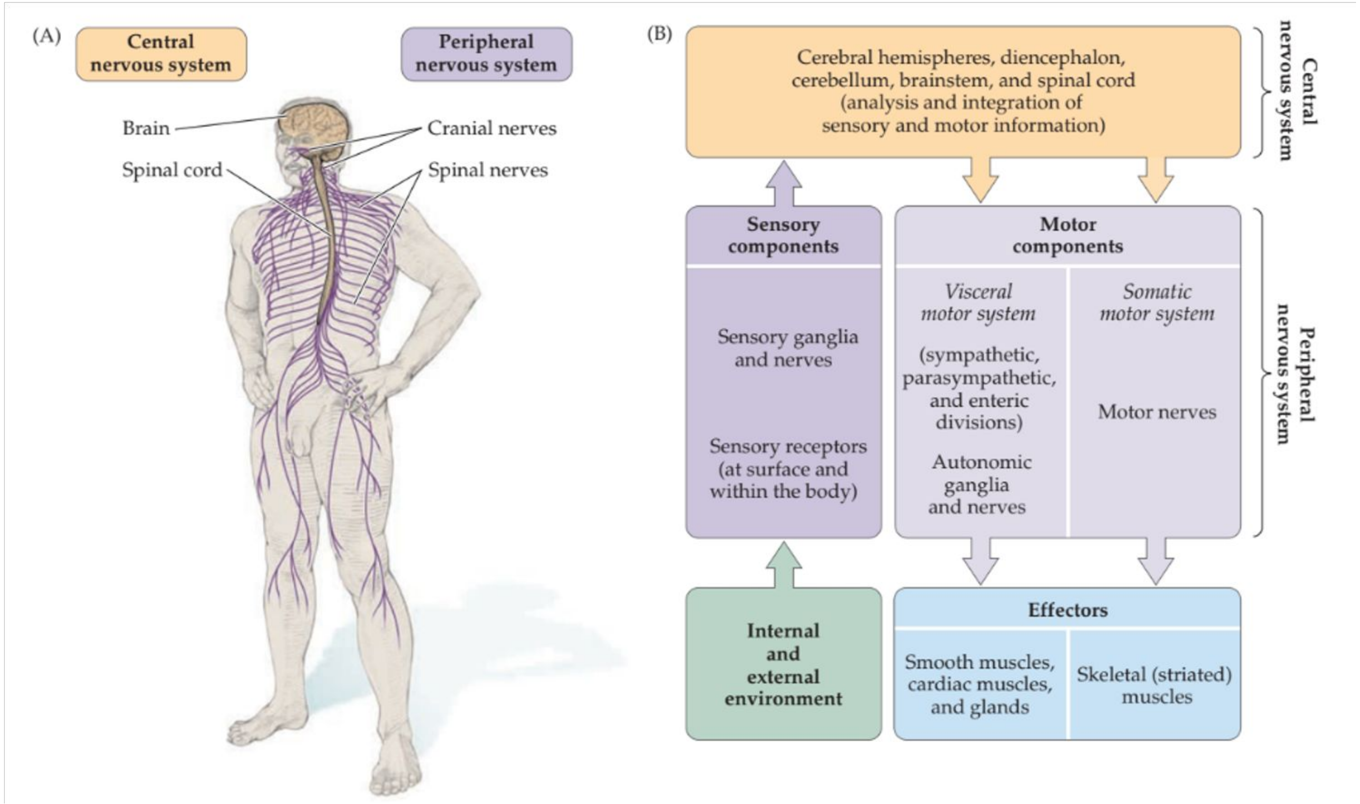
NX-435

Review

Intro & memory

The Nervous System



Engrams can be studied at the region, neural population, cell or even nucleus level...

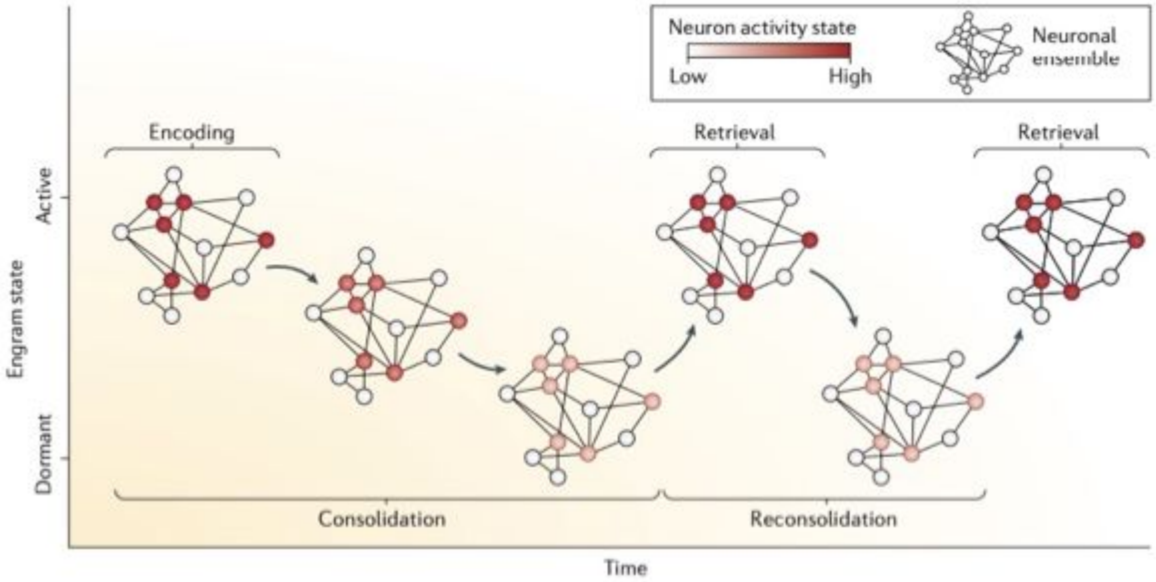
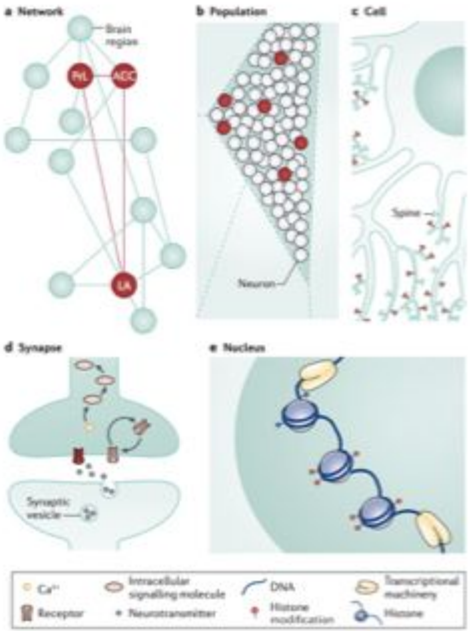


Figure 1 | **The lifetime of an engram.** The formation of an engram (encoding) involves strengthening of connections between collections of neurons (neuronal ensemble) that are active (red) during an event. Consolidation further strengthens the connections between these neurons, which increases the likelihood that the same activity pattern can be recreated at a later time, allowing for successful memory retrieval. During consolidation, the engram enters a mainly dormant state. Memory retrieval returns the engram back to an active state and transiently destabilizes this pattern of connections. The engram may be restabilized through a process of reconsolidation and re-enter a more dormant state. Therefore, an engram may exist in a dormant state between the active processes of encoding and retrieval required to form and recover the memory. In this way, an engram is not yet a memory, but provides the necessary conditions for a memory to emerge.



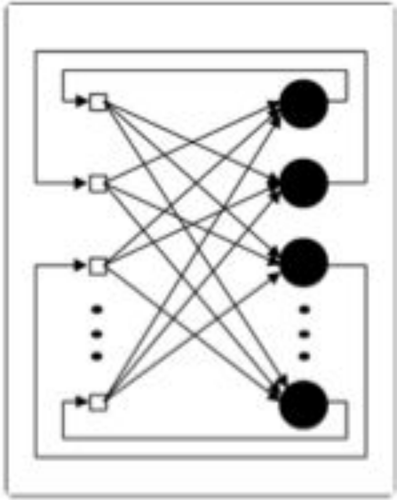
Finding the engram

Sheena A. Josselyn[1–4], Stefan Köhler[5,6] and Paul W. Frankland[1–4]

https://core.ac.uk/download/pdf/289079817.pdf

EPFL

Associative memory

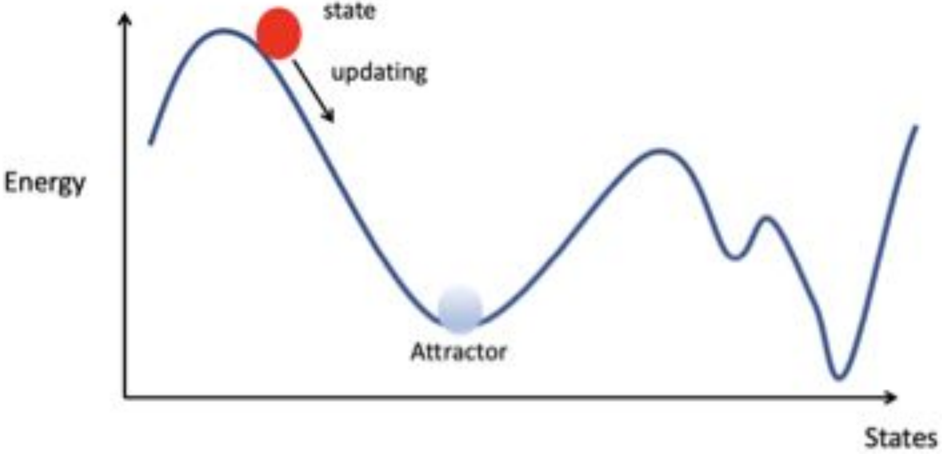


Dynamics

$$\mathbf{p}^{(t+1)} = \sigma\left(\mathbf{W}\mathbf{p}^{(t)}\right)$$

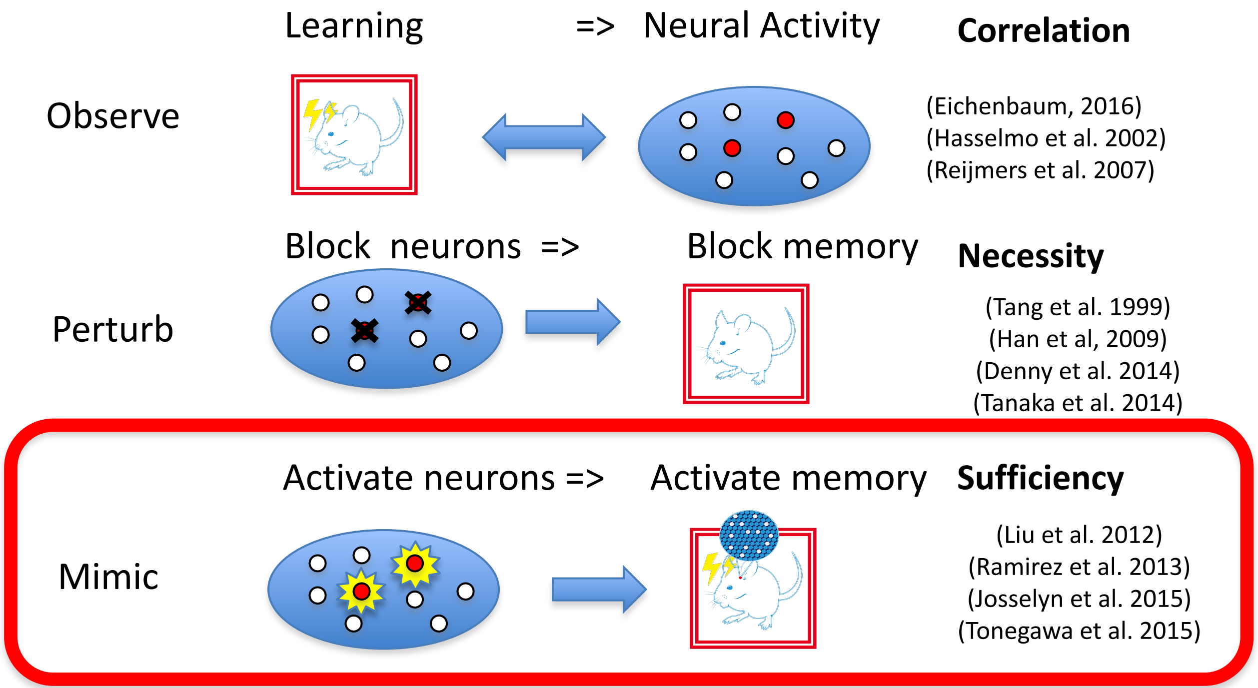
Storage of memories:

$$w_{ij} = \frac{1}{M}\sum_{\mu=1}^{M} p_i^{\mu} p_j^{\mu} \quad i \neq j$$



Hopfield PNAS 1982

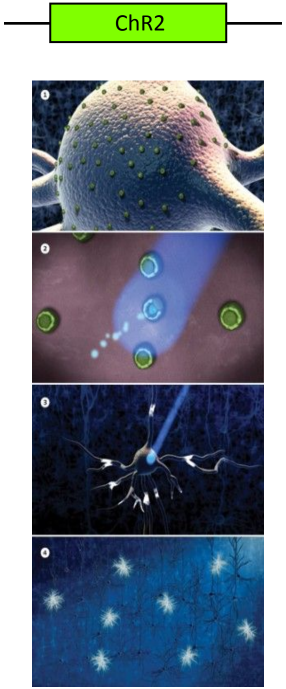
Criteria for identifying correlation, necessity & sufficiency of memory



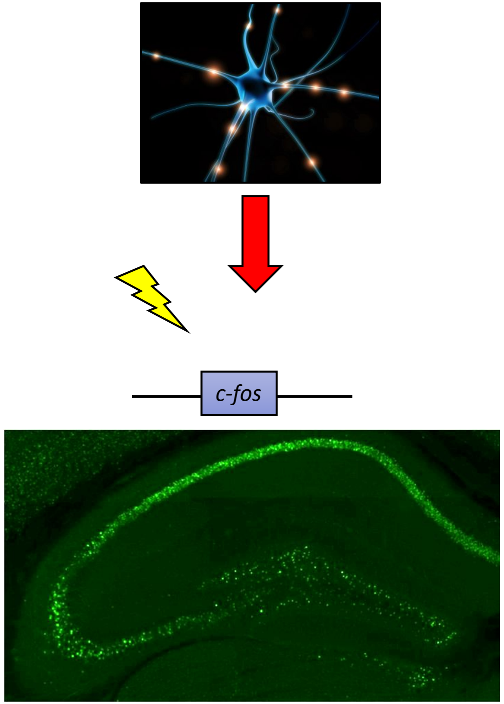
(Martin and Morris, 2000)

Slide courtesy of Prof. Steve Ramirez (BU)

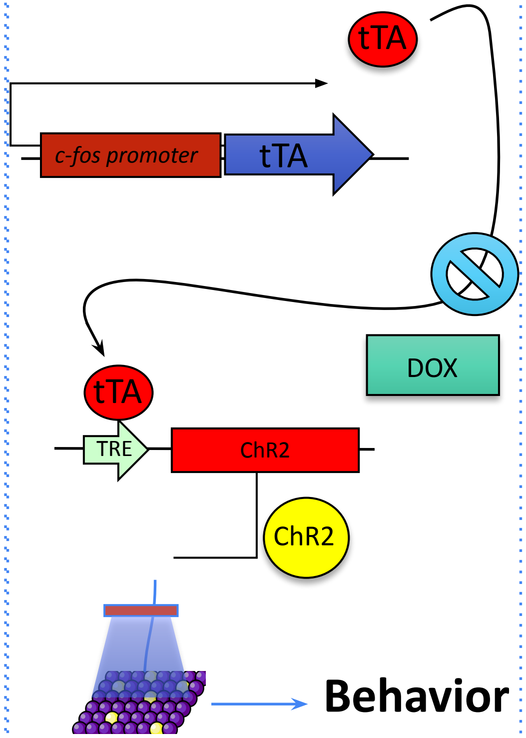
Activity-dependent and inducible optogenetics



ChR2 makes cells responsive to light



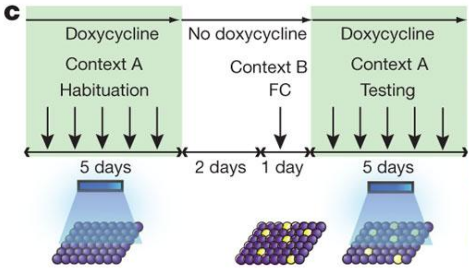
c-Fos is only expressed in active neurons

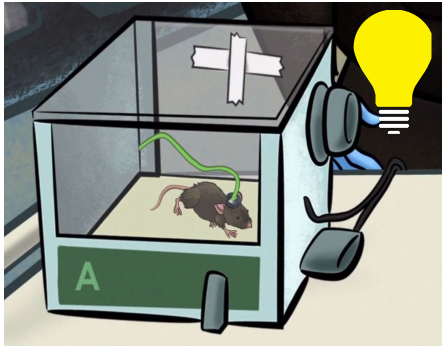


Dox can open and close windows for expressing a given gene

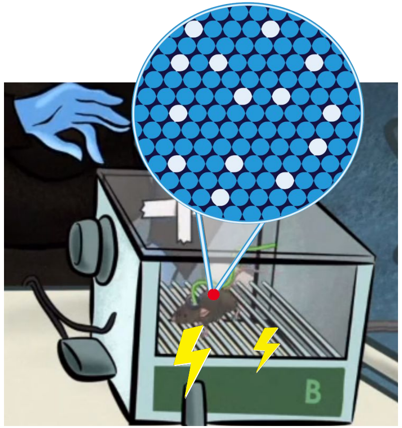
Slide courtesy of Prof. Steve Ramirez (BU)

Using c-fos-tTA mice & ChR2 during fear conditioning

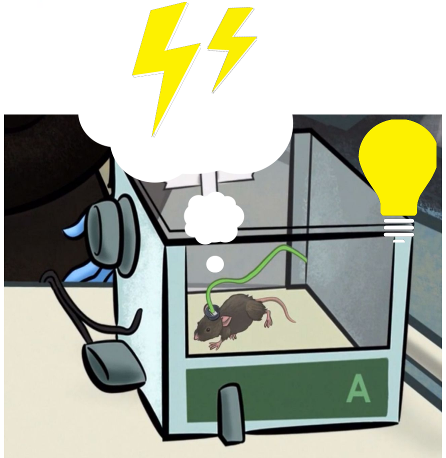




Mice were habituated in context A with light stimulation **while on Dox for 5 days**



Then taken **off Dox** for 2 days and fear conditioned (FC) in context B



Mice were put **back on Dox** and tested for 5 days in context A with light stimulation

What behavioral readout are we going to look for?

Cartoons by Prof. Steve Ramirez (BU)

Ramirez et al. 2013, Round-Up:

"In particular, a **hypothesis** of great interest is whether artificially activating a previously formed contextual memory engram while simultaneously delivering foot shocks can result in the creation of a false fear memory for the context in which foot shocks were never delivered."

- They established a paradigm to genetically tag active neurons with optogenetics
- They show that DG (in comparison to CA1) holds memories and they can indeed implant a false memory.
- The optogenetic stimulation drives activity in neurons (remarkably robust!)
- They do this during fear-learning, a powerful innate response in animals
- They additionally show that in a decision-making task the mice can act on their false memory

Reward & Reinforcement Learning

Summary

- Marr’s 3 levels provide a computational formulation for studying computations in the brain
- Decision-making is hard: the “credit assignment problem”, delayed rewards, uncertain outcomes
- Perceptual and value-based decision-making can help refine how to study and where in the brain to study
    - Reminder for the neuro-anatomy that supports visually guided decisions
    - Encoding & decoding is critical
- Decision variables (DV), evidence accumulation, and how to use decoding to closed-loop test how DV are related to actions → Change of mind in decisions – how did they test this?
- Operant and classical conditioning
- PSTH
- Dopamine (DA) neurons in VTA
- RPEs
- RL & TD learning
- How to formalize finding computations: mapping TD to DA
- Inputs to DA neurons show distributed information and even (possibly) partially computed RPEs
- Distributional RL in the DA population better fits the data

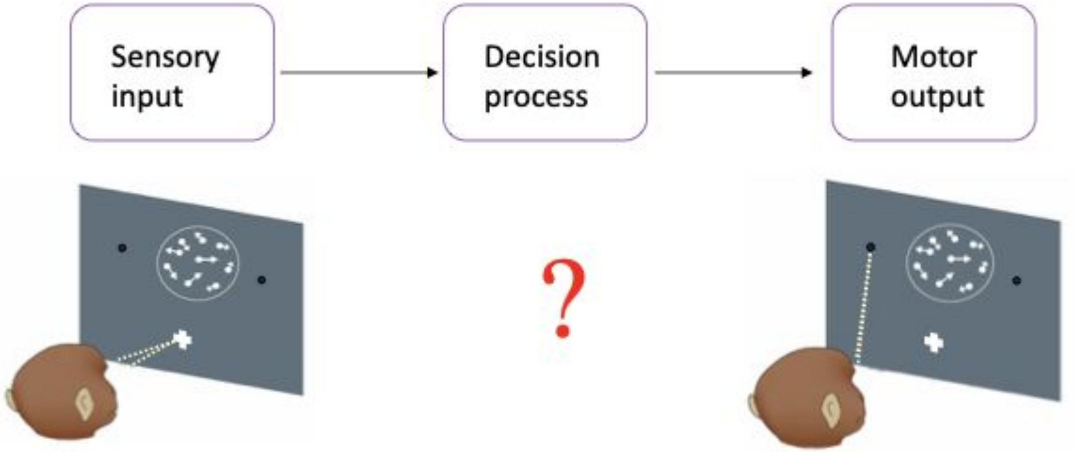
Rewards (punishment) and decision making

- **Perceptual & Value-based Decision Making**

- Operant Conditioning

- Classical Conditioning

Perceptual decision making

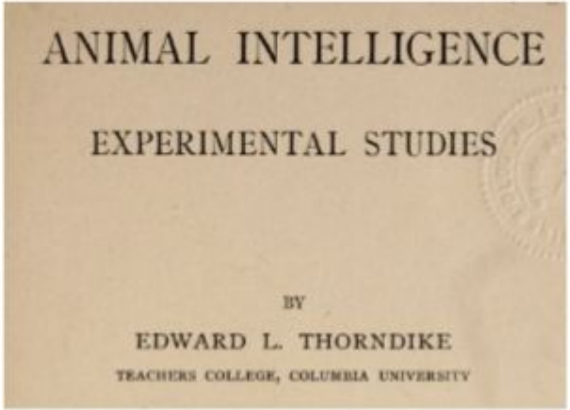


Random dot motion task

- It takes up to 1-2 seconds to decide
- Decisions unfold gradually by accumulating noisy evidence.

Operant Conditioning (also called trial-and-error learning)

Edward Thorndike
(Wikipedia, 1912)


ANIMAL INTELLIGENCE

EXPERIMENTAL STUDIES

BY

EDWARD L. THORNDIKE

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

- operant conditioning can be considered as the formation of a **predictive relationship between an action and an outcome**
- **classical conditioning is the formation of a predictive relationship between two stimuli (the CS and the US)*

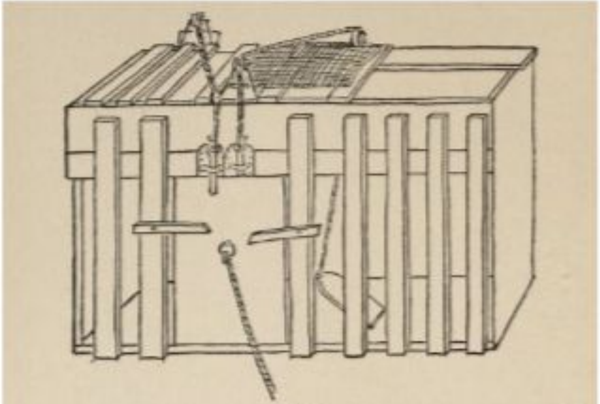
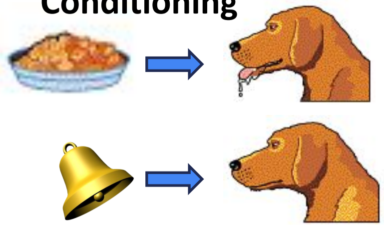


FIG. 1.

https://archive.org/details/animalintelligen00thor/page/20/mode/2up

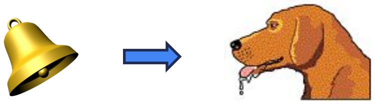
Pavlov’s classical conditioning

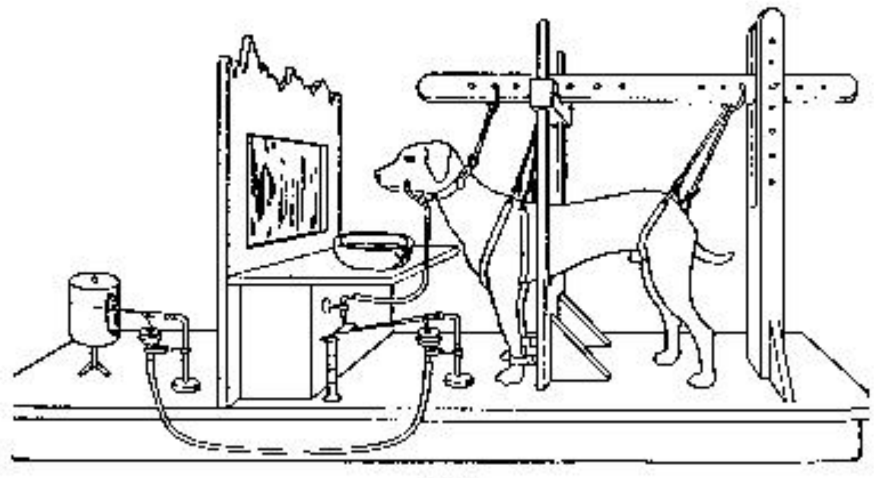
Before Conditioning



During Conditioning

After Conditioning



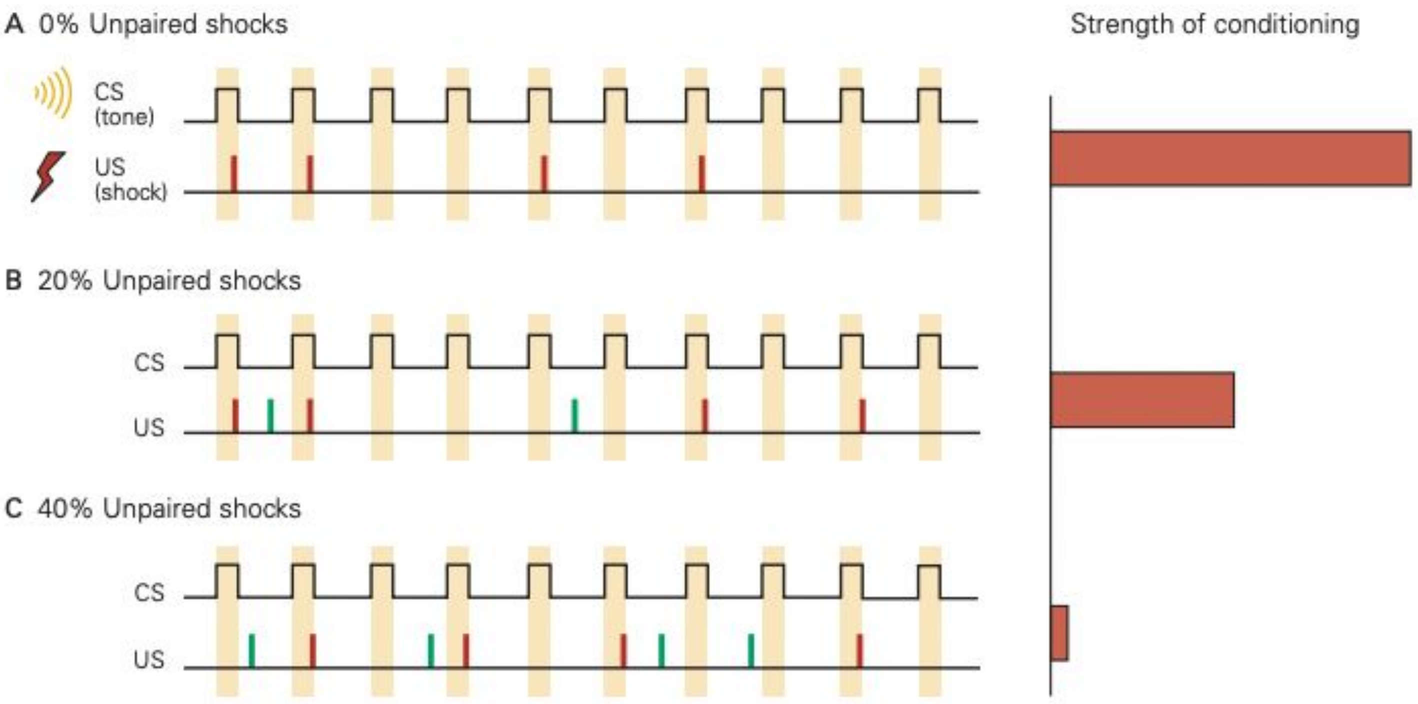


https://en.wikipedia.org/wiki/Classical_conditioning#/media/File:Ivan_Pavlov_research_on_dog's_reflex_setup.jpg

