

Centre for Cognitive Science

Bratislava



# Computational cognitive neuroscience:

## 6. Perception

Lubica Beňušková  
Centre for Cognitive Science, FMFI  
Comenius University in Bratislava

# Eyes: How humans see versus how animals see

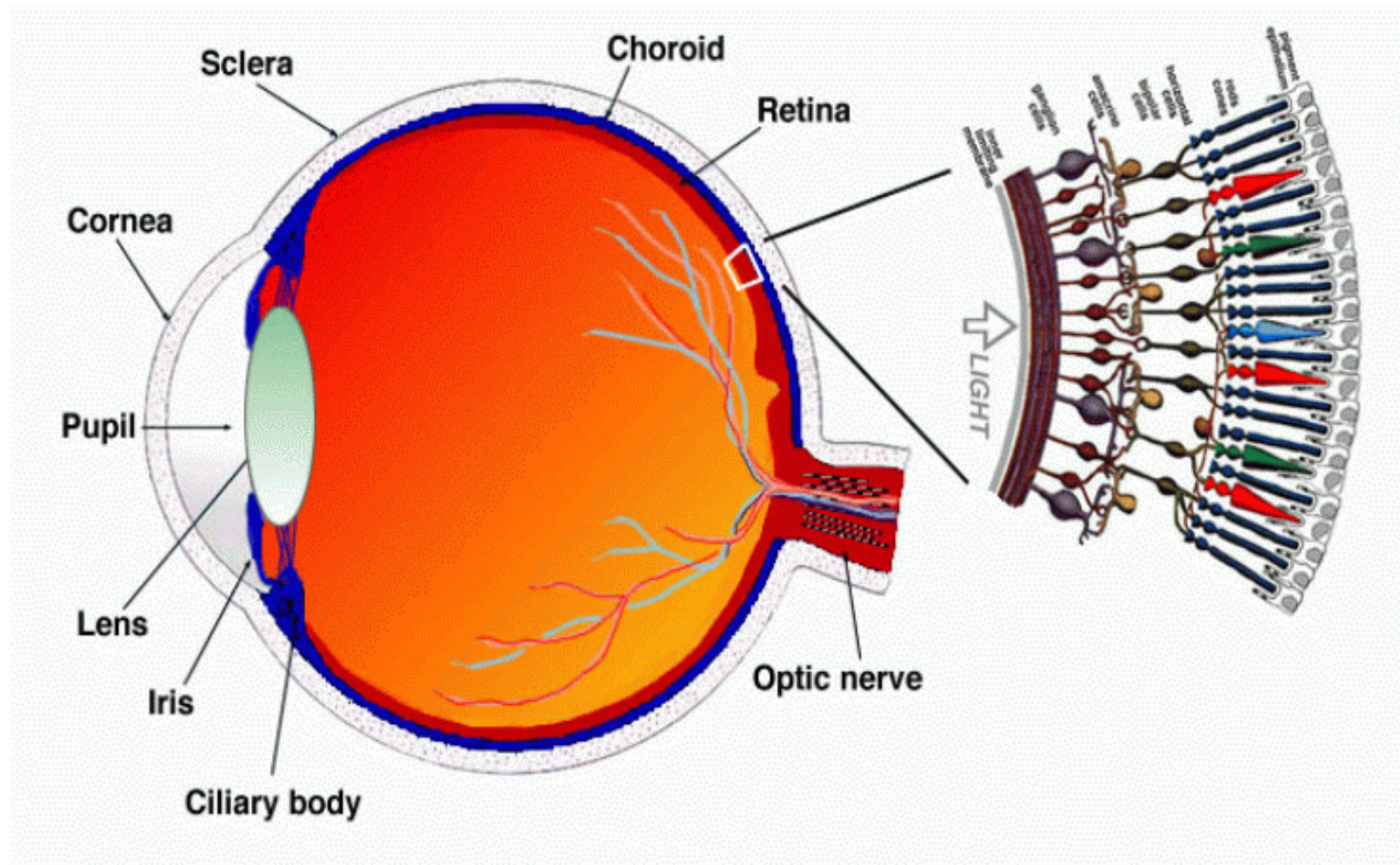
- We are confined to our own perceptions and version of reality we see and perceive, which in turn depends on the properties and resolution of our senses. Perception (from Latin 'perceptio' meaning gathering, receiving).





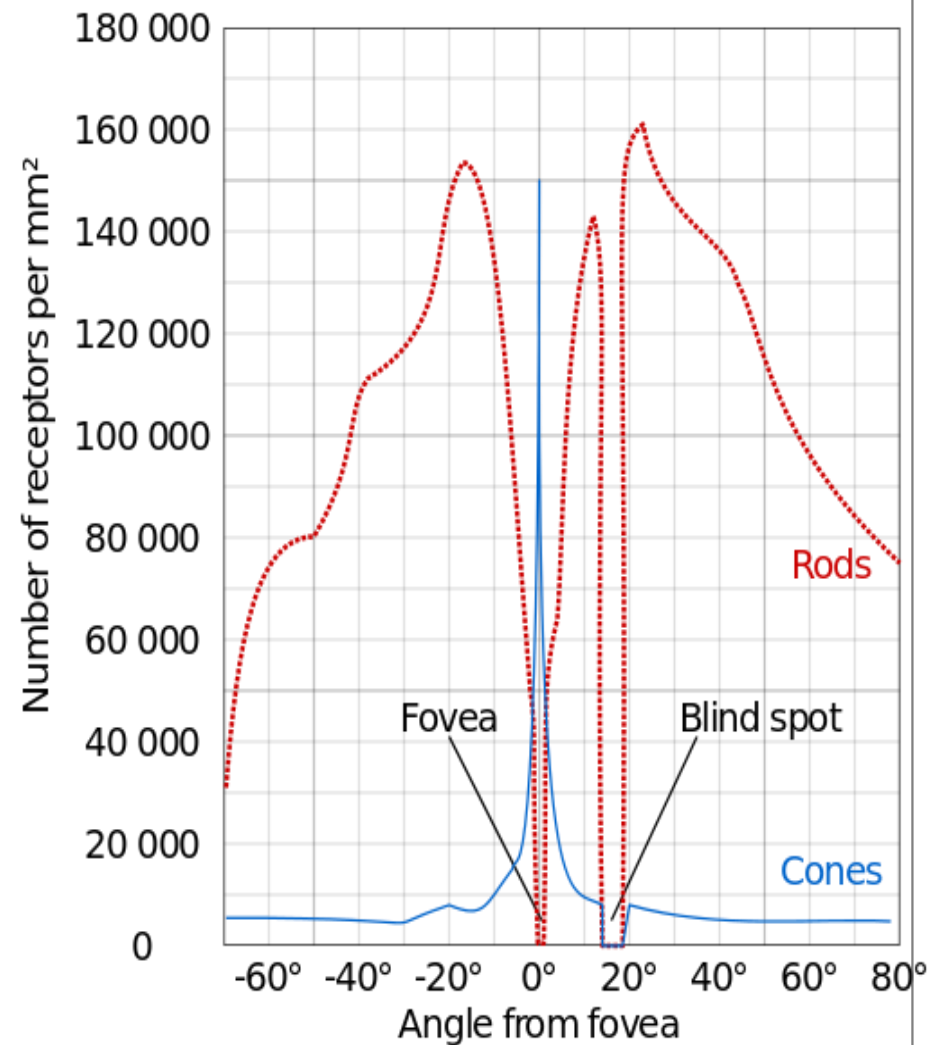
# The primate eye

- The act of seeing starts when the cornea and then the pupil and lens of the eye focuses an image onto the retina.
- **The retina** is in fact part of the brain that contains photoreceptive cells which convert patterns of light into neuronal signals relayed to the cortex.



## Fovea, rods and cones

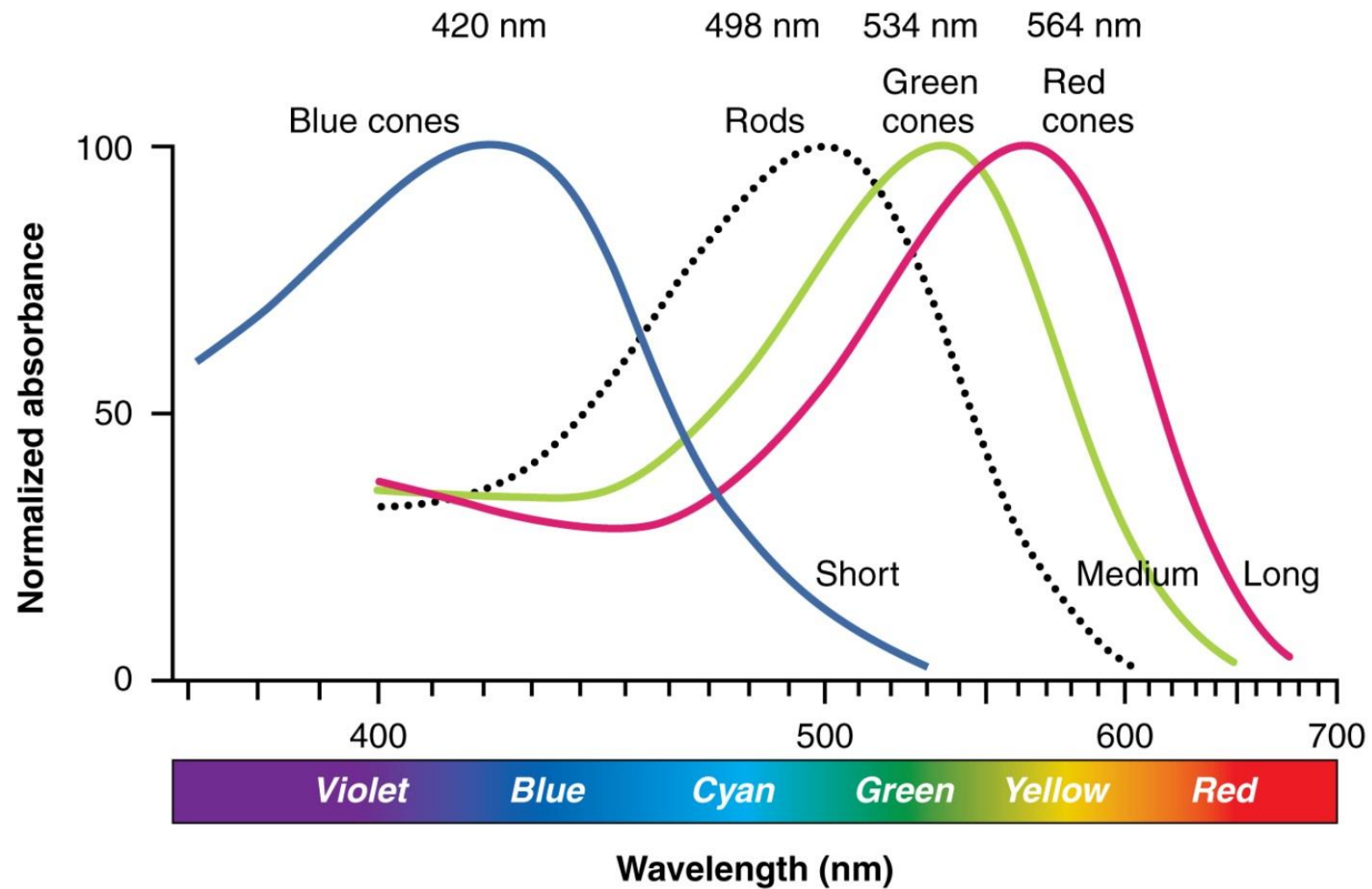
- At the "centre" of the retina lies the **fovea** (macula), which contains only **cone** cells; and
- Fovea is the region capable of producing the highest visual acuity and highest resolution.
- Across the rest of the retina, rods and cones are intermingled.
- Signals from the rods and cones converge on ganglion and bipolar cells.
- No photoreceptors are found at the blind spot, the area where ganglion cell axons are collected into the optic nerve and leave the eye towards the LGN.



## Comparison of rods and cones (for humans)

RODS	CONES
Very light sensitive	Sensitive only to direct light
One type of photosensitive pigment	3 types of photosensitive pigment
Confer achromatic vision	Confer colour vision
Have more pigment than cones	Have less pigment than rods
Scotopic vision (for low light intensity)	Photopic vision (for high light intensity)
Loss causes night blindness	Loss causes colour blindness
Low visual acuity (resolution)	High visual acuity (resolution)
Not present in fovea	Concentrated in fovea
Slow response to light	Fast response to light
About 120 million rods in each retina	About 6 million cones in each retina

# Light wavelength for cones and rods

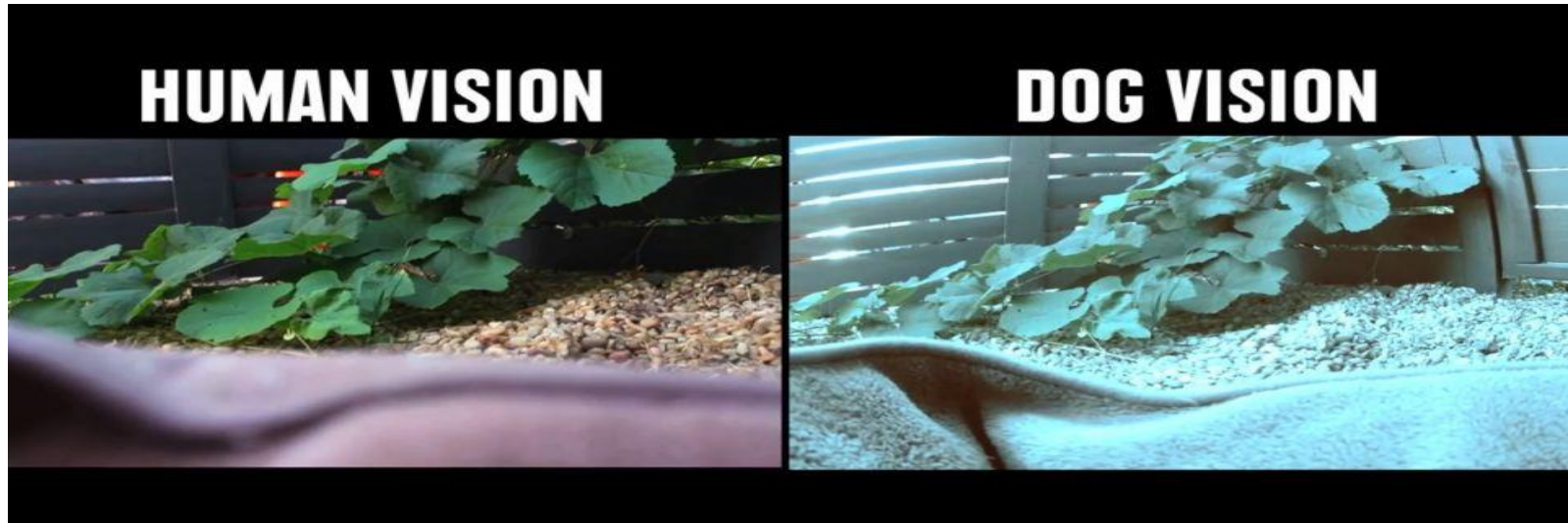


- The spectrum of the wavelength for 3 types of cones and 1 type of rods.

