Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 130/230. Harvard College/GSAS 78454

Web site: http://tinyurl.com/visionclass

→ Class notes, Class slides, Readings Assignments

Location: Biolabs 2062

Time: Mondays 03:00 – 05:00

Lectures:

Faculty: Gabriel Kreiman and invited guests

TA: Emma Giles

Contact information:

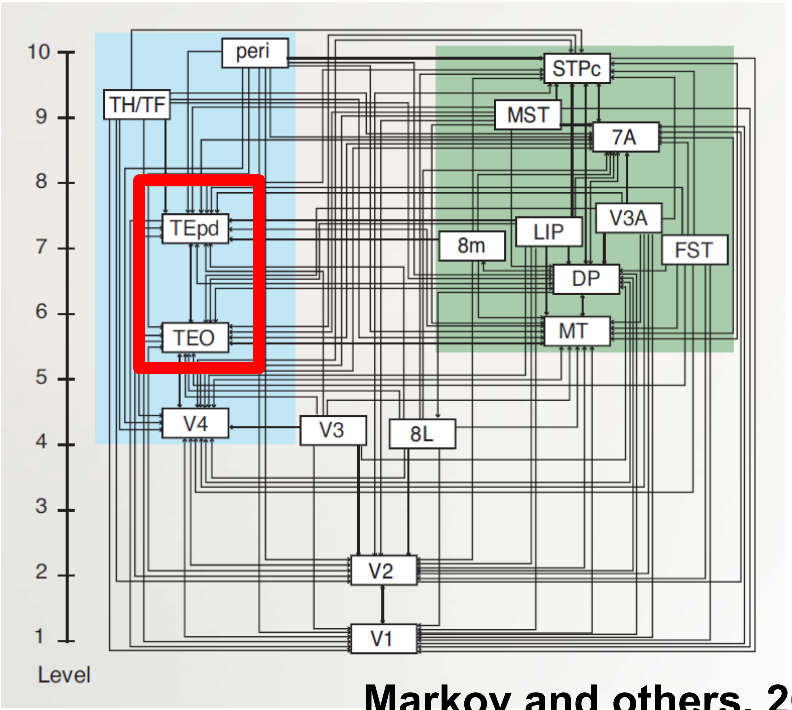
Gabriel Kreiman
gabriel.kreiman@tch.harvard.edu
617-919-2530
Office Hours: After Class. Mondays 5pm, or by appointment

Emma Giles
emmagiles@g.harvard.edu

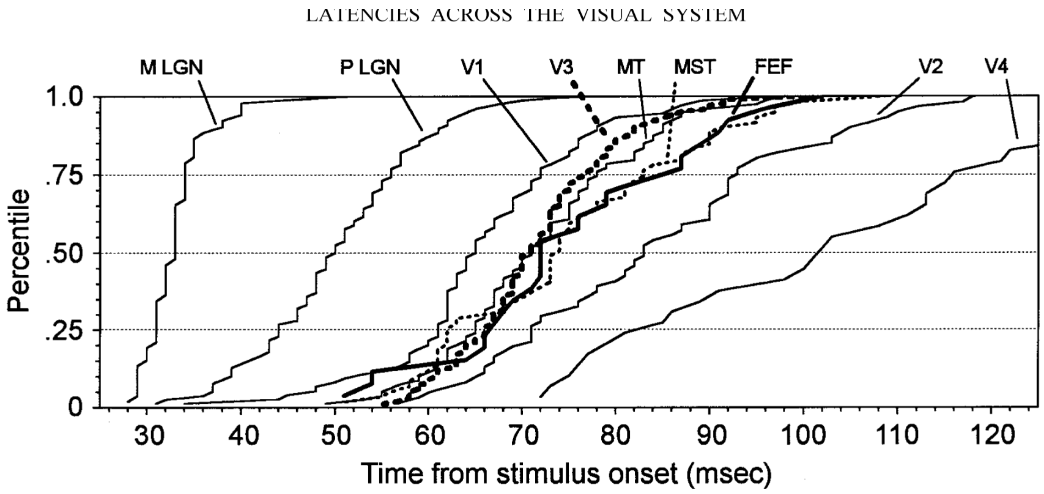
Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 230. Harvard College/GSAS 78454

Class 1 [09/10/2018]. Introduction to pattern recognition [Kreiman]
Class 2 [09/17/2018]. Why is vision difficult? Natural image statistics. The retina. [Kreiman]
Class 3 [09/24/2018]. Lesions and neurological studies [Kreiman].
Class 4 [10/01/2018]. Psychophysics of visual object recognition [Sarit Szpiro]
October 8: University Holiday
Class 5 [10/15/2018]. Primary visual cortex [Hartmann]
Class 6 [10/22/2018]. Adventures into *terra incognita* [Frederico Azevedo]
Class 7 [10/29/2018]. High-level visual cognition [Diego Mendoza-Haliday]
Class 8 [11/05/2018]. Correlation and causality. Electrical stimulation in visual cortex [Kreiman]
Class 9 [11/12/2018]. Visual consciousness [Kreiman]
Class 10 [11/19/2018]. Computational models of neurons and neural networks. [Kreiman]
Class 11 [11/26/2018]. Computer vision. Artificial Intelligence in Visual Cognition [Bill Lotter]
Class 12 [12/03/2018]. The operating system for vision. [Xavier Boix]
FINAL EXAM, PAPER DUE 12/13/2018. No extensions.

The hierarchy of ventral visual cortex, recap



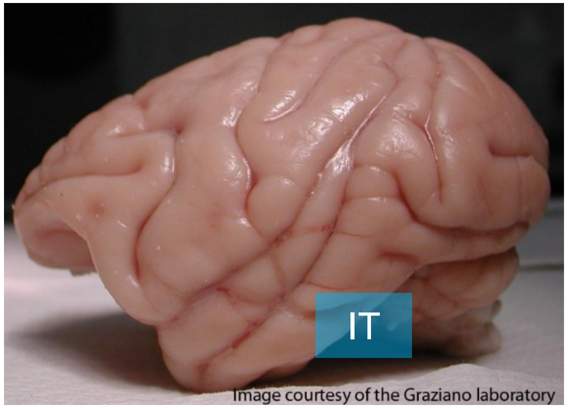
Markov and others, 2013

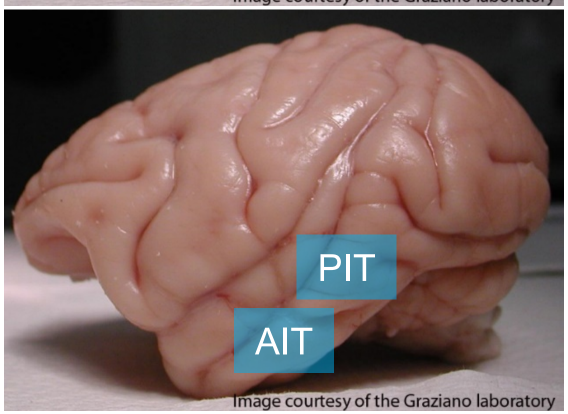


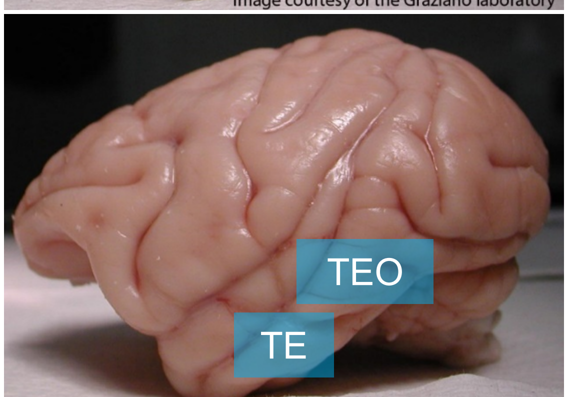
Increase in receptive field size

Increase in feature preference complexity

Inferior temporal cortex – Neuroanatomical notes







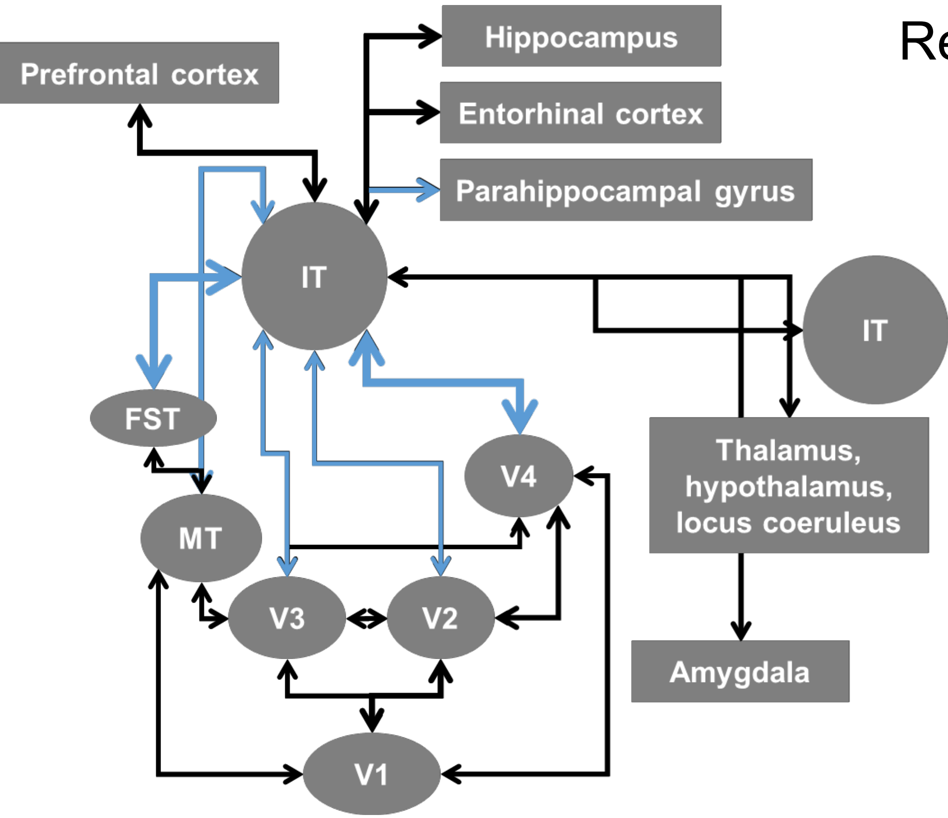
Nomenclature and subdivisions

- ~Brodmann cytoarchitectonic areas 20 and 21
- Area TE + TEO (Von Bonin and Bailey)
- Areas PIT, CIT, AIT (Felleman and Van Essen, 1991)

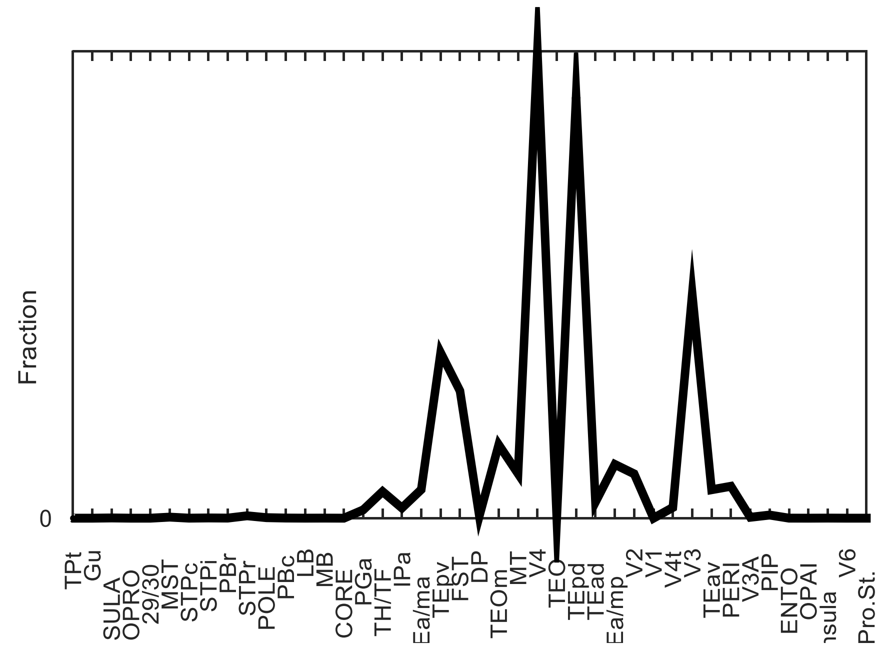
Non-exhaustive connectivity list

- Receives feed-forward topographically organized inputs from V2, V3, V4
- Fewer inputs from V3A, MT
- Backprojections from ITC, parahippocampal gyrus.
- IT sends back-projections to V2, V3, V4
- IT projects to parahippocampal gyrus, frontal eye fields, pre-frontal cortex, amygdala, hippocampus (mostly through entorhinal cortex and parahippocampal gyrus).
- Interhemispheric connections via the corpus callosum (splenium and anterior commissure)
- Subcortical inputs from the thalamus, hypothalamus, locus coeruleus.