Ten of the more photogenic of Pavlov's dogs. Krasavietz (upper left), Beck, Milkah, Ikar, Joy, Tungus, Arleekin, Ruslan, Toi and Murashka (bottom right). The rest of Pavlov's dogs and their corresponding *Drosophila* memory mutants can be found on the author's webpage at www.cshl.org.

https://www.sciencedirect.com/science/article/pii/S0960982203000666

Classical conditioning depends on degree of stimulus-outcome correlation



Kandel. Figure 65–12

Kamin’s blocking experiment

1. Conditioning 2. After conditioning 3. 2nd conditioning 4. Test

Kamin, L. J. (1969). Predictability, Surprise, Attention, and Conditioning. In B. A. Campbell, & R. M. Church (Eds.), Punishment Aversive Behavior (pp. 279-296). New York: Appleton- Century-Crofts


predicts food already.
No surprise…

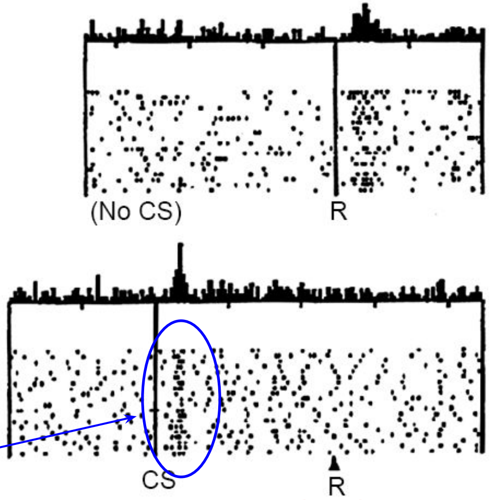
“Blocking ”

- **Learning occurs only when expectation is violated!**
- *What is the neural basis of this?*

Dopamine neurons in the ventral tegmental area

Before conditioning

After conditioning

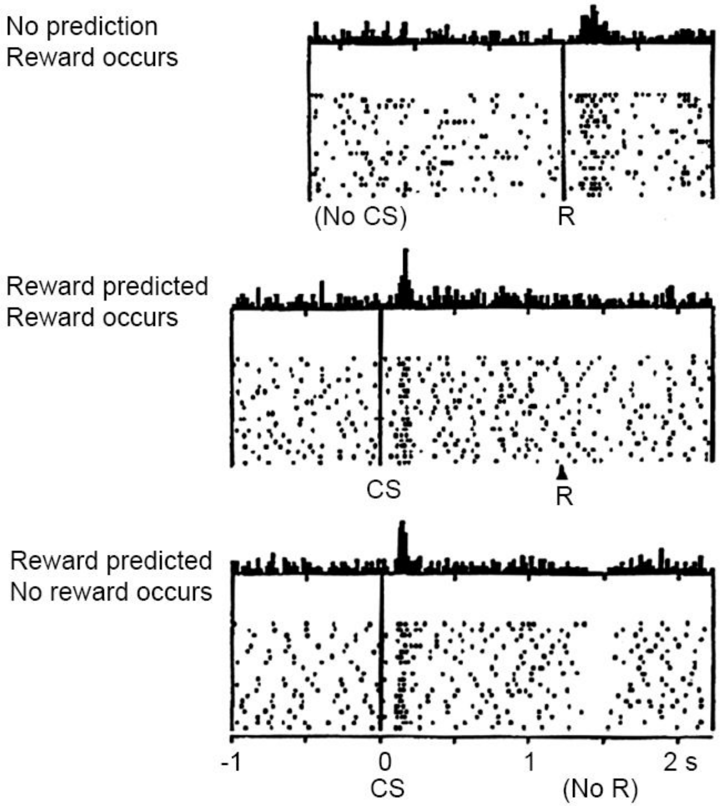


What is this for?

- Lack of reward responses when the reward was fully predicted

(Schultz, Dayan, Montague, 1997)

Dopamine as ~~reward~~ temporal difference (TD) error: reward prediction errors!



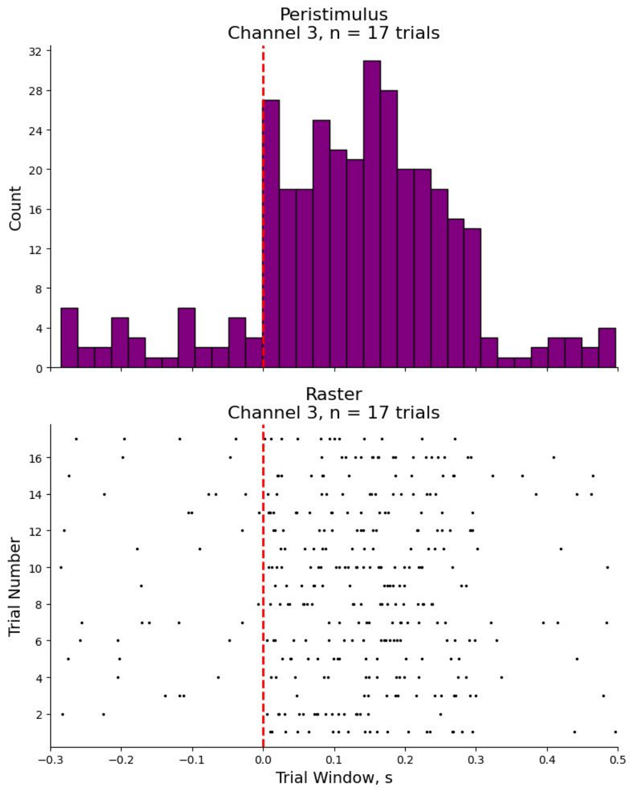
- Dopaminergic (DA) neurons fire phasically (100–500 ms) after unpredicted rewards or cues that predict reward.
- Their response to reward is reduced when a reward is fully predicted (the phasic firing happens at cue presentation).
- DA activity is suppressed when a predicted reward is omitted (negative prediction error).

(Schultz, Dayan, Montague, 1997)

Key concept: peri-stimulus time histogram

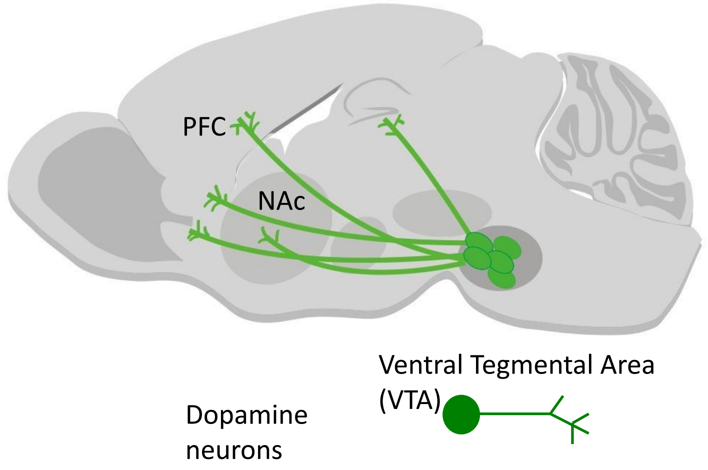
The Peri-Stimulus Time Histogram (PSTH) plots the average firing rate of a neuron over time relative to the onset of a stimulus. Here's how it's typically calculated:

1. Define a time window around the onset of the stimulus.
2. Divide this time window into small bins.
3. Count the number of spikes (action potentials) that occur within each bin across multiple trials.
4. Average the spike counts across trials for each bin.
5. Plot the average spike count (firing rate) for each bin as a function of time.

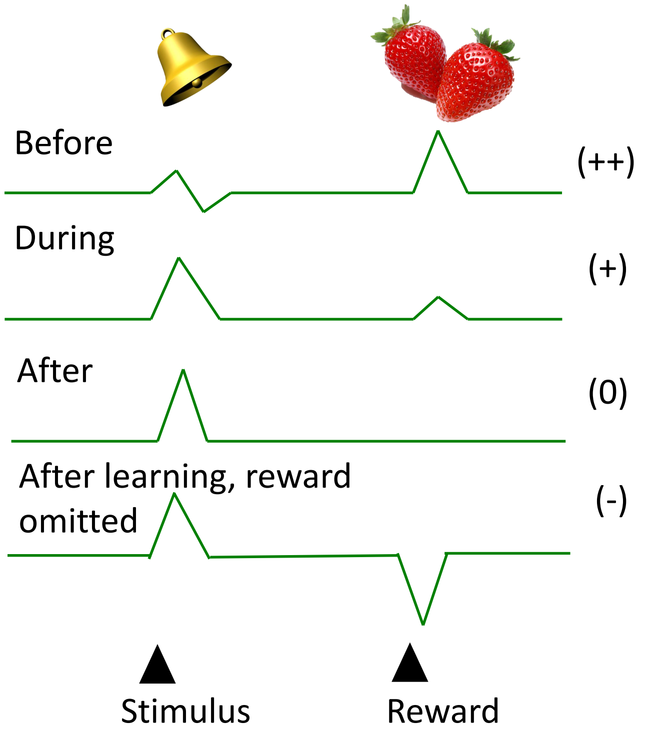


https://colab.research.google.com/github/MMathisLab/Nx-435_EPFL/blob/main/Notebooks/Demo_PSTH.ipynb

Dopamine circuitry of the brain



- Dopaminergic neurons are ~55–65% of VTA neurons
- The rest are mostly GABAergic inhibitory neurons or Glutamatergic neurons

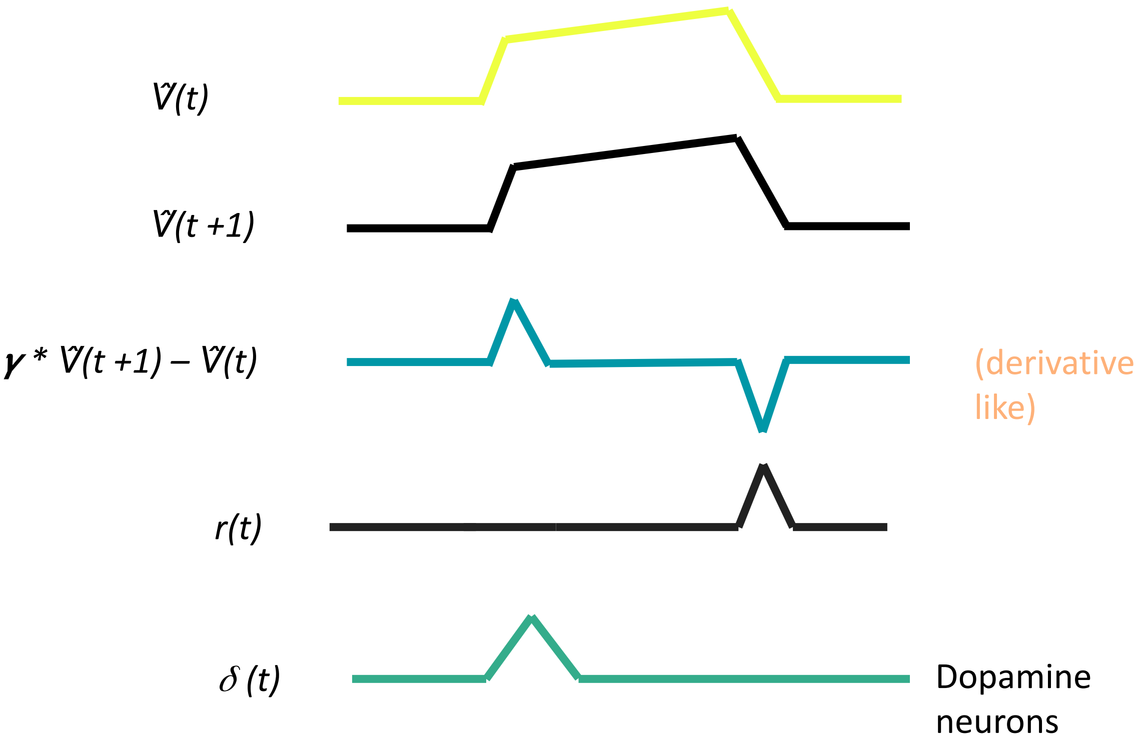


How could a system encode a temporal difference (TD) error

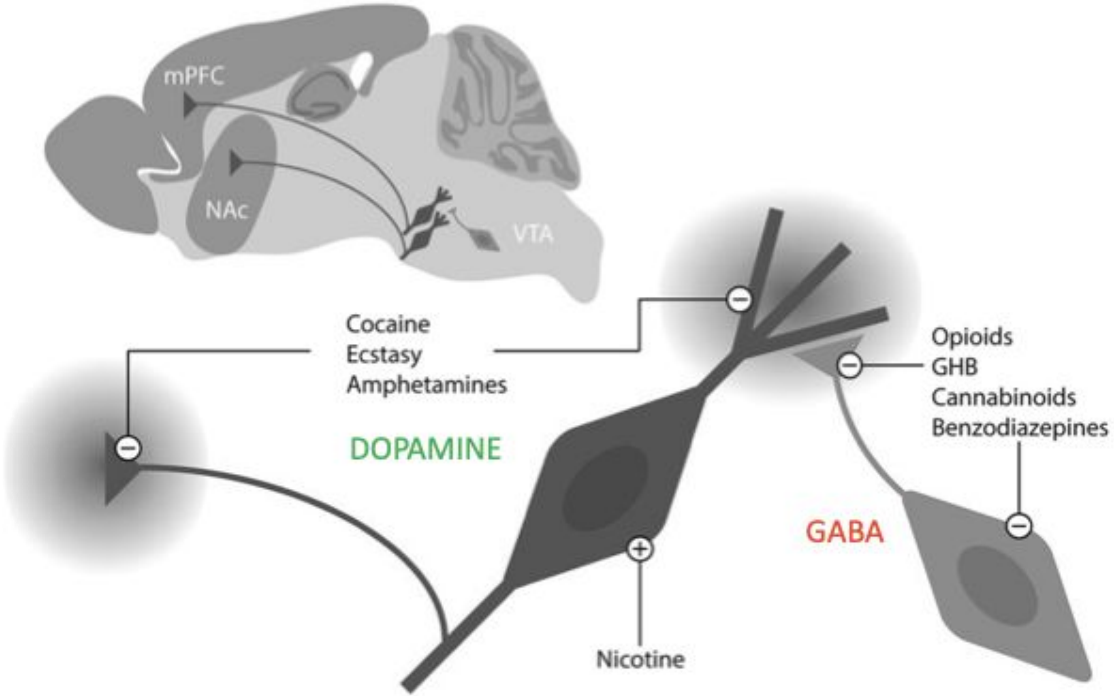
TD error as a derivative-like computation:
(neurally doable!)

$\delta\,(t) = r(t) + \gamma * \hat{V}(t+1) - \hat{V}(t)$

t = time
r = reward
$V(t)$ = value
γ = discount factor
δ = prediction error



Dopamine circuitry of the brain: drugs have strong effects



Addictive drugs cause an increase in mesocorticolimbic dopamine through three distinct cellular mechanisms:

- (1) direct activation of dopamine neurons (e.g., nicotine)
- (2) indirect disinhibition of dopamine neurons [opioids, gamma-hydroxybutyric acid (GHB), cannabinoids, and benzodiazepines]
- (3) interference with dopamine reuptake (cocaine, ecstasy, and amphetamines).

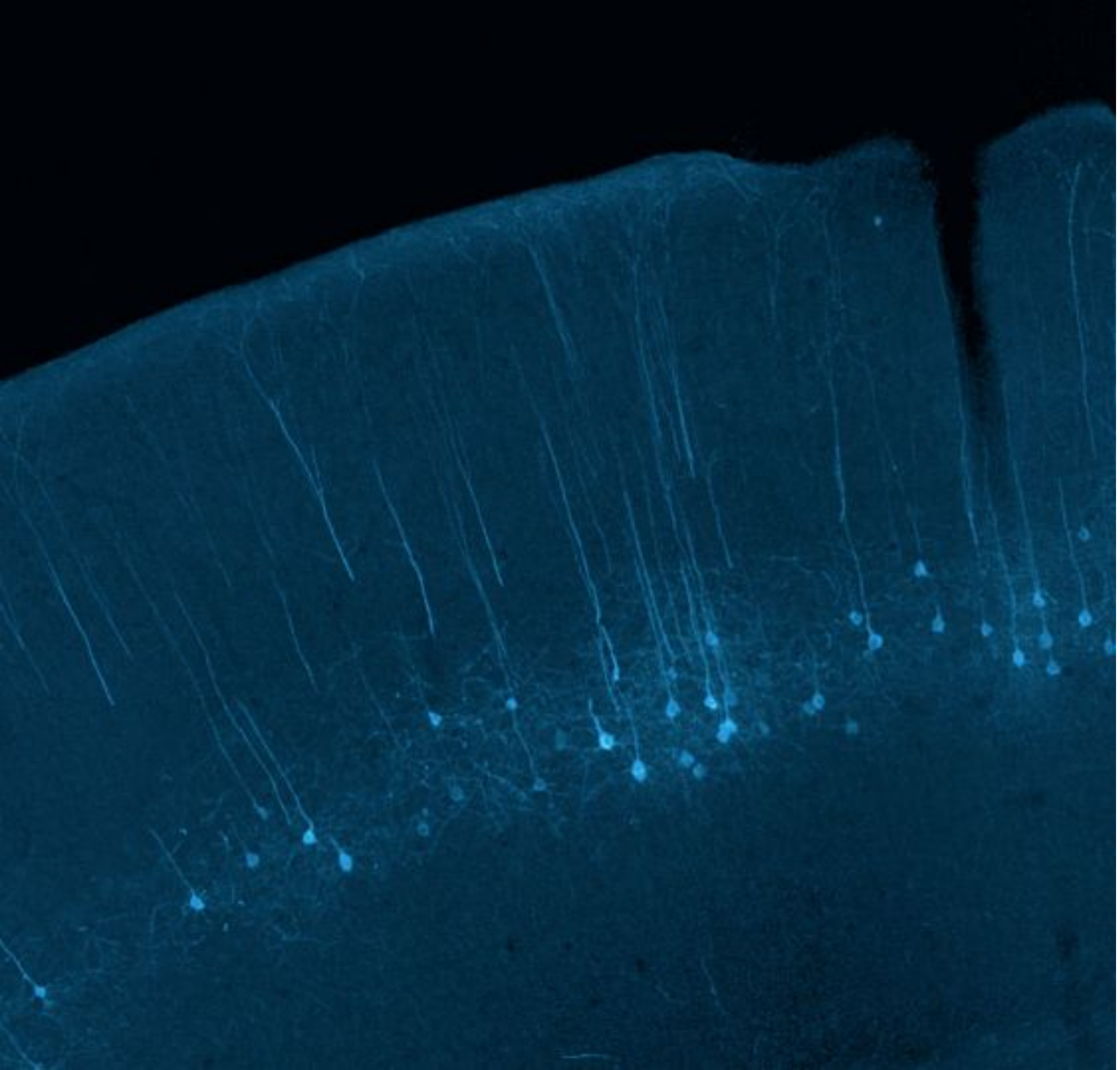
Drug-Evoked Synaptic Plasticity Causing Addictive Behavior

Christian Lüscher

Journal of Neuroscience 6 November 2013, 33 (45) 17641-17646; DOI: https://doi.org/10.1523/JNEUROSCI.3406-13.2013

Cohen et al. 2012,
Paper round-up

- They identify 3 types of neurons in the ventral tegmental area.
- They differentiate dopaminergic and GABAergic neurons using optogenetic tools.
- They characterize dopaminergic neurons diversity (excited by either reward, reward-predicting CS or both) which seems to be related to the effect of training.
- They show that some dopaminergic neurons might not strictly follow canonical RPE coding.
- They show that GABAergic neurons parametrically encoded the value of upcoming outcomes.



Motor learning & neuromodulation

EPFL

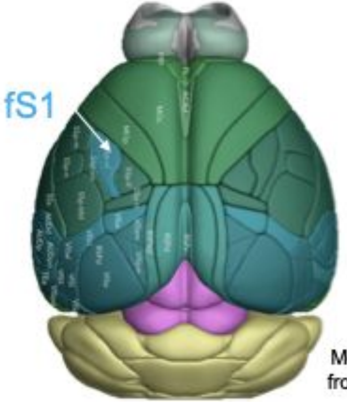
Part 1 conclusions

- **Mice can learn to rapidly learn a new sensorimotor mapping (motor adaptation)**

- **Forelimb S1 is essential to** adaptation (in this task), but inactivation of S1 did not effect motor control

- Theory-guided experiments suggest S1 does not exclusively house an internal model, and sensory prediction errors (vs. reward prediction errors) drive learning

- Ongoing work: what are neurons in S1 encoding …



Mouse brain diagram from the Allen Institute

What other systems are can modulate motor learning?

Summary Part 2

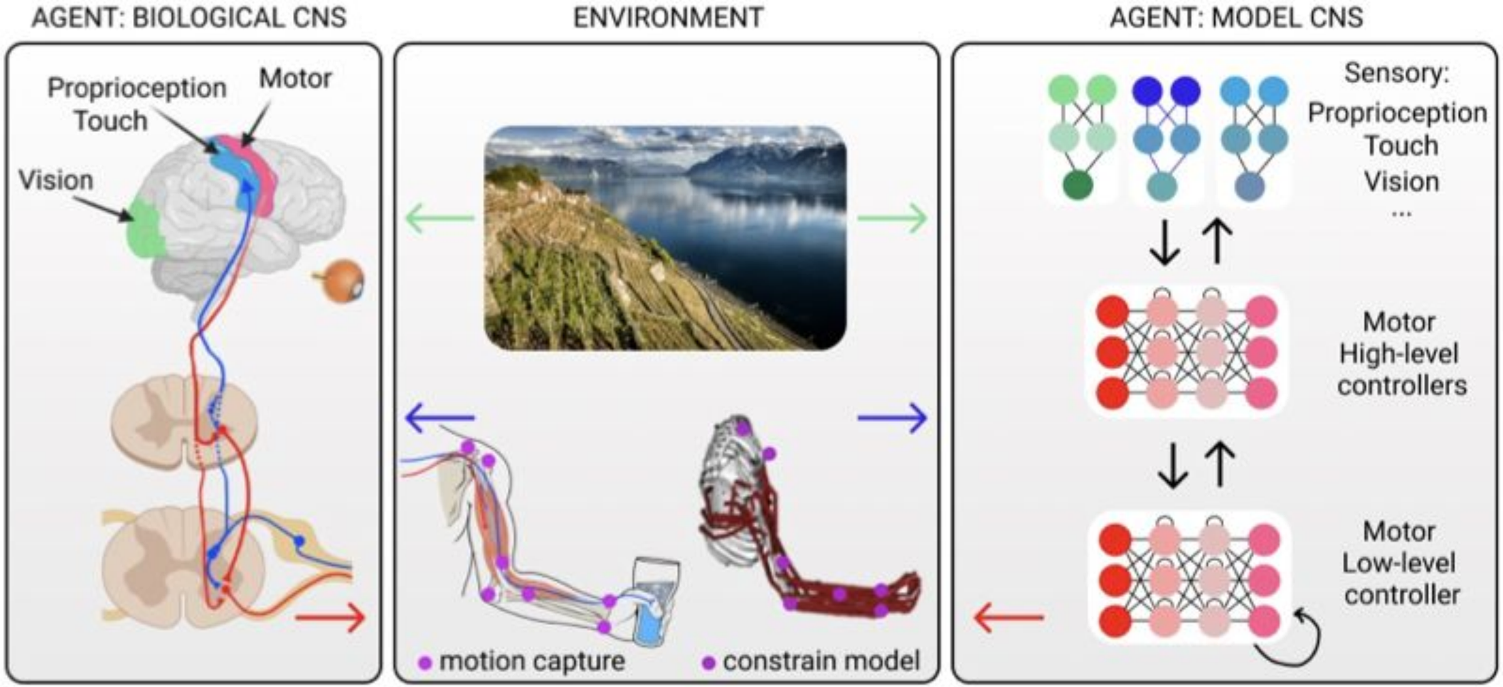
- Neuromodulators refer to neurotransmitters that act primarily though G-protein couple receptors, rather than ligand-gated excitation and inhibition.
  - Neuromodulators can have diverse effects due to the variety of their receptors.

- Acetylcholine is one commonly studied NM.
  - It is associated with mediating plasticity and arousal, as well as encoding cues and outcomes.

- Bioelectric interfaces are a tool for manipulating NMs that can act on a more rapid timescale than pharmaceuticals.
  - They also have high potential for targeted treatment due to closed-looping.

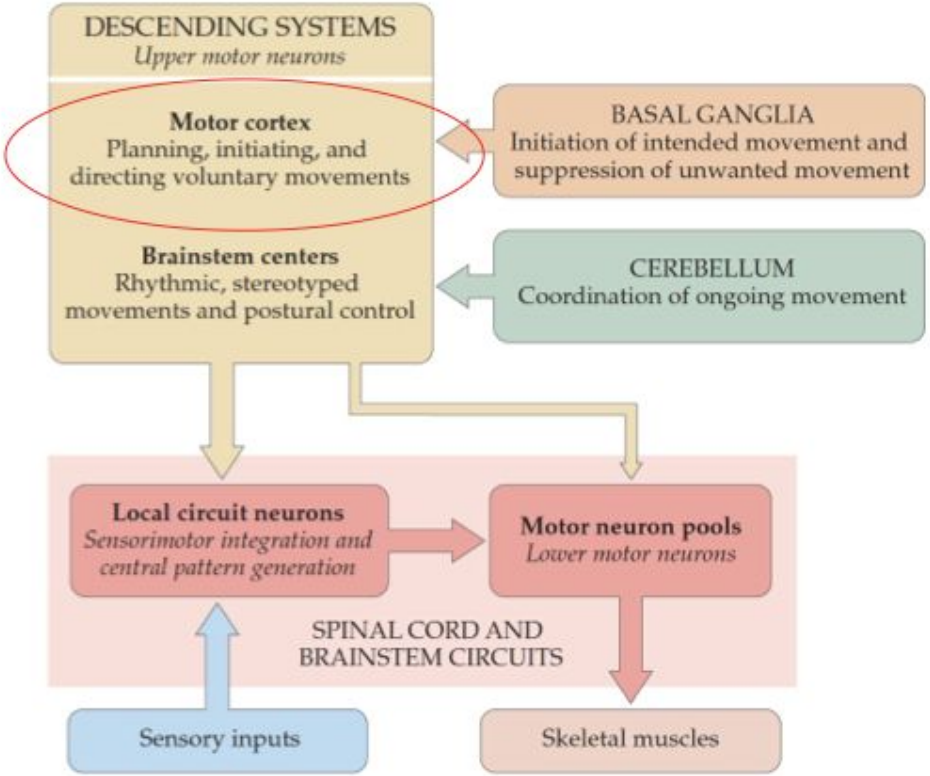
- VNS is a BMI that can enhance rehabilitation after stroke through closed-loop stimulation.
  - There is evidence that this effect is mediated, in part, by activating cholinergic neuromodulation.