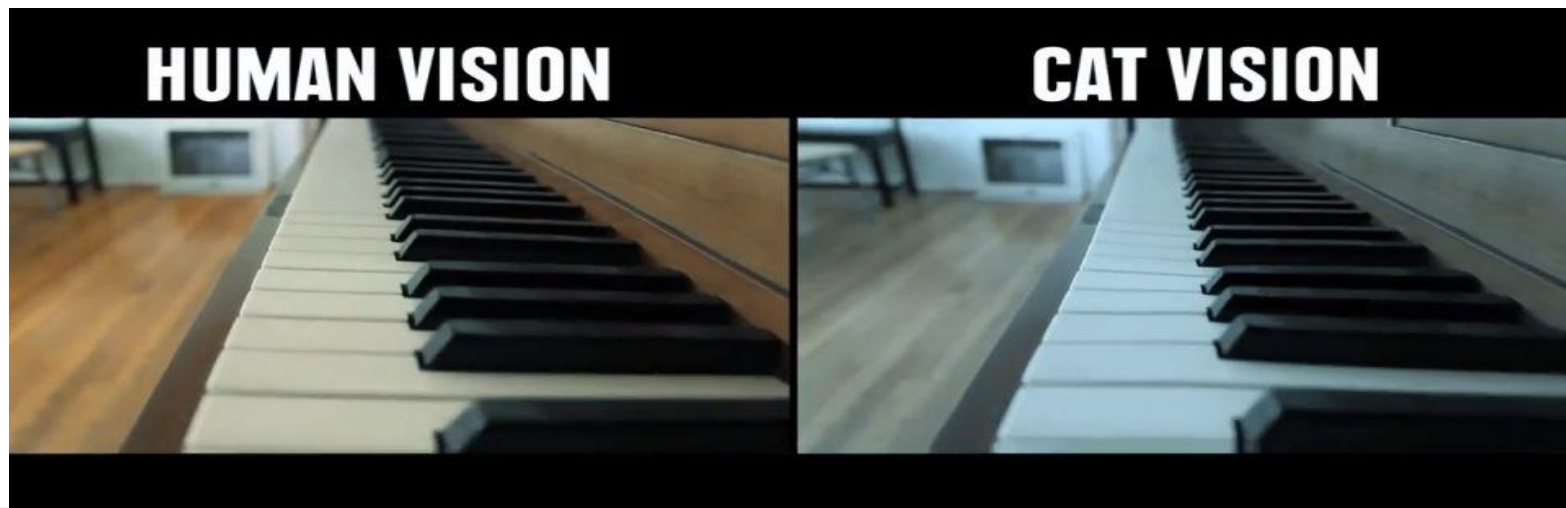


## Photoreceptor sensitivity

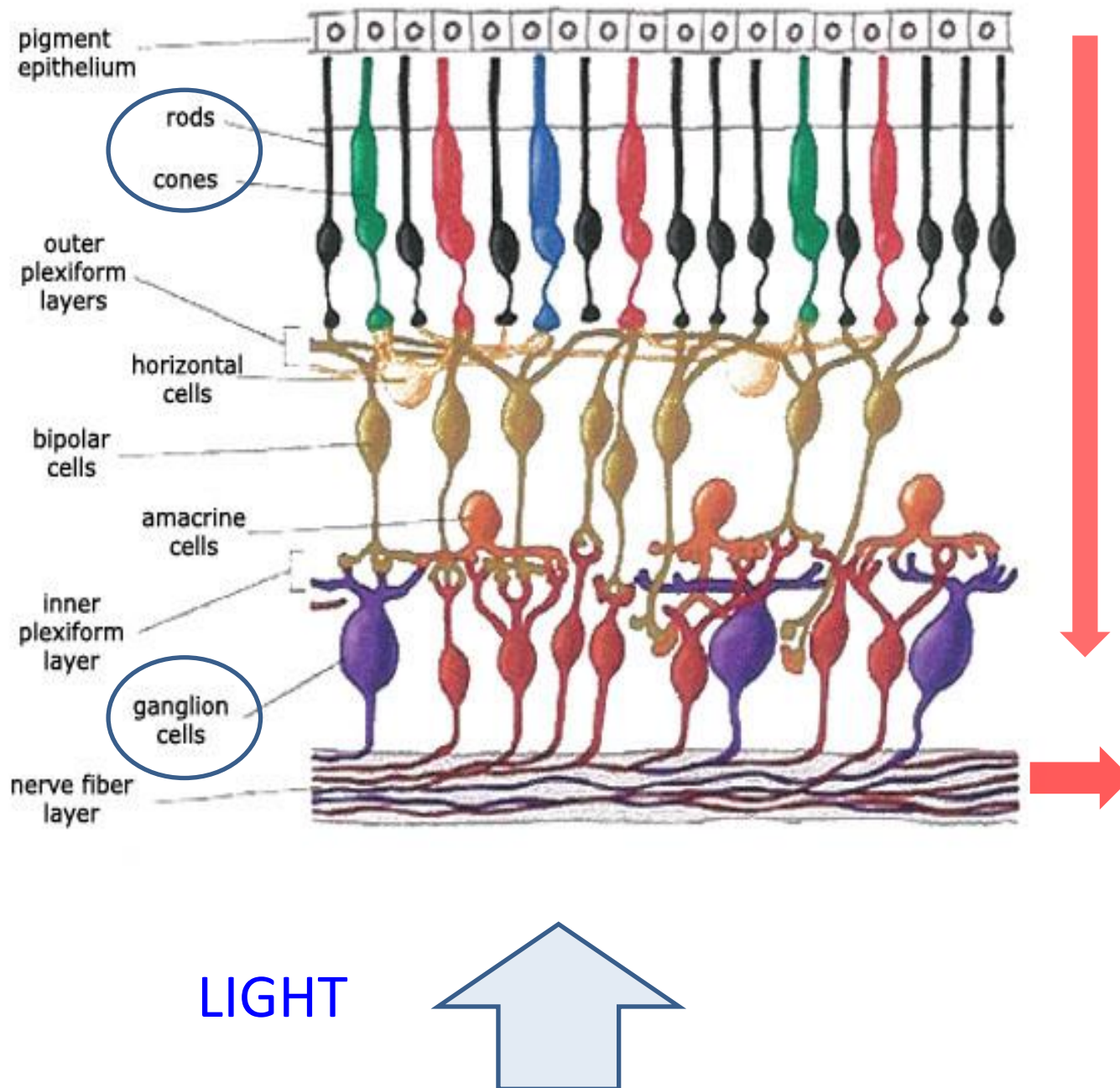
- Dogs have only 2 cones, blue and yellow. They see like colour blind people.



- Cats have many more rods, thus an excellent night vision, but have fewer cones than humans, thus less colourful vision .



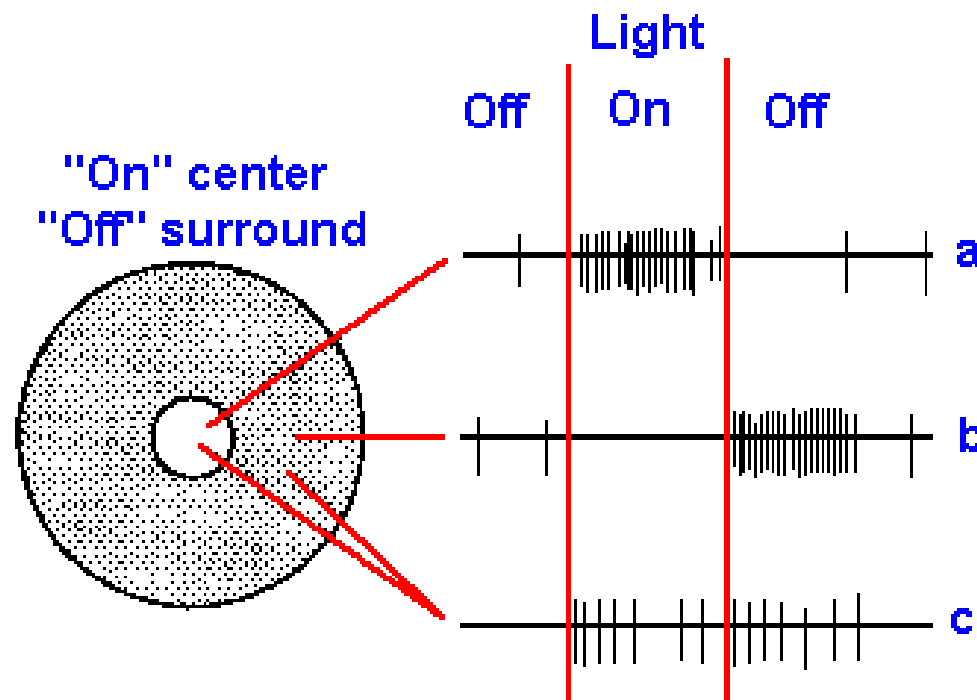
# The retina



- The flow of electric signals goes from photoreceptors to ganglion cells.
- Axons of ganglion cells in the very back of the retina form the optical nerve, which travels to the Lateral Geniculate Nucleus (LGN) in the thalamus.

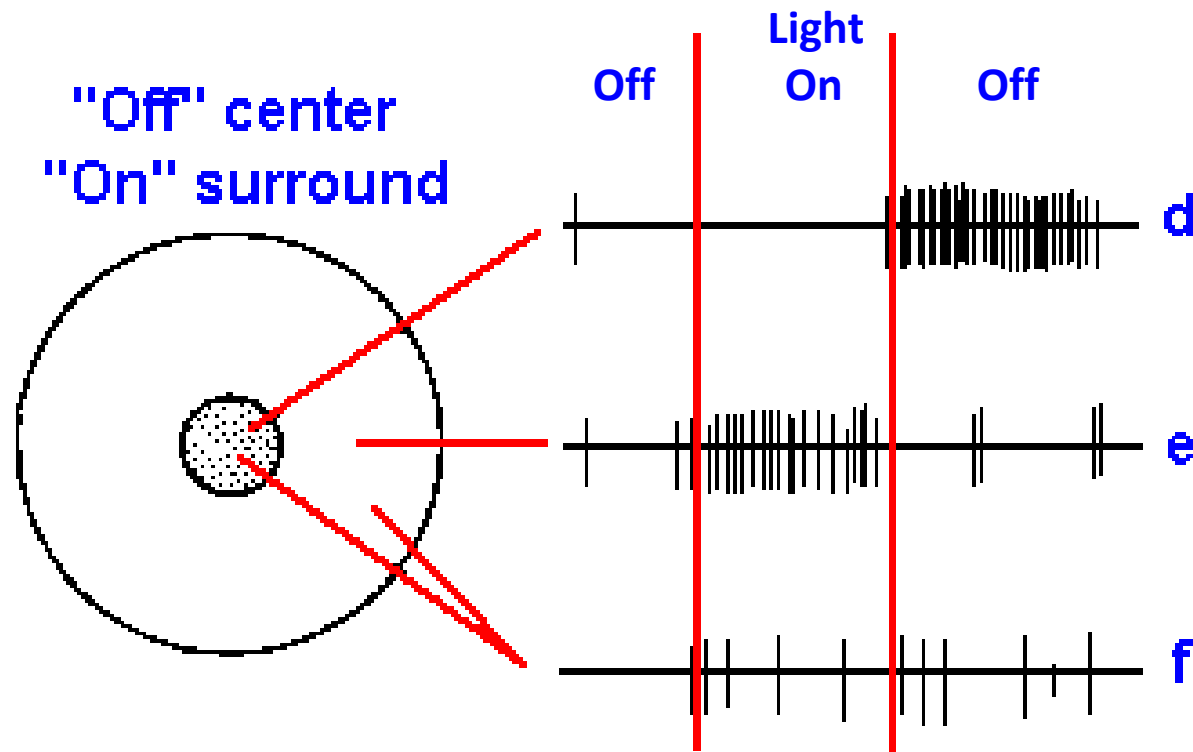


## Receptive fields of ganglion and LGN neurons: **ON-OFF** cells



- Receptive fields are circular areas of the retina surrounded by an annulus of different properties. The ON-OFF cell responds (a) by an increase in the spiking rate when the centre is illuminated (on centre) and (b) by an increase in spiking rate when the surround is darkened (off surround). (c) The cell gives no difference in response when both the centre and surround are illuminated (Hubel DH: *Sci. Amer.* 209:54-62, 1963)

