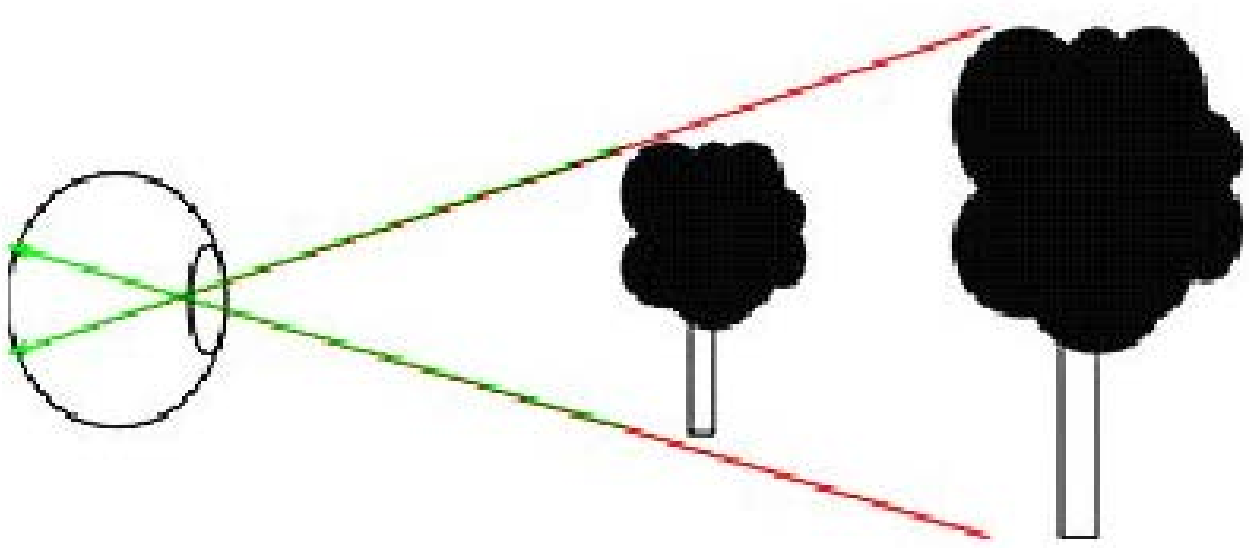
Learning to See

Cognitive and Computational Perspectives

amitabha mukerjee

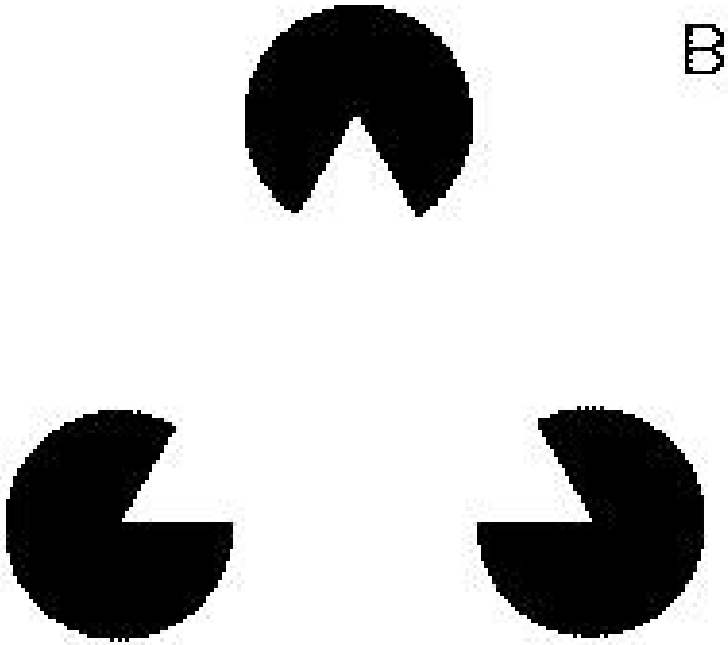
www .cse .iitk. ac.in/ ~amit

Indeterminacy in Vision

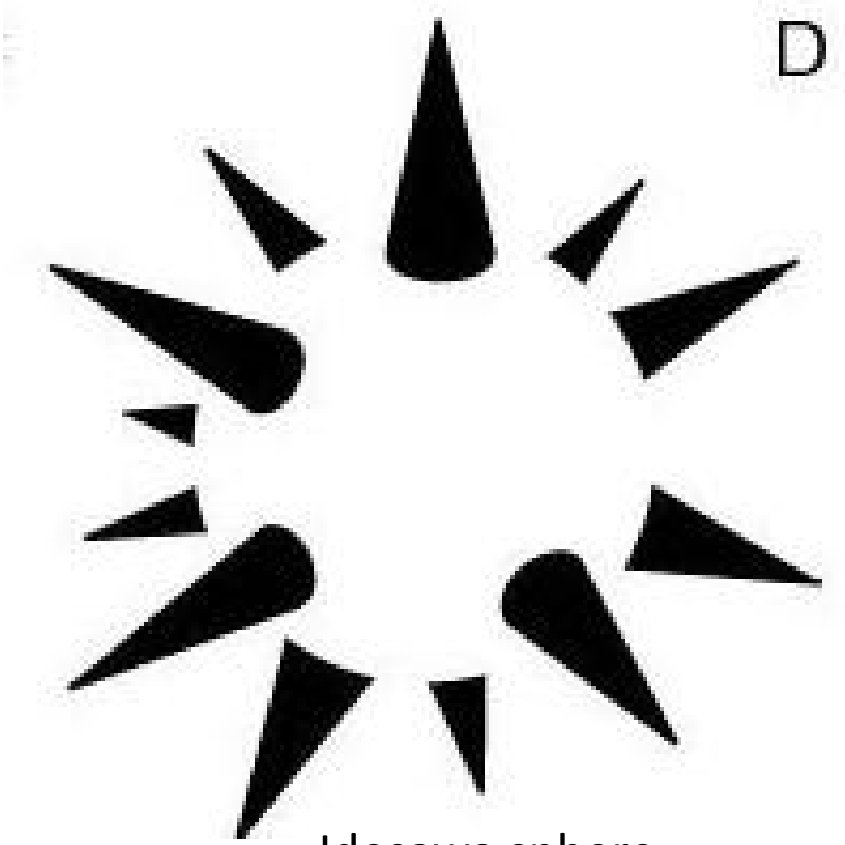


Seeing Objects

Perception of wholes / Good Continuation



kanizsa triangle



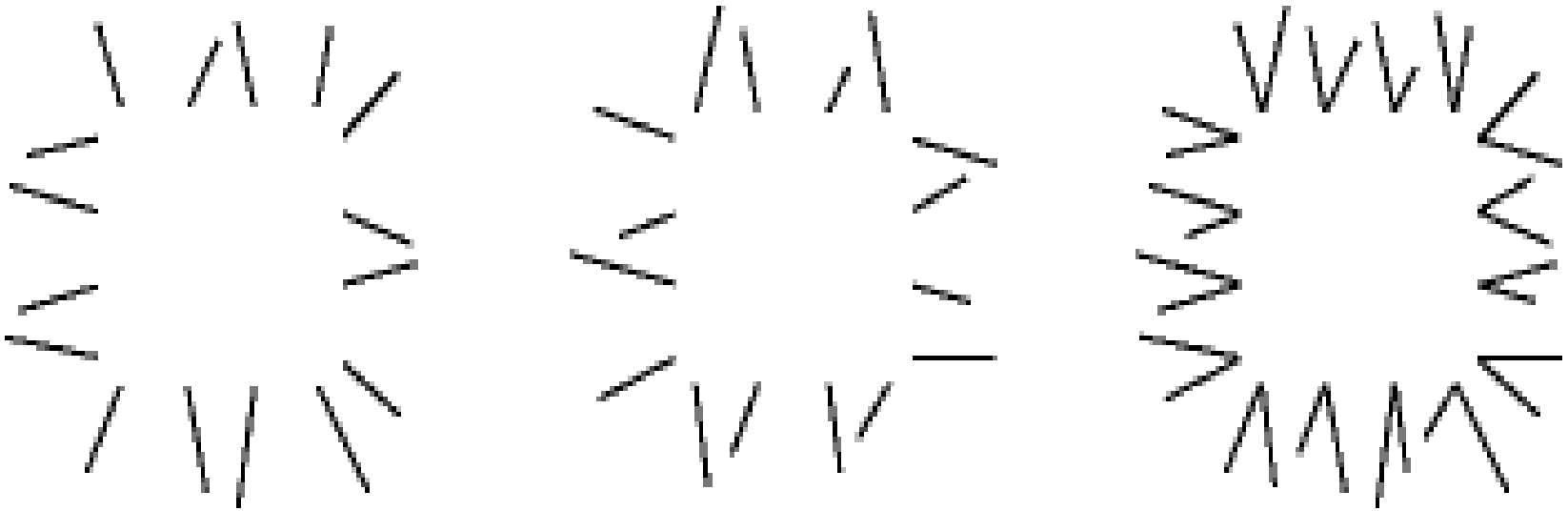
Idesawa sphere

Manifolds in vision

images: 100 x 100 pixels

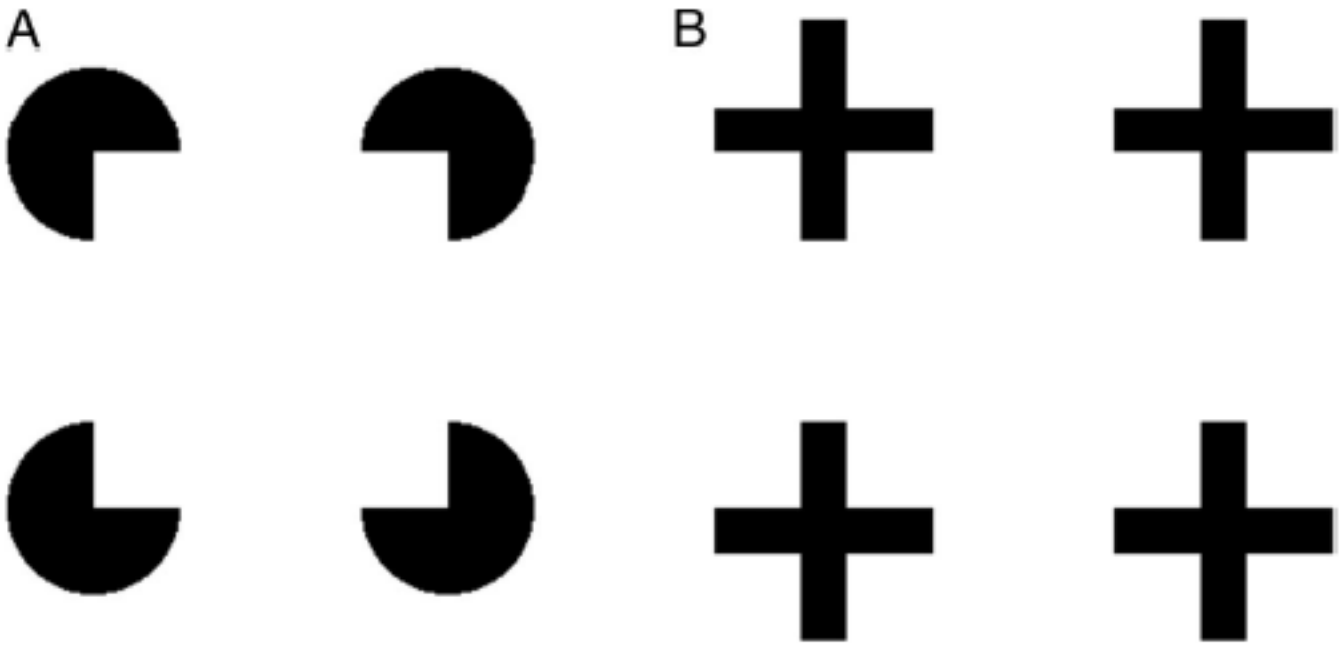
from: A. Efros, original images from hormel corp.

