Perception & Attention

Perception is effortless but its underlying mechanisms are incredibly sophisticated.

- Biology of the visual system
- Representations in primary visual cortex and Hebbian lea
- Object recognition
- Attention: Interactions between systems involved in objec recognition and spatial processing

Some motivating questions:

1. Why does primary visual cortex encode oriented bars of l

Perception & Attention

Some motivating questions:

1. Why does primary visual cortex encode oriented bars of l

2. Why is visual system split into what/where pathways?

Perception & Attention

Some motivating questions:

1. Why does primary visual cortex encode oriented bars of l

2. Why is visual system split into what/where pathways?

3. Why does parietal damage cause attention problems (neg

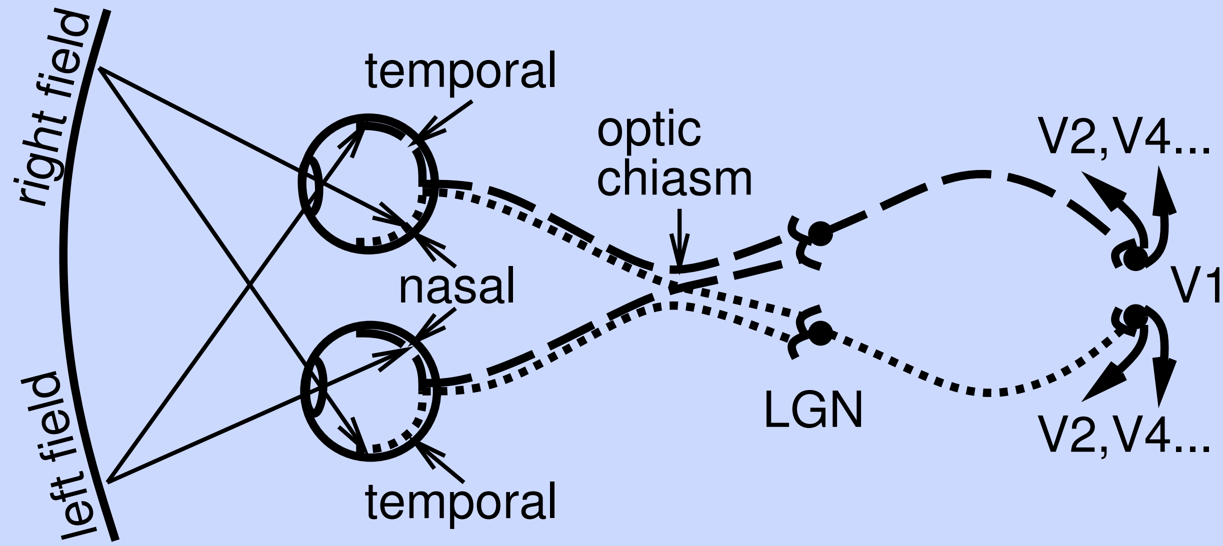
Perception & Attention

Some motivating questions:

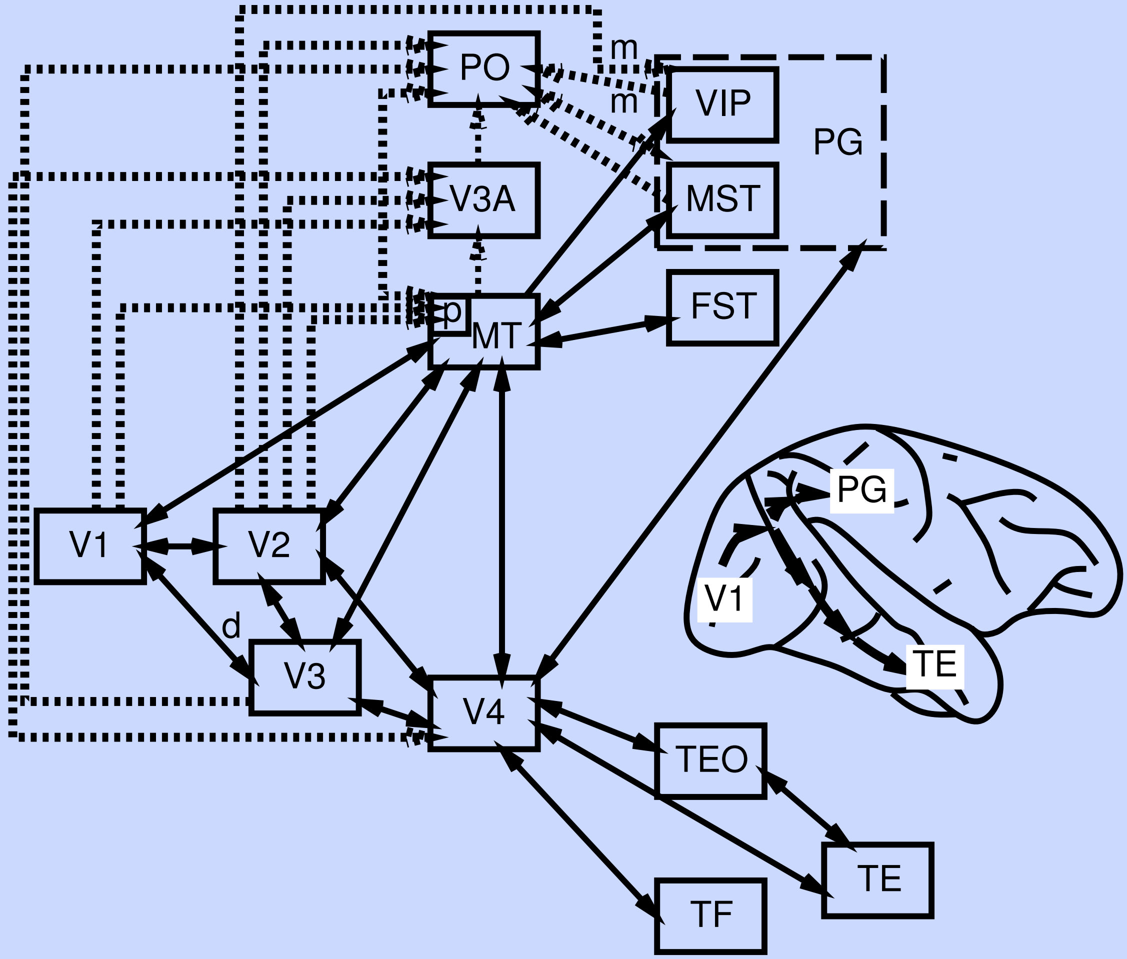
1. Why does primary visual cortex encode oriented bars of l
2. Why is visual system split into what/where pathways?
3. Why does parietal damage cause attention problems (neg
4. How do we recognize objects (across locations, sizes, rota
   with wildly different retinal images)?

Overview of the Visual System

Hierarchies of specialized visual pathways, starting in retina, t
LGN (thalamus), to V1 & up:



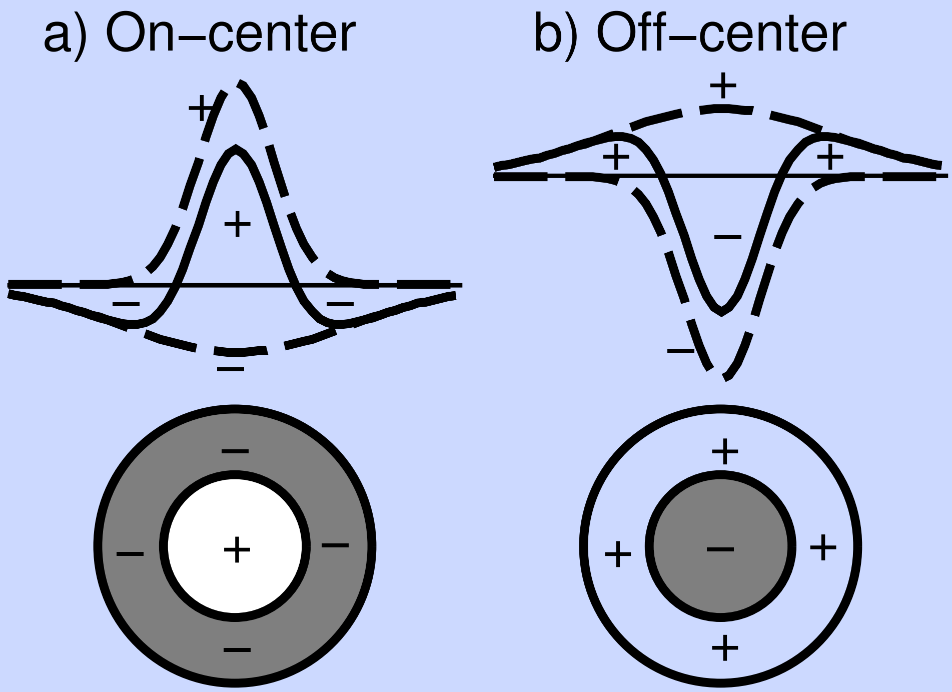
Two Streams: Ventral "what" vs. Dorsal "where"



The Retina

Retina is *not* a passive “camera”

Key principle: *contrast enhancement* that emphasizes *changes*
space & time.



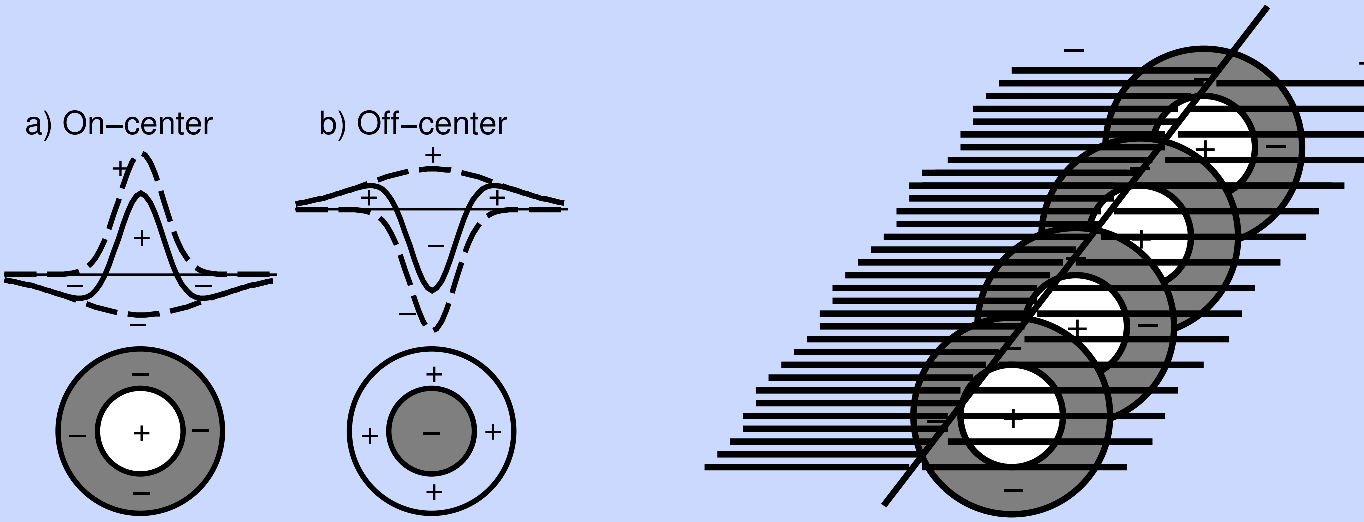
LGN of the Thalamus

A "relay station", but so much more!

- Organizes different types of information into different lay
- Performs *dynamic* processing: magnocellular motion processing cells, *attentional* processing.
- On- and off-center information from retina is preserved in

Primary Visual Cortex (V1): Edge Detectors

V1 combines LGN (thalamus) inputs into oriented *edge detec*

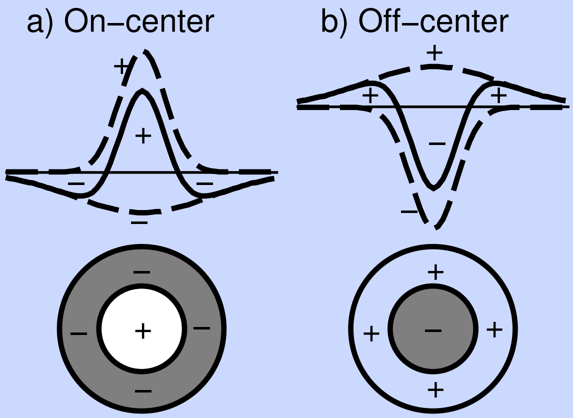


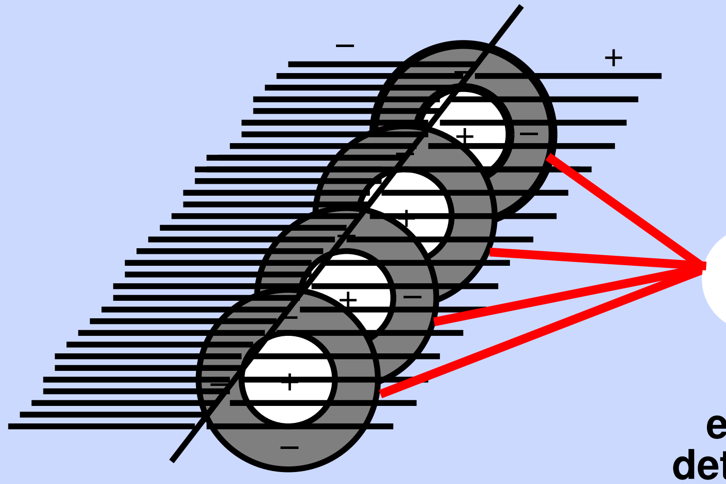
- Edges differ in orientation, size (spatial frequency), and position.

- For coherent vision, need to detect varying degrees of all

Primary Visual Cortex (V1): Edge Detectors

V1 combines LGN (thalamus) minputs into oriented *edge detectors*:

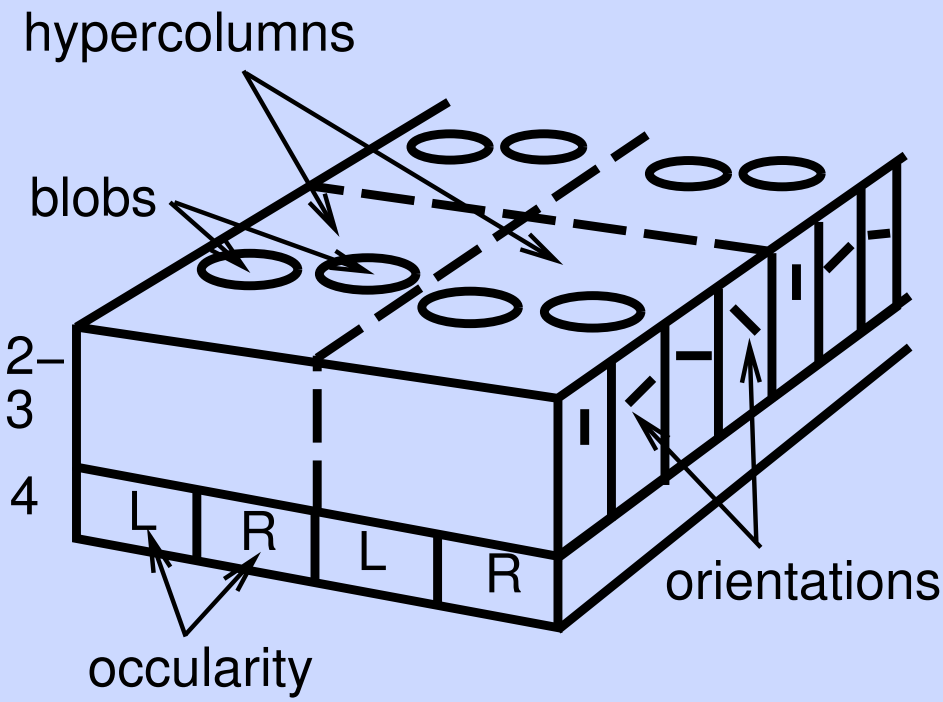


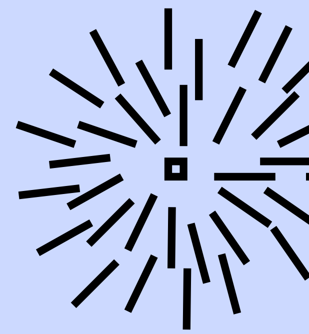


- Edges differ in orientation, size (spatial frequency), and position.

