Stat405

Effective visualisation

Hadley Wickham

1. Roadmap
2. Graphics
3. Perception

Roadmap

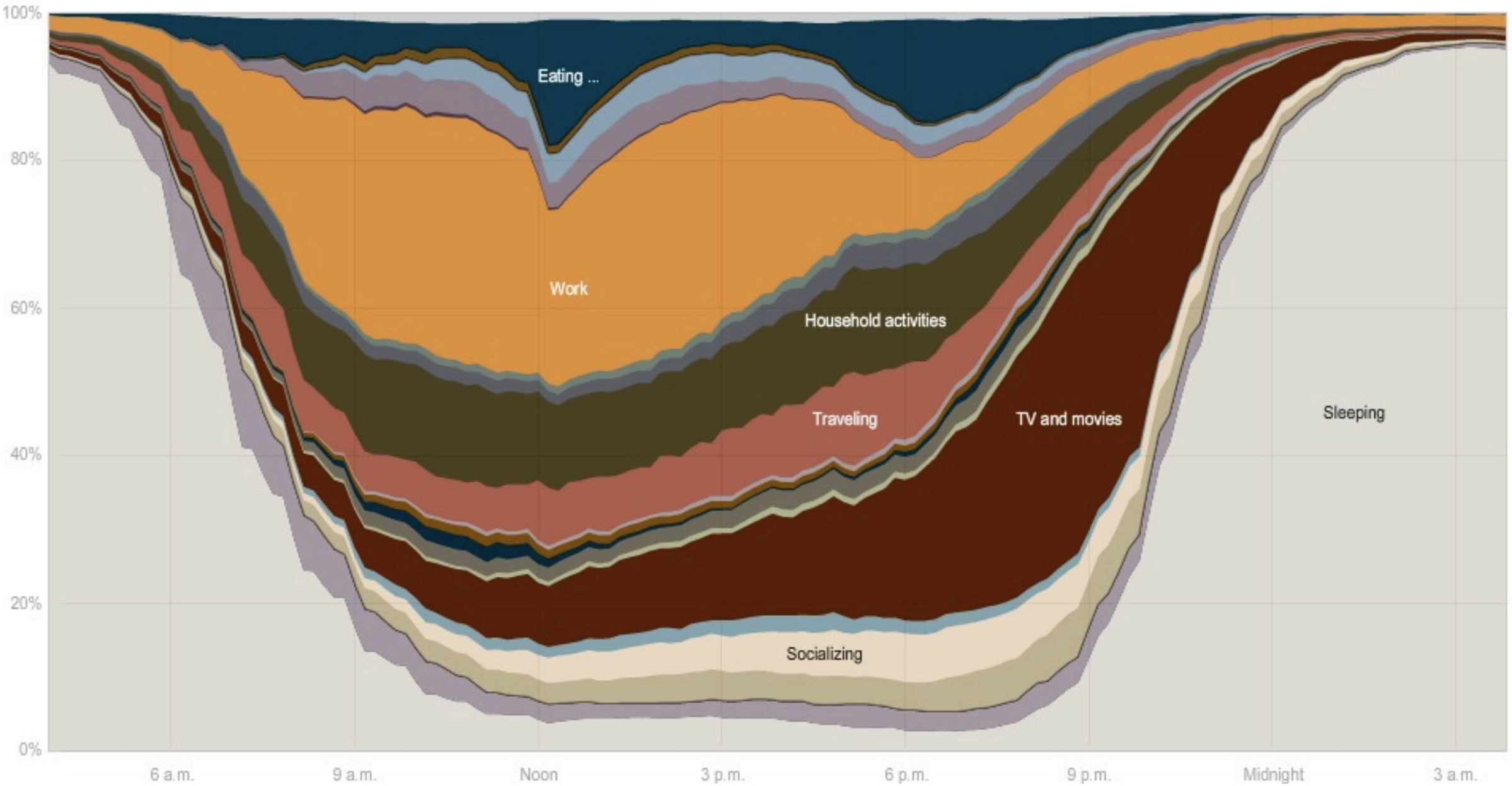
- Lectures 20-22: advanced graphics
- Lectures 23-25: advanced data handling, incl. modelling
- Lectures 26-27: advanced functions
- Lecture 28: Final poster presentation

Graphics

Everyone

Sleeping, eating, working and watching television take up about two-thirds of the average day.

Everyone	Employed	White	Age 15-24	H.S. grads	No children
Men	Unemployed	Black	Age 25-64	Bachelor's	One child
Women	Not in lab...	Hispanic	Age 65+	Advanced	Two+ children

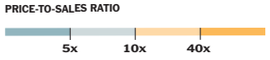


http://nyti.ms/np29Yk

What Happens After the I.P.O.?

There have been about 2,400 technology, Internet and telecom I.P.O.'s since 1980. On the first day of trading, the average stock rose 32 percent above its offer price. But in the three years after that, most companies had negative returns, according to statistics compiled by Jay Ritter, a professor of finance at the University of Florida. Companies with higher values compared with their revenue before the I.P.O. have fared especially poorly.

Circles are sized by value at the end of the first trading day, in today's dollars

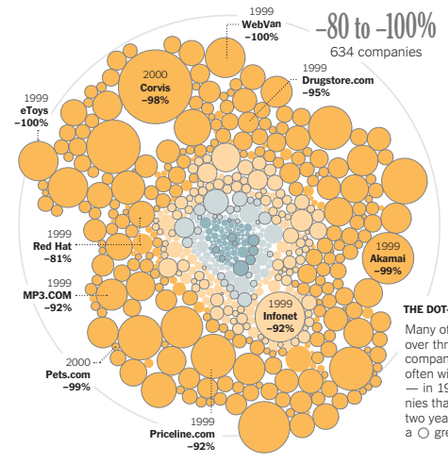


Colors show the ratio of the company's value to its revenue in the 12 months before the I.P.O.

HOW FACEBOOK COMPARES

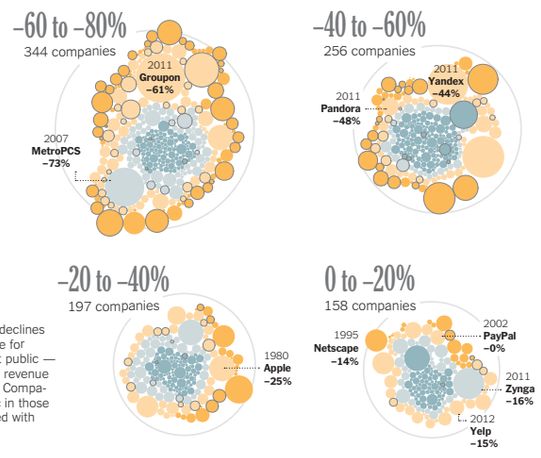
At its offer price, Facebook's market value is $tk billion, more than four times that of Google at its I.P.O. in 2004. Facebook had revenue of about 4 billion in the last year, meaning it will have one of the higher price-to-sales ratios, especially outside of the dot-com bubble.

Return three years after the I.P.O.: The decliners . . .

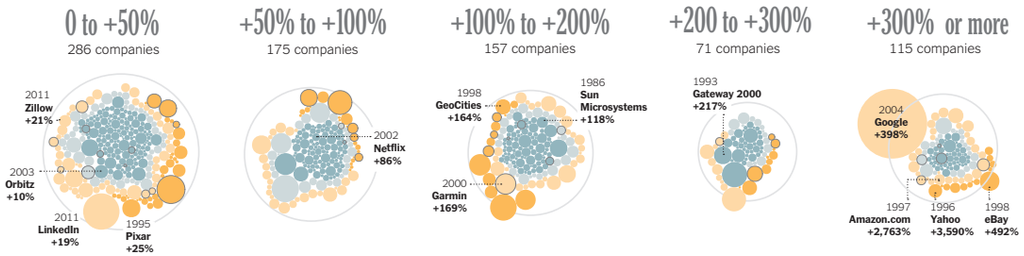


THE DOT-COM BUST

Many of the biggest declines over three years were for companies that went public — often with little or no revenue — in 1999 or 2000. Companies that went public in those two years are outlined with a ○ grey stroke.



. . . and the gainers.



Over the Long Haul

Performance after three years, however, is not necessarily indicative of a company's future. Yahoo skyrocketed only to plummet, while Apple took decades to rise. A look at how Facebook's current market value and revenue compare to five other notable technology I.P.O.'s.

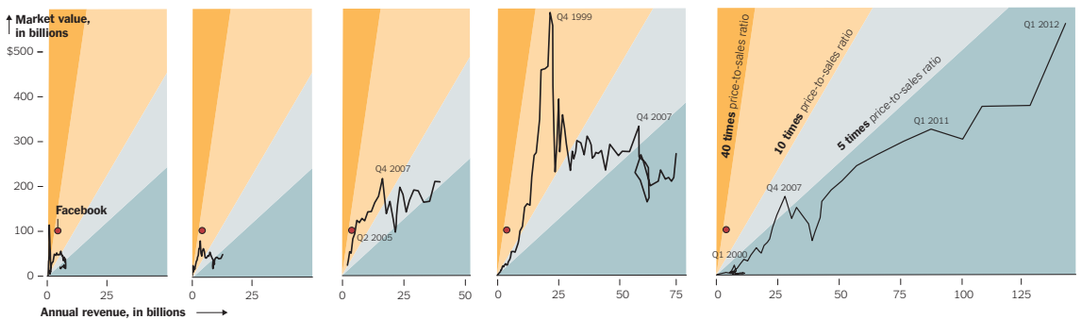
Yahoo (1996 I.P.O.)
Valued above $125 billion in early 2000, Yahoo was worth one-tenth that amount two years later.

eBay (1998)
In 2004, eBay had sales and value similar to Facebook now. Sales have nearly quadrupled since, but its value has fallen.

Google (2004)
Google currently has a price-to-sales ratio around 5. But in June 2005, it had values similar to Facebook's current levels.

Microsoft (1986)
Although Microsoft's revenue has increased most years, its stock price has been relatively flat since it fell in 2000.

Apple (1980)
Three years after its I.P.O., investors in Apple had lost 25 percent. But nearly three decades later in May 2010, it surpassed Microsoft in market capitalization and it is now the largest company by market capitalization.

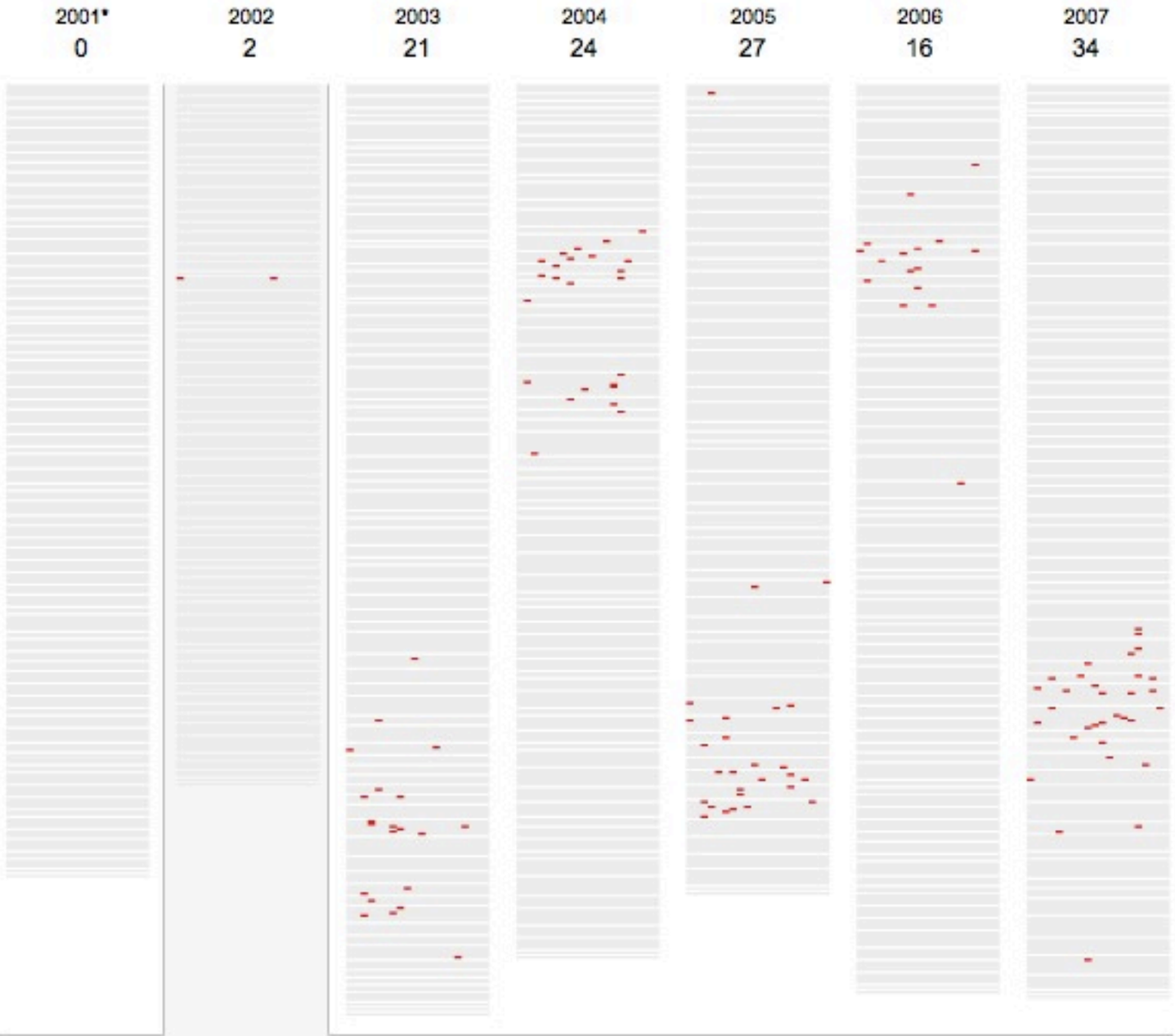


Sources: Jay Ritter, University of Florida; Compustat; Bloomberg

Note: Returns through Monday are shown for companies with I.P.O.'s in May 2009 or later.

AMANDA COX AND SETH W. FEASTER/THE NEW YORK TIMES

Use of the phrase "Iraq" in past State of the Union Addresses



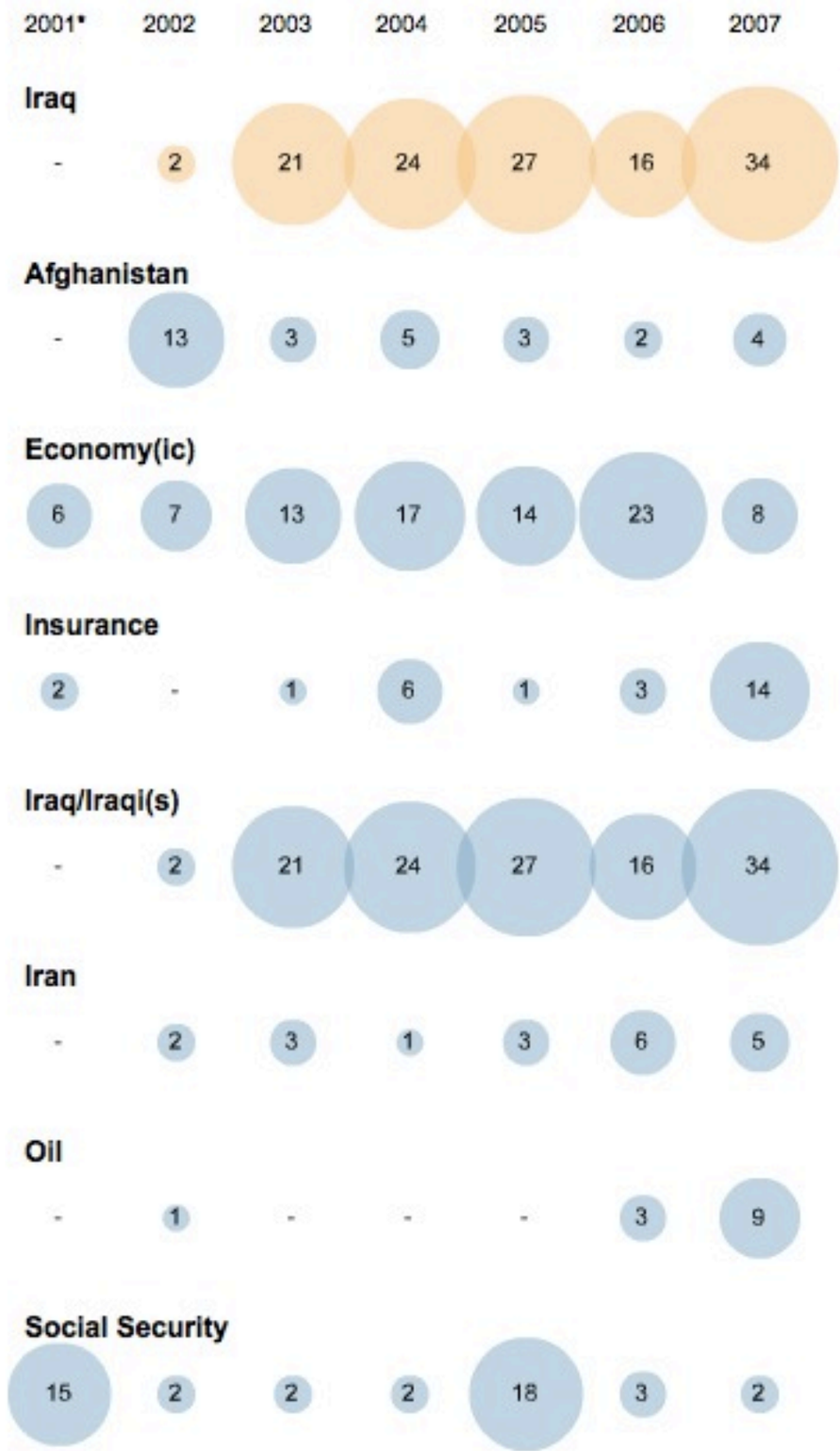
The word in context

Next Instance of 'Iraq'

IRAQ continues to flaunt its hostility toward America and to support terror. The Iraqi regime has plotted to develop anthrax, and nerve gas, and nuclear weapons for over a decade. This is a regime that has already used poison gas to murder thousands of its own citizens -- leaving the bodies of mothers huddled over their dead children. This is a regime that agreed to international inspections -- then kicked out the inspectors. This is a regime that has something to hide from the civilized world.

-- 2002 (Paragraph 20 of 67)

Compared with other words



http://nyti.ms/r8KdvU

Your turn

In small groups, identify the data and non-data in each of the three plots. Which features are the most important? Which are just useful background information?