Interference



[donald hoffman] : Visual intelligence 2000

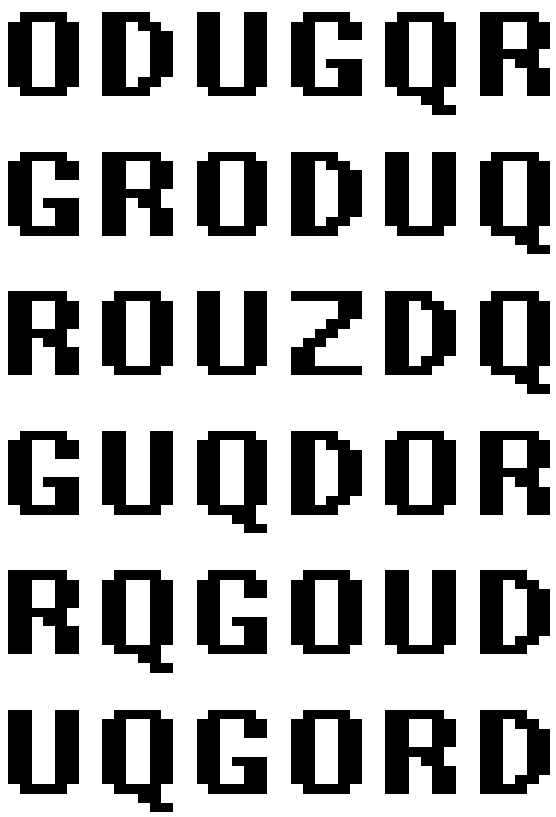
Interference



[wagemans elder 12] : *a century of gestalt perceptual grouping*

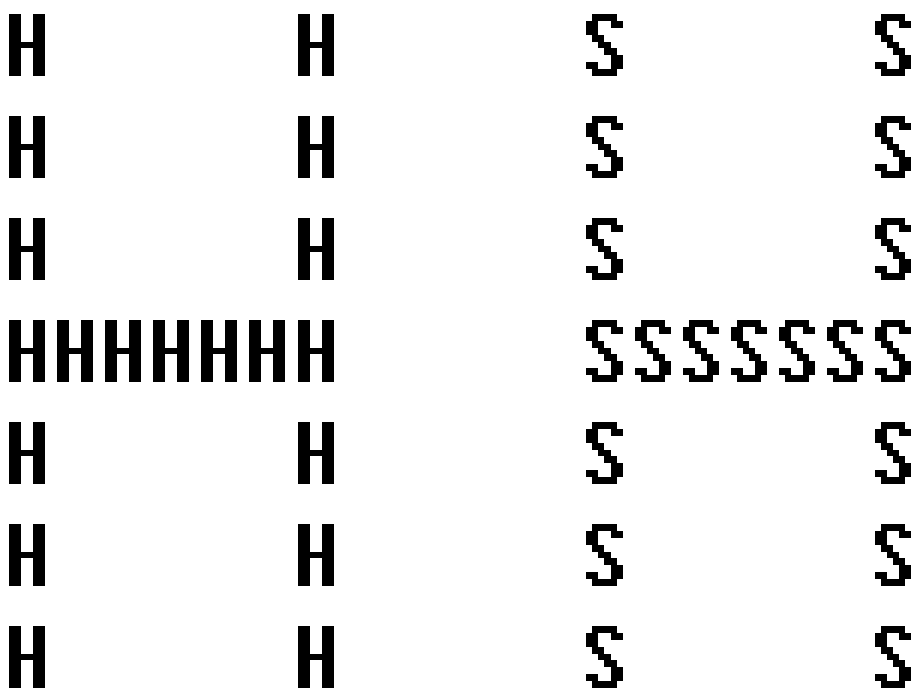
Theories of Perception

Feature Theory



[Selfridge 1959]

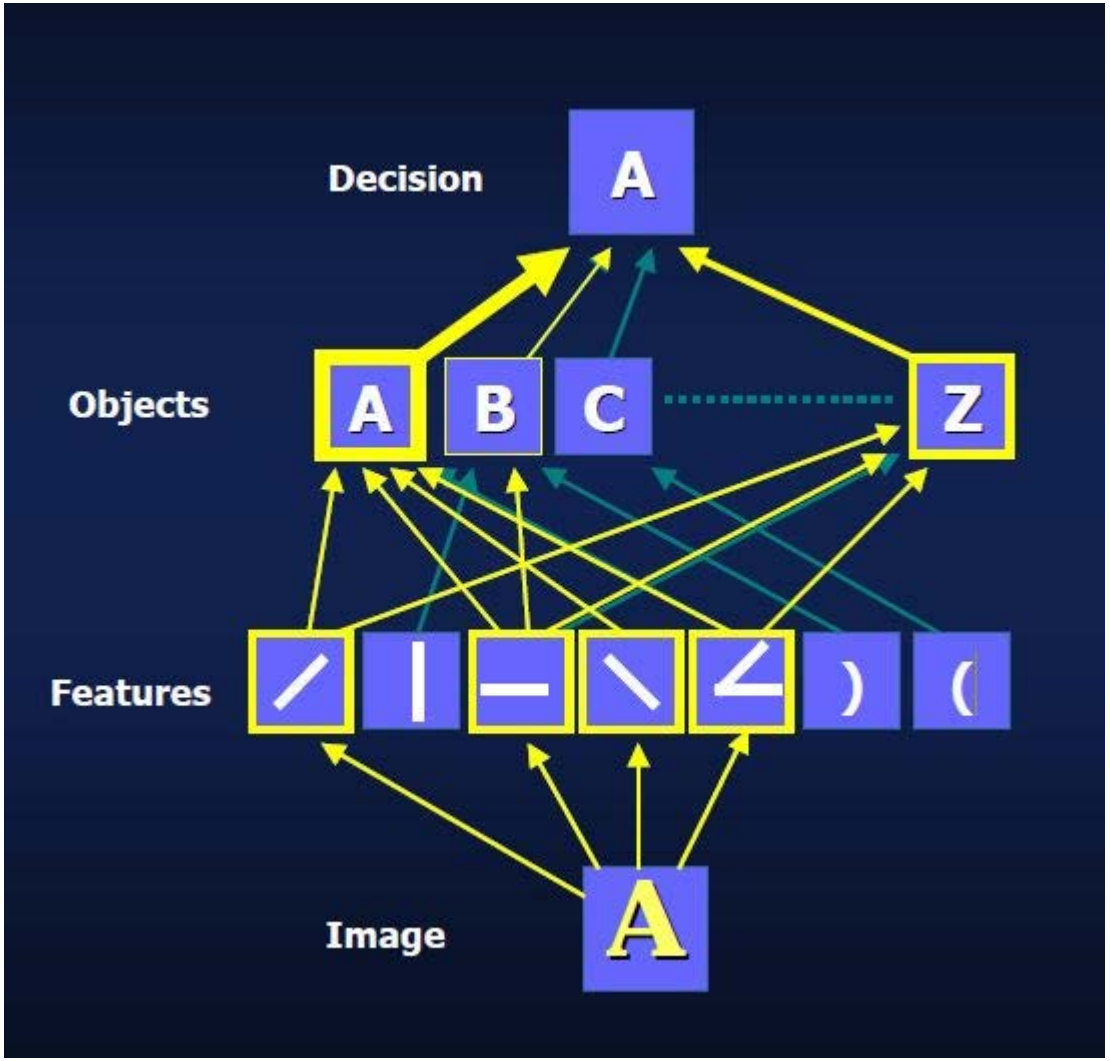
Local vs Global Perception



"letter S" at right
Takes longer to
respond to.

[Navon 1977]

Feature Integration

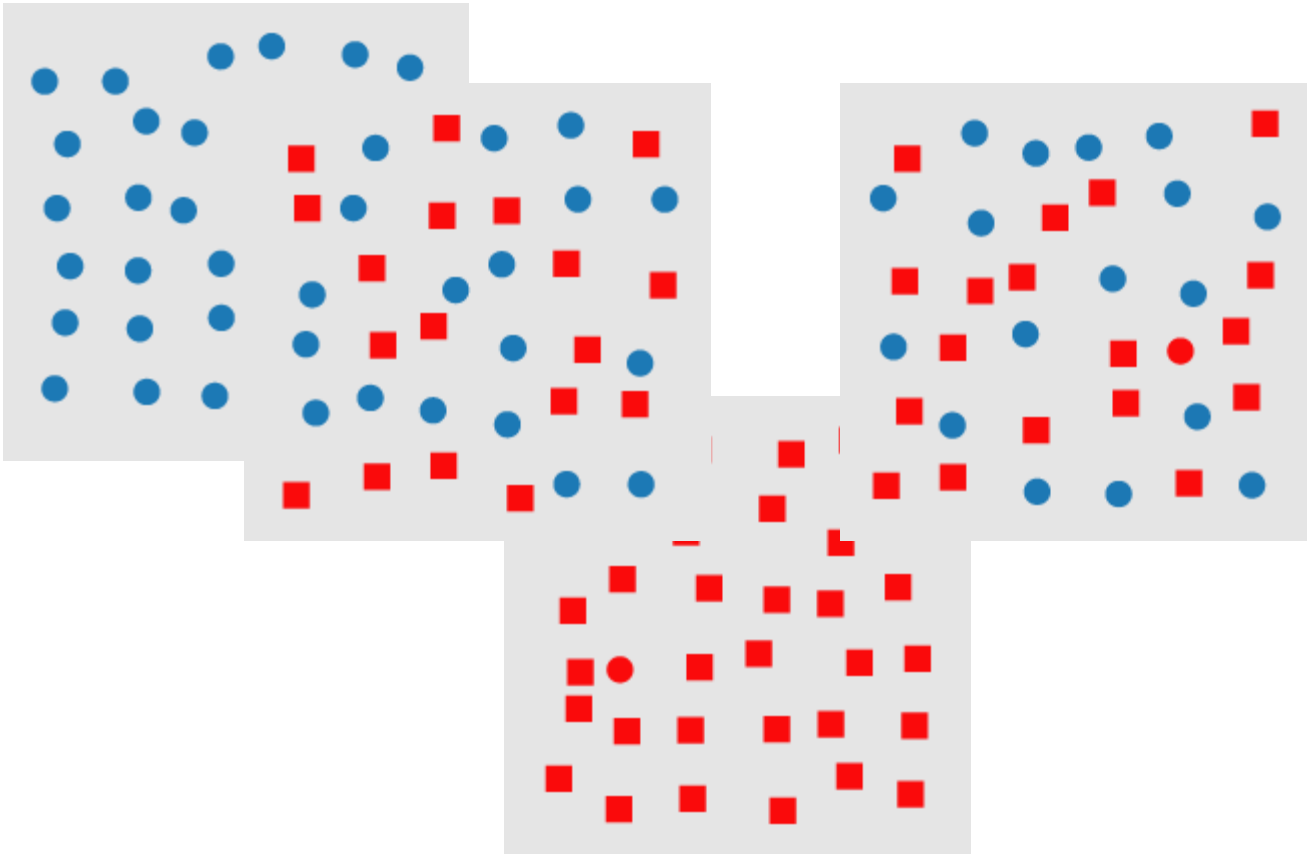


- Features detected at lowest level: lines, corners
- Combined upwards
- Construct object models

"pandemonium" theory, [O Selfridge 59]

