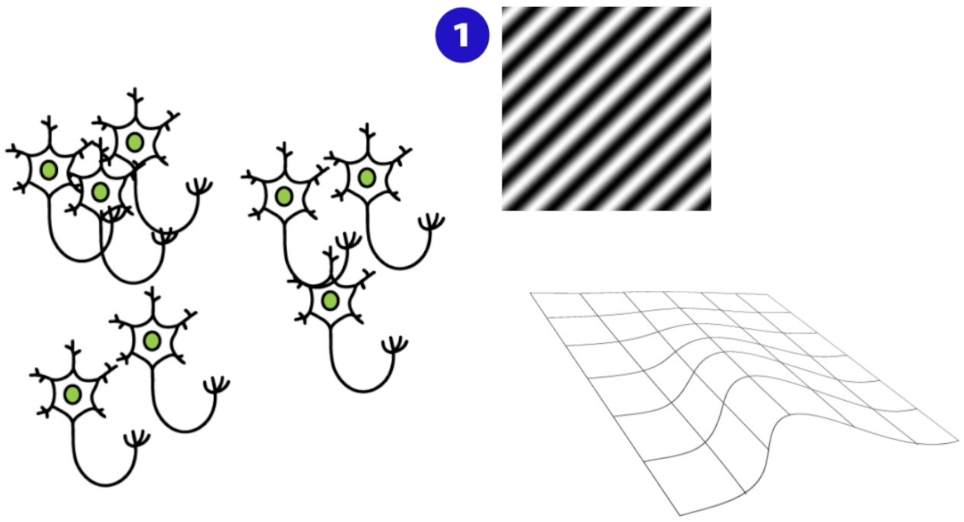
Building stable maps between the brain and behavior becomes even more challenging with complex behavior!

Challenges



When viewed at the **population level**,
neural responses may be **far more stable!**

Challenges



When viewed at the **population level**,
neural responses may be **far more stable!**

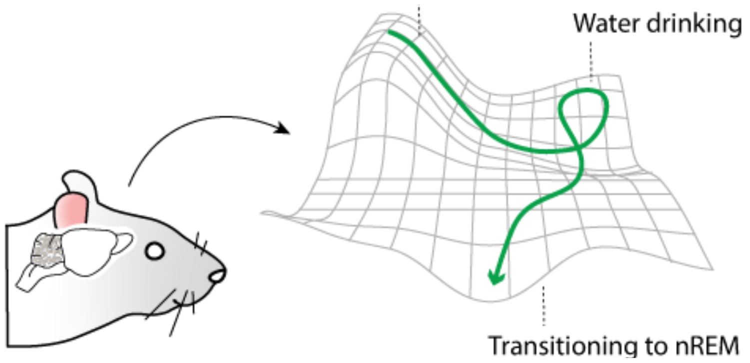
But we still have many sources of variability due to which neurons are sampled, and drift in neural states

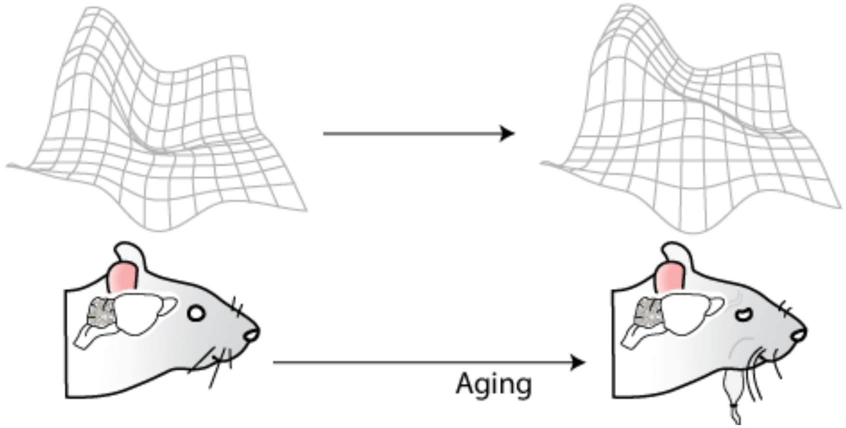
Approach

To tackle these challenges, we need:

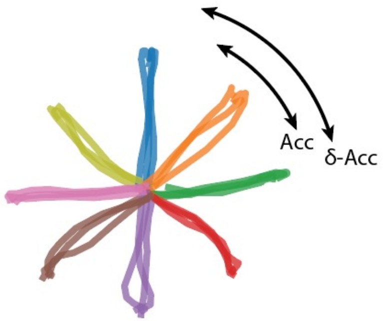
(1) Ways to **learn representations** and build inferences from neural activity – from unlabeled complex behavior

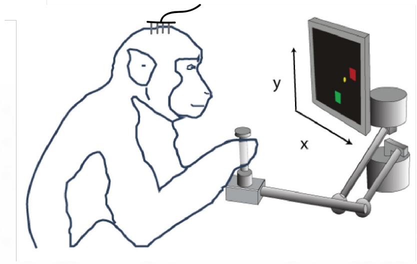
(2) Approaches for **comparing representations**
- How is the brain's representation changing over time and with learning?





Ingredient (1): Learning

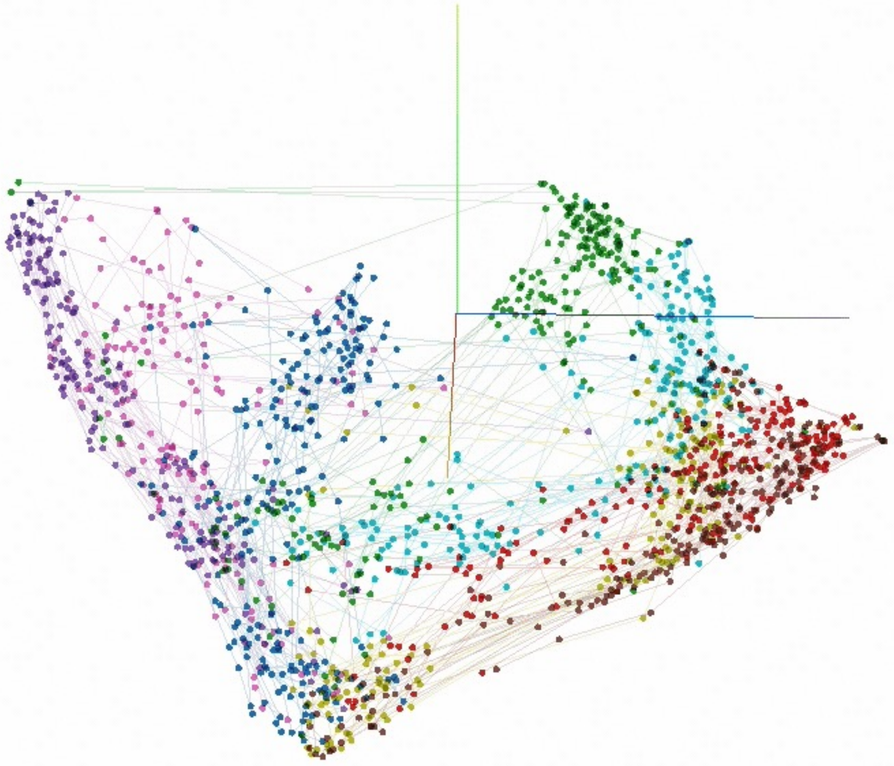




Goal: Learn robust representations of brain activity

Challenges:

- Neural activities shift over time
- Limited or no “true” labels



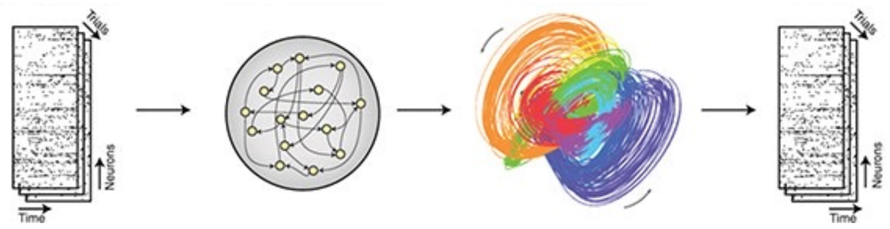
Unsupervised learning approaches

Generative approach:

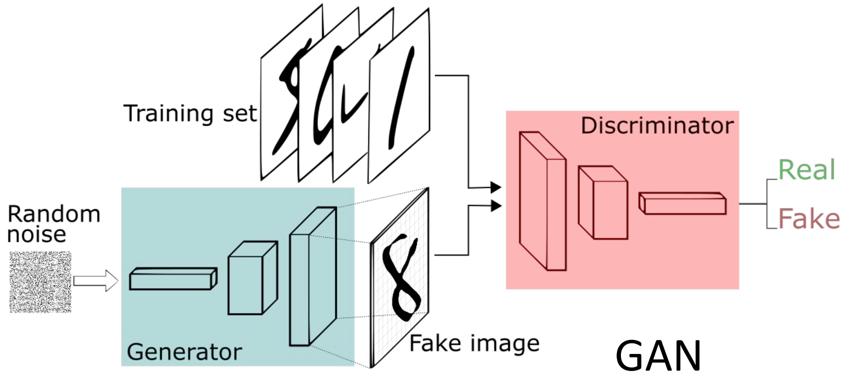
- Find a latent space that allows for efficient recovery
- Ex: Autoencoders, GANs, VAEs

A lot of success in using reconstruction and synthesis as the main objective!

- jPCA (Churchland), GPFA (Yu)
- LFADS (Pandarinath)



Autoencoder



GAN

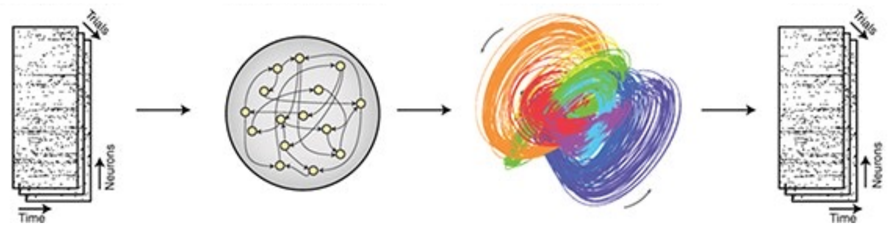
Unsupervised learning approaches

Reconstruction-based approach:

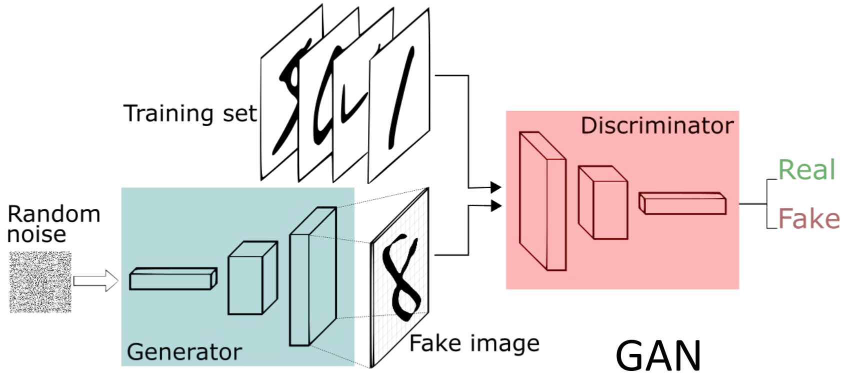
- Find a latent space that allows for efficient recovery
- Ex: Autoencoders, GANs, VAEs

Challenges:

- Sources of high variance will take precedence
- Noise can be easily captured
- <u>Generative models may fail to generalize</u>



Autoencoder



GAN

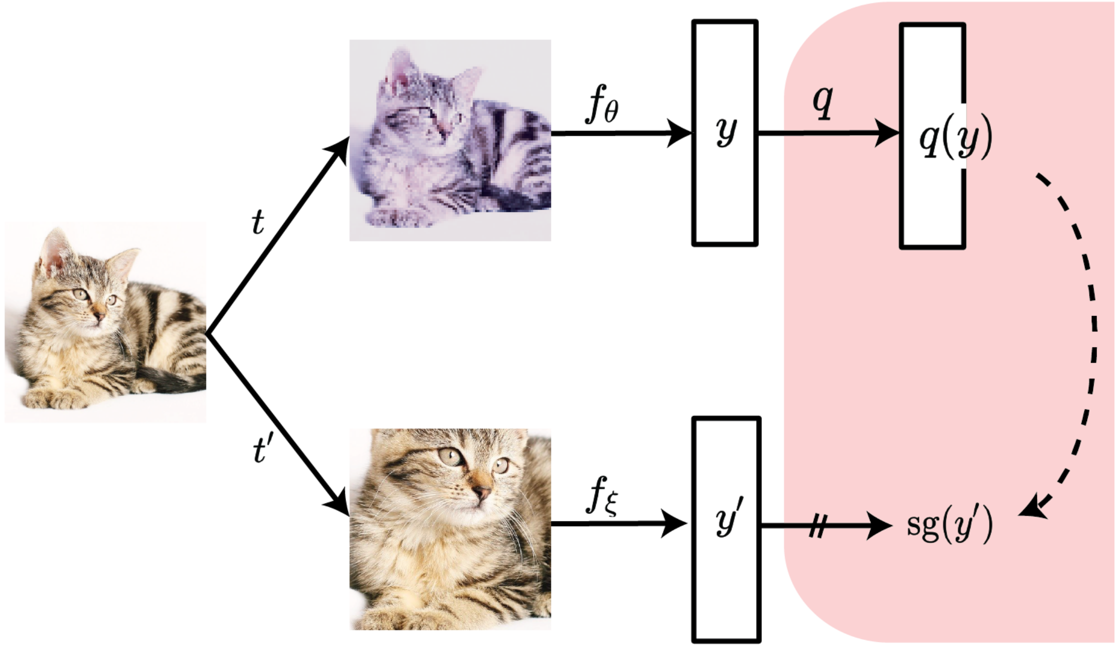
Building invariance through augmentations

Approach:

Learn a representation that *maps augmented views to nearby points* in the representation space

