Computer vision
Light and human visual system

Le Thanh Ha, Ph.D
Assoc. Prof. at University of Engineering and Technology,
Vietnam National University
ltha@vnu.edu.vn; lthavnu@gmail.com; 0983 692 592

Lecture content

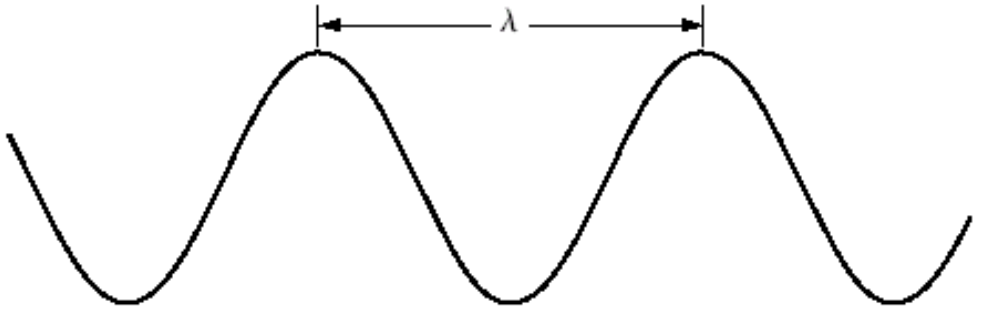
- Light
- Human eye
- Perceiving objects and scenes
- Illutions

HMI
HUMAN MACHINE INTERACTION LABORATORY

LIGHT

Light

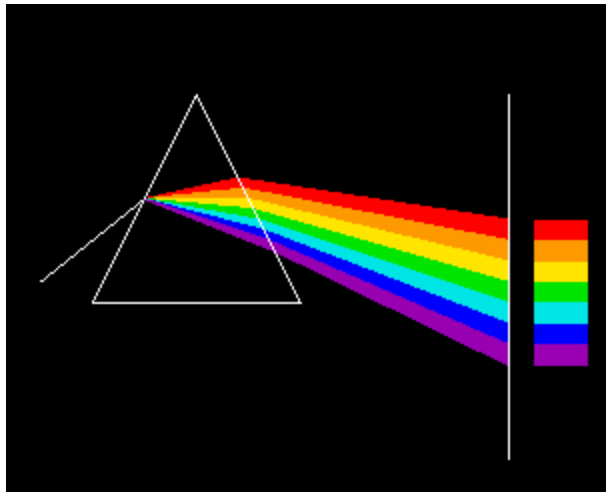
FIGURE 2.11
Graphical representation of one wavelength.



$$\lambda = \frac{c}{f}$$

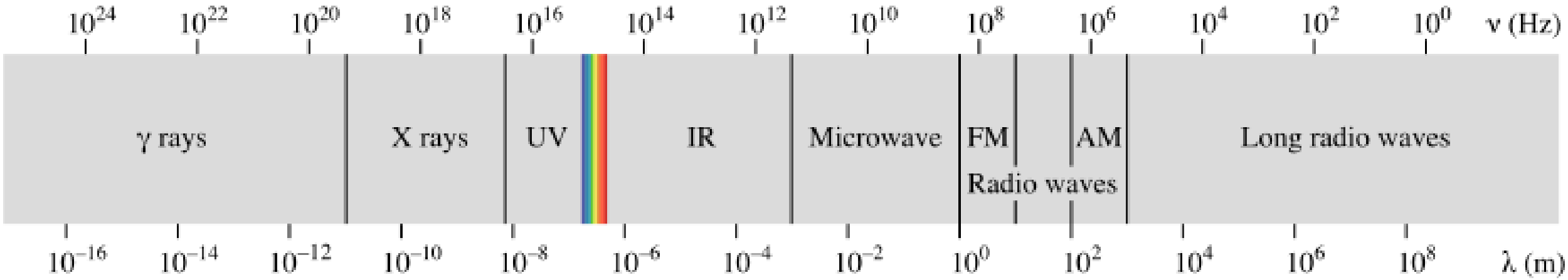
- λ: wavelength
- c: speed of light (2.998×10^8 m/s)
- f: frequency

Isaac Newton prism



- Sunlight consists of continuous spectrum of colors ranging from violet to red

Light and Electromagnetic Spectrum



Light properties

- Light is a particular type of EM radiation:
  - Can be presented by frequency or wavelength
  - Natural white light (include all frequencies) comes from the sun (S. I. Newton's glass prism)
  - When an object is radiated by while light:
    - Object absorbs some ranges of light's frequency and reflects the others
    - Human eyes perceive the reflected light and assume it as the object's color.
    - Eg: green objects reflect ligth with wavelengths primarily in [500,570] nm range and absorb most of the energy at other wavelengths.

Why we still perceive scene around even thought there is no sun?

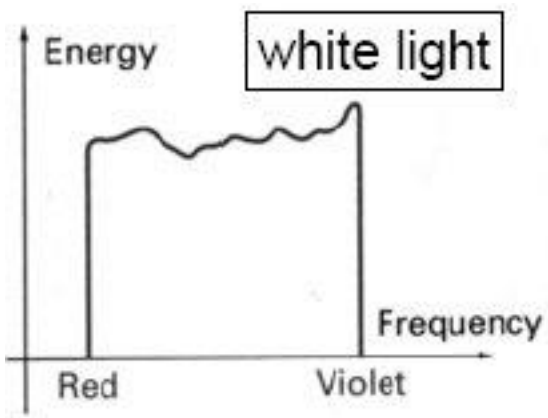
Light properties

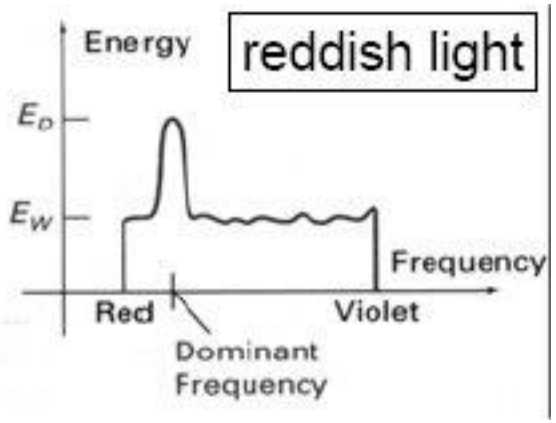
- Light that is advoid of color is called monochromatic light (ML).
  - Intensity of ML varies from black to gray, and white → gray level, or just intensity.

- Color light:
  - Luminance: amount of energy perceived.
  - Chrominance:
    - Frequencies of wavelight
    - Hue
    - Purity

Light properties

- Frequency (hue, color)
- Luminance (sum of area under the freq. curve)
- Purity: $(E_D-E_W)/ E_D$





HMI
HUMAN MACHINE INTERACTION LABORATORY

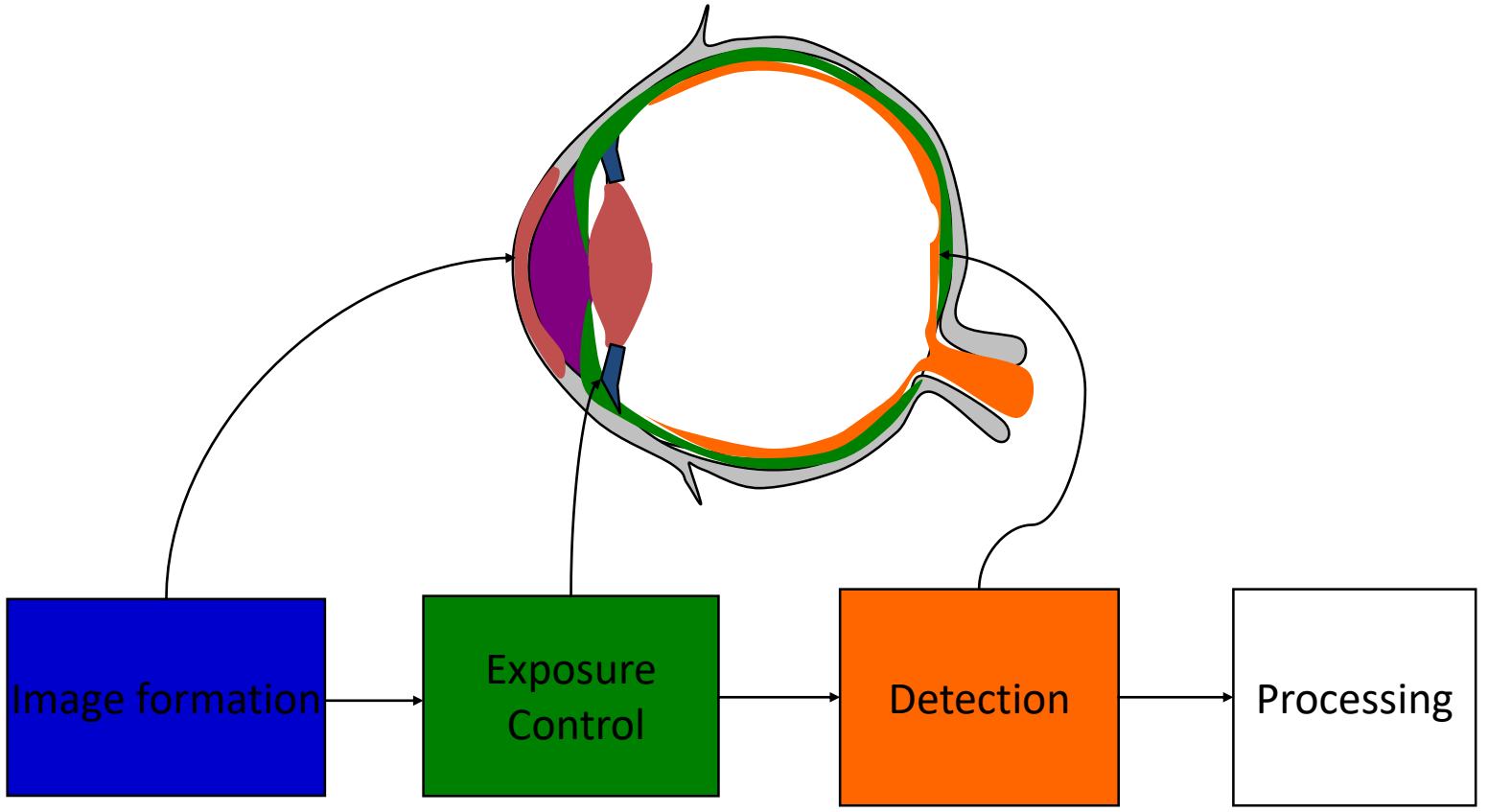
Human Visual Perception

- Human perception encompasses both
  - physiological aspects
  - psychological aspects.

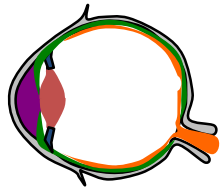
HMI
HUMAN MACHINE INTERACTION LABORATORY

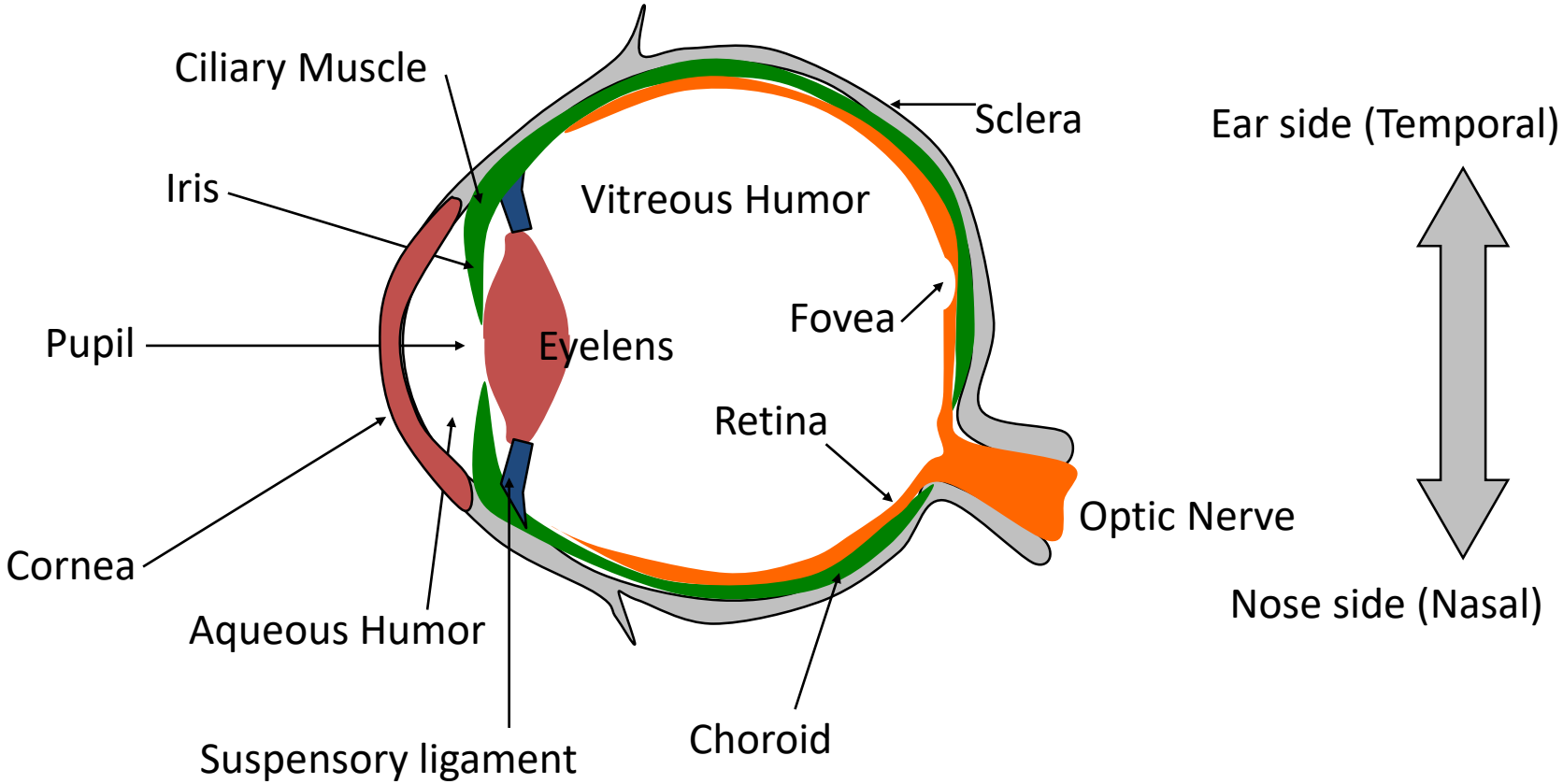
HUMAN EYE

Human Visual System



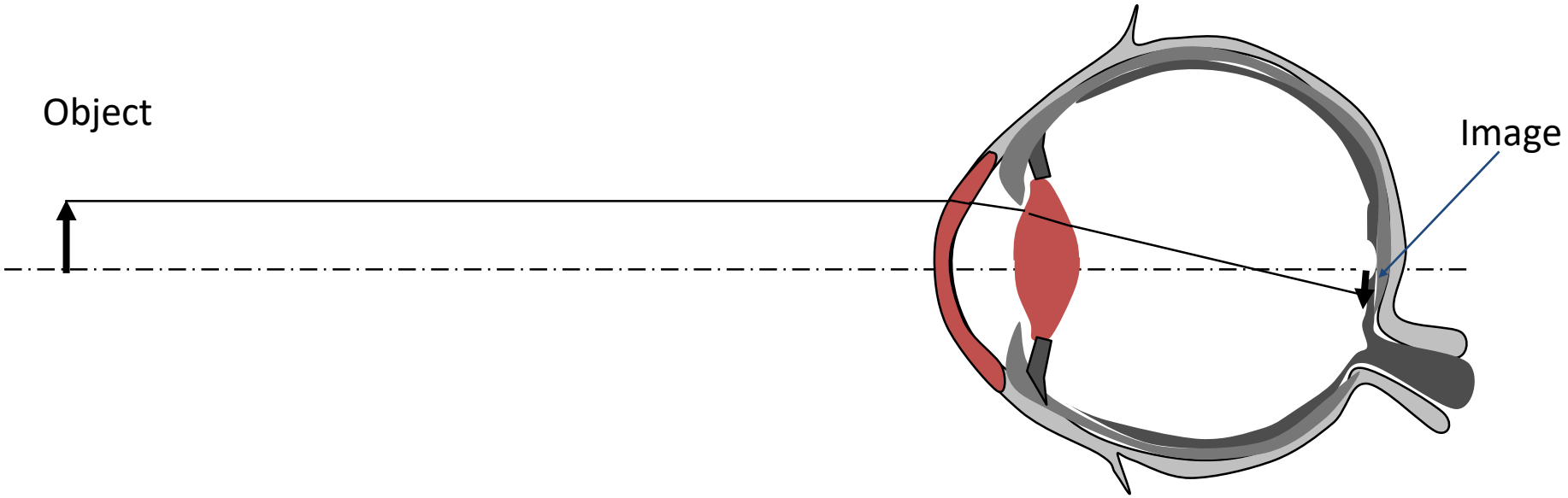
Human Eye





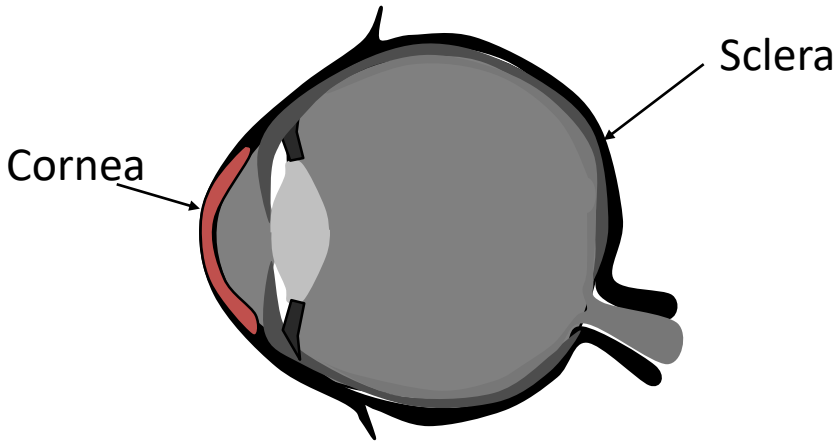
- Human eye is a complete imaging system.