- For coherent vision, need to detect varying degrees of all t

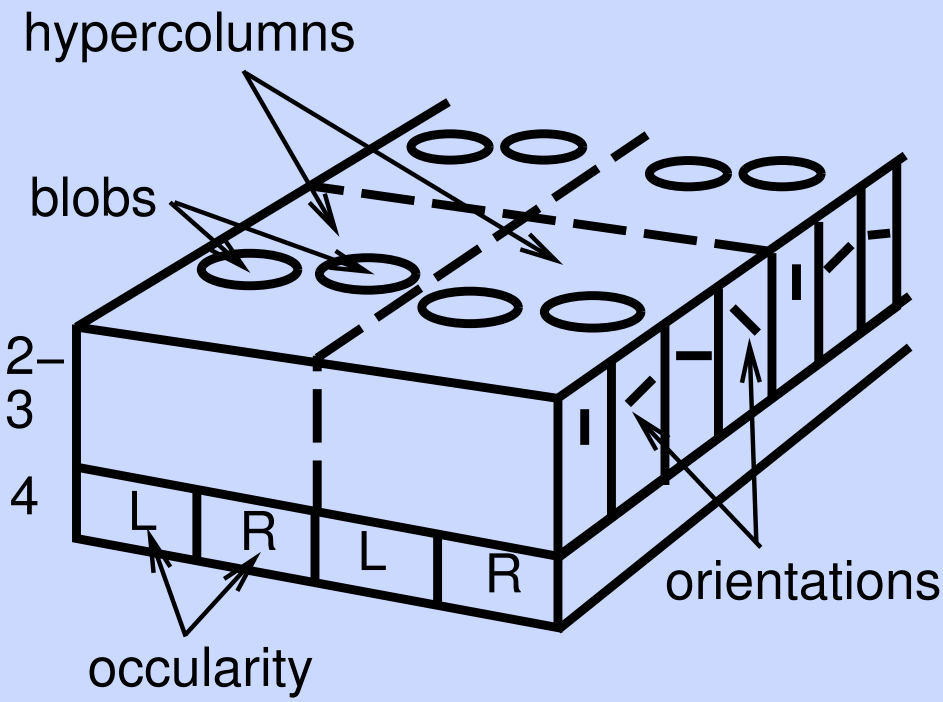
Primary Visual Cortex (V1): Topography

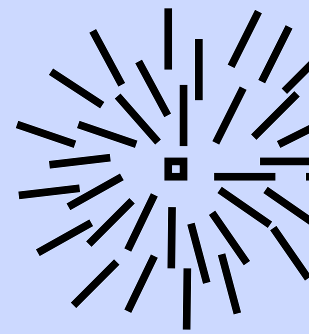




Pinwheel

Primary Visual Cortex (V1): Topography



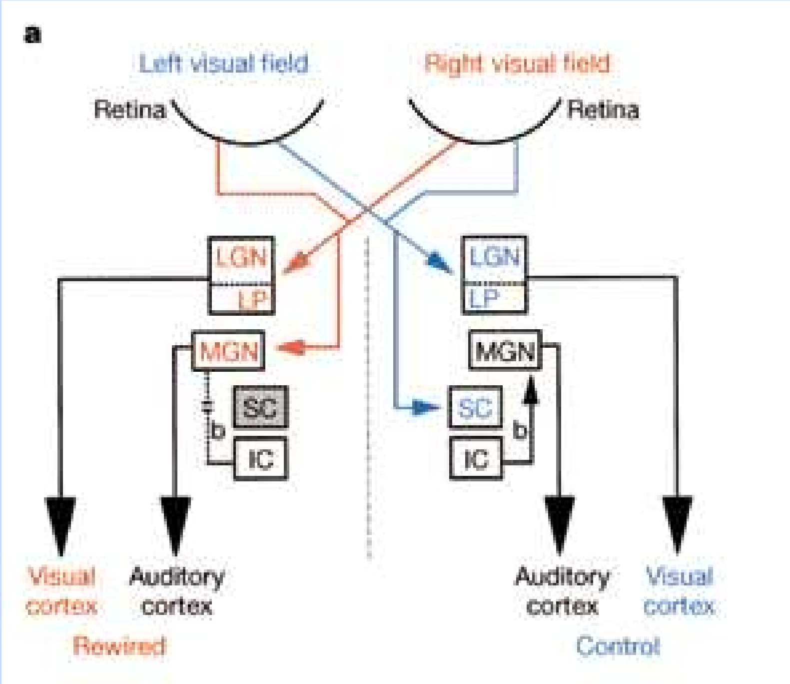


Pinwheel

Pinwheel can arise from *learning* and lateral connectivity: not hard-wired!

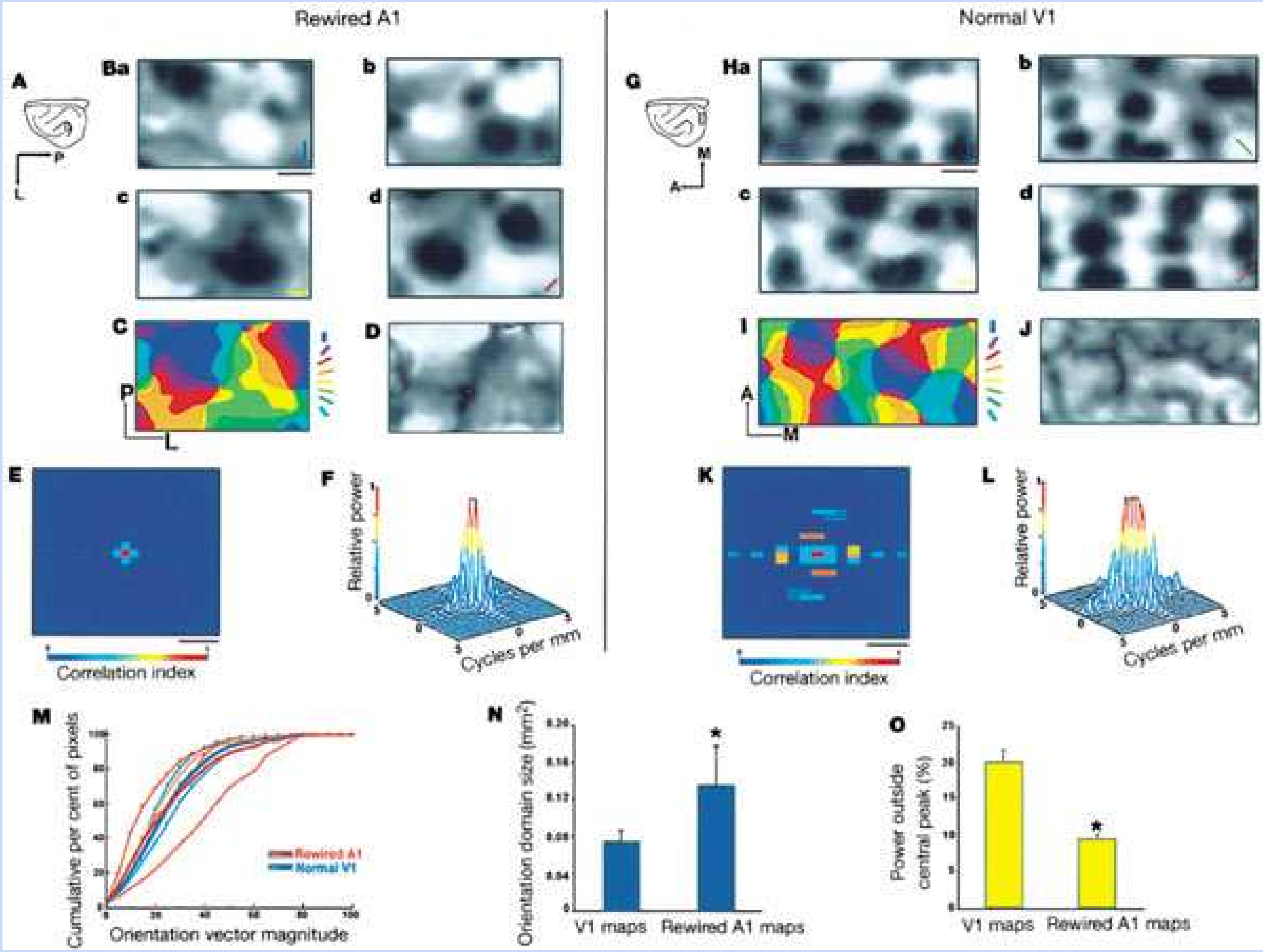
Rerouting of Visual Info to Auditory Cortex

- Sharma, Angelucci & Sur (2000), *Nature*
  Rerouted fibers from Retina→ auditory thalamus (MGN)



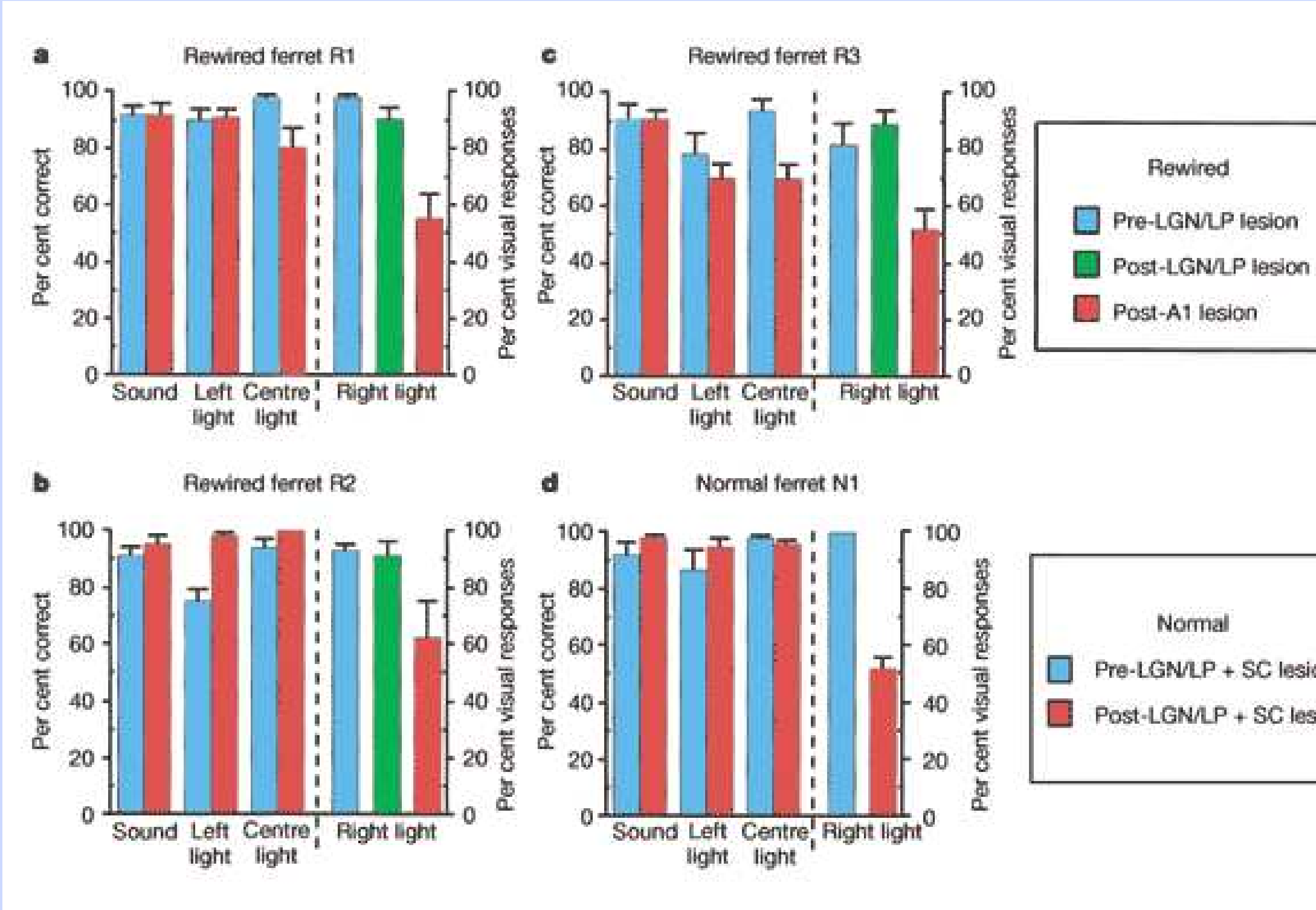
- If visual properties are learned, they should develop in

Rerouting of Visual Orientation Modules in A1

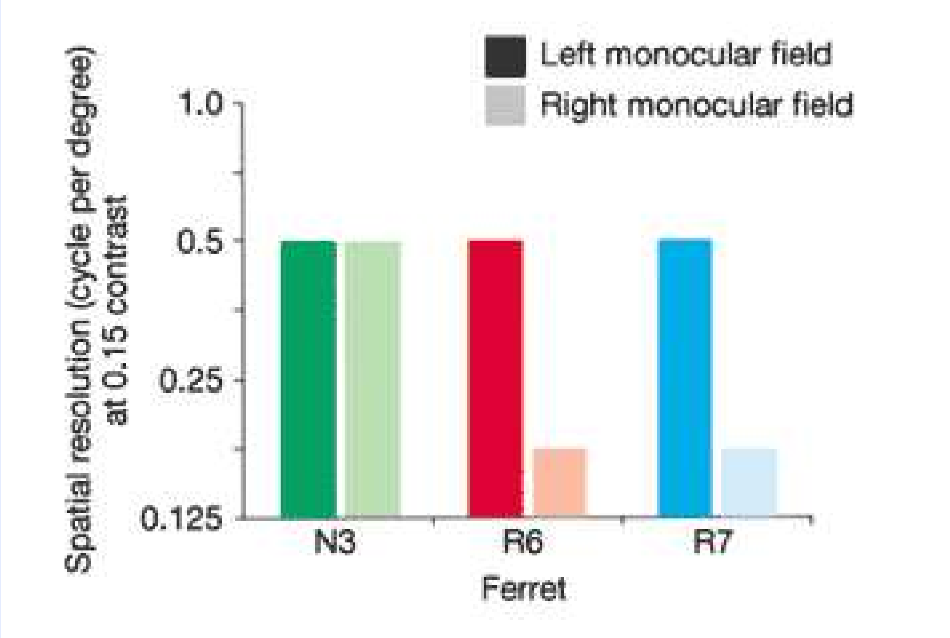


Visual Behavior After Rerouting

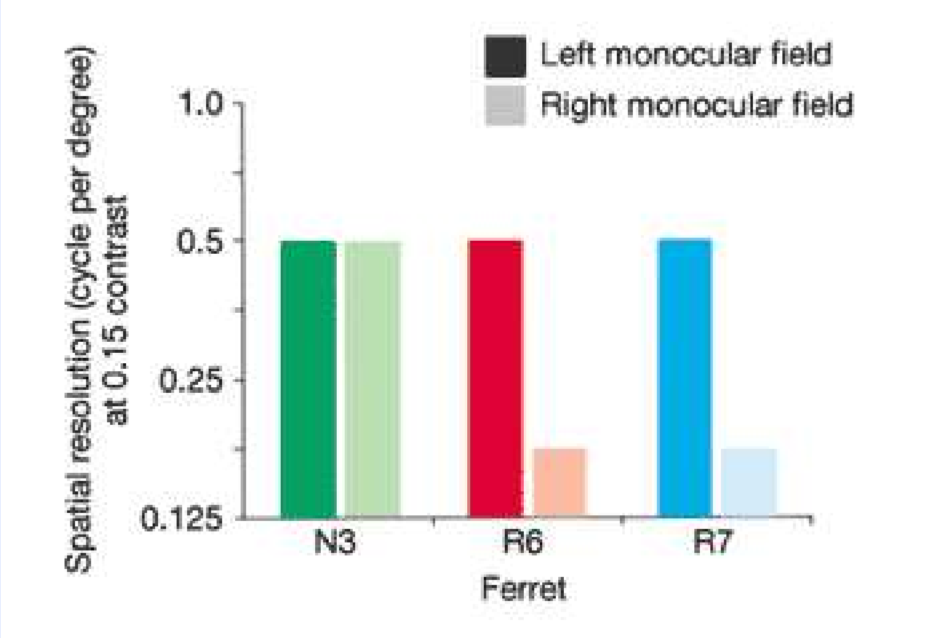
von Melchner, Pallas & Sur (2000)



Visual Acuity After Rerouting



Visual Acuity After Rerouting



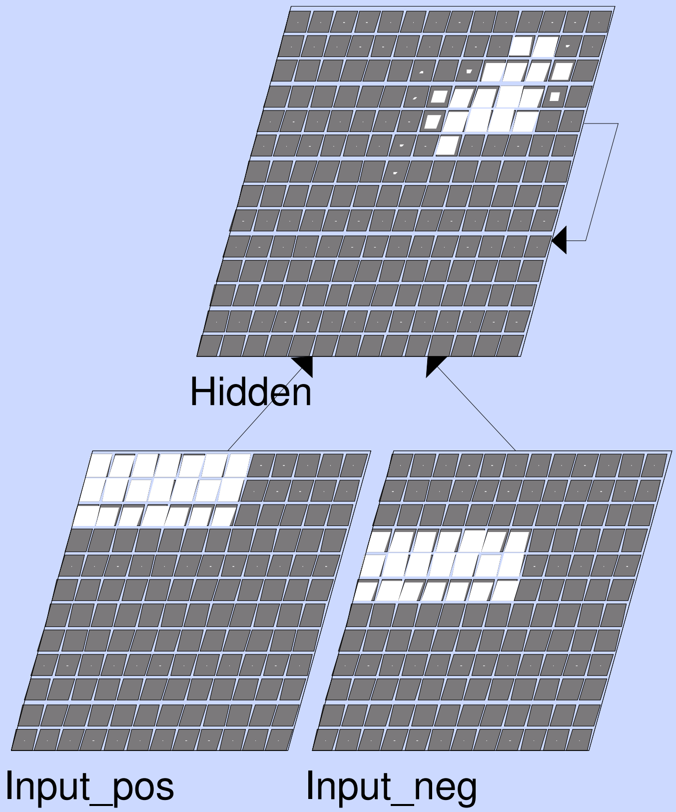
→ So learning is powerful, but so is evolution!

A Question

What makes visual cortex visual cortex? Why does it represen
oriented bars of light?

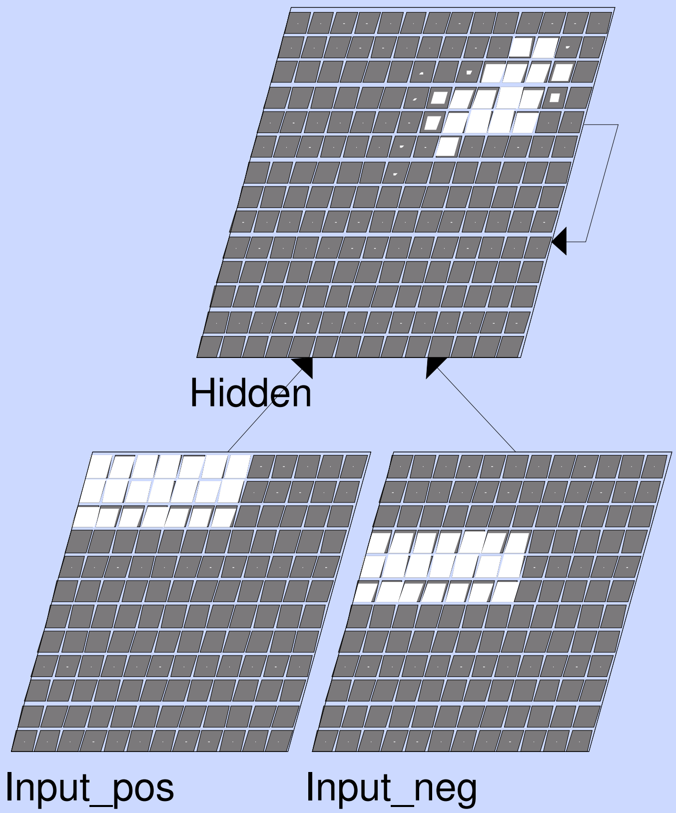
Key idea: Oriented edge detectors can develop from Hebbian correlational learning based on natural visual scenes.