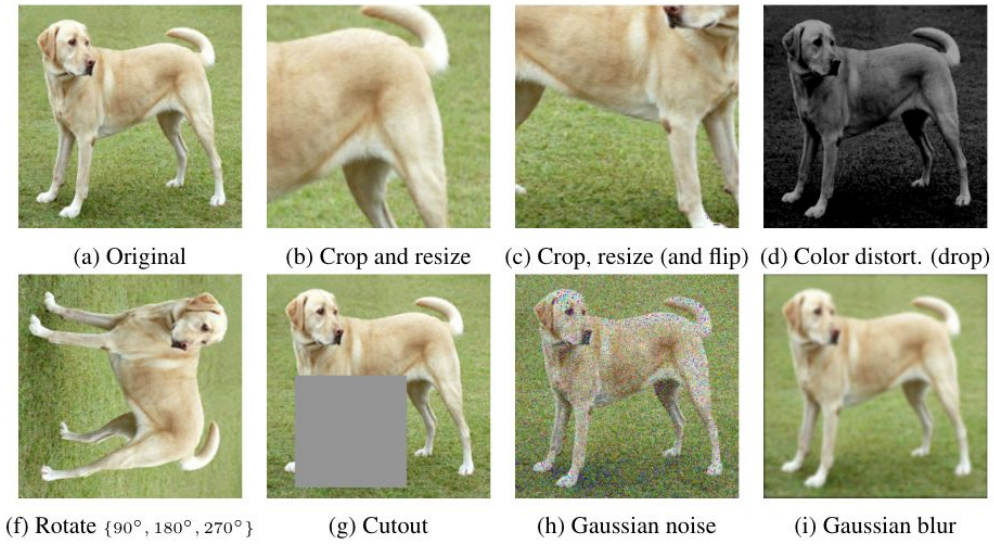
Outcome:

Find common information across different augmentations

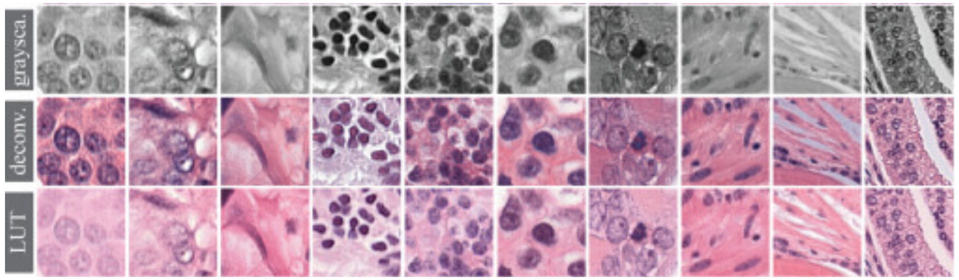


BYOL (Grill, 2020)

Rich augmentations are key to success in vision



Chen, SimCLR, 2020



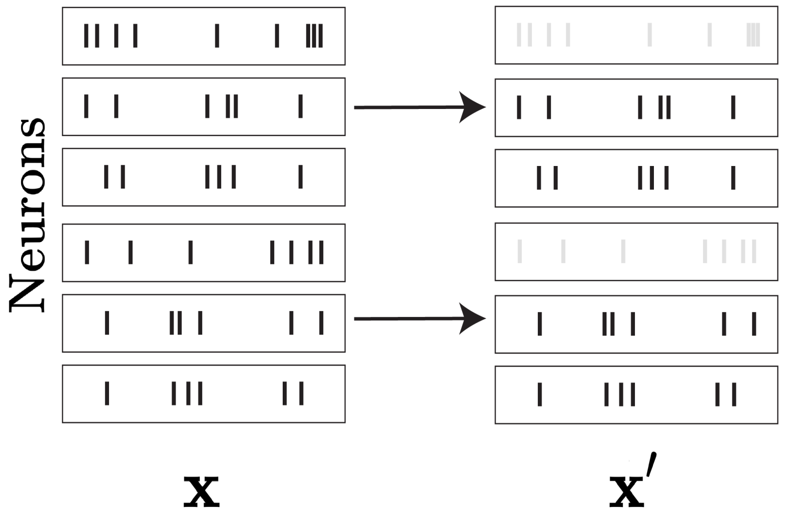
Tellez, 2019

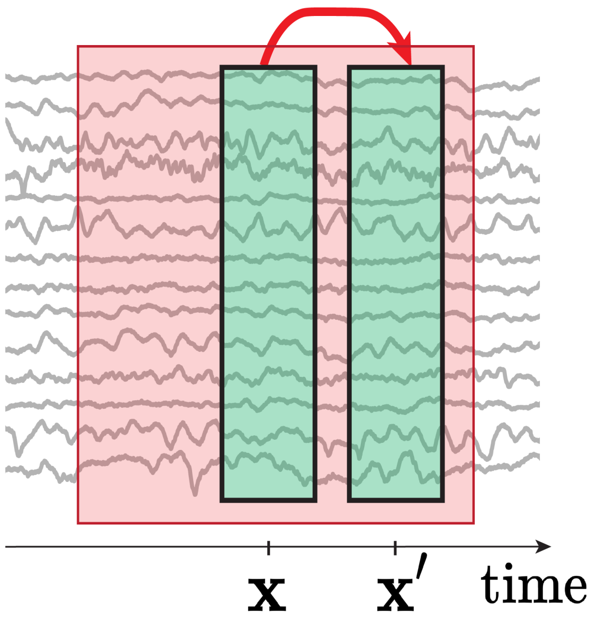
- The success of current SSL methods relies heavily on the choice of augmentations

- When moving to new domains, it is unclear how to **design augmentations** that will build in the types of invariances needed to solve different tasks

Q: What types of invariances do we need when building representations of **neural population activity**?

Self-supervised learning for neural decoding



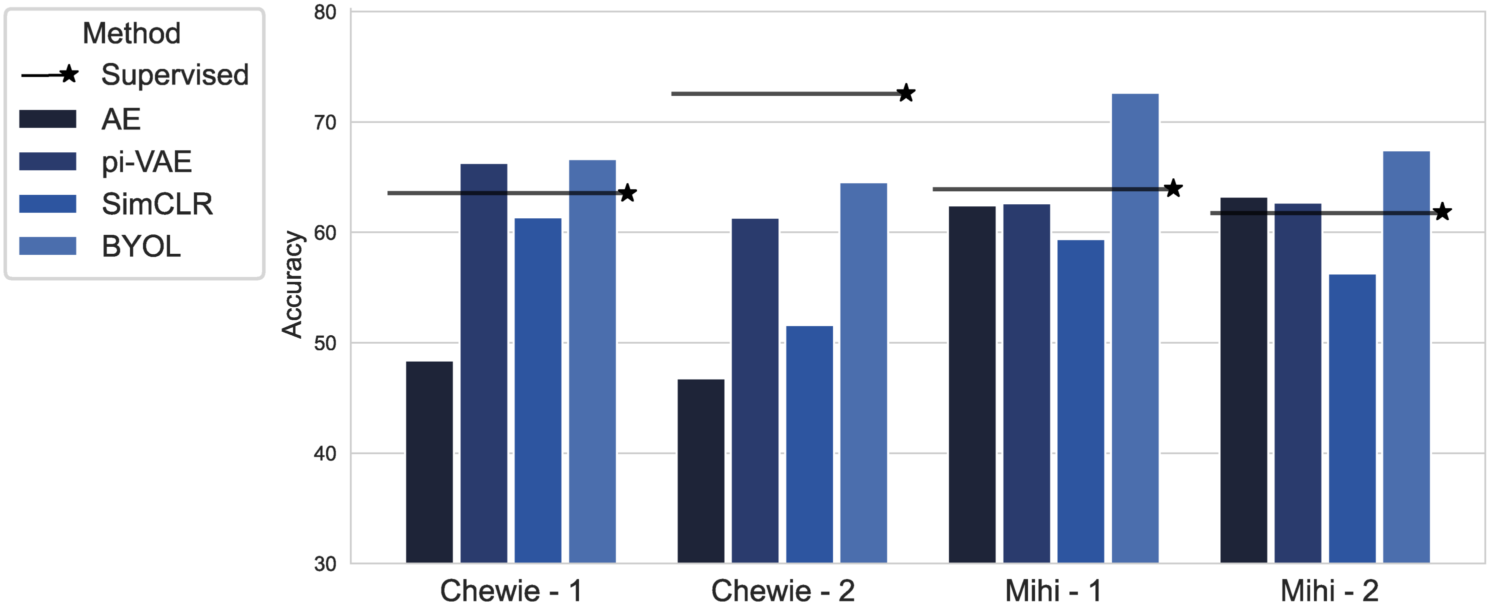


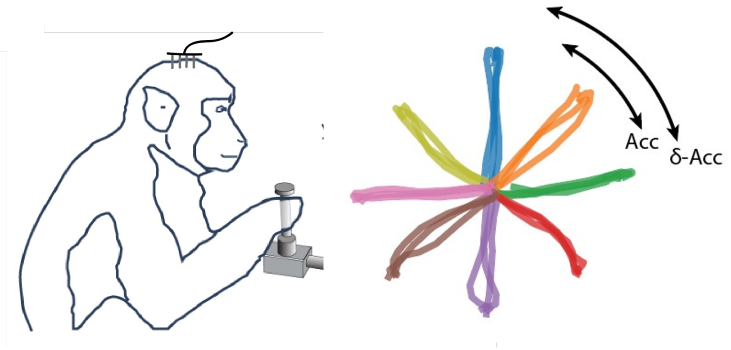
Challenge: We need diverse augmentations and predictive challenges

What are good augmentations for neural data?

- **Temporal prediction** – predict nearby brain states
- **Neuron dropout** – predict responses of different masked neurons

Self-supervised learning for neural decoding

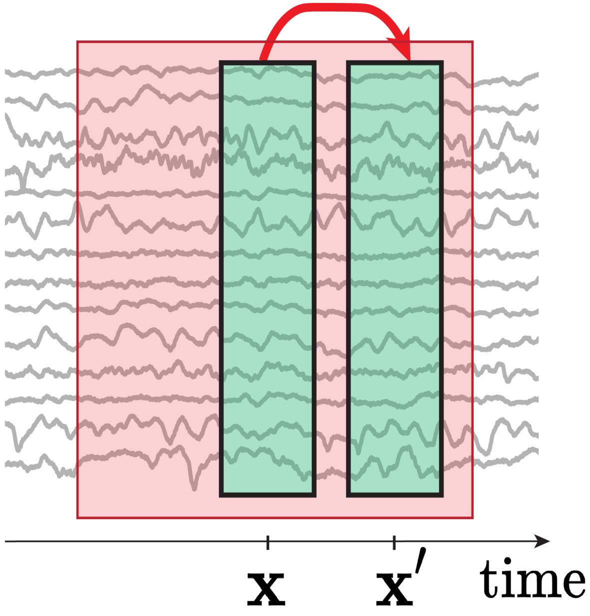


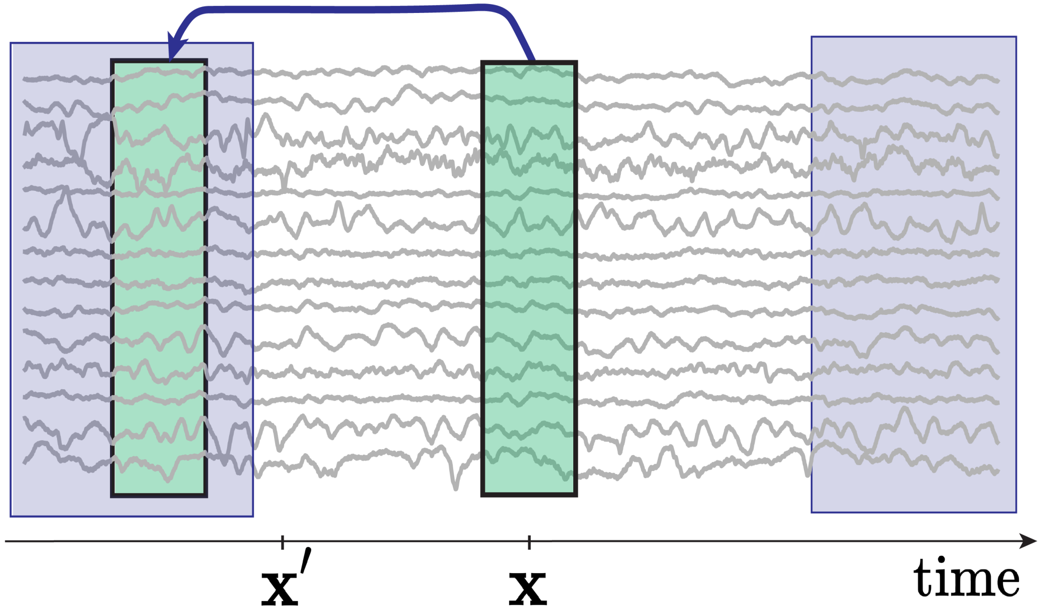


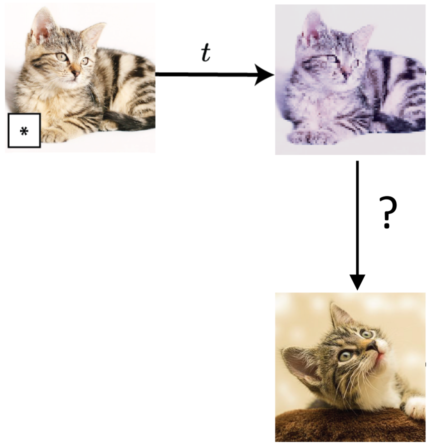
Task: Decode reach direction from motor cortex (M1)

- Train w/ unsupervised loss, freeze network, and train linear layer
- Dropout and temporal predictions provide good predictive signal
- BYOL often **beats the supervised baseline**!

