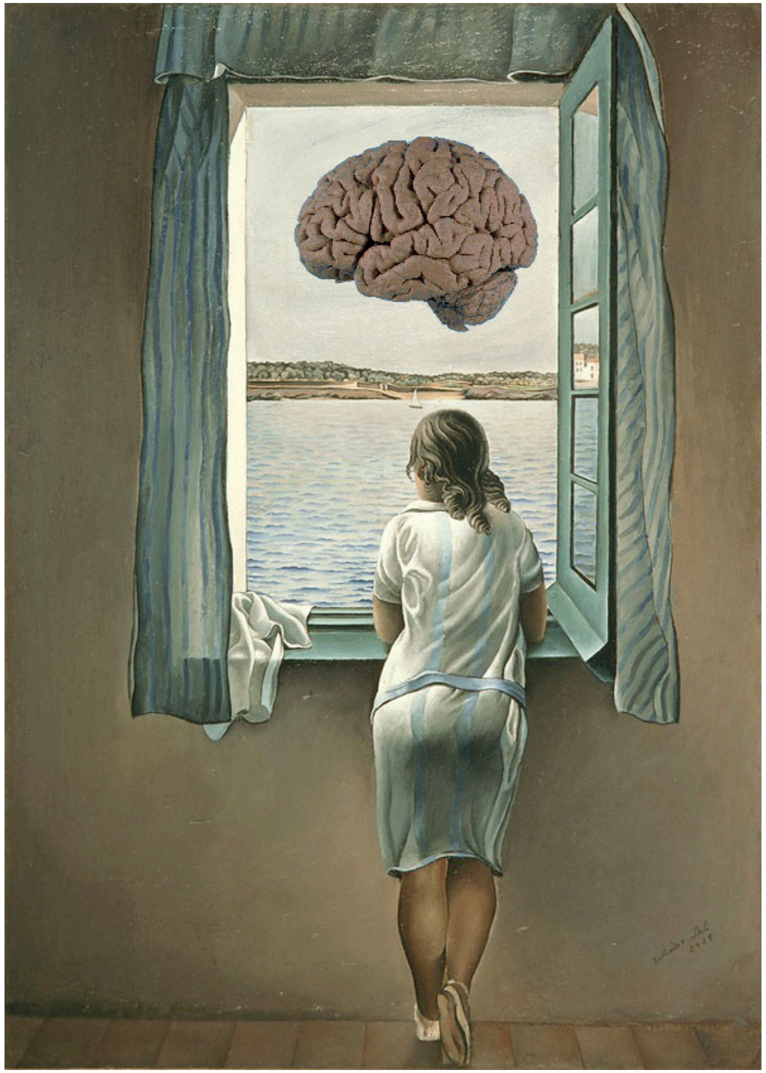
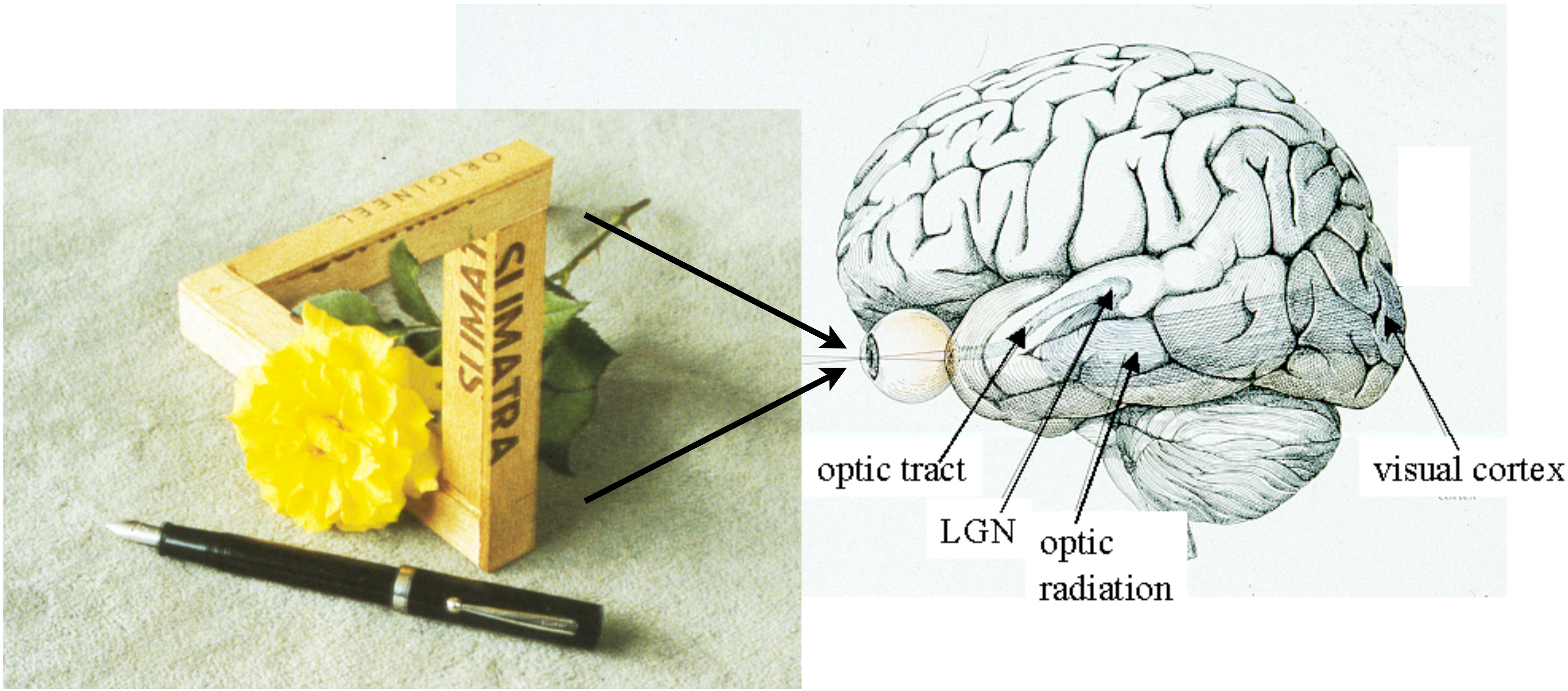
Visual Processing



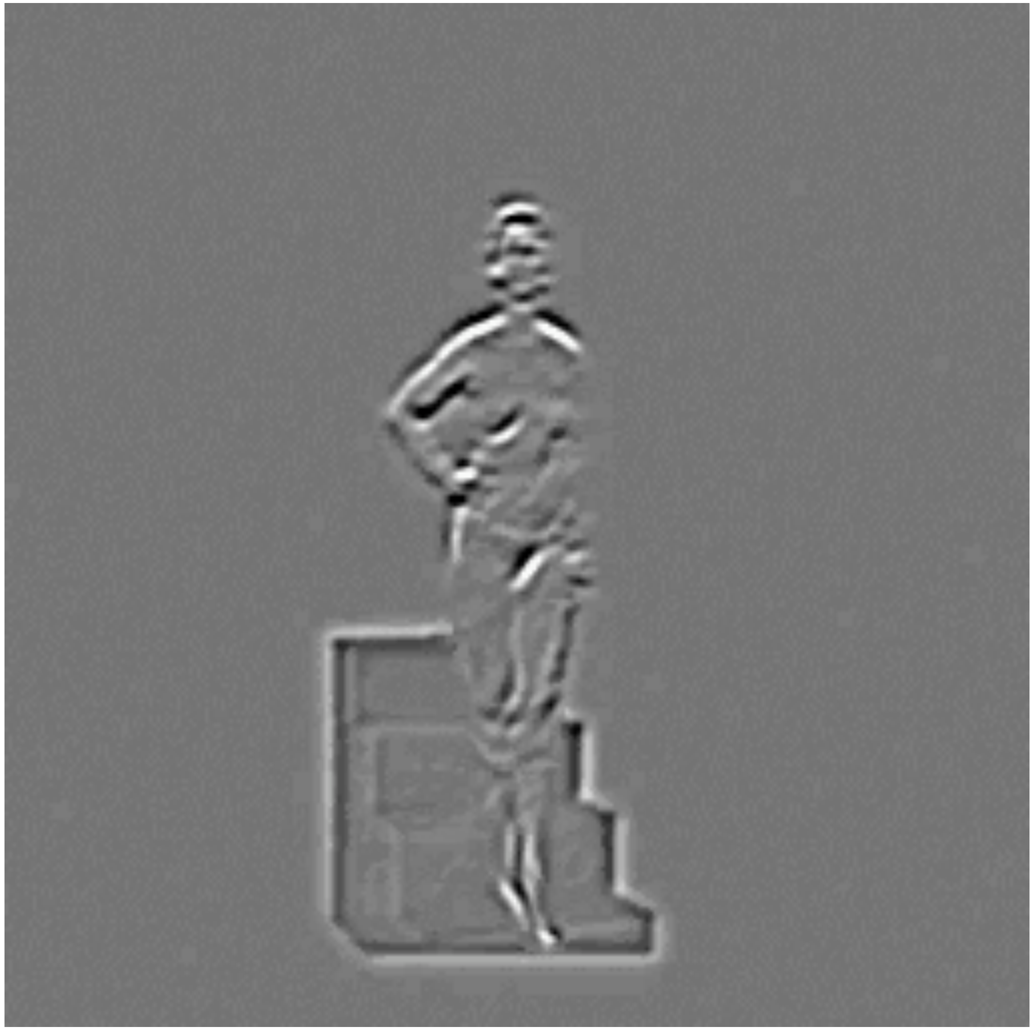
David J. Heeger
New York University

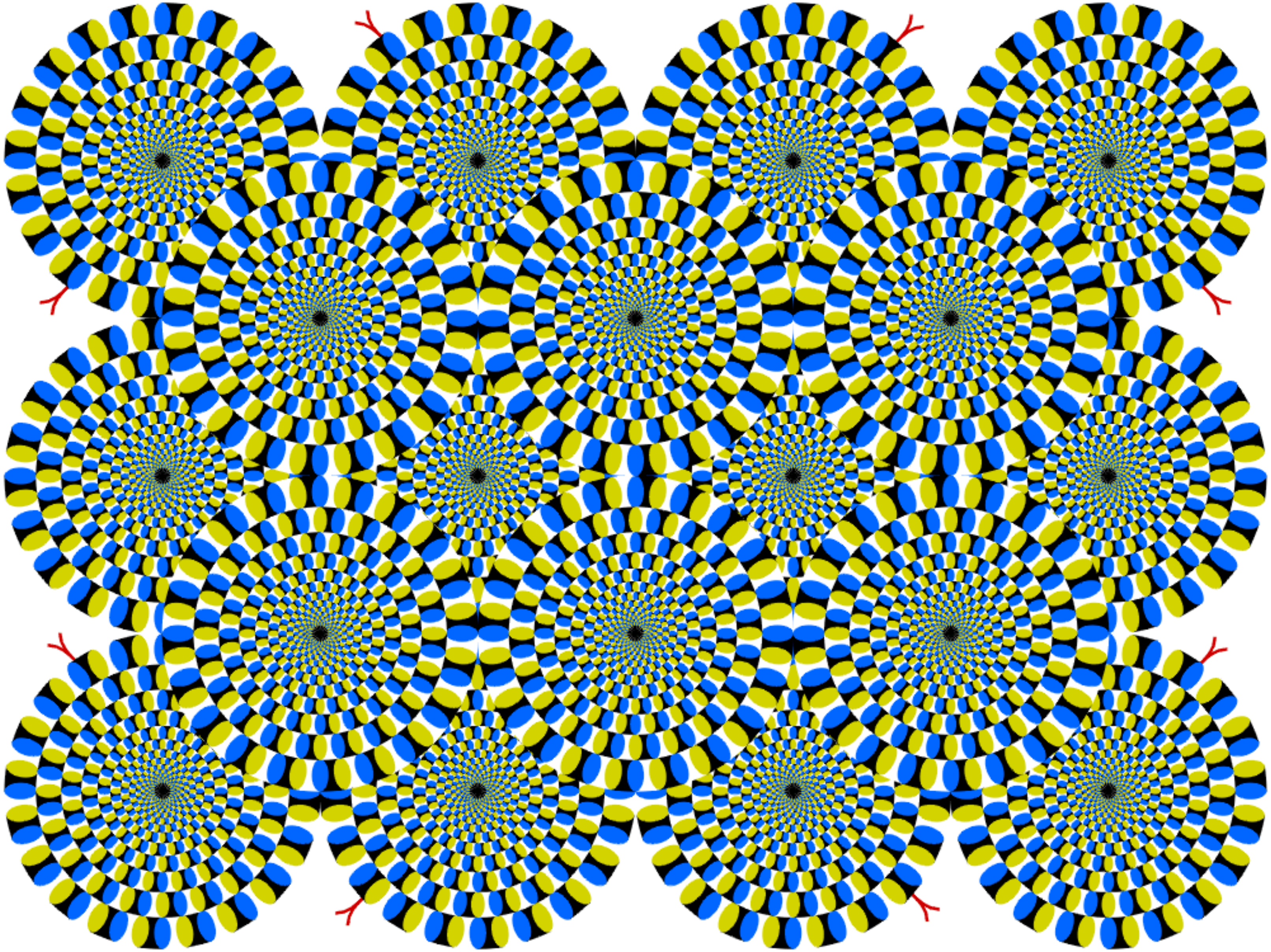
david.heeger@nyu.edu

Vision is an unconscious inference



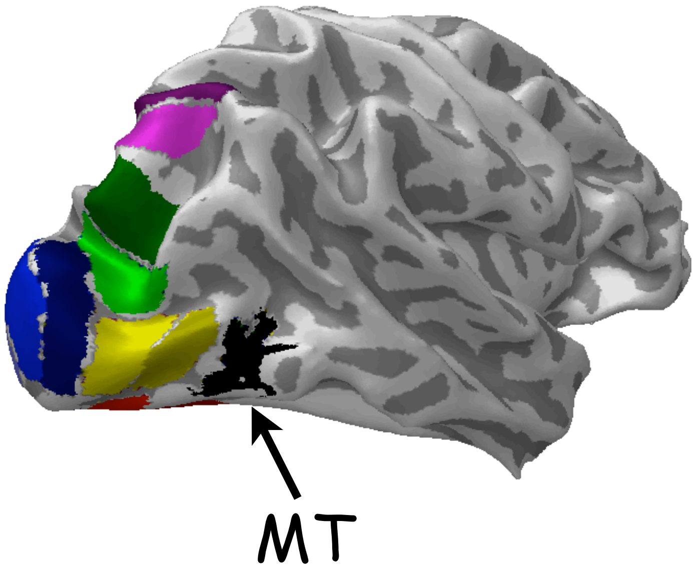
Visual inference: motion perception



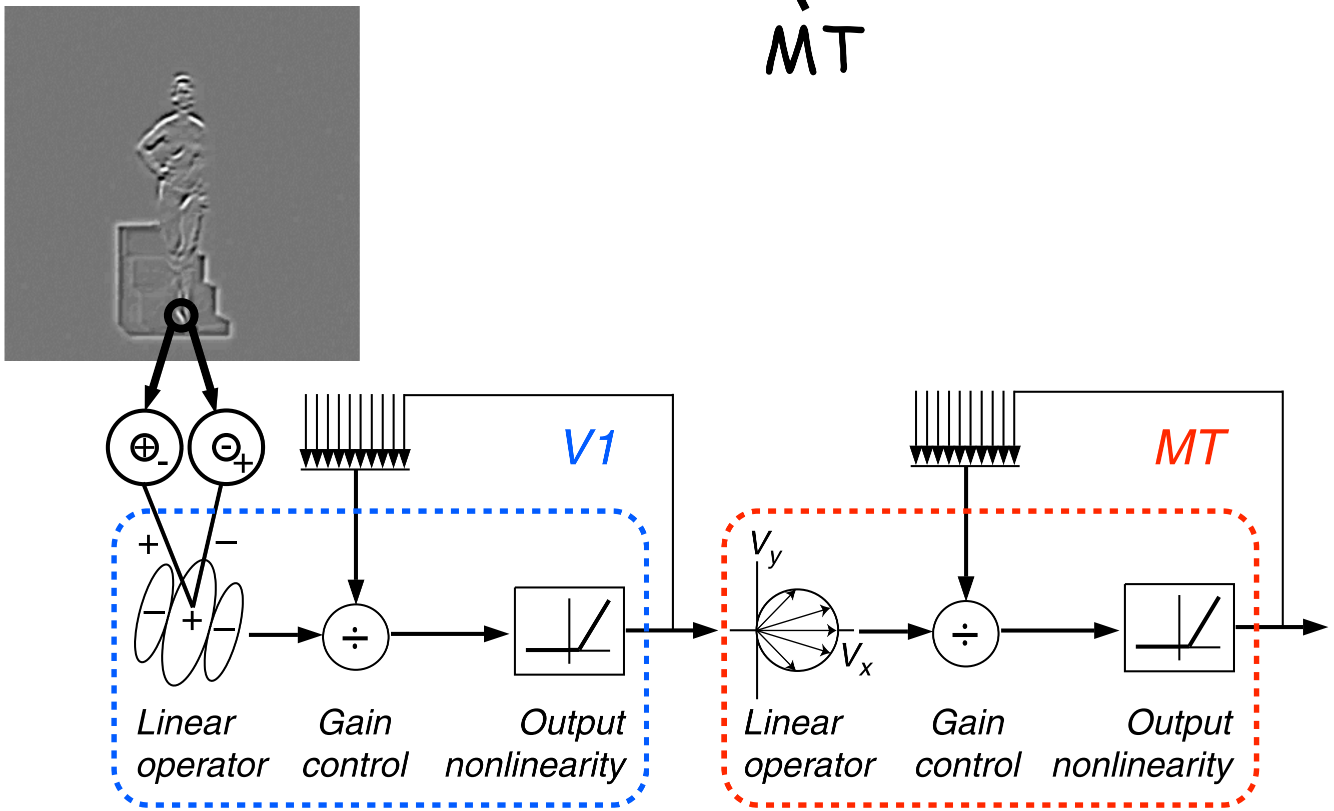


Two guiding principles

Functional specialization



Computational theory

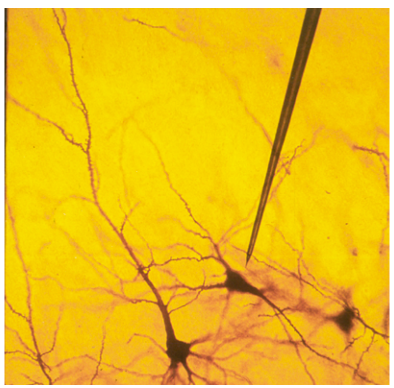


Visual pathways and functional specialization

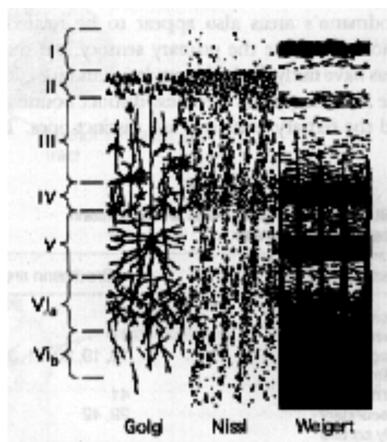
Defining visual cortical areas

PhACT

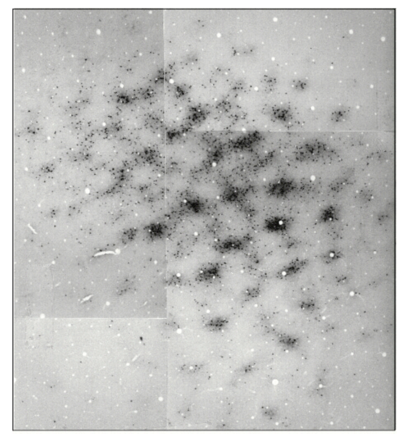
Physiology



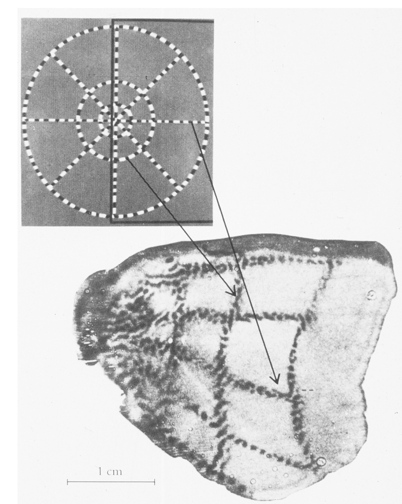
Architecture



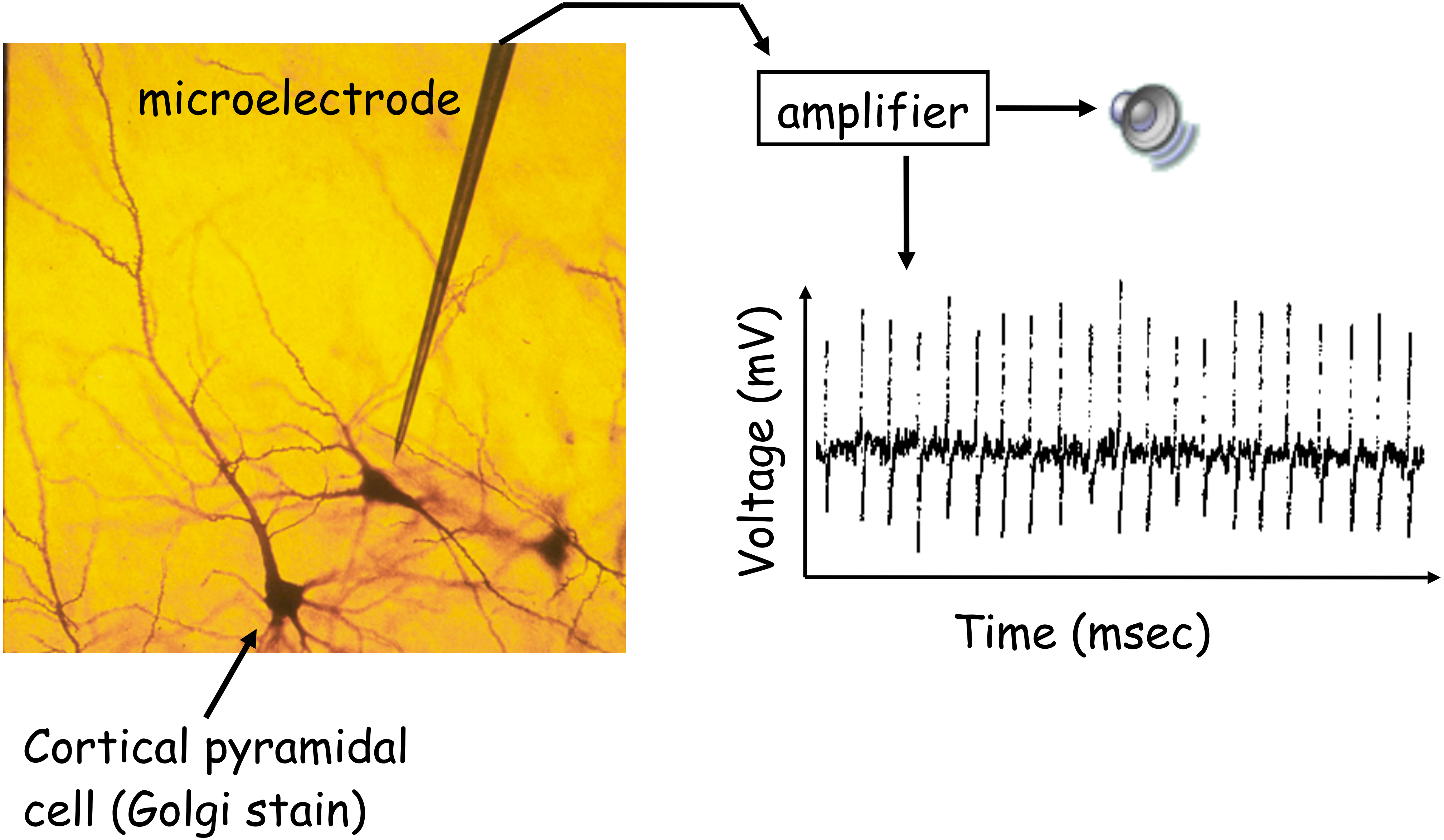
Connections



Topography



Electrophysiology (action potential)



Cytoarchitecture: Brodmann's areas

Korbinian Brodmann
(1868-1918)

~50 cytoarchitectural areas defined by cell size, cell density, number of layers, density of myelinated axons.

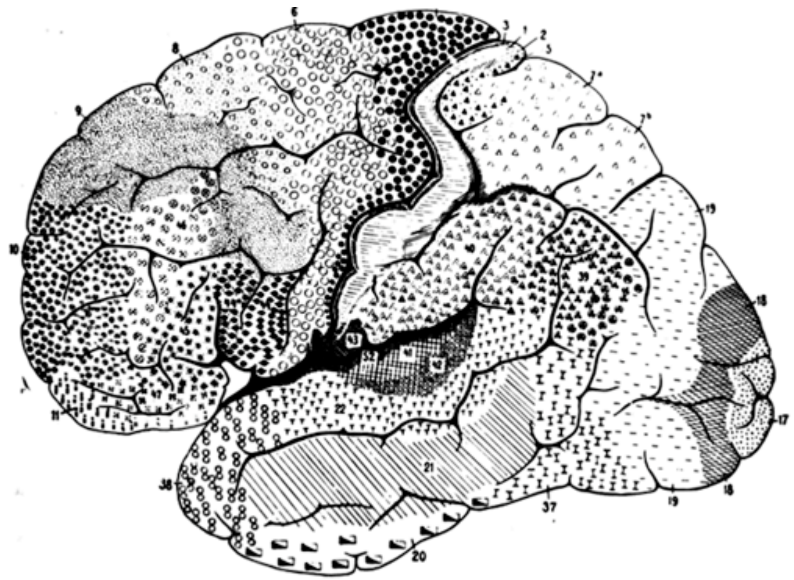


Abb. 6.

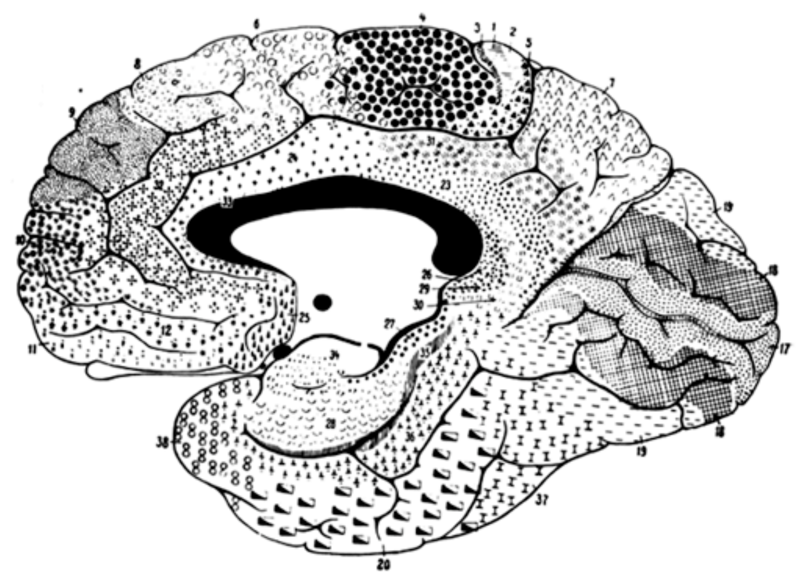
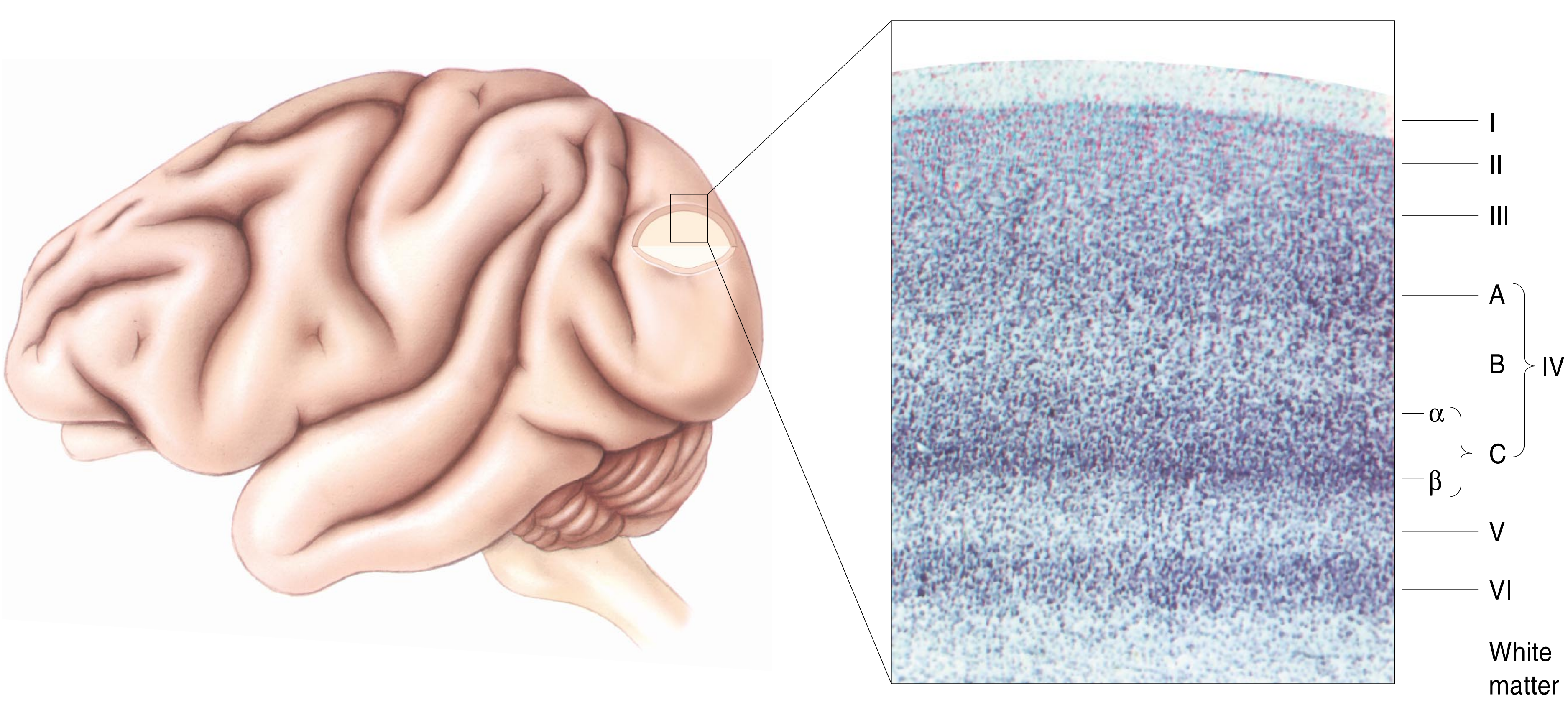


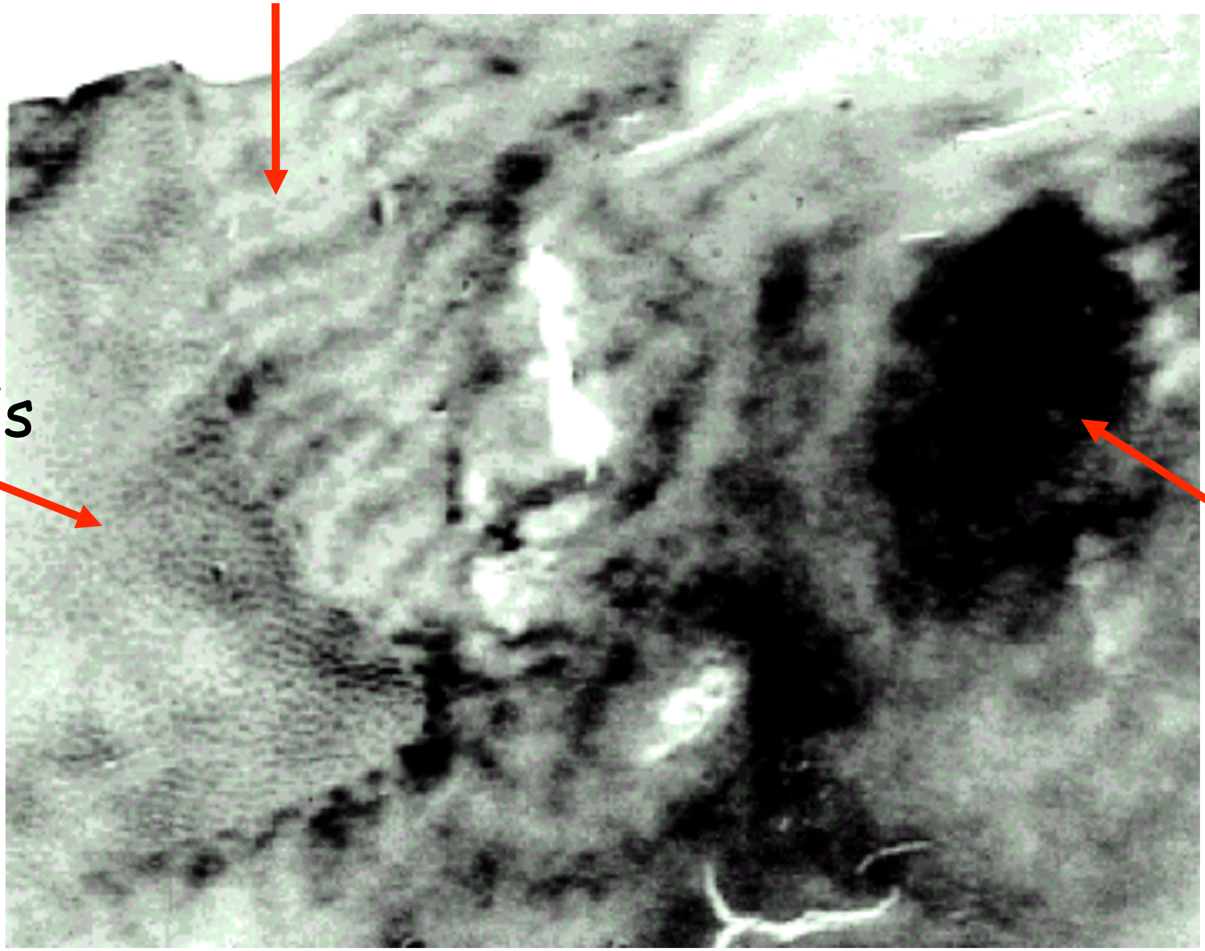
Abb. 7.

Architecture: cortical layers



Primary visual cortex slice (Nissl stain)

Architecture: cytochrome oxidase

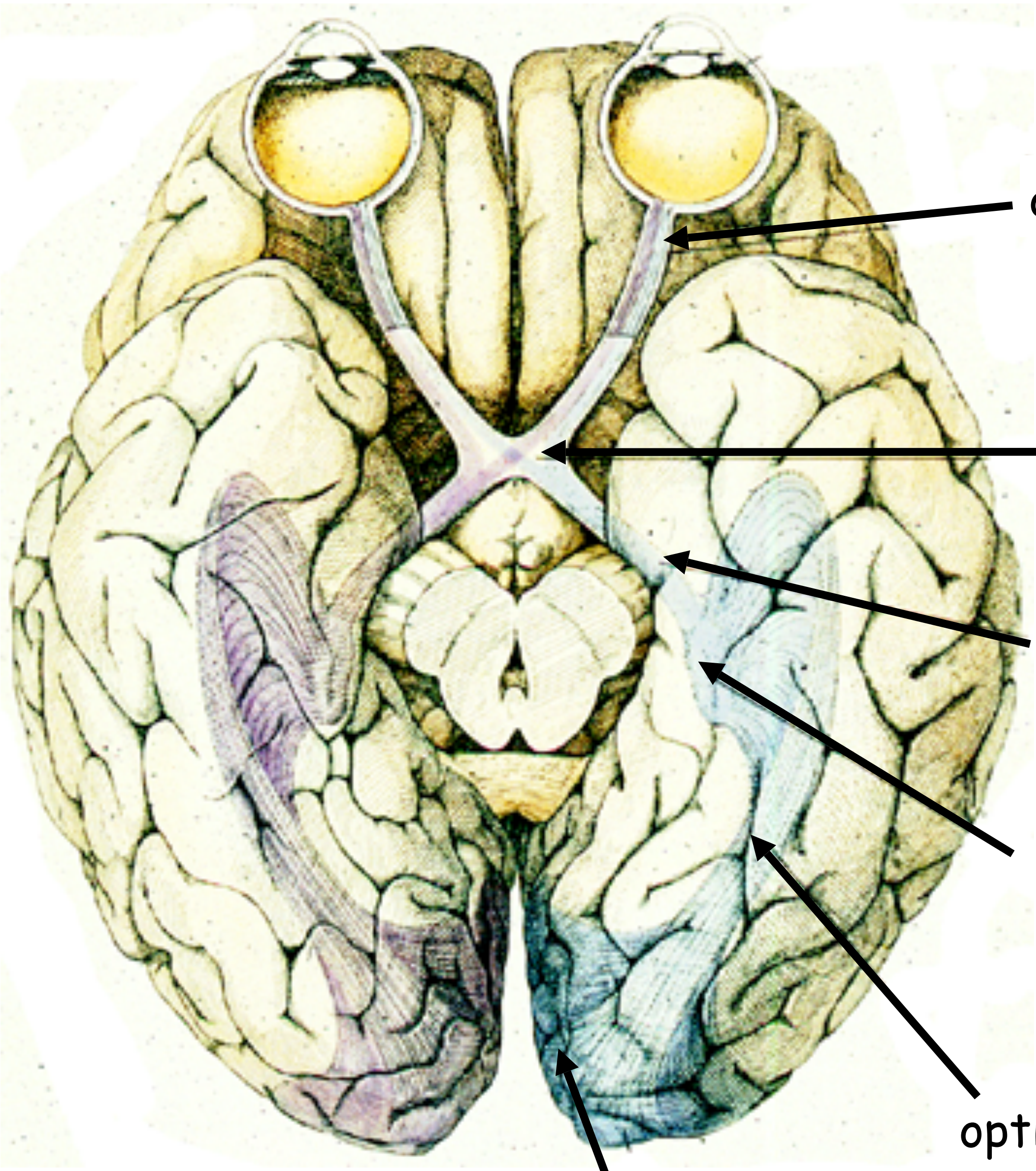


Cytochrome oxidase staining in human visual cortex

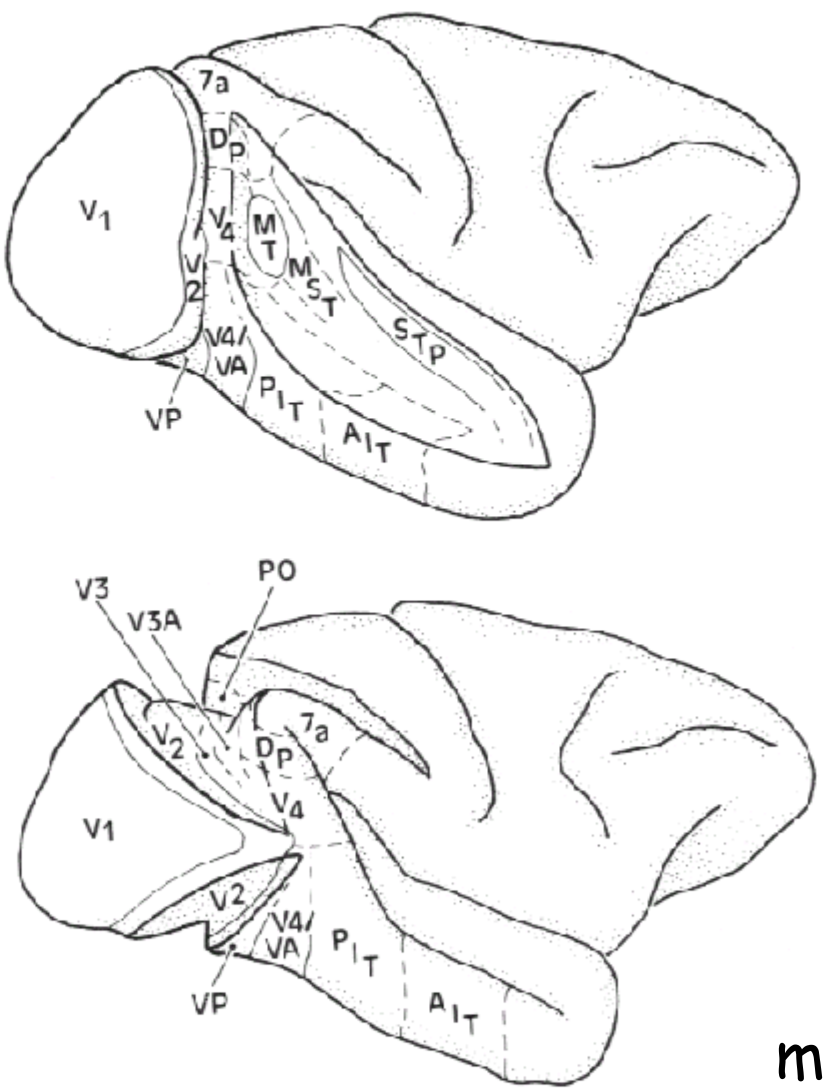
Tootell et al (1995)

Connections: retinogeniculate visual pathway

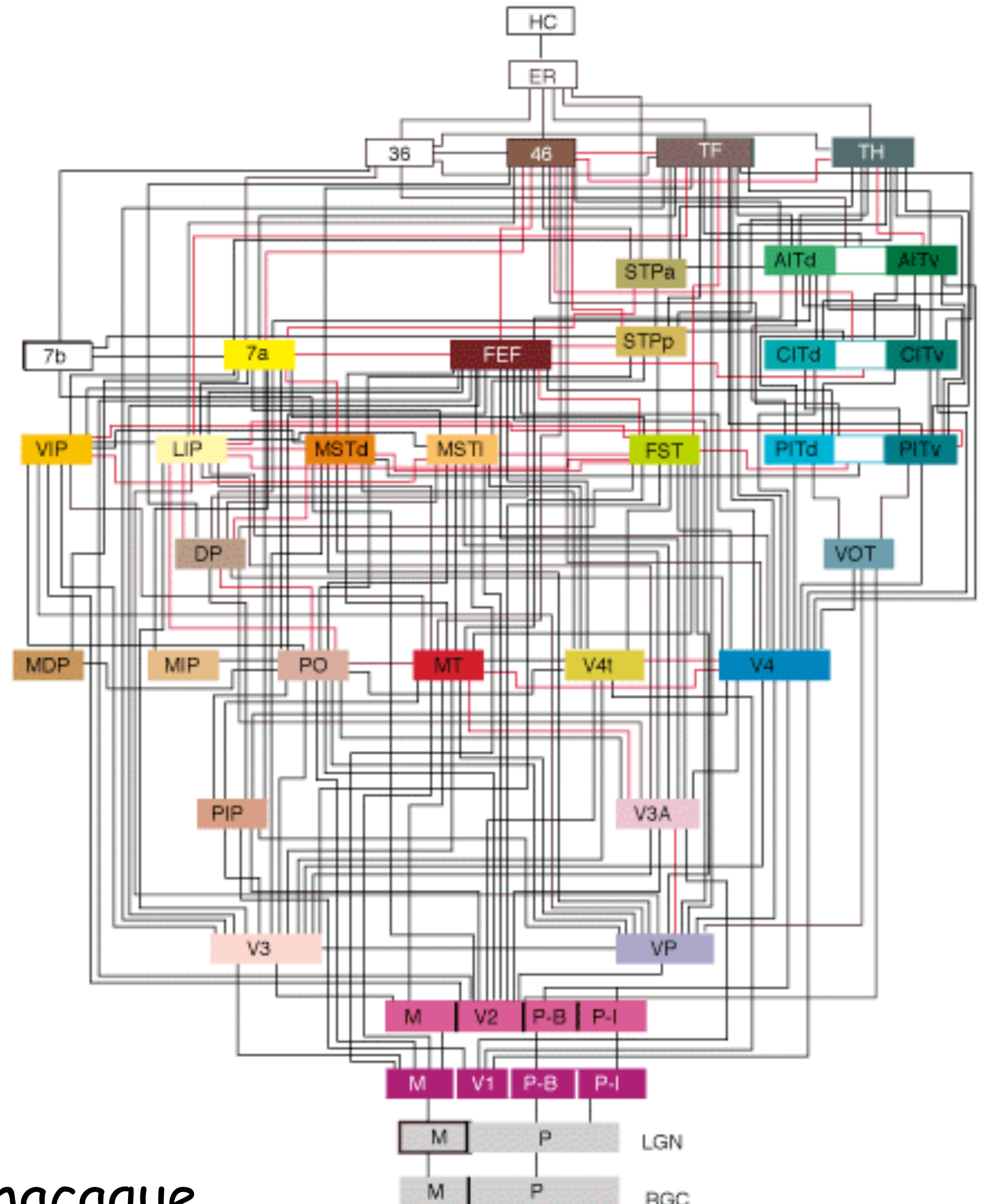
Ventral view



Network of visual cortical areas



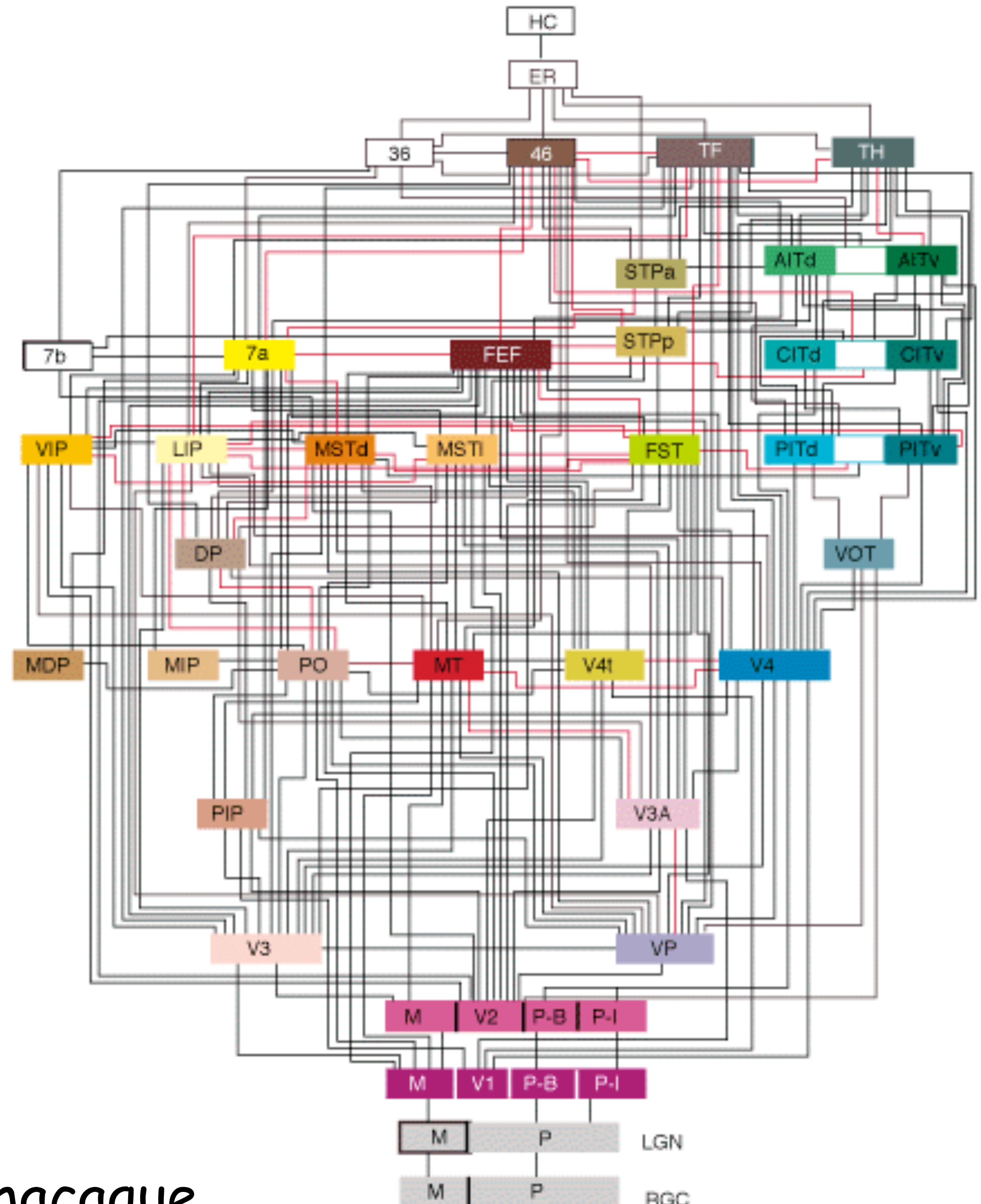
macaque
monkey brain

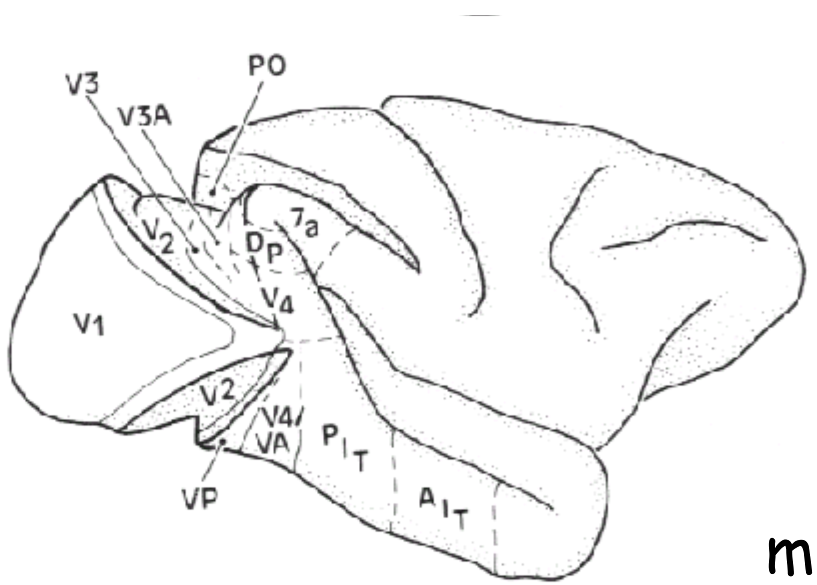


Felleman & Van Essen (1991)

