

Paul Rosen

paul.rosen@utah.edu
@paulrosenphd
<https://cspaul.com>



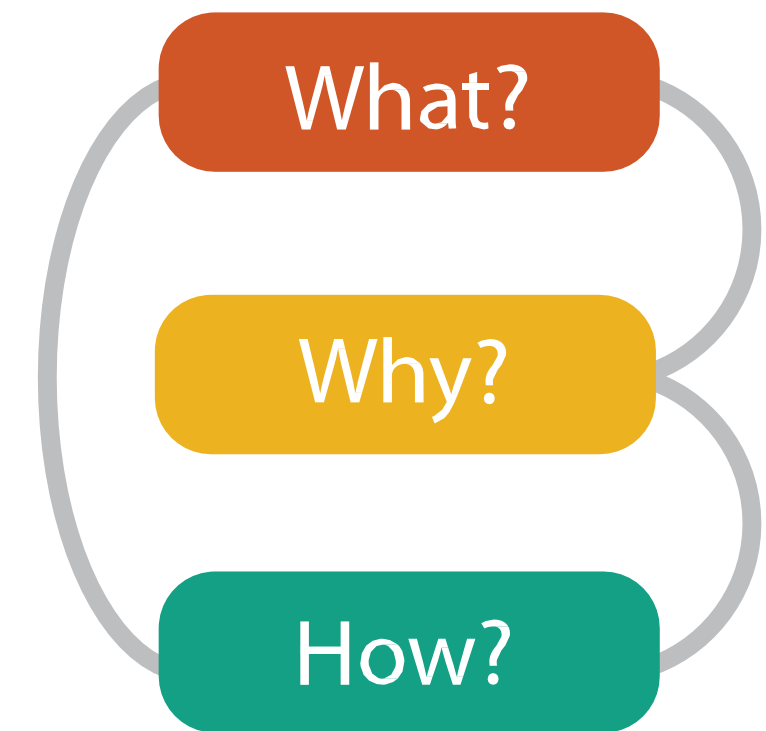
Visualization for Data Science

DS-4630 / CS-5630

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



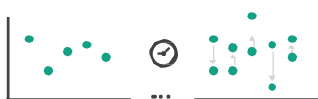
➔ Motion

Direction, Rate, Frequency, ...



Manipulate

➔ Change



➔ Select



➔ Navigate



Facet

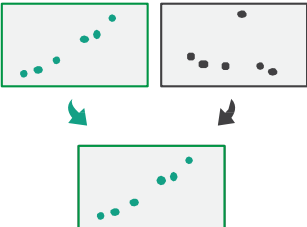
➔ Juxtapose



➔ 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)

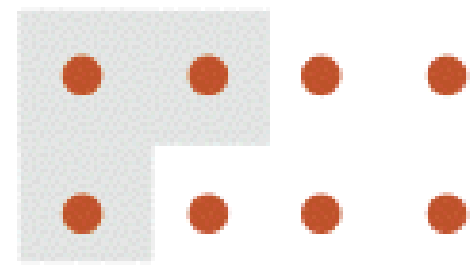


lines (1D)

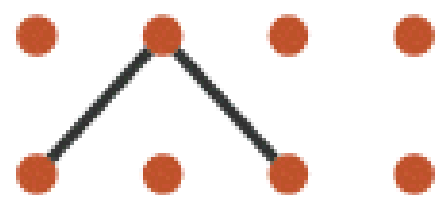


areas (2D)

marks as links



containment



connection

CHANNELS

- parameters that control the appearance of marks

➔ Position

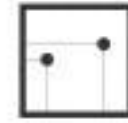
➔ Horizontal



➔ Vertical



➔ Both



➔ Color



➔ Shape



➔ Tilt



➔ Size

➔ Length



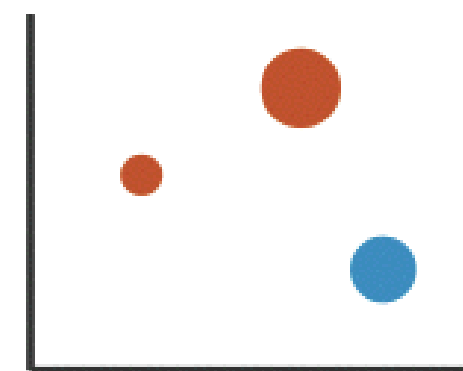
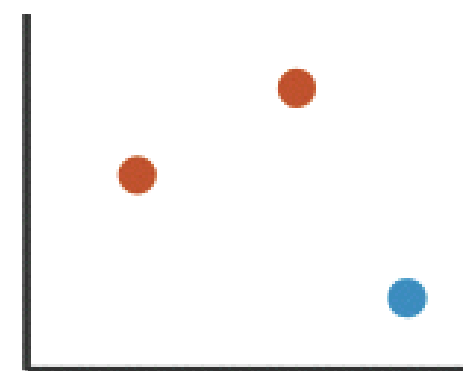
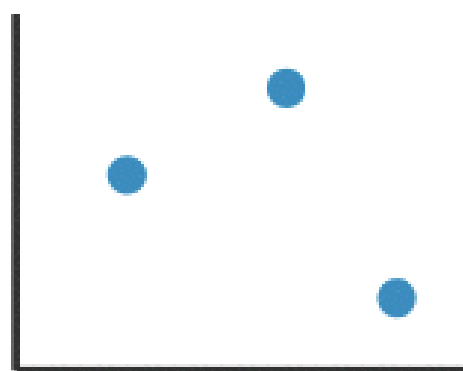
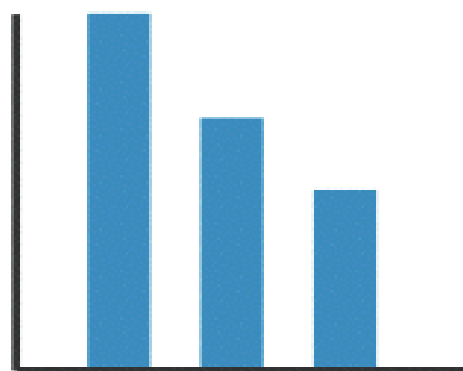
➔ Area



➔ Volume



name that mark and channel



CHANNEL TYPES

identity (what or where)

magnitude (how much)

→ Position

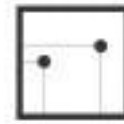
→ Horizontal



→ Vertical



→ Both



→ Color



→ Shape



→ Tilt



→ Size

→ Length



→ Area

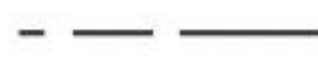
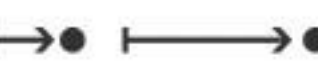


→ Volume



expressiveness & effectiveness

➔ **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)	

(how much)

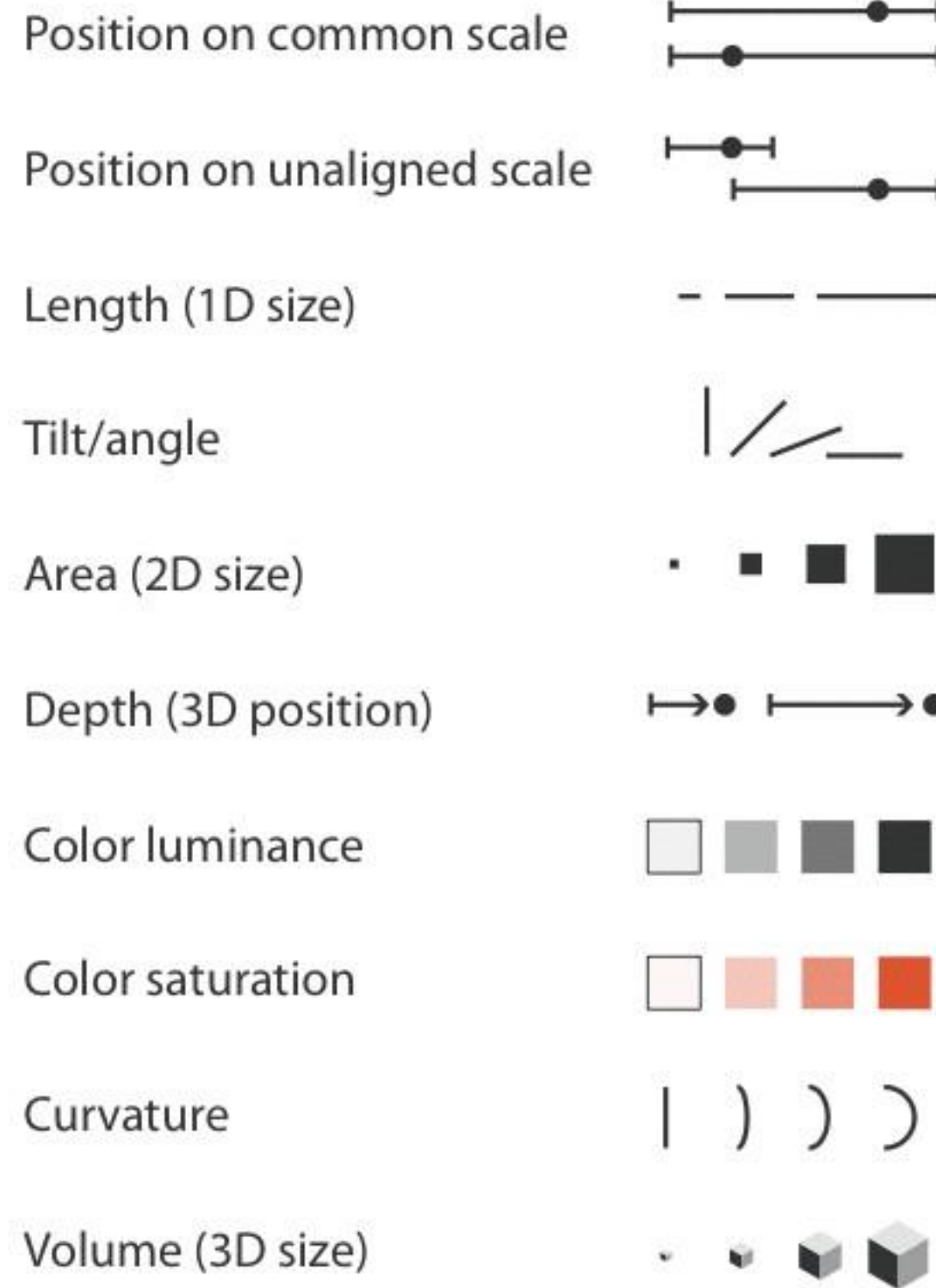
➔ **Identity Channels: Categorical Attributes**

Spatial region	
Color hue	
Motion	
Shape	

(what or where)

expressiveness

➔ **Magnitude Channels: Ordered Attributes**



Same

➔ **Identity Channels: Categorical Attributes**



effectiveness

Most
Effectiveness
Least

WHERE DO RANKINGS COME FROM?

