4

Perceiving and Recognizing Objects

Chapter 4 Perceiving and Recognizing Objects

- *What* and *Where* Pathways
- The Problems of Perceiving and Recognizing Objects
- Middle Vision
- Object Recognition

Introduction

How do we recognize objects?

- Retinal ganglion cells and LGN = Spots
- Primary visual cortex = Bars

How do spots and bars get turned into objects and surfaces?

- Clearly our brains are doing something pretty sophisticated beyond V1.

What and *Where* Pathways

Objects in the brain

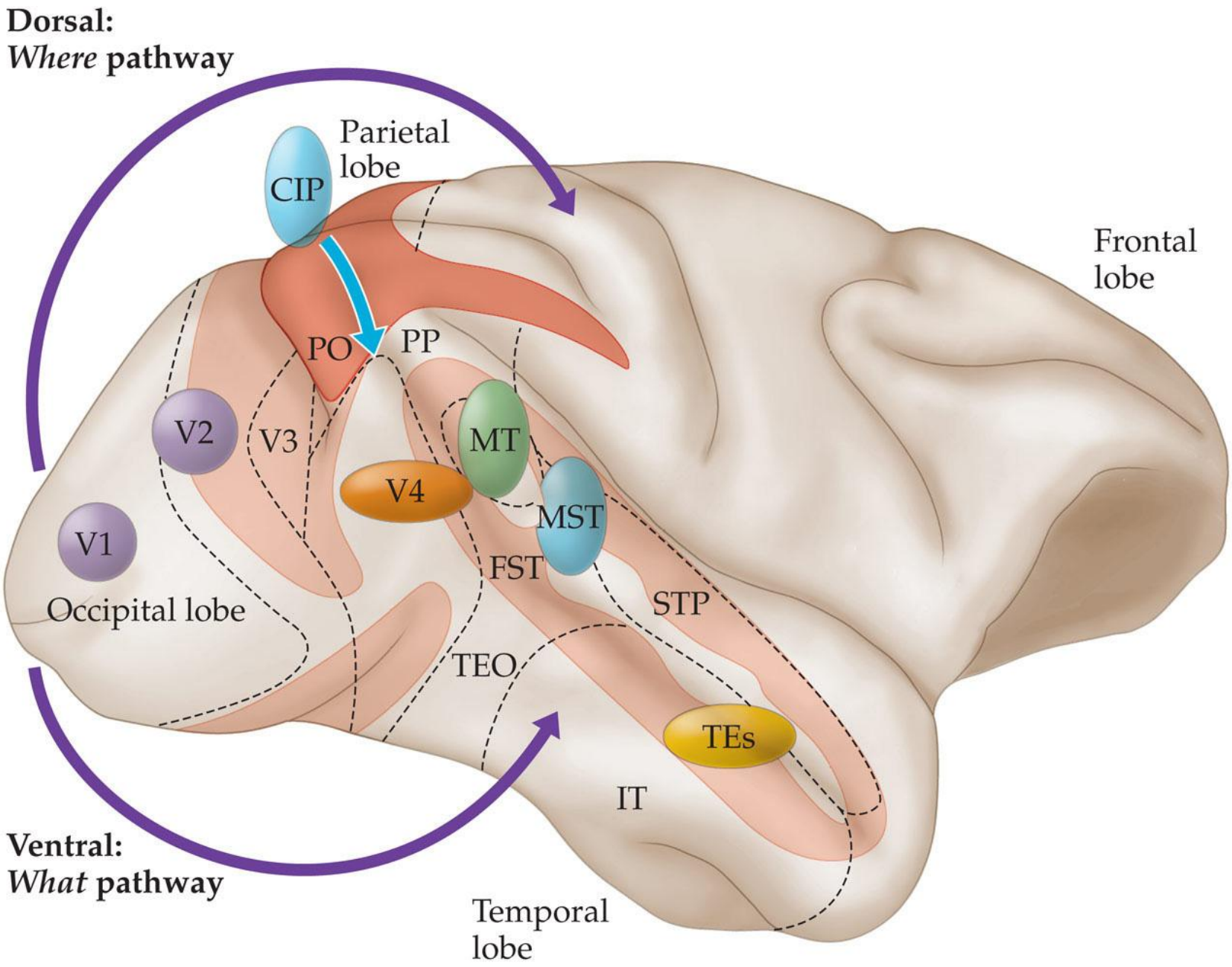
- Extrastriate cortex: Brain regions bordering primary visual cortex that contains other areas involved in visual processing.
  - V2, V3, V4, etc.

What and *Where* Pathways

Objects in the brain (*continued*)

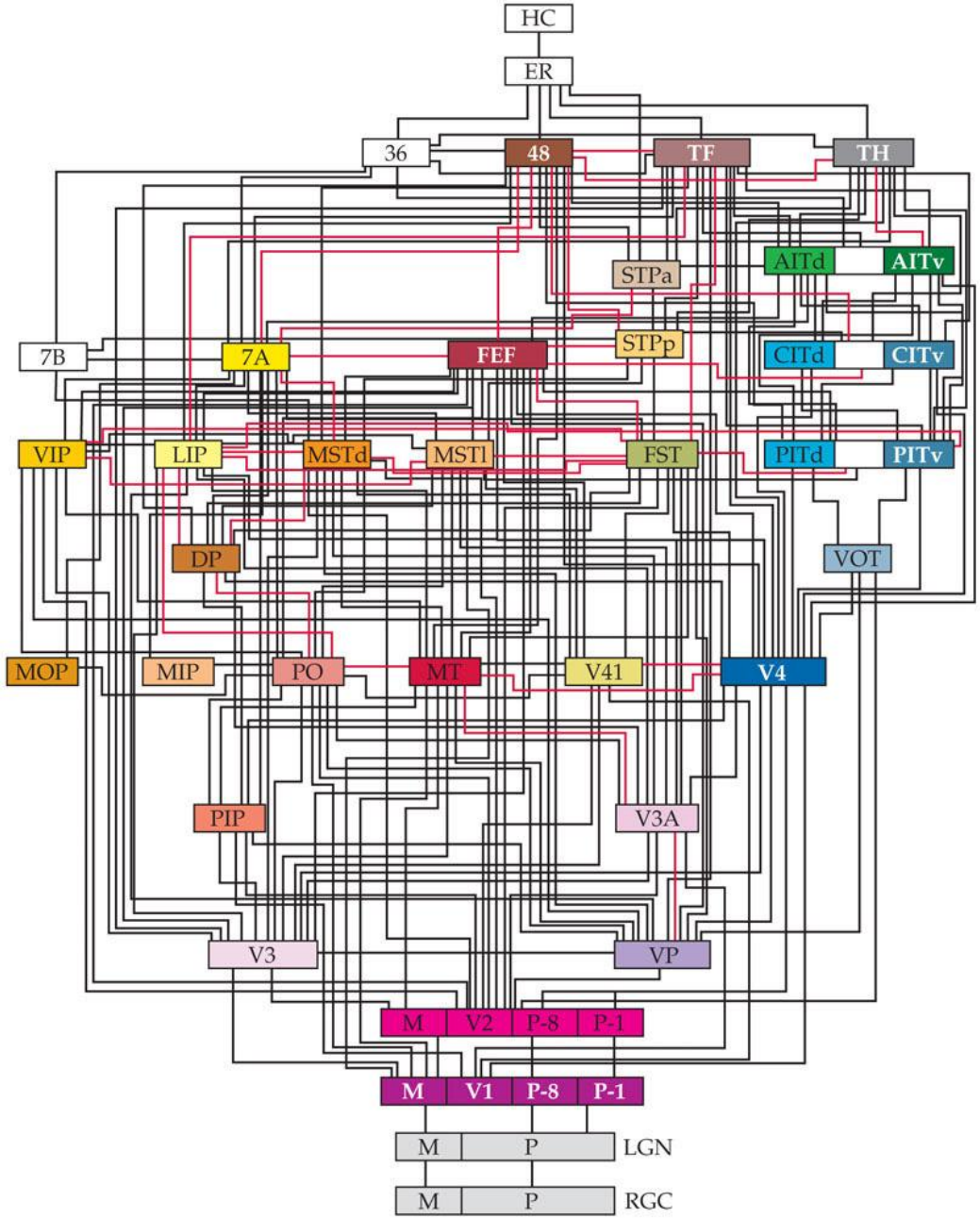
- After extrastriate cortex, processing of object information is split into a “what” pathway and a “where” pathway.
  - “Where” pathway is concerned with the locations and shapes of objects but not their names or functions.
  - “What” pathway is concerned with the names and functions of objects regardless of location.

Figure 4.2 The main visual areas of the macaque monkey cortex



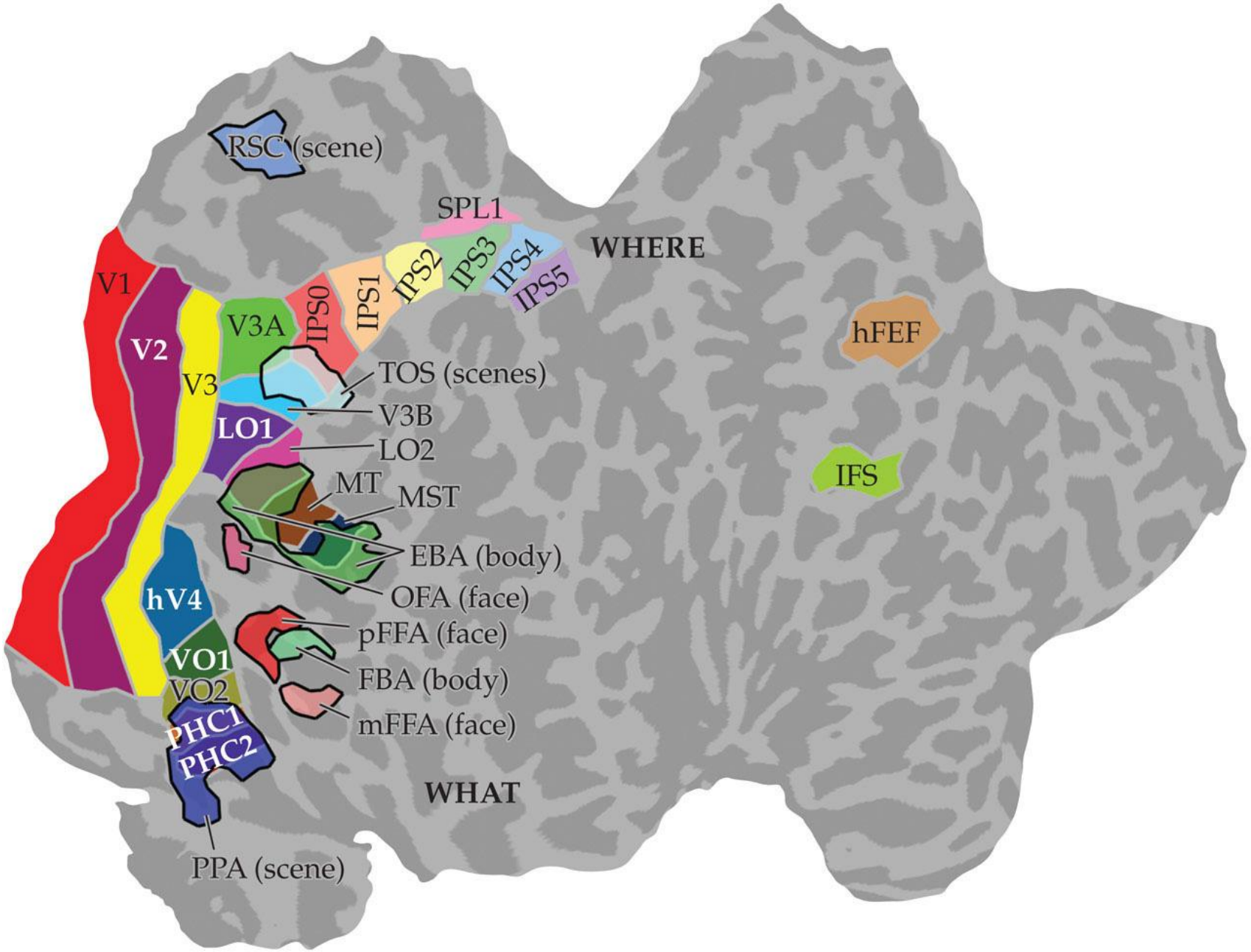
***SENSATION & PERCEPTION 4e*, Figure 4.2**