Network of visual cortical areas

Each "feedforward" connection has a corresponding "feedback" connection.





macaque monkey brain

Felleman & Van Essen (1991)

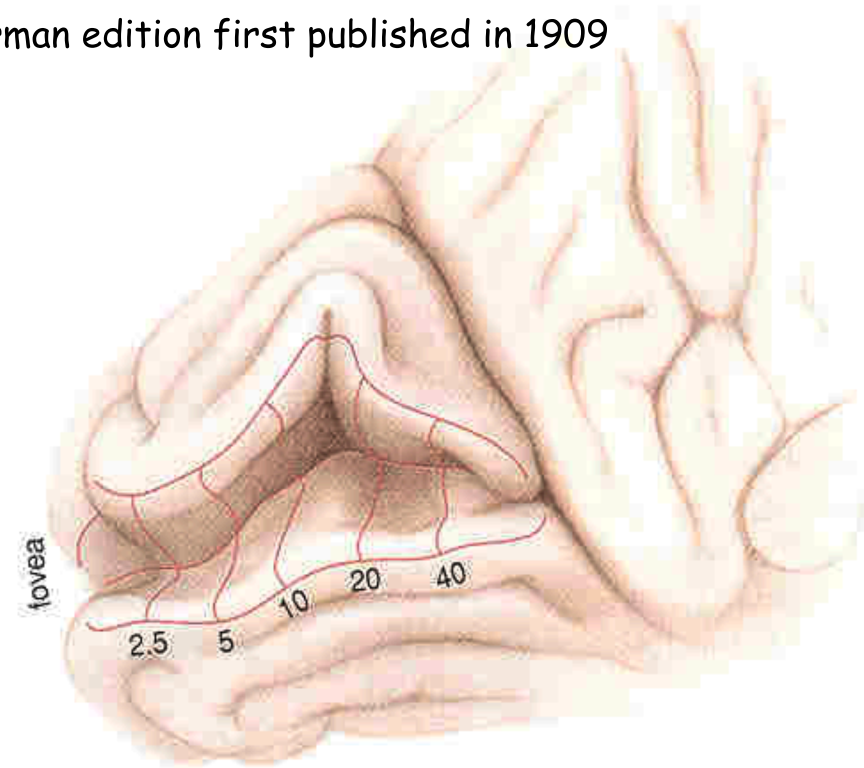
Topography: retinotopy in human V1

Tatsuji Inouye
(1880-1976)

Visual Disturbances Following Gunshot Wounds of the Cortical Visual Area

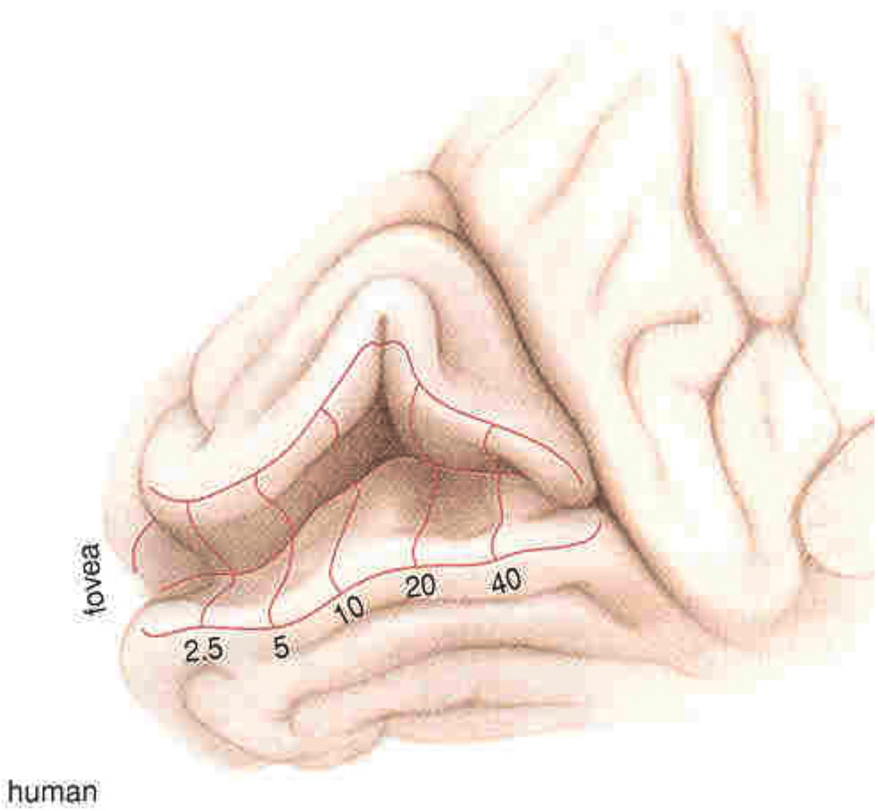
Based on observations of the wounded in the recent Japanese wars