## Receptive fields of ganglion and LGN neurons: **OFF-ON** cells



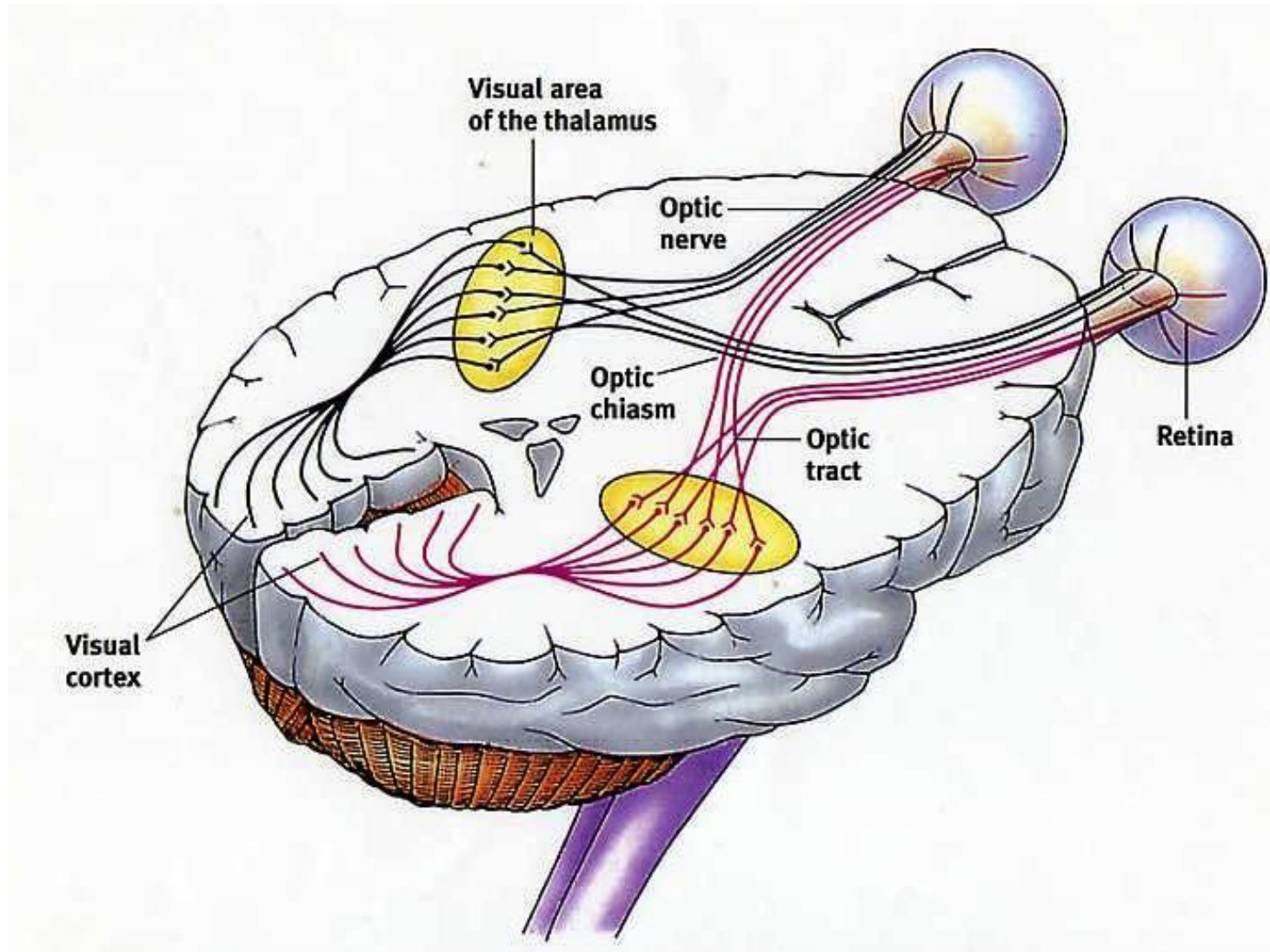
- Receptive fields are circular areas of the retina surrounded by an annulus of different properties. The OFF-ON cell responds (d) by an increase in the spiking rate when the annulus is dark (off centre) and (e) by an increase in spiking rate when the surround is illuminated (on surround). (f) The cell gives no difference in response when both the centre and surround are illuminated (Hubel DH: *Sci. Amer.* 209:54-62, 1963)

Stare at the 4 dots in the middle for about 30 sec then close your eyes





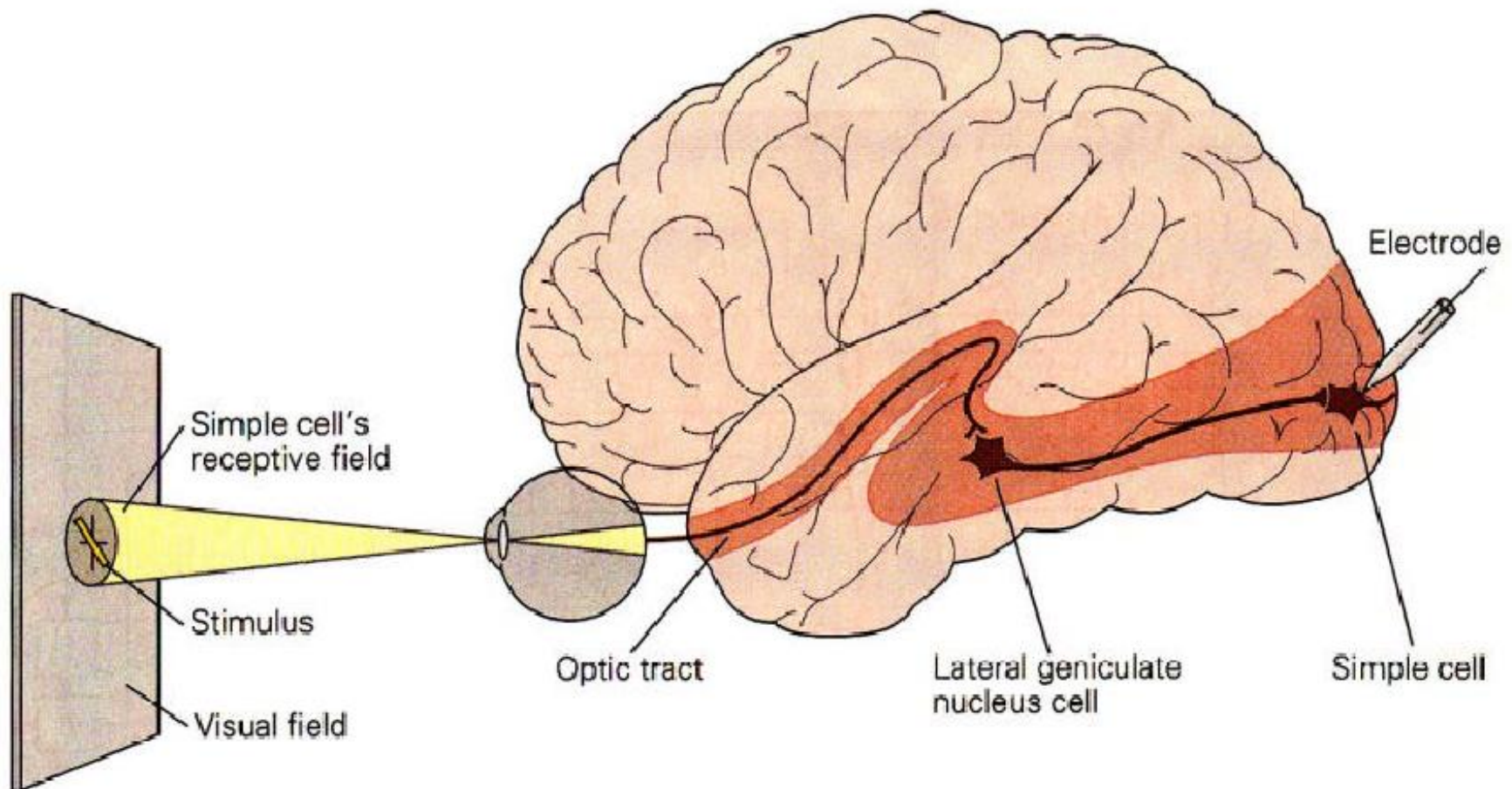
# Optical tract in the human brain



- The signals in the form of trains of spikes go from the retinal ganglion cells to the LGN and then to the primary visual cortex V1.

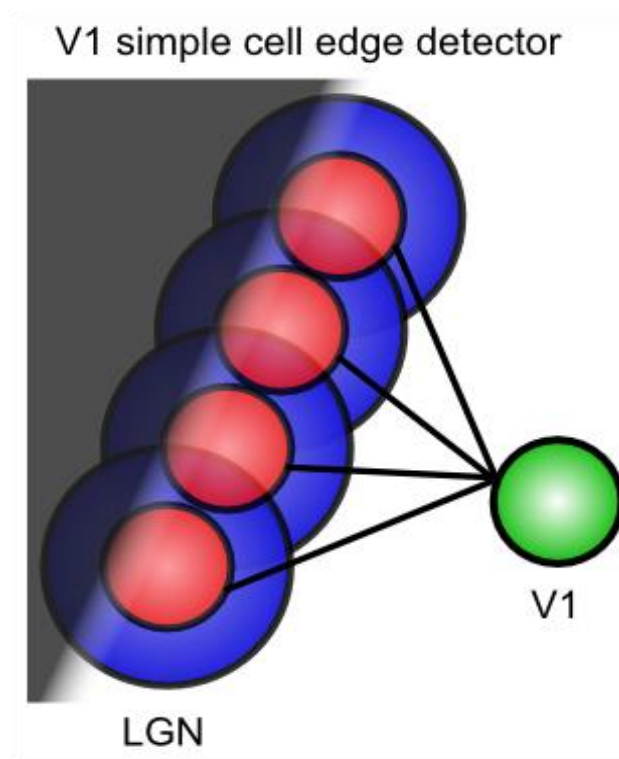
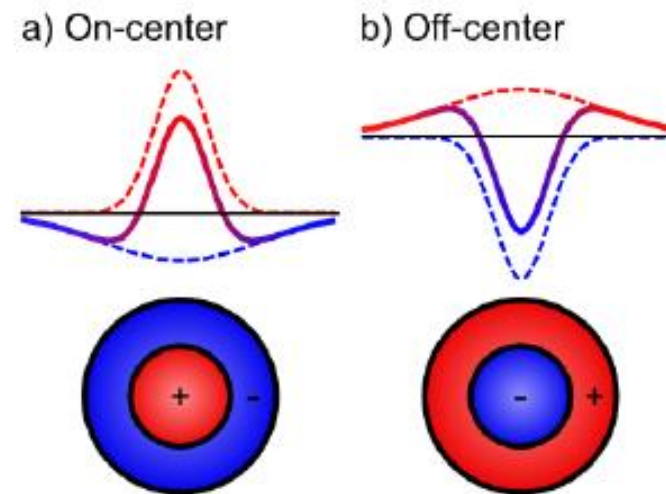
## What happens between the LGN and V1

- Recording the neural activity in V1 in response to stimuli of different shapes revealed the V1 cells respond to the bar-like stimuli of light or **edges of different contrast** (Hubel and Wiesel).



## What happens between the LGN and V1

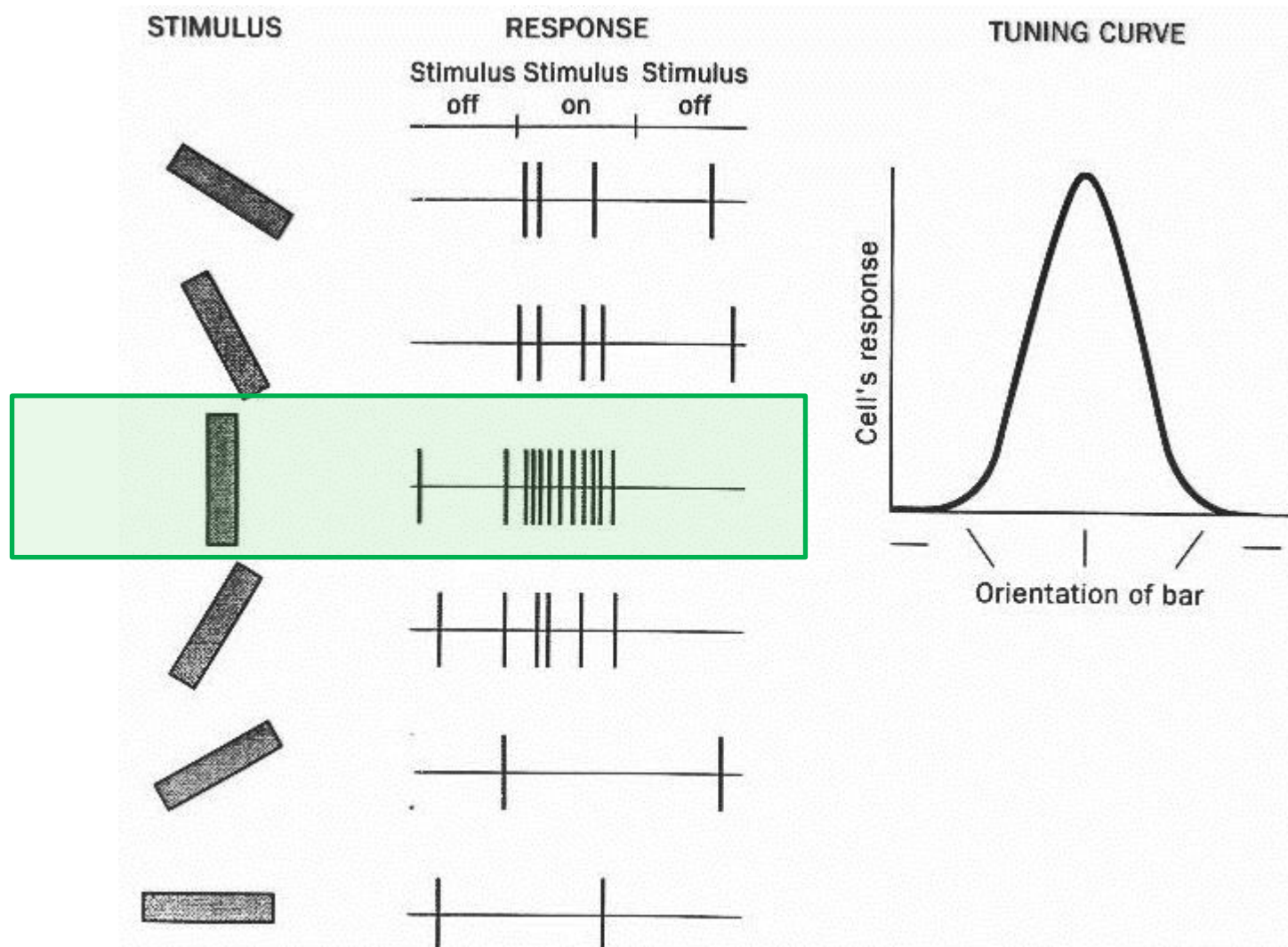
- The response properties of retinal cells can be summarized by the Difference-of-Gaussian (DoG) filters, with a narrow central region and a wider surround (also called centre-surround receptive fields).
- A V1 simple cell detects an oriented edge of contrast in the image thanks to receiving spikes from a series of LGN on-centre cells aligned along the edge.