Figure 4.3 A partial “wiring” diagram for the main visual areas of the brain



***SENSATION & PERCEPTION 4e*, Figure 4.3**

Figure 4.4 The main visual areas of the human cortex



***SENSATION & PERCEPTION 4e*, Figure 4.4**

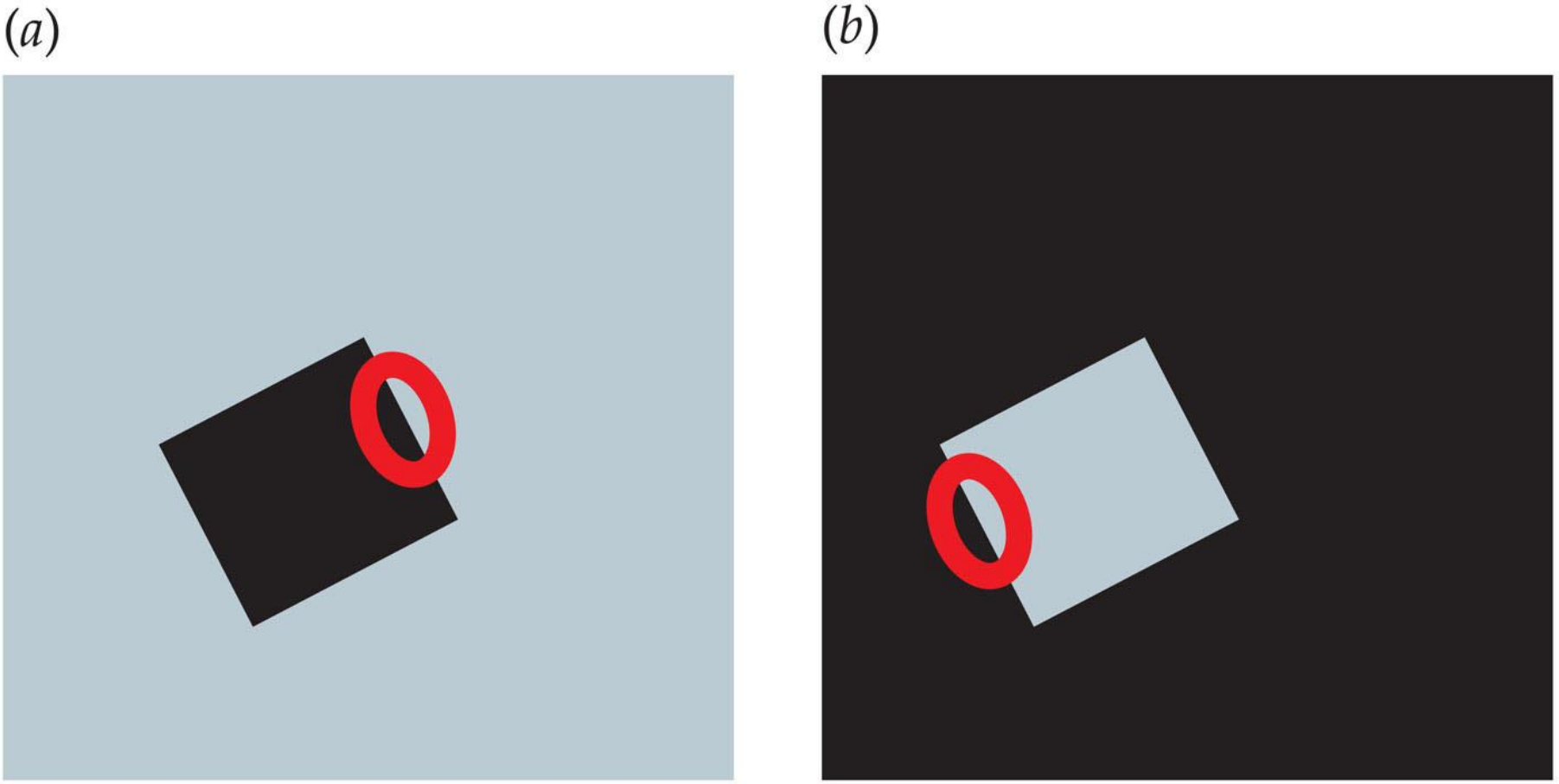
What and *Where* Pathways

The receptive fields of extrastriate cells are more sophisticated than those in striate cortex.

They respond to visual properties important for perceiving objects.

- For instance, “boundary ownership.” For a given boundary, which side is part of the object and which side is part of the background?

Figure 4.5 Notice that the receptive field sees the same dark edge in both images



***SENSATION & PERCEPTION 4e*, Figure 4.5**

What and *Where* Pathways

The same visual input occurs in both (*a*) and (*b*) and a V1 neuron would respond equally to both.

A V2 neuron might respond more to (*a*) than (*b*) because the black edge is owned by the square in (*a*) but not in (*b*).

What and *Where* Pathways

Inferotemporal (IT) cortex: Part of the cerebral cortex in the lower portion of the temporal lobe, important for object recognition.

- Part of the “what” pathway

Lesion, in neuropsychology:

1.(*n*.) A region of damaged brain.

2.(*v*.) To destroy a section of the brain.

What and *Where* Pathways

When IT cortex is lesioned, it leads to agnosias.

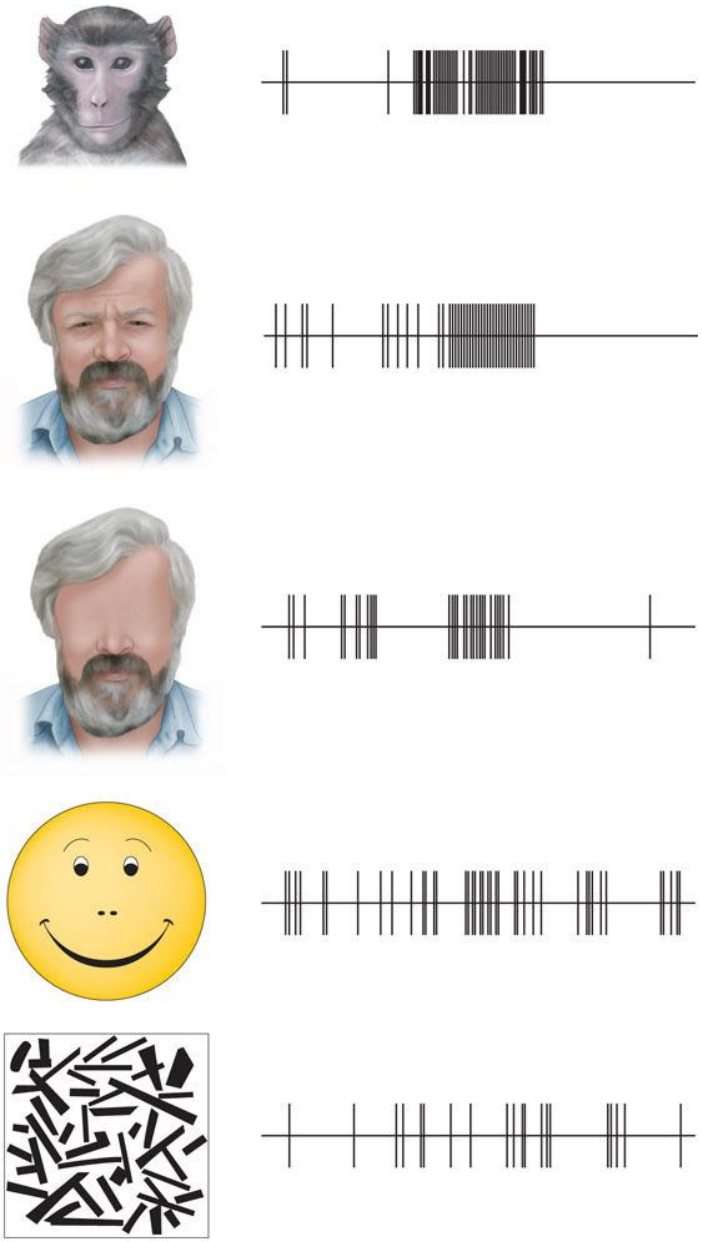
- Agnosia: Failure to recognize objects in spite of the ability to see them.

What and *Where* Pathways

Receptive field properties of IT neurons

- Very large—some cover half the visual field
- Don’t respond well to spots or lines
- Do respond well to stimuli such as hands, faces, or objects

Figure 4.6 Cells in the inferotemporal cortex of macaque monkeys are interested in very specific stimuli



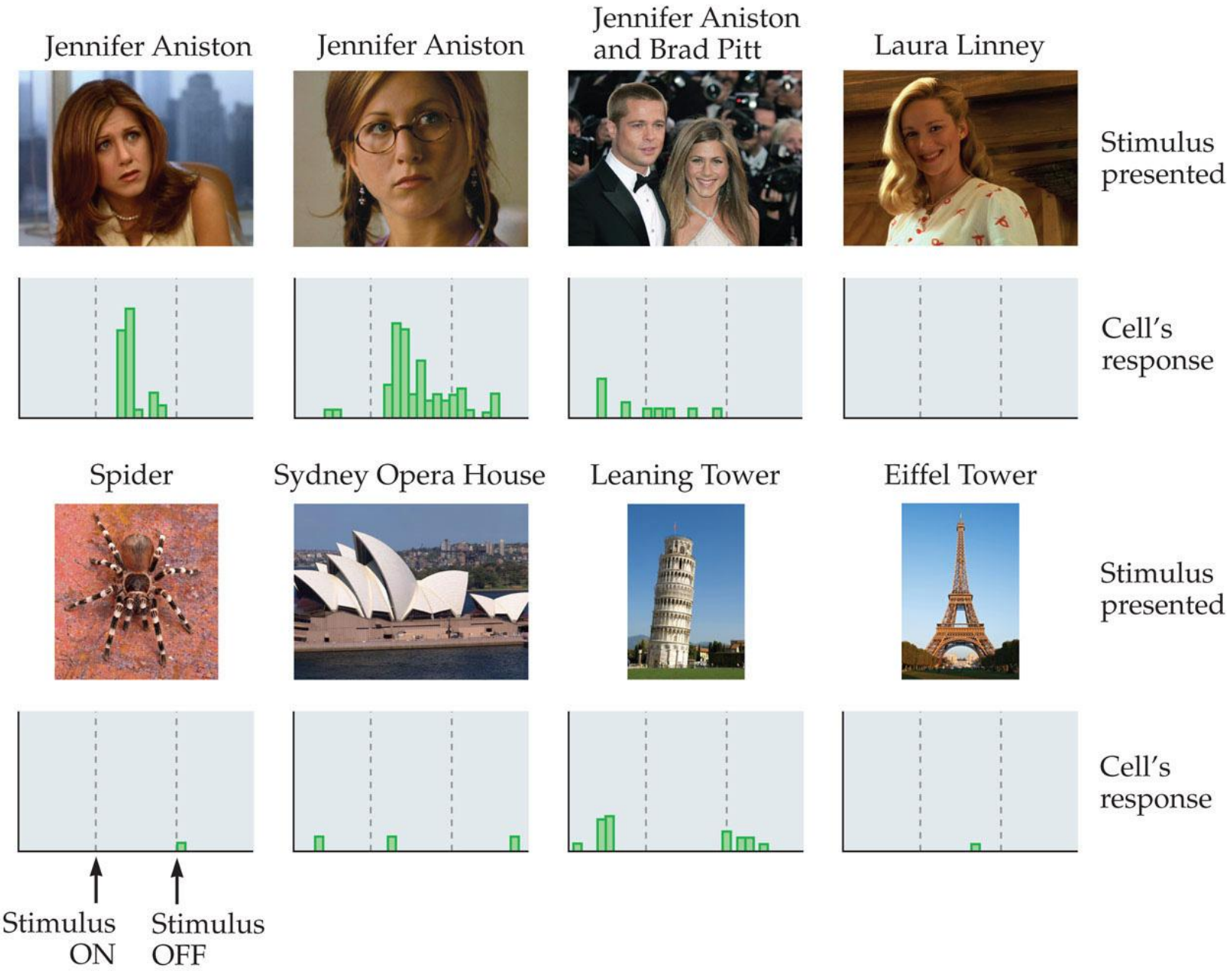
***SENSATION & PERCEPTION 4e*, Figure 4.6**

What and *Where* Pathways

Grandmother cells

- Could a single neuron be responsible for recognizing your grandmother?
- Perhaps!
- Quiroga et al. (2005) identified a cell that responds specifically to Jennifer Aniston.

Figure 4.7 Results of recording the activity of one cell in the temporal lobe of a human patient



***SENSATION & PERCEPTION 4e*, Figure 4.7**

What and *Where* Pathways

Object recognition is fast!

- Studies indicate that object recognition happens in as little as 150 ms.
- This is such a short time that there cannot be a lot of feedback from higher brain areas.
  - Feed-forward process: A process that carries out a computation (e.g., object recognition) one neural step after another, without the need for feedback from a later stage to an earlier stage.

Flexibility of Object Recognition: Recognize the Flowers in spite of variations, recognize the sketches in spite of paucity of information.

A

B

The Problems of Perceiving and Recognizing Objects

What do you see?

(a)

The Problems of Perceiving and Recognizing Objects

What do you see?

(b)

The Problems of Perceiving and Recognizing Objects

What do you see?

(c)

The Problems of Perceiving and Recognizing Objects

Is this the same red house that was shown before?

(d)

The Problems of Perceiving and Recognizing Objects

The problem of object recognition

- The pictures were just a bunch of pixels on a screen, but in each case you perceived a *house.*
- How did you recognize all four images as depicting a house?

The Problems of Perceiving and Recognizing Objects

The problem of object recognition (*continued*)

- How did you recognize the first and third images as depicting the same house, but from different viewpoints?
- How does your visual system move from points of light, like pixels, to whole entities in the world, like houses?