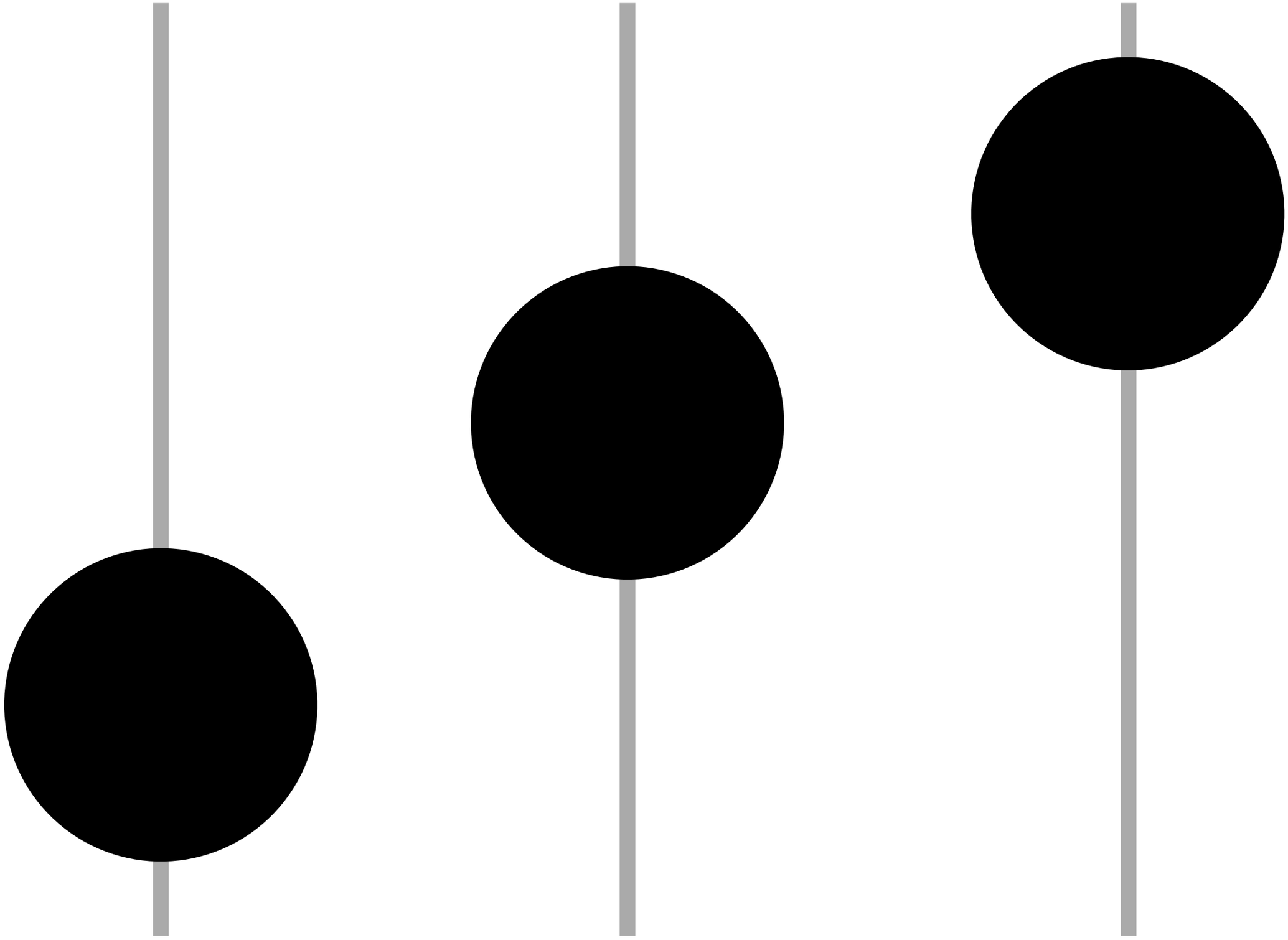
Perception

http://www.dreamsystemsmedia.com/blog/index.php/social-media-statistics-of-the-day/

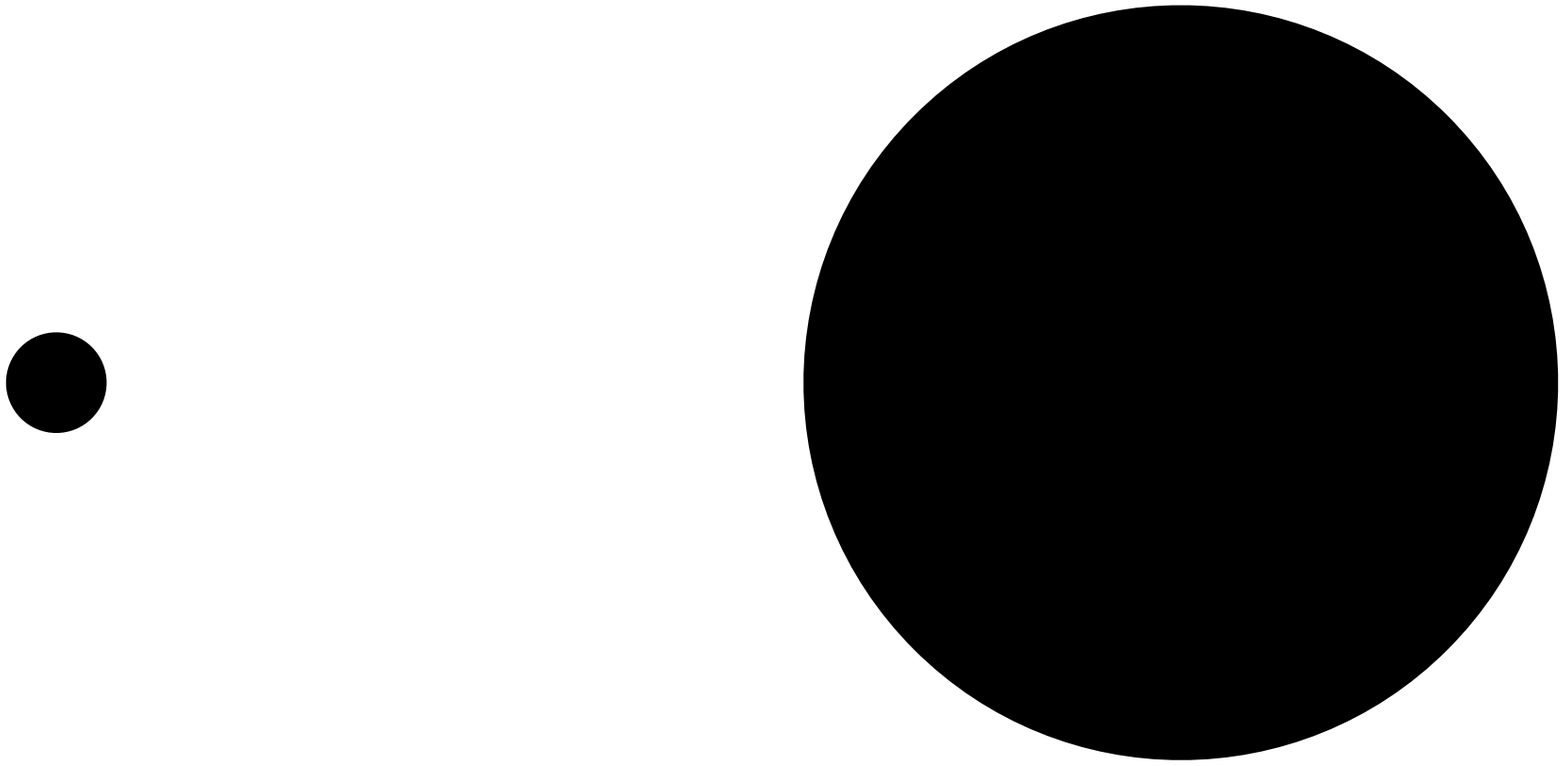
1

Match perceptual and data topology

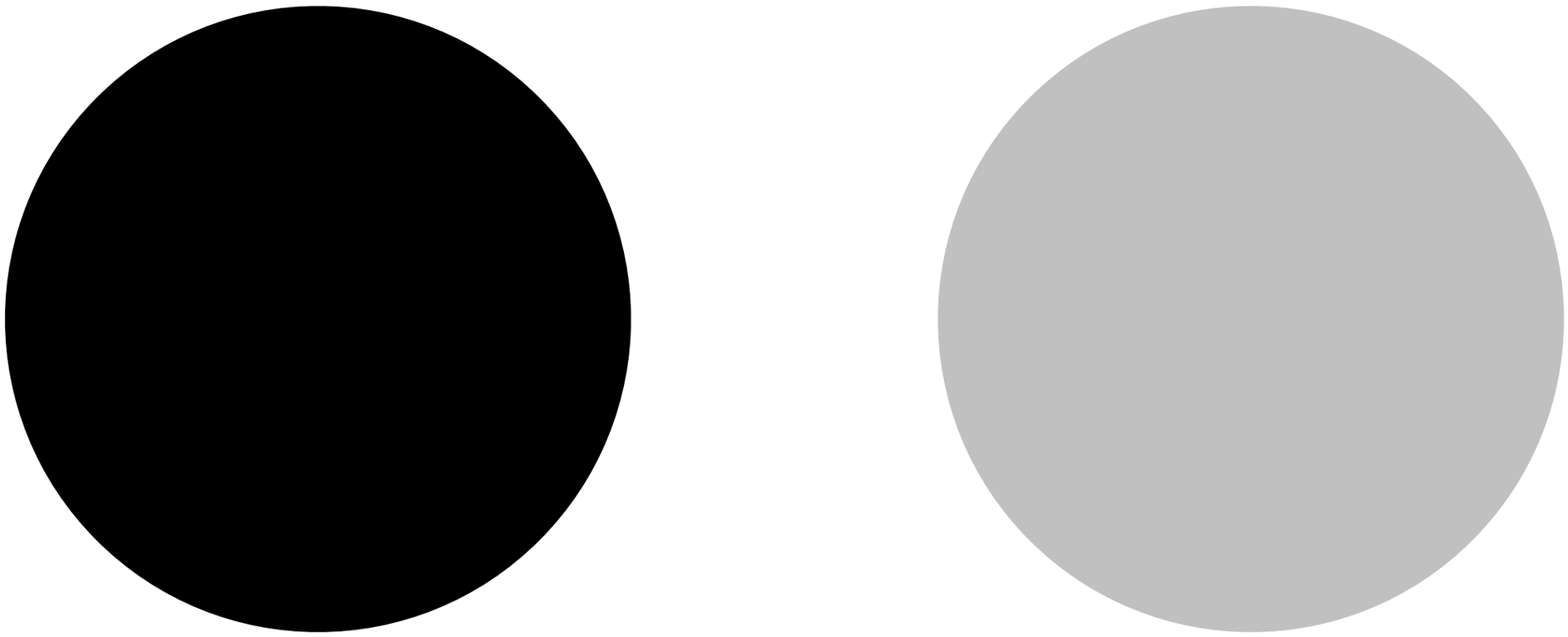
Which represents the larger value?



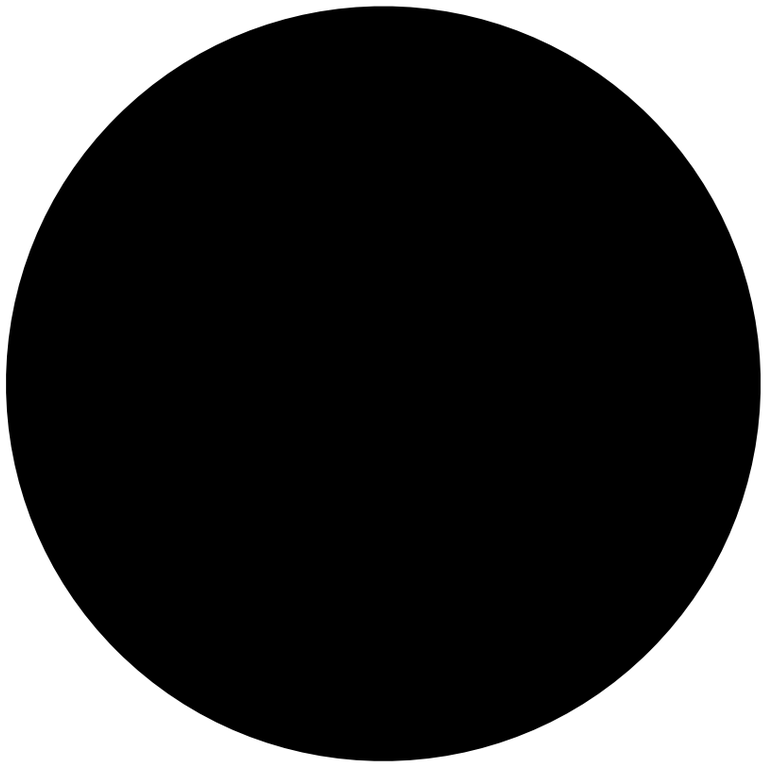
Which represents the larger value?

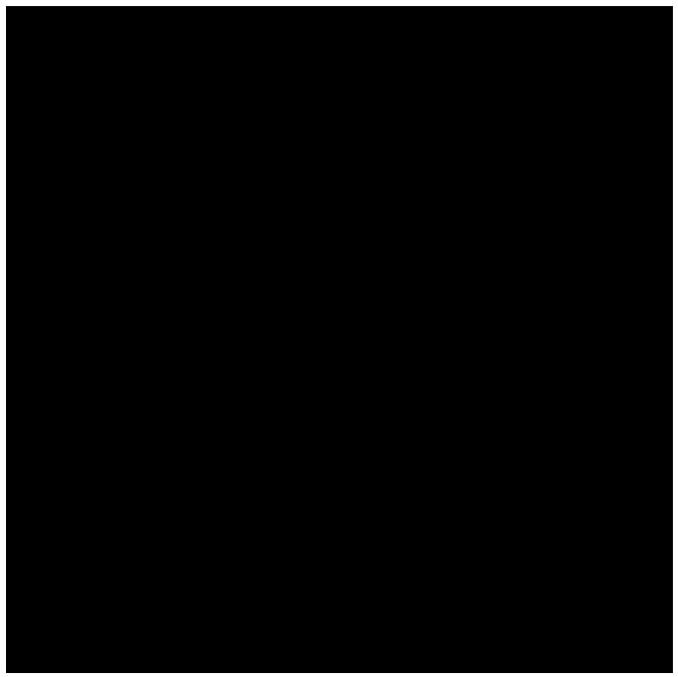


Which represents the larger value?



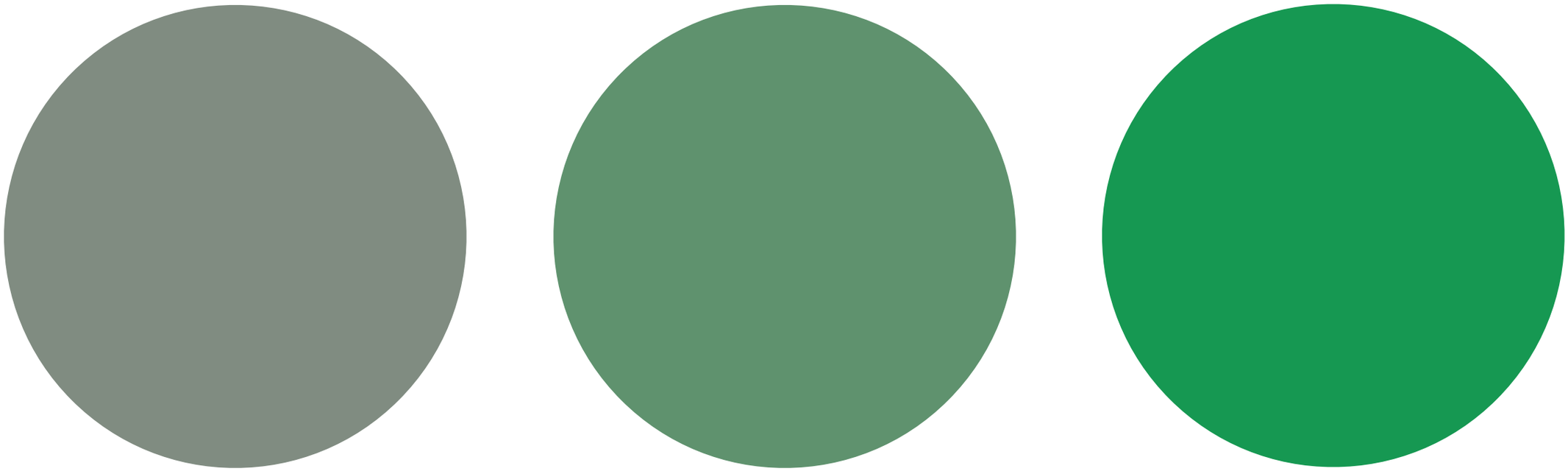
Which represents the larger value?



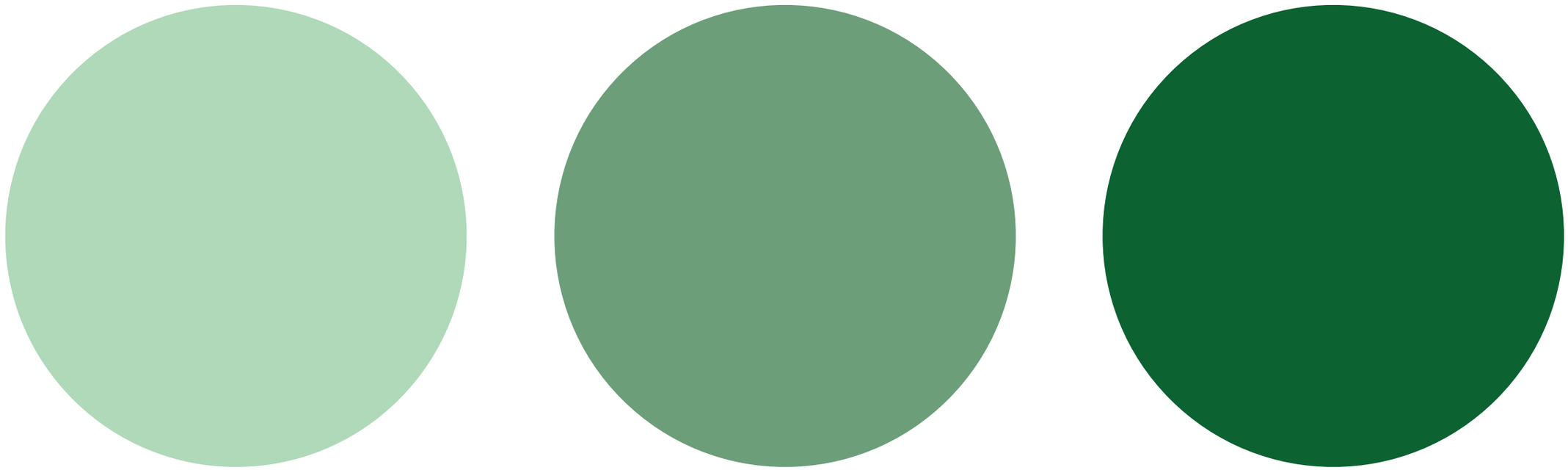


Which represents the larger value?

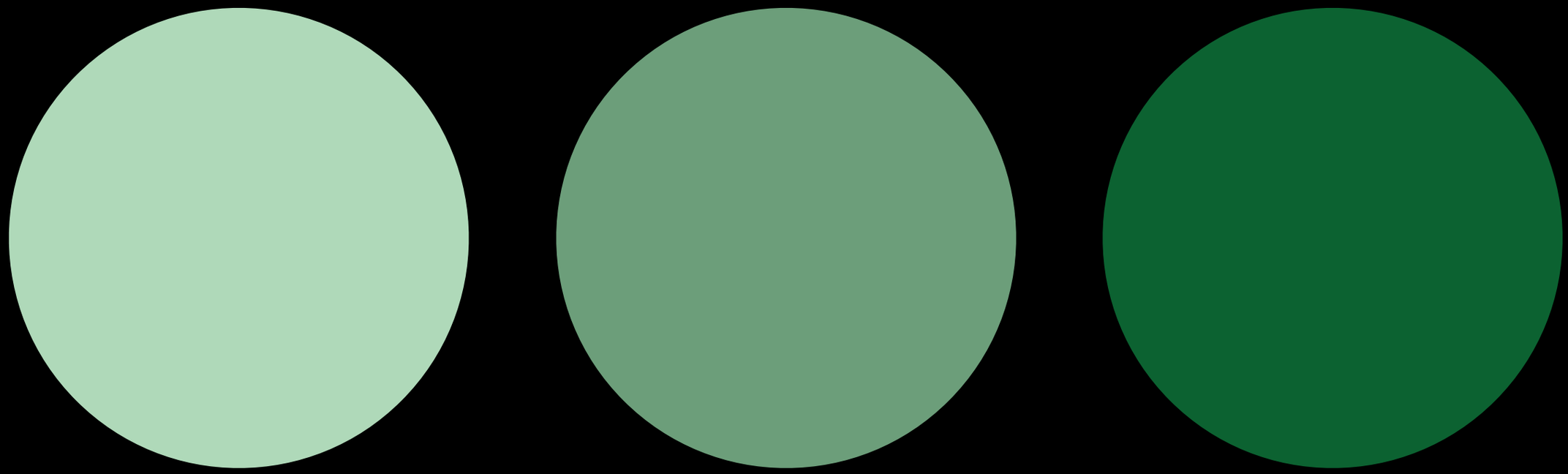
Which represents the larger value?



Which represents the larger value?



Which represents the larger value?



What are the three components of colour?