Local vs Global Perception

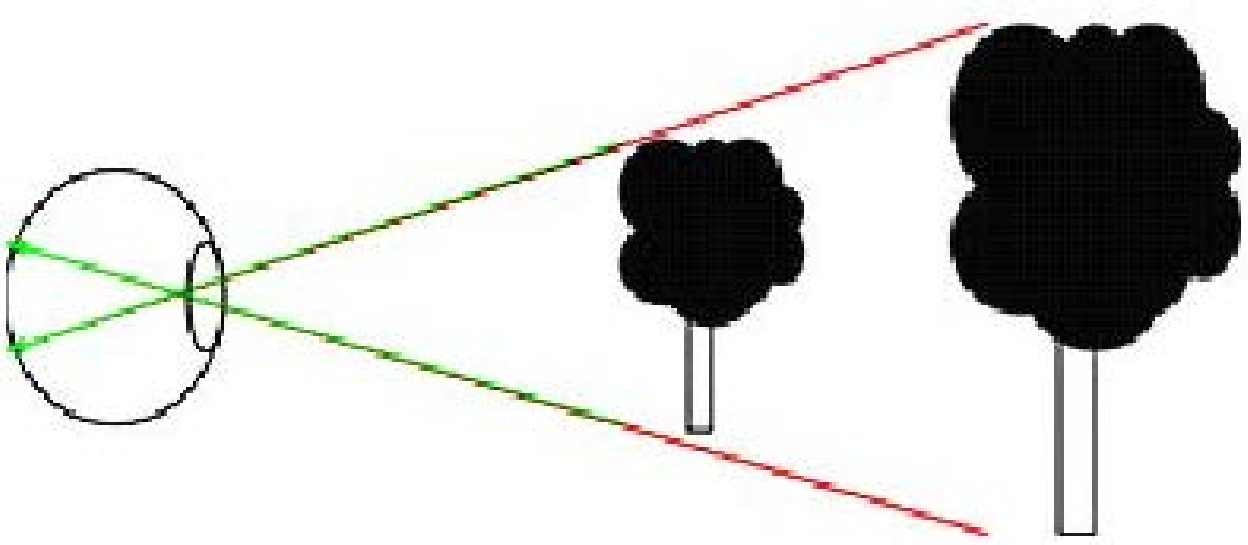
Find the red circle



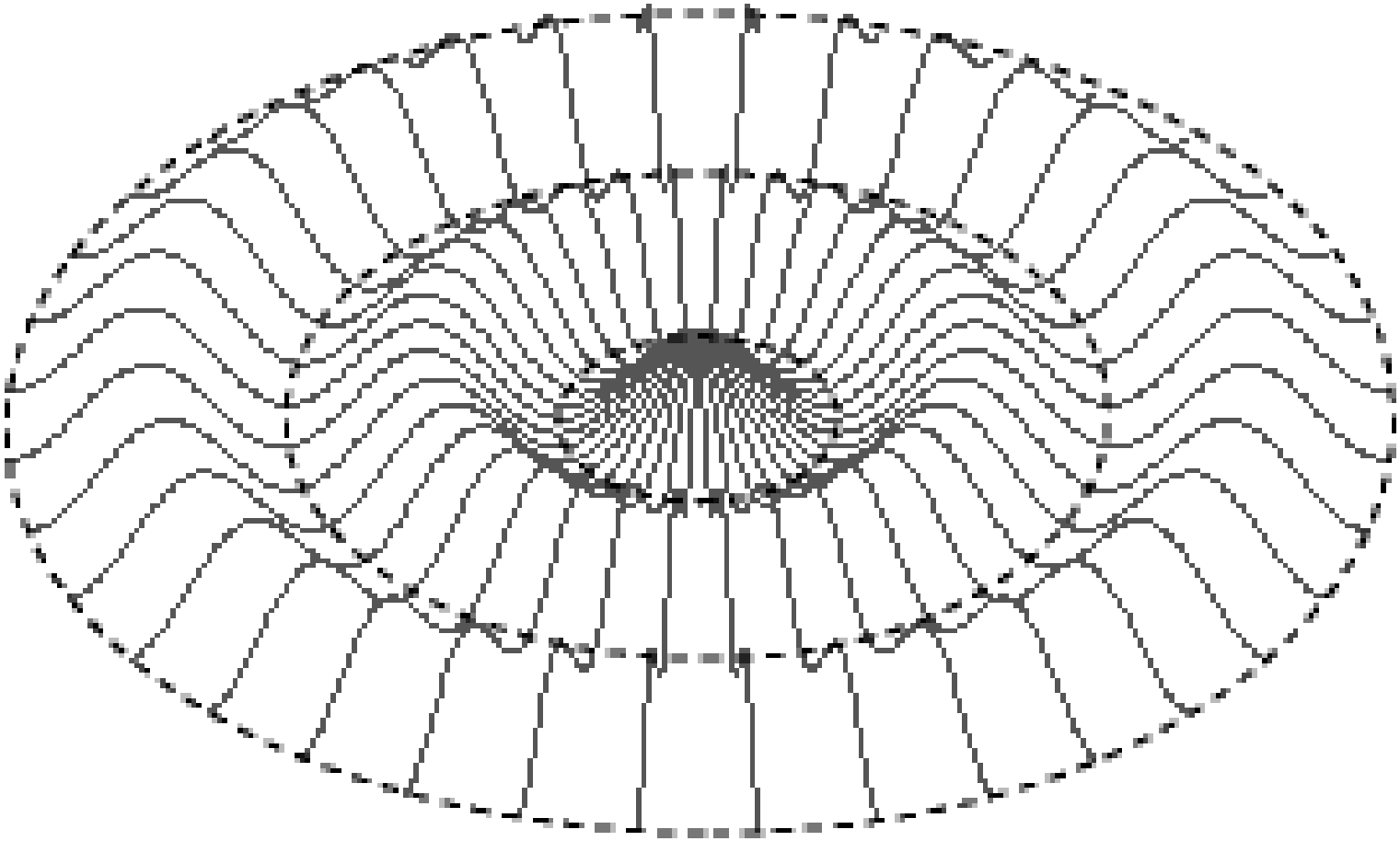
[http://www.csc.ncsu.edu/faculty/healey/PP/]

3-D Perception

Perception of Form

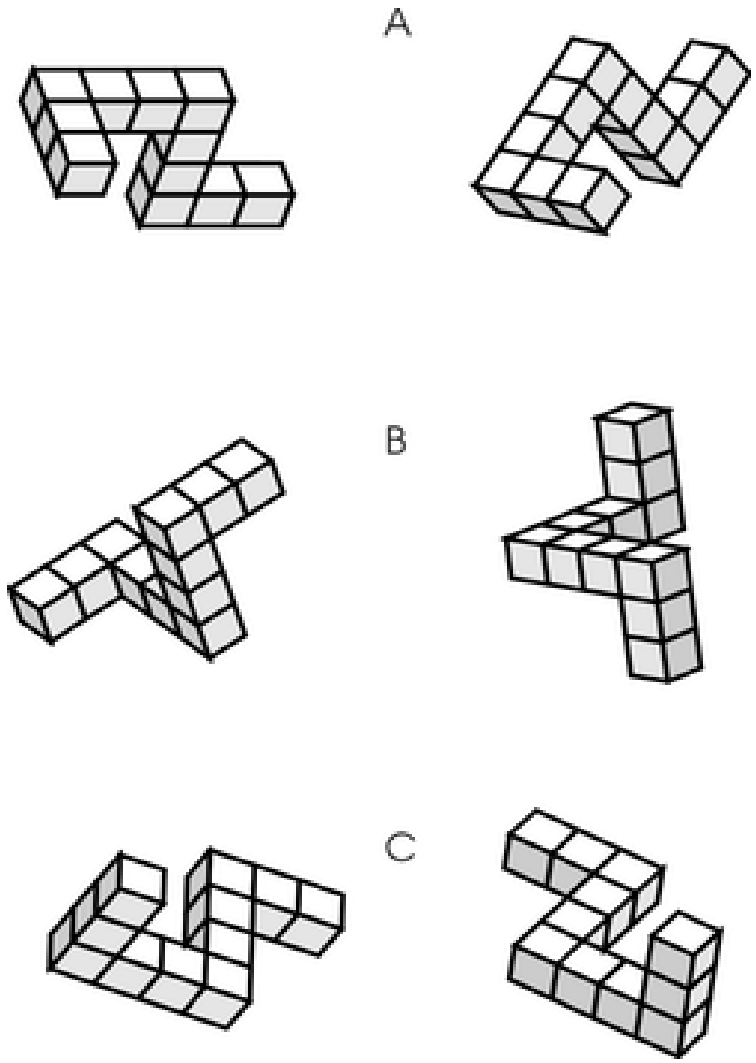


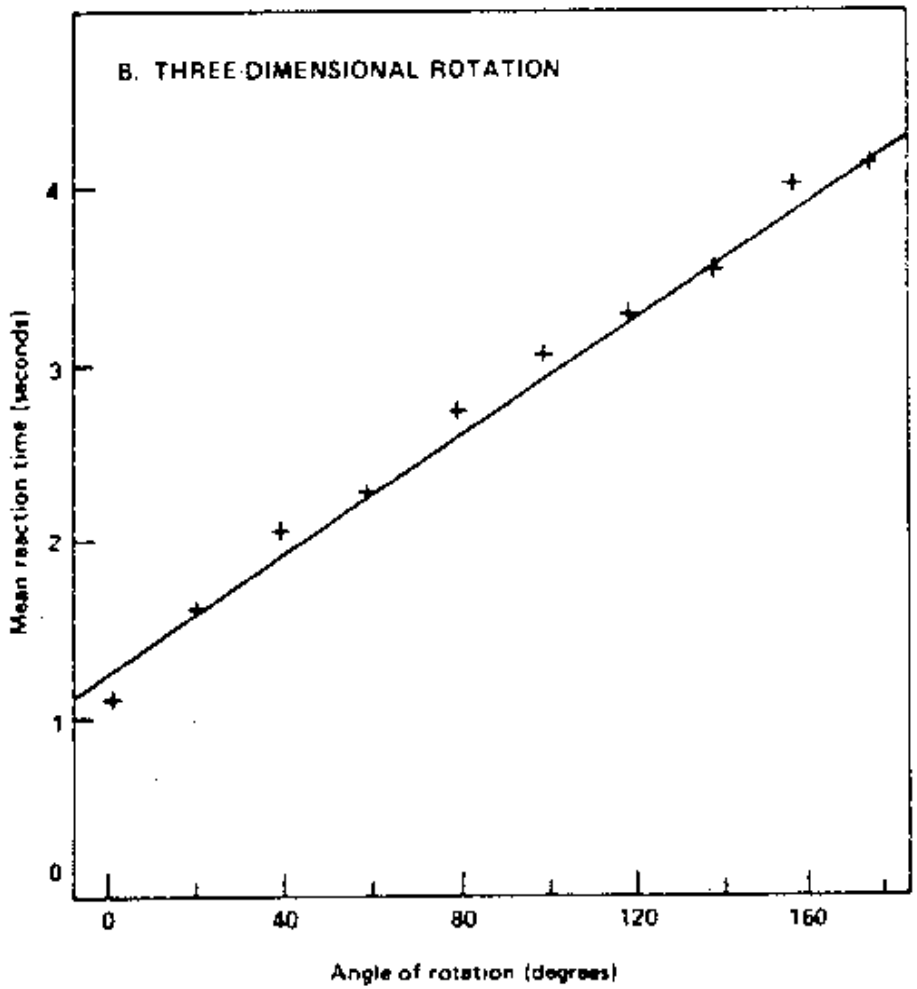
Perception of Form



[donald hoffman] : Visual intelligence 2000

Mental Rotation

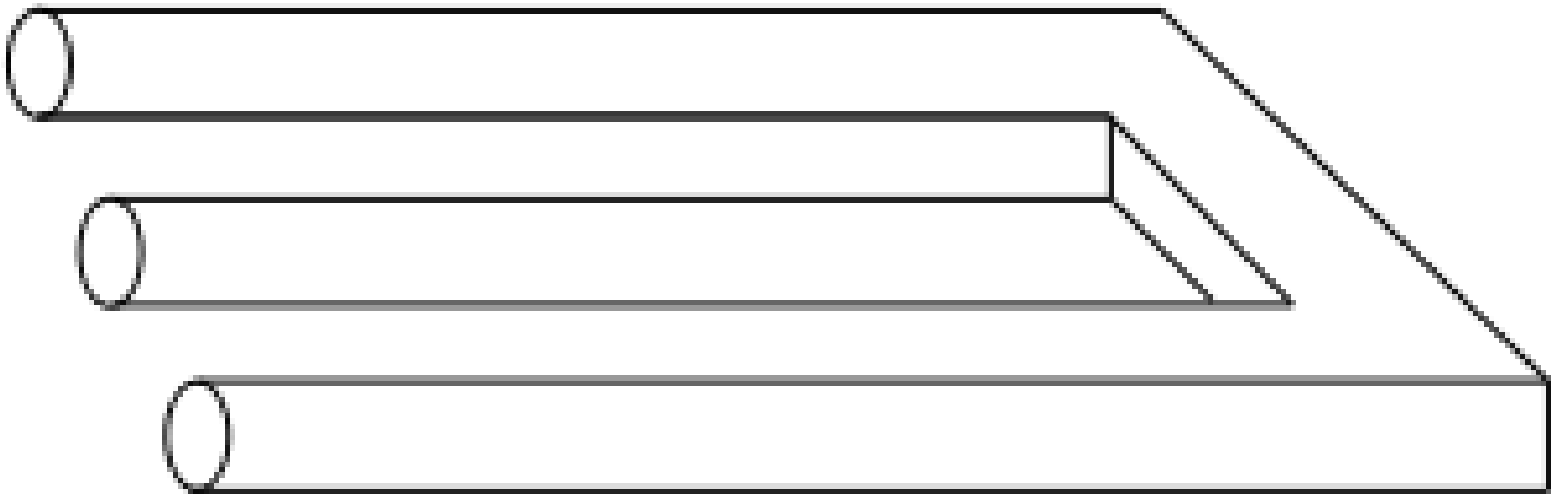




Shepard and Metzler (1971)