The Model: Simulating one Hypercolumn



- Natural visual scenes are preprocessed by passing them (separately) through layers of on-center and off-center inputs

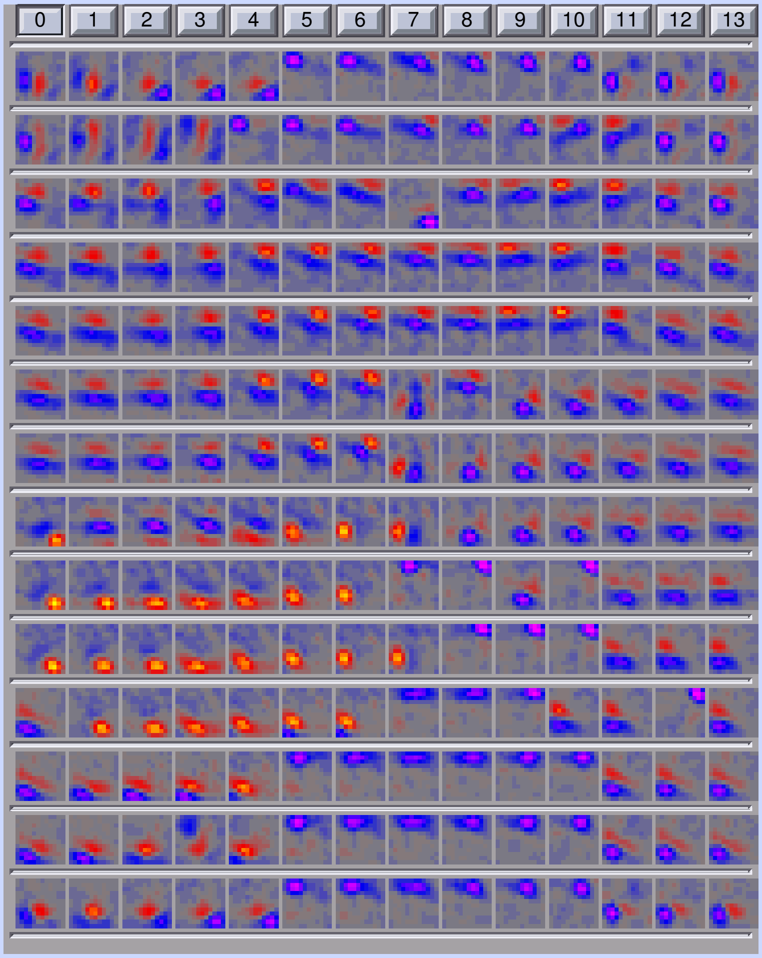
- Hidden layer: edge detectors seen in layers 2/3 of V1; Layer 4 (input) represents unoriented on/off inputs like LGN (modulated by attentio



- Hebbian learning only
- KWTA inhib competition for specialization (see Ch 4)

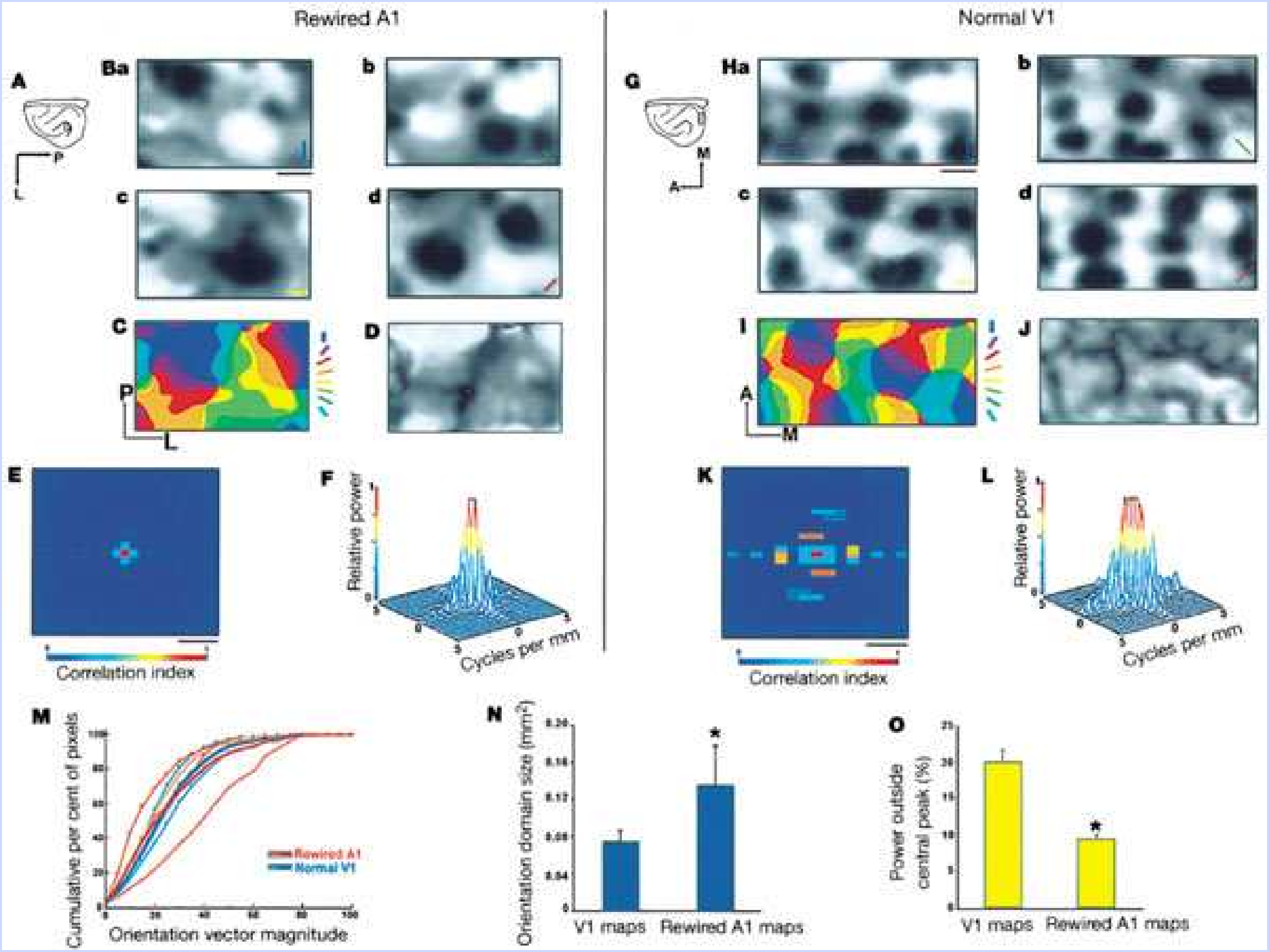
[v1rf.proj.gz]

The Receptive Fields

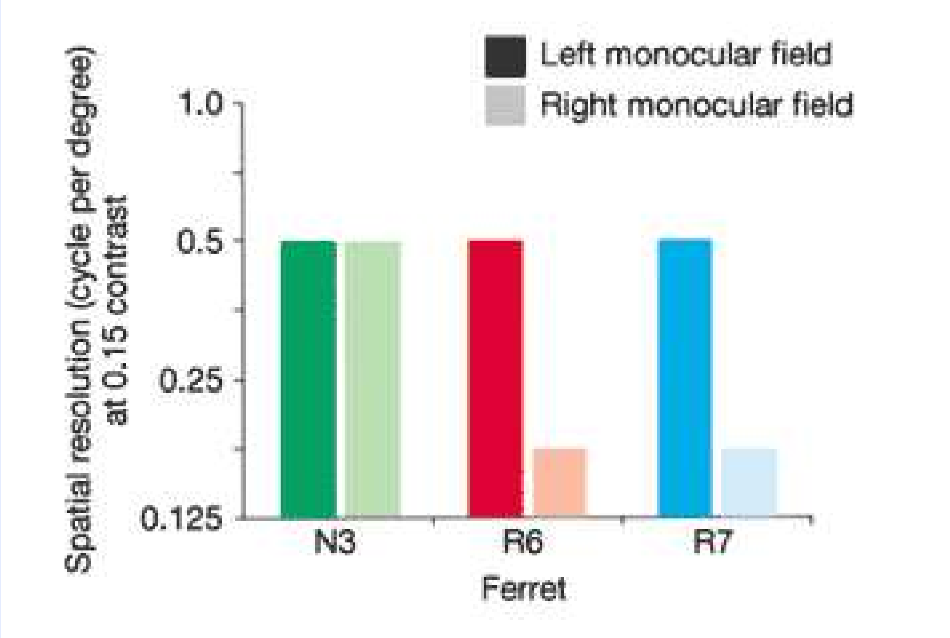


Red = on-center > off-center, Blue = off-center > on-center

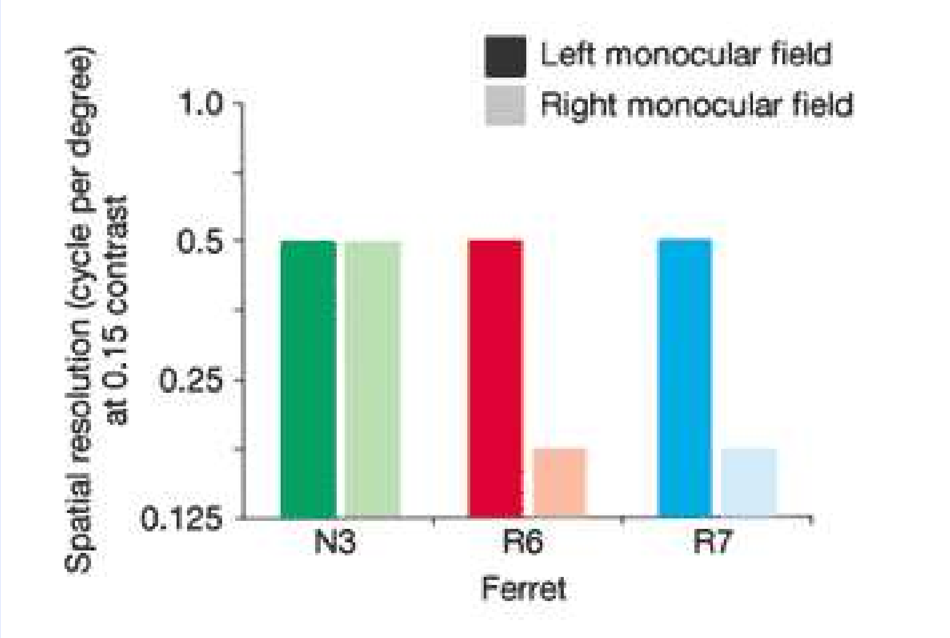
Rerouting of Visual Orientation Modules in A1



Visual Acuity After Rerouting

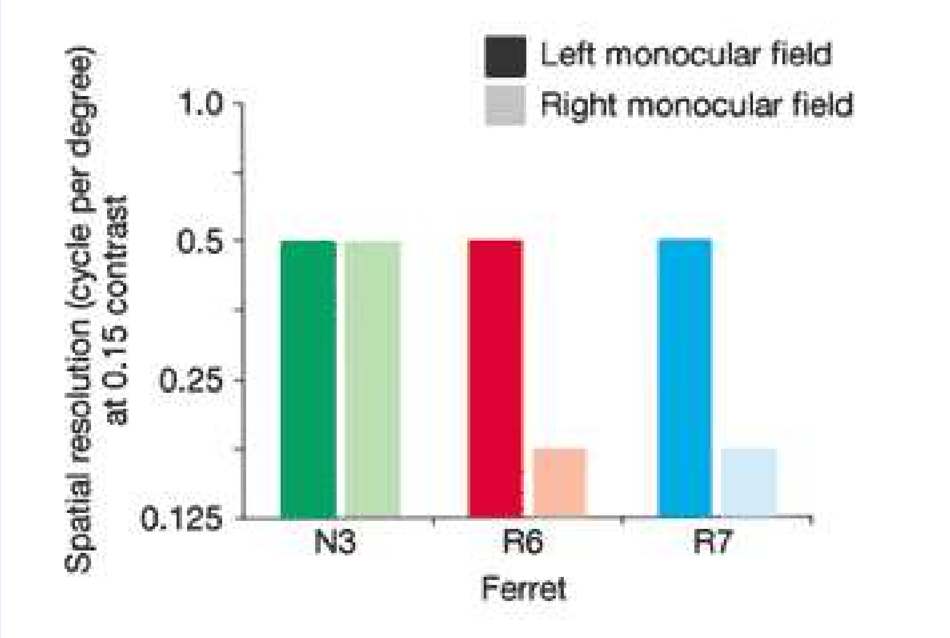


Visual Acuity After Rerouting



→ So learning is powerful, but so is evolution!

Visual Acuity After Rerouting



→ So learning is powerful, but so is evolution!

How to account for evolution of visual specialization in mod

1. Why does primary visual cortex encode oriented bars of l
   Correlational learning based on natural visual scenes.

Reflects reliable presence of edges in natural images, which va
size, position, orientation and polarity.

→ model shows how documented V1 properties can result fro
interactions between learning, architecture (connectivity), and
structure of environment.

Reading Reactions

- Brad: Do perception models make the same errors people
with visual illusions? This seems like a critical test of a vis
model.

- Anastasia: How would such models bind color to an obje
that isn't always presented in the same color? For exampl
how would these models resolve an input where a red cir
and a blue square are presented?

- Jim: [re: exemplar theories] that the brain stores some sor
ideal form for input comparison is overly simplistic and
ultimately grounded in fundamentals of cognitive theory
rather than principles of neural systems... [but] the book d
not account for the sheer volume of information the corte

must simultaneously handle in order to utilize parallel transformations to represent unique objects.

Perception and Attention

1. Why does primary visual cortex encode oriented bars of l
   Correlational learning based on natural visual scenes.

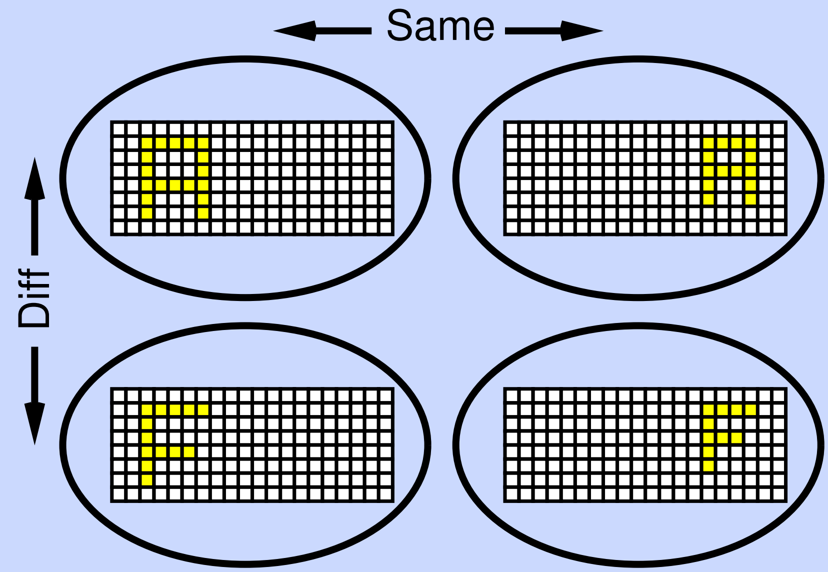
2. How do we recognize objects (across locations, sizes, rota
   with wildly different retinal images)?

3. Why is visual system split into what/where pathways?

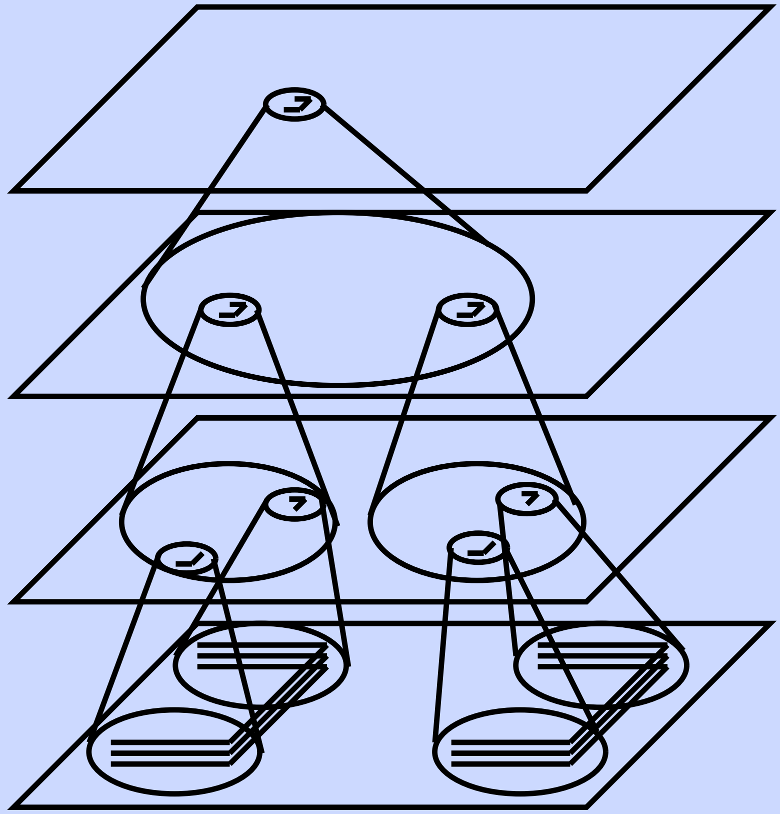
4. Why does parietal damage cause attention problems (neg

The Object Recognition Problem

Problem: Recognize object regardless of: location, size, rotatio

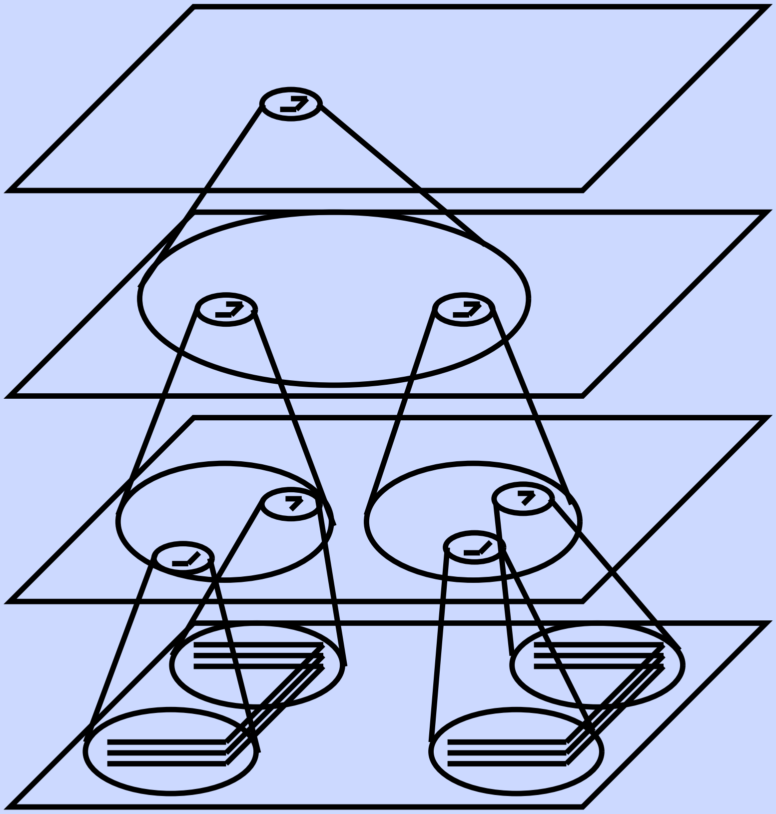


This is hard because different patterns in same location can ov
a lot, while the same patterns in different locations/sizes/rota
can not overlap at all!