RGB

r

+50/255 R

+100/255 R

g

+50/255 G

+100/255 G

b

+50/255 B

+100/255 B

~~RGB~~

HSV HSL

HCL

(aka polar LUV)


http://en.wikipedia.org/wiki/HSV_color_space#Disadvantages


hcl


hsv


hsl

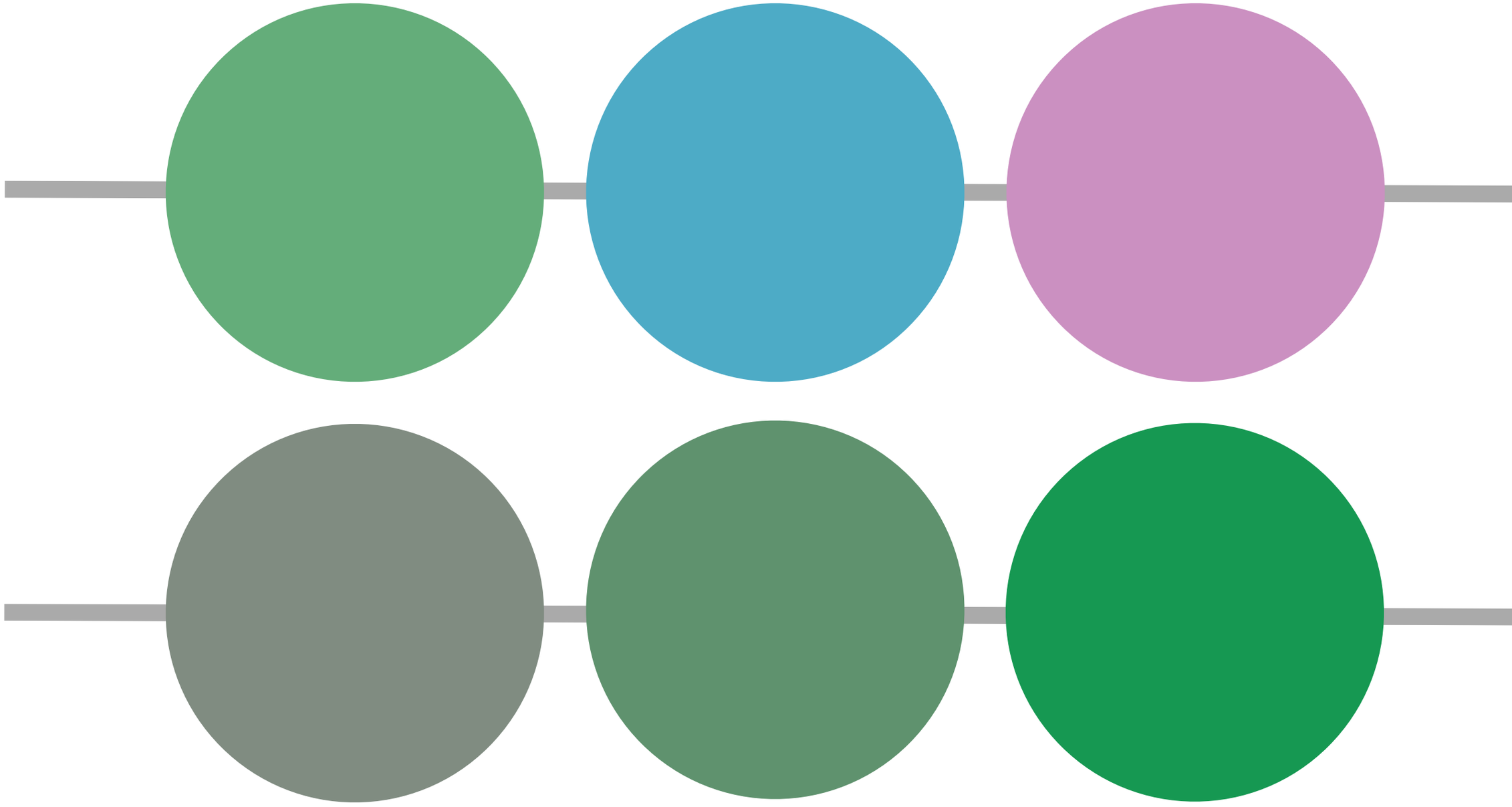
~~RGB~~

~~HSV~~ ~~HSL~~

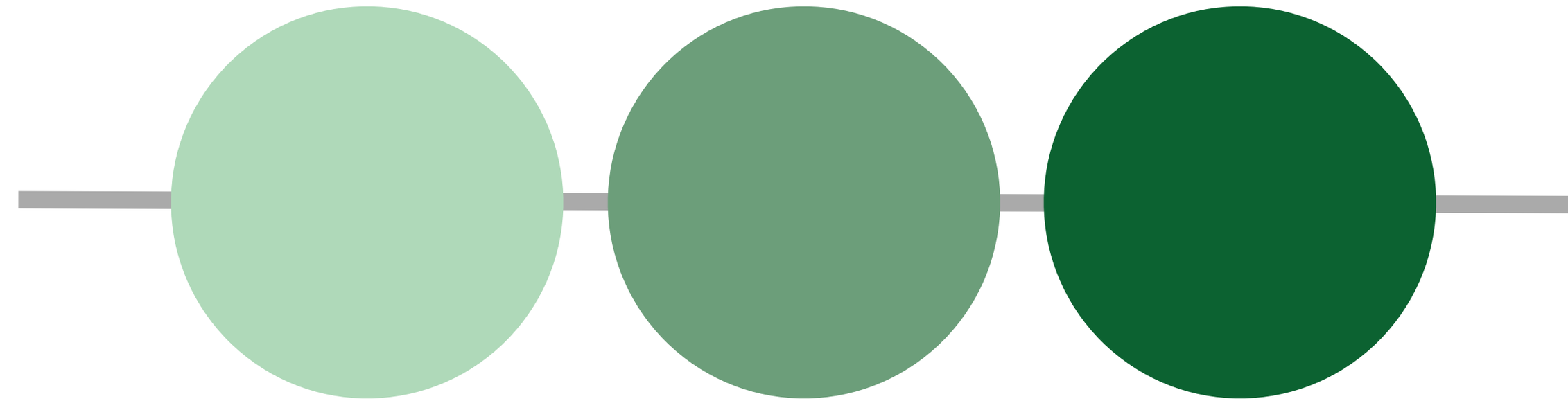
HCL

(aka polar LUV)

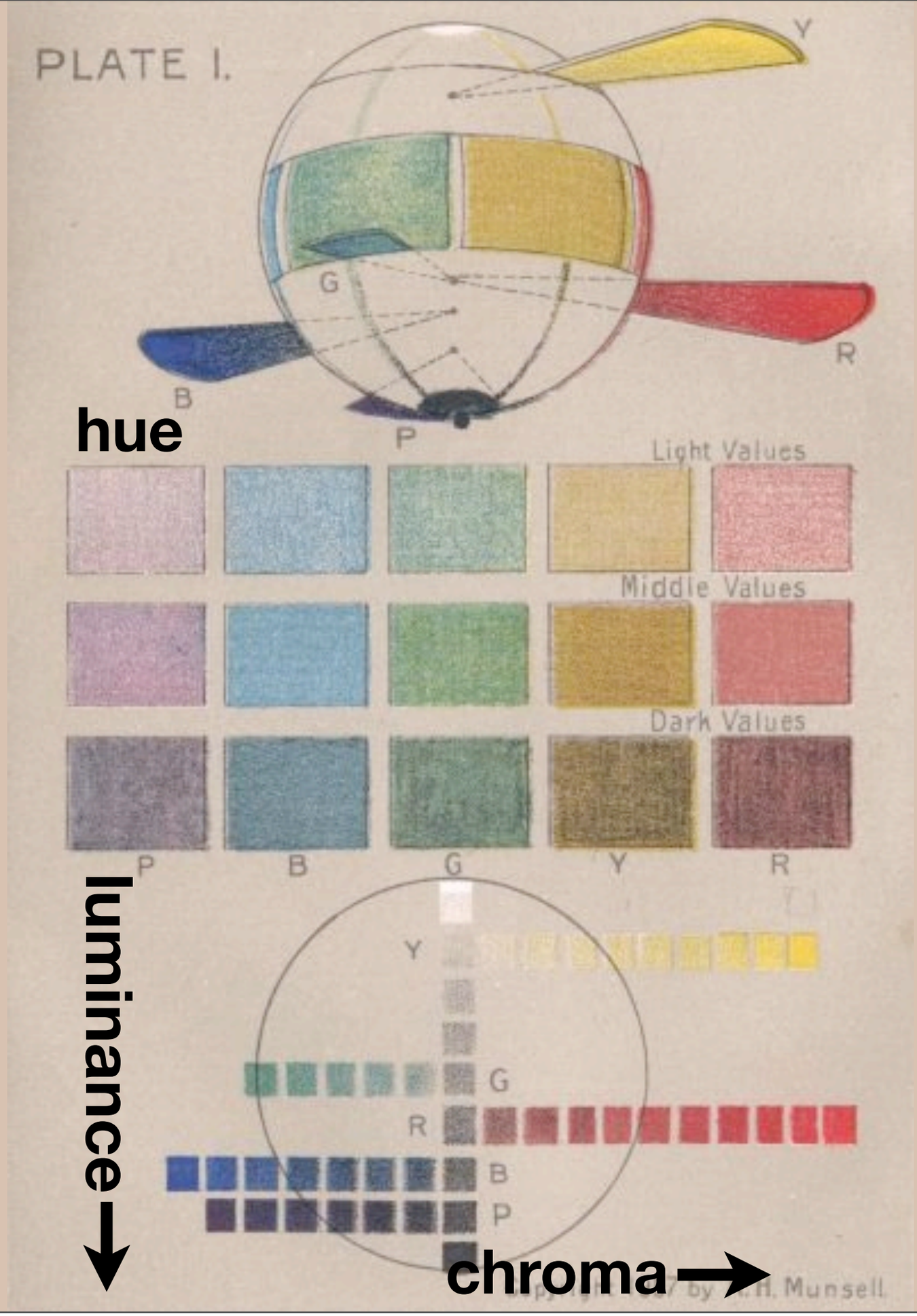
h



C



I



Munsell, Albert H. (1905). *A Color Notation.* Boston: G. H. Ellis Co.

Why care?

Perceptually uniform

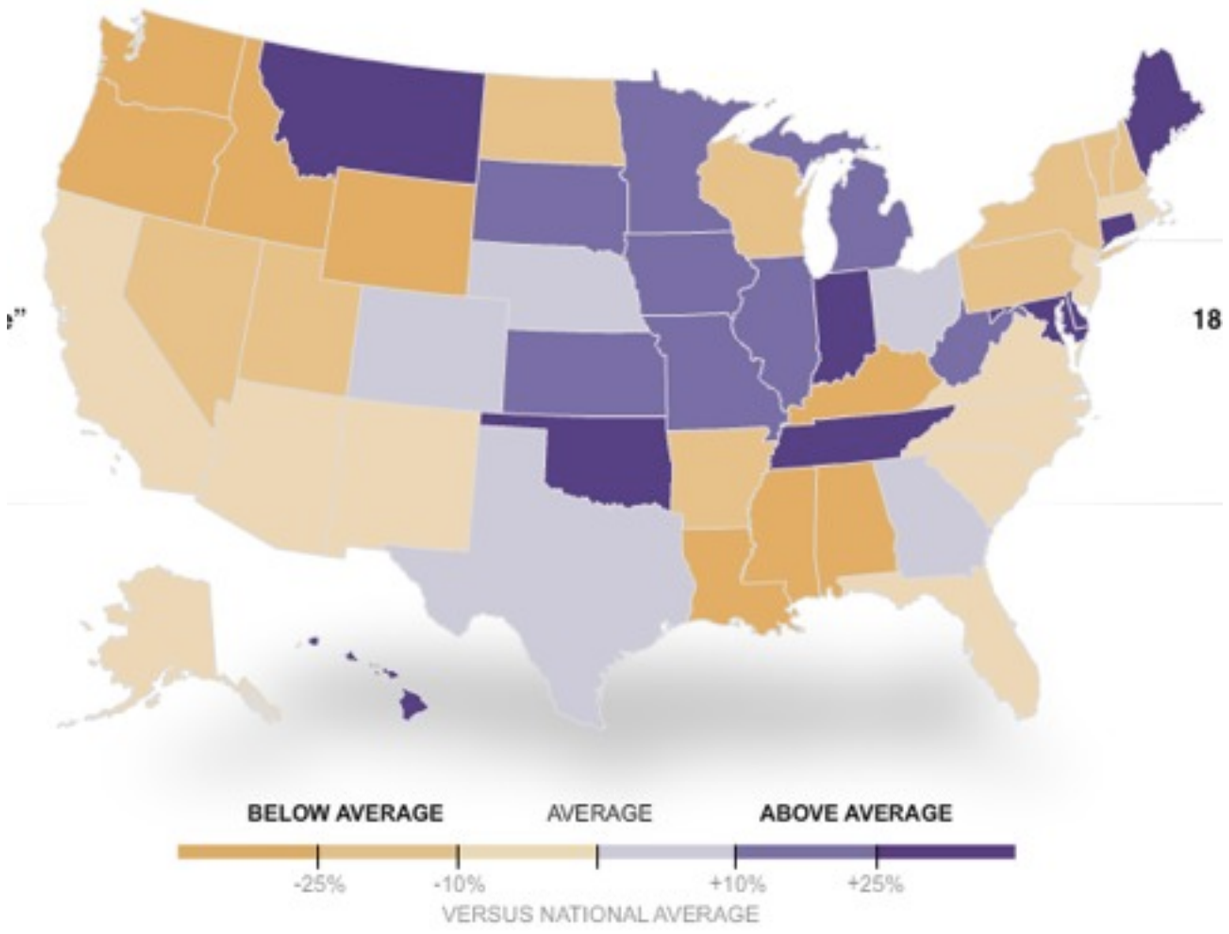
Hue is unordered. Use evenly spaced hues with equal chroma and luminance to make aesthetically pleasing discrete palettes.

Chroma and **luminance** are ordered. Easy to make perceptually uniform gradients by varying either (or both). Never use rainbow scales again!

Aesthetic	Topology
Position	Ordered
Size	Ordered
Luminance	Ordered
Chroma	Ordered
Shape	Unordered
Hue	Unordered

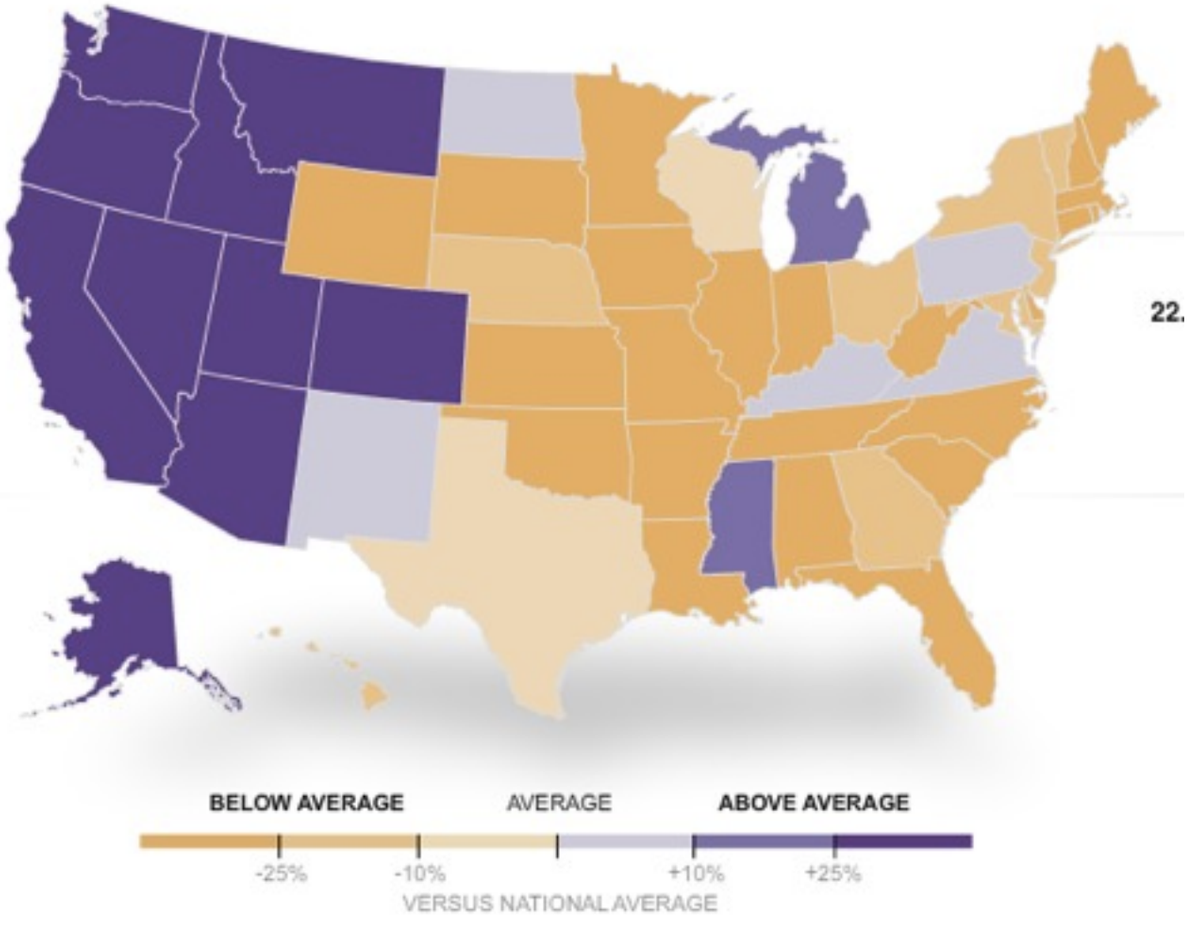
RANK SEARCH TERM PER 10,000 SEARCHES

17. “sweet potato” 47



RANK SEARCH TERM PER 10,000 SEARCHES

21. “yams” 40



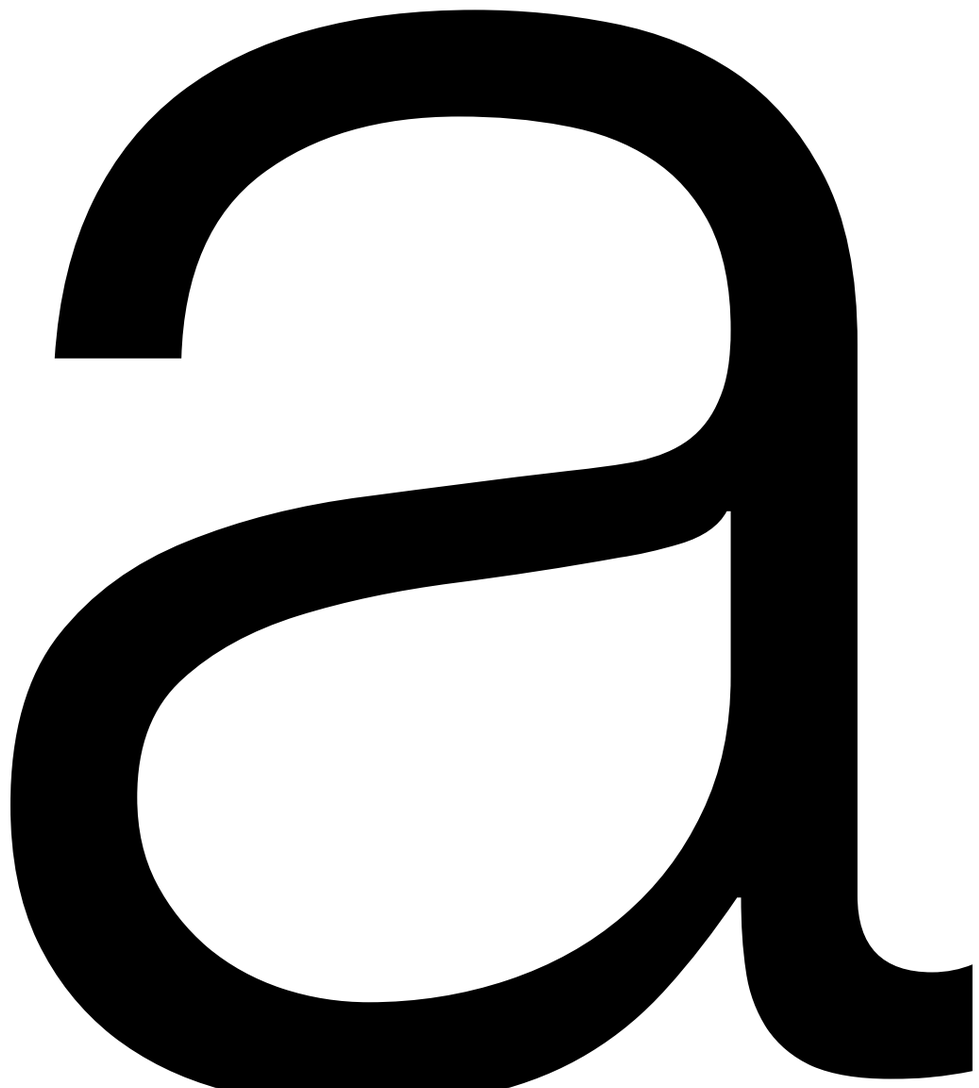
http://www.nytimes.com/interactive/2009/11/26/us/20091126-search-graphic.html

Your turn

In small groups, work through each of the three graphics. Does the data topology match the perceptual topology?