Going beyond local views…



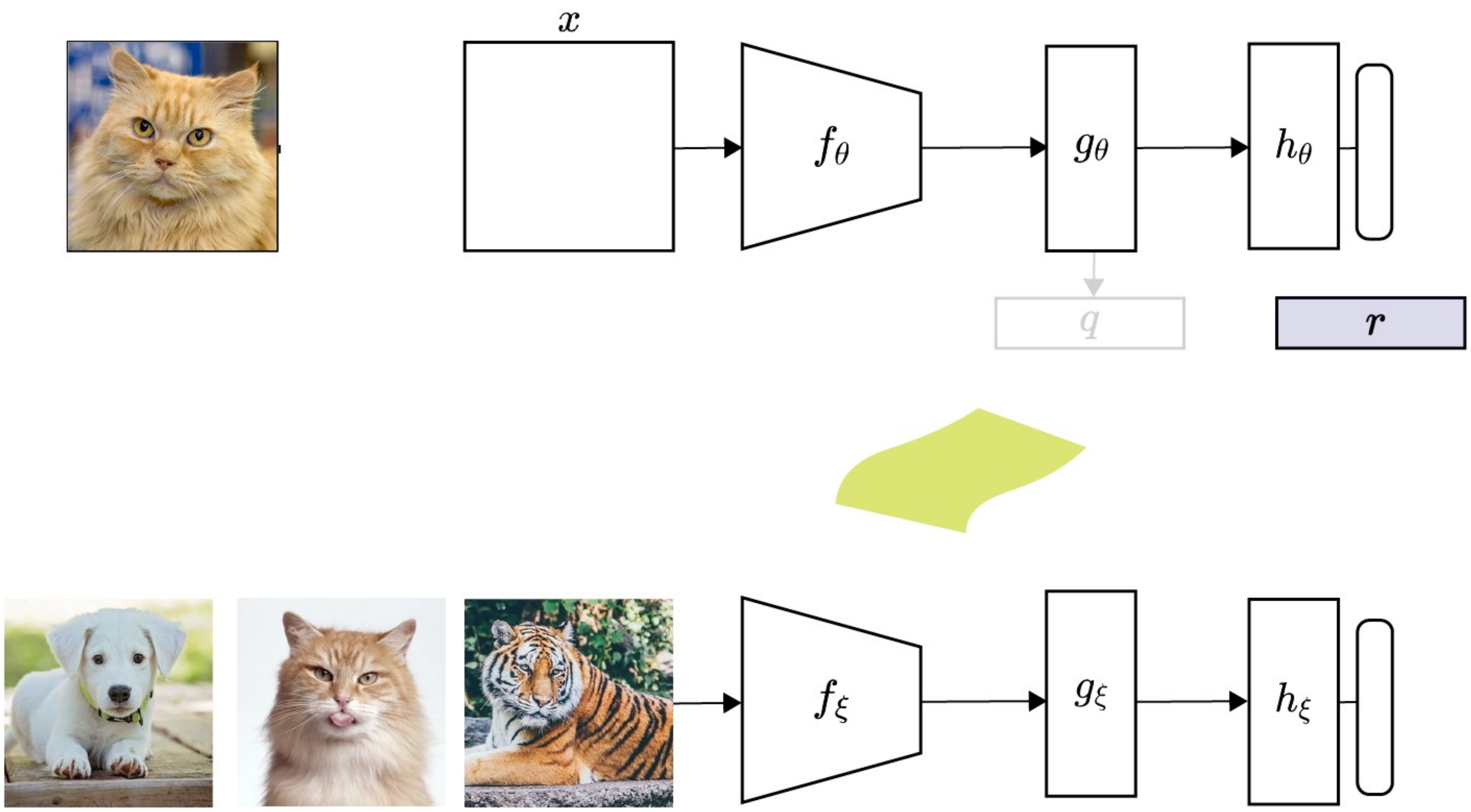




How can we go further and introduce more diversity into our augmentations/ prediction targets?

Goal: find “nonlocal” points to predict – mine views or other samples to predict

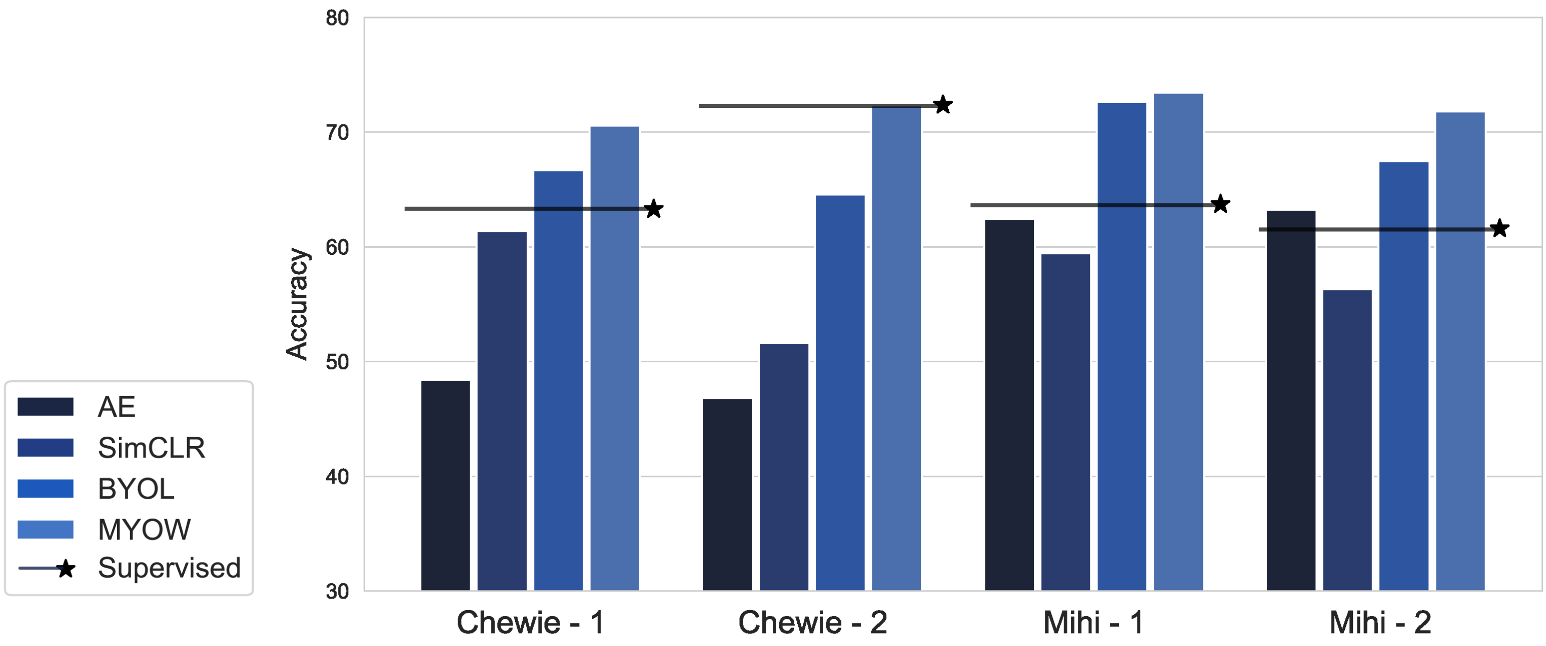
Mine Your Own vieW (MYOW)

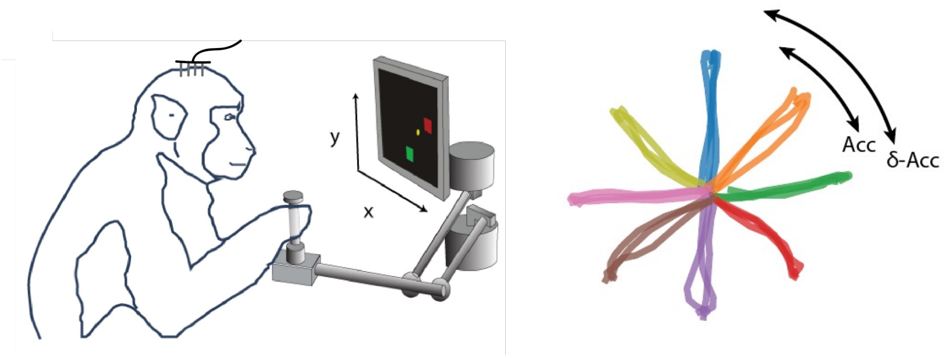


MYOW architecture

$$\mathcal{L}_{\text{aug}} = \mathrm{d}(q(z), z')$$

Results – Movement decoding from cortex





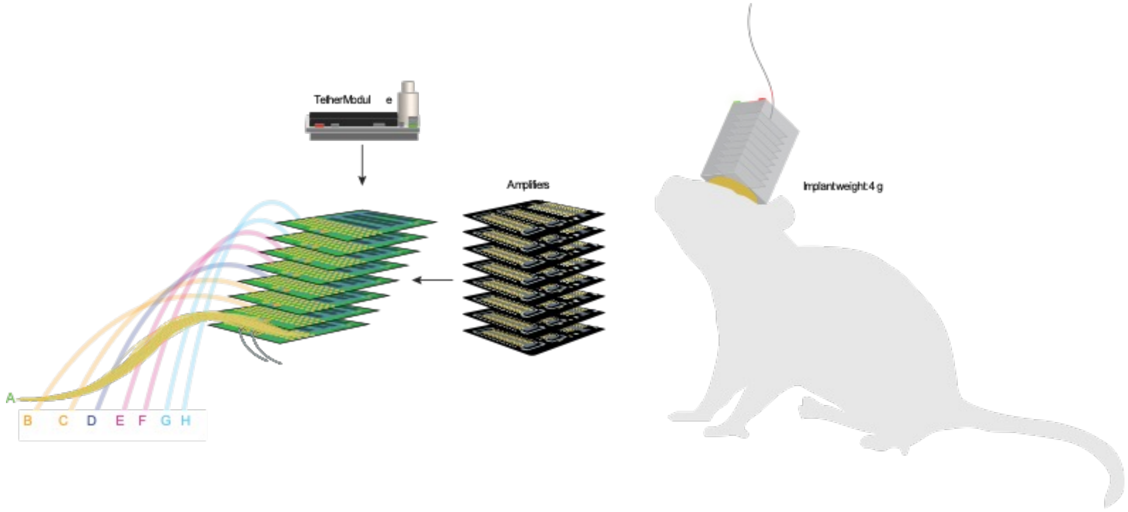
Task: Decode reach direction
from the primary motor cortex

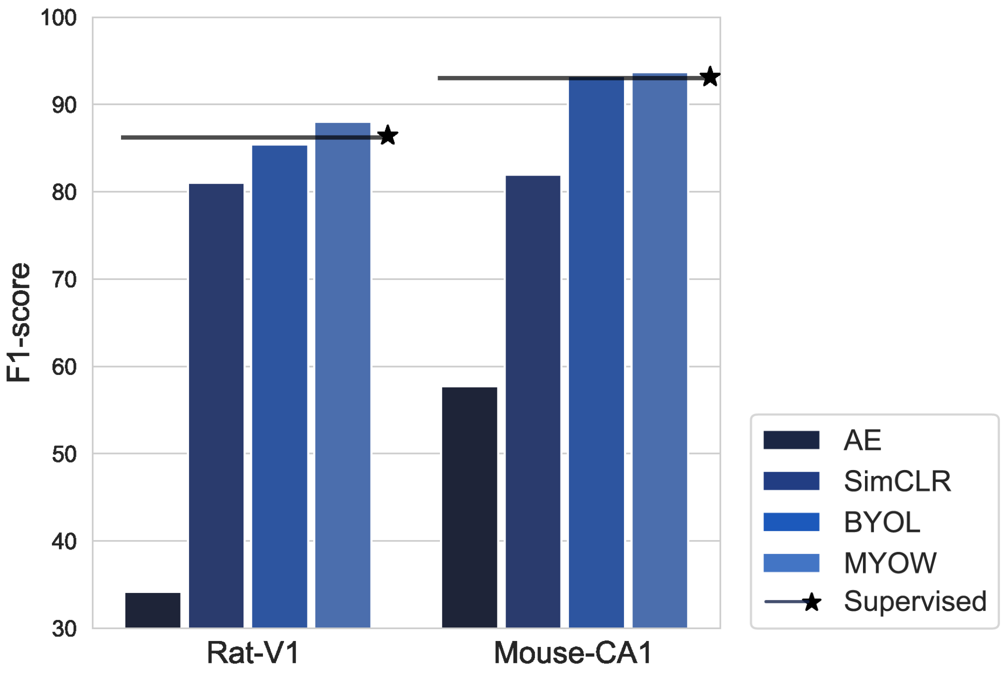
- Self-supervised methods (BYOL, MYOW) outperform the supervised baseline!