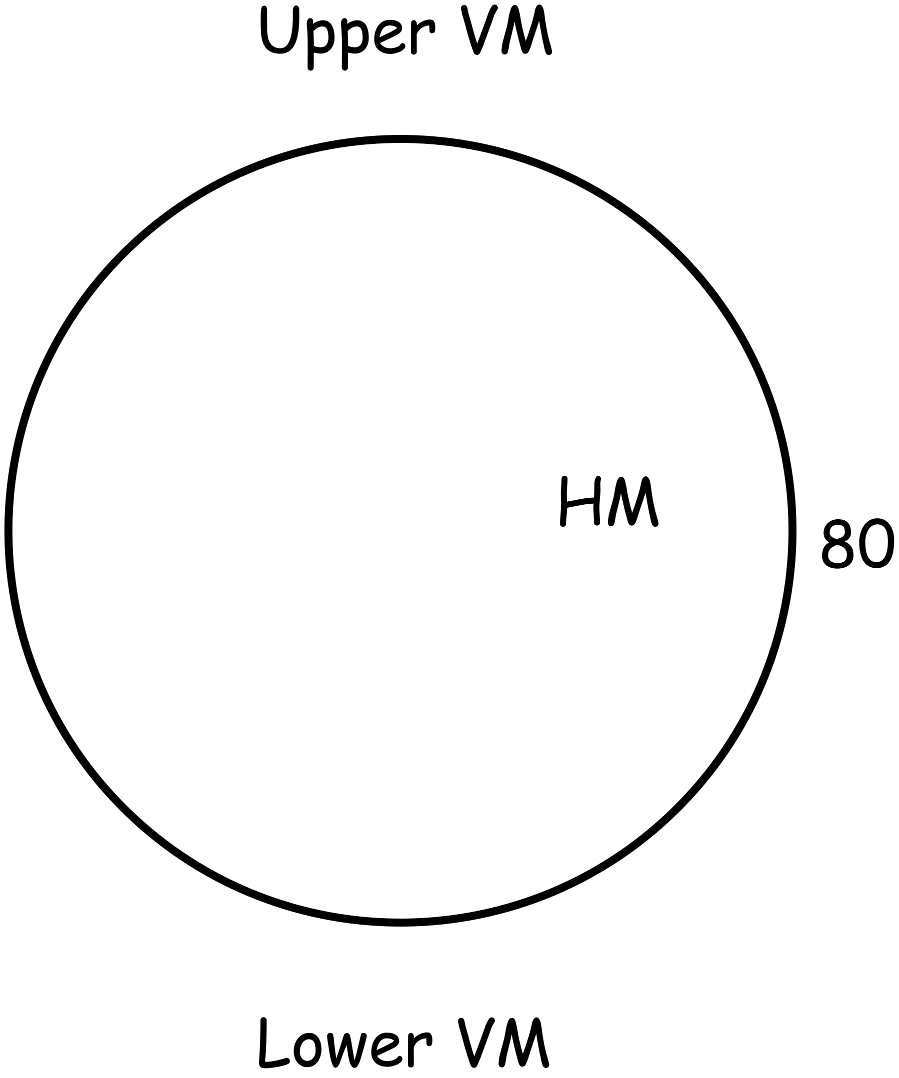
German edition first published in 1909



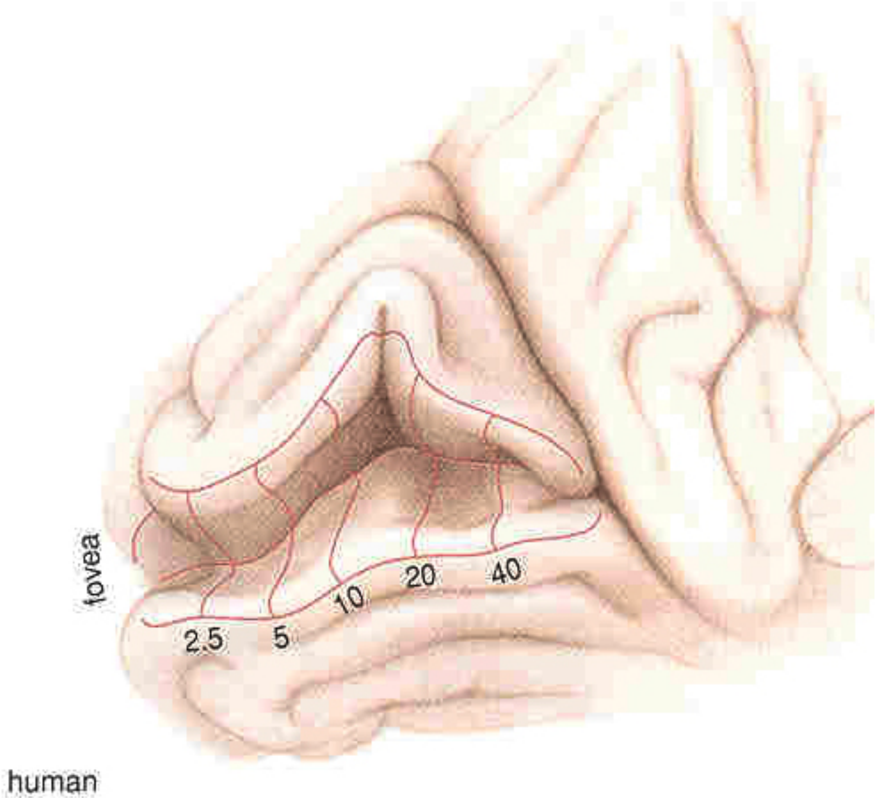
Horton & Hoyt, 1991

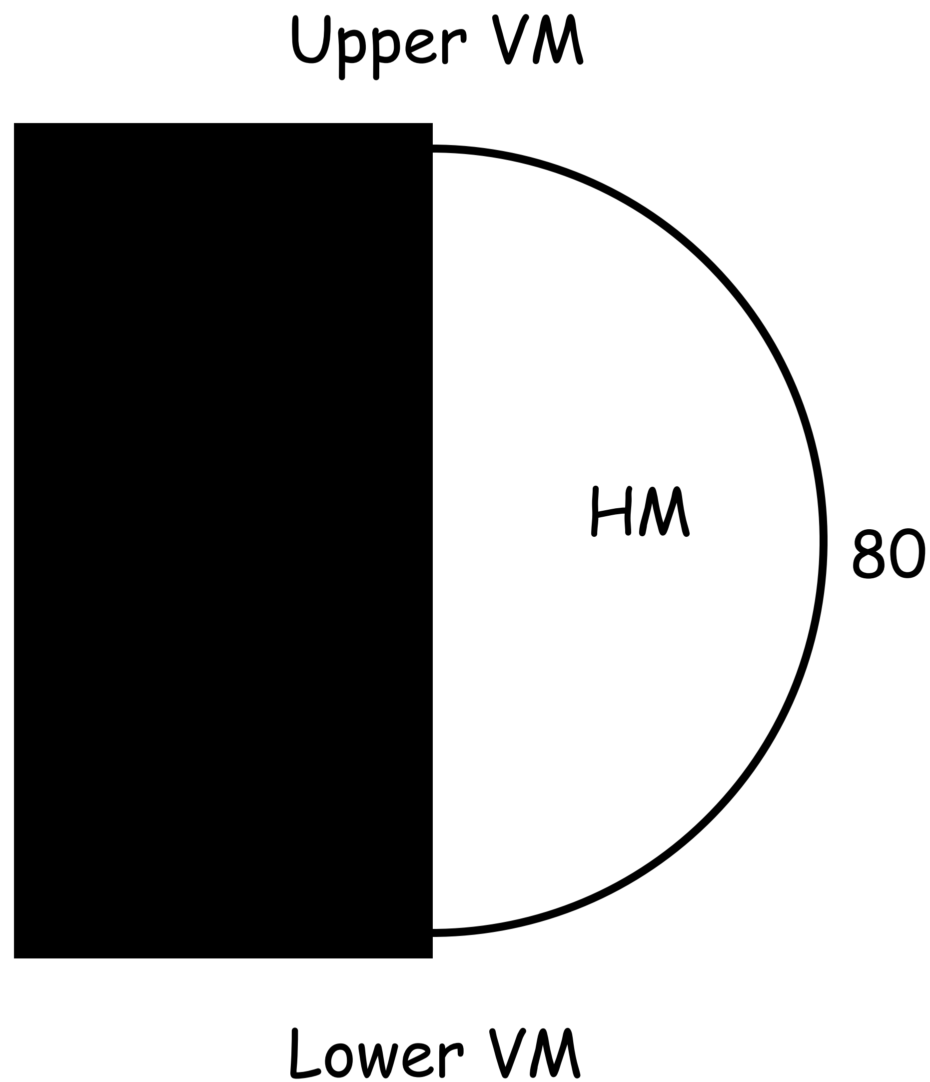
V1 retinotopy



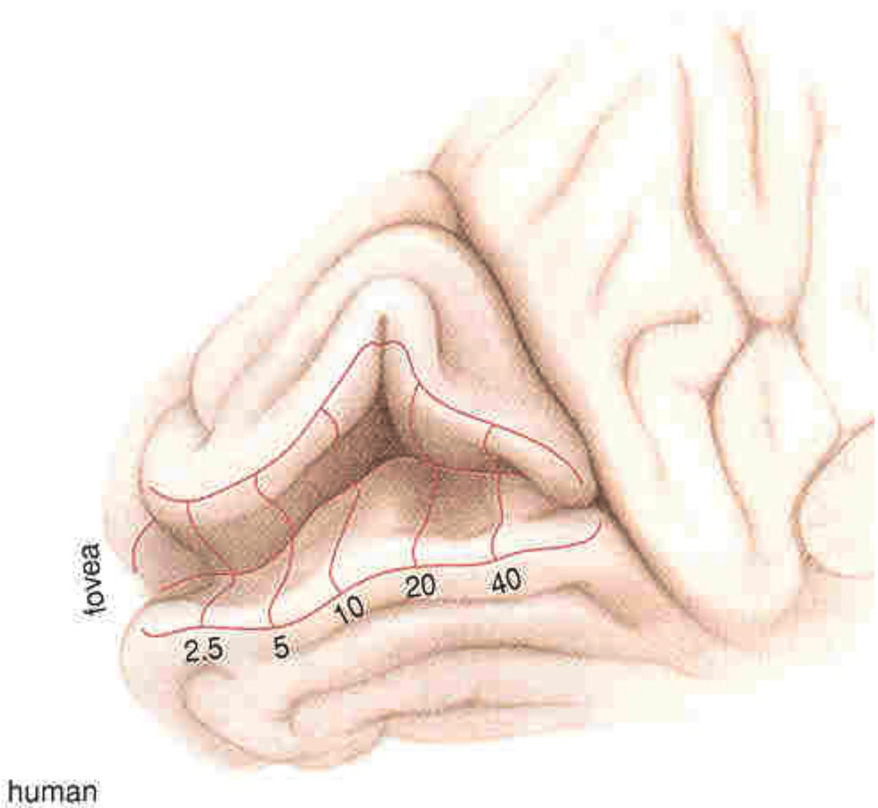


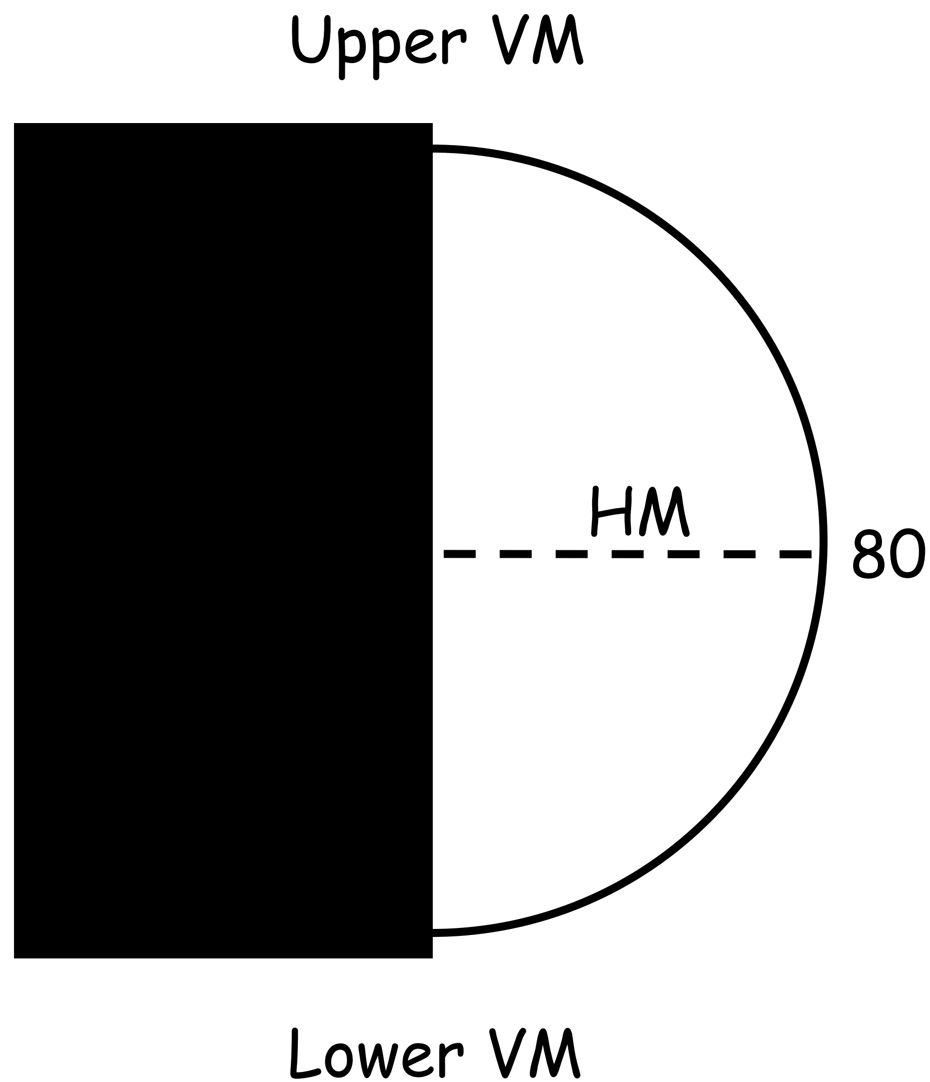
V1 retinotopy



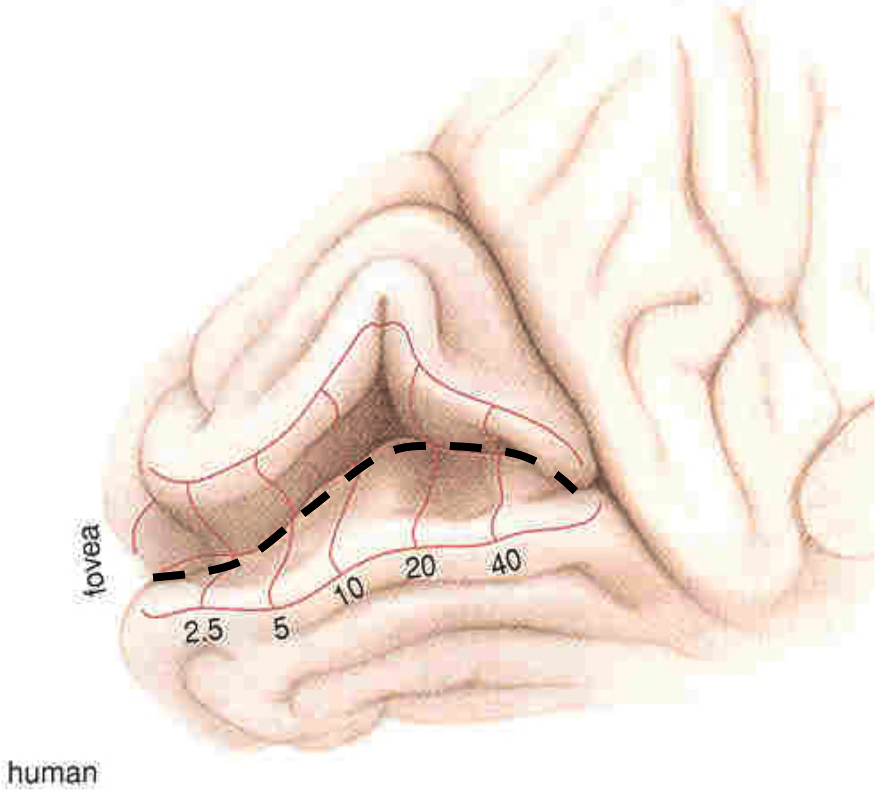


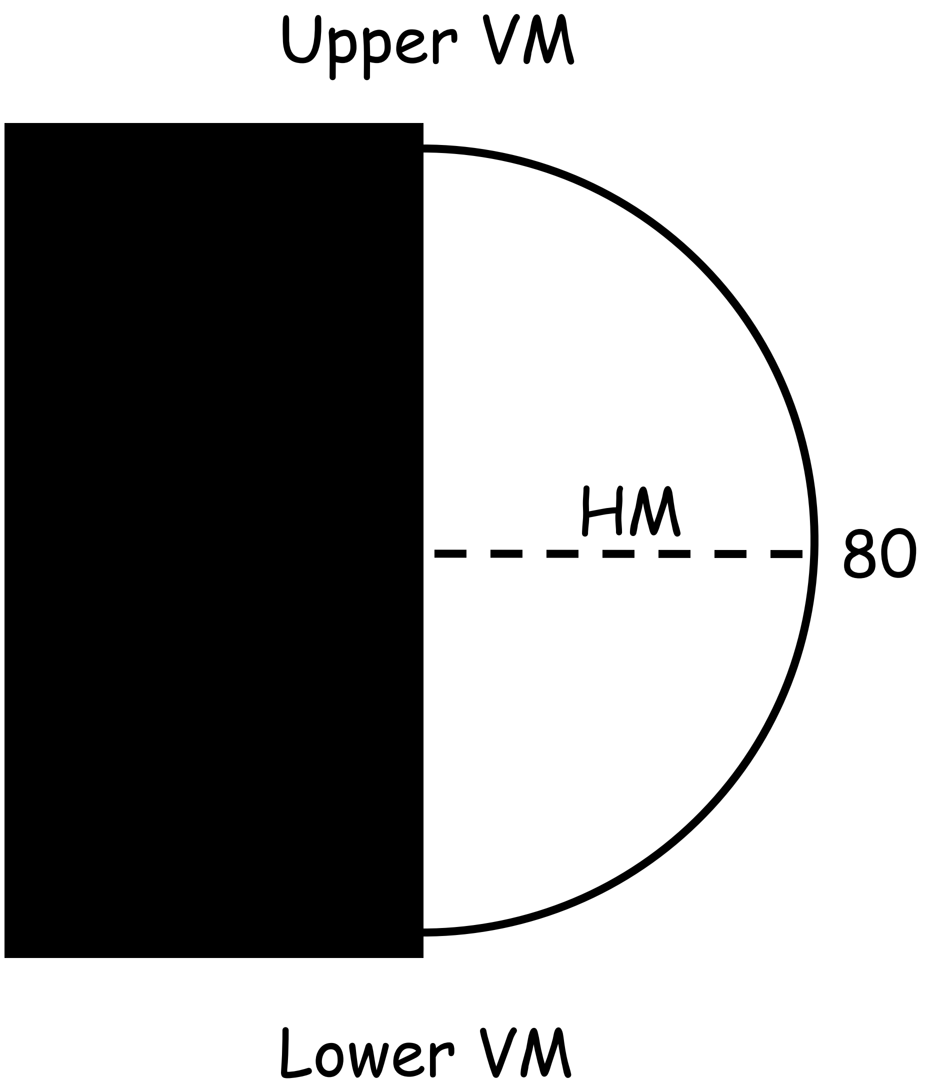
V1 retinotopy



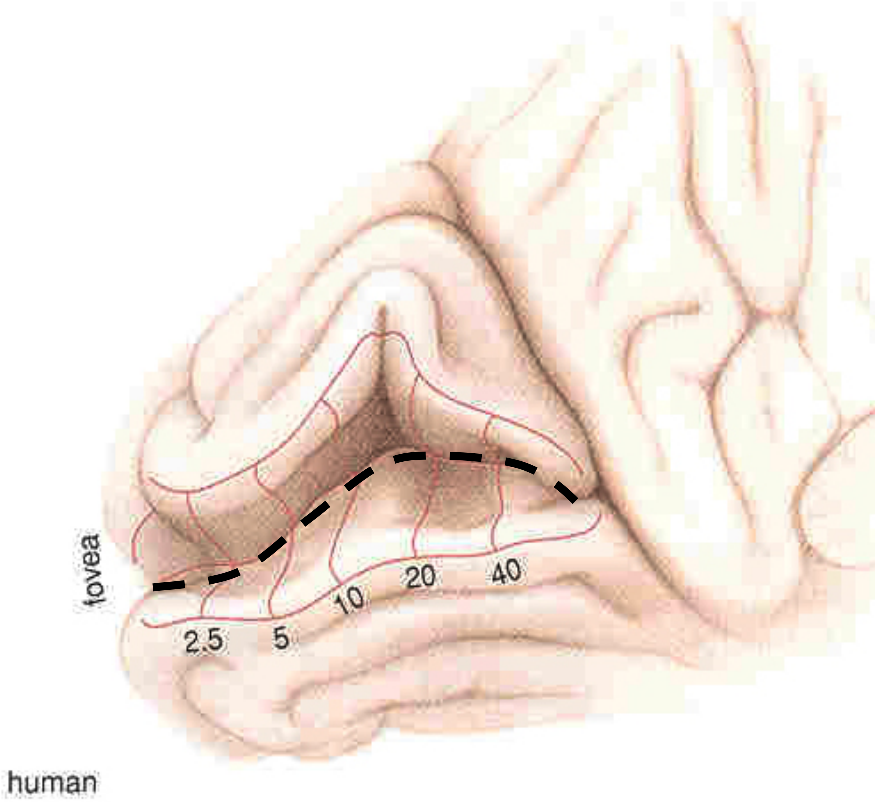


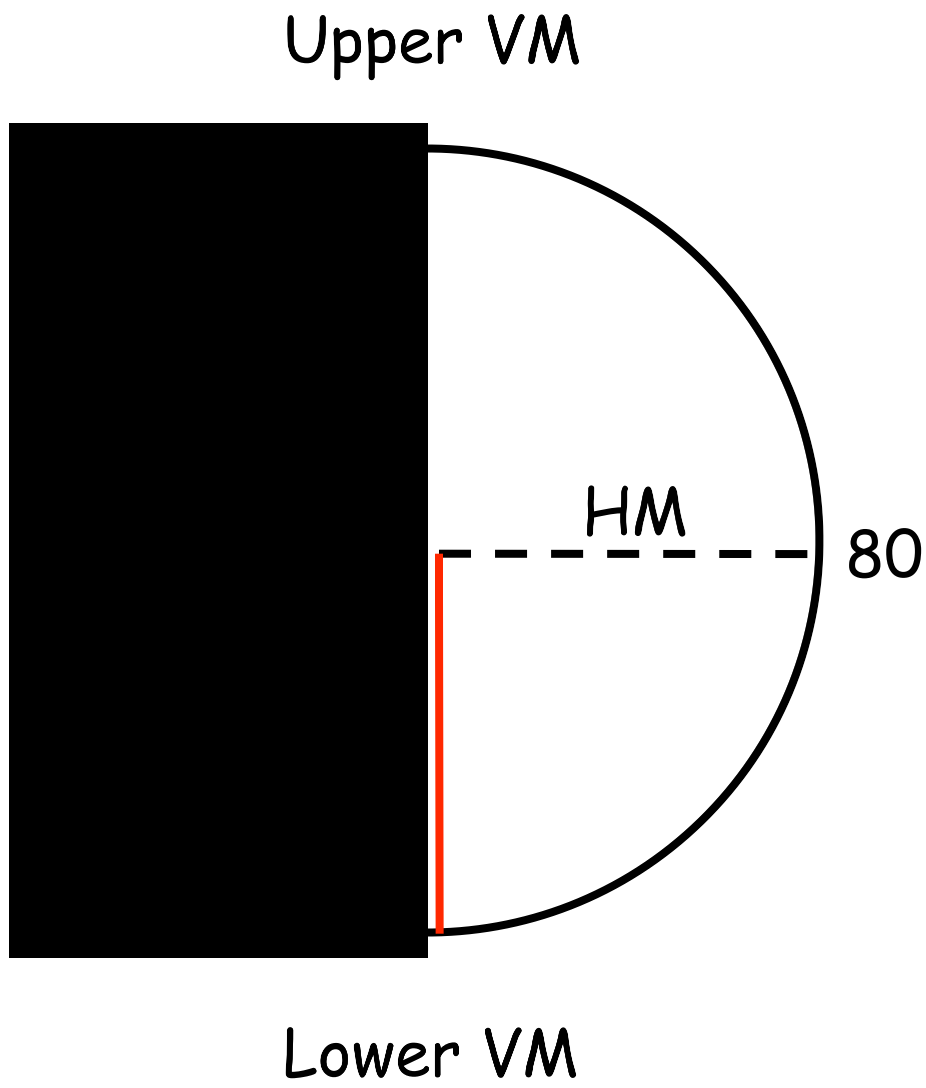
V1 retinotopy



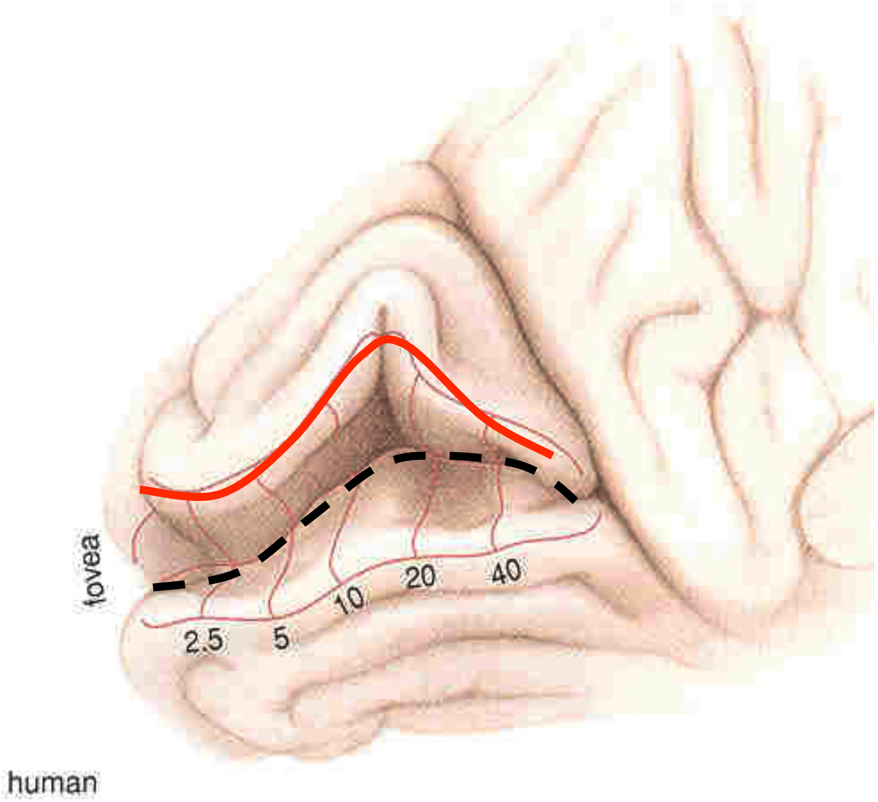


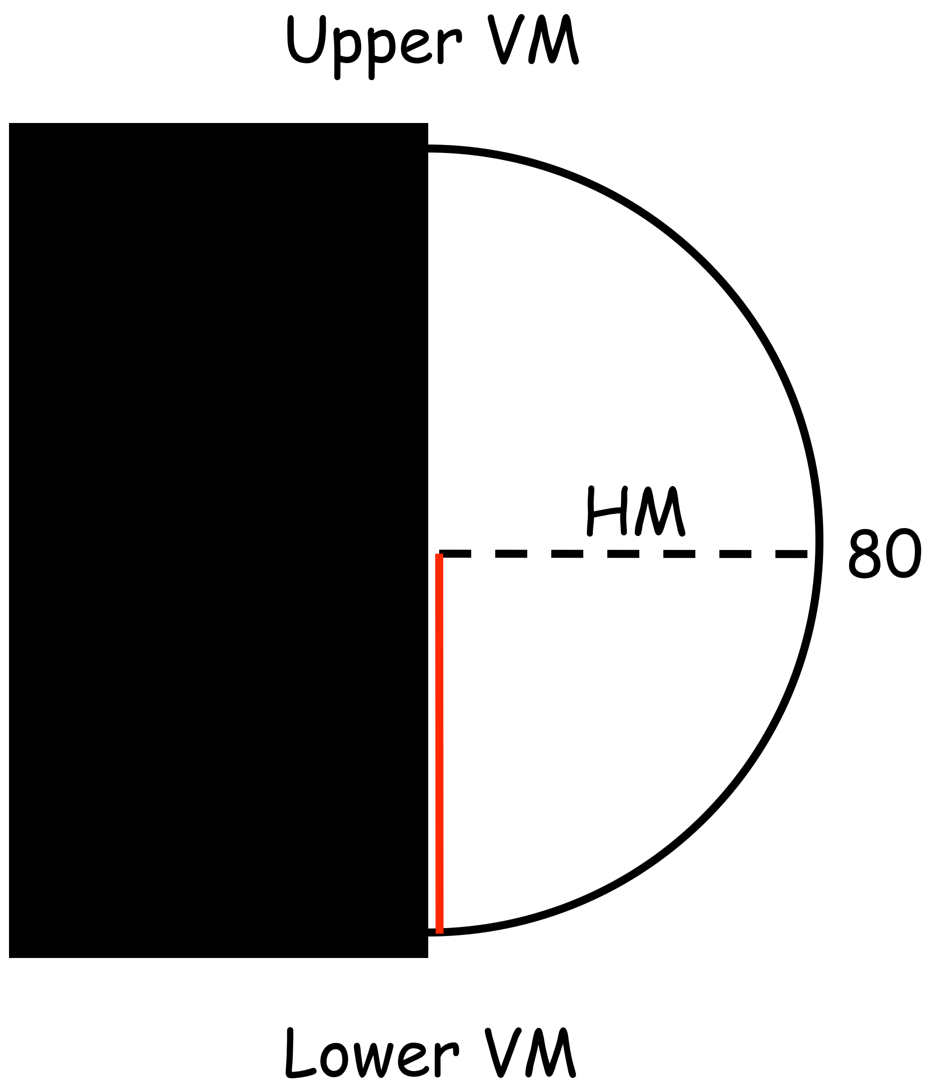
V1 retinotopy



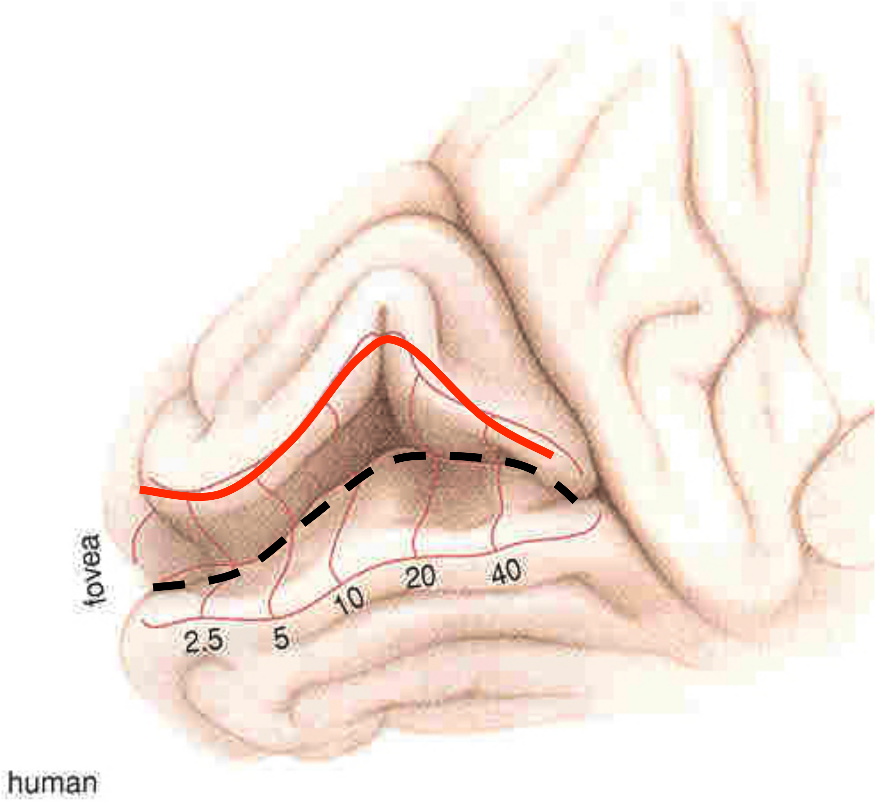


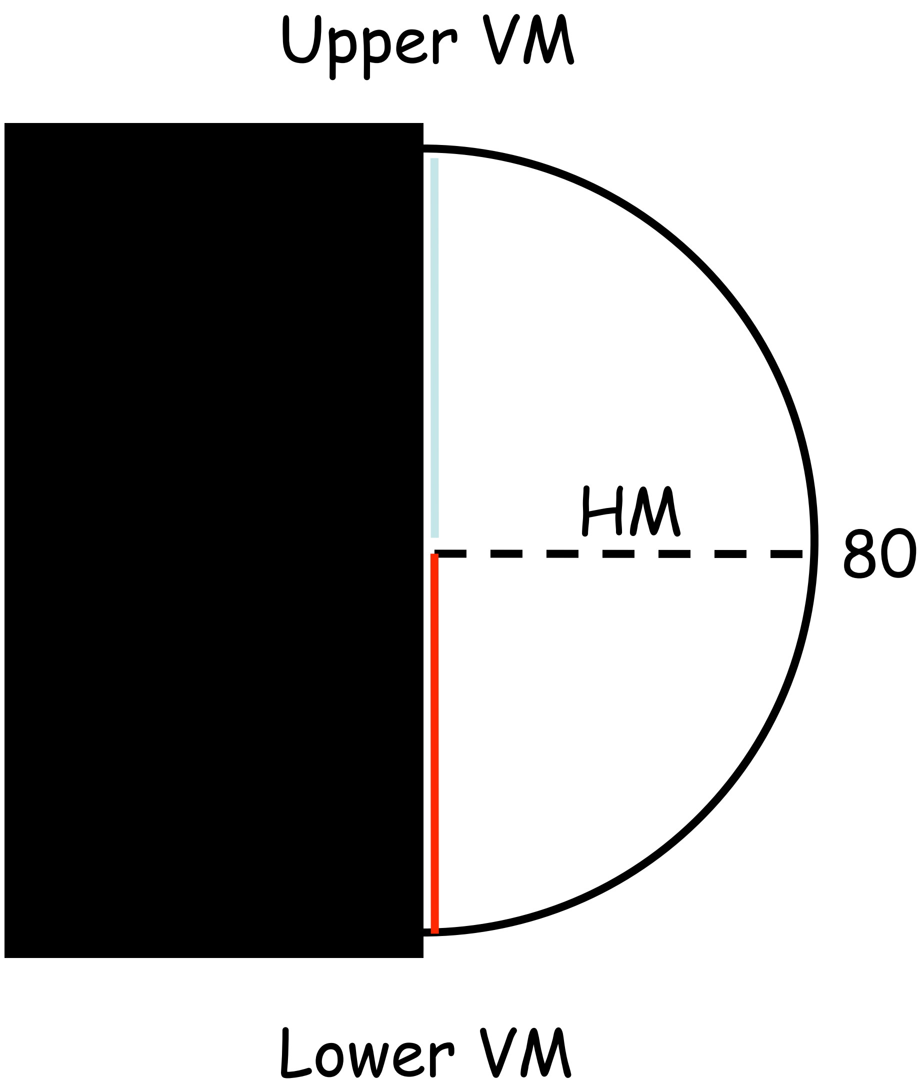
V1 retinotopy



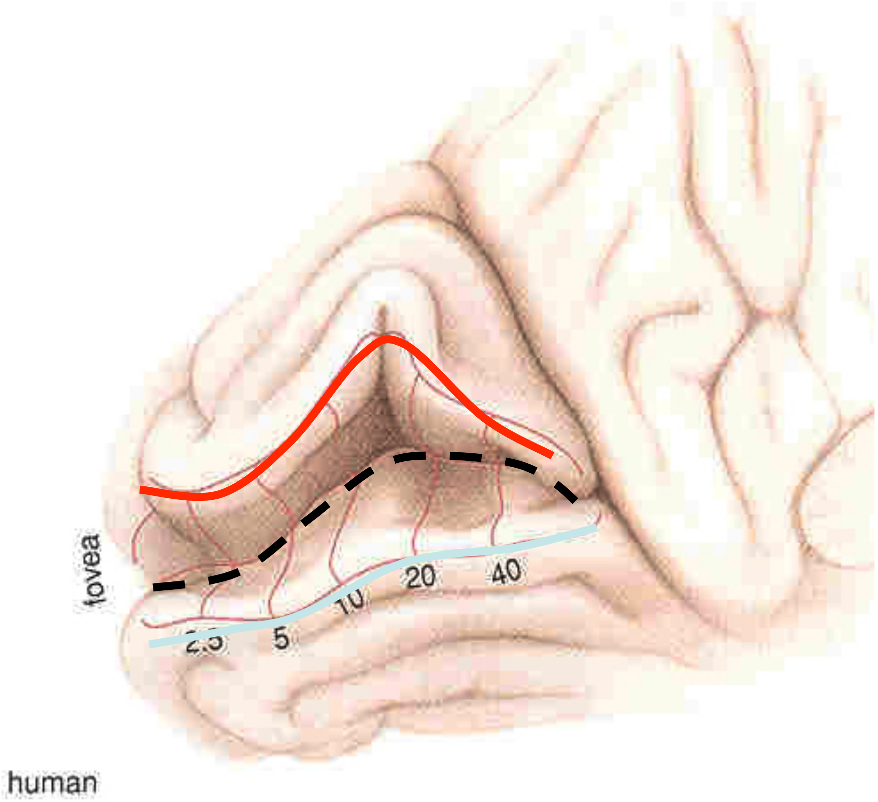


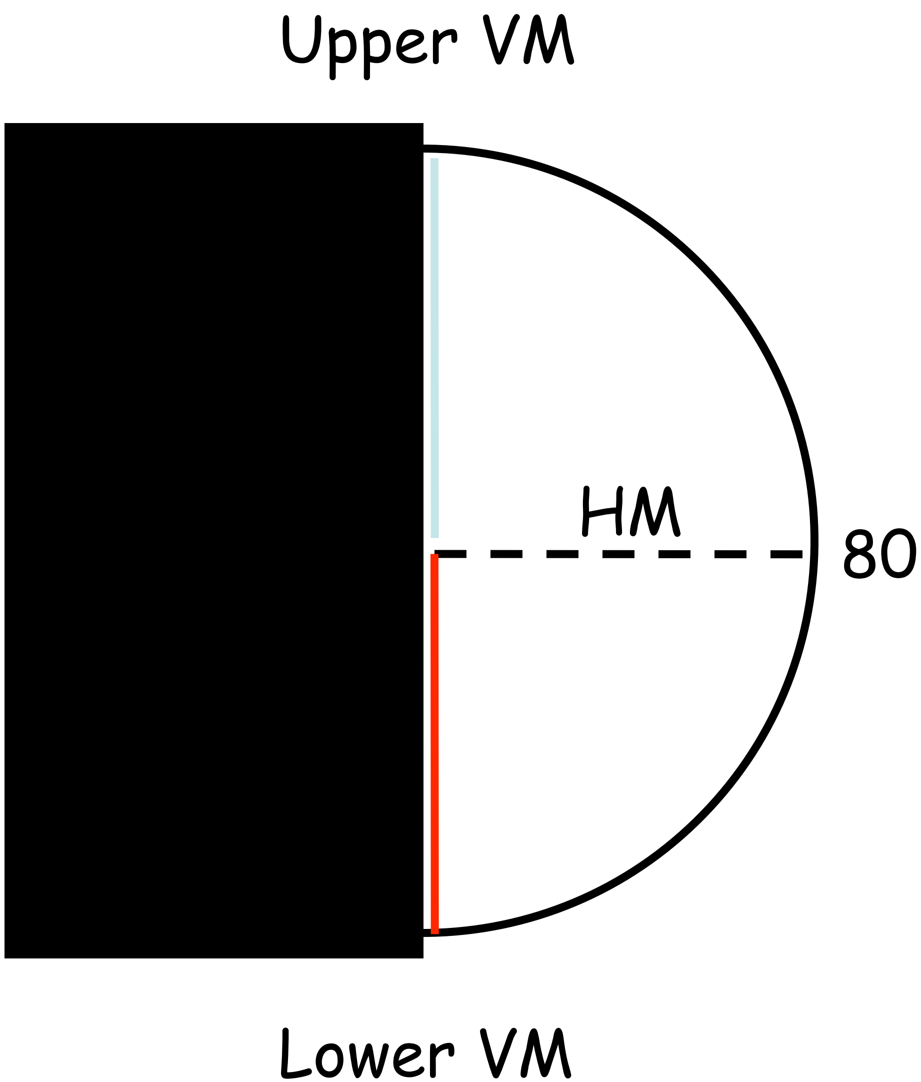
V1 retinotopy



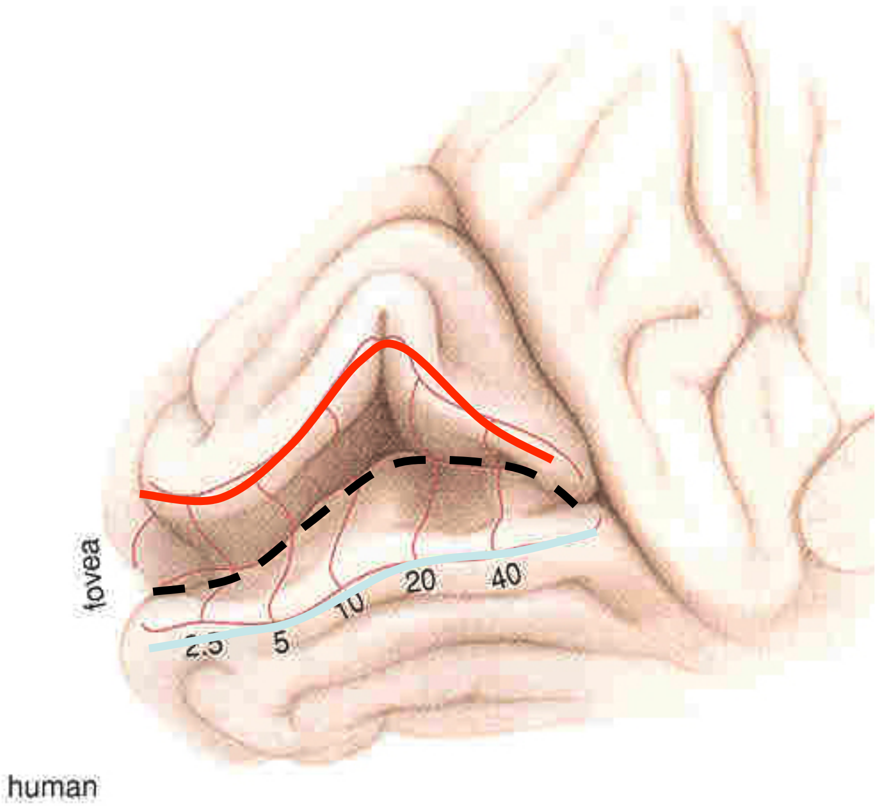


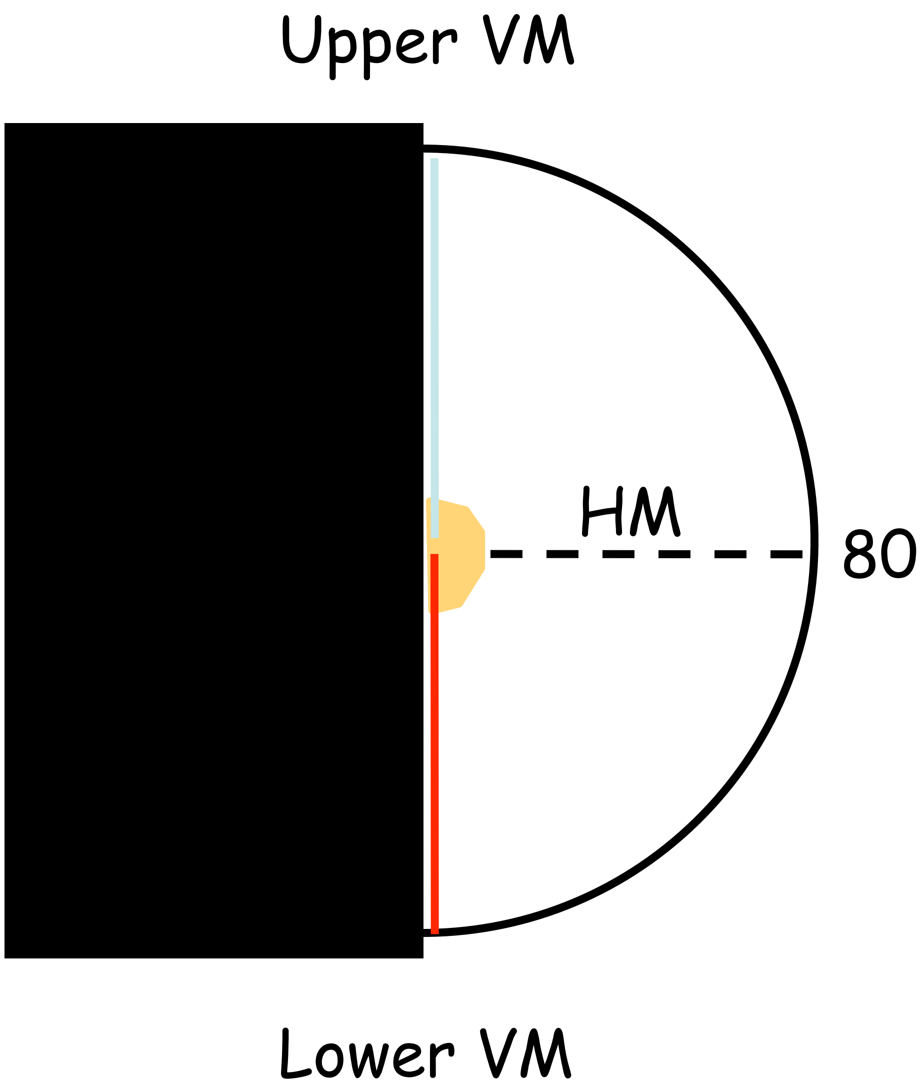
V1 retinotopy



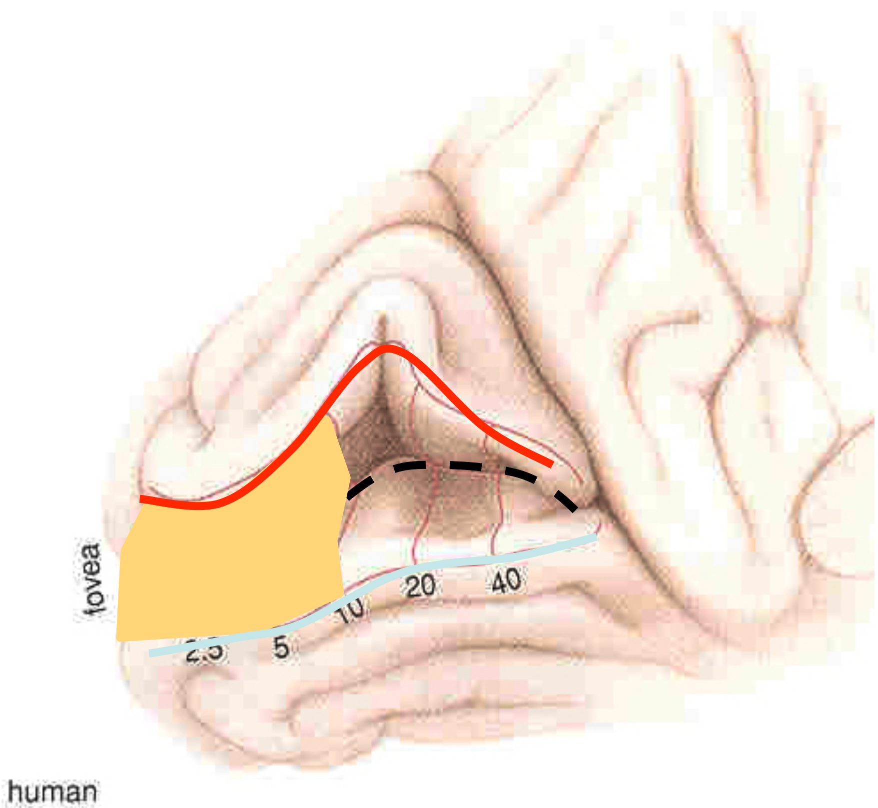


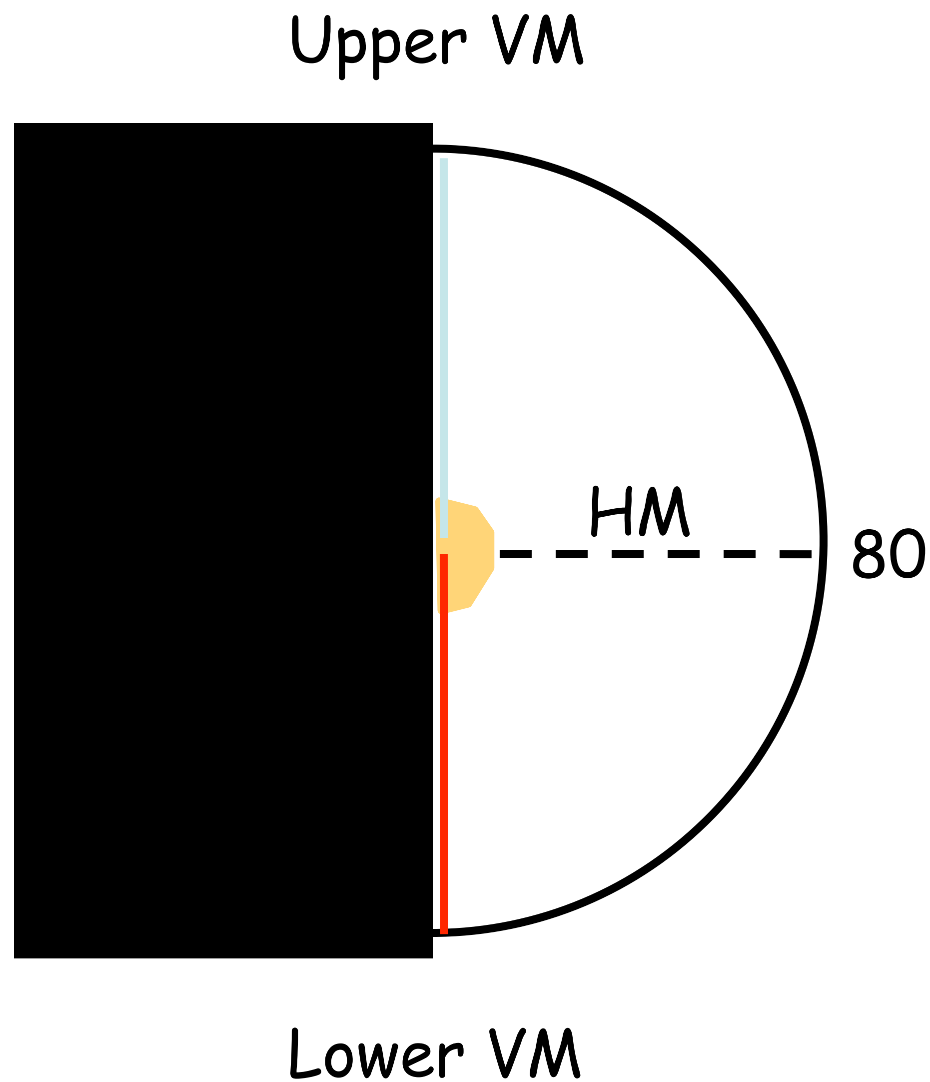
V1 retinotopy



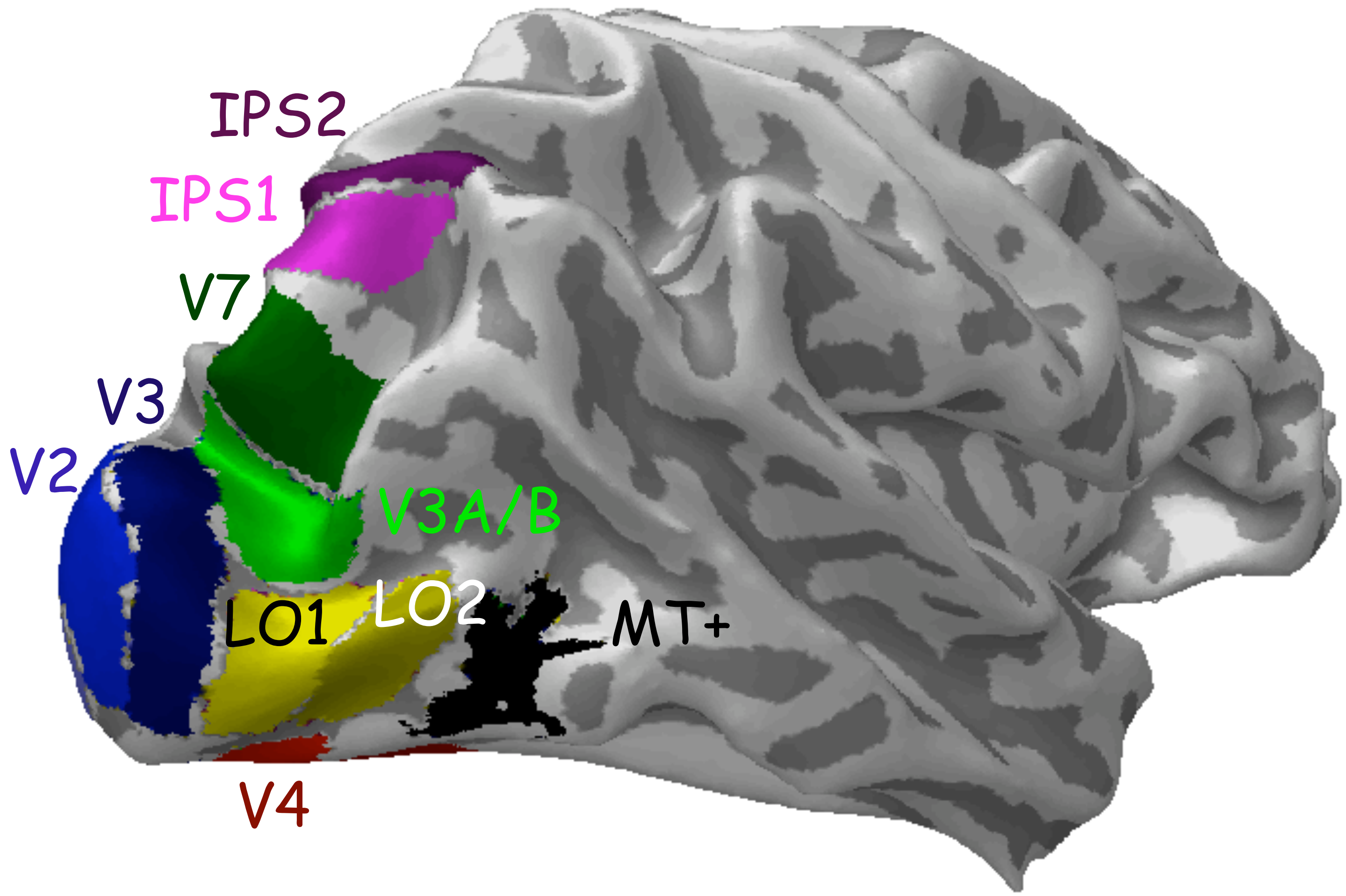


V1 retinotopy



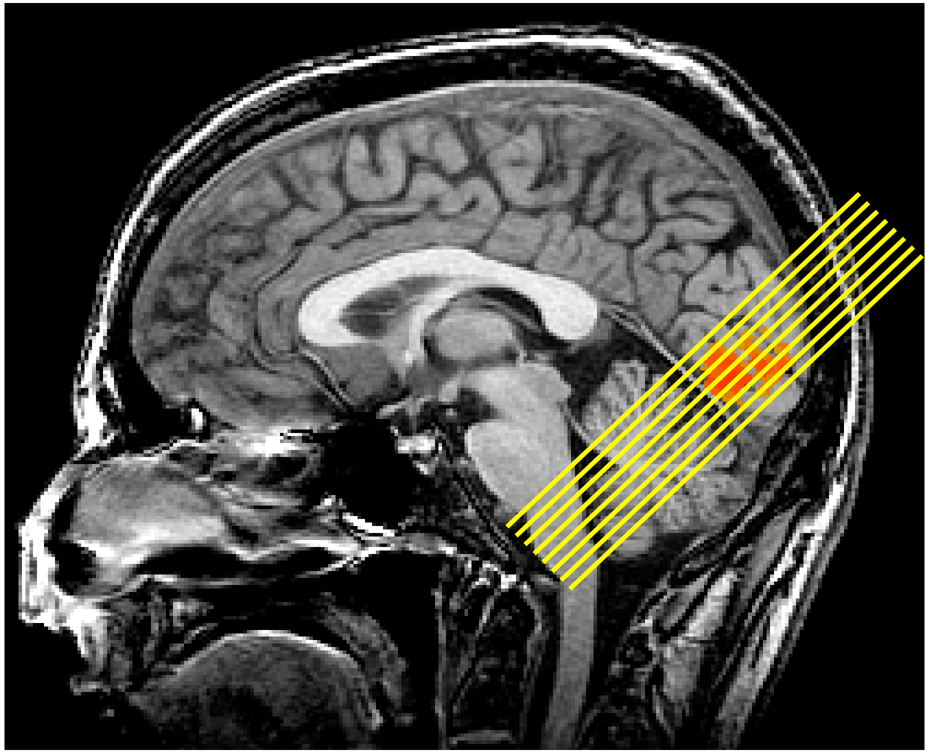


Visual maps in the brain



Each visual brain area contains a map of the visual world and performs a different function.

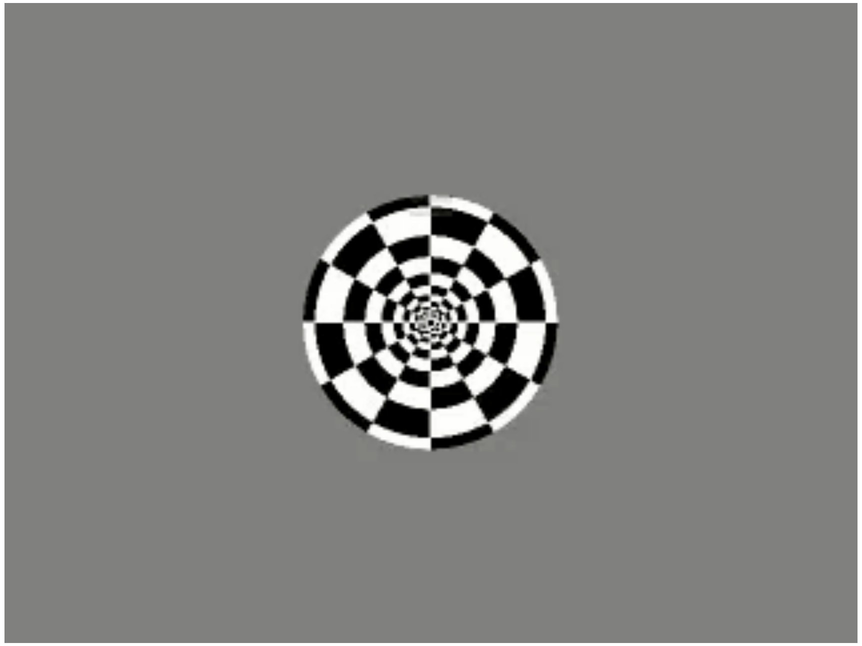
Functional magnetic resonance imaging



Revolution in psychology and neuroscience: > 1000 papers published per year!

Measuring retinotopic maps

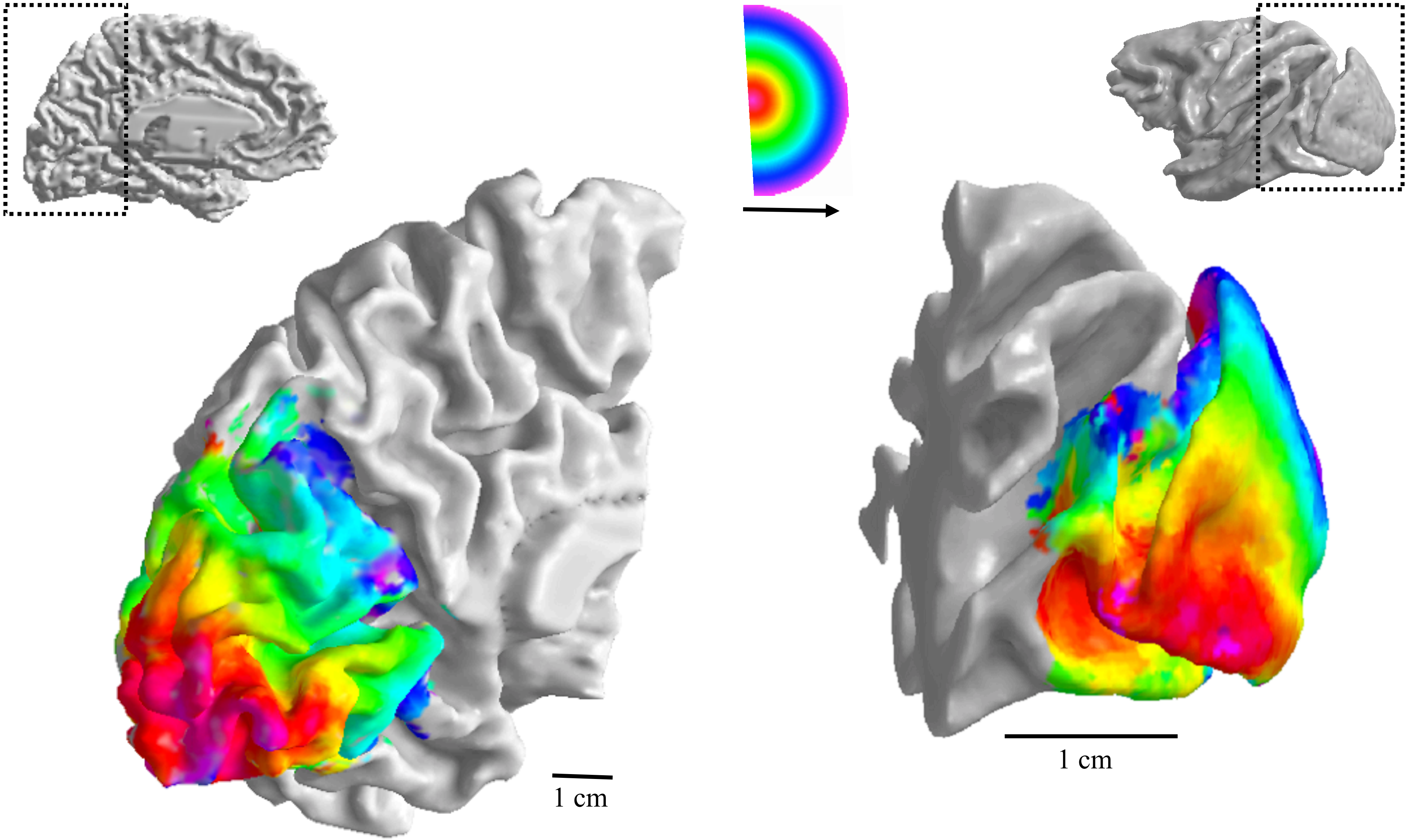
Radial component



Angular component

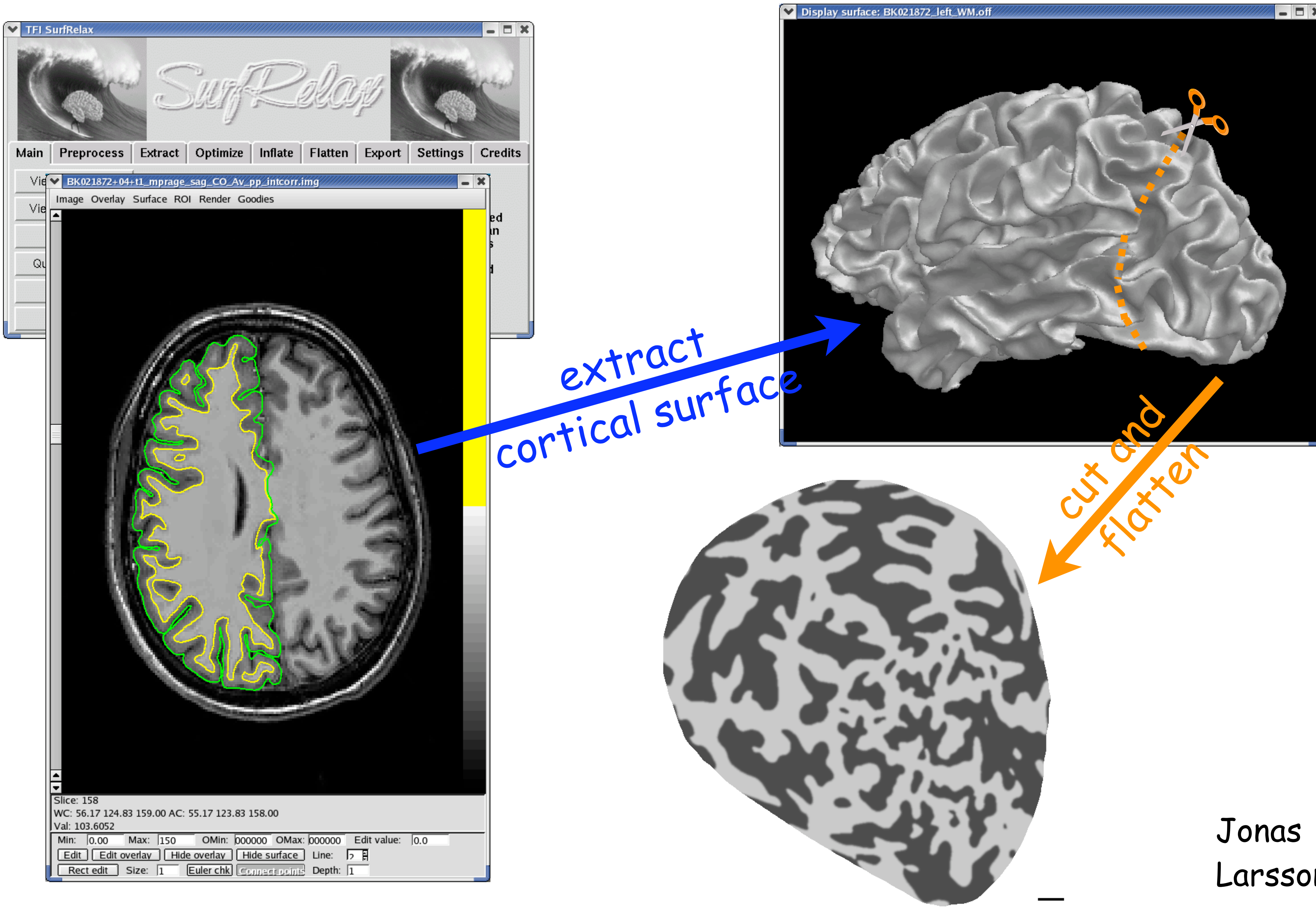
Engel et al, Nature (1994)

Retinotopy: radial component



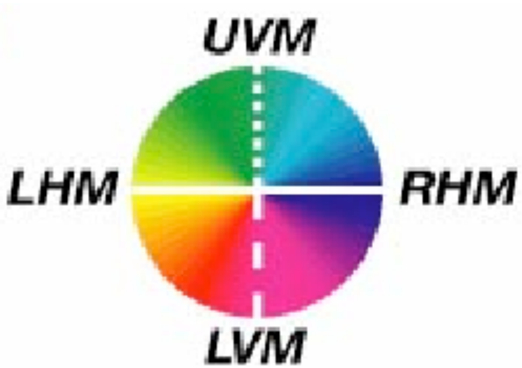
Brewer, Press, Logothetis & Wandell (2002)

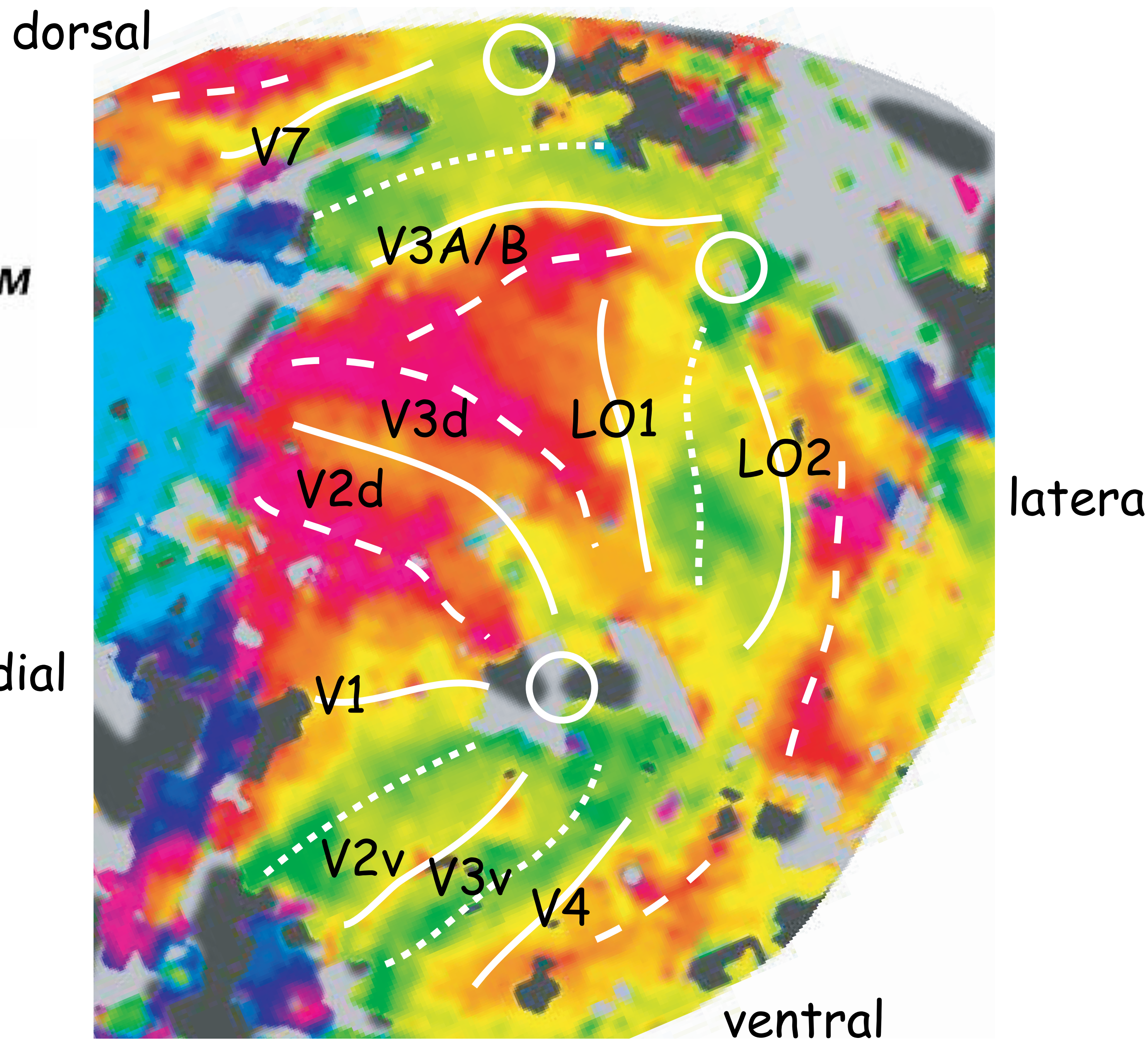
Cortical segmentation & flattening



Jonas Larsson

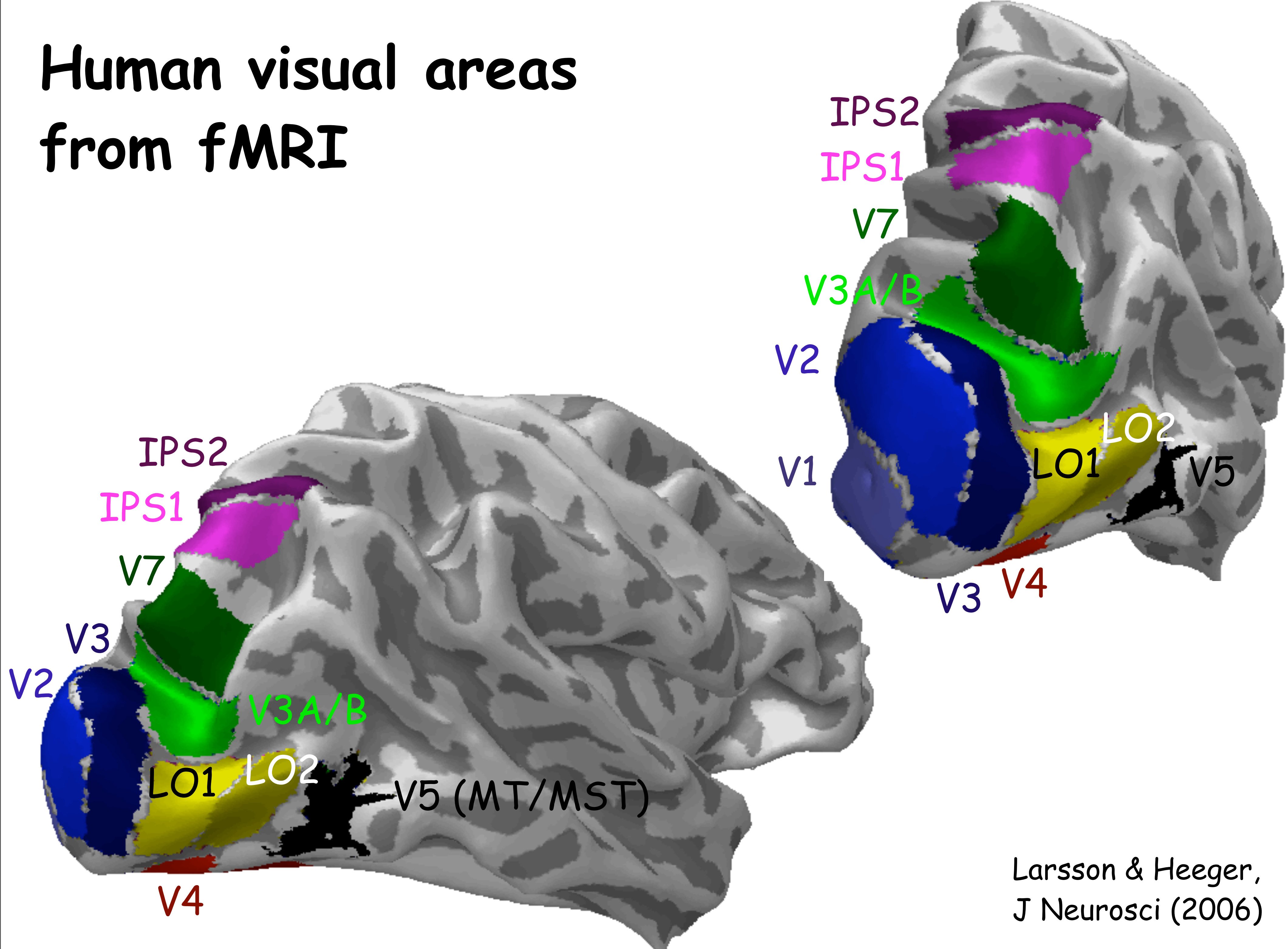
Retinotopy: angular component





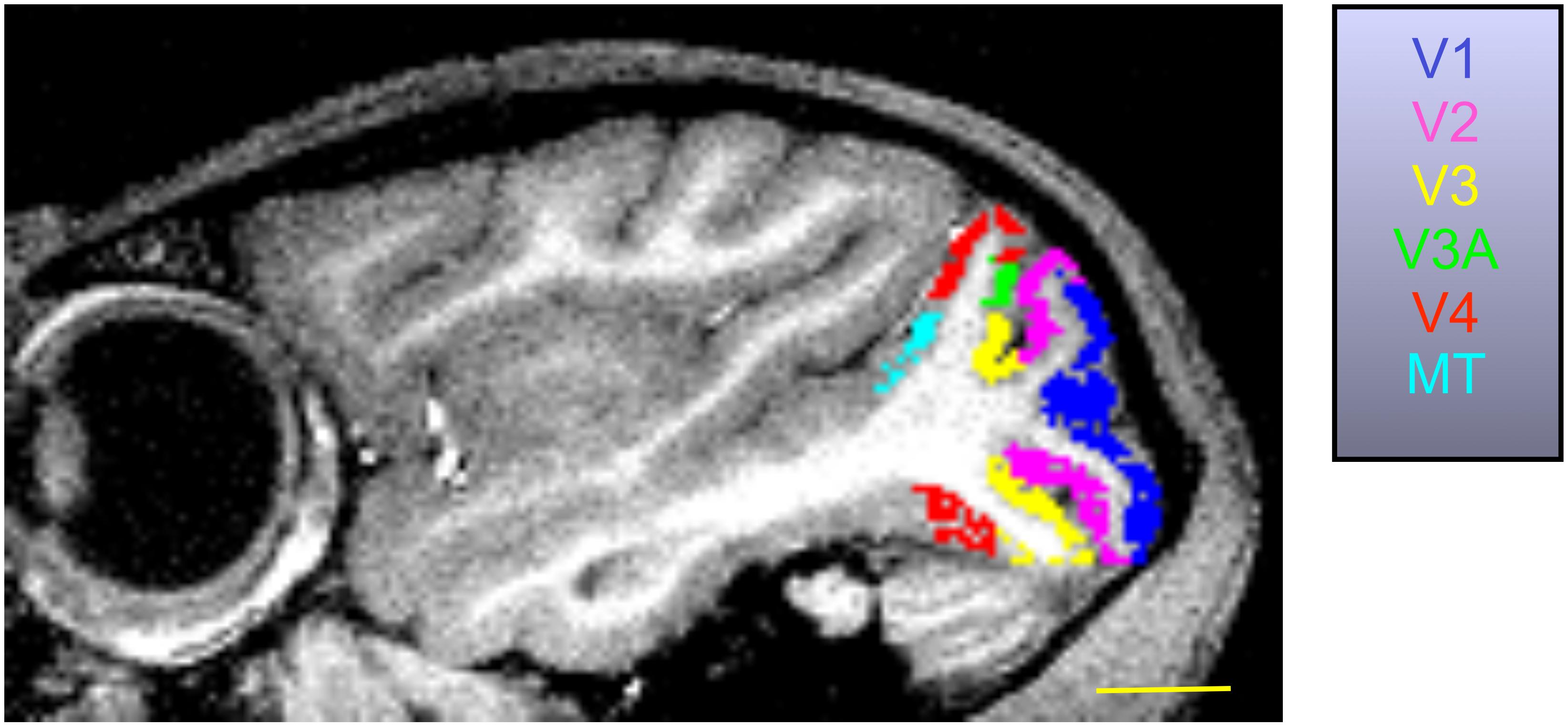
Larsson & Heeger,
J Neurosci, 2006

Human visual areas from fMRI



Larsson & Heeger, J Neurosci (2006)

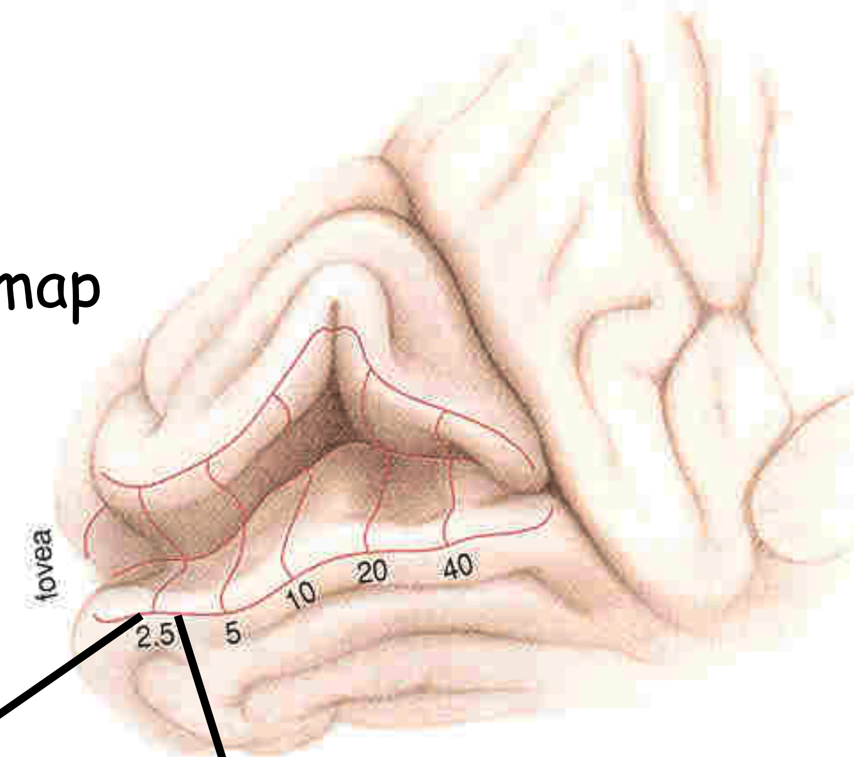
Monkey visual areas from fMRI



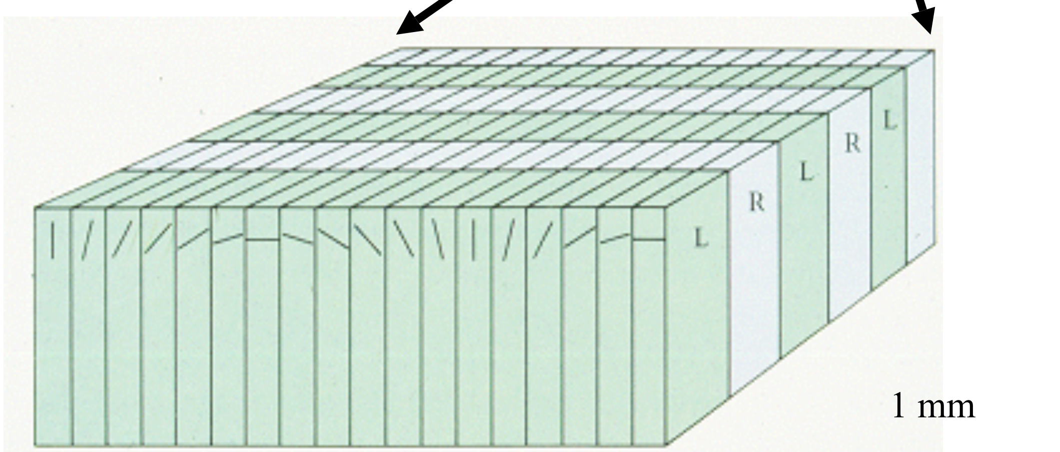
Brewer, Press, Logothetis & Wandell (2002)

Topography: columnar architecture in V1

Retinotopic map



Columnar architecture

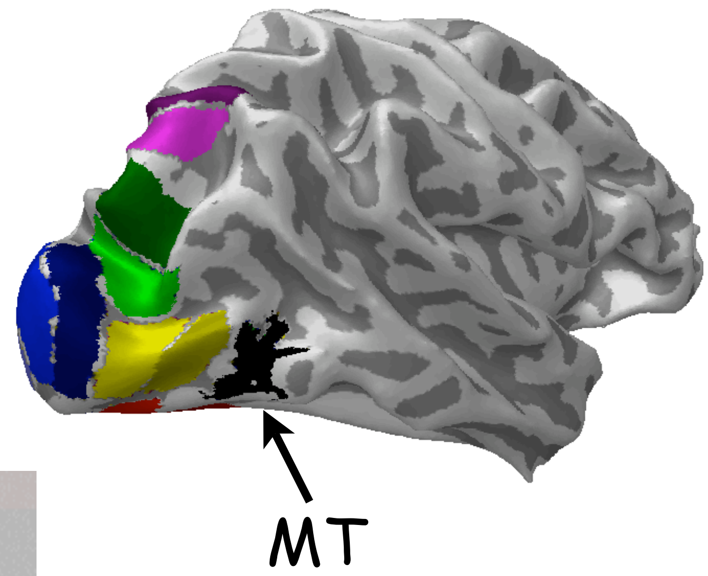


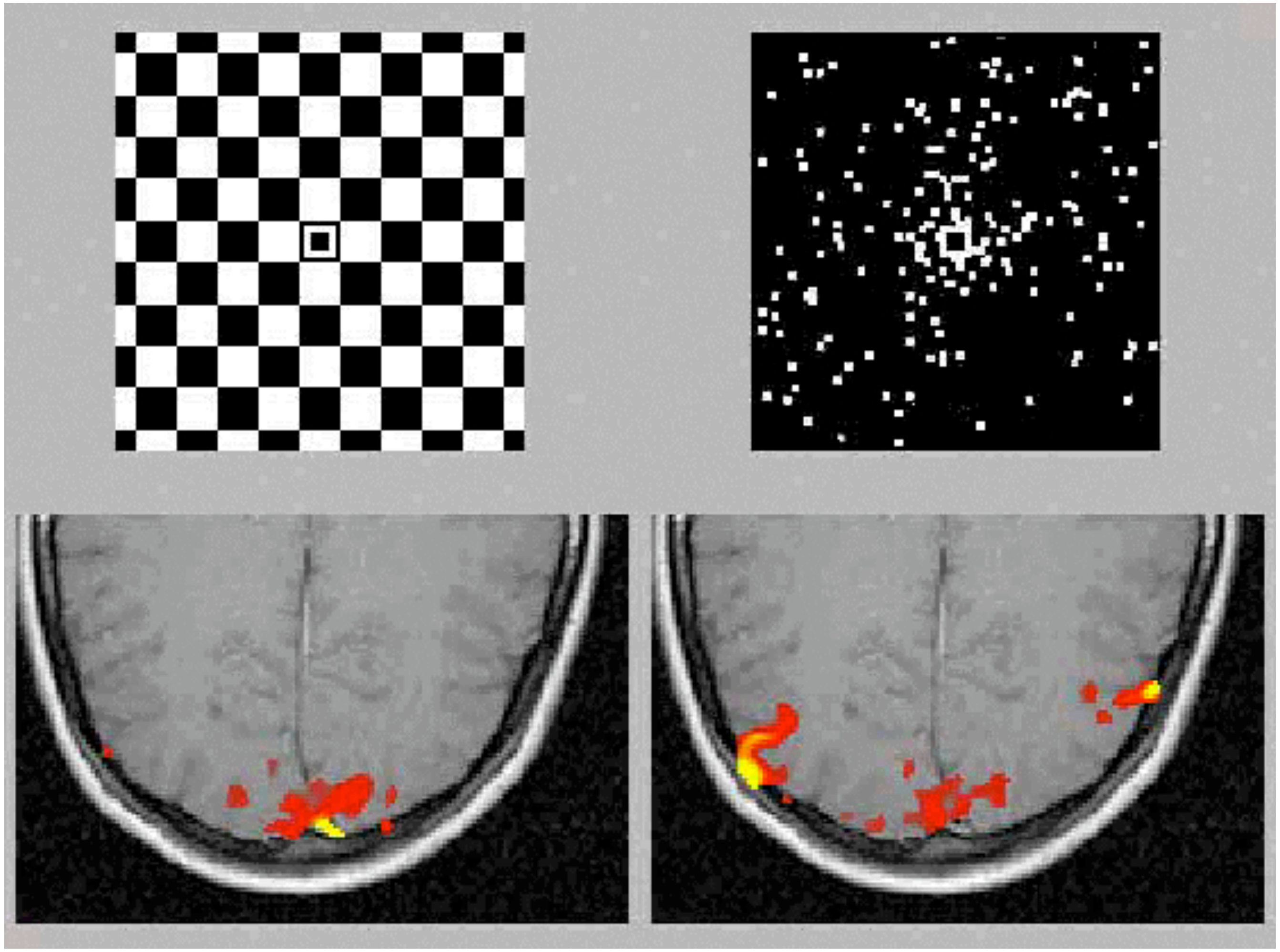
Functional specialization

Match each cortical area to its corresponding function:

V1	Motion
V2	Stereo
V3	Color
V3A	Texture
V3B	Segmentation, grouping
V4	Recognition
V5	Attention
V7	Working memory
LO1	Mental imagery
IPS1	Decision-making
IPS2	Sensorimotor integration
Etc.	Etc.

