# Neural Data Science
## The ‘Big Data’ Revolution


Mark Reimers, PhD
Michigan State University


Course on Neural Data Science
Cold Spring Harbor 2019

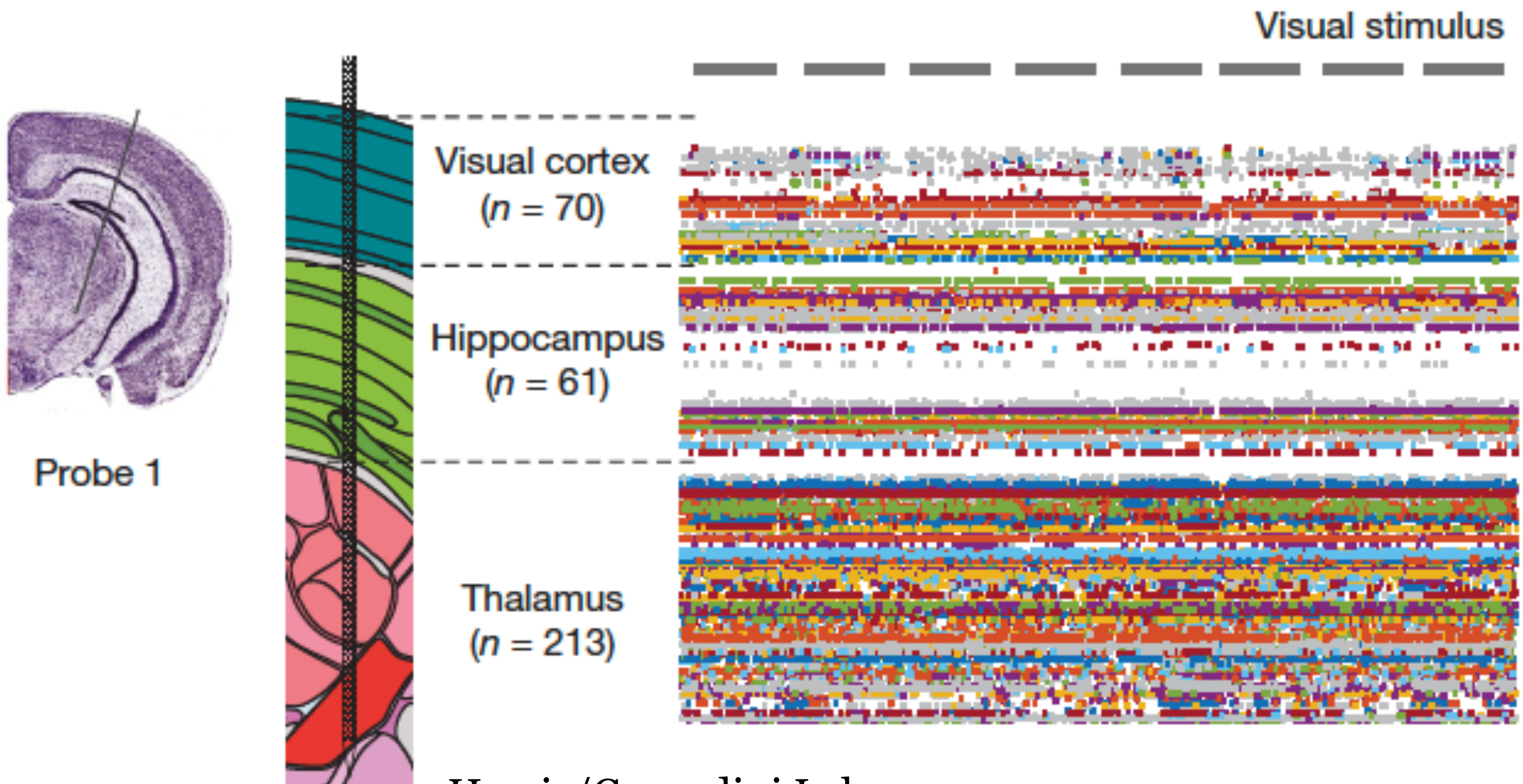
# Neuroscience is Changing

New recording technologies – especially optical imaging – enable us to gather data in quantities unimagined ten years ago.
The challenge is to make sense of it.

These new technologies will enable us to see how the brain is working as a system, and to address the deeper questions about brain function that brought many of us into neuroscience