Image Formation



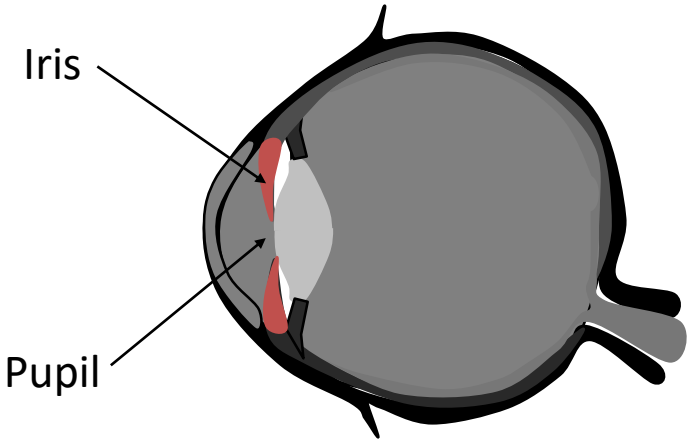
- The curved surfaces of the eye focus the image onto the back surface of the eye.

Cornea

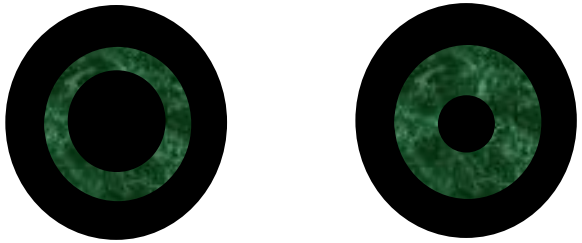


- The outer wall of the eye is formed by the hard, white *sclera*.
- *Cornea* is the clear portion of the sclera.
- 2/3 of the refraction takes place at the cornea.

Iris and Pupil



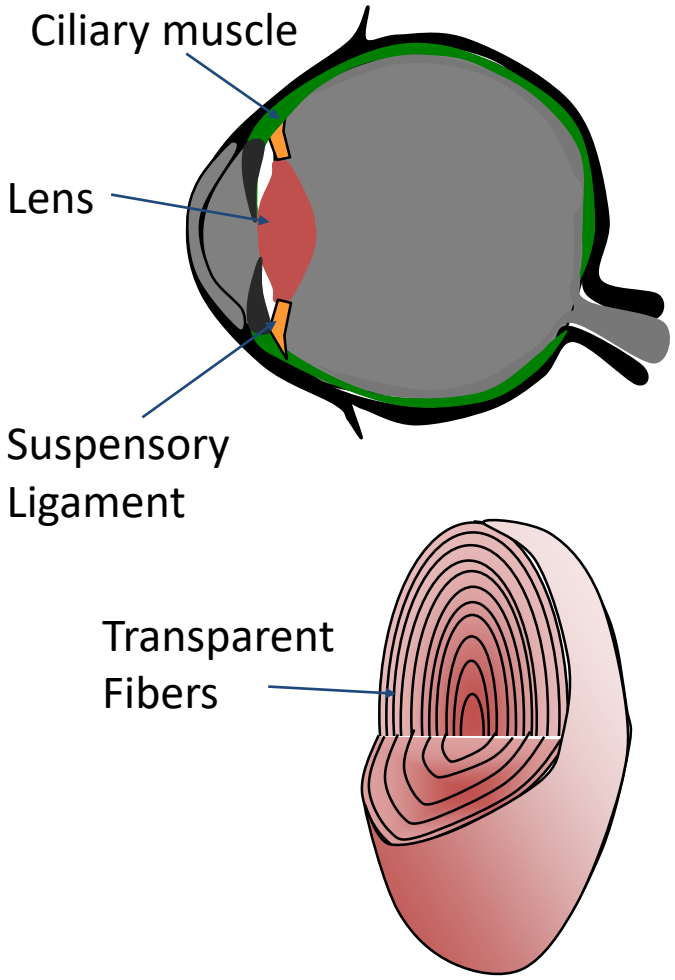
- Colored *iris* controls the size of the opening (*pupil*) where the light enters.
- Pupil determines the amount of light, like the aperture of a camera.



Iris open
Dilated pupil

Iris closed
Constricted pupil

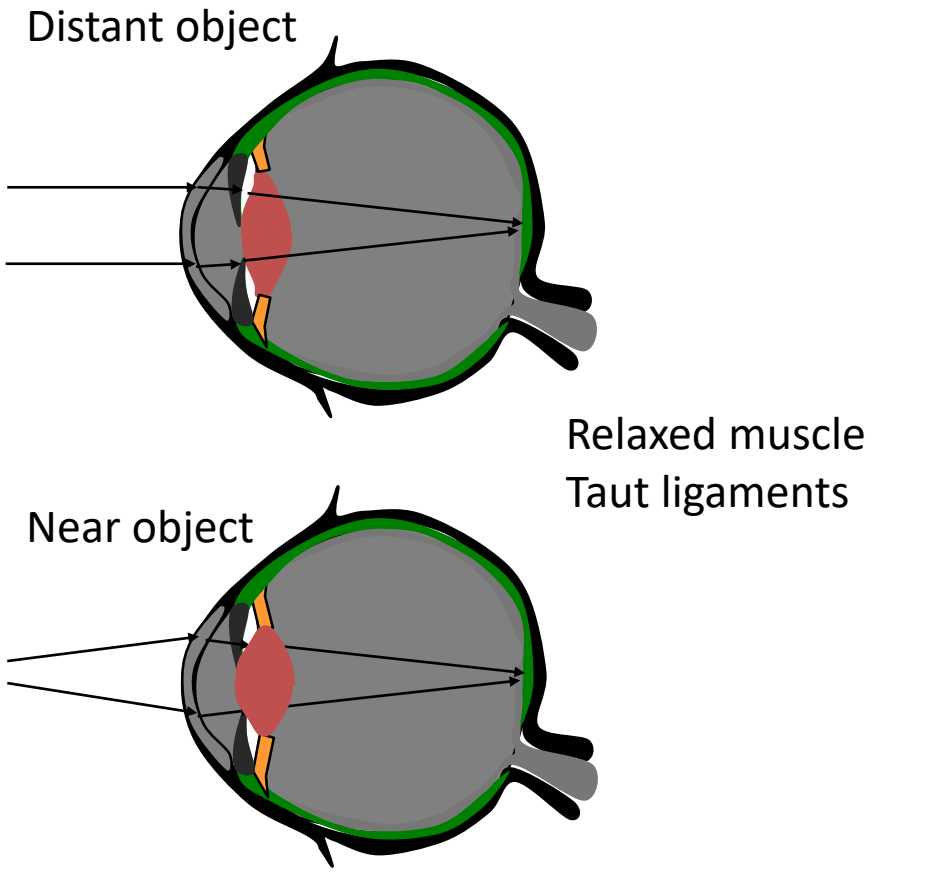
Lens



Cross section of the eye lens

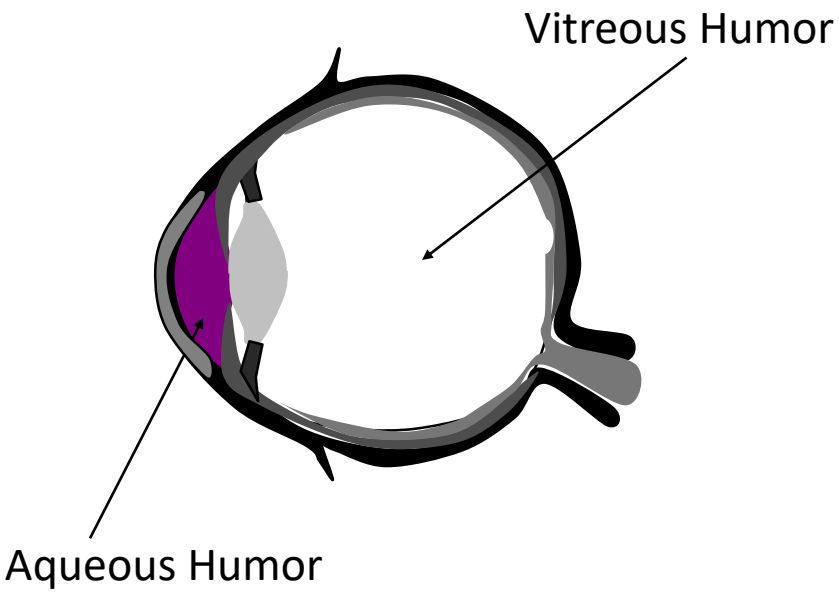
- *Eye lens* is made of transparent fibers in a clear membrane.
- Suspended by suspensory ligament.
- Used as a fine focusing mechanism by the eye; provides 1/3 of eye's total refracting power.
- Non-uniform index of refraction.

Accommodation



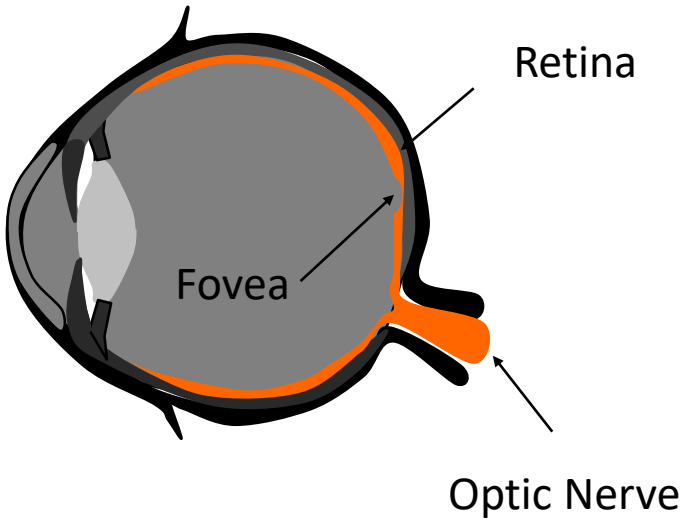
- The *suspensory ligaments* attach the lens to the *ciliary muscle*.
- When the muscle contracts, the lens bulges out in the back, decreasing its focal length.
- The process by which the lens changes shape to focus is called *accommodation*.

Aqueous Humor and Vitreous Humor



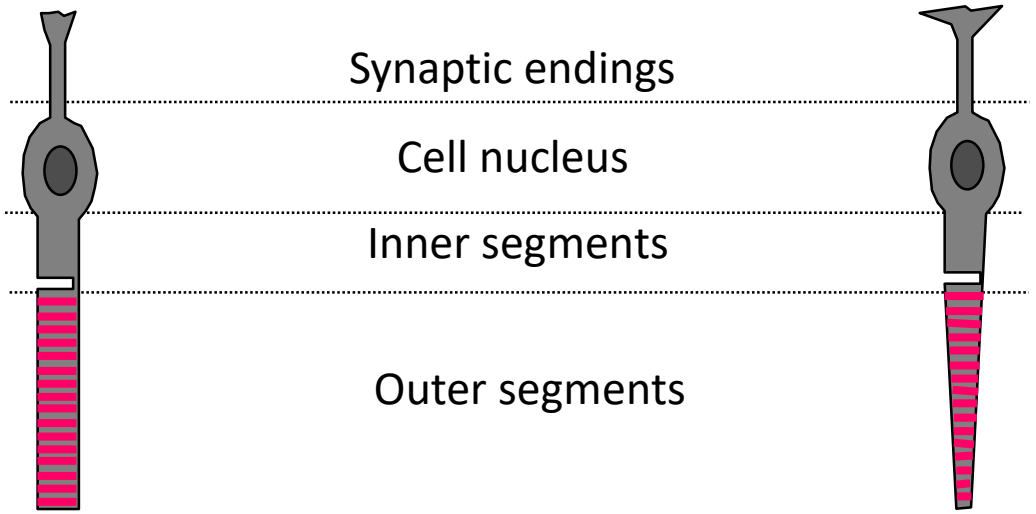
- Transparent gelatinous liquid filling the eye.
- Provides nutrients to the cornea and eye lens.
- Also helps maintain the eyeball shape with its pressure.

Retina



- *Retina* is the photosensitive “detector” for the eye.
- Two types of receptors in the retina: *rods* for low light level, and *cones* for color.
- Located at the center of the retina, *fovea* contains a greater concentration of cones.
- Signals from the receptors leave through the *optic nerve* to the brain.

Rods and Cones



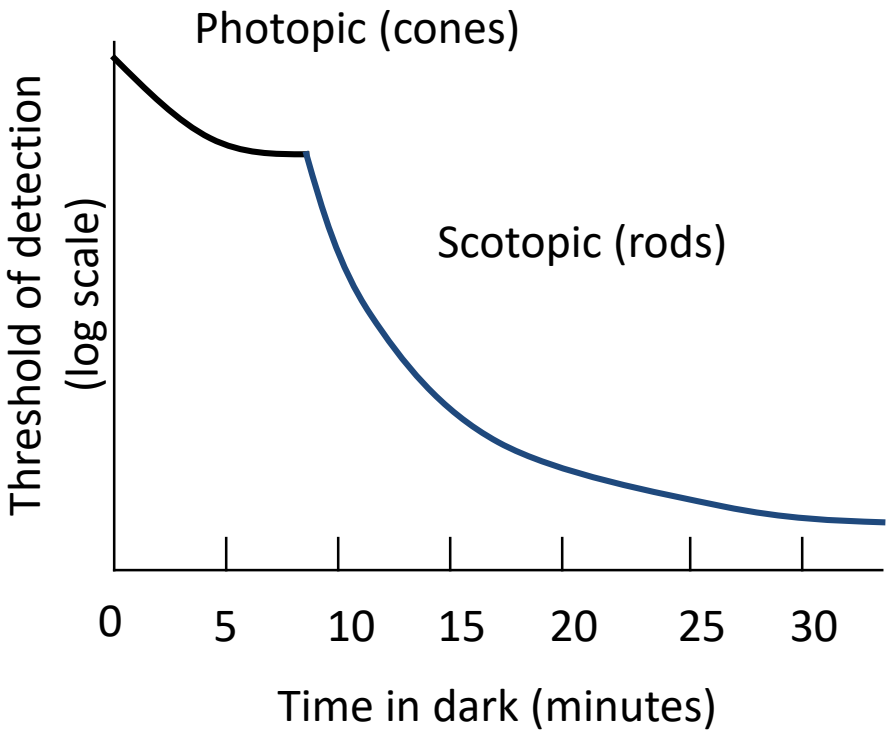
Rod

- Highly sensitive to low light level or *scotopic* conditions.
- Black and white.
- Dispersed in the periphery of the retina.

Cone

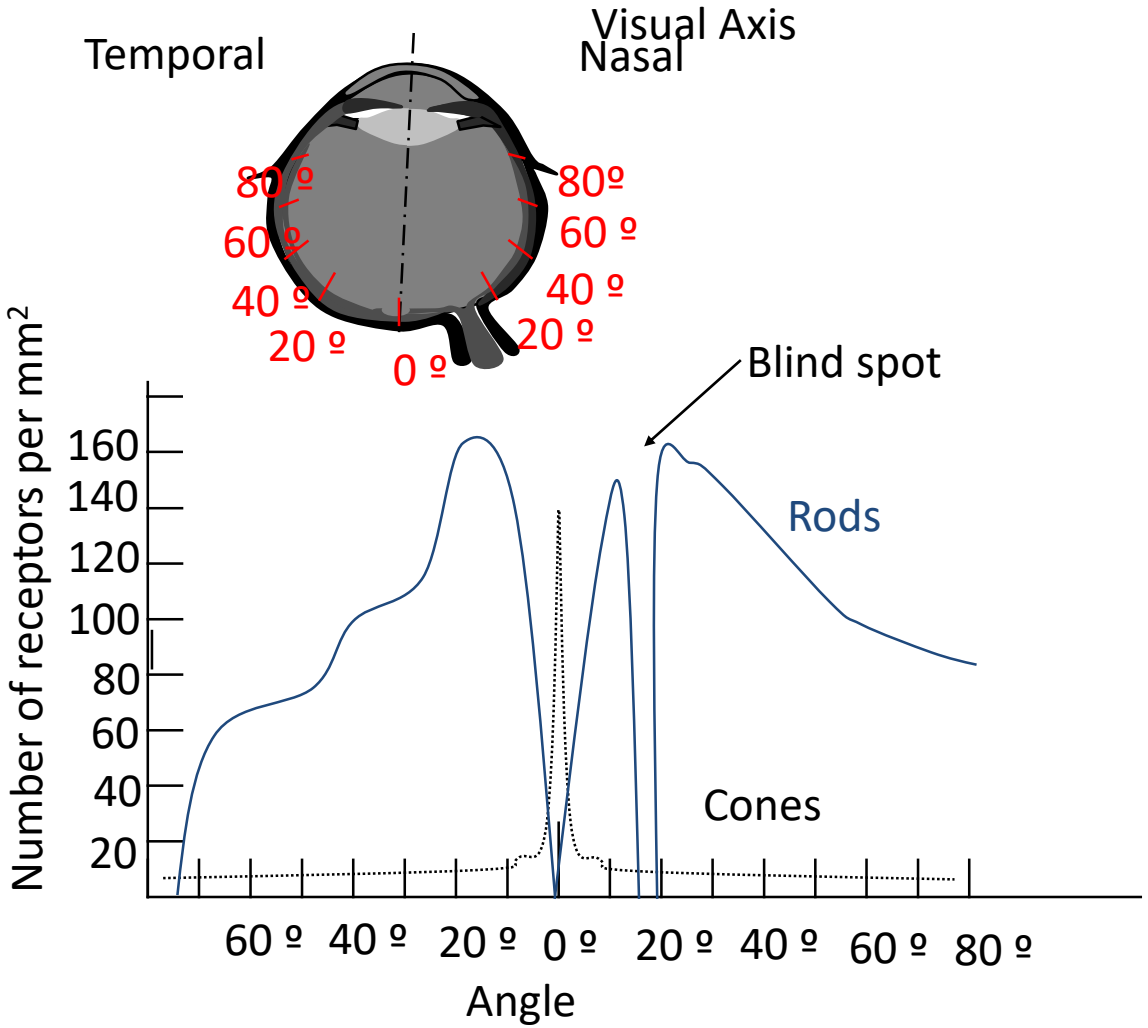
- Sensitive to high light level or *photopic* conditions.
- Three types of cones responsible for color vision.
- Concentrated in the fovea.