- Both **temporal** and **dropout** augmentations provides sufficient diversity and preservation of task semantics

Results – Free behavior over 12 hours

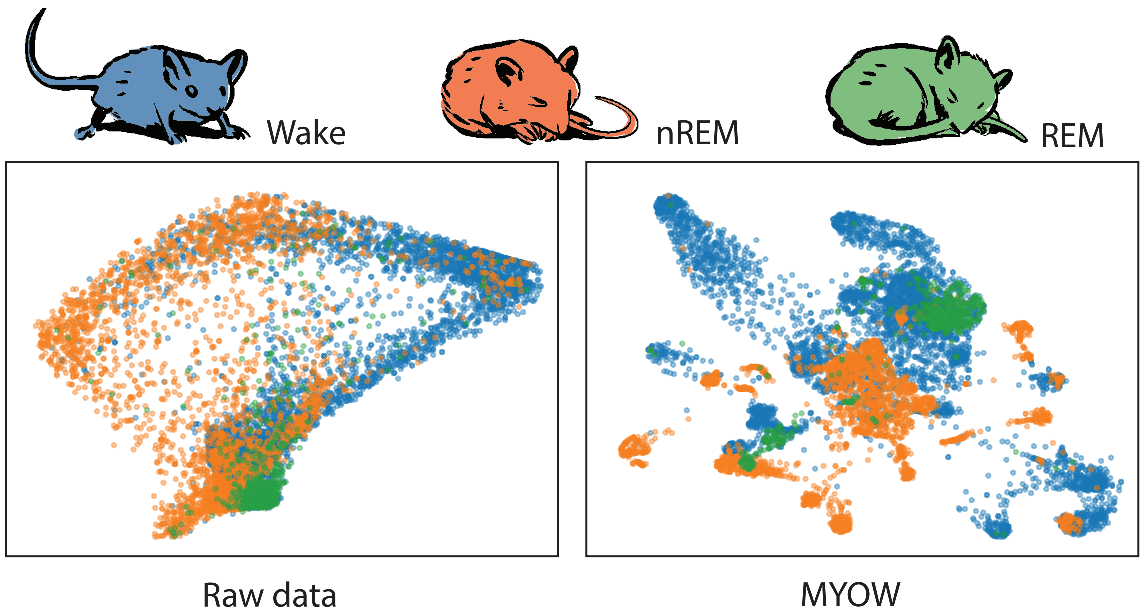
Task: Decoding REM, nREM, wake from mouse CA1 and rat V1 during free behavior (Hengen Lab)

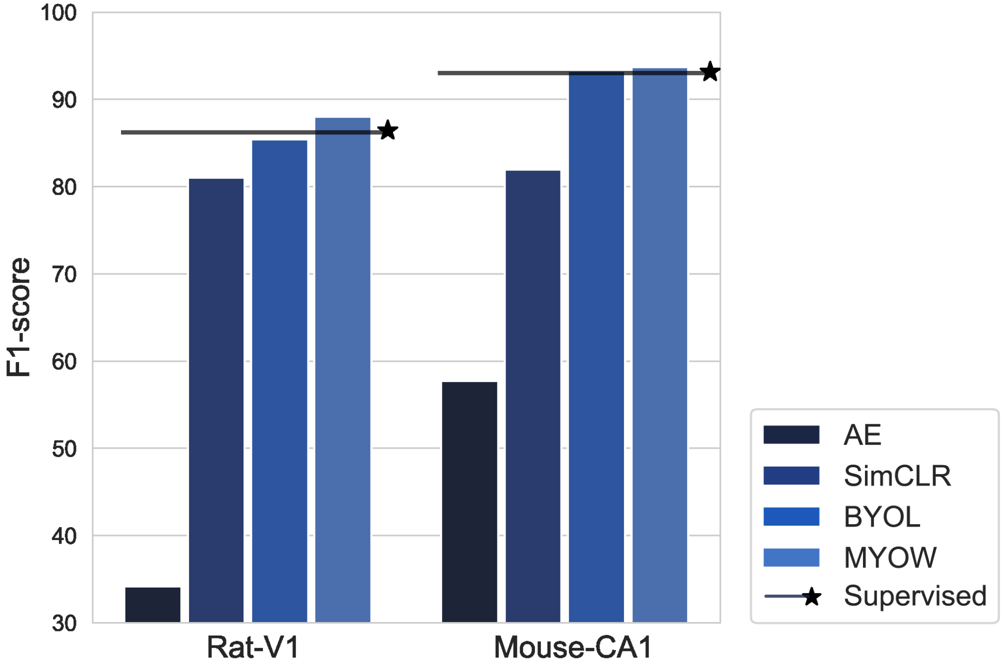




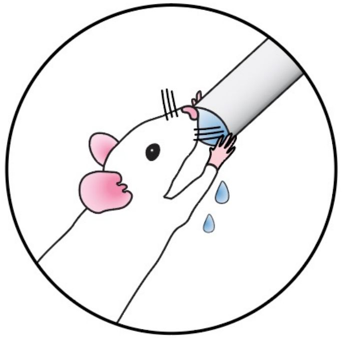
- Similar augmentation strategies are effective in diverse domains!
- Interesting finding: **Dropout as an augmentation** is sufficient to seed learning

Results – free behavior over 12 hours

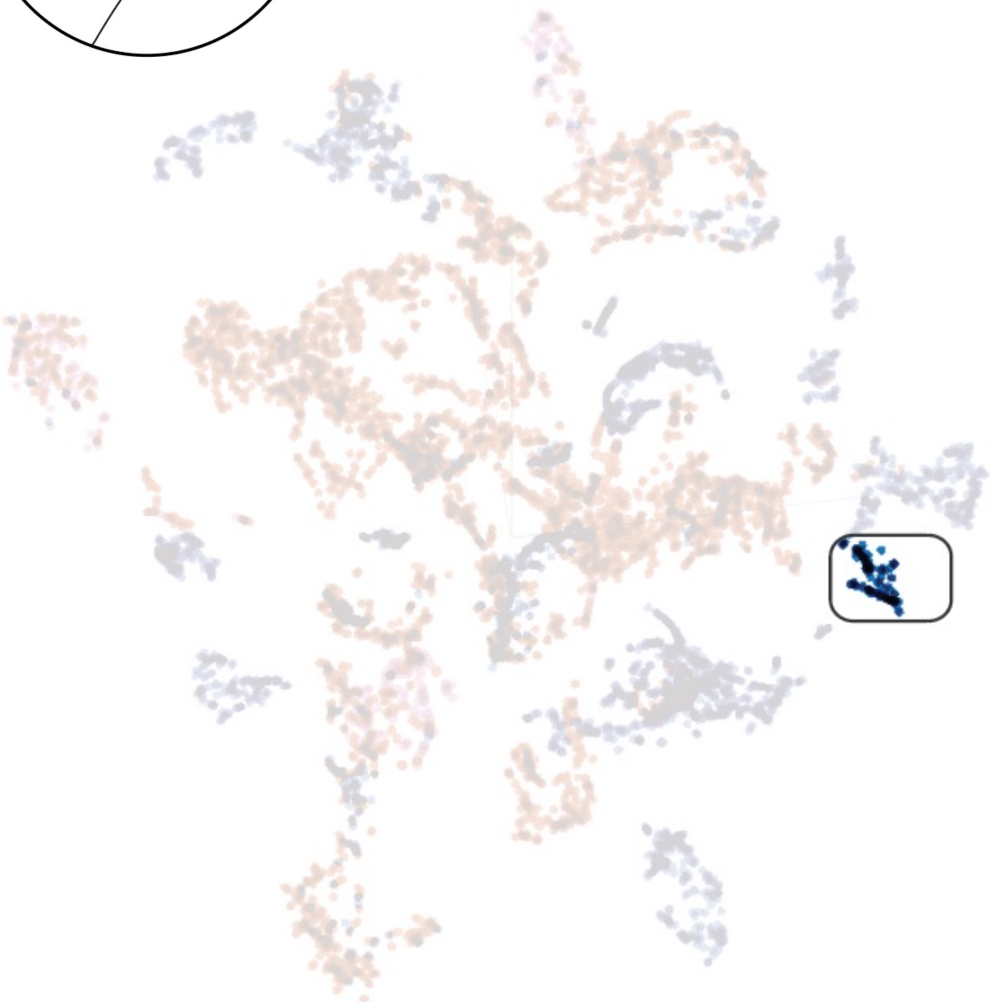


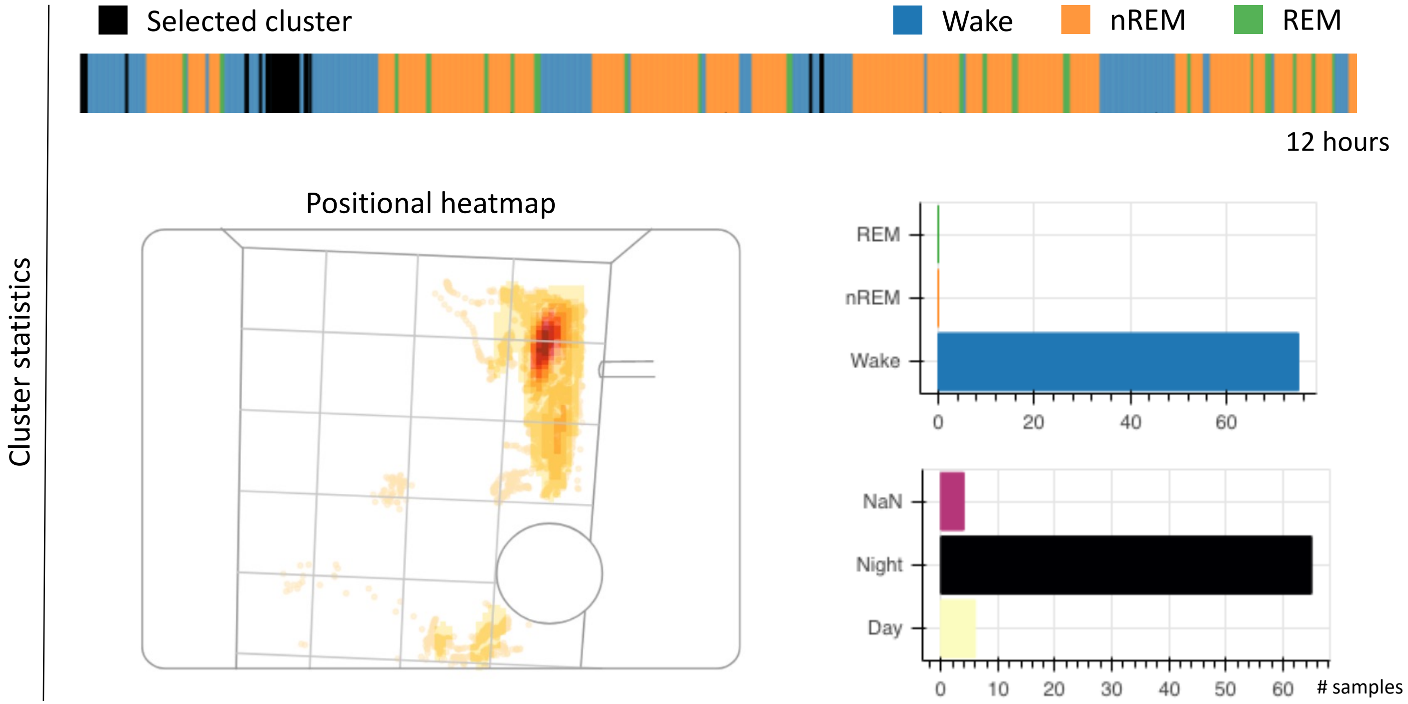


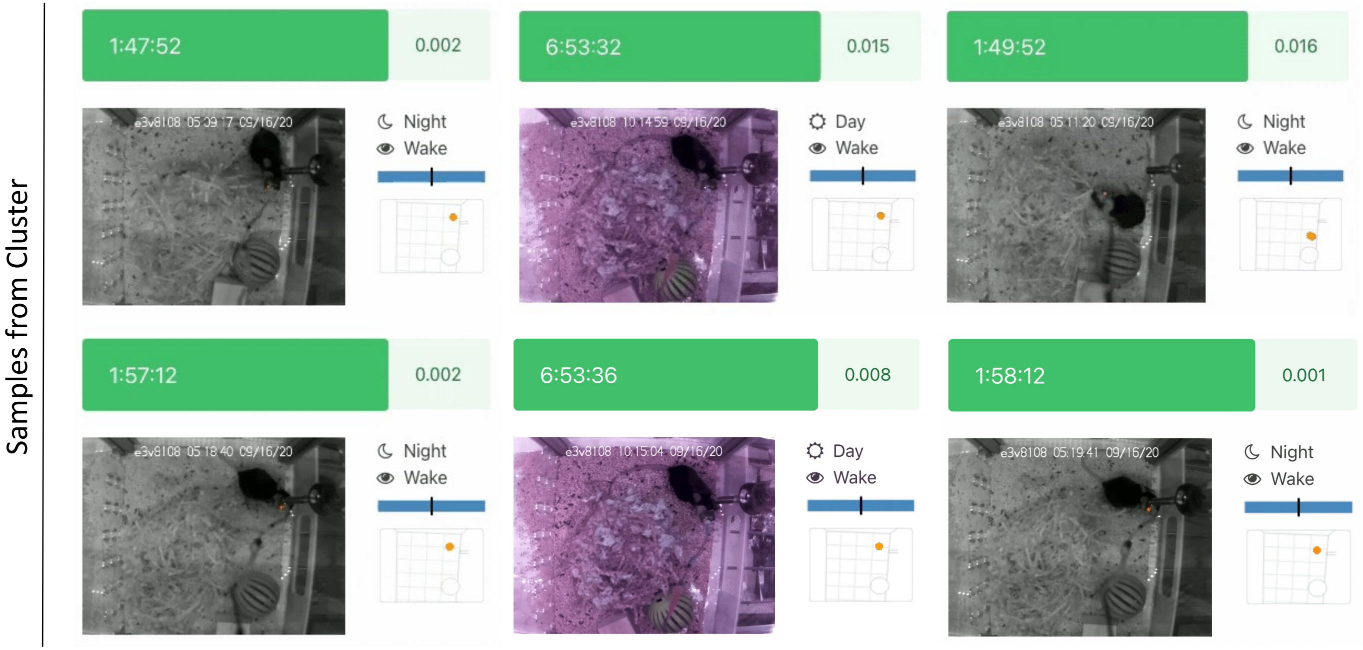
- Similar augmentation strategies can be applied despite key differences between datasets!
- MYOW reveals additional structure in the neural activities during wake and sleep



