# Visualization through human eyes

# Readings for today

- Franconeri, S., L. Padilla, P. Shah, J. Zacks, & J. Hullman (2021). The science of visual data communication: What works. Psychological Science in the Public Interest.

# Topics

1. What the human visual system “sees”

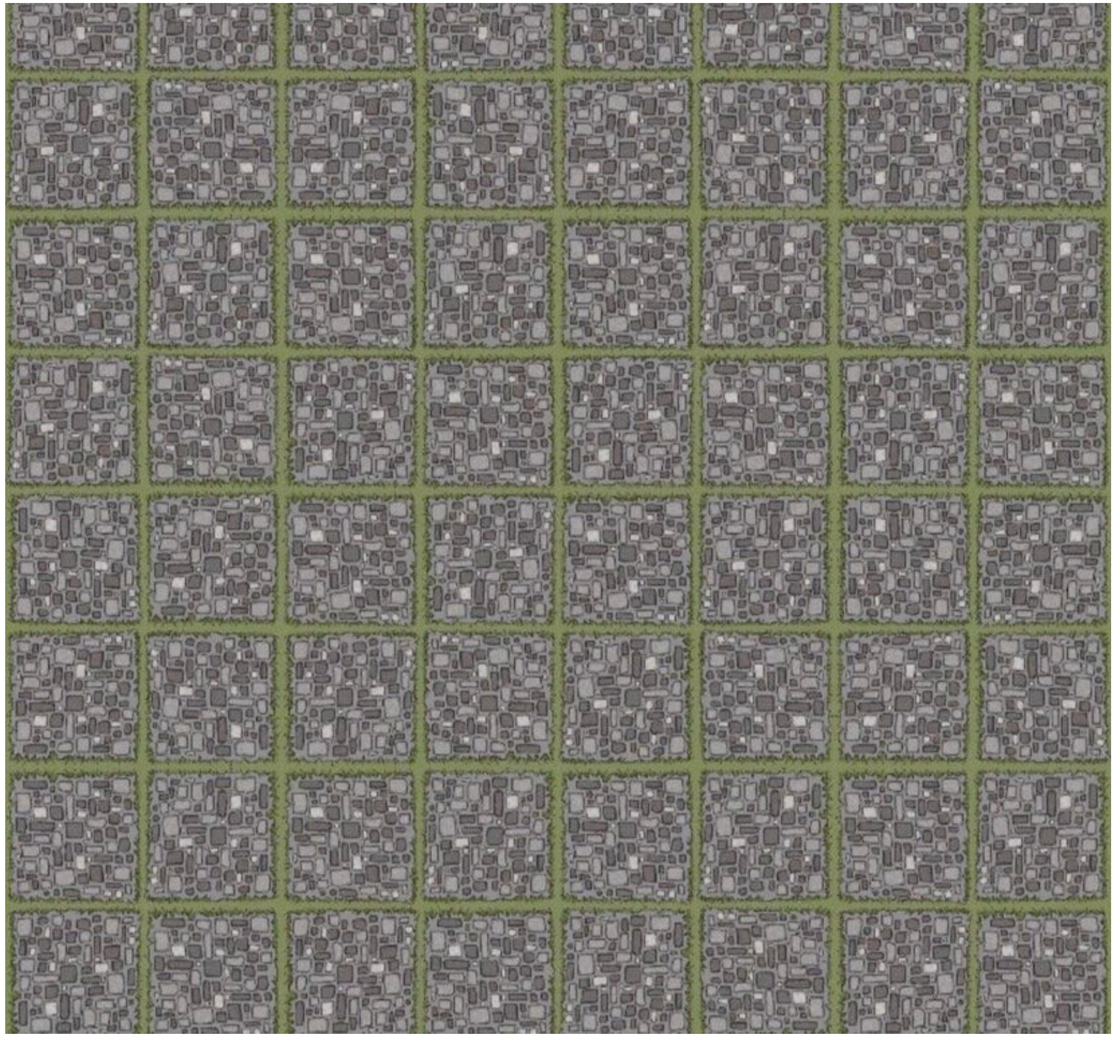
2. Working with visual biases

# What the human visual system “sees”

# Seeing is not believing

Humans see the world biased by their internal expectations (i.e., biases/priors), not based purely on the senses themselves.

There are no curved lines

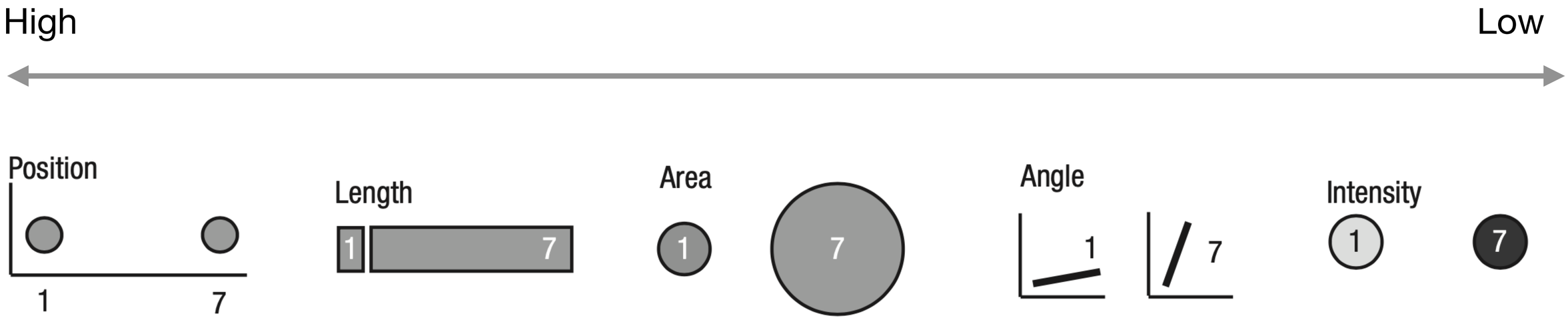




(Ditzinger 2021)

# Precision of visual channels

**Precision:** The ability by which viewers can verbally state the relations between 2 values when visually presented.

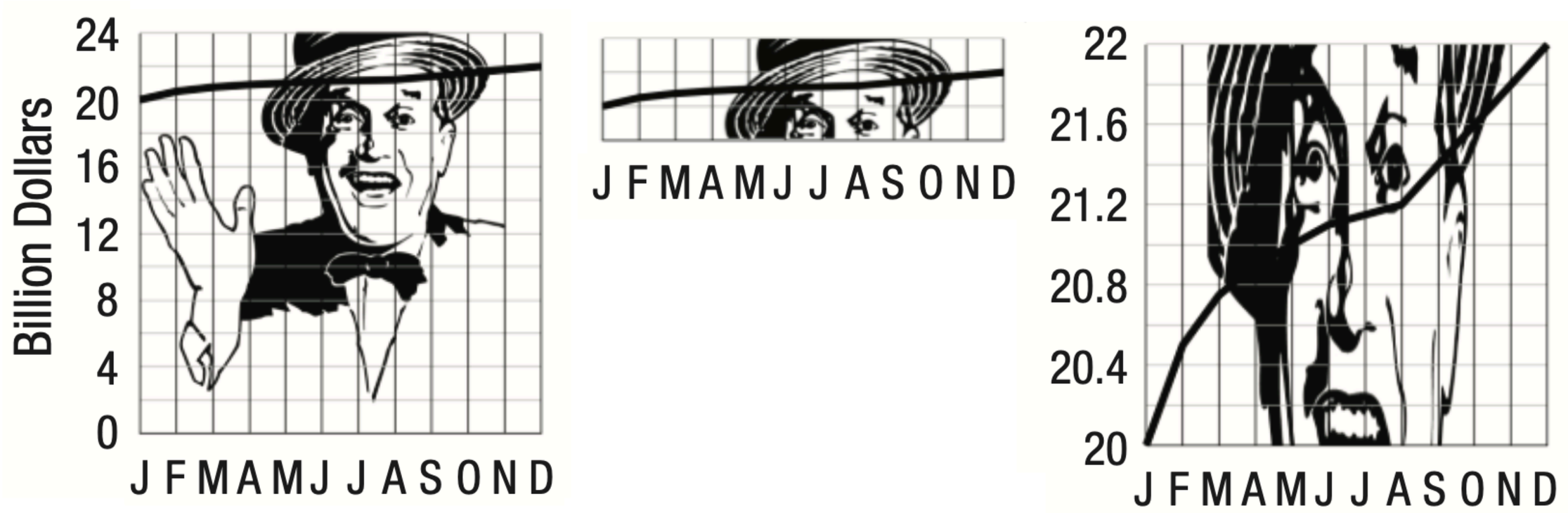





(Francoreni et al. 2021)

# Reference frames → understanding

Mapping visual ratios to numbers can lead to perceptual & inferential errors.