Reverse engineering adaptive behavior



Hausmann et al. Current Opinion in Neurobiology 2021

The neural control of movement



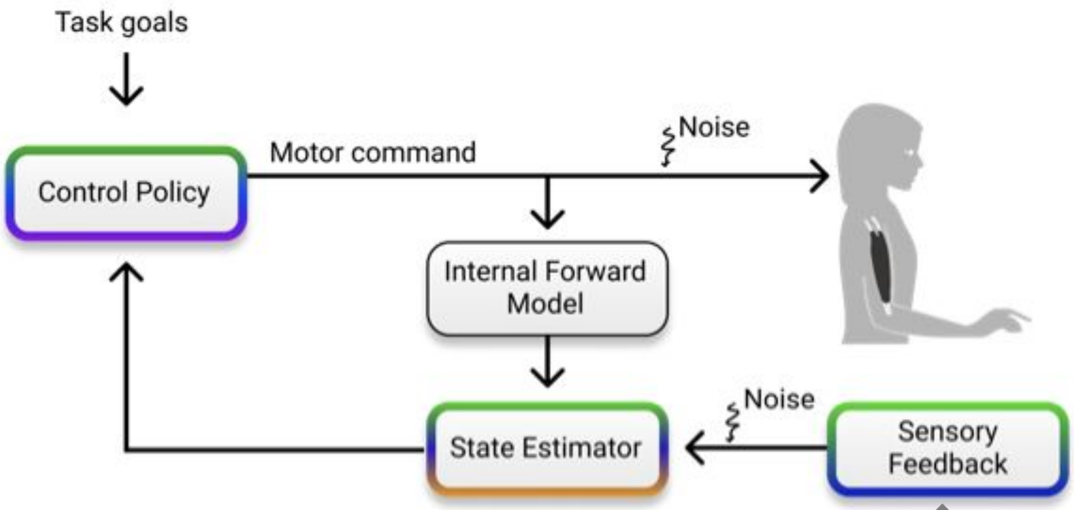
Purves, Figure 16.1

Four systems make essential and distinct contributions to motor control:

- The **spinal cord** (and brainstem circuits)
- The **cerebellum**
- Descending control centers in the cerebral cortex and brainstem
- The basal ganglia

Theory-guided framework for studying motor learning

How do animals (and agents) learn to adapt?



Sensory prediction error

$$m_{k+1} = \underbrace{\widehat{m}_k}_{\text{“= 0” (memory hyp.)}} + \underbrace{K_k \, (s - \hat{s})}_{}$$

- Wolpert et al. Science 1995, Todorov & Jordan 2002
  Izawa & Shadmehr 2011, Kawato & Gomi 1992, …, Scott 2004

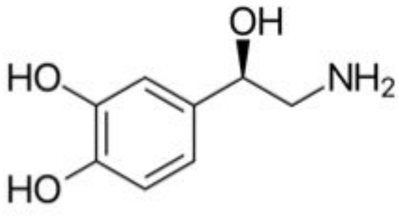
What are neuromodulators?

Neurotransmitters (NTs) refer to any chemical released from neurons that activate receptors on other neurons.

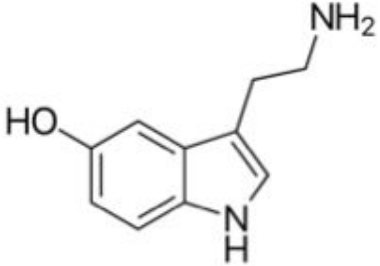
Glutamate and GABA are the most common NTs, accounting for approximately 90% of all neurons!

Neuromodulators (NMs) refer to a subset of NTs that alter do not directly activate ion-channel receptors, but instead alter neural responses to excitation and inhibition.

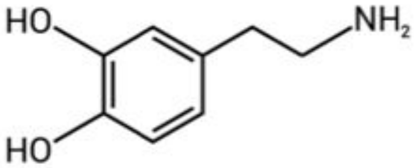
Noradrenaline



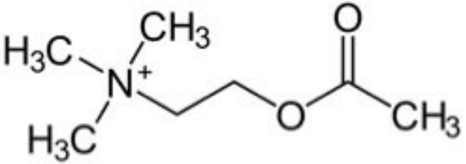
Serotonin



Dopamine

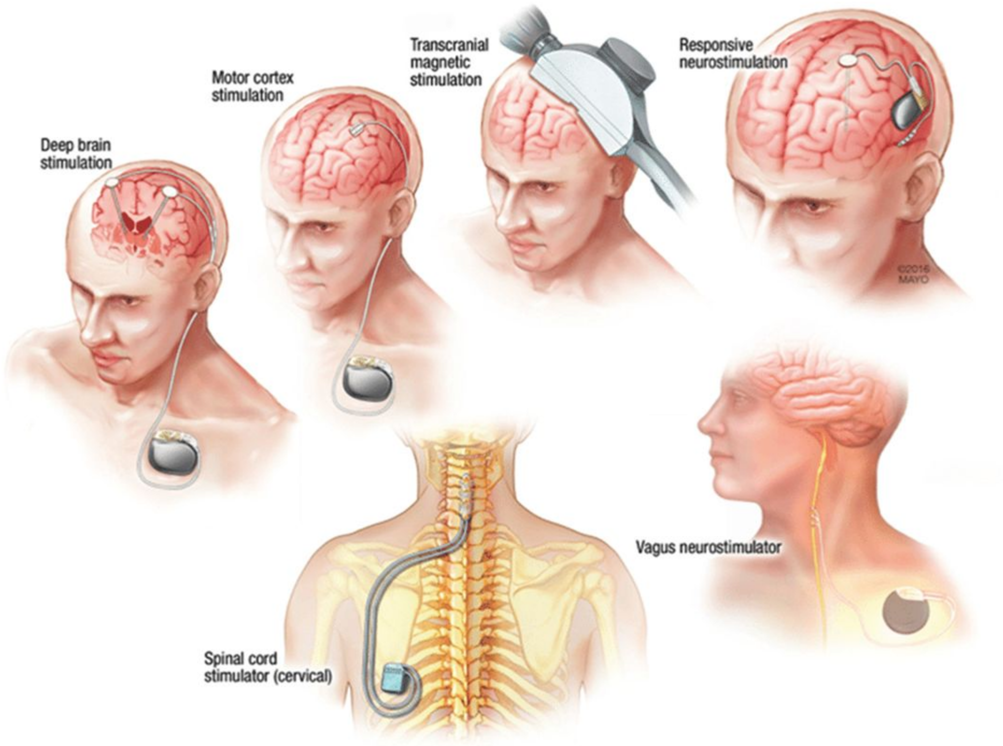


Acetylcholine



Avery & Krichmar, 2017

Neurostimulation devices can alter CNS activity across broad timescales



DBS for Parkinson's disease and essential tremor.

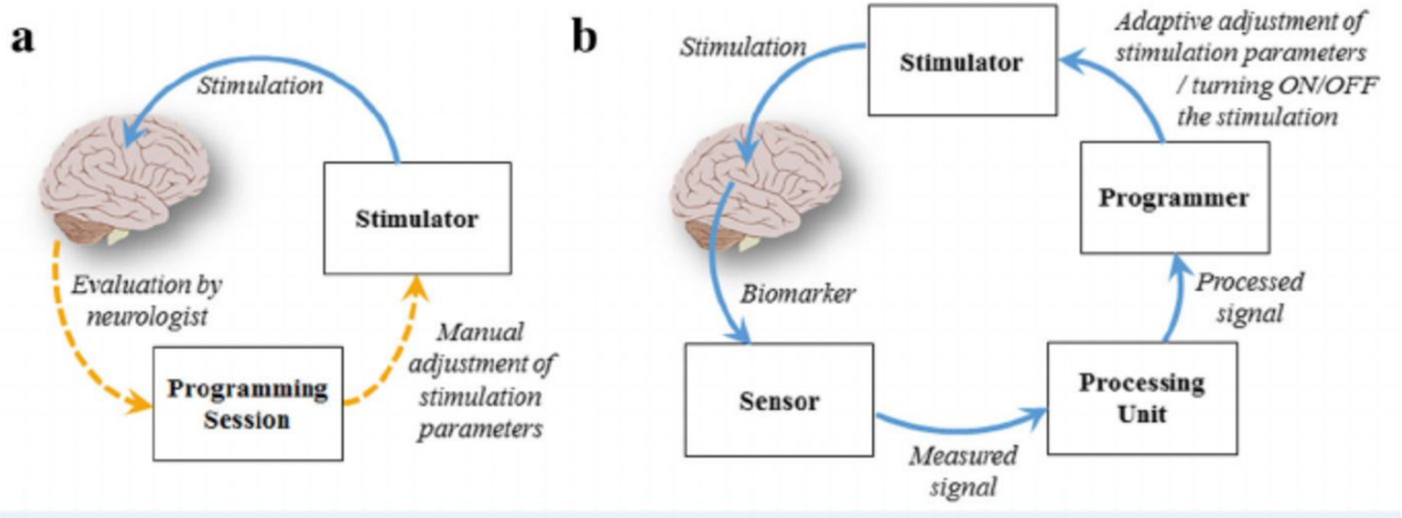
SCS for chronic pain.

TMS for depression, OCD, and migraines.

Closed-loop stimulation may increase neurostimulation efficacy and reduce side effects

Open-loop stimulation

Closed-loop stimulation



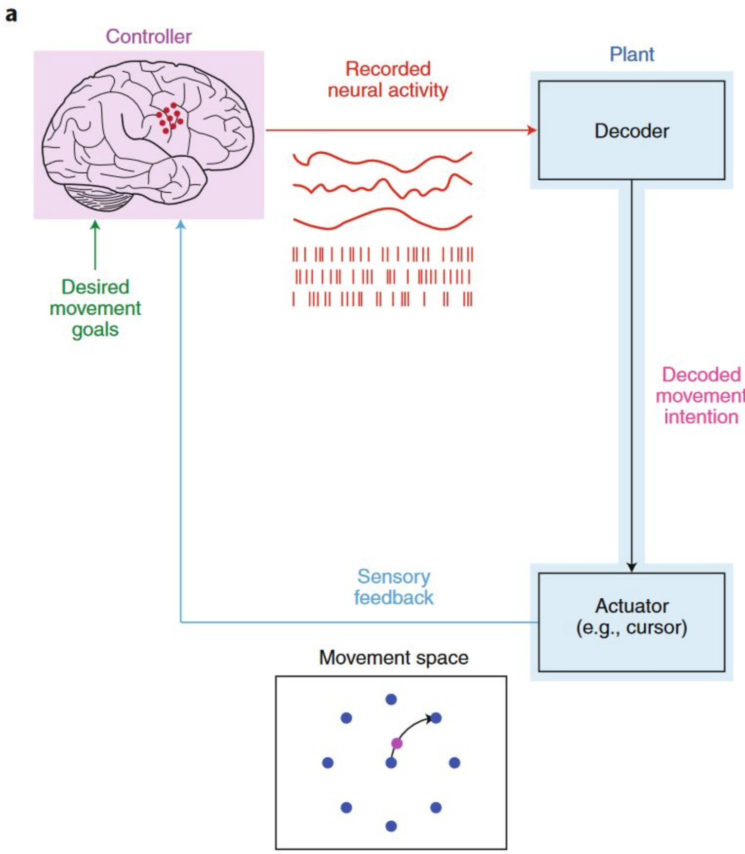
How can we develop relevant closed loop stimulation paradigms?

Closed-loop paradigms can be targeted towards many types of triggers:

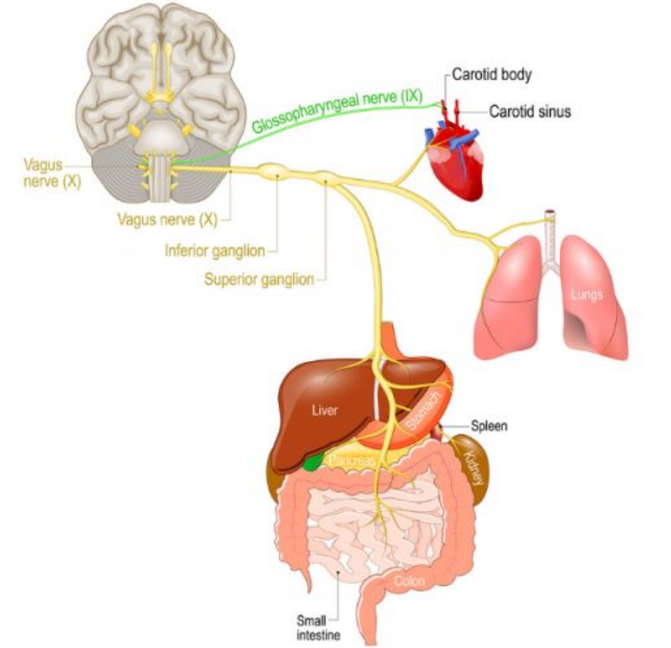
Environment triggers - tones, or task outcomes

Biomechanical triggers - certain movements, or tactile sensations

Physiological triggers - neural activity, muscle activity, hormones

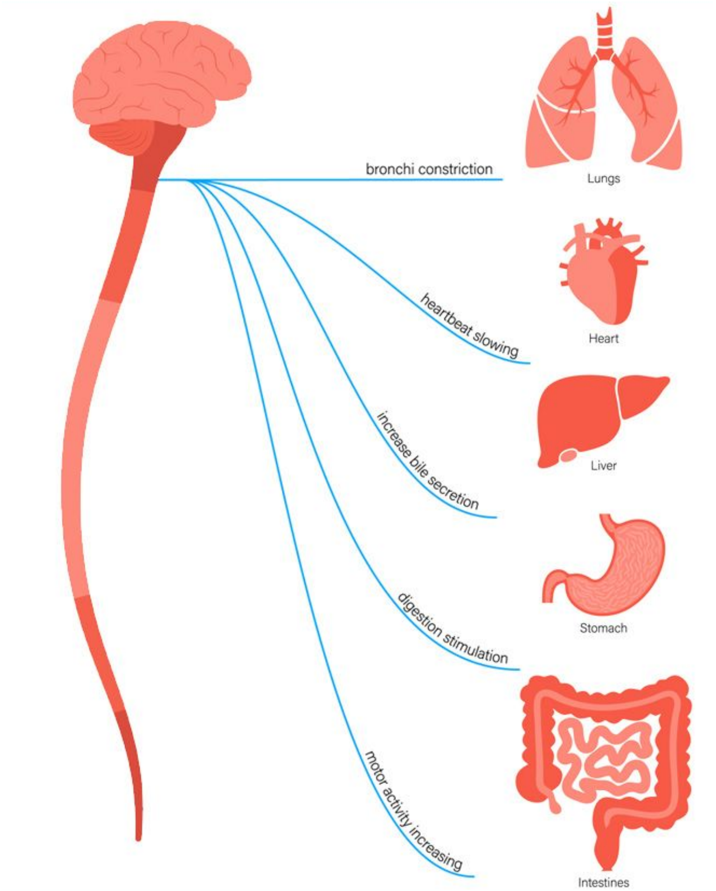


Vagus nerve function and anatomy

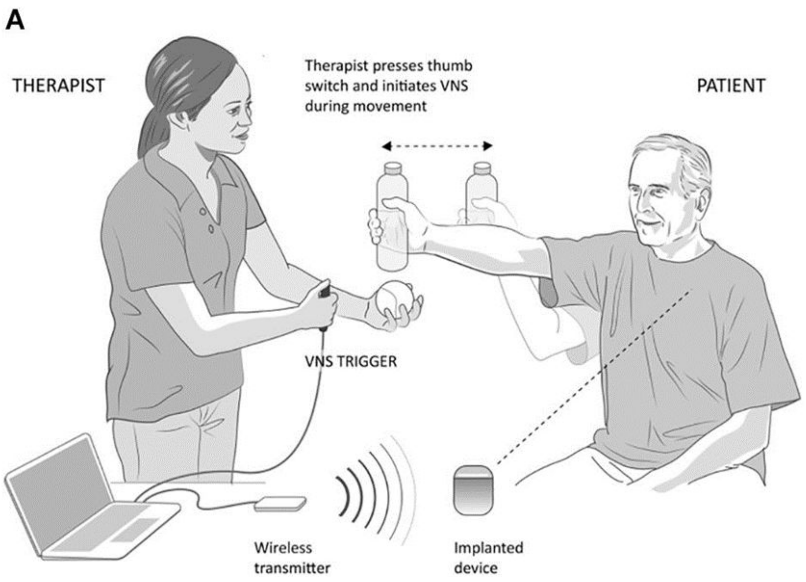


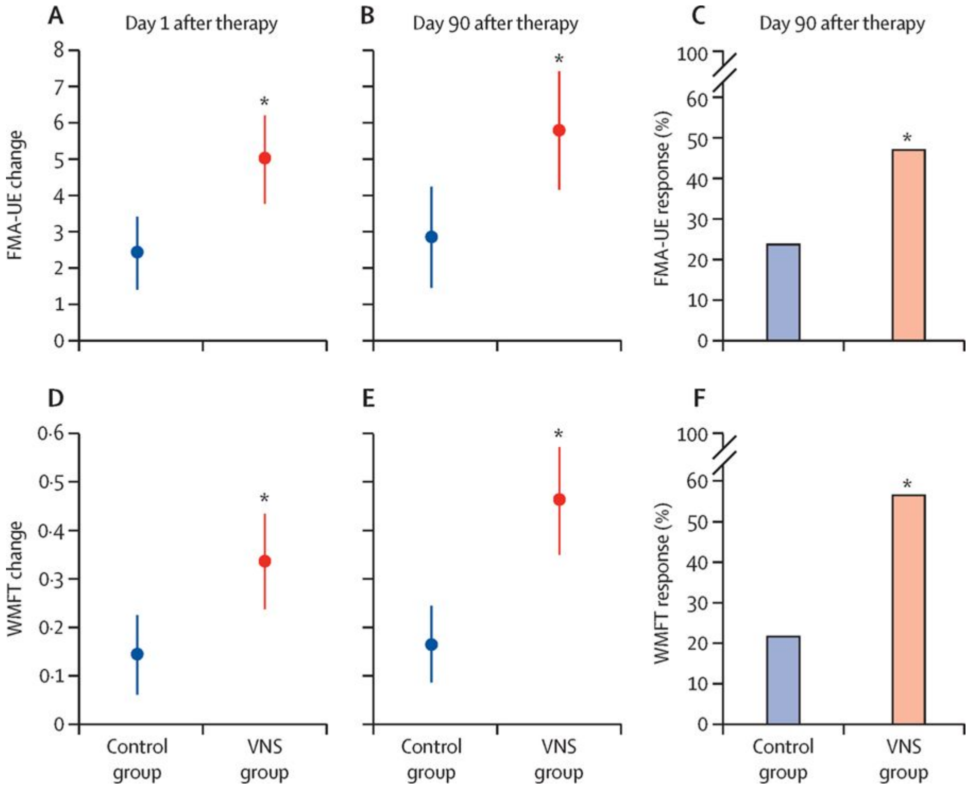
The VN innervates most visceral organs

Stimulating the VN activates the parasympathetic nervous system



VNS enhances stroke rehabilitation

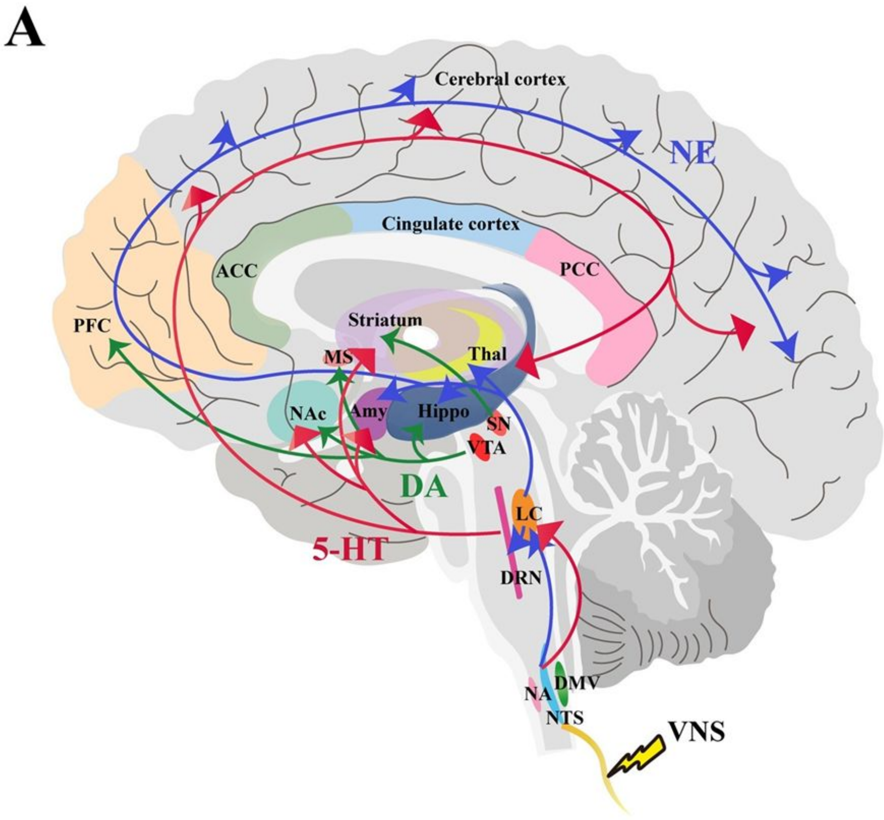




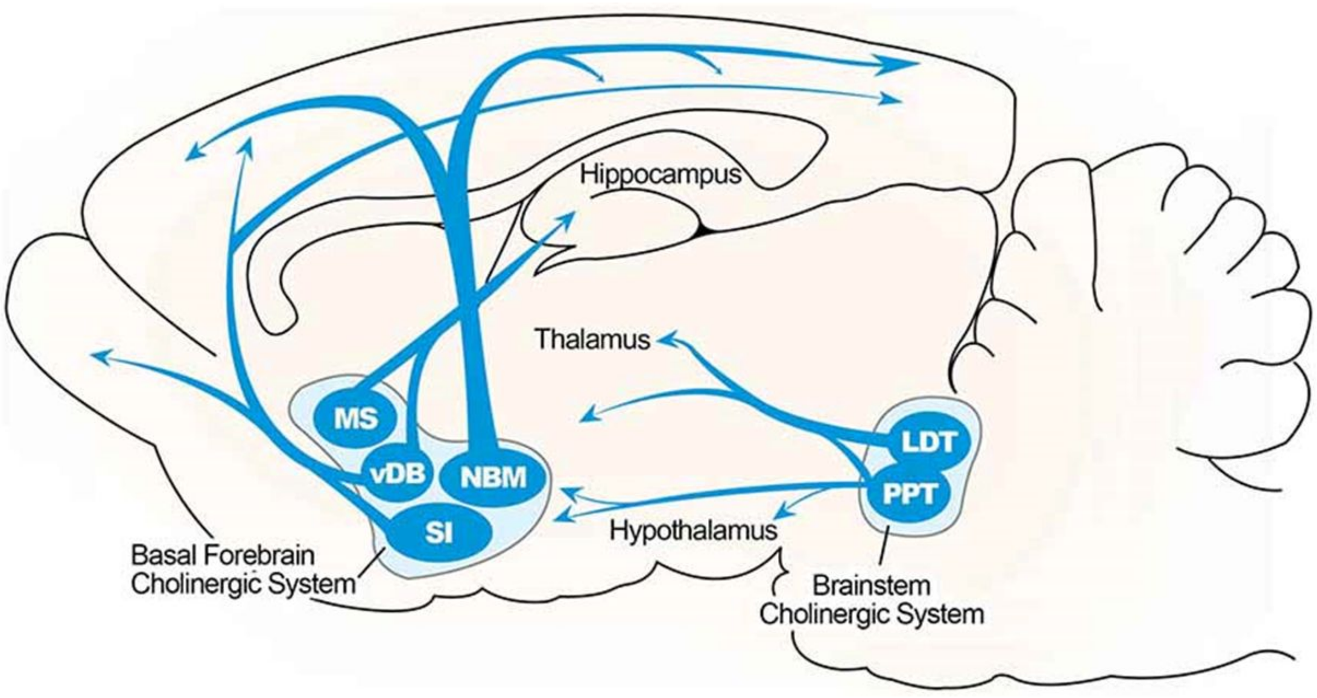
VNS activates multiple neuromodulatory systems

VNS is a “messy” stimulus:

Serotonergic, dopaminergic, noradrenergic, and cholinergic systems are all activated.