Adaptation



- Why can’t you see immediately after you enter a movie theater from daylight?
- The threshold of detection changes with overall light level.
- The switch is quite gradual, until the sensitivities of cones and rods cross over at about 7 minutes in the dark.

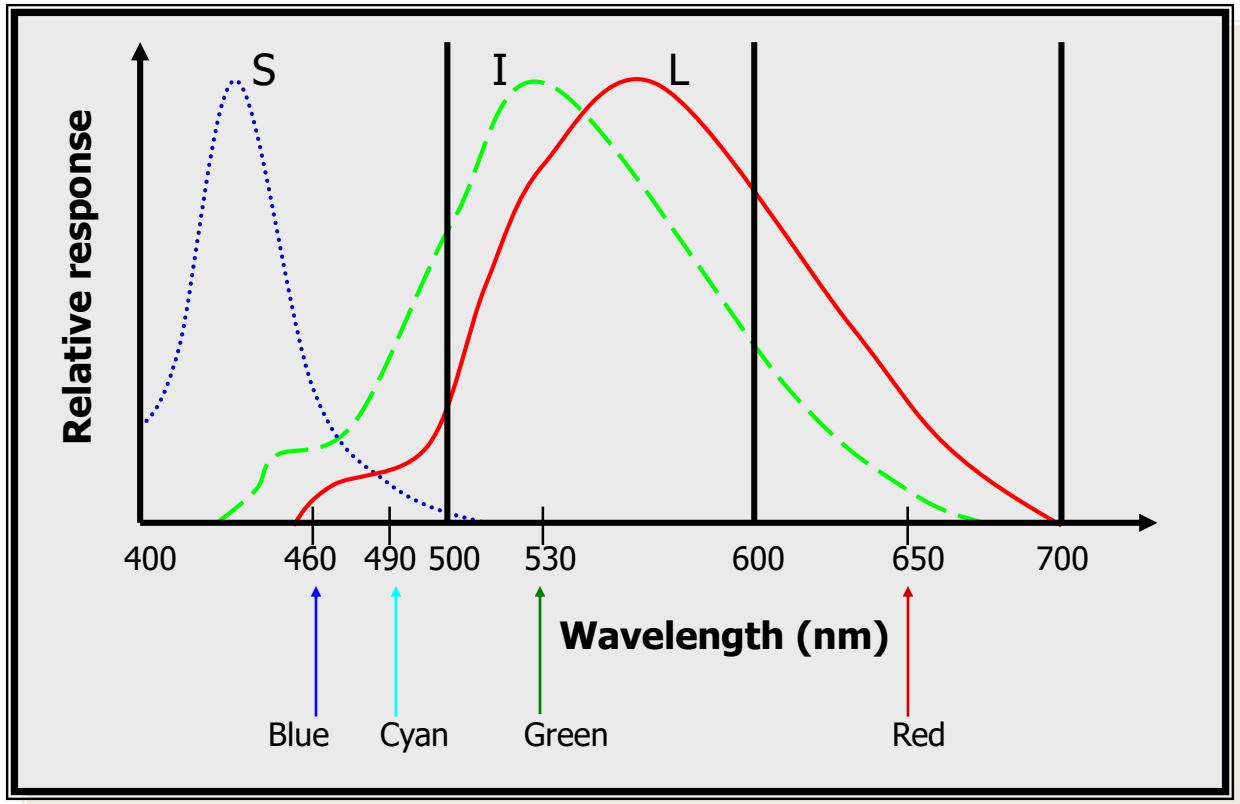
Distribution of Photoreceptors



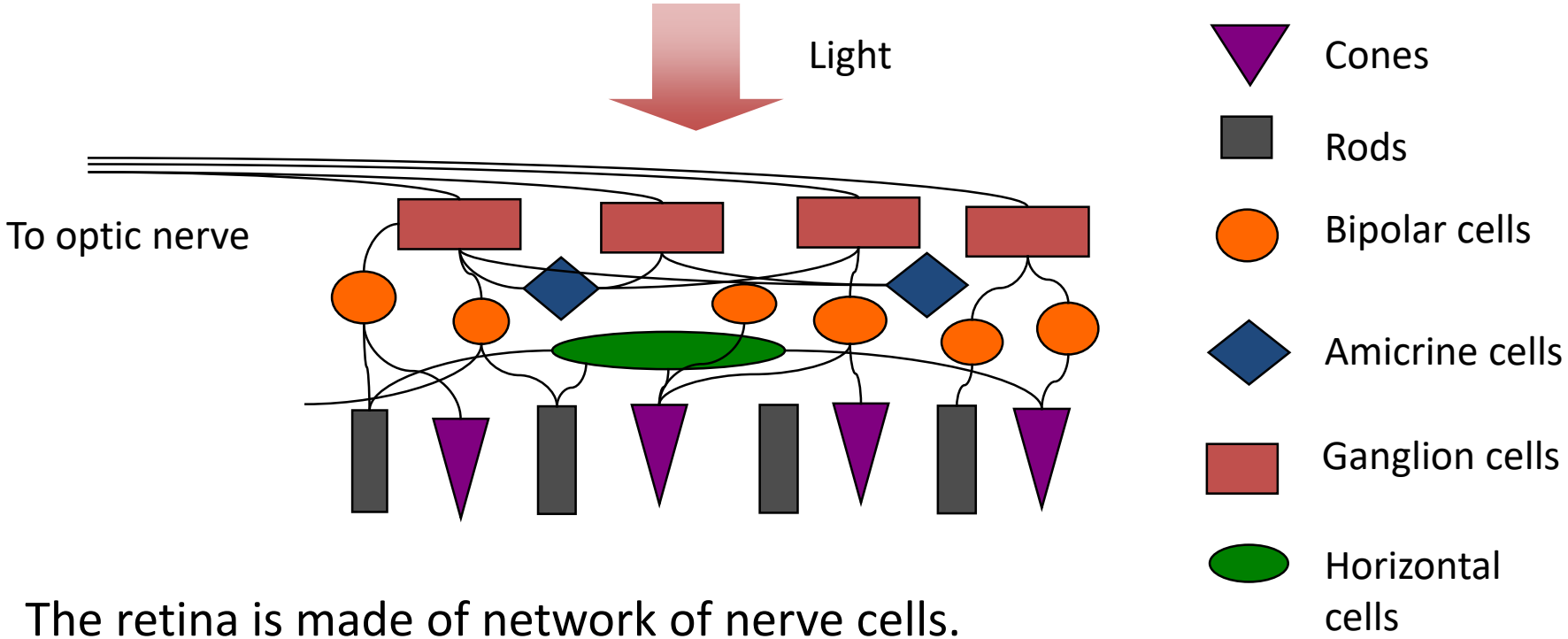
- Cones are concentrated in the fovea.
- Rods predominate the periphery.
- There is a blind spot where there are no photoreceptors, at the point where the nerves exit the eye (optic nerve).

Human Vision

- Human Cone Response to Color
  - three cone types (S,I,L) correspond to B,G,R

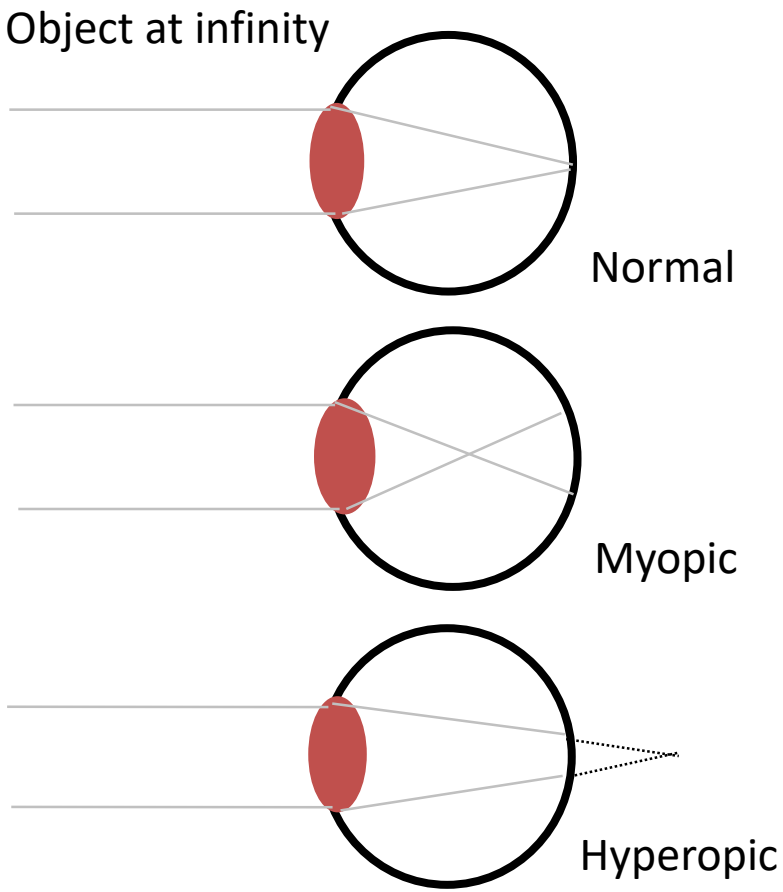


Retina



- The retina is made of network of nerve cells.
- The network works together to reduce the amount of information in a process called *lateral inhibition*.

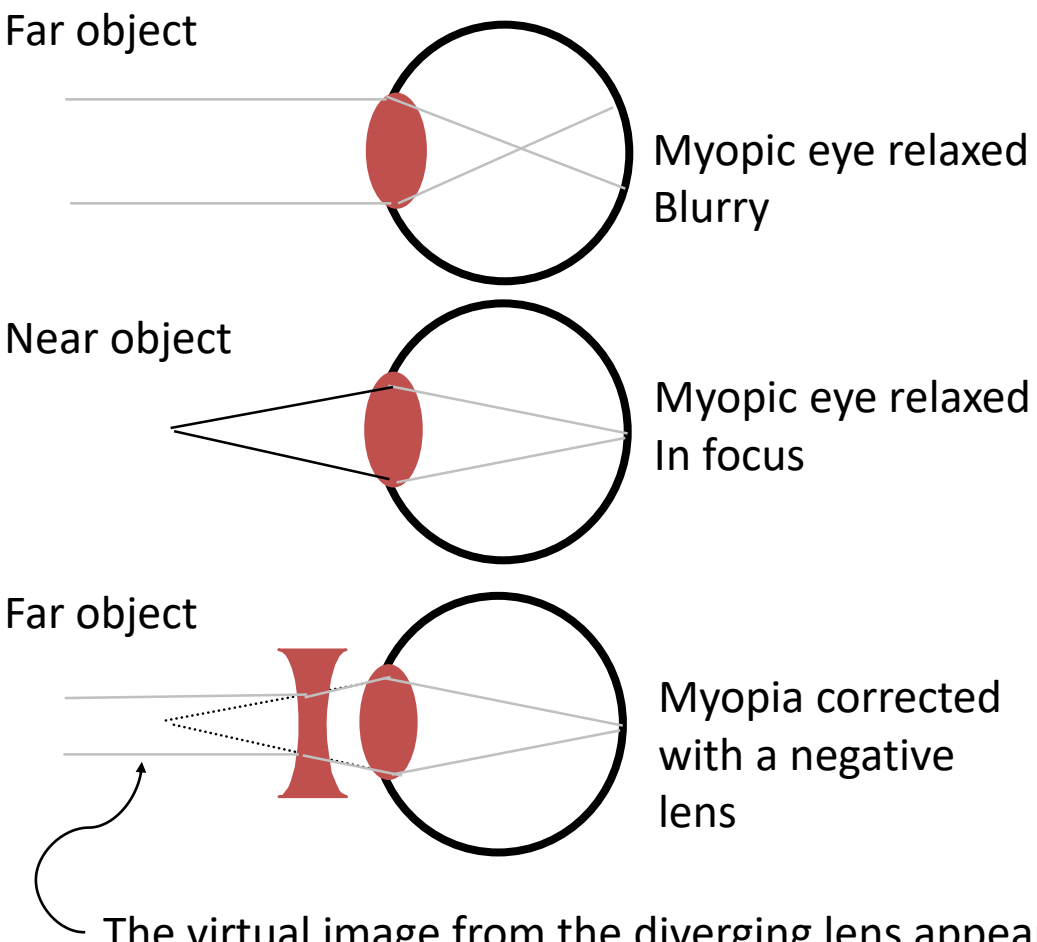
Eye Defects



Eyes at relax state.

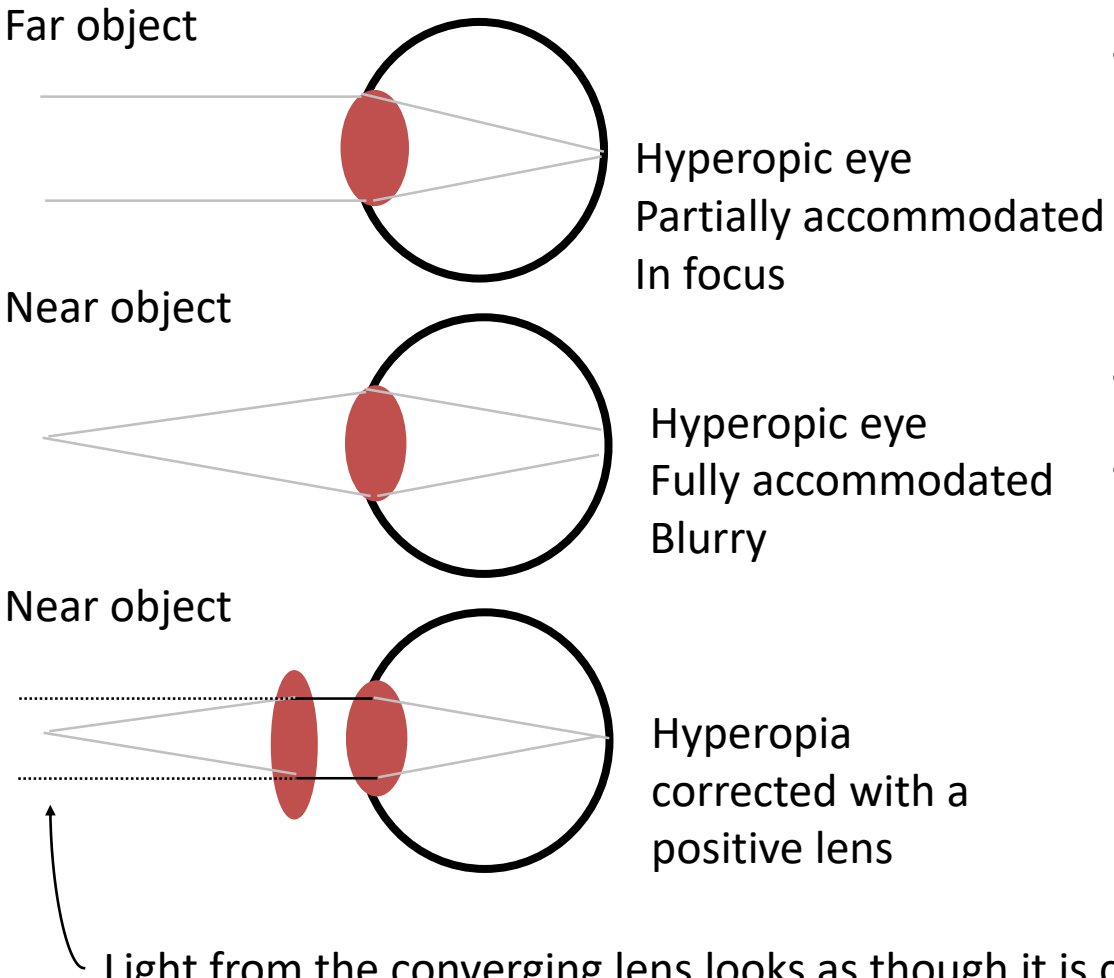
- Image focuses on the retina for a normal eye.
- Distant objects look blurry for a *myopic* (near sighted) eye.
- Near objects look blurry for a *hyperopic* (far sighted) eye.

Myopia - Near sightedness



- Distant objects look blurry because the eye cannot relax any farther so that the image is focused before the retina.
- Near object in focus without accommodation.
- Corrected with a negative lens.

Hyperopia - Far sightedness



- Near objects look blurry because the eye cannot accommodate enough for near objects.
- Far object in focus.
- Corrected with a positive lens.

