MTTTS17 Dimensionality reduction and visualization

Spring 2020

Lecture 5: Human perception

Jaakko Peltonen

jaakko dot peltonen at tuni dot fi

Human perception (part 1)

Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

Human perception (part 1)

Most strikingly, a recent paper showed only an 11% slowing when people read words with reordered internal letters:

WILEY InterScience®

Home / Psychology / Psychology (general)

Psychological Science — RESEARCH, THEORY, & APPLICATION IN PSYCHOLOGY AND RELATED SCIENCES — A JOURNAL OF THE ASSOCIATION FOR PSYCHOLOGICAL SCIENCE aps

Psychological Science

Volume 17 Issue 3, Pages 192 - 193

Published Online: 23 Feb 2006

- Get Sample Copy
- Recommend to Your Librarian
- Save journal to My Profile
- Set E-Mail Alert
- Email this page
- Print this page
- RSS web feed (What is RSS?)

Save Article to My Profile | Download Citation | Request Permissions

Abstract | References | Full Text: HTML, PDF (Size: 44K) | Related Articles | Citation Tracking

Short Report

Raeding Wrods With Jubmled Lettres

There Is a Cost

Keith Rayner [1], Sarah J. White [2], Rebecca L. Johnson [1] and Simon P. Liversedge [2]

[1] University of Massachusetts, Amherst, and [2] University of Durham, Durham, United Kingdom

Address correspondence to Keith Rayner, Department of Psychology, University of Massachusetts, Amherst, MA 01003, e-mail: rayner@psych.umass.edu.

Human perception (part 1)

Human perception (part 1)

Human perception (part 1)

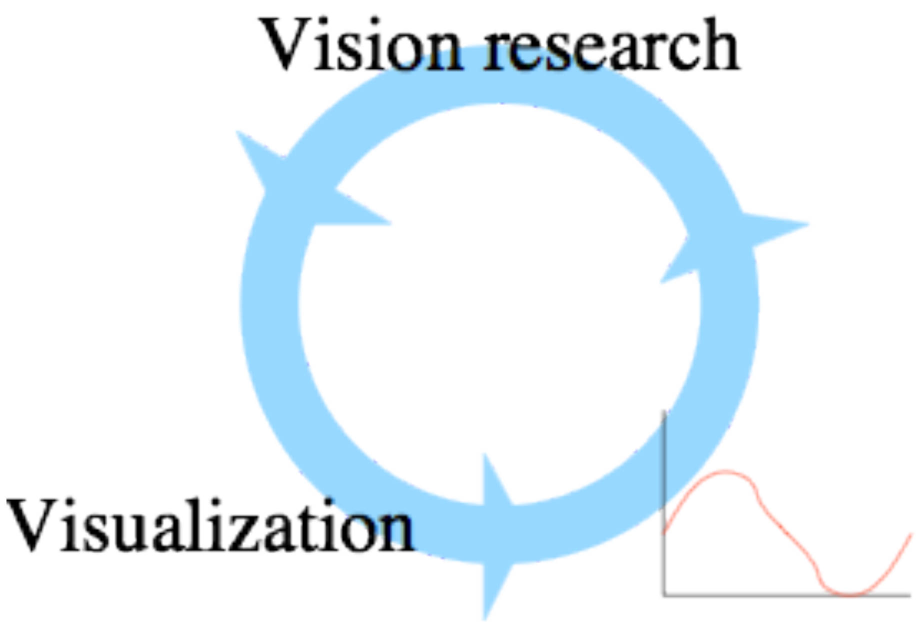
Human perception (part 1)

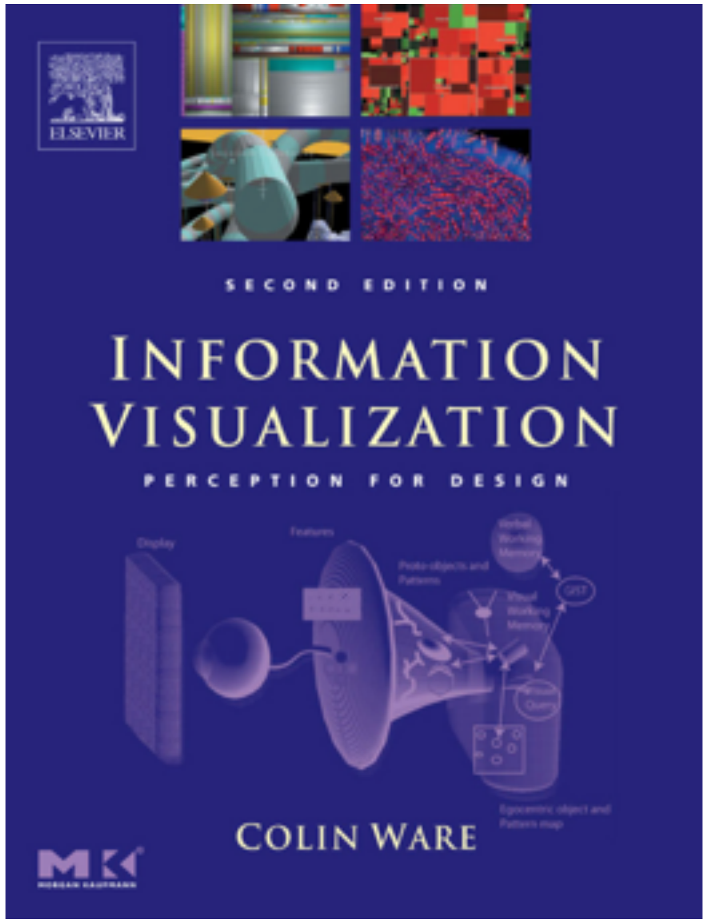
Human perception (part 1)

Human perception (part 1)

Human perception and visualization

- Visualization is young as a science
- The conceptual framework of the science of visualization is based on the human perception
- If care is not taken bad designs may be standardized





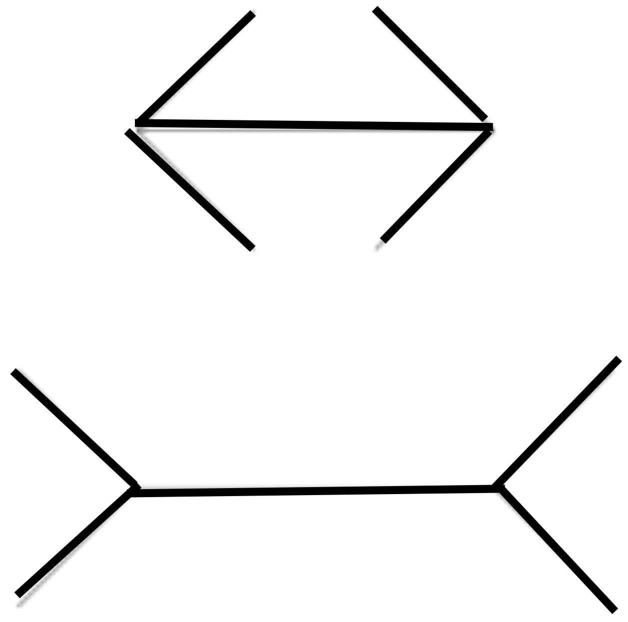
Gibson’s affordance theory

- “We perceive in order to operate.”
- We do not perceive elementary units (points of light, . . .)
- We perceive possibilities for action, or *affordances* (where to step etc.).
- Gibson claimed that we perceive affordances directly, by the visual system as a whole. Visual system *resonates* to respond to properties of the environment.
- Influental theory, but it is not to be taken too literally, unless we want to dump results of years of vision research (e.g. what we know of colors).

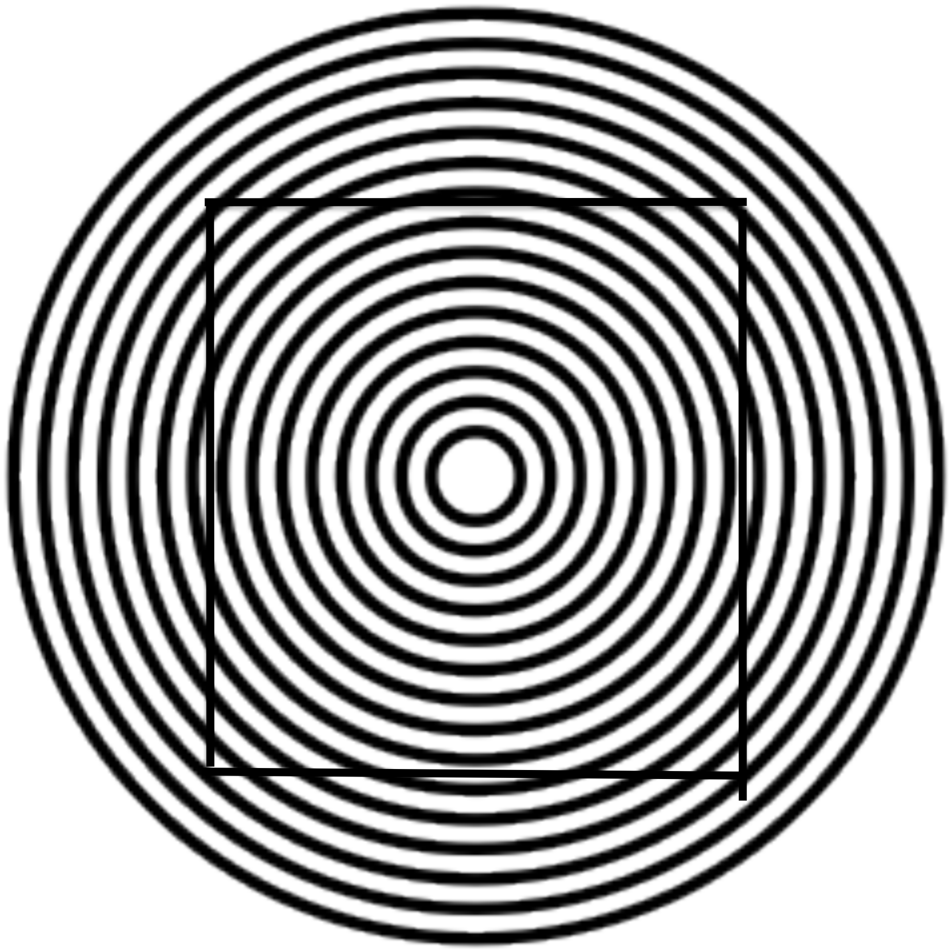
Sensory and arbitrary symbols

- Sensory symbols
  - Understandable without learning
  - Processing is hard-wired and fast
  - Resistant to instructional bias (results of millions of years of evolution)
  - Cross-cultural
- Arbitrary symbols
  - Hard to learn and and easy to forget (except when overlearned)
  - Formally powerful
  - Capable of rapid change
  - Culture-specific

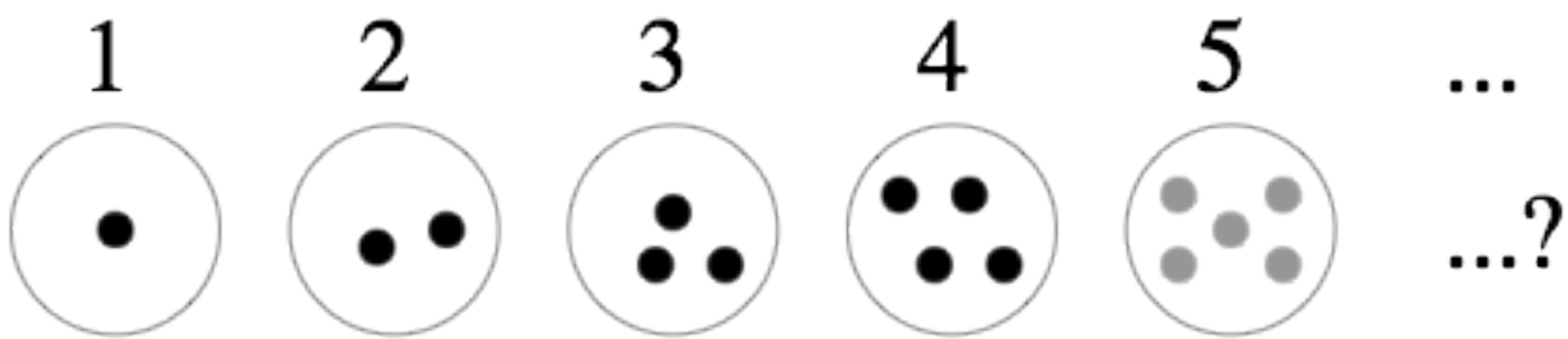
Sensory symbols: resistance to instructional bias



Müller-Lyer illusion



Arbitrary symbols



Arbitrary symbols are harder to learn, but easier to extend.

(Could you tell the difference between 10000 dots and 9999 dots?)

- Distinction to sensory and arbitrary symbols is difficult
- If all symbols were arbitrary, the problem of visualization would reduce to establishing a consistent notion (*standardization*)

Stages of perceptual processing

1. Parallel processing to extract low-level properties of the visual scene

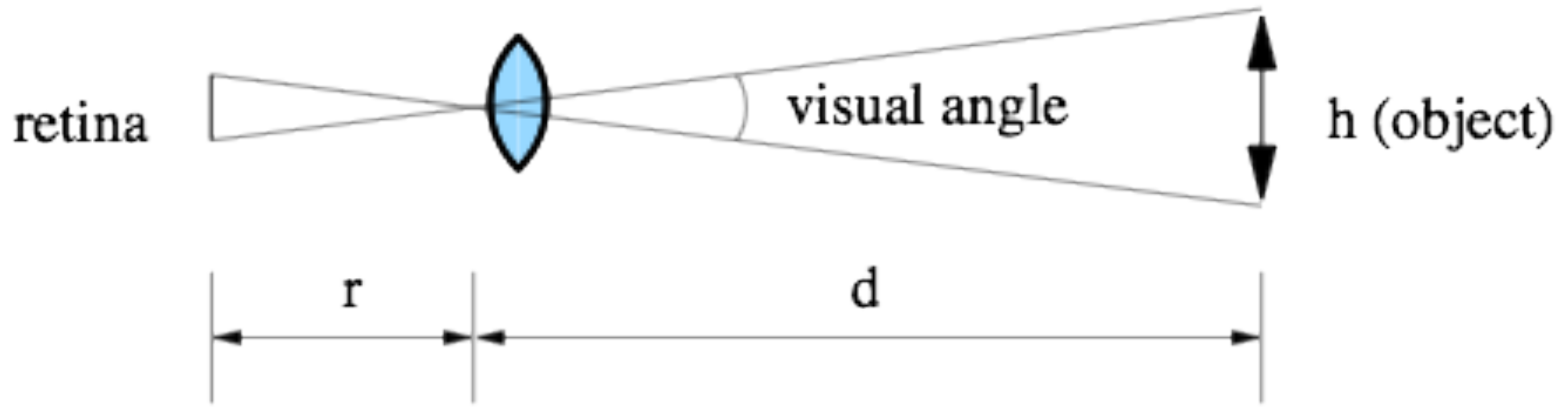
- rapid parallel processing
- extraction of features, orientation, color, texture, and movement patterns
- iconic store
- bottom-up, data driven processing

2. Pattern perception

- slow serial processing
- involves both working memory and long-term memory
- arbitrary symbols relevant
- different pathways for object recognition and visually guided motion

3. Sequential goal-driven processing

Optics

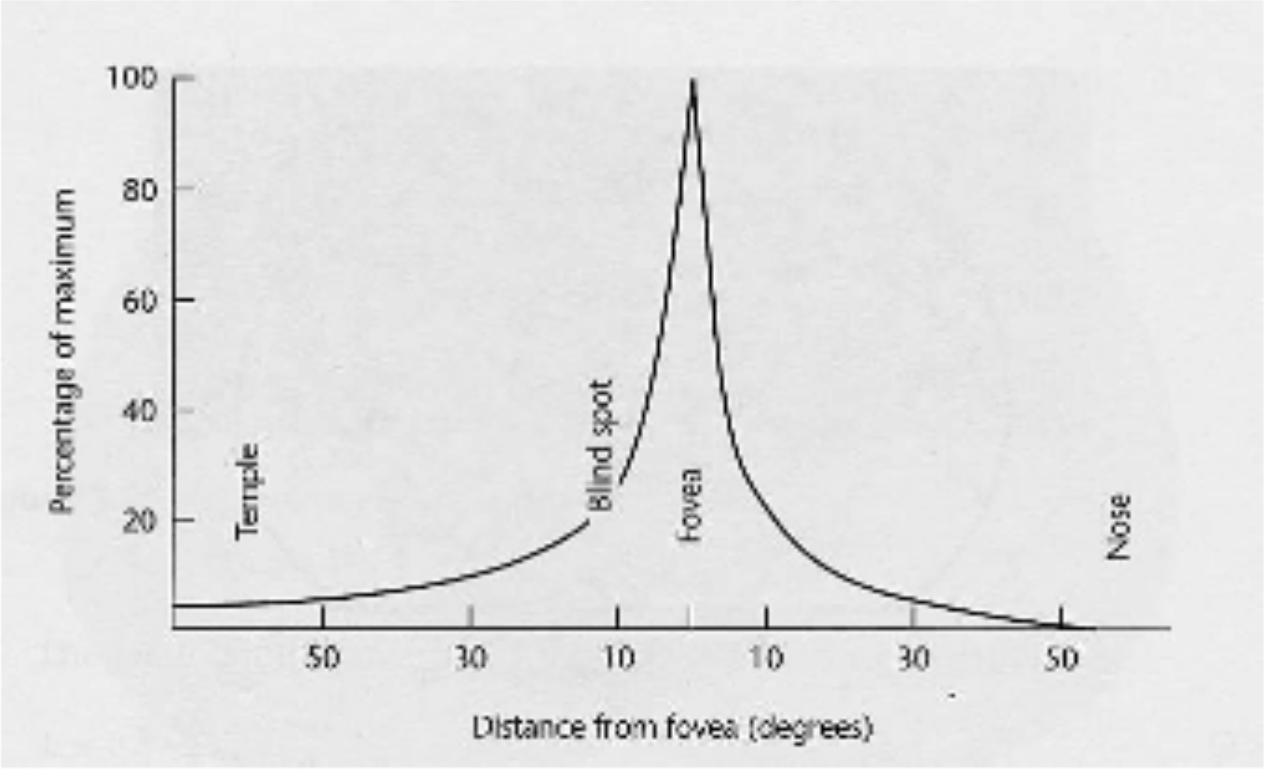


$$\frac{1}{f} = \frac{1}{d} + \frac{1}{r}$$

- The human lens system has a *focal length* of about $r = 0.017$ m.
- The power of the lens is $1/f = 59m^{-1}$ diopters (corresponding to $d = \infty$).
- Young children are capable of adjusting the lens 12 diopters or more, which means that they can focus to about $d = 0.08$ m.
- Flexibility drops by age at rate of about diopter per five years (the lens of people of the age over 60 is almost completely rigid!).

Acuity

- Visual acuities are measurements of our ability to see detail
- Acuity is at maximum at the center of the fovea
- Peripheral acuity drops rapidly; at 10° the acuity is 20 % of the maximum
- The point acuity at the center of the fovea (about 1 minute of arc) corresponds to the separation of receptors there

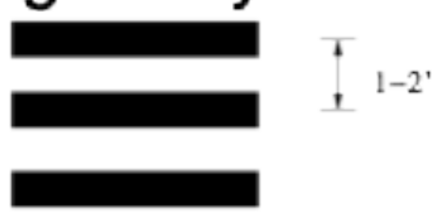


[W 2.18].