Bertin, “Semiology of Graphics”, 1967

O = Ordinal, Q = Quantitative
≠ = Differences = = Similarities

VARIABLES OF THE IMAGE		POINT			LINE		AREA (ZONE)	
Z	XY 2 DIMENSIONS OF THE PLANE	OQ	≠					
	SIZE	OQ	≠					
	VALUE	O	≠					
	TEXTURE	O	≠					
	COLOR		≠					
DIFFERENTIAL VARIABLES								
	ORIENTATION		≠					
	SHAPE		≠					

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

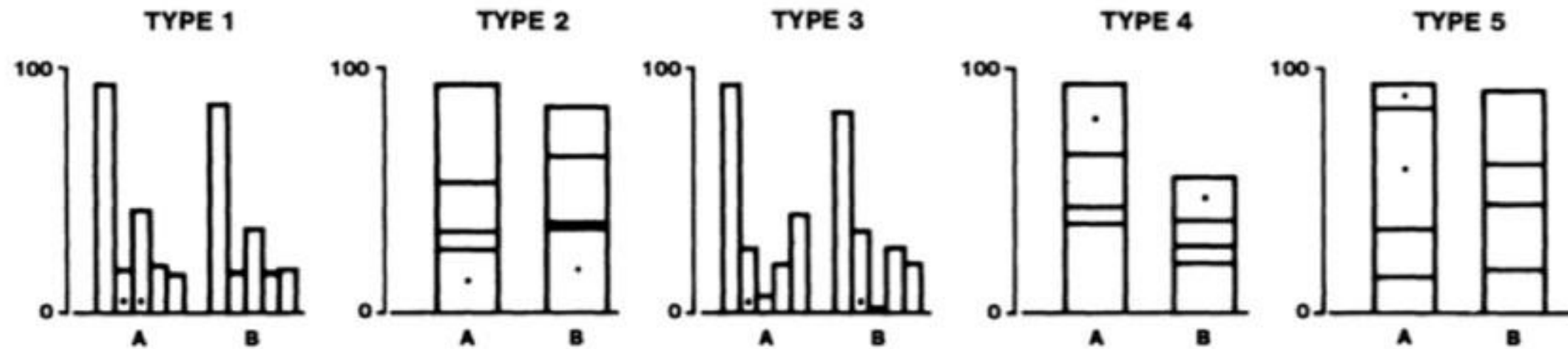


Figure 4. Graphs from position-length experiment.

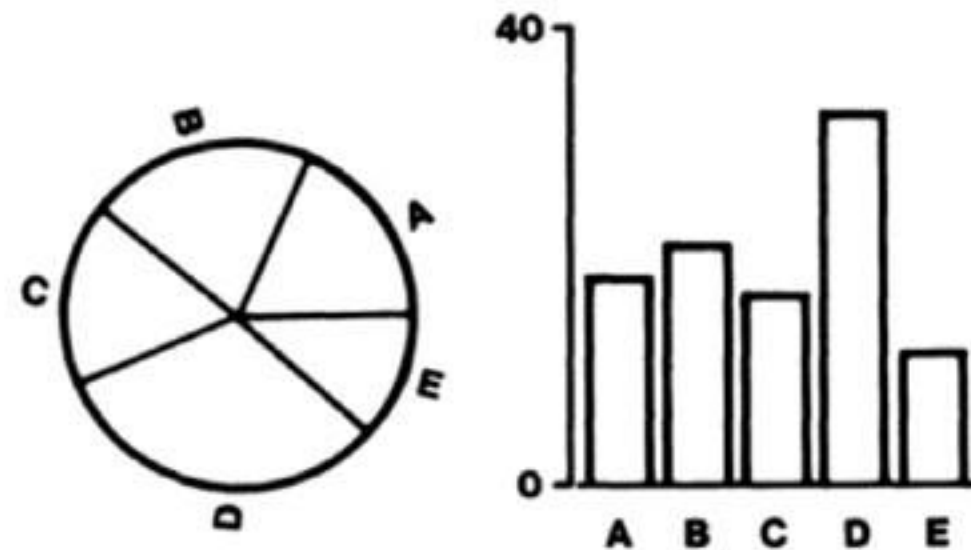
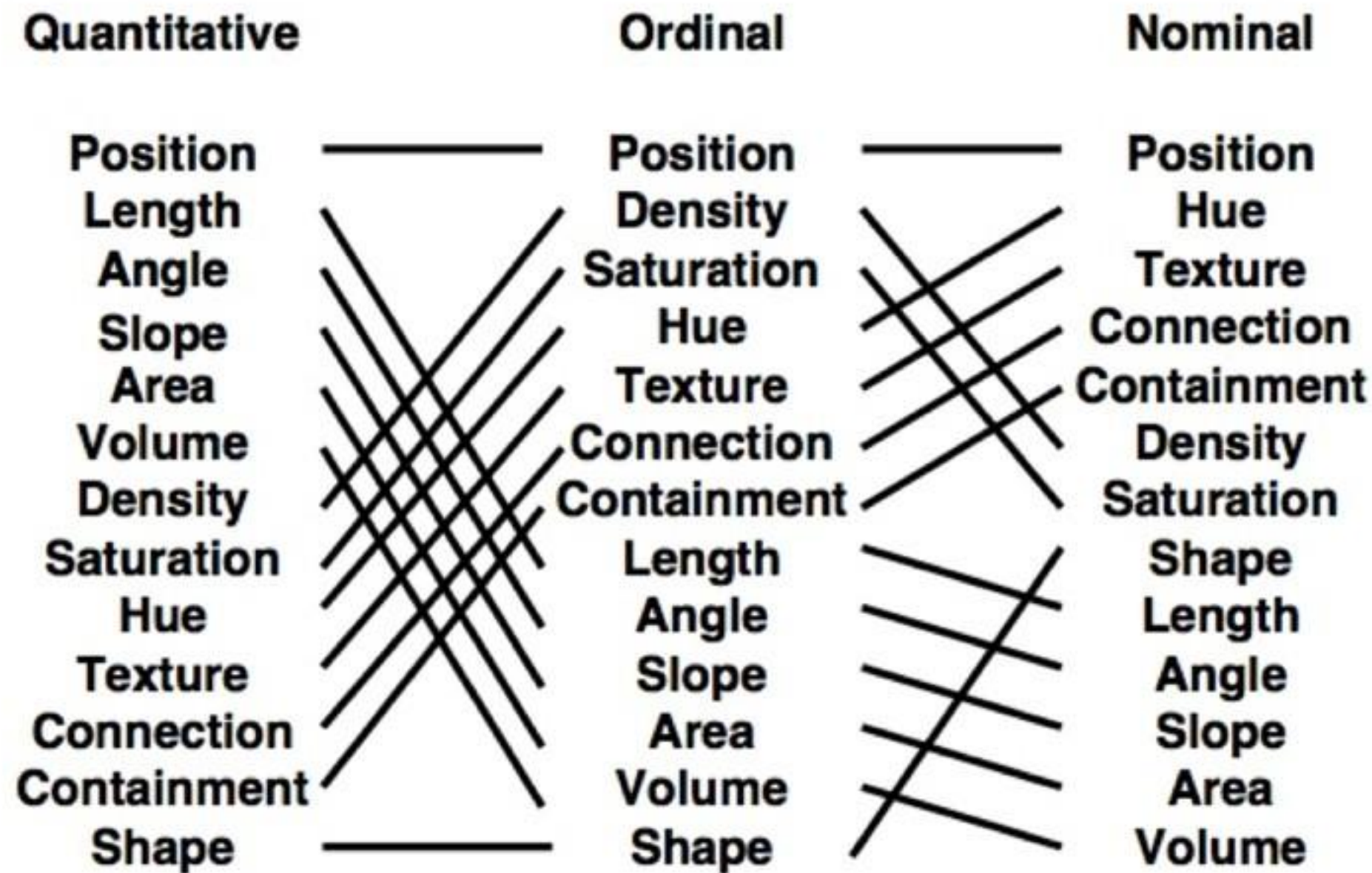


Figure 3. Graphs from position-angle experiment.

