Computation and the Brain 2019

welcome to Week 2

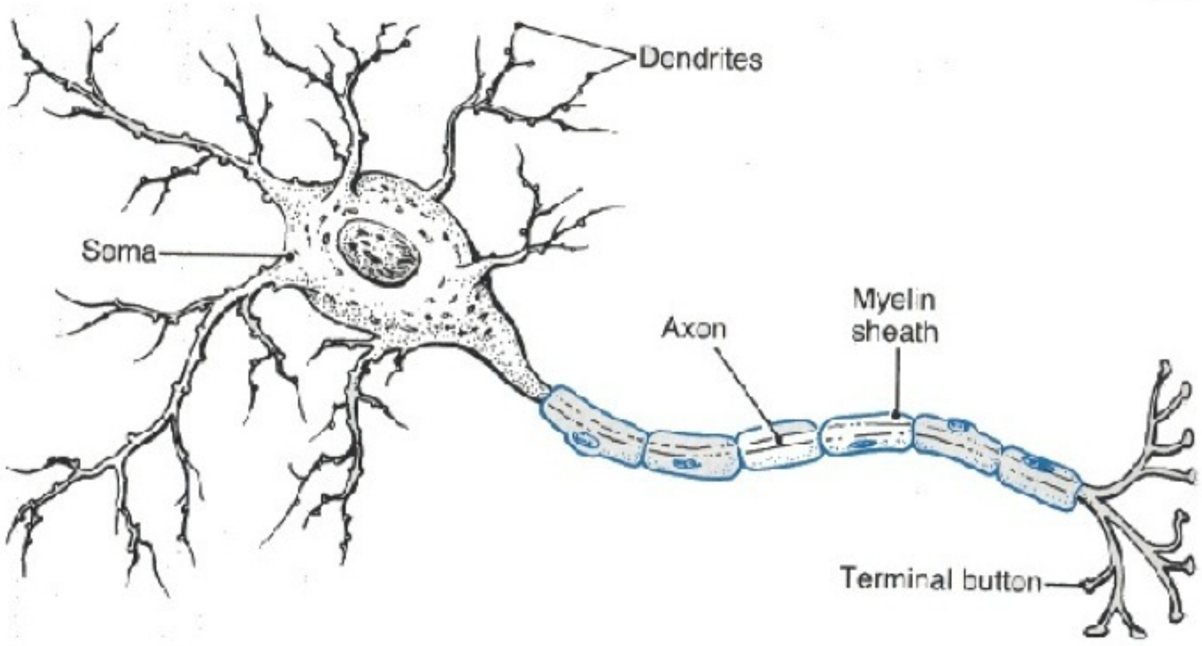
First: What happened last Wednesday

COMPUTER	THE BRAIN
Inorganic	Alive
Exact	Error prone
70 – 7,000,000 Watts	20 Watts
Chess, Go, Jeopardy	Survival, creativity, dreams
$10^{10 -- 18}$ operations/s 1 – 40,000 processors	50 operations/s, 10^{11} components
Connects to others	If in the mood
Was designed	Has evolved
Cool machine	Rather emotional
Behavior is deliberately programmed	Behavior is emergent

Administrivia

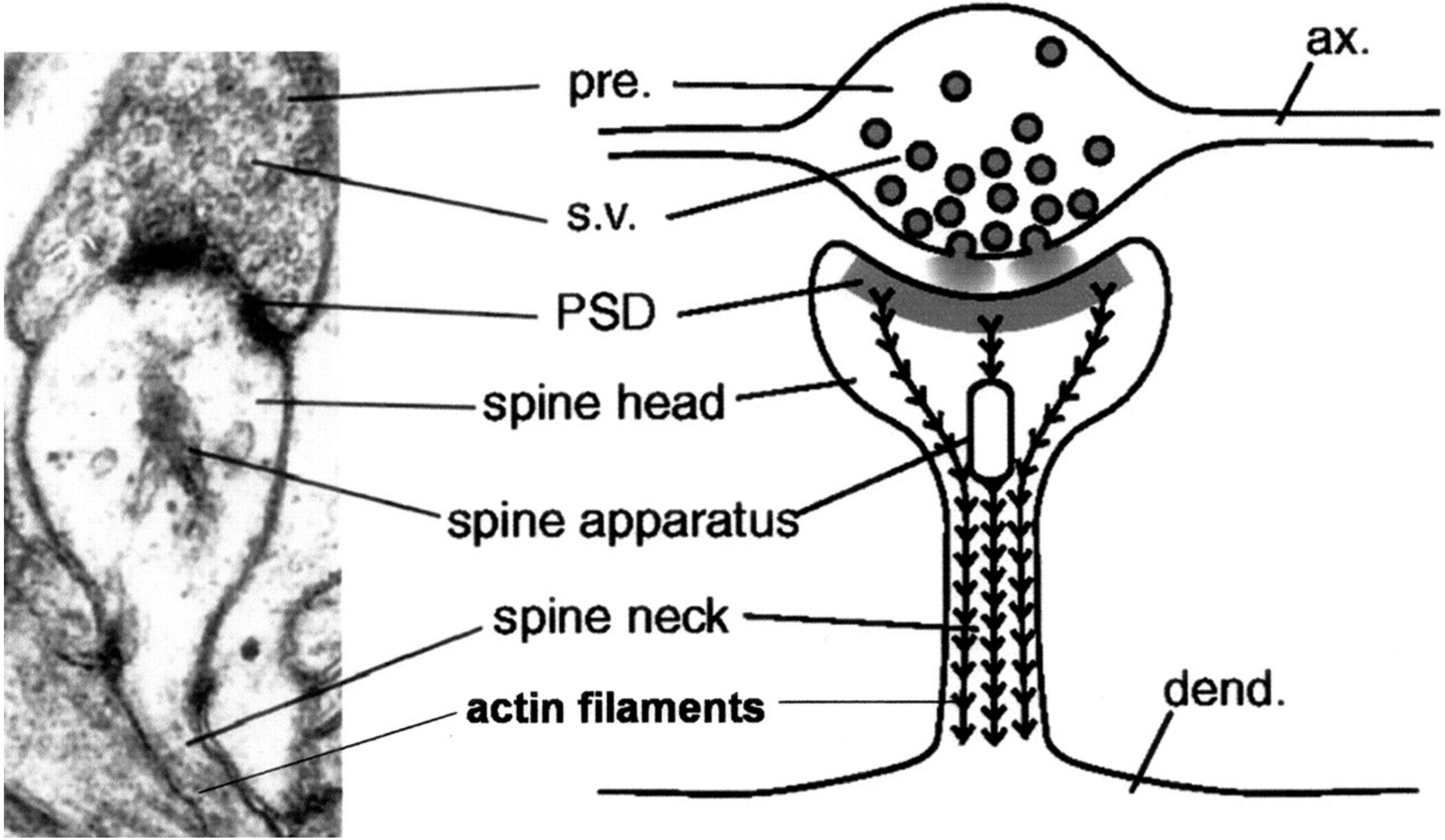
- The nature of the course
- Homework
- Participation and attendance
- The project
- ***The wait list***

The Neuron



- soma
- dendrite
- axon
- cone
- myeline
- boutons and spines
- membrane and its potential (-70 mV in steady state)
- ***many*** kinds

The synapse



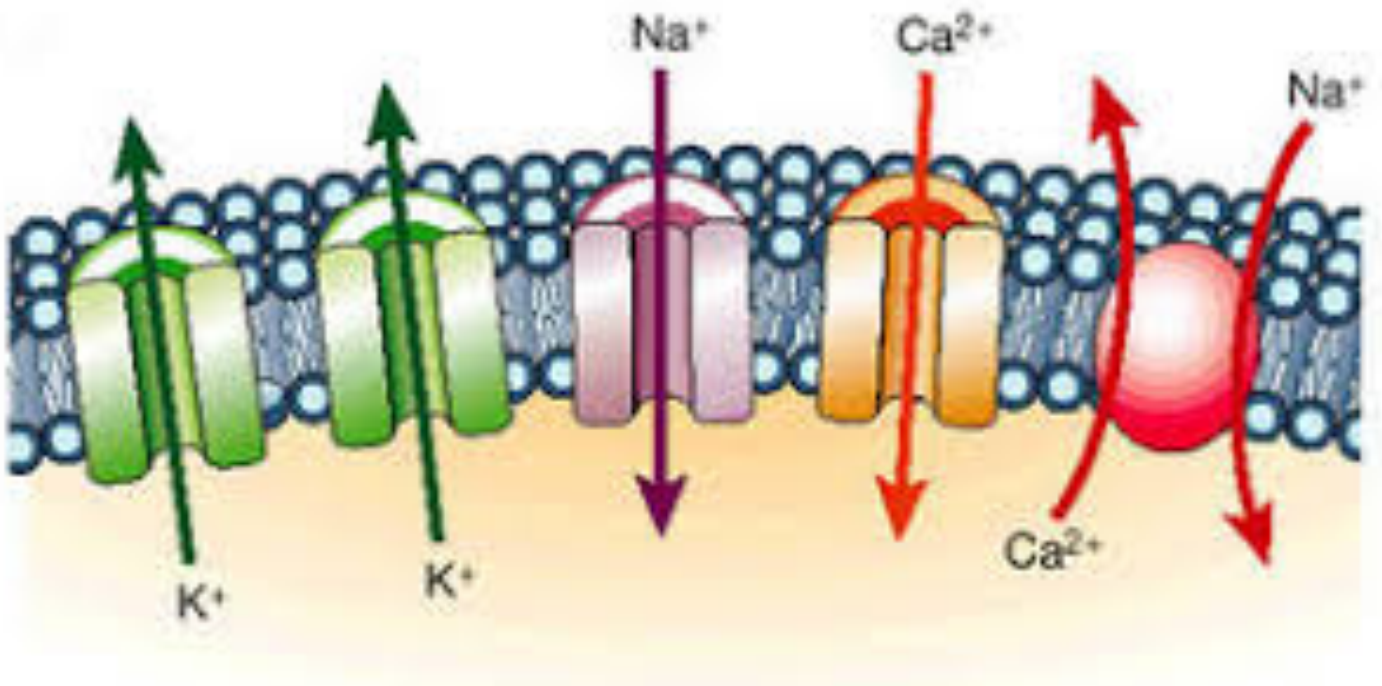
- Bouton
- Spine
- Cleft
- ***Synaptic weight and plasticity***

Neurotransmitters: many kinds

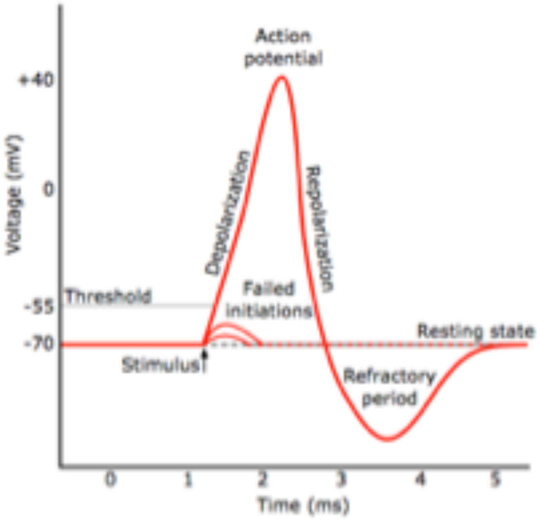
- **Amino acids**: glutamate, aspartate, D-serine, γ-aminobutyric acid (GABA), glycine
- **Gasotransmitters**: nitric oxide (NO), carbon monoxide (CO), hydrogen sulfide (H_2S)
- **Monoamines**: dopamine (DA), norepinephrine (noradrenaline; NE, NA), epinephrine (adrenaline), histamine, serotonin (SER, 5-HT)
- **Trace amines**: phenethylamine, *N*-methylphenethylamine, tyramine, 3-iodothyronamine, octopamine, tryptamine, etc.
- **Peptides**: somatostatin, substance P, cocaine and amphetamine regulated transcript, opioid peptides[11]
- **Purines**: adenosine triphosphate (ATP), adenosine
- Others: acetylcholine (ACh), anandamide, etc.

Ion channels

- Voltage gated, or
- Ligand gated (controlled by other neurotransmitters…)
- Glu lets X^+ in: excitatory
- GABA lets X^- in: inhibitory
- Ions define the membrane potential of the neuron

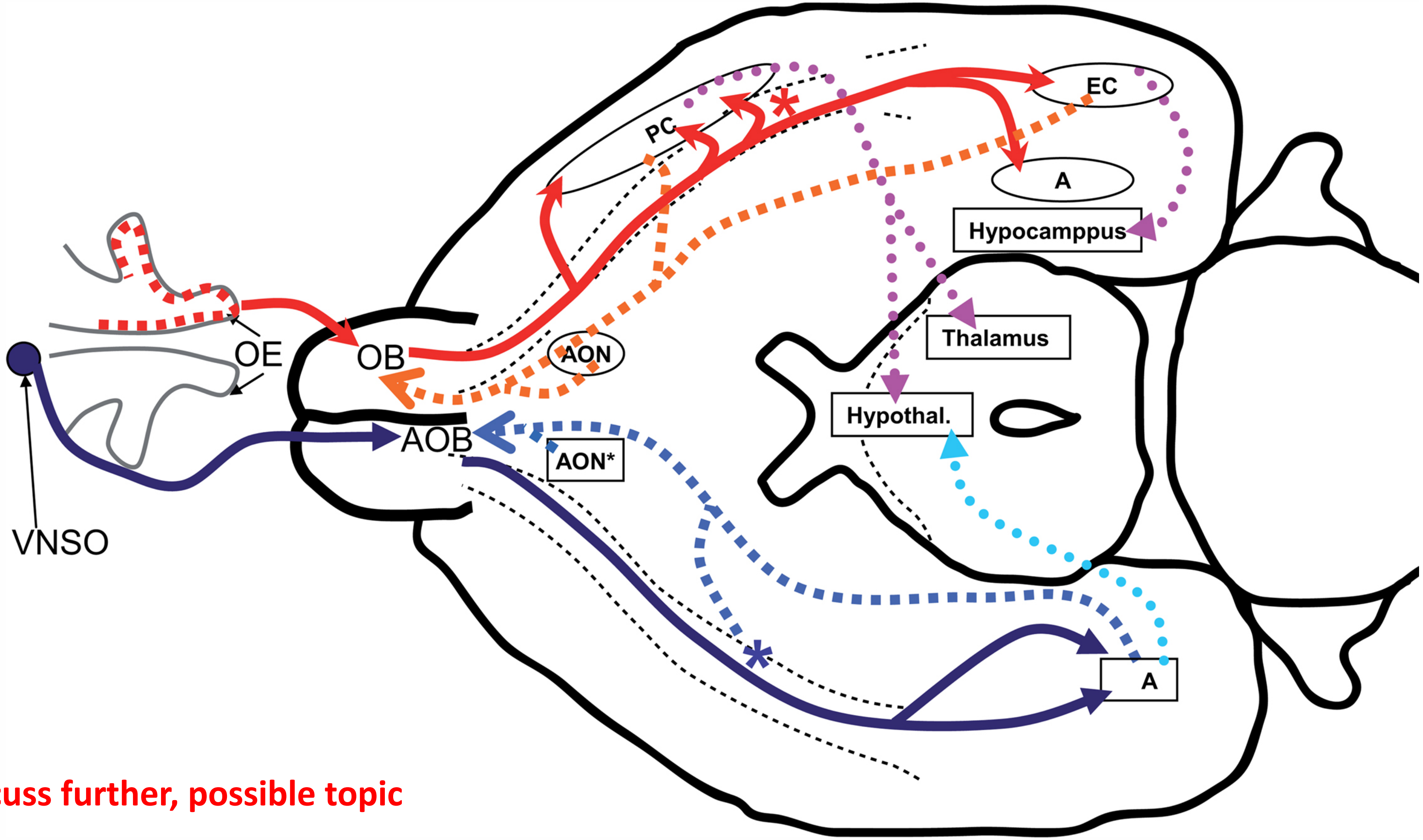


The spike, how?