HMI
HUMAN MACHINE INTERACTION LABORATORY

PERCEIVING OBJECTS AND SCENES

HMI
HUMAN MACHINE INTERACTION LABORATORY

The Challenge of Object Perception

- Human perceptual system
- The stimulus on the receptors is ambiguous.
  - Inverse projection problem: An image on the retina can be caused by an infinite number of objects. How many round objects are there? How large? How far away?
- Objects can be hidden or blurred.
  - Occlusions are common in the environment.

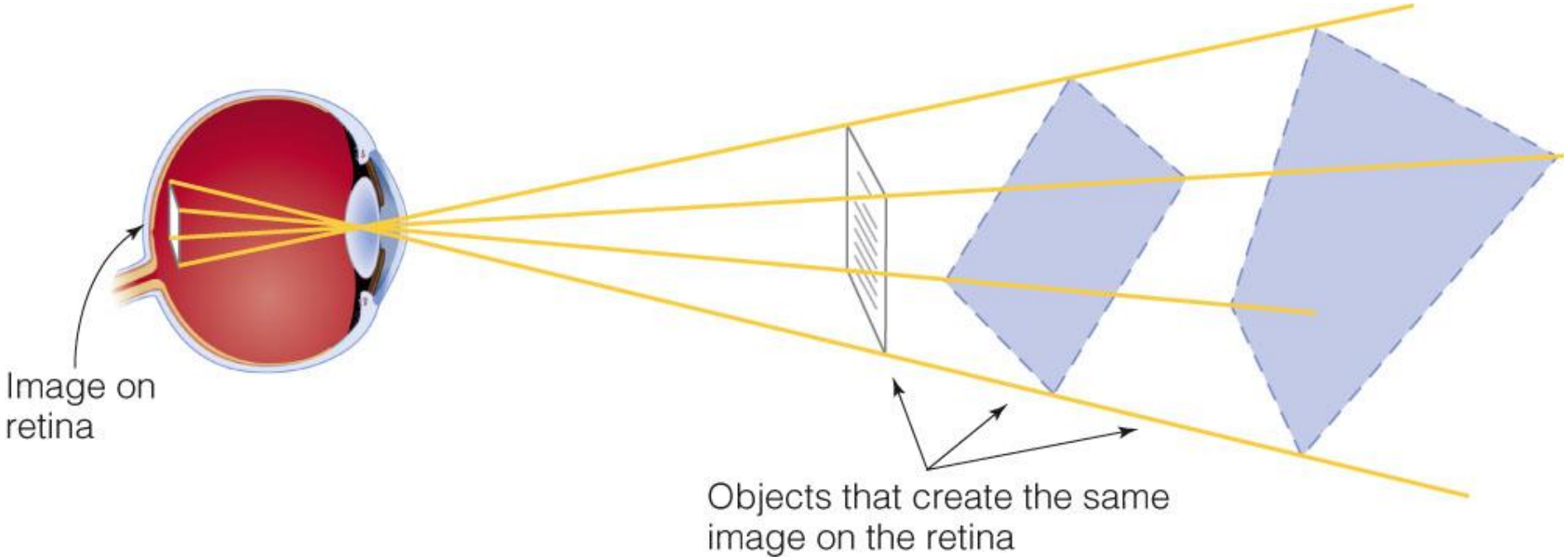


Figure 5.3 The principle behind the inverse projection problem. The page of the book that is near the eye creates a rectangular image on the retina. However, this image could also have been created by the larger more distant square, by the tilted trapezoid and by many other stimuli. This is why we say that the image on the retina is ambiguous.

HMI
HUMAN MACHINE INTERACTION LABORATORY

Viewpoint changes image

- Objects look different from different viewpoints
  - Viewpoint invariance: the ability to recognize an object regardless of the viewpoint
  - This is a difficult task for computers to perform

The Structuralist Approach

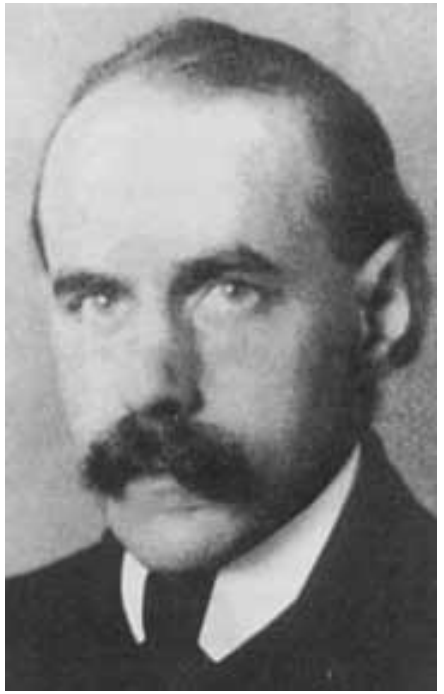
- Approach established by Wundt (late 1800s)
  - States that perceptions are created by combining elements called sensations

Figure 5.9 According to structuralism, a number of sensations (represented by the dots) adds up to create our perception of the face.

Birth of Gestalt Psychology

Structuralism could not explain apparent movement produced by a toy stroboscope.

Stimulated the founding of Gestalt psychology in the 1920s by Wertheimer, Koffka, and Kohler



Max Wertheimer

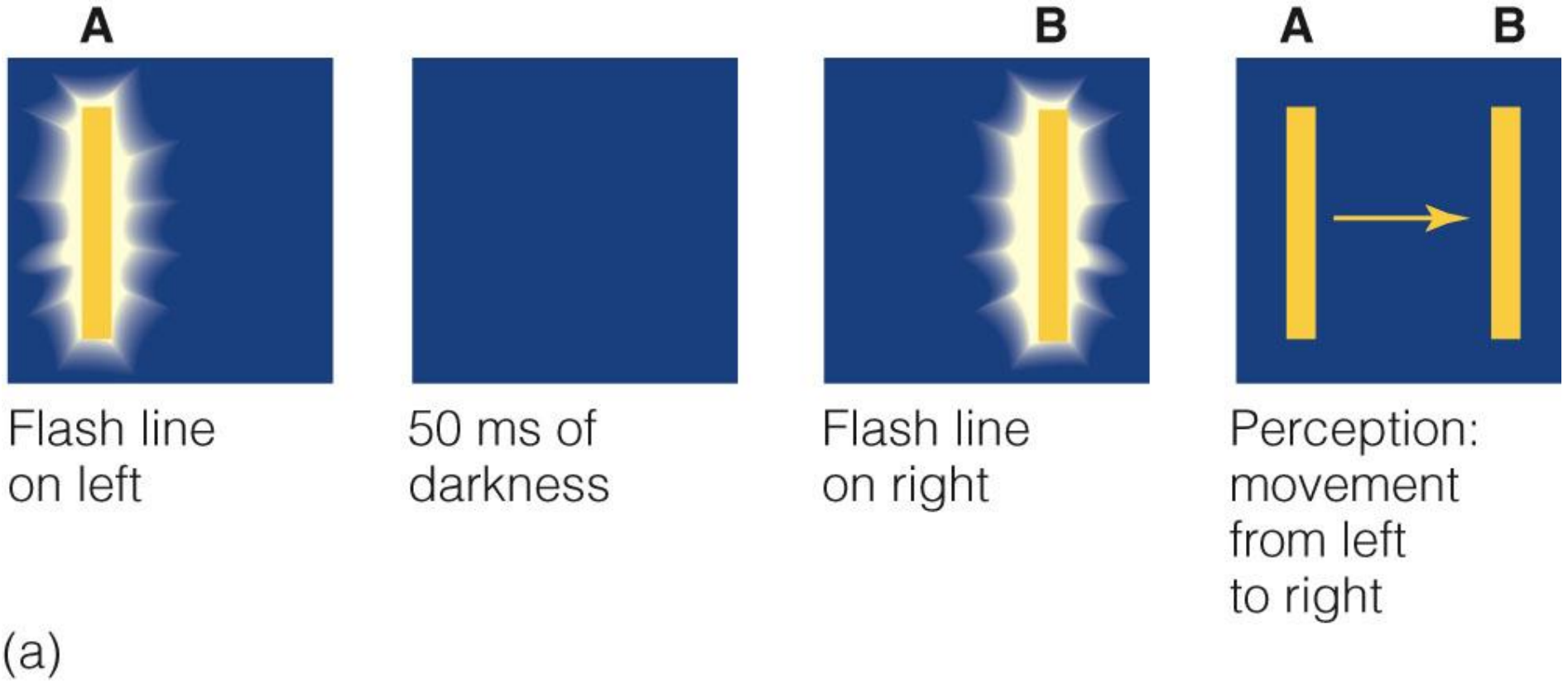
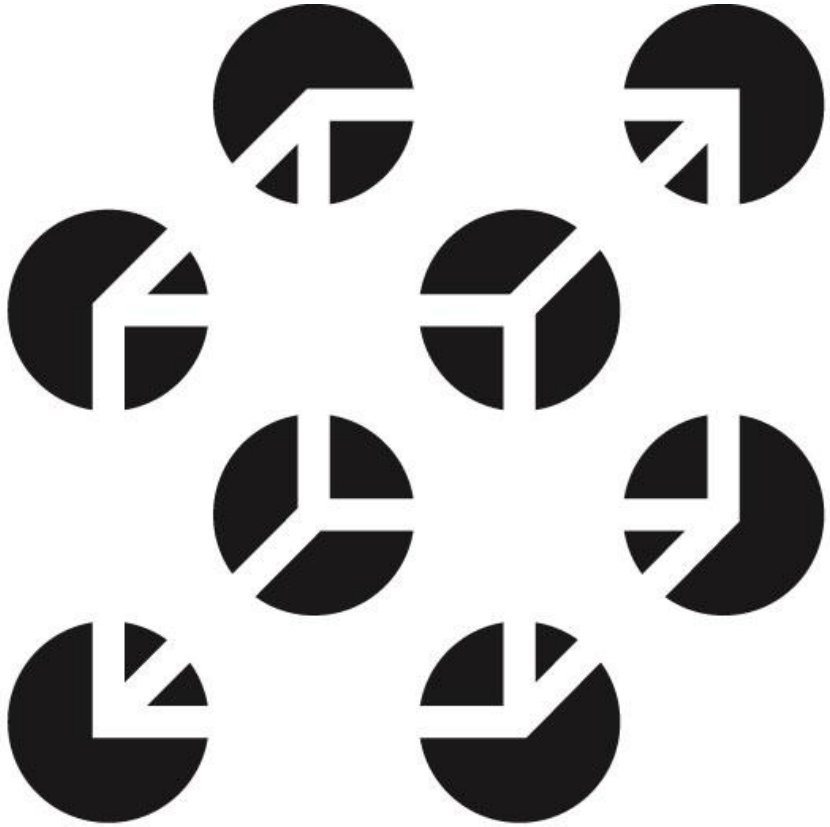


Figure 5.10 (a) Wertheimer's demonstration of movement perception.

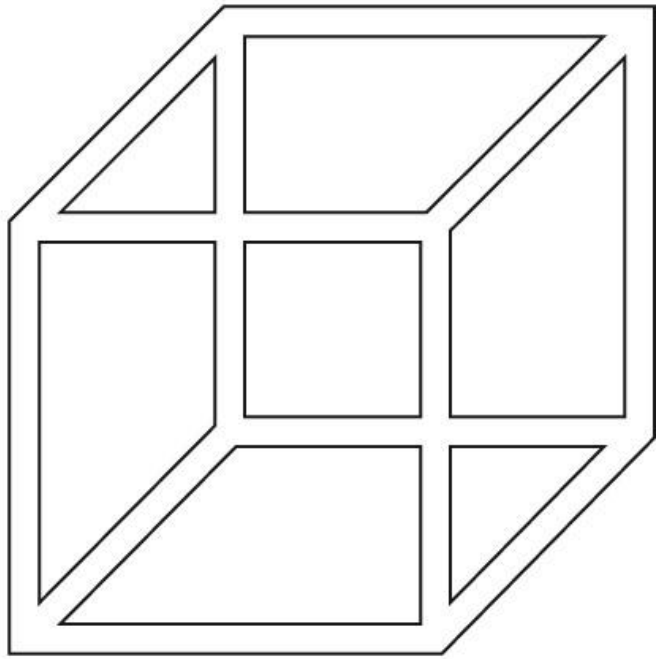
HMI
HUMAN MACHINE INTERACTION LABORATORY

The Gestalt Approach

- The whole differs from the sum of its parts.
  - Perception is not built up from sensations, but is a result of perceptual organization.
- Develop some principles of perceptual organization.



(a)



(b)

Figure 5.11 (a) This can be seen as a cube floating in front of eight discs or as a cube seen through eight holes. In the first case, the edges of the cube appear as illusory contours.
(b) The cube without the black circles.

Some Gestalt “Laws”

- Pragnanz – “good figure”, every stimulus is seen as simply as possible
- Similarity - similar things are grouped together

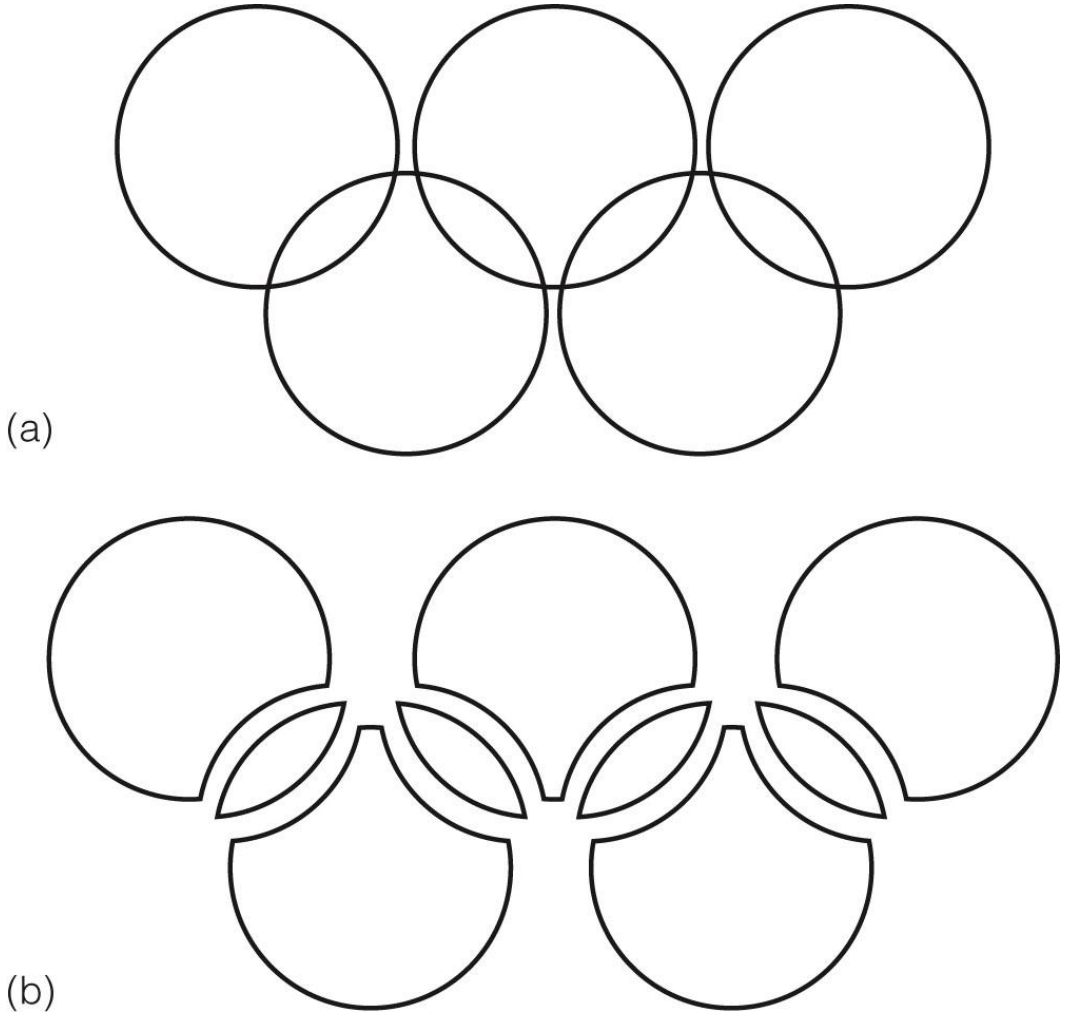


Figure 5.13 (a) This is usually perceived as five circles, not as the nine shapes in (b). An example of good figure.