Simple acuities

- Simple acuities are limited by spacing of receptors at the center of the fovea (about one minute of arc)
- Point acuity describes the ability to resolve two distinct point targets:

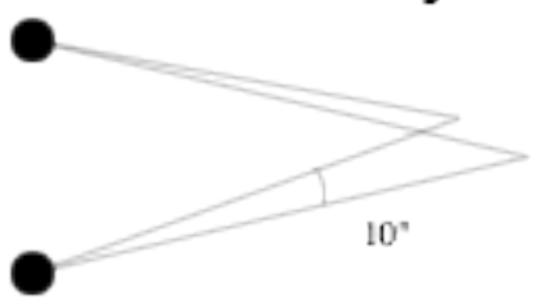
- Grating acuity describes ability to distinguish grated pattern from uniform gray:



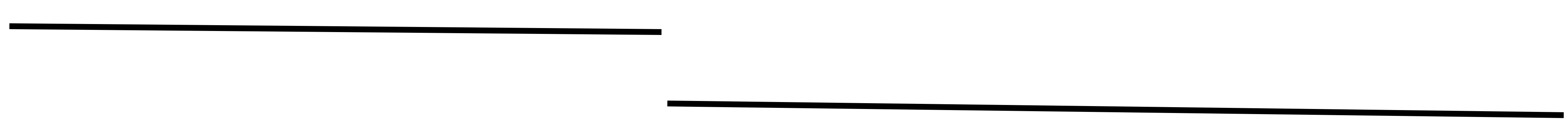
- Letter acuity describes ability to resolve letters:

Superacuities

- Superacuity is ability to achieve better resolution by integrating information over space (or time!)
- Stereo acuity is describes the ability to resolve objects in depth:

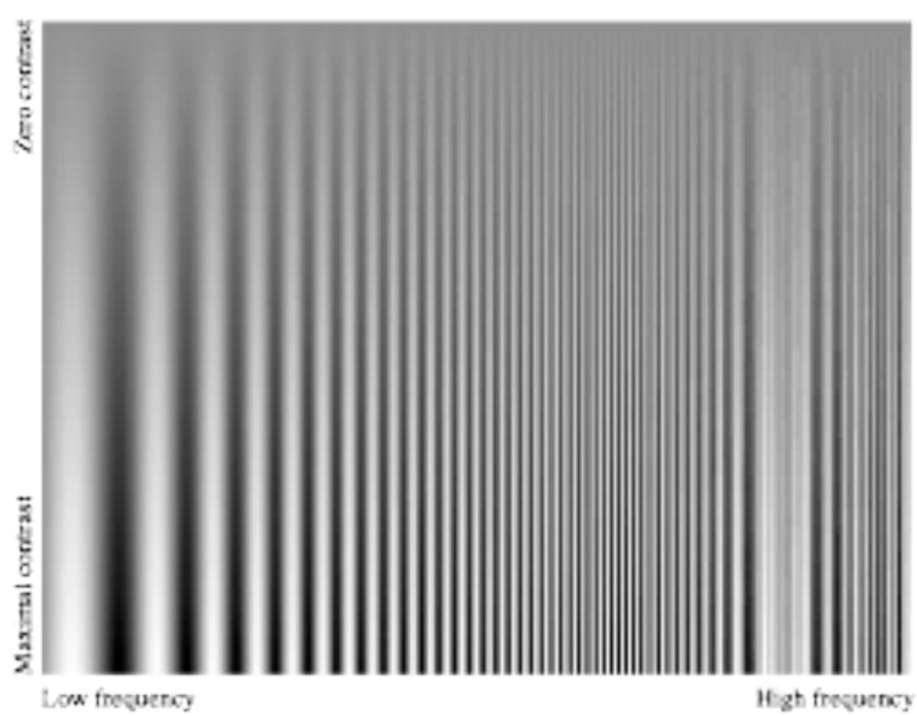


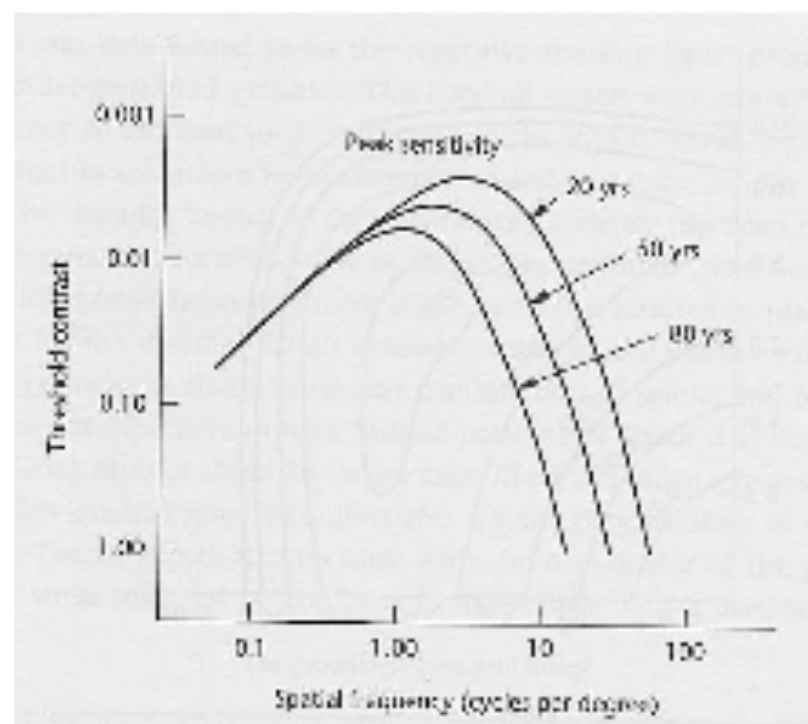
- Vernier acuity describes ability to see if two line segments are colinear:



Contrast sensitivity

- Contrast sensitivity is lowest at high frequencies (60 cycles per degree for young people)
- Sensitivity also falls off at low frequencies. Sensitivity is highest at about 2 or 3 cycles per degree.
- Contrast sensitivity falls of by age



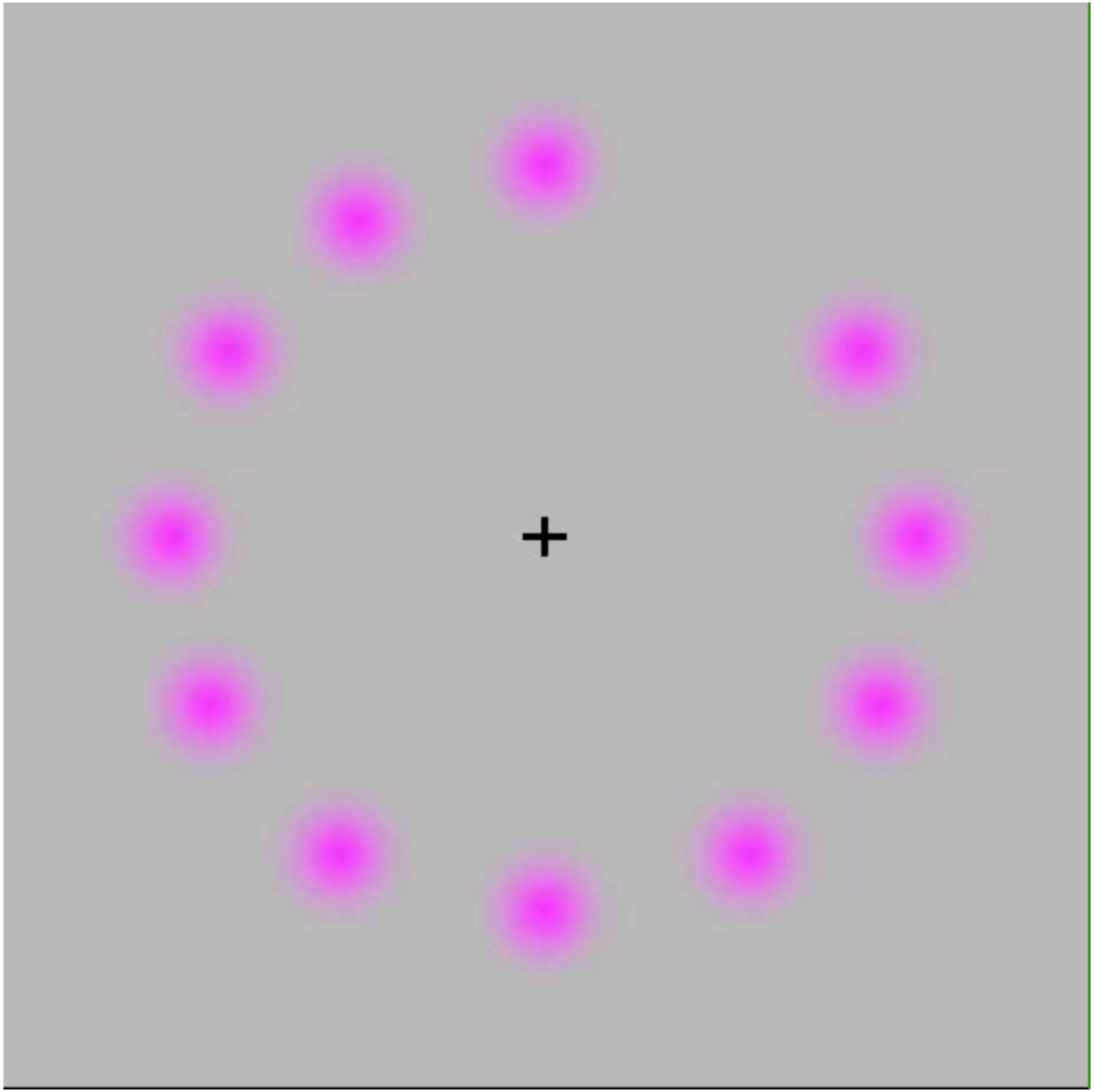


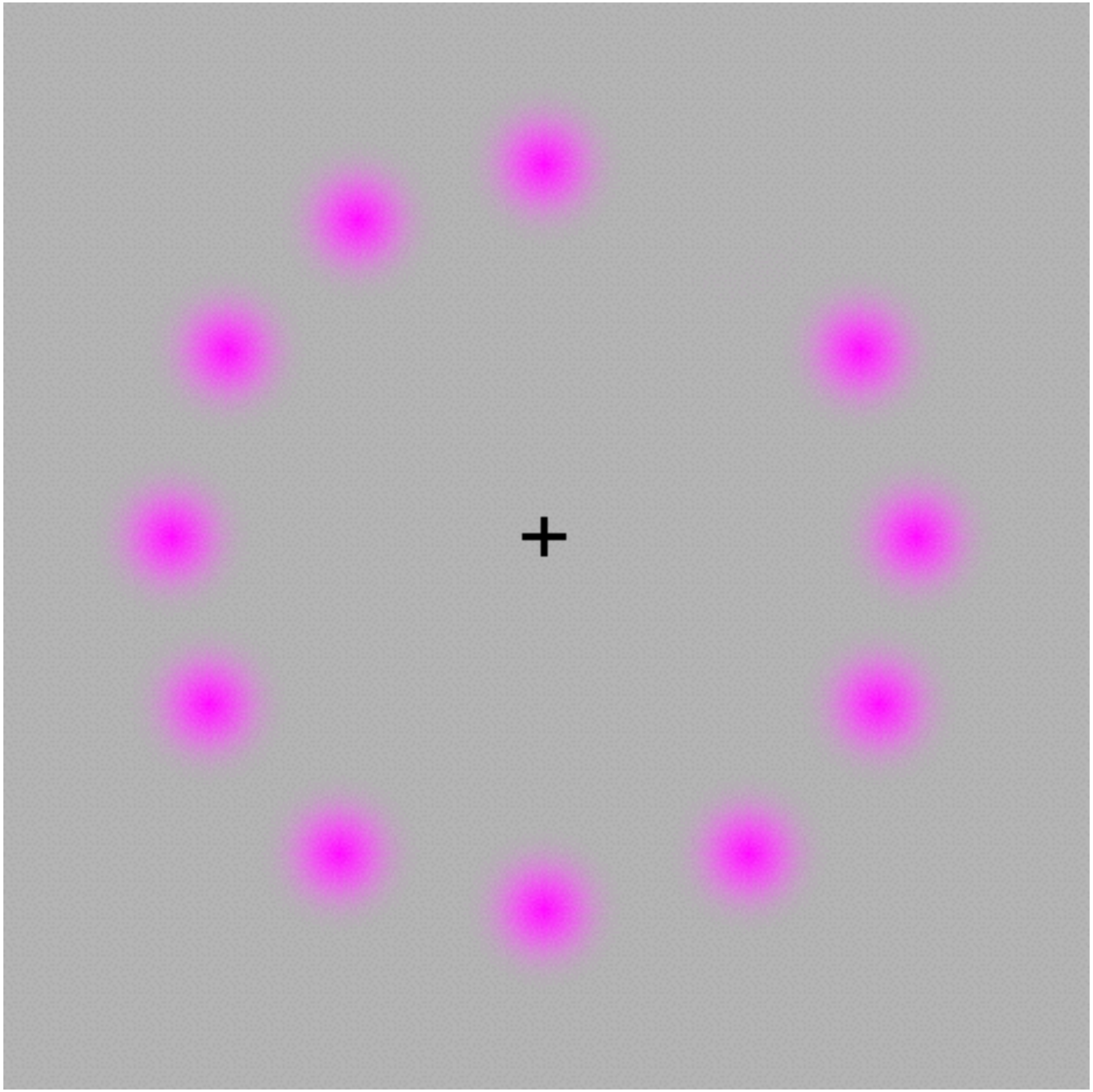
[W 2.22].

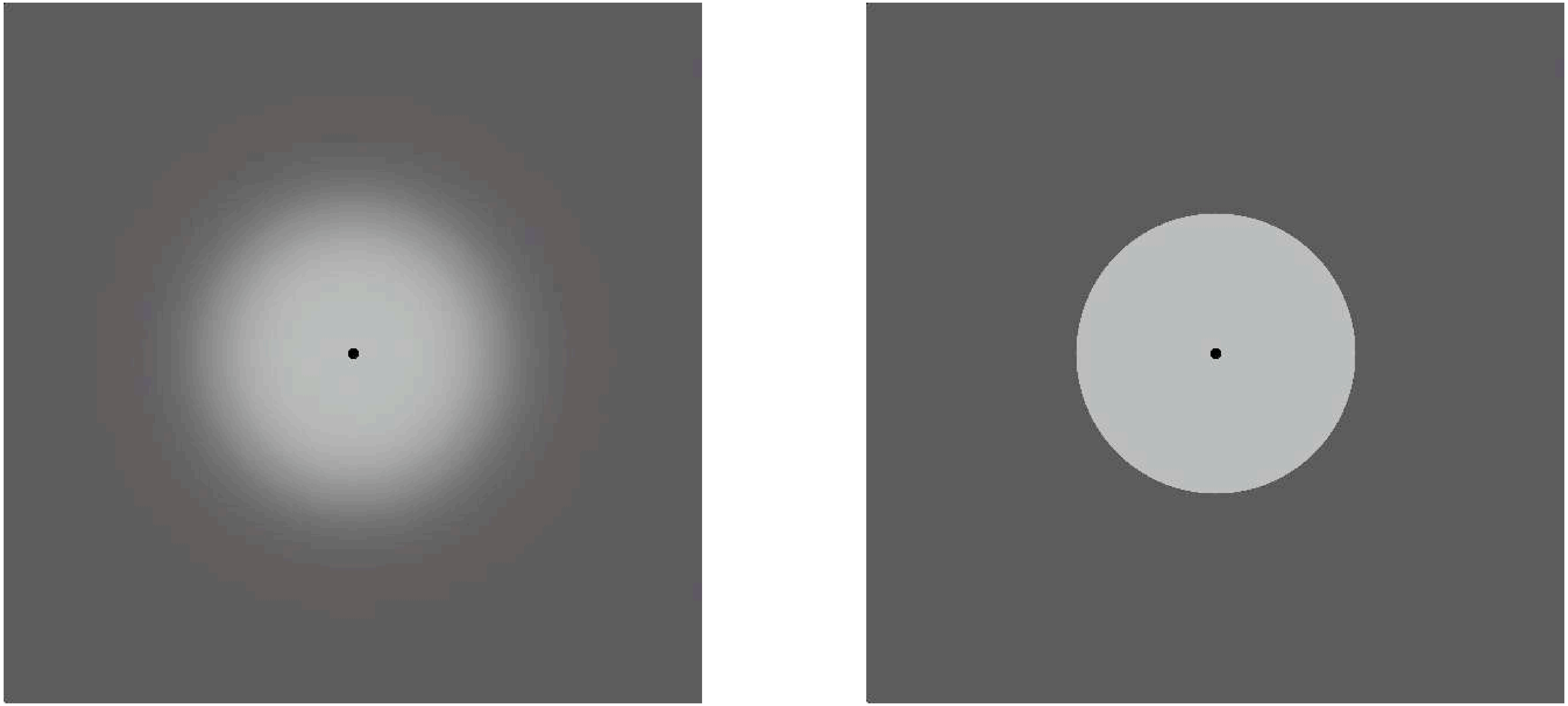
Properties of the retina

- The retina is like the film in camera. However, there are important differences. Roughly:
  - Retina must be able to work at varying light levels ➜ only relative differences in luminance are perceived
  - All constant stimuli are ignored
- For example, *contrast* is the *perception* of difference in the brightness
- *Perceptual contrast* and *physical contrast* (actual difference in the luminance) does not necessarily correspond to each other

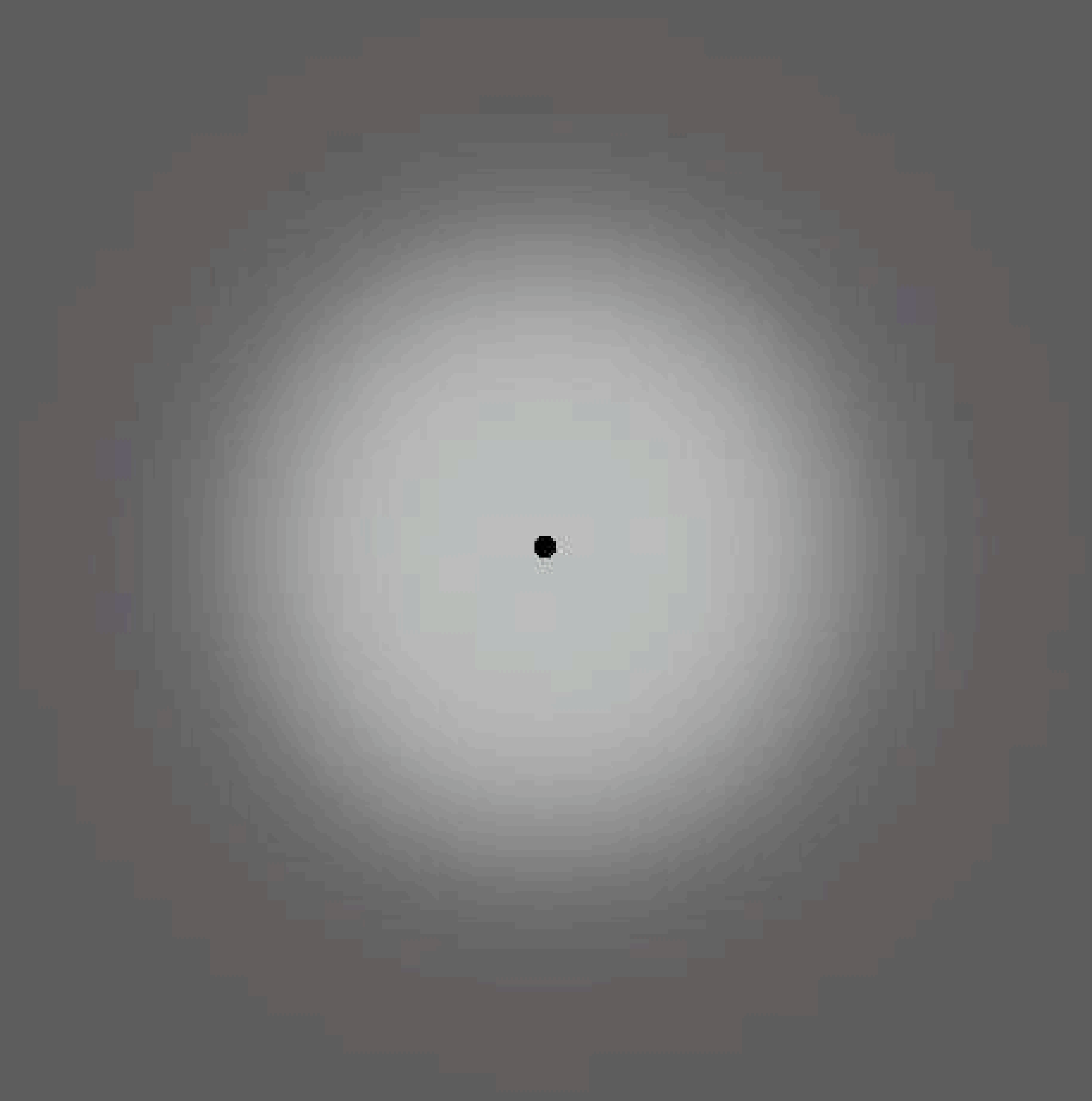
- Close one eye. Follow the rotating pink dot with your eye for at least 30 seconds.
- Keep the other eye closed. Now, keep your eye fixed on the black cross (+) at the center of the picture for at least 30 seconds.
- You should see at least two strange phenomena

