Data Visualization

INF552 (2023-2024)

Session 02
Fundamental Principles of Visual Perception for Visualization

emmanuel.pietriga@inria.fr

Individual Project - Milestone #1

- By session 03 (*i.e.*, Friday next week), identify a dataset (or collection thereof) that you are interested in visualizing (based on your own interests).

- Validate it with me (just a sanity check to confirm feasibility) by uploading a short text description in the relevant Assignment on Moodle (will post a link to it on Slack).

- You can use the catalogs and search engines referenced on slide 6 of the course slide deck (INF552-2023-course-s01) to search for data, or use any other means.

Datasets

➔ Data Types

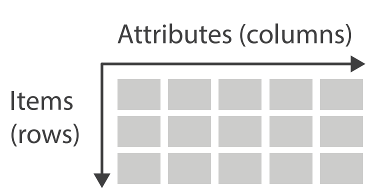
➔ Items ➔ Attributes ➔ Links ➔ Positions ➔ Grids

➔ Data and Dataset Types

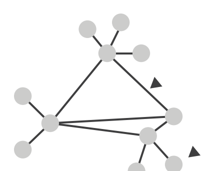
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

➔ Dataset Types

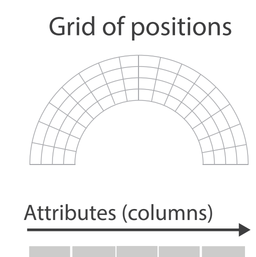
➔ Tables



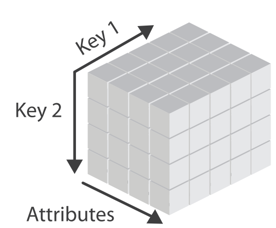
➔ Networks



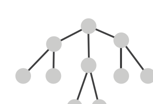
➔ Fields (Continuous)



➔ *Multidimensional Table*



➔ *Trees*



➔ Geometry (Spatial)

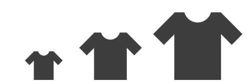
Attributes

➔ Attribute Types

➔ Categorical

➔ Ordered

➔ *Ordinal*



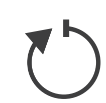
➔ *Quantitative*

➔ Ordering Direction

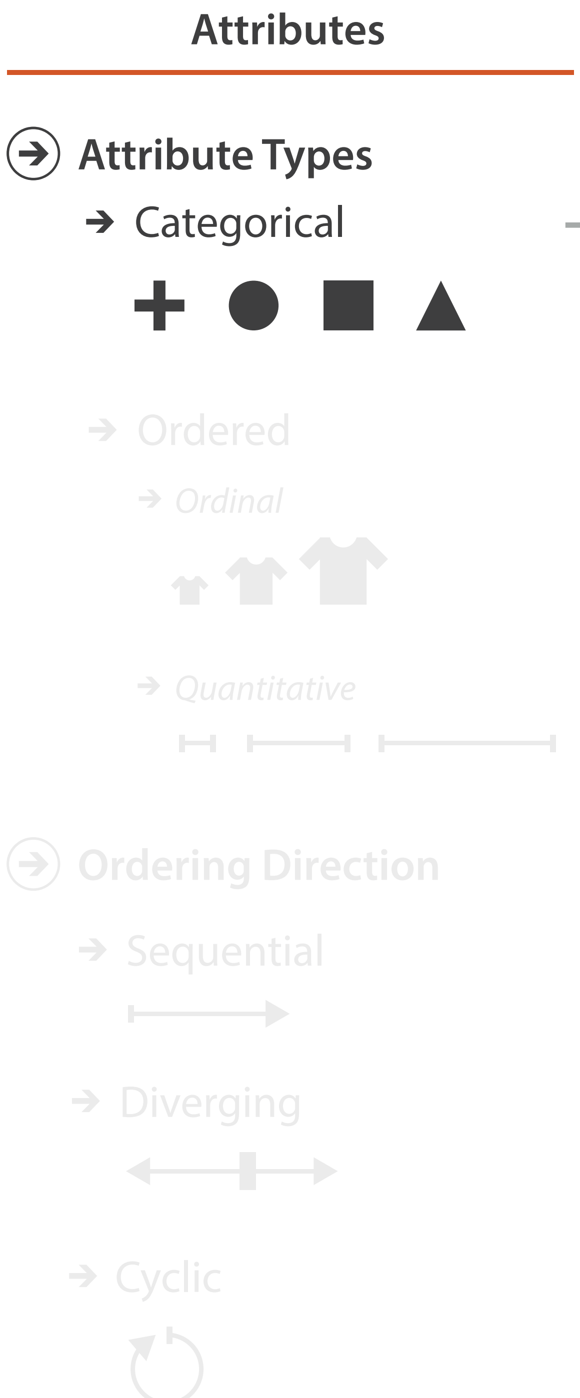
➔ Sequential

➔ Diverging

➔ Cyclic



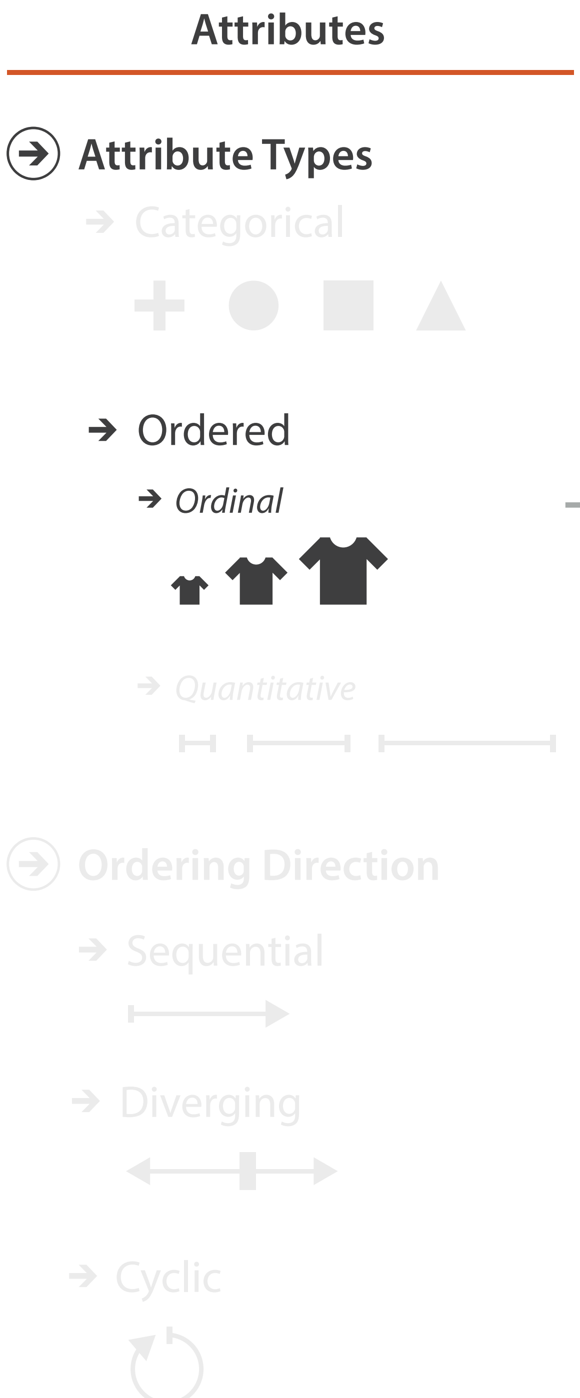
… we need to talk about data and dataset types.



- or nominal (labels)
- can be counted
- examples:
  - fruits (apple, oranges, …)
  - car makers (Audi, Renault, …)
  - group membership
- operations:
  - =, ≠

A categorical scale divides items into classes that have no particular order.

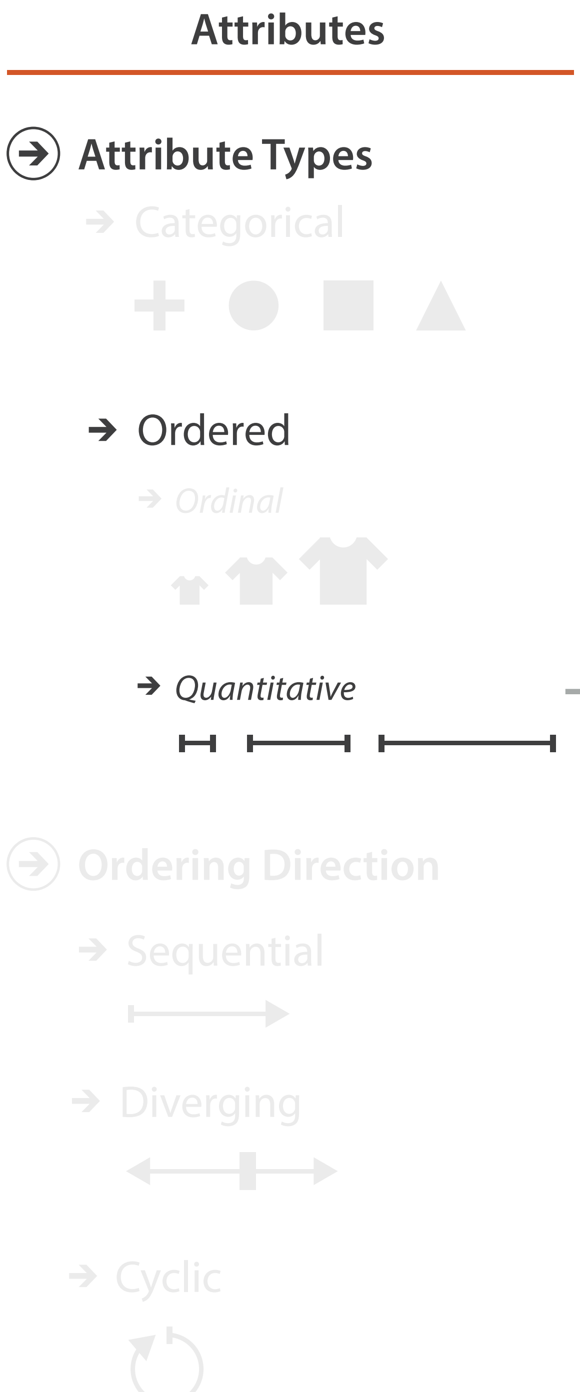
[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]



- can be counted, ordered
- examples:
  - 5-star rating scheme (★★★★☆)
  - shirt sizes (XS, S, M, L, XL, XXL)
  - Likert scale
  - Military ranks
- operations:

  =, ≠, <, >

An ordinal scale says nothing about the magnitude of differences between classes.



- can be counted, ordered, measured

Ratio

meaningful zero

- examples:
  - physical measurements: length, mass, luminance, temperature (°K)…
  - financial indicators
- operations: =, ≠, <, >, - , +, ×, ÷

Items on an interval scale have an order and a relative distance between them.

Interval

no clear zero, or an arbitrary one

- examples:
  - dates
  - lat/lon
  - temperature (°C or °F)
- operations: =, ≠, <, >, - , +

Items on a ratio scale have an order and a distance to a non-arbitrary absolute zero.

Quantitative scales can be linear or logarithmic.

Attributes

Attribute Types

- Categorical
- Ordered
  - *Ordinal*
  - *Quantitative*

Ordering Direction

- Sequential — altitudes, speeds, …
- Diverging — temperature, elevation (terrain+bathymetry), profit & loss…
- Cyclic — time measurements: hour of the day, month of the year, …

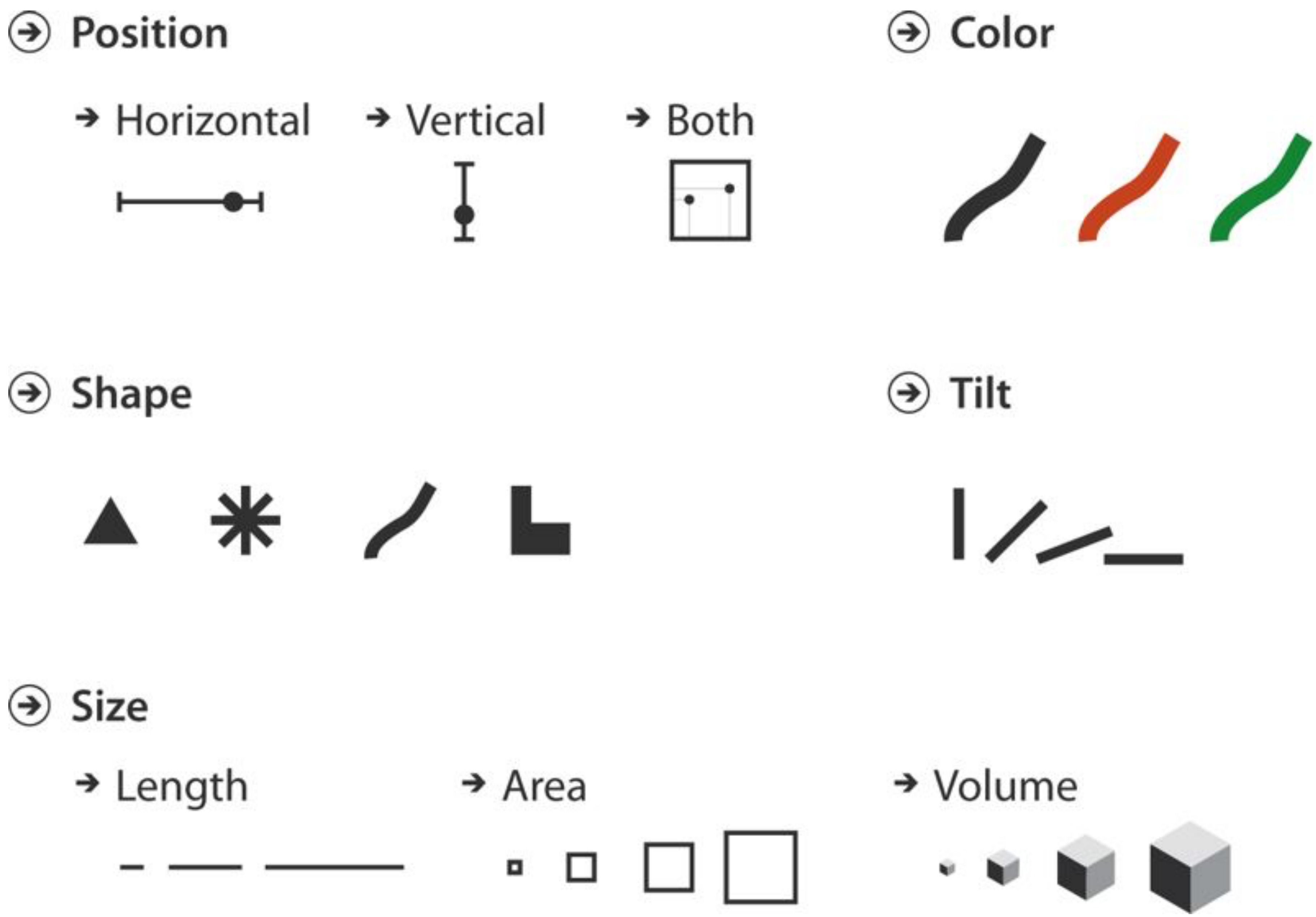
Why is this important?

- Because nominal, ordinal, and quantitative data are best expressed in different ways visually.
- Because the ordering direction will also have an influence on the choice of visual variable and its range.
- Because different data types are often associated with different tasks:
  - temporal data (comparison of events)
  - trees (understand parent-child relationships)
  - tabular data: identify correlations, understand distribution, …
  - …
- and the design space, again, is *enormous*.

[Source: adapted from InfoVis introductory course, Petra Isenberg, 2013]

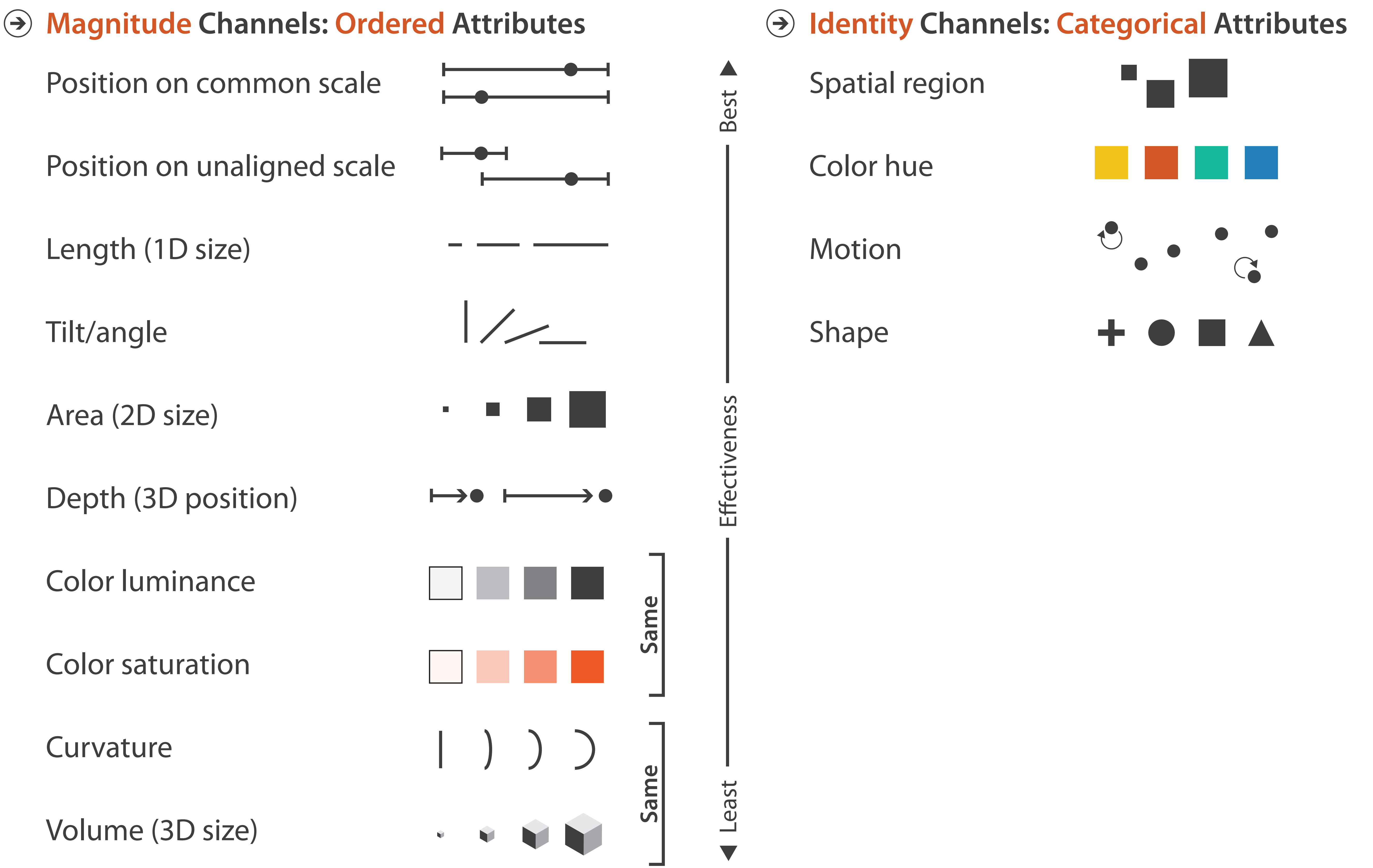
Marks are geometric primitives

[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]



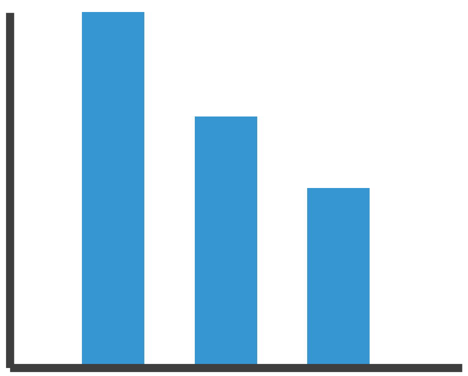
Visual (encoding) channels control the appearance of marks

Channels: Expressiveness Types and Effectiveness Ranks



[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]

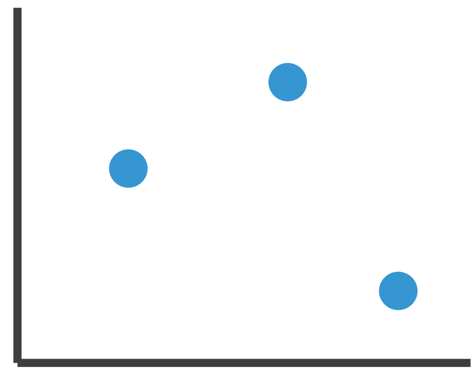
Using marks and channels



Mark line

Vertical spatial position quantitative

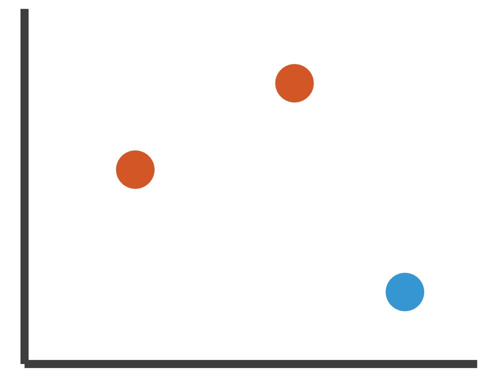
Horizontal spatial position categorical



Mark point

Vertical spatial position quantitative

Horizontal spatial position quantitative

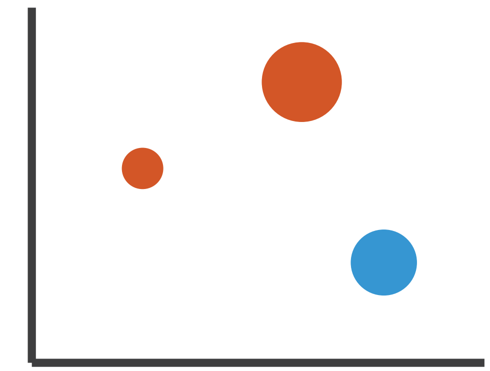


Mark point

Vertical spatial position quantitative

Horizontal spatial position quantitative

Color hue categorical



Mark point

Vertical spatial position quantitative

Horizontal spatial position quantitative

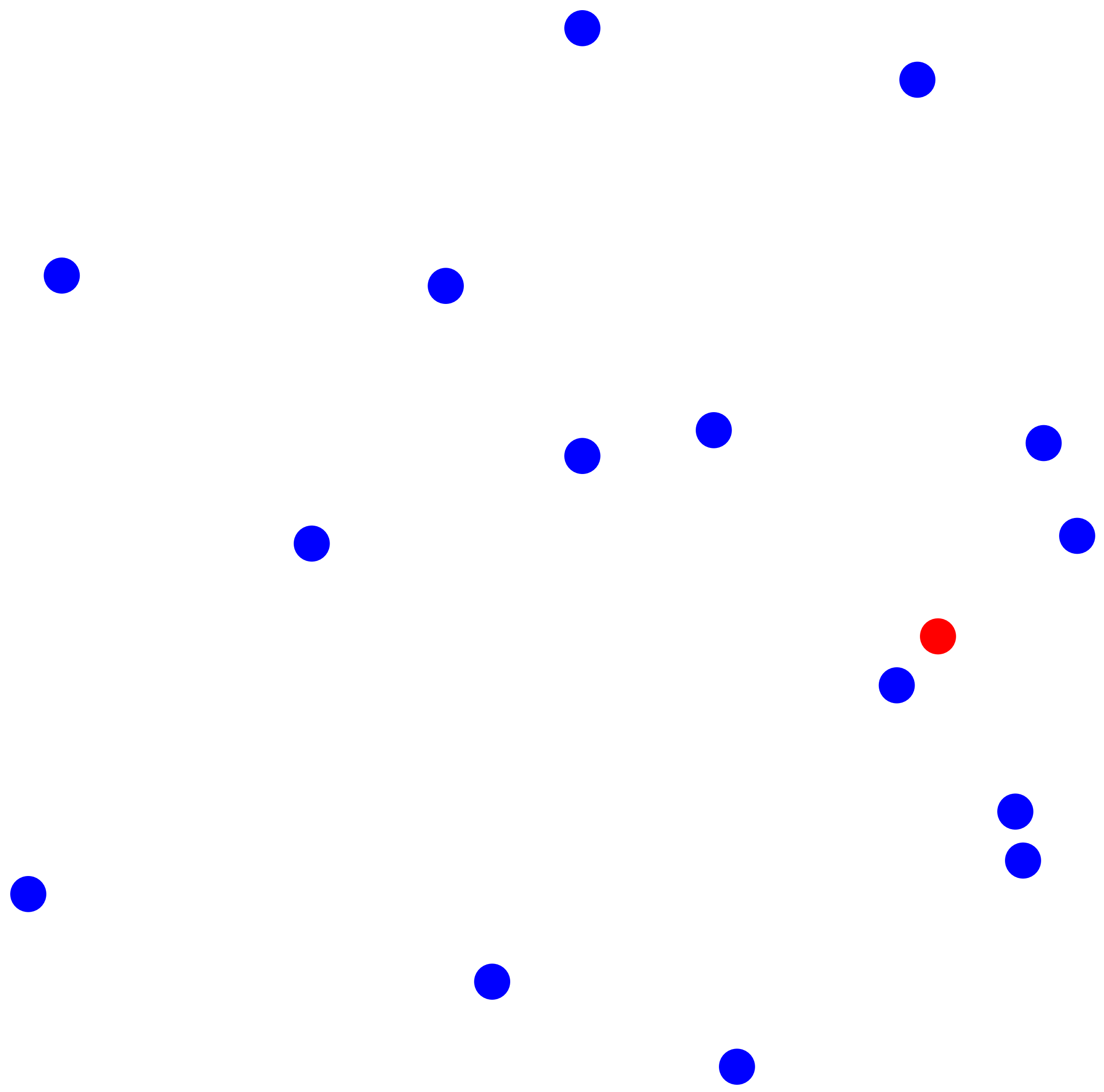
Color hue categorical

Size quantitative

Assign variables of lesser importance to weaker visual channels.
(Variables that help interpret the data by giving context rather than enable making precise judgment.)

