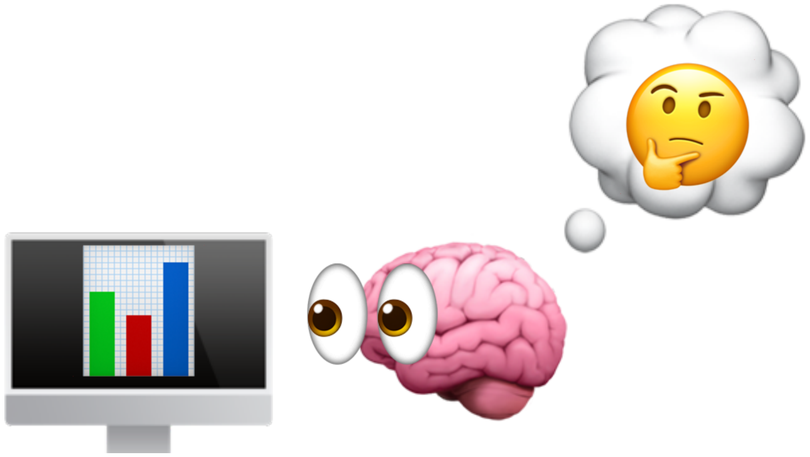
MANCHESTER 1824
The University of Manchester



Cognitive Processing of Magnitude In Data Visualisations

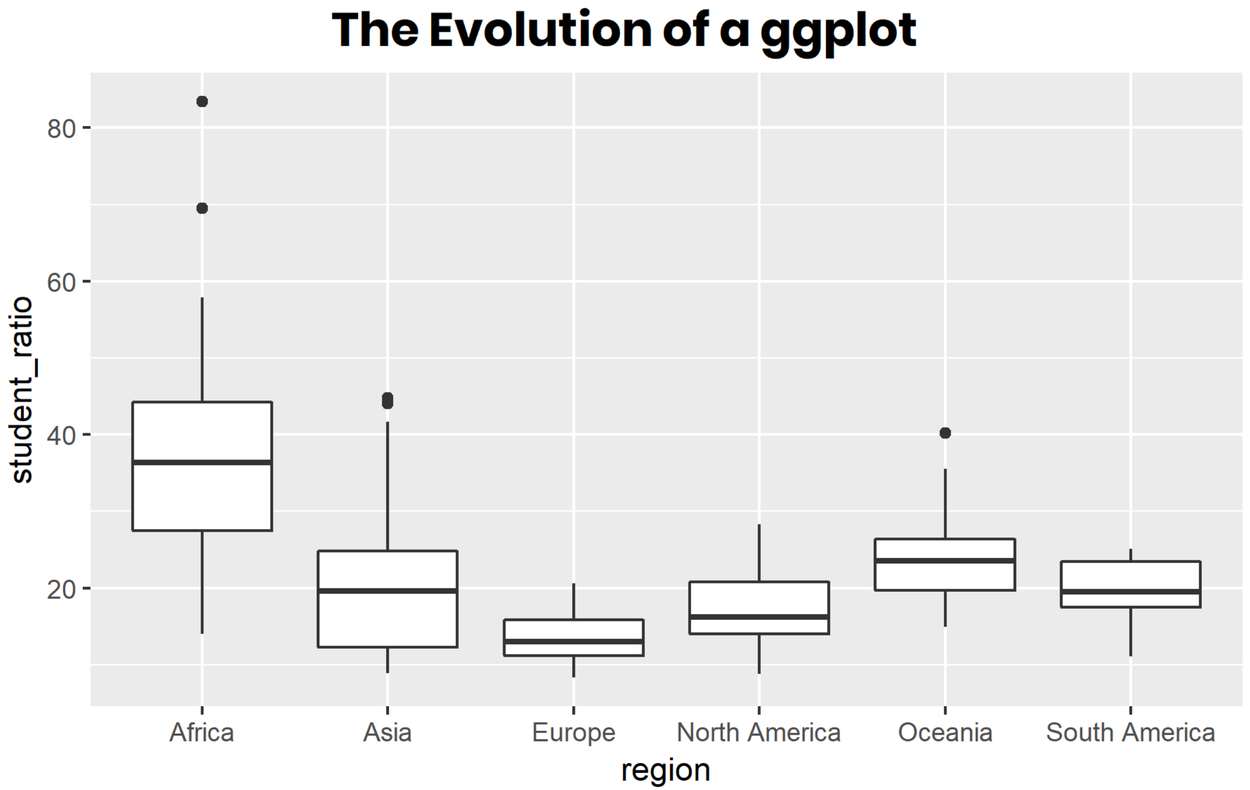
Duncan Bradley

The University of Manchester

@duncanbradley_

Data Visualisation Design: Good and Bad

- Encode values using visual features
- Countless ways to visualise the same dataset
- Some choices can result in misleading charts

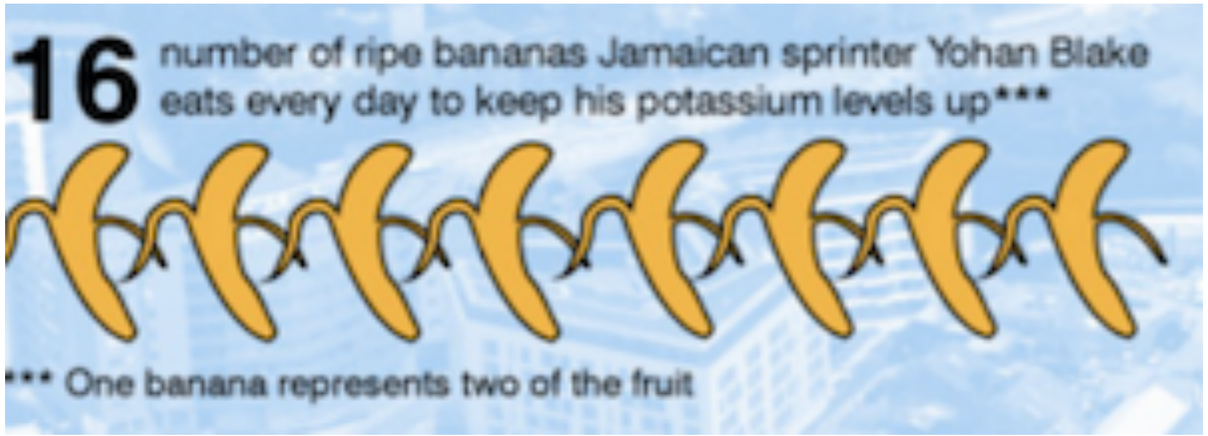


Data: UNESCO Institute for Statistics
Visualization by Cédric Scherer

Cédric Scherer

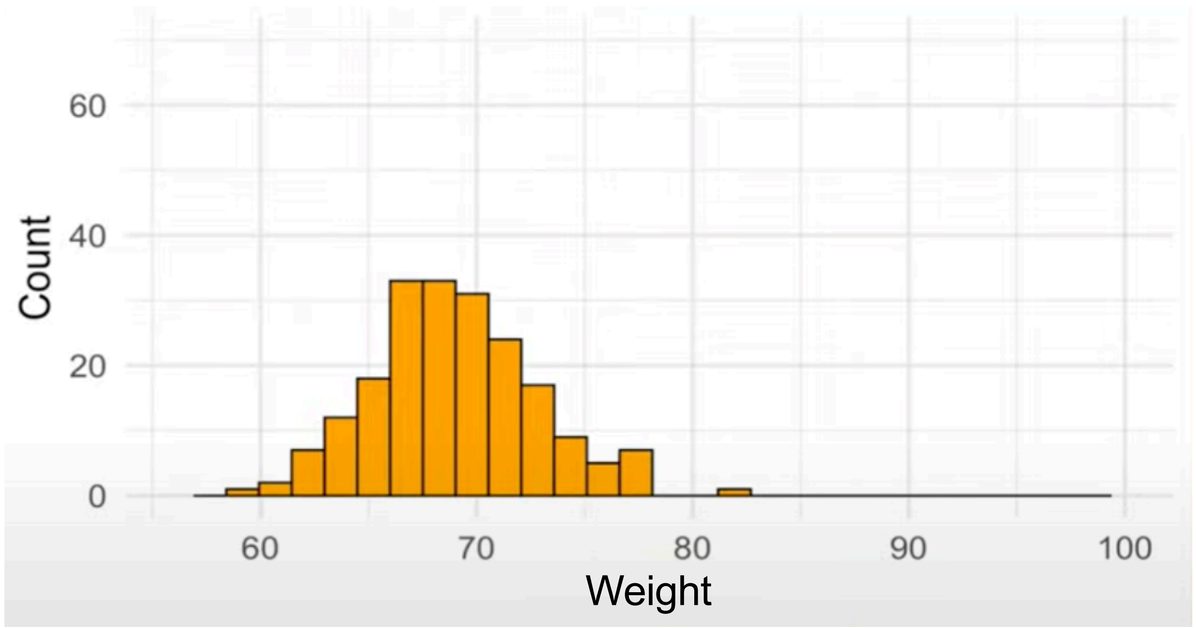
Data Visualisation Design: Good and Bad

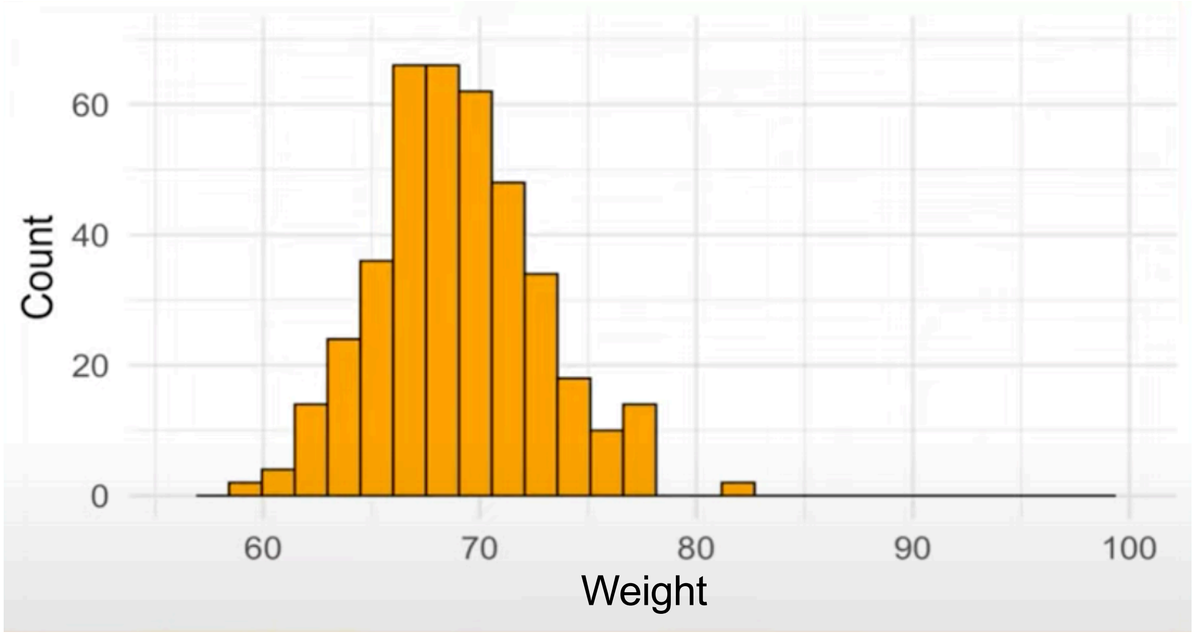
- Encode values using visual features
- Countless ways to visualise the same dataset
- Some choices can result in misleading charts
- And cognitive biases can interfere with interpretation…



CNN
http://edition.cnn.com/2012/08/03/sport/olympics-nutrition-phelps-blake/index.html

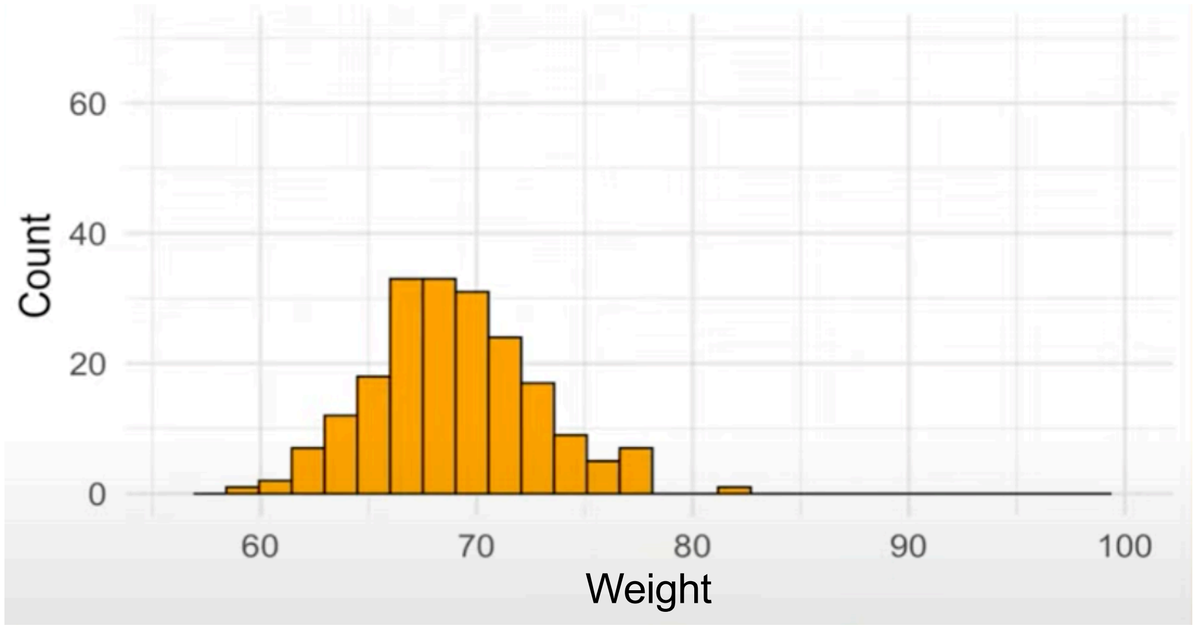
Histograms

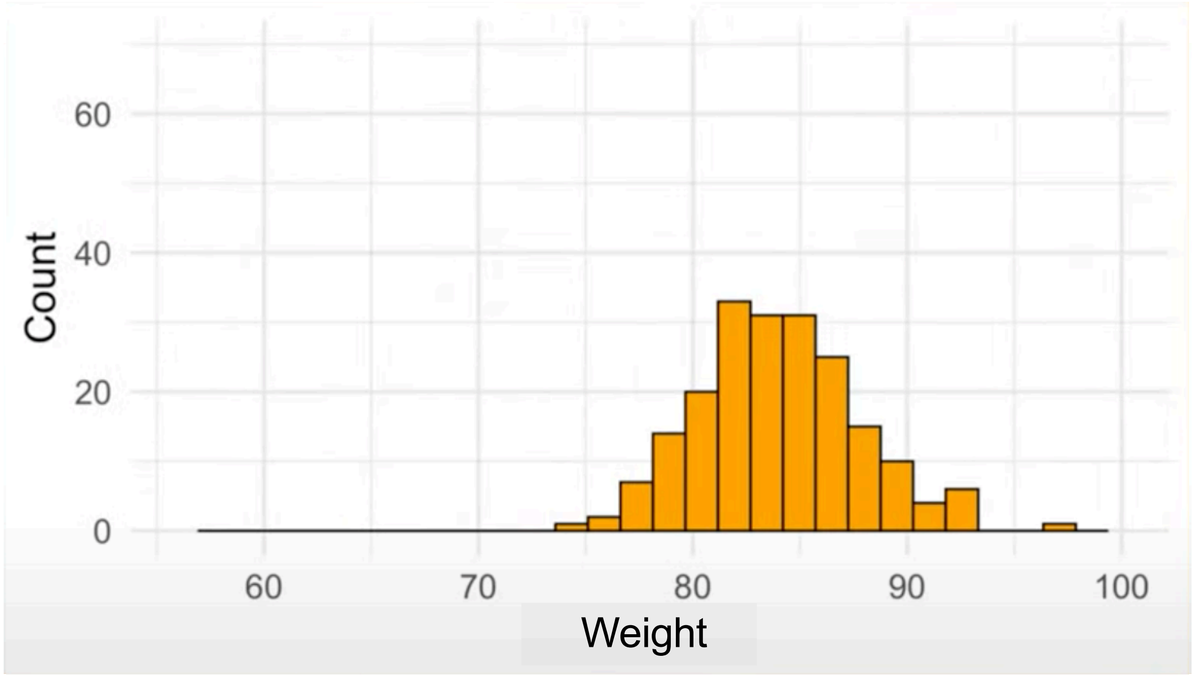




Lem et al. (2014)

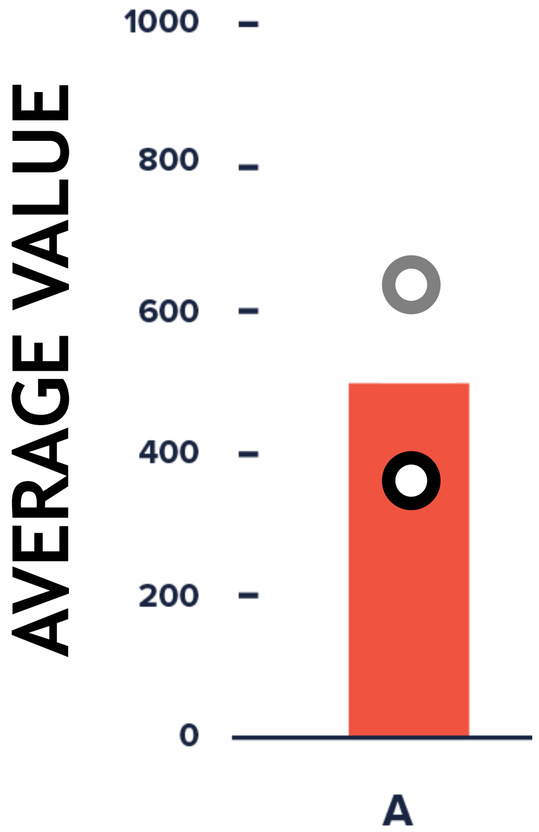
Histograms





Lem et al. (2014)

Bar Charts Depicting Averages



"Within-bar bias"

Newman and Scholl (2014)

Studying Cognitive Processing

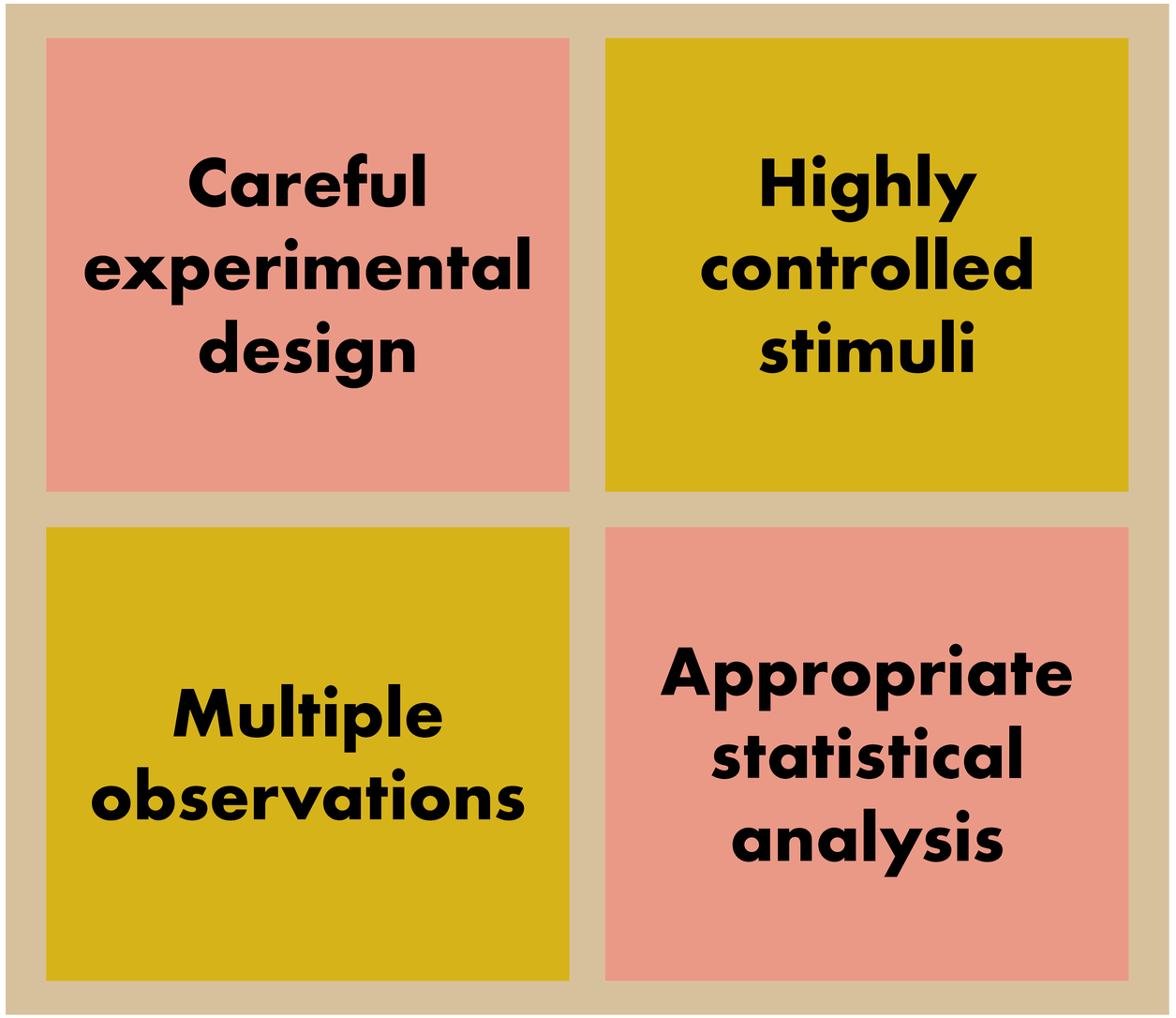
- Deliberate deception and improper designs aren't the only issues

Misleading Charts 🤝 Ineffective Charts

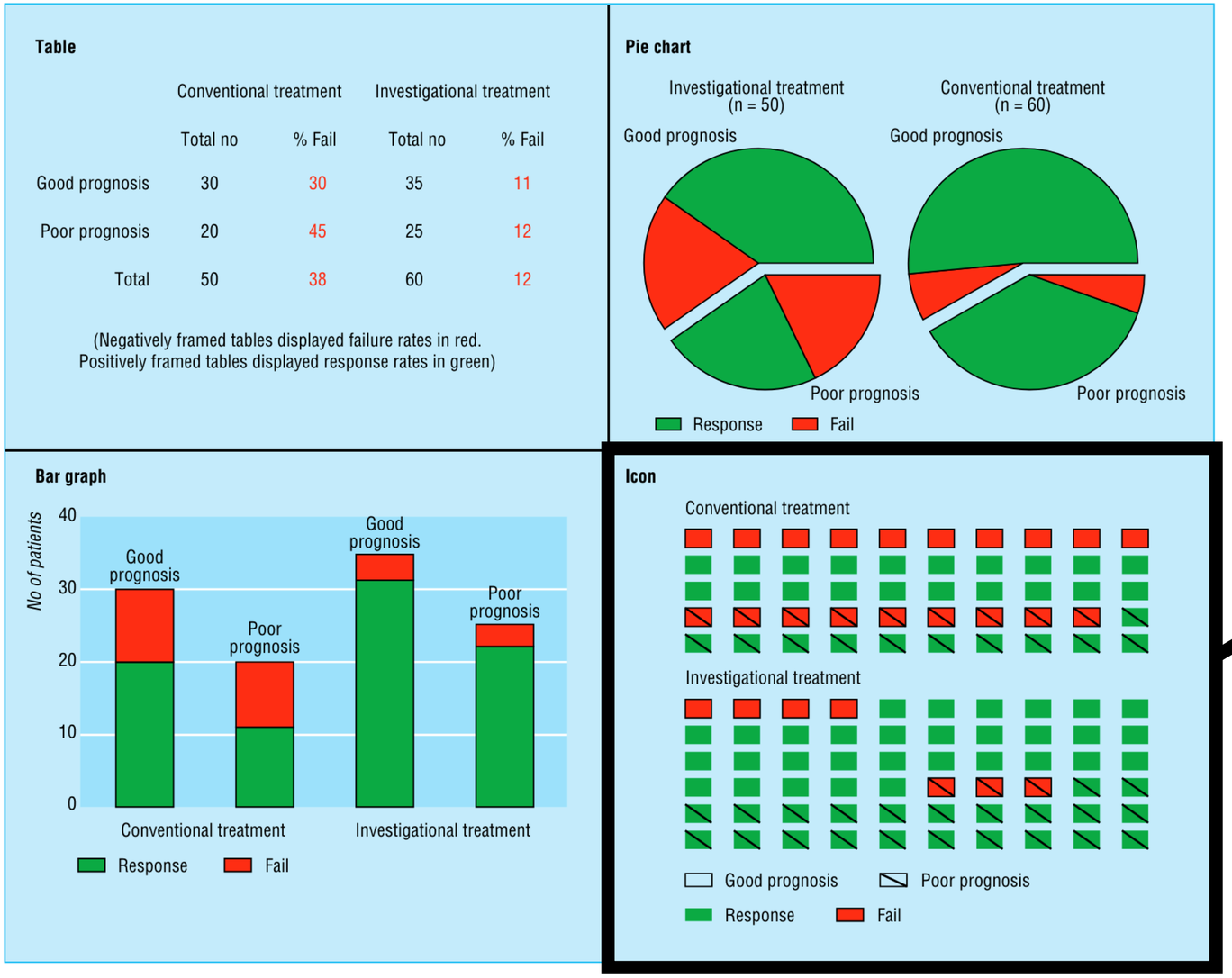
Don't take into account cognitive processing

- Study cognitive processes - how is data comprehended?
- Identify what might mislead and what might communicate effectively

Studying Cognitive Processing



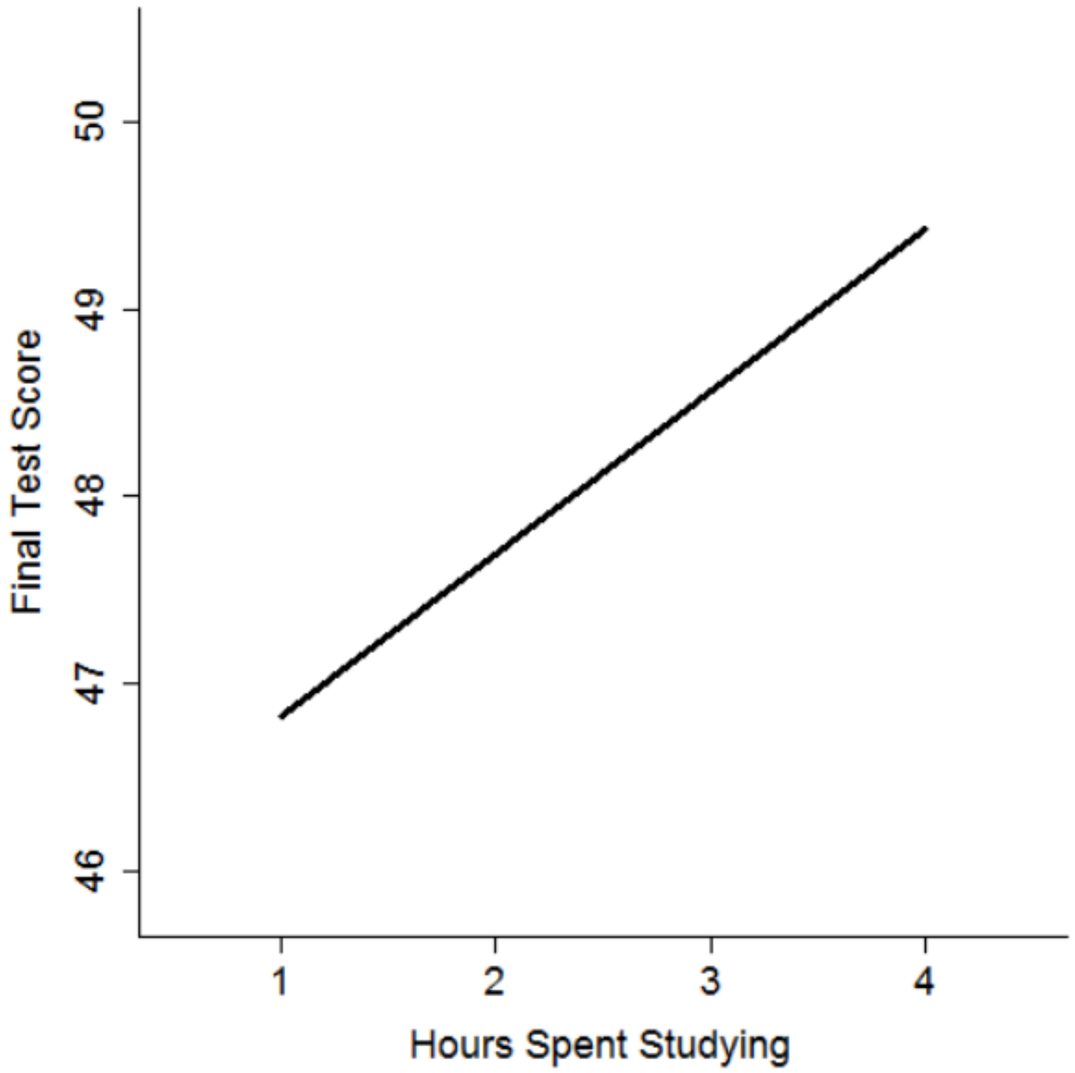
Why don’t we just ask people what they like?

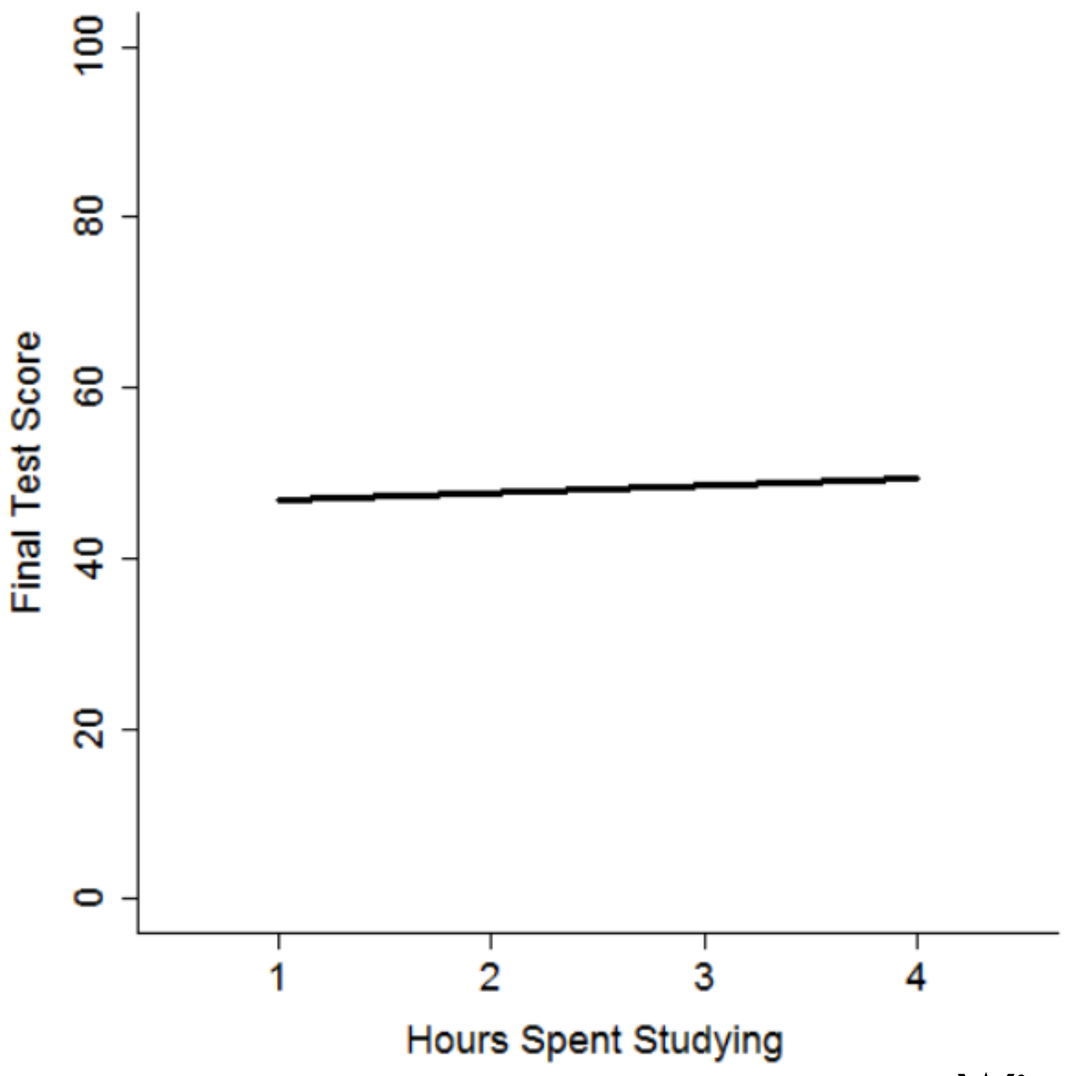


- Most accurate
- 0/34 participants preferred this method
- Eight expressed ‘considerable contempt’ 😤

Elting et al. (1999)

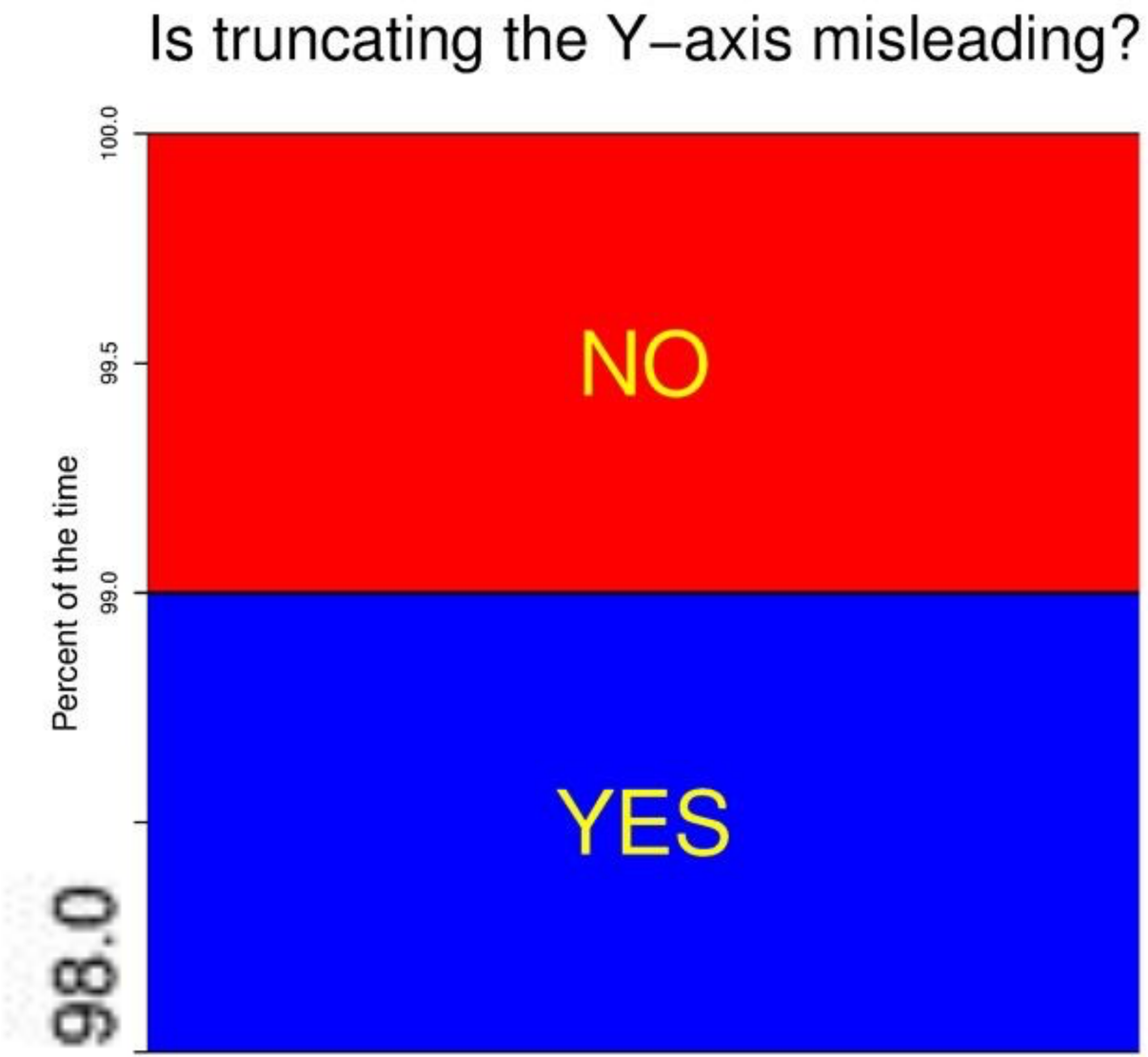
Design Choice: Axis Range



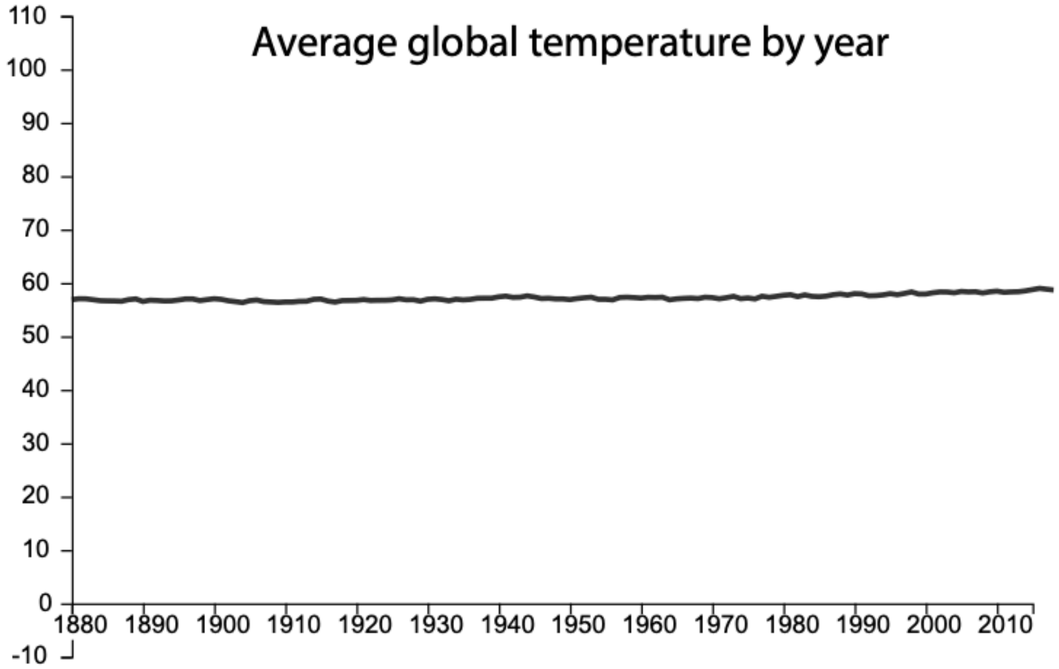


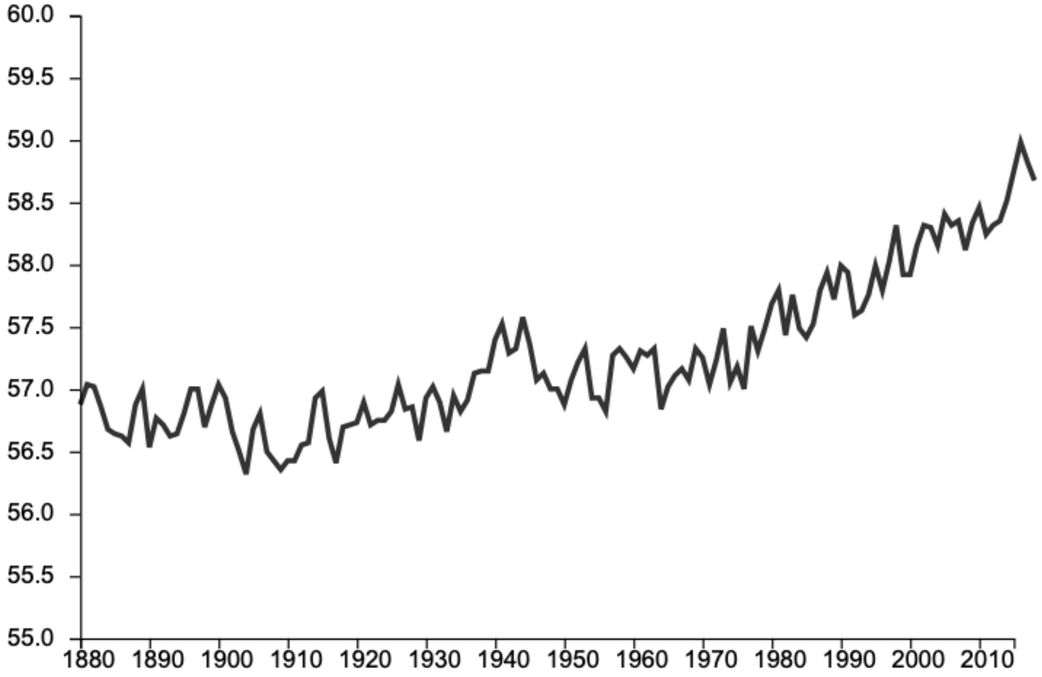
Witt (2019)

Design Choice: Axis Range



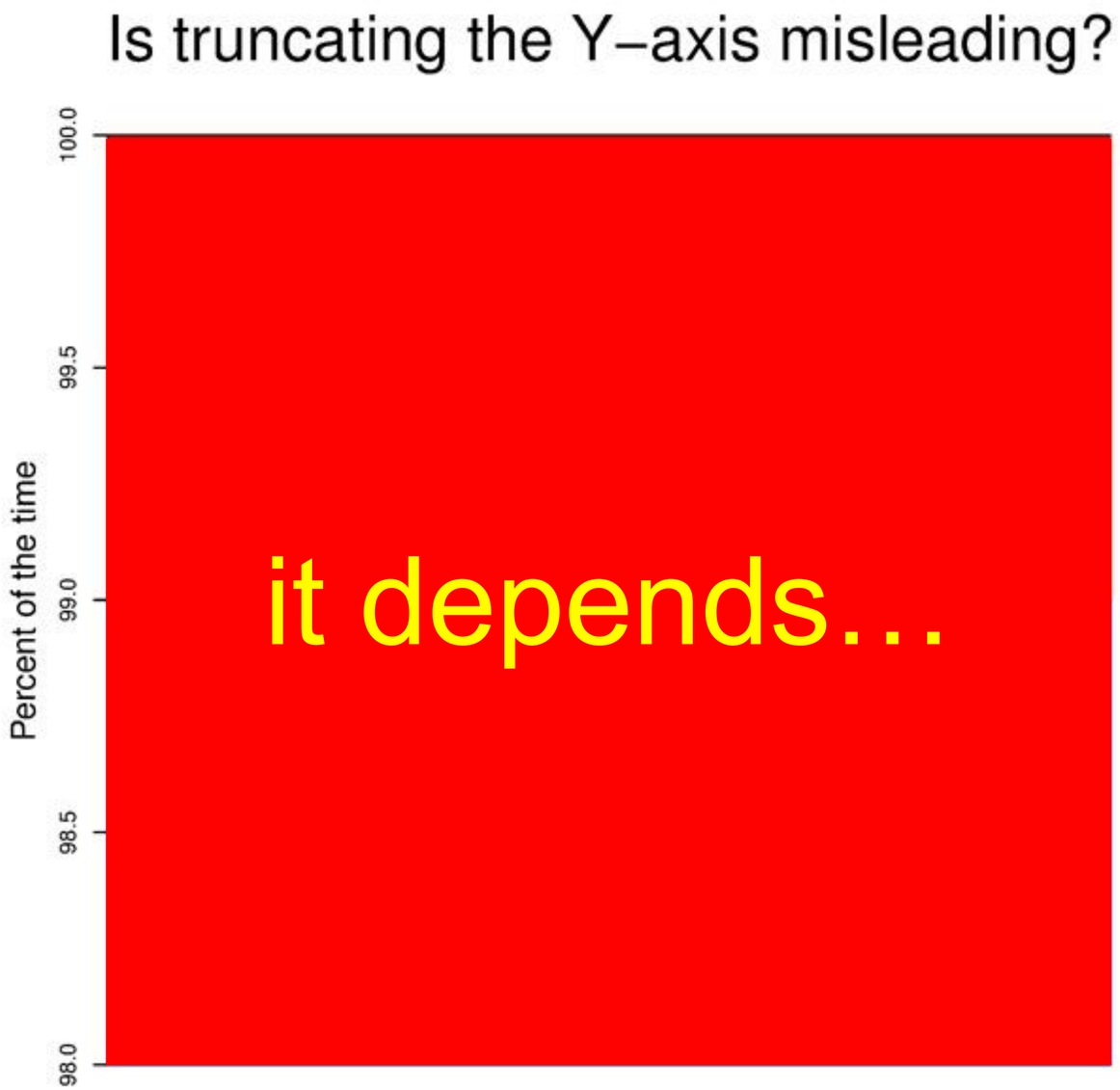
Design Choice: Axis Range





Correll et al. (2020)

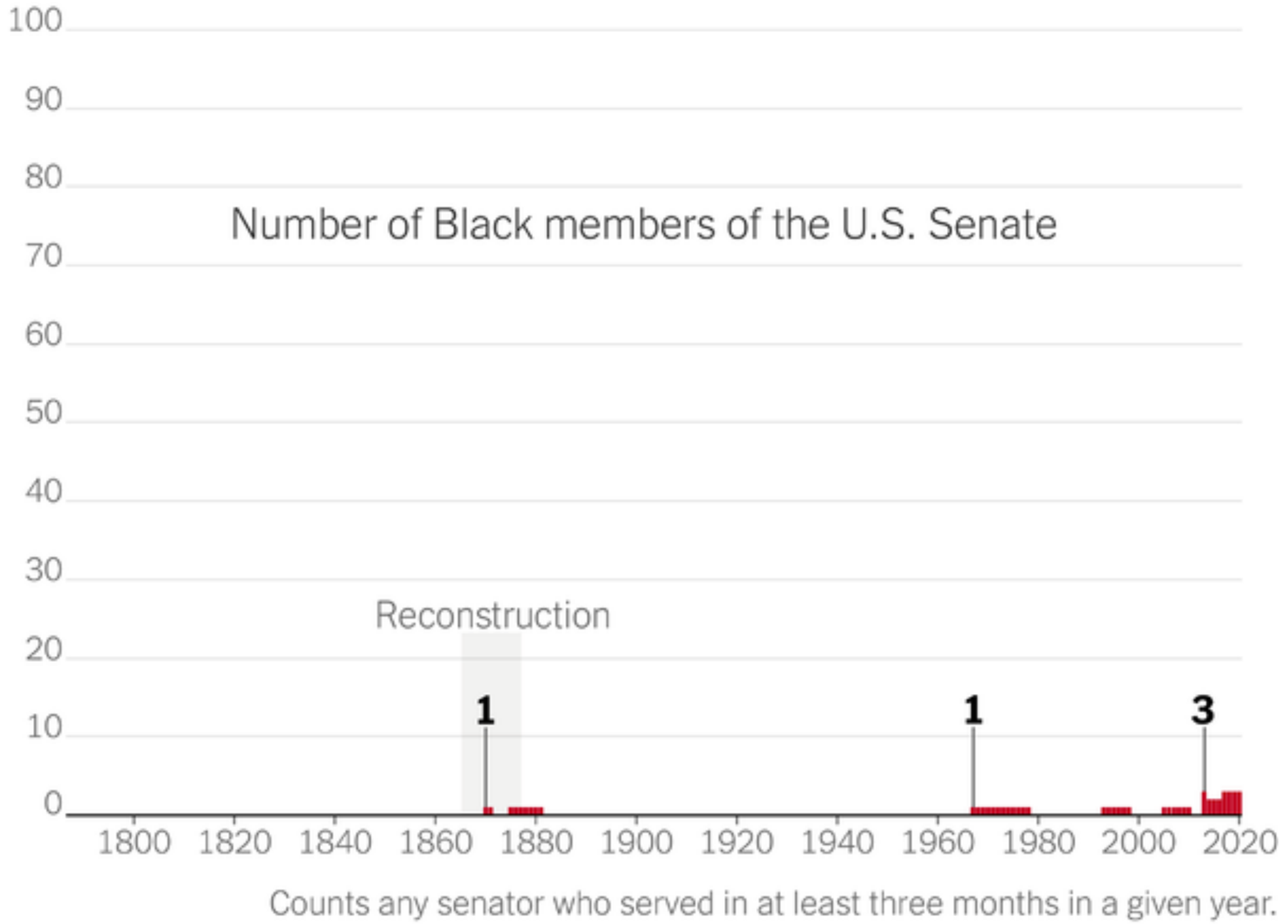
Design Choice: Axis Range



Design Choice: Axis Range

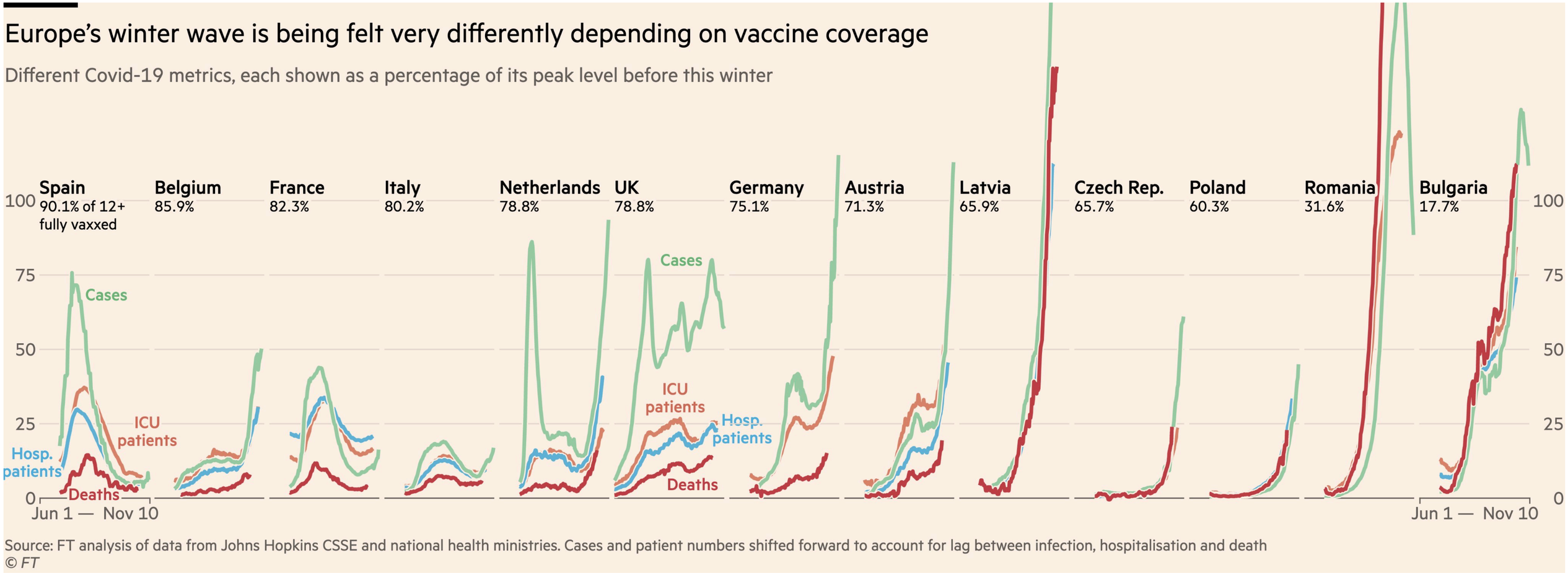
- Influences interpretations of the *magnitude of differences between values*
- Line charts and bar charts (Correll et al. 2020)
- Not eliminated by warnings (Yang et al. 2021)

Changing the Axes to communicate magnitude



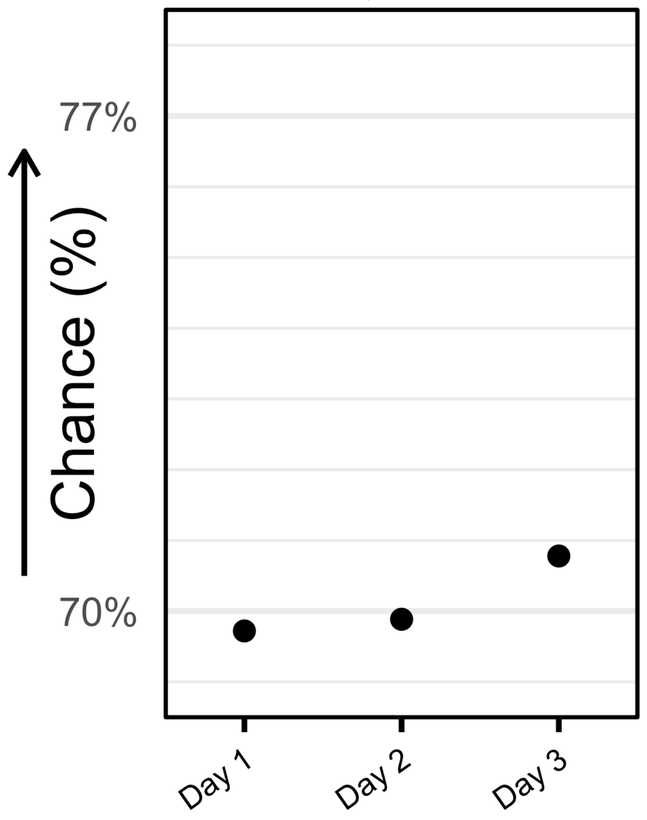
The New York Times

Not just small magnitudes...

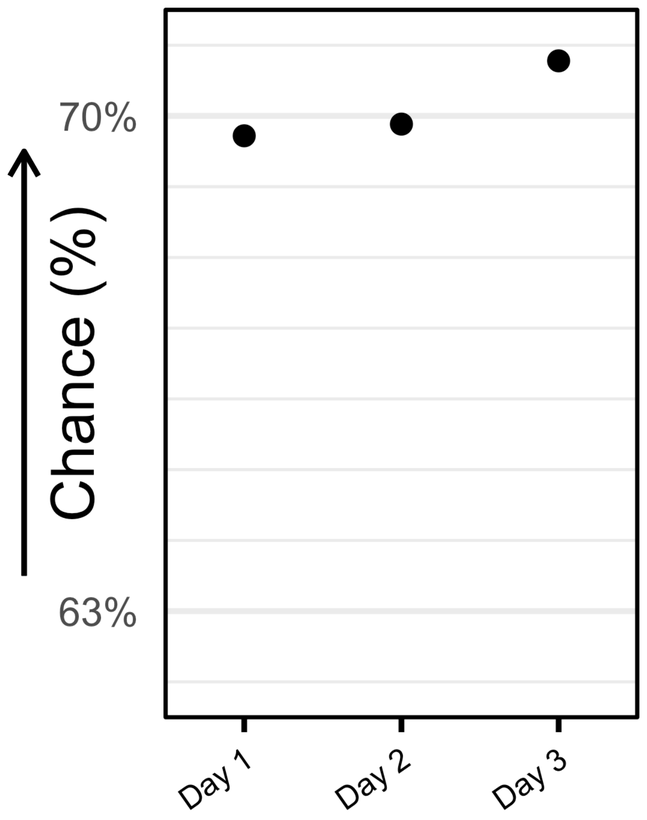


Financial Times

Same dataset in both versions of each graph



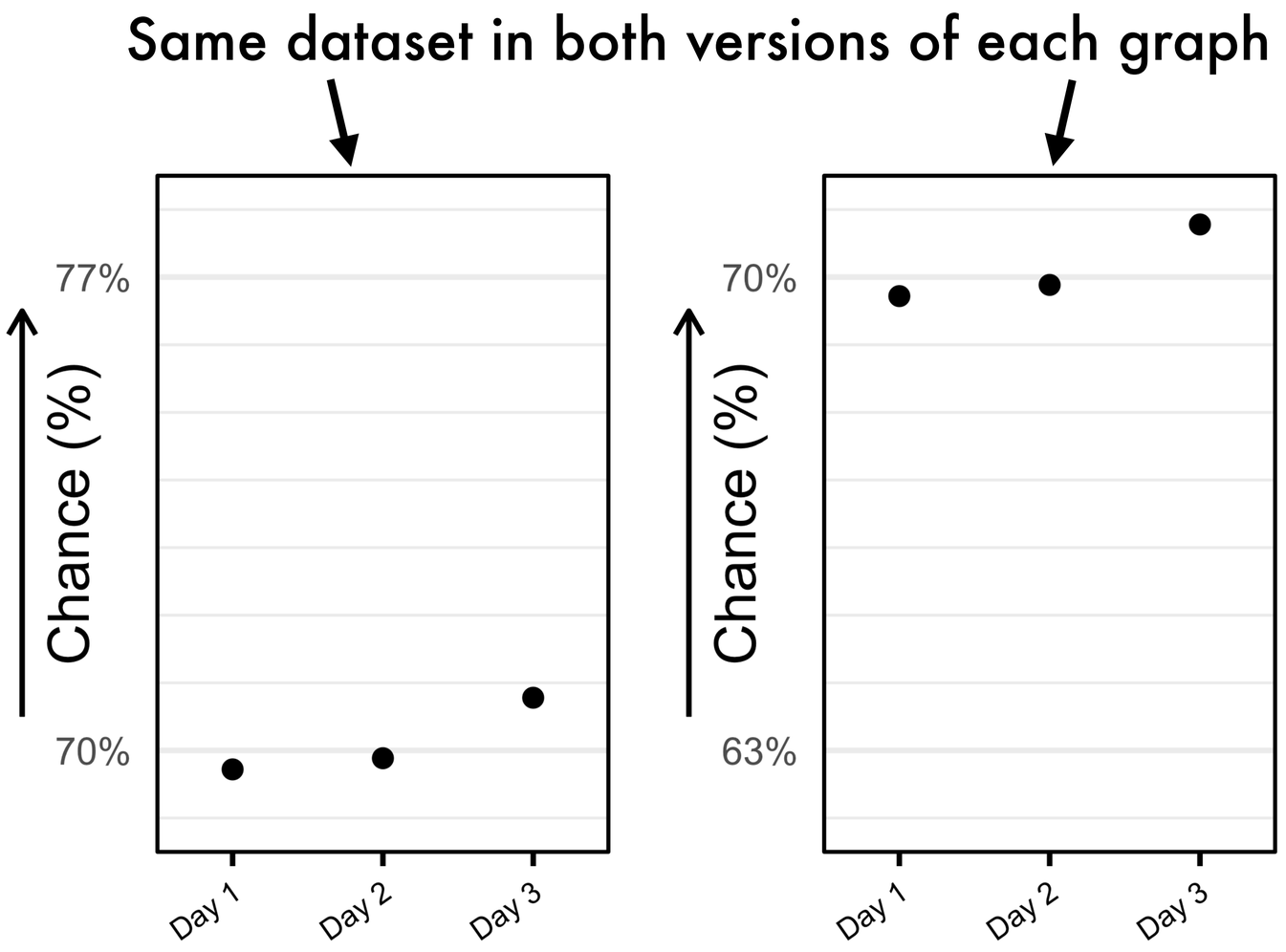
Data points with **low physical position**



Data points with **high physical position**

Experiment 1

- Q: **How are interpretations of magnitude affected by axis range?**
- Two versions of each graph (40 experimental trials)
- 150 participants - prolific.co
- Risk scenarios



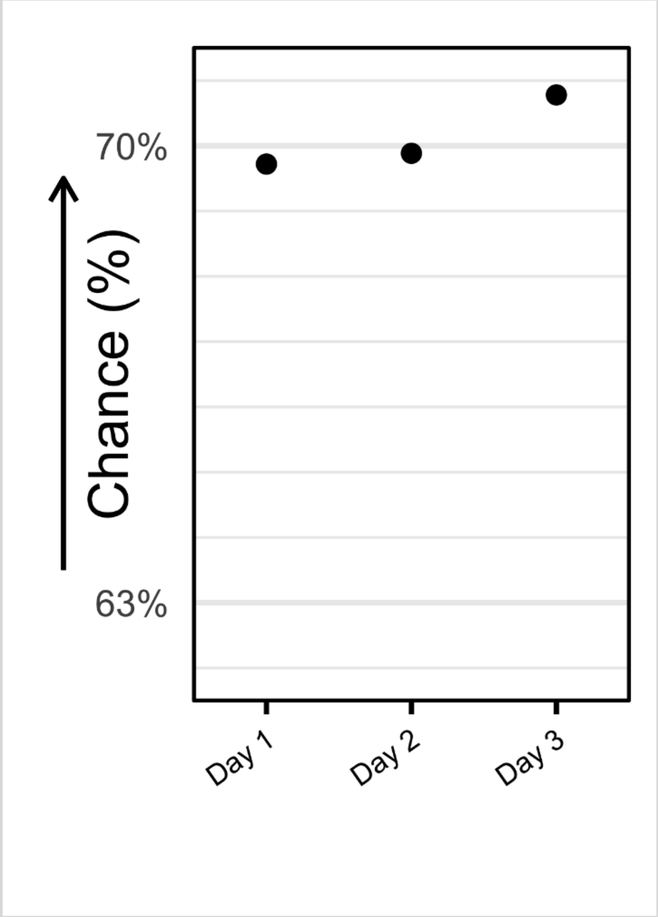
Data points with **low physical position**

Data points with **high physical position**

Heavy Rainfall Trial 15 of 60

You are going on a camping trip next week.

The graph below shows the chance of heavy rainfall for three randomly selected days next week.



If you camp on one of these days...

What is the chance you experience heavy rainfall?

Very unlikely **Very likely**

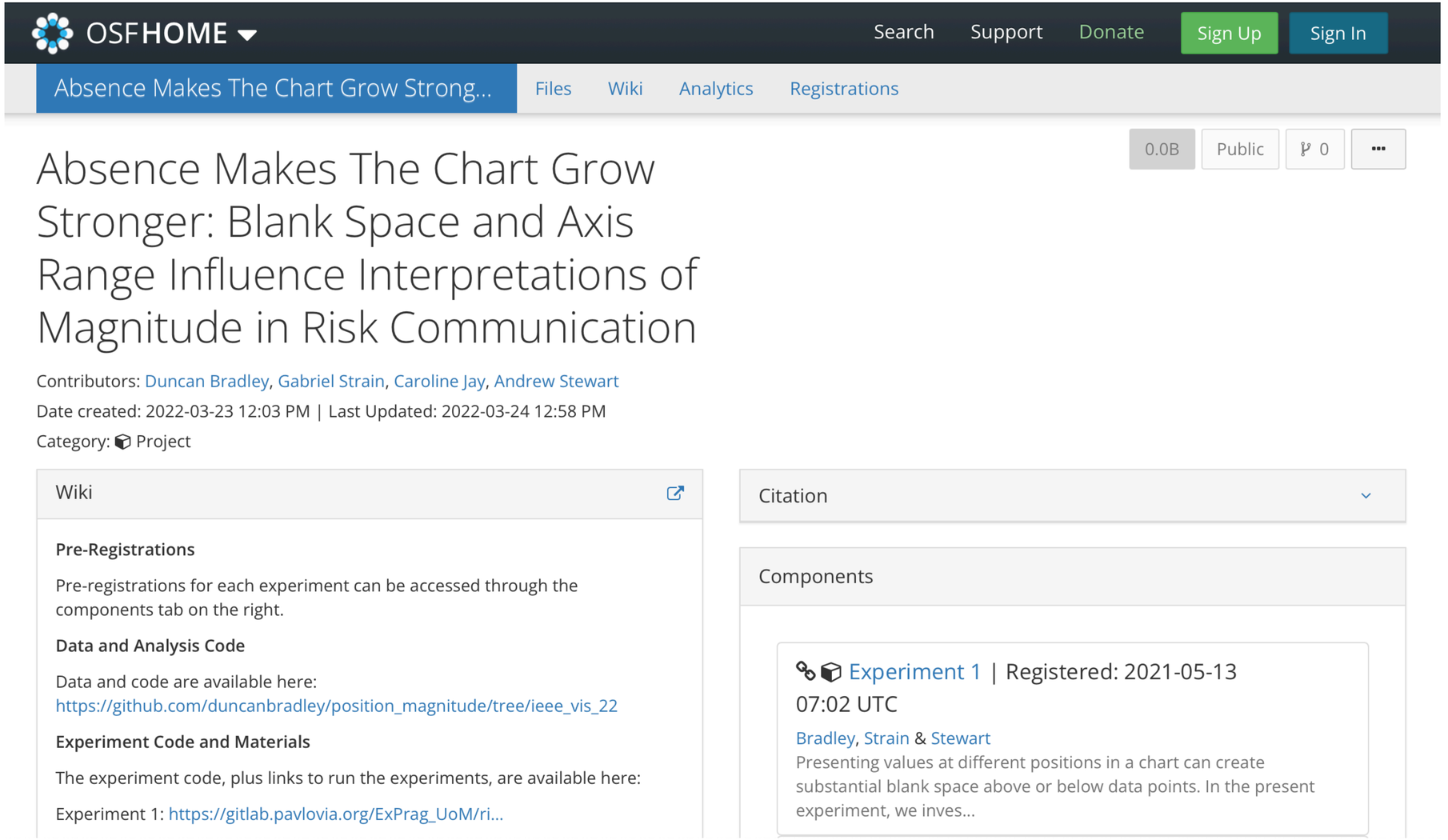
What is the severity of consequences if you experience heavy rainfall?

Very mild **Very severe**

Press the spacebar to continue when you have made your response.

Hypothesis

- Pre-registered hypotheses and analysis plan: https://osf.io/qn46s/



Hypothesis

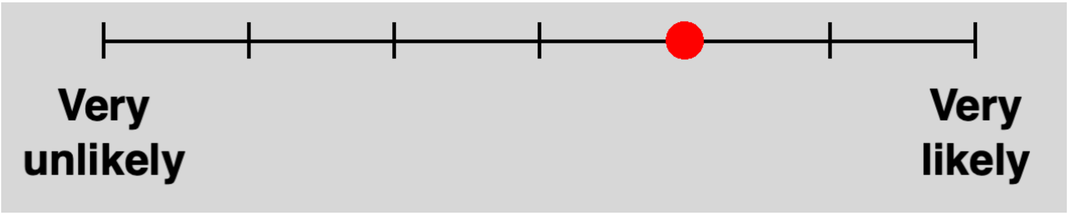
- Pre-registered hypotheses and analysis plan: https://osf.io/qn46s/

Hypothesis

The same data point can be located at different vertical positions in different data visualisations. Hypothesis: When data points are positioned higher on the y-axis, judgements of the likelihood and/or the severity of negative outcomes associated with the presented information will increase, compared to when data points are positioned lower on the y-axis.

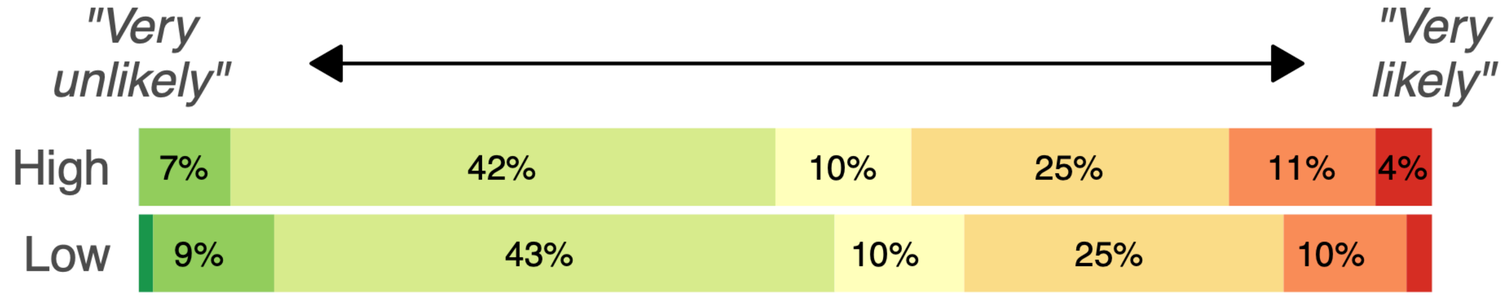
Analyses

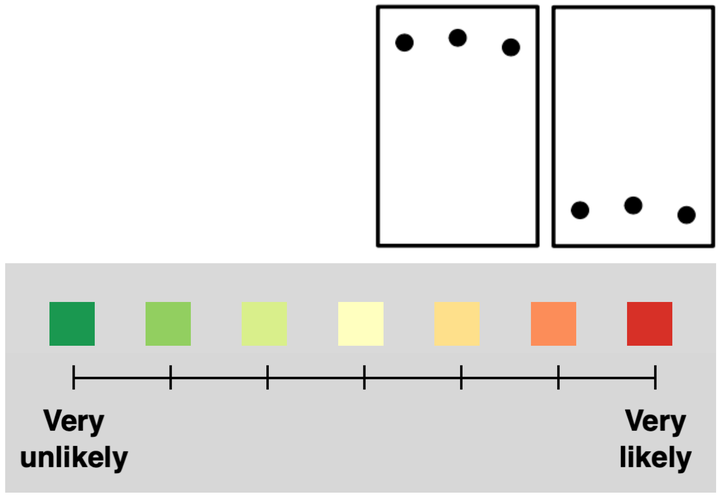
- Cumulative link mixed-effects model analysis in R using the `ordinal` package (Christensen, 2019), for Likert scale data (Kruschke & Liddell, 2018)



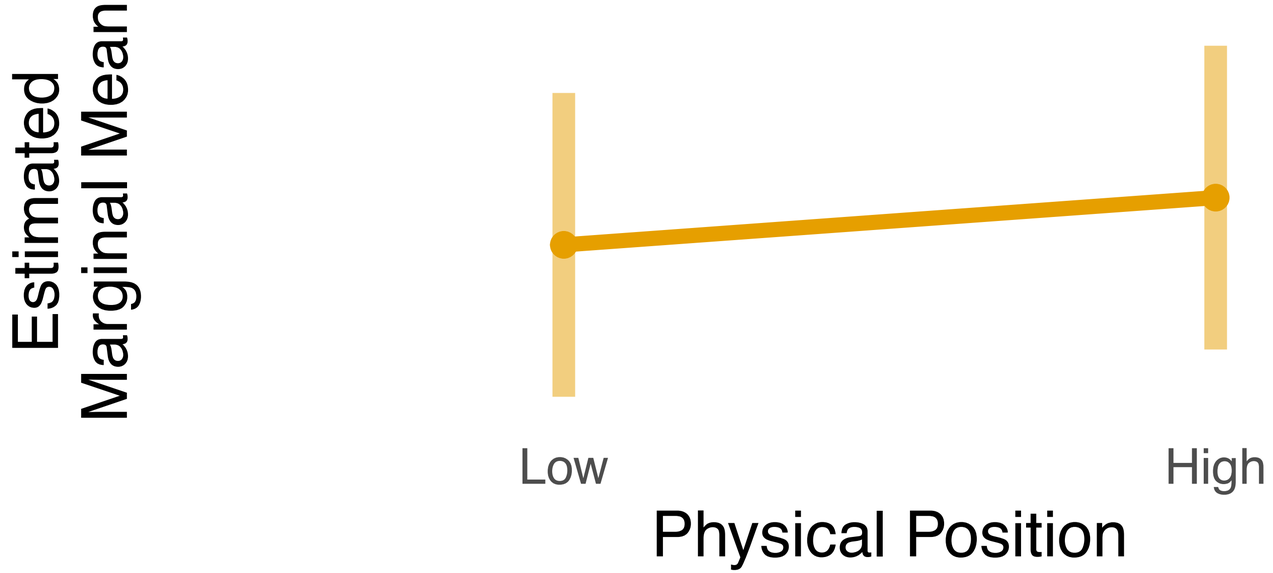
- `buildmer` package (Voeten, 2022):
  - identified the most complex random effects structures that would successfully converge
  - subsequently removed terms which did not contribute substantially to explaining variance in ratings

Experiment 1:
Ratings of Data Points' Magnitudes





Experiment 1:
Ratings of Data Points' Magnitudes (Modeled)



ANOVA: $\chi^2(1) = 74.21$, $p < .001$

Physical Position: $z = 8.57$, $p < .001$

What's Driving This Effect?

Magnitude

Absolute Position

'Up is more'

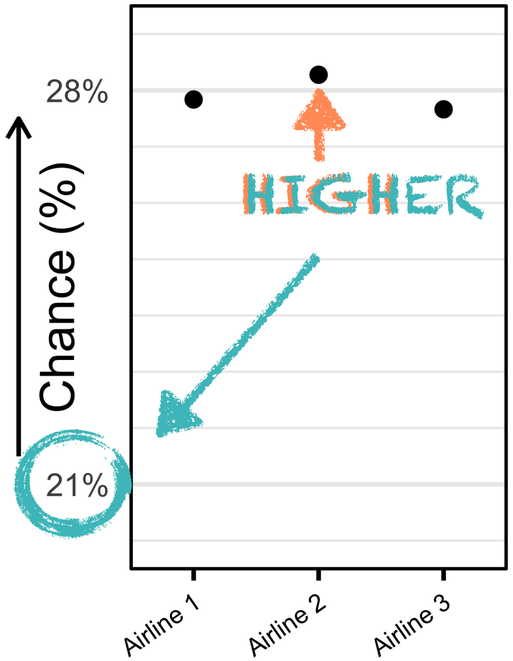
- Position in physical space

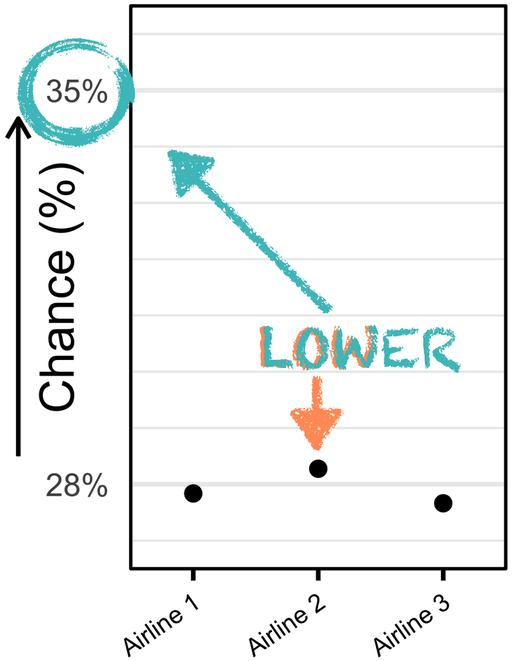
Relative Position

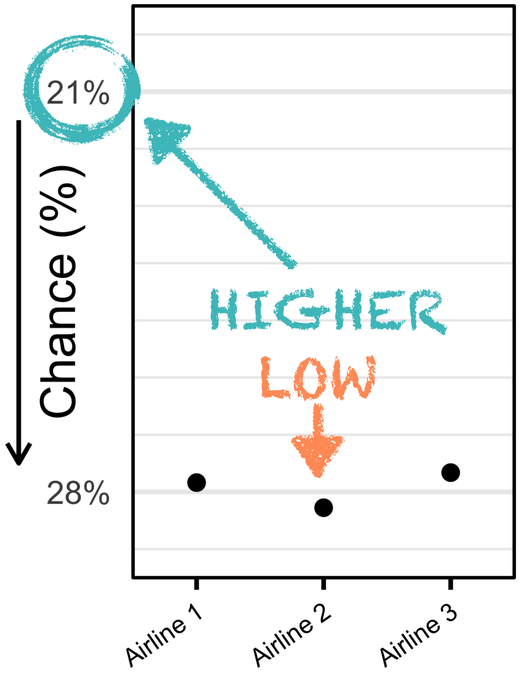
Axis Range

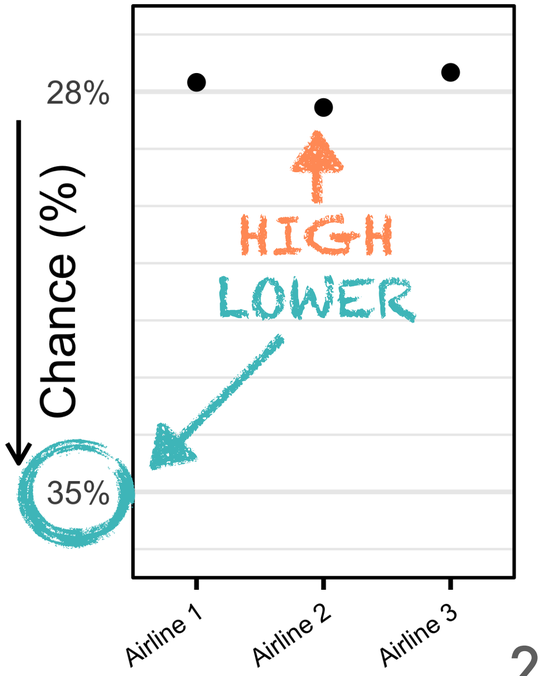
- Position relative to other plausible (but absent) values

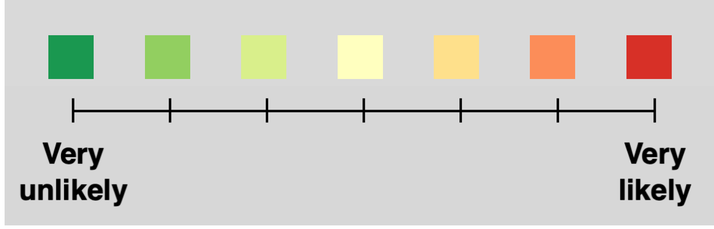
- **Experiment 2:** physical position (high/low) and axis orientation (conventional/inverted) - 2x2 design
- 120 participants; 24 experiment trials
- Absolute position = no interaction
  Relative position = crossover interaction









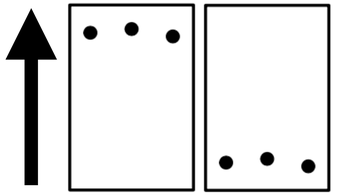


Experiment 2:
Ratings of Data Points' Magnitudes

"Very unlikely" ⟷ *"Very likely"*

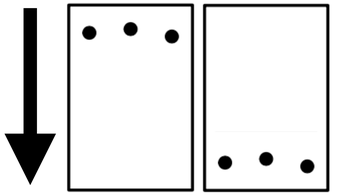
Conventional

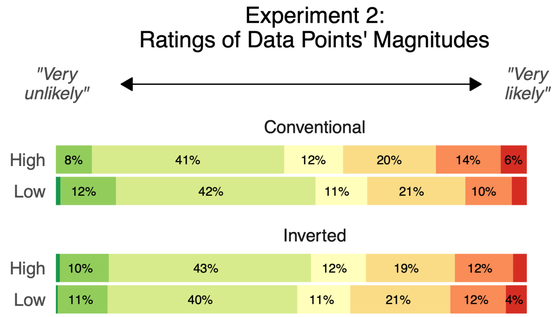
	8%	41%	12%	20%	14%	6%
High						
Low	12%	42%	11%	21%	10%	



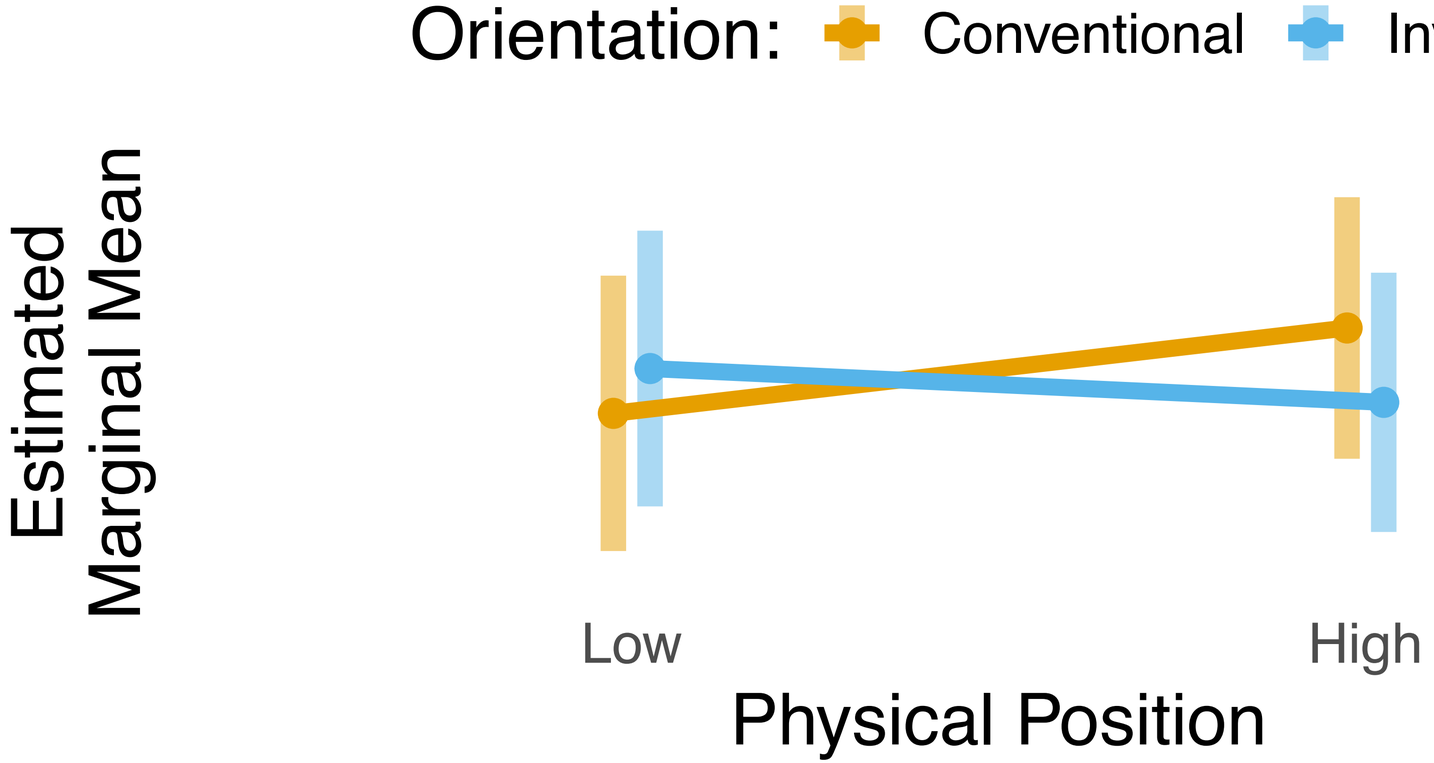
Inverted

	10%	43%	12%	19%	12%	
High						
Low	11%	40%	11%	21%	12%	4%





Experiment 2:
Interaction in Ratings of Data Points' Magnitudes (Modeled)



ANOVA: $\chi^2(1) = 8.22$, $p = .004$

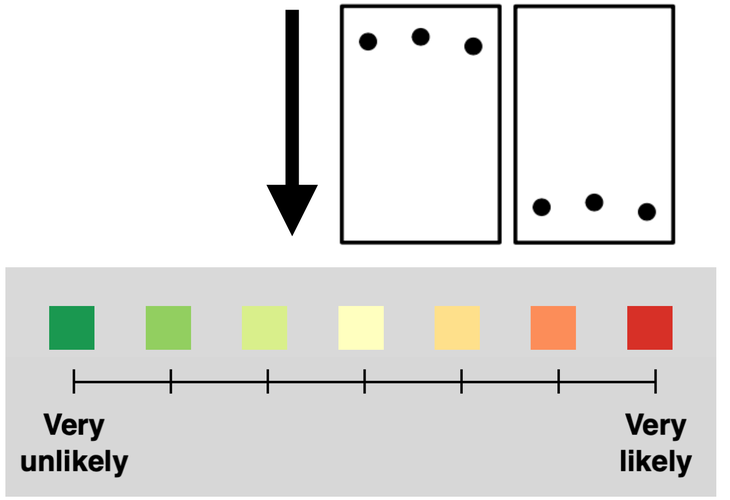
Interaction: $z = 2.91$, $p = .004$

Pairwise:

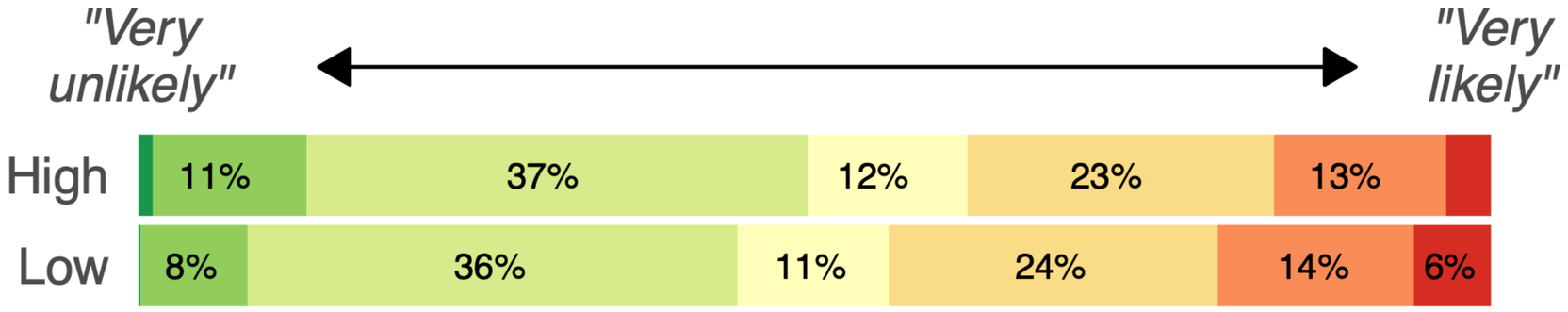
Conventional: $z = 3.56$, $p = .001$

Inverted: $z = 1.39$, $p = .512$

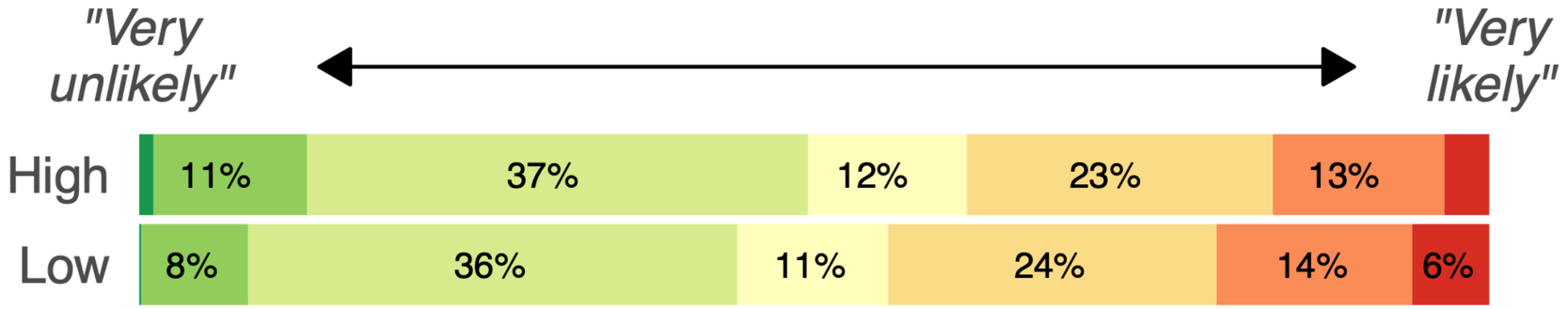
Experiment 2:
Inverted

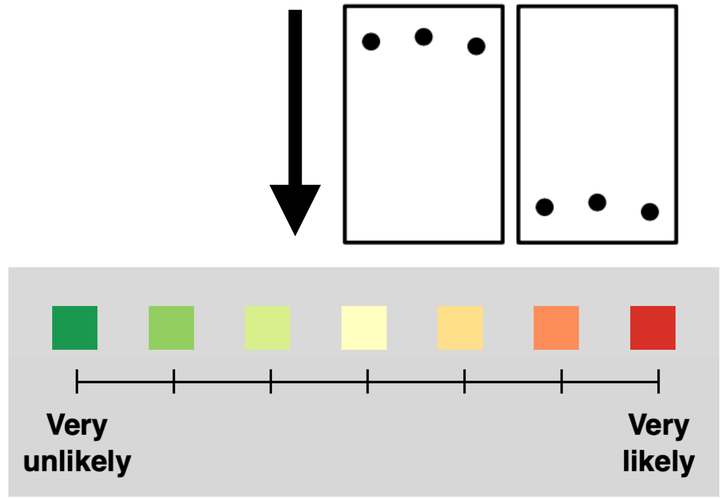


Experiment 3:
Ratings of Data Points' Magnitudes

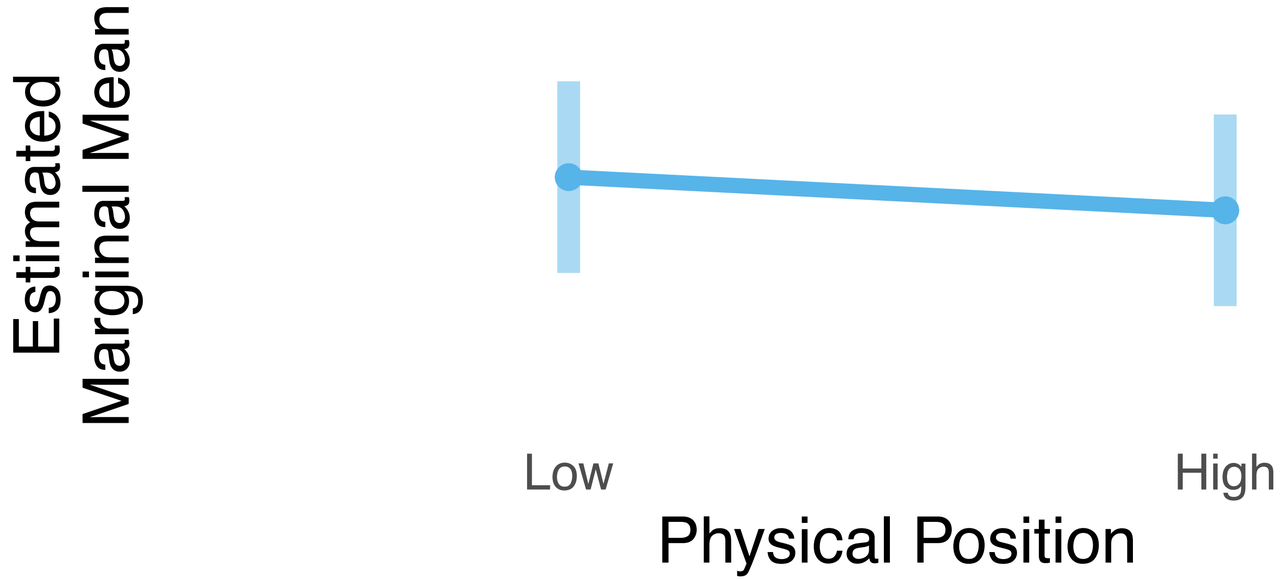


Experiment 3:
Ratings of Data Points' Magnitudes





Experiment 3:
Ratings of Data Points' Magnitudes (Modeled)



ANOVA: $\chi^2(1) = 46.45$, $p < .001$

Physical Position: $z = 6.80$, $p < .001$

What's Driving This Effect?

Magnitude

Absolute Position

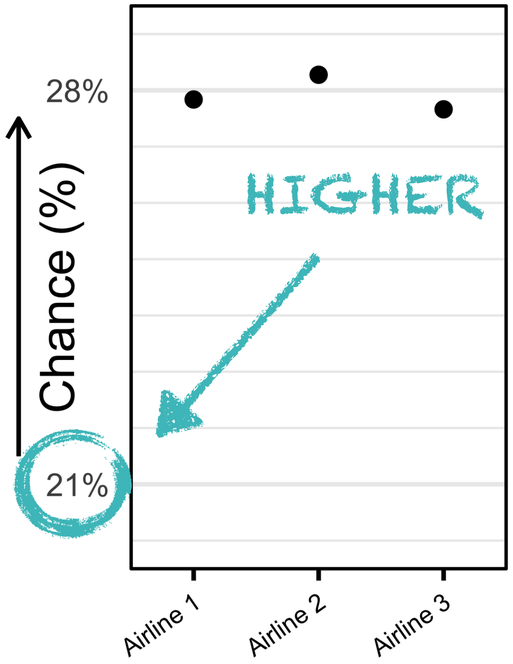
'Up is more'

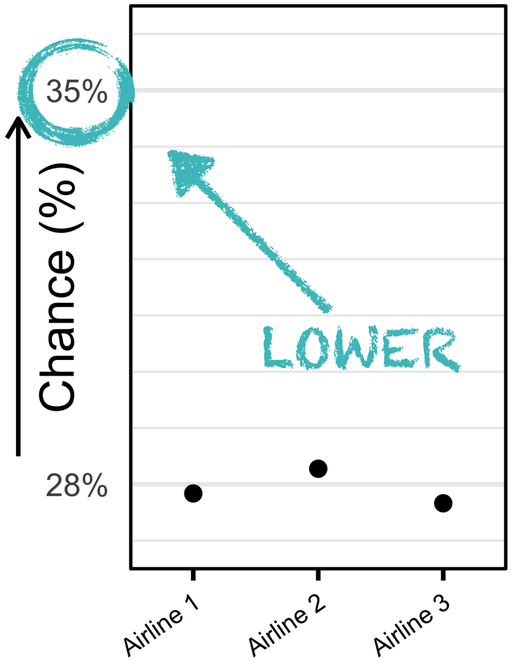
- Position in physical space

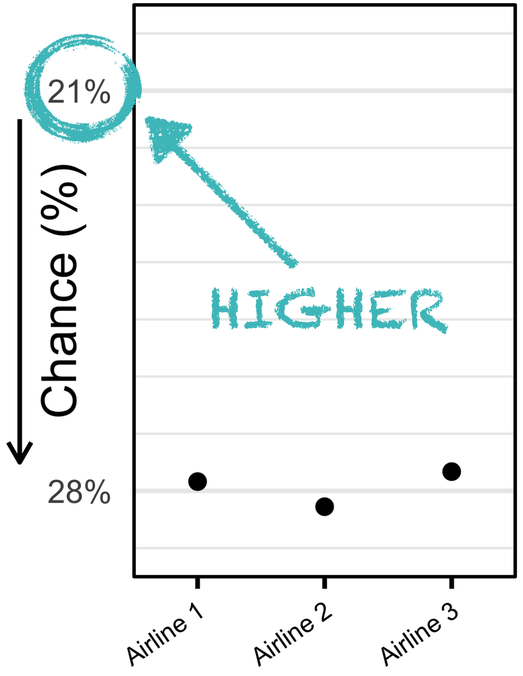
Relative Position

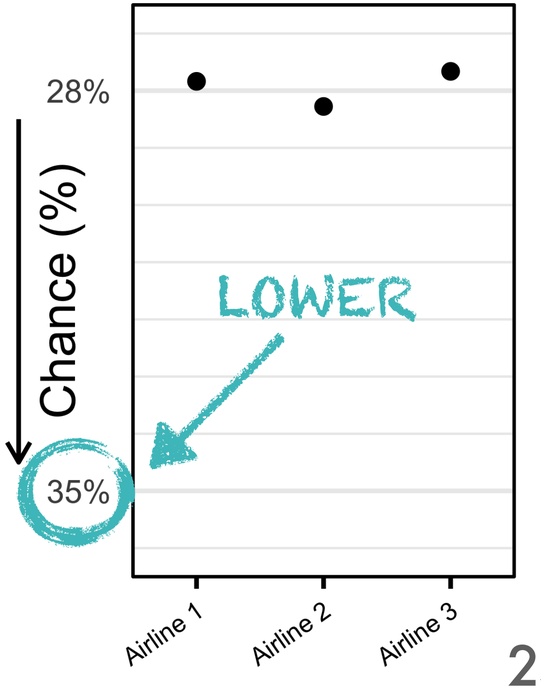
Axis Range

- Position relative to other plausible (but absent) values









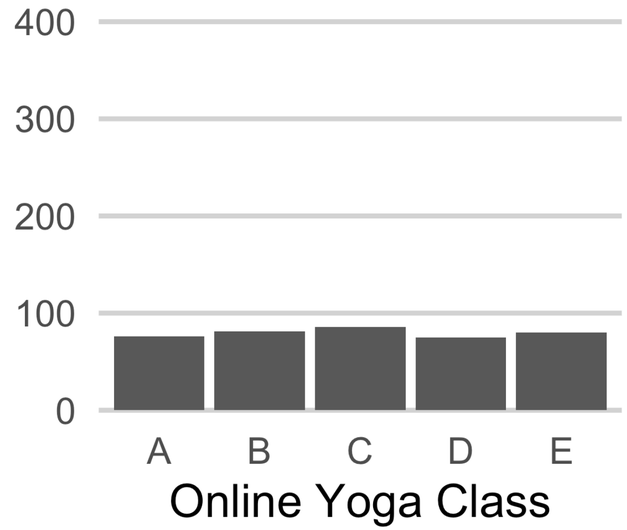
Extending This Finding: Bar Charts

- **Issues:**
  - Arbitrary axis limits
  - Risk scenarios only
- **Follow-up experiment:**
  - More realistic axis limits
  - Range of scenarios
  - 150 participants - prolific.co
  - 32 experimental trials

ggplot2 default

TRUNCATED VERSION

Custom upper limit



EXTENDED VERSION

Online Yoga Classes

A study was conducted to investigate the effectiveness of five online yoga classes. For each class, 400 people were surveyed.

The graph shows, for each online yoga class, the number of people who felt more flexible after the class.

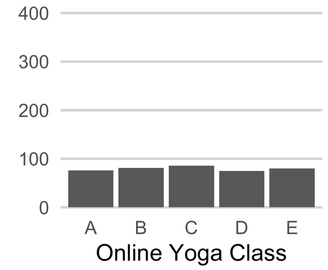
How effective were the yoga classes at improving flexibility?

Very ineffective — Very effective

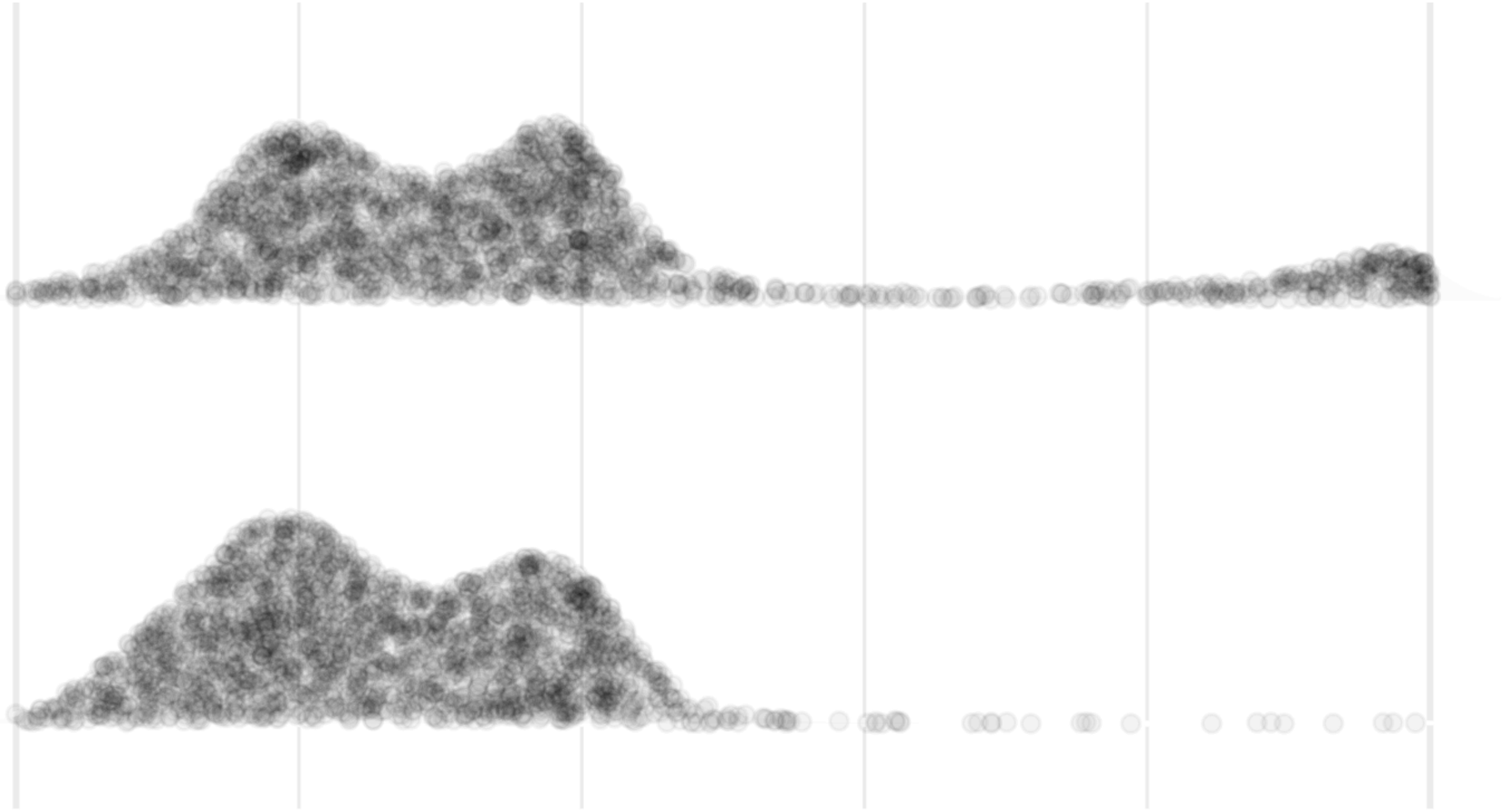
Press the spacebar to continue when you have made your response.

Distribution of Magnitude Ratings

Default Range



Extended Range

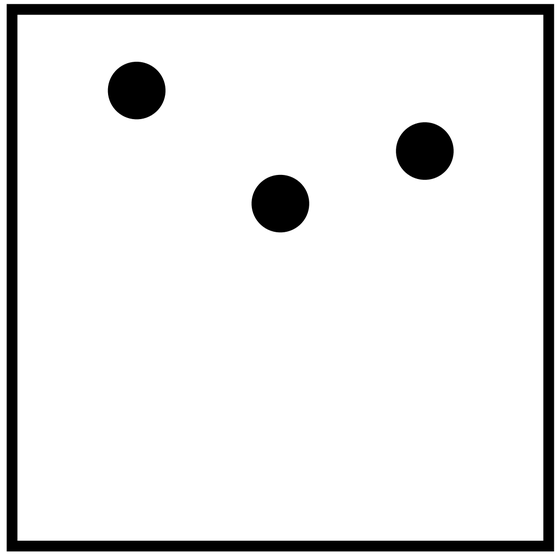


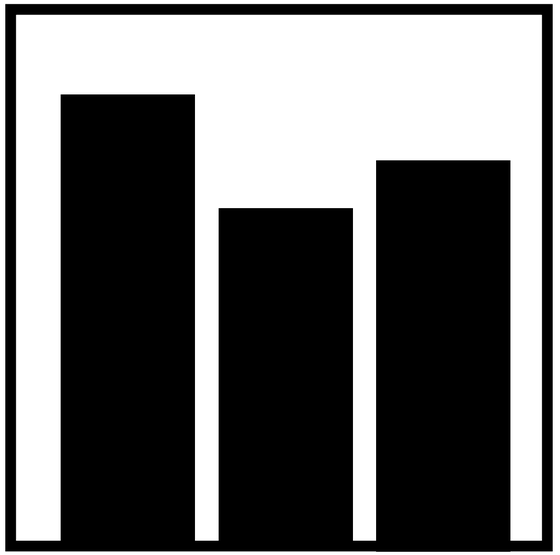
Very low magnitude

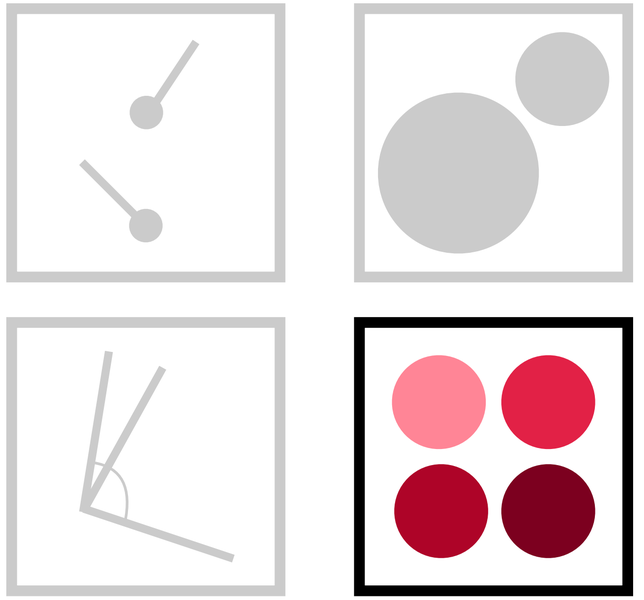
ANOVA: $\chi 2(1) = 39.66$, $p < .001$

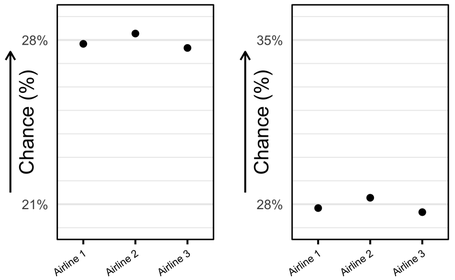
Very high magnitude

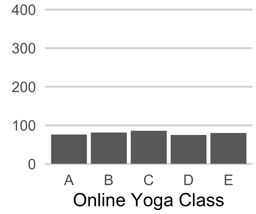
Encoding Types



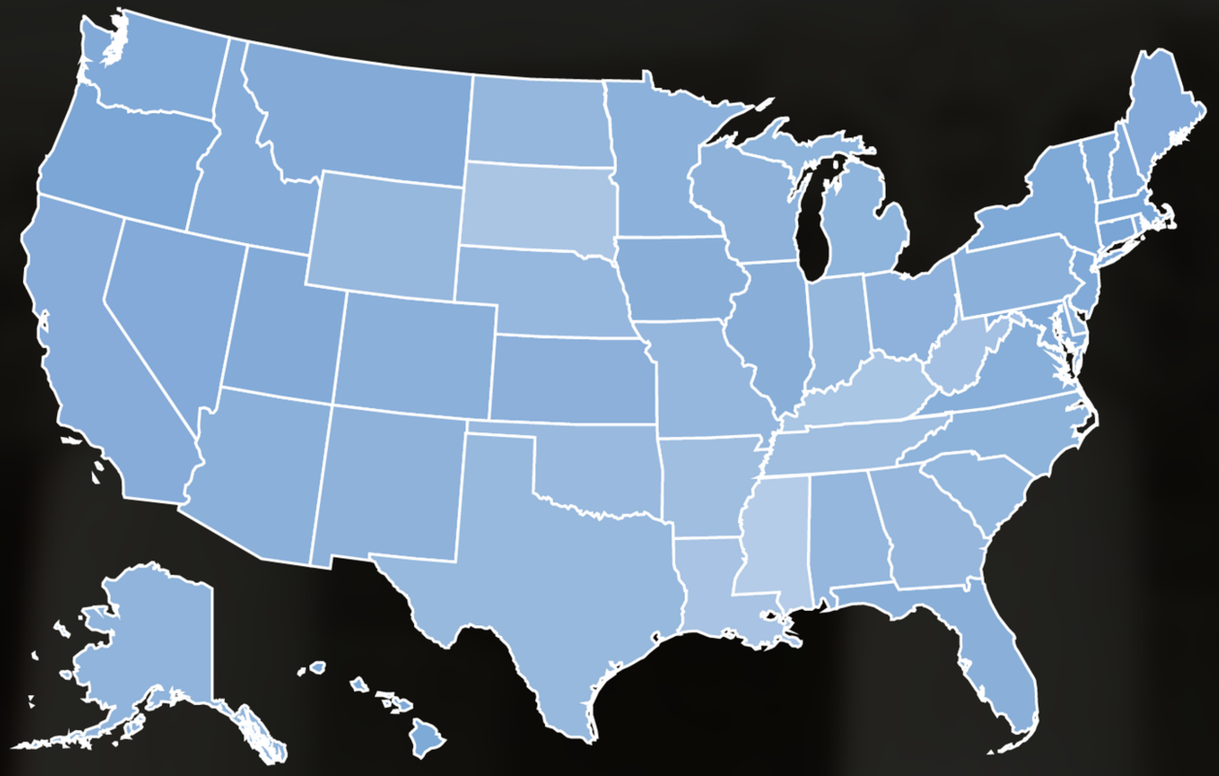








There is not a single state where support for a federal ban on abortion has more than 30% support among the public.

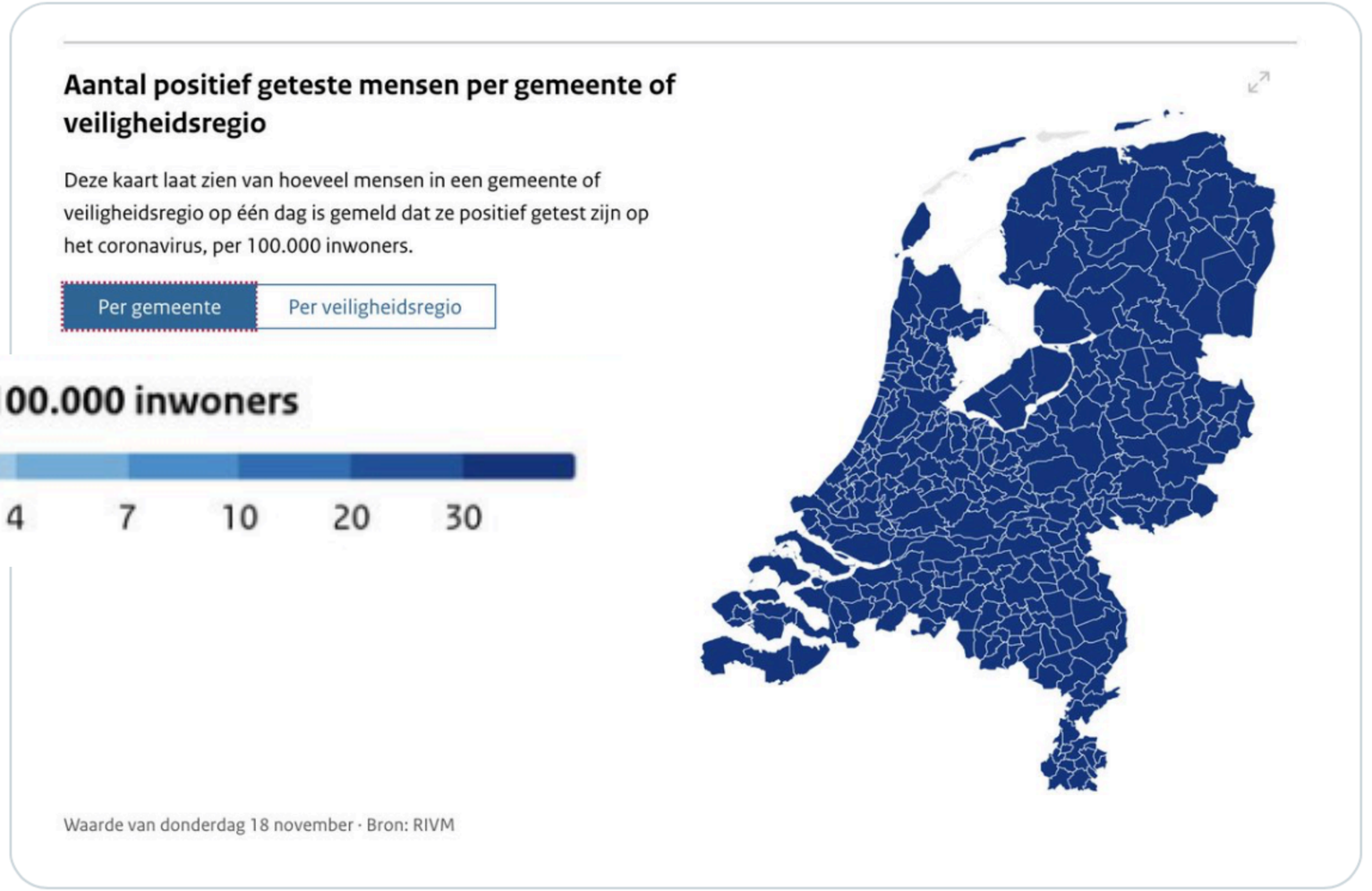


Analysis of the 2020 Cooperative Congressional Election Study

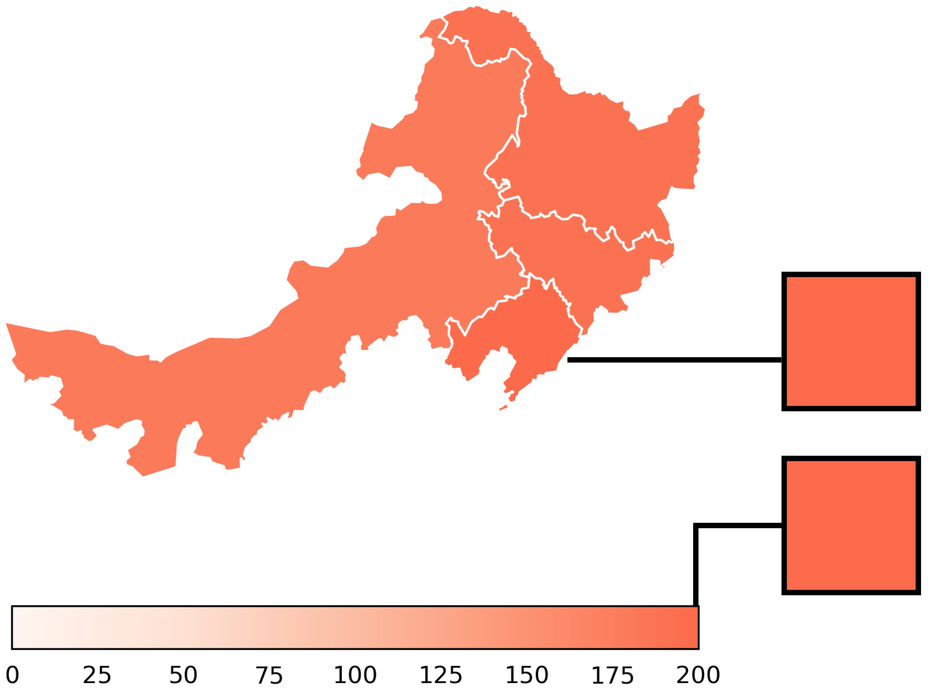
DATA FOR *PROGRESS*

alexandre afonso
@alexandreafonso

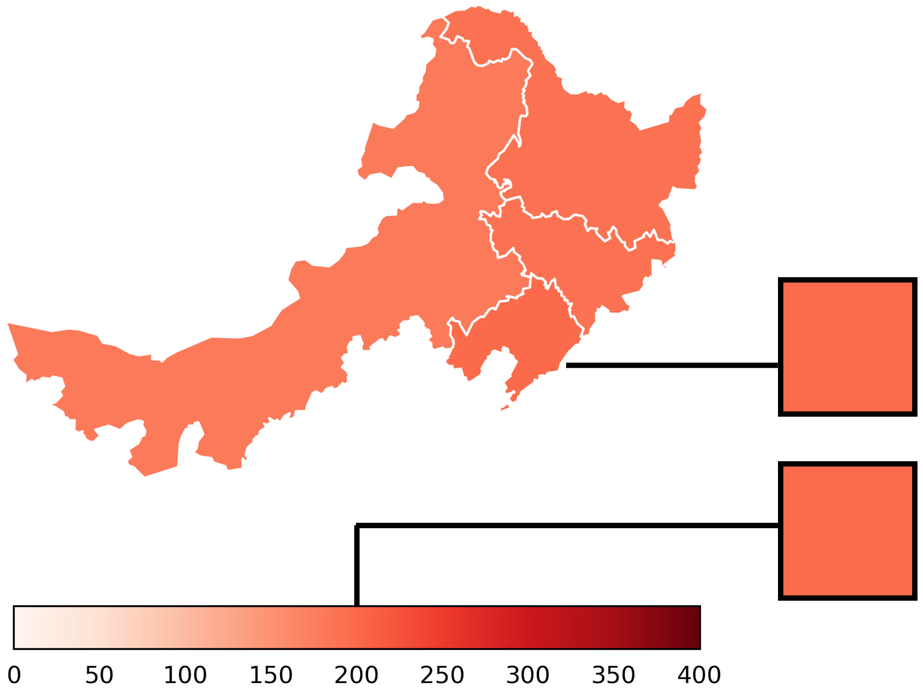
The colour code of the Dutch covid map goes from 0 to 30 positive tests per 100'000 inhabitants.
The current average rate is 135.



Experiment: Interpreting magnitude using colour legends



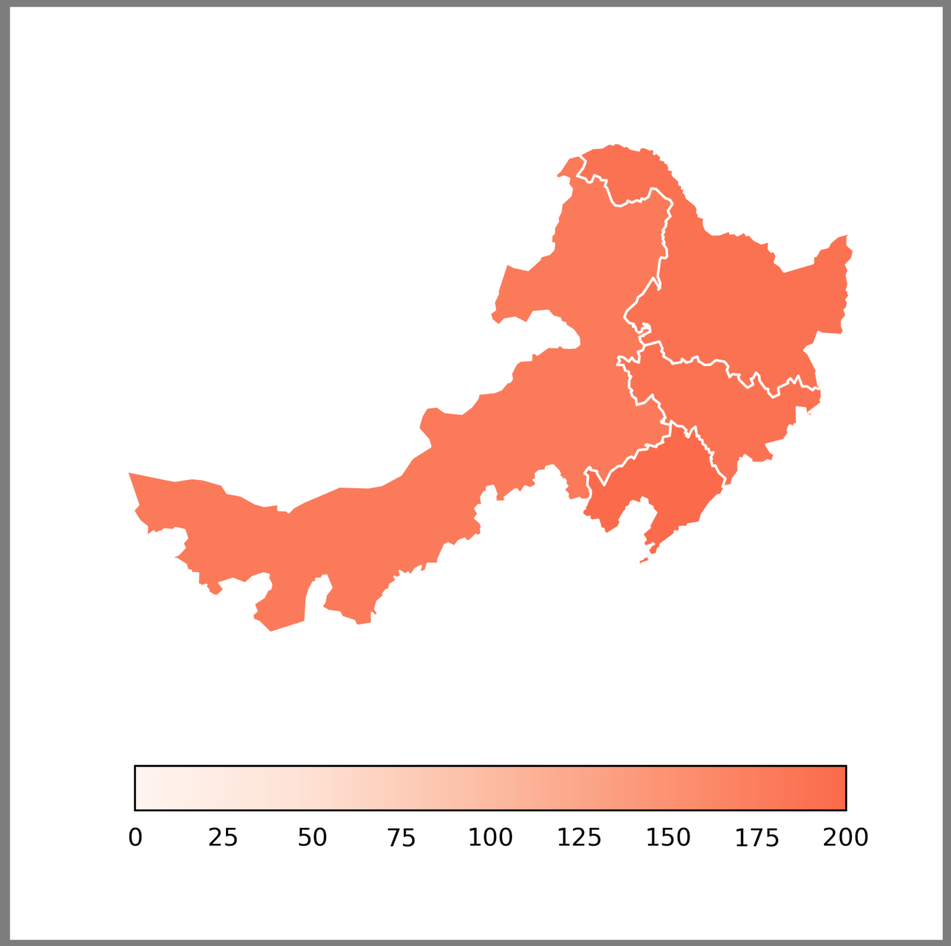
TRUNCATED VERSION



EXTENDED VERSION

- 160 participants - prolific.co
- 48 experimental trials
- Scenario: pollution data

This map shows the levels of a certain type of pollution, in four geographic regions:

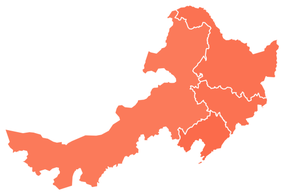


How urgently should pollution levels in these regions be addressed?

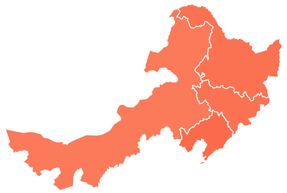
Not very urgently **Very urgently**

Press the spacebar to continue when you have made your response.

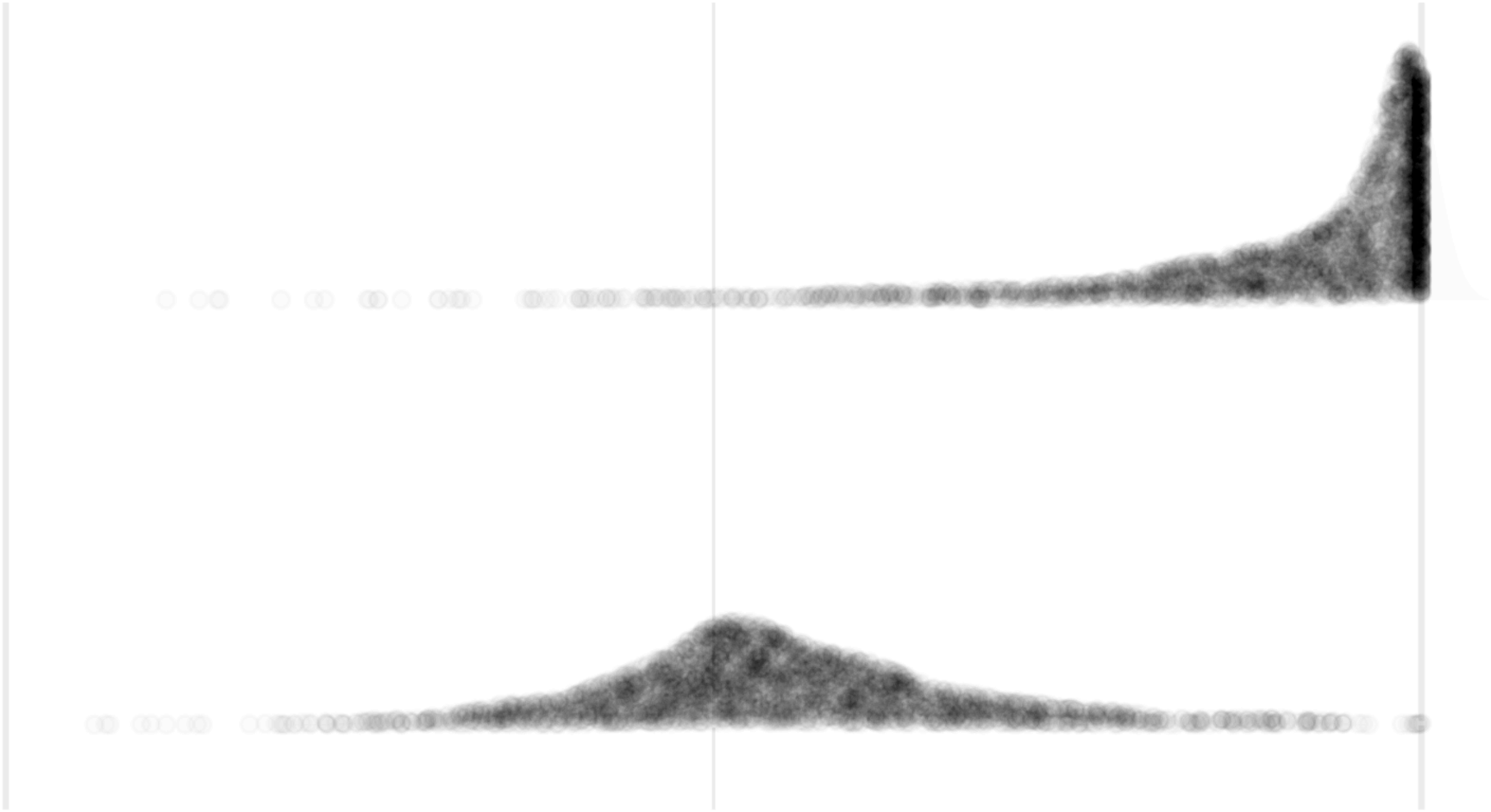
Distribution of Urgency Ratings



Truncated Range



Extended Range



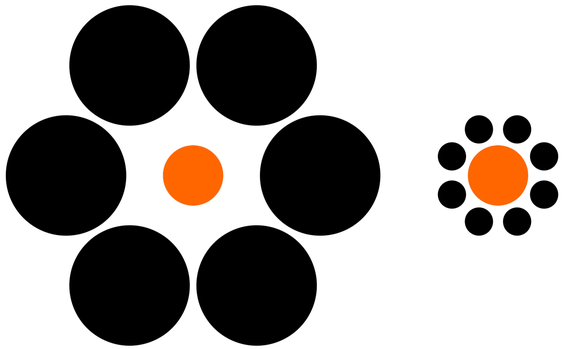
"Not very urgently"

"Very urgently"

ANOVA: $\chi 2(1) = 225.41$, $p < .001$

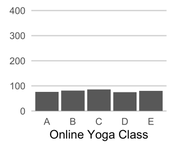
Seems Familiar?

- Framing effects (Tversky & Kahneman, 1974)
- Influence of surrounding context
- Vision: Ebbinghaus Illusion

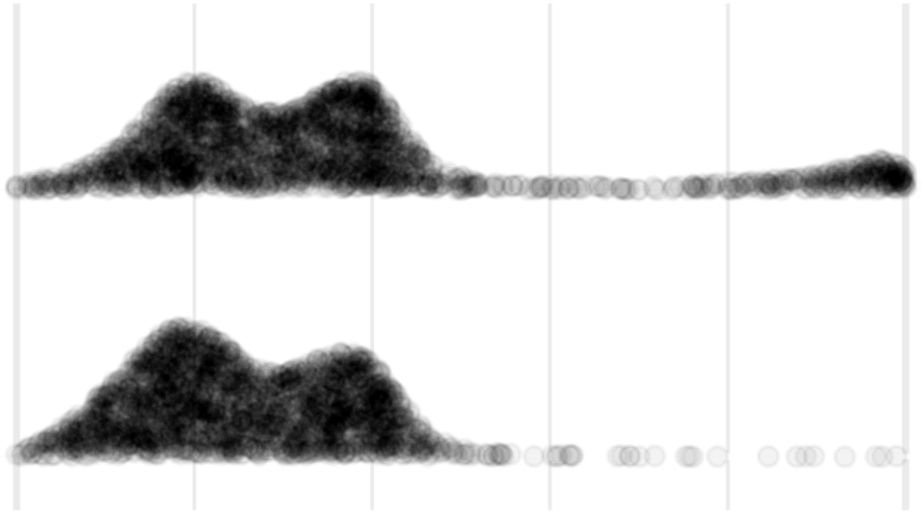


- Language: *"Almost"* vs. *"Only"* (Lundquist & Jarvella, 1994)
- What other cognitive biases might affect interpretation of information presented in data visualisations?

Default Range



Extended Range

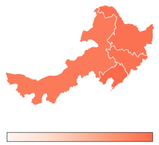


"out of 400..."

Very low magnitude

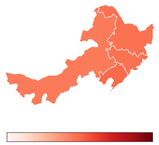
Very high magnitude

?



Truncated Range

" "



Extended Range

"Not very urgently"

"Very urgently"

How much does axis range affect judgements of magnitude?

Possible explanation: awareness of denominator

Key Points 🔑

- Different displays of the same data can provoke different interpretations
- Studying cognitive processing provides insight into comprehension
- Inferences about magnitude informed by axis limits
- Judgements influenced by relative position of data points on axis
- But strength of association seems to vary

Questions? 🤔🧐

The University of Manchester

Co-authors:
Gabriel Strain
Harvey Schneider
Boshuo Zhang
Caroline Jay
Andrew Stewart

Economic and Social Research Council

This work was supported by the Economic and Social Research Council [Grant Number ES/P000665/1].

the british psychological society

This work was supported by a Postgraduate Rapid Project Grant from the BPS Cognitive Psychology Section

@duncanbradley_