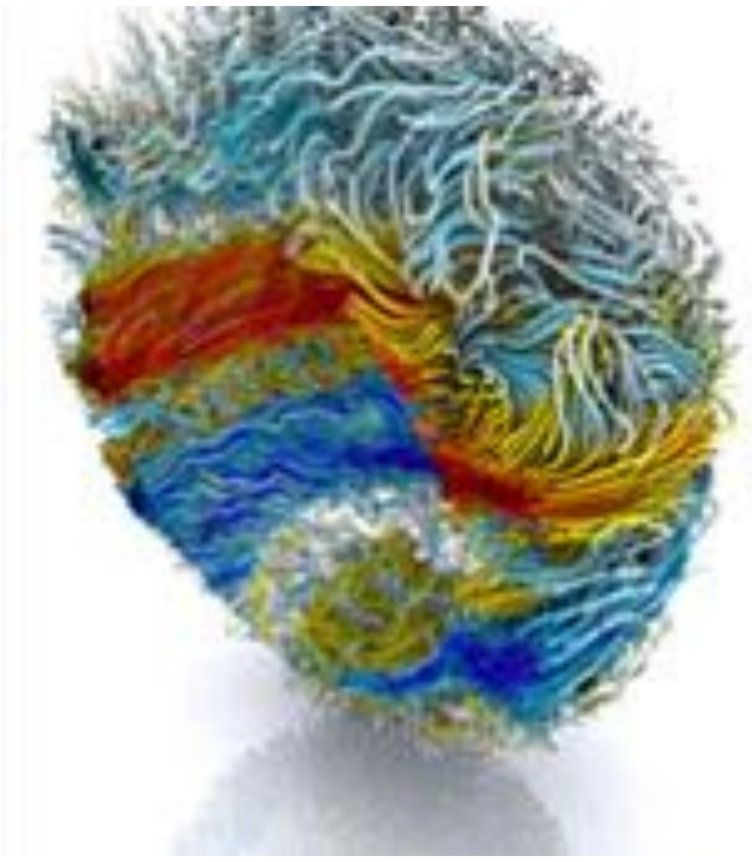
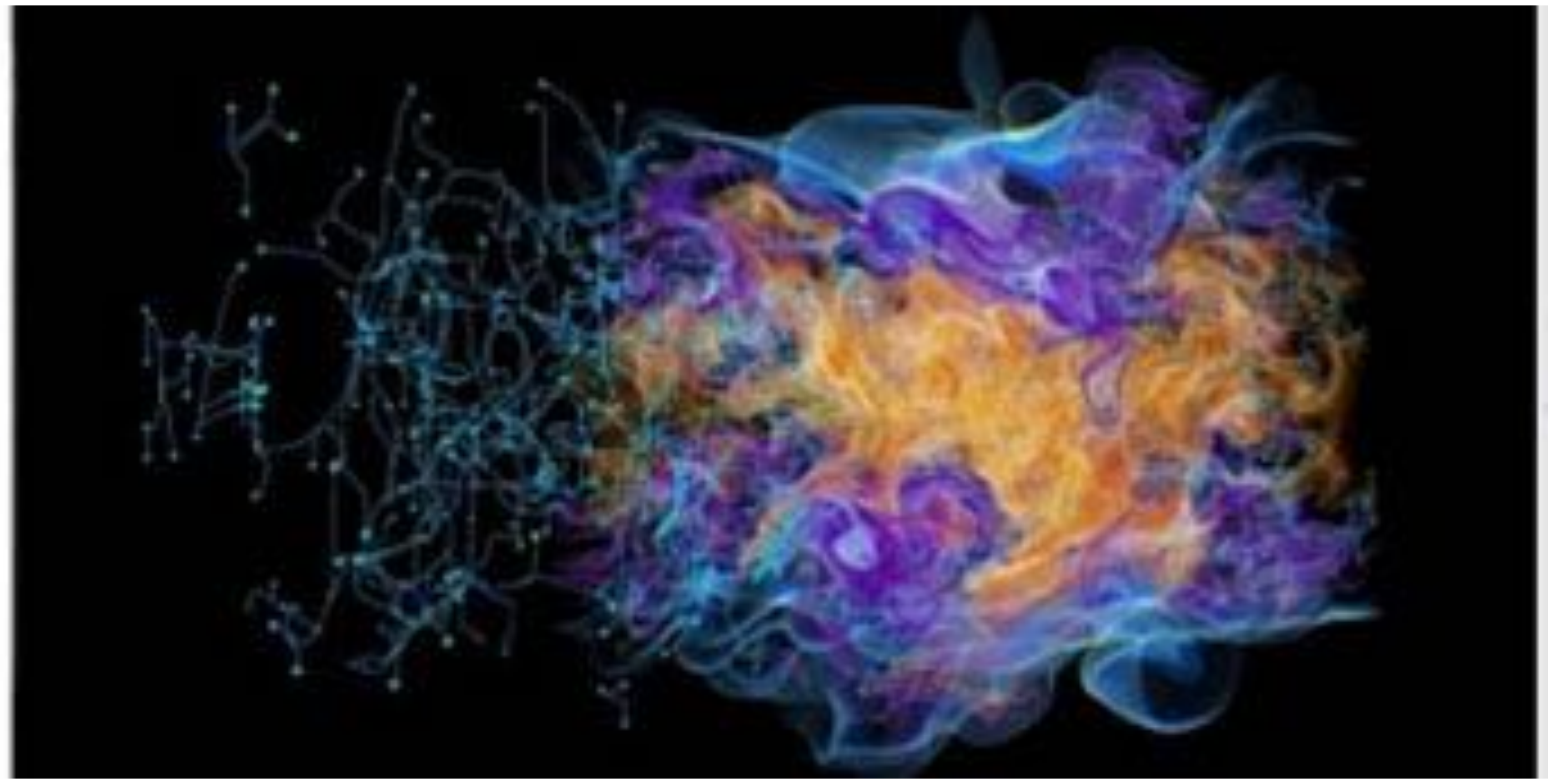
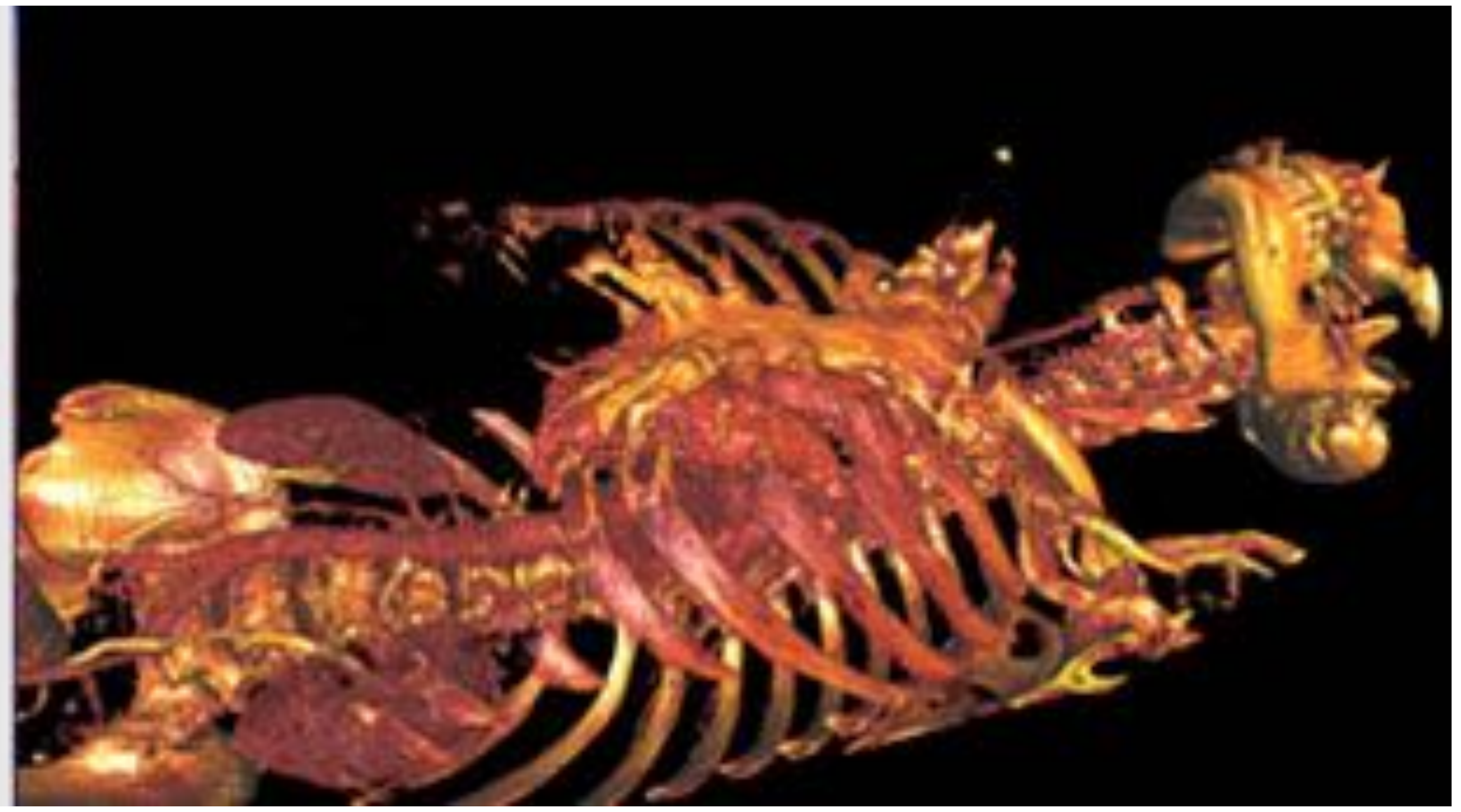





Visualization

Slides adapted from Alex Lex (Univ. of Utah),
Joshua A. Levine (Univ. of Arizona), Carlos Scheidegger (Univ. of Arizona)
and others.

What is Visualization?

“a cognitive process performed by humans in forming a mental image of a domain space. In computer and information science it is, more specifically, the ***visual*** representation of a domain space using ***graphics, images, animated sequences***, ~~and sound augmentation~~ to present the data, structure, and dynamic behavior of large, complex data sets that represent systems, events, processes, objects, and concepts”

J. G. Williams, K. M. Sochats, and E. Morse. “Visualization.” Annual Review of Information Science and Technology (ARIST) 30 (1995), 161–207

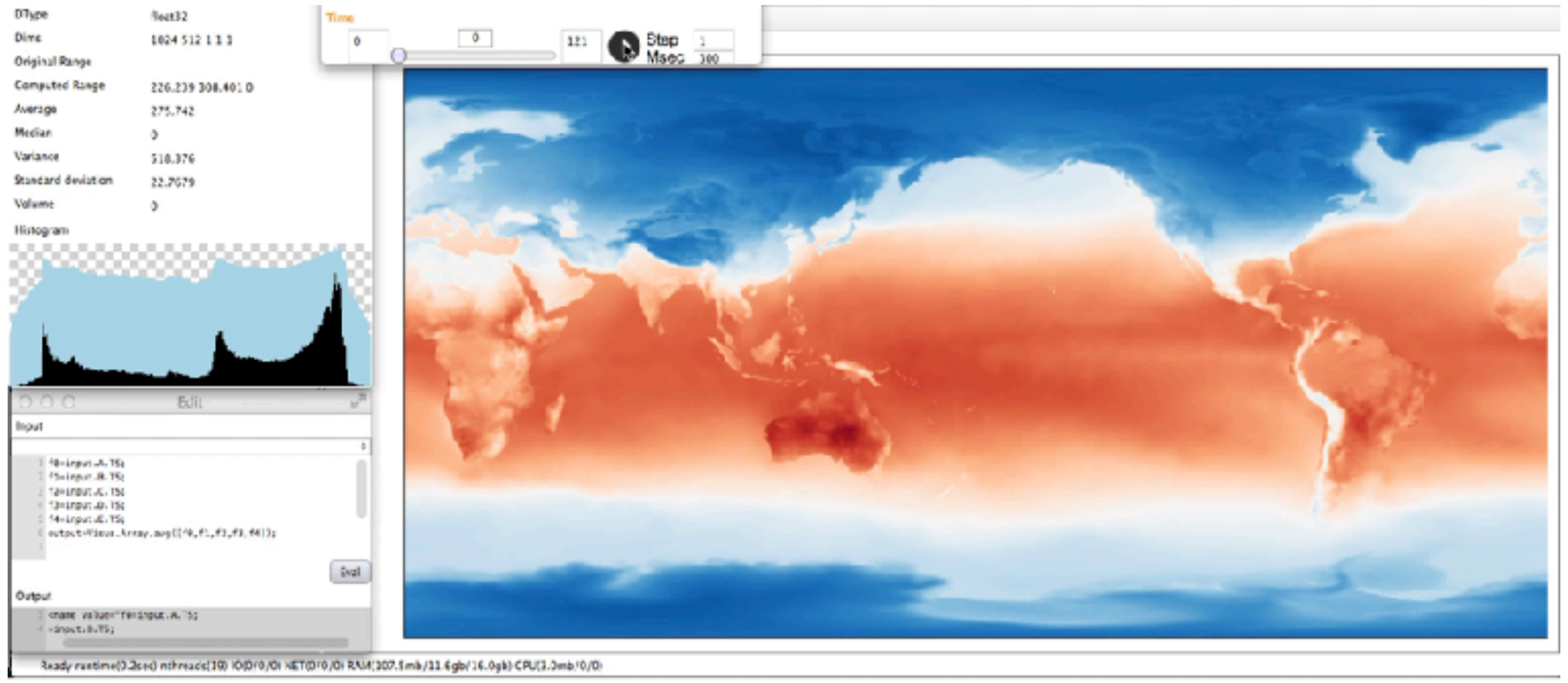
Computer-based visualization systems provide ***visual*** representations of datasets designed to help people carry out tasks more effectively.

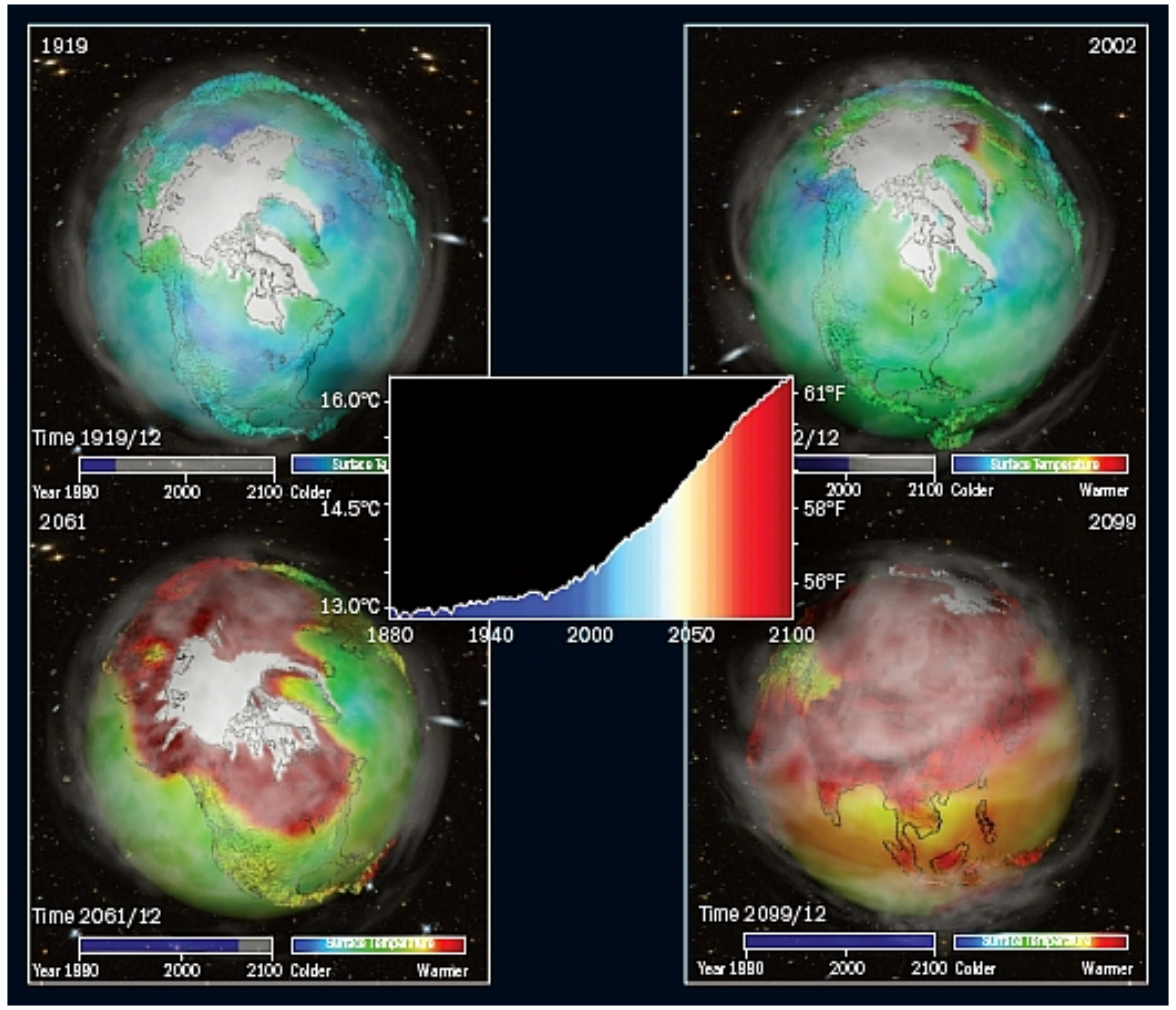
- Our book

Visualization is computer graphics to aid understanding of data

Why do we use visualization?

- To inform
- To communicate
- To explore

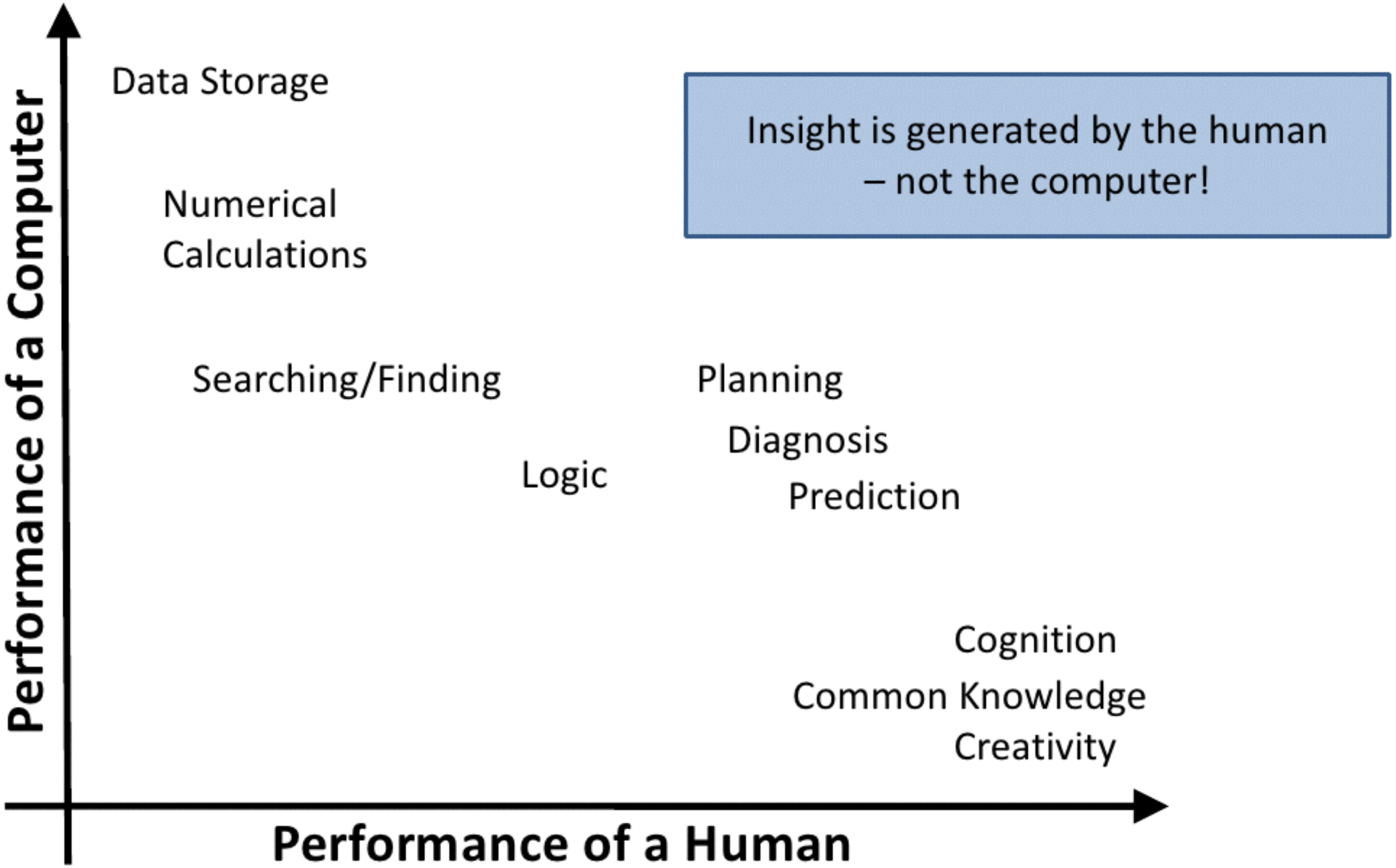




Why have a human in the loop?

- don’t need vis when fully automatic solution exists and is trusted
  - many analysis problems are ill-specified
- don’t know exactly what questions to ask in advance
- possibilities
  - long-term use for end users (e.g. exploratory analysis of scientific data)
  - presentation of known results
  - stepping stone to better understanding of requirements before developing models
  - help developers of automatic solution refine/debug, determine parameters
  - help end users of automatic solutions verify, build trust

The Ability Matrix



Why have a human in the loop?

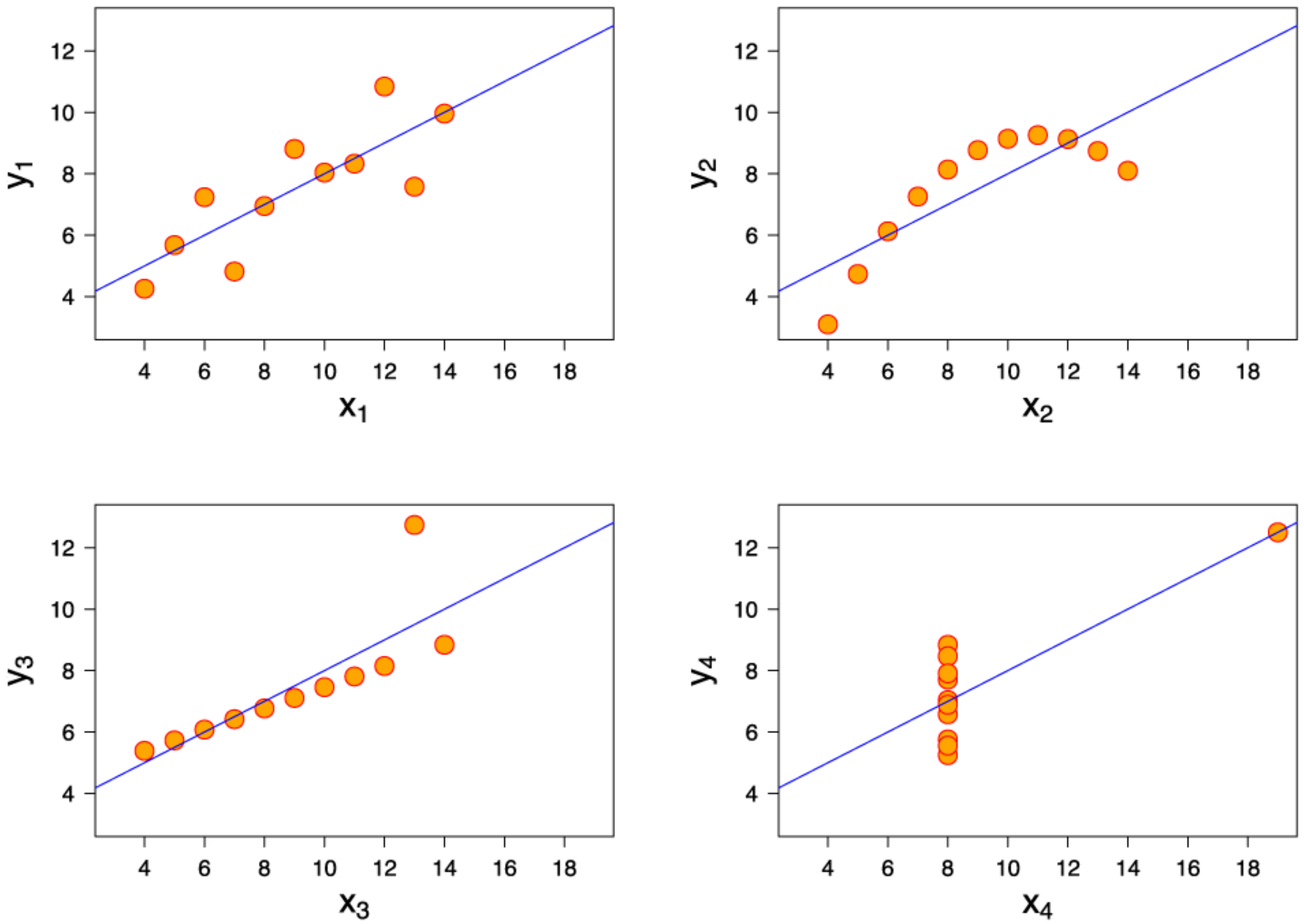
Machine Learning is just Math. Sometimes it makes sense …

Google Translate

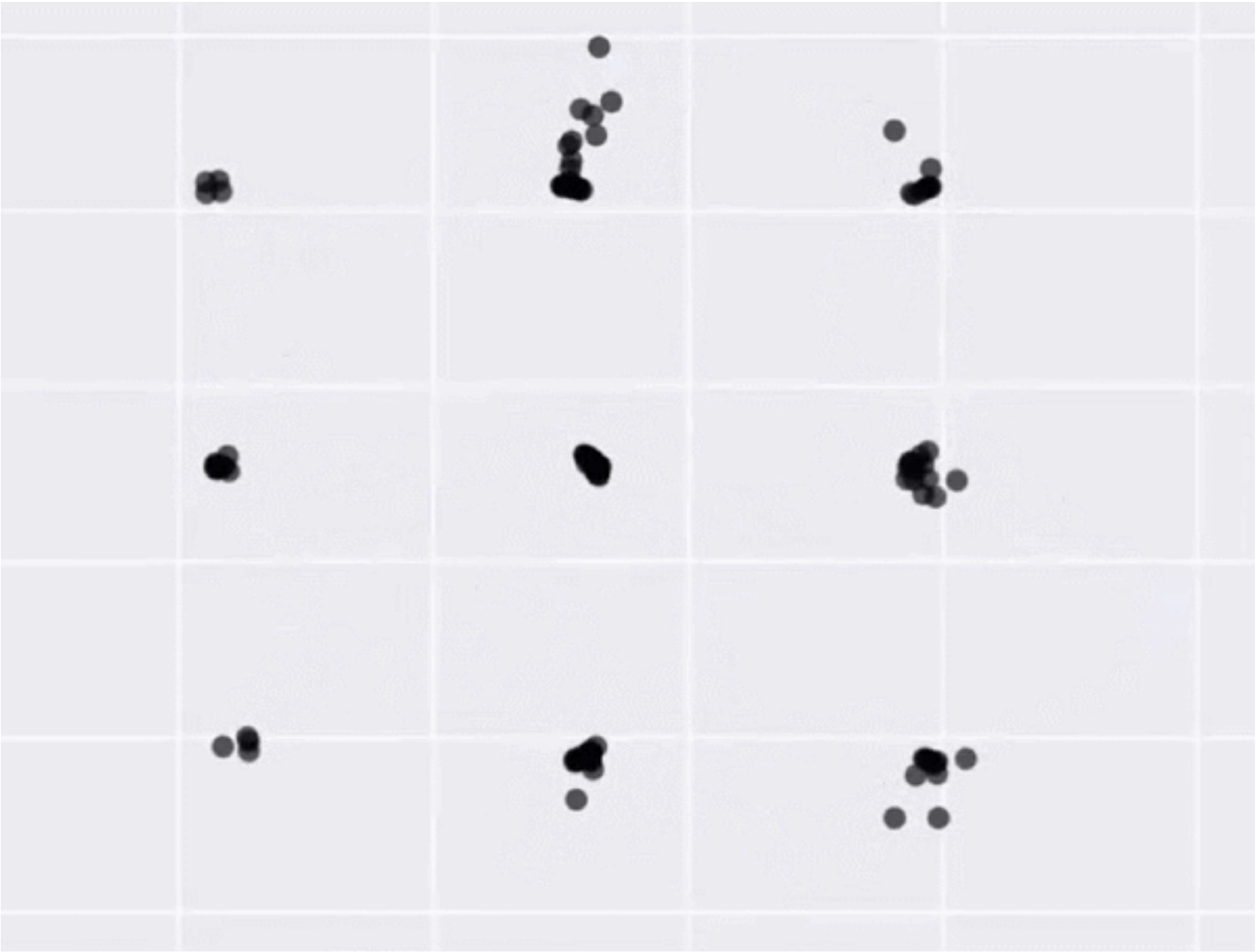
Francis Anscombe's Quartet (1973)

	Set A		Set B		Set C		Set D	
	X	Y	X	Y	X	Y	X	Y
0	10	8.04	10	9.14	10	7.46	8	6.58
1	8	6.95	8	8.14	8	6.77	8	5.76
2	13	7.58	13	8.74	13	12.74	8	7.71
3	9	8.81	9	8.77	9	7.11	8	8.84
4	11	8.33	11	9.26	11	7.81	8	8.47
5	14	9.96	14	8.10	14	8.84	8	7.04
6	6	7.24	6	6.13	6	6.08	8	5.25
7	4	4.26	4	3.10	4	5.39	19	12.50
8	12	10.84	12	9.13	12	8.15	8	5.56
9	7	4.82	7	7.26	7	6.42	8	7.91
10	5	5.68	5	4.74	5	5.73	8	6.89
mean	9.00	7.50	9.00	7.50	9.00	7.50	9.00	7.50
std	3.32	2.03	3.32	2.03	3.32	2.03	3.32	2.03
corr	0.82		0.82		0.82		0.82	
lin. reg.	$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$	

Anscombe's Quartet



Datasaurus Dozen



$$\mu_x = 54.02$$
$$\mu_y = 48.09$$
$$\sigma_y = 14.52$$
$$\sigma_y = 24.79$$
$$corr = \pm 0.32$$

Matejka, Justin, and George Fitzmaurice. "Same stats, different graphs: generating datasets with varied appearance and identical statistics through simulated annealing." Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 2017.