Middle Vision

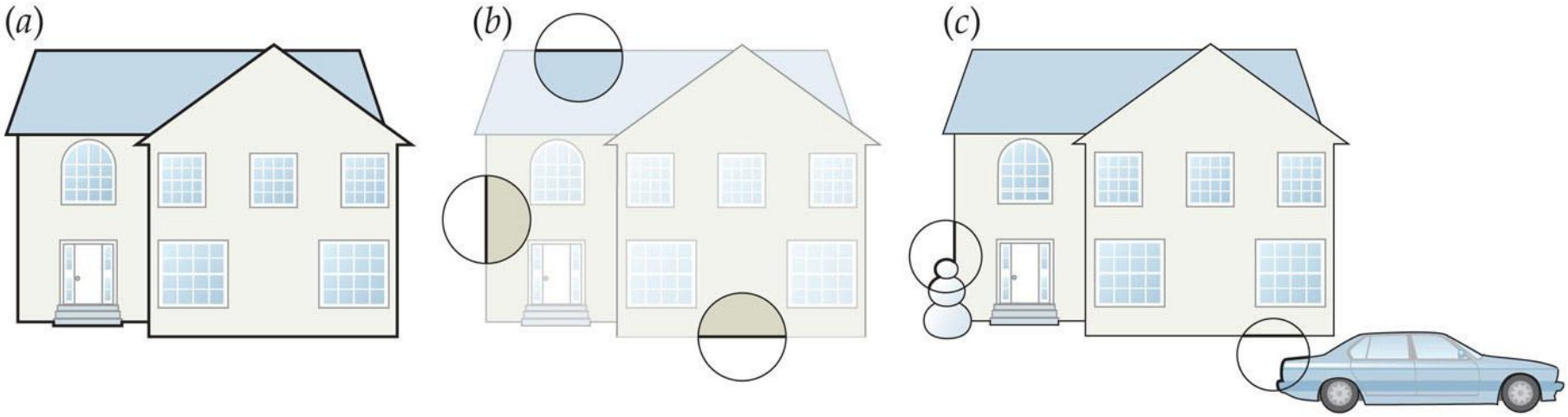
Middle vision: A loosely defined stage of visual processing that comes after basic features have been extracted from the image (low-level vision) and before object recognition and scene understanding (high-level vision).

- Involves the perception of edges and surfaces
- Determines which regions of an image should be grouped together into objects

Finding edges

- How do you find the edges of objects?
- Cells in primary visual cortex have small receptive fields.
- How do you know which edges go together and which ones don’t?

Figure 4.9 The problem continued



***SENSATION & PERCEPTION 4e*, Figure 4.9**

Middle Vision

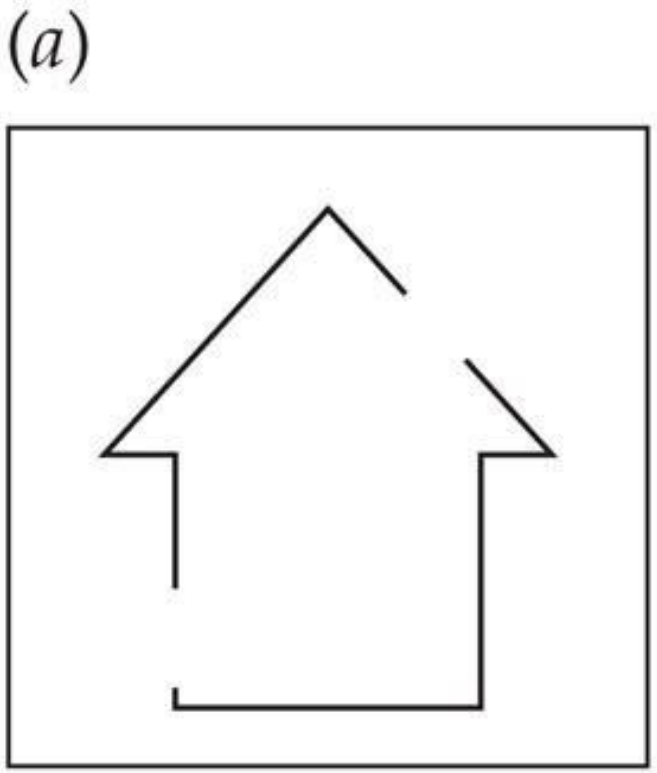
Computer-based edge detectors are not as good as humans.

- Sometimes computers don’t find edges that humans see easily.

Figure 4.10 In some places this object is darker than the background; in other places it is lighter

***SENSATION & PERCEPTION 4e*, Figure 4.10**

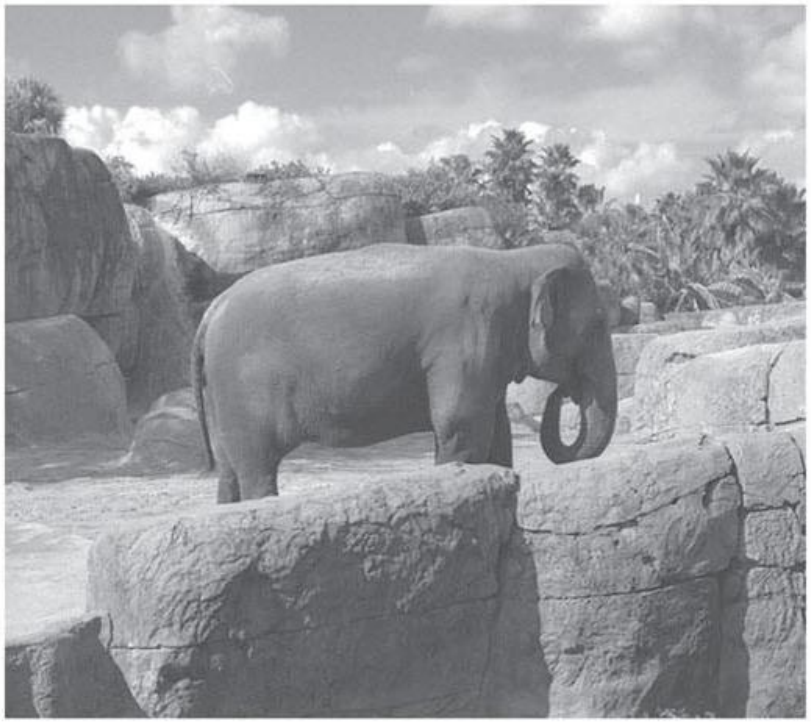
Figure 4.11 (*a*) The “find edges” function in a graphics program finds gaps in the borders of an image (Part 1)



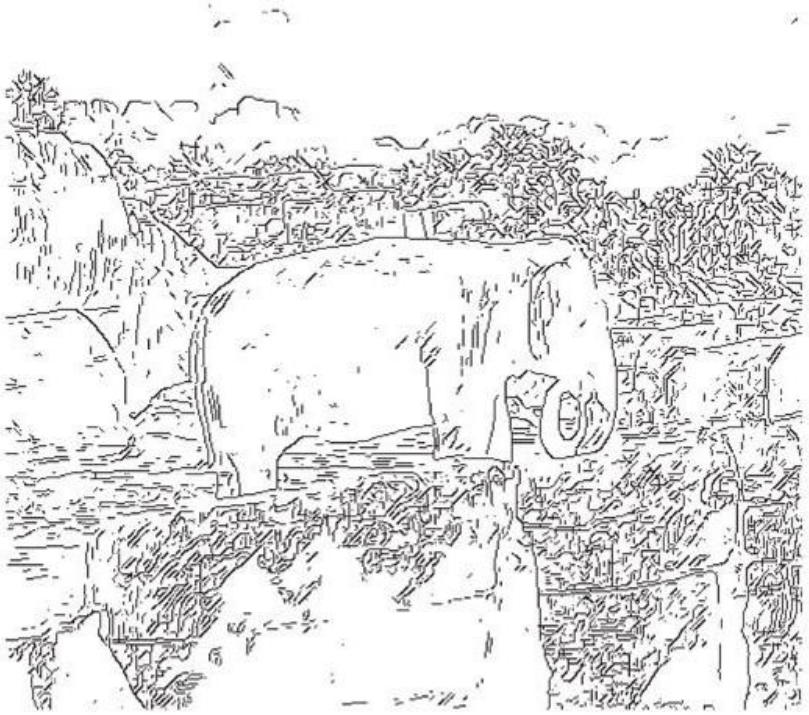
Middle Vision

- Sometimes computers find too many edges.

(b)



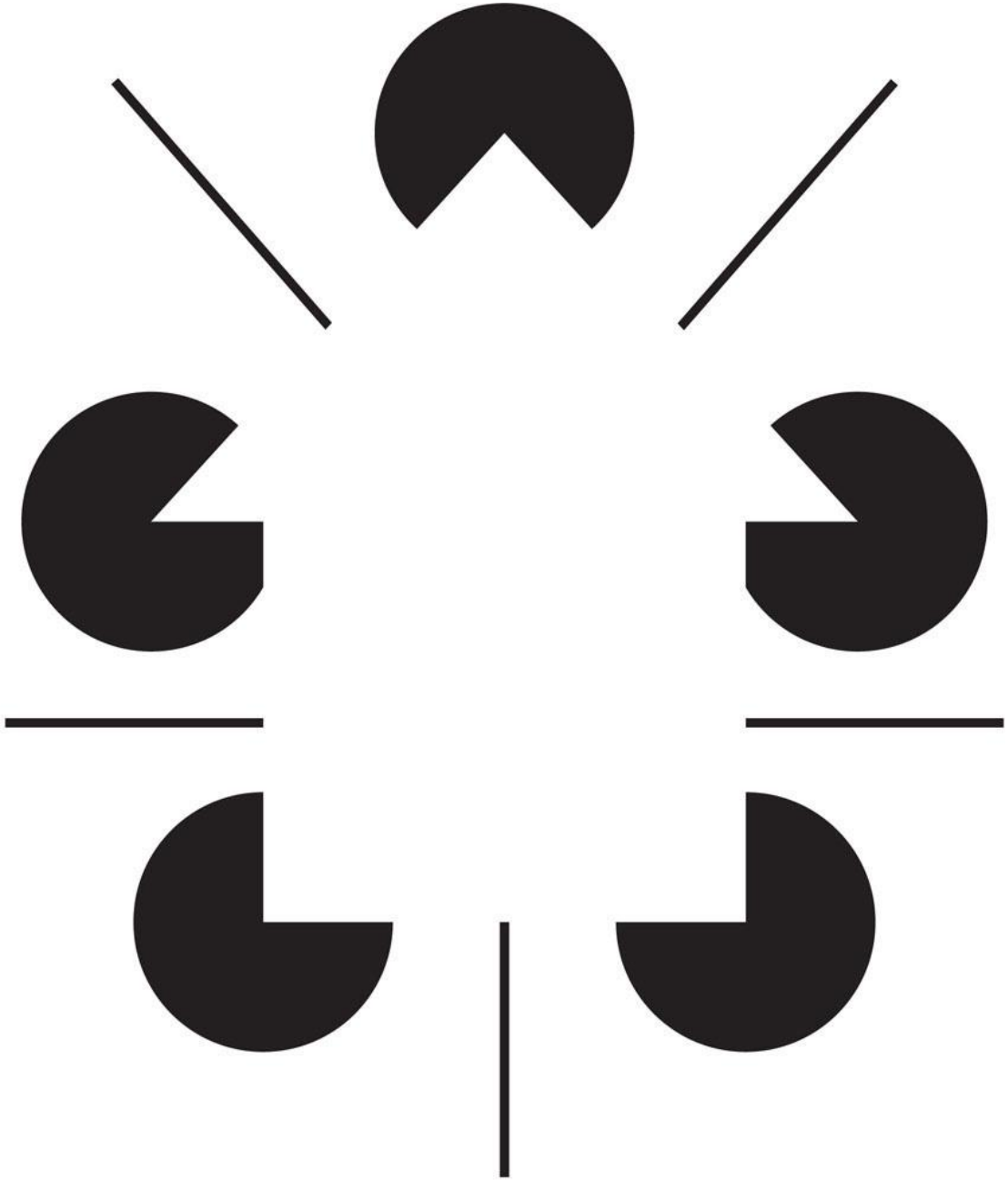
(c)



Middle Vision

Illusory contour: A contour that is perceived even though nothing changes from one side of the contour to the other.

Figure 4.12 This “house” outline is constructed from illusory contours



***SENSATION & PERCEPTION 4e*, Figure 4.12**

Middle Vision

Do you see a white house sitting on top of some circles?

There is no house! Just some “Pac-Men” and disconnected lines.

Middle Vision

Rules that make contours

- Gestalt: In German, “form” or “whole.”
- Gestalt psychology: “The whole is greater than the sum of its parts.”
  - Opposed to other schools of thought, such as structuralism, that emphasize the basic elements of perception
- Gestalt grouping rules: A set of rules that describe when elements in an image will appear to group together.

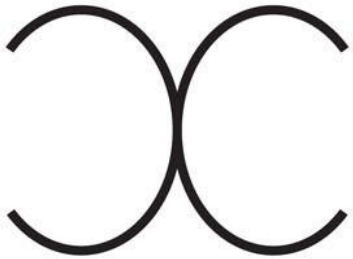
Rules that make contours (*continued*)

- Good continuation: A Gestalt grouping rule stating that two elements will tend to group together if they lie on the same contour.

Figure 4.14 The Gestalt principles of good continuation and closure

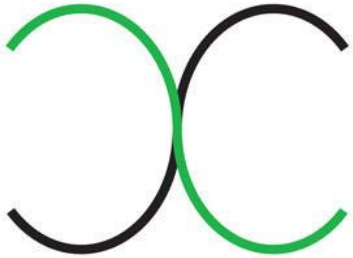
(a)

This…



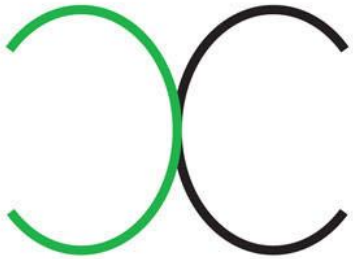
(b)

…looks more like this…



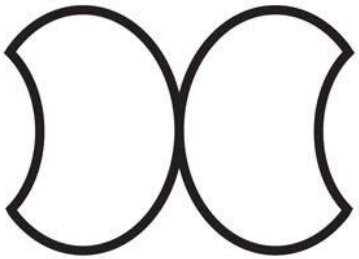
(c)

…not like this.