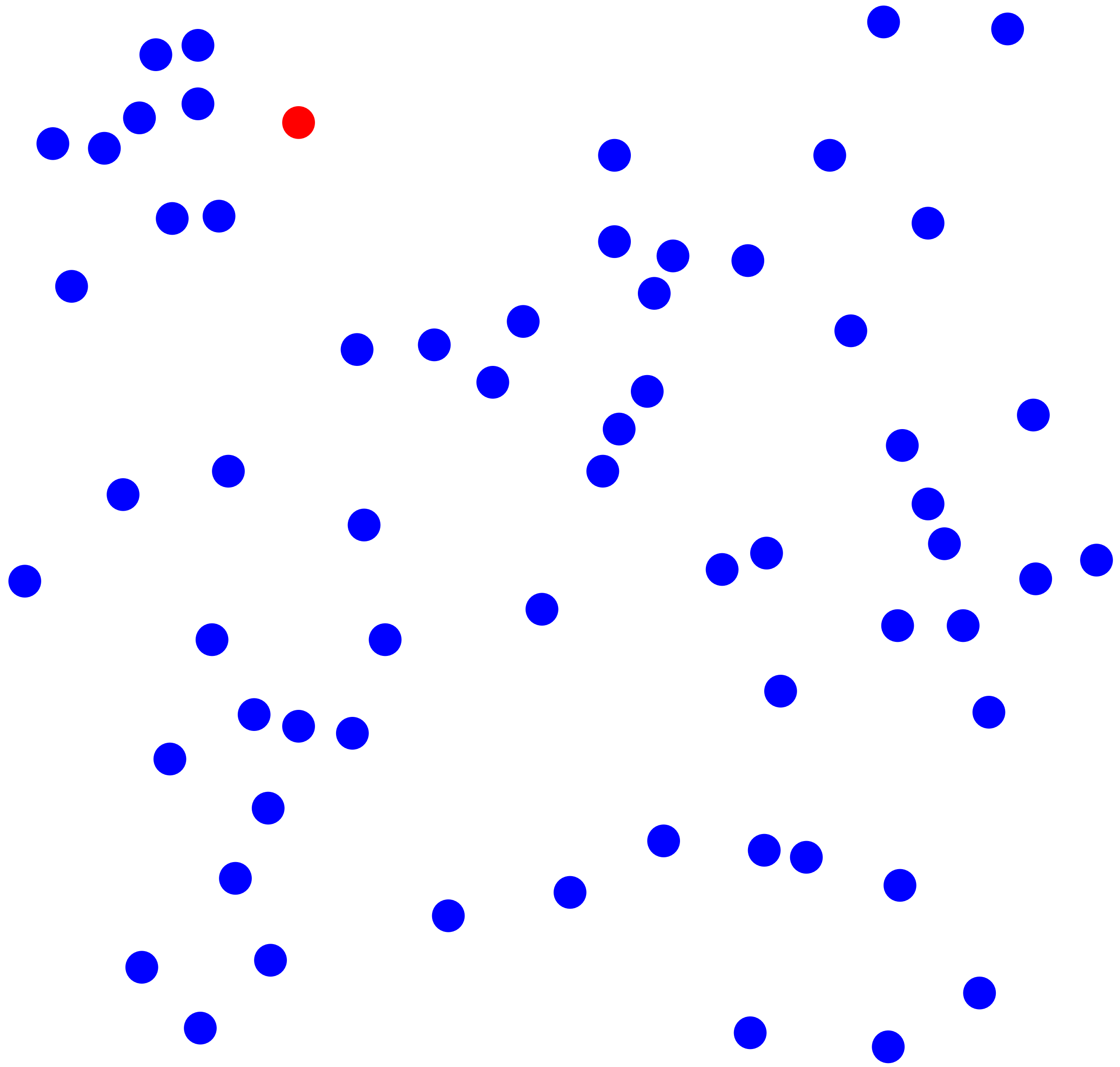
[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]

Visual pop out



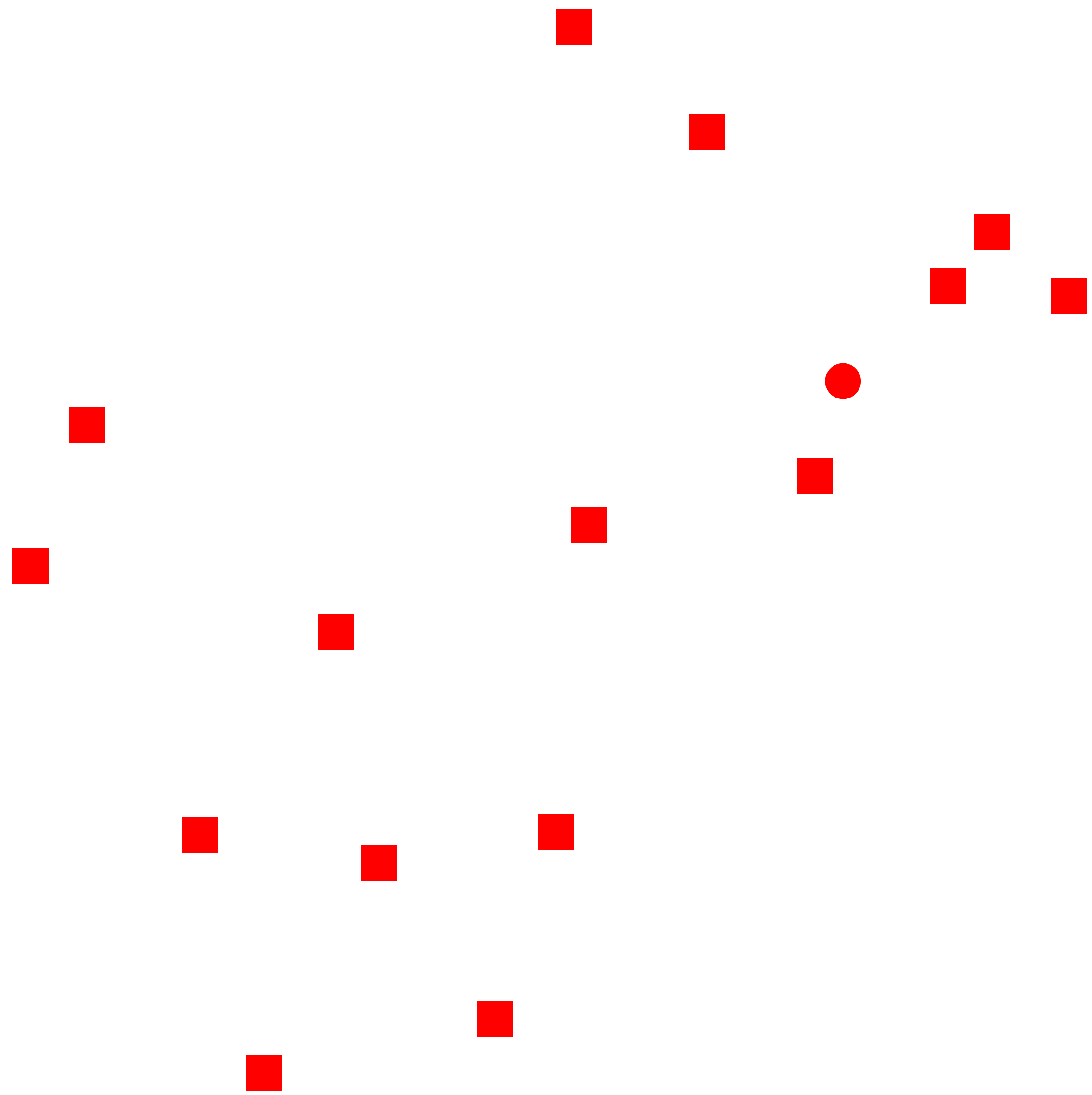
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



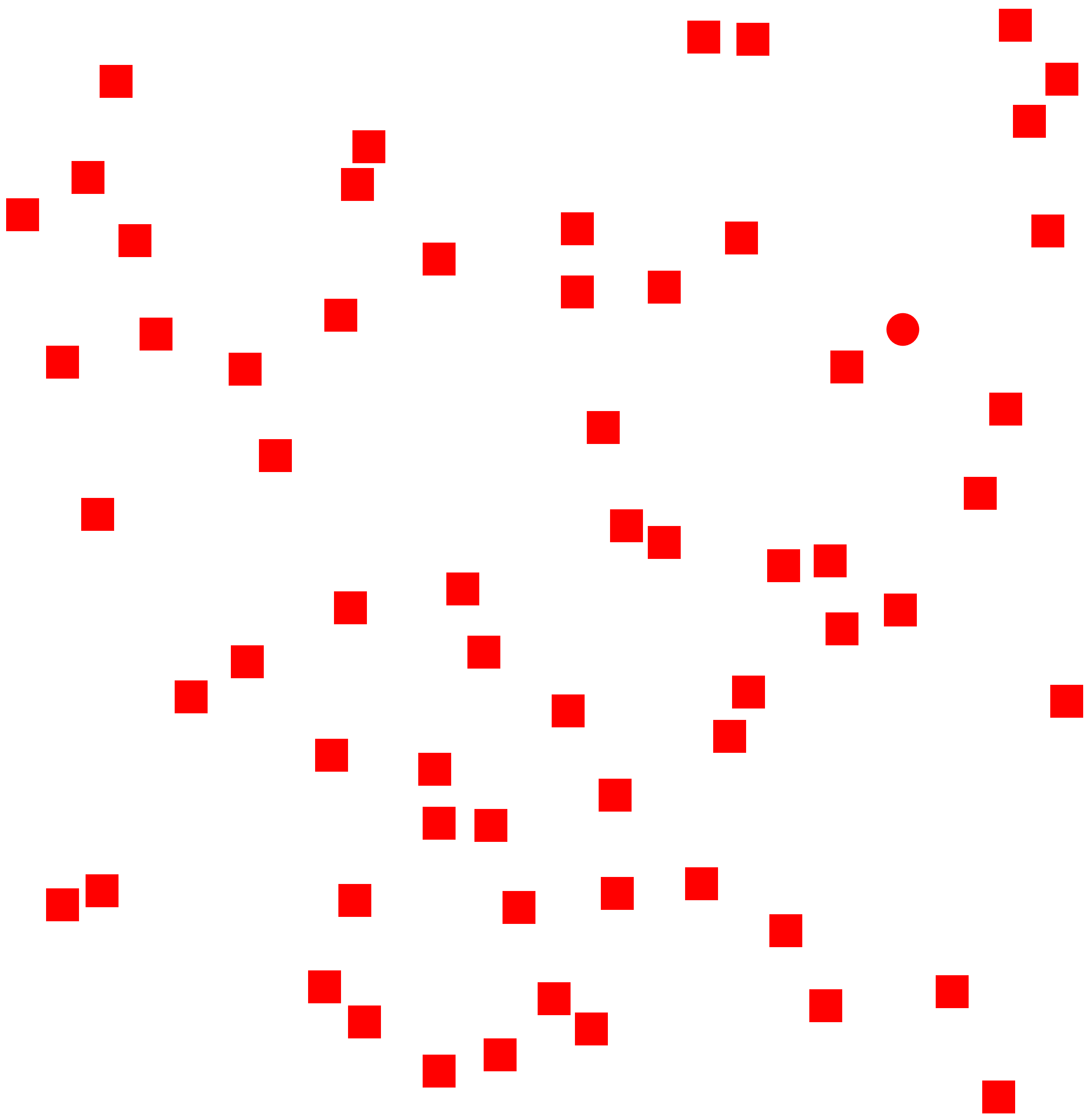
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



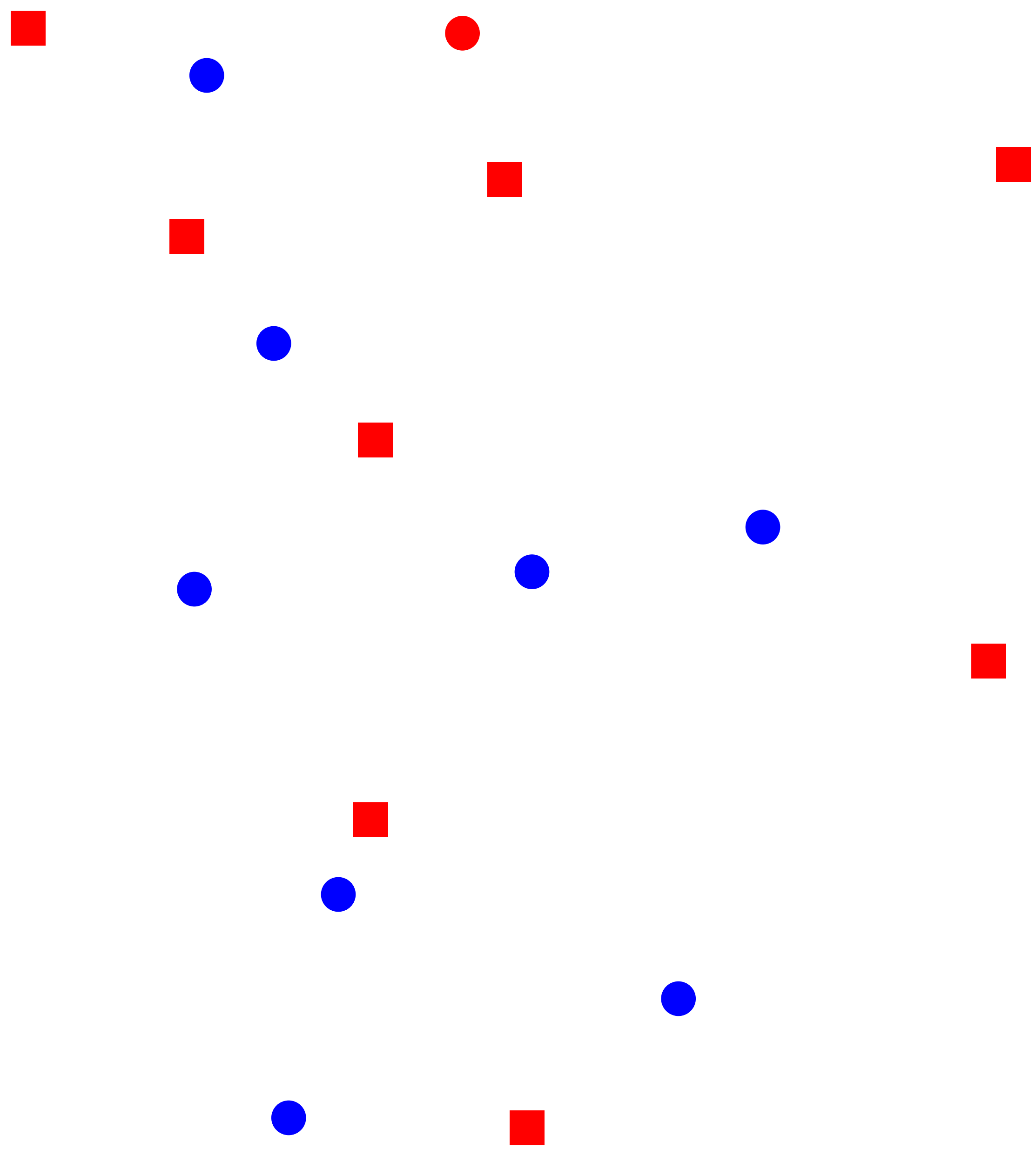
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



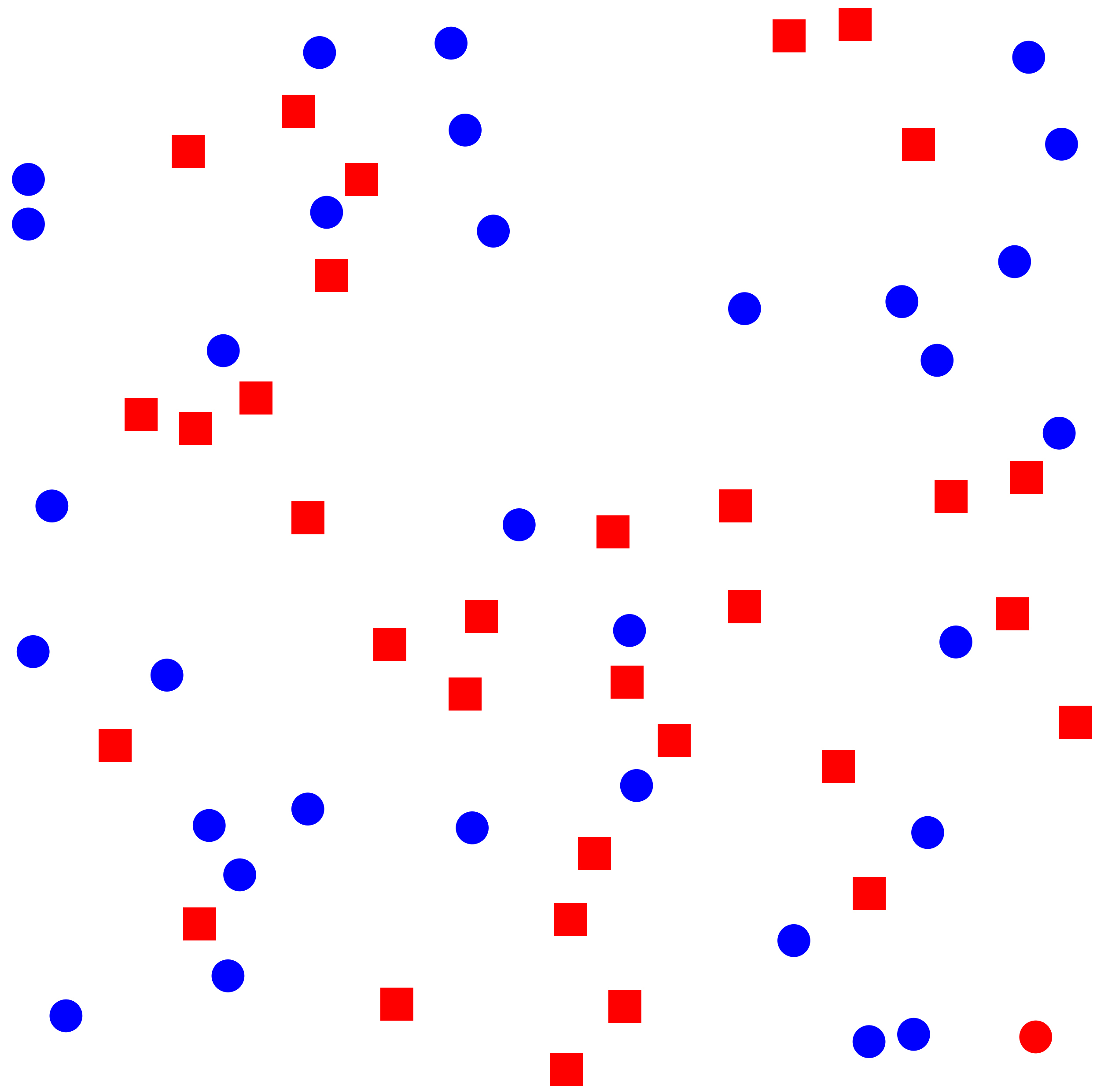
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



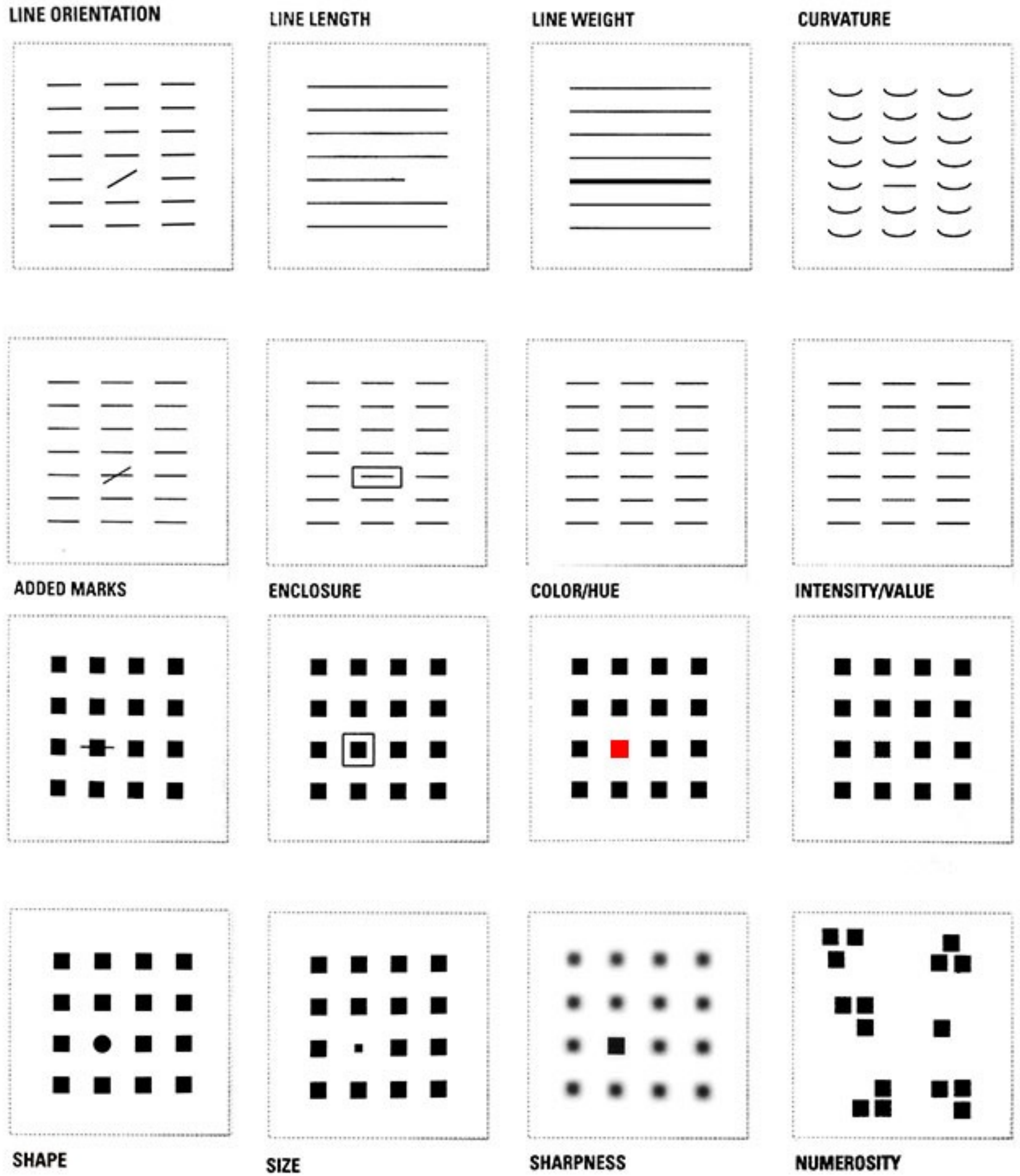
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



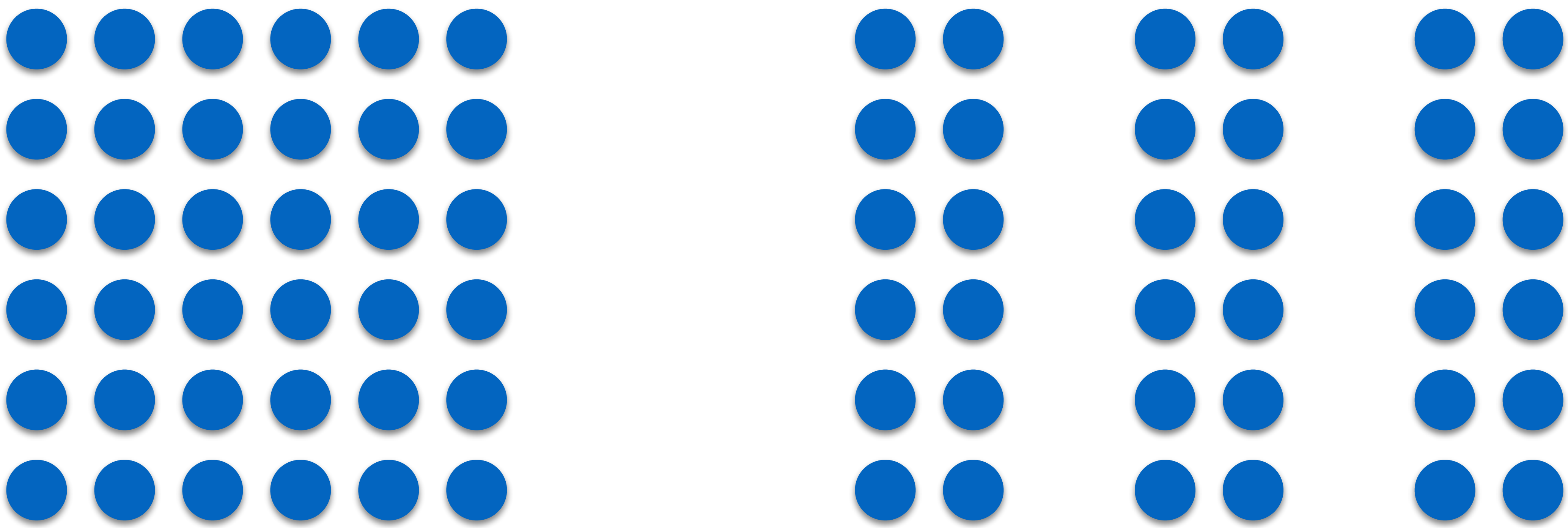
[Inspired by: Visualization Analysis and Design. Tamara Munzner. 2014]