Inferior temporal cortex – Projections



Relative weights of posterior IT inputs



Markov et al 2014

Inferior temporal cortex is composed of many subareas

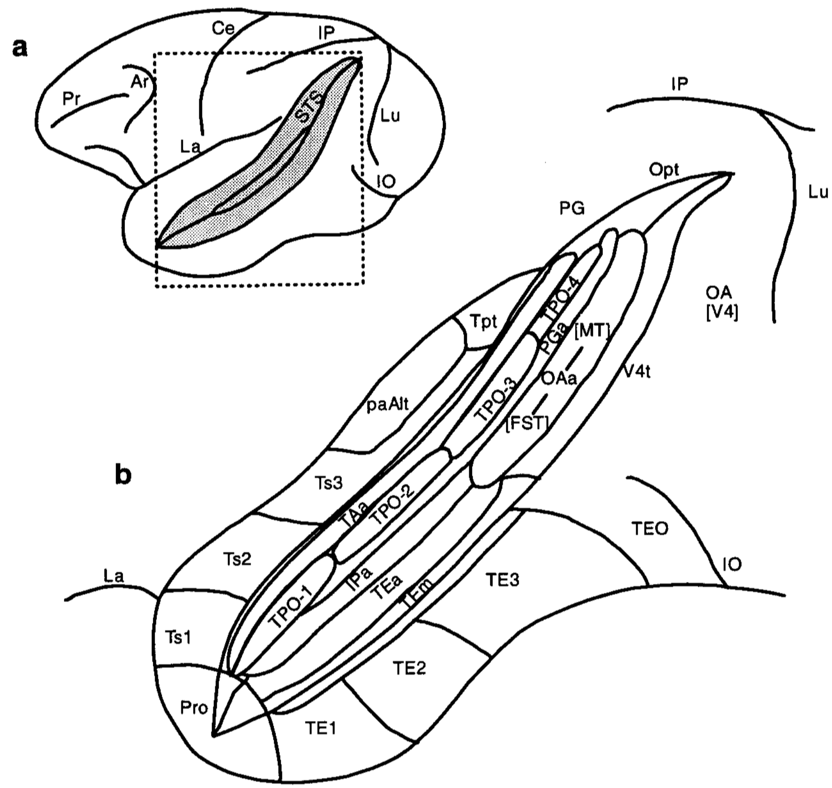


Figure 3 **Subdivision of monkey inferior temporal lobe centered around the superior temporal sulcus (STS). (*a*) Lateral view of the cortical surface with major visible sulci labeled: inferior occipital (IO), lunate (Lu), intraparietal (IP), central (Ce), lateral (Sylvian) fissure (La), arcuate (Ar), and principal (Pr). (*b*) Expanded view of the inferior temporal areas surrounding the STS. [Adapted from Seltzer & Pandya (1994).]**

Logothetis and Sheinberg. Annual Review of Neuroscience 1996

ITC is the last exclusively visual area along the ventral visual pathway

• There are visually responsive areas beyond IT (e.g. perirhinal, entorhinal, hippocampus, amygdala, superior temporal sulcus, striatum, prefrontal cortex).

• However, these other areas are not purely visual.

• Most ITC neurons show visual responses and are either excited or inhibited by simple or complex visual stimuli

•ITC neurons respond to color, orientation, texture, direction of movement or shape

Posterior ITC

- Coarse retinotopic organization

- Almost complete representation of the contralateral visual field

- Receptive field sizes (1.5-2.5 degrees) > V4

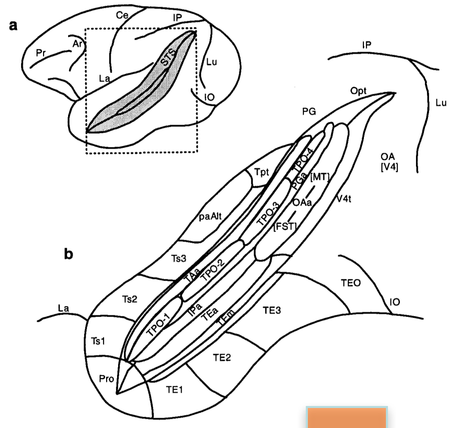
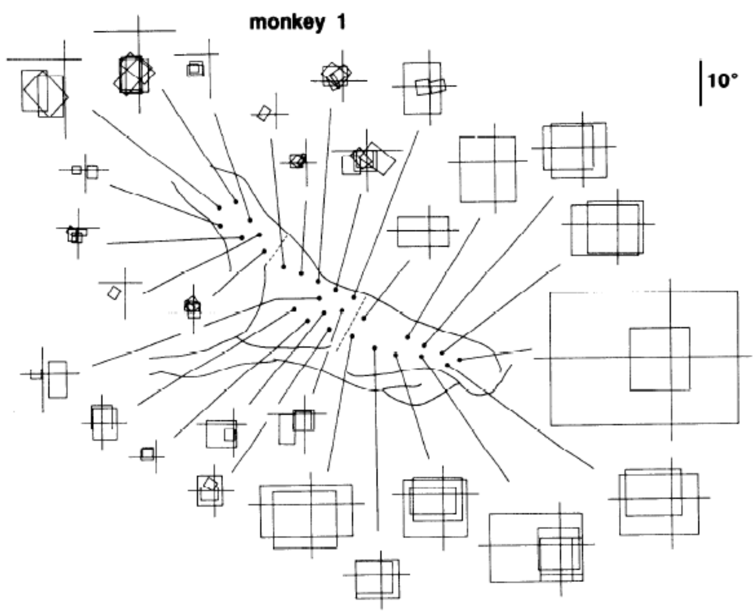


Figure 3 Subdivision of monkey inferior temporal lobe centered around the superior temporal sulcus (STS). (*a*) Lateral view of the cortical surface with major visible sulci labeled: inferior occipital (IO), lunate (Lu), intraparietal (IP), central (Ce), lateral (Sylvian) fissure (La), arcuate (Ar), and principal (Pr). (*b*) Expanded view of the inferior temporal areas surrounding the STS. [Adapted from Seltzer & Pandya (1994).]

EUCALY KOBATAKE AND KEIJI TANAKA



IT cells closer to V1 (more posterior) have smaller receptive fields.

RFs frequently include the fovea, and may extend to the contralateral hemifield.

Central and anterior ITC

No clear retinotopy

However, there are have been suggestions of feature topography (Tanaka 1996)

Large receptive fields (ipsilateral, contralateral or bilateral).

Most receptive fields include the fovea. Foveal stimuli elicit stronger responses.

Receptive fields of up to 30 degrees (Tanaka 2003; see however DiCarlo et al 2004)

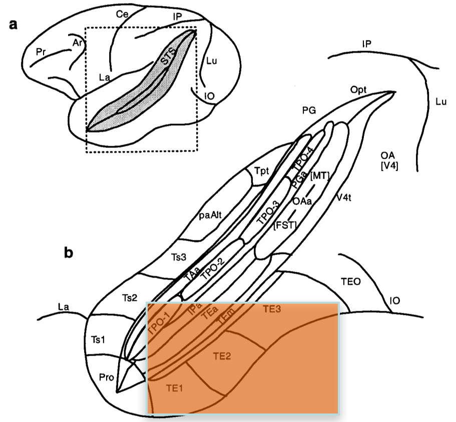
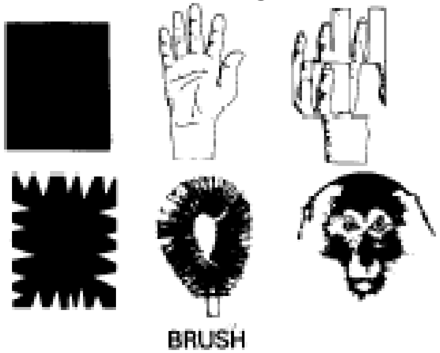


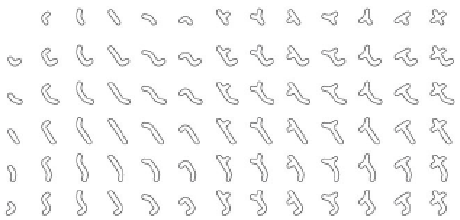
Figure 3 Subdivision of monkey inferior temporal lobe centered around the superior temporal sulcus (STS). (*a*) Lateral view of the cortical surface with major visible sulci labeled: inferior occipital (IO), lunate (Lu), intraparietal (IP), central (Ce), lateral (Sylvian) fissure (La), arcuate (Ar), and principal (Pr). (*b*) Expanded view of the inferior temporal areas surrounding the STS. [Adapted from Seltzer & Pandya (1994).]