# Many Spiking Neurons

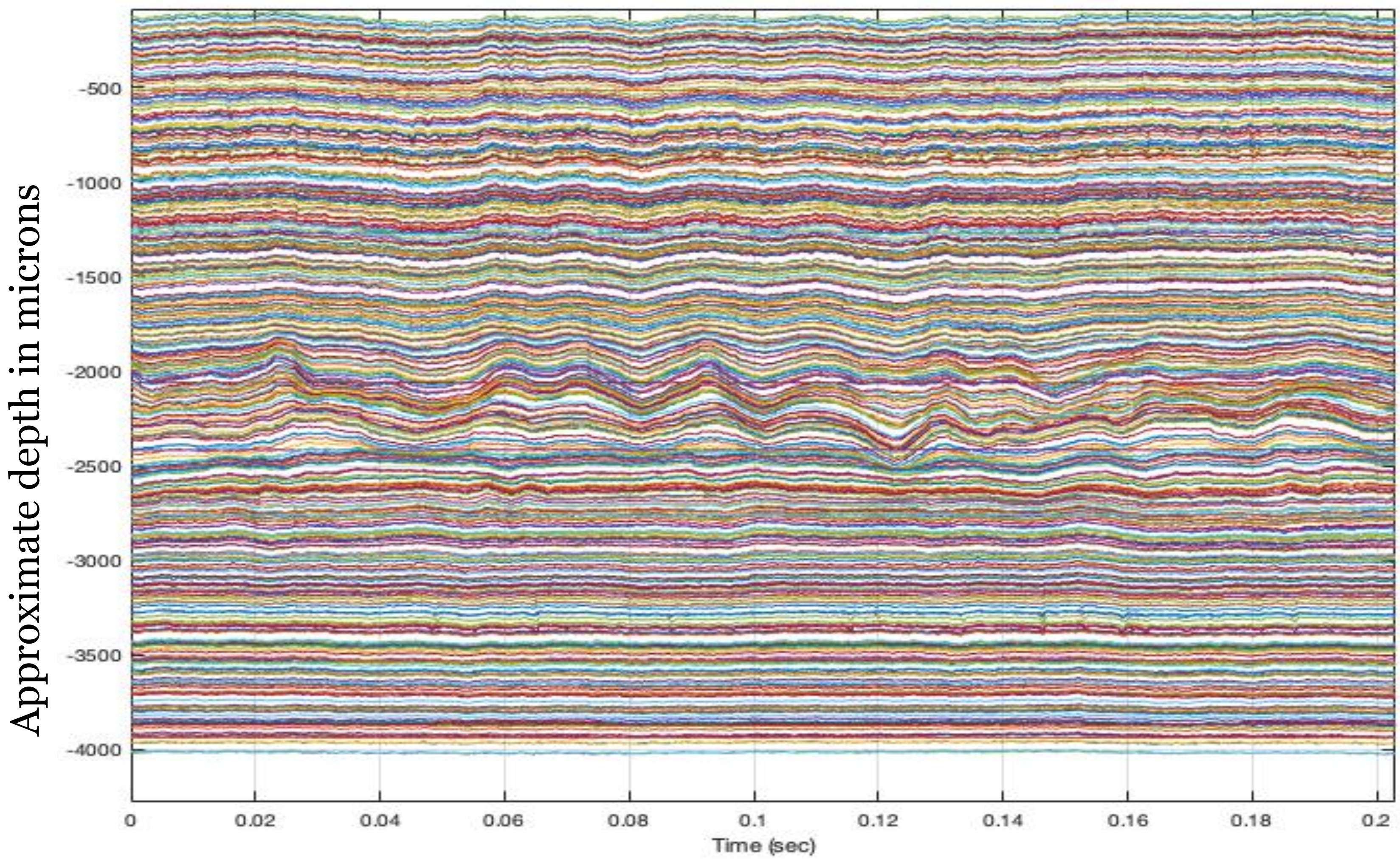
- Neuropixels probes enable recording of hundreds of spiking neurons at ms resolution