Local vs Global Perception

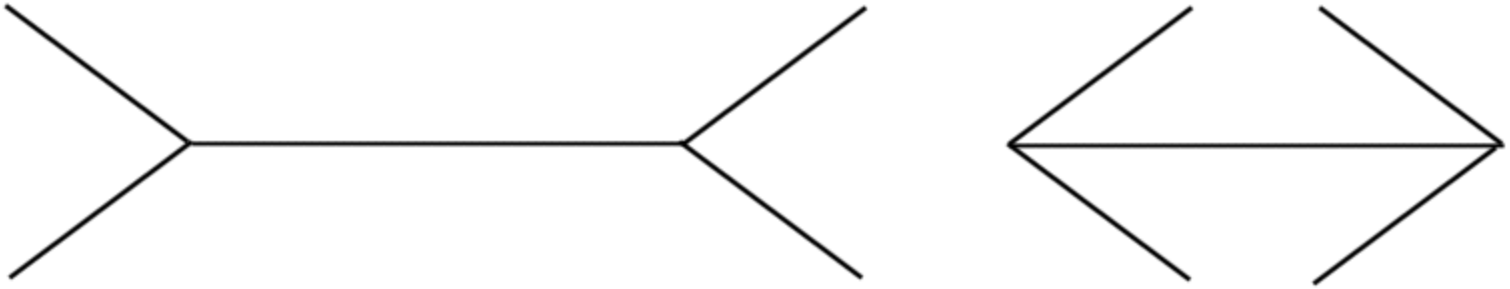
Devil's tuning fork



Local to Global conflict

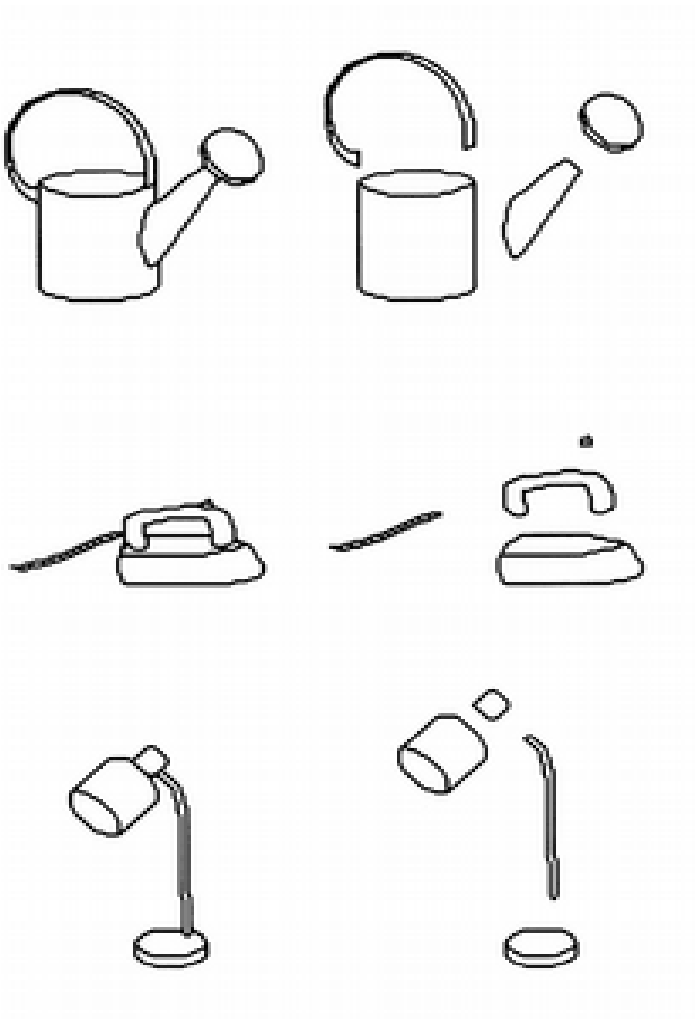
Cultural differences in perception

- ‘Carpenteredness’
  (Segall, Campbell & Herskovits, 1966):
  - Muller-Lyer illusion



3-D Shape perception: Geon theory

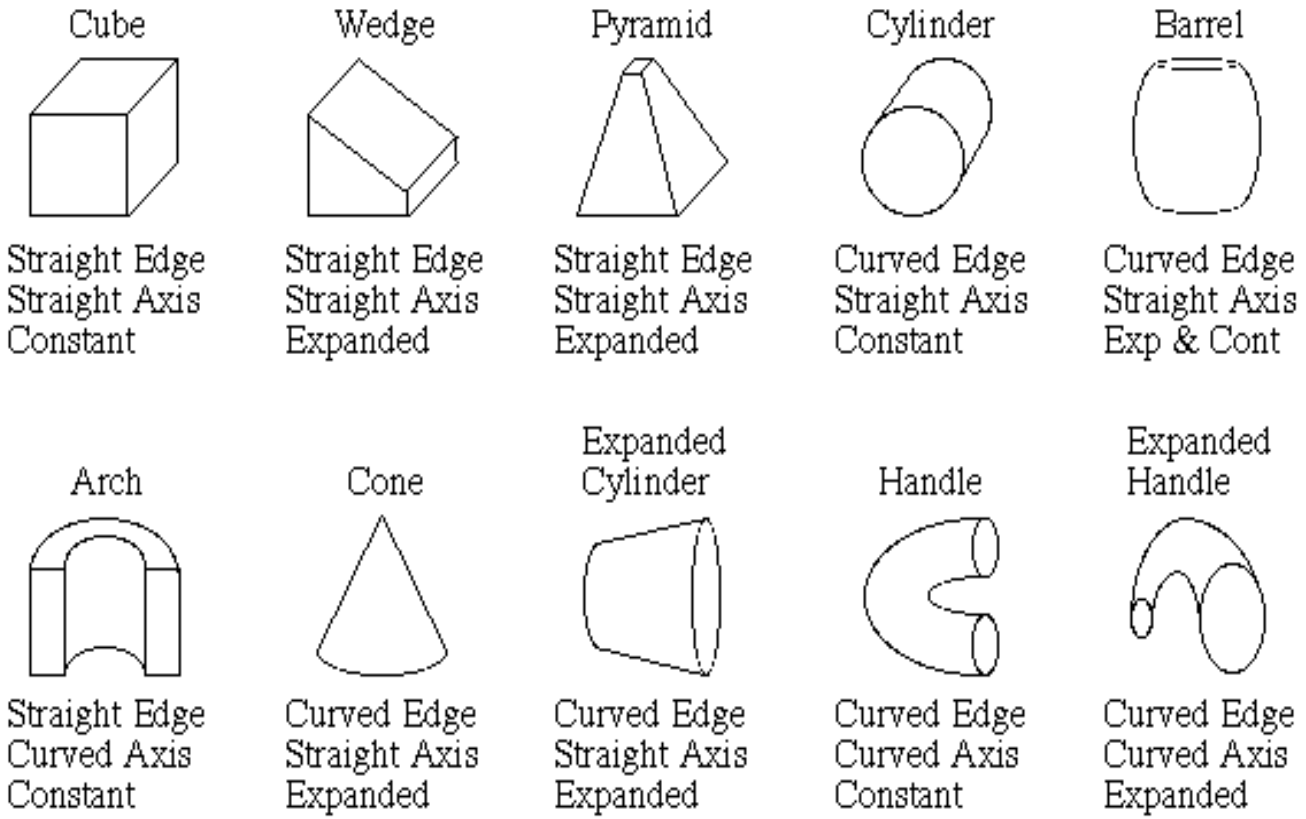
- Volume Decomposition :
- Geon theory
  - Recognition by parts
  - Each part = simple 3Dshape / volume



Recognition-By-Components (RBC) theory; [Biederman, 1987]

Object perception: Geon Theory

Geons:



Recognition-By-Components (RBC) theory; [Biederman, 1987]

Binocular Rivalry: Hole in the hand

look through tube
with left eye and
at right hand w right eye

From perception to categories

Classical Categories

- Categories defined by set of properties common to all members

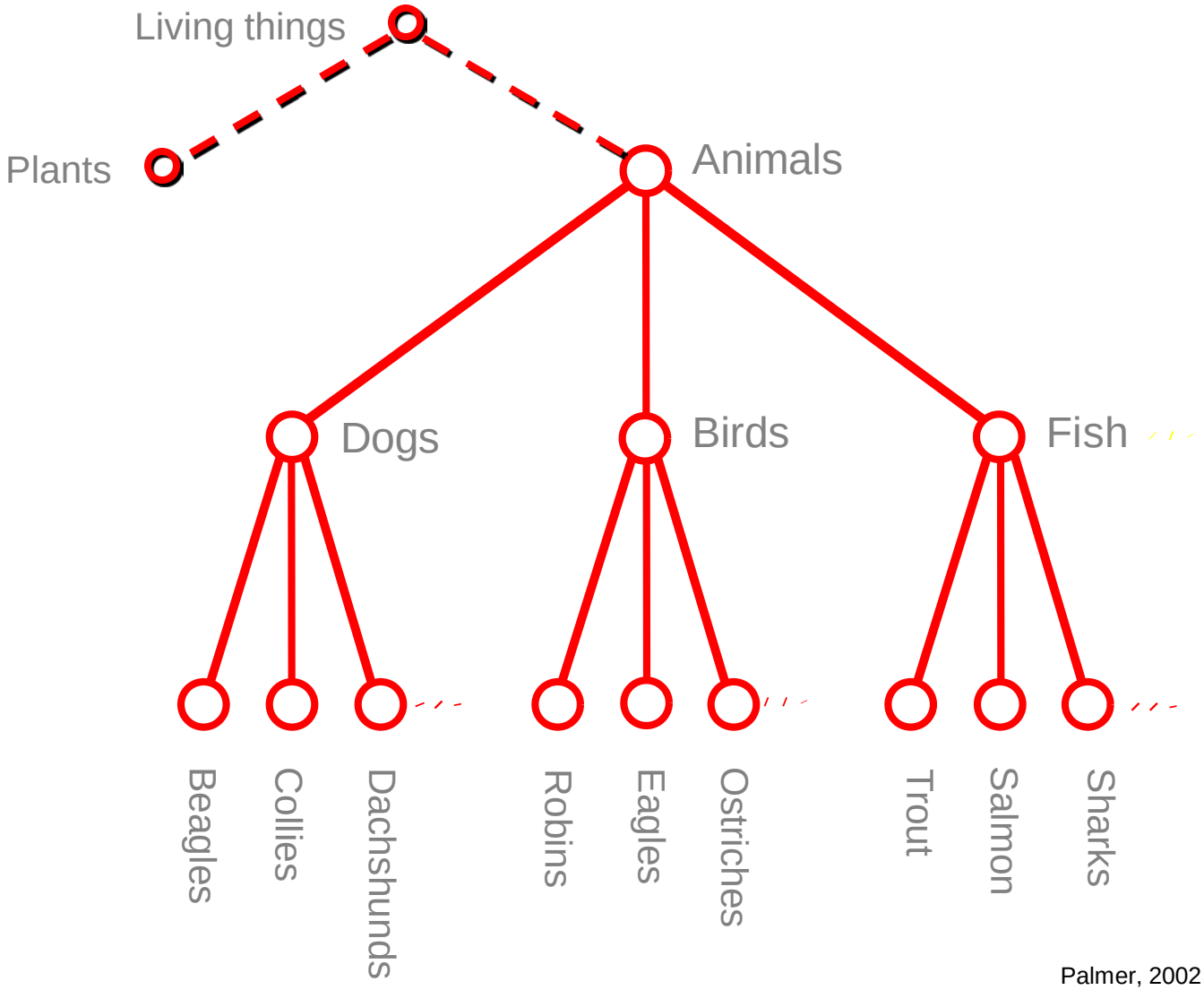
  f(x) = {x | has-prop (x,[a,b,c])}

- Category membership is True / False [f(x) is boolean]

- Every member in a category is equal

Aristotle

Classical category hierarchy



Palmer, 2002

Prototype Theory

- Classical theory does not explain human categories
  - What is a prototypical bird?
    a) sparrow
    b) Penguin

- Prototype Theory [Rosch, 70s]:
  - membership in a category is **graded**
  - properties of the category ← best prototypes

- Categories = clusters of shared features

- Boundaries are not crisp

Naming pictures

"Say the first name that comes to mind"

Prototypical

"Bird"
not "Robin"

"Bird"
not "Sparrow"

