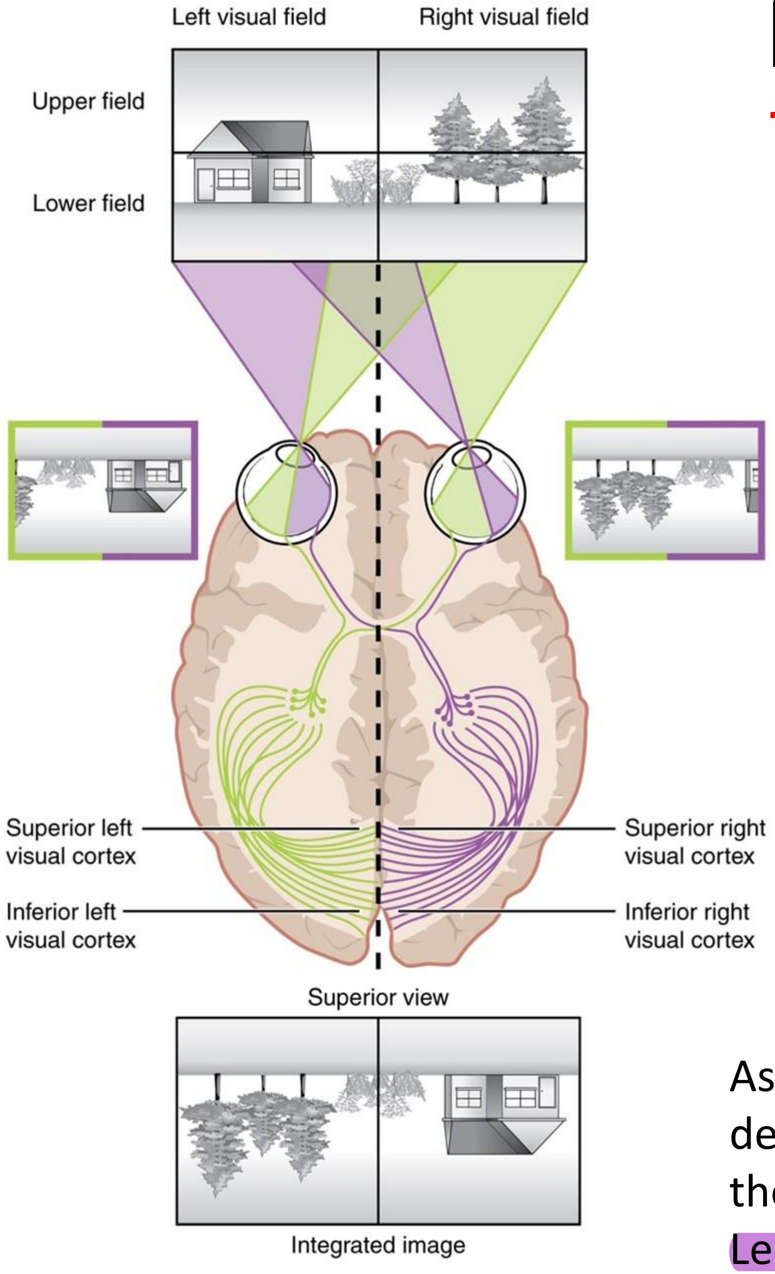
Cognitive Science and AI

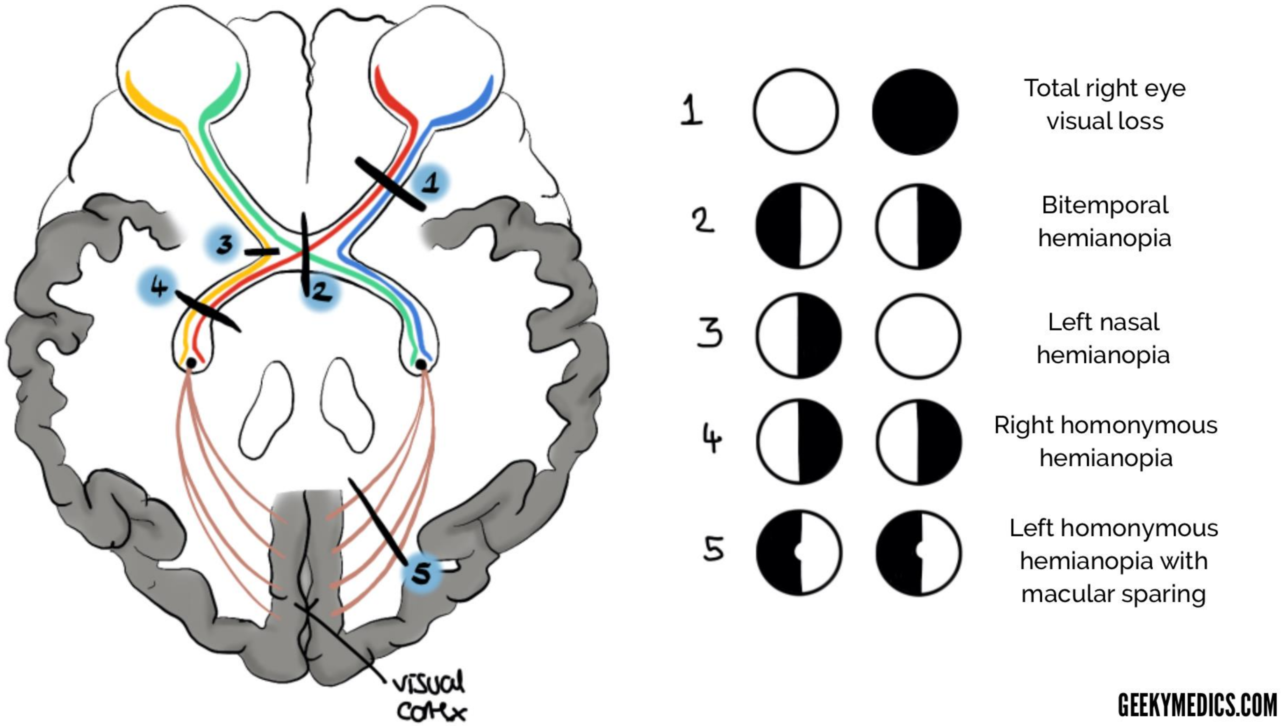
Intermediate-Level Perception: Representation Spaces

Presentation from: Shimon Edelman, Computing the Mind, 2008 (Chapter 5: Perception) & PNS (6 Ed) Ch 23