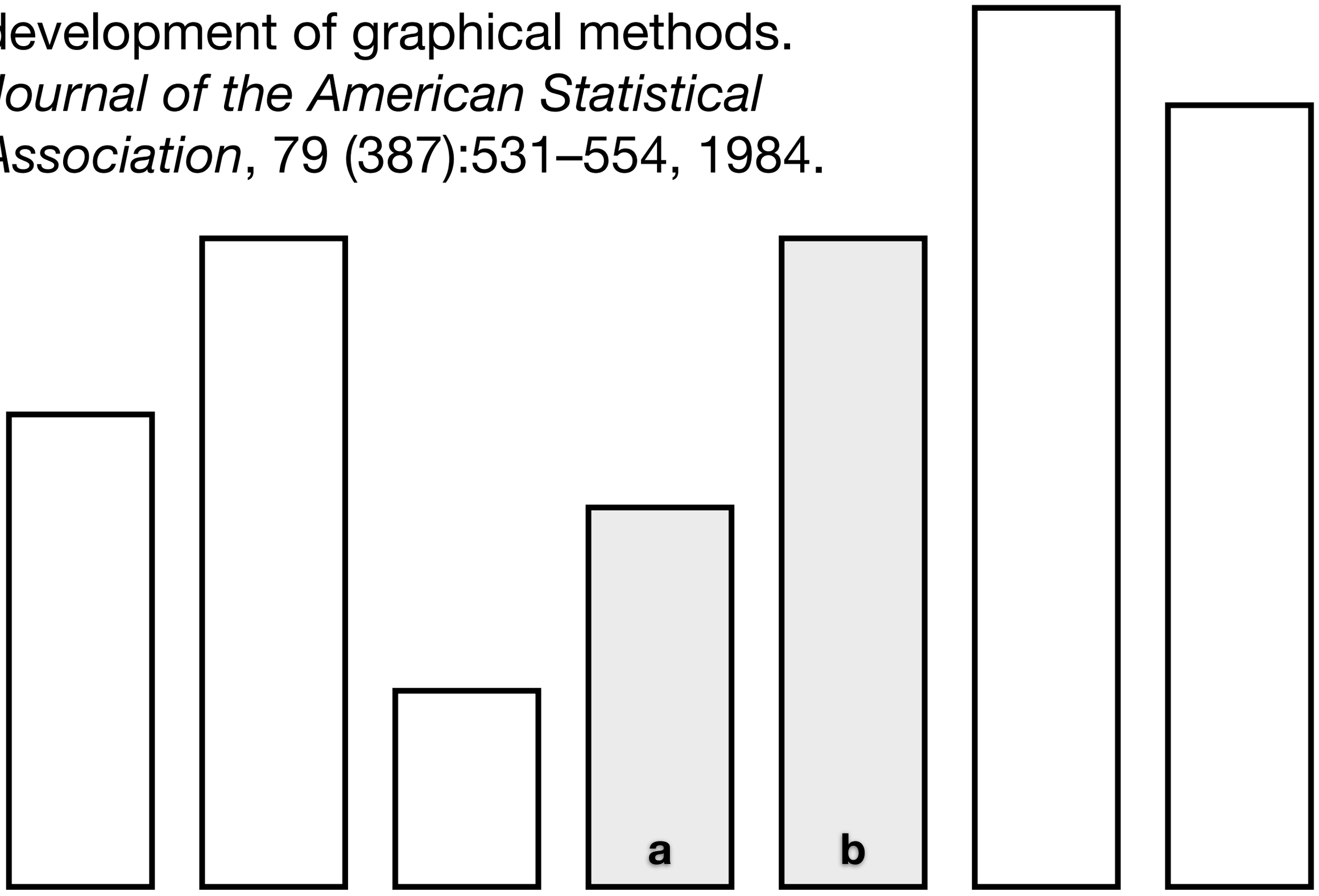
2

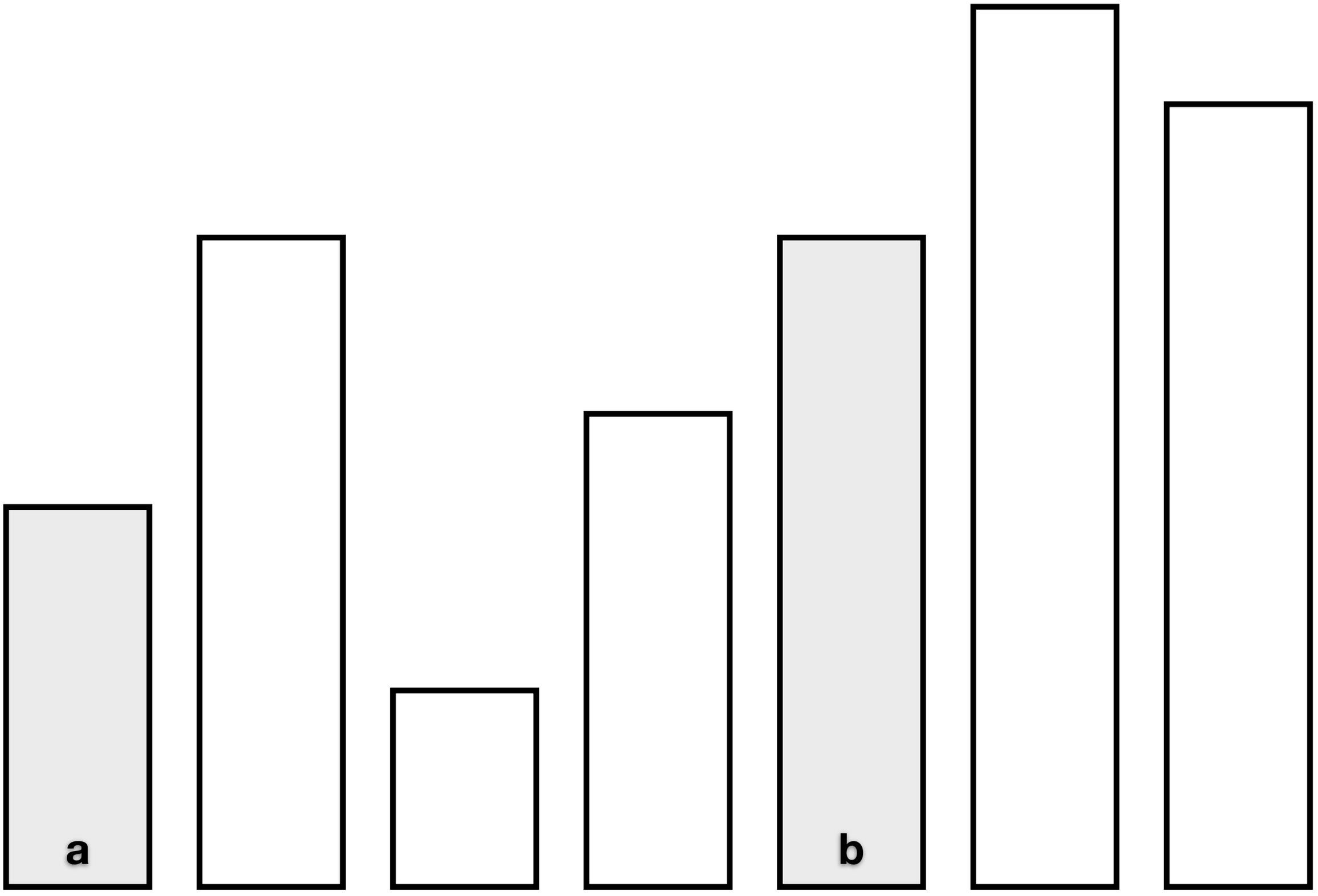
Make important comparisons easy

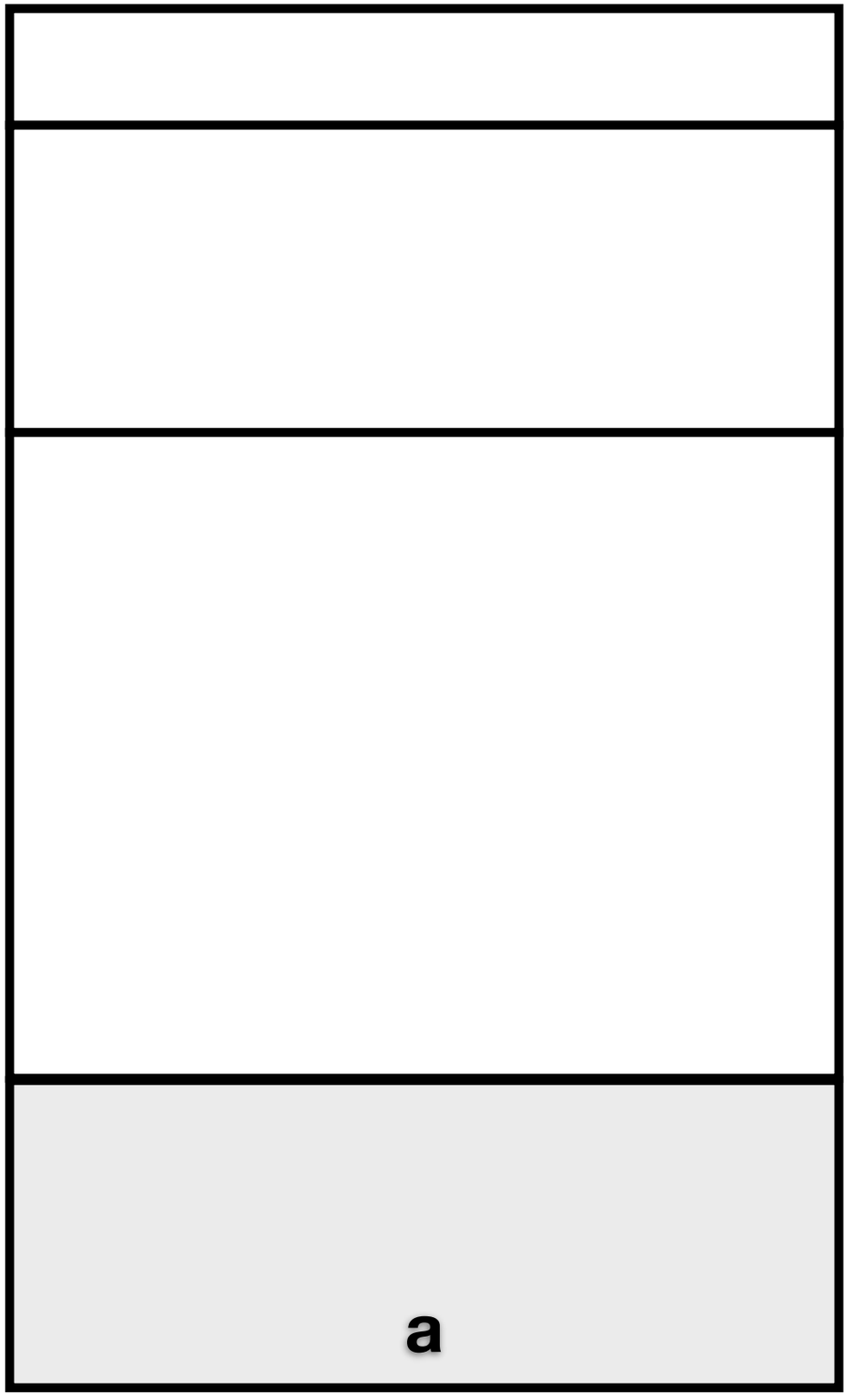


Some comparisons
are easier than others

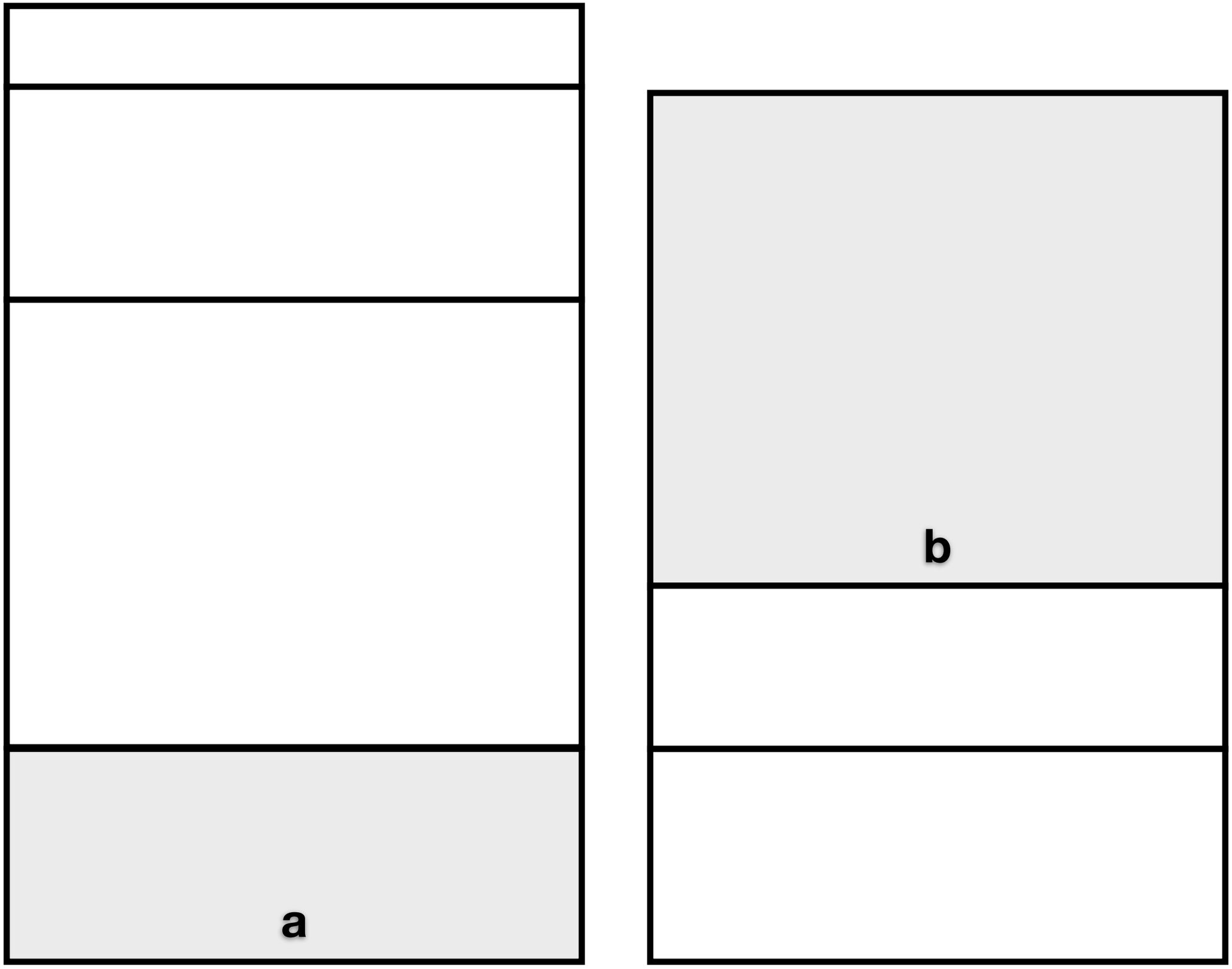
W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, 79 (387):531–554, 1984.