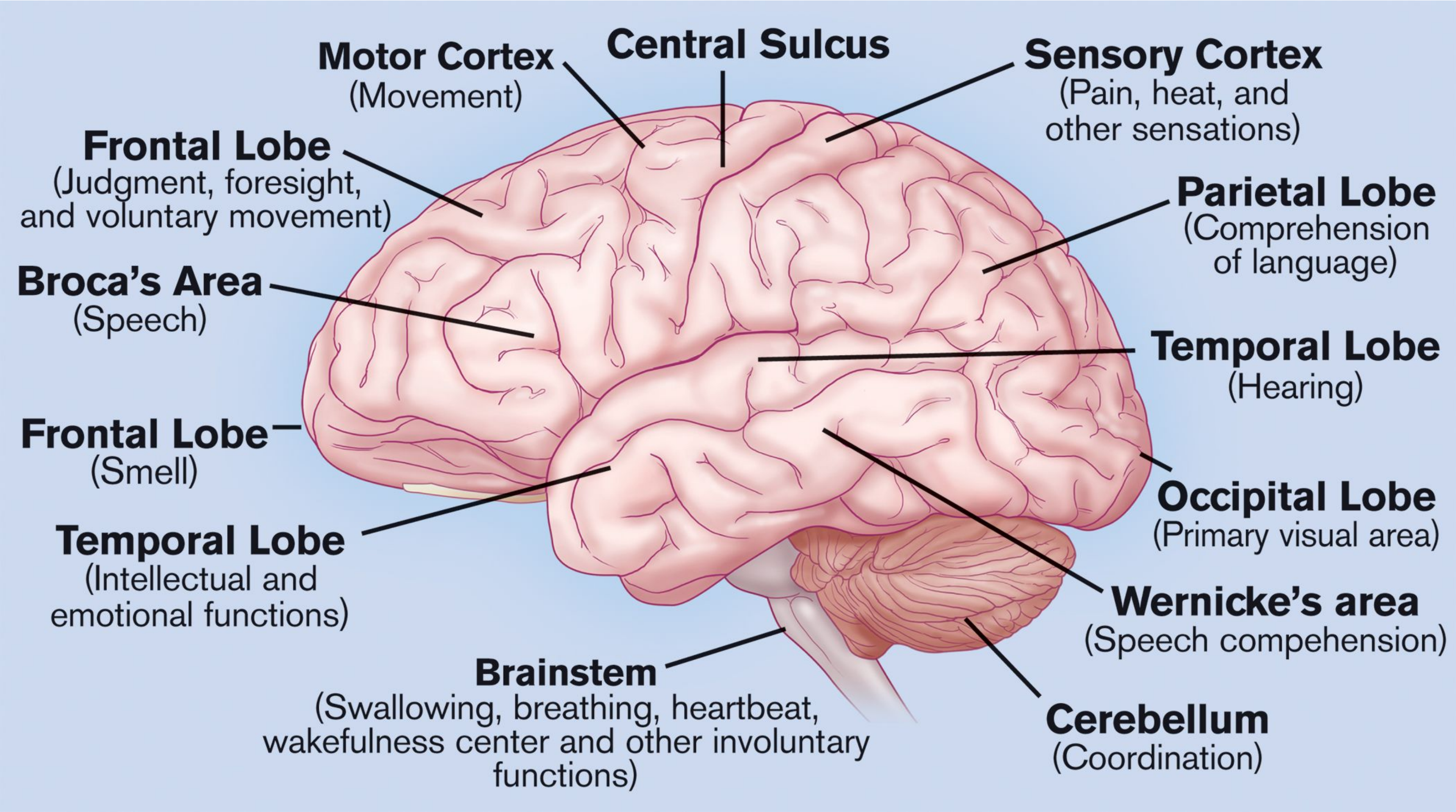
- Resting potential = K^+ balance
- Accumulation of excitatory input raises potential
- When threshold passed, voltage-gated influx of Na^+ : depolarization
- 1ms later, a voltage-gated massive outflux of K^+: back to resting potential and beyond: hyperpolarization
- Refractory period ~1ms

Olfaction*

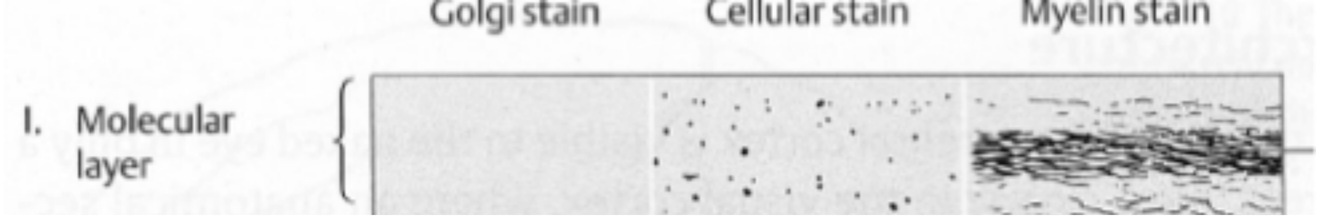


* Will not discuss further, possible topic

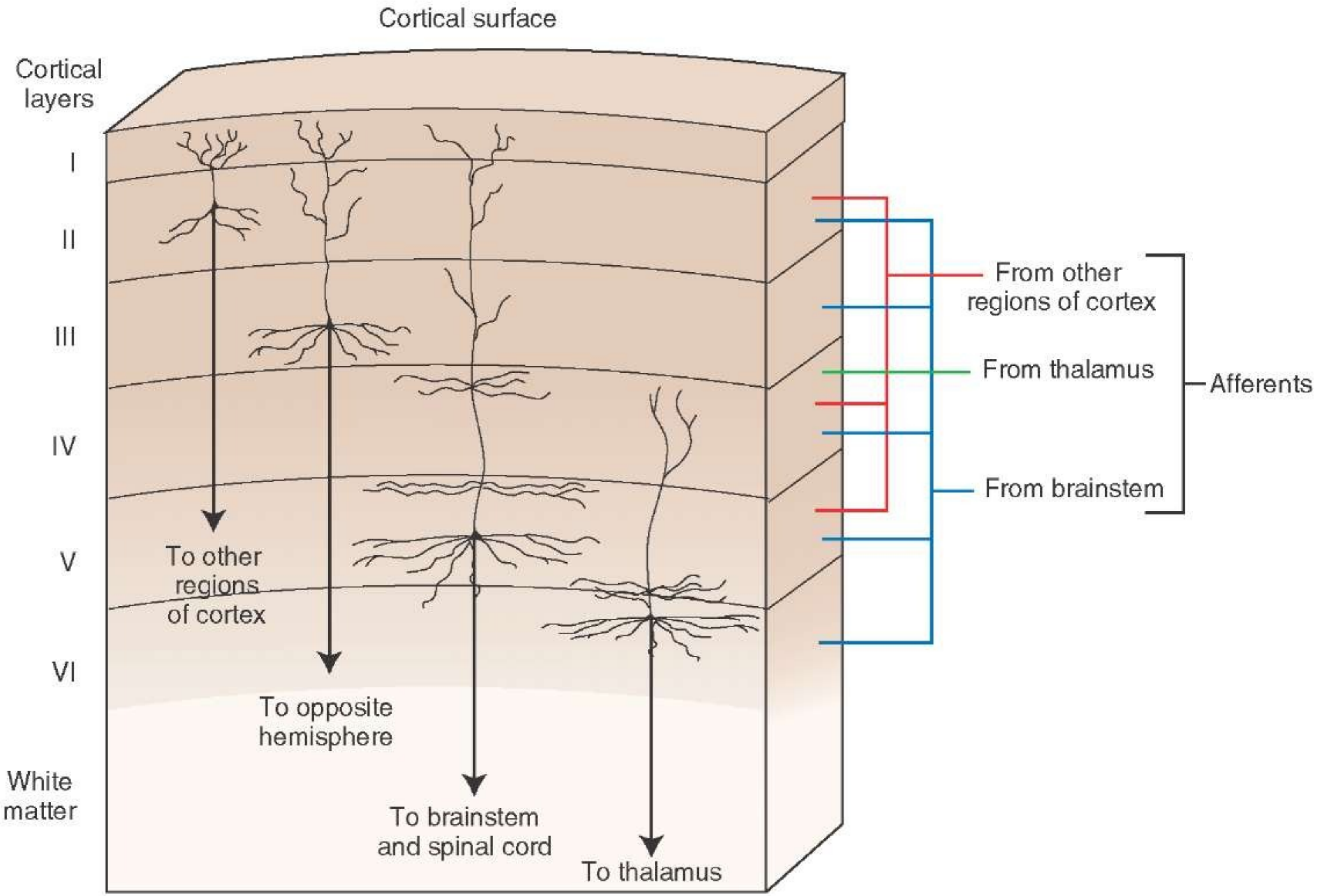
parts of the brain



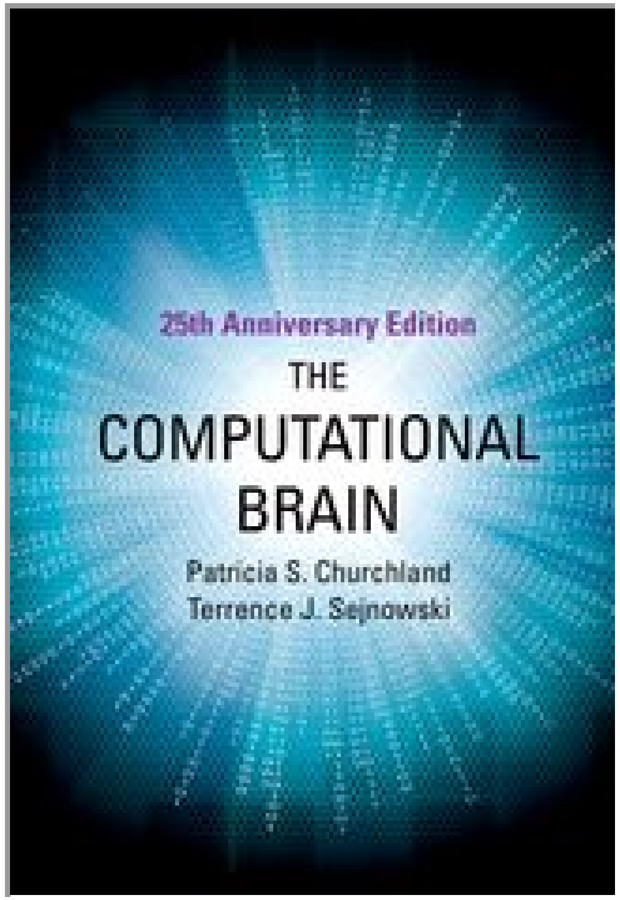
(Cerebral) Cortex: the engine

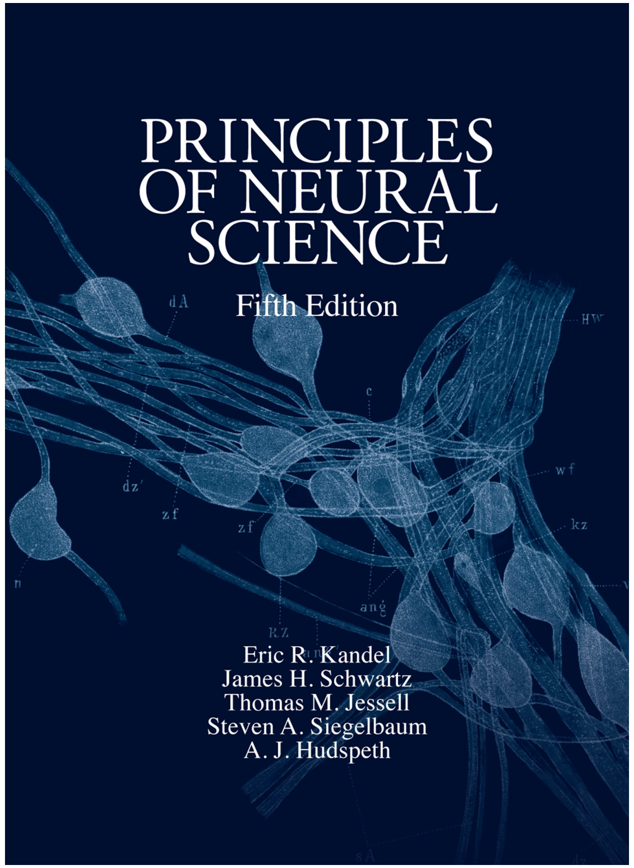


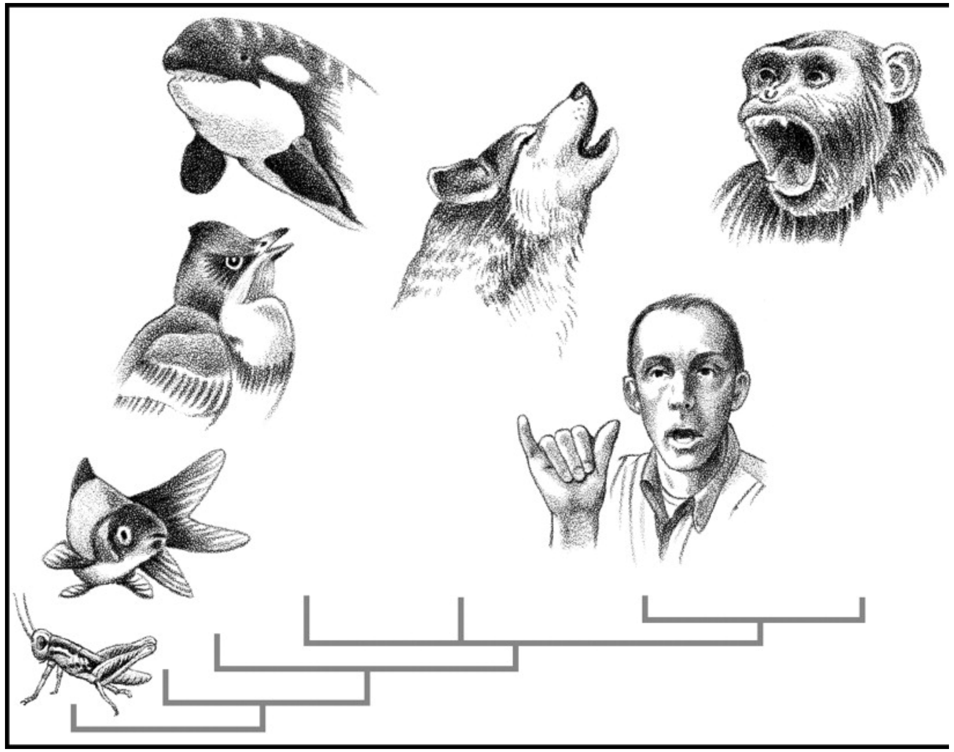
- 1.5 mm
- 2500 cm^2
- six layers
- 20B neurons



The readings







Some of your questions

- Can the brain be ”overclocked”?
- Does the brain function like a gradient descent?
- Are there any NN approaches that recycle an image through the network until reaching a confident result?
- What kinds of computational tasks are learned over generations by “fixed” circuits for each person, and what kinds are by individuals using flexible hardware?
- Language is fundamentally discrete; does a theory of cognition centered around language preclude a continuous representation of ontology? What is the role of recursion in ontological representations of thought?
- What’s tistory of “biolinguistics” and the ways in which the fields have worked (or failed to work) together.
- Is the fact that there are 10 times as many recurrent connections as there are forward connections something that we should be attempting to include in artificial neural networks? Are the recurrent connections used for learning synaptic weights in the brain? Can this be an argument for using additional nets for error propagation rather than backprop?
-

Some of your questions

- Is computational neuroscience about discovering how the human mind and body work, or for using this information to develop a better technology?
- Can natural language be fully replicated via computers and technology?
- If specific sensory nerves are injured, thereby silencing the corresponding part of the brain, then that “silenced” brain part may become responsive to different, uninjured nerves in nearby regions of the body. How is the new mapping decided?
- In the second reading, why are there much less signs of activity in the late bilingual image?
- Do we think that the “recursion” ability proposed by The Faculty of Language as unique to humans more likely to be the result of some communication processing component in our brains that other animals lack, or is it that thousands of years of communication have simply evolved this part of our brain to handle more
- I found the part about how the brain processes sign language being very similar to how it processes spoken language interesting, perhaps showing that it is more about the capacity to communicate than to speak
- Provided there is a way (that does not alter anything else within this species) to give this species vocal cords, could it learn to speak as well?
-