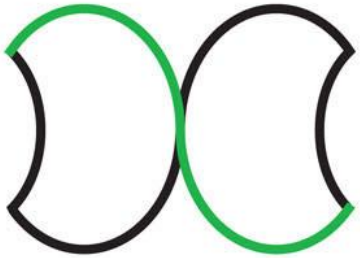
(d)

But this…



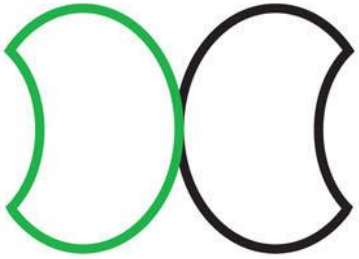
(e)

…looks *less* like this…



(f)

…than like this.



(g)

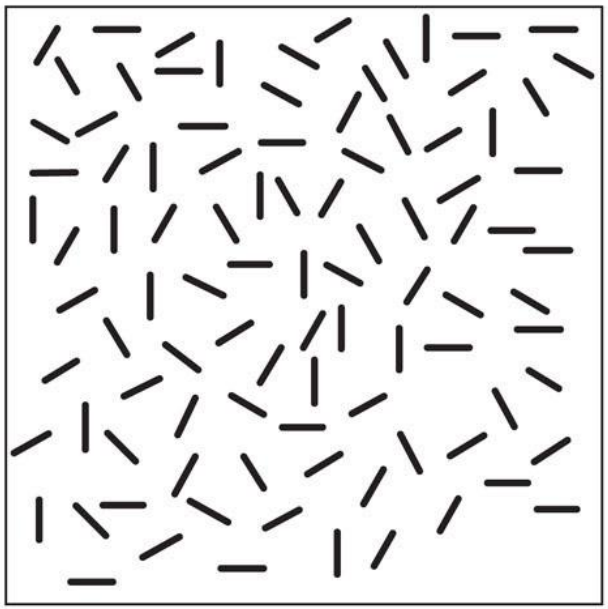
***SENSATION & PERCEPTION 4e*, Figure 4.14**

Rules that make contours (*continued*)

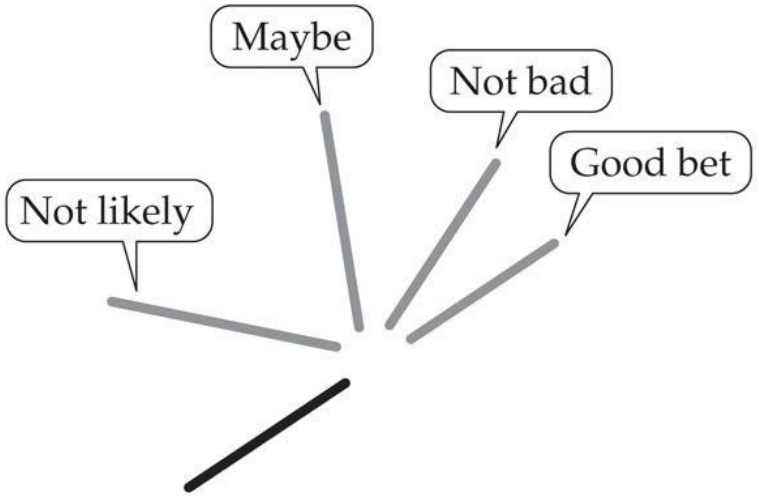
- Some contours in an image will group because of good continuation.
- Can you find the shape embedded within the field of lines in (*a*) on the next slide?

Figure 4.13 Contour completion

(a)



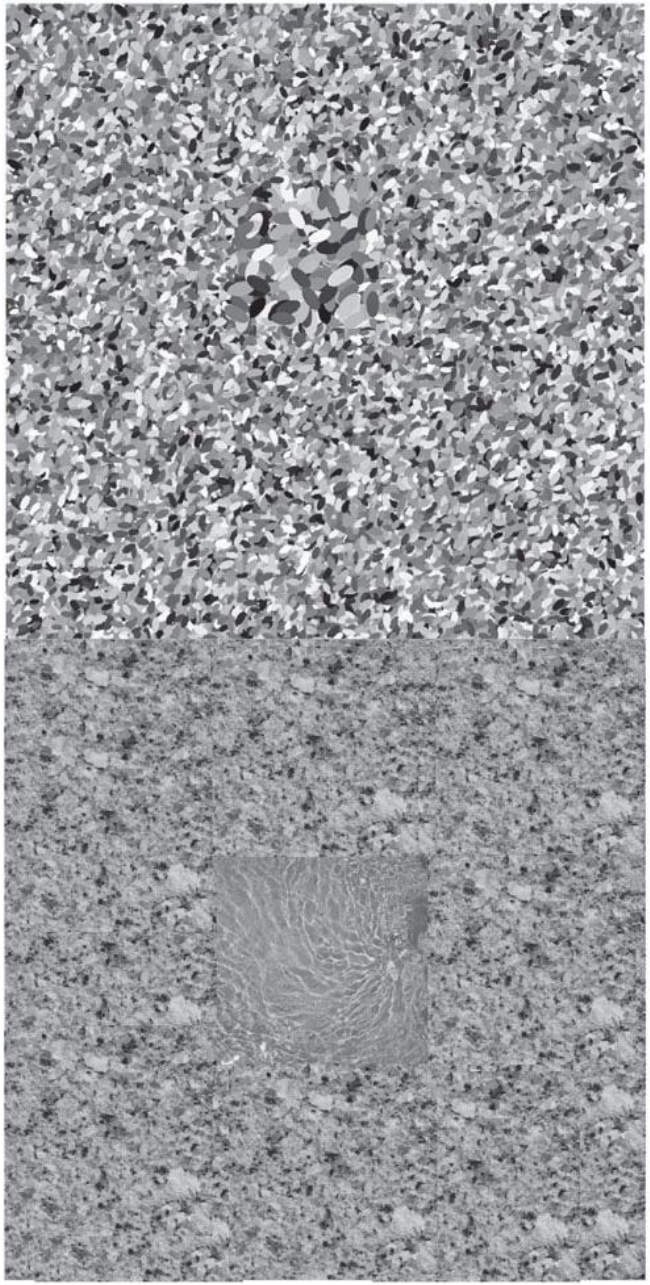
(b) Which gray line is a likely continuation of the black line?



Texture segmentation and grouping

- Texture segmentation: Carving an image into regions of common texture properties.
- Texture grouping depends on the statistics of textures in one region versus another.

Figure 4.16 The problem of texture segmentation



***SENSATION & PERCEPTION 4e*, Figure 4.16**

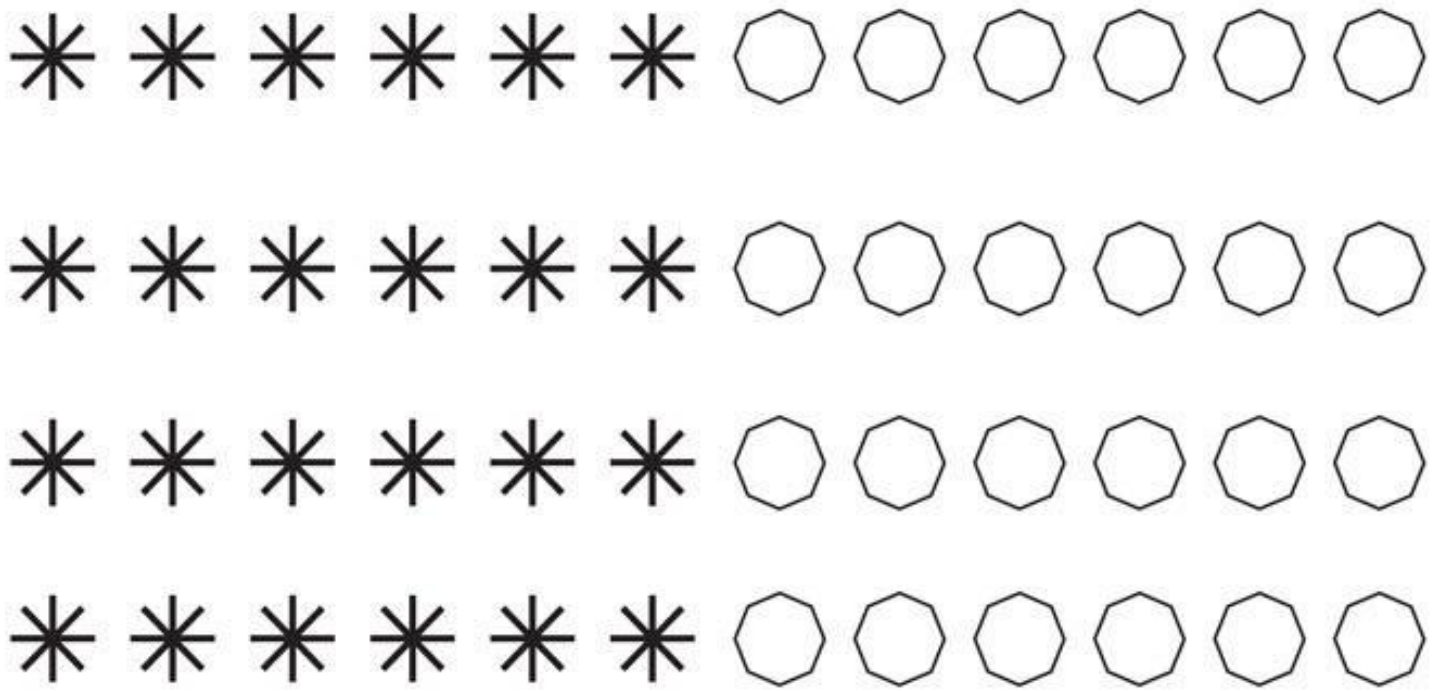
Middle Vision

Gestalt grouping rules

- Similarity: Similar looking items tend to group.
- Proximity: Items that are near each other tend to group.

Figure 4.18 Similarity and proximity

(a)



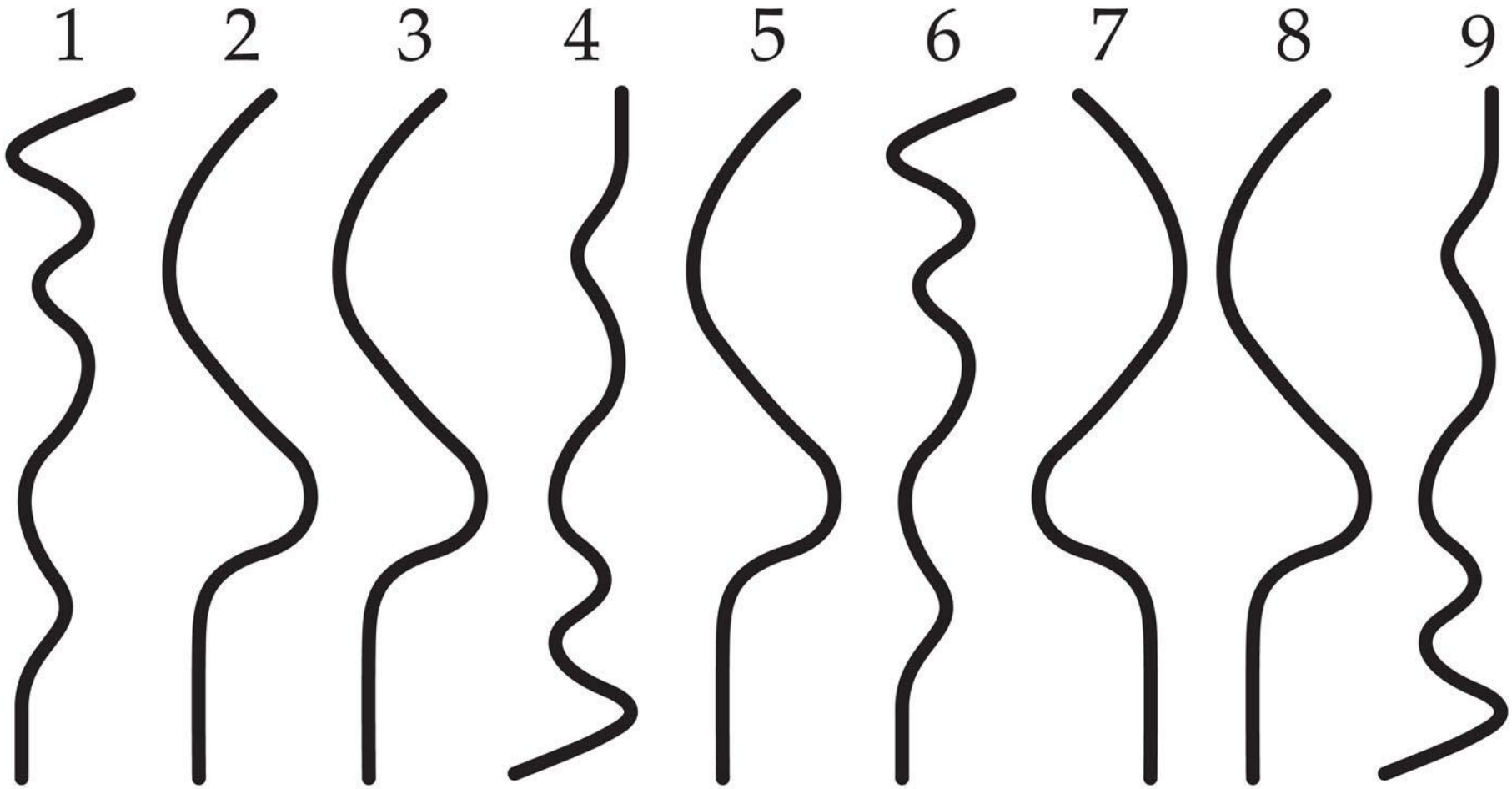
***SENSATION & PERCEPTION 4e*, Figure 4.18**

Middle Vision

Texture segmentation and grouping

- Parallelism: Parallel contours are likely to belong to the same group.
- Symmetry: Symmetrical regions are more likely to be seen as a group.