Harris/Carandini Lab

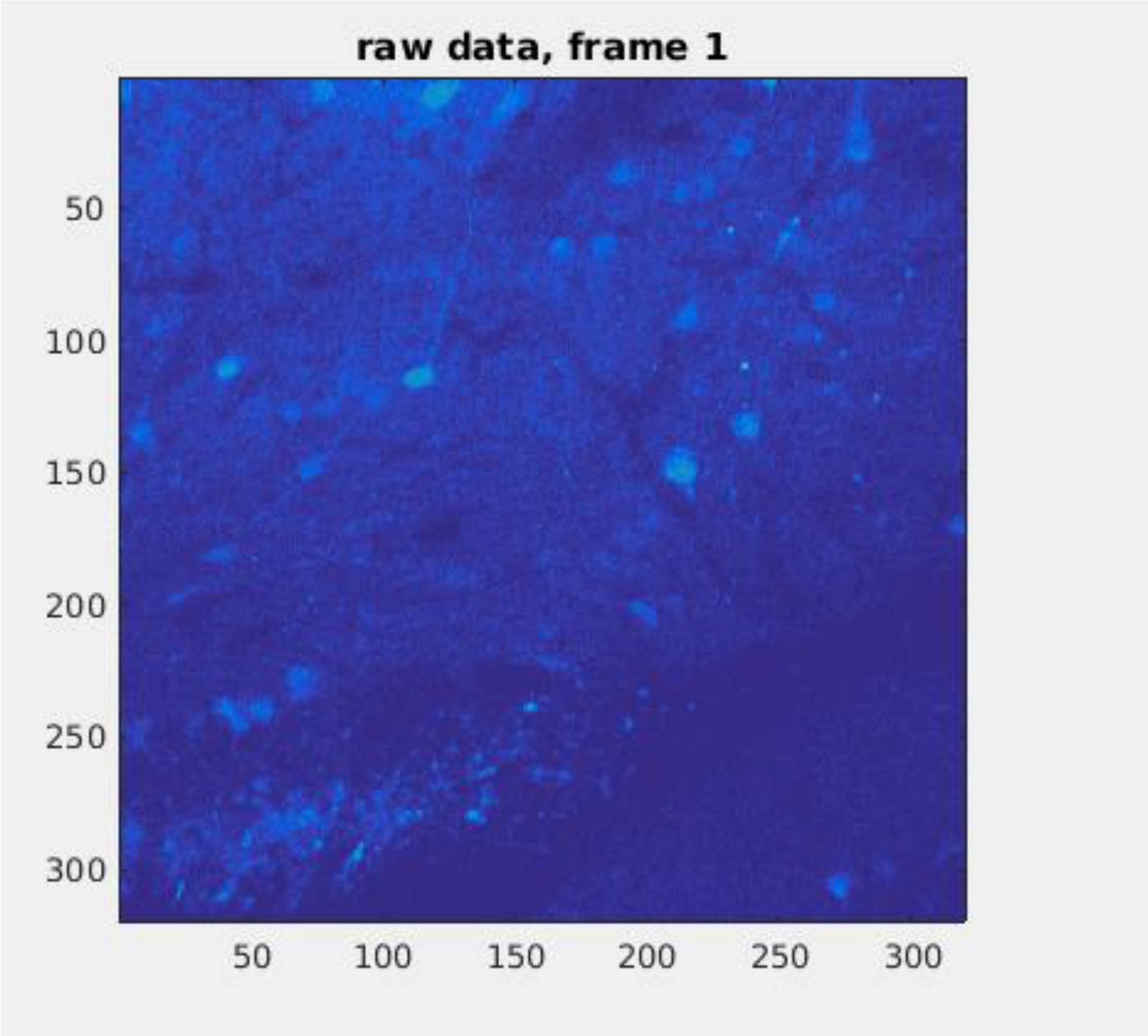
# What the Voltage Profile Looks Like

- Voltages at 384 depths from pia surface over 200ms



# Activity of Motor Cortex in Motion

- Visualizing activity across layers in motor cortex using prism with 2P imaging




Courtesy Khirug Lab, Helsinki

# Twelve Thousand Bursting Neurons

- Video from Marius

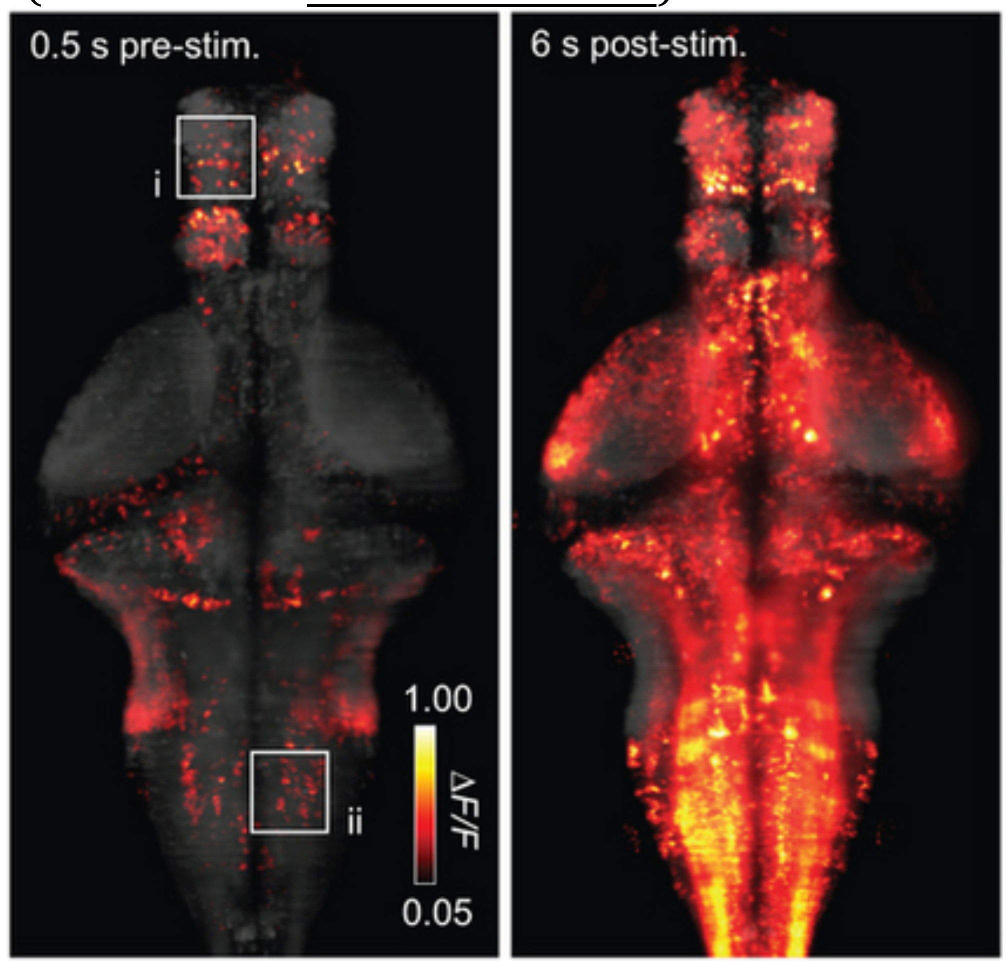
# Really Big Data: Cells in Whole Brain

Calcium imaging can record activation of thousands of individual neurons

Chronic recordings go for hours at 1 – 20 frame/sec

From 20 GB to 1 TB of image data to analyze

Zebrafish brain activity imaged with calcium at two times. The brain is 700 x 300 microns. 85,000 individual neurons are measured (Ahrens et al *Nature Methods*)

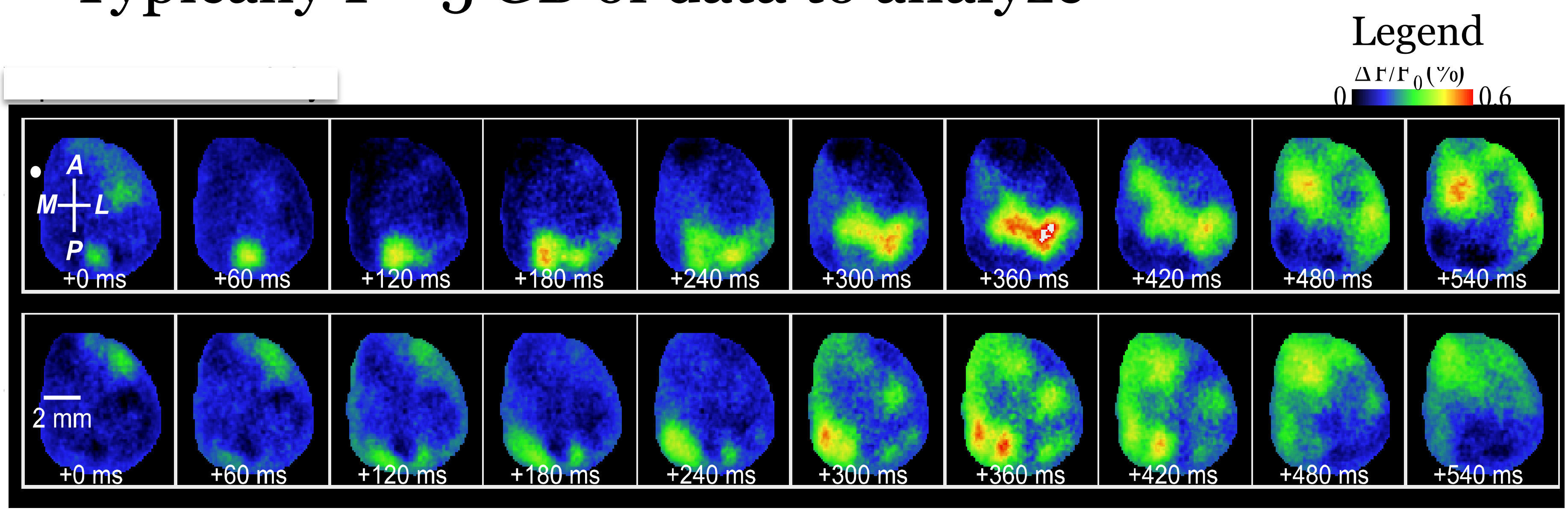


# Imaging of Cortical Surface

Voltage-sensitive dyes give very fast response to voltage changes in neurons

Three episodes below show patterns of spontaneous activity moving over one hemisphere of mouse cortex