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



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

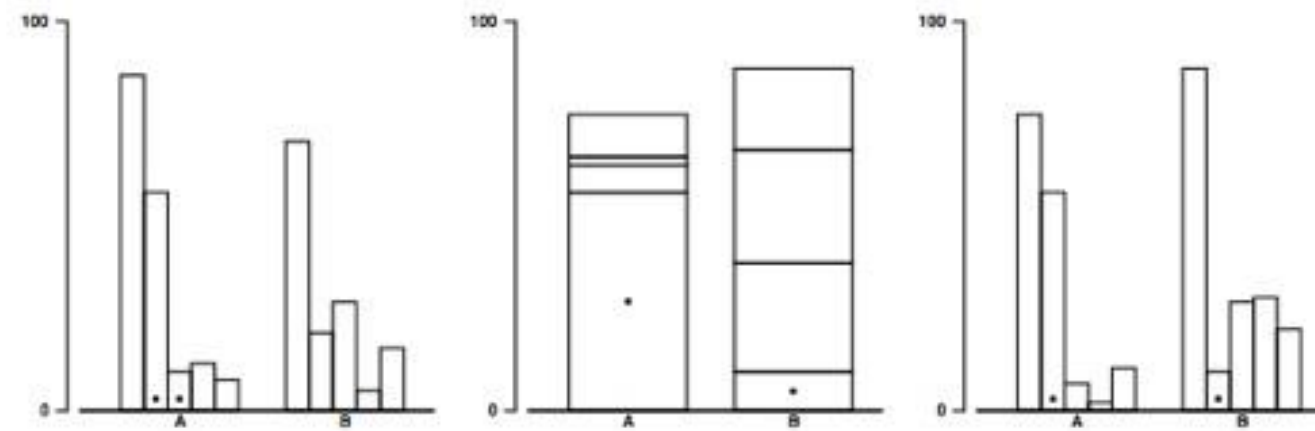


Figure 1: Stimuli for judgment tasks T1, T2 & T3. Subjects estimated percent differences between elements.

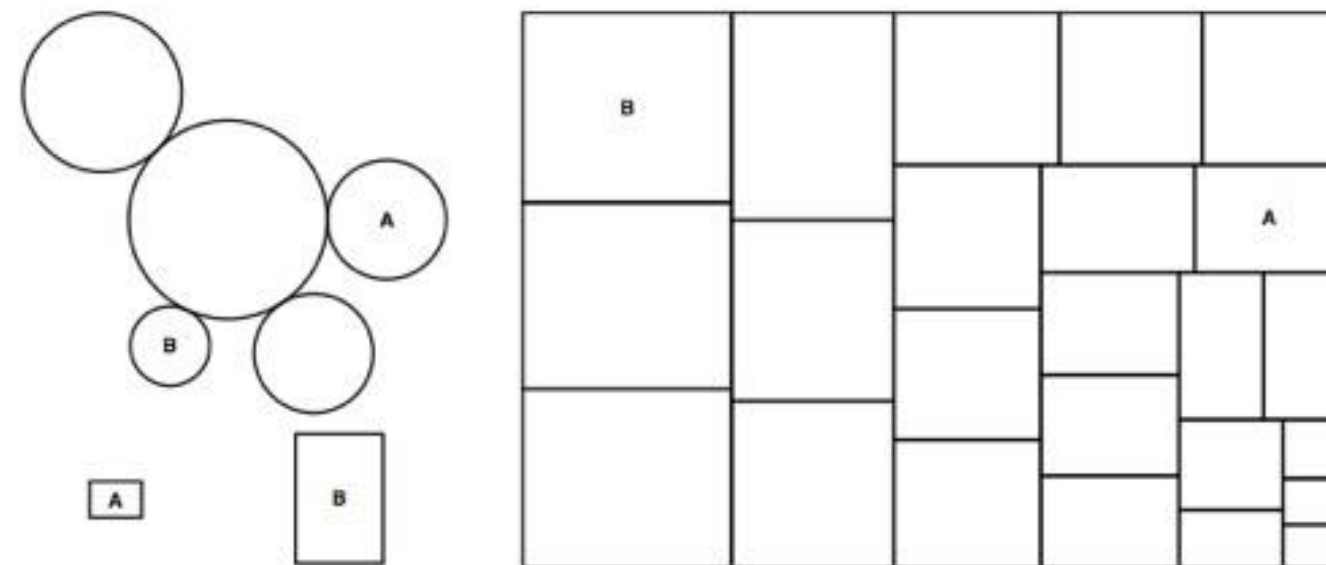
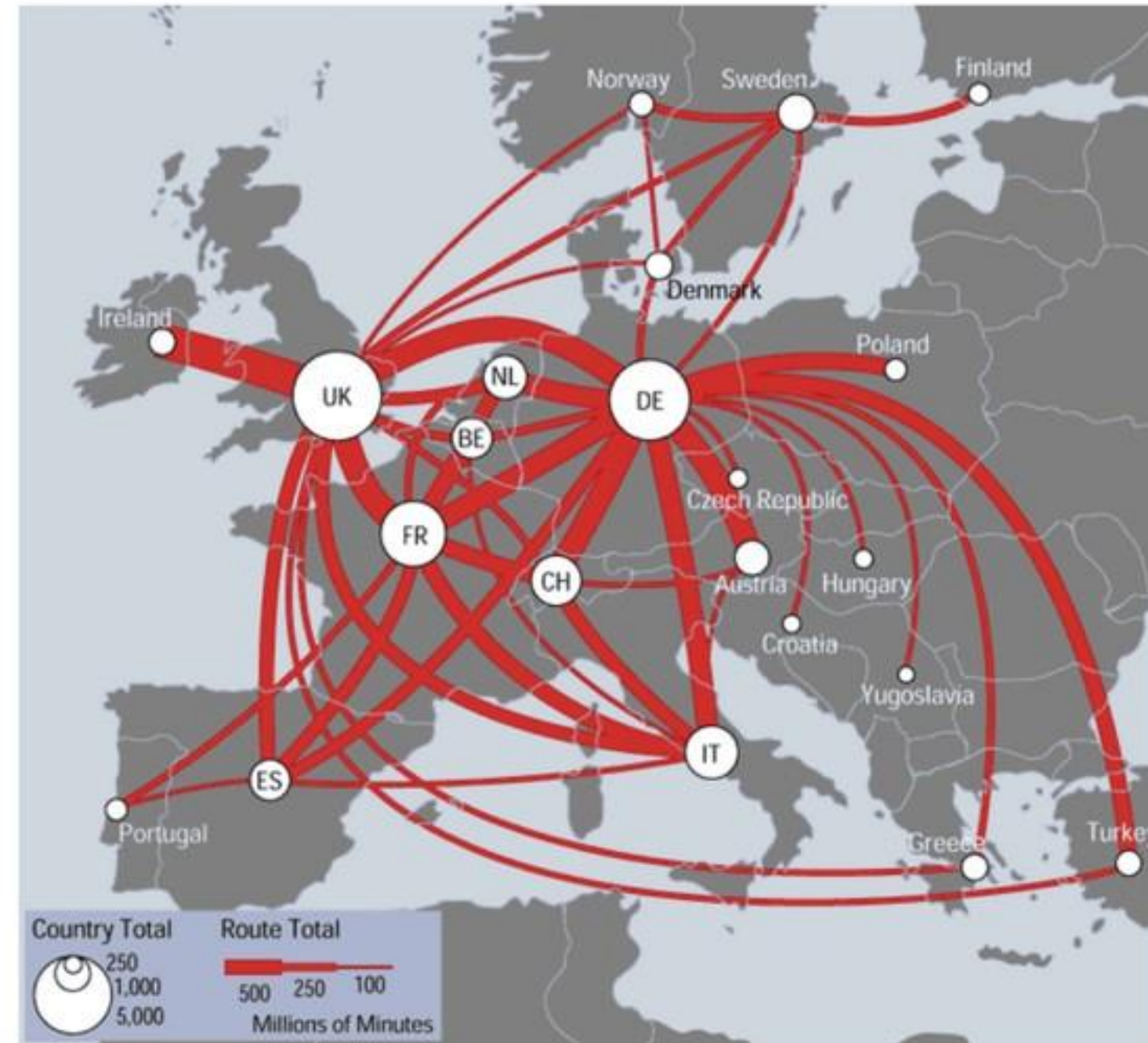


Figure 2: Area judgment stimuli. Top left: Bubble chart (T7), Bottom left: Center-aligned rectangles (T8), Right: Treemap (T9).



DISCRIMINABILITY

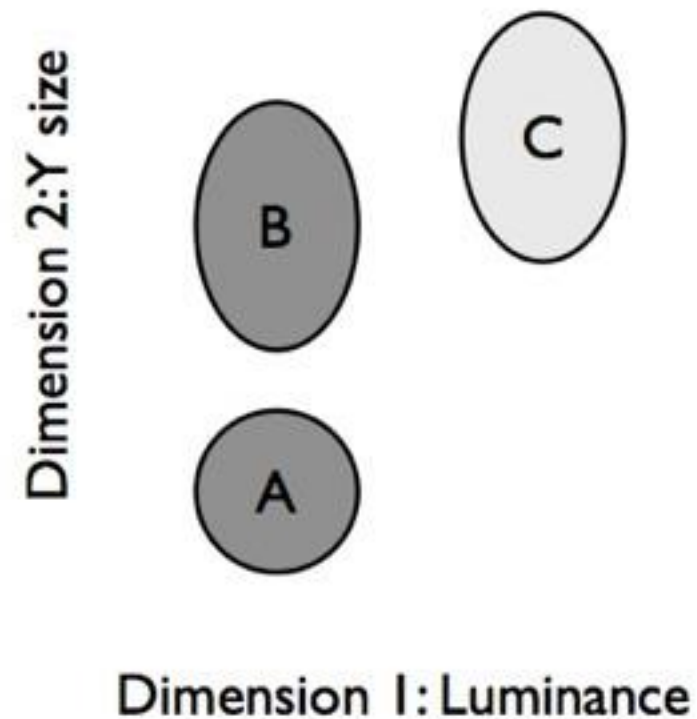
- can channel differences be discerned?