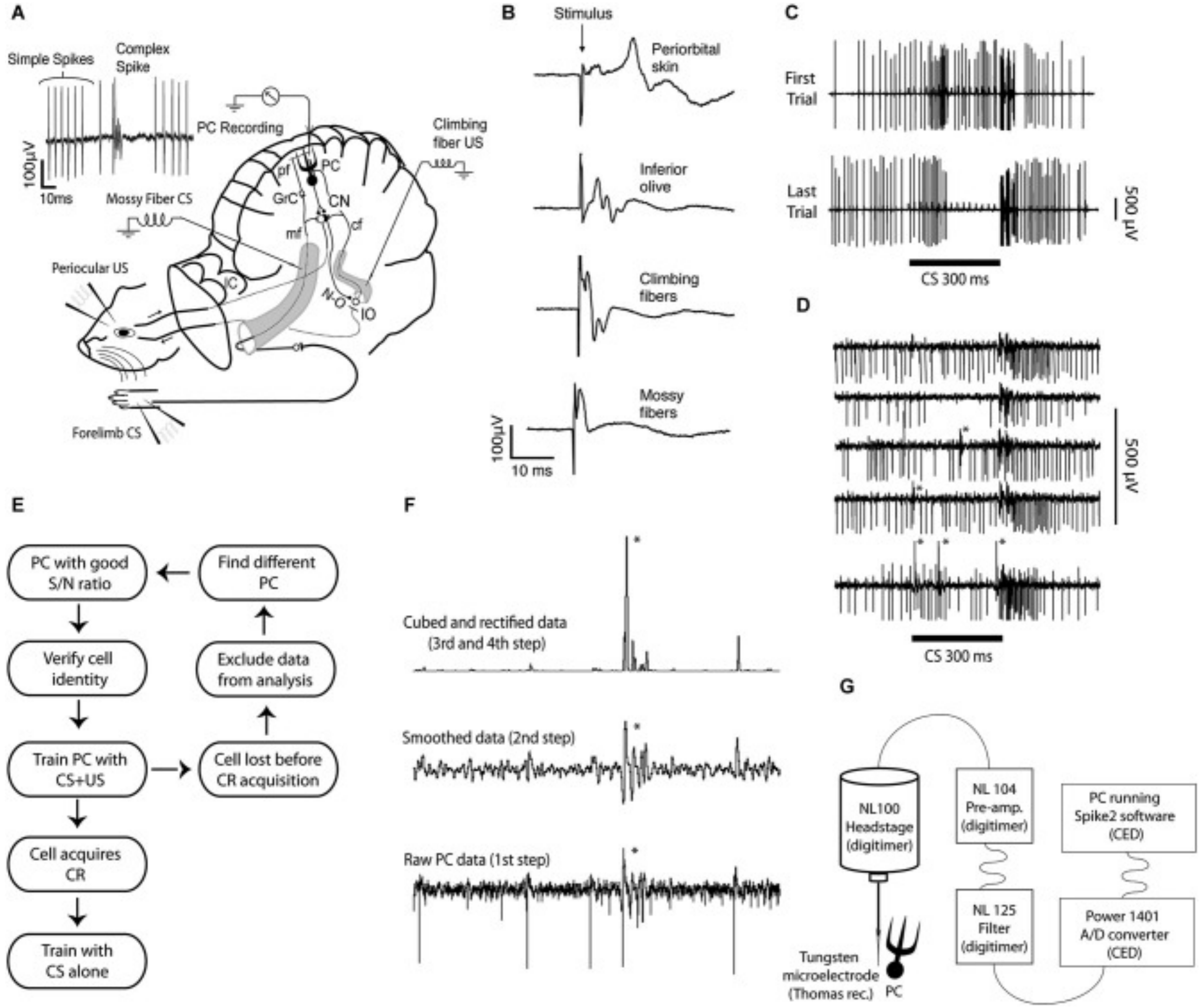
Some of your questions

- This is just a written thought – I am quite fascinated by what happens at the level of a chemical synapse. The length scale is so small that we don't have any pretty real image of what's there (as far as I searched around). The synaptic cleft is so tiny that only so few molecules fit in there, yet it does so amazing things on a human-size scale. This is just fascinating to me and I thought it's worth mentioning and perhaps talking about the level of accuracy and alignment needed for the ions channels to work properly, to successfully transmit the action potential.
- In studying epilepsy, Sperry, Gazzaniga, and Bogen discovered that in sectioning the corpus callosum, the two hemispheres were able to function independently of one another. If each hemisphere has its own "consciousness", how is our sense of self thus reconciled and accepted in an unsevered corpus callosum? ("The Brain and Behavior")
- In studying Neuroscience, it is noted that Neurobiologists concentrate on understanding the neural correlates of consciousness. With Neuroscience dealing with the mind and behavior, should we be wary of its ethical implications?

Our topics:

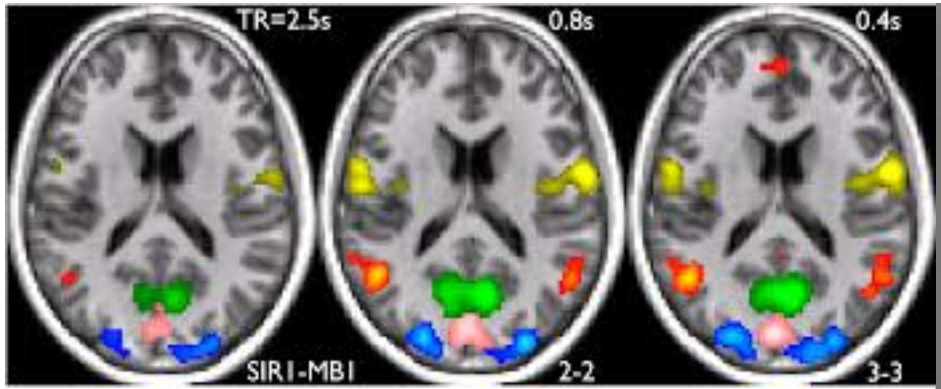
- ***Today (continued to next time):*** Vision, ANN, gradient descent, PCA
- Synapses and connectomics
- Learning and plasticity
- Motion and sensing
- Reinforcement learning
- Assemblies of neurons
- ***Language in the brain (~6 weeks)***
- On evolution and development
- ***Anything else?***

how do we record from the brain?
Single unit: ~ 1 neuron, ~ 1ms

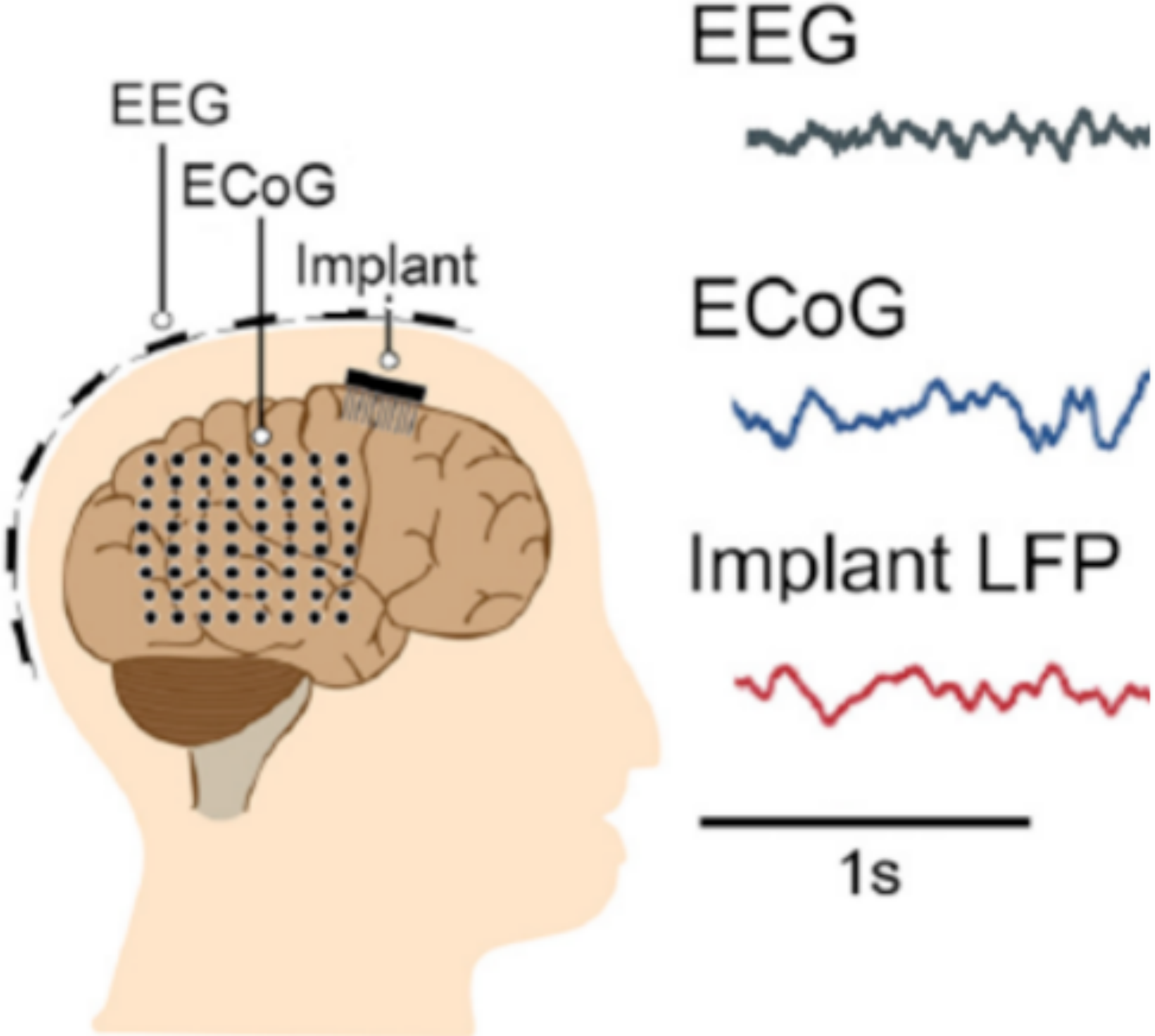


how do we record from the brain?
EEG: 1cm, 1ms

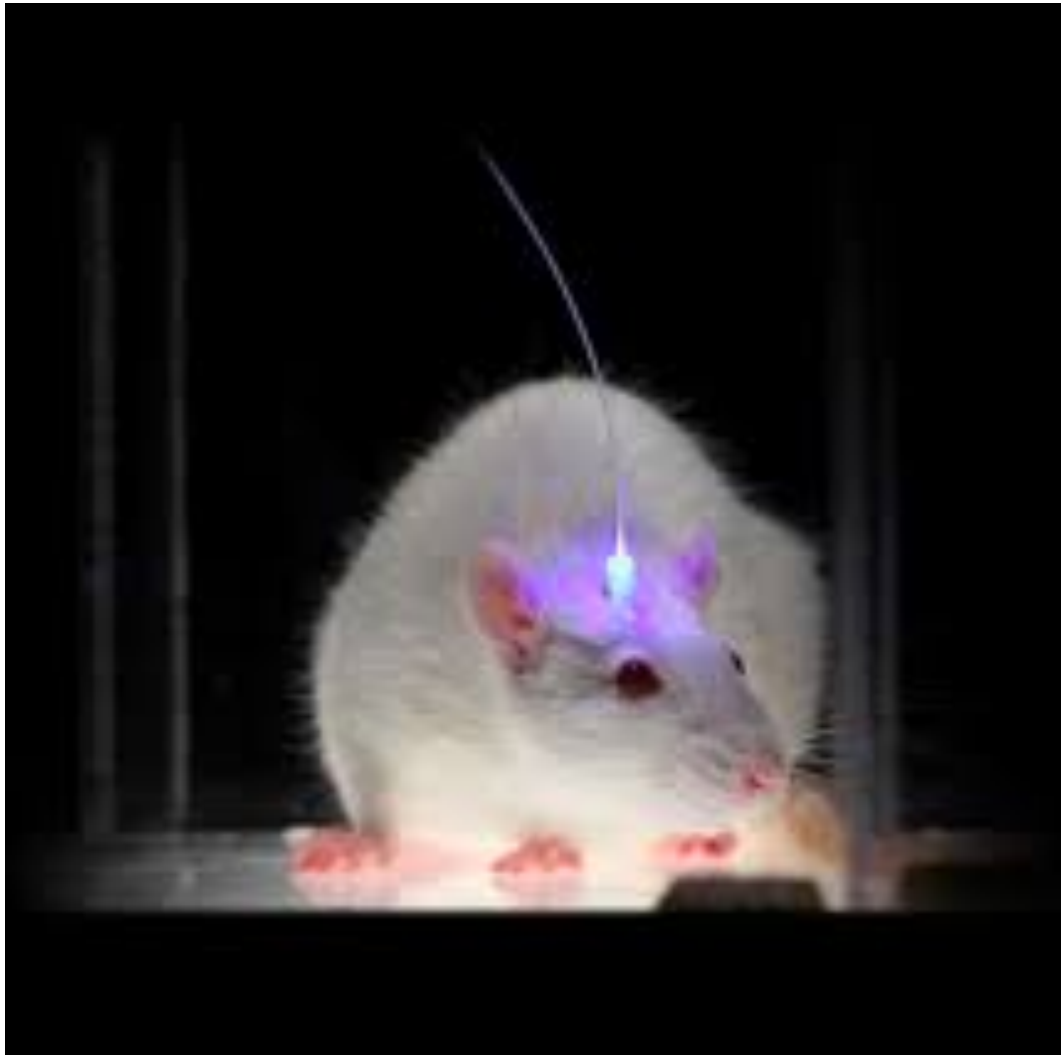
how do we record from the brain?
fMRI: ~ 1mm, ~ 1s (+ ML!)



how do we record from the brain?
ECoG



Optogenetics – Ca^{+} imaging



Recall the Neuro Nobel Prizes

- 1906: Ramon y Cajal and Golgi: ***the neuron***
- 1932: C. Sherrington: ***the synapse***
- 1936: H. Dale and O. Loewi: ***neurotransmitters***
- 1963: Eccles, Hodgkin, Huxley: ***the spike***
- 1967 Granit, Hartline, and Wald; 1981 Sperry, Hubel, Wiesel: ***vision***
- 1970: Katz, von Euler, Axelrod: ***synaptic vesicles***
- 1991: Neher,Sakmann; 2000: Carlsson, Greengard, Kandel: ***ion channels***
- 2004: R. Axel and L. Buck: ***olfaction***
- 2014: O'Keefe, M. Moser, and E. Moser: ***place cells and grid cells***

Place cells

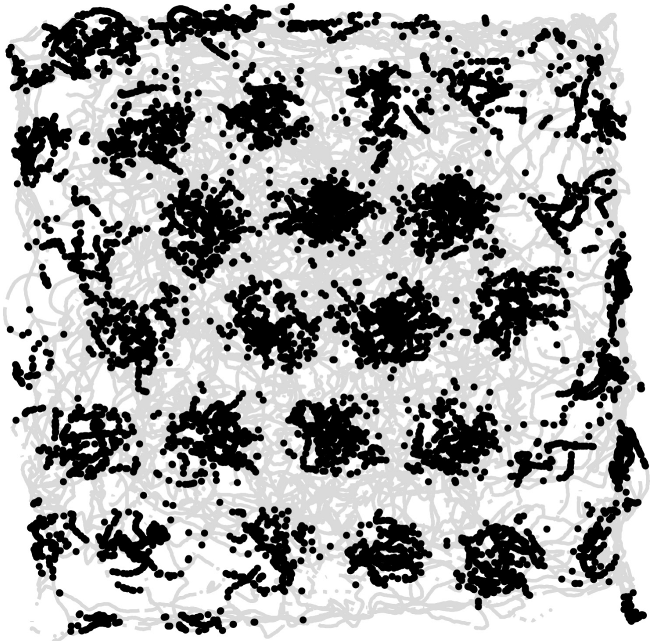
- Tuning curve of a place cell:

- A neuron in the hippocampus that fires when the animal is in a particular place