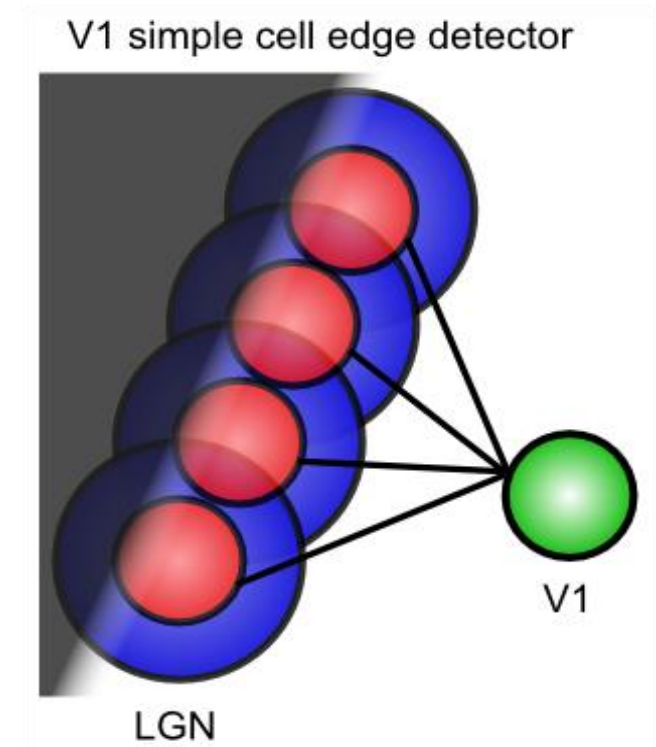
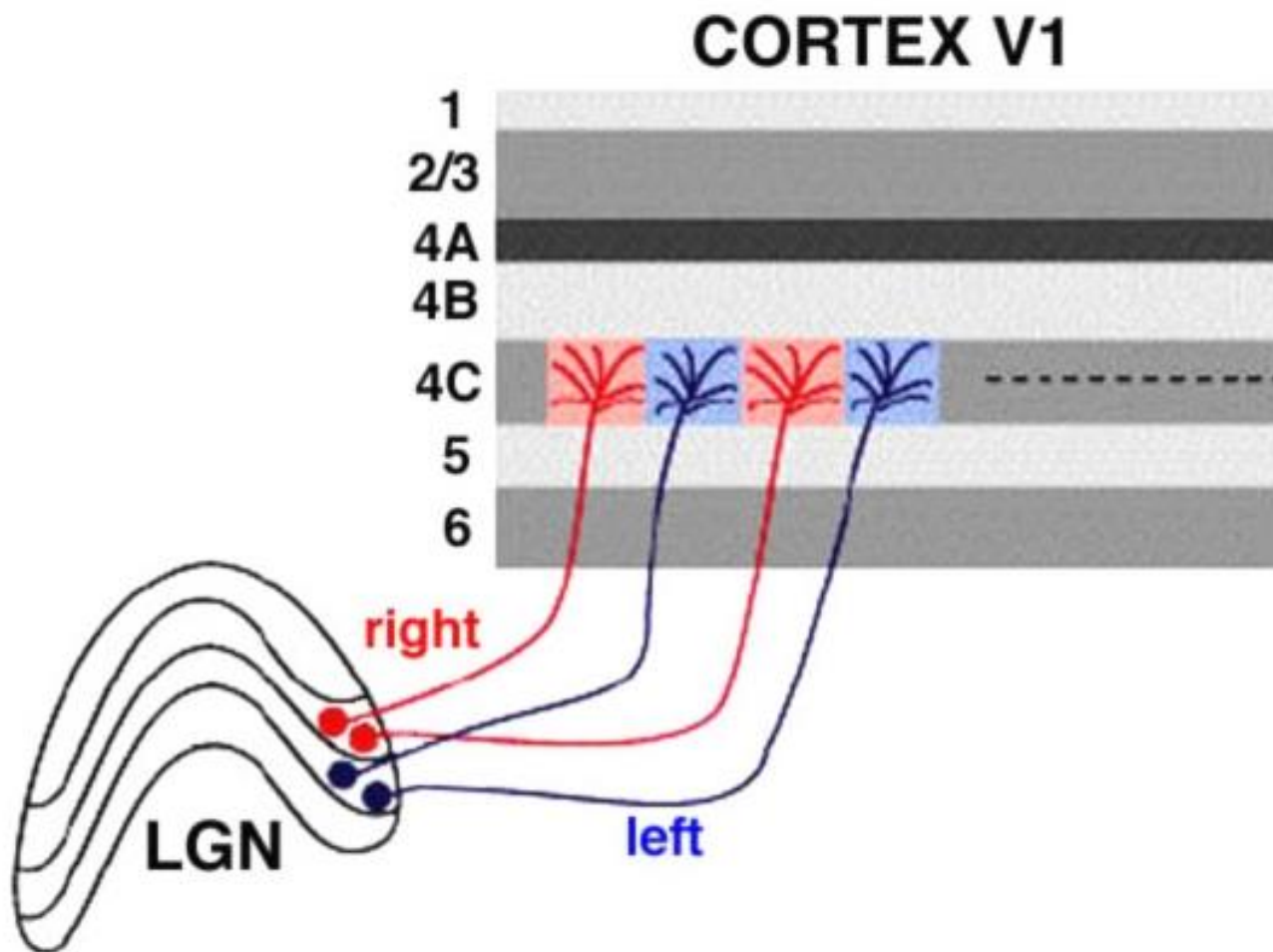
## Orientation tuning data in V1 neurons

- Orientation tuning of an individual V1 neuron in response to bar-like stimuli at different orientations -- this neuron shows a preference for vertically oriented stimuli.



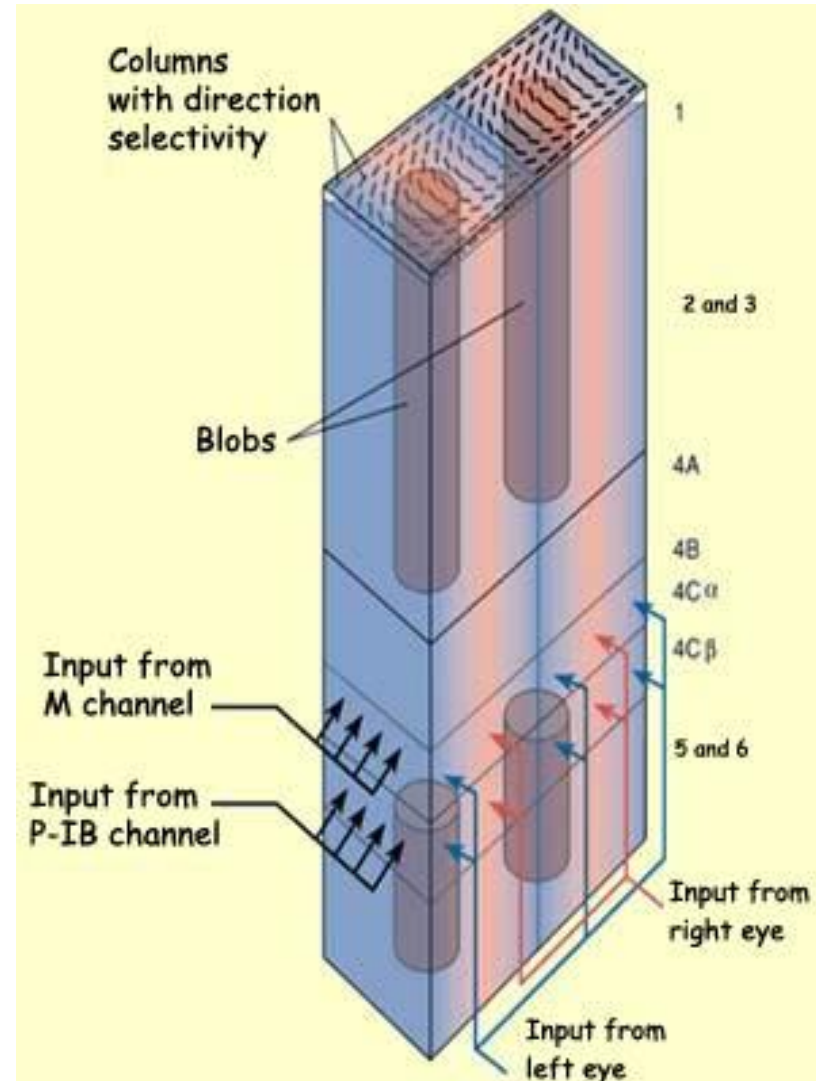
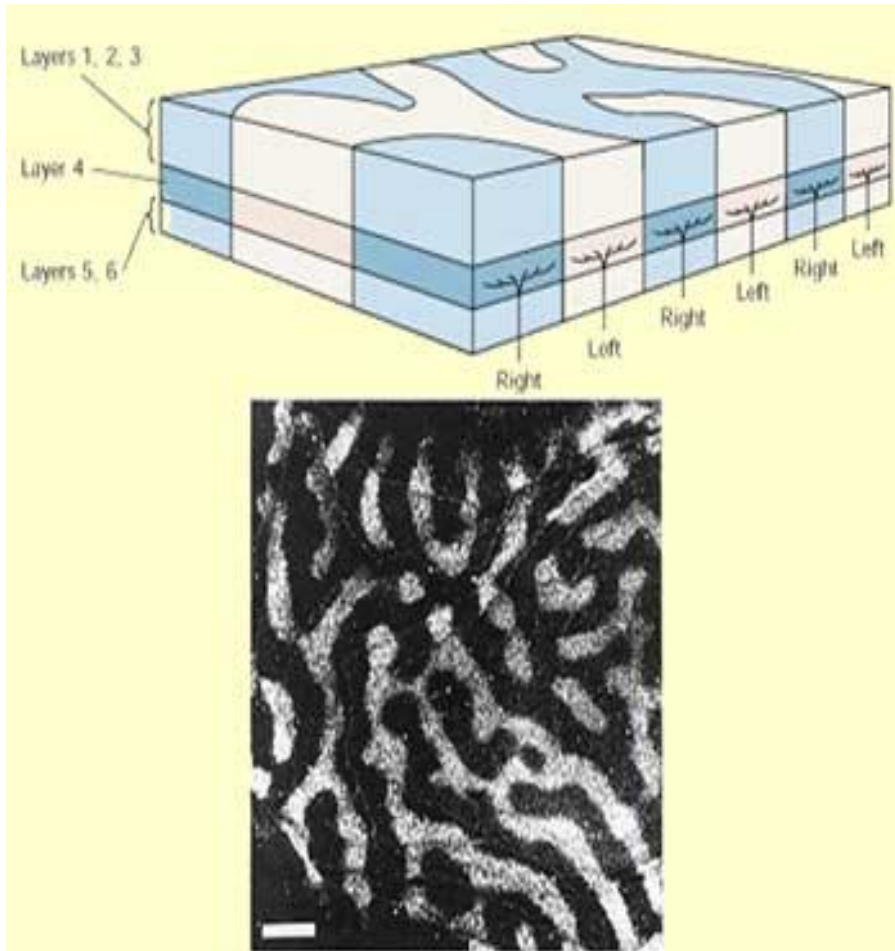
## What happens between the LGN and V1

- The signals from each eye are segregated within LGN and go into different ocular dominance columns within area V1.