Visual pop out



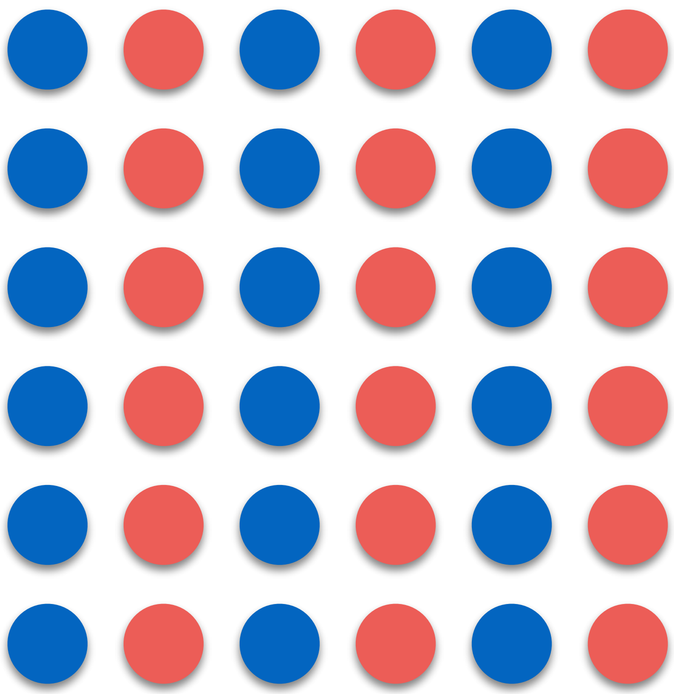
[Source: Visual Thinking for Design, Ware 2008]

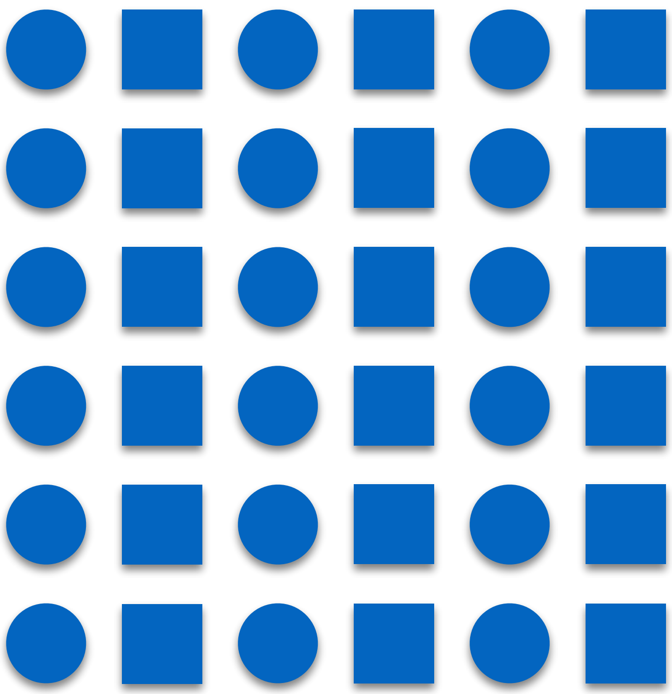
Gestalt Laws of Perceptual Organization



Law of proximity: objects that are close together are likely to be perceived as a group.

Gestalt Laws of Perceptual Organization

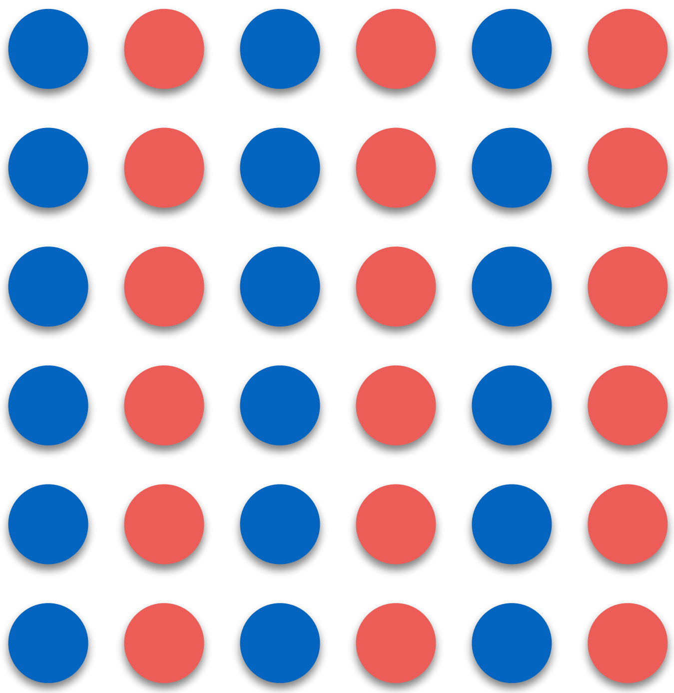


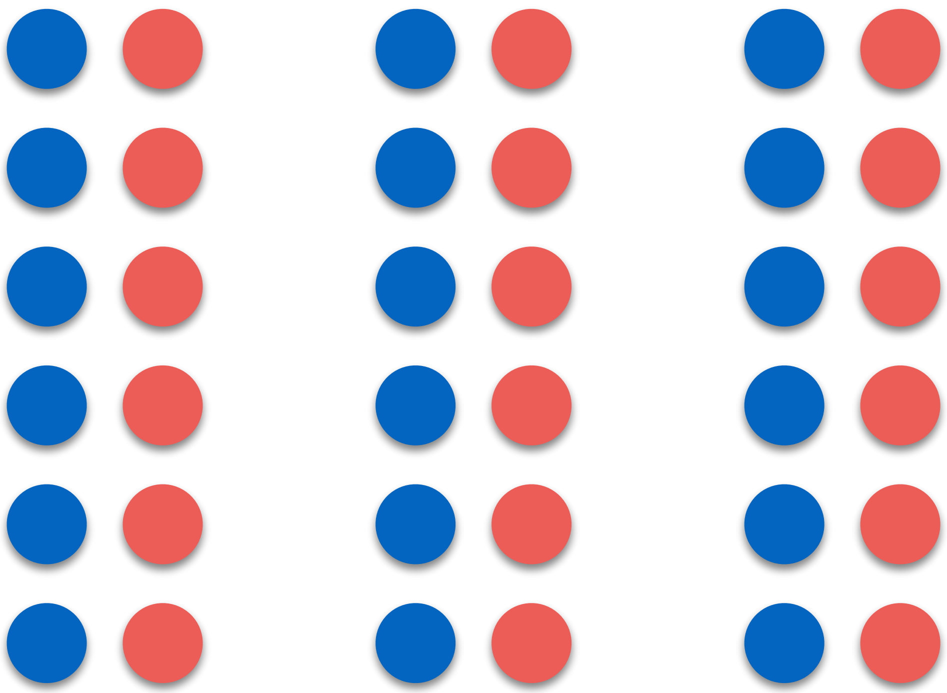


Law of similarity: objects that look similar are likely to be perceived as a group.

Similarity can occur in the form of color, shape, shading, *etc*.

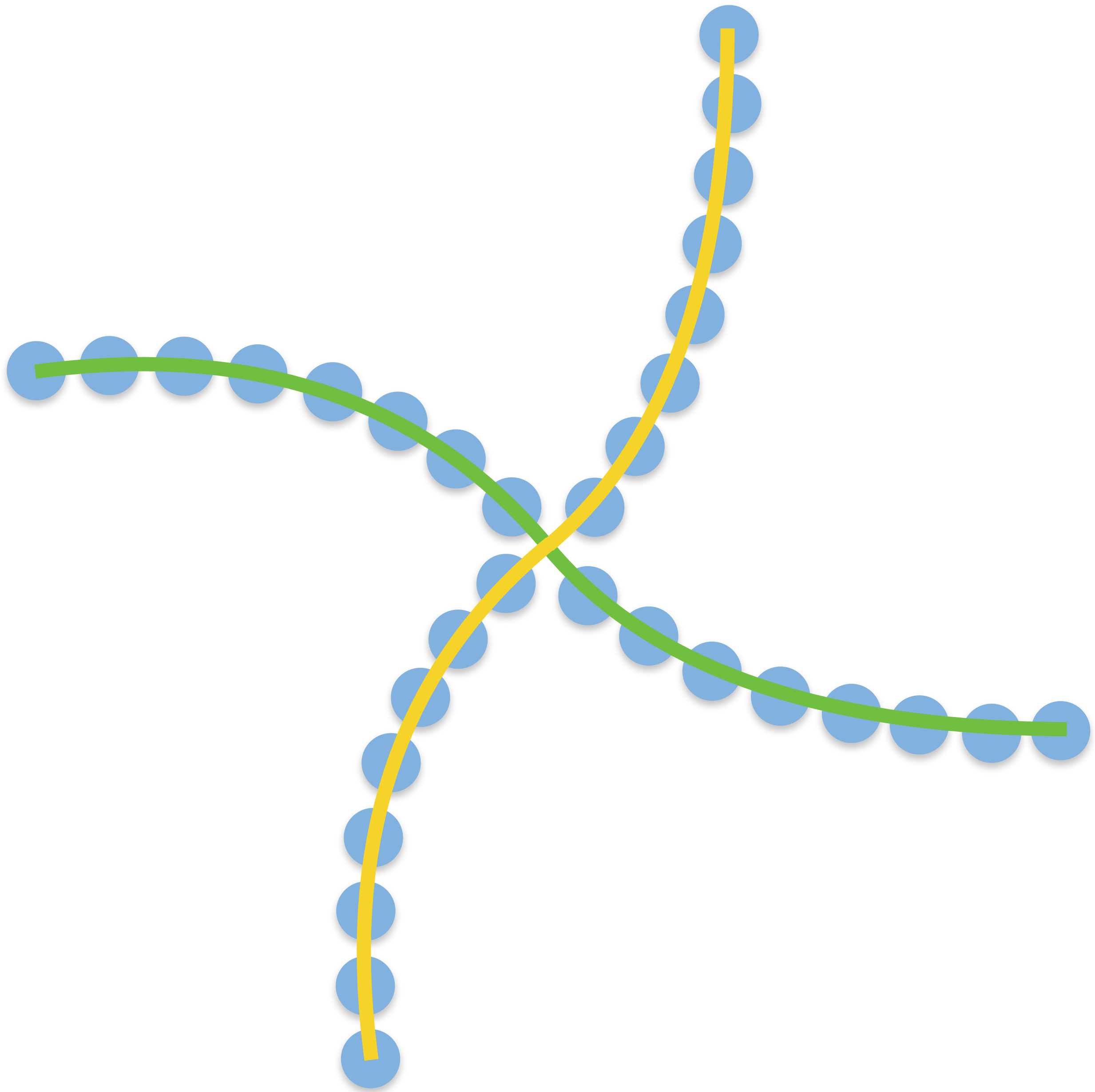
Gestalt Laws of Perceptual Organization





Proximity is stronger than color similarity

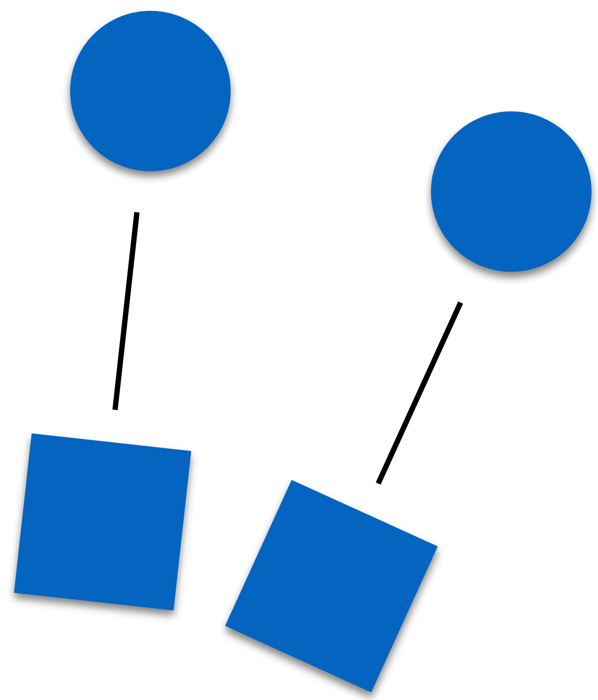
Gestalt Laws of Perceptual Organization



Law of continuity: objects that follow straight or curved lines without abrupt changes are perceived as belonging together.
In cases where there is an intersection between objects, individuals tend to perceive the two objects as two single uninterrupted entities.
Stimuli remain distinct even with overlap.

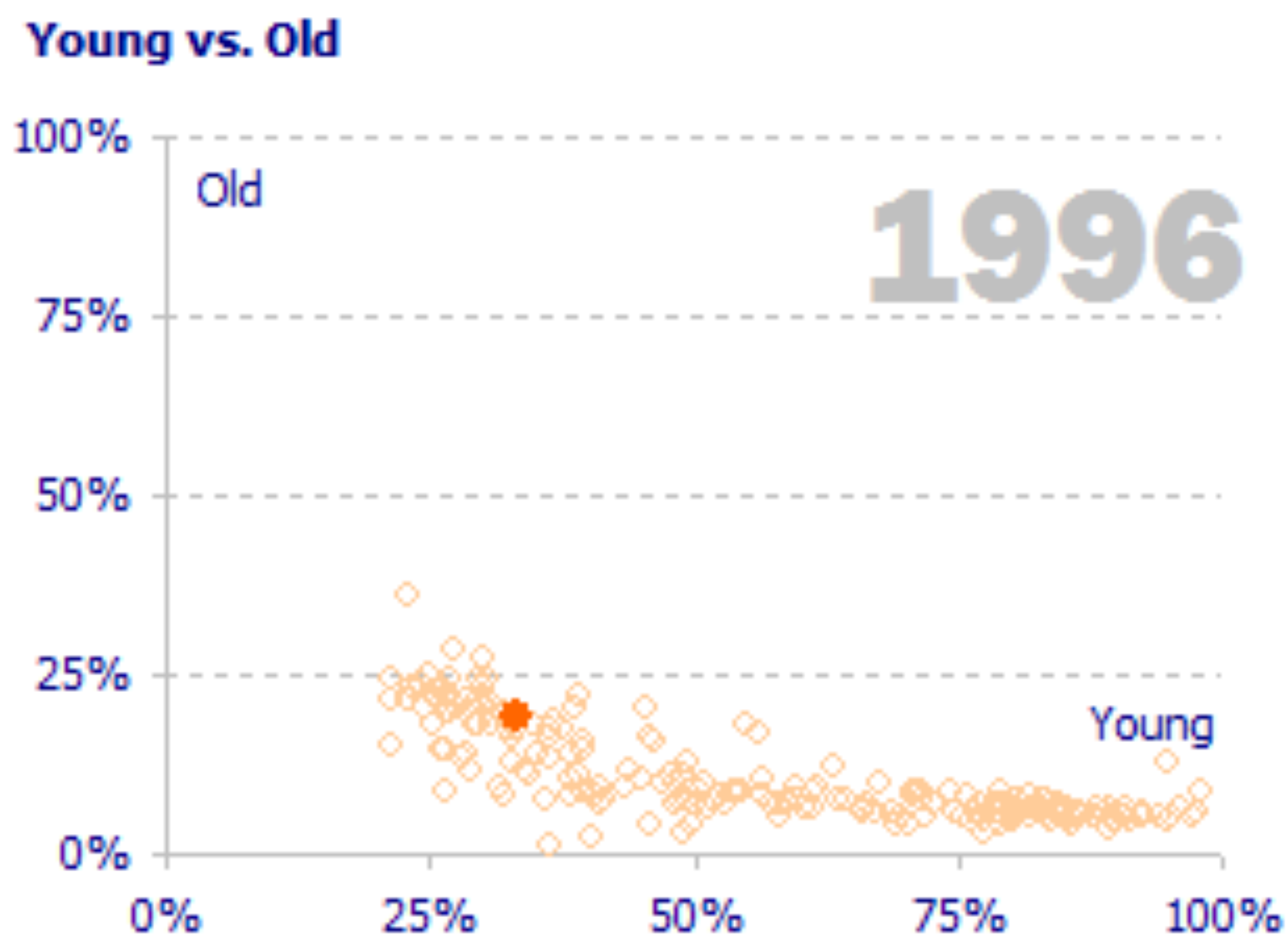
[Source: https://en.wikipedia.org/wiki/Gestalt_psychology]

Gestalt Laws of Perceptual Organization



Law of Connectedness: objects that are connected by lines are likely to be perceived as a group. This is true even if the resulting combinations are in conflict with other Gestalt laws.

Gestalt Laws of Perceptual Organization



Law of Common Fate: objects that move together are likely to be perceived as a group.

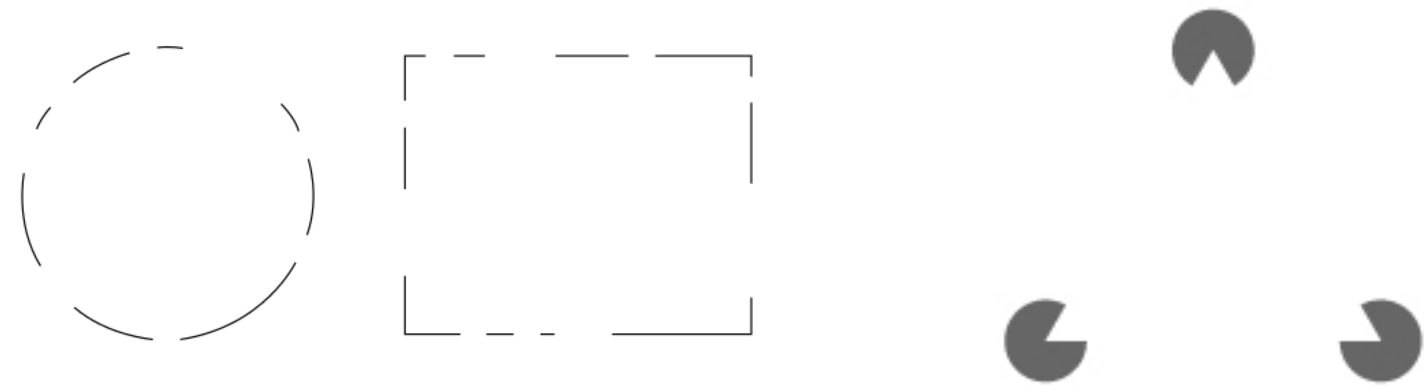
Movement of elements of an object produce paths that individuals perceive that the objects are on. Objects that have the same trend of motion are perceived as being on the same path.

[Source: https://en.wikipedia.org/wiki/Gestalt_psychology]

Additional Gestalt Laws of Perceptual Organization

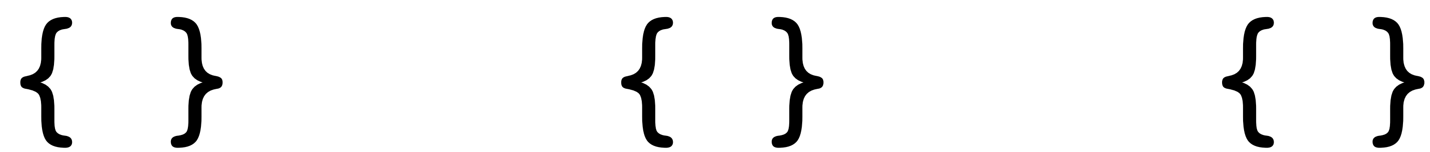
Law of closure: individuals perceive objects such as shapes, letters and images as being whole when they are not complete.

When parts of a whole picture are missing, our perception fills in the visual gap.



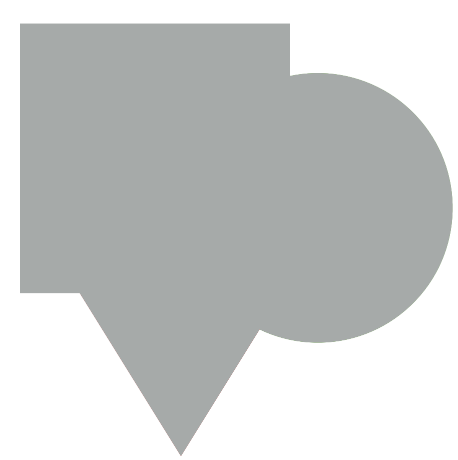
Law of symmetry: individuals perceive objects as being symmetrical and forming around a center point.

When two symmetrical elements are unconnected the mind perceptually connects them to form a coherent shape.



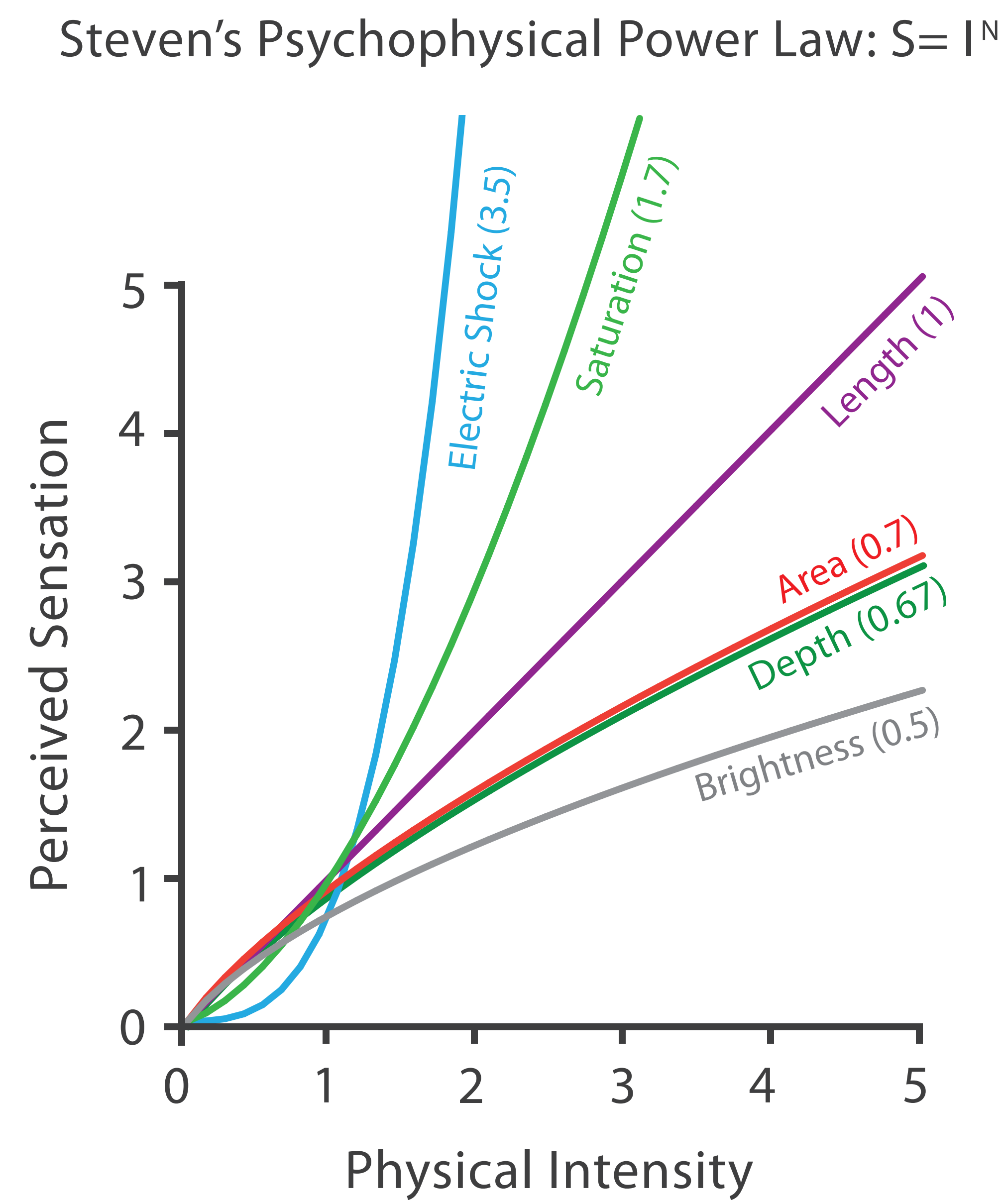
Law of Good Figure / Prägnanz principle: complex object are perceptually decomposed into simple geometries, based on the identification of regular, simple, orderly patterns.