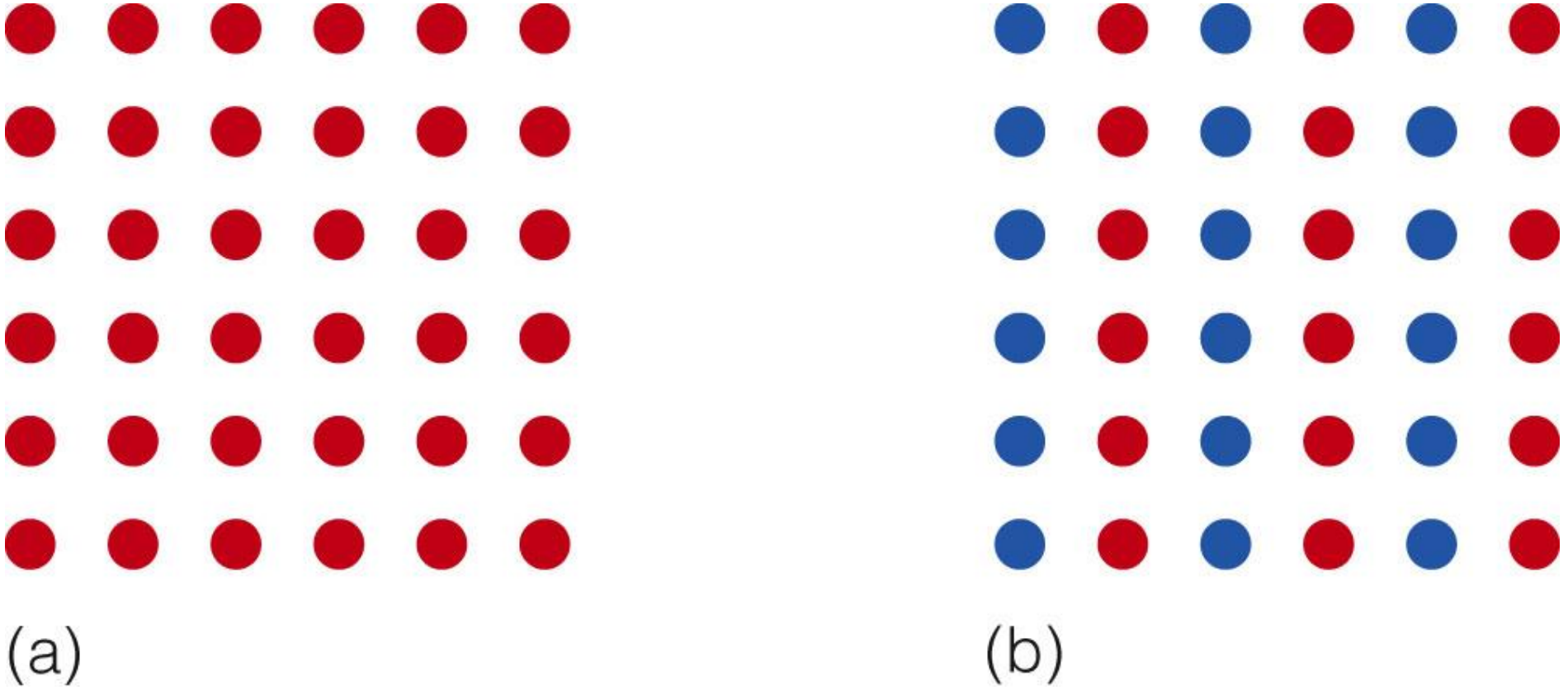


Figure 5.14 (a) Perceived as horizontal rows or vertical columns or both. (b) Perceived as vertical columns. An example of similarity.

More Gestalt Laws

- Good continuation - connected points resulting in straight or smooth curves belong together
  - Lines are seen as following the smoothest path
- Proximity - things that are near to each other are grouped together
- Common region - elements in the same region tend to be grouped together

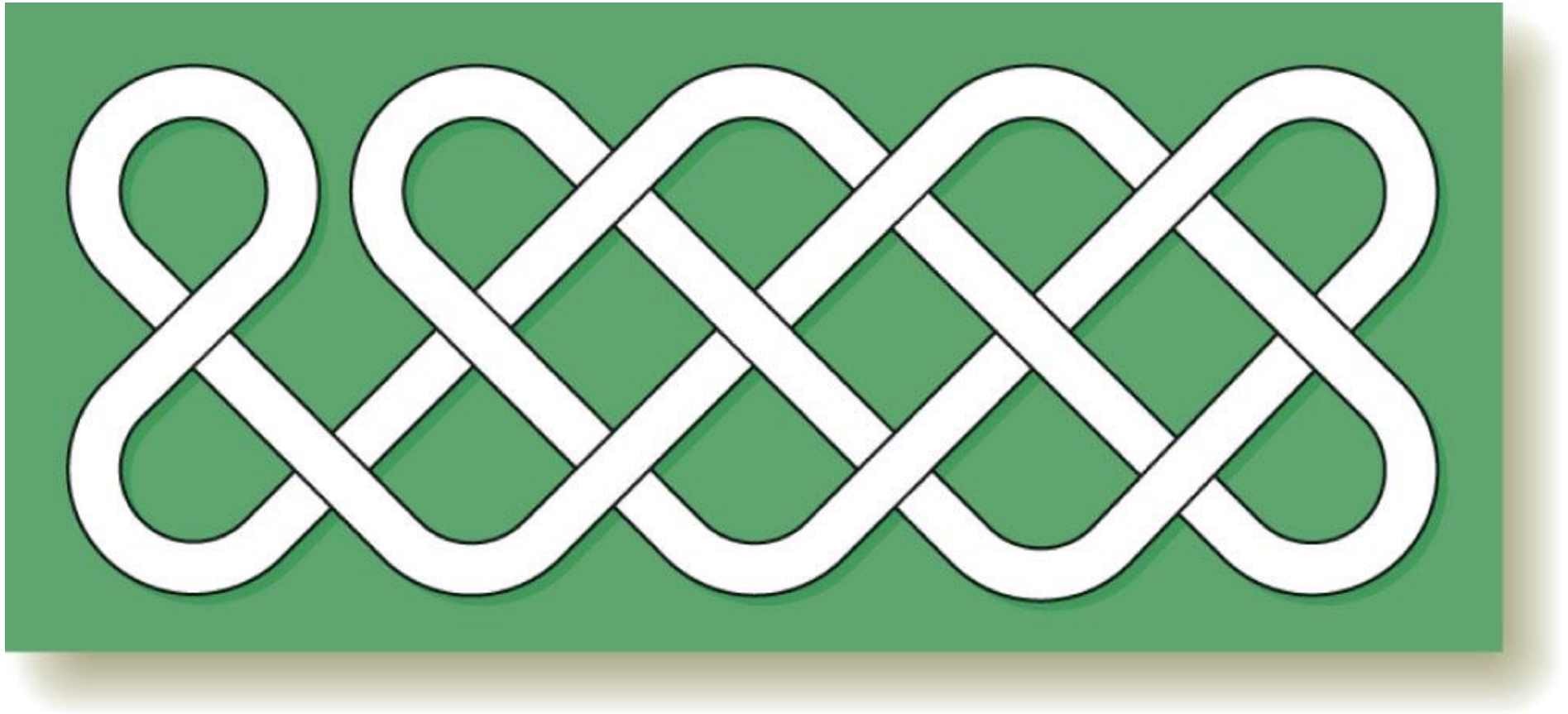


Figure 5.17 Because of good continuation, we perceive this pattern as continuous interwoven strands.

HMI
HUMAN MACHINE INTERACTION LABORATORY

Principles of Perceptual Organization - continued

- Meaningfulness or familiarity - things form groups if they appear familiar or meaningful
- Uniform connectedness - connected region of visual properties are perceived as single unit
- Synchrony - elements occurring at the same time are seen as belonging together
- Common fate - things moving in same direction are grouped together

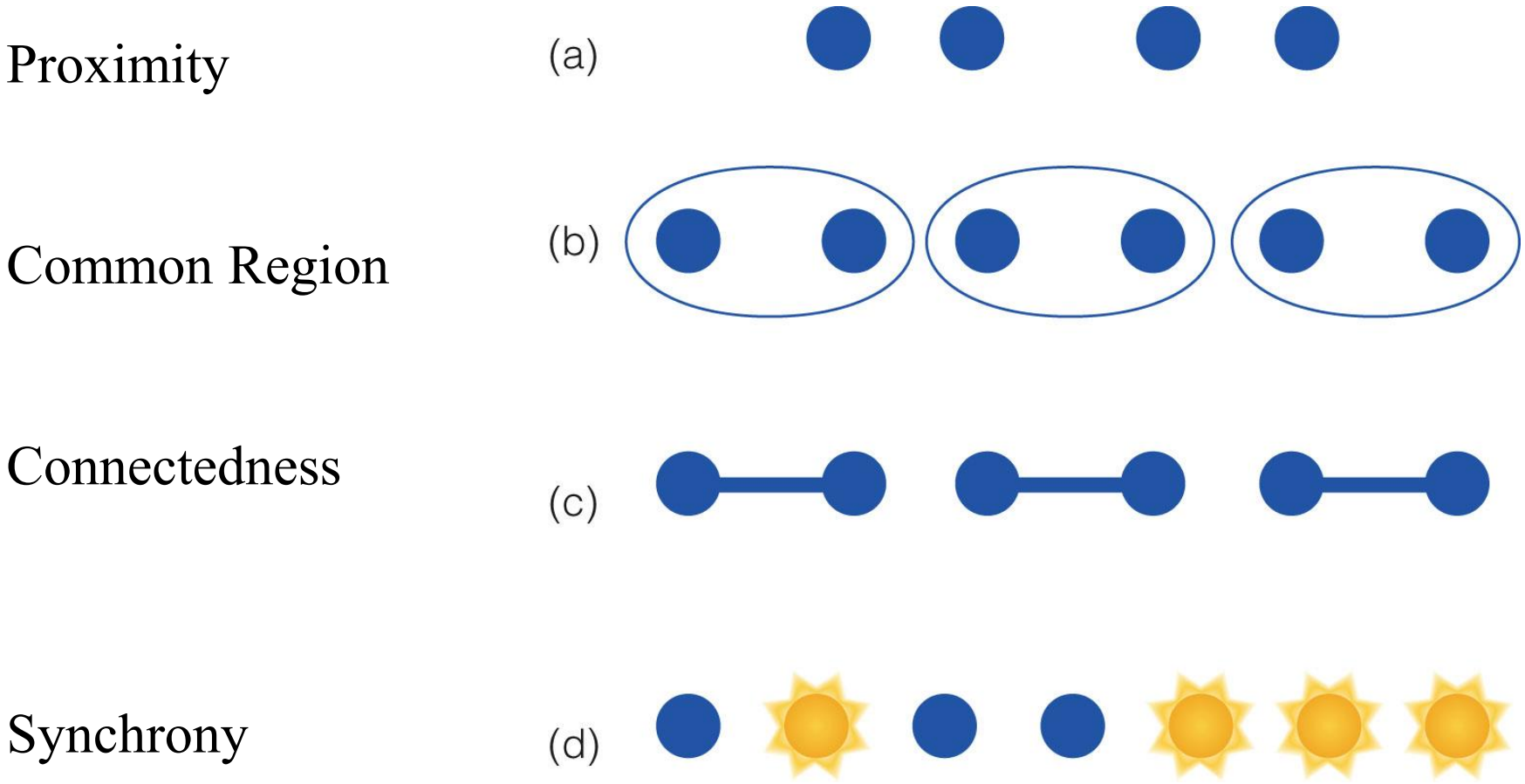


Figure 5.18 Grouping by (a) proximity; (b) common region; (c) connectedness; and (d) synchrony. Synchrony occurs when yellow lights blink on and off together.

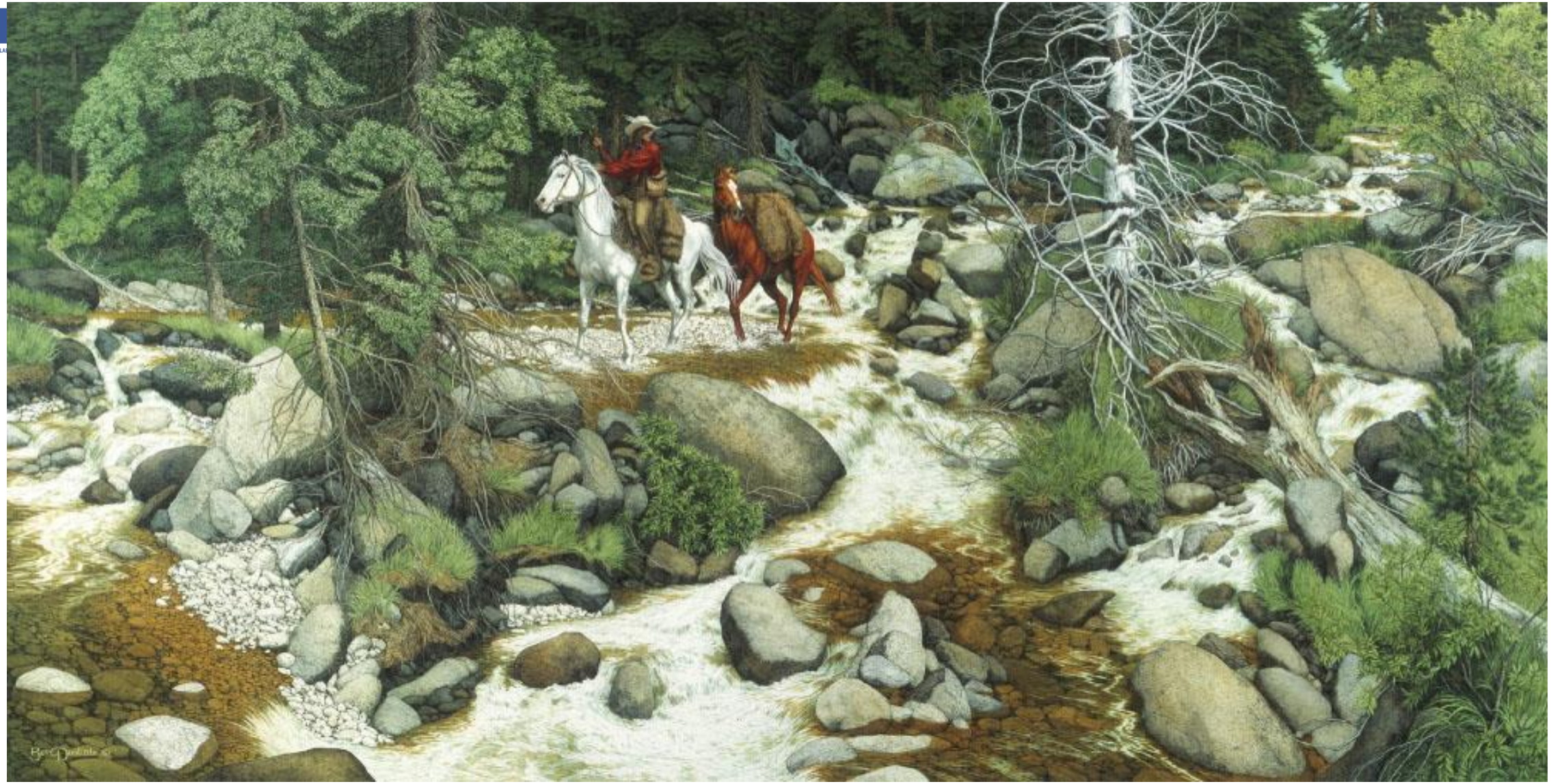


Figure 5.20 *The Forest Has Eyes* by Bev Doolittle (1984). Can you find the 12 faces in this picture?

HMI
HUMAN MACHINE INTERACTION LABORATORY

Perceptual Segregation

- Figure-ground segregation - determining what part of environment is the figure so that it “stands out” from the background

Figure 5.21 A version of Rubin’s reversible face-vase figure.

Properties of figure and ground

- The figure is more “thinglike” and more memorable than ground.
- The figure is seen in front of the ground.
- The ground is more uniform and extends behind figure.
- The contour separating figure from ground belongs to the figure (border ownership).

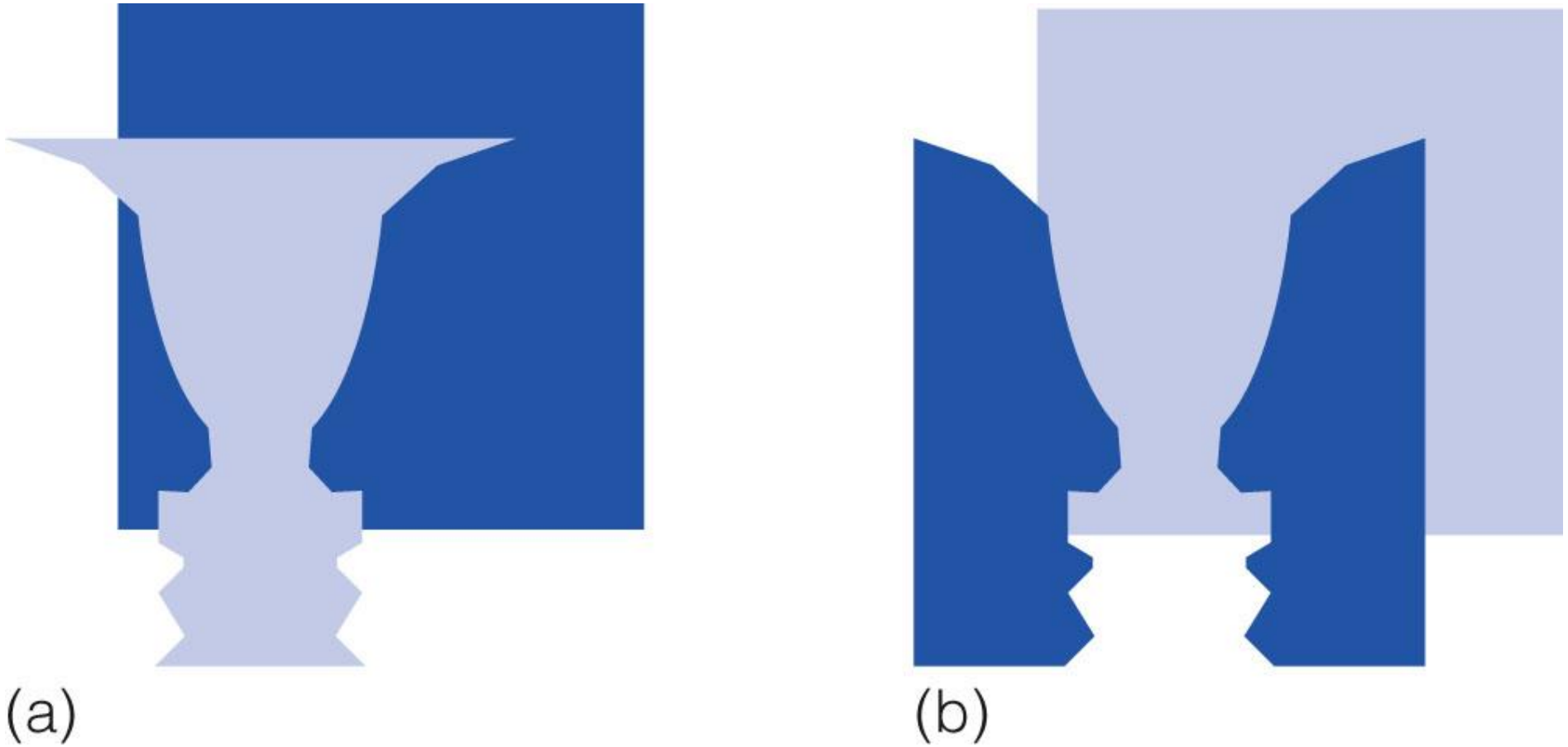


Figure 5.22 (a) When the vase is perceived as figure, it is seen in front of a homogeneous dark background.

