DT2350, Lecture 4: Introduction to vision, including perception of objects and scenes

Dr. Ginevra Castellano

Department of Information Technology

Uppsala University

Email: ginevra.castellano@it.uu.se

Textbooks

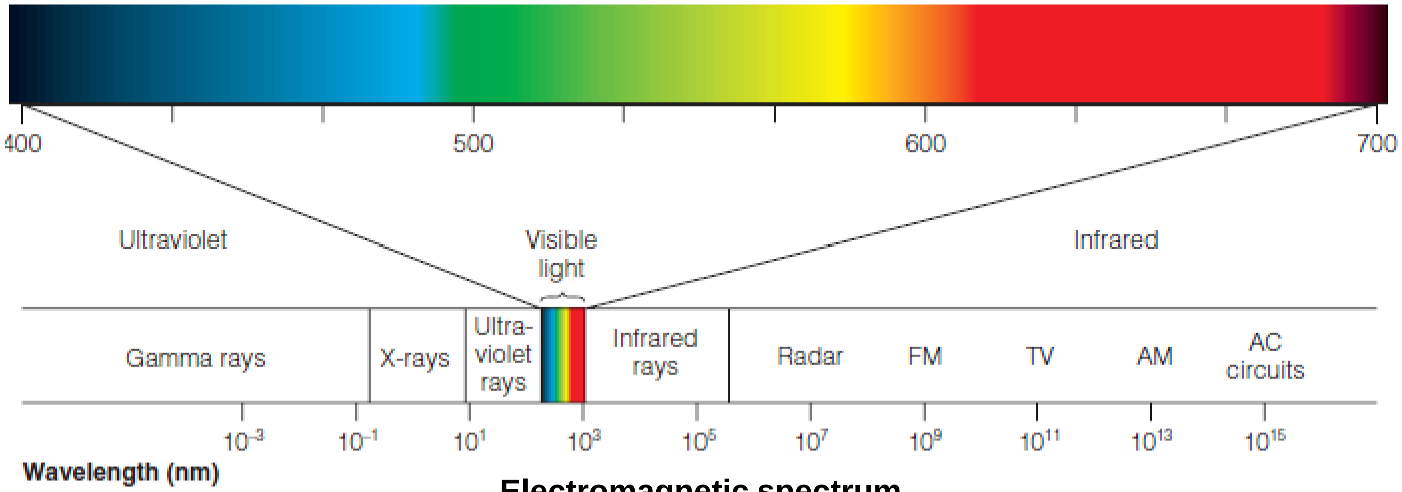
- Goldstein, E. (2009). Sensation and Perception.
- Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

Overview

- Introduction to vision
  - What happens at the very beginning of the visual system
- The visual cortex and beyond
  - Processes that occur in the visual areas of the brain
- Perceiving objects and scenes
  - How do we distinguish objects from their background?

Introduction to vision - Focusing light onto the retina

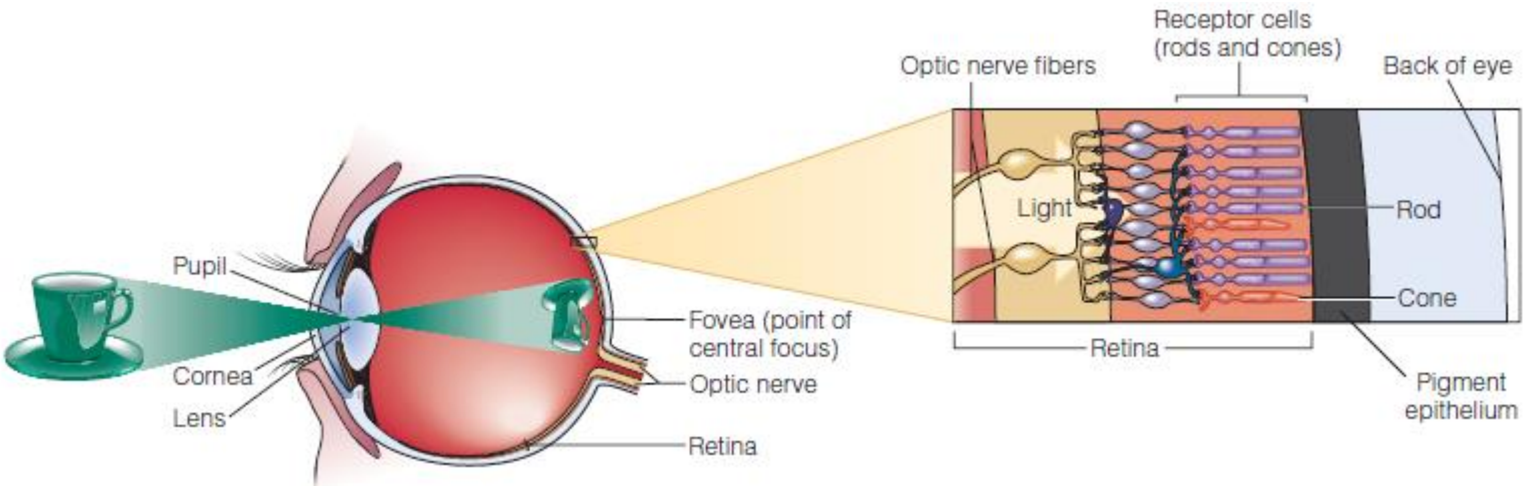
- Vision begins when visible light is reflected from objects into the eye
  - Wavelength
  - Photons



Electromagnetic spectrum

Introduction to vision - Focusing light onto the retina

- Light entering through the pupil is focused by the cornea and lens to form images on the retina
  - The lens can change its shape to adjust the eye's focus for stimuli located at different distances (*accomodation*)



Introduction to vision - Transforming light onto electricity

- Transduction is carried out by receptors (rods and cones) via visual pigments molecules

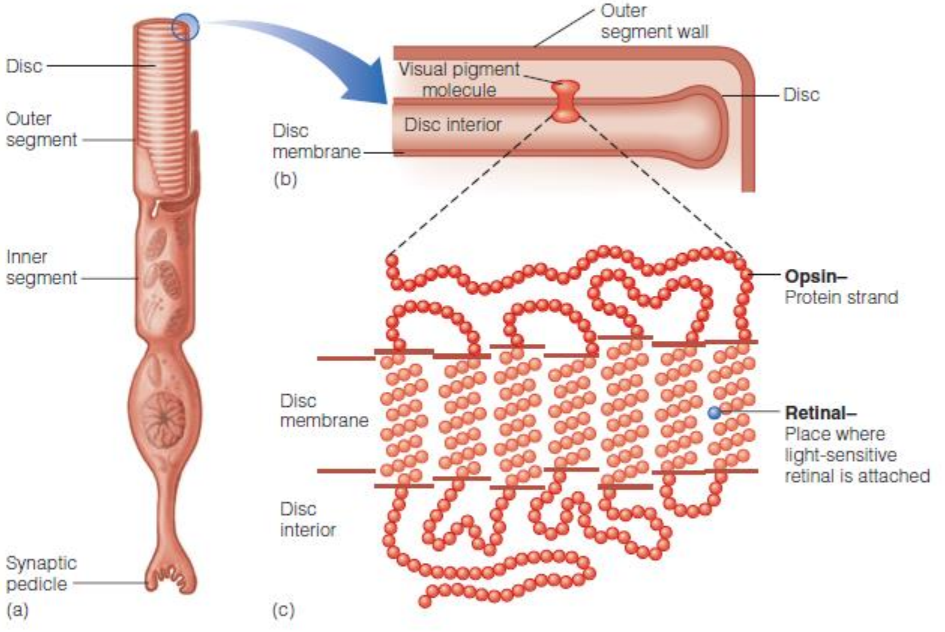
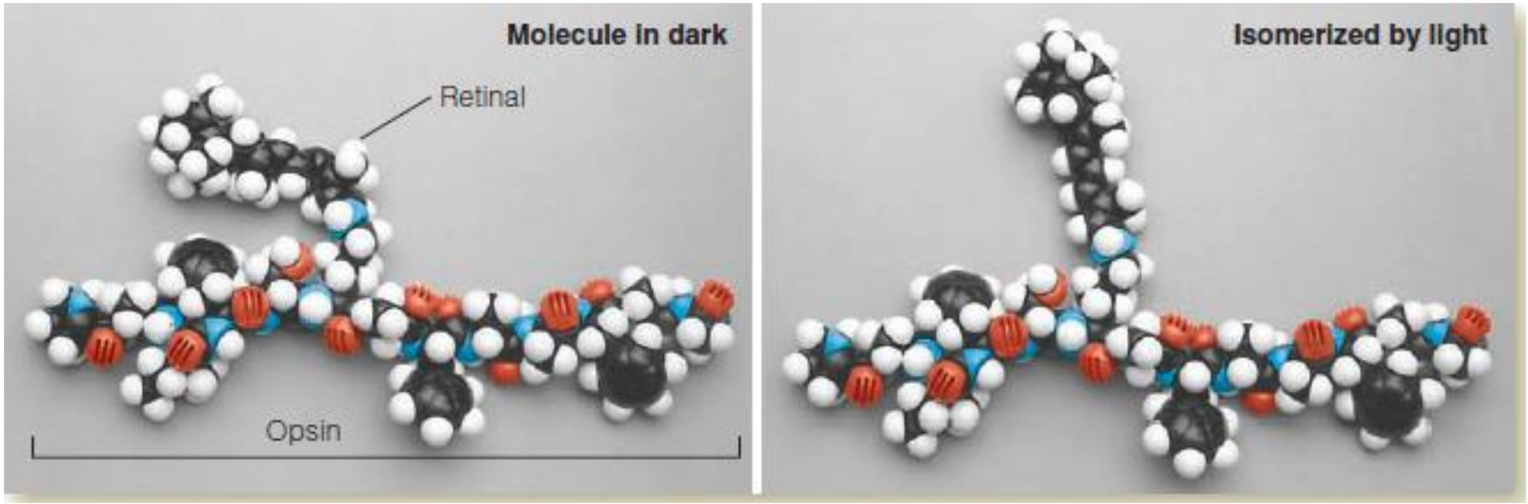