- https://www.youtube.com/watch?v=STyd1qJr3yM

Grid cells

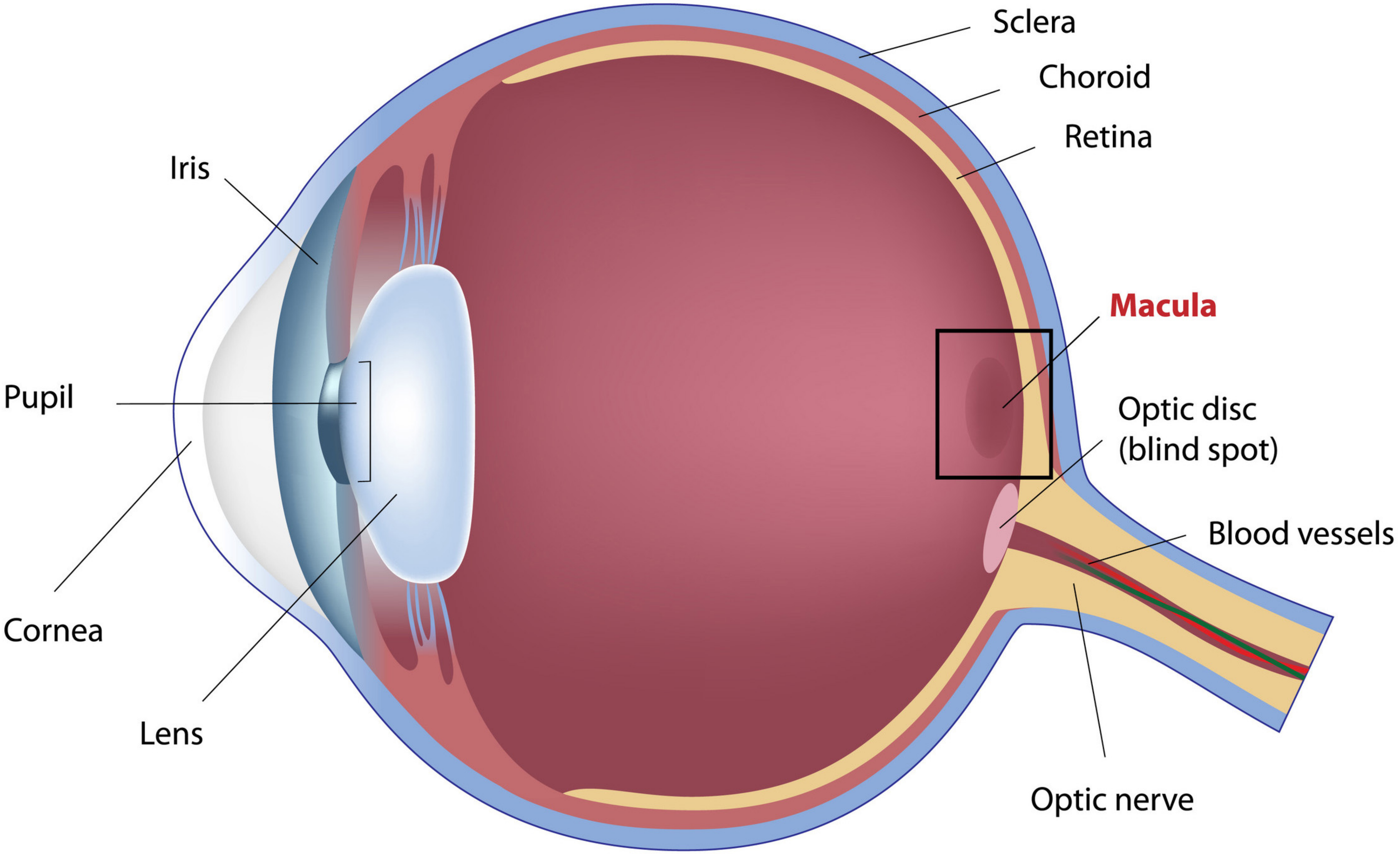
- Tuning curve of a grid cell

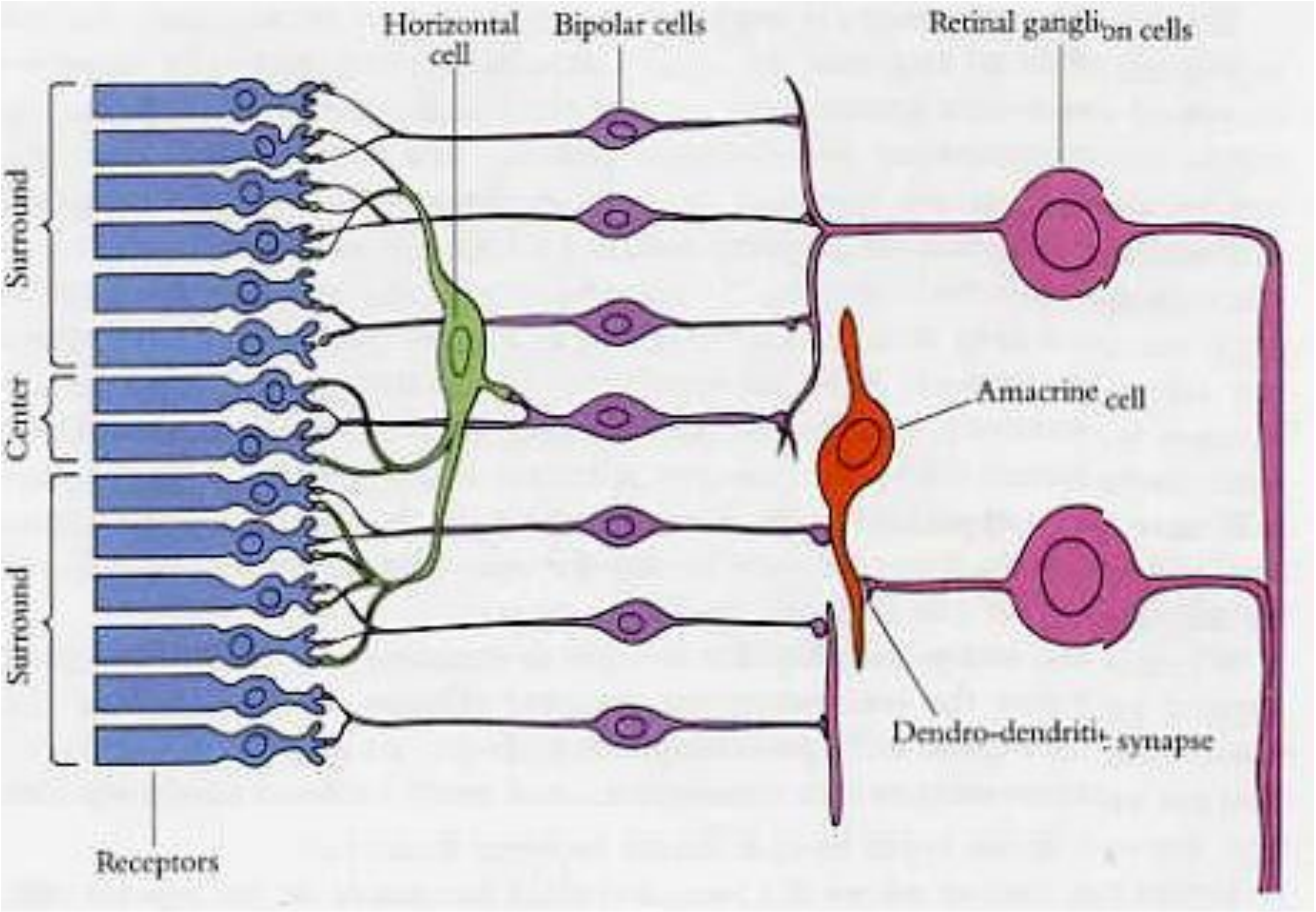


- A neuron in the entorhinal cortex that fires when the animal is in any one of the nodes of a triangular grid
- The movie:
- https://www.youtube.com/watch?v=i9GiLBXWAHIons

Next: Vision

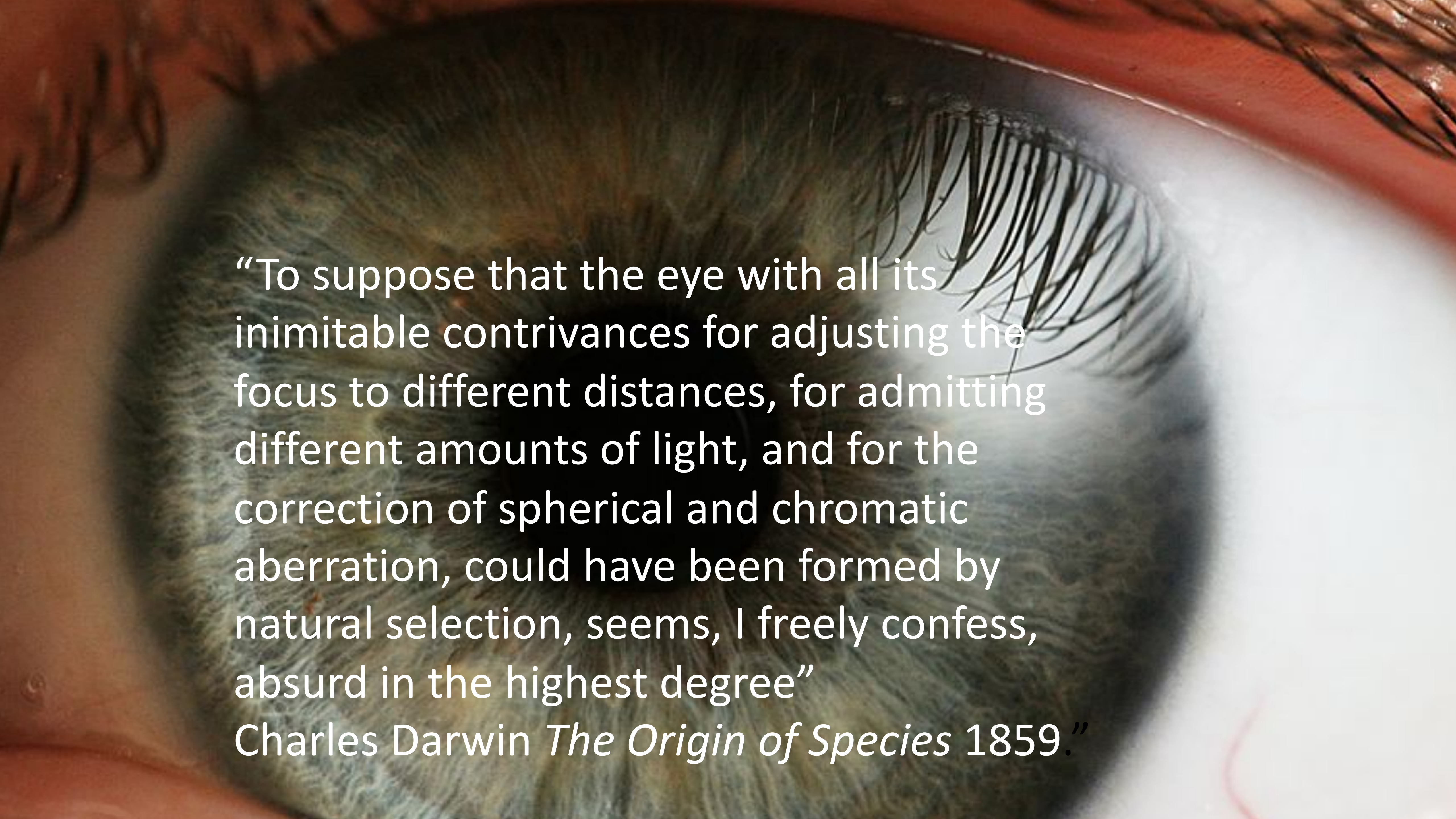
- The retina circuit
- The visual cortex
- The Hubel - Wiesel cat
- Gradient descent and convolutional neural networks
- Supervised vs unsupervised learning, RP and PCA
- The Olshausen-Field paper


Sclera
Choroid
Retina
Iris
Macula
Pupil
Optic disc
(blind spot)
Blood vessels
Cornea
Lens
Optic nerve


Horizontal cell
Bipolar cells
Retinal ganglion cells
Surround
Center
Surround
Amacrine cell
Dendro-dendritic synapse
Receptors

How do we know all this?

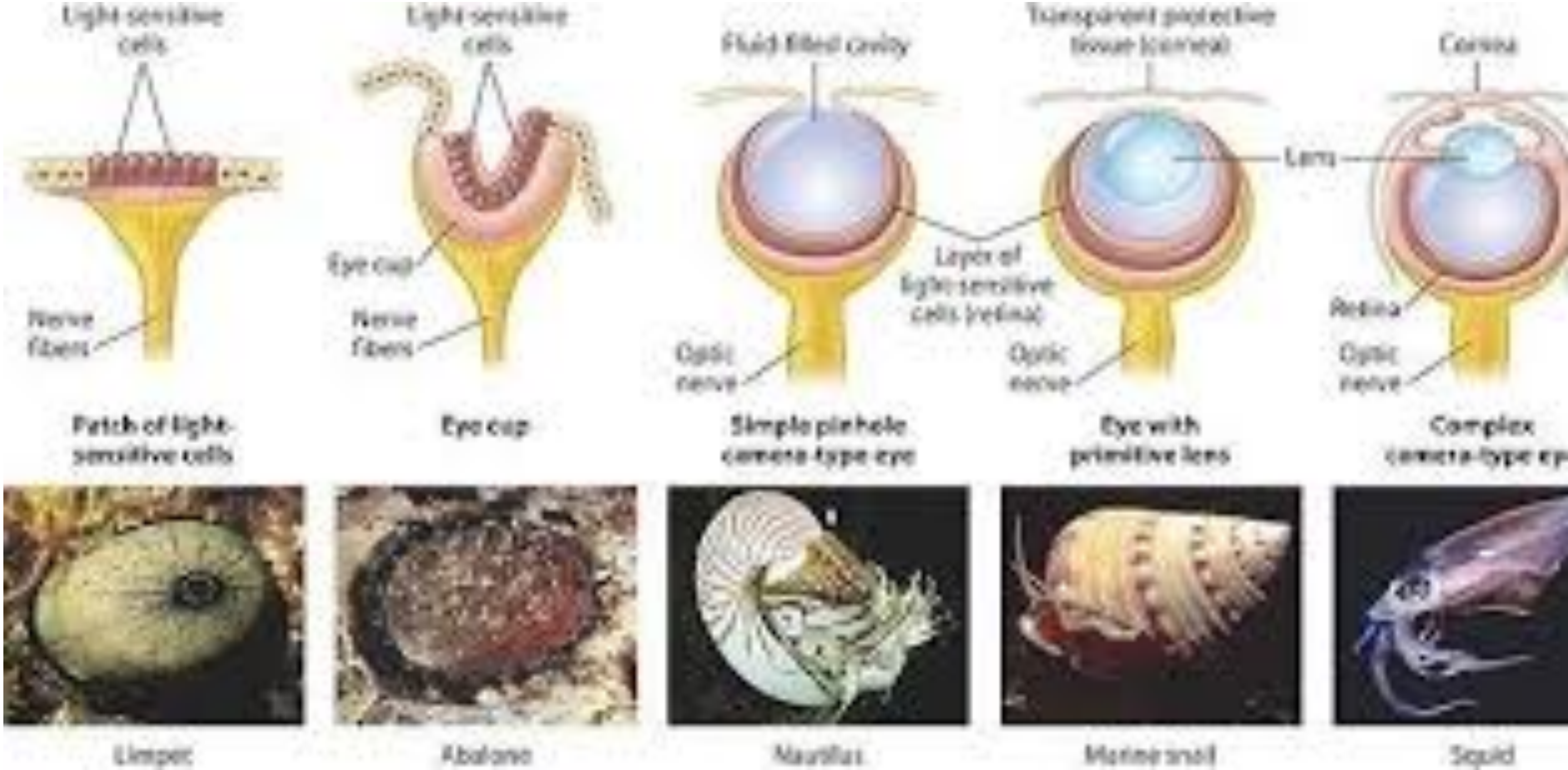
- 1967 Nobel prize for the physiology and chemistry of vision: George Wald, Ragnar Granit, and Haldan Keffer Hartline



“To suppose that the eye with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest degree”

Charles Darwin *The Origin of Species* 1859.”

