Data Visualization (CSCI 627/490)

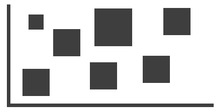
Color

Dr. David Koop

Arrange Tables

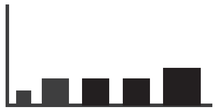
- **Express Values**

- **Separate, Order, Align Regions**
  - Separate



  - Order

  - Align



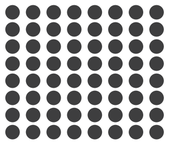
    - *1 Key*
      List

    - *2 Keys*
      Matrix

    - *3 Keys*
      Volume

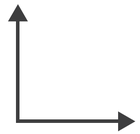
    - *Many Keys*
      Recursive Subdivision

- **Layout Density**
  - Dense



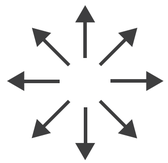
  - Space-Filling

- **Axis Orientation**
  - Rectilinear



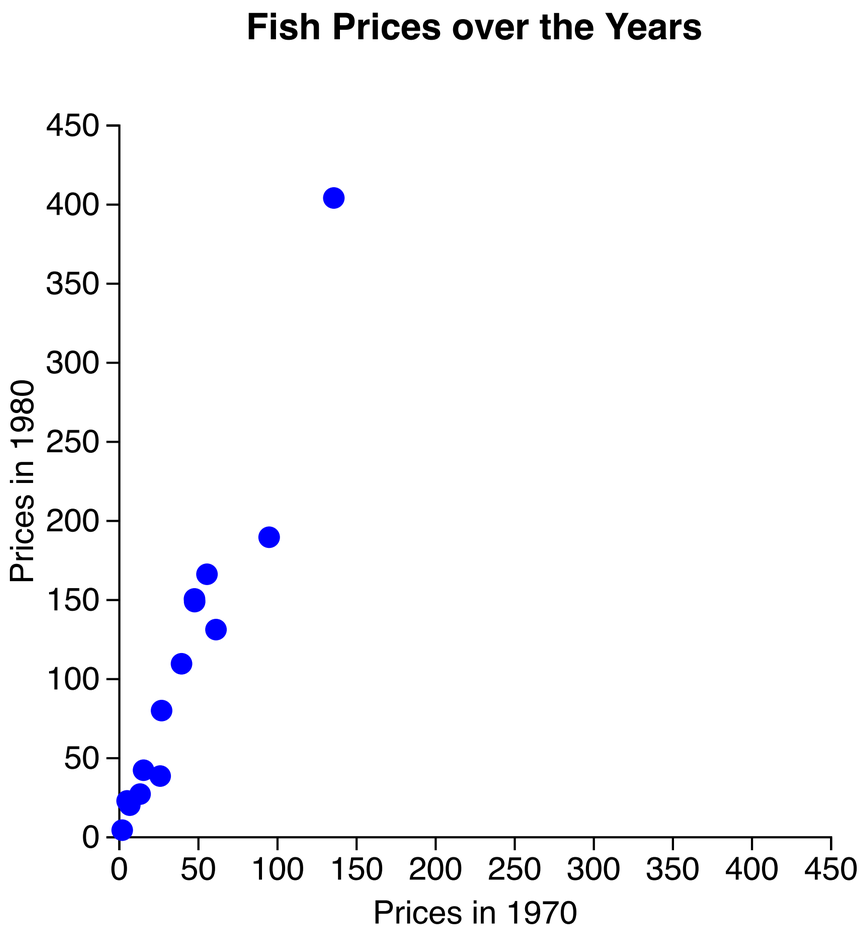
  - Parallel

  - Radial



[Munzner (ill. Maguire), 2014]

Express Values: Scatterplots

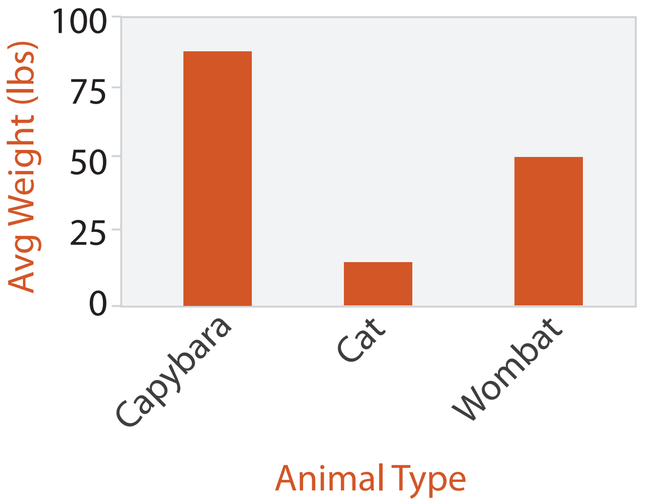


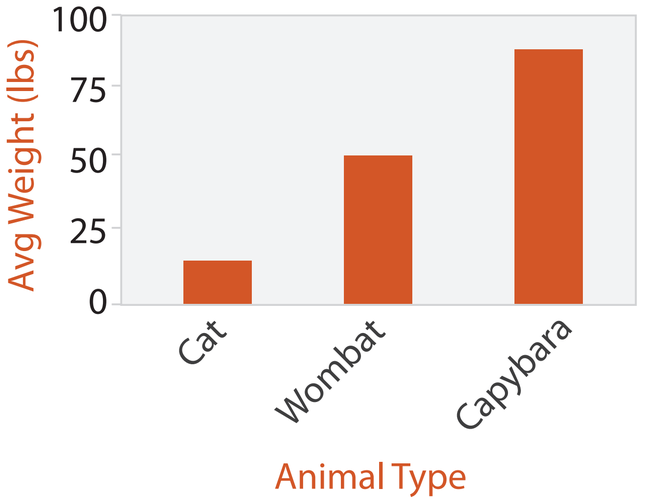
- Data: two quantitative values
- Task: find trends, clusters, outliers
- How: marks at spatial position in horizontal and vertical directions

- Correlation: dependence between two attributes
  - Positive and negative correlation
  - Indicated by lines
- Coordinate system (axes) and labels are important!

List Alignment: Bar Charts

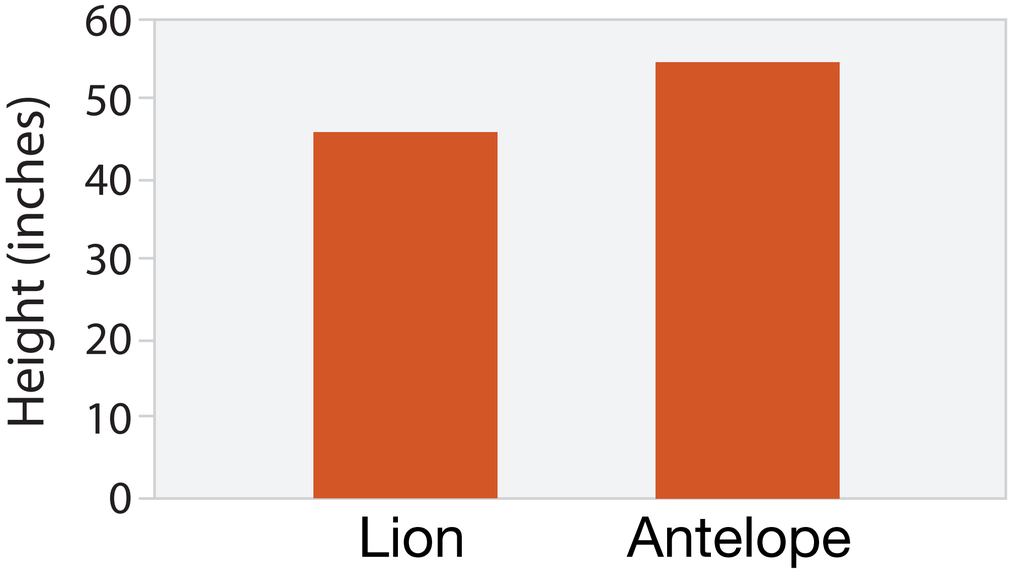
- Data: one quantitative attribute, one categorical attribute
- Task: lookup & compare values
- How: line marks, vertical position (quantitative), horizontal position (categorical)
- What about **length**?
- Ordering criteria: alphabetical or using quantitative attribute
- Scalability: distinguishability
  - bars at least one pixel wide
  - hundreds

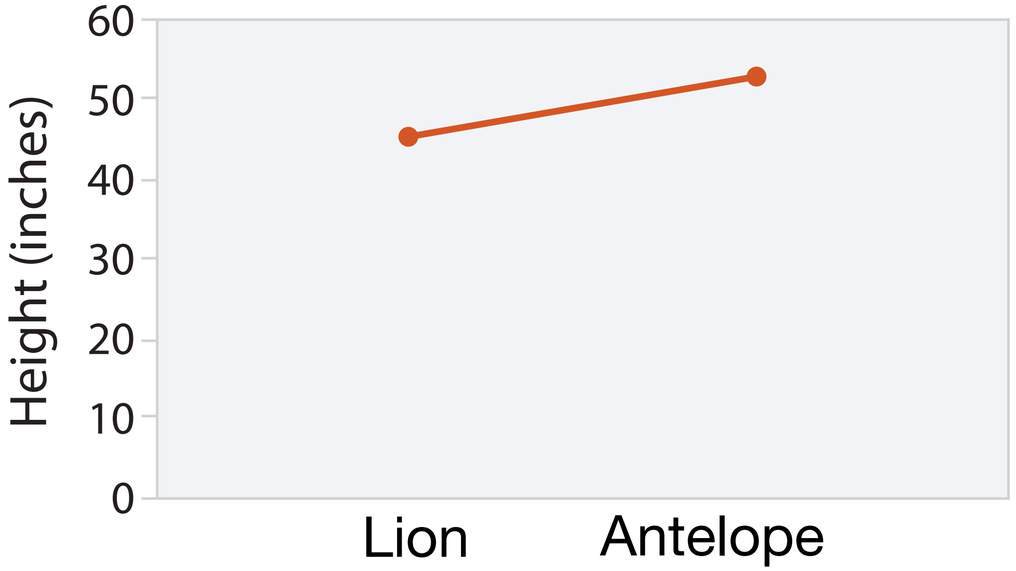




[Munzner (ill. Maguire), 2014]

Proper Use of Line and Bar Charts



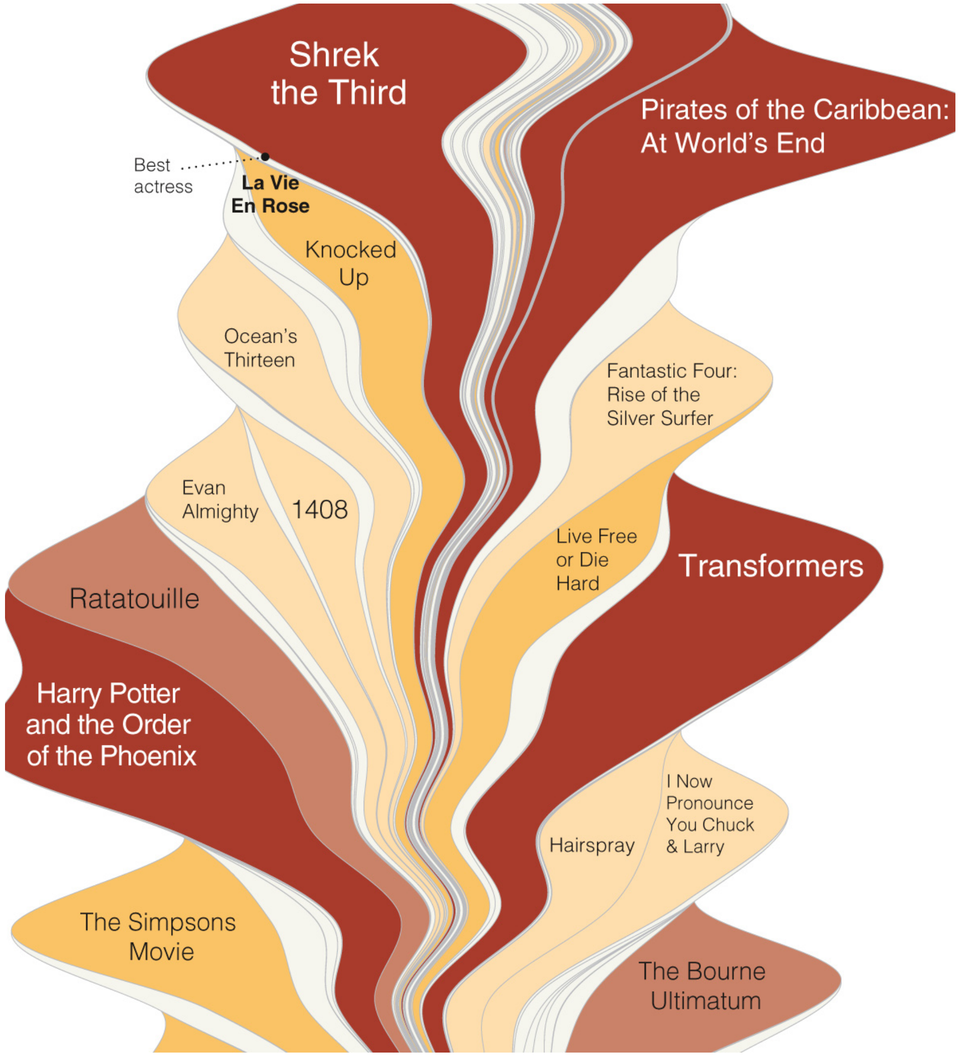


- What does the line indicate?
- Does this make sense?

[Adapted from Zacks and Tversky, 1999, Munzner (ill. Maguire), 2014]

Streamgraphs

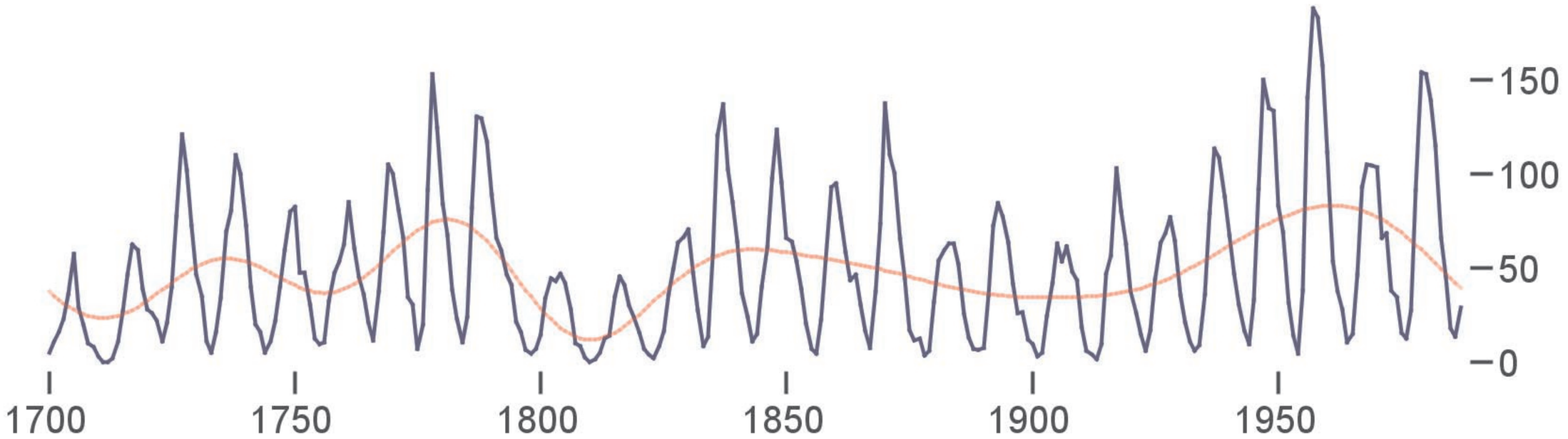
- Include a time attribute
- Data: multidimensional table, one quantitative attribute (count), one ordered key attribute (time), one categorical key attribute
- + derived attribute: layer ordering (quantitative)
- Task: analyze trends in time, find (maxmial) outliers
- How: derived position+geometry, length, color



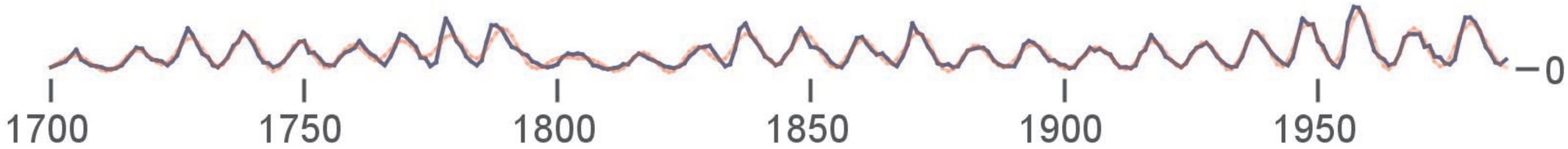
[Byron and Wattenberg, 2012]

Multiscale Banking

Aspect Ratio = 3.96

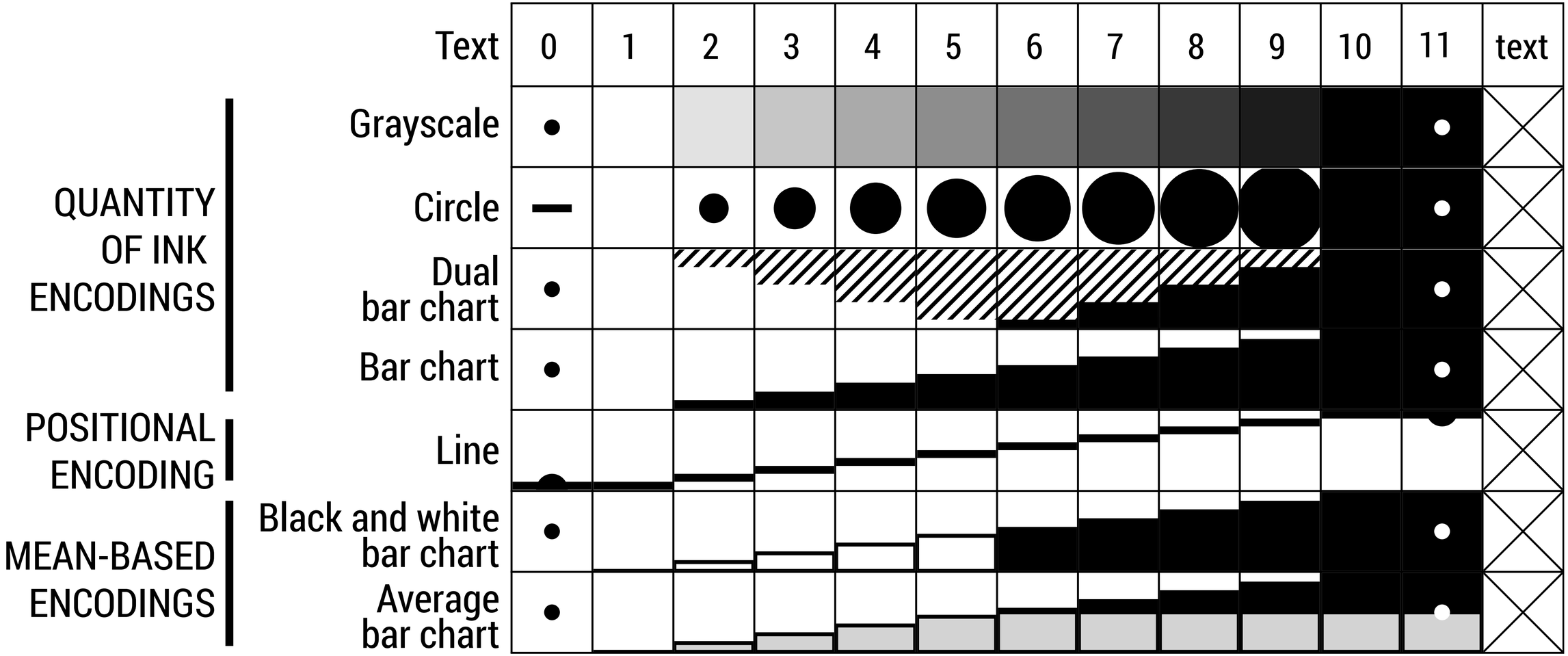


Aspect Ratio = 22.35



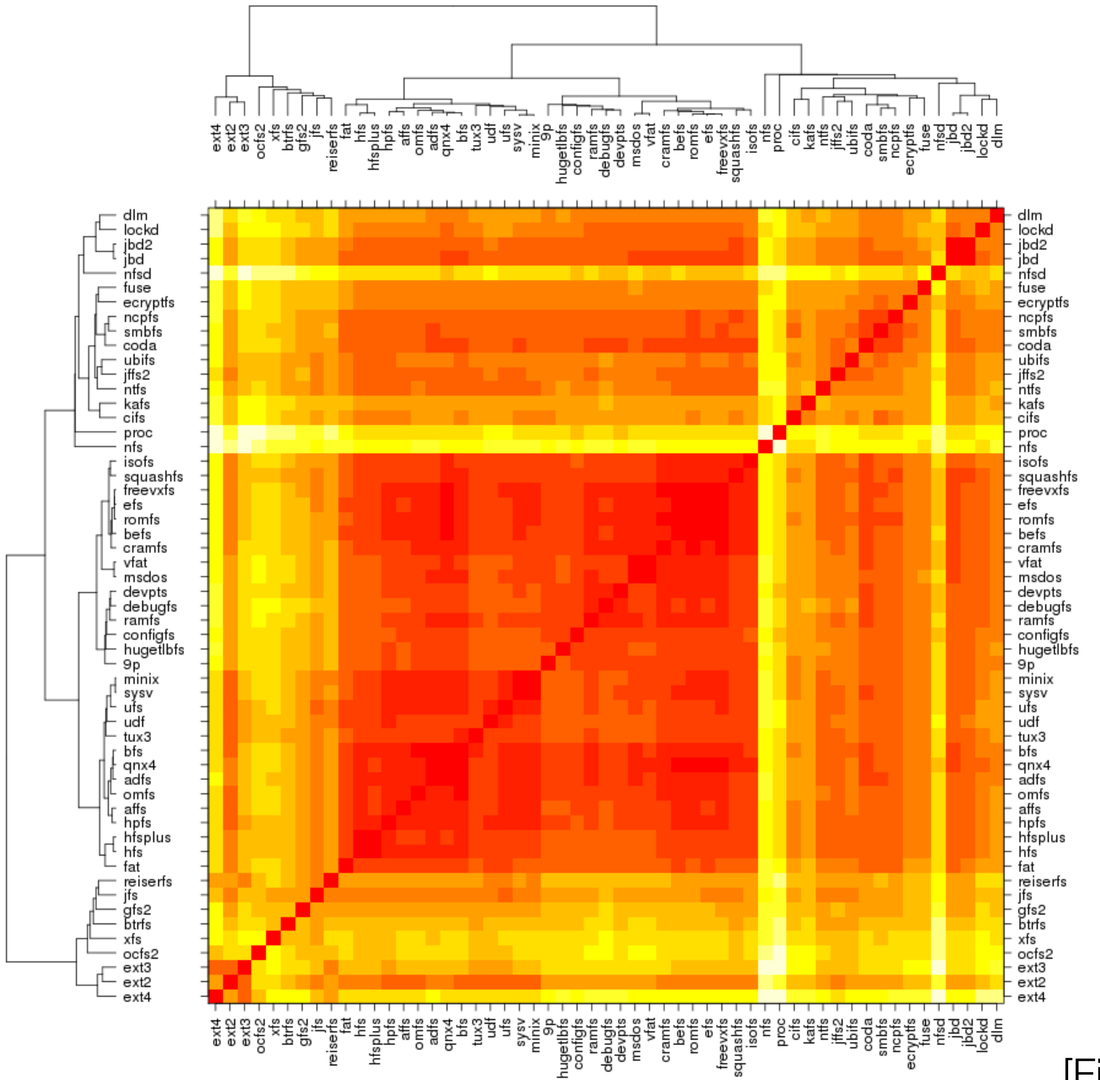
[Heer and Agrawala, 2006]

Bertin's Encodings



[C.Perrin et al., 2014]