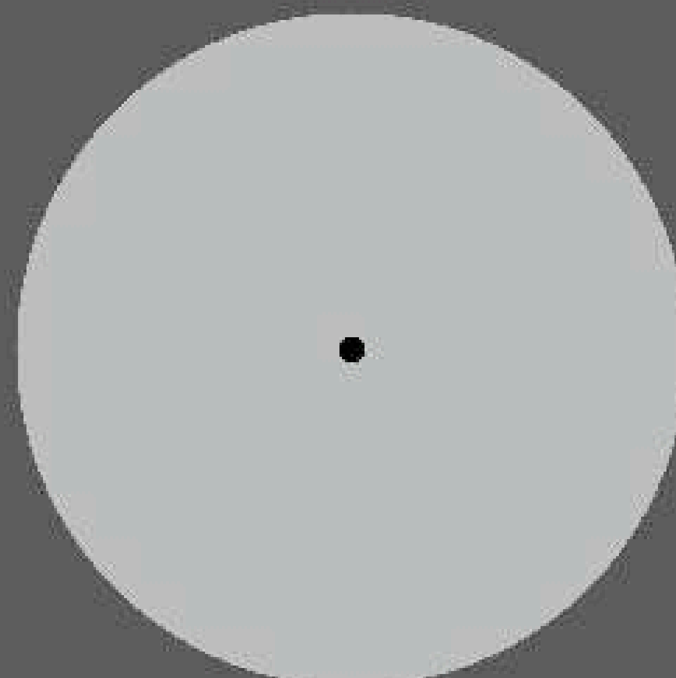




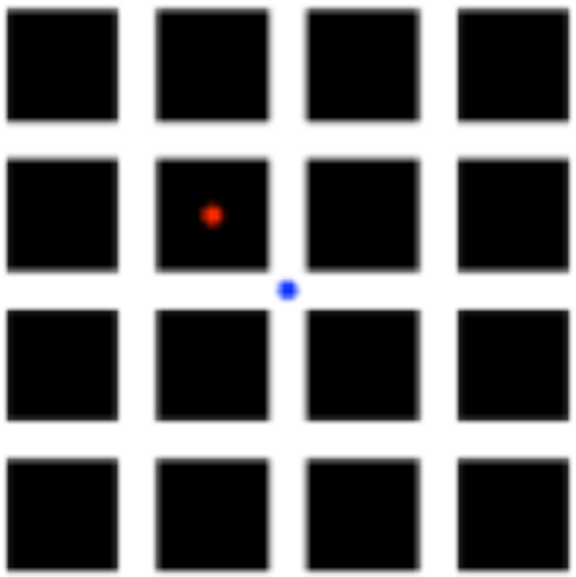


Close one eye. Look at the black dot at the center of the fuzzy disc for at least 30 seconds. Then look at the center of the sharp disc. Is there any difference?





Contrast

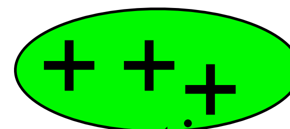


Look at the blue dot for 60 seconds. Then look at the red dot. You should see the white afterimage jiggle.

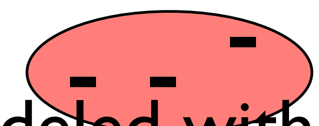
- The disc with sharp contours does not (start to) disappear because the eye jiggles involuntarily. The amount of light to the receptors near to the contour is thus constantly changing.
- Jiggling of the fuzzy contours induces only slight changes to the receptors.
- Summary: retina responds poorly, if at all, to constant stimulus.

Difference of Gaussians model

- Retinal ganglion cells are organized with circular receptive fields
- When light falls at the center of receptive field it emits pulses at increased rate (*excitation)*

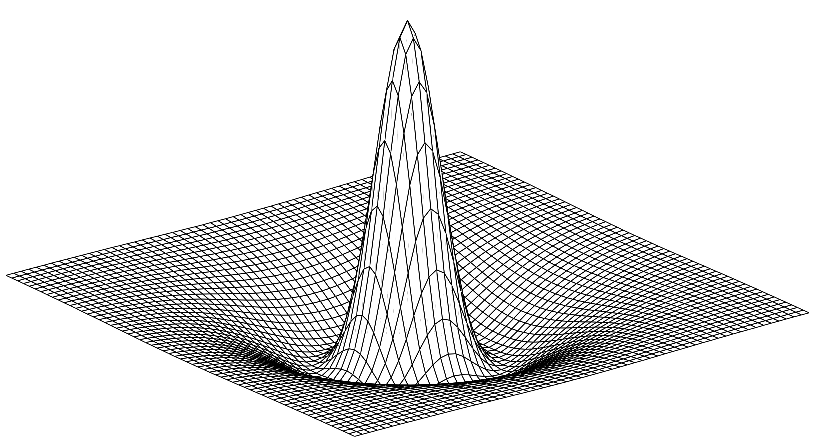


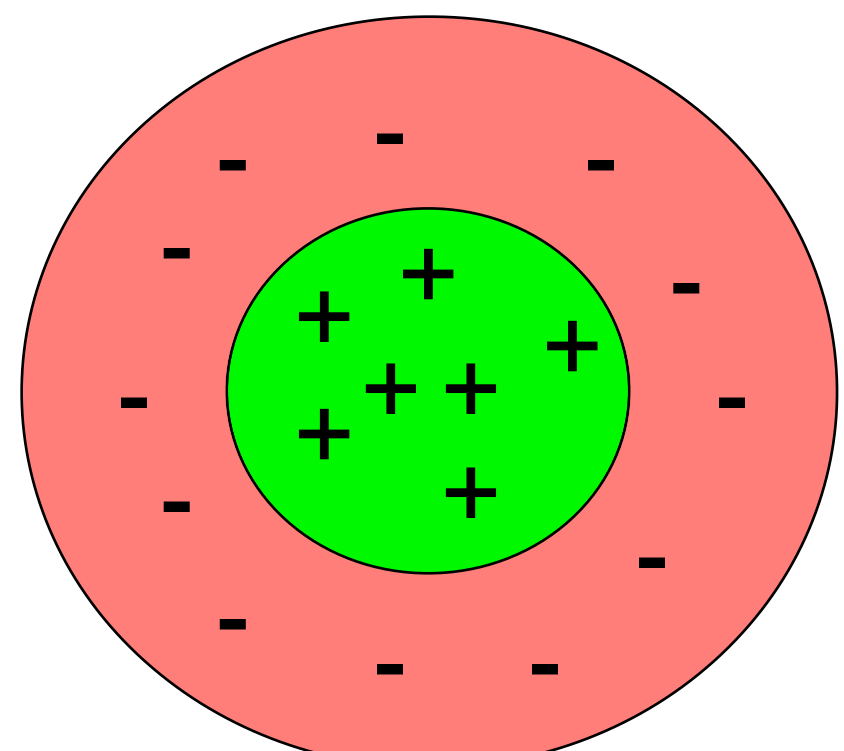
- When light falls off center of receptive field it emits pulses at lower rate (lateral *inhibition)*



- The receptive fields can be modeled with *Difference of Gaussians* (DOG) model

$$\text{Response} = K_e e^{-\left(\frac{2r}{a}\right)^2} - K_i e^{-\left(\frac{2r}{b}\right)^2}$$



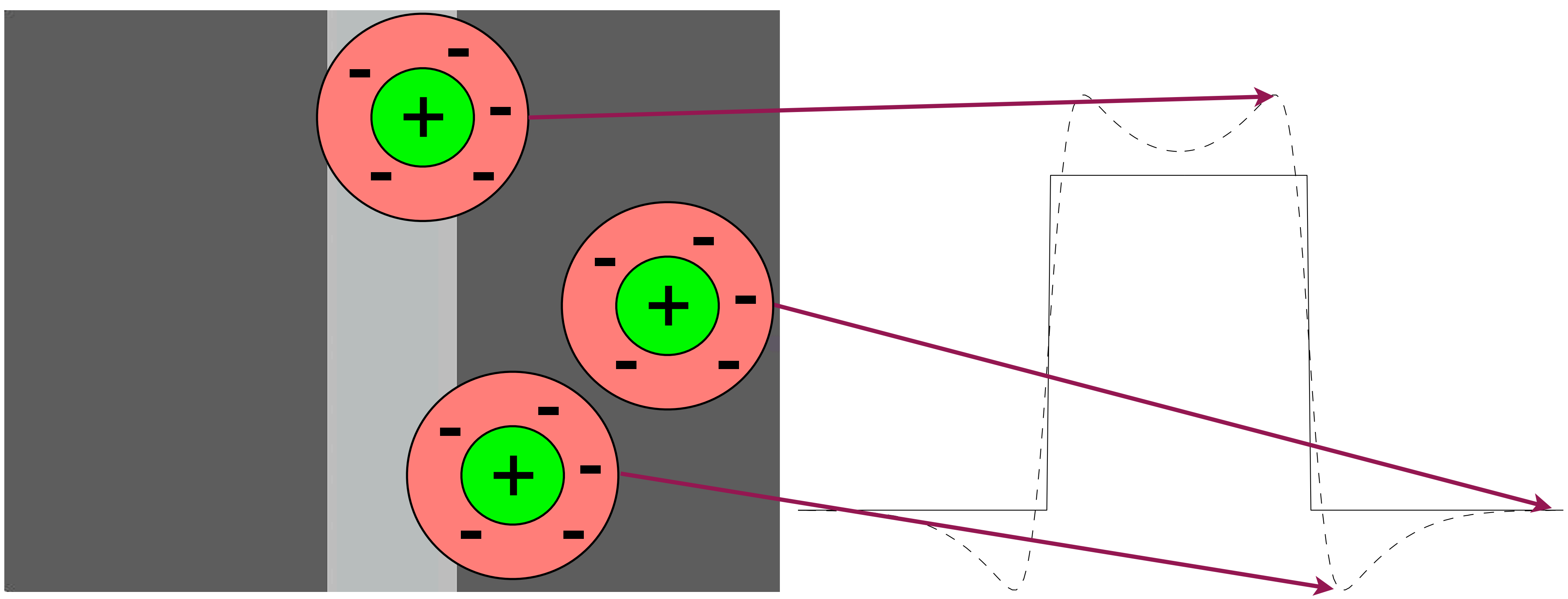


Difference of Gaussians model

- The DOG model can be used to explain the difference between physical *luminance* and perceived *brightness*
- Discontinuous lightness profiles generate dark and light bands near the discontinuities (*Chevreul illusion*)
- *Mach bands* appear if there are discontinuities in the first derivative of the lightness profile
- *Simultaneous brightness contrast:* A gray patch placed on a dark background looks brighter than the same patch on a light background

Discontinuous lightness profiles

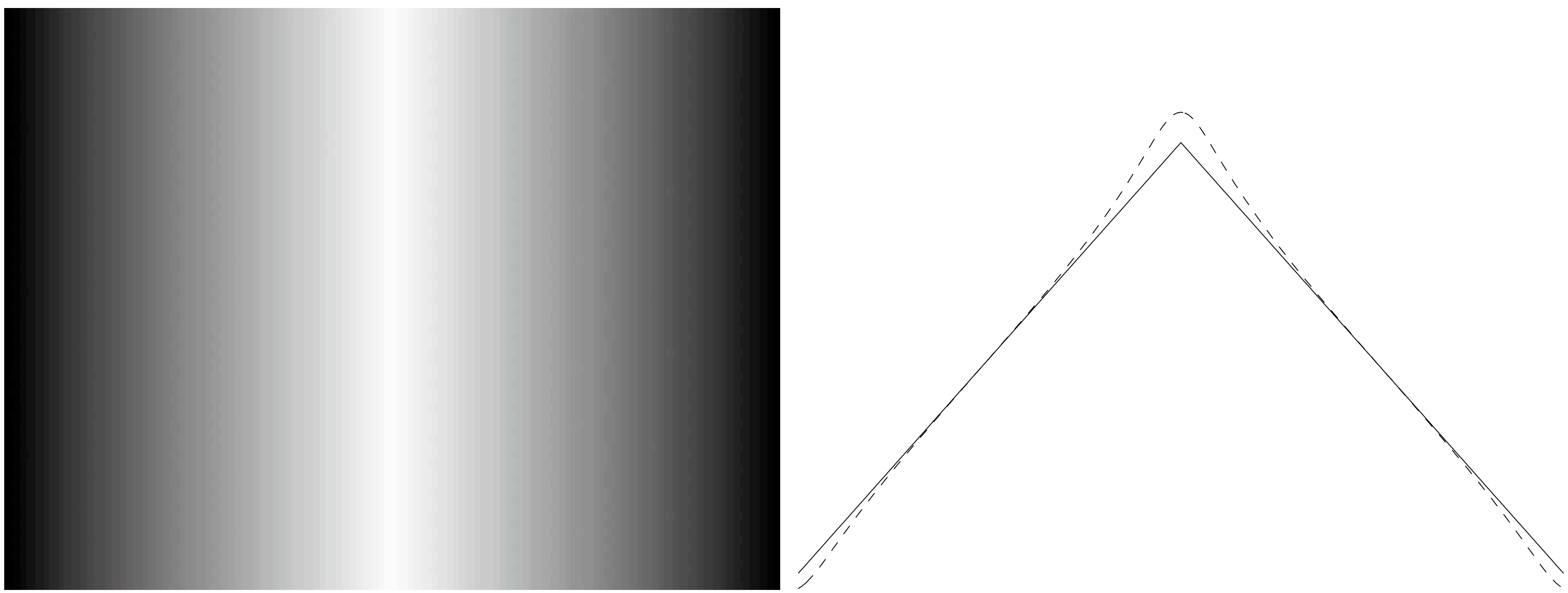
Discontinuous lightness profiles generate dark and light bands near the discontinuities (*Chevreul illusion*)



Solid lines describe (physical) *luminance* and dashed lines (perceived) *brightness*.

Mach bands

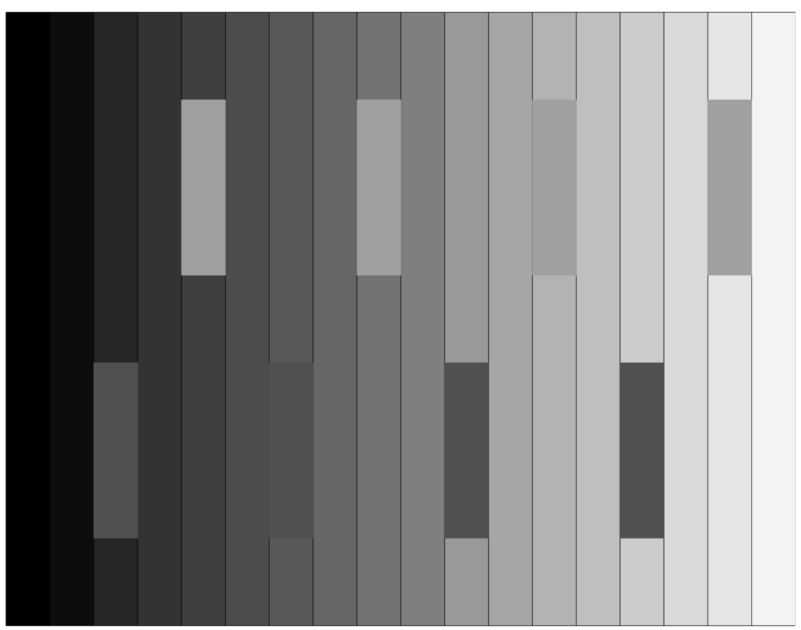
Mach bands appear at the discontinuities of the first derivative of the continuous brightness profile.

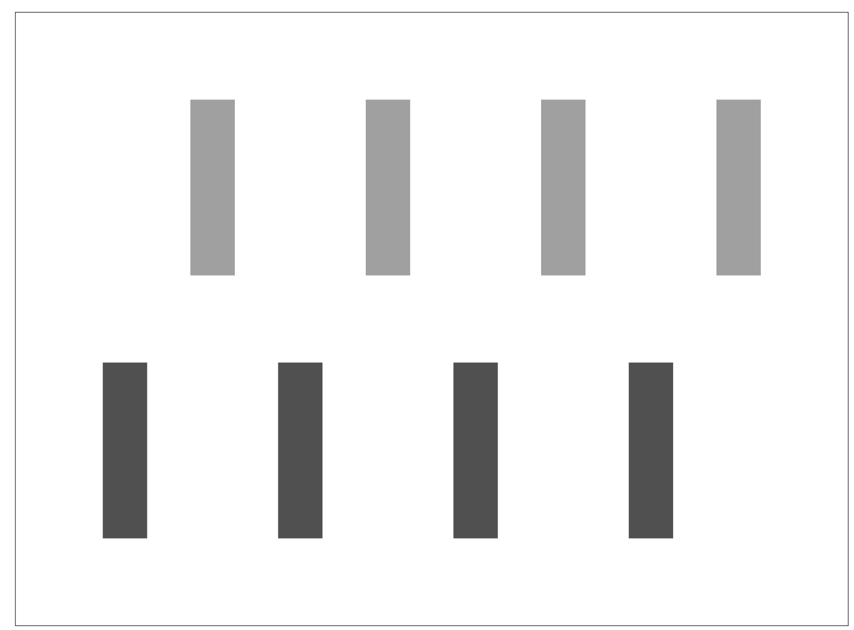


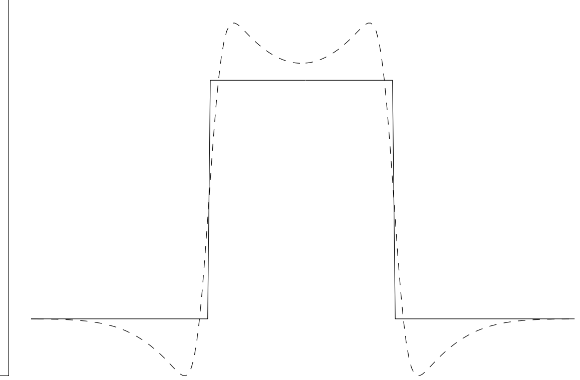
Solid lines describe physical *luminance* and dashed lines perceived *brightness*.

Simultaneous brightness contrast

A gray patch placed on a dark background looks brighter than the same patch on a light background.





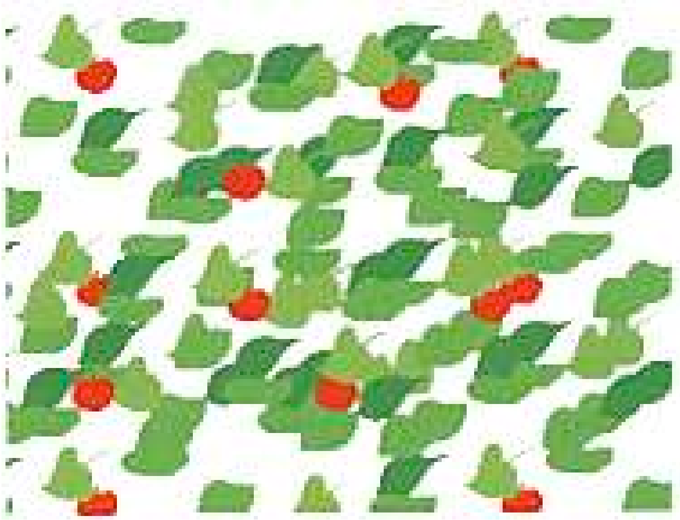


Why do we have color vision?

A modern man can go through large parts of his life without knowing that he is color blind.

