

Perception: Psychophysics and Modeling

07 | Object recognition |

Felix Wichmann



Neural Information Processing Group
Eberhard Karls Universität Tübingen

Overview

*The Problems of Perceiving and Recognising Objects (**VLo7-Object Recognition 1**)*

*Mid-level vision (**VLo7-Object Recognition 1**)*

- What are “edges” and (illusionary) “contours”?
- Gestalt psychology and “Gestalt laws” of perceptual organisation

*More on mid-level vision (**VLo8-Object Recognition 2**)*

- Accidental viewpoint and non-accidental features
- Figure-ground, occlusion, wholes and parts
- Texture segmentation, grouping and camouflage

*Neuroscience of object recognition (**VLo8-Object Recognition 2**)*

*Object representation (**VLo9-Object Recognition 3**)*

- Structural description models
- View-based models

*Object recognition by algorithms: DNNs (**VL10-Object Recognition 4**)*

Overview

The Problems of Perceiving and Recognising Objects (**VLo7-Object Recognition 1**)

Mid-level vision (**VLo7-Object Recognition 1**)

- What are “edges” and (illusionary) “contours”?
- Gestalt psychology and “Gestalt laws” of perceptual organisation

More on *mid-level vision* (**VLo8-Object Recognition 2**)

- Accidental viewpoint and non-accidental features
- Figure-ground, occlusion, wholes and parts
- Texture segmentation, grouping and camouflage

Neuroscience of object recognition (**VLo8-Object Recognition 2**)

Object representation (**VLo9-Object Recognition 3**)

- Structural description models
- View-based models

Object recognition by algorithms: DNNs (**VL10-Object Recognition 4**)

Literature

Wolfe, J.M. et al. (2015). "Perceiving and recognizing objects, ch. 4, pp. 88–120. *Sensation and Perception*. 4th edition, Sunderland, Massachusetts: Sinauer.

Supplementary literature

Peissig, J.J. and Tarr, M.J. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, 58: 75–96.

Supplementary Literature on Gestalt Psychology

Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M. and von der Heydt, R. (2012). A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure-Ground Organization. *Psychological Bulletin*, 138(6): 1172–1217.

Wagemans, J., Feldman, J., Gepshtain, S., Kimchi, R., Pomerantz, J.R., van der Helm, P.A. and van Leeuwen, C. (2012). A Century of Gestalt Psychology in Visual Perception: II. Conceptual and Theoretical Foundations. *Psychological Bulletin*, 138(6): 1218–1252.

Green, C.D. (2000). Introduction to: "Perception: An introduction to the Gestalt-Theorie" by Kurt Koffka (1922). *Classics in the History of Psychology*. An internet resource at York University, Toronto, CA.
<http://psychclassics.yorku.ca/Koffka/Perception/intro.htm>

Jäkel, F., Singh, M., Wichmann, F. A., and Herzog, M. H. (2016). An overview of quantitative approaches in Gestalt perception. *Vision Research*, 126:3–8.

Background

"At a functional level, visual object recognition is at the center of understanding how we think about what we see. Object identification is a primary end state of visual processing and a critical precursor to interacting with and reasoning about the world."

(Peissig & Tarr, 2007, p. 76)

Introduction

Object recognition is fast! Studies indicate that object recognition can happen in as little as 150 ms.

This is such a short time that—under such circumstances—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.

Object recognition is remarkably flexible ...



A46



The Problems of Perceiving and Recognising Objects

What do you see?

(a)



The Problems of Perceiving and Recognising Objects

What do you see?

(b)



The Problems of Perceiving and Recognising Objects

What do you see?

(c)



The Problems of Perceiving and Recognising Objects

Is this the same red house that was shown before?

(d)



The Problems of Perceiving and Recognising Objects

The pictures were just a bunch of pixels on a screen, but in each case you perceived a house.

How did you recognise all four images as depicting a house?

How did you recognise the first and third images as depicting the same house, but from different viewpoints?

Perceptual organisation

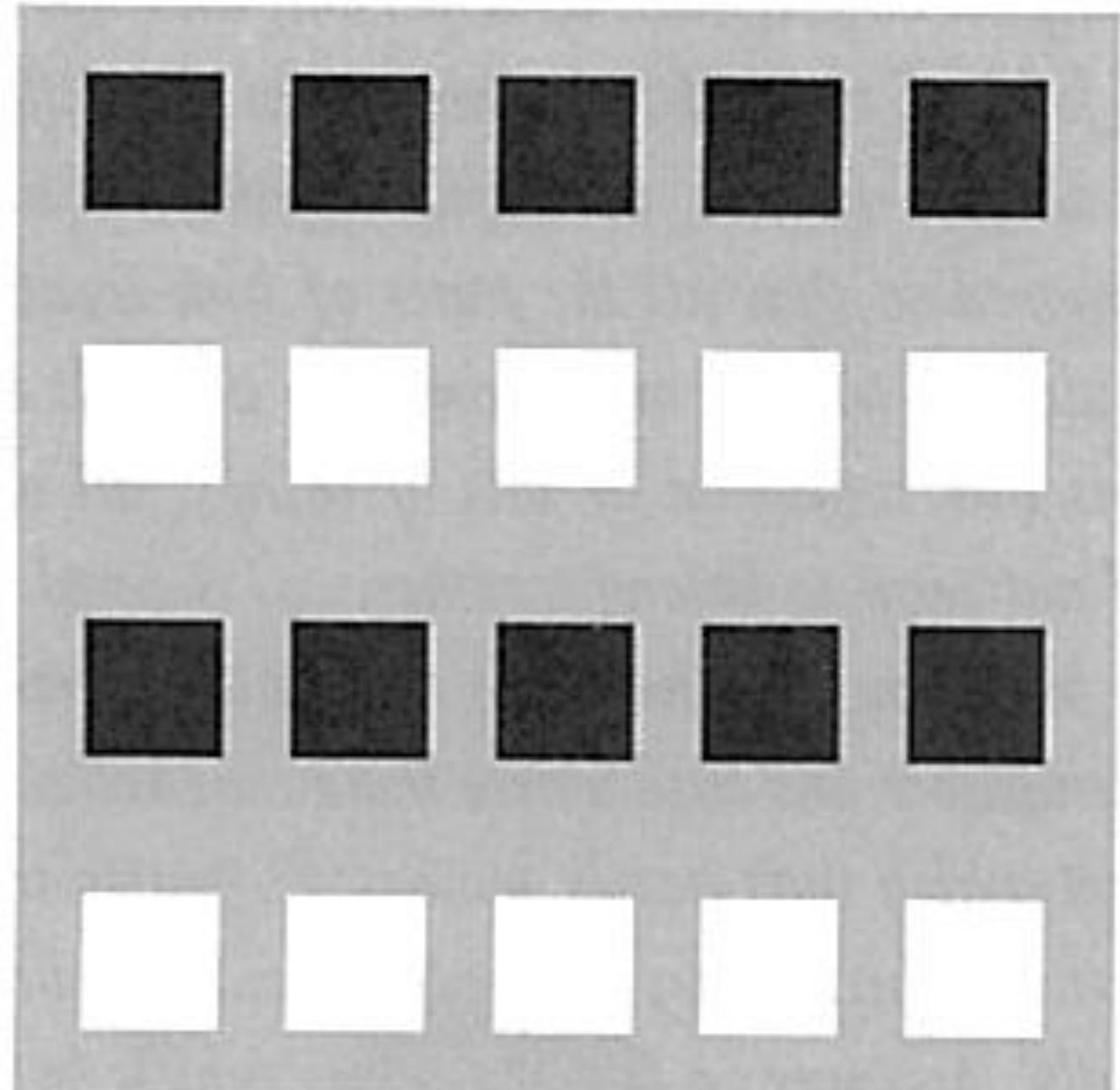
How does your visual system move from points of light, like pixels, to whole entities in the world, like houses?

What structure is present in the image to the left?

Now do you see?

The same information is present in both images. The top image is a lot like the one the visual system faces: the structure we perceive so effortlessly is not explicitly given in the retinal input, but must be discovered by the visual system itself.

5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	2	2	5	2	2	5	2	2	5	2	2	5	2	2	5	2	2
5	2	2	5	2	2	5	2	2	5	2	2	5	2	2	5	2	2
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	8	8	5	8	8	5	8	8	5	8	8	5	8	8	5	8	8
5	8	8	5	8	8	5	8	8	5	8	8	5	8	8	5	8	8
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	2	2	5	2	2	5	2	2	5	2	2	5	2	2	5	2	2
5	2	2	5	2	2	5	2	2	5	2	2	5	2	2	5	2	2
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	8	8	5	8	8	5	8	8	5	8	8	5	8	8	5	8	8
5	8	8	5	8	8	5	8	8	5	8	8	5	8	8	5	8	8
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5



From Palmer, Vision Science, p. 256

Mid-level vision

Mid-level 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.

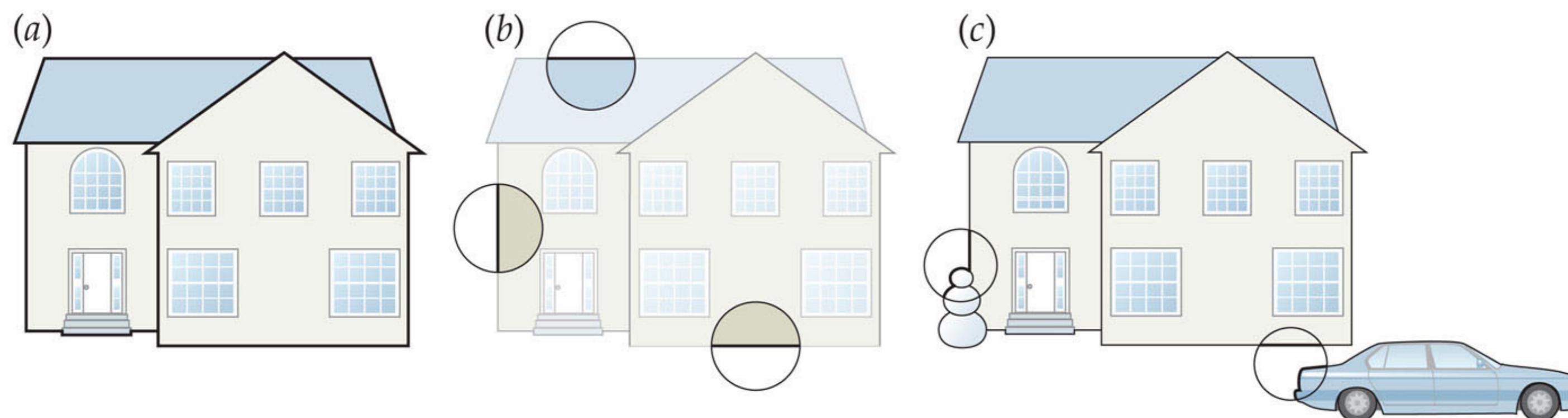
Determines which regions of an image should be grouped together into objects.

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?

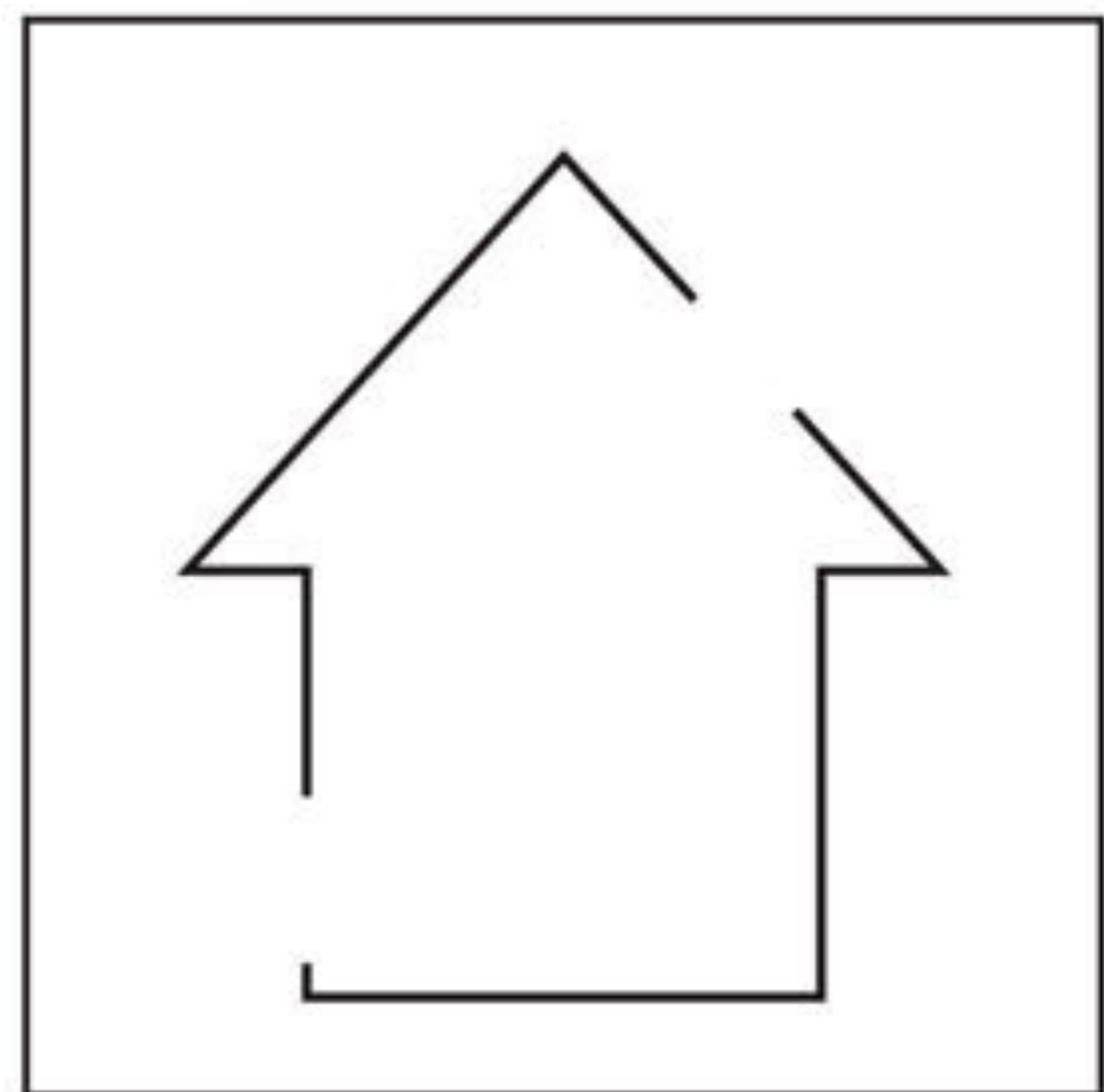


SENSATION & PERCEPTION 4e, Figure 4.9
© 2015 Sinauer Associates, Inc.