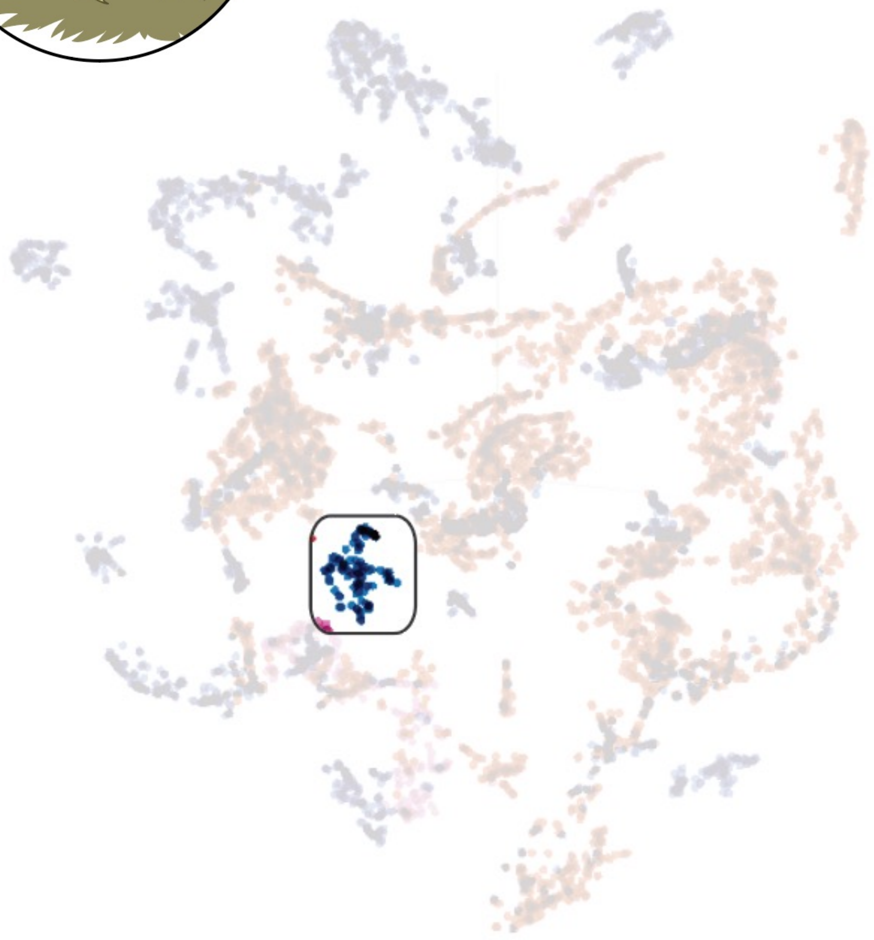
Water drinking

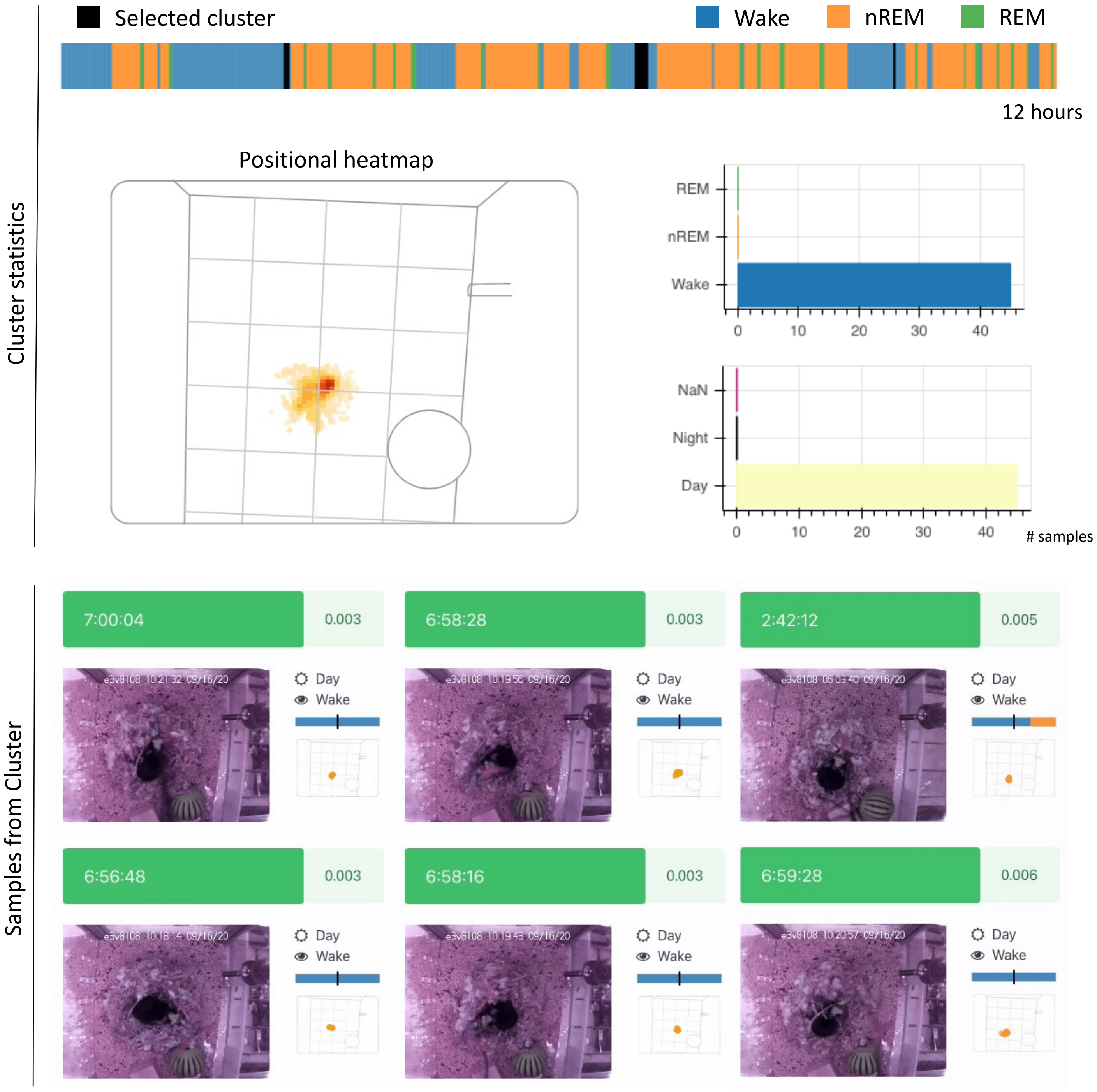




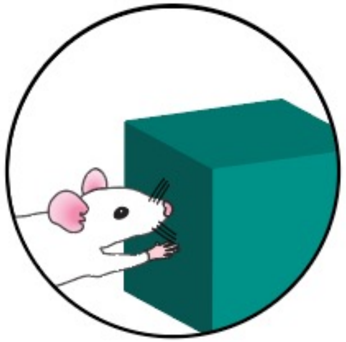


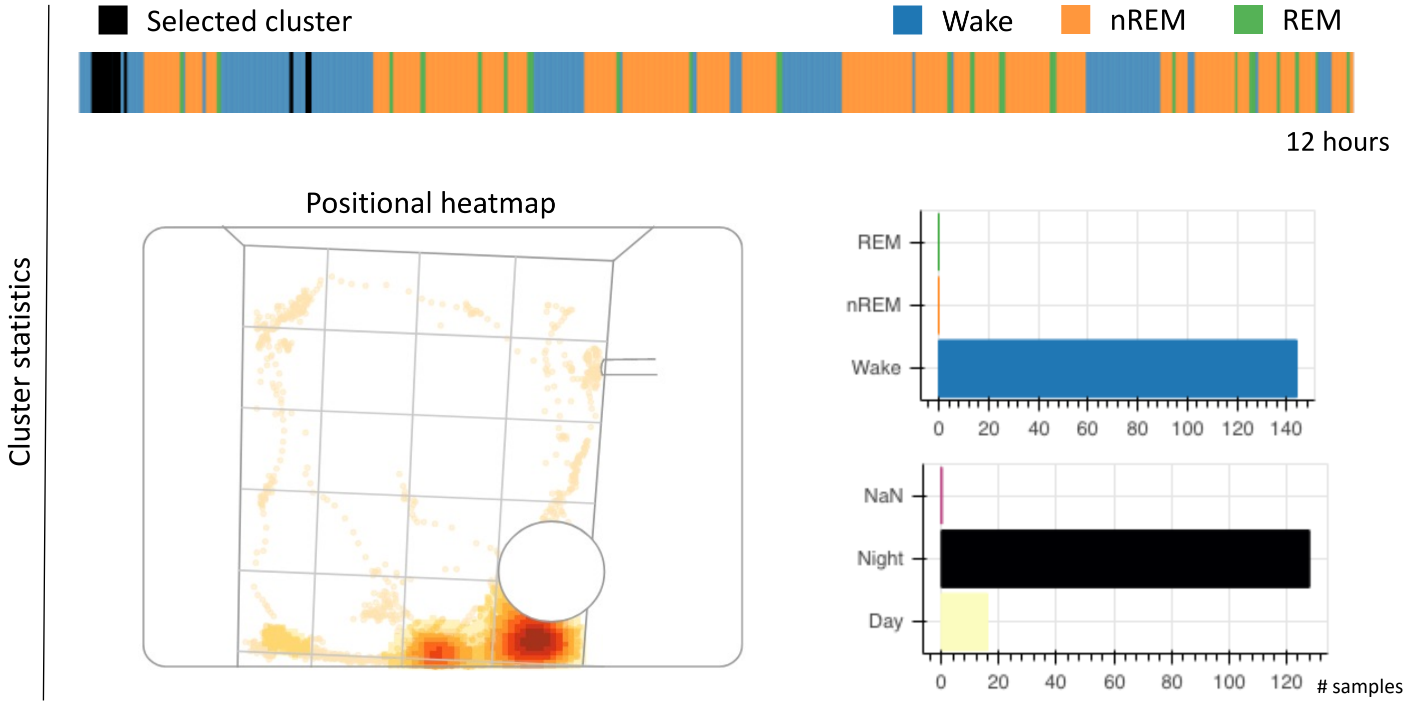
Getting comfy

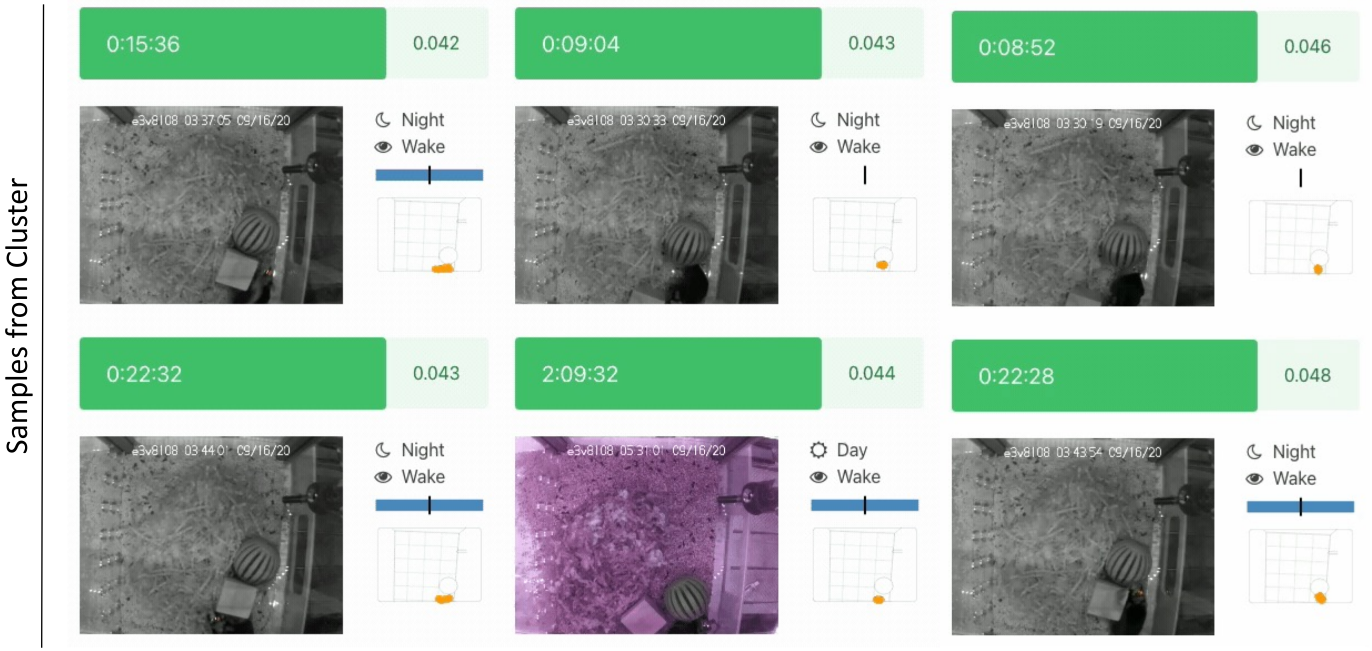




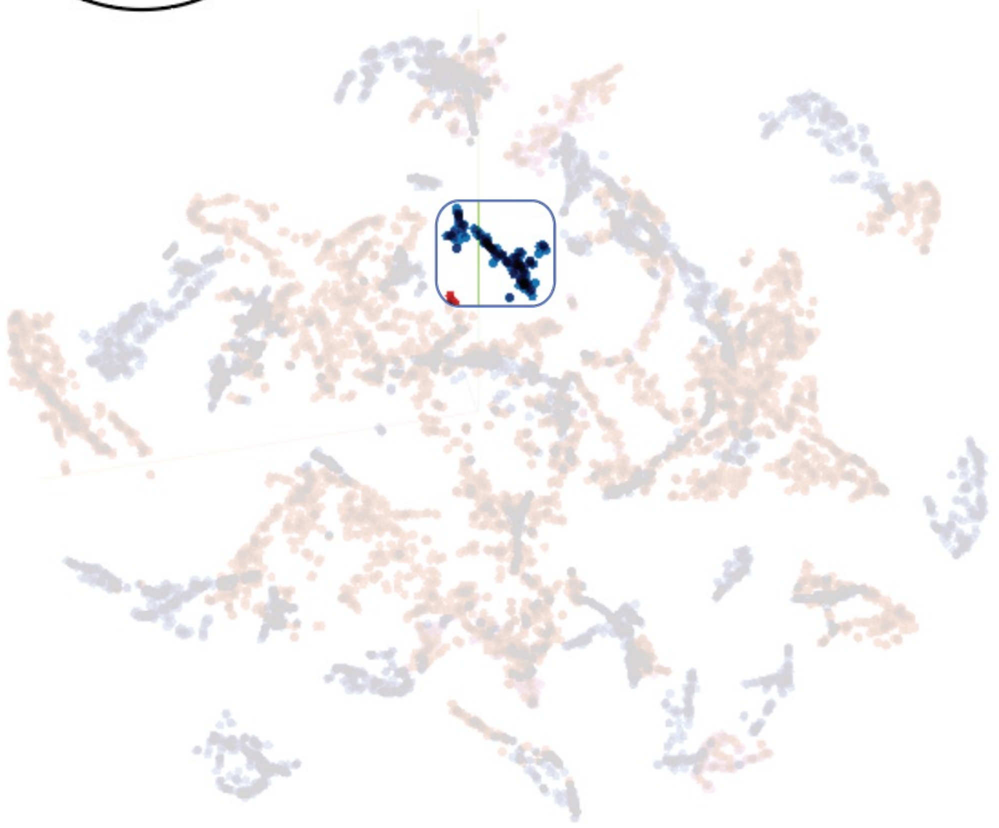
Moving box

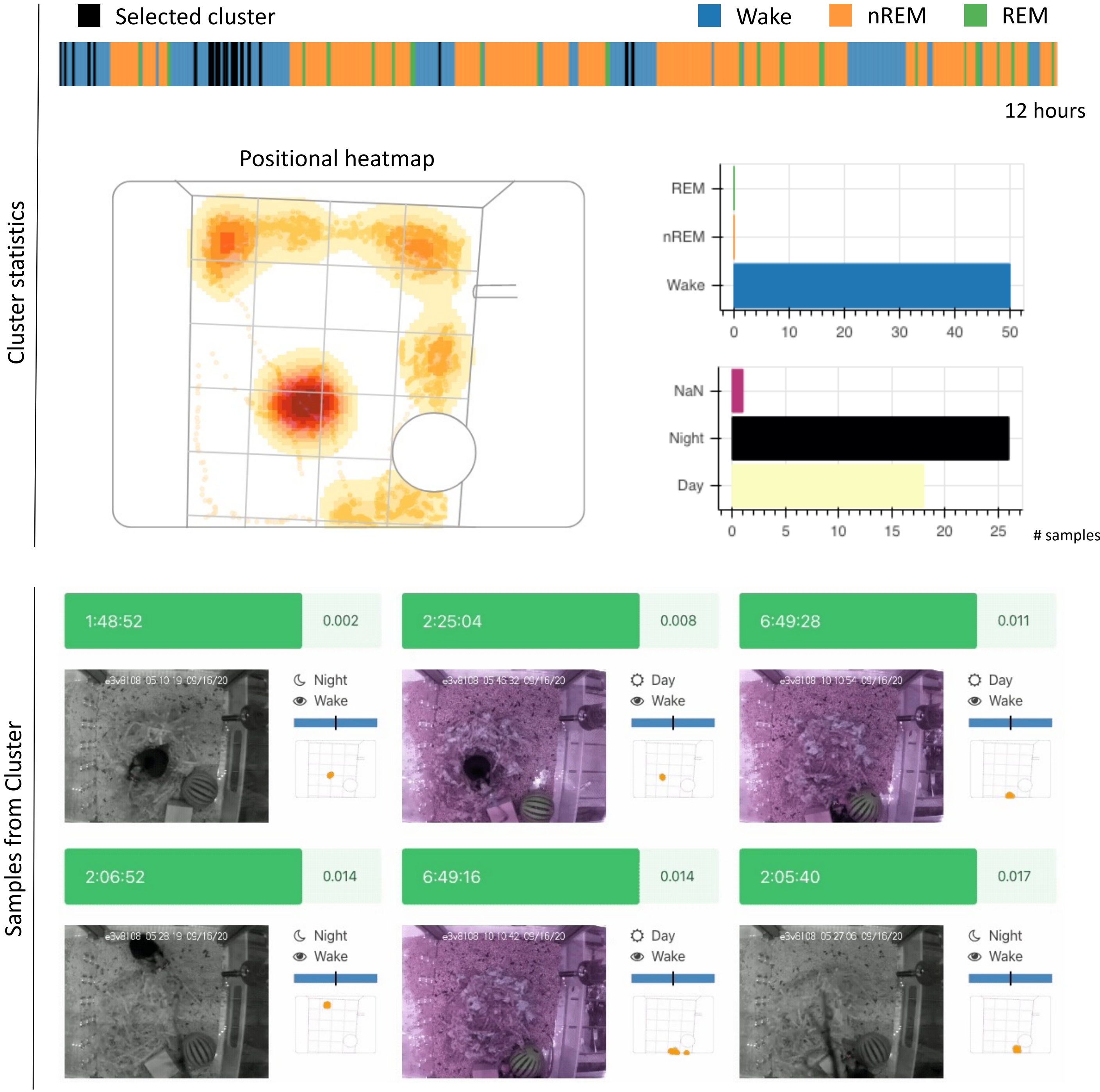



Cluster statistics
Selected cluster
Wake
nREM
REM
12 hours
Positional heatmap
REM
nREM
Wake
0 20 40 60 80 100 120 140
NaN
Night
Day
0 20 40 60 80 100 120