Recap



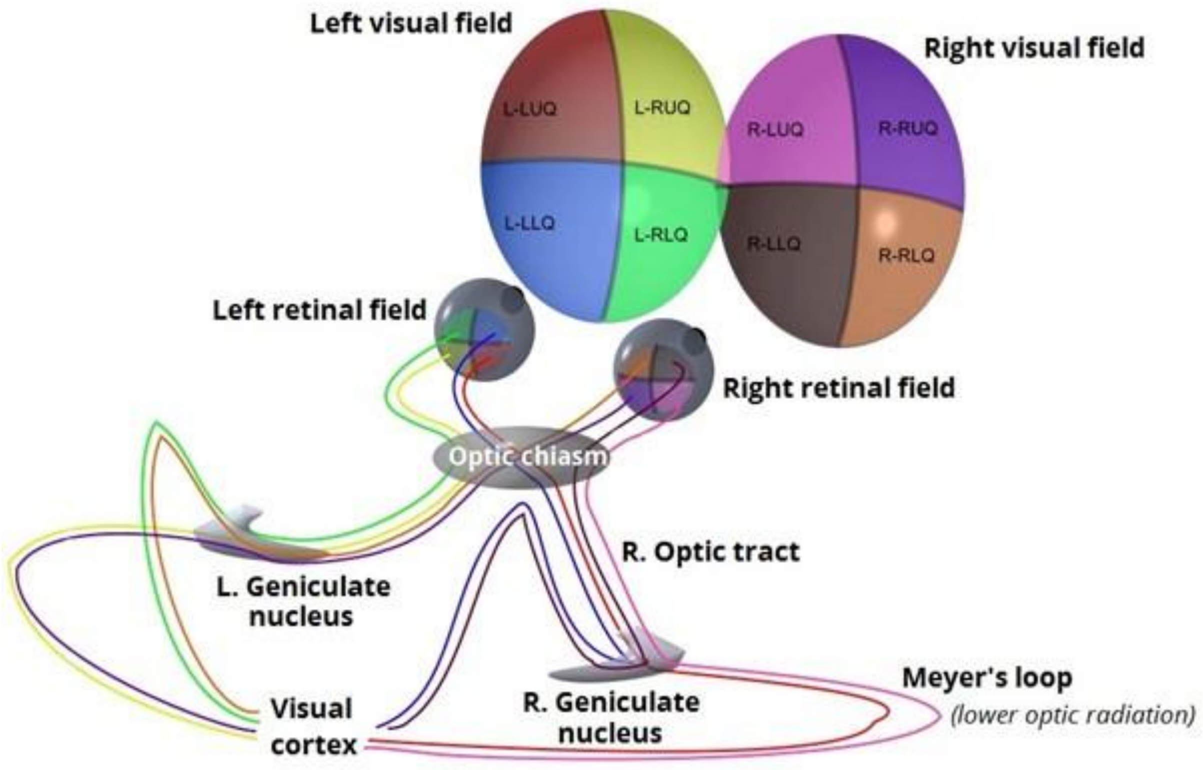
Visual field defects

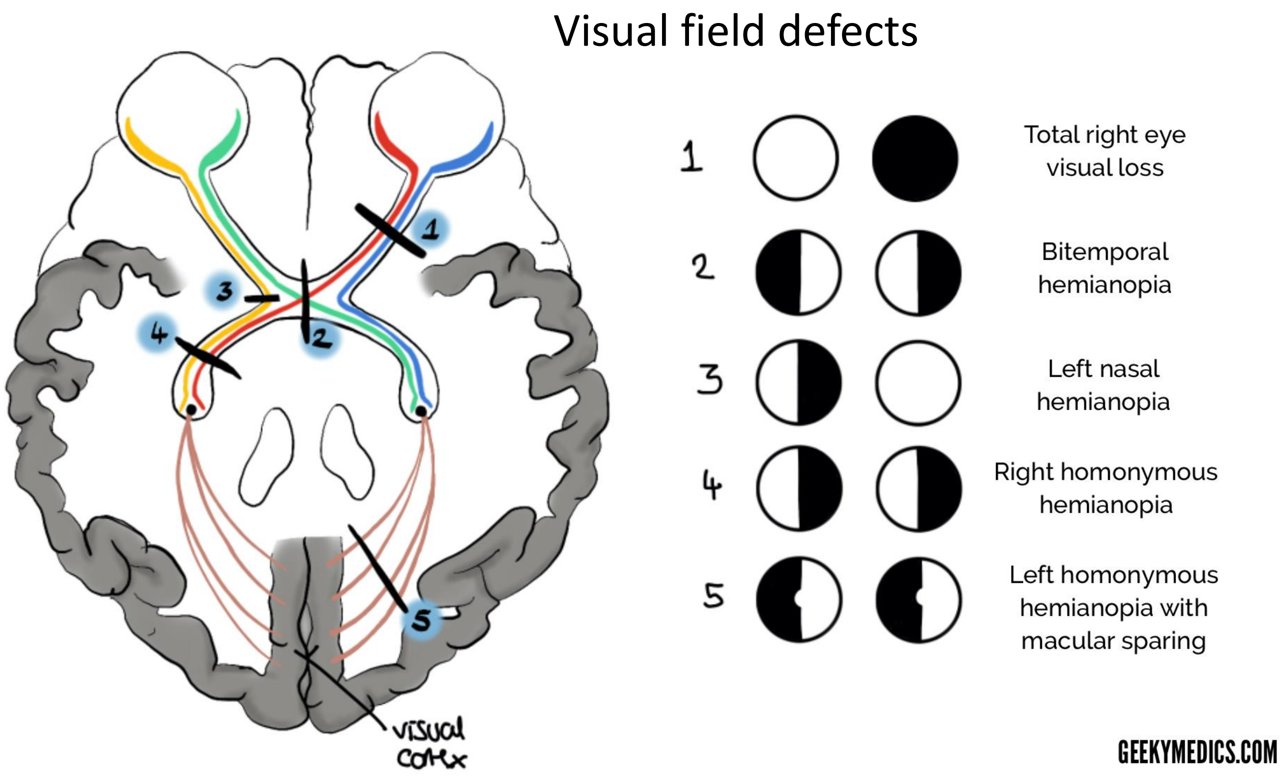


As a rule, pre-chiasmal lesions will result in an ipsilateral monocular visual field defect. Post-chiasmal lesions will result in homonymous visual field defects of the contralateral side.

Lesions of the chiasm most commonly result in bitemporal hemianopia

Recap



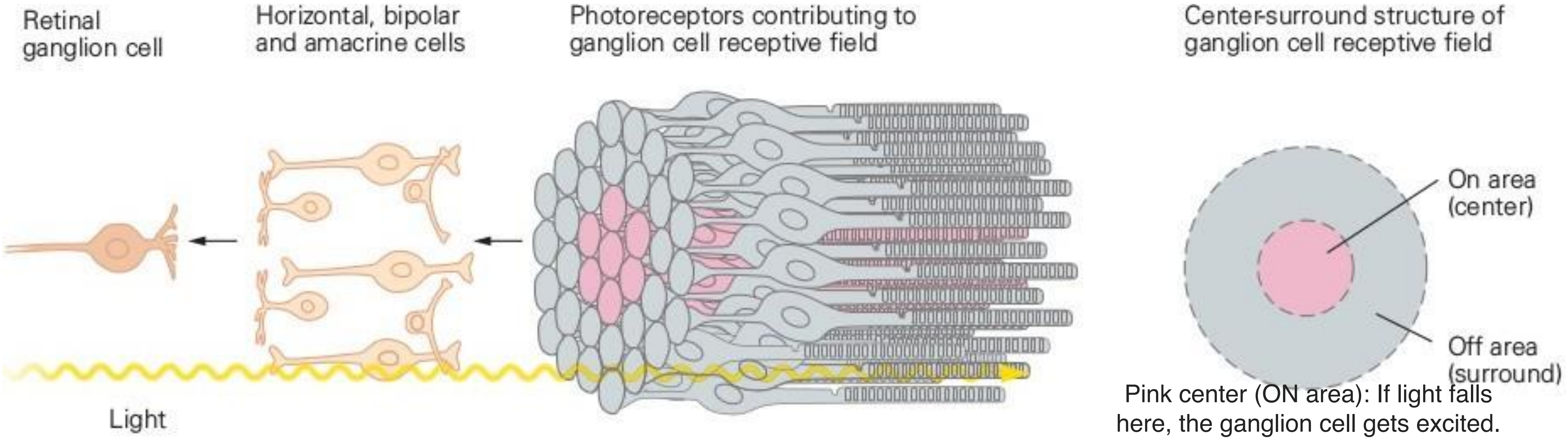


As a rule, pre-chiasmal lesions will result in an ipsilateral monocular visual field defect. Post-chiasmal lesions will result in homonymous visual field defects of the contralateral side.
Lesions of the chiasm most commonly result in bitemporal hemianopia

Receptive Fields

A receptive field is the area in the visual space (or retina) that affects the activity of a single neuron. In this case, it’s the Retinal Ganglion Cell (a type of neuron in the eye).

Receptive Field of a Retinal Ganglion Cell



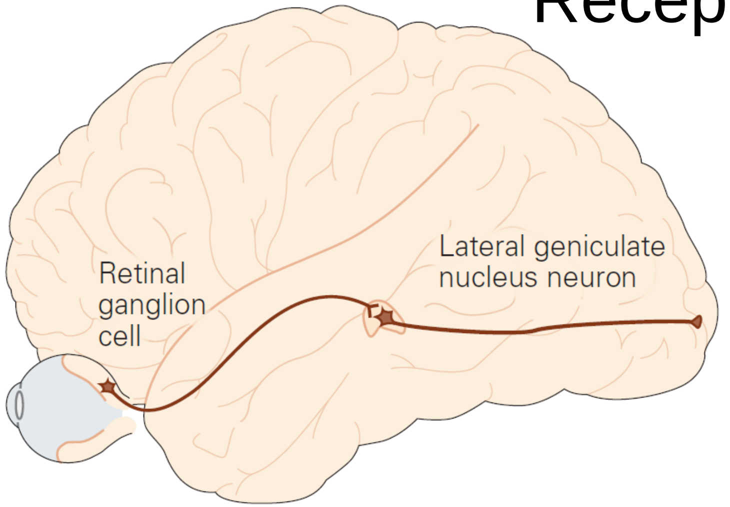
Pink center (ON area): If light falls here, the ganglion cell gets excited.

Gray surrounding area (OFF area): If light falls here, the ganglion cell gets inhibited (reduced activity).