Cluster Heatmap



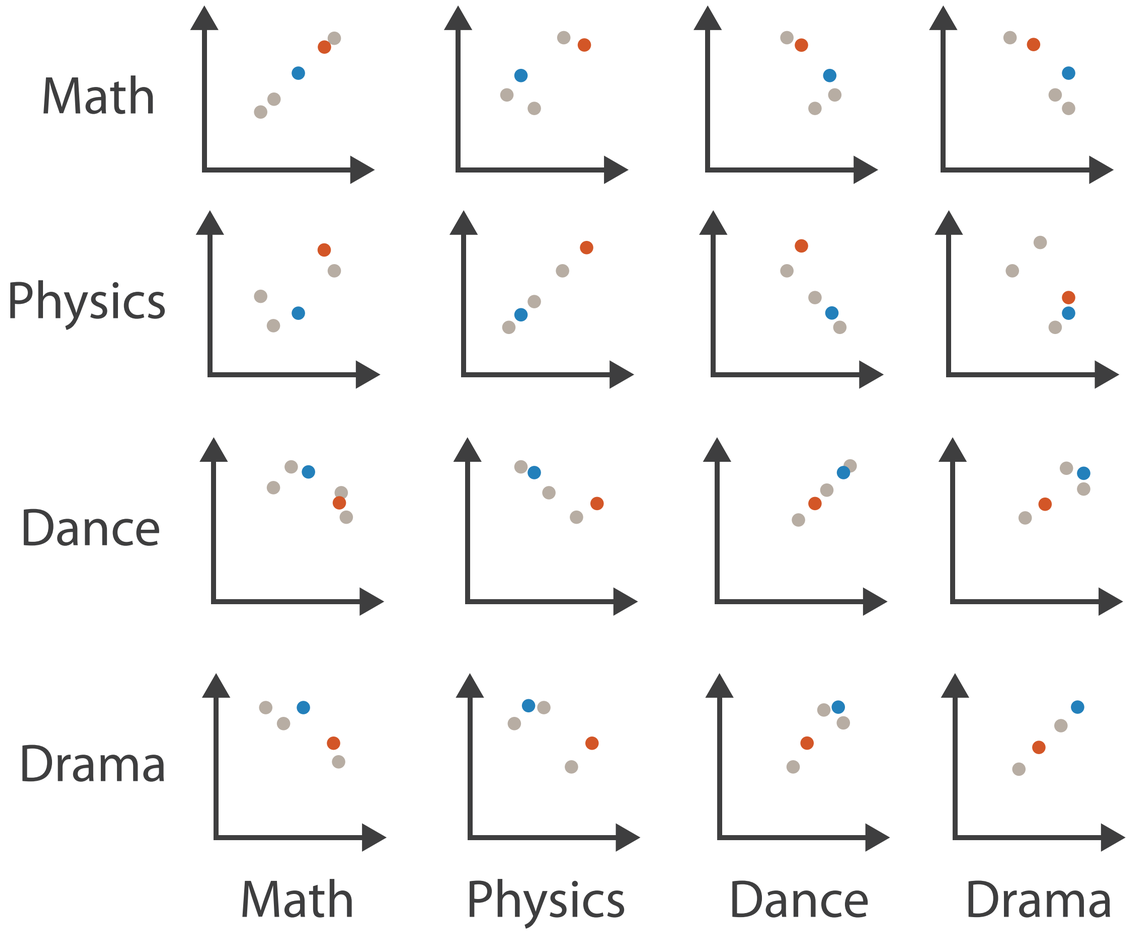
[File System Similarity, R. Musăloiu-E., 2009]

Scatterplot Matrices and Parallel Coordinates

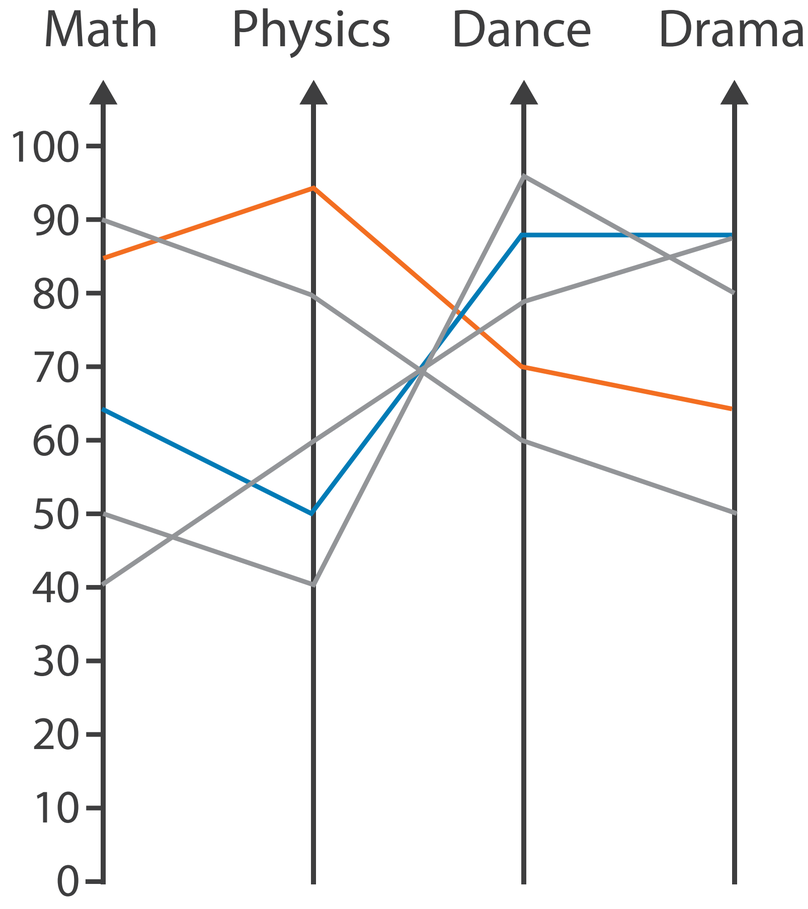
[Munzner (ill. Maguire), 2014]

Scatterplot Matrices and Parallel Coordinates

Scatterplot Matrix

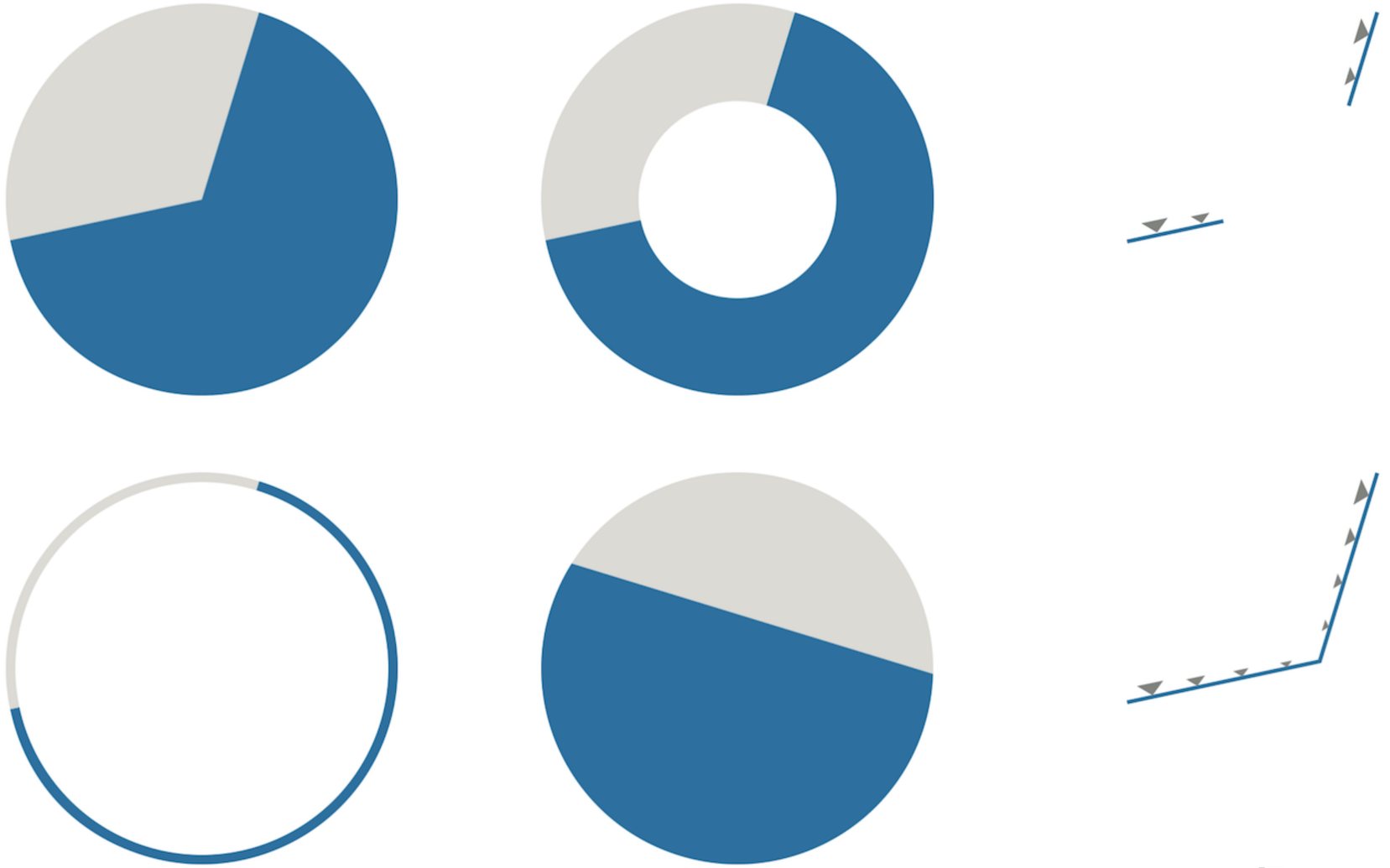


Parallel Coordinates



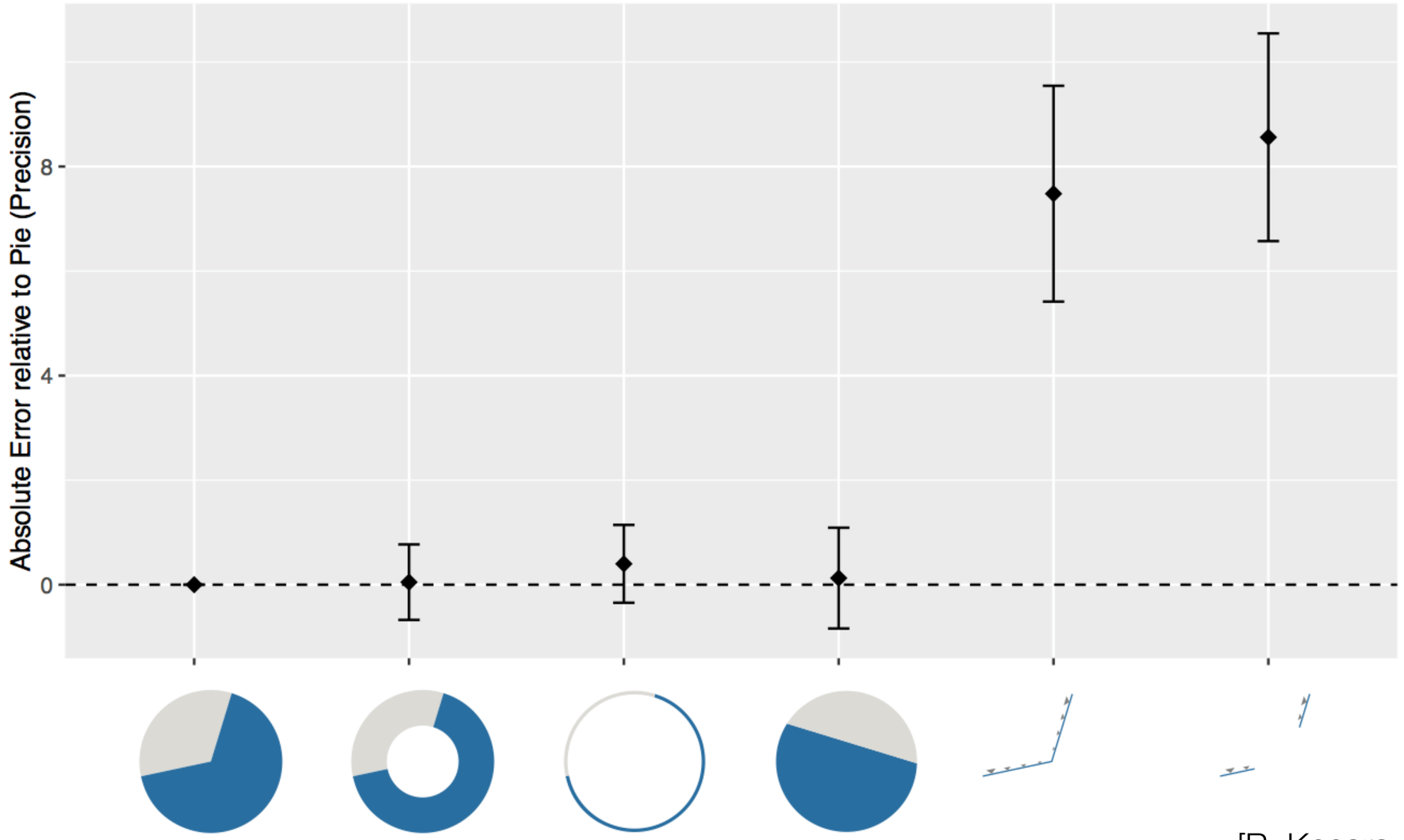
[Munzner (ill. Maguire), 2014]

Arcs, Angles, or Areas?



[R. Kosara and D. Skau, 2016]

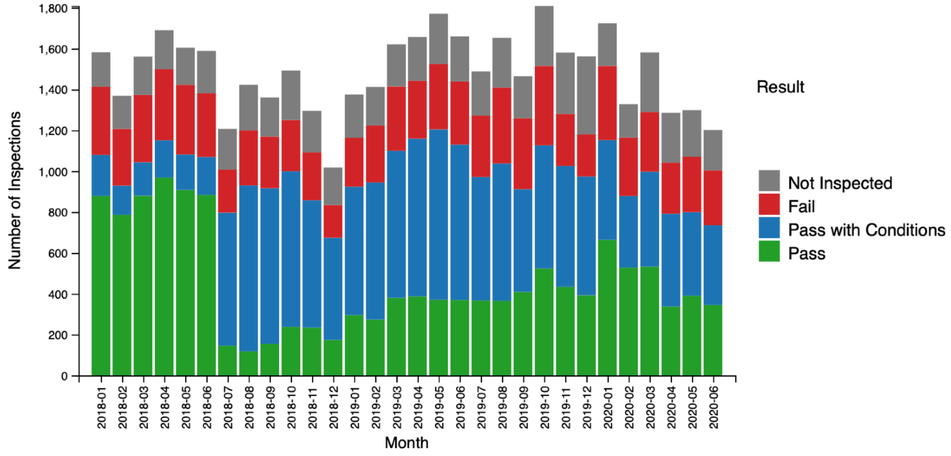
User Studies: Absolute Error Relative to Pie Chart



[R. Kosara and D. Skau, 2016]

Assignment 3

- Same stacked bar chart visualization
- Three tools
  - Tableau (free academic license)
  - Vega-Lite
  - D3
- For Vega-Lite, use the online editor
- For D3, use template files so the data is properly loaded
- [CS 490] Only need to do a standard bar chart in D3



Color

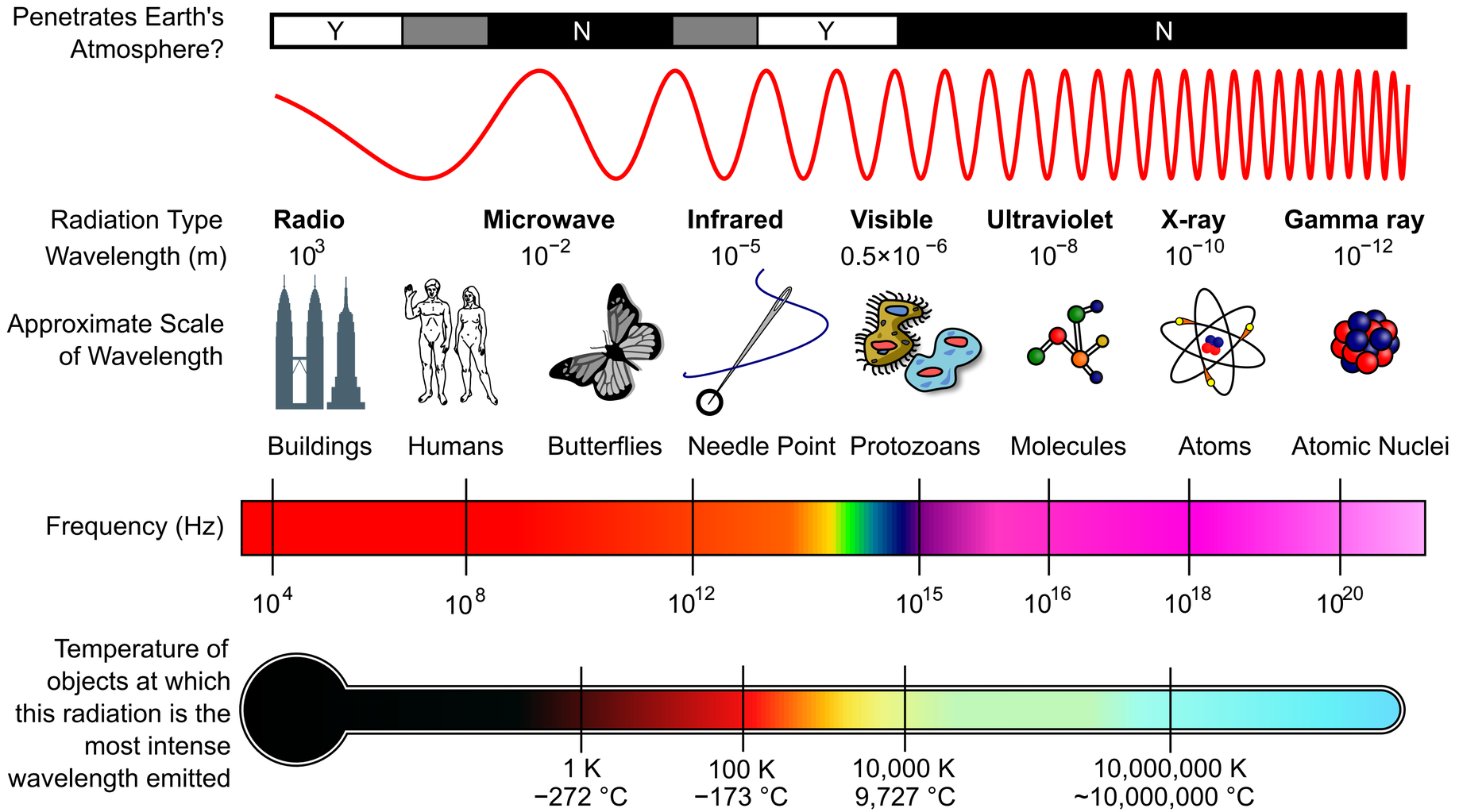
Color

Color

Color and Light

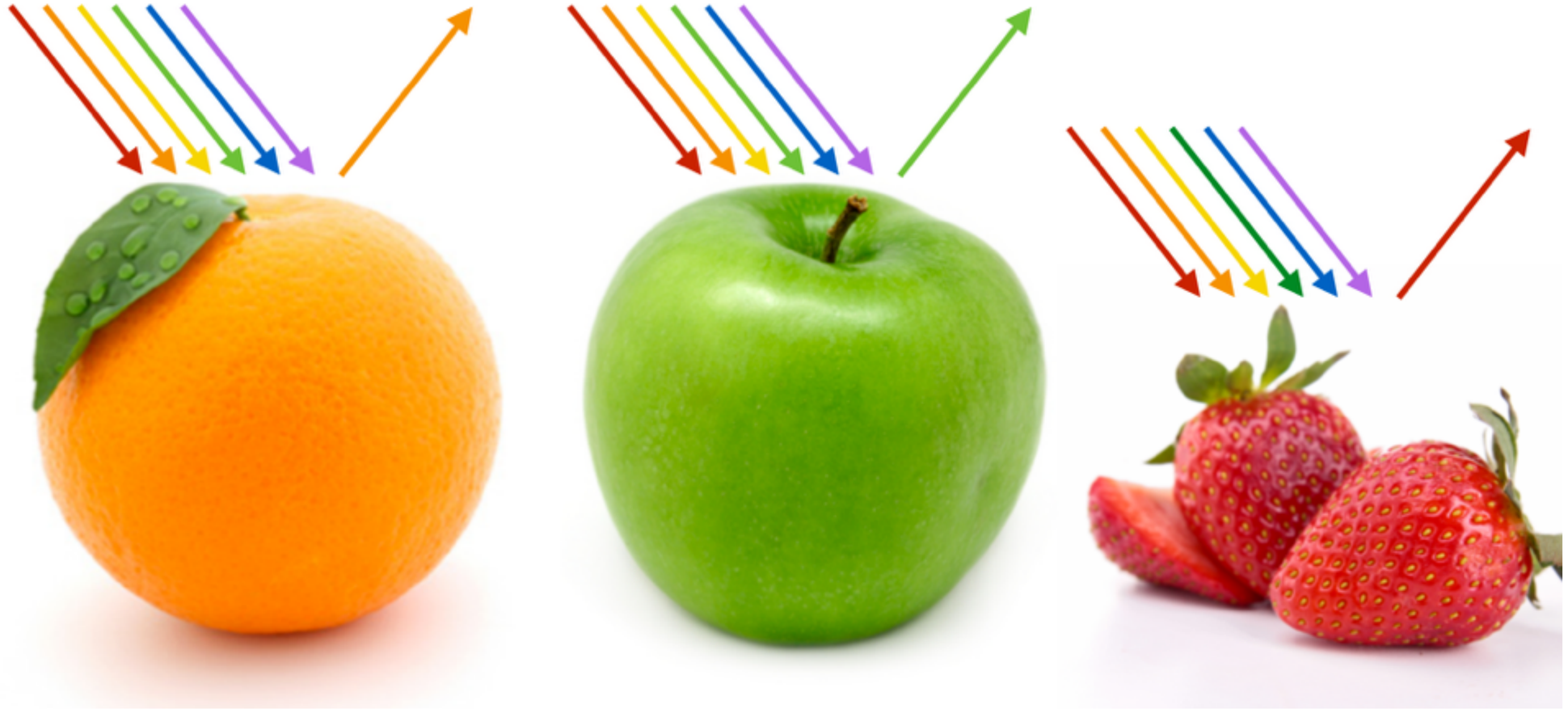
- Color is a **perceptive** property: color depends on the eyes and brain
- Visible light is a small portion of the **electromagnetic spectrum** which is composed of waves that at various frequencies (wavelengths), all traveling at the speed of light