Figure 3.6 (a) Rod receptor showing discs in the outer segment. (b) Close-up of one disc showing one visual pigment molecule in the membrane. (c) Close-up showing how the protein opsin in one visual pigment molecule crosses the disc membrane seven times. The light-sensitive retinal molecule is attached to the opsin at the place indicated.

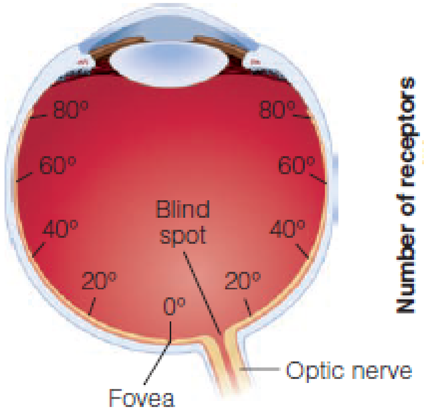
Introduction to vision - Transforming light onto electricity

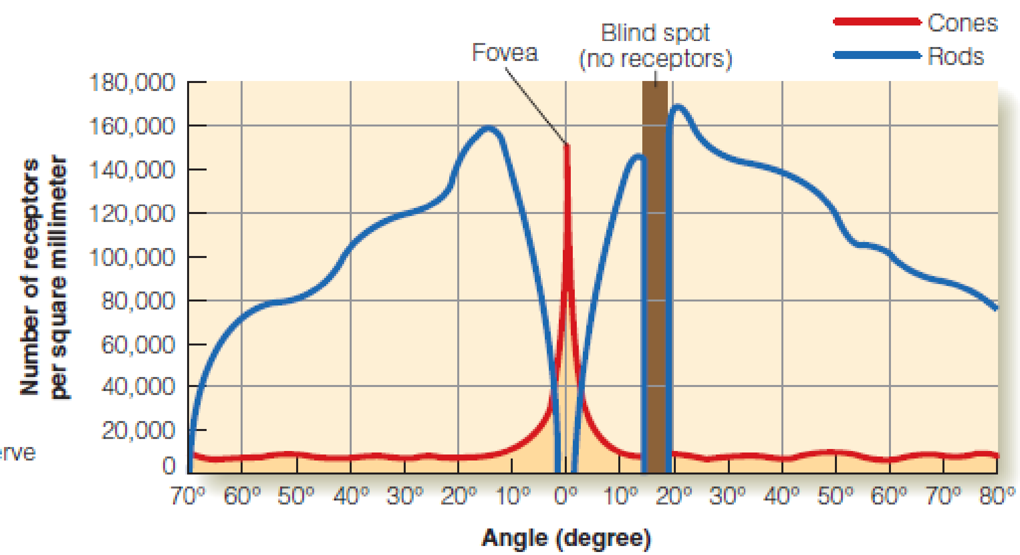
- Retinal is the part of the visual pigment that is sensitive to light
- Isomerization -> activation of the entire receptor



Introduction to vision - Pigments and perception

- Distribution of rods and cones (120 vs. 6 million)
  - Fovea contains only cones
  - When we look directly at an object, its image falls on the fovea
  - Blind spot: no receptors where the optic nerve leaves the eye