Increasing receptive field size enables:
Conjunction of features (to form more complex objects); and
Collapsing over location information ("spatial invariance")

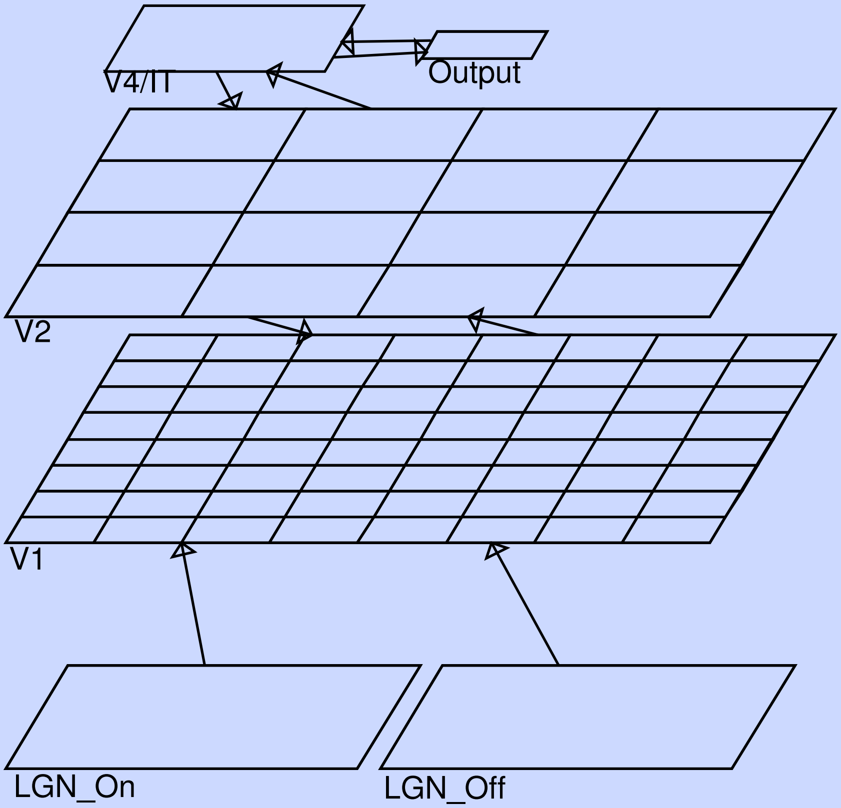
Gradual Invariance Transformations



if did spatial invariance in one fell swoop: binding problem - can't tell T fr

Goal: Units at the top of the hierarchy should represent compl
object *features* in a location and size invariant fashion

The Model



V1 = oriented line (edge) detectors, hard-coded

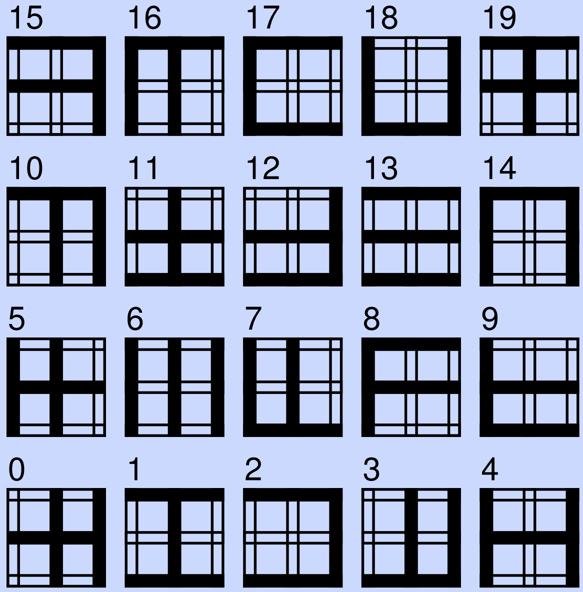
V2 units encode conjunctions of V1 edges across a subset of sp

Each V4 unit pays attention to all of V2

Reading Reaction

- Zaneta: [In the 1st model], the V1 layer used Hebbian lear
to develop orientation pinwheels, but when it was connec
to the other layers it was fixed, and no longer learned by e
mechanism. If it was allowed to keep learning for longer,
would the neurons change their orientation selectivity
gradually over time, since Hebbian learning continues to
occur? .. in the development of real brains, there are critic
periods for learning in different brain regions, after which
point the amount of learning that can occur in that structu
greatly reduced. It would be interesting if this were actua
required to occur in an hierarchical fashion, in order for th
higher layers to learn effectively.

The Objects

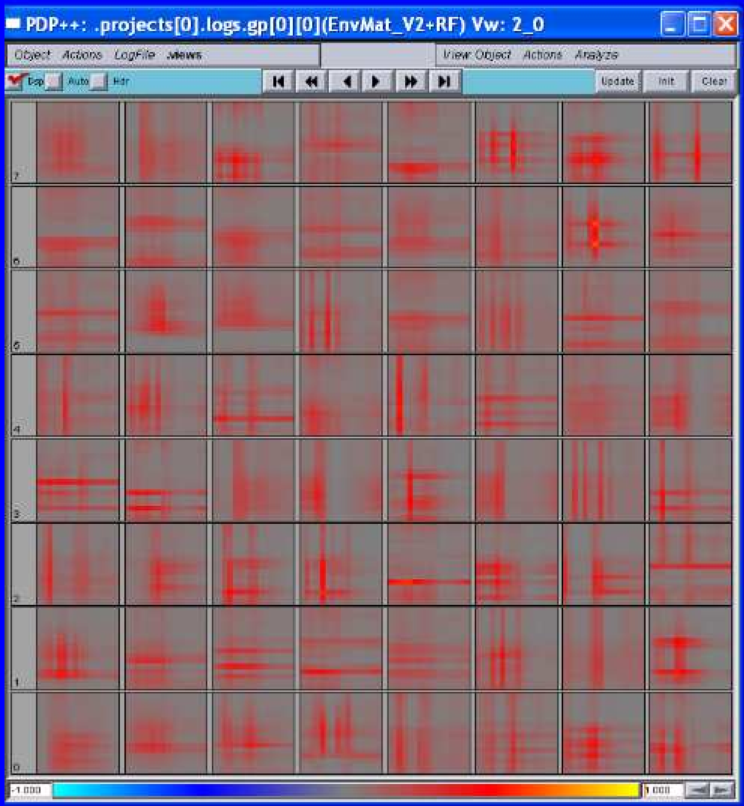


Each object is presented at multiple locations, sizes

Network's job is to activate the appropriate Output unit (0-19) for each object, regardless of location and size

[objrec.proj.gz]

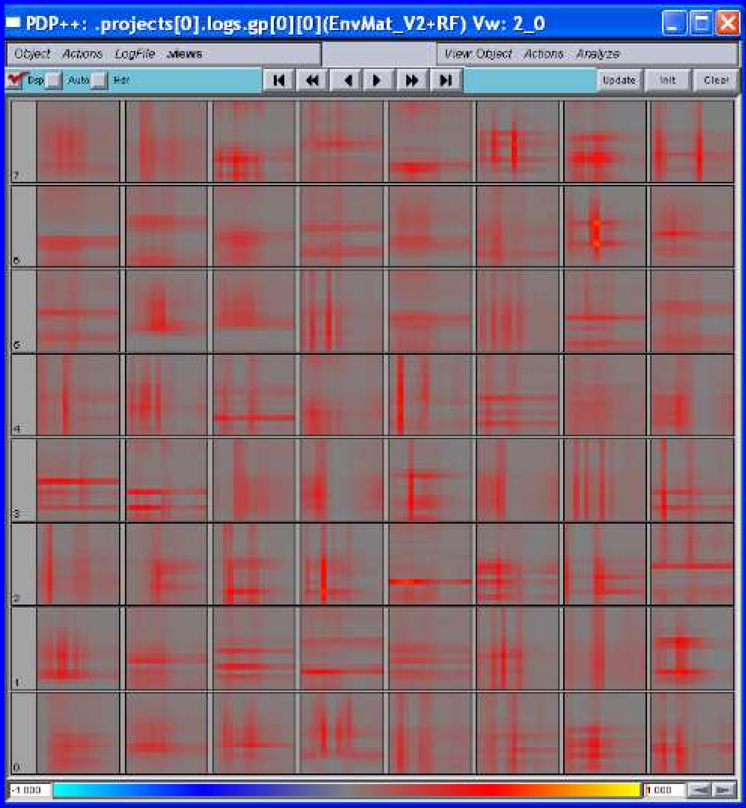
Activation-Based Receptive Fields



- How do we plot receptive fields for V2? Receiving weights show which V1 units a V2 unit responds to but they don't show what **thing in the world** the unit responds to

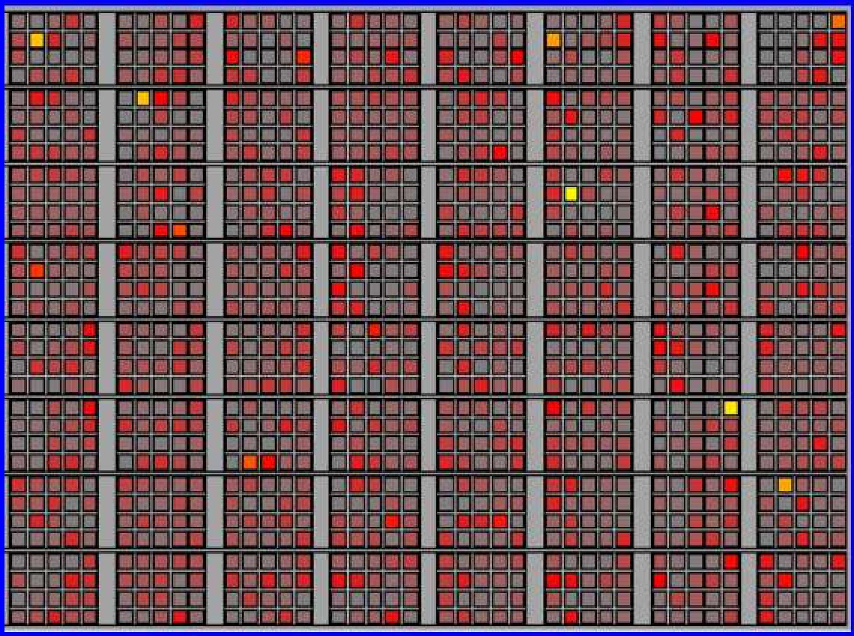
- Solution: Show the networ lots of input patterns. Display a **composite** of all of the input patterns that activate the unit.

V2 Receptive Fields from On-Center Input



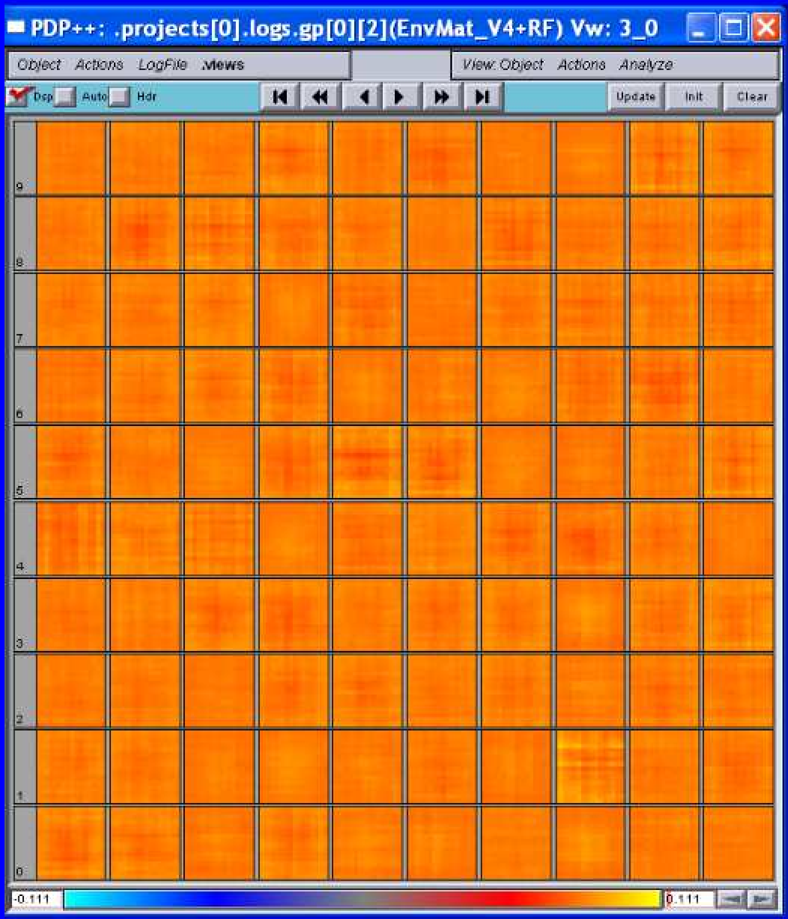
- Some units code for location-specific conjunctions of V1 features (lines)
- This shows up as a sharp receptive field
- Some units code for simple V1 features in a location-invariant way
- This shows up as smeary parallel lines

V2 “Receptive Fields” for Output



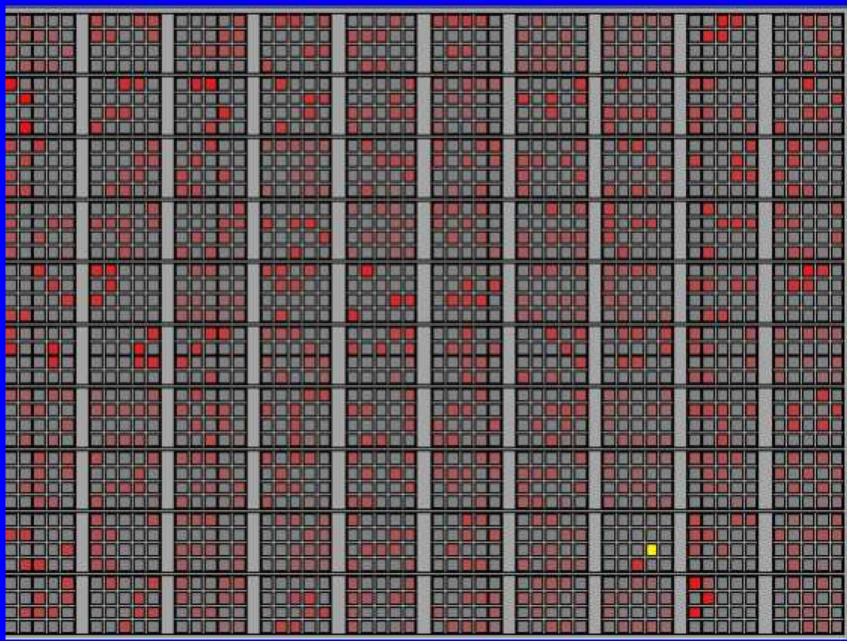
- Present all possible input patterns
- Plot which output units are active when a particular V2 unit is active
- Do V2 units participate in representing multip objects?
- Yes!