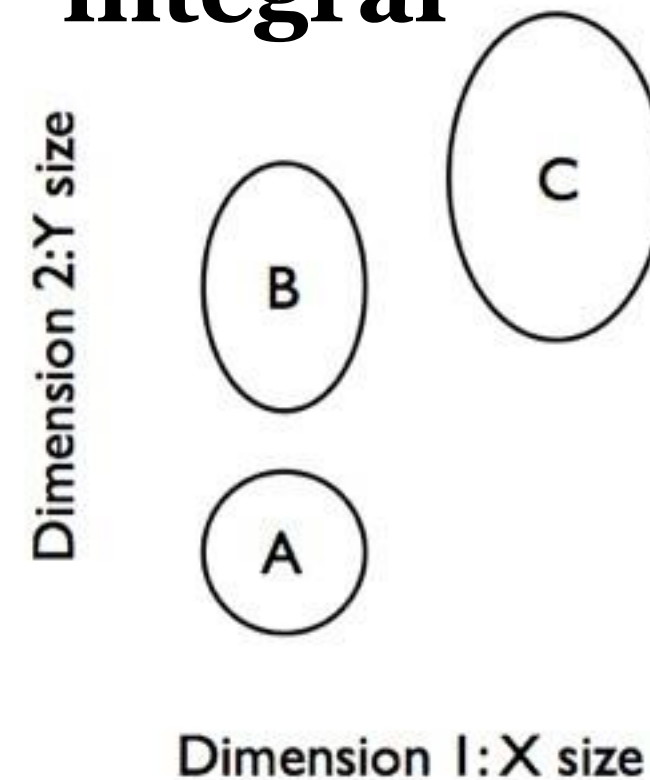
SEPARABLE vs INTEGRAL

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

separable



integral



SEPARABLE vs INTEGRAL

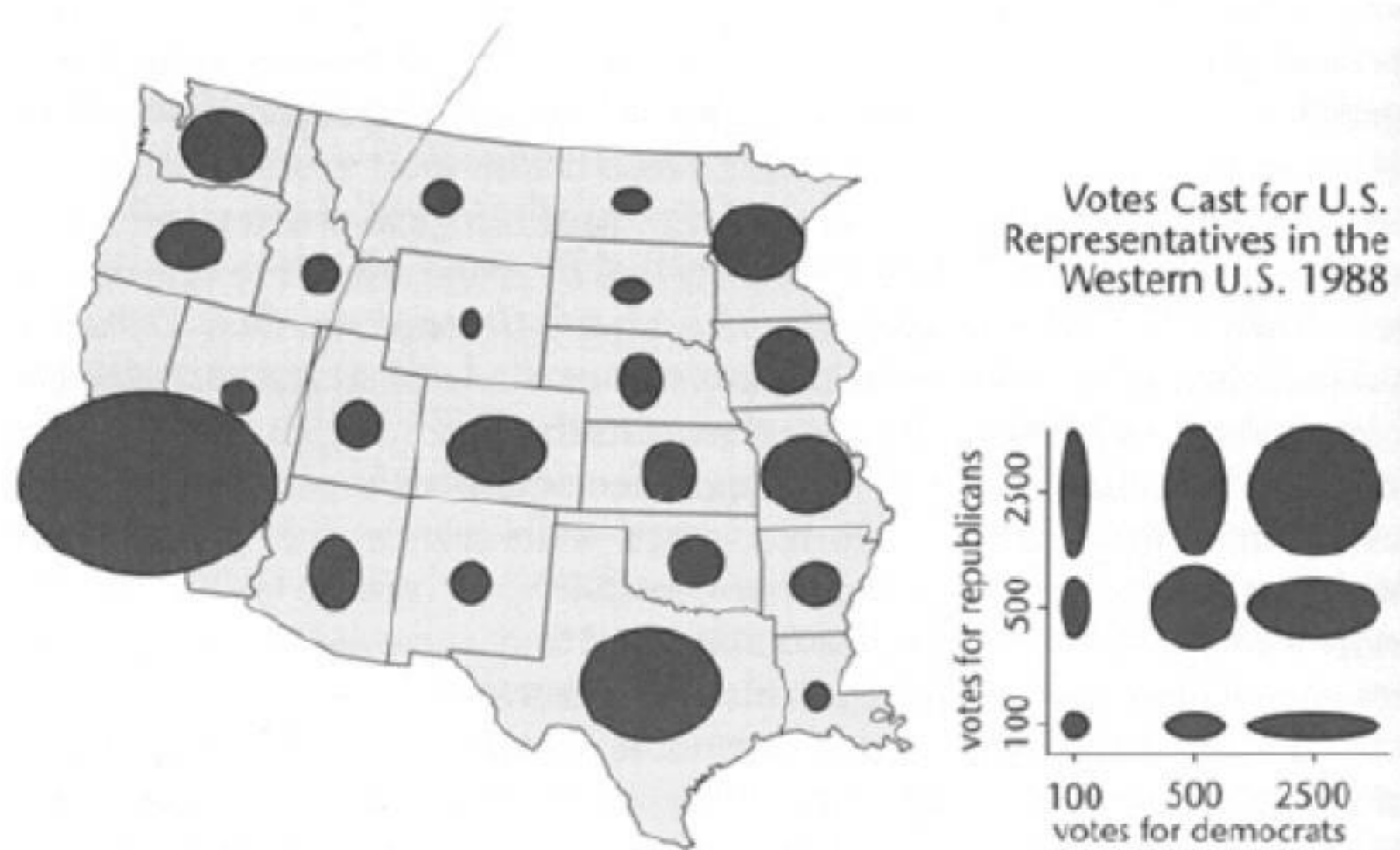
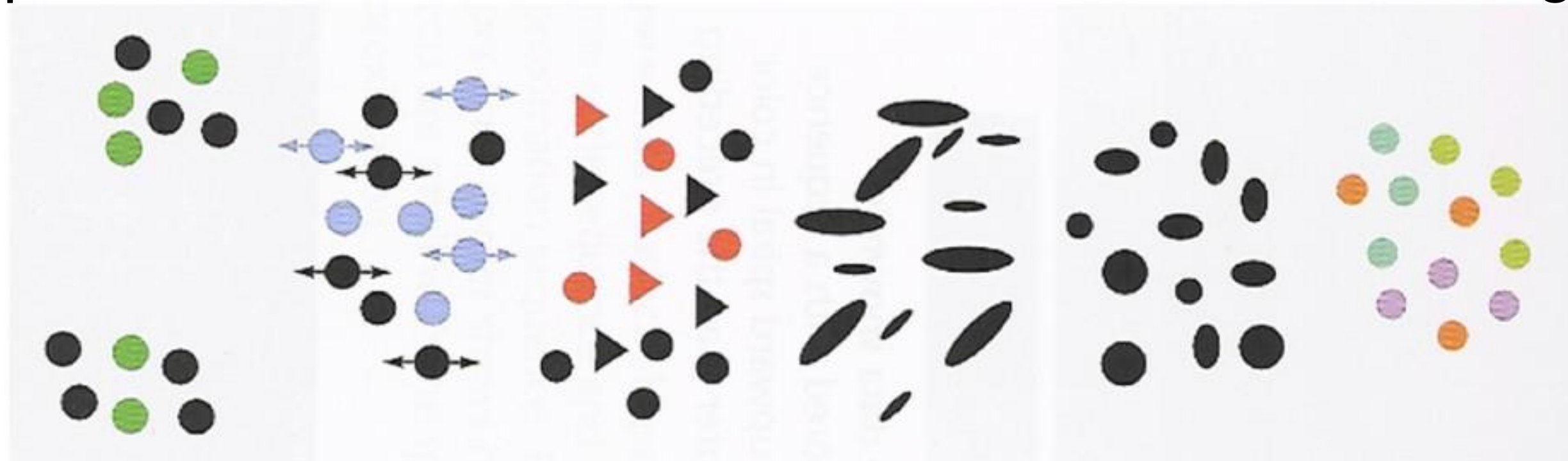


FIGURE 3.38. An example of the use of an ellipse as a map symbol in which the horizontal and vertical axes represent different (but presumably related) variables.

SEPARABLE vs INTEGRAL

separable ← → integral



color | location

color | motion

color | shape

size | orientation

x-size | y-size

red-green | yellow-blue

READING, WRITING, AND EARNING MONEY

The latest data from the U.S. Census's American Community Survey paints a fascinating picture of the United States at the county level. We've looked at the data about achievement and the median income of the entire nation, to see where people are going to school, where they're earning money, and if there is any correlation.