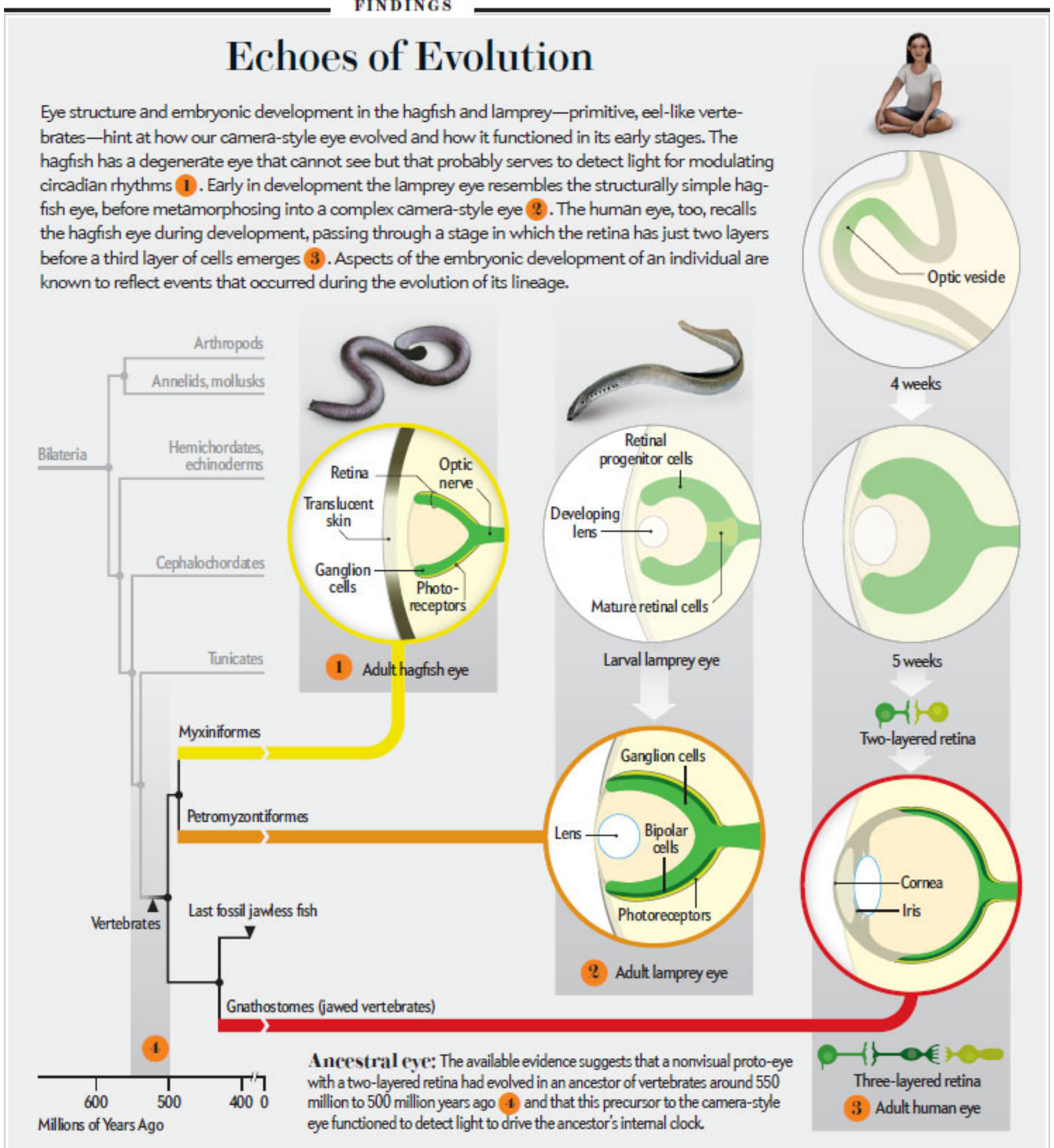
Absurd things do happen…



Early Cambrian (530 Mya): vision + locomotion = acceleration

Many independent discoveries/early branches

One of them is the vertebrate eye…

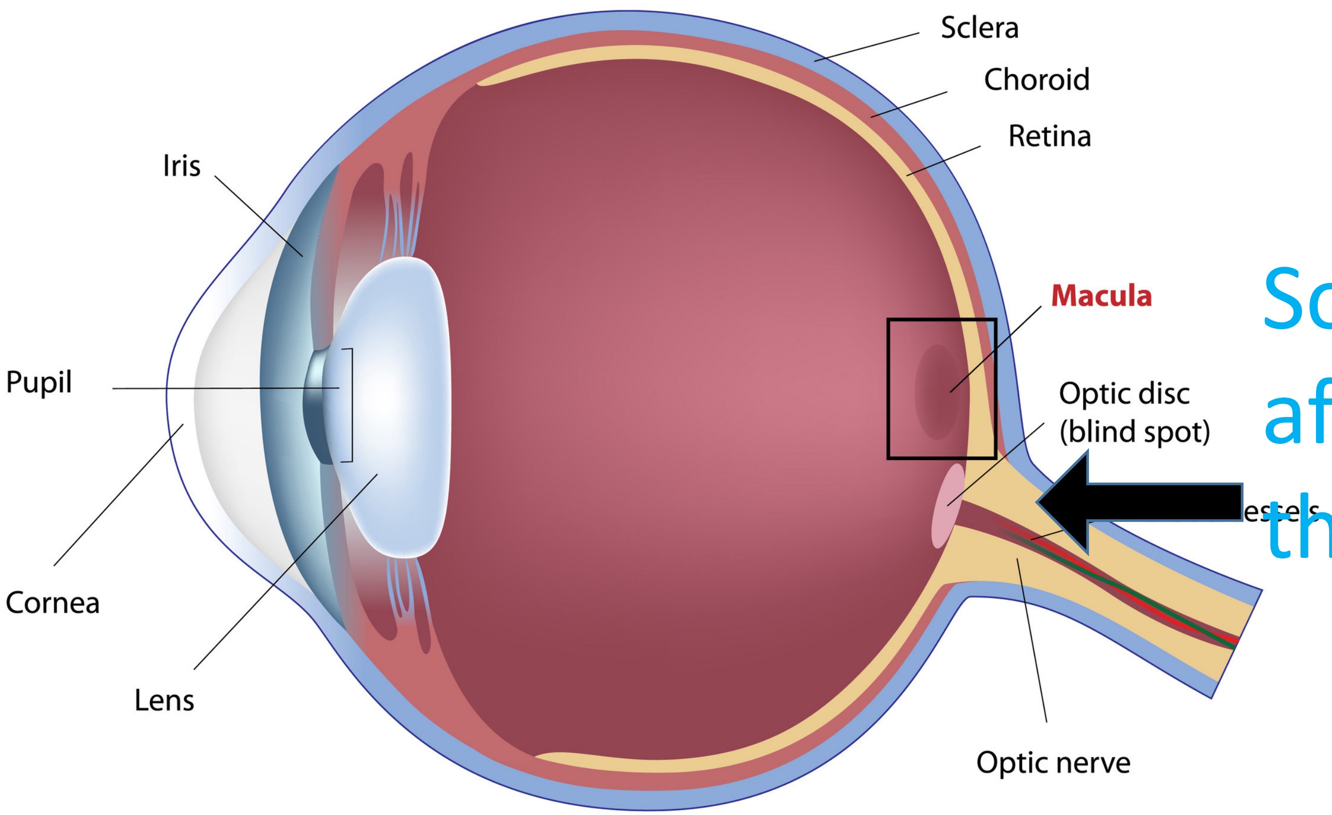


FINDINGS

Echoes of Evolution

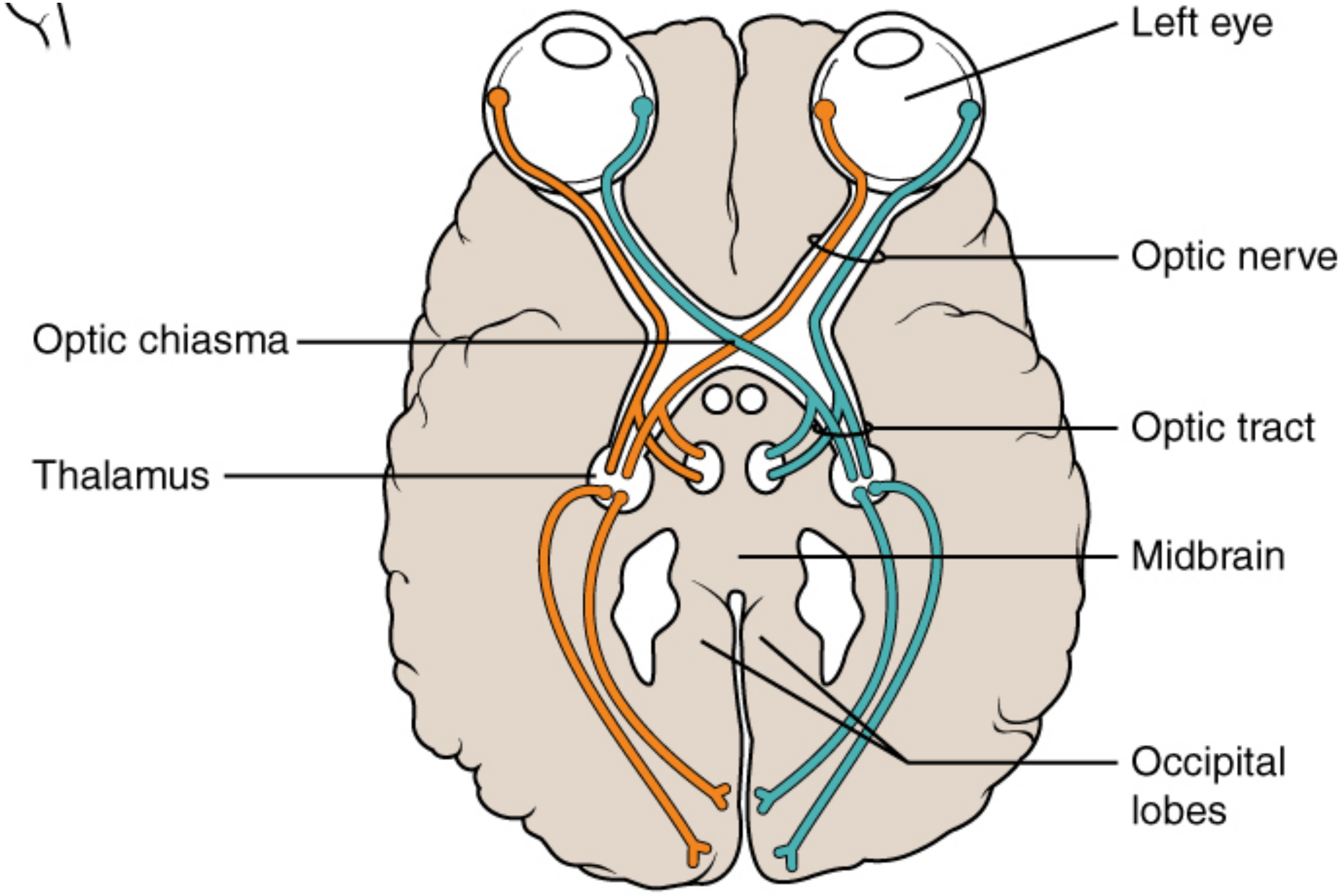
Eye structure and embryonic development in the hagfish and lamprey—primitive, eel-like vertebrates—hint at how our camera-style eye evolved and how it functioned in its early stages. The hagfish has a degenerate eye that cannot see but that probably serves to detect light for modulating circadian rhythms 1. Early in development the lamprey eye resembles the structurally simple hagfish eye, before metamorphosing into a complex camera-style eye 2. The human eye, too, recalls the hagfish eye during development, passing through a stage in which the retina has just two layers before a third layer of cells emerges 3. Aspects of the embryonic development of an individual are known to reflect events that occurred during the evolution of its lineage.

Ancestral eye: The available evidence suggests that a nonvisual proto-eye with a two-layered retina had evolved in an ancestor of vertebrates around 550 million to 500 million years ago 4 and that this precursor to the camera-style eye functioned to detect light to drive the ancestor's internal clock.



So, what happens after
the retinal ganglia?

The chiasm


Left eye
Optic nerve
Optic chiasma
Optic tract
Thalamus
Midbrain
Occipital
lobes

The archetypical deep net

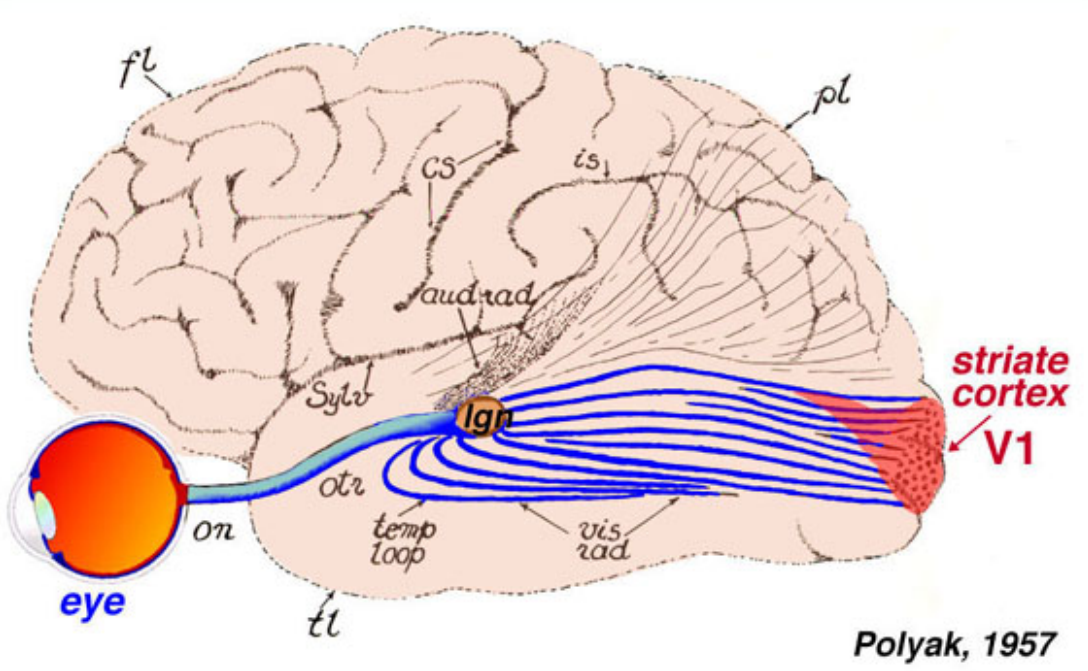
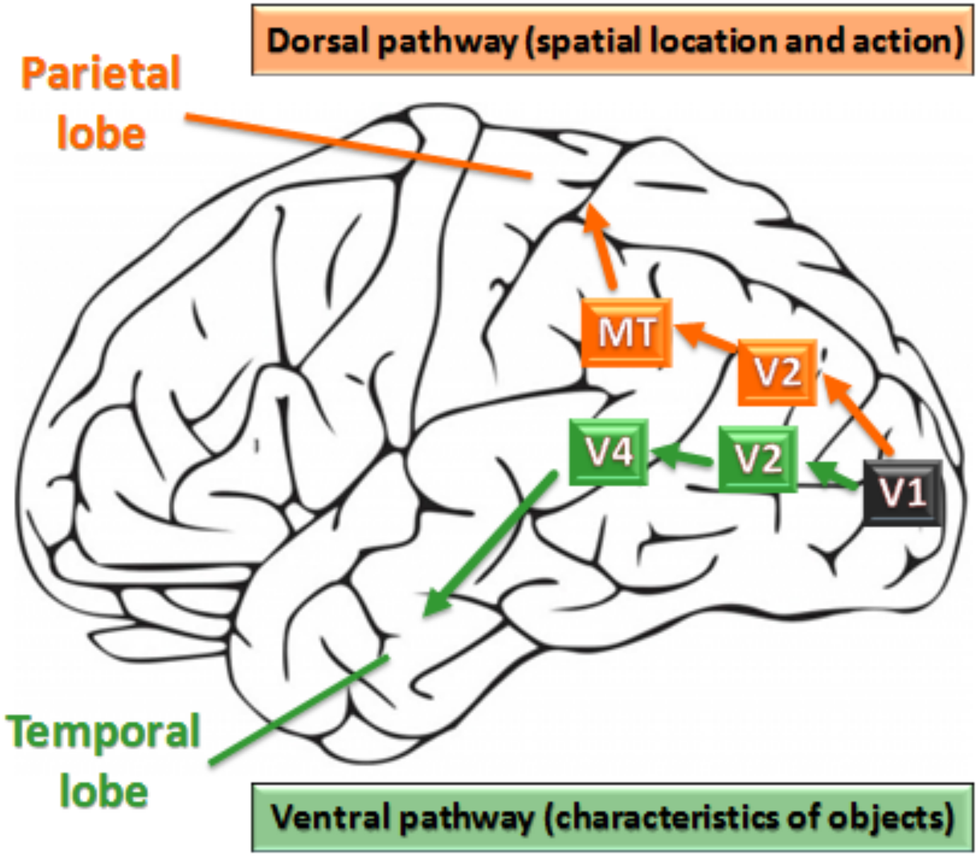
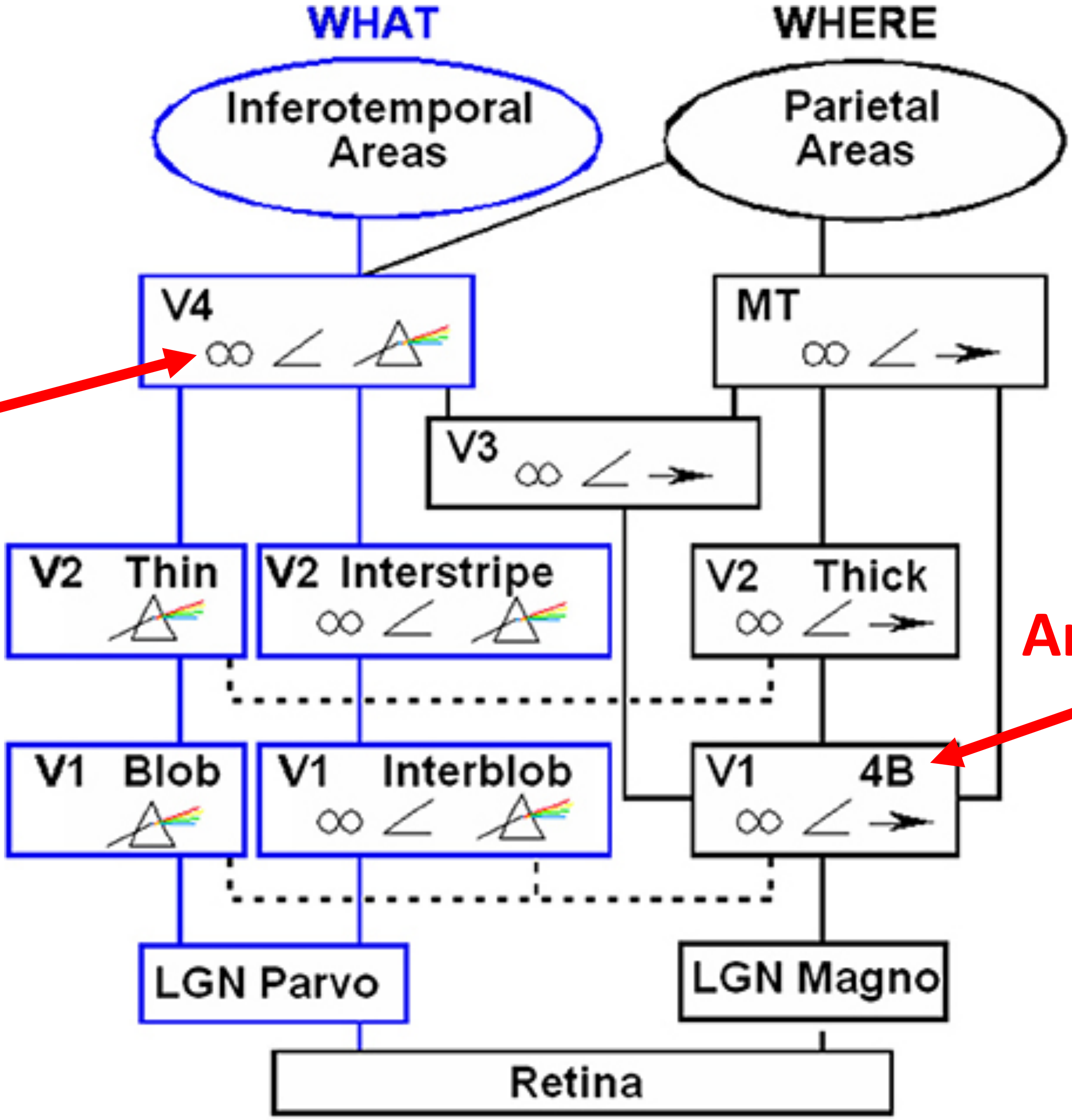