Electromagnetic Spectrum



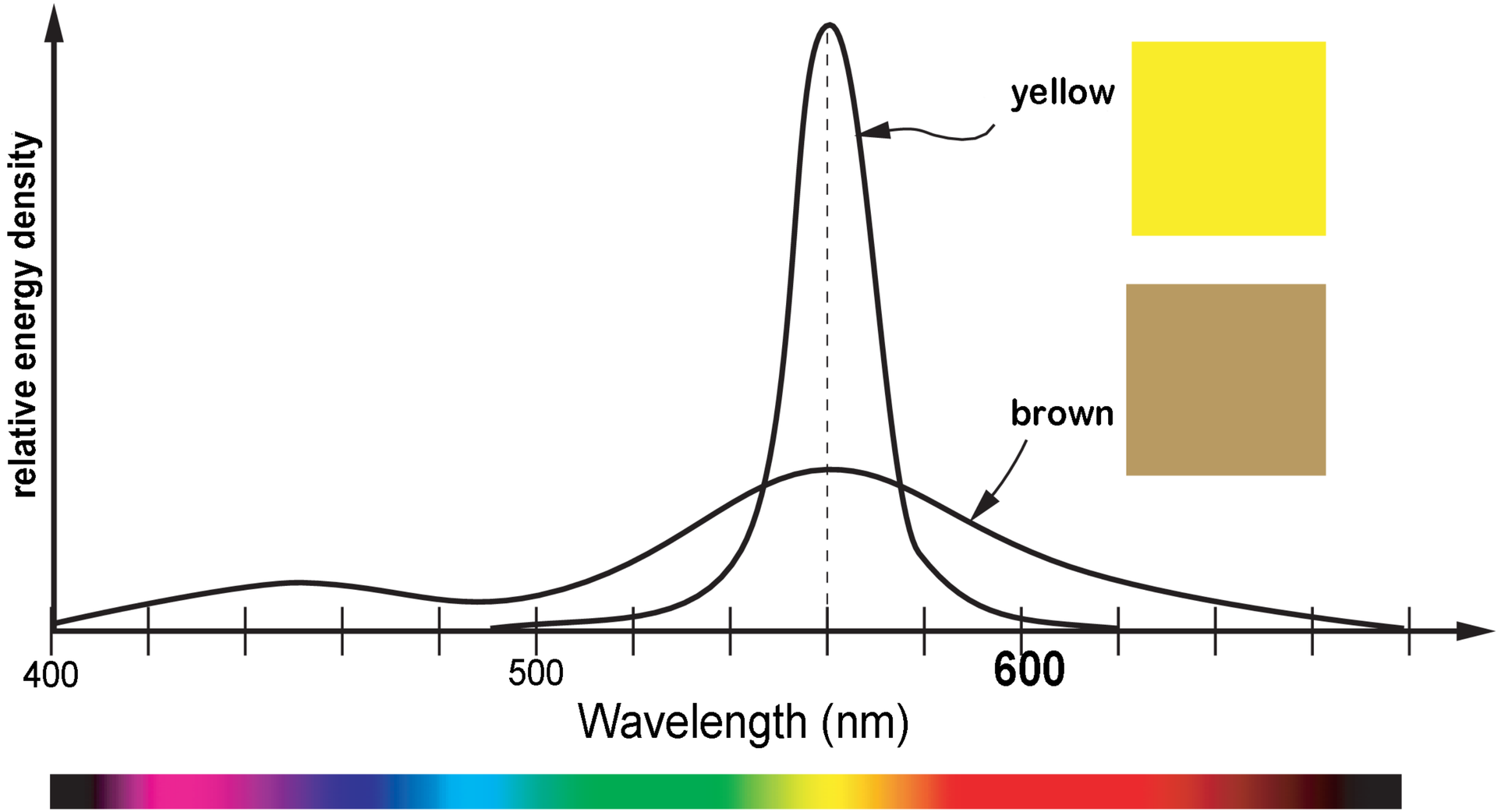
[Wikimedia, NASA]

Light Reflection & Absorption



[via M. Meyer]

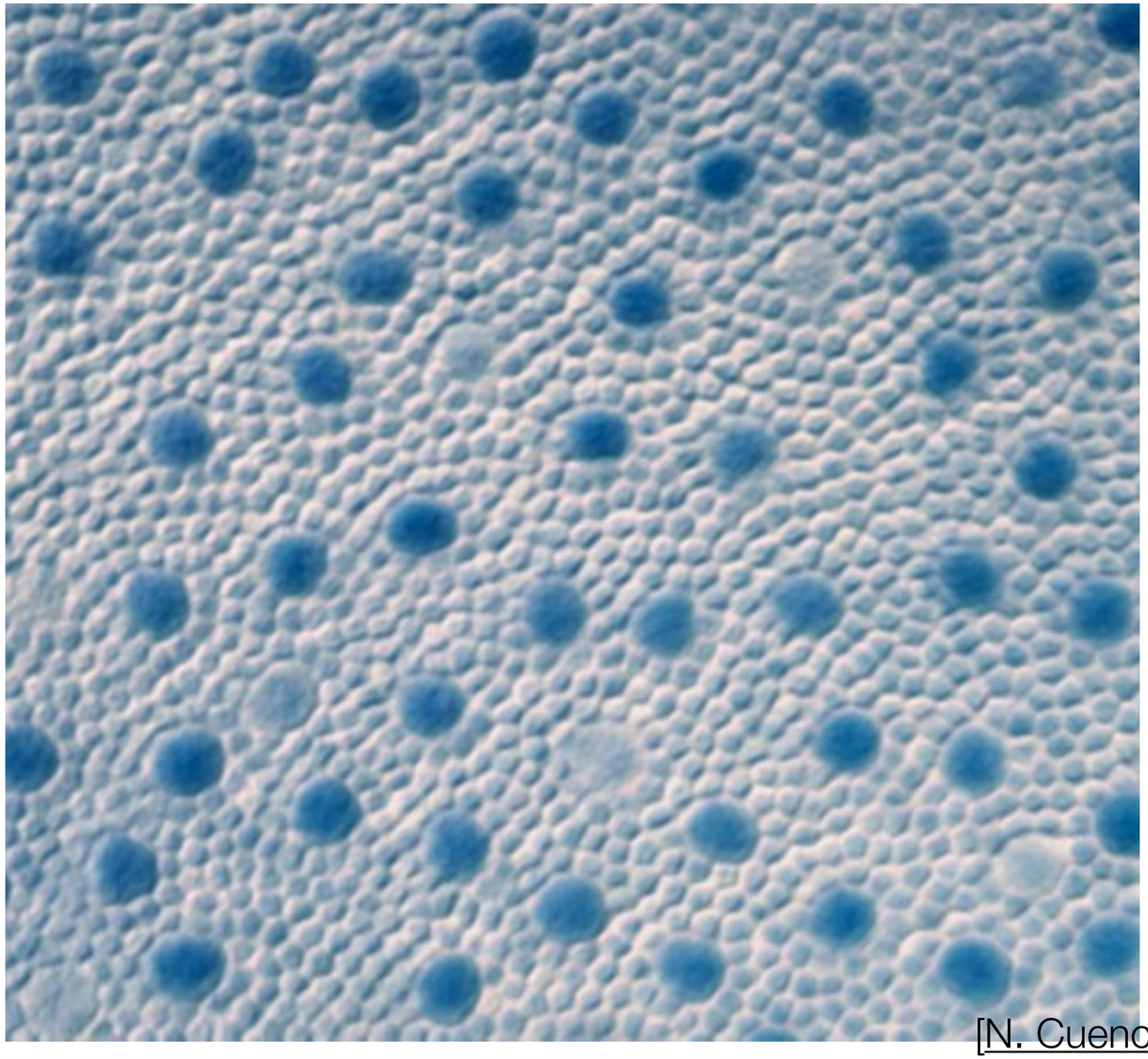
Color != Wavelength



[via M. Meyer]

Human Color Perception

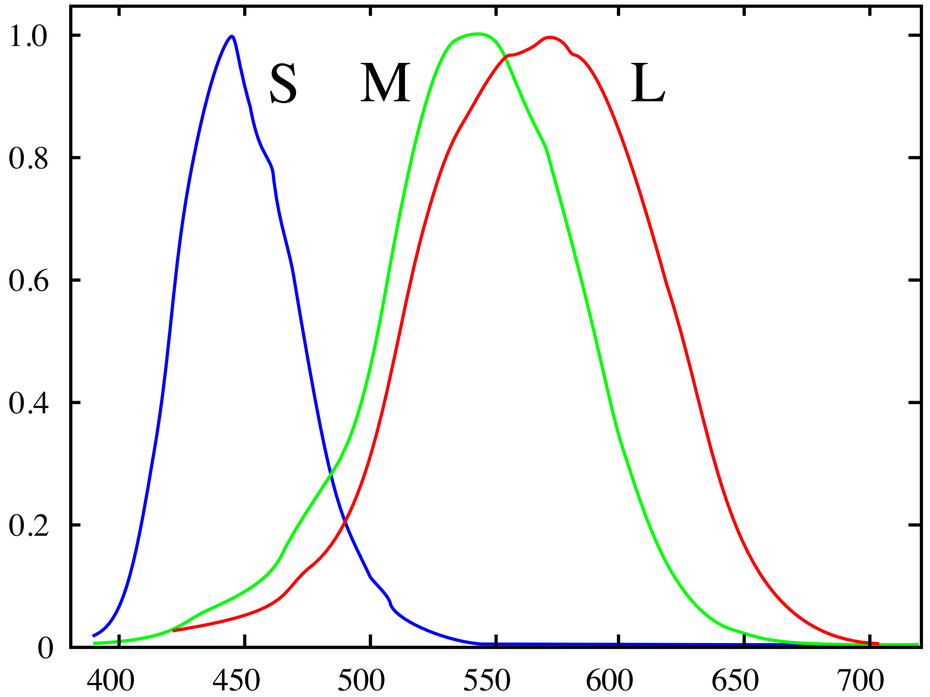
- Humans **do not** detect individual wavelengths of light
- Use **rods** and **cones** to detect light
  - rods capture intensity
  - cones capture color



[N. Cuenca]

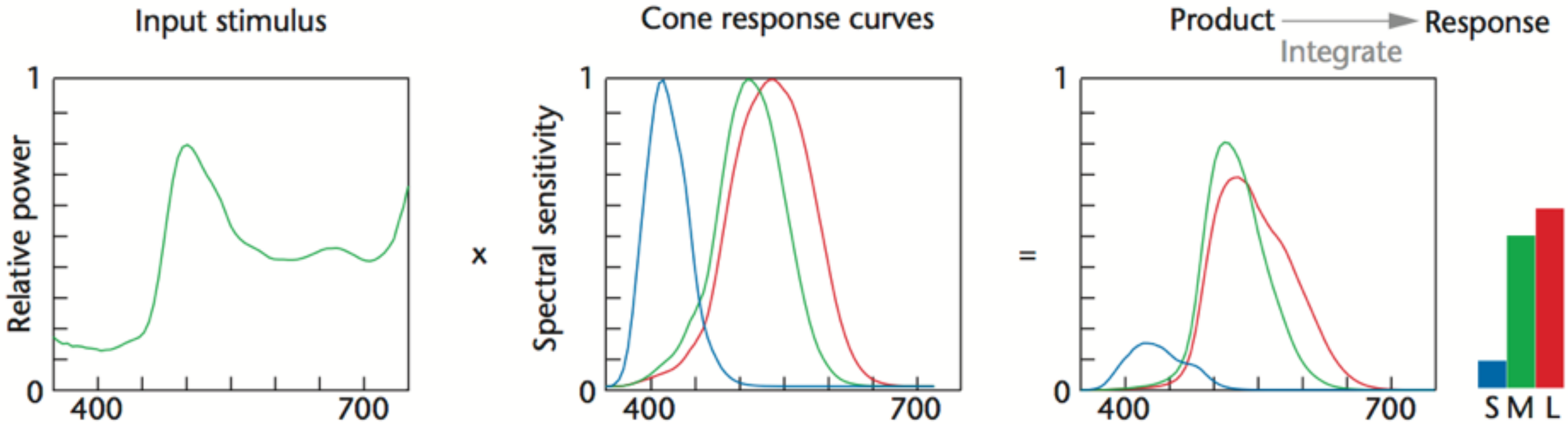
Human Color Perception

- Humans are **trichromatic**—we have three different types of cones
  - S (430nm): blue
  - M (540nm): green
  - L (570nm): "red"
- Note that the response curves **overlap**
- Spectra of visible light are "covered" by these responses
- Three numbers -> color



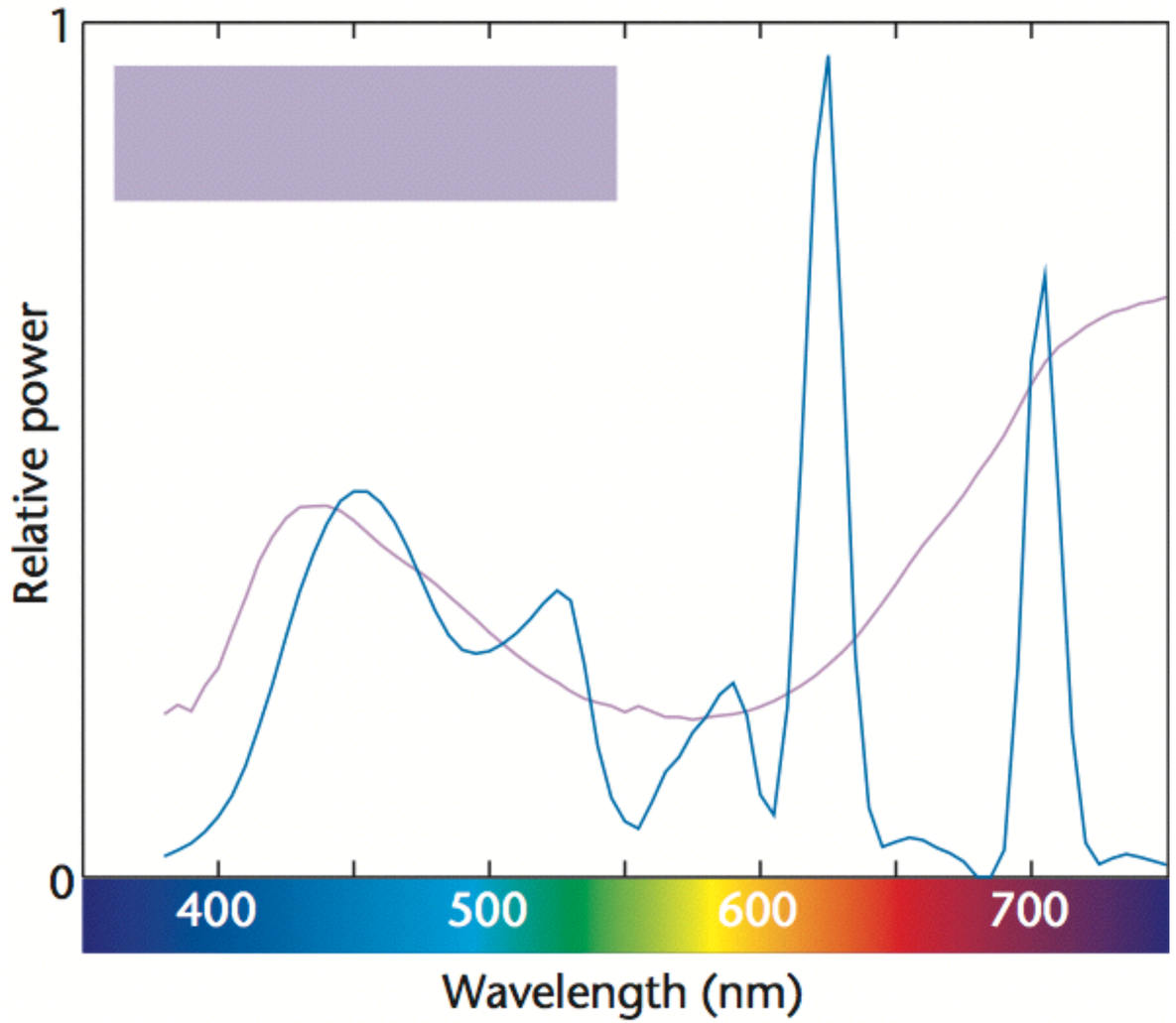
[Vanessaezekowitz at en.wikipedia]

Human Color Perception



[via M. Meyer]

Metamerism



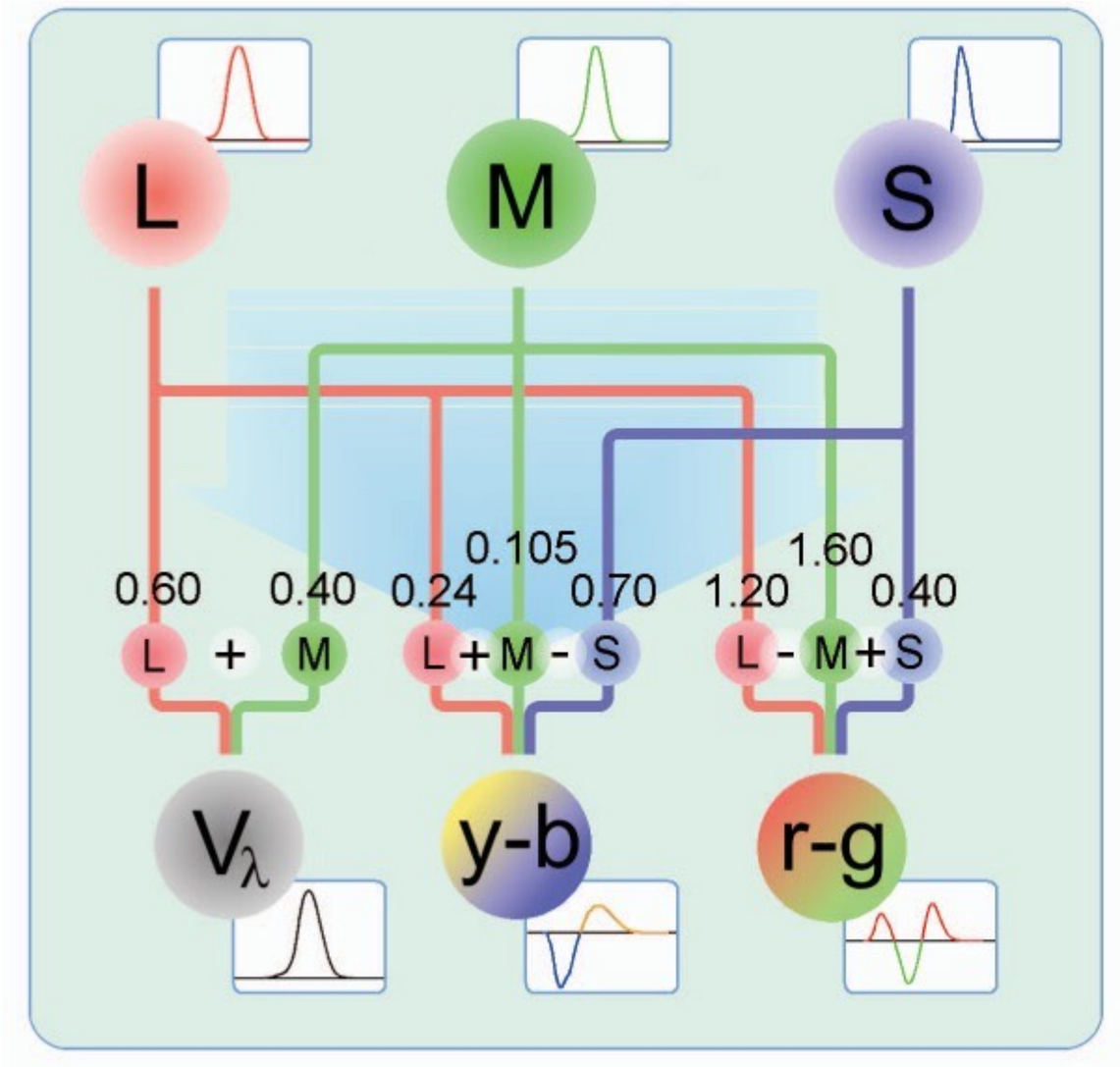
- Same responses == same color
- Humans are not spectrometers
- Do not get the whole function
- Three responses

[via M. Meyer]

Color

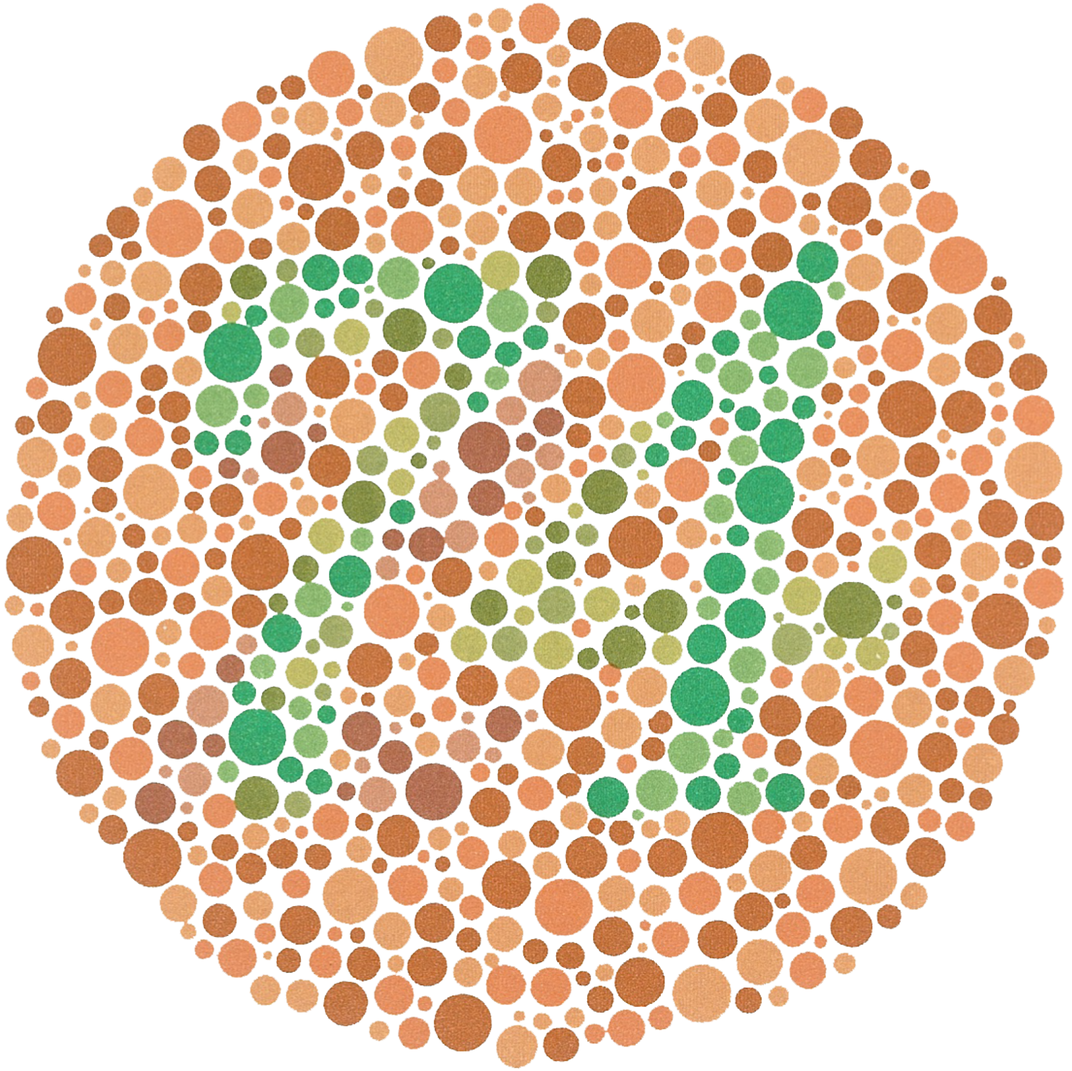
- Cones respond to different areas of the visible light spectrum
- Cover all wavelengths but certain wavelengths generate greater responses
- Color is determined by calculations based on the responses from the different cones
- Opponent Process Theory: three "opponent" channels
  - Light/Dark
  - Blue/Yellow
  - Red/Green
- Opposite colors are not perceived together

Opponent Process Theory



[Machado et. al, 2009]