ITC neurons respond to a large variety of complex shapes

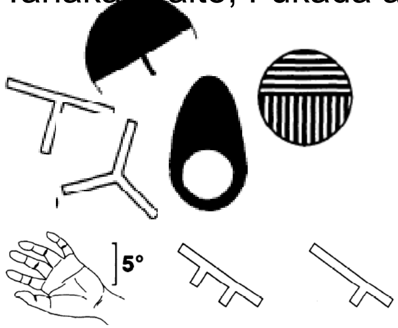
Desimone, Albright, Gross and Bruce



Connor and others

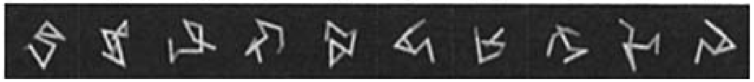


Tanaka, Saito, Fukada and Moriya



Kiani, Esteky, Mirpour and Tanaka

Logothetis, Pauls and Poggio



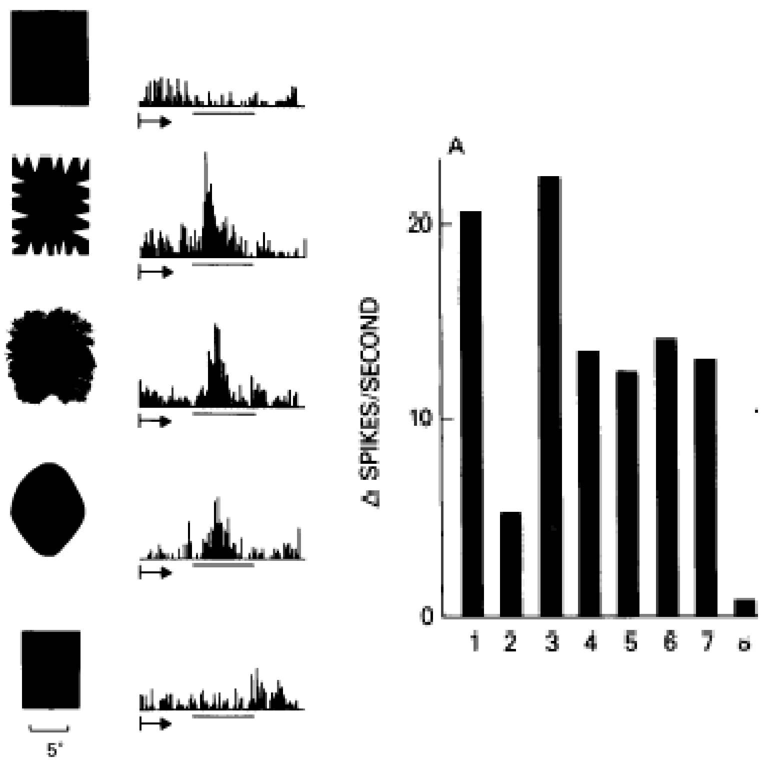
Hung, Kreiman, Poggio and DiCarlo

Selective responses to almost every kind of stimulus tried. For example:

- Faces, hands, body parts (Gross et al 1969; Desimone et al 1984; Perrett et al 1982; Young and Yamane 1992; Rolls 1984)
- Parametric shape descriptors (Schwartz et al 1983)
- “Abstract” 2D patterns (Richmond et al 1987)
- Paperclips (Logothetis et al 1995)

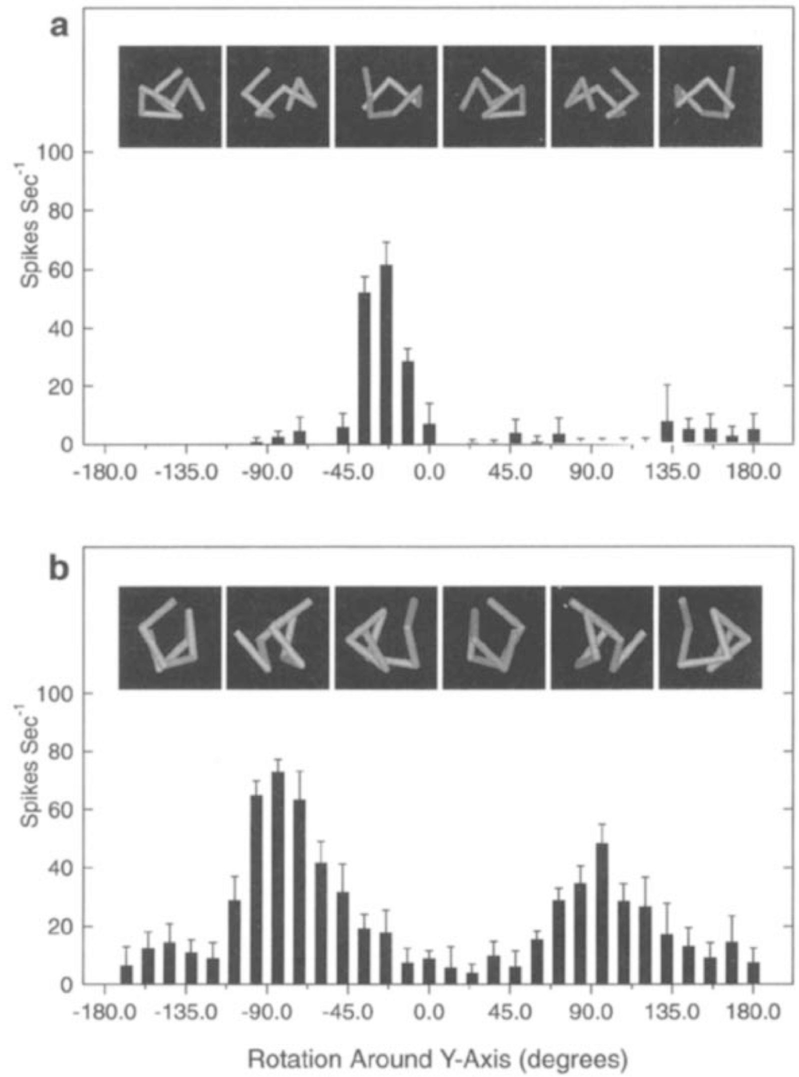
Shape can be defined in many different ways.

Example ITC responses



IT cells emit different number of action potentials (“spikes”) in response to different images...

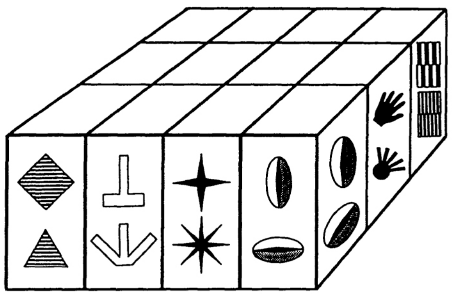
Desimone et al. 1984



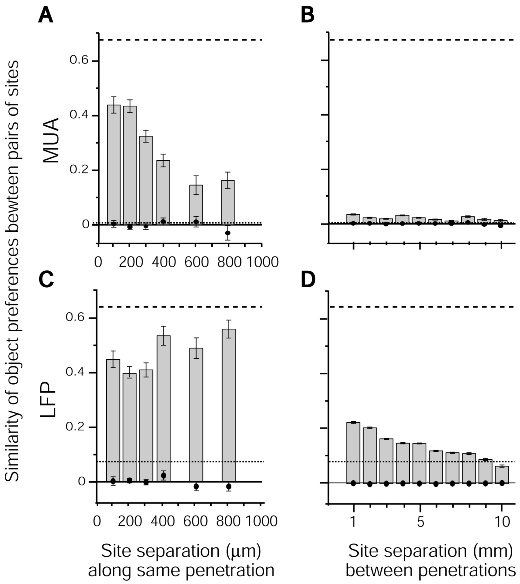
Logothetis and Sheinberg.
Annual Review of Neuroscience 1996

Feature topography in ITC

Tanaka. Science 1993



Kreiman et al, Neuron 1996



In earlier visual areas, retinotopy refers to an organized mapping of the visual field

The columnar organization in earlier visual areas refers to the similar feature preferences of neurons along a penetration tangential to cortex

In ITC Tanaka and others have argued for similar object feature preferences along columns tangential to the cortical structure (Fujita et al 1992; Tanaka 1993; Young 1993; Gawne and Richmod 1993; Kobatake and Tanaka 1994).

Size invariance in ITC

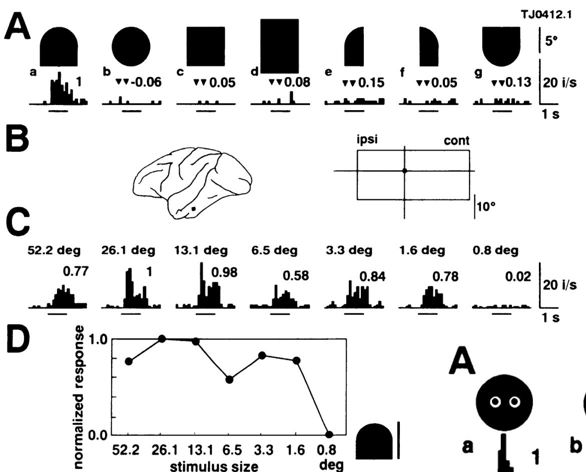


FIG. 4. Example of broadly tuned cells. *A*: optimal stimulus was a dome-shaped structure shown in *a*. *B*: recording site, the receptive field, and site of the stimulus presentation. *C*: responses to different sizes of the optimal stimulus. Comparably strong responses were evoked at a wide range of size (52.2–1.6° in height). *D*: magnitude of the responses normalized by that of the response to the stimulus 26.1° in size.

Ito et al .
J. Neurophys.
1995

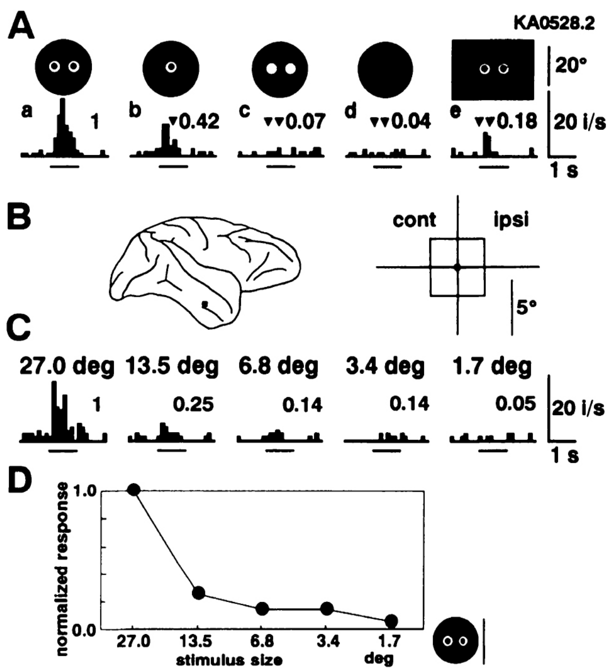
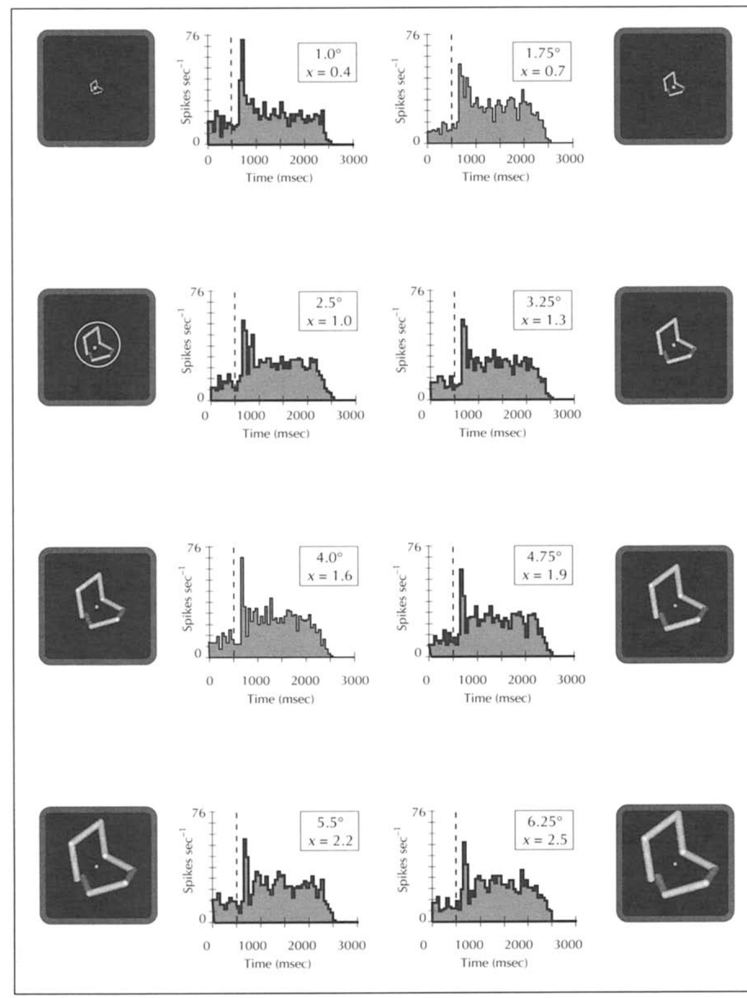