Human Eye Bandwidth Analogy

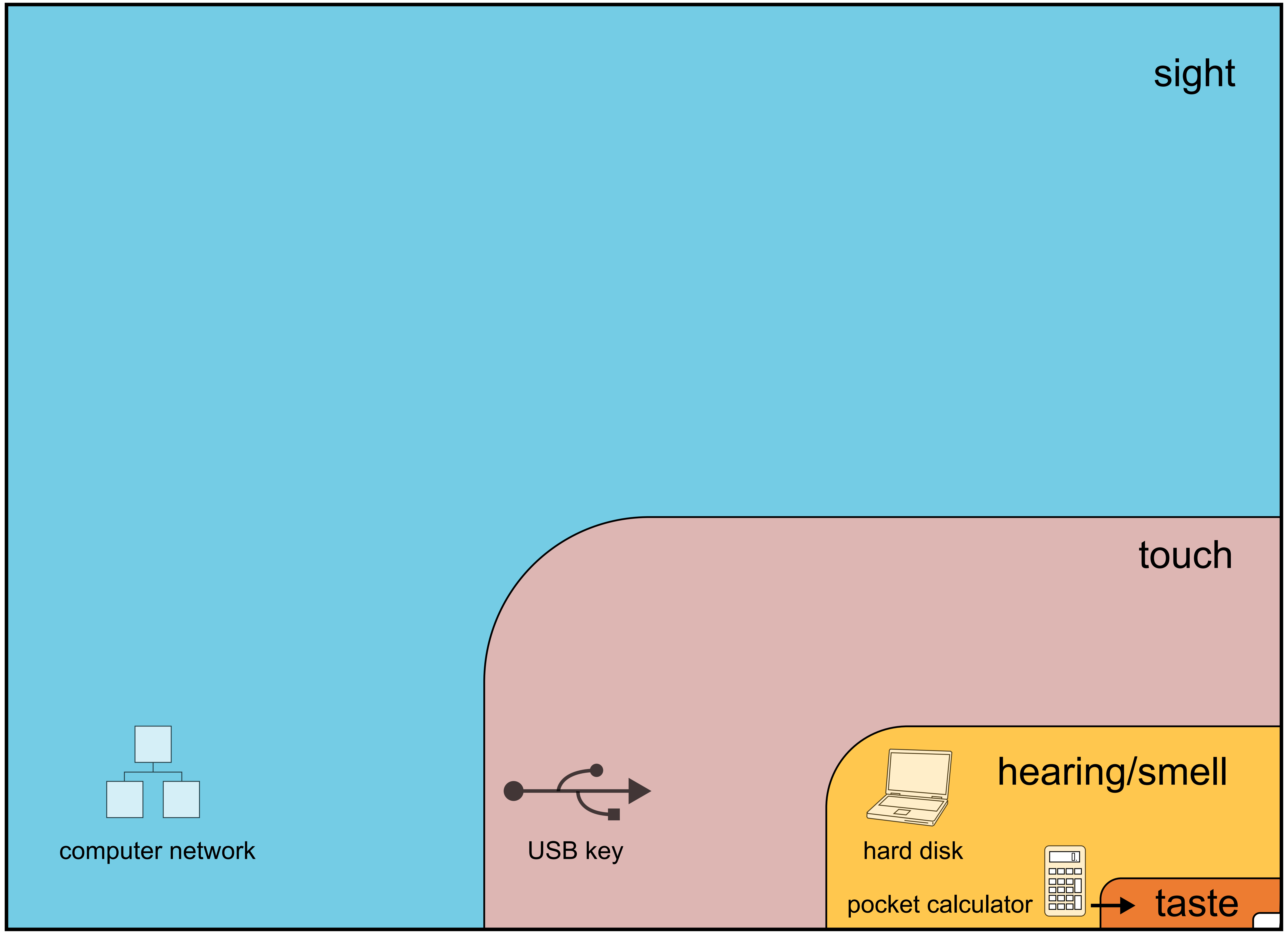


Image courtesy Tor Norretranders

amount we're actually aware of (0.7%)

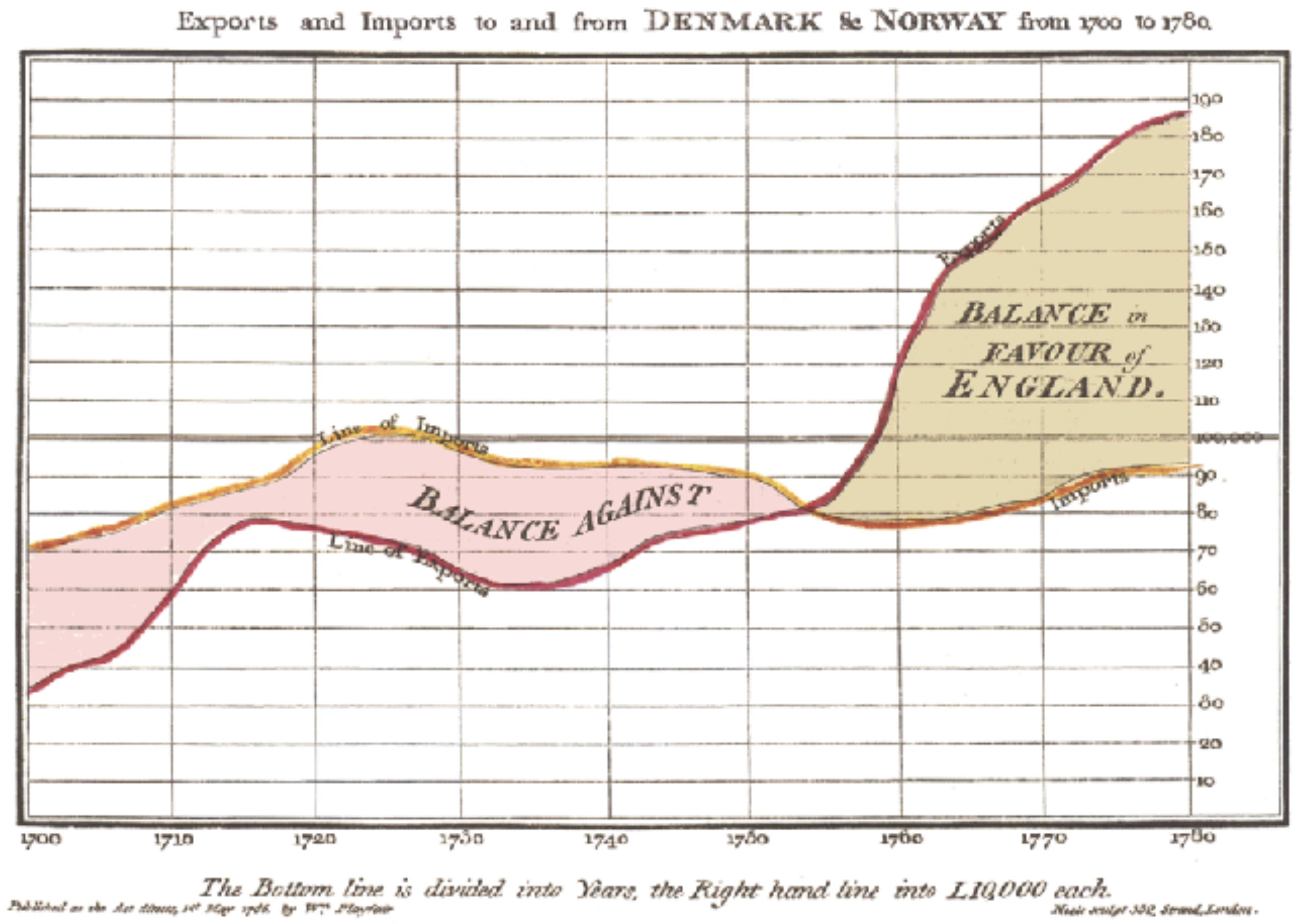
History

Edmond Halley (1686) Current Maps

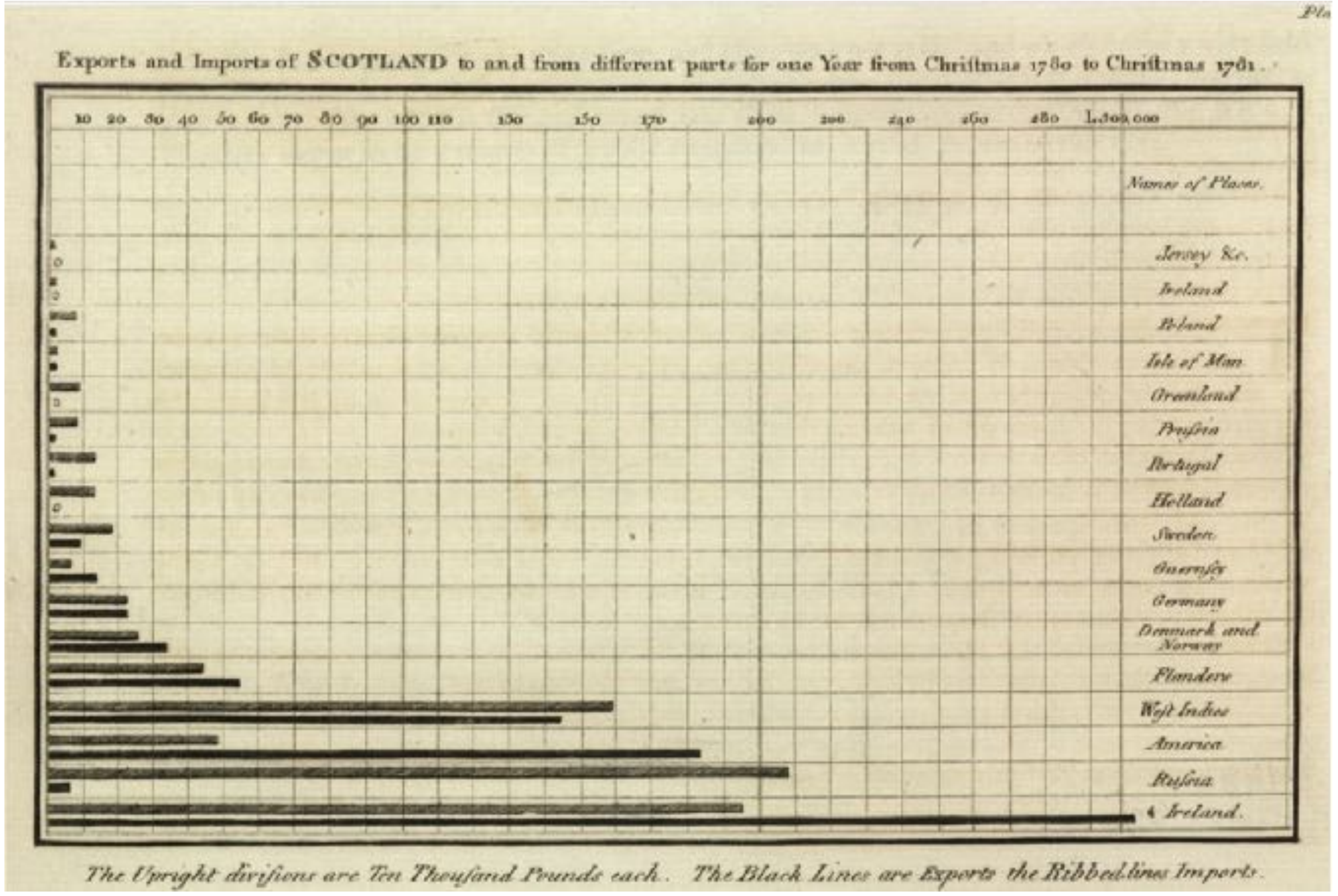
- Edmond Halley 1686

William Playfair

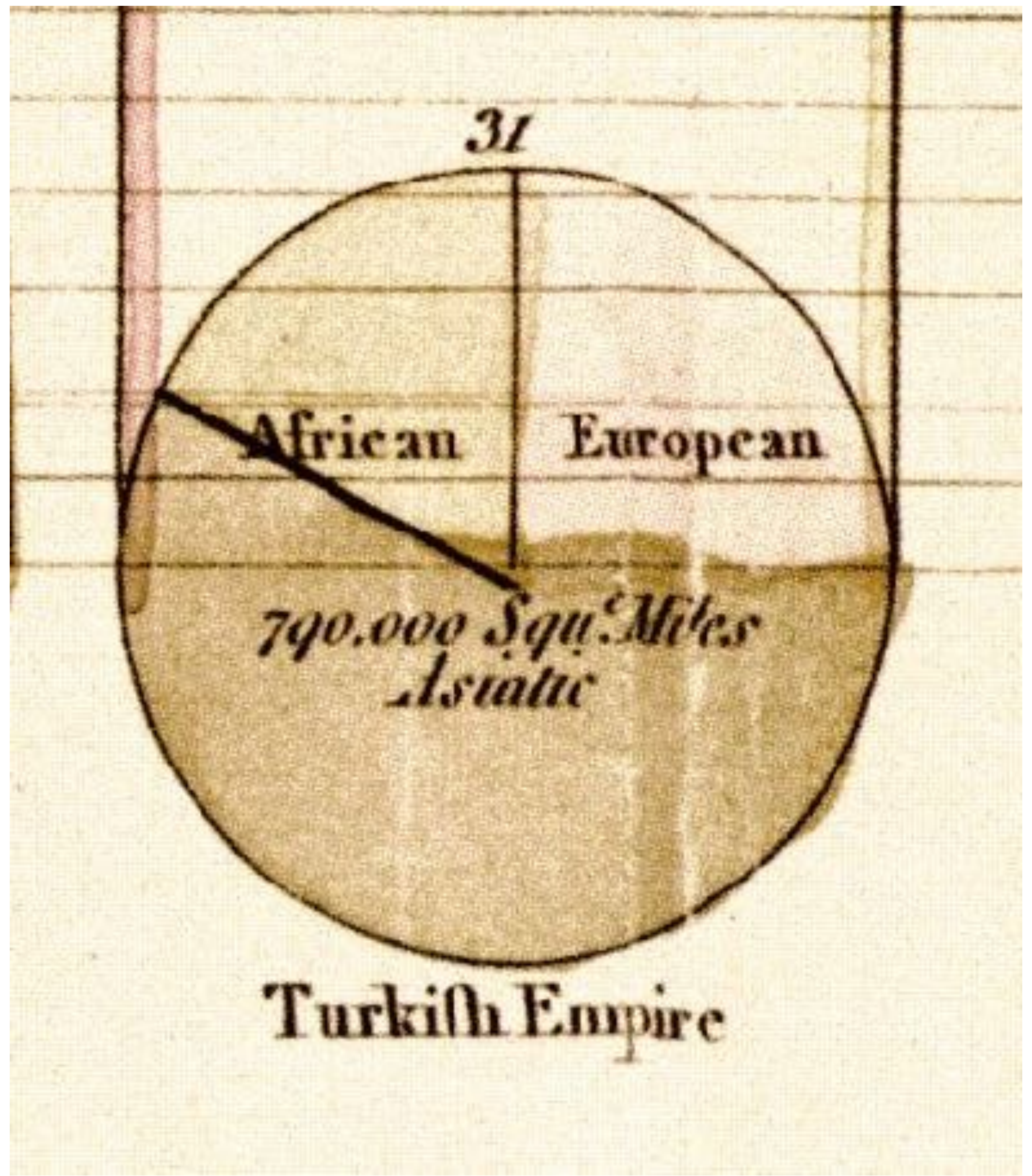
Time series



Bar

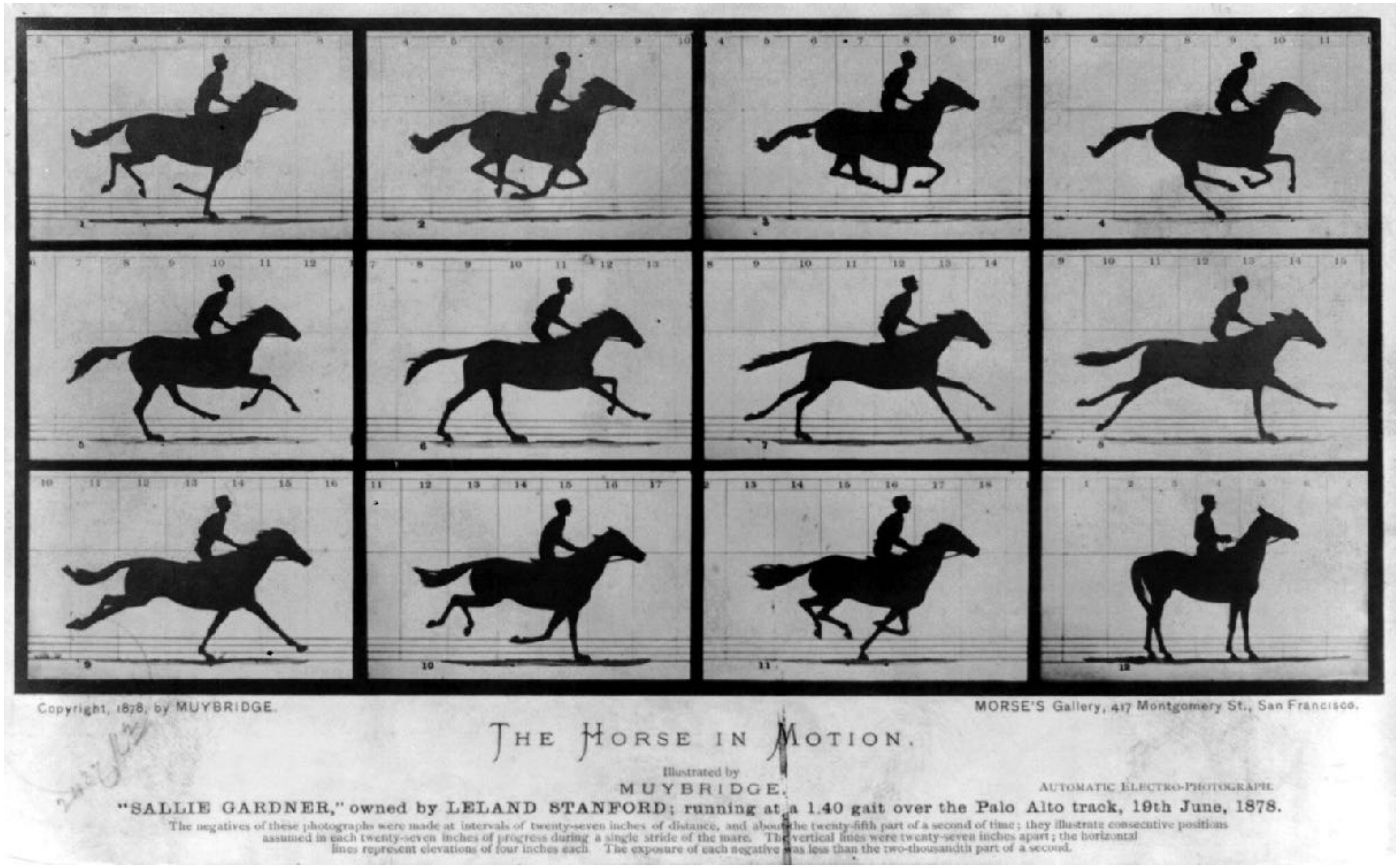


Pie



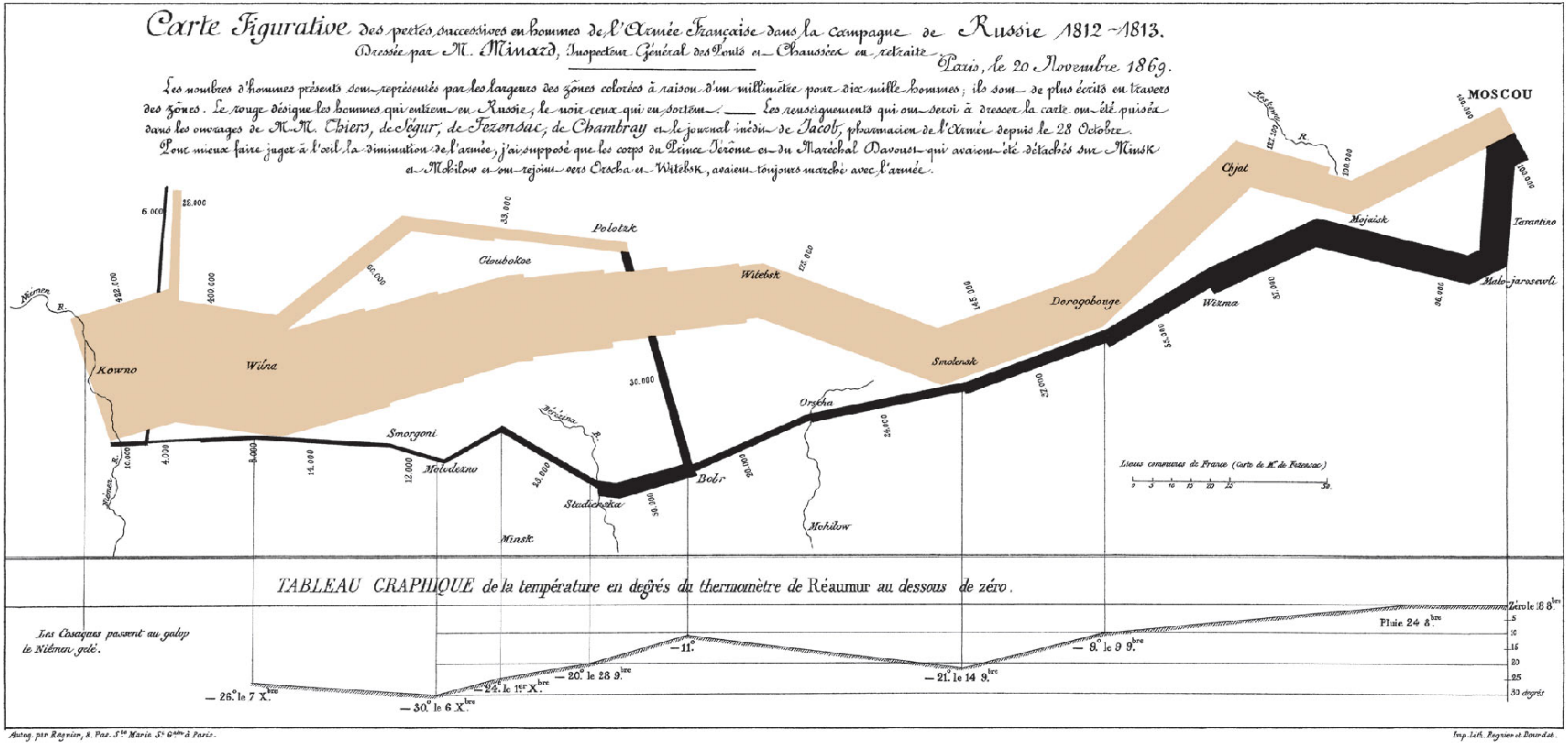
- 1786, 1789, 1801

Eadweard Muybridge (1878), The Horse in Motion



John Snow (1854), Cholera Epidemic

Charles Minard (1869), Napoleon's Russian Campaign



- Encodes troop numbers, temperature, distances, location, directions, and time

Perception

What are we good and bad at seeing?

(mostly we'll talk about bad)

Perception vs. Cognition

Perception

- Eye, optical nerve, visual cortex
- Basic perception
- First processing
- (edges, shapes)
- Not conscious
- Reflexes

Cognition

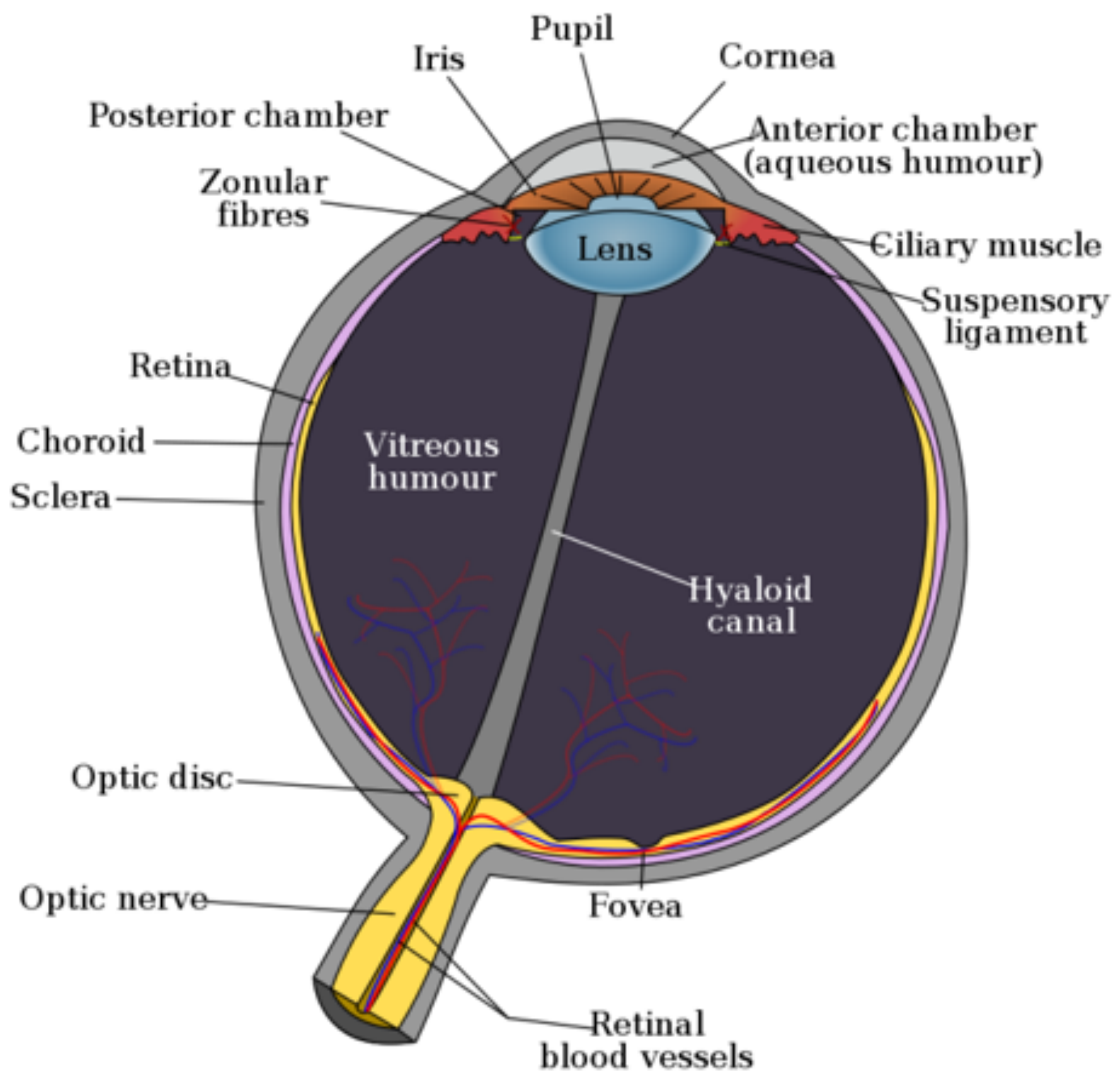
- Recognizing objects
- Relations between objects
- Conclusion drawing
- Problem solving
- Learning, …

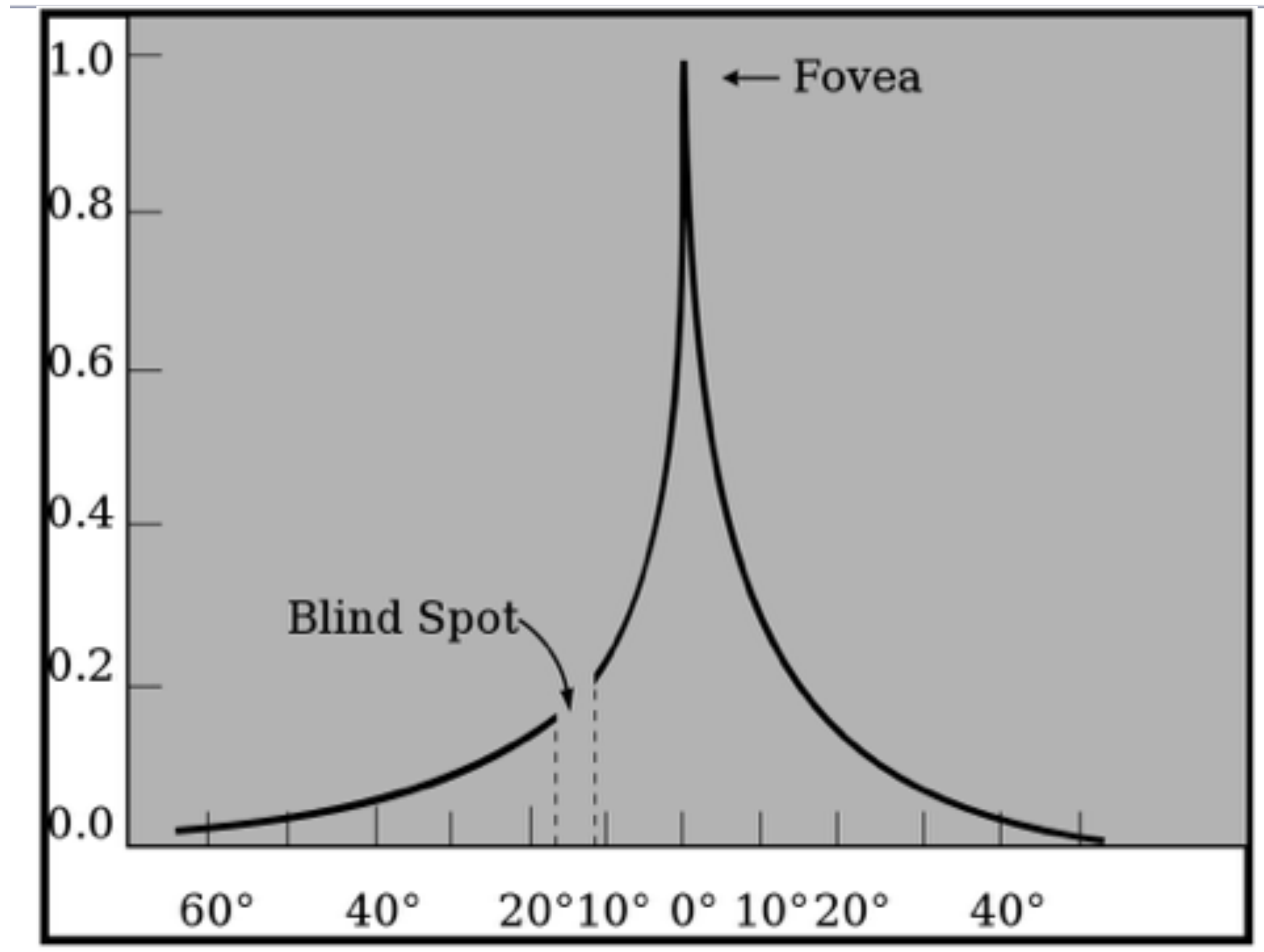
Vision is “constructed” top down from the input

“What you see when you see a thing depends on what the thing ***is***. What you see the thing ***as*** depends on what you know about what you are seeing.” - Zenon Pylyshyn, Canadian Cognitive Scientist and Philosopher

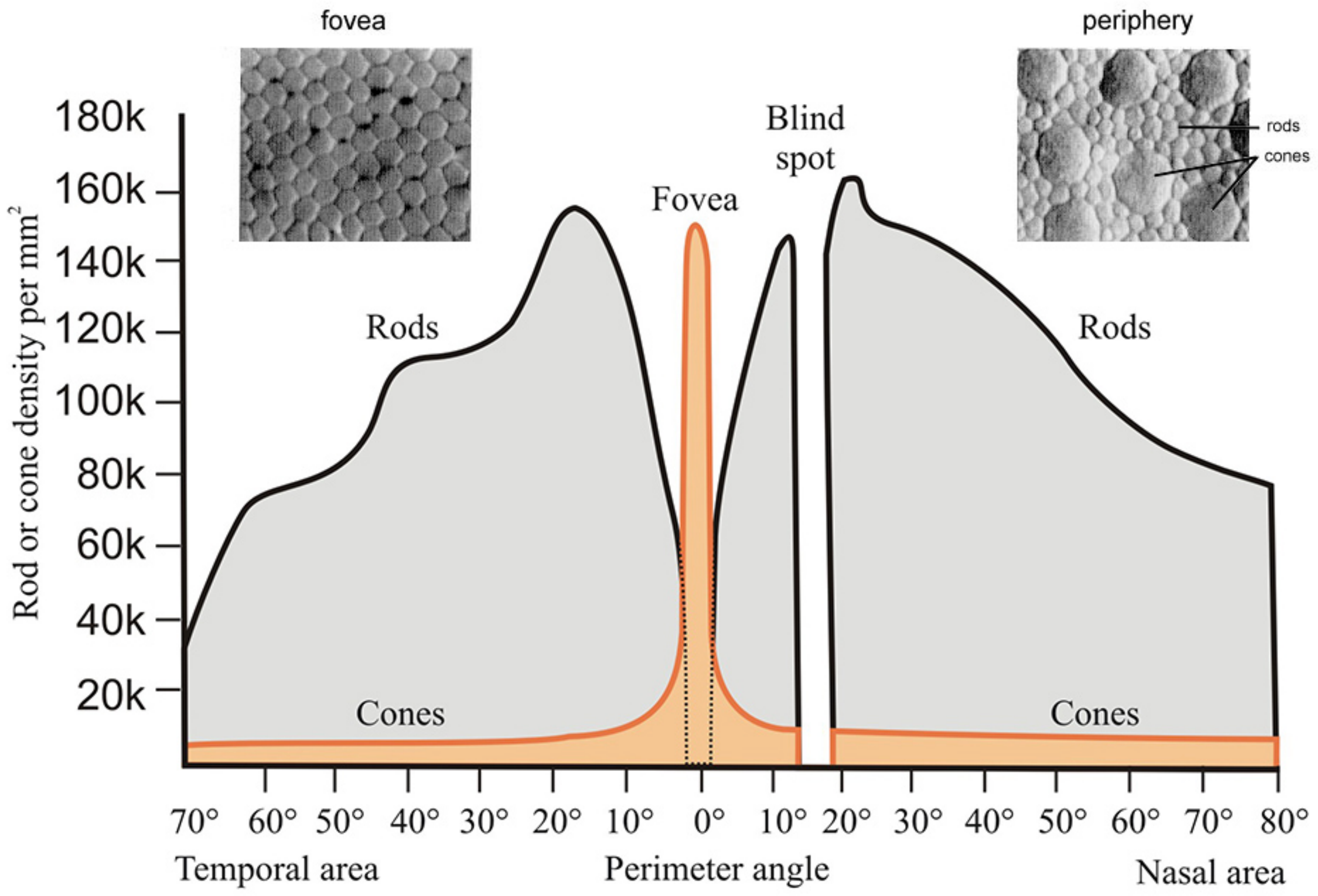
Human Visual System

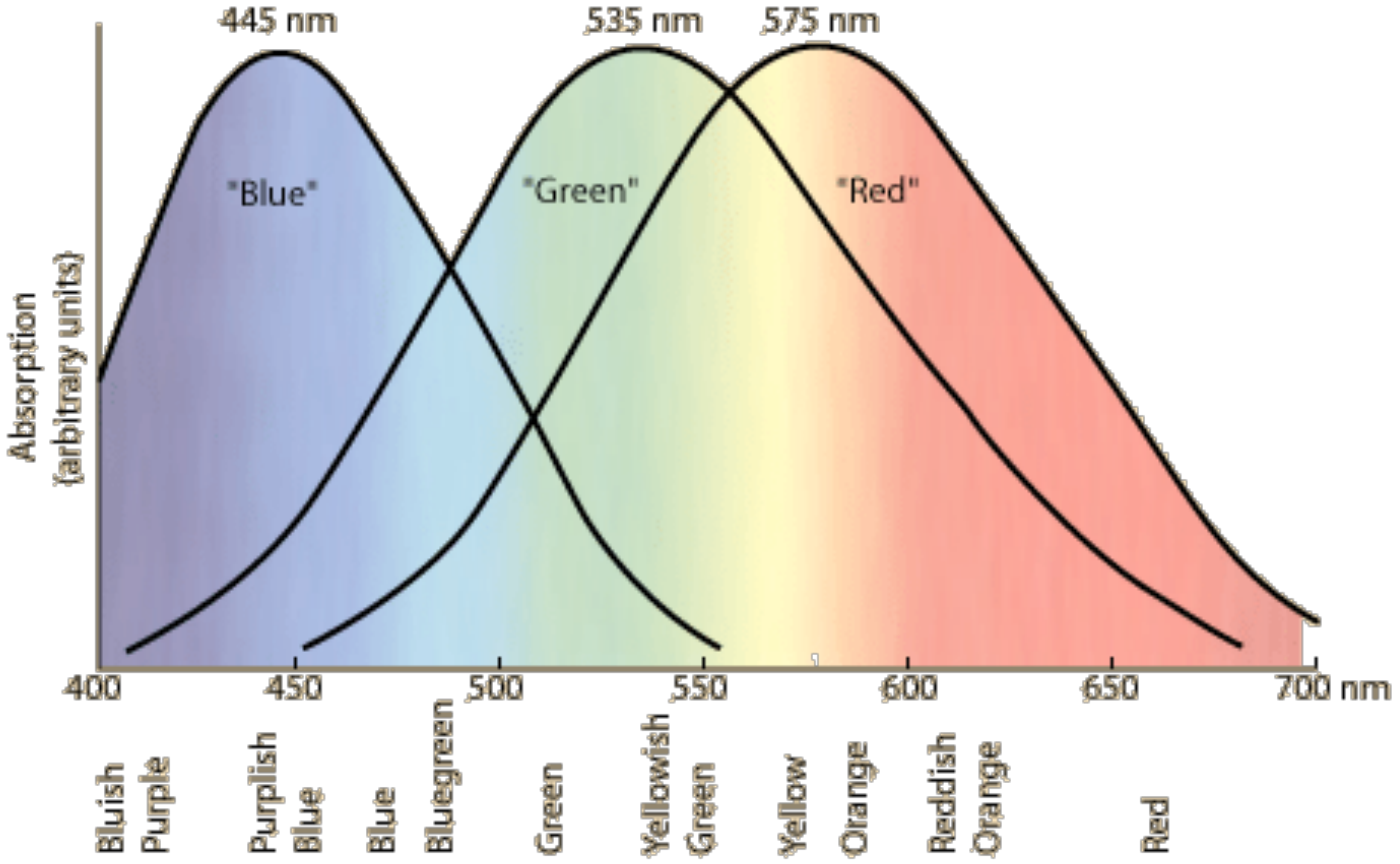
- 5-6 million cones
  - color vision
  - dense in the center
- ~120 million rods
  - light/dark
- Fovea: 27 times the density
  - responsible for sharp central vision
  - only cone cells