# Organisation of primary visual cortex

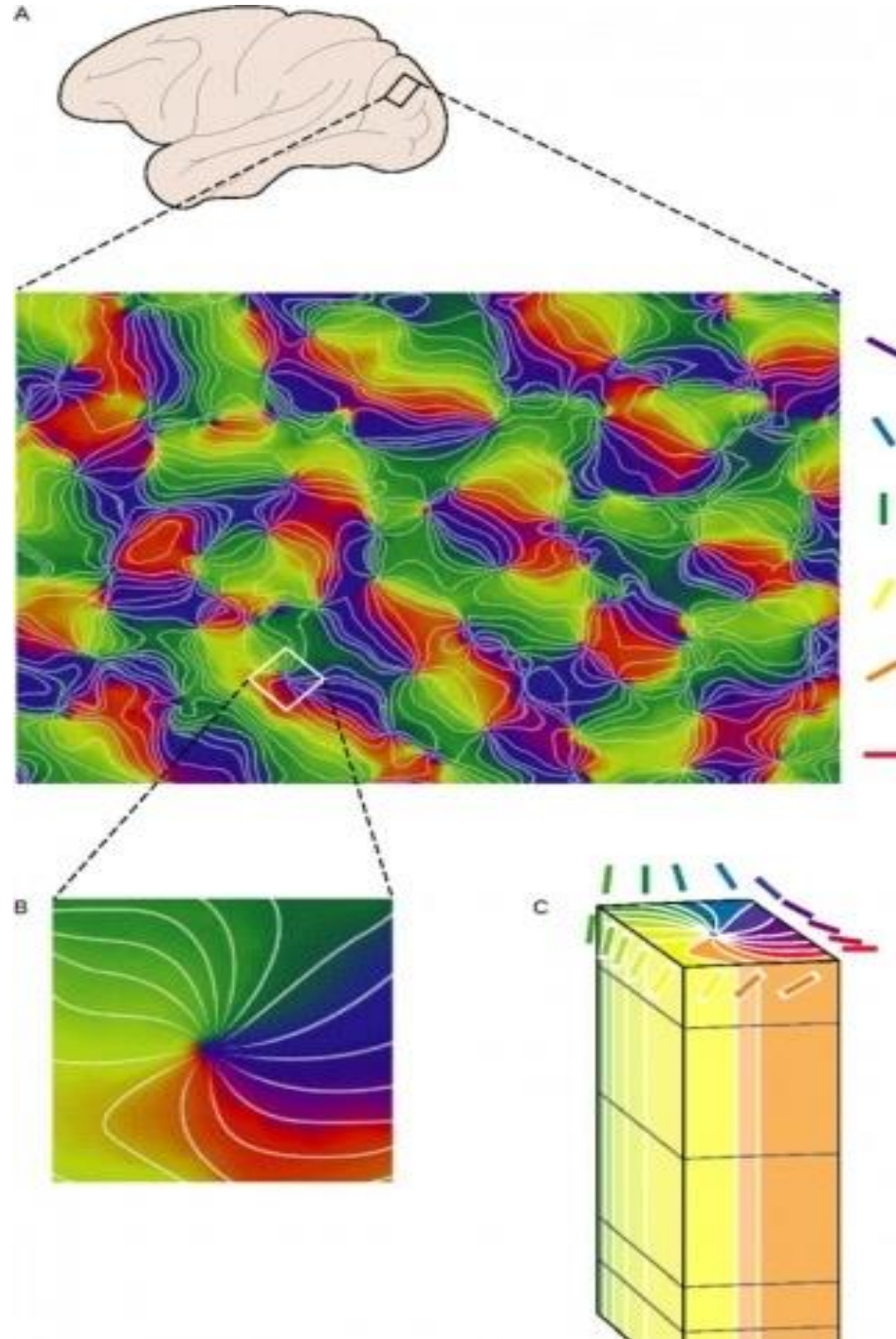
- Ocular dominance stripes and orientation selective cells in V1.





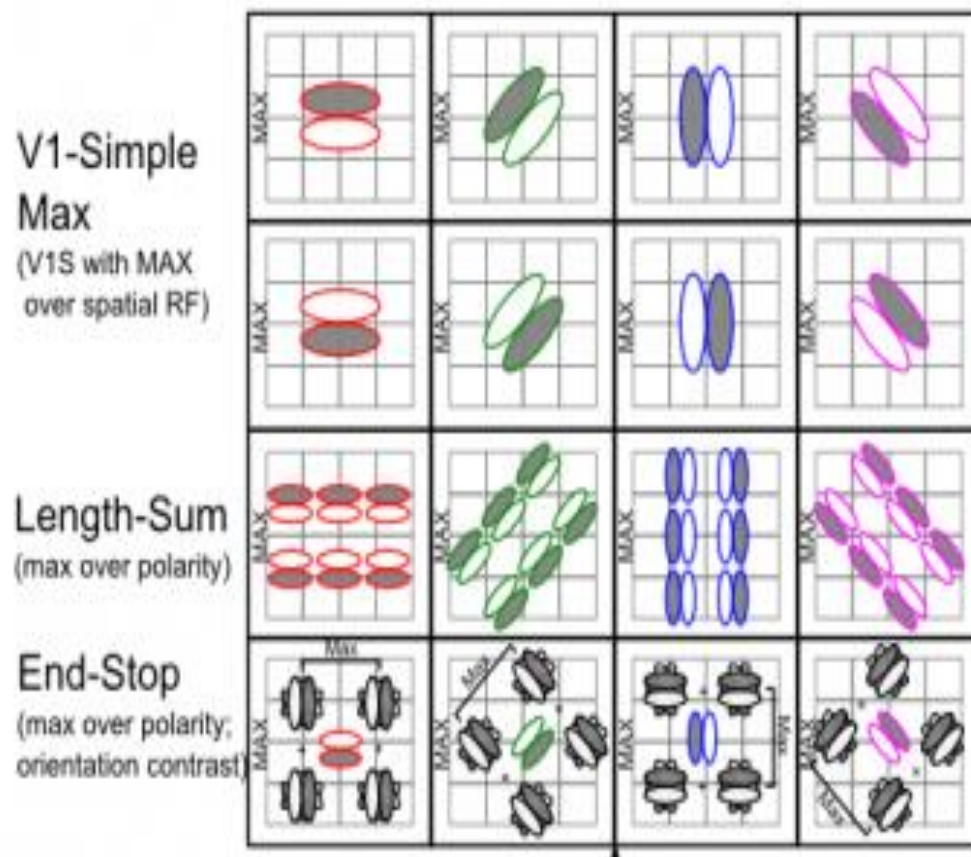
# Topographic organization of oriented edge detectors in V1

- Neighbouring regions of neurons have similar orientation tuning, as shown in this colored map.
- Different colours indicate orientation preference as shown in panel C.
- Panel B shows how a full 360° loop of orientations arrange around a central point -- these are known as **pinwheel structures**.



# Neurons with more complex RF in V1

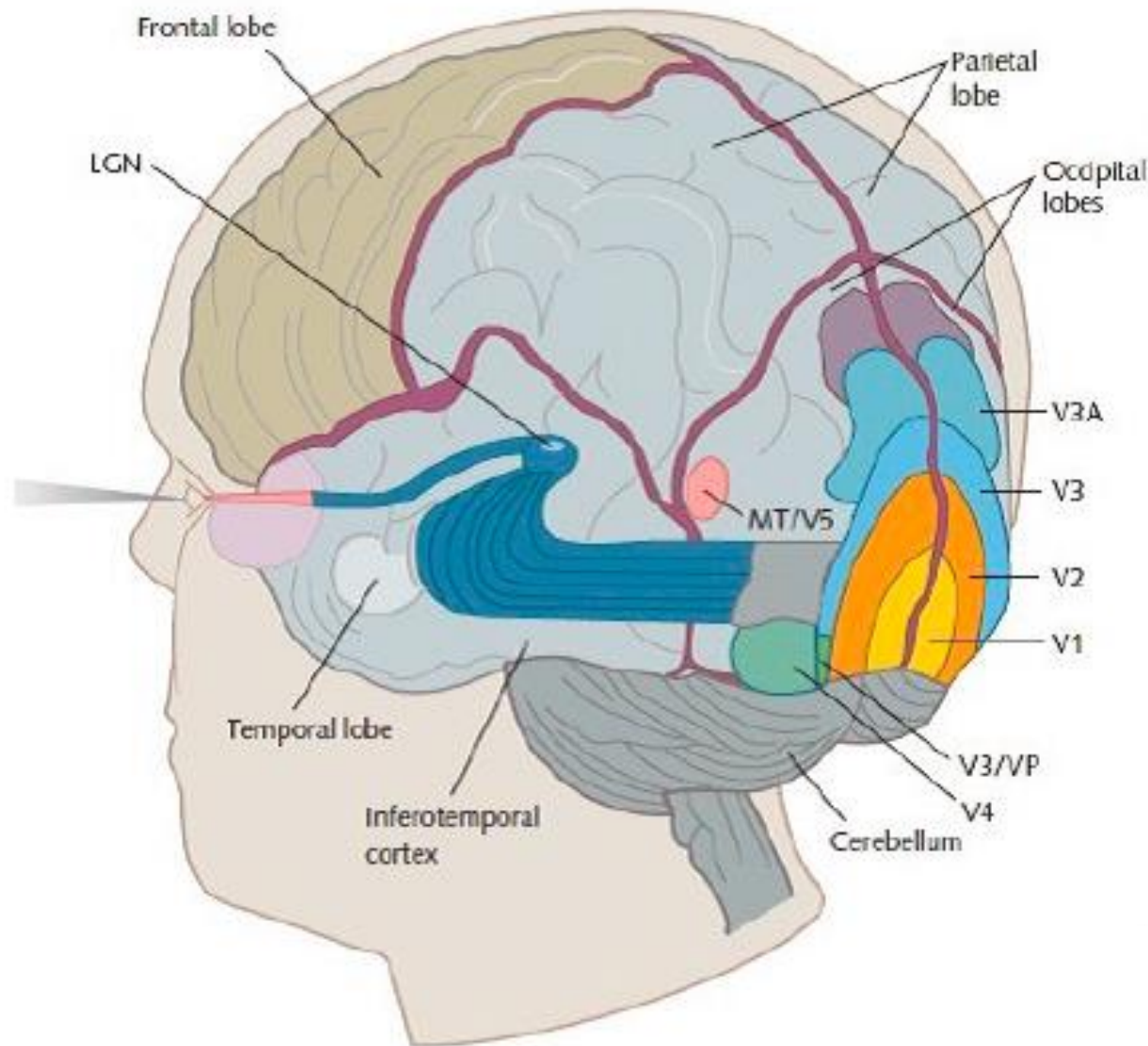
## V1 Complex: & MAX over larger spatial RF



- The **complex cells** integrate over the simple cell properties, including **abstracting** across the polarity (positions of the on vs. off coding regions), and creating larger receptive fields by integrating over multiple locations
- The **end stop cells** are the most complex, detecting any form of contrasting orientation adjacent to a given simple cell.

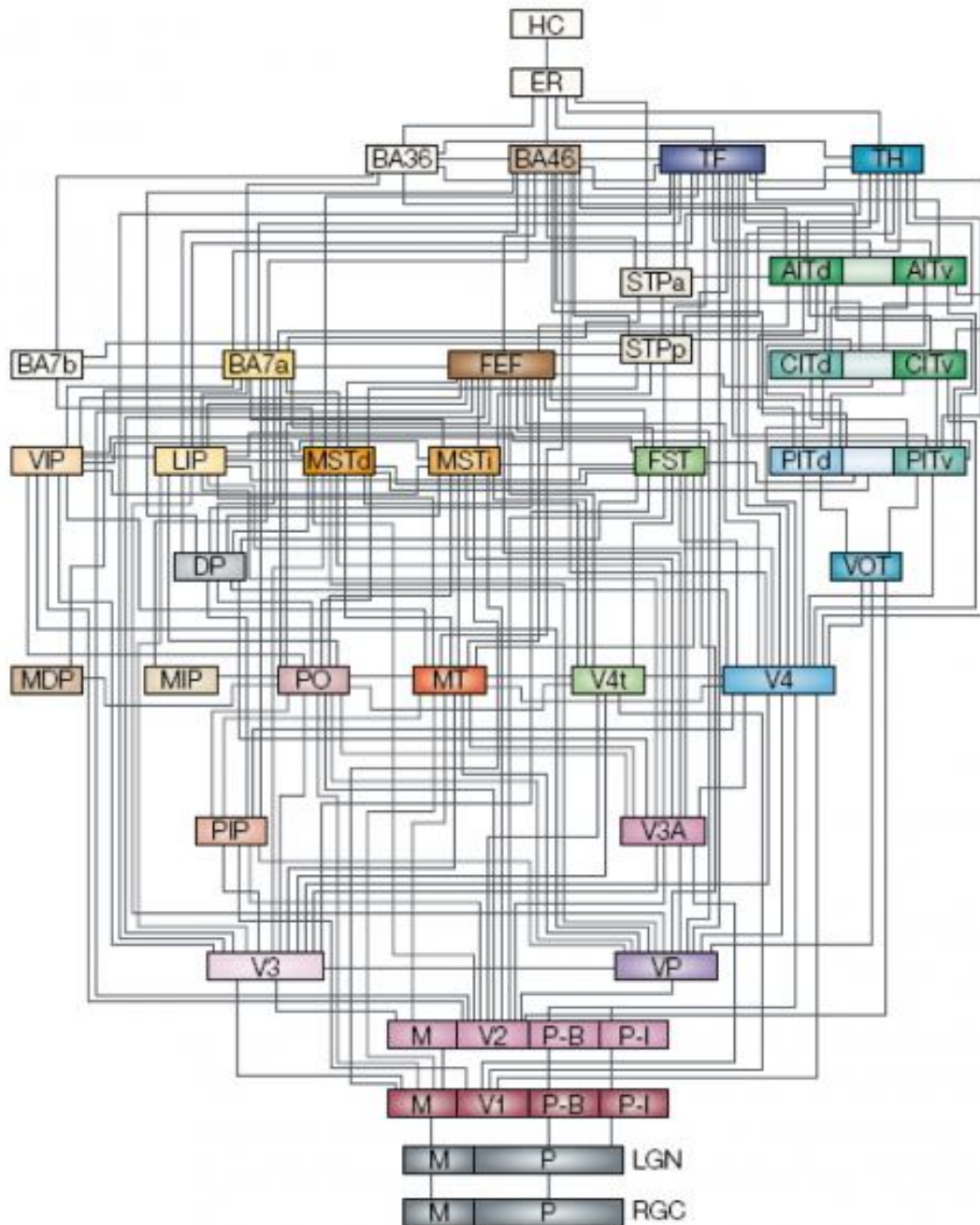
## Cortical visual system is comprised of many areas

- Visual cortical areas hierarchically higher than V1 span the occipital lobe as well as parietal lobe and part of the temporal lobe of the brain.