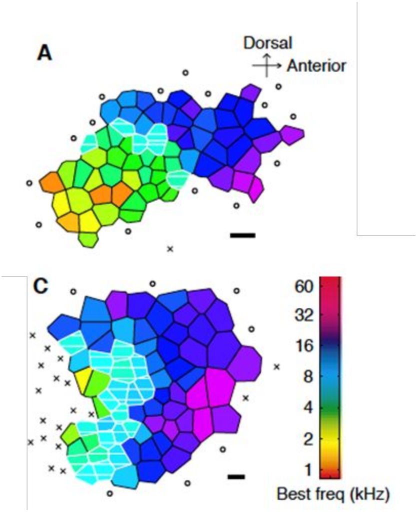


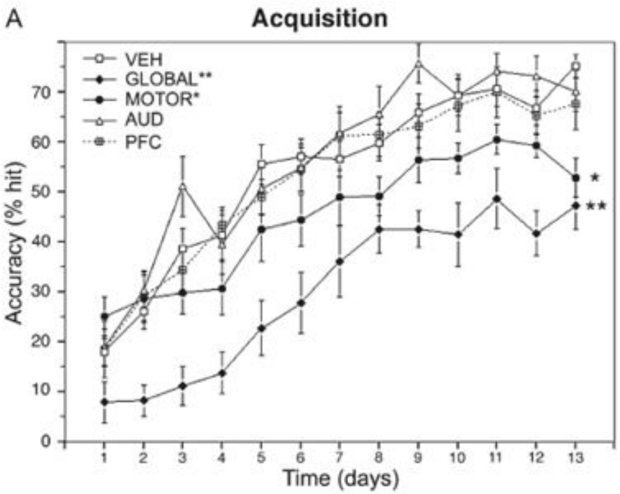
The cholinergic neuromodulatory system

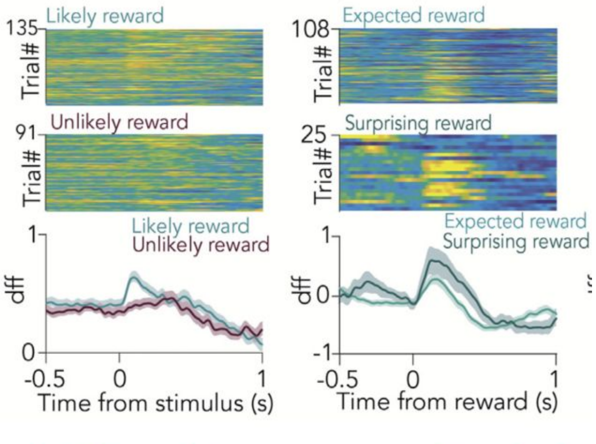


Cholinergic neuromodulation is closely linked with learning and plasticity

Cholinergic neuromodulation is a strong candidate for mediating VNS effects







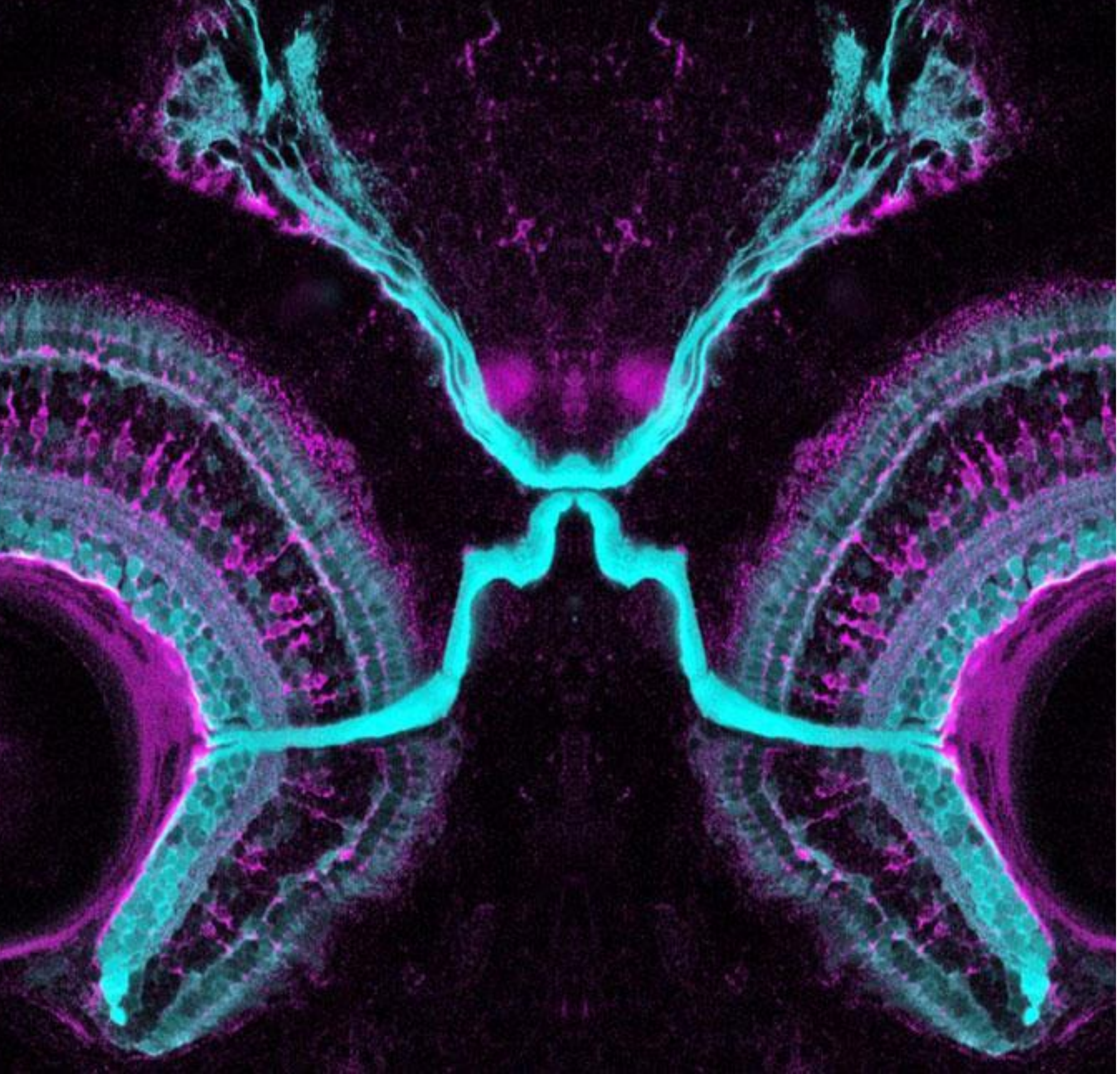
Cholinergic stim. enhances plasticity

Cholinergic activity impacts motor learning

Cholinergic neurons encode cues and outcomes

Bowles et al., 2021
Paper round-up

- VNS paired with success enhances skilled motor learning in healthy animals
- Enhanced motor performance is due to accelerated consolidation of an expert motor plan
- Enhanced motor learning depends on cholinergic neural activity in the basal forebrain
- In primary motor cortex, VNS specifically modulates outcome-activated neurons

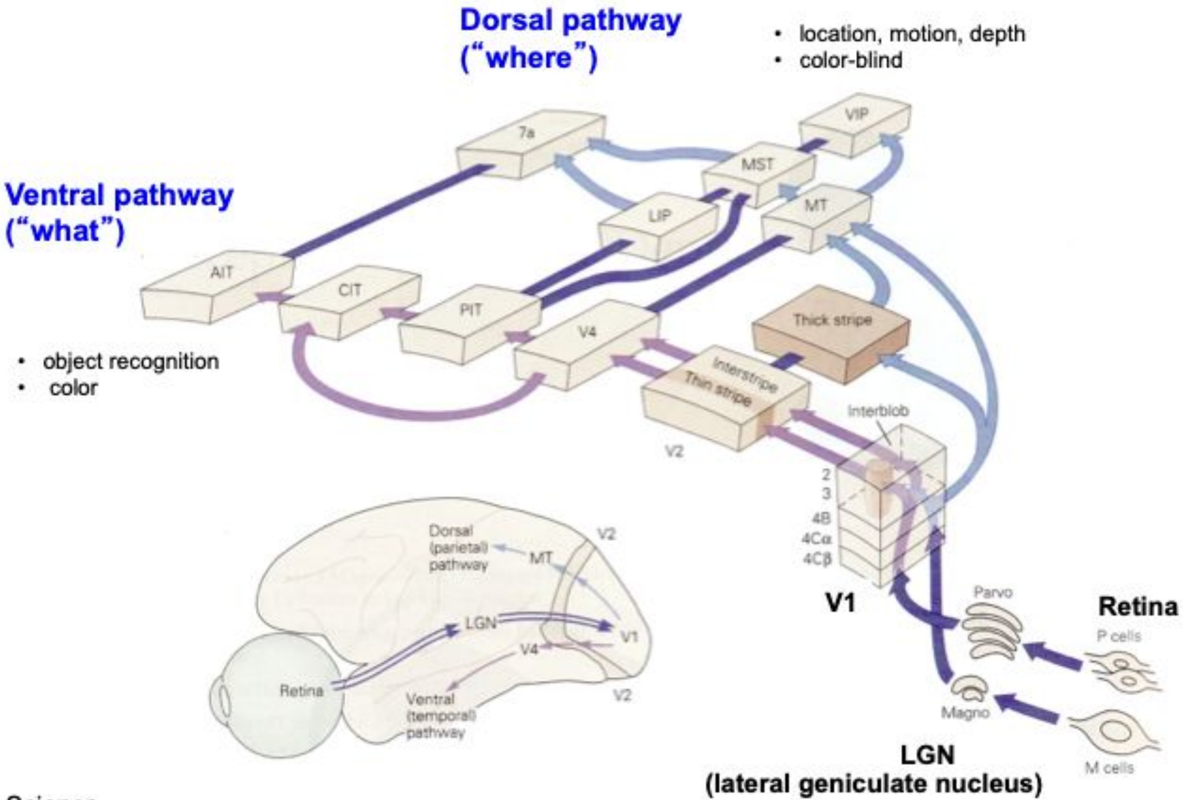


Visual System Neuroscience

Summary

- Center-surround model and it’s ethological relevance (motion)
- Anatomy of the visual pathway, and in particular, cortical layers
- The retina is an evolutionary old structure, and adapted to the niche of the animal
- Zebrafish are a great model systems neuroscience due to their small size and optical transparency
- Zebrafish display a wide diversity of visually driven behaviour such as the OMR, prey capture and predator avoidance
- The organisation of the zebrafish retina – 4 cones with UV cone being integral for detecting prey
- Retinal ganglion cells act as feature detectors, providing parallel processing streams to the brain
- The tectum has a highly organised structure and acts as a local motion detector classifying prey and predators
- The tectum uses this information to trigger approach and avoidance behavior

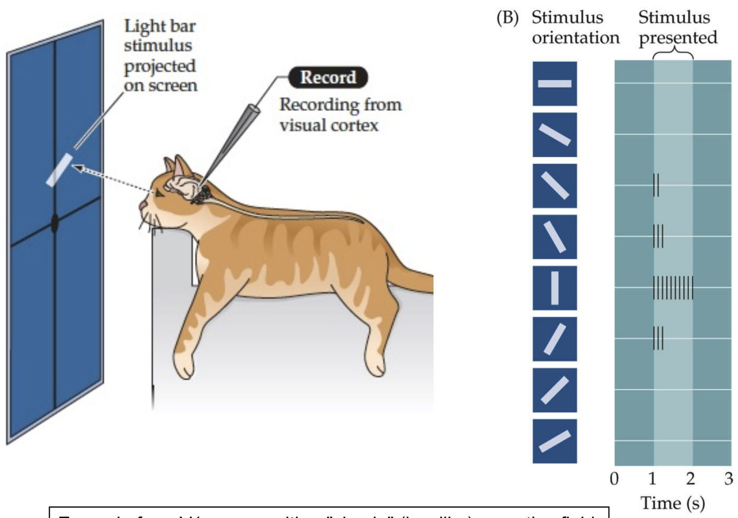
The Visual pathway



Principles of Neural Science

What are neurons in V1 encoding?

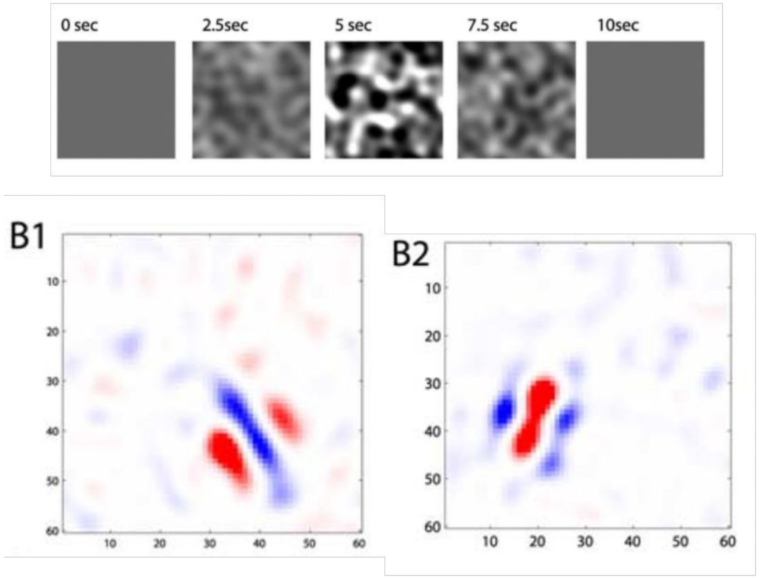
Orientation selectivity (Hubel & Wiesel)



Example for a V1 neuron with a "simple" (bar-like) receptive field

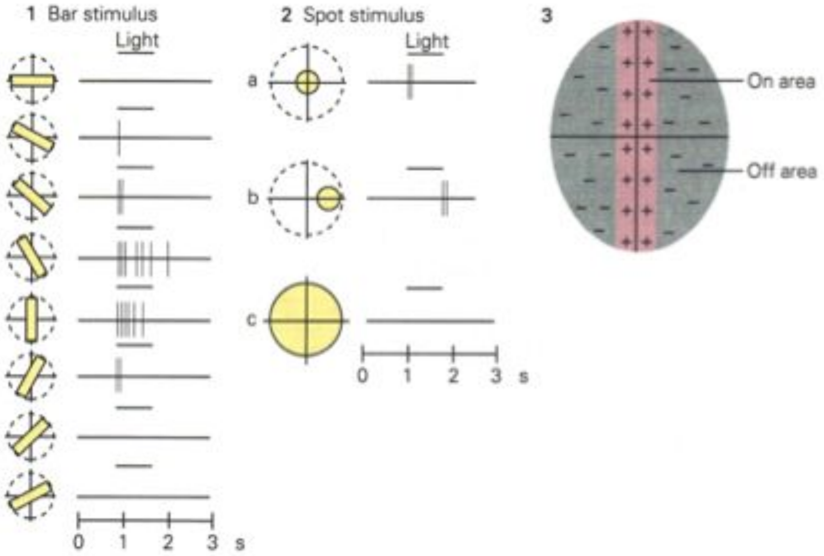
Purves Fig. 12.8

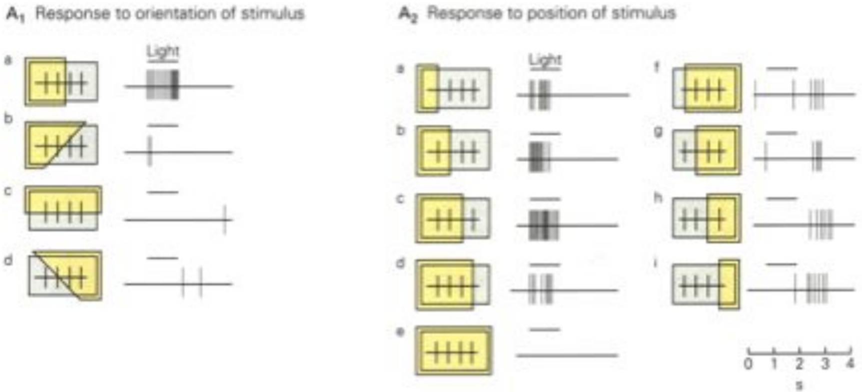
Spike triggered average



Niell and Stryker, 2008

V1 simple & complex cells

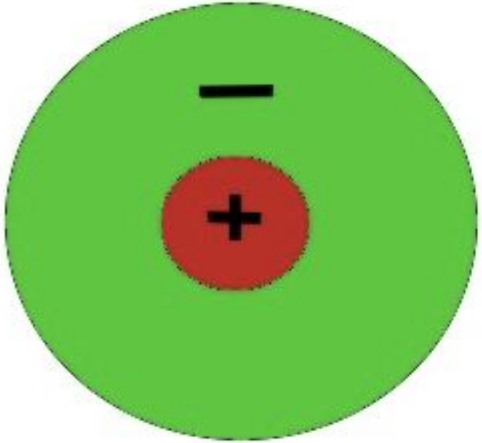




- Orientation specificity!

- Orientation specificity!
- Less sensitive to exact locations

The beginning of the visual system: retinal ganglion cells and the Center-Surround Model



Center-Surround Receptive Fields
The center-surround model is based on the organization of the RGCs' receptive fields, which are the specific areas of the retina where light stimuli can influence the firing rate of the cell. These receptive fields are structured in a center-surround arrangement, consisting of two distinct parts:

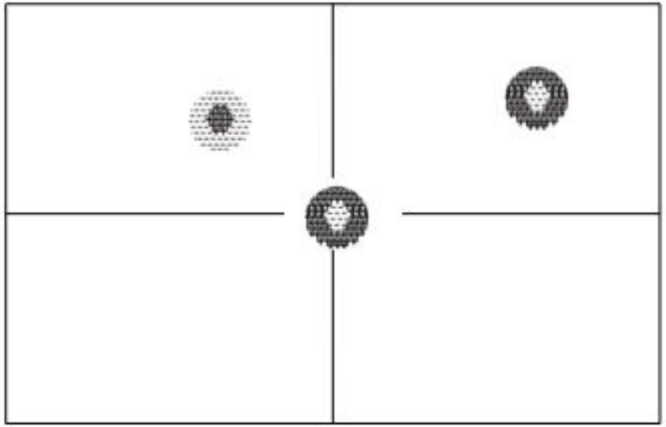
•**Center**: The central part of the receptive field can either be excitatory (increasing the cell's firing rate when stimulated by light) or inhibitory (decreasing the cell's firing rate when stimulated by light).

•**Surround**: The surrounding part of the receptive field has the opposite effect to the center. If the center is excitatory, the surround is inhibitory, and vice versa.

Image adapted from https://muse.union.edu/visualmotion/center-surround/

The beginning of the visual system: retinal ganglion cells and the Center-Surround Model

Field of view:



If you record from a retinal ganglion cell (RGC).
They fire APs with generally two types of responses:

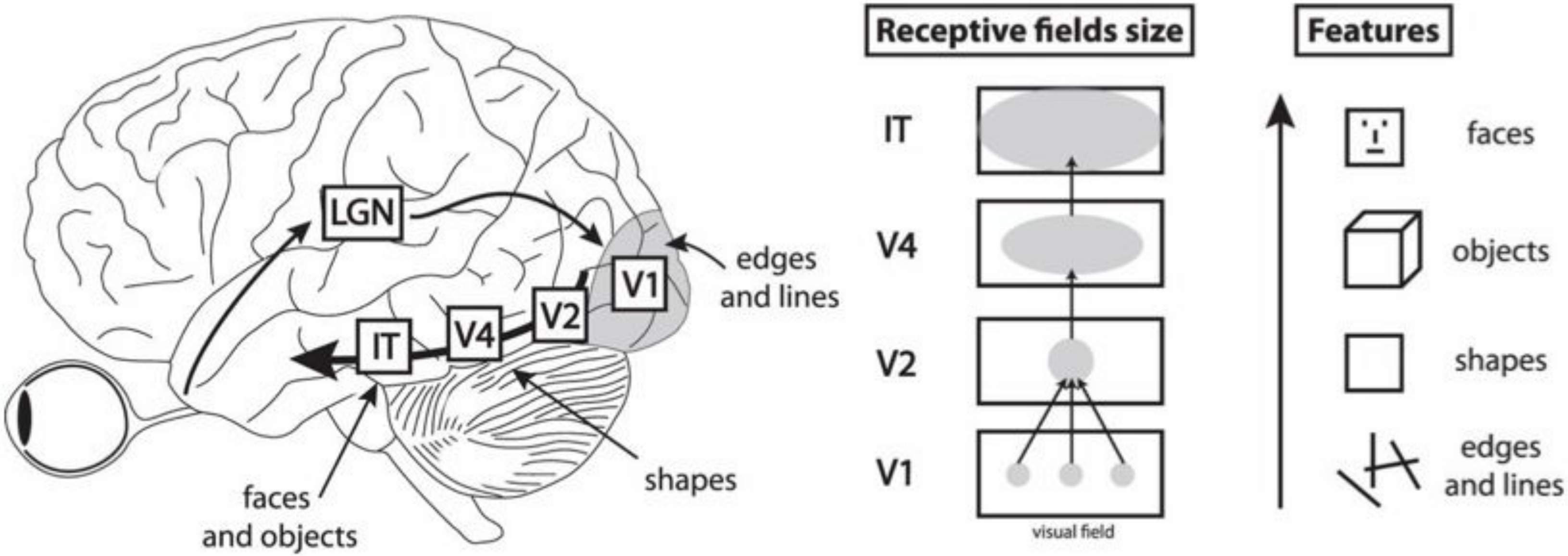
On-center ganglion cell

"ON-center":
↑ AP- frequency
in the center of the receptive field (RF)

Off-center ganglion cell

"OFF center"
↓ AP-frequency
in the center of the RF

Hierarchical visual processing

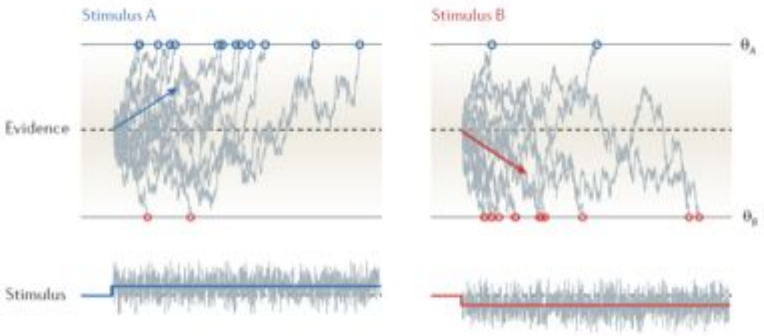


- V1 neurons are most sensitive to low-level features, such as edges and lines.
- In higher visual areas, like V4 and IT, receptive fields are larger, and neurons are sensitive to complex features, such as shapes and objects.
- Responses of high-level neurons are fully determined by the neural firing of lower-level neurons. For example, the neural firing to a square is determined by the neural firing for two vertical and two horizontal lines.

From: When crowding of crowding leads to uncrowding
Journal of Vision. 2013;13(13):10. doi:10.1167/13.13.10

Drift diffusion models: accumulating noisy evidence

- Variability in response times and judgments

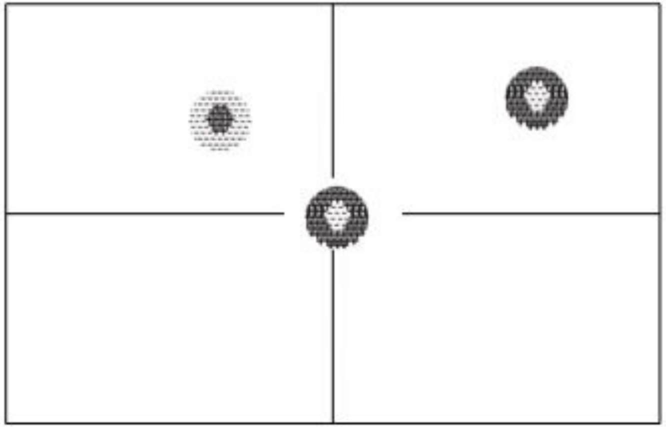


- Effect of difficulty on response times

(Uchida et al., 2006)

The beginning of the visual system: retinal ganglion cells and the Center-Surround Model

Field of view:



If you record from a retinal ganglion cell (RGC).
They fire APs with generally two types of responses:

On-center ganglion cell

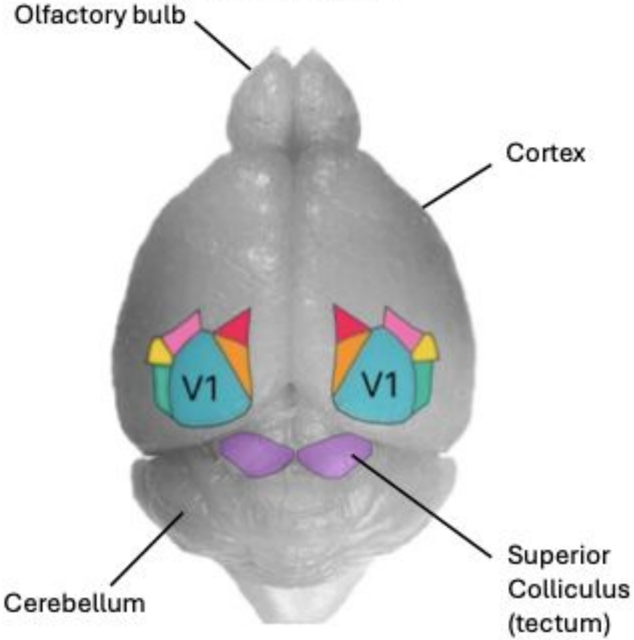
"ON-center":
↑ AP- frequency
in the center of the receptive field (RF)

Off-center ganglion cell

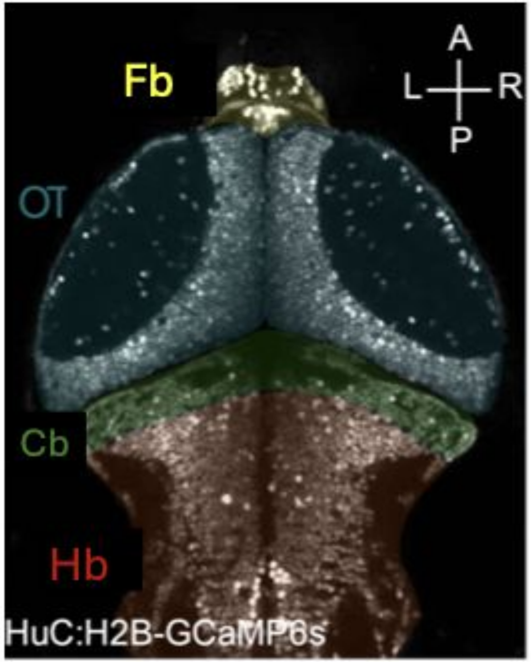
"OFF center"
↓ AP-frequency
in the center of the RF

Zebrafish general neuroanatomy

The mouse brain



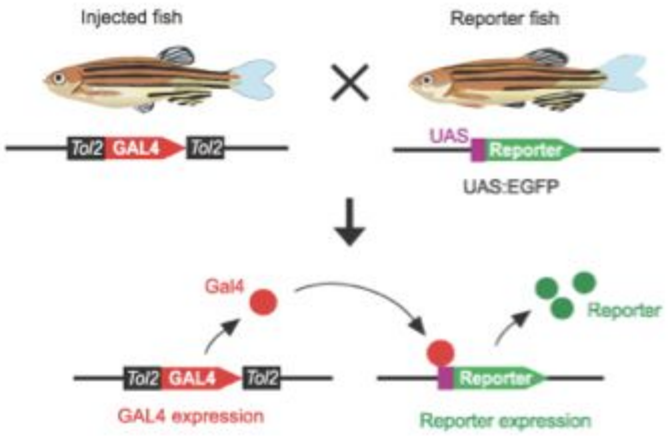
Zebrafish brain



Fb = forebrain
OT = optic tectum
Cb = cerebellum
Hb = Hind brain

- No cortex!
- Optic tectum is the main visual area and sits on the dorsal surface of the midbrain
- The optic tectum is homologous to the superior colliculus in mammals
- Contains a large neuropil region where neurons from the retina provide visual input.