① HIGH SCHOOL GRADUATES
60% 70% 80% 90%



② COLLEGE GRADUATES
10% 20% 30% 40%

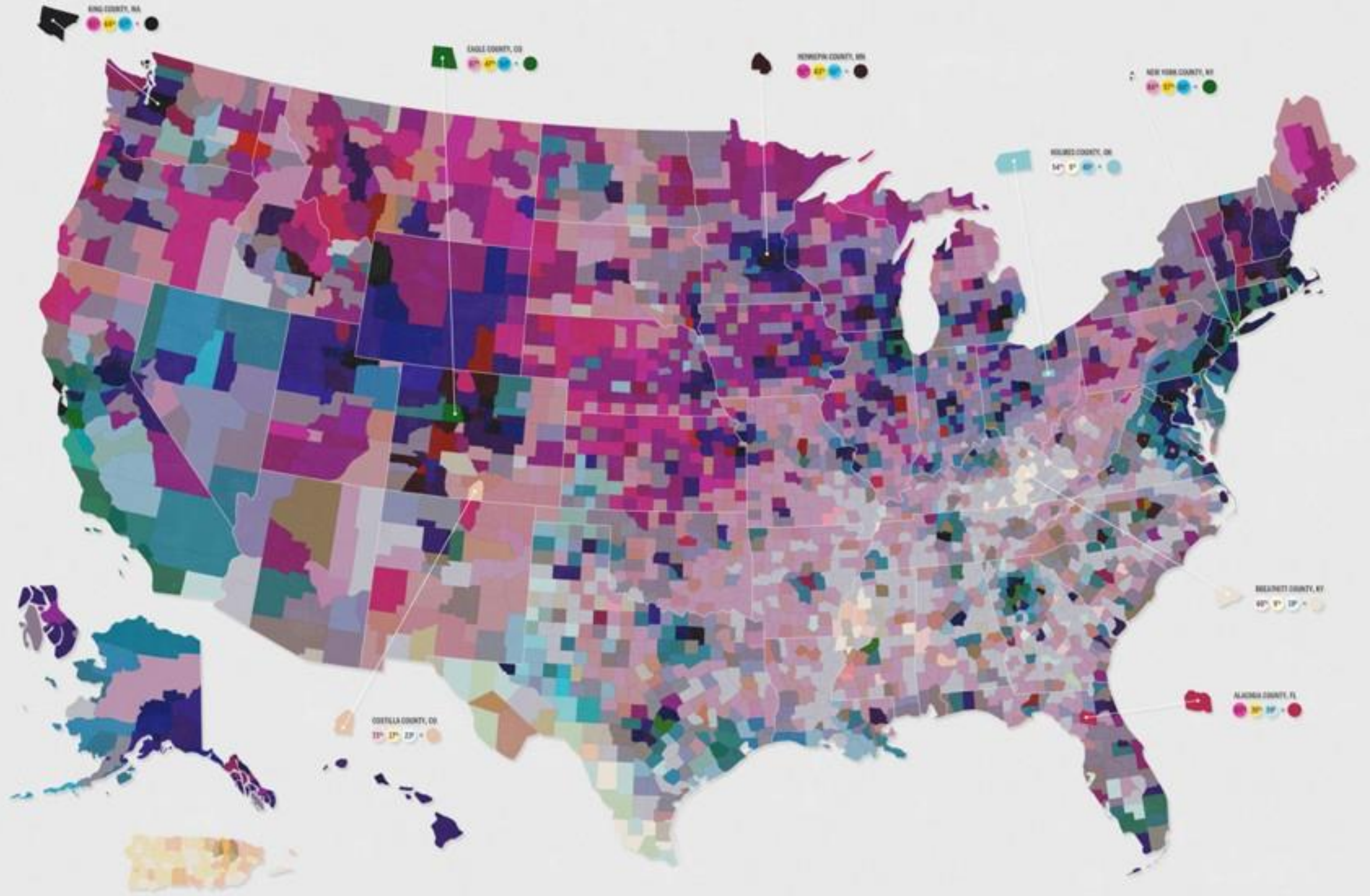


③ MEDIAN HOUSEHOLD INCOME
20K 40K 60K 80K

The map at right is a product of overlaying the three sets of data. The variations in hue and value has been produced from the data shown above. In general, darker counties represent a more educated, better paid population, while lighter areas represent communities with lower graduates and lower incomes.



A collaboration between GOOD and Gregory McNamee
GOOD, WI Census



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

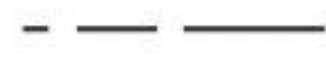
planar position

- 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**

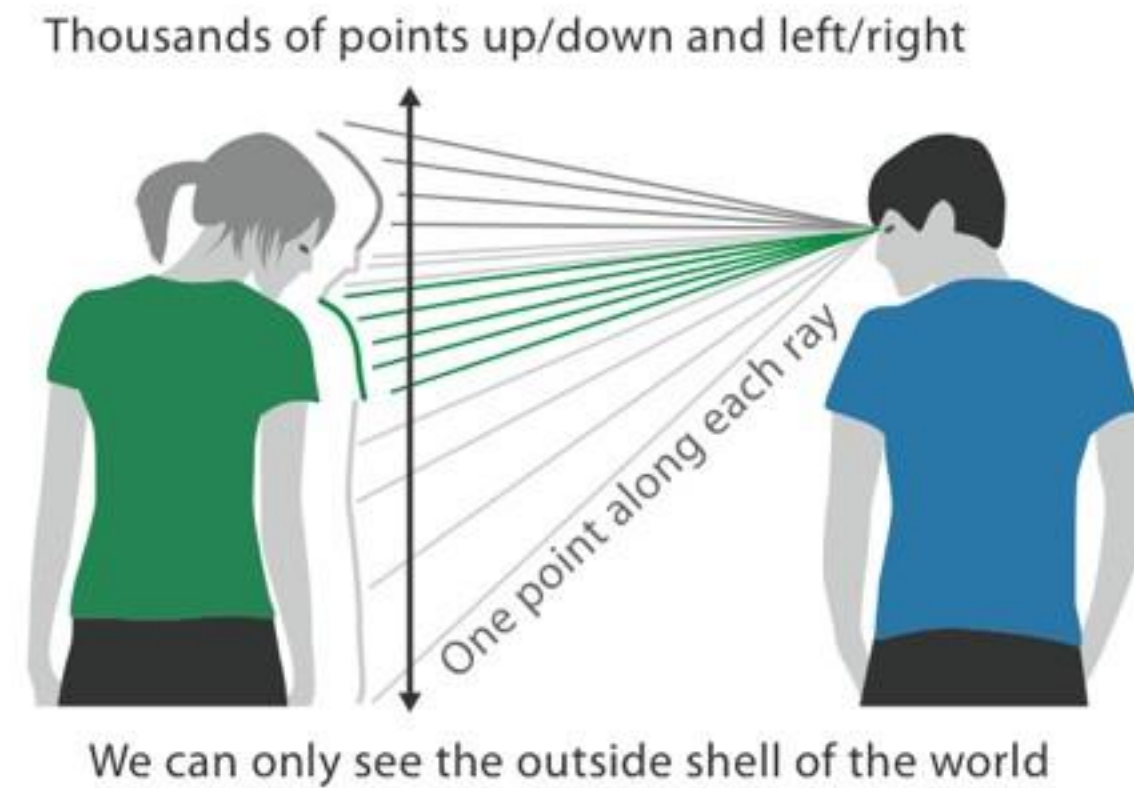
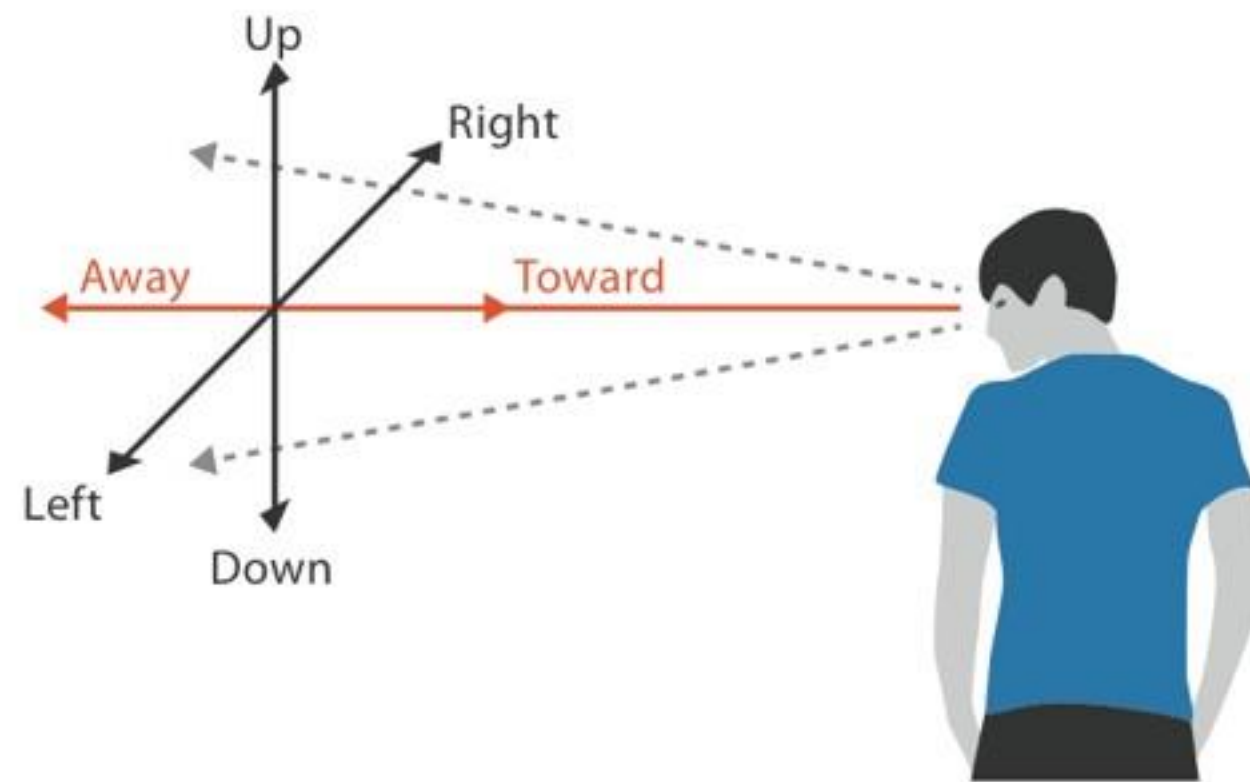
Spatial region 