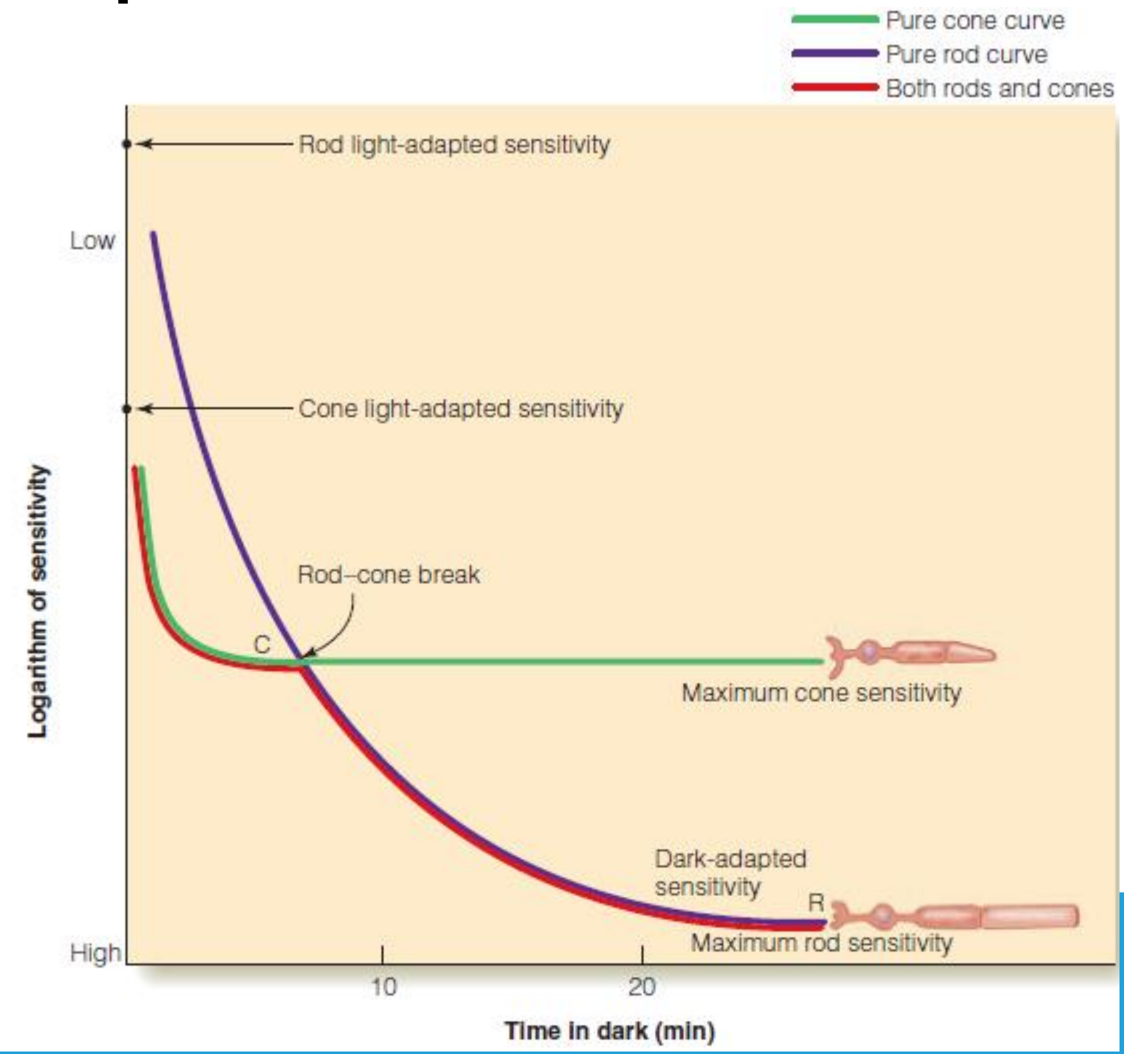




Introduction to vision - Pigments and perception

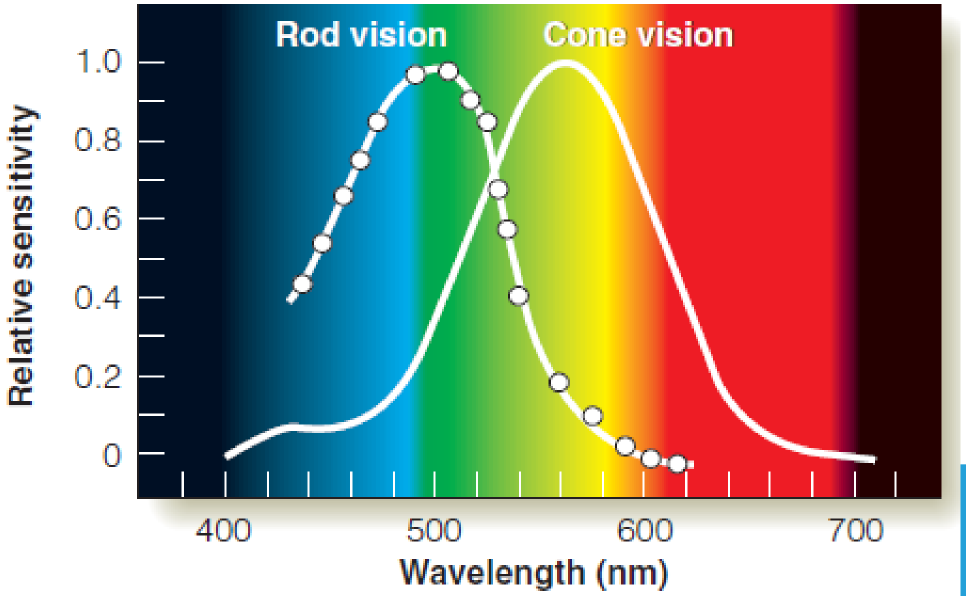
- Dark adaptation of the rods and cones
  - Increase eye sensitivity in the dark
- Two different stages
  - Initial rapid stage
    - Cone receptors
  - Second slower stage
    - Rod receptors

Introduction to vision - Pigments and perception



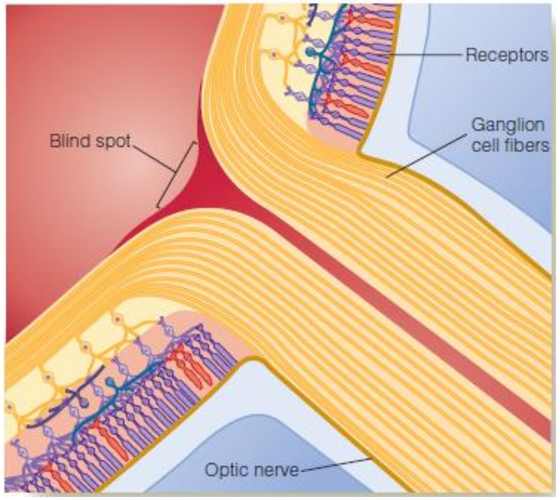
Introduction to vision - Pigments and perception

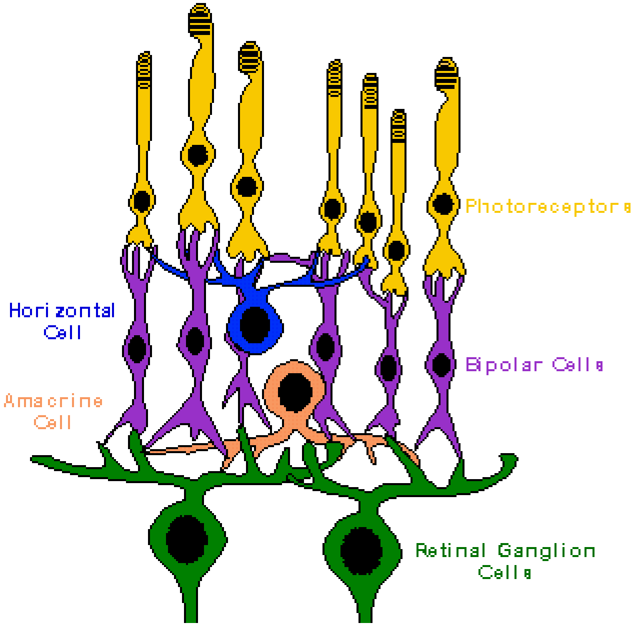
- Spectral sensitivity curves for rods and cones
  - Rods are more sensitive to short wavelength light than are the cones
  - The cones are more sensitive in the light, and the rods are more sensitive in the dark



Introduction to vision - Neural convergence and perception

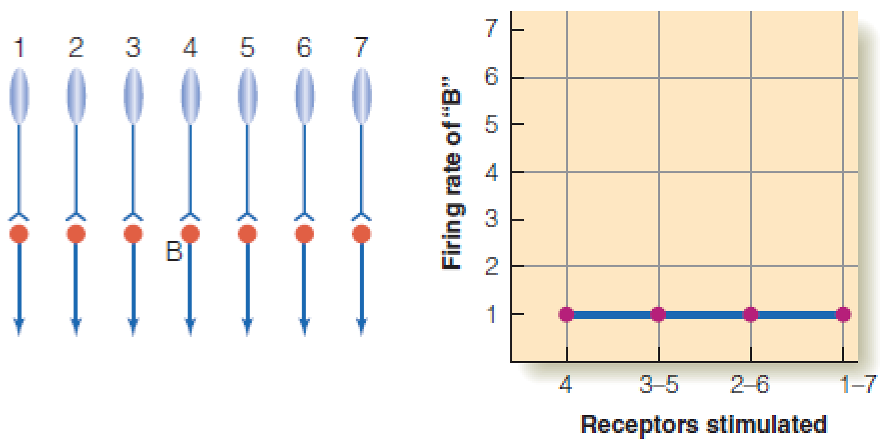
- Neural convergence
  - From receptors to ganglion cells



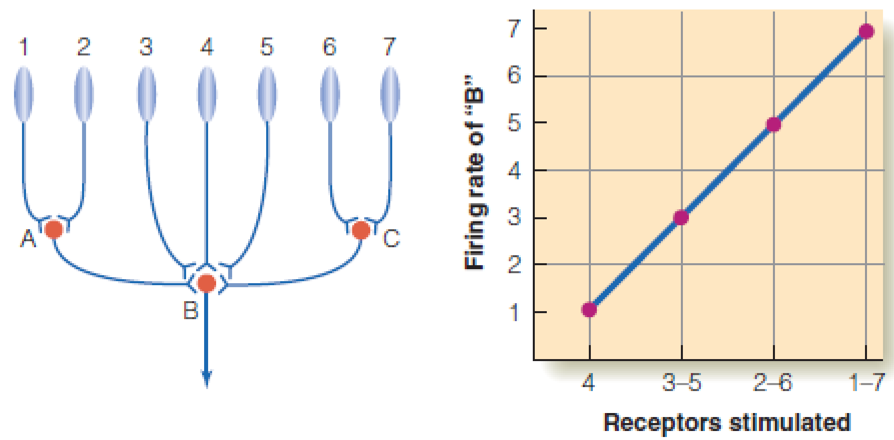


Introduction to vision - Neural convergence and perception

- Neural convergence
  - When one neuron receives signals from many other neurons