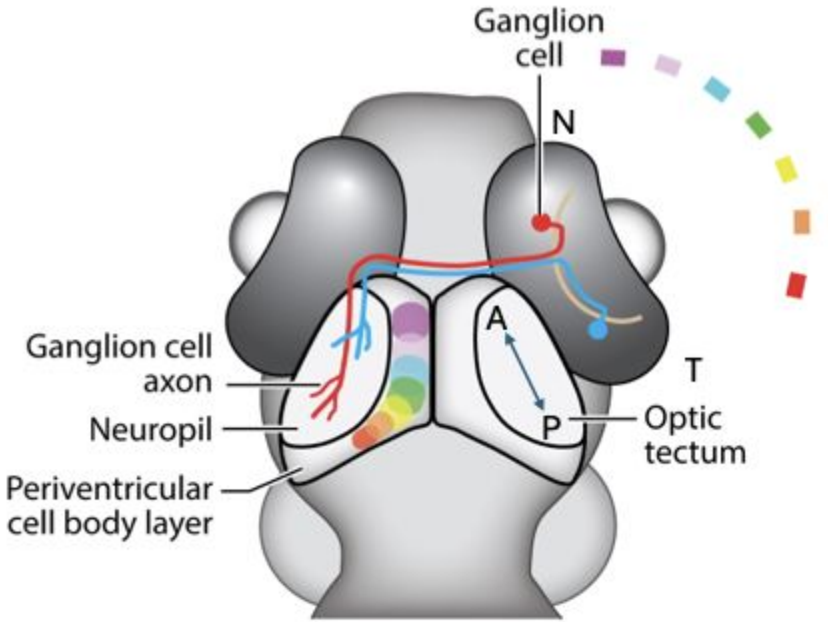
Key concept: the Gal4-UAS system



(Asakawa et al., 2008)

- The GAL4-UAS works in a similar way to the cre-recombinase system that is used in mice
- GAL4 is a transcription factor that binds to an upstream activator signal (UAS) causing transcription of the downstream reporter (such as GFP).
- If the Gal4-is placed downstream of a particular endogenous promotor then this can restrict expression to a single neuron subtype

Retinotopy: RGCs preserve a map of visual space in the tectum

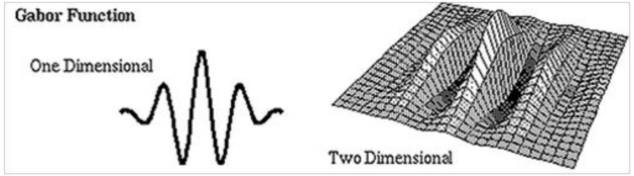


(Bollman 2019)

- This means that within the tectum there is a map of visual space (A retinotopic map)
- When neurons are active in a particular region of the tectum = possible locate the position of the stimulus in visual space

V1 RFs resemble Gabor filters and neural response is sparse

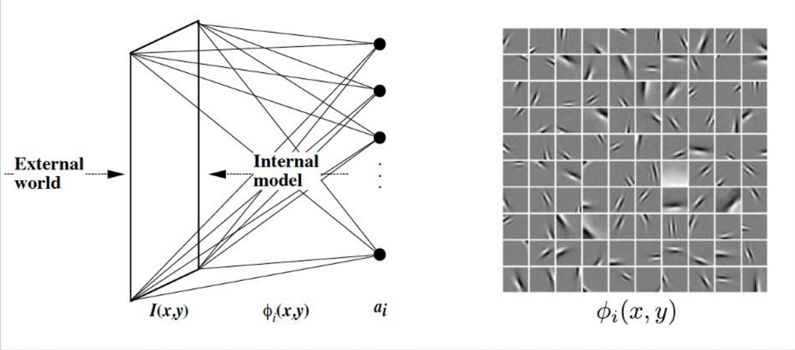
Gabor filters



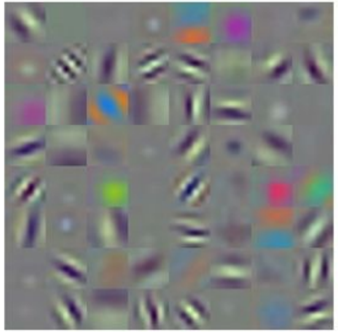
edge detector neurons can be explained with sparse autoencoding

$$I(x,y) = \sum_i a_i \, \phi_i(x,y) + \epsilon(x,y)$$

Adapted from A. Mathis

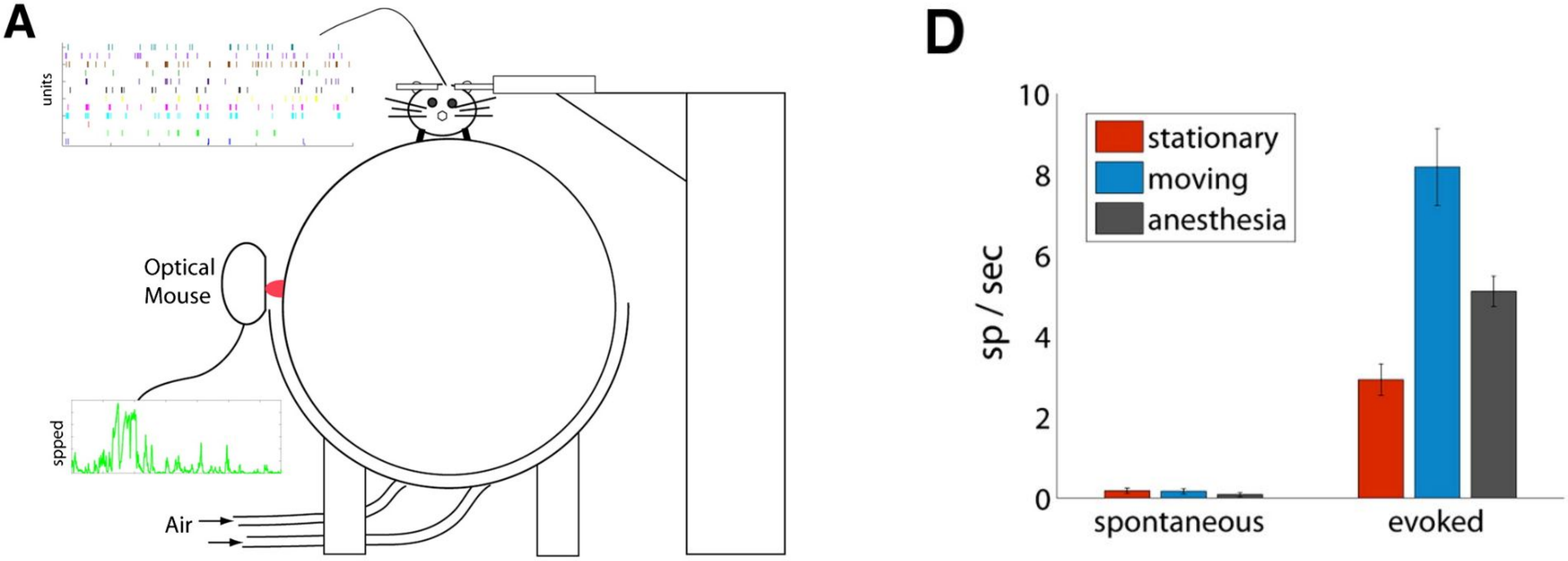


Olshausen & Field, 1996 Nature



Representations in ImageNet trained (CNN)

Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex



Niell & Stryker, 2010

Dunn et al. 2016
Paper round-up

- They provide the first detailed description of a rapid escape behavior elicited by a visual stimulus in freely swimming larval zebrafish.
- They suggest that the circuits processing looming stimuli may primarily use stimulus size information when determining when and if an escape should be initiated.
- They show that the optic tectum (OT) might serve as a primary nucleus involved in looming detection within the larval zebrafish brain, by encoding a critical looming visual angle as an ensemble.
- They establish a necessary role of the M-system in the sensorimotor transformation from looming stimuli to escape behavior, providing a functional scaffold for the zebrafish to quickly evade threats identified with their eyes alone.

Neural analysis

Summary

- **Neural encoding** and **neural decoding** are fundamental descriptions of neural (coding) processing and data analysis.
- A fundamental goal is: how much information does **K** have about **x**
- We mathematically model this as P(K|x), where the neural response of population **K** to a stimulus (or event) **x**. **K** is a vector representing the activity of *N* neurons, and each entry represents, e.g., the number of spikes in some time bin or the rate response of that particular neuron.
- **Generalized Linear Models** (GLMs) are very attractive for both individual neurons and populations, yet assume **linear** $\boldsymbol{\theta}$ dynamics (careful: despite having a nonlinear parameter).
- Modern hardware advances continue to push the upper limit on the # of neurons we can record, and therefore we need new mathematical tools for understanding neural coding.
- Manifold of behavioral and neural data hypothesis comes into play…
- Two large classes of approaching modeling a system: data-driven or hypothesis (task)--driven
- Modern methods for mapping the statistical properties of neurons to a stimulus/behavior are fully-observable models and latent variable models.
- Latent variable models infer hidden (i.e., latent) variables that capture the underlying structure of the observed data through a joint probability distribution.
- VAEs and contrastive learning approach to neural analysis; contrastive learning (CEBRA) has highly attractive properties like combining across datasets and producing consistent latent embeddings.

EPFL

What information is our brain trying to encode & decode?

Stimulus

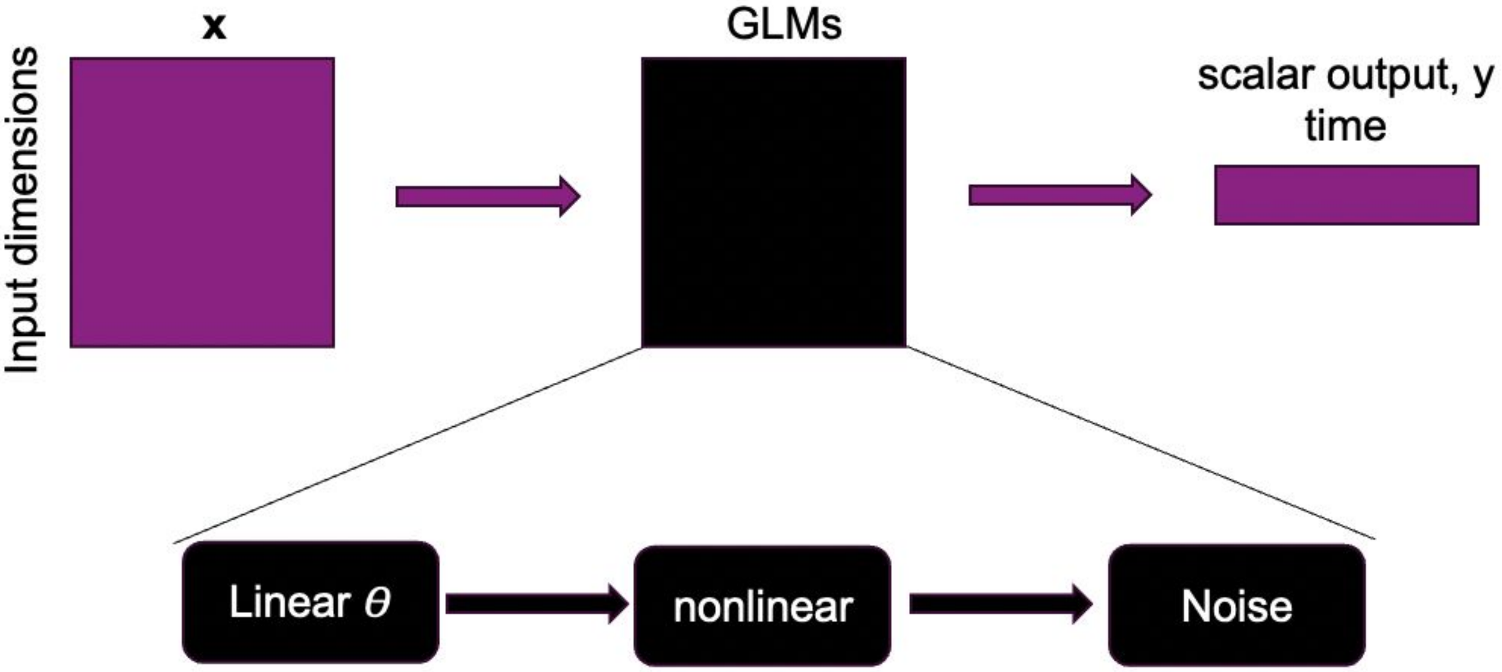
Spikes

• How do neurons (**K**) respond to a certain stimulus (**x**)?

• Our brain needs to determine what is going on in the real world from patterns of spikes.

- We mathematically model this as P(**K|x**), where the neural response of population **K** to a stimulus (or event) **x**.
- **K** is a vector representing the activity of *N* neurons, and each entry represents, e.g., the number of spikes in some time bin or the rate response of that particular neuron.

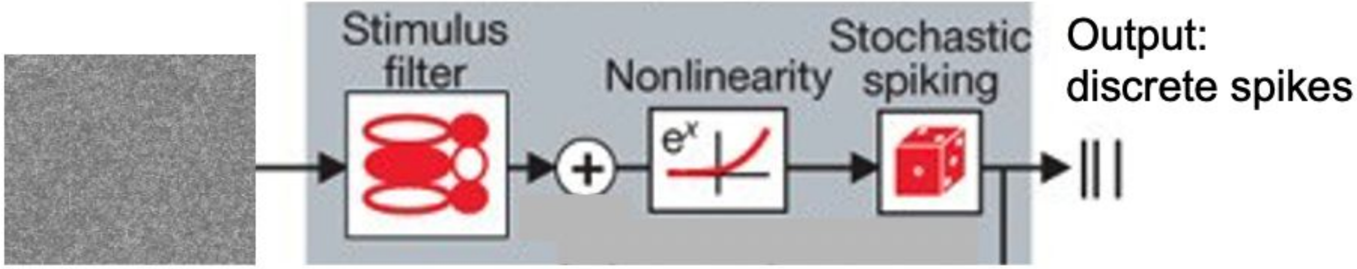
Generalized Linear Models



Poisson GLM

θ

$$P(y_t = y) = \frac{\lambda^y e^{-\lambda}}{y!}$$



- **In the model, each neuron's input is described by a set of linear filters:**
  - a stimulus filter, or spatial receptive field (θ)

$$p_t = exp\left(\sum_i \theta_i \; x_{t-i}\right)$$

Details: Poisson GLM

Poisson Distribution: Single Event

- Probability of events y_t at time *t*
- Formula:

$$P(y_t = y) = \frac{\lambda^y e^{-\lambda}}{y!}$$

- λ: rate parameter (average number of events)
- *y*!: factorial of *y* (number of events)

Link Function and Predictors

- λ linked to predictors x_t
- Canonical link function: natural log
- Formula:

$$\lambda = \exp(\theta^T x_t)$$

- θ : model parameters (as a vector)
- x_t : predictors vector

Likelihood: All Data Points

- Joint probability as the product of individual probabilities
- Formula

$$P(y_1 : T) = \prod_t P(y_t)$$

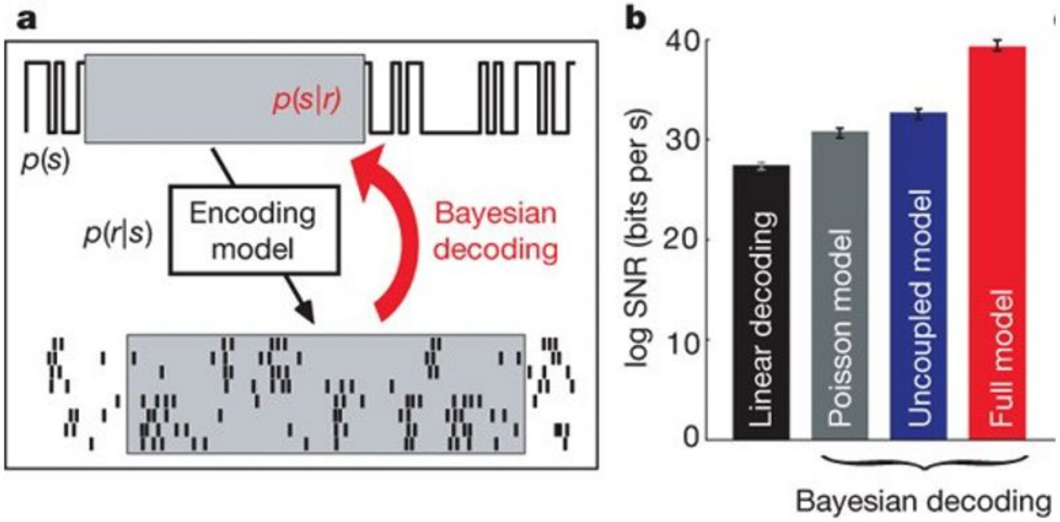
- Assumes independence between data points

Simplify!!

Log Likelihood

$$\log \mathcal{L} = \sum_t \log P(y_t)$$

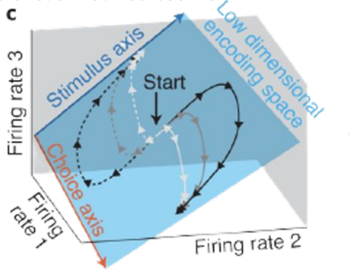
GLMs in action: Pillow et al. 2008



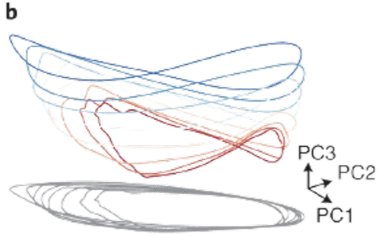
GLMs with coupling filters were shown to capture 40% more visual information from the retina than optimal linear decoding, indicating that GLMs can model additional details in the activity that are relevant for representing the stimulus!

Population analysis can reveal core principles of neural coding

Urai et al Nat Neurosci 2022



Behaviorally relevant neural variance within a small number of dimensions

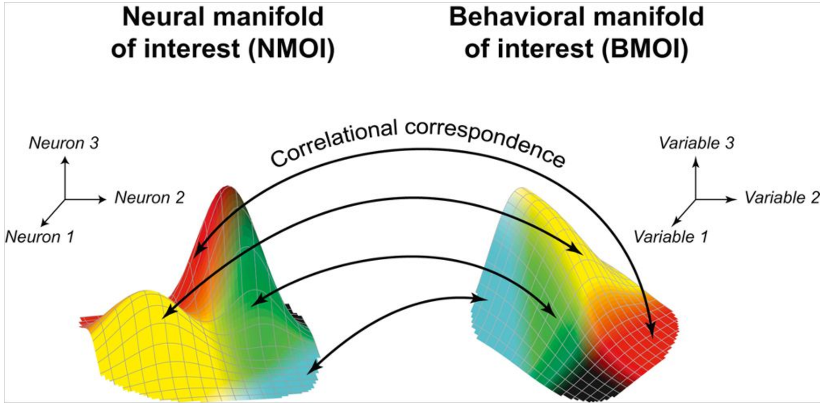


Neural computations at population dynamics but not in single-neuron firing rates

Intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep

Chaudhuri, R. et al. Nat Neurosci 2019

Manifolds for measuring neural trajectories



Mehrdad Jazayeri and Arash Afraz Neuron 2017

How can we (consistently) extract the behaviorally-relevant latent dimensions from neural population activity?

Nonlinear embeddings via linear dynamical system (LDS)

Dynamics of *n* neurons are modulated by LDS w/ *m*-dim latent state (z) that evolves:

$$\mathbf{z}_{r1} \sim \mathcal{N}(\mu_1, \mathbf{Q}_1)$$
$$\mathbf{z}_{r(t+1)}|\mathbf{z}_{rt} \sim \mathcal{N}(\mathbf{A}\mathbf{z}_{rt}, \mathbf{Q}),$$

A = linear dynamics matrix (m x m)
Q_1 = covariance of initial states
Q = Gaussian noise

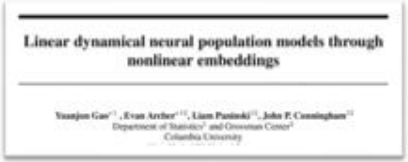
Observation model:

$$x_{rti}|\mathbf{z}_{rt} \sim \mathcal{P}_\lambda\left(\lambda_{rti} = [f(\mathbf{z}_{rt})]_i\right).$$

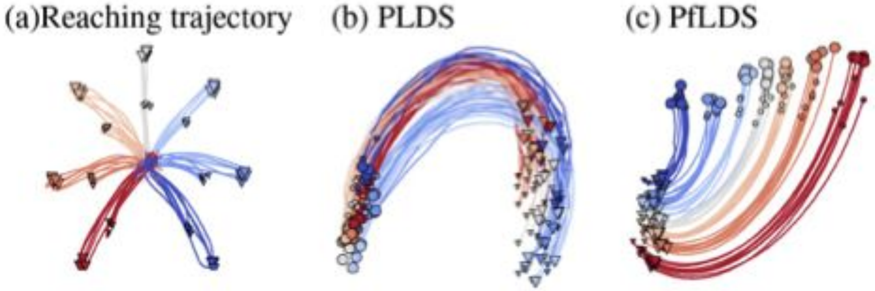
fLDS: exchange observation model such that each neuron as a separate nonlinear dep. on latent variable:

$$x_{rti}|\mathbf{z}_{rt} \sim \mathcal{P}_\lambda\left(\lambda_{rti} = [f_\psi(\mathbf{z}_{rt})]_i\right),$$

where $[f(\mathbf{z}_{rt})]_i$ is the i^{th} element of a deterministic "rate" function $f(\mathbf{z}_{rt}) : \mathbb{R}^m \to \mathbb{R}^n$, and $\mathcal{P}_\lambda(\lambda)$ is a noise model with parameter λ.


Linear dynamical neural population models through nonlinear embeddings

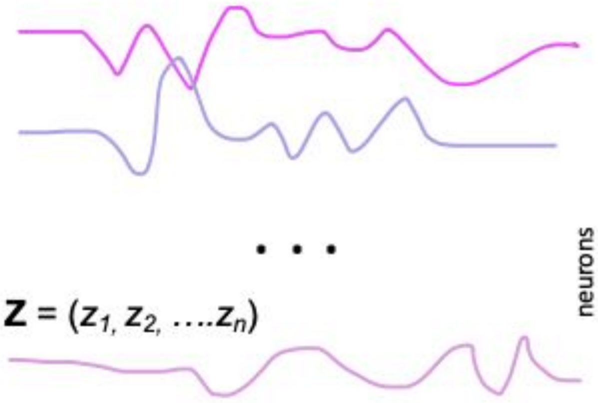
Yuanjun Gao, Evan Archer, Liam Paninski, John P. Cunningham
Department of Statistics and Grossman Center
Columbia University



Identifiable non-linear ICA: the problem setting....

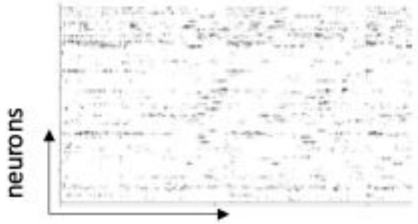
Latent (hidden) underlying brain-state factors (z)

Observable neural data



Z = $(z_1, z_2, \ldots. z_n)$

X = $(x_1, x_2, \ldots. x_n)$

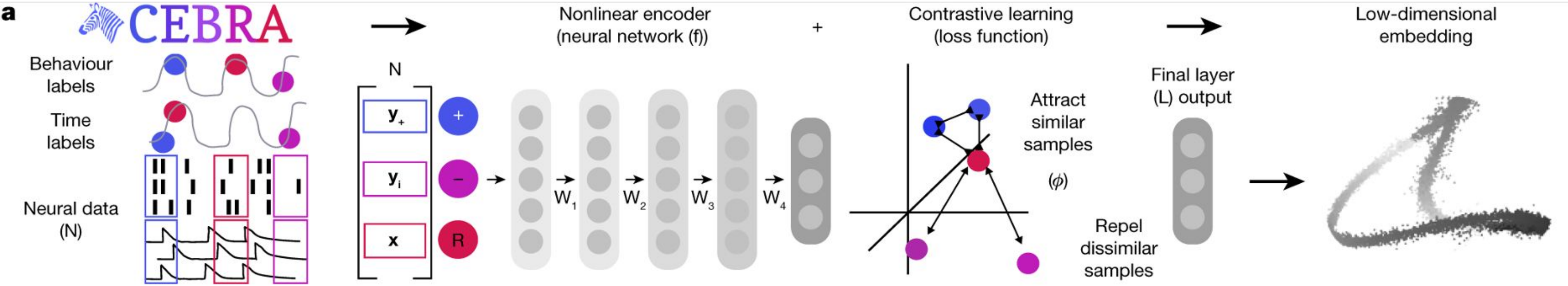


$$\mathbf{z} \xrightarrow{\text{Mixing function}} \mathbf{x} = \mathbf{g}(z)$$

- Non-linear ICA attempts to find non-linear components such that they correspond to a well-defined generative model (Hyvärinen et al., 2001; Jutten et al., 2010).
- The aim is to recover the inverse function **g** as well as the independent components **z** based on observations of x alone.

▪ Identifiable: "the sources can be separated"- Aapo Hyvärinen

CEBRA : an algorithm for joint modeling of auxiliary & times series data



$$\underset{\substack{\mathbf{x}\sim p(\mathbf{x})\ \mathbf{y}_+\sim p(\mathbf{y}|\ \mathbf{x}) \\ \mathbf{y}_1,\ldots,\mathbf{y}_n\sim q(\mathbf{y}|\mathbf{x})}}{\mathbb{E}}\left[-\psi(\mathbf{x},\mathbf{y}_+) + \log\sum_{i=1}^{n} e^{\psi(\mathbf{x},\mathbf{y}_i)}\right]$$

Schneider, Lee, Mathis Nature 2023

Behavioral analysis

Deep learning in the laboratory: leveraging transfer learning

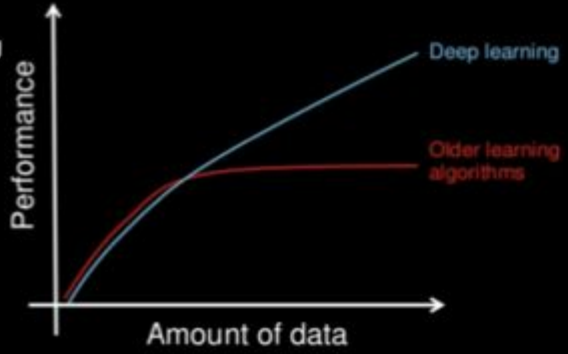
DeepPose
DeeperCut
OpenPose
Conv. PoseMachines
...
HRNet

deep neural networks

image → Predictor → pose

train **A lot of labeled images (>10^6 joints!)**

Andrew Ng



DATA hungry algorithms… how to bring this to the lab?

Transfer Learning: take a trained network and ask it to learn a new task

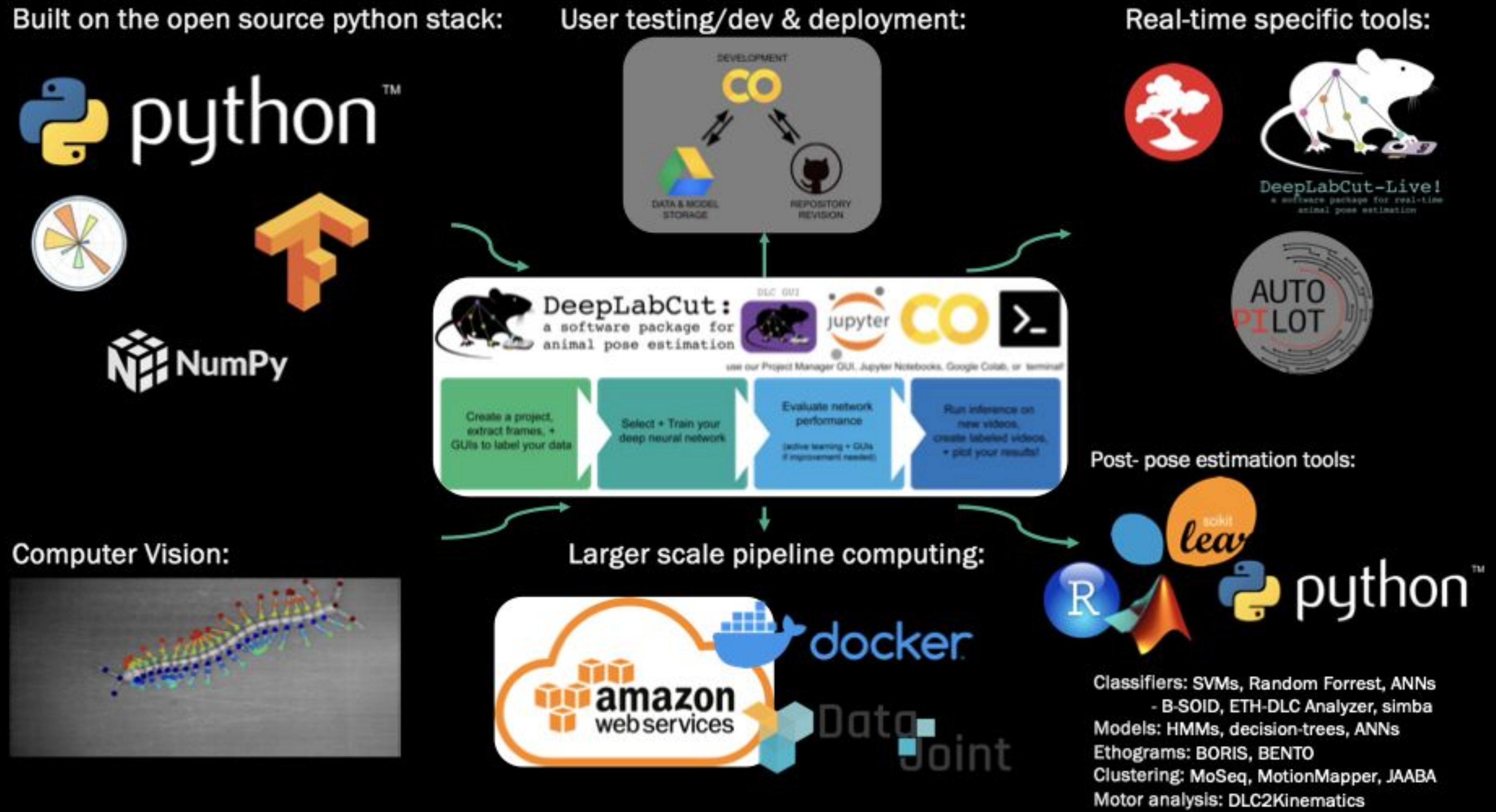
...