a
b


a

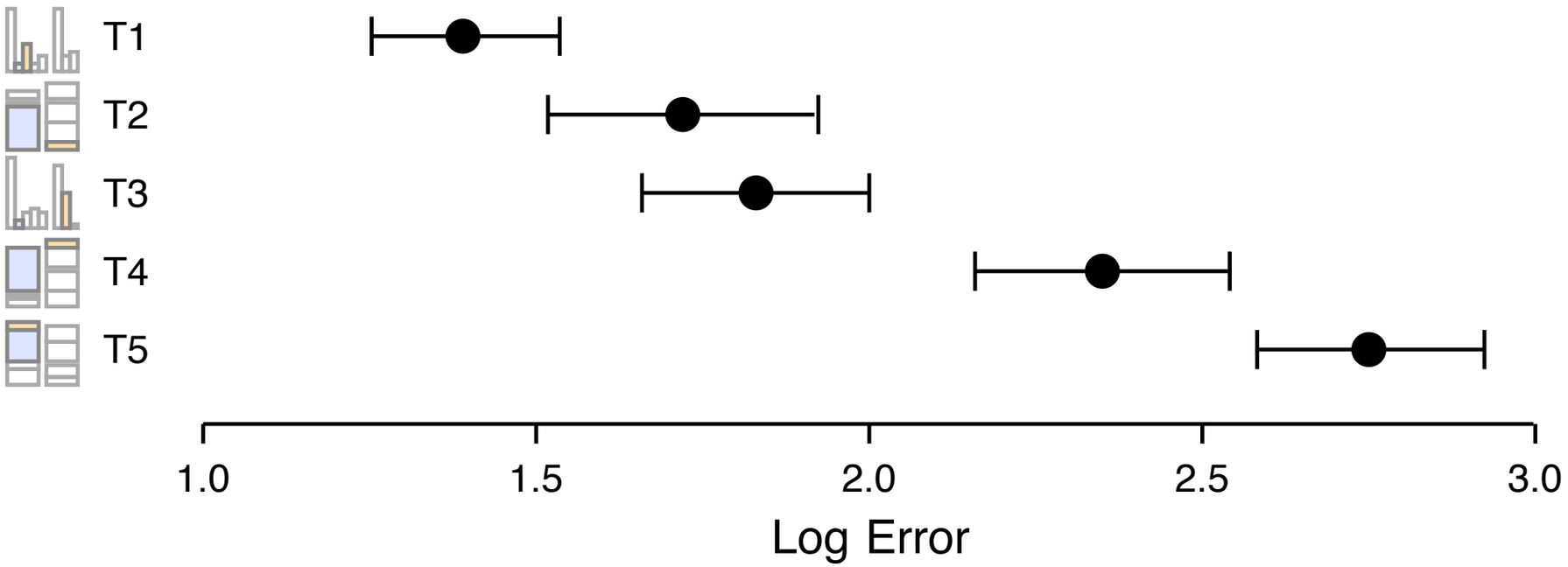

b


a
b

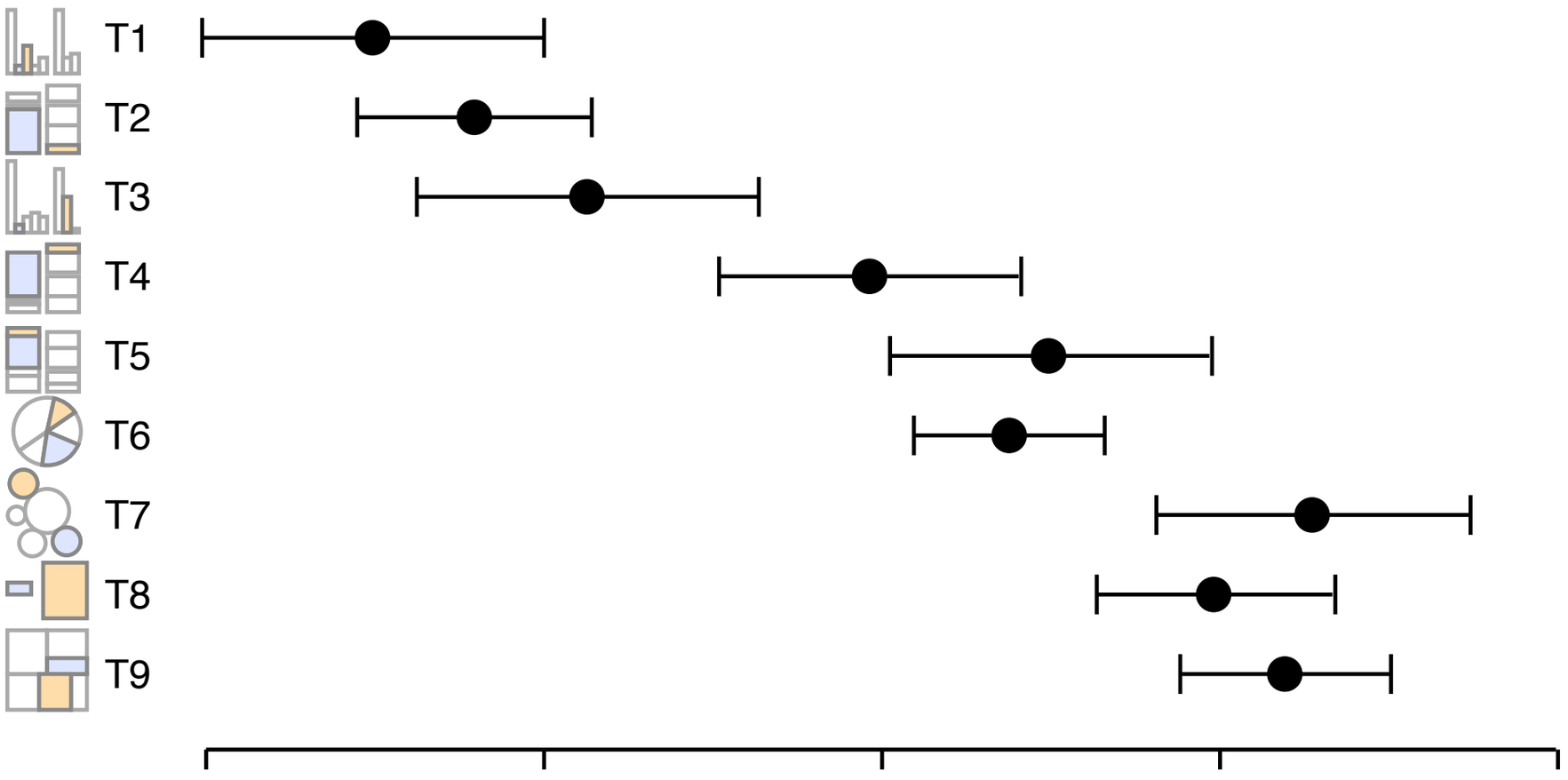
b

a

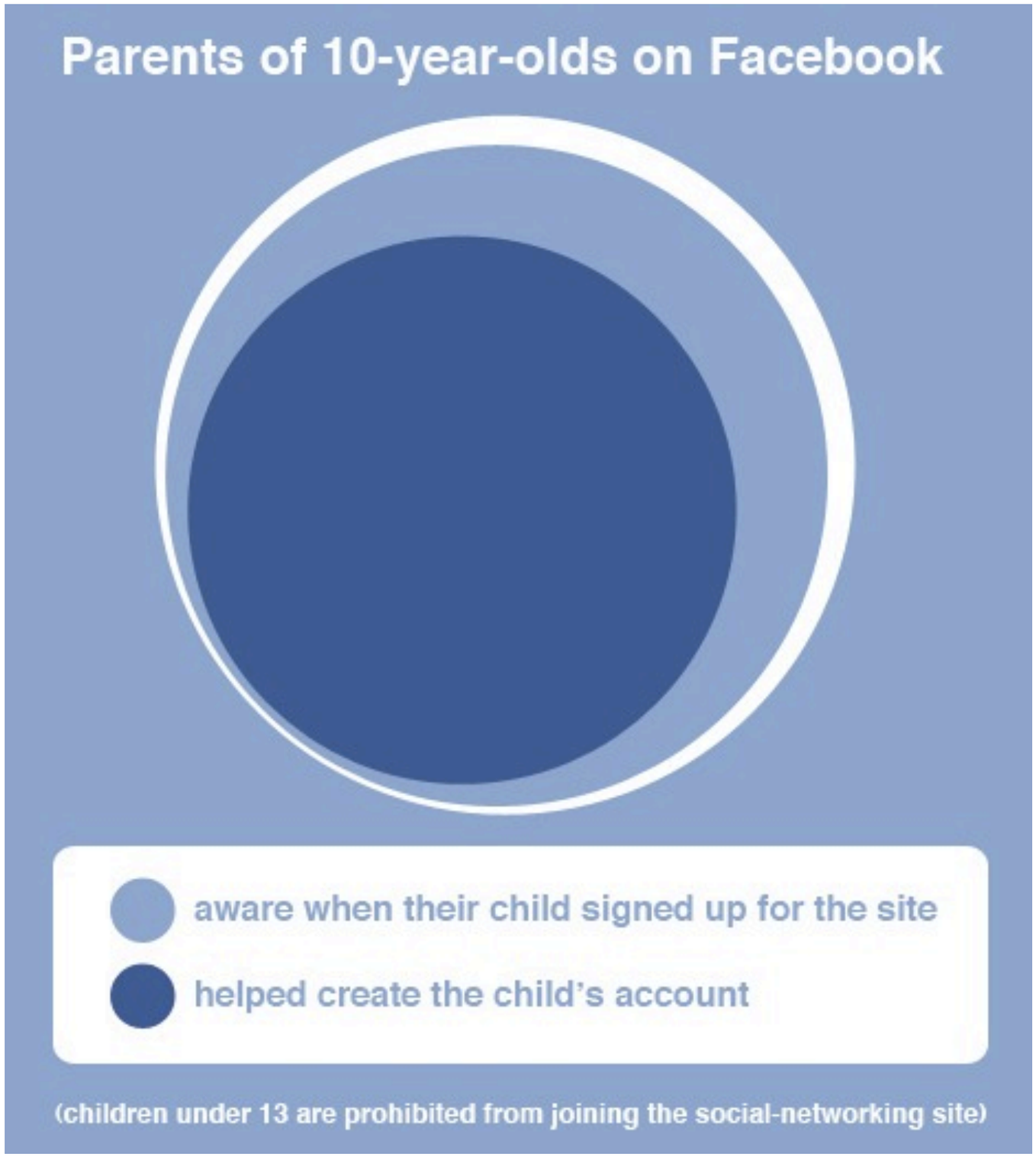
Cleveland & McGill's Results



Crowdsourced Results



J. Heer and M. Bostock. Crowdsourcing graphical perception: Using mechanical turk to assess visualization design. In *CHI 2010*, 2010.



http://www.dreamsystemsmedia.com/blog/index.php/social-media-statistics-of-the-day/

Common misunderstanding

Pie charts are bad! **Die pie chart, DIE**

Common misunderstanding

Pie charts are bad! **Die pie chart, DIE**

Pie charts are bad when you want to accurately compare two numbers

Common misunderstanding

Pie charts are bad! **Die pie chart, DIE**

Pie charts are bad when you want to accurately compare two numbers

But:

As good as bars for estimating percentage of whole.

Better than bars for comparing compound proportions (A + B vs C + D)

Common misunderstanding

Pie charts are bad! **Die pie chart, DIE**

Pie charts are bad when you want to accurately compare two numbers

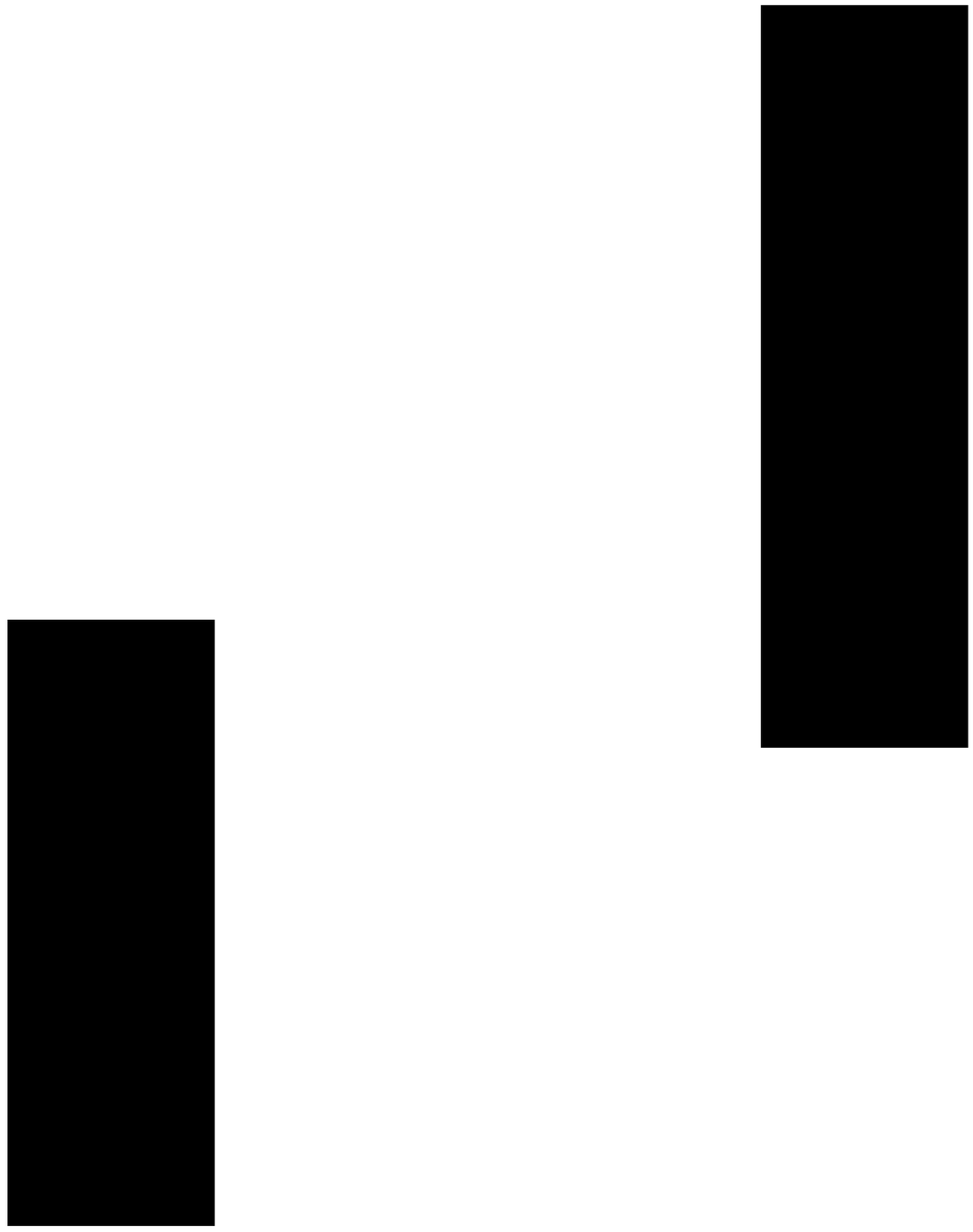
But:

As good as bars for estimating percentage of whole.
Better than bars for comparing compound proportions (A + B vs C + D)

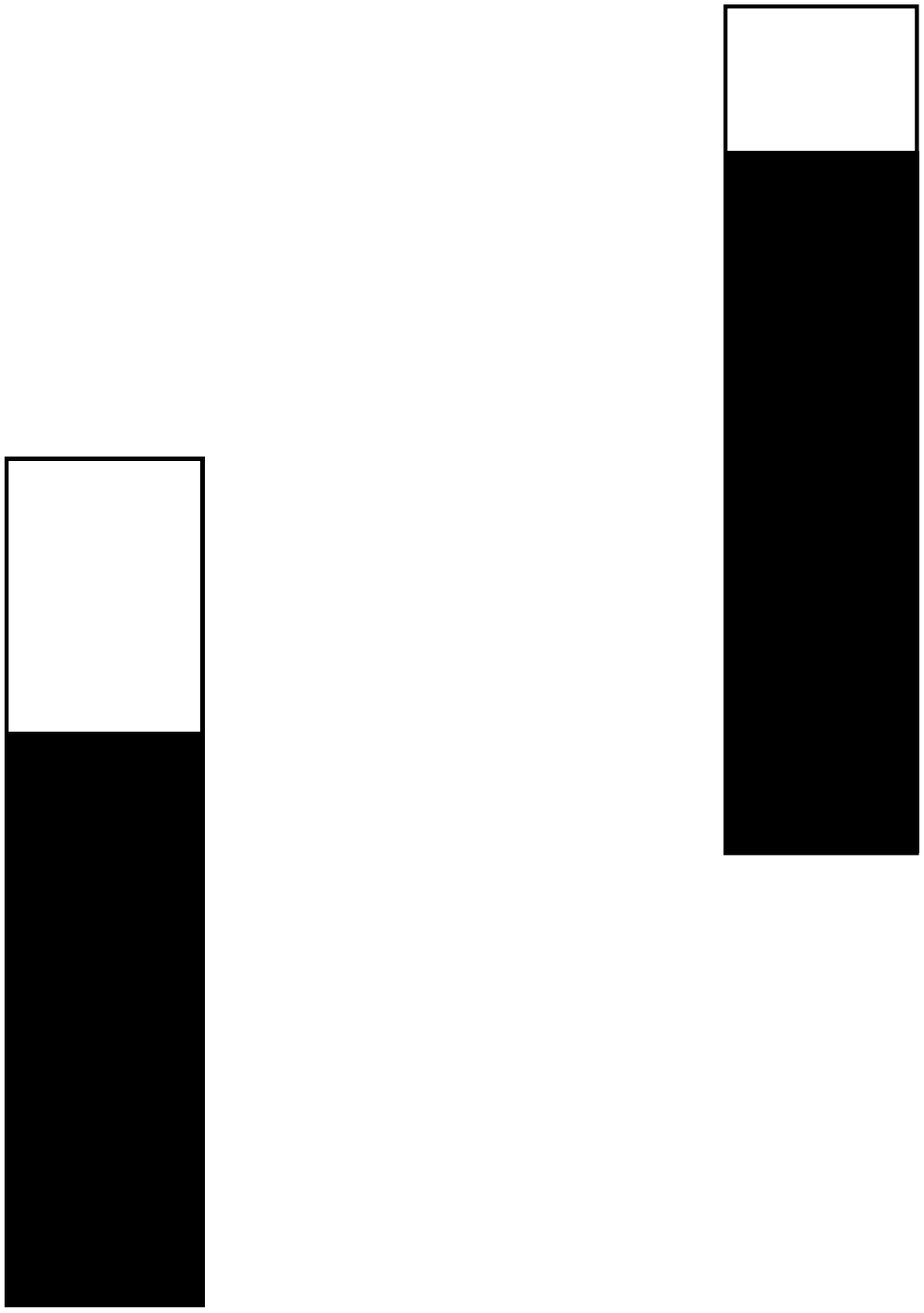
I. Spence. No Humble Pie: The Origins and Usage of a Statistical Chart. *Journal of Educational and Behavioral Statistics*, 30:353–368, 2005.