samples


Samples from Cluster
0:15:36
0.042
Night
Wake
0:09:04
0.043
Night
Wake
0:08:52
0.046
Night
Wake
0:22:32
0.043
Night
Wake
2:09:32
0.044
Day
Wake
0:22:28
0.048
Night
Wake

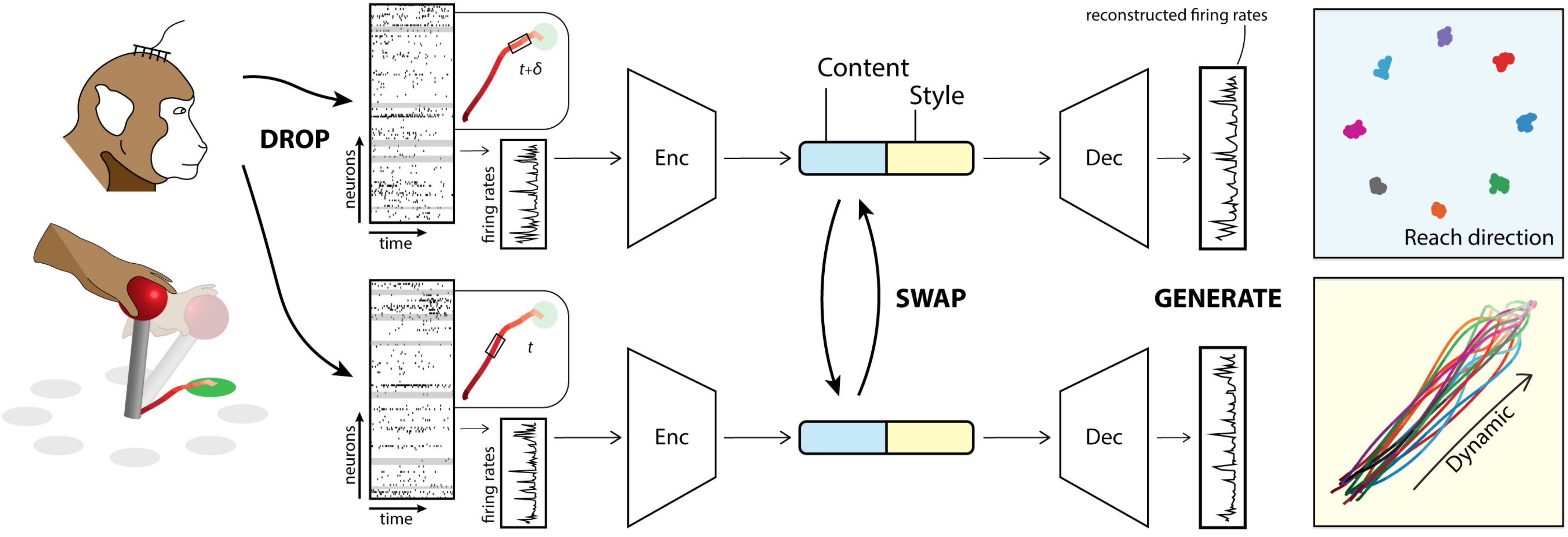
Nest building





MYOW provides insights into complex behaviors

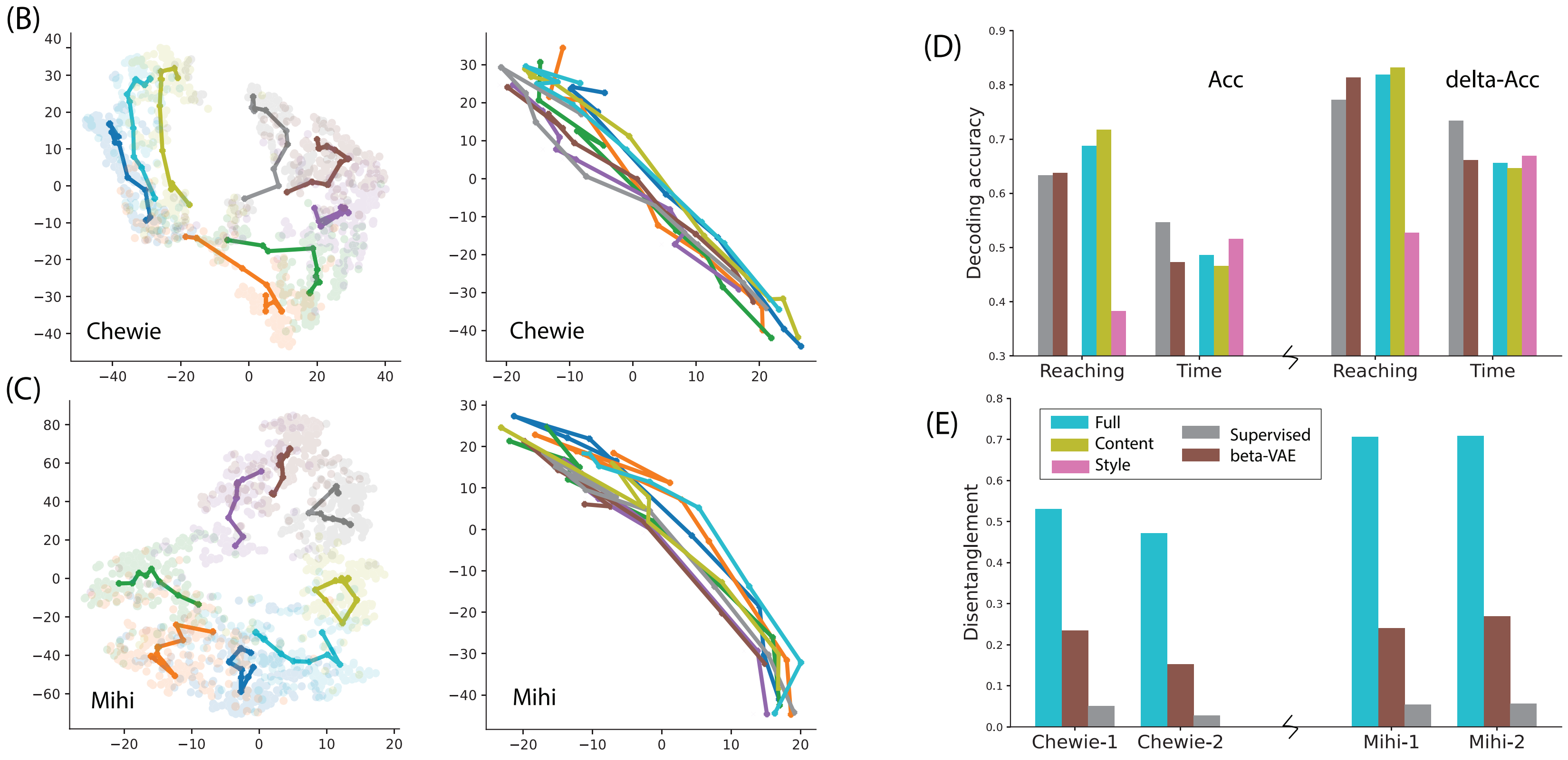
Combining SSL and generative modeling



By incorporating a the reconstruction loss in the style space, we can simplify our architecture (only one network needed!)