Mid-level vision: Too few edges

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

Sometimes computers don't find edges that humans see easily.

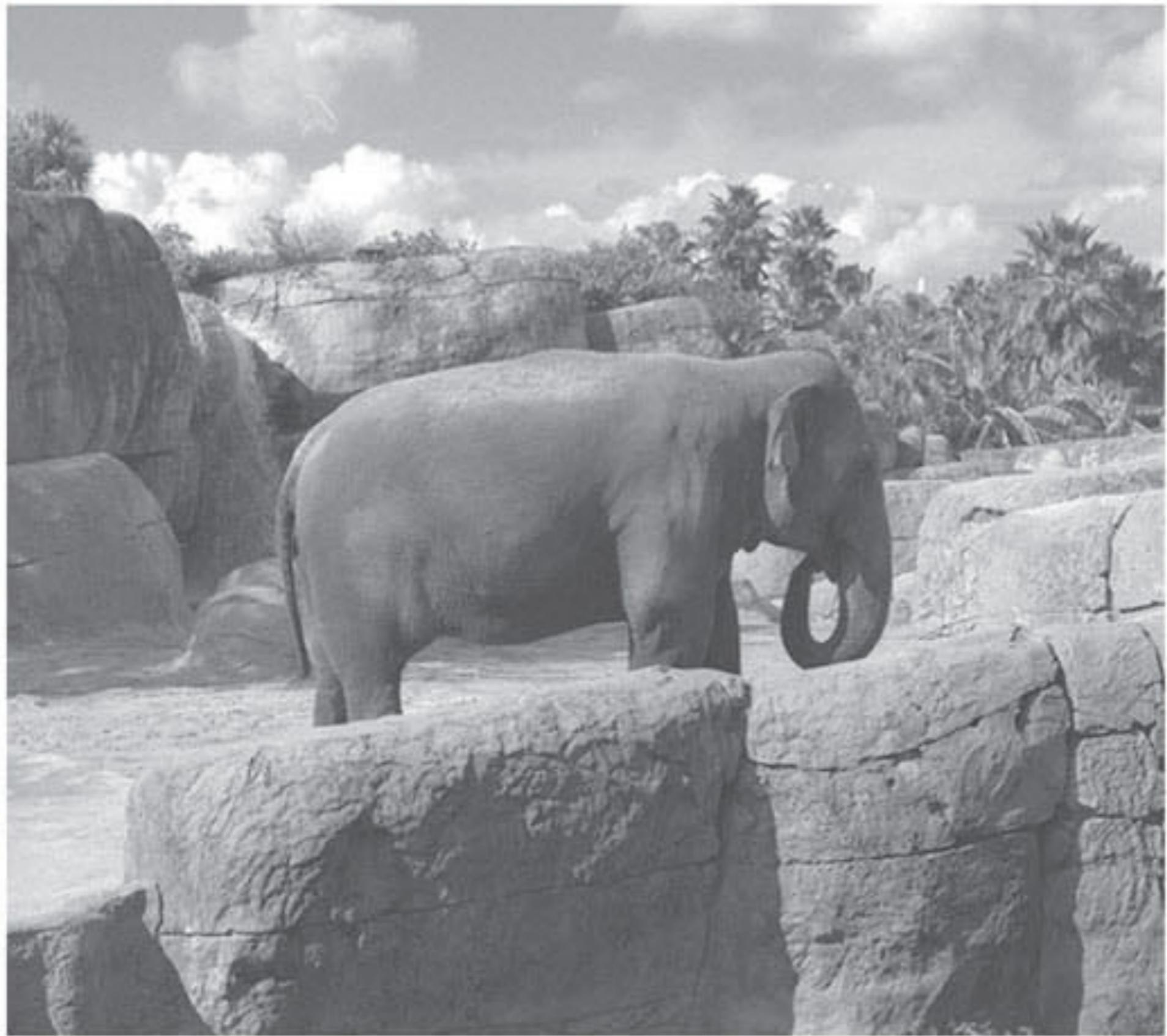


SENSATION & PERCEPTION 4e, Figure 4.10
© 2015 Sinauer Associates, Inc.

Mid-level vision: Too many edges

Sometimes computers find too many edges.

(b)



(c)



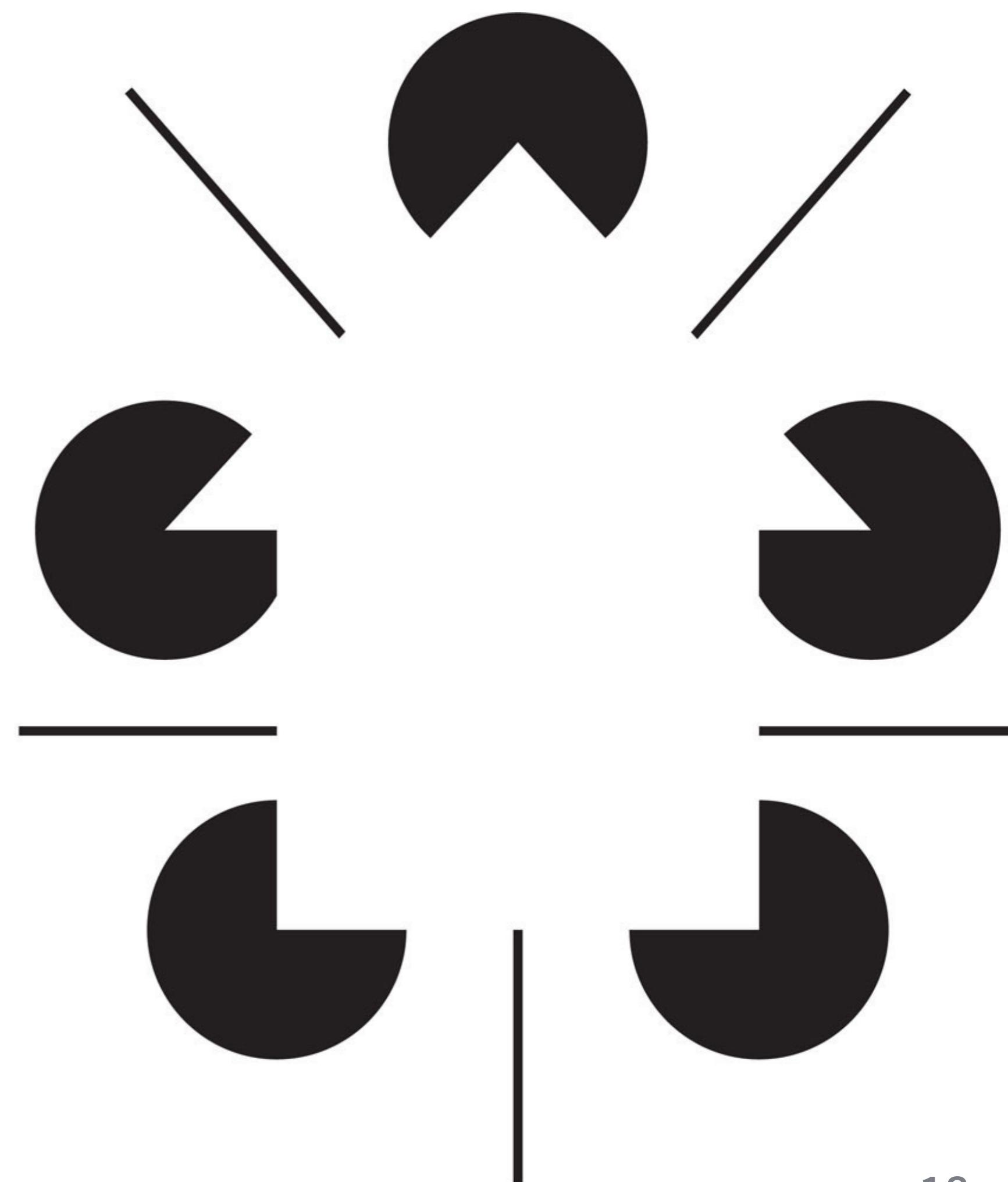
Mid-level vision: Illusory contours

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

This “house” outline is constructed from illusory contours.

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

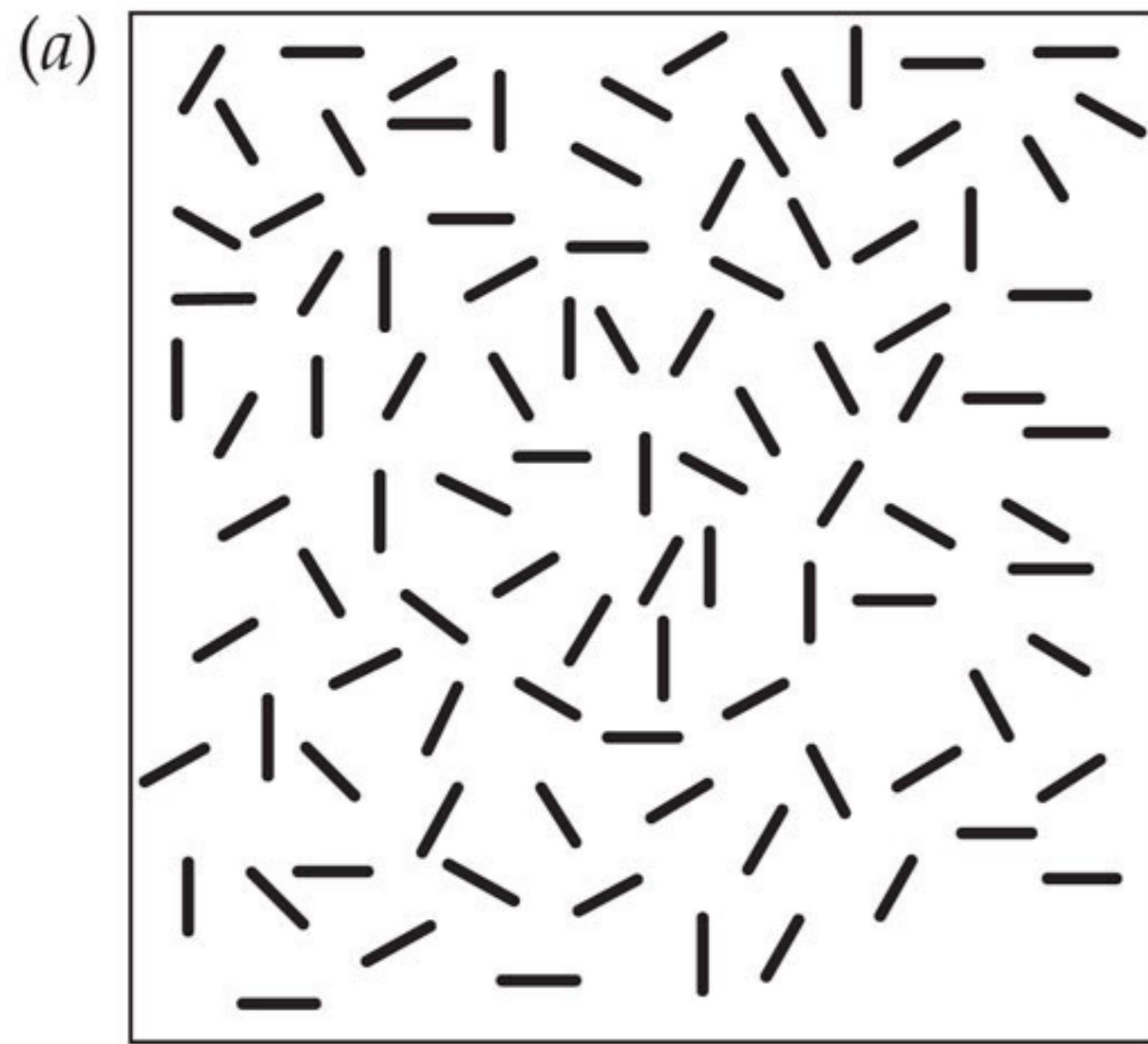
There is no house! Just some “Pac-Men” and disconnected lines.