Circuit with no convergence



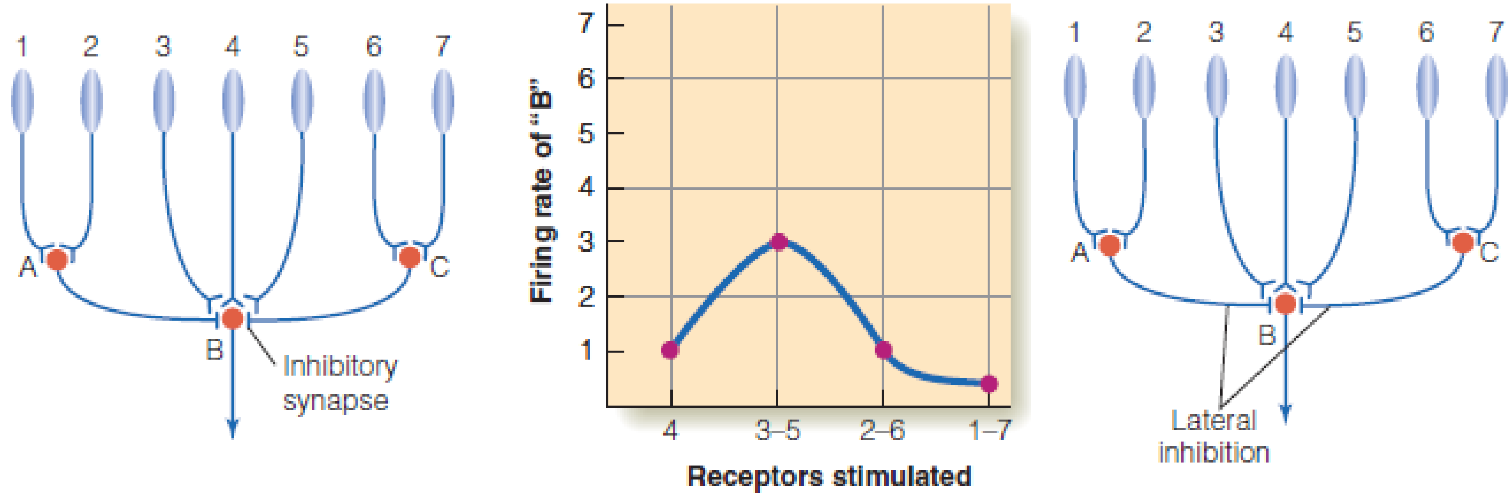
Circuit with convergence

Introduction to vision - Neural convergence and perception

- Signals from the rods converge more than do the signals from the cones
  - 120 rods to one ganglion cell vs 6 cones to a ganglion cell
  - The rods result in better sensitivity than the cones
- Signals from the cones converge less than do the signals from the rods
  - The cones result in better detail vision (visual acuity) than the rods
  - Visual acuity is highest in the fovea (rich in cones); objects that are imaged on the peripheral retina are not seen as clearly

Introduction to vision - Lateral inhibition and perception

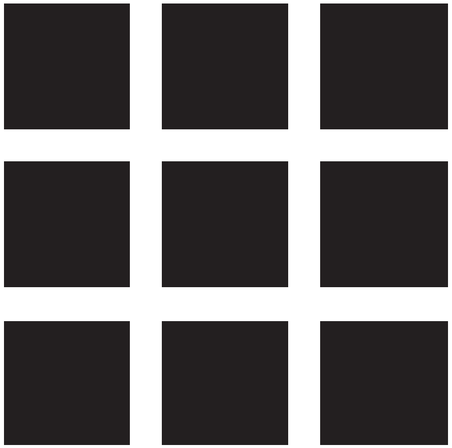
- Lateral inhibition



Lateral inhibition arrives at neuron B from A and from C;
neuron B responds best when just the center receptors (3–5) are stimulated

Introduction to vision - Lateral inhibition and perception

- Perceptual effects of lateral inhibition
  - Perception of lightness
    - Perception of shades ranging from white to grey to black
- The Hermann Grid: seeing spots at intersections



Introduction to vision - Lateral inhibition and perception

- How the dark spots at the intersections can be explained by lateral inhibition

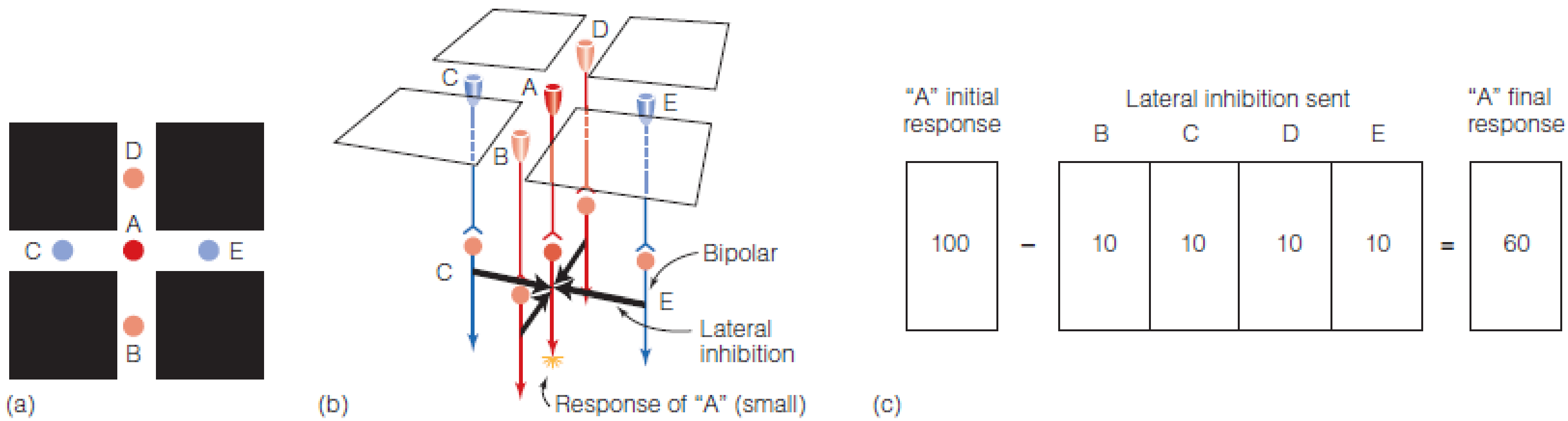


Figure 3.33 (a) Four squares of the Hermann grid, showing five of the receptors under the pattern. Receptor A is located at the intersection, and B, C, D, and E have a black square on either side. (b) Perspective view of the grid and five receptors, showing how the receptors connect to bipolar cells. Receptor A's bipolar cell receives lateral inhibition from the bipolar cells associated with receptors B, C, D, and E. (c) The calculation of the final response of receptor A's bipolar cell starts with A's initial response (100) and subtracts the inhibition associated with each of the other receptors.

Assumption: The lateral inhibition sent by each receptor's bipolar cell is one-tenth of each receptor's response

The visual cortex and beyond - Following the signal from retina to cortex

- Overview of the visual system

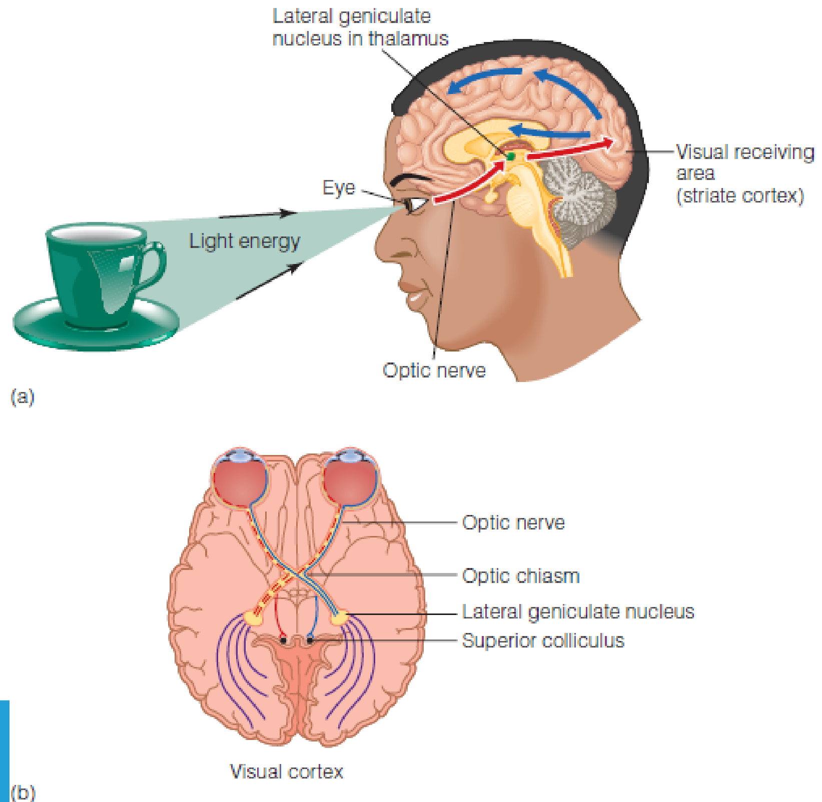


Figure 4.1 (a) Side view of the visual system, showing the three major sites along the primary visual pathway where processing takes place: the eye, the lateral geniculate nucleus, and the visual receiving area of the cortex. (b) Visual system seen from underneath the brain showing how some of the nerve fibers from the retina cross over to the opposite side of the brain at the optic chiasm.