Typically 1 – 5 GB of data to analyze



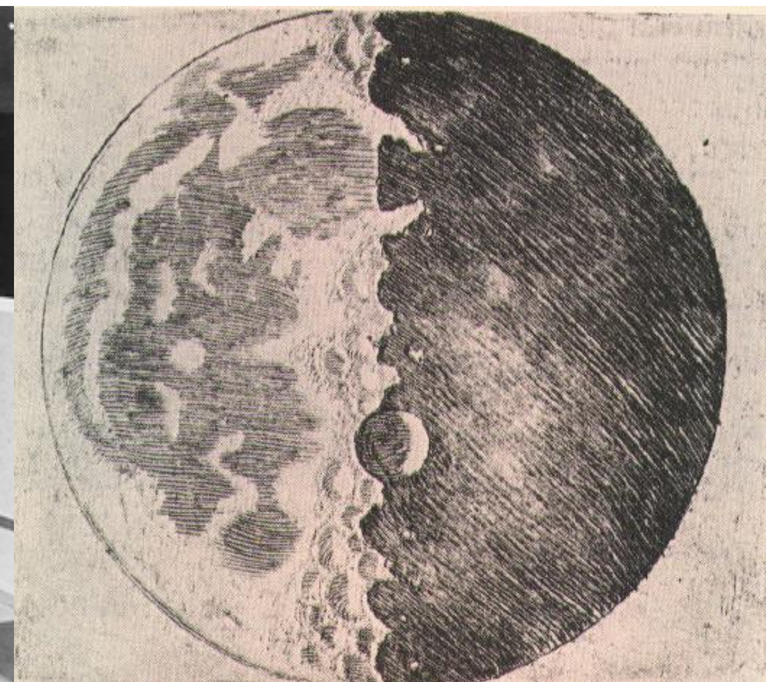

Mohajerani et al 2013

# What Spatial Resolution Do We Need?


Courtesy Tony Kim (Schnitzer Lab)
1 min (500 frames at 9.5fps) of resting-state data

# Rich Data is Revolutionary

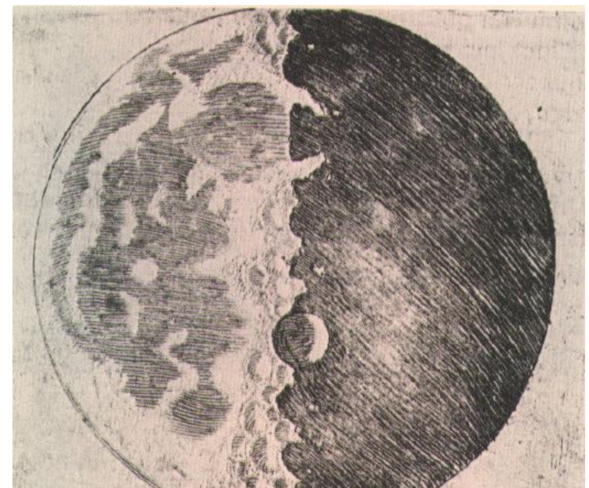
Galileo's telescope revealed unknown worlds, and unexpected detail in worlds astronomers thought they knew well

One of Galileo's drawings of the moon showing mountains

# Rich Data is Revolutionary

Galileo's telescope revealed unknown worlds
….with a bit of embellishment

# Neural Data Analysis is Changing

From Tuning to Dynamics

# What Can You Do with One Neuron?

- Neurons spike irregularly
- Want to find some correlate in the real world
- Characterize conditions under which a neuron is more likely to fire
- Works fairly well for sensory systems
- John O'Keefe won Nobel for finding location correlate for deeper cells

# Neural Data Analysis is Changing

- More discussion of ‘patterns’ in activity of many neurons or regions
- Less focus on ‘tuning’ or ‘representing’ or regional activation
- Introduction of ‘state space’ and dimension reduction methods
- Focus on characterizing variation in activity many neurons over single trials rather than averaging one over many trials
- Comparison with more complex simulation models

# New Statistical Techniques

- High-dimensional methods (Lasso, etc) to characterize response profiles
- New approaches to dimension reduction: ICA, NMF, various kinds of 'guided' PCA
  - Yu, Mante, Machens, Pillow, ... picture the dynamics of neural processes
- Network characterizations
  - Trying to infer active communication networks
  - Network metrics: degree, clustering
- Machine learning methods
  - (Controversy: What do they tell us?)

# How is Infrastructure Changing?

- Bigger machines with more RAM
- Networked CPUs
- GPUs
- Data in the cloud
- Bandwidth to transfer big data
- Need convenient sockets to connect processes (e.g. in Python or MATLAB on HPCC)

# How is Data Analysis Culture Changing?

- More use of creative graphics to represent high-dimensional data
- Integrating data of different types
- More cross-disciplinary collaborations
- Asking questions about dynamics
- Data sharing is becoming more common
    - CRCNS
    - NeuroData Without Borders
    - CodeNeuro
- Journals and funders ask for open data

# What New Questions Are We Asking?

- How are different regions in the brain communicating with each other?
- What is relation of brain states to behavior?
  - How variable are brain states with a behavior?
  - What are typical progressions of brain states?
- What are intrinsic dynamics of brain states?
- How are regional dynamics affected by inputs?
- How does learning affect communication between brain regions?
- Brain as ‘dynamical system’ metaphor
  - Shenoy et al Ann. Rev Neuro 2013
  - Balaguer-Ballester et al, PLoS Comp Bio 2011

# Metaphors for the Brain

Aristotle thought the brain was like a wax tablet
– the ‘killer app’ information technology of the ancient world

