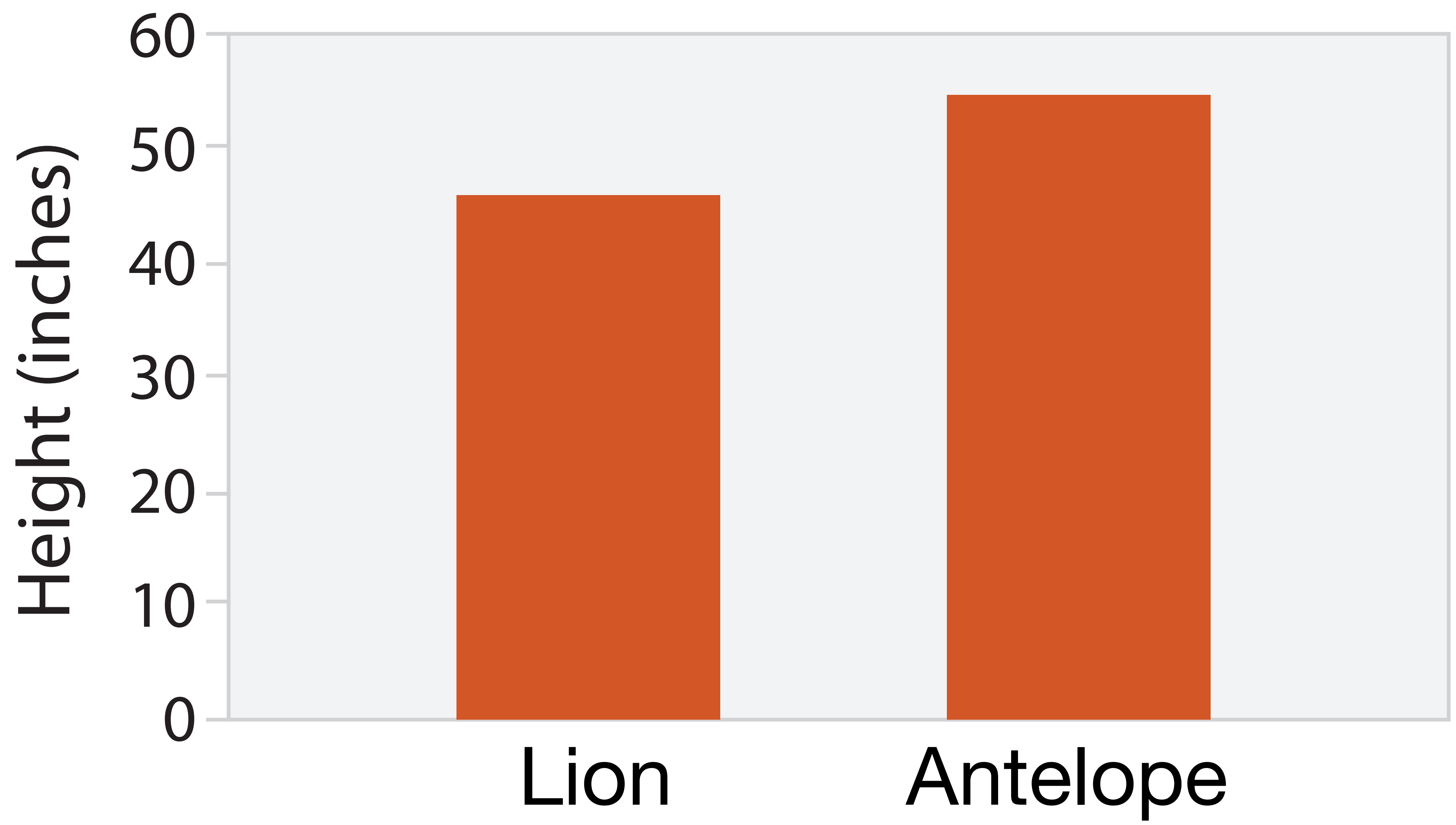
Data Visualization (CIS 490/680)

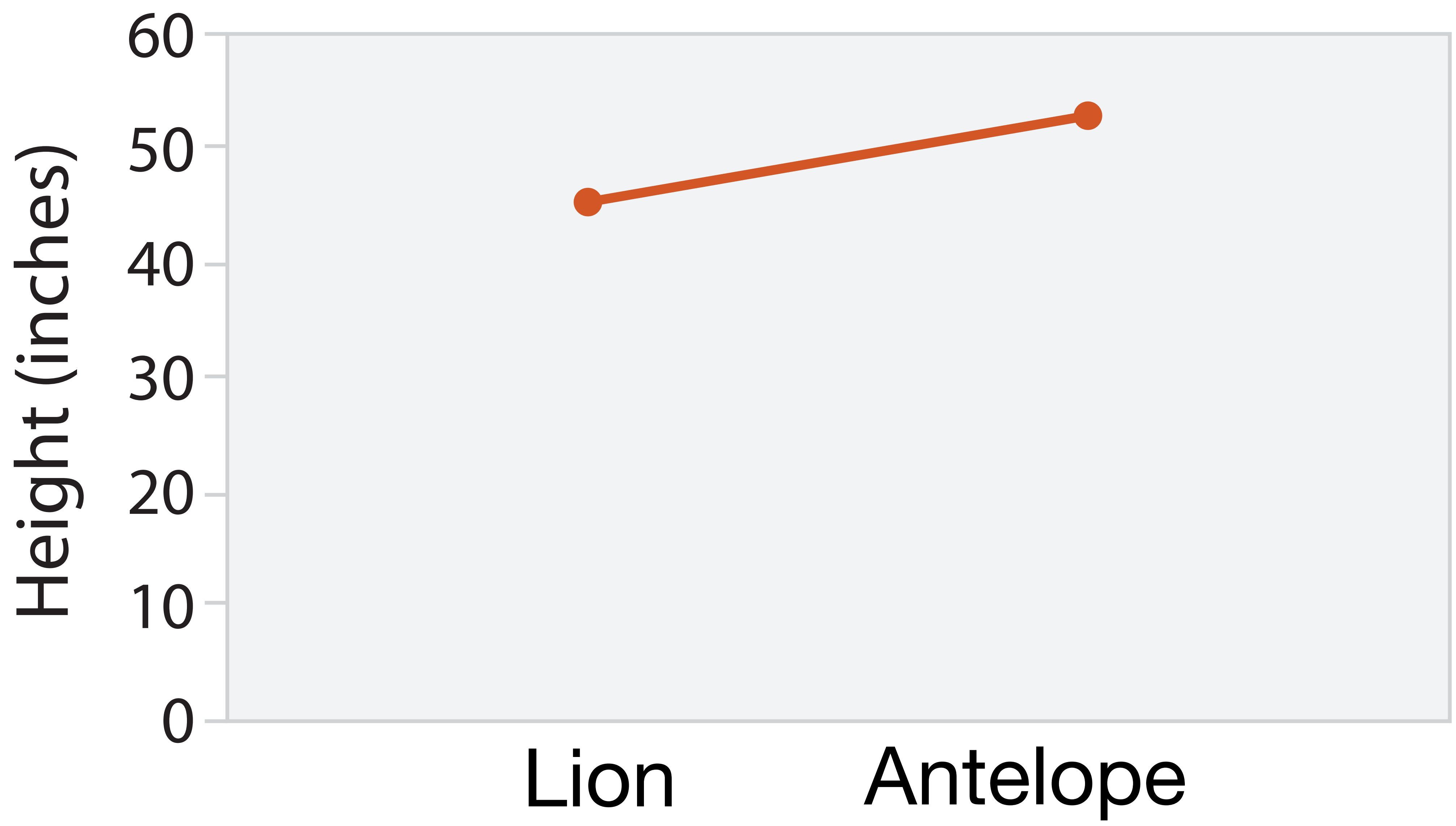
Color

Dr. David Koop

Northern Illinois University

Proper Use of Line and Bar Charts



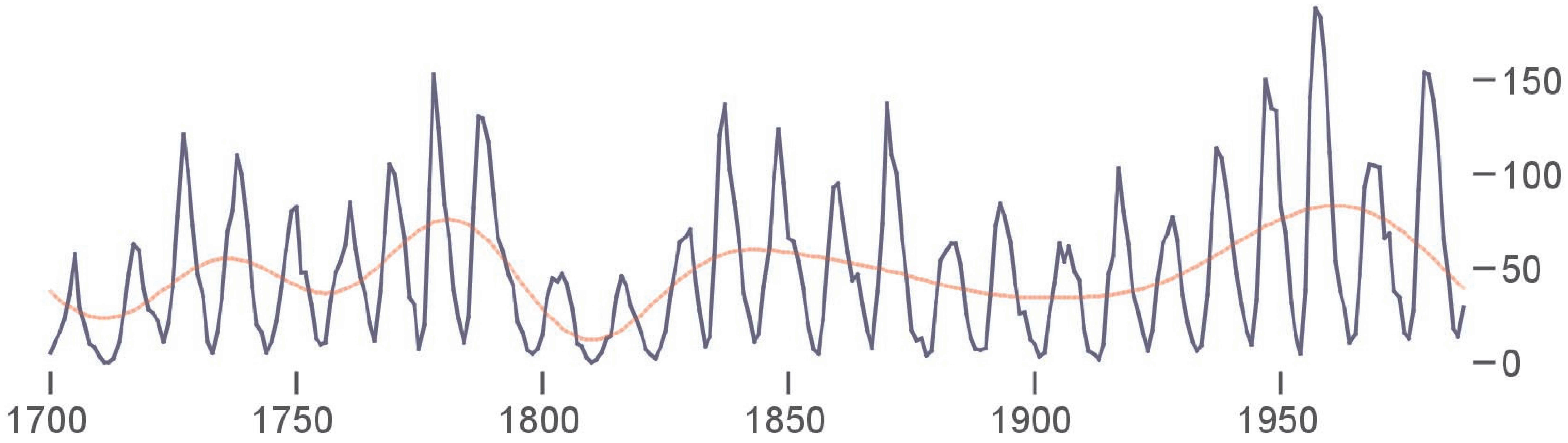


- What does the line indicate?
- Does this make sense?

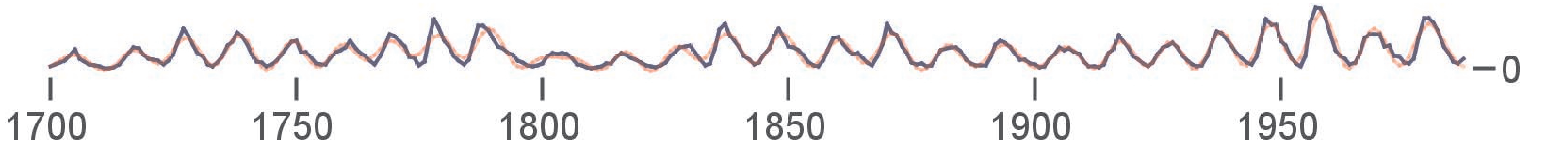
[Adapted from Zacks and Tversky, 1999, Munzner (ill. Maguire), 2014]

Effect of Aspect Ratio in Line Charts

Aspect Ratio = 3.96

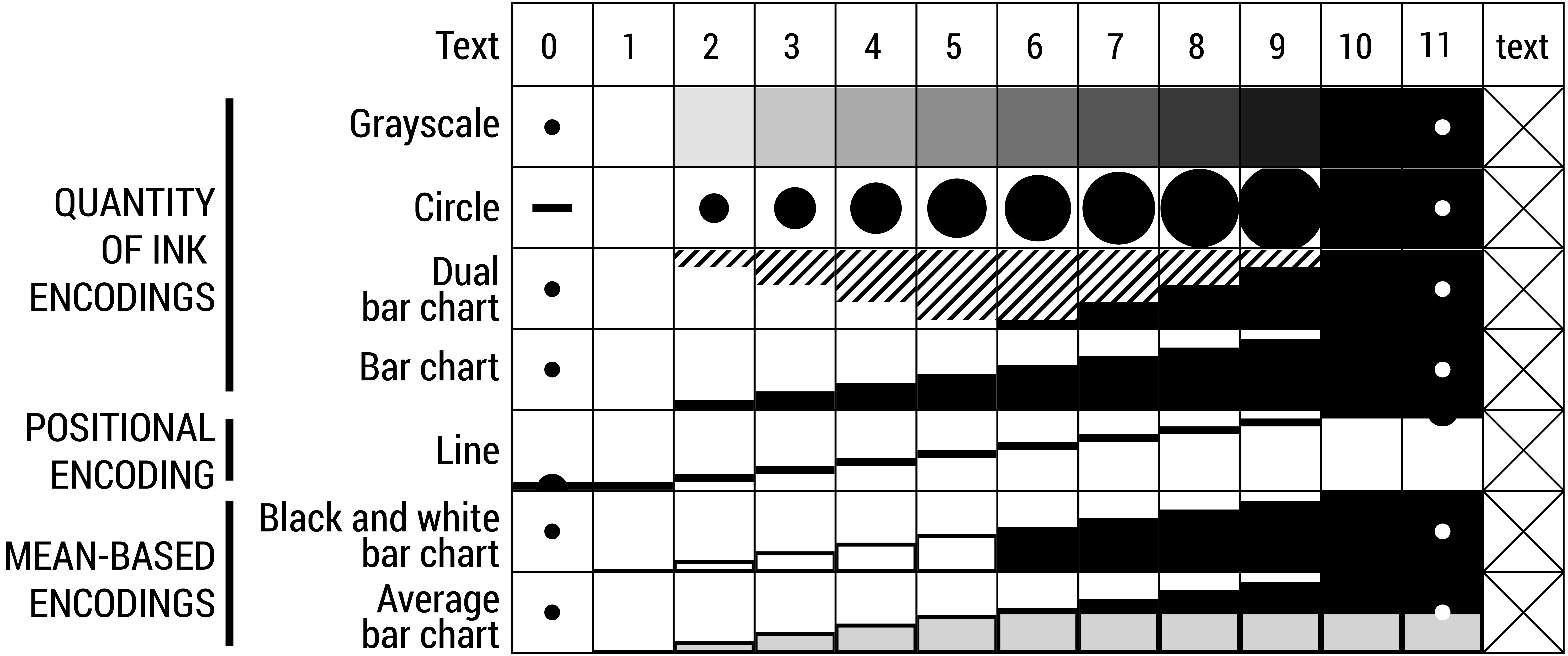


Aspect Ratio = 22.35



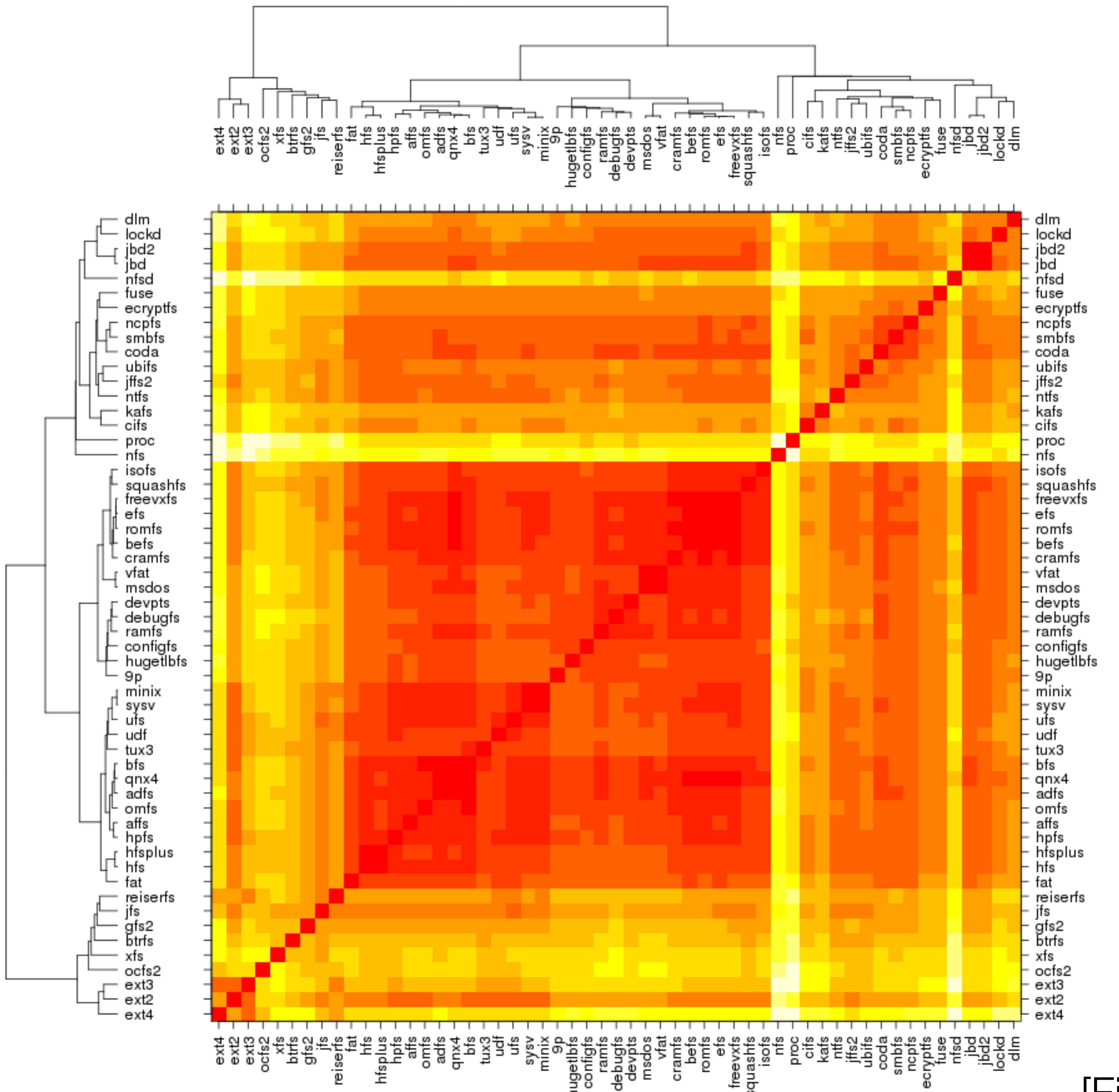
[Heer and Agrawala, 2006]

Bertin’s Matrix Encodings



[C.Perrin et al., 2014]

Cluster Heatmap (with color)

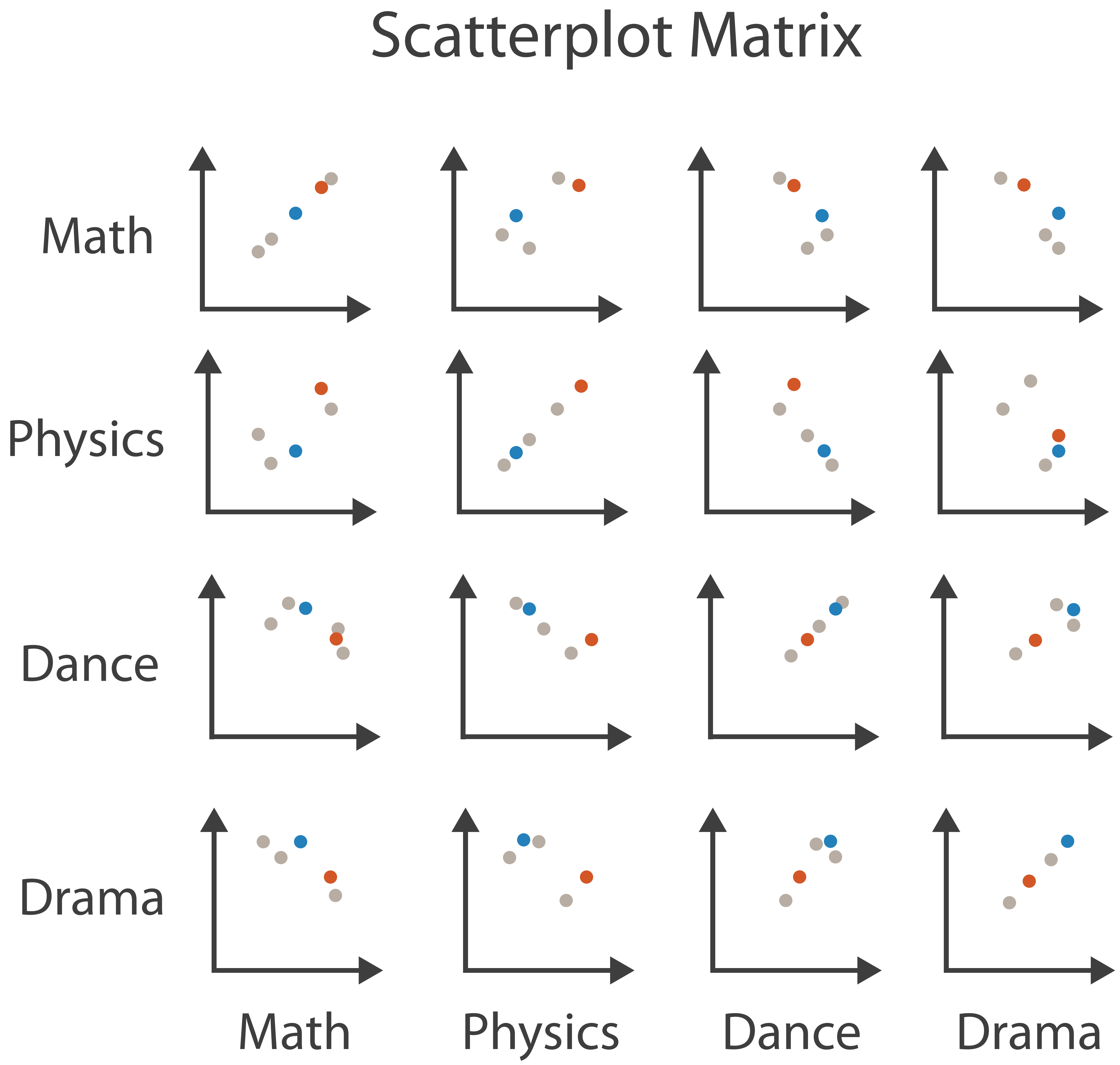


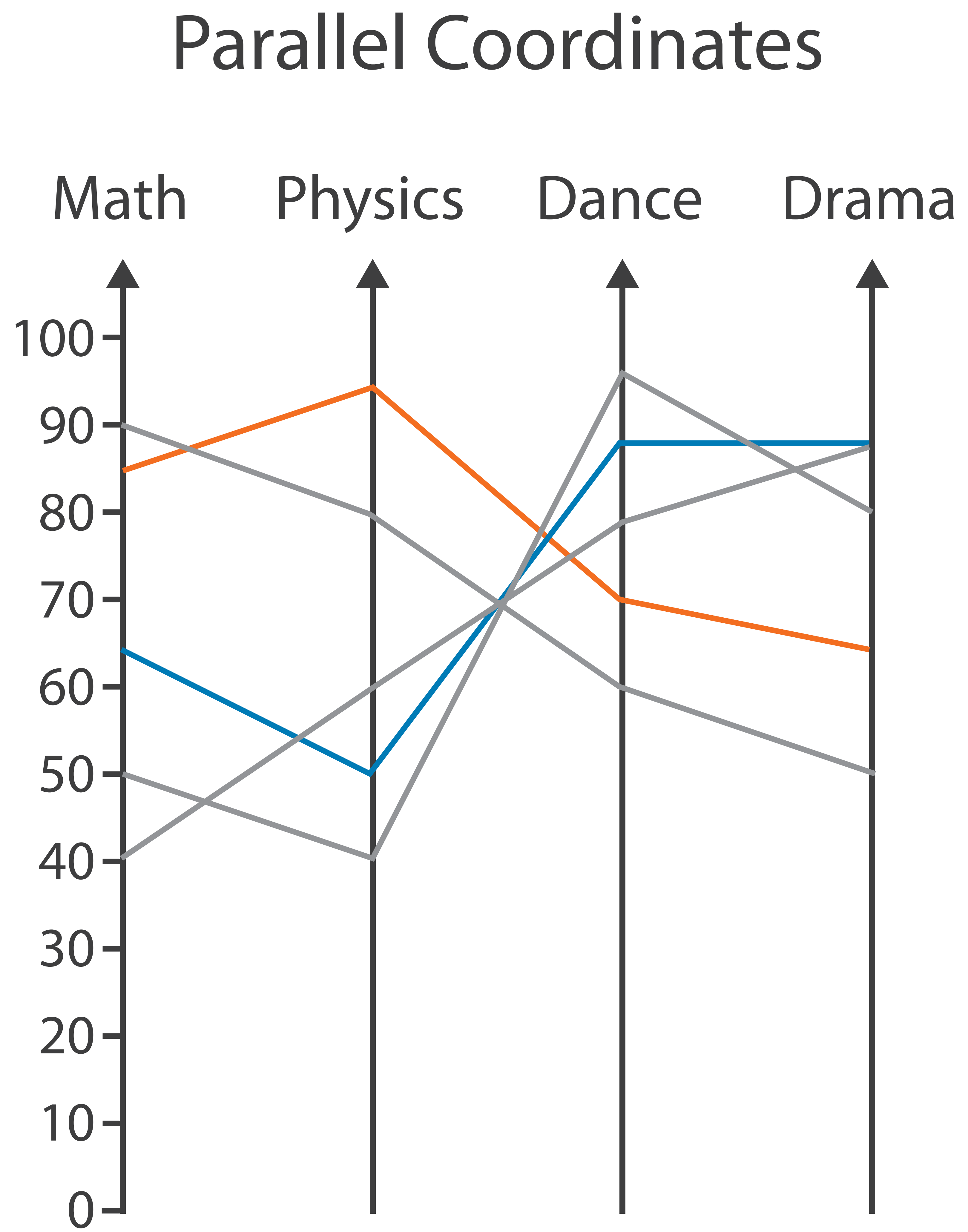
[File System Similarity, R. Musăloiu-E., 2009]

Scatterplot Matrices and Parallel Coordinates

[Munzner (ill. Maguire), 2014]

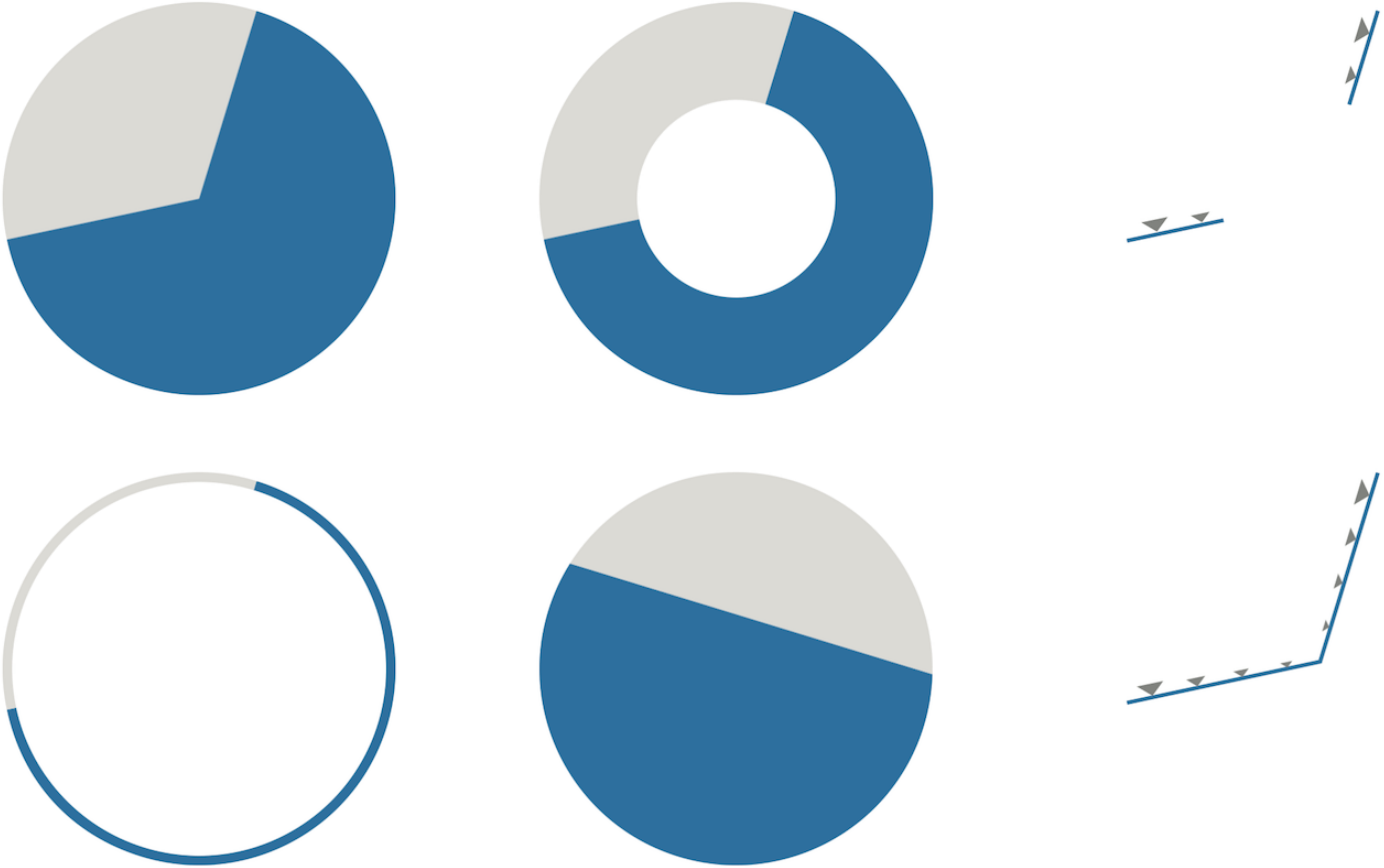
Scatterplot Matrices and Parallel Coordinates