The visual cortex and beyond - Following the signal from retina to cortex

- Processing in the lateral geniculate nucleus (LGN)

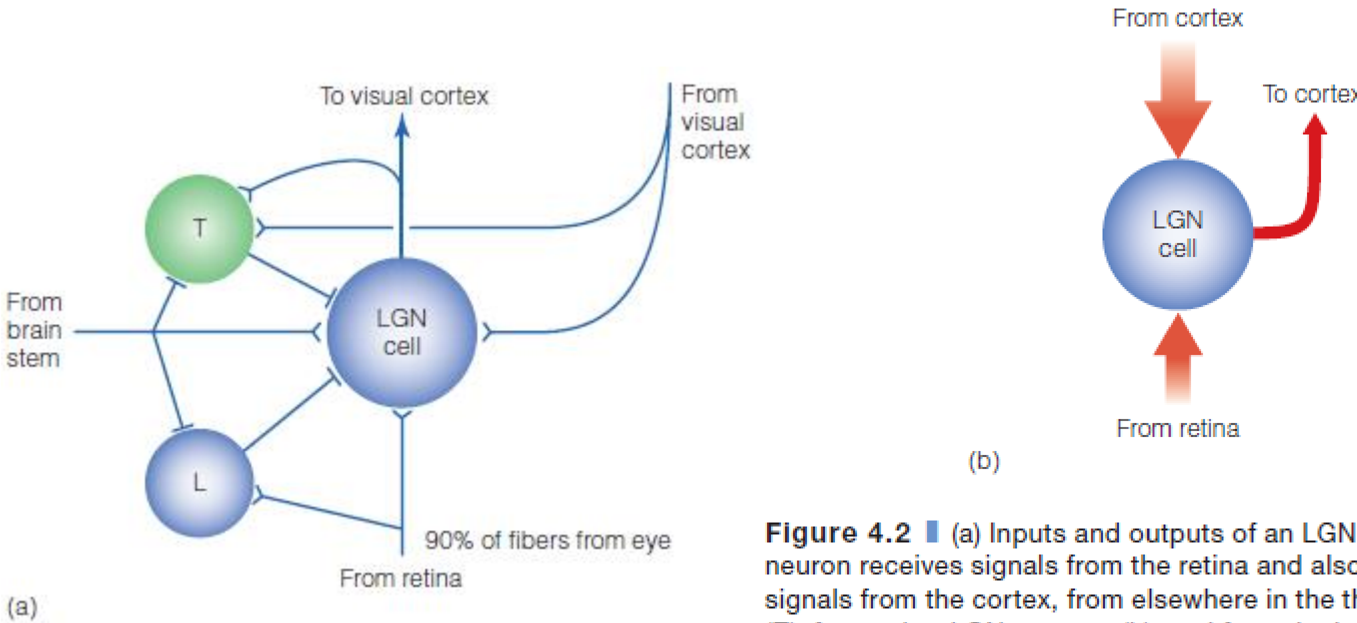


Figure 4.2 (a) Inputs and outputs of an LGN neuron. The neuron receives signals from the retina and also receives signals from the cortex, from elsewhere in the thalamus (T), from other LGN neurons (L), and from the brain stem. Excitatory synapses are indicated by Y's and inhibitory ones by T's. (b) Information flow into and out of the LGN. The sizes of the arrows indicate the sizes of the signals. *(Part a adapted from Kaplan, Mukherjee, & Shapley, 1993.)*

The visual cortex and beyond - Following the signal from retina to cortex

- Signals arriving at the LGN are sorted and organized based on
  - the eye they came from
  - the receptors that generated them
  - the type of environmental information that is represented in them
- LGN is a bilateral structure
  - one LGN in the left hemisphere and one in the right hemisphere

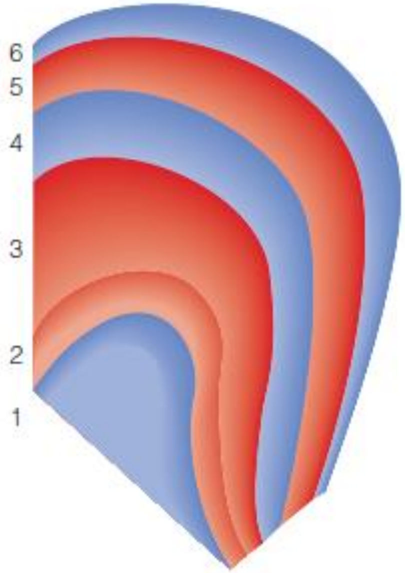


Figure 4.3 ▌ Cross section of the LGN showing layers. Red layers receive signals from the ipsilateral (same side of the body) eye. Blue layers receive signals from the contralateral (opposite side) eye.

The visual cortex and beyond - Following the signal from retina to cortex

- Organisation as a spatial map
  - Retinotopic map
  - A map in which each point on the LGN corresponds to a point on the retina

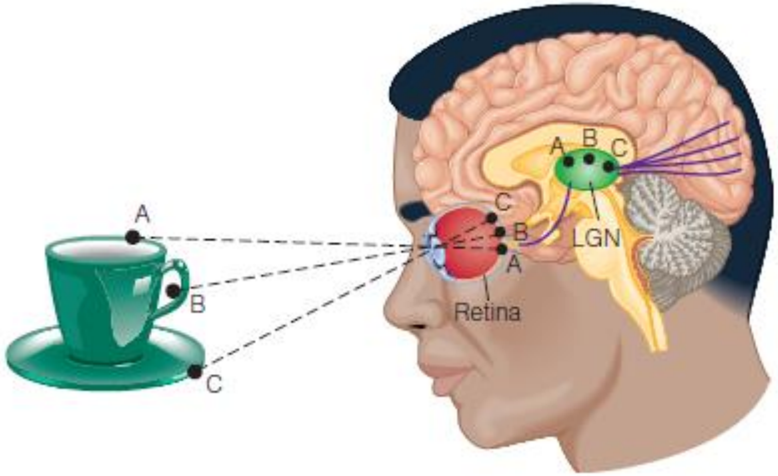


Figure 4.4 Points A, B, and C on the cup create images at A, B, and C on the retina and cause activation at points A, B, and C on the lateral geniculate nucleus (LGN). The correspondence between points on the LGN and retina indicates that there is a retinotopic map on the LGN.

The visual cortex and beyond - Following the signal from retina to cortex

- Receptive Fields of Neurons in the Striate Cortex
  - *Simple*, *complex* and *end-stopped cells*
    - Fire in response to specific features of the stimulus
    - E.g., orientation or direction of movement
    - *Feature detectors*

TABLE 4.1 Properties of Neurons in the Optic Nerve, LGN, and Cortex

TYPE OF CELL	CHARACTERISTICS OF RECEPTIVE FIELD
Optic nerve fiber (ganglion cell)	Center-surround receptive field. Responds best to small spots, but will also respond to other stimuli.
Lateral geniculate	Center-surround receptive fields very similar to the receptive field of a ganglion cell.
Simple cortical	Excitatory and inhibitory areas arranged side by side. Responds best to bars of a particular orientation.
Complex cortical	Responds best to movement of a correctly oriented bar across the receptive field. Many cells respond best to a particular direction of movement.
End-stopped cortical	Responds to corners, angles, or bars of a particular length moving in a particular direction.

The visual cortex and beyond - Do feature detectors play a role in perception?

- When we view a stimulus with a specific property, neurons tuned to that property fire
- *Selective adaptation*
  - if the neurons fire for long enough, they become fatigued, or adapt
    - the neuron’s firing rate decreases
    - the neuron fires less when that stimulus is immediately presented again

The visual cortex and beyond - Maps and columns in the striate cortex

- Retinotopic mapping indicates that information about objects near each other in the environment is processed by neurons near each other in the cortex
- Maps in the striate cortex

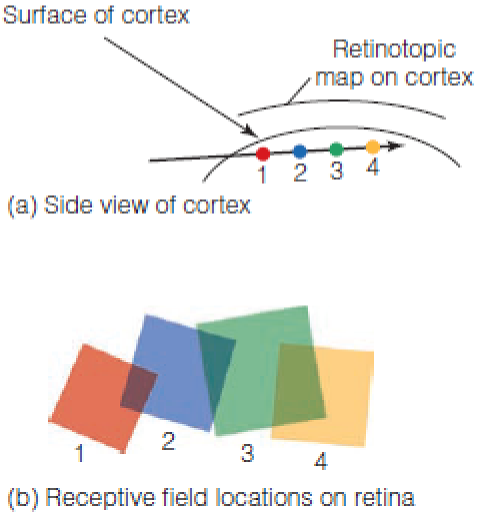
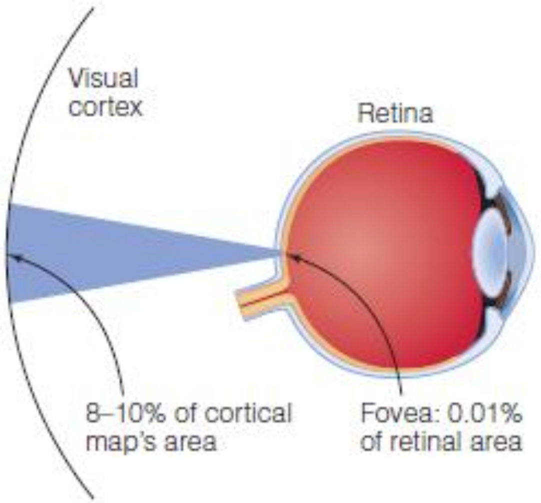


Figure 4.13 Retinotopic mapping of neurons in the cortex. When the electrode penetrates the cortex obliquely, the receptive fields of neurons recorded from the numbered positions along the track are displaced, as indicated by the numbered receptive fields; neurons near each other in the cortex have receptive fields near each other on the retina.