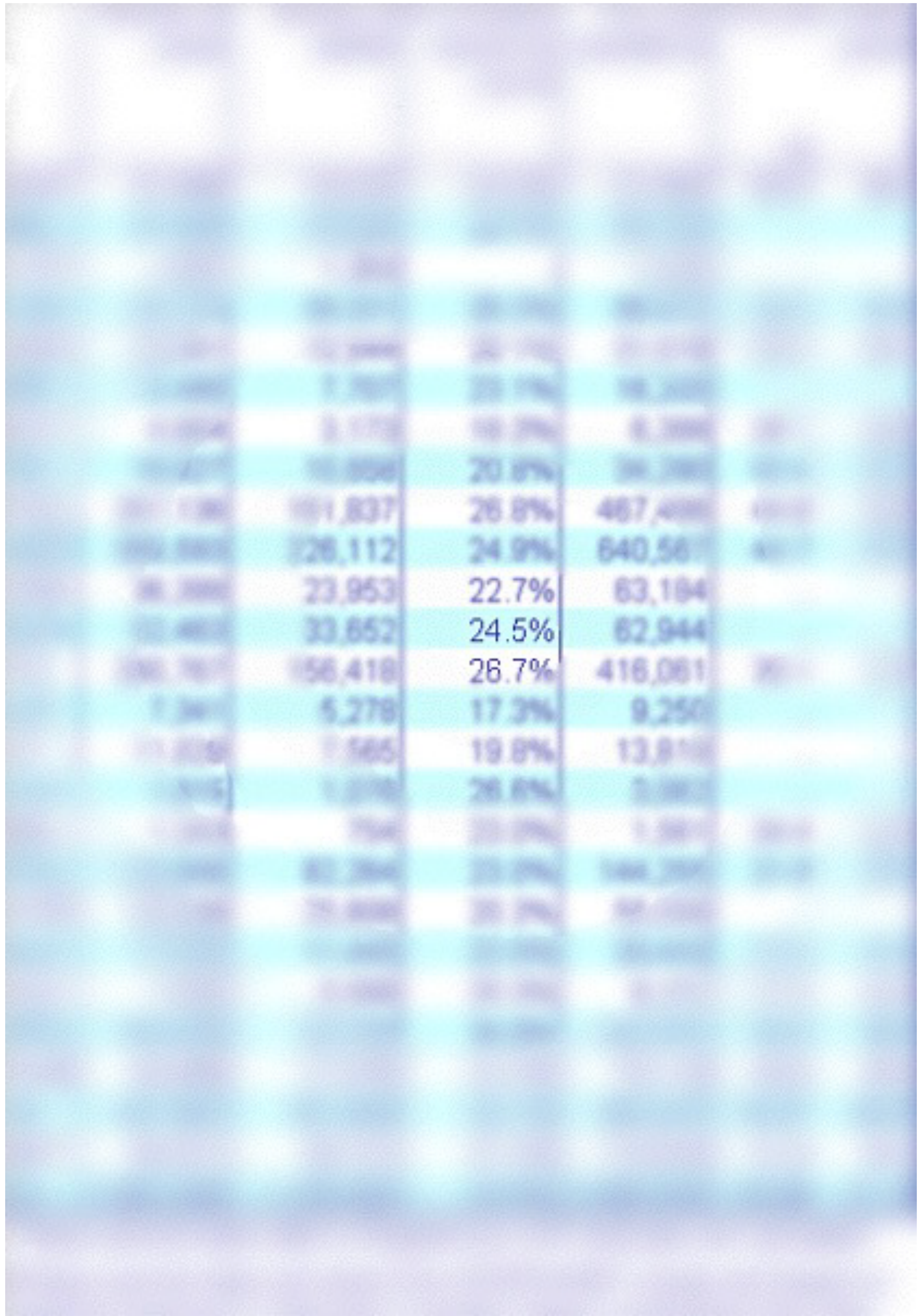




Cone Response



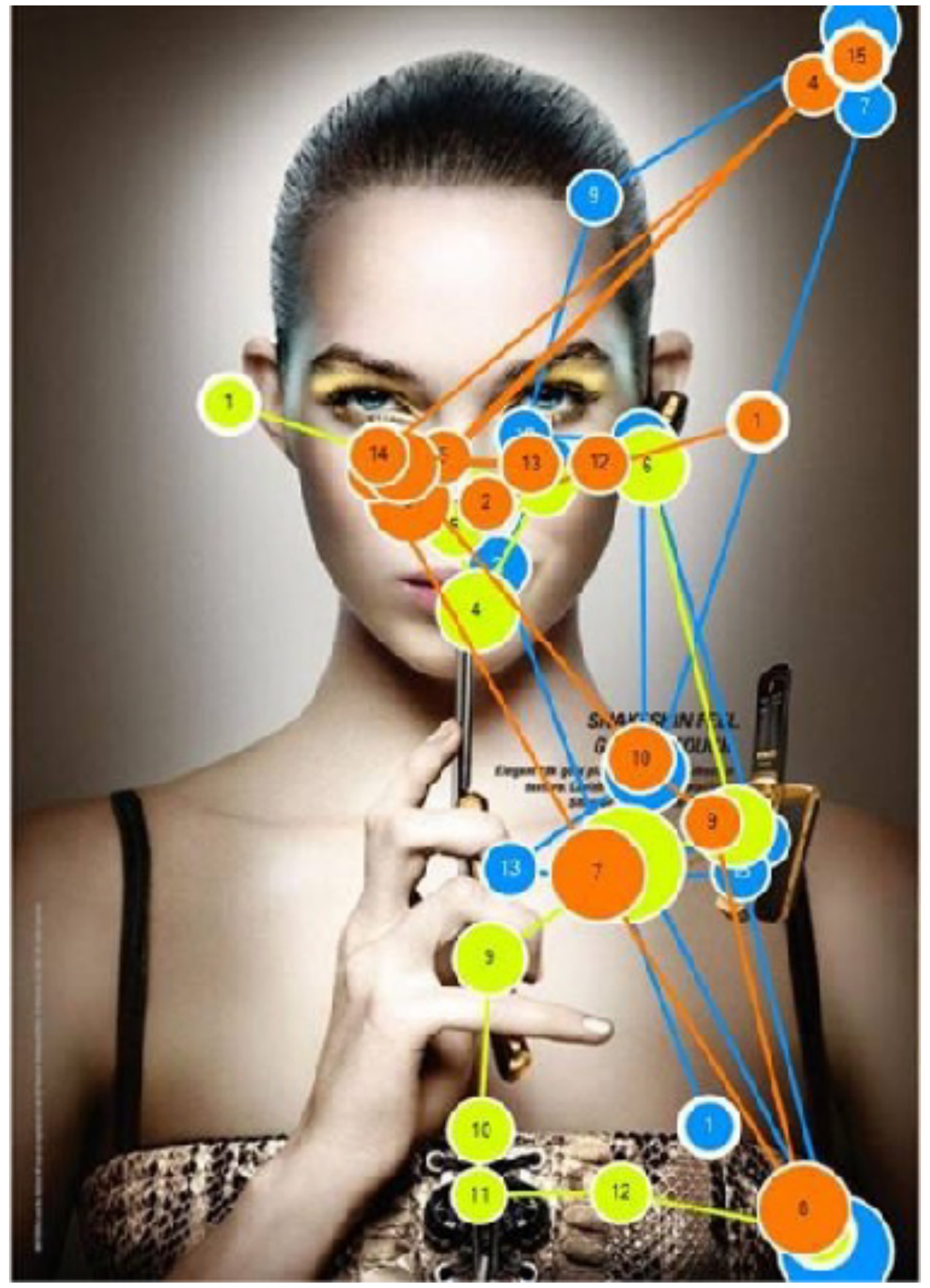


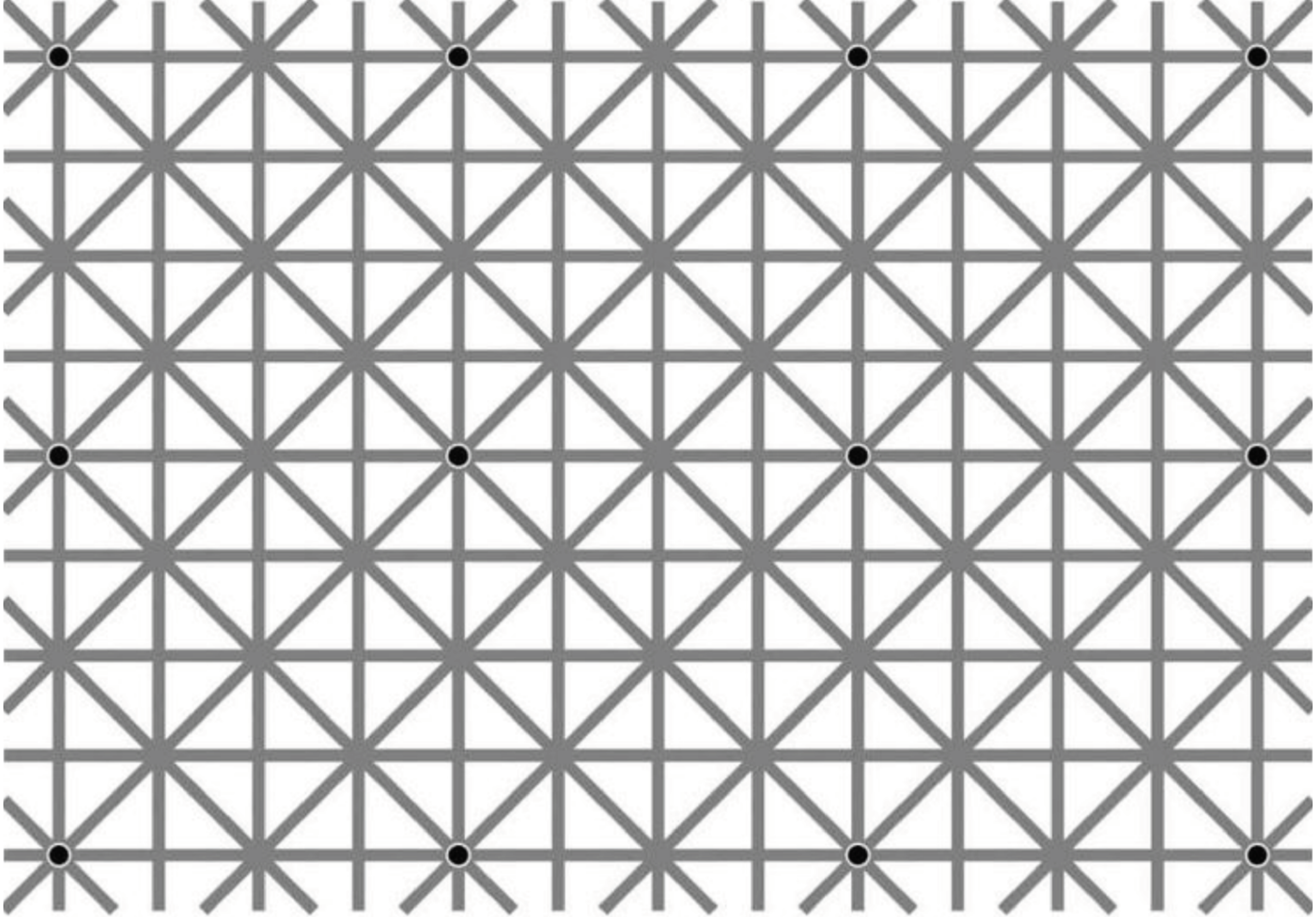


C. Ware, “Visual Thinking for Design”

Human Visual System

- Vision works as sequence of ***fixations*** and ***saccades***
  - ***fixations***: maintaining gaze on single location (200-600 ms)
  - ***saccades***: moving between different locations (20-100 ms)
- Vision not similar to a camera
  - More similar to a dynamic and ongoing construction project

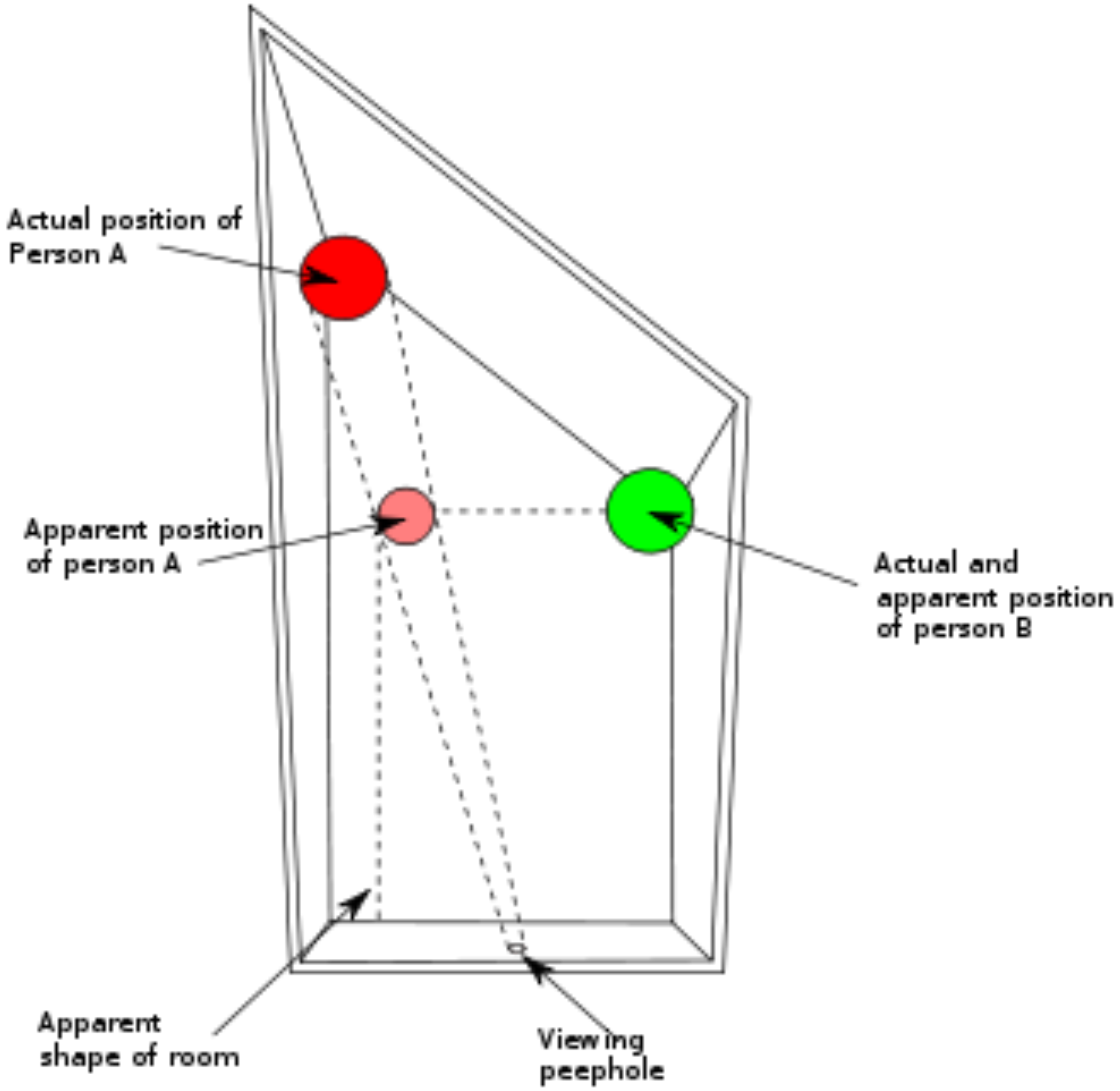


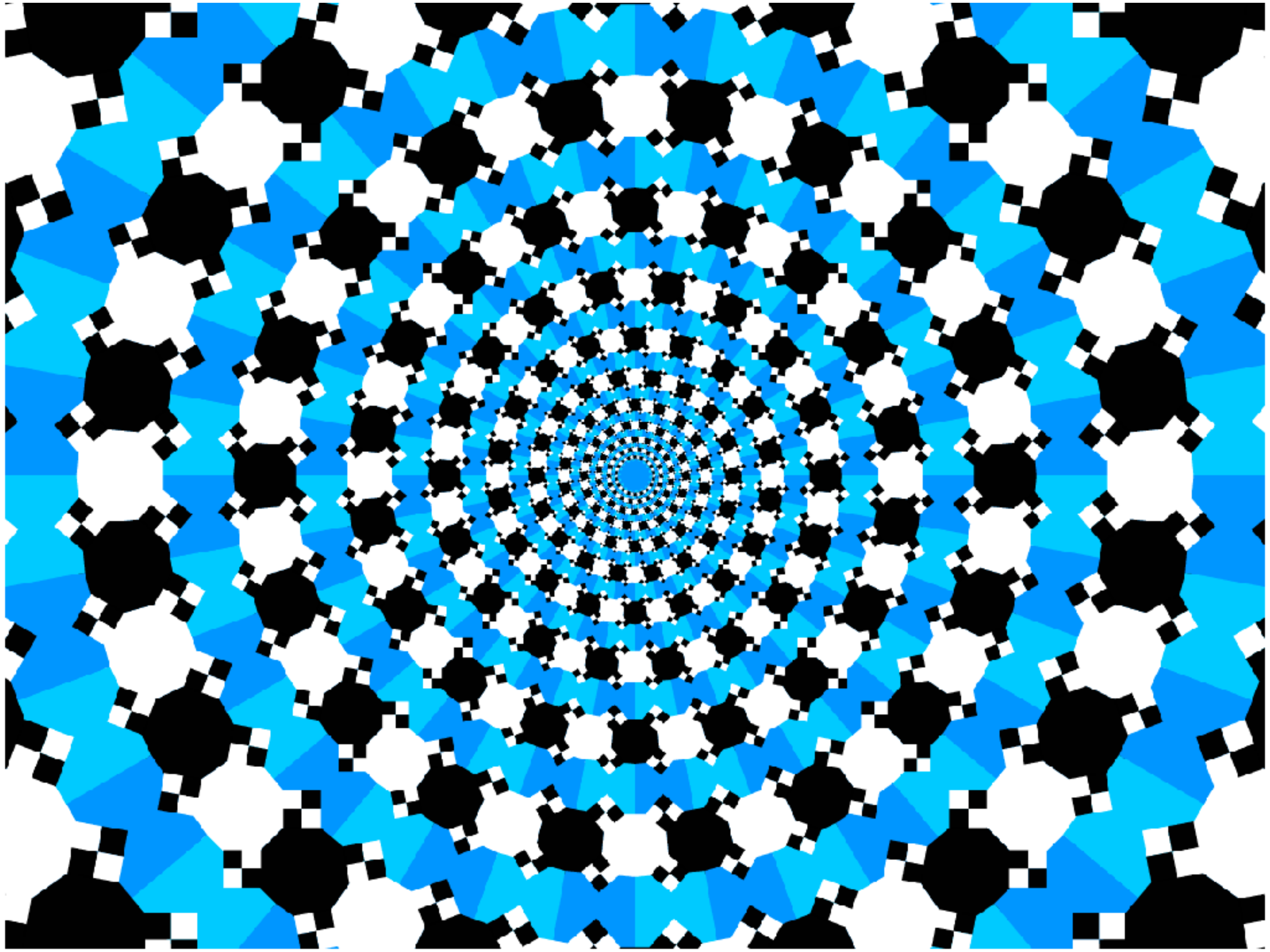


Ninio’s extinction illusion

Human Visual System

- No general purpose vision
  - What we see depends on our goals and expectations
- Relative judgments: strong
- Absolute judgments: weak

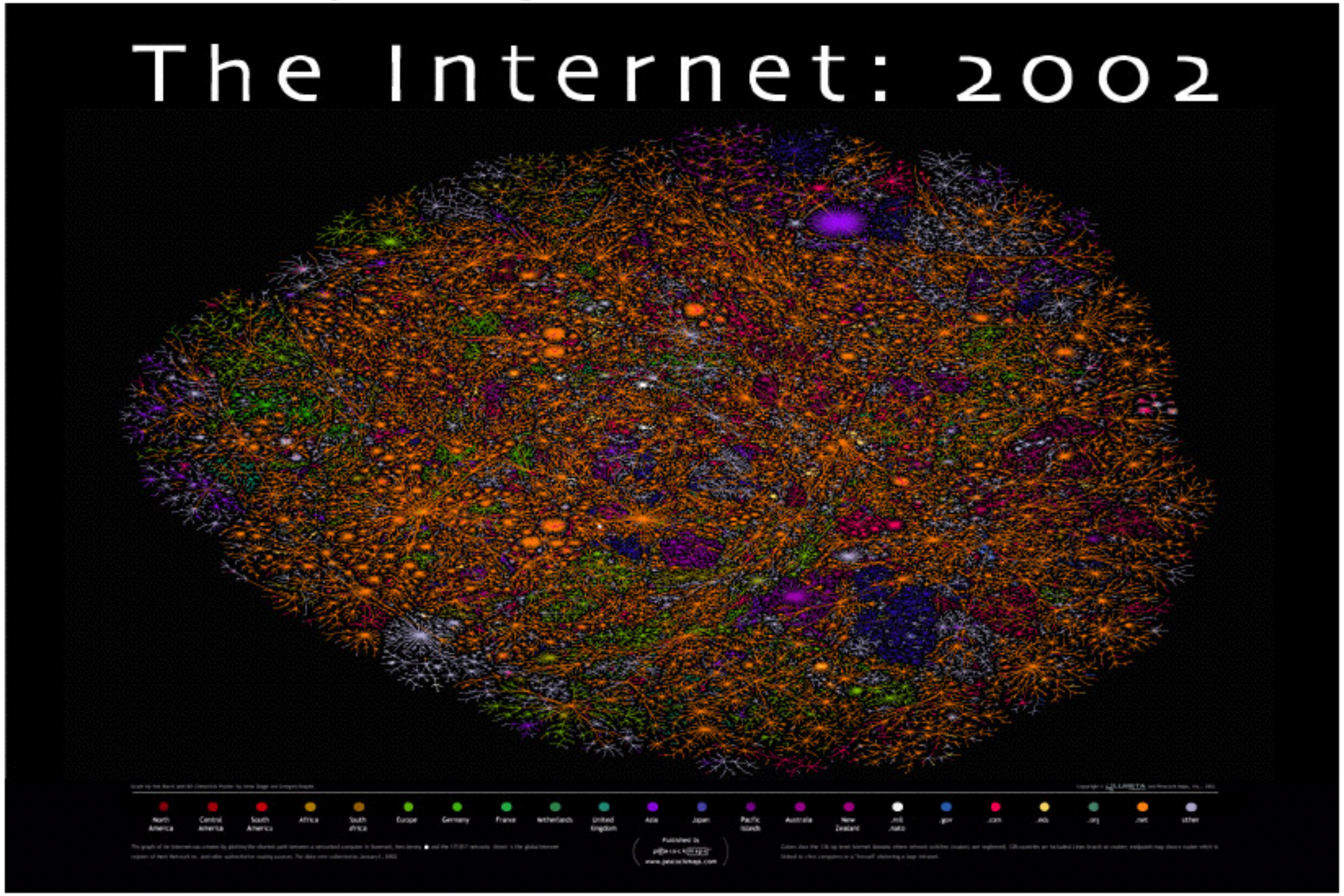




Qualitative Data Vis

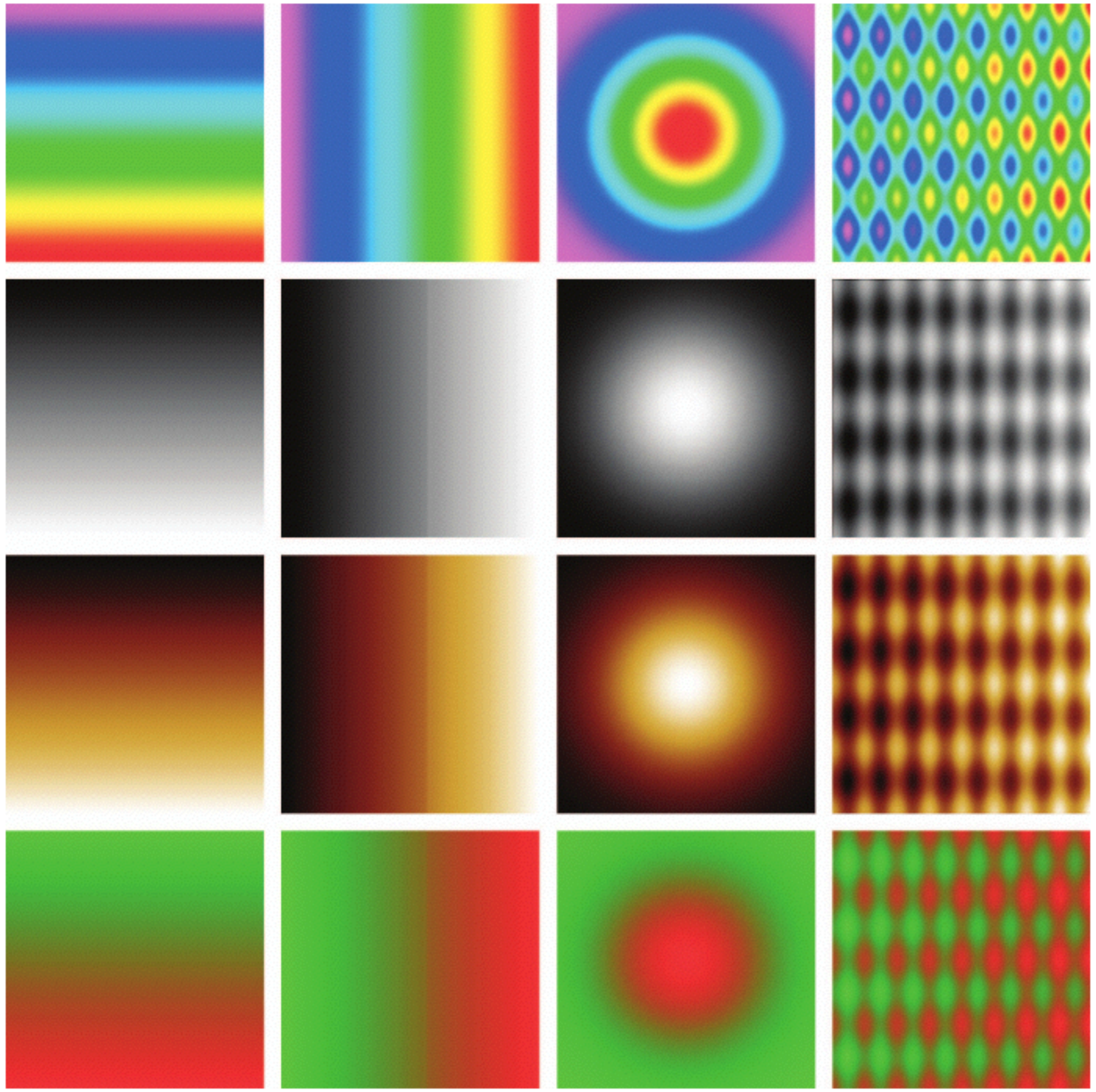
- Color labeling (nominal information coding)
  - recommended: about 6, no more than 10

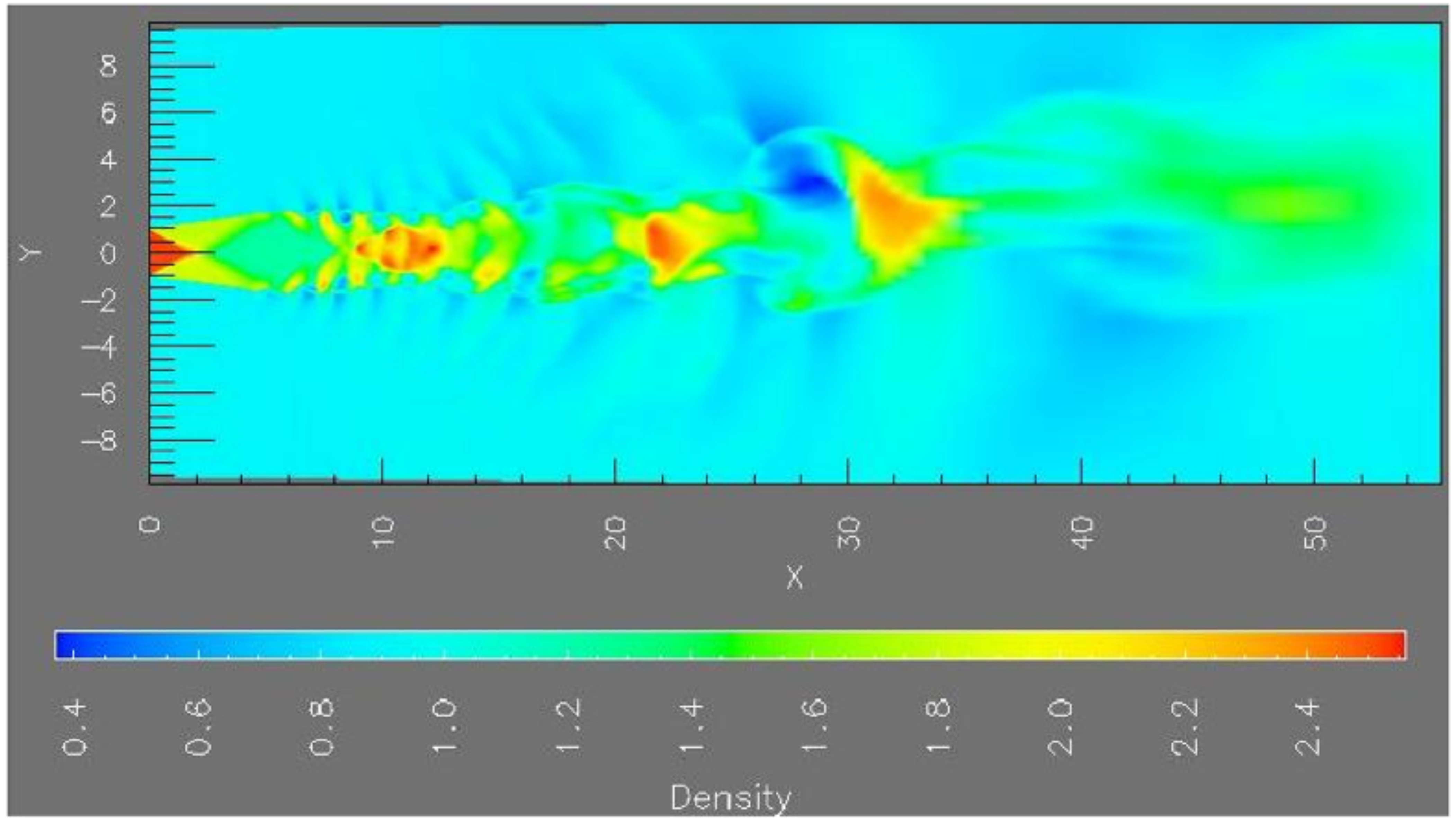
22 colors, but only 8 distinguishable



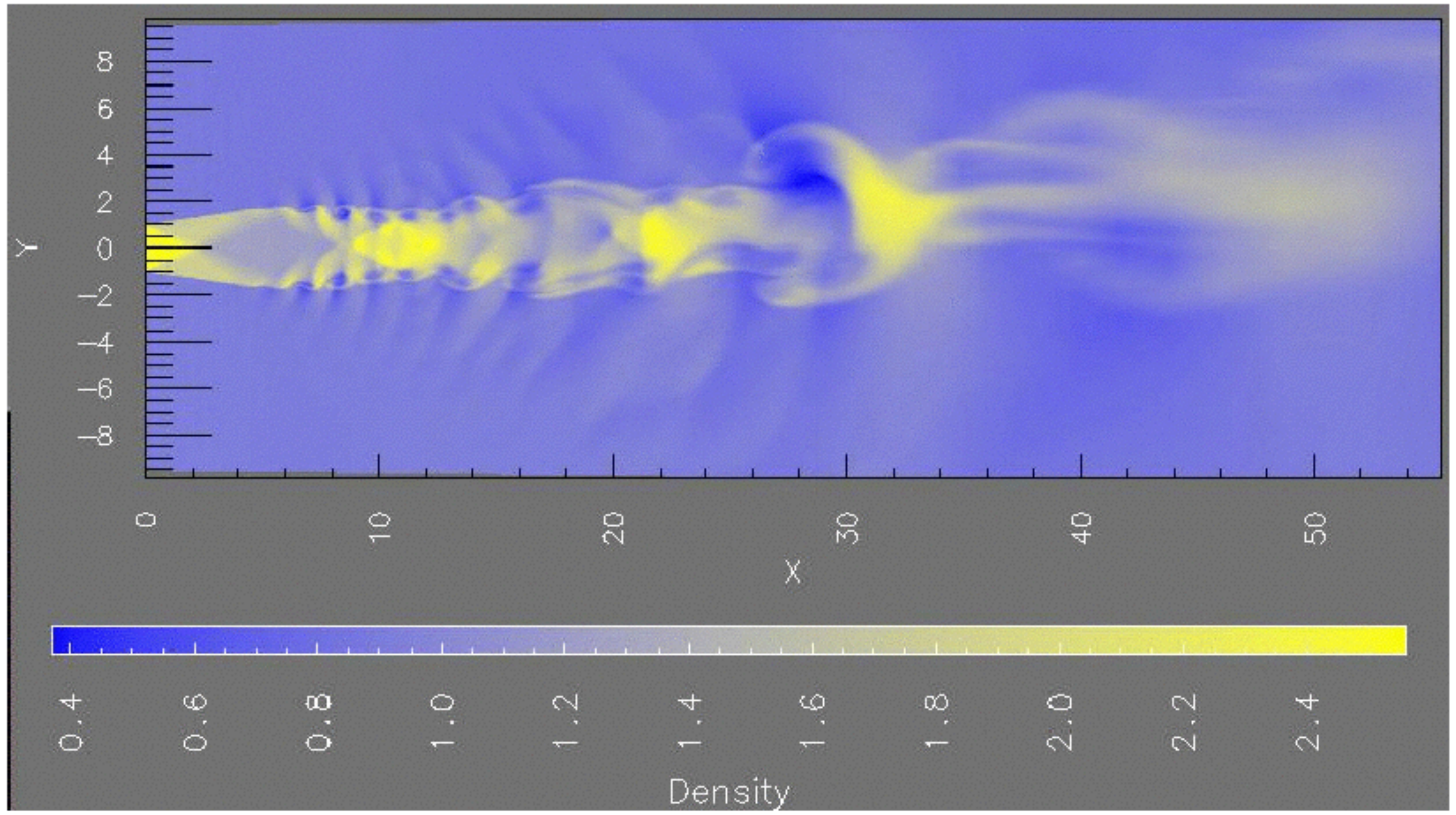
Quantitative Data Vis

- use value
- saturation works but not as good
- avoid hue
- Danger: rainbow color map



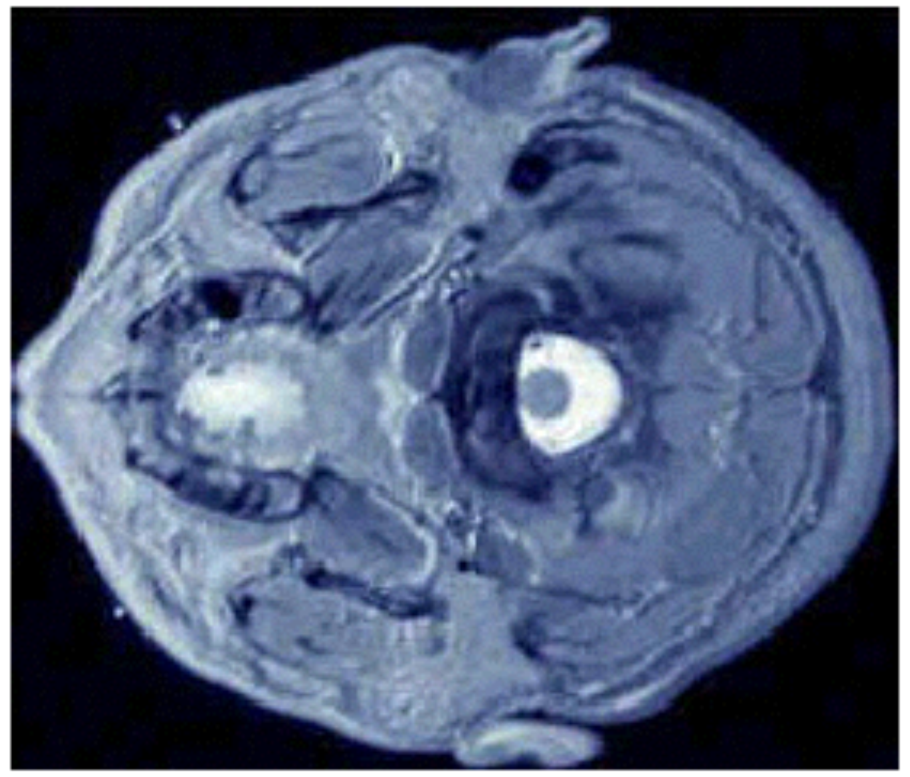


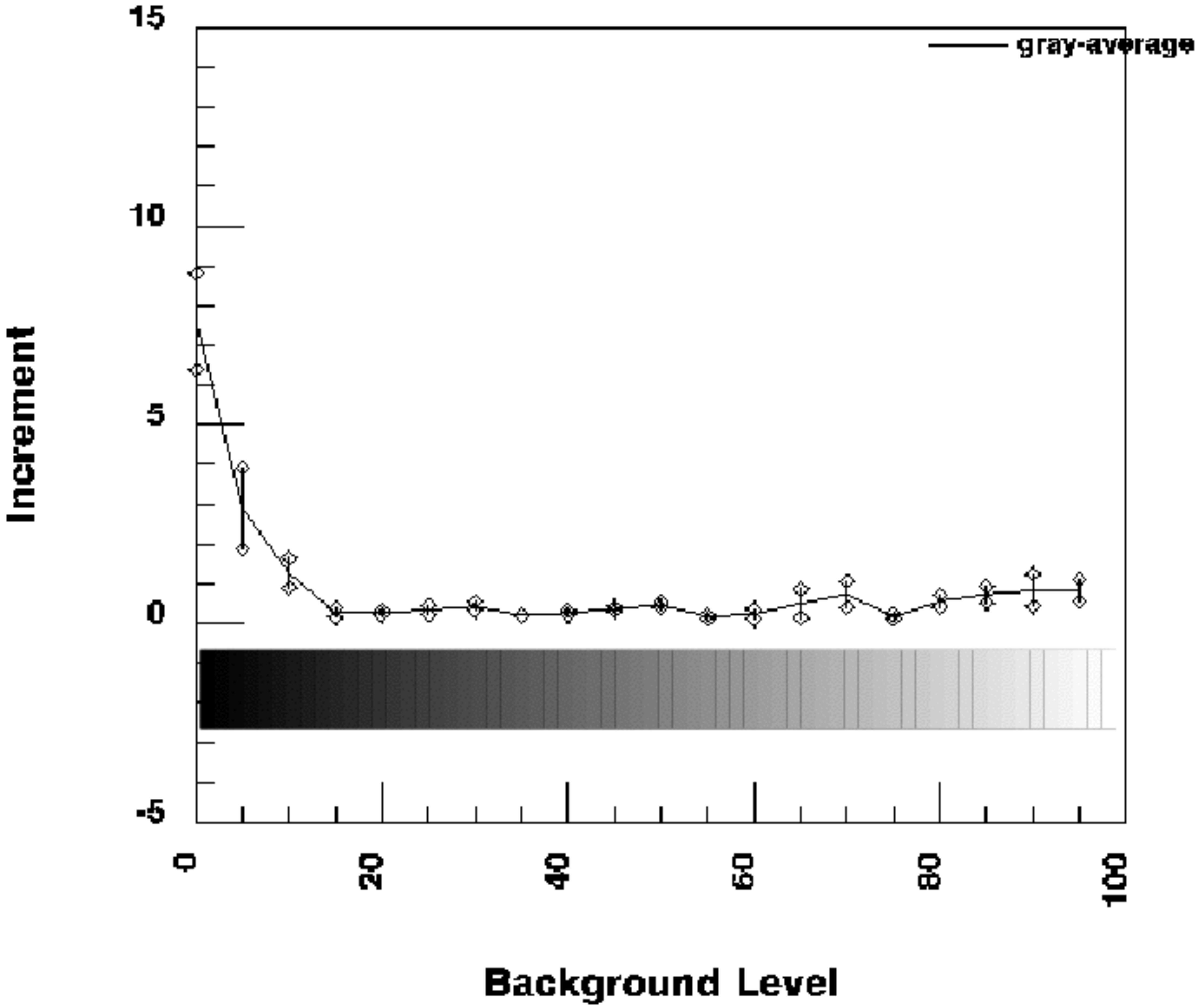
[Rogowitz and Treinish, Why Should Engineers and Scientists Be Worried About Color? http://www.research.ibm.com/people/l/lloydt/color/color.HTM]



