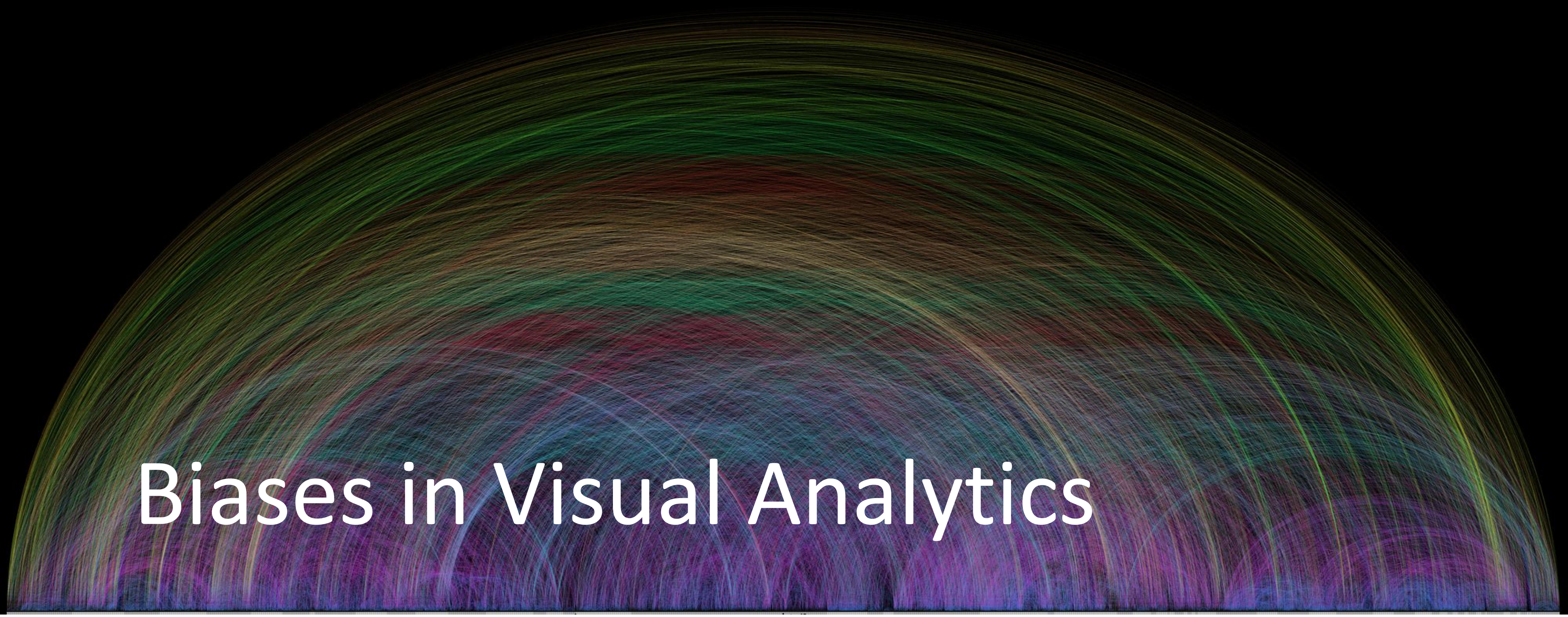
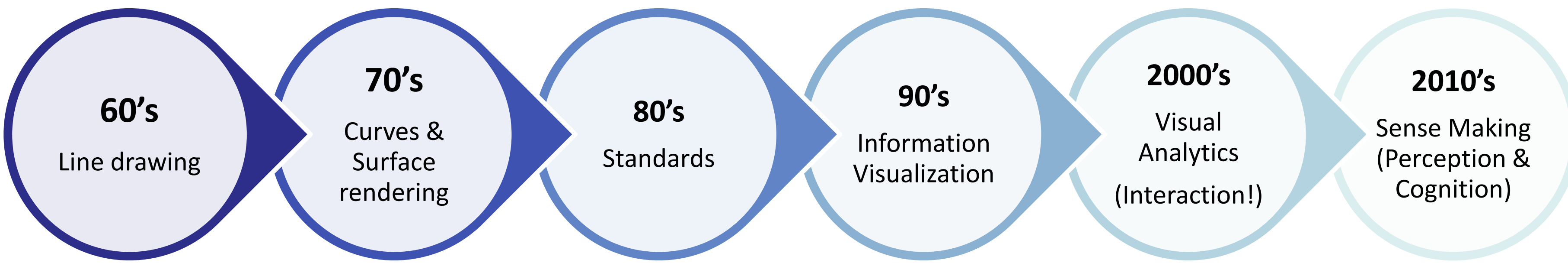


Biases in Visual Analytics

BAIS 6140 – Information Visualization

L. Miguel Encarnação

A brief history of Visualization



What is a bias?

- Bias is a **disproportionate weight** in favor of or against an idea or thing.
- In science and engineering, a bias is a **systematic error**.

Wikipedia

Biases in Data Analysis

Biases can come in at any step along the data analysis pipeline.

Note the general focus on data, not methods!

Data Source

- **Functional:** biases due to platform affordances and algorithms
- **Normative:** biases due to community norms
- **External:** biases due to phenomena outside social platforms
- **Non-individuals:** e.g., organizations, automated agents

Data Collection

- **Acquisition:** biases due to, e.g., API limits
- **Querying:** biases due to, e.g., query formulation
- **Filtering:** biases due to removal of data “deemed” irrelevant

Data Processing

- **Cleaning:** biases due to, e.g., default values
- **Enrichment:** biases from manual or automated annotations
- **Aggregation:** e.g., grouping, organizing, or structuring data

Data Analysis

- **Qualitative Analyses:** lack generalizability, interpret. biases
- **Descriptive Statistics:** confounding bias, obfuscated measurements
- **Prediction & Inferences:** data representation, perform. variations
- **Observational studies:** peer effects, select. bias, ignorability

Evaluation

- **Metrics:** e.g., reliability, lack of domain insights
- **Interpretation:** e.g., contextual validity, generalizability
- **Disclaimers:** e.g., lack of negative results and reproducibility

Biases in *Social* Data Analysis

Issues When Working With Social Data.

Organized based on effects and triggers.

General Challenges

- **Population Biases:** differences in demographics
- **Behavioral Biases:** differences in user behavior
- **Content Biases:** lexical, syntactic, and semantic biases in user content
- **Linking Biases:** differences in network connections, interactions
- **Temporal Biases:** changing biases over time
- **Redundancy:** duplicates, near duplicates

Challenges During Data Analysis

- **Data Source:** biases at the source of social data
- **Data Collection:** biases due to data collection
- **Data Processing:** biases due to data preprocessing
- **Data Analysis:** validity threats due to methods selection and usage
- **Evaluation:** metrics selection, interpretation pitfalls

Ethical Boundaries

- **General Concepts and Principles**
- **Individual Autonomy:** ensure informed consent
- **Beneficence and Non-maleficence:** actions are beneficial and do not cause harm
- **Justice:** risks and benefits justly shared

– Benjamin Haydon

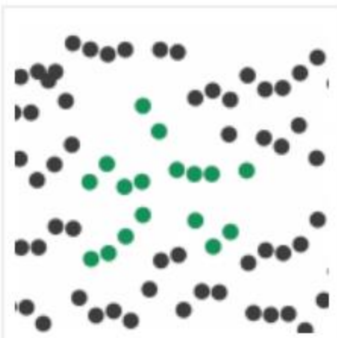
“FORTUNATELY FOR SERIOUS MINDS, A BIAS RECOGNIZED IS A BIAS STERILIZED.”

Statistical Bias Types

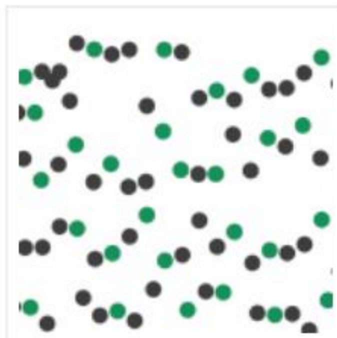
- Selection bias
- Self-selection bias
- Recall bias
- Observer bias
- Survivorship bias
- Omitted variable bias
- Cause-effect bias
- Funding bias
- Cognitive bias

https://data36.com/statistical-bias-types-explained/

Selection Bias



selection bias



proper random sampling

Example

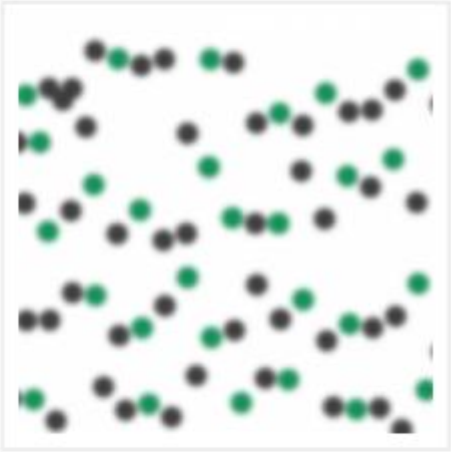
- Student experiments

Self-selection Bias

Example

- Polls

Recall Bias



Example

- Nostalgia
  - Good Ol' times
  - Make America Great Again
  - First love

Observer Bias

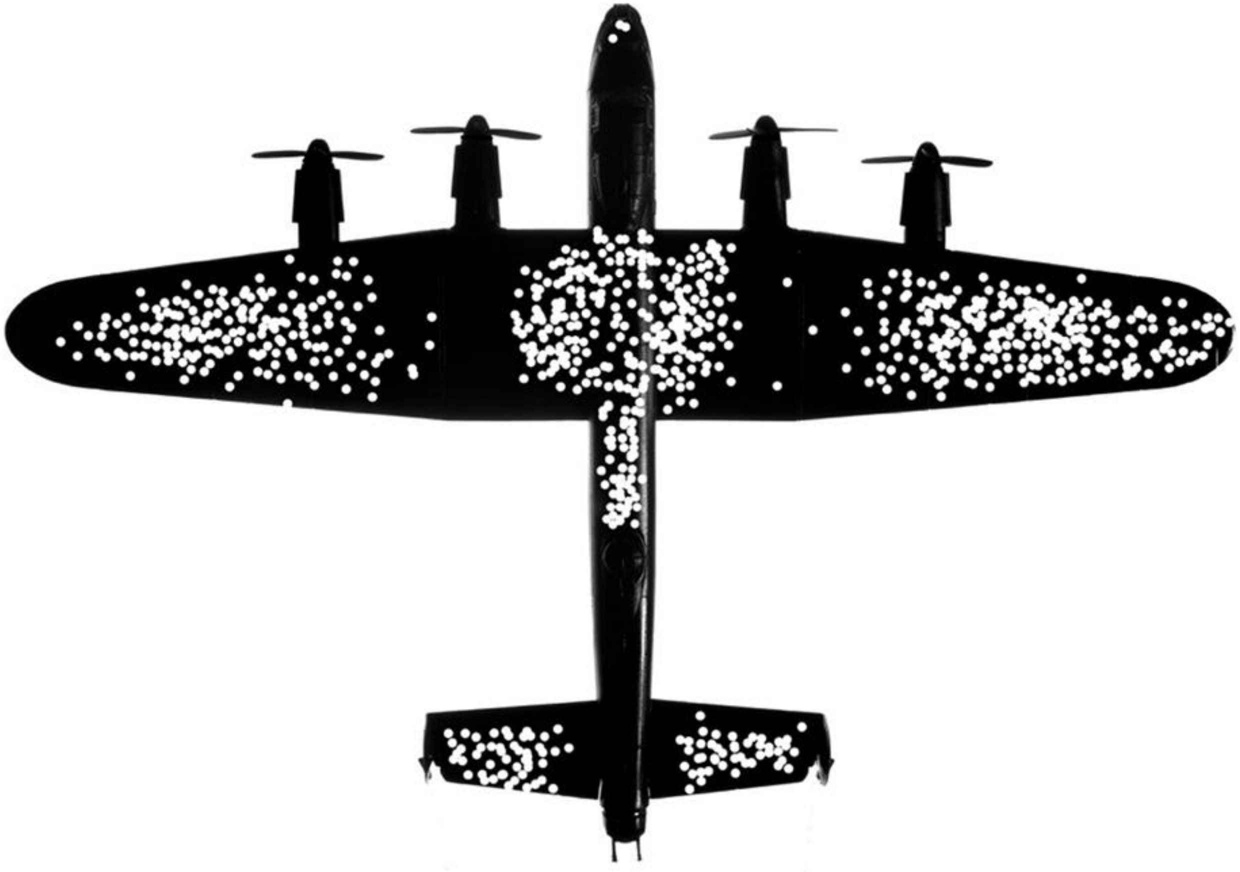
Example

- Conscious & unconscious prejudices

United Features Syndicate, Inc.

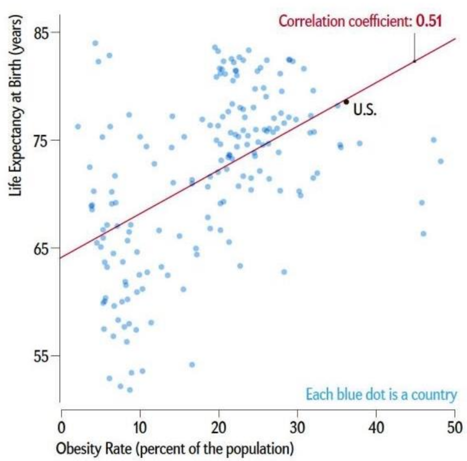
Survivorship Bias

https://www.youtube.com/watch?v=ZyLVIvBidIA



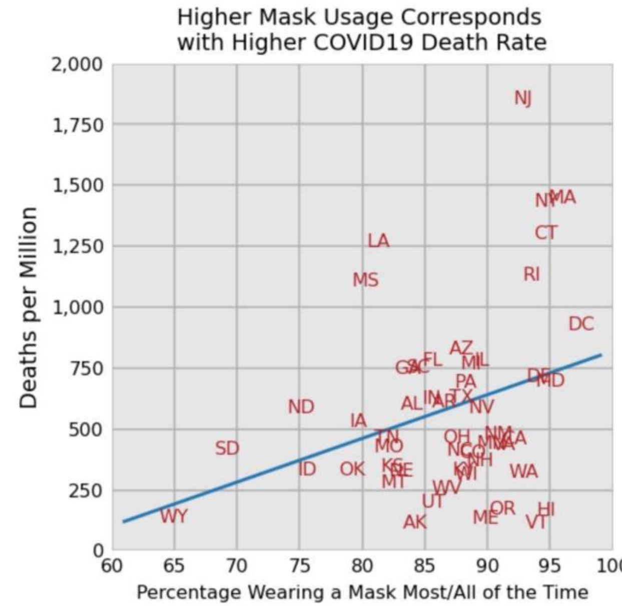
Omitted Variable Bias

	A and B are positively correlated	A and B are negatively correlated
B has a positive effect on Y	*Positive bias*	*Negative bias*
B has a negative effect on Y	*Negative bias*	*Positive bias*

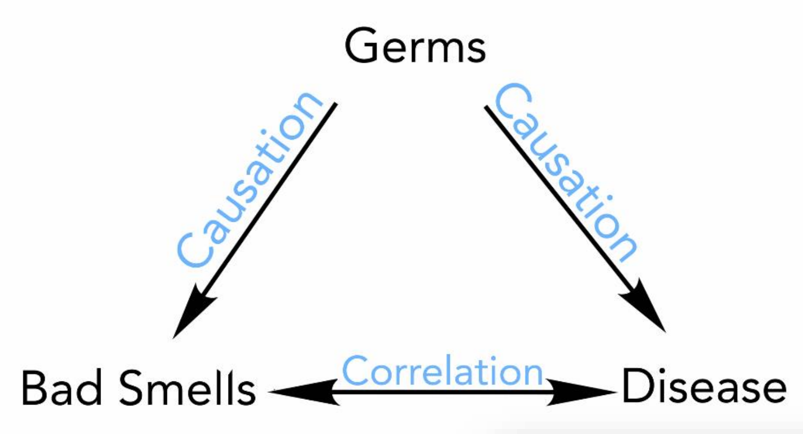


Example

- Omitting cost of <u>not</u> innovating from innovation risk calculation

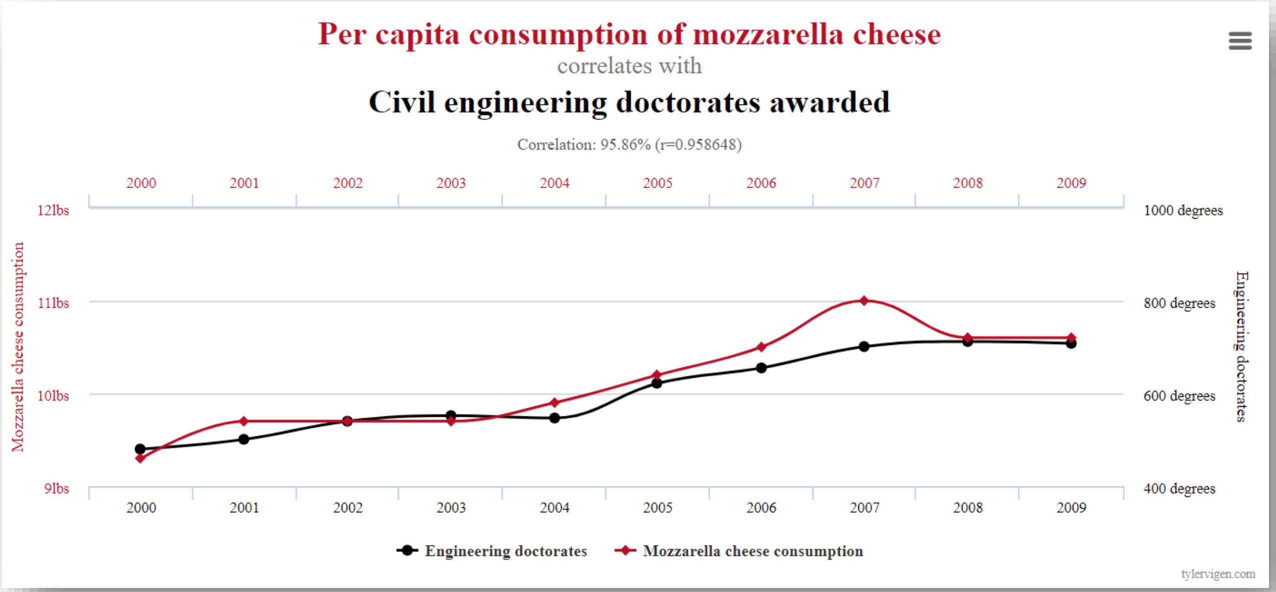


Cause-Effect Bias



Example

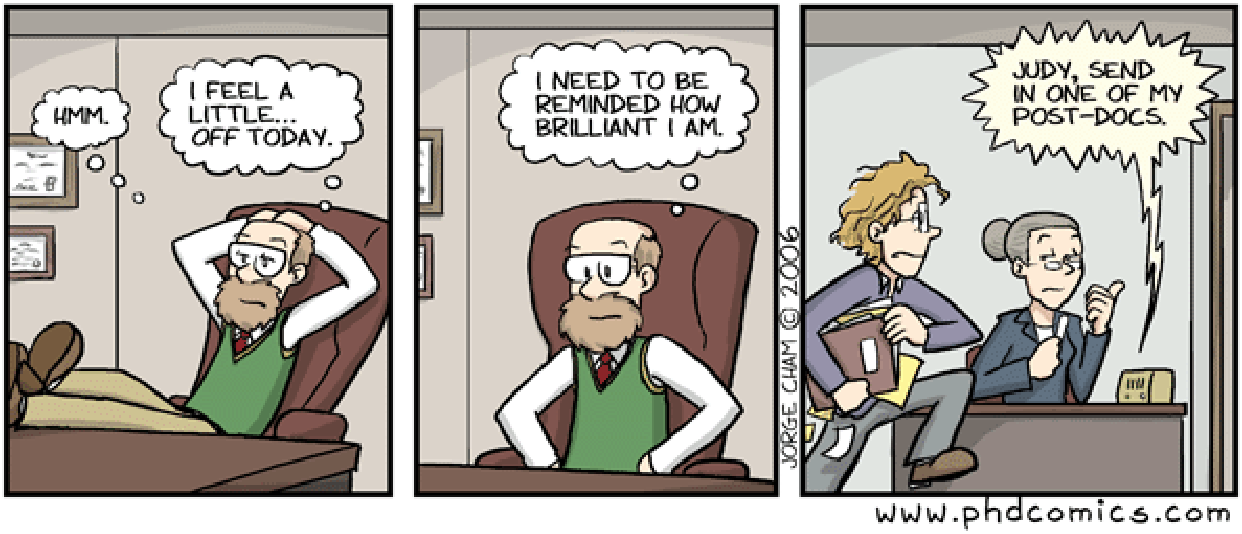
- Increase of autism diagnoses due to vaccinations

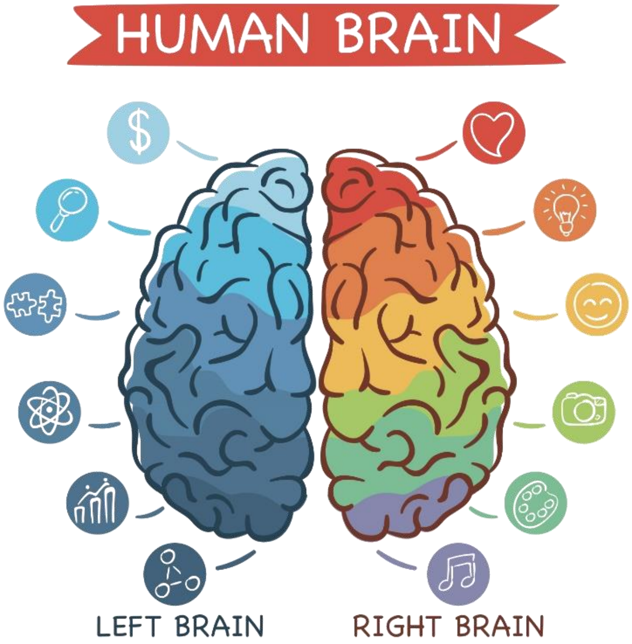


Funding Bias

Examples

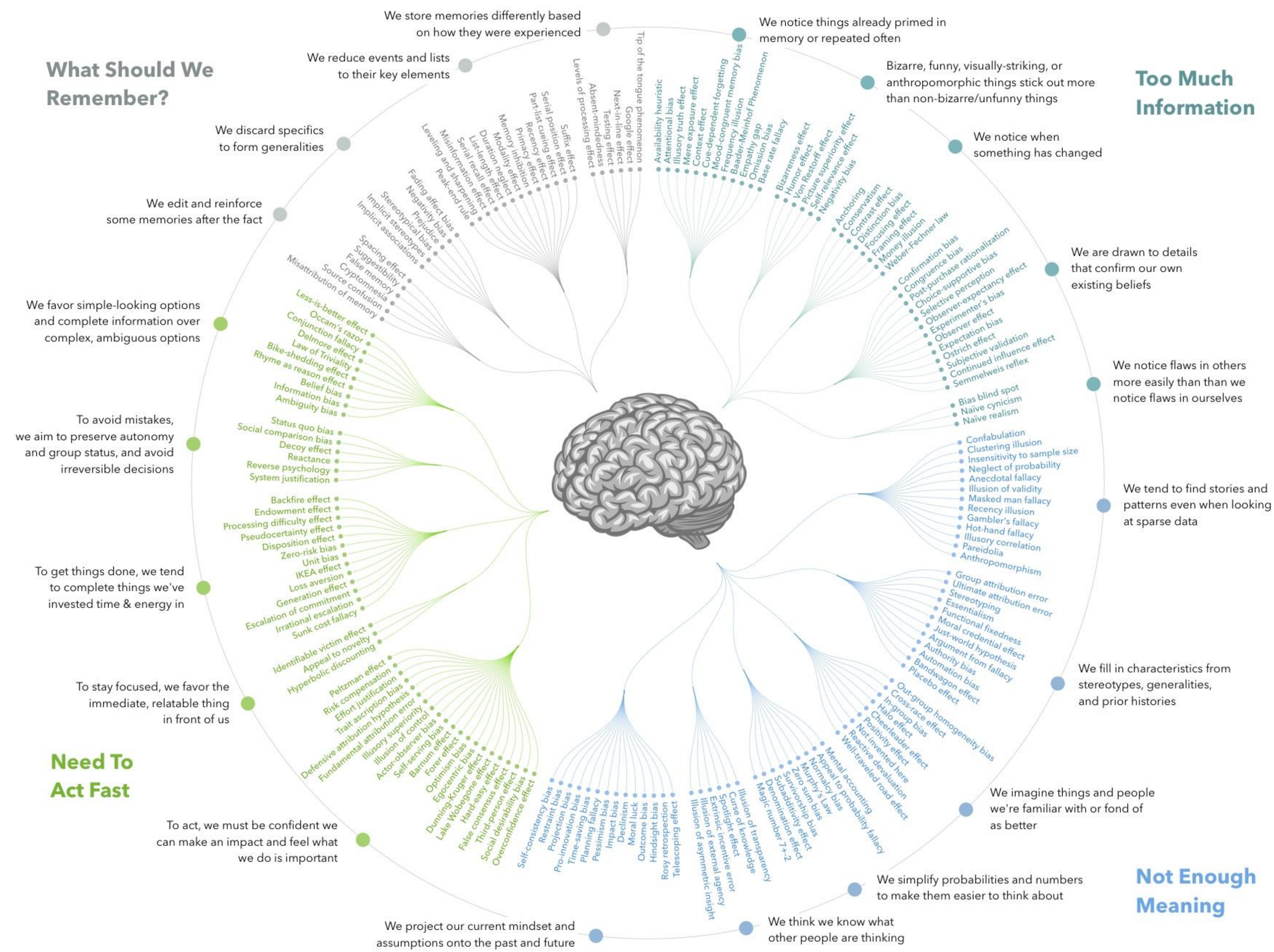
- Partisan News
- PhD research
  - also Survivorship Bias once published





COGNITIVE BIASES

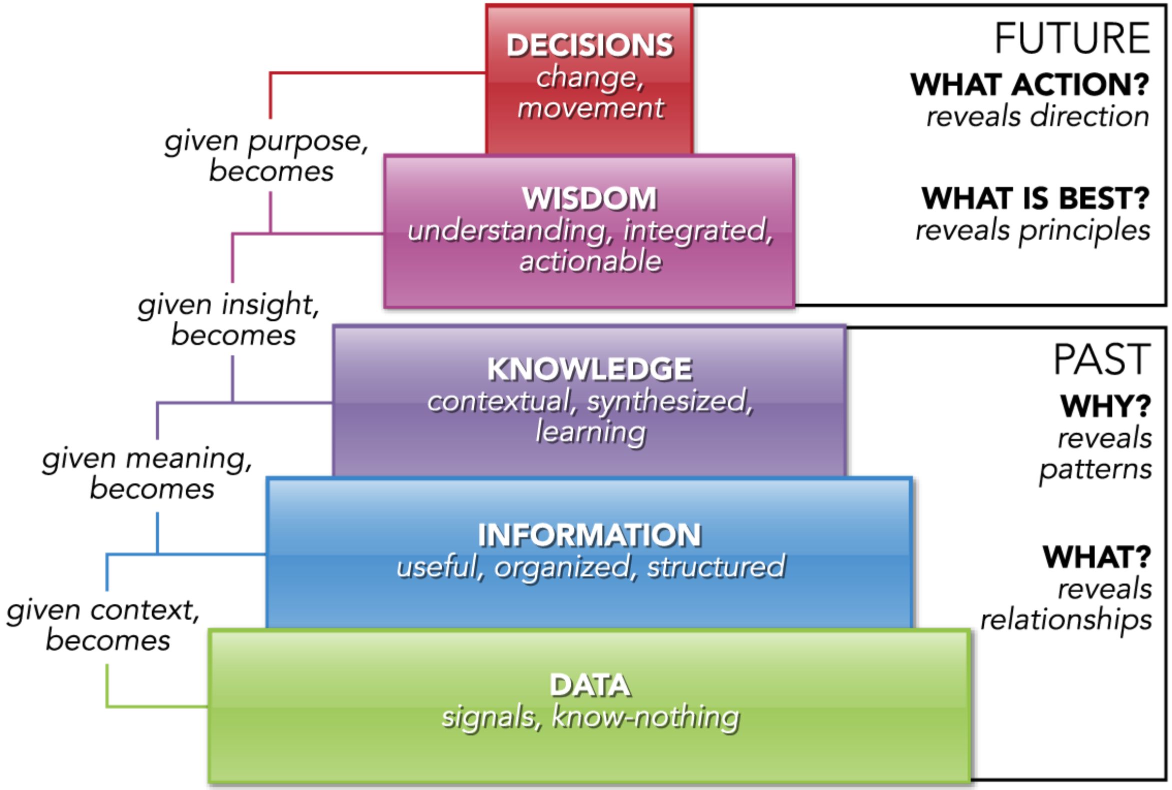
COGNITIVE BIAS CODEX



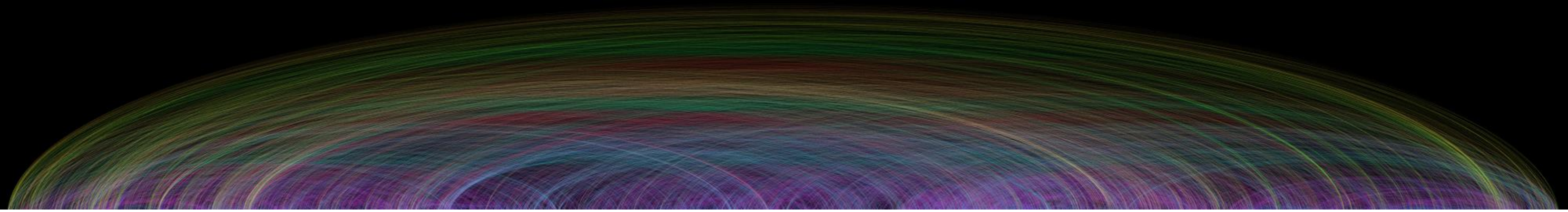
Recap: Analytic Purposes & Decision Making

*

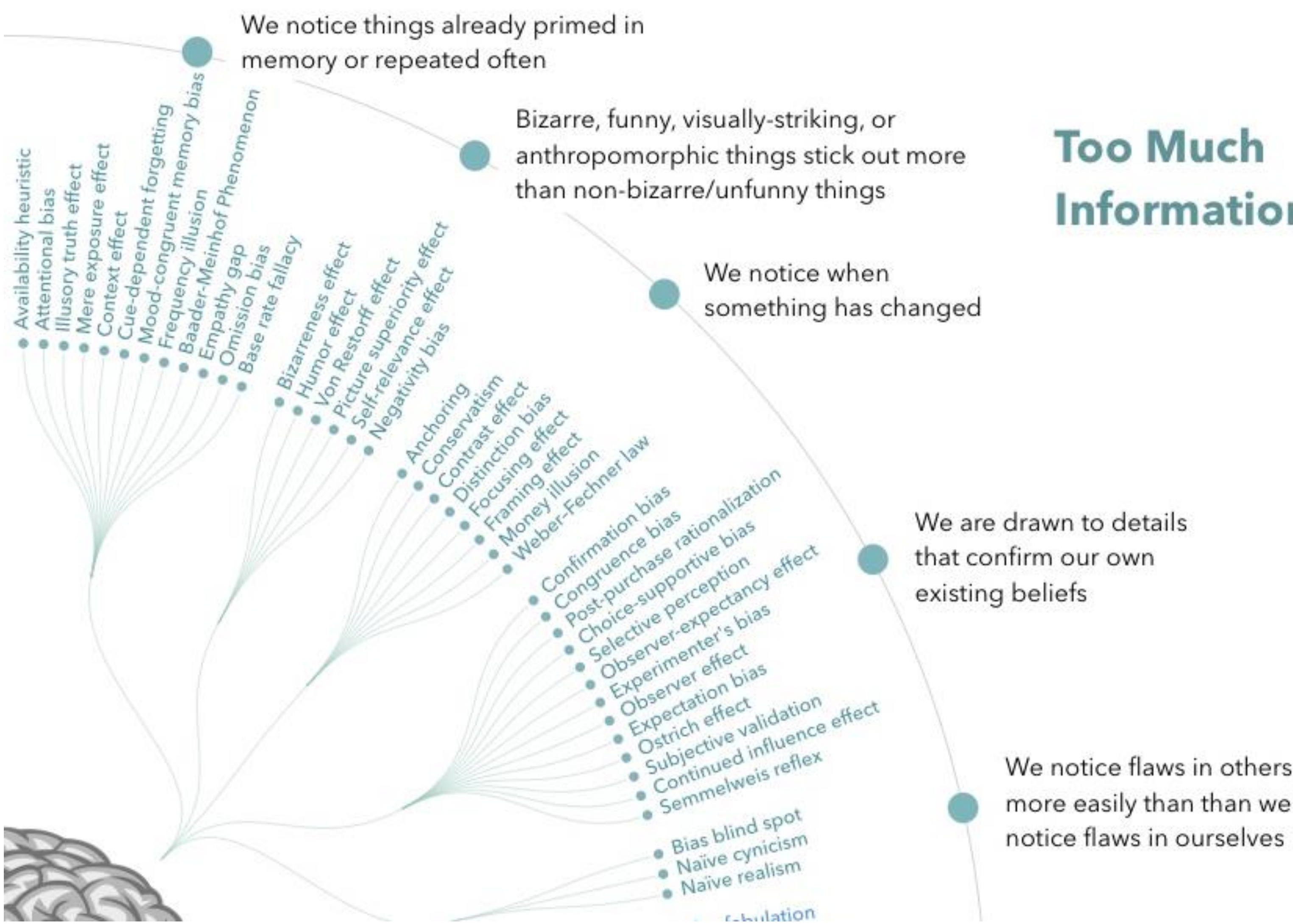
DIKW pyramid

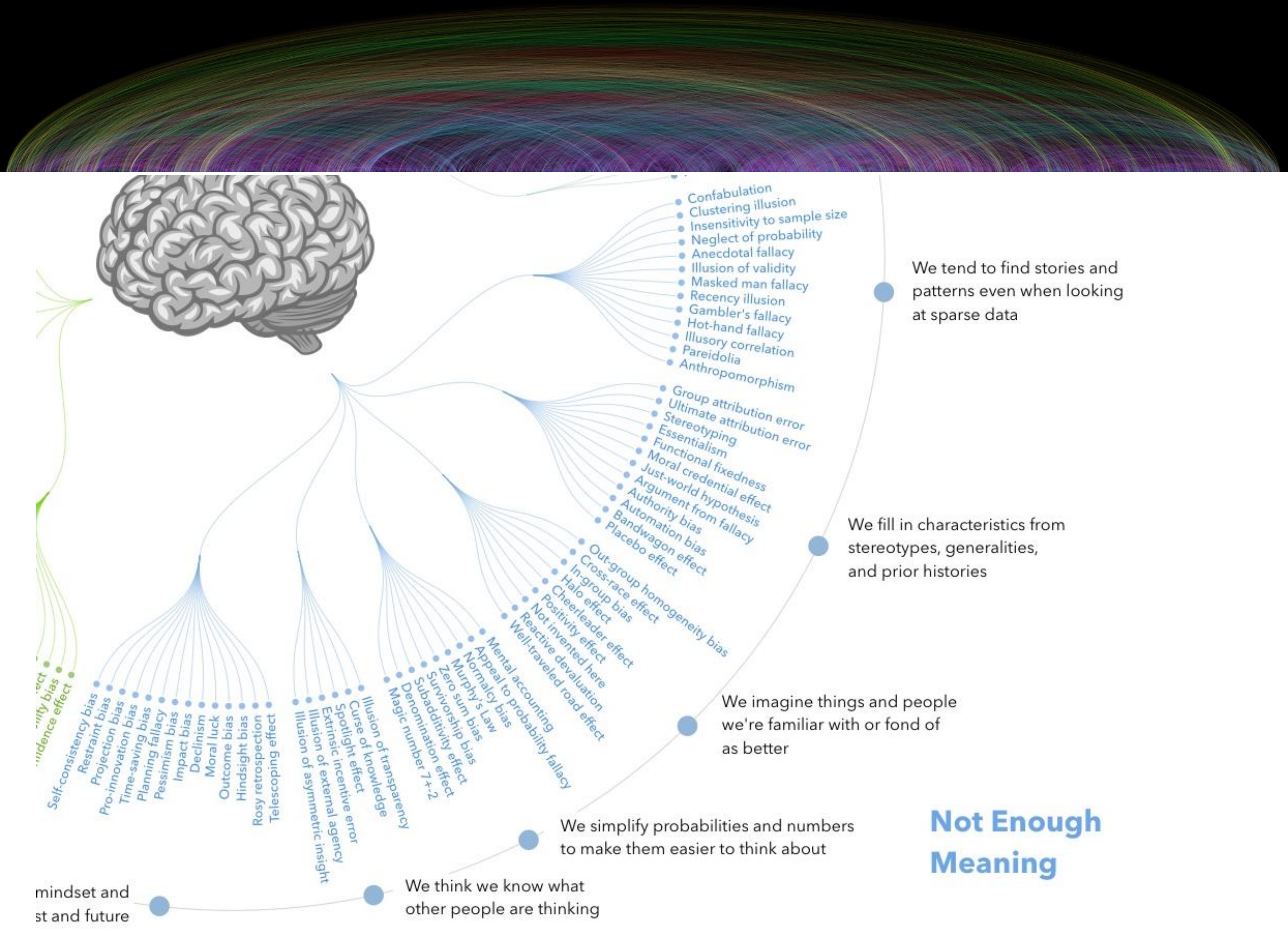


What Should We Remember?

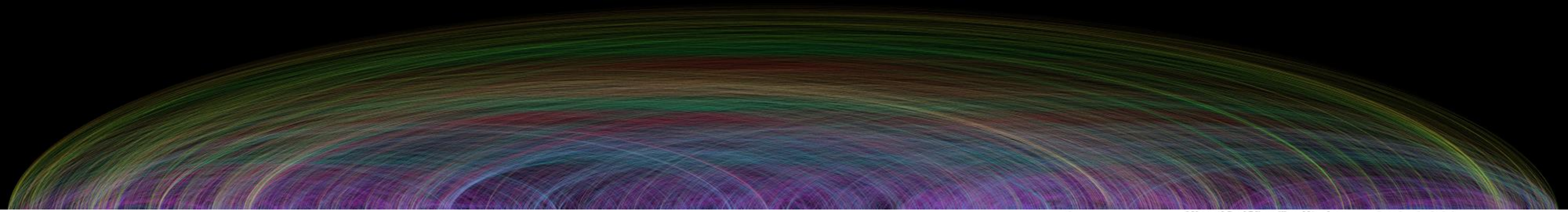


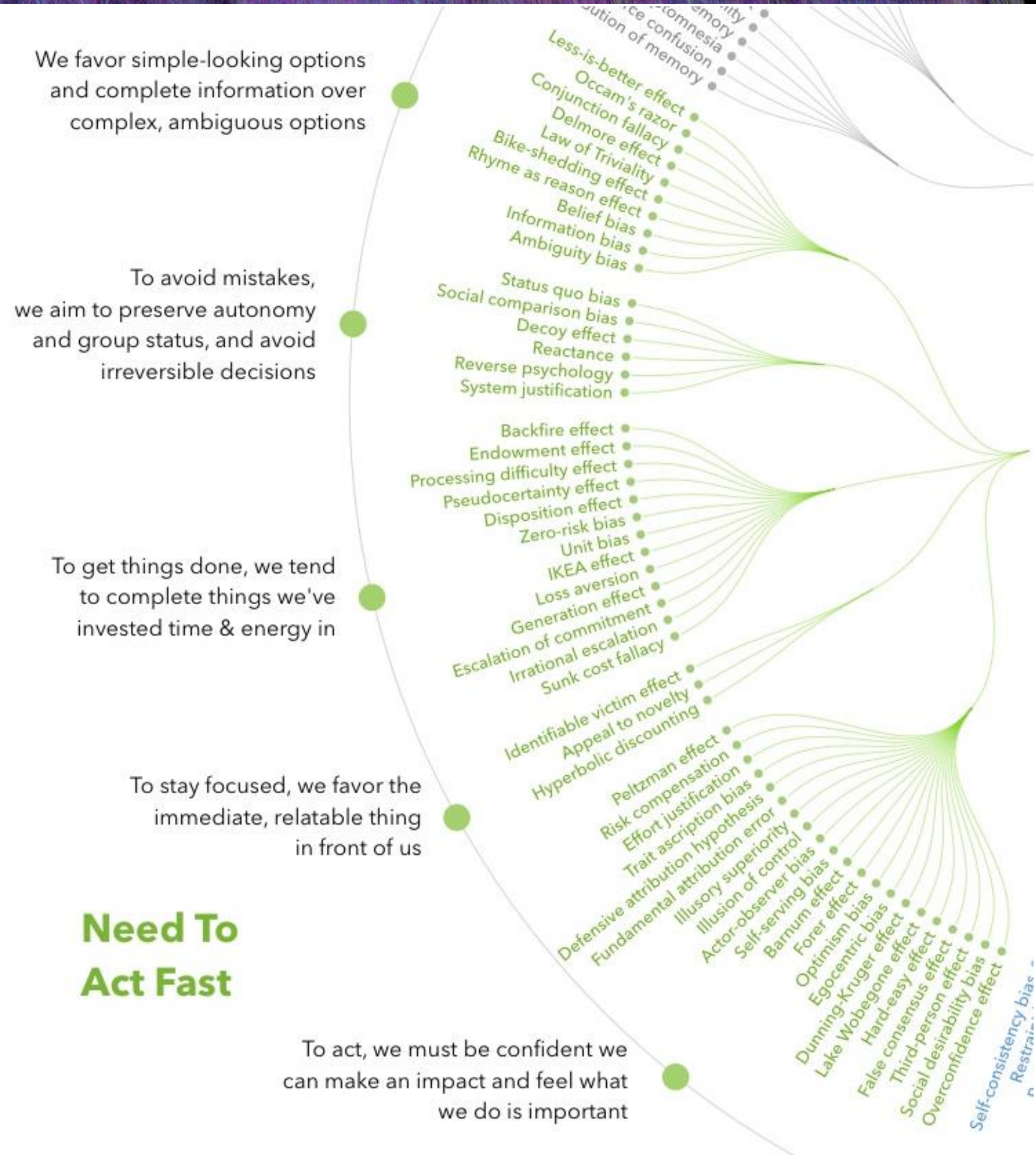
Too Much Information





Confabulation
Clustering illusion
Insensitivity to sample size
Neglect of probability
Anecdotal fallacy
Illusion of validity
Masked man fallacy
Recency illusion
Gambler's fallacy
Hot-hand fallacy
Illusory correlation
Pareidolia
Anthropomorphism
We tend to find stories and patterns even when looking at sparse data
Group attribution error
Ultimate attribution error
Stereotyping
Essentialism
Functional fixedness
Moral credential effect
Just-world hypothesis
Argument from fallacy
Authority bias
Automation bias
Bandwagon effect
Placebo effect
We fill in characteristics from stereotypes, generalities, and prior histories
Out-group homogeneity bias
Cross-race effect
In-group bias
Halo effect
Cheerleader effect
Positivity effect
Not invented here
Reactive devaluation
Well-traveled road effect
We imagine things and people we're familiar with or fond of as better
Mental accounting
Appeal to probability fallacy
Normalcy bias
Murphy's Law
Zero sum bias
Survivorship bias
Subadditivity effect
Denomination effect
Magic number 7+-2
We simplify probabilities and numbers to make them easier to think about
Illusion of transparency
Curse of knowledge
Spotlight effect
Extrinsic incentive error
Illusion of external agency
Illusion of asymmetric insight
We think we know what other people are thinking
Self-consistency bias
Restraint bias
Projection bias
Pro-innovation bias
Time-saving bias
Planning fallacy
Pessimism bias
Impact bias
Declinism
Moral luck
Outcome bias
Hindsight bias
Rosy retrospection
Telescoping effect
mindset and
Not Enough Meaning



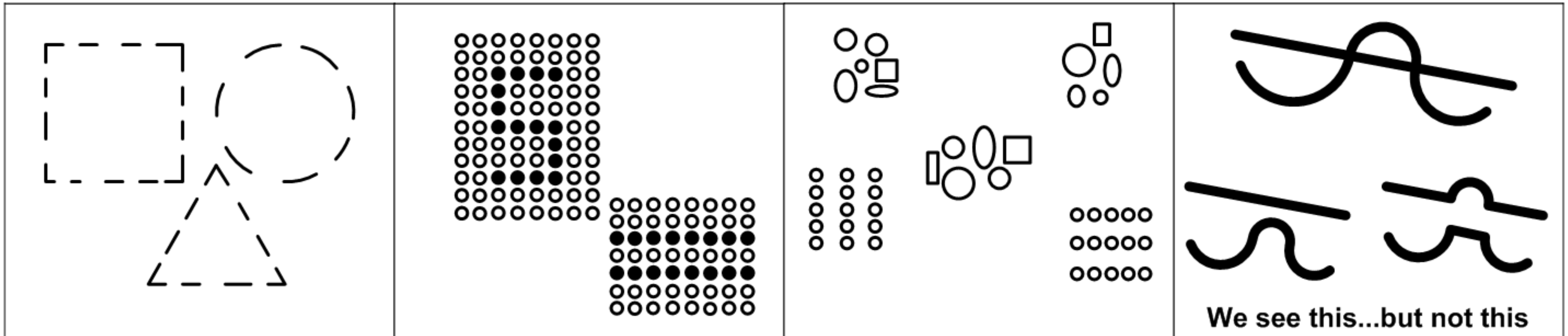


We favor simple-looking options and complete information over complex, ambiguous options
Less-is-better effect
Occam's razor
Conjunction fallacy
Delmore effect
Law of Triviality
Bike-shedding effect
Rhyme as reason effect
Belief bias
Information bias
Ambiguity bias
To avoid mistakes, we aim to preserve autonomy and group status, and avoid irreversible decisions
Status quo bias
Social comparison bias
Decoy effect
Reactance
Reverse psychology
System justification
Backfire effect
Endowment effect
Processing difficulty effect
Pseudocertainty effect
Disposition effect
Zero-risk bias
Unit bias
IKEA effect
Loss aversion
Generation effect
Escalation of commitment
Irrational escalation
Sunk cost fallacy
To get things done, we tend to complete things we've invested time & energy in
Identifiable victim effect
Appeal to novelty
Hyperbolic discounting
To stay focused, we favor the immediate, relatable thing in front of us
Peltzman effect
Risk compensation
Effort justification
Trait ascription bias
Defensive attribution hypothesis
Fundamental attribution error
Illusory superiority
Illusion of control
Actor-observer bias
Self-serving bias
Barnum effect
Forer effect
Optimism bias
Egocentric bias
Dunning-Kruger effect
Lake Wobegone effect
Hard-easy effect
False consensus effect
Third-person effect
Social desirability bias
Overconfidence effect
Self-consistency bias
Need To Act Fast
To act, we must be confident we can make an impact and feel what we do is important

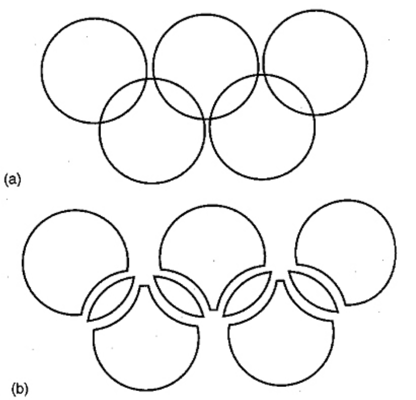
(Nice catalogue at http://www.psy.ritsumei.ac.jp/~akitaoka/cataloge.html)

EXAMPLES OF (VISUAL) BIAS

It starts with Gestalt …

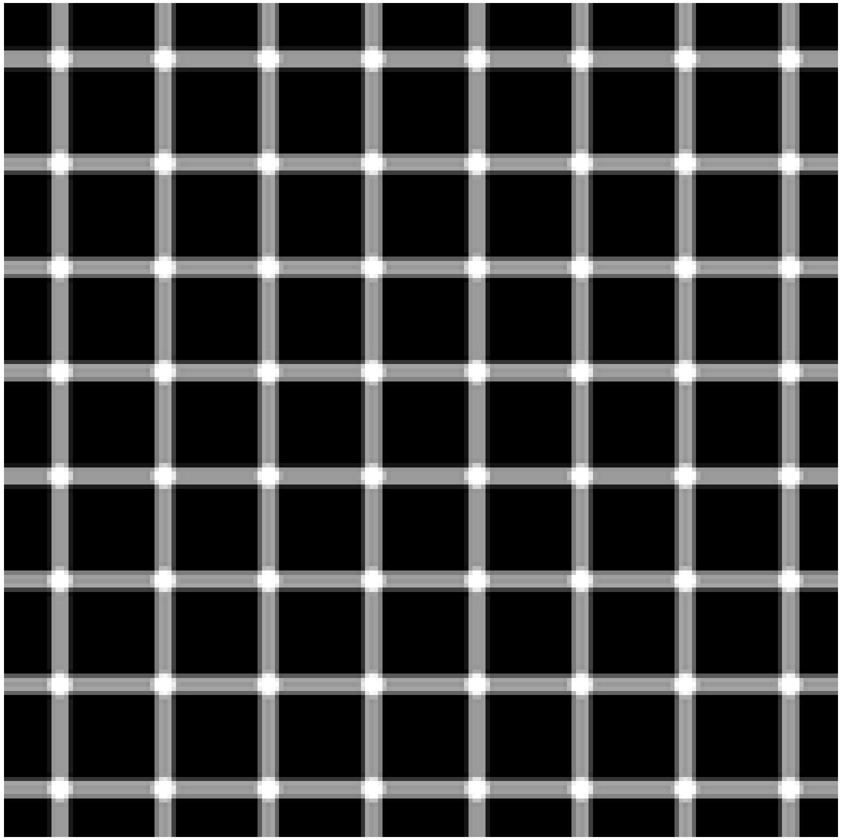


Closure | Similarity | Proximity | Continuity



Prägnanz

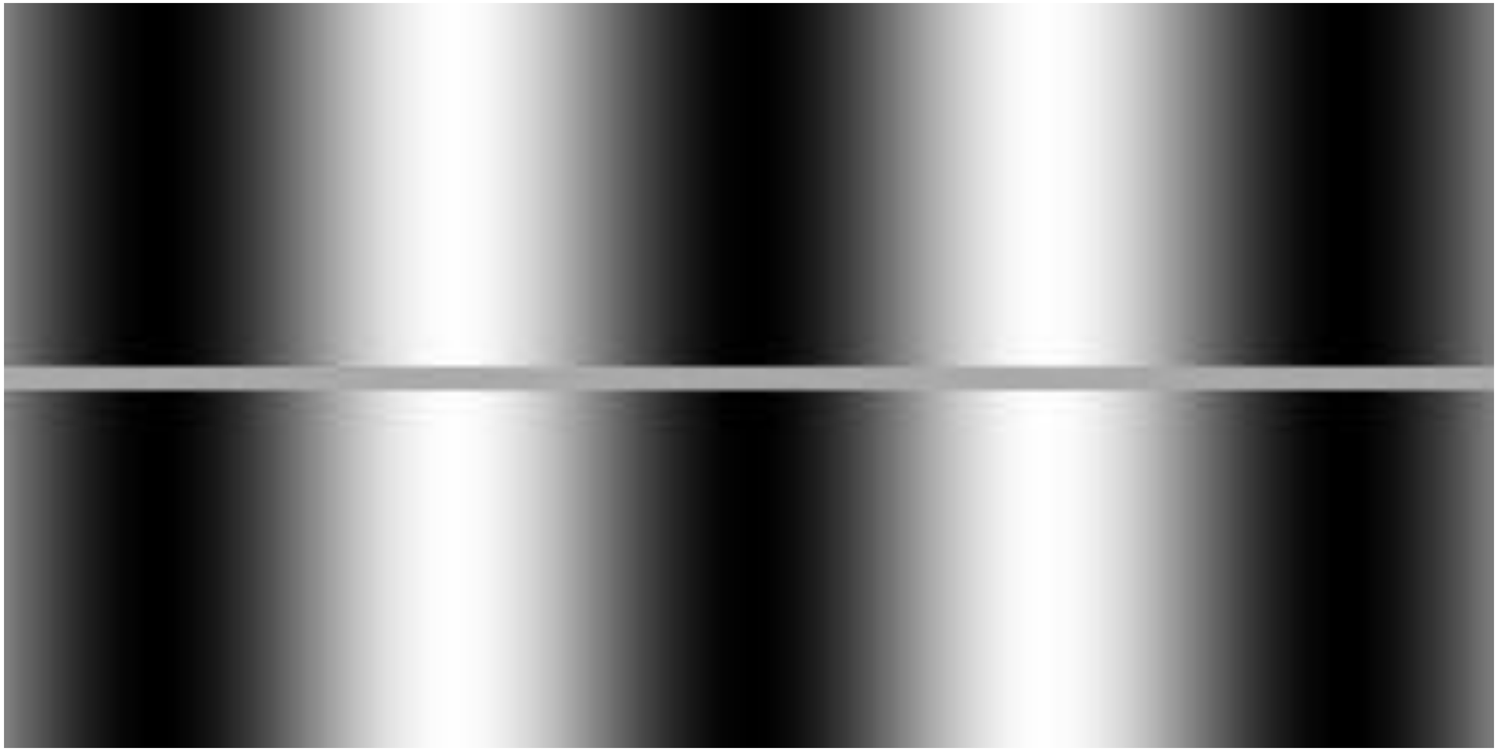
Brightness constancy illusion



(Scintillating Grid Illusion)

Brightness constancy illusion

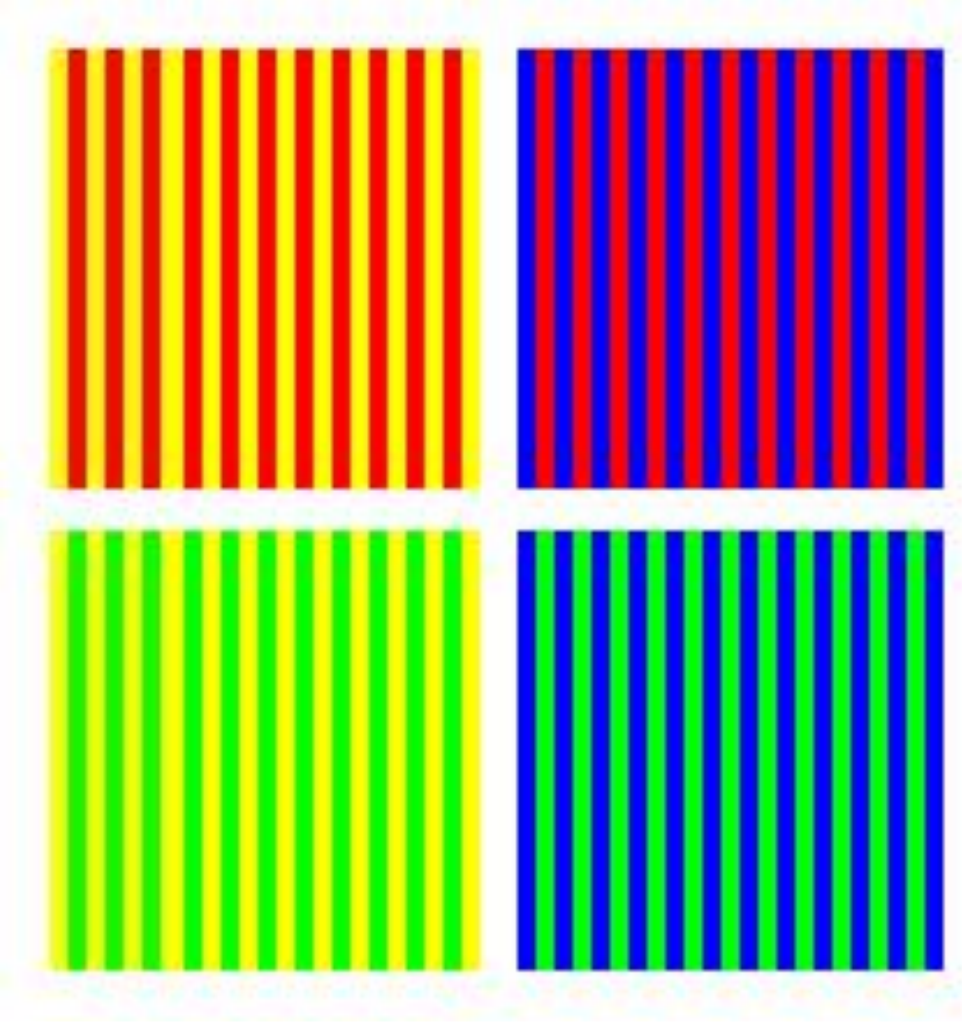
Brightness constancy illusion



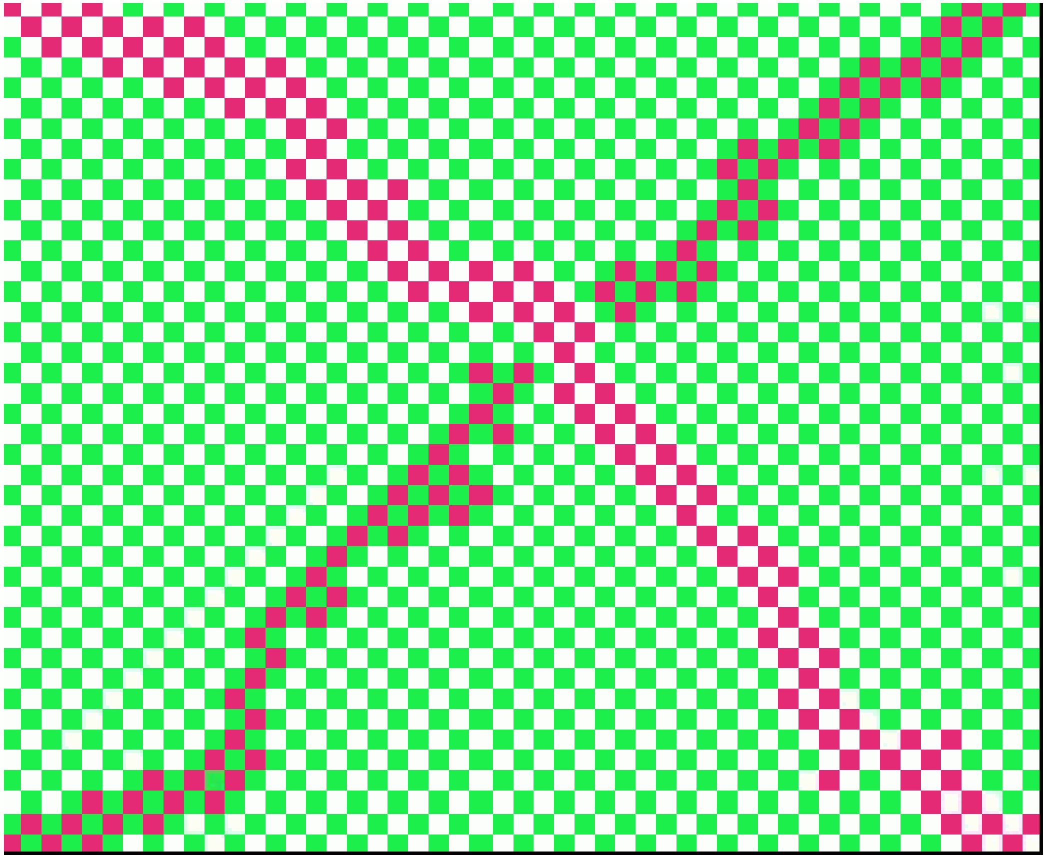
February 27, 2015

- White/Gold or Black/Blue?

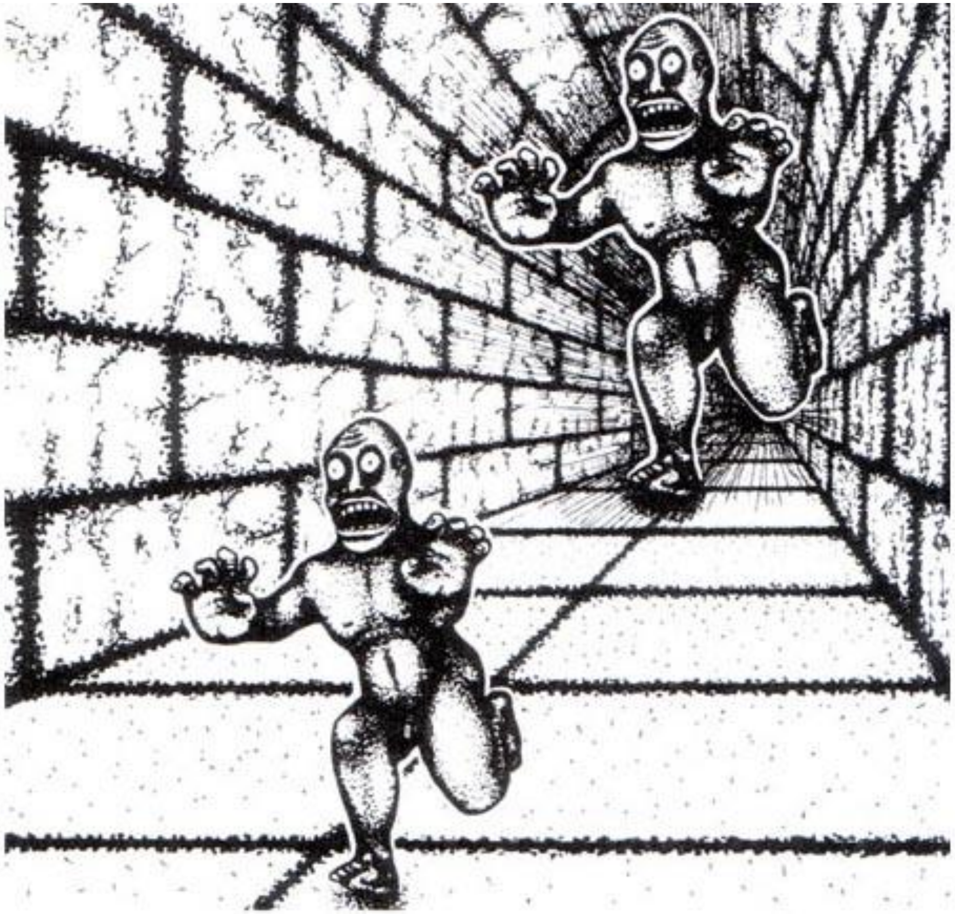
Color constancy illusion



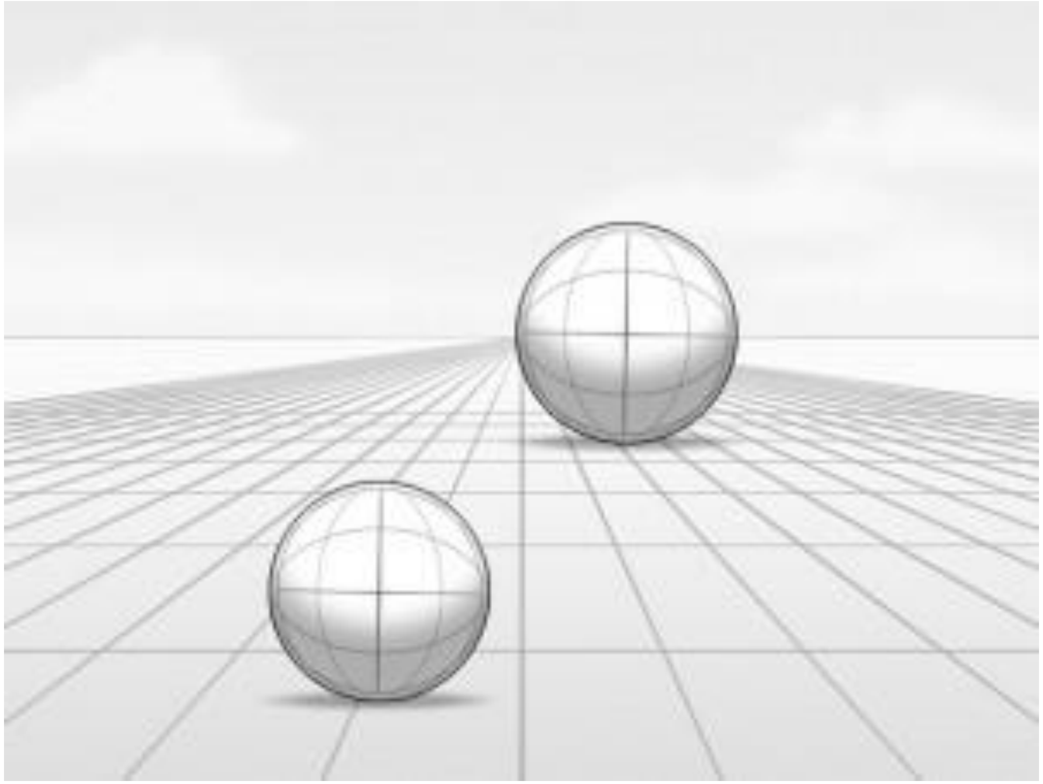
Color constancy illusion



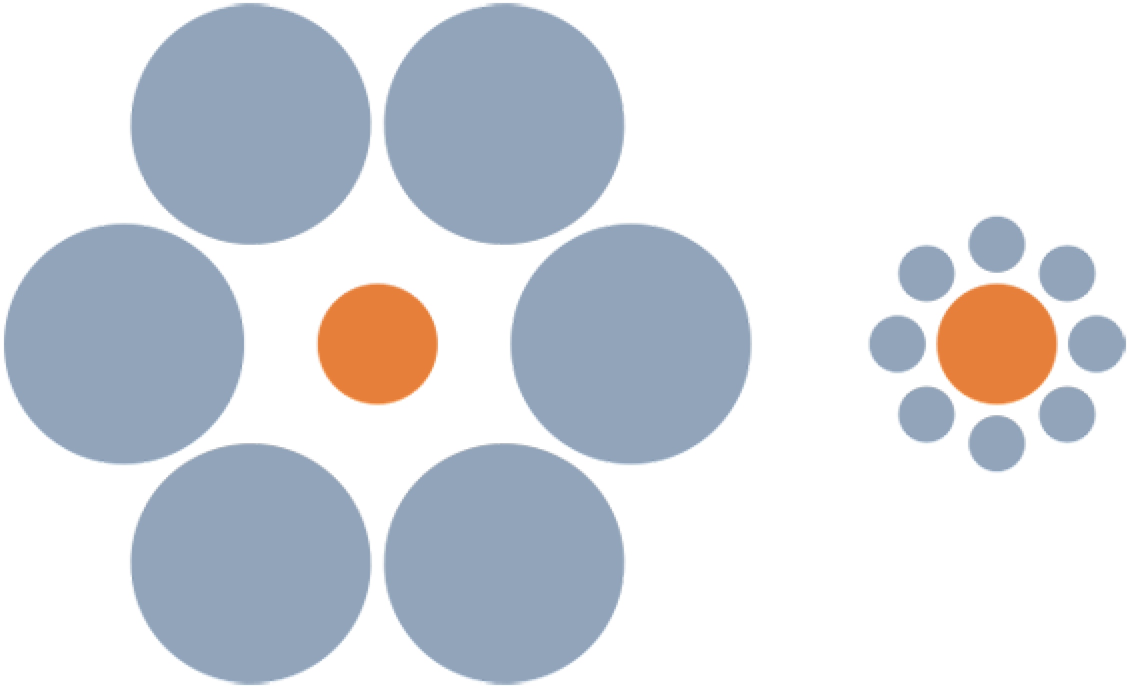
Size constancy illusion



Size constancy illusion

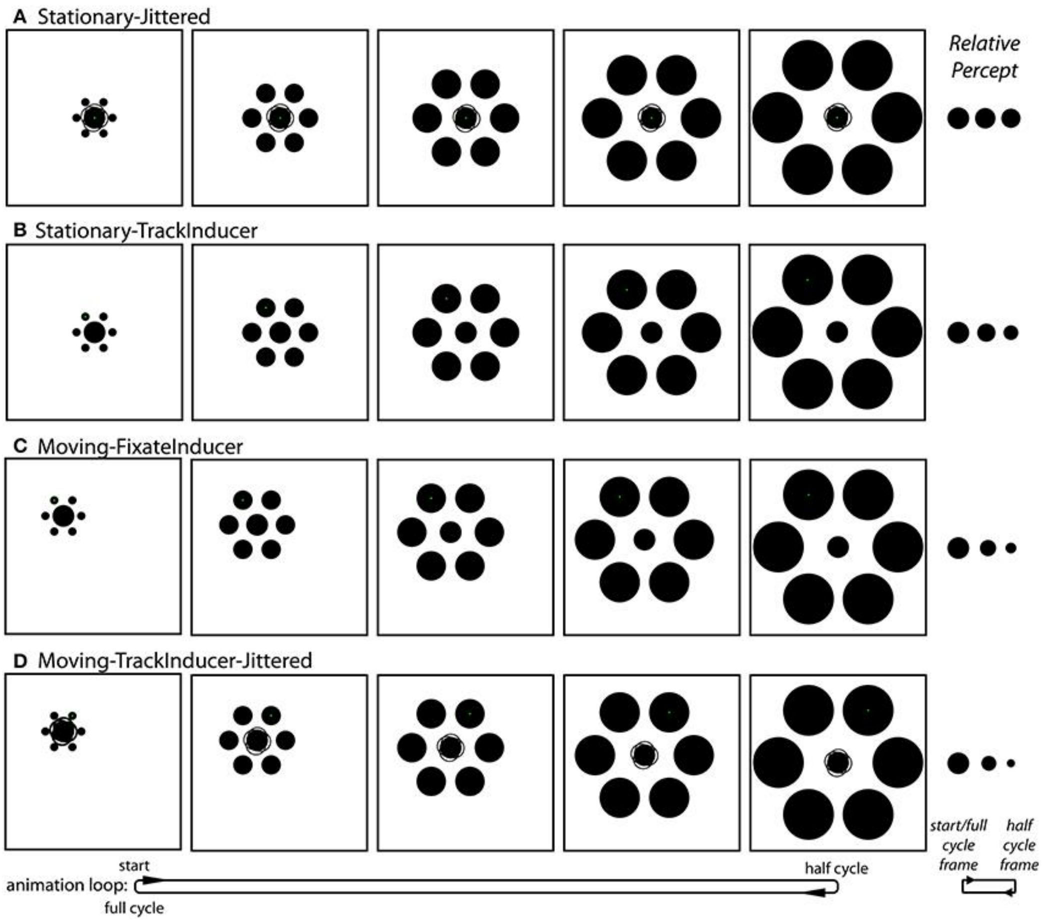


Size constancy illusion

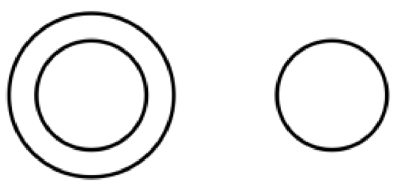
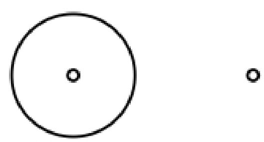


(Ebbinghaus illusion)

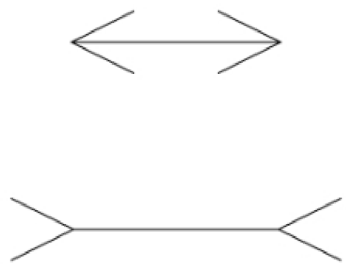
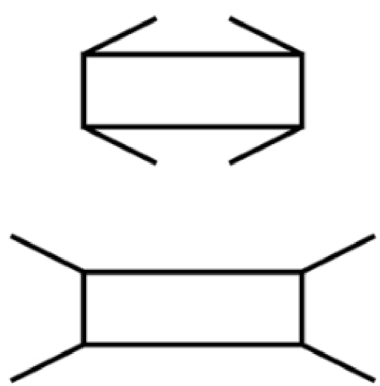
Size constancy illusion



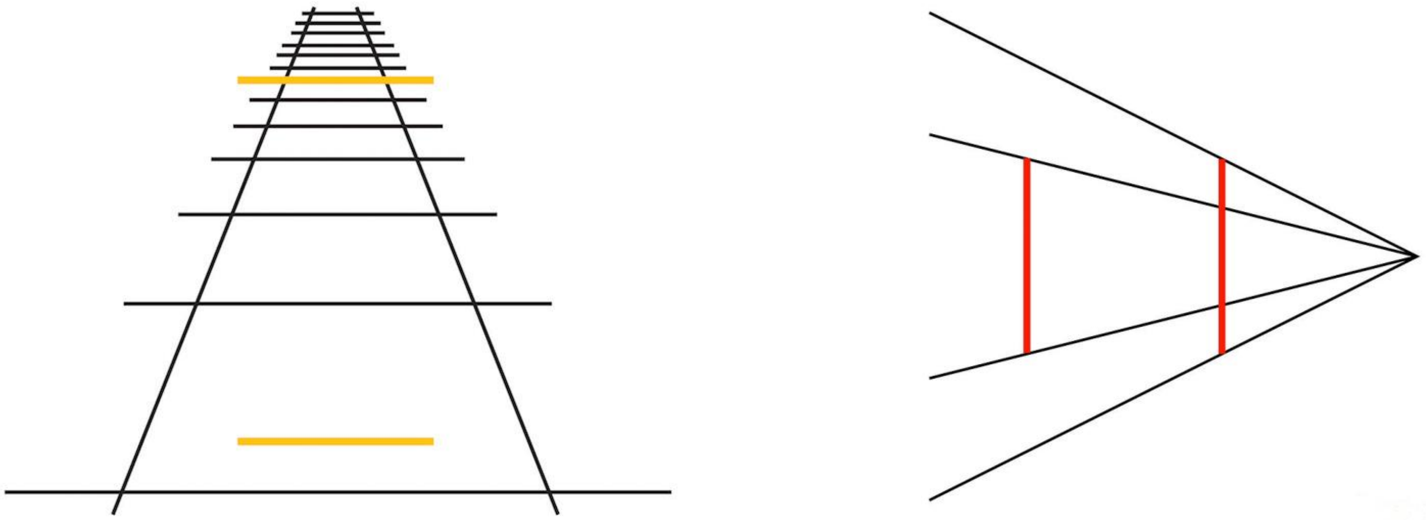
Size constancy illusion

	Delboeuf illusion, underestimation of the outer circle
	Delboeuf illusion, overestimation of the inner circle
	Delboeuf illusion, underestimation of the inner circle

Size constancy illusion

