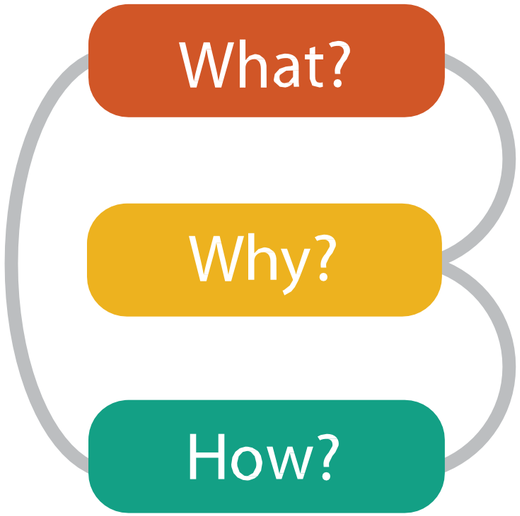
Paul Rosen
paul.rosen@utah.edu
@paulrosenphd
https://cspaul.com

Visualization for Data Science
DS-4630 / CS-5630 / CS-6630

VISUAL ENCODING

analysis: what, why, and how

- what is shown?
- why is the user looking at it?
- how is it shown?



- abstract vocabulary avoids domain-specific terms
- what-why-how analysis framework as scaffold to think systematically about design space

visual encoding

How?

Encode

- Arrange
  - Express
  - Separate
  - Order
  - Align
  - Use

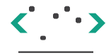
- Map
  from categorical and ordered attributes
  - Color
    - *Hue*
    - *Saturation*
    - *Luminance*
  - Size, Angle, Curvature, ...
  - Shape
  - Motion
    Direction, Rate, Frequency, ...

Manipulate

- Change

- Select

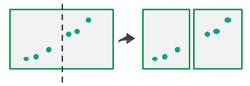
- Navigate



Facet

- Juxtapose

- Partition



- Superimpose

Reduce

- Filter

- Aggregate

- Embed

What?

Why?

How?

Visual Encoding

- marks and channels
- planar position
- time

MARKS

- graphical element in an image
- classified according to number of spatial dimensions required

points (0D) | lines (1D) | areas (2D)

marks

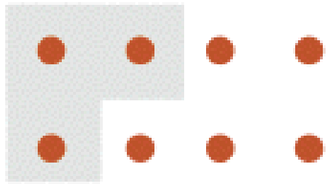
marks as nodes (items)

points (0D)

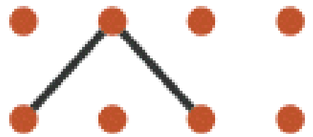
lines (1D)

areas (2D)