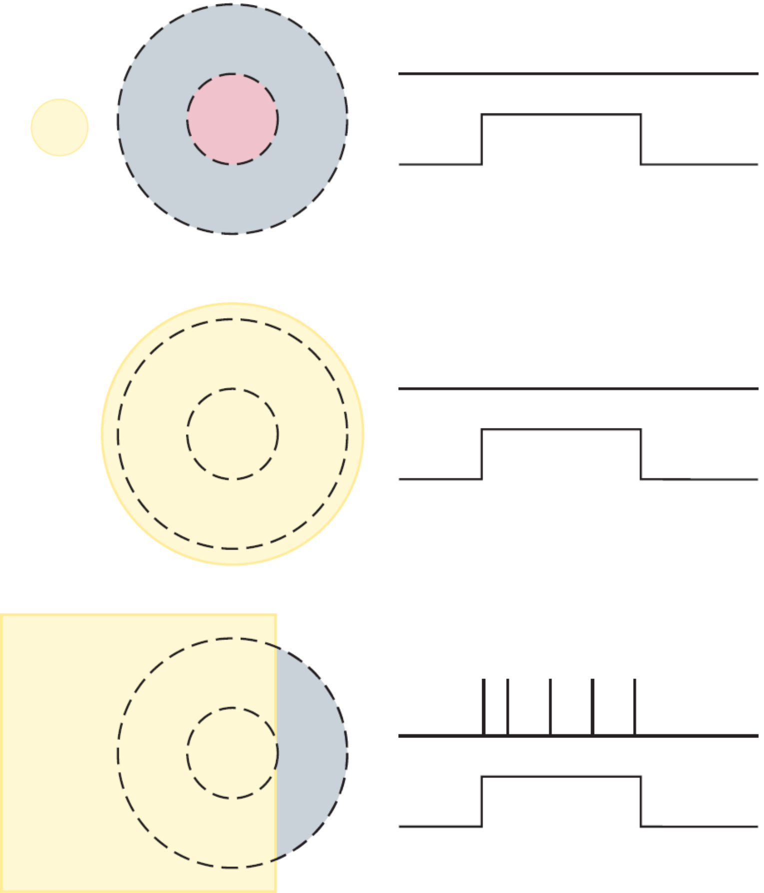
Gray surrounding area (OFF area): If light falls here, the ganglion cell gets inhibited (reduced activity).

Receptive Fields

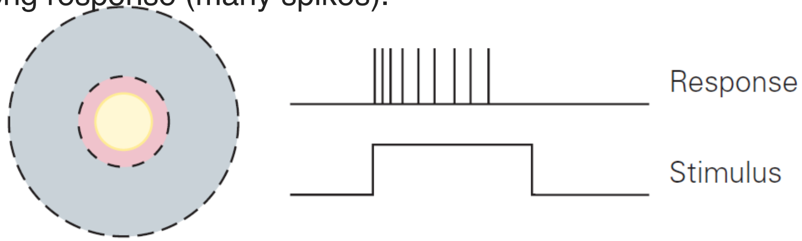
The yellow part represents the light stimulus,



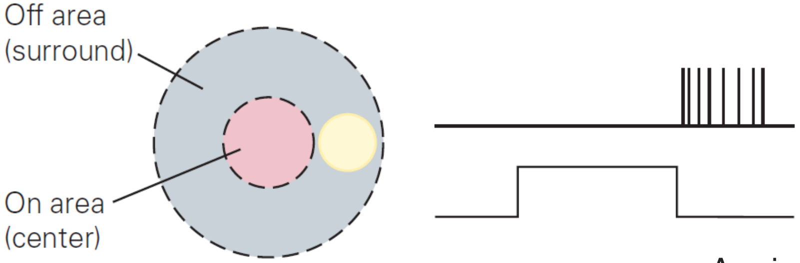
Contrast (light-dark boundary) Specificity



A light spot in the center excites the neuron → strong response (many spikes).



Neurons in early vision respond to small spots of light/dark.

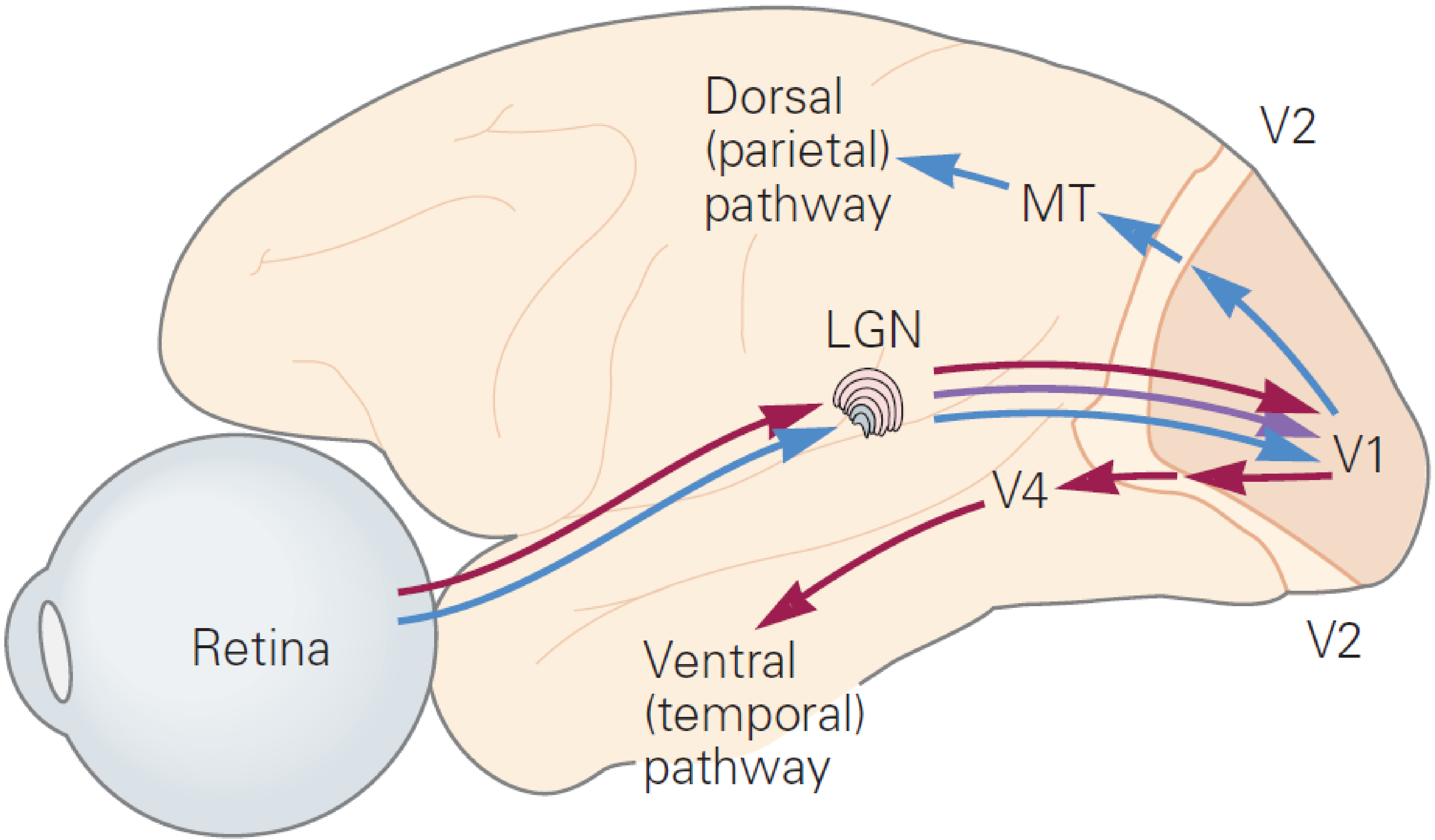


As visual processing continues, receptive fields get larger and focus on patterns and edges.

Receptive Field size increases through the visual processing hierarchy

Dorsal & Ventral Visual Pathways

Dorsal pathway (blue, to the parietal lobe): Focuses on “where” objects are—motion and spatial relationships.



Ventral pathway (red, to the temporal lobe): Focuses on “what” objects are—recognition and identification.

Serial and Parallel Processing in Visual Pathways

Visual information is processed in two ways:

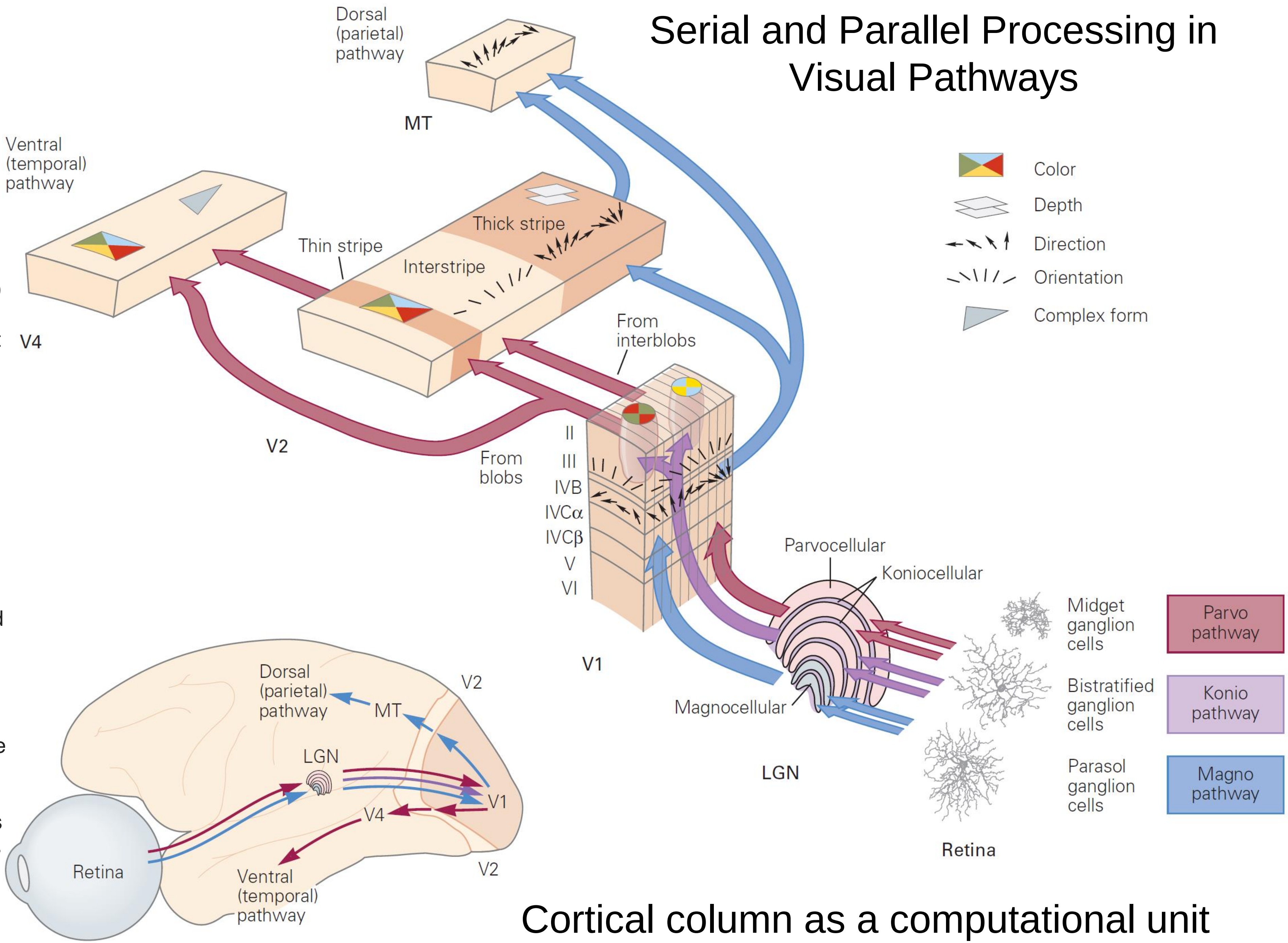
Serial processing: Information moves step by step through different brain areas.

Parallel processing: Multiple features (color, motion, shape) are processed at the same time in different pathways.•

The dorsal (“where”) pathway (blue) handles motion and spatial location.•

The ventral (“what”) pathway (red) processes object recognition, color, and shape.

The LGN (Lateral Geniculate Nucleus) relays signals from the eye to the visual cortex (V1), then different brain regions extract different types of visual details.



Cortical column as a computational unit

Cortical column as a computational unit

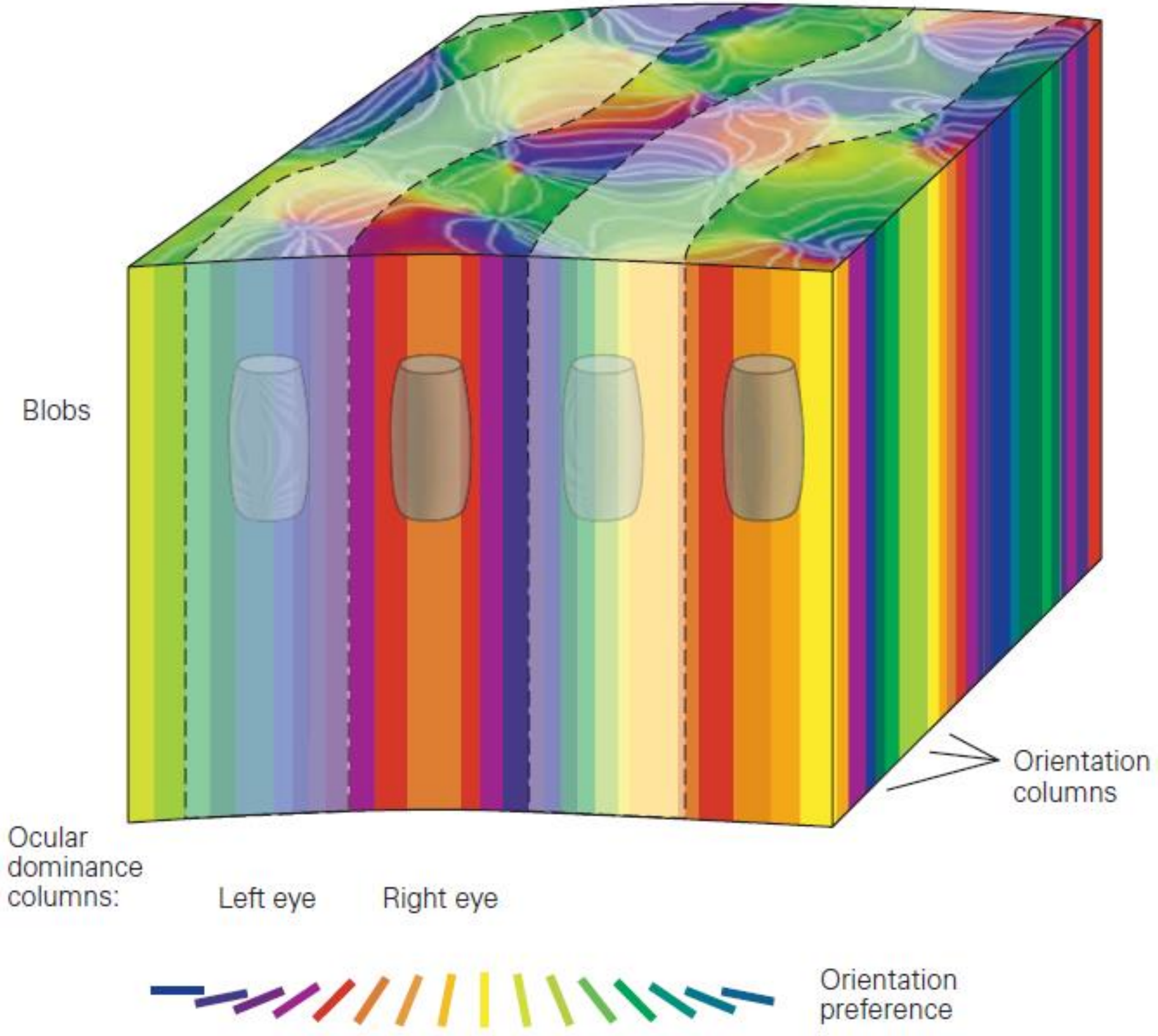
The visual cortex (V1) is organized into columns that process specific information.

Blobs (gray areas) –

process color.

Ocular dominance columns (striped areas) – neurons prefer input from one eye or the other.

Orientation columns (color-coded) – neurons respond to edges in specific directions.