Color Blindness

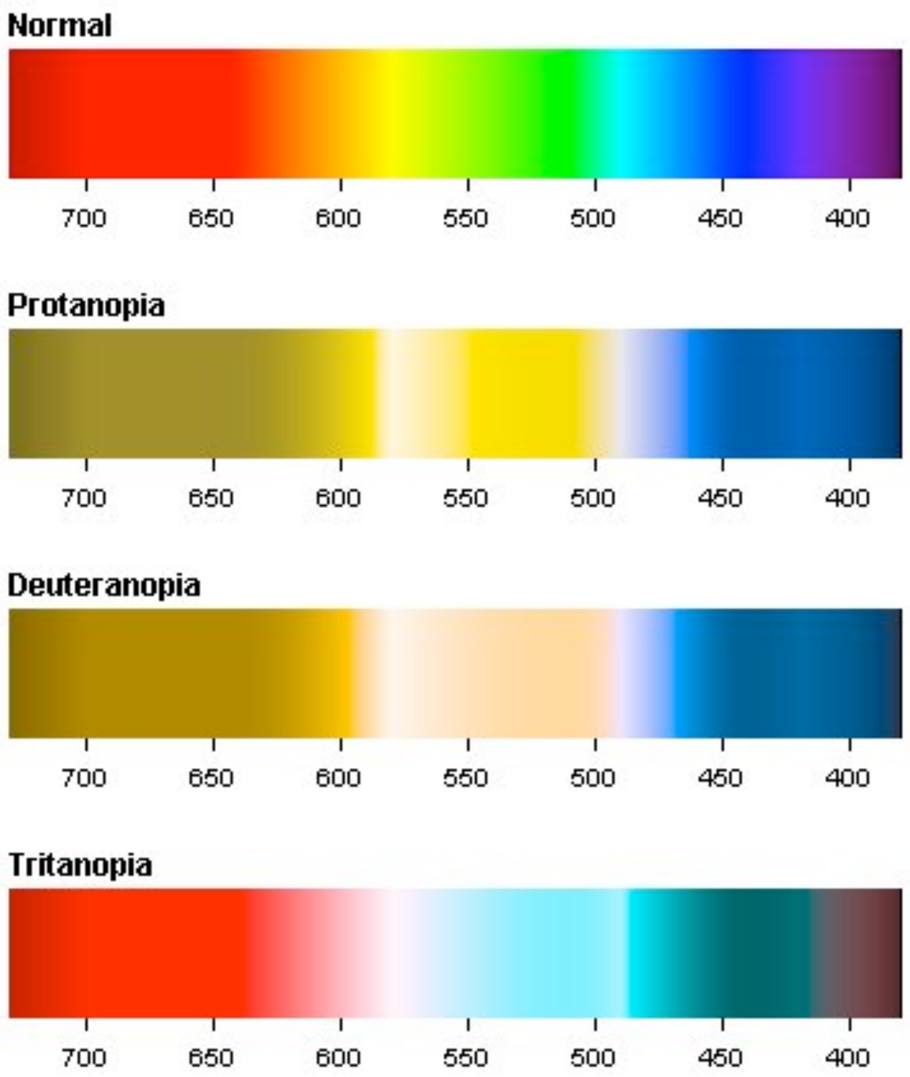


[Ishihara (Plate 9) via Wikipedia]

Color Blindness

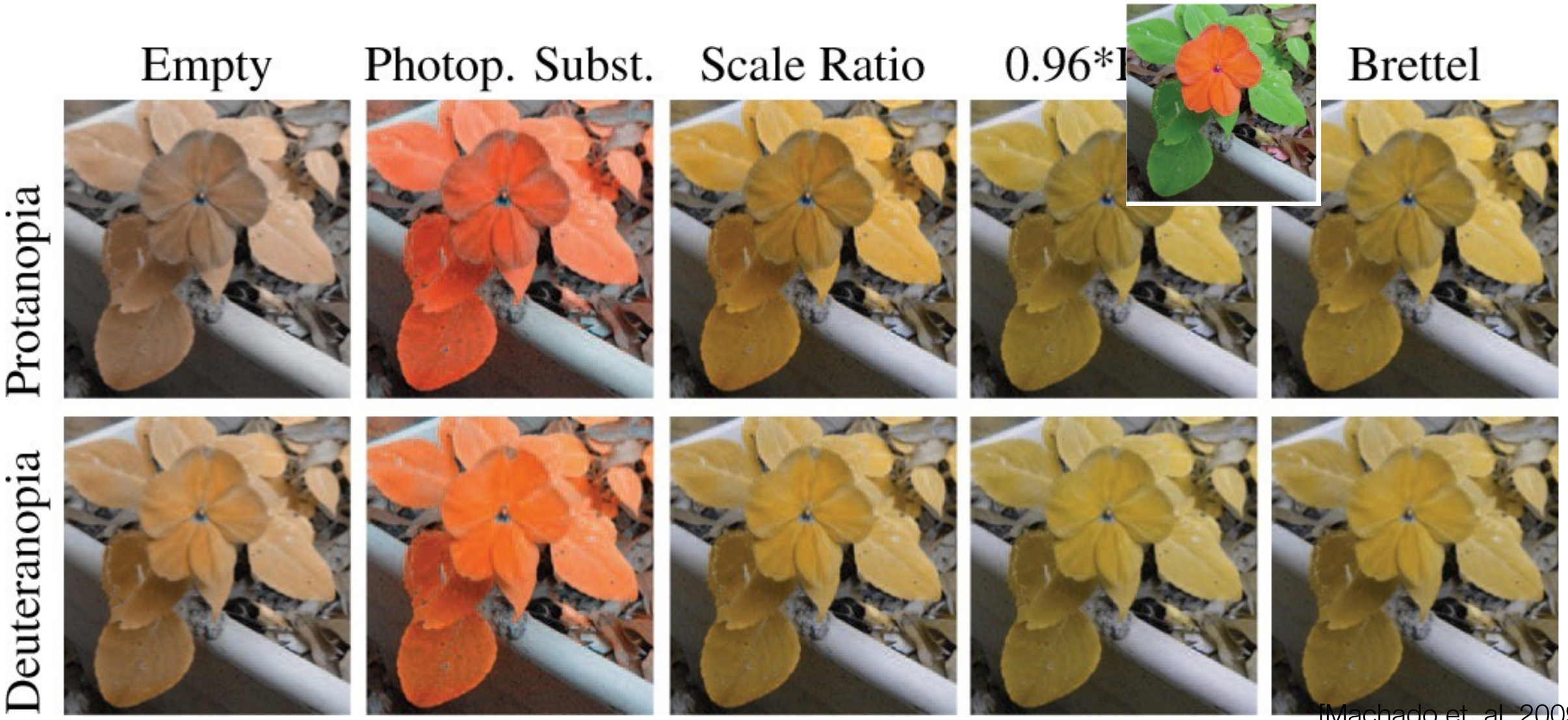
- Sex-linked: 8% of males and 0.4% of females of N. European ancestry
- Abnormal distribution of cones (e.g. missing the S, M, or L types)
- Either dichromatic (only two types of cones) or anomalous trichromatic (one type of cones has a defect)
  - Protanopia (L missing), Protanomaly (L defect)
  - Deuteranopia (M missing), Deuteranomaly (M defect) [Most Common]
  - Tritanopia (S missing), Tritanomaly (S defect) [Rare]
- Dichromacy is rarer than anomalous trichromacy
- Opponent process model explains why colors cannot be differentiated

Color Blindness



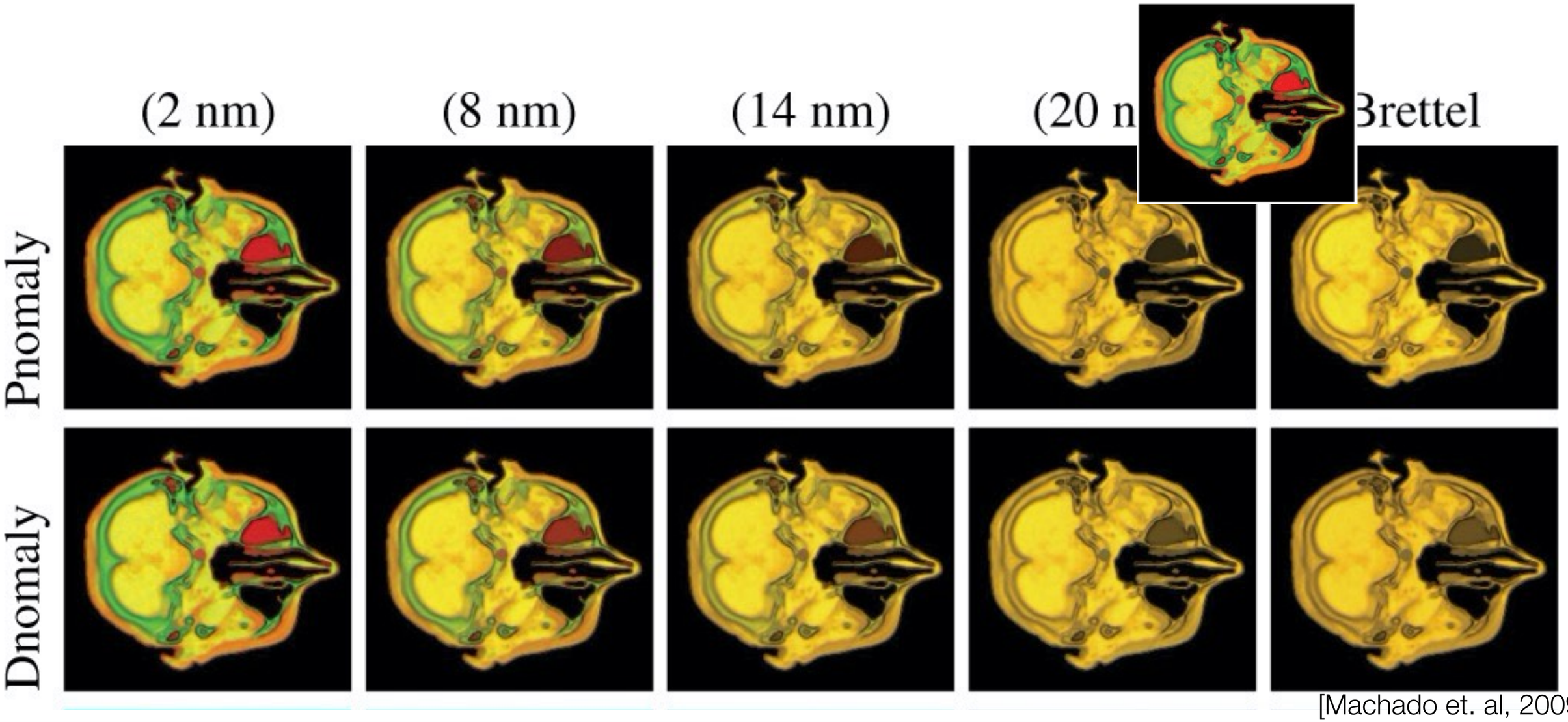
[via M. Meyer]

Simulating Color Blindness



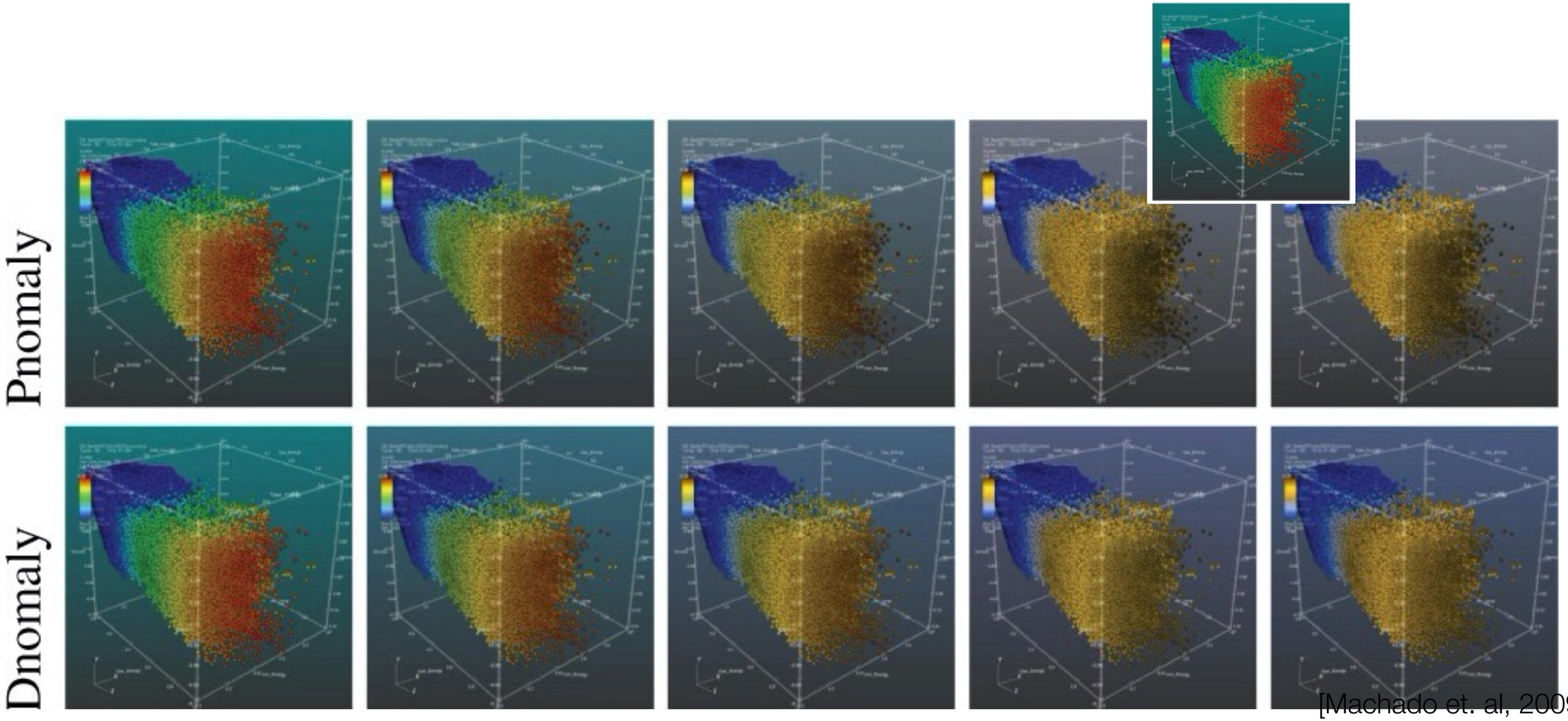
[Machado et. al, 2009]

Simulating Color Blindness



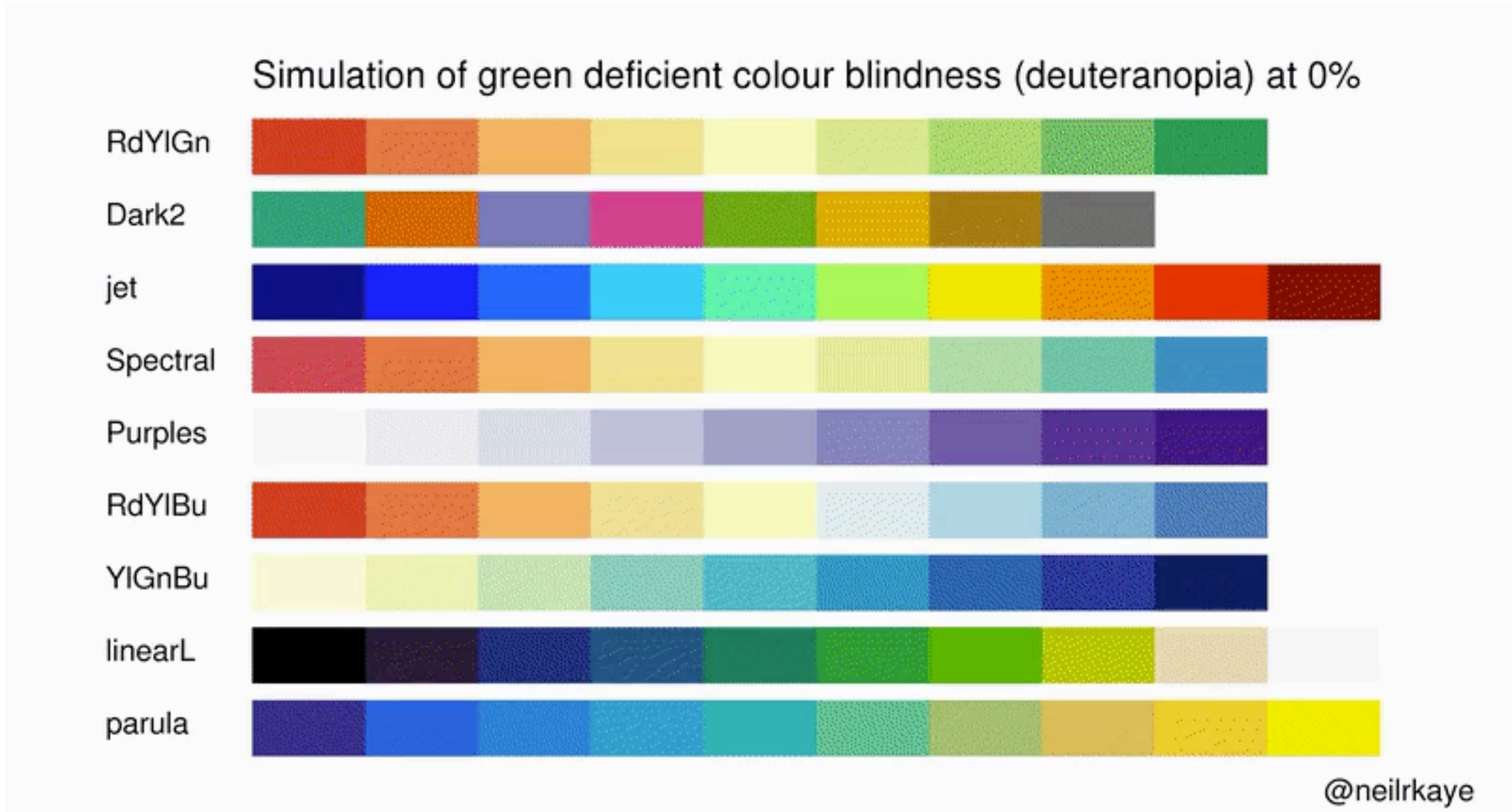
[Machado et. al, 2009]

Simulating Color Blindness



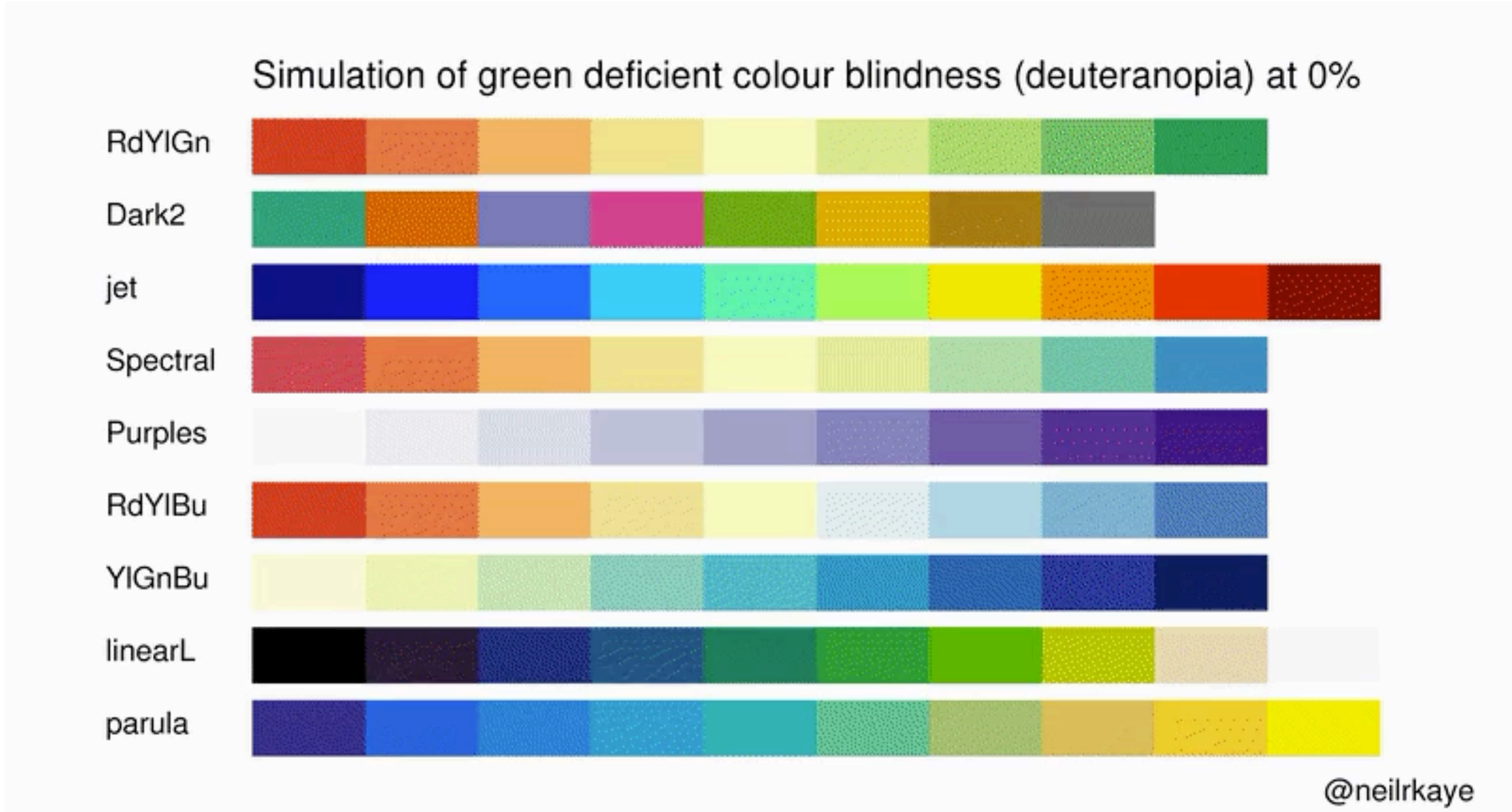
[Machado et. al, 2009]

Simulating Deuteranopia (Colormaps)



[@neilrkaye, reddit]

Simulating Deuteranopia (Colormaps)



[@neilrkaye, reddit]

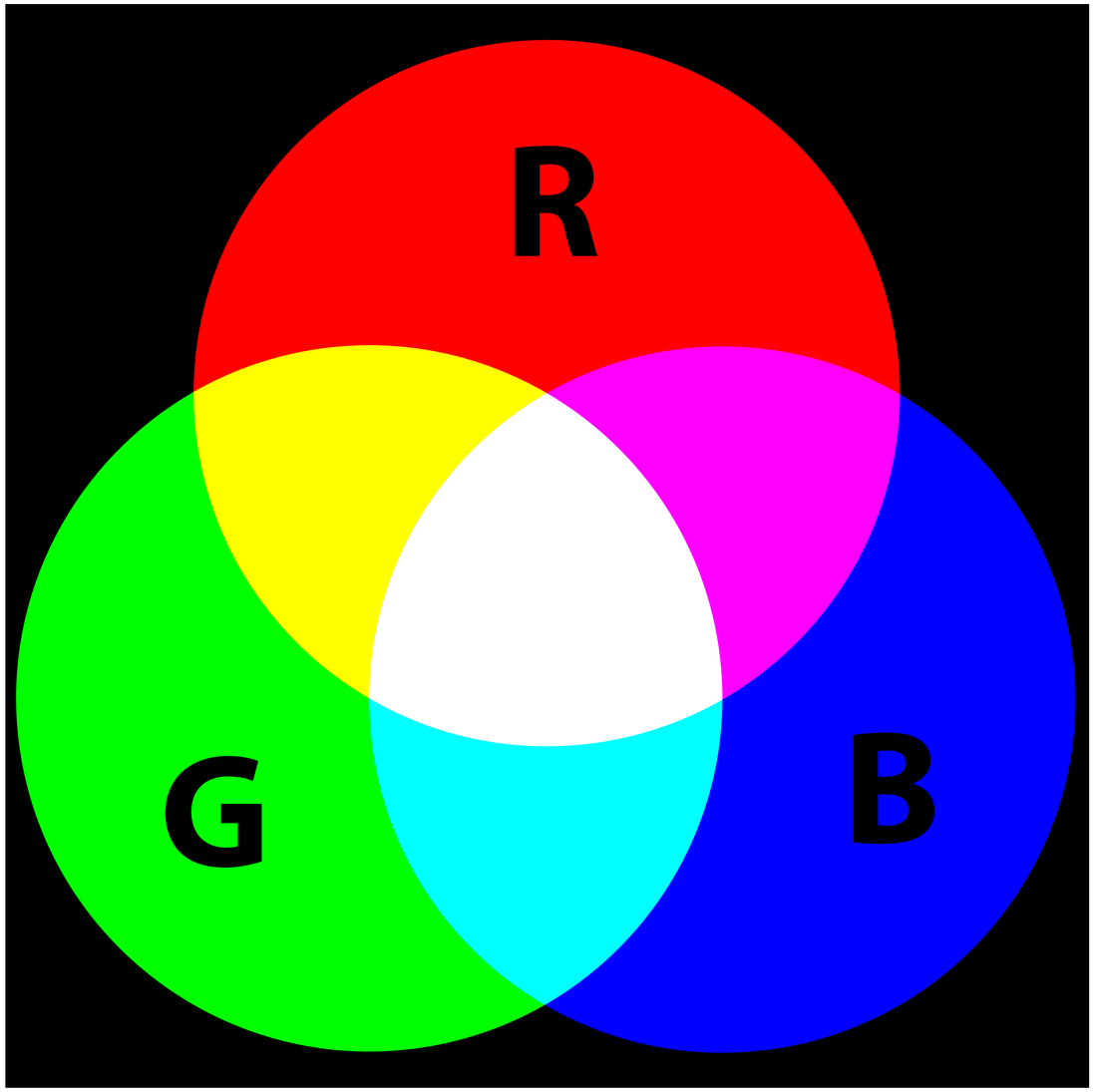
Primary Colors?