Individuals eliminate complexity and unfamiliarity so they can observe a reality in its most simplistic form.



[Source: https://en.wikipedia.org/wiki/Gestalt_psychology]

Stevens' Psychophysical Power Law

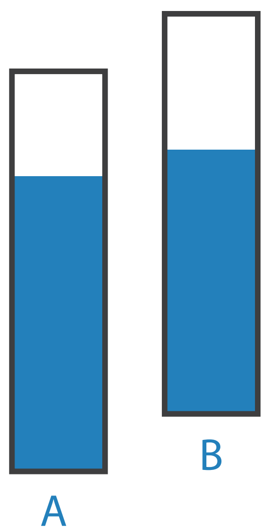


[Source: Tamara Munzner's course slides, 2017]

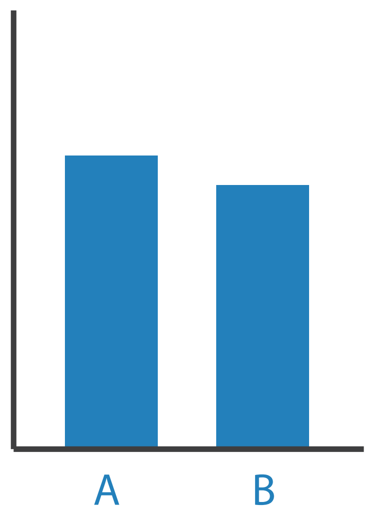
Relative *vs.* absolute judgements

- Our perceptual system mostly operates with relative judgements, not absolute ones.
- Accuracy increases with common frame/scale and alignment.

Length



Position along unaligned common scale



Position along aligned common scale

[Source: adapted from InfoVis course, T. Munzner, 2017]

Characterizing visual variables

From Bertin

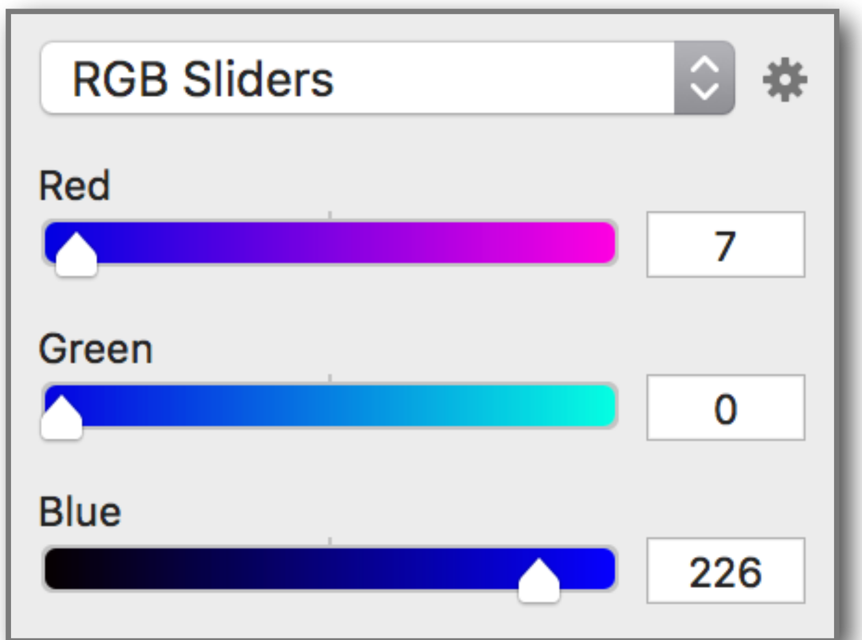
- Selective (related to visual popout):
  Can this variable allow us to spontaneously differentiate/isolate items from groups?
- Associative:
  Can this variable allow us to spontaneously put items in a group?
- Ordered:
  Can this variable allow us to spontaneously perceive an order?
- Quantitative:
  Is there a numerical reading obtainable from changes in this variable?
- Length, or rather, Discriminability (resolution):
  Across how many changes in this variable are distinctions possible?

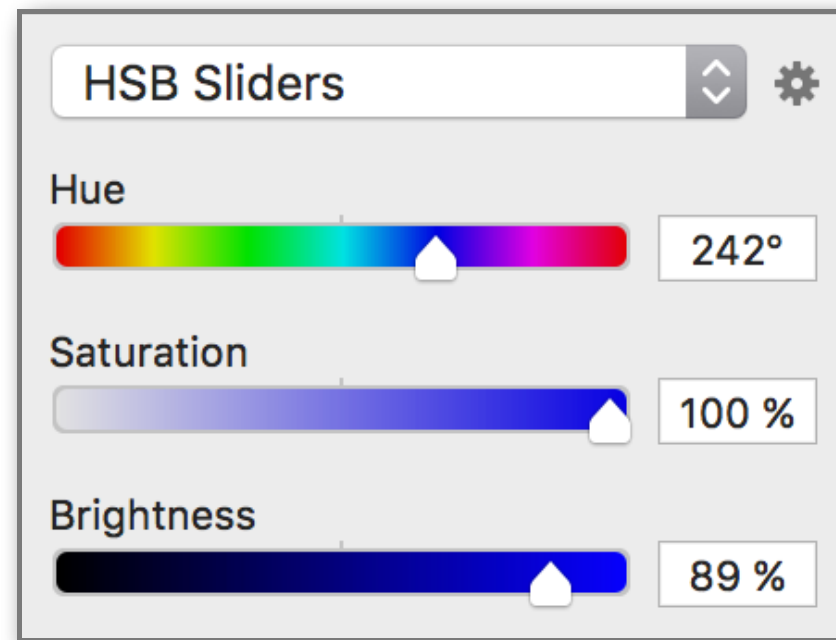
- Accuracy:
  Can we make accurate judgments about this variable in terms of perception?
- Separability vs. Integrality:
  Are two (or more) variables interfering with each other?

Color hue

Color saturation

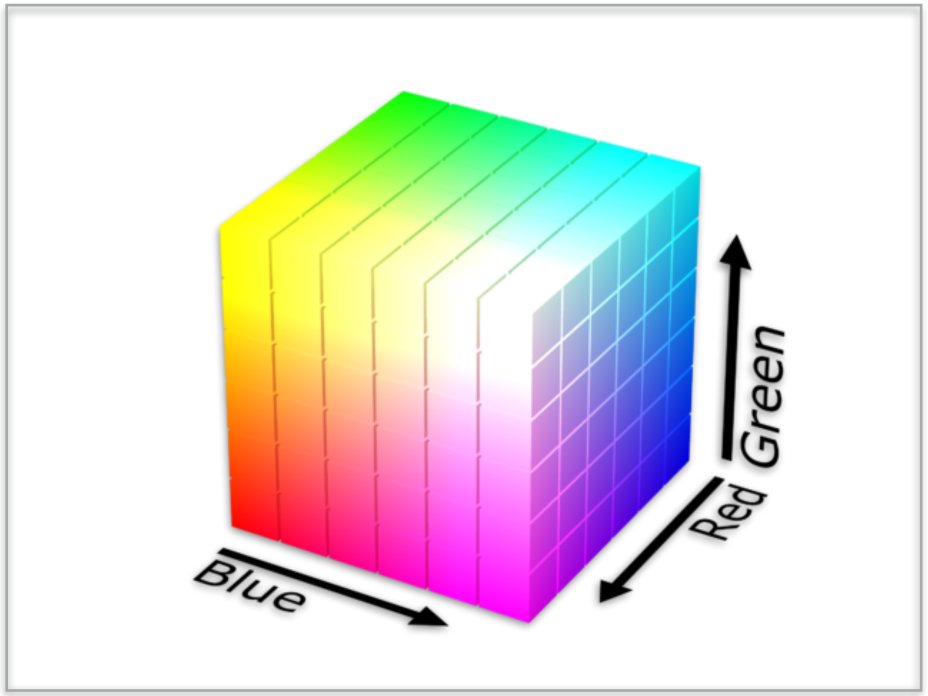
Color luminance



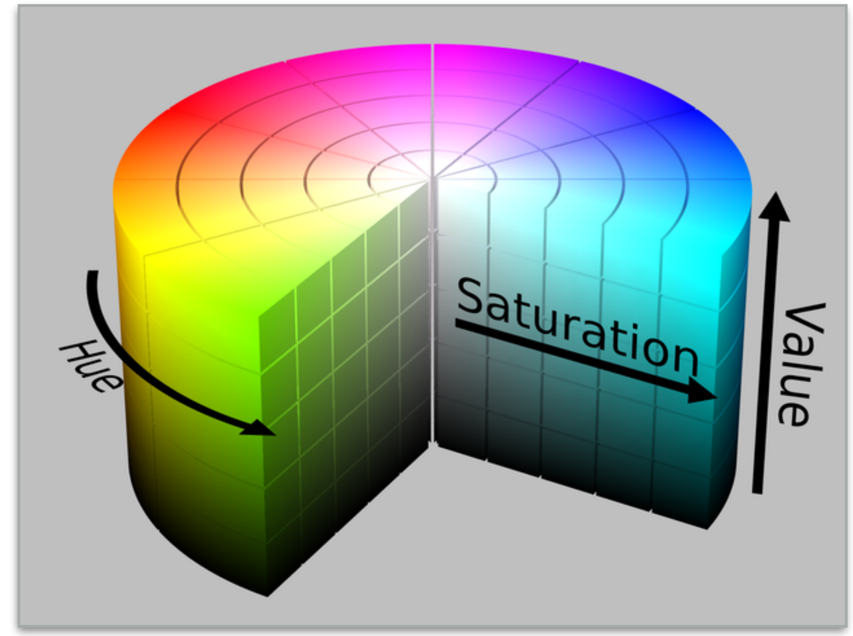


[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]

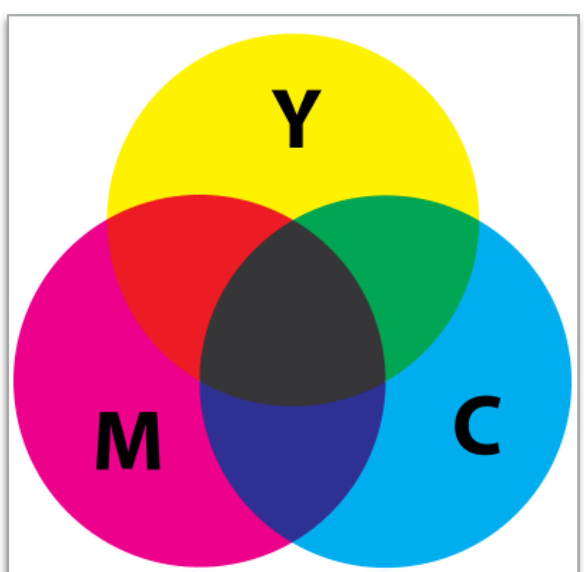
Color hue Color saturation Color luminance



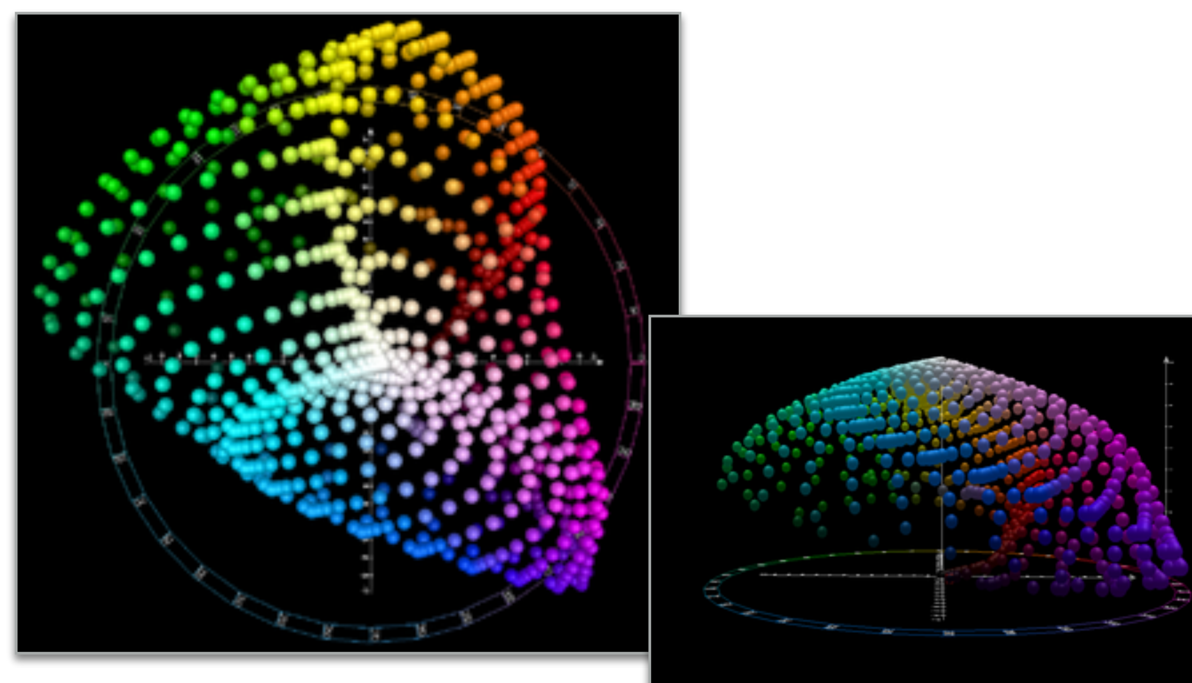
RGB



HSV



CMYK

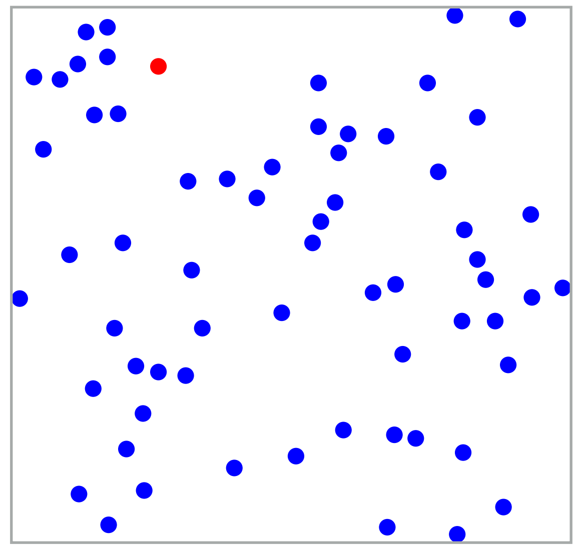


CIE L*a*b and L*C*h

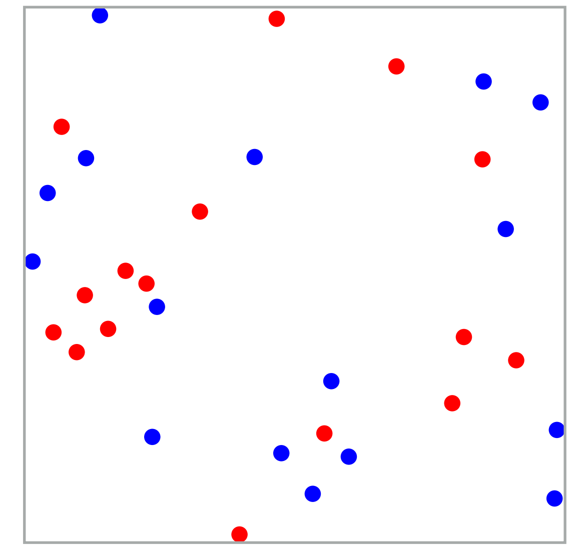
[Image source: WikiMedia]

Color hue

- Selective: ✓



- Associative: ✓

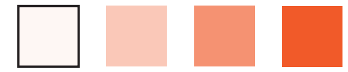


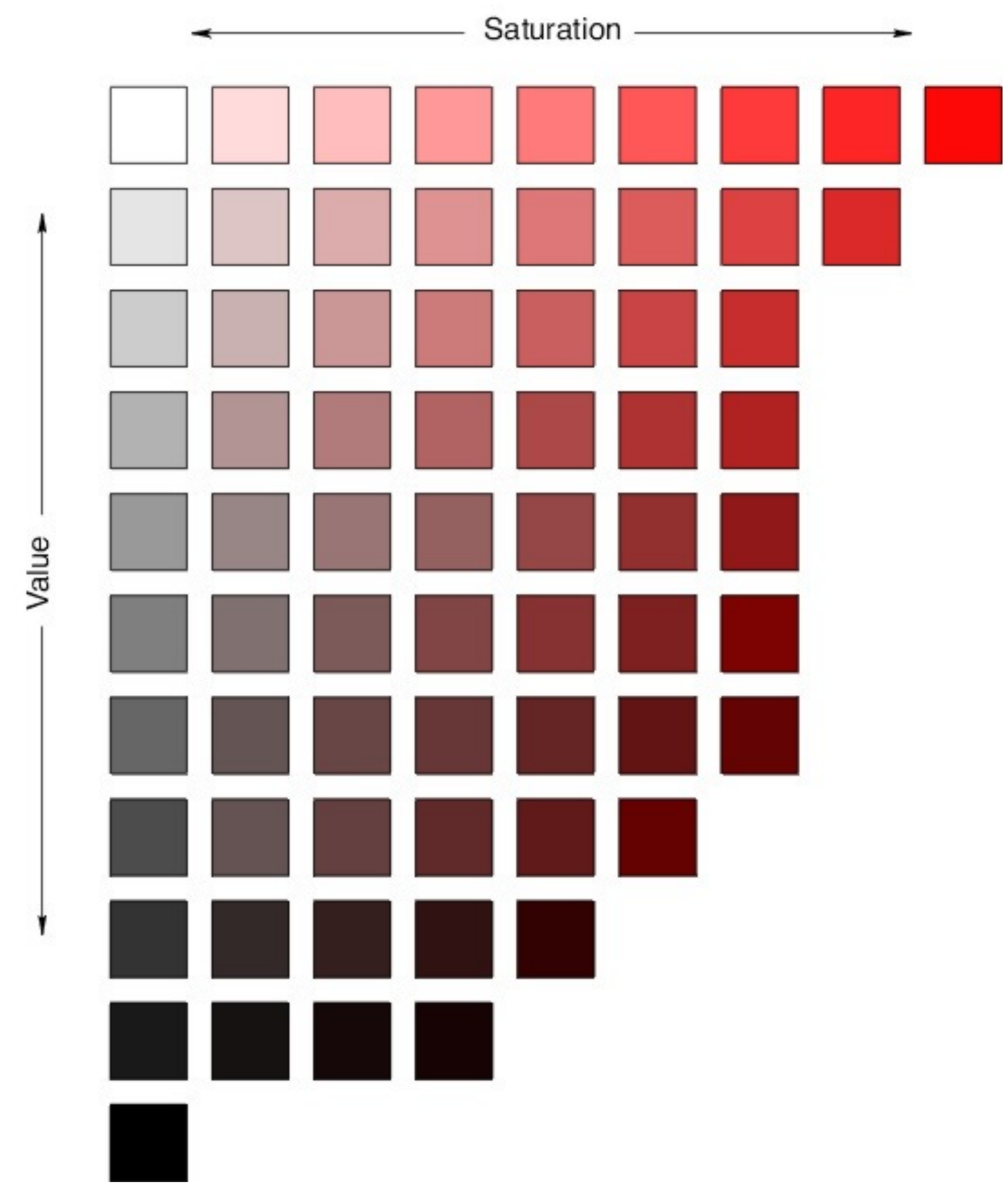
- Ordered: ✗

$\ngeq \quad \ngeq \quad \ngeq$

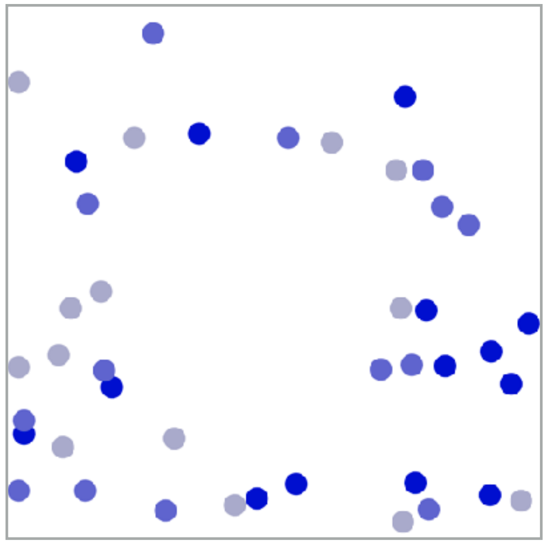
- Quantitative: ✗

- Discriminability (resolution):
  - in theory: ∞
  - in practice: ≲7 for selection and association, ≲10 for distinction **but…**

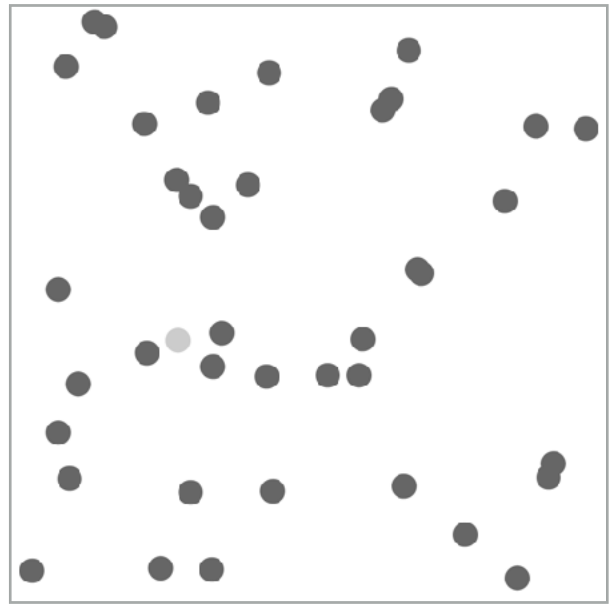
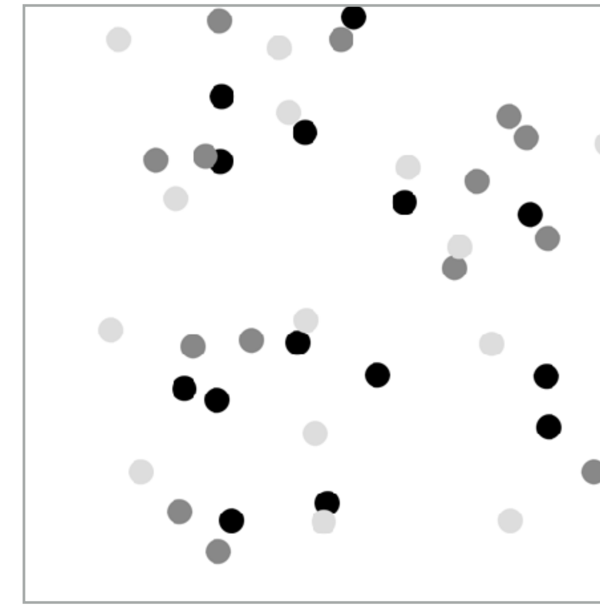
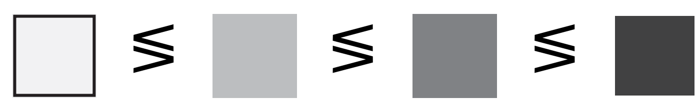
Color saturation  Color luminance 