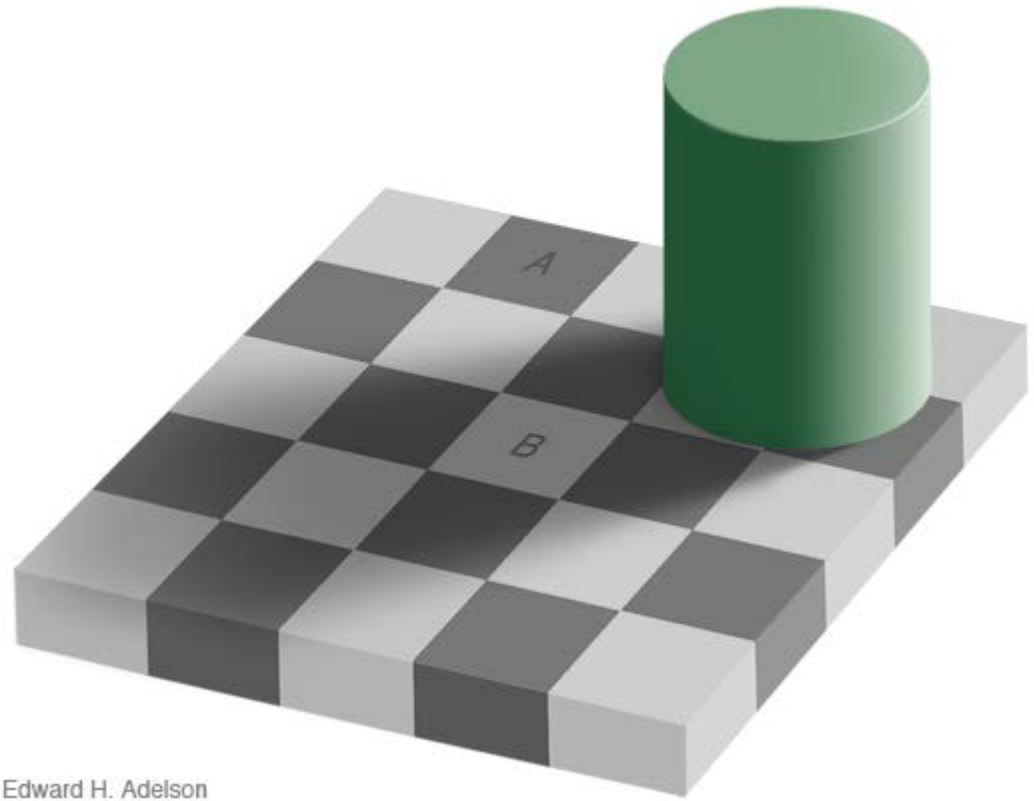
Atypical

"Ostrich"
not "Bird"

"Penguin"
not "Bird"

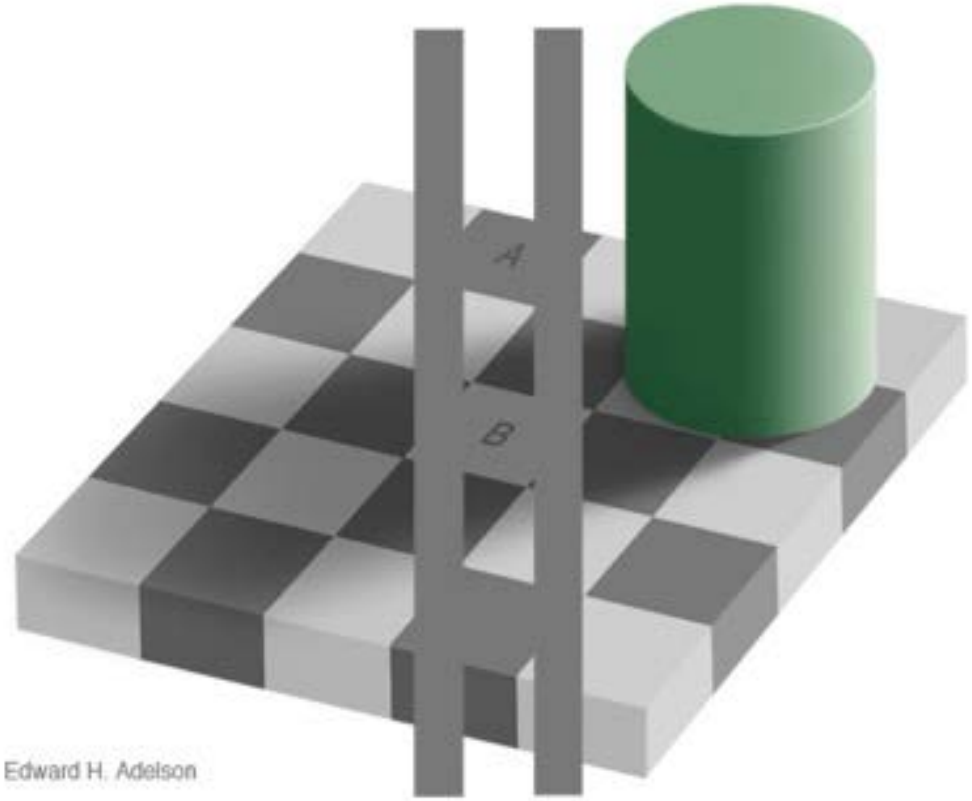
[Jolicoeur, Gluck & Kosslyn (1984)]

Colours



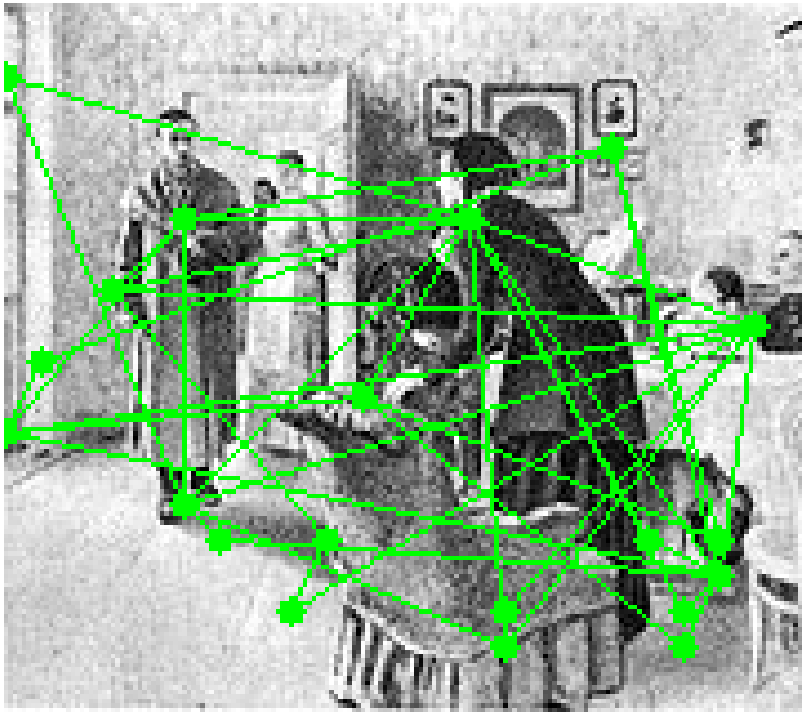
Edward H. Adelson

Colours



Purpose in Vision: Attention

Sense out of Blooming Buzzing Confusion



Ilya Repin: Unexpected Visitor

1

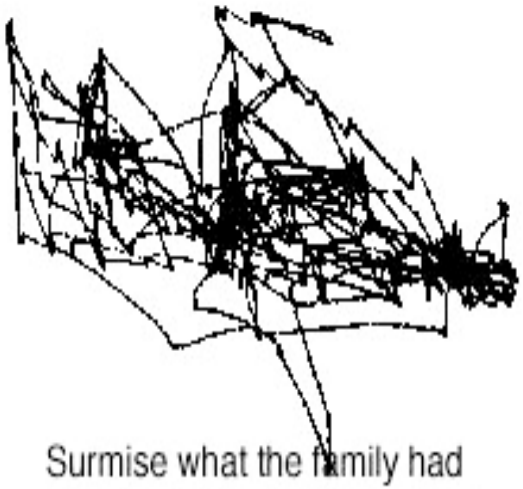
Free examination.

2

Estimate material circumstances of the family

3

Give the ages of the people.



4

Surmise what the family had been doing before the arrival of the unexpected visitor.

5

Remember the clothes worn by the people.

4. Aus: A. L. Yarbus: *Eye Movements and Vision*, New York 1967 S. 174 (Fig. 109)

Reading

हिंदी के आख्यान पढ़ने का हमेशा से मुझे बहुत शौक था। ये शौक मुझे धरोहर में अपनी माँ से लिया है, जो कि स्वयं हिंदी के आख्यान किताबी कीड़ो कि तरह चाट डालती थी। परन्तु हिंदी के आख्यान ढूंढ पाना बहुत कष्टदायक कार्य है। माँ ने कभी पुस्तके खरीदी ही नहीं थी, वे ग्रंथालय से पुस्तके ला ला के पढ़तीं और उन्हें लौटा देती थी। इस कारणवश पुस्तको का कभी संग्रह न हो सका। जैसे तैसे मैंने रबिन्द्रनाथ टैगोर की "काबुलीवाला" पढ़ी थी और आनंद से प्रफुल्लित हो उठी थी। परन्तु उसके बाद से कोई अवसर ही न मिल सका। लेकिन भगवान के आगे कहीं कुछ रुक पाया है? कुछ ही दिनो पहले मुझे एक ऐसा ही स्वर्णिम अवसर मिल गया। ऐसा जिसकी मुझे हमेशा से तलाश थी। लैंडमार्क में हिंदी पुस्तको का ढेर लगा था जिसे मैं उठा लायी ! अब एक एक कर के मैं वो सारी किताबें पढ़ूंगी, जो पहले न पढ़ सकी थी। इनमे से सब से ऊपर थी मुंशी प्रेमचंद की "नैराश्य लीला" और हरिवंश राइ बच्चन की "बचपन के साथ"। इस उपन्यास के बारे में मैंने माँ से बहुत सुना था।

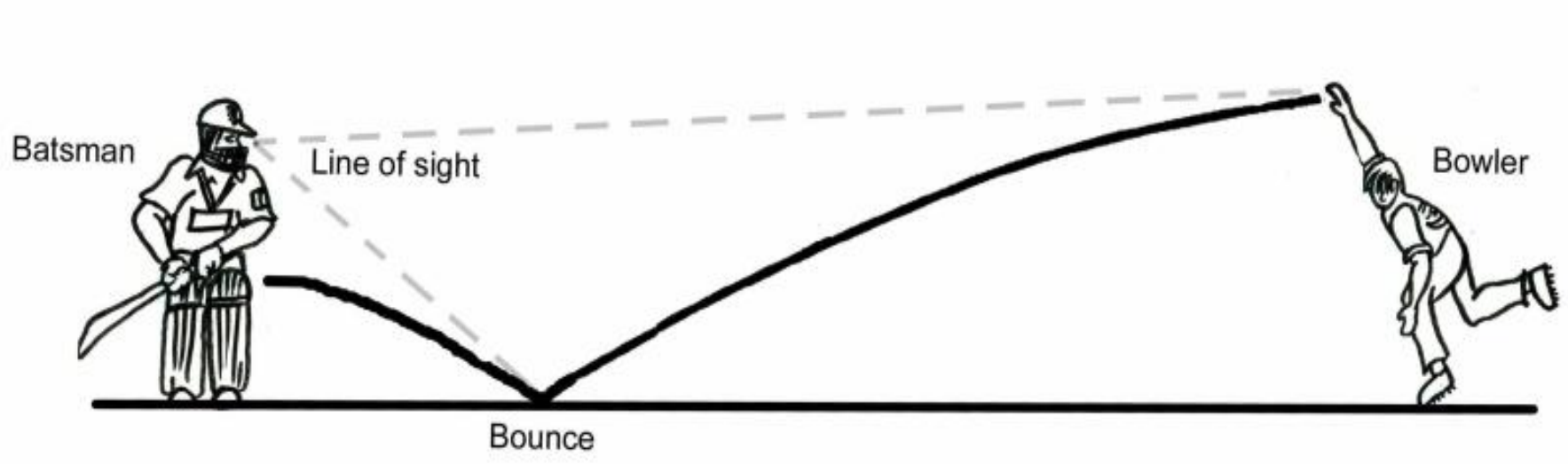
Task-driven Attention

Attention is determined by the task

Q. Is the task-gaze relation innate, or is it learned?

Two models of attention
- bottom-up
    (independent of task, data driven, innate)
- top down
    (depends on task, possibly acquired)

Visual Attention in Cricket



mukerjee : *Sehwag's Brain*, 2003

Visual Attention in Cricket

In a perfect world, you will see the ball early and play it late.

- Geoffrey Boycott

What information is needed to play the ball?