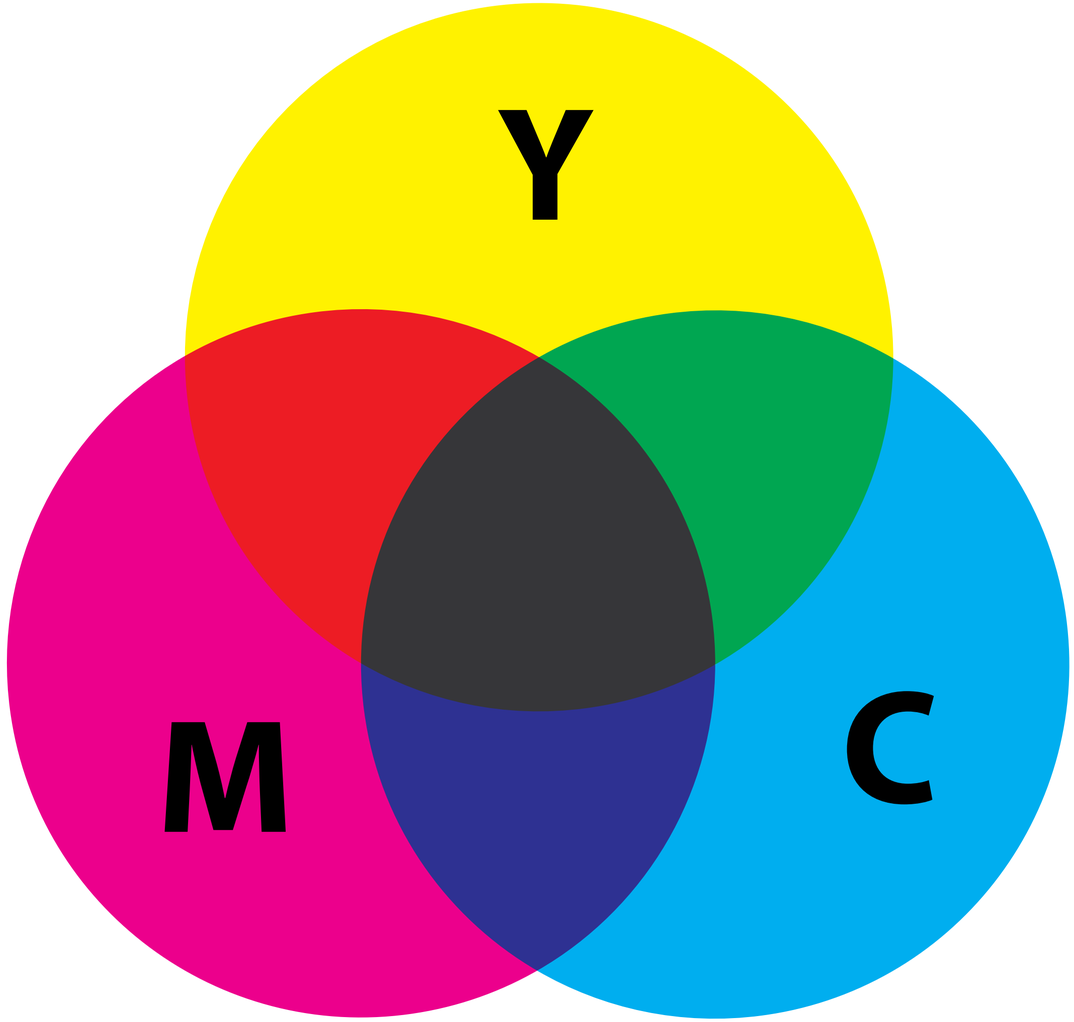
- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow

Primary Colors?

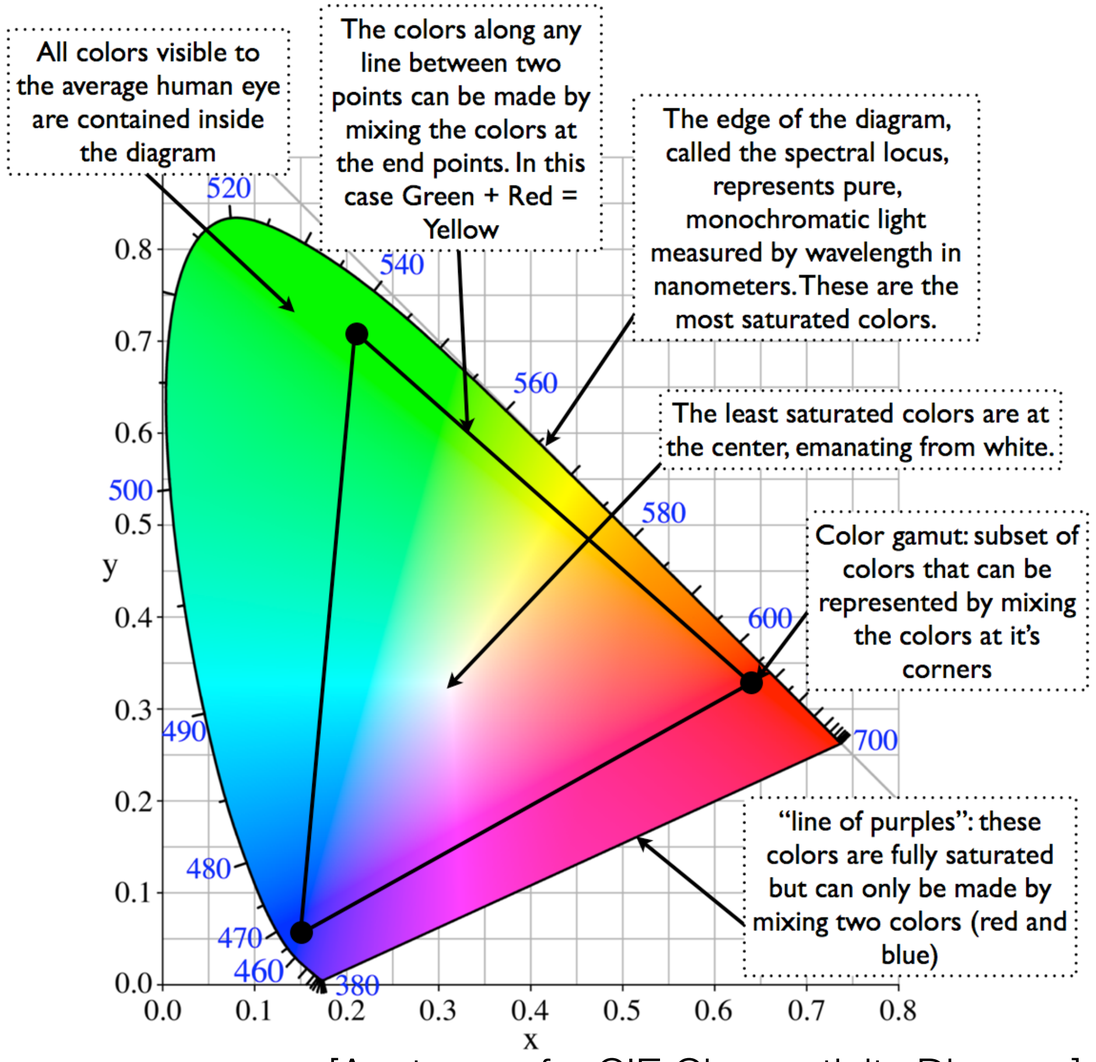
- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
- **All of the above!**

Color Addition and Subtraction





Color Spaces and Gamuts



[Anatomy of a CIE Chromaticity Diagram]

- **Color space**: the organization of all colors in space
  - Often human-specific, what we can see (e.g. CIELAB)
- **Color gamut**: a subset of colors
  - Defined by corners of color space
  - What can be produced on a monitor (e.g. using RGB)
  - What can be produced on a printer (e.g. using CMYK)
  - The gamut of your monitor != the gamut of someone else's or a printer

Color Models

- A **color model** is a representation of color using some basis
- RGB uses three numbers (red, blue, green) to represent color
- Color space ~ color model, but there can be many color models used in the same color space (e.g. OGV)
- Hue-Saturation-Lightness (HSL) is more intuitive and useful
  - Hue captures pure colors
  - Saturation captures the amount of white mixed with the color
  - Lightness captures the amount of black mixed with a color
  - HSL color pickers are often circular
- Hue-Saturation-Value (HSV) is similar (swap black with gray for the final value), linearly related

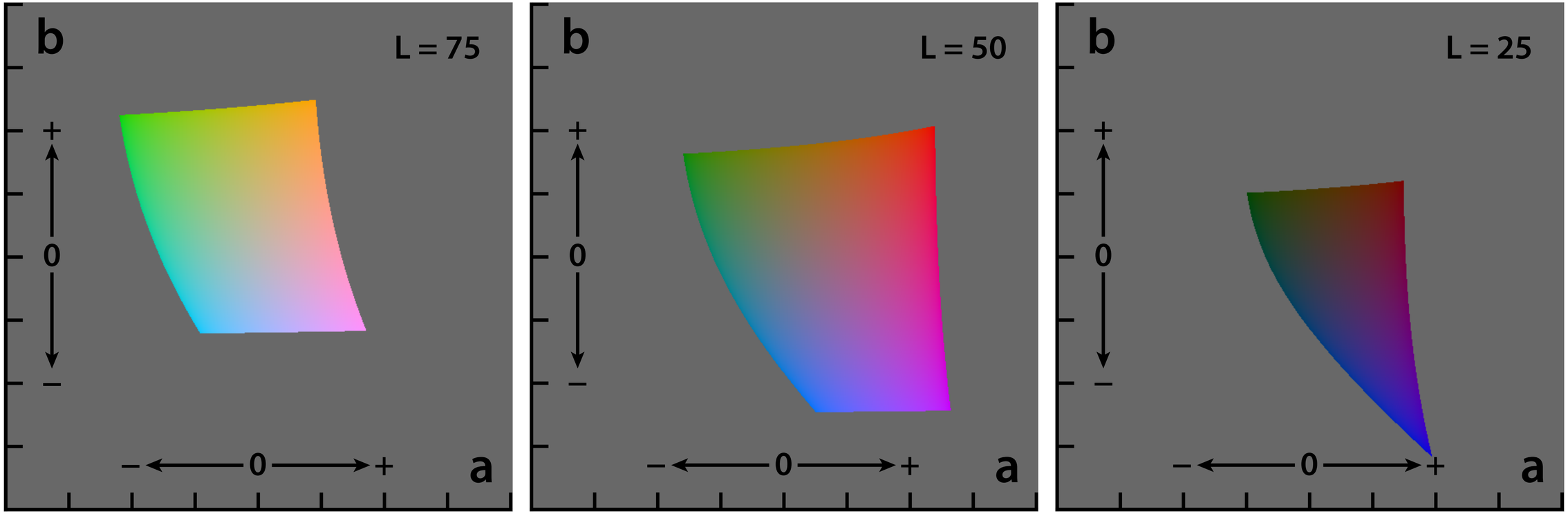
Luminance

- HSL does not truly reflect the way we perceive color
- Even though colors have the same lightness, we perceive their luminance differently
- Our perception (L*) is **nonlinear**

[Munzner (ill. Maguire), 2014 (based on Stone, 2006)]

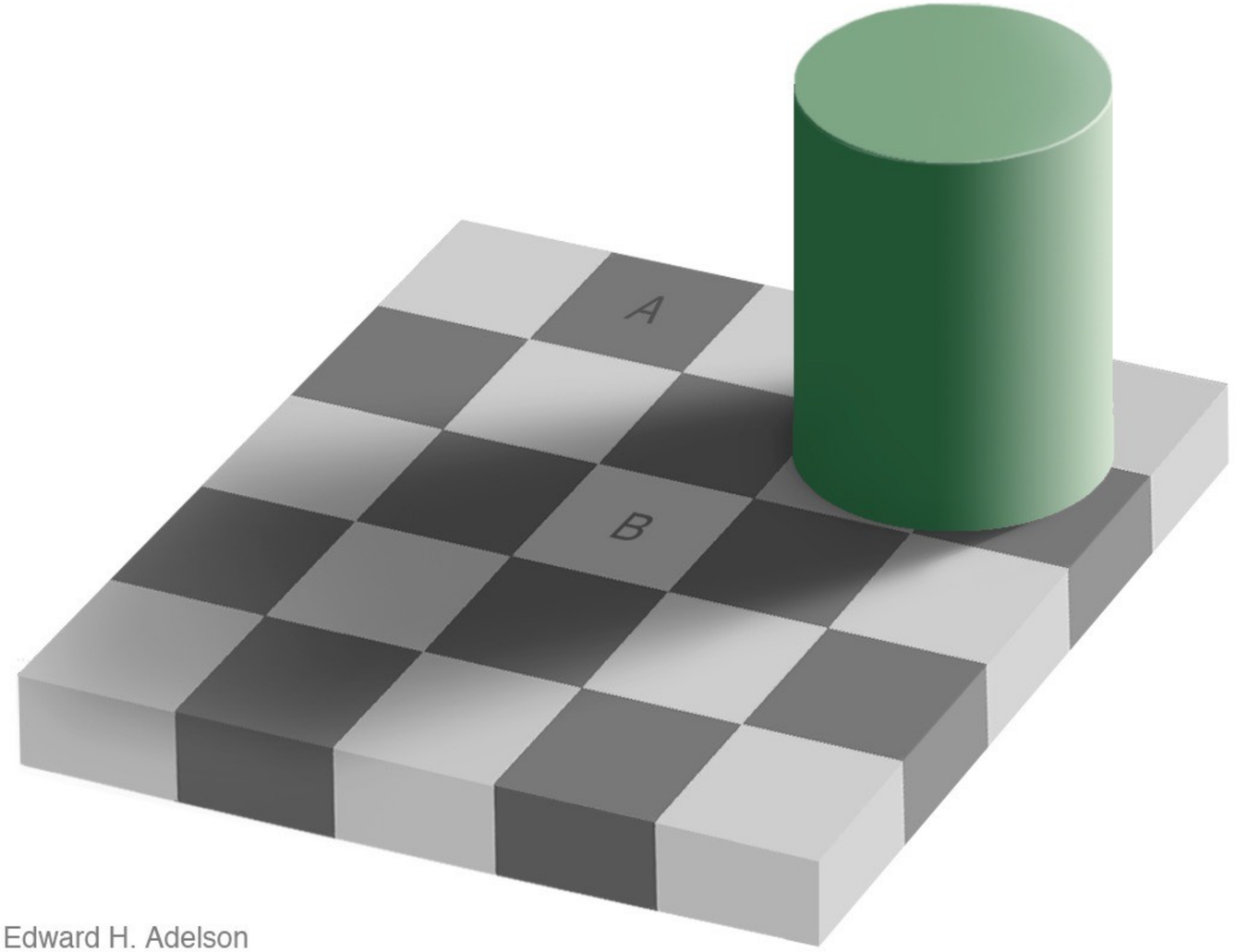
Perceptually Uniform Color Spaces

- L*a*b* allows perceptually accurate comparison and calculations of colors



[J. Rus, CC-BY-SA (changed to horizontal layout)]

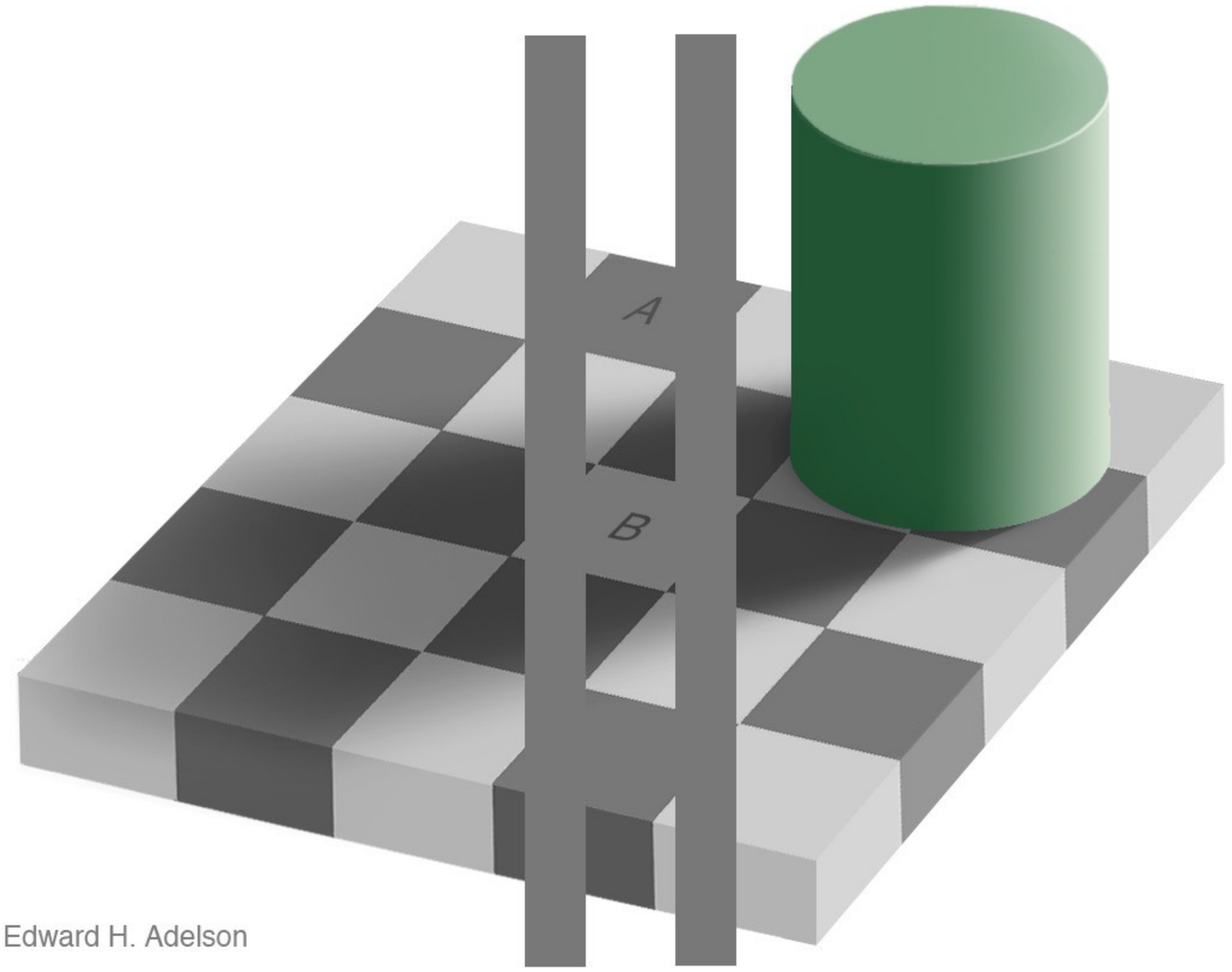
Luminance Perception (Spatial Adaption)



Edward H. Adelson

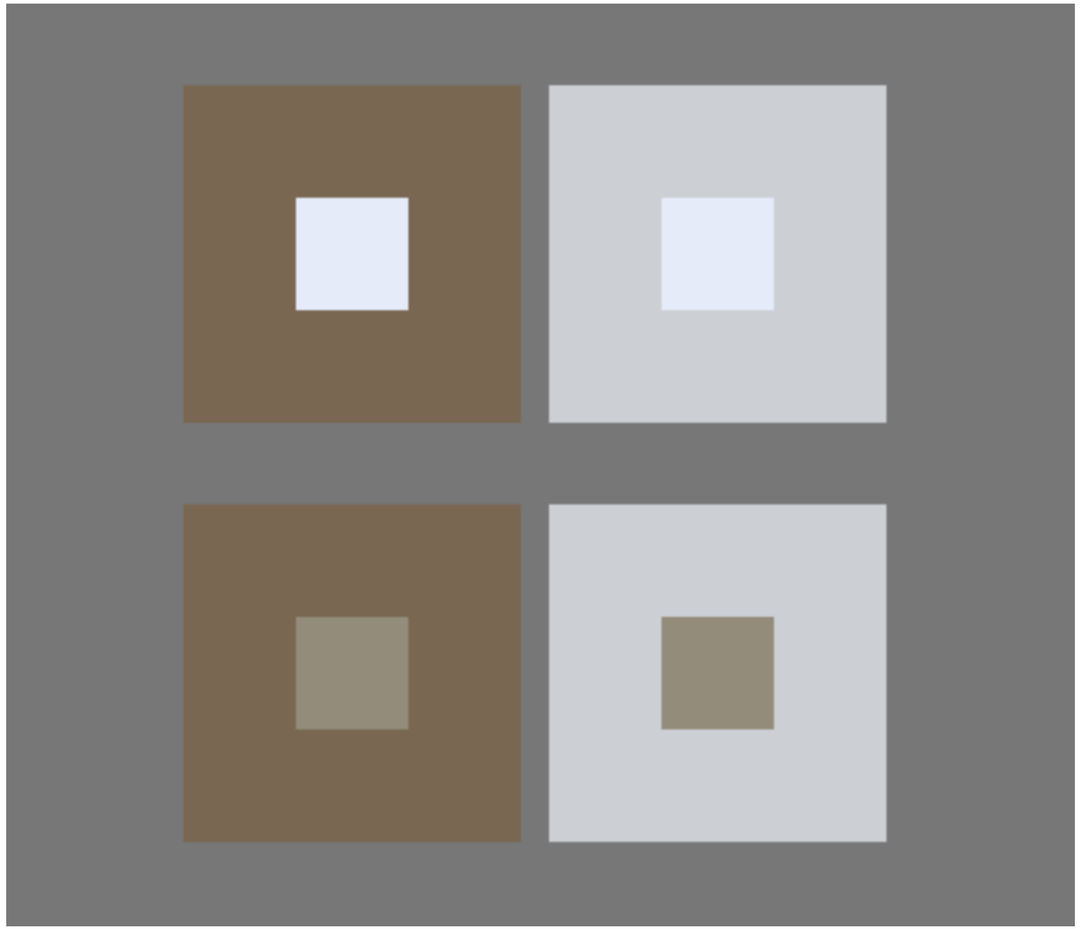
[E. H. Adelson, 1995]

Luminance Perception (Spatial Adaption)