The visual cortex and beyond - Maps and columns in the striate cortex

- *Cortical magnification factor*
- The area representing the cone-rich fovea is much larger than one would expect from the fovea's small size
- Even though the fovea accounts for only 0.01 percent of the retina's area, signals from the fovea account for 8 to 10 percent of the retinotopic map on the cortex



The visual cortex and beyond - Maps and columns in the striate cortex

- The cortex is organized into a number of different kinds of columns

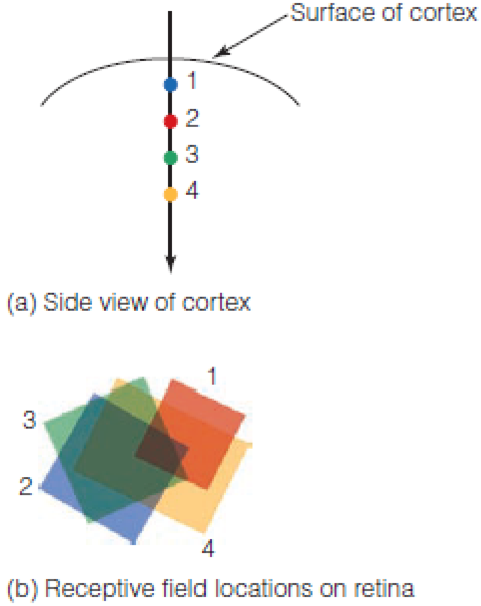


Figure 4.19 When an electrode penetrates the cortex perpendicularly, the receptive fields of the neurons encountered along this track overlap. The receptive field recorded at each numbered position along the electrode track is indicated by a correspondingly numbered square. *(This figure was published in* Neuron, *56, Wandell, B. A., Dumoulin, S. O., & Brewer, A. A., Visual field maps in human cortex, 366–383. Copyright Elsevier, 2007.)*

The visual cortex and beyond - Maps and columns in the striate cortex

- How is an object represented in the striate cortex?

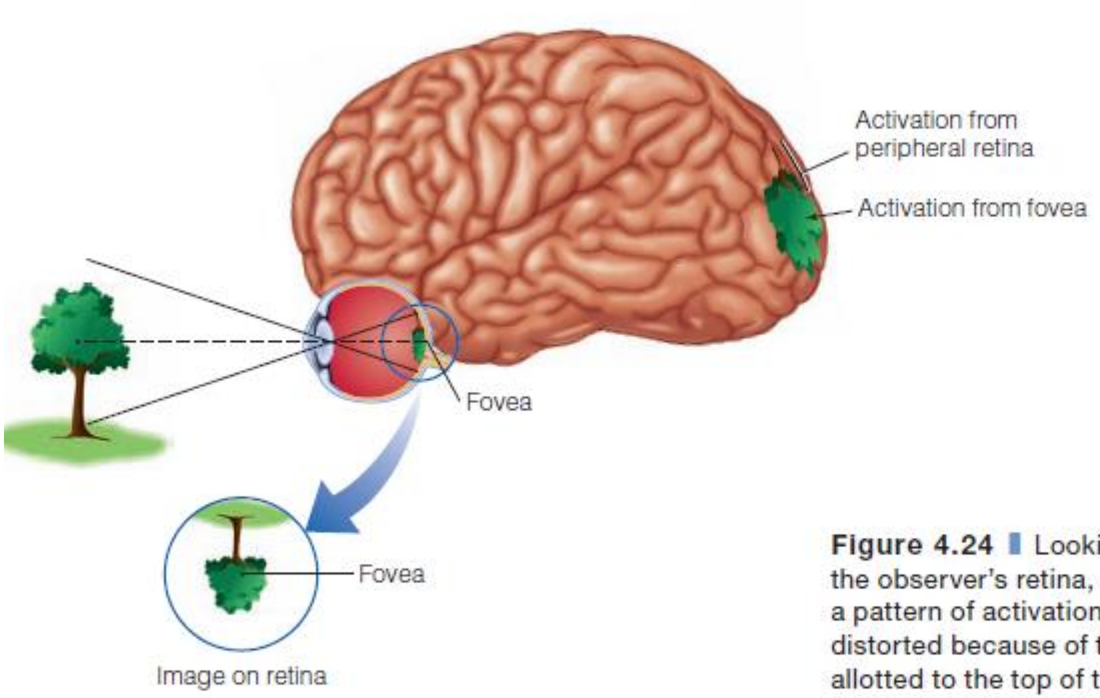


Figure 4.24 Looking at the tree creates an image on the observer's retina, and this image on the retina causes a pattern of activation on the visual cortex. This pattern is distorted because of the magnificaton factor (more space is allotted to the top of the tree, where the observer is looking.)

The visual cortex and beyond - Streams: pathways for what, where and how

- Streams that transmit information from the striate cortex to other areas in the brain
  - The *what* pathway (the ventral pathway)
    - From the striate cortex to the temporal lobe
    - Identifying objects
  - The *where (how)* pathway (the dorsal pathway)
    - From the striate cortex to the parietal lobe
    - Locating objects
    - Taking action

The visual cortex and beyond - Modularity: structures for faces, places and bodies

- Fusiform face area (FFA)
  - Activated by faces
  - Located in the Fusiform Gyrus

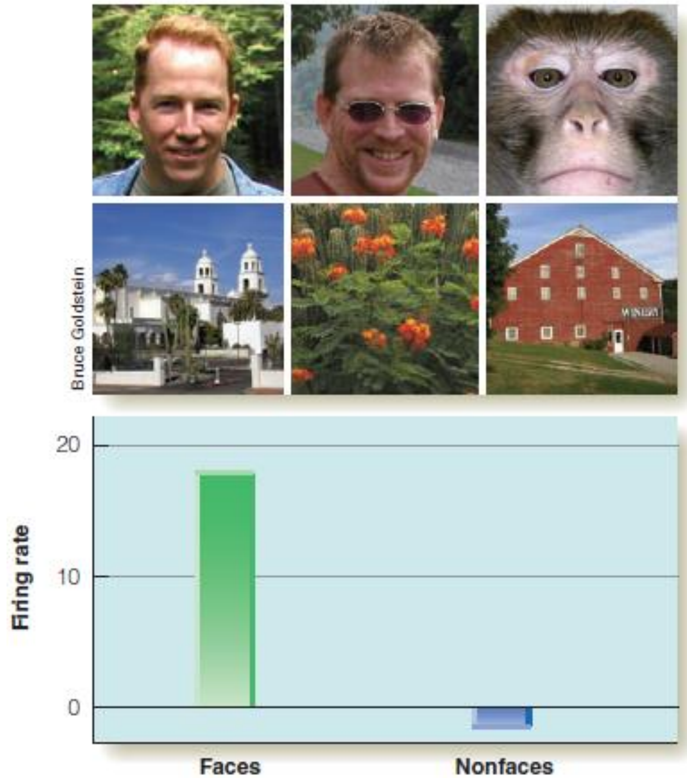


Figure 4.33 Size of response of a neuron in the monkey's IT cortex that responds to face stimuli but not to nonface stimuli. *(Based on data from Rolls & Tovee, 1995.)*

The visual cortex and beyond - Modularity: structures for faces, places and bodies

- Parahippocampal place area (PPA)
  - Sensitivity to indoor and outdoor scenes
- Extrastriate body area (EBA)
  - Activated by pictures of bodies and parts of bodies (but not by faces)
- Located in the temporal cortex