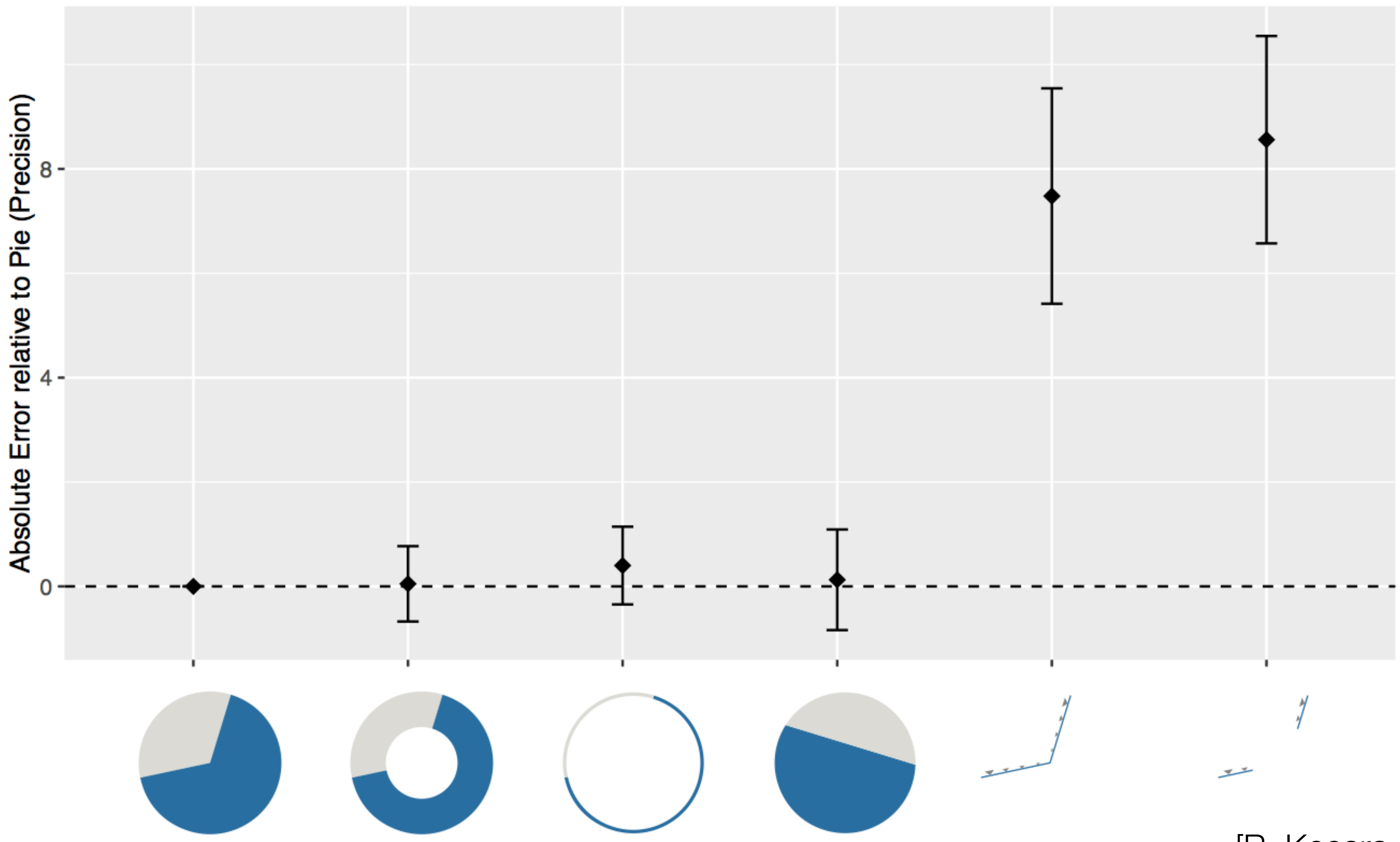
[Munzner (ill. Maguire), 2014]

Pie Charts: Arcs, Angles, or Areas?



[R. Kosara and D. Skau, 2016]

User Studies: Absolute Error Relative to Pie Chart



[R. Kosara and D. Skau, 2016]

Project Proposal

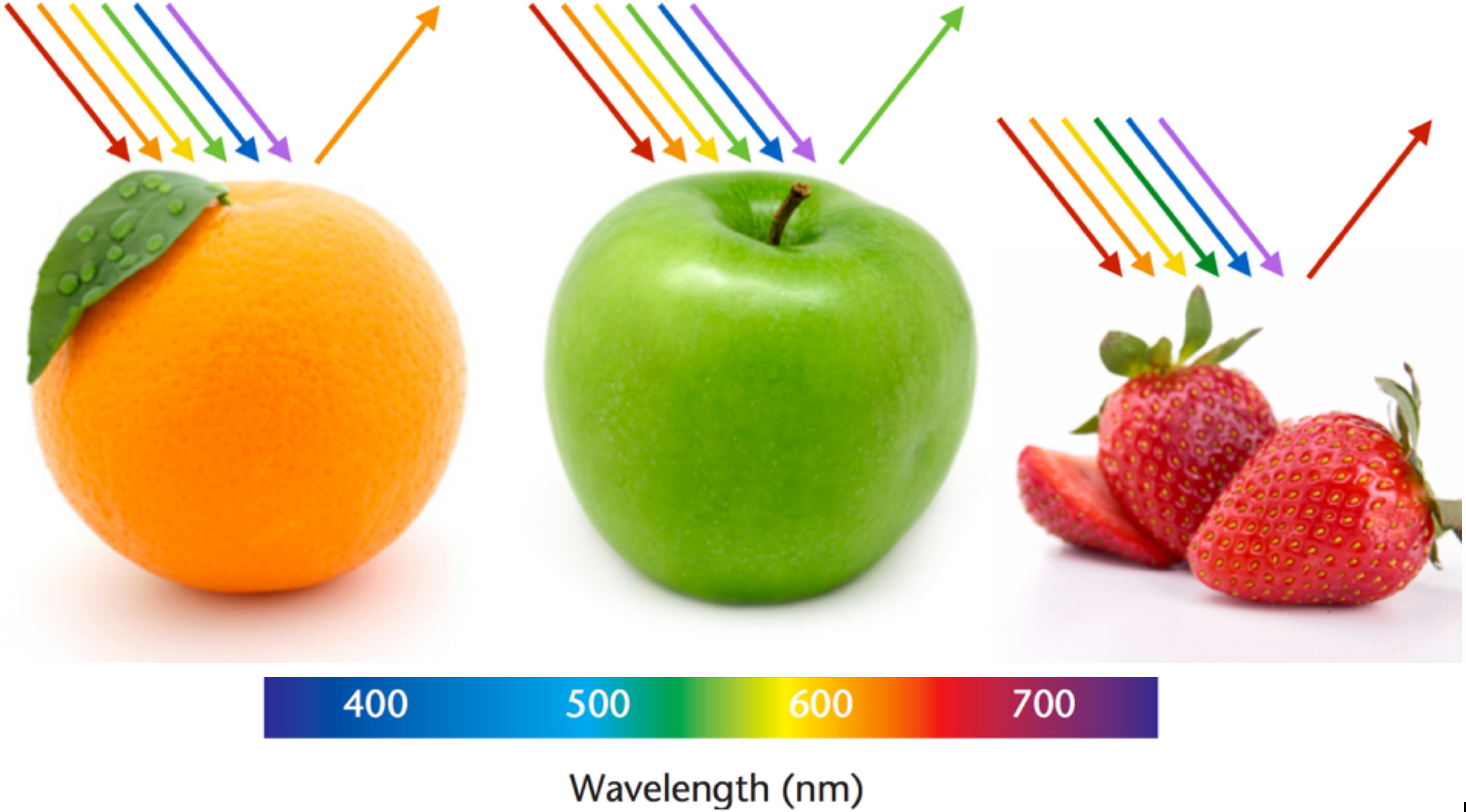
- Find an interesting subject or dataset
  - see List of lists of datasets [B. Keegan]
- Understand the data available (format, types, semantics)
- Figure out some interesting questions and tasks
- Start brainstorming about visualizations and interactions
- Inspiration:
  - Information Is Beautiful Awards
  - MBTA Viz
- Due Friday, October 11, 2019

Midterm

- Two weeks
- Thursday, October 17

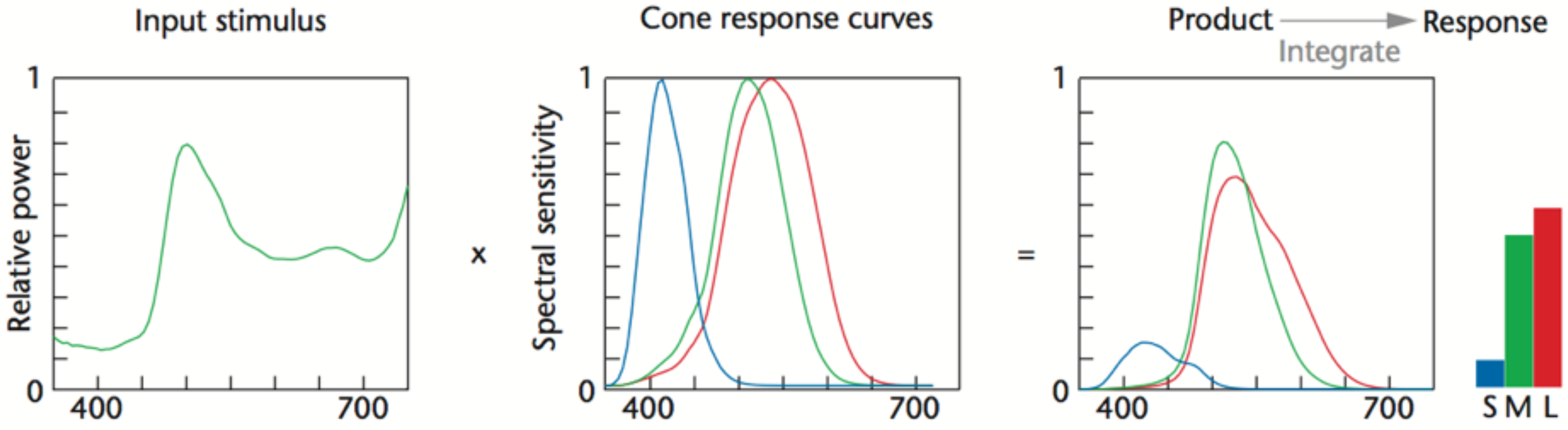
Color

Light Reflection & Absorption



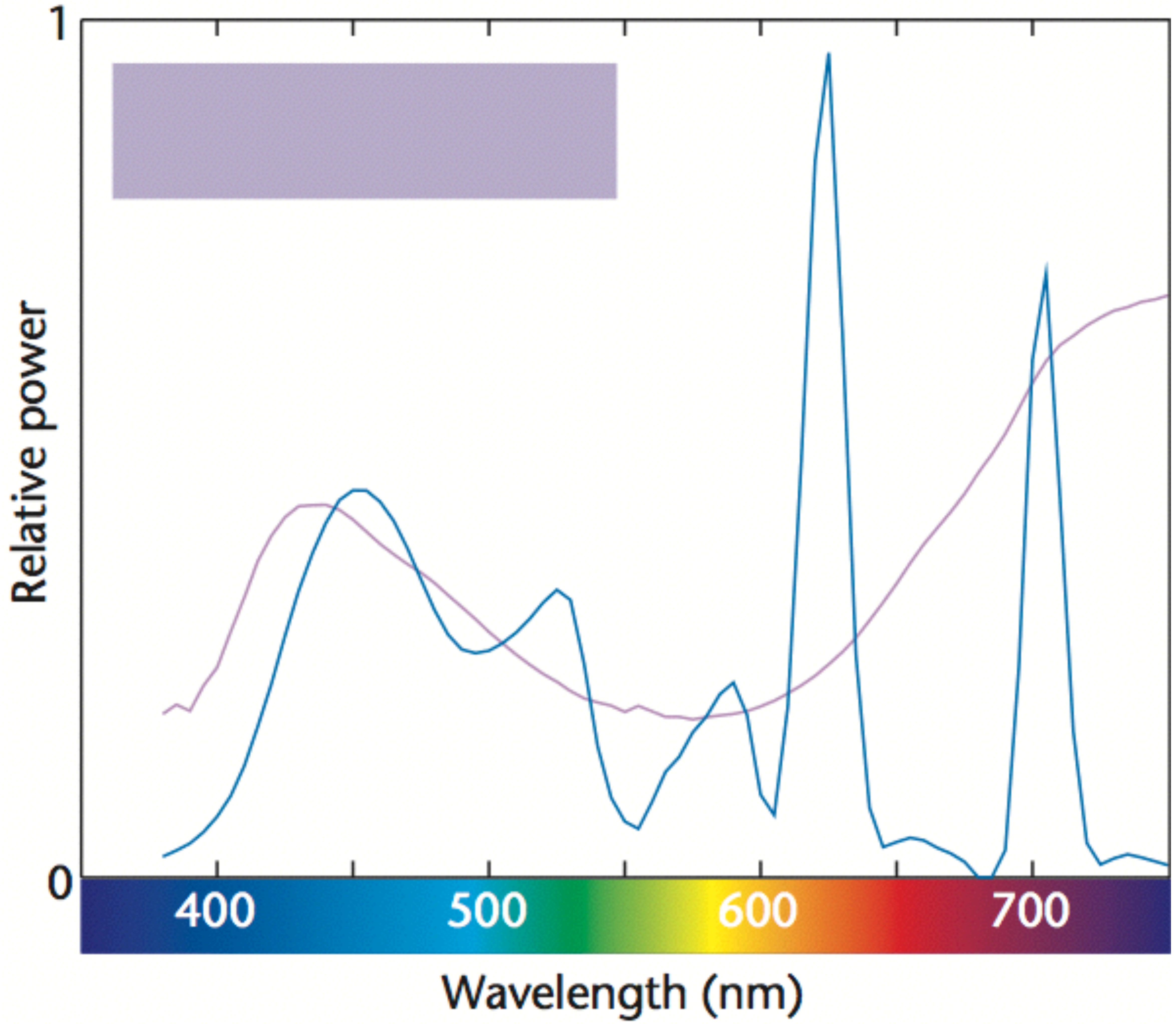
[via M. Meyer]

Human Color Perception



[via M. Meyer]

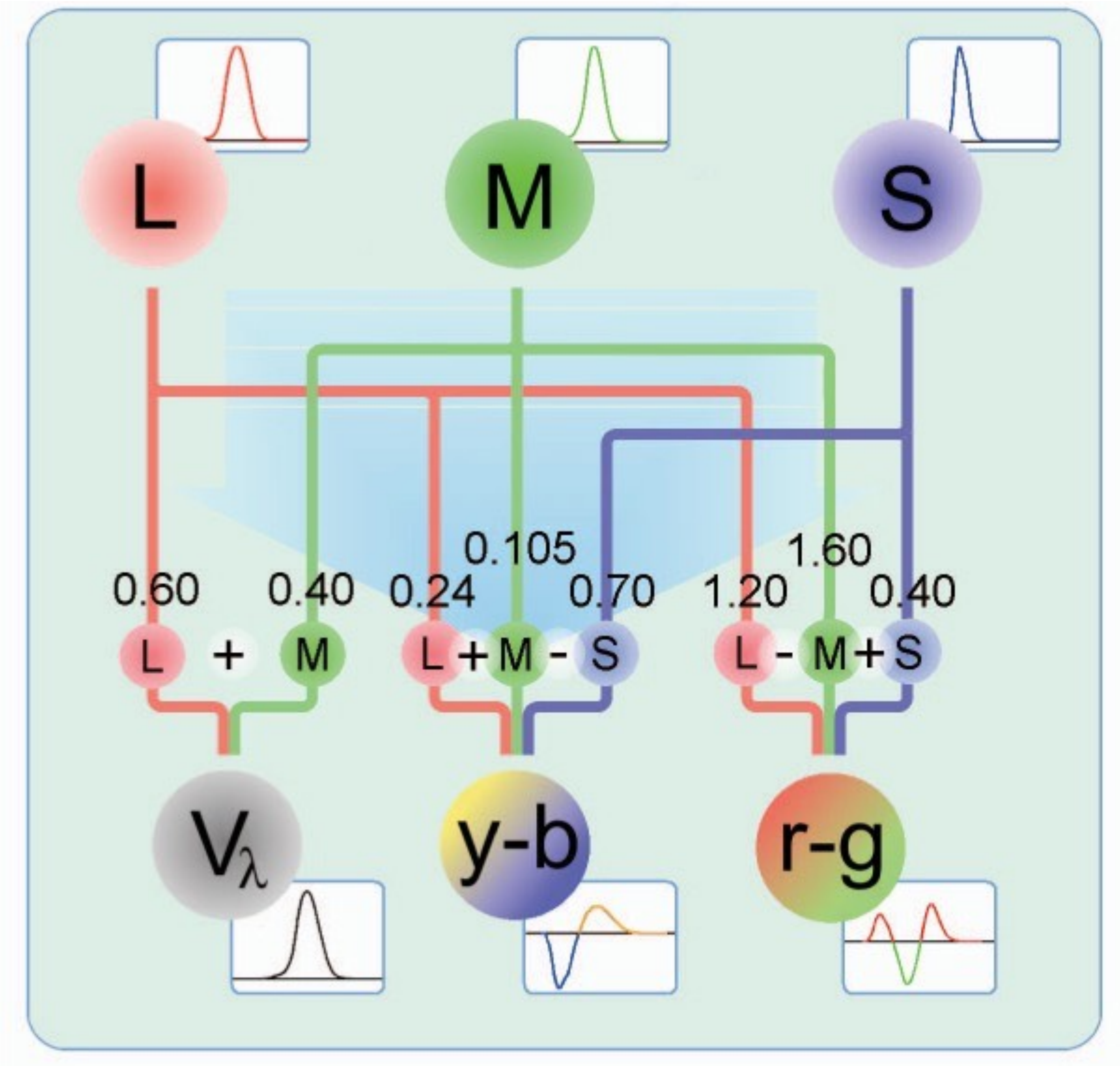
Metamerism



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

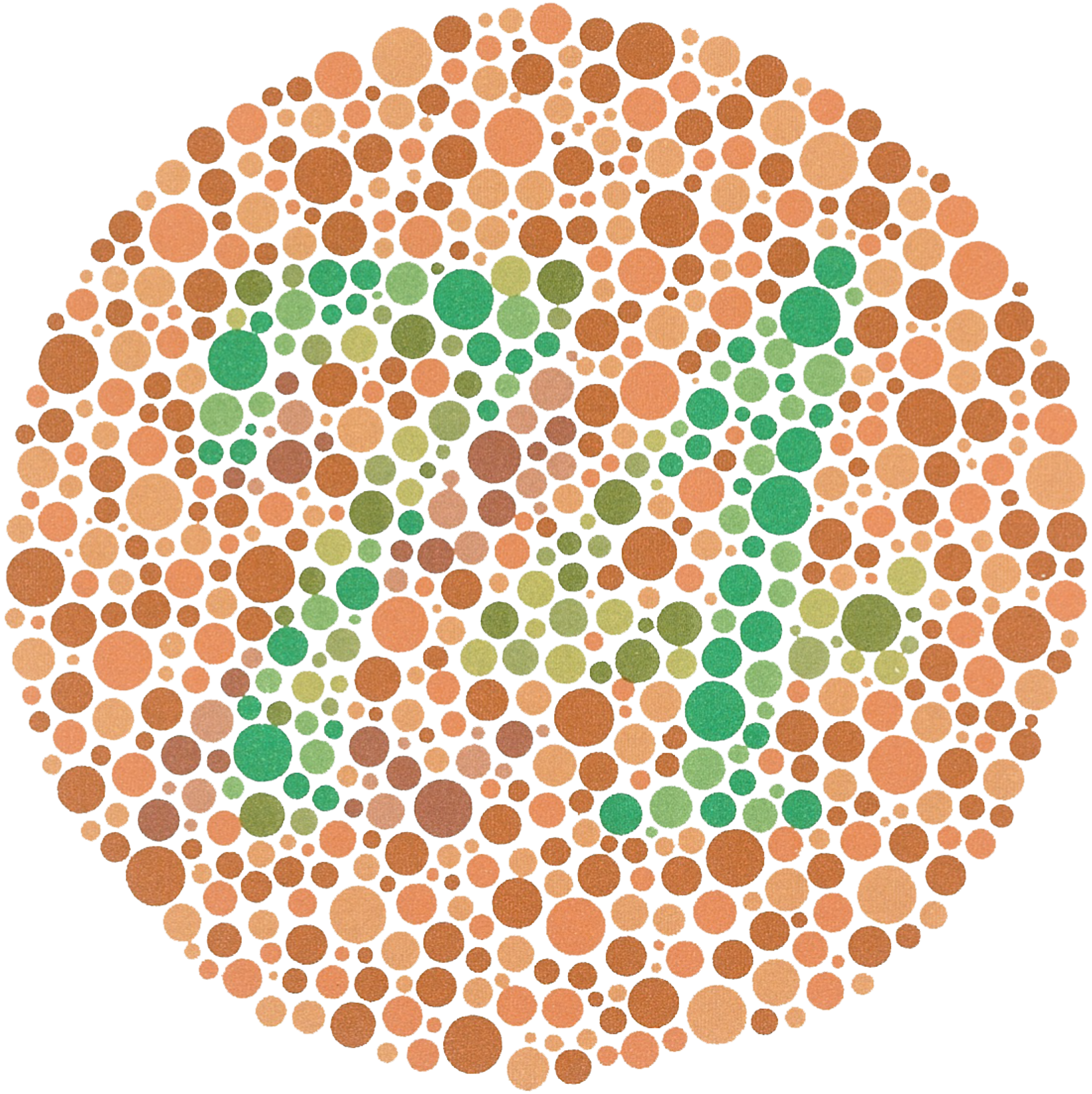
[via M. Meyer]

Opponent Process Theory



[Machado et. al, 2009]

Color Blindness

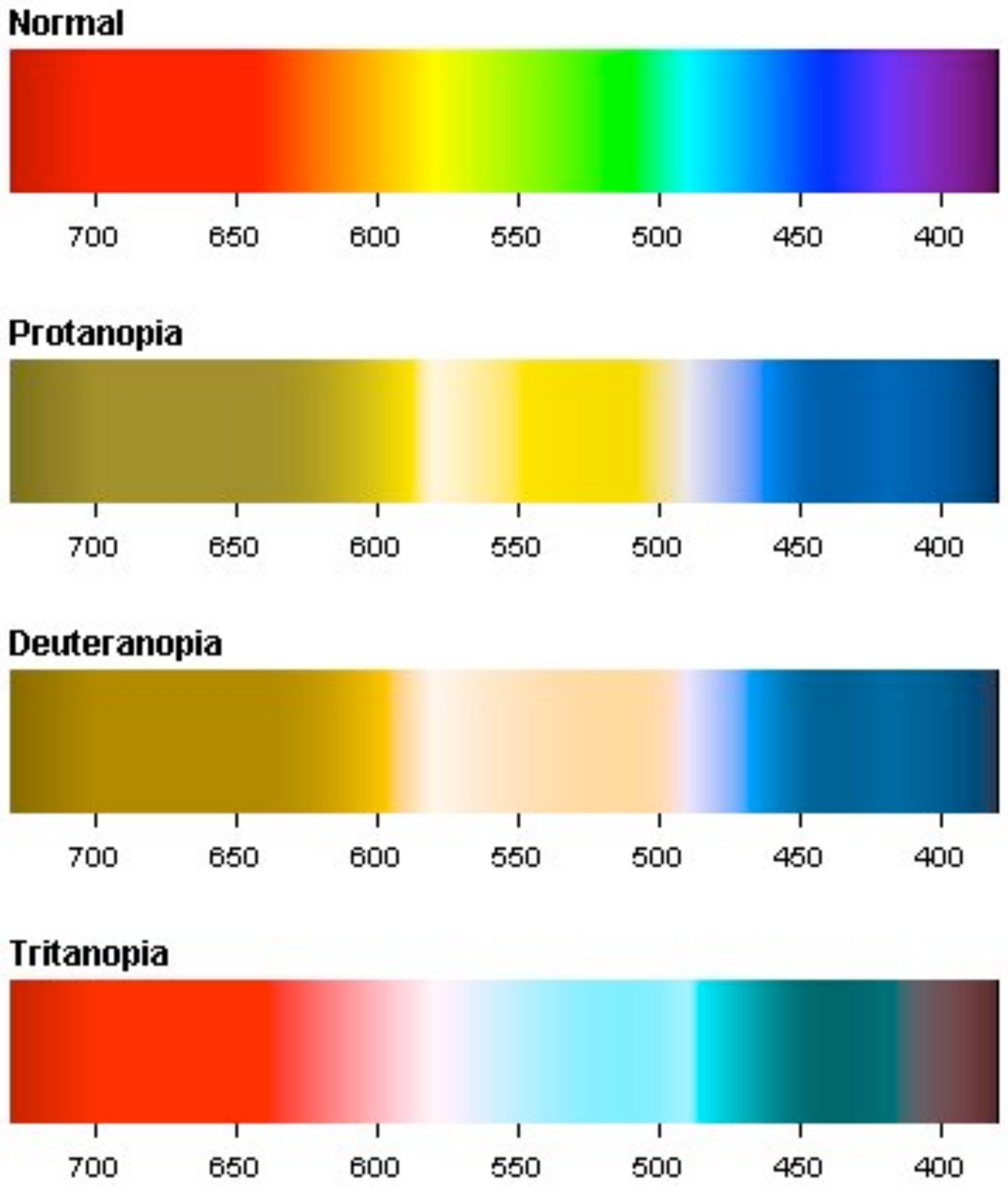


[Ishihara (Plate 9) via Wikipedia]