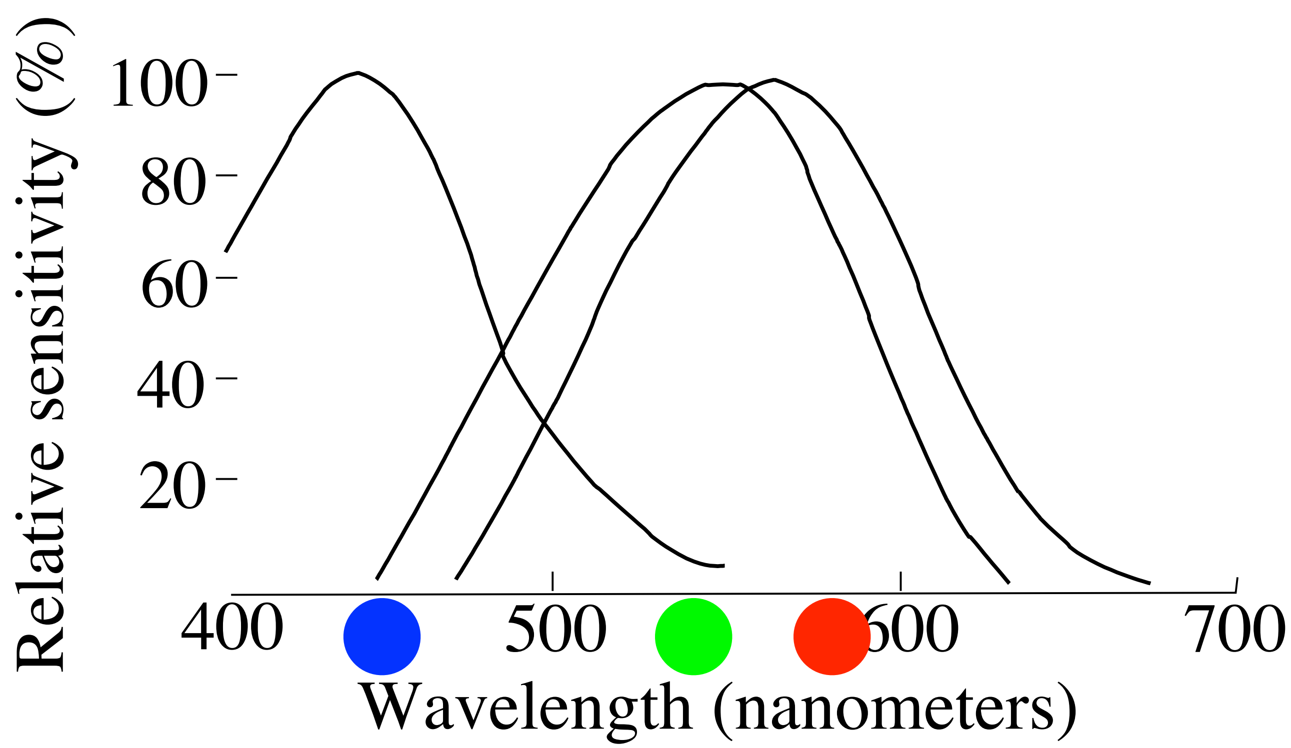
Why then do we need color vision at all?

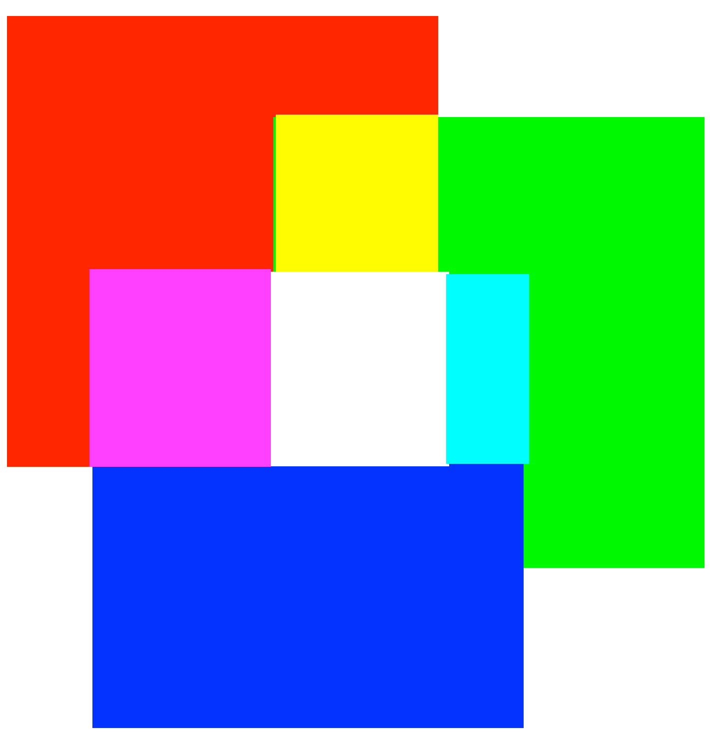
- Color breaks camouflage.
- Color tells us about material properties of objects. (Which fruit are ripe and which food has gone bad etc.)
- Color is an attribute of an object that helps us distinguish it from others.

Trichromacy theory

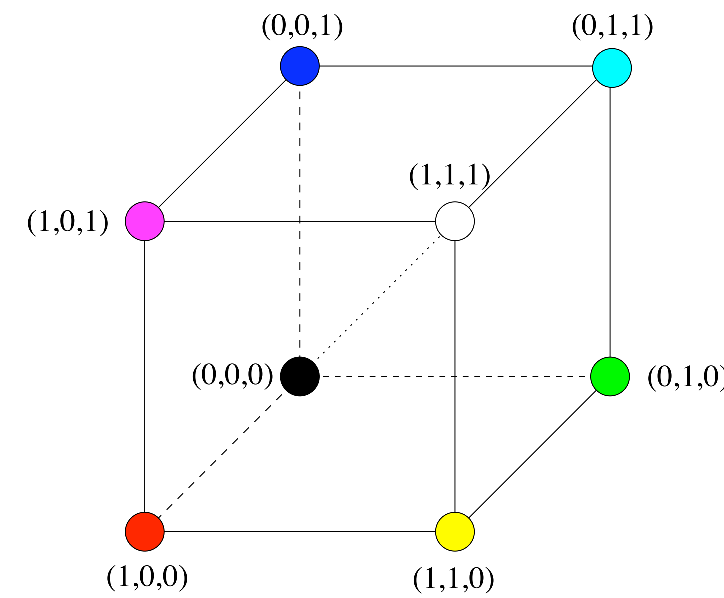
- The human eye has three different types of cone cells that respond differently to different light wavelengths (chickens have 12)
- The color perception is therefore fundamentally three dimensional
- Because of only three receptor types, any color can be represented as a mixture of three colors
- Color blindness can be seen as the collapsing of three dimensional color space into two-dimensional one





Some color spaces

- *RGB color space* based on the voltages of electron guns in a CRT-monitor. Axes Red, Green, Blue form a cube.



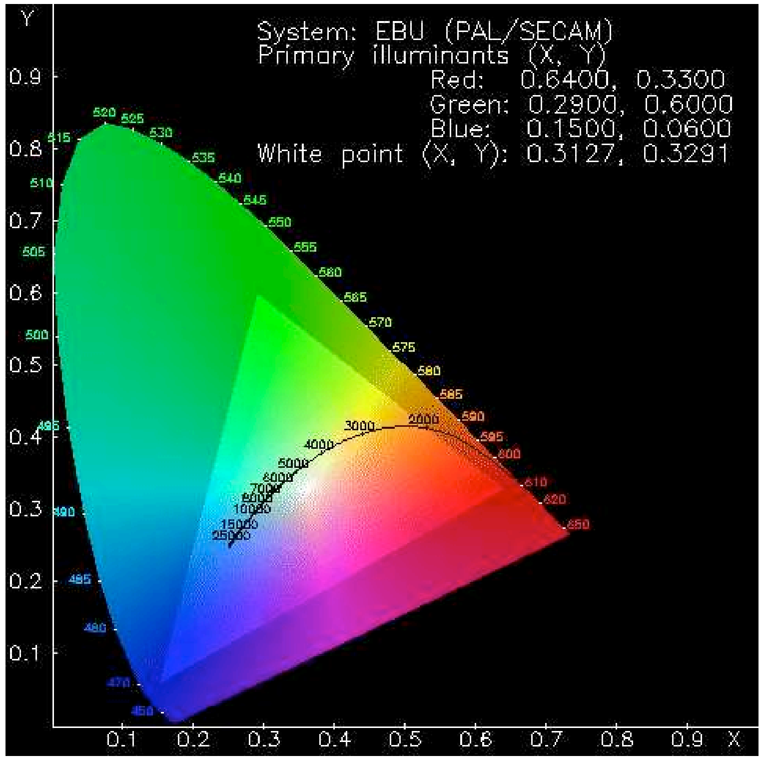
- *CMY color space* based on printer inks, C=1-R, M=1-G, Y=1-B

- *Standard observer* is a hypothetical person whose color sensitivity is held to be that of a typical person.

- *CIE XYZ tristimulus space* is a color space based on the standard observer. It uses abstract primary colors XYZ
  - Primaries X, Y, Z do not match any actual lights
  - Y corresponds to the luminance
  - All colors visible to humans can be expressed with positive values of X, Y and Z

CIE chromaticity diagram

- Chromaticity coordinates are XYZ values that are normalized to the amount of light.



$$x = X/(X+Y+Z)$$
$$y = Y/(X+Y+Z)$$
$$z = Z/(X+Y+Z)$$
$$1 = x+y+z$$

- Colors are usually specified by giving x, y and Y.

- Chromaticity diagram is the plot of the xy plane of the color space.

1. All colors on a line between 2 colored lights can be created as a mix of the 2 colors
2. Any set of three colored lights specifies a triangle. All points within the triangle can be represented as a mixture of the given lights.
3. All realizable colors fall within the *spectrum locus* (the set of chromaticity coordinates representing single wavelength colors)

Uniform color spaces

- CIELab
- CIELuv - better for specifying large color differences
- Difference between colors is given by (1 is approximately *just noticeable difference*)

$$\Delta E_{uv} = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}$$

- Tries to be perceptually uniform
- Only approximates perceptual differences: e.g. size of color patches affects perception, can't see difference in small patches (small field color blindness)

Conversion from CIE (X, Y, Z) to CIE (L^*, u^*, v^*) is defined as follows:

$$L^* = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & (\text{for } Y/Y_n > 0.008856) \\ 903.3(Y/Y_n) & (\text{otherwise}) \end{cases}, \quad (1)$$

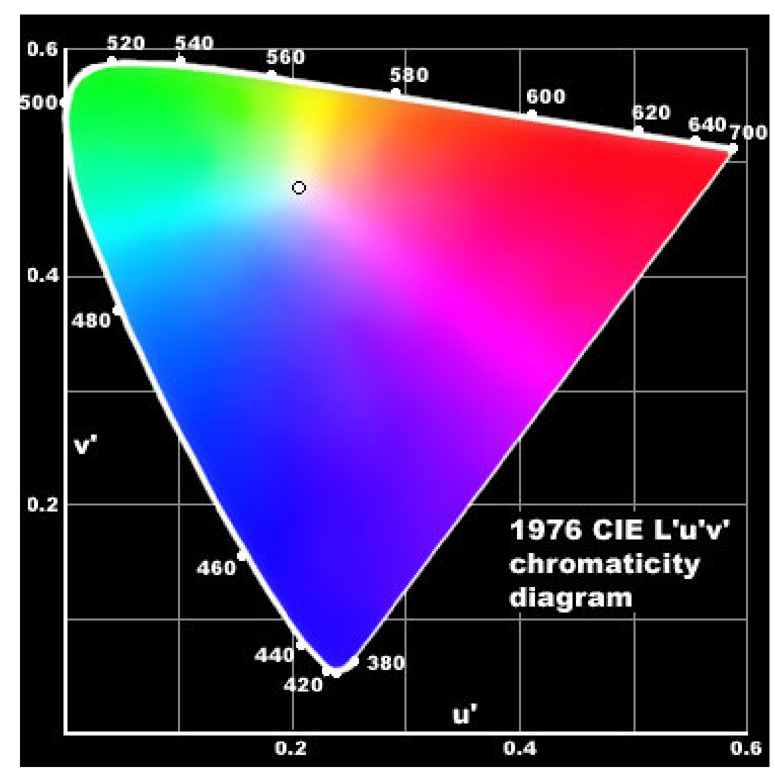
$$u' = \frac{4x}{-2x + 12y + 3}, \quad v' = \frac{9y}{-2x + 12y + 3}, \quad (2)$$

$$u^* = 13L^*(u' - u'_n), \quad v^* = 13L^*(v' - v'_n), \quad (3)$$

$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}, \quad (4)$$

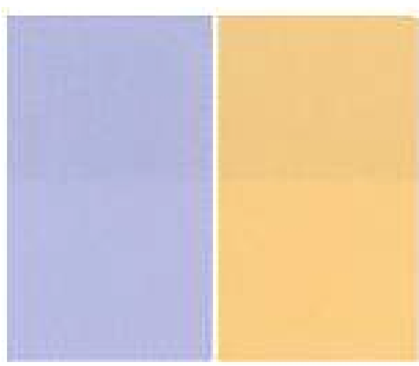
where Y_n, u'_n and v'_n are Y, u' and v' of reference white respectively. Spectrum locus, sRGB gamut and MacAdam's

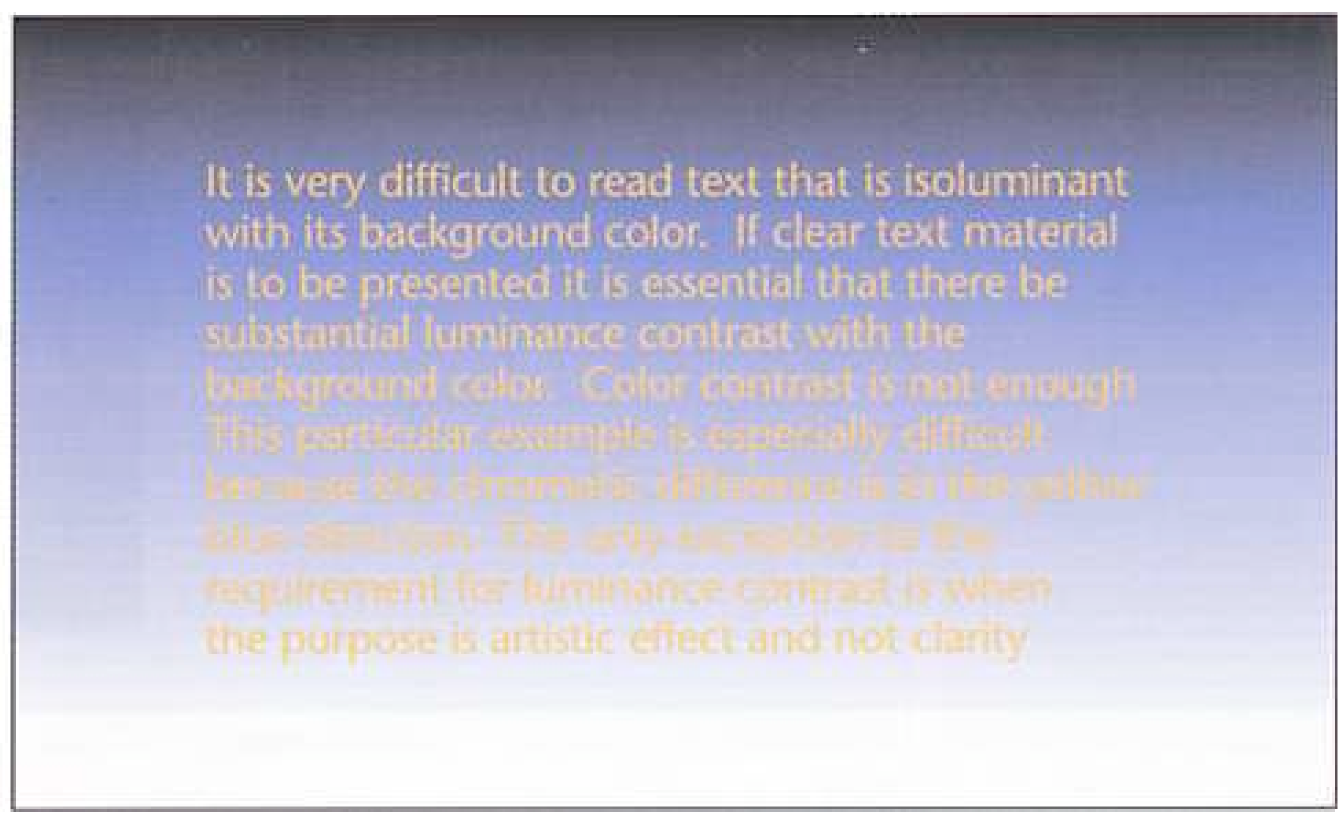
Takamura, Kobayashi 2002



Color vs. luminance

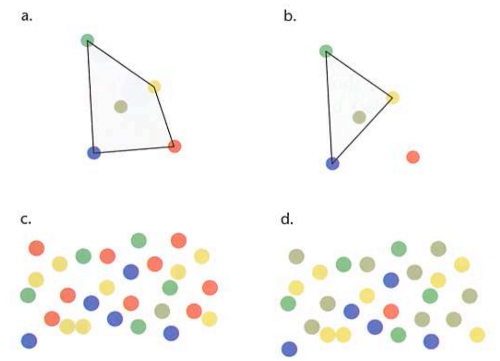
- Color channels have less spatial resolution than luminance channel
- Perception of shape or motion is due to mainly luminance channel
- Color channels are better (only?) for labeling

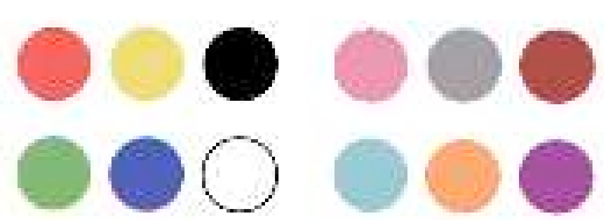




Color for labeling

- Choosing color for labels (*nominal information coding*):
  - distinctness (e.g., CIELuv, preattentivity)
  - unique hues
  - contrast with background
  - color blindness (avoid red-green distinctions)
  - number (5-10 codes can be rapidly distinguished)
  - field size
  - convention (in west red=hot,danger, blue=cold etc.)
- Recommendation (first six before the remaining six):





Color scales (sequences)

Important issues in designing color scales:

- Some common color scales are not perceived by people who are color blind.
- Perceptually orderable color scales are in general formed from the six color opponent channels. Some other color orderings that might work are cold-hot and dark-light.
- Different color scales might work with different levels of detail in the data. With high levels of details the color scale should be mostly based on luminance.
- Uniform color spaces can be used to create color scales that have perceptively constant color steps.
- If it is useful to read values back from the color scale, the scale should cycle through many colors. This will minimize color contrast effects.
- In many cases the best color scale may be a spiral in color space.
- Color category boundaries might cause miscategorization of data.

Color for Exploring Multidimensional Discrete Data

Variable 1 → *x*-axis position

Variable 2 → *y*-axis position

Variable 3 → amount of red

Variable 4 → amount of green

Variable 5 → amount of blue

Visual attention

- *Visual attention* is the process of seeking our visual stimuli and then focusing on them
- We do not perceive much if we do not have at least some expectation or need to see it
- We perceive "*visual objects*" (next week's topic), not "light pixels" (NB: Gibson's theories of last week!)
- *Searchlight model of attention: useful field of view* (UFOV) is the area where we can rapidly take in information (1-15°, size depends e.g. on the target density)
- As cognitive load goes up, UFOV shrinks (*tunnel vision*)
- Eye movements and visual attention are somehow related
- Attention can be attracted with motion, salient features (maybe later lectures)

Eye movements

- The eye moves according to three basic strategies:
  - *Saccadic movements*. Eye movements consist of *fixations* (duration 0.2-0.6 s), during which eye is relatively stable. Eye moves from fixation to fixation with *saccades* (duration 0.02-0.1 s, velocities up to 900°/s). Saccadic movements are pre-programmed (*ballistic*). We are practically blind during the saccade (*saccadic suppression*). Refocusing (*accommodation*) takes about 0.2 s.
  - *Smooth-pursuit movements*. We can track smoothly moving visual objects.
  - *Convergent movements*. When objects move closer or further away, our eyes converge or diverge.
- During this lecture we only consider saccadic movements and make (over?)simplification: information comes into visual system as a series of discrete snapshots.