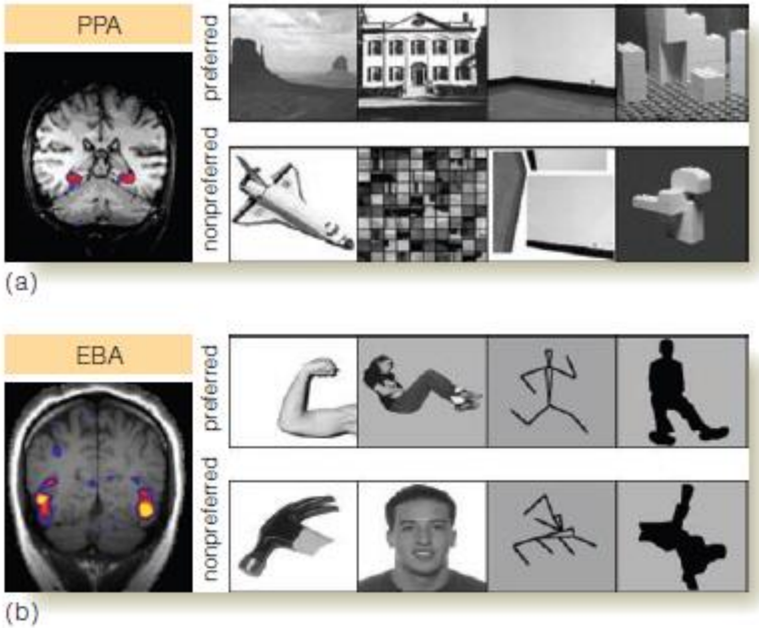
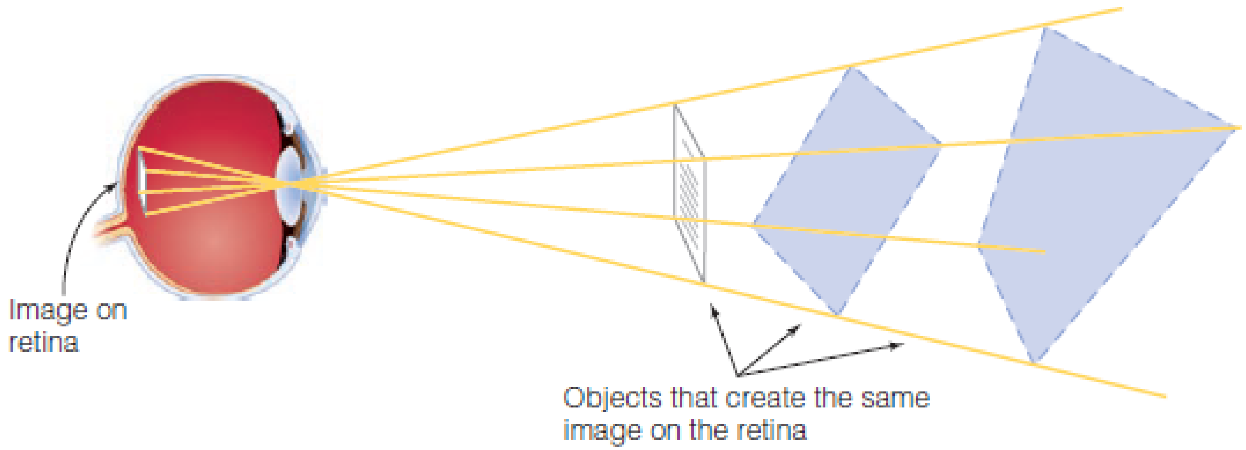


Figure 4.34 (a) The parahippocampal place area is activated by places (top row) but not by other stimuli (bottom row). (b) The extrastriate body area is activated by bodies (top), but not by other stimuli (bottom). *(From Kanwisher, N., The ventral visual object pathway in humans: Evidence from fMRI. In* The Visual Neurosciences, *2003, pp. 1179–1189. Edited by Chalupa, L., & Werner, J., MIT Press.)*

Perceiving objects and scenes - Why is it so difficult to design a perceiving machine?

- The stimulus on the receptors is ambiguous



Perceiving objects and scenes - Why is it so difficult to design a perceiving machine?

- Objects can be hidden or blurred

Perceiving objects and scenes - Why is it so difficult to design a perceiving machine?

- Objects look different from different viewpoints

Perceiving objects and scenes - The Gesthalt approach to object perception

- Rejecting the idea that perception is built up of sensations
- Perceptual organisation
  - Grouping elements in an image to create larger objects
  - Six laws to explain how *perceptual grouping* occurs

Perceiving objects and scenes - The Gesthalt approach to object perception

- Law of pragnanz
  - *Every stimulus pattern is seen in such a way that the resulting structure is as simple as possible*

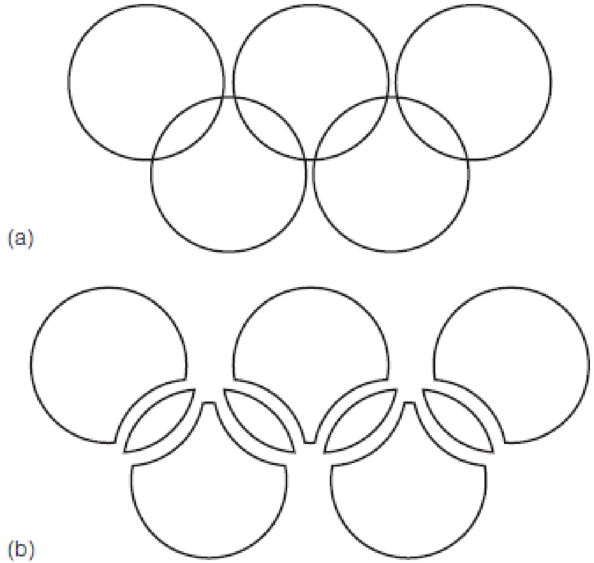


Figure 5.13 (a) This is usually perceived as five circles, not as the nine shapes in (b).

Perceiving objects and scenes - The Gesthalt approach to object perception

- Law of similarity
  - *Similar things appear to be grouped together*

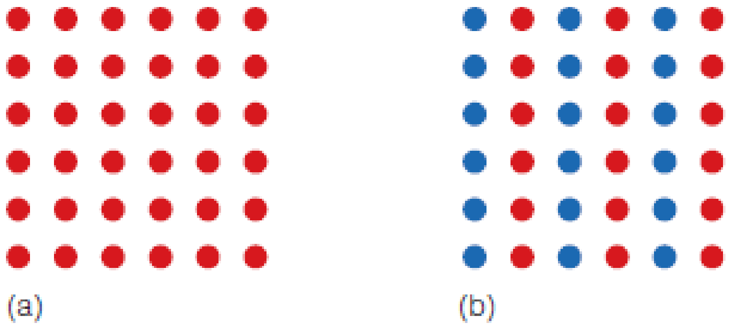


Figure 5.14 (a) Perceived as horizontal rows or vertical columns or both. (b) Perceived as vertical columns.

Perceiving objects and scenes - The Gesthalt approach to object perception

- Law of good continuation
  - *Points that, when connected, result in straight or smoothly curving lines are seen as belonging together, and the lines tend to be seen in such a way as to follow the smoothest path*

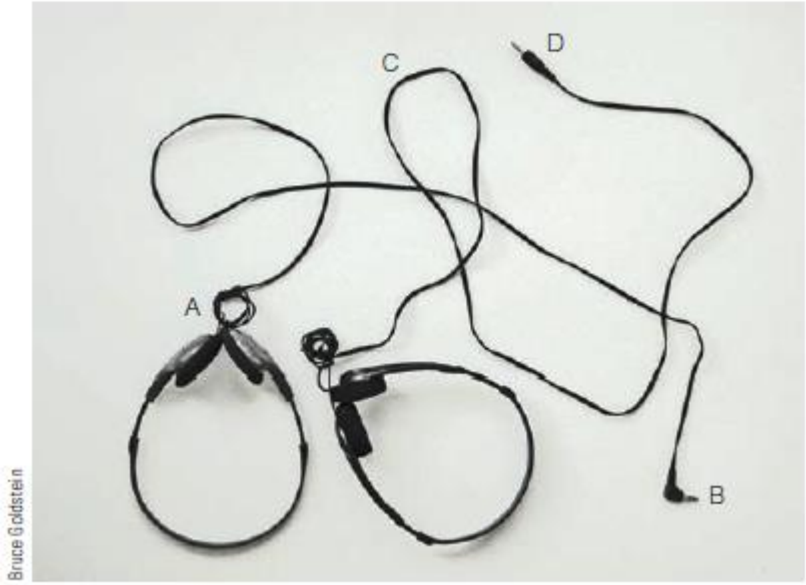
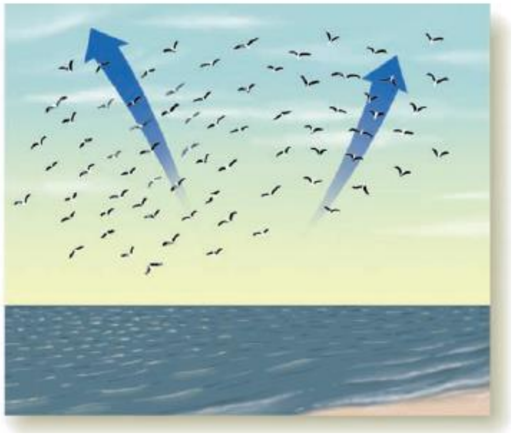


Figure 5.16 Good continuation helps us perceive two separate wires, even though they overlap.