Figure 4.19 Parallelism and symmetry

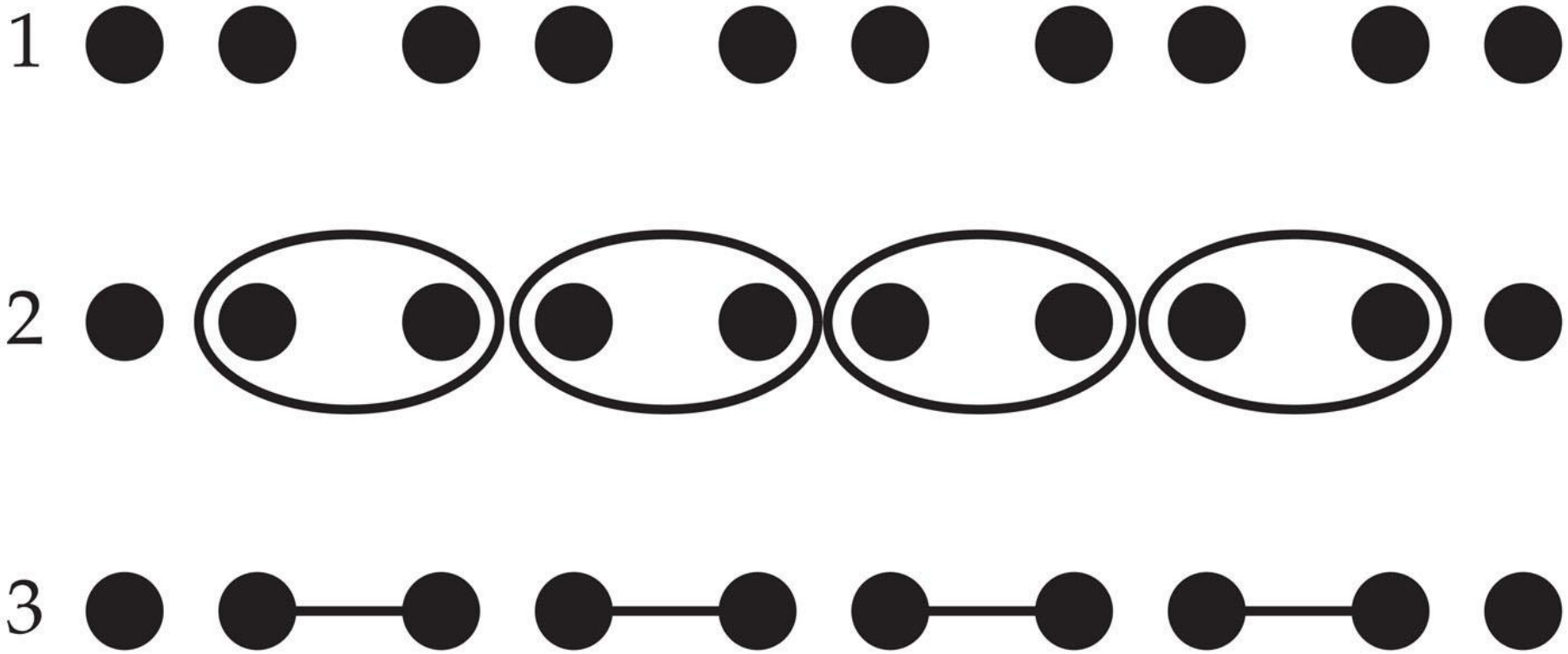


***SENSATION & PERCEPTION 4e*, Figure 4.19**

Texture segmentation and grouping (*continued*)

- Common region: Items will group if they appear to be part of the same larger region.
- Connectedness: Items will tend to group if they are connected.

Figure 4.20 Proximity grouping (line 1) can be overruled by grouping by common region (line 2) or grouping by connectedness (line 3)



***SENSATION & PERCEPTION 4e*, Figure 4.20**

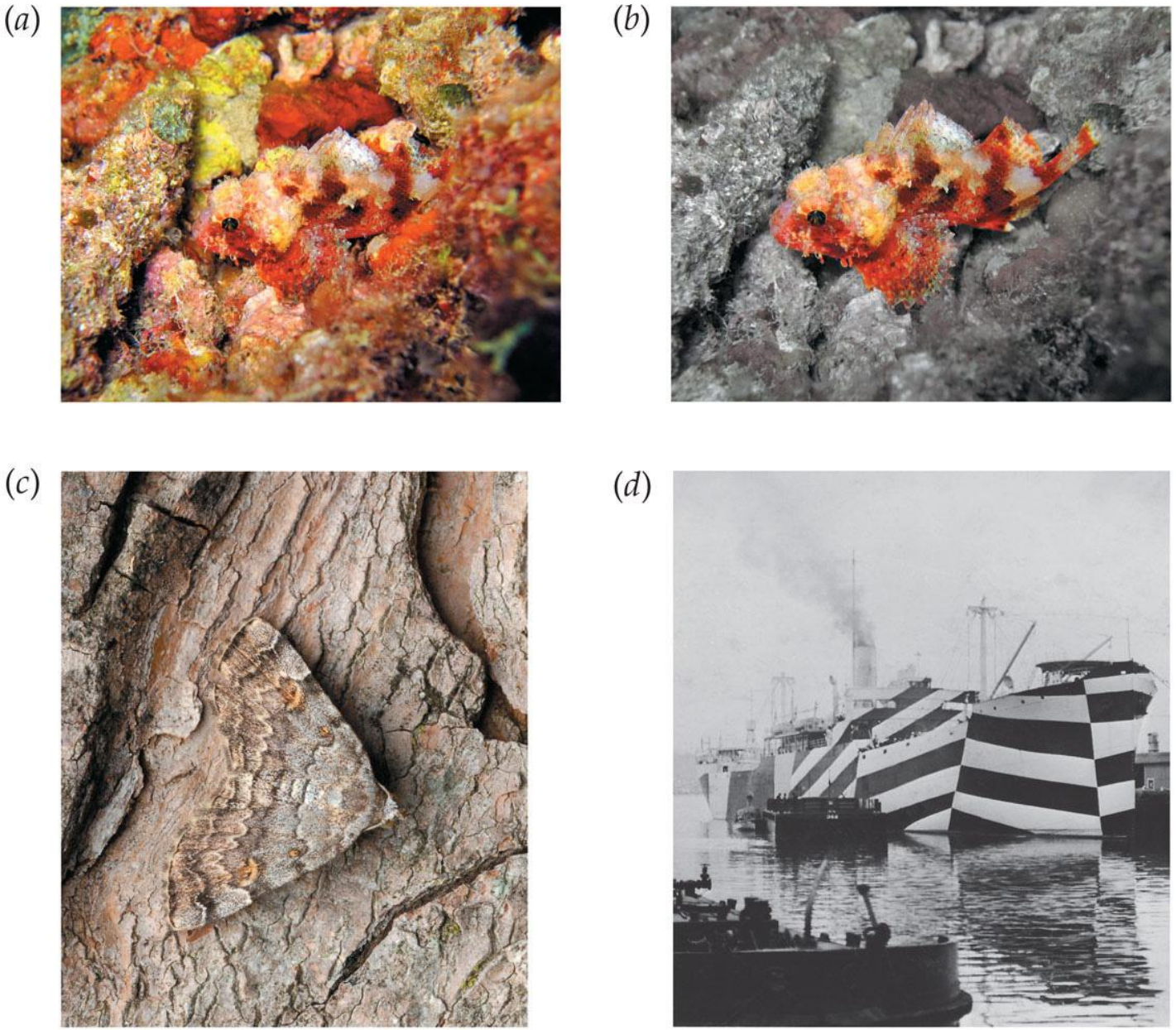
Texture segmentation and grouping (*continued*)

- Dynamic grouping principles
  - Common fate: Elements that move in the same direction tend to group together.
  - Synchrony: Elements that change at the same time tend to group together.

Camouflage

- Animals exploit Gestalt grouping principles to group into their surroundings.
- Sometimes camouflage is used to confuse the observer.

Figure 4.21 Panel (*b*) shows the hidden animal in (*a*). We leave (*c*) as an exercise for the reader. (*d*) Sometimes camouflage is meant to confuse the viewer, rather than to hide the object



***SENSATION & PERCEPTION 4e*, Figure 4.21**

Middle Vision

Perceptual committees revisited

- A metaphor for how perception works
  - Committees must integrate conflicting opinions and reach a consensus.
- Many different and sometimes competing principles are involved in perception.
- Perception results from the consensus that emerges.

The pandemonium model

- Oliver Selfridge's (1959) simple model of letter recognition
- Perceptual committee made up of "demons"
  - Demons loosely represent neurons.
  - Each level is like a different brain area.

Middle Vision

Committee rules: Honor physics and avoid accidents

- Ambiguous figure: A visual stimulus that gives rise to two or more interpretations of its identity or structure.
  - Perceptual committees tend to obey the laws of physics.

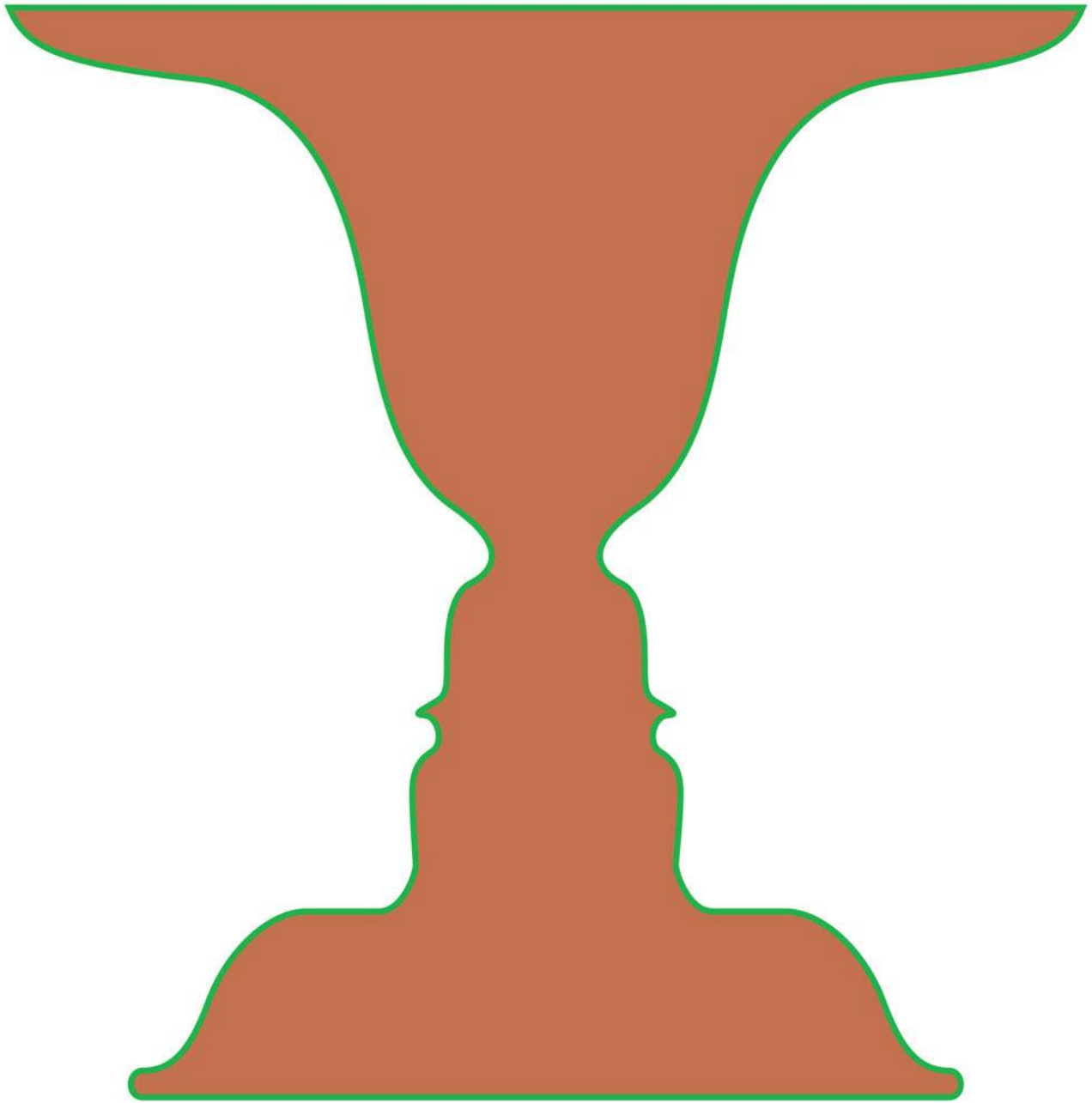
Committee rules: Honor physics and avoid accidents (*continued*)

- Accidental viewpoint: A viewing position that produces some regularity in the visual image that is not present in the world.
  - Perceptual committees assume viewpoints are not accidental.

Figure and ground

- Figure-ground assignment: The process of determining that some regions of an image belong to a foreground object (figure) and other regions are part of the background (ground).