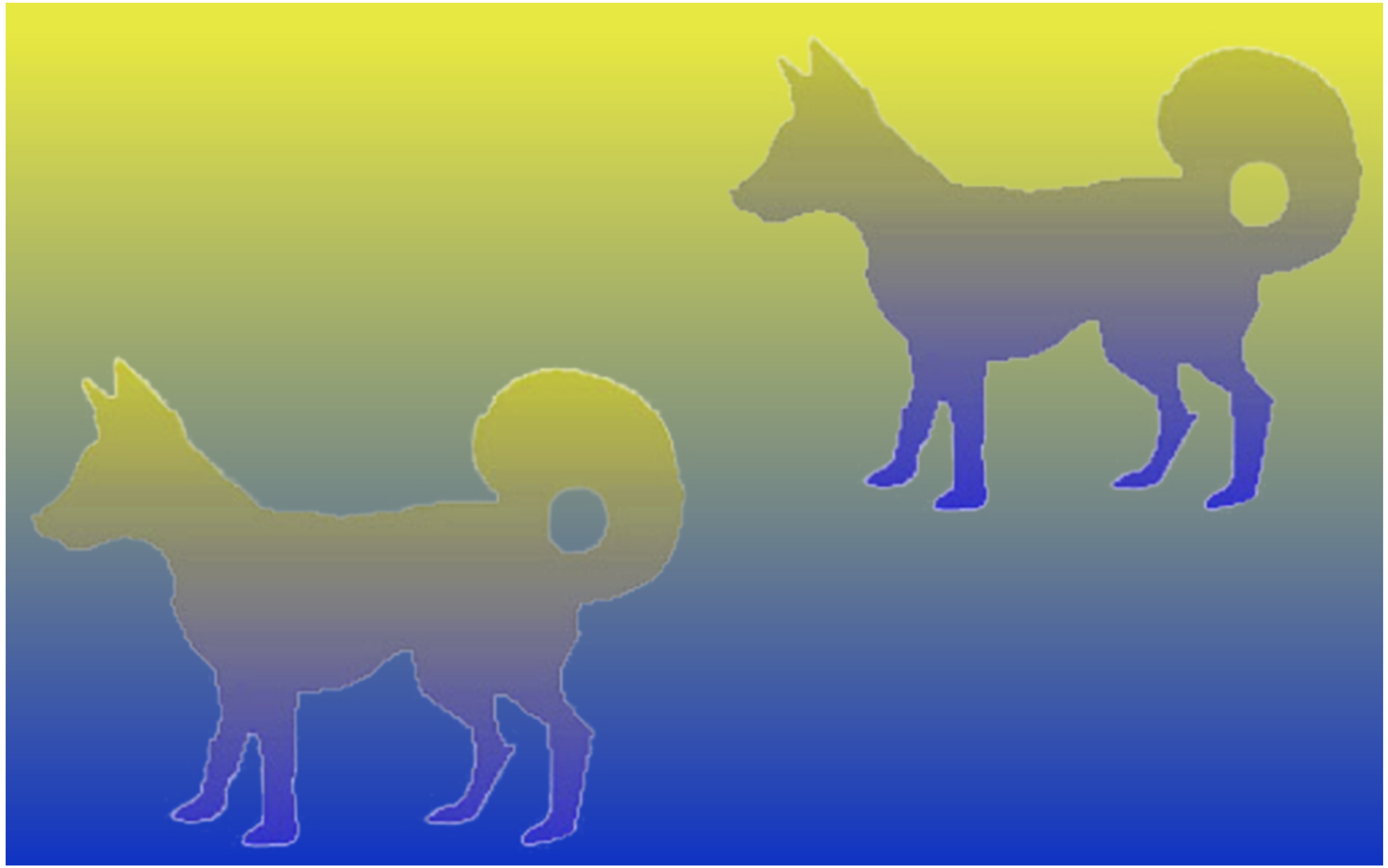


[E. H. Adelson, 1995]

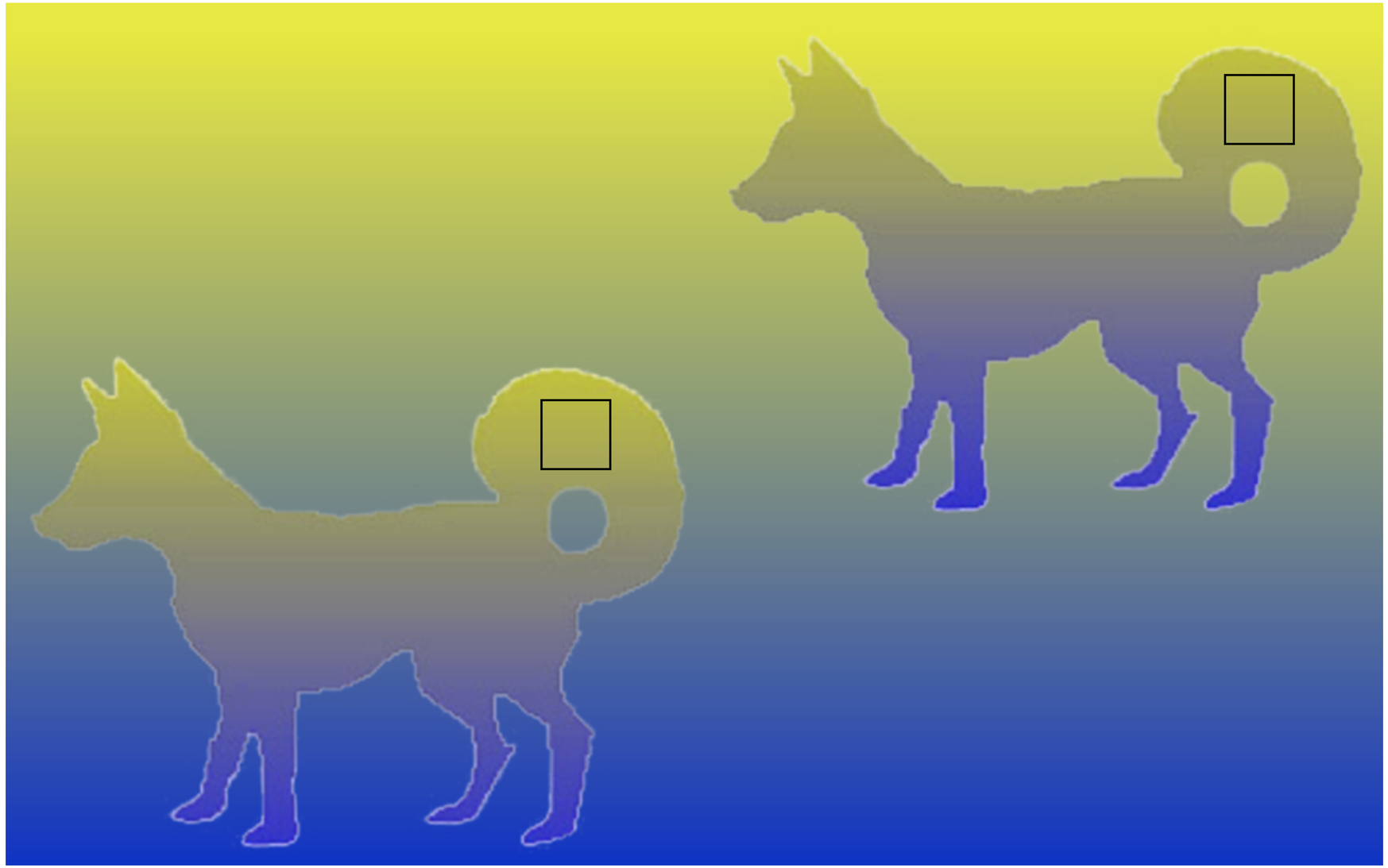
Simultaneous Contrast



Simultaneous Contrast

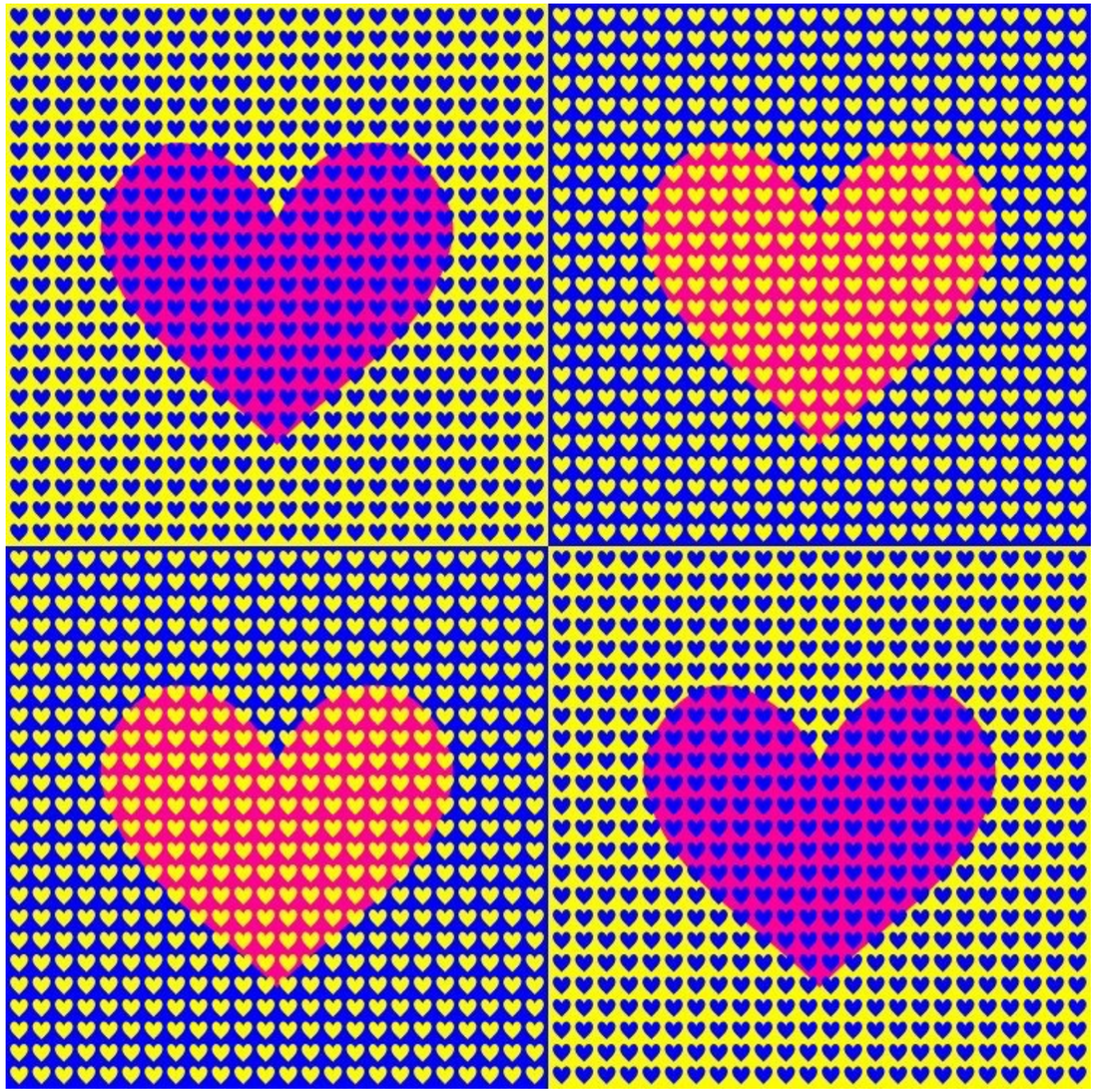


Simultaneous Contrast



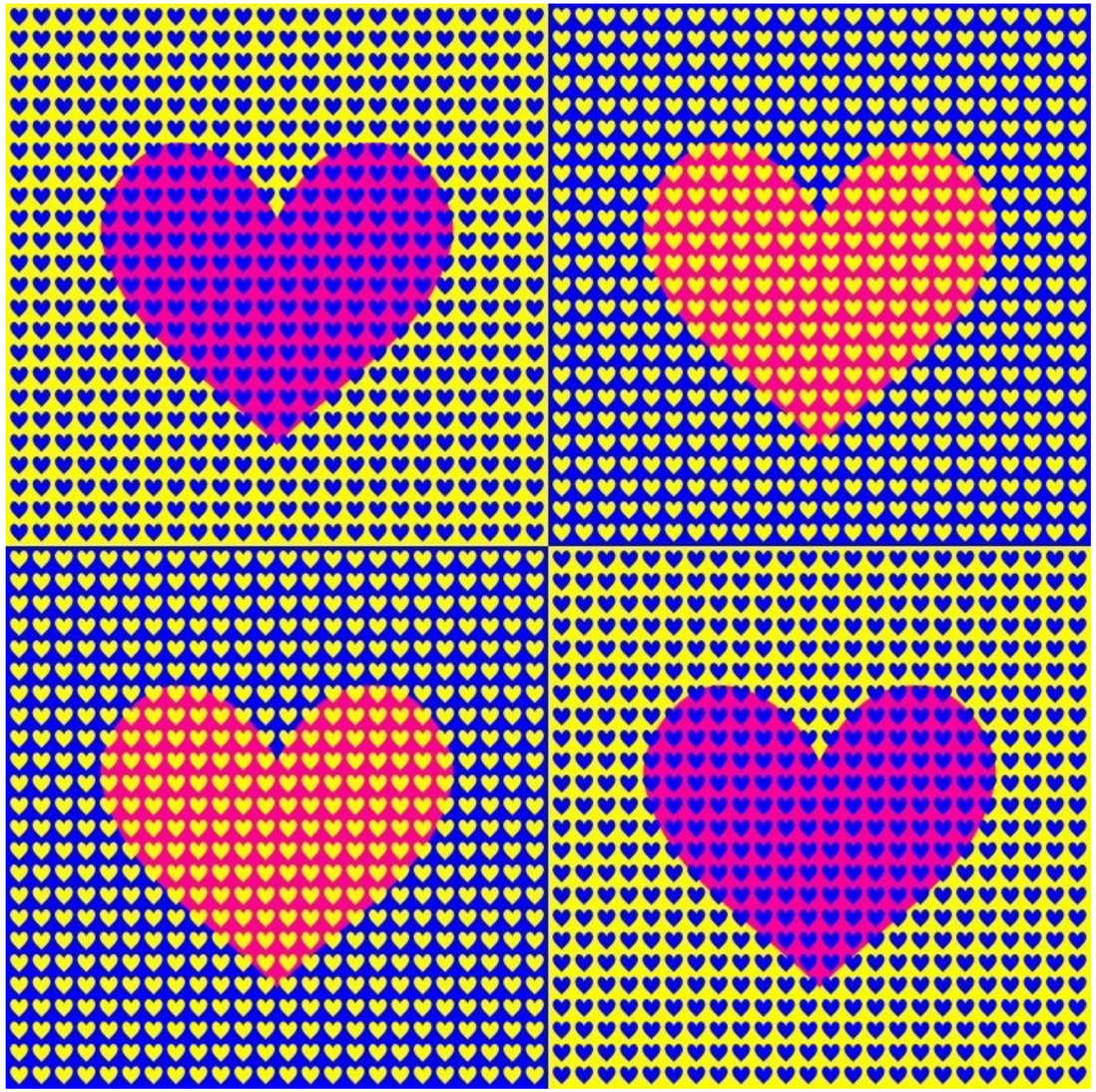
Simultaneous Contrast

What colors?



[A. Kitaoka]

What colors?



Red, yellow, blue

Purple, orange
do not exist!

[A. Kitaoka]

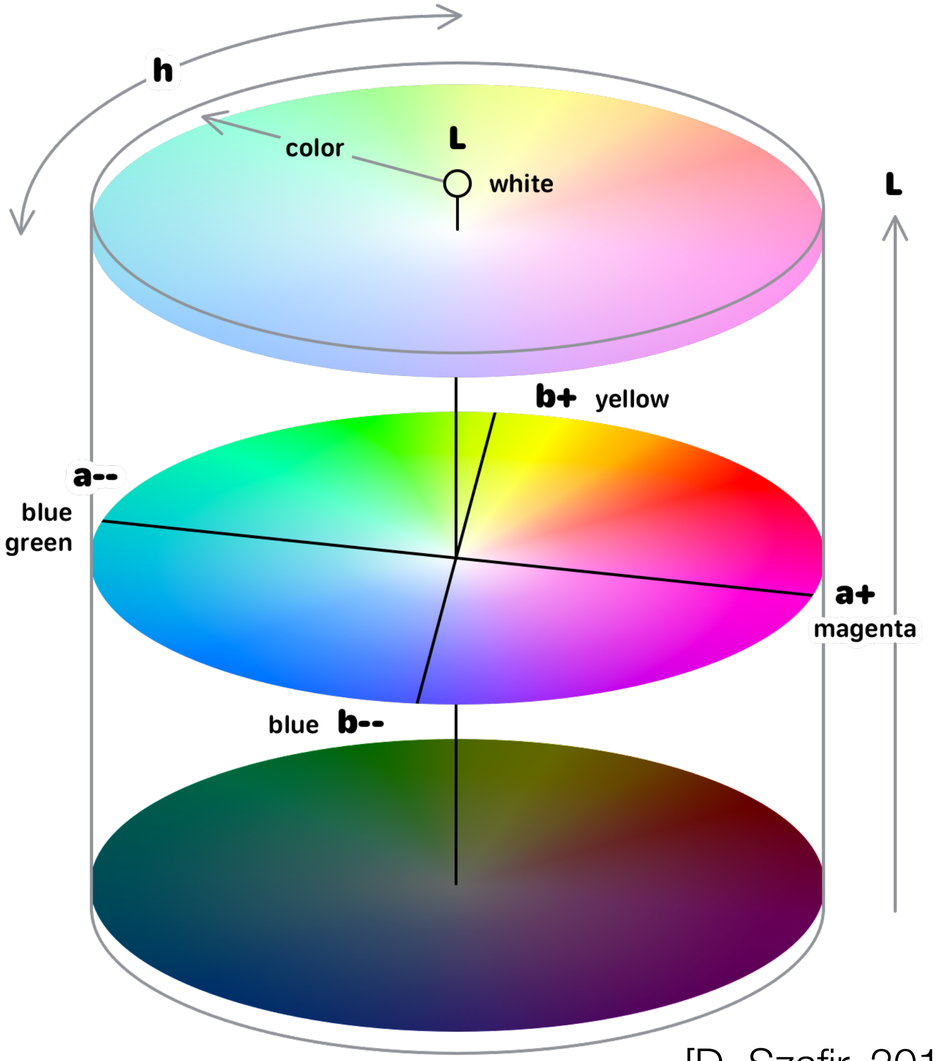
What does this mean for visualization?

What does this mean for visualization?

- We need to be aware of colorblindness when encoding via color
- Our brains may misinterpret color (surrounding colors matter!) even if we aren't colorblind
- Be careful! Don't assume that adding color always works the way you intended
- Use known colormaps when possible

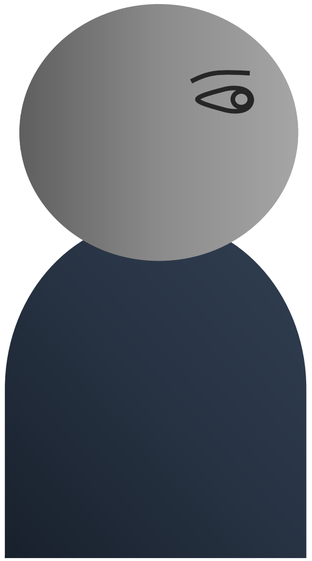
Violations of CIELAB Assumptions

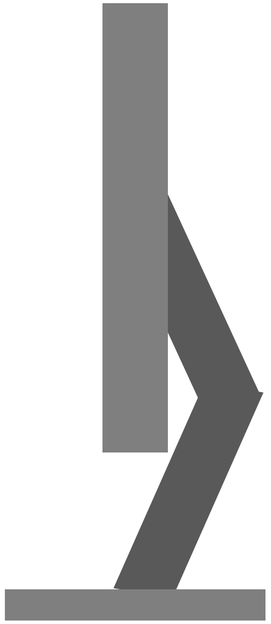
- CIELAB:
  - Approximately perceptually linear
  - 1 unit of Euclidean distance = 1 Just Noticeable Difference (JND)
  - JND: people detect change at least 50% of the time
- Assumptions CIELAB makes:
  - Simple world
  - Isolation
  - Geometric



[D. Szafir, 2017]

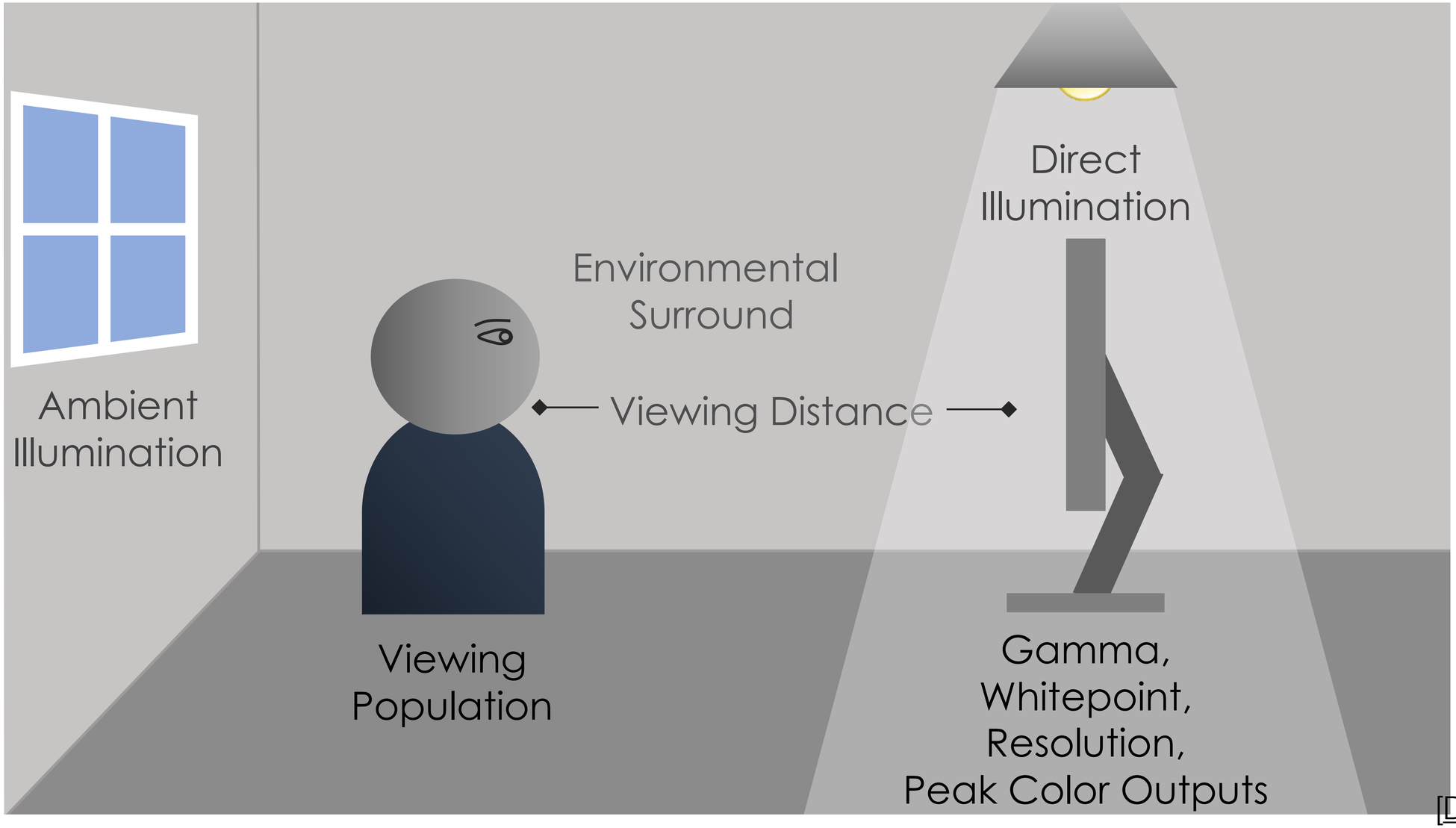
Simple World Assumption





[D. Szafir, 2017]