marks as links



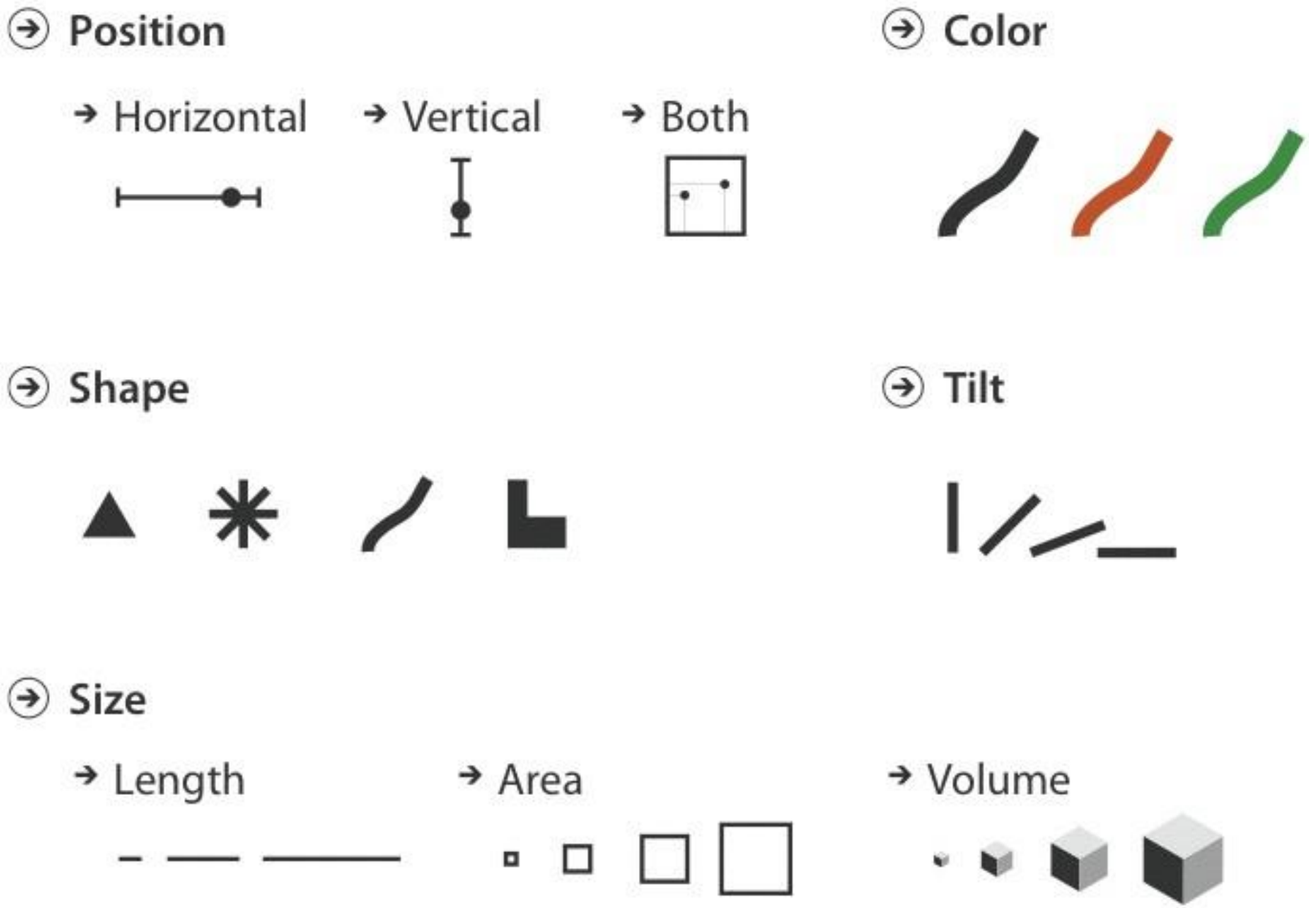
containment



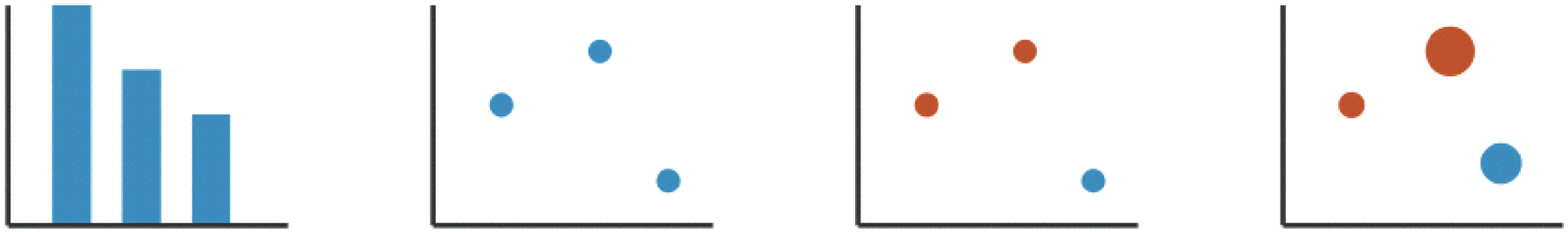
connection

CHANNELS

- parameters that control the appearance of marks

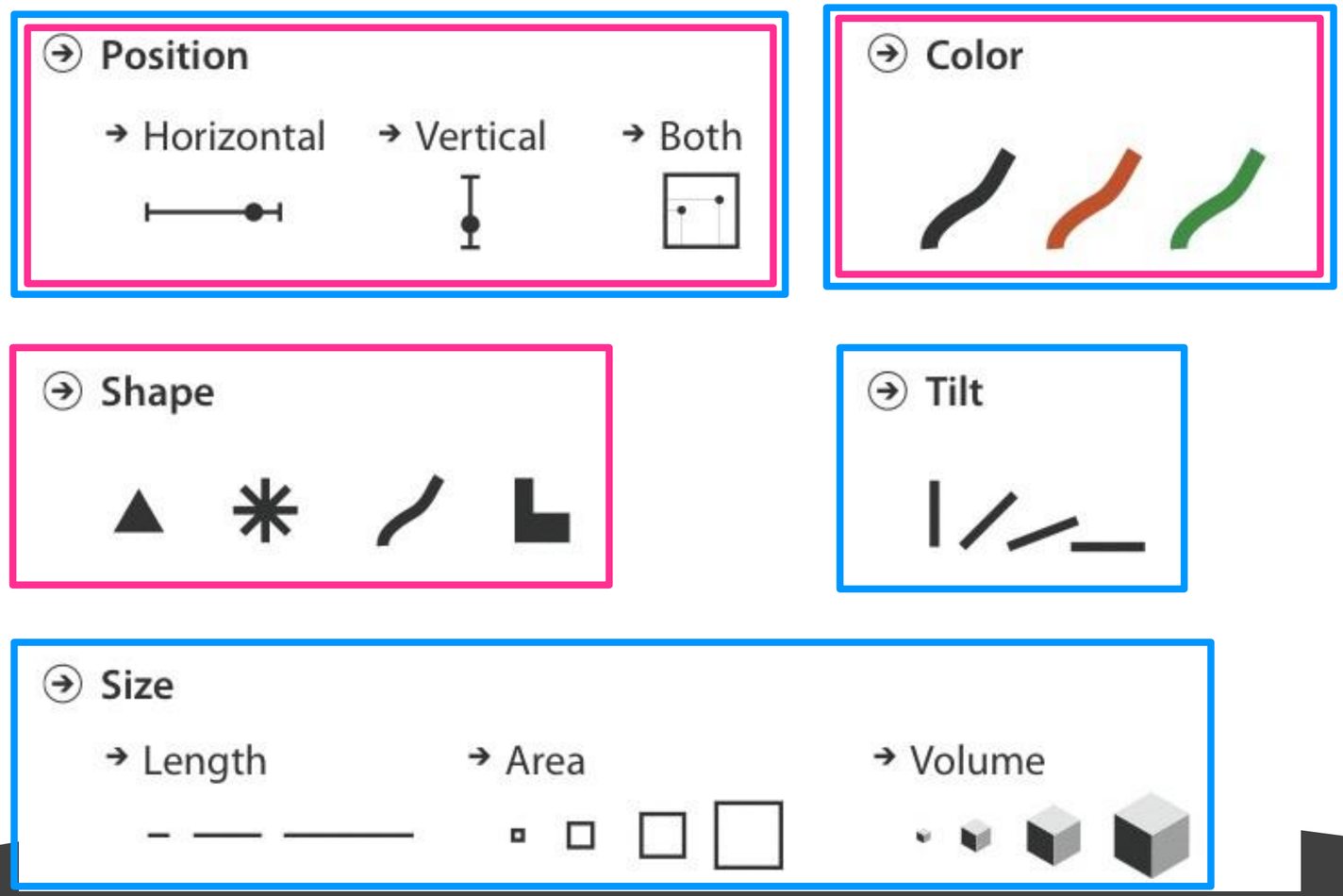


name that mark and channel

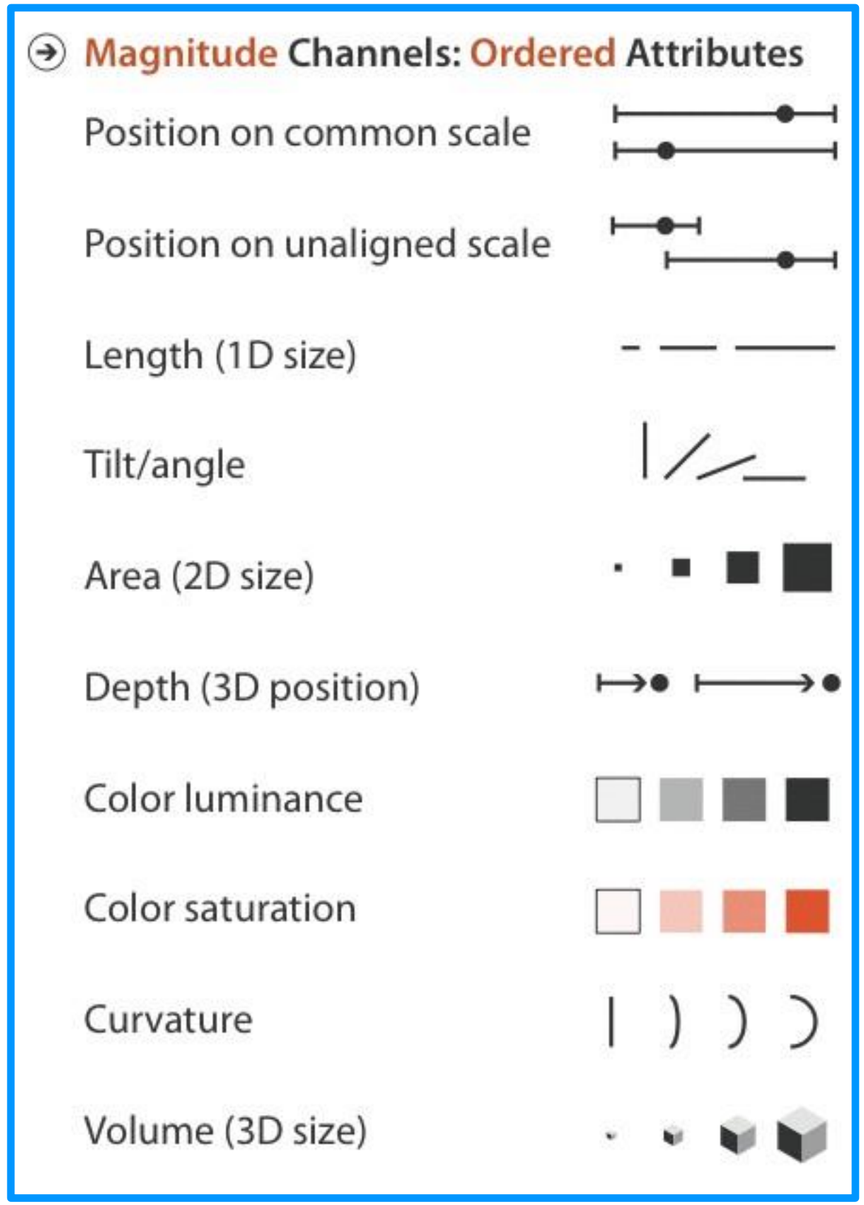


CHANNEL TYPES

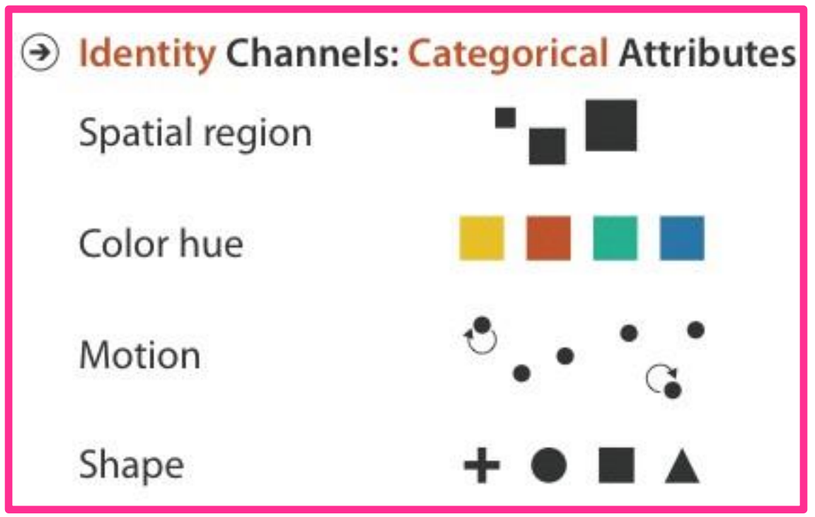
identity (what or where)
magnitude (how much)



expressiveness & effectiveness

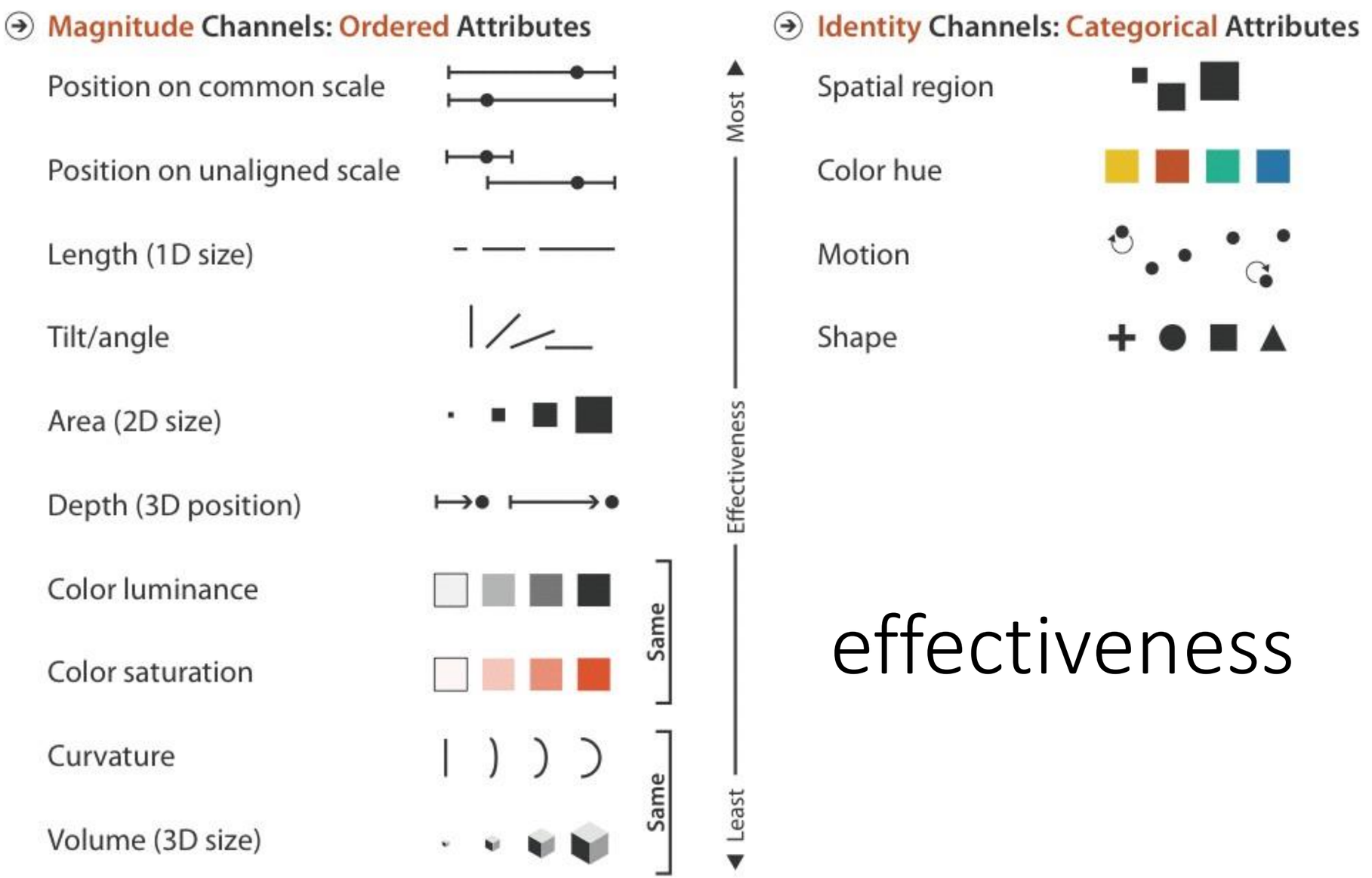


(how much)