Perceiving objects and scenes - The Gesthalt approach to object perception

- Law of proximity
  - *Things that are near each other appear to be grouped together*

- Law of common fate
  - *Things that are moving in the same direction appear to be grouped together*



Perceiving objects and scenes - The Gesthalt approach to object perception

- Law of familiarity
  - *Things that form patterns that are familiar or meaningful are likely to become grouped together*

Perceiving objects and scenes - The Gesthalt approach to object perception

- Perceptual segregation: how objects are separated from the background

Figure 5.21 ▌ A version of Rubin's reversible face–vase figure.

Perceiving objects and scenes - The Gesthalt approach to object perception

- What are the properties of figure and ground?
- The figure is more “thinglike” and more memorable than the ground
- The figure is seen as being in front of the ground
- The ground is seen as unformed material and seems to extend behind the figure
- The contour separating the figure from the ground appears to belong to the figure

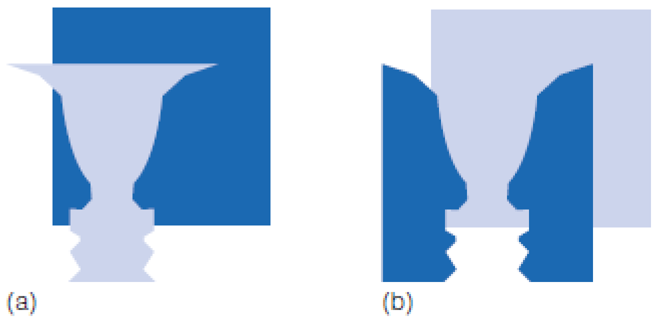
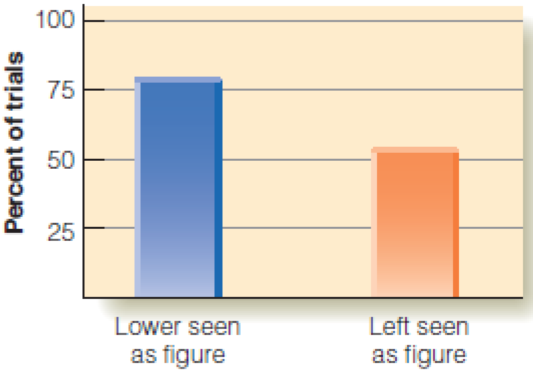


Figure 5.22 (a) When the vase is perceived as figure, it is seen in front of a homogeneous dark background. (b) When the faces are seen as figure, they are seen in front of a homogeneous light background.

Perceiving objects and scenes - The Gesthalt approach to object perception

- What factors determine which area is figure?
  - Regions in the lower part of a display are more likely to be perceived as figure than regions in the upper part


(a)



(b)

Perceiving objects and scenes - Perceiving scenes and objects in scenes

- A scene is a view of a real-world environment that contains (1) background elements and (2) multiple objects that are organized in a meaningful way relative to each other and the background
- Perceiving the *gist of a scene* is possible within a fraction of a second

Perceiving objects and scenes - Perceiving scenes and objects in scenes

- What enables observers to perceive the gist of a scene so rapidly?
  - Global image features (Oliva and Torralba, 2006)
    - Degree of naturalness; degree of openness; degree of roughness; degree of expansion; colour

Perceiving objects and scenes - Perceiving scenes and objects in scenes

- We easily use our knowledge of regularities in the environment to help us perceive, even though we may not be able to identify the specific information we are using
  - *Physical regularities*
    - Regularly occurring physical properties of the environment people can perceive horizontals and verticals more easily than other orientations
  - *Semantic regularities*
    - Meaning of a scene

Perceiving objects and scenes - The physiology of object and scene perception

- Neurons that respond to perceptual grouping

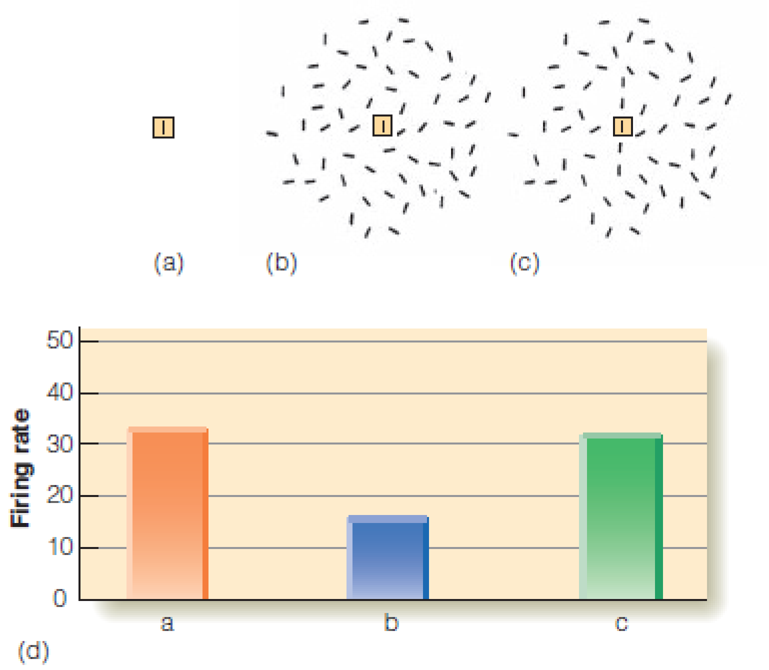


Figure 5.43 How a neuron in the striate cortex (V1) responds to (a) an oriented bar inside the neuron's receptive field (the small square); (b) the same bar surrounded by randomly oriented bars; (c) the bar when it becomes part of a group of vertical bars, due to the principles of similarity and good continuation. *(Adapted from Zapadia, M. K., Ito, M., Gilbert, C. G., & Westheimer, G. (1995). Improvement in visual*

Perceiving objects and scenes - The physiology of object and scene perception

- How does the brain respond to objects?
- Distributed activity across the brain
- Fusiform face area (FFA)
  - A face might cause a large amount of activity in the FFA, but also cause activity in other areas as well
  - Firing is, therefore, distributed in two ways:
    - across groups of neurons within a specific area
    - across different areas in the brain

"Design" case study: There's a special part of the brain just for recognising faces

- There's a special part of the brain just for recognising faces
- People are born for a preference for faces
- Fusiform face area (FFA) (Kanwisher, 1997; 2000)
  - Outside visual cortex
  - Recognise faces
  - Helps us identify faces more quickly than objects
  - People with autism do not use FFA to identify faces (Pierce, 2001)

Kanwisher, N. (2000). Domain specificity in face perception. *Nature, 3,* 759–763.

We look where the face looks

- Eye tracking research shows that if a picture of a face looks away from us and toward a product on a web page, we tend to also look at the product
- Establish emotional connection vs directing attention

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

People decide who and what is alive by looking at the eyes

- Looser and Wheatley (2010) took pictures of people and then morphed them in stages into inanimate mannequin faces

- Subjects say the pictures no longer show someone who is alive at about the 75 percent mark
- People primarily use the eyes to decide if a picture shows someone who is human and alive
- Video 1; video 2

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

Take away messages

- People recognize and react to faces on Web pages faster than anything else on the page (at least by those who are not autistic)

- Faces looking right at people will have the greatest emotional impact on a Web page, probably because the eyes are the most important part of the face

- If a face on a Web page looks at another spot or product on the page, people will also tend to look at that product. This doesn't necessarily mean that they paid attention to it, just that they physically looked at it.

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

References

- Goldstein, E. (2009). Sensation and Perception.
- Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

Contact

Dr. Ginevra Castellano

Department of Information Technology

Uppsala University

Email: **ginevra.castellano@it.uu.se**