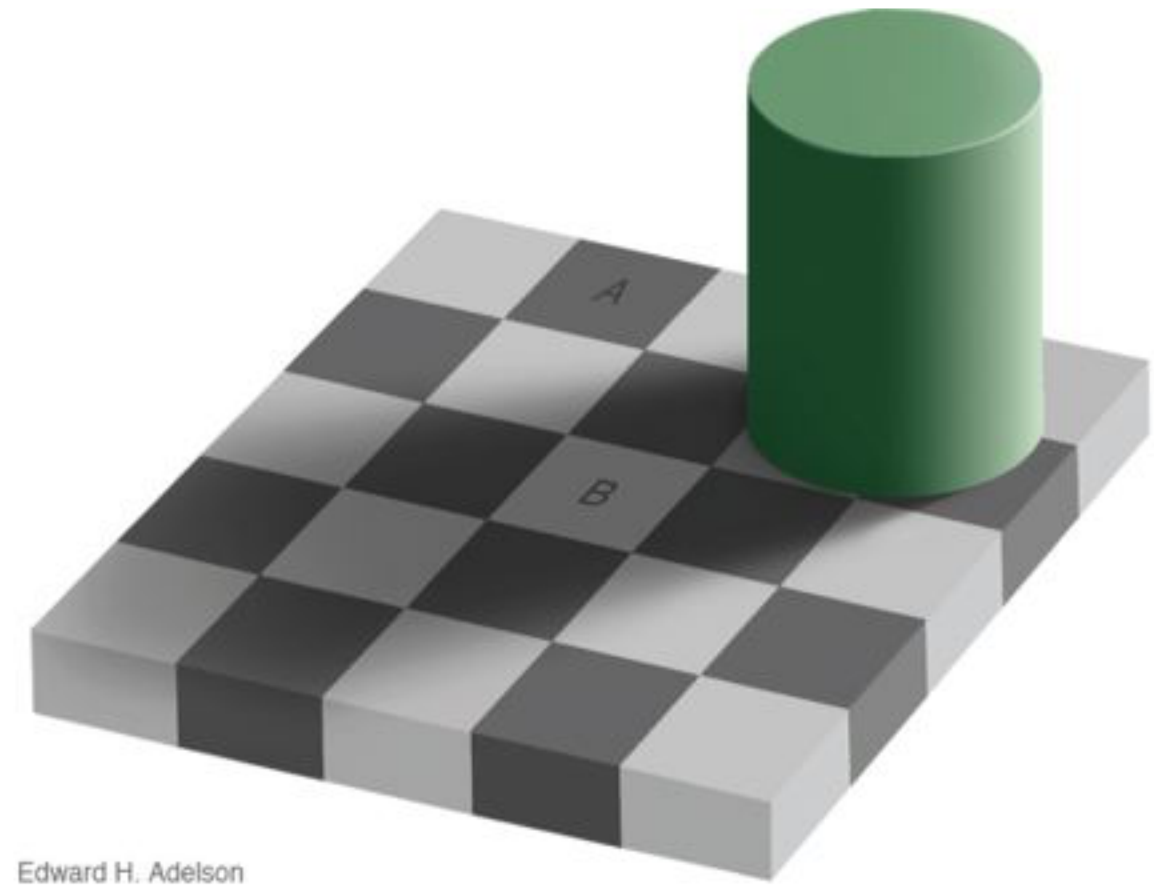


Color saturation 

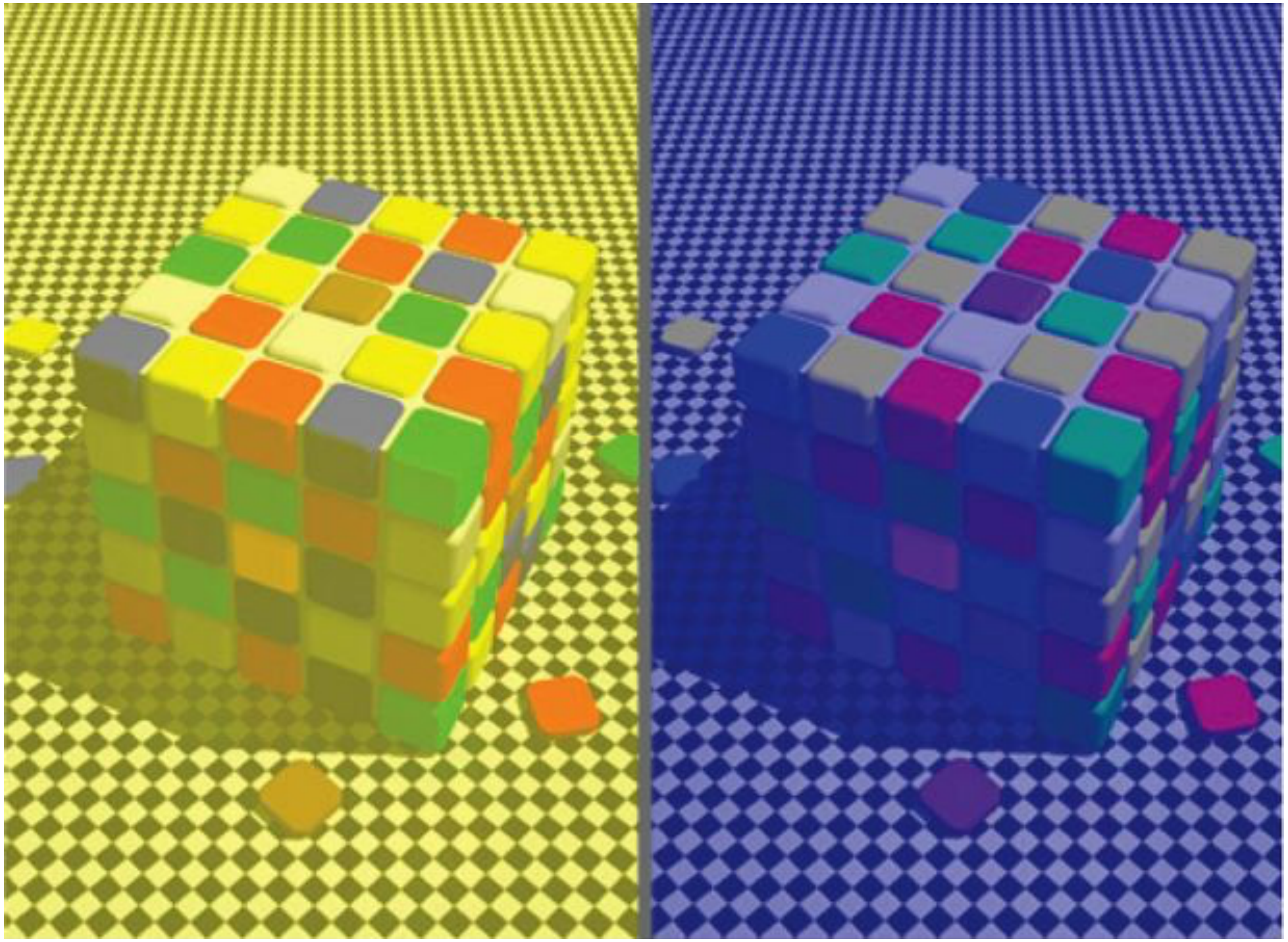
- Selective: ✓ 
- Associative: ✓ 
- Ordered: ✓  (will often be <, but depends on the semantics)
- Quantitative: ✕
- Discriminability (resolution):
  - in theory: ∞
  - in practice: ≲7 for selection and association, ≲10 for distinction

Color luminance 

- Selective: ✓ 
- Associative: ✓ 
- Ordered: ✓  (highly depends on the semantics)
- Quantitative: ✕
- Discriminability (resolution):
  - in theory: ∞
  - in practice: ≲7 for selection and association, ≲10 for distinction



Luminance perception is based on relative, not absolute, judgements.



Color perception is also relative to surrounding colors and depends on context. Both cubes have tiles that appear to be red.

Sequential color scales

Variation in lightness is the most important parameter

[Figure source: Robert Simmon, Subtleties of Color, NASA Earth Observatory, 2013]

Sequential color scales

Variation in lightness is the most important parameter

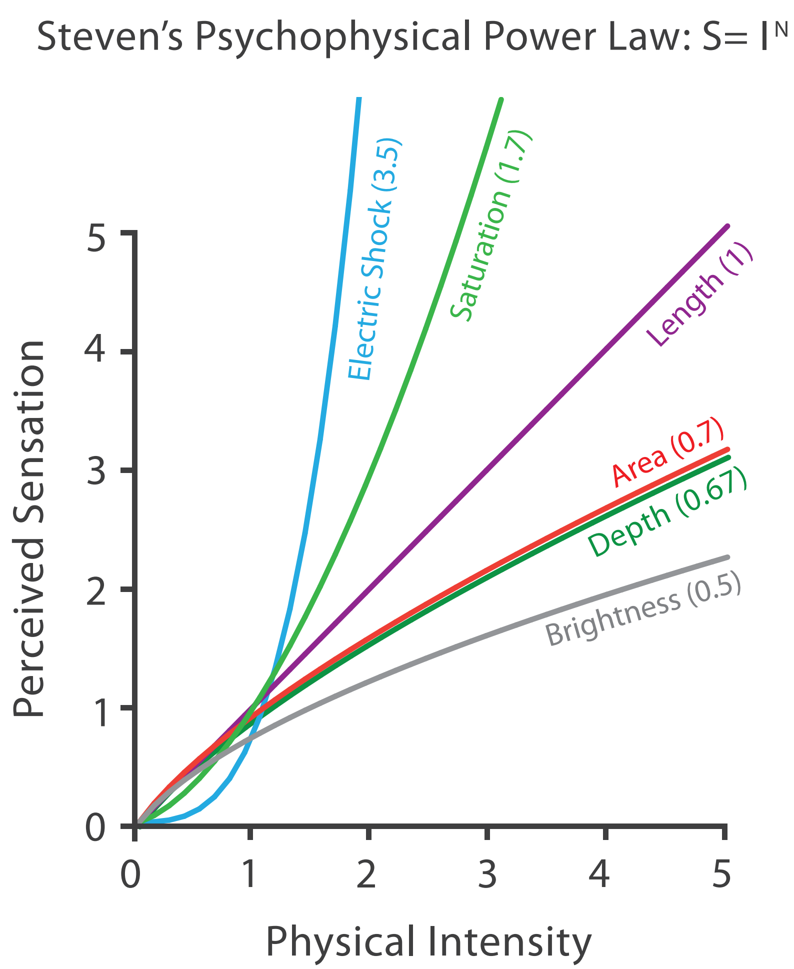
Varying the hue will provide further color contrast

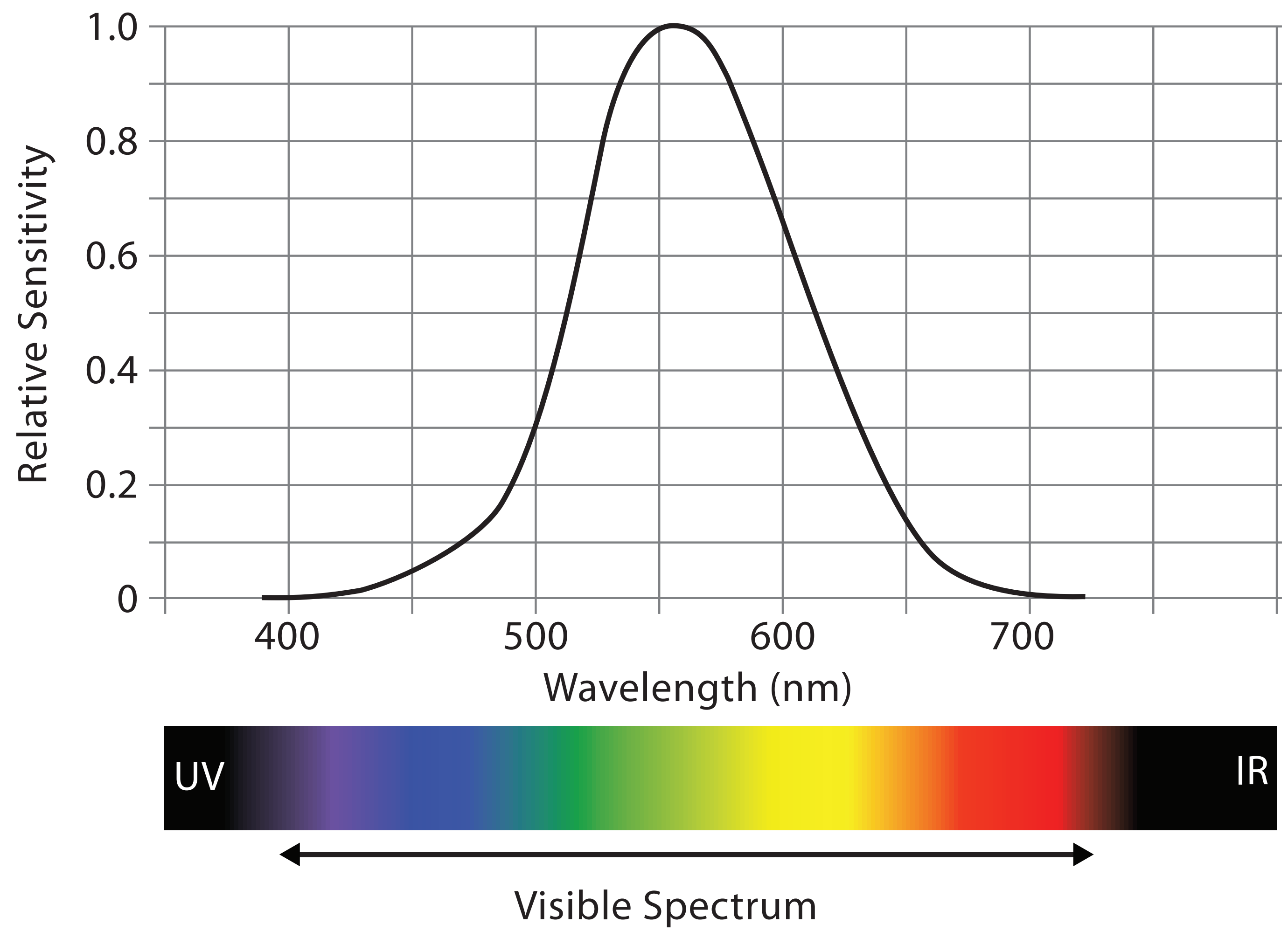
[Figure source: Robert Simmon, Subtleties of Color, NASA Earth Observatory, 2013]

Photopic Spectral Sensitivity

More sensitive to changes at low light levels

More sensitive to green than red light, and even less sensitive to blue light





Color hue Color saturation Color luminance

Sequential color scales

Variation in lightness is the most important parameter

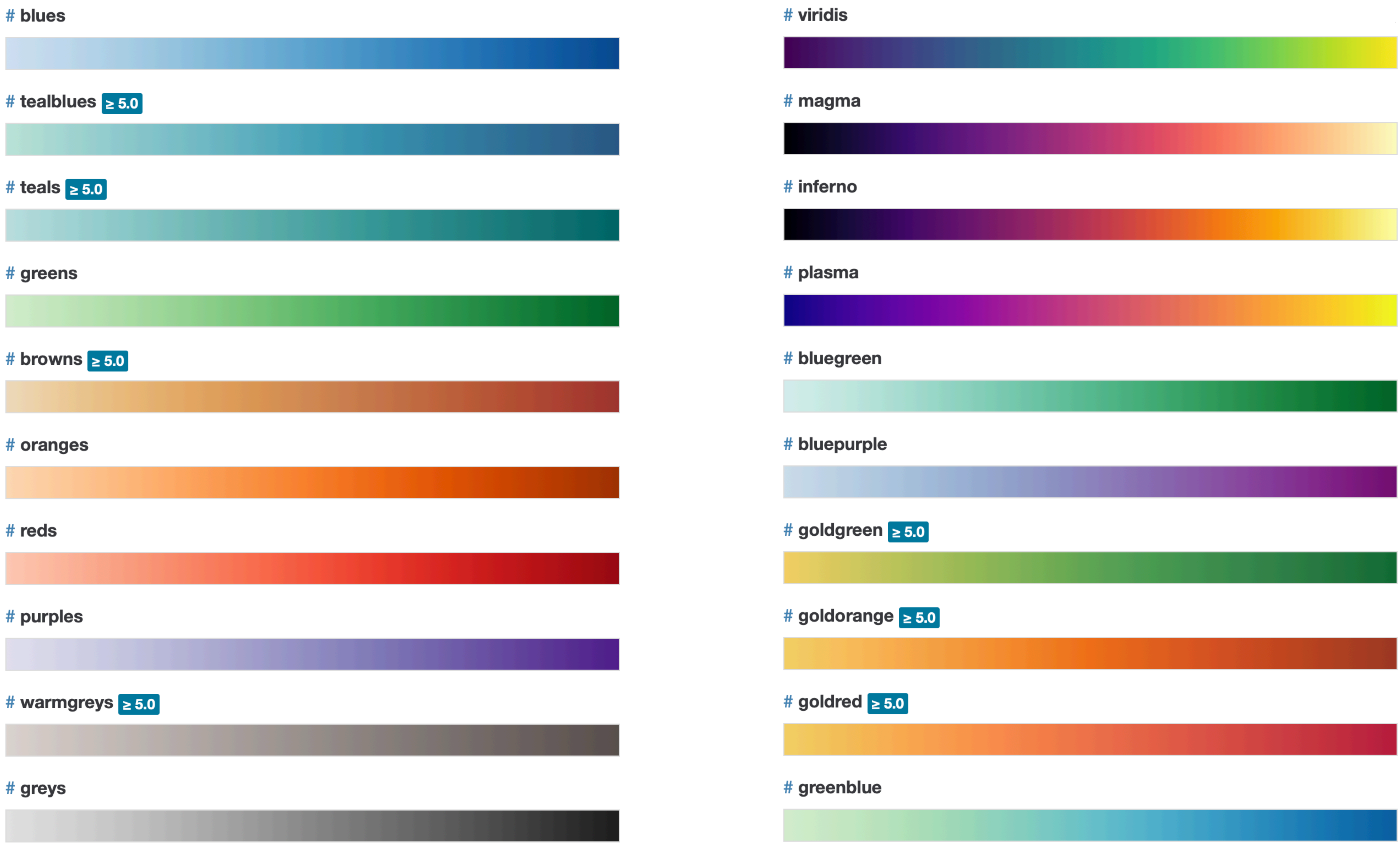
Varying the hue will provide further color contrast

Create *perceptually* linear variations. Counter example:

[Figure source: Robert Simmon, Subtleties of Color, NASA Earth Observatory, 2013]

Color hue   Color saturation   Color luminance

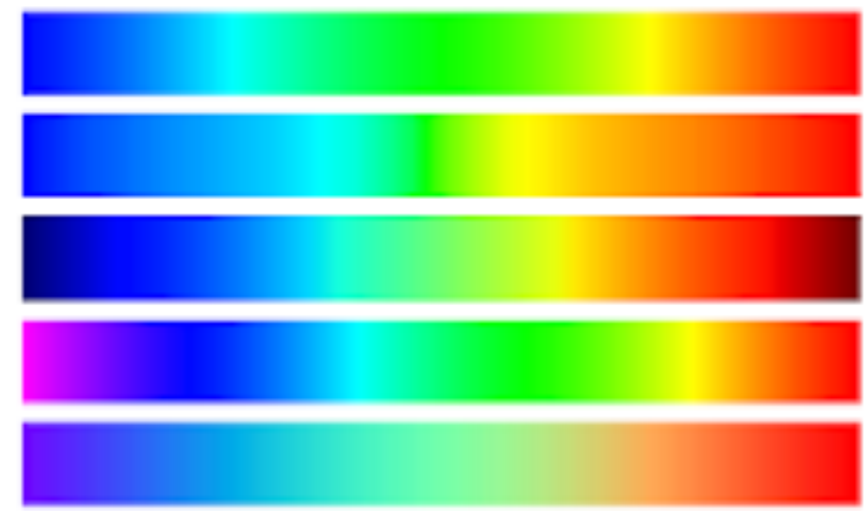
Sequential color scales



[Captured from https://vega.github.io/vega/docs/schemes/]

Rainbow color scale(s)

We do not see it as ordered.



Luminance suggests depth (bright => peak)

Some scales are highly non-monotonic.

It tends to hightlight arbitrary aspects of the data.

It makes us see categorical boundaries in otherwise continuous data (illusory discretization).

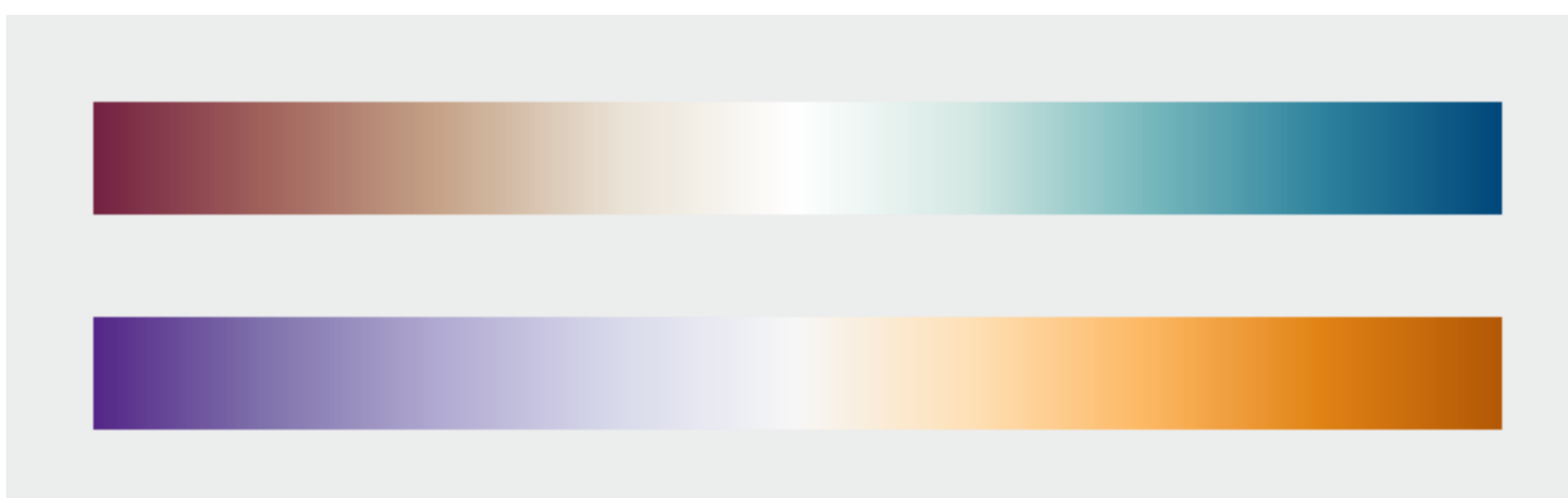
Bifurcated color scales

Divergent (or bipolar) data:

- varies from a central value
- implies a qualitative change as it crosses a threshold

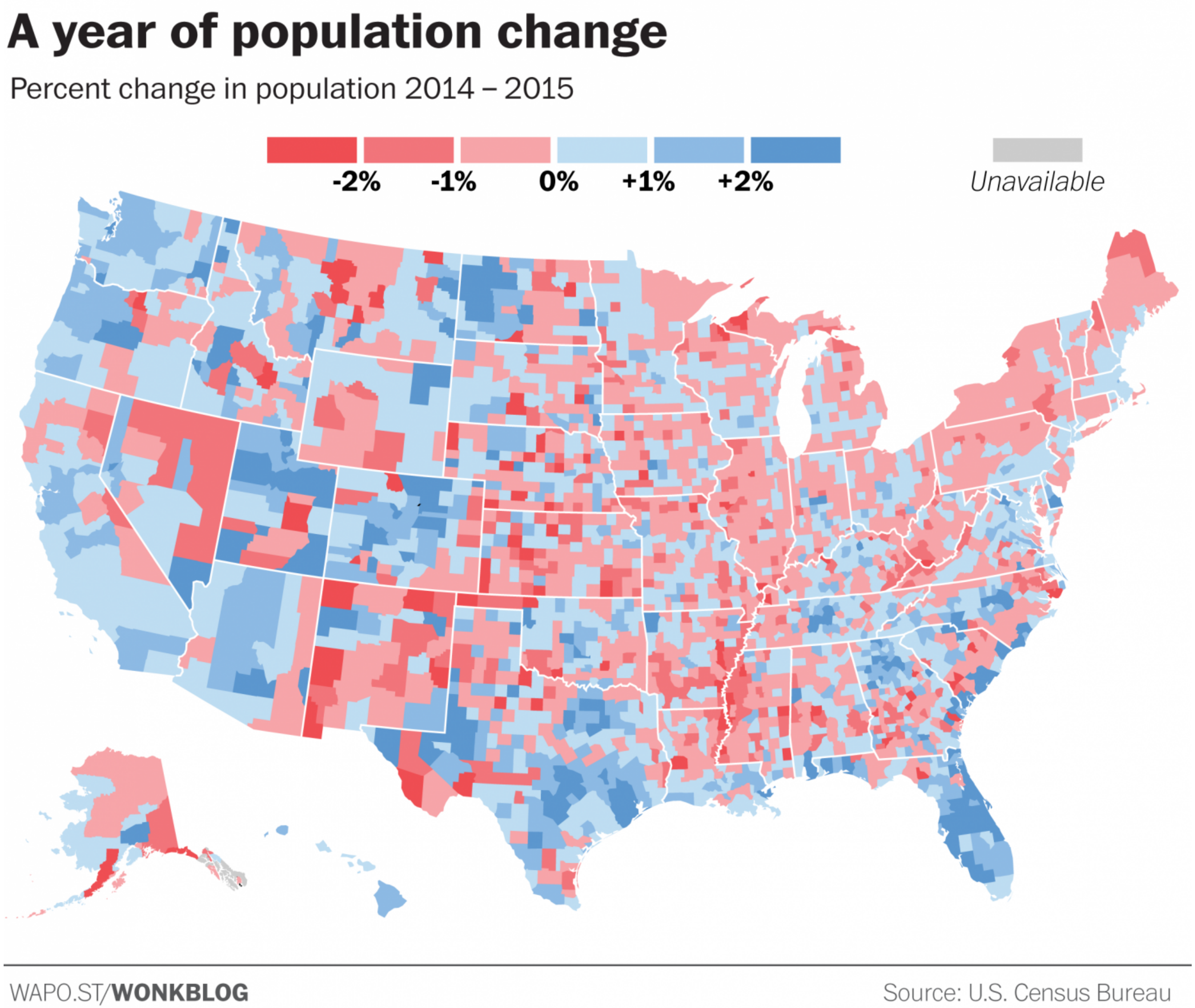
Best represented with a bifurcated color palette