(what or where)

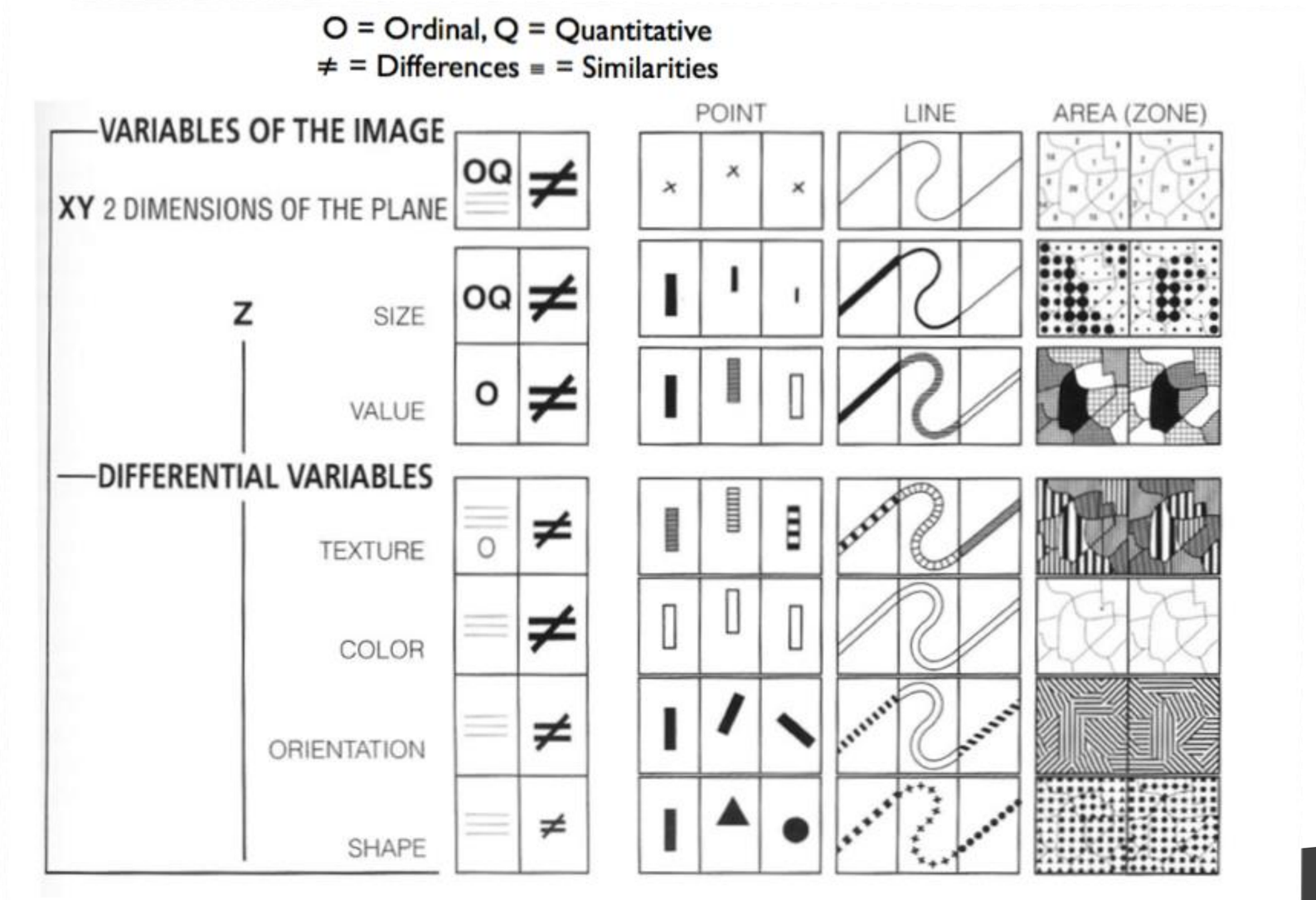
expressiveness



effectiveness

WHERE DO RANKINGS COME FROM?

Bertin, “Semiology of Graphics”, 1967



Cleveland & McGill, “Graphical Perception and Graphical Methods for Analyzing Scientific Data”, 1985

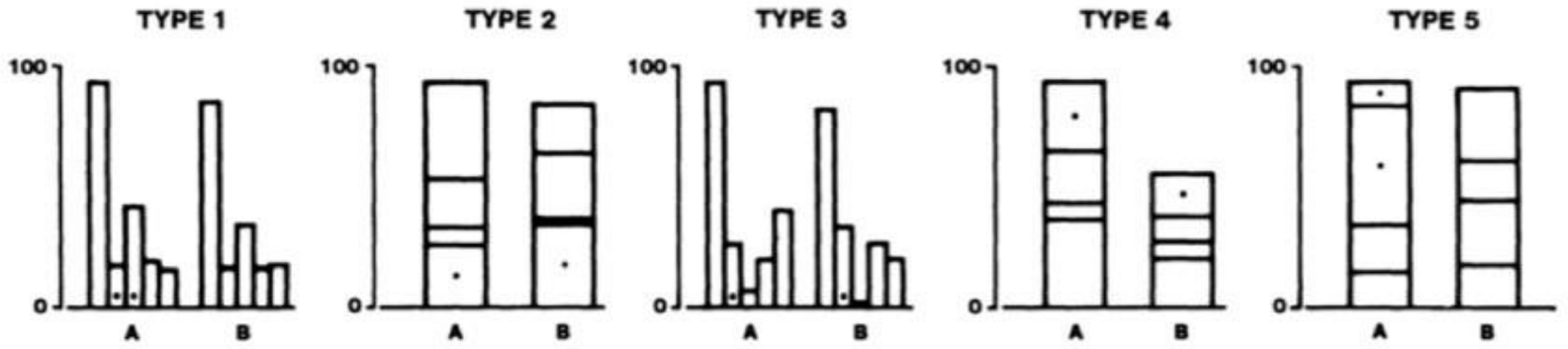


Figure 4. Graphs from position–length experiment.

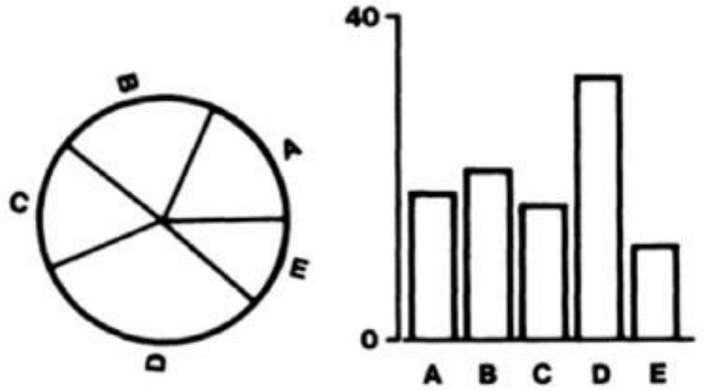
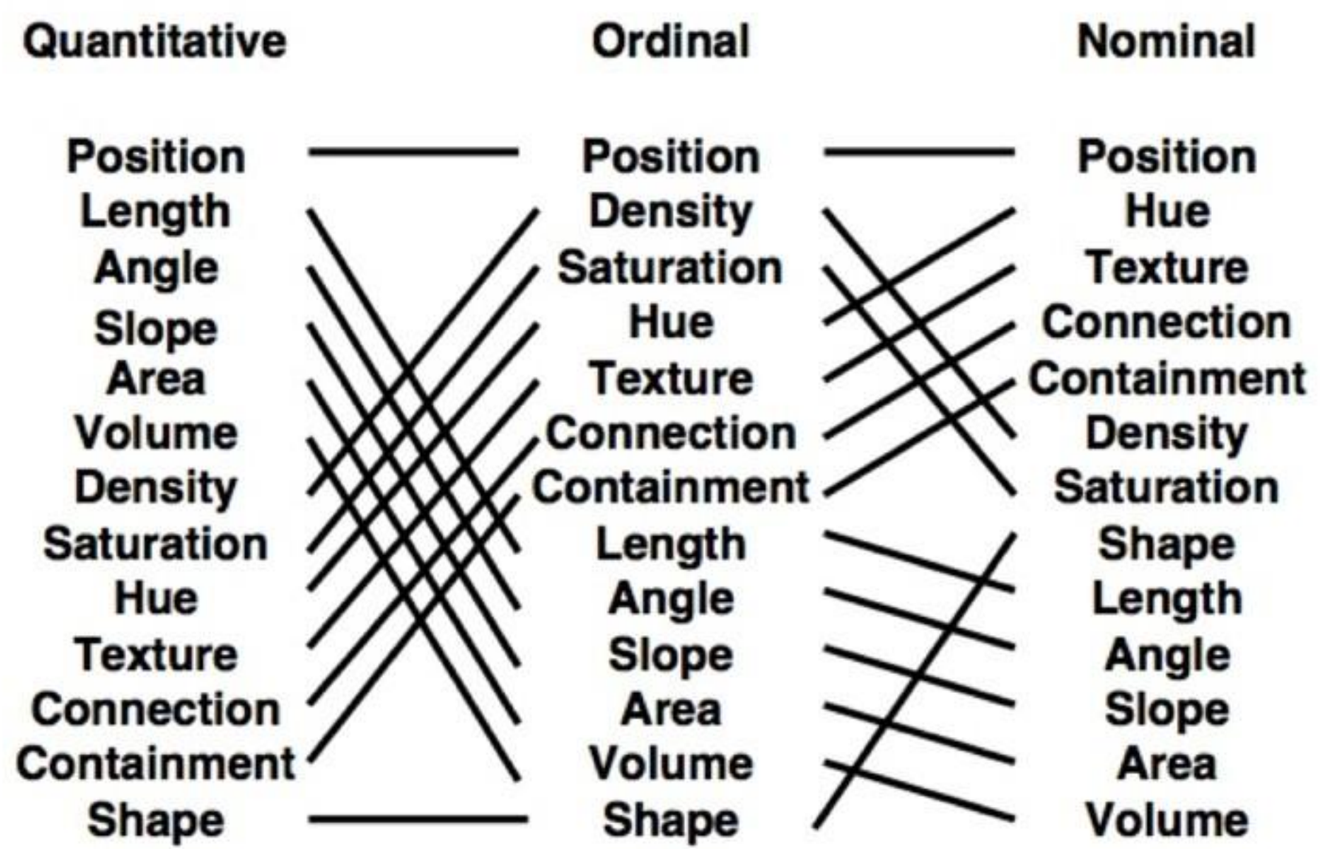