Problems with Simple World Assumption

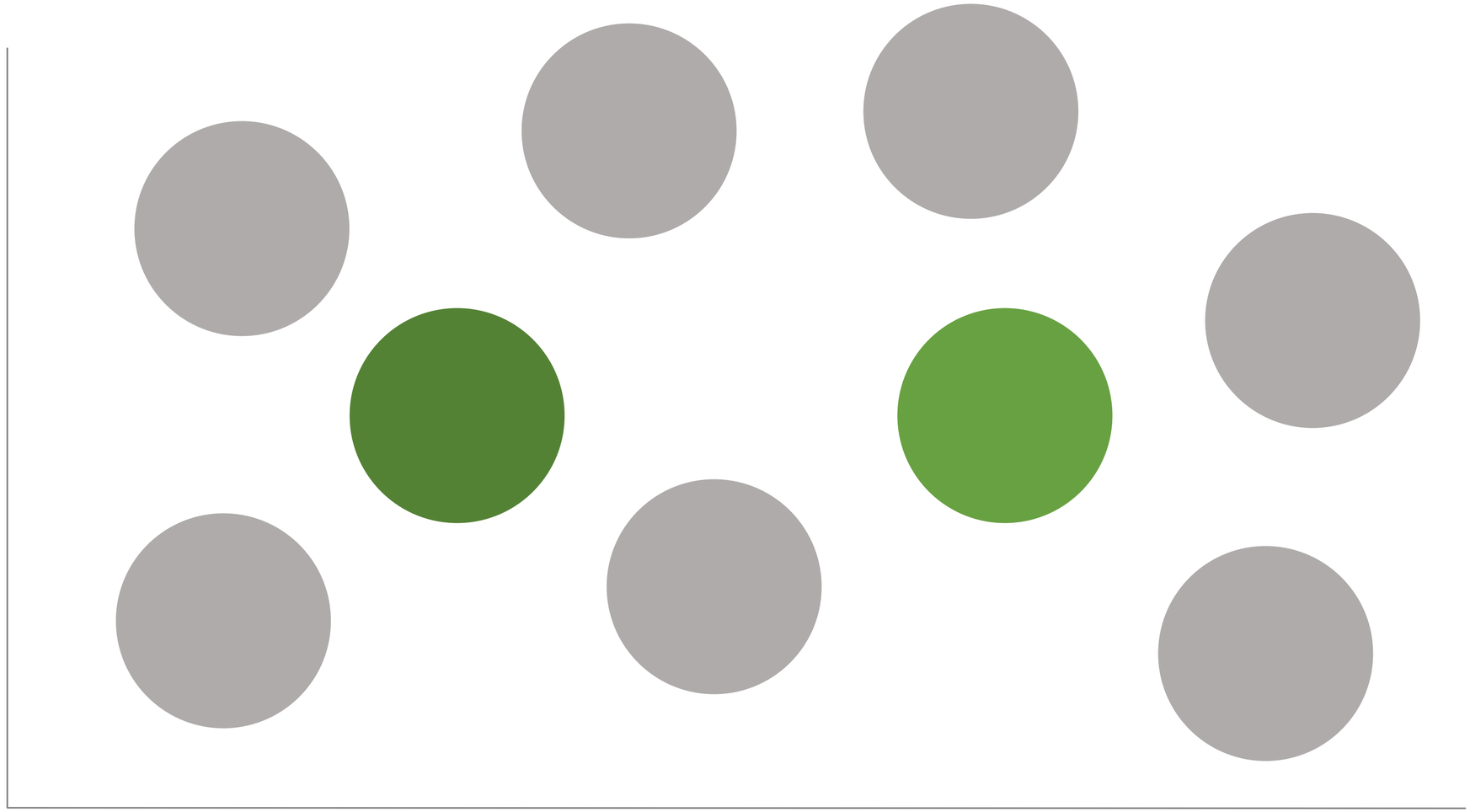


[D. Szafir, 2017]

Isolation Assumption

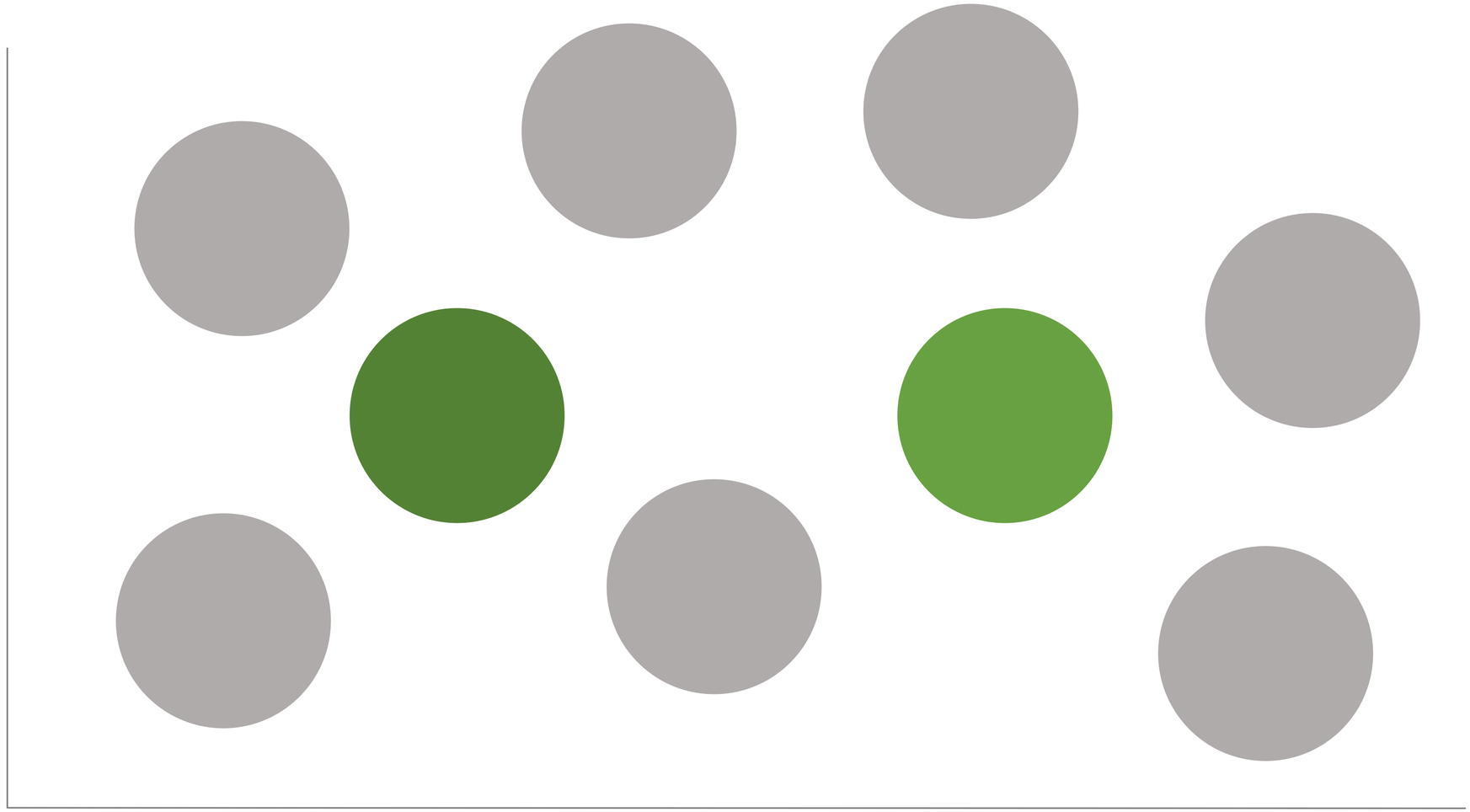
[D. Szafir, 2017]

Problems with Isolation Assumption



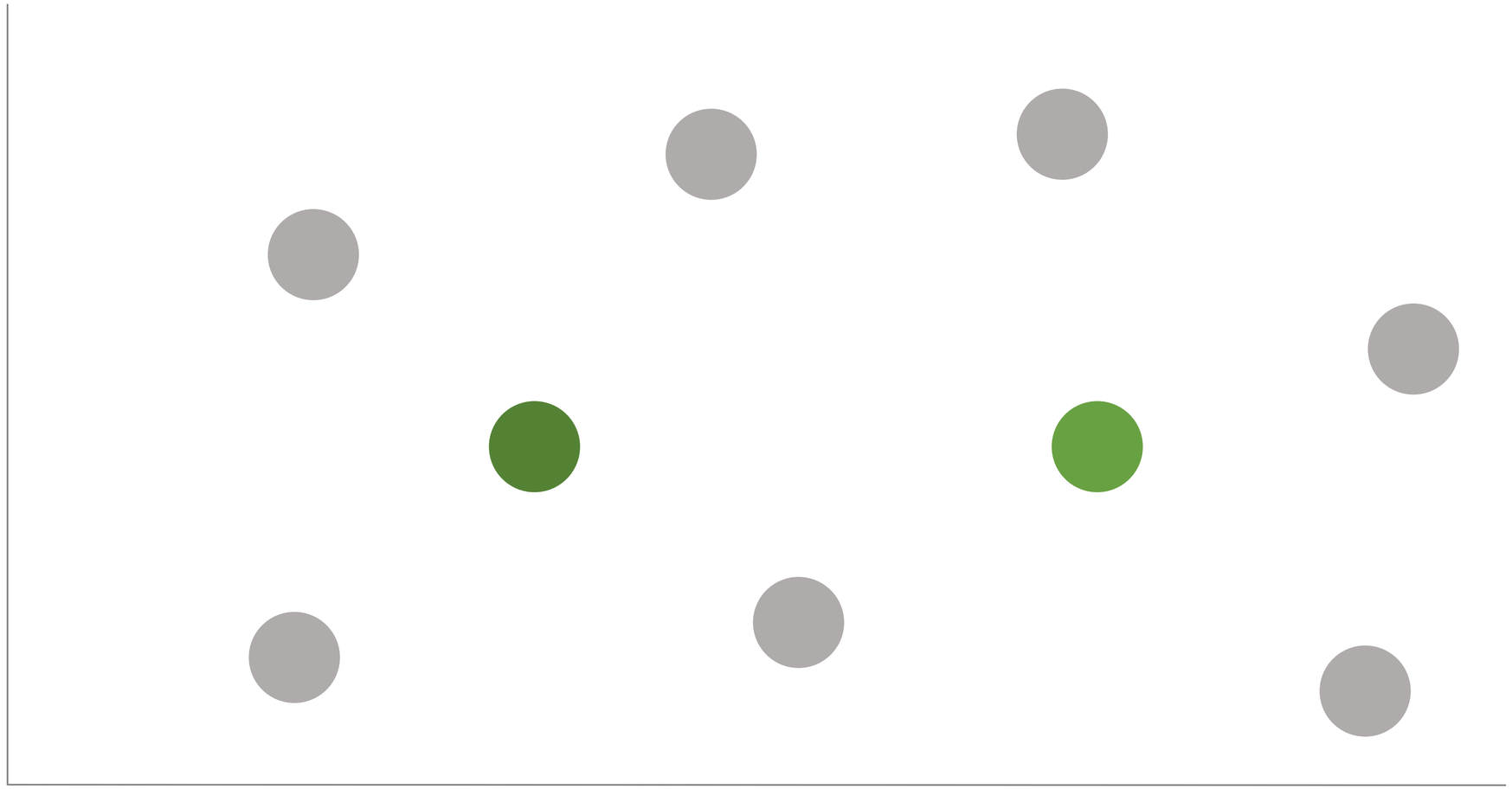
[D. Szafir, 2017]

Geometric Assumption



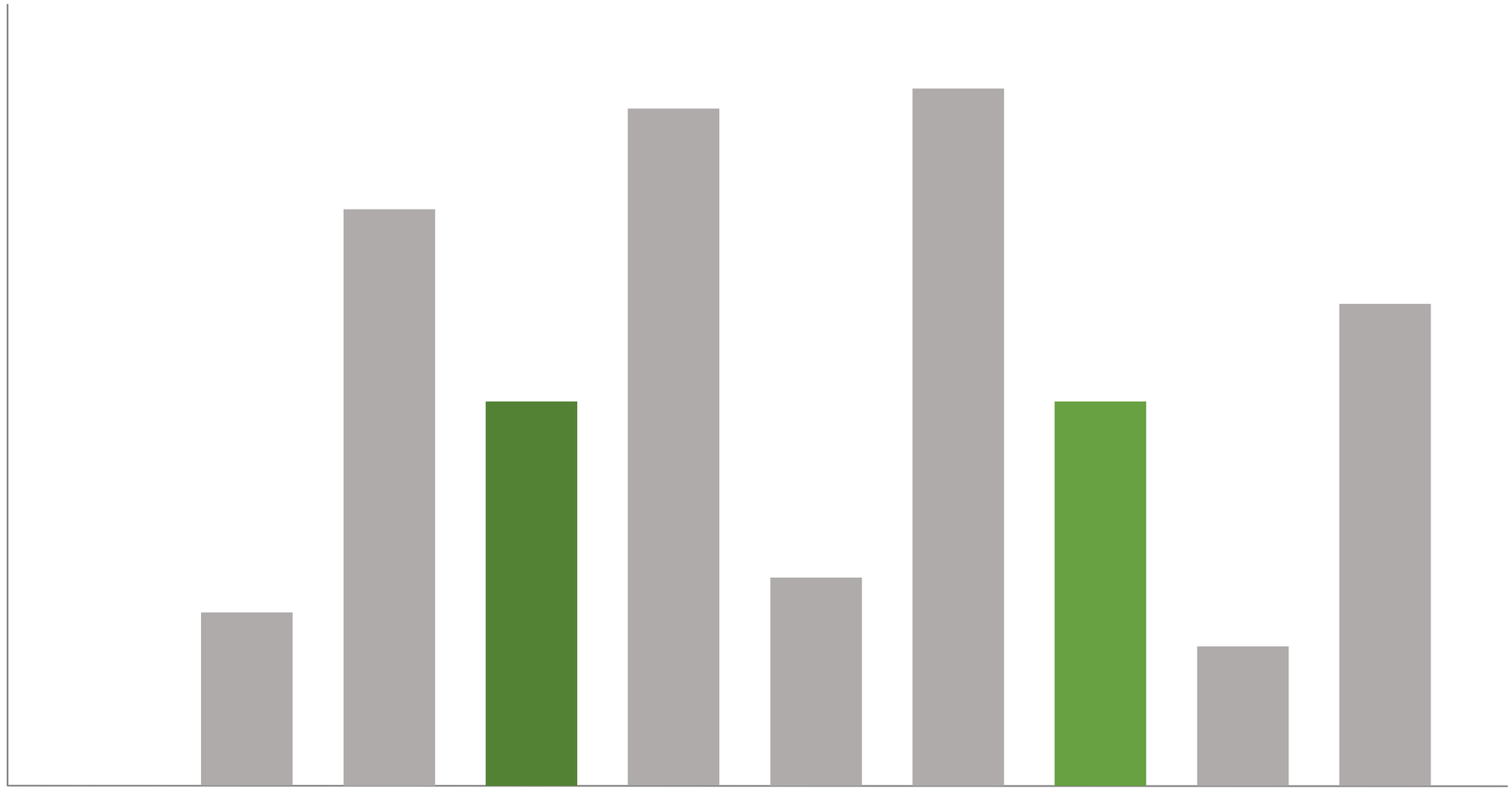
[D. Szafir, 2017]

Size Problem with Geometric Assumption



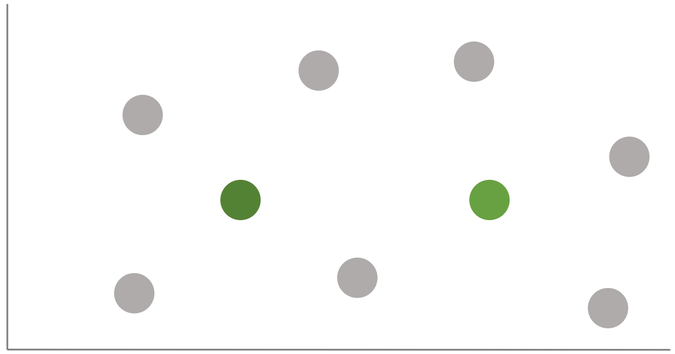
[D. Szafir, 2017]

Shape Problem with Geometric Assumption

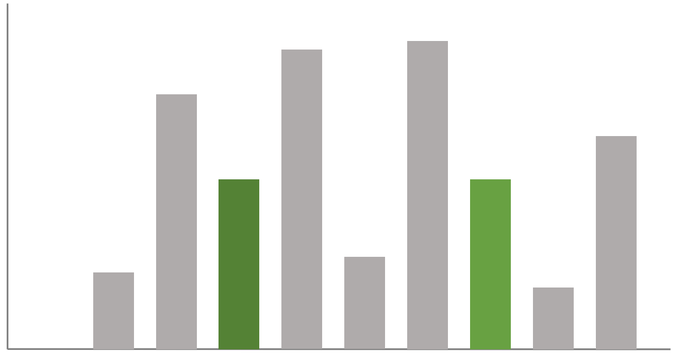


[D. Szafir, 2017]

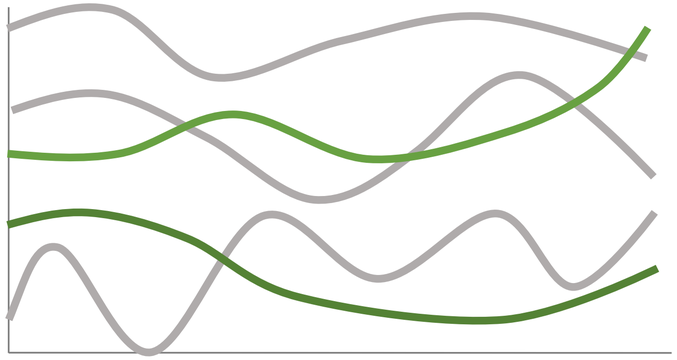
Types of Geometry



Diagonally Symmetric Marks



Elongated Marks

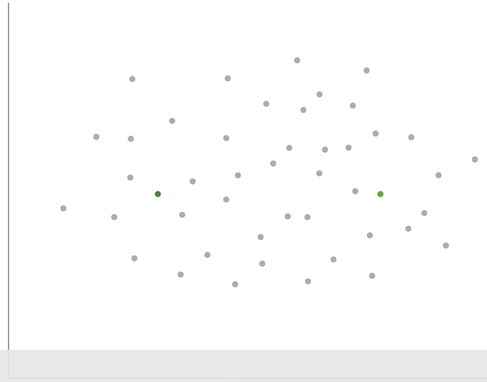


Asymmetric Marks

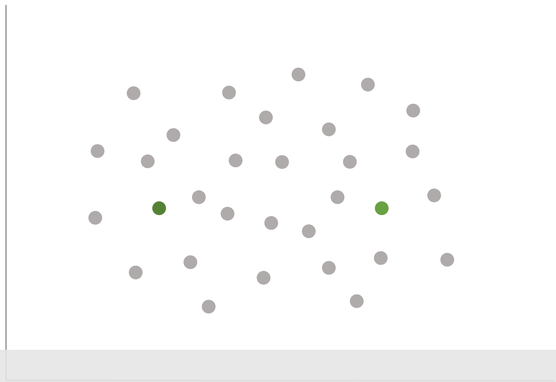
Area Marks

[D. Szafir, 2017]

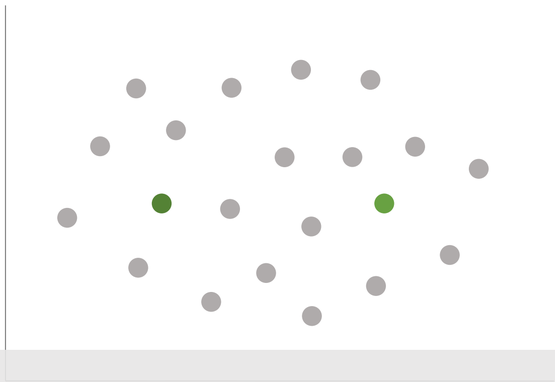
Run the tests!



6 pixels
0.25°



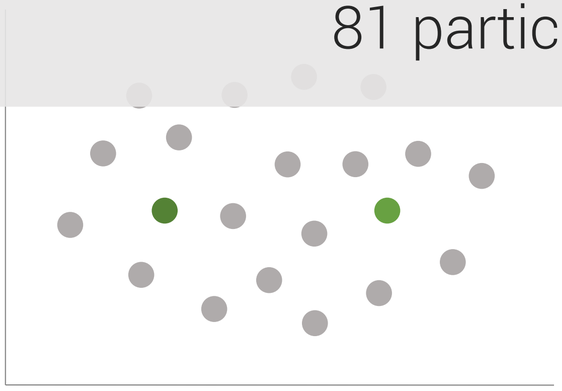
12 pixels
0.5°



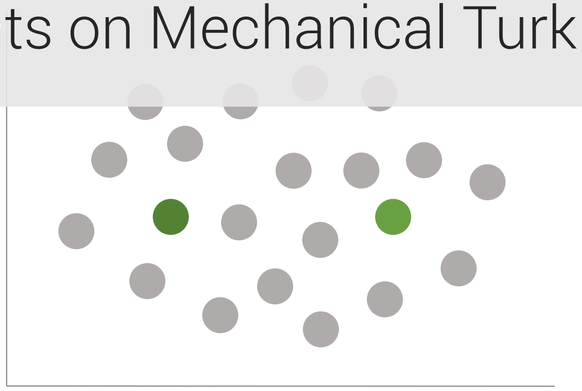
18 pixels
0.75°

6 (diameters, within) × 6 (color differences, within) × 3 (color axis, between)

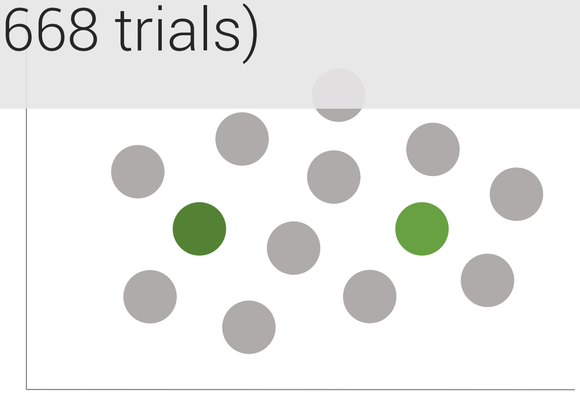
81 participants on Mechanical Turk (5,668 trials)