It was soft enough to take impressions and erasable

# Metaphors for the Brain

Descartes thought the nervous system was composed of hydraulic tubes

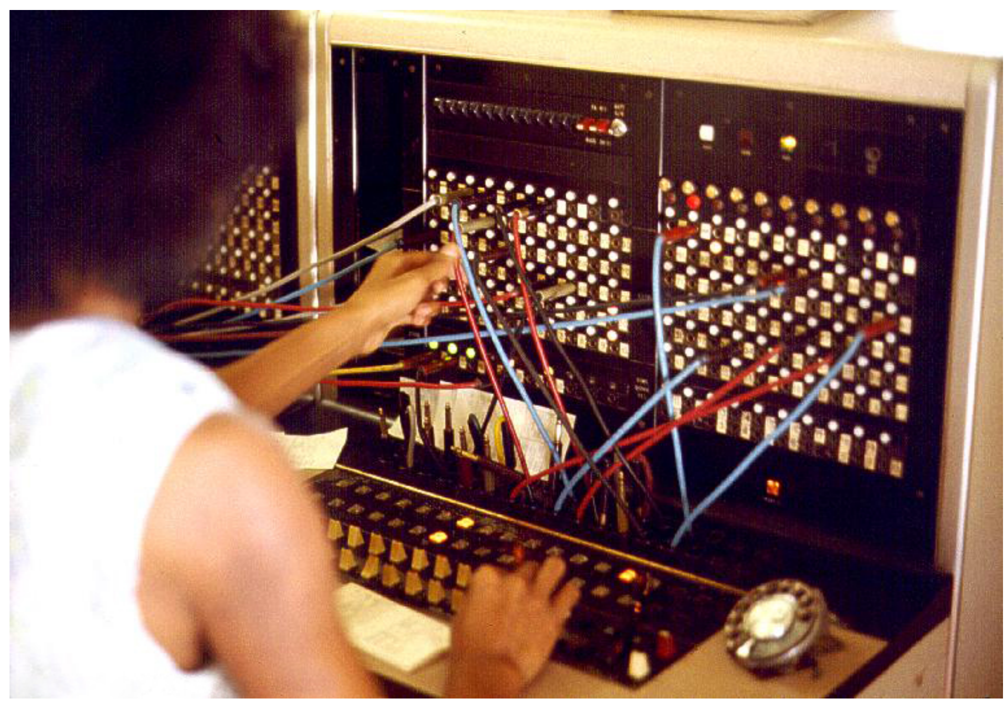
"animal spirits" flow from the ventricles of the brain, through the nerves, to inflate the muscles


Foot-withdrawal reflex according to Descartes

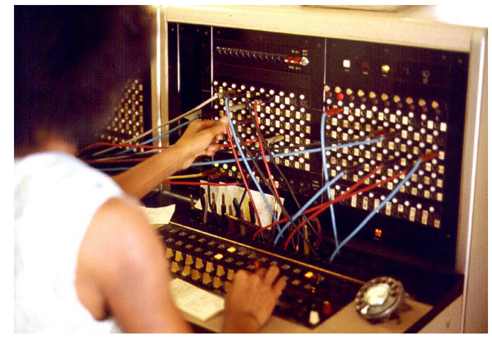
# Metaphors for the Brain

The glamorous information technology of the mid-twentieth century was the telephone switchboard

The tangled spaghetti of cords even looked like the neuropil



# Metaphors for the Brain



Now we ‘know’ that the brain is like a computer
Its ‘job’ is ‘information processing’ using algorithms like those used by programmers

Gary Marcus:
*Face It, Your Brain Is a Computer*
NY Times June 27, 2015

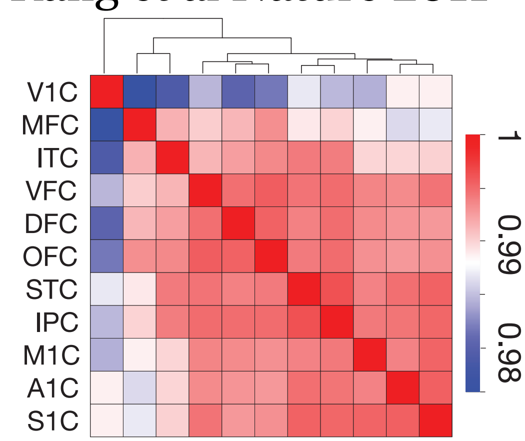
# Metaphors Shape Research Questions

- If information processing is what the brain does, then it makes sense to monitor information transformations within regions and flows between regions
- This strategy has been moderately successful in understanding neural activity in visual cortex

But…

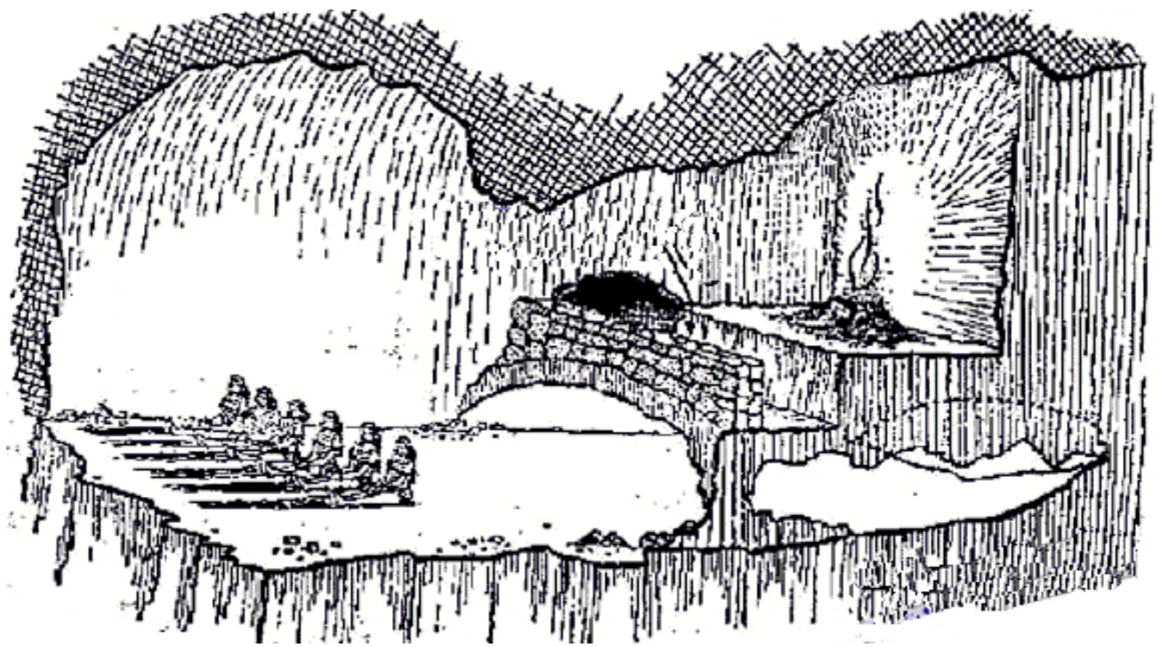
- Vision isn't typical of brain function
- Visual cortex is not like other cortex
  - Gene expression patterns are most different from other regions
- The approach isn't working well in motor cortex or PFC