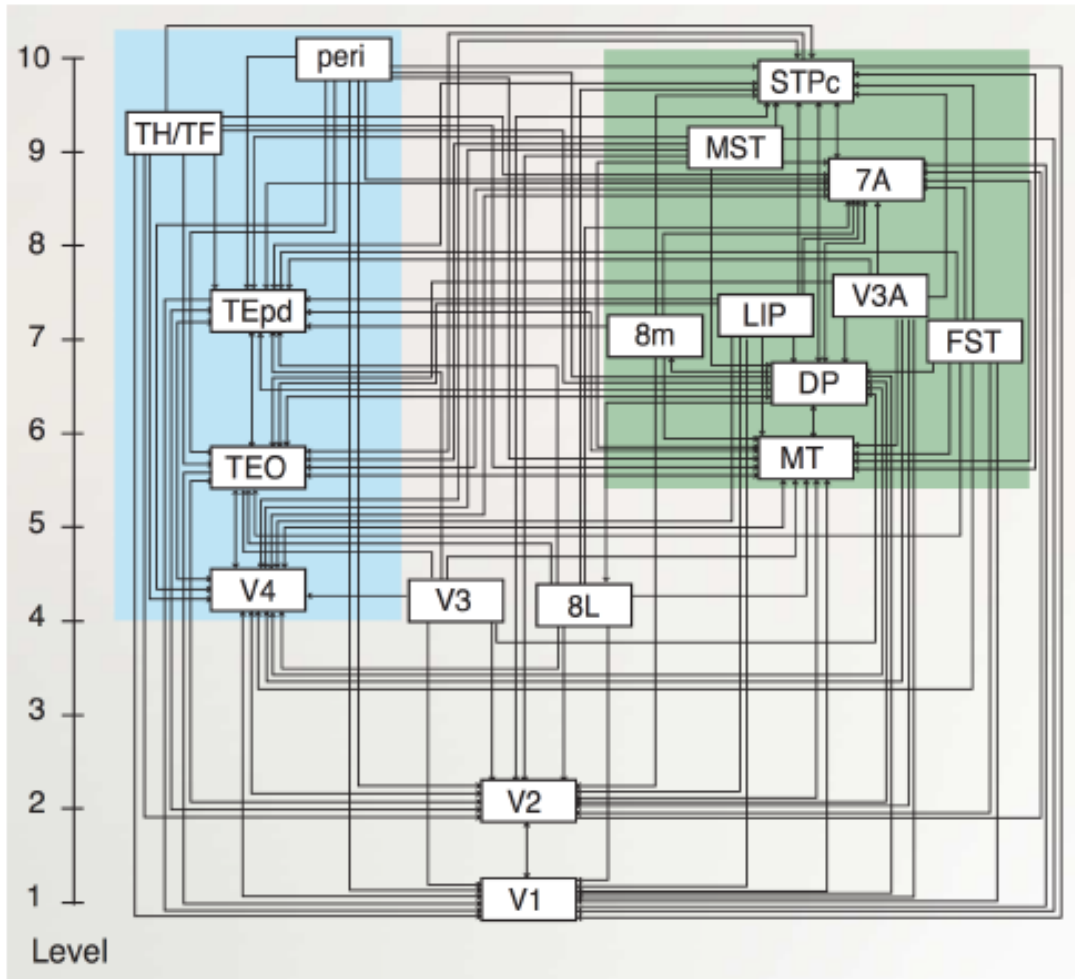


# Hierarchically organized visual system



- Felleman & Van Essen's (1991) diagram of the anatomical connectivity of visual processing pathways, starting with retinal ganglion cells (RGC) to the LGN, then primary visual cortex (V1) and on up.
- Scientists distinguish **about 40 hierarchically organised visual areas in the brain.**

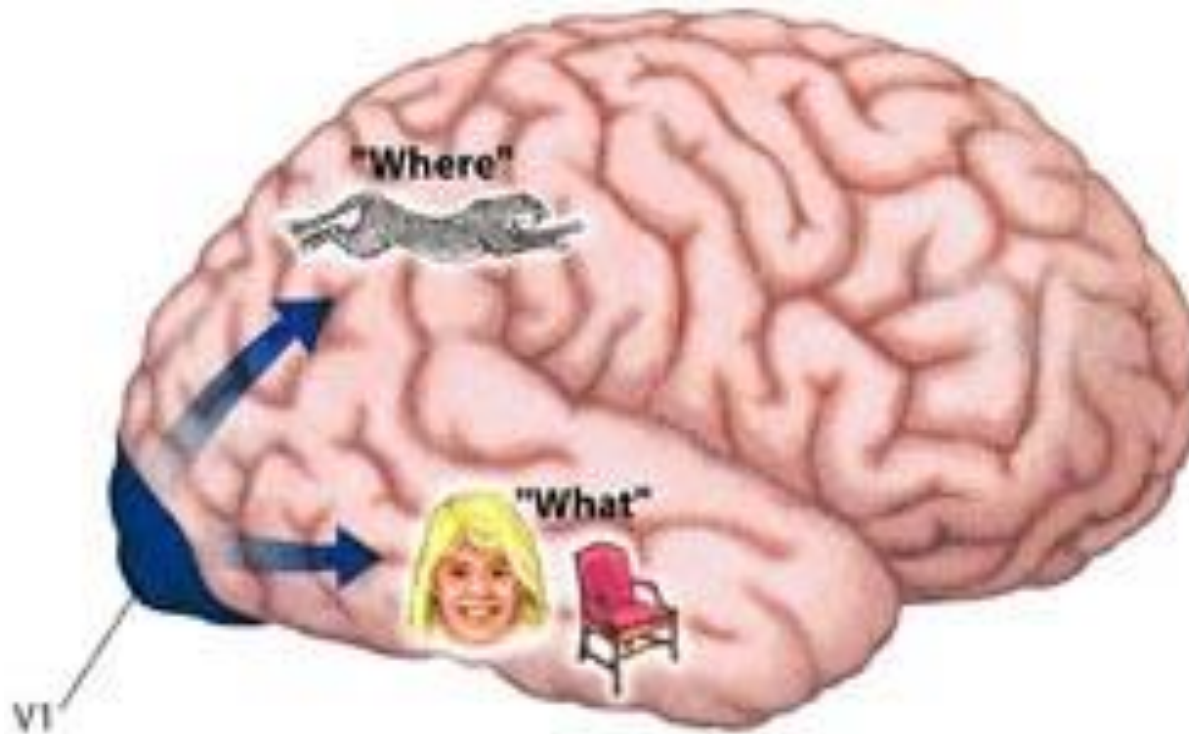
# Hierarchically organized visual system



- Updated and simplified version of Felleman & Van Essen's (1991) diagram of the anatomical connectivity of visual processing pathways, starting with primary visual cortex (V1) and on up.
- The blue-shaded areas comprise the ventral 'What' pathway, and green-shaded are the dorsal 'Where'.  
(Reproduced from Markov et al. 2014).

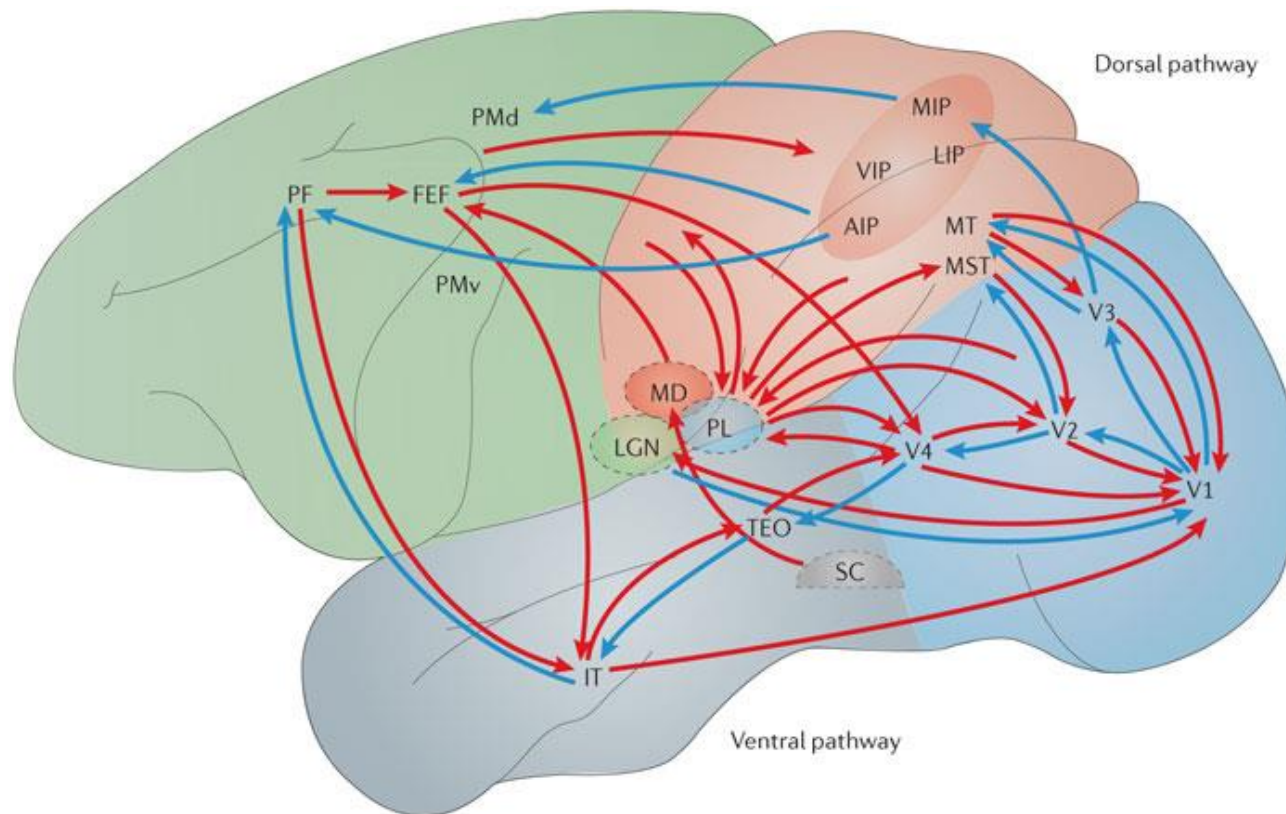
# What vs where (ventral vs dorsal) split in visual processing

- The projections going in a ventral direction from V1 are important for recognizing the identity ("what") of objects.
- The projections going up in a dorsal direction from V1 through parietal cortex extract spatial ("where") information: if a puma was jumping at us from the tree, we need to react fast, therefore where is separated from what.



## What vs where (ventral vs dorsal) split in visual processing

- The projections going from V1 to V4 to areas of inferotemporal cortex (IT) via TE and TEO, are important for recognizing the identity ("what") of objects.
- The projections going up through parietal cortex extract spatial ("where") information, including motion signals in area MT and MST.















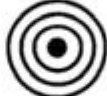





















## The main areas in the WHAT pathway

- **V1** – encodes oriented edge detectors that respond to edges (transitions in illumination) along different angles of orientation.
- **V2** – encodes combinations of edge detectors to respond to intersections and junctions, along with many other basic visual features (e.g., 3D depth selectivity, basic textures, etc.) along with a broader range of locations.
- **V4** – detects more complex shape features, over an even larger range of locations (and sizes, angles, etc).
- **IT-posterior (PIT)** – detects entire object shapes, over a wide range of locations, sizes, and angles. IT neurons are able to recognize an object regardless of where it appears in the visual field (i.e., invariant object recognition).
  - In addition, there is an area near the fusiform gyrus on the bottom surface of the temporal lobe, called the **fusiform face area (FFA)**, that appears especially responsive to faces.
- **IT-anterior (AIT)** – this is where visual information becomes extremely abstract and semantic in nature – it can encode all manner of important information about different people, places and things.