Figure 8. Visual input to the brain goes from eye to LGN and then to primary visual cortex, or area V1, which is located in the posterior of the occipital lobe.
Adapted from Polyak (1957).



The paths



Symbols for:
Binocular
Angle
Color
Motion

Areas within V1

Total time for the path: 0.15s

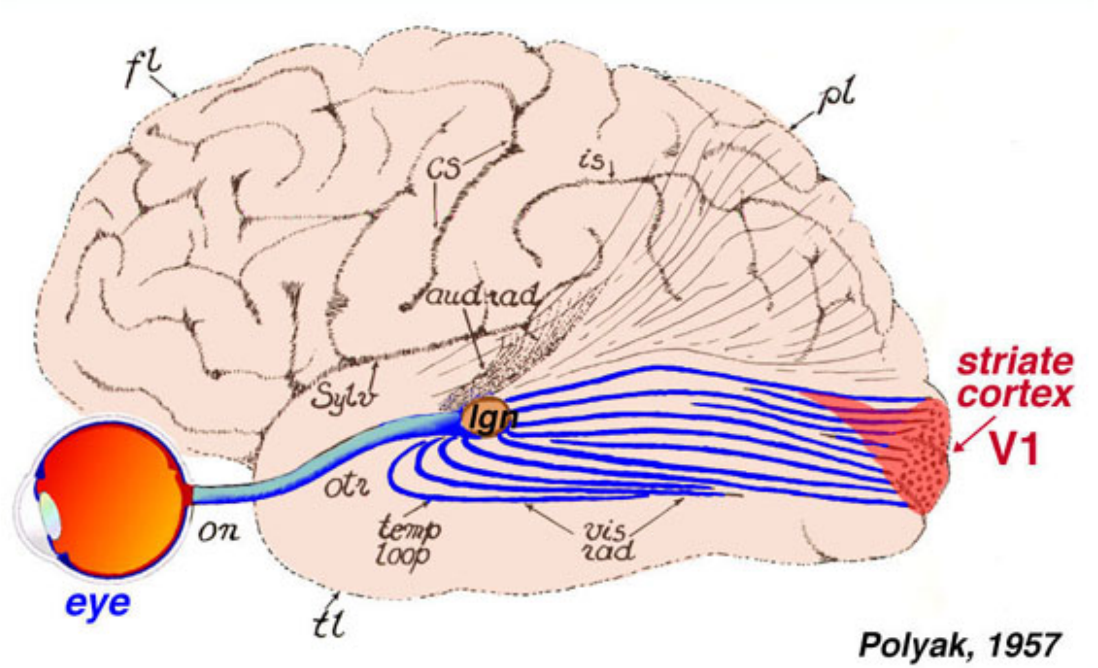
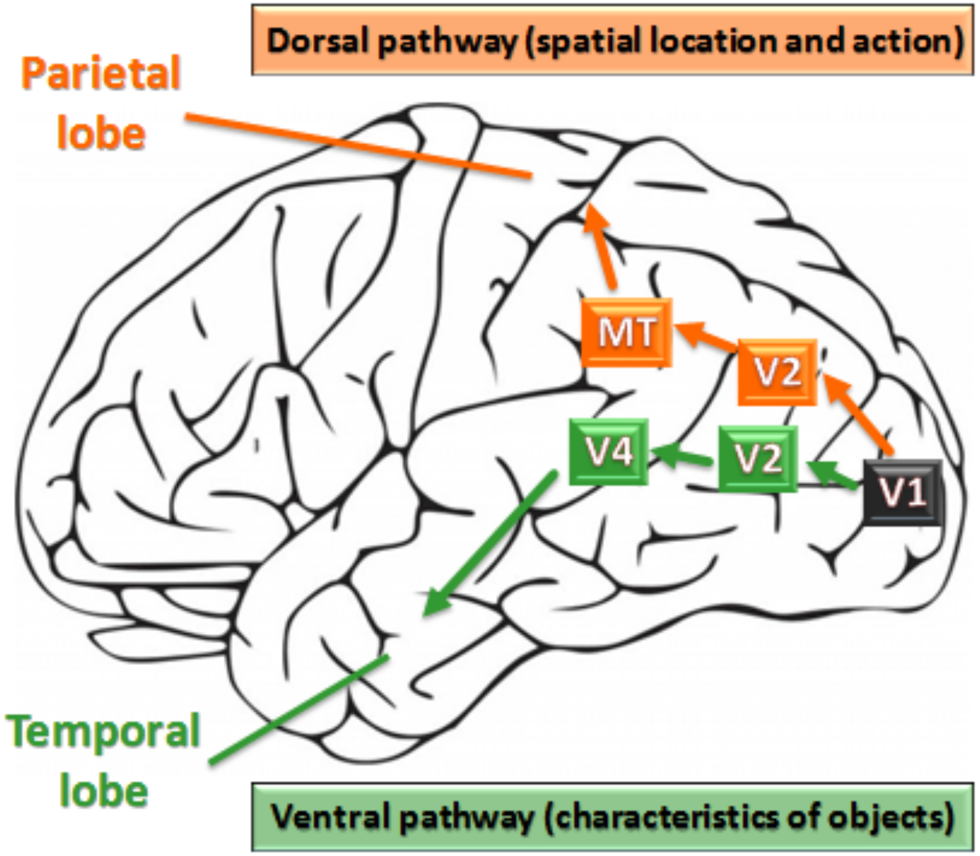


Figure 8. Visual input to the brain goes from eye to LGN and then to primary visual cortex, or area V1, which is located in the posterior of the occipital lobe. Adapted from Polyak (1957).



Total time for the path: 150 ms

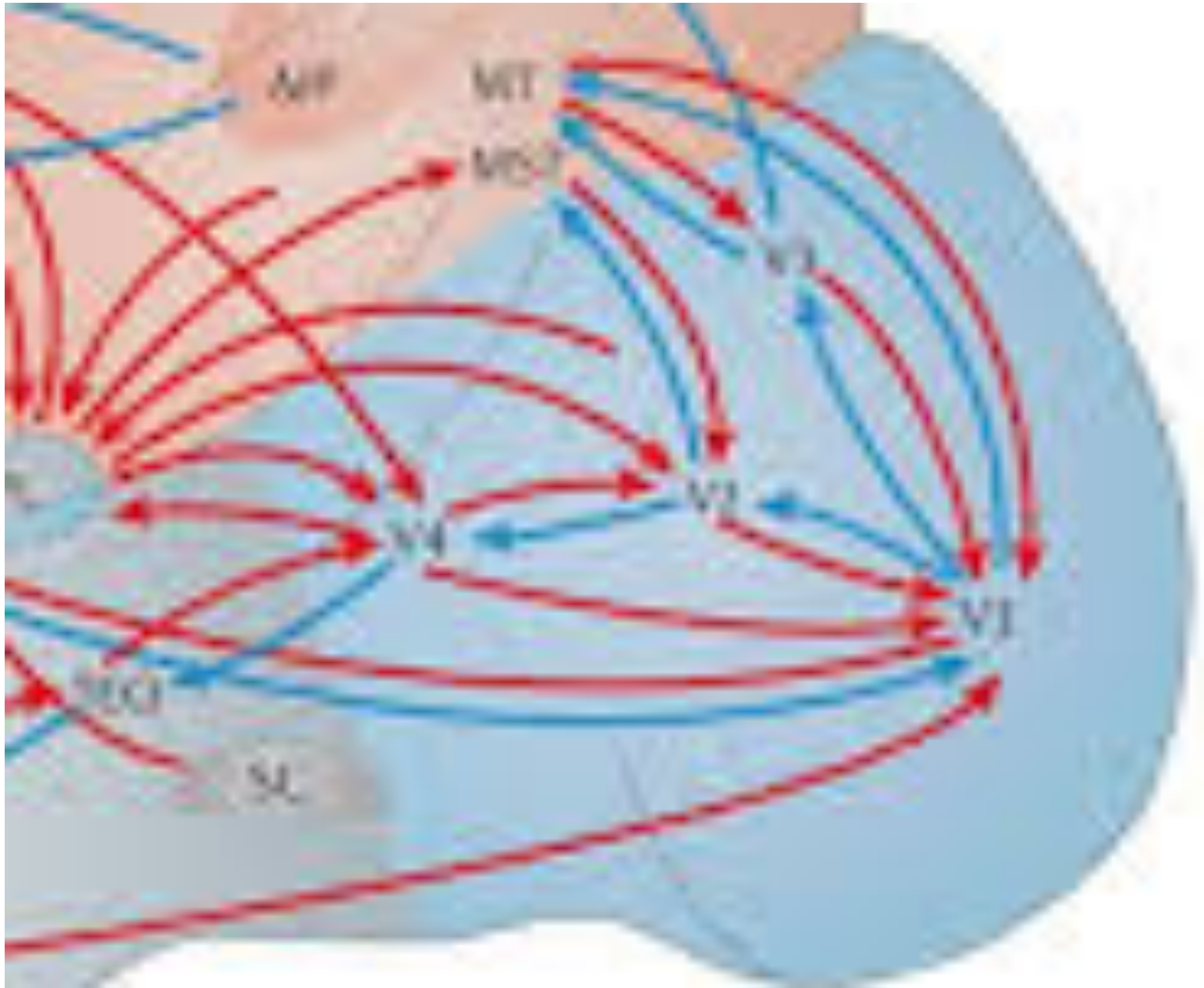
- Q: Too fast? Too slow?
- What is the depth of the network?
- Total axon distance traveled ~10ms
- Each neuron firing every 20ms
- Human reaction to visual stimulus: 250ms
- Audio: 170ms. Touch: 150ms
- btw, Break: 1,500 ms

Not the Thinker!

- https://www.dailymotion.com/video/x278zke

Back to the path: not so simple…

There are *many*
feedback axons!
(see Jeff Hawkins
“On Intelligence”)



The path: not so simple….
All upward links also go through the thalamus!
(see Simons talk by Murray Sherman)

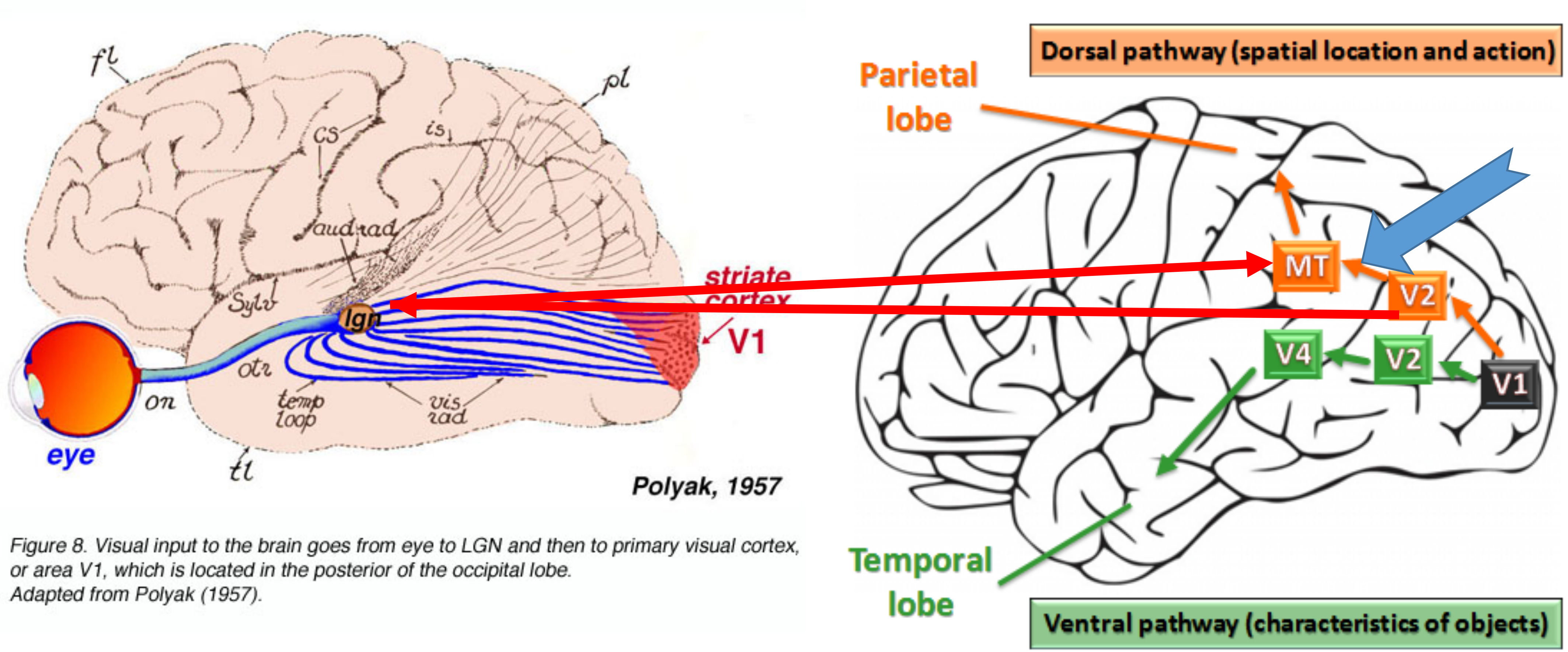
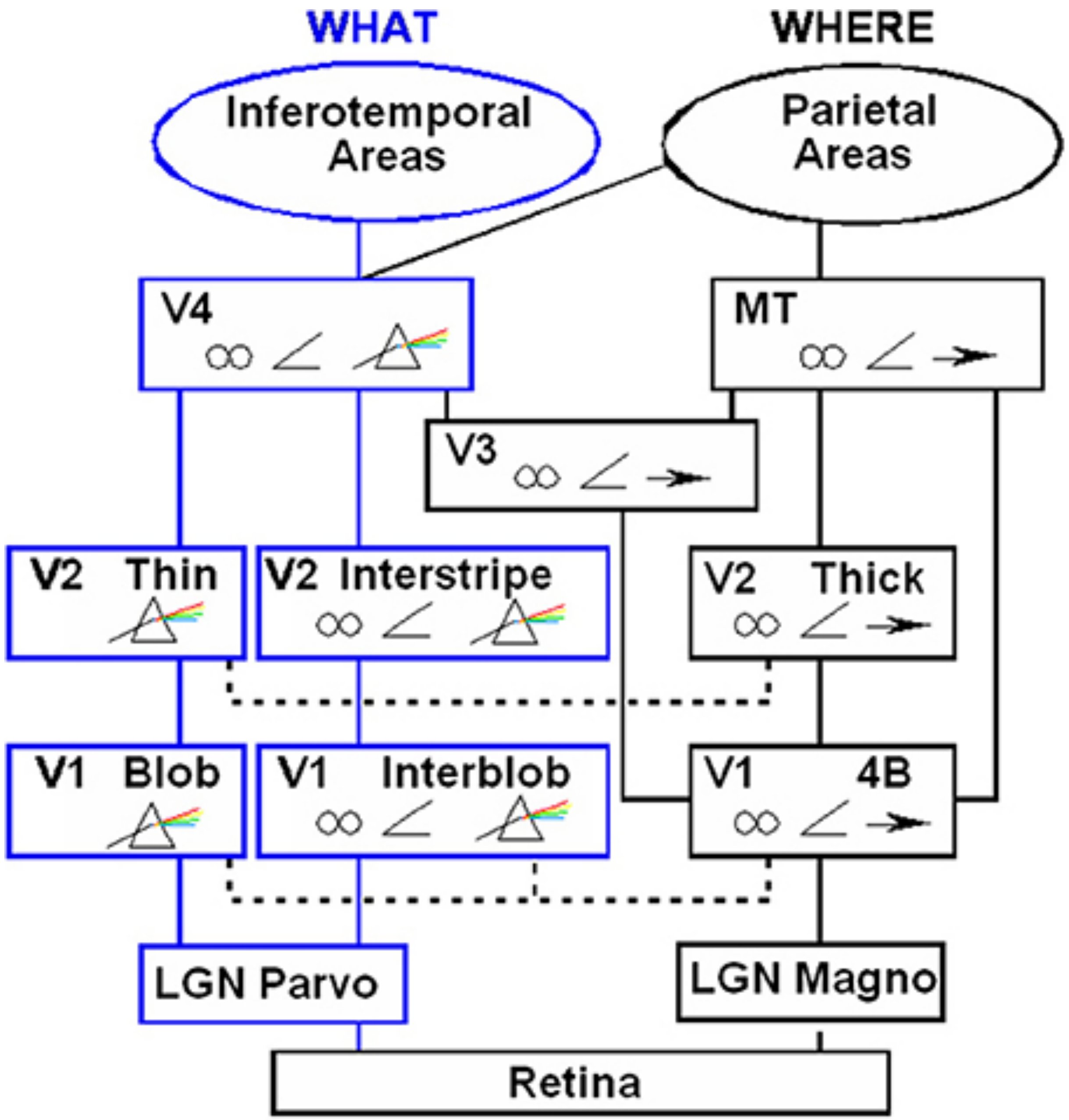