Functional specialization: motion perception





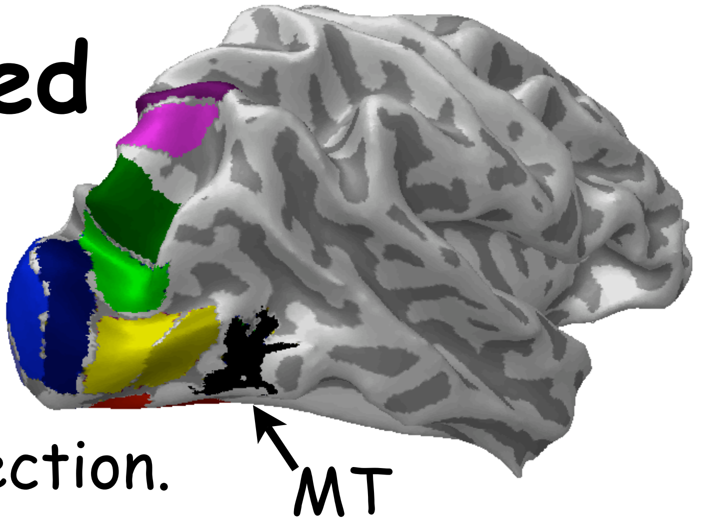
Geoff Boynton
Alex Huk

Beware of circular reasoning in functional specialization

1. Hypothesize that there is a particular visual process that is localized to a functionally specialized brain area.
2. Design an experiment with two stimuli/tasks, one of which you believe imposes a greater demands on that visual process.
3. Run the experiment and find sure enough that there are some neurons in a brain area that respond more strongly during trials with high demand on that visual process then low demand trials.

What can you conclude from this?

Cortical area MT is specialized for visual motion perception



- Neurons in MT are **selective** for motion direction.
- Neural responses in MT are **correlated** with the perception of motion.
- Damage to MT or temporary inactivation **causes** deficits in visual motion perception.
- Electrical stimulation in MT **causes** changes in visual motion perception.
- Computational **theory** quantitatively explains both the responses of MT neurons and the perception of visual motion.
- Well-defined **pathway** of brain areas (cascade of neural computations) underlying motion specialization in MT.

Is MT specialized for only visual motion perception?

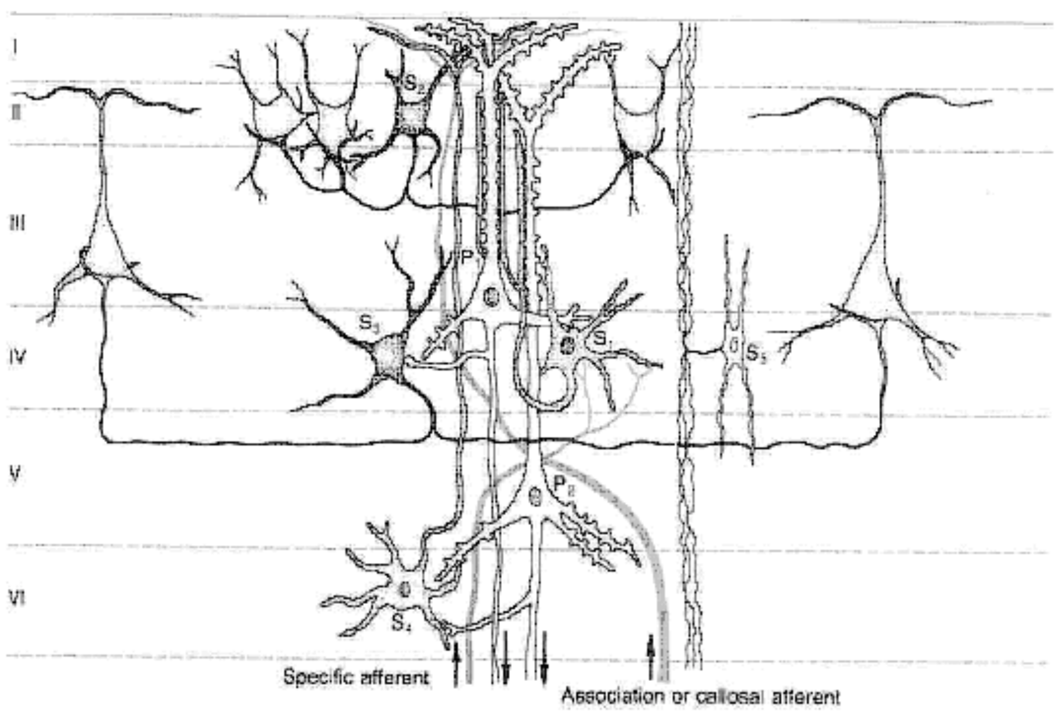
- Neurons in MT are also selective for binocular disparity.
- Neural responses in MT are also correlated with the perception of depth.
- Motion discrimination performance mostly recovers following carefully circumscribed lesions to MT in monkeys.
- Electrical stimulation in MT causes changes in stereo depth perception.

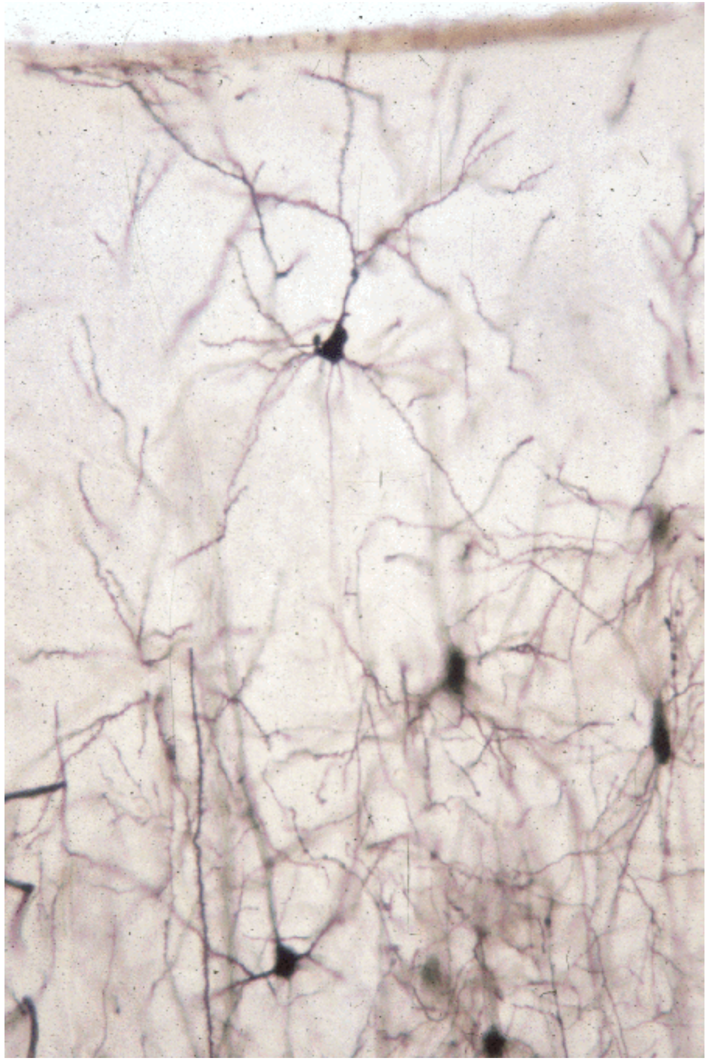
Is MT specialized for only visual motion perception?

- Neurons in MT are also selective for binocular disparity.
- Neural responses in MT are also correlated with the perception of depth.
- Motion discrimination performance mostly recovers following carefully circumscribed lesions to MT in monkeys.
- Electrical stimulation in MT causes changes in stereo depth perception.

Even so... **computational theory** quantitatively explains the responses of MT neurons.

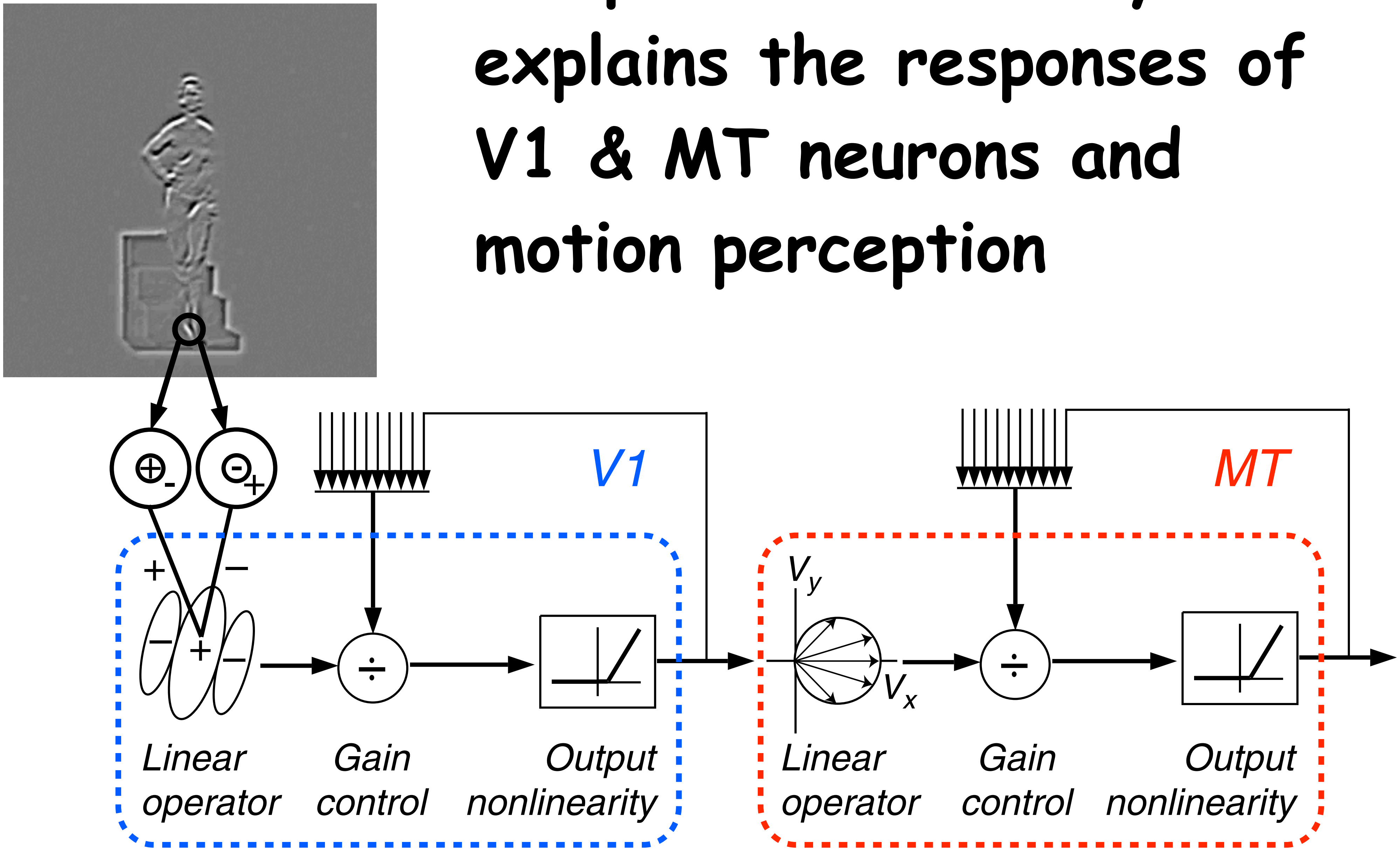
Neural circuits perform computations





~50,000 neurons per cubic mm
~6,000 synapses per neuron
~10 billion neurons & ~60 trillion synapses in cortex

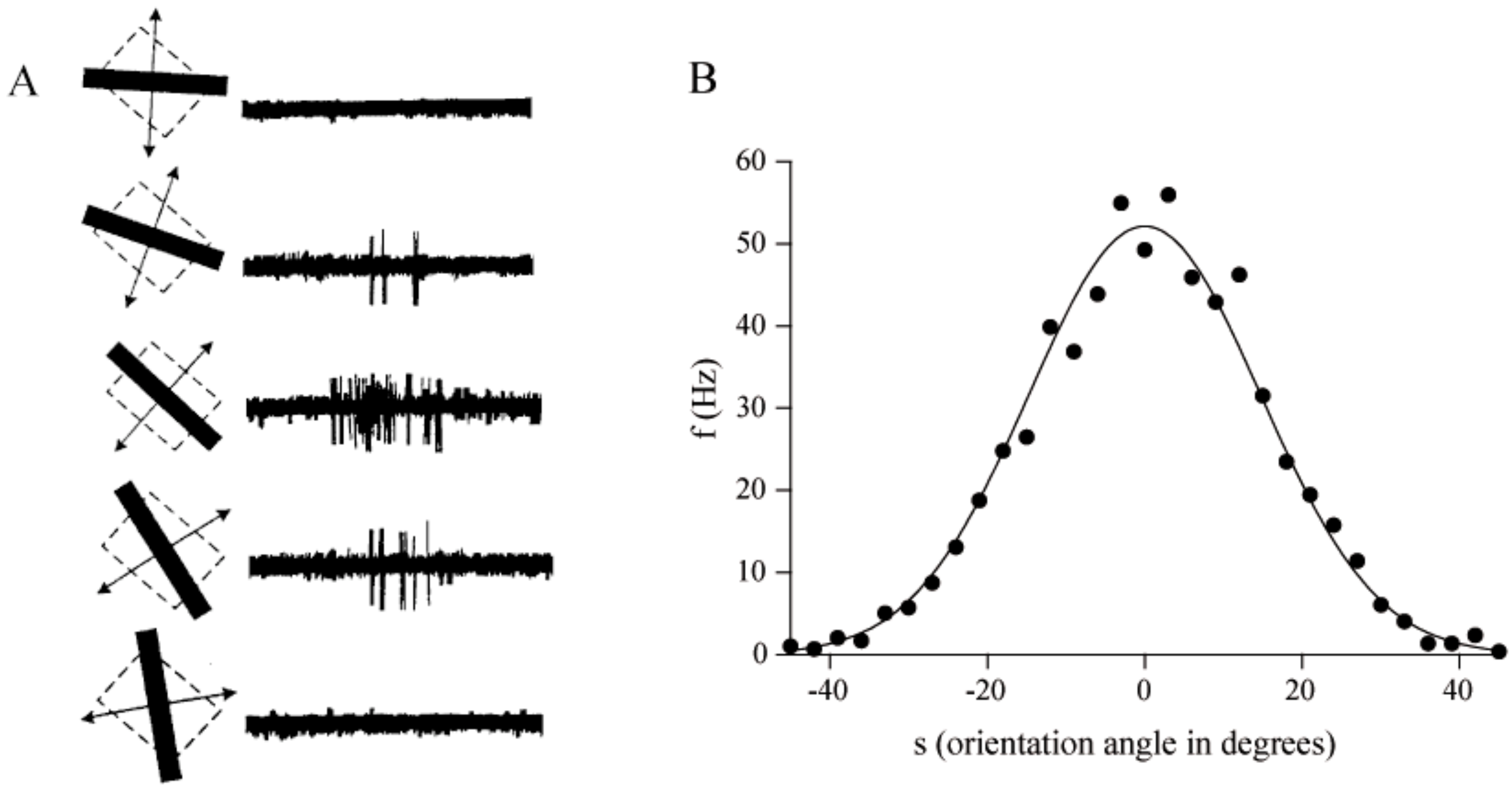
Computational theory explains the responses of V1 & MT neurons and motion perception



with Matteo Carandini, Tony Movshon, Eero Simoncelli

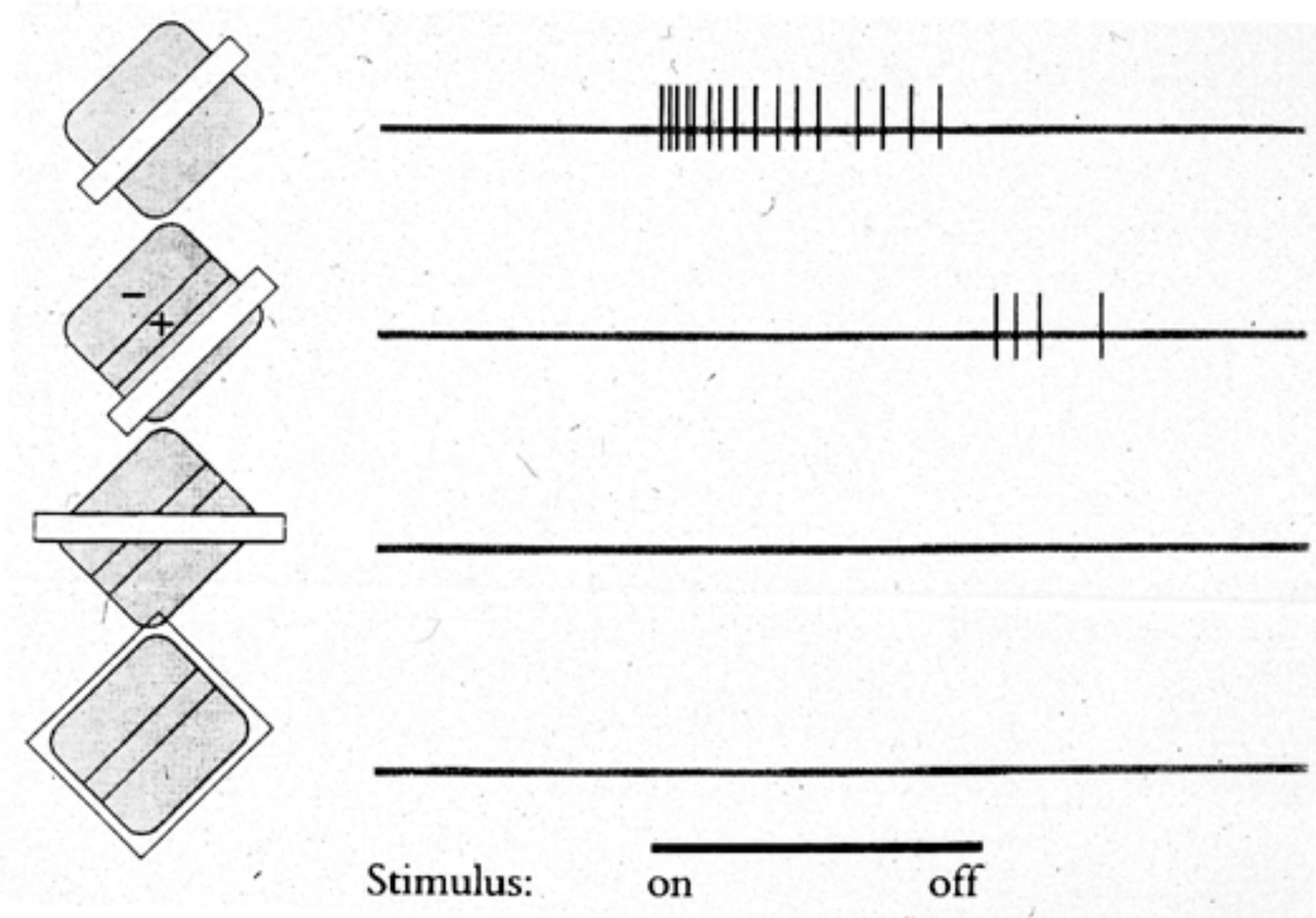
V1 physiology and computational theory

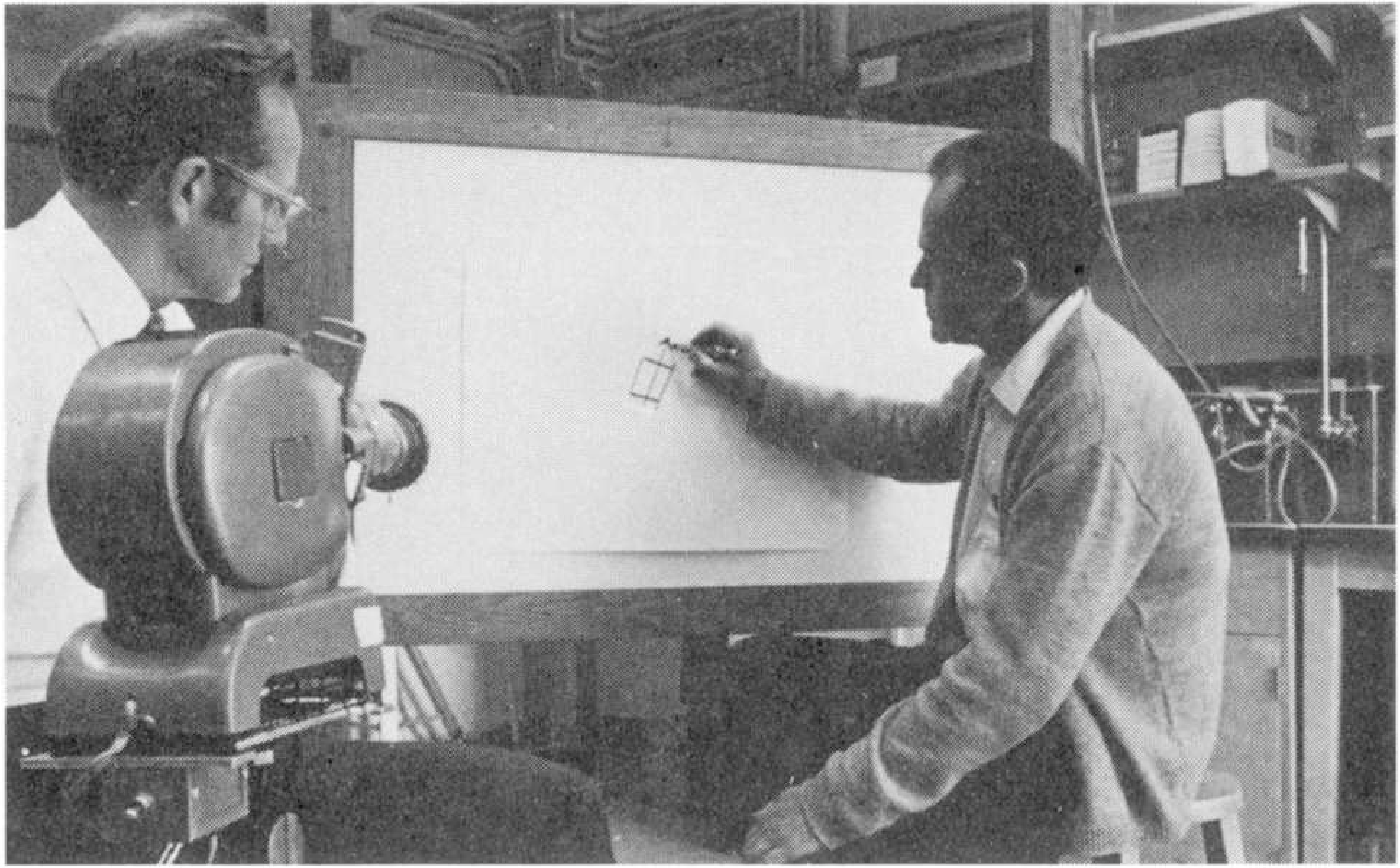
V1 orientation selectivity



Hubel & Wiesel (1968)

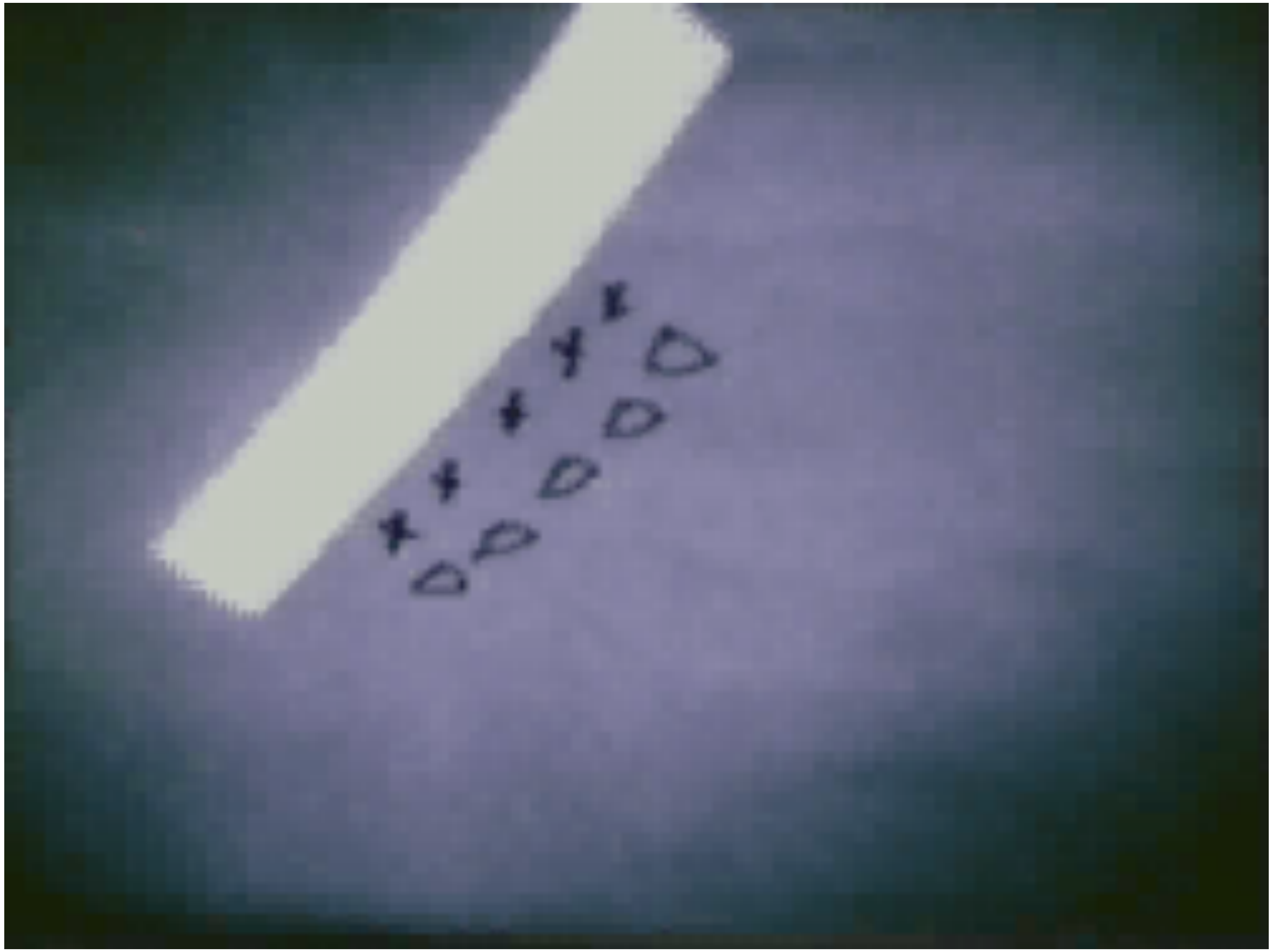
Simple cell



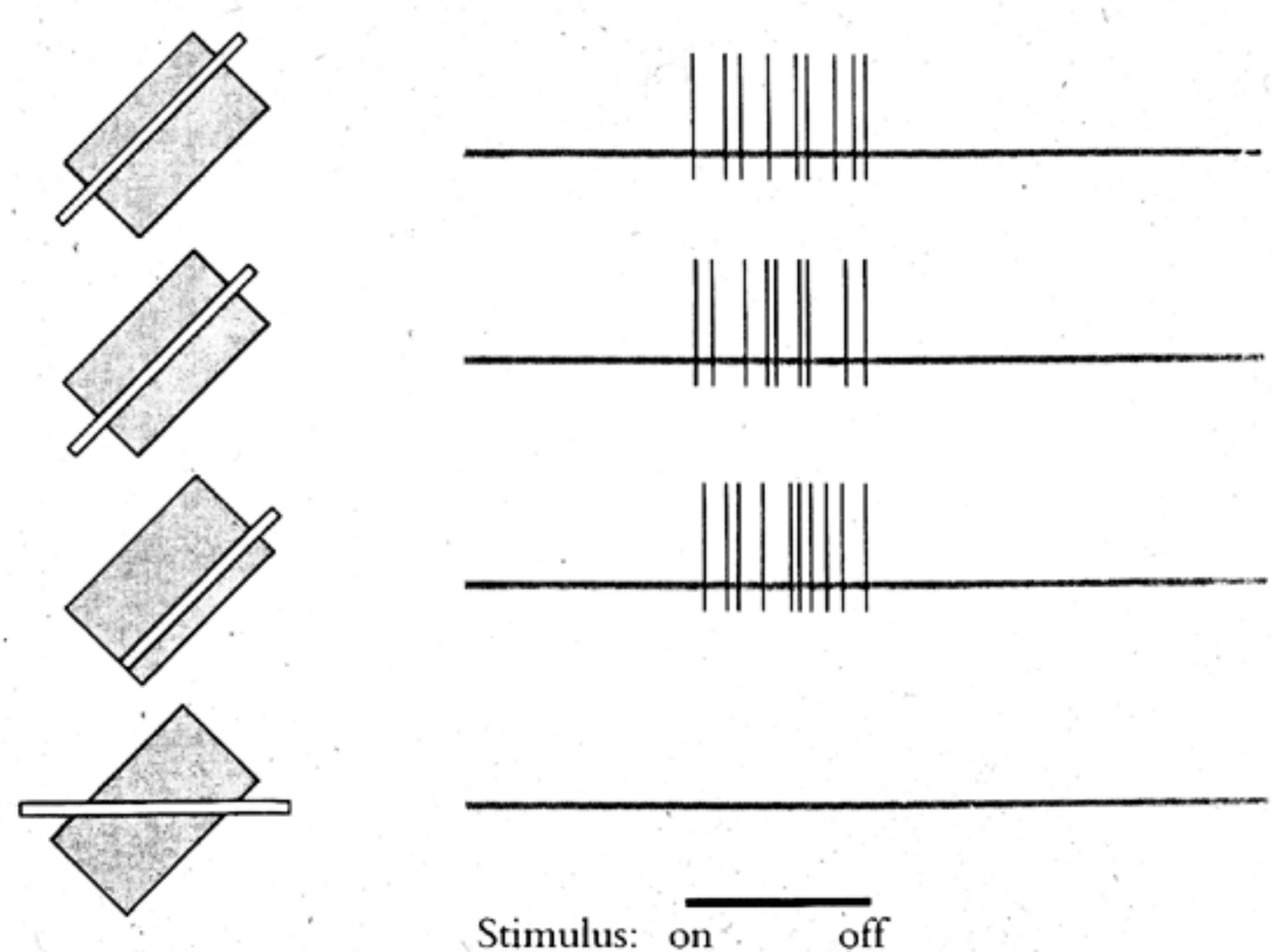


Hubel & Wiesel movie

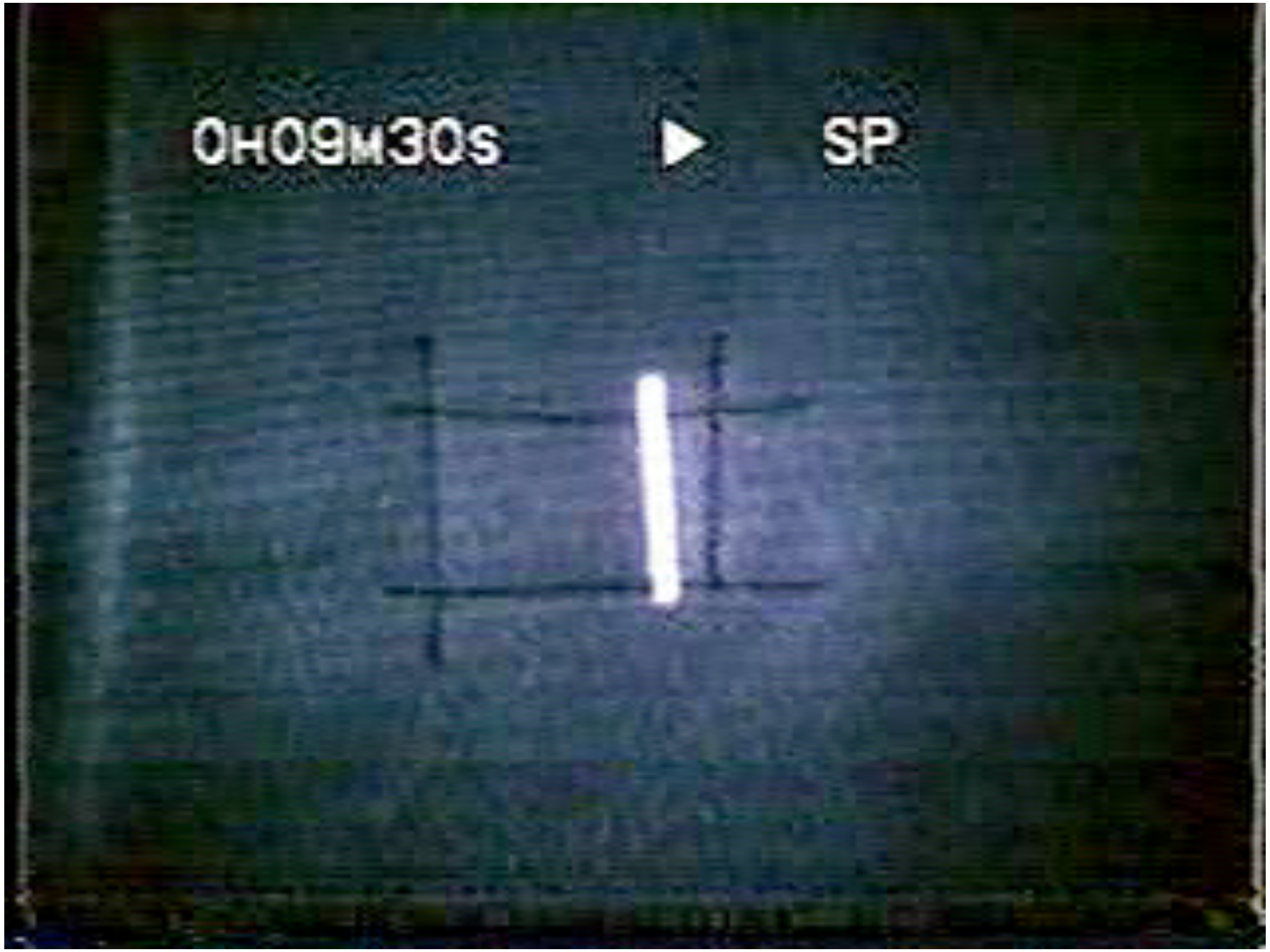
Simple cell



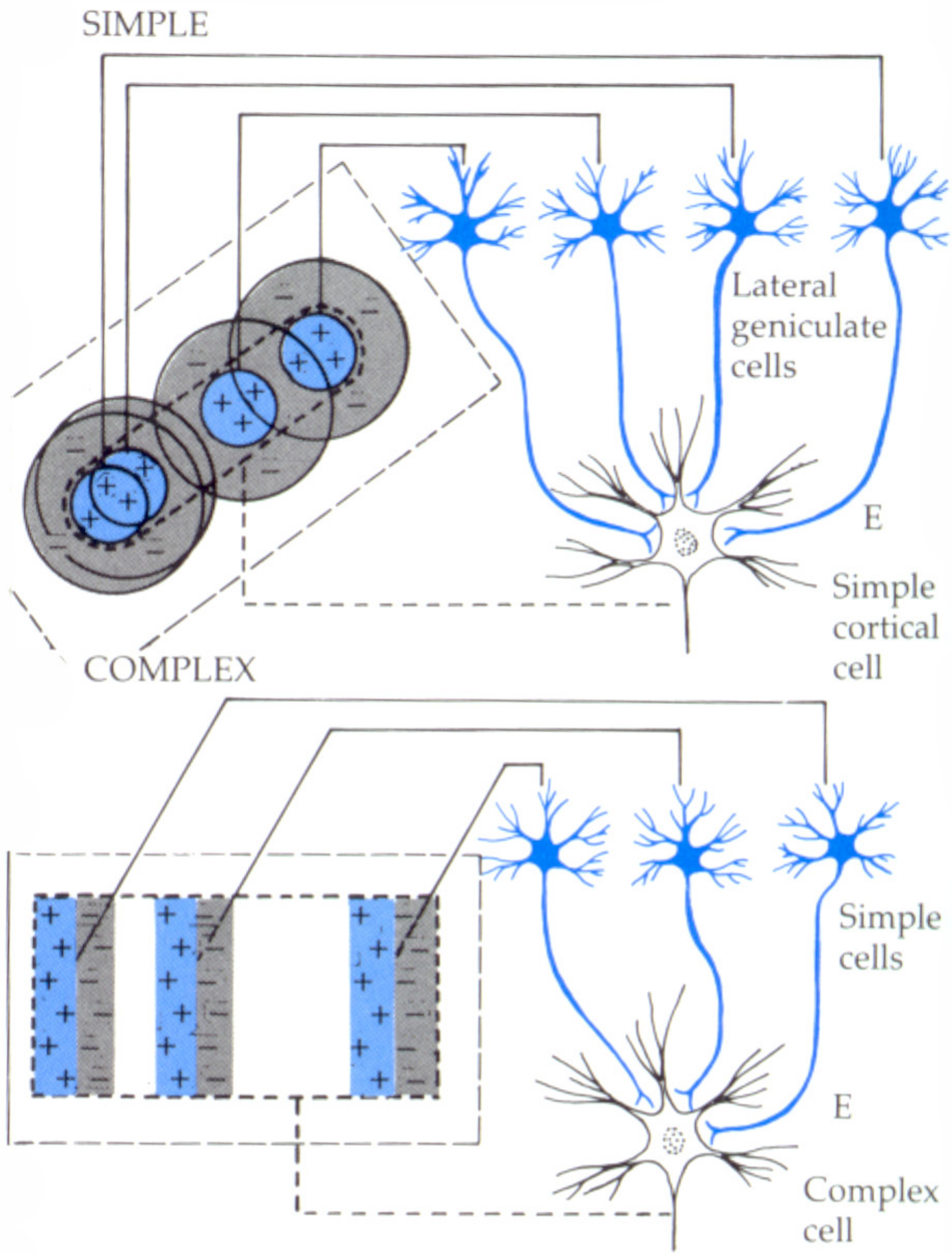
Complex cell



Complex cell

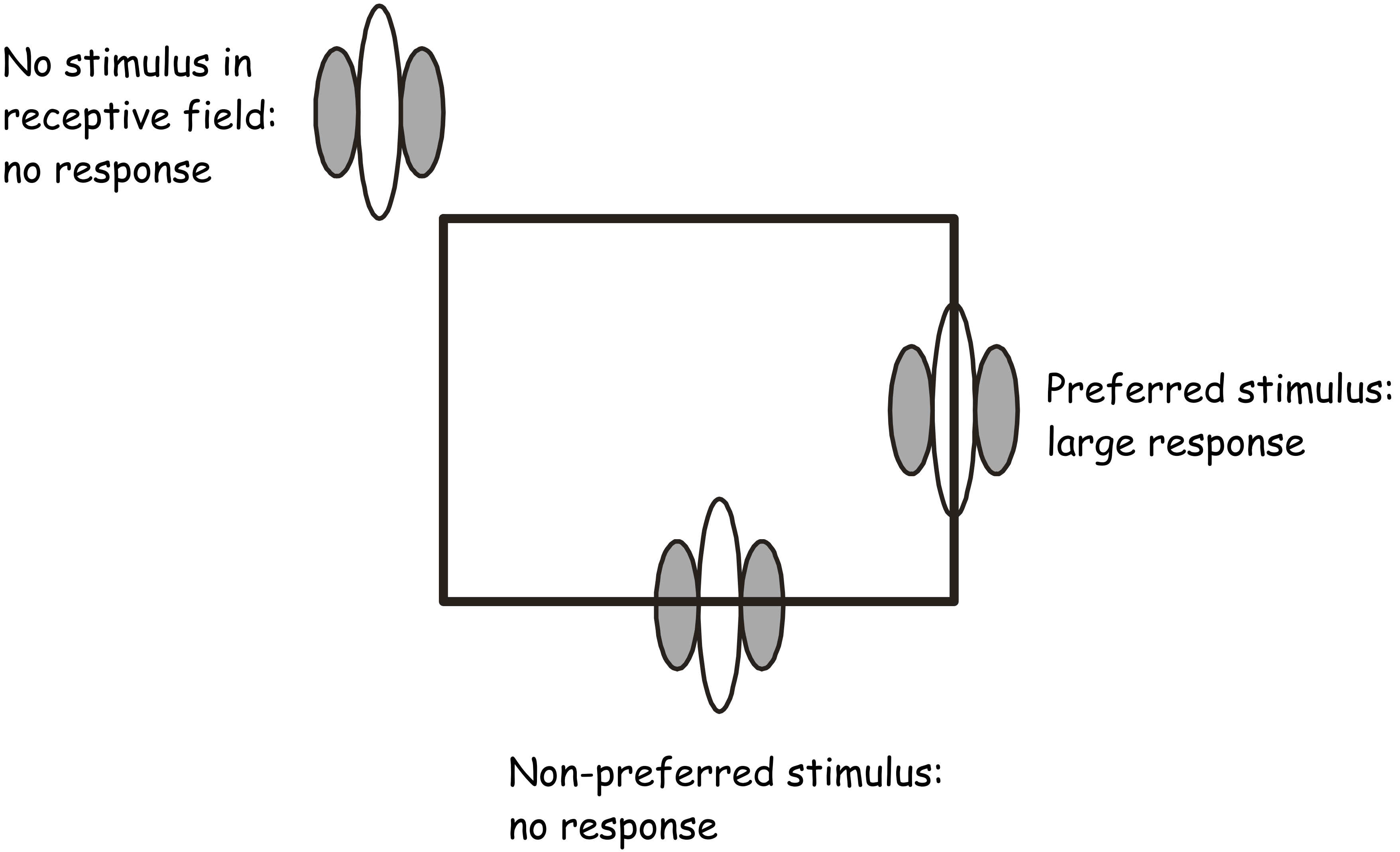


Classical view: summation & spike threshold

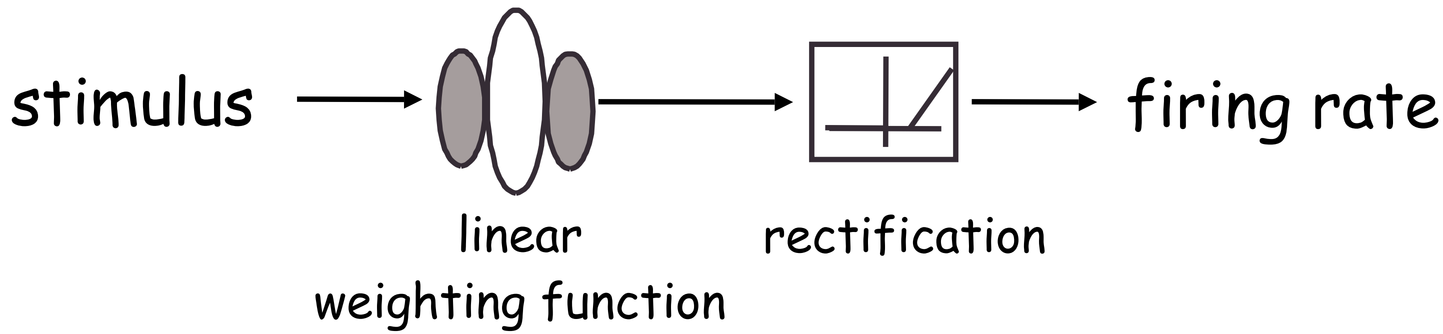


Hubel & Wiesel (1962)