Color hue 

Motion 

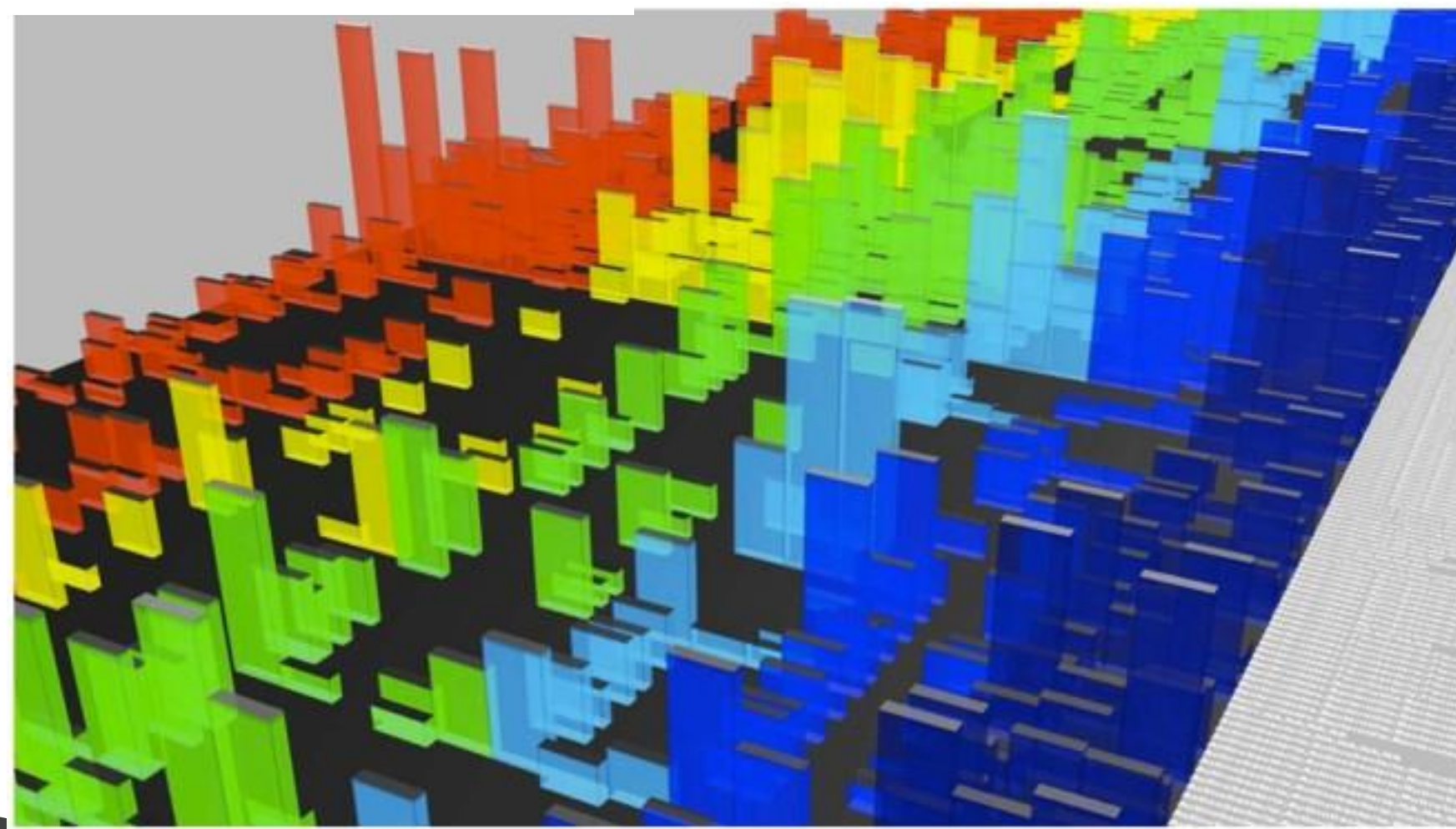
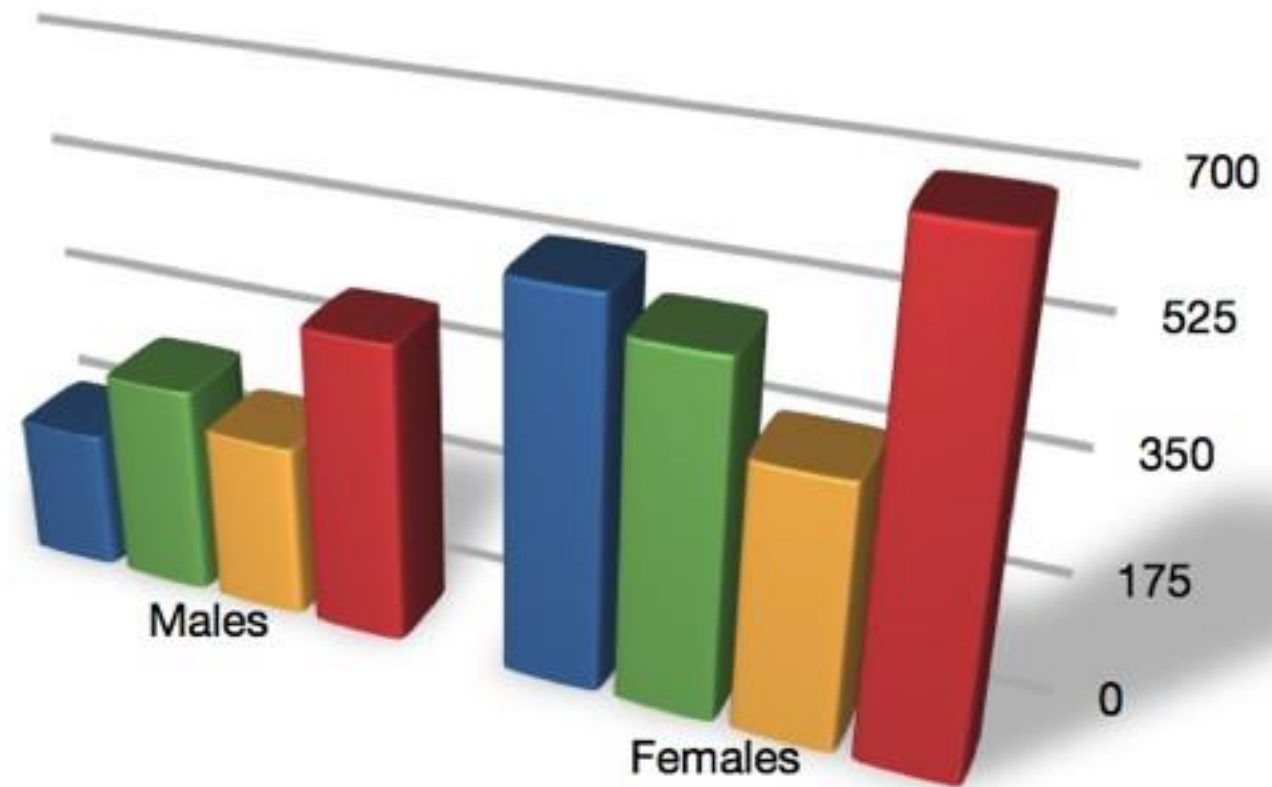
Shape 