Neural Codes and Visual Information Representation

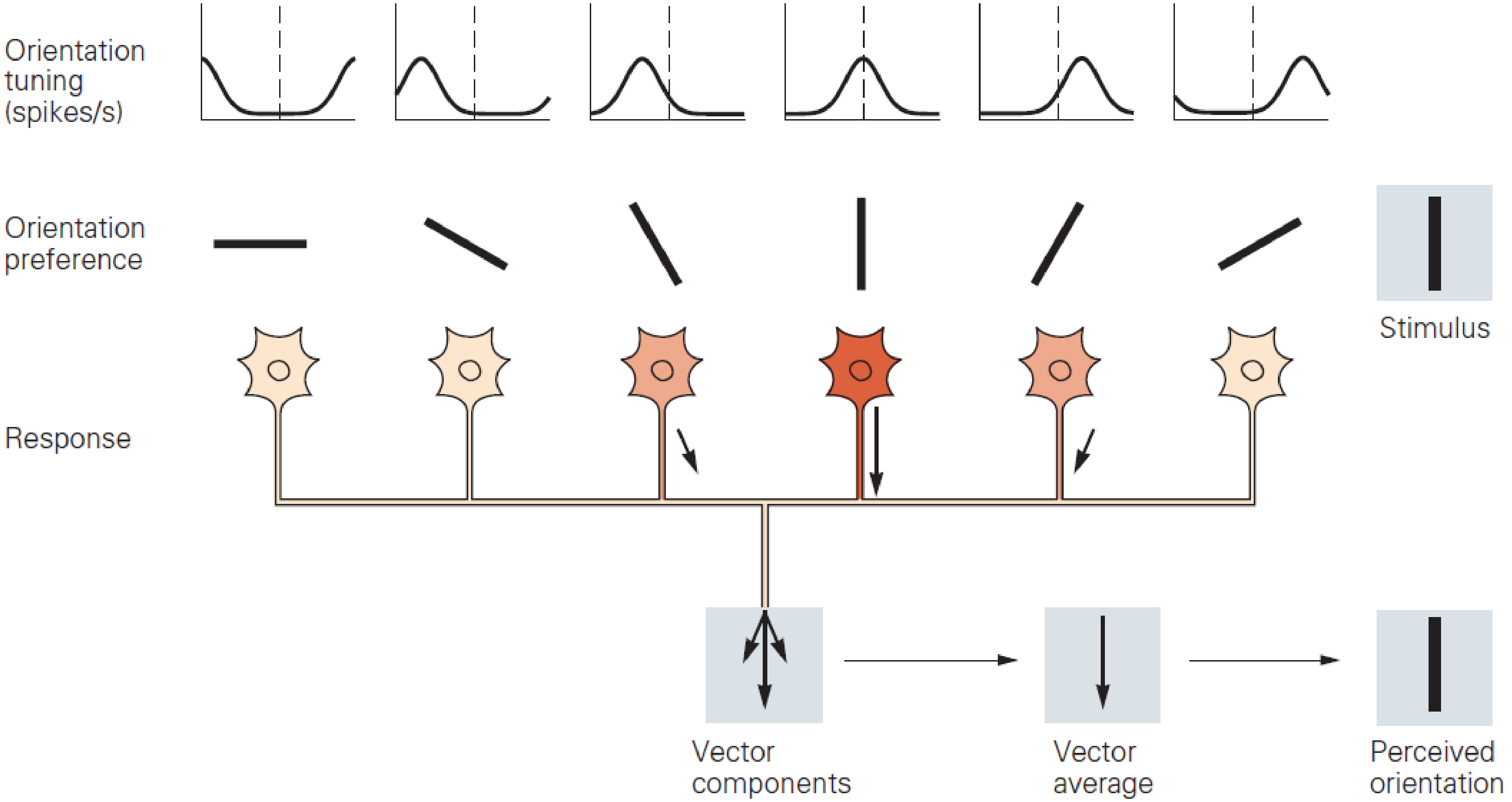
he brain represents visual information using groups of neurons.

Each neuron responds to a specific orientation (angle) of an edge.

Instead of a single neuron encoding a shape, the brain uses a group of neurons working together (population coding).

The vector averaging method helps combine responses to determine the overall perception of orientation.
Localist (grandmother) vs. distributed coding.

Localist coding = a single neuron recognizes a specific object (e.g., one neuron fires only for your grandmother).
Distributed coding = multiple neurons work together to recognize an object.



Population code: Vector Averaging

Localist (Grand mother) vs Distributed Code

An ensemble of neurons represents an entire object, each member may participate in different ensembles that are activated by different objects

Summary

- Form, Color, Motion, and Depth are processed in Discrete Areas of the Cerebral Cortex
- The receptive fields of neurons at successive relays in an afferent pathway provide clues to how the brain analyzes visual form
- The visual cortex is organized into columns of specialized neurons
- Intrinsic cortical circuits transform neural information
- Visual Information is represented by a variety of neural codes

Perception

- The phenomenal contents of perception — defy conceptualization – ineffable!
- To understand perception (as well as its relation to memory) we must, therefore, understand what the raw, phenomenal, pre-conceptual feel of sensing the world consists of, and how the sensory stimulation gets computed into that feel.

What is Perception?
Perception is how we experience the world through our senses.

It is difficult to fully describe because it feels so direct and natural.

Perceptual measurements and representations

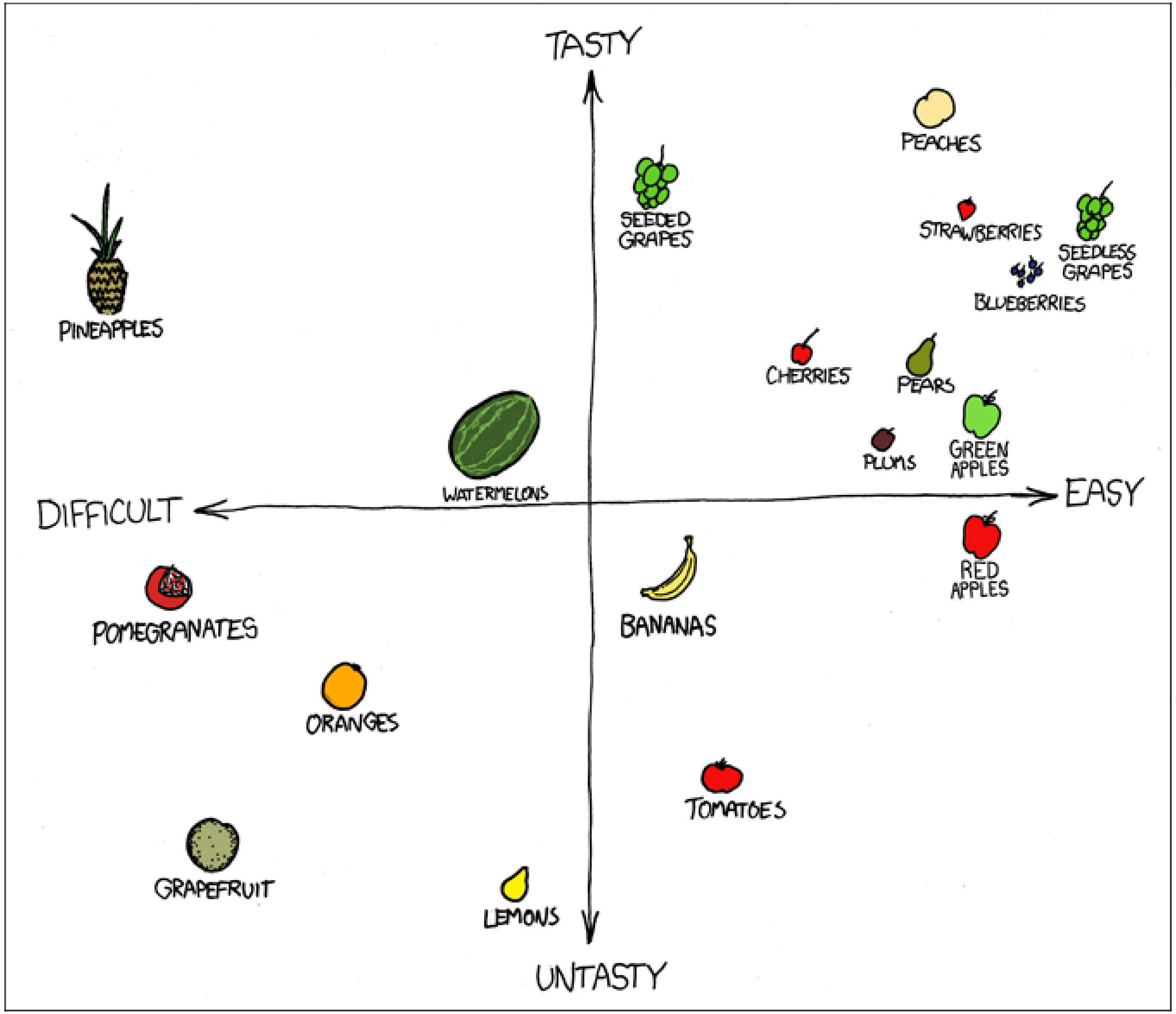
- Perception in the service of behaviour involves the brain performing many measurements on the outside world:
- The measurements are structured in space and time, and they carry information about the space-time structure of the world.
  - Example: The use of Inter-aural time difference (ITD) in sound localization in Barn Owls (intro lecture!)
  - Another one is shown in the next slide and also later.
- The resulting representation spaces are also structured.

How many dimensions?

- Dimensionality of representation space

= # of measurements per point

- The measurements populate a representation space.
  - It's a topological space (not merely a set) because intermediate points in it make sense too (like morphing).
  - Example: the "fruit space" has 2 dimensions – represented as 2 numbers per point.
  - Each point represents a single kind of item (apples, bananas, etc.).