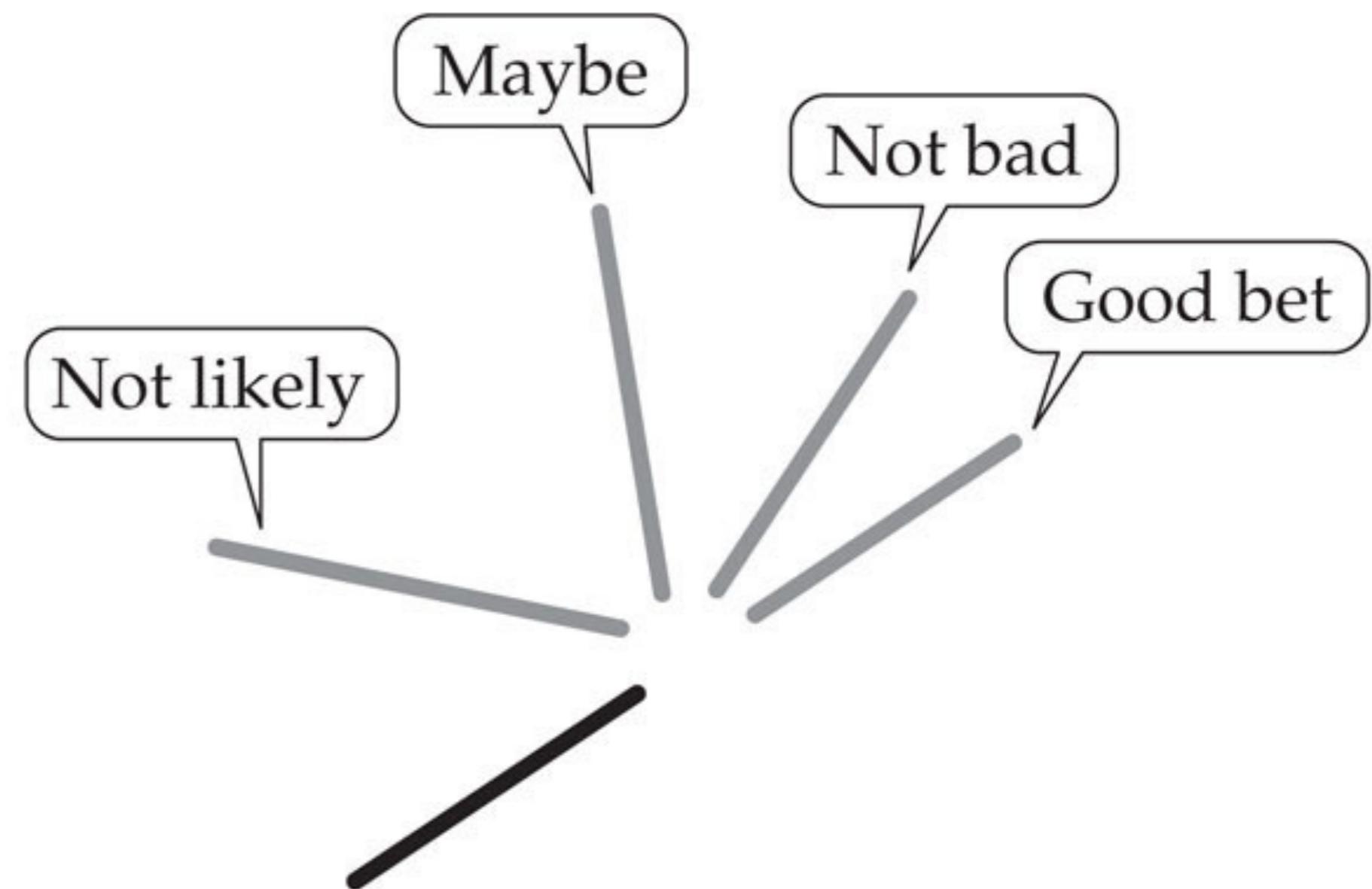
Mid-level vision: Rules that make contours

Some contours in an image will group because of good continuation.

Can you find the shape embedded within the field of lines in (a) in the next figure?



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



Mid-level vision: Rules that make contours

Gestalt: In German, “form” or “whole.”

Gestalt grouping rules: A set of rules that describe when elements in an image will appear to group together.

Previous examples show “good continuation”: A Gestalt grouping rule stating that two elements will tend to group together if they lie on the same contour.

A brief history of the Gestalt movement in psychology

Early psychologists and physicists* approached perception from a *structuralist* perspective:

Sensory atoms: primitive, indivisible elements of experience

Perception arises from rapid, unconscious associations between sensory atoms (memory-like linking of experiences)

Associations were simply added together: e.g. joining points on the retina (concatenation)

As observers learned more about the world via associations, their perceptions become richer and more complex

* e.g. Wilhelm Wundt (1832-1920), Ernst Mach (1838-1916), Christian von Ehrenfels (1859-1932)

A brief history of the Gestalt movement in psychology

von Ehrenfels: in addition to sense atoms, there is an extra element that emerges from their association (the “Gestalt-Qualität”).

For example: when hearing a melody, you perceive the notes but also something in addition, which binds them together into a tune (the Gestalt-Qualität).

It is the presence of this Gestalt-Qualität which, according to von Ehrenfels, allows a tune to be transposed to a new key, using completely different notes, but still retain its identity.

A brief history of the Gestalt movement in psychology

The Gestalt movement arose as a reaction against structuralism.

Inverted the structuralist theory. The Gestalt-Qualität *is* the immediate perception, and it cannot be reduced to mere atoms.

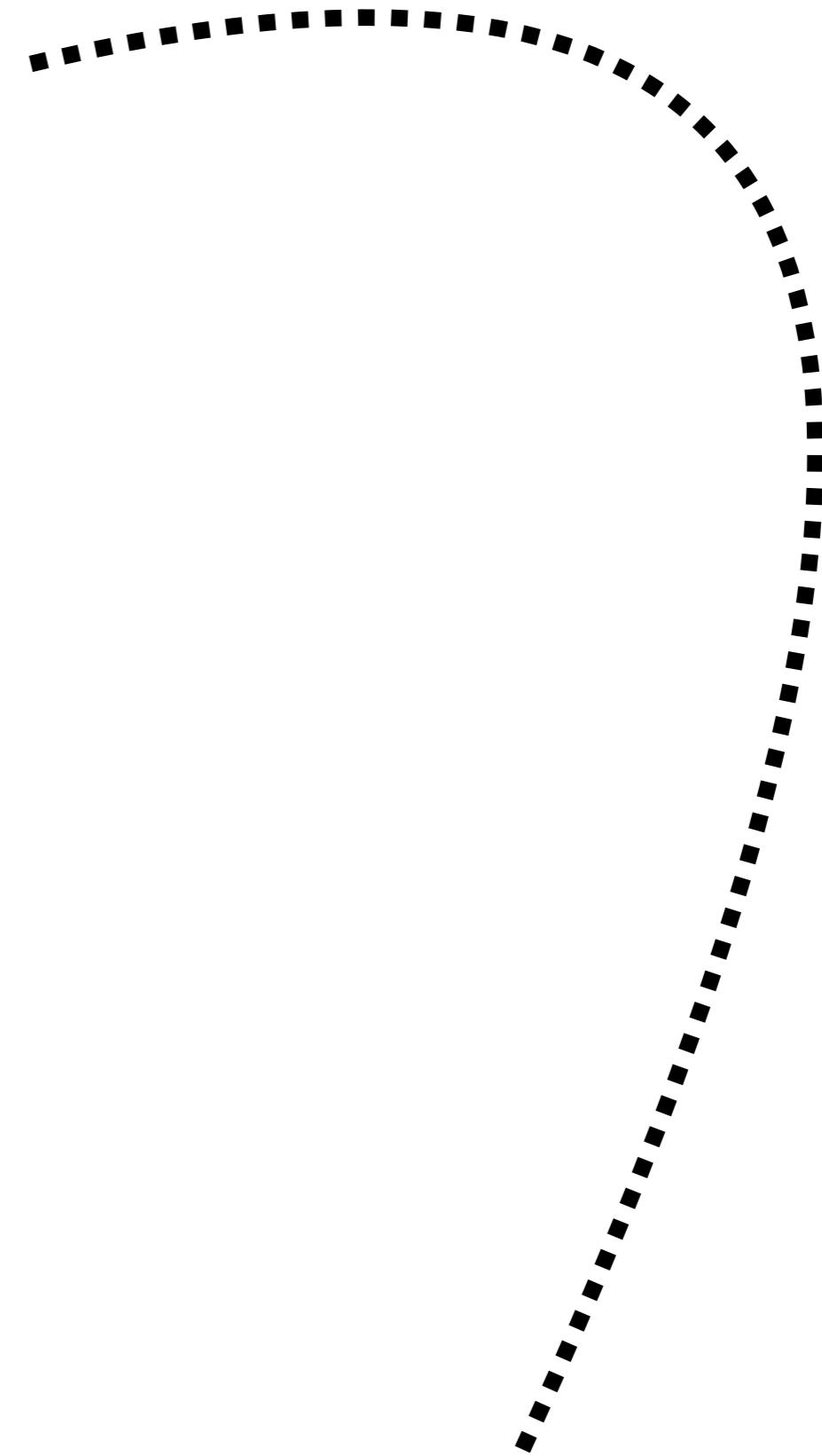
The primary proponents of the Gestalt theory were Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) and Kurt Koffka (1886-1941), who together founded the Berlin school.

Max Wertheimer: "what is given me by the melody does not arise ... as a secondary process from the sum of the pieces as such. Instead, what takes place in each single part already depends upon what the whole is," (1925/1938).

You hear the melody first, and *only then* can it be divided into individual notes.

A brief history of the Gestalt movement in psychology

Similarly for vision: you see the line first, and *only* then do you notice that it's made up of small squares.



Emergent properties

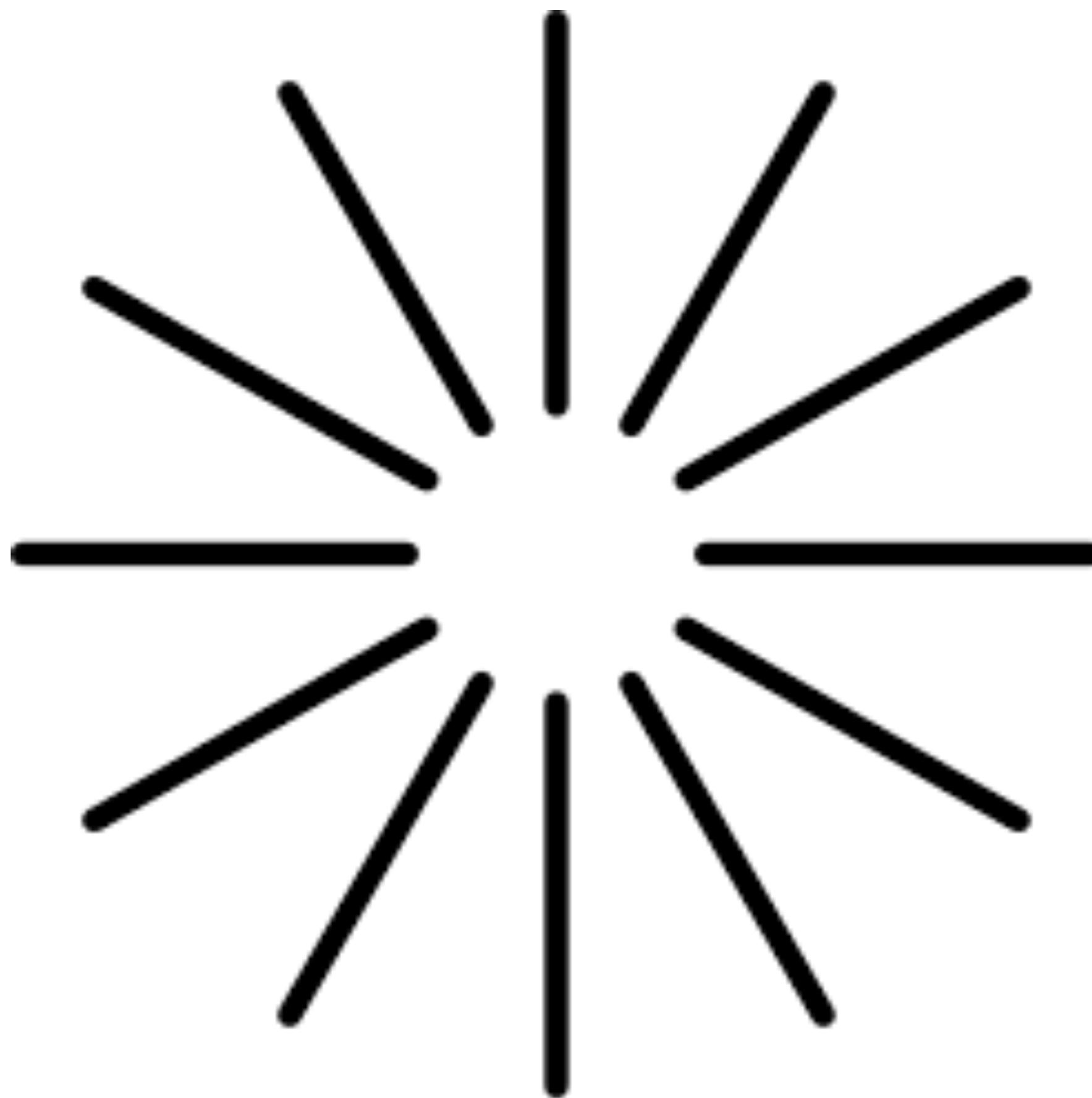
The squares themselves have just three perceptual properties: colour, size and position.