Figure 4.27 The Rubin vase/face figure

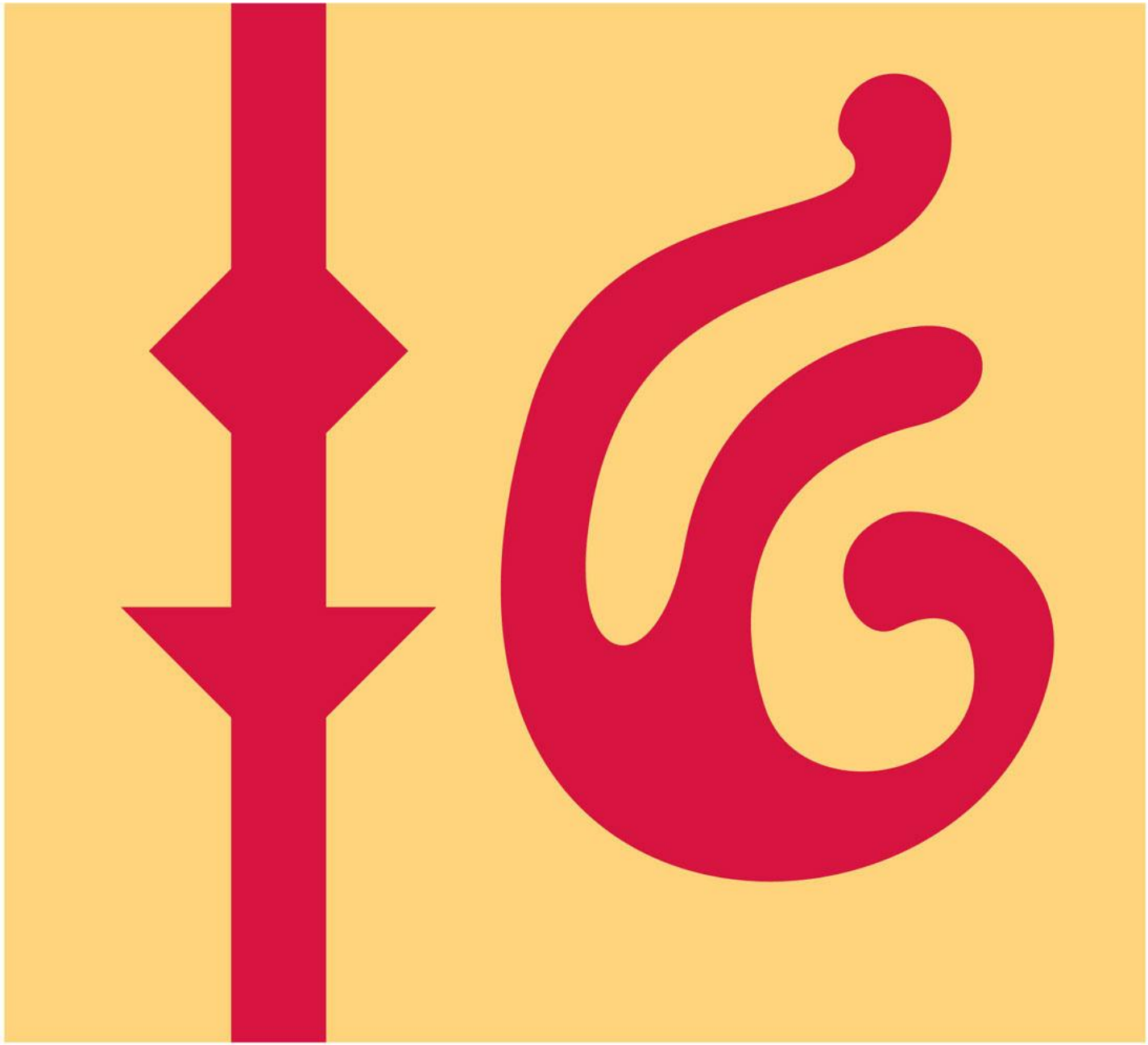


***SENSATION & PERCEPTION 4e*, Figure 4.27**

Gestalt figure-ground assignment principles

- Surroundedness: The surrounding region is likely to be ground.
- Size: The smaller region is likely to be figure.
- Symmetry: A symmetrical region tends to be seen as figure.
- Parallelism: Regions with parallel contours tend to be seen as figure.

Figure 4.26 What is figure and what is ground and why?

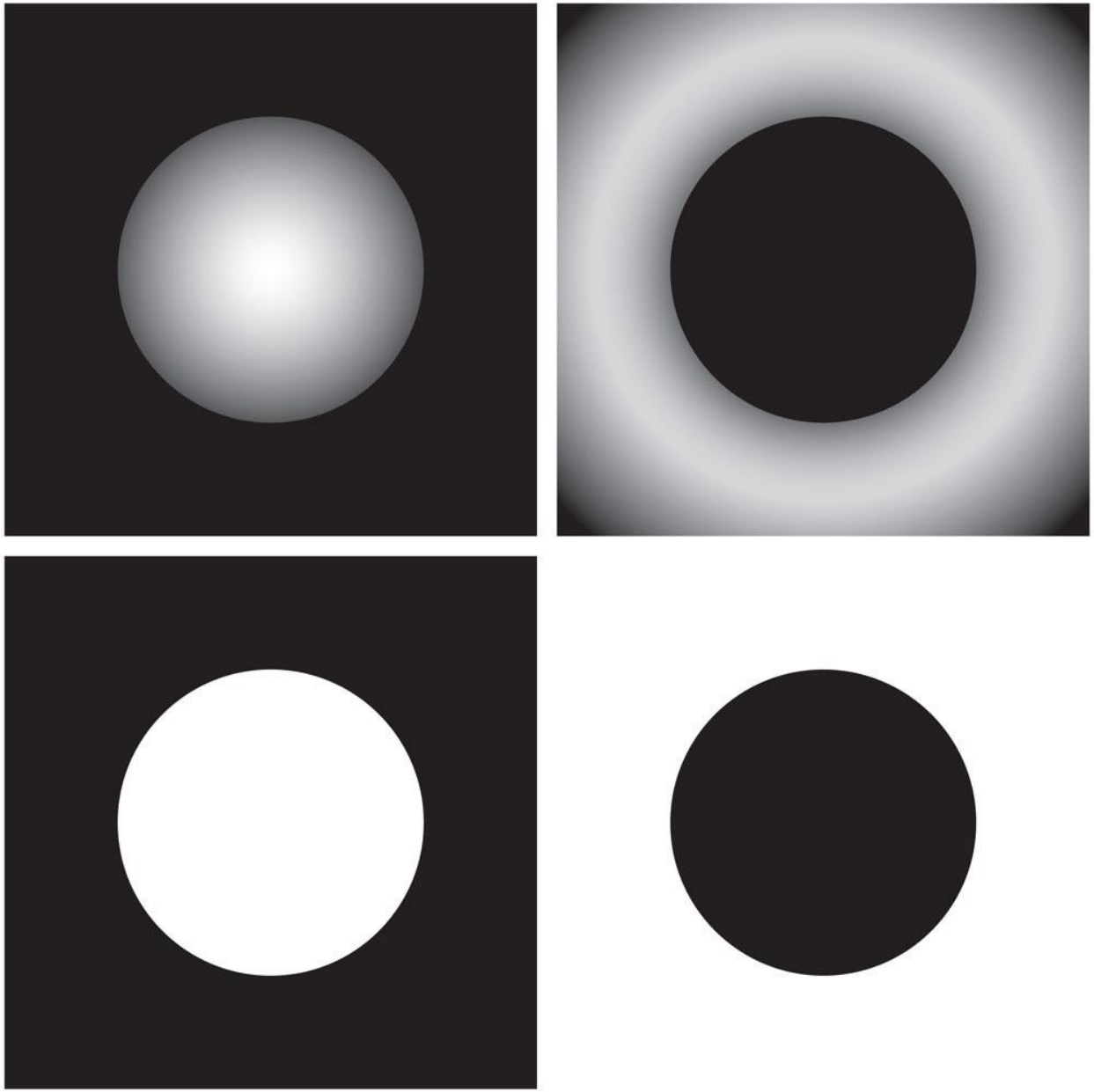


***SENSATION & PERCEPTION 4e*, Figure 4.26**

Gestalt figure-ground assignment principles (*continued*)

- Extremal edges: If edges of an object are shaded such that they seem to recede in the distance, they tend to be seen as figure.
- Relative motion: If one region moves in front of another, then the closer region is figure.

Figure 4.29 In the top left panel, the “extremal edge” cue makes the image look like a sphere on a background

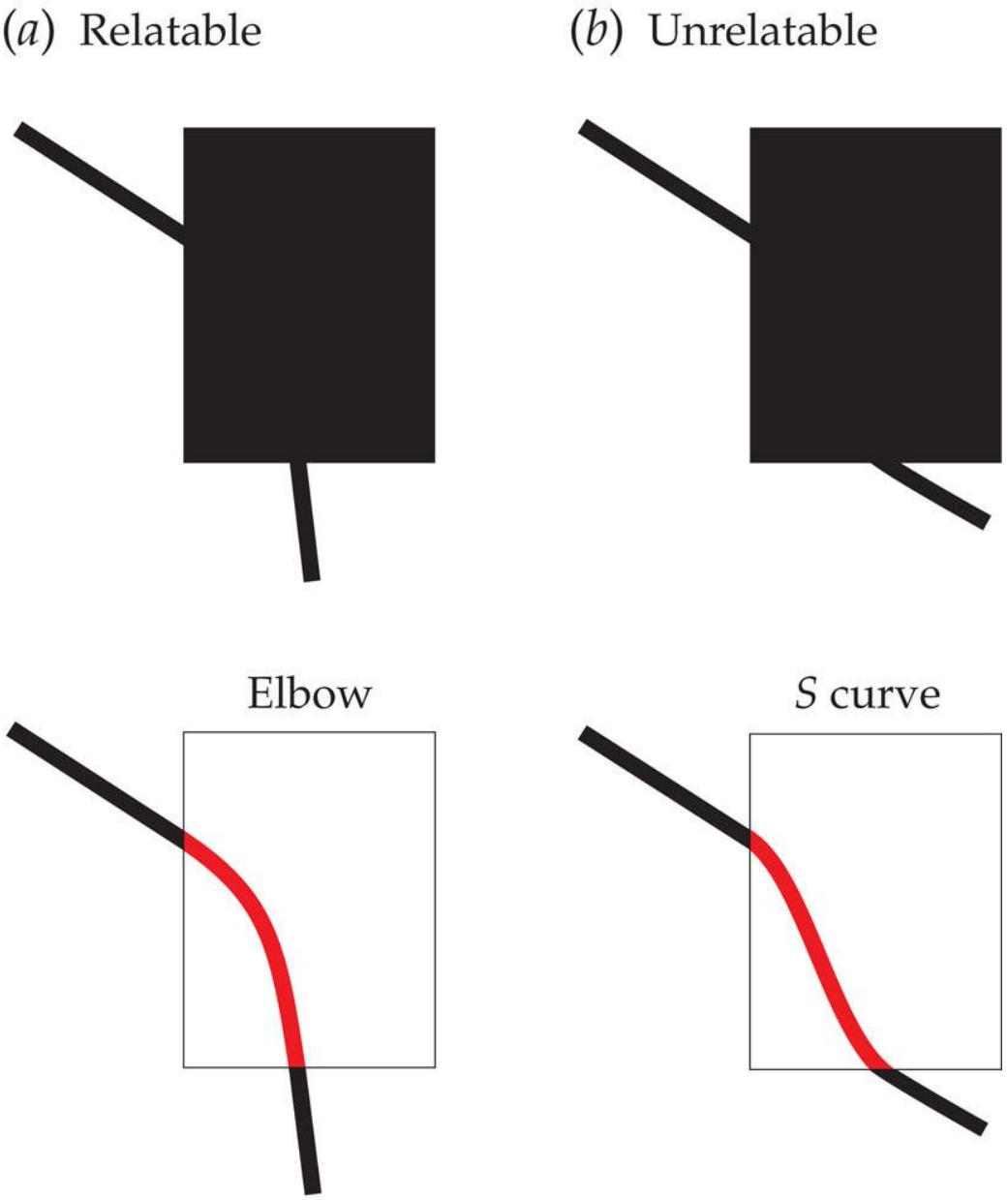


***SENSATION & PERCEPTION 4e*, Figure 4.29**

Dealing with occlusion

- Relatability: The degree to which two line segments appear to be part of the same contour.

Figure 4.31 Two edges are relatable if they can be connected with a smooth convex or smooth concave curve (*a*), but not if the connection requires an *S* curve (*b*)

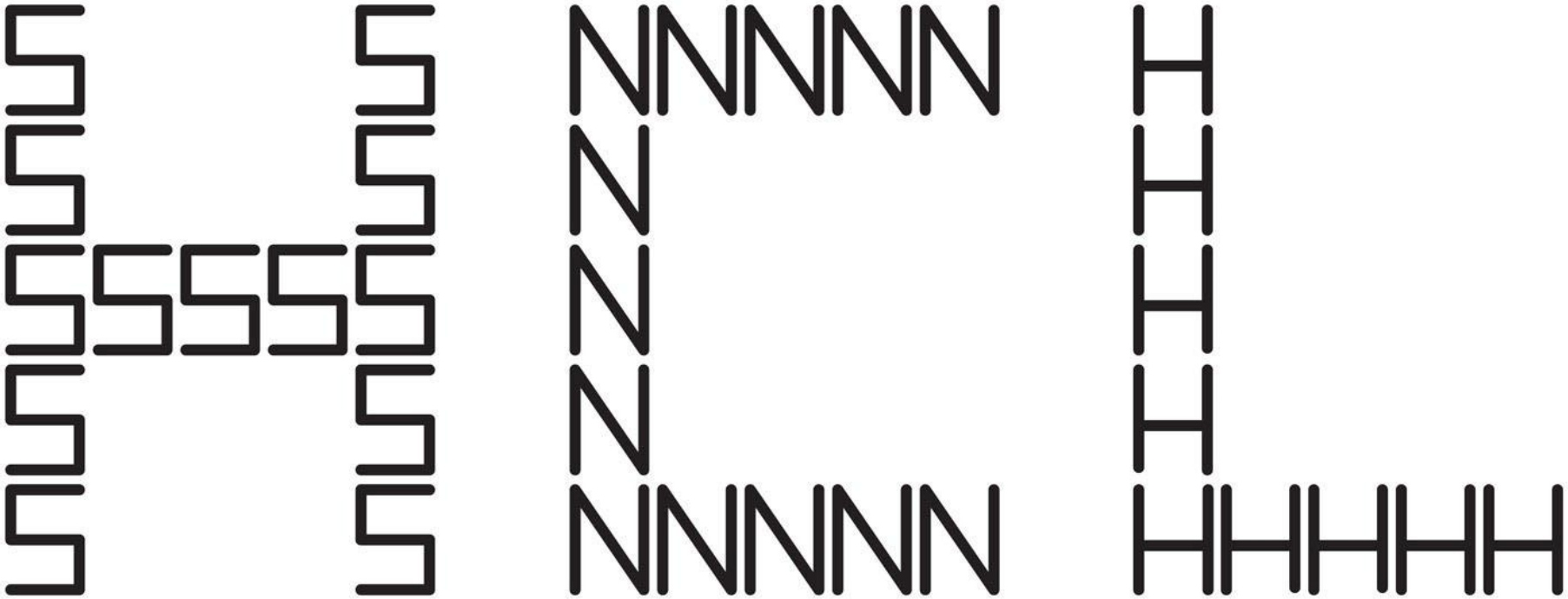


***SENSATION & PERCEPTION 4e*, Figure 4.31**

Parts and wholes

- Global superiority effect: The properties of the whole object take precedence over the properties of parts of the object.