How many 3s?

45586587686456574928655558476529874230984724947324732487942714957238974298247928074293874256487564765490284296847674546427478467457384764856248478984798 5

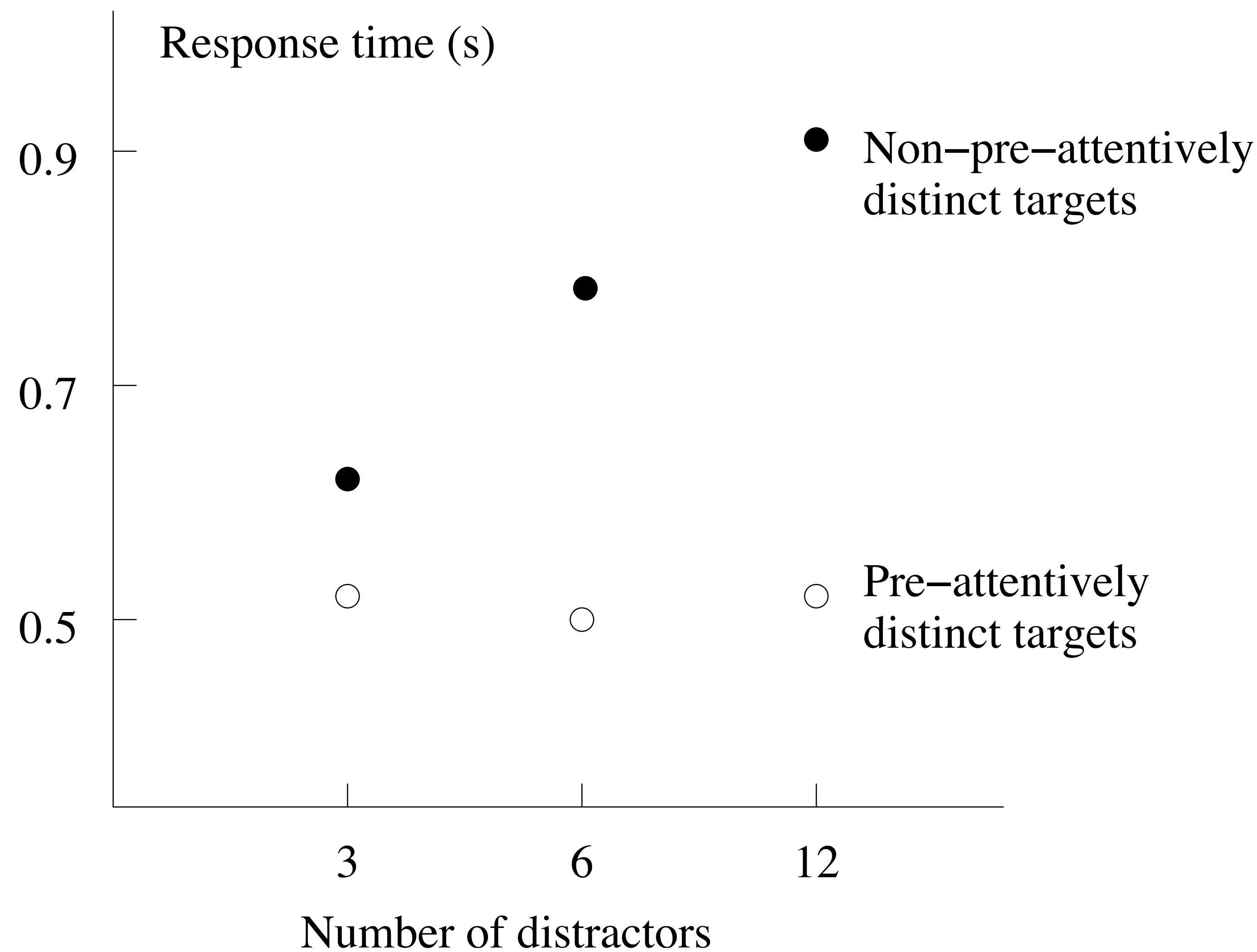
45586587686456574928655558476529874**2**3098472494**7**3247

32487942714957**2**38974298247928074**2**9**3**874256487564765**4**

902842968476745464274784674**5**7**3**84764856248478984798**5**

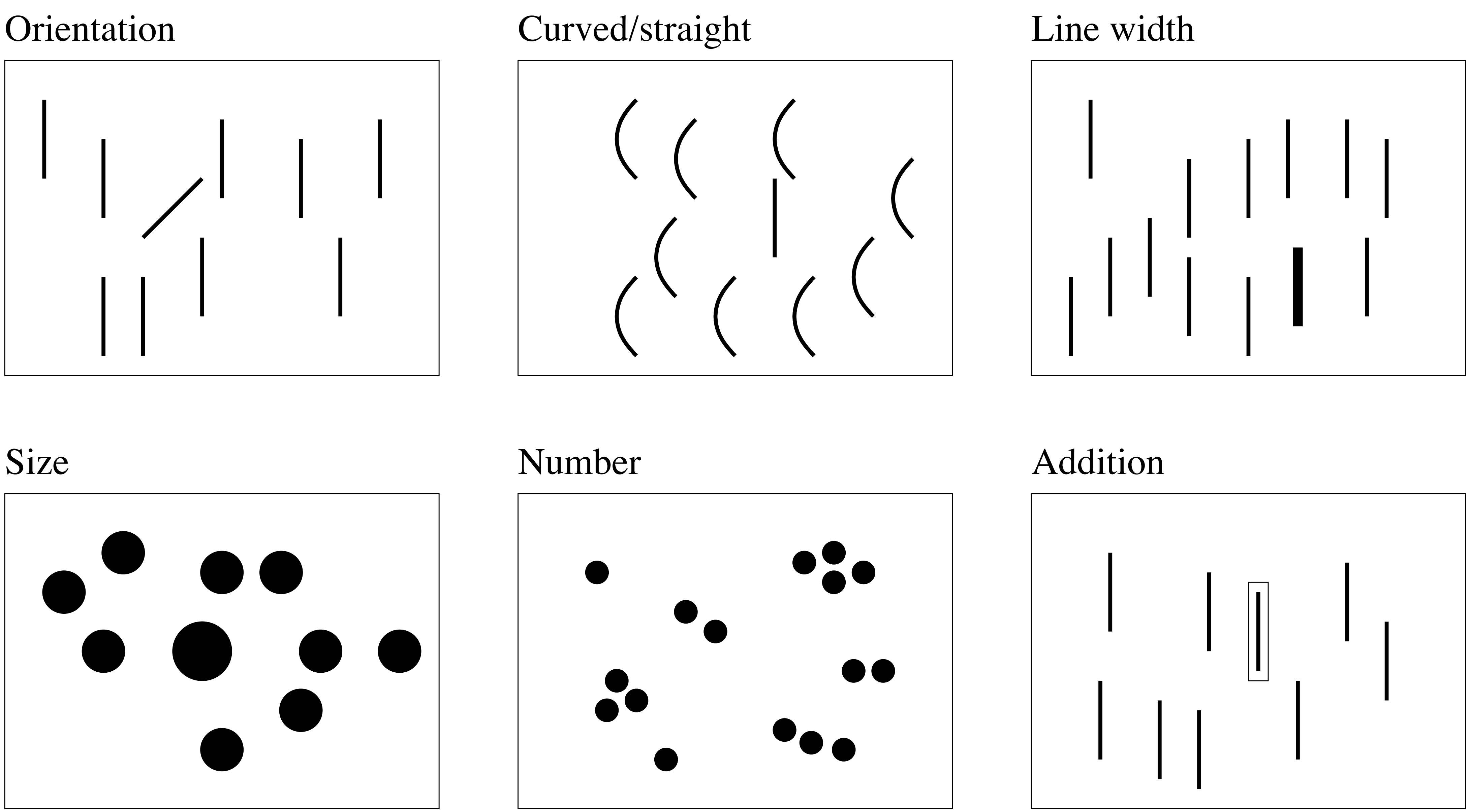
Pre-attentive processing

- Some visual objects are processed *pre-attentively*, before the conscious attention
- Pre-attentive features "pop out"
- Pre-attentive processing speed is independent of the number of distractors
- Processing speed of non-pre-attentive features is slower and the speed decreases as the number of distractors increases (i.e., you must go through all numbers to find 3s)



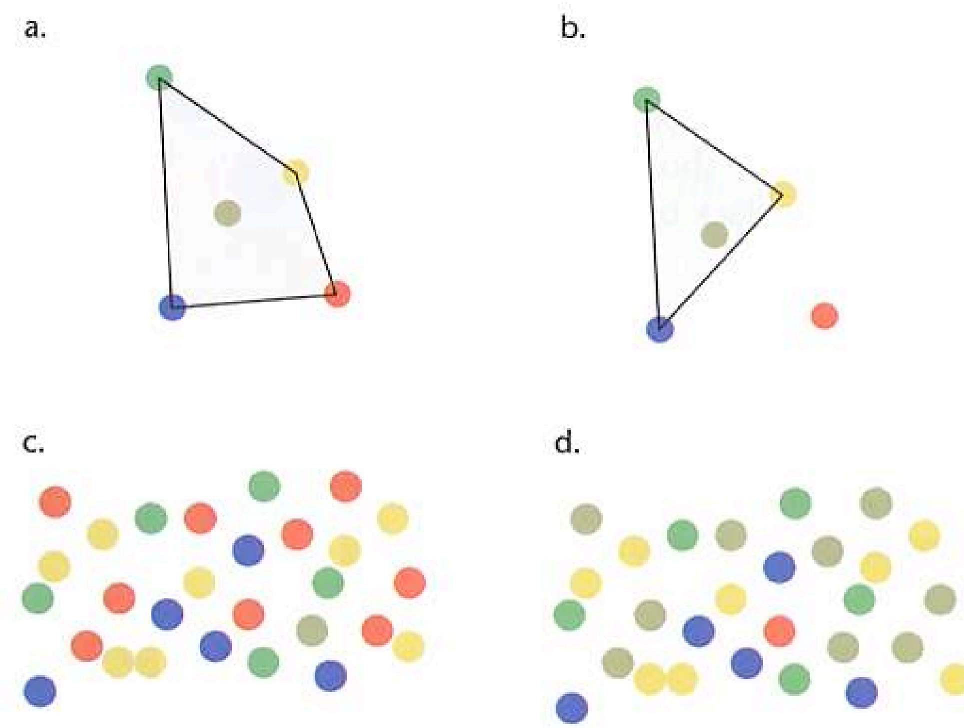
Pre-attentively distinct properties

- **Form** (Line orientation, length, width and collinerity, size, curvature, spatial grouping, added marks, numerosity [up to four])



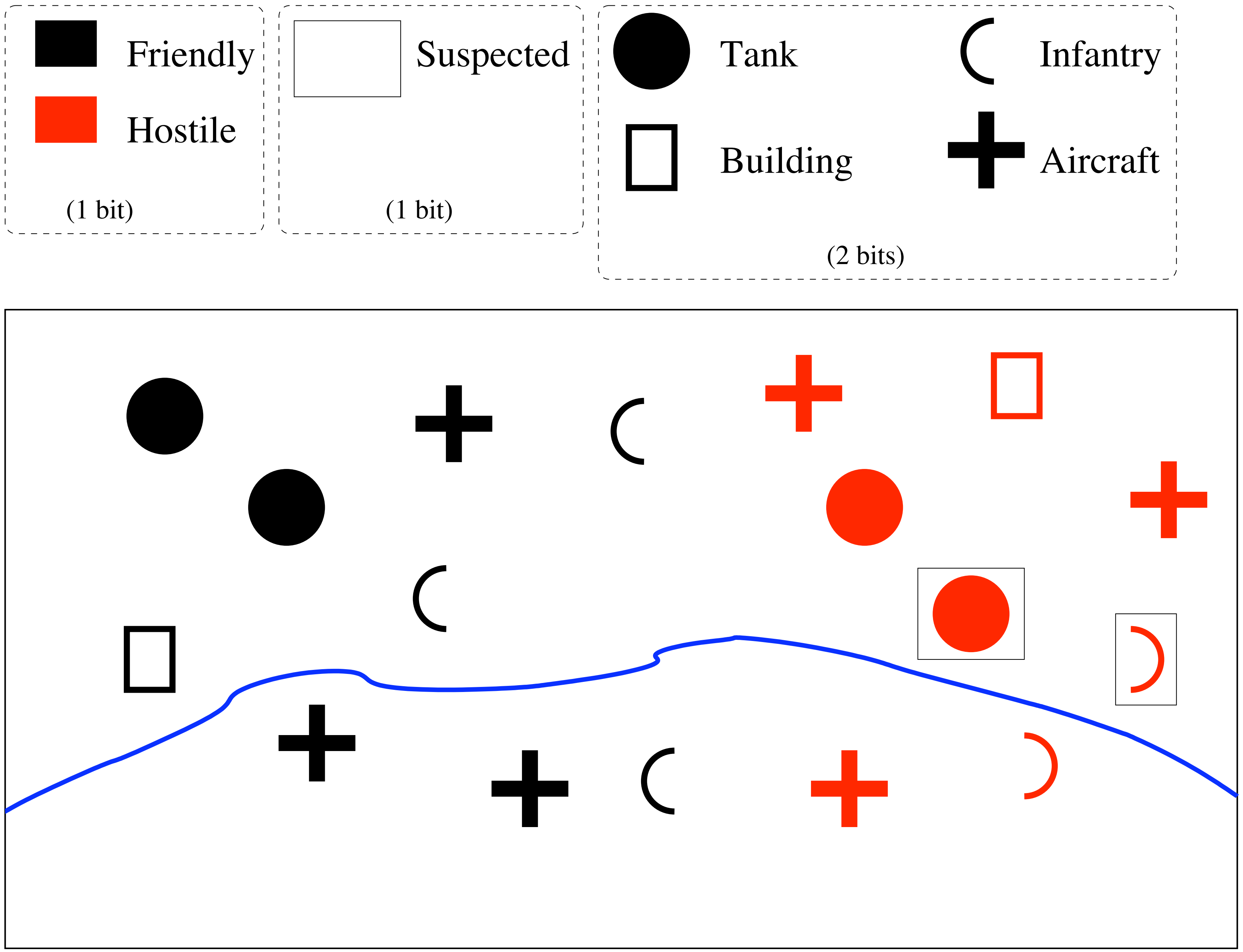
Pre-attentively distinct properties

- **Color** (hue, intensity [if outside CIE convex defined by other colors])



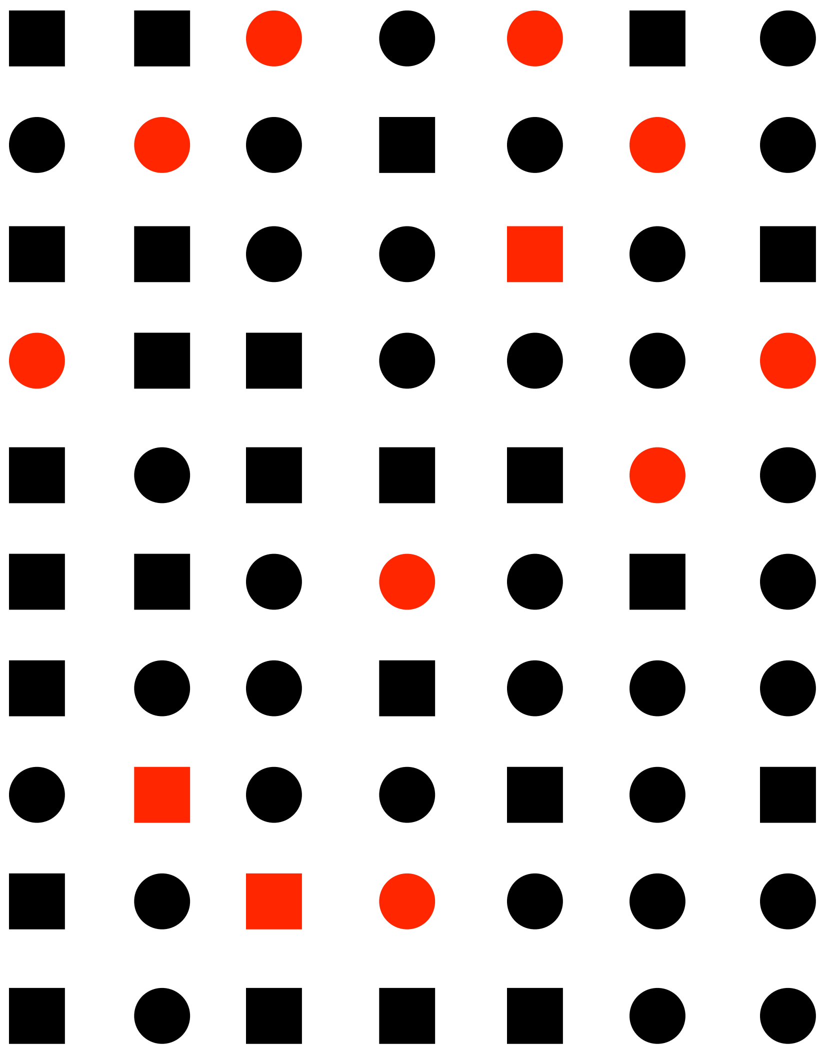
- **Motion** (flicker, direction of motion)
- **Spatial position** (2D position, stereoscopic depth, convex/concave form from shading)

Example: a pre-attentively separable symbol set



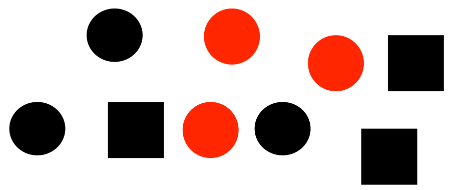
Conjunction searches

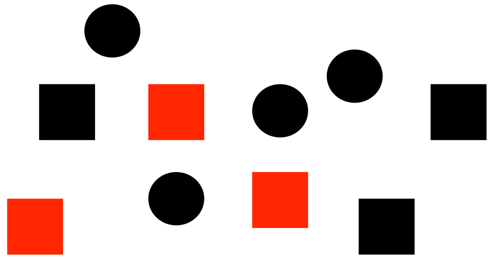
- *Conjunction search* is a visual search that involves searching a specific *conjunction* of several (e.g., 2) visual attributes
- Conjunction searches are usually not pre-attentive, even if the individual features are
- Examples:
  - "*Find red **and** circular objects*" is **not** pre-attentive search (conjunction search)
  - "*Find red objects*" is pre-attentive search
  - *"Find circular objects"* is pre-attentive search