V4 Receptive Fields from On-Center Input



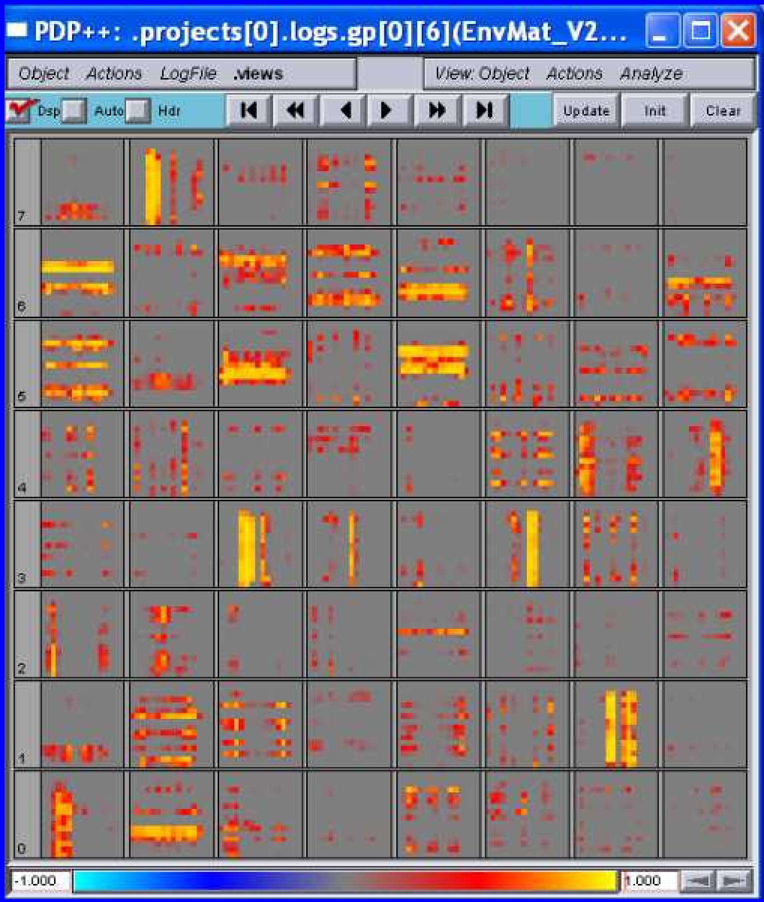
- Results are consistent wi
there being a high degree
spatial invariance (althou
it's hard to say…)

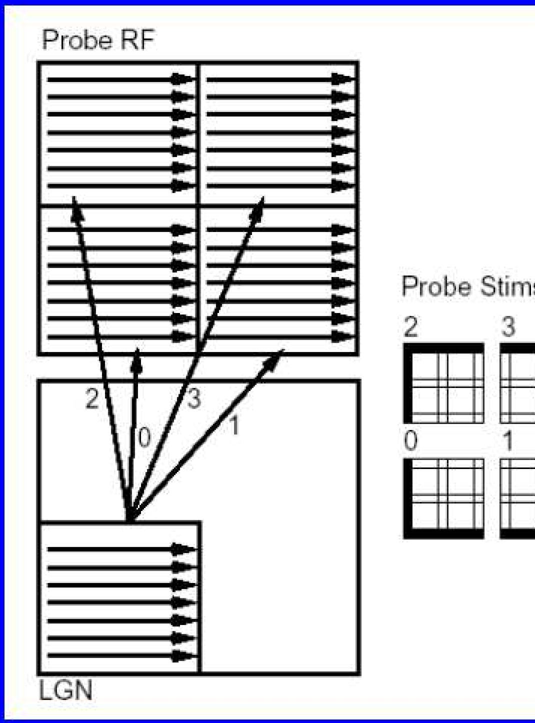
V4 "Receptive Fields" for Output



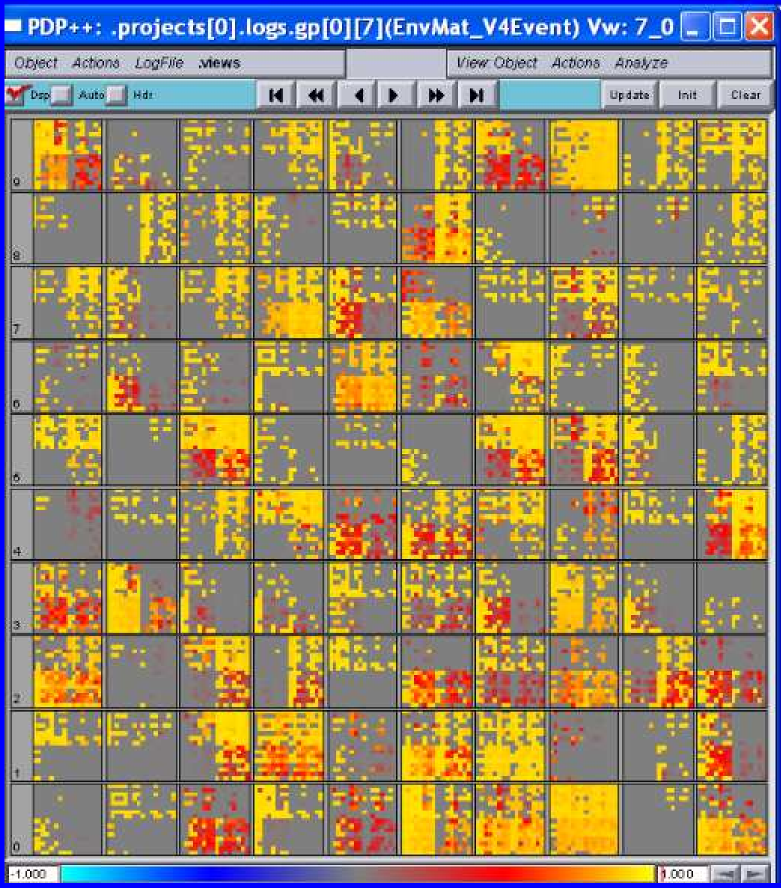
- Present all possible in patterns
- Look at which output units are active when particular V4 unit is active
- V4 units participate in representing multiple objects
- V4 units represent **features**, not whole objects

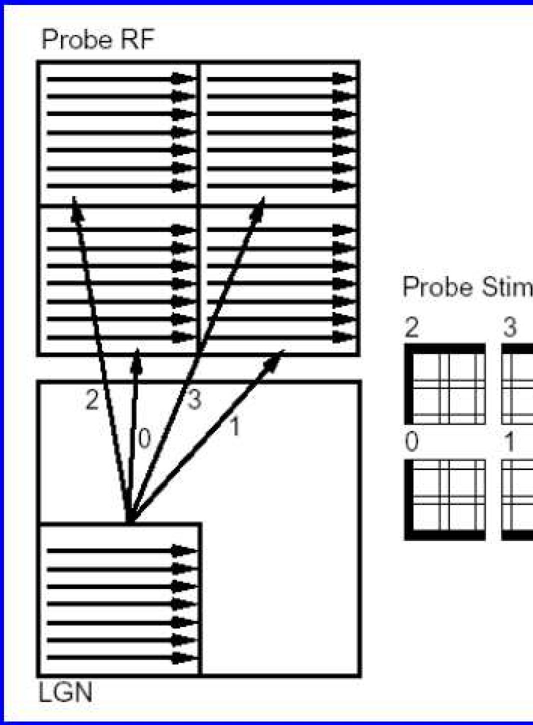
V2 Probe Tests



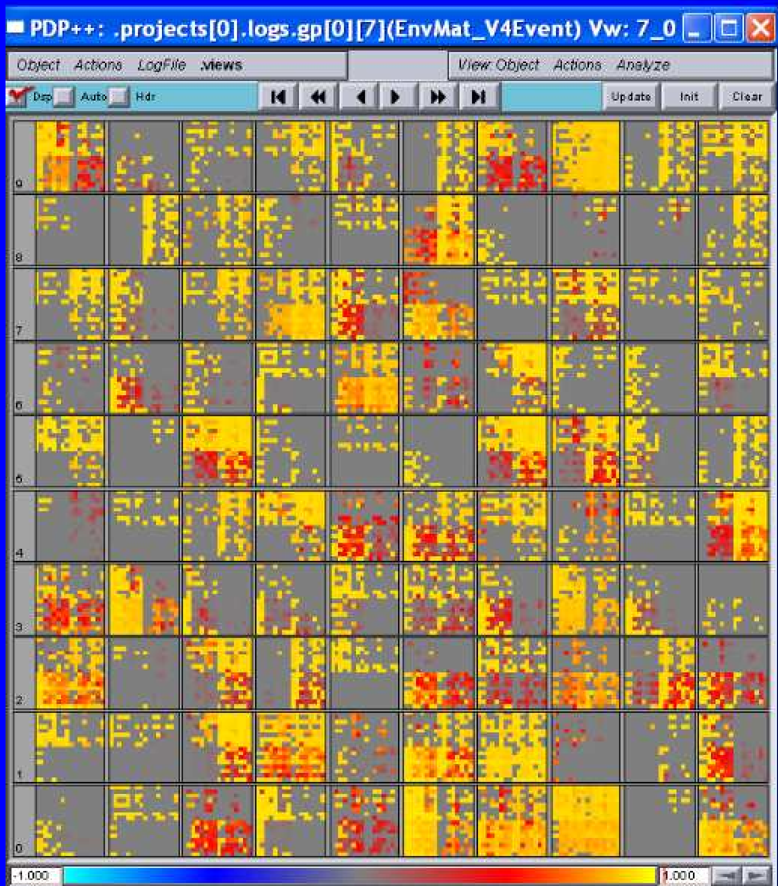


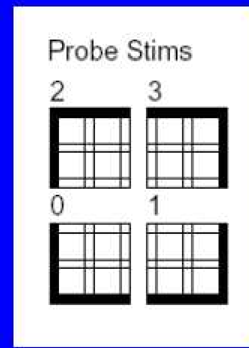
V4 Probe Tests





V4 Probe Tests

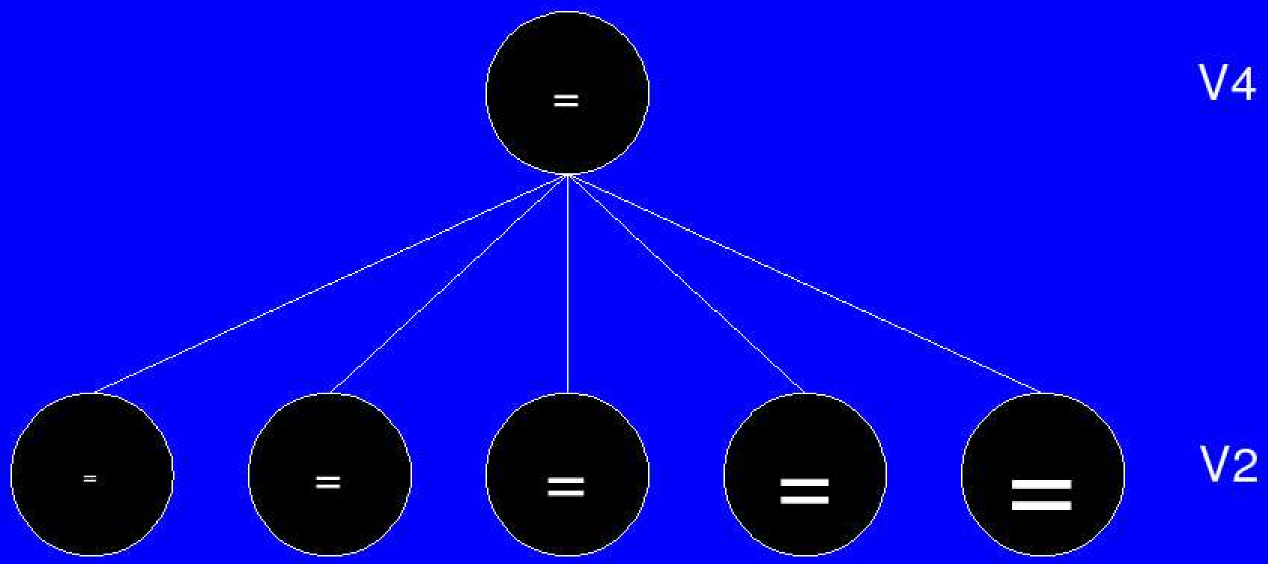




- V4 units represent featur
in a location-invariant wa
- What about size invarian

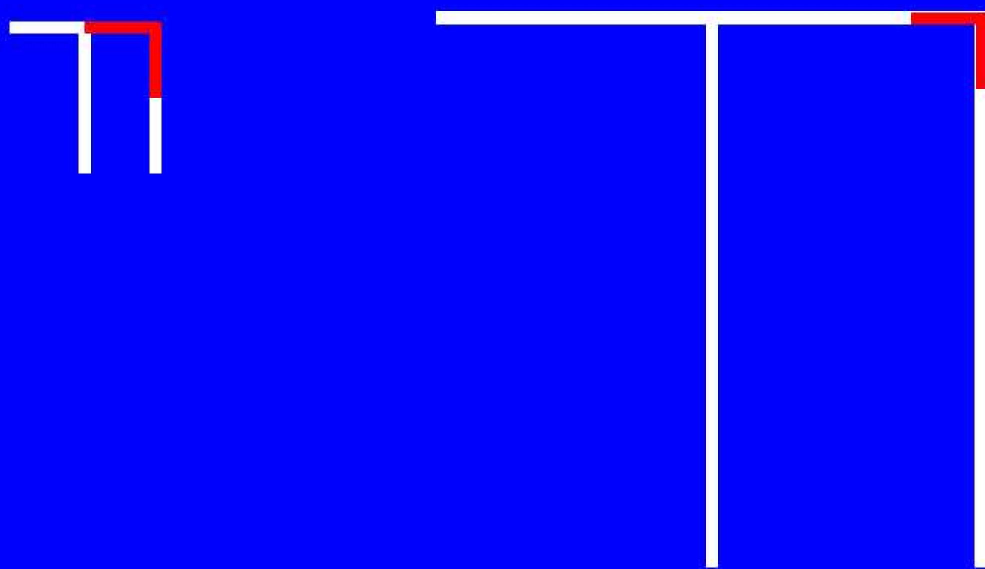
Size Invariance

- One approach to this problem is to have V4 units respond t
  all of the V2 units that represent a feature (regardless of siz

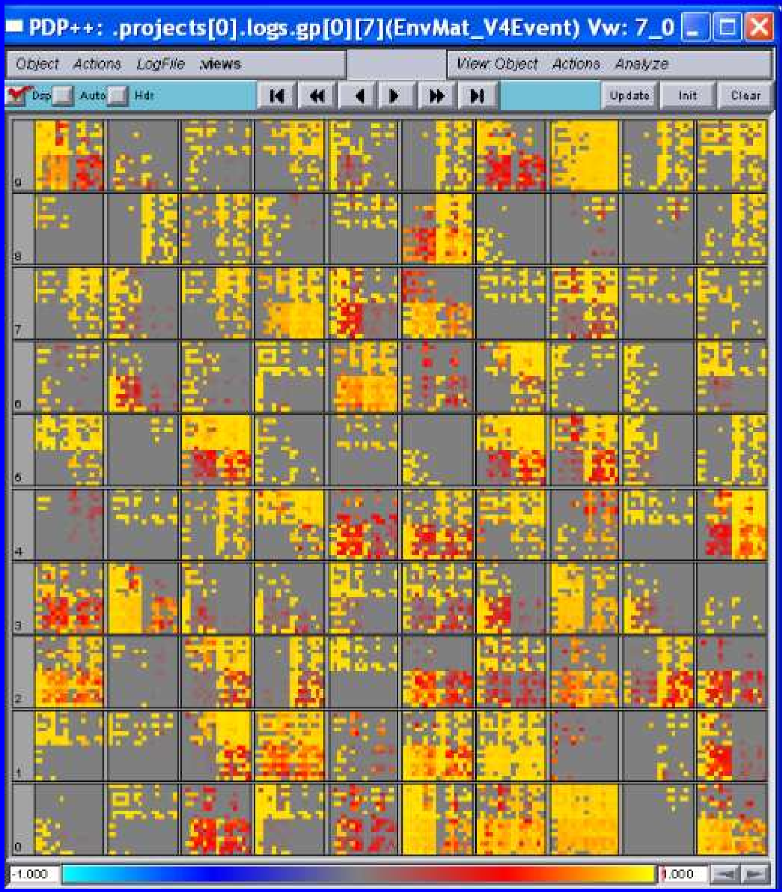


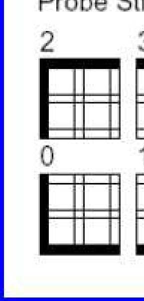
Size Invariance

- Another approach to this problem is to **pick features that a invariant across size transformations**
- e.g., for this set of objects, corners are good!



V4 Probe Tests





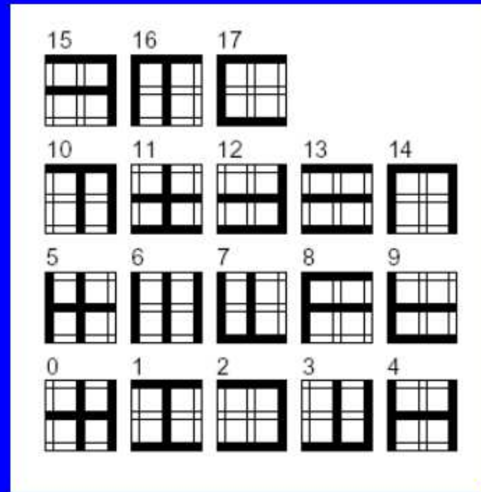
- This diagram shows that units respond to corners (among other things)
- The fact that V4 respond corners helps explain siz invariance...