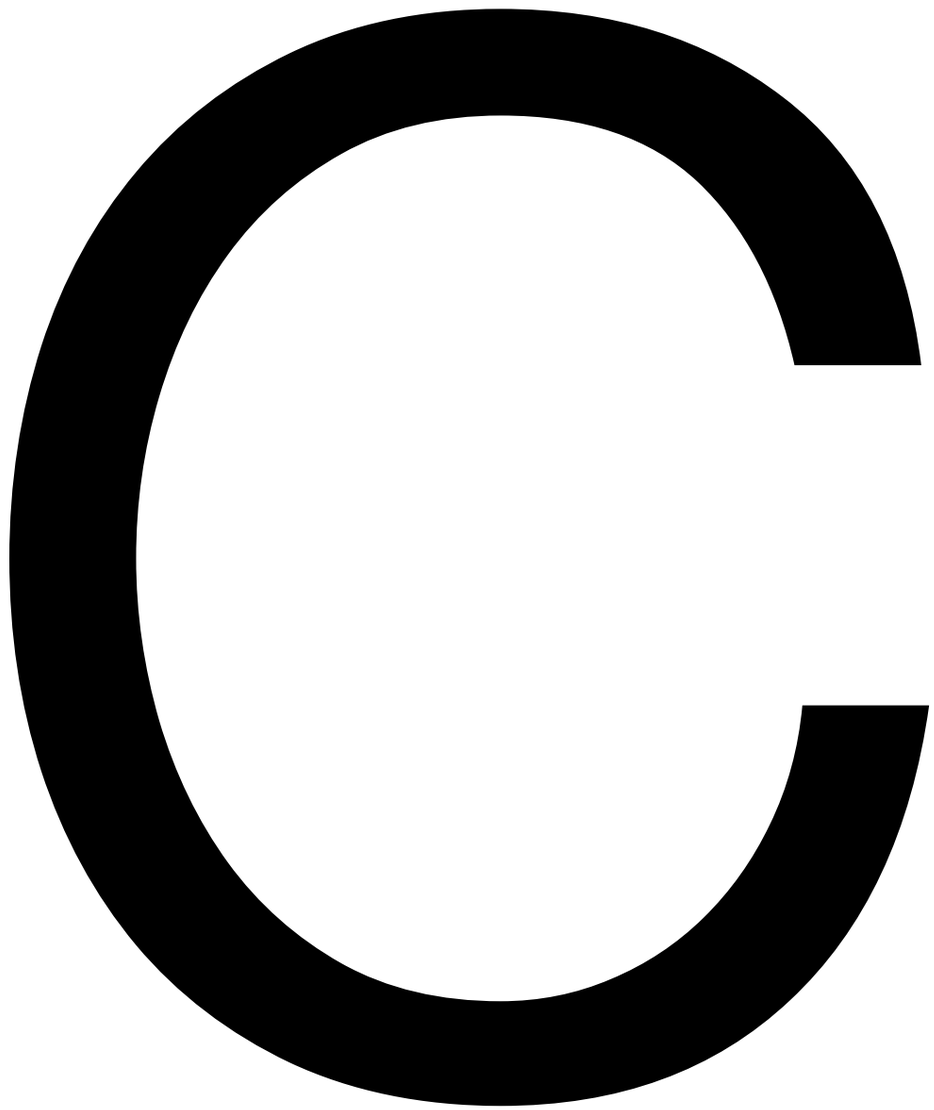
lo

We perceive
relative differences

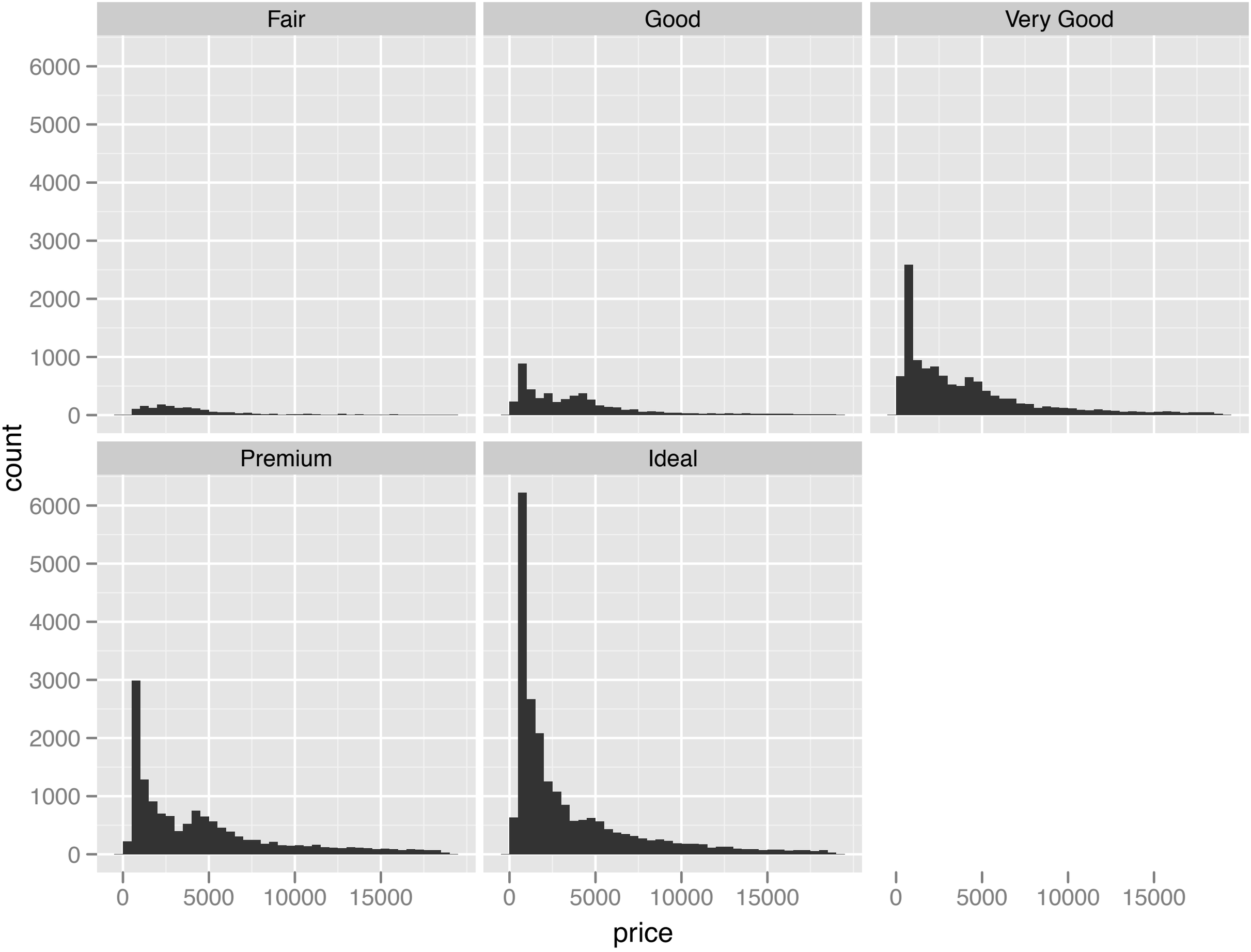


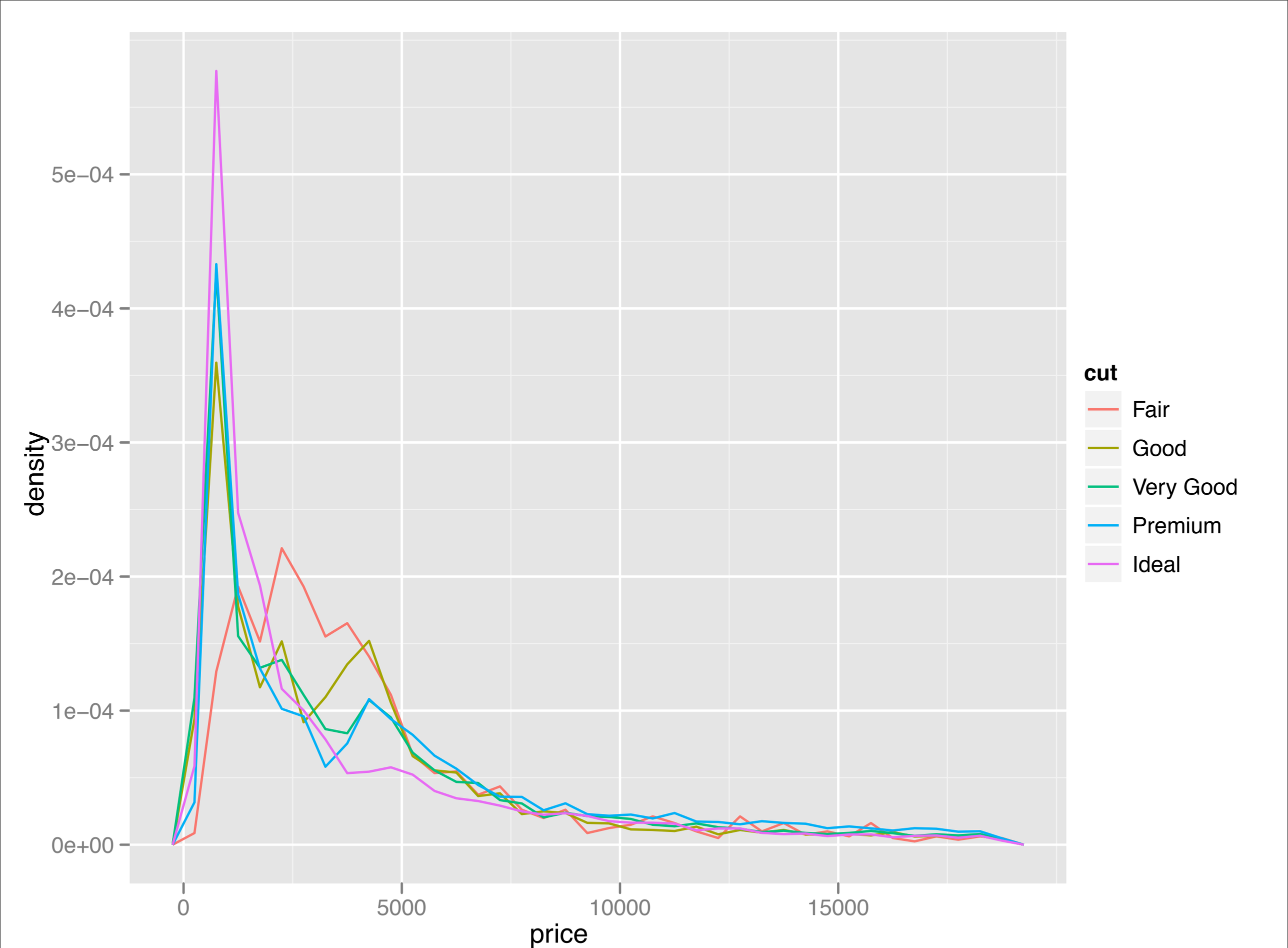
We perceive relative differences





It’s easier to make nearby comparisons


Fair
Good
Very Good
Premium
Ideal
count
6000
5000
4000
3000
2000
1000
0
0
5000
10000
15000
price


density
5e−04
4e−04
3e−04
2e−04
1e−04
0e+00
0
5000
10000
15000
price
cut
Fair
Good
Very Good
Premium
Ideal

Position

Length / Angle

Area

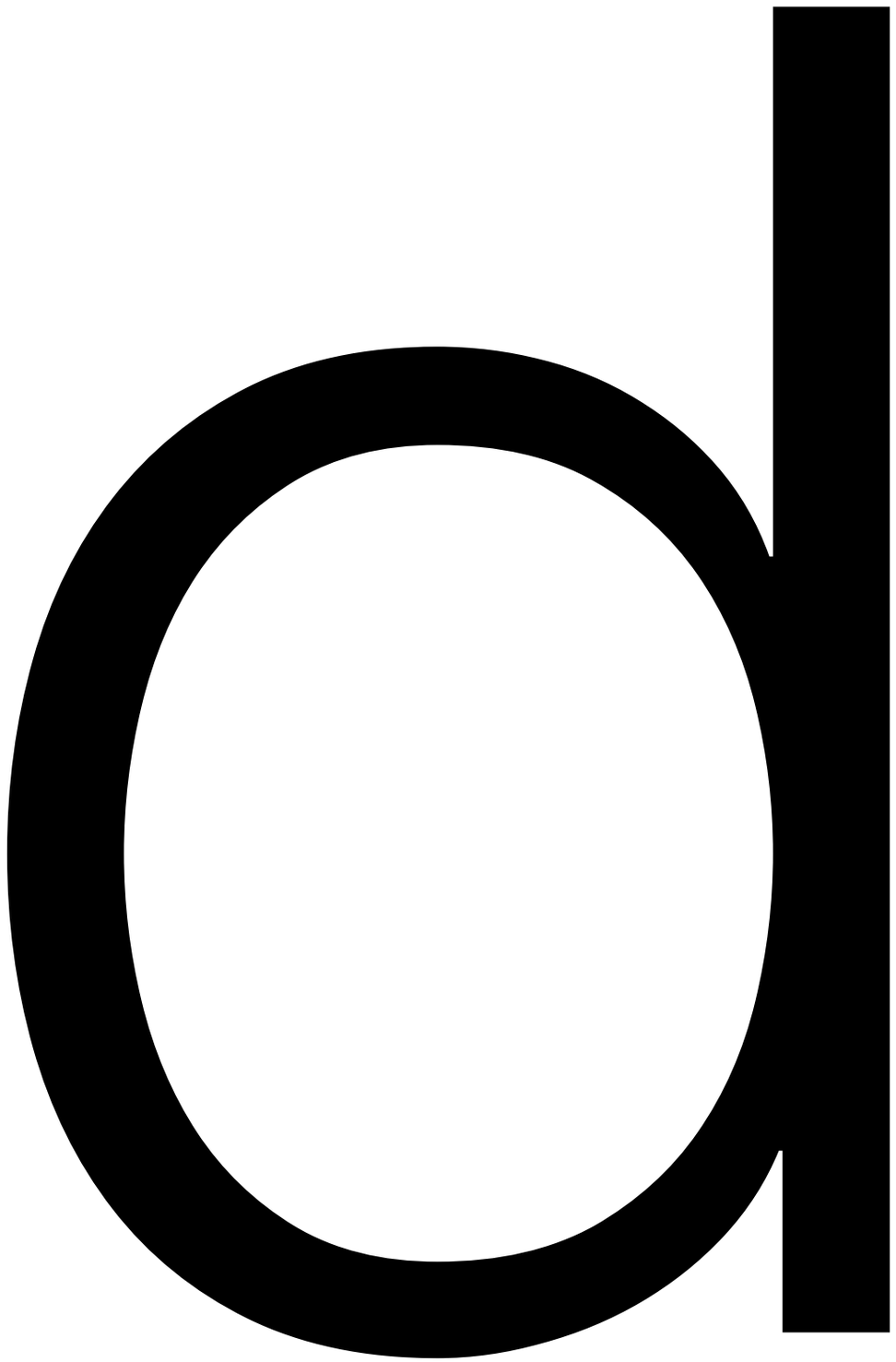
Volume / Chroma / Luminance

×

Perception is relative

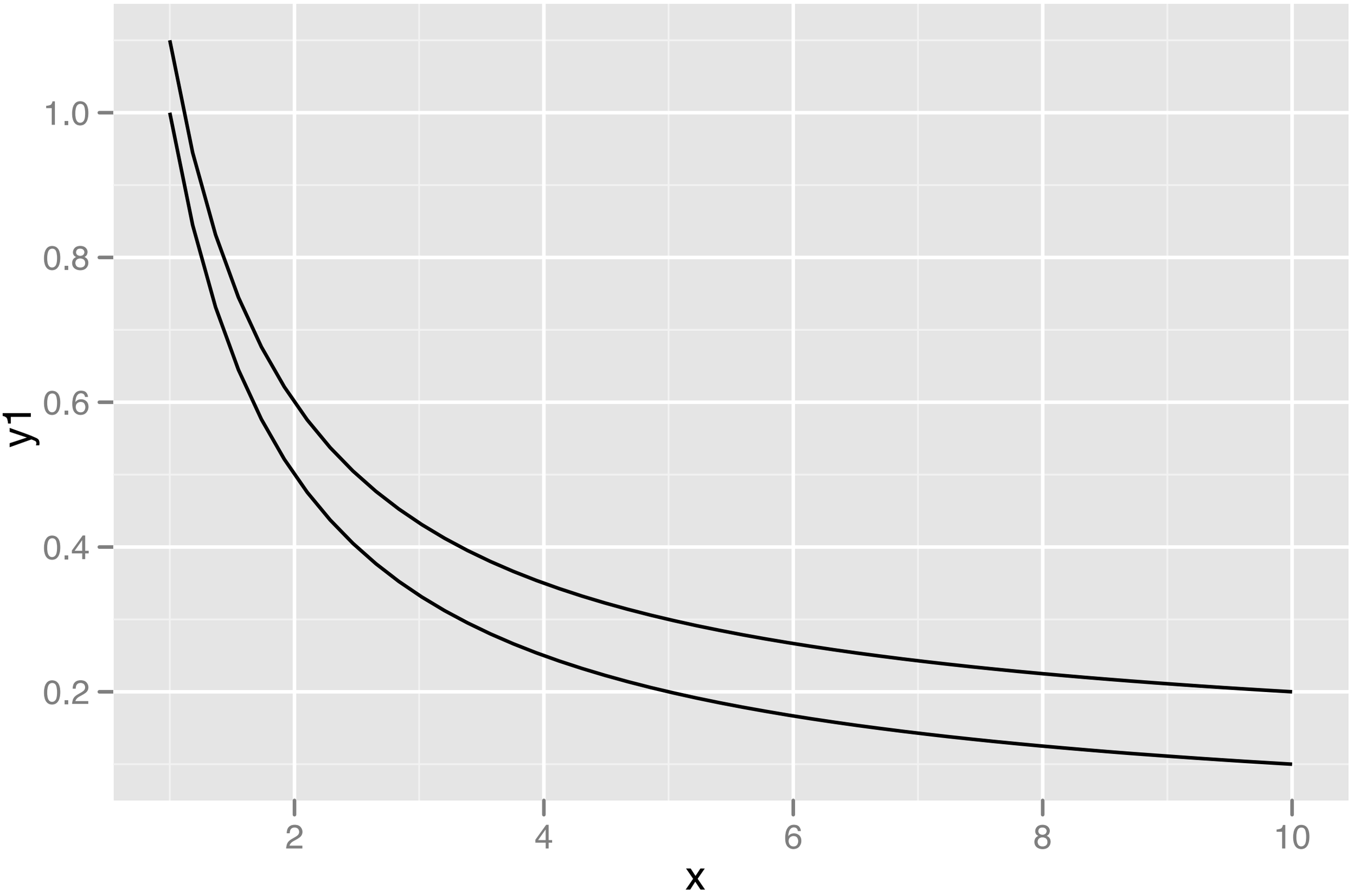
×

Close objects are easier to compare than distant objects



Beware visual illusions

Beware



Beware

Sine Illusion

www.michaelbach.de/ot/sze_sinelllusion/index.html

next→

←prev

Sine Illusion

From Michael's "Visual Phenomena & Optical Illusions"

What to do & see

On the right there are many small upright bars which are vertically displaced following a sine-wave. The question is: do all bars look alike in height? Many would agree they don't. You can use the control "compensate" on the top right, for me a value of 15–20 is about right, agreed?

As you will already have guessed, they are all identical in height, but don't look like it. If you press the button "amp Ø" on the left, the amplitude of the sine-wave is reduced to zero, revealing the veridical height of the vertical lines.

Comment

The original authors (Day & Stecher, 1991) write in their abstract: "The illusion is explained in terms of a perceptual compromise between the vertical extent and the greater overall dimensions of the section at the turn of the sine-wave figure and is thereby held to be the same in principle as the Müller-Lyer illusion." Hmm... When I have obtained the original article (which is non-trivial for me) I will comment more.

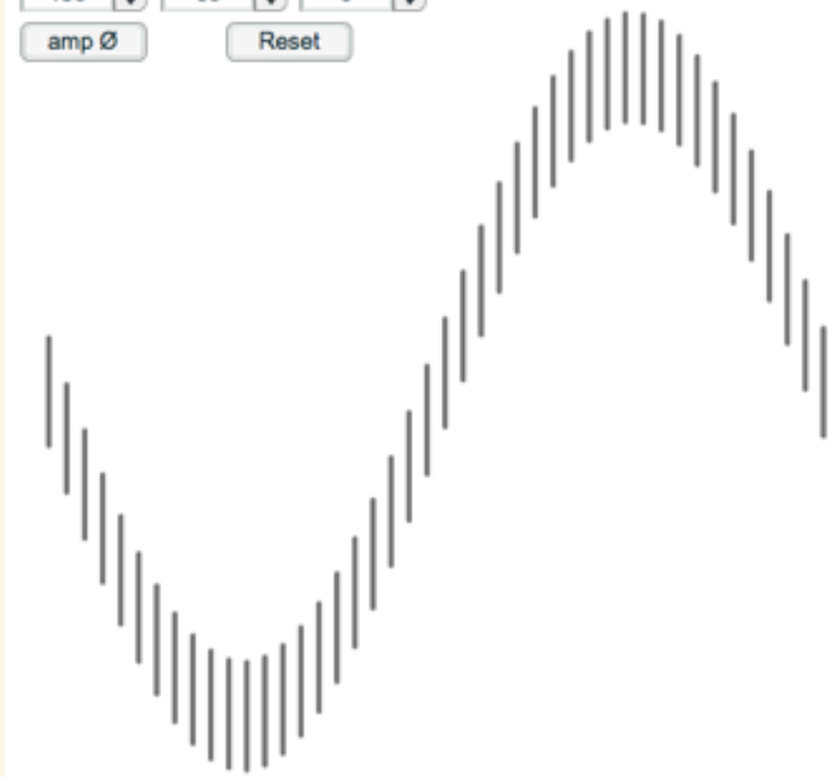
The phenomenon has occurred in my laboratory a number of times, when we were looking at visual evoked responses in the EEG. Small structures that ride on a rising or falling edge are easily misjudged in size.

Sources

Day RH, Stecher EJ (1991) Sine of an illusion. Perception 20;49–55

sine ampl. 180 | line length 30 | compensate 0

amp Ø | Reset



http://www.michaelbach.de/ot/sze_sinelllusion/index.html

Beware

Sine Illusion

www.michaelbach.de/ot/sze_sinelllusion/index.html

next→

←prev

Sine Illusion

From Michael's "Visual Phenomena & Optical Illusions"

What to do & see

On the right there are many small upright bars which are vertically displaced following a sine-wave. The question is: do all bars look alike in height? Many would agree they don't. You can use the control "compensate" on the top right, for me a value of 15–20 is about right, agreed?

As you will already have guessed, they are all identical in height, but don't look like it. If you press the button "amp Ø" on the left, the amplitude of the sine-wave is reduced to zero, revealing the veridical height of the vertical lines.

sine ampl. 0 | line length 30 | compensate 0

amp Ø | Reset

©2007 M.Bach

Comment

The original authors (Day & Stecher, 1991) write in their abstract: "The illusion is explained in terms of a perceptual compromise between the vertical extent and the greater overall dimensions of the section at the turn of the sine-wave figure and is thereby held to be the same in principle as the Müller-Lyer illusion." Hmm... When I have obtained the original article (which is non-trivial for me) I will comment more.

The phenomenon has occurred in my laboratory a number of times, when we were looking at visual evoked responses in the EEG. Small structures that ride on a rising or falling edge are easily misjudged in size.

Sources

Day RH, Stecher EJ (1991) Sine of an illusion. Perception 20;49–55

http://www.michaelbach.de/ot/sze_sinelllusion/index.html

Beware

Sine Illusion

www.michaelbach.de/ot/sze_sinelllusion/index.html

next→

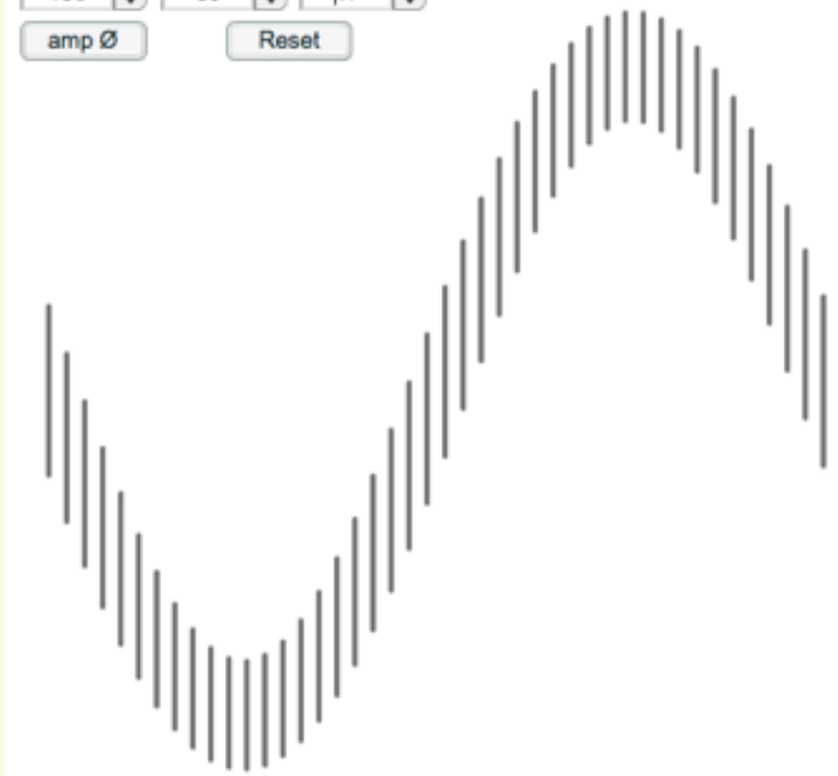
←prev

Sine Illusion

From Michael's "Visual Phenomena & Optical Illusions"

sine ampl. 180 | line length 30 | compensate 17

amp Ø | Reset



©2007 M.Bach

What to do & see

On the right there are many small upright bars which are vertically displaced following a sine-wave. The question is: do all bars look alike in height? Many would agree they don't. You can use the control "compensate" on the top right, for me a value of 15–20 is about right, agreed?