When the squares are arranged, new properties of length, orientation and curvature emerge.

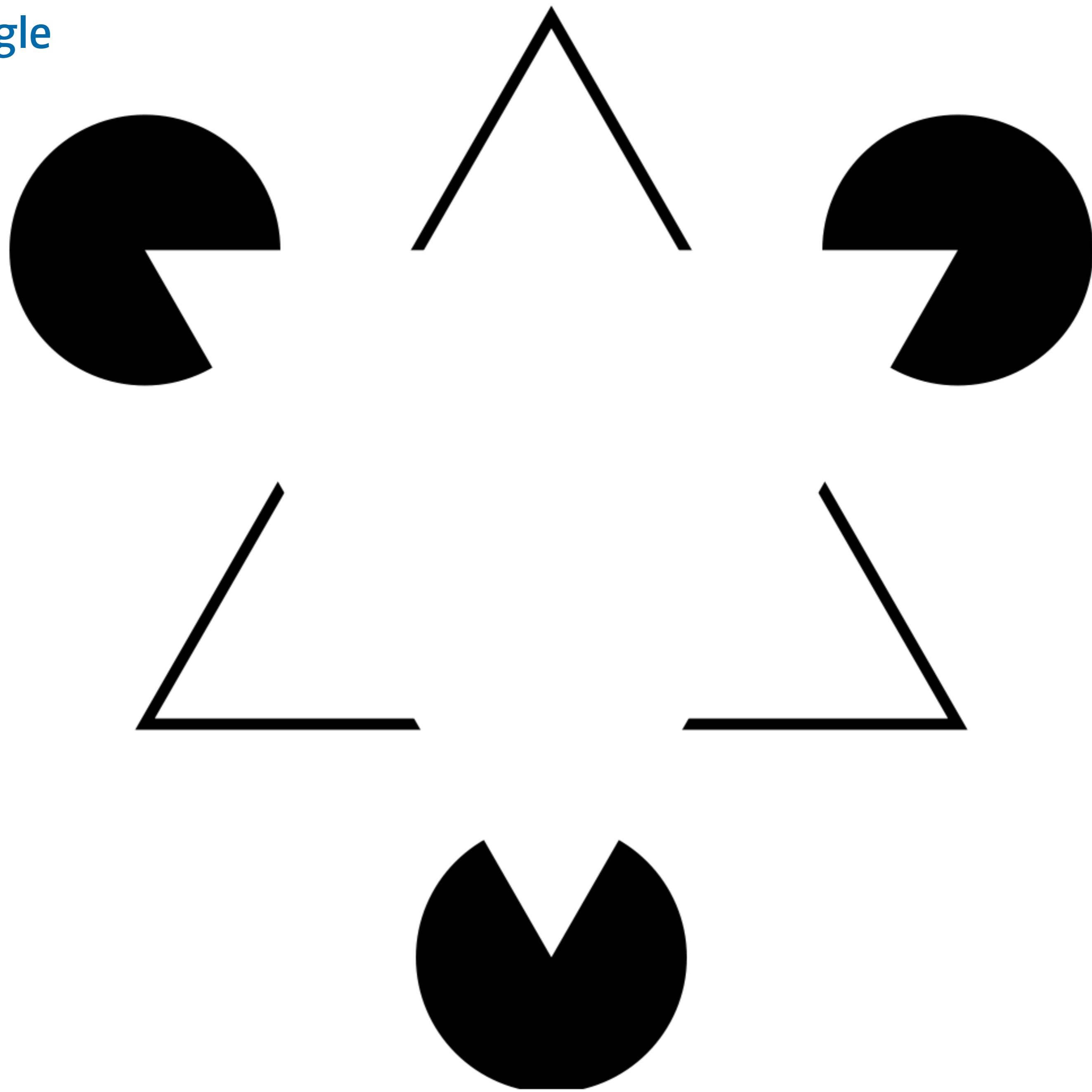
Emergent properties



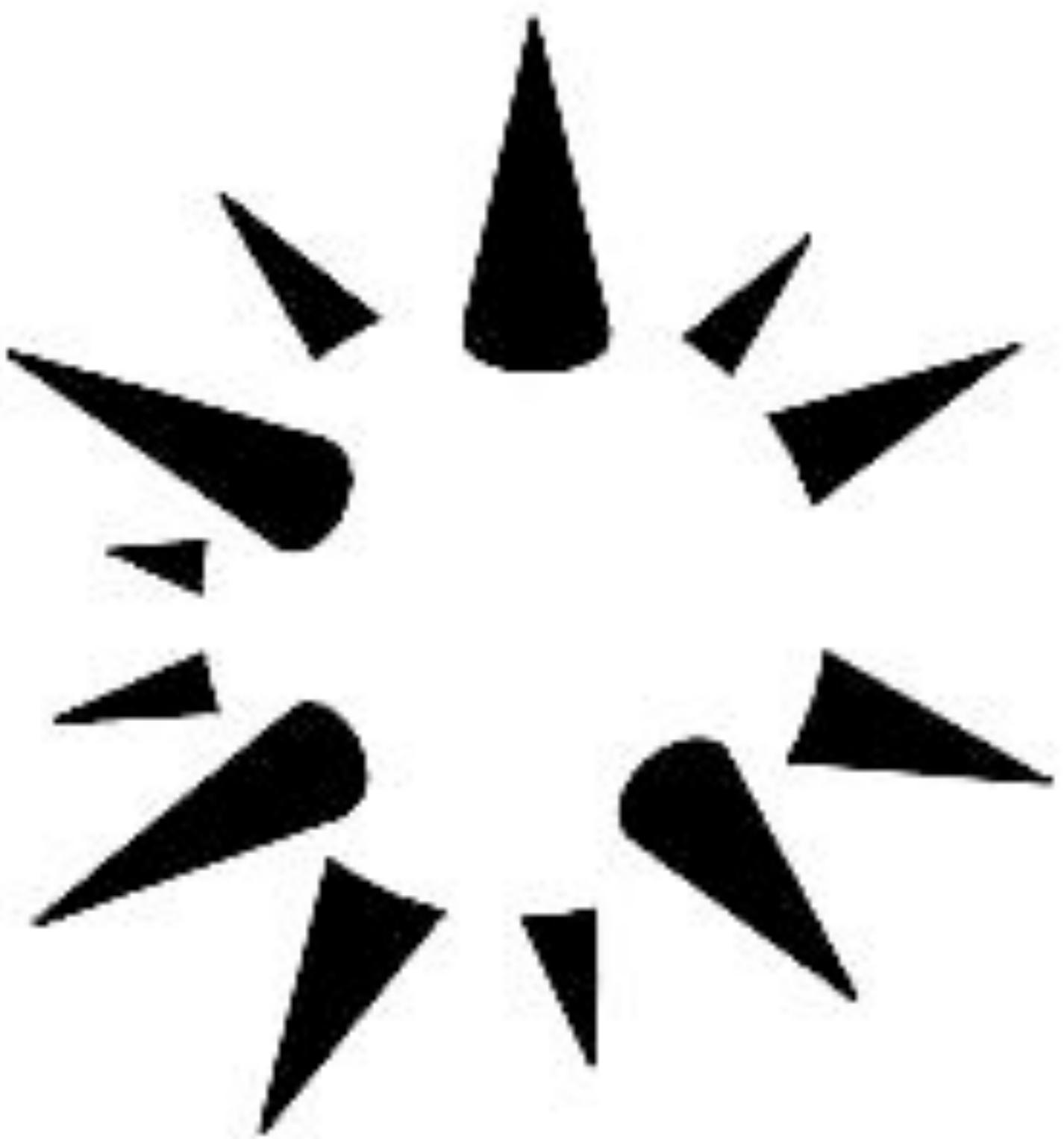
Ehrenstein Illusion: Reification



Kanizsa Triangle



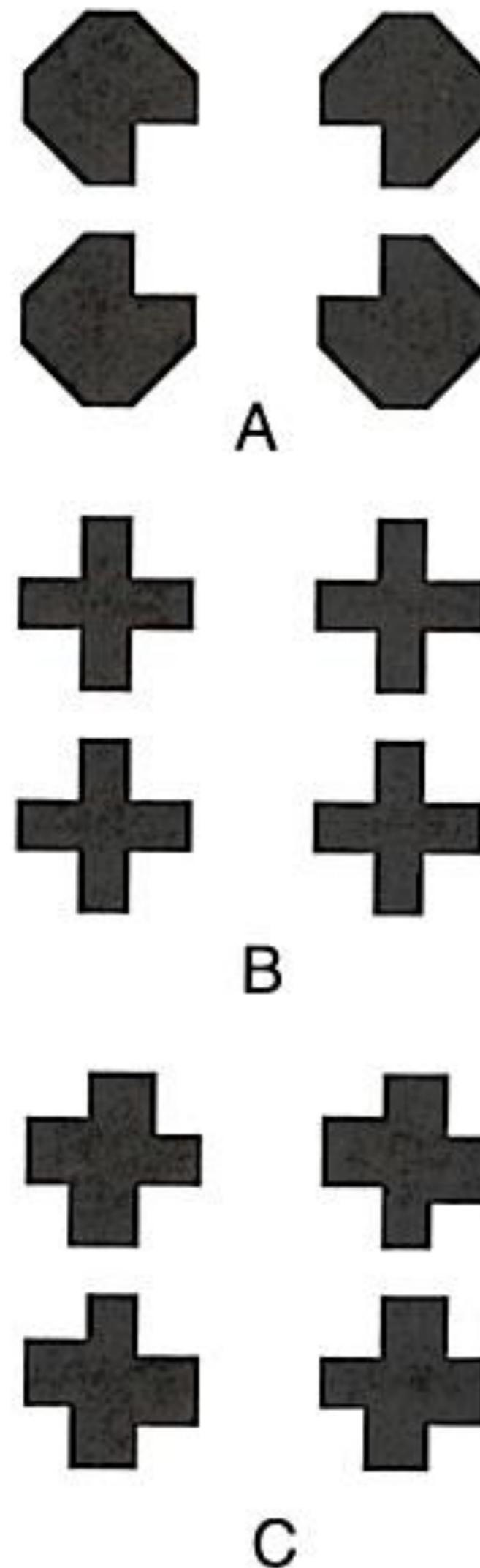
Idesawa's Spiky Sphere



Illusory contours

The perception of illusory contours can depend crucially on seemingly-subtle manipulations of the display.

Note that in all three cases, the local contour information in the middle region is the same.