ConvNets (such as ResNet-50, etc)

cat

IMAGENET

Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) **ImageNet Large Scale Visual Recognition Challenge**. *International Journal of Computer Vision*, 2015.



Built on the open source python stack:
python™
NumPy
User testing/dev & deployment:
DEVELOPMENT
CO
DATA & MODEL STORAGE
REPOSITORY REVISION
Real-time specific tools:
DeepLabCut-Live!
a software package for real-time animal pose estimation
AUTO PILOT
DeepLabCut:
a software package for animal pose estimation
DLC GUI
jupyter
CO
use our Project Manager GUI, Jupyter Notebooks, Google Colab, or terminal!
Create a project, extract frames, + GUIs to label your data
Select + Train your deep neural network
Evaluate network performance
(active learning + GUIs if improvement needed)
Run inference on new videos, create labeled videos, + plot your results!
Computer Vision:
Larger scale pipeline computing:
amazon web services
docker
DataJoint
Post- pose estimation tools:
scikit learn
R
python™
Classifiers: SVMs, Random Forrest, ANNs
- B-SOID, ETH-DLC Analyzer, simba
Models: HMMs, decision-trees, ANNs
Ethograms: BORIS, BENTO
Clustering: MoSeq, MotionMapper, JAABA
Motor analysis: DLC2Kinematics

Advances:

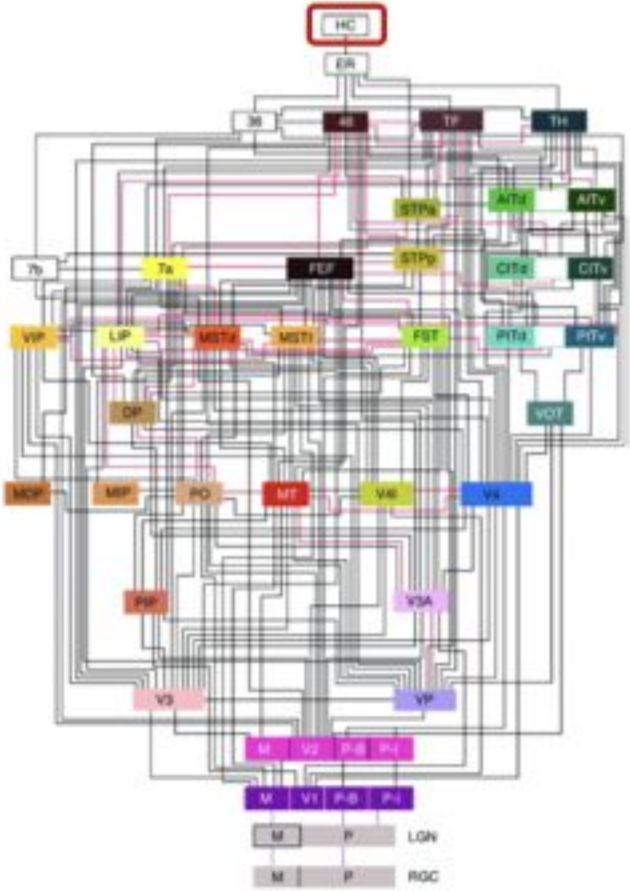
- Zero training from scratch could be required (huge energy savings & time/compute!)
- Zero-shot inference, or only tens of images for rapid fine-tuning required
- *(networks: gradient masking, memory replay, semantic mapping)*
- Zero-shot video inference, or 1.3x video inference w/test time aug.
- Tops OOD pose benchmarks

Still (more) challenges:

- TopView rodents & quadrupeds are not all animals in neuroscience
- Do we build centralized models, or groups build their own SuperAnimals?
- good data sharing practices // central resources?
- Is this really foundational?

Encoding of space in the brain

Recap - The Hippocampus



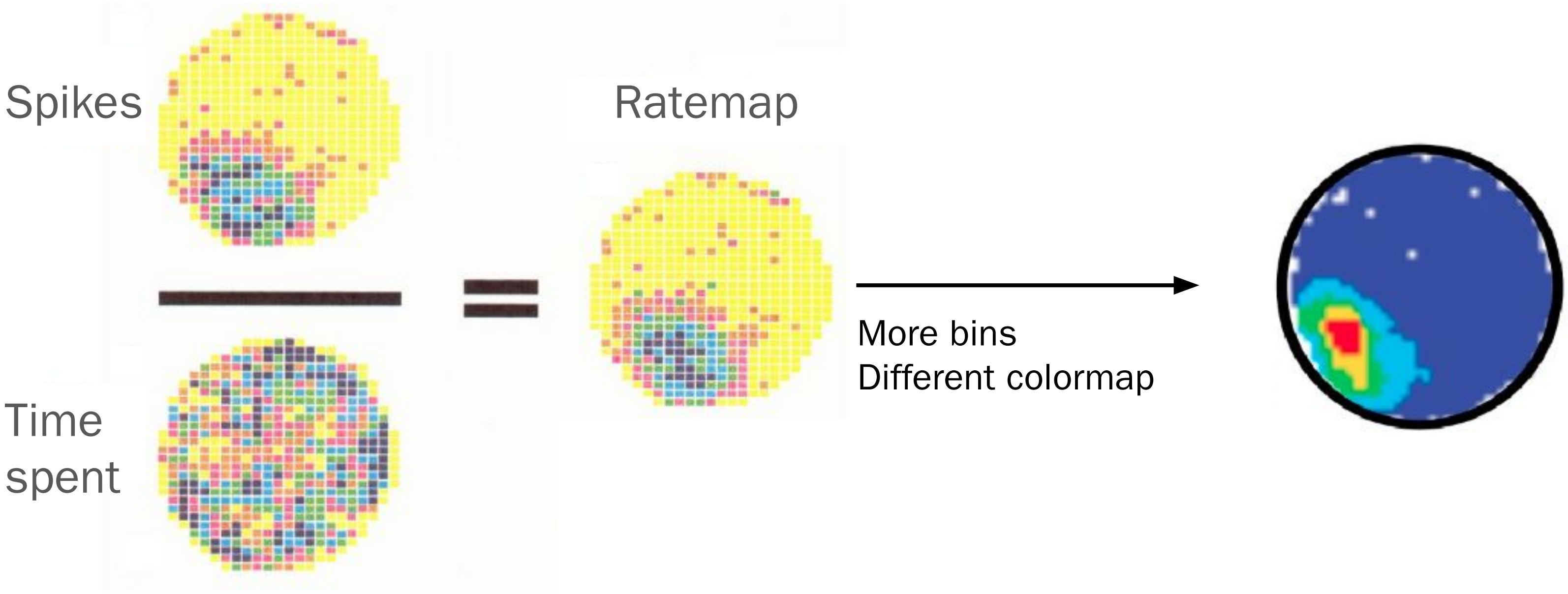
Huge amount of visual processing until any external sensory information reaches the hippocampus

In other senses (auditory, somatosensory) there is similarly complex processing upstream of the hippocampus – except olfactory inputs that reach the hippocampus much more directly (olfactory bulb -> entorhinal cortex)

Such high-level brain areas are expected to be **notoriously difficult to understand**: Presumably, responses must be extremely complex?

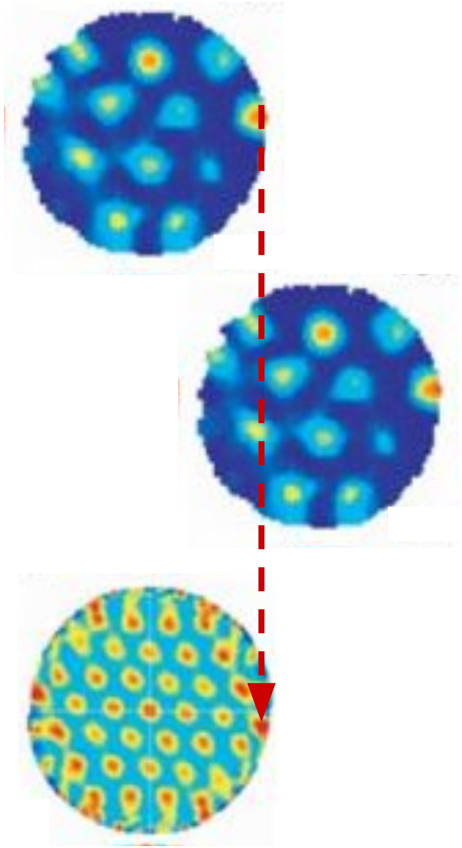
Fellemann & Van Essen 1991

Ratemaps



Autocorrelogram

Correlate ratemap to a shifted version of itself and then visualize the correlation coefficient

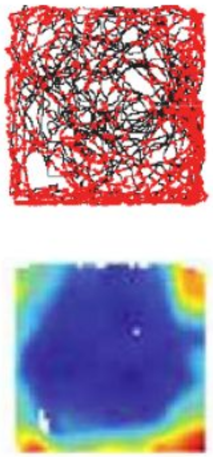


The cognitive map

Medial entorhinal cortex

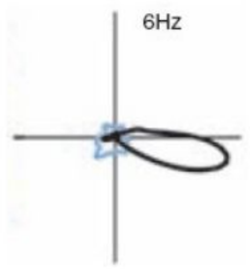
Hippocampus

Border cells



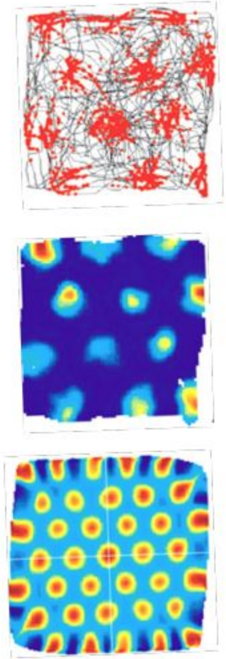
Mosers, O'Keefe, Knierim 2008

Head-direction cells



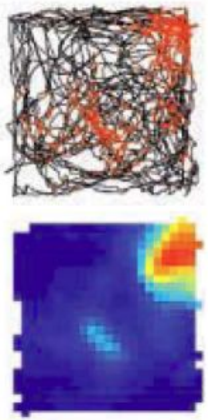
Ranck, Taube 1980s

Grid cells



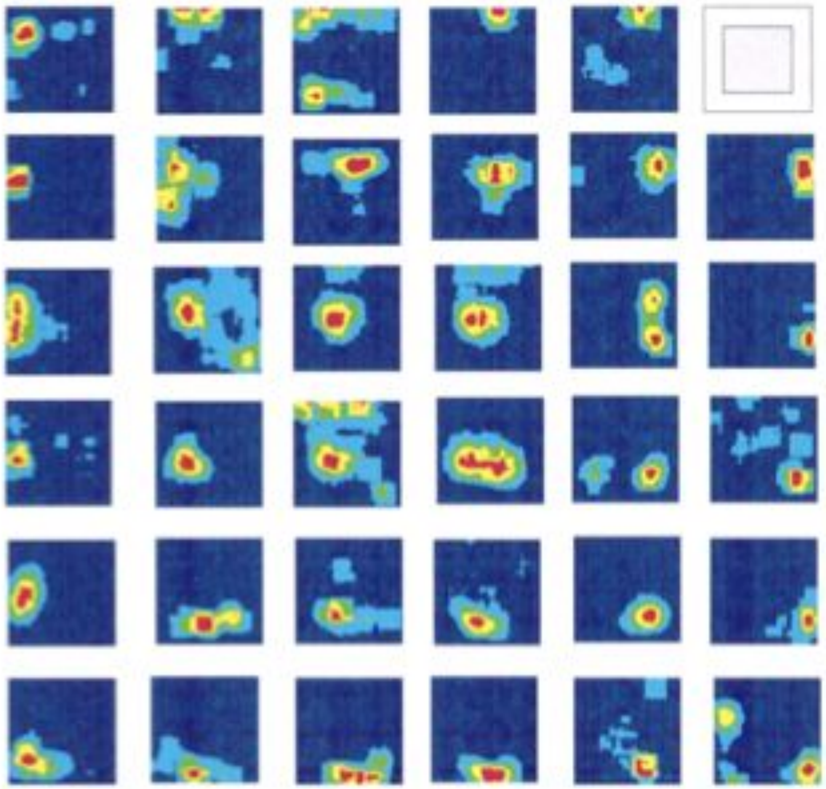
Mosers 2005

Place cells



O'Keefe 1971

Place cells in hippocampal subfield CA1



Many place cells together **tile the whole environment**

They provide a map of the environment, in the sense that the combination of currently active cells is sufficient to read out **precisely where the animal is** in the environment

Physical space is encoded in reference to the world (allocentric) - it is **fixed with respect to a point in the outside world**

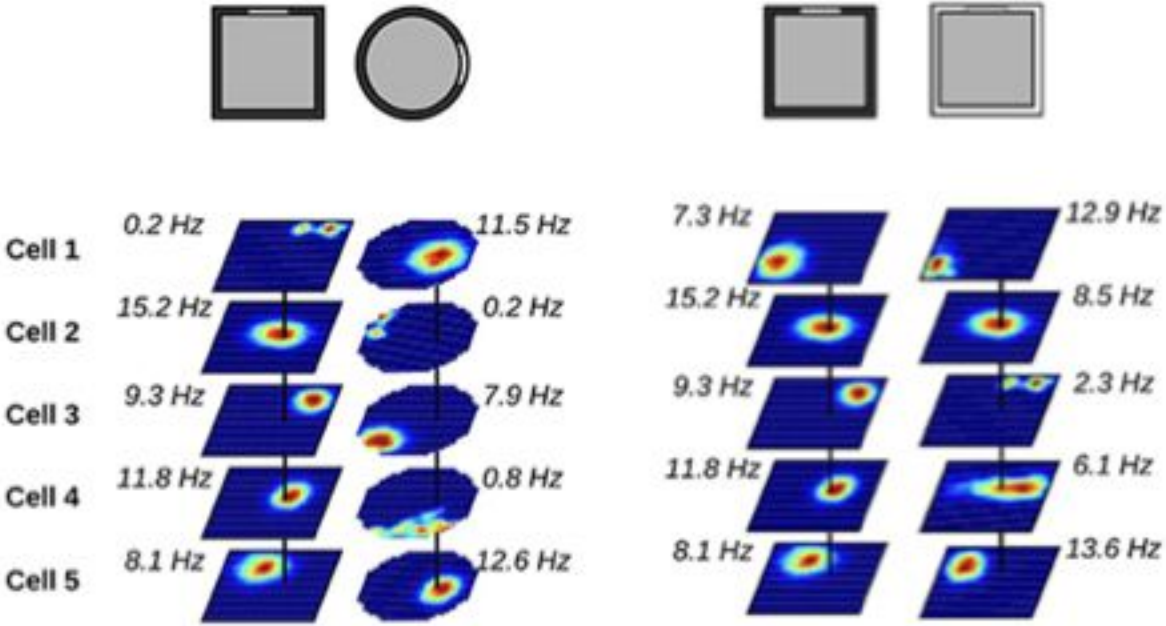
O'Keefe et al. 1999

Place cell remap in novel contexts

Place cell locations remap when context changes drastically (*global remapping*)

Smaller context changes are encoded as changes in firing rate (*rate remapping*)

Allows place cells to **encode multiple spaces and adapt to new environments**



Global remapping

Rate remapping

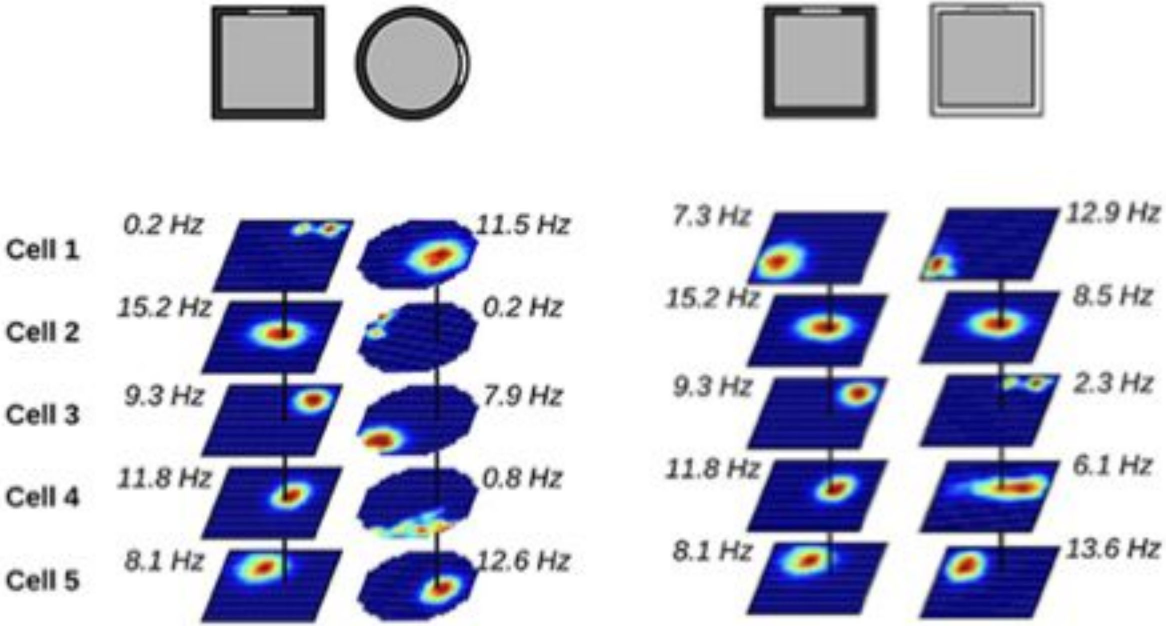
Latuske et al. 2018

Place cells remap in novel contexts

Place cell locations remap when context changes drastically (*global remapping*)

Smaller context changes are encoded as changes in firing rate (*rate remapping*)

Allows place cells to **encode multiple spaces and adapt to new environments**

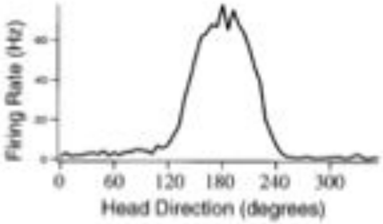


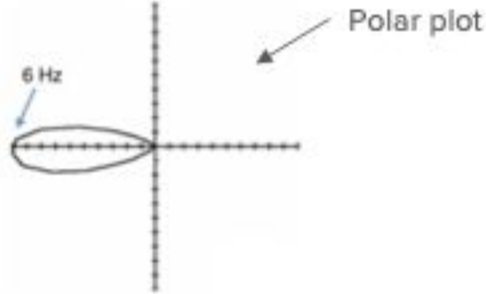
Global remapping

Rate remapping

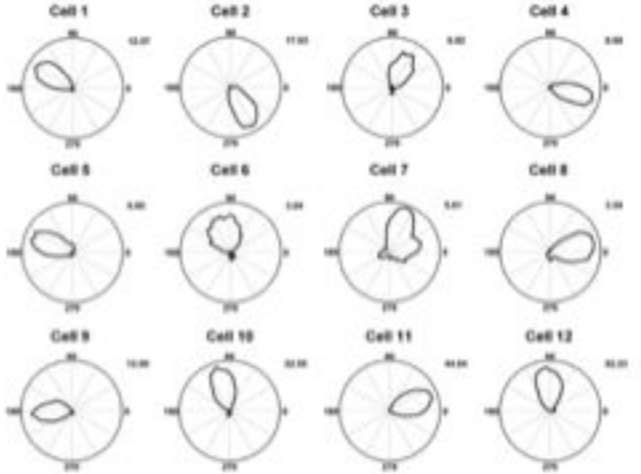
Latuske et al. 2018

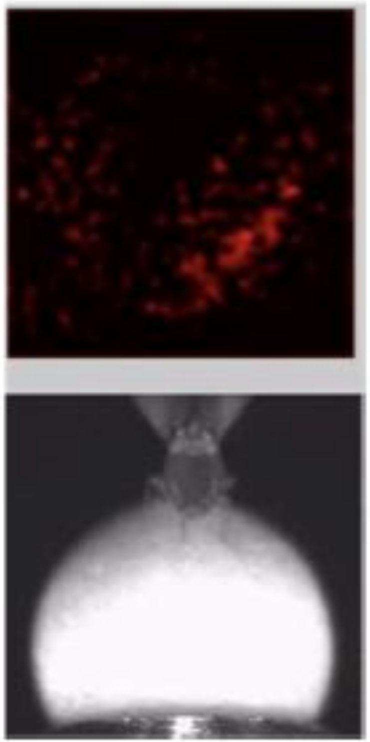
Head direction cells





A given population of head direction cells encodes the full 360 degrees



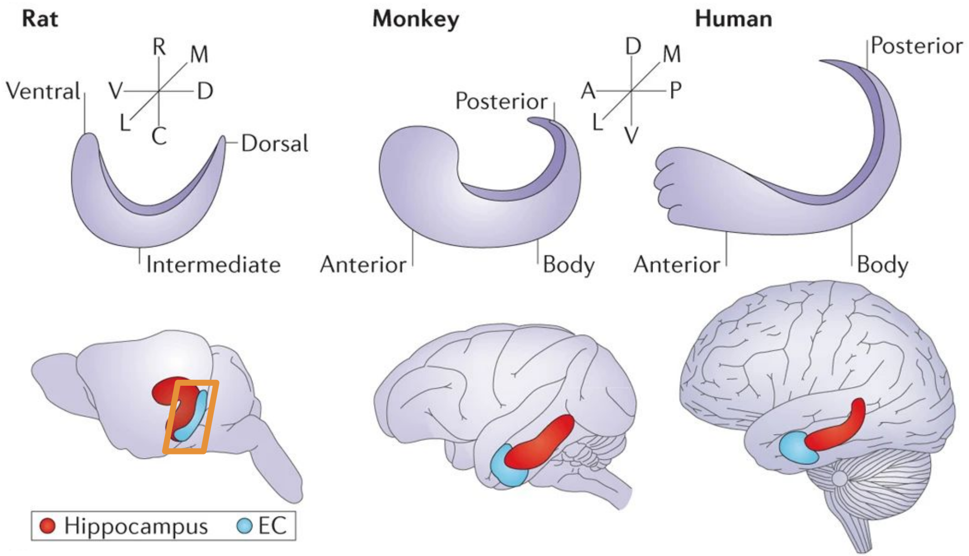


In drosophila the head direction circuit is arranged in **a topographical ring** with nearby cells encoding nearby angles

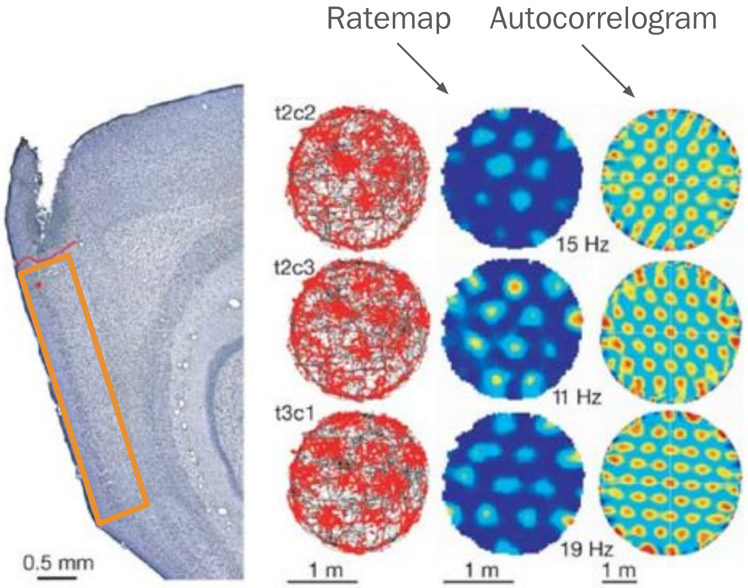
Head direction cells are mostly driven by vestibular input and visual landmarks

Taube et al. 1990
Preston-Ferrer et al. 2016

Grid cells in medial entorhinal cortex



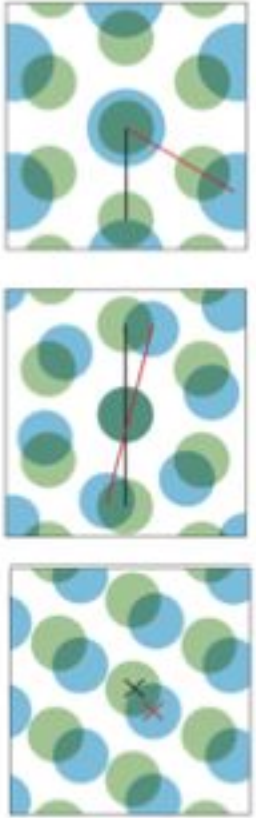
Strange et al. 2014



Entorhinal cortex (EC) is a major form of input to the hippocampal formation and is further split into medial (MEC) and lateral (LEC) entorhinal cortex

Animal trajectory is visualized in grey and the spikes are overlayed in red

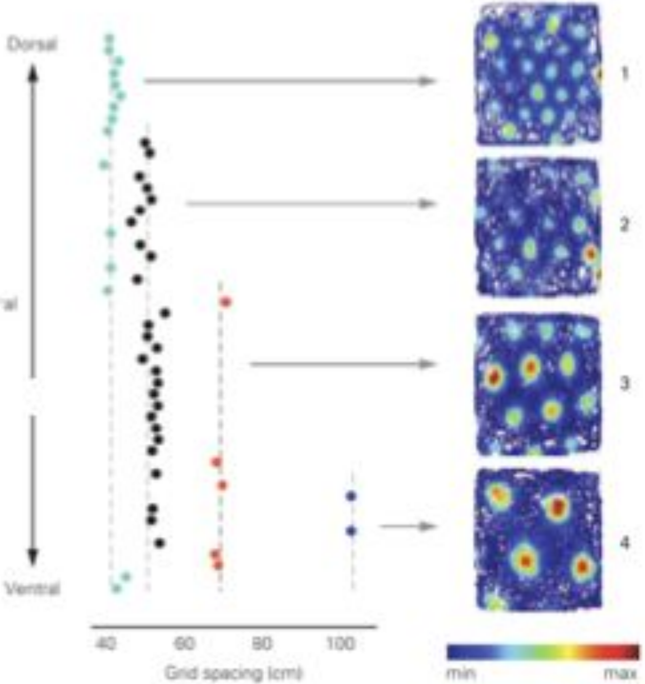
Grid cells form modules along the dorsoventral axis of EC

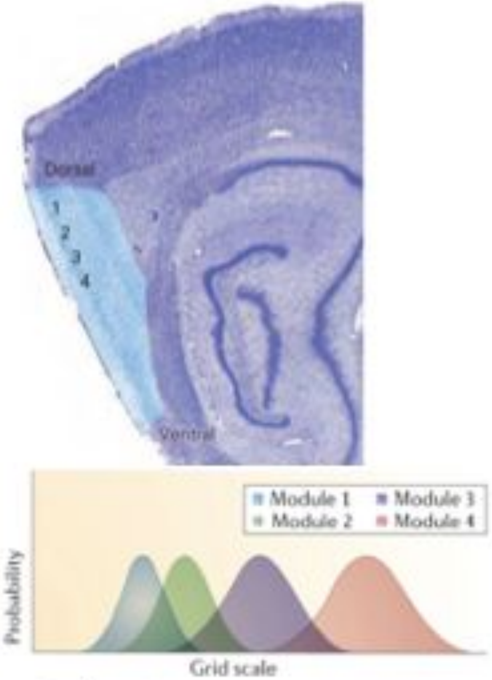


Scale / Spacing

Orientation

Phase





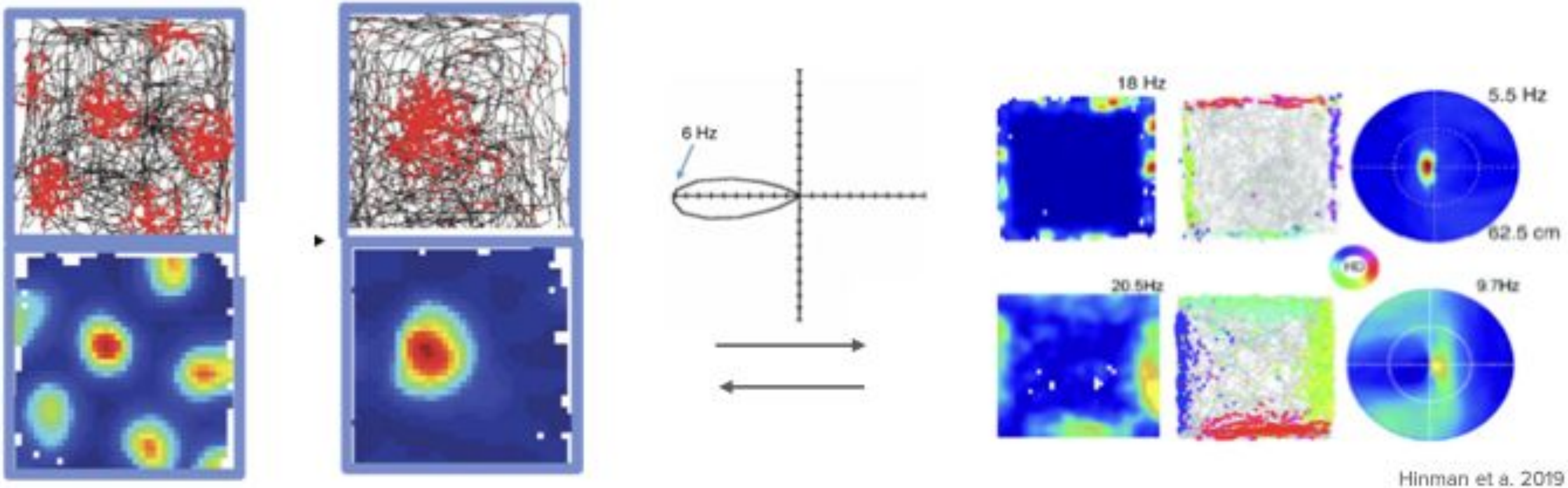
Scale of grid cells **increases topographically from dorsal to the ventral** part (~30cm dorsal to several metres ventral)

The expansion is **not linear but step-like**, suggesting that the grid-cell network is modular.