As you will already have guessed, they are all identical in height, but don't look like it. If you press the button "amp Ø" on the left, the amplitude of the sine-wave is reduced to zero, revealing the veridical height of the vertical lines.

Comment

The original authors (Day & Stecher, 1991) write in their abstract: "The illusion is explained in terms of a perceptual compromise between the vertical extent and the greater overall dimensions of the section at the turn of the sine-wave figure and is thereby held to be the same in principle as the Müller-Lyer illusion." Hmm... When I have obtained the original article (which is non-trivial for me) I will comment more.

The phenomenon has occurred in my laboratory a number of times, when we were looking at visual evoked responses in the EEG. Small structures that ride on a rising or falling edge are easily misjudged in size.

Sources

Day RH, Stecher EJ (1991) Sine of an illusion. Perception 20;49–55

http://www.michaelbach.de/ot/sze_sinelllusion/index.html

Beware

Sine Illusion

www.michaelbach.de/ot/sze_sinelllusion/index.html

next→

←prev

Sine Illusion

From Michael's "Visual Phenomena & Optical Illusions"

What to do & see

On the right there are many small upright bars which are vertically displaced following a sine-wave. The question is: do all bars look alike in height? Many would agree they don't. You can use the control "compensate" on the top right, for me a value of 15–20 is about right, agreed?

As you will already have guessed, they are all identical in height, but don't look like it. If you press the button "amp Ø" on the left, the amplitude of the sine-wave is reduced to zero, revealing the veridical height of the vertical lines.

sine ampl. 0 | line length 30 | compensate 17

amp ↑ | Reset

©2007 M.Bach

Comment

The original authors (Day & Stecher, 1991) write in their abstract: "The illusion is explained in terms of a perceptual compromise between the vertical extent and the greater overall dimensions of the section at the turn of the sine-wave figure and is thereby held to be the same in principle as the Müller-Lyer illusion." Hmm... When I have obtained the original article (which is non-trivial for me) I will comment more.

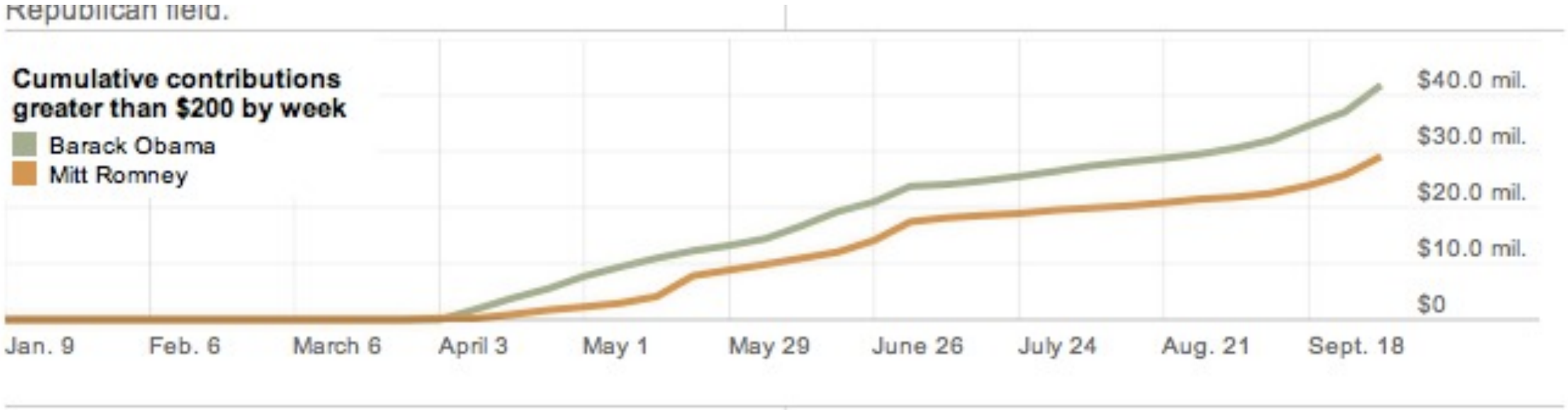
The phenomenon has occurred in my laboratory a number of times, when we were looking at visual evoked responses in the EEG. Small structures that ride on a rising or falling edge are easily misjudged in size.

Sources

Day RH, Stecher EJ (1991) Sine of an illusion. Perception 20;49–55

http://www.michaelbach.de/ot/sze_sinelllusion/index.html

Republican field.



http://elections.nytimes.com/2012/campaign-finance

Use position, then length/area, then chroma/luminance

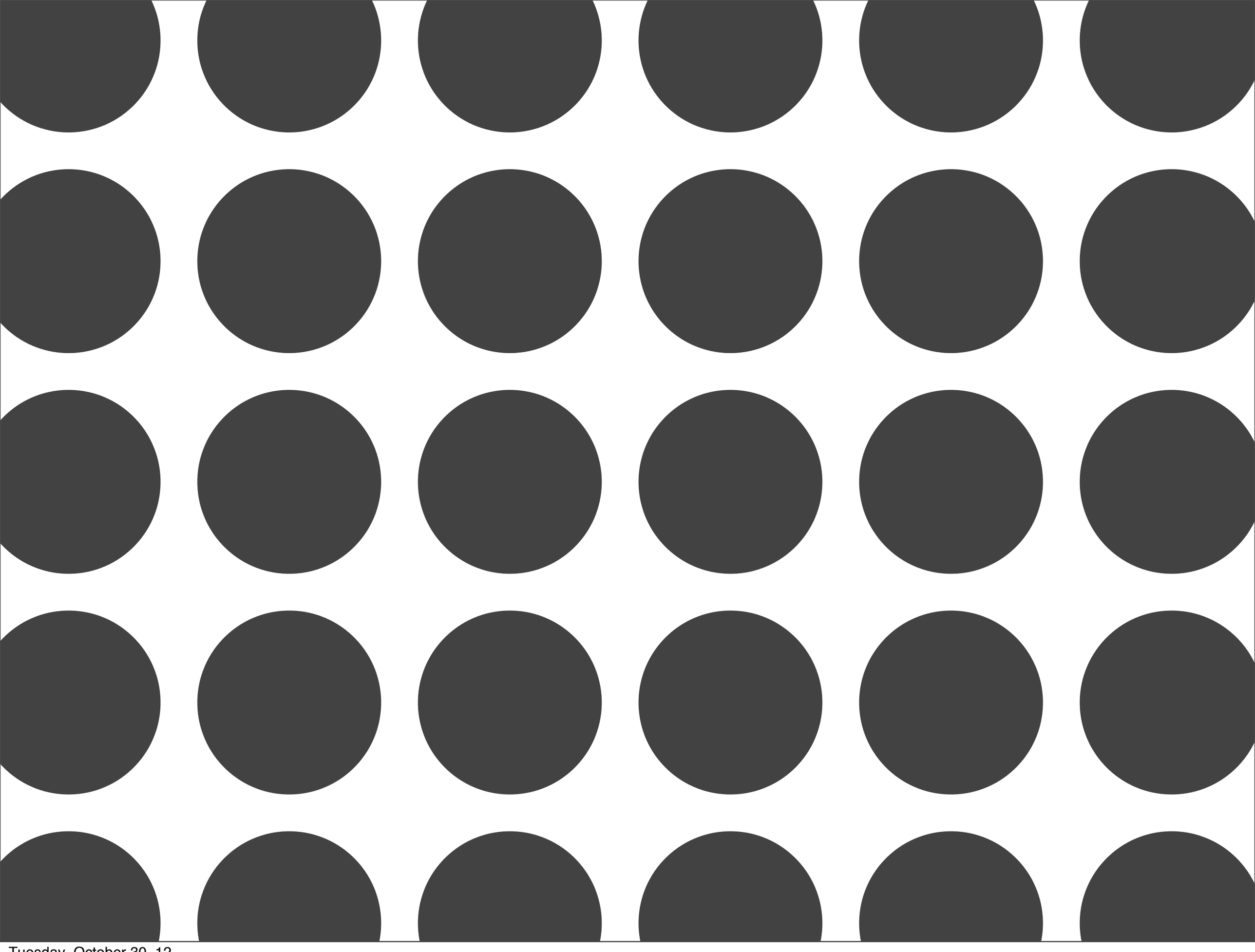
Ensure important comparisons are nearby

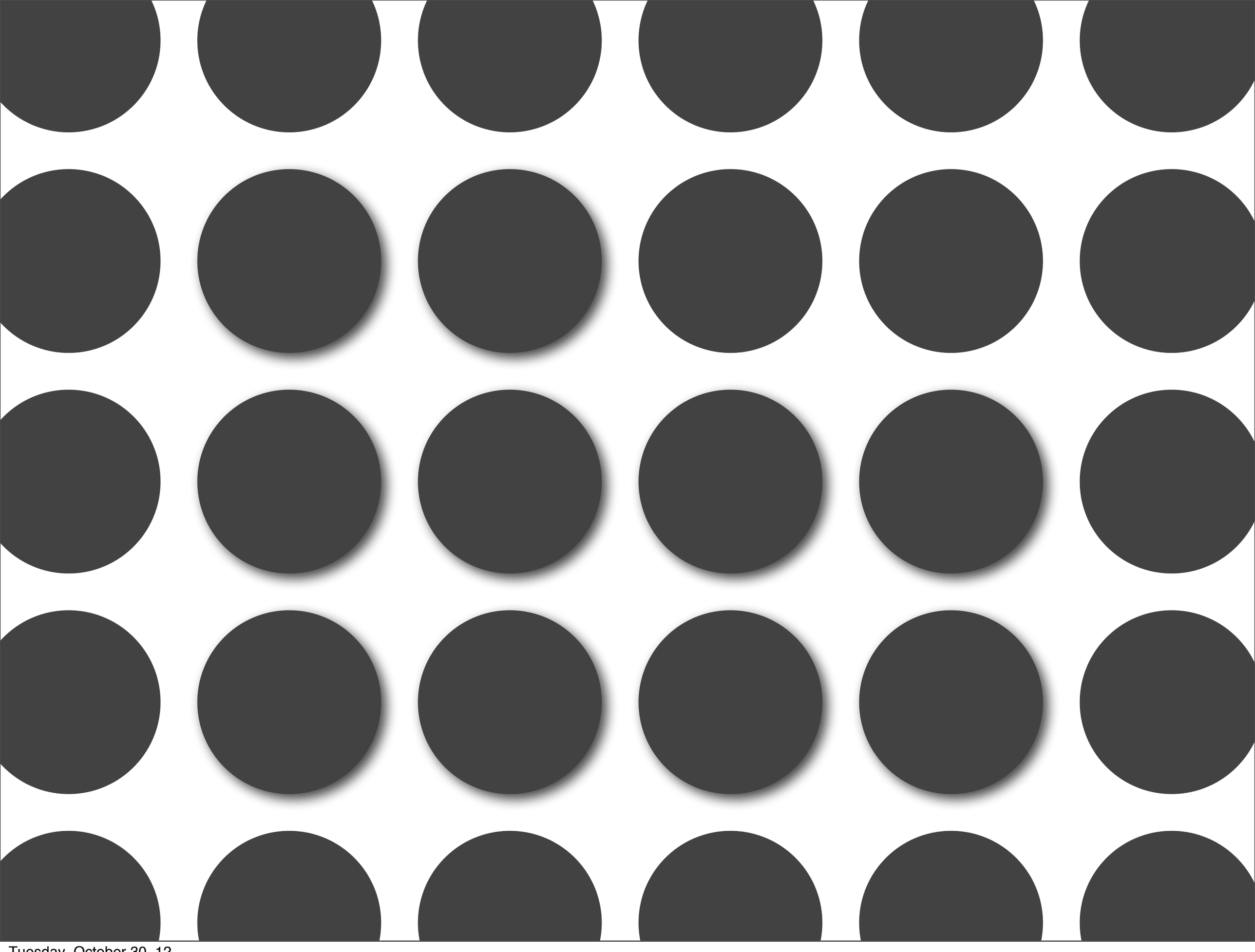
If possible, display comparisons directly

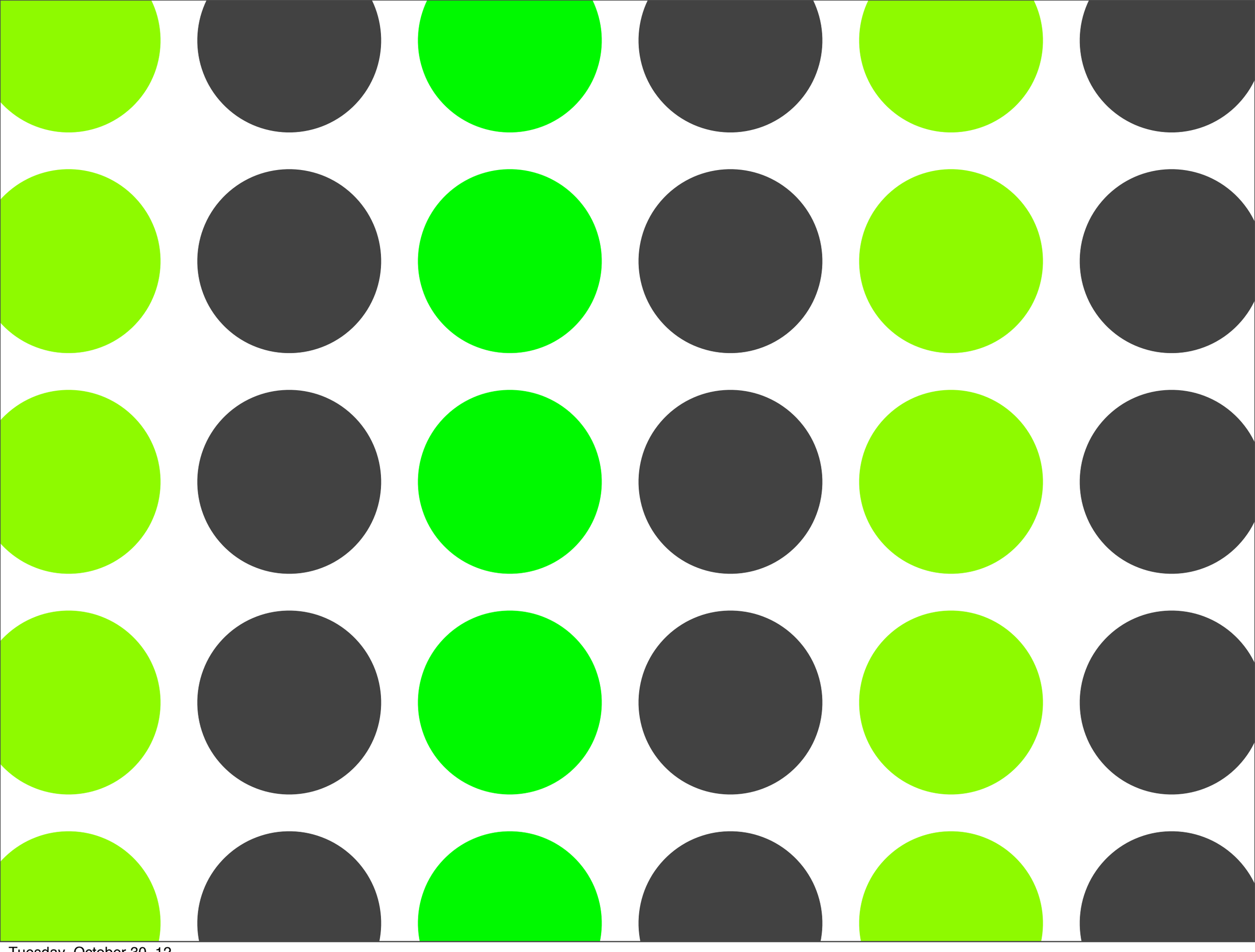
Your turn

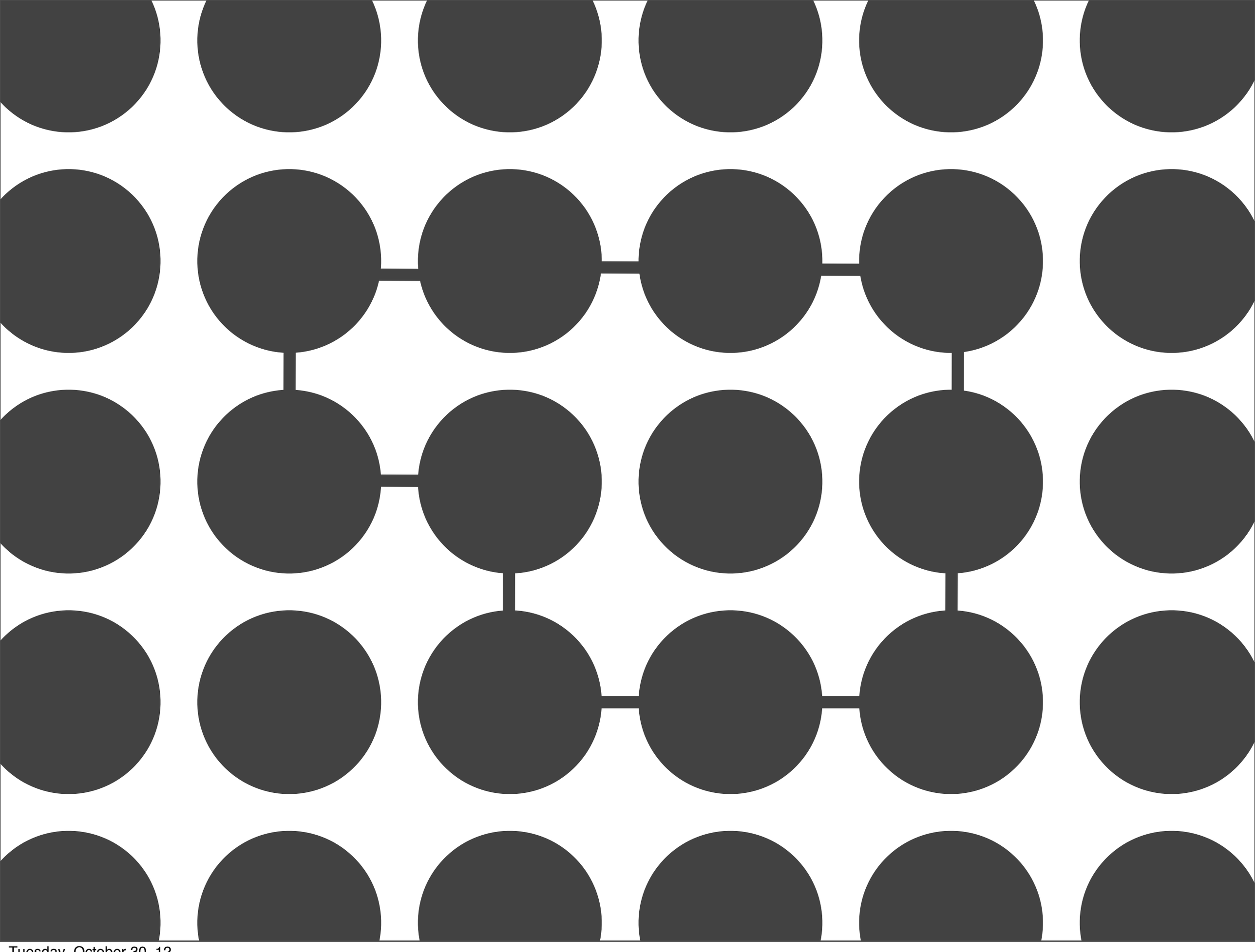
In small groups, work through each of the three graphics. What are the important comparisons? What's easy to do and what's hard to do?