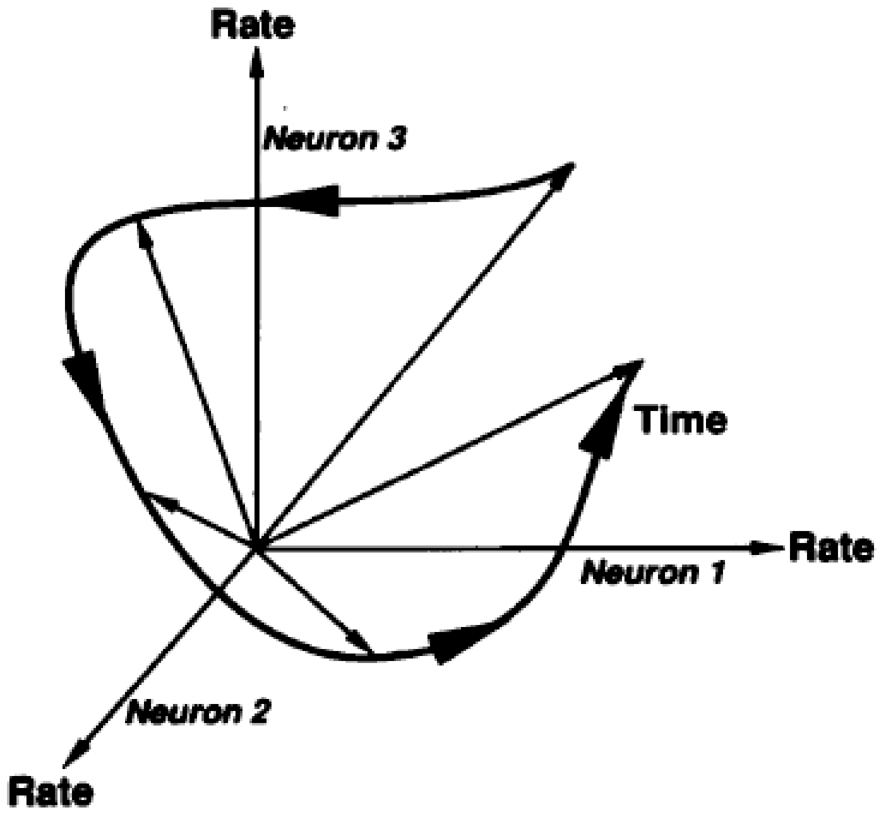
How many measurements?

- Consider a skier skiing down a hill...

- To represent this process, one must represent a function that maps
  - Time [the domain of the function].
- to
  - the state of the skier [the range of the function].

- How many dimensions does this function's range possess?

the state space of a 3-neuron brain



The trajectory — state plotted against time — of a three-neuron dynamical system through the space of its possible states.

Nominal and Effective dimensionality

- How many dimensions are there in the data that the eye sends to the brain?

Two-dot acuity task:

Measures how close two dots can be before they look like one dot instead of two.

The limit (threshold) for humans is about 30 arcseconds (a very small angle!).

- About 1,000,000!

Vernier acuity task:

Measures how well we detect a slight misalignment between two lines.

- Luckily, throughout cognition,

Humans can detect misalignments as small as 5 arcseconds—much finer than two-dot acuity!

EFFECTIVE dimensionality << NOMINAL dimensionality

More measurements = sharper vision.

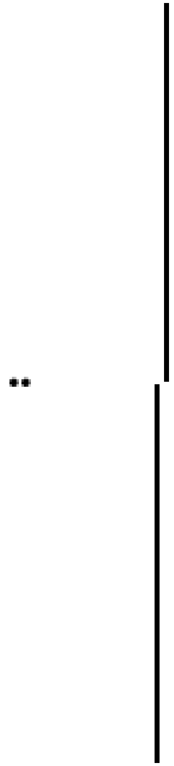
The way measurements are structured in space and time affects perception accuracy.

Vernier acuity is much more precise than simple two-dot acuity!

Importance of spatially structured measurements

- Dimensionality is about the number of measurements!
- Importance of spatially structured measurements: a visual task that illustrates this is acuity (sharpness)
- The spatial structure of the measurements is very important (as is their temporal structure).

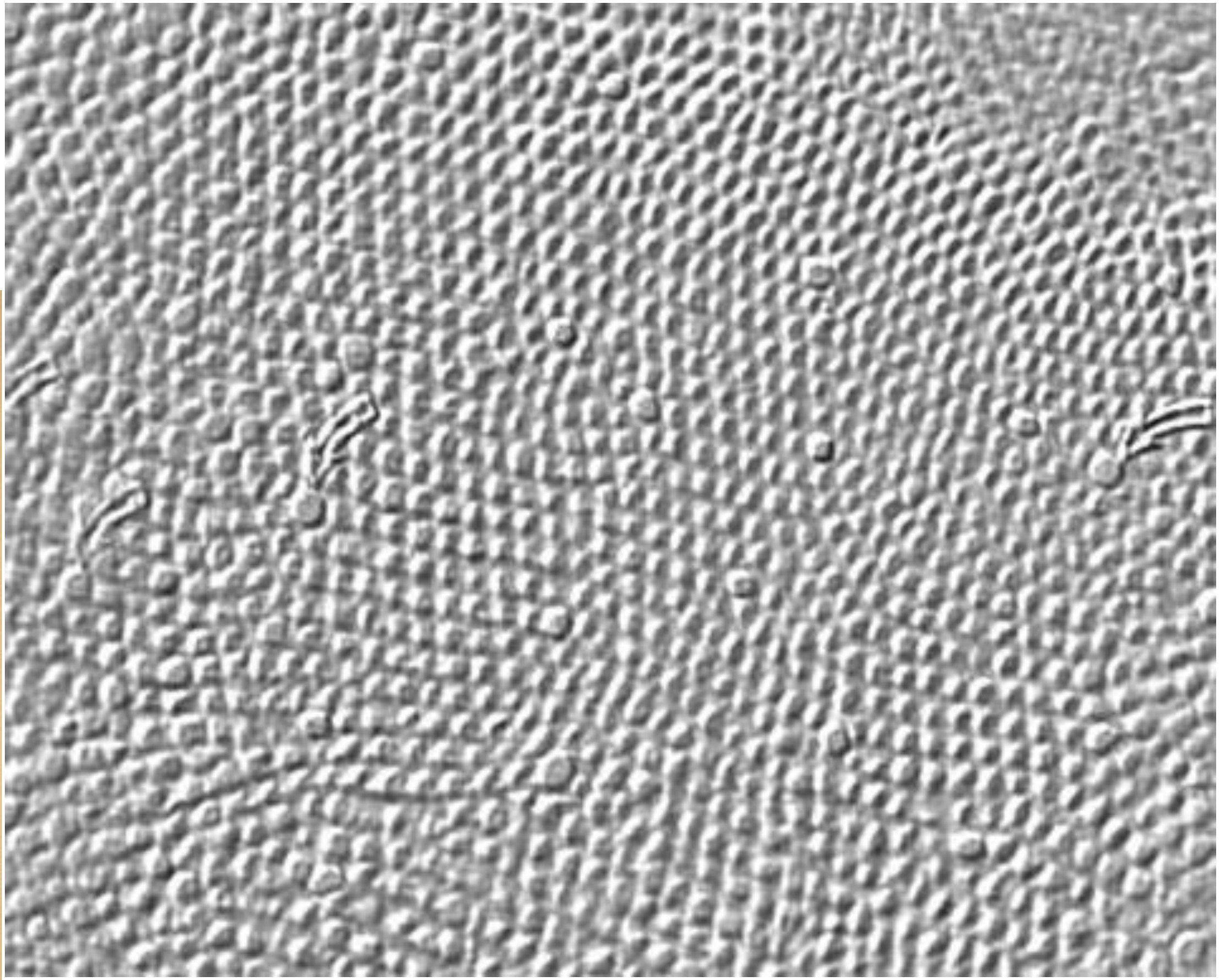
  - Two spatial resolution tasks, illustrating two-dot and vernier acuity tasks —
    - The perception threshold for the two-dot task: about 30'' (seconds of arc).
    - The perception threshold for the vernier task: can be as low as 5''!