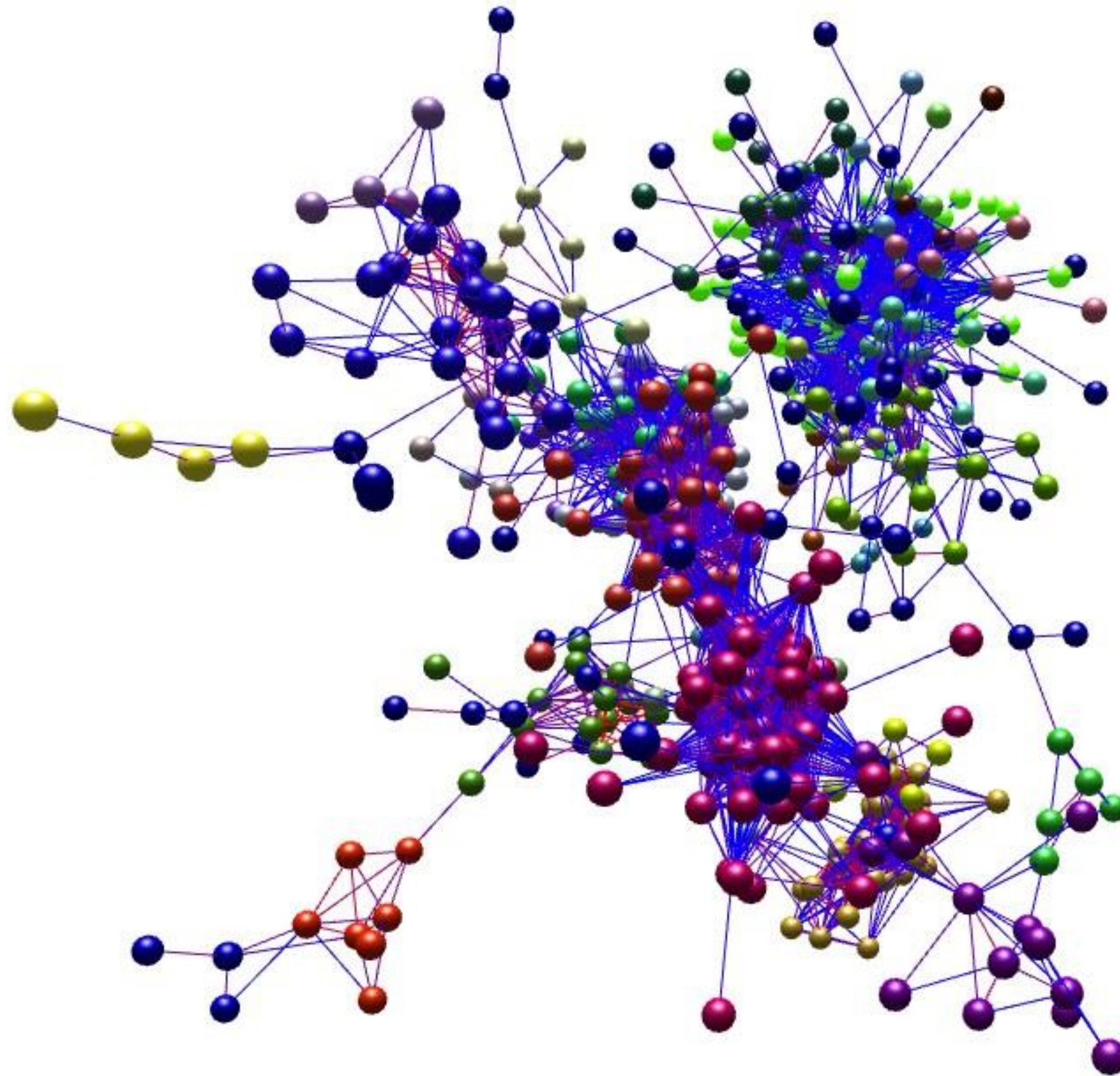
we see the world as a ~~2.5D~~^{2.05D} space



power does not extend to 3D

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





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?



www.comparotpuzzles.com © 2001 Bonnie J. Malcolm

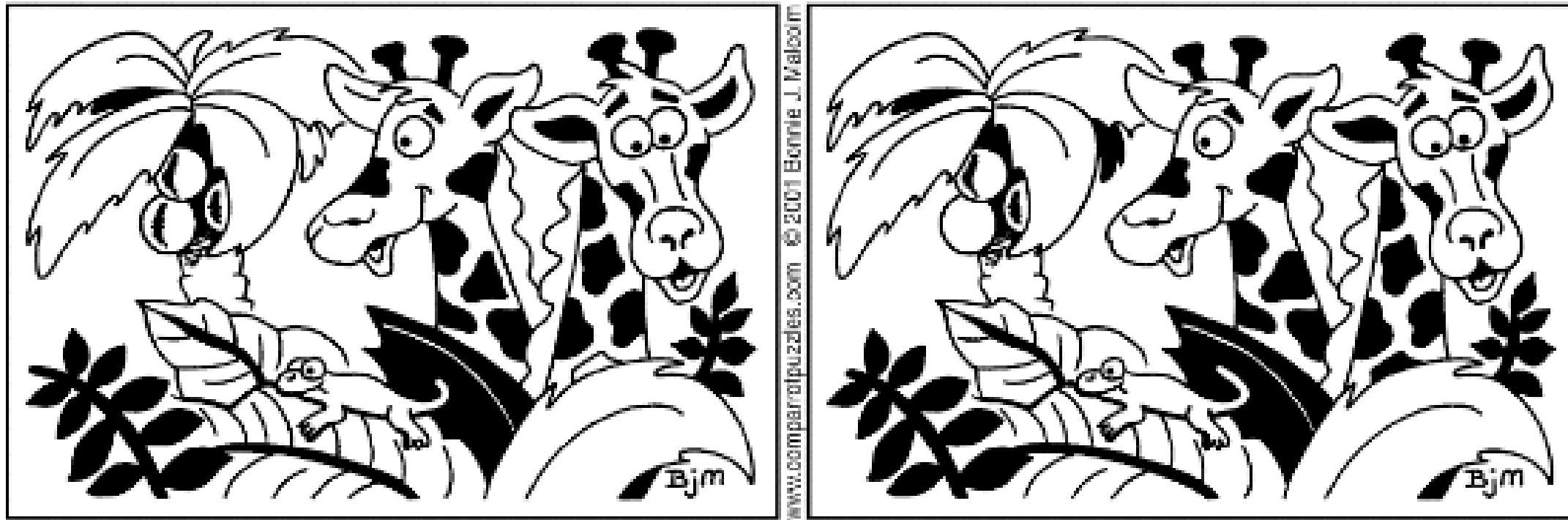
ComParrot
by Bonnie J. Malcolm

Can you spot 12 differences between these pictures?



ComParrot
by Bonnie J. Malcolm

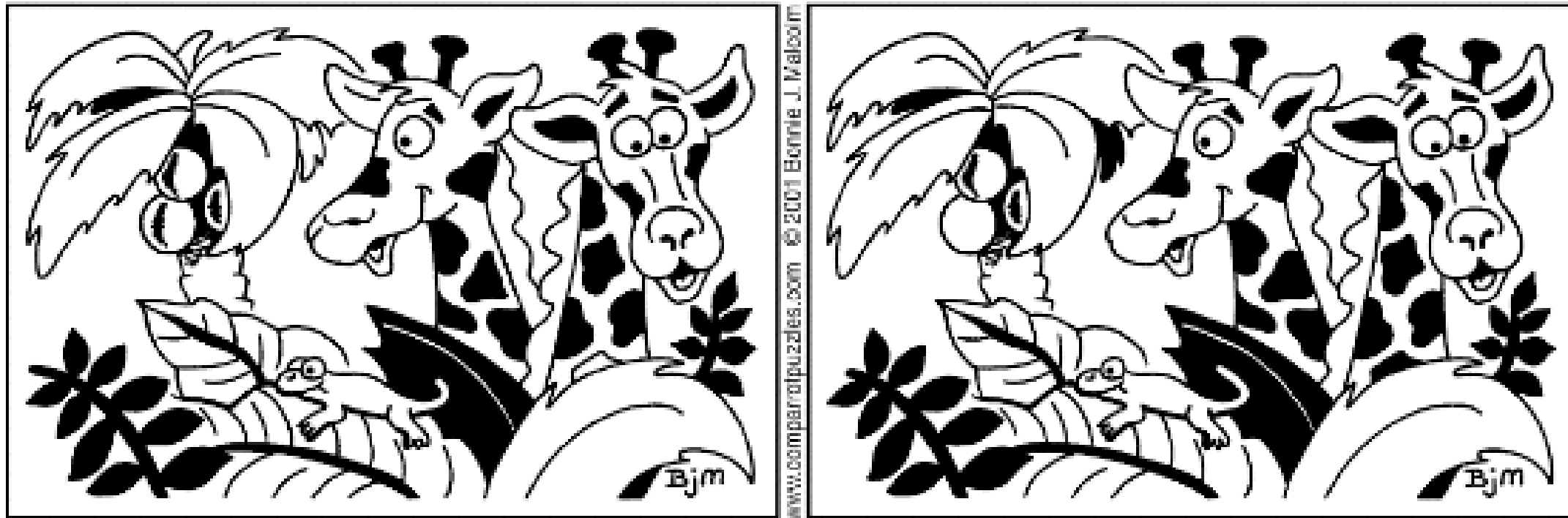
Can you spot 12 differences between these pictures?



Solution: 1. Top tree leaf removed. 2. Nose line on left giraffe removed. 3. Shadow on lower left coconut removed. 4. Leaf vein below geko 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. Gekko tail longer. 12. Gekko eye missing.

ComParrot
by Bonnie J. Malcolm

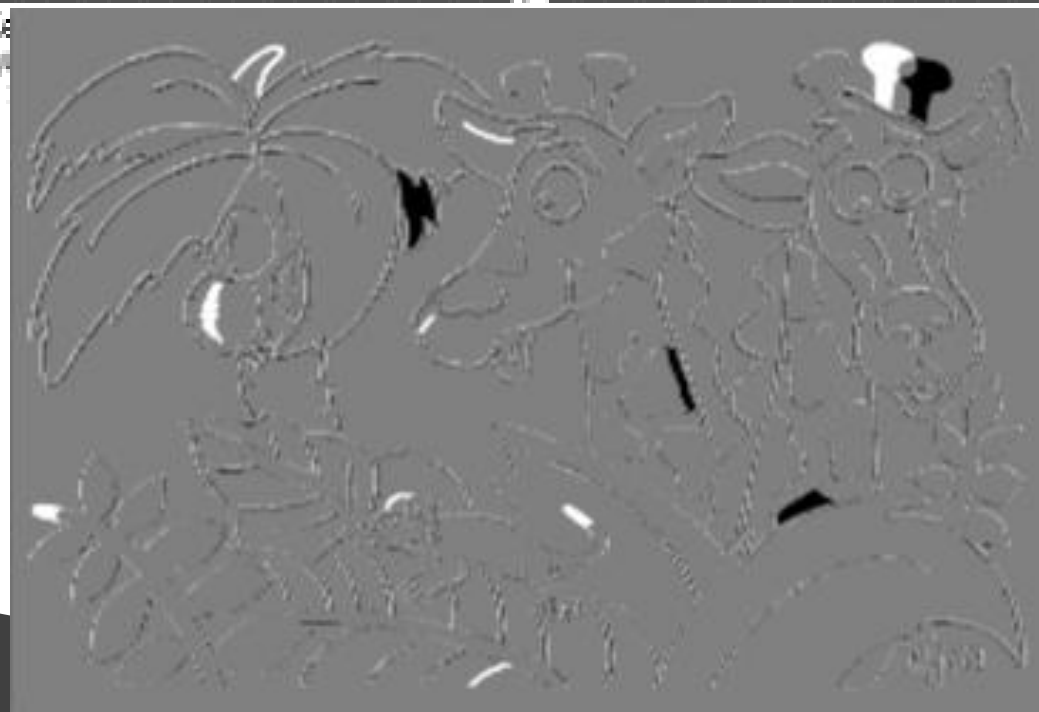
Can you spot 12 differences between these pictures?



www.comparotpuzzles.com © 2001 Bonnie J. Malcolm

removed. 4. Leaf vein below gecko
of tree colored in. 5. Horn on right
eye missing.