(Liu et al., NeurIPS 2021 – https://nerdslab.github.io/SwapVAE/)

Disentangling neural population activity



- Mihi has a wait period before hearing auditory feedback that gives the ‘Go’ cue
- Target location and dynamics are **more easily disentangled** in datasets from Mihi

Summary so far…

- **Mine Your Own vieW (MYOW)**
  - A new approach for self-supervised representation learning
  - General approach that can be used in diverse domains
- When applied to neural population activity:
  - New augmentations for neural activity that work in rodent and macaque and across brain areas
  - Across-sample prediction w/ MYOW often beats the supervised baseline
- **Swap-VAE**:
  - Combining self-supervision with a generative model
  - Disentanglement of content from dynamics

Summary so far…

- Mine Your Own vieW (MYOW)
  - A new approach for self-supervised representation learning
  - General approach that can be used in diverse domains
- When applied to neural population activity:
  - New augmentations for neural activity that work in rodent and macaque and across brain areas
  - Across-sample prediction w/ MYOW often beats the supervised baseline
- Swap-VAE:
  - Combining self-supervision with a generative model
  - Disentanglement of content from dynamics

Summary so far…

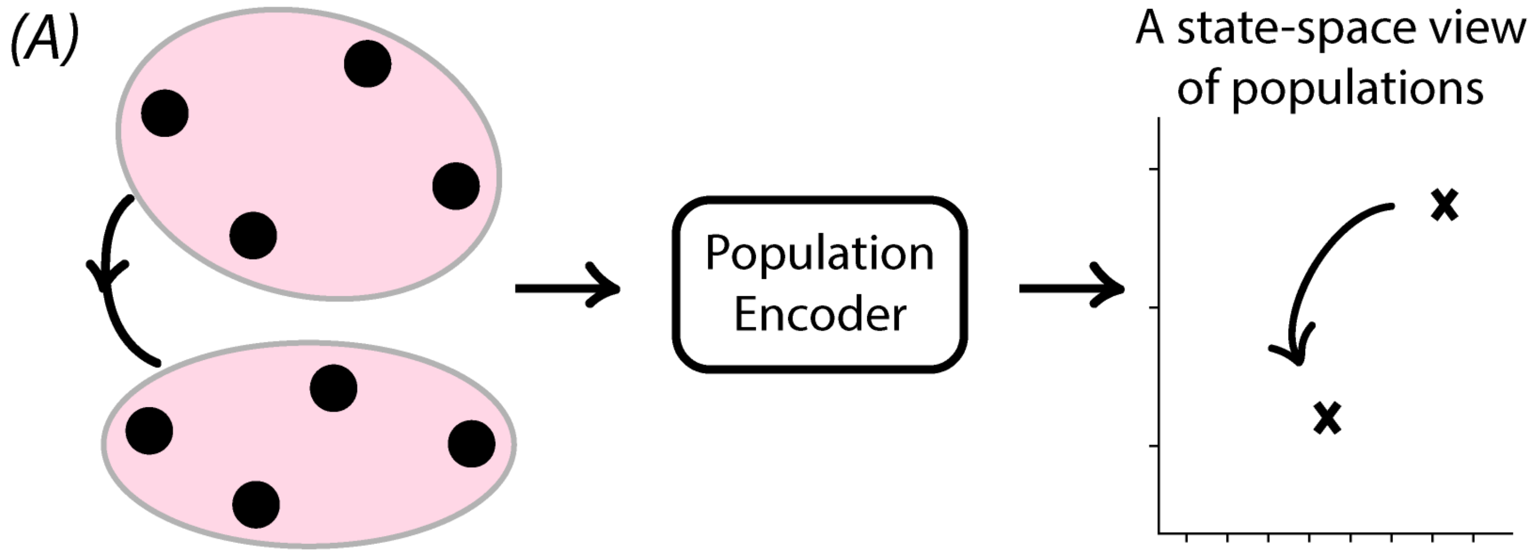
- **Mine Your Own vieW (MYOW)**
  - A new approach for self-supervised representation learning
  - General approach that can be used in diverse domains

- When applied to neural population activity:
  - New augmentations for neural activity that work in rodent and macaque and across brain areas
  - Across-sample prediction w/ MYOW often beats the supervised baseline

- **Swap-VAE**:
  - Combining self-supervision with a generative model
  - Disentanglement of content from dynamics

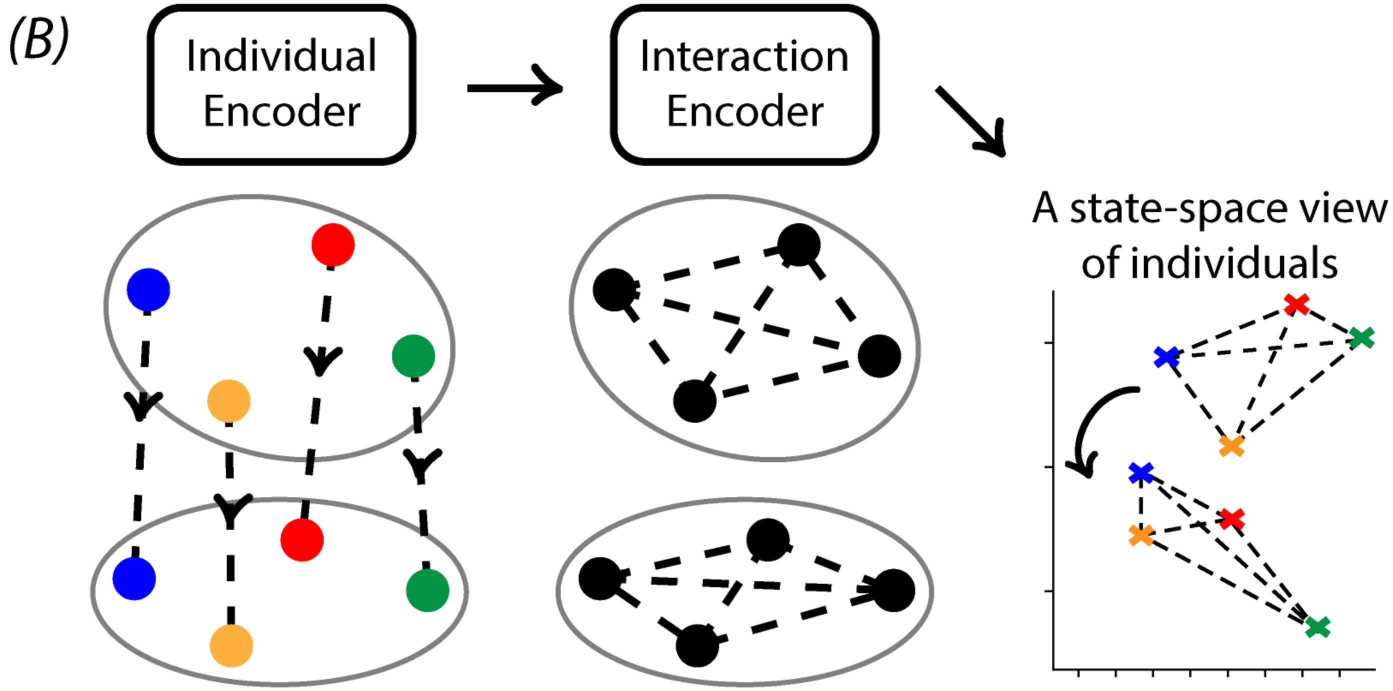
Zooming back in on neurons…

But how are different neurons contributing to the overall picture?



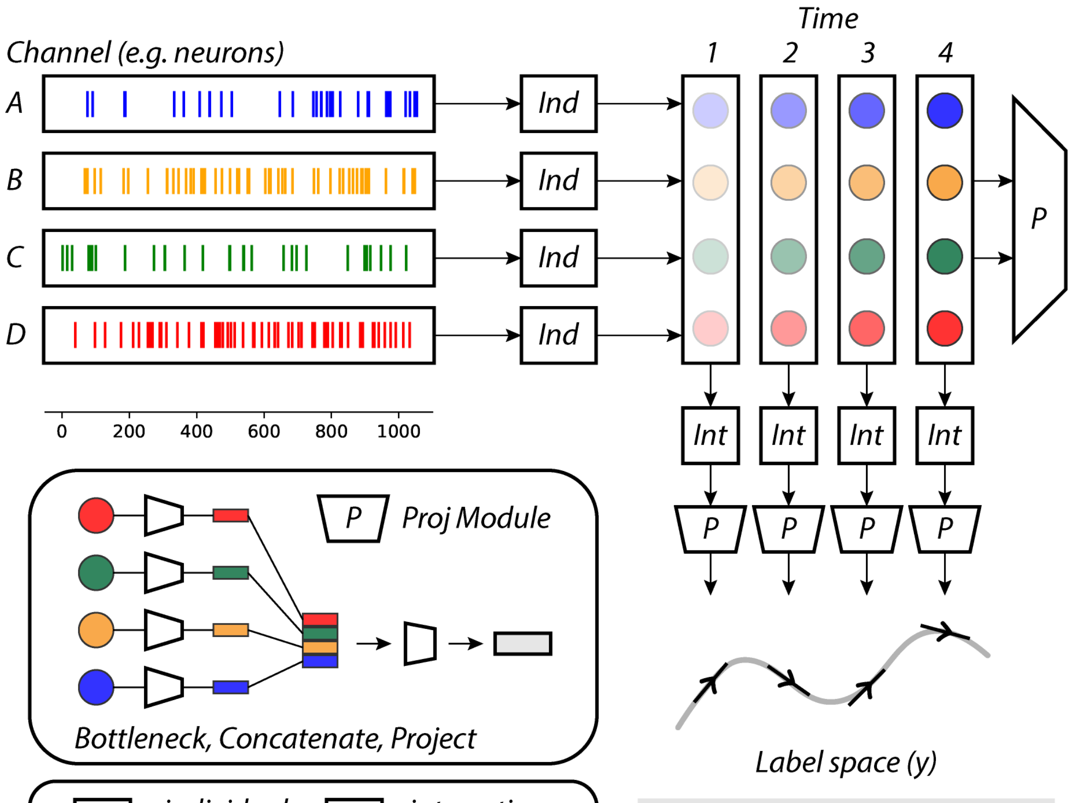
Liu, 2022: https://arxiv.org/abs/2206.06131

But how are different neurons contributing to the overall picture?



Liu, 2022: https://arxiv.org/abs/2206.06131

Embedded Interaction Transformer (EIT)



(C) EIT Model Training

Embedded Interaction Transformer (EIT)

1. Learn an embedding **for each neuron's** firing rate patterns over time

2. After embedding**, combine neurons into a population-level** representation

Results – Reach Decoding

Table 2: *Performance on behavioral decoding from populations of neurons in the motor cortex.*

I. Decoding performance on neural datasets

	Mihi-Chewie ($T = 2$)				*Mihi-Chewie* ($T = 6$)				*Maze* (R^2x100)	
	C-1	C-2	M-1	M-2	C-1	C-2	M-1	M-2	Vel	Pos
MLP	74.22	74.54	78.17	74.42	78.91	90.74	87.90	84.11	60.64	78.59
GRU	75.78	75.46	79.96	74.22	84.72	90.12	85.98	78.81	79.97	94.01
NDT	59.90	58.56	73.02	70.74	64.93	63.73	80.56	71.96	76.01	90.07
NDT-Sup	**80.47**	80.56	83.93	80.79	87.33	**94.29**	**96.83**	91.47	82.02	**94.88**
EIT (T)	76.04	81.25	81.15	71.71	83.33	88.27	93.65	86.82	71.18	87.09
EIT	79.69	**82.41**	**86.51**	**81.61**	**88.36**	92.59	95.24	**91.57**	**82.15**	94.77

Liu, 2022: https://arxiv.org/abs/2206.06131

But how are different neurons contributing to the overall picture?

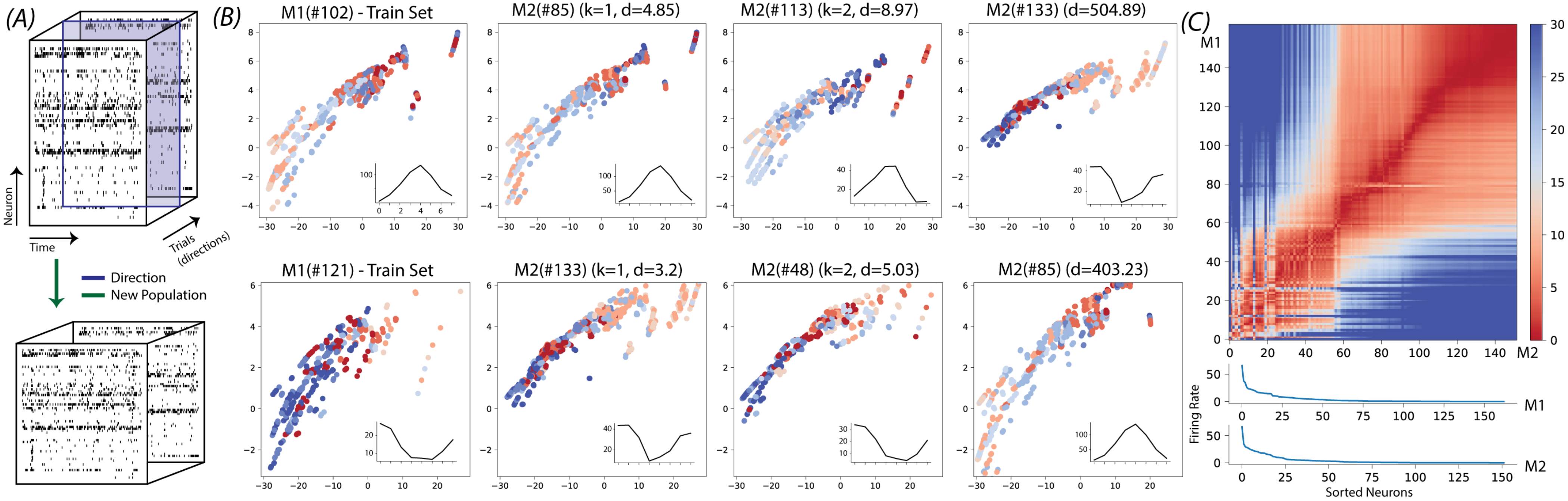
Table 2: *Performance on behavioral decoding from populations of neurons in the motor cortex.*

	I. Decoding performance on neural datasets									
	Mihi-Chewie ($T = 2$)				*Mihi-Chewie* ($T = 6$)				*Maze* (R^2x100)	
	C-1	C-2	M-1	M-2	C-1	C-2	M-1	M-2	Vel	Pos
MLP	74.22	74.54	78.17	74.42	78.91	90.74	87.90	84.11	60.64	78.59
GRU	75.78	75.46	79.96	74.22	84.72	90.12	85.98	78.81	79.97	94.01
NDT	59.90	58.56	73.02	70.74	64.93	63.73	80.56	71.96	76.01	90.07
NDT-Sup	**80.47**	80.56	83.93	80.79	87.33	**94.29**	**96.83**	91.47	82.02	**94.88**
EIT (T)	76.04	81.25	81.15	71.71	83.33	88.27	93.65	86.82	71.18	87.09
EIT	79.69	**82.41**	**86.51**	**81.61**	**88.36**	92.59	95.24	**91.57**	**82.15**	94.77
	II. Generalization performance - trained on one population, tested on another (more in Appendix 3.2)									
NDT retrain(C-2)	75.52	-	77.78	73.06	85.24	-	92.46	86.18	×	×
NDT retrain(M-2)	74.48	70.37	76.78	-	**86.28**	89.20	91.99	-	×	×
EIT (C-2)	**79.17**	-	82.94	**75.24**	81.42	-	92.33	**91.34**	×	×
EIT (M-2)	78.13	**81.02**	**84.13**	-	84.72	**91.06**	**93.25**	-	×	×

Wow!