PoNS 6th Ed. 54–13
Moser et al. 2015

Allocentric and egocentric coding

Allocentric encoding cells | **Egocentric encoding cells**

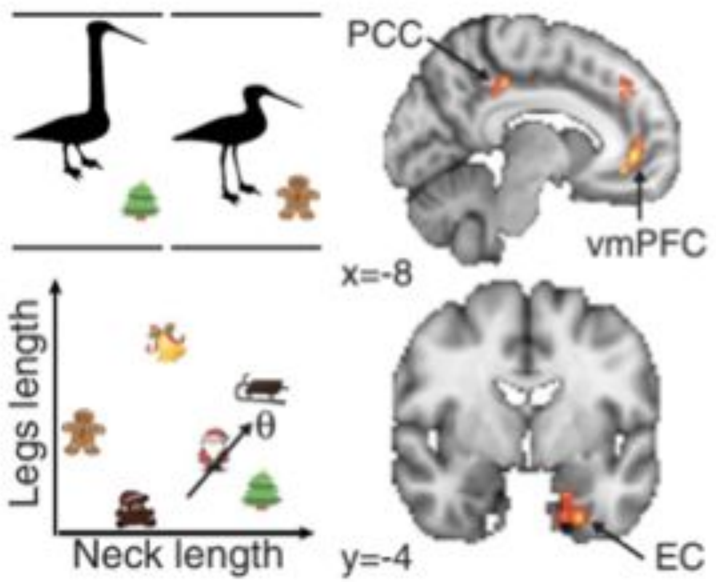


Hinman et a. 2019

The transformation between egocentric and allocentric coding cells is governed by head direction cells which anchor egocentric coding cells to the world

sensory information (egocentric) is processed and transformed into a stable, map-like (allocentric) representation

Encoding of abstract spaces



Participants in an fMRI scanner learned **association between objects and birds** (with variable neck and leg length) -> a novel abstract 2D space

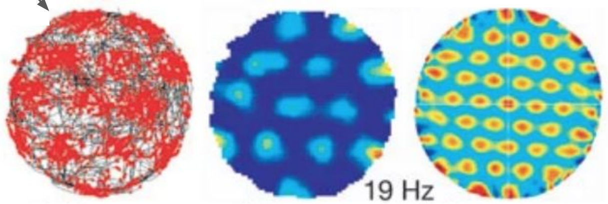
During trajectories through that space **grid-like activity** in entorhinal cortex can be observed

This shows that the cells underlying **physical space also encode abstract space**

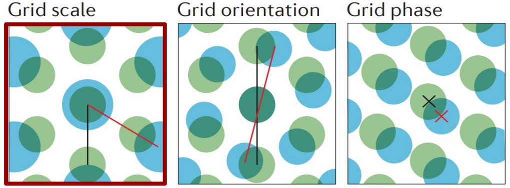
Constantinescu et al. 2016
Bellmund et al. 2018

Summary

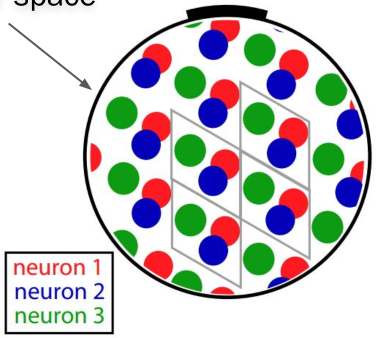
Periodic firing



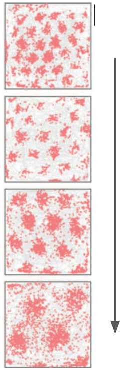
Grid cell parameters



Tiling of space

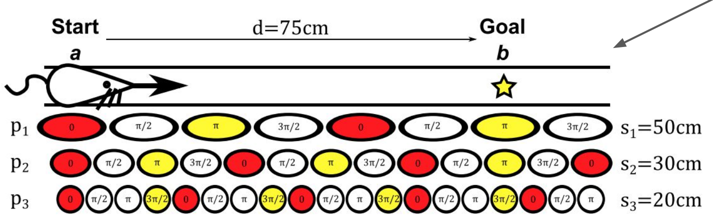


Dorsal

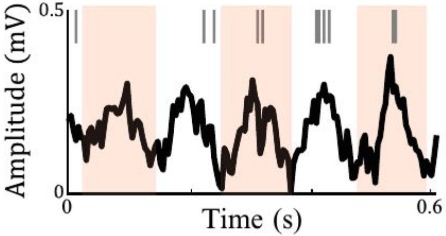


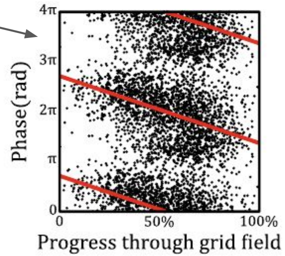
Ventral

Vector computation

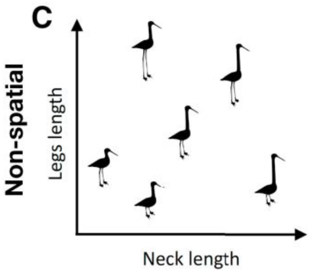


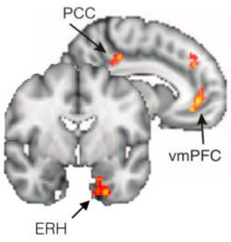
Phase precession





Encoding of abstract spaces!





Hafting et al. 2005
Fiete et al. 2008
Bush et al. 2015
Stemmler et al. 2015
Behrens et al. 2018

Summary

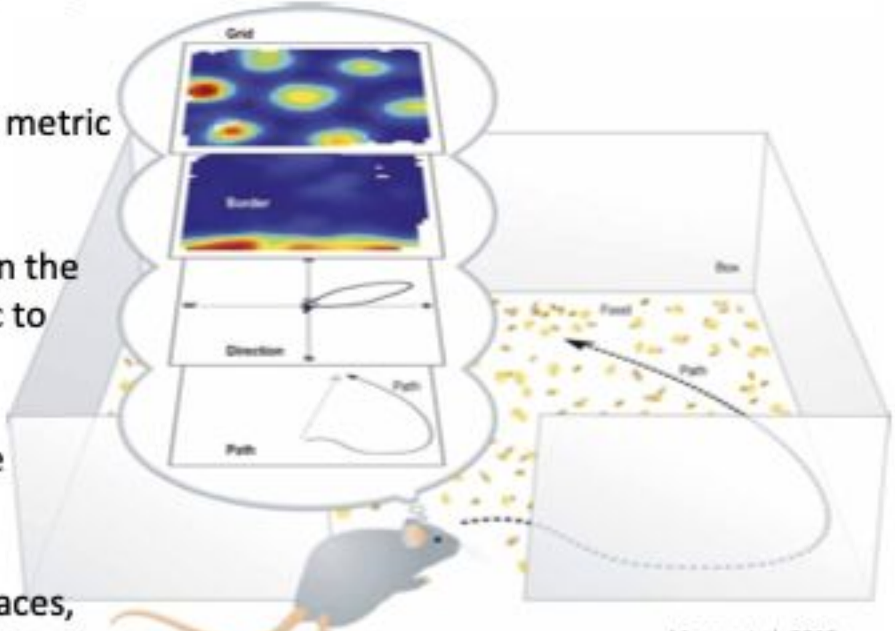
Place cells serve as landmarks in the brain's cognitive map, firing when an animal is in a specific location

Grid Cells create a hexagonal grid of spatial firing, acting as metric for the space (estimating of distances or vectors)

Head direction cells act as a neural compass, firing based on the animal's head direction, crucial for transforming egocentric to allocentric signals

The collective activity of these cells forms a comprehensive cognitive map for navigating complex environments.

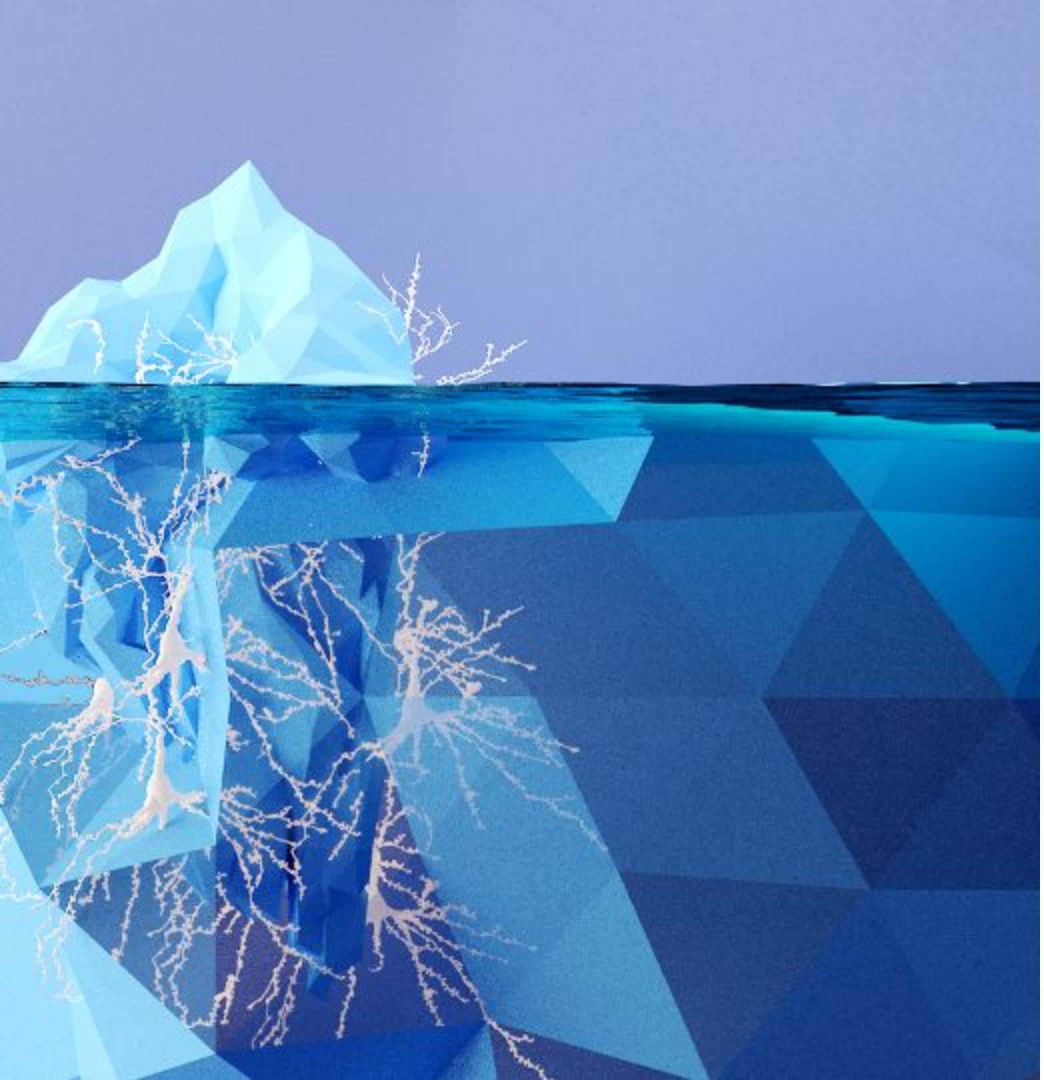
These mechanisms also enable the encoding of abstract spaces, suggesting a fundamental role in imagination, planning, and memory.



Moser et al. 2012

Hafting et al. 2005
Paper round-up

- They propose that the dMEC is part of a neural map of the spatial environment.
- They find a novel cell type in the dMEC that would be the basic unit of the map: the grid cell, which shows periodic firing as a response to non-periodic behavior.
- They show that the grid spacing, orientation and field size are topographically arranged from dorsal to ventral entorhinal cortex.
- They show that the grid phase vary randomly among co-localized cells, so that the full surface of the environment is represented within a local cell ensemble, suggesting a modular local organization of the spacial map.
- They suggest that grid cells are aligned to external landmarks but also persist in darkness (further work challenged that last point!).
- They find that the grid structure in the dMEC is expressed instantly in a novel environment, suggesting that the periodic structure is encoded by default, and the phase and orientation are set in relation to context-specific landmarks.



NeuroAI

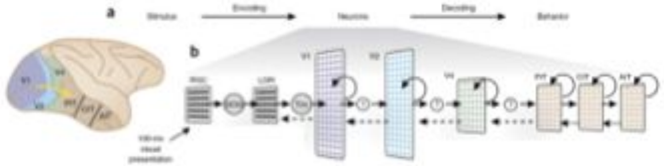
Intermediate Take Homes:

- NN models can be trained on different visual tasks to make hypotheses about the goal of the visual system
- Better NN models at the categorization task predicted IT neurons better
- Task mattered more than architecture or depth of networks
- Three points to consider when comparing:

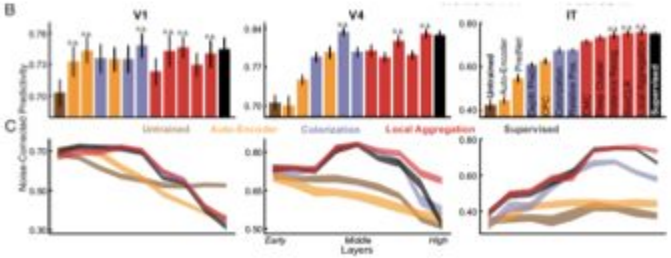
Task information consistency

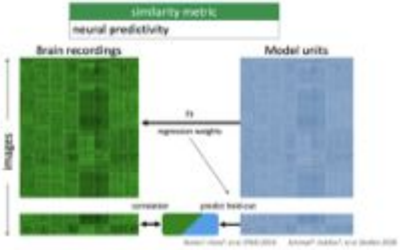
Single-unit response predictivity

Population representational similarity



Yamins and Di Carlo, Nat Neuro 2016



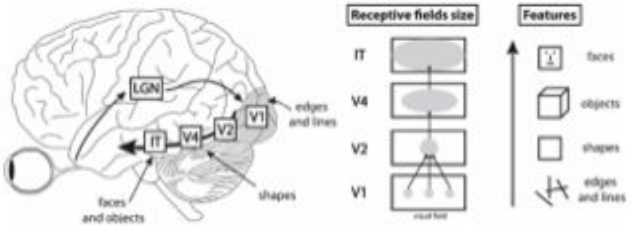


Summary

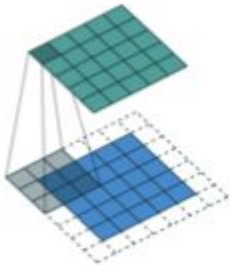
- NeuroAI is an emerging discipline that crosses across systems neuroscience and computer science
- It’s goal is broadly to use neuro insights to build AI, and to develop AI for understanding the brain (neuro)
- It is needed as it is still VERY hard to develop embodied AI, human-like movement into robotics, and we still lack generally intelligent systems (although LLMs for language are impressive …)
- Key example in Neuroscience inspiring AI: convolutional neural networks (likely transformers too “attention”): this is a hot area in industry – using cognitive neuro approaches to study NN btw!
- Interestingly, CNNs developed representations similar to the brain
- Key examples of AI influencing neuro: better behavioral analysis tools, better neural analysis tools (see also BCI week soon!)
- What is missing? NNs are very simple “neurons,” that lack the complexity of what we find in the real brain: an opportunity awaits!
- Data-driven and task-driven modeling: key approaches in neuroAI
- How do we model sensory systems: examples in vision and proprioception
- What to consider: both how close they are at single cell, task performance, and population level similarity
- Ongoing efforts: Brain-Score, **Inception Loops …**

What is neuroAI?

Neuroscience ⟷ Artificial Intelligence (AI)

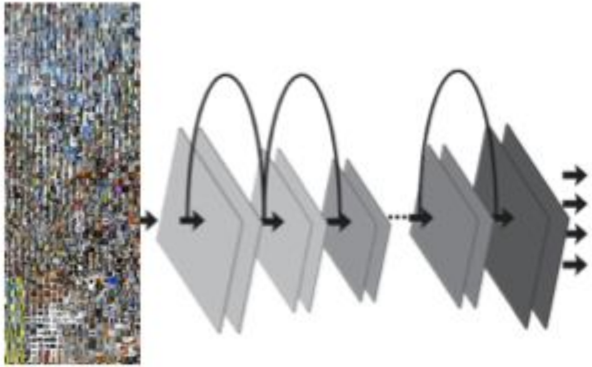


Hubel & Wiesel discoveries in cat V1 inspired convolutional neural networks



Convolutions (CNN)

Representations in ImageNet trained (CNN)



cat

What is neuroAI:

- Many definitions, but widely accepted that it is the **new inter-disciplinary field of merging neuroscience and AI research (← →)**

- Others define is more narrowly as using neuroscience (→) to shape research in AI

nature neuroscience

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature neuroscience › perspectives › article

Perspective | Published: 28 October 2019

A deep learning framework for neuroscience

Blake A. Richards, Timothy P. Lillicrap, Philippe Beaudoin, Yoshua Bengio, Rafal Bogacz, Amelia Christensen, Claudia Clopath, Rui Ponte Costa, Archy de Berker, Surya Ganguli, Colleen J. Gillon, Danijar Hafner, Adam Kepecs, Nikolaus Kriegesk
Richard Naud, Christopher C. Pack, Pana
Saxe, Benjamin Scellier, ... Konrad P. Kord

neuron

Review

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis,[1,2,*] Dharshan Kumaran,[1,3] Christopher Summerfield,[1,4] and Matthew Botvinick[1,2]
[1]DeepMind, 5 New Street Square, London, UK
[2]Gatsby Computational Neuroscience Unit, 25 Howland Street, London, UK
[3]Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK
[4]Department of Experimental Psychology, University of Oxford, Oxford, UK
*Correspondence: dhcontact@google.com

nature communications

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature communications › perspectives › article

Perspective | Open access | Published: 22 March 2023

Catalyzing next-generation Artificial Intelligence through NeuroAI

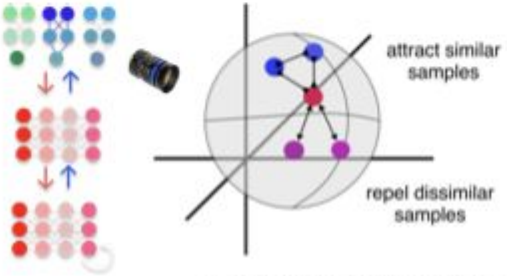
Anthony Zador, Sean Escola, Blake Richards, Bence Ölveczky, Yoshua Bengio, Kwabena Boahen, Matthew Botvinick, Dmitri Chklovskii, Anne Churchland, Claudia Clopath, James DiCarlo, Surya Ganguli, Jeff Hawkins, Konrad Körding, Alexei Koulakov, Yann LeCun, Timothy Lillicrap, Adam Marblestone, Bruno Olshausen, Alexandre Pouget, Cristina Savin, Terrence Sejnowski, Eero Simoncelli, Sara Solla, David Sussillo, Andreas S. Tolias & Doris Tsao — Show fewer authors

Data-driven modeling

GLMs, PCA, Sussillo et al. 2015 Nat Neuro
State-space models, ...

Record from neural data during a behavioral task

GLMs
.
.
.
ANNs
(CEBRA)



Schneider Lee Mathis 2023 Nature

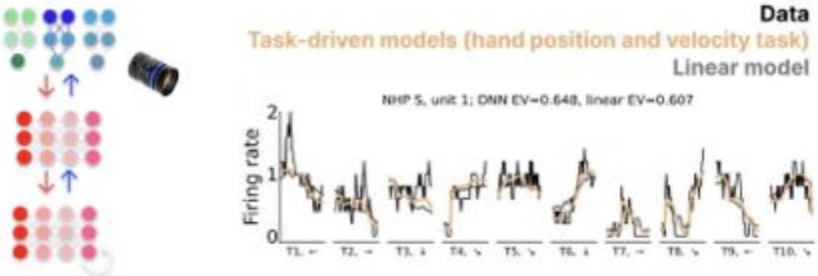
Joint models that describe
neural variance & representations

Task-driven modelling

Yamins et al. PNAS 2014, Kell et al. 2018 Neuron,
Banino et al. 2018 Nature ....

Constrain ANN based on behavioral task to test hypotheses about a system

Sandbrink et al. 2023 eLife



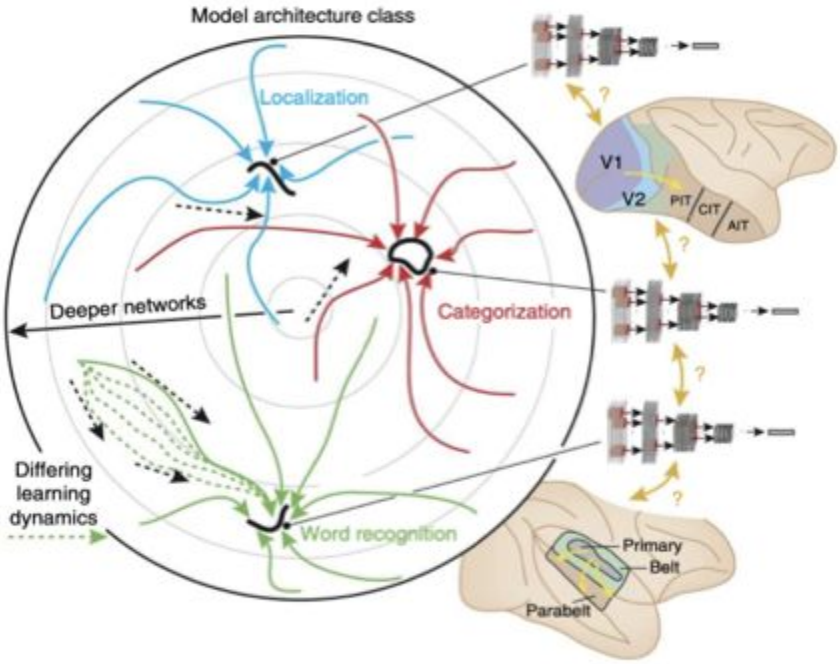
Marin Vargas et al. 2024 Cell

NN models that describe
neural variance & computationally
constrain system

Building models of visual pathway: the ingredients

Task-driven deep neural network models are built from three basic components:

1. **model architecture class** from which the system is built, formalizing knowledge about the brain's anatomical and functional connectivity;
2. a **behavioral goal** that the system must accomplish, such as object categorization; and
3. **a learning rule** that optimizes parameters within the model class to achieve the behavioral goal.



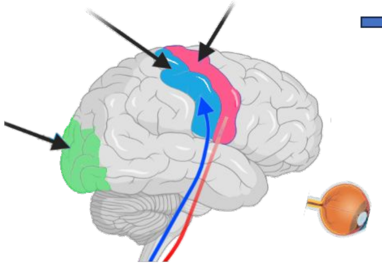
Yamins & DiCarlo (2016)

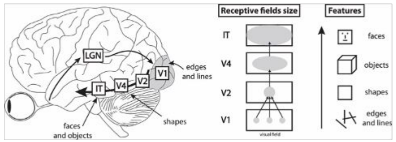
What other tasks? What other stimuli is the brain (visual) encoding?