(b) When the faces are seen as figure, they are seen in front of a homogeneous light background.

Horizontal vs Vertical. Which is figure and which is ground for each?

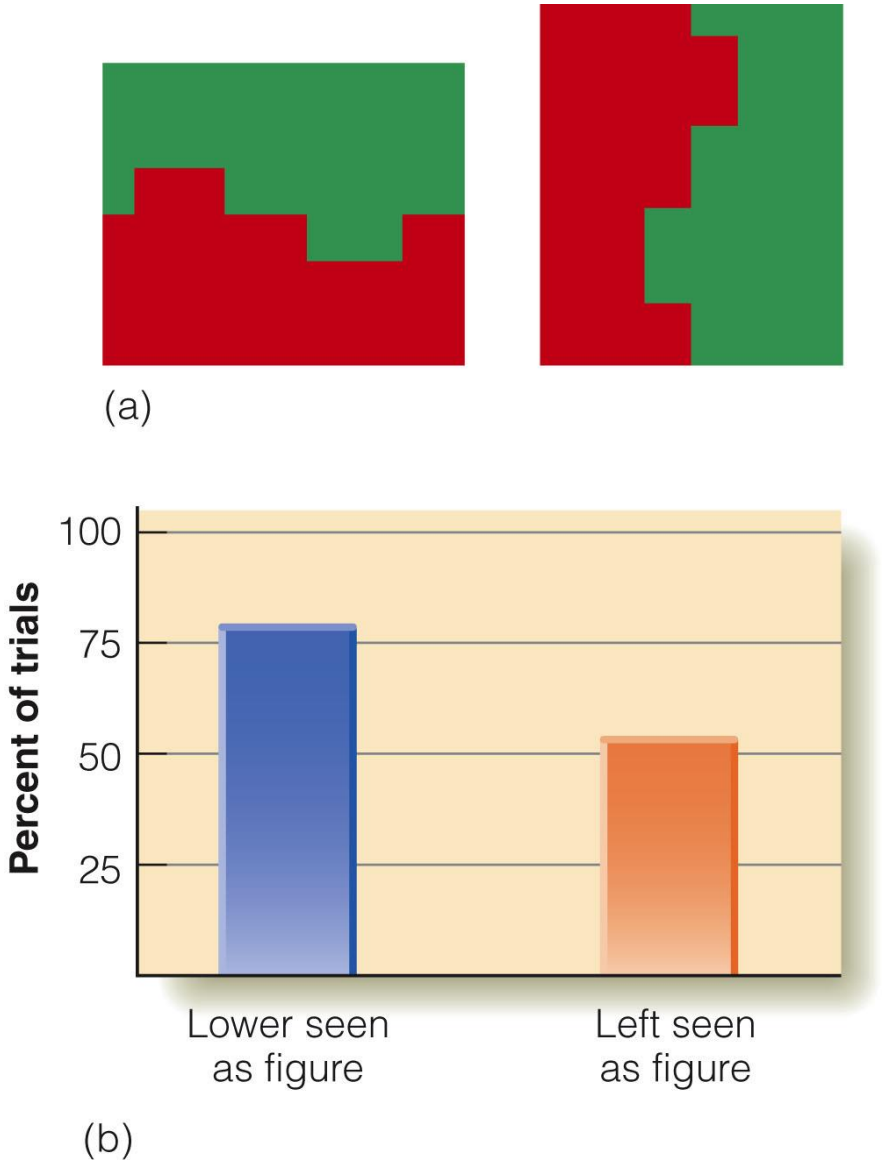


Figure 5.24 (a) Stimuli from Vecera et al. (2002).
(b) Percentage of trials on which lower or left areas were seen as figure.

Figure-Ground Segregation - continued

- Factors that determine which area is figure:
  - Elements located in the lower part of displays
  - Units that are symmetrical
  - Elements that are small
  - Units that are oriented vertically
  - Elements that have meaning

Perceiving Scenes and Objects in Scenes

- A scene contains:
  - background elements.
  - objects organized in meaningful ways with each other and the background.
- Difference between objects and scenes
  - A scene is *acted within*
  - An object is *acted upon*

Perceiving Scenes and Objects in Scenes - continued

- Research on perceiving gists of scenes
  - Potter showed that people can do this when a picture is only presented for 1/4 second
  - Fei-Fei used masking to show that the overall gist is perceived first followed by details.
  - VL 5.17

Global image features of scences

- Global image features of scenes
  - Degree of naturalness (forest vs street)
  - Degree of openness (visible horizon)
  - Degree of roughness (beach vs forest)
  - Degree of expansion (or convergence)
  - Color (blue for ocean, green for lawn)
- Such features are holistic and perceived rapidly

HMI
HUMAN MACHINE INTERACTION LABORATORY

Regularities in the Environment

- Physical regularities - regularly occurring physical properties
  - Oblique effect - people perceive horizontals and vertical more easily than other orientations
  - Uniform connectedness - objects are defined by areas of the same color or texture
  - Light-from-above heuristic - light in natural environment comes from above us

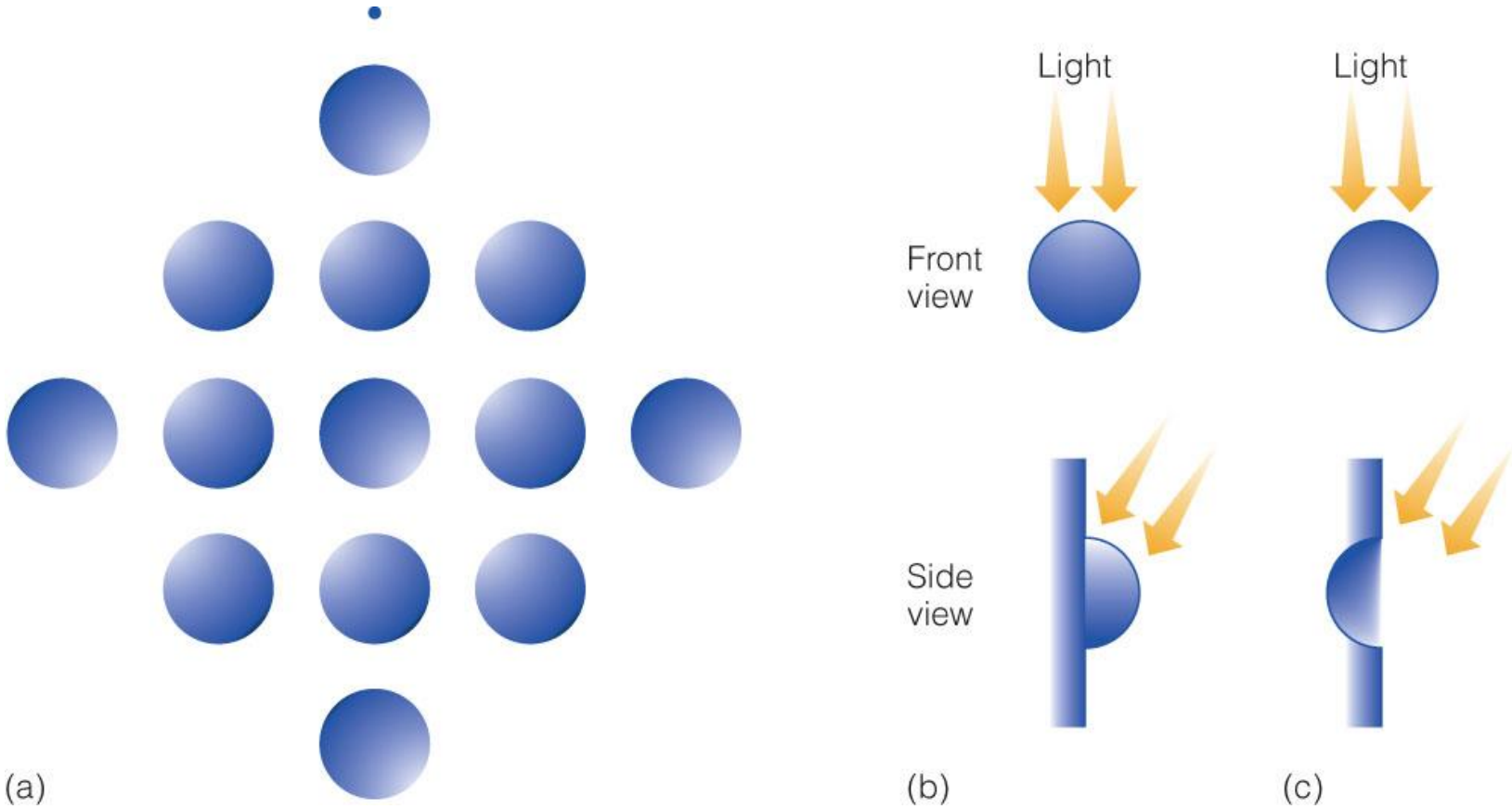


Figure 5.37 (a) Some of these discs are perceived as jutting out and some are perceived as indentations. Why? Light coming from above would illuminate (b) the top of a shape that is jutting out and (c) the bottom of the indentation.

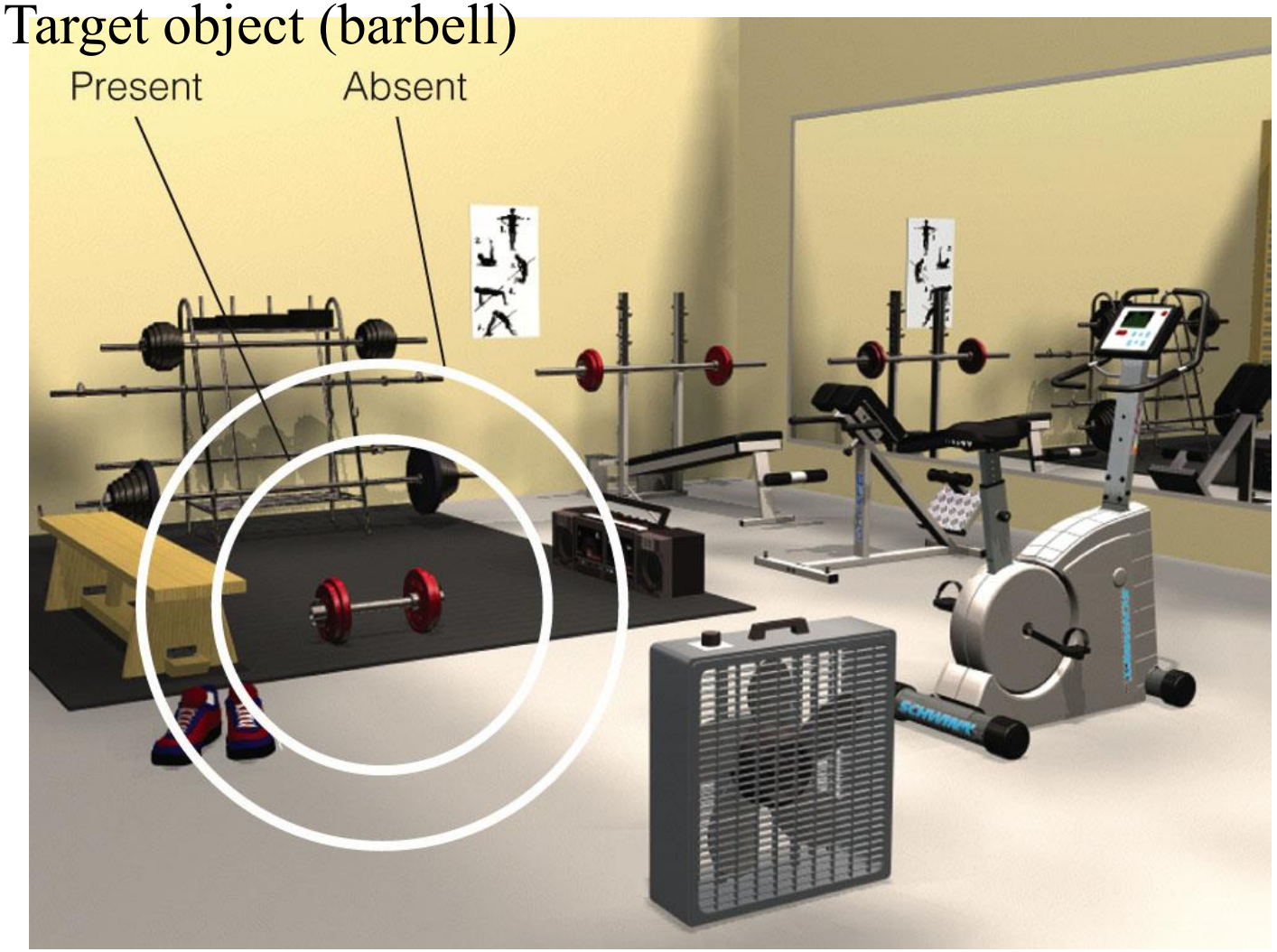


Figure 5.39 (a) One of the scenes shown to Hollingworth's (2005) observers. In this picture the target object is the barbell, although observers do not know this when they are viewing the scene. "Non-target" scenes are the same but do not include the target. Ignore circles.

Semantic regularities

--characteristics associated with the functions of scenes

– Study by Hollingworth (2005)

- Observers were presented with a scene either with or without a target object.
- They then saw the target followed by a blank screen and were asked where the object was in the scene or where they would expect it to be.
- Both groups could accurately predict where the object would be.

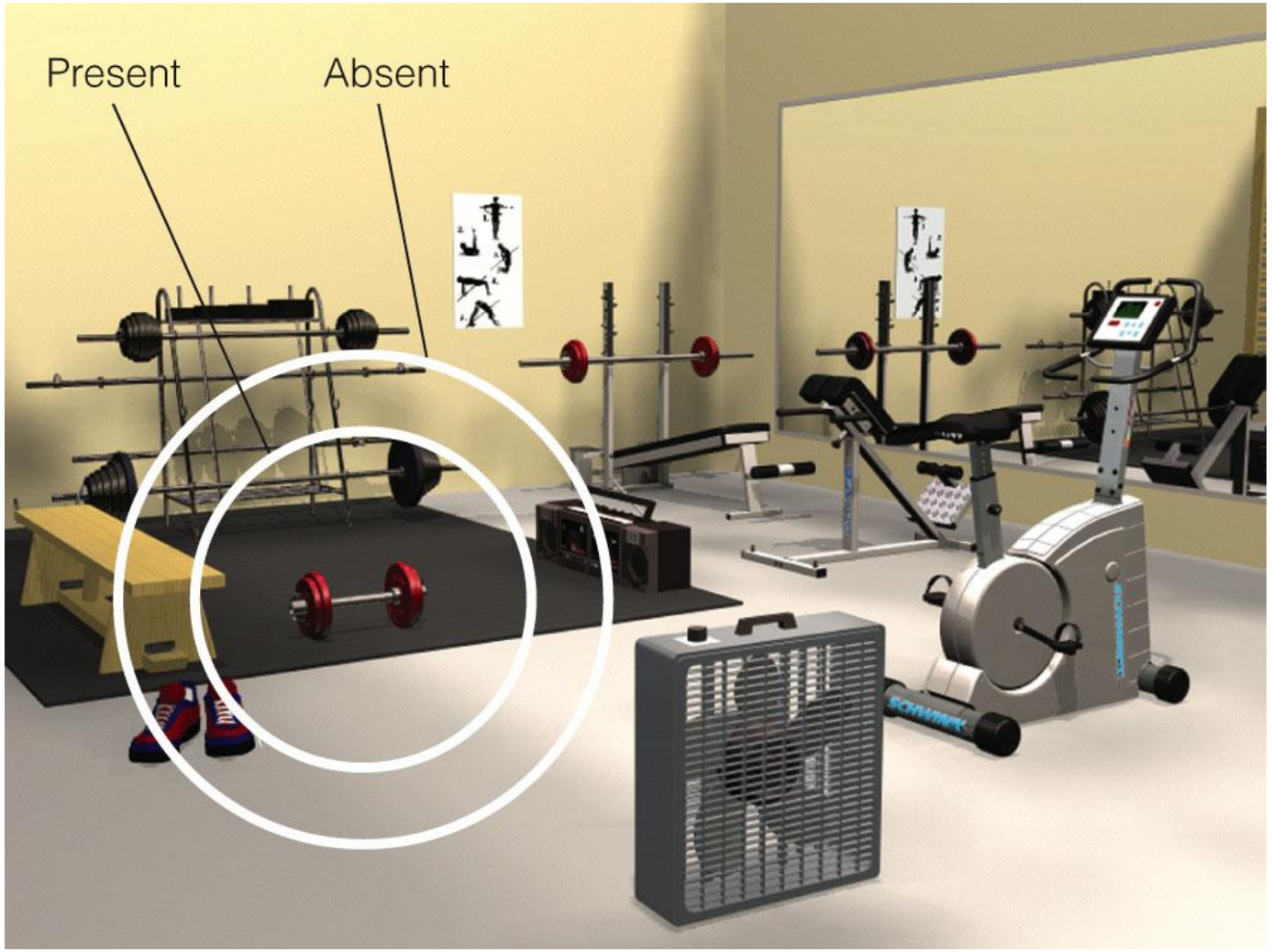


Figure 5.39 Circles indicate the observers' judgments of the position of the target object for trials in which they had seen the object in the scene (small circle) and trials in which the object had not appeared in the scene (larger circle).