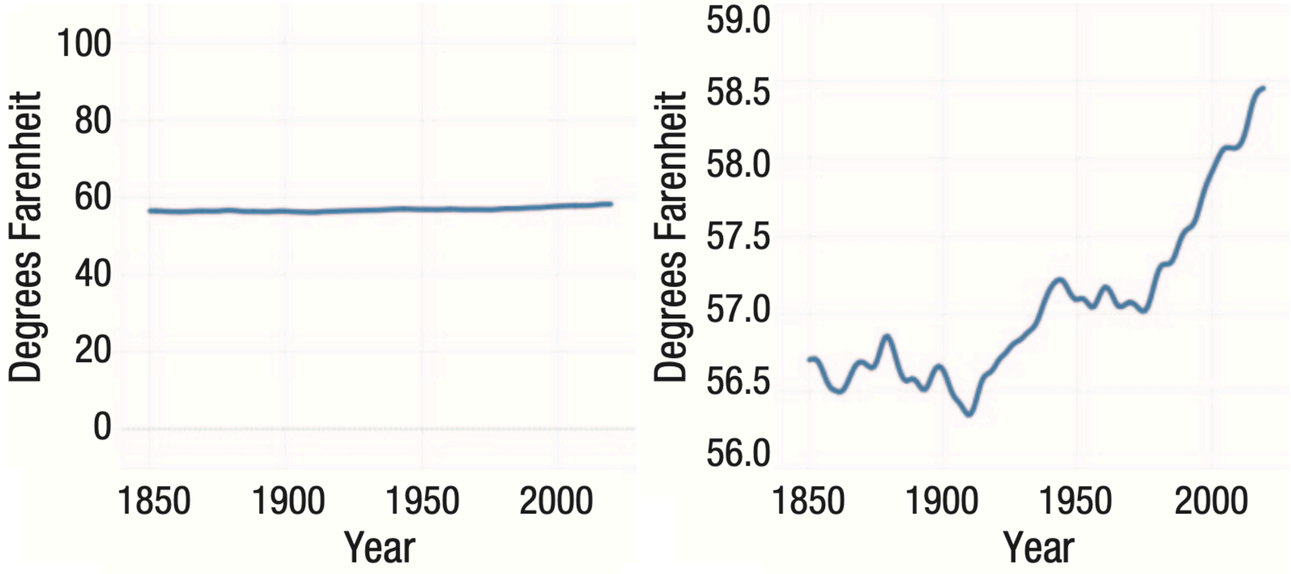





(Francoreni et al. 2021)

# Reference frames → understanding

Mapping visual ratios to numbers can lead to perceptual & inferential errors.



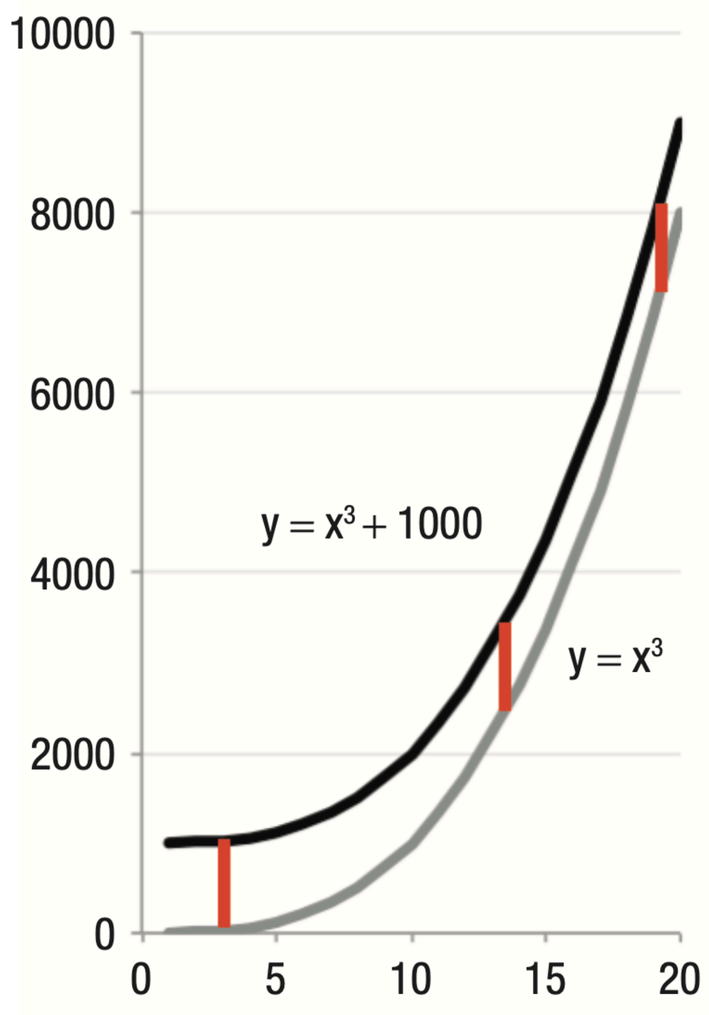

Designers should consider the practical relevance of scale, reference frame, etc. when thinking of how they want to the viewer to encode the information.



(Francoreni et al. 2021)

# Common illusions & misperceptions

- Hue & position lead to perceived distance/difference errors.





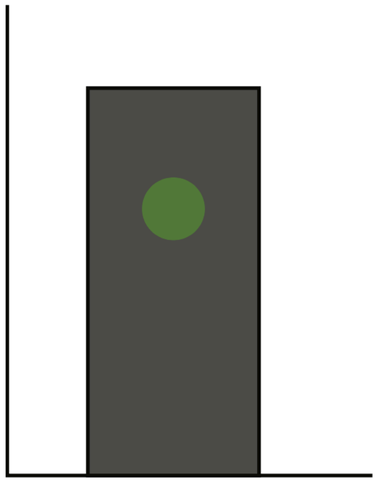

(Francoreni et al. 2021)

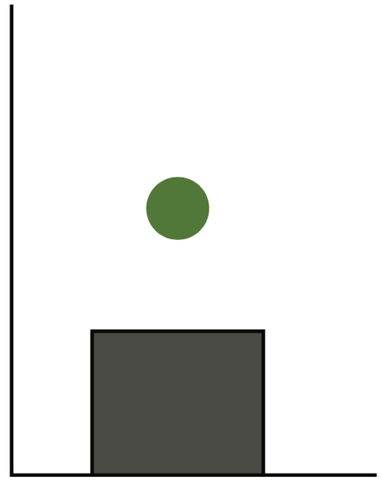
# Common illusions & misperceptions

- Hue & position lead to perceived distance/difference errors.
- Contrast & context skew perceptions of intensity.


Intensity values can look different depending their backgrounds.

Do not plot intensities on intensities.