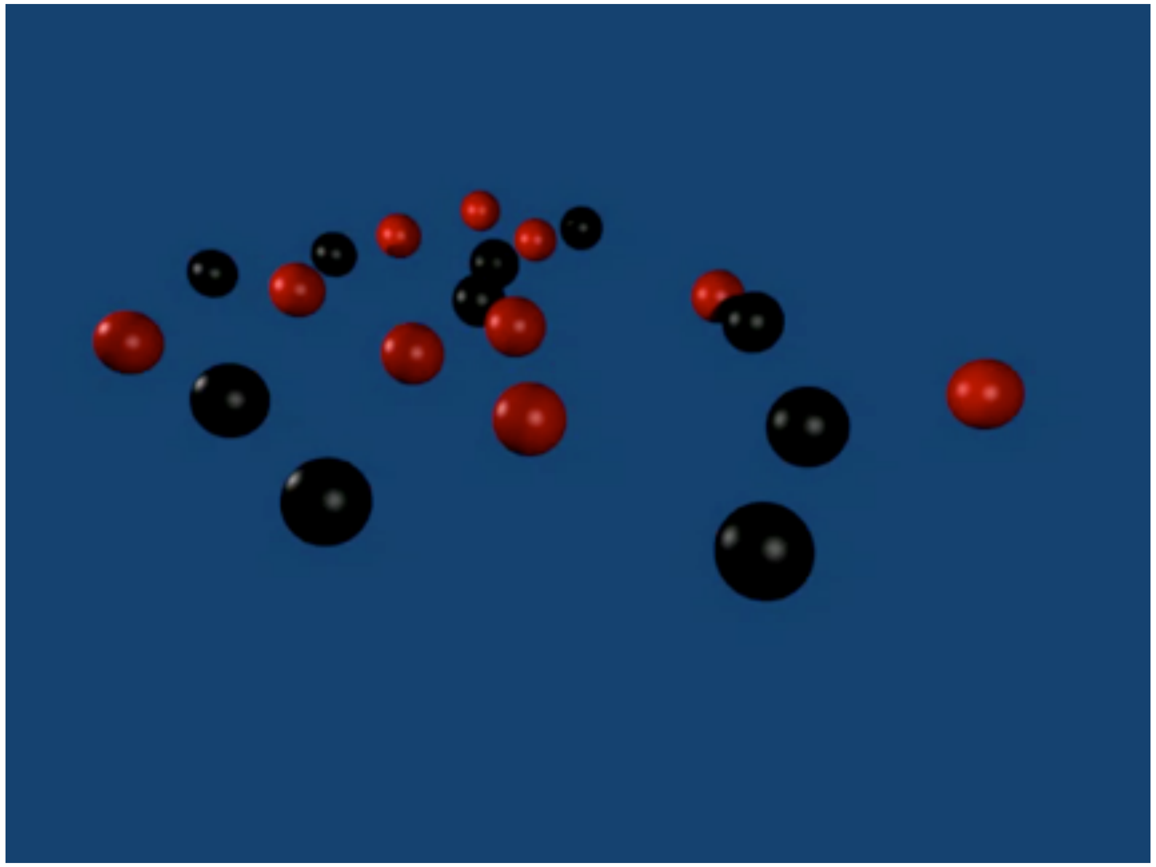


Conjunction searches

- Conjunction searches are usually not pre-attentive
- Some exceptions:
  - spatial grouping on the XY plane (“*find red objects in the lower group*”)
  - motion (“*find red moving things*”)
  - stereoscopic depth
  - combination of convexity/ concavity and color

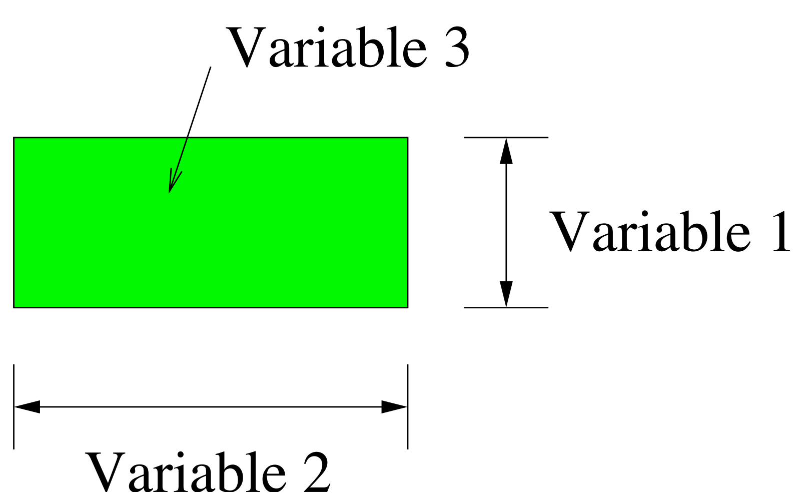


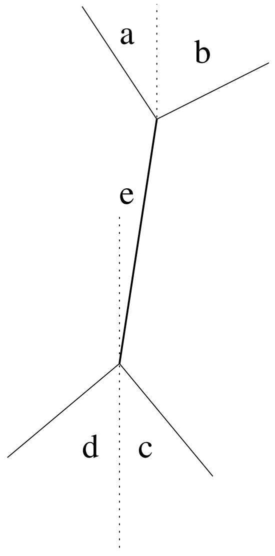


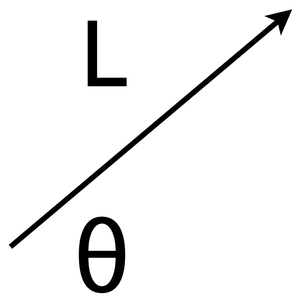


Glyphs

- Glyphs are symbols used to describe multivariate discrete data
- Single glyph corresponds to one sample in a data set
- Data values are mapped to the visual properties of the glyph
- Problem: how to design a glyph so that the data values can be perceived pre-attentively?

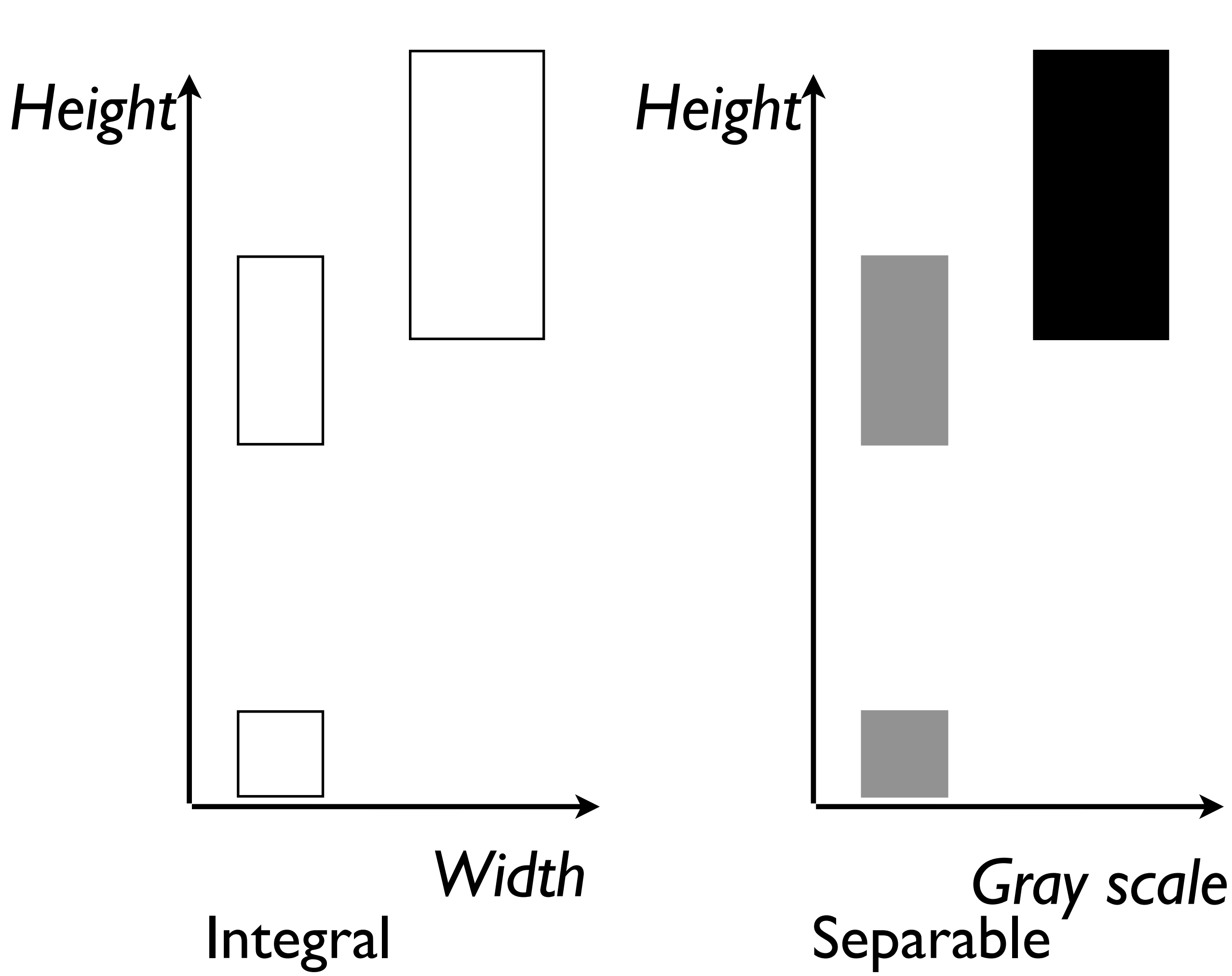




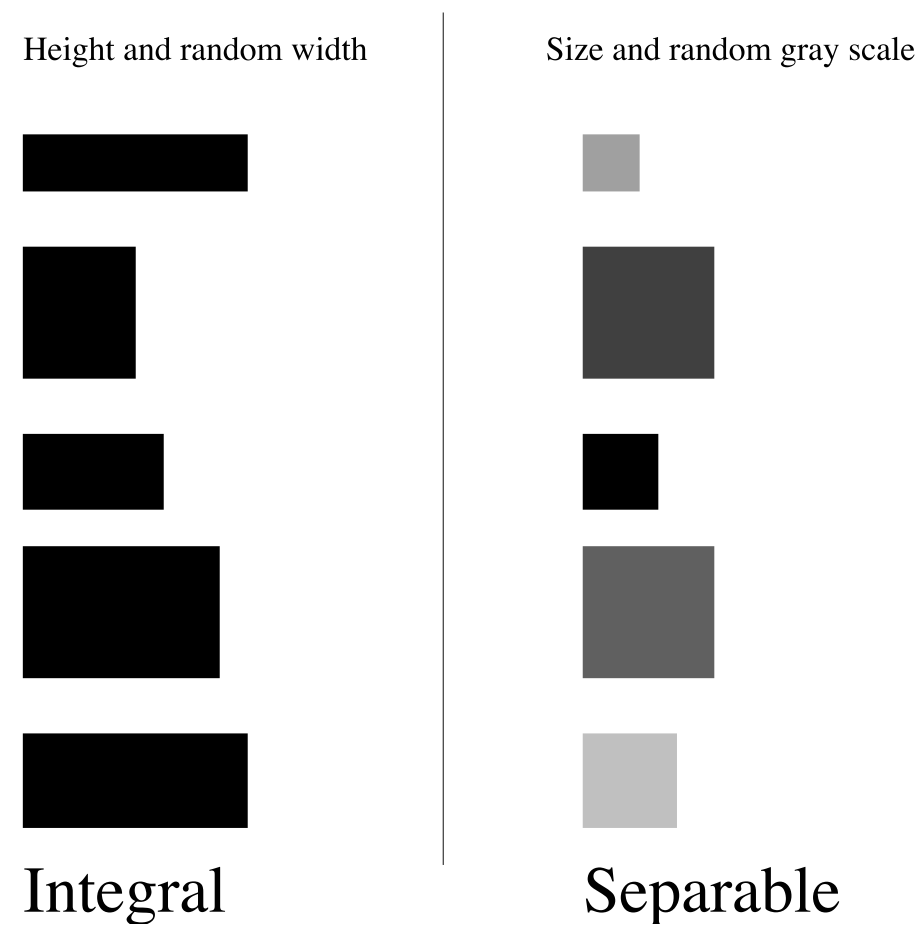


Integral and separable dimensions

- Pre-attentive features are *separable*, if they are perceived independent of each other (e.g., size and gray scale)
- Pre-attentive features are *integral*, if they are perceived holistically (e.g., width and height)
- Lesson: use separable features in glyphs



Restricted classification task



Find this high rectangle.

Speeded classification task

Integral and separable dimensions

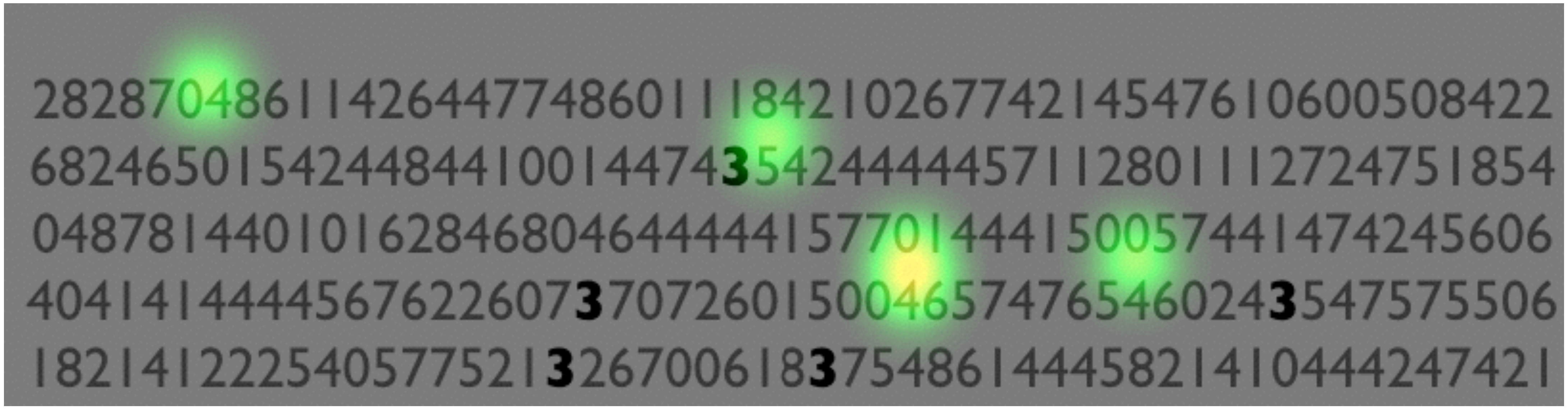
- Pairs of integral and separable dimensions have been determined in a number of ways
- Little work has been done on interactions among three or more display variables

Integral

red–green	yellow–blue
red–green	black–white
height	width
shape	size
color	size
direction of motion	shape
color	shape
color	direction of motion
XY position	size, shape or color

Separable

Summary: Pre-attentive processing

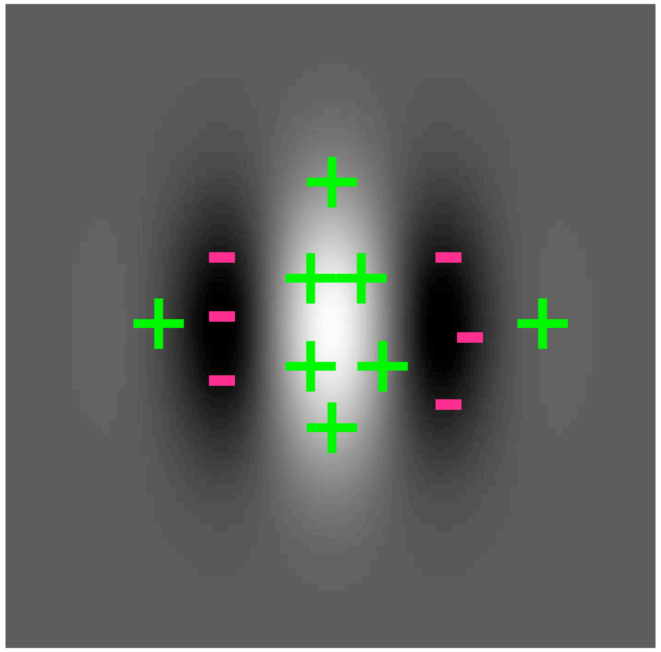


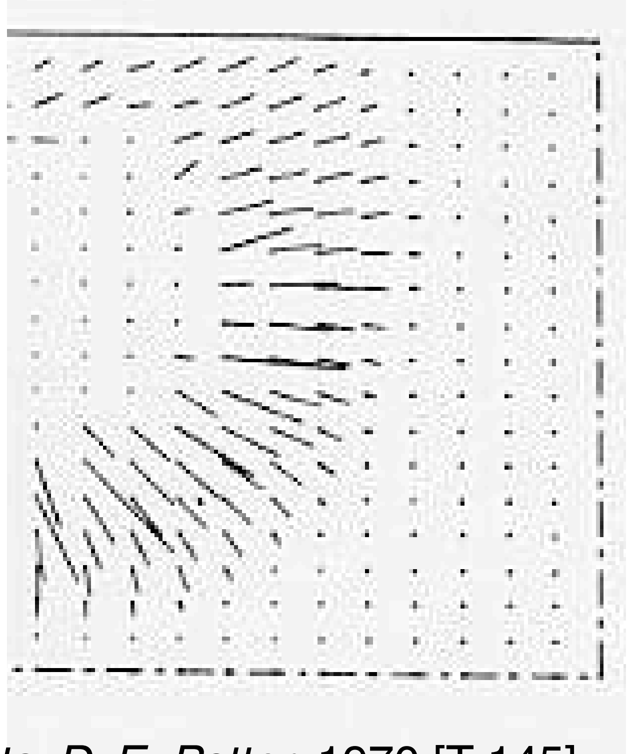
- Some visual objects are processed *pre-attentively,* before the conscious attention
- Pre-attentive features “pop out”
- Pre-attentive processing speed is independent of the number of distractors
- Some limitations: conjunction search usually not possible, some channels interfere with each other, limited resolution on channels

Gabor textures

- Some (but not all) pre-attentive features can be explained by the properties of early visual processing
- One example: visual areas 1 and 2 (V1&V2) contain large arrays of neurons that filter for orientation and size information at each point of the visual field
- These arrays can be modeled by the *Gabor model*
- Locally changing textures/glyphs can become a *texture field*

$$\text{Response} = C\cos\left(\frac{\pi}{\sqrt{2}}\frac{Ox}{S}\right)\exp\left(-\frac{1}{2}\frac{\|x\|^2}{S^2}\right)$$



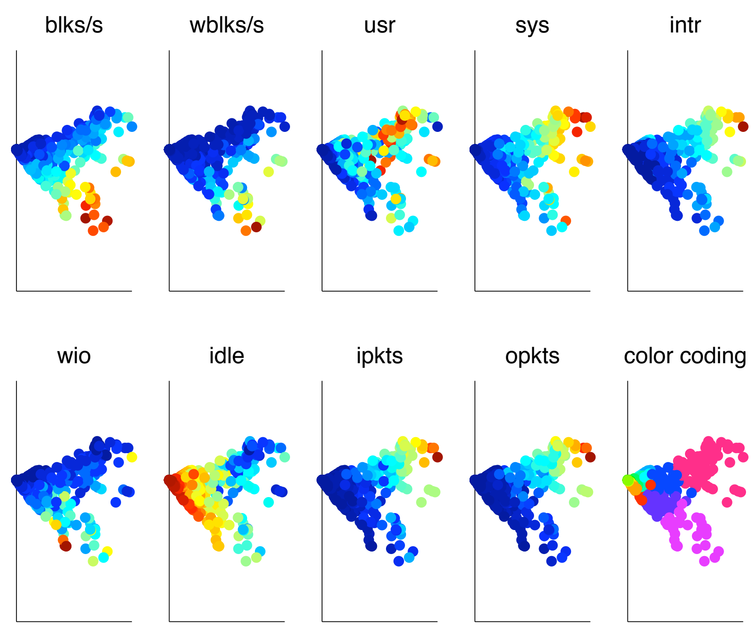


K. V. Roberts, D. E. Potter, 1970 [T 145].

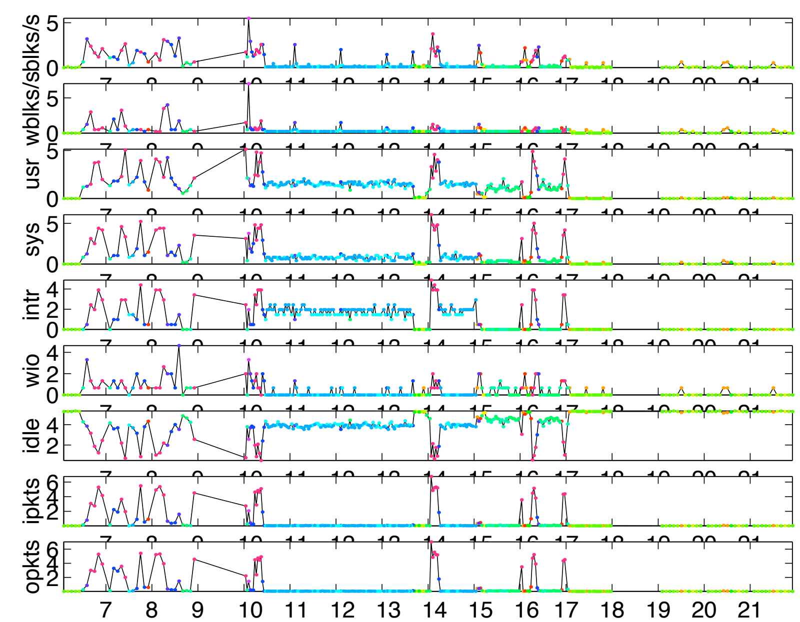
Patterns in 2D data

- Exploratory visualization is based on finding patterns from data
- Oversimplification: the patterns are recognized between pre-attentive processing and higher level object perception
- Relevant questions:
  - How do we see groups?
  - How can 2D space be divided into perceptually distinct regions?
  - When are two patterns similar?
  - When do two different elements appear to be related?
- Patterns may be perceived even where there is only visual noise