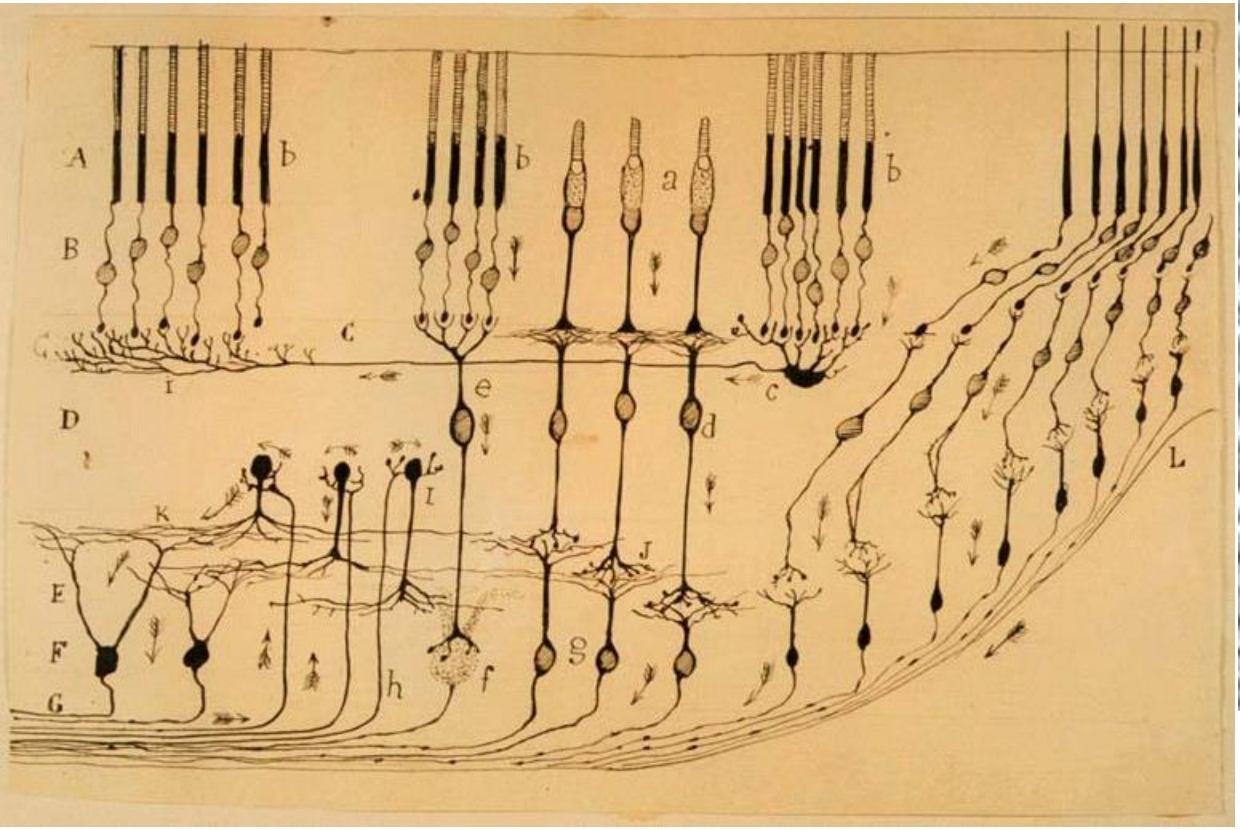


the measurement device

A magnified image of the retinal mosaic

This is the fovea, hence no rods — only cones.



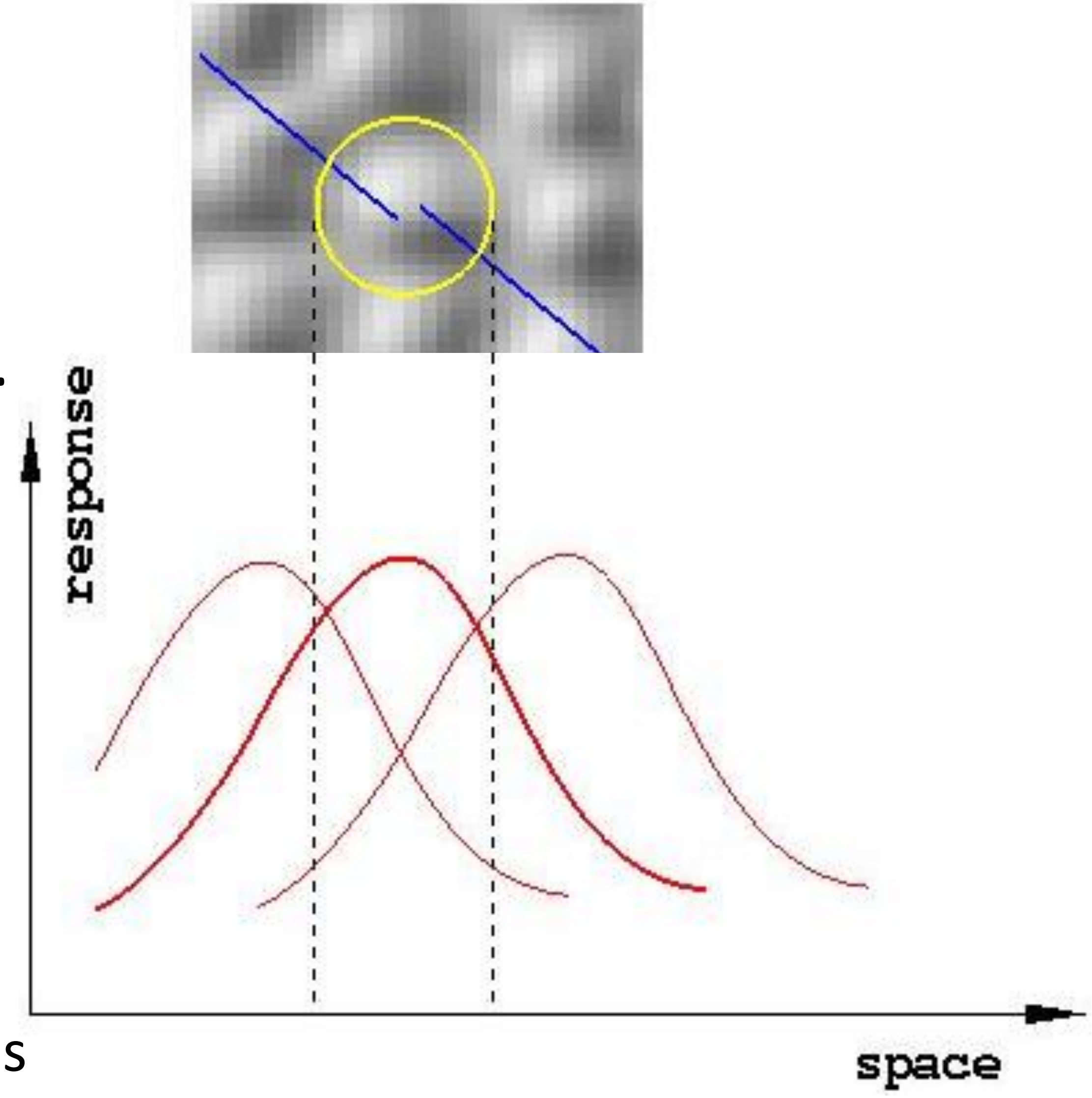


Hyperacuity

Hyperacuity means detecting visual details that are smaller than the size of photoreceptors in the eye.
The displacement is much smaller than a single photoreceptor, yet the brain can still detect it!

Simple receptive fields (RFs) aren’t good at this because they treat nearby points as the same if they fall into the same RF.

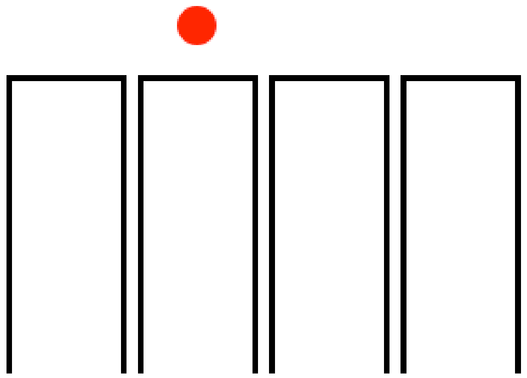
- The smallest discernible vernier, as it projects onto the retinal mosaic —
- Note that the vernier displacement is much smaller than photoreceptor size.
- This is an example of hyperacuity-level performance.
- This receptive field (RF) coding isn't very good
  - This measurement device is too **insensitive**: two close-by dots will likely fall under the same RF and their locations will be perceived as the same.



A cross-section of the receptive fields of three adjacent receptors.

High-Resolution Coding?

- "high-resolution" coding isn't very good either
  - This measurement device is also suboptimal: dot locations get "digitized", but some information is still lost.



What if we just make really small receptive fields to improve resolution?
Problem:

If each RF detects a very tiny area, some details might be lost (like “pixelation” in low-resolution images).

The system may “digitize” locations but miss subtle differences.

overlapping coding is better

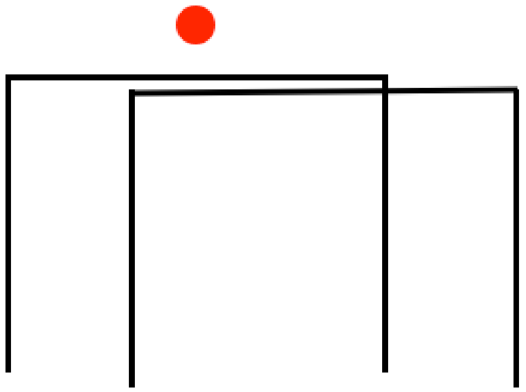
- overlapping "tabletop" coding is better, why?

•Instead of having small, separate RFs, we overlap them

Why is this better?