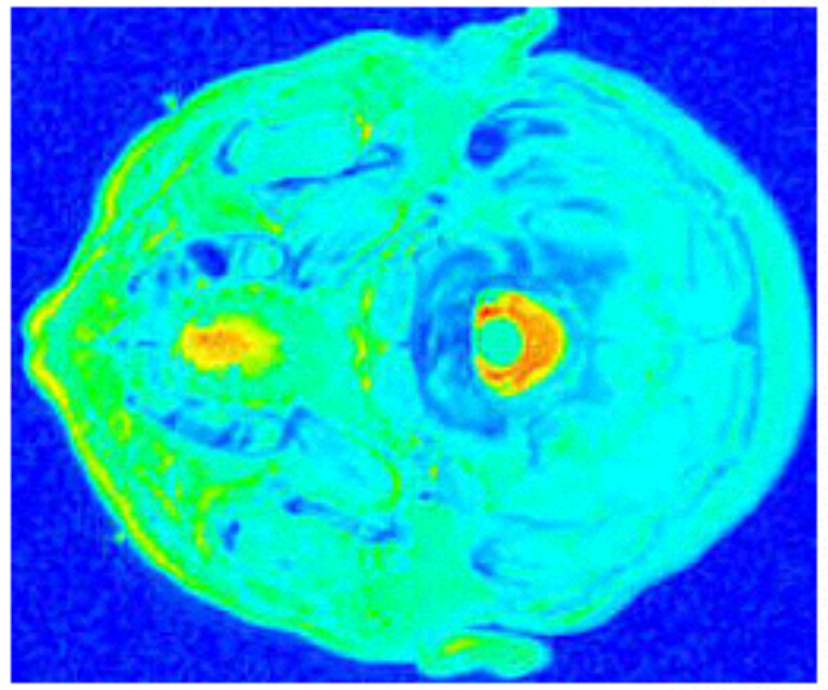
[Rogowitz and Treinish, How NOT to Lie with Visualization, www.research.ibm.com/dx/proceedings/pravda/truevis.htm]

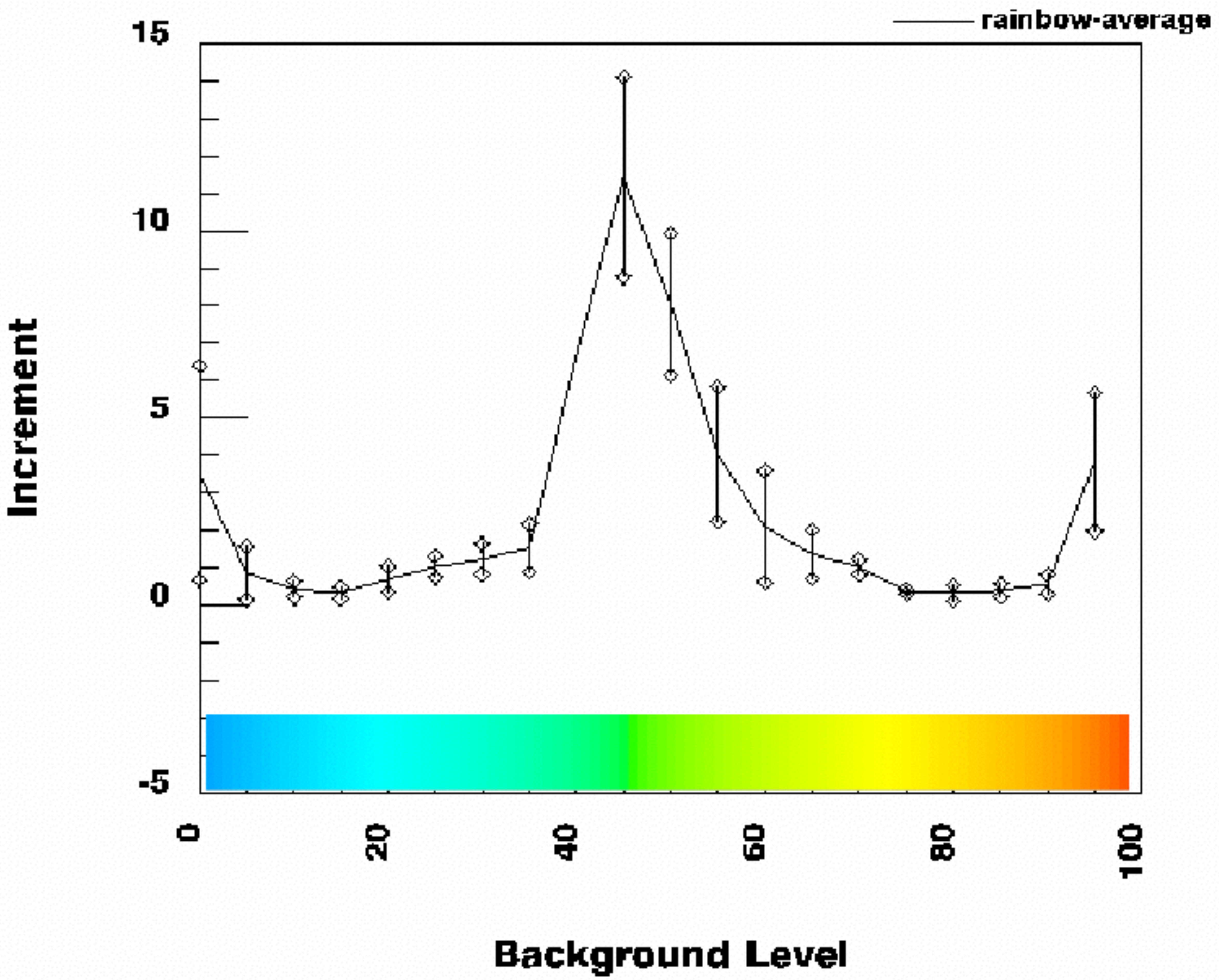
Luminance




gray-average
Increment
Background Level
15
10
5
0
-5
0
20
40
60
80
100

Rainbow Color Scale





Color Blindness

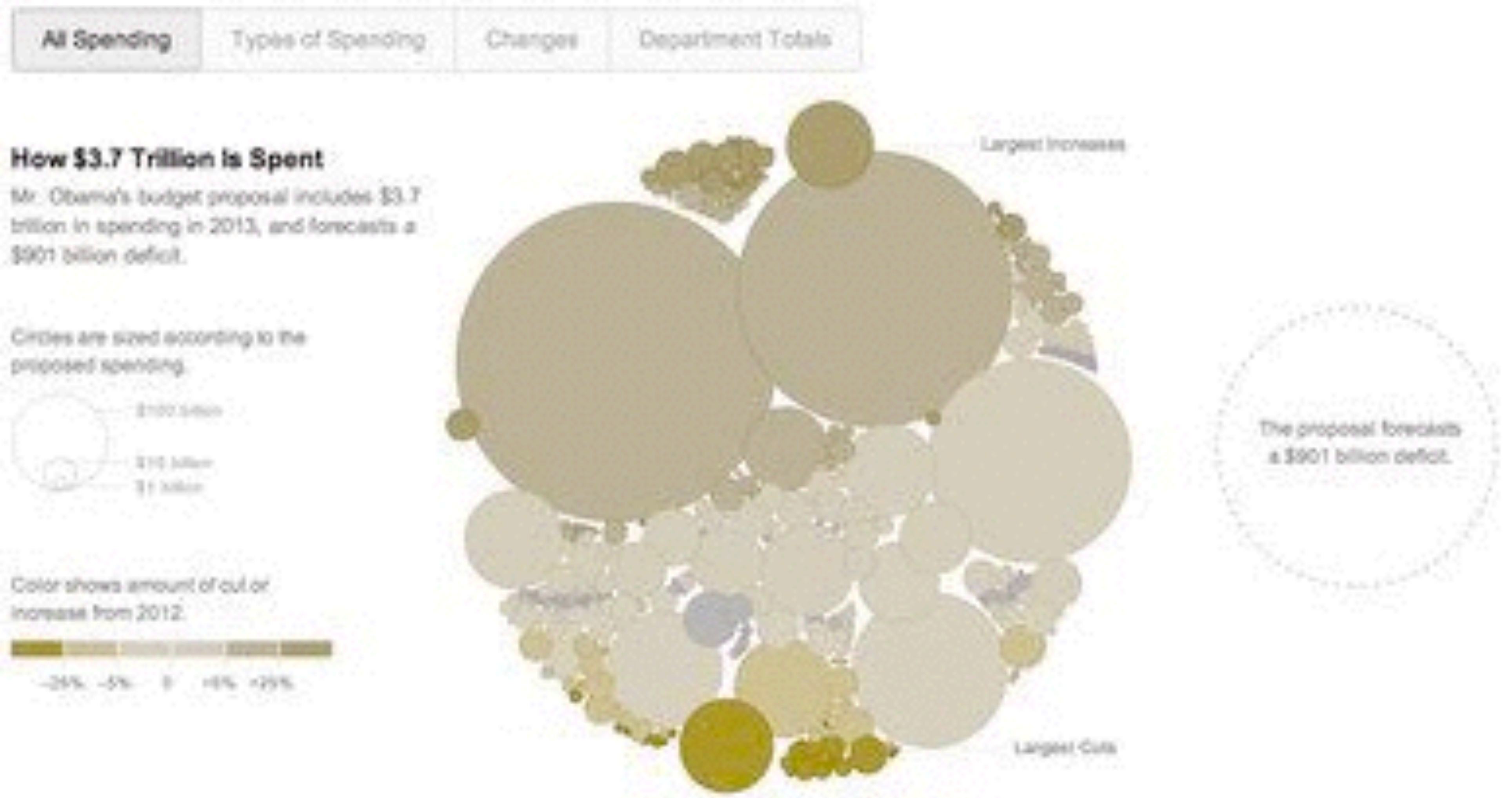
- 10% of males, 1% of females (probably due to x-chromosomal recessive inheritance)
- Most common: red-green weakness / blindness
- Reason: lack of medium or long wavelength receptors, or altered spectral sensitivity (most common: green shift)

Normal Color Perception

Deuteranopia (no green receptors)

Protanopia (no red receptors)

New York Times, Feb 12, 2012

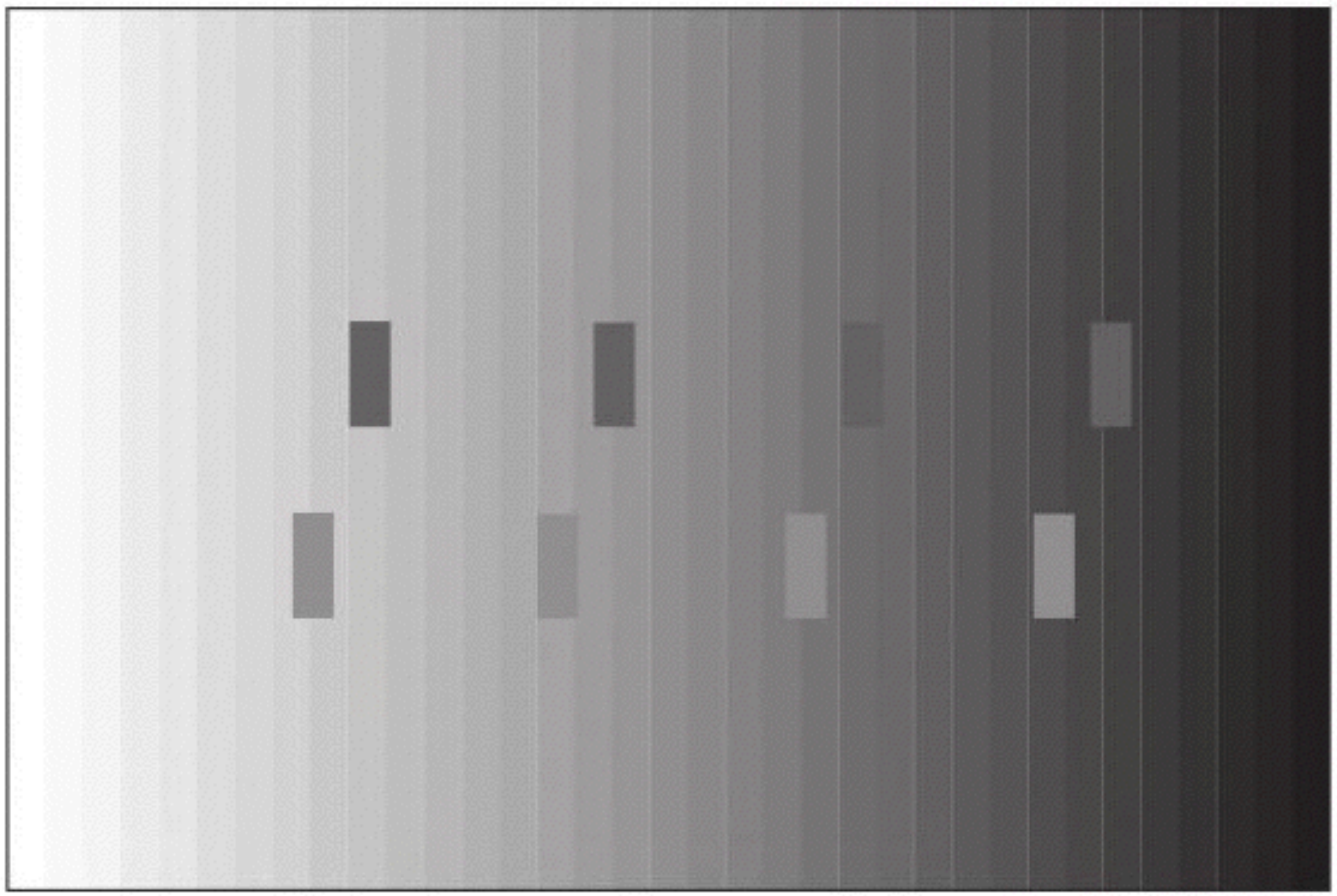


Luminance, Brightness, Lightness

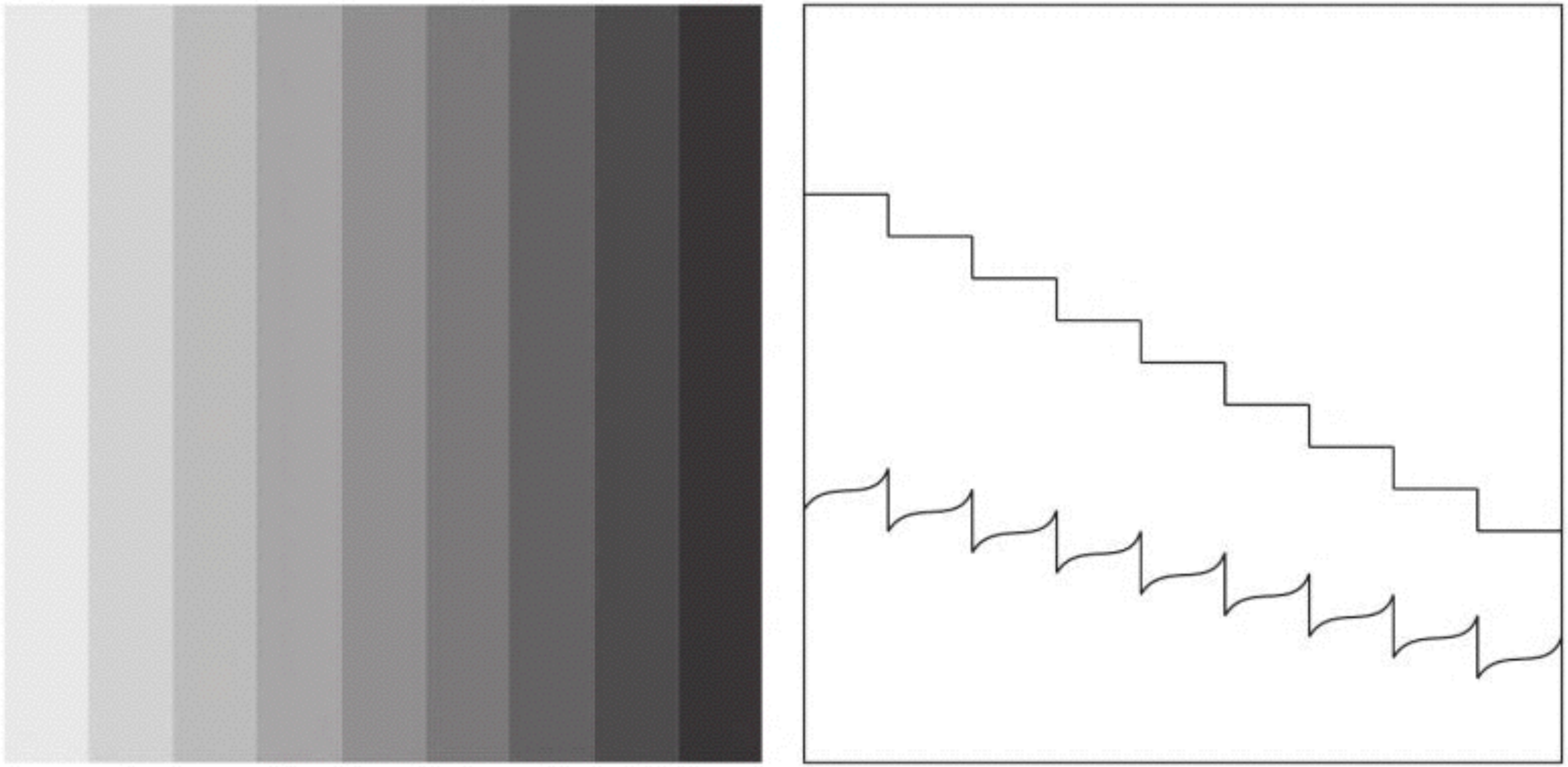
- Luminance
  - measured amount of light (luminous intensity per area)
- Brightness
  - perceived amount of light
- Lightness
  - perceived reflectance of a surface

Simultaneous Brightness Contrast

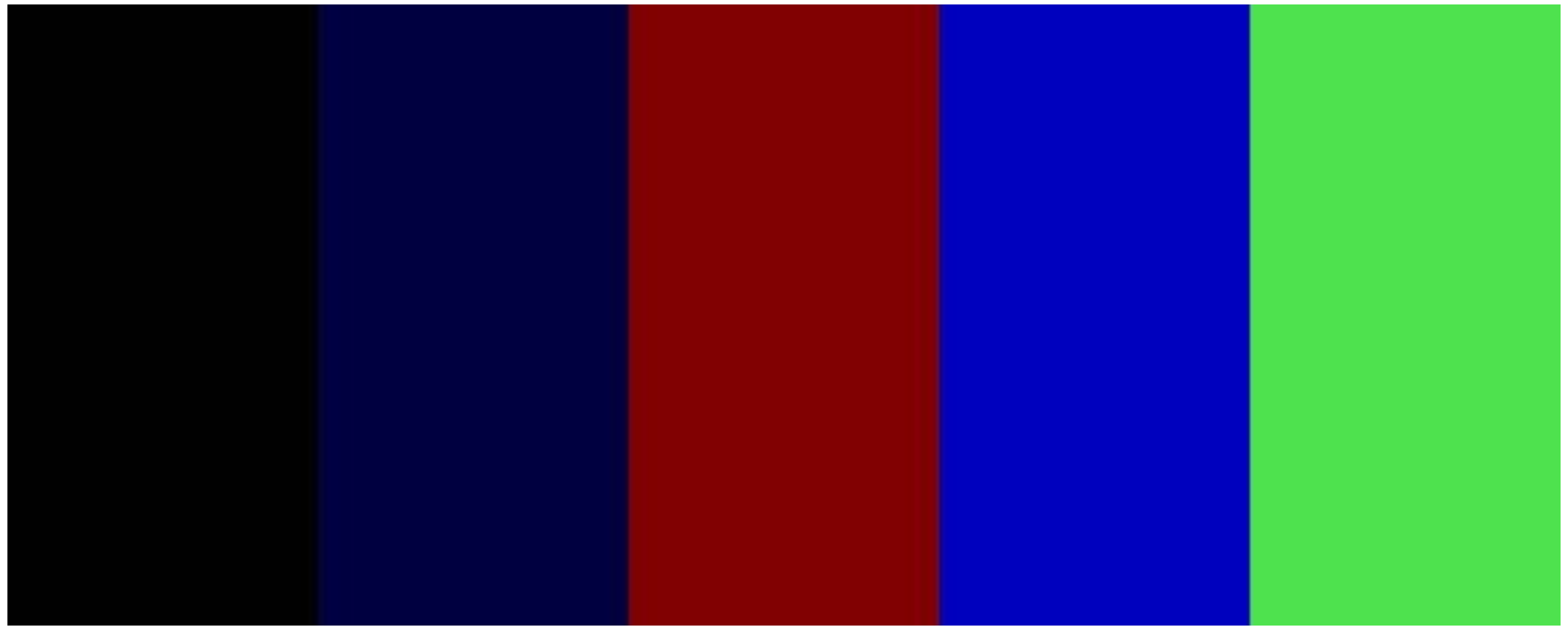
- The perceived brightness of an object is relative to it’s background