Color Blindness

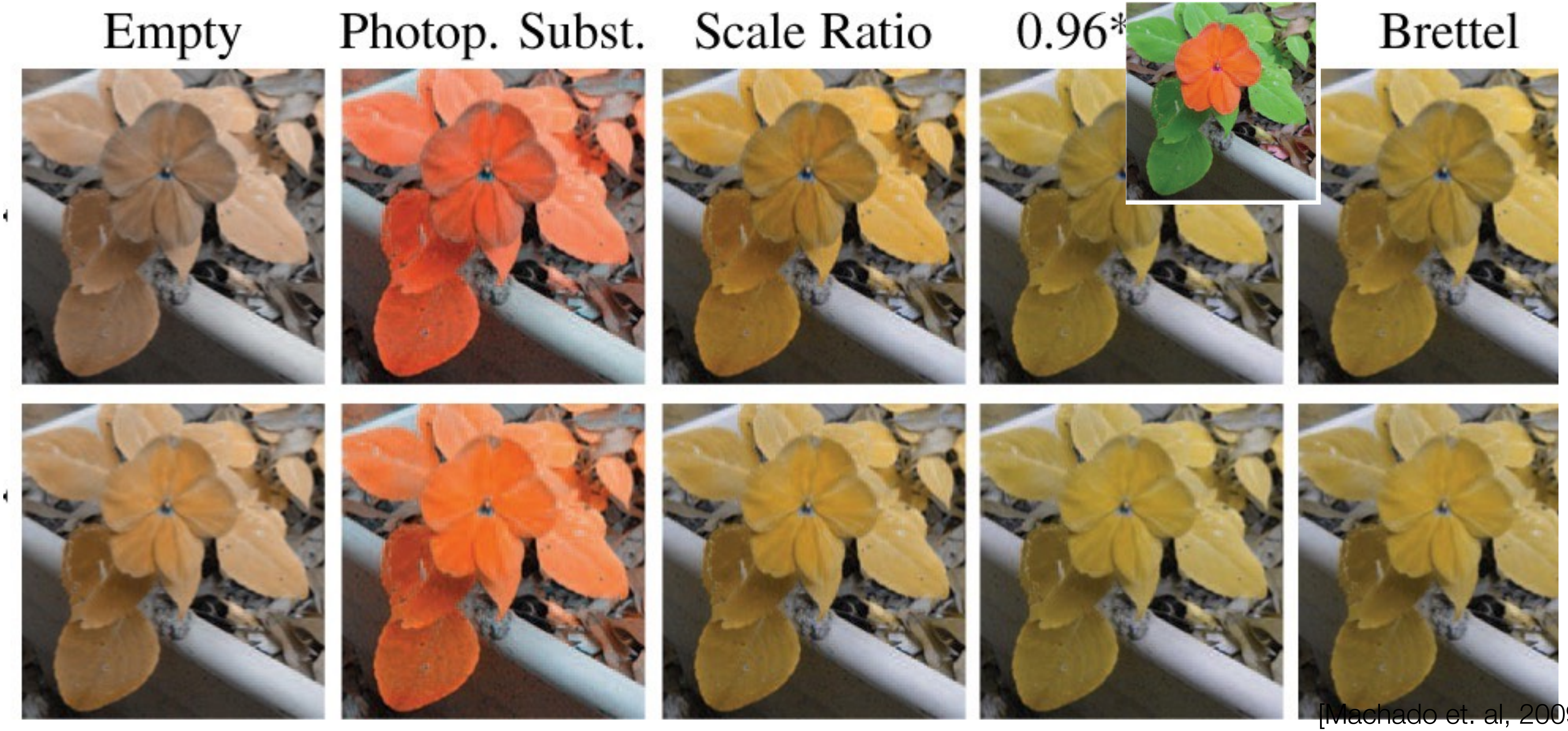
- Sex-linked: 8% of males and 0.4% of females of N. European ancestry
- Abnormal distribution of cones (e.g. missing the S, M, or L types)
- Either dichromatic (only two types of cones) or anomalous trichromatic (one type of cones has a defect)
  - Protanopia (L missing), Protanomaly (L defect)
  - Deuteranopia (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



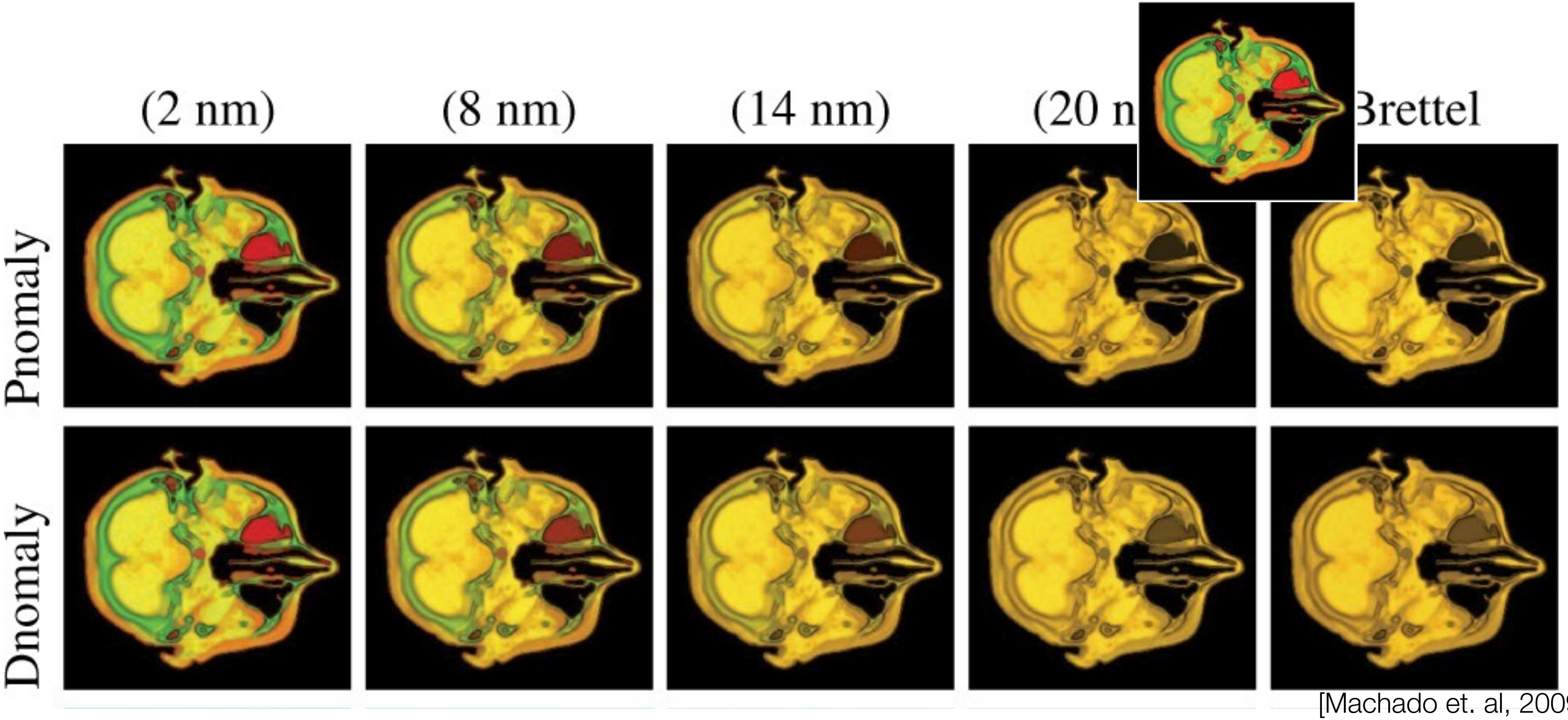
[via M. Meyer]

Simulating Color Blindness



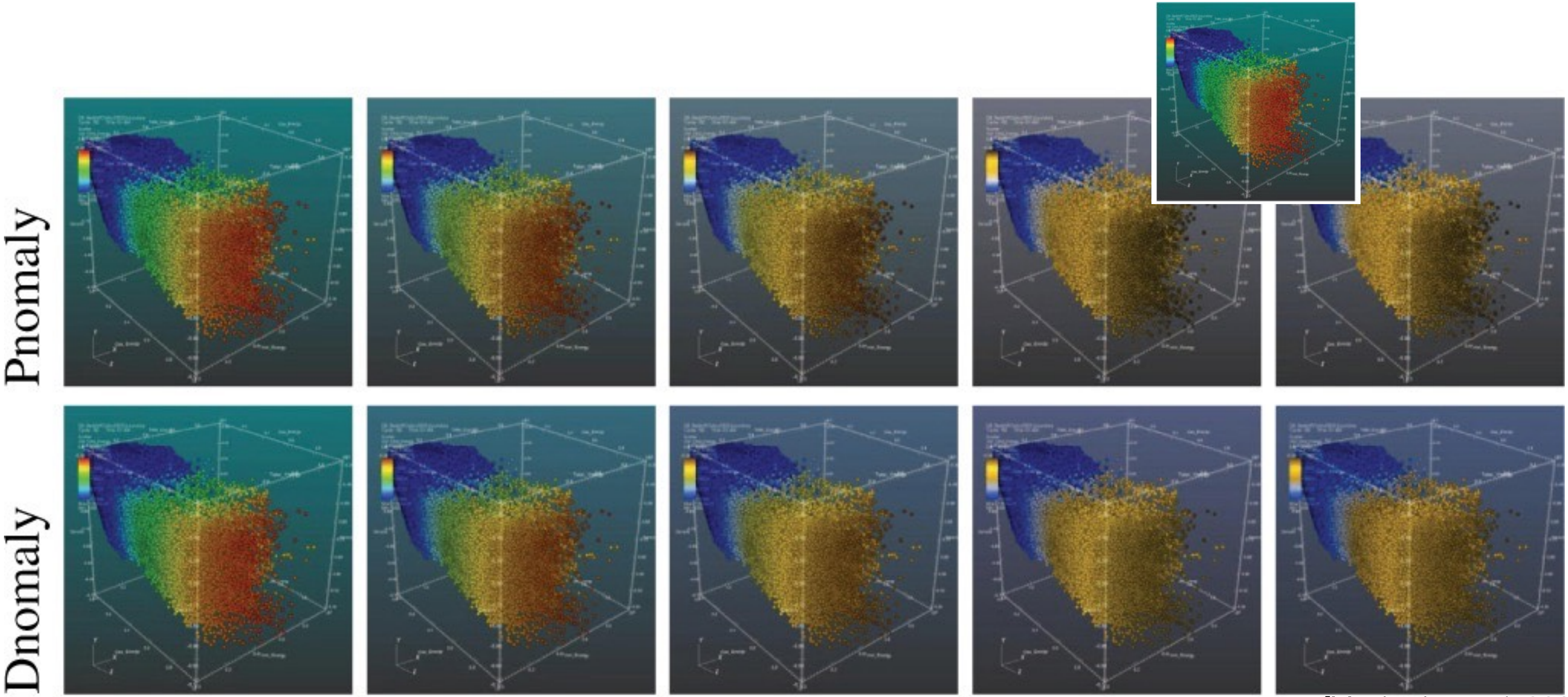
[Machado et. al, 2009]

Simulating Color Blindness



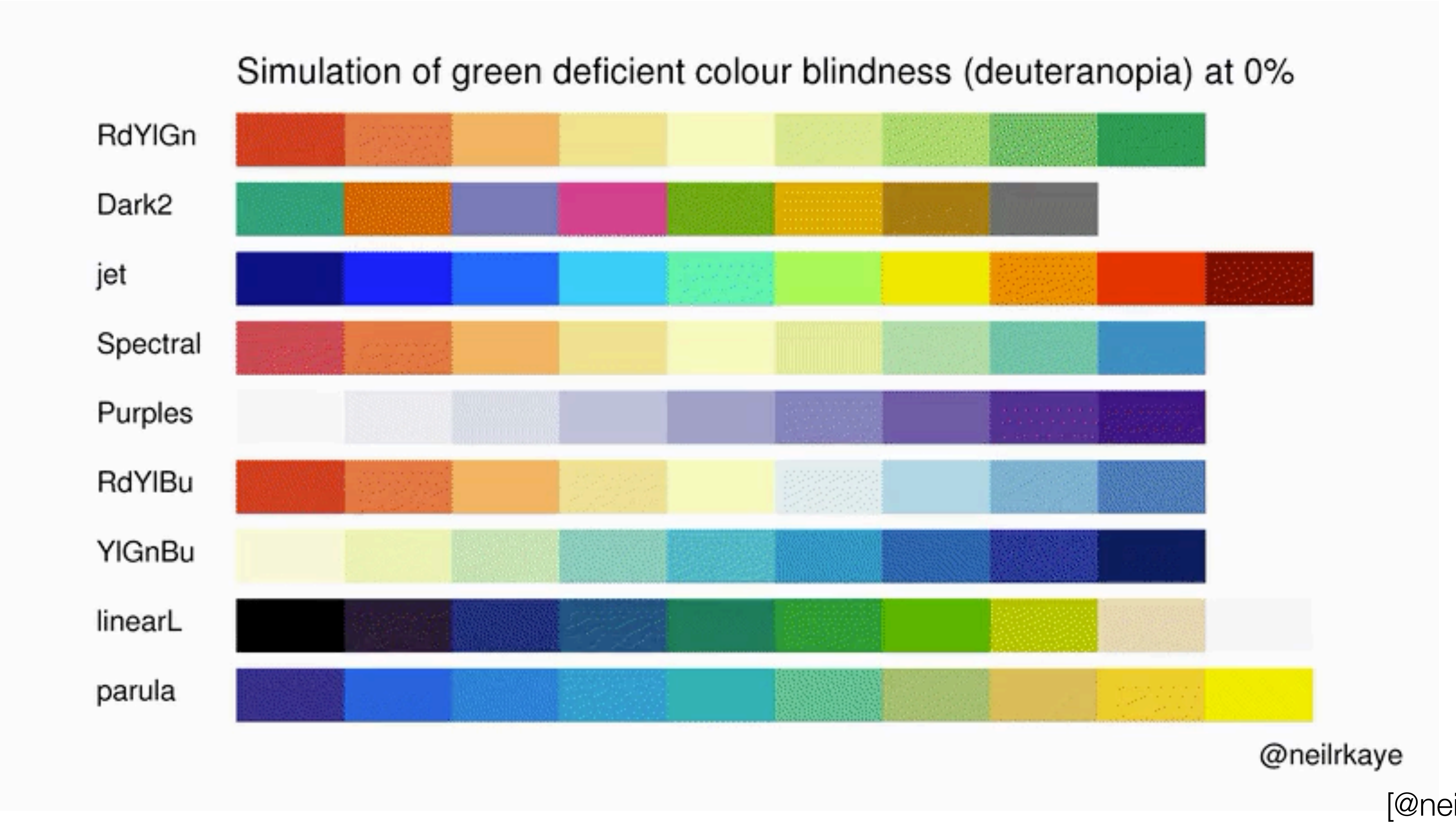
[Machado et. al, 2009]

Simulating Color Blindness



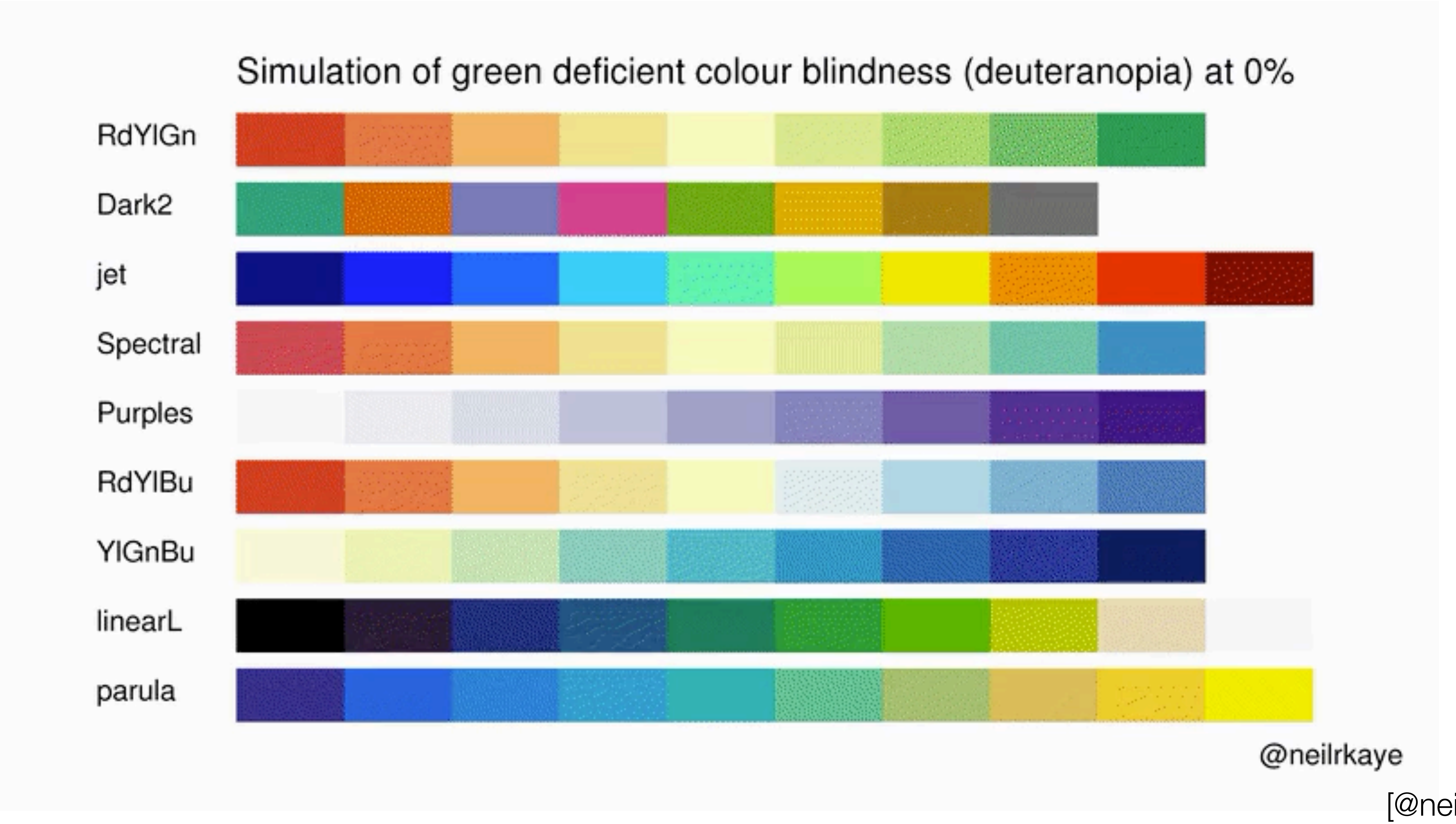
[Machado et. al, 2009]

Simulating Deuteranopia (Colormaps)



[@neilrkaye, reddit]

Simulating Deuteranopia (Colormaps)



[@neilrkaye, reddit]

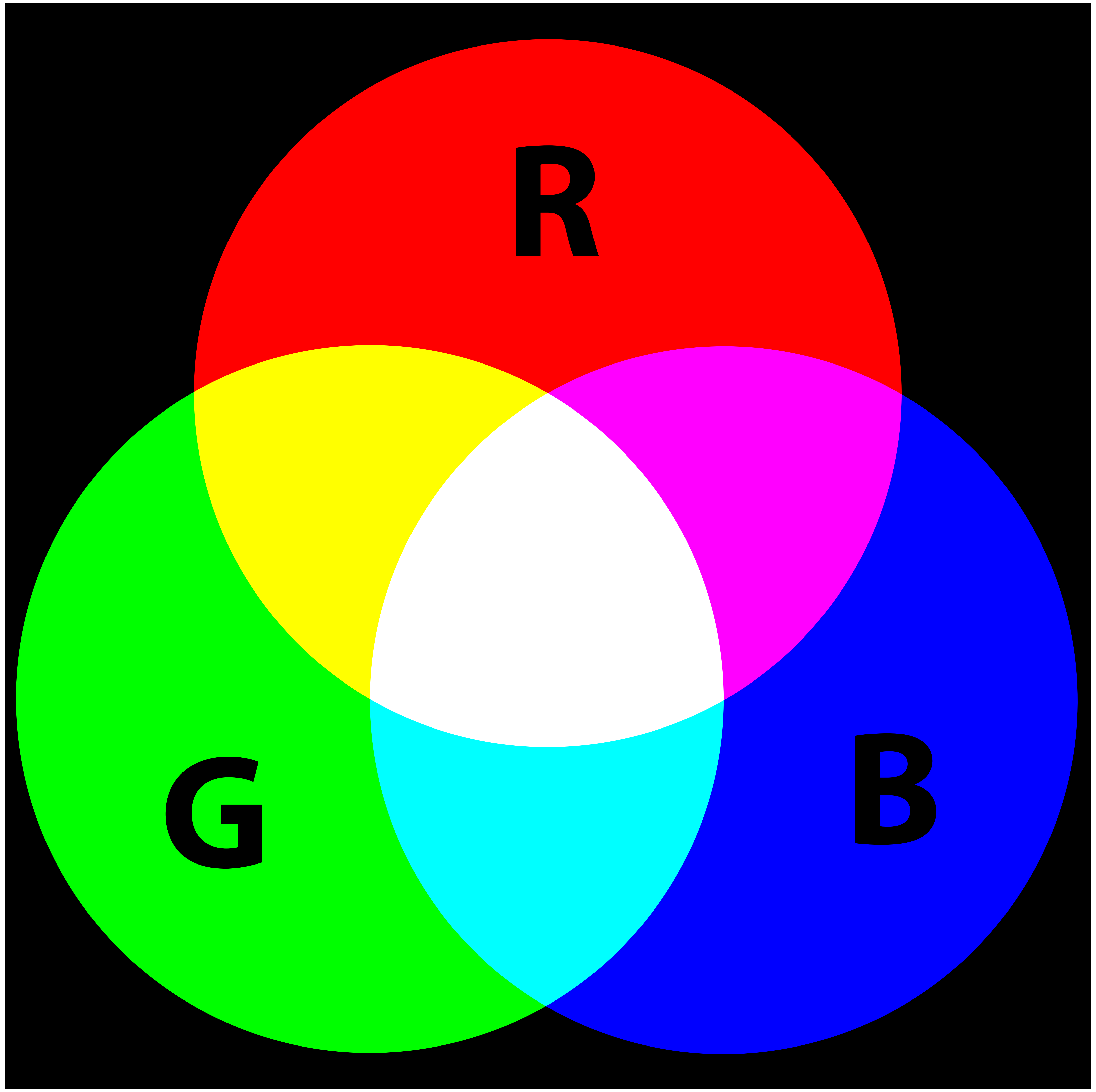
Primary Colors?

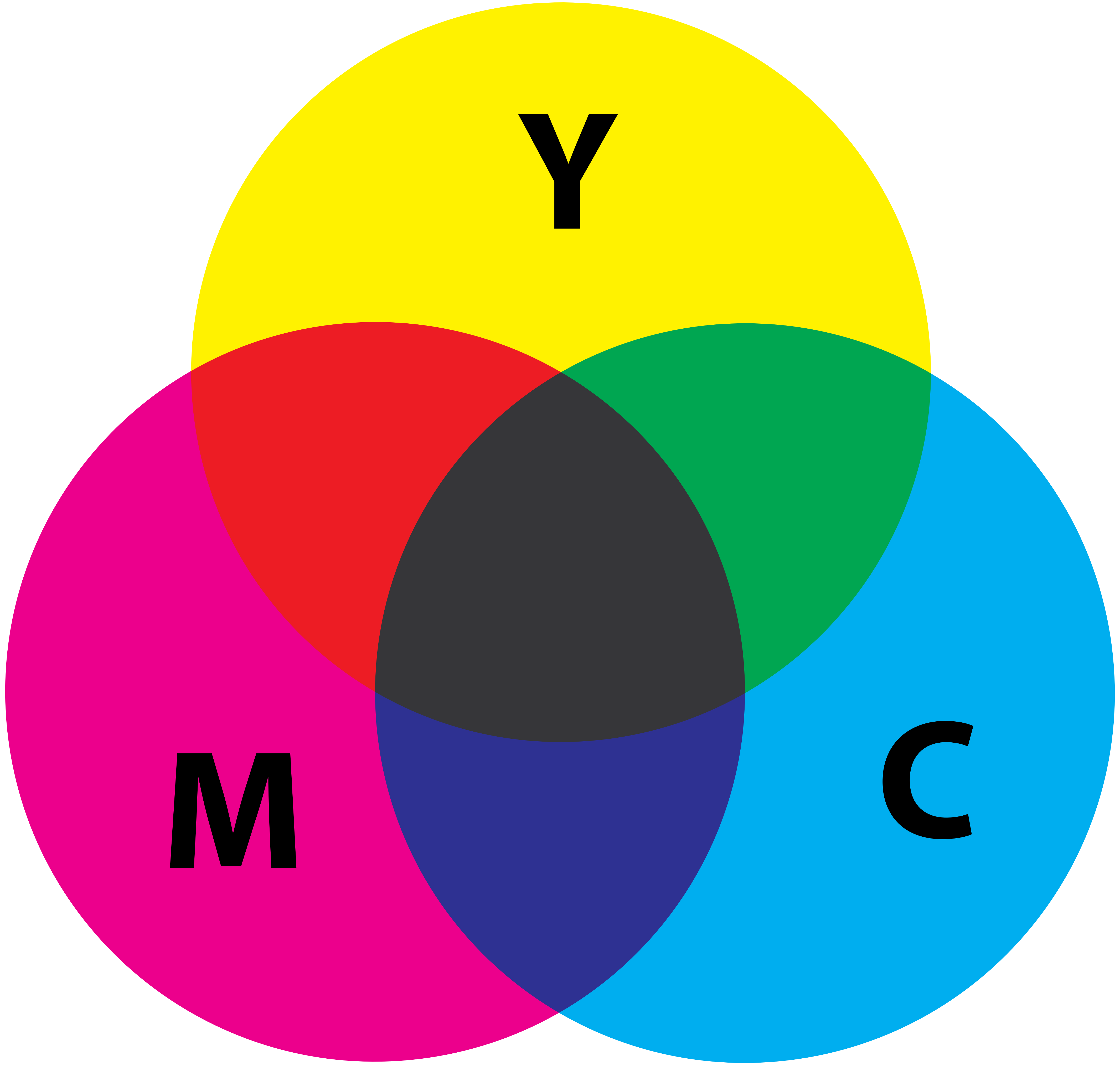
- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow

Primary Colors?