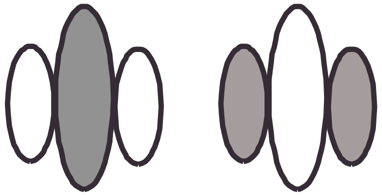
Orientation selectivity model



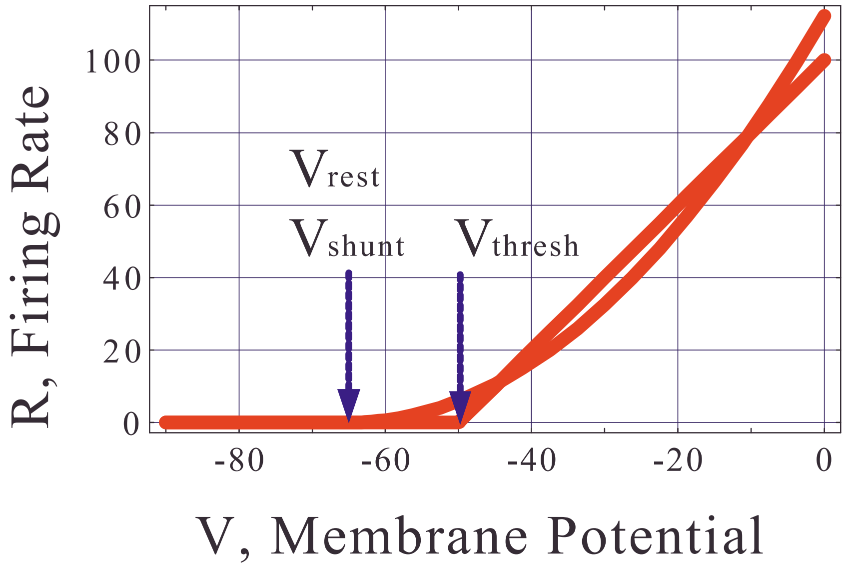
Rectification and spiking threshold



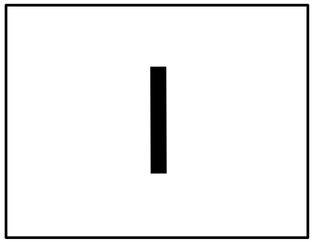
Complementary receptive fields



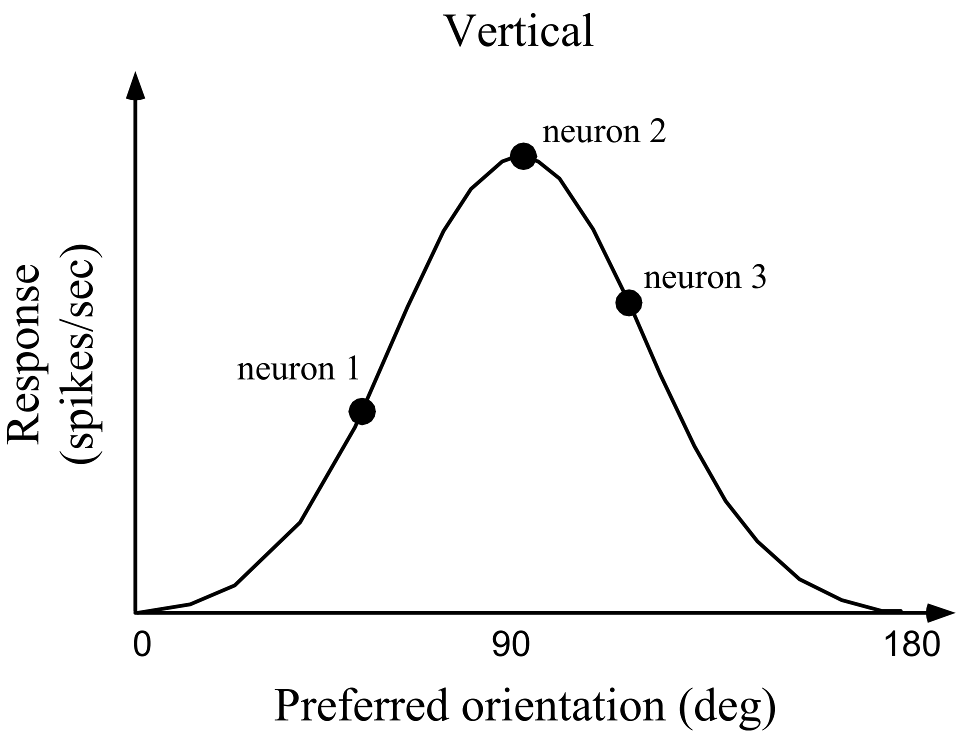
Rectification and squaring



Distributed representation of orientation



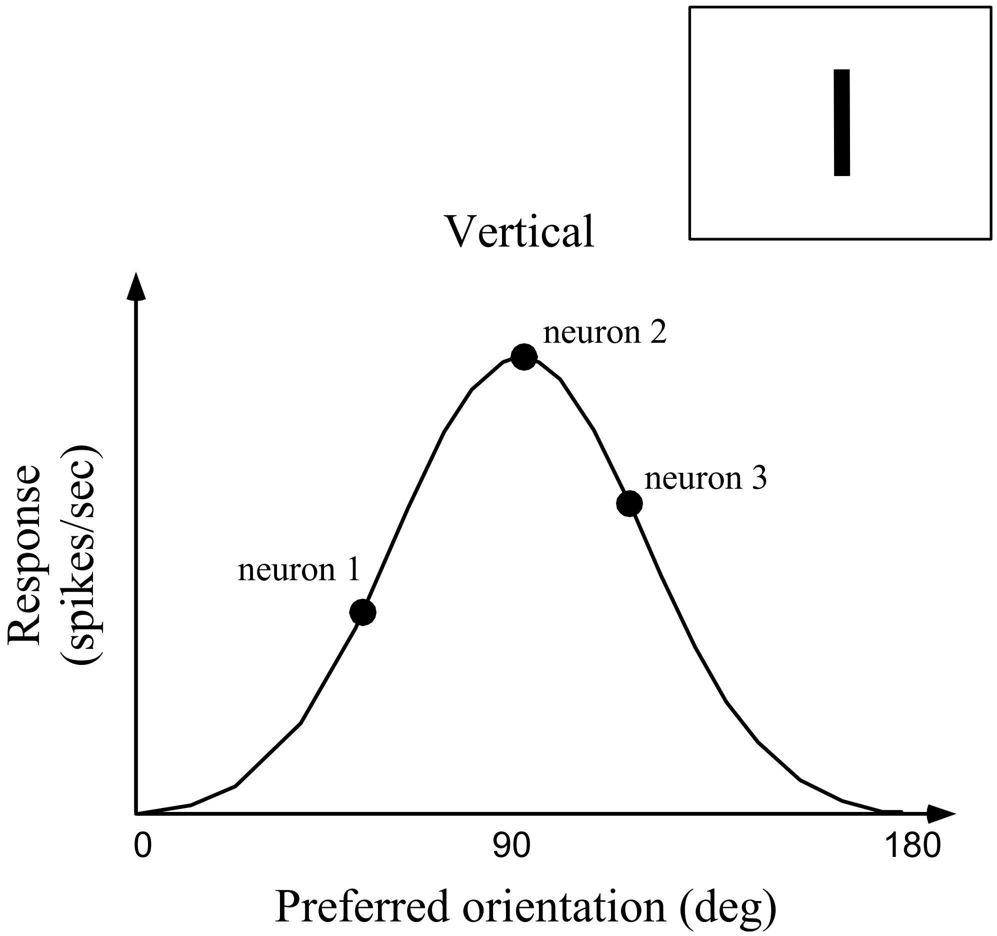
Stimulus: vertical bar

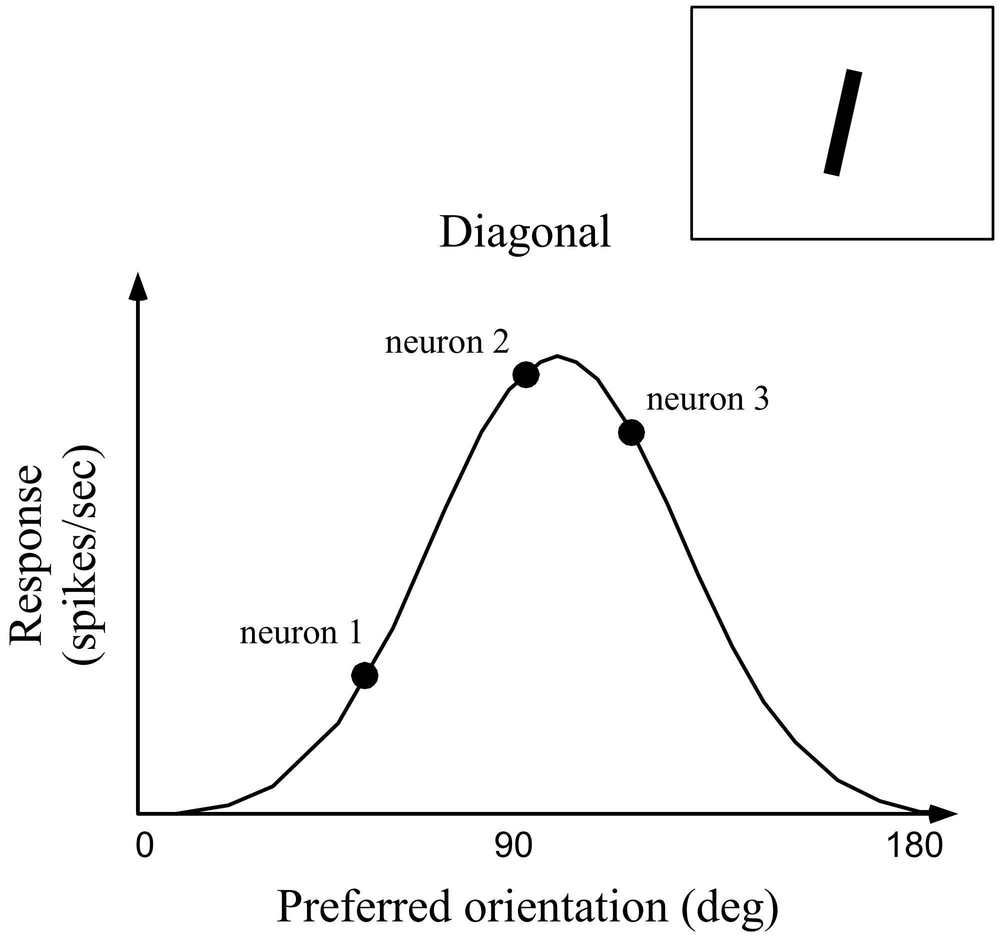


Responses of each of several orientation tuned neurons.

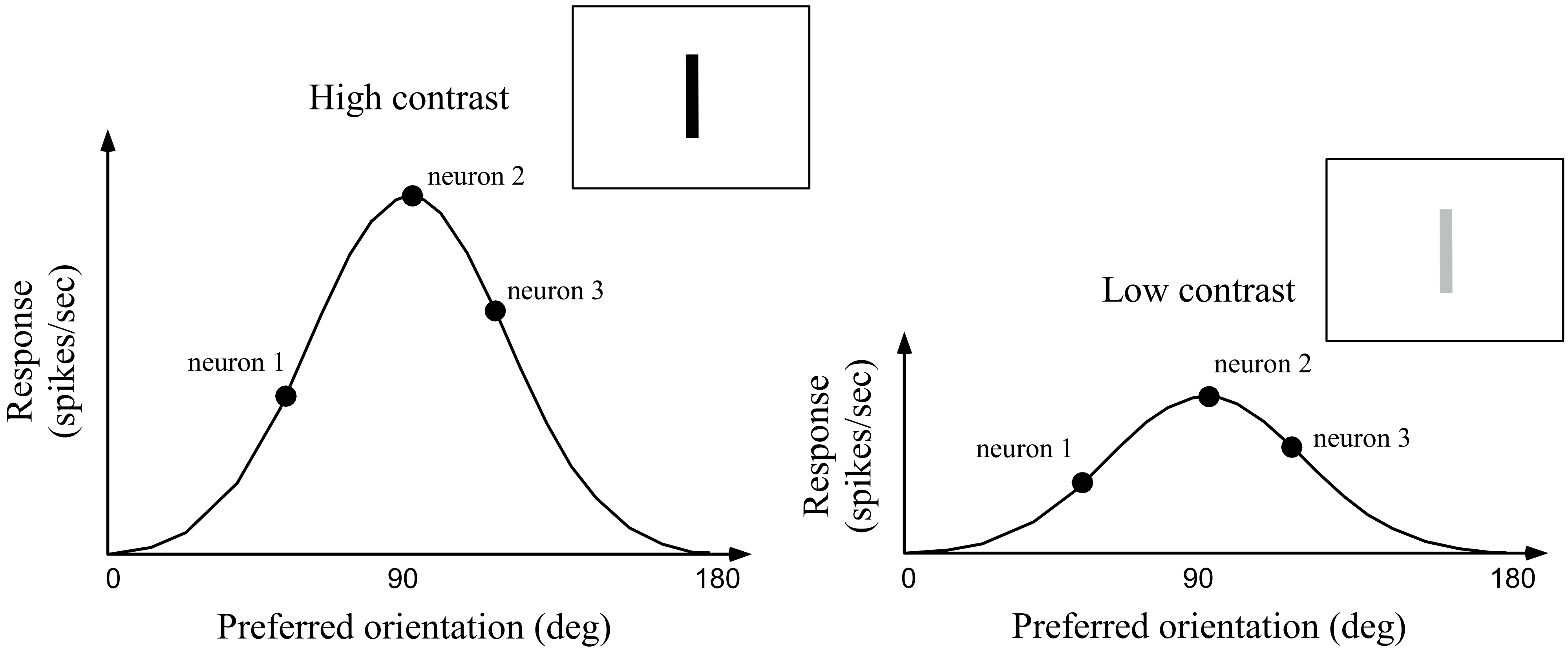
Peak (distribution mean) codes for stimulus orientation.

Broad tuning can code for small changes

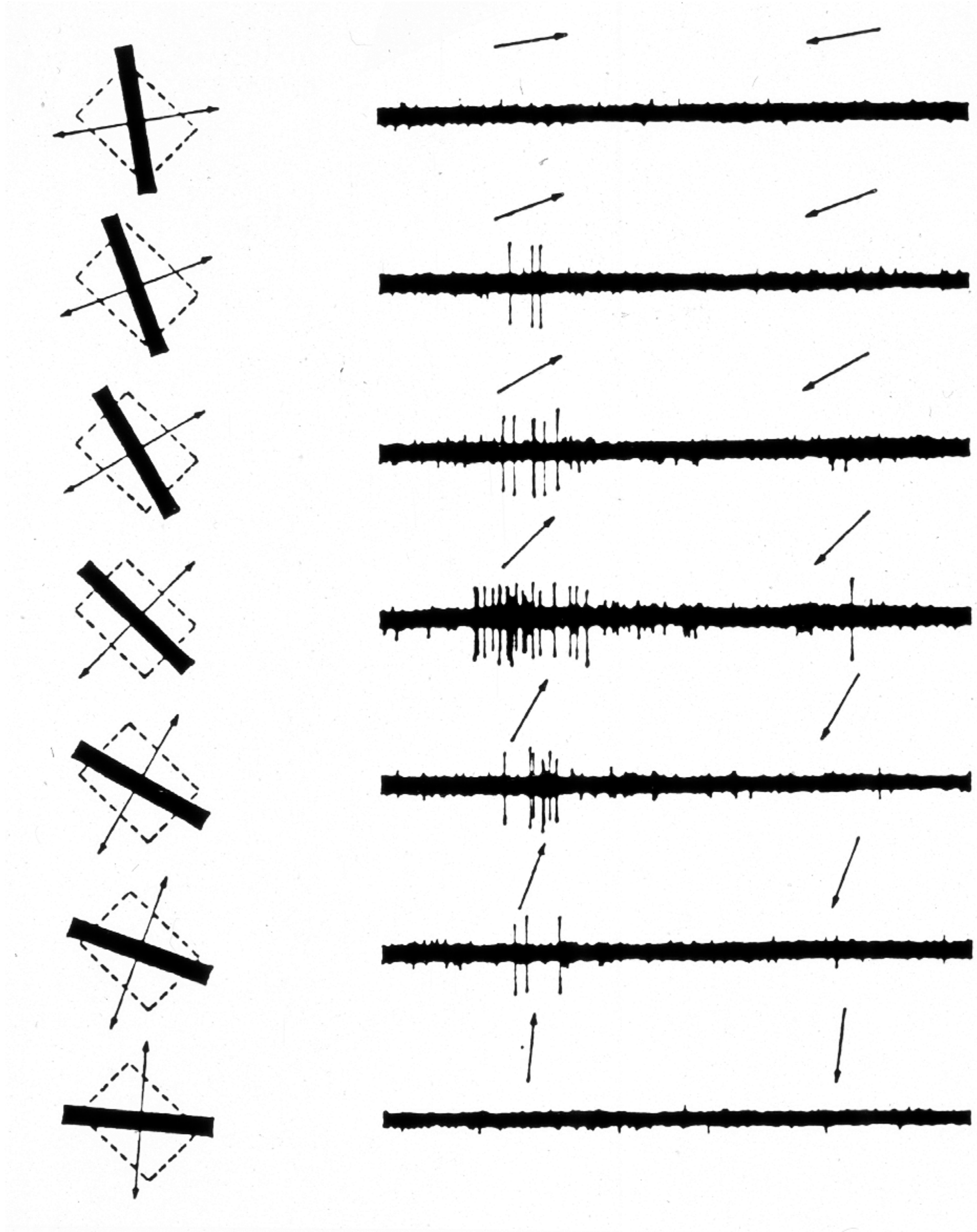




Neural code depends on multiple factors



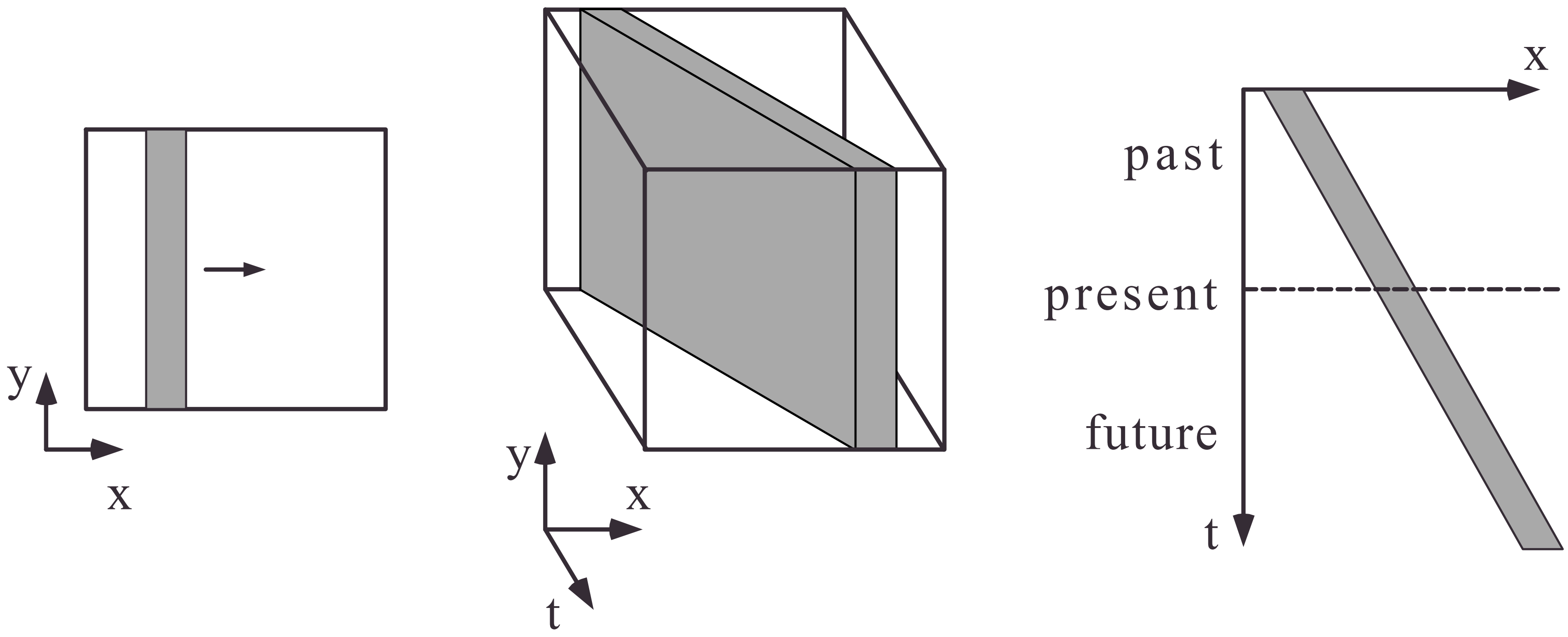
Direction selectivity



Hubel & Wiesel (1968)

Direction-selective complex cell

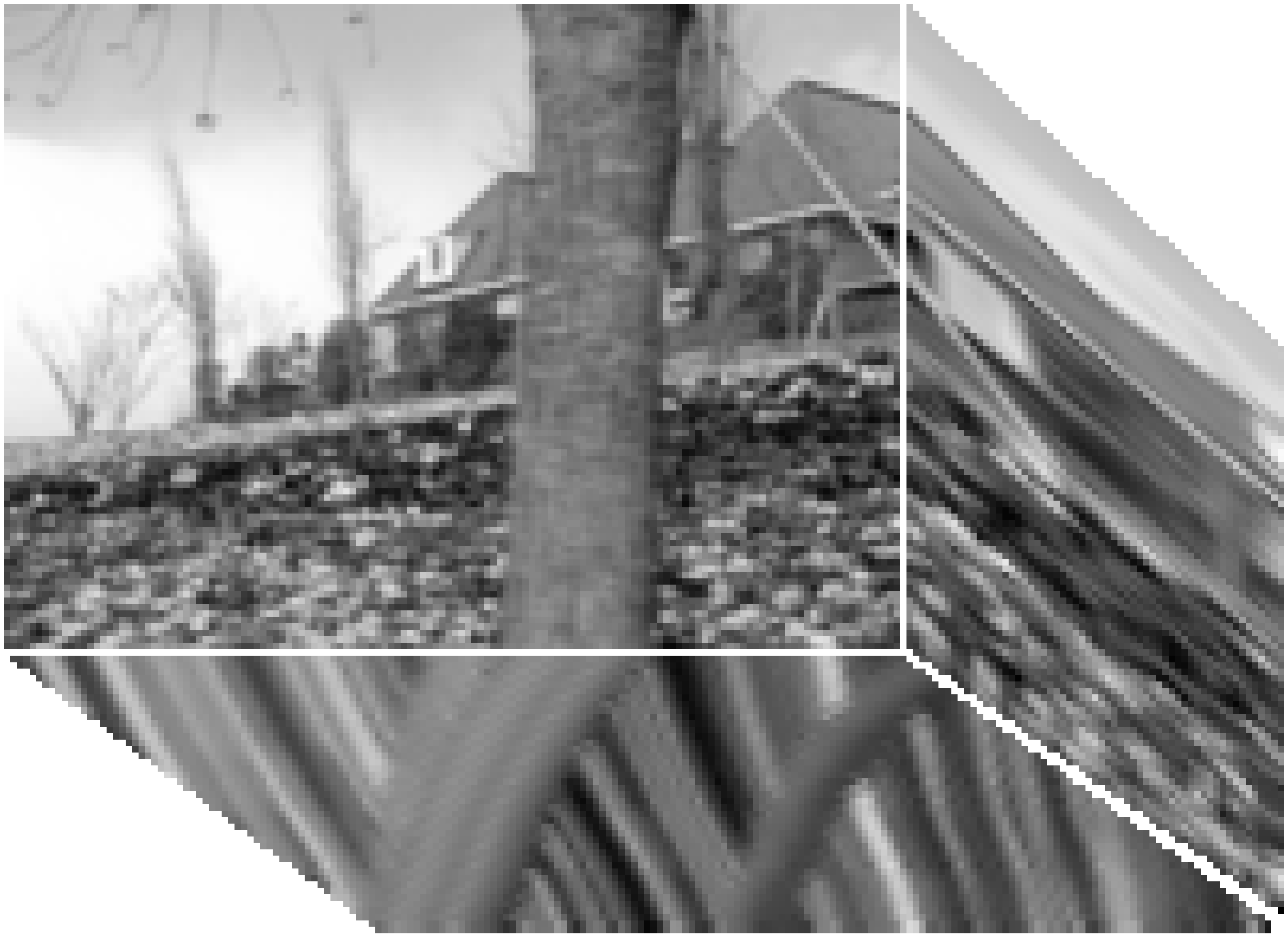
Orientation in space-time



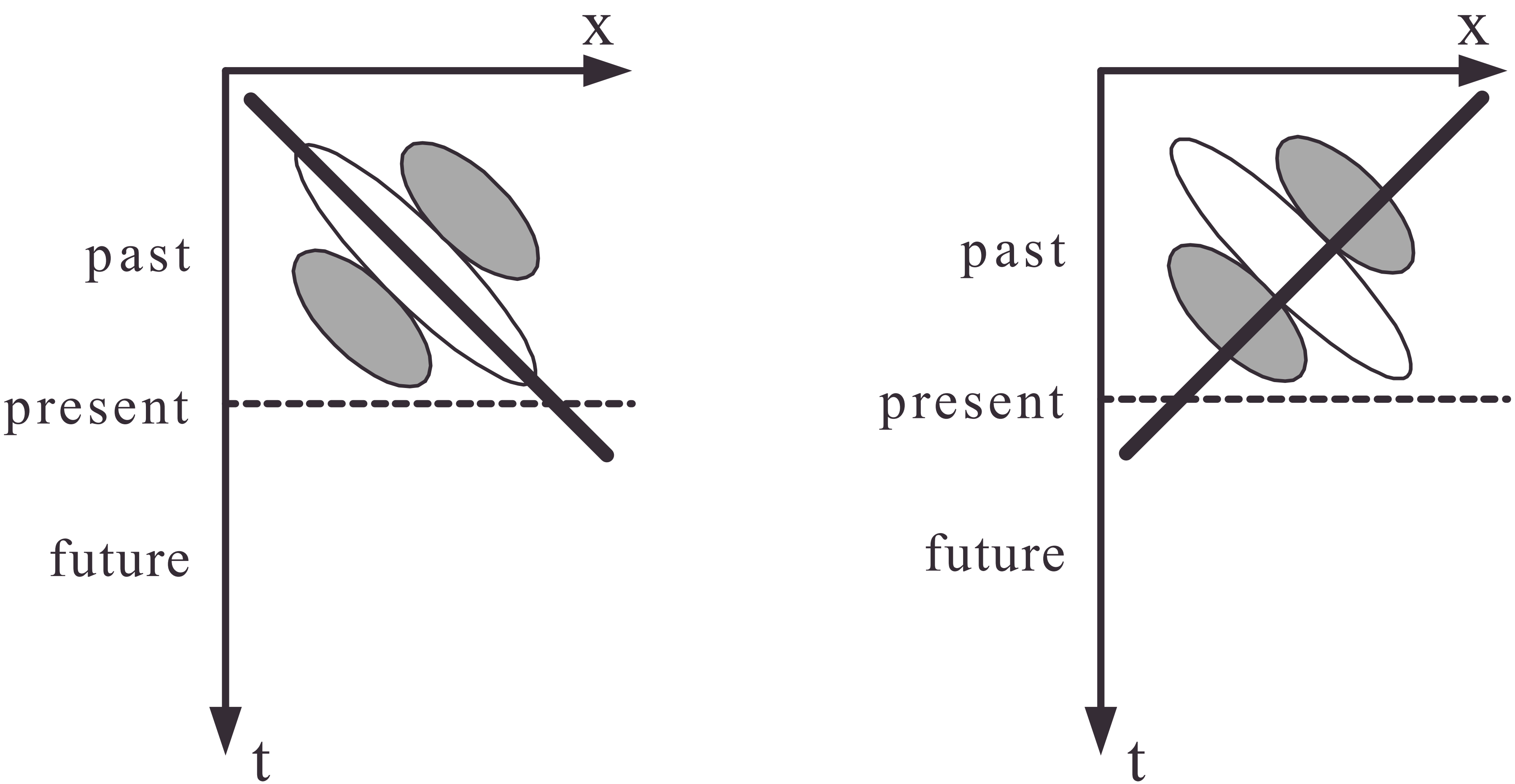
Motion is like orientation in space-time and spatiotemporally oriented filters can be used to detect and measure it.

Adelson & Bergen (1985)

Motion is orientation in space-time



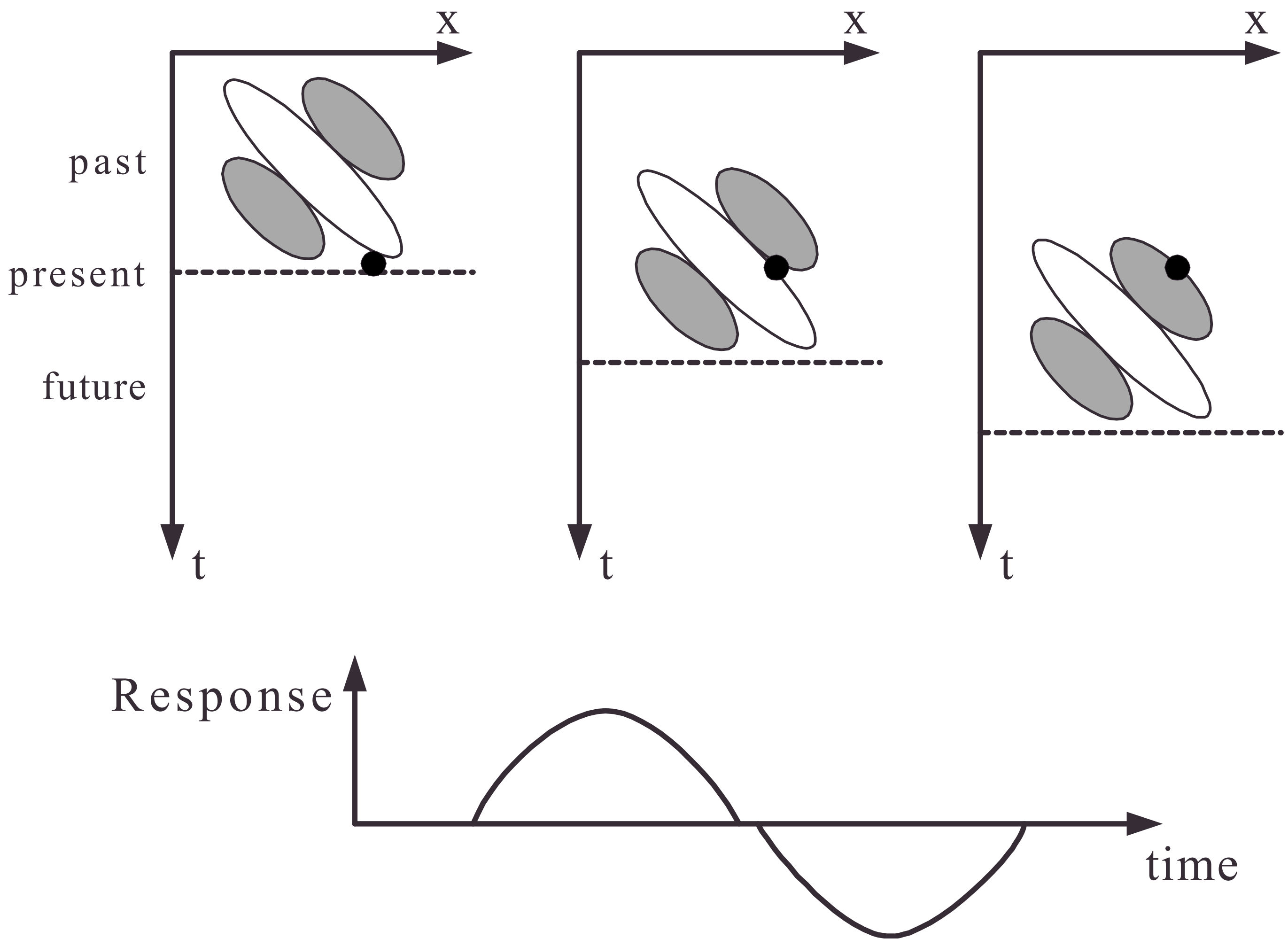
Direction selectivity model



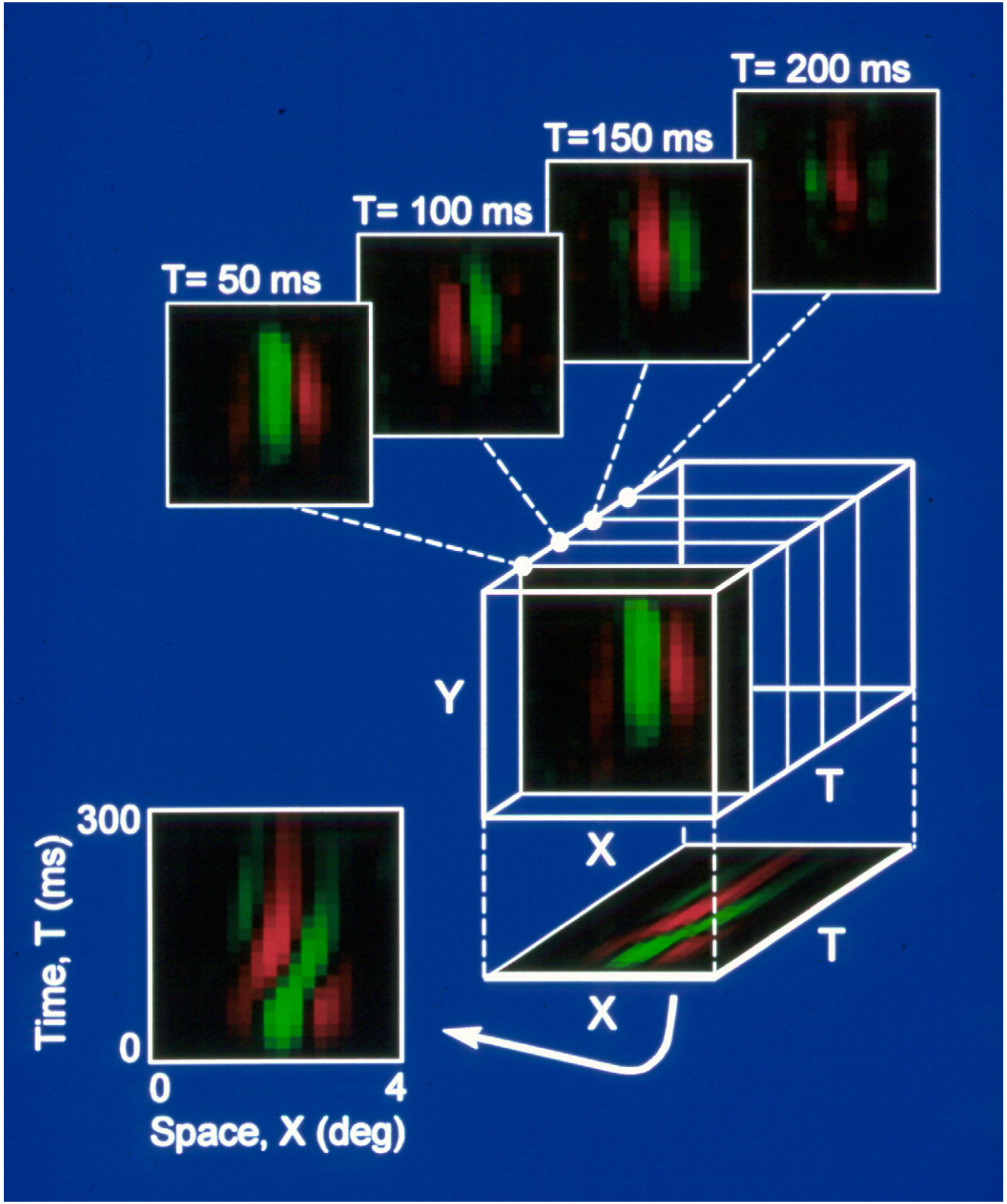
Strong response for motion in preferred direction.

Weak response for motion in non-preferred direction.

Impulse response



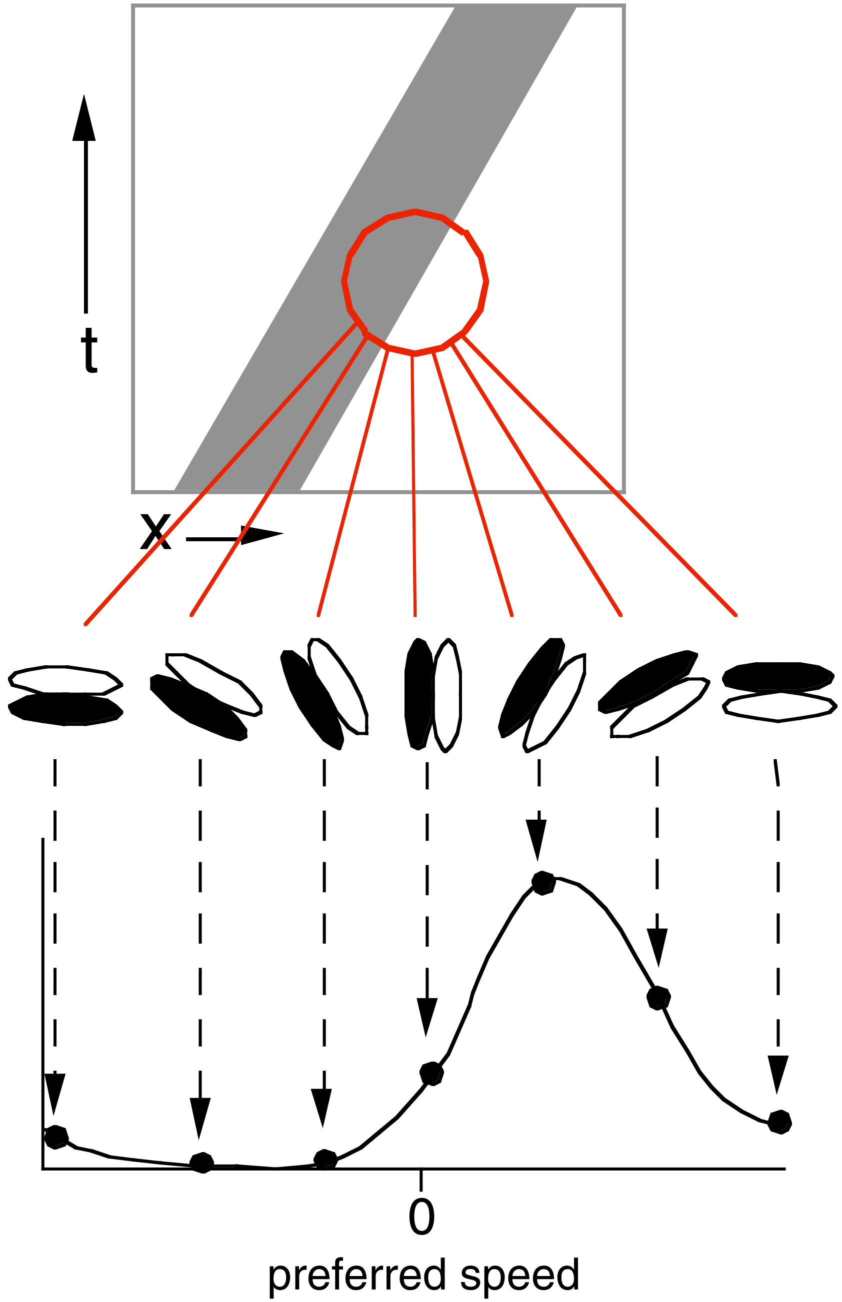
Space-time receptive field



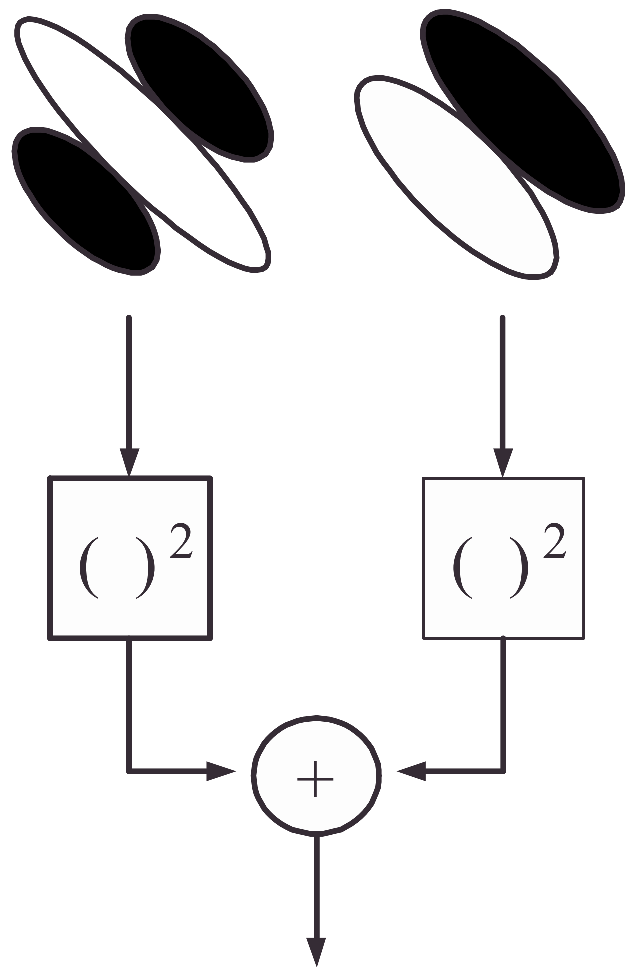
Ohzawa, DeAngelis, & Freeman (1995)

Distributed representation of speed

Each spatiotemporal filter computes something like a derivative of image intensity in space and/or time. "Perceived speed" is the orientation corresponding to the gradient in space-time (max response).

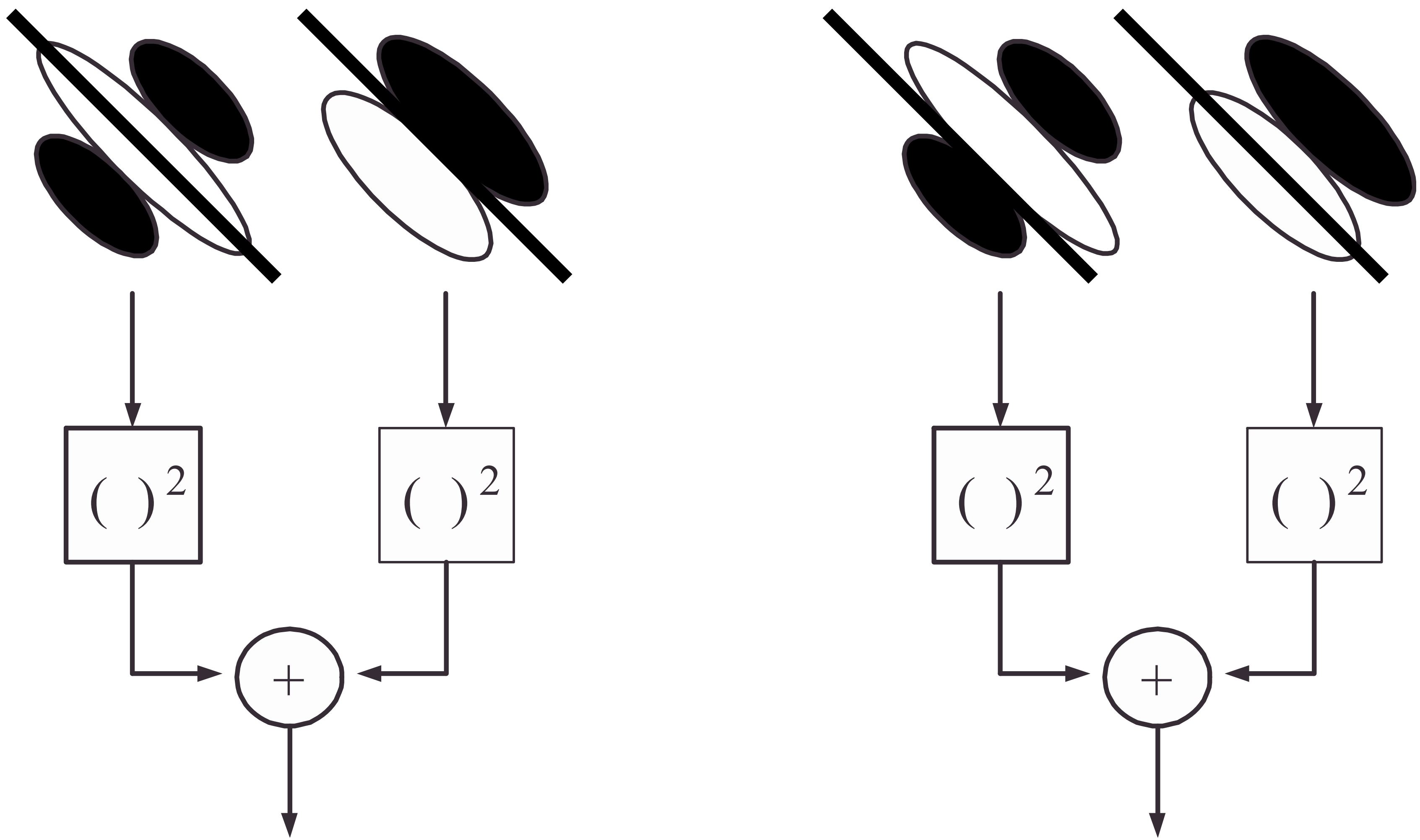


Complex cells: motion energy

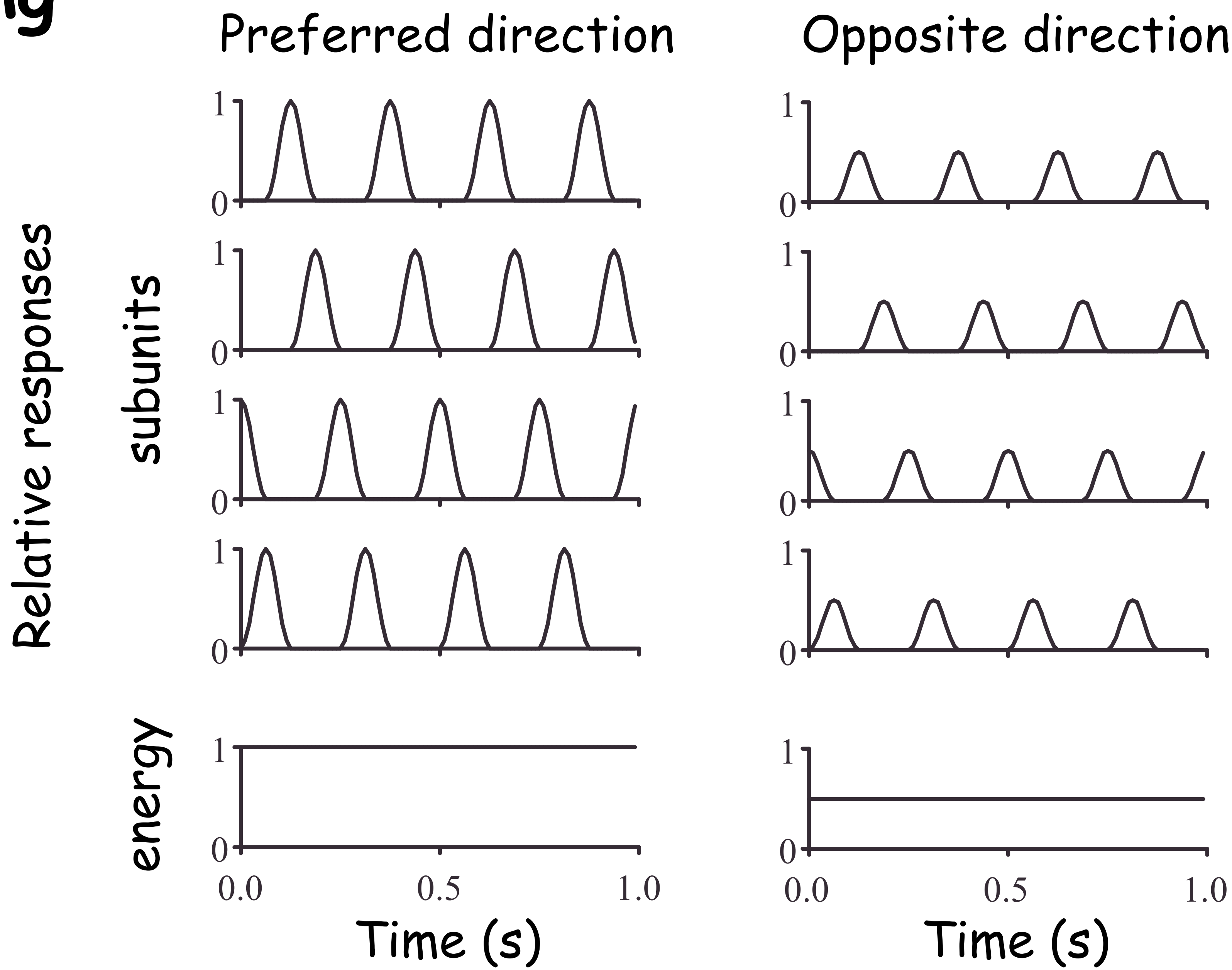


Motion energy & position invariance

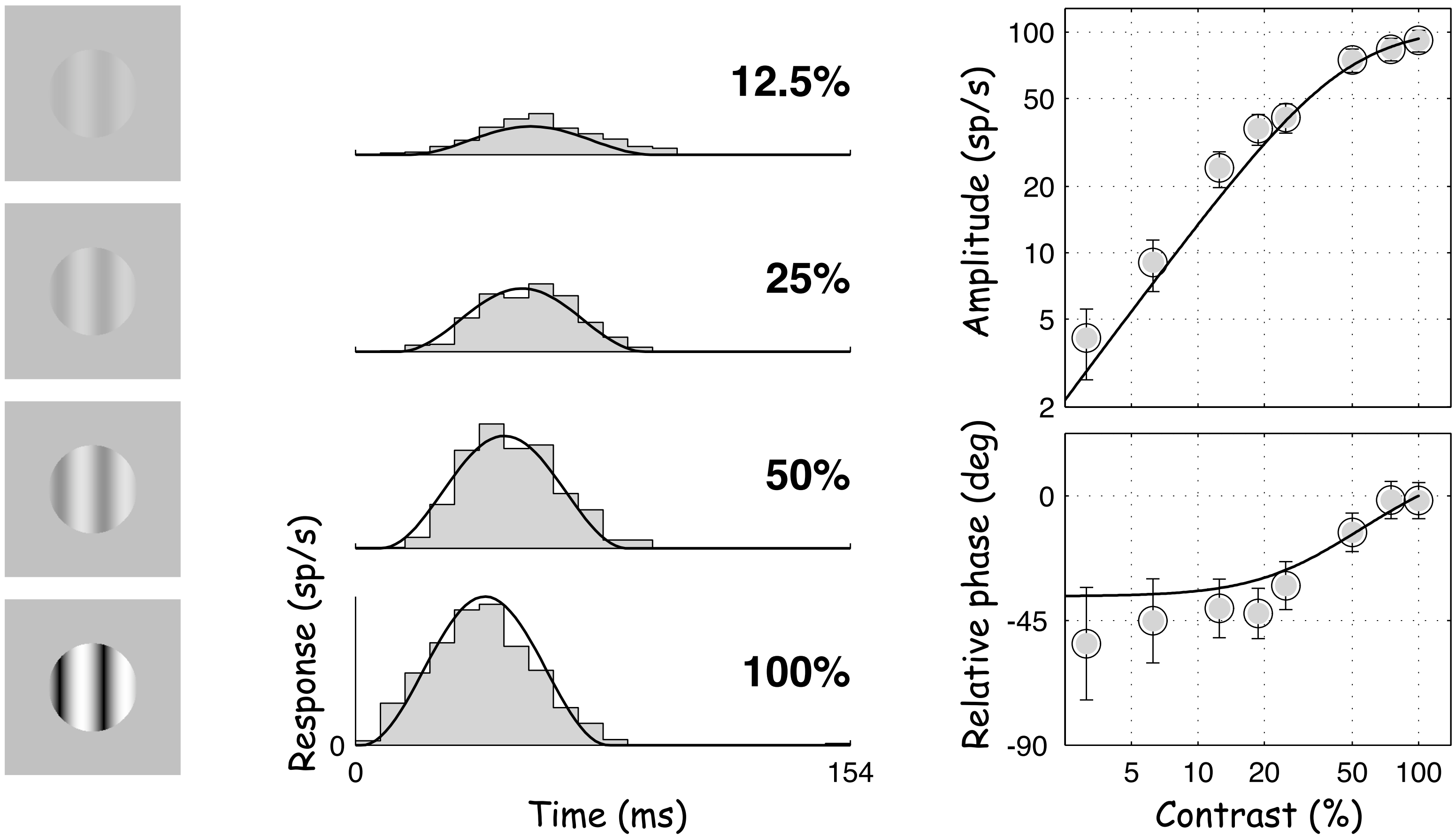
Moving stimulus as seen by both subunits at two different moments in time:



Motion energy responses to moving grating

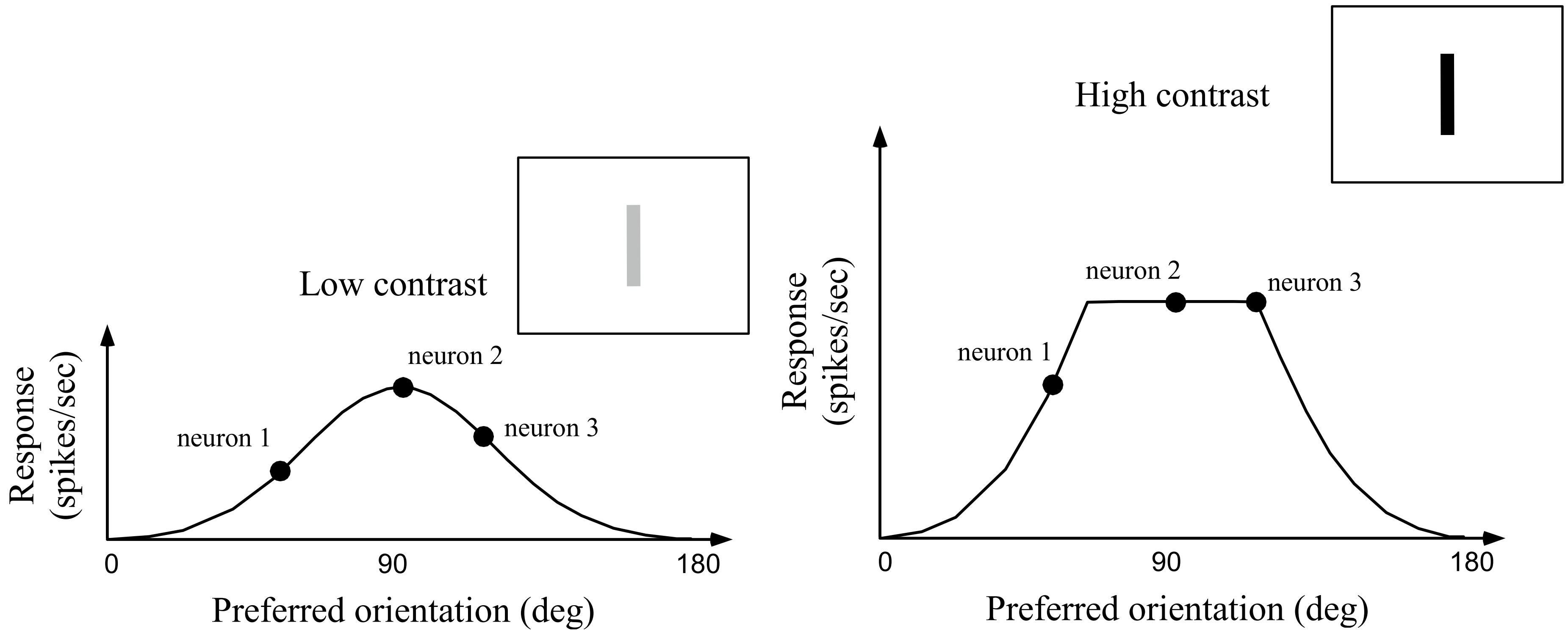


Response saturation and phase advance



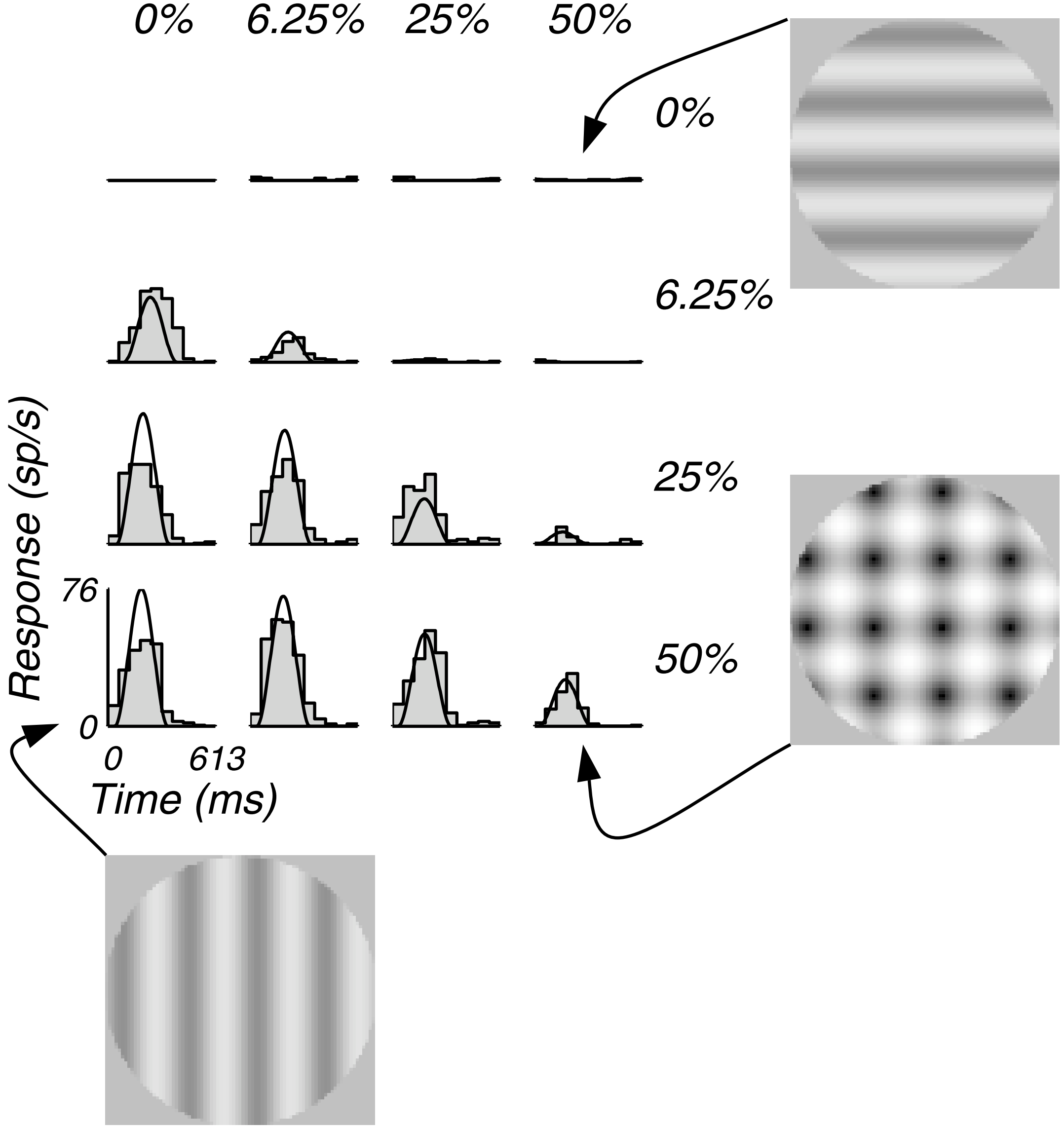
Carandini, Heeger & Movshon, J Neurosci, 1997

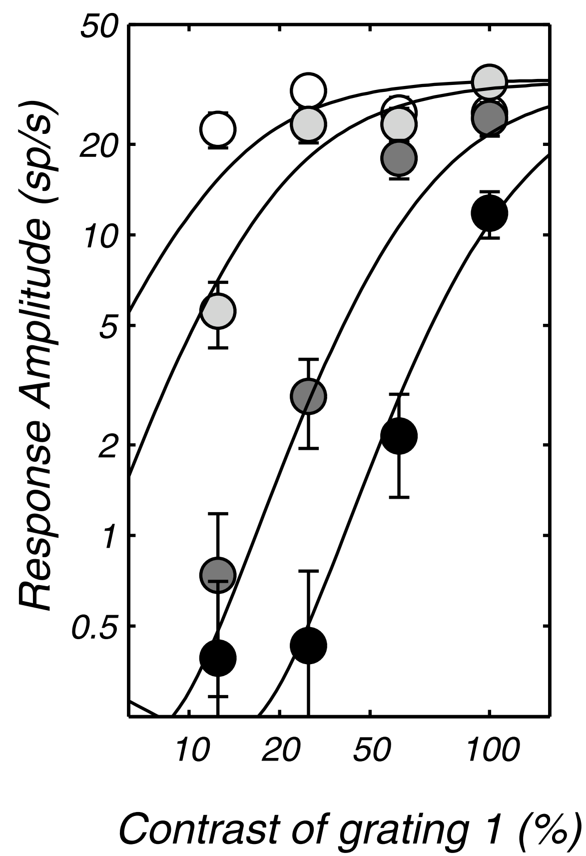
Failure of invariance with saturation?



Can no longer discriminate orientations near vertical

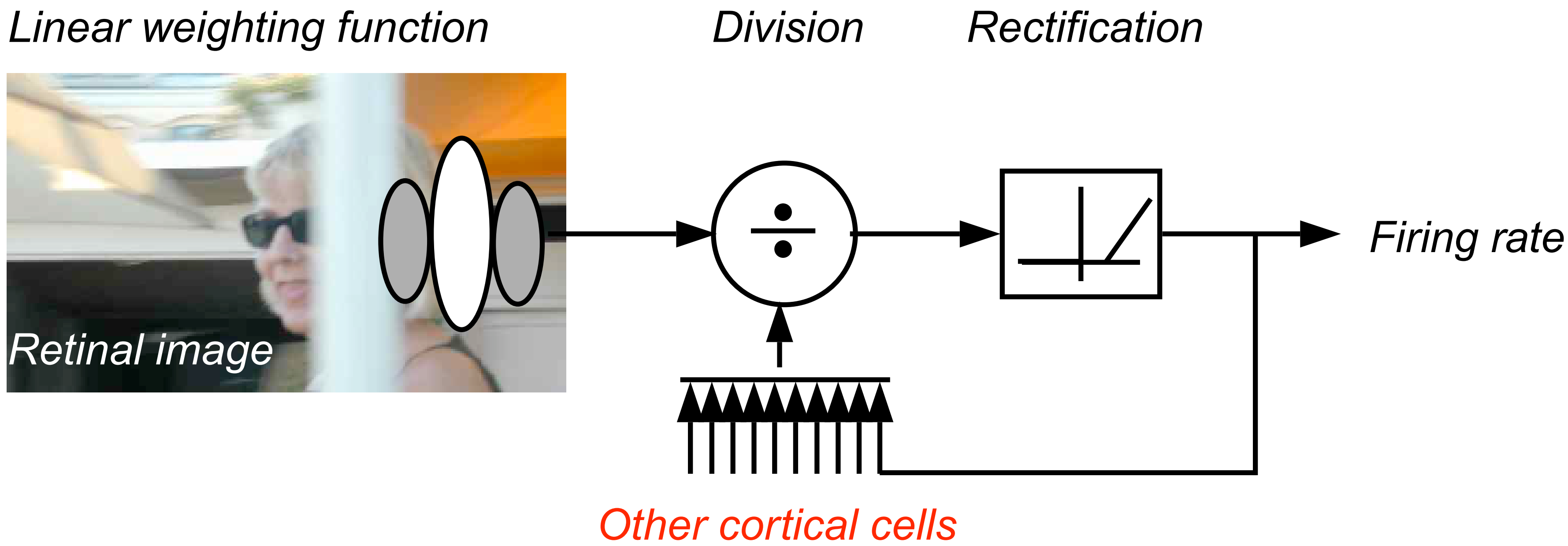
Masking





Carandini, Heeger & Movshon, J Neurosci, 1997

Normalization model

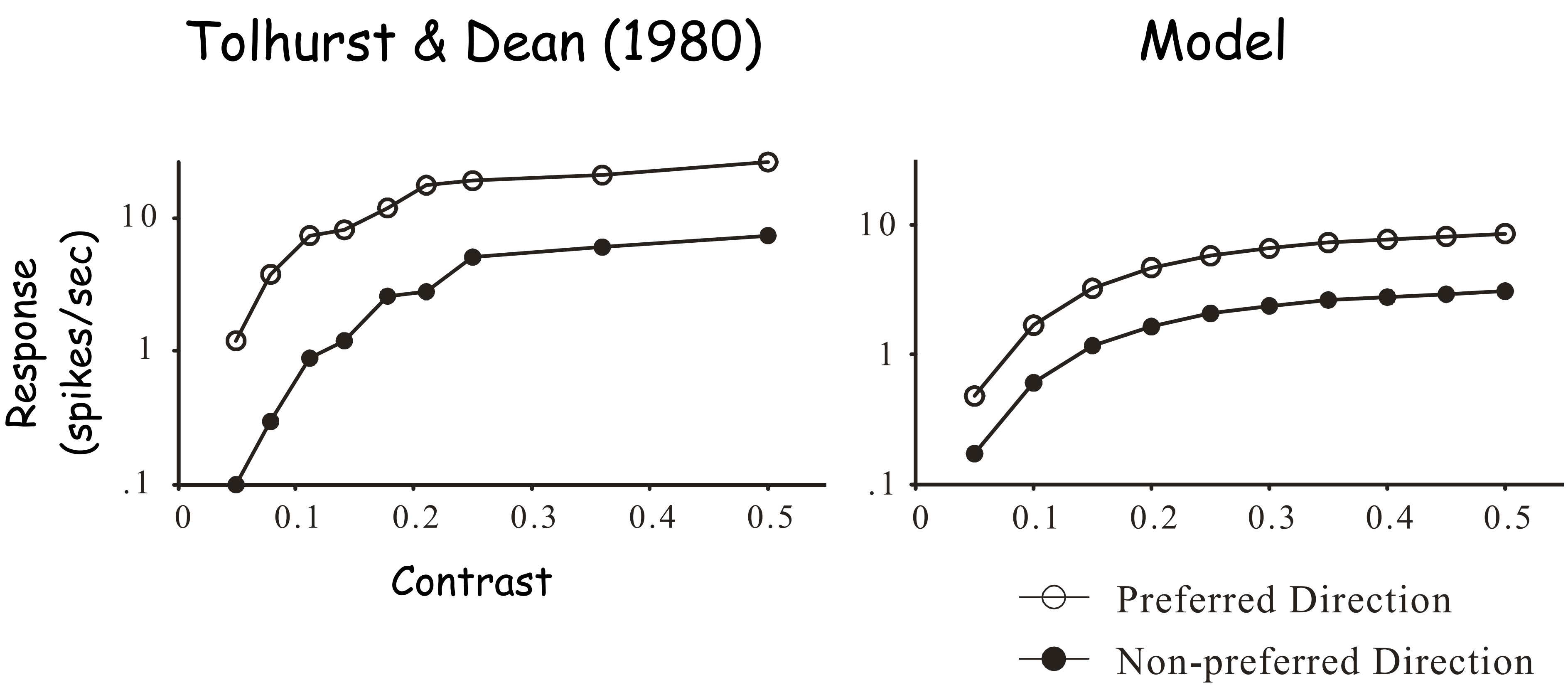


$$\text{normalized response} = \frac{\text{unnormalized response}}{\sum \text{unnormalized responses} + \sigma}$$

Heeger, Vis Neurosci, 1992

Contrast invariance

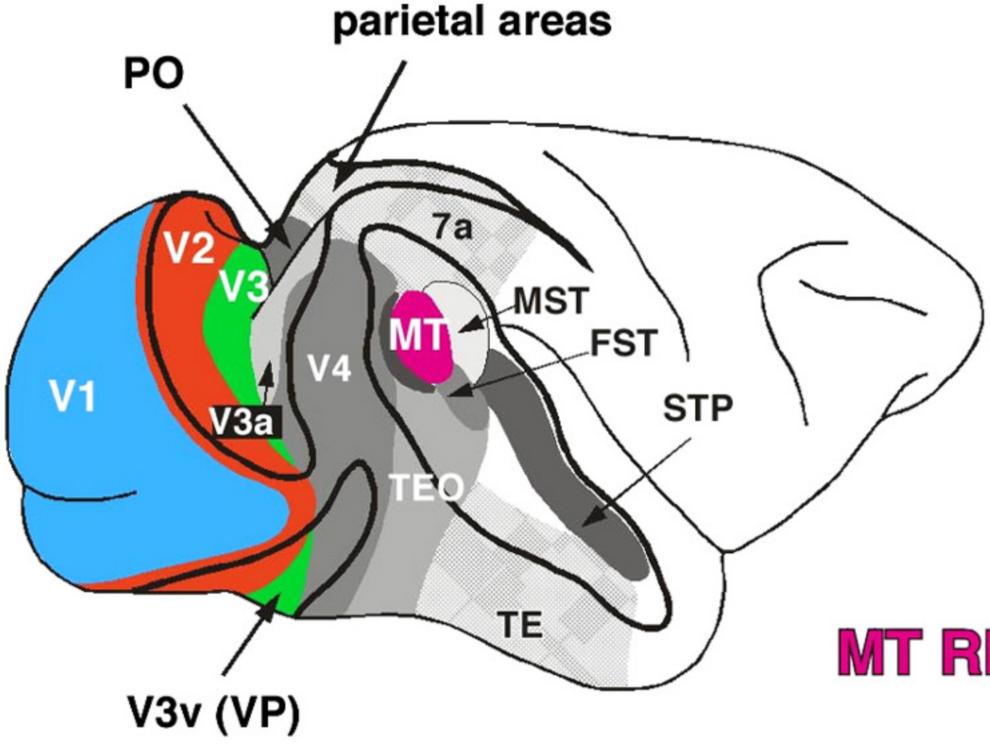
Ratio of responses to pref and non-pref directions constant over full range of contrasts.

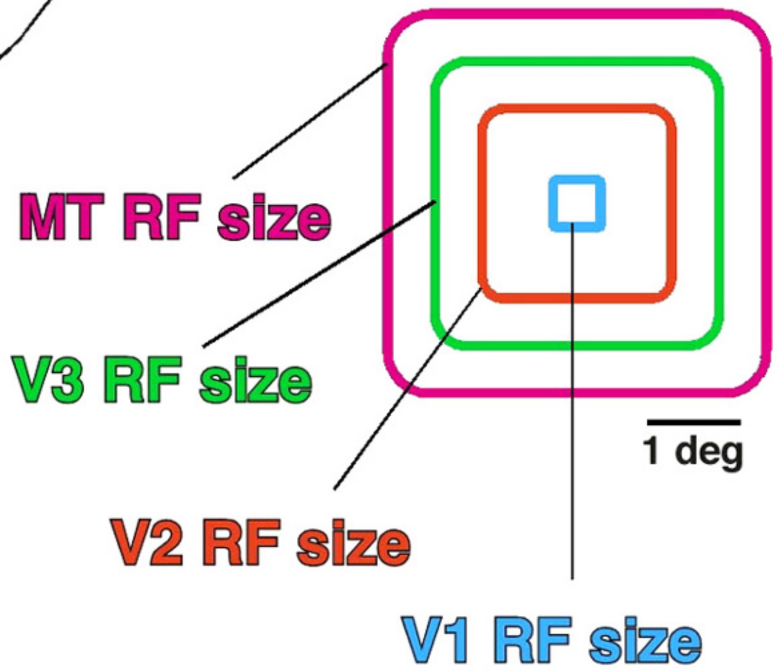


Tolhurst & Heeger, Vis Neurosci, 1997

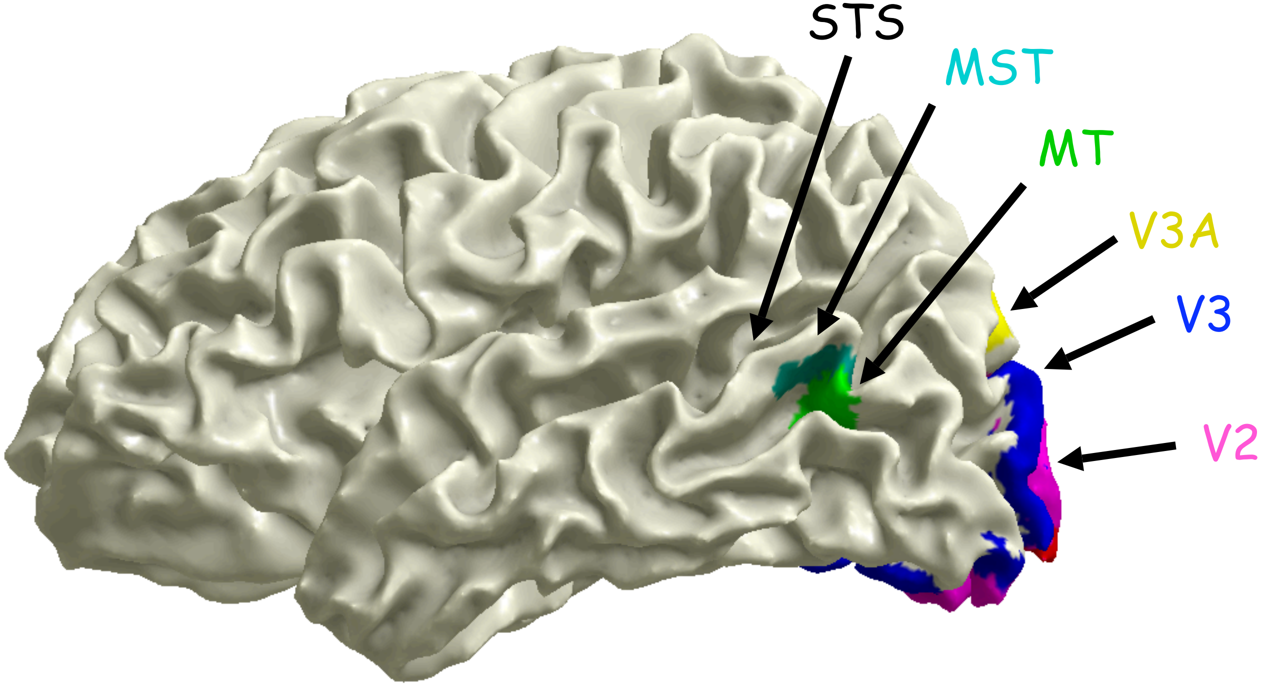
MT physiology and computational theory

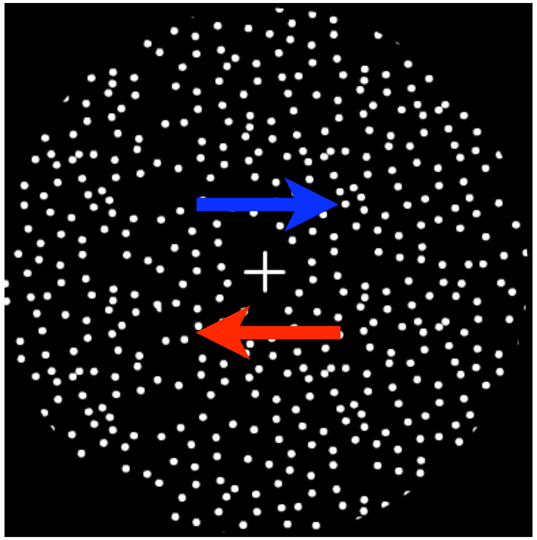
Increasing receptive field size





Increasingly complex selectivity



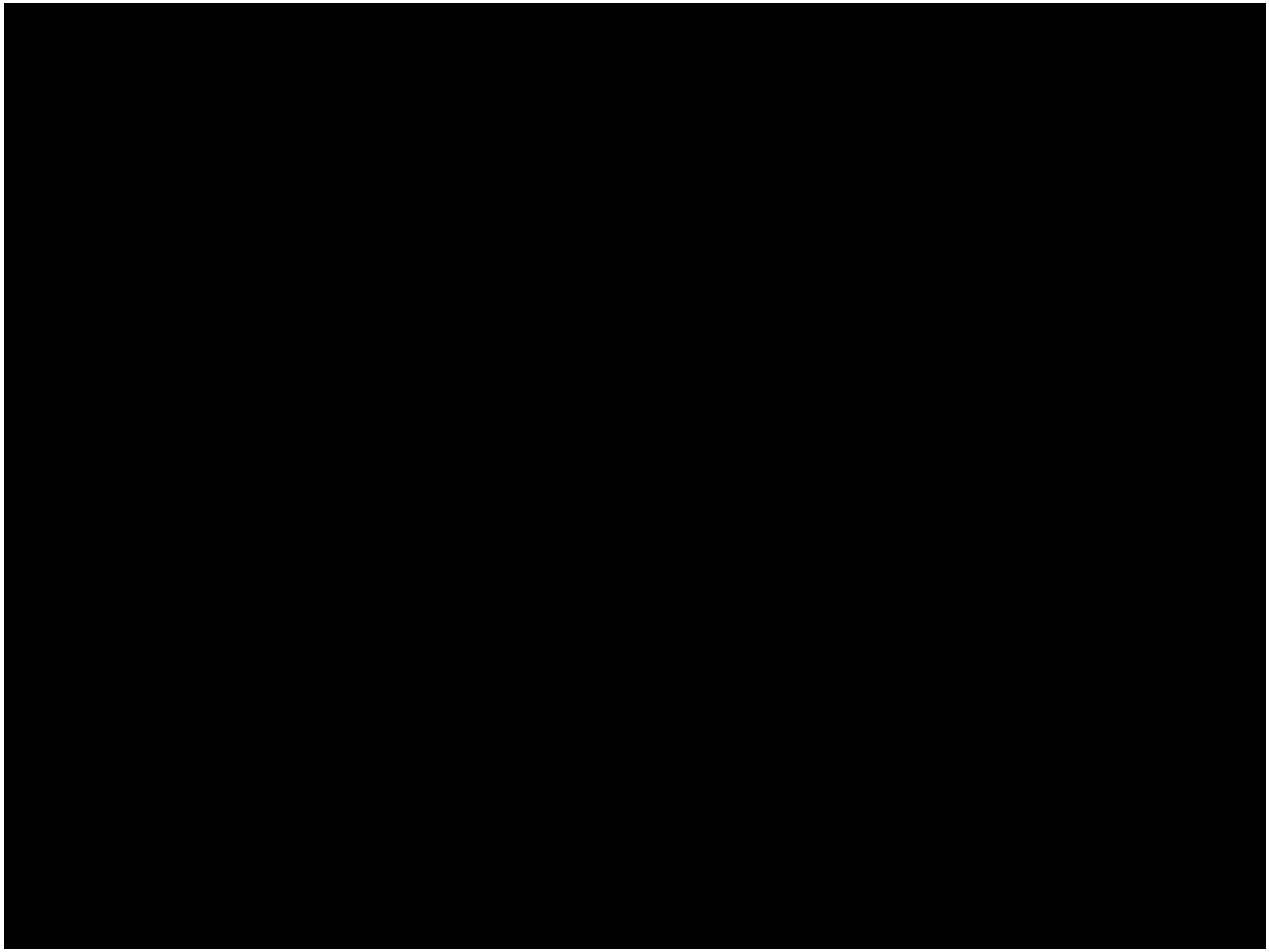


MT

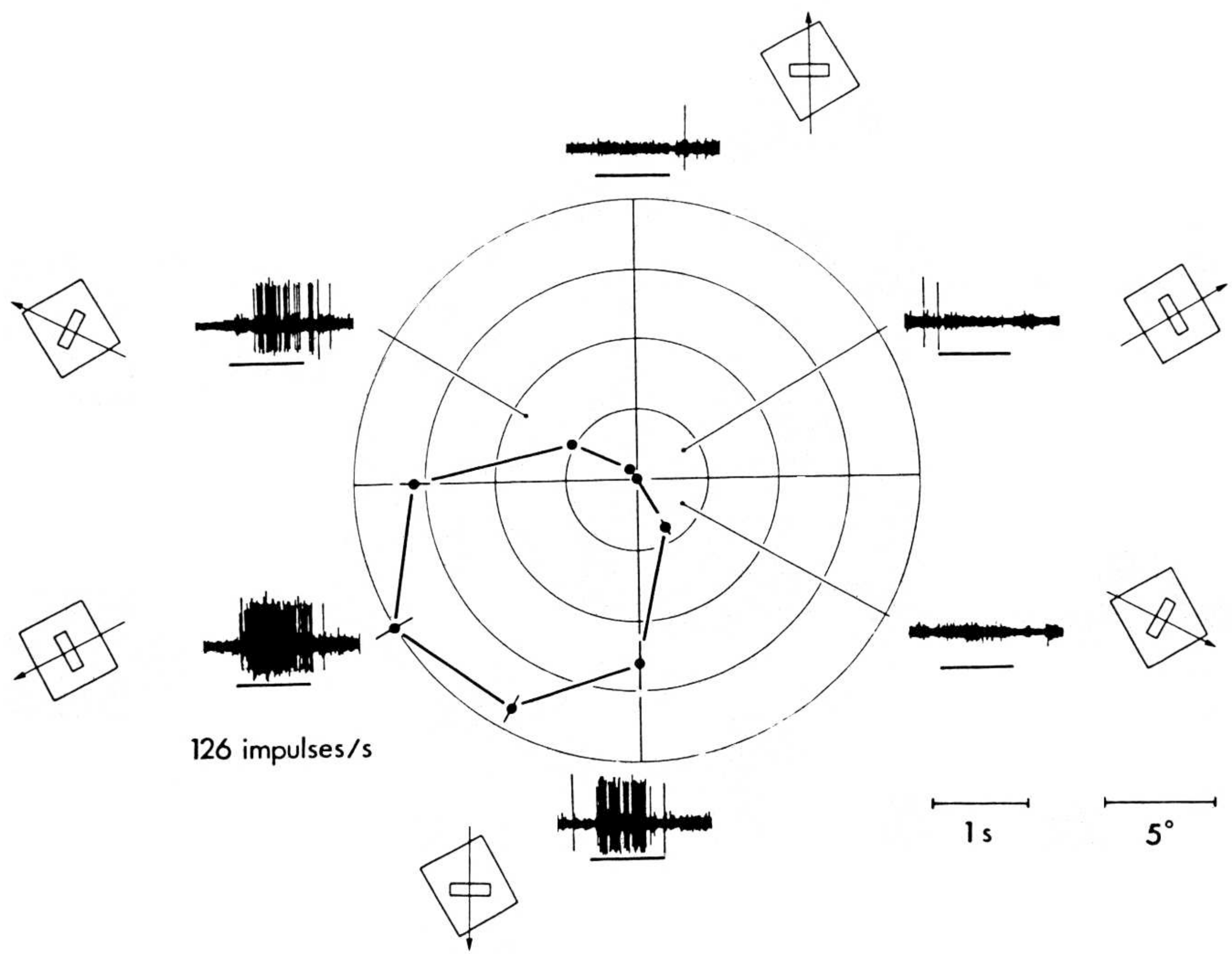
MST

STS

Neurons in MT are selective for motion direction

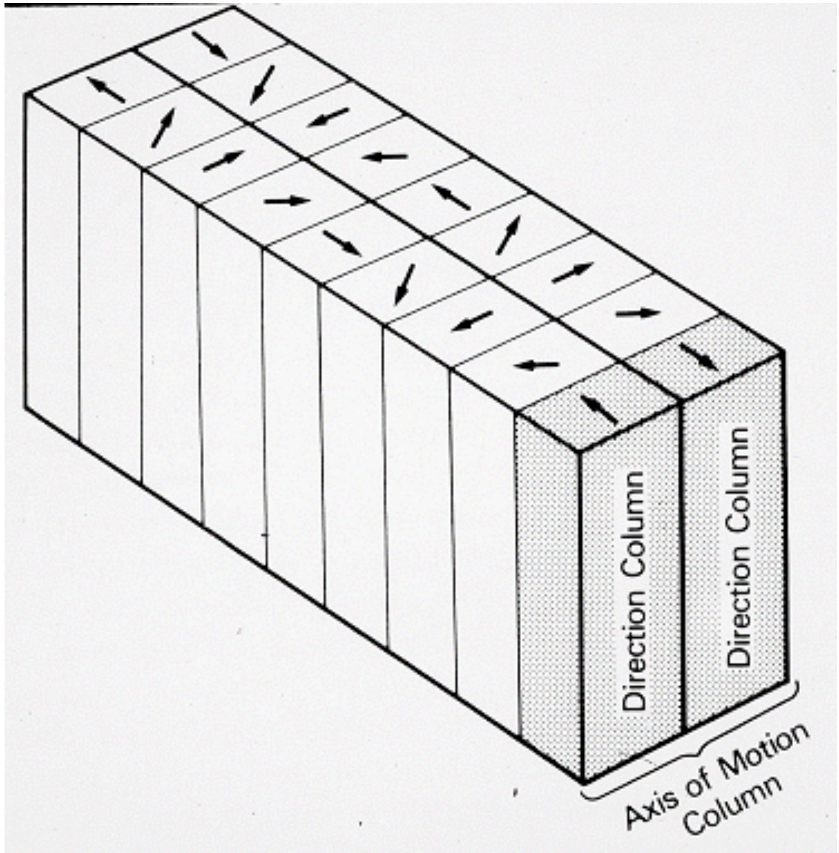


Neurons in MT are selective for motion direction



Maunsell and Van Essen, 1983

Columnar architecture in MT

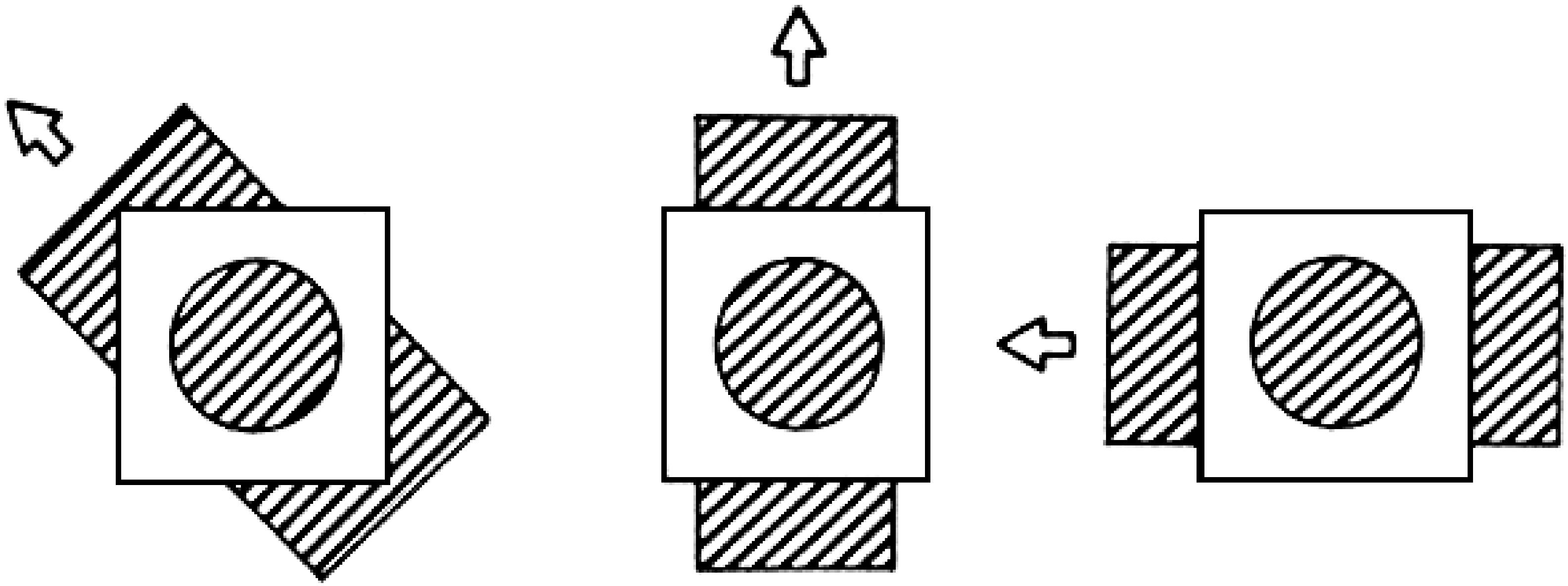


Direction columns in MT

Albright, Desimone & Gross,
J Neurophysiol (1984)

The “aperture problem”

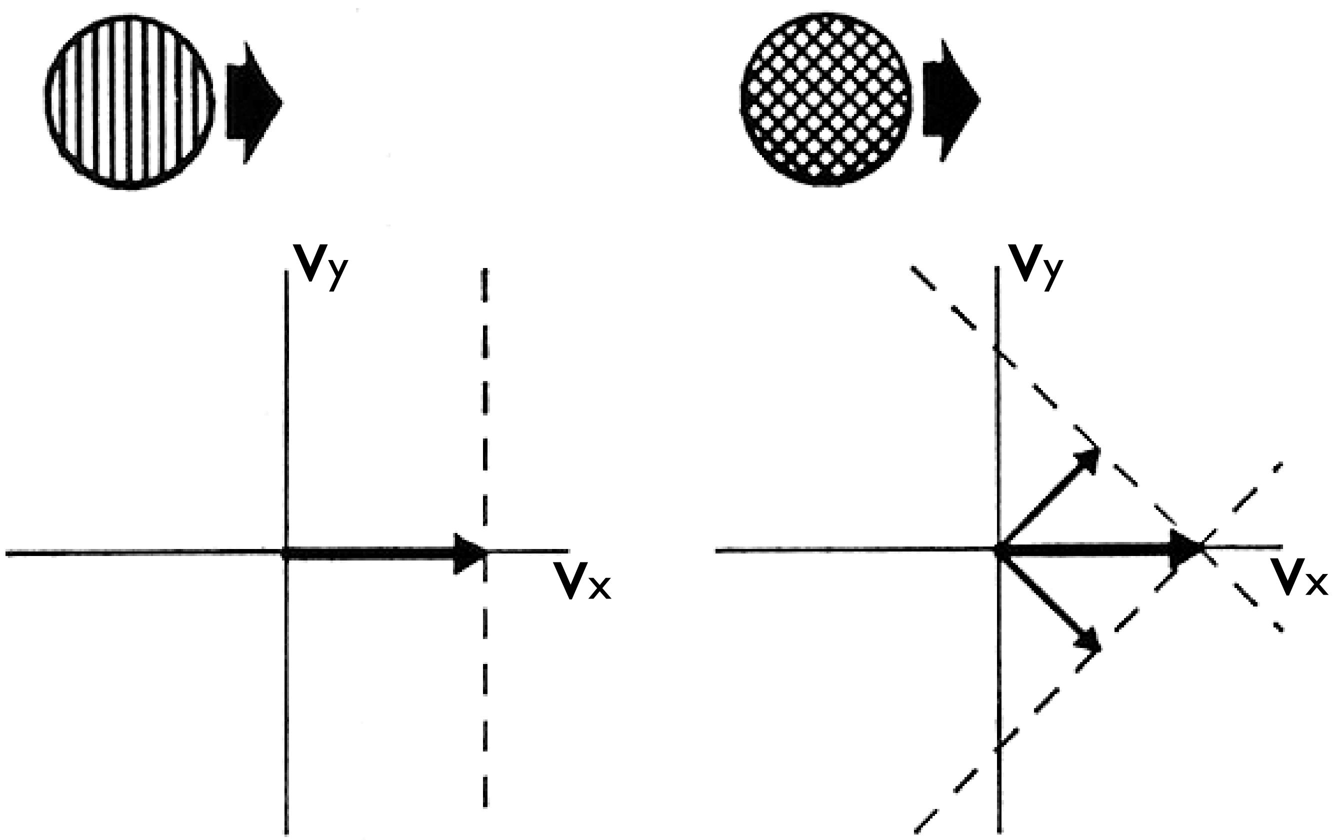
These three motions are different but look the same when viewed through a small aperture (i.e., that of a direction-selective receptive field).



Wallach (1935)

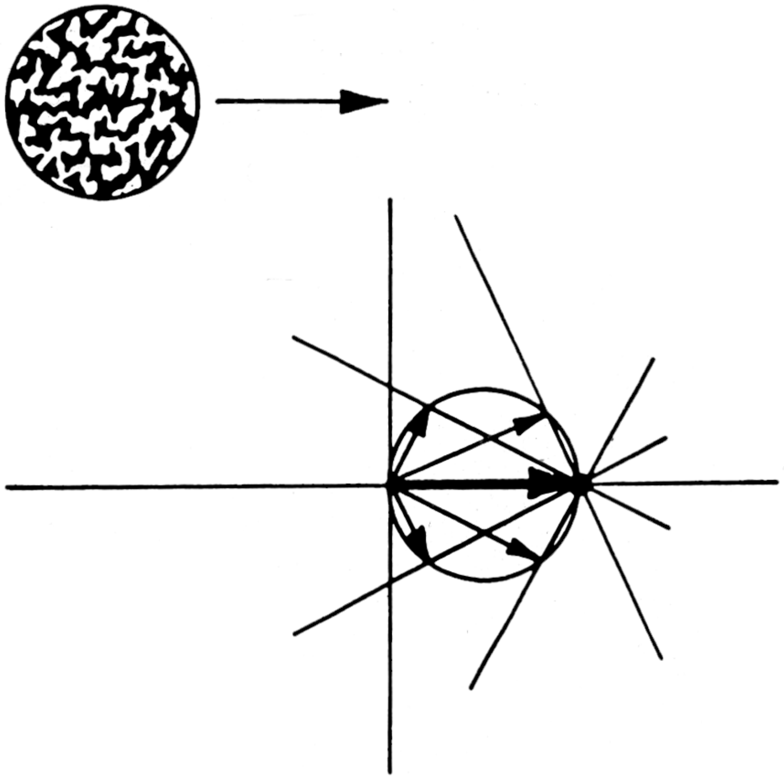
Intersection of constraints

With two different motion components within the aperture, there is a unique solution:



Adelson & Movshon (1981)

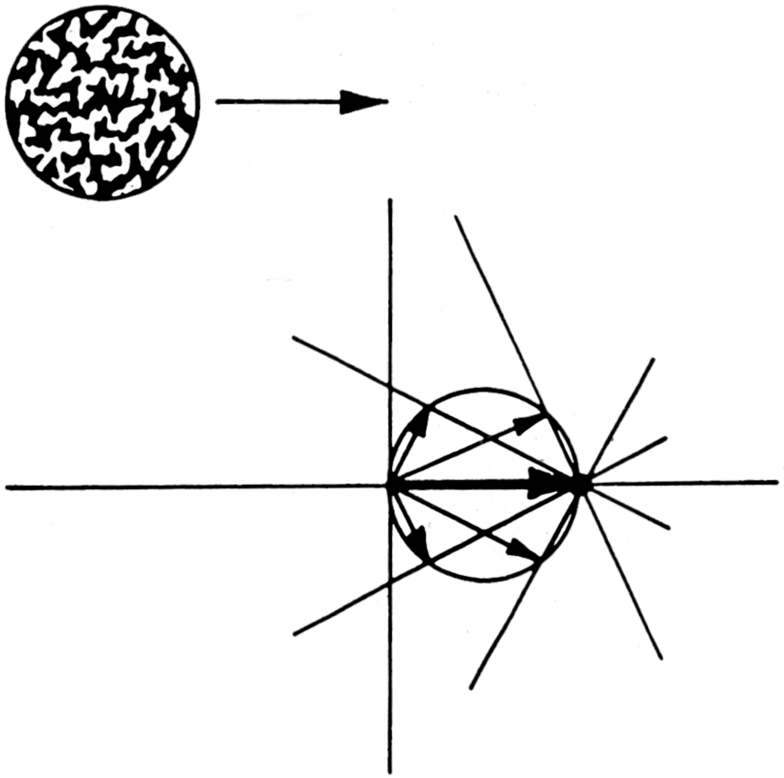
Intersection of constraints (many components)



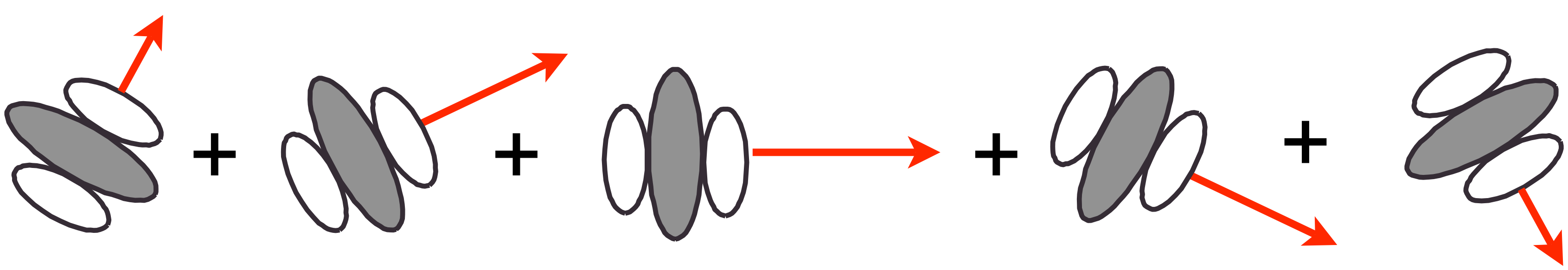
Each component activates a different V1 neuron, selective for a different orientation and speed.