Solution: 1. Top tree leaf removed. 2.
removed. 5. Ear line on left giraffe re-
giraffe moved. 8. Spot on left giraffe m-



WHEN TO USE ANIMATION?

GOOD: STORYTELLING

Hans Rosling shows the best stats you've ever seen | Video on TED.com

http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html

Sign In | Register

TED Ideas worth spreading

Talks
Speakers
Themes
Translations

TED Conferences
TEDx Events
TED Prize
TED Fellows

TED Conversations **NEW**
TED Community
TED Blog
TED Initiatives

Search

TALKS

Hans Rosling shows the best stats you've ever seen

TED2006, Filmed Feb 2006; Posted Jun 2006



3,471,109 Views

Like 33k

INTERACTIVE TRANSCRIPT

ABOUT THE SPEAKER

ABOUT THIS TALK

You've never seen data presented like this. With the drama and urgency of a sportscaster, statistics guru Hans Rosling debunks myths about the so-called "developing world."

THE ROLEX ARTS INITIATIVE PAIRS ESTABLISHED MENTORS WITH EMERGING PROTÉGÉS FOR A YEAR OF CREATIVE COLLABORATION

WHAT TO WATCH NEXT

00:17 | 19:53

Share Rate

GOOD: TRANSITIONS

Hans Rosling shows the best stats you've ever seen | Video on TED.com

http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html

Sign In | Register

TED Ideas worth spreading

Talks
Speakers
Themes
Translations

TED Conferences
TEDx Events
TED Prize
TED Fellows

TED Conversations **NEW**
TED Community
TED Blog
TED Initiatives

Search

TALKS

Hans Rosling shows the best stats you've ever seen

TED2006, Filmed Feb 2006; Posted Jun 2006



3,471,109 Views

Like 33k

INTERACTIVE TRANSCRIPT

ABOUT THE SPEAKER

ABOUT THIS TALK

You've never seen data presented like this. With the drama and urgency of a sportscaster, statistics guru Hans Rosling debunks myths about the so-called "developing world."

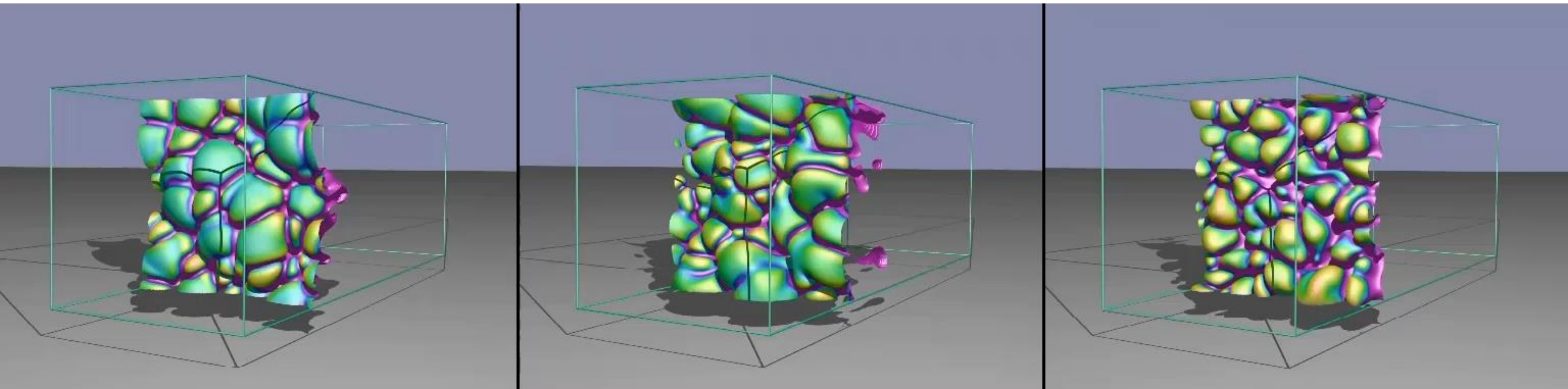
THE ROLEX ARTS INITIATIVE PAIRS ESTABLISHED MENTORS WITH EMERGING PROTÉGÉS FOR A YEAR OF CREATIVE COLLABORATION

WHAT TO WATCH NEXT

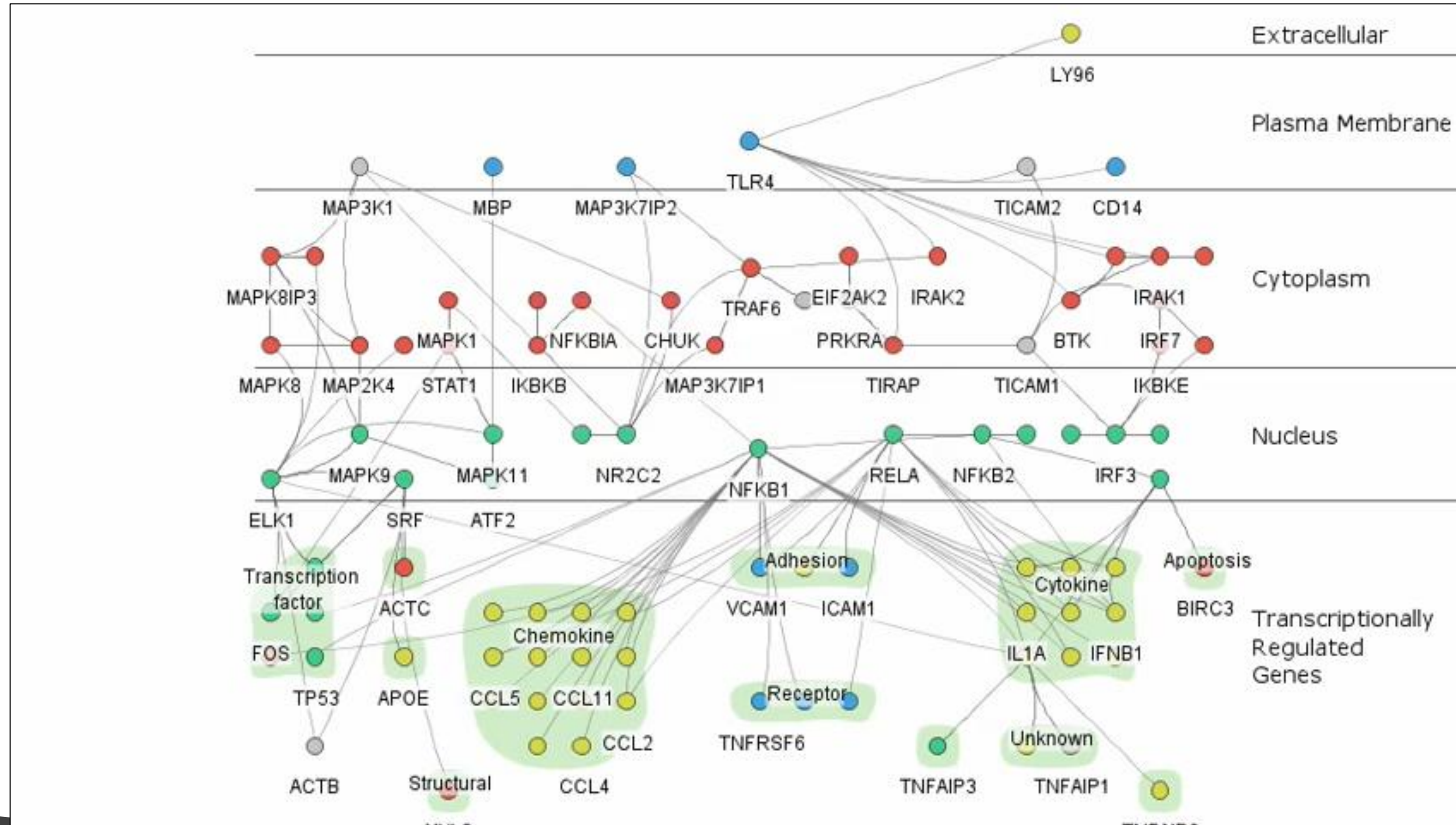
00:17 | 19:53

Share Rate

BAD: COMPARING COMPLEX STATE CHANGES OVER TIME



BAD: MULTIPLE STATES WITH MULTIPLE CHANGES



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