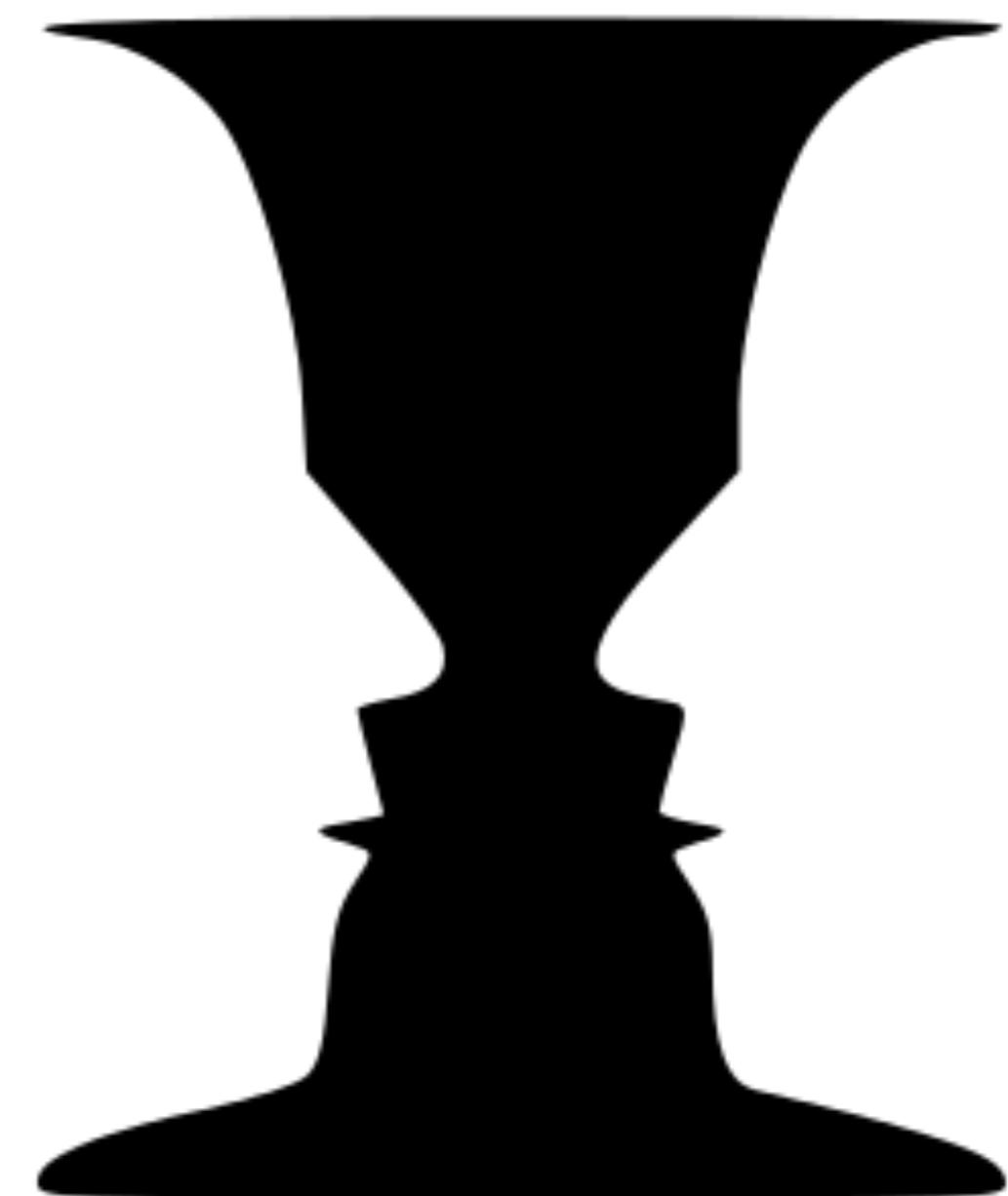
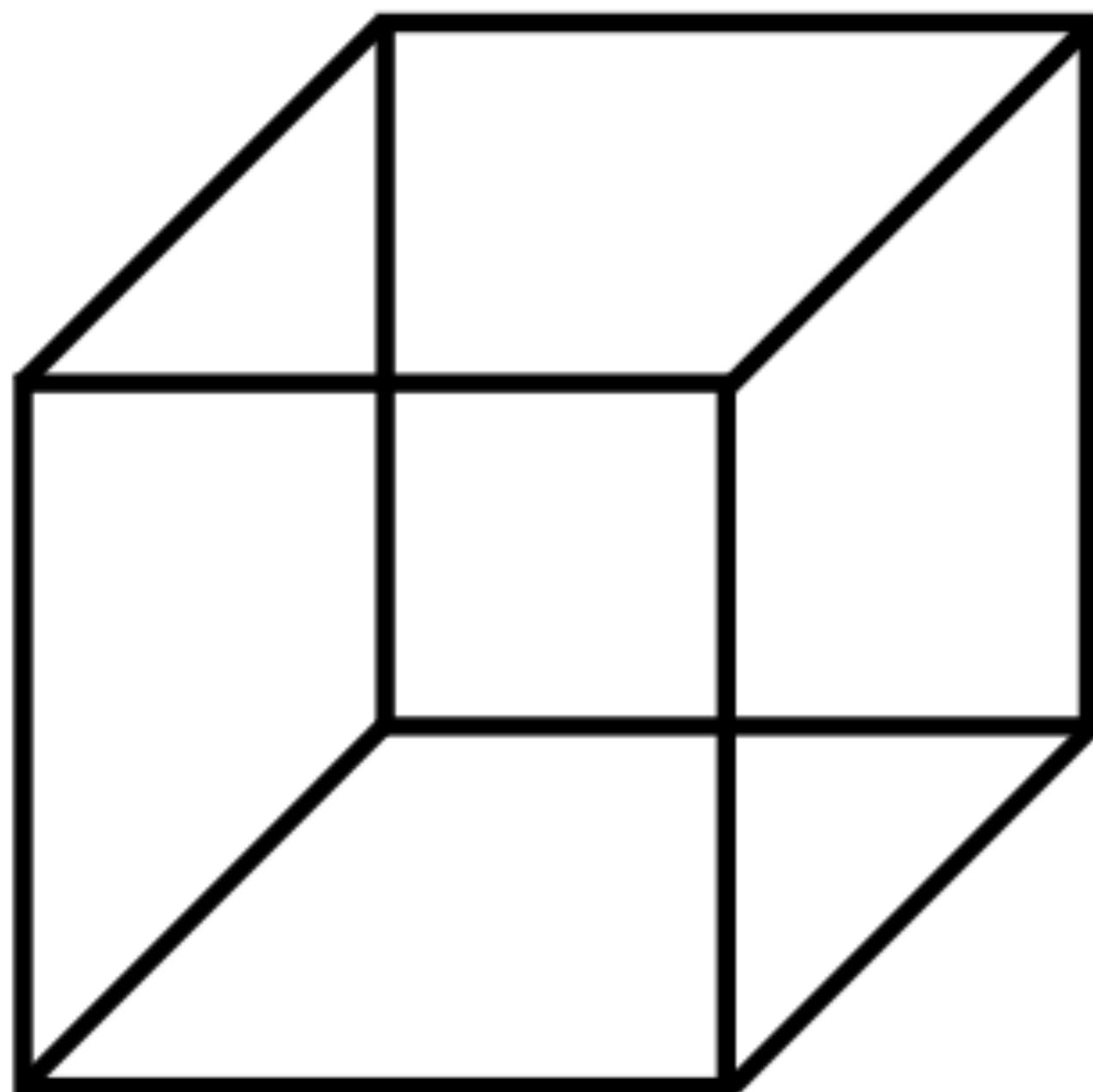


Multistability

Some of their primary evidence that sensory atoms cannot account for perception came from “reversible” or “multistable” figures.

If one were to list of all the individual sensations involved when one sees the “vase”, it would be identical to a list of all the sensations when one sees “faces”. Nevertheless, the two perceptions are very different.

Therefore the perception can't be reconstructed one-to-one from “sensory atoms”.



The whole is *different* from the sum of its parts

A brief history of the Gestalt movement in psychology

They tried to extend Gestalt principles to create a whole theory of psychology.

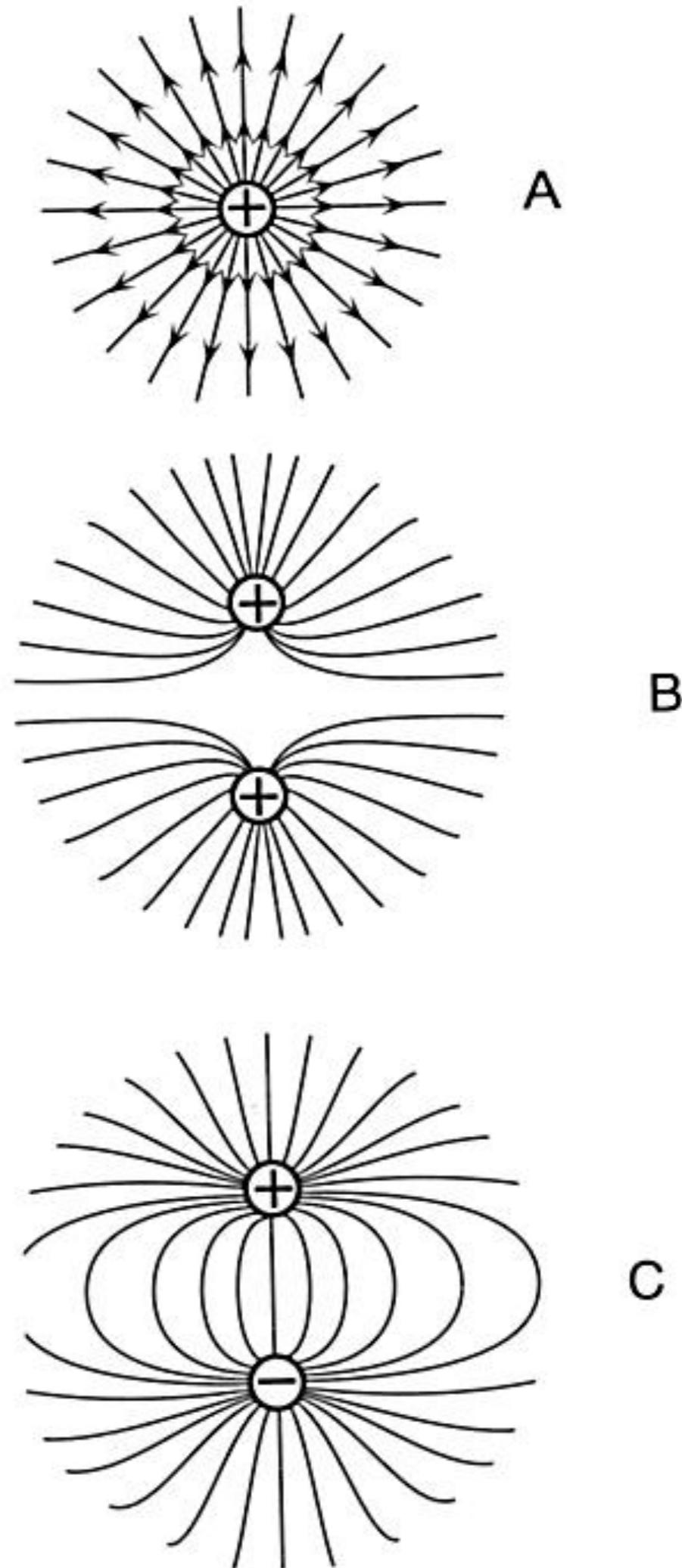
For example, Köhler studied learning in Chimpanzees, showing that contrary to most other theories of learning, Chimpanzees can learn via “sudden insight” into the structure of the problem, rather than incrementally via trial and error (e.g. Pavlov & Thorndike).

A brief history of the Gestalt movement in psychology

The Gestaltists thought of mental processes as analogous to force fields in physics (such as magnetic fields).

e.g. a single charged particle creates a certain field, but if a second particle is added, the field changes in a way that depends on the relationship between the particles.

The field depends on the structure of configurations rather than the sum of its parts.



From Palmer, *Vision Science*, p. 52

A brief history of the Gestalt movement in psychology

Koffka's 1935 book *Principles of Gestalt Psychology* formulated the basic question of vision research as "Why do things look as they do?". He rejected the idea that only visual illusions required explanation: why does a 2 cm line look longer than a 1 cm line? Not just "because it is longer". Perception requires explanation regardless of whether it is veridical.

A Gestalt theory of Science would "integrate the facts of inanimate nature, life, and mind into a single scientific structure".

Gestalt “laws” of perceptual
organisation

Gestalt principles of perceptual organisation

Mostly comes down to *perceptual grouping*: how are elements of a complex display perceived as “going together”?

Proximity: nearby items tend to group

Similarity: similar-looking items tend to group

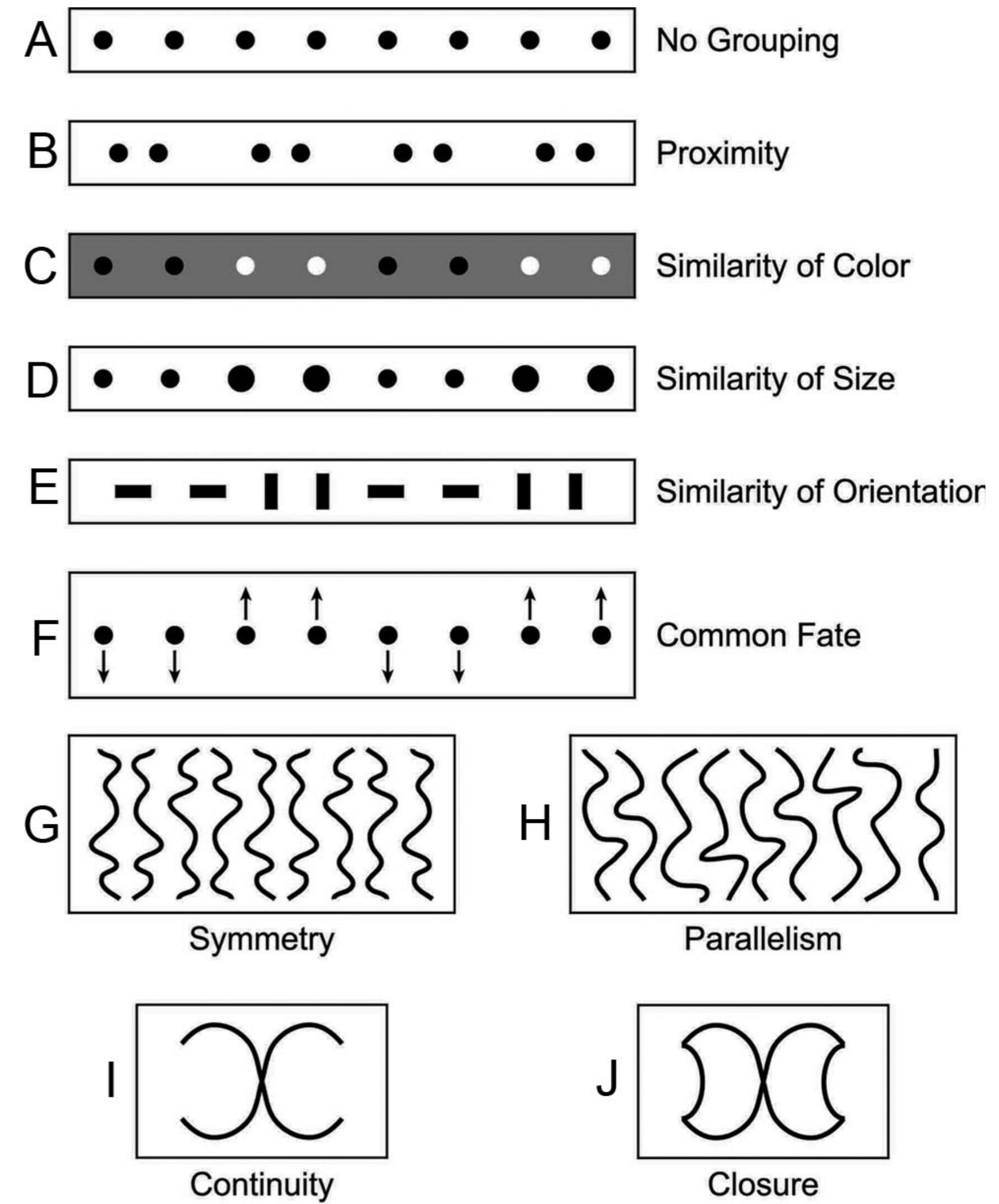
Common Fate: items that move together tend to group