Correlations between regional gene expressions Kang et al Nature 2011



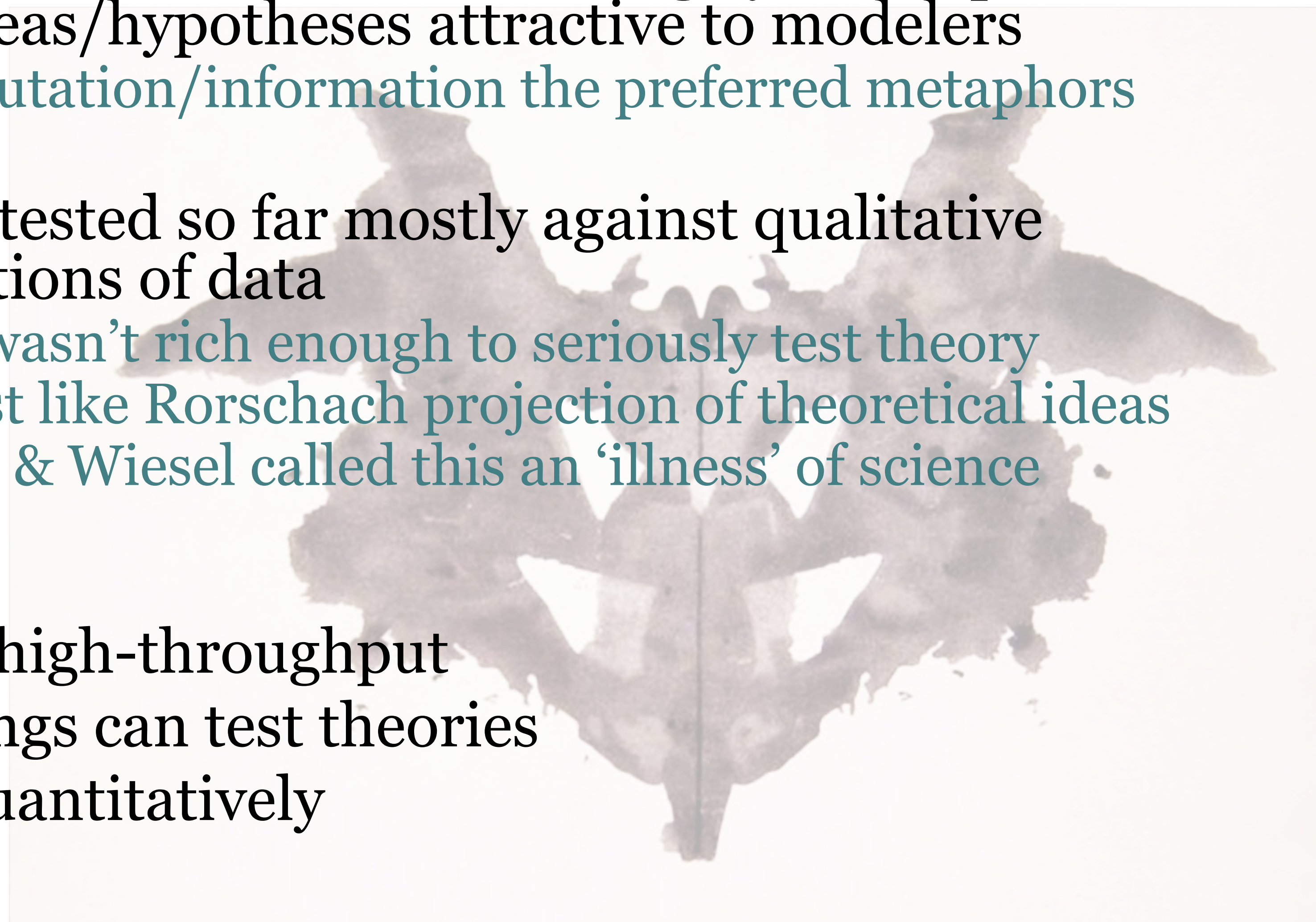
# Reality, Measures and Metaphors

- We are like the prisoners in Plato's Cave
- Data are the shadows on the cave wall, by which we infer the reality
- Data limited by our technologies and we are shackled by our metaphors



# Theory and Data in Neuroscience

- Theoretical neuroscience has largely developed from ideas/hypotheses attractive to modelers
  - Computation/information the preferred metaphors
- Theory tested so far mostly against qualitative descriptions of data
  - Data wasn’t rich enough to seriously test theory
  - Almost like Rorschach projection of theoretical ideas
  - Hubel & Wiesel called this an ‘illness’ of science
- Recent high-throughput recordings can test theories more quantitatively