FIG. 5. Example of cells that maximally responded to the largest size of the optimal stimulus. *A*: optimal stimulus of the cell was a pair of white rings on a black base. *B*: recording site, the receptive field, and site of the



Logothetis et al 1995

Object selective responses can be invariant to size even if absolute rates are not

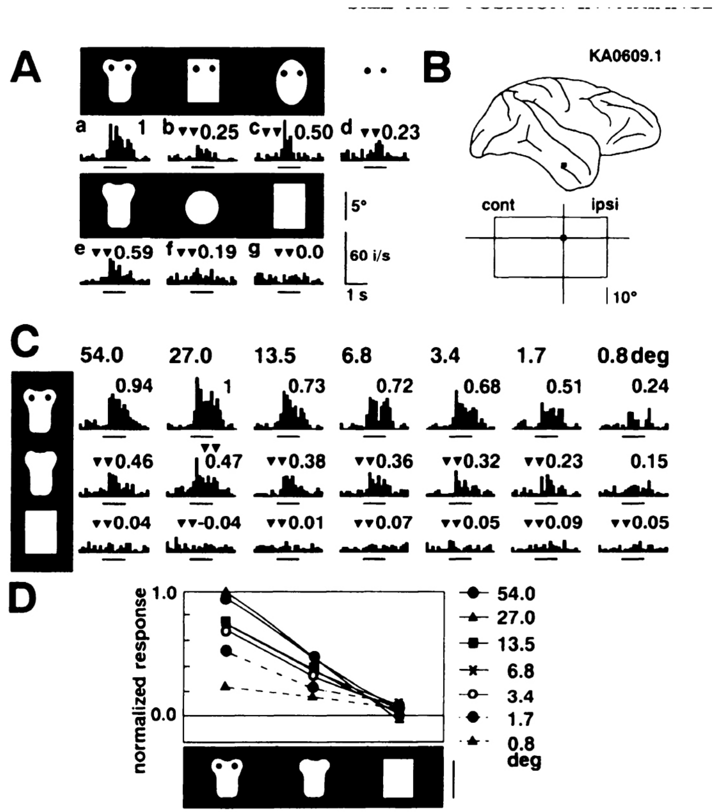
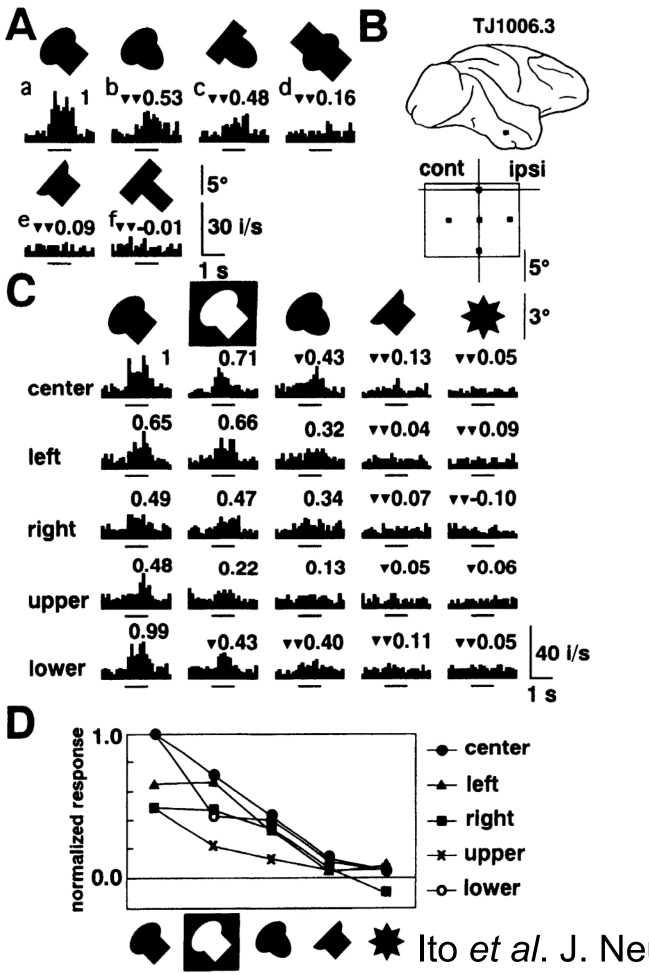


FIG. 9. Example of cells in which the selectivity for shape was preserved throughout the entire range of size changes. *A*: optimal stimulus of the cell was a pair of black dots on a white base of a complex contour shape as shown in *a*. *B*: recording site, the receptive field, and site of the stimulus presentation. *C*: the cell showed response to the optimal stimulus in a wide range of size (54.0–0.8° in height). Response to the white contour of complex shape without dots was ~ ½ of that to the optimal stimulus for all sizes. Arrowheads indicate the results of statistical comparison of responses with different stimuli at each size. *D*: magnitude of the responses normalized by that of the response to the optimal stimulus at 27°.

Ito et al . J. Neurophys. 1995

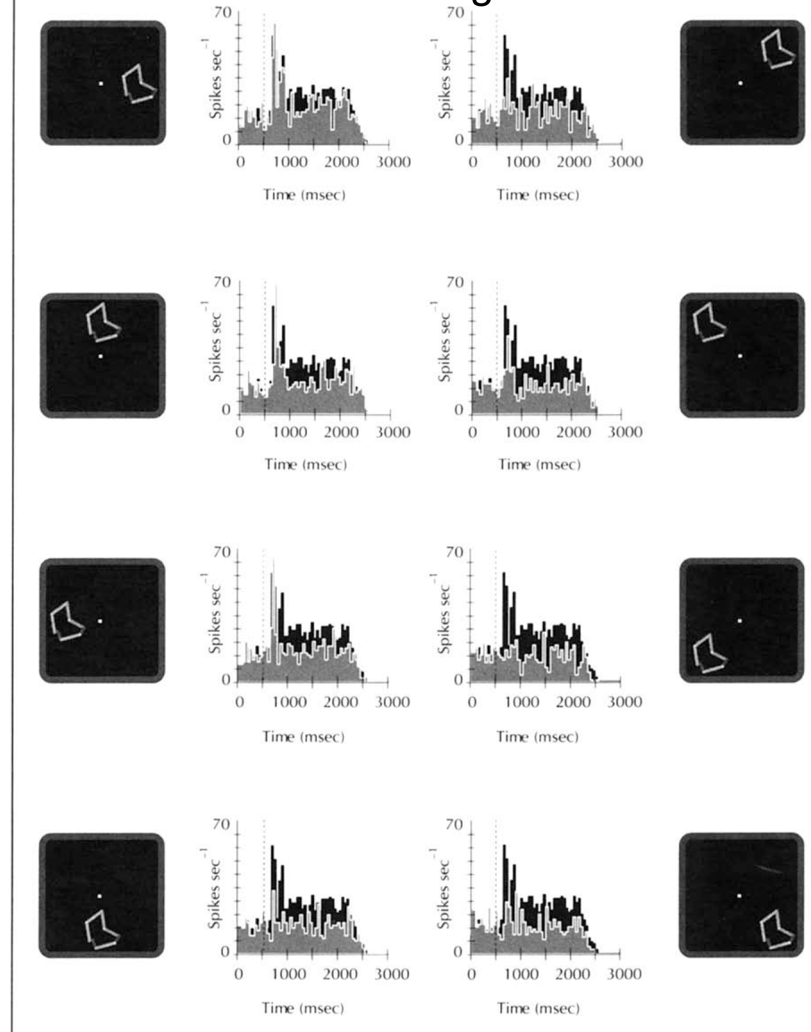
Position invariance in ITC



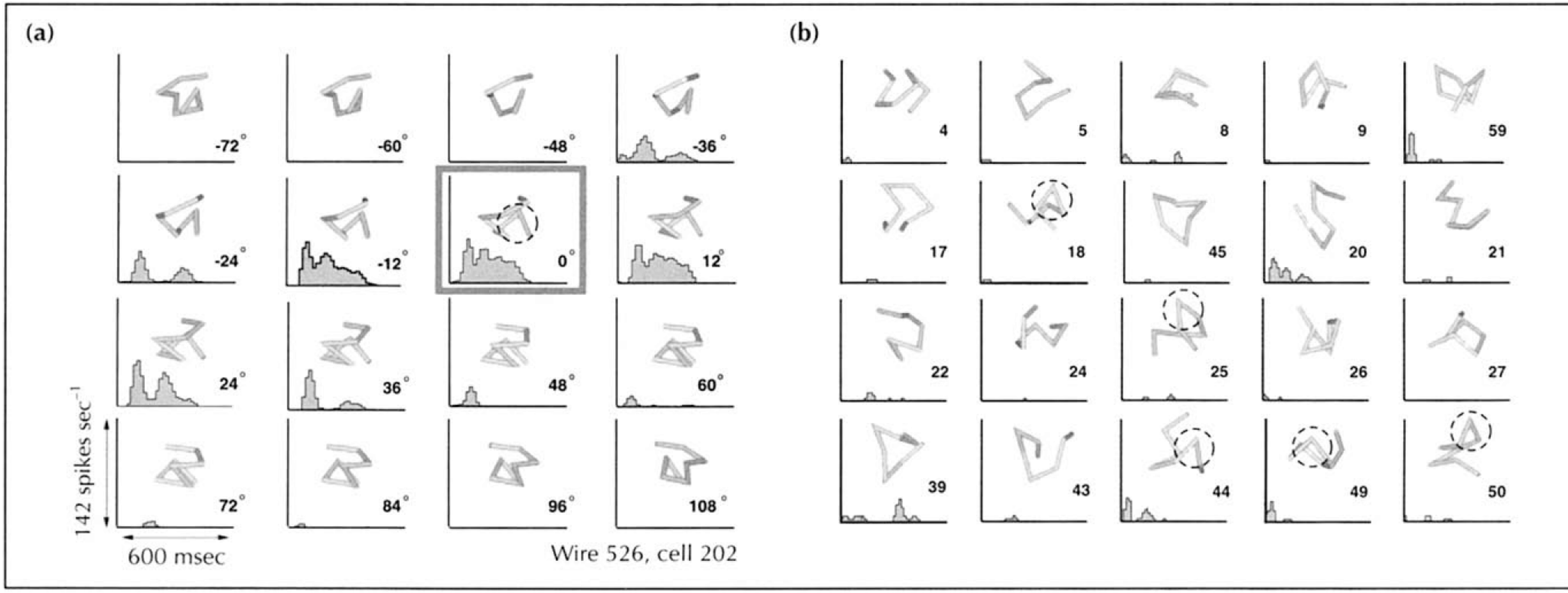
Ito *et al*. J. Neurophys. 1995

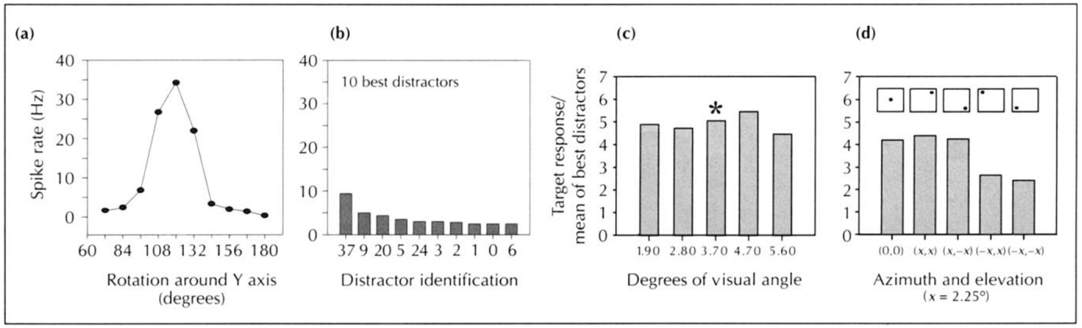
FIG. 11. Example of cells in which the shape selectivity was preserved throughout the receptive field. *A*: optimal stimulus of the cell was a combination of an ellipse in the top left and a rectangle in the bottom right. *B*: recording site, the receptive field, and the 5 sites of the stimulus presenta-