Symmetry: symmetrical items tend to group

Parallelism: parallel items tend to group

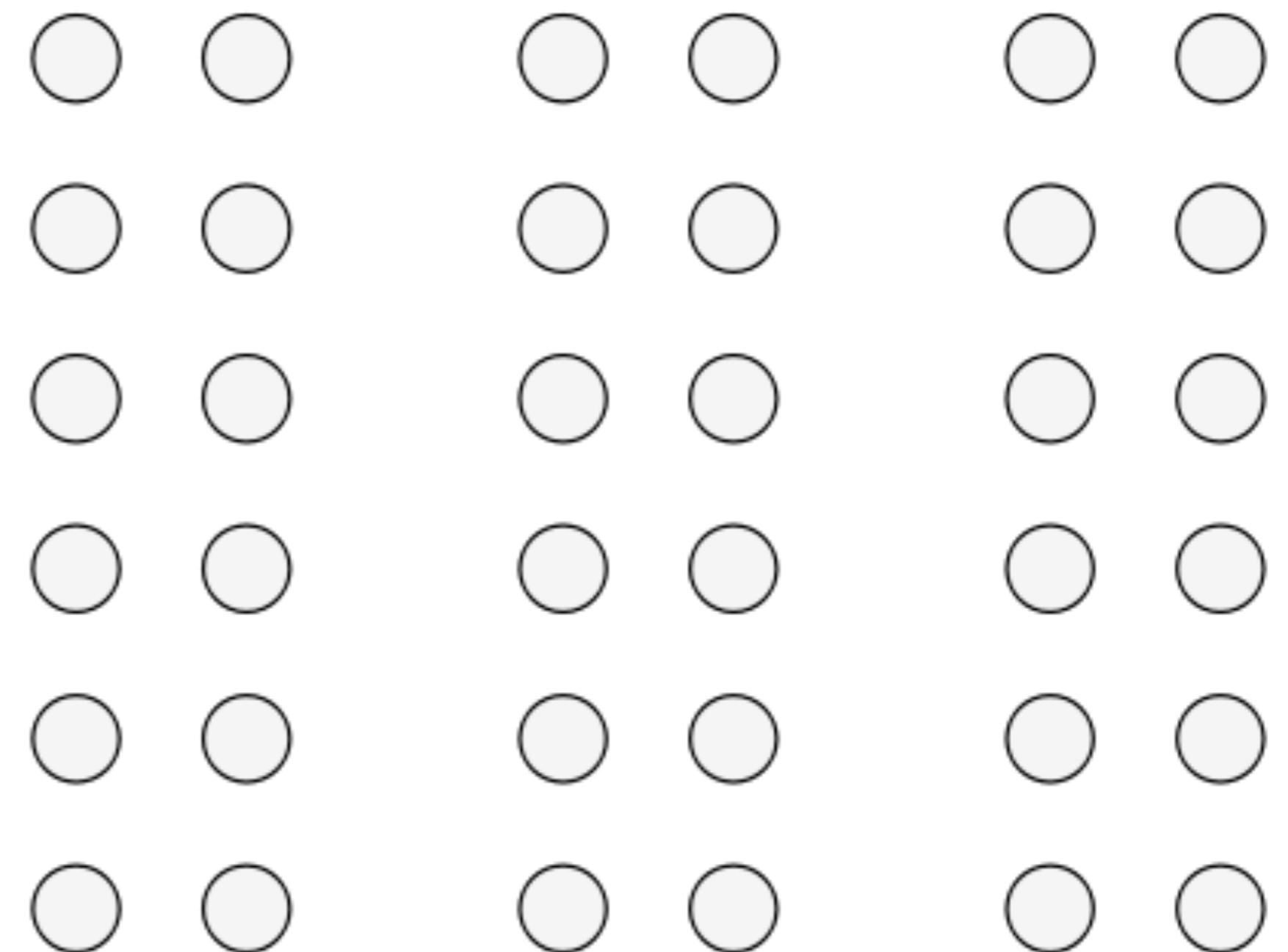
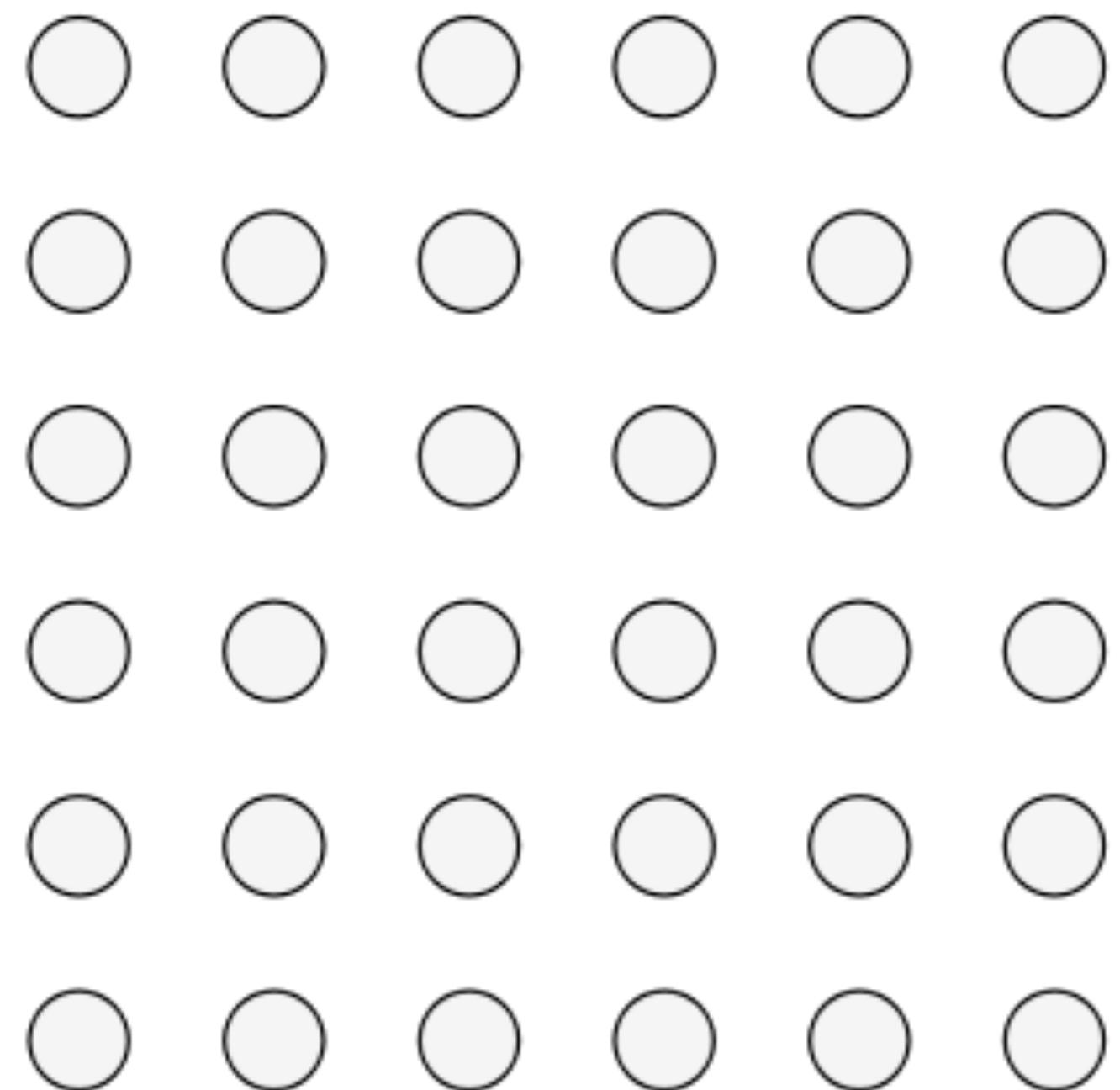
Good continuation (continuity): smooth curvatures tend to group