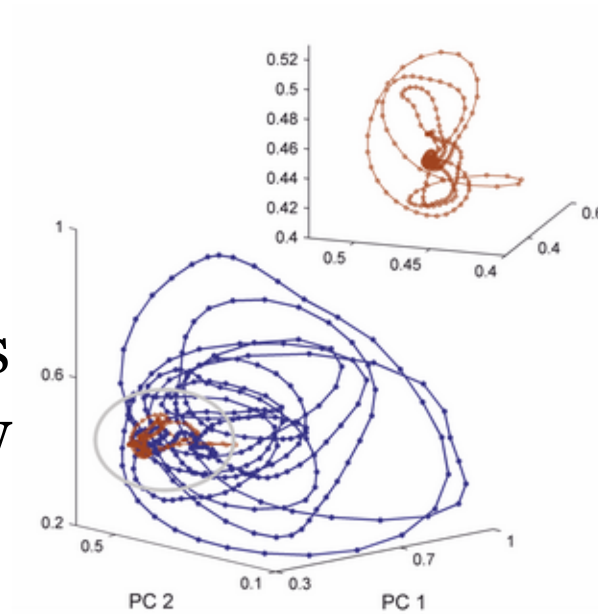
# Neural Data Analysis Overthrows Our Favored Ideas

What new analyses of rich data have done for theory in neuroscience

# State Space

- The activity of all neurons at one time can be represented as one point (vector) in a high-dimensional space of all possible activity measures
- The activities change during a behavior, defining trajectories
- Often recurrent coordinated changes in activity will be apparent in narrow repeated trajectories through same small regions of state-space

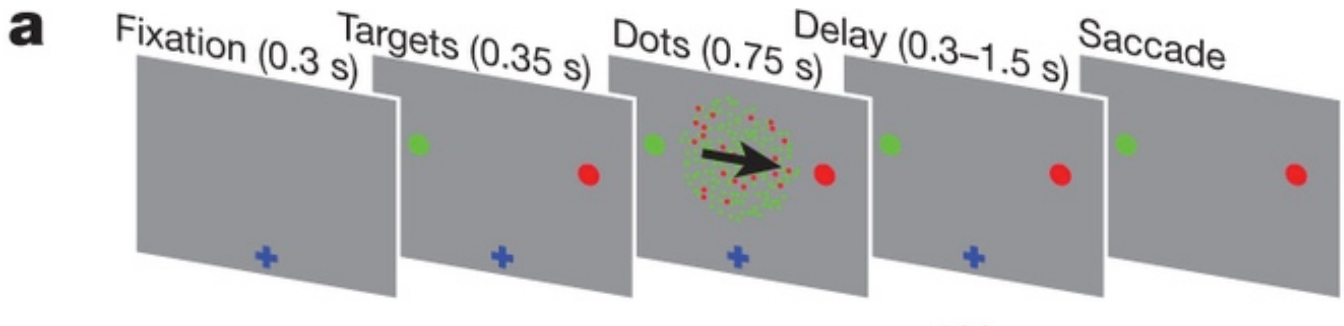
Example trajectories of neural activity during activity




Balaguer-Ballester et al
PLoS Comp. Bio., 2011

# An Example: Representing State-Space

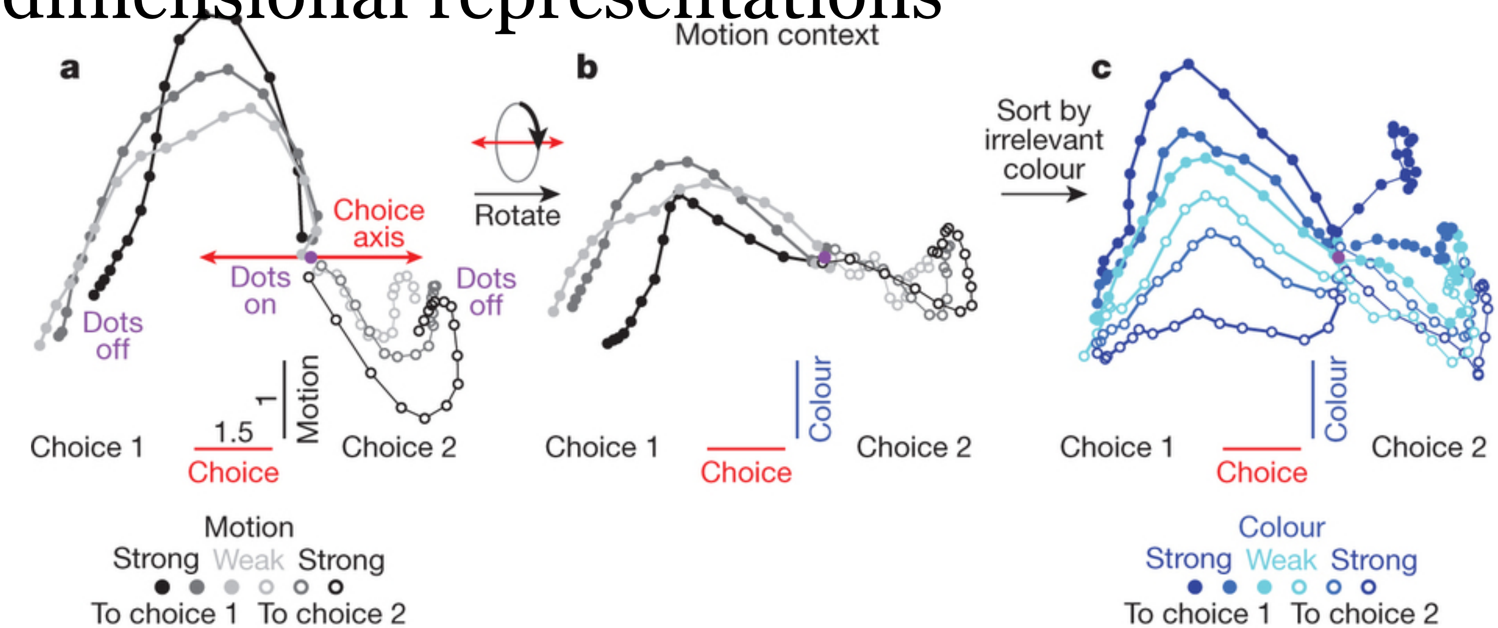
- (Mante et al, 2013) asked monkeys to discriminate color or motion of moving dots




- They recorded from ~ 200 neurons and wanted to represent how neural states change during discrimination