How do you get selectivity for the moving pattern as a whole, not the individual components?

Neural implementation of IOC

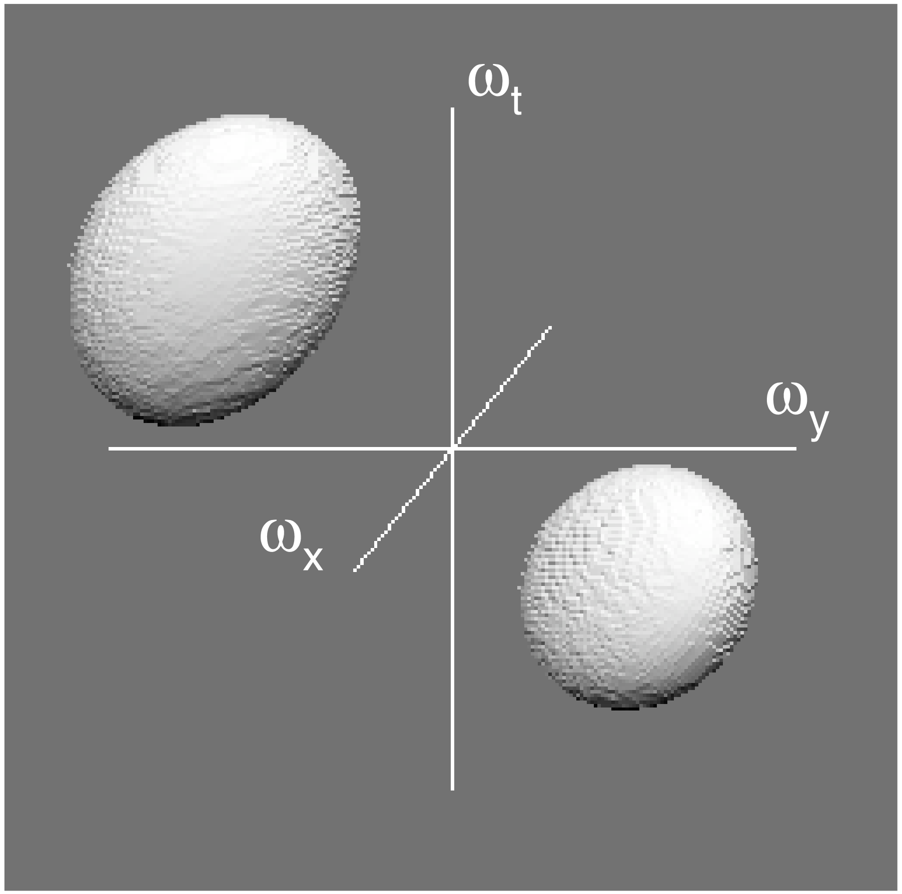


Answer: For each possible 2D velocity, add up the responses of those V1 neurons whose preferred orientation and speed is consistent with that 2D velocity.

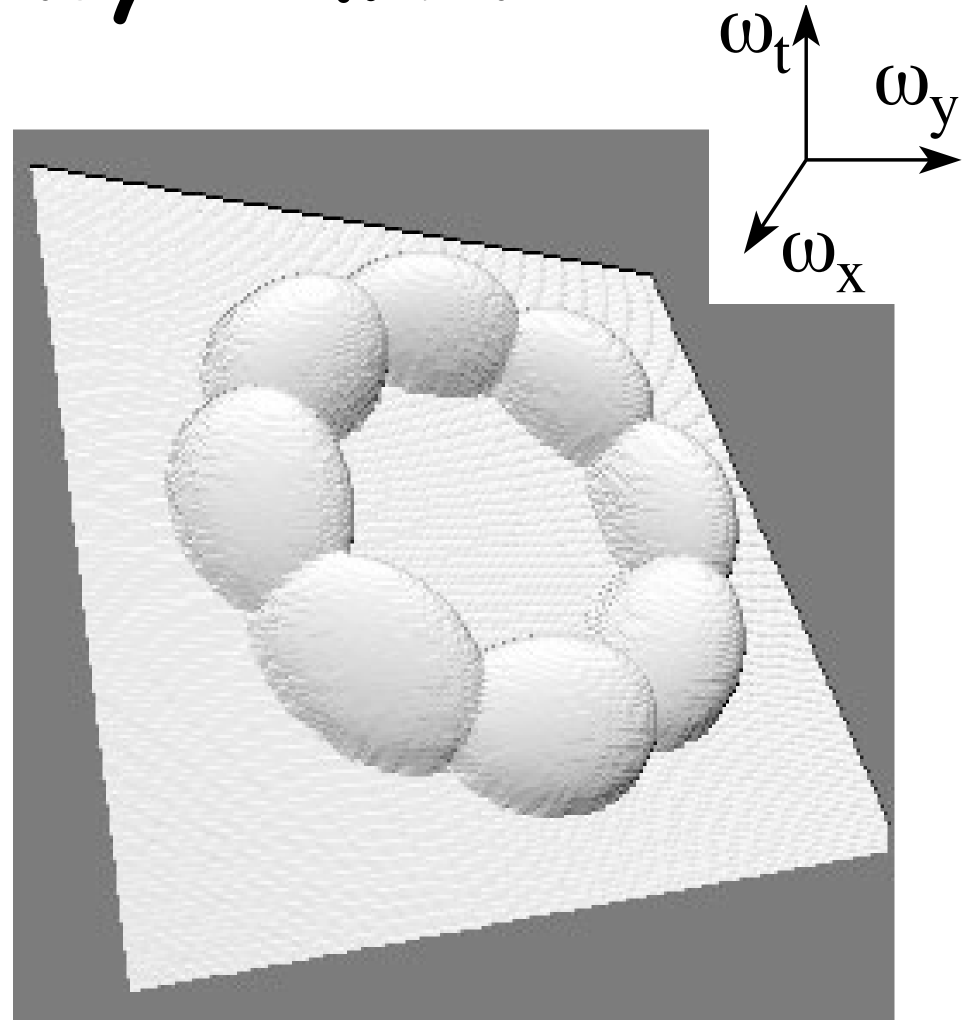


Simoncelli & Heeger, Vis Res, 1998

Spatiotemporal frequency domain



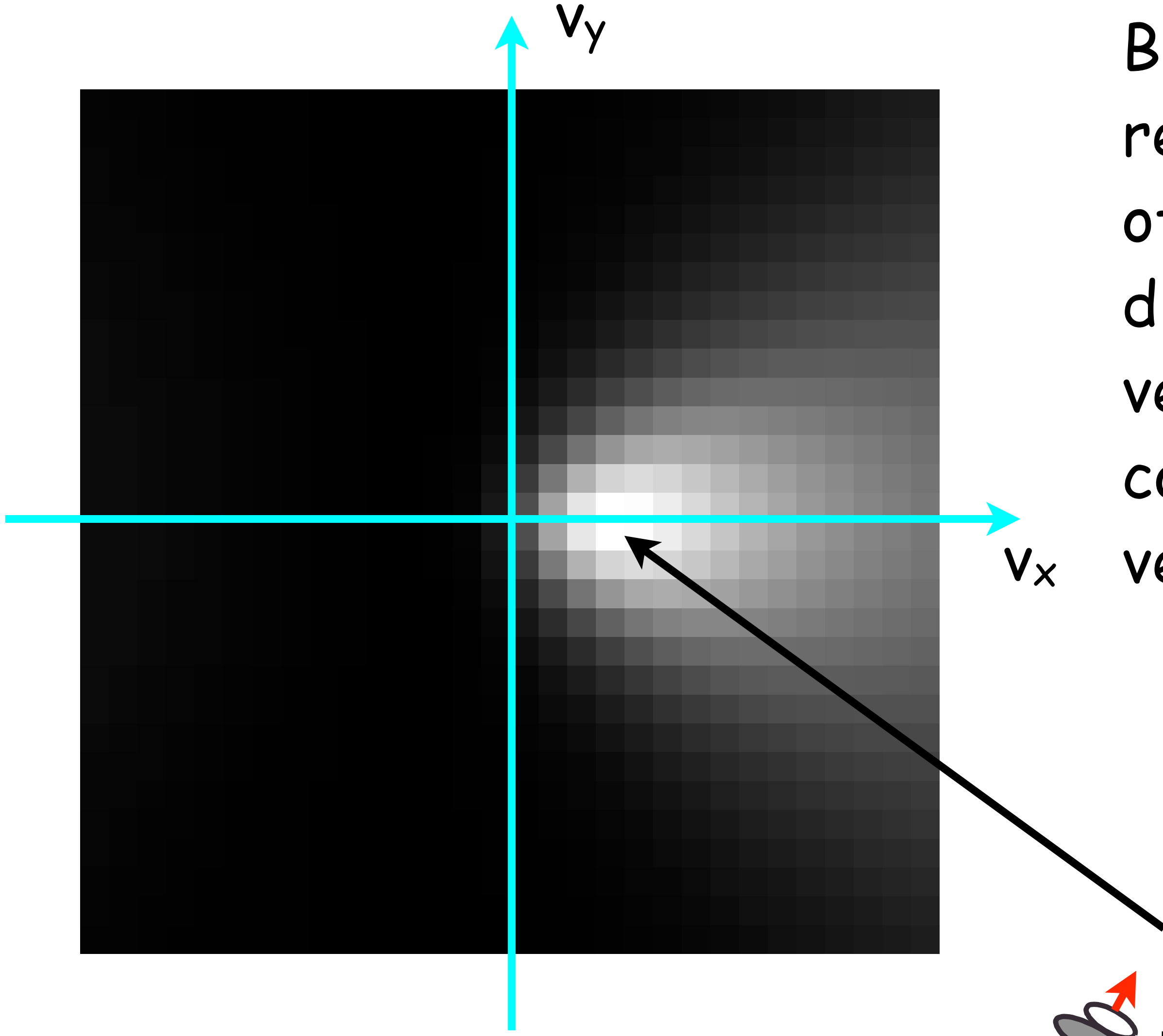
Spatiotemporal frequency response of space-time oriented linear filter.



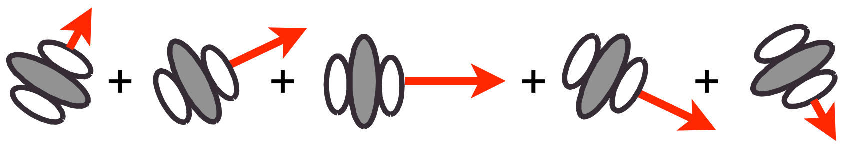
Frequency responses of filters that are all consistent with one velocity.

Simoncelli & Heeger, Vis Res, 1998

Distributed representation of 2D velocity



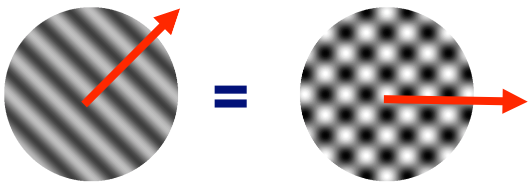
Brightness at each location represents the firing rate of a single MT neuron with a different preferred velocity. Location of peak corresponds to perceived velocity.



Simoncelli & Heeger, Vis Res, 1998

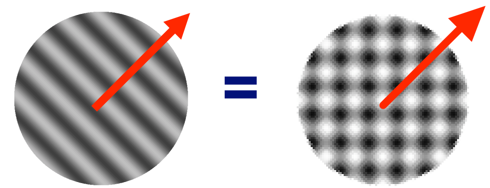
Component vs. pattern motion selectivity

component-motion cell

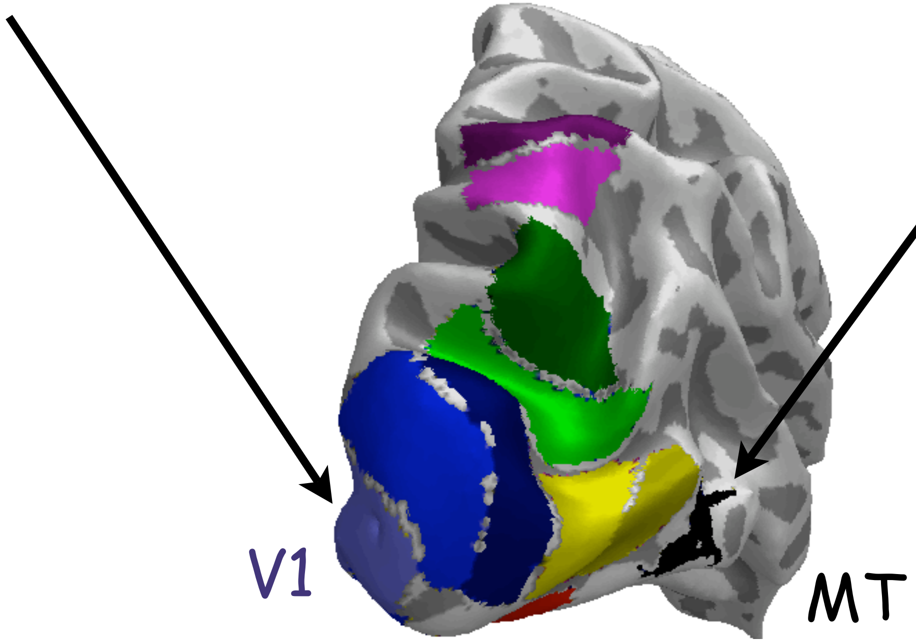


grating component moving
up-right => strong response

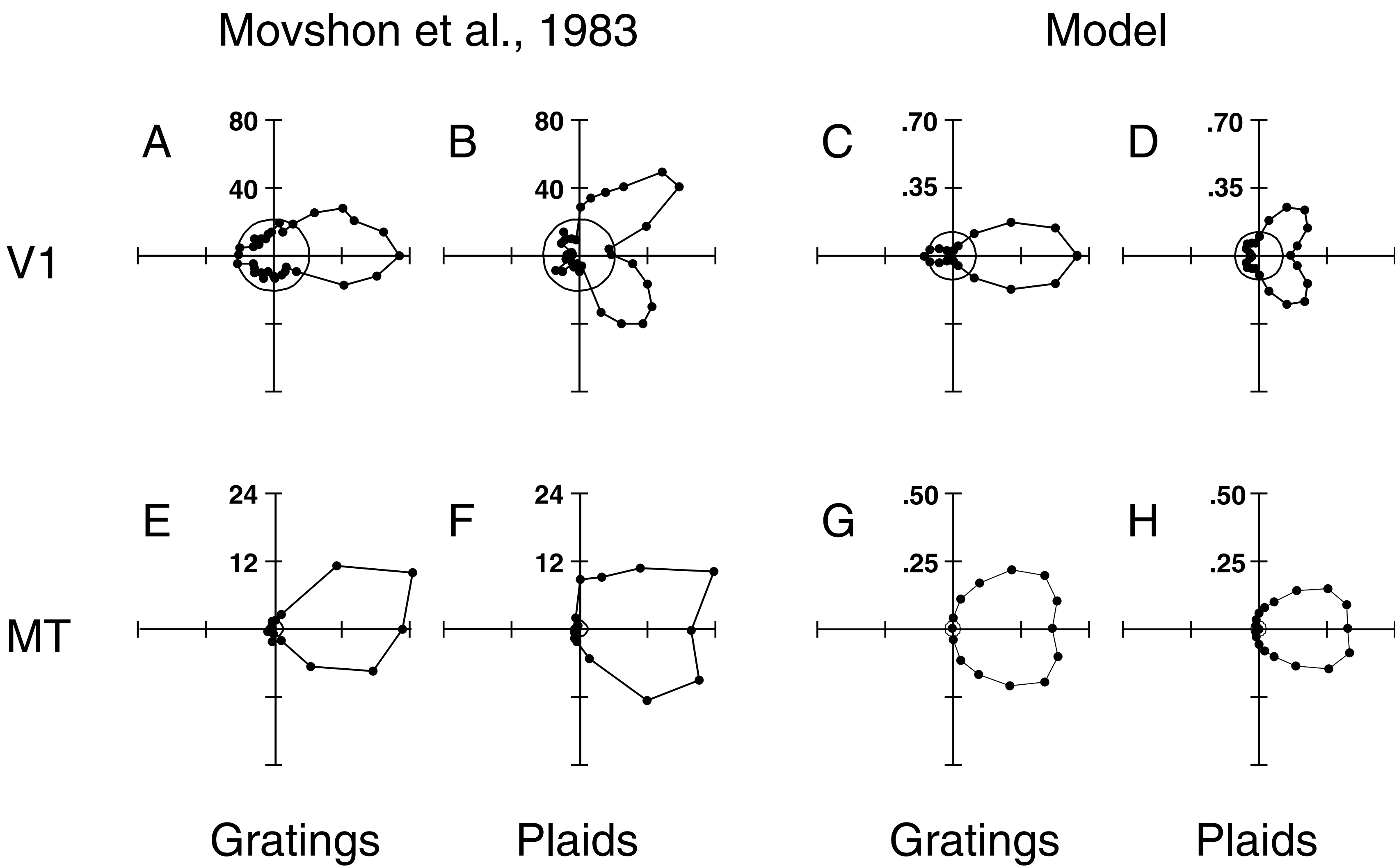
pattern-motion cell



pattern moving up-right
strong response

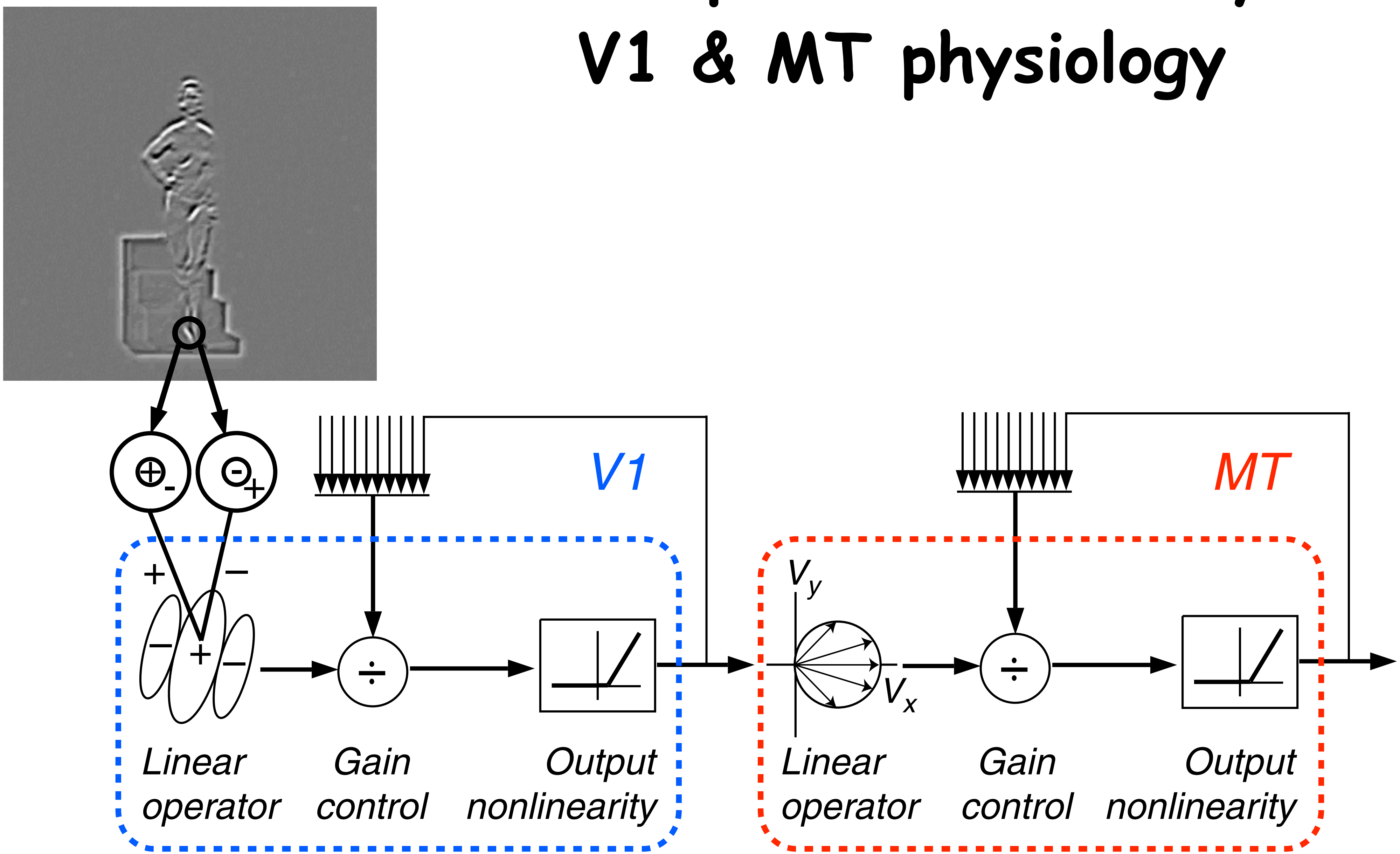


Component vs. pattern motion: single neurons

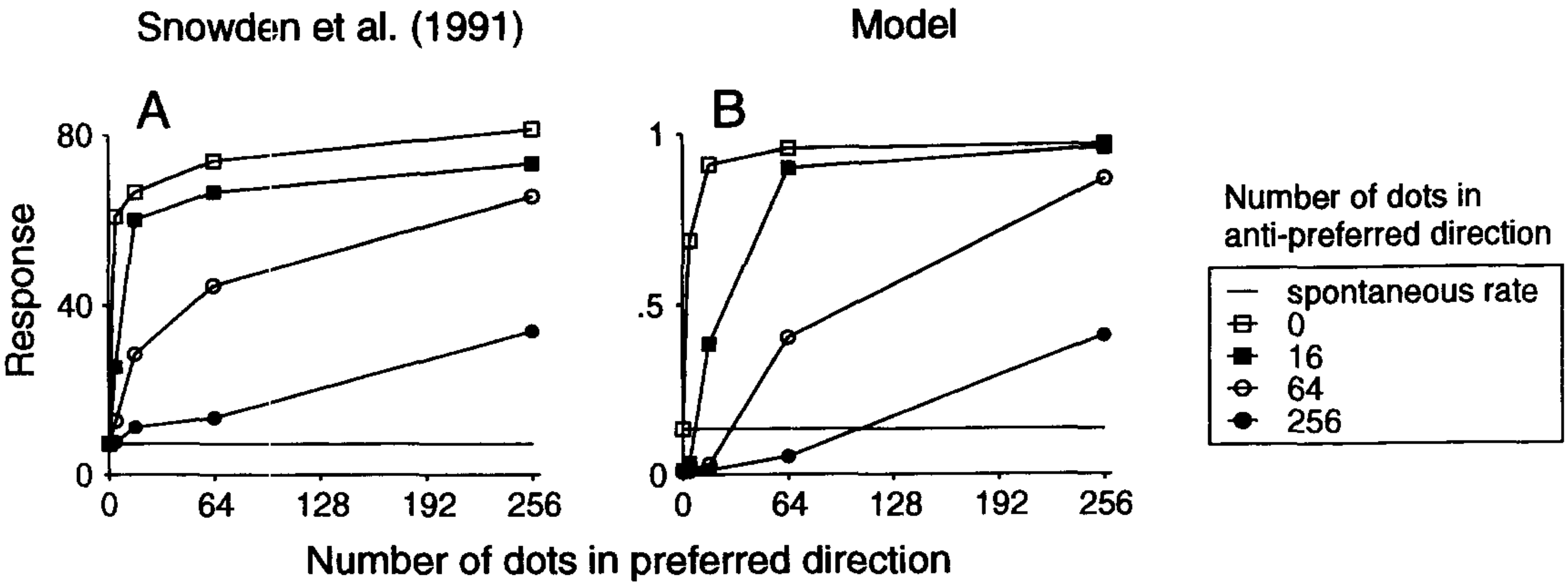


Simoncelli & Heeger, Vis Res, 1998

Computational theory of V1 & MT physiology

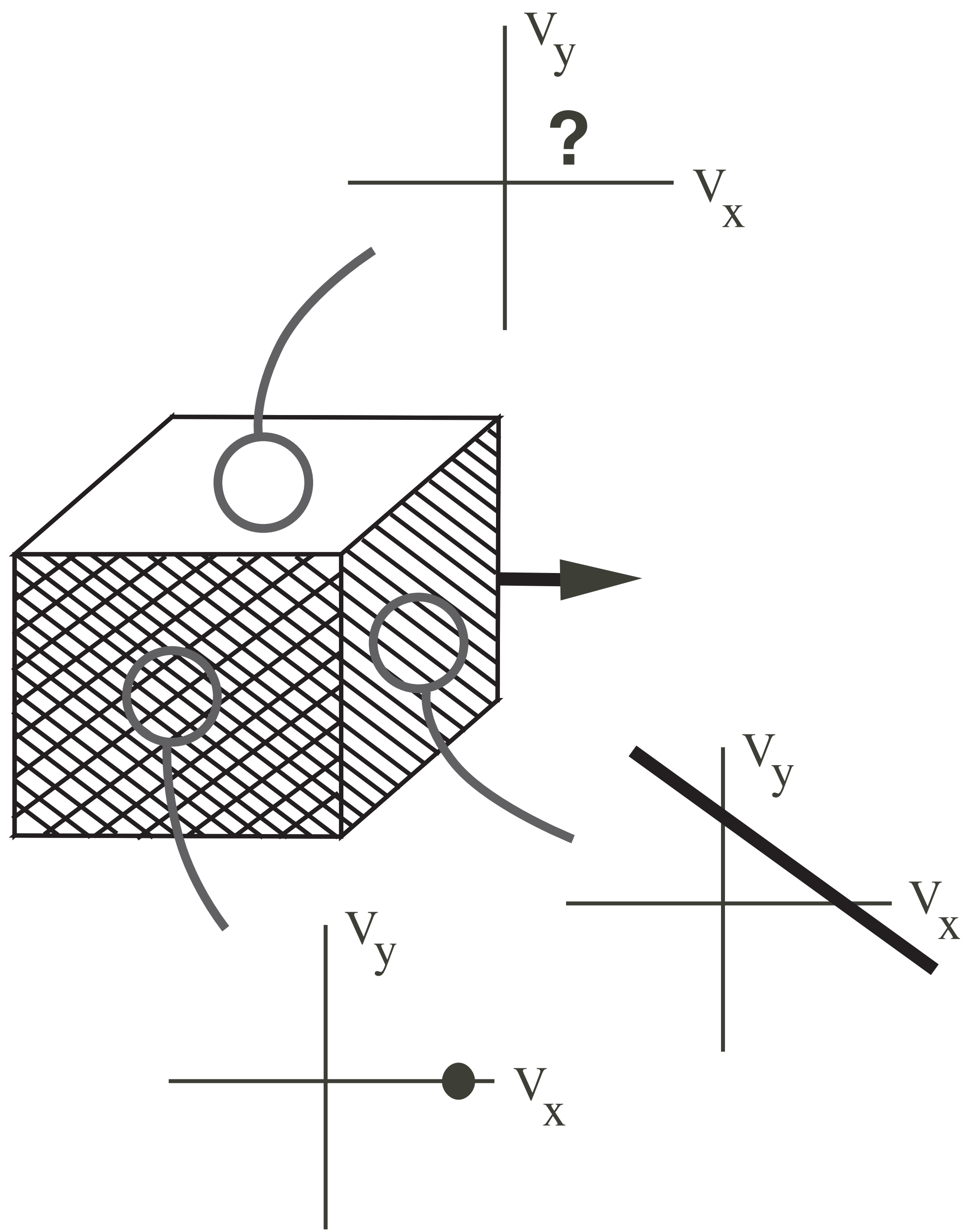


Normalization in MT

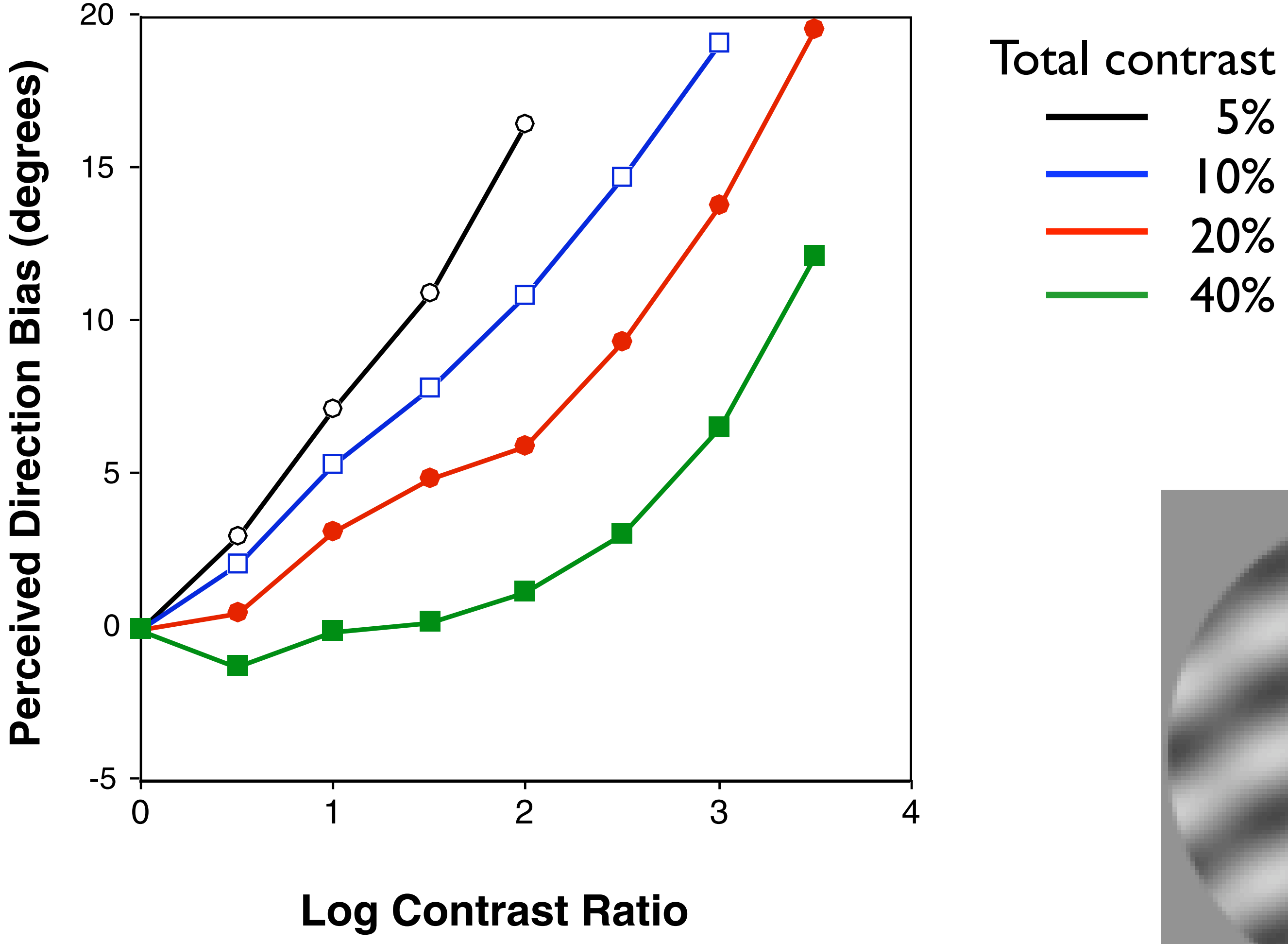


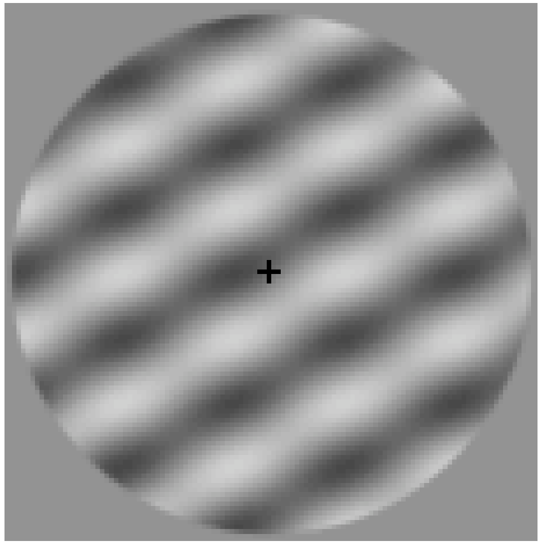
Simoncelli & Heeger, Vis Res, 1998

Visual motion ambiguity



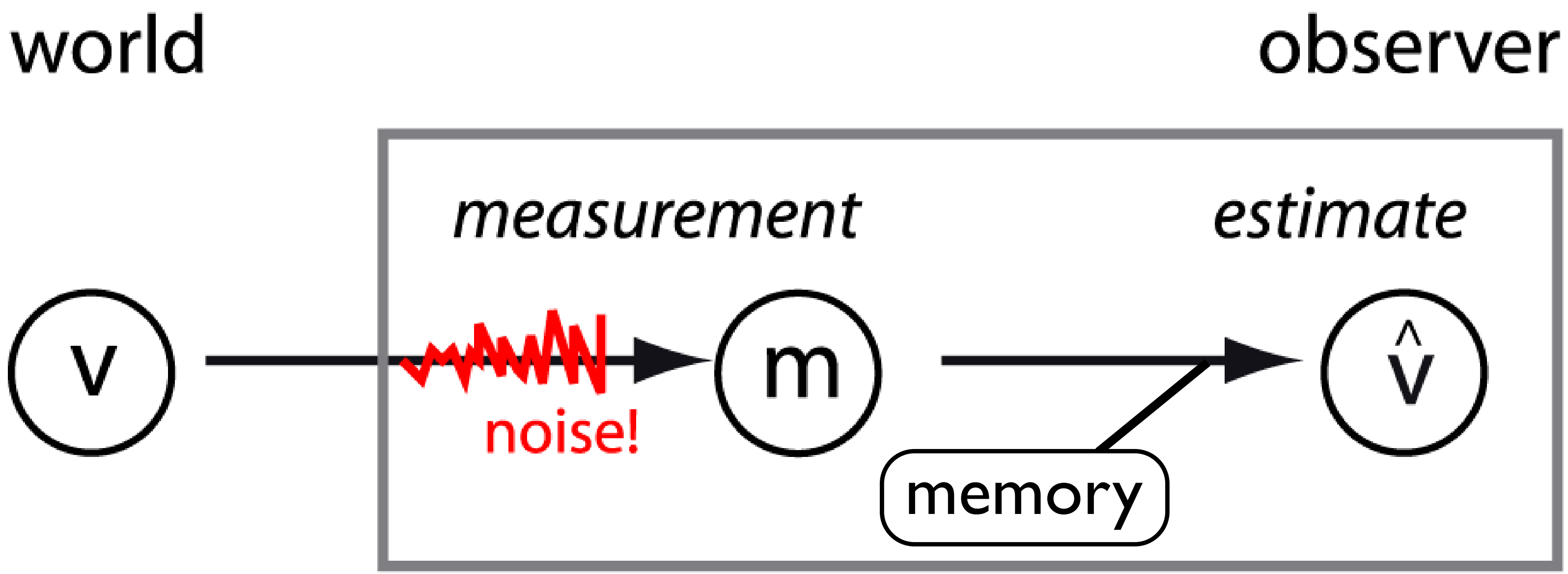
Bias in perceived velocity





Stone, Watson, & Mulligan (1990)

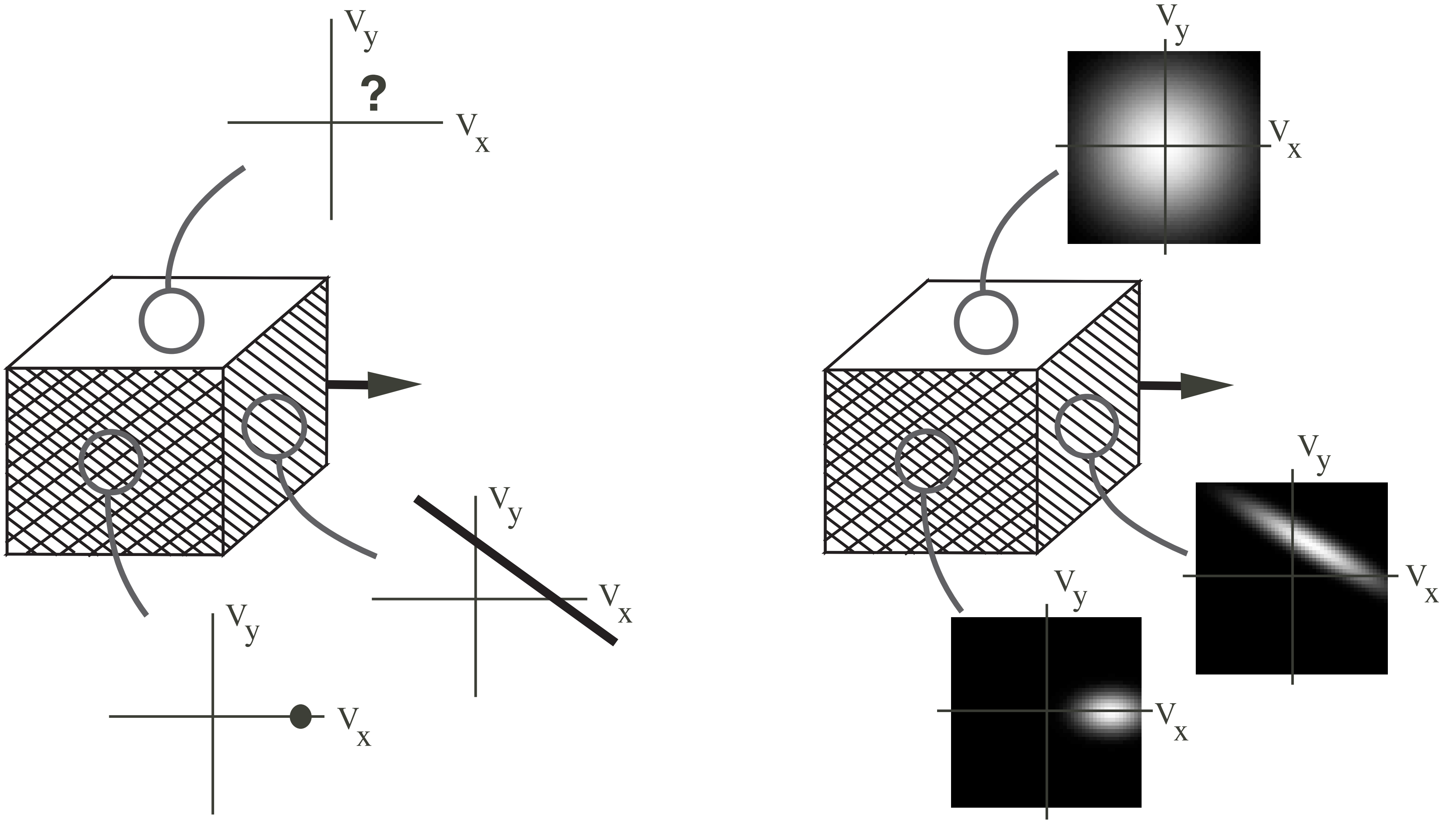
Bayesian models of perception



Perception is our best guess as to what is in the world, given our current sensory input and our prior experience (Helmholtz, 1866).

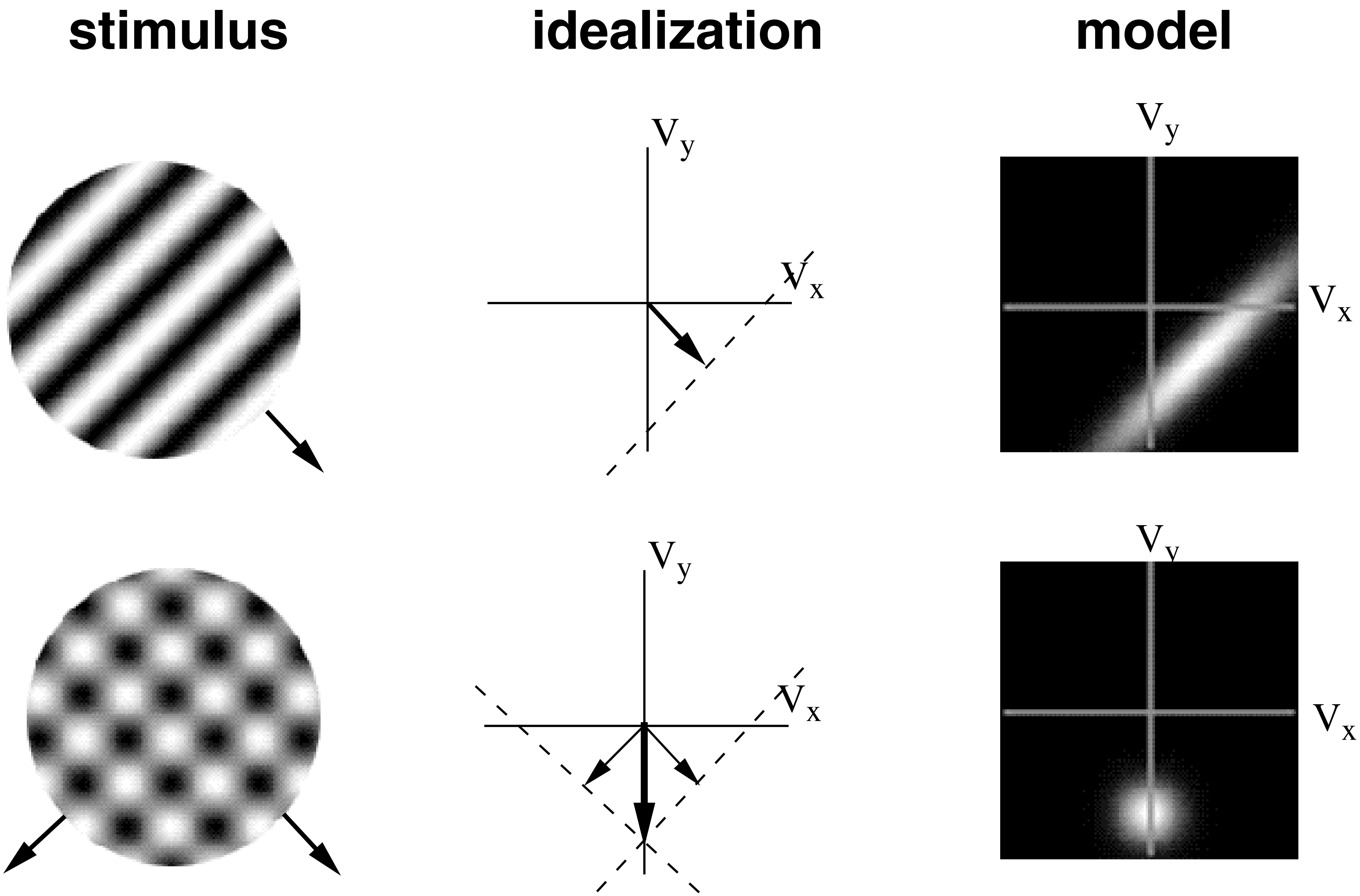
Goal: explain "mistakes" in perception as "optimal" solutions given the statistics of the environment.