Logothetis et al 1995



Tolerance to viewpoint and illumination changes





Logothetis et al 1995

Rotation invariance in ITC

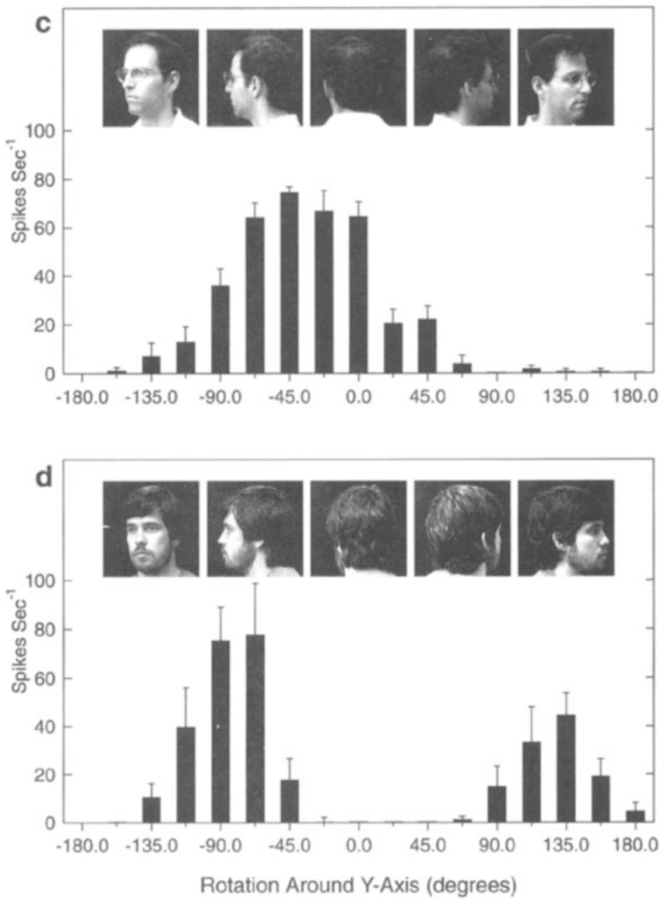
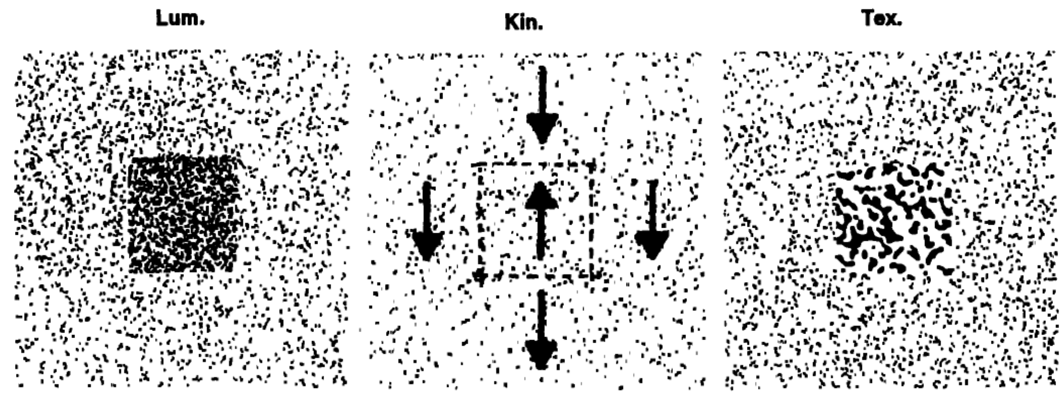
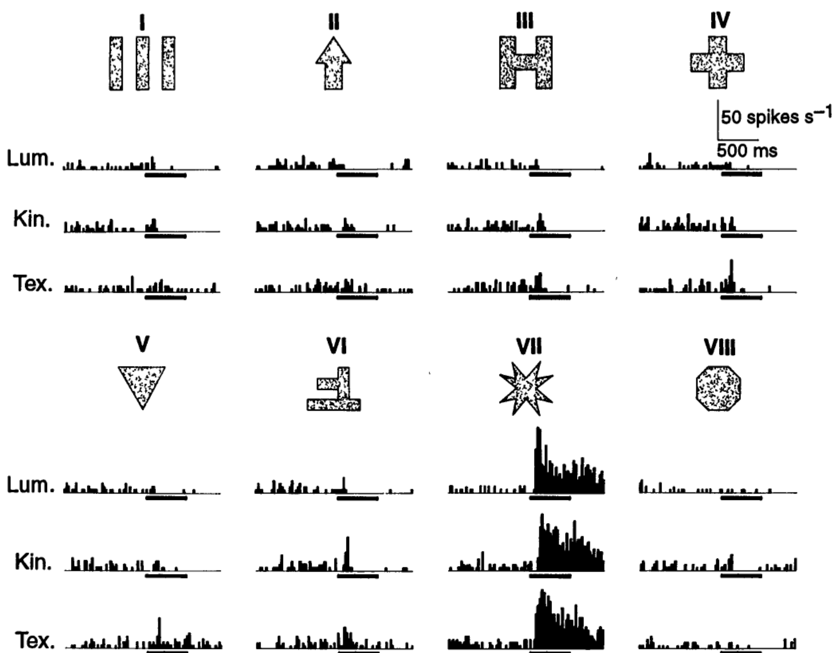


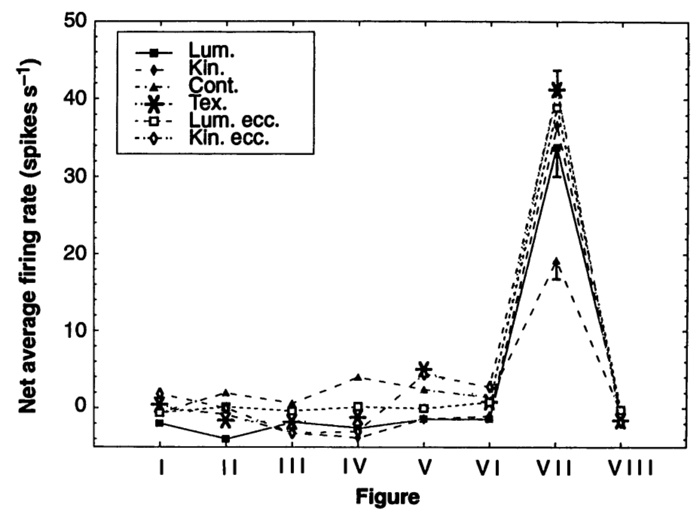
Figure 4 Four different IT neurons selective for views of wires and faces. (*a–d*) Two of the neurons shown here responded maximally for a single view of an object (*a*, *c*), and response magnitude decreased gradually as the object was rotated in depth away from the preferred view. Figure 4*b* shows an example of a cell responding to two views of a wire object separated by 180°, and Figure 4*d* shows data from a cell that exhibited its maximum response for the left-facing profile of a head and nearly the same response for the right-facing profile. (Error bars indicate standard deviations of mean response rates.)