Figure 8. Visual input to the brain goes from eye to LGN and then to primary visual cortex, or area V1, which is located in the posterior of the occipital lobe. Adapted from Polyak (1957).

But what does this circuit compute?



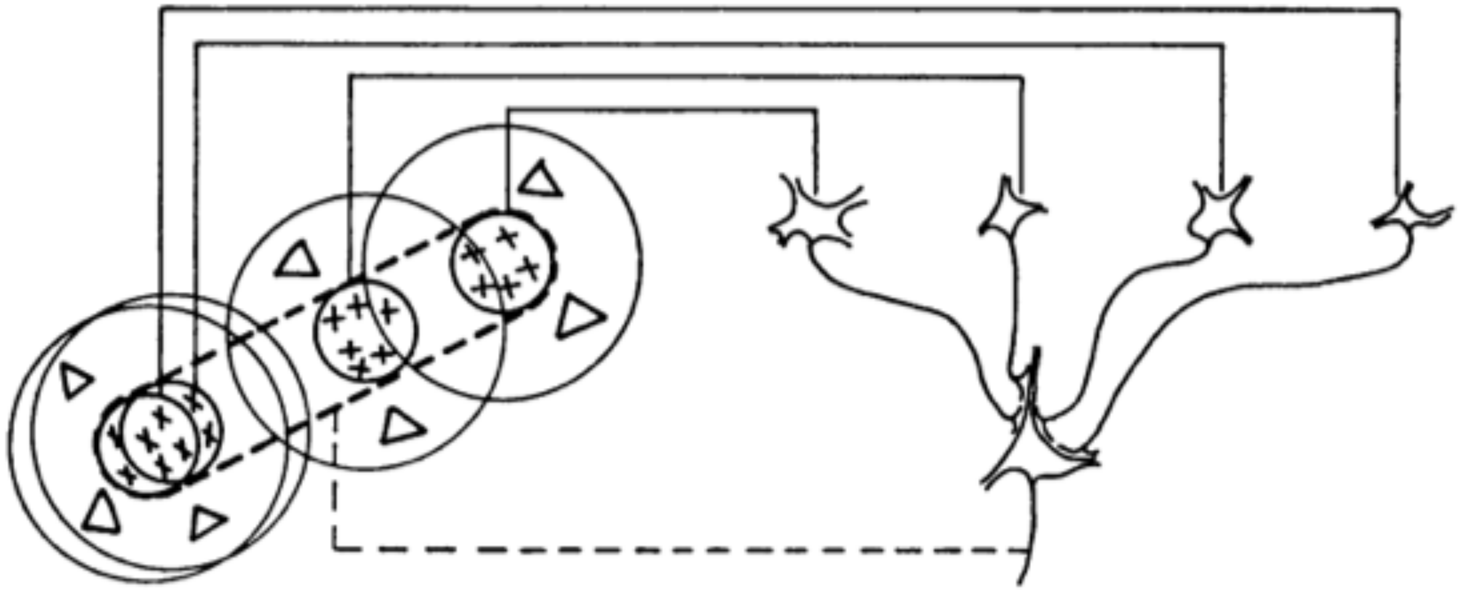
The Hubel and Wiesel experiment (1959)

- Recorded from neurons in the visual cortex (and LGN) of an immobilized, awake, sedated, locally anesthetized cat
- Showed various visual stimuli
- Changed a stimulus parameter, such as angle or horizontal offset
- Calculated firing rate, ultimately the receptive field, of each neuron (~ the max of the tuning curve) as a function of the parameter(s)
- https://www.youtube.com/watch?v=8VdFf3egwfg

What Hubel and Wiesel did

- Pioneered the experimental methodology (1959)
- Receptive fields beyond the retina (cf. Hartline 1940)
- The importance of edge detection (1959)
- Reflected on the experiment, conjectured how edges can be detected by a hierarchical neural circuit (1962)
- Nobel (1980)

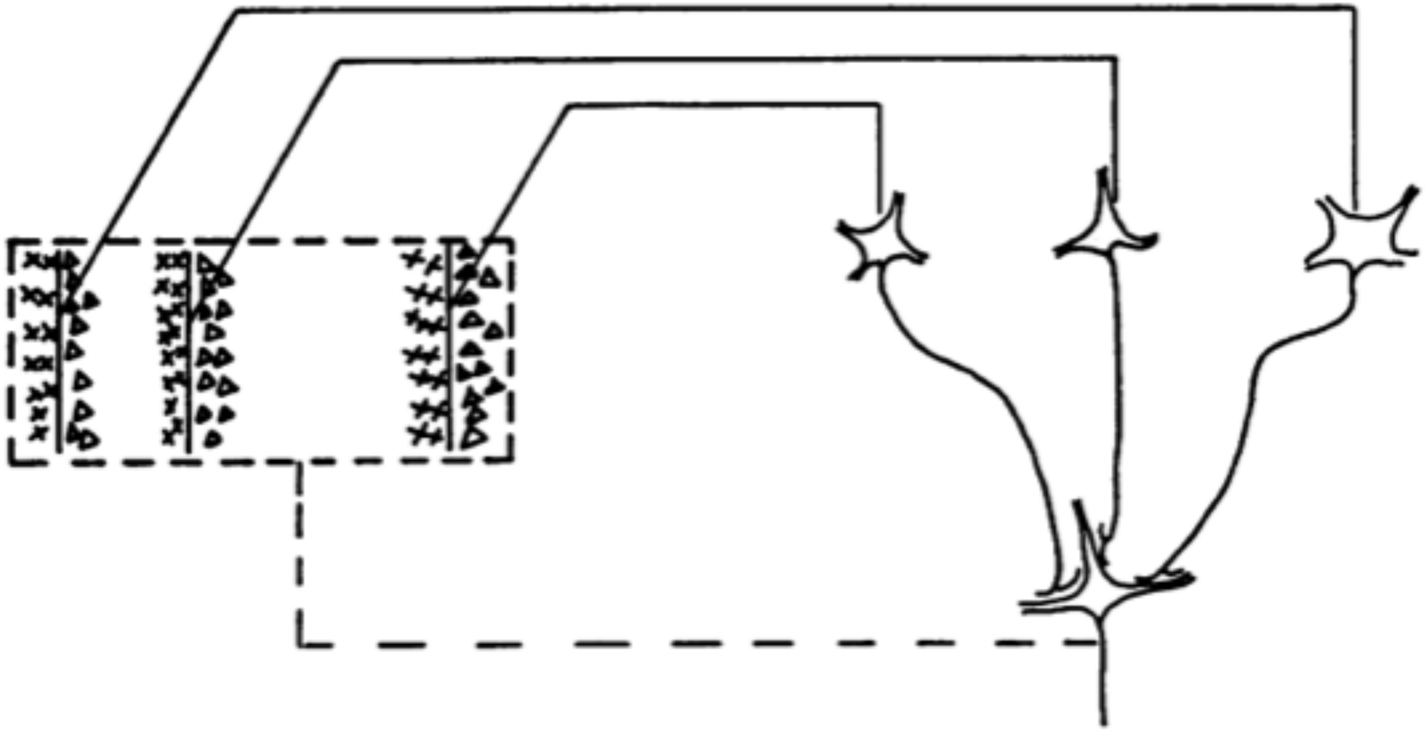
The Hubel – Wiesel conjecture: simple cells



Text-fig. 19. Possible scheme for explaining the organization of simple receptive fields. A large number of lateral geniculate cells, of which four are illustrated in the upper right in the figure, have receptive fields with 'on' centres arranged along a straight line on the retina. All of these project upon a single cortical cell, and the synapses are supposed to be excitatory. The receptive field of the cortical cell will then have an elongated 'on' centre indicated by the interrupted lines in the receptive-field diagram to the left of the figure.

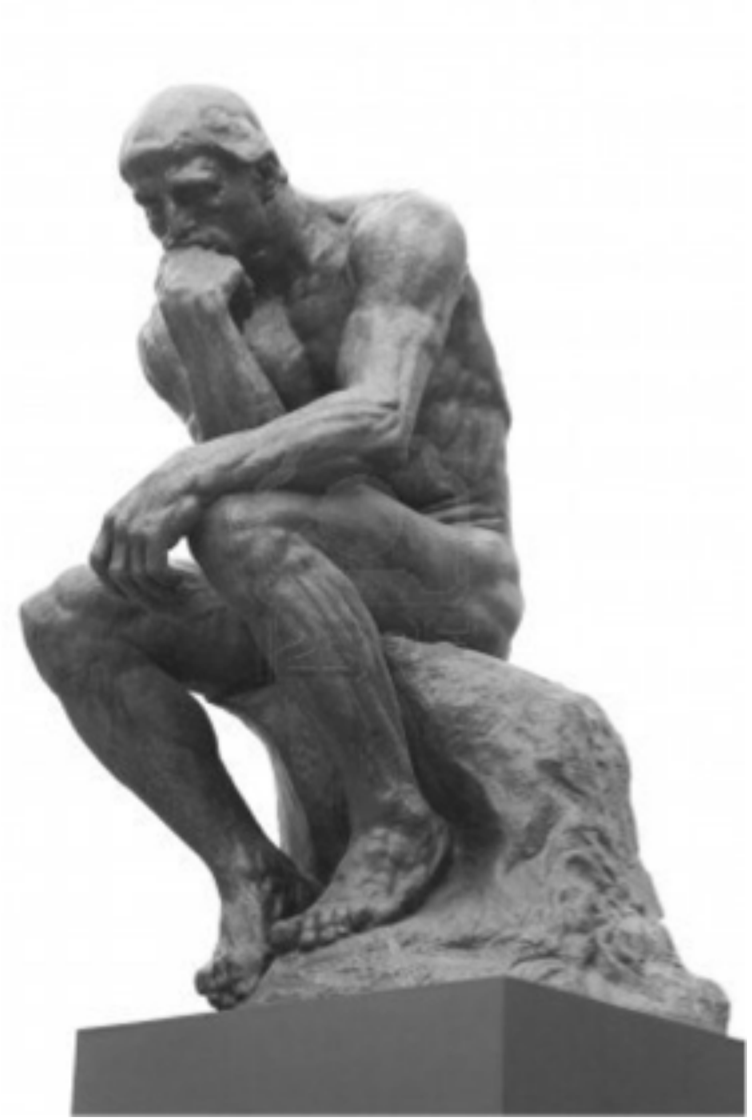
The Hubel – Wiesel conjecture: complex cells

-
-



Text-fig. 20. Possible scheme for explaining the organization of complex receptive fields. A number of cells with simple fields, of which three are shown schematically, are imagined to project to a single cortical cell of higher order. Each projecting neurone has a receptive field arranged as shown to the left: an excitatory region to

Wow!



Incidentally: brief history of the killer discussion section maneuver

- Crick and Watson, 1953 → Nobel 1962

"It did not escape our attention that a certain connection that we immediately postulated offers a possible mechanism for copying genetic material."

- Hubel and Wiesel 1962 → Nobel 1981
- Abel and Buck 1991 → Nobel 2006

Recall the study of neurons in CS

- 1944: McCulloch – Pitts neuron
- 1957: The Rosenblatt perceptron
- *"the embryo of an electronic computer that the Navy expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence."* NYT 1957
- 1960s: XOR and the Minsky – Papert perceptron (1972)
- The first AI winter
- 1975 K. Fukushima: the Cognitron

K Fukushima 1980: The Neocognitron

Explicitly inspired by Hubel and Wiesel

”unsupervised learning” of weights

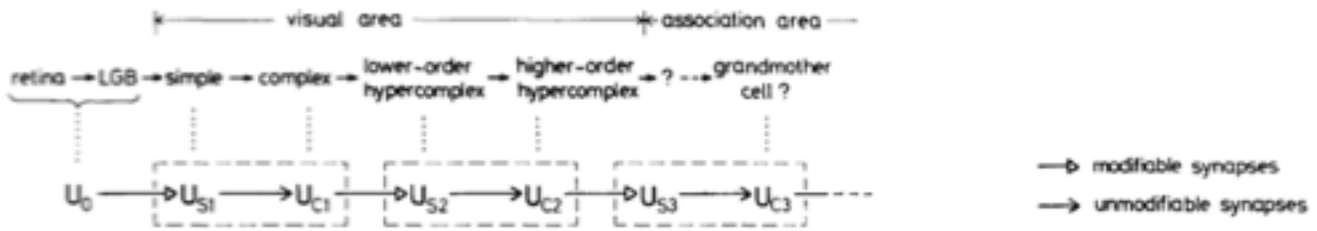


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of the neocognitron

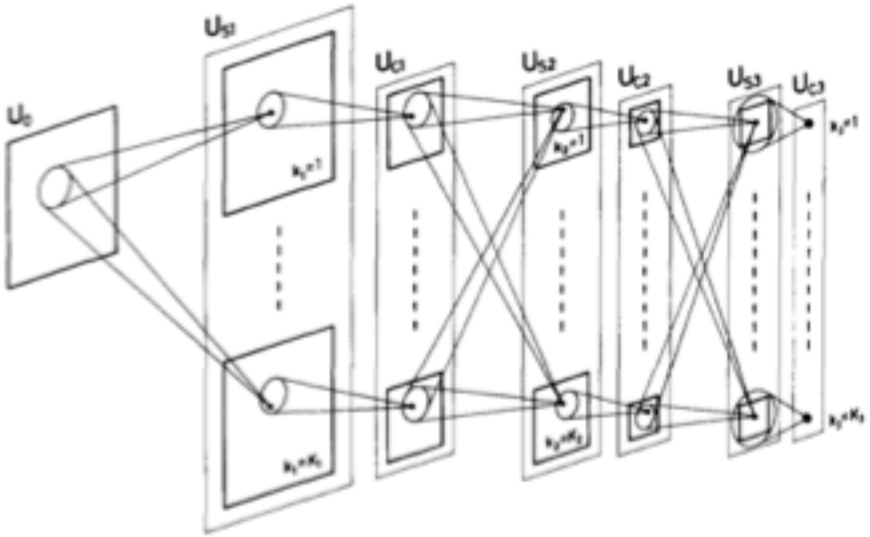


Fig. 2. Schematic diagram illustrating the interconnections between layers in the neocognitron