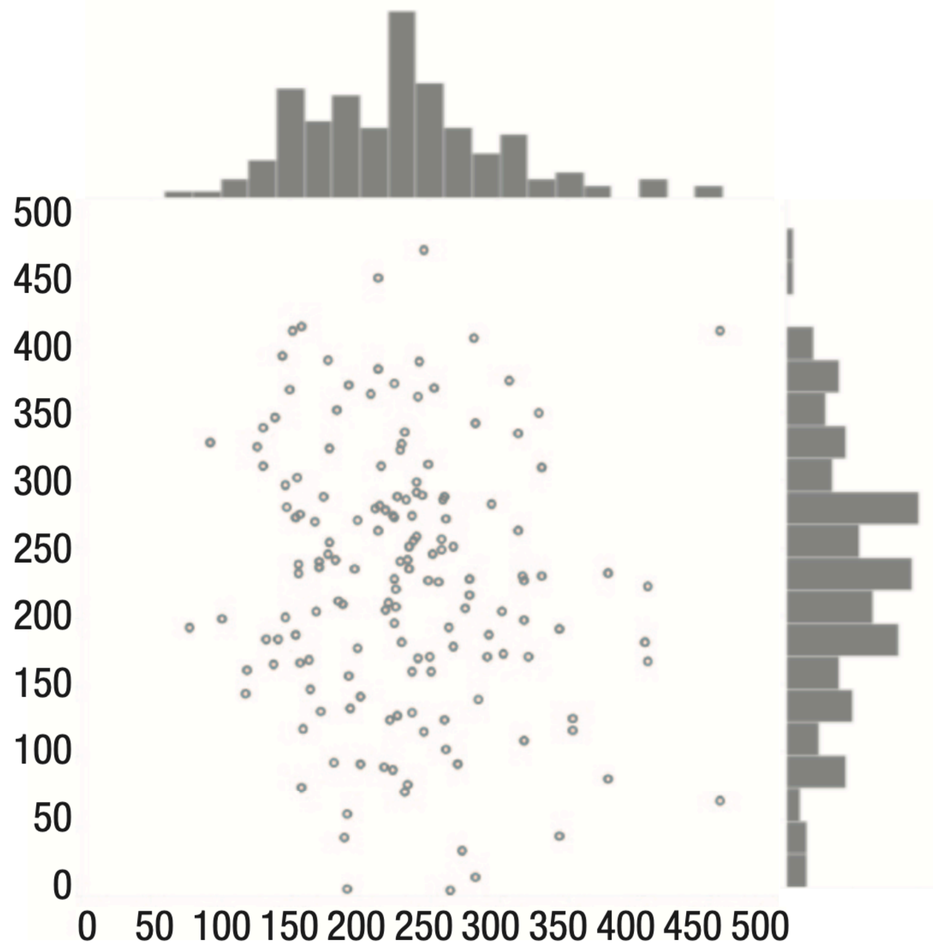




(Francoreni et al. 2021)

# Common illusions & misperceptions

- Hue & position lead to perceived distance/difference errors.
- Contrast & context skew perceptions of intensity.
- Combining across dimensions can lead to spurious inferences of associations.

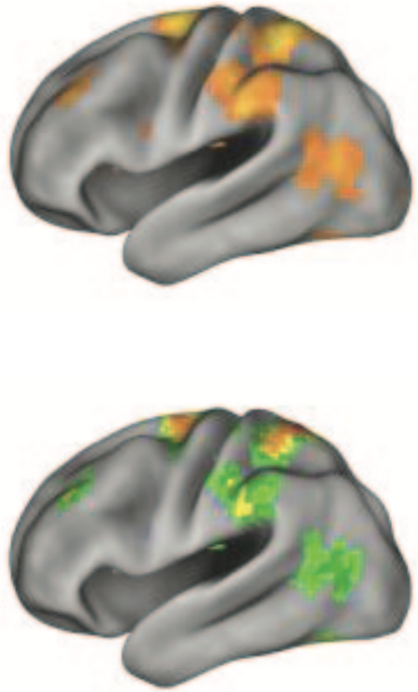





(Francoreni et al. 2021)

# Common illusions & misperceptions

- Hue & position lead to perceived distance/difference errors.
- Contrast & context skew perceptions of intensity.
- Combining across dimensions can lead to spurious inferences of associations.
- Humans tend towards categorical biases, even when no categories are present.



(Francoreni et al. 2021)

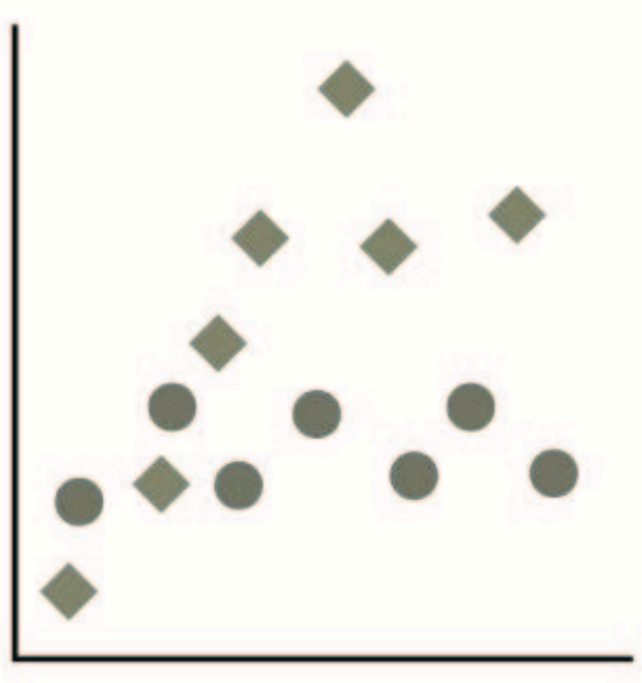
# Working with visual biases

# Use visualizations to compute fast, simple statistics

Vision quickly extracts simple statistics from data encoded via:

1. position
2. one dimensional length
3. two dimensional area
4. angles
5. intensities

**Question:** Are $X$ and $Y$ associated?

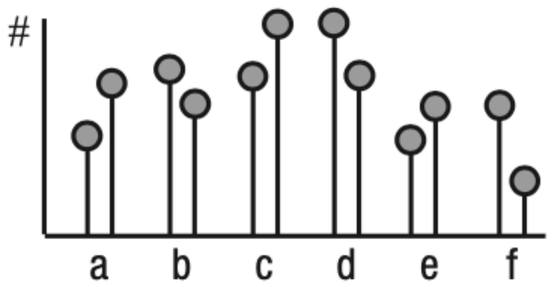




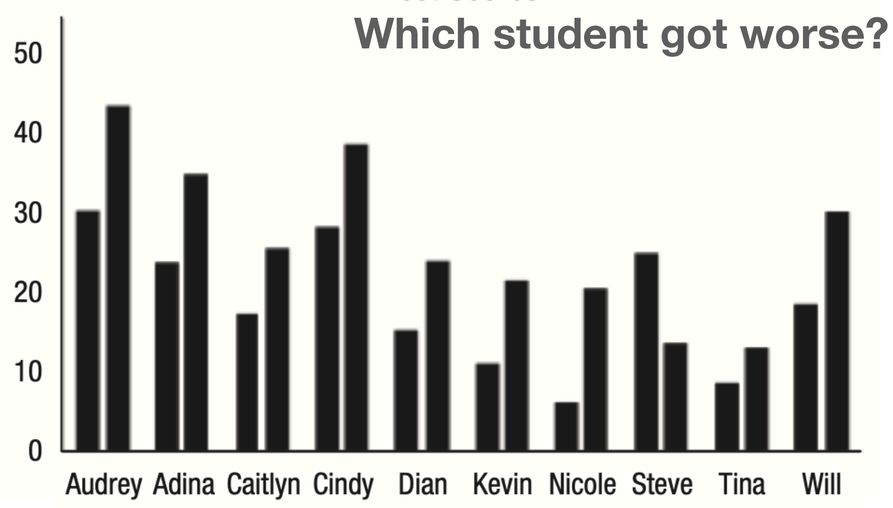
(Francoreni et al. 2021)

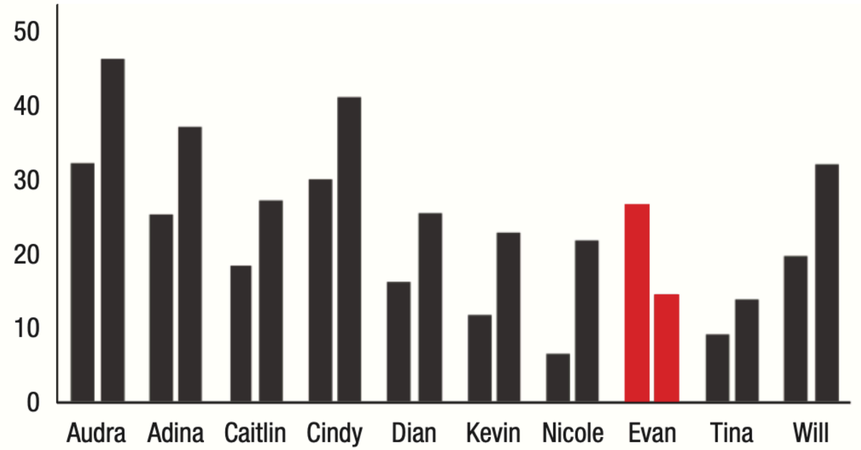
# Comparisons as a processing limit

- Seeing general statistical relations (e.g., value, distribution) is ***fast***.
- Extracting relationships (and differences) is ***slow***.