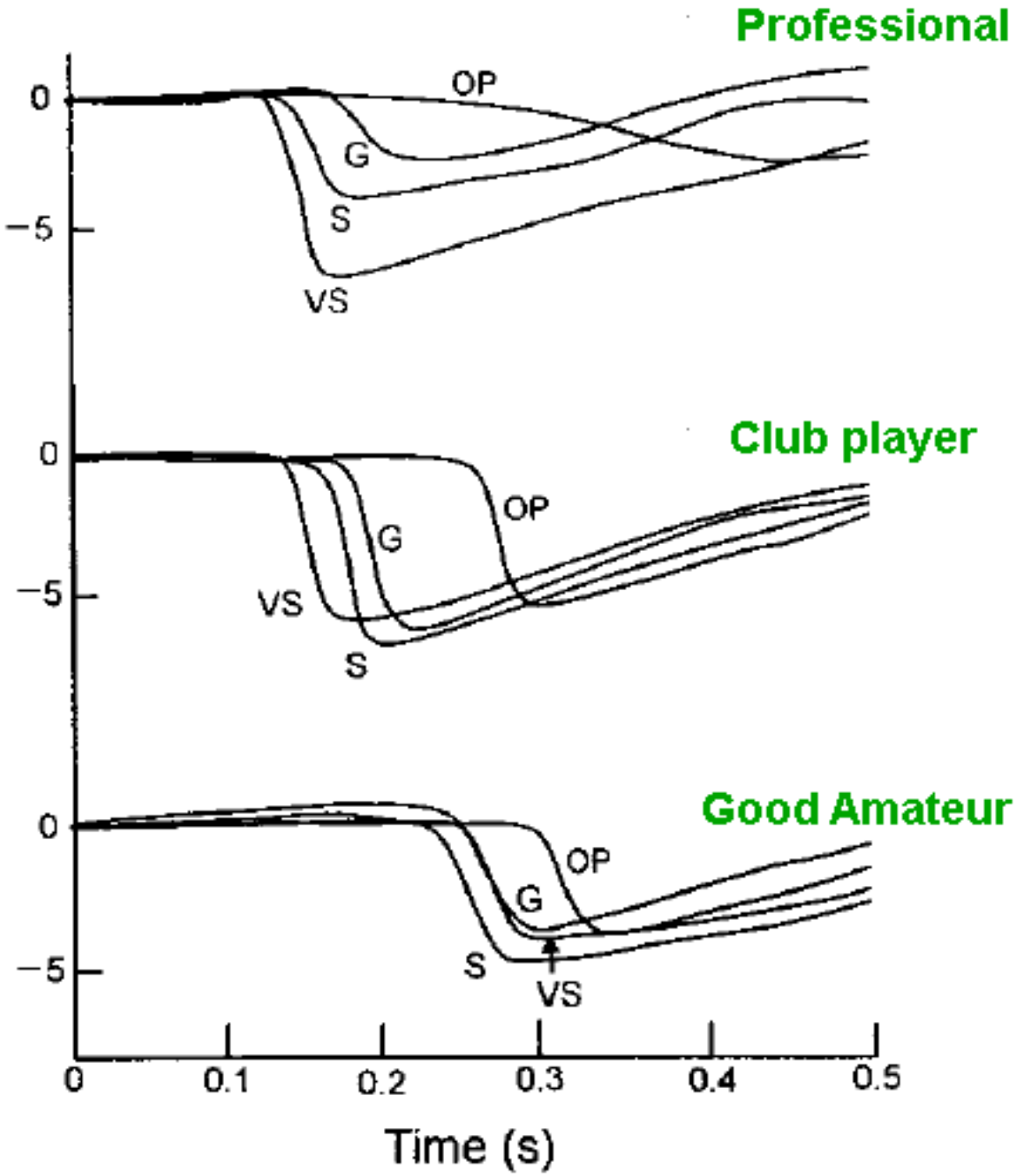
- line and length at bat
- speed / timing

Time : less than ½ sec for fast balls

Visual inputs: time from delivery to bounce
position of bounce

Saccades

VS, very short;
S, short;
G, good length;
OP, over pitched.

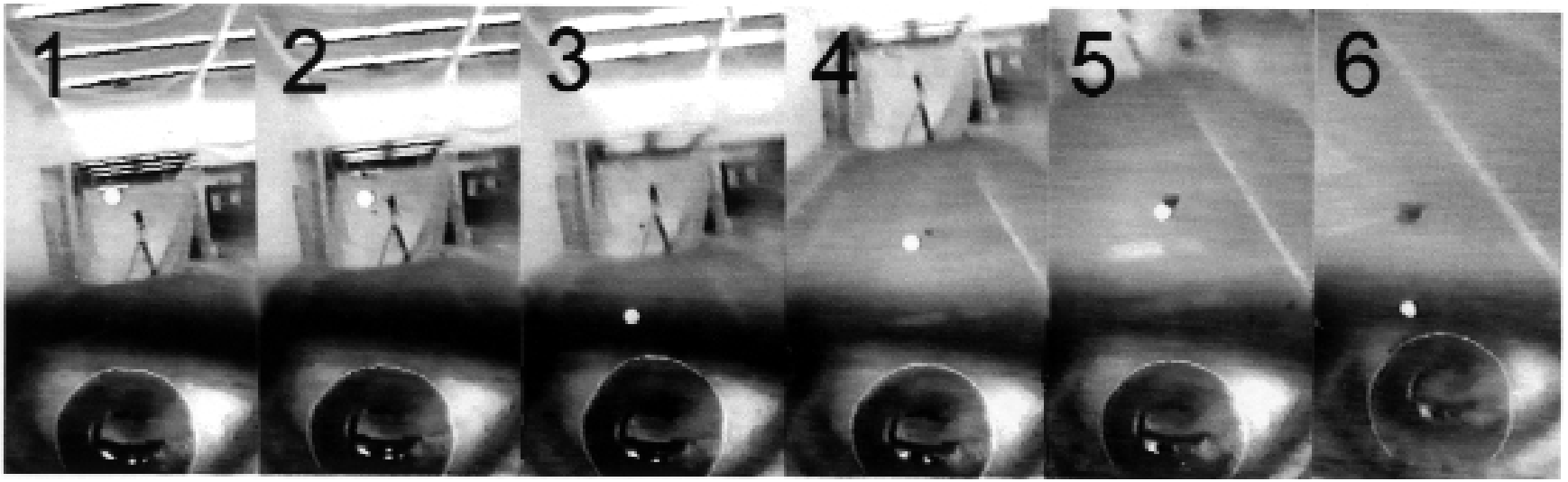


on short deliveries, a professional shifts his gaze to the bounce point 100ms before the amateur

[land-mcleod-00_eye-movements-for-batsmen]

Saccades

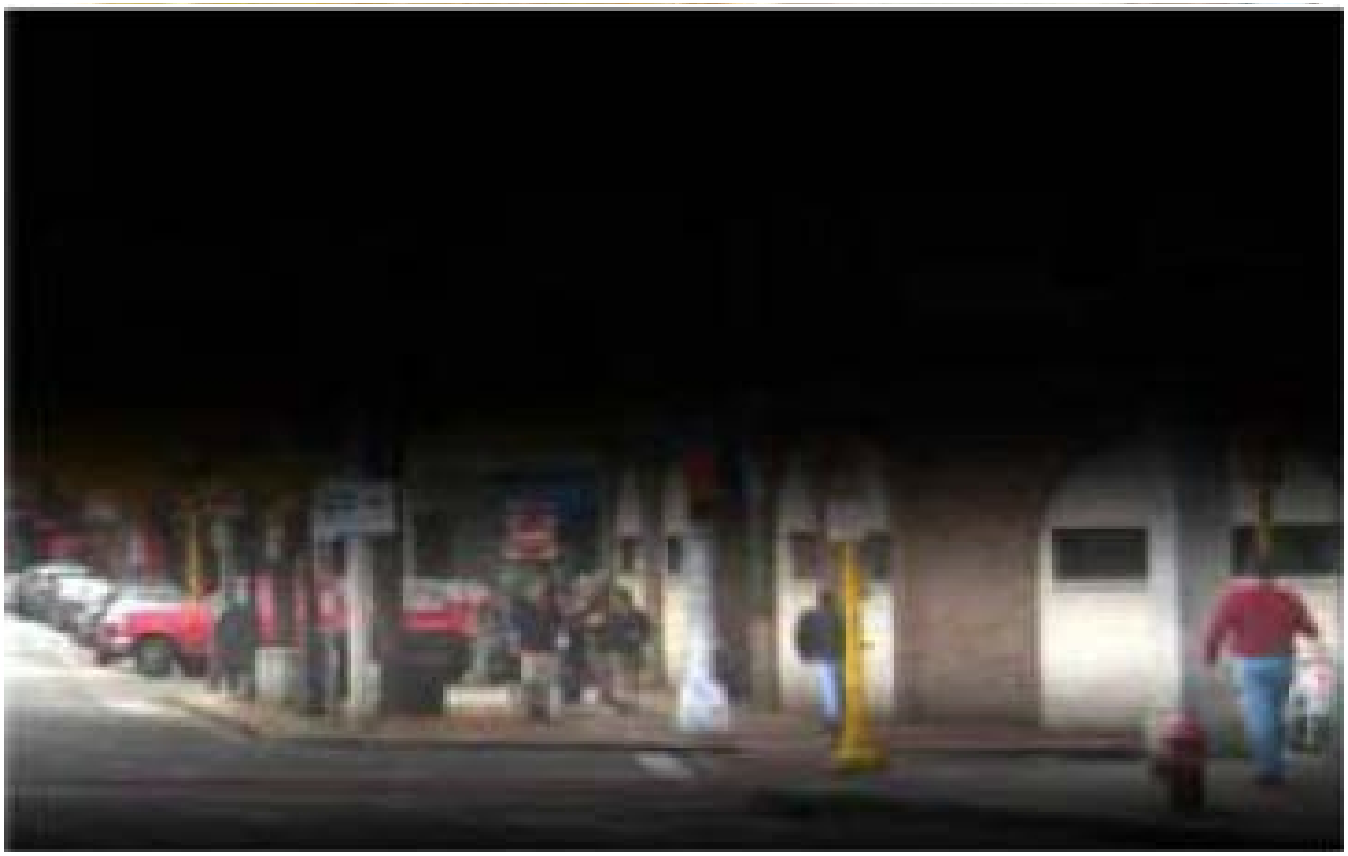
batsman's view



Foveal gaze shifts from the delivery (2) to the expected bounce point (4) before the ball arrives there (5). It tracks the ball for the next 200ms

[land-mcleod-00_eye-movements-for-batsmen]

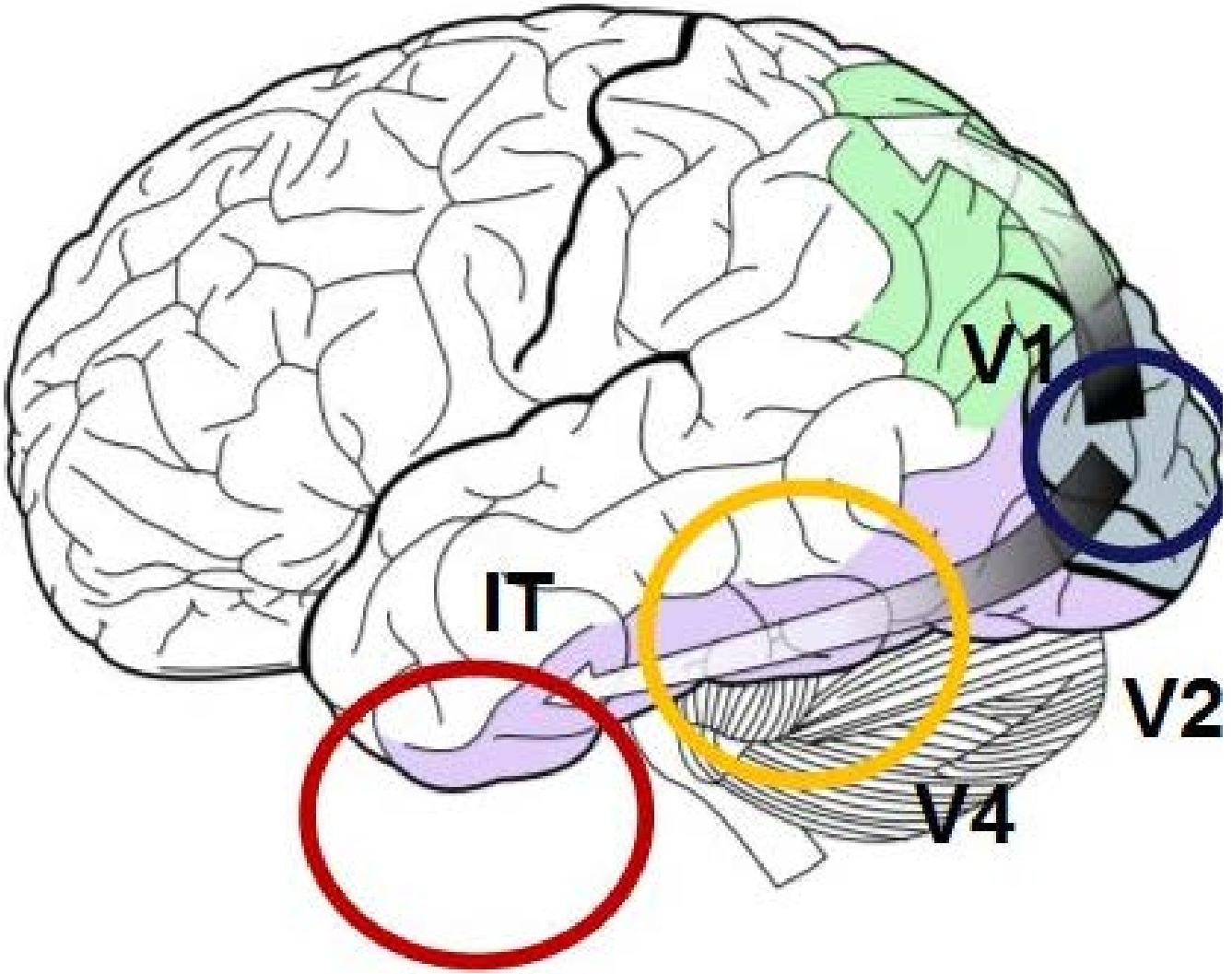
Top-down Attention: Category-driven context