Figure 3. Graphs from position–angle experiment.

Mackinlay, “Automating the Design of Graphical Presentations of Relational Information”, 1986



Heer & Bostock, “Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design”, 2010

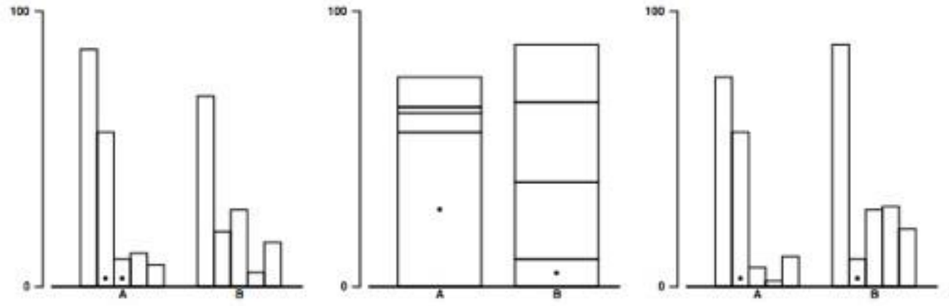


Figure 1: Stimuli for judgment tasks T1, T2 & T3. Subjects estimated percent differences between elements.

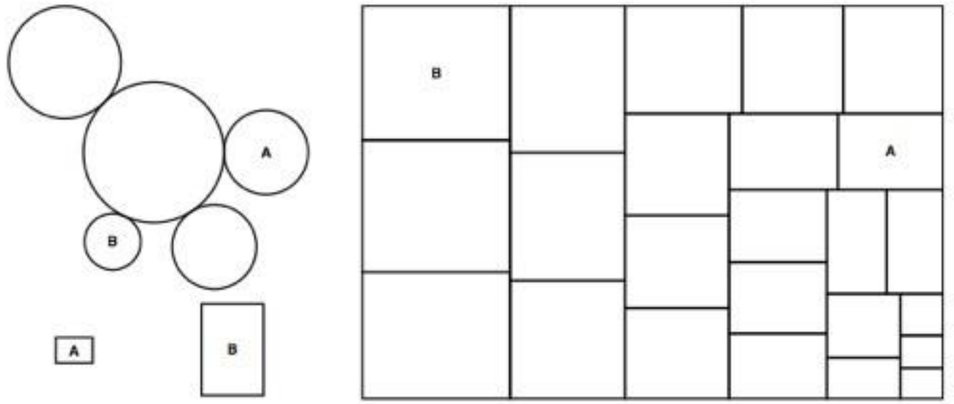
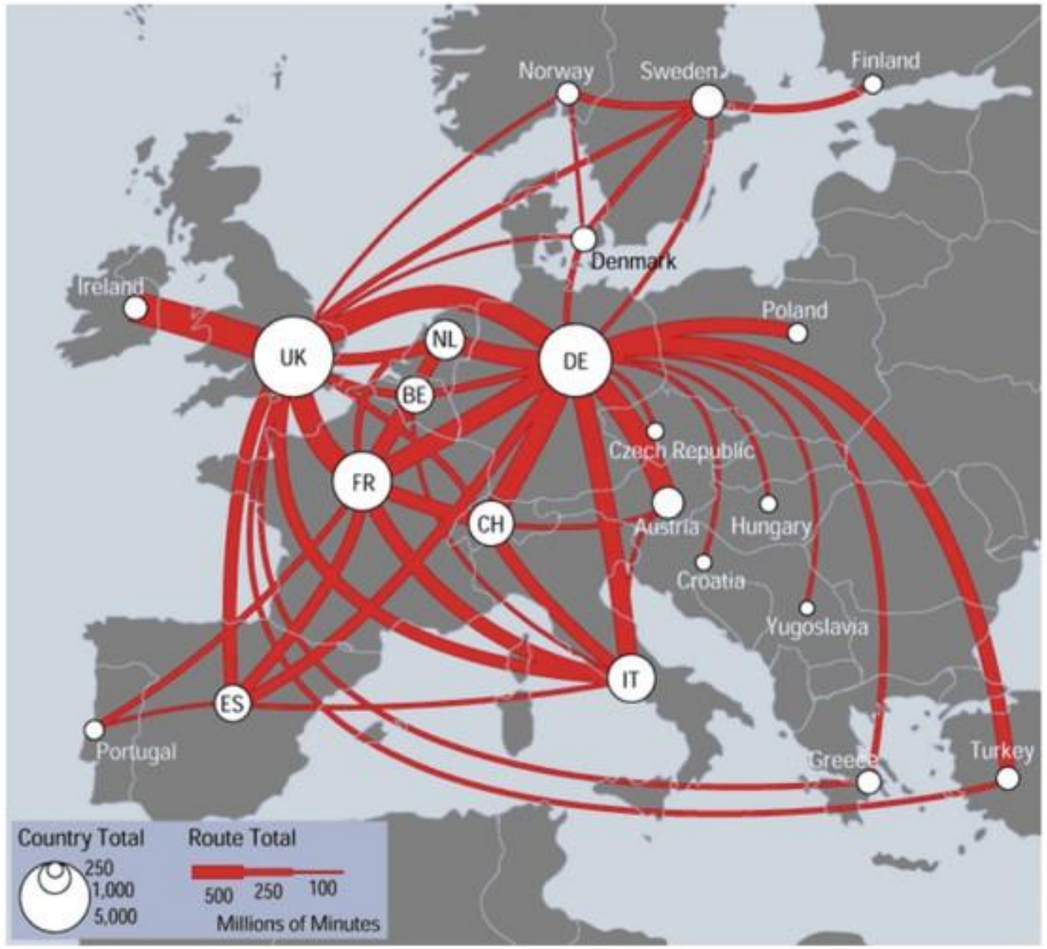


Figure 2: Area judgment stimuli. Top left: Bubble chart (T7), Bottom left: Center-aligned rectangles (T8), Right: Treemap (T9).

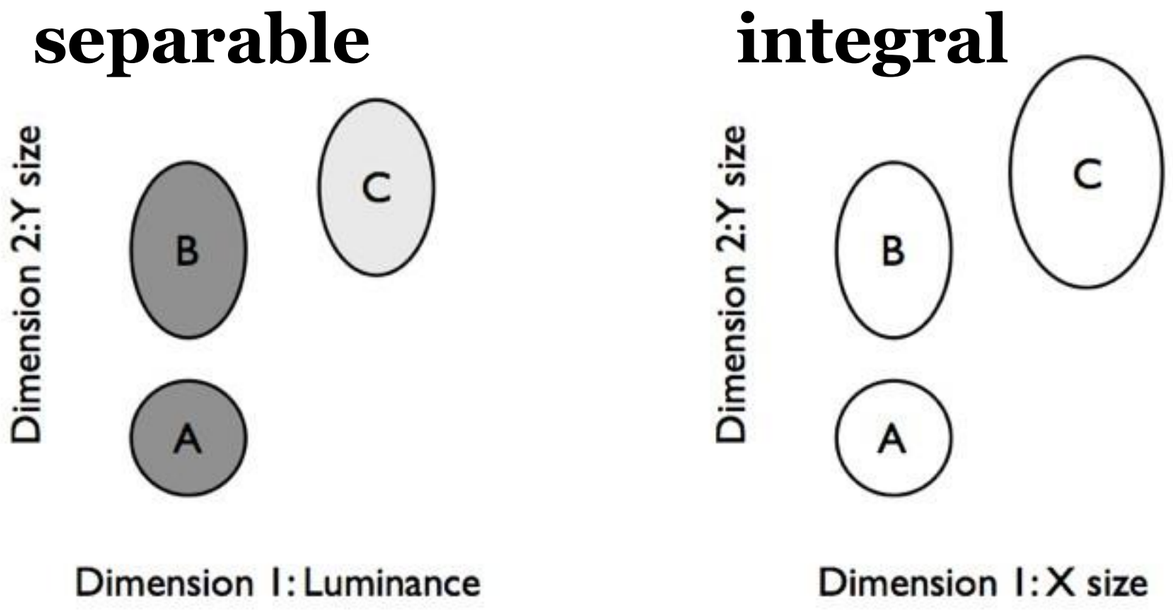
DISCRIMINABILITY

- can channel differences be discerned?