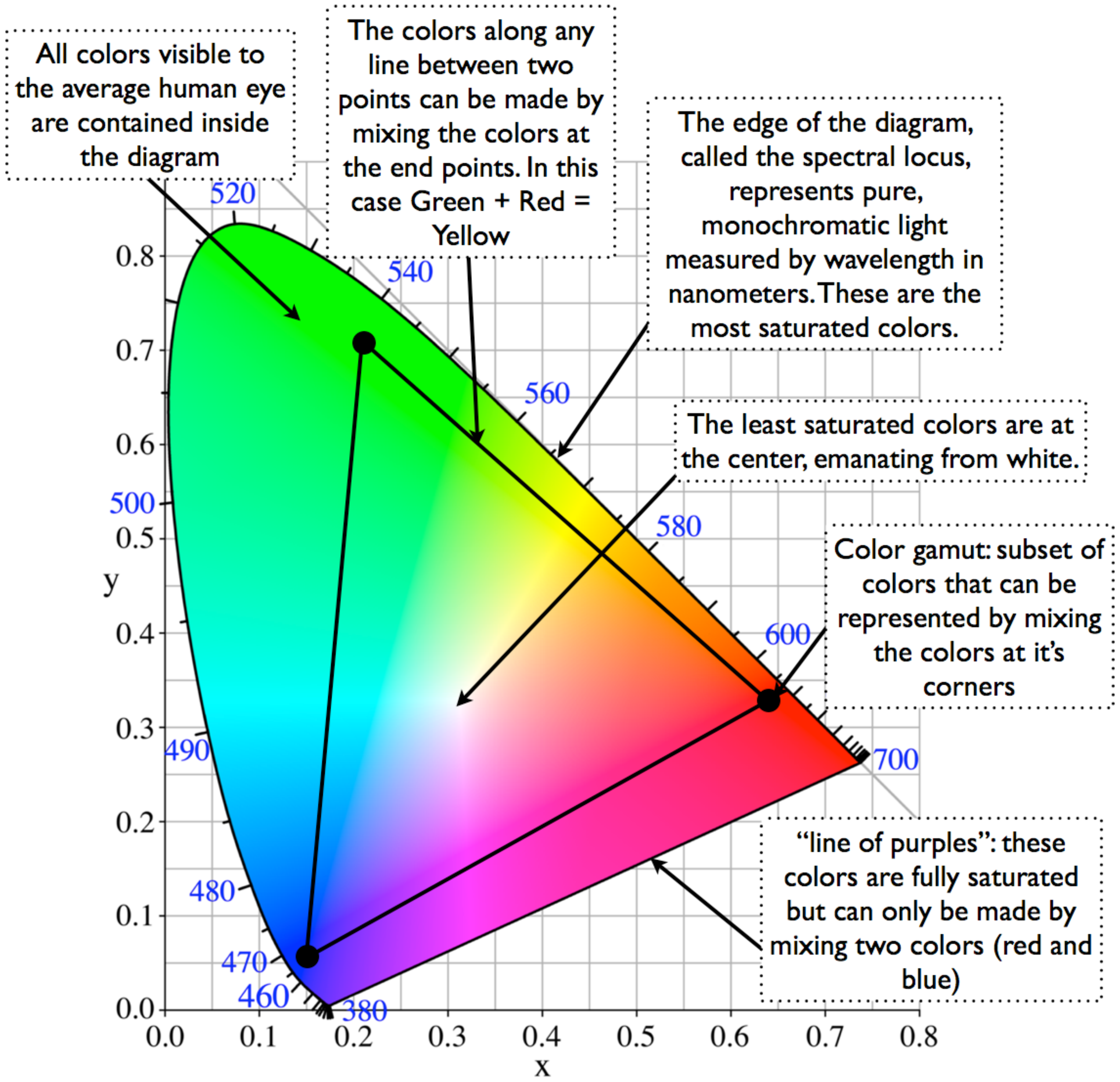
- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
- **All of the above!**

Color Addition and Subtraction





Color Spaces and Gamuts



[Anatomy of a CIE Chromaticity Diagram]

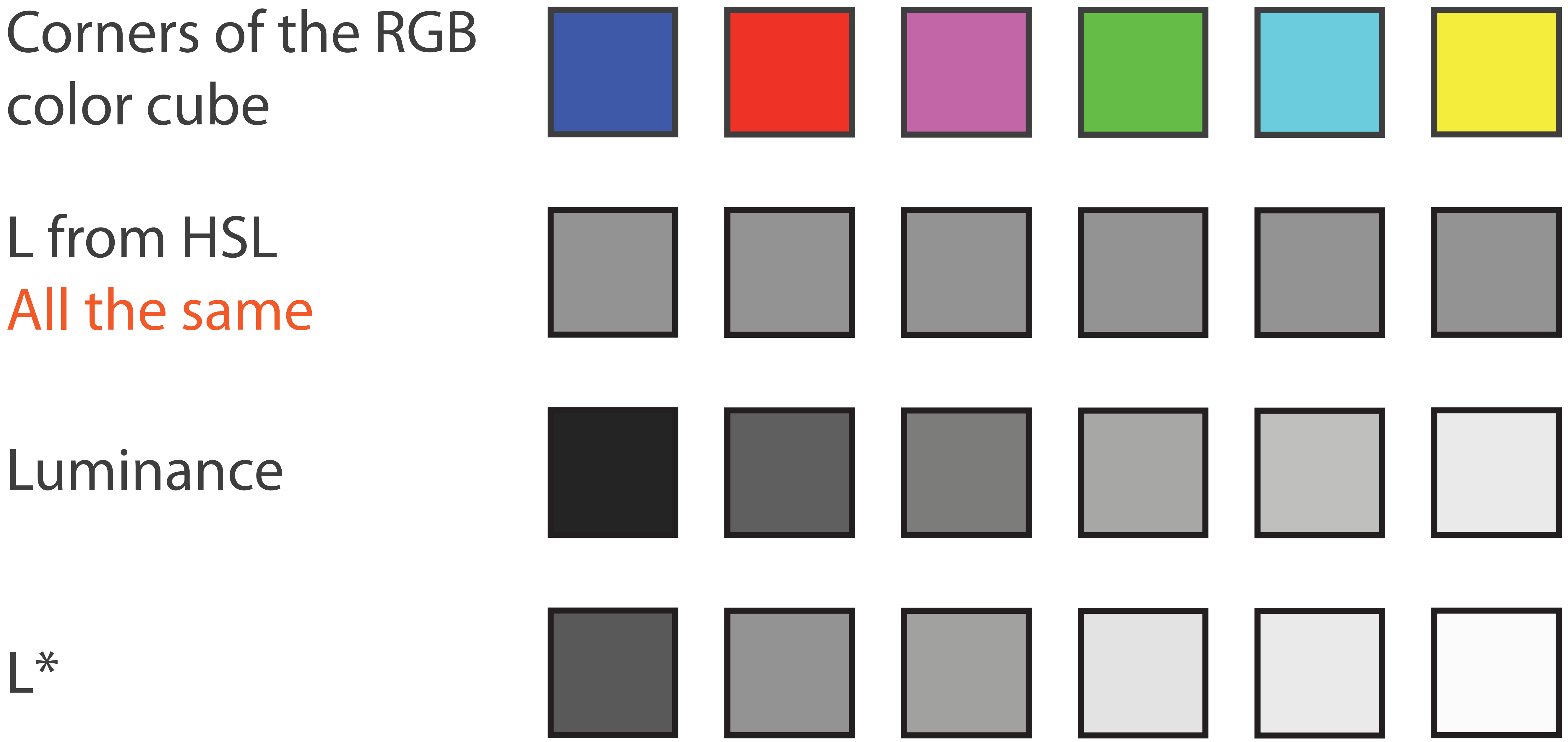
- **Color space**: the organization of all colors in space
  - Often human-specific, what we can see (e.g. CIELAB)
- **Color gamut**: a subset of colors
  - Defined by corners of color space
  - What can be produced on a monitor (e.g. using RGB)
  - What can be produced on a printer (e.g. using CMYK)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

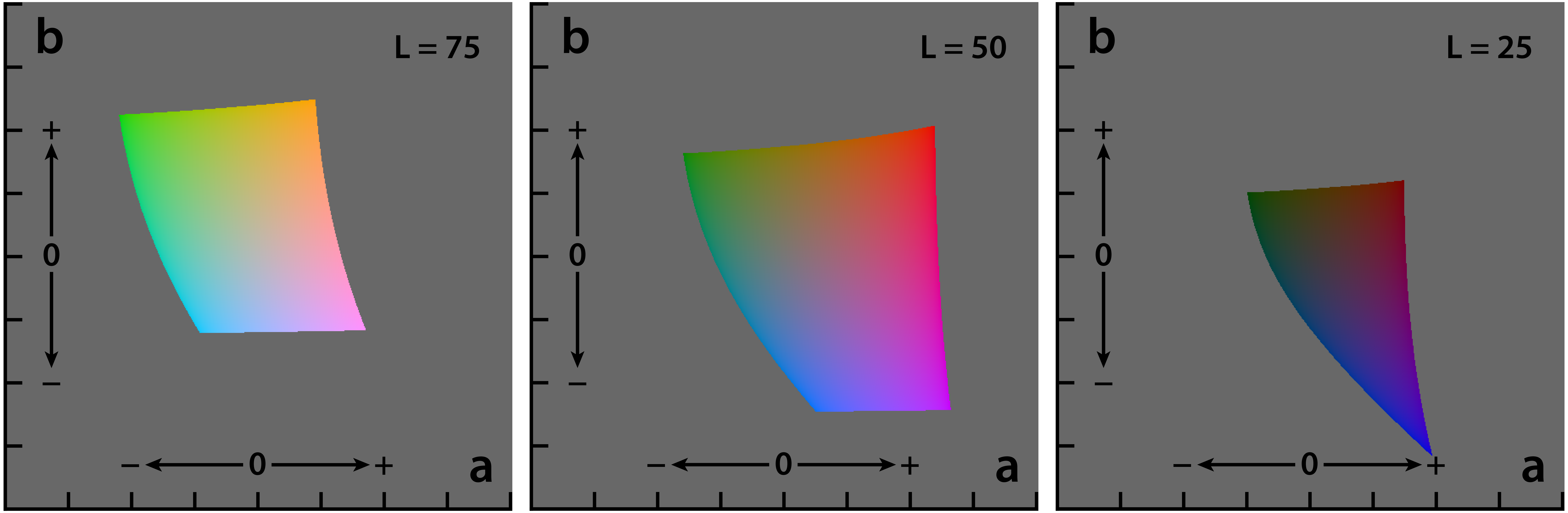
- HSL does not truly reflect the way we perceive color
- Even though colors have the same lightness, we perceive their luminance differently
- Our perception (L*) is **nonlinear**



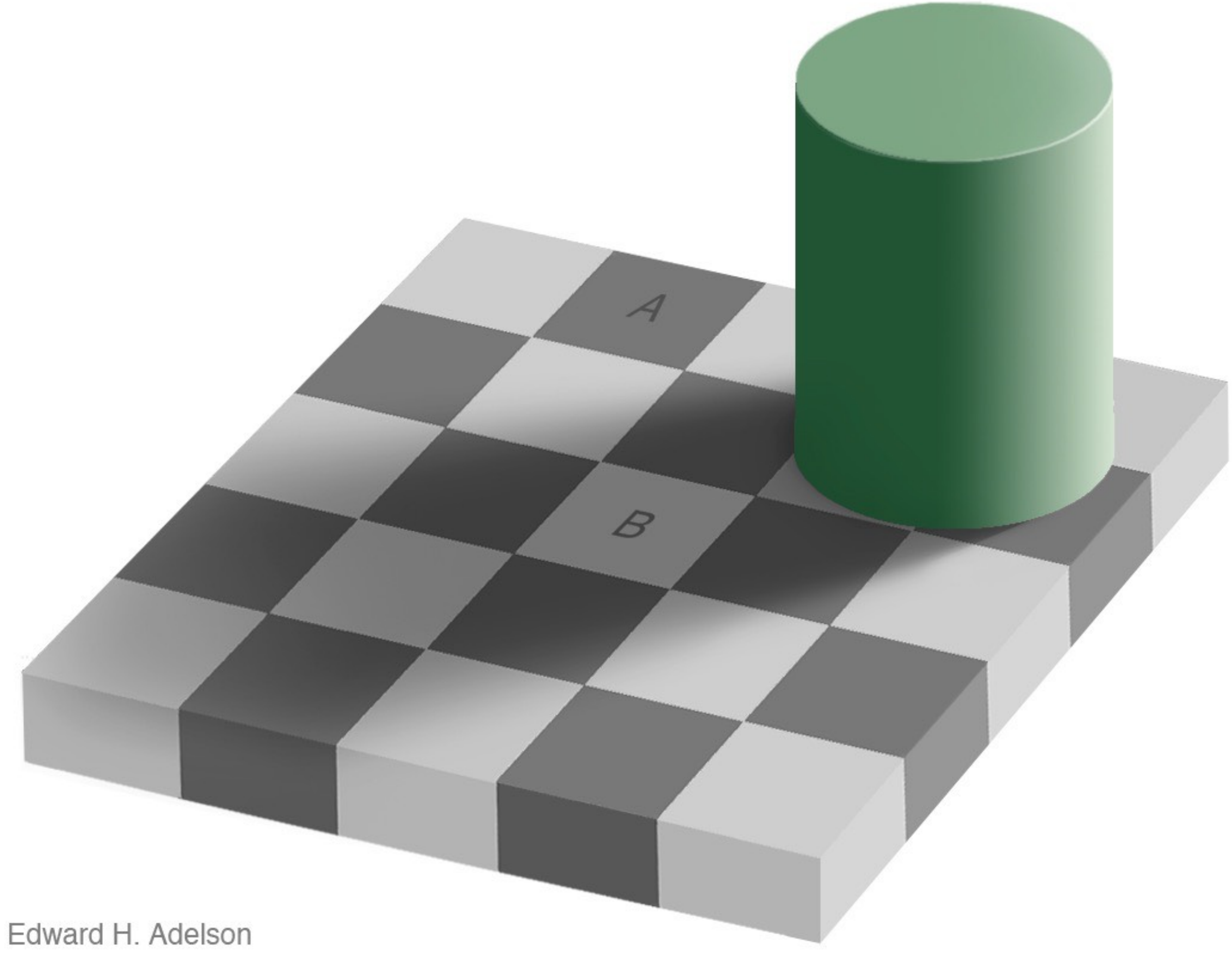
[Munzner (ill. Maguire), 2014 (based on Stone, 2006)]

Perceptually Uniform Color Spaces

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

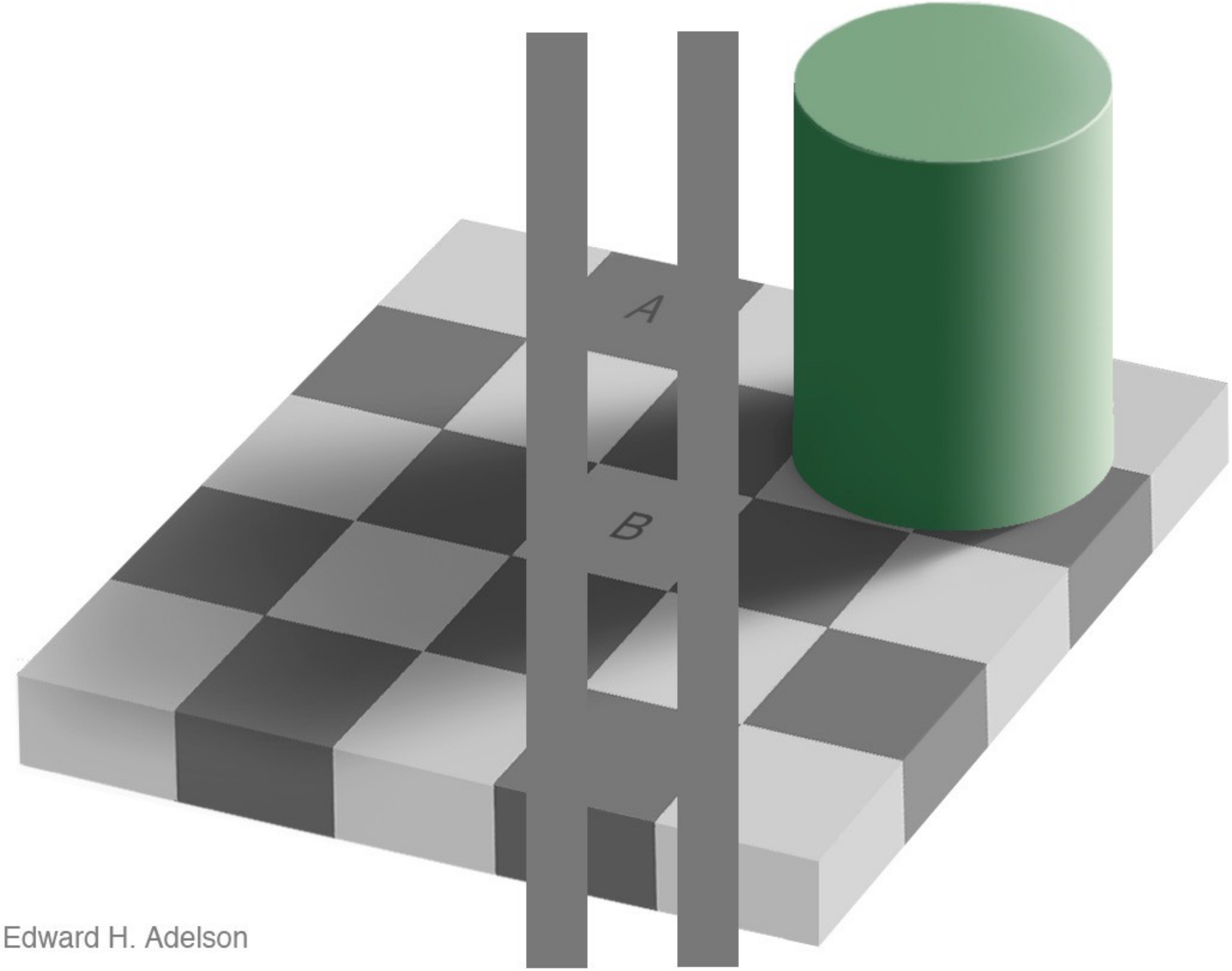


Luminance Perception (Spatial Adaption)



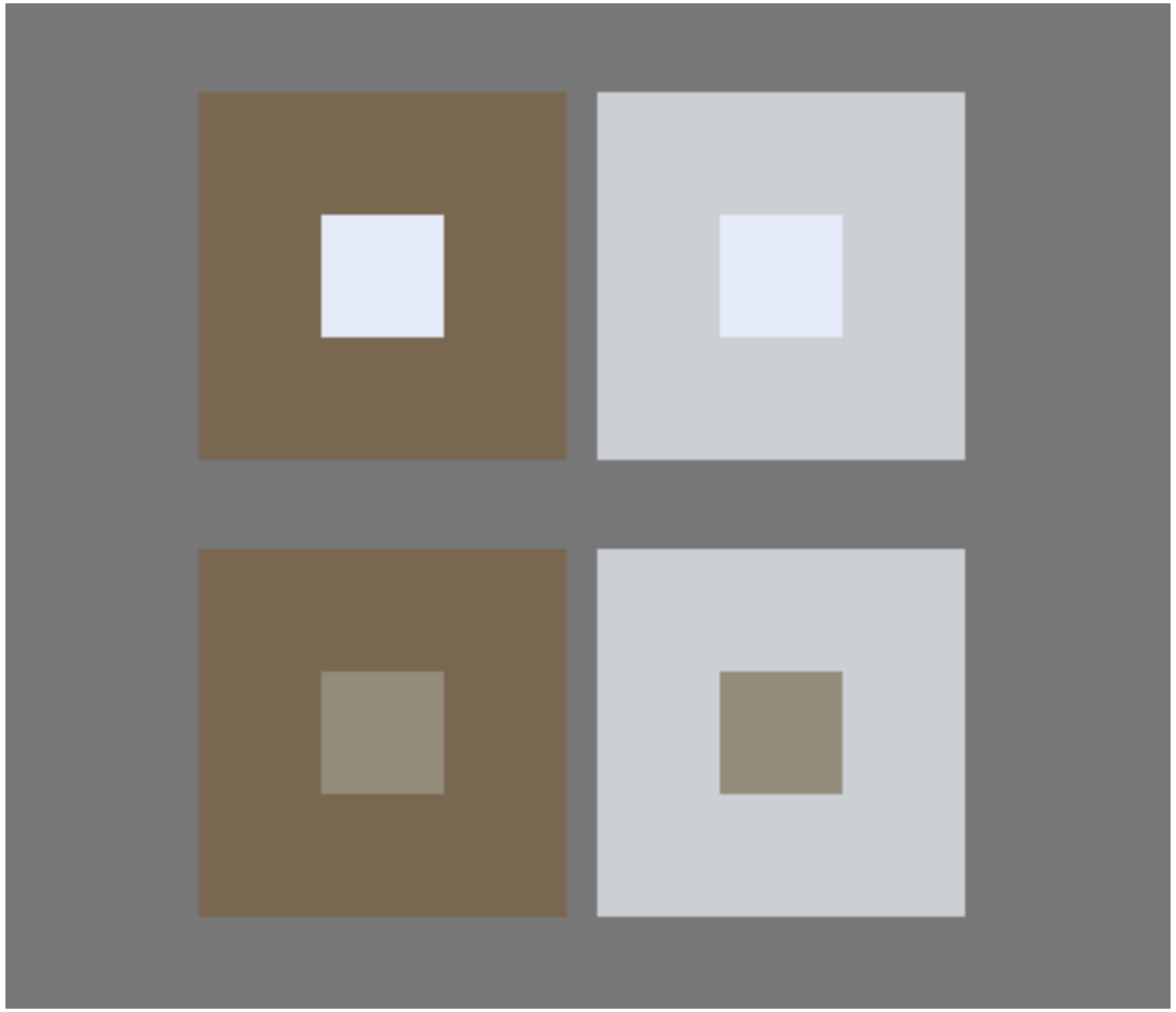
[E. H. Adelson, 1995]

Luminance Perception (Spatial Adaption)

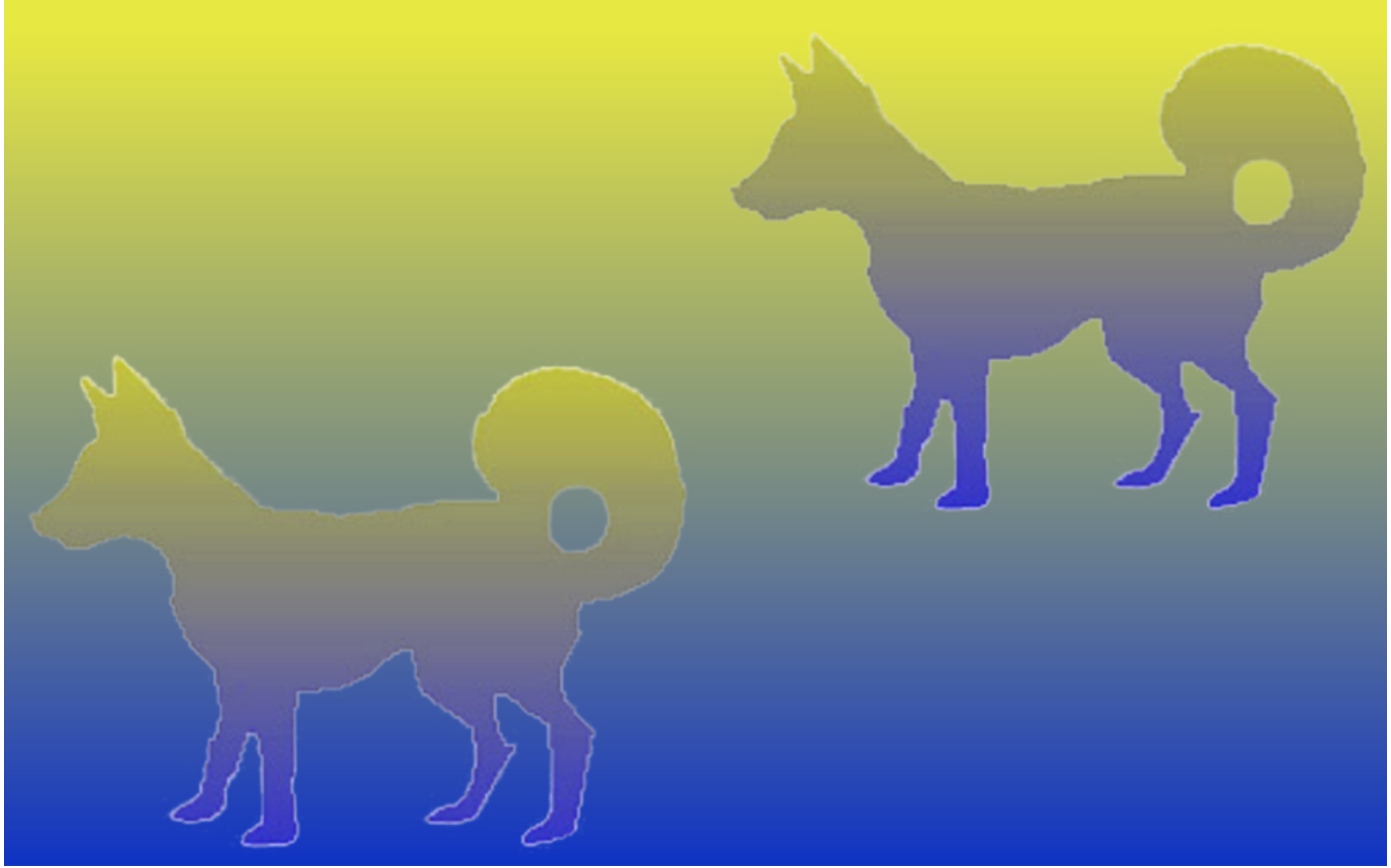


[E. H. Adelson, 1995]