Logothetis and Sheinberg.
Annual Review of Neuroscience 1996

Cue invariance in the responses of ITC

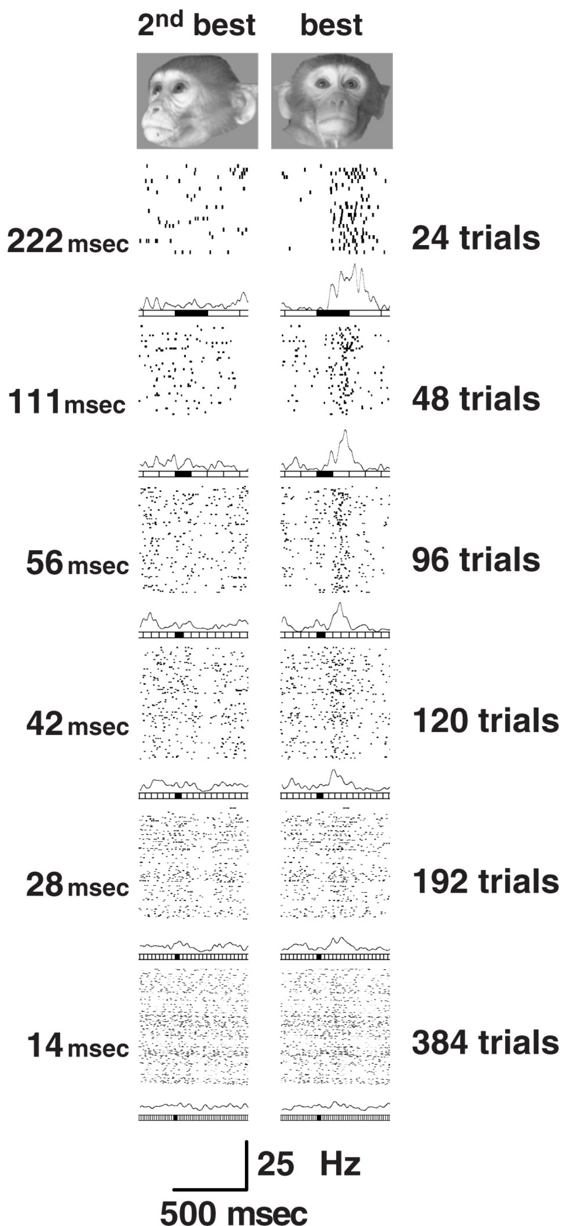




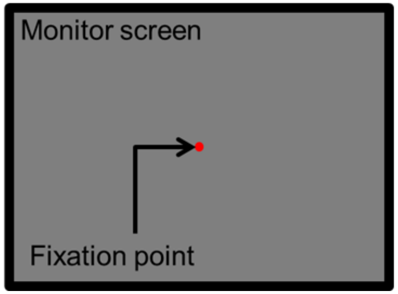


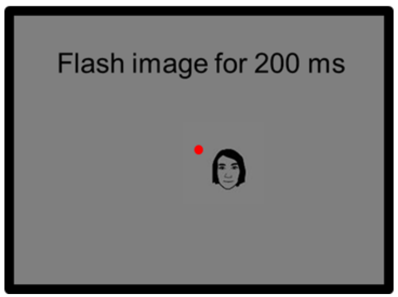
Sary et al Science 1993

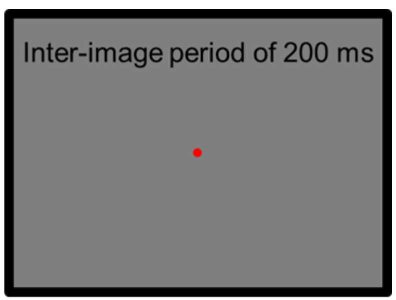
Very fast responses in ITC

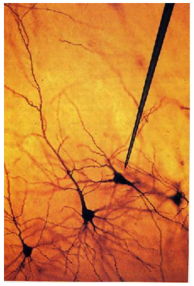


Reading out the mind

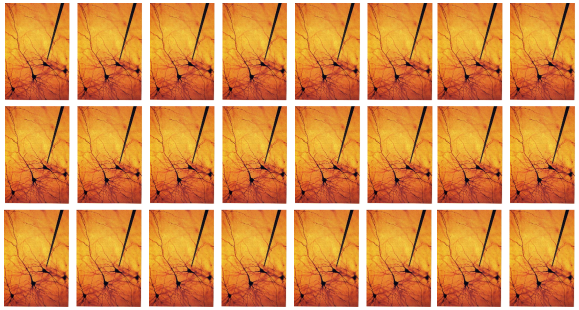






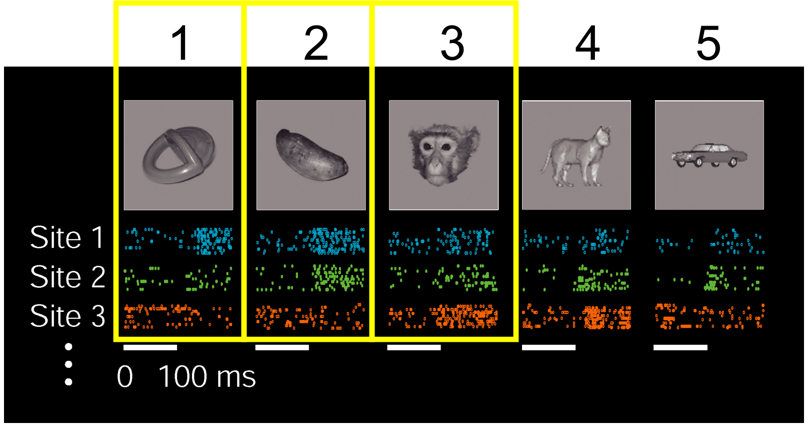


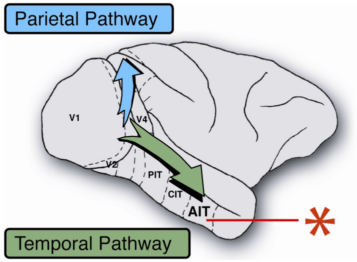
Virtually all studies above were conducted using single-electrode experiments



What do we do when we have many, many electrodes?

Reading out object information from a small population of inferior temporal cortex neurons

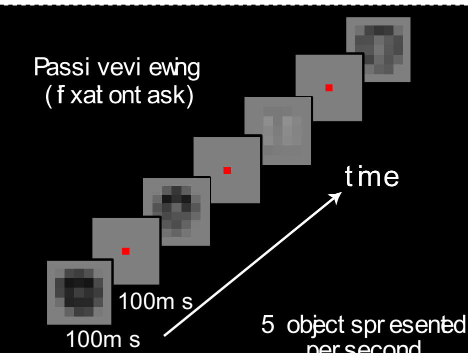


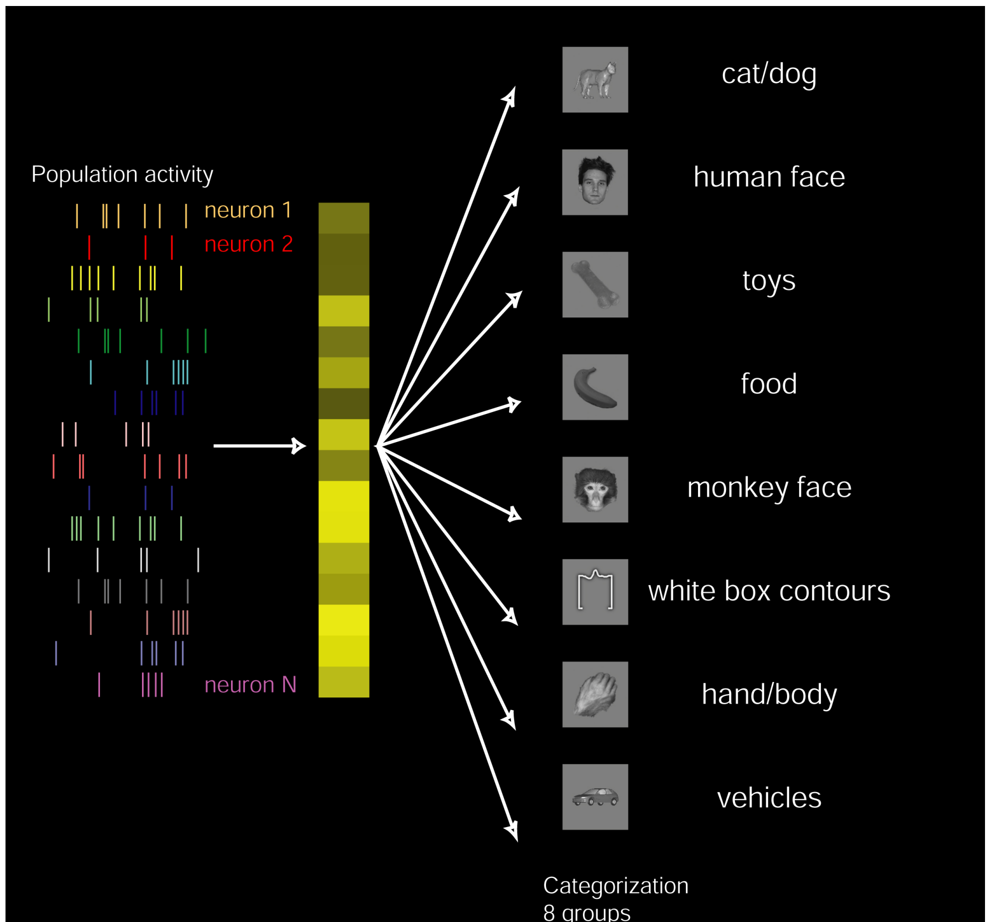


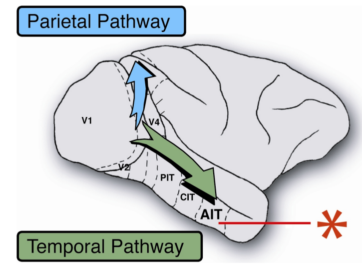
Neuron 1	Neuron 2	Neuron 3	Object
Yes	No	No	1
Yes	Yes	No	2
Yes	Yes	Yes	3

Chou Hung, Gabriel Kreiman, Jim DiCarlo, Tomaso Poggio

Using machine learning to read out the “monkey’s mind”



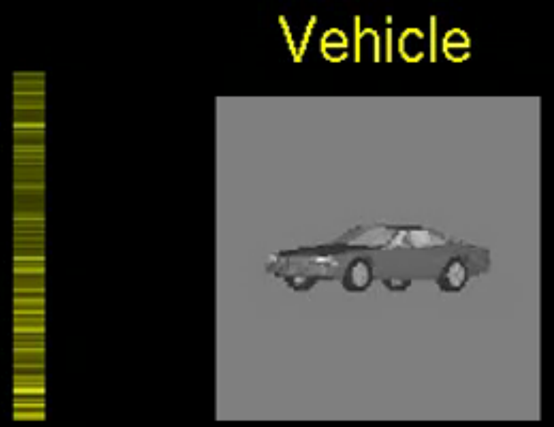




Video: decoding performance

Neuronal population activity

Classifier prediction



256 units
Categorization: ~90% (chance = 12.5%)
Identification: ~70% (chance = 1.3%)

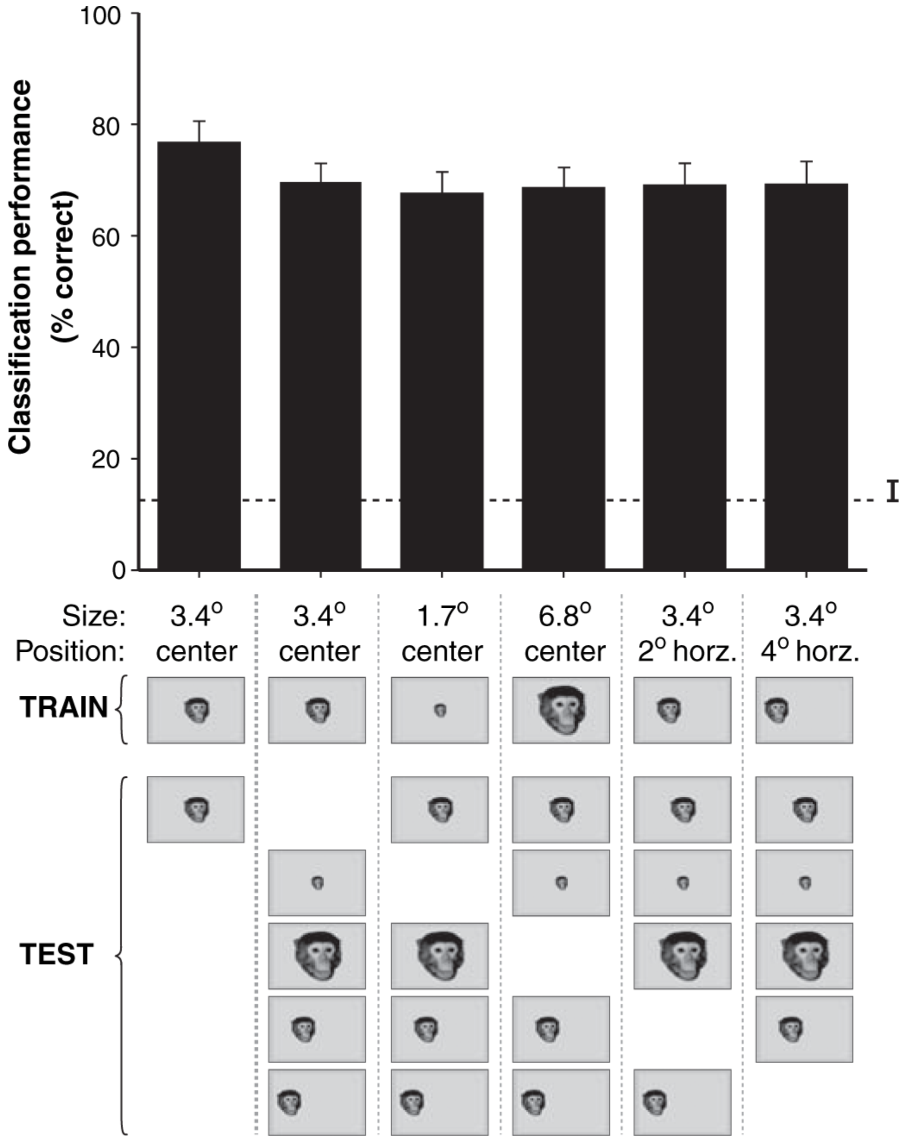
Video speed: 1 frame/sec
Actual presentation rate: 5 objects/sec

Categorization

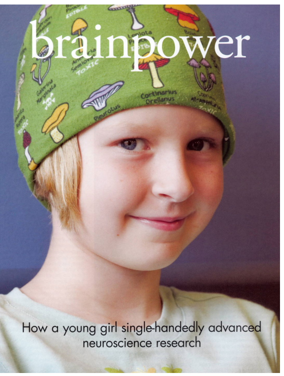
- Toy
- Body
- Human Face
- Monkey Face
- Vehicle
- Food
- Box
- Cat/Dog