SEPARABLE vs INTEGRAL

- separable: can judge each channel individually
- integral: two channels are viewed holistically



- Ware 2004

SEPARABLE vs INTEGRAL

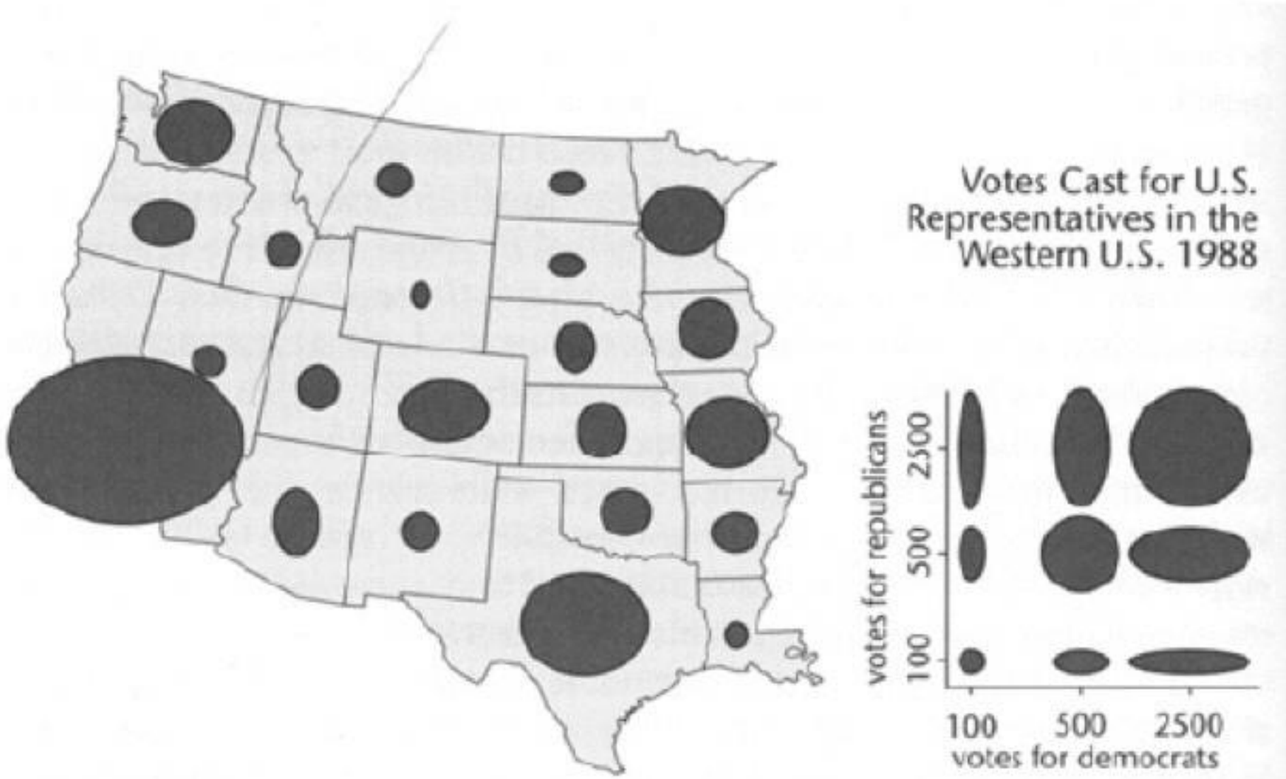
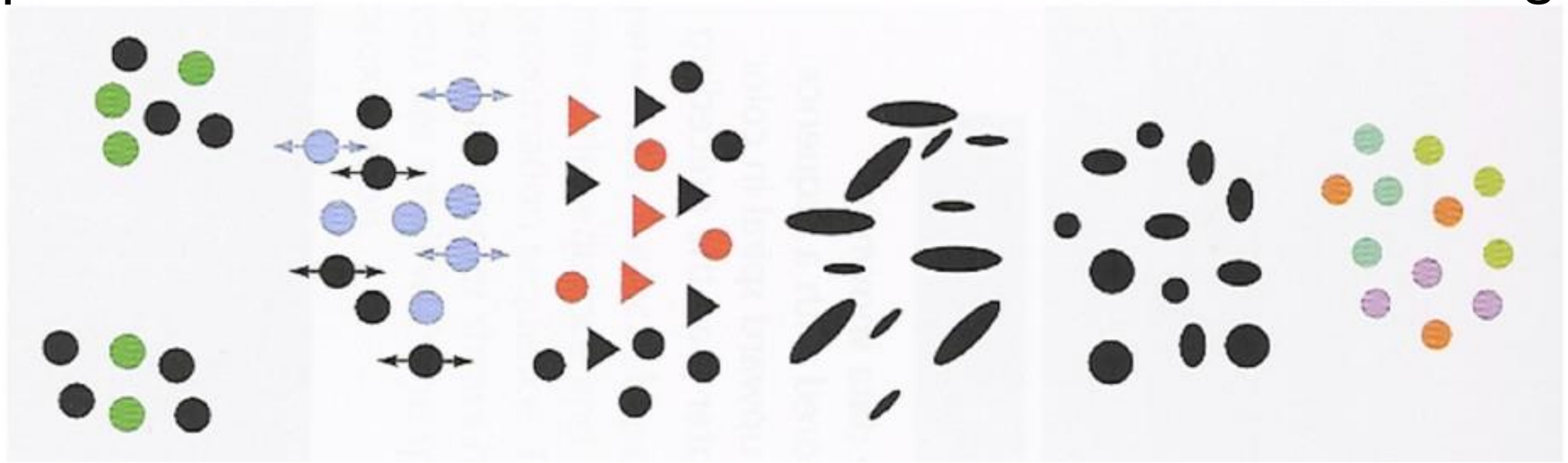


FIGURE 3.38. An example of the use of an ellipse as a map symbol in which the horizontal and vertical axes represent different (but presumably related) variables.

SEPARABLE vs INTEGRAL

separable ⟷ integral



color | location

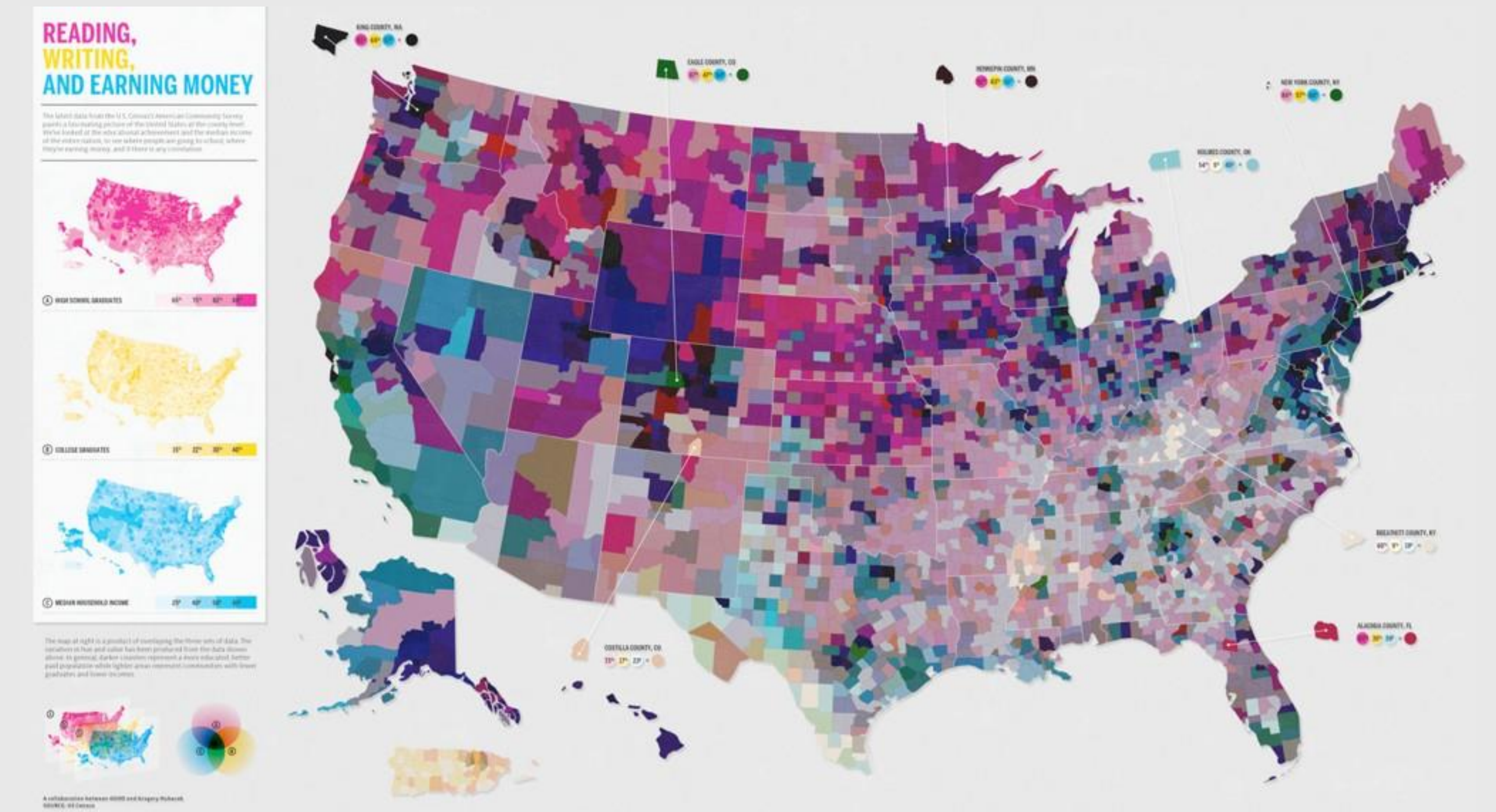
color | motion

color | shape

size | orientation

x-size | y-size

red-green | yellow-blue



http://www.good.is/post/america-s-richest-counties-and-best-educated-counties/

Encoding semantics

Graphical Code	Semantics
Small shapes defined by closed contour, texture, color, shaded solid.	Object, idea, entity, node.
Spatially ordered graphical objects.	Related information or a sequence. In a sequence the left-to-right ordering convention borrows from the western convention for written language.
Graphical objects in proximity.	Similar concepts, related information.
Graphical objects having the same shape, color, or texture.	Similar concepts, related information.
Size of graphical object Height of graphical object.	Magnitude, quantity, importance.
Shapes connected by contour.	Related entities, path between entities.
Thickness of connecting contour.	Strength of relationship.
Color and texture of connecting contour.	Type of relationship.
Shapes enclosed by a contour, or a common texture, or a common color.	Contained entities. Related entities.
Nested regions, partitioned regions.	Hierarchical concepts.
Attached shapes.	Parts of a conceptual structure.