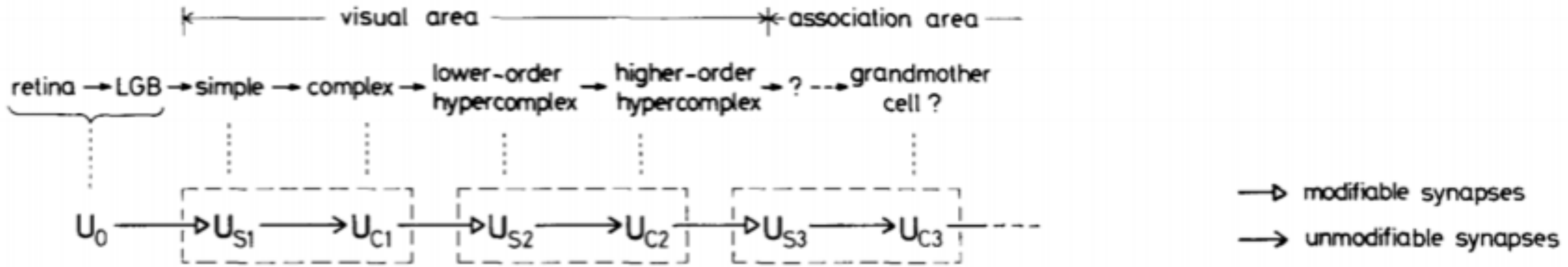


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of the neocognitron

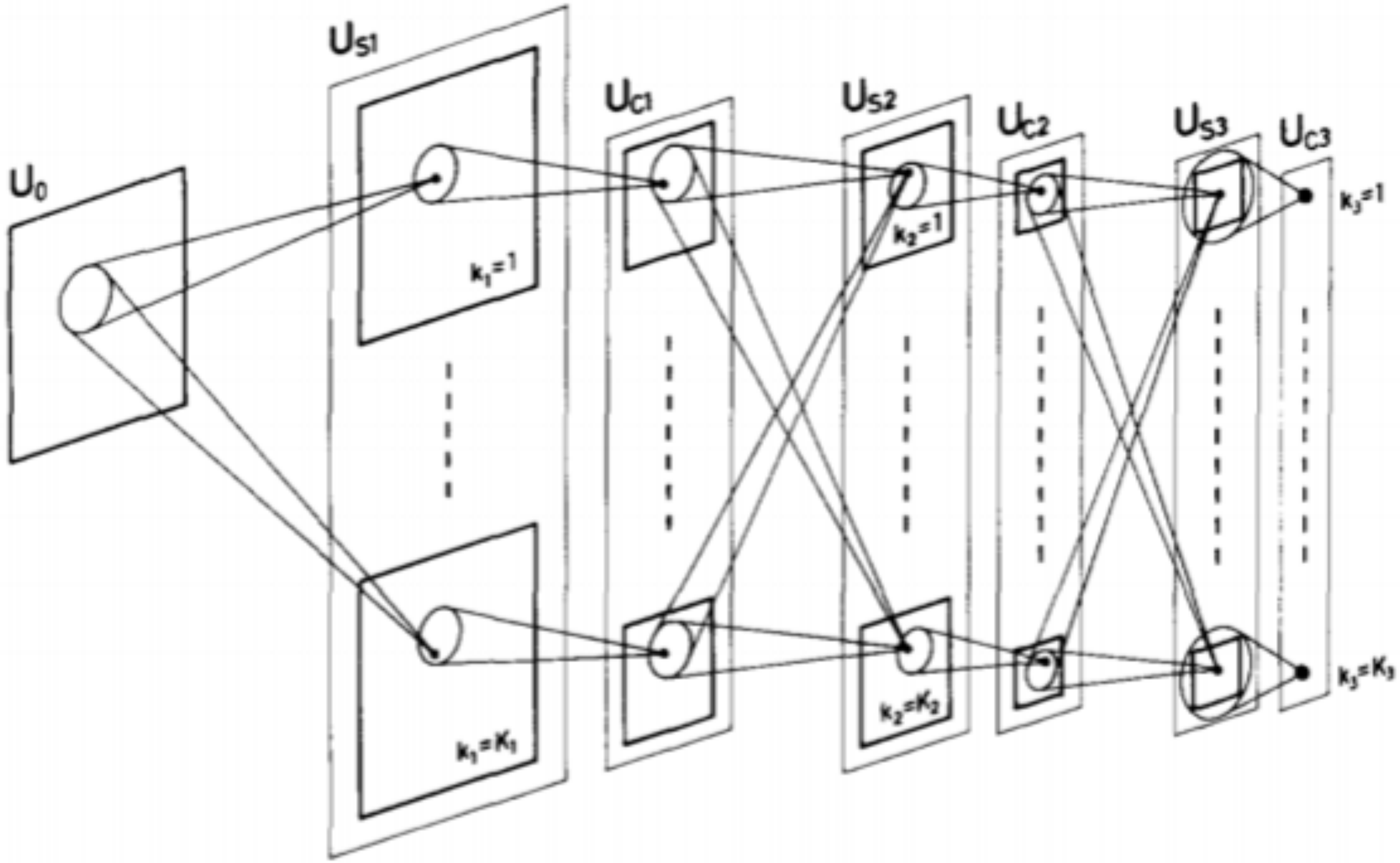


Fig. 2. Schematic diagram illustrating the interconnections between layers in the neocognitron

But how are the weights learned?

”The unreasonable effectiveness of gradient descent in learning”

cf: **The unreasonable effectiveness of mathematics in the natural sciences** by Eugene Wigner

Gradient descent

To approximate the minimum of a differentiable function $f: R^d \rightarrow R$

Set $t = 0$, and set x^0 to some initial point

Repeat

calculate the gradient $\nabla f(x^t)$

and the ``learning rate'' $\boldsymbol{\alpha}_t$

$x^{t+1} = x^t - \boldsymbol{\alpha}_t \nabla f(x^t)$

Until termination condition (typically, $|x^{t+1} - x^t|$ very small)

This is the science/ dark art of the matter

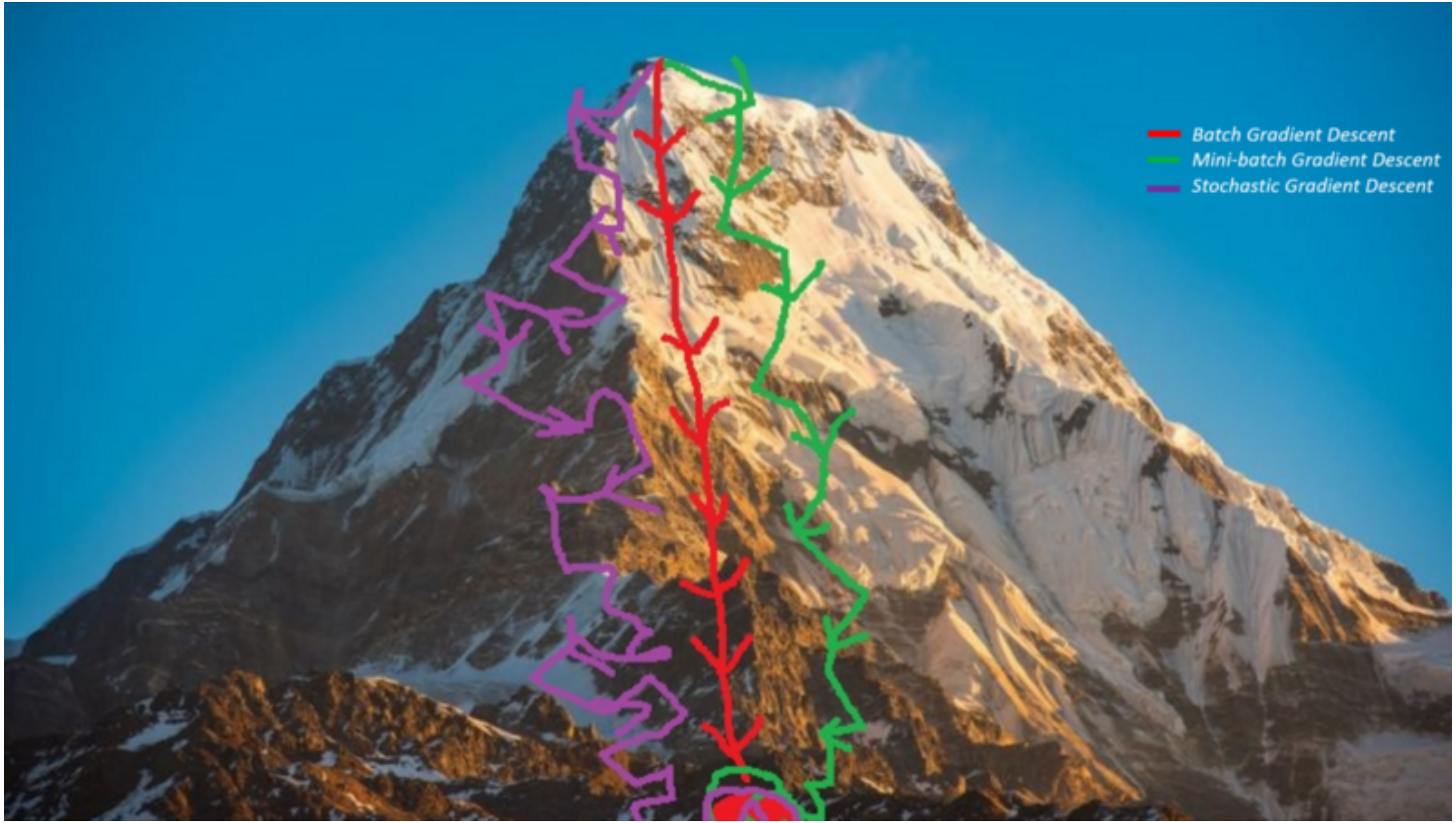
What is known about gradient descent

Theorem: If f is a strictly convex quadratic function $x^TA\,x + b\,x$, and L is the largest eigenvalue of A, taking $\boldsymbol{\alpha}_t = 1/L$ guarantees logarithmic convergence (the distance from the optimum decreases exponentially fast).

Theorem: Ditto if f is a general strictly convex function, and L is an upper bound on the Hessian's eigenvalues.

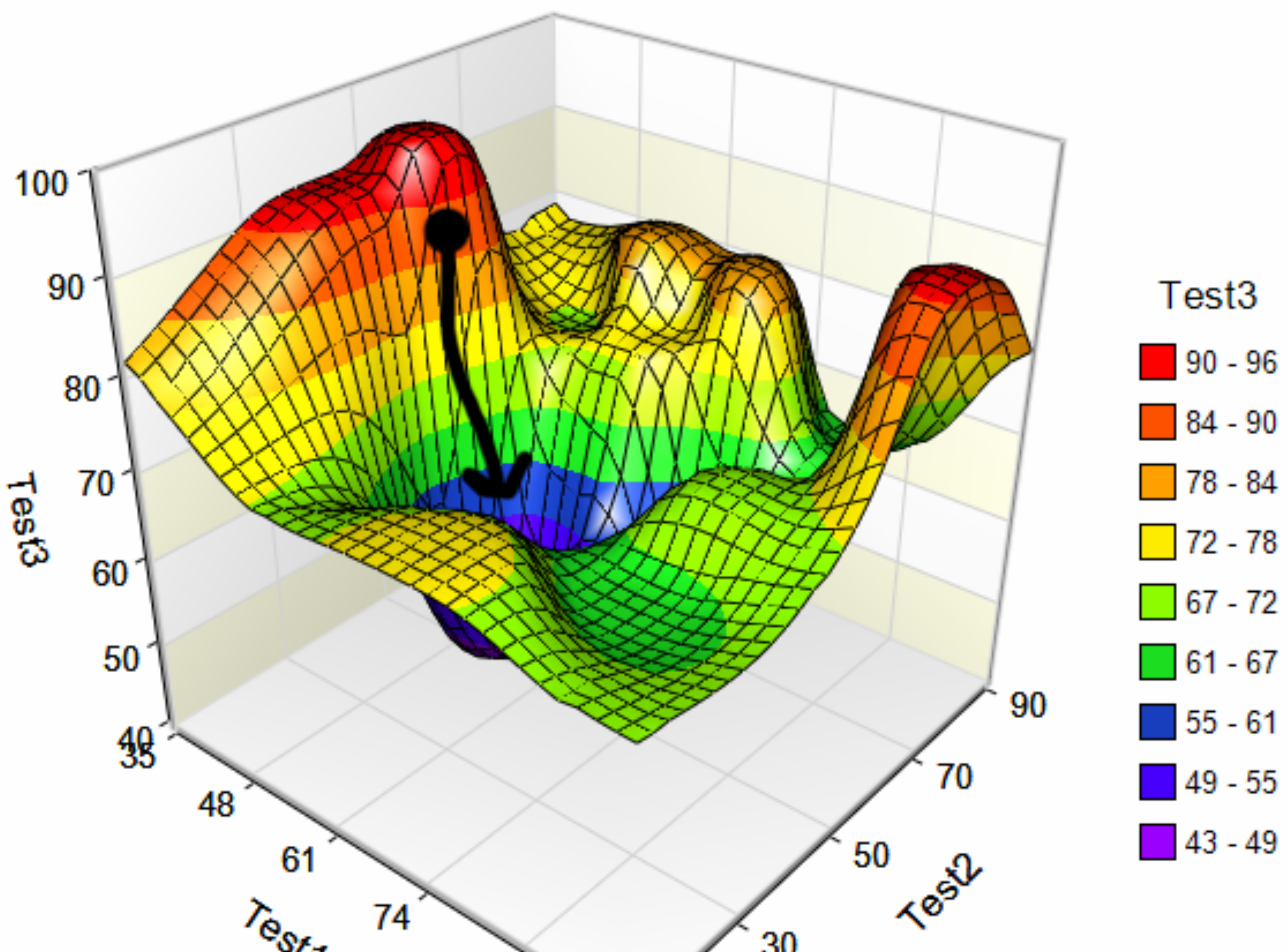
Theorem: If f is not convex, then no convergence can be guaranteed whatsoever

It doesn't matter!


Batch Gradient Descent
Mini-batch Gradient Descent
Stochastic Gradient Descent

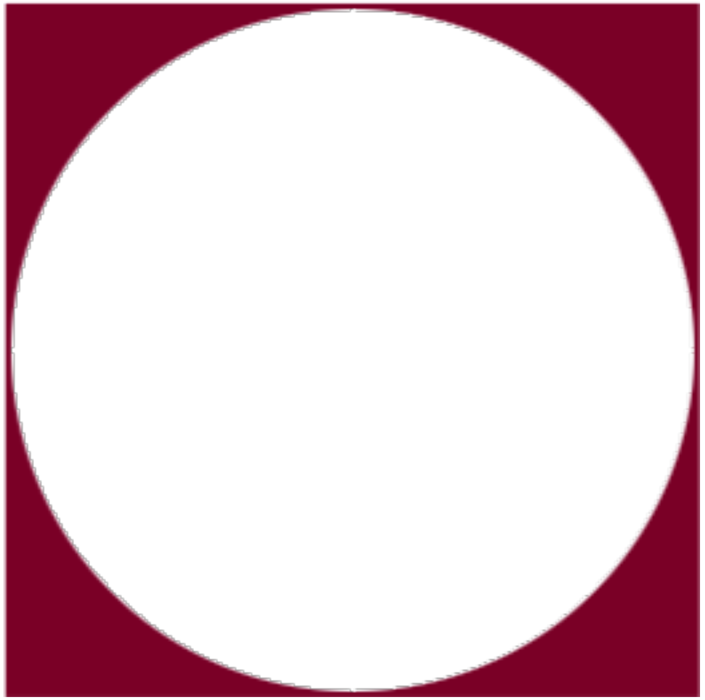
Stochastic gradient descent

- What if $f(x) = \sum_k g(x)$ is the sum of millions of functions?
- (The loss function of a NN is an example.)
- Computing the gradient takes millions of steps
- Except if you sample it!
- At every step you add the gradients of the next few functions (and scale up)



100
90
80
70
60
50
40
Test3
35
48
61
74
Test1
30
50
70
90
Test2
Test3
90 - 96
84 - 90
78 - 84
72 - 78
67 - 72
61 - 67
55 - 61
49 - 55
43 - 49

Btw: Can you visualize 10-dimensional data? Are you sure?



...the amazing shrinking inscribed ball

1-d	1.00
2-d	0.79...
3-d	0.52...
4-d	0.31...
5-d	0.16...
...	
10-d	0.002...
16-d	0.000004...