Generalization

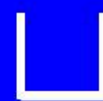
- Can the network generalize to unseen views of studied ob
- In other words: Does training the net to recognize a set of objects in a size/location invariant fashion help it recogni new objects in a size/location invariant fashion?
- Procedure:
  - Take a net trained on 18 objects
  - Train with 2 new objects in only some locations/sizes
  - Test the net with nonstudied "views" (sizes/locations) new objects

Generalization

- Train on these using multiple sizes/locations



- Then train on two new objects (using a limited number of sizes/locations)

 = 18  = 19

- Test on new sizes/locations:

Generalization

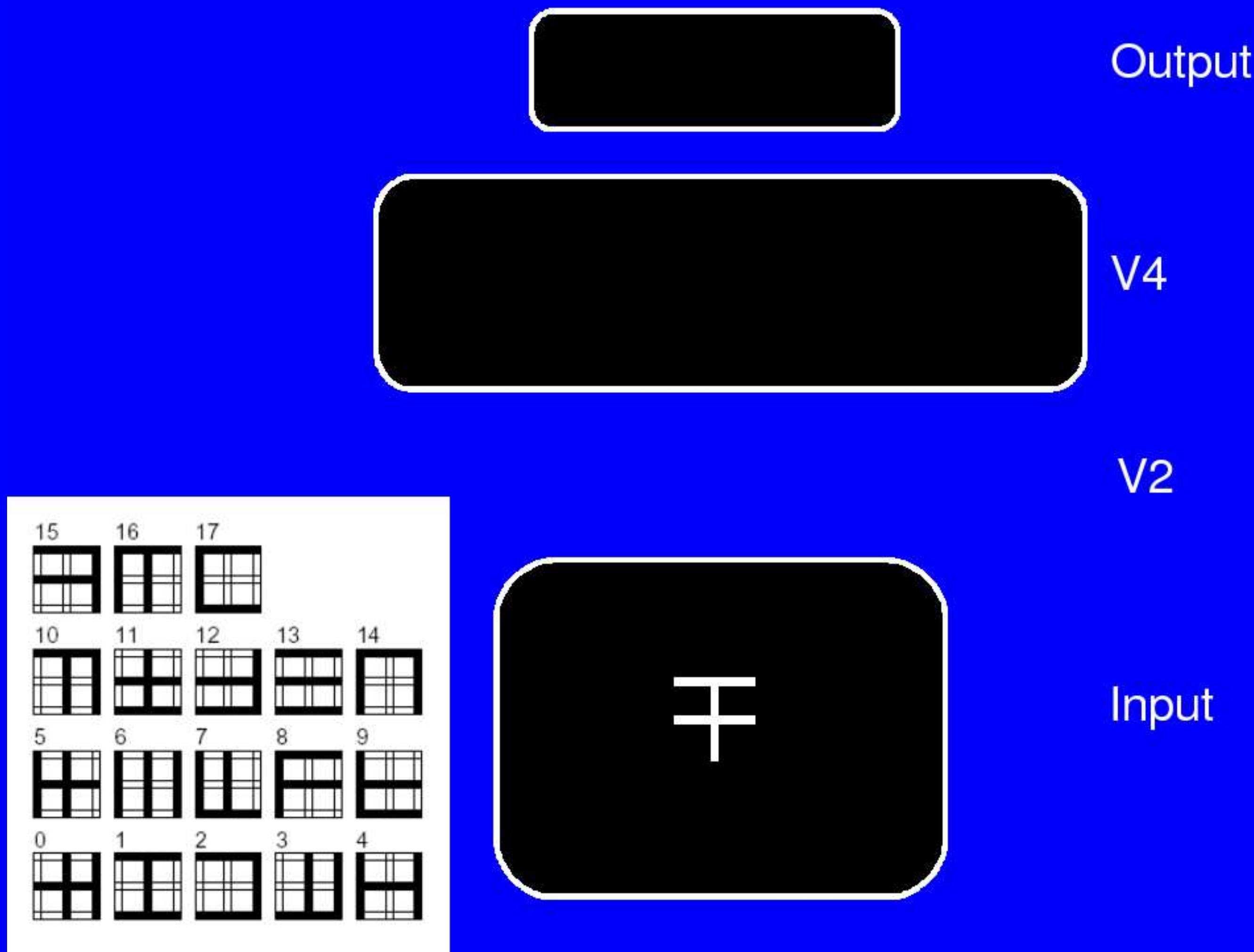
- Can the network generalize to unseen views of studied objects? *yes*
- Approx. 75% correct on novel views following training or
  of possible sizes/locations

Generalization

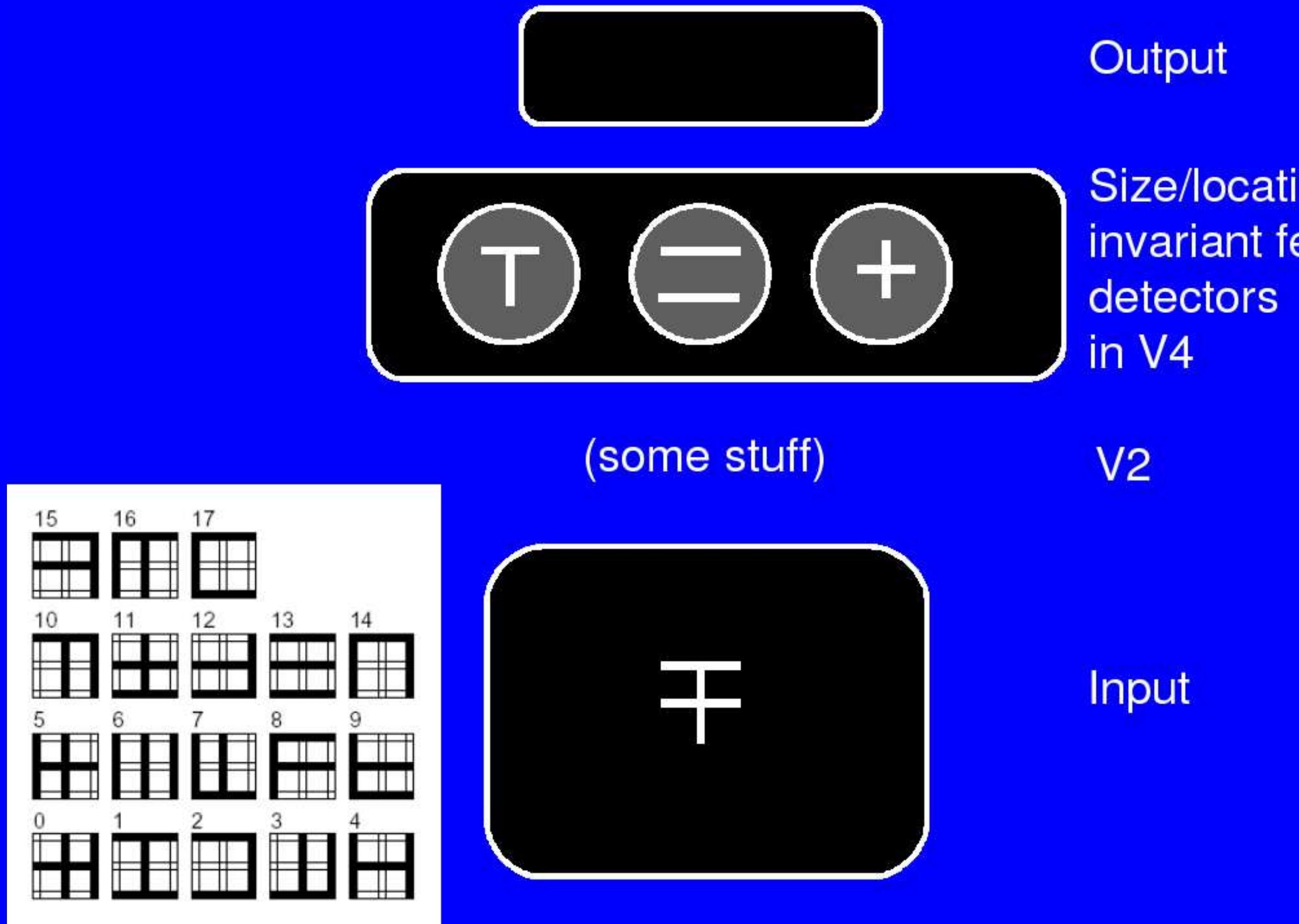
- Can the network generalize to unseen views of studied objects? *yes*
- Approx. 75% correct on novel views following training on
of possible sizes/locations

  Explanation: Distributed representations!
- V4 represents object **features** in a location/size invariant
- Each object activates a distributed pattern of these invaria
feature detectors