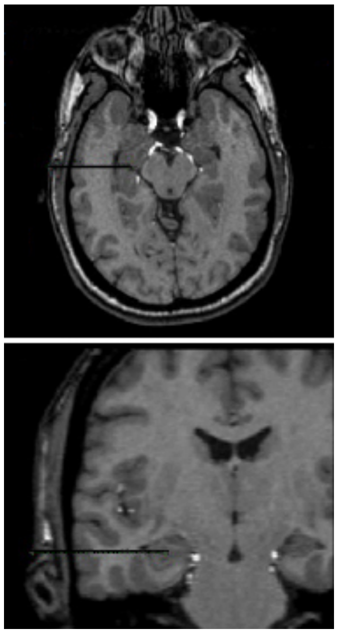
Generalization in read-out performance

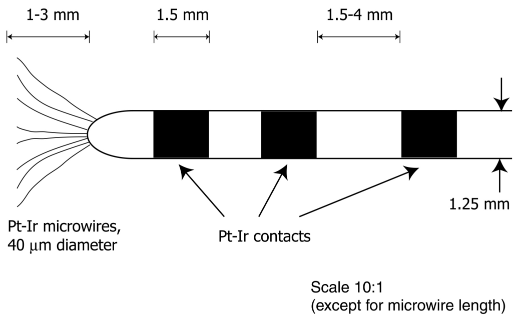
"To learn is to forget a difference"
JL Borges

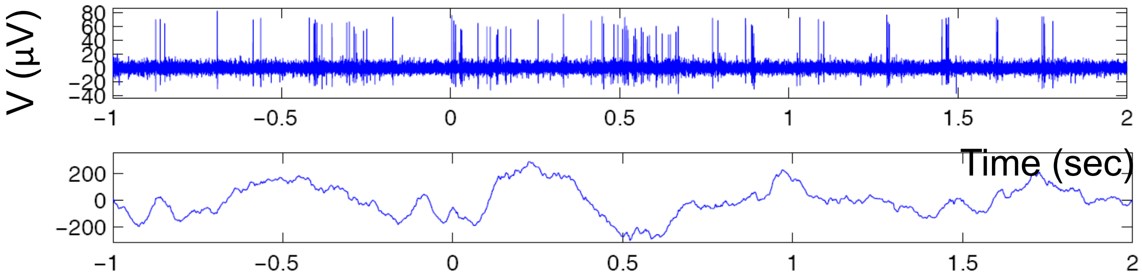


Neurophysiological recordings in the human brain



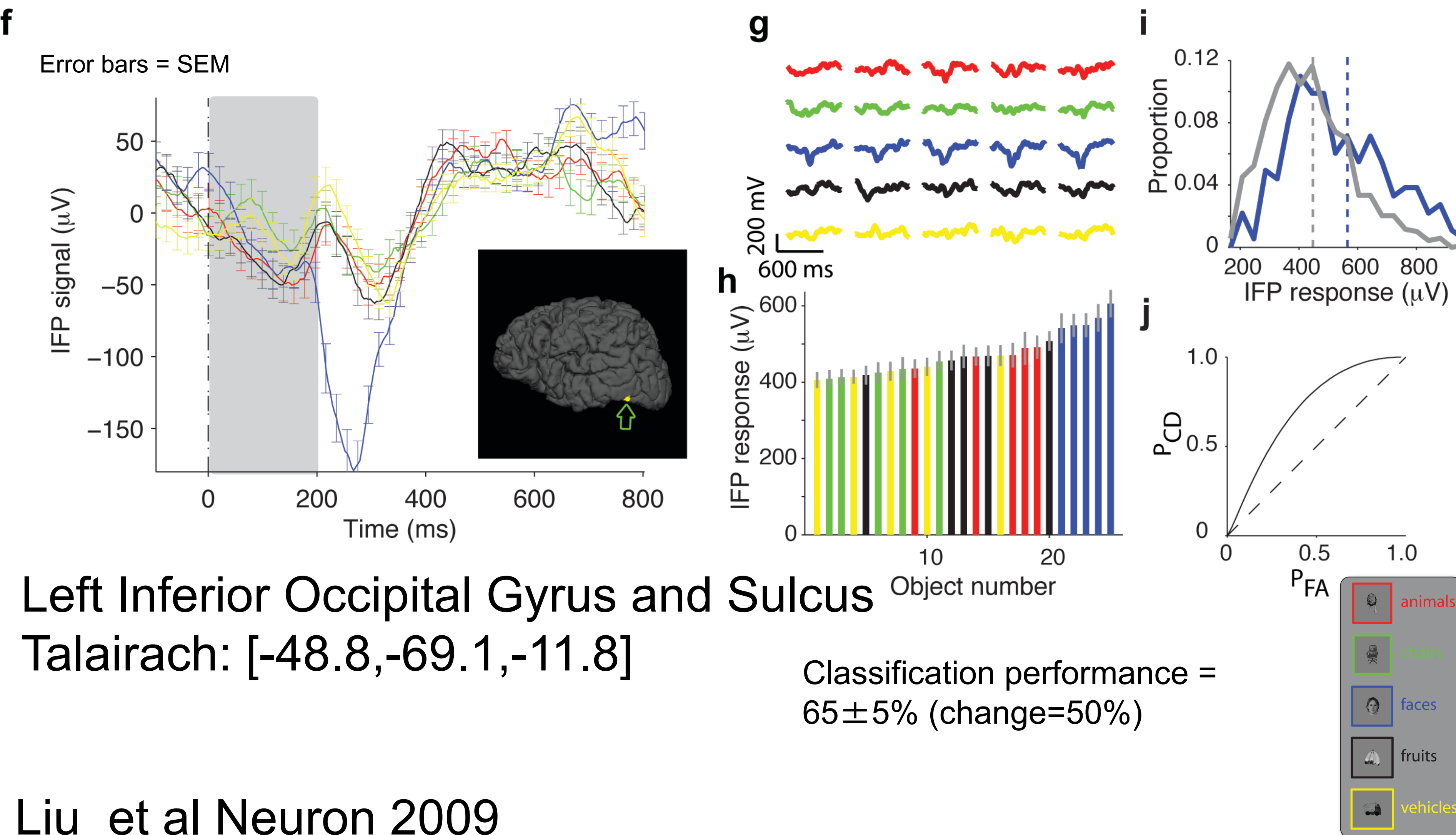






- Patients with pharmacologically intractable epilepsy
- Multiple electrodes implanted to localize seizure focus
- Targets typically include the temporal lobe (inferior temporal cortex, fusiform gyrus), medial temporal lobe (hippocampus, entorhinal cortex, amygdala and parahippocampal gyrus)
- Patients stay in the hospital for about 7-10 days

Selectivity in human visual cortex - Example



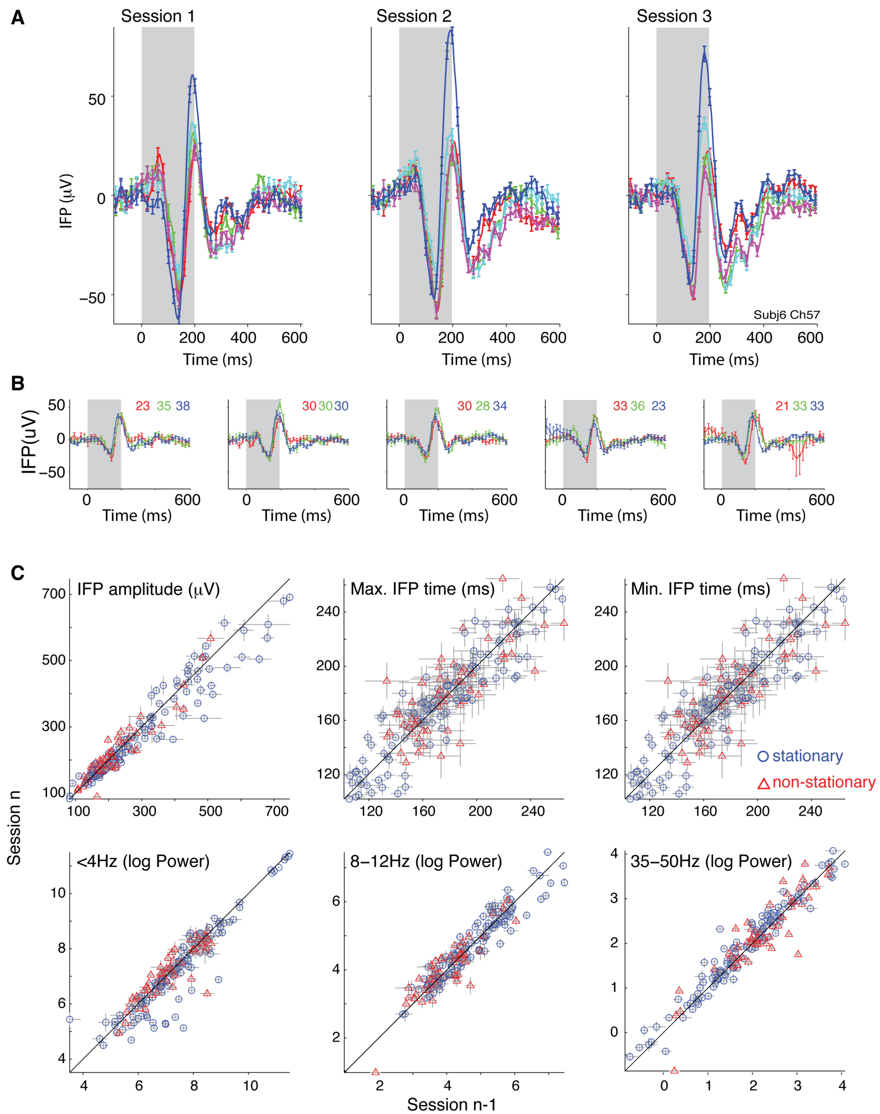
Left Inferior Occipital Gyrus and Sulcus

Talairach: [-48.8,-69.1,-11.8]

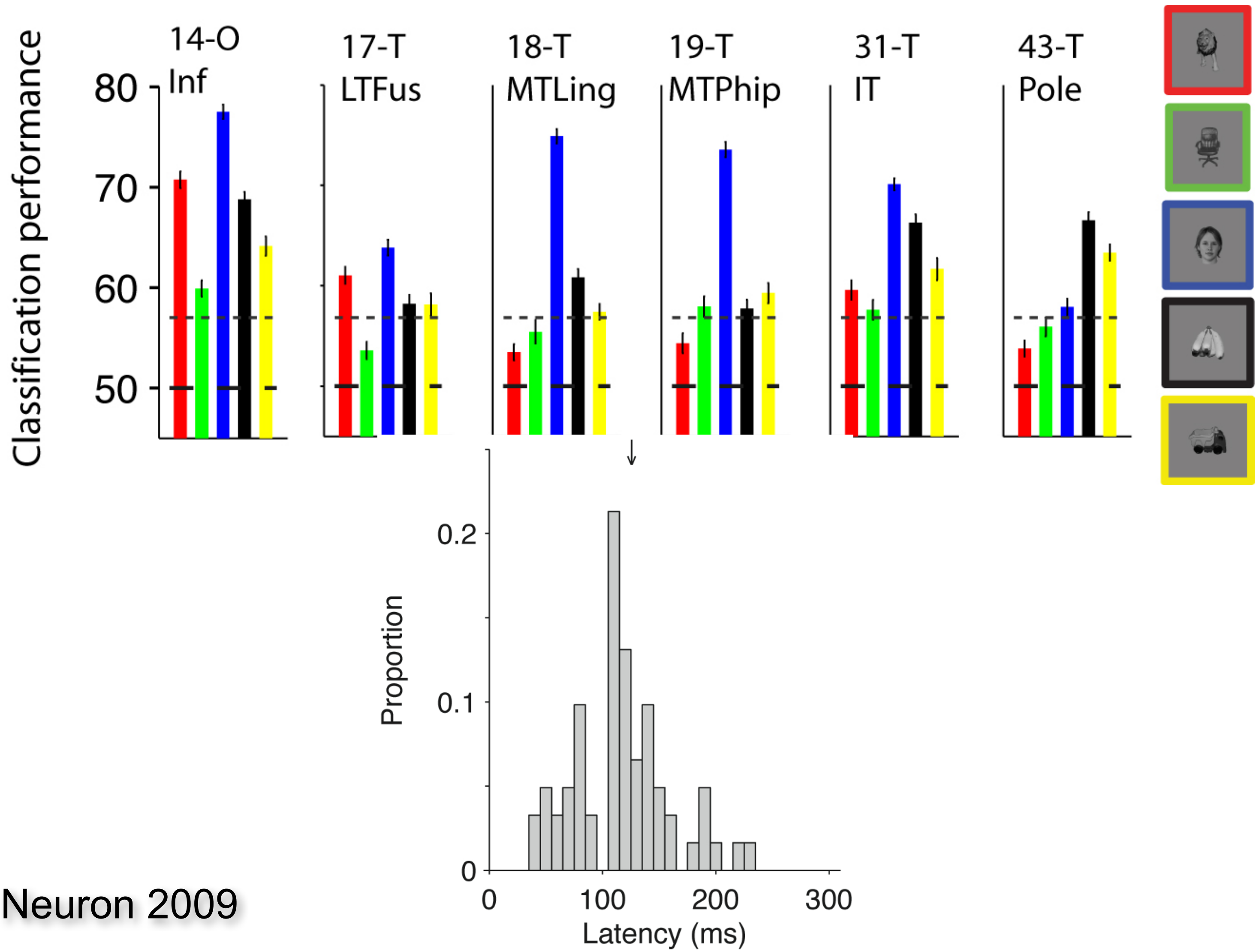
Classification performance = 65±5% (change=50%)