Biological: recordings in visual system

Building “digital twins”, NN models of the system

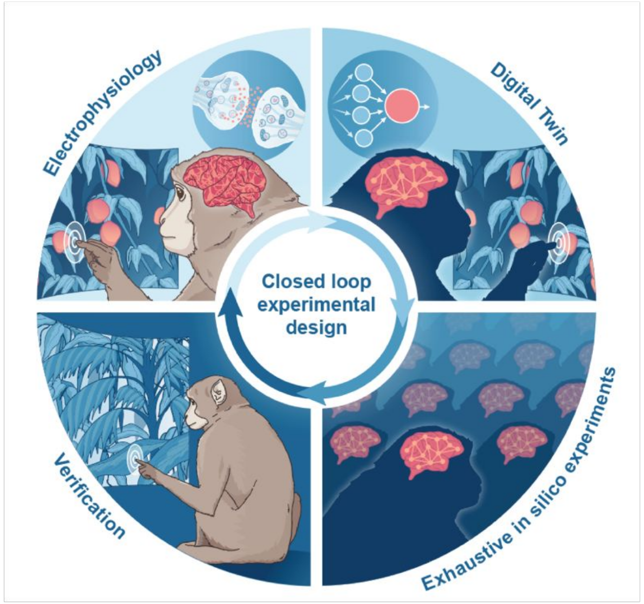
vision

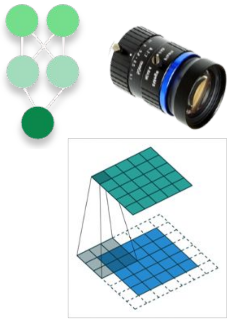




Hubel & Wiesel discoveries in cat V1 inspired convolutional neural networks

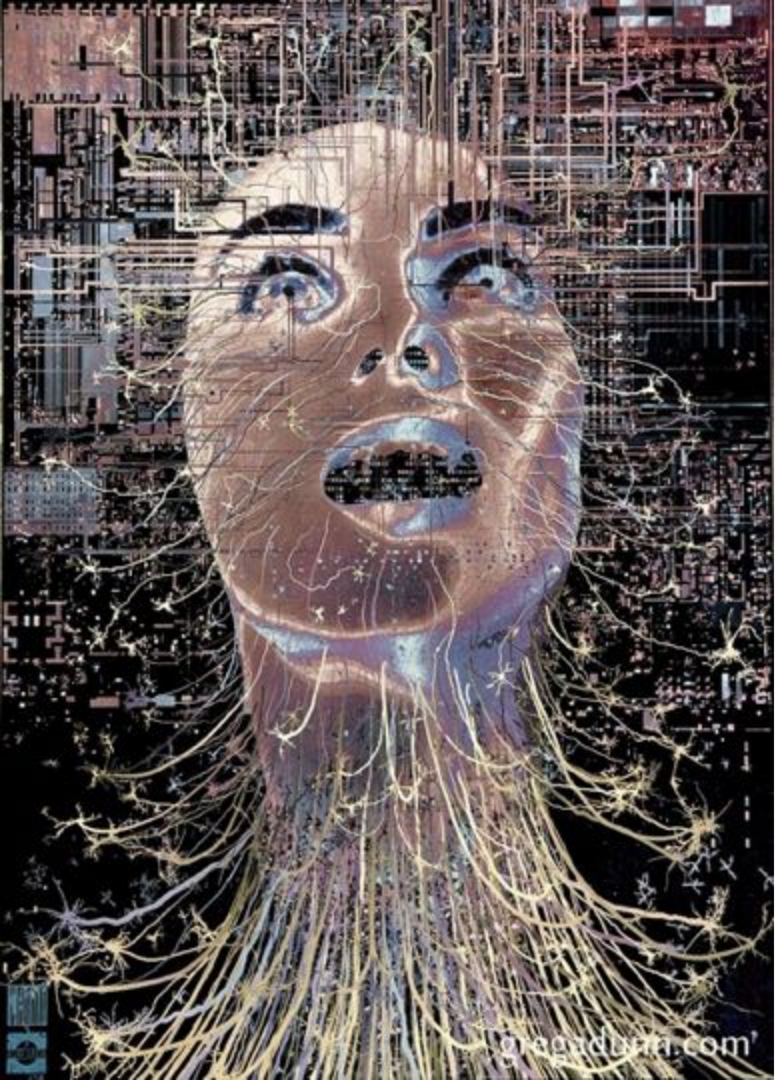
We now now a lot more (faces, motion, value coding) but we never can give enough stimuli ….
What would the ideal stimulus be for a given neuron?





Can we use our NN to produce predictions of optimal stimuli? Can this help reveal anew computational principle, or validate a discovered rule?

Mathis, Perez Rotondo, Chang, Tolias, Mathis (unpublished)



Brain Machine Interface for systems neuroscience

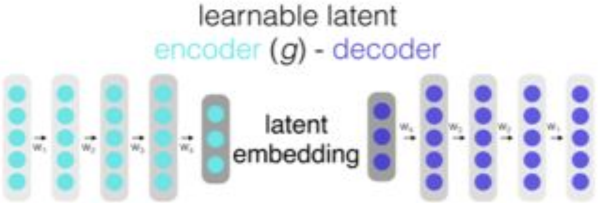
Summary

Overview:
•BCIs, or Brain Computer Interfaces, are systems that facilitate a **direct communication pathway between a brain and an external device.** This technology enables individuals to control devices using only their brain signals.
•**Recording neural activity is the foundation of how BCIs operate**. Specialized algorithms, known as decoders, are then employed to interpret these signals into commands that can control devices or computer systems.
•The importance of (**encoder-) decoder** algorithms lies in their ability to translate neural activity into actionable instructions for external devices, making them integral to the functionality of BCIs.
Systems Neuroscience Contributions:
•Instrumental in identifying optimal brain areas for signal recording, understanding neural subtypes, and designing effective sensory feedback within BCIs.
•Insights into neural dynamics, such as the relationship between neural firing and sensory stimuli or motor actions, thereby informing the development of more advanced BCIs.
•Current research in systems neuroscience contributes to BCIs by examining the principles of encoding sensory information into neural activity and decoding it back into the brain.

Current Technologies in BCIs:
•**Microelectrode arrays** are a key technology in BCIs, allowing for the stable recording of neural activity over extended periods. These arrays can be implanted and have been used in both research settings and, to a lesser extent, in clinical applications to assist individuals with paralysis.
•**Two-photon holographic optogenetics** represents a cutting-edge approach in BCI technology. It enables precise manipulation and recording of neural activity using light (calcium imaging and optogenetics).
•Technological advancements in BCI include increased recording stability and longevity, more biocompatible materials for implants, and higher throughput in signal recording. These improvements are crucial for the reliability and user-friendliness of BCIs, ultimately enhancing their applicability and integration into various aspects of life and healthcare. Ethics are also deeply important to consider.

Simple overview

Record Neural Activity → Build a Decoder Algorithm → Calculate requisite control parameters for external device



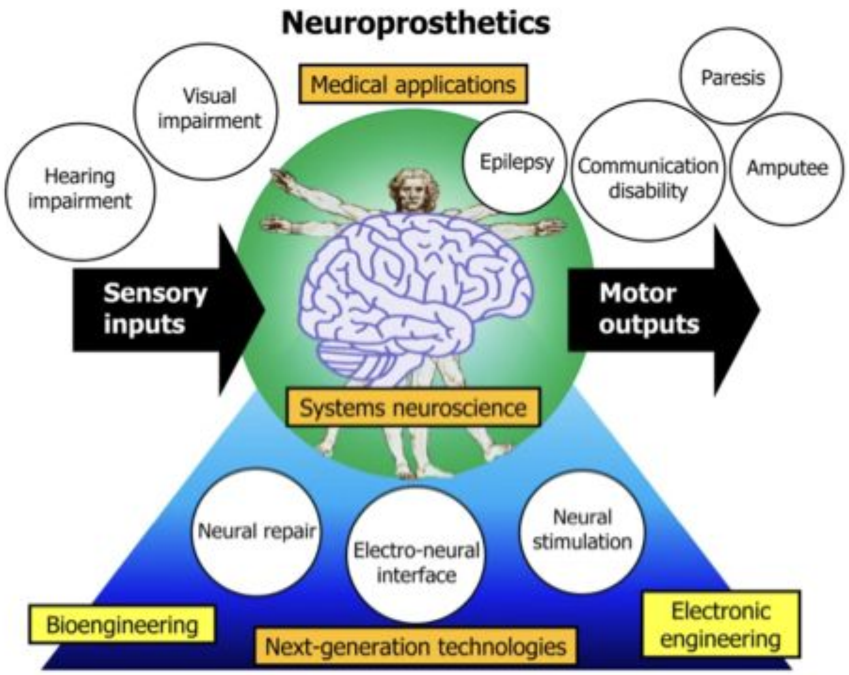
How systems neuroscience is enabling advancements in neuroprosthetics & BCIs

Neuroprosthetics in systems neuroscience and medicine

Our accumulating knowledge in systems neuroscience combined with the development of innovative technologies may enable brain restoration for patients with nervous system disorders. This Collection provides a platform for interdisciplinary research in neuroprosthetics. It will gather studies investigating medical applications of systems neuroscience, informatics, and engineering in the development of neural prostheses. Submissions with a clinical focus on nervous system diseases and brain repair in either humans or animals are also included.

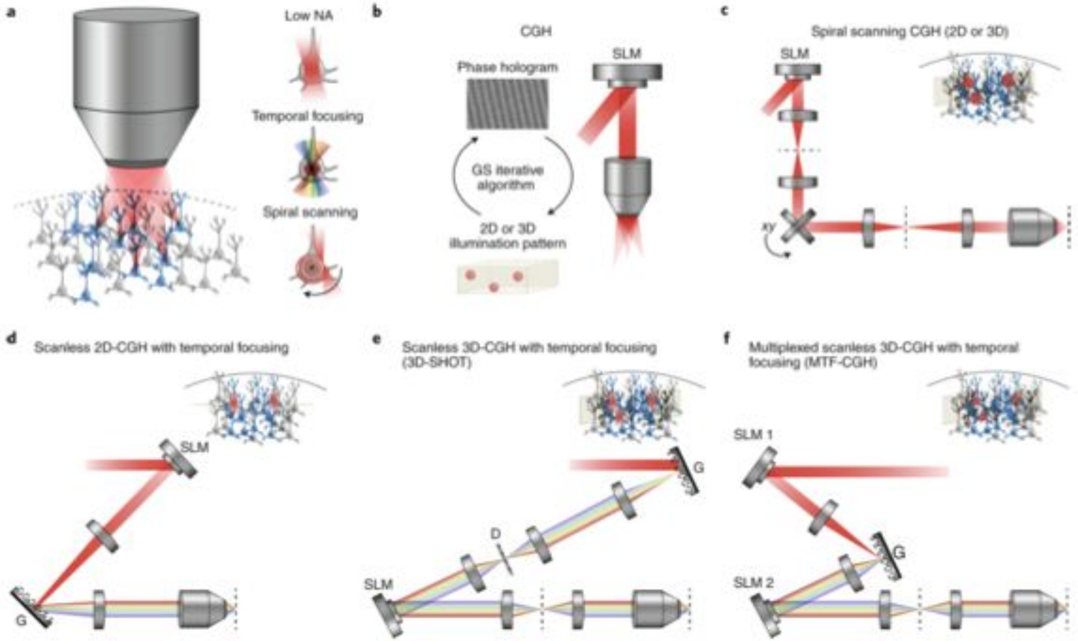
https://www.nature.com/collections/hjcgcjcach

- Which brain areas to record from
- Need to understand neural subtypes
- How to give appropriate sensory feedback
- How do we enable adaptation and learning



Reading & Writing into the brain: all optical studies

- **Calcium imaging + optogenetics allows for "all optical" access the neural circuits.**
- We can design closed-loop experiments to measure and perturb neural activity.
- We can design these such that we "closed-loop" record neural activity and have the animal use this activity to complete a task.



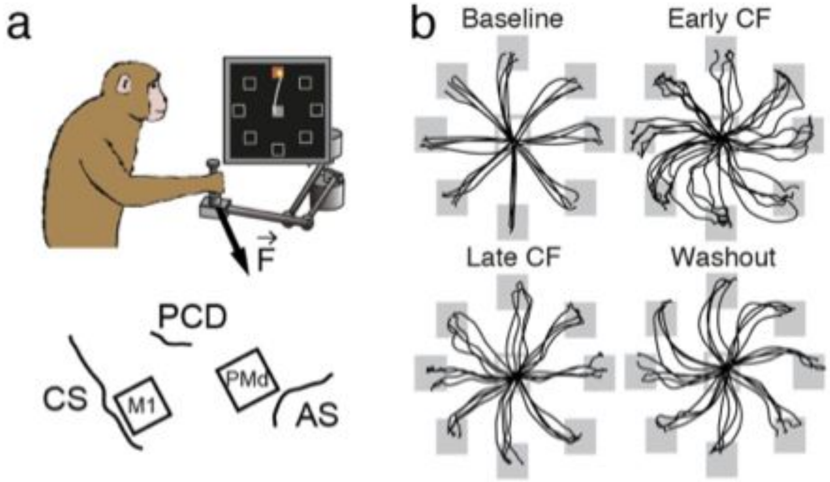
Adesnik, H., Abdeladim, L. Probing neural codes with two-photon holographic optogenetics. *Nat Neurosci* **24**, 1356–1366 (2021). https://doi.org/10.1038/s41593-021-00902-9

Center-out motor task

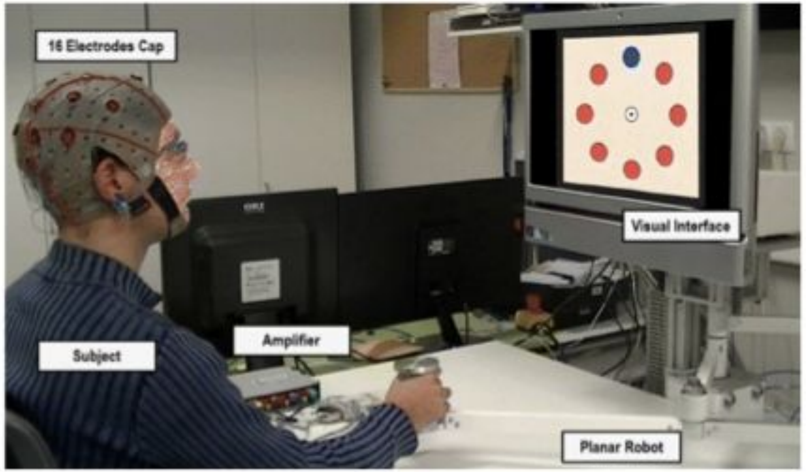
The center-out task is frequently used in motor control studies.

The task is very separatable yet constrained.

It can be performed with a limb, a cursor, or a robotic manipulandum.



Perich, Gallego & Miller, 2018 *Neuron*

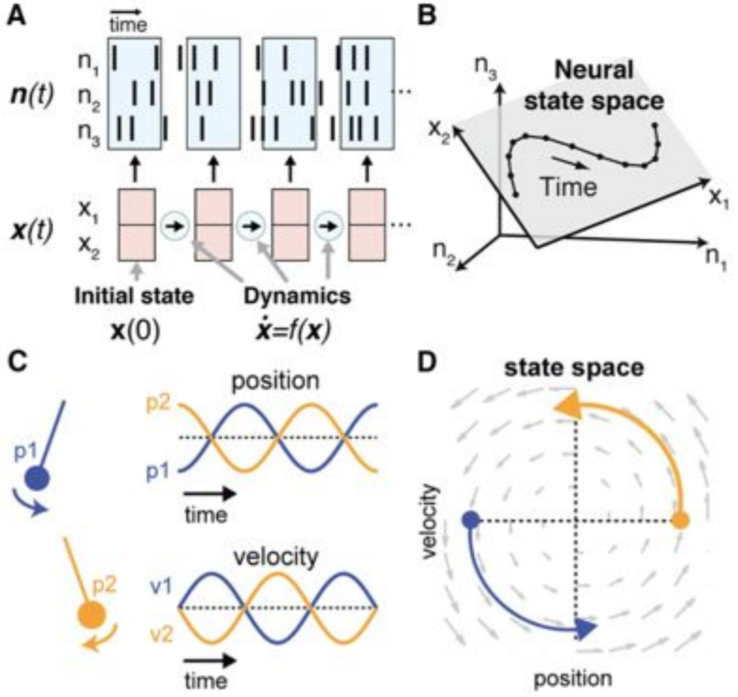


Úbeda et al., 2017 *J. NeuroEng & Rehab.*

What are dynamical systems?

"A set of coordinates, often represented as a vector, describing the instantaneous configuration of a dynamical system and that is sufficient to determine the future evolution of that system and its response to inputs."

Churchland Shenoy 2013
Ann. reviews in neuroscience



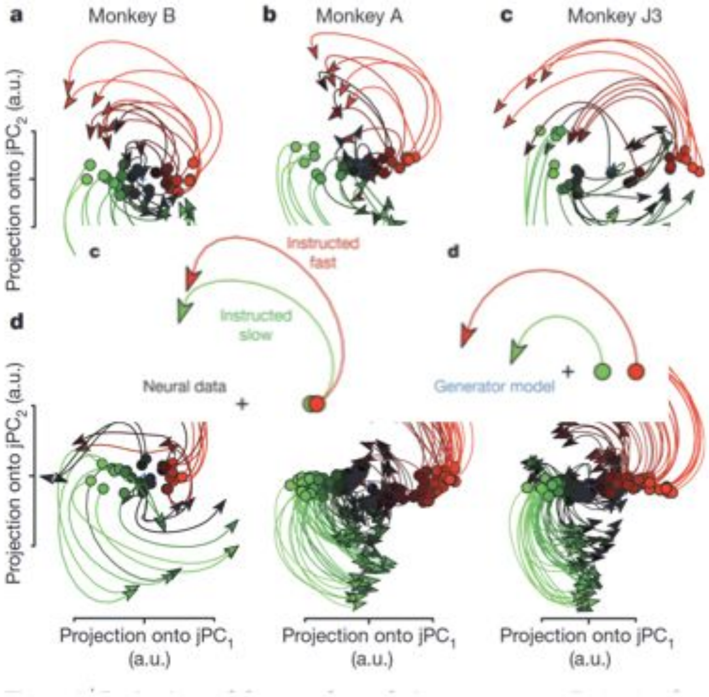
Pandarinath et al., 2018 *J neuroscience*

Rotational dynamics

- Low dimensional projections of neural activity during center-out reaching tasks produce highly consistent neural trajectories.

- These cyclical trajectories appear to show organization based on movement kinematics including direction and velocity of movements.

- This finding suggests that motor cortex acts as a dynamical system, with neural activity evolving over time based on local dynamics and external inputs.



Chruchland et al., 2012 *Nature*

Athalye et al. 2023,
Paper round-up

- Monkeys were trained to control a cursor in several tasks using a BMI based on neural activity recorded in motor cortex.
- Invariant dynamics in the recorded neurons could predict the neural activity that was used to produce a motor command, even when task inputs were removed from the model.
- Invariant dynamics alter neural activity in dimensions relevant to the decoder, demonstrating a causal link between invariant dynamics and motor commands (at least in this BMI setting).
- Adding an optimal feedback controller to an *in silica* model of invariant dynamics trained to perform the center out tasks reduced the amount of inputs needed for successful execution.

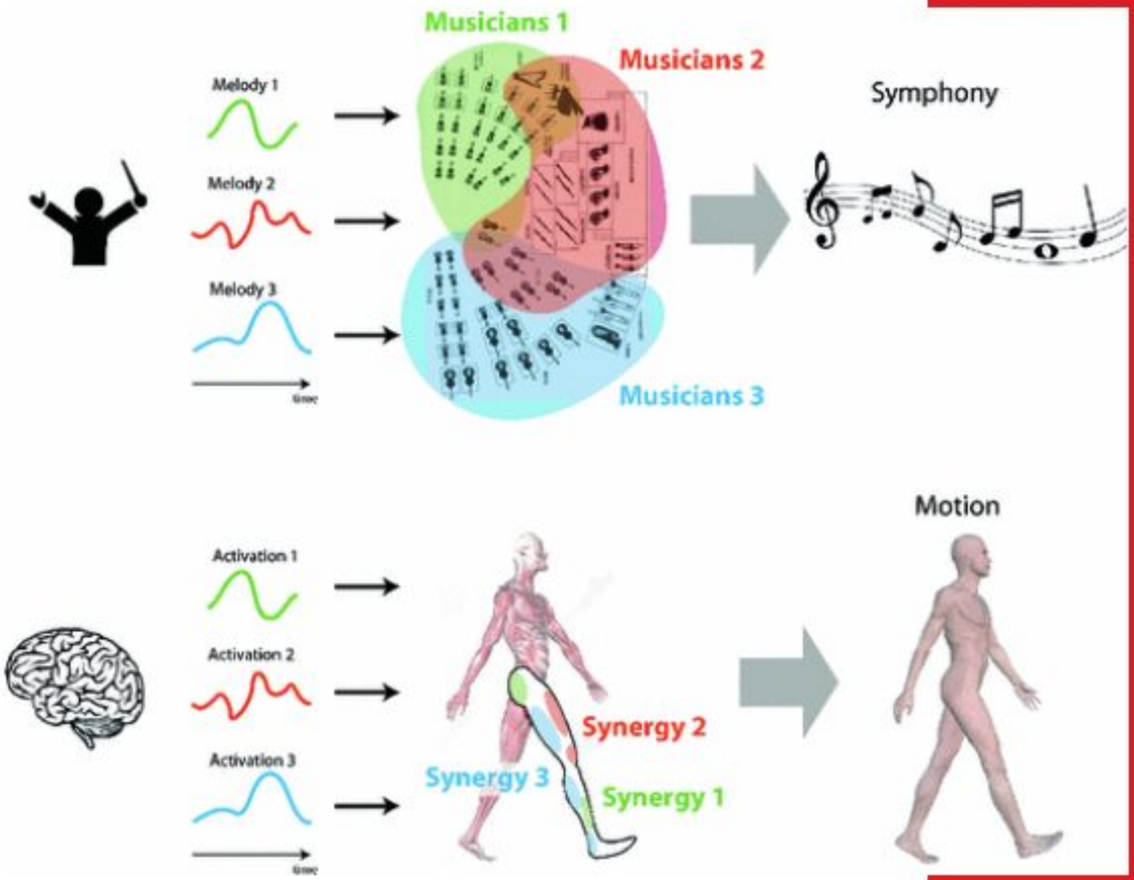
Cognition & skill learning

Summary

- Muscle synergies have been proposed as a key principle for motor control
- Yet, low-dimensional nature might be underestimated with existing techniques!
- For the hand -- learned muscle synergies are highly task-specific, and thus generalize poorly
- This suggests that low-dimensional control is an emergent property (of the task/biomechanics/distributed circuits) rather than the mechanism of control (not a simplifying strategy)
- Neural networks are ideal for taming complex biomechanics
- Training neural network with a curriculum based learning leads to better performance
- Muscle synergies from artificial agents resemble closely the ones from humans

EPFL

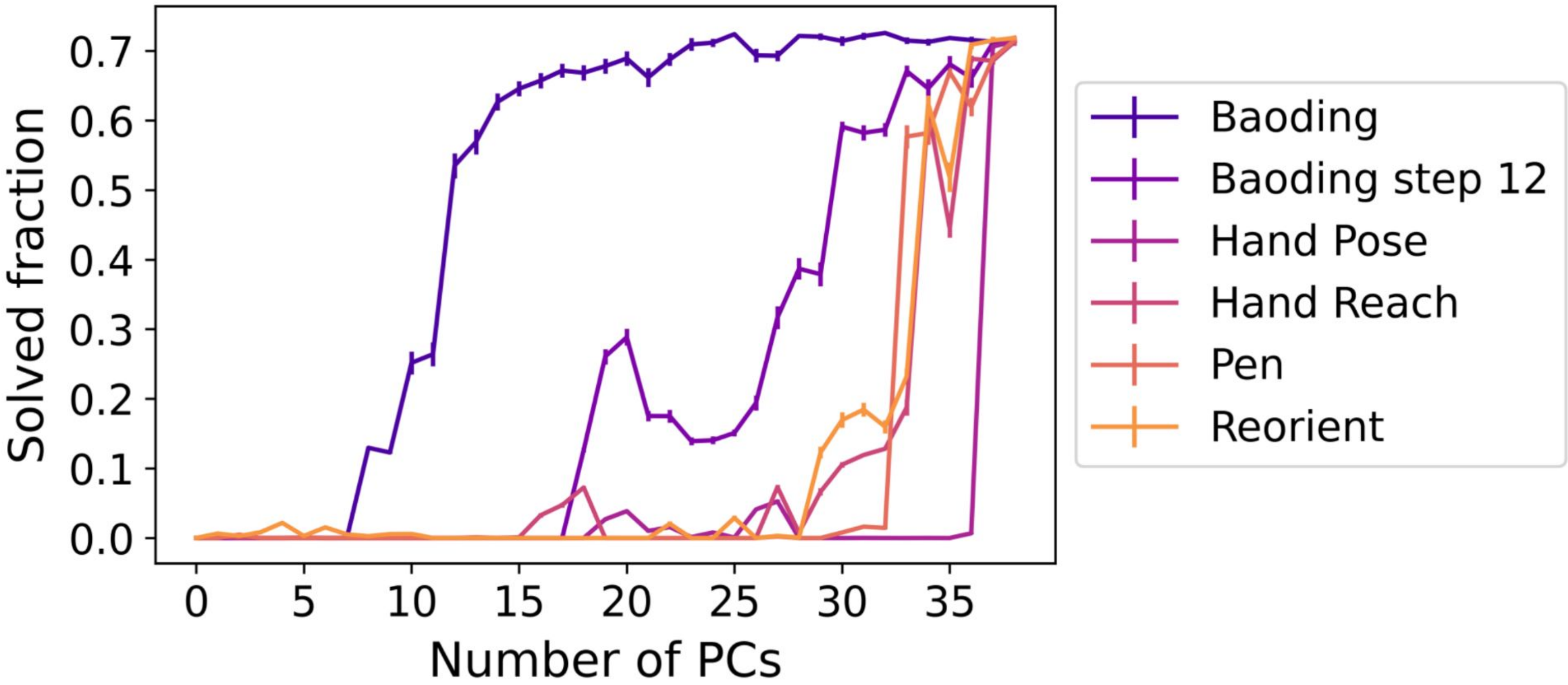
Muscle synergies as principle for motor control



Torricelli et al., Emerging Therapies in Neurorehabilitation 2015

EPFL

Control spaces are highly task-dependent & transfer poorly

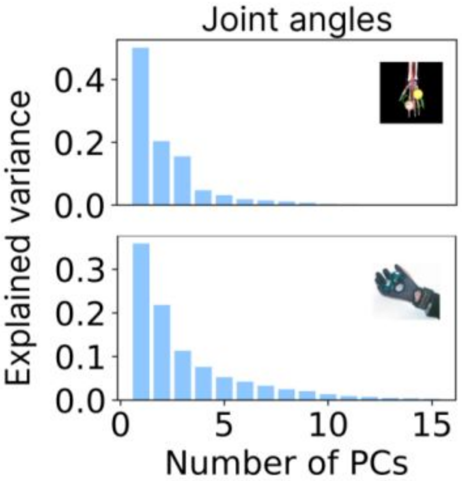


SDS also discovers a low-dimensional control space

SDS

Human

		Position	Muscle act.
SDS	Baoding	4.5	12
	Control	8	7
Human	Baoding	5	
	Control	8.5	



This notion of muscle/kinematic synergy is purely based on reconstruction error!

Chiappa et al 2024 , Paper round-up

- They succeeded in training a musculoskeletal model on an object-manipulation task.
- They propose a static to dynamic stabilization (SDS) curriculum, inspired by coaching practice.
- They show that, akin to experimental data, SDS learns low-dimensional kinematic and kinetic spaces.
- They show that muscle synergies are highly task specific and thus generalize poorly.
- They found that more dimensions contributed to the task performance than suggested by traditional synergy analysis.
- They found lower tangling of the dynamics in the controller state space than in the action state space, consistent with previous observations that motor cortical dynamics avoiding tangling more than muscle dynamics.