Regularities in the Environment

- Palmer experiment
  - Observers saw a context scene flashed briefly, followed by a target picture.
  - Results showed that:
    - Targets congruent with the context were identified 80% of the time .
    - Targets that were incongruent were only identified 40% of the time.

Scene presented first

Which object?

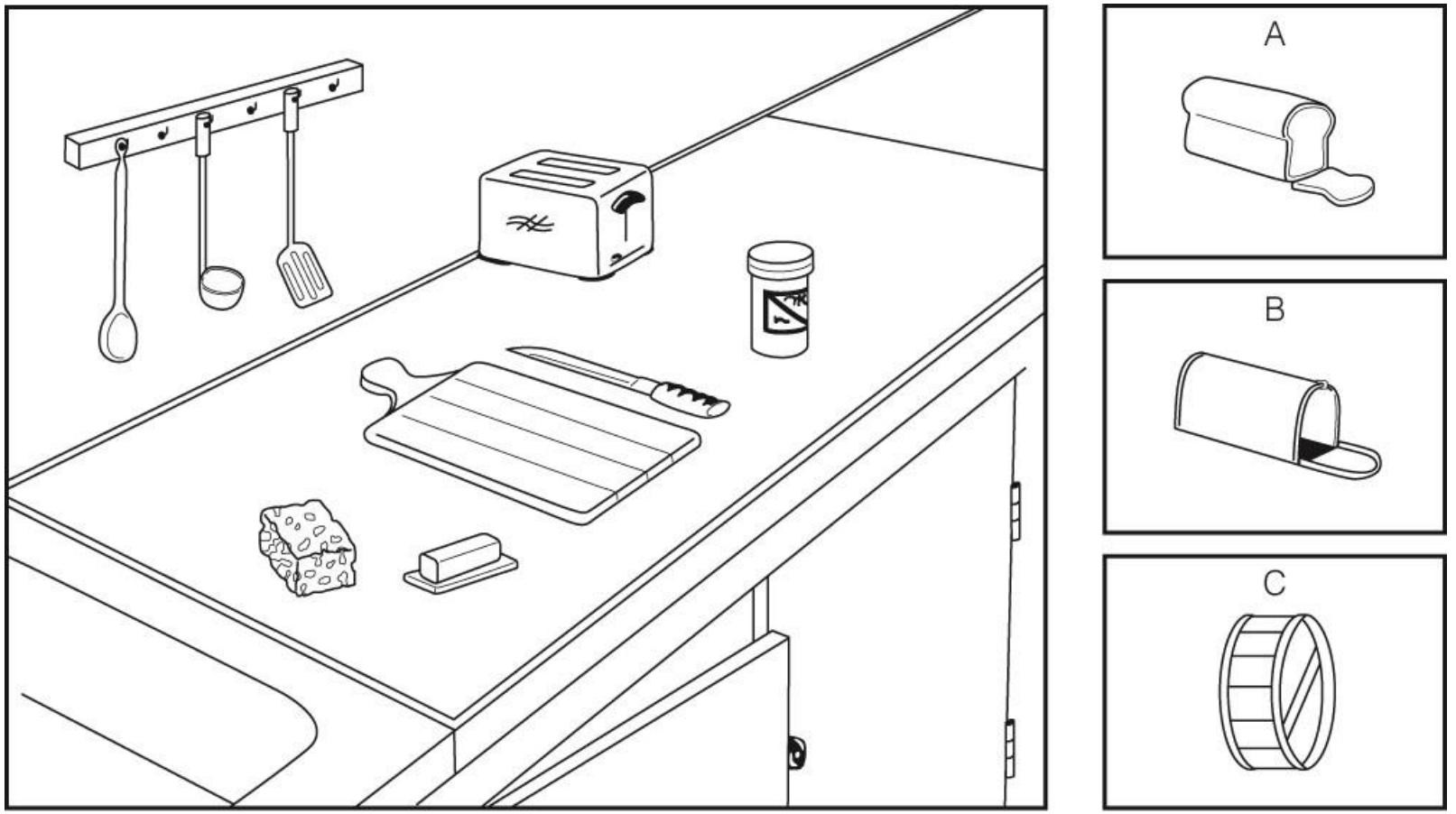


Figure 5.40 Stimuli used in Palmer's (1975) experiment. The scene at the left is presented first, and the observer is then asked to identify one of the objects on the right briefly flashed.

Role of Inference in Perception

- Theory of “unconscious” inference
  - Created by Helmholtz (1866/1911) to explain why stimuli can be interpreted in more than one way
  - Main Principle - perceptions are result of “unconscious assumptions” or inferences about the environment
  - Likelihood principle - objects are perceived based on what is most likely to have caused the pattern

Which if the best (most likely) display?

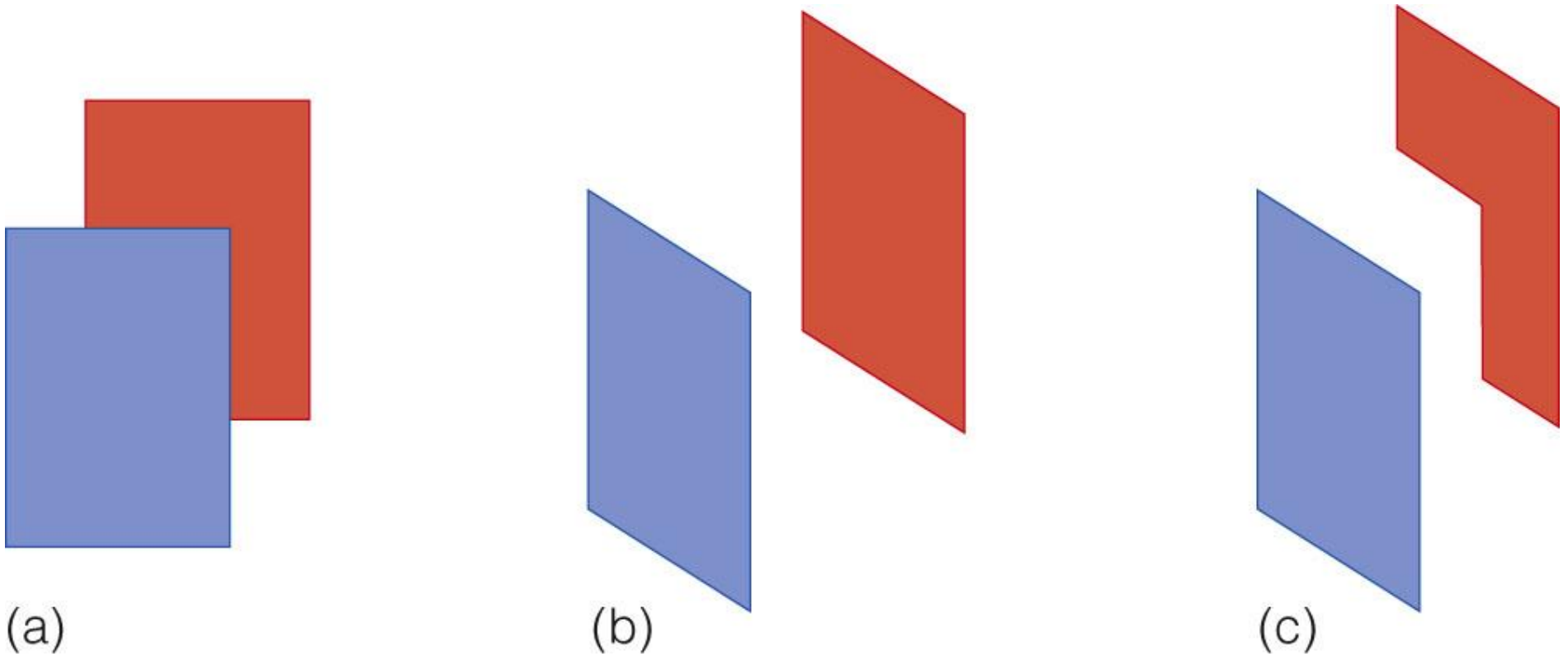
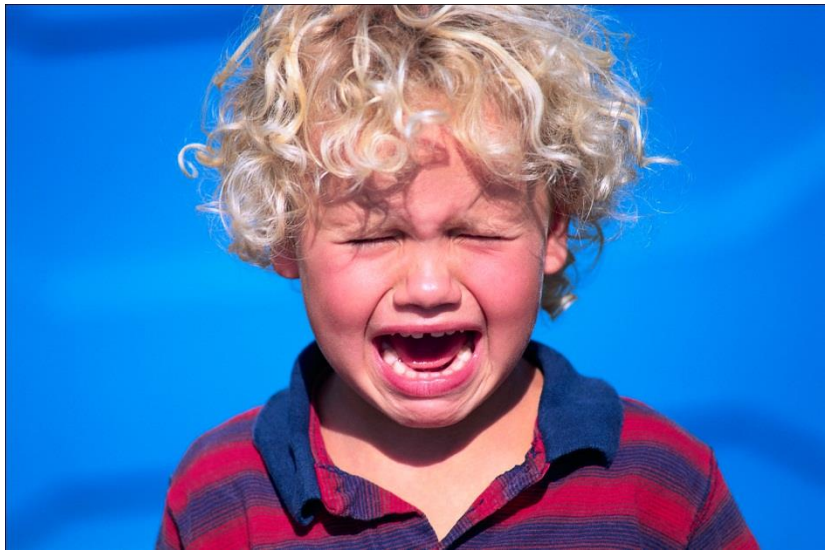


Figure 5.42 The display in (a) is usually interpreted as being (b) -- a blue rectangle in front of a red rectangle. It could, however, be (c) -- a blue rectangle and an appropriately positioned six-sided red figure.

HMI
HUMAN MACHINE INTERACTION LABORATORY

Percieving Faces



HMI
HUMAN MACHINE INTERACTION LABORATORY

Brain Modules Involved in Perceiving Faces

- Fusiform face area (FFA) - responds only to faces
- Amygdala (AG) - activated by emotional aspects of faces
- Superior temporal sulcus (STS) - responds to where the person is looking and to mouth movements
- Frontal Cortex (FC) - activated when evaluating facial attractiveness

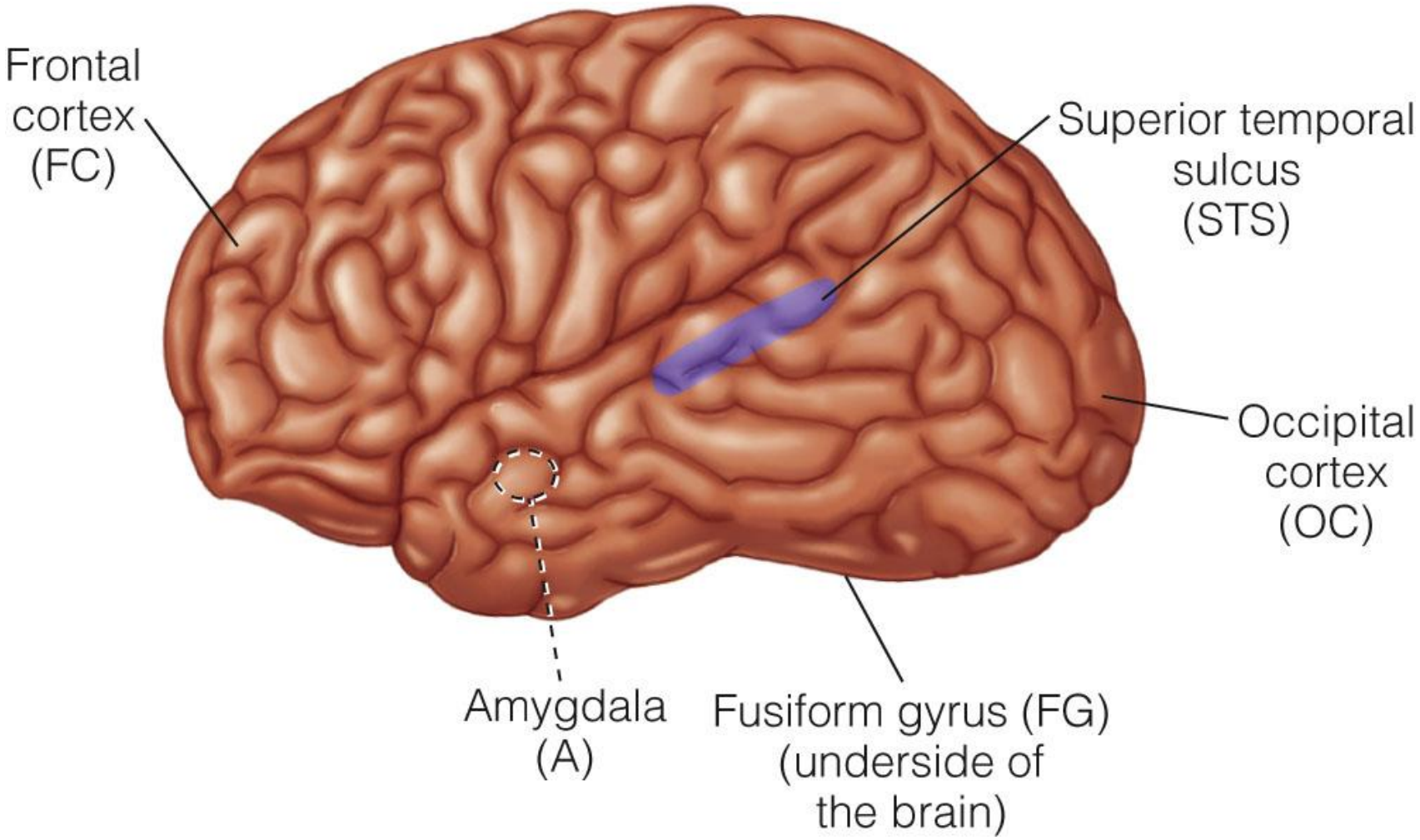
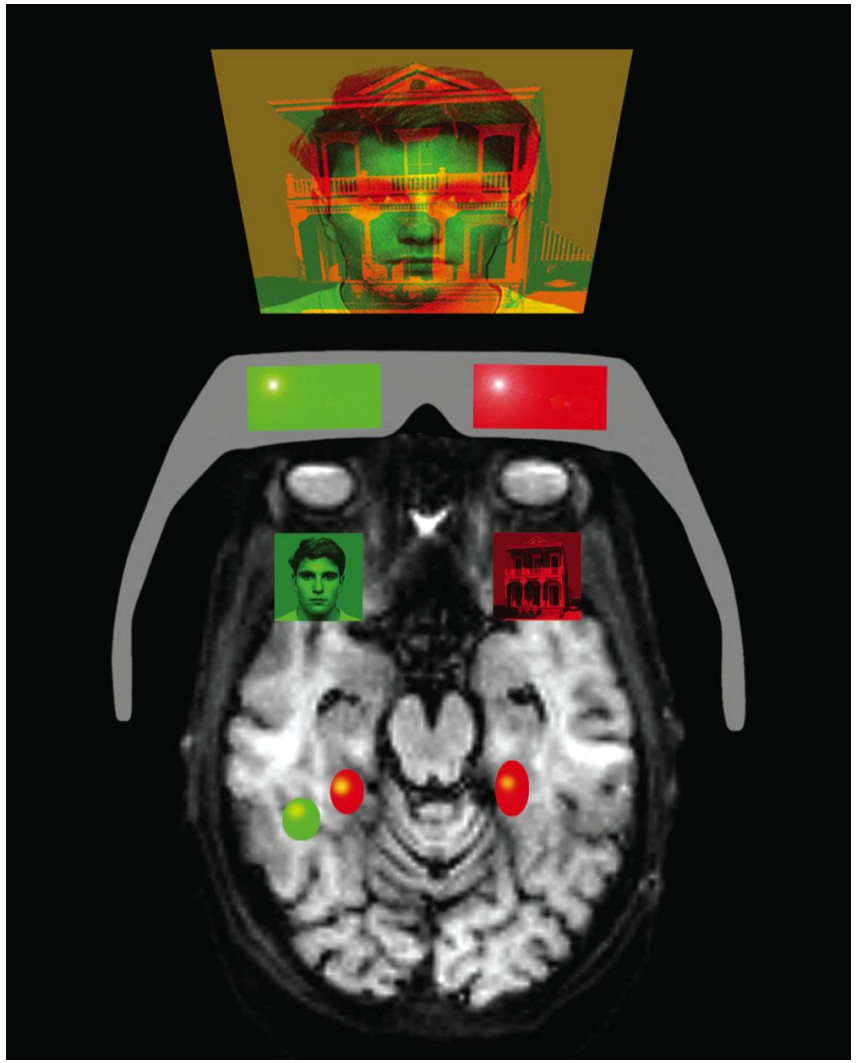


Figure 5.45 The human brain, showing some of the areas in involved in perceiving faces: FC = frontal cortex; STS = superior temporal sulcus; FG = fusiform gyrus; OC = occipital cortex; AG = amygdala. Note that the labels indicate a general area of cortex, but not the overall extent of the area. Also, the amygdala is located deep inside the cortex, approximately under the label shown here.

Perceiving an Object - Face or House?