Chevreul Illusion - Same color different intensities

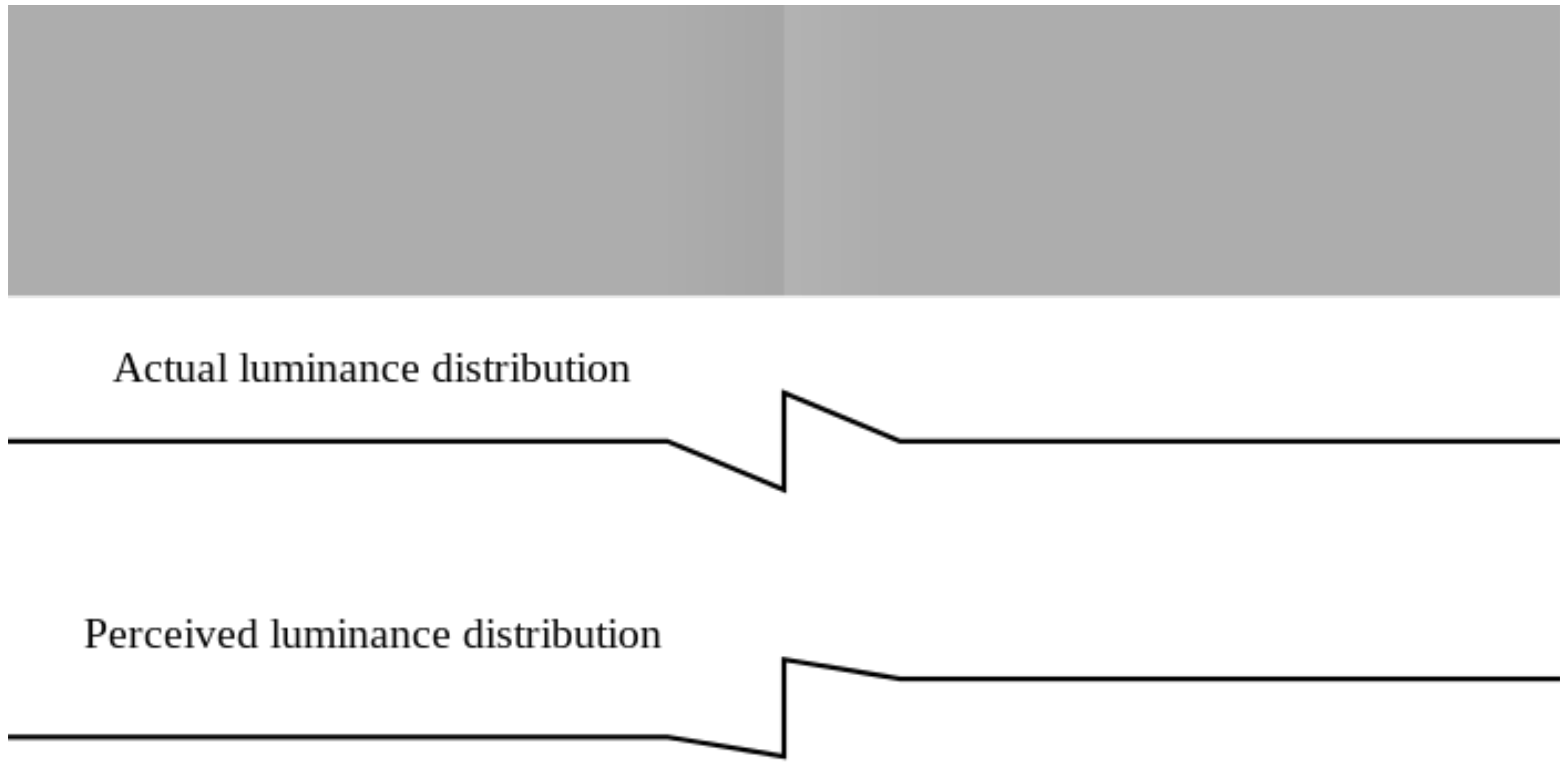


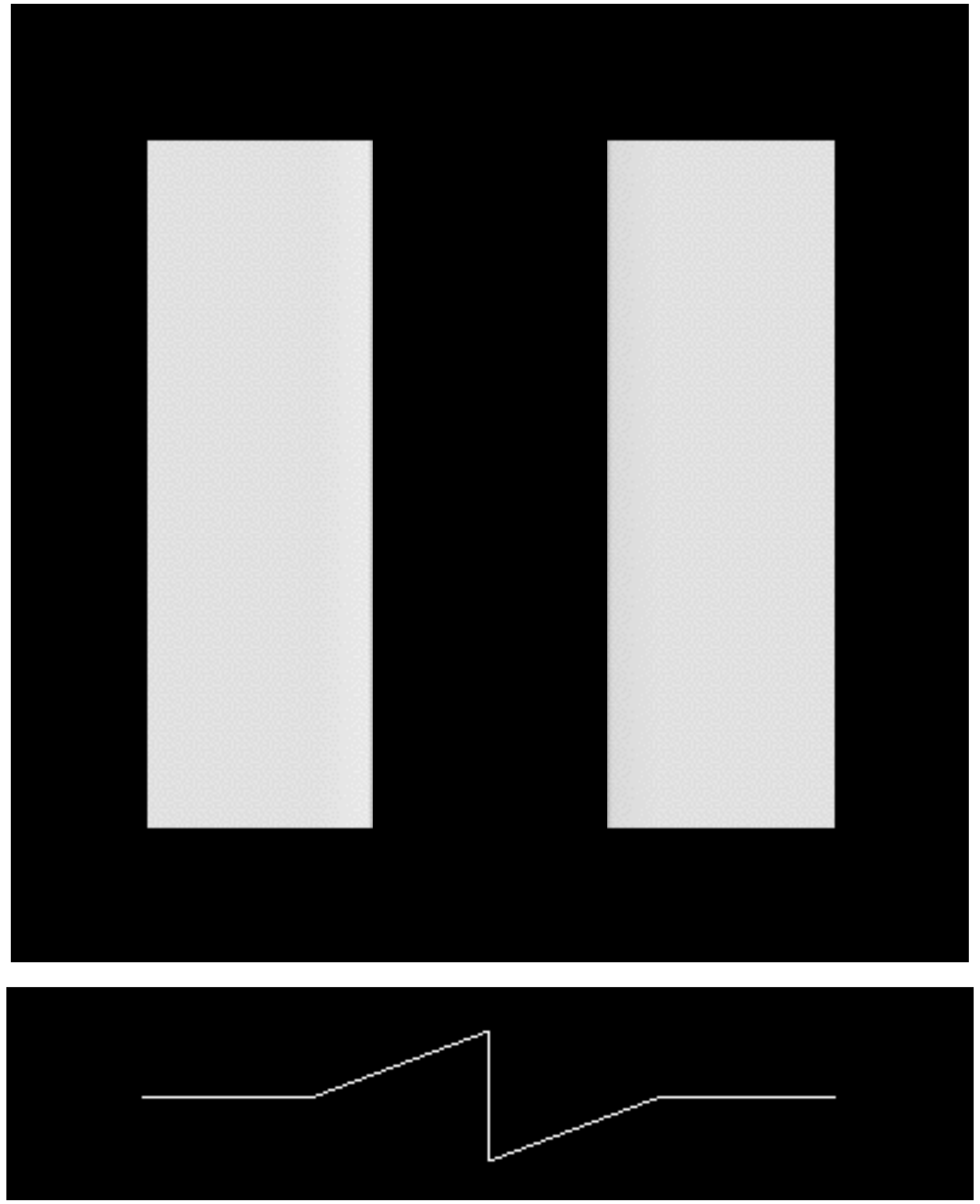
Chevreul Illusion



Edge Enhancement

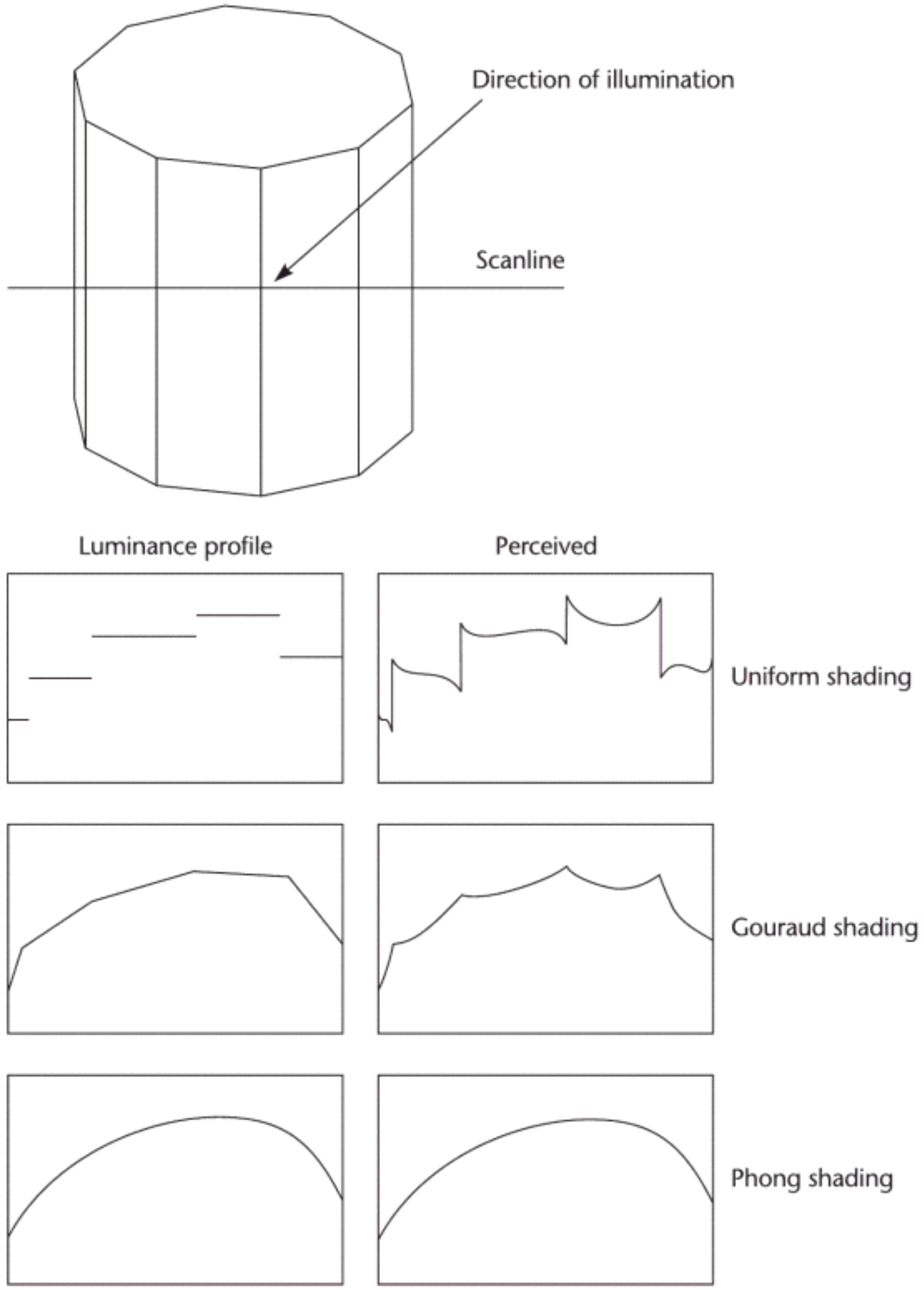
- Cornsweet effect

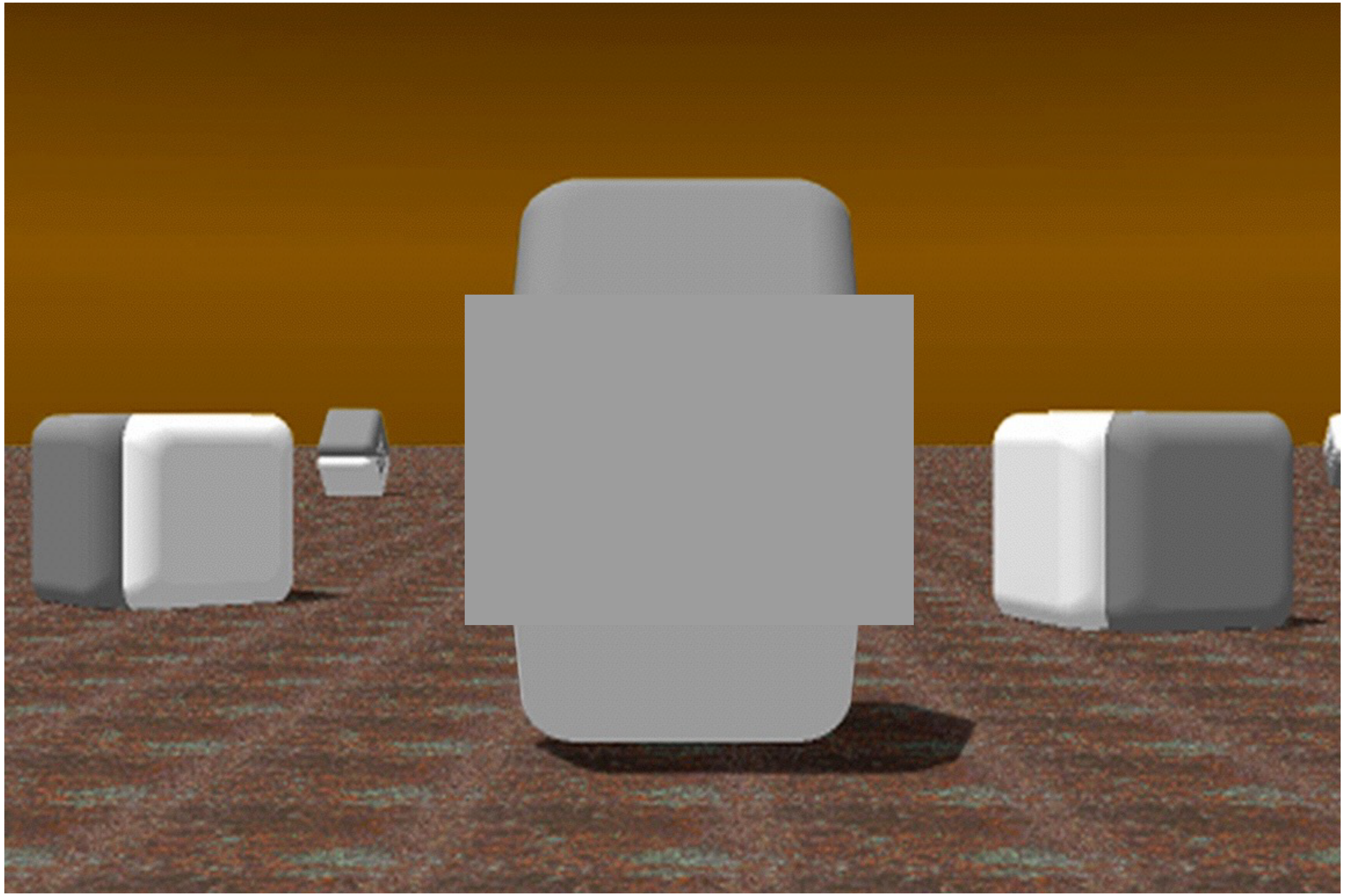




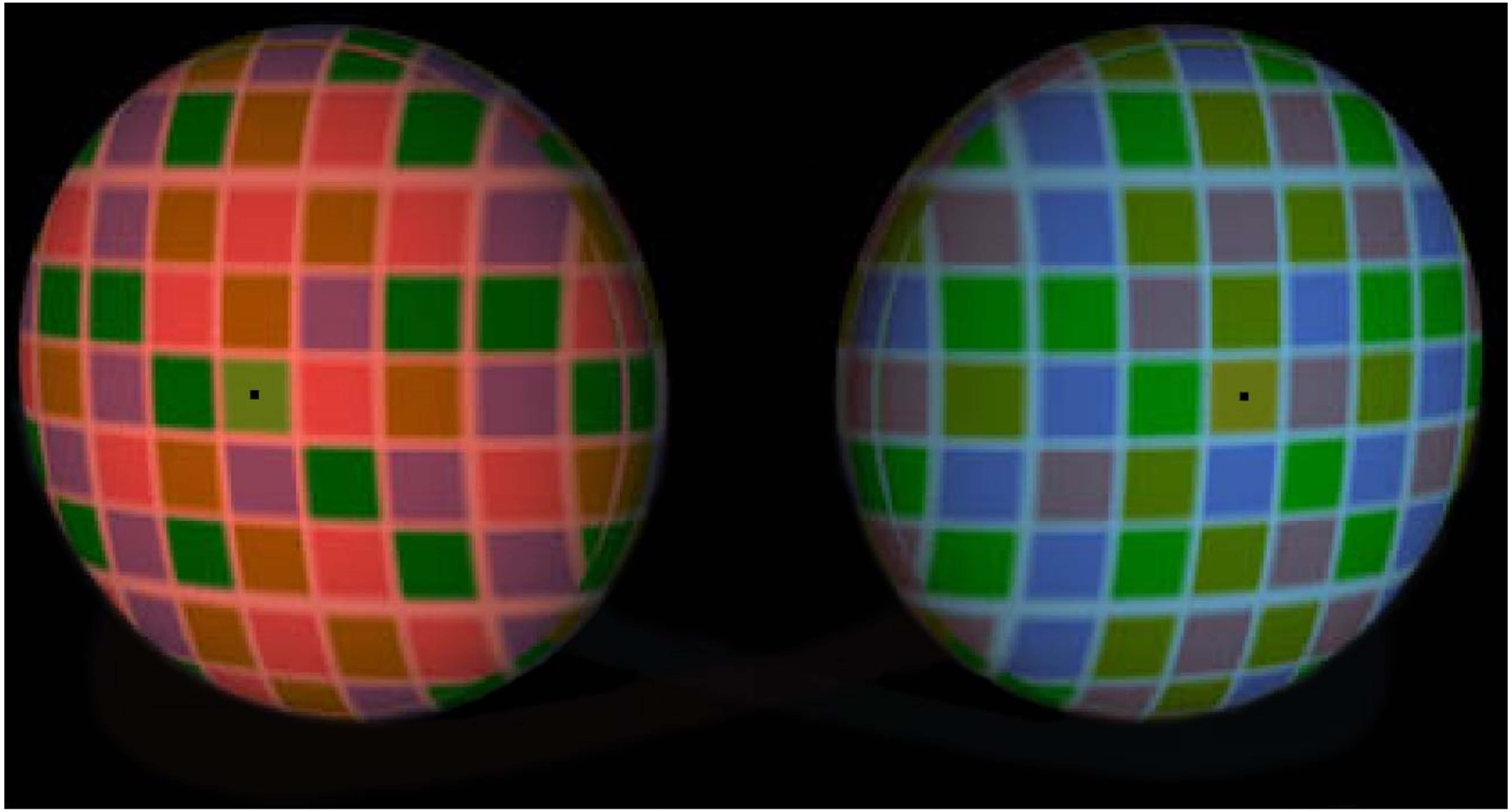
Why is this an issue?

- Can result in large errors of judgment
- Amplifies artifacts in computer graphics shading





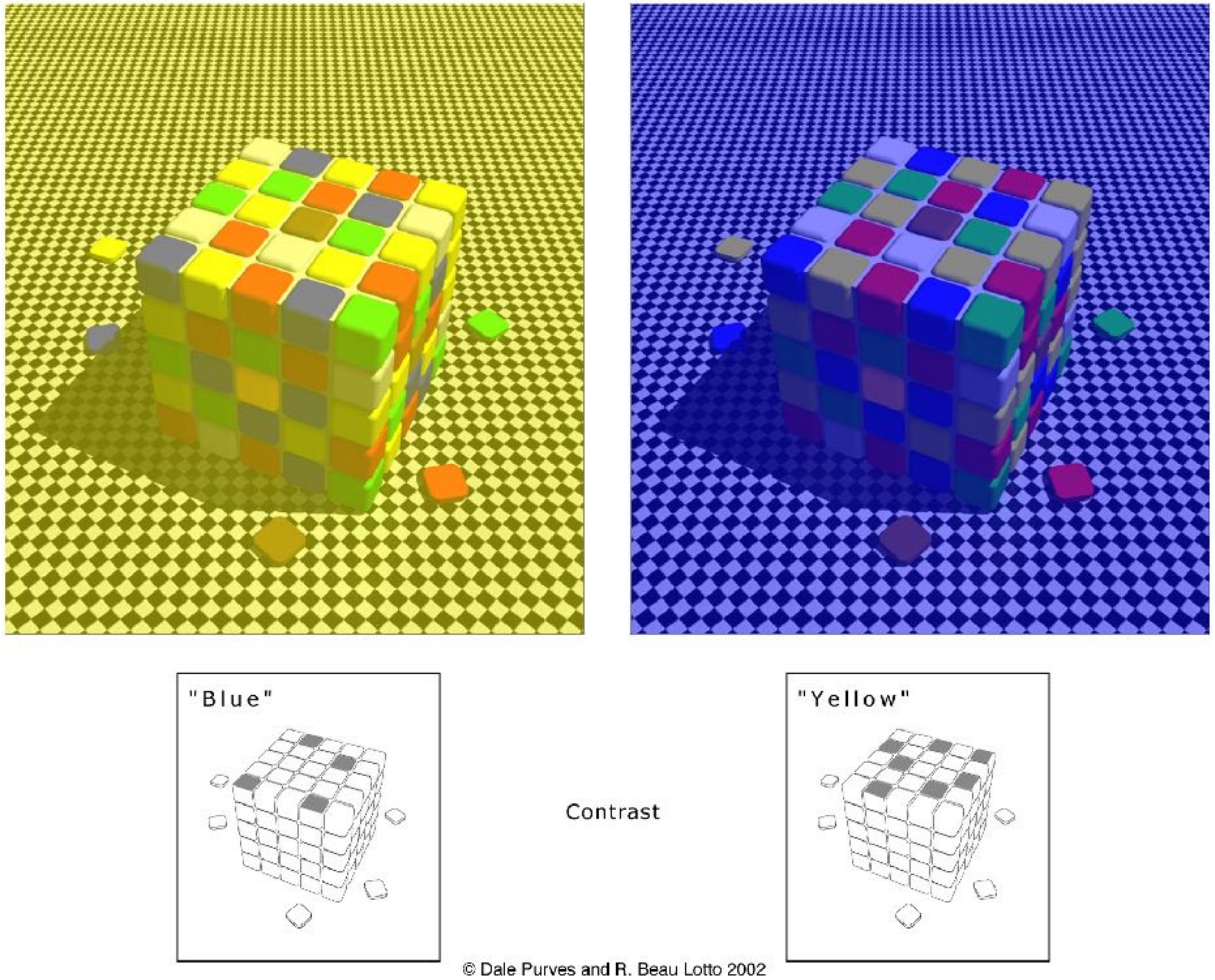
Color contrast: multiple cues



© Dale Purves and R. Beau Lotto 2002

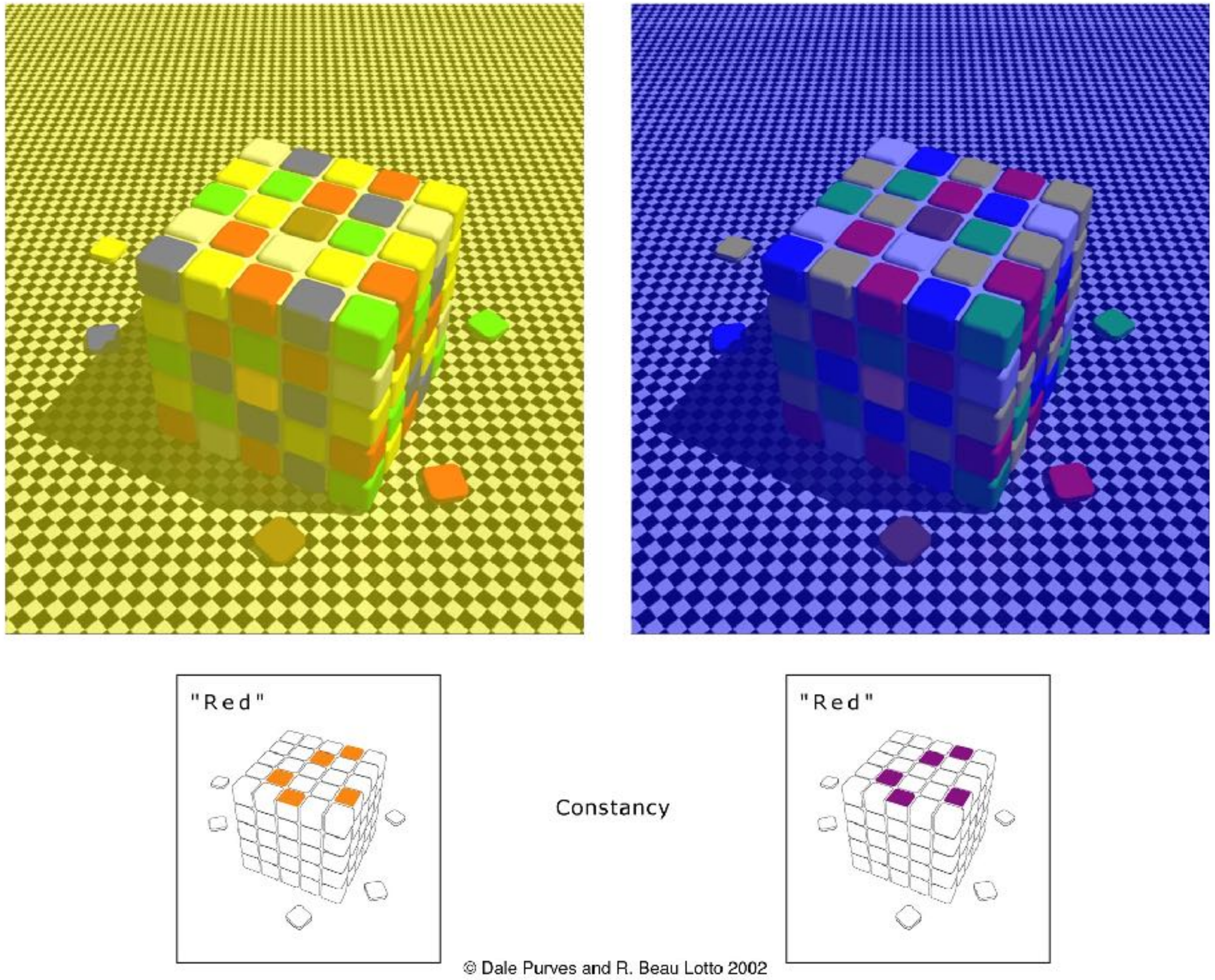
- Color noted is the same

Color Contrast



- Same colors can look different (gray)

Color Consistency



- Colors can made to look the same (orange and purple to red)

Stroop Effect - Task Interference

Change Blindness

- Details of an image cannot be remembered across separate scenes – except in areas with focused attention
- Interruption (e.g. a blink, eye saccade or blank screen) amplifies this effect
- No failure of vision system, failure based on inappropriate attentional guidance

Change Blindness

- Various theories about causes
  - Overwriting: Information that was not abstracted is lost
  - First Impression: Only initial view is abstracted
  - Nothing is Stored: Only abstract concepts are committed to memory
  - Everything is Stored, Nothing is Compared: We compare only when we are forced to
  - Feature Combination: scenes are combined as long as they make sense
- Influencing factors
  - attention
  - expectation (knowing something will change)
  - semantic importance of changed object
  - low level object properties overlooked more easily

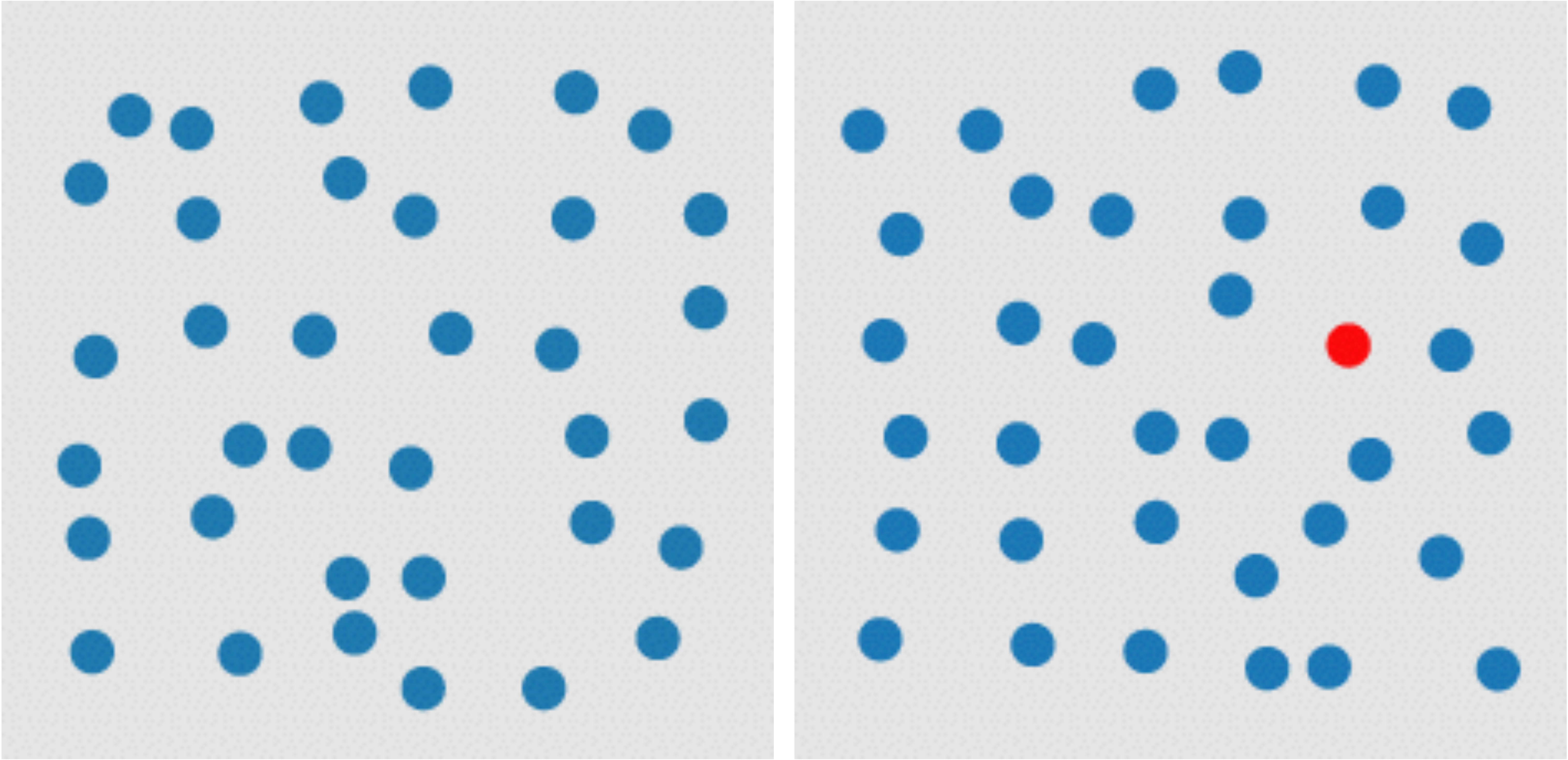
Instructions

Count how many times the players wearing white pass the basketball.

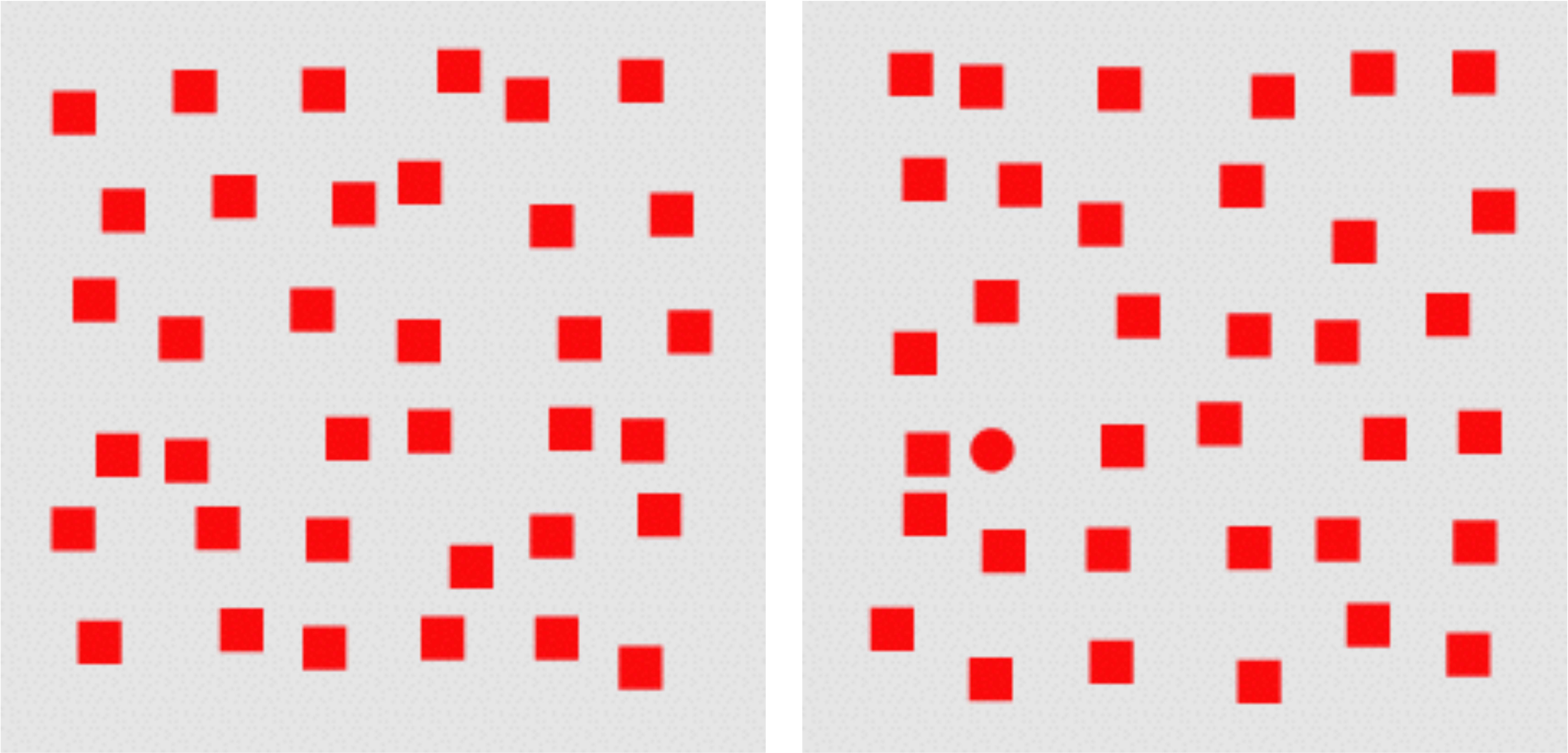
OK, what are we really good at? - Preattentive Processing

- Properties detected by the low-level visual system
    - very rapid - 200-250 milliseconds
    - very accurate
    - processed in parallel
- happens before focused attention -> “pre”attentive
- attention is very important for cognition
- Independent of the number of distractors!
- Opposite: sequential search (processed serially)

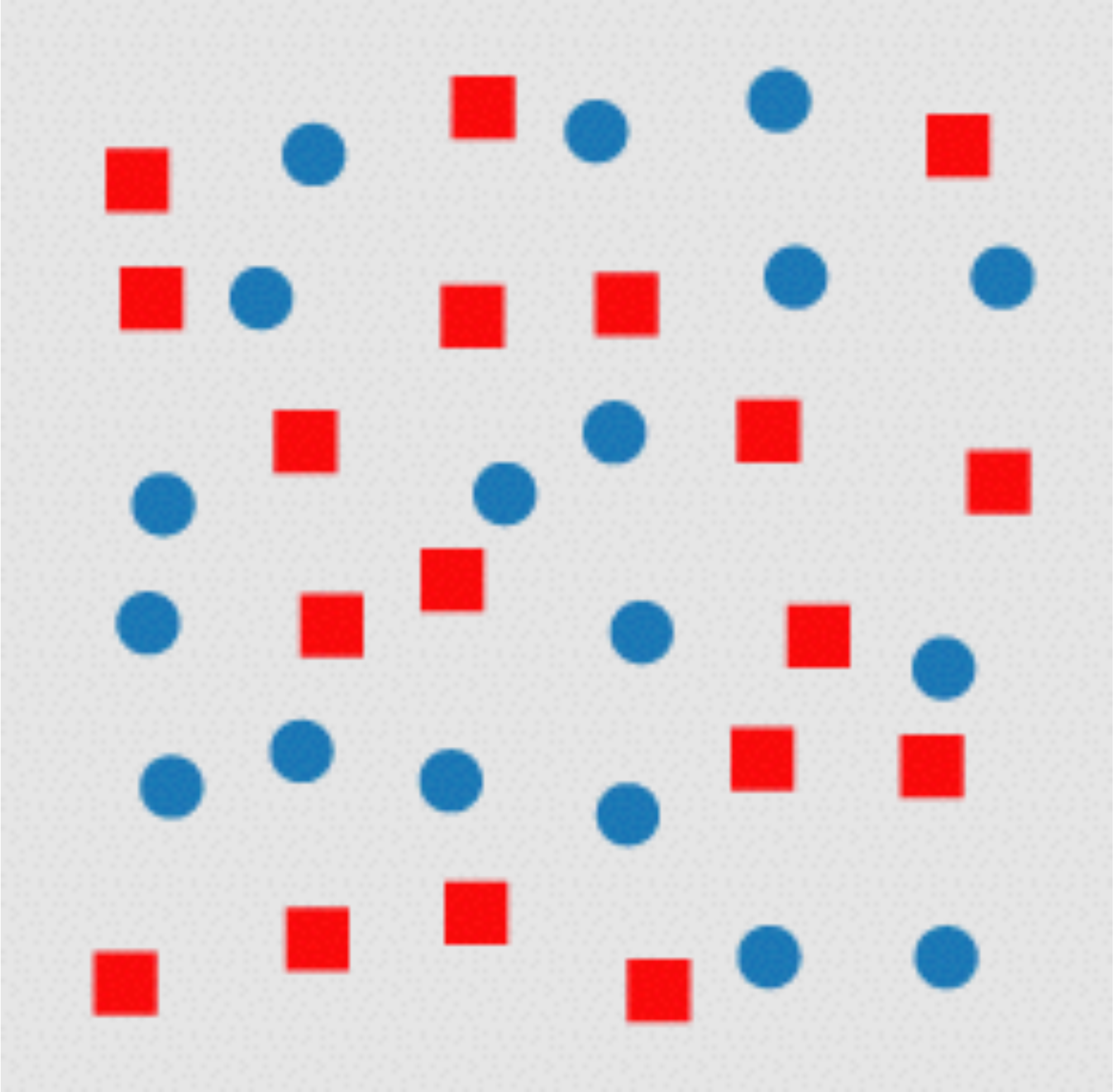
Difference in Hue

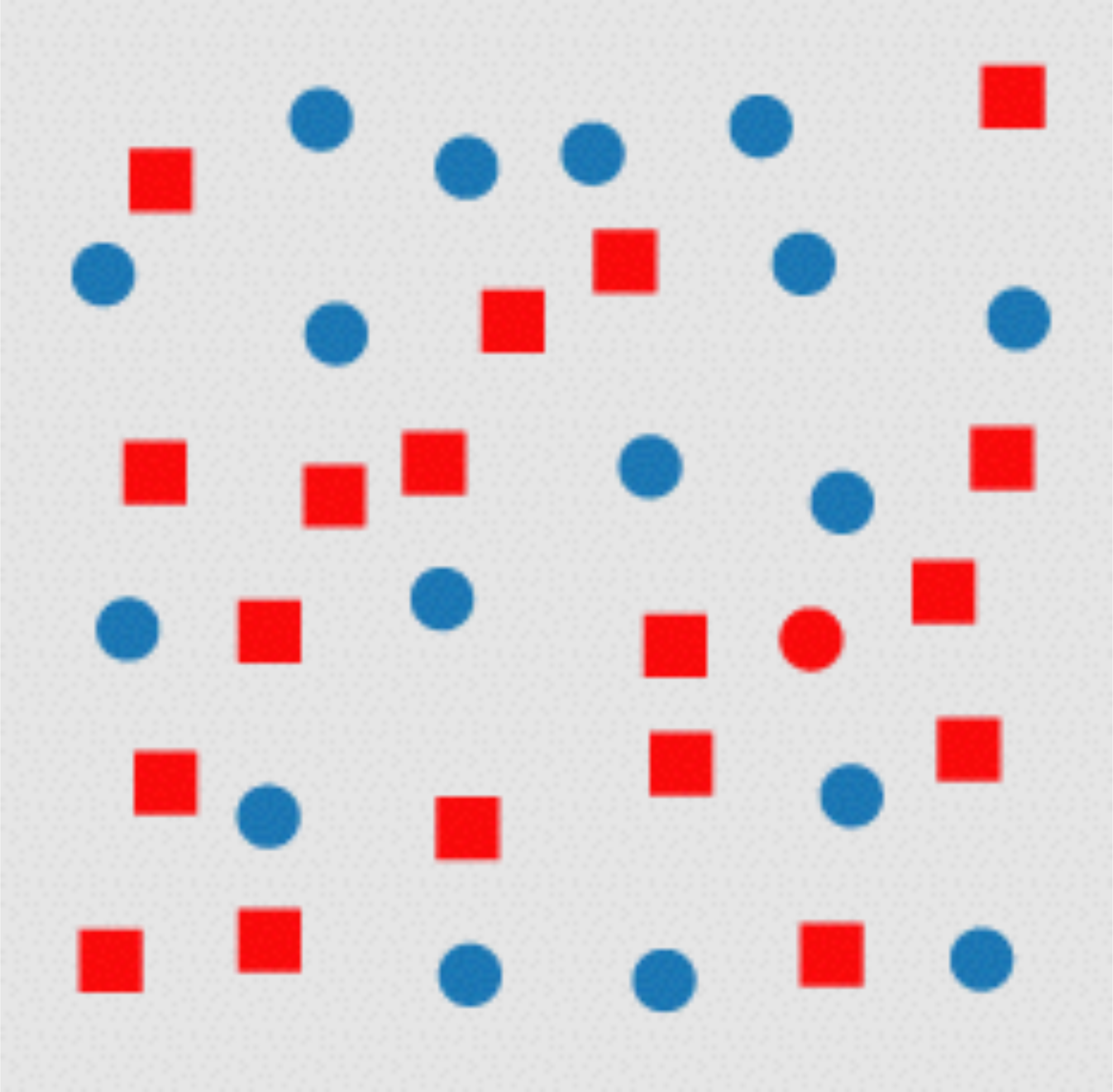


Difference in Curvature / Form



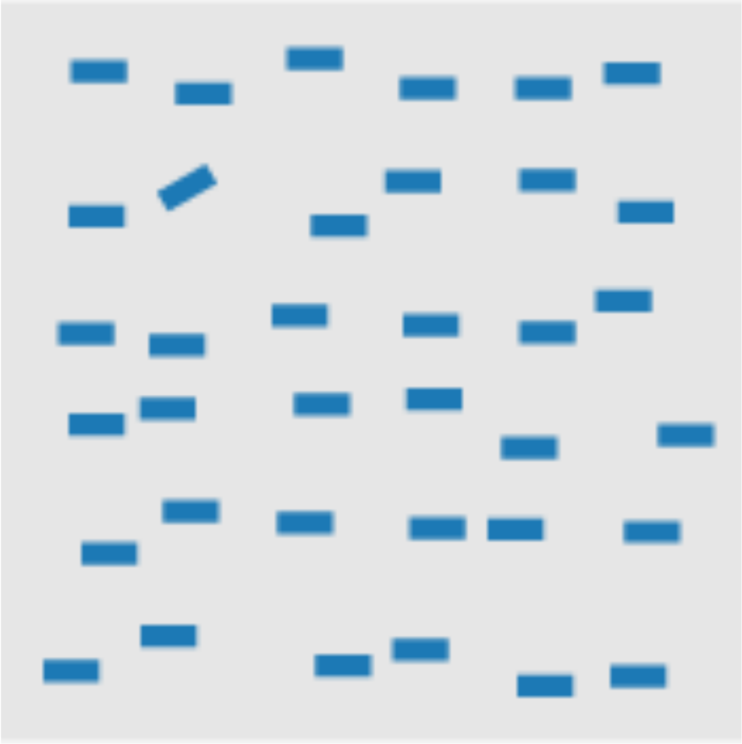
Not Valid for Combinations

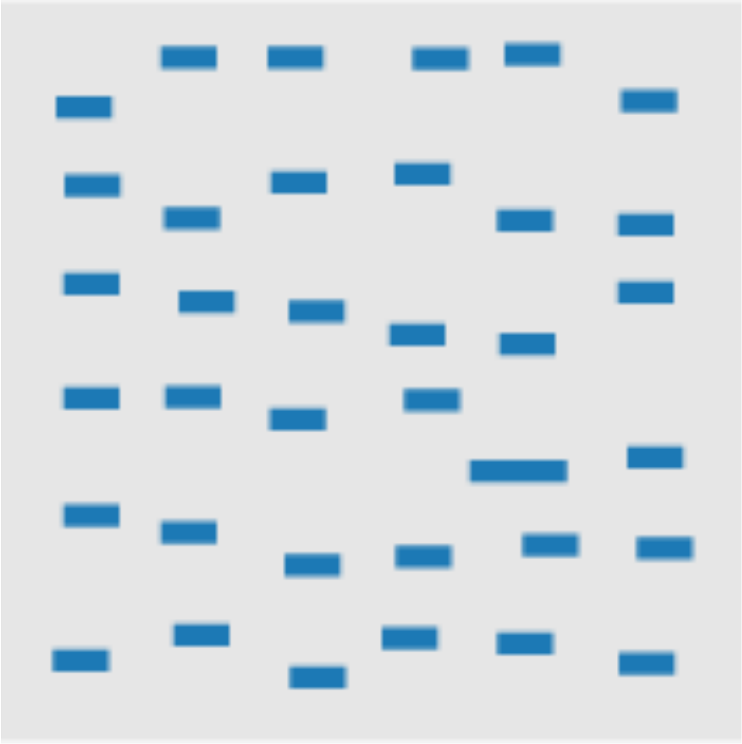




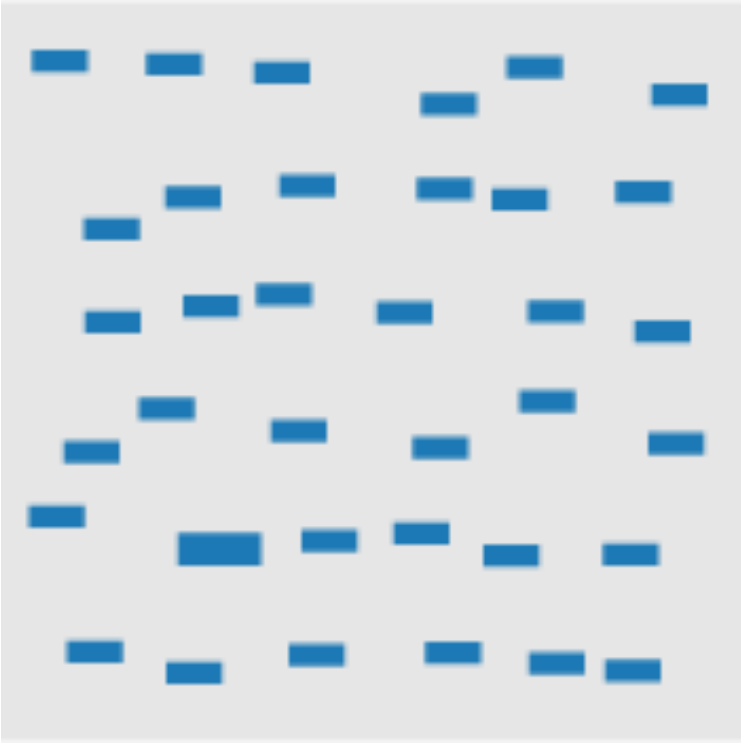
- Conjunction Targets – no unique visual property
- target: red, circle
- distractor objects have both properties