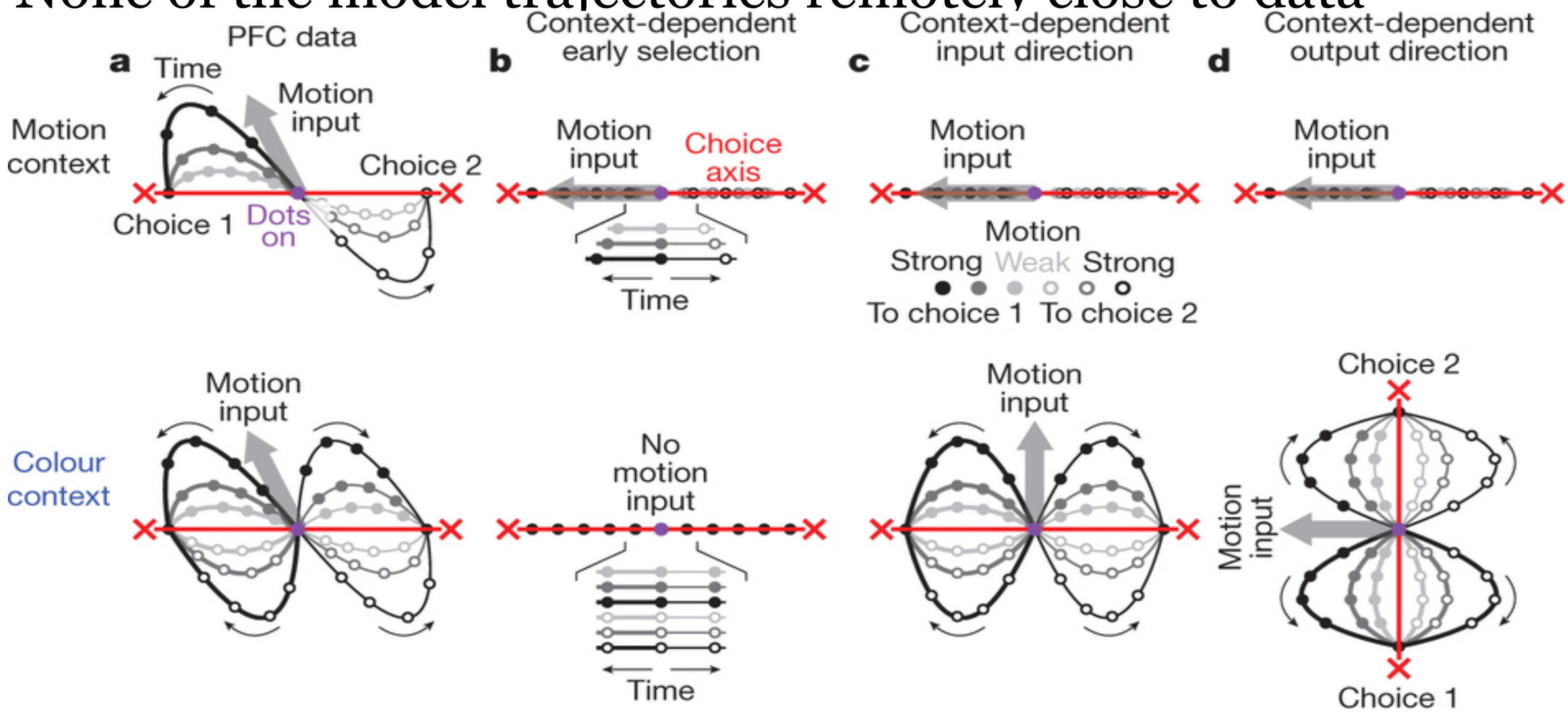
# Representing State-Space

- (Mante et al, 2013) used GLM to obtain regression (tuning) parameters for each neuron
- They then projected regression coefficients onto the subspace of first 12 PCs to obtain low-dimensional representations



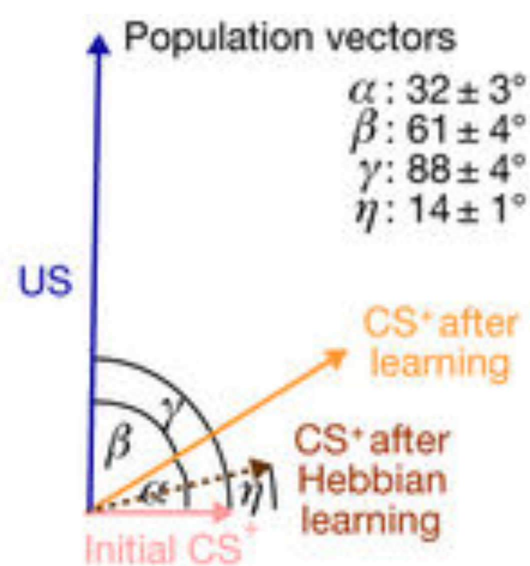
# Testing Ideas with Rich Data

- Mante et al, (2013) computed trajectories of spike recordings from > 100 neurons monkey PFC during perceptual choice task (left column)
- They simulated three leading models (studied for decades) of PFC activity during choice
  - For each they calculated the trajectories in state-space during the choice and compared to data (Right three columns)
- None of the model trajectories remotely close to data



# Classical Conditioning not Hebbian

- Grewe et al (Schnitzer lab) did classical fear conditioning (CC) while recording hundreds of neurons in BLA with calcium imaging
- They found few neurons whose changes follow Hebb's model (intended to explain CC!)
- Rather the pattern of activity (represented by population vector) in response to the CS changed (rotated) toward that activated by the US




B F Grewe *et al. Nature* 1–6 (2017)

# ‘Voyages of Discovery’

Up to the 18th century Europeans had (seemingly) adequate dogmatic ideas about plants and animals

The ‘voyages of discovery’ opened up a new world and gave rise to what we now think of as ‘biological science’

These are revolutionary times in neuroscience…

Data science will play a central role

# Themes of Neural Data Science

Pre-processing

- How to get from the raw recording – voltages, fluorescence, etc. – to accurate measures of activity that we seek

Characterizing joint activity

- Dimension reduction, dynamics, networks

Integration

- What kinds of neurons are acting in specific patterns

# Our Aims in This Course

- Introduce you to key methods in neural data analysis as practiced now
- Expose you to the ideas we think will become important
  - Get you thinking about how things will change over your career
- Our approach to data analysis is deliberately exploratory rather than model-driven
- Stimulate you to think about how data analysis can get at big questions of neuroscience