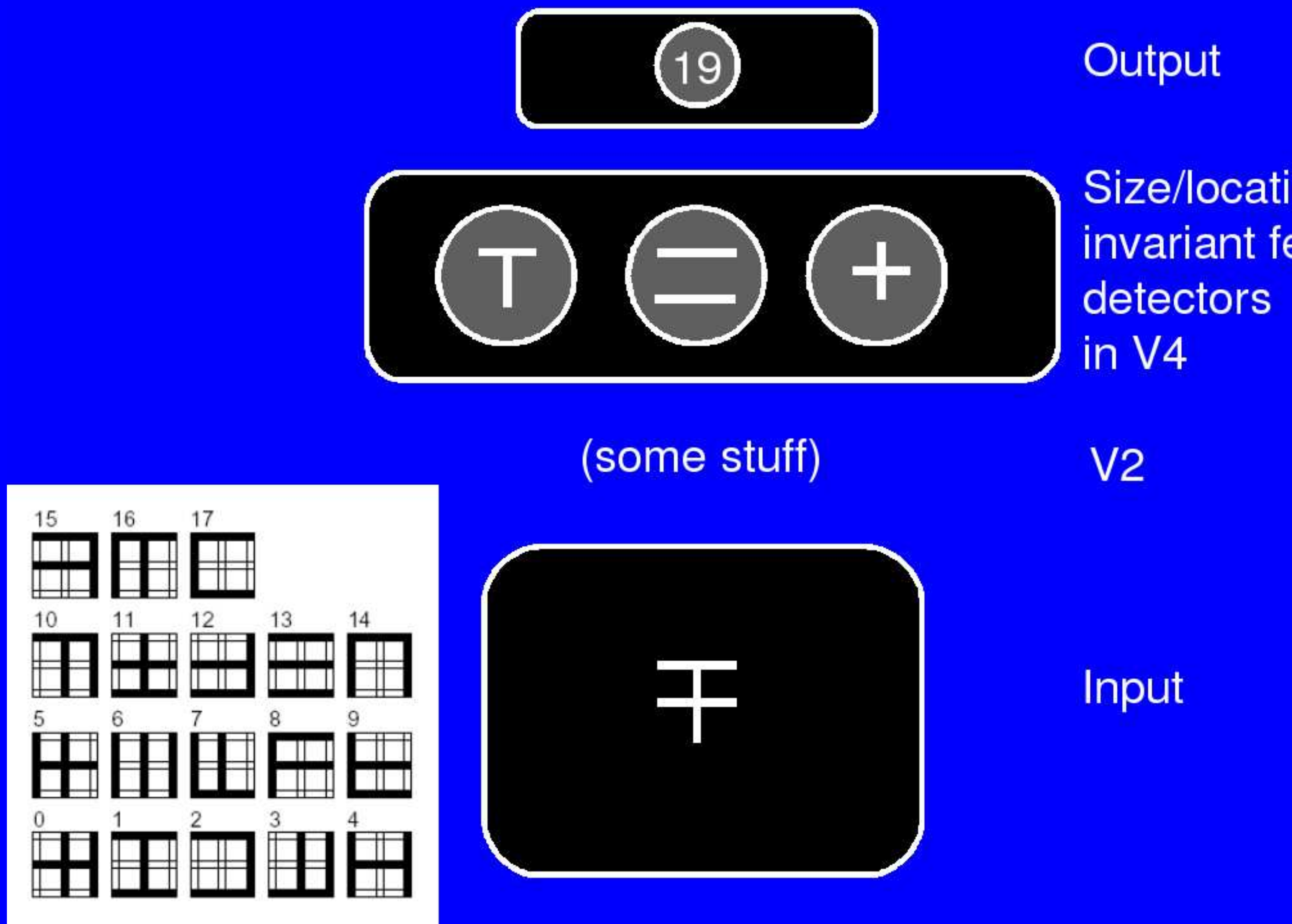
Generalization



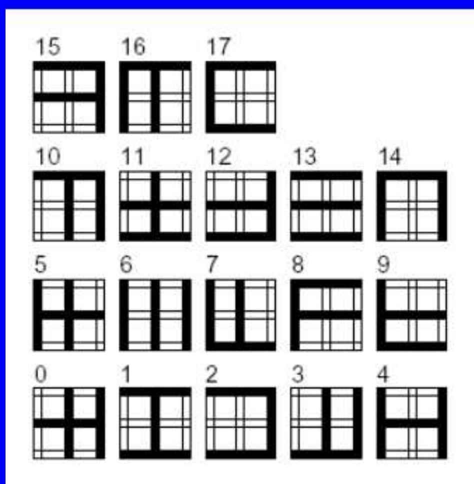
Generalization

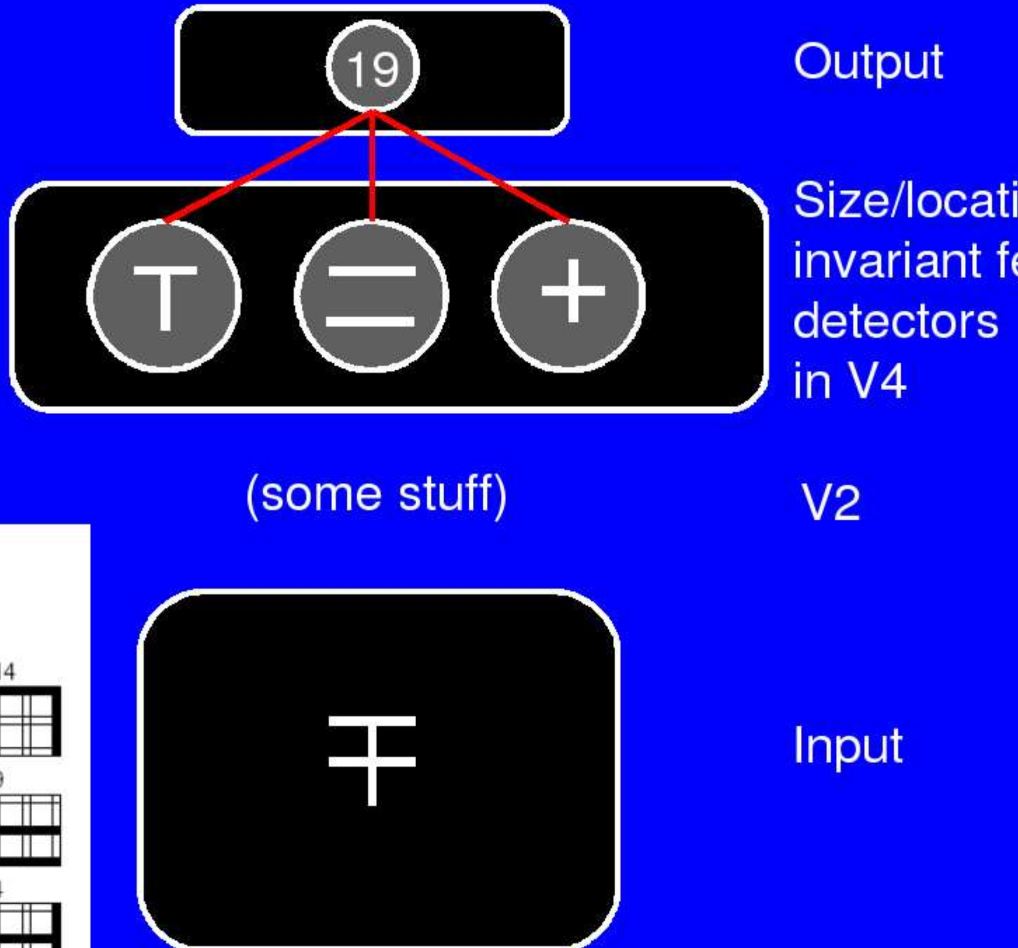


Generalization

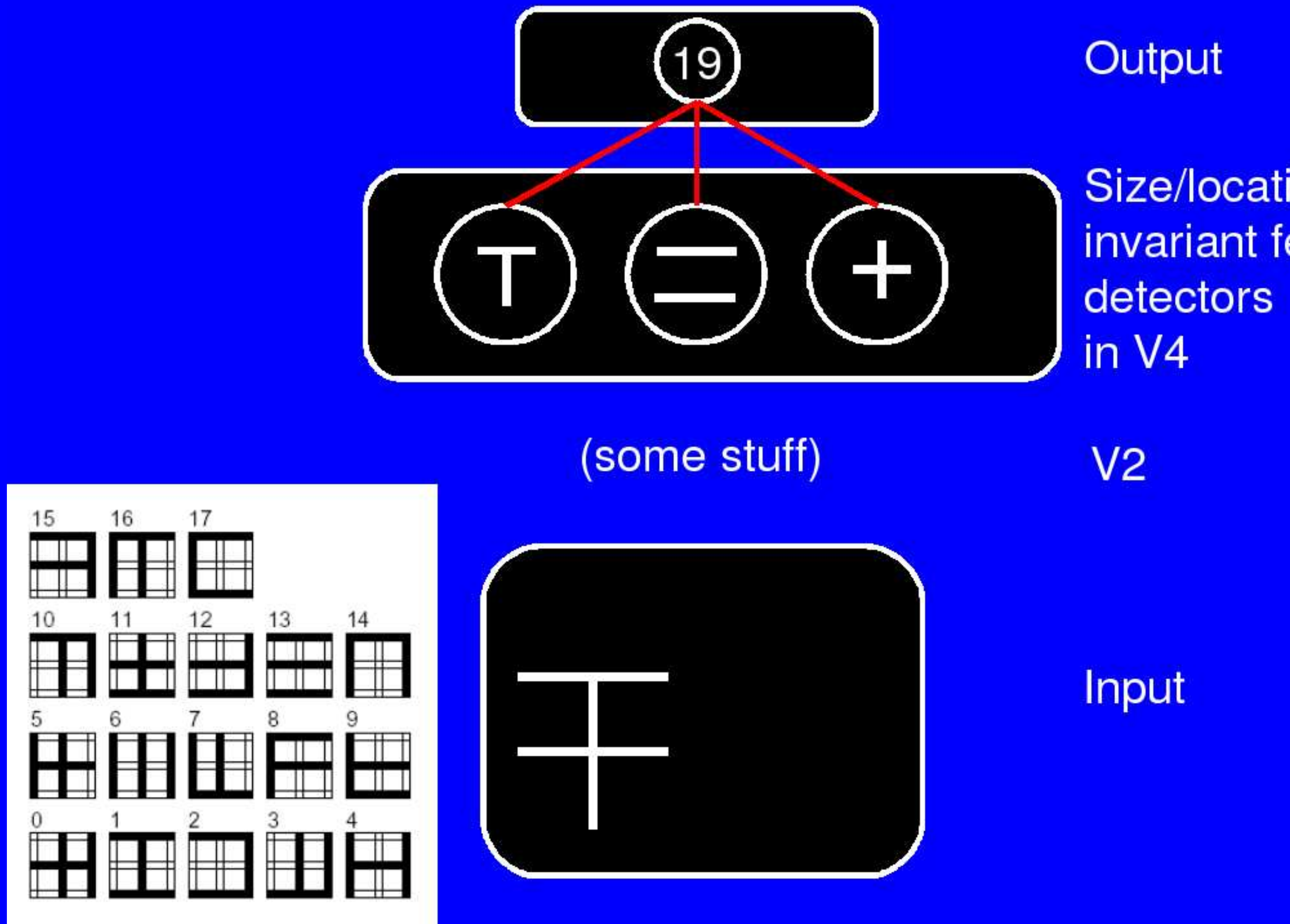


Generalization

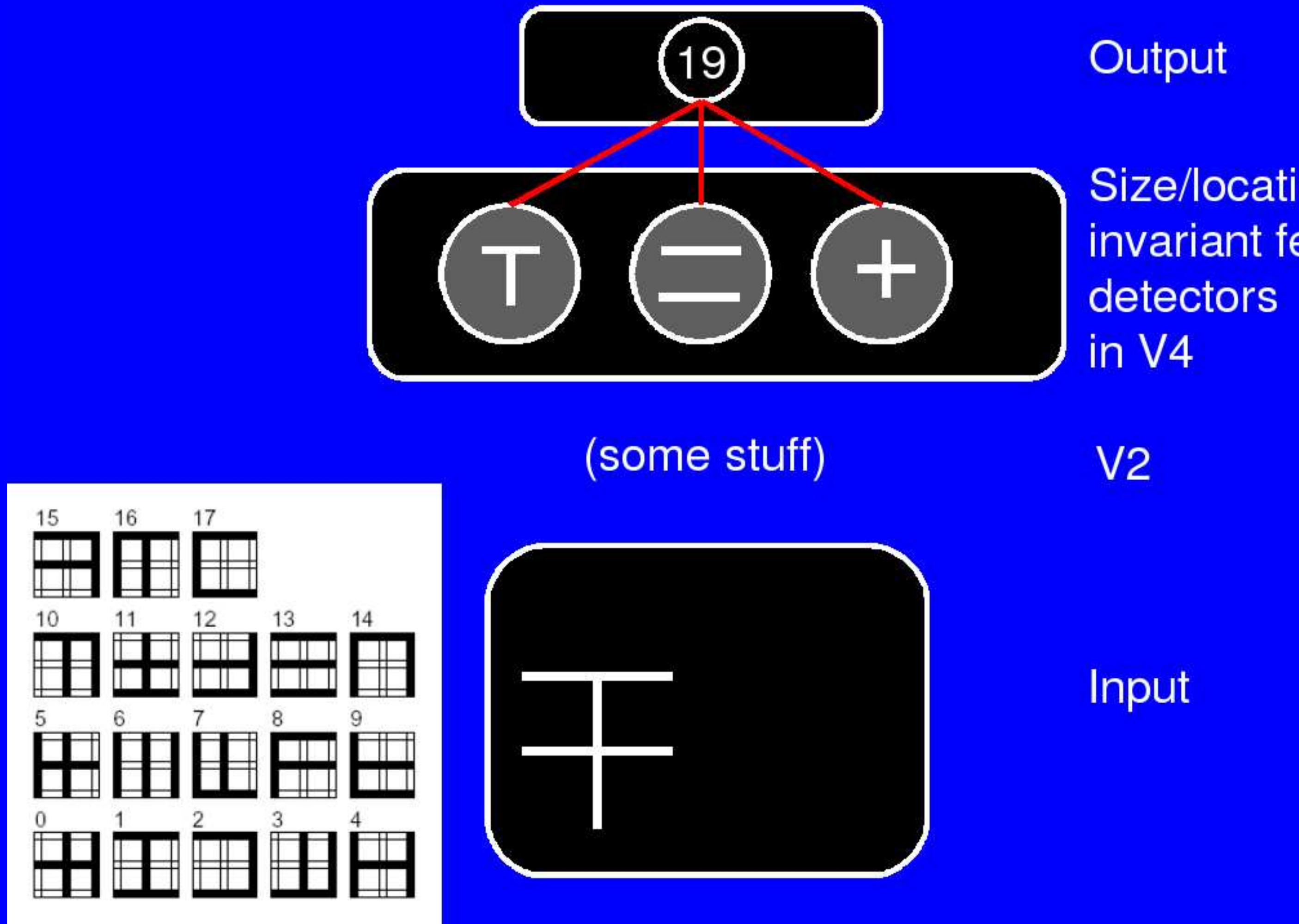




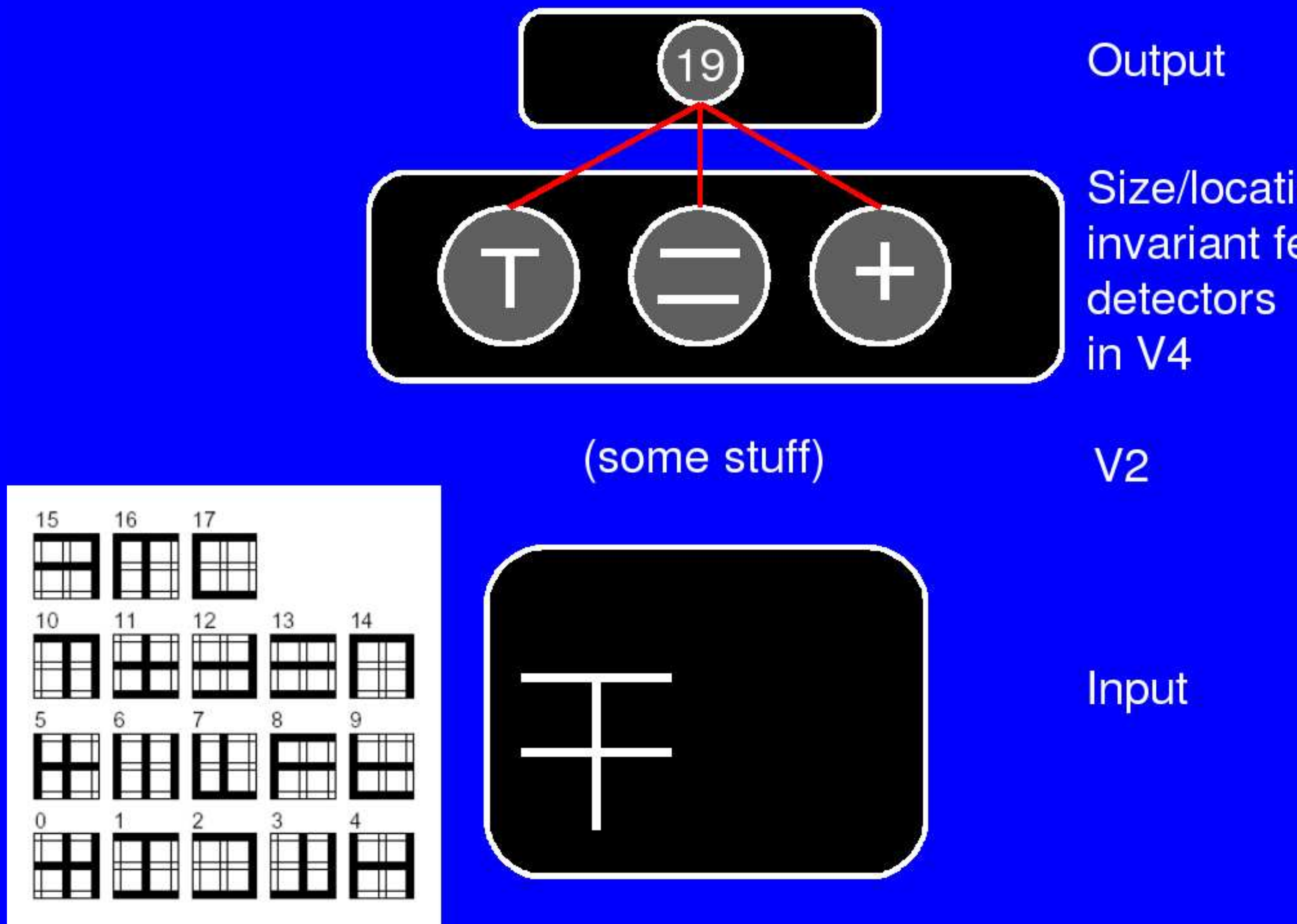
Generalization



Generalization



Generalization



Demo: Recognizing Airplanes!

[hvs.obja1.demo_airplane.mpg]

Perception and Attention

1. Why does primary visual cortex encode oriented bars of l
   Correlational learning based on natural visual scenes.

2. How do we recognize objects (across locations, sizes, rota
   with wildly different retinal images)? *Transformations:*
   increasingly complex featural encodings, increasing levels
   spatial invariance; Distributed representations.

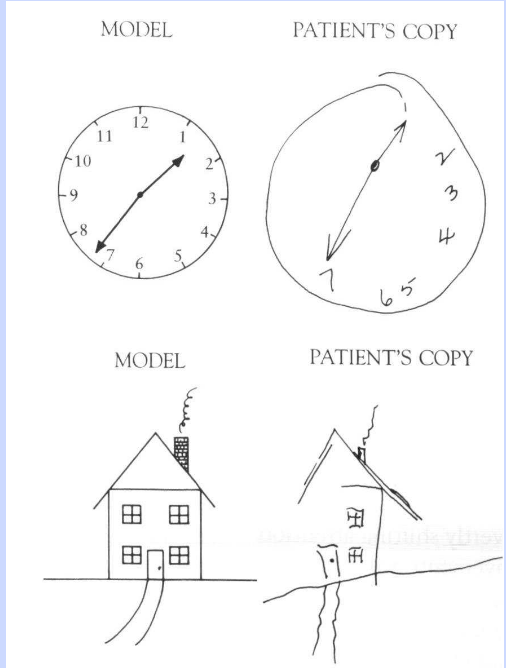
3. Why is visual system split into what/where pathways?

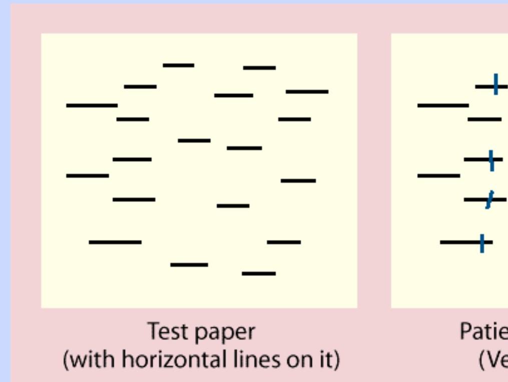
4. Why does parietal damage cause attention problems (neg

- Vanessa: hemispatial neglect: patients have difficulty focu attention in the damaged half of the visual space. Would t be similar to children that have ADHD because they act without thinking, are hyperactive, and have trouble focus they can't sit still, pay attention, or attend to details.

- Anastasia: if attention is considered to be "an emergent property of constraint satisfaction under the limits of inhibition", then what would consciousness/awareness b emergent property of? Text states that "conscious awaren requires an activation pattern that is sufficiently strong to activation elsewhere in the network?" However, it explai neither why it emerges from such activation patterns, nor its function is.

Spatial Attention: Unilateral Neglect



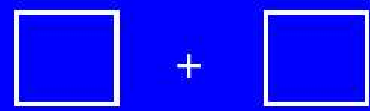


Self portrait, copying, line bisection tasks:
In all cases, patients with parietal/temporal lesions seem to fo
about 1/2 of space! *but they still see it!*

Posner Spatial Cuing Task

Valid cue

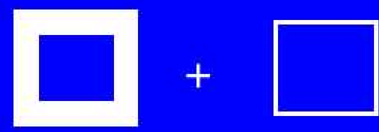
- Fixation



Posner Spatial Cuing Task

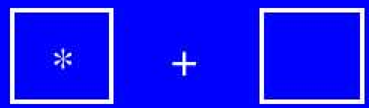
Valid cue

- Cue appears



Posner Spatial Cuing Task

Valid cue

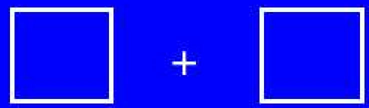


- Target appears, respond with target location

Posner Spatial Cuing Task

Invalid cue

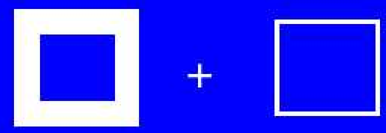
- Fixation



Posner Spatial Cuing Task

Invalid cue

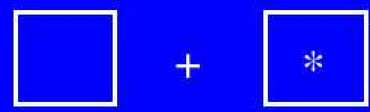
- Cue appears



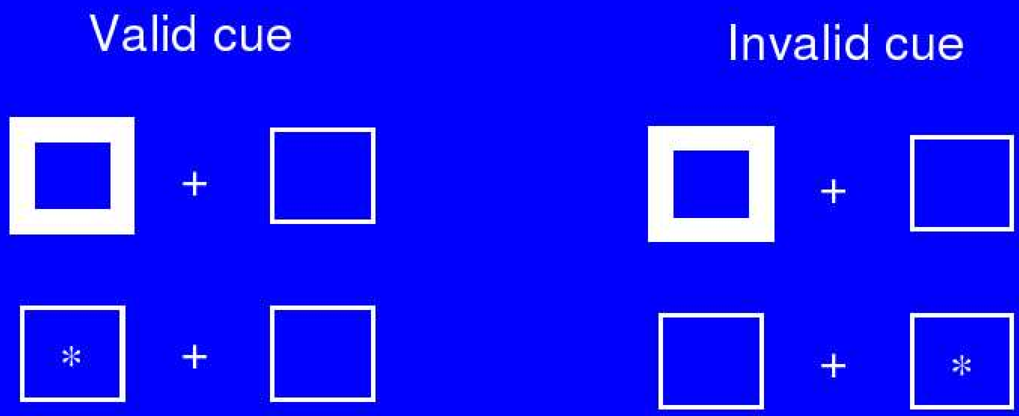
Posner Spatial Cuing Task

Invalid cue

- Target appears, respond with target location



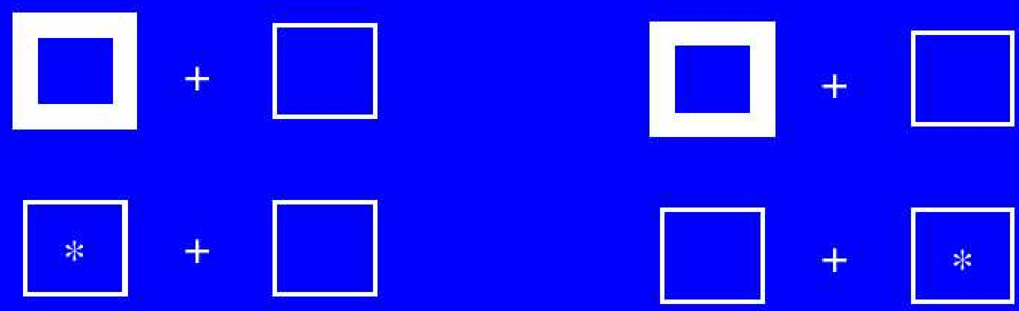
Posner Spatial Cuing Task



Posner Spatial Cuing Task

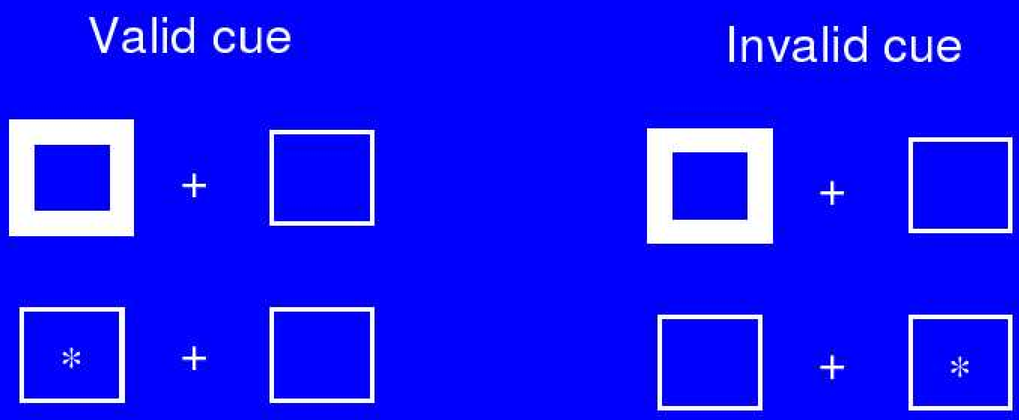
Valid cue

Invalid cue



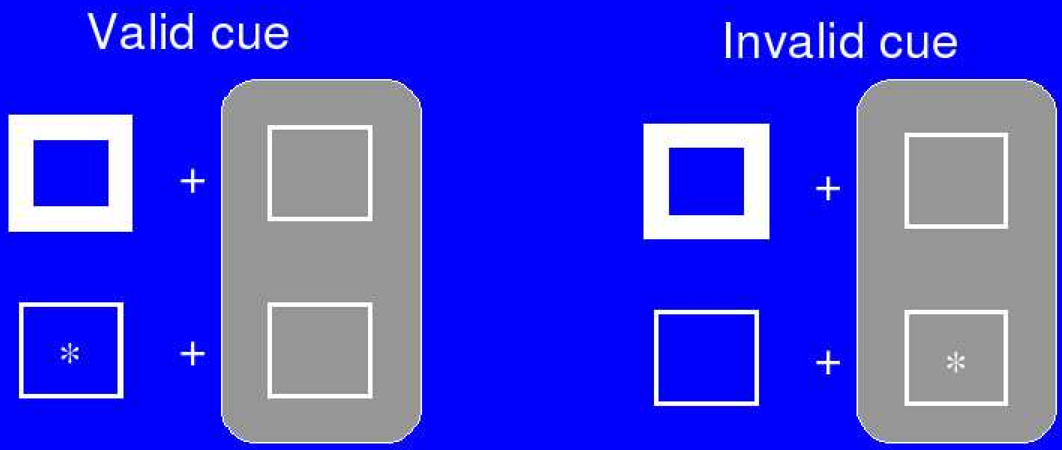
- Valid cues speed up performance (relative to “no cue” condition)
- Invalid cues slow down performance (relative to “no cue” condition)

Effects of Parietal Lesions on Posner Task



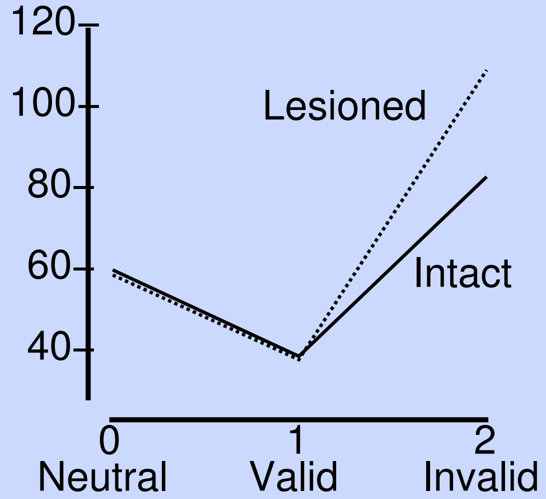
- Large, unilateral parietal lesions result in **neglect** of the opposite (contralateral) side of space
- Subjects do not respond to targets in the neglected hemifield
- What about smaller, unilateral parietal lesions?