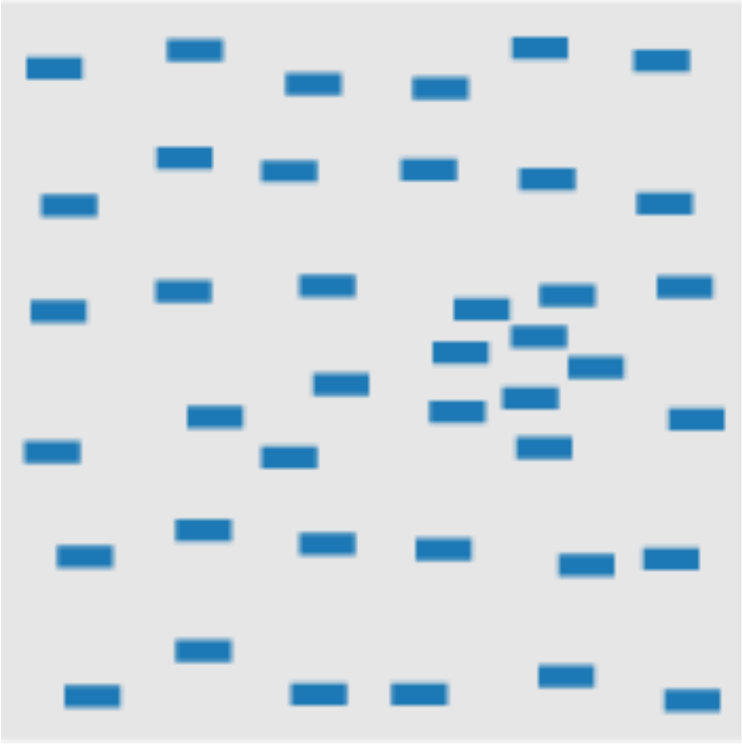
Some Preattentive Properties

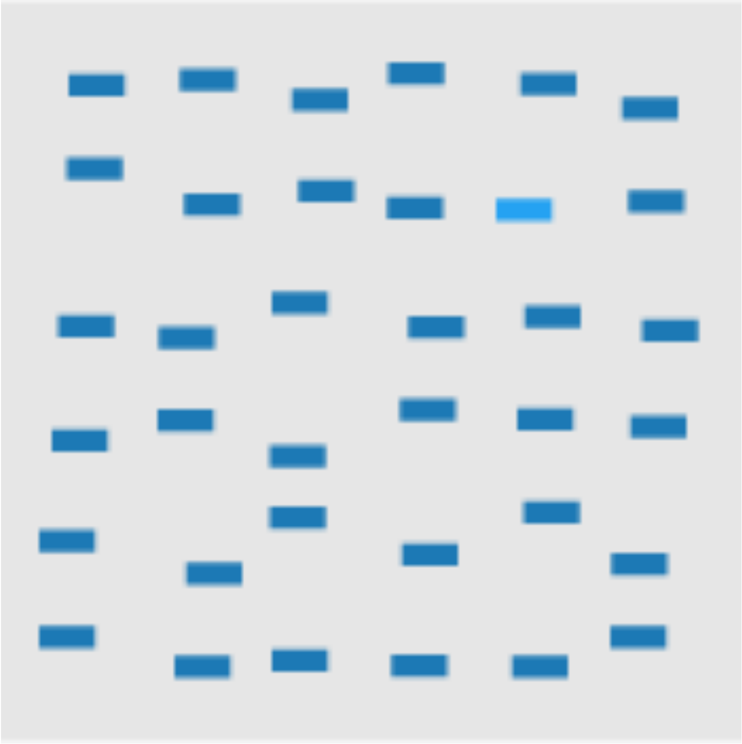

line (blob) orientation

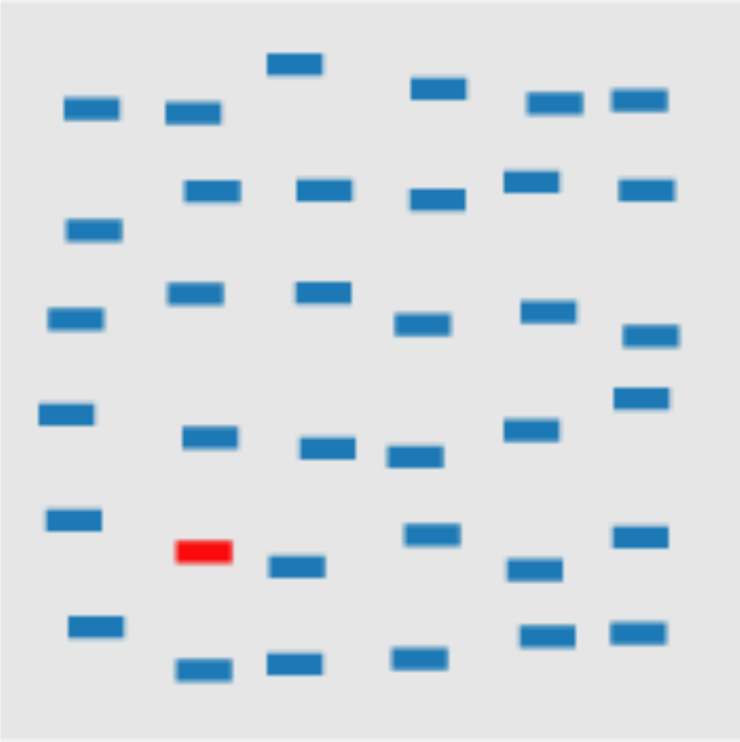

length/width

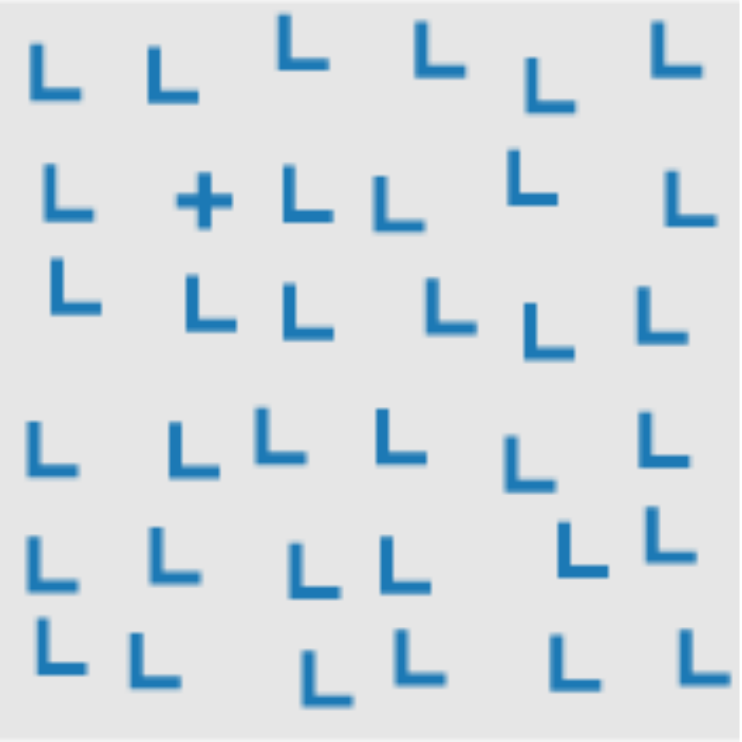

closure

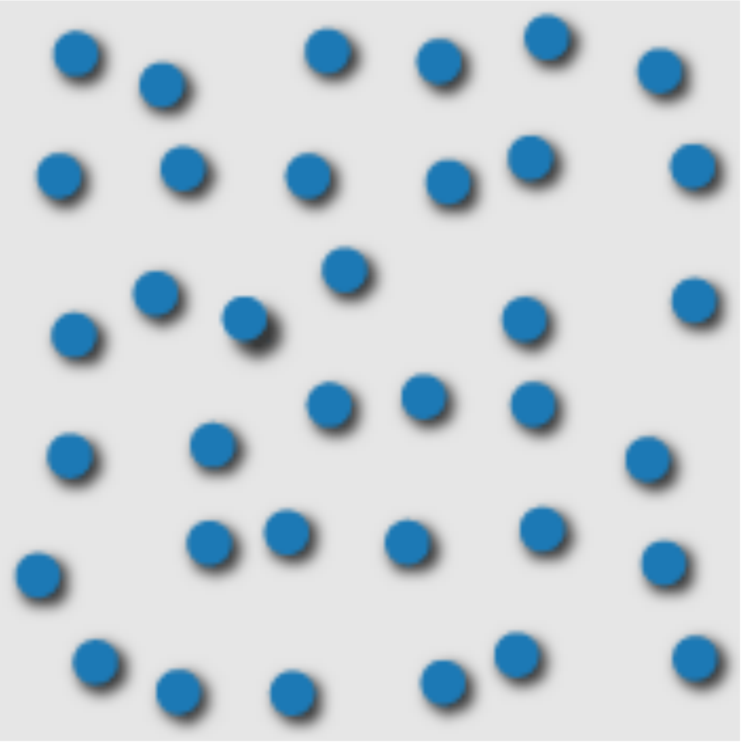

size

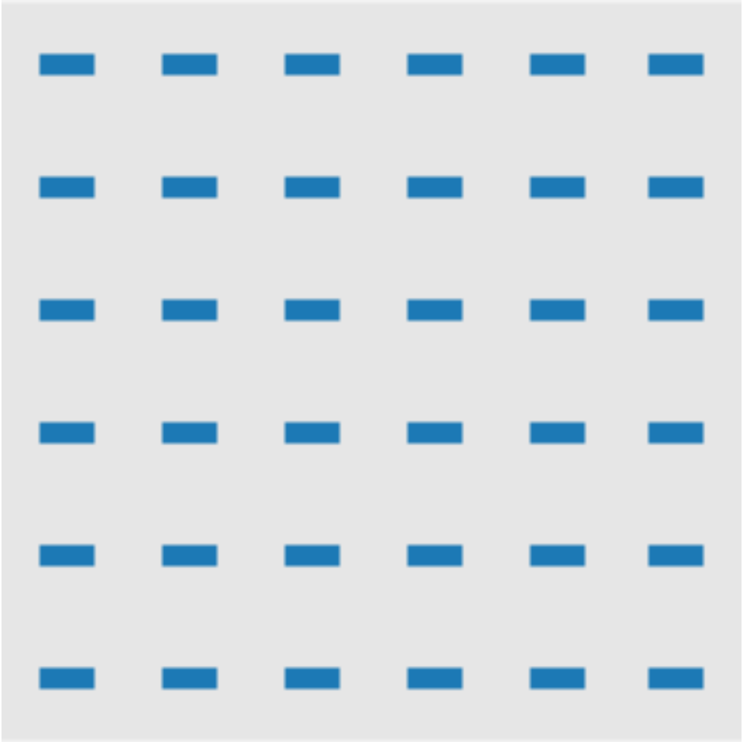

clustering

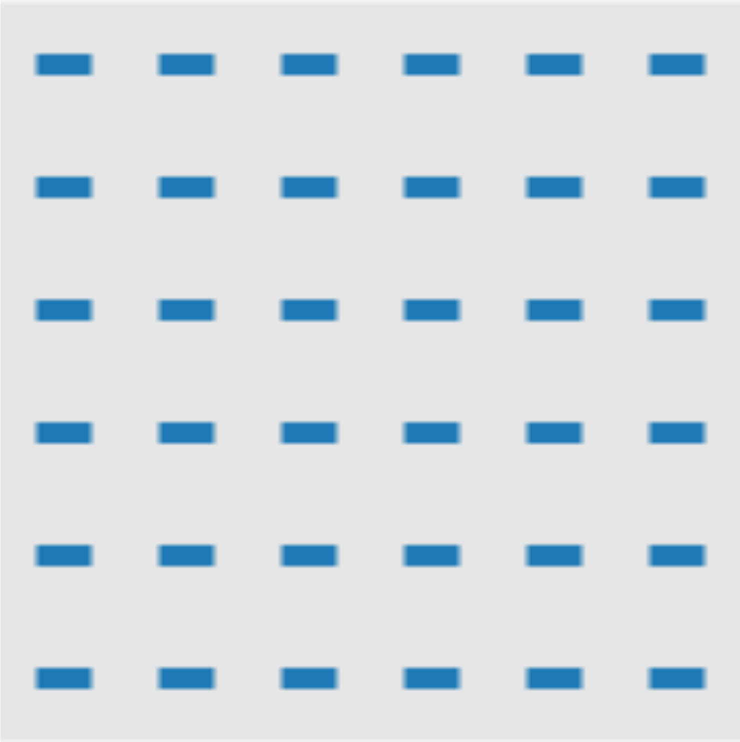

intensity

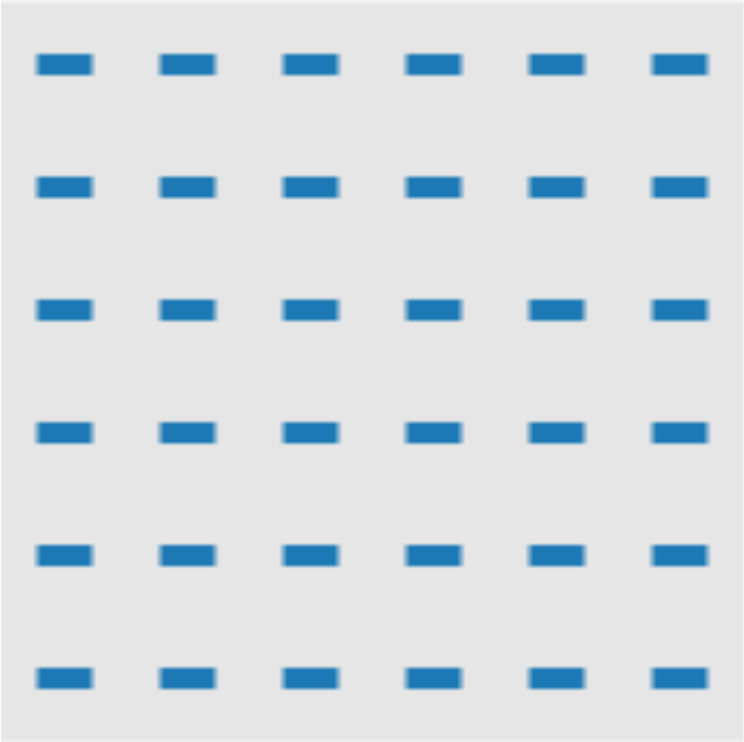

hue


intersection


3D depth cues


flicker


direction of motion


velocity of motion

Best Practices for Encoding Info

Value/Luminance/Saturation

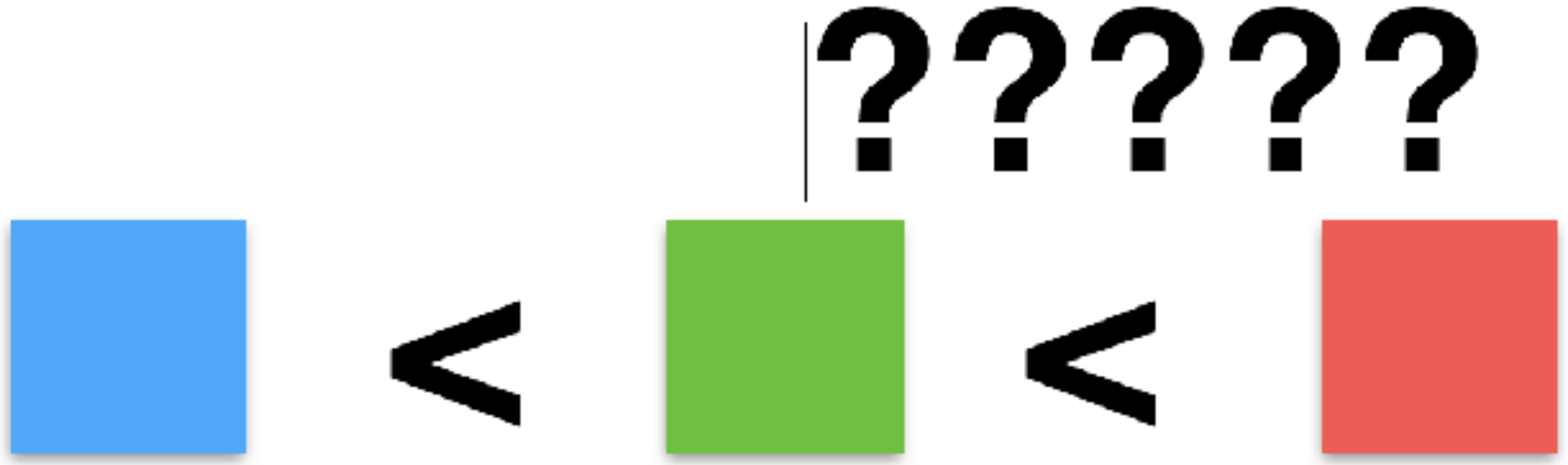
- OK for quantitative data.
- Not very many shades recognizable

Selective: yes
Grouping: yes
Quantitative: somewhat (with problems)
Order: yes
Scales: limited

http://colorbrewer2.org

Color

- Good for qualitative data (identity channel)
- Limited number of classes/length (~7-10!)
- Can be misleading for quantitative data

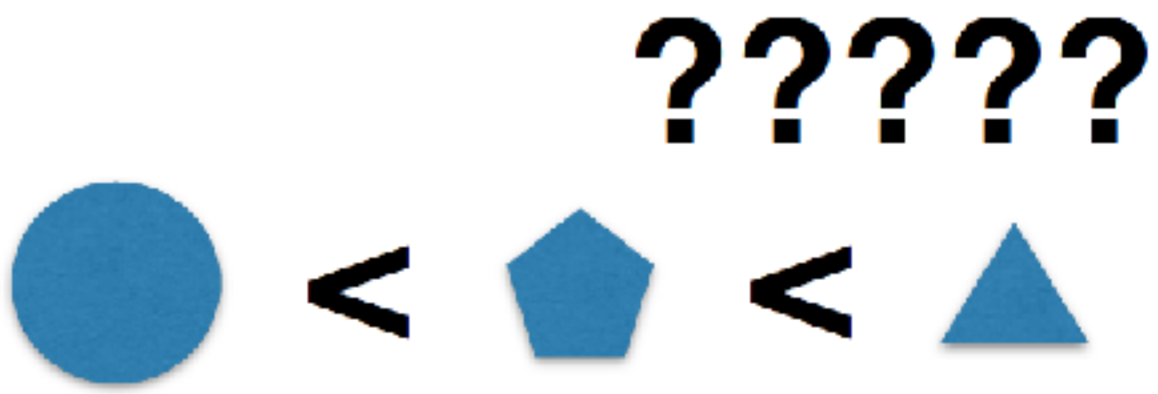


Selective: yes
Grouping: yes
Quantitative: no
Order: no
Scales: limited

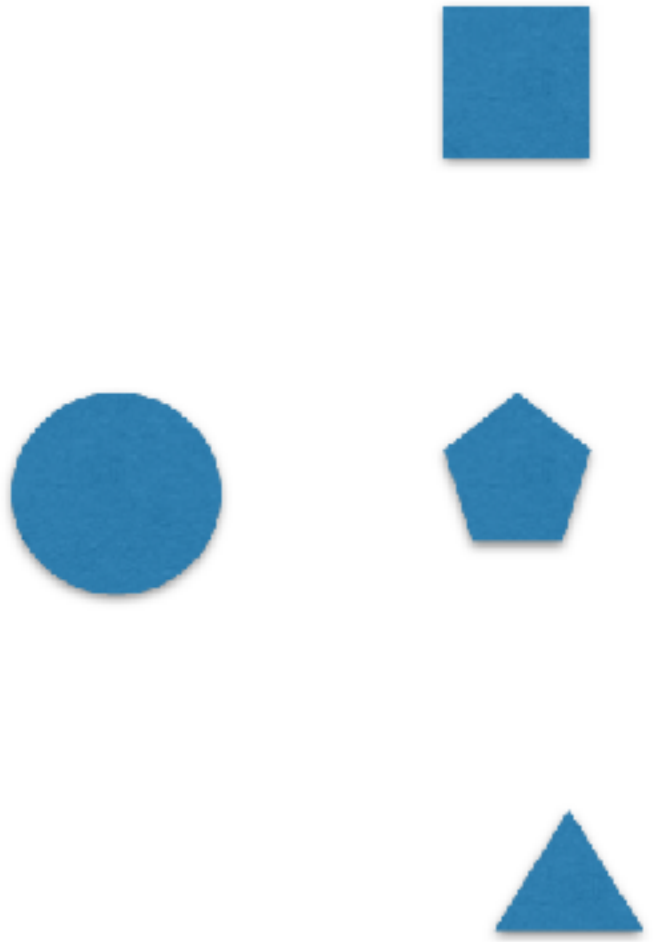
http://colorbrewer2.org

Shape

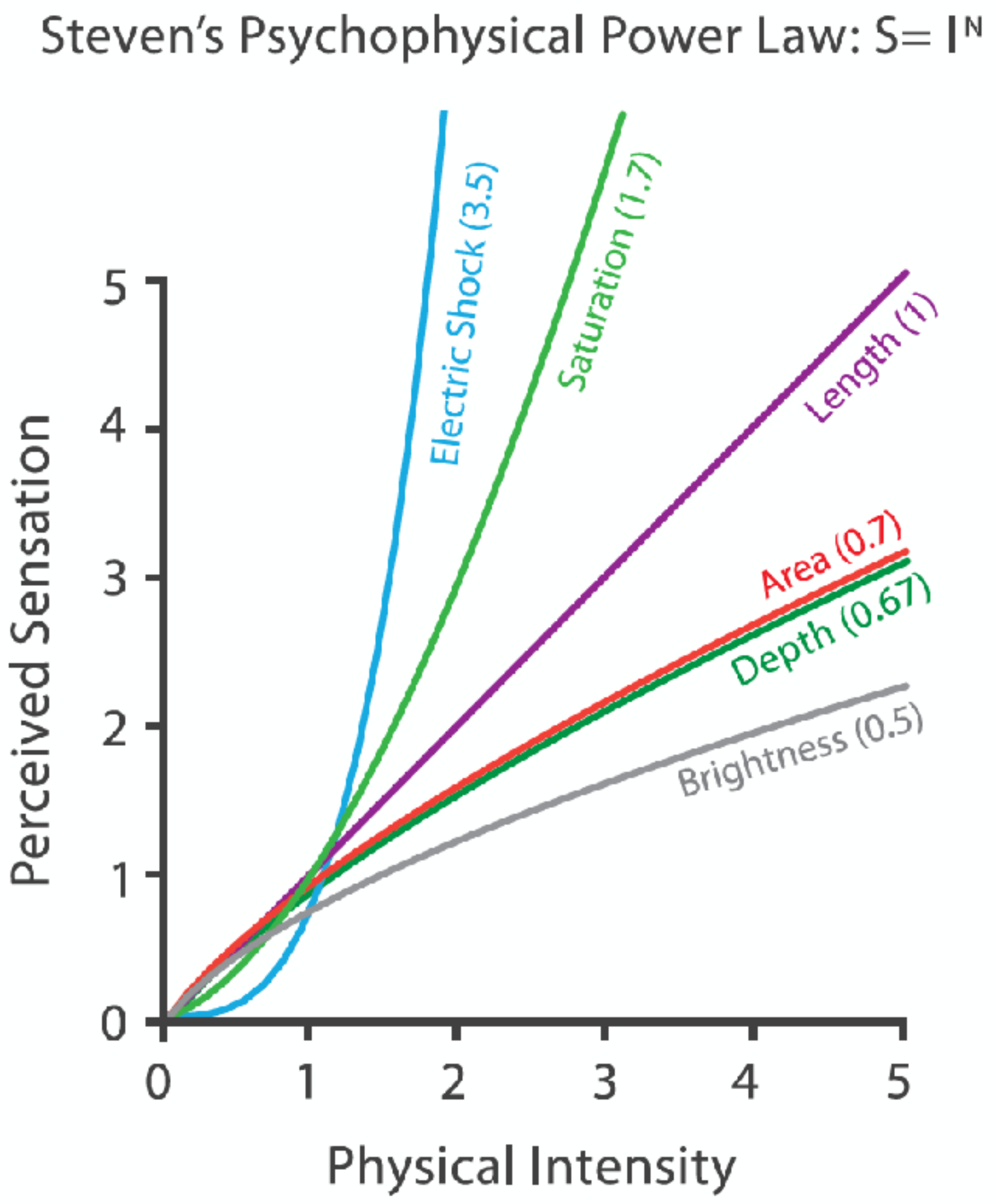
- Great to recognize many classes.
- No grouping, ordering.



Selective: yes
Grouping: limited
Quantitative: no
Order: no
Scales: big



Why are quantitative channels different?



S = sensation

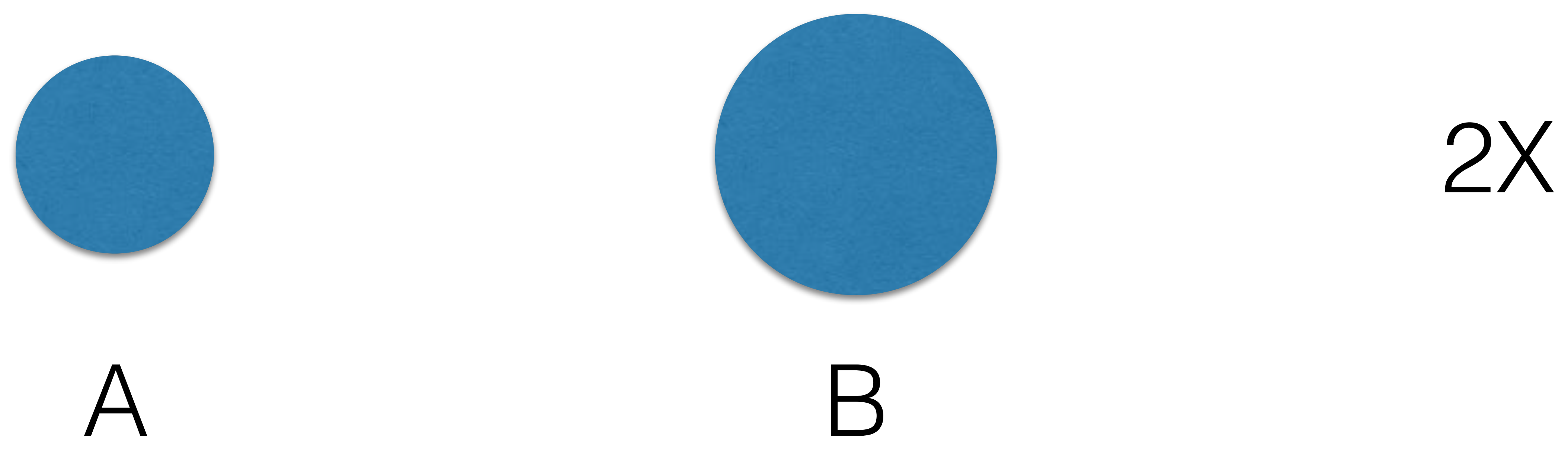
I = intensity

How much longer?

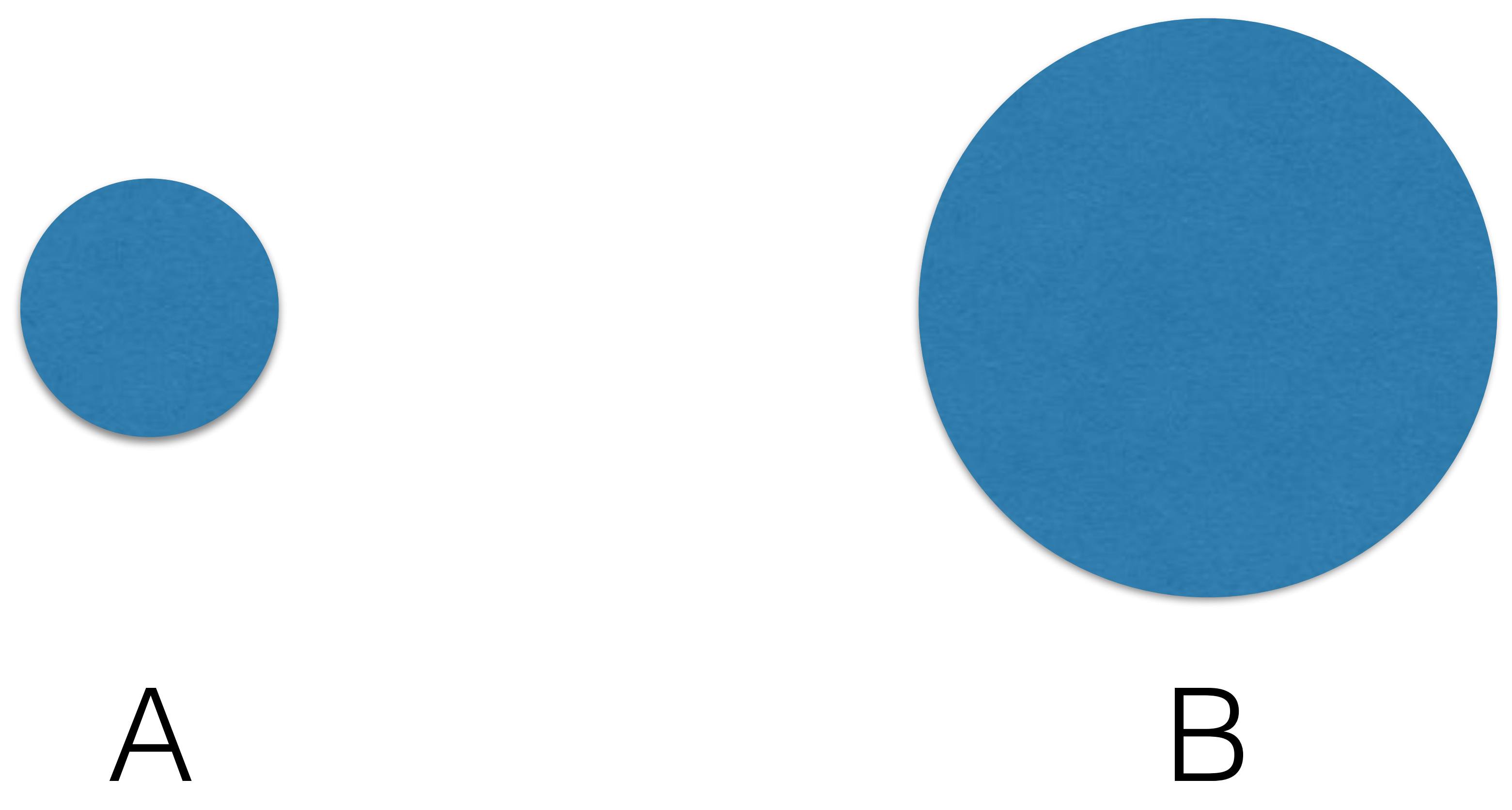
2X

How much longer?

How much larger (area)?

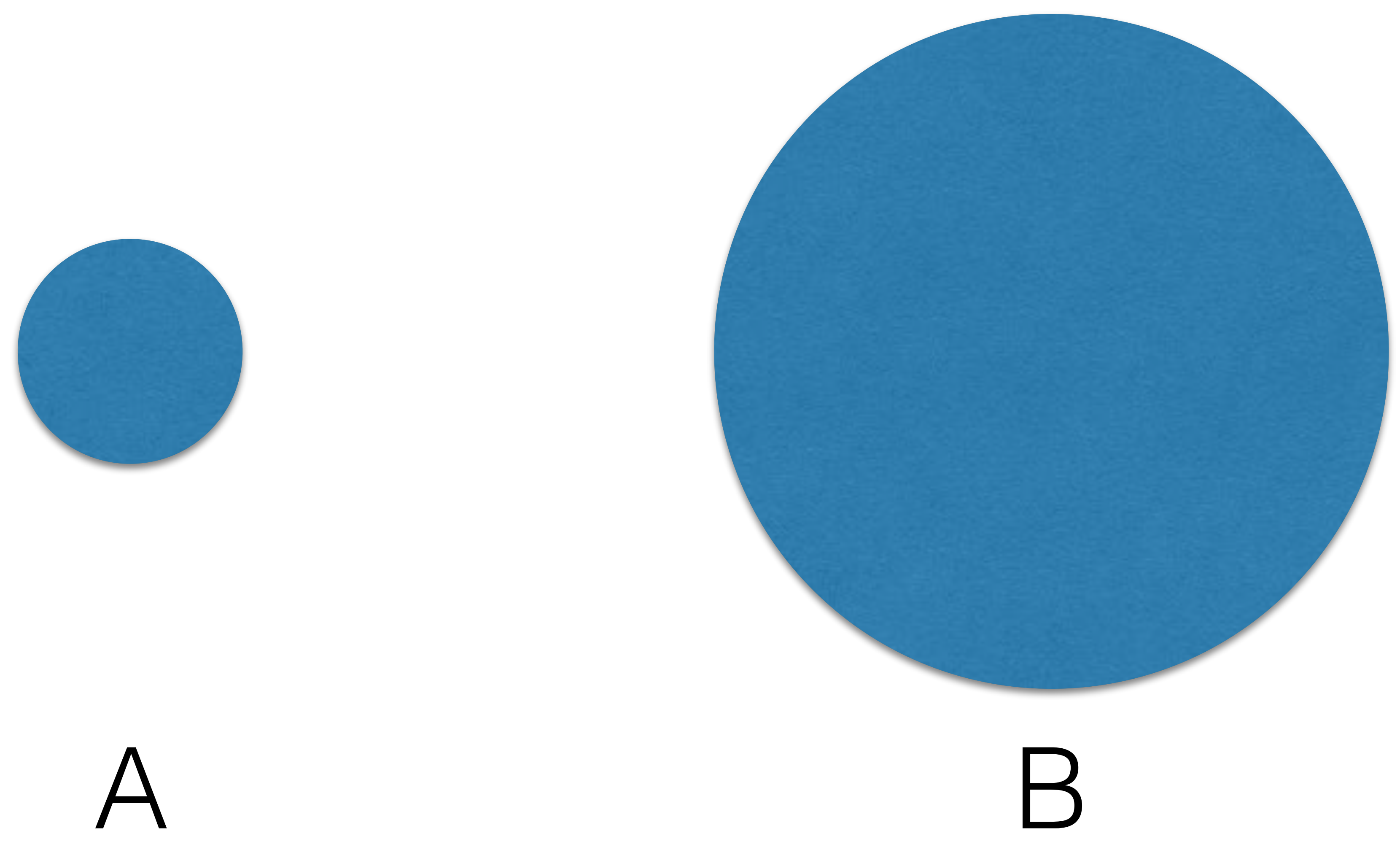


How much larger (area)?



5X

How much larger (diameter)?



3X

How much darker?