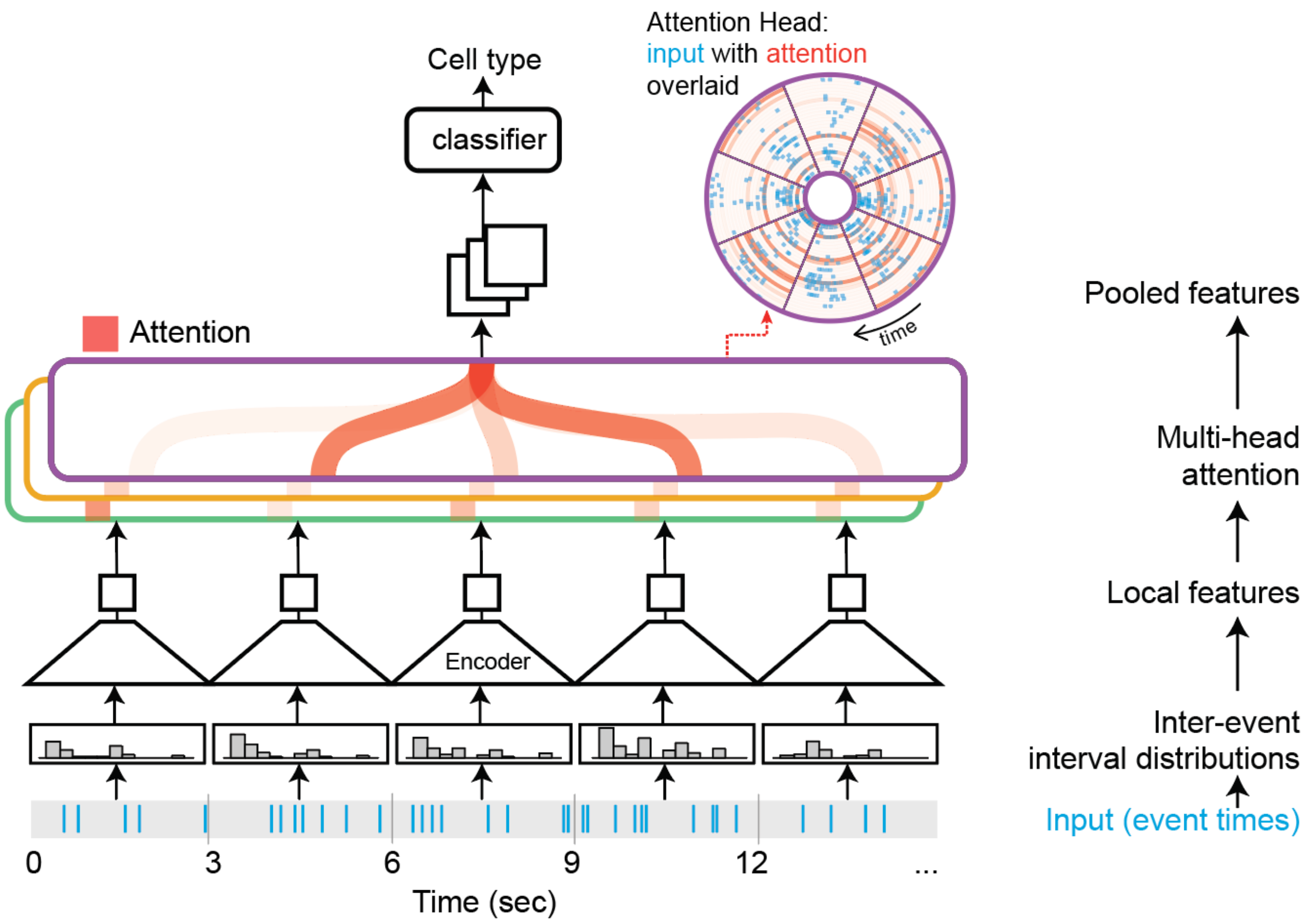
Across animal transfer without retraining!!

But how are different neurons contributing to the overall picture?



Liu, 2022: https://arxiv.org/abs/2206.06131

Decoding cell types from in vivo activity



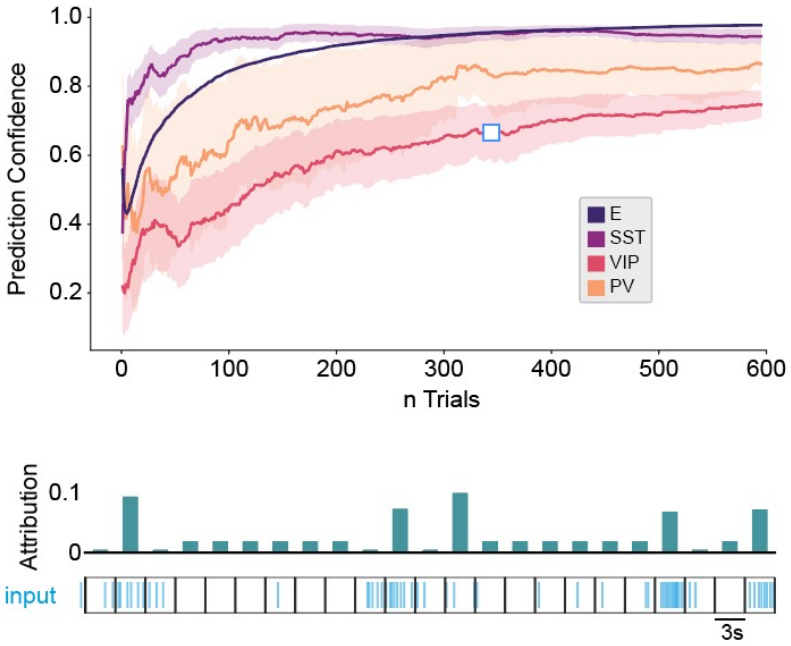
Local Latent Attention (LOLCAT)

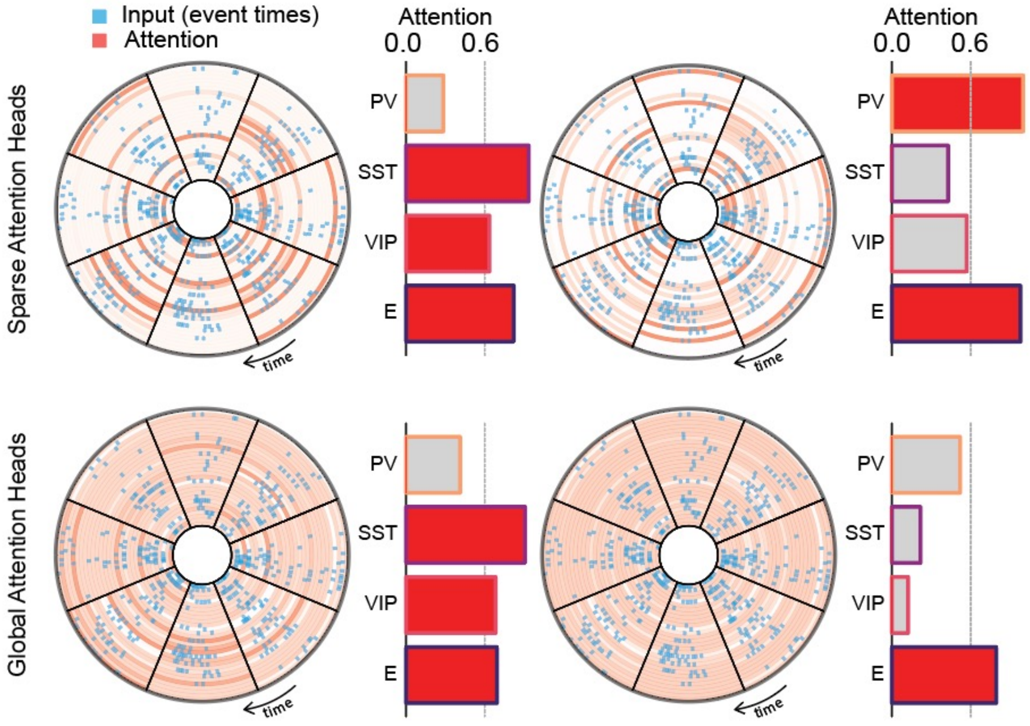
Idea: Build a rich signature of cell type through attention!

Need **multiple timescales** to resolve cell types in vivo

Schneider, Azabou, 2022: https://www.biorxiv.org/content/10.1101/2022.07.10.499487v1.full.pdf

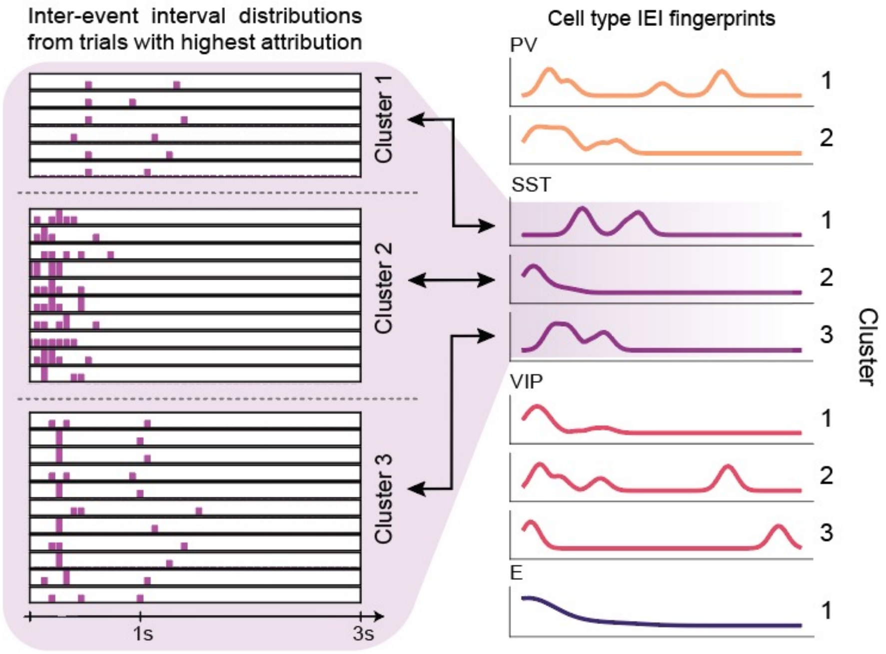
Decoding cell types from in vivo activity

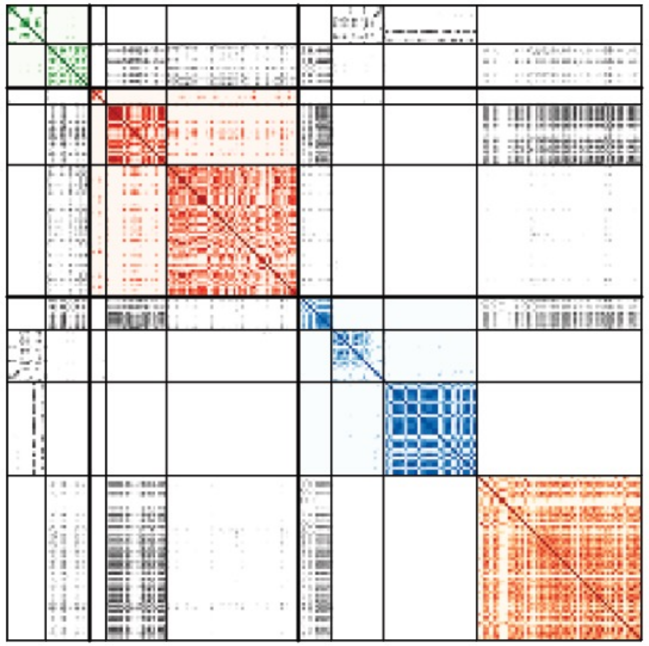




Schneider, Azabou, 2022: https://www.biorxiv.org/content/10.1101/2022.07.10.499487v1.full.pdf

Visualizing salient trials across the different cell classes





Schneider, Azabou, 2022: https://www.biorxiv.org/content/10.1101/2022.07.10.499487v1.full.pdf

Summary

- **Embedded Interaction Transformer**
  - Instead of building a population level description from the start, we decompose representation learning into two parts
  - We can transfer between animals and sessions –we think this could be the type of advance needed to build **foundational models**

- **Local Latent Attention (LOLCAT) for decoding cell type in vivo**
  - Leverage snapshots of interevent intervals in neurons + a global attention pooling to capture longer timescales
  - Can reliably decode cell type in both electrophysiology and calcium imaging data (Allen Brain Observatory)

NerDS Lab + Collaborators


Mehdi Azabou


Ran Liu


Chi-Heng Lin

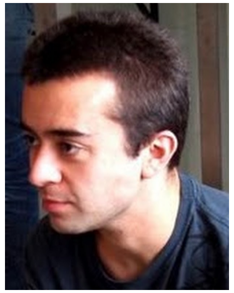

Max Dabagia



Keith Hengen
Kiran Bhaskaran-Nair
Aidan Schneider

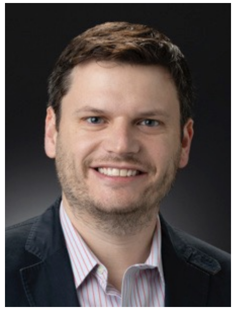
Wash U St. Louis





Mohammad Gheshlaghi Azar
Michal Valko
Bernardo A. Pires

DeepMind

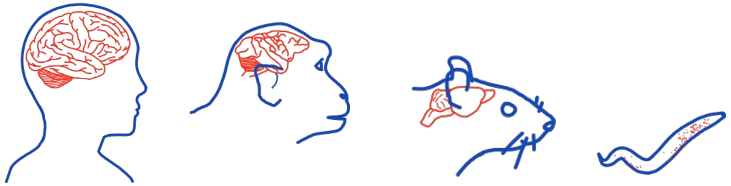


Will Gray-Roncal
Erik Johnson
Lindsey Kitchell

Johns Hopkins APL

NerDS Lab - GaTech

dyerlab.gatech.edu
github.com/nerdslab



Thank you for your attention!

dyerlab.gatech.edu - github.com/nerdslab