Figure 4.33 Global letters composed of local letters



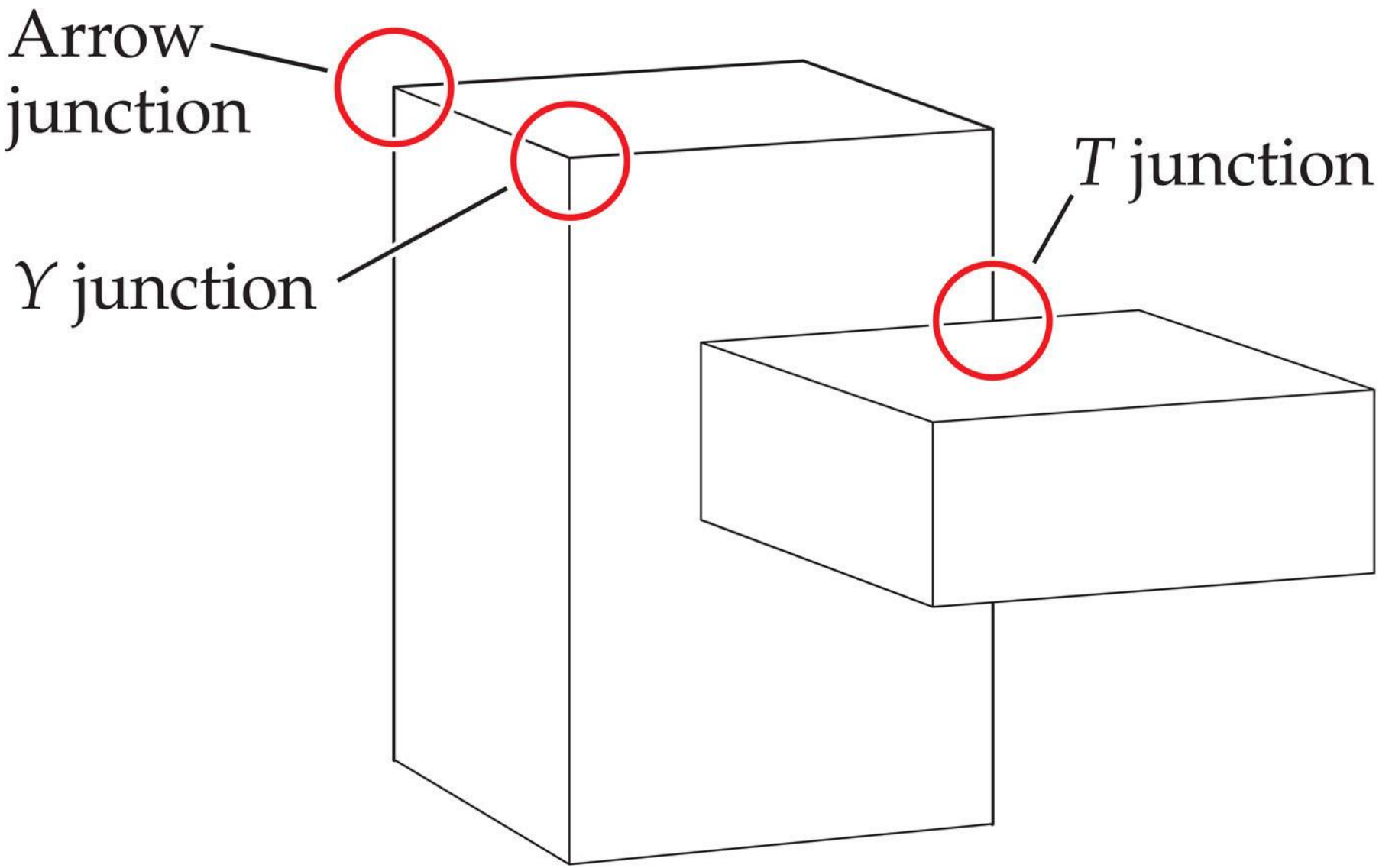
***SENSATION & PERCEPTION 4e,* Figure 4.33**

Middle Vision

Nonaccidental feature: A feature of an object that is not dependent on the exact (or accidental) viewing position of the observer.

- *T* junctions: Indicate occlusion. Top of *T* is in front and stem of *T* is in back.
- *Y* junctions: Indicate corners facing the observer.
- Arrow junctions: Indicate corners facing away from the observer.

Figure 4.32 Different sorts of line junctions have different meanings



***SENSATION & PERCEPTION 4e*, Figure 4.32**

Middle Vision

Summarizing middle vision

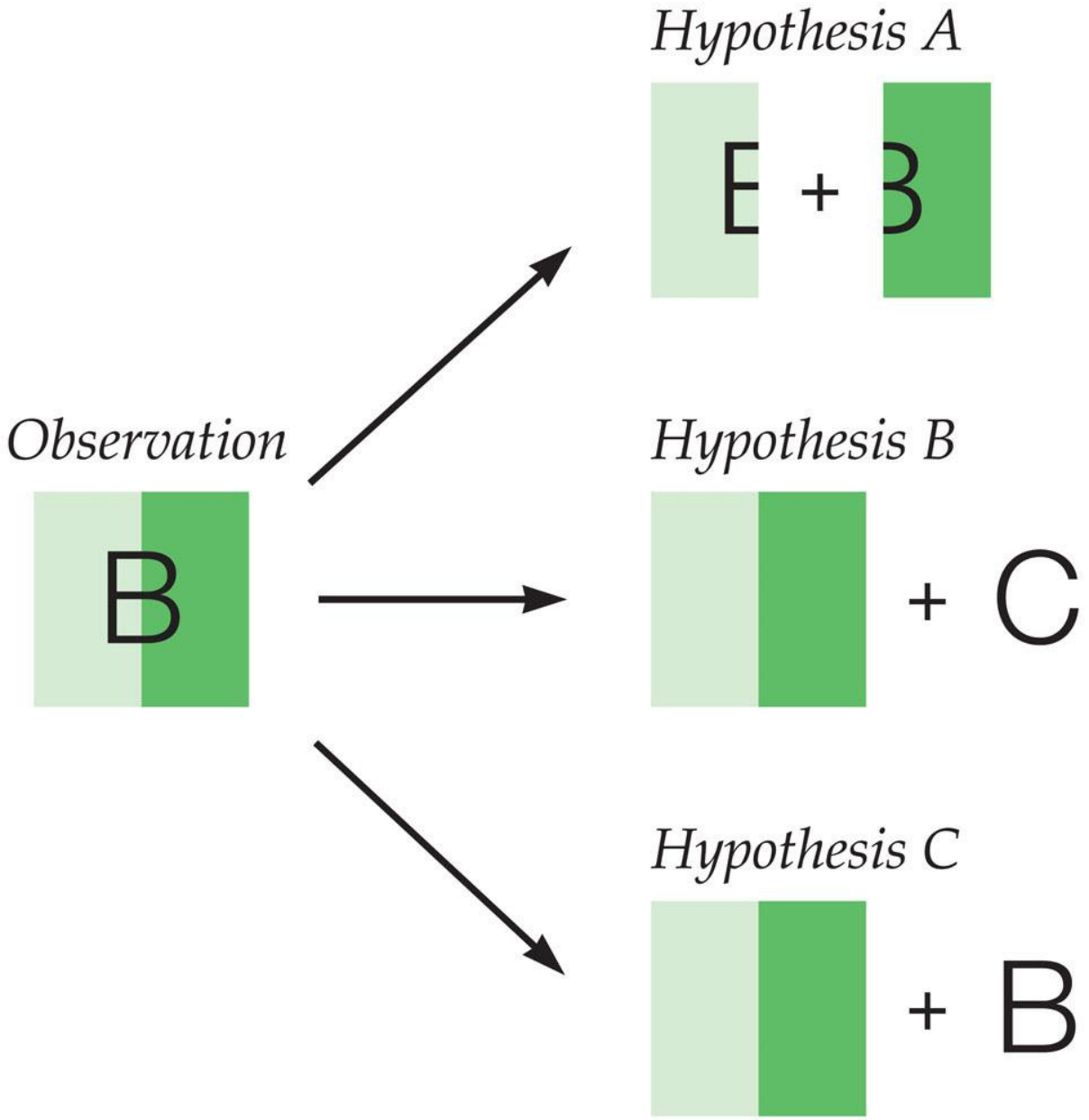
- Five principles of middle vision:
  1. Bring together that which should be brought together
  2. Split asunder that which should be split asunder
  3. Use what you know
  4. Avoid accidents
  5. Seek consensus and avoid ambiguity

Middle Vision

From metaphor to formal model

- “Perceptual committees” is a metaphor, but there are formal, mathematical models that can be used.
- The Bayesian approach: A formal, mathematical system that combines information about the current stimulus with prior knowledge about the world.

Figure 4.35 When the visual system is faced with a stimulus, it tries to figure out the most likely situation in the world that has produced this particular pattern of activity



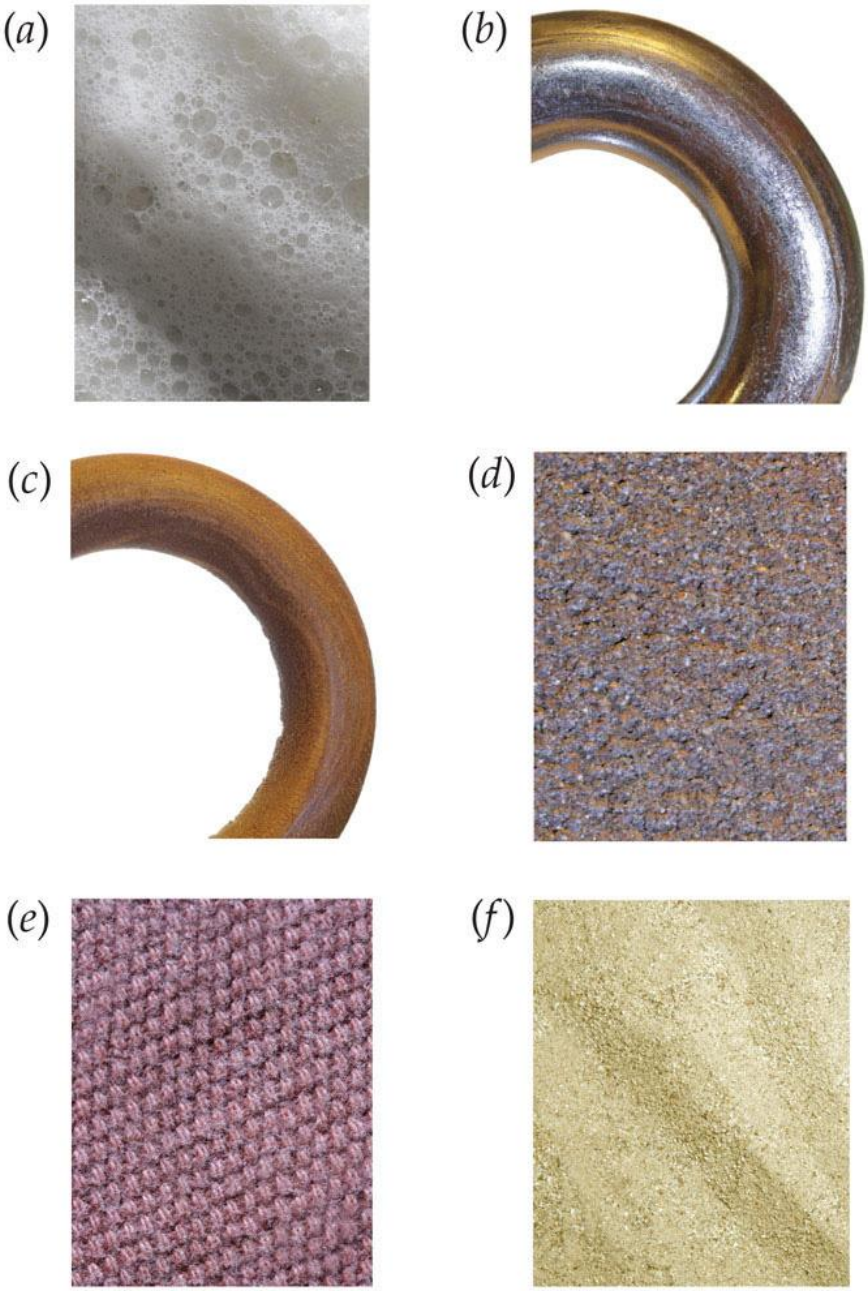
***SENSATION & PERCEPTION 4e*, Figure 4.35**

Sensation & Perception in Everyday Life: Material Perception: The Everyday Problem of Knowing What It Is Made Of

Material Perception

- We know what substances objects are made of at a glance.
- Independent of object recognition
- Has to do with reflectance properties of surfaces

Figure 4.36 You are remarkably good at figuring out what things are made of



***SENSATION & PERCEPTION 4e*, Figure 4.36**

Object Recognition

Moving from V1 to IT in the *what* pathway, neurons respond to more and more complex stimuli.