2X

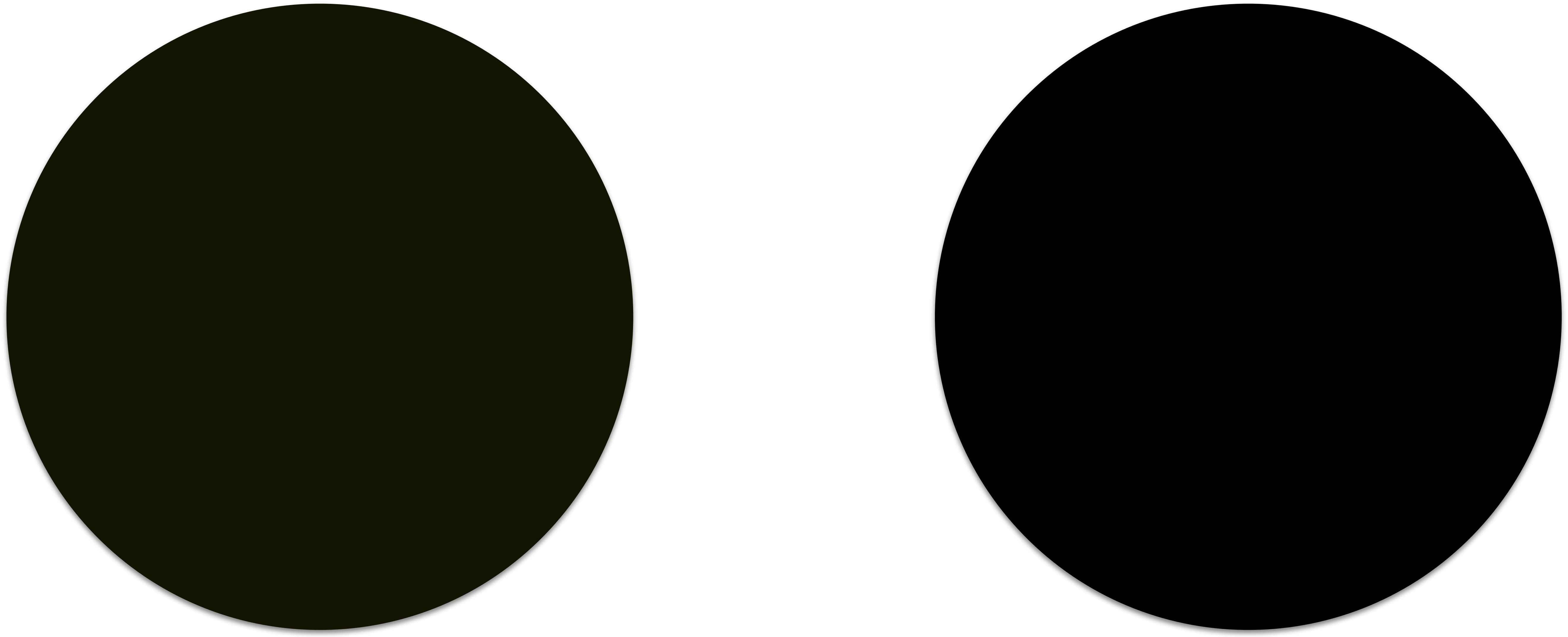
- 200 vs 100

How much darker?

3X

- 150 vs 50

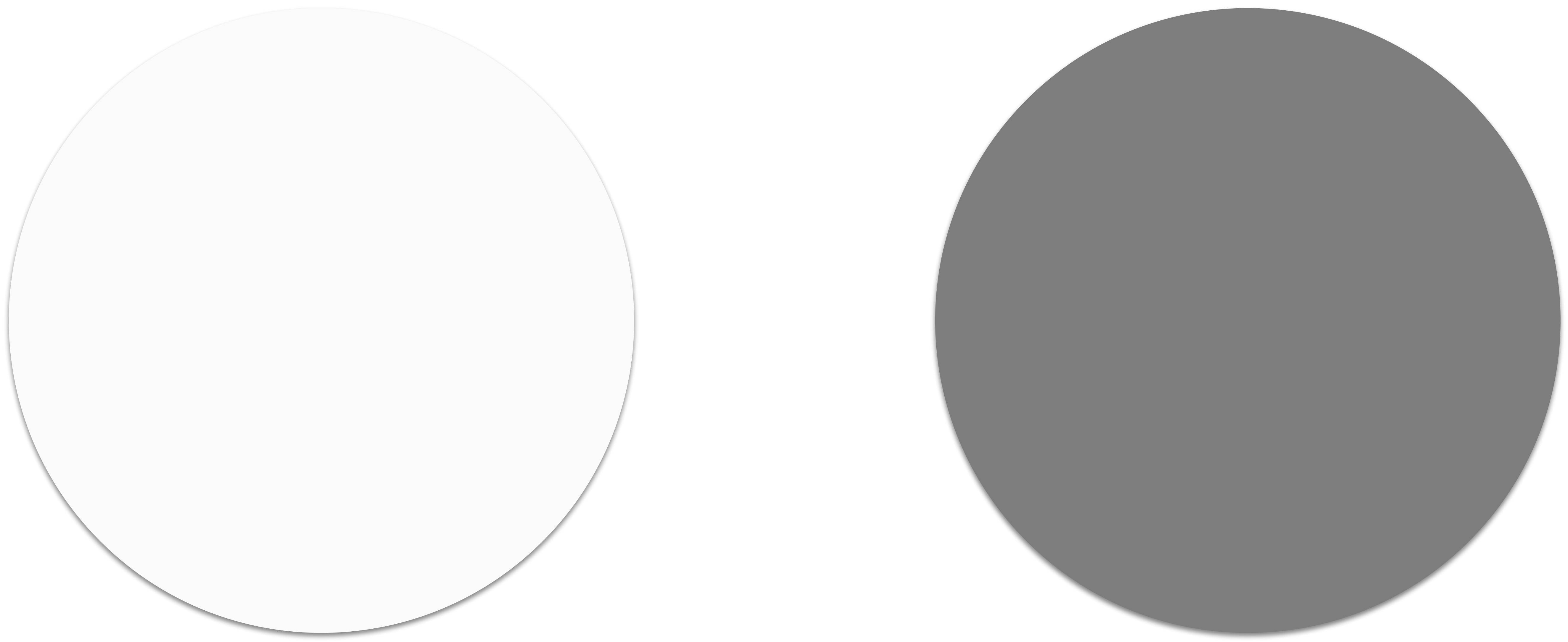
How much darker?



10X

- 20 vs 2

How much darker?

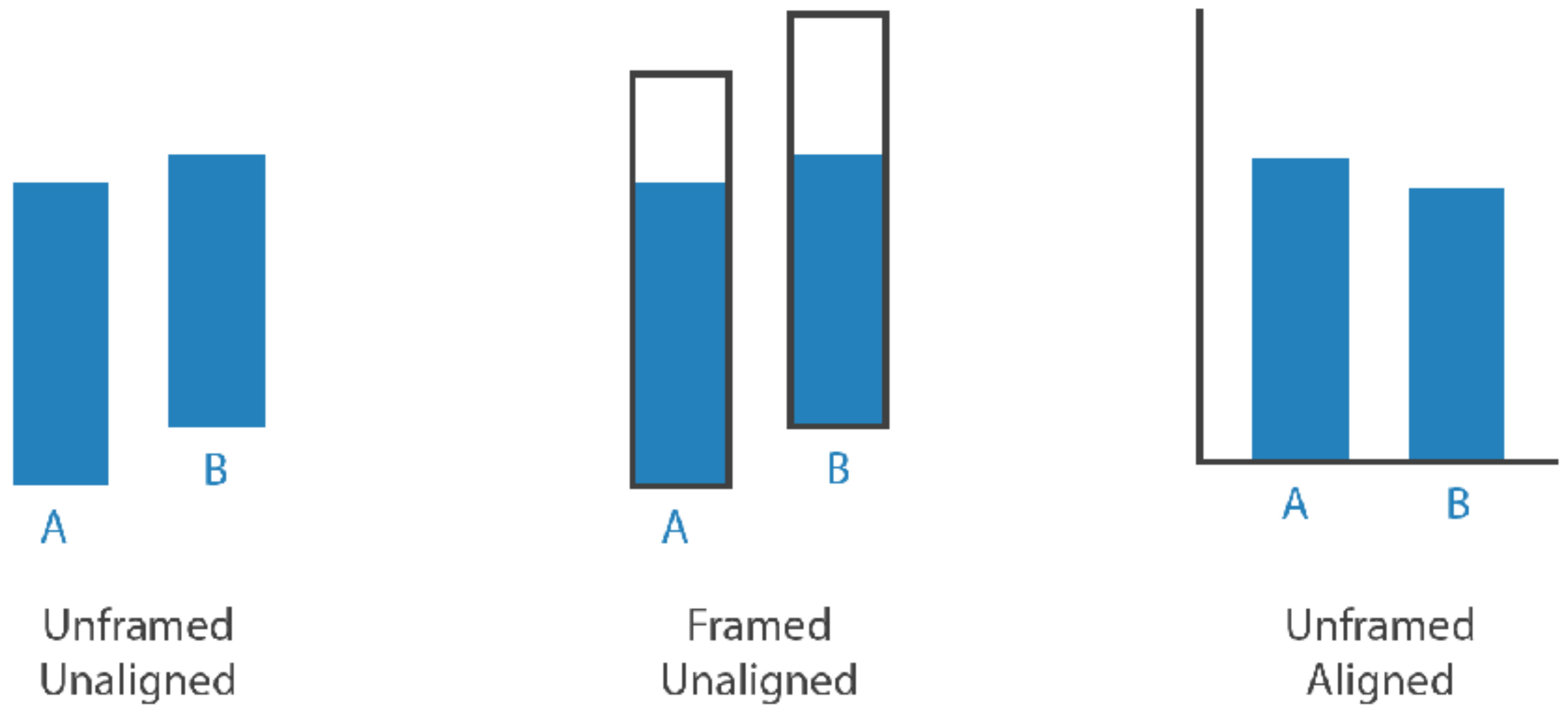


2X

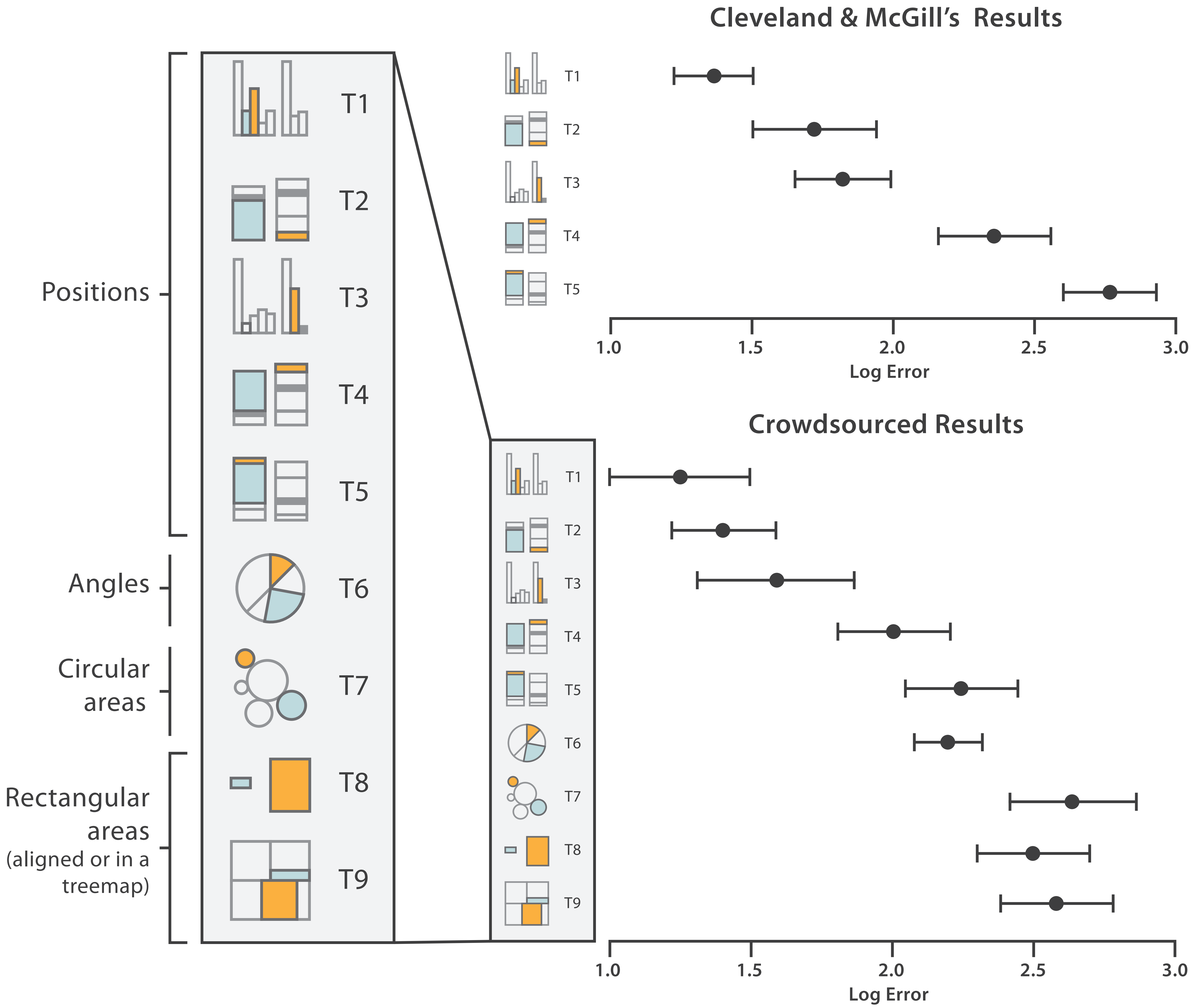
- 250 vs 125

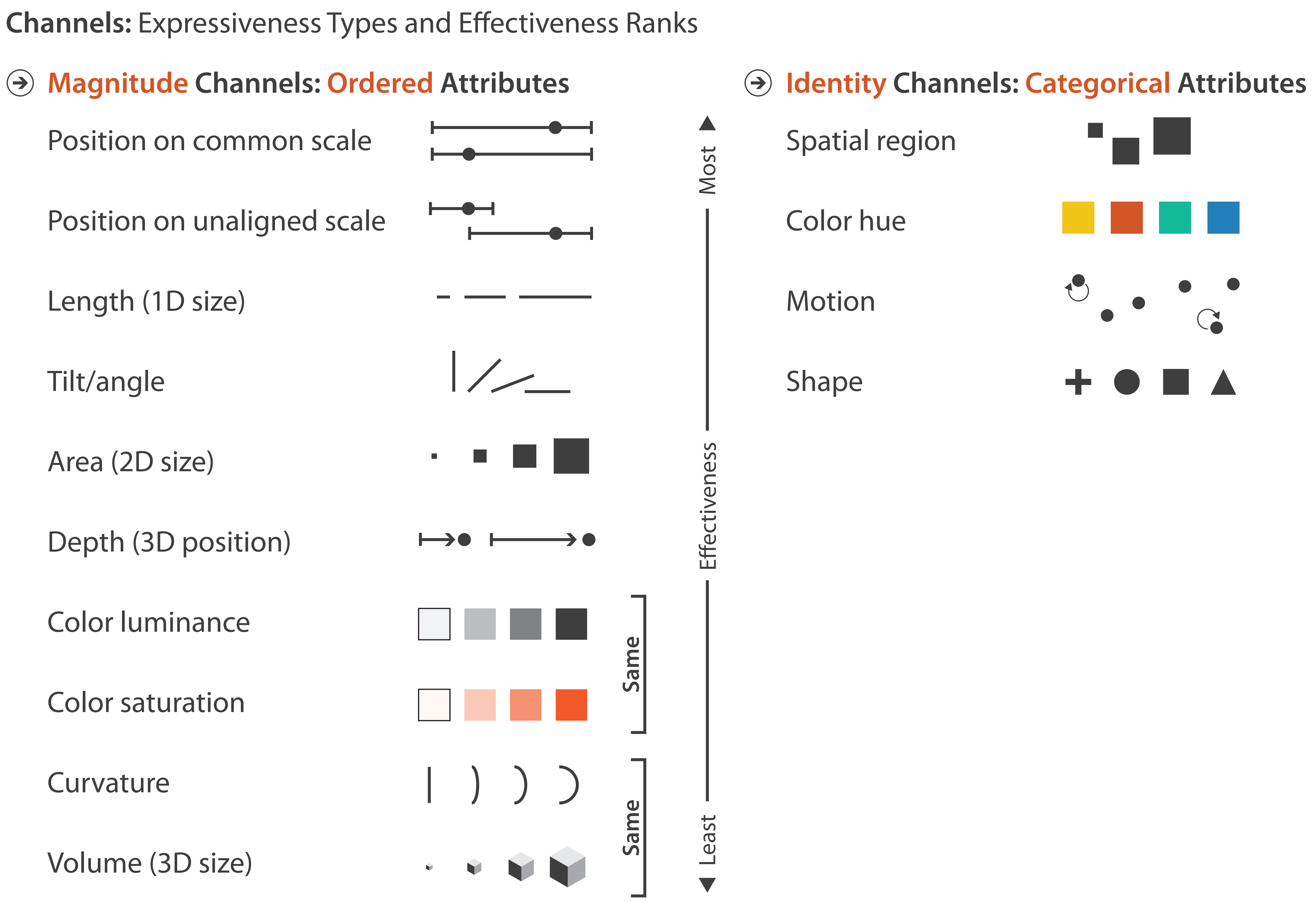
Other Factors Affecting Accuracy

- Alignment
- Distractors
- Distance
- Common scale
- …



Heer & Bostock, 2010   Cleveland / McGill, 1984

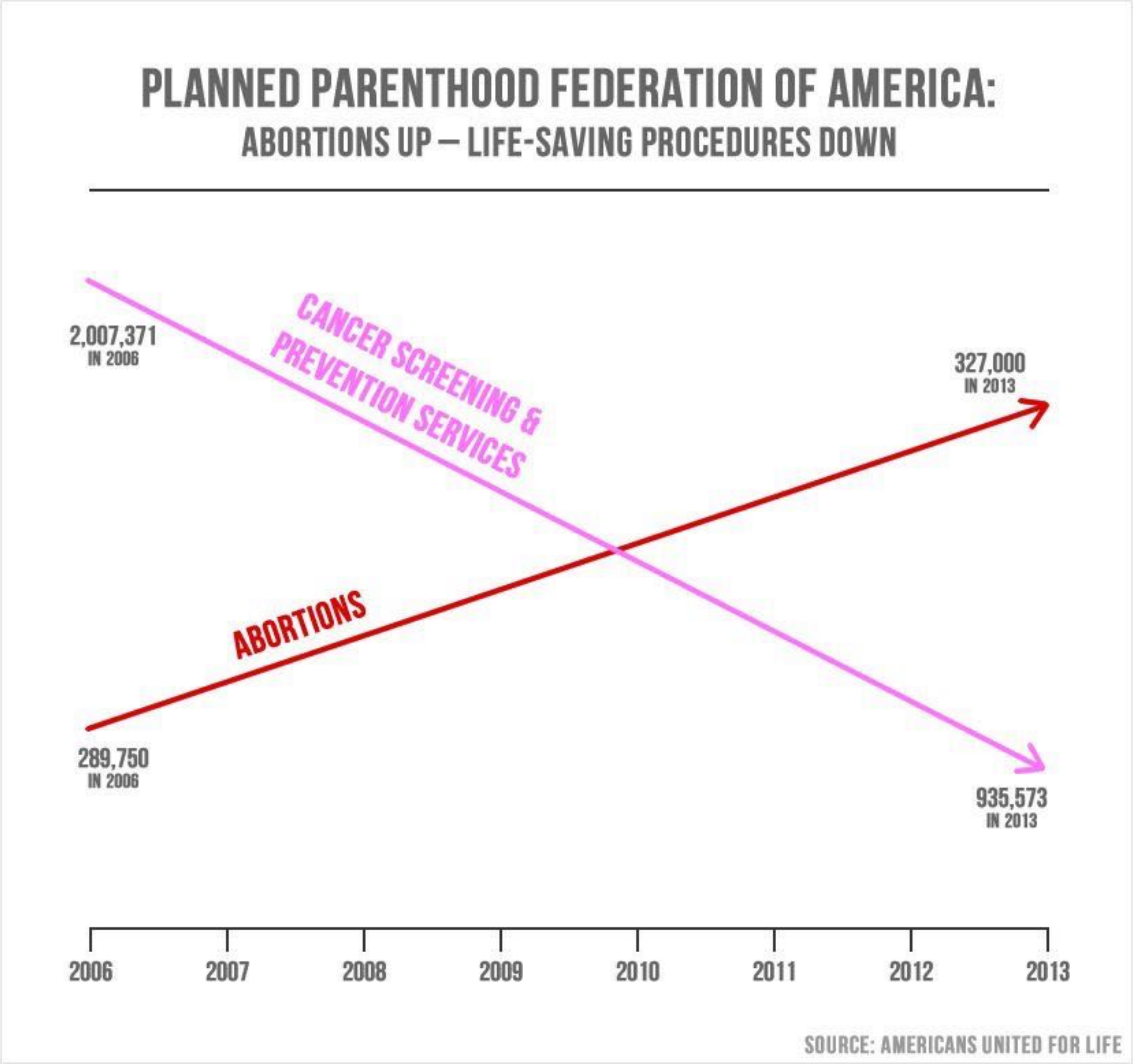



Channels: Expressiveness Types and Effectiveness Ranks
Magnitude Channels: Ordered Attributes
Position on common scale
Position on unaligned scale
Length (1D size)
Tilt/angle
Area (2D size)
Depth (3D position)
Color luminance
Color saturation
Curvature
Volume (3D size)
Same
Same
Most
Effectiveness
Least
Identity Channels: Categorical Attributes
Spatial region
Color hue
Motion
Shape

Some Fun Examples of Bad Vis

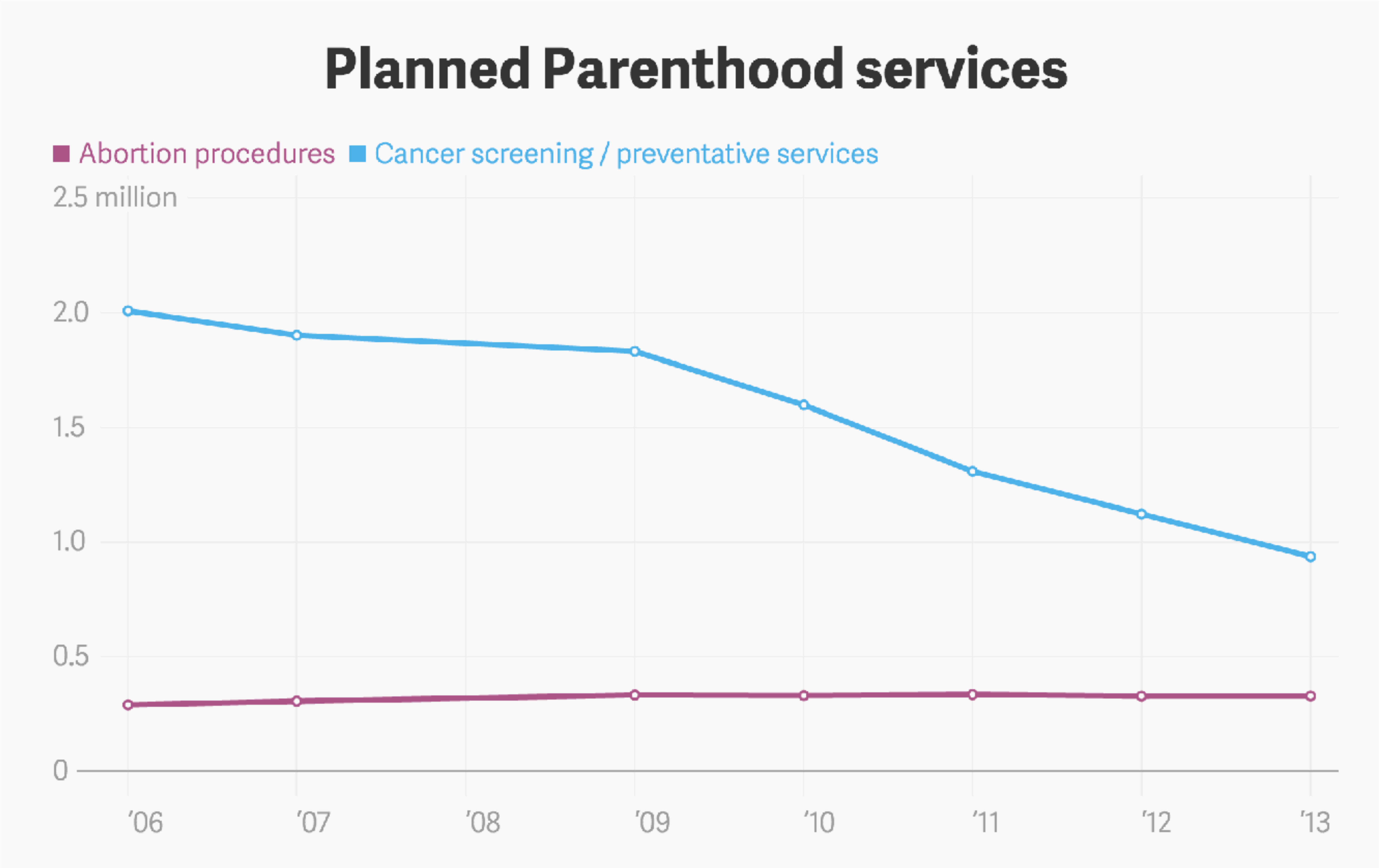
"There are three kinds of lies: lies, damned lies, and statistics."

Bad Charts (http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/)

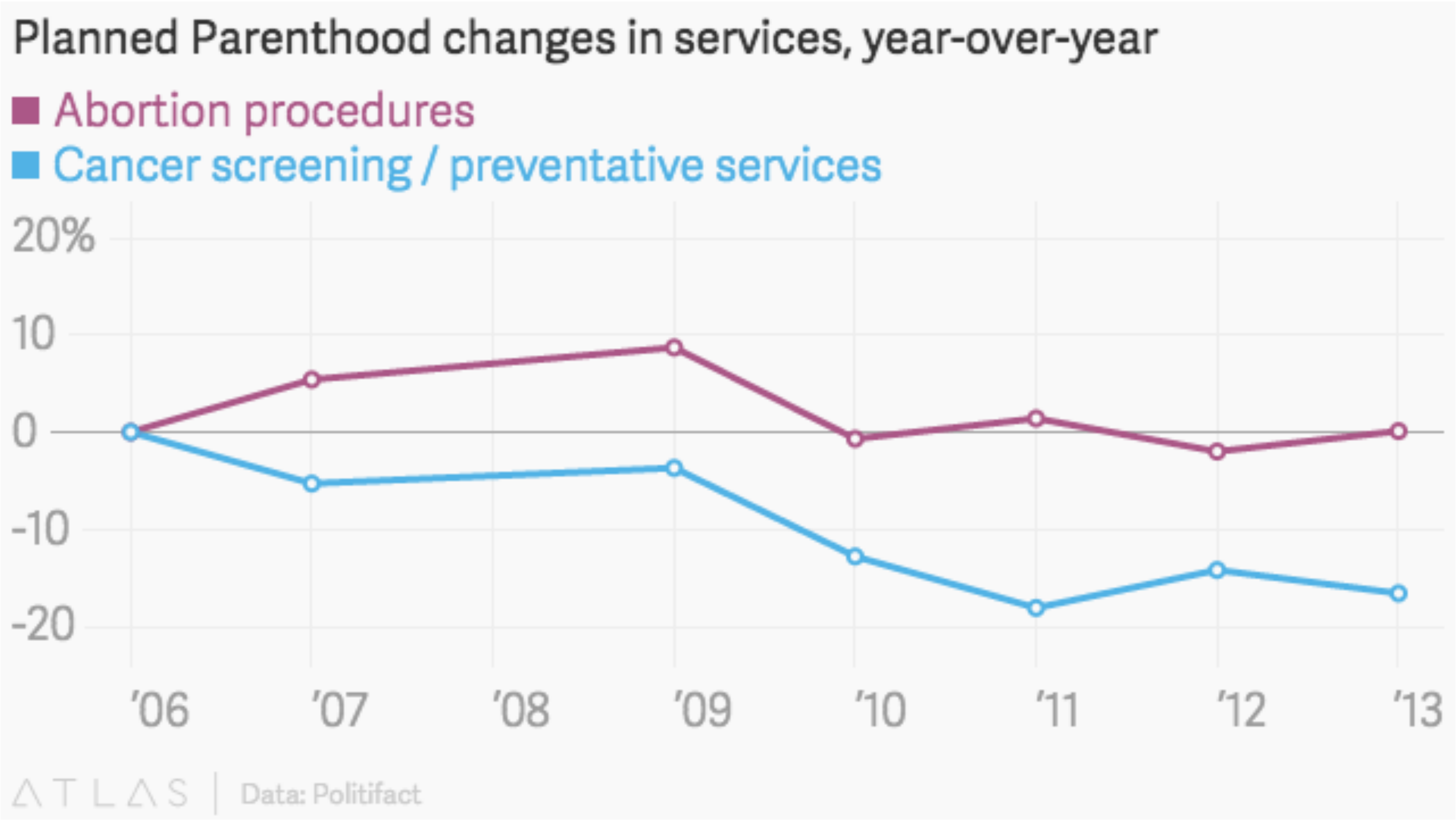


- Rep. Jason Chaffetz of Utah (R) Sept. 29, 2015

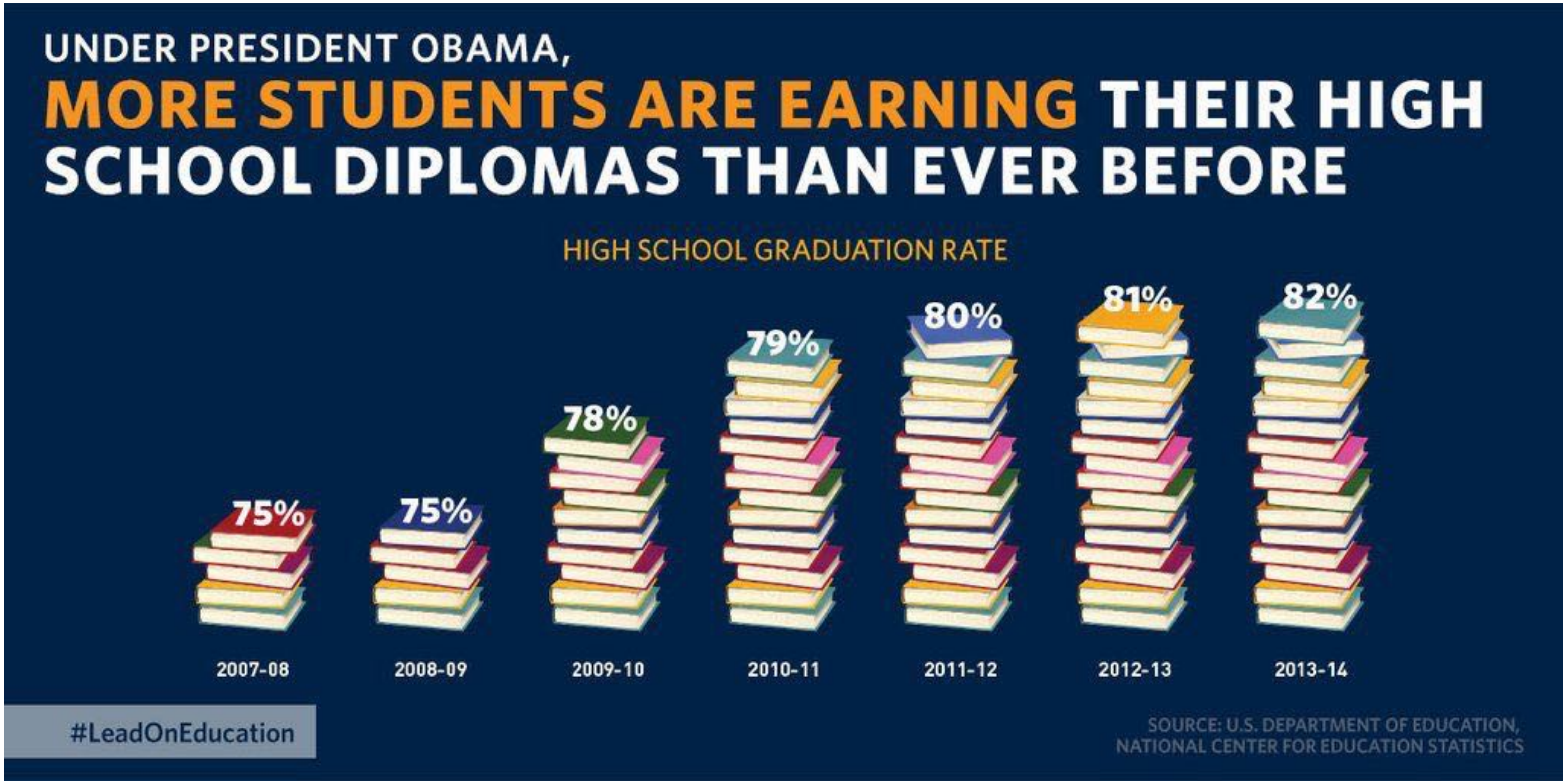
Better Charts (http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/)



Better Charts (http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/)

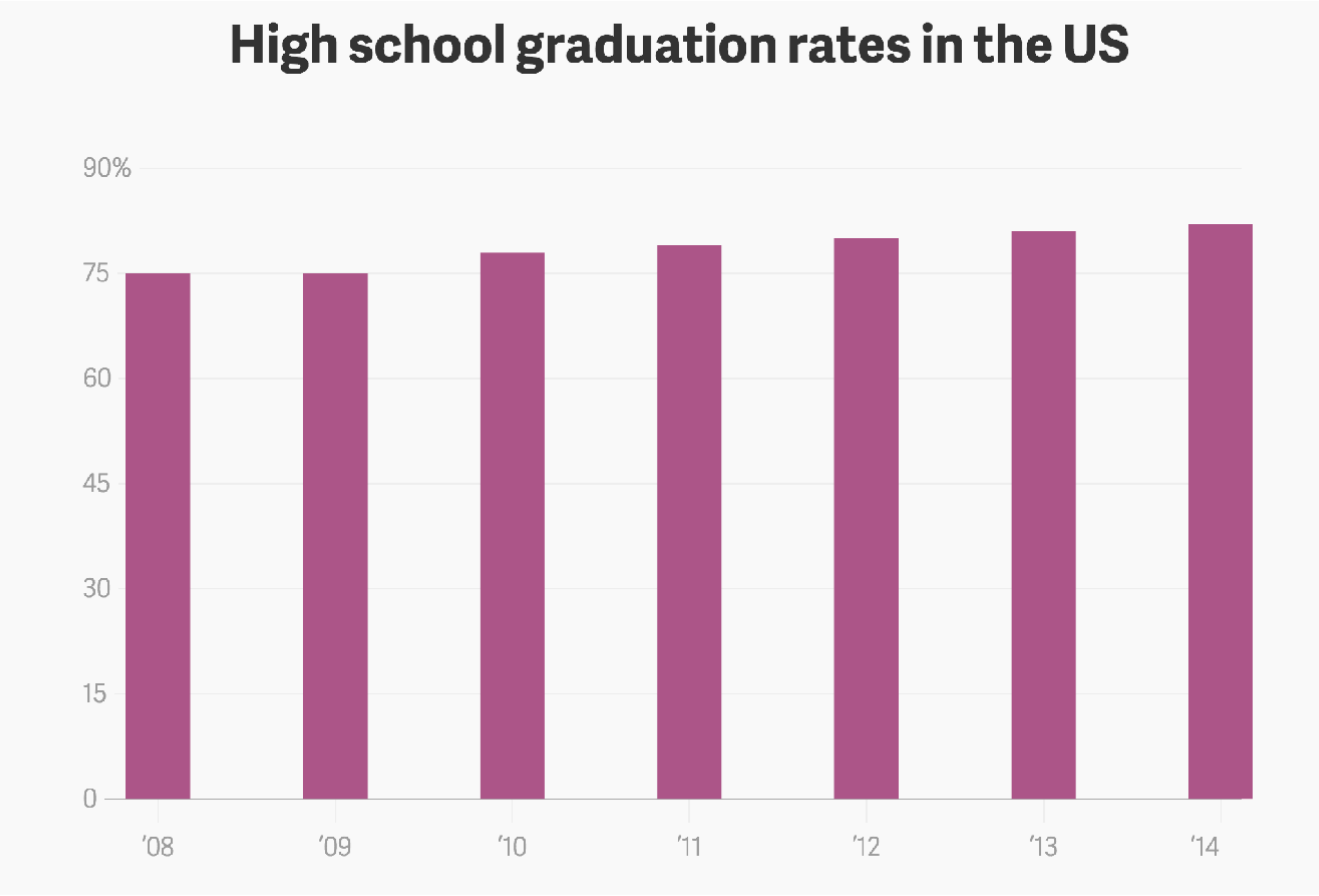


Bad Charts (http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/)