25 pixels
1.0°



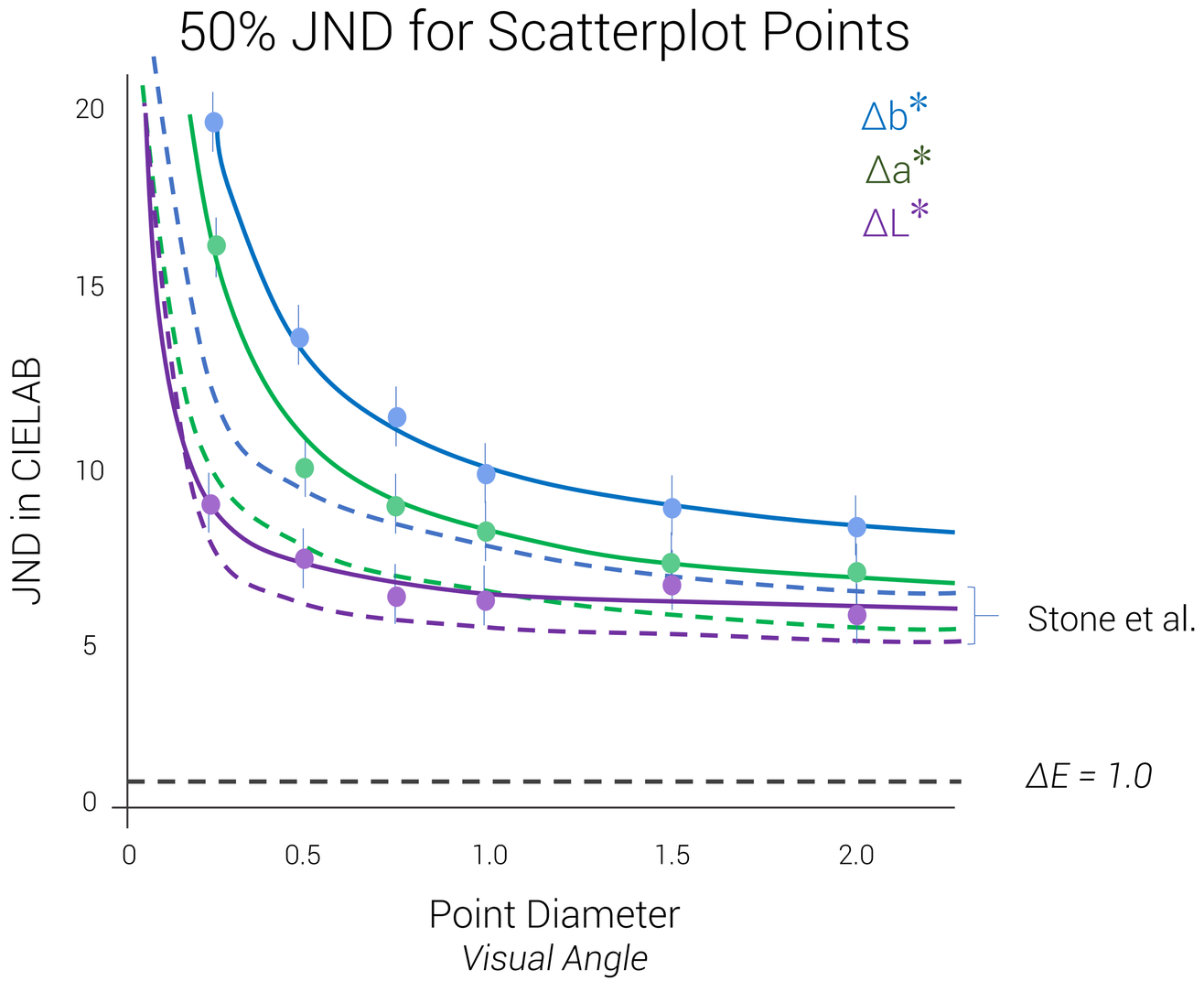
37 pixels
1.5°



50 pixels
2.0°

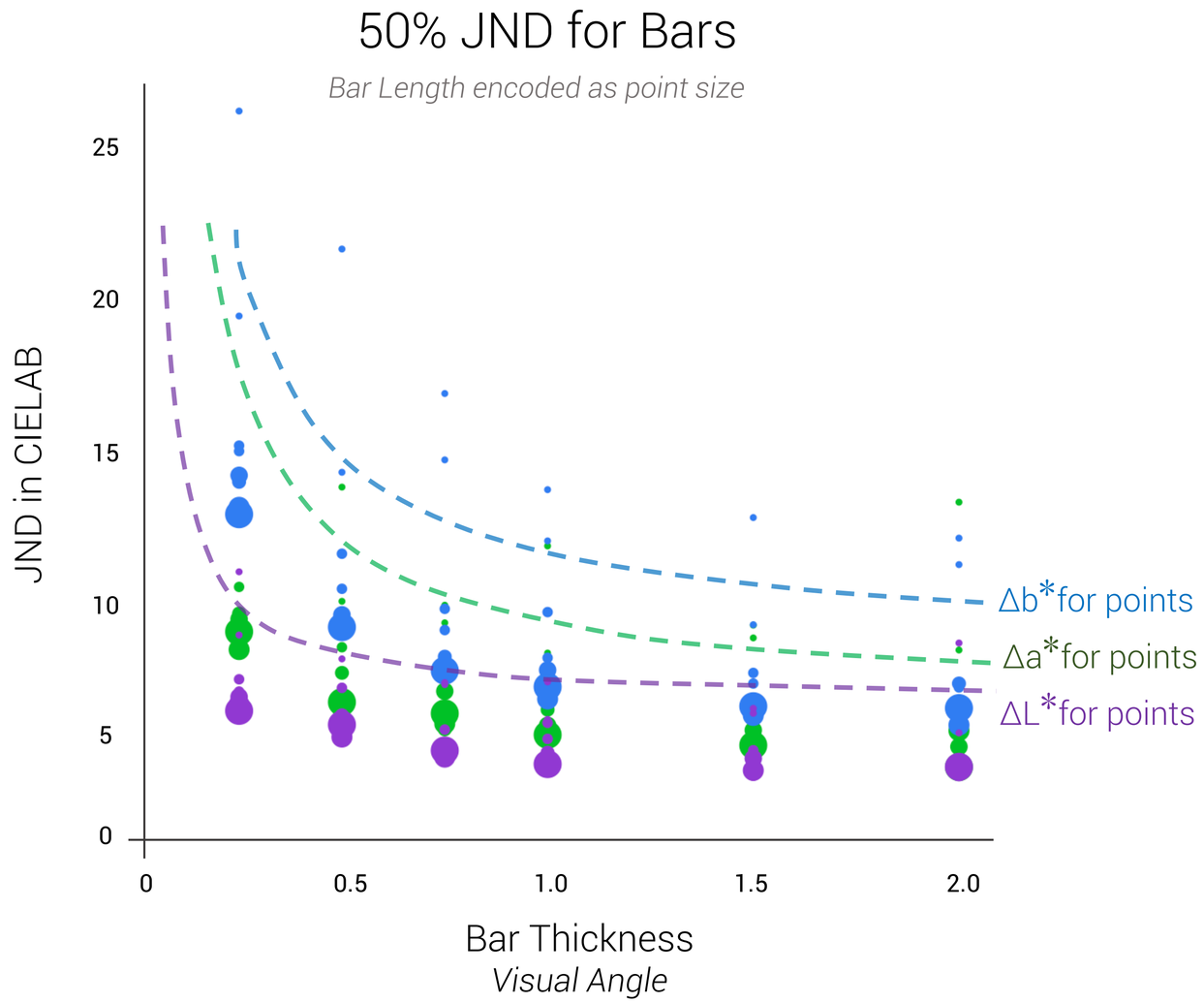
[D. Szafir, 2017]

Point Size: consistent with previous results



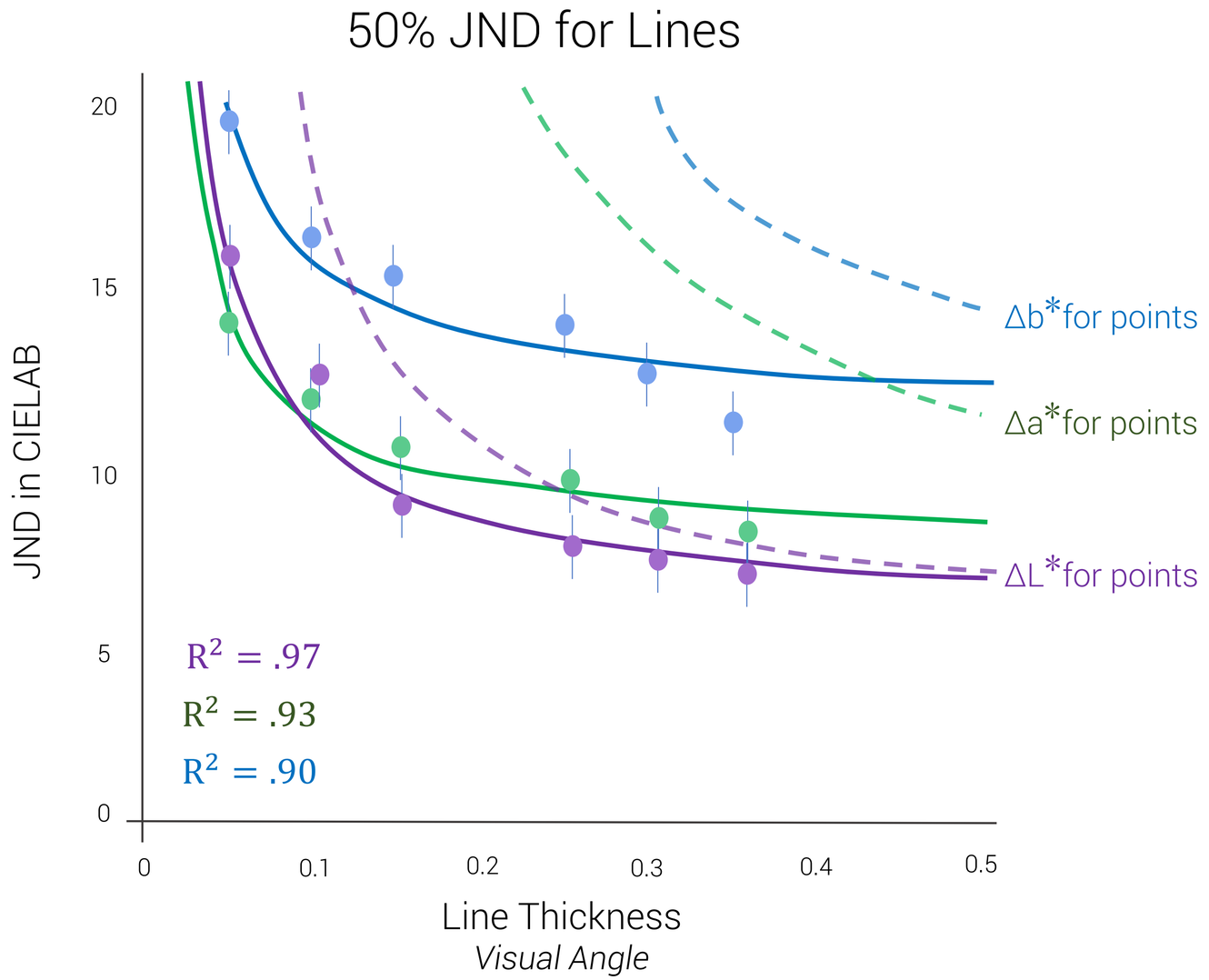
[D. Szafir, 2017]

Bar Thickness and Length: longer bars help



[D. Szafir, 2017]

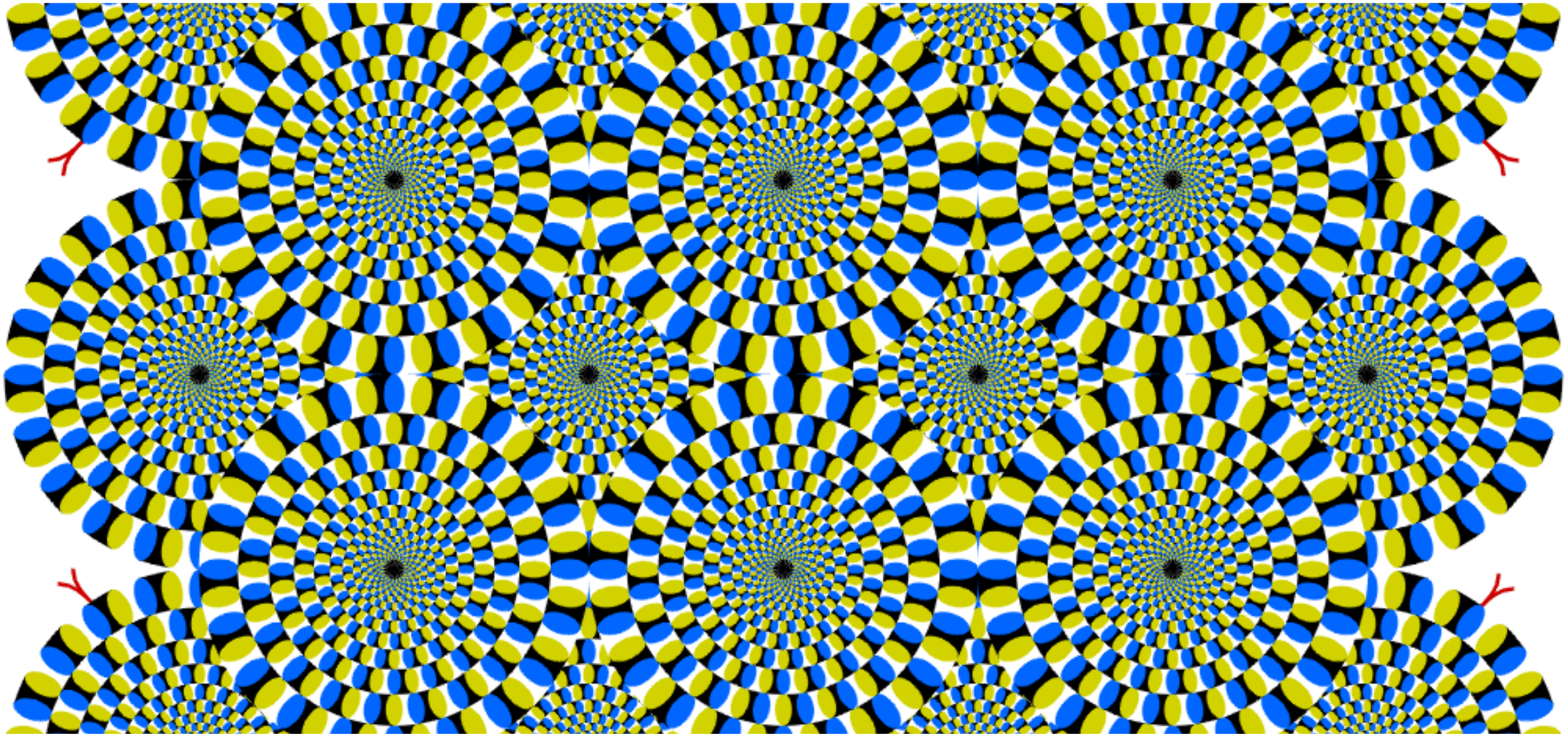
Line Thickness: better than points



[D. Szafir, 2017]

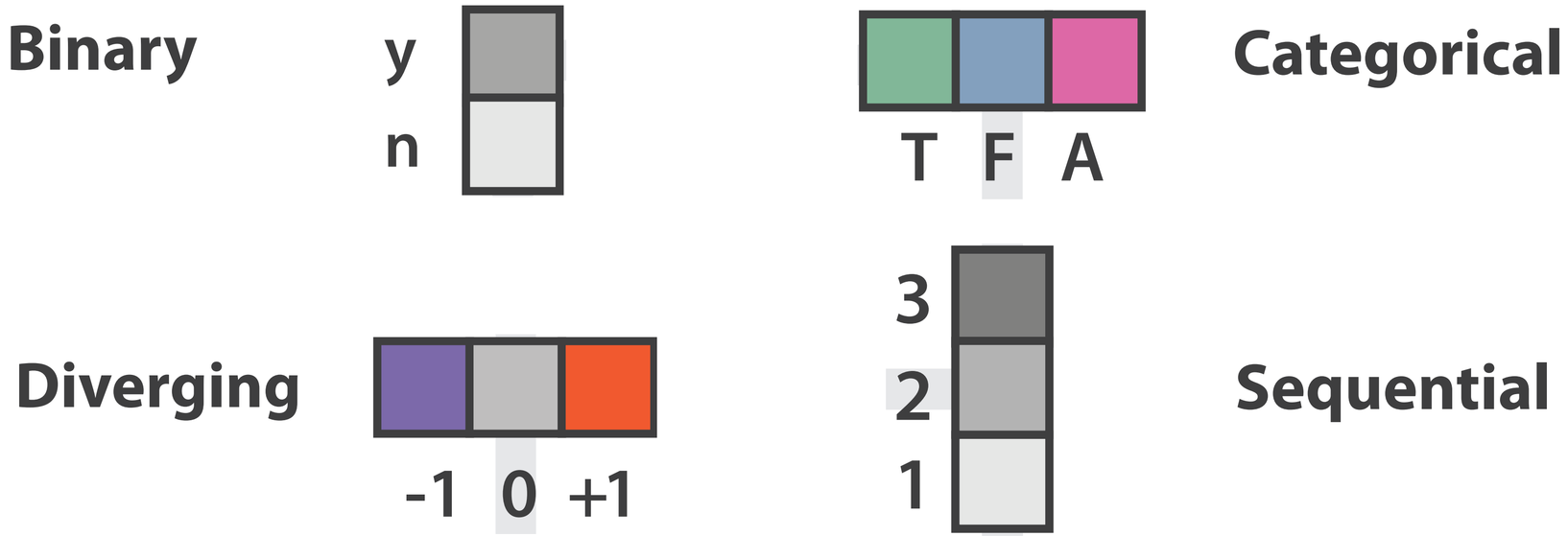
Color perception in real-world visualizations is complicated

Akiyoshi Kitaoka's Illusion pages



Colormap

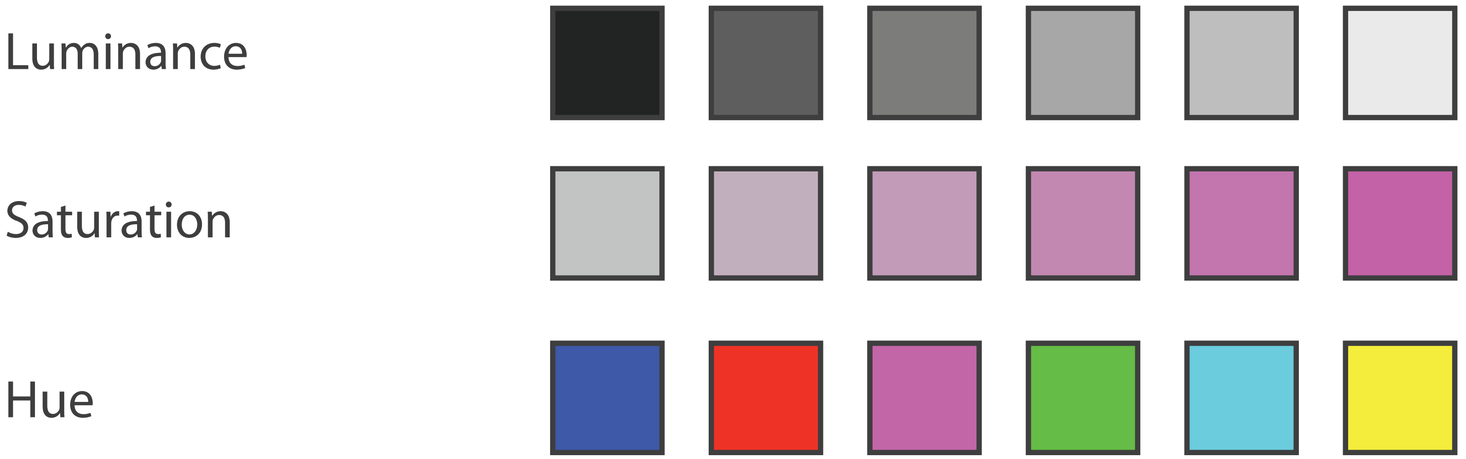
- A colormap specifies a mapping between colors and data values
- Colormap should follow the expressiveness principle
- Types of colormaps:



[Munzner (ill. Maguire), 2014]

Categorical vs. Ordered

- Hue has no implicit ordering: use for categorical data
- Saturation and luminance do: use for ordered data

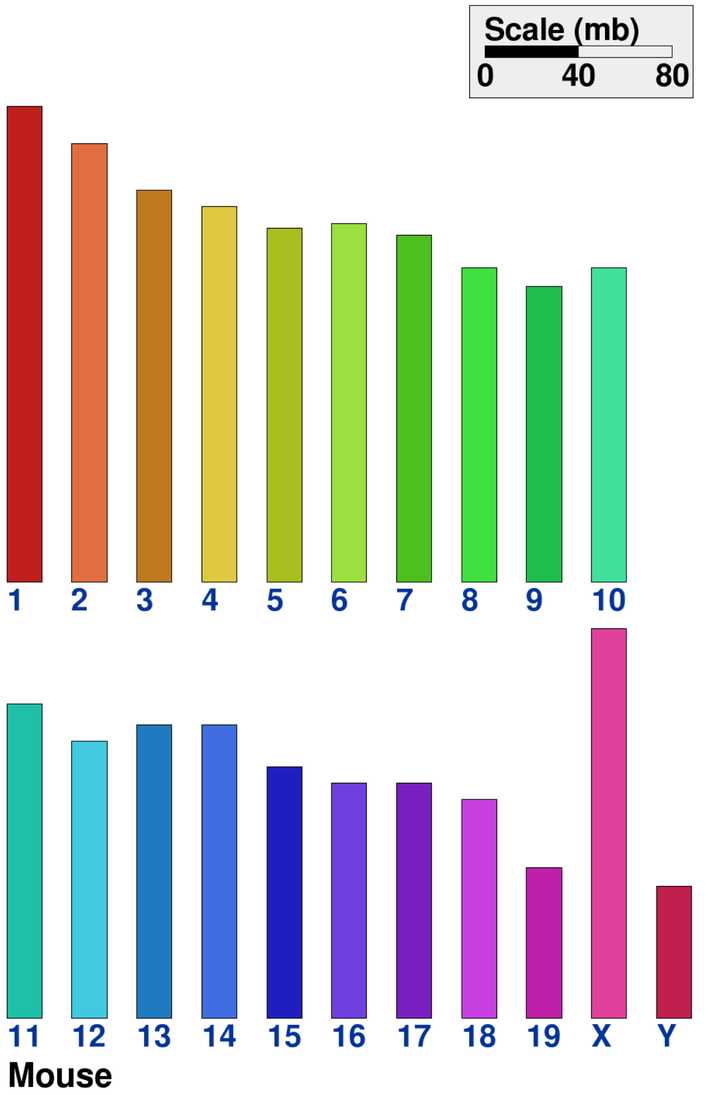


[Munzner (ill. Maguire), 2014]

Categorical Colormap Guidelines

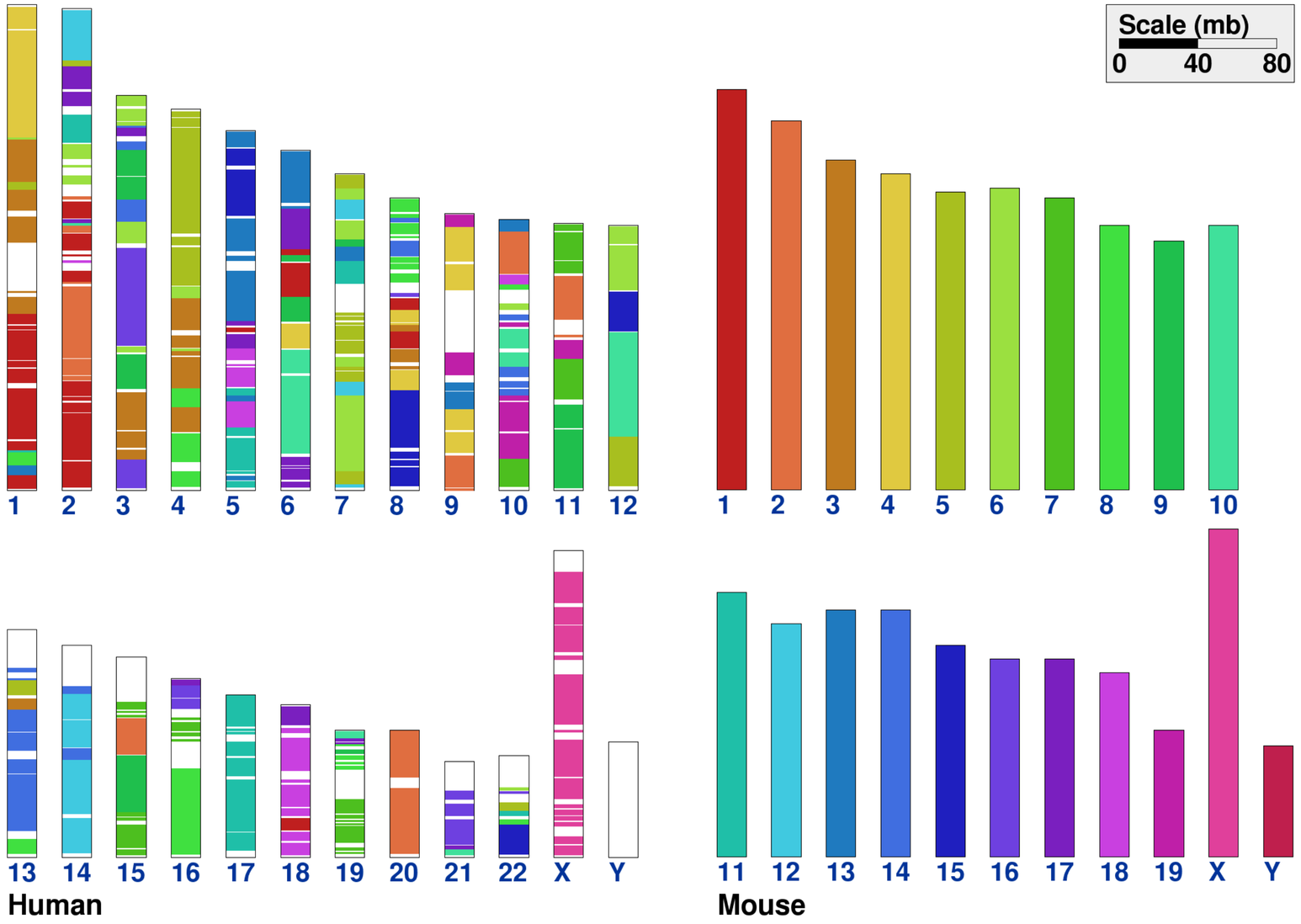
- Don't use too many colors (~12)
- Remember your background has a color, too
- Nameable colors help
- Be aware of luminance (e.g. difference between blue and yellow)
- Think about other marks you might wish to use in the visualization

Number of distinguishable colors?



[Sinha & Meller, 2007]

Number of distinguishable colors?



[Sinha & Meller, 2007]