Liu et al Neuron 2009

Selective responses are stable over time scales of days

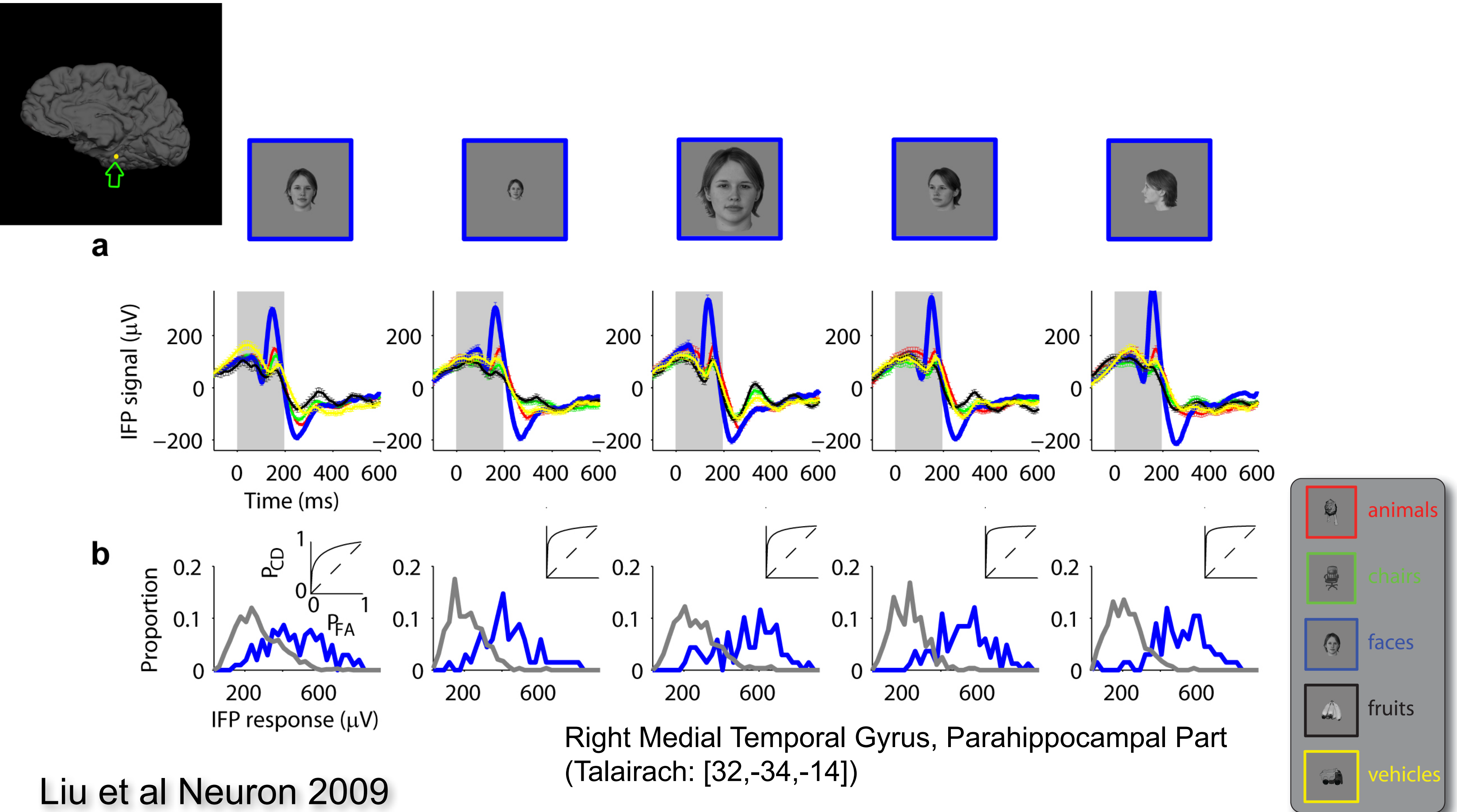


Locations and timing in human ventral visual cortex



Liu et al Neuron 2009

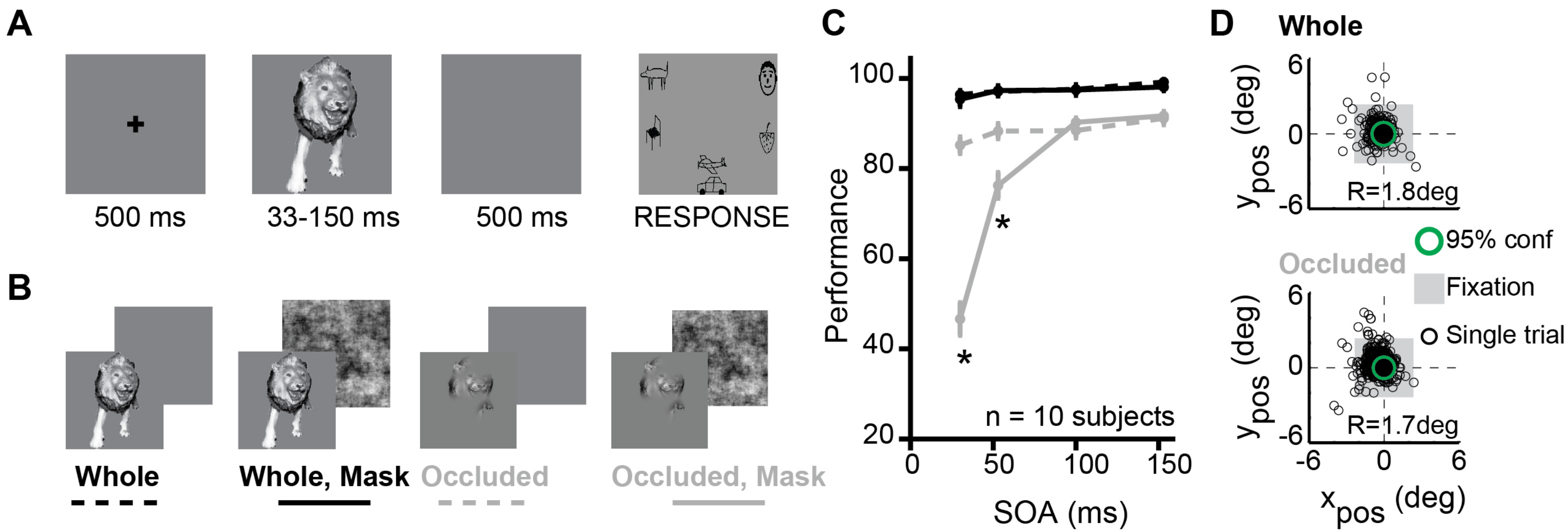
Tolerance to scale and rotation changes - Example

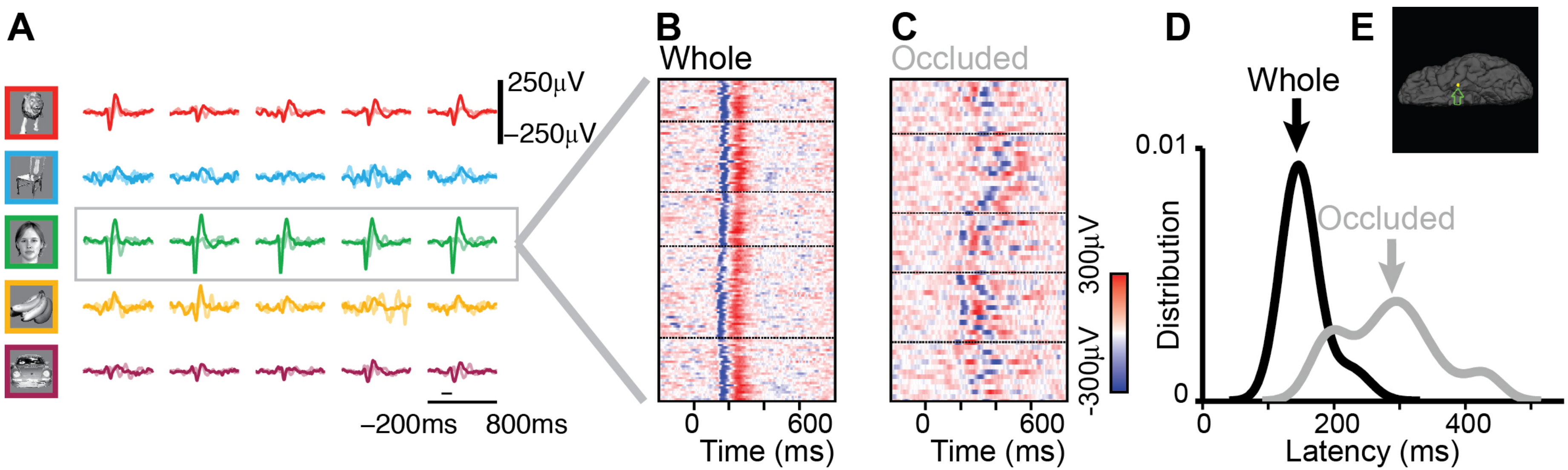


Right Medial Temporal Gyrus, Parahippocampal Part
(Talairach: [32,-34,-14])

Liu et al Neuron 2009

Tolerance to object occlusion - Example





Further reading

- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. Annual Review of Neuroscience, 19, 577-621.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. Annual Review of Neuroscience, 19, 109-139.

<u>Original articles cited in class (see lecture notes for complete list)</u>

- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. Current Biology, 5(5), 552-563.
- Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. J Neurophysiol, 73(1), 218-226.
- Hung, C., Kreiman, G., Poggio, T., & DiCarlo, J. (2005). Fast Read-out of Object Identity from Macaque Inferior Temporal Cortex. Science, 310, 863-866.
- Liu H, Agam Y, Madsen J, Kreiman G. (2009) Timing, timing, timing: Fast decoding of object inforrmation from intracranial field potentials in human visual cortex. Neuron 62:281-290