- By area V4, cells are interested in stimuli such as fans, spirals, and pinwheels.
- It is difficult to know exactly what V4 neurons like, but it is something more complicated than spots or bars of light.

Figure 4.37 Response of V4 cells to different shapes

(*a*)

(*b*)

***SENSATION & PERCEPTION 4e*, Figure 4.37**

Object Recognition

Some areas of human cortex are specialized to process certain types of stimuli.

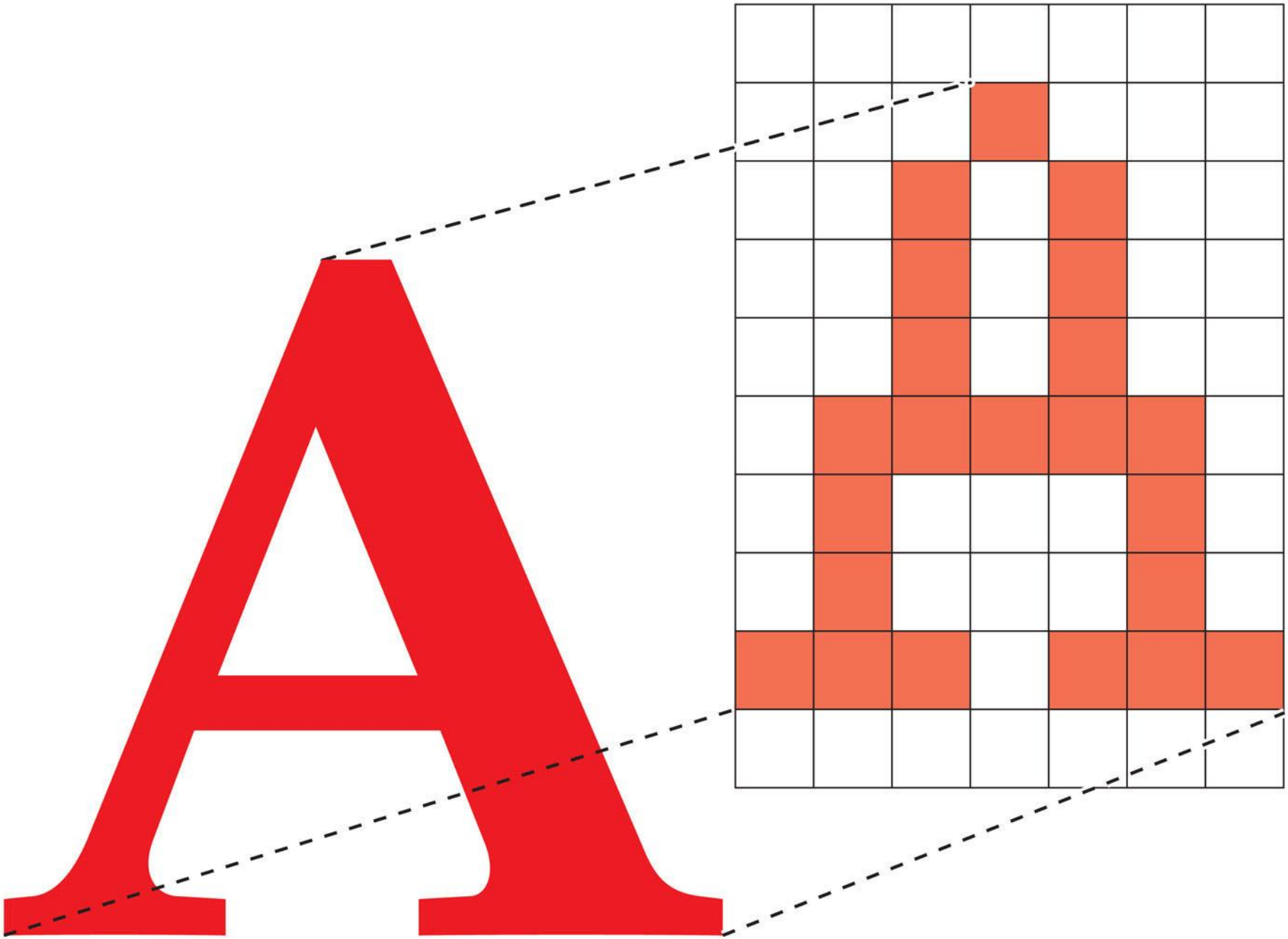
- PPA: Parahippocampal place area—responds preferentially to places, such as pictures of houses.
- FFA: Fusiform face area—responds to faces more than other objects.
- EBA: Extrastriate body area—specifically involved in the perception of body parts

Object Recognition

Templates versus structural descriptions

- Naïve template theory: The proposal that the visual system recognizes objects by matching the neural representation of the image with a stored representation of the same “shape” in the brain.
- Structural description: A description of an object in terms of its parts and the relationships between those parts.

Figure 4.40 A basic template



SENSATION & PERCEPTION 4e, Figure 4.40

Figure 4.41 The problem with templates is that we need a lot of them

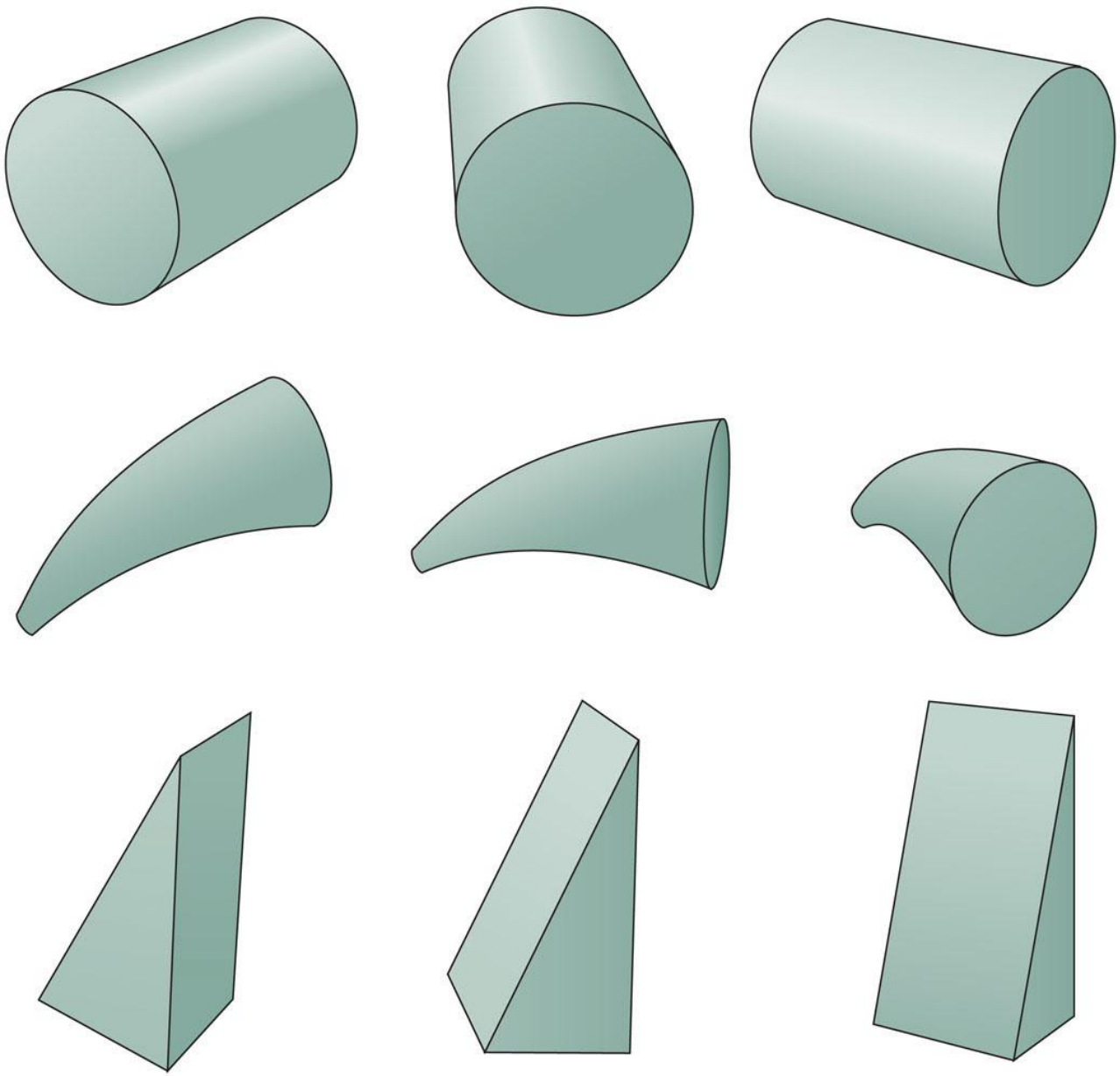
***SENSATION & PERCEPTION 4e*, Figure 4.41**

Object Recognition

Recognition-by-components

- Biederman's model of object recognition: Holds that objects are recognized by the identities and relationships of their component parts.
- Geons: The "geometric ions" out of which objects are built.

Figure 4.42 Three of the 36 or so geons in Biederman's recognition-by-components model of object recognition



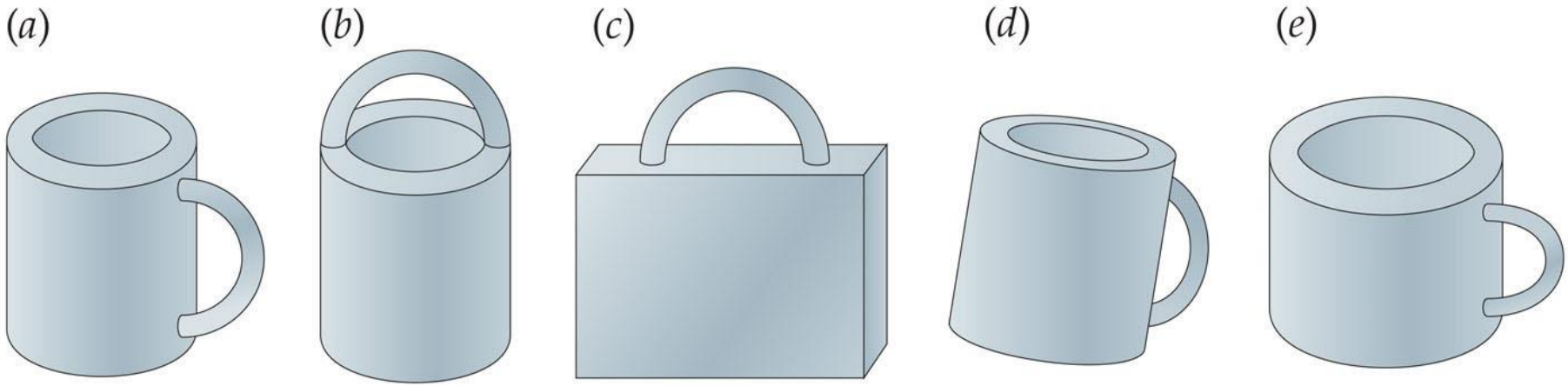
***SENSATION & PERCEPTION 4e*, Figure 4.42**

Object Recognition

Viewpoint invariance

1. A property of an object that does not change when an observer changes viewpoint
2. A class of theories of object recognition that proposes representations of objects that do not change when viewpoint changes

Figure 4.43 Combining geons can create a wide variety of object representations



***SENSATION & PERCEPTION 4e*, Figure 4.43**

Object Recognition

Problems with structural-description theories

- Object recognition is not completely viewpoint-invariant.
- Observers show viewpoint effects in object recognition.
  - The farther an object is rotated away from a learned view, the longer it takes to recognize.
- Geons aren’t always the best descriptions of objects.

Object Recognition

Multiple recognition committees?

- Perhaps there are several object recognition processes, depending on the category level.
  - Entry-level category: For an object, the label that comes to mind most quickly when we identify the object.
  - Subordinate-level category: A more specific term for an object.
  - Superordinate-level category: A more general term for an object.

Object Recognition

Faces: An illustrative special case

- Face recognition seems to be special and different from object recognition.
- Prosopagnosia: An inability to recognize faces.
- Agnosia: A failure to recognize objects despite being able to see them.
- Double dissociation: When one perceptual function can be damaged without affecting the other.

Figure 4.46 Which of these two photos has been altered?

***SENSATION & PERCEPTION 4e*, Figure 4.46**

Figure 4.46 Which of these two photos has been altered?

Object Recognition

Perception is a two-way street: Feedback and reentrant processing

- Initial object recognition can occur very quickly (150 ms), but that's not the end of the story.
- The brain continues to process information, sending signals up and down the *where* pathway.
- Object recognition should be seen as a conversation among many parts of the brain rather than as a one-way progression.