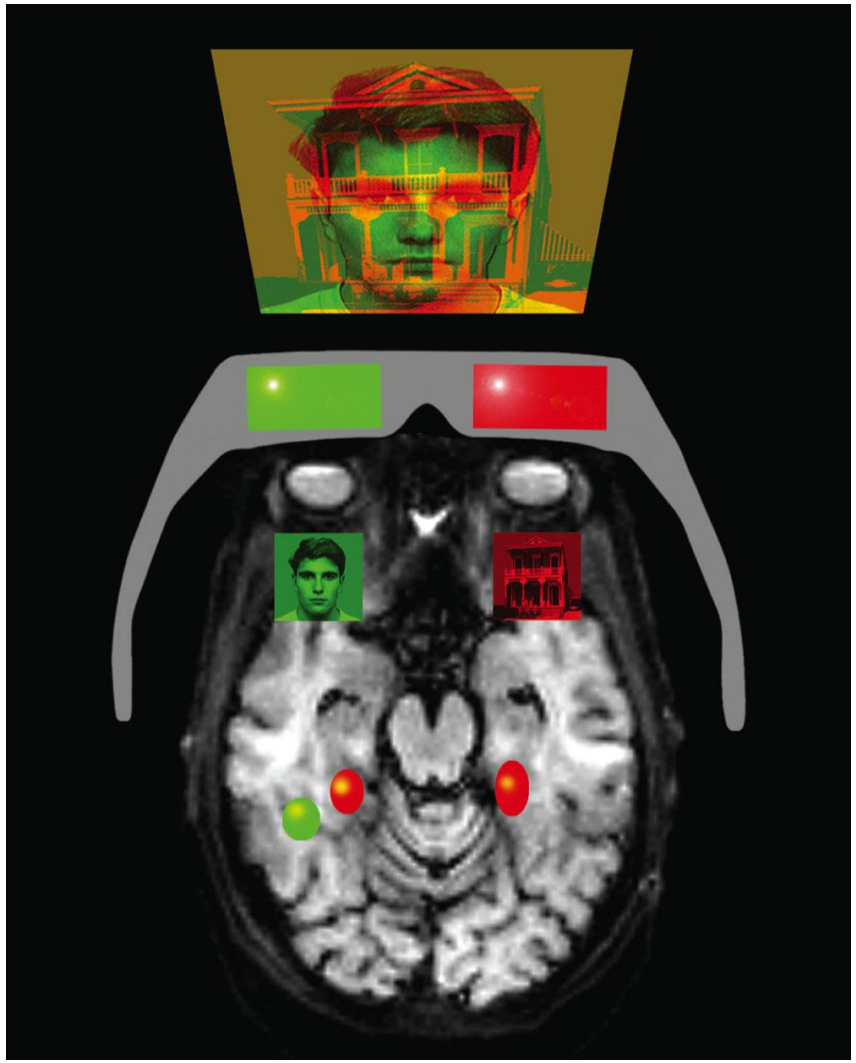
- Experiment by Tong et al.
  - Binocular rivalry used again with people
  - Picture of a house shown to one eye and a face to another
  - Participants pushed button to indicate perception.
  - fMRI showed an increase in activity in
    - Parahippocampal place area for the house
    - Fusiform face area for the face

Observers in Tong and coworkers (1998) experiment viewed the overlapping red house and green face through red-green glasses, so the house image was presented to the right eye and the face image to the left eye.



House→ PPA (right and left)
Face-→ FFA (left hemi only)

When the observers perceived the house, activity occurred in the parahippocampal place area (PPA), in the left and right hemispheres (red ellipses). When the observers perceived the face, activity occurred in the fusiform face area (FFA) in the left hemisphere (green ellipse).



House→ PPA (right and left)
Face-→ FFA (left hemi only)

ILLUTION

HMI
HUMAN MACHINE INTERACTION LABORATORY

GIÁC QUAN + NIỀM TIN = THỰC TẠI

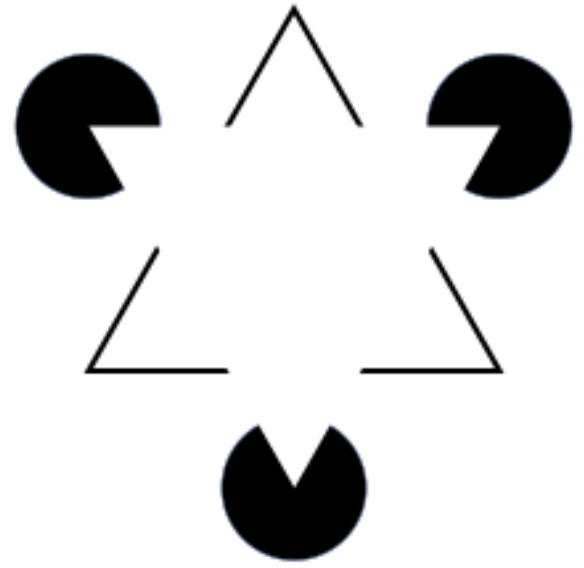
GIÁC QUAN + NIỀM TIN = ẢO GIÁC

HMI
HUMAN MACHINE INTERACTION LABORATORY

Các dạng ảo ảnh thị giác

- Ảo ảnh liên quan đến tổ chức nhận thức
- Ảo ảnh độ sâu và nhận thức chuyển động
- Ảo ảnh mù độ cong
- Ảo ảnh nhận thức tương lai
- Ảo ảnh bất biến mầu sắc – độ sáng

Ảo ảnh liên quan đến tổ chức nhận thức (Perceptual organization)



Kanizsa ‘s triangle

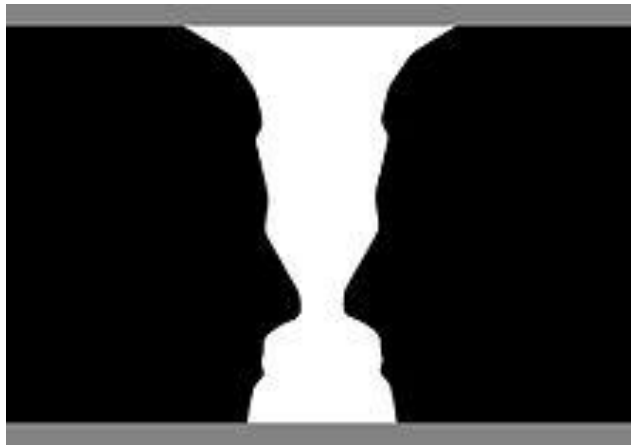
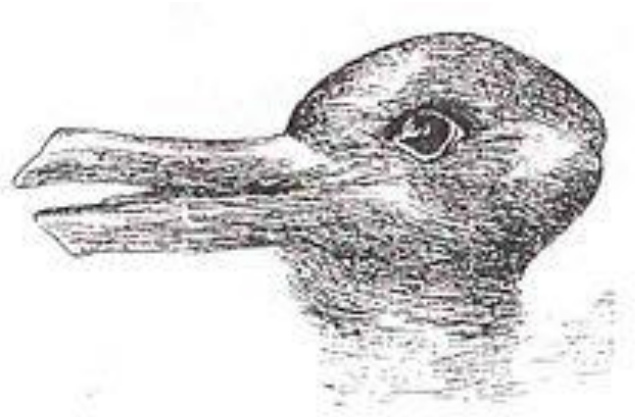
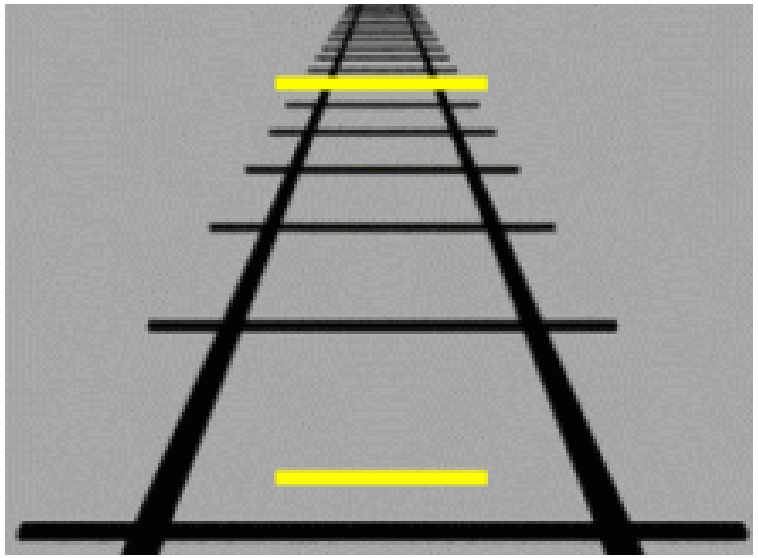


figure groud illusion

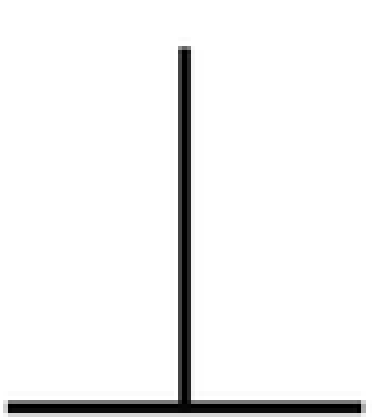


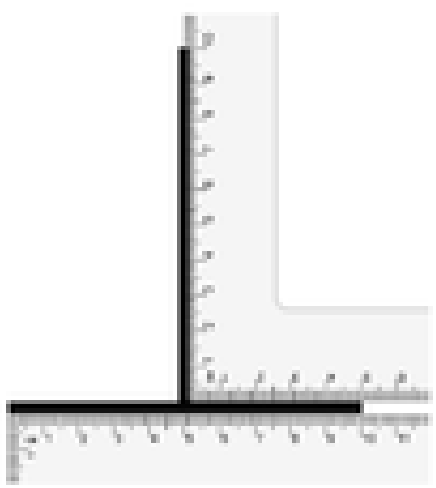
Rabbbit-duck illusion

Độ sâu và nhận thức chuyển động (Depth and motion perception)



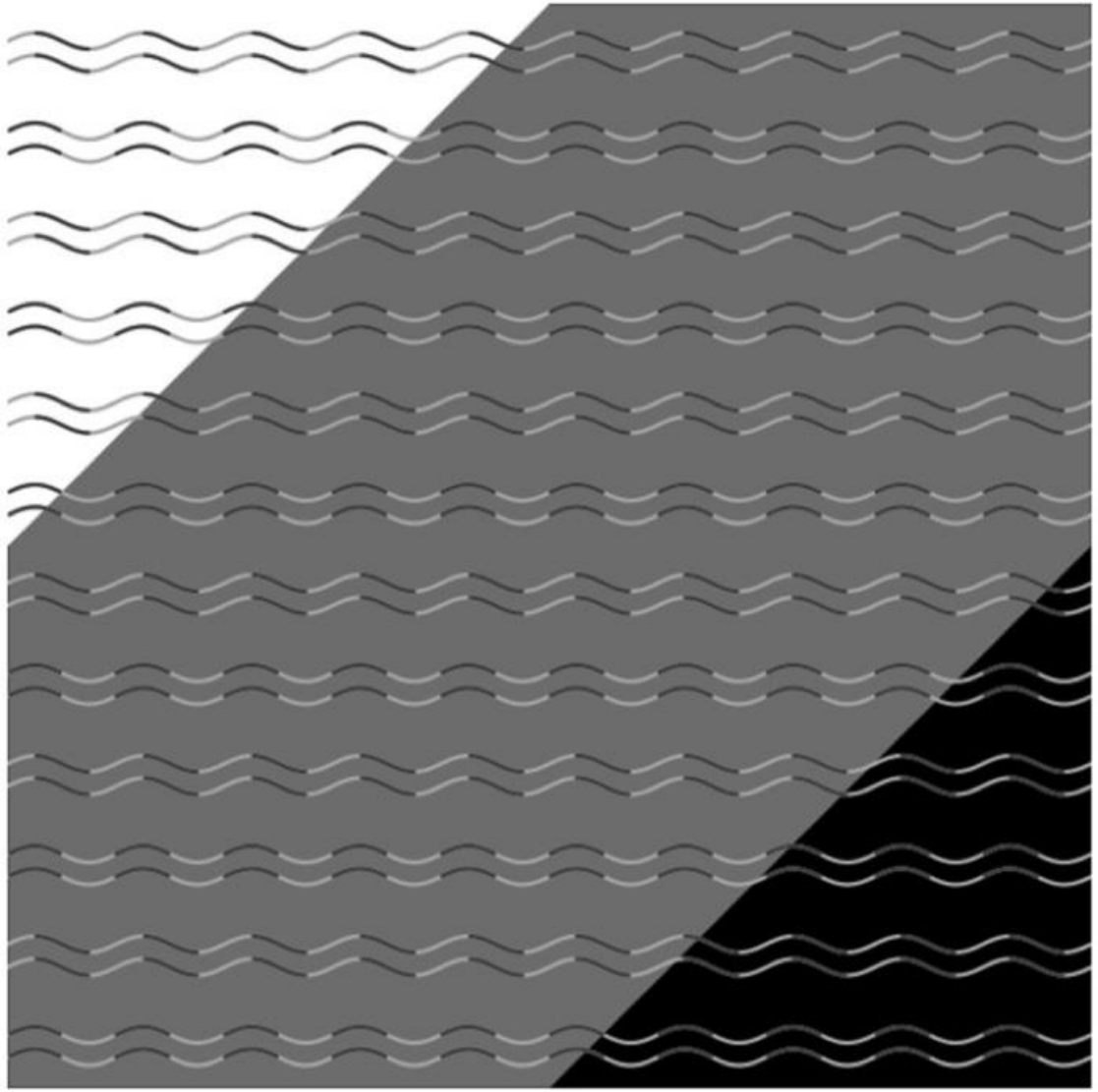
Ponzo illusion



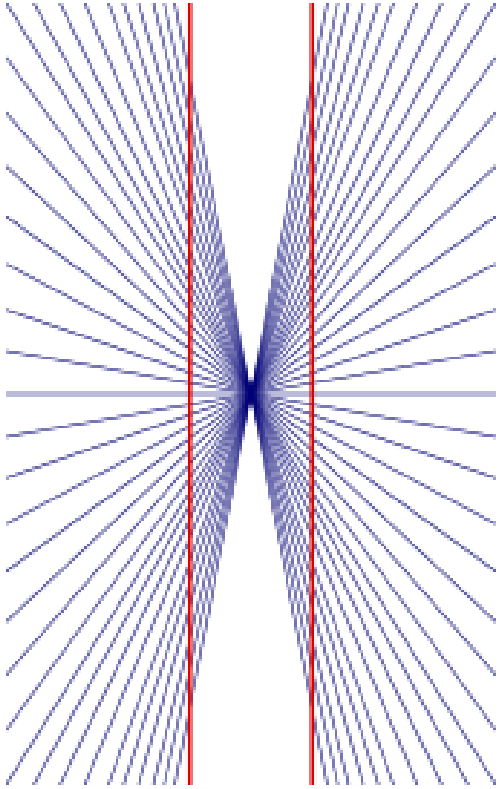


Vertical Horizontal illision

Ảo ảnh mù độ cong
(the illumination "Curvature Blindness" illusion)



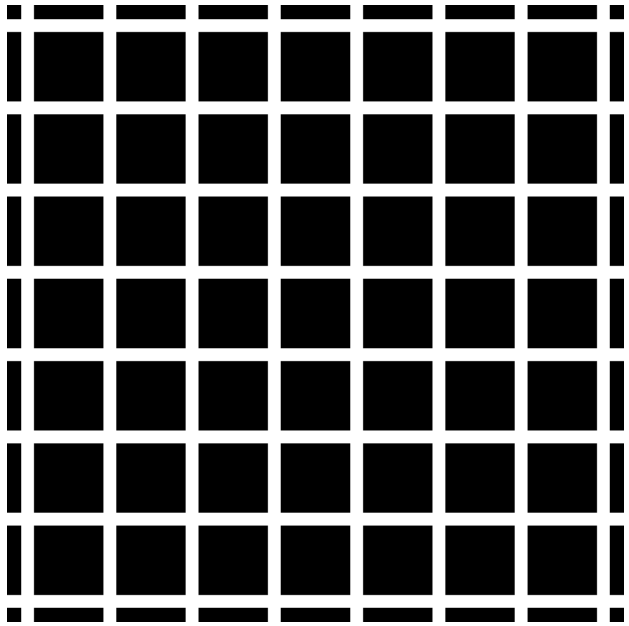
Ảo ảnh nhận thức tương lai (Future perception)



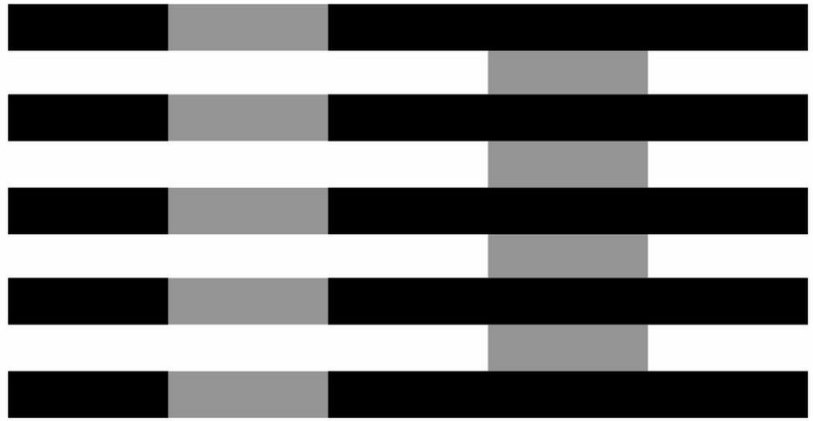
Hering illusion

Ảo ảnh bất biến mầu sắc – độ sáng (color and brightness constancies)

Simultaneous Contrast Illusion



Hermann Grid



While’s Effect