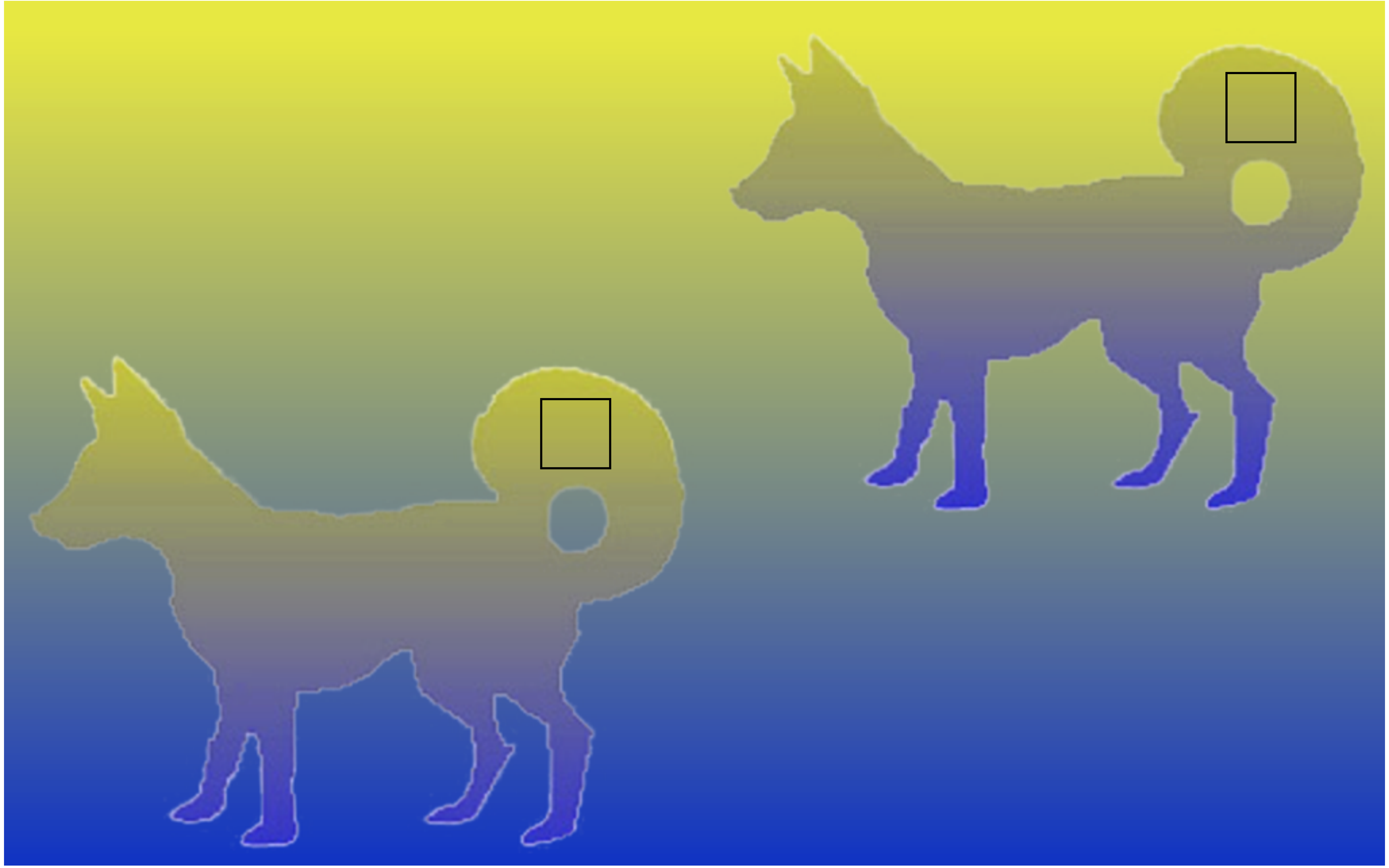
Simultaneous Contrast



Simultaneous Contrast

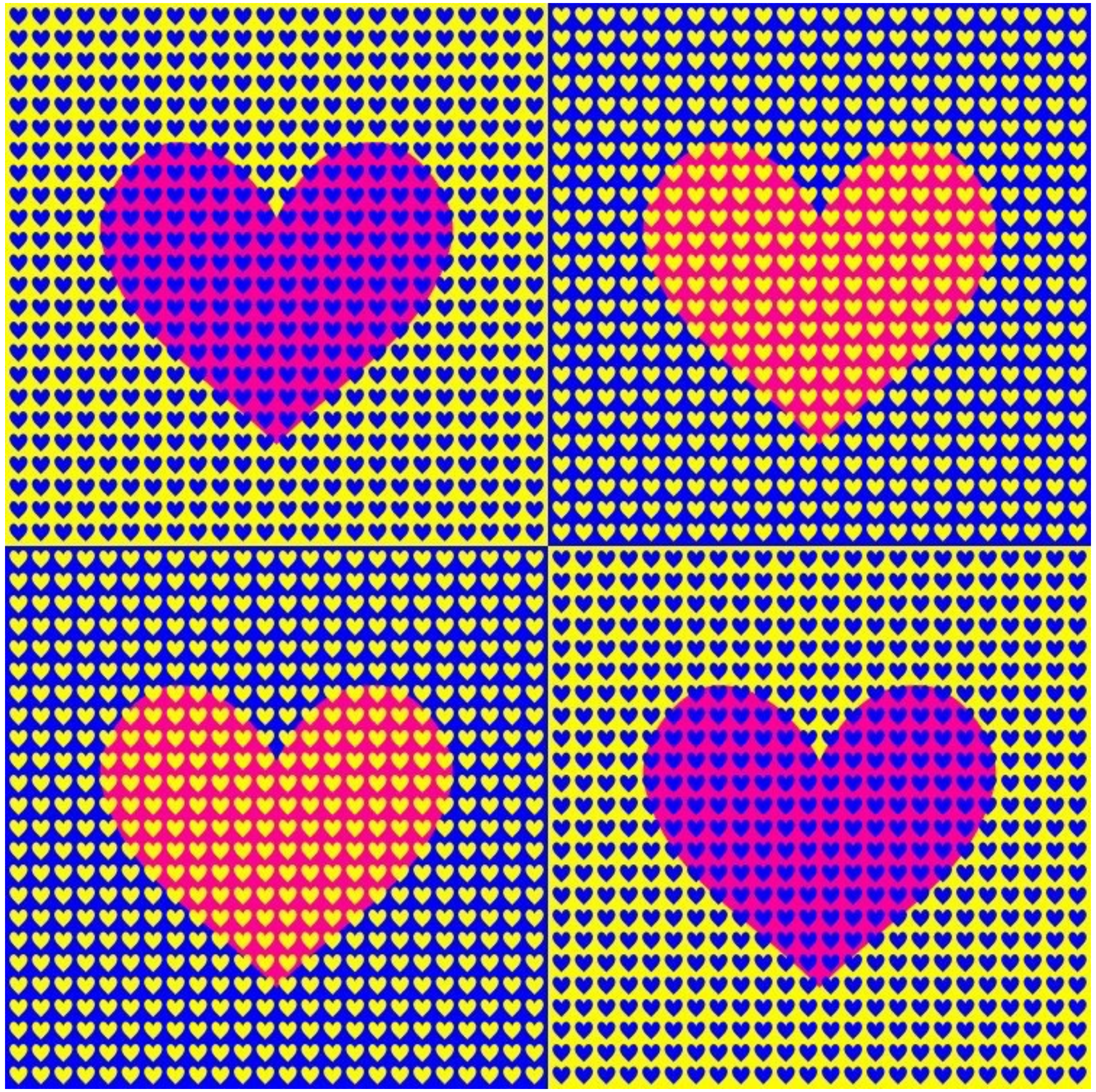


Simultaneous Contrast



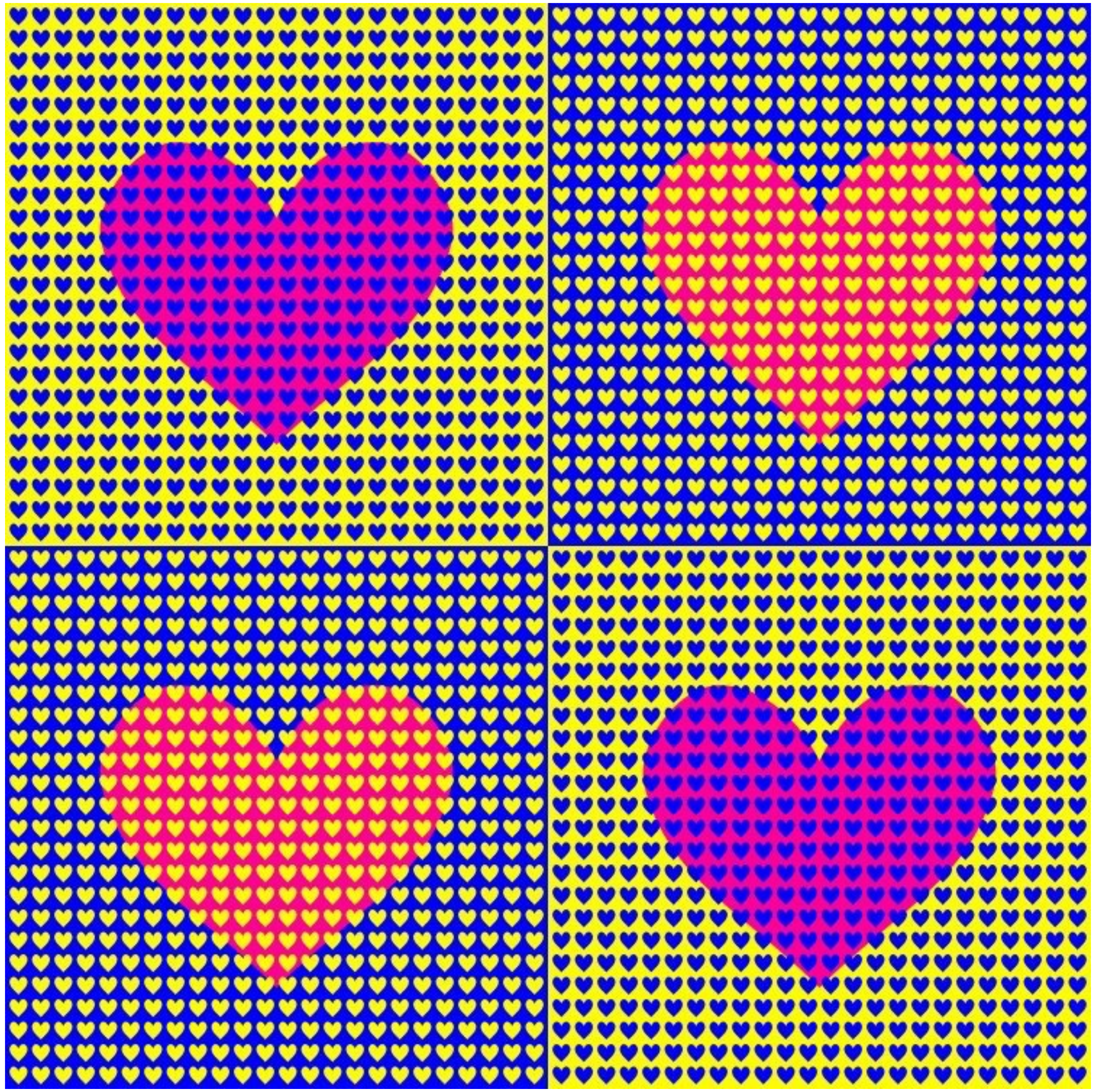
Simultaneous Contrast

What colors?



[A. Kitaoka]

What colors?



Red, yellow, blue

Purple, orange do not exist!

[A. Kitaoka]

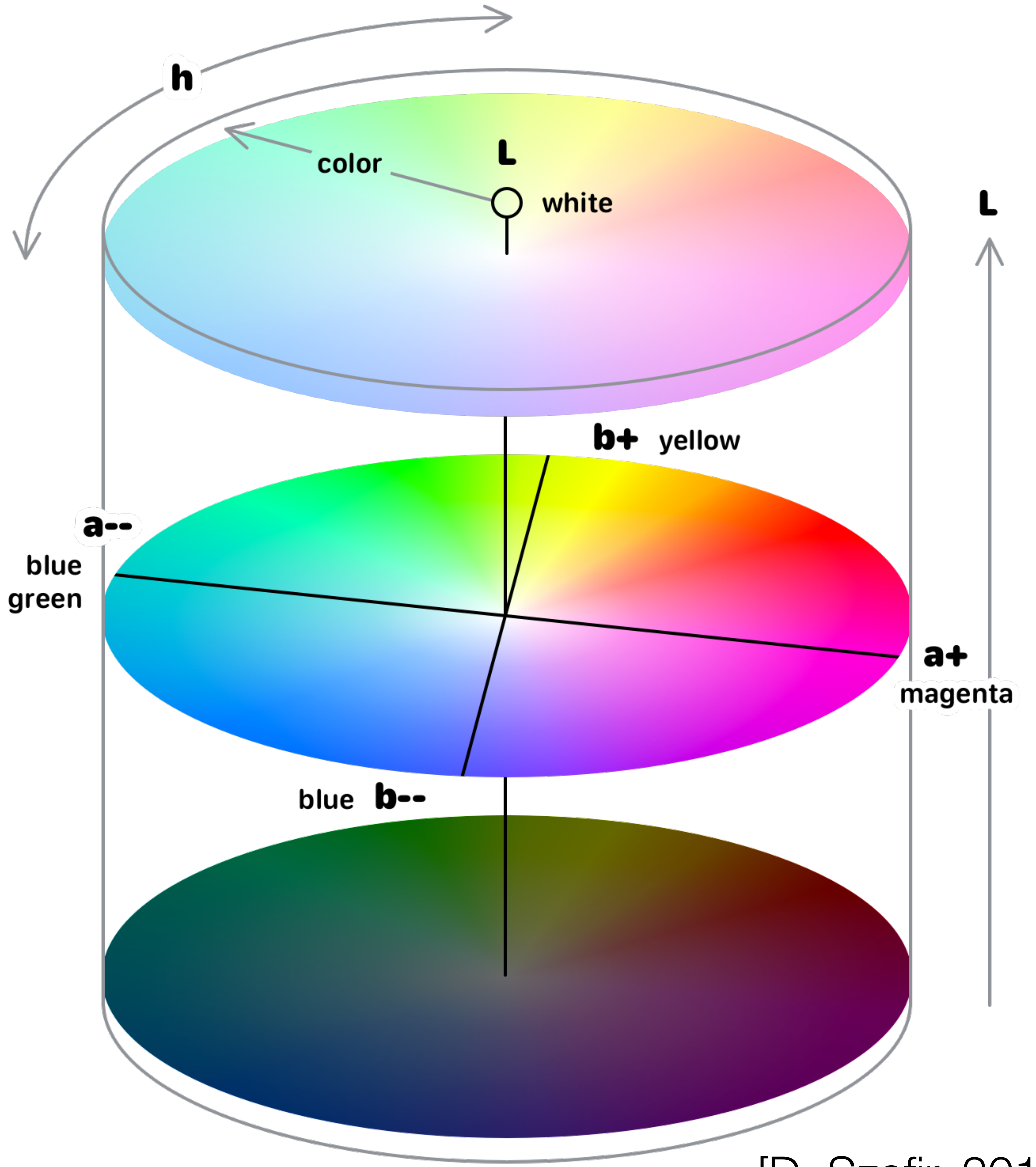
What does this mean for visualization?

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

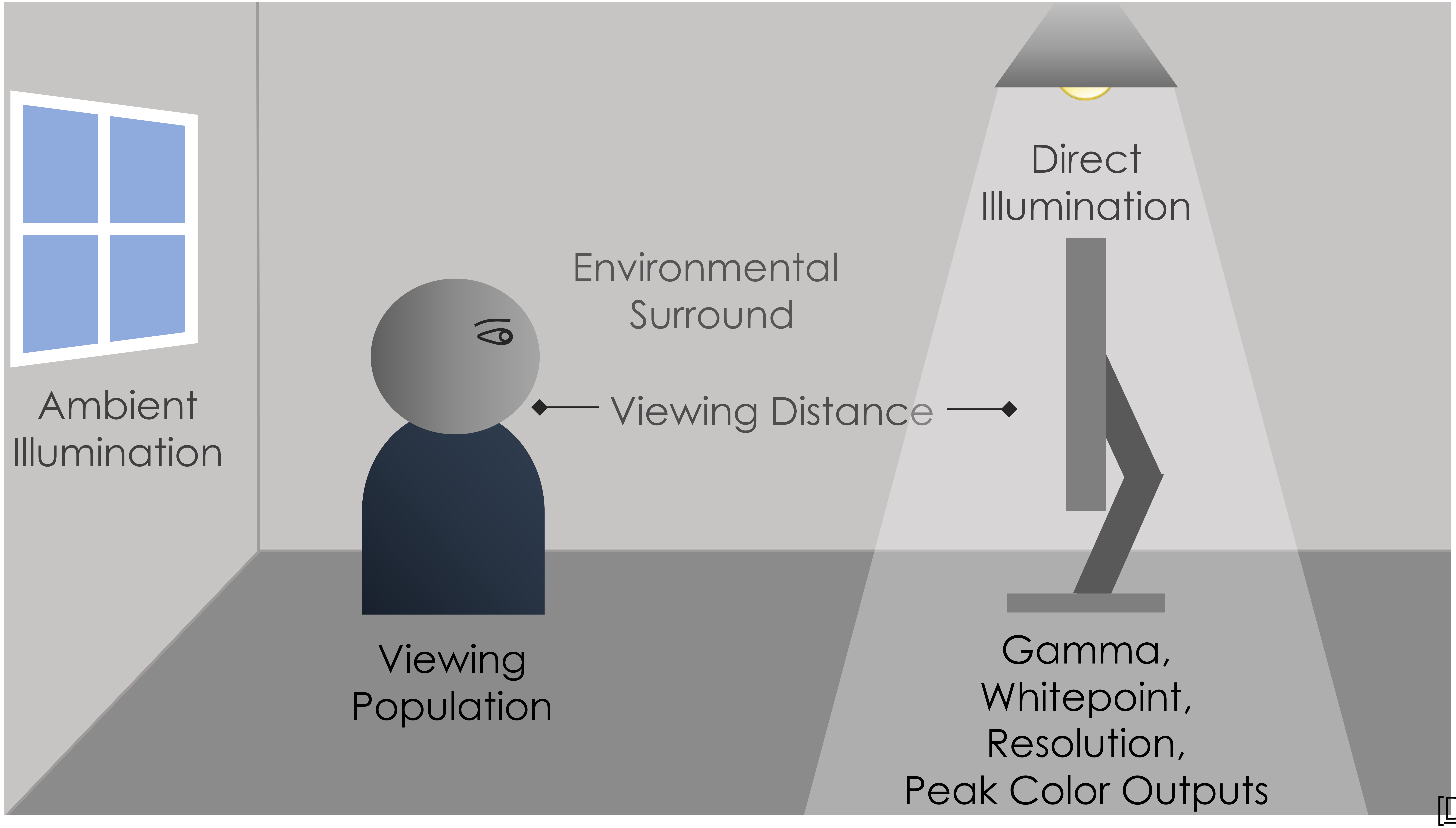


[D. Szafir, 2017]

Simple World Assumption

[D. Szafir, 2017]

Problems with Simple World Assumption

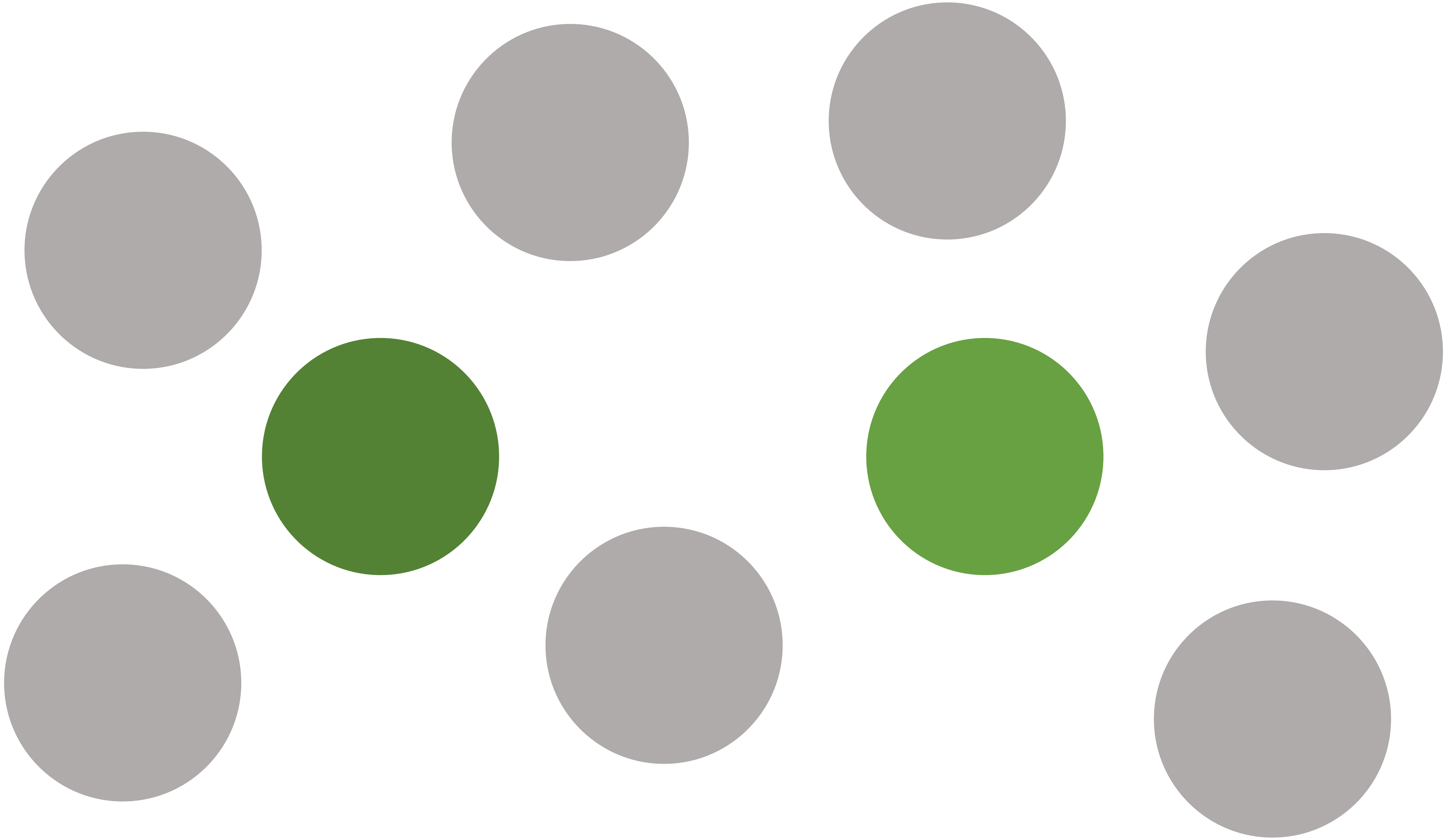


[D. Szafir, 2017]

Isolation Assumption

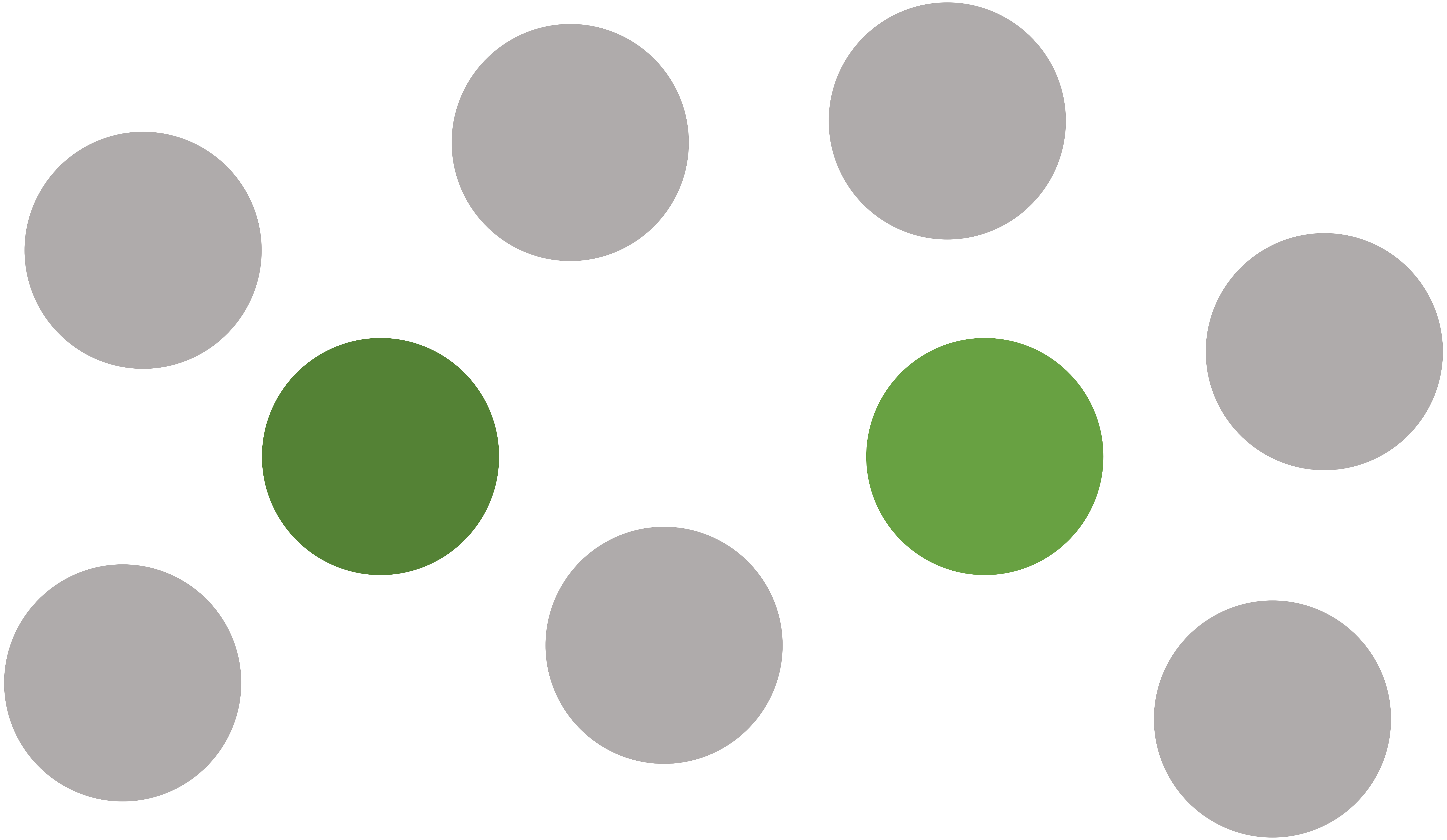
[D. Szafir, 2017]

Problems with Isolation Assumption



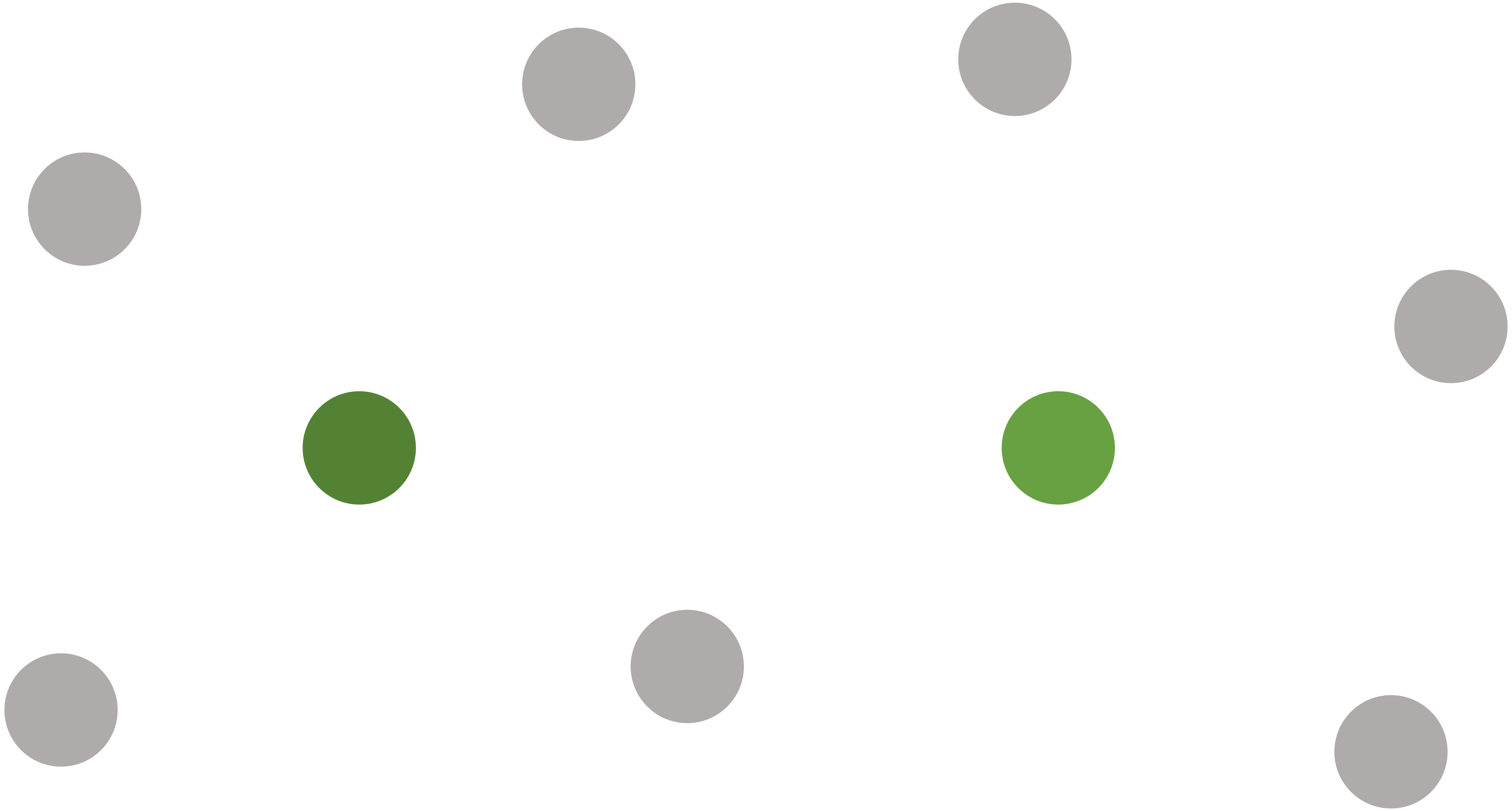
[D. Szafir, 2017]

Geometric Assumption



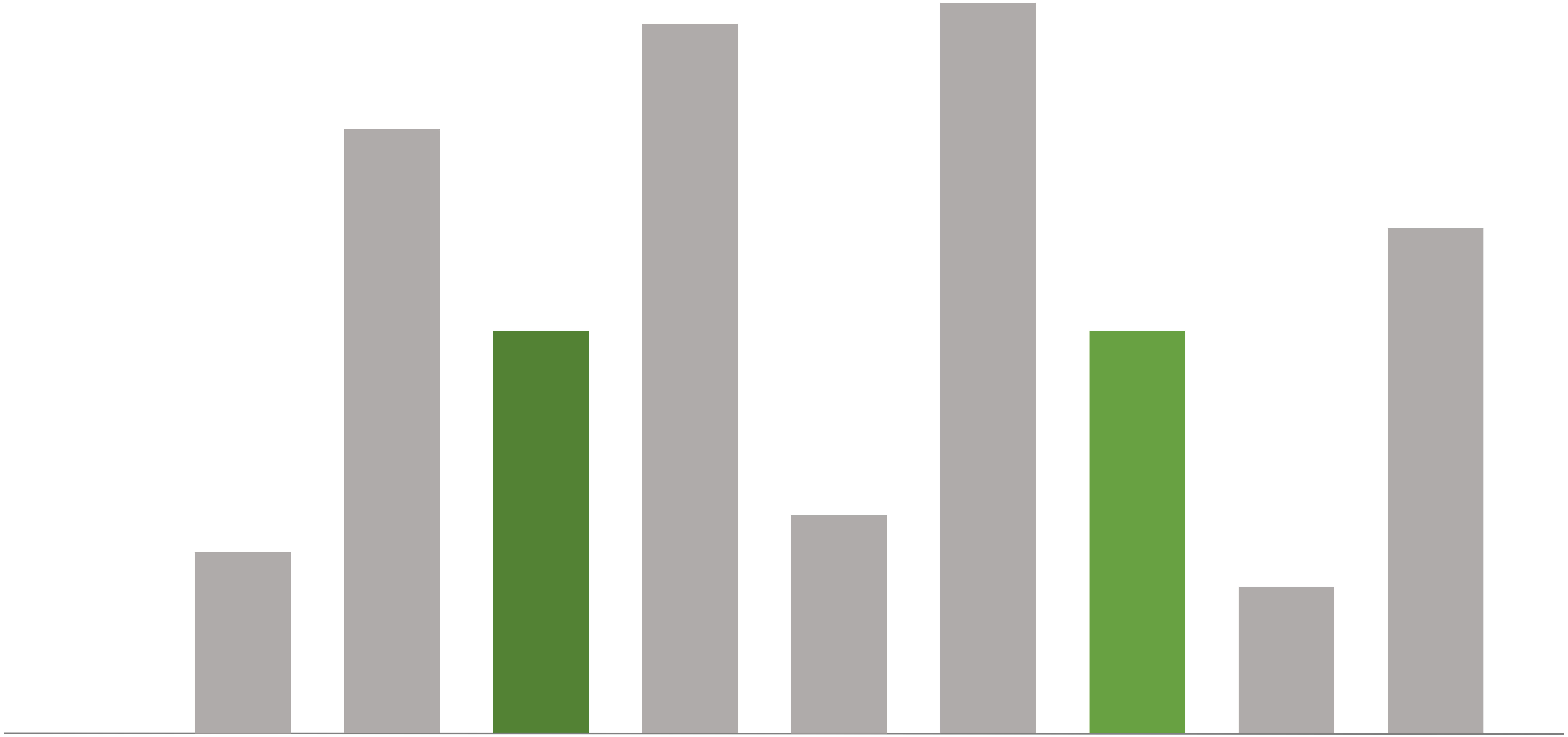
[D. Szafir, 2017]