+ perceptual effects (several of which we already discussed)

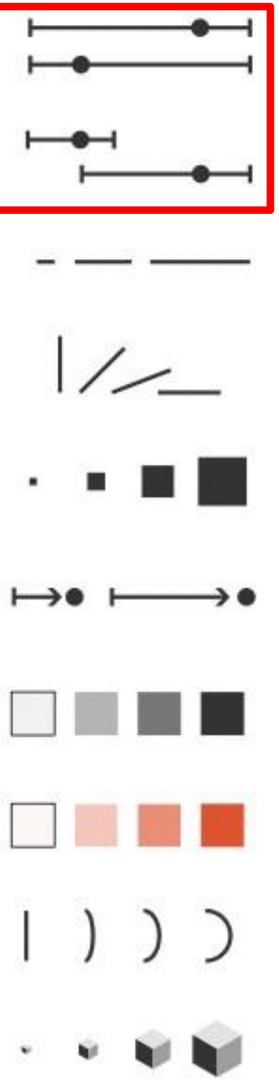
- pop-out
- steven’s power law
- weber’s law
- gestalt principles

planar position

- what‘s so special about the plane?

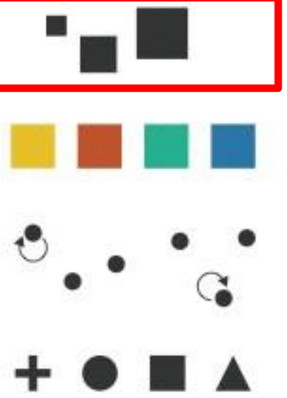
Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

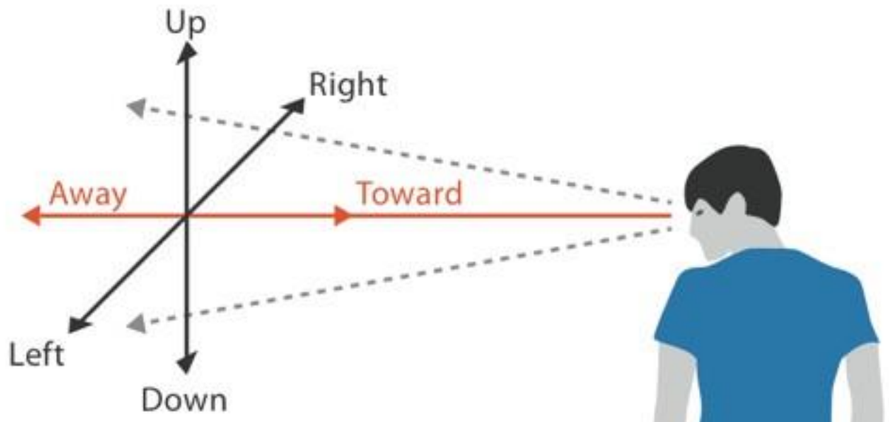


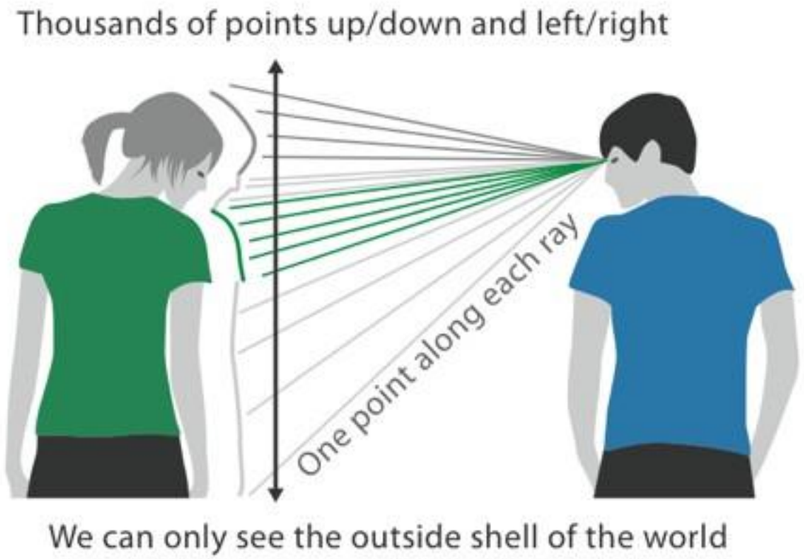
Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape



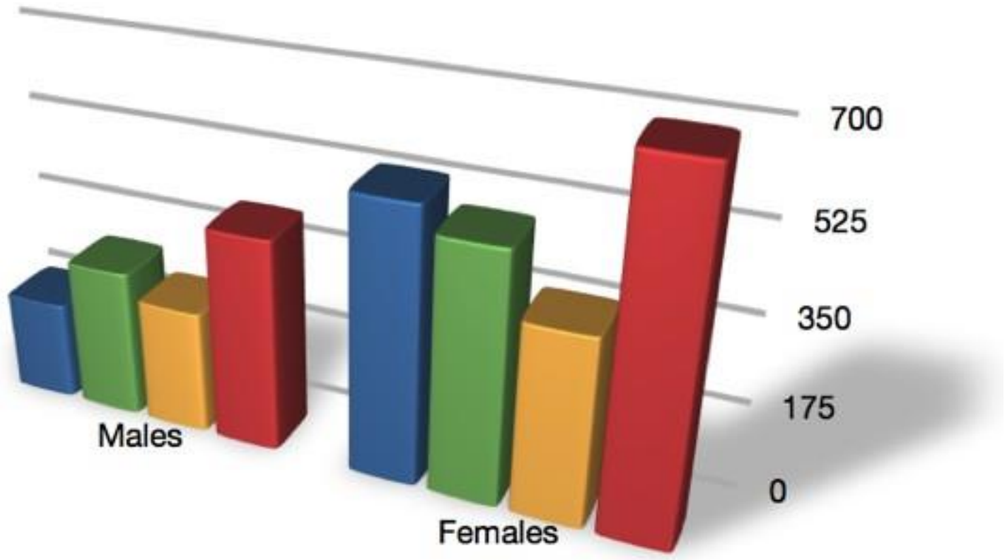
we see the world as a ~~2.5D~~ 2.05D space