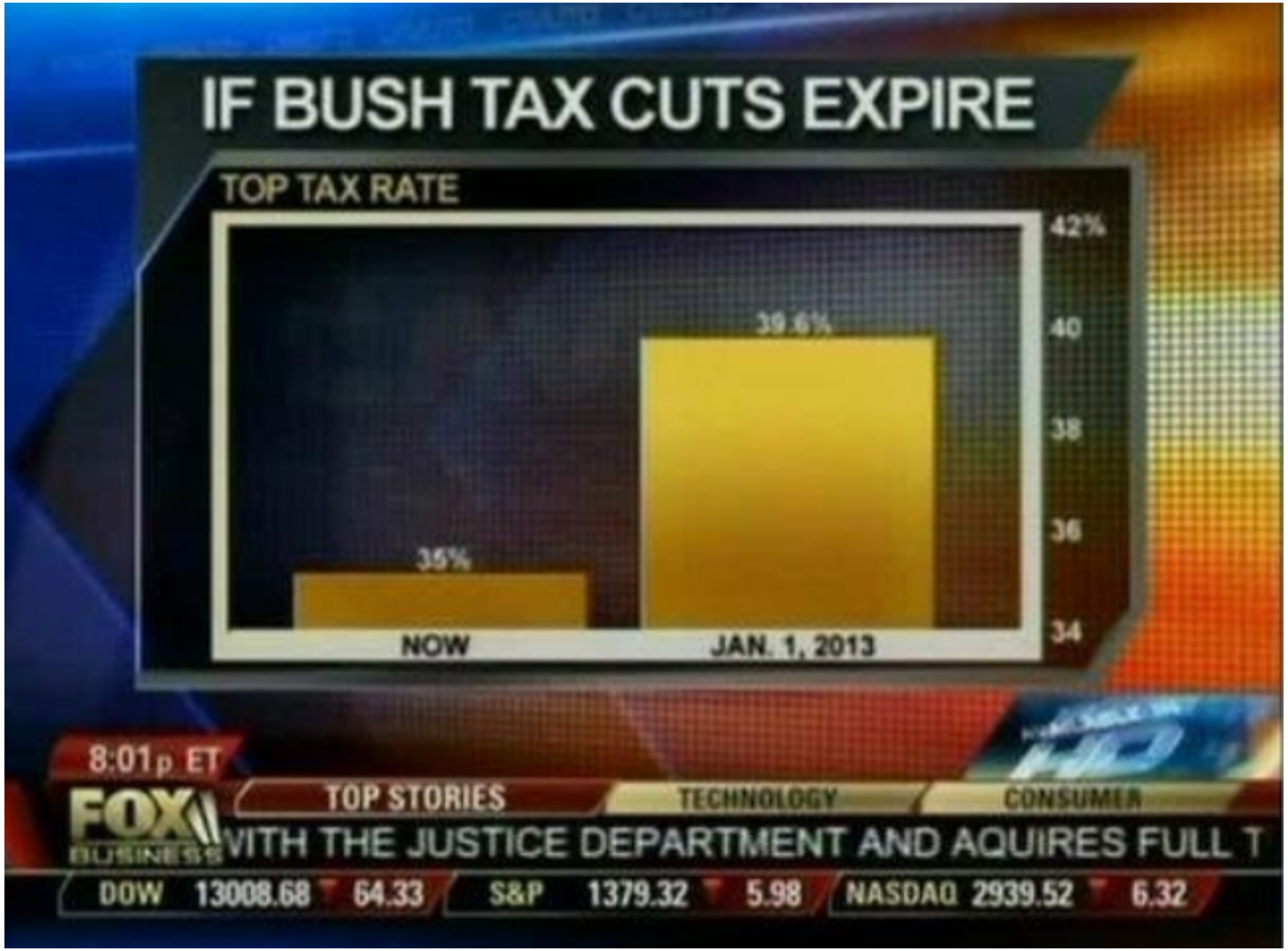


- White House Twitter Account, 2015

Better Charts (http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/)

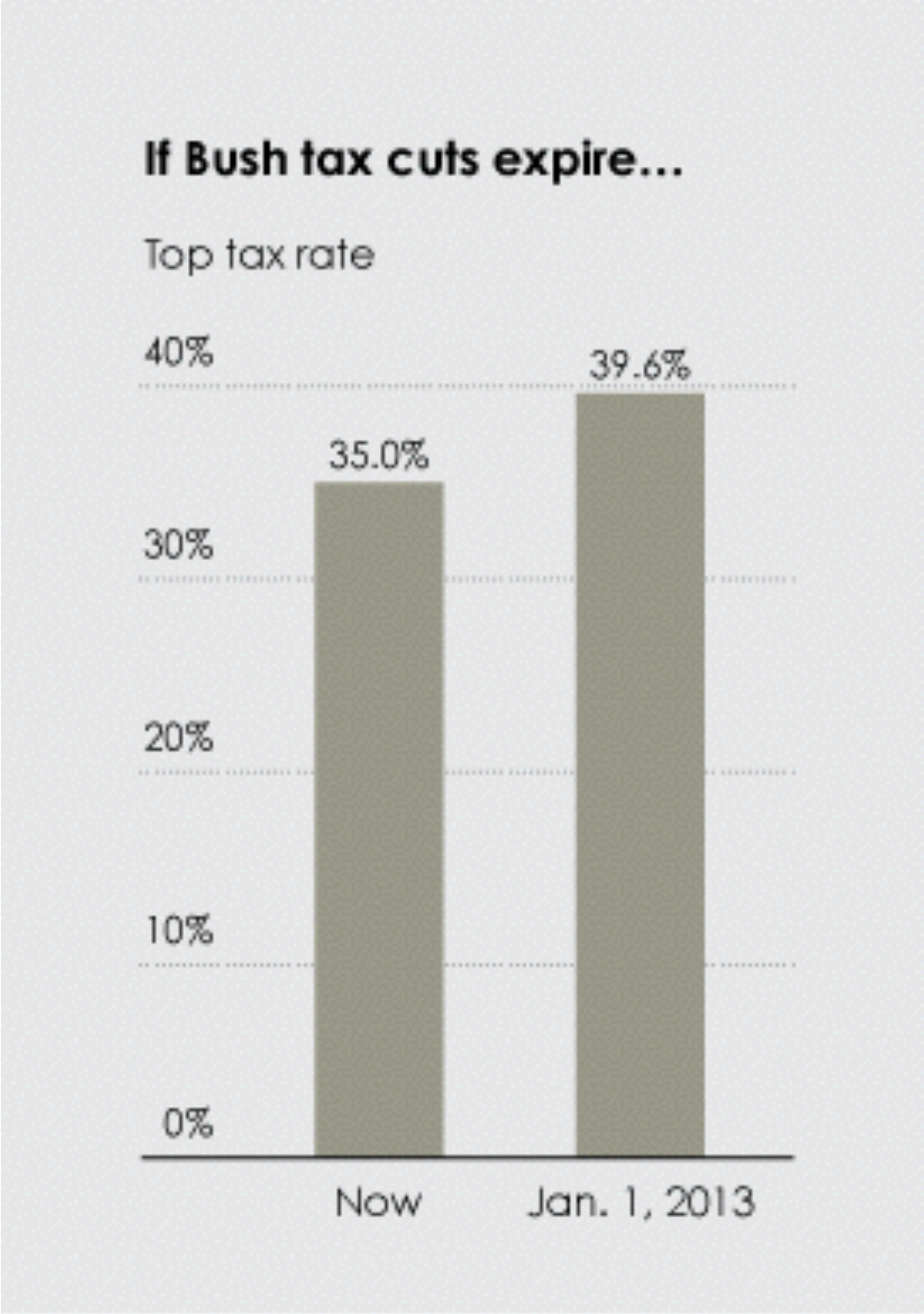


Graphical Integrity



http://flowingdata.com/2012/08/06/fox-news-continues-charting-excellence/

Scale Distortions



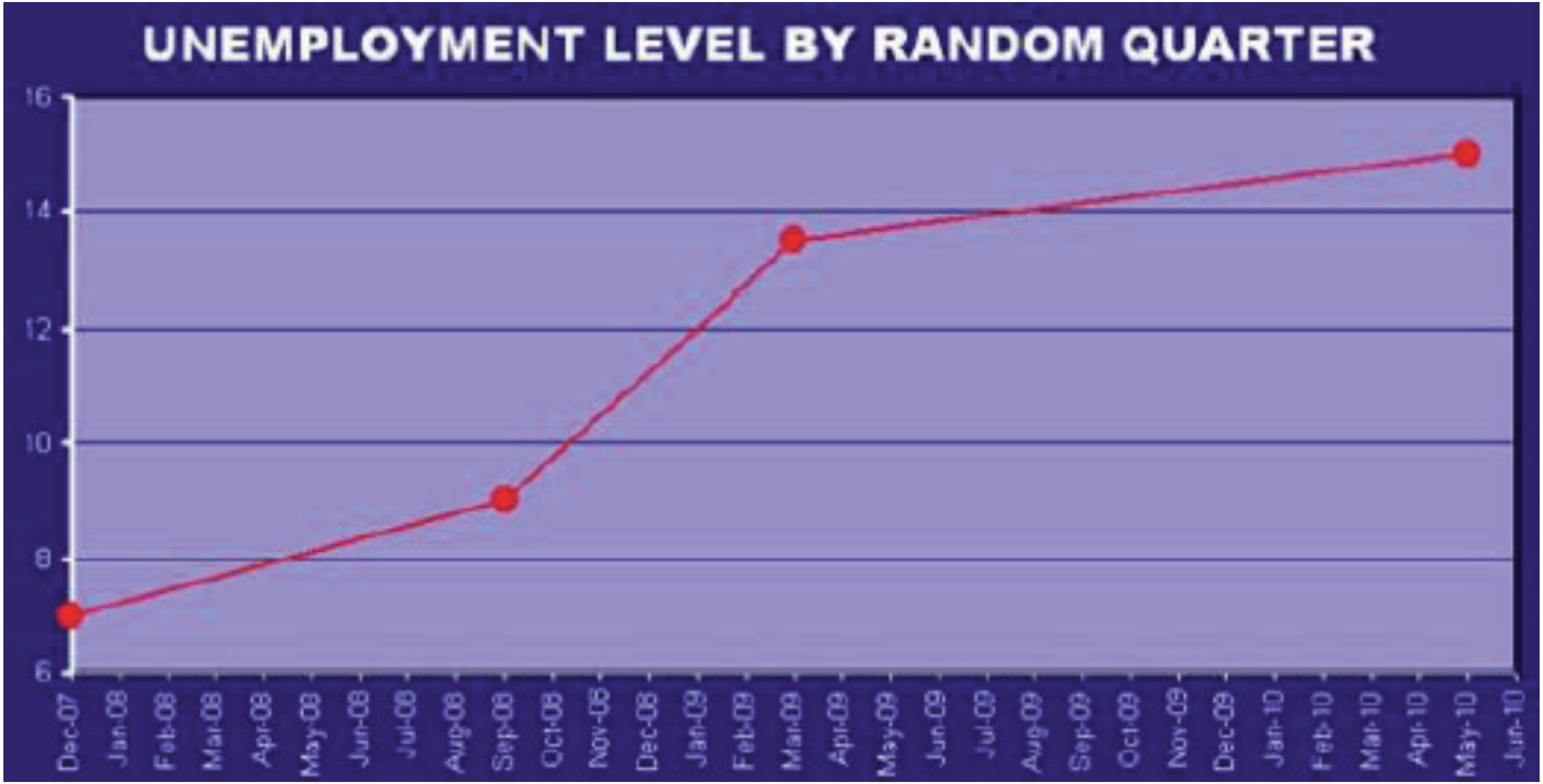
http://flowingdata.com/2012/08/06/fox-news-continues-charting-excellence/

What’s wrong?

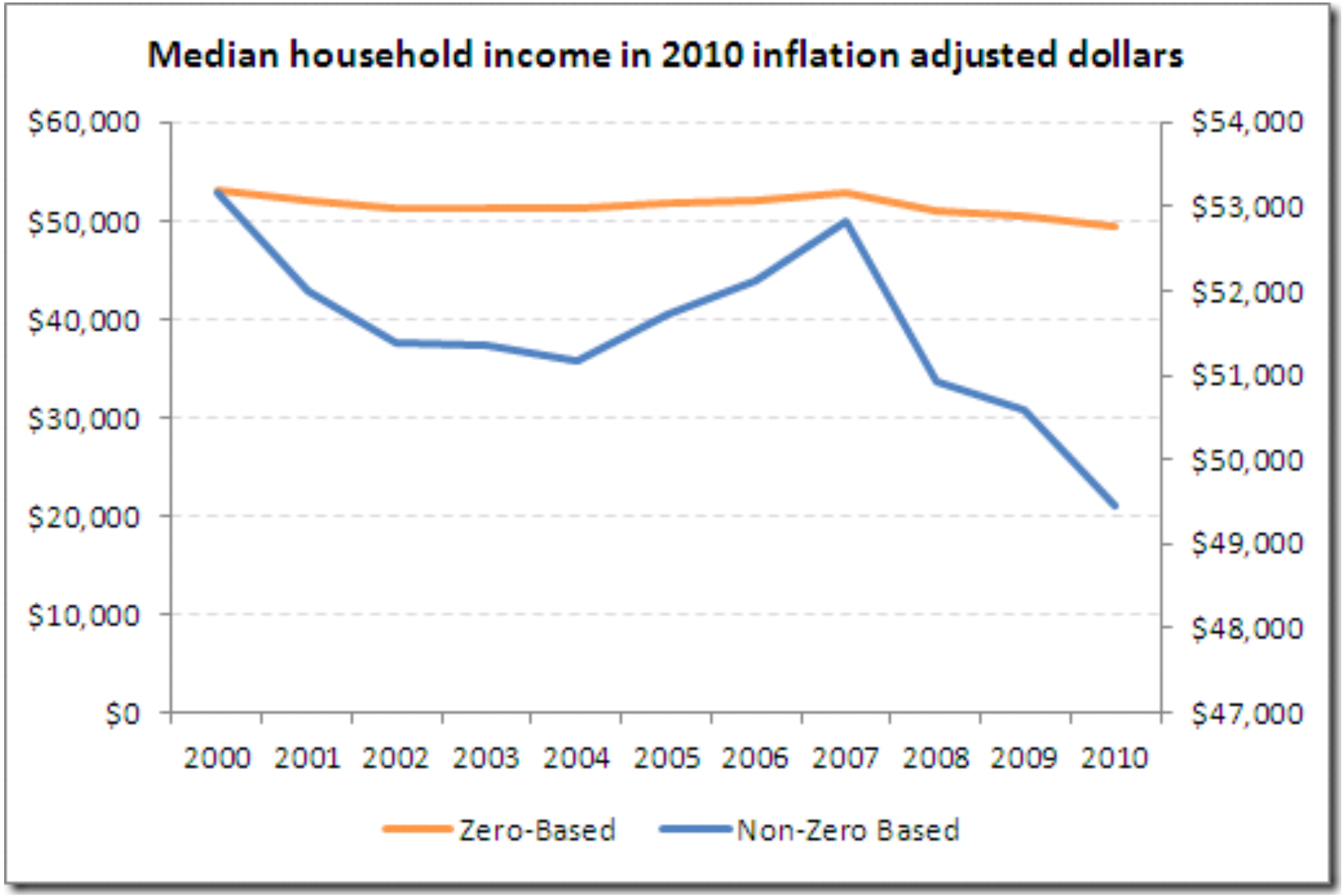

How 2012 Stacks Up
The Warmest Years on Record
Contiguous U.S.
53.81
53.93
54.14
54.31
54.32
55.3
1921
1999
1934
2006
1998
2012
55.5°
55.0°
54.5°
54.0°
53.5°
53.0°
Source: NOAA's National Climatic Data Center - State of the Climate National Overview
CLIMATE CENTRAL

Scale Distortions

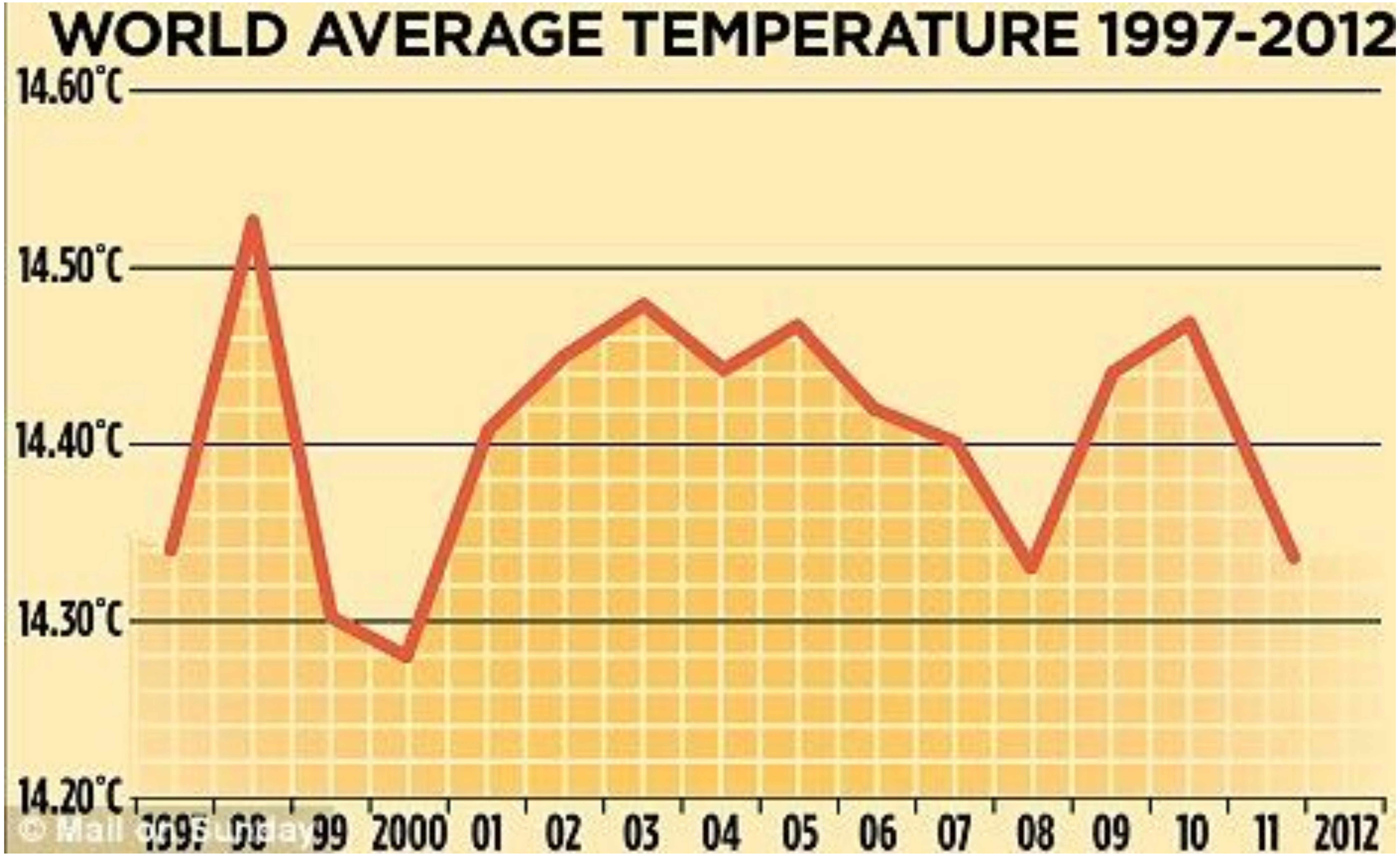
Scale Distortions



Start Scales at 0?

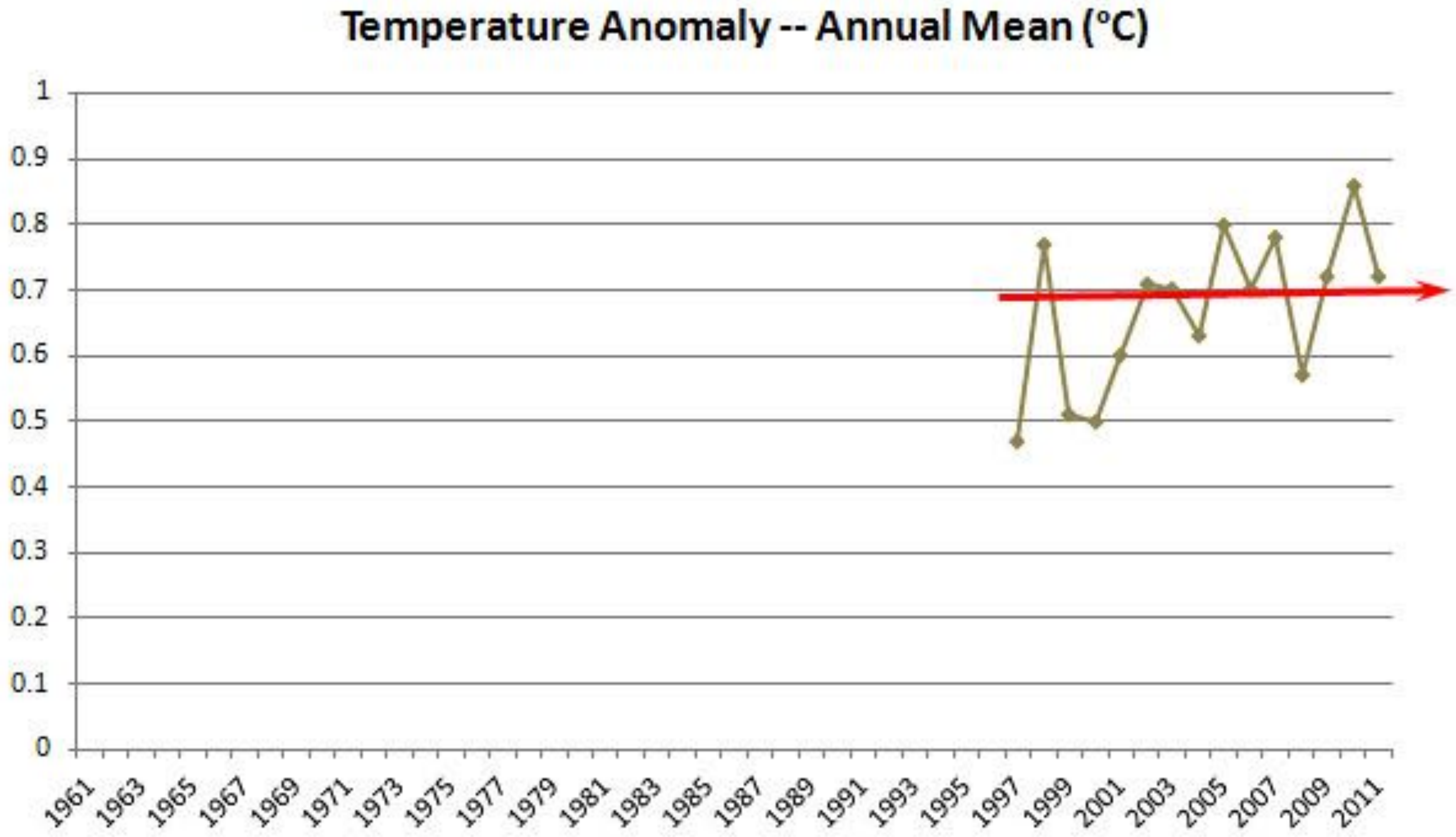


Global Warming?



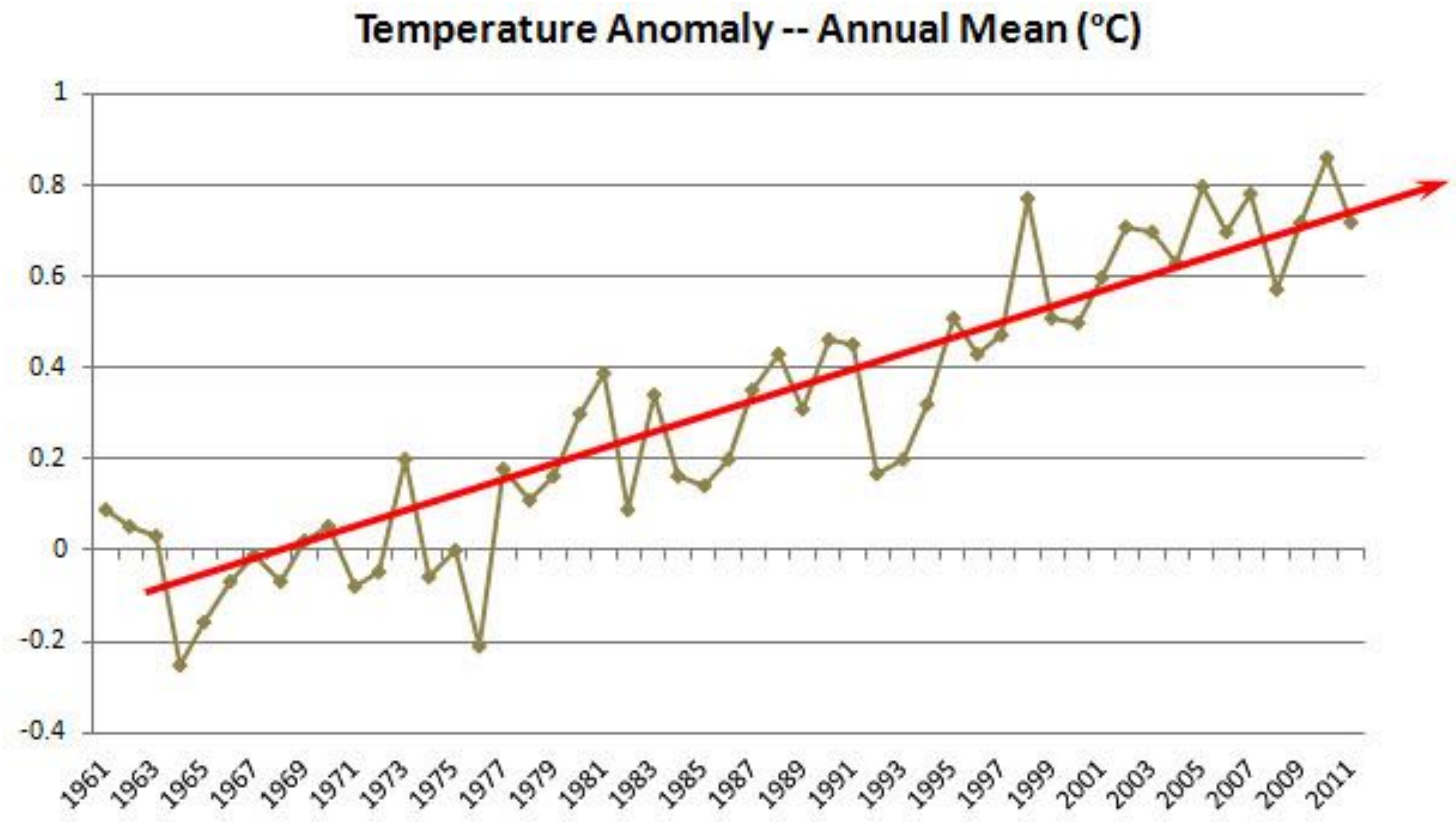
http://www.dailymail.co.uk/sciencetech/article-2093264/Forget-global-warming--Cycle-25-need-worry-NASA-scientists-right-Thames-freezing-again.html

Framing



Temperature Anomaly -- Annual Mean (°C)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011

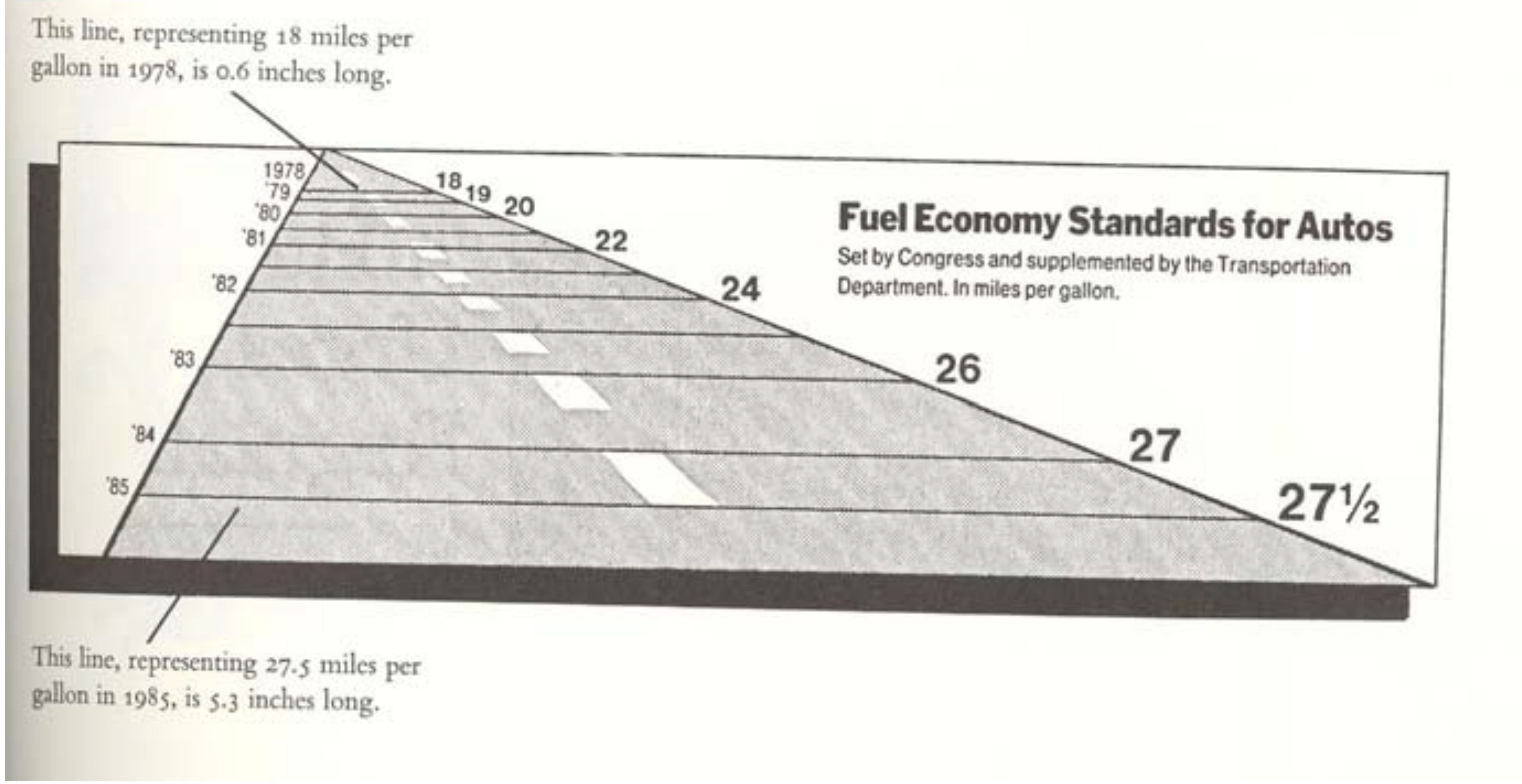
Framing



Temperature Anomaly -- Annual Mean (°C)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011

The Lie Factor

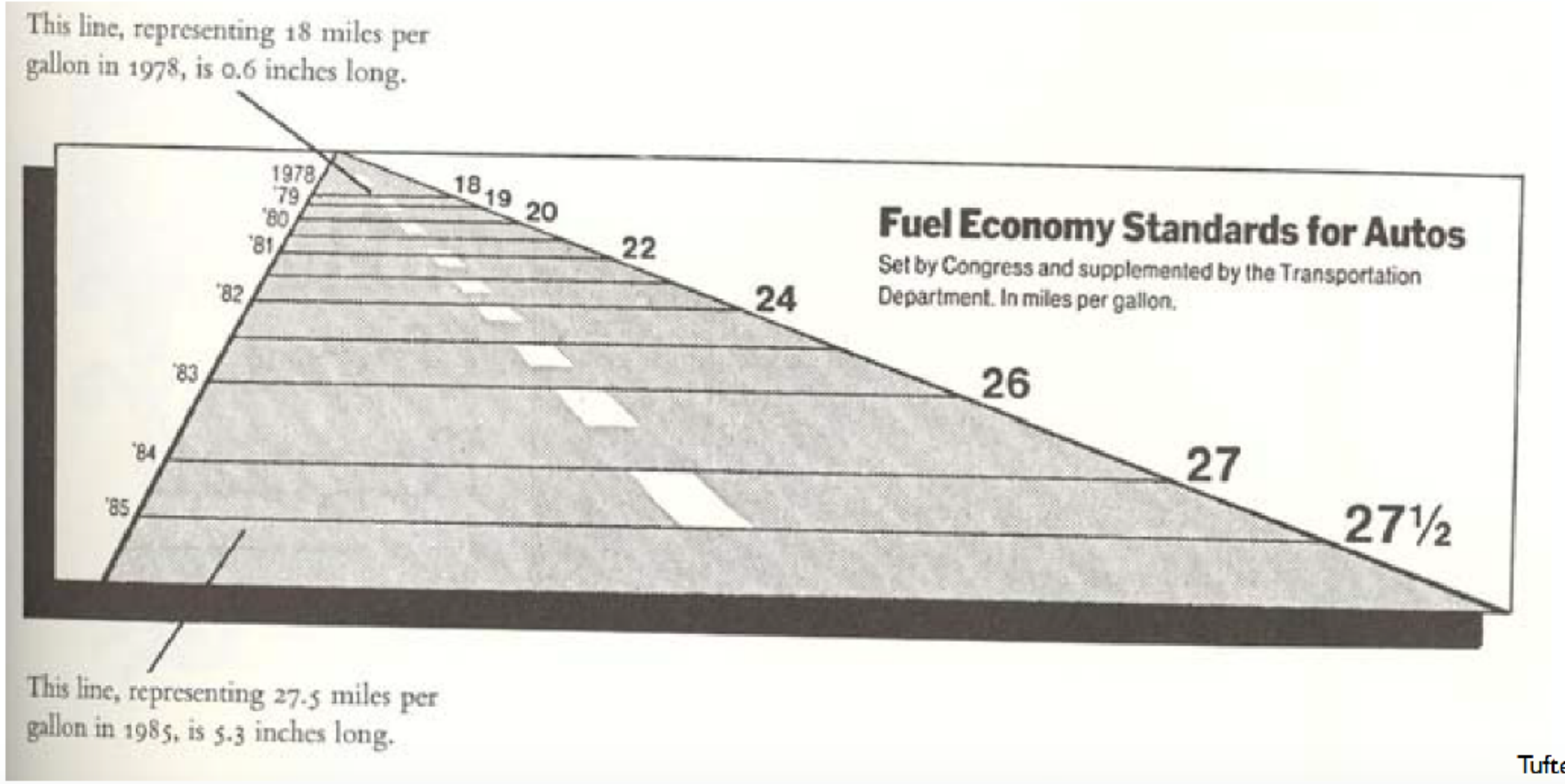
$$\frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}$$



The Lie Factor

$$\frac{5.3-0.6}{0.6} \Big/ \frac{27.5-18}{18} = 14.8$$

(Size of effect in graphic)/(size of effect in data)



Tufte, VDQI

https://viz.wtf or #wtfviz