Size Problem with Geometric Assumption



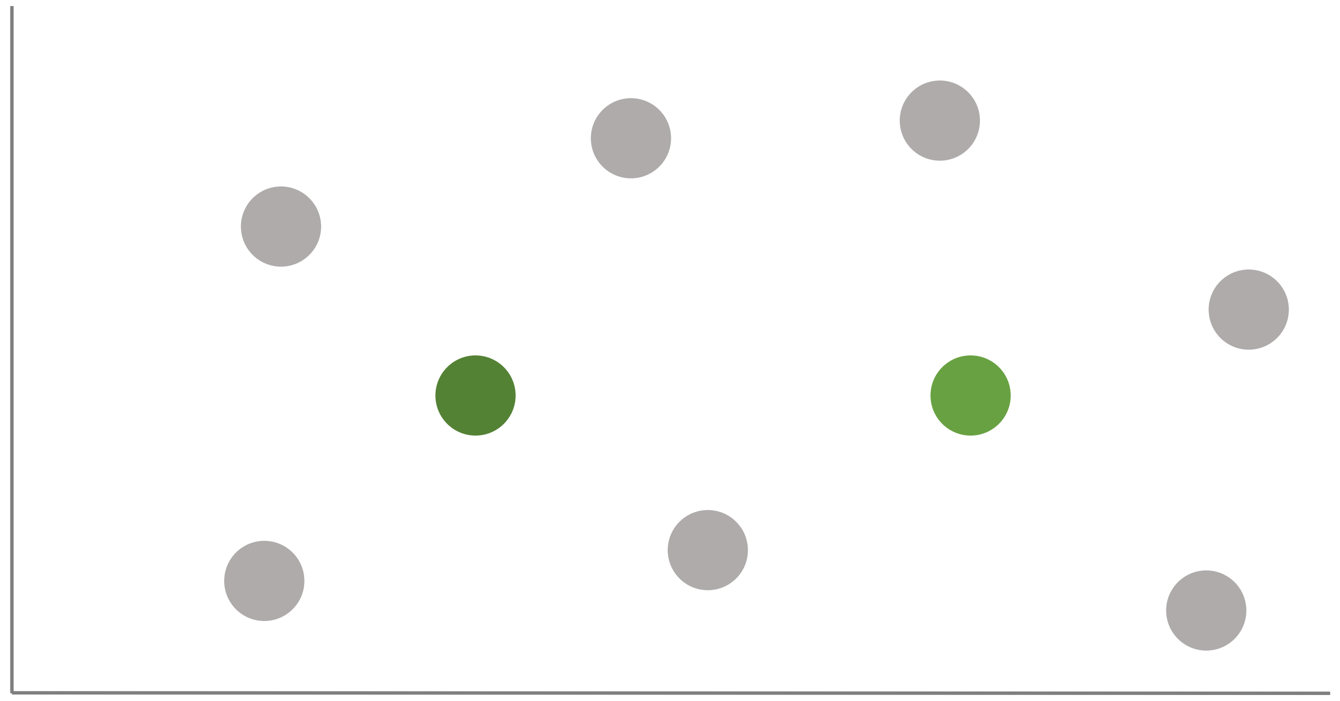
[D. Szafir, 2017]

Shape Problem with Geometric Assumption

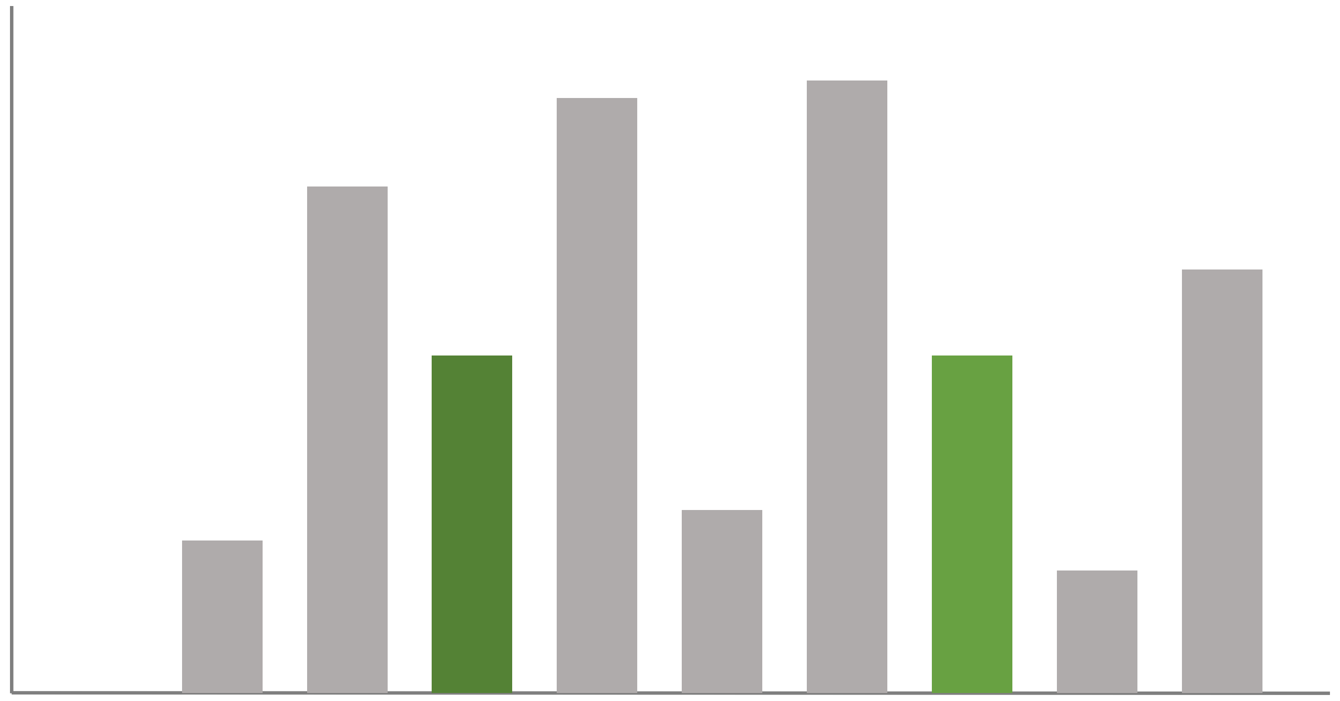


[D. Szafir, 2017]

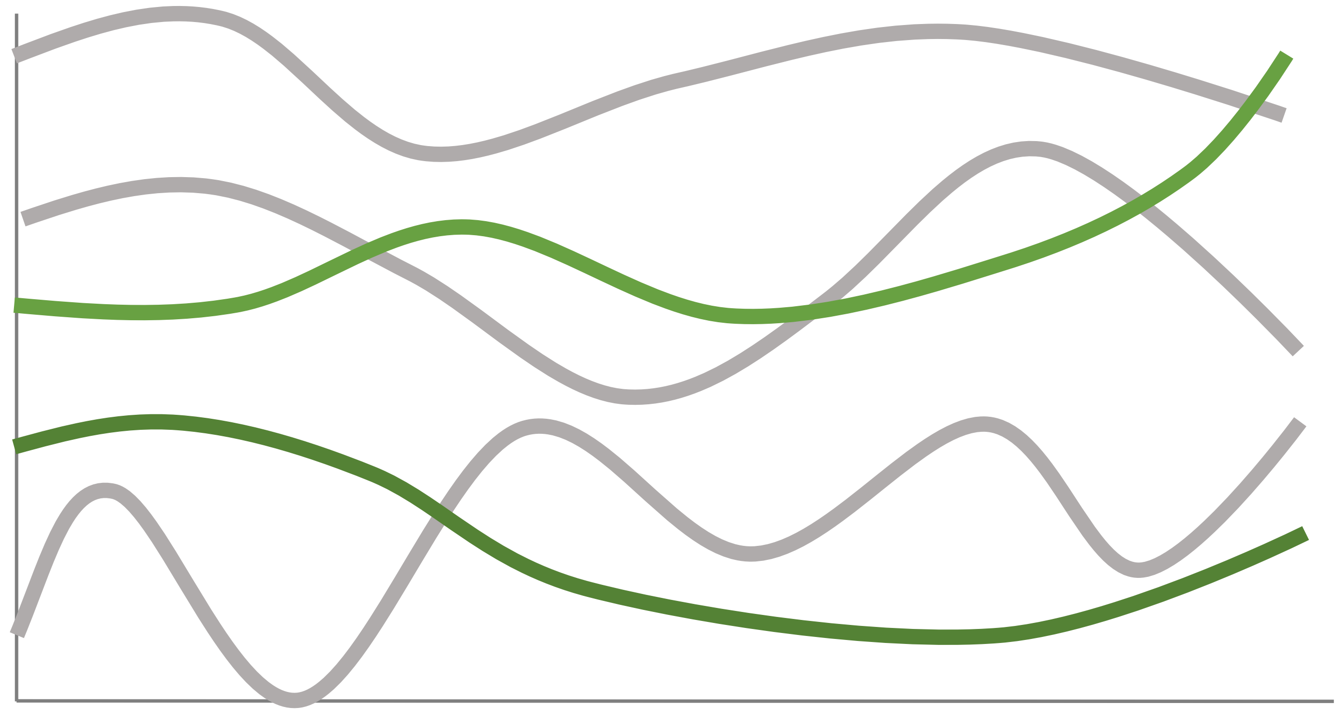
Types of Geometry



Diagonally Symmetric Marks



Elongated Marks

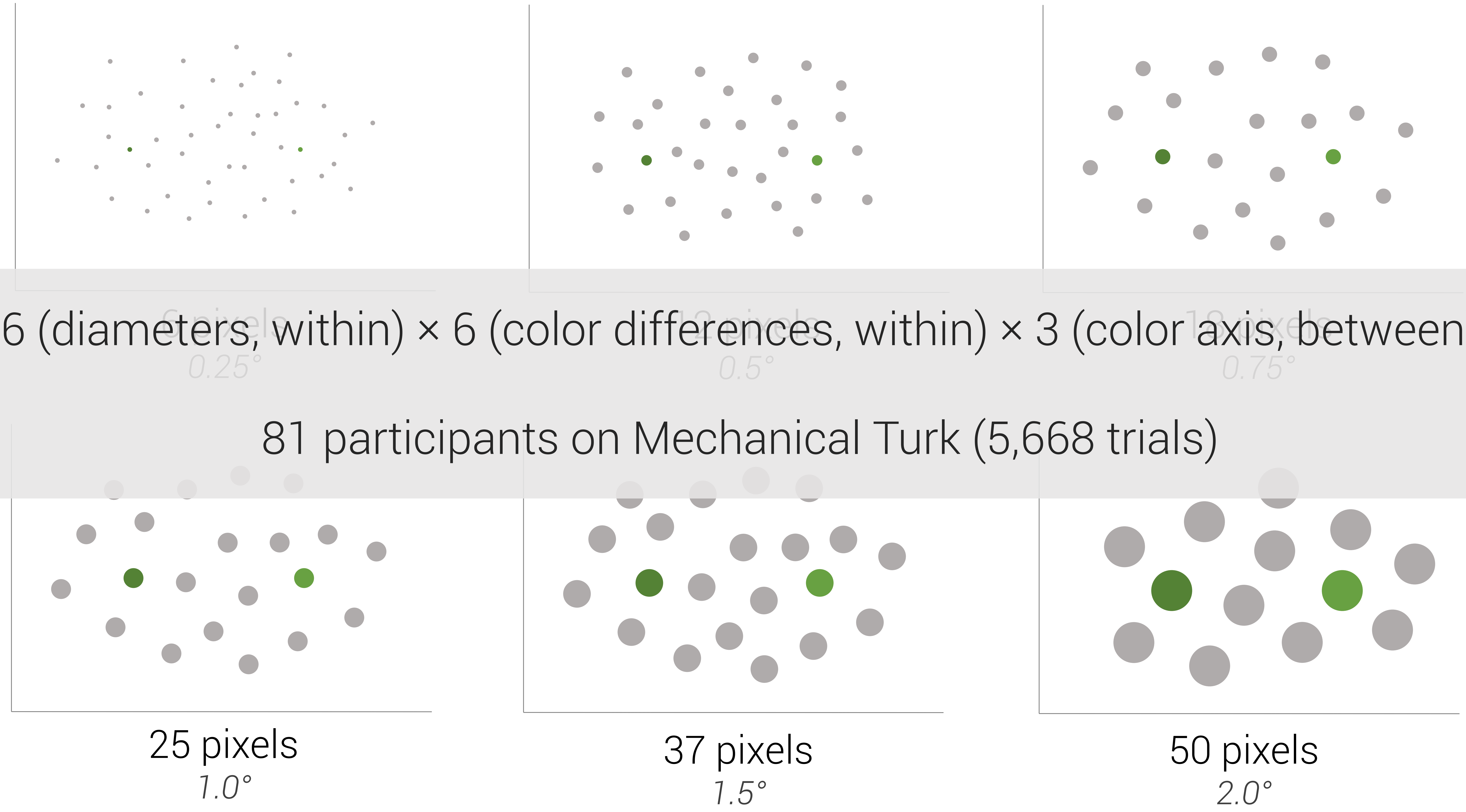


Asymmetric Marks

Area Marks

[D. Szafir, 2017]

Run the tests!

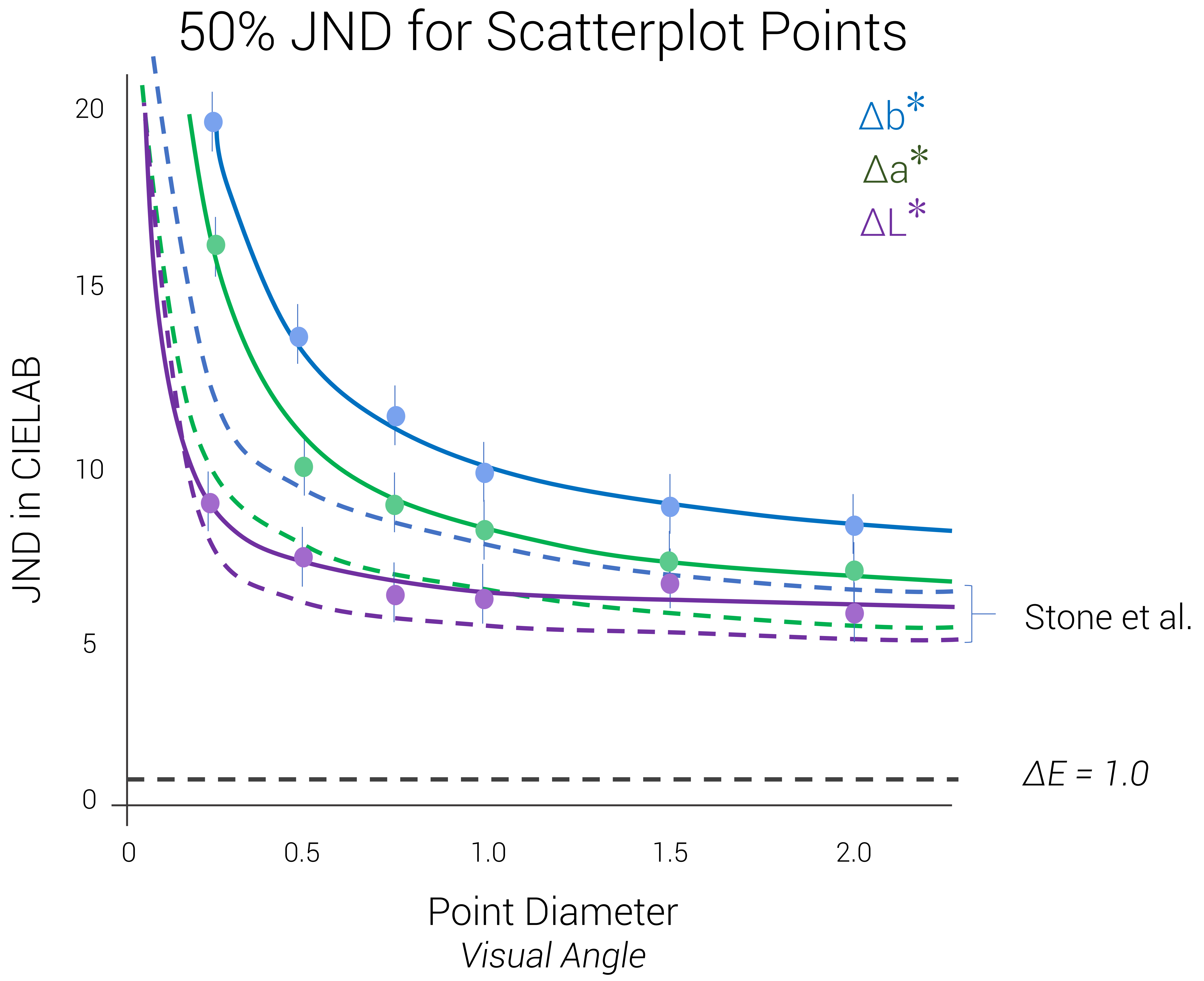


6 (diameters, within) × 6 (color differences, within) × 3 (color axis, between)

81 participants on Mechanical Turk (5,668 trials)

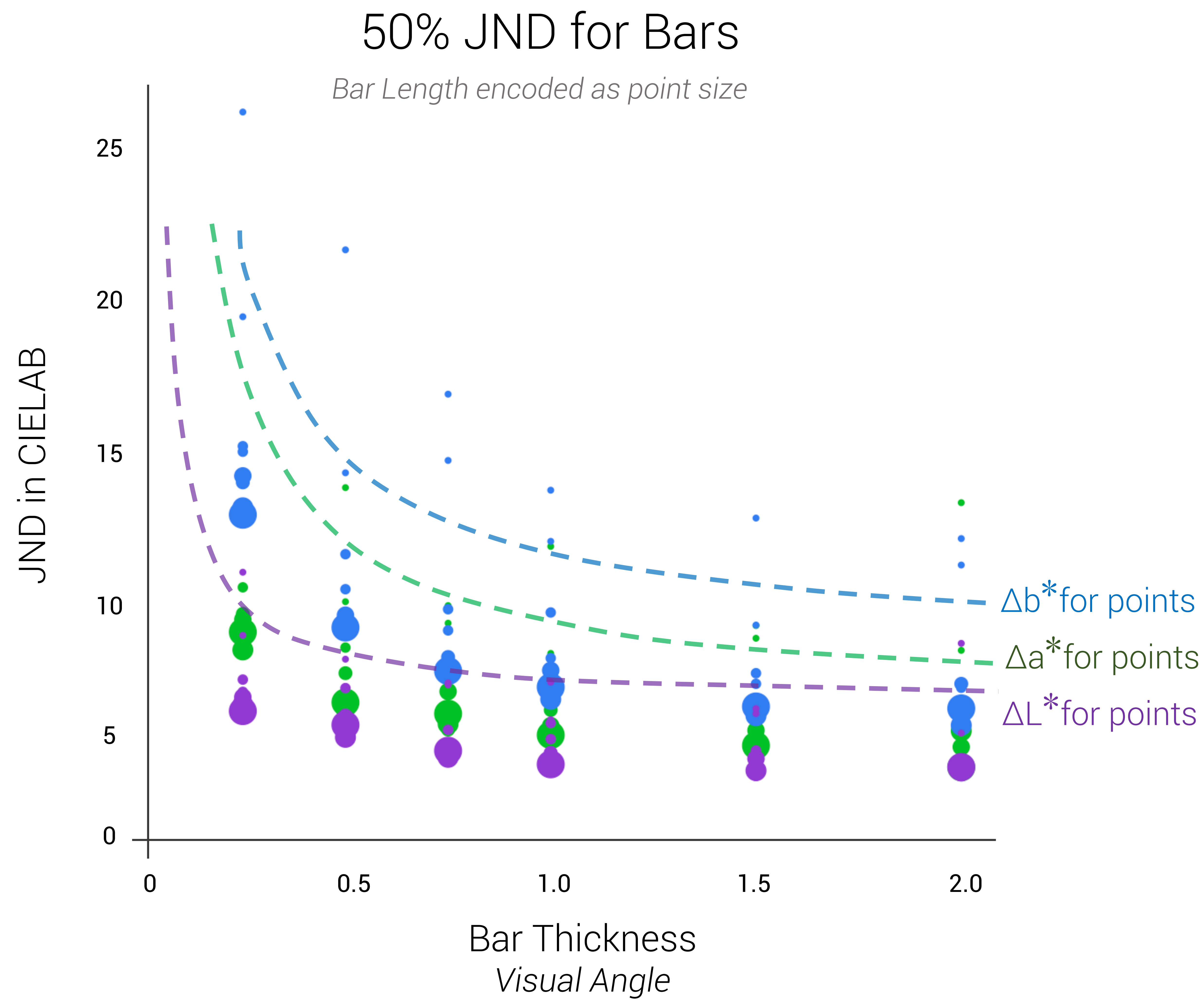
[D. Szafir, 2017]

Point Size: consistent with previous results



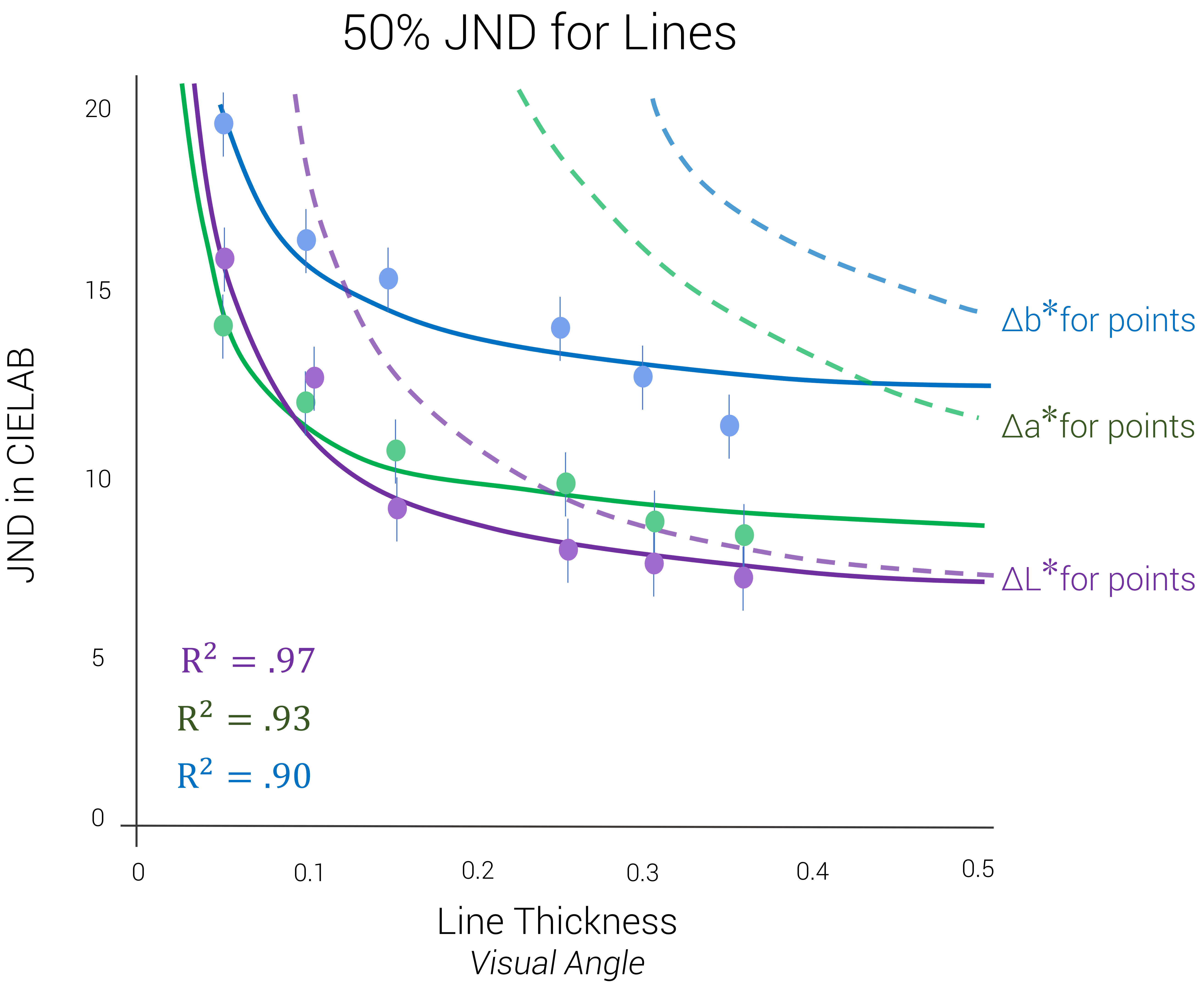
[D. Szafir, 2017]

Bar Thickness and Length: longer bars help



[D. Szafir, 2017]

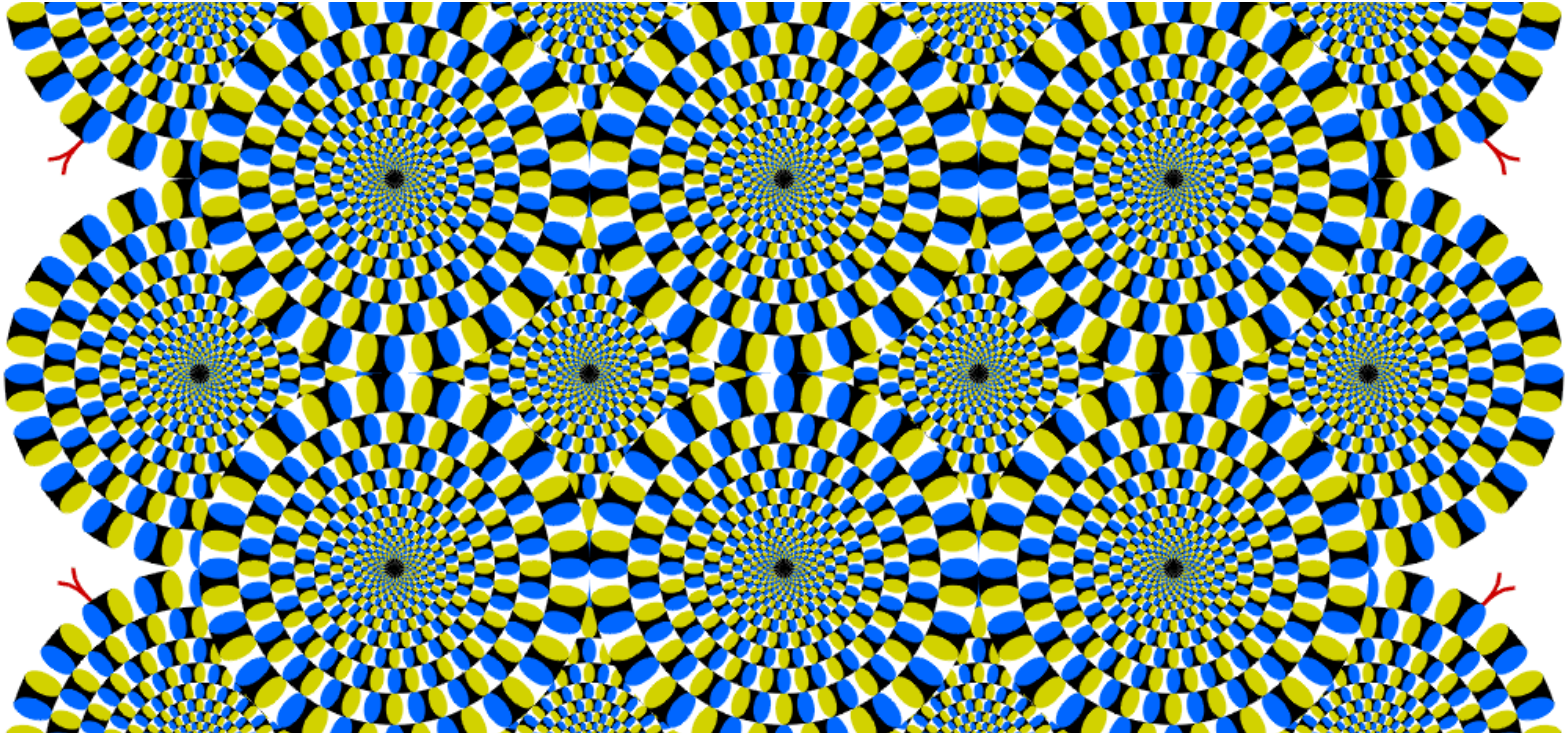
Line Thickness: better than points



[D. Szafir, 2017]