Tool: Shortcut comparisons by adding direct depictions of the deltas, as below

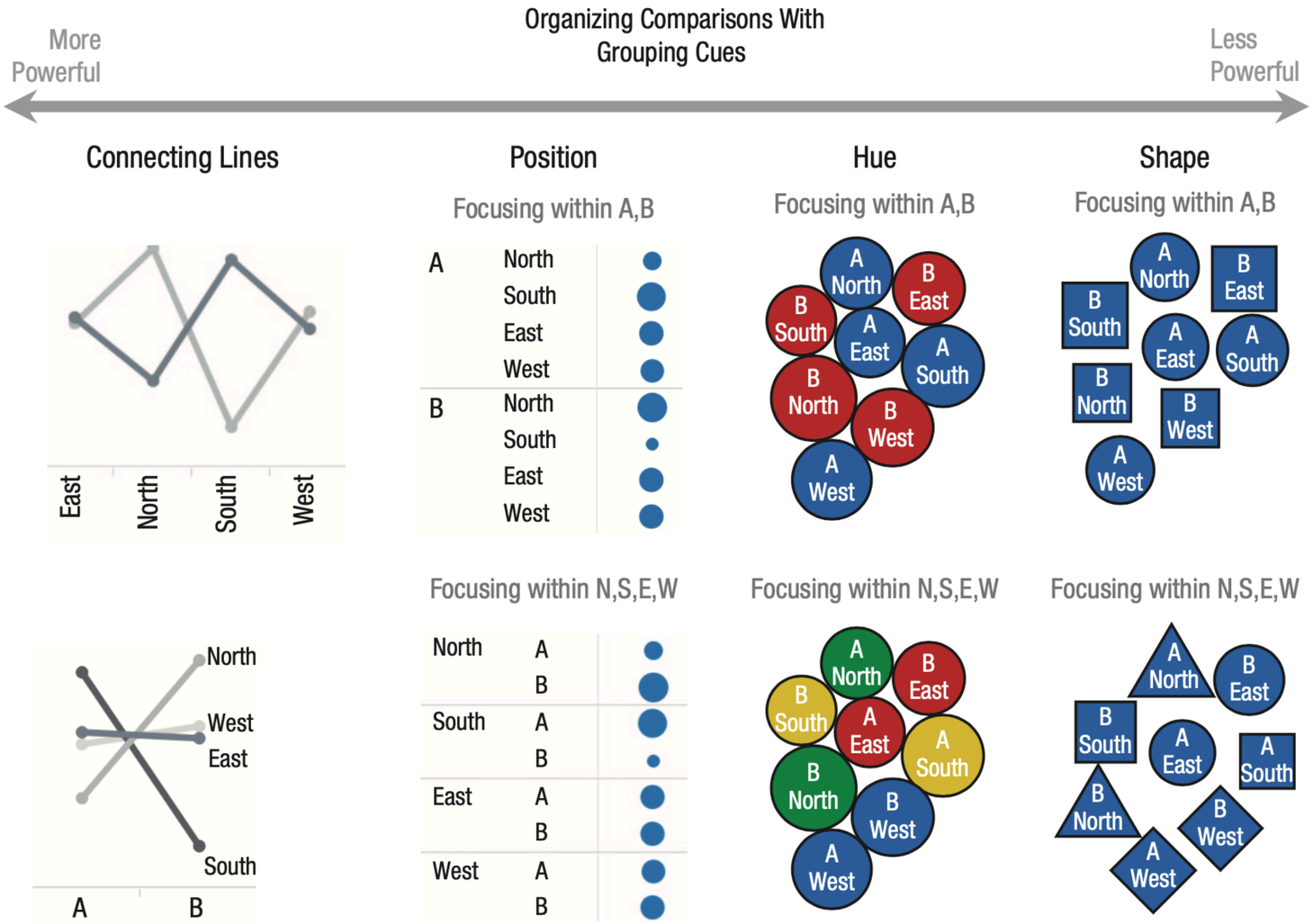






(Francoreni et al. 2021)

# Grouping to facilitate comparisons

Proximity (position) and hue/ shape are natural cues for similarity and grouping respectively.





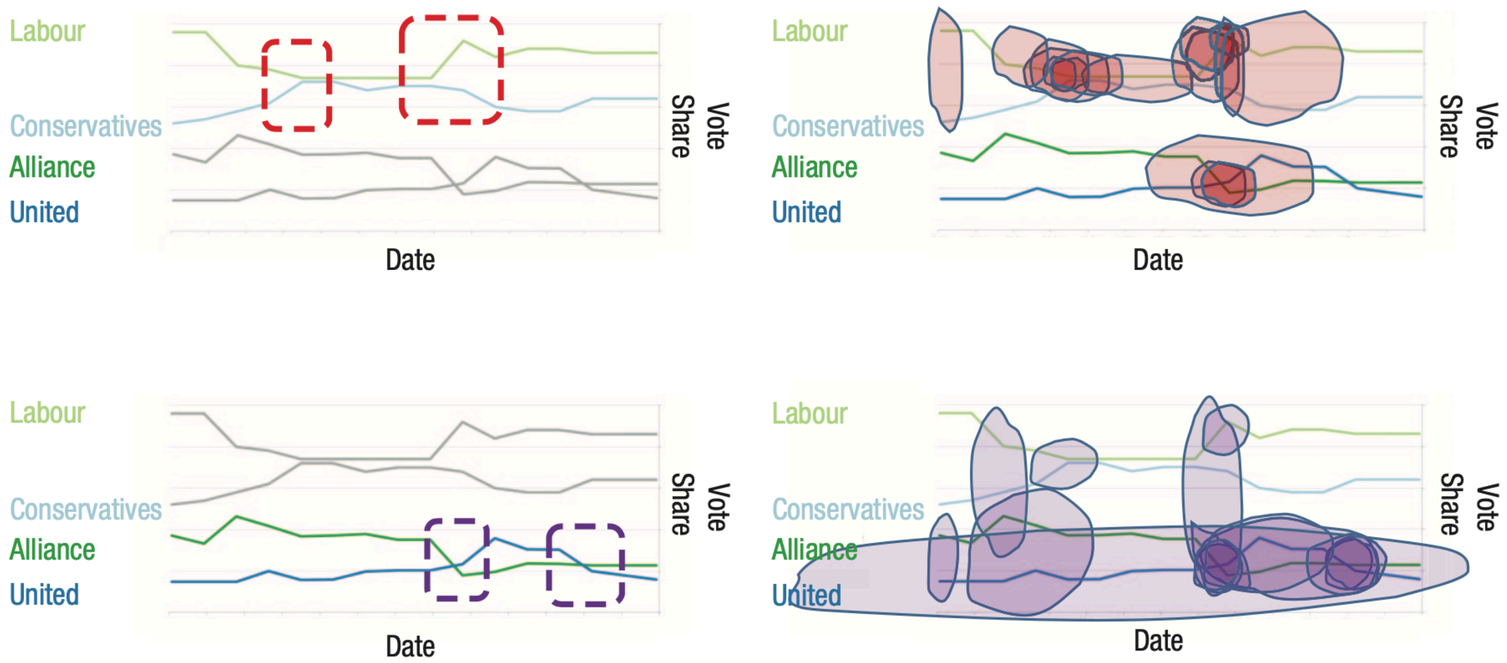

(Francoreni et al. 2021)

# Curse of knowledge

**Curse of Knowledge**: Inability to simulate the perspective of another viewer.

- Prior knowledge guides how viewer's *see* graphics.

(Xiong et al. 2020)



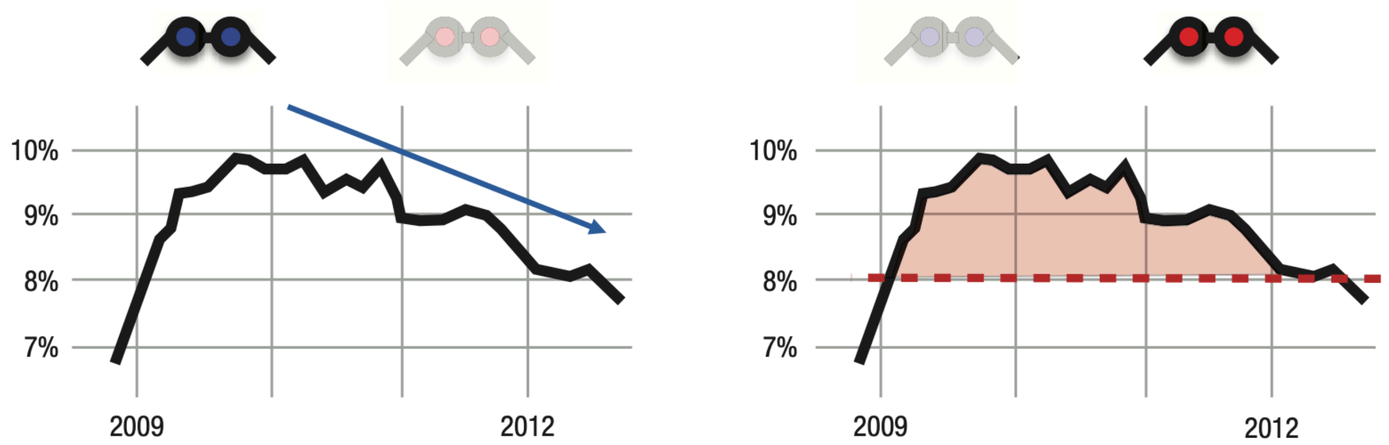


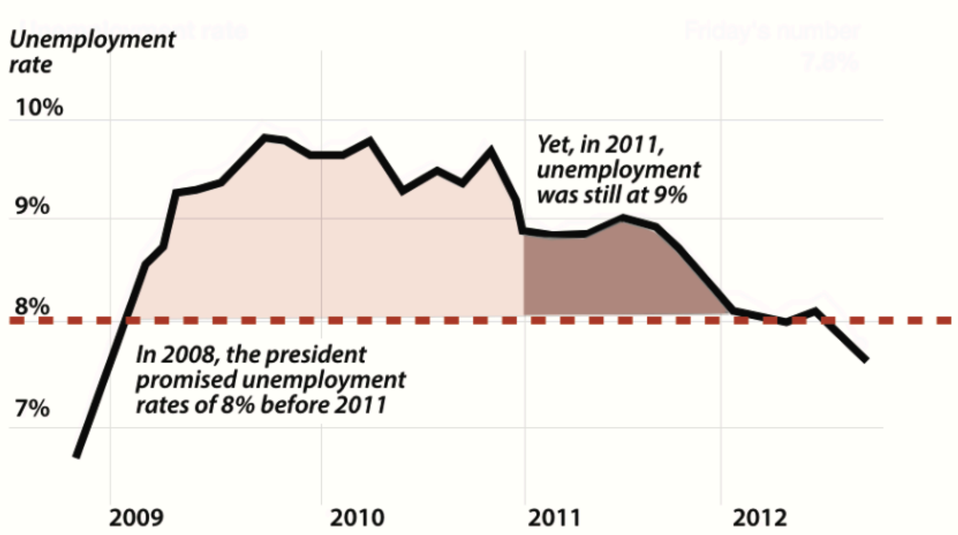

(Francoreni et al. 2021)

# Curse of knowledge

**Curse of Knowledge**: Inability to simulate the perspective of another viewer.

- Prior knowledge guides how viewer's *see* graphics.
- Use visual features (e.g., color, shading) to guide viewer's eye.
- Use annotations for direct inference.



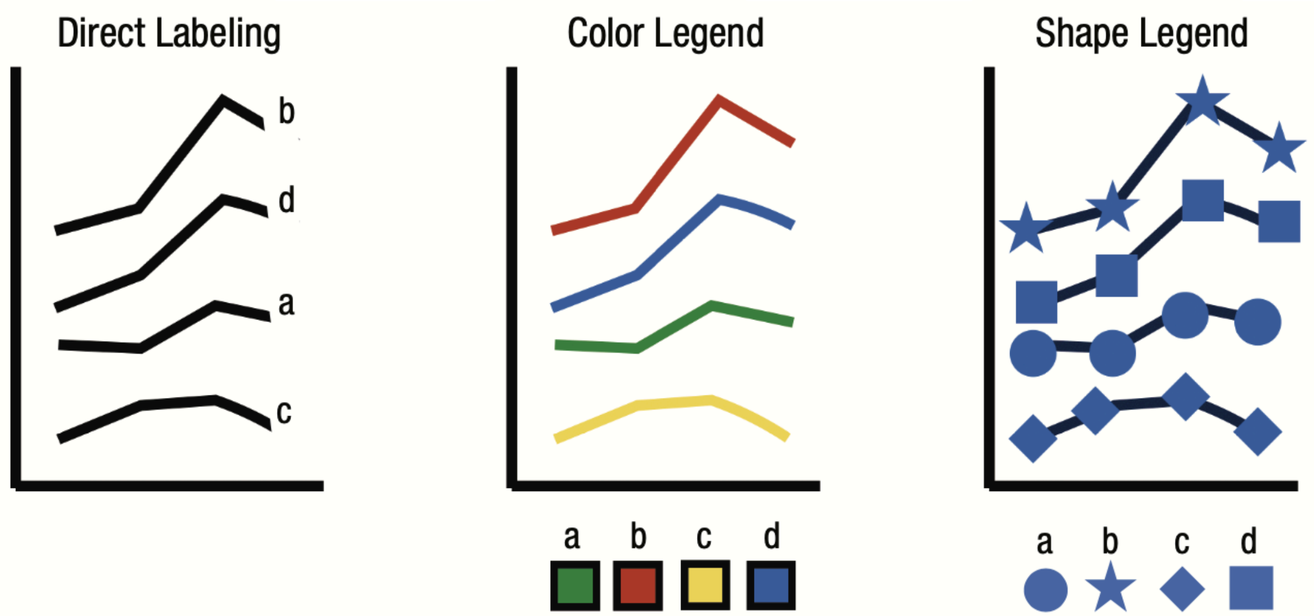

Inspired by:
http://www.nytimes.com/interactive/2012/10/05/business/economy/one-report-diverging-perspectives.html



(Francoreni et al. 2021)

# Avoid taxing working memory

People answer questions about data faster and more accurately when data are directly labeled in graph, rather than on legends or captions.

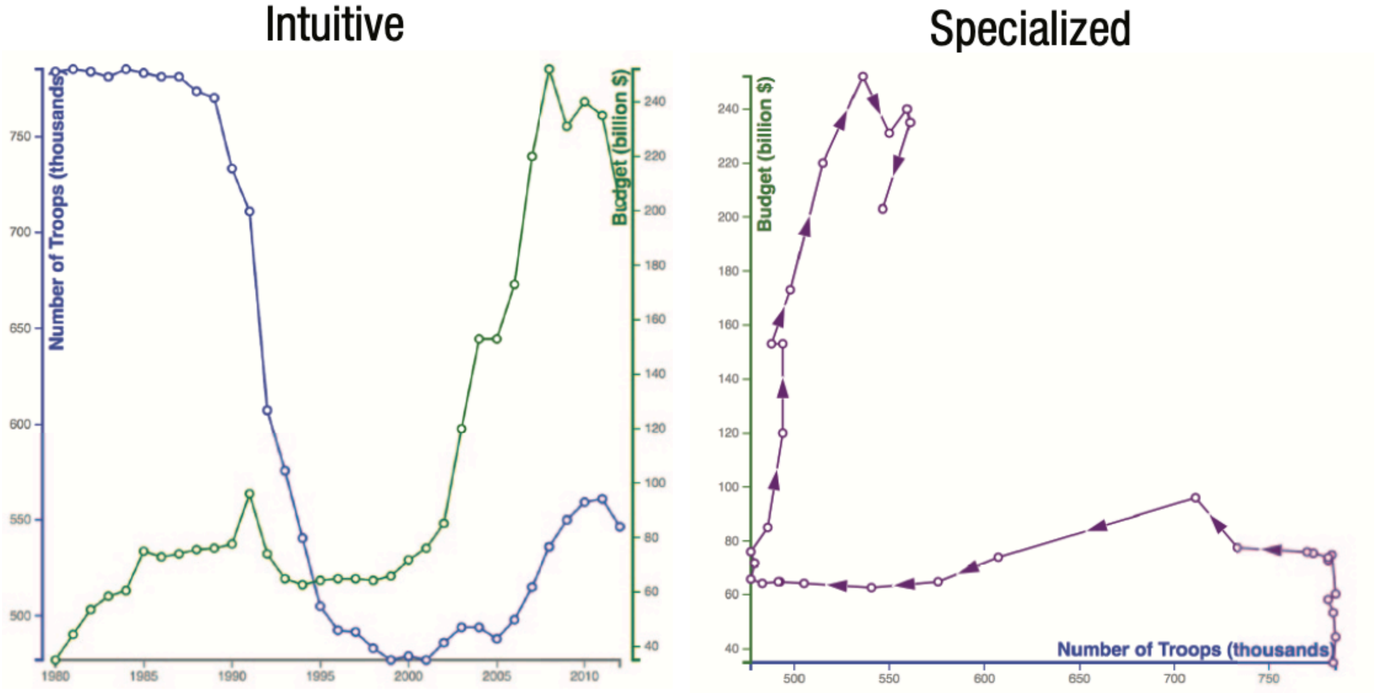





(Francoreni et al. 2021)

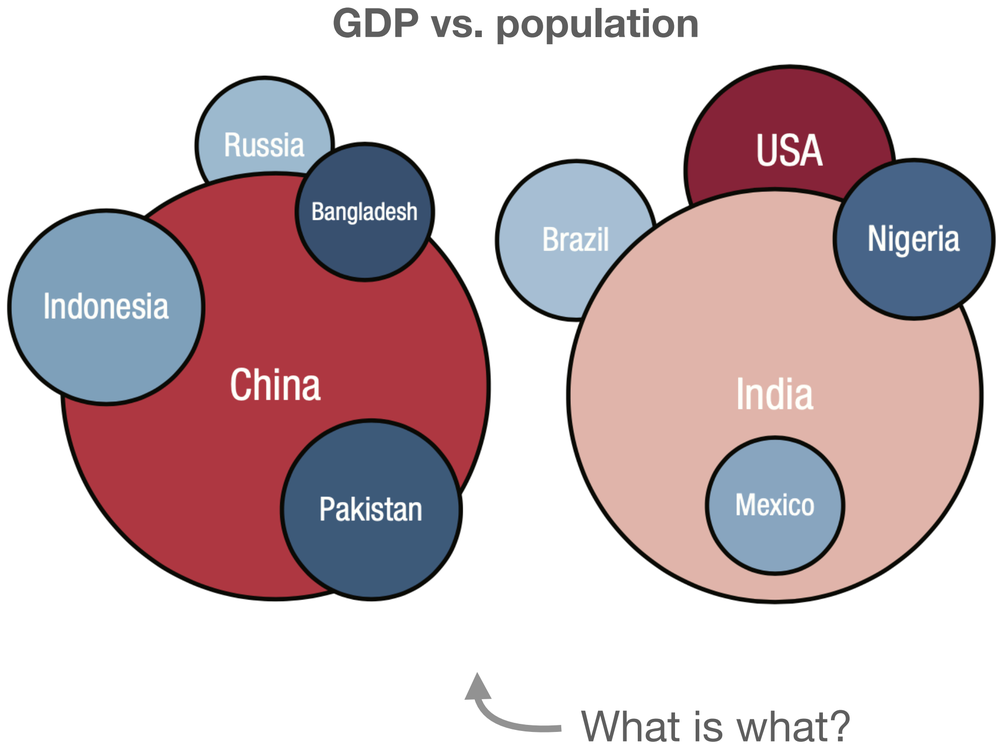
# Familiarity facilitates comprehension

**Graph schema**: A knowledge structure that includes default expectations, rules, and associations that a viewer uses to extract conceptual information from the data visualization.









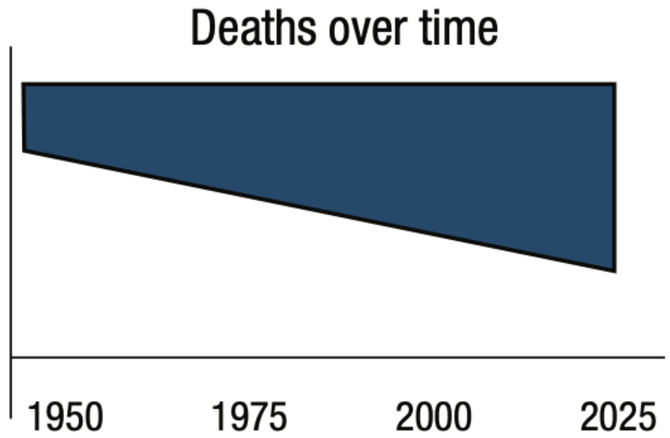



(Francoreni et al. 2021)

# Respect the viewer's assumptions

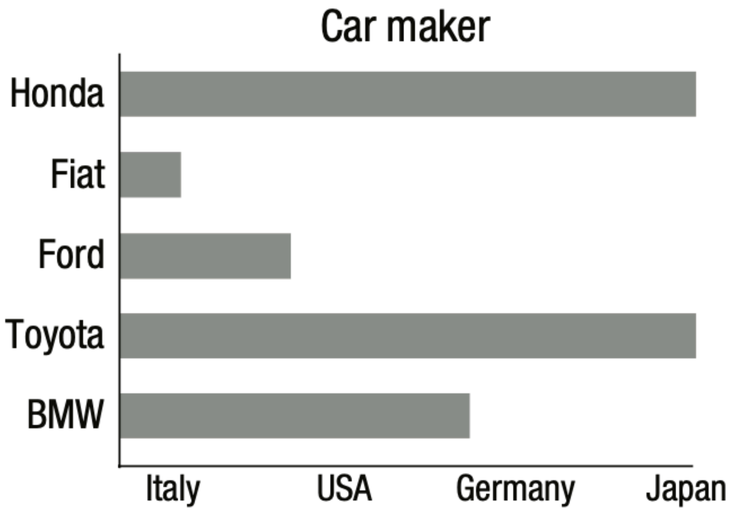
- Some assumptions are universal

$$\text{Length or area} \begin{cases} \uparrow = \text{more} \\ \downarrow = \text{less} \end{cases}$$




**General assumptions**

- **metric data** →position, length, area, angle, or intensity.
- **nominal data** →position, color, shape



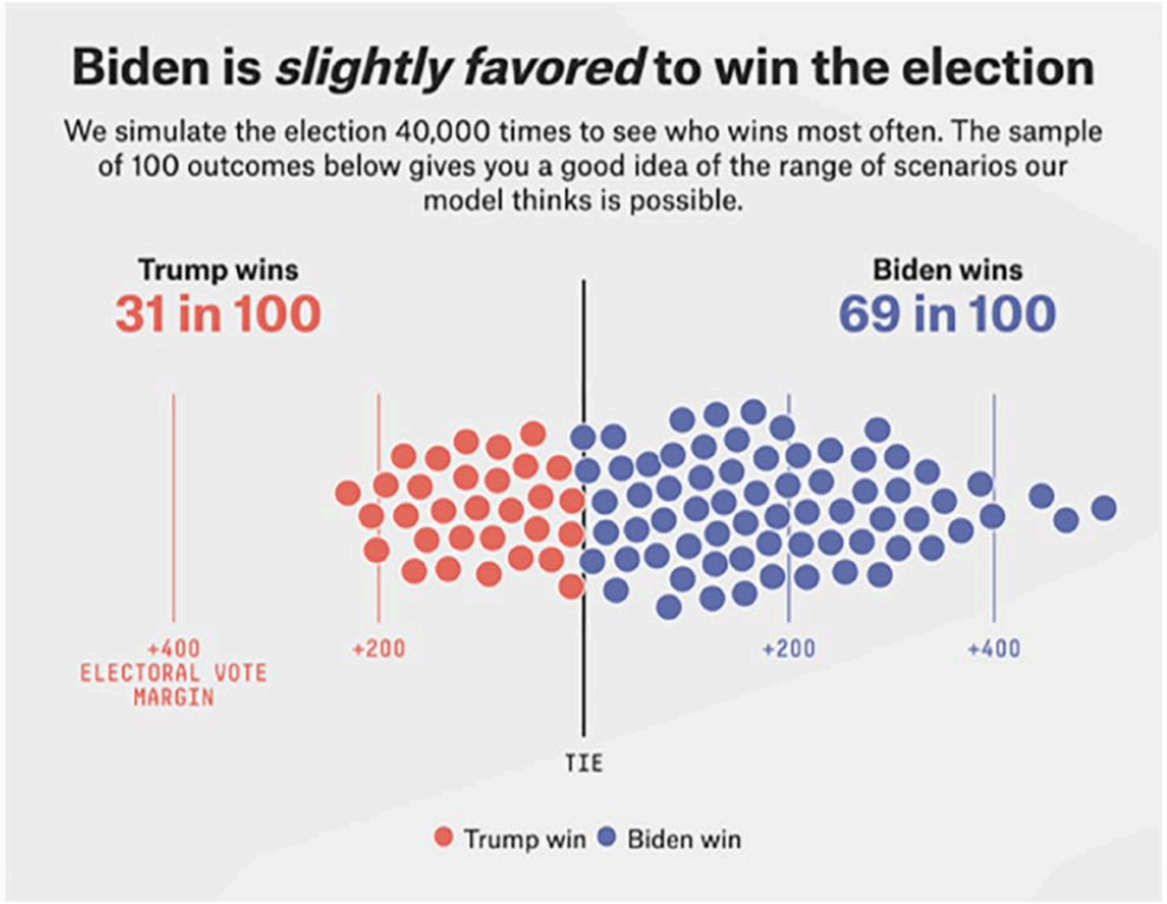



(Francoreni et al. 2021)

# Visual processing → encoding

- **Bottom-up** → saliency
- **Top-down** → features

(FiveThirtyEight.com)



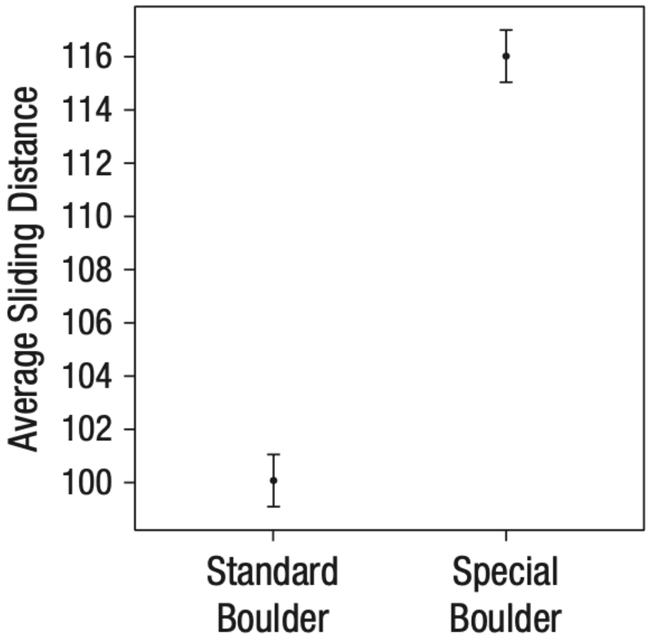



(Francoreni et al. 2021)

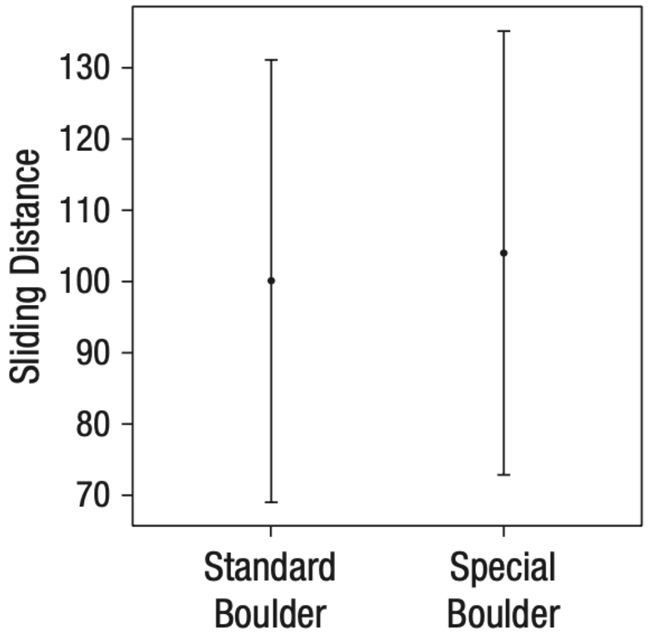
# Communicate uncertainty effectively

Errorbars are highly specialized and prone to misunderstanding.

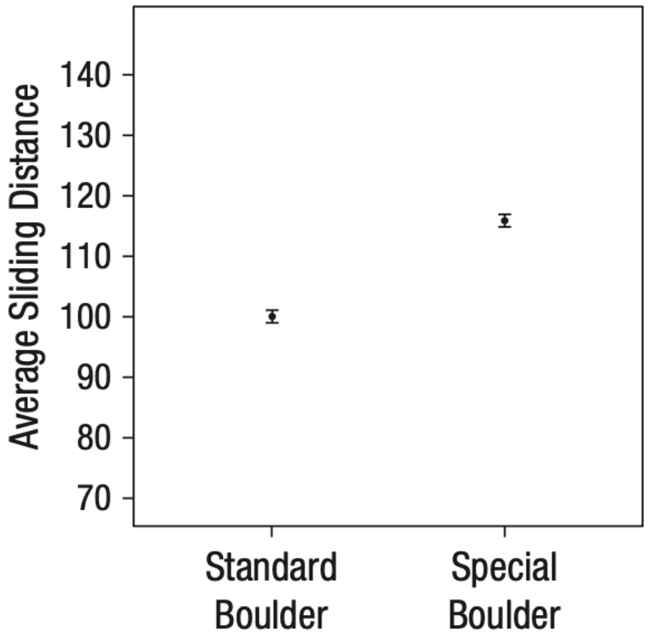
$1.95 \pm SE$




$1.95 \pm SD$




$1.95 \pm SE$ (scaled axis)

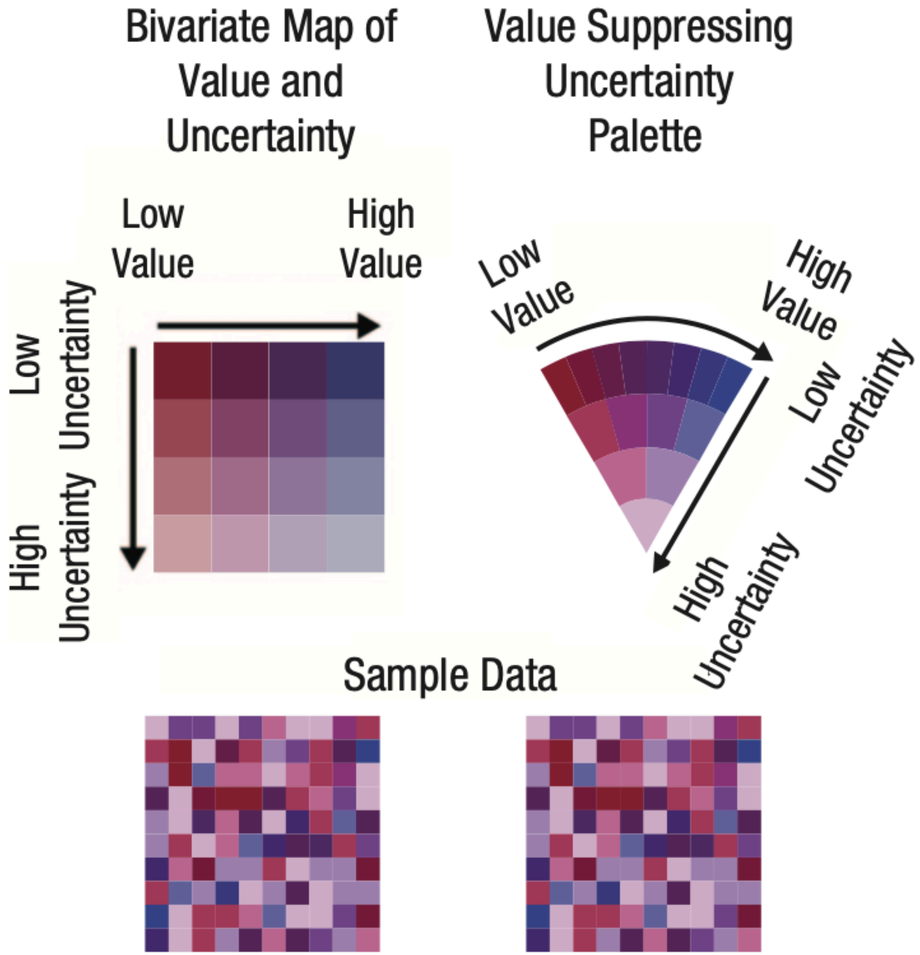





(Francoreni et al. 2021)

# Communicate uncertainty effectively

Some visual channels provide subjective impression of uncertainty, but at lower perceptual precision.



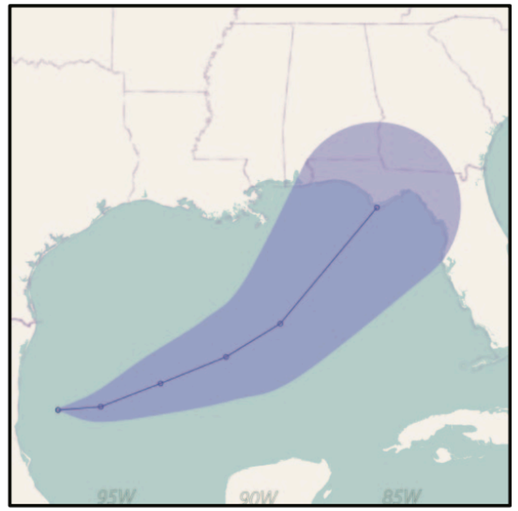


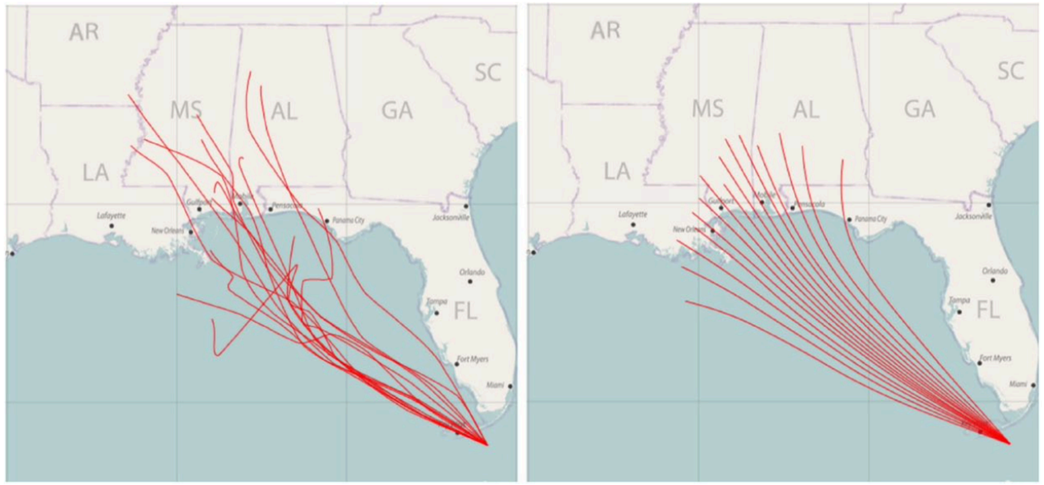

(Francoreni et al. 2021)

# The problem of probabilities

Humans struggle with understanding probabilities.

**Solution**: present probabilities as frequencies or samples





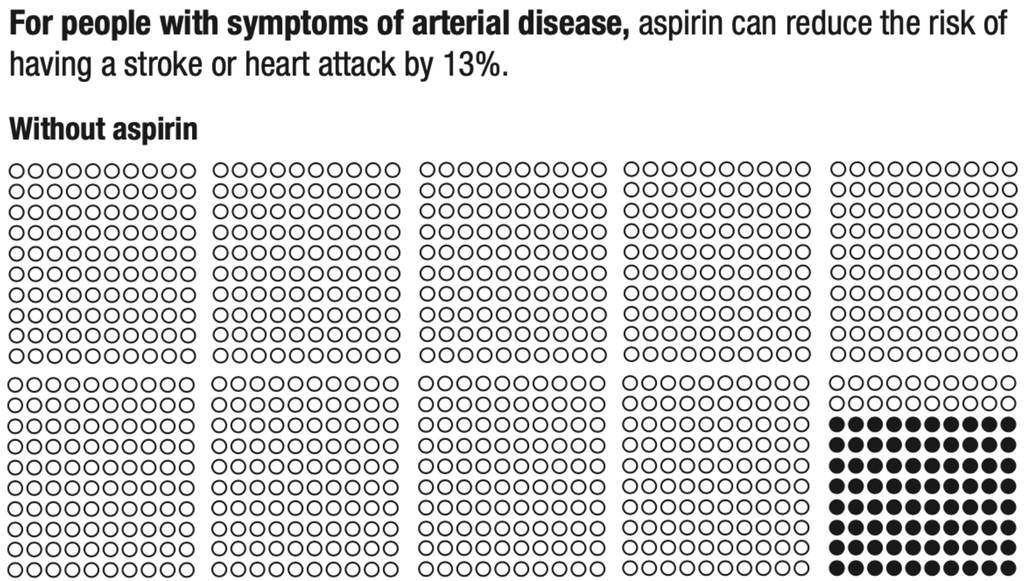

(Francoreni et al. 2021)

# The problem of risk

Problems with understanding probabilities lead to problems in understanding risk

- **absolute risk:** overall probability of an experience.
- **relative risk:** comparison of absolute risk between 2 categories or groups.

  **denominator neglect :** people pay more attention to differences in numerators than denominators.






(Francoreni et al. 2021)

# Guidelines for dataviz for humans

1. Visualize data *before* applying statistics.
2. Beware of visual illusions & misperceptions.
3. Comparisons are slow, global statistics are fast.
4. Avoid taxing working memory.
5. New visualizations must be learned as new schemas.
6. Graph comprehension depends on both bottom-up & top-down visual processing.
7. Communicate your uncertainty effectively.
8. Convey probabilities as frequencies and highlight absolute, not relative, risk.



(Francoreni et al. 2021)

# Take home message

- For the foreseeable future data visualization is used to convey information to humans, so understand human perceptual abilities (and biases) when making decisions about graphical form.