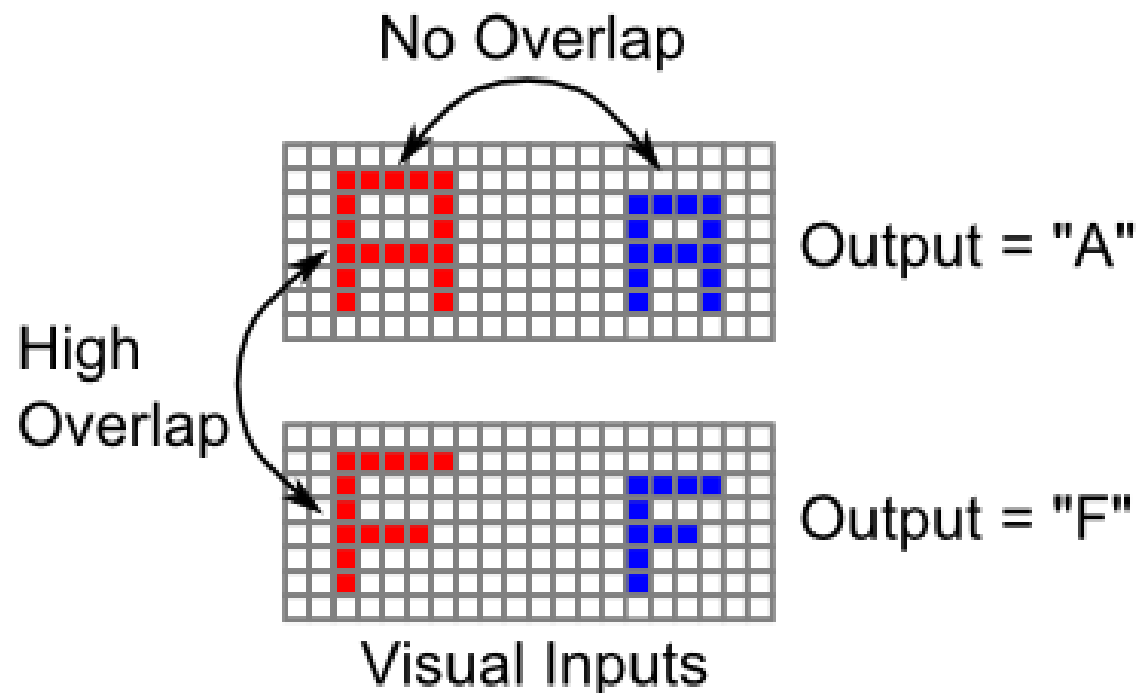
## The WHAT visual processing

- Complex stimuli that evoke a maximal response from neurons in V2, V4, and IT, providing some suggestion for what kinds of complex features these neurons can detect. These areas have something like a “library of shapes” (Reproduced from Kobatake & Tanaka 1994).

V2	V4	posterior IT	anterior IT
 	 	 	 
 	 	 	 
 	 	 	 
 	 	 	 

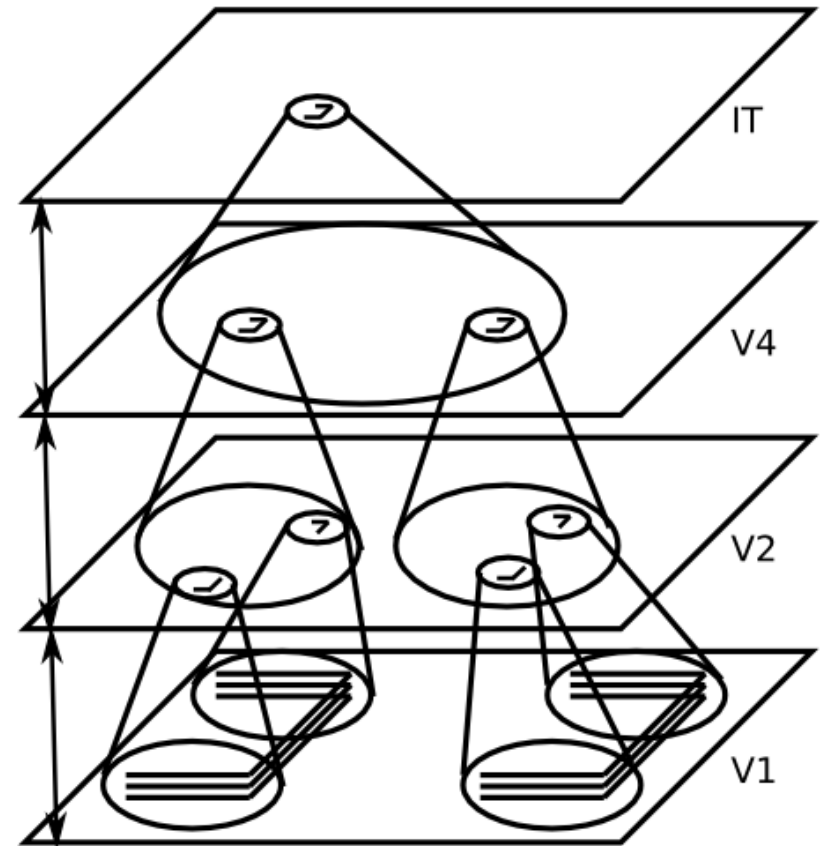
## Why object recognition is hard?

- Things that should be categorized as the same (i.e., have the same output label) often have no overlap in their retinal input features when they show up in different locations, sizes, etc., but things that should be categorized as different often have high levels of overlap when they show up in the same location.
- Thus, the bottom-up similarity structure is directly opposed to the desired output similarity structure, making the problem very difficult.



# Invariant object recognition in the WHAT pathway

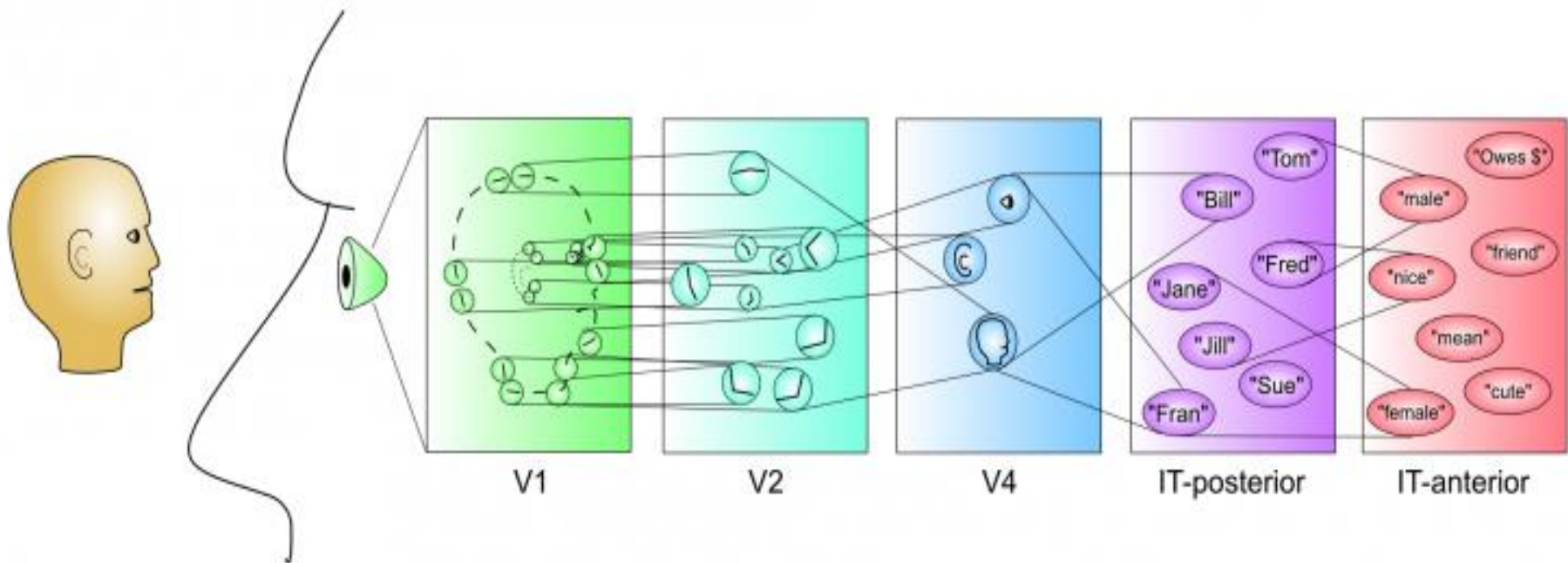
- Schematic for how multiple levels of processing can result in invariant object recognition, where an object can be recognized at any location.
- Each level of processing incrementally increases the complexity of features and spatial invariance of what it detects.
- Doing this incrementally allows the system to appropriately bind together features and their relationships, while also gradually building up overall spatial, size, etc. invariance.





# The WHAT visual processing

- Another way of representing the hierarchy of the ventral WHAT visual pathways. V1 has elementary feature detectors (oriented edges). Next, these are combined into combinations of lines in V2, followed by more complex visual features in V4.
- Individual faces (and other objects) are recognized at the next level in IT. Finally, at the highest level are important semantic categories – this level corresponds to more anterior areas of IT.



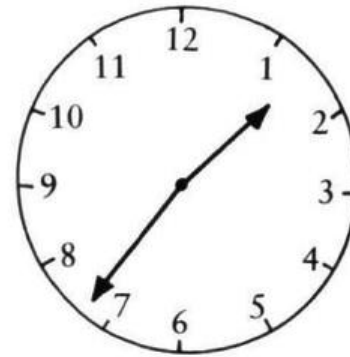
# The WHERE visual processing

- This neural pathway mediates the ability to direct the spotlight of attention to a particular spatial location.
- This depends on an overall **representations of space in the parietal cortex**, which interacts with lower levels of the object recognition pathway (V1, V2, V4) to constrain the inputs to those visual features that come from within this spotlight of attention.
  - i.e. the WHAT and WHERE pathways work together. WHERE focuses attention and WHAT identifies what's there.
- One crucial function of spatial attention is to enable object recognition in visual scenes that have multiple different objects present (e.g. a lion in the bush, an enemy among the crowd, etc.).

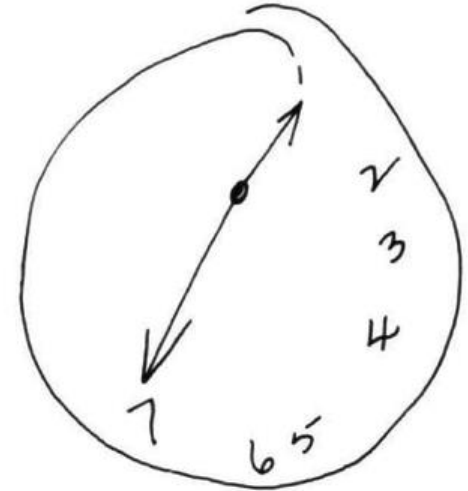
# Hemi-spatial neglect (Half-space ignorance)

- The most striking evidence that the **parietal cortex** (WHERE pathway) is important for spatial attention comes from patients with **hemispatial neglect**, who tend to ignore (neglect) one side of space.
- This condition arises from a damage affecting the parietal cortex, which then gives rise to a neglect of the opposite half of space.

MODEL



PATIENT'S COPY



MODEL



PATIENT'S COPY



## Hemi-spatial neglect (Half-space ignorance)

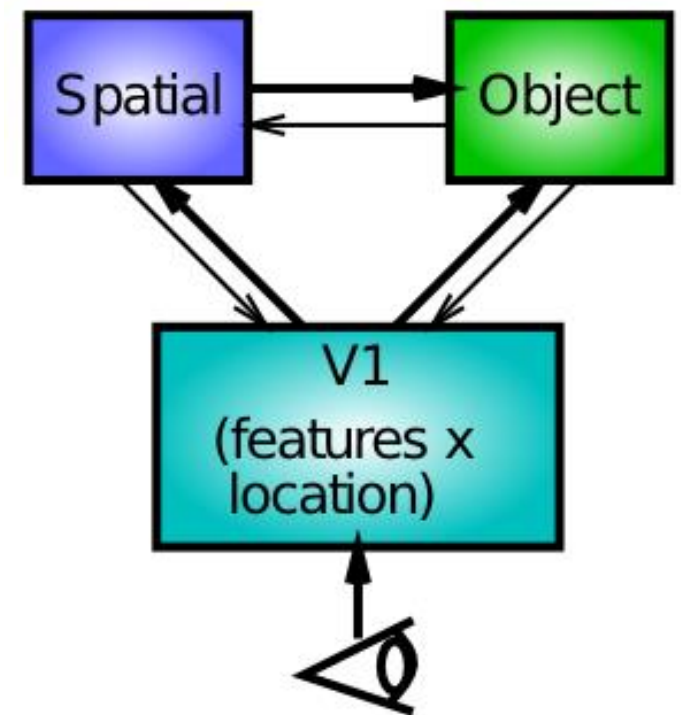
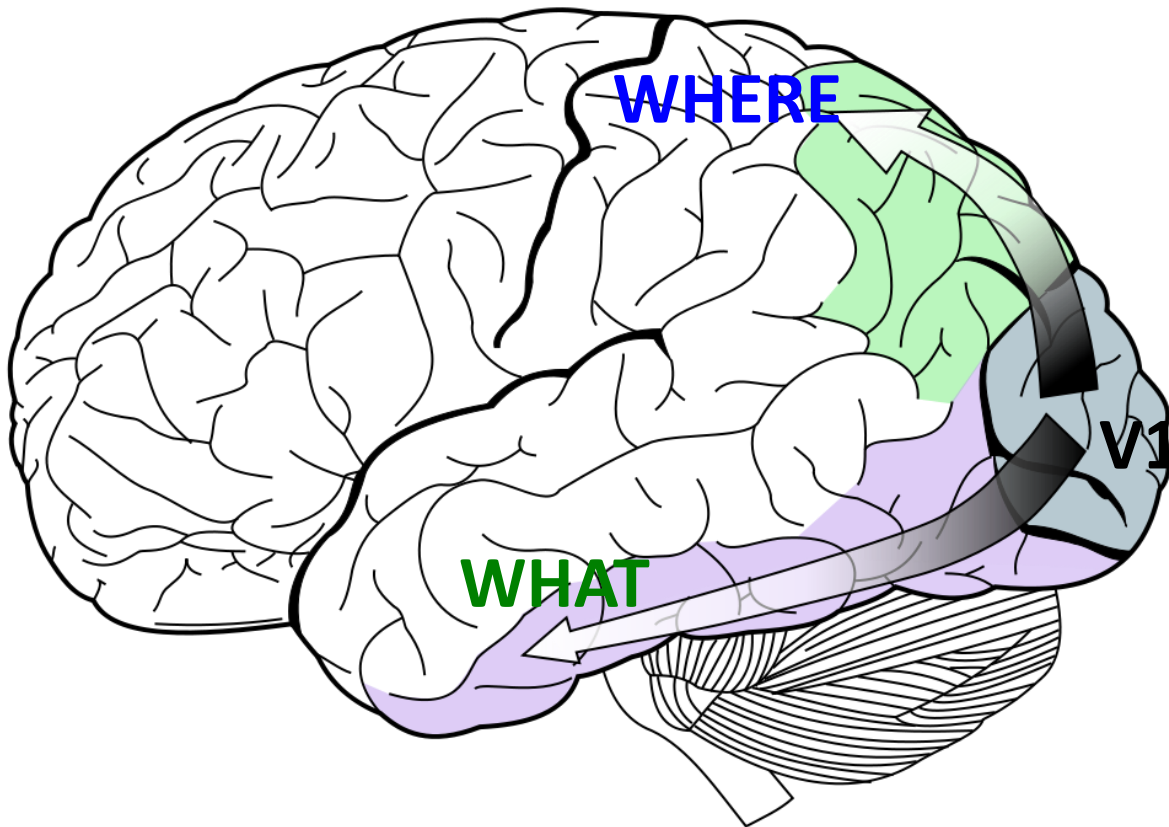
- The patient, an artist, had a stroke damaging the right parietal cortex.
- Thus, at first he was unable to draw the left side of himself.
- As he gradually recovered, and brain plasticity healed the damaged parietal cortex, he regained the capacity to perceive the left side of space.