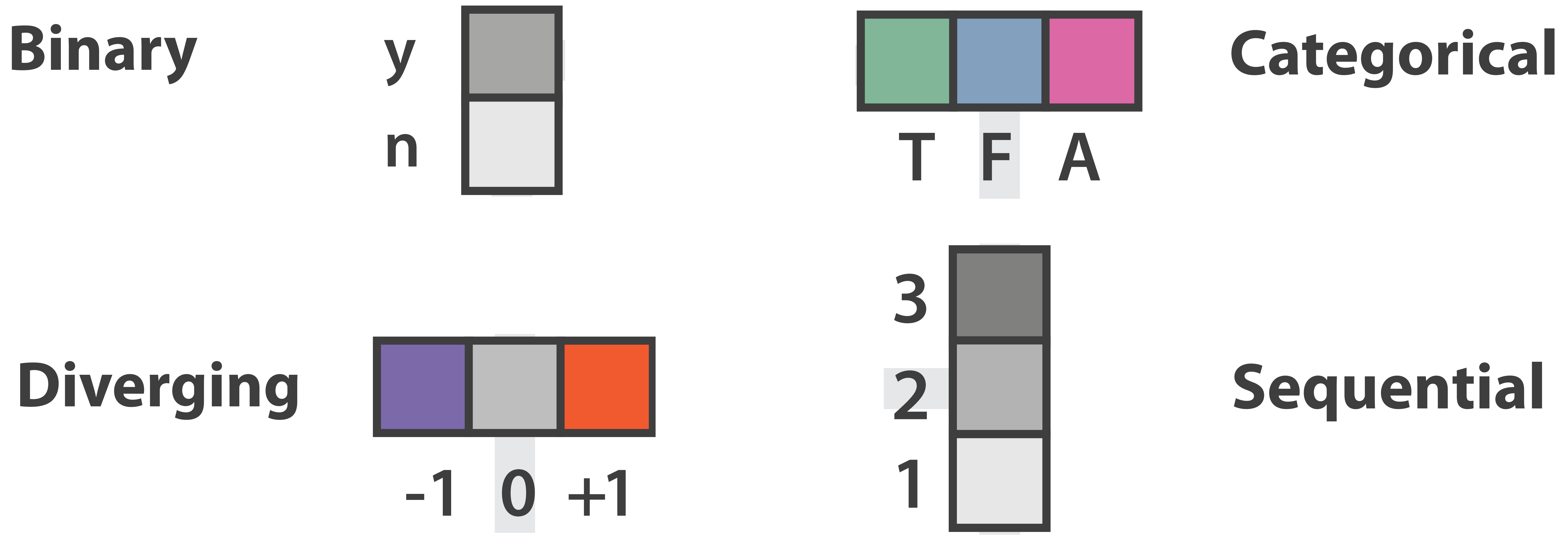
Color perception in real-world visualizations is complicated

Akiyoshi Kitaoka's Illusion pages



Colormap

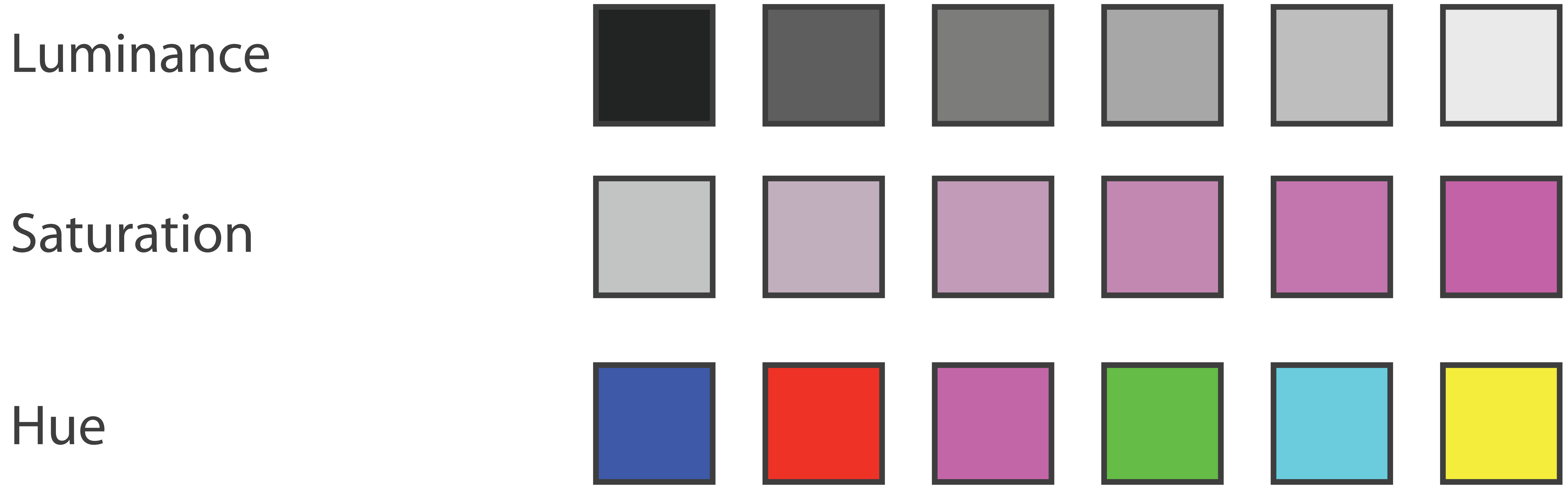
- A colormap specifies a mapping between colors and data values
- Colormap should follow the expressiveness principle
- Types of colormaps:



[Munzner (ill. Maguire), 2014]

Categorical vs. Ordered

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

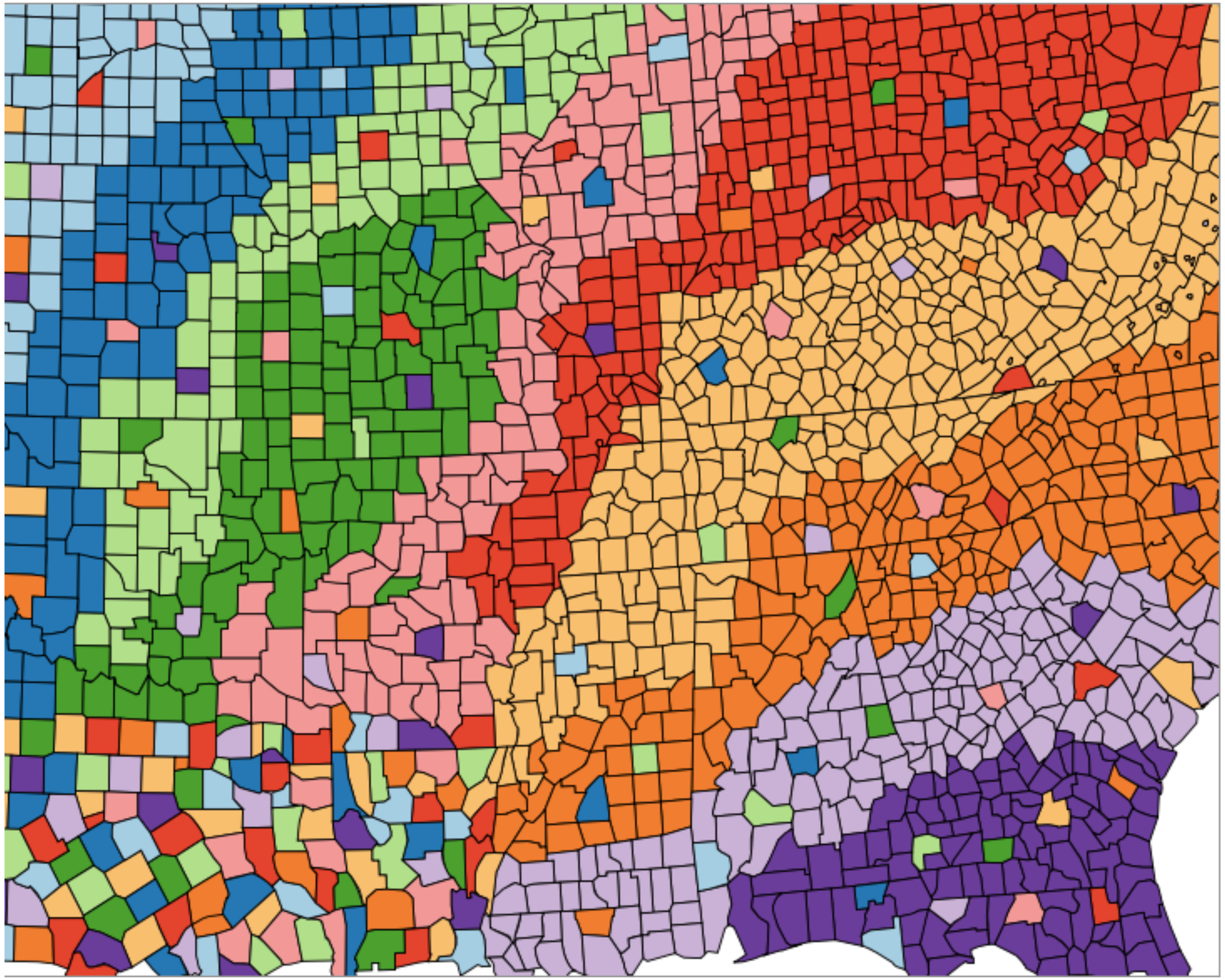


[Munzner (ill. Maguire), 2014]

Categorical Colormap Guidelines

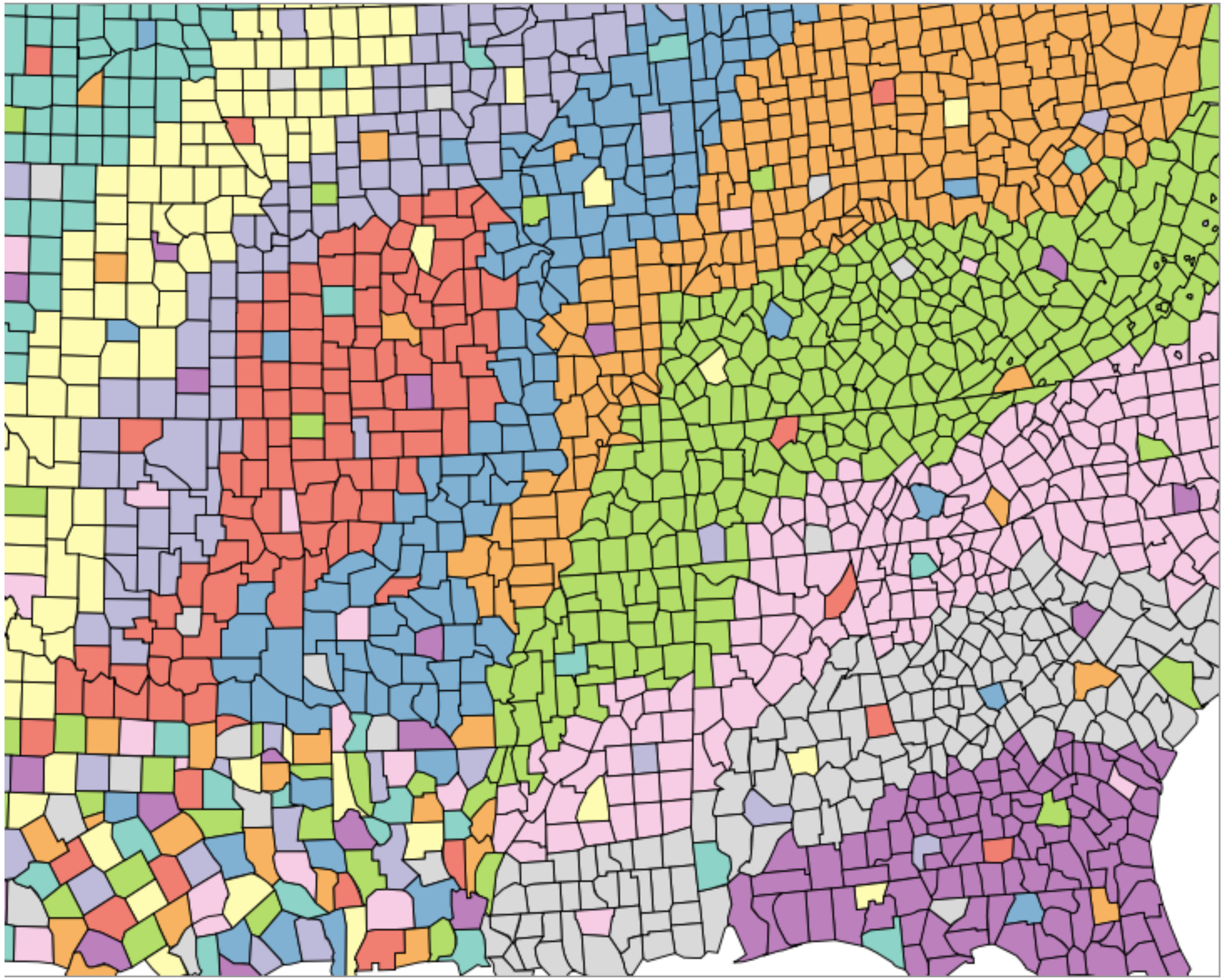
- Don't use too many colors (~12)
- Remember your background has a color, too
- Nameable colors help
- Be aware of luminance (e.g. difference between blue and yellow)
- Think about other marks you might wish to use in the visualization

Categorical Colormaps



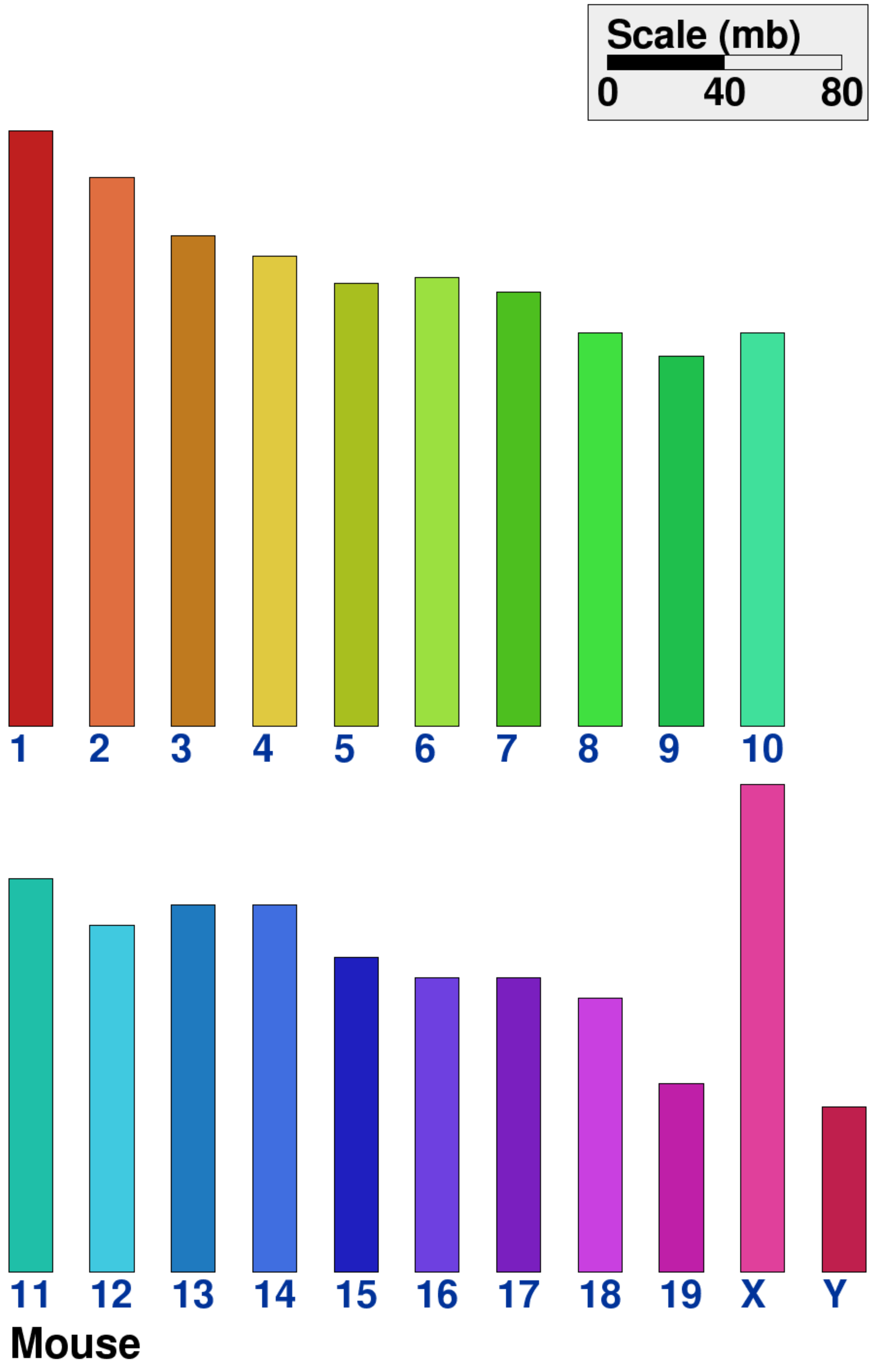
[colorbrewer2.org]

Categorical Colormaps



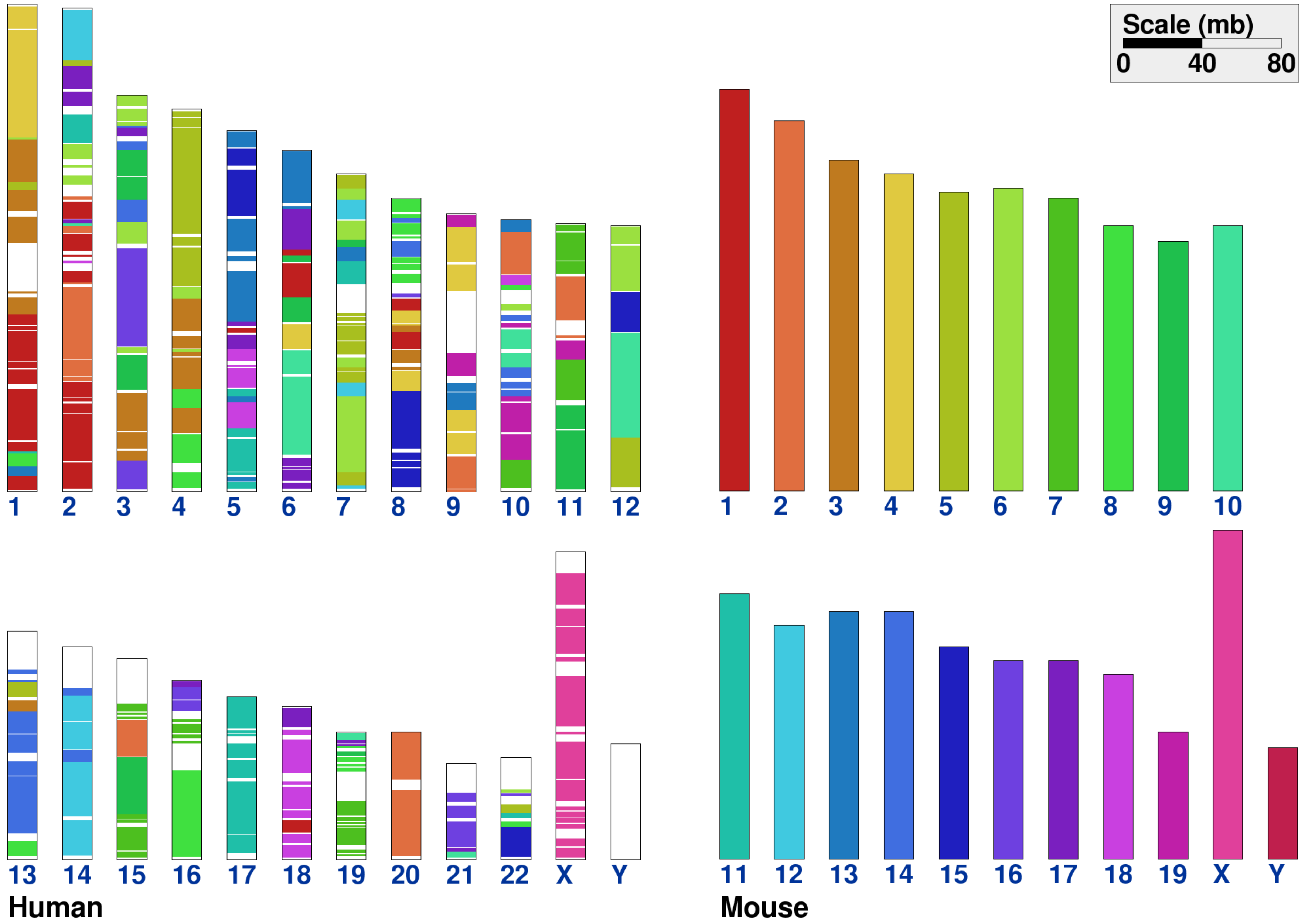
[colorbrewer2.org]

Number of distinguishable colors?



[Sinha & Meller, 2007]

Number of distinguishable colors?



[Sinha & Meller, 2007]

Discriminability

- Often, fewer colors are better
- Don't let viewers combine colors because they can't tell the difference
- Make the combinations yourself
- Also, can use the "Other" category to reduce the number of colors

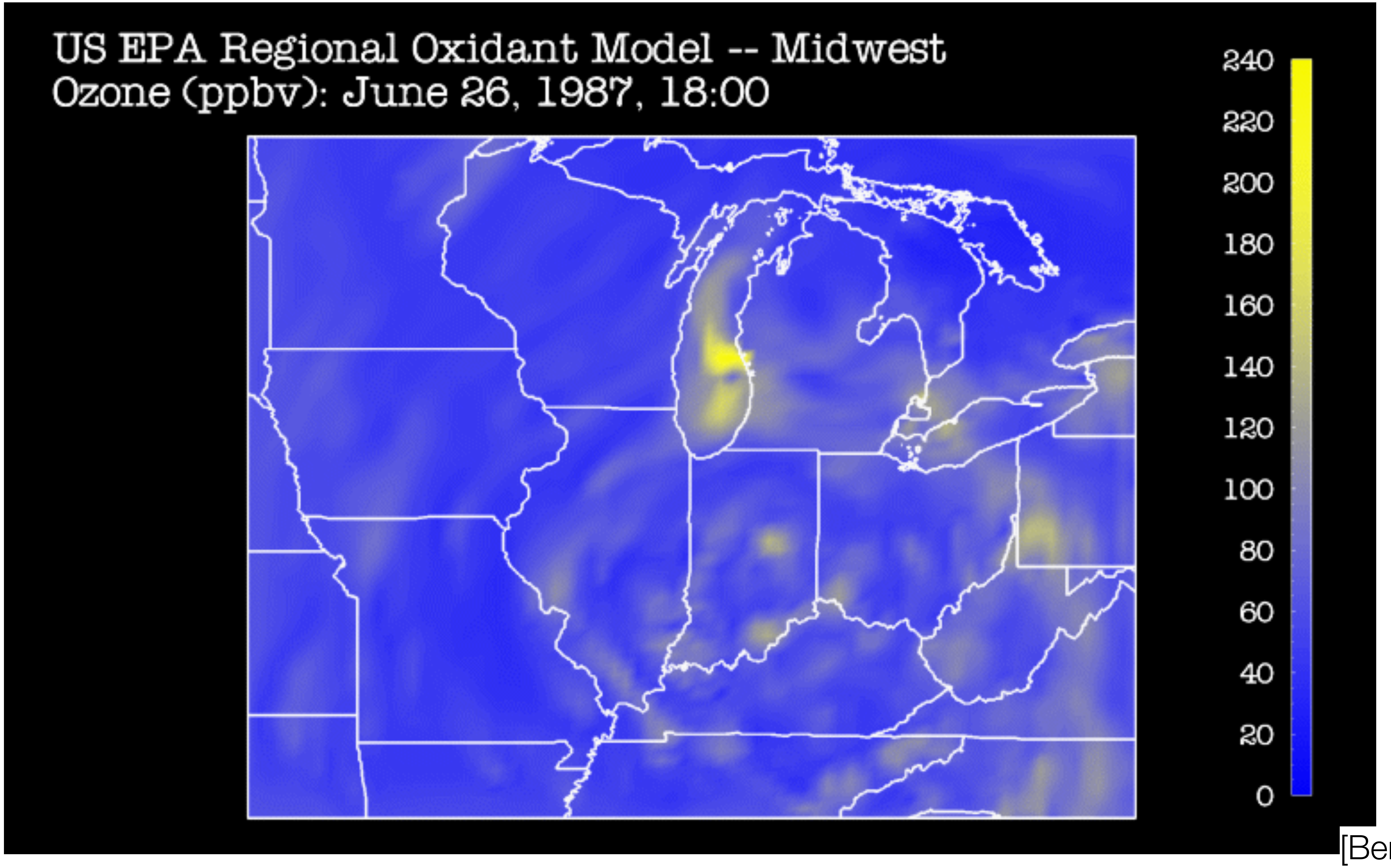
Discriminability

- Often, fewer colors are better
- Don't let viewers combine colors because they can't tell the difference
- Make the combinations yourself
- Also, can use the "Other" category to reduce the number of colors

Ordered Colormaps

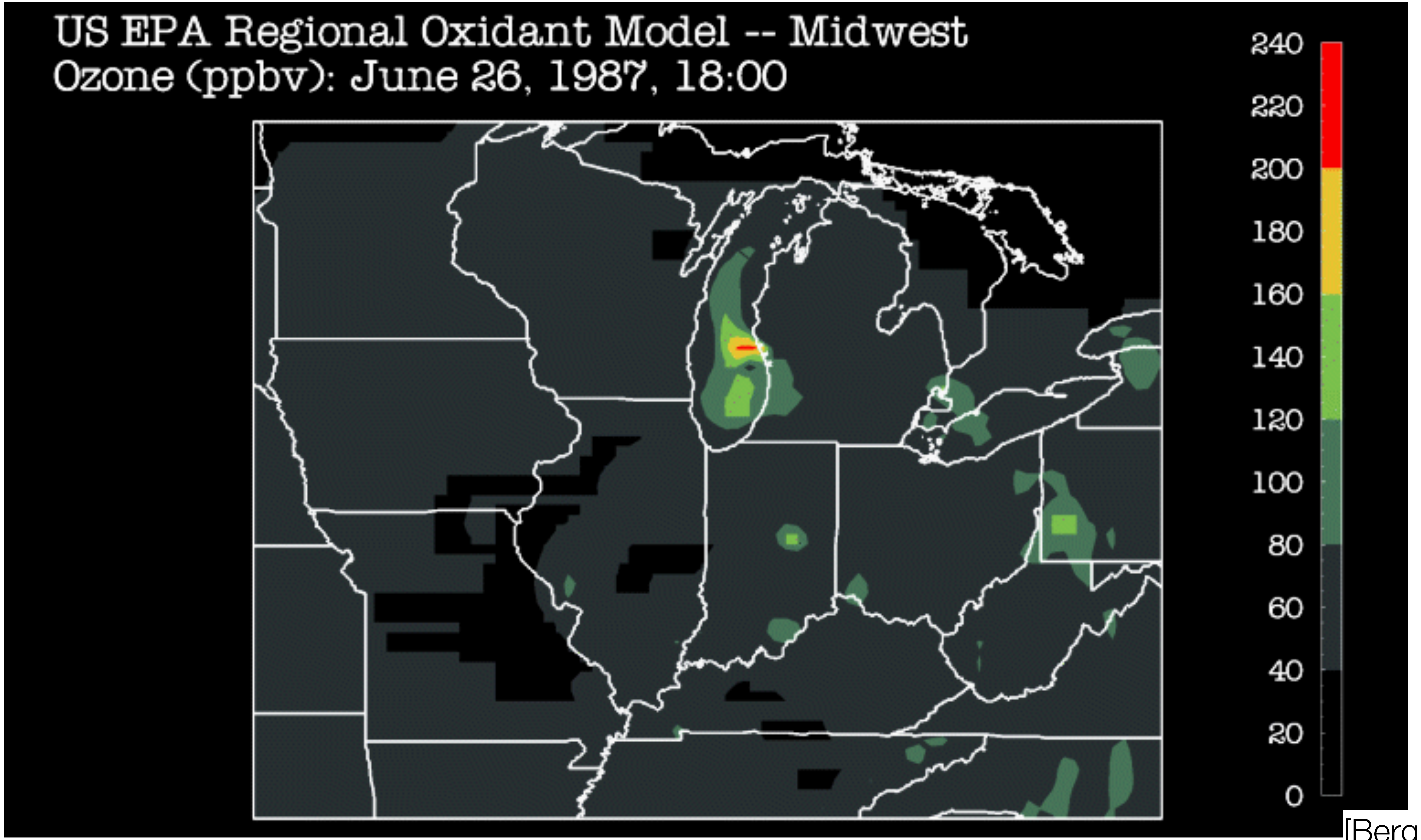
- Used for ordinal or quantitative attributes
- [0, N]: Sequential
- [-N, 0, N]: Diverging (has some meaningful midpoint)
- Can use hue, saturation, and luminance
- Remember hue is not a magnitude channel so be careful
- Can be **continuous** (smooth) or **segmented** (sharp boundaries)
  - Segmented matches with ordinal attributes
  - Can be used with quantitative data, too.