Prior bias for slower speeds

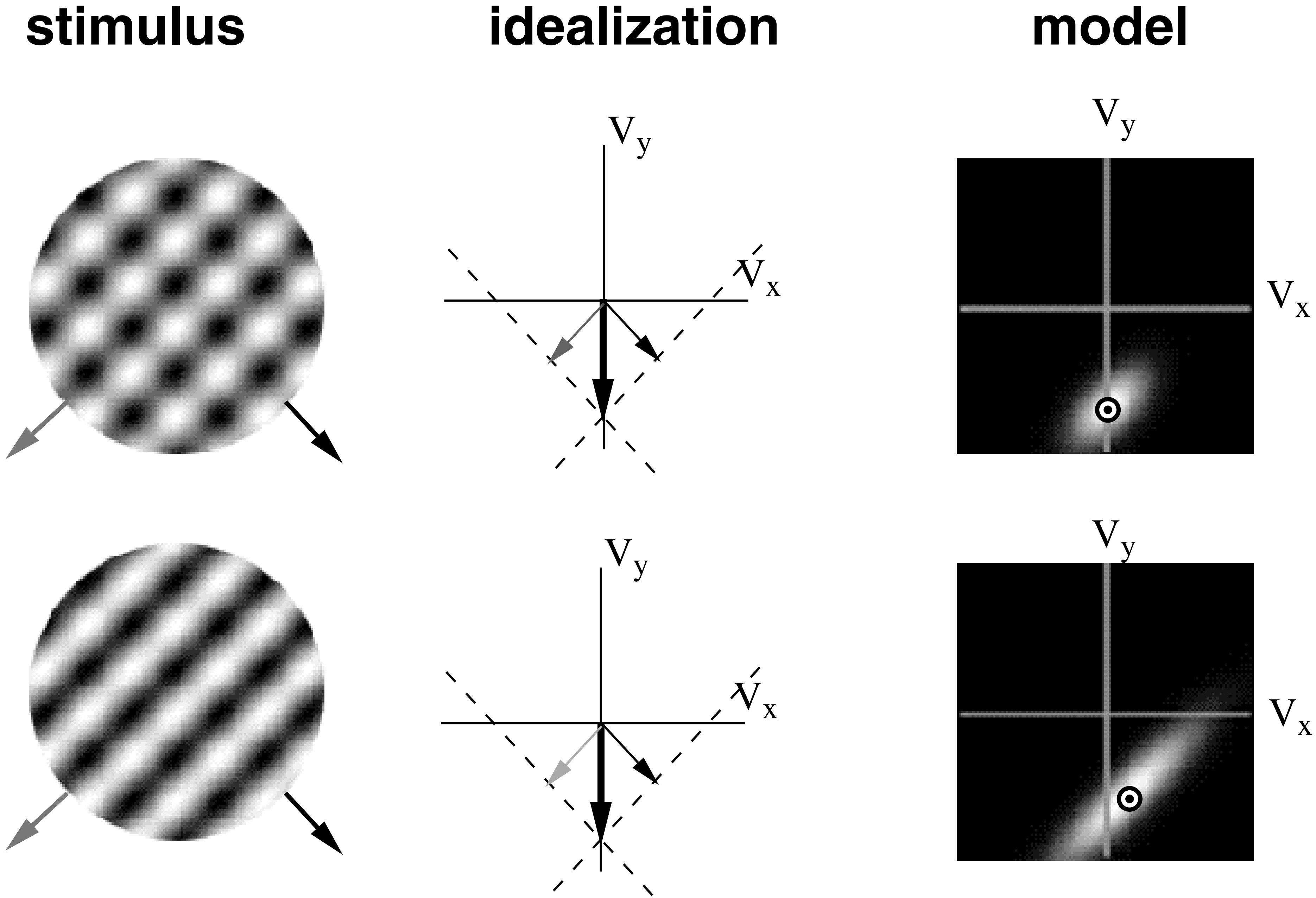


Simoncelli (1993)

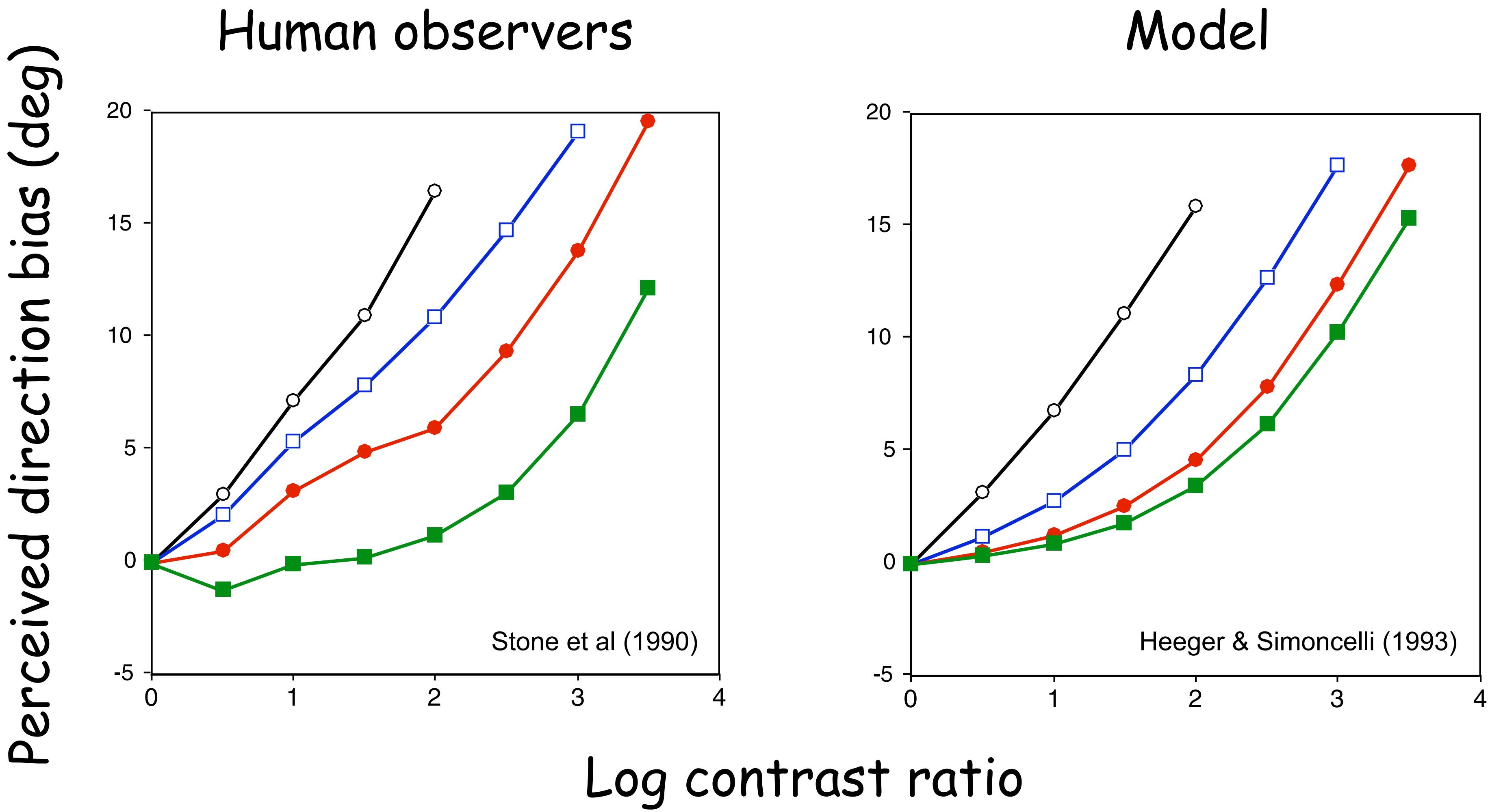
Bayesian model predictions



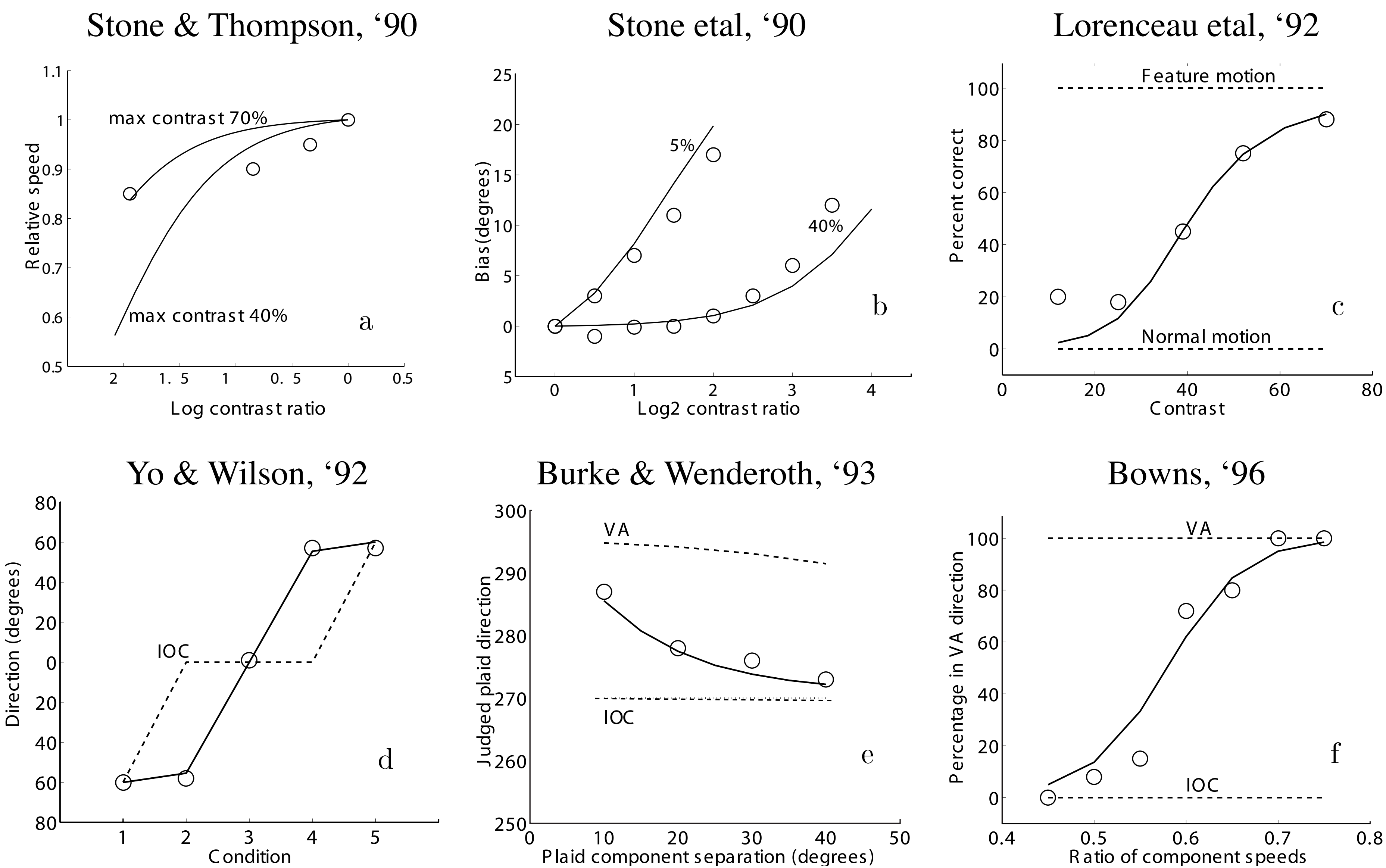
Bayesian model predictions



Prior for slow speeds explains bias in perceptual bias



Theory fits lots of behavioral data

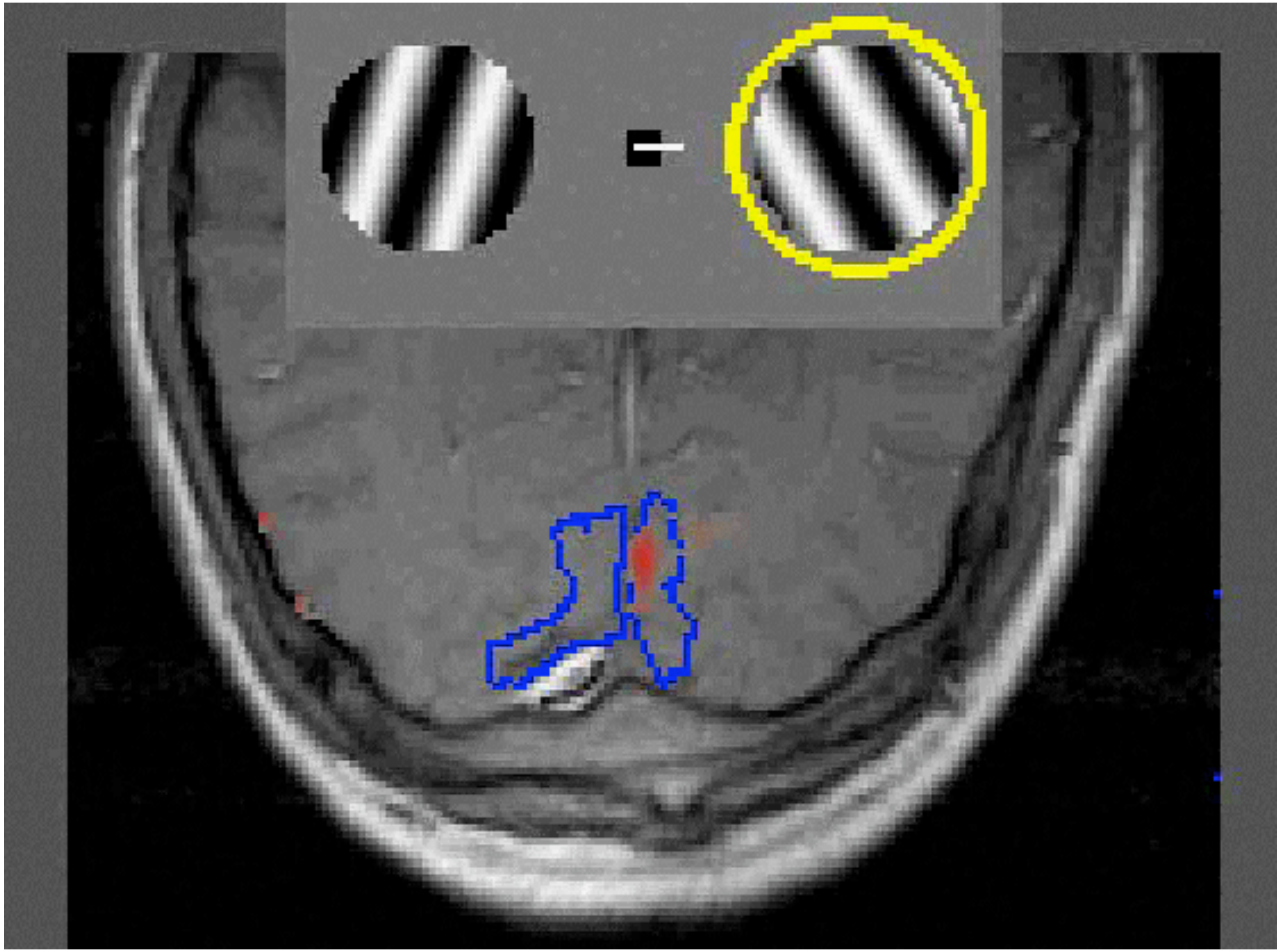


Weiss, Simoncelli, & Adelson (2002)

see also Stocker & Simoncelli (2006)

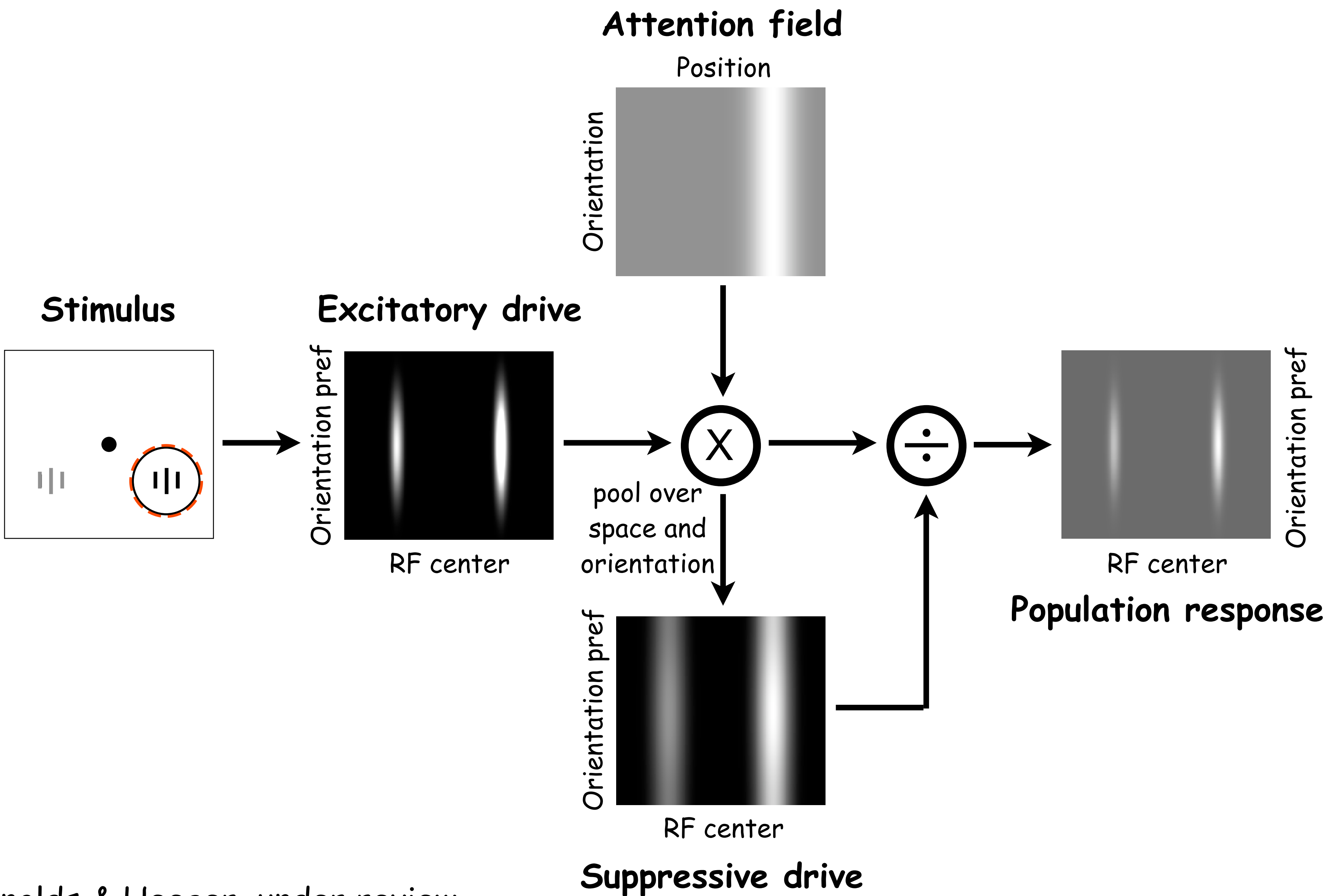
Attention and recurrent (top-down) processing

Attention



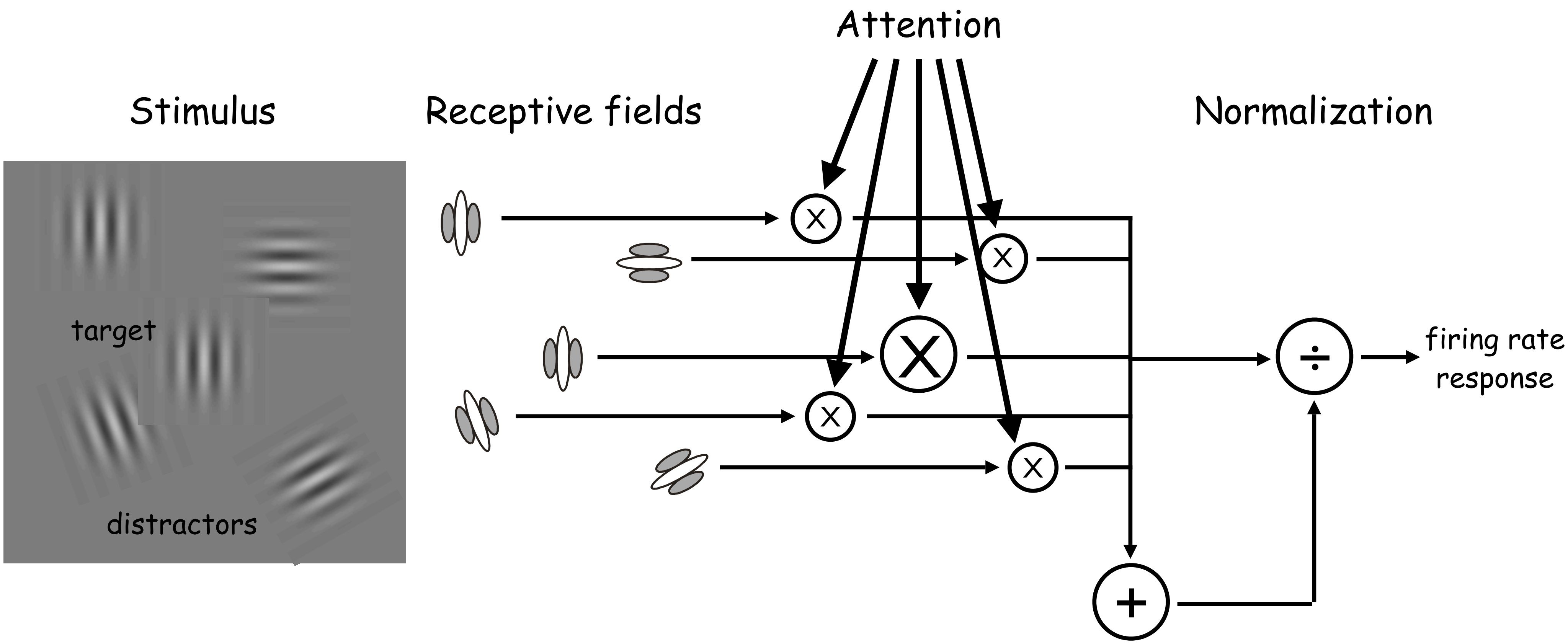
Gandhi, Heeger, & Boynton, PNAS, 1999

Normalization model of attention



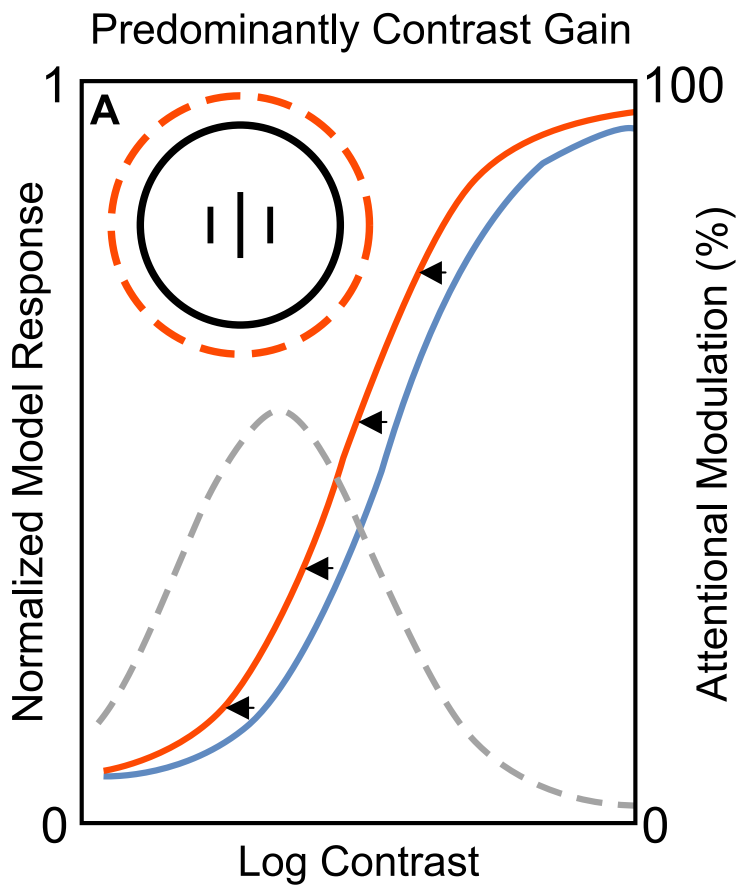
Reynolds & Heeger, under review

Normalization model of attention



Reynolds & Heeger, under review

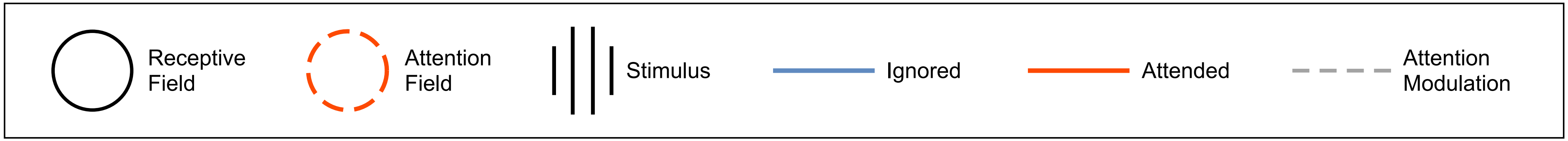
Small stimulus, large attention field



$$R = \alpha \frac{c}{c+\sigma}$$

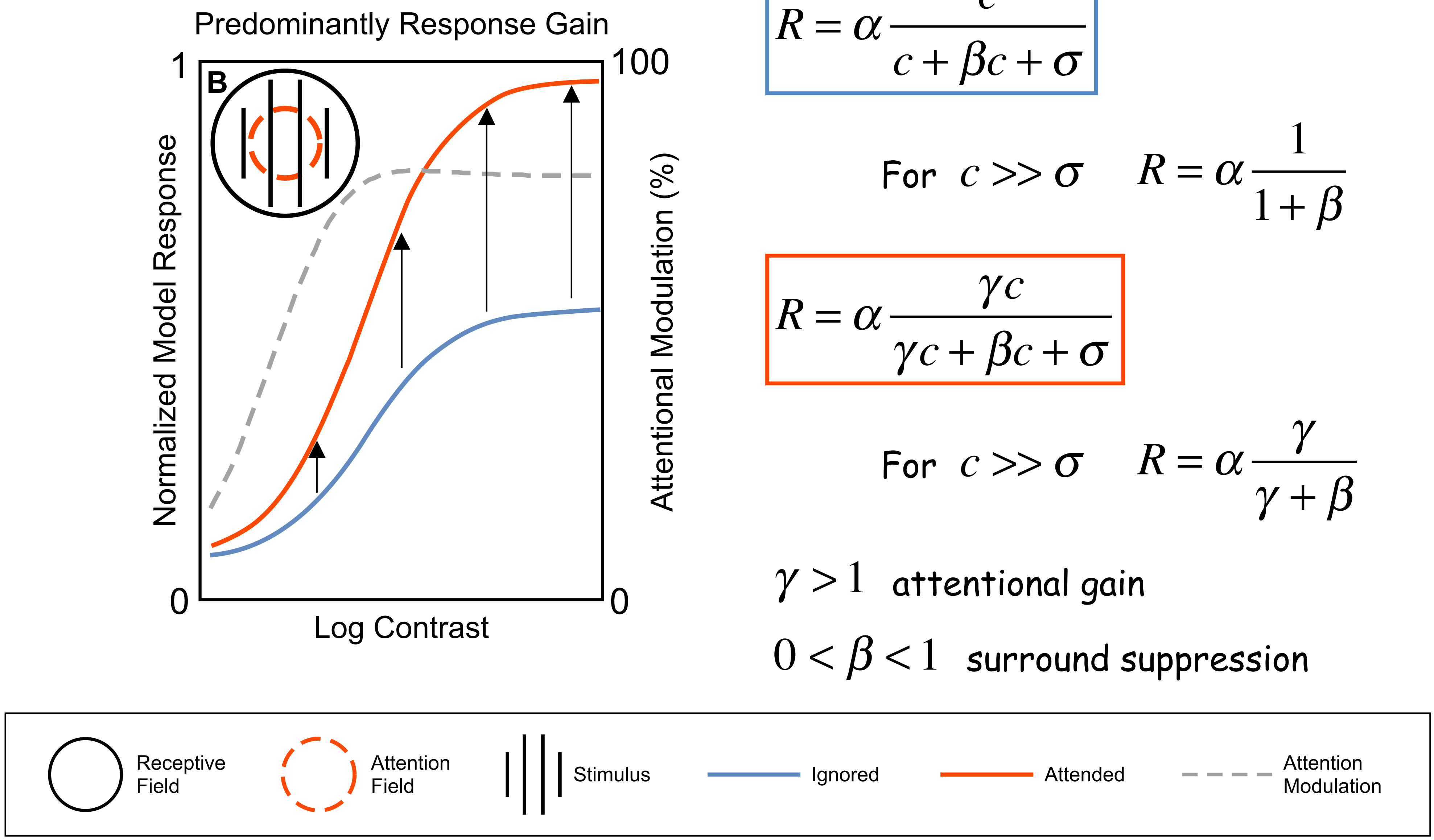
$$R = \alpha \frac{\gamma c}{\gamma c+\sigma} = \alpha \frac{c}{c+\sigma/\gamma}$$

$\gamma > 1$ attentional gain affects excitatory drive and suppressive drive equally



Reynolds & Heeger, under review

Large stimulus, small attention field



$$R = \alpha \frac{c}{c + \beta c + \sigma}$$

For $c >> \sigma$ $\quad R = \alpha \frac{1}{1 + \beta}$

$$R = \alpha \frac{\gamma c}{\gamma c + \beta c + \sigma}$$

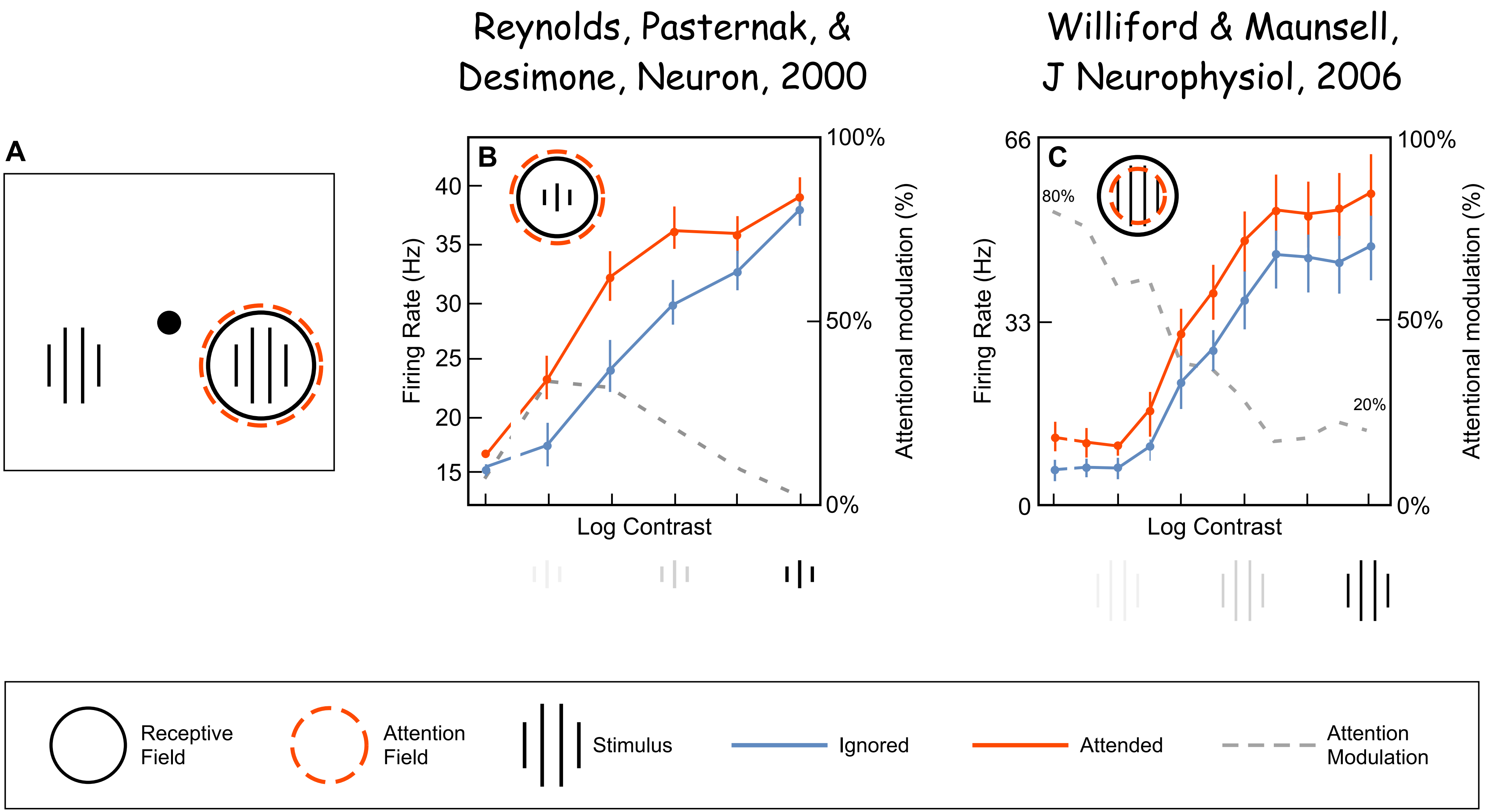
For $c >> \sigma$ $\quad R = \alpha \frac{\gamma}{\gamma + \beta}$

$\gamma > 1$ attentional gain

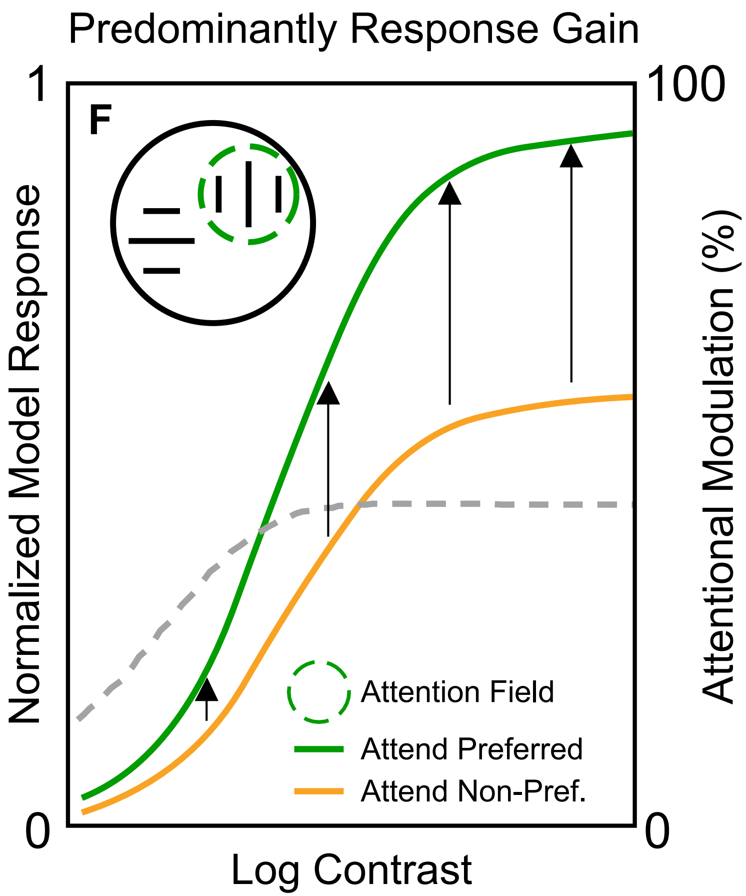
$0 < \beta < 1$ surround suppression

Reynolds & Heeger, under review

Contrast gain & response gain



Attentional selection: gain change with two stimuli in the receptive field



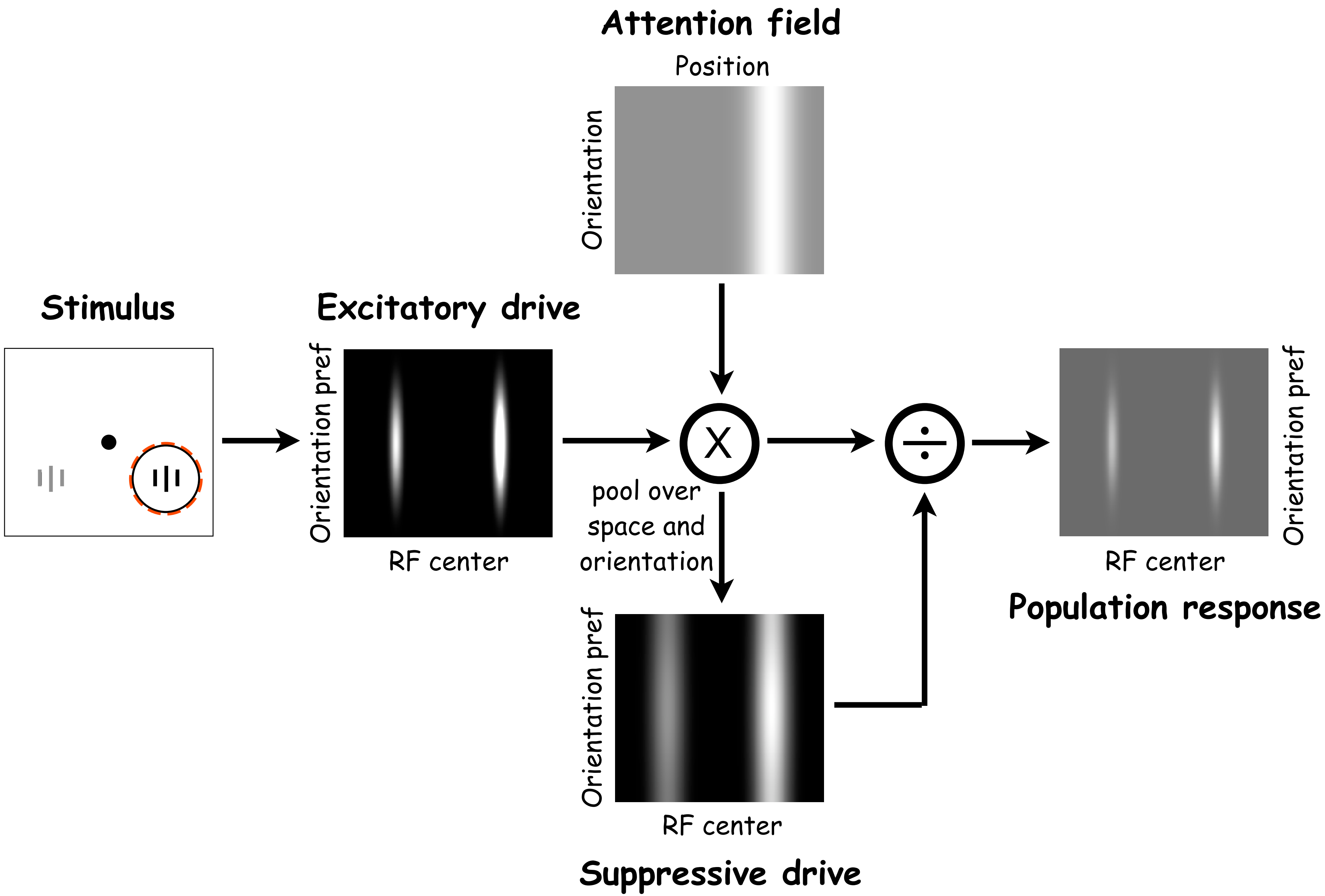
Two stimuli within receptive field. One preferred orientation. Other non-preferred (orthogonal).

$$R_p = \alpha \frac{\gamma c}{\gamma c + c + \sigma}$$

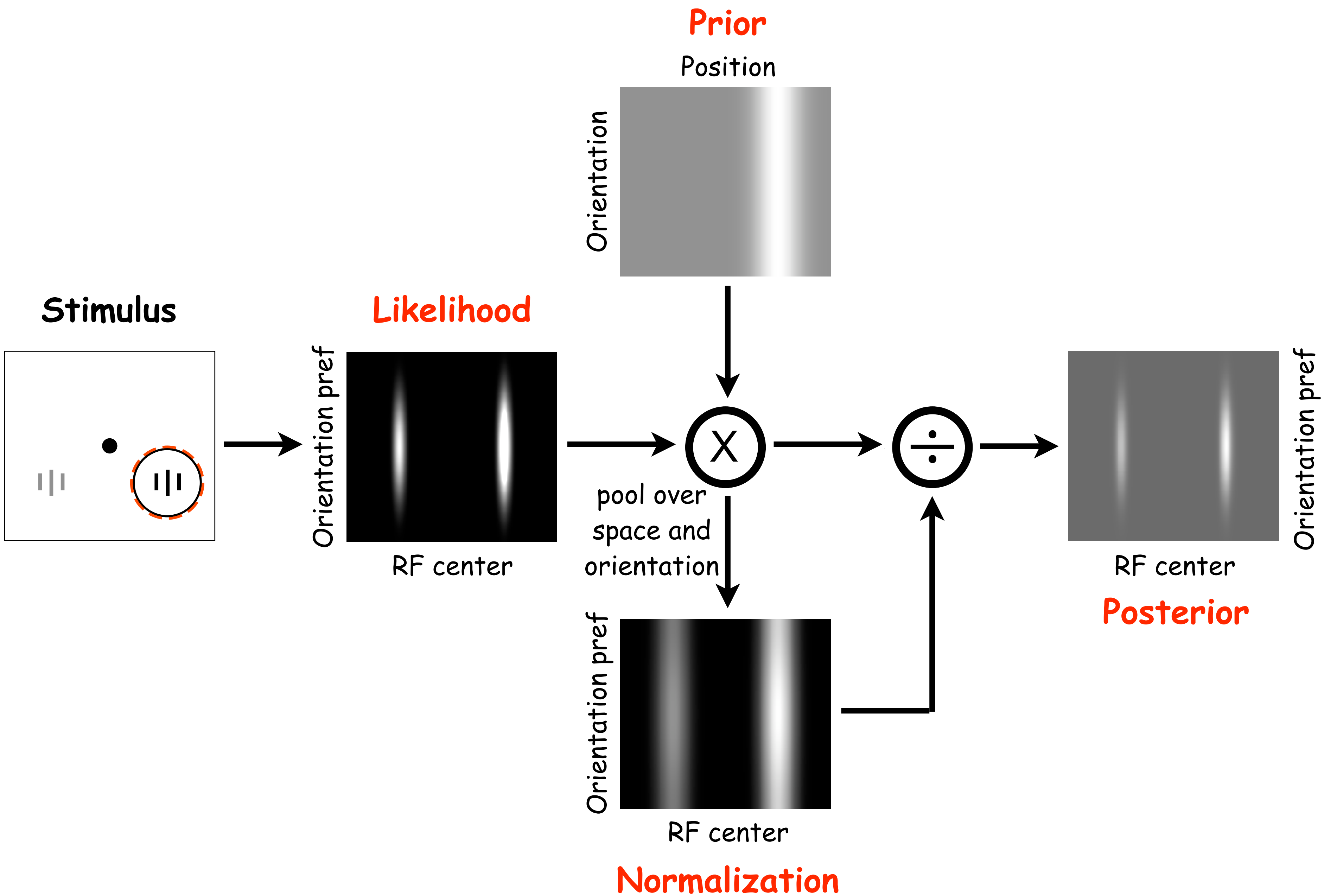
$$R_n = \alpha \frac{c}{c + \gamma c + \sigma}$$

Reynolds & Heeger, under review

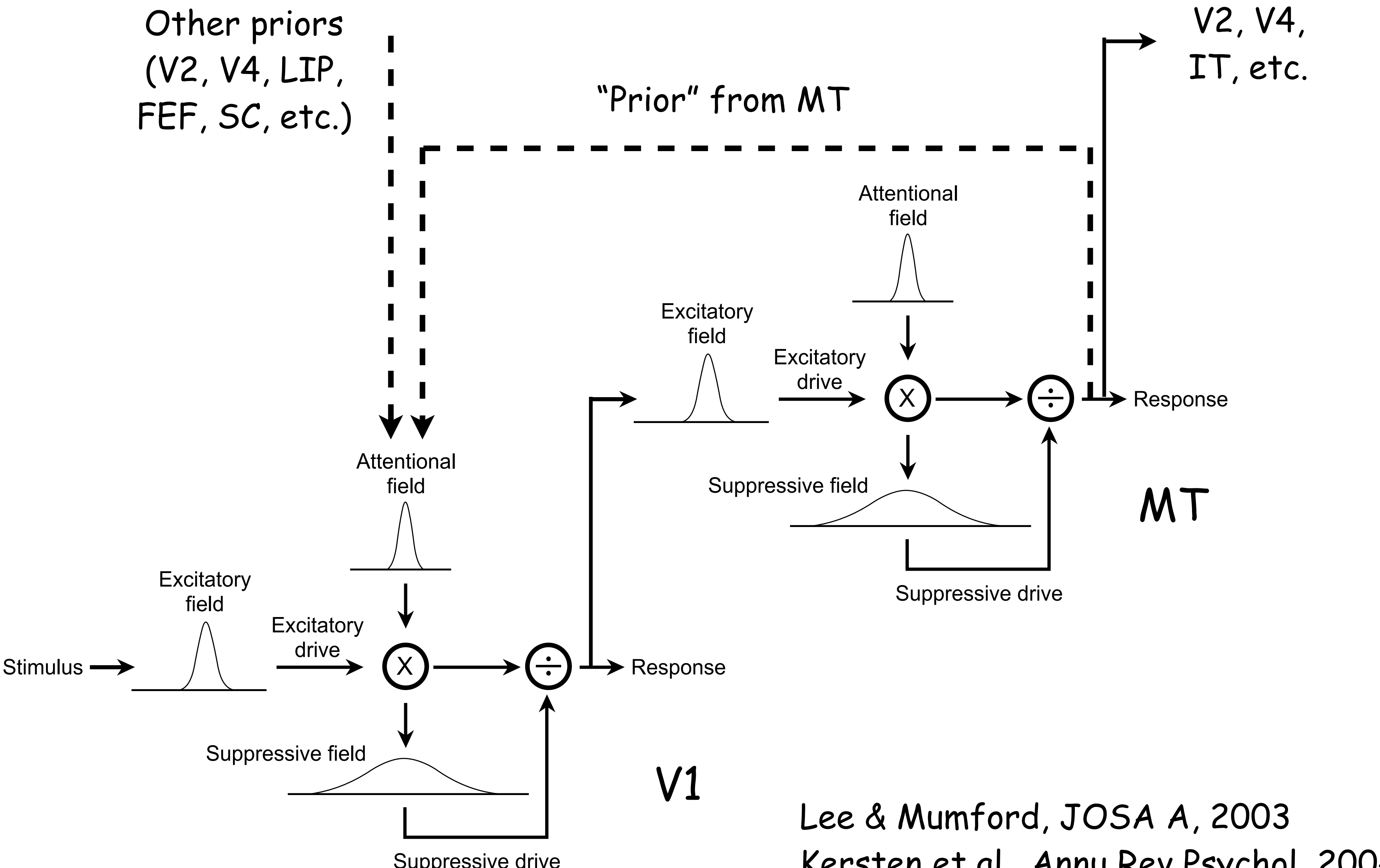
Bayesian inference



Bayesian inference



Hierarchical Bayesian Inference



Lee & Mumford, JOSA A, 2003
Kersten et al., Annu Rev Psychol, 2004

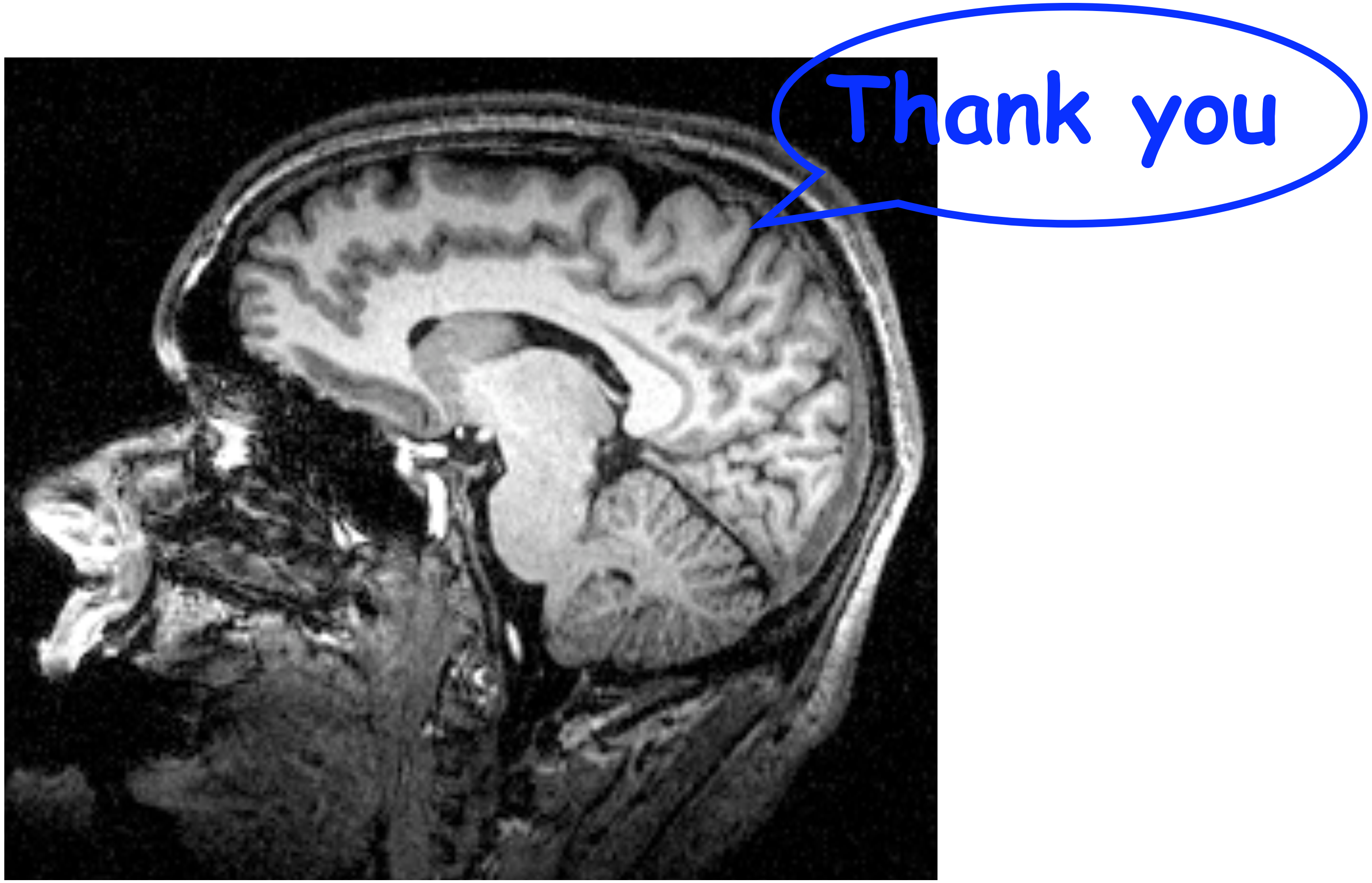
What distinguishes neural activity that underlies conscious visual appearance?

- Neural activity in certain brain areas.
- Activity of specific subtypes of neurons.
- Particular temporal patterns of neural activity (e.g., oscillations).
- Synchronous activity across groups of neurons in different brain areas.
- Neural activity that is driven by a coherent combination of bottom-up sensory information and top-down recurrent processing (e.g., linked to attention).
- Nothing. Once you know the computations, you're done!

Summary

- Functional specialization and computational theory: two balancing principles in the field.
- Visual cortical areas: physiology, architecture, connections, topography.
- Parallel pathways: hierarchy of processing with increasingly complex selectivity, increasing invariance, and increasing RF size.
- Canonical computation: linear sum, halfwave rectification or squaring, normalization, adaptation.
- Recurrent/feedback/top-down processing: attention and hierarchical Bayesian inference.
- Visual awareness?

Thank you


Thank you