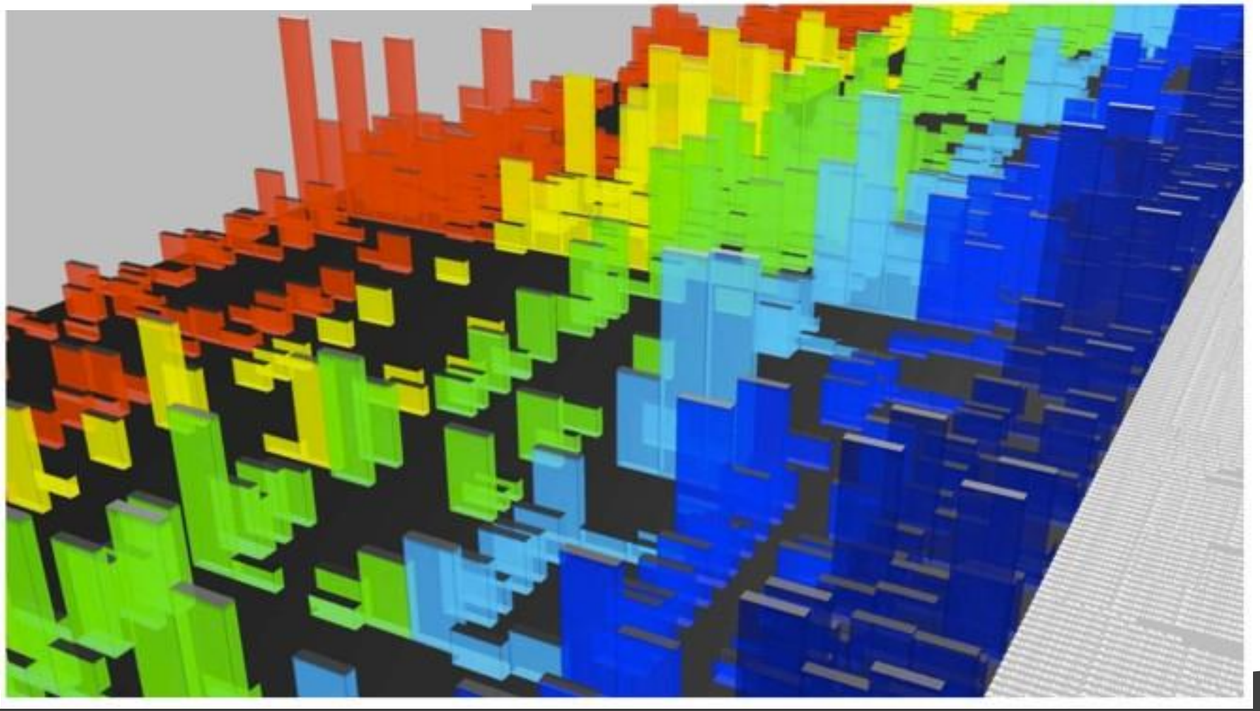


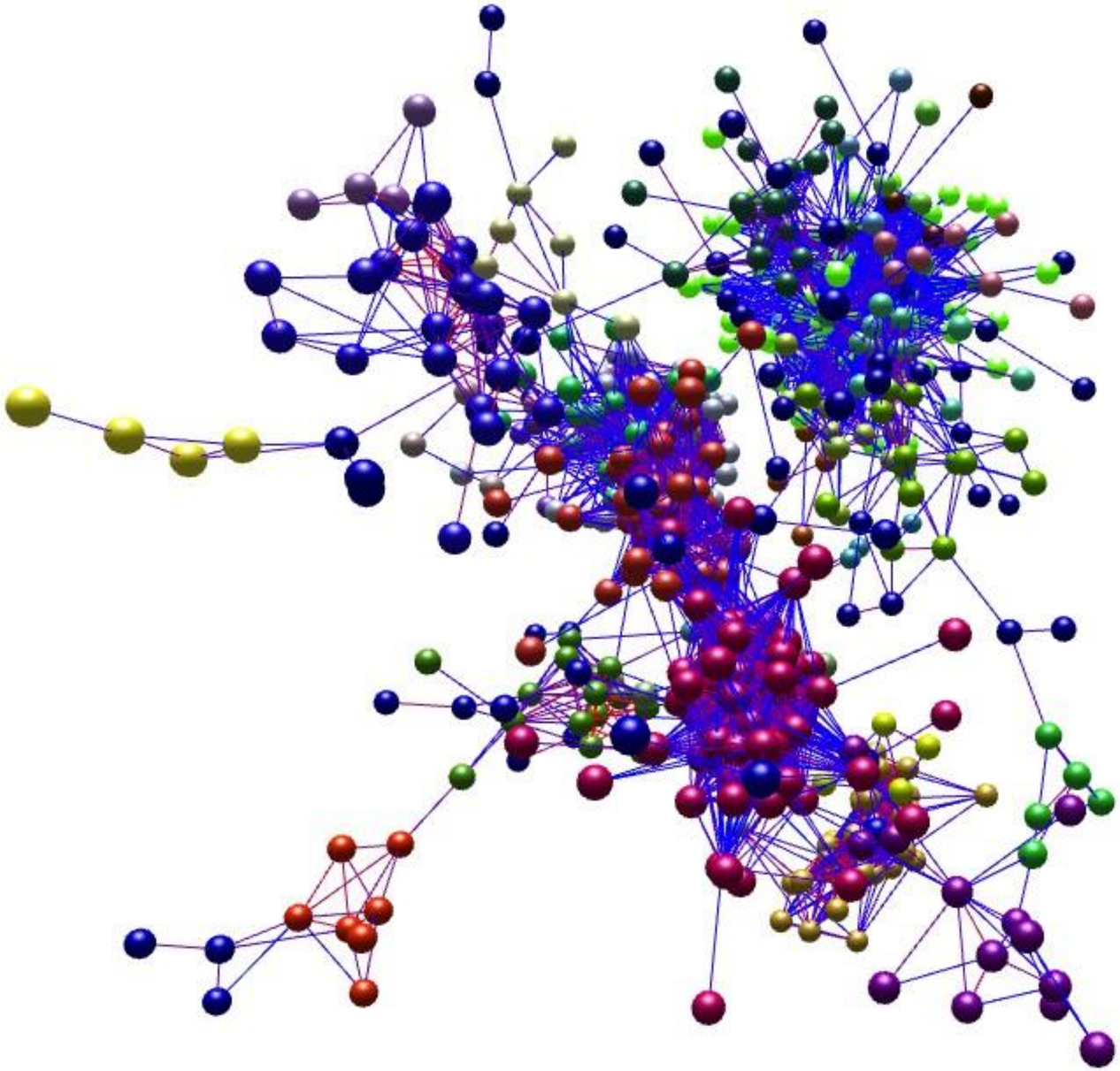


power does not extend to 3D

- perspective cues
  - interfere with color and size channels
- occlusion of data
- text legibility


700
525
350
175
0
Males
Females





TIME AS ENCODING CHANNEL

- You'll remember, visualization uses pictures to enhance working memory
- external versus internal memory
  - *easy to compare views by moving eyes*
  - *hard to compare view to memory of what you saw*

ComParrot by Bonnie J. Malcolm

Can you spot 12 differences between these pictures?

ComParrot
by Bonnie J. Malcolm

Can you spot 12 differences between these pictures?

ComParrot

by Bonnie J. Malcolm

Can you spot 12 differences between these pictures?

Solution: 1. Top tree leaf removed. 2. Nose line on left giraffe removed. 3. Shadow on lower left coconut removed. 4. Leaf vein below gecko removed. 5. Ear line on left giraffe removed. 6. Bottom spot on right giraffe colored in. 7. Small leaf at right of tree colored in. 8. Horn on right giraffe moved. 9. Spot on left giraffe moved. 10. Branch on left side shorter. 11. Gecko tail longer. 12. Gecko eye missing.

ComParrot

by Bonnie J. Malcolm

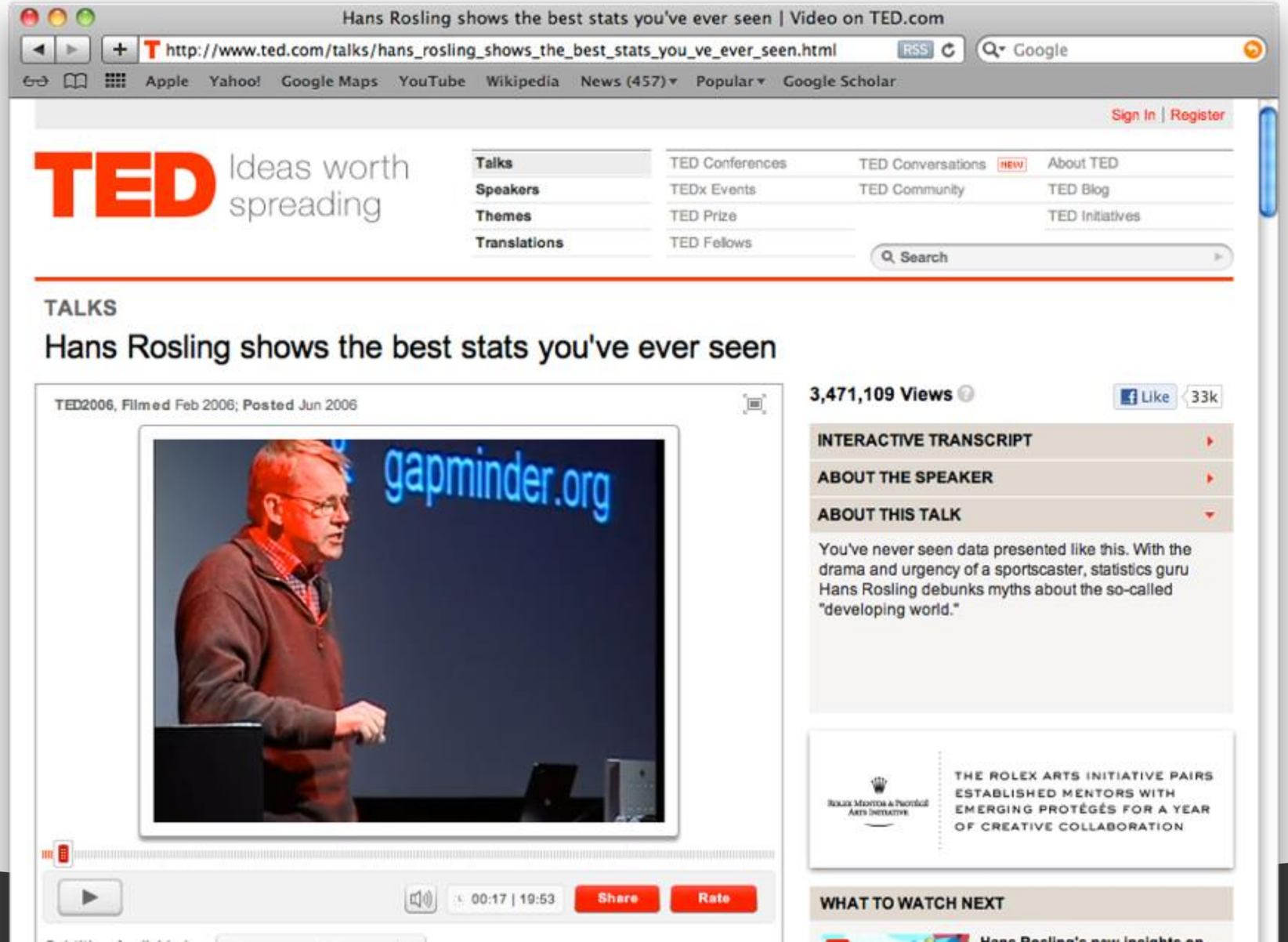
Can you spot 12 differences between these pictures?