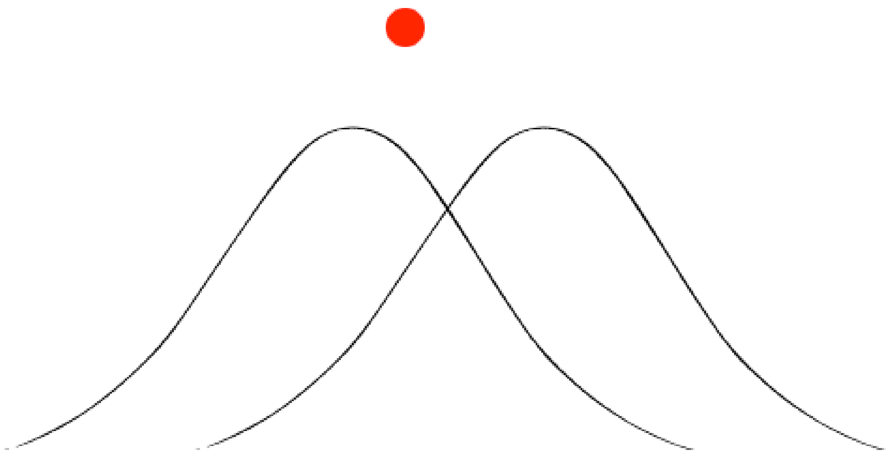
Overlapping RFs capture more fine details by combining information from multiple regions.

Key Idea: Overlapping receptive fields improve precision by making sure small changes don’t go unnoticed.

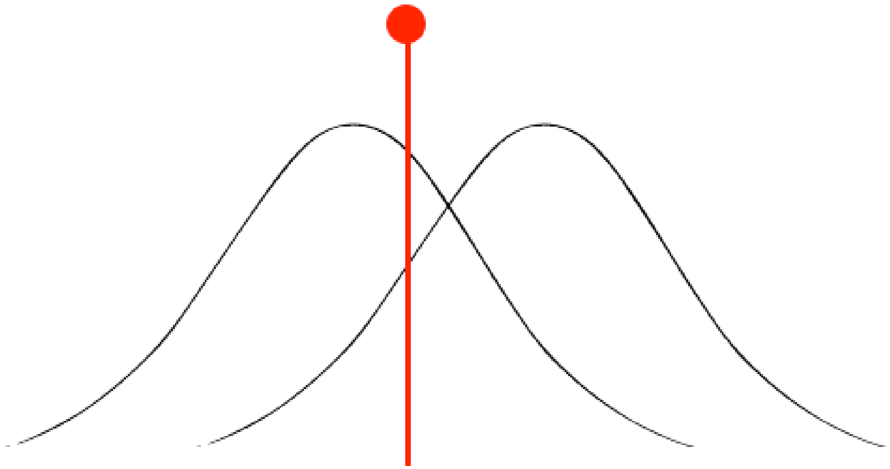


overlapping graded coding is better

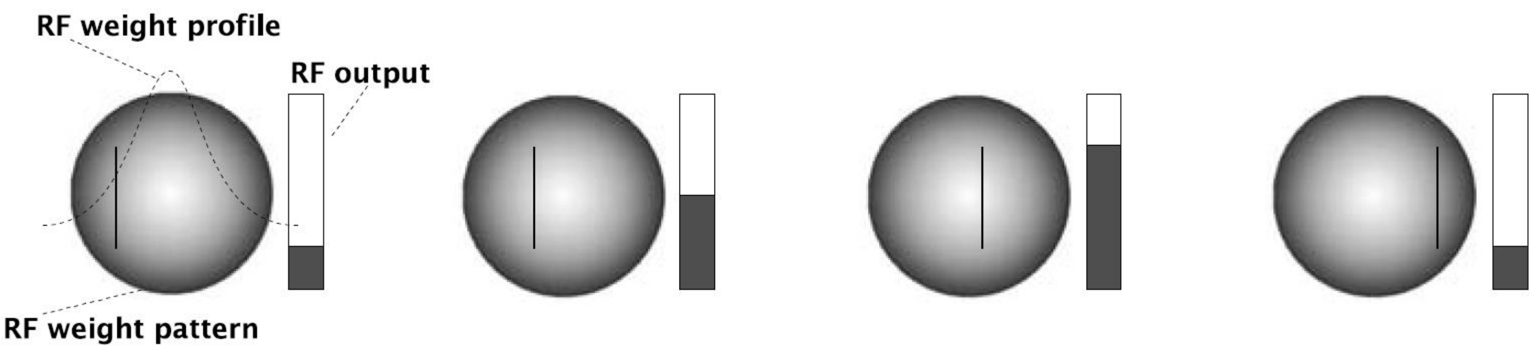
- overlapping, graded RF coding is even better, why?



- Even small lateral displacements of the dot will not go unnoticed:
  - they get transduced into measurable changes in the outputs of the RFs.



Computational Basis of Hyperacuity



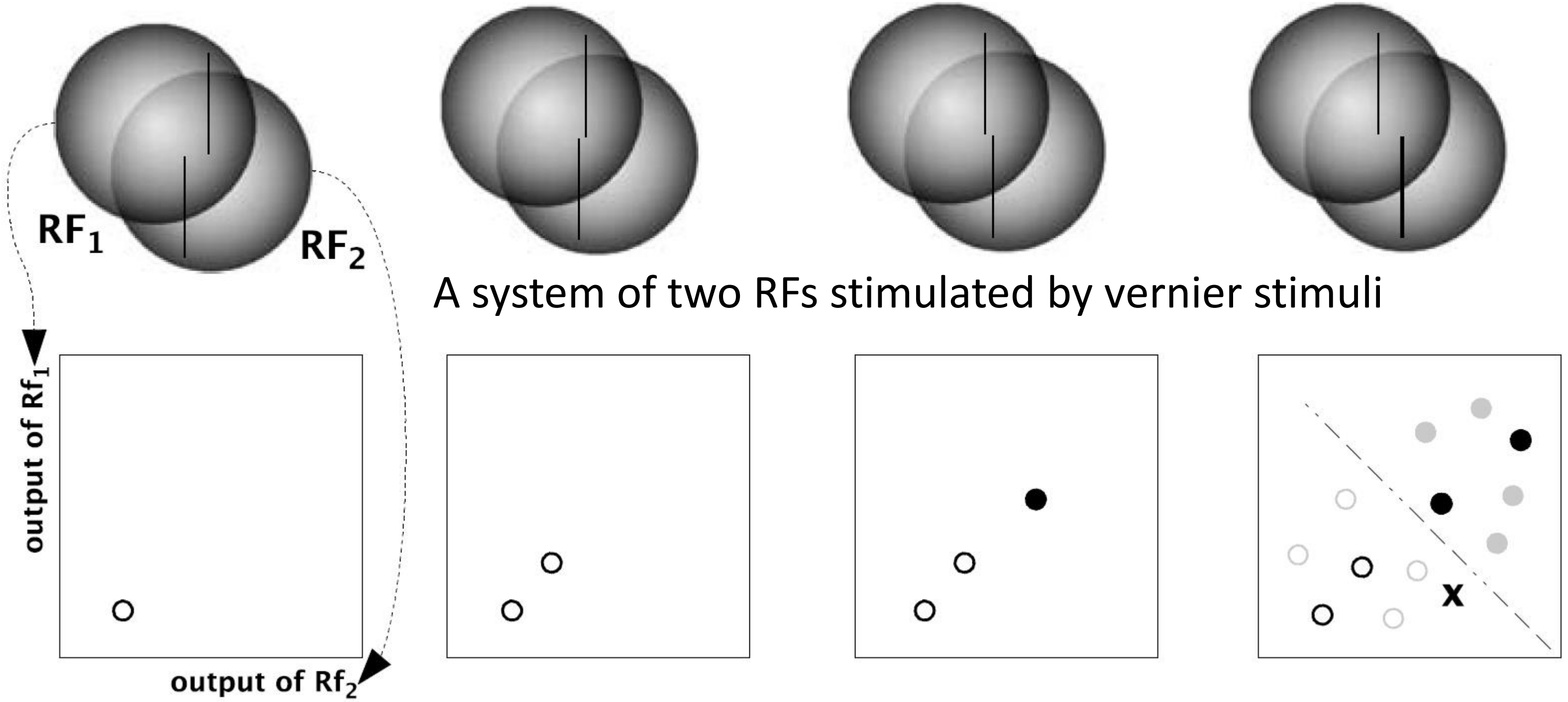
Responses of a single receptive field (RF) stimulated with a vertical line

Slide 5: Computational Basis of Hyperacuity

Shows how multiple receptive fields work together to detect very fine details.

•Even though each RF on its own is not precise, their combined response creates a more accurate perception.

Key Idea: Hyperacuity is possible because multiple RFs contribute to fine visual details.



A system of two RFs stimulated by vernier stimuli

the activation states evoked by these stimuli. Open and filled circles represent sign flipping

Poggio et al. (1992)

Broad, overlapping, graded receptive fields effective

Summary:

hyperacuity-level performance is possible because

- the RFs are graded, and
- the RFs are broad and overlapping in space.
- Perceptual learning (implicit) and specific to the task

Instead of just increasing resolution, the brain combines overlapping signals to enhance precision.

Three-line bisection task