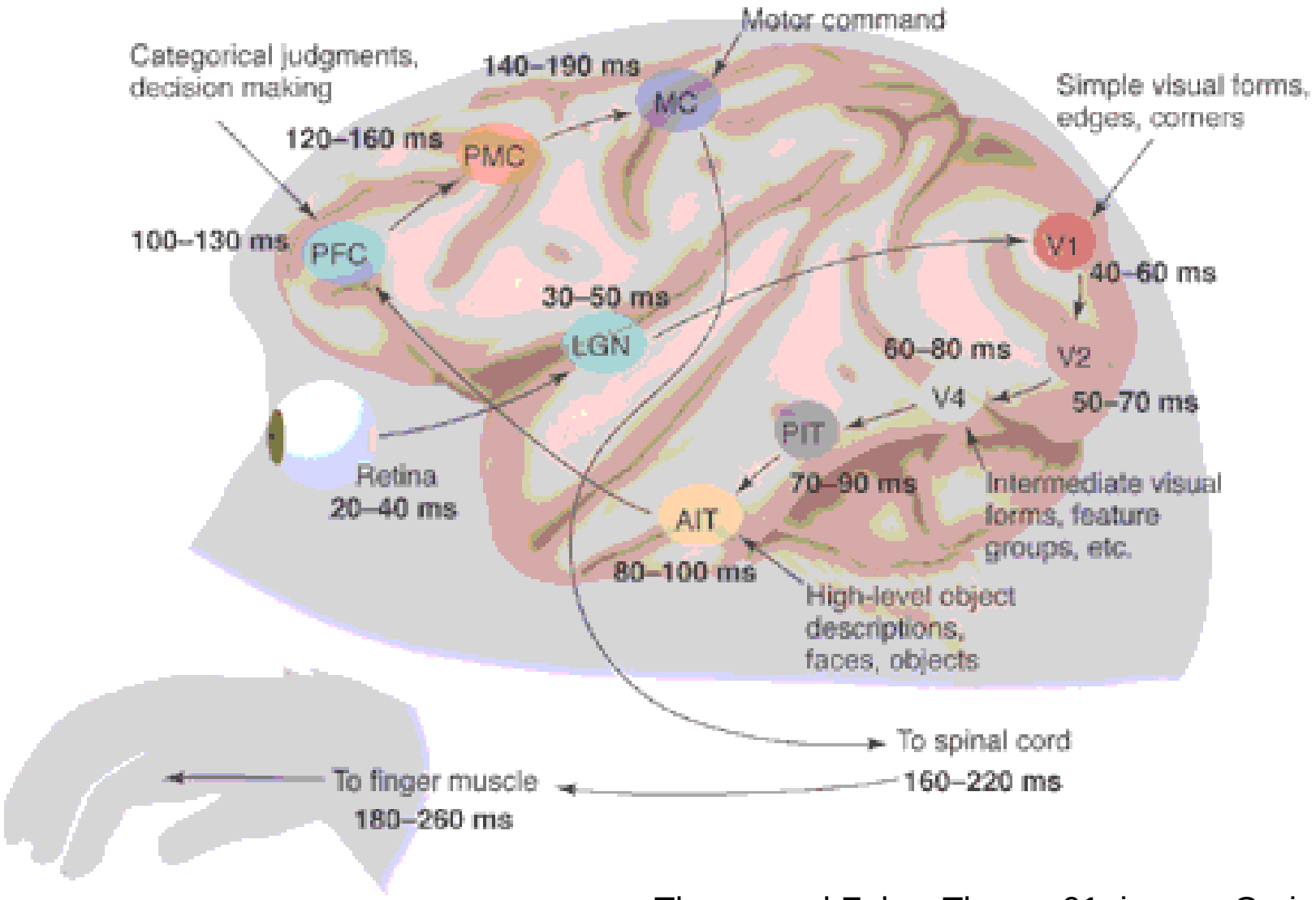
[judd-torralba-09_learning-to-predict-where-humans-look]

Neural Processes for Object Recognition

Perception in the Brain

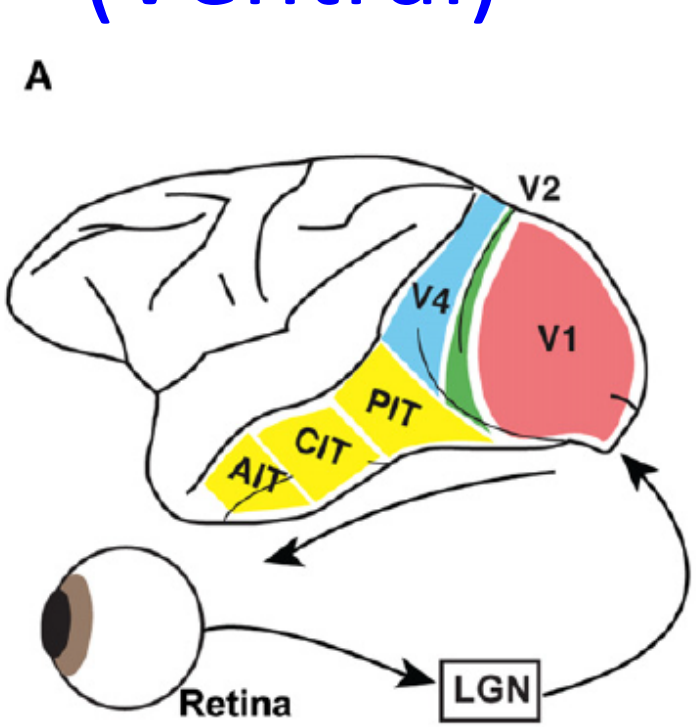


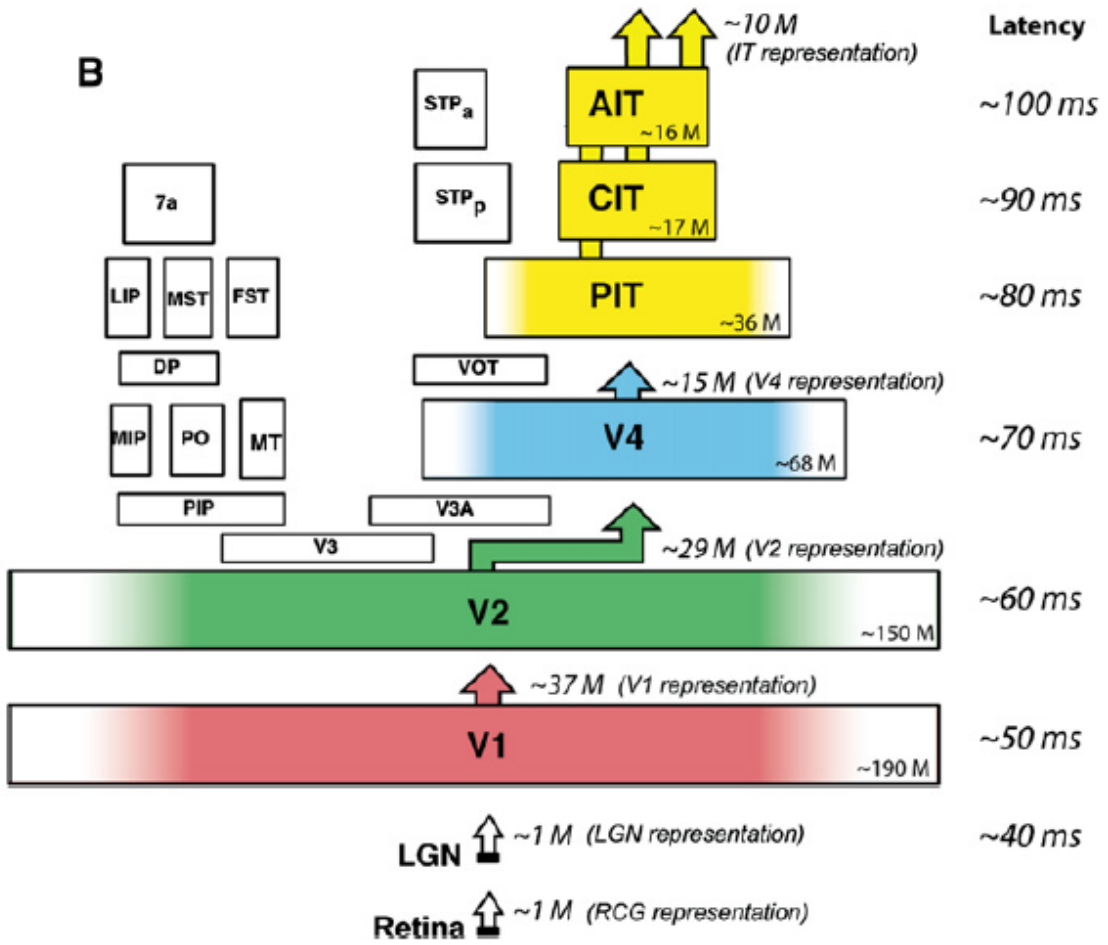
From Perception to Action



Thorpe and Fabre-Thorpe 01; image: Carin Cain

Visual Recognition Pathway (Ventral)