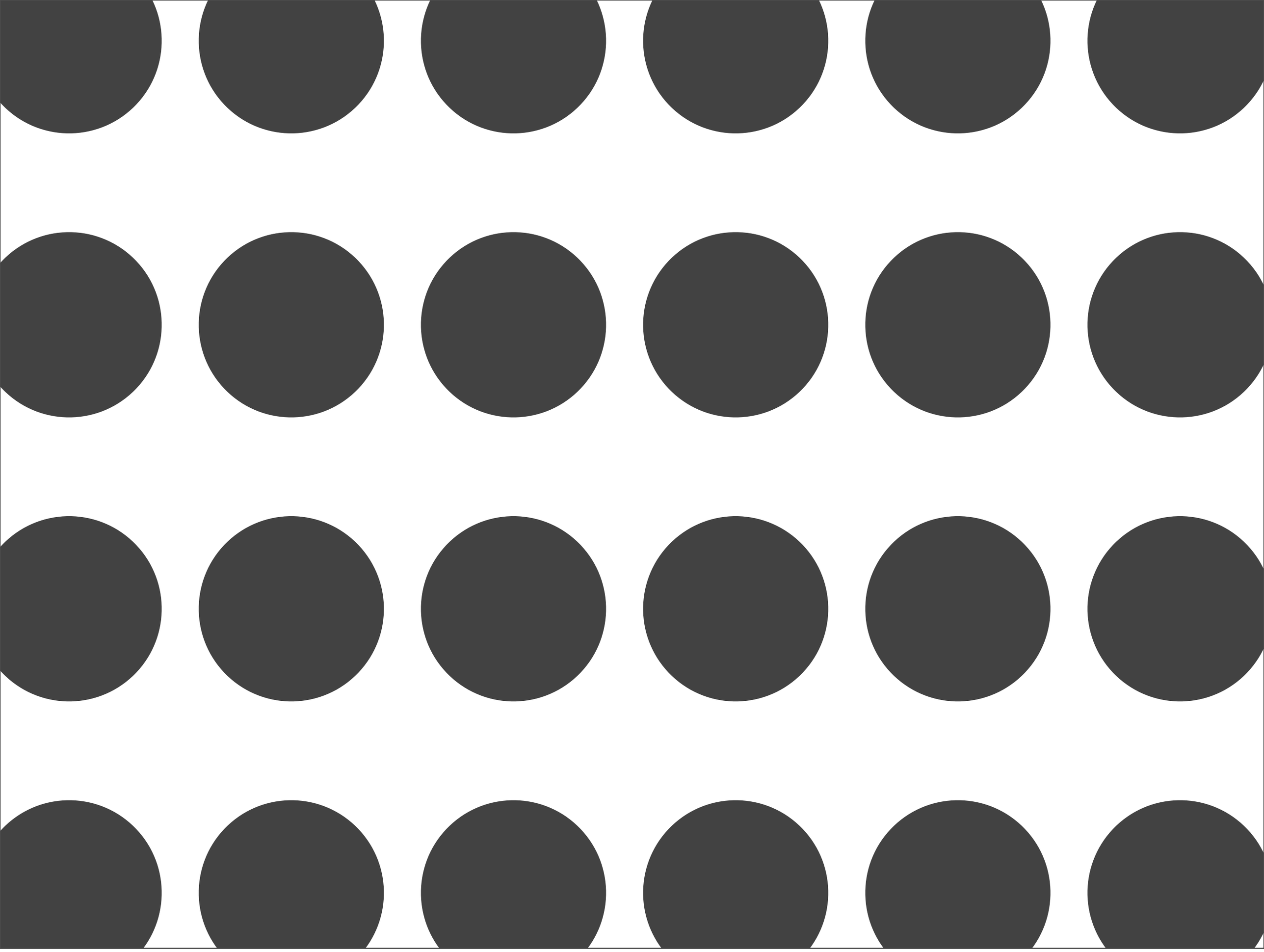
Visual connections should reflect real connections

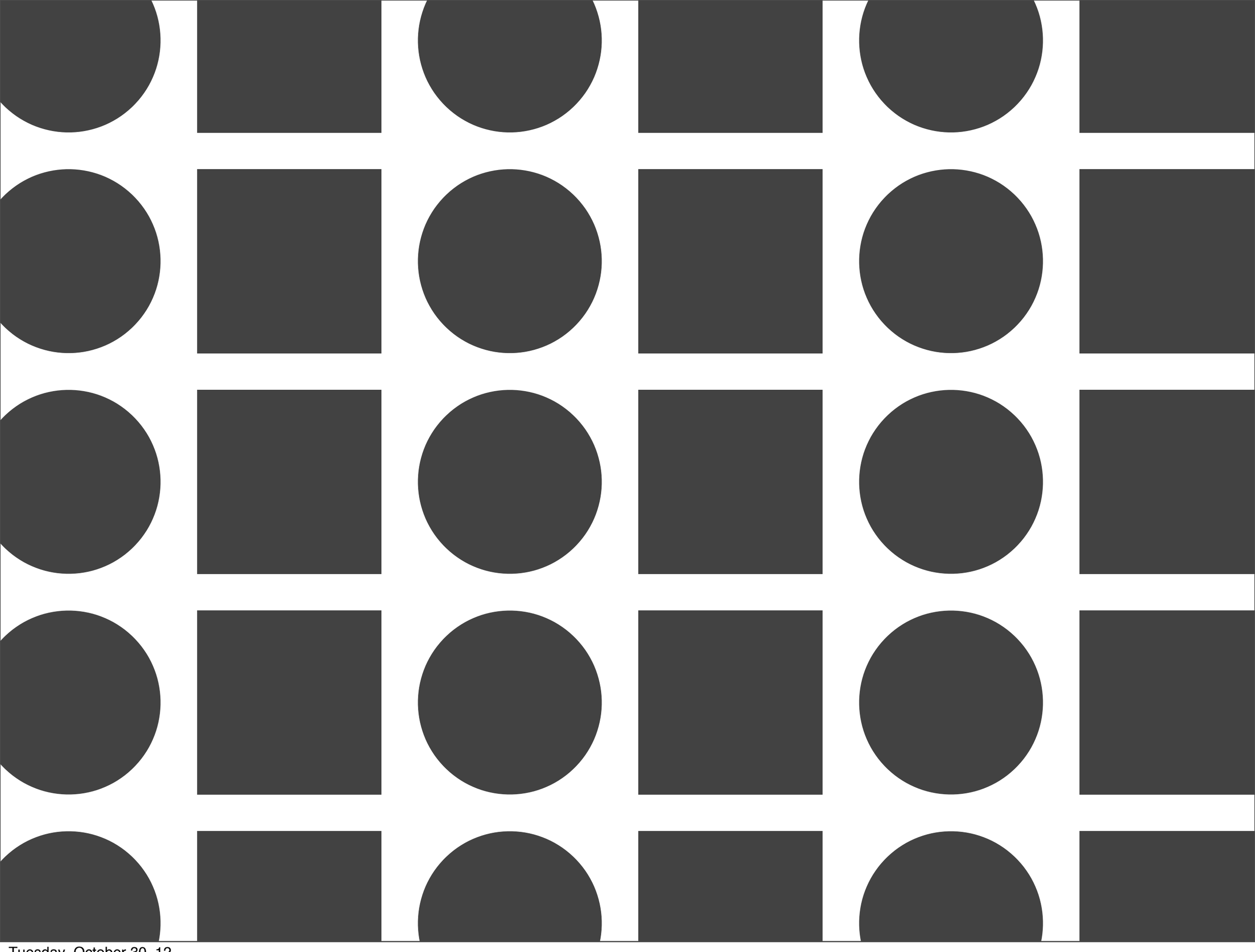


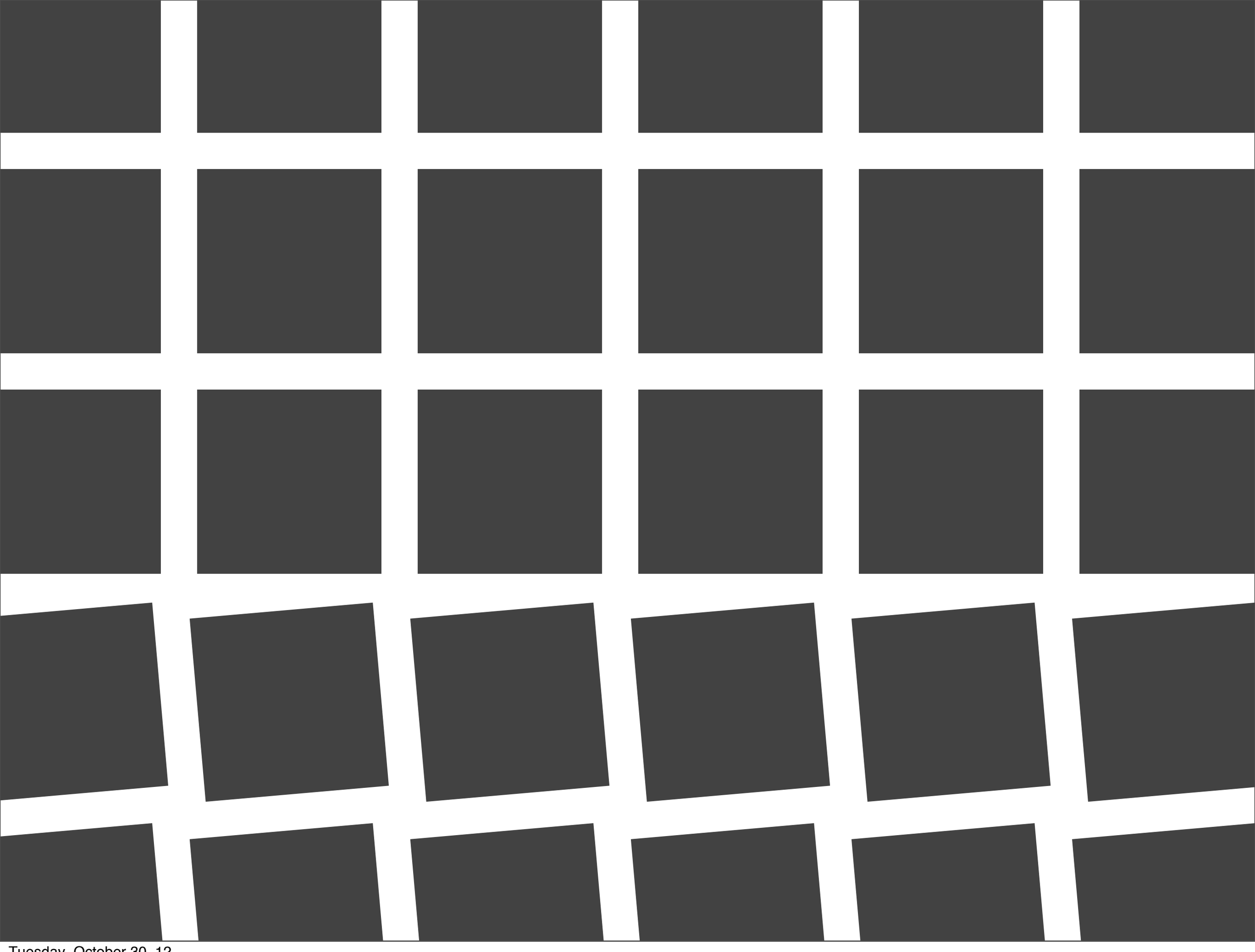




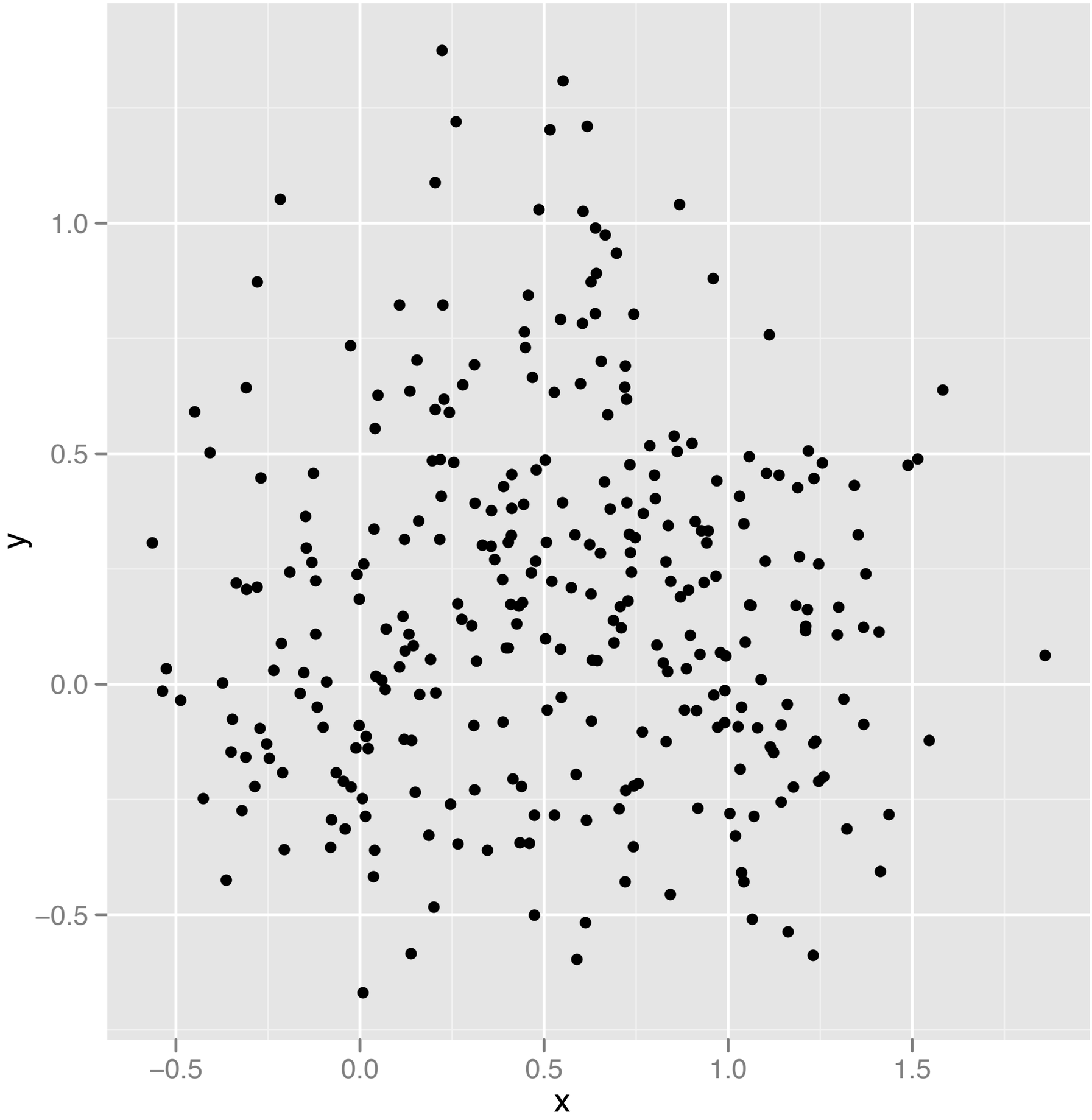




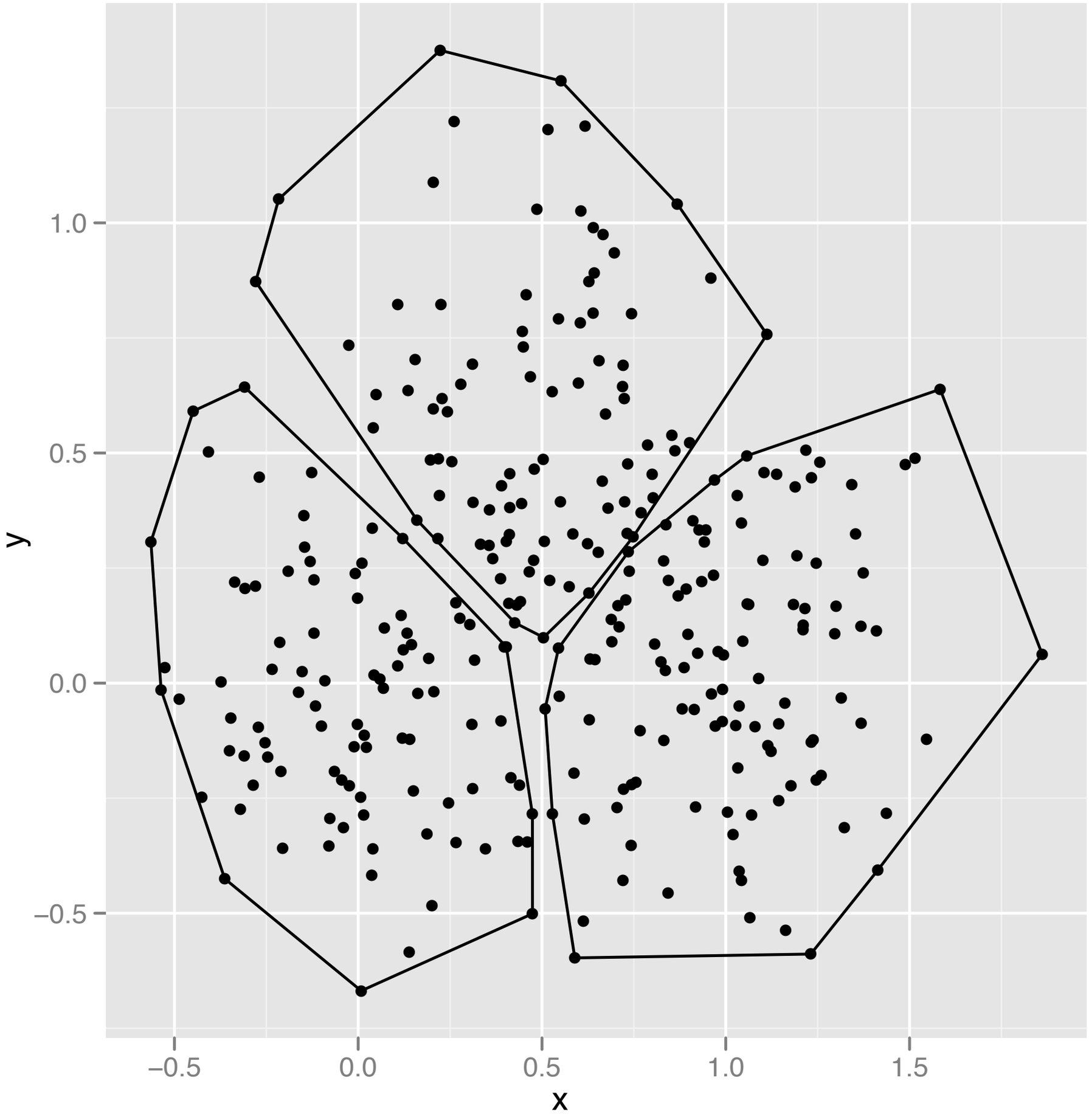




Beware



Beware



Your turn

In small groups, work through each of the three graphics. Are components of the graphics appropriately connected?

4

Beware of animation!

4

Beware of animation!

(Compare in space, not time)

We often don’t notice abrupt changes **http://youtu.be/FWSxSQsspiQ?t=0m12s**

We often miss gradual changes too **http://youtu.be/1nL5ulsWMYc**

And movement makes us miss other changes

http://visionlab.harvard.edu/silencing/

Your turn

In small groups, work through the three graphics. (Use the online version of the facebook graphic at **http://nyti.ms/NEgIDh**) How has animation been used? Is it effective or ineffective?

1

Match perceptual and data topology

2

Make important comparisons easy

Visual connections should reflect real connections

4

Beware of animation!

5*

Visualisation is only one part of data analysis