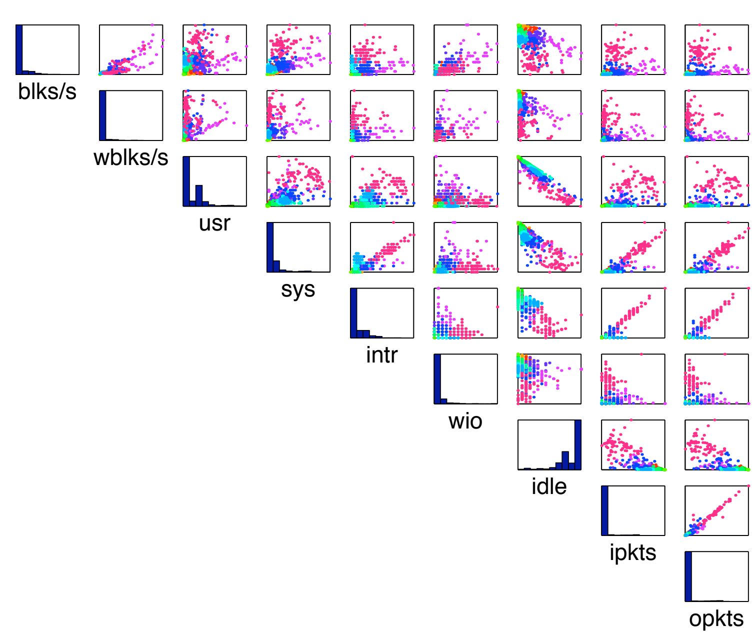
Patterns in 2D data



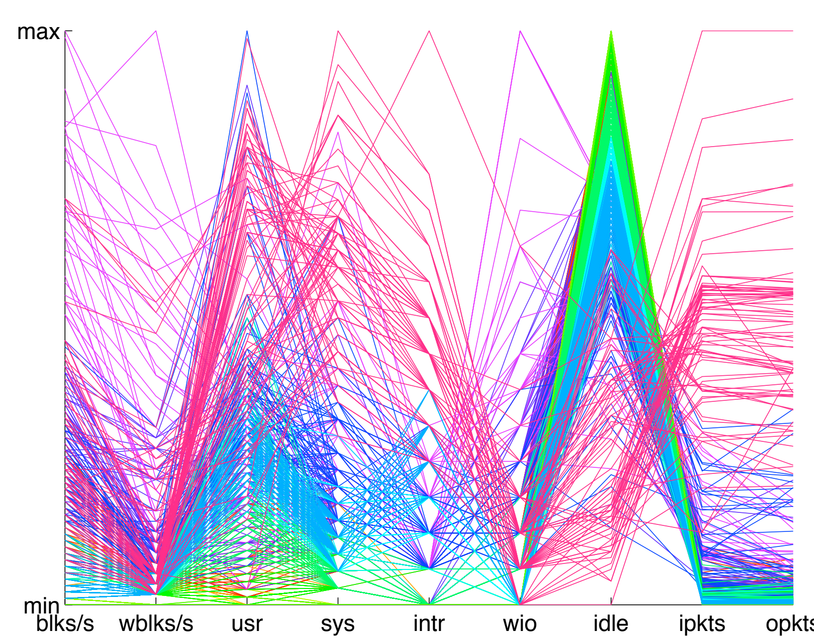
(a) Component planes



(b) Time-series



(c) Scatterplot matrix

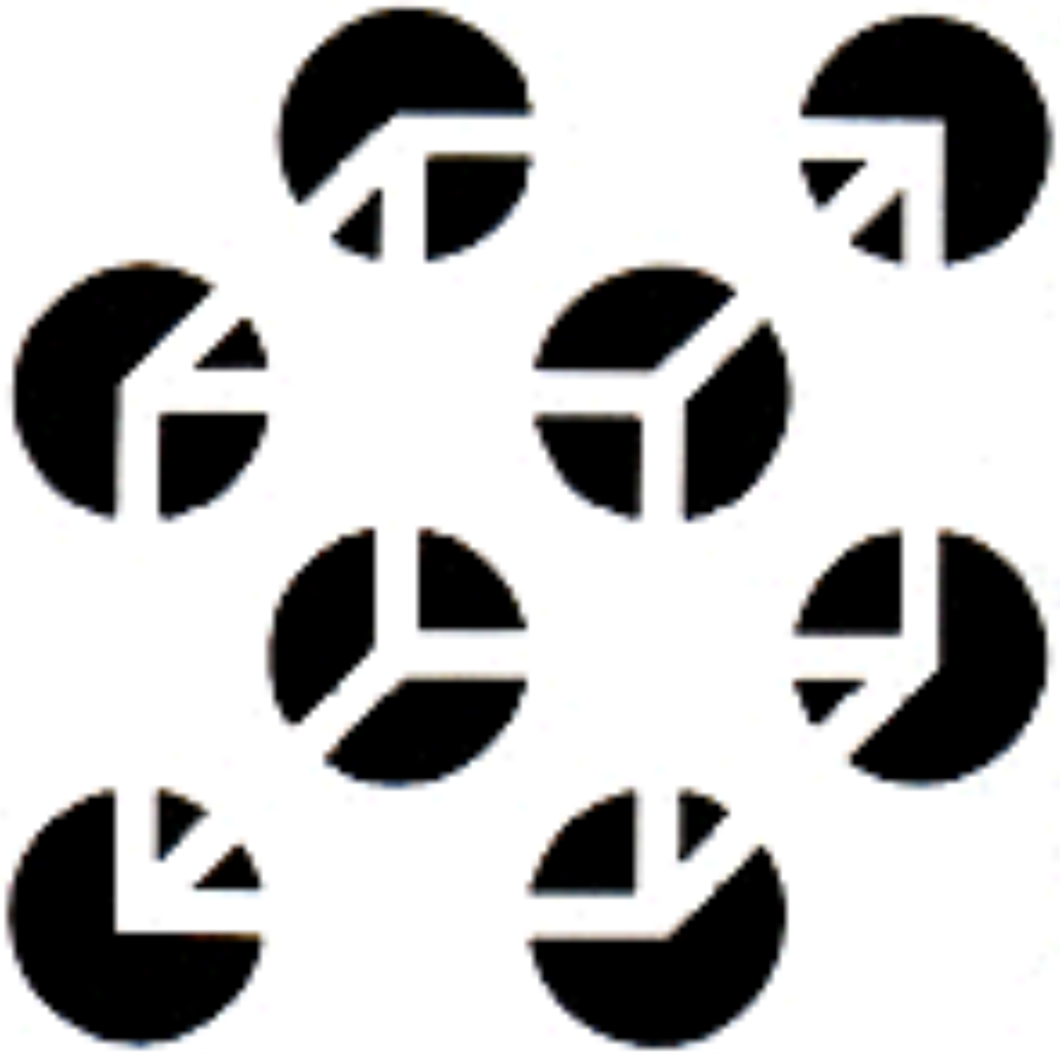


(d) Parallel axis

[V p. 38]

Gestalt laws

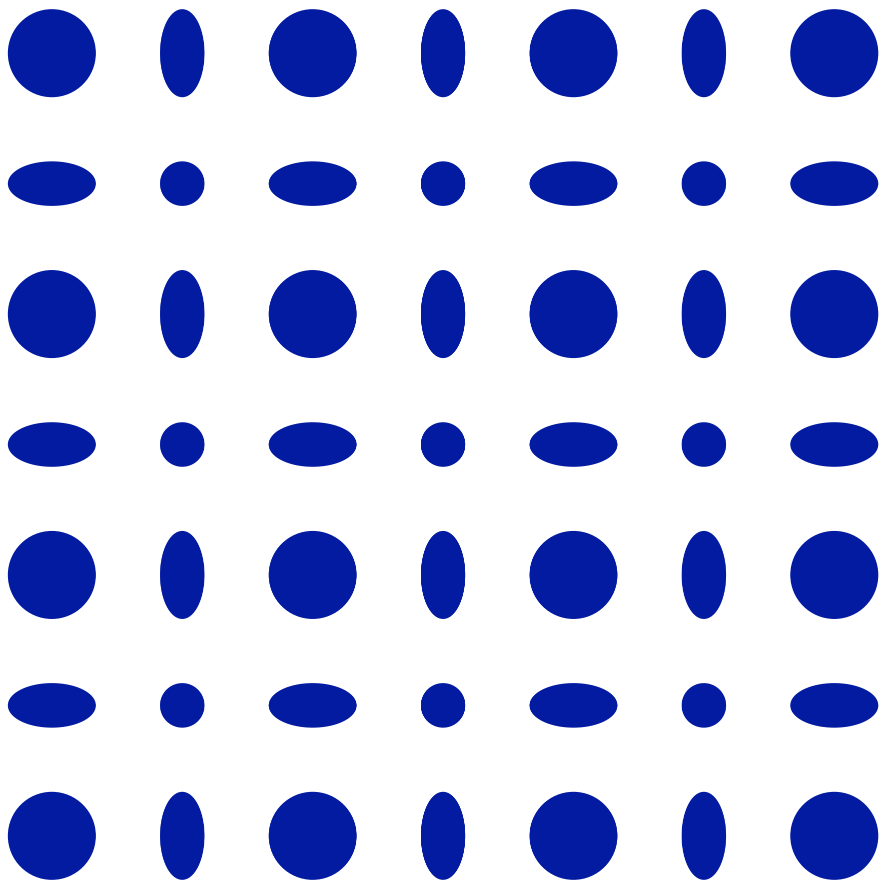
- *Gestalt* is *form* in German
- The Gestalt School of Psychology (1912 onwards) investigated the way we perceive form
- They produced several *Gestalt laws* (laws of organization) of pattern perception
- The Gestalt laws translate directly into design principles of visual displays
- Many of the rules seem obvious, but they are violated often



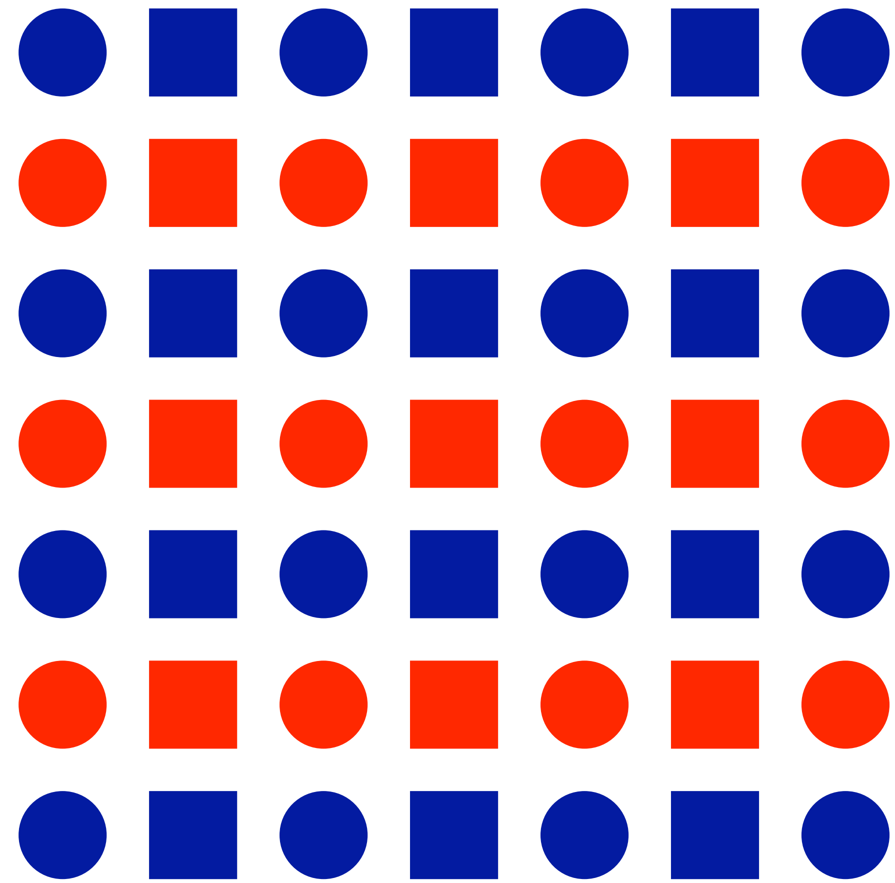
Bradley&Petry 1977

Similarity

Similar things appear to be grouped together.



Integral dimensions are used to delineate rows and columns

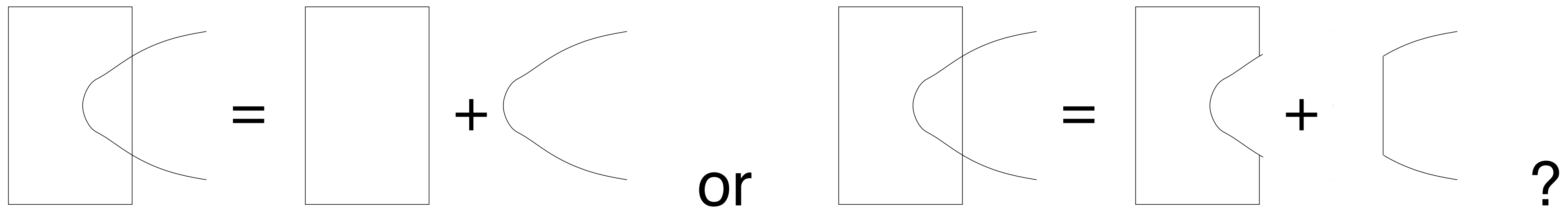


Separable dimensions are used to delineate rows and columns

- The use of integral dimensions emphasizes the overall pattern
- The use of separable dimensions segment the rows and columns

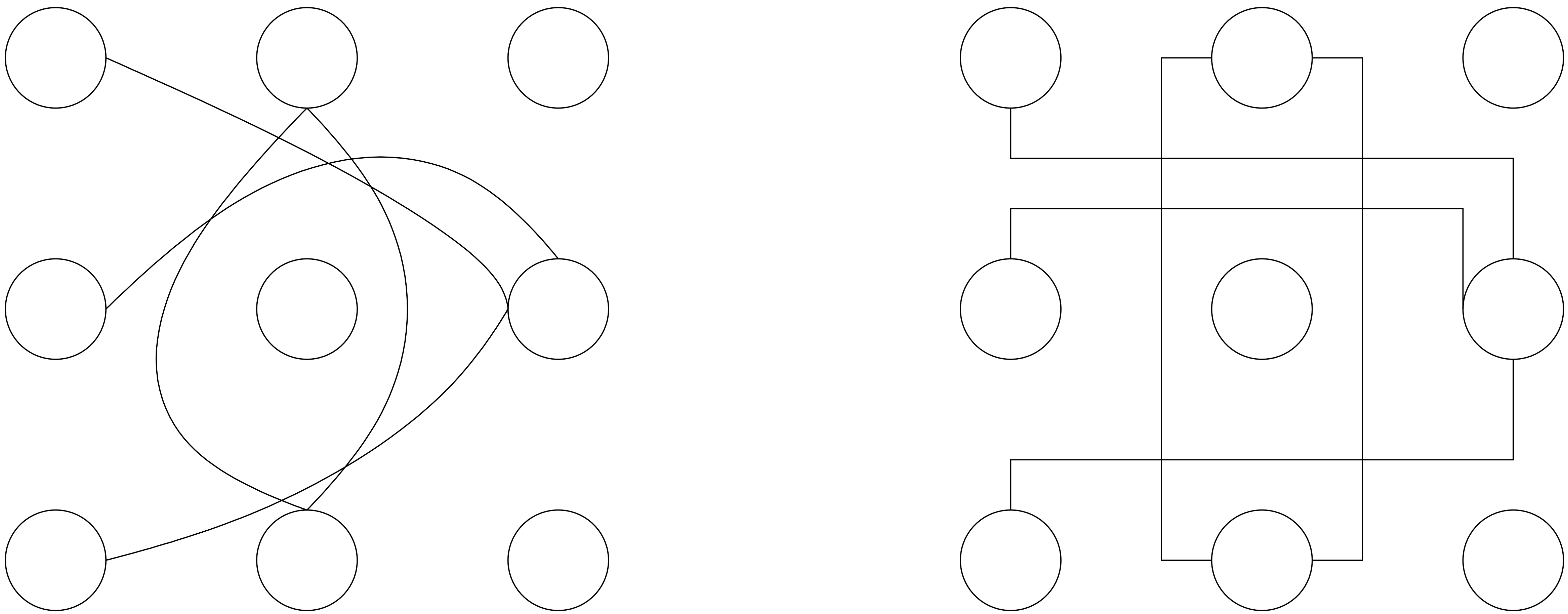
Good continuation

Points that, when connected, result in straight or smoothly curving lines are seen belonging together, and lines tend to be seen in such way as to follow the smoothest path.