www.comparrotpuzzles.com © 2001 Bonnie J. Malcolm

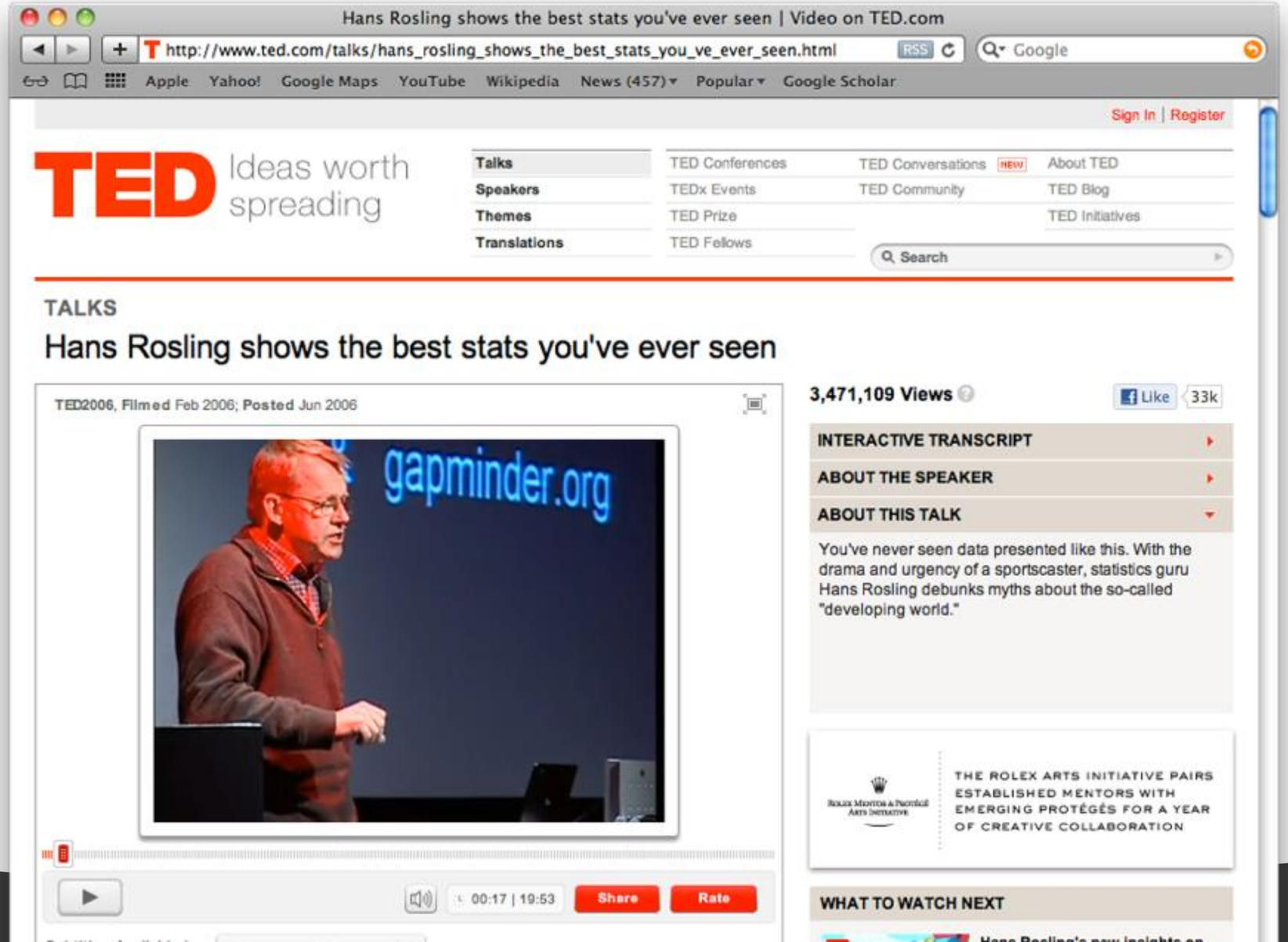
Solution: 1. Top tree leaf removed. 2 removed. 4. Leaf vein below gecko
removed. 5. Ear line on left giraffe re of tree colored in. 6. Horn on right
giraffe moved. 8. Spot on left giraffe m ye missing.

WHEN TO USE ANIMATION?

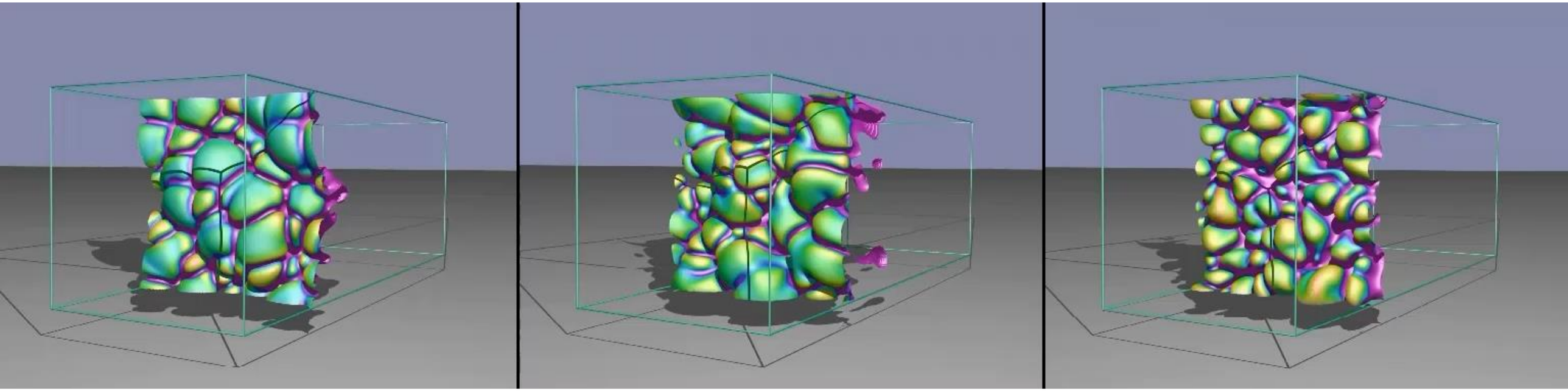
GOOD: STORYTELLING



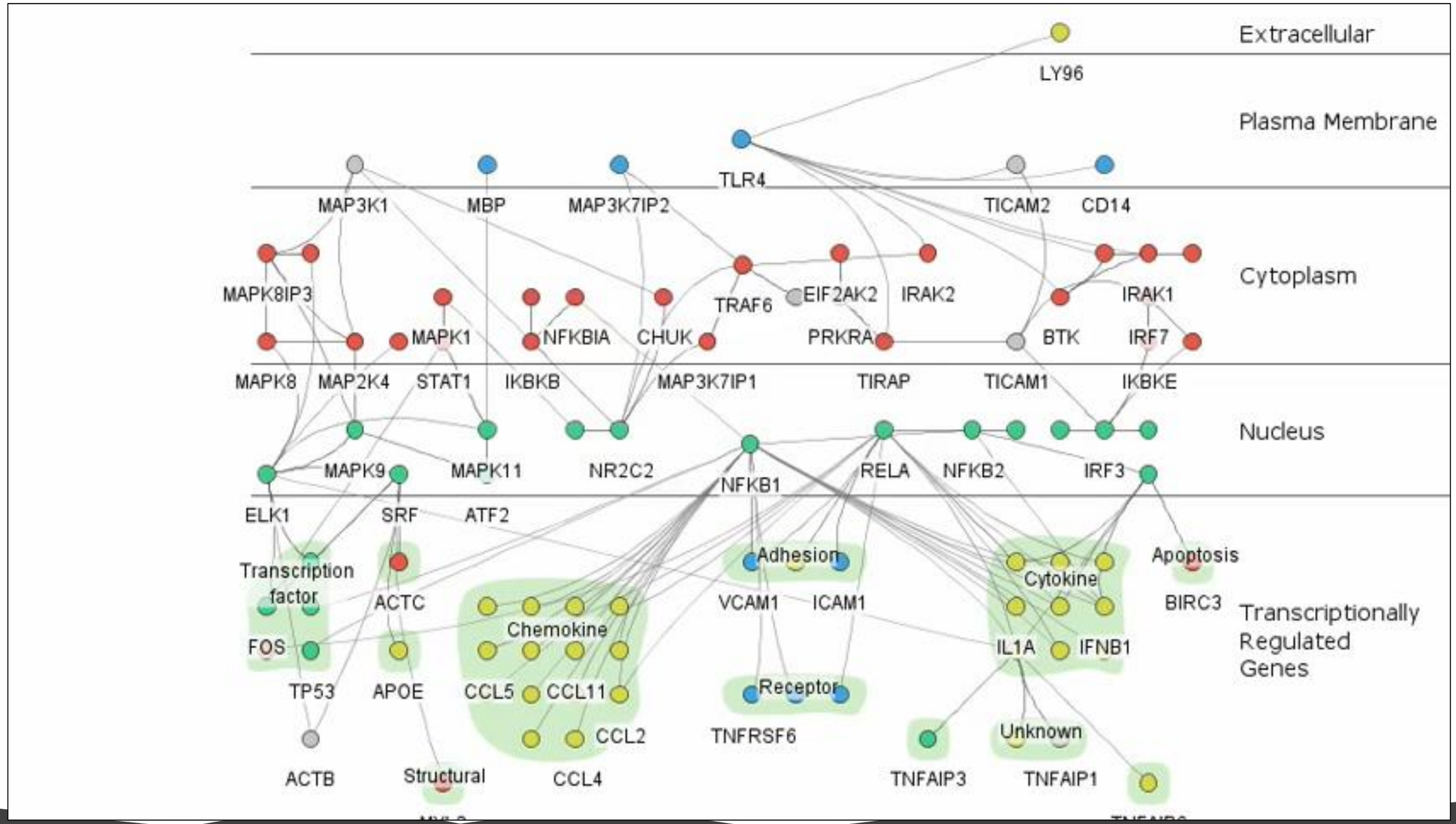
GOOD: TRANSITIONS



BAD: COMPARING COMPLEX STATE CHANGES OVER TIME



BAD: MULTIPLE STATES WITH MULTIPLE CHANGES



BAD: MULTIPLE STATES WITH MULTIPLE CHANGES (use small multiples instead)