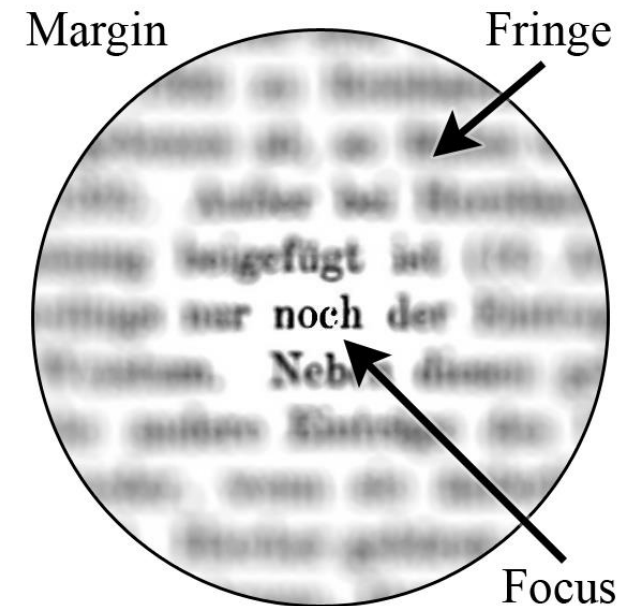
# Summary

- The ventral stream (the "WHAT pathway") is involved with object visual identification and recognition. The dorsal stream (the "WHERE pathway") is involved with processing the object's spatial location relative to the viewer. They interact together.



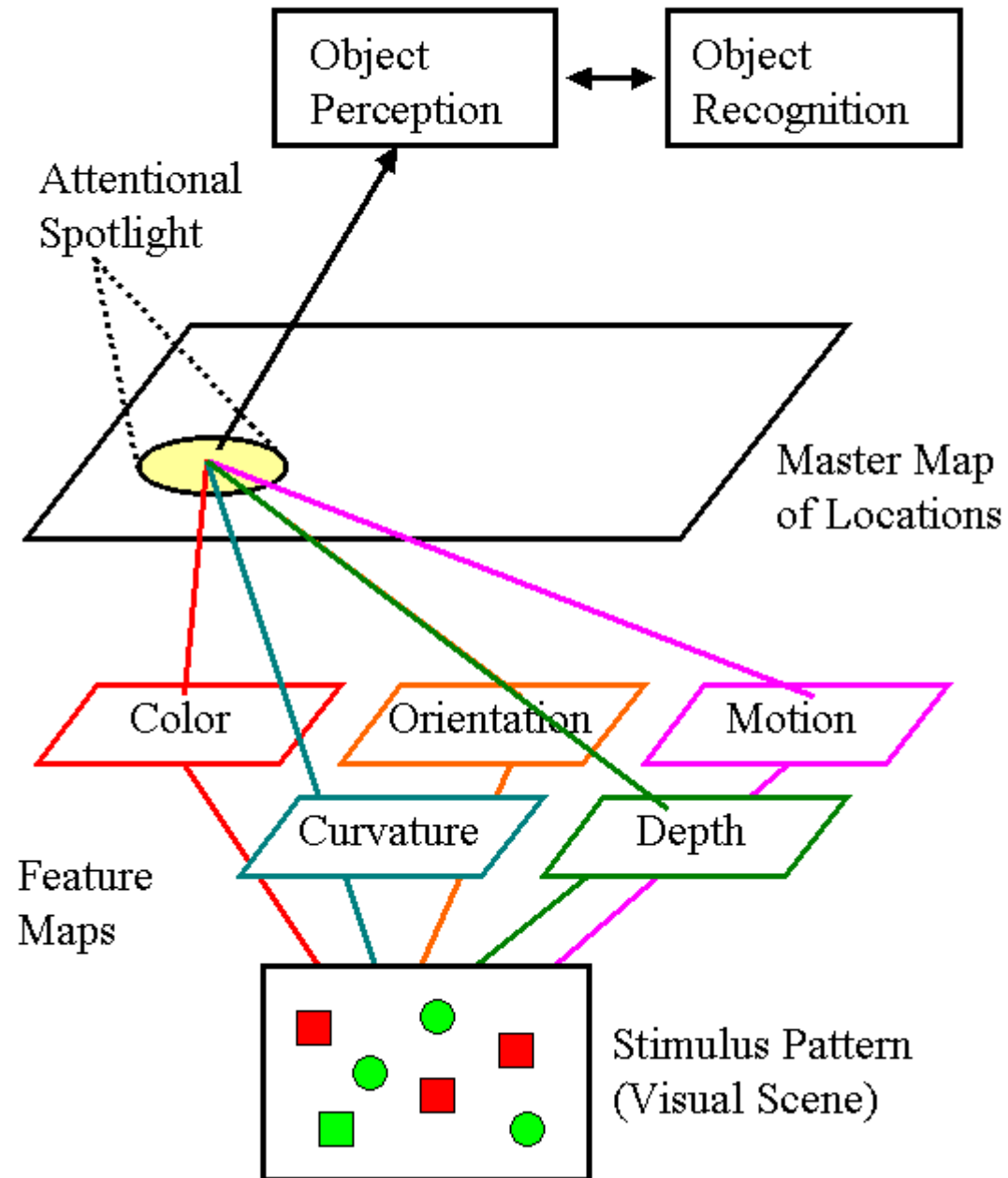
# Attention

- It is a cognitive process of selectively concentrating on a discrete aspect of information, whether subjective or objective, while ignoring other perceivable information.
- The term **"spotlight"** was introduced by William James, who described attention as having a focus, a margin, and a fringe. The focus is an area that extracts information from the visual scene with a high-resolution, the geometric center of which being where visual attention is directed.
- Surrounding the focus is the fringe of attention, which extracts information in a much more crude fashion (i.e., low-resolution).



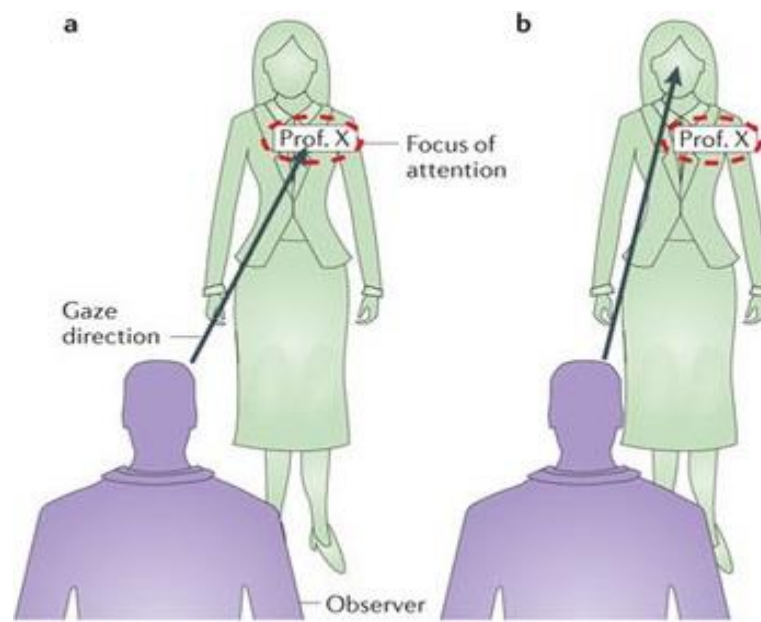
## Feature Integration Theory (Treisman)

- Objects are retrieved from scenes by means of selective spatial attention that picks out objects' features, forms feature maps, and integrates those features that are found at the same location into forming objects.
- **Preattentive Stage:** The unconscious detection and separation of features of an item (color, shape, size).
- **Focused Attention Stage:** The combining of all feature identifiers to perceive all parts as one whole. This is possible through prior knowledge.



## Overt and covert orienting

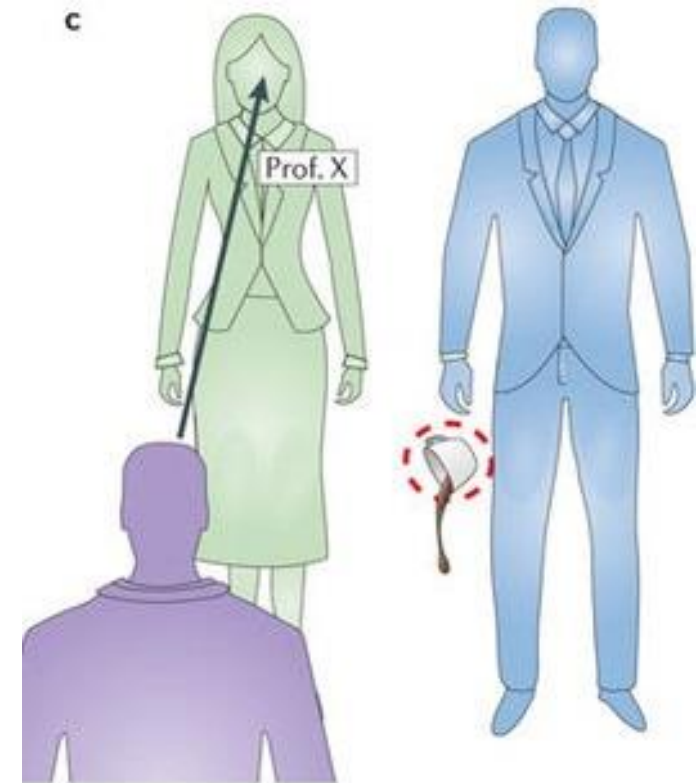
- **Overt orienting** is the act of selectively attending to an item or location over others by moving the eyes to point in that direction. Two types of eye movements; **reflexive and controlled**. Reflexive movements are commanded by the superior colliculus of the midbrain. These movements are fast and are activated by the sudden appearance of a stimulus. In contrast, controlled eye movements are commanded by areas in the frontal lobe. These movements are slow and voluntary.
- **Covert orienting** is the act of mentally shifting one's focus without moving one's eyes (V4 neurons).





# Exogenous and endogenous orienting

- **Exogenous orienting** is described as being under control of a stimulus. Exogenous orienting is considered to be reflexive and automatic and is caused by a sudden change in the periphery (peripheral cues). This often results in a reflexive saccade.
- **Endogenous orienting** is the intentional allocation of attentional resources to a predetermined location or space. Endogenous orienting occurs when attention is oriented according to an observer's goals or desires, allowing the focus of attention to be manipulated by the demands of a task (central cues).



## Processes in attention (Knudsen)

- General neural correlate of attention is an enhanced firing. If a neuron has a certain response to a stimulus when the animal is not attending to the stimulus, then when the animal does attend to the stimulus, the neuron's response will be enhanced even if the physical characteristics of the stimulus remain the same.
- **Working memory** temporarily stores information for detailed analysis.
- **Competitive selection** is the process that determines which information gains access to the working memory (lateral inhibition).
- **Top-down control**: higher cognitive processes can regulate signal intensity in information channels that compete for access to working memory, and thus give them an advantage in competitive selection.
- **Bottom-up saliency filters** automatically enhance the response to infrequent stimuli, or stimuli of instinctive or learned biological relevance.

## Neural correlates (Knudsen)

- At the top of the hierarchy are the **frontal eye fields (FEF)** and the dorsolateral prefrontal cortex. Microstimulation in the FEF induces monkeys to make a saccade to the relevant location.
- At the next lower level, a variety of spatial maps are found in the **parietal cortex**. In particular, the lateral intraparietal area (LIP) contains a saliency map and is interconnected both with the FEF and with sensory areas.
- Exogenous attentional guidance in humans and monkeys is by a **bottom-up saliency** map in the **primary visual cortex**.
- Certain automatic responses that influence attention, like orienting to a highly salient stimulus, are mediated subcortically by the **superior colliculus**.

## Other attributes of attention

- **Focused attention:** The ability to respond discretely to specific visual, auditory or tactile stimuli.
- **Sustained attention** (vigilance and concentration): The ability to maintain a consistent behavioral response during continuous and repetitive activity.
- **Selective attention:** The ability to maintain a behavioral or cognitive set in the face of distracting or competing stimuli.
- **Alternating attention:** The ability of mental flexibility that allows individuals to shift their focus of attention and move between tasks having different cognitive requirements.
- **Divided attention:** This refers to the ability to respond simultaneously to multiple tasks or multiple task demands.
- **Mindfulness** is the practice of purposely bringing one's attention in the present moment without judgment, a skill one develops through meditation or other training.