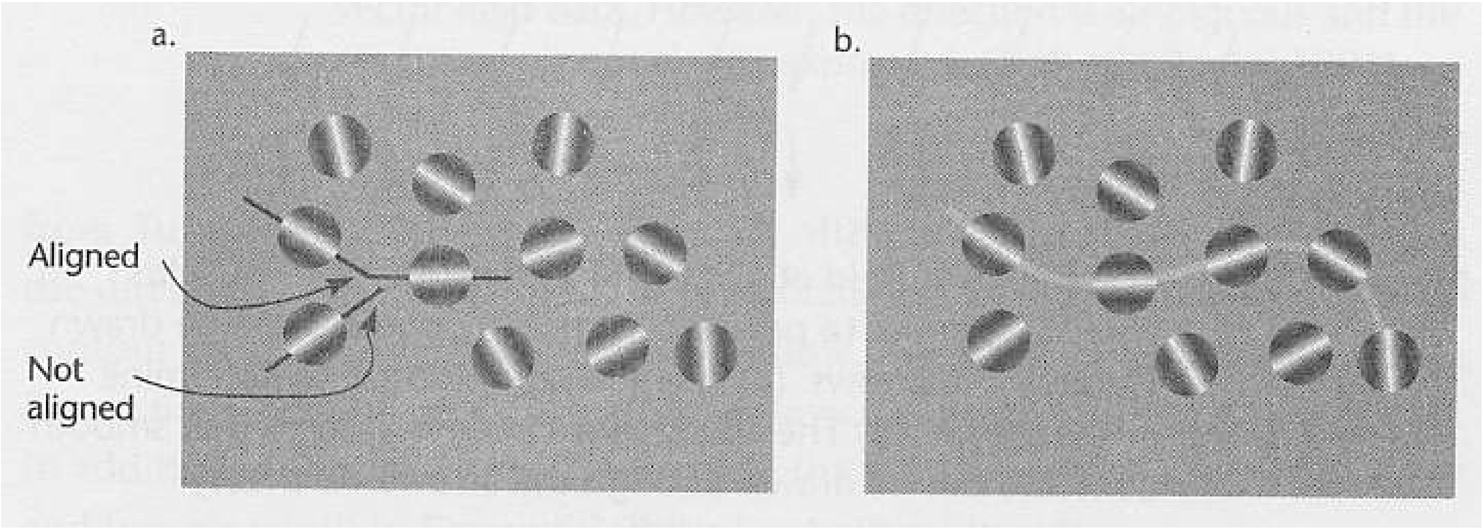
Good continuation

- Connectedness is maybe the most powerful of the grouping principles
- It is easier to perceive connections when contours connect smoothly:

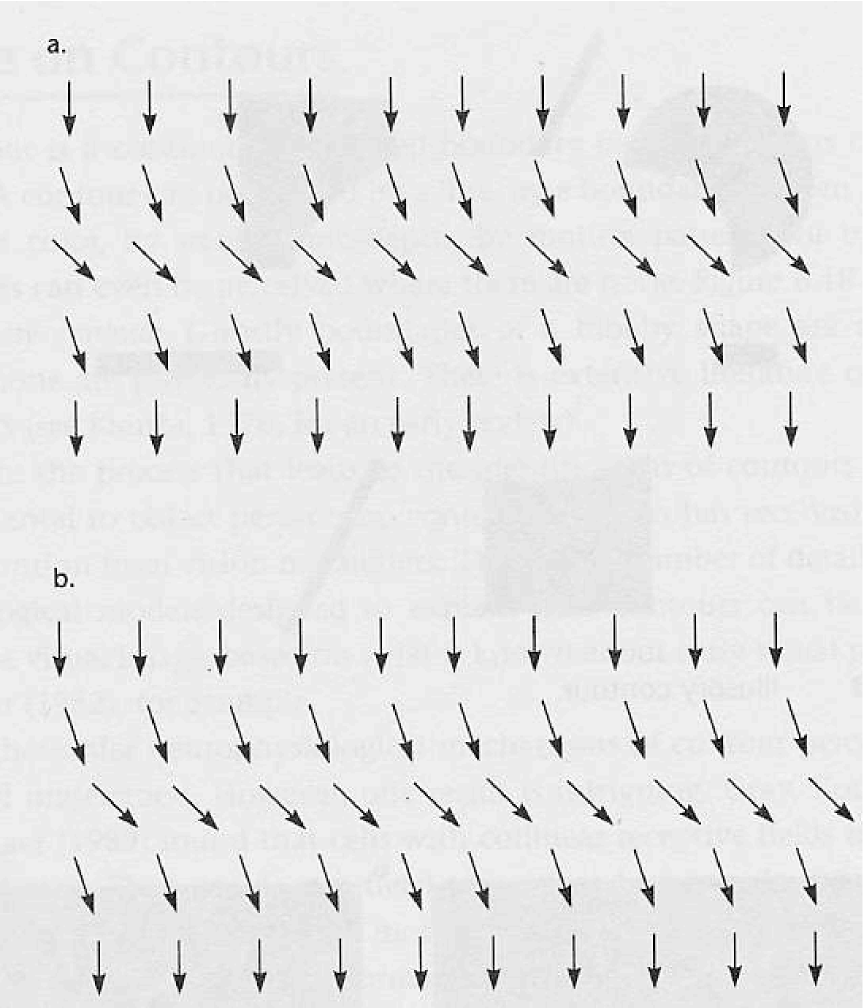


Good continuation

- People detect presence of even quite wiggly paths through a sequence of Gabor patches if a smooth curve can be drawn through them



[W 6.19]



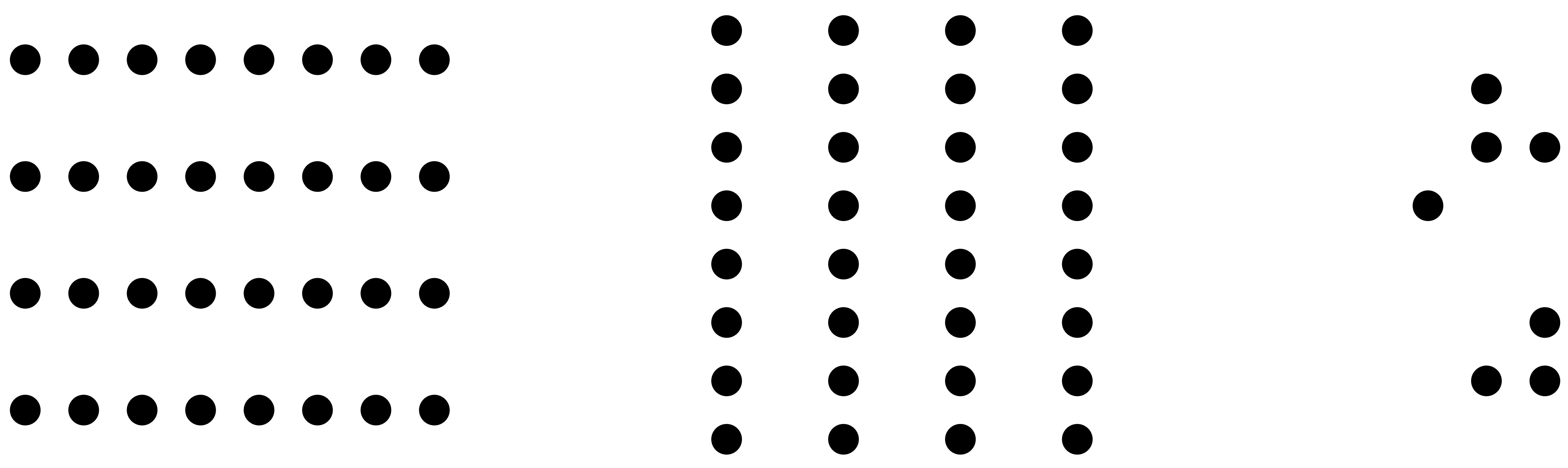
The arrows have been shifted so that smooth countours can be drawn through them (lower right).

[W 6.20]

Proximity or nearness

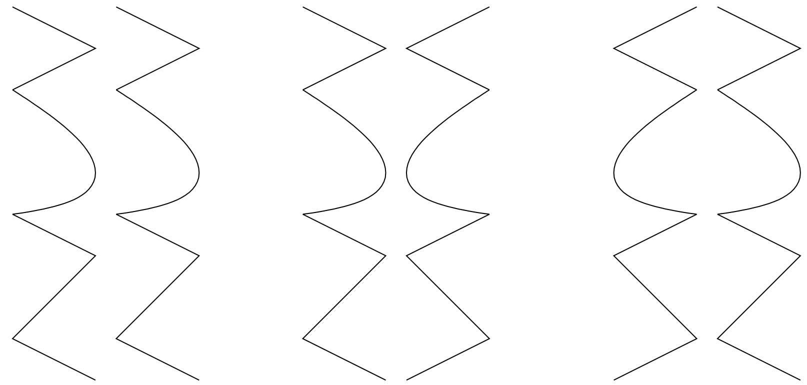
Things that are near to each other appear to be grouped together.

- Proximity is one of the most powerful of the Gestalt laws
- Place the data elements in proximity in a display to emphasize relationships between them

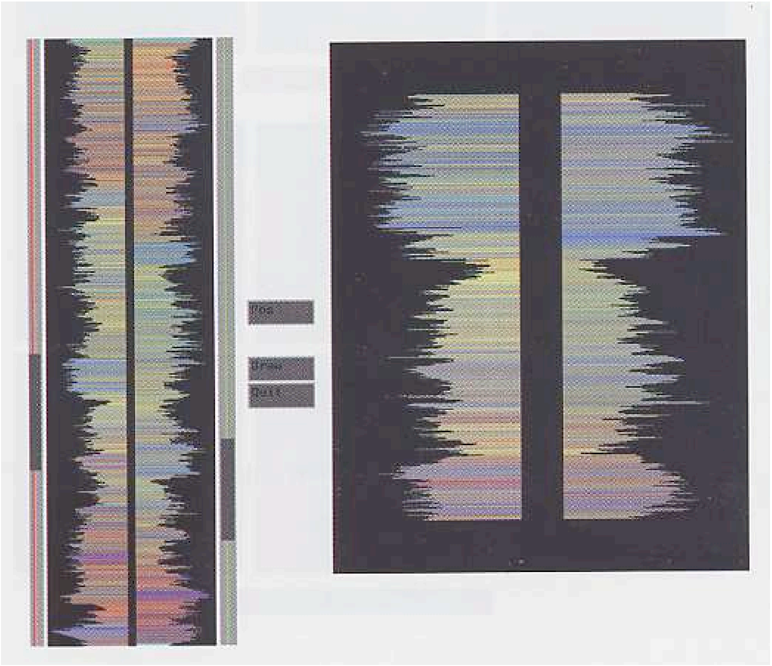


Symmetry

Symmetrically arranged pairs are perceived strongly together.



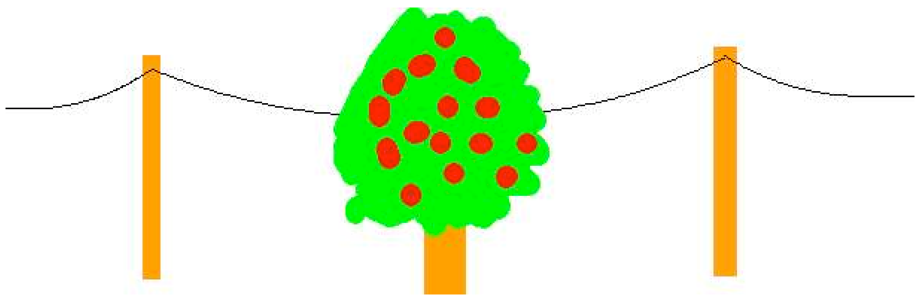
Example: it is easier to spot differences in two data sets by using the vertical symmetry:



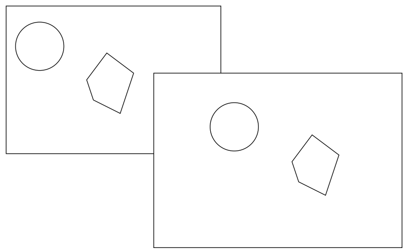
[W 6.11].

Closure

Closed controus tend to be seen as objects. There is a tendency to close controus that have gaps in them.

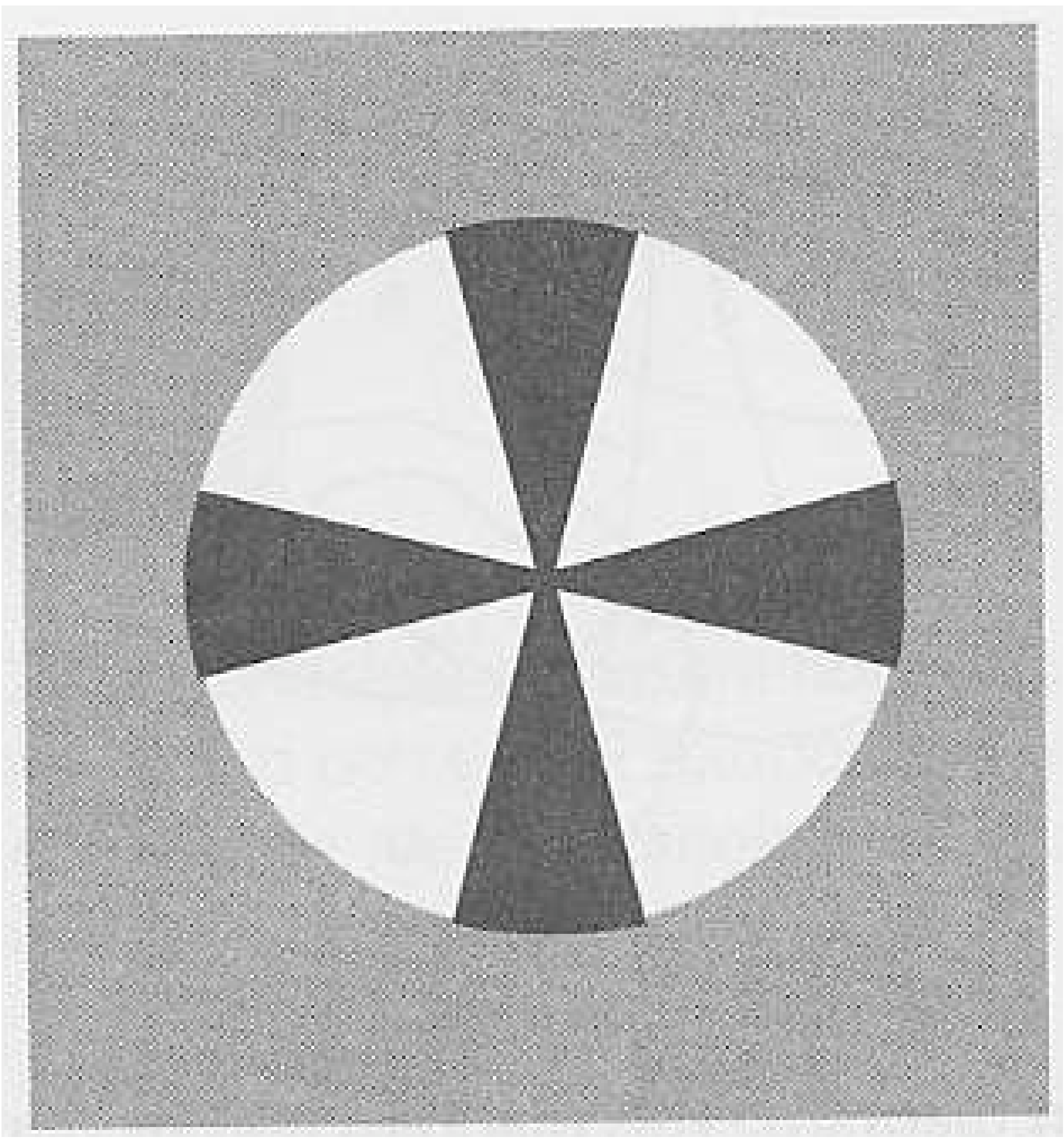


- Closed controus are extremely important in segmenting windows-based interface
- The strong framing effect inhibits between-window-comparisons: related information should not be placed in separate windows



Relative size

Smaller components of a pattern tend to be perceived as objects.

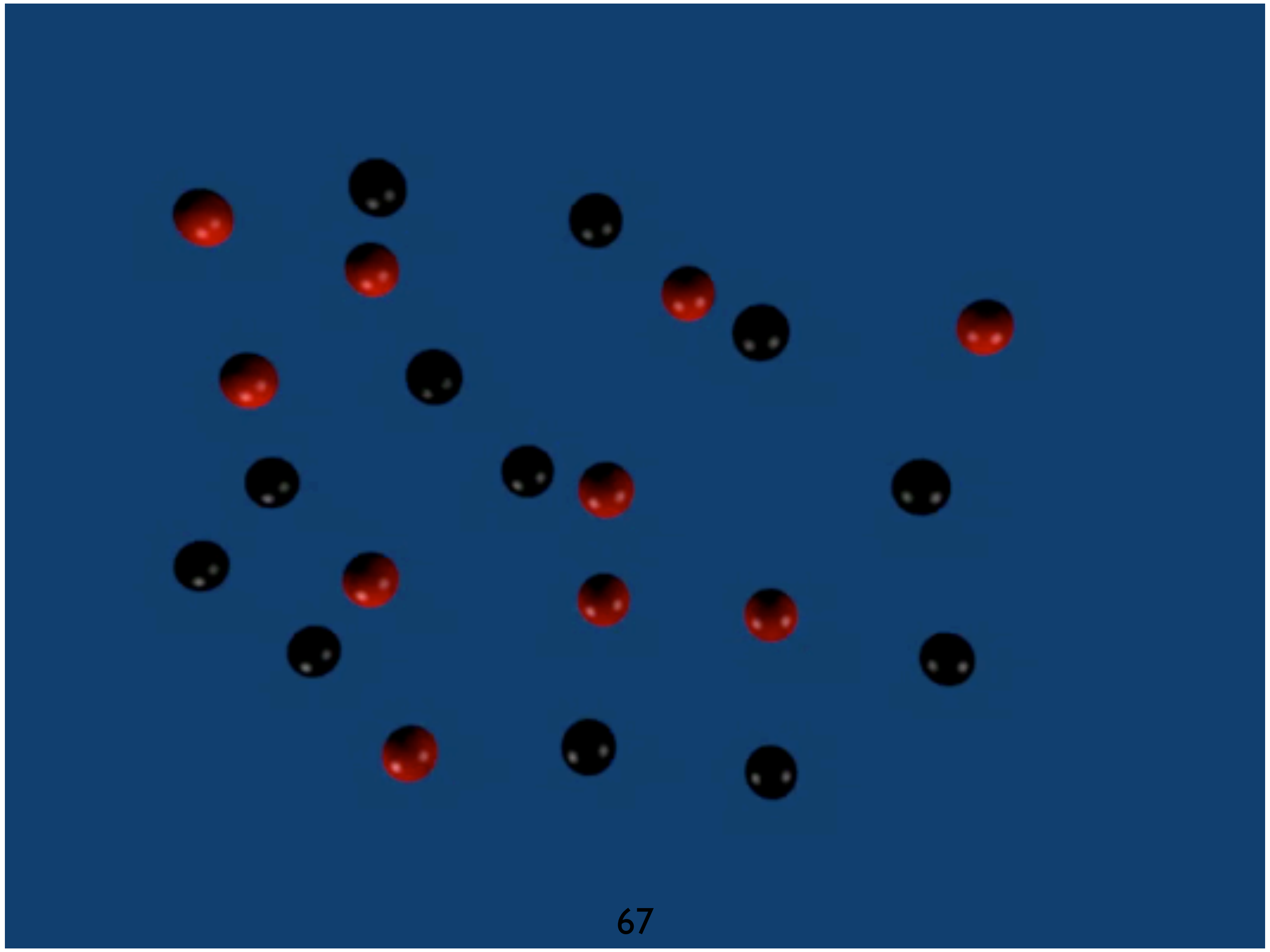


[W 6.15]

Rubin's reversible face-vase figure
(multistability)

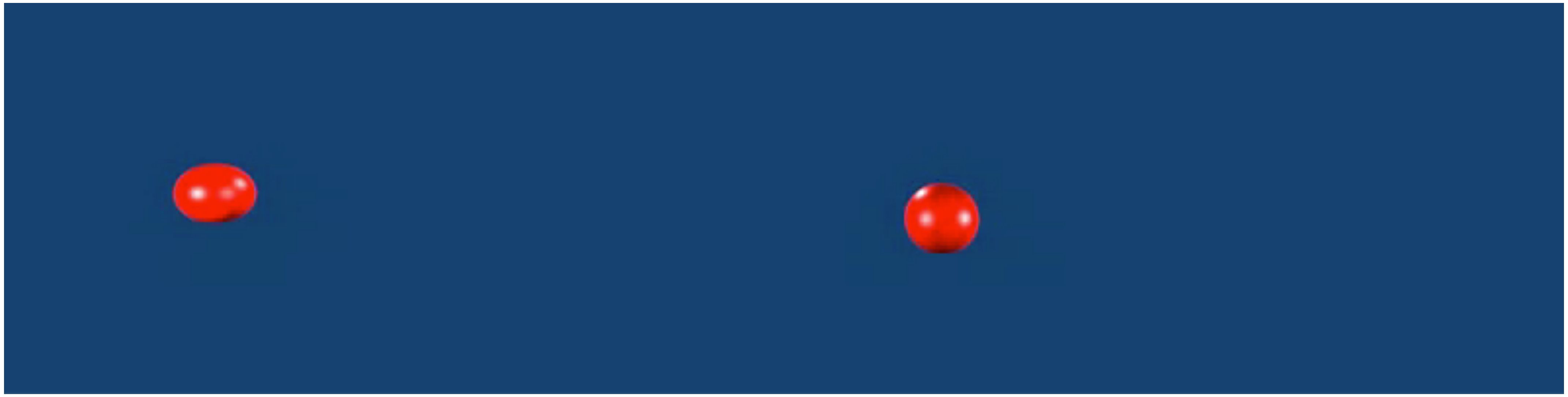
Patterns from motion

- Relative motion is an extremely efficient method of showing patterns from data
- Demo:
  - Data points oscillate around center point
  - Variables: frequency, phase, amplitude of motion
  - Phase is the most effective variable



Animation and perception of shapes

- Gestalt laws also work for animated images: structures and patterns are seen from partial data (as with static images)
- Mystery lights in the dark:



No delay

Causality

- *Launching:* an object is perceived to set another into motion
- Perception of launching requires precise timing (delays less than 0.07-0.16 s)
- Already infants can perceive causal relations, such as launching

Delay of 0.2 s

Summary of Gestalt laws

- Similarity
- Good continuation
- Proximity
- Symmetry
- Closure
- Relative size
- Some "new" motion-based Gestalt(-like) laws (after examples):
  - Patterns from motion
  - Animation and perception of shapes
  - Causality