Closure: closed figures tend to group

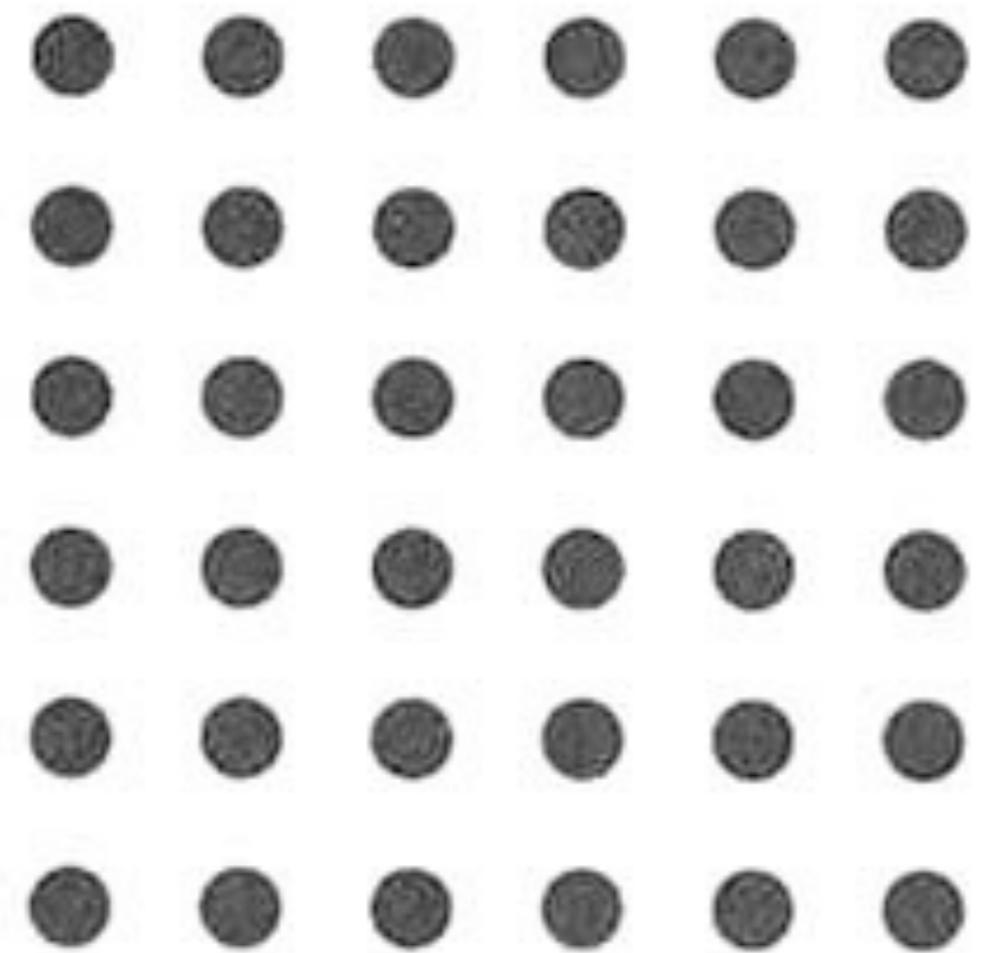


From Palmer, *Vision Science*, p. 258

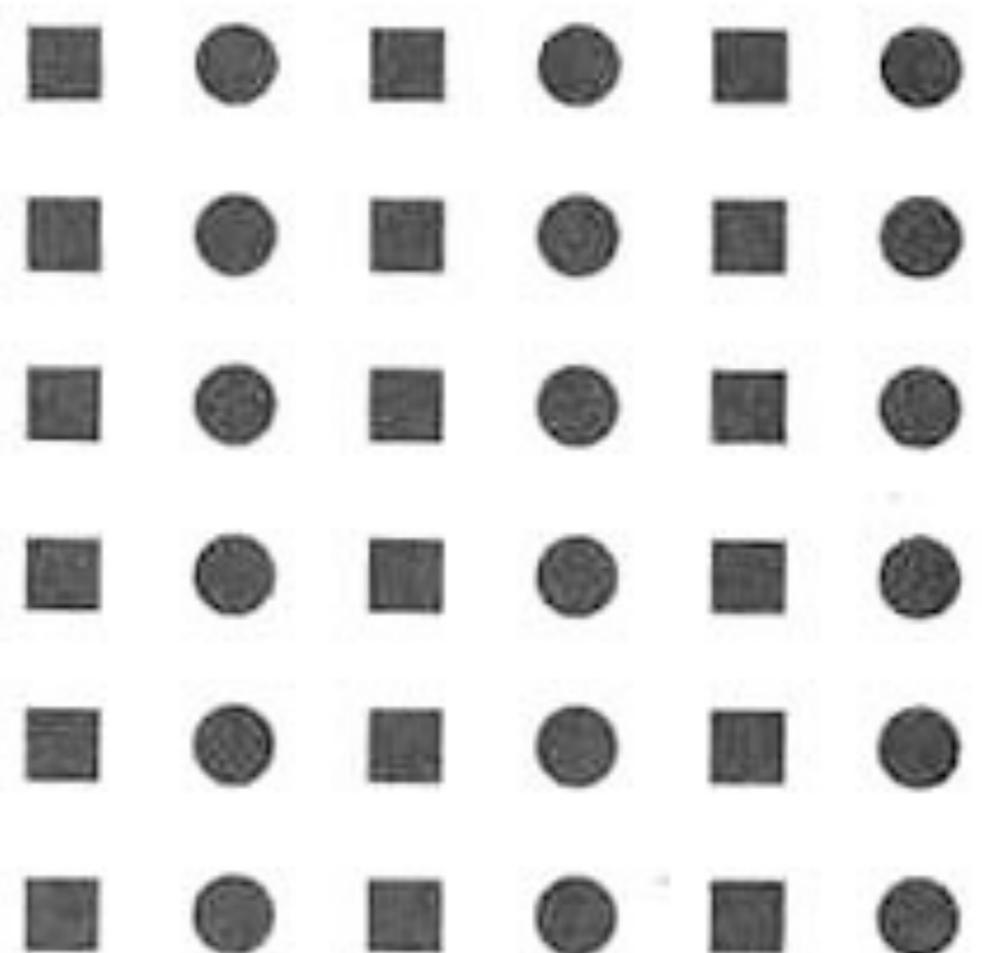
Proximity



Similarity

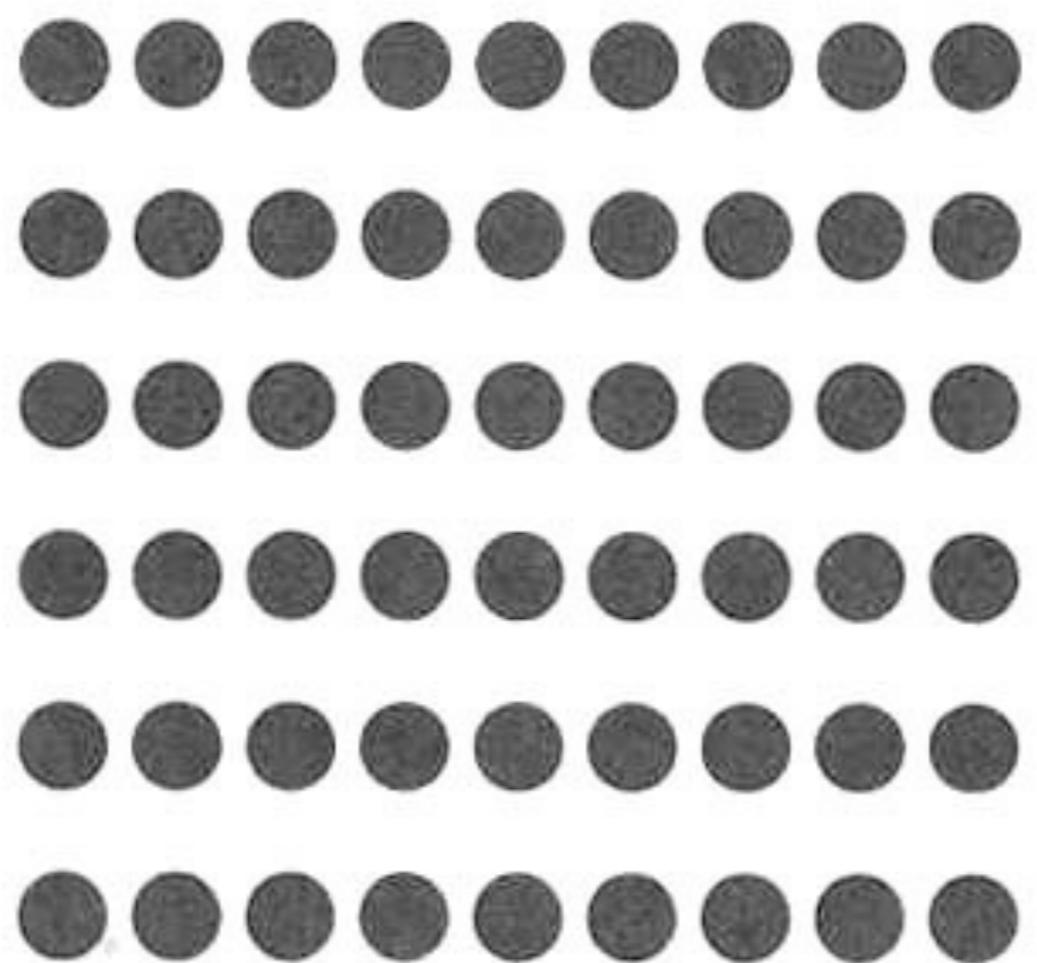


(a)

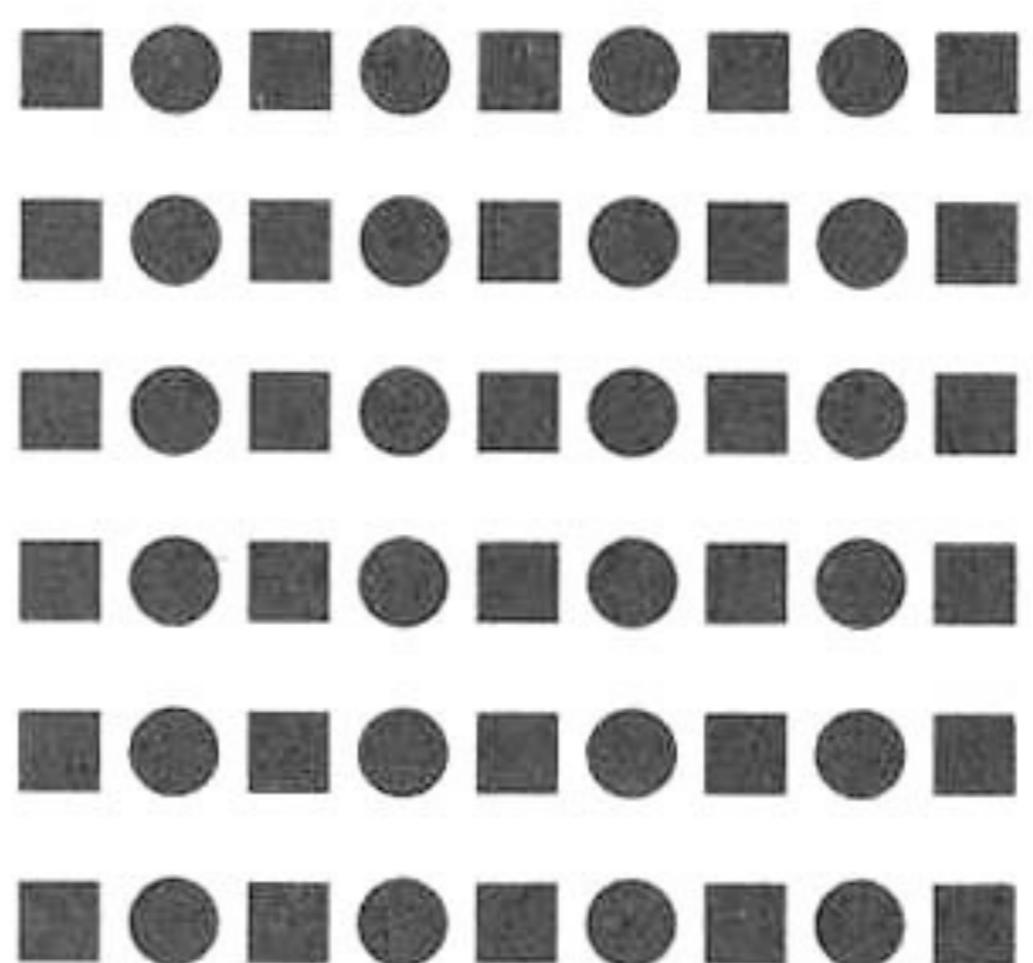


(b)

Proximity in competition with Similarity

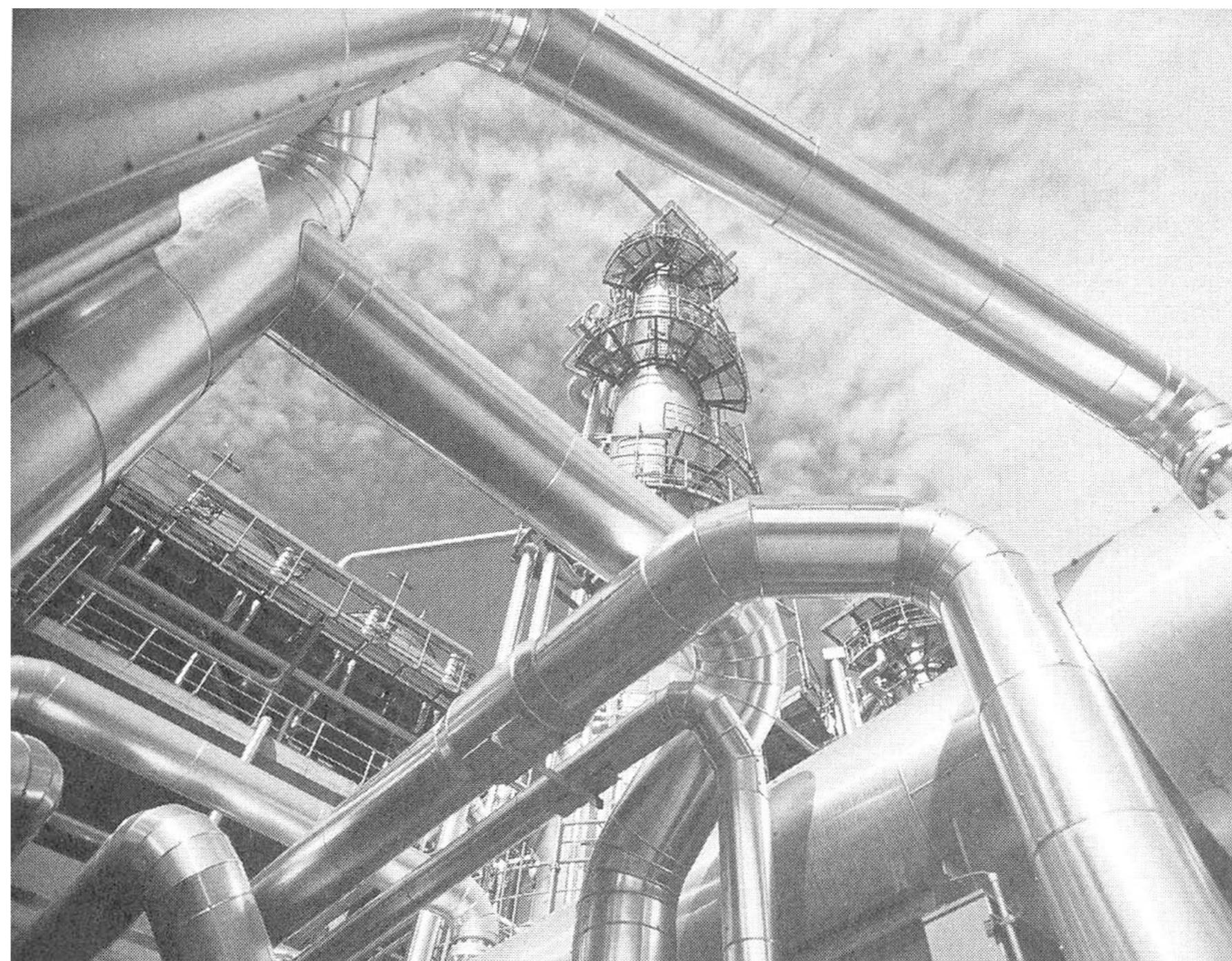
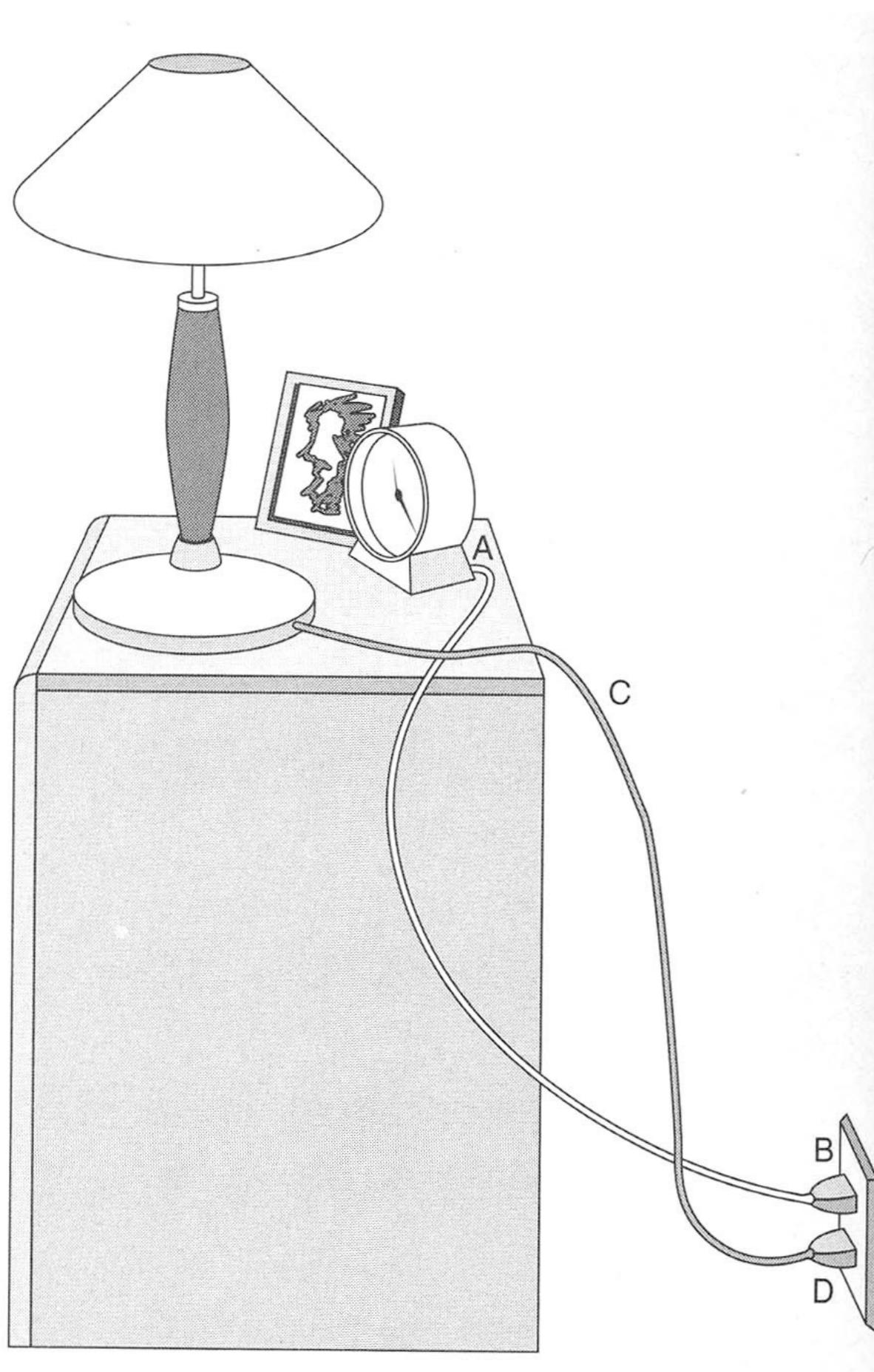