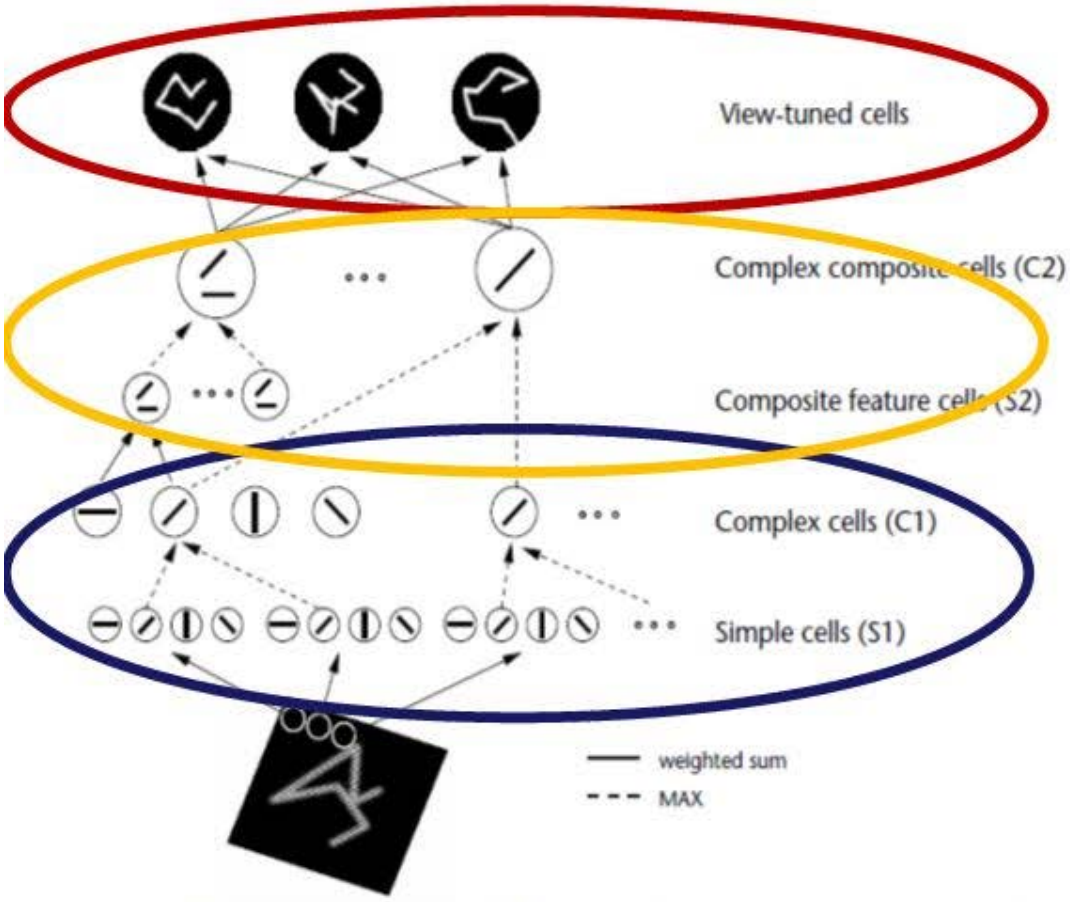
Ventral pathway, Macaque brain

Deep colours: process central 10 deg

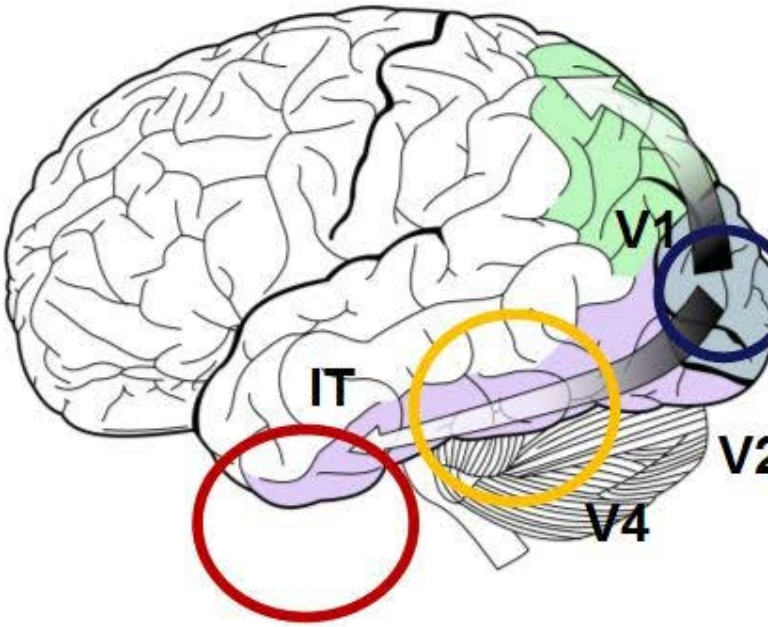
dicarlo-zoccolan-12-visual-object-recognition

Computational model of vision

(primate visual cortex, ventral stream)



Riesenhuber and Poggio, Hierarchical models of Object recognition in cortex, (1999)



: http://en.wikipedia.org /wiki/Visual_perception (July, 2011)

Recognition

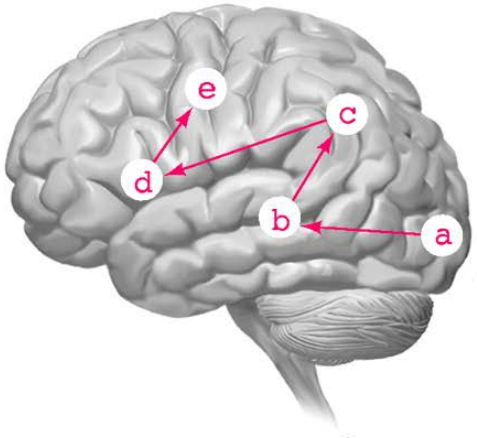
vs

Imitation

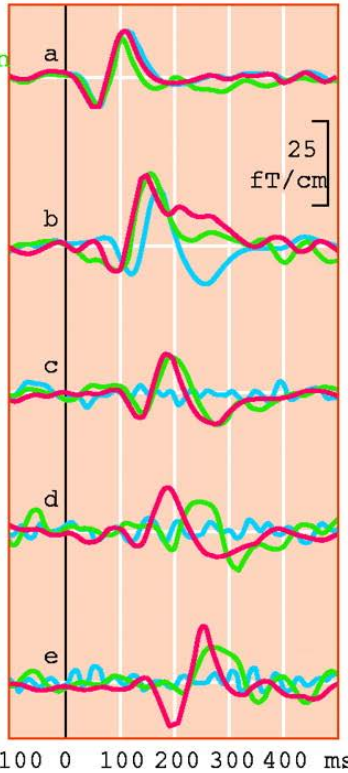
Whole-scalp neuromagnetometer

Nonverbal lip-form stimuli

Activation sequence



Responses

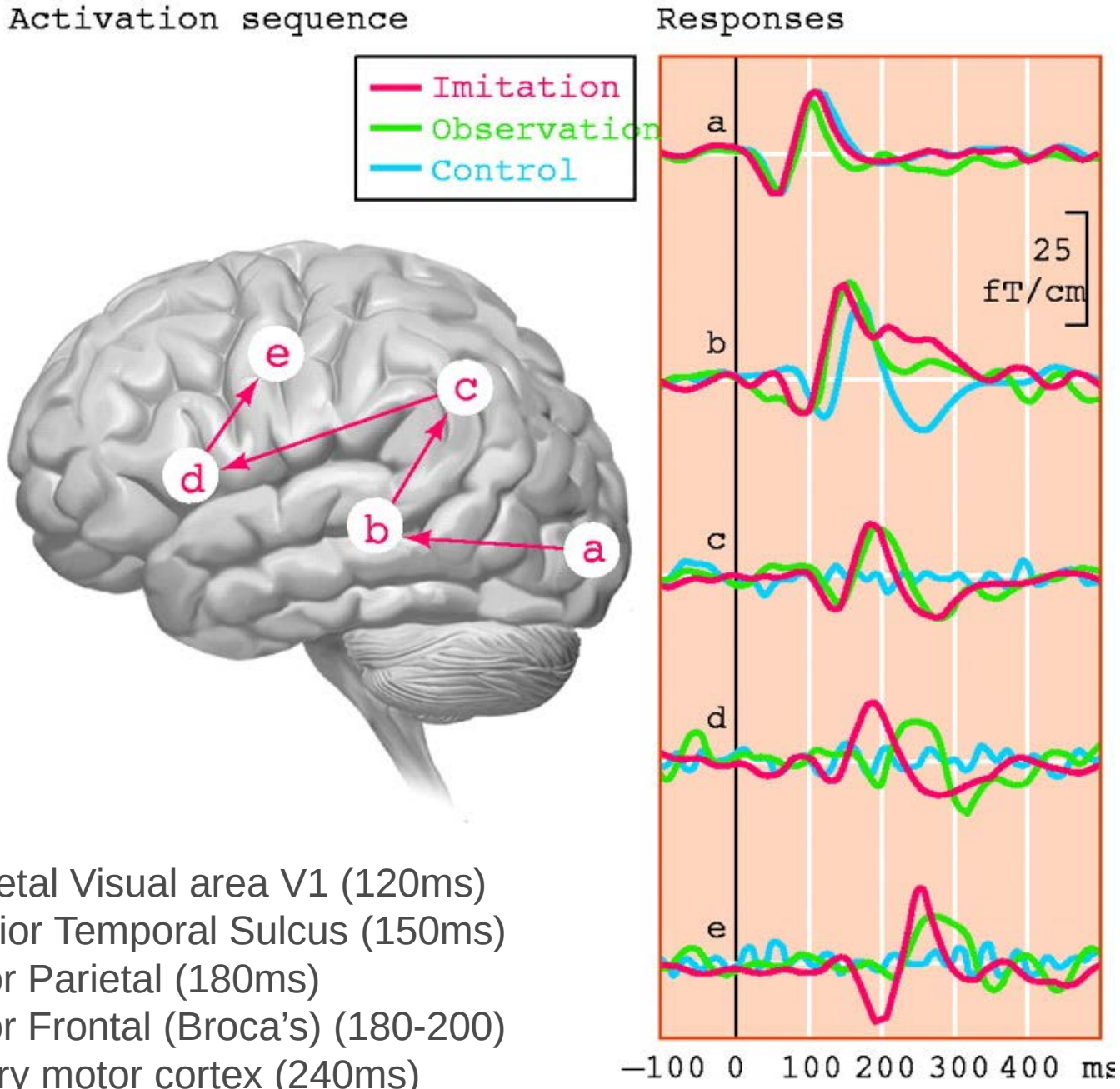


[Nishitani and Hari 02]

Recognition

vs

Imitation

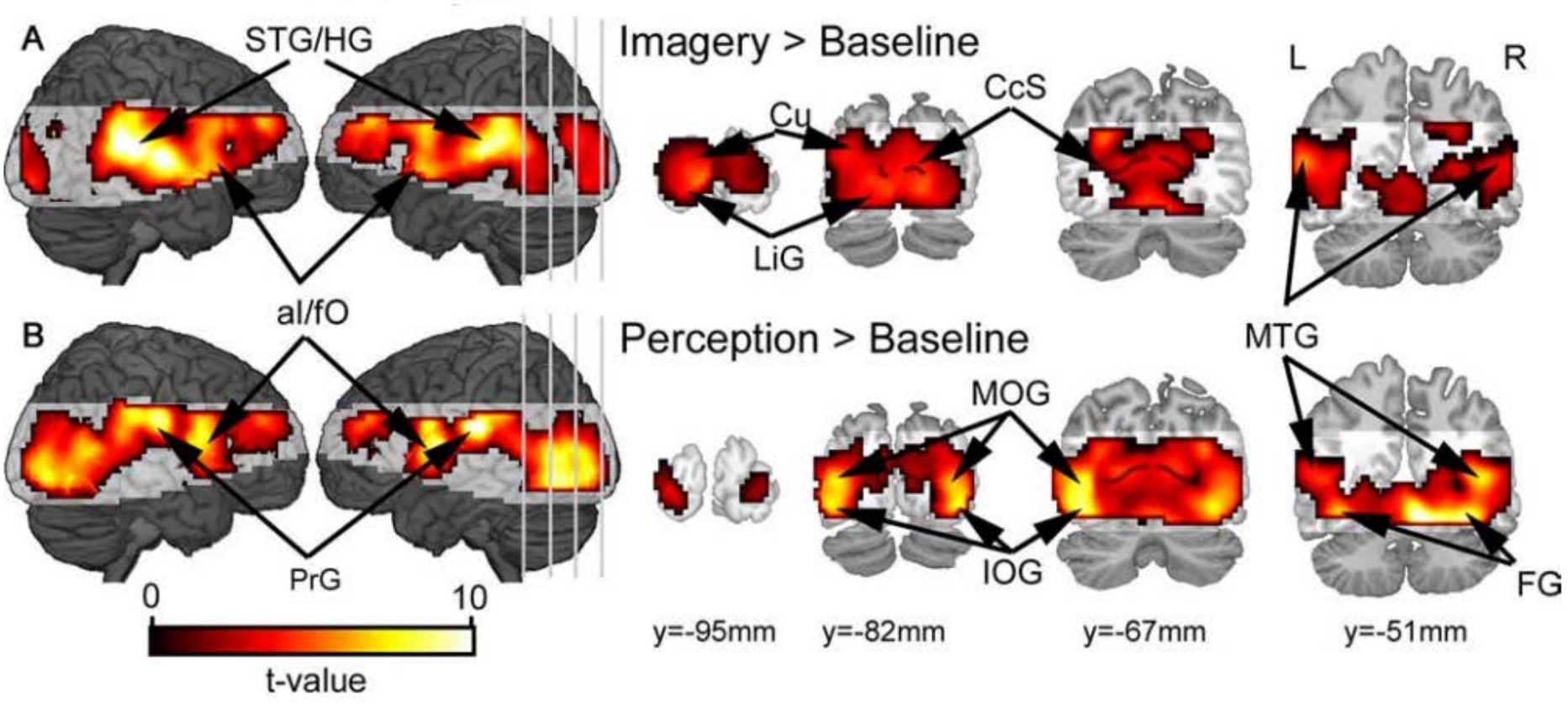


a. Occipetal Visual area V1 (120ms)
b. Superior Temporal Sulcus (150ms)
c. Inferior Parietal (180ms)
d. Inferior Frontal (Broca's) (180-200)
e. Primary motor cortex (240ms)

[Nishitani and Hari 02]

Mental Imagery vs Recognition

Imagery



aI/fO: anteriorinsula/frontal operculum Cu: cuneus FG: fusiform gyrus IOG: inferior occipital gyrus LiG lingual gyrus MOG: middle occipital gyrus MTG: middle temporal gyrus MTG: middle temporal gyrus STG/HG: superior temporal sulcus/Heschl's gyrus

[Stokes etal 1990