Effects of Parietal Lesions on Posner Task



- Say that you have a small, left parietal lesion, so the right side is affected
- Run the Posner task with cues in the ipsilateral (left) side of space

Effects of Parietal Lesions on Posner Task

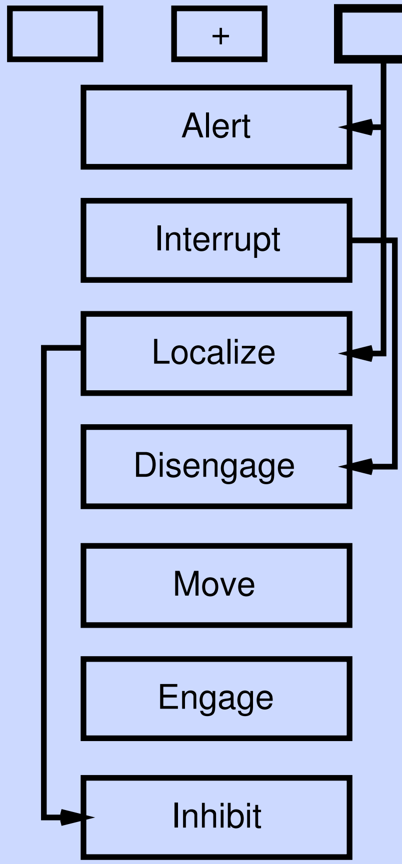


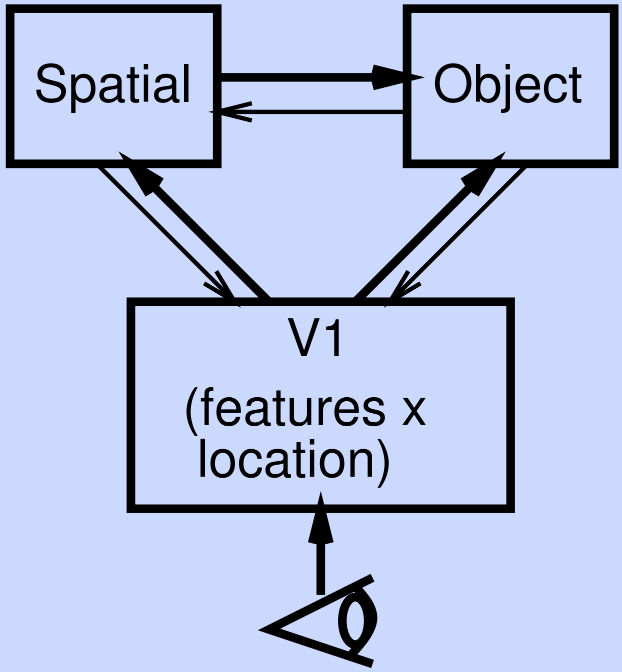
- Patients perform normally in the "neutral" (no cue) condi *regardless* of where the target is presented

- Patients benefit just as much as controls from valid cues

- Patients are hurt more than controls by invalid cues

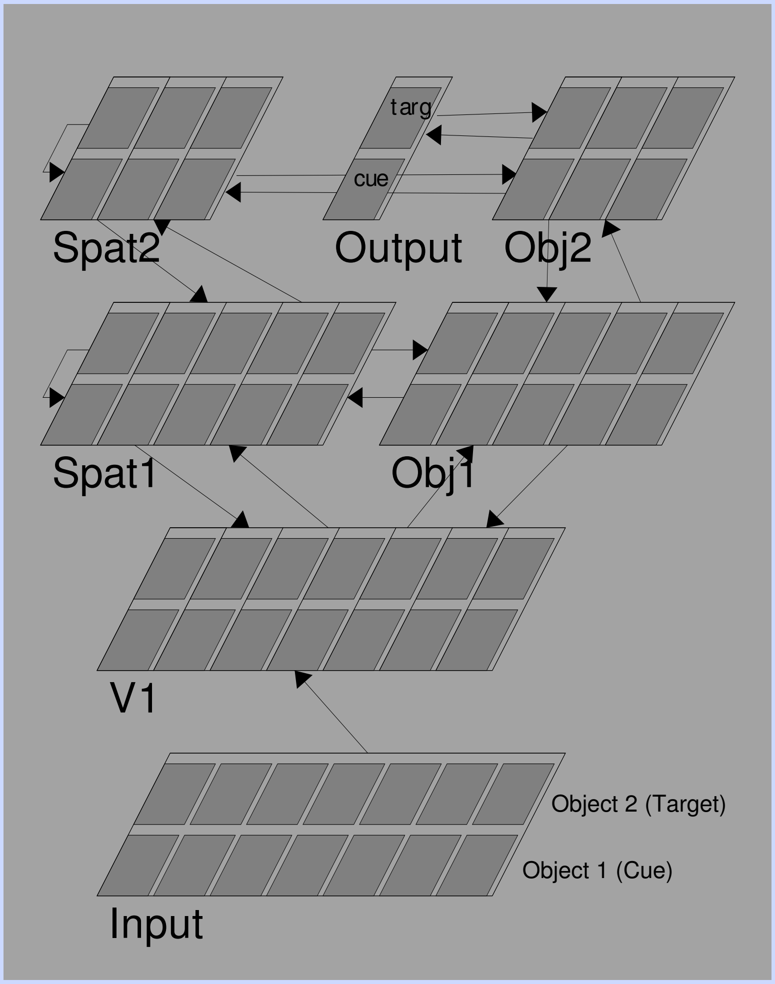
Possible Models





Attention *emerges* from bidirectional constraint satisfaction & inhibitory competition.

Simple Model



[attn_simple.proj.gz]

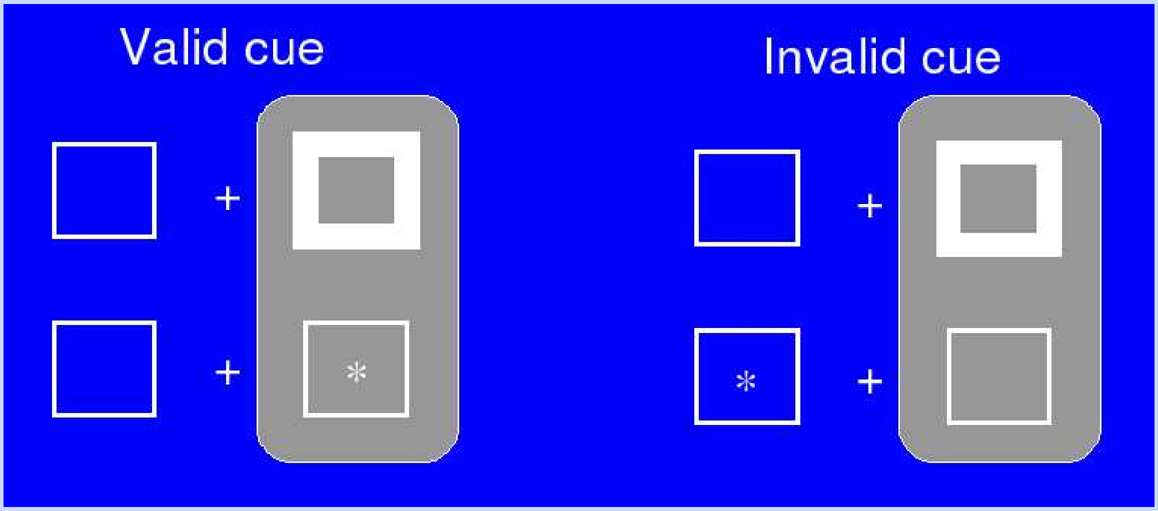
Posner Task Data

	Valid	Invalid	Diff
Adult Normal	350	390	40
Elderly Normal	540	600	60
Patients	640	760	120
Elderly normalized (*.65)	350	390	40
Patients normalized (*.55)	350	418	68

Posner Task Sims

- The model explains the basic finding that valid cues speed target processing, while invalid cues hurt
- Also explains finding that patients with small unilateral parietal lesions benefit normally from valid cues in ipsilat field but are disproportionately hurt by invalid cues.
- No need to posit “disengage” module!
- Also explains finding of **neglect** of contralateral visual fie after large, unilateral parietal lesions when some stimulus present in ipsilateral field (“extinction”)

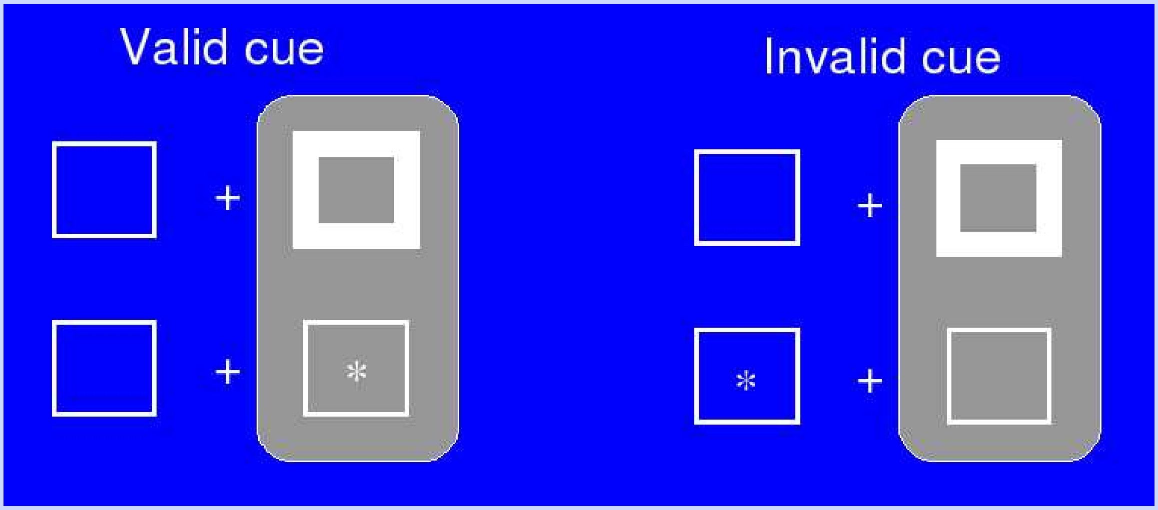
More Posner Lesion Fun



- Returning to patient with left parietal lesion...
- What happens if cues are presented in **contralateral** (affe
hemifield?

[attn_simple.proj.gz]

More Posner Lesion Fun



Returning to patient with left parietal lesion...

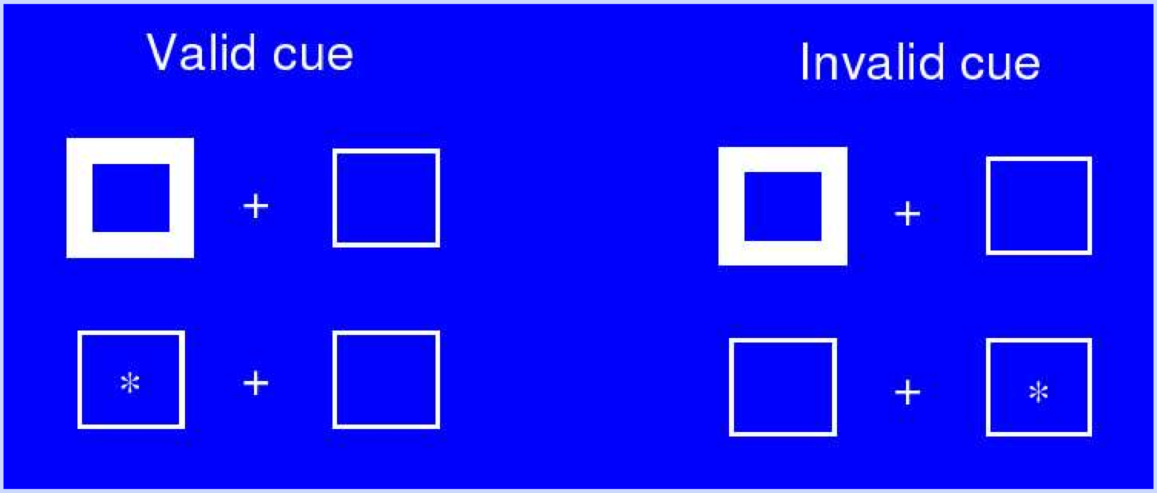
- What happens if cues are presented in **contralateral** (affe
  hemifield?

  Predictions:

- Smaller benefit for valid cues

- Patients should be hurt less than controls by invalid cues.

Inhibition of Return



- Typically, target detection is faster on trials with valid vs invalid cues
- **However**, if the cue is presented for a longer time (eg. 500
performance is faster on ***invalid*** vs valid trials
- Can explain in terms of **accommodation** (neural fatigue)

[attn_simple.proj.gz]

Simple model: too simple?

- Has unique one-to-one mappings between low-level visu
features and object representations (not realistic)

- Does not address issue of spatial attention when trying to
perceive multiple objects simultaneously

Simple model: too simple?

- Has unique one-to-one mappings between low-level visu features and object representations (not realistic)

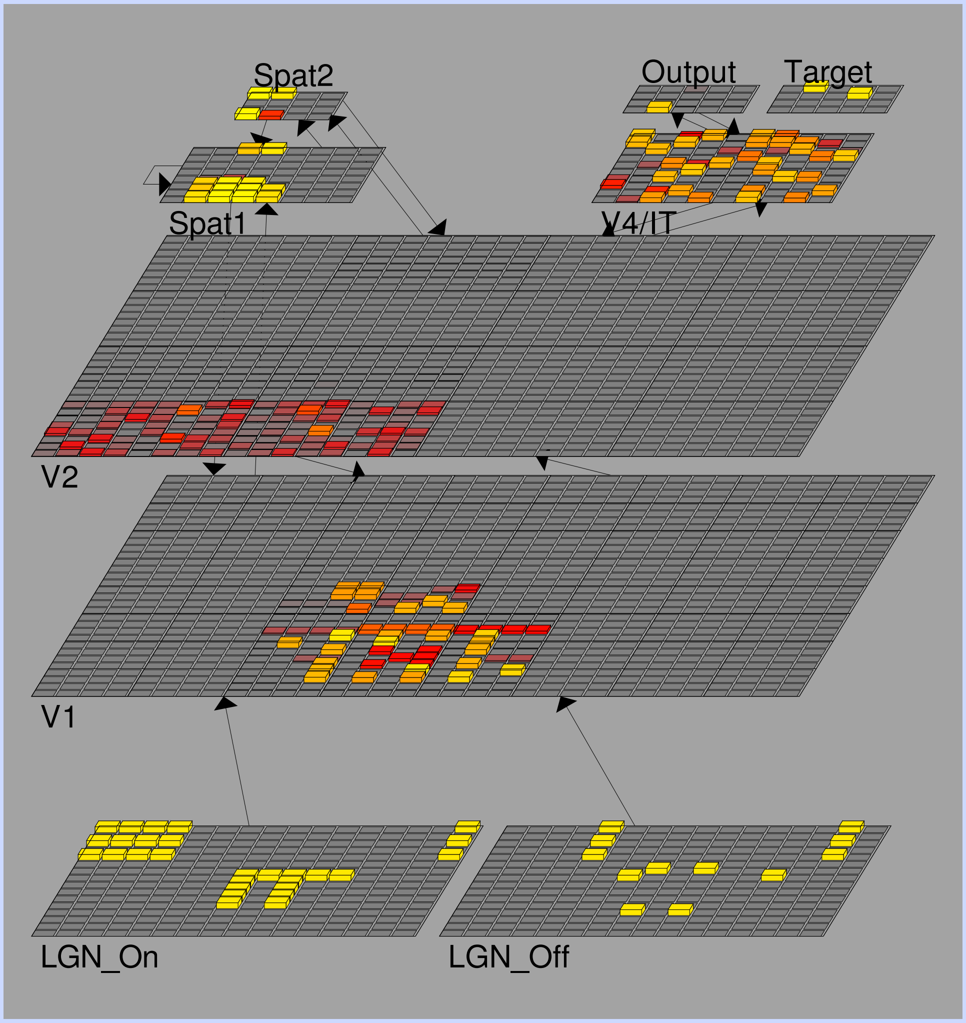
- Does not address issue of spatial attention when trying to perceive multiple objects simultaneously

- "Complex" model combines more realistic model of objec recognition (starting from LGN) with simple attention mo

  → Can use spatial attention to restrict object processing pathway to one object at a time, enabling it to sequentiall process multiple objects.

- Lesions of entire spatial pathway cause *simultanagnosia*: inability to concurrently recognize two objects

Complex Model



[objrec_multiobj.proj.gz]

Perception and Attention

1. Why does primary visual cortex encode oriented bars of l
   Correlational learning based on natural visual scenes.

2. How do we recognize objects (across locations, sizes, rota
   with wildly different retinal images)? *Transformations:*
   increasingly complex featural encodings, increasing level
   spatial invariance; Distributed representations.

3. Why is visual system split into what/where pathways?
   Transformations: emphasizing and collapsing across diffe
   distinctions

4. Why does parietal damage cause attention problems (neg
   Attention as an emergent property of competition

General Issues in Attention

Attention:

- Prioritizes processing.
- Coordinates processing across different areas.
- Solves binding problems via coordination.