Continuous Colormap



[Bergman et al., 1995]

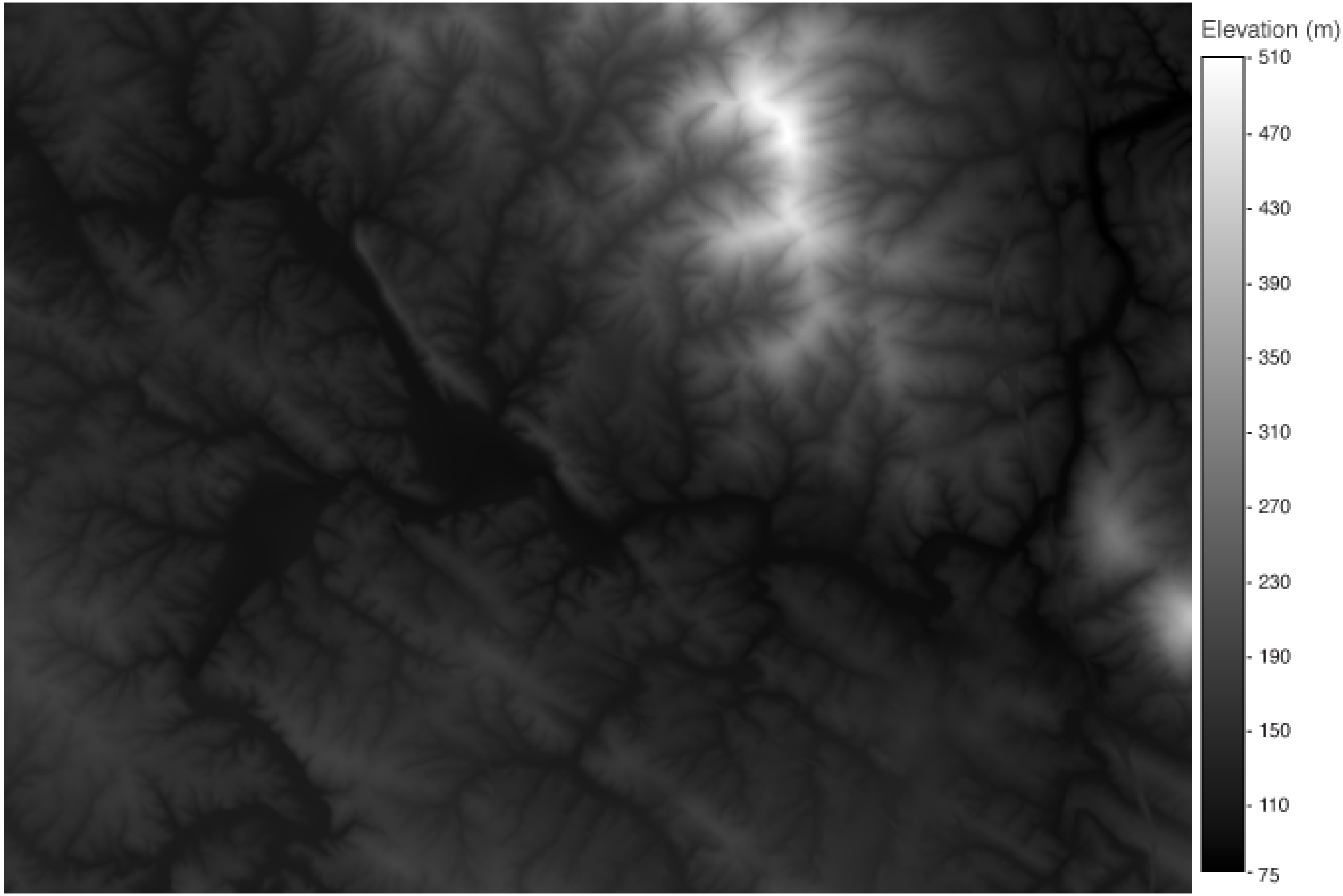
Segmented Colormap



[Bergman et al., 1995]

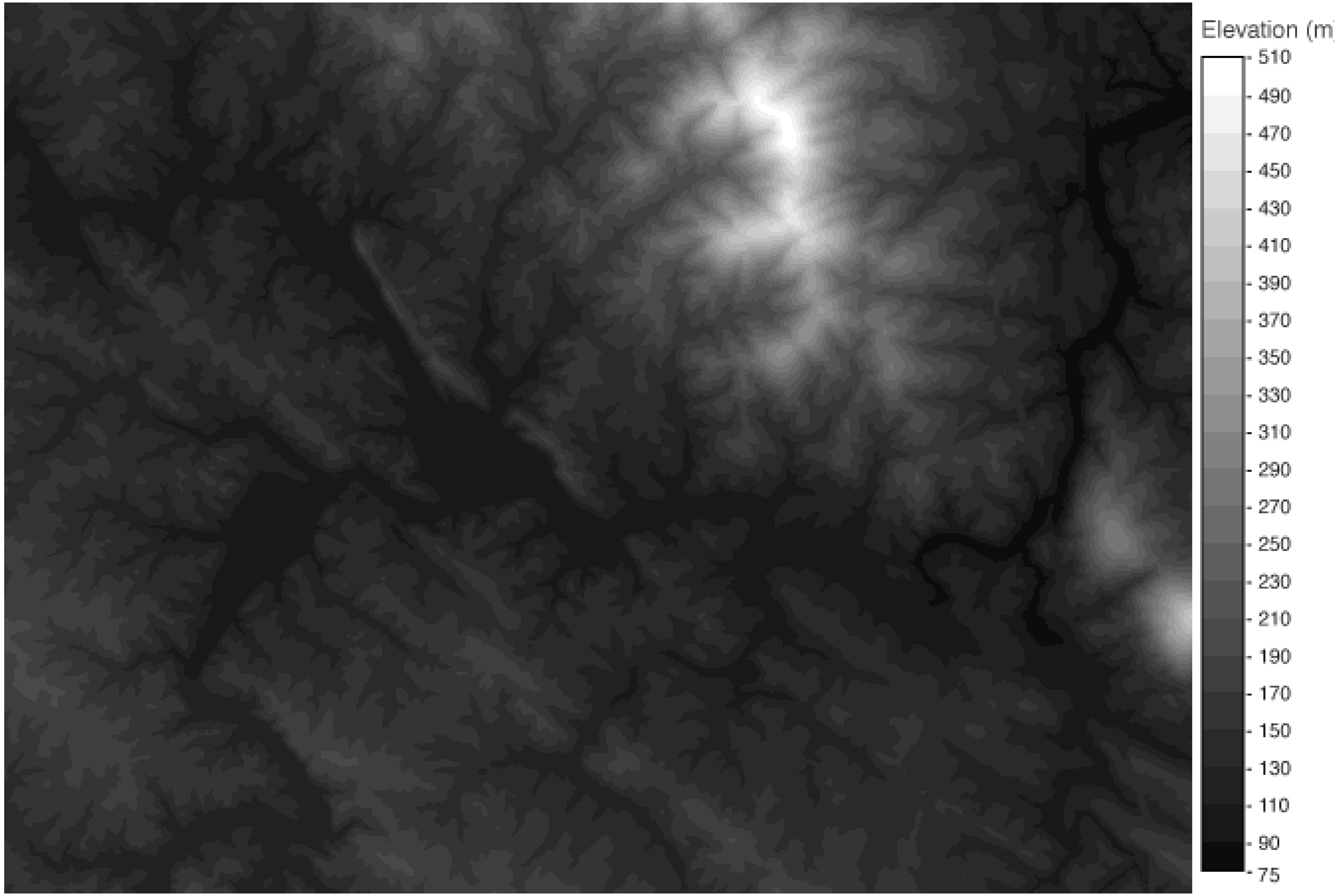
Is continuous better than segmented?

Continuous



[Padilla et al., 2017]

Many Segments



[Padilla et al., 2017]

Fewer Segments

[Padilla et al., 2017]

Types of Tasks

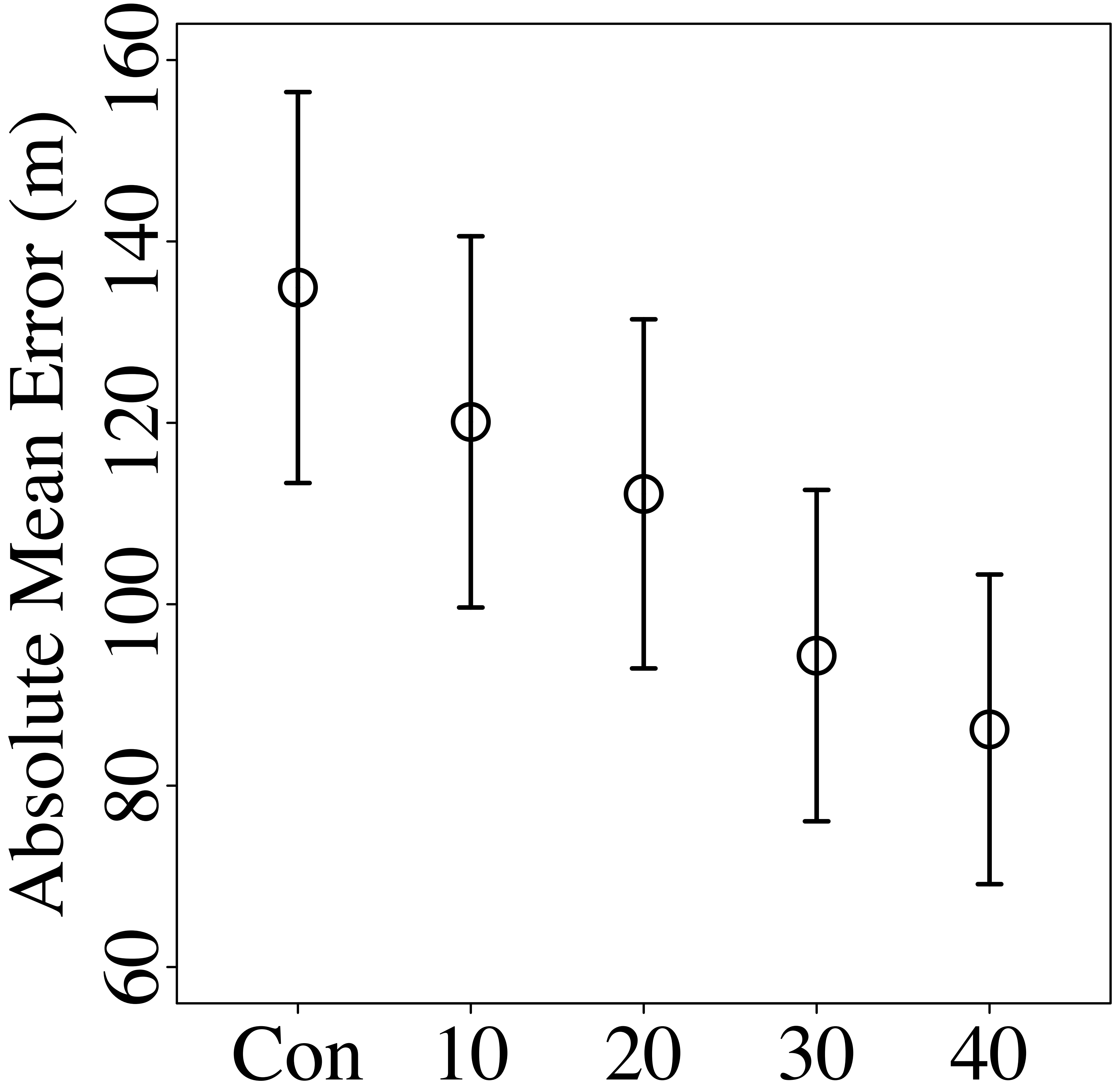
- Locate/Explore & Identify: Highest Point (Global, In Region), 275m
- Locate/Explore & Compare: Height Compare/Rank
- Explore & Identify: Steepest
- Lookup & Identify: Lookup
- Explore & Compare: Steepness Compare/Rank
- Browse & Summarize: Average Height
- Browse & Compare: Compare Average Height
- Combination: Steepest at 355m

[Padilla et al., 2017]

Results

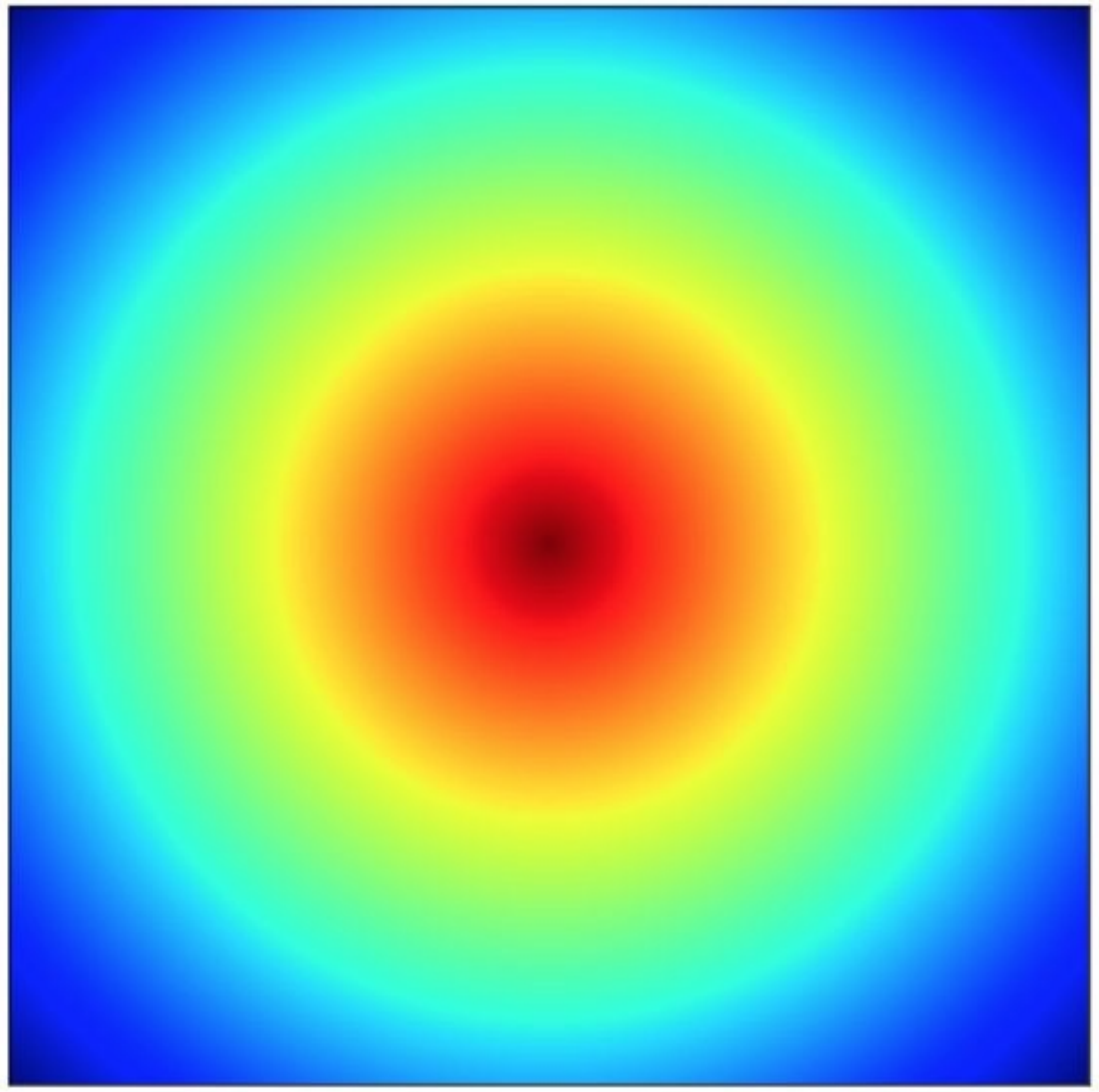
- "[C]ontrary to the expressiveness principle, no cases were found in which a continuous encoding of 2D scalar field data was advantageous for task accuracy, and for some tasks, specific binned encodings facilitated accuracy."
- "[S]upport for the counterintuitive finding that decisions with binned encoding were slower than those made with continuous encoding"
- Word of caution: single image!

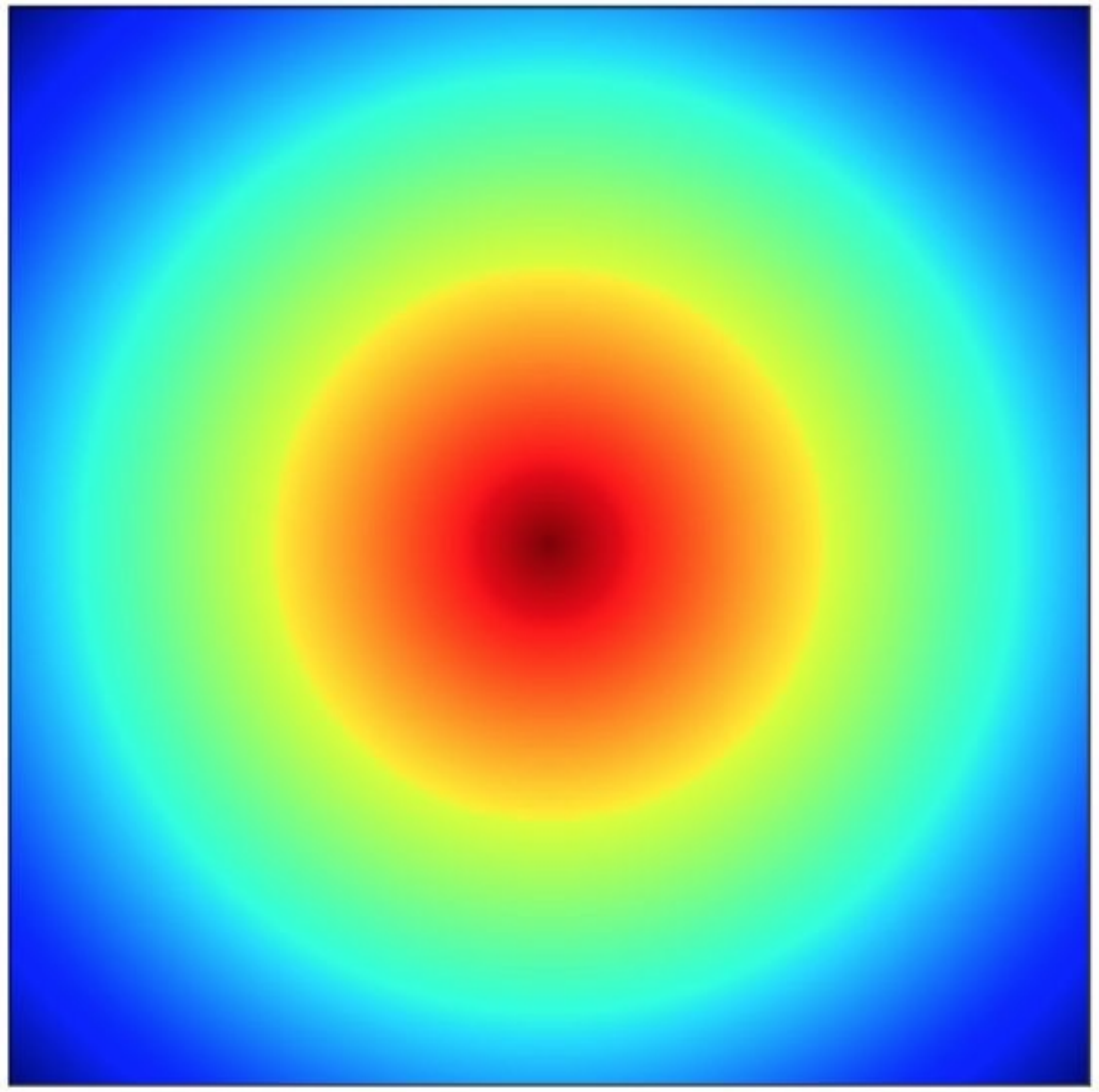
Lookup Task (Lower)



[Padilla et al., 2017]

Don't Use Rainbow Colormaps

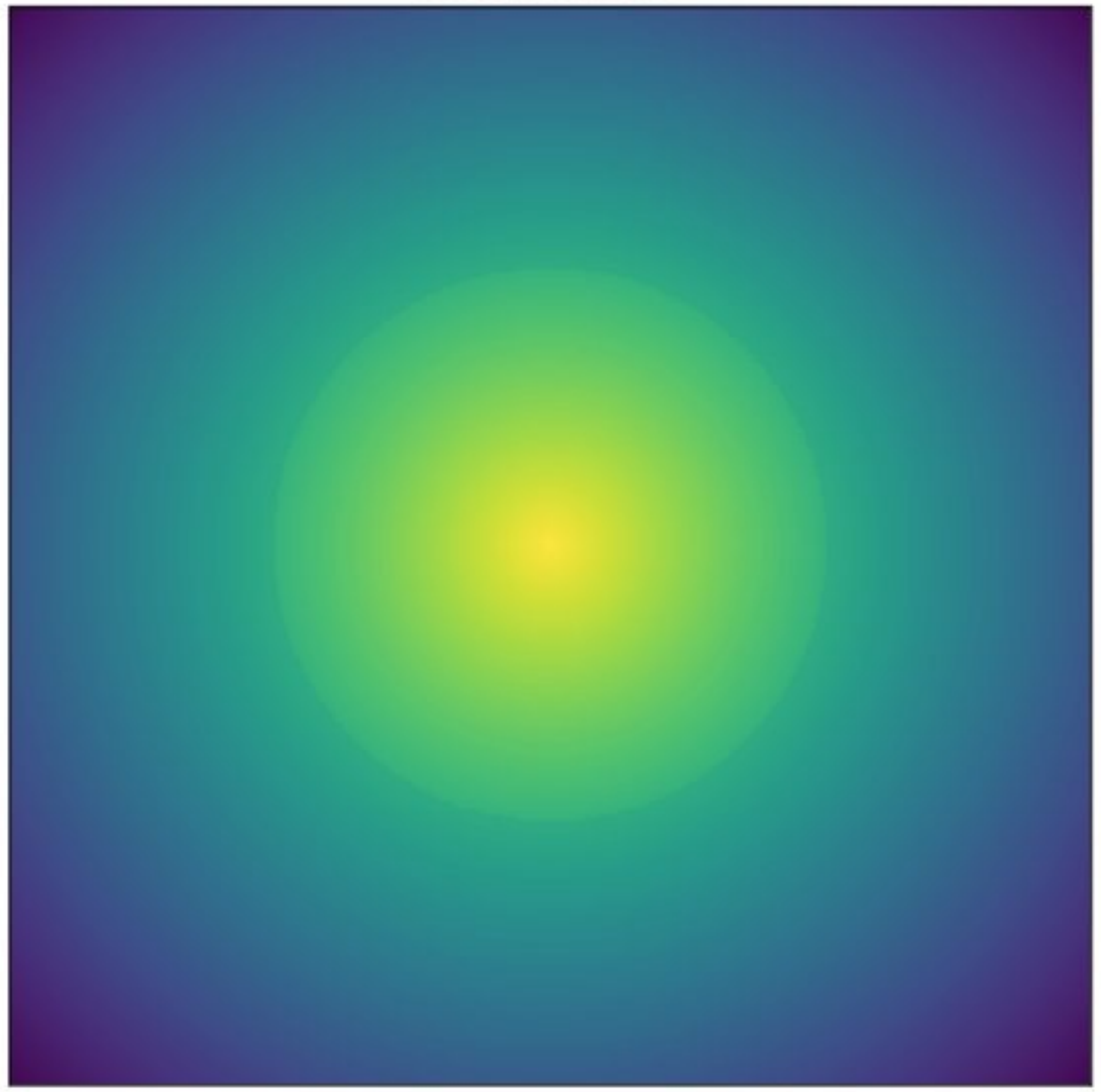


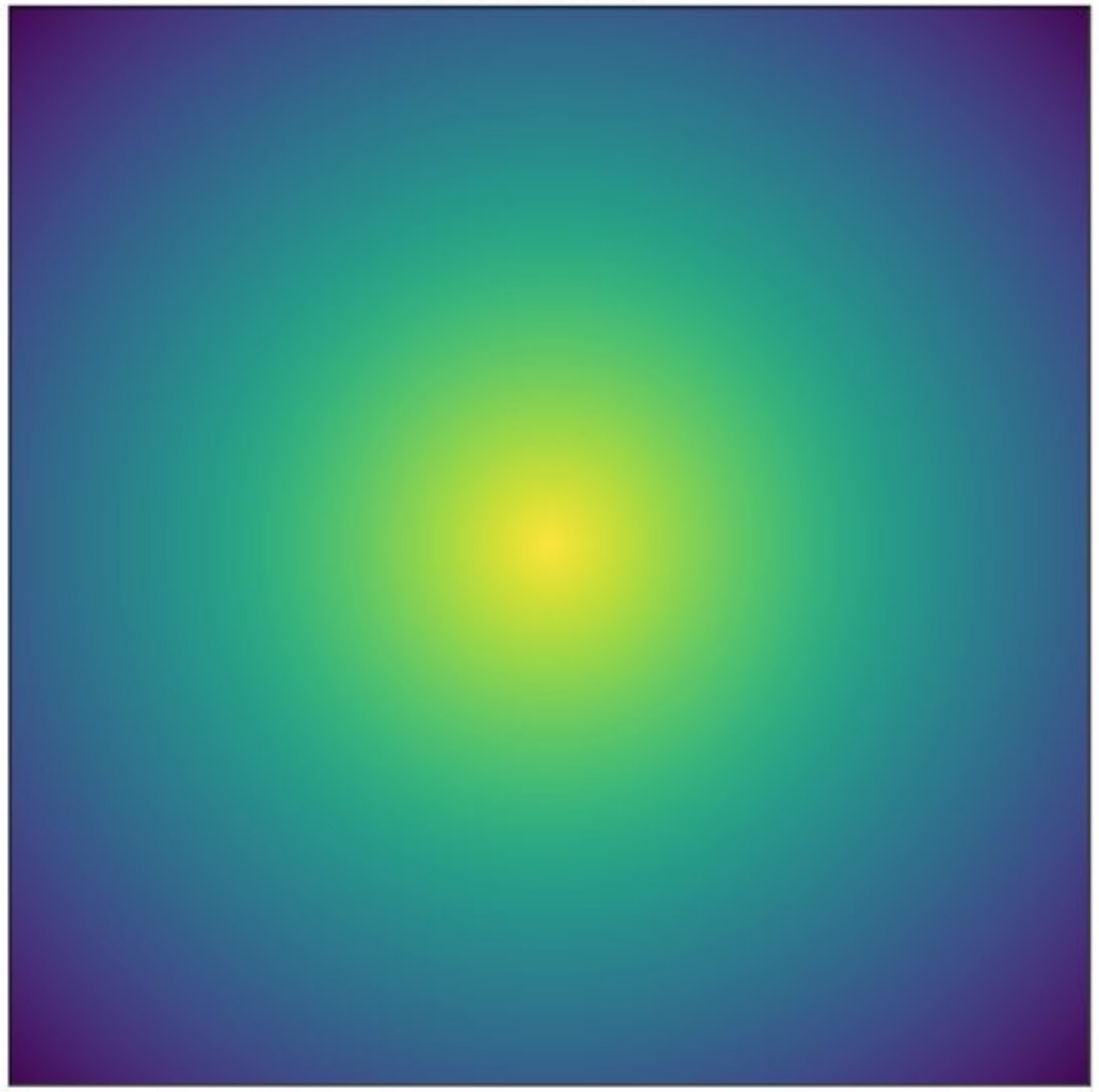


Which has a discontinuity?

[M. Bussonnier]

Other Colormaps Work Better





Which has a discontinuity?

[M. Bussonnier]