(a)



(b)

Good Continuation



<https://www.facebook.com/laurawilliamsphotos/photos/>



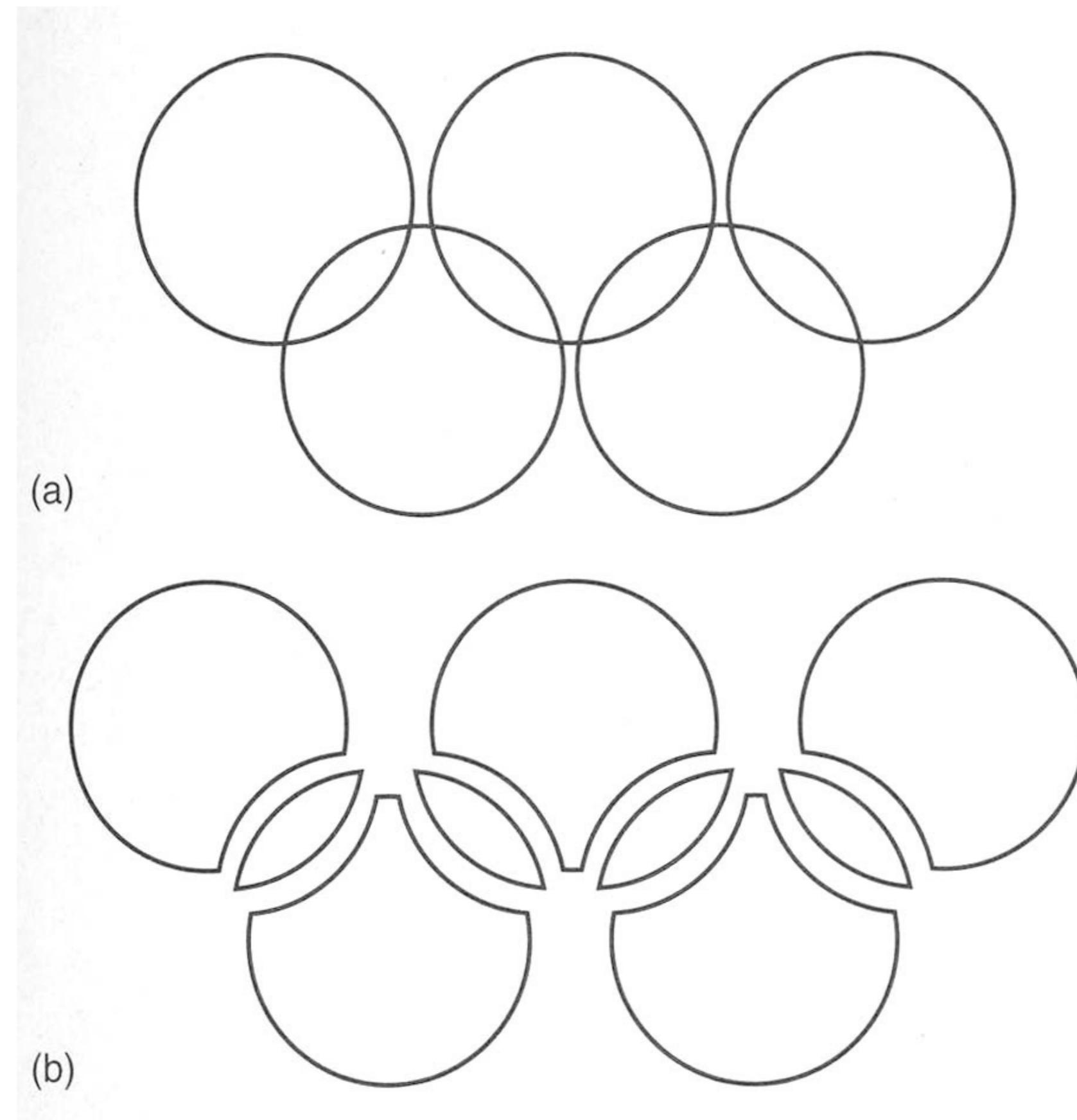
Law of Prägnanz ("Good Form" or "Good Gestalt")

Wertheimer tried to unify the Gestalt Laws under a general principle: the law of Prägnanz.

The perceptual field will take on the simplest and most encompassing structure permitted by the given conditions.

Similar to other processes in physical systems: try to achieve the maximum stability (homogeneity, simplicity) with the minimum expenditure of energy.

Law of Prägnanz ("Good Form" or "Good Gestalt")



Criticism of the Gestalt Theory of Perception

The physiological theory of the gestaltists has fallen by the wayside, leaving us with a set of descriptive principles, but without a model of perceptual processing. Indeed, some of their "laws" of perceptual organisation today sound vague and inadequate. What is meant by a "good" or "simple" shape, for example?

- Bruce, Green & Georgeson (1996)

Laws of Perception?

Gestalt Laws are certainly not like the laws of natural science in physics.

Furthermore, often unclear whether the “laws” are really independent (“similarity” vs. “good form” vs. “good continuation” ...), and often they are not based on experiments but on “demos” and (more or less) “convincing stimuli” only.

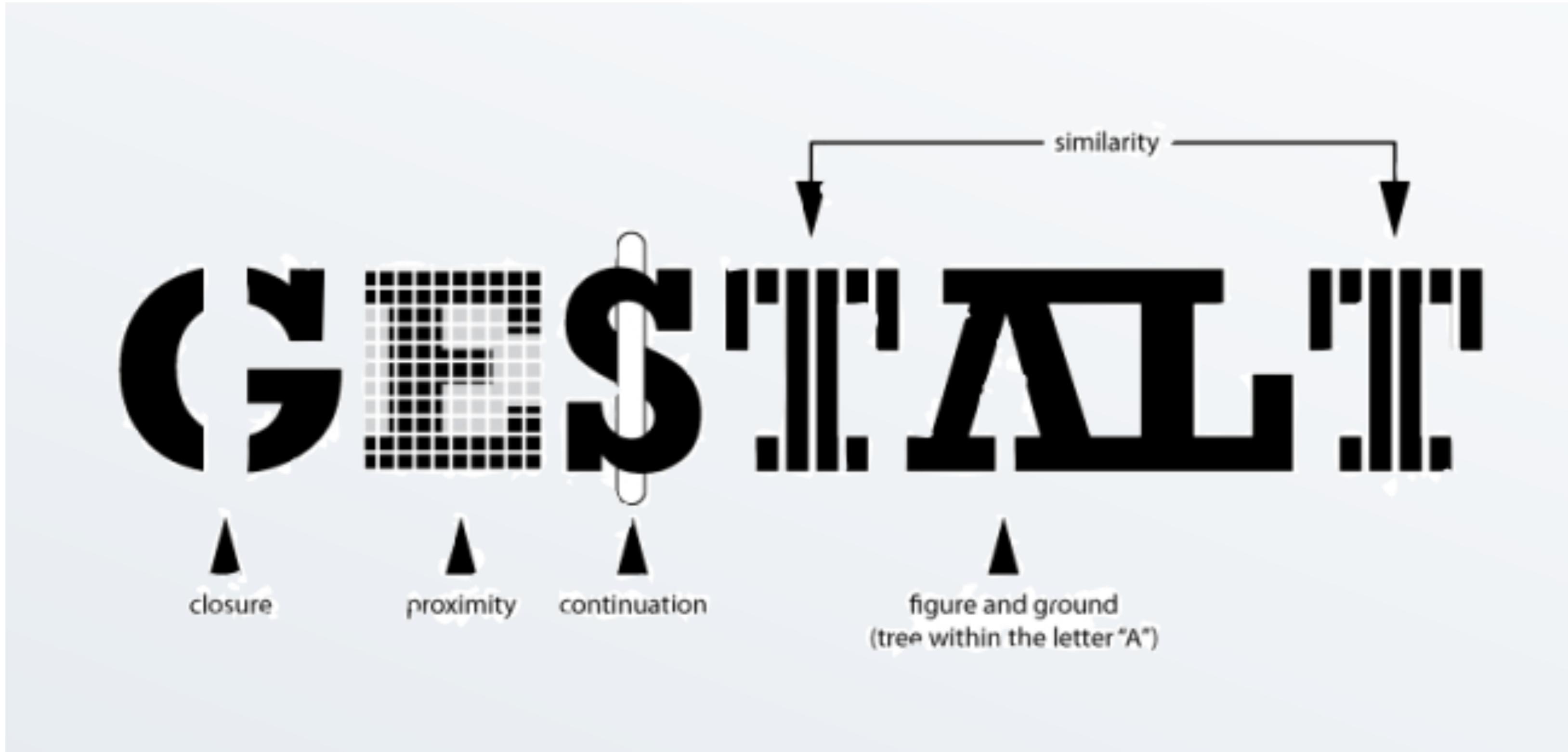
Even worse, they typically do not allow the prediction of behaviour in novel (unseen) circumstances—the danger of circularity looms large over much of the Gestalt literature!

Nonetheless, Gestalt phenomena and the radical stance adopted by the early (Berlin) Gestalt School has had a profound impact on some researchers in visual perception, and the influence continues to this day.

Much of contemporary Gestalt-inspired research tries to find mechanistic explanations (algorithms) of Gestalt laws in terms of image features etc., e.g. the dot lattice research. While this is laudable and helps to counter the circularity criticism, it is “atomistic” in building up the Gestalt from smaller elements: The Berlin Gestalt School founders would presumably not have liked this

A second modern approach tries to relate Gestalt laws to natural image statistics, e.g. whether the Law of Good Continuation may be related to contour statistics in photographs of natural scenes.

Gestalt principles applied to design



<http://michaeldain.com/2014/08/gestalt-theory-for-interface-designers-1-figureground/>

<http://michaeldain.com/2014/10/learn-gestalt-theory-in-5-minutes-or-less/>

The End

Felix Wichmann



Neural Information Processing Group
Eberhard Karls Universität Tübingen