- neutral color at the center



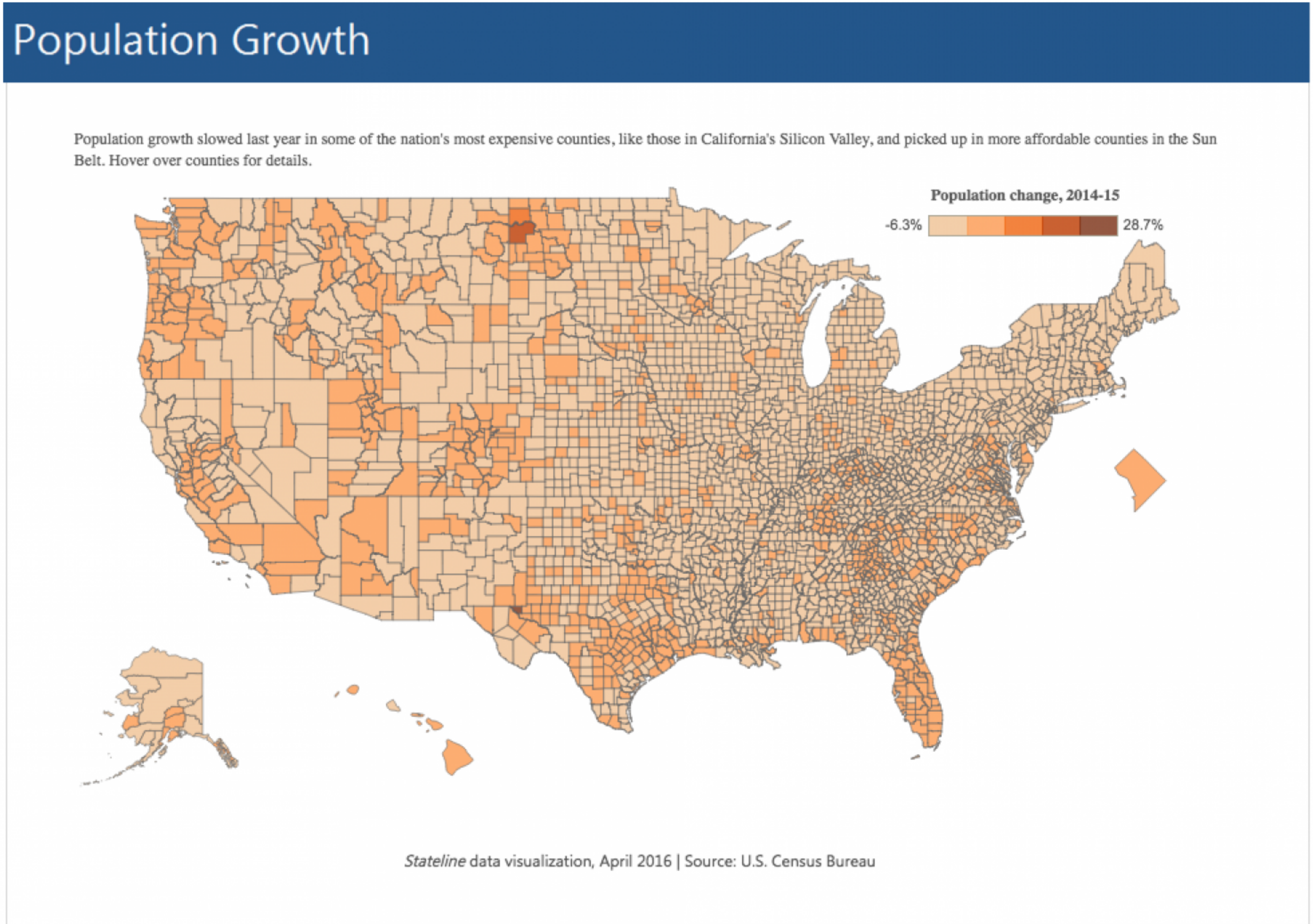
[Figure source: Robert Simmon, Subtleties of Color, NASA Earth Observatory, 2013]

Proper use of color



[Source: Christopher Ingraham, Washington Post, 2016-04-11]

Improper use of color (exact same data as previous map)

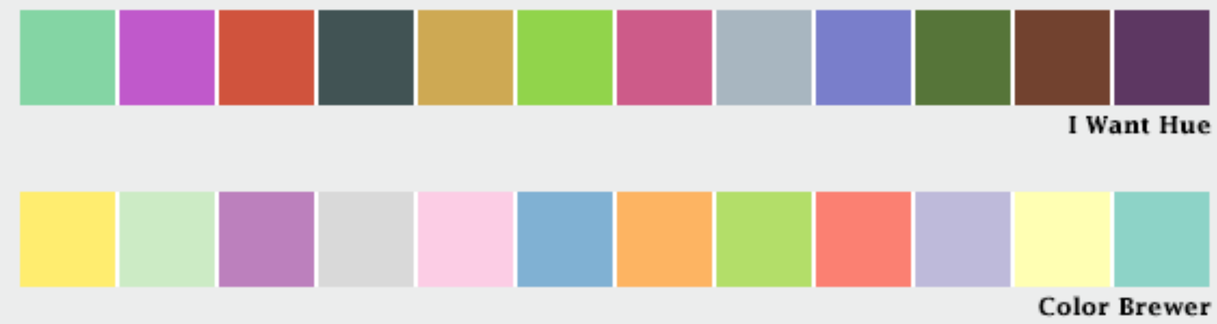


[Source: Christopher Ingraham, Washington Post, 2016-04-11]

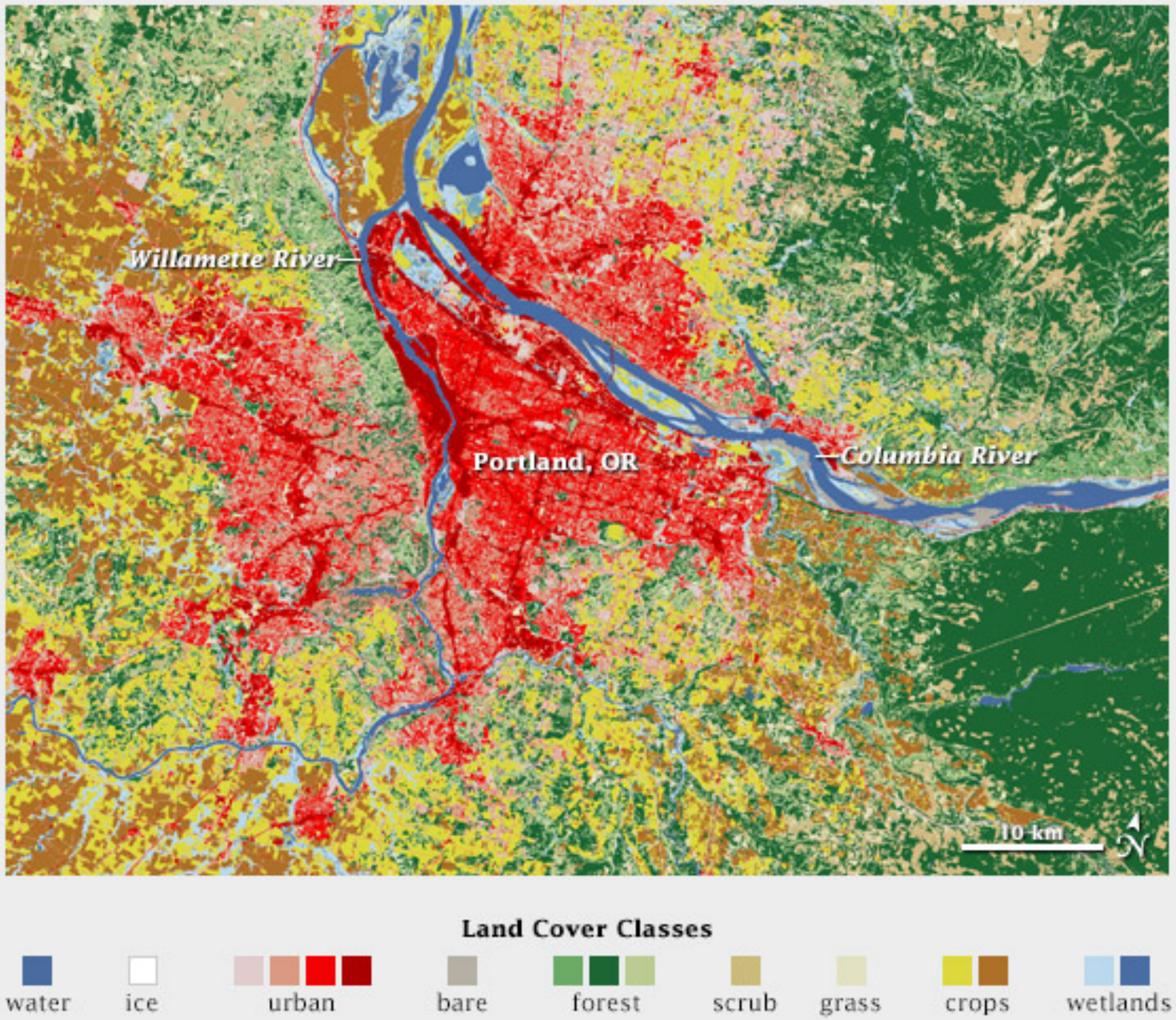
Color hue Color saturation Color luminance

Qualitative color schemes

Colors as distinct from one another as possible.

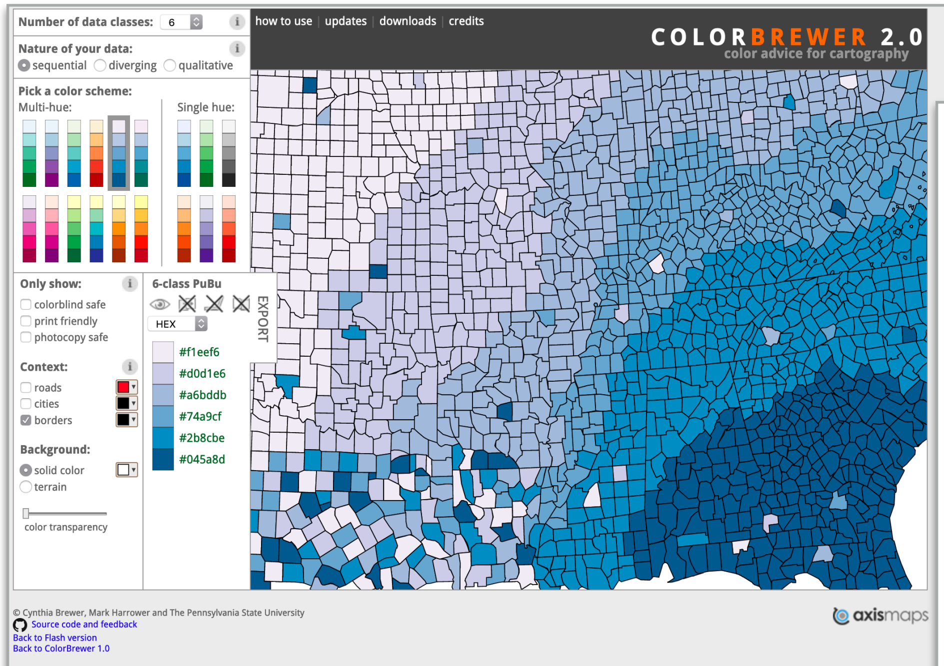


If too many categories, try to group them:

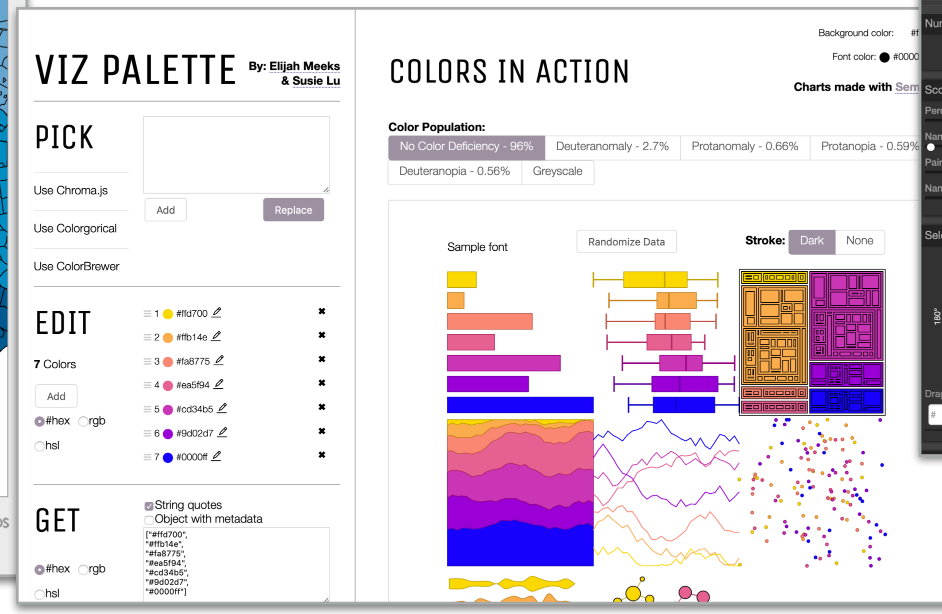


[Figure source: Robert Simmon, Subtleties of Color, NASA Earth Observatory, 2013]

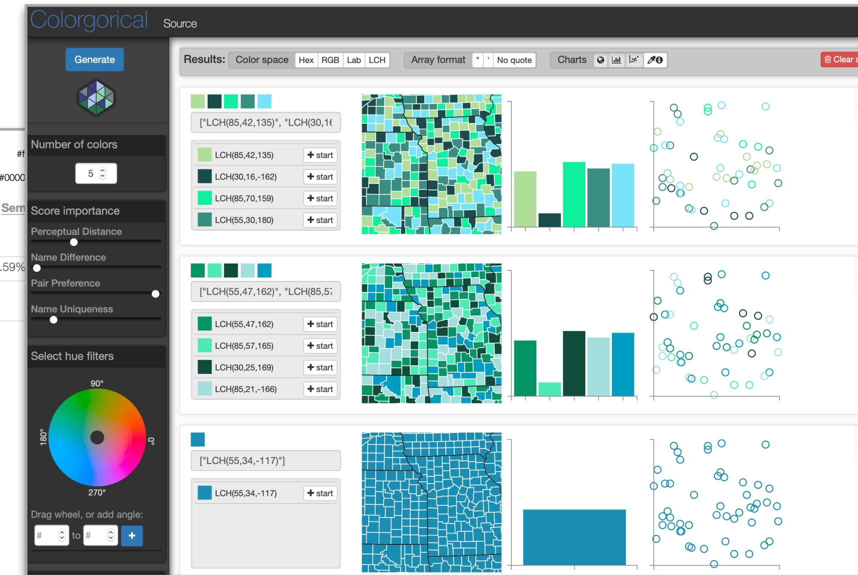
Color Palettes for data visualization



http://colorbrewer2.org

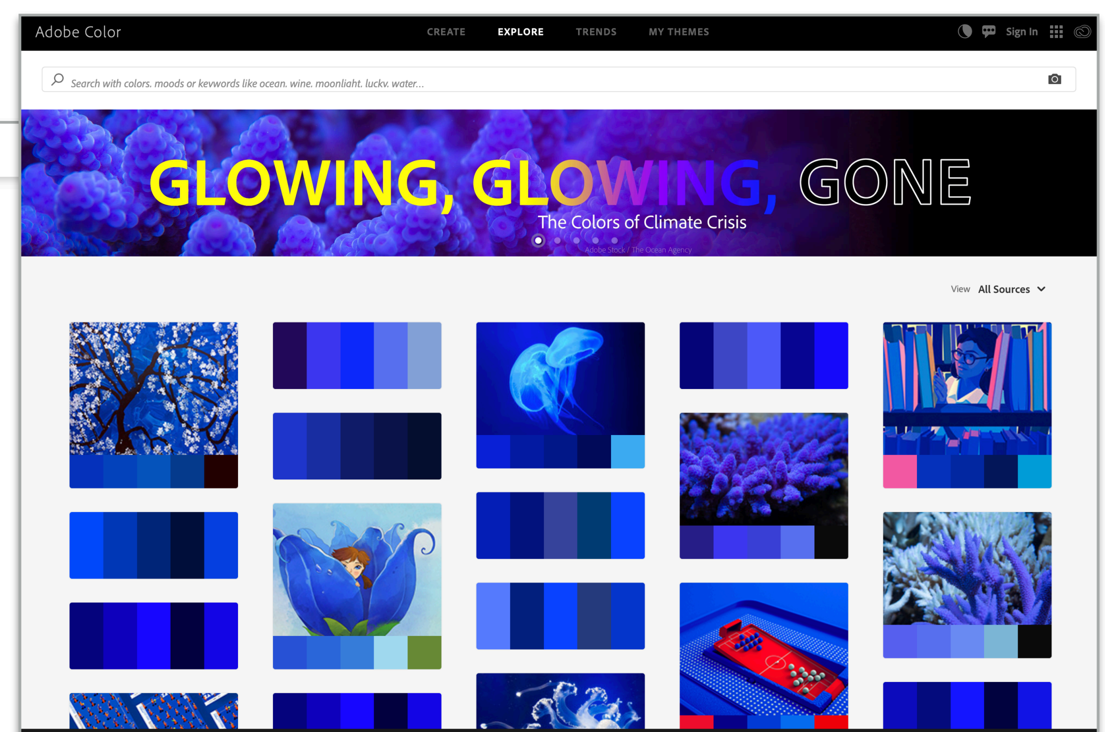


https://projects.susielu.com/viz-palette



http://vrl.cs.brown.edu/color

Color Palettes for UI design



https://color.adobe.com/explore

https://material.io/design/color/#tools-for-picking-colors

Color vision deficiency

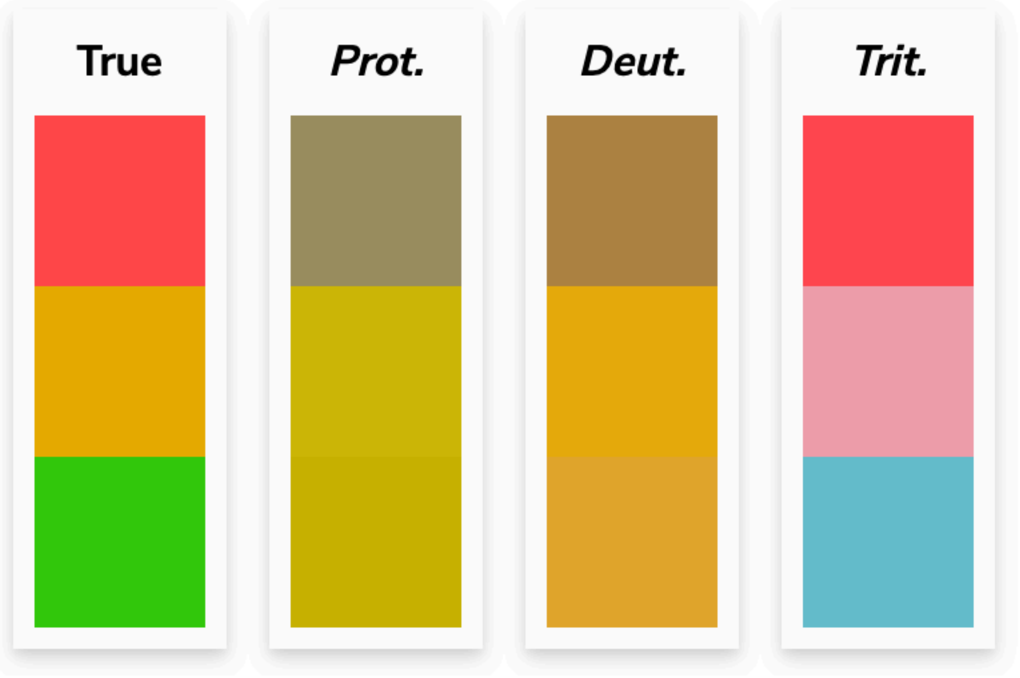
Trichromacy: response functions of cones to light waves - blue (S), green (M), red (L)

1 in 20 people suffer from some type of color vision deficiency:
protanopia, deuteranopia, or tritanopia, …

8% of men, 0.5% of women

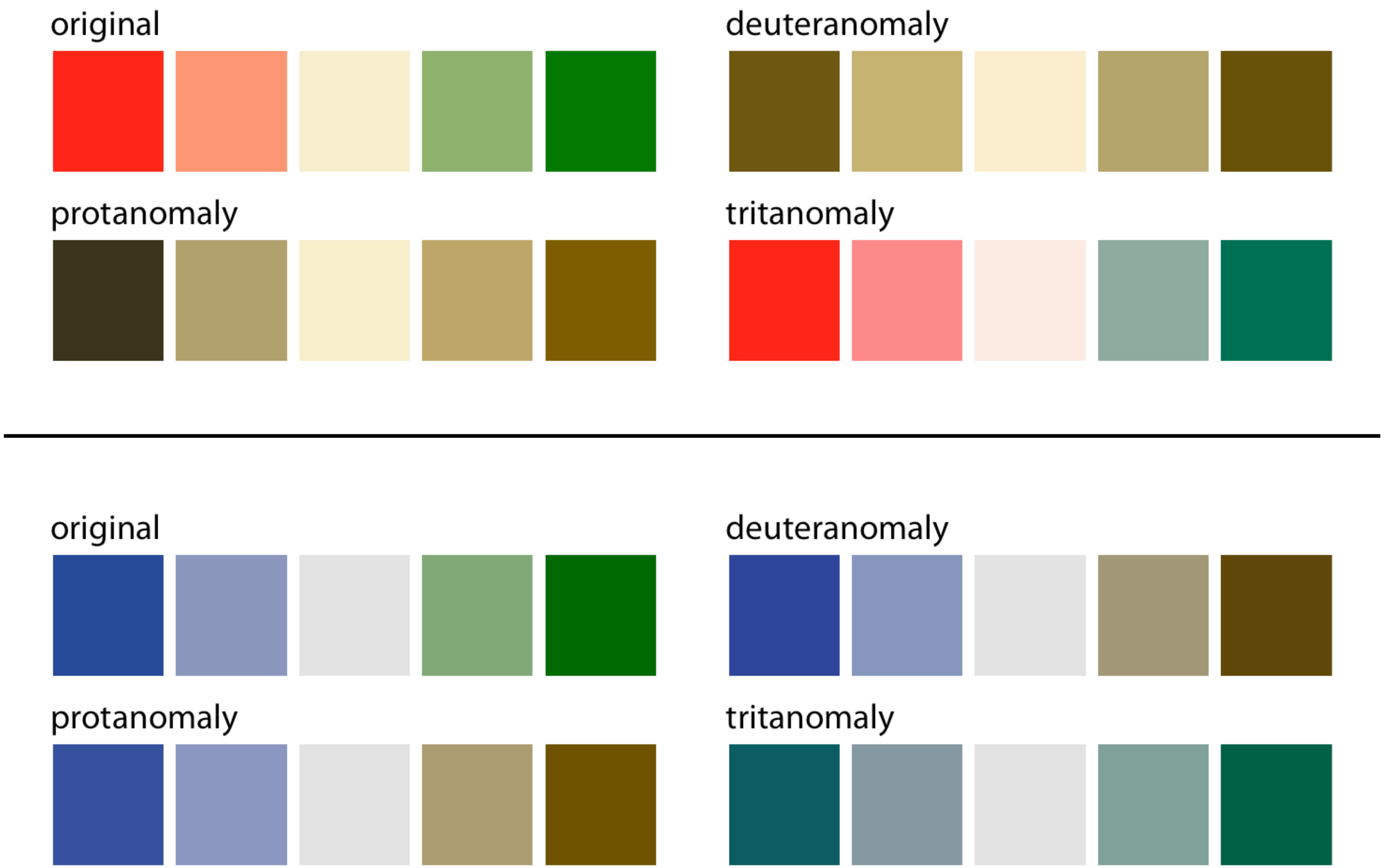
Anomalies in the cone cells, which are responsible for detecting primary colors (red,green,

Intentionally bad palette:



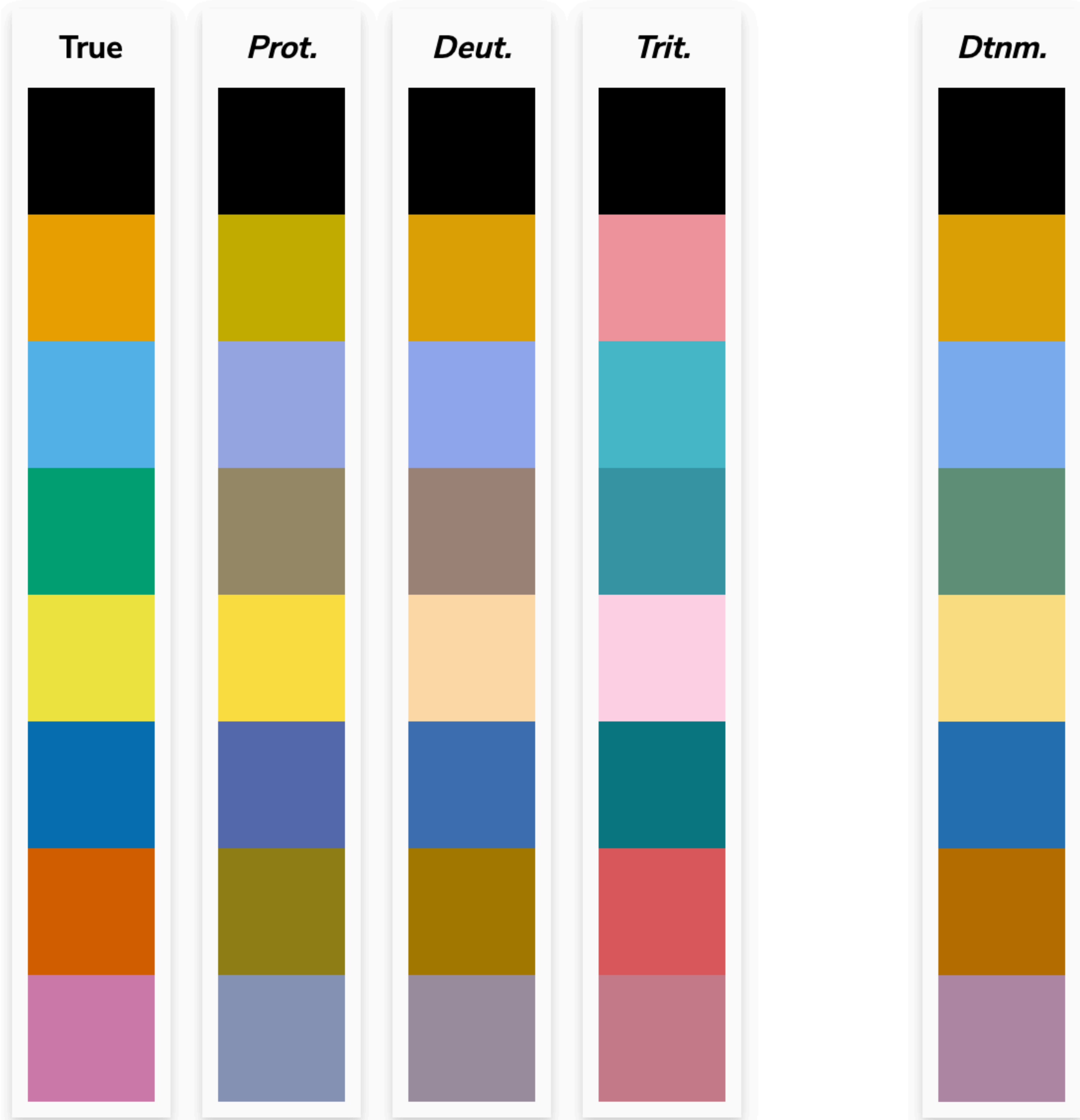
Made with https://davidmathlogic.com/colorblind/

Color vision deficiency



Color vision deficiency

Conservative 7-color palette adapted for color blindness



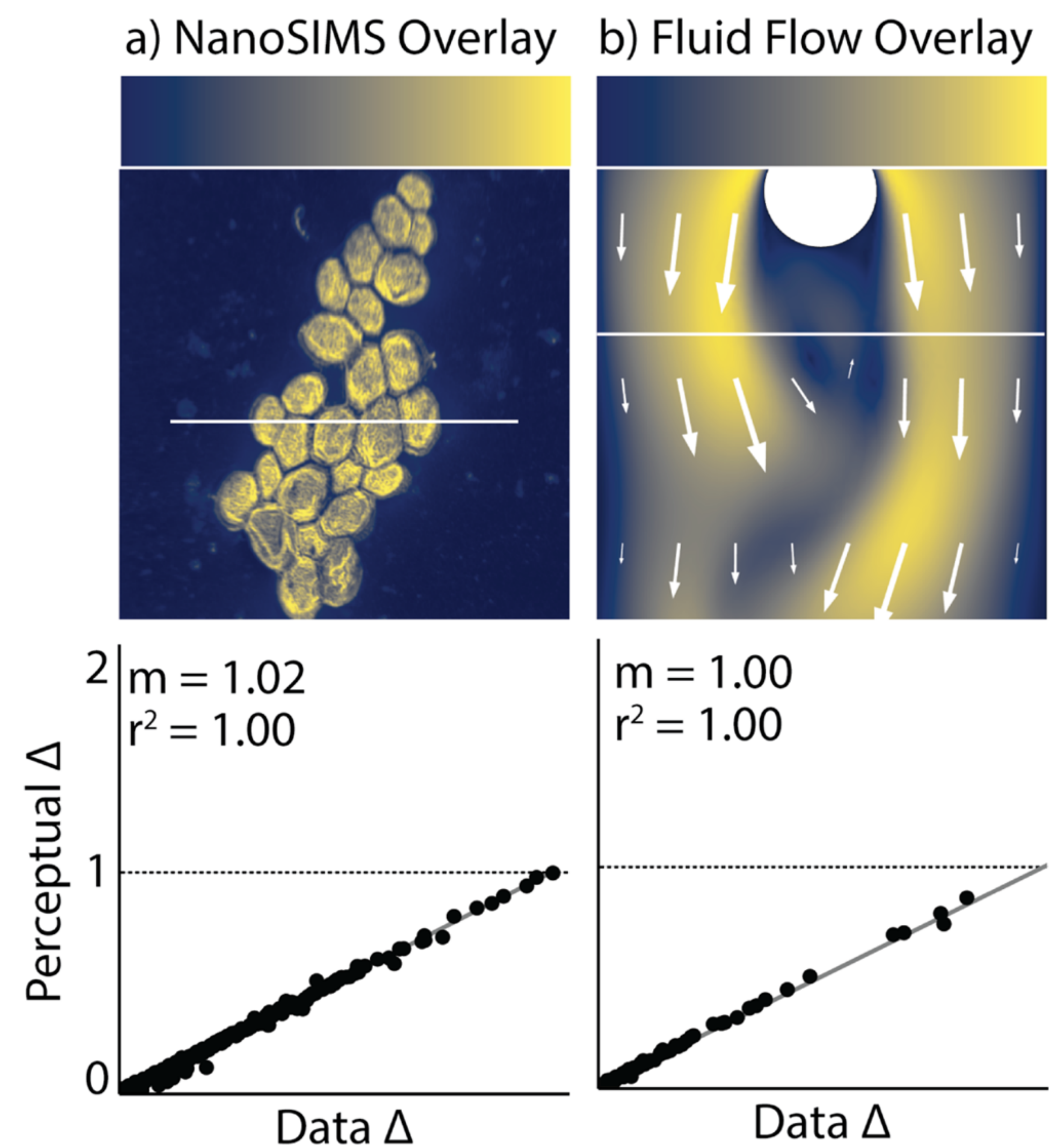
Made with https://davidmathlogic.com/colorblind/

Color hue Color saturation Color luminance

Multi-hue scheme with two colors that have a clear brightness hierarchy

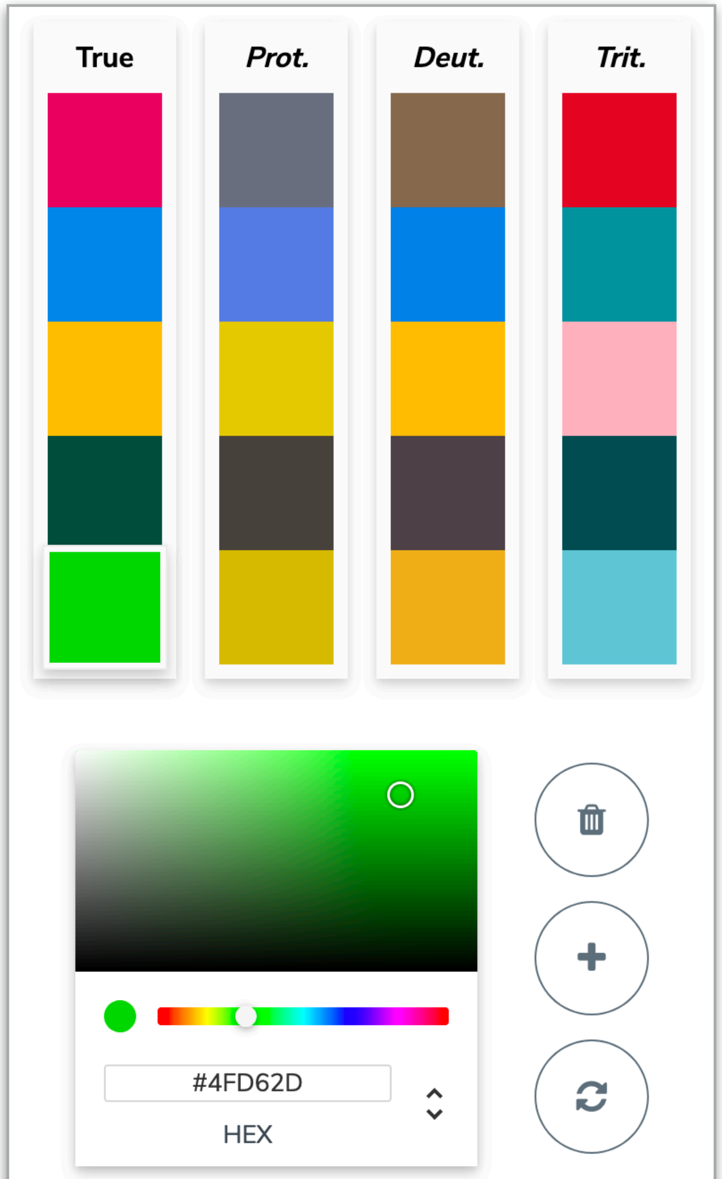
Both color and brightness vary at a steady rate through the entire scale

Perceptually consistent among people who are color blind and those with normal color vision.



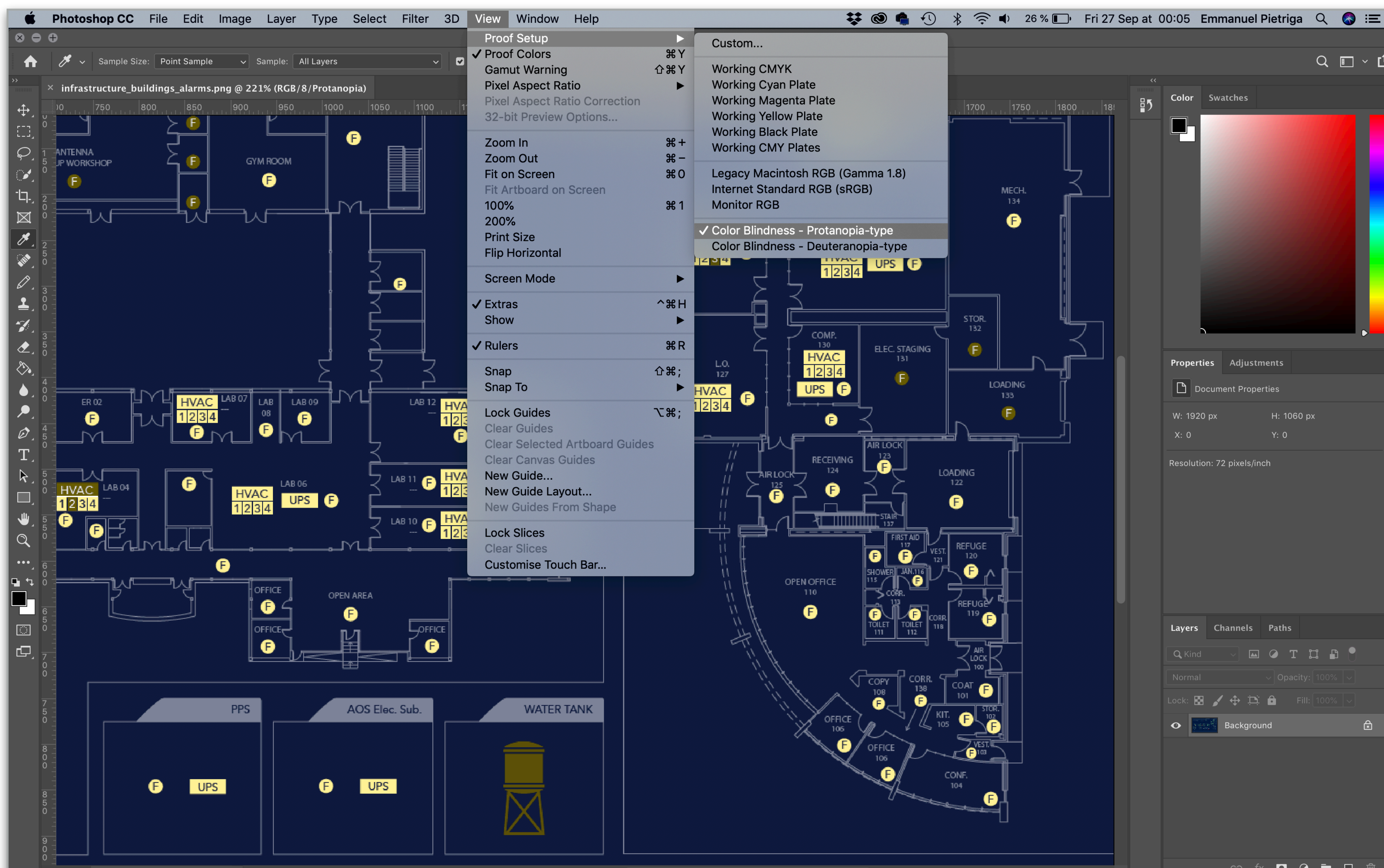
Nuñez JR, Anderton CR, Renslow RS (2018) Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. PLOS ONE 13(7): e0199239. https://doi.org/10.1371/journal.pone.0199239

Color vision deficiency



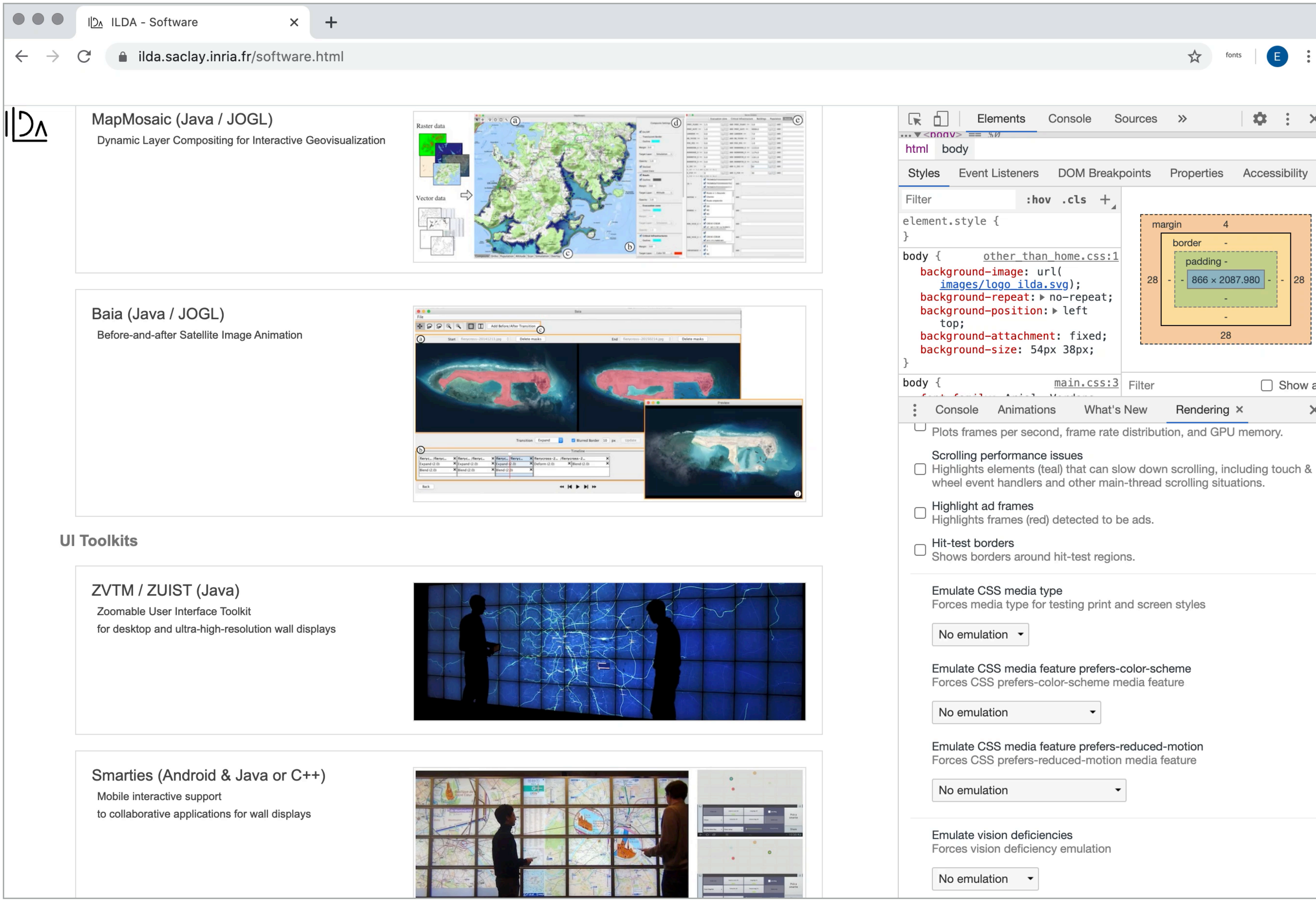
Early design phase, use interactive tools like
https://davidmathlogic.com/colorblind/

In later stages, tools such as Adobe Photoshop are convenient



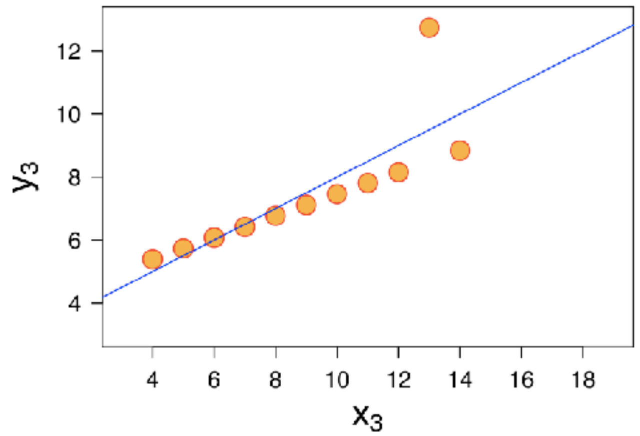
Color vision deficiency

Live simulation in Chrome

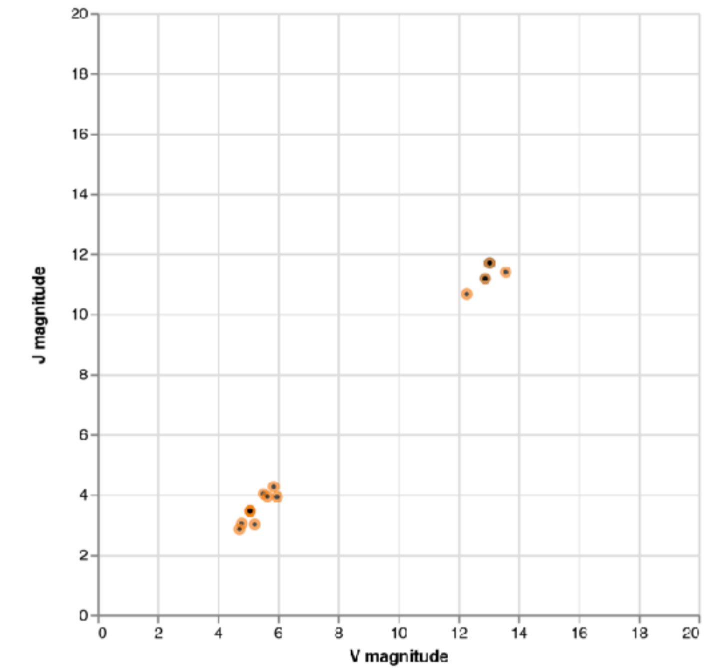


Position

- Selective: ✓



- Associative: ✓



- Ordered: ✓

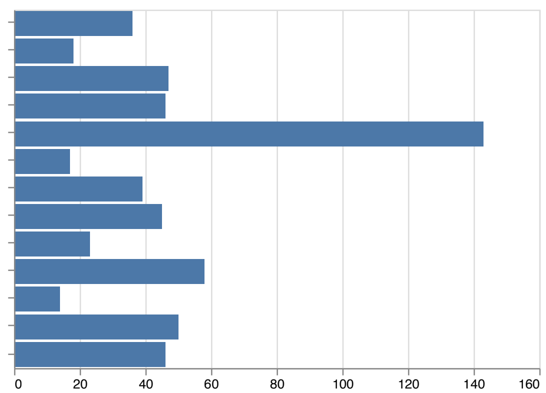
- Quantitative: ✓

- Discriminability (resolution): ✓
  - All the way down to individual pixels
  - Can make very accurate judgments

Length (1D size)

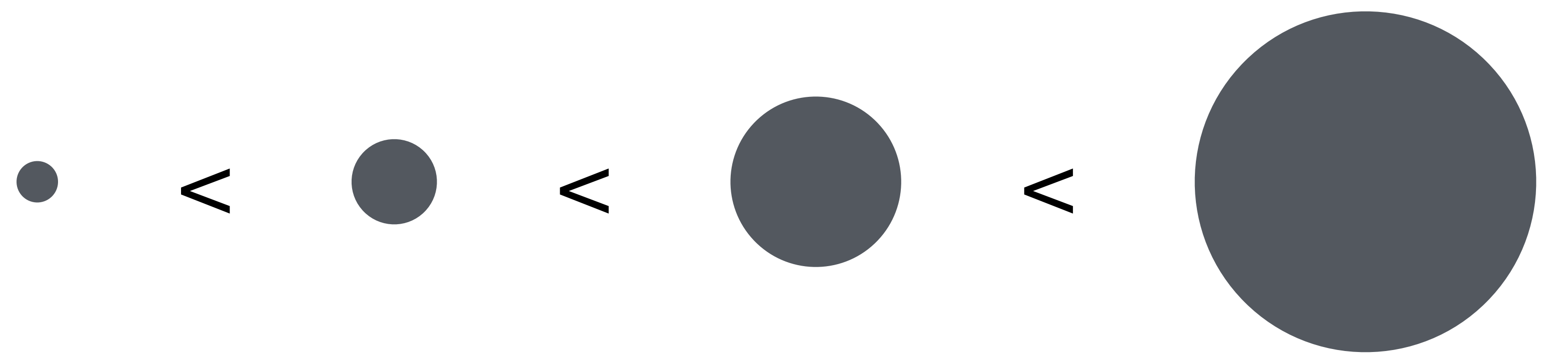
Area (2D size)

- Selective: ✓



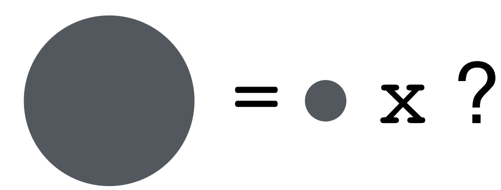
- Associative: ✓

- Ordered: ✓



- Quantitative: ✓

relative judgement
quadratic for areas (amplifies differences)



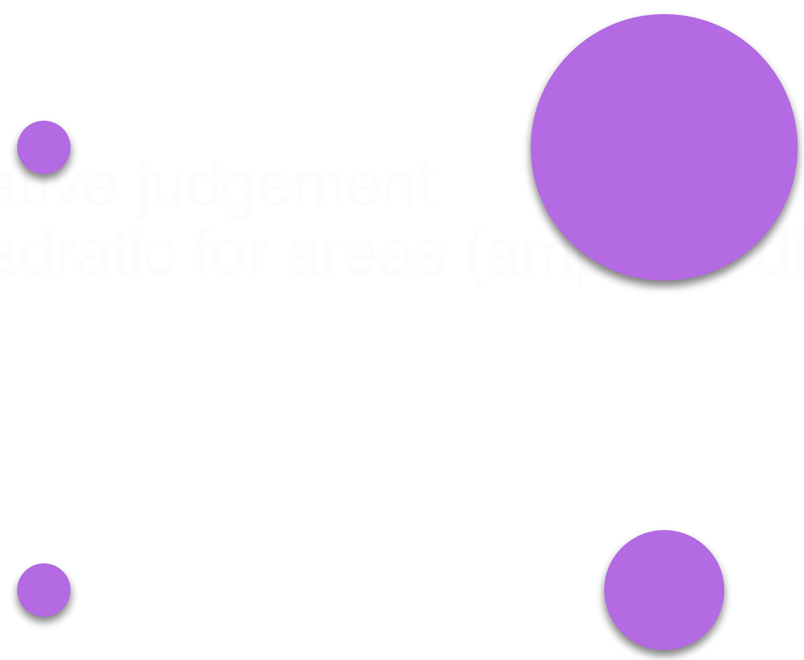
- Discriminability (resolution):
  - in theory: ∞
  - in practice: ≲5 for selection and association, ≲20 for distinction

Length (1D size)

Area (2D size)

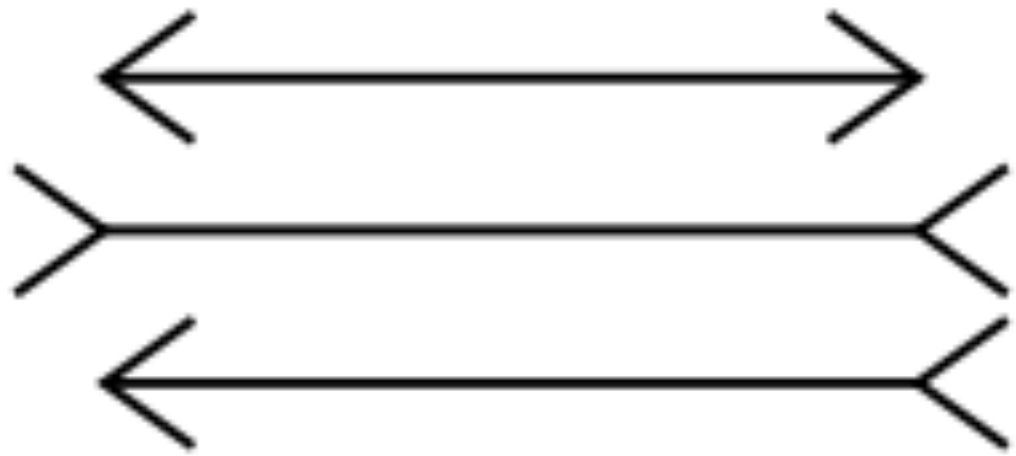
Properly mapping values to areas:

v_1 $\qquad$ $v_2 = 5 \cdot v_1$



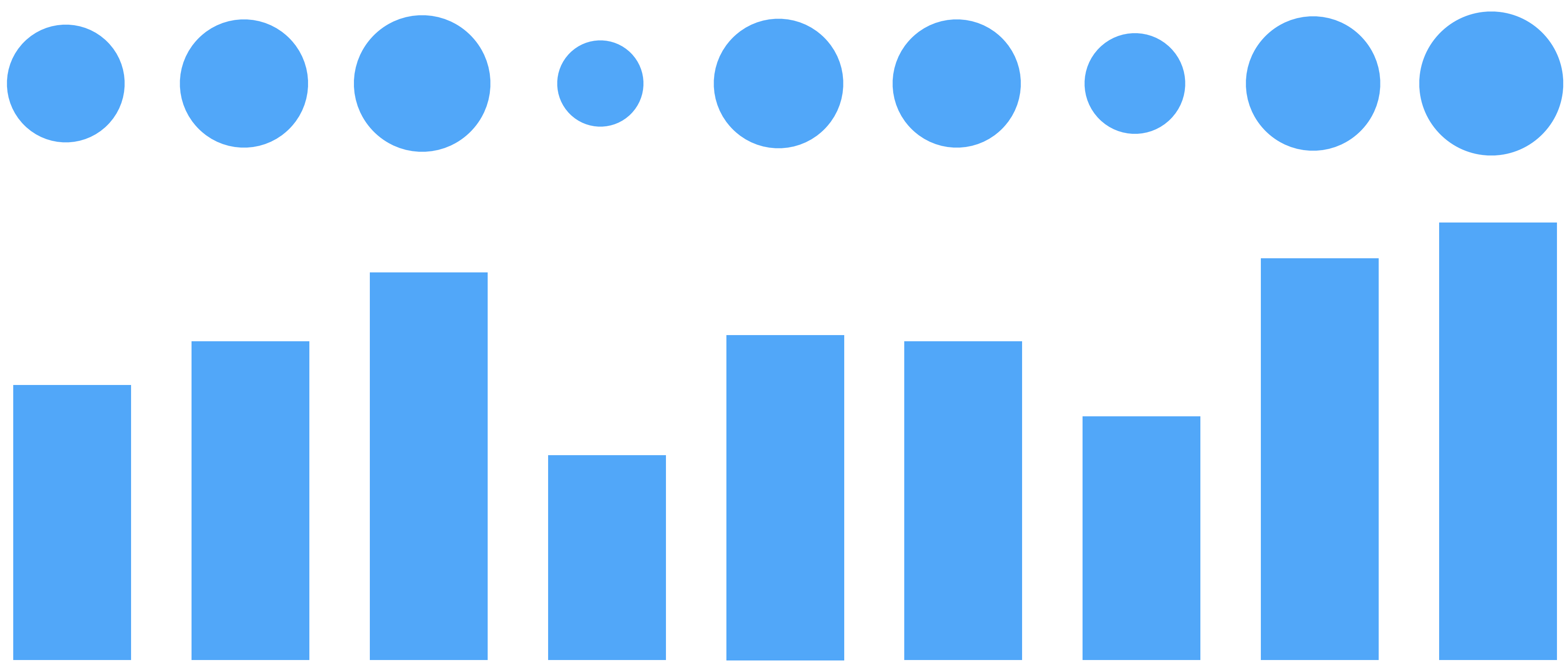
$r = v$ $\qquad$ $(= 25 \cdot)$

$r = \sqrt{v} \cdot a$ $\qquad$ $(= 5 \cdot)$



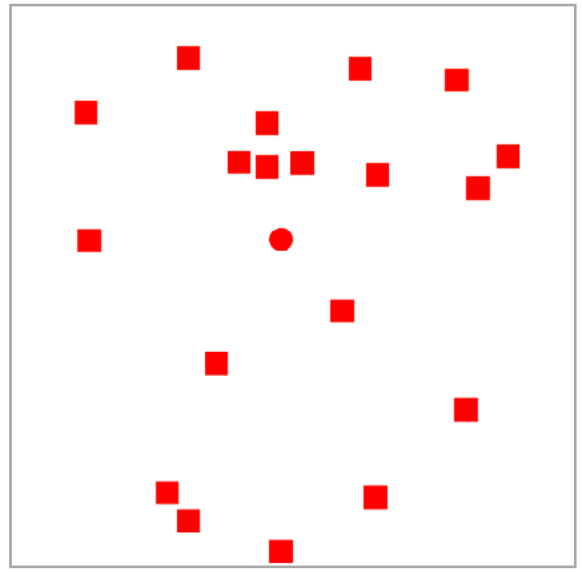
[Source: http://en.wikipedia.org/wiki/M%C3%BCller-Lyer_illusion]

Length (1D size) vs. Area (2D size)

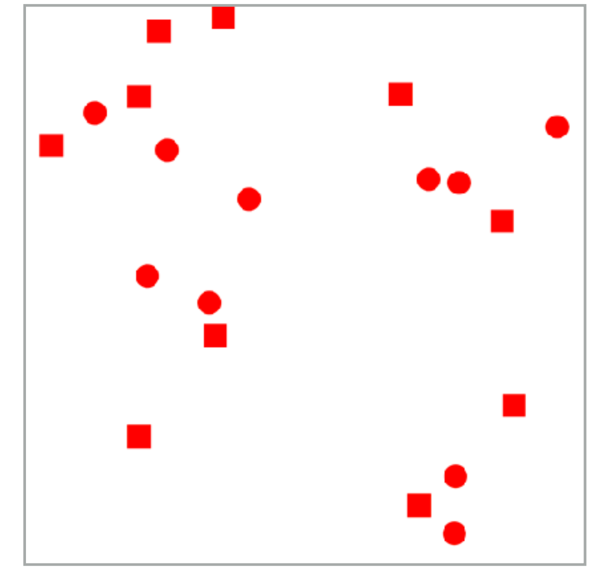


Shape 

- Selective: ✓

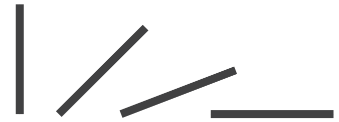


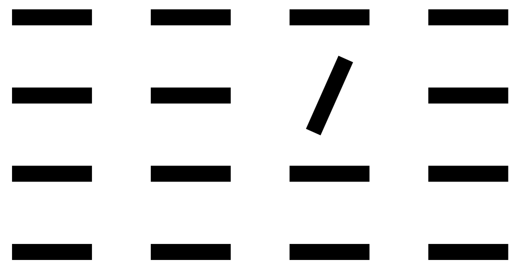
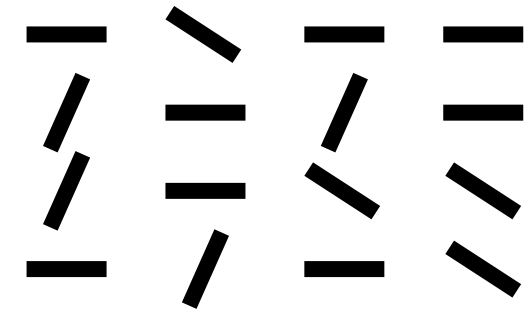
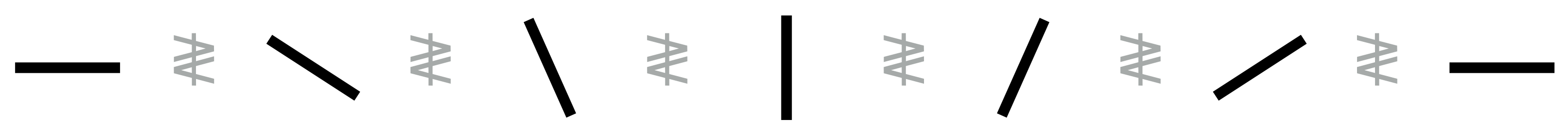
- Associative: ✓



- Ordered: ✗

- Quantitative: ✗
- Discriminability (resolution): ✓
  - in theory: ∞
  - in practice: a lot (enough, for most practical purposes)

Tilt/angle 

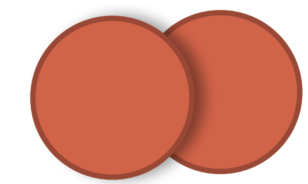
- Selective: ✓ 
- Associative: ✓ 
- Ordered: ✕

- Quantitative: ✕

most of the time…

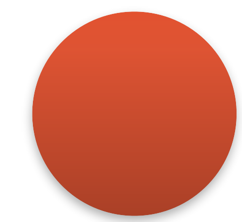
- Discriminability (resolution): ✓
  - ~ 5 in 2D
  - ? in 3D

Additional variables

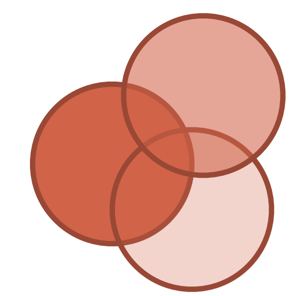
- Depth (with shadows)



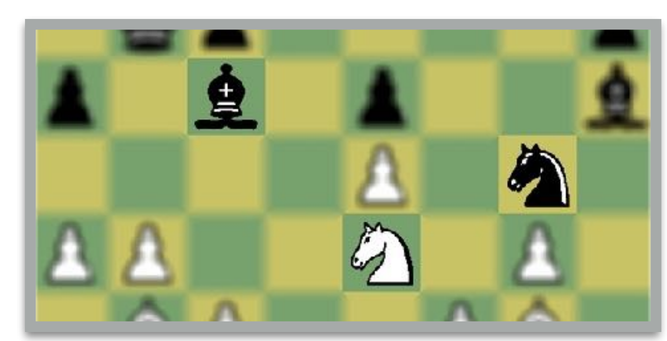
- Illumination (with gradients)



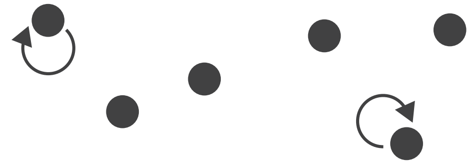
- Transparency (with alpha channel / compositing)
  - Can interfere with other color-based encodings
  - But also a great way to handle overlapping objects / dense regions



- Blur (sharpness)
  - Can be useful to guide attention
  - Uncertainty?

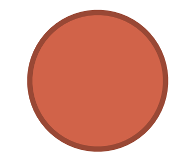


[Image Source: R. Kosara, Eagereyes, Feb 2011]

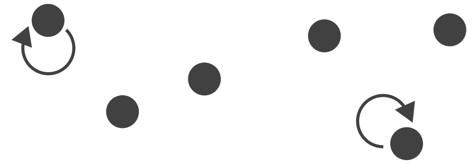


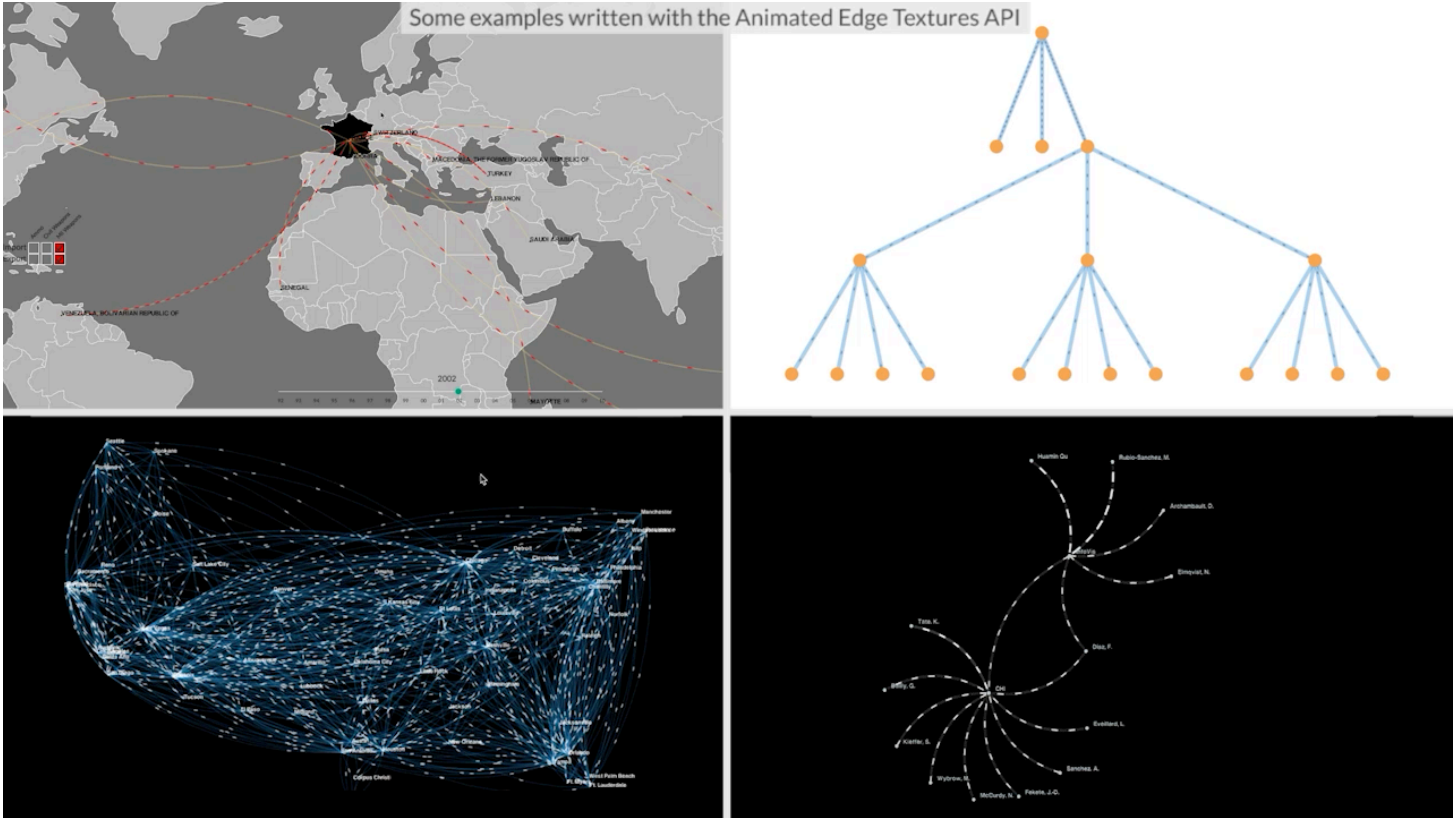
… and other types of animation

- Motion
  - direction, acceleration, speed, frequency, trajectory (e.g. wiggling to represent uncertainty)



- Flicker
  - frequency, rhythm, appearance/disappearance

Motion  … and other types of animation



[H. Romat et al., Animated Edge Textures in Node-Link Diagrams: a Design Space and Initial Evaluation, CHI '18: Proceedings of the 36th SIGCHI conference on Human Factors in computing systems, pages 187:1--187:13, April 2018]

Characterizing visual variables

From Bertin

- Selective (related to visual popout):
  Can this variable allow us to spontaneously differentiate/isolate items from groups?

- Associative:
  Can this variable allow us to spontaneously put items in a group?

- Ordered:
  Can this variable allow us to spontaneously perceive an order?

- Quantitative:
  Is there a numerical reading obtainable from changes in this variable?

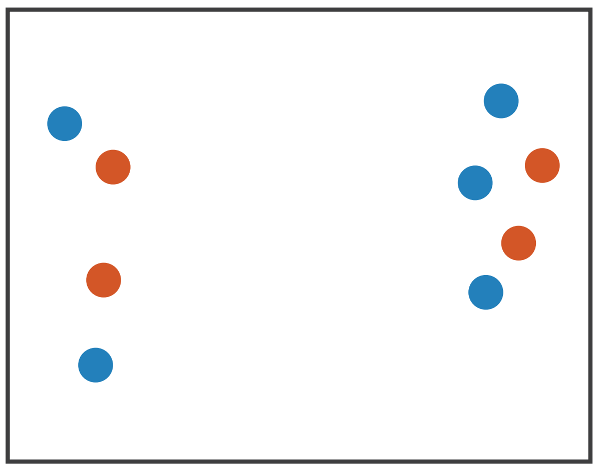
- Length, or rather, Discriminability (resolution):
  Across how many changes in this variable are distinctions possible?

- Accuracy:
  Can we make accurate judgments about this variable in terms of perception?

- Separability vs. Integrality:
  Are two (or more) variables interfering with each other?

Separability vs. Integrality, Selective Attention

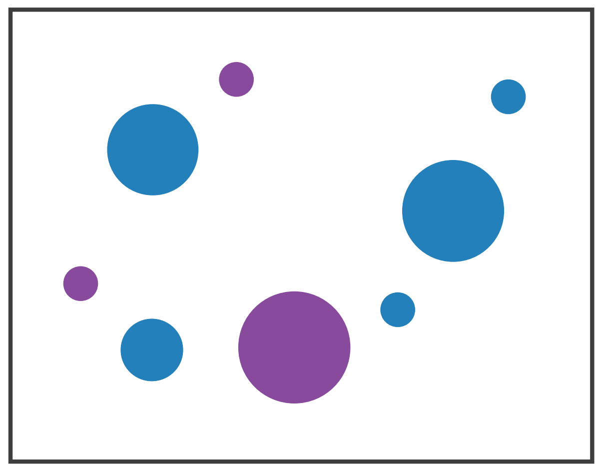
Position
+ Hue (Color)



Fully separable

2 groups each

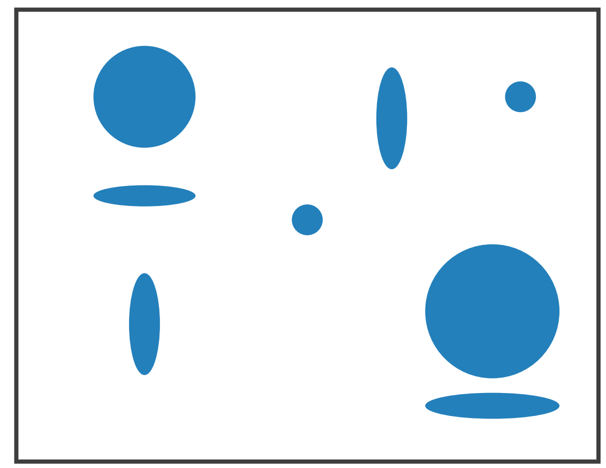
Size
+ Hue (Color)



Some interference

2 groups each

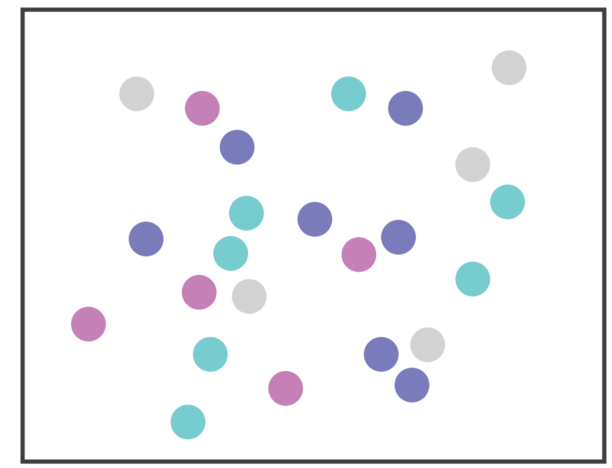
Width
+ Height



Some/significant interference

3 groups total: integral area

Red
+ Green

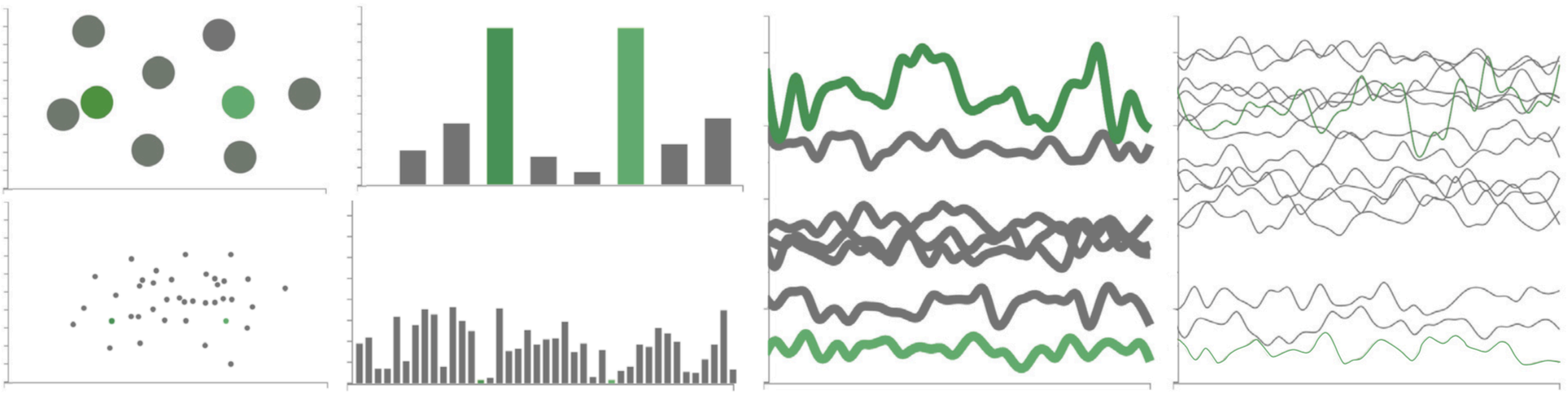


Major interference

4 groups total: integral hue

[Source: adapted from InfoVis course, T. Munzner, 2017]

[Image source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]



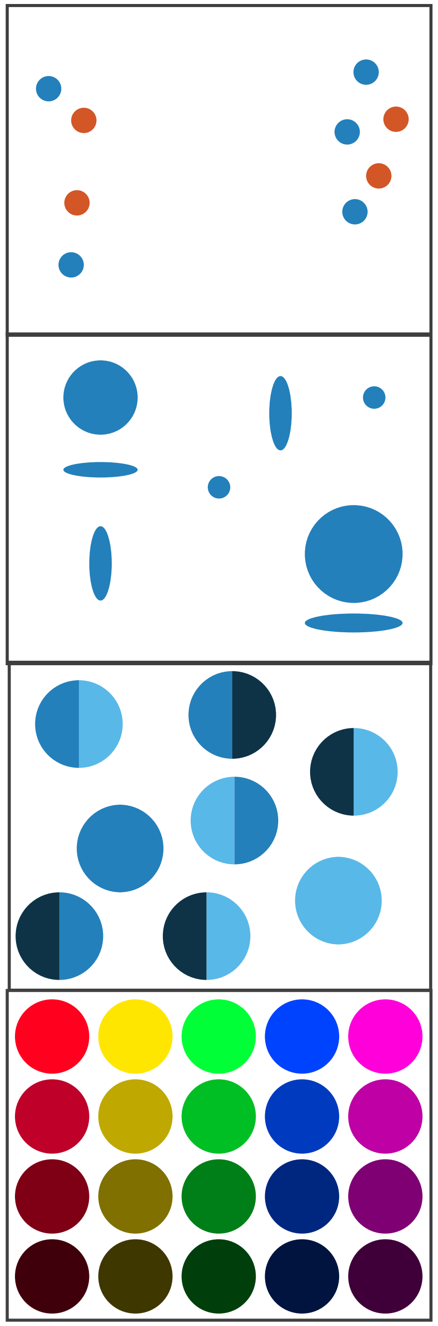
[Danielle A. Szafir. 2018. Modeling color difference for visualization design. IEEE Trans. on Visualization and Comp. Graphics 24, 1 (2018), 392–401.]

Selective Attention and Emergent Features

Ability to attend to one stimulus while ignoring the confounding influence of others

Emergent features arise from the combination of two visual variables

Separable combination	minimal perceptual interference
Integral combination	attention to individual variables is inhibited
Configural combination	features arise when the combined variables are congruent
Asymmetrical combination	interaction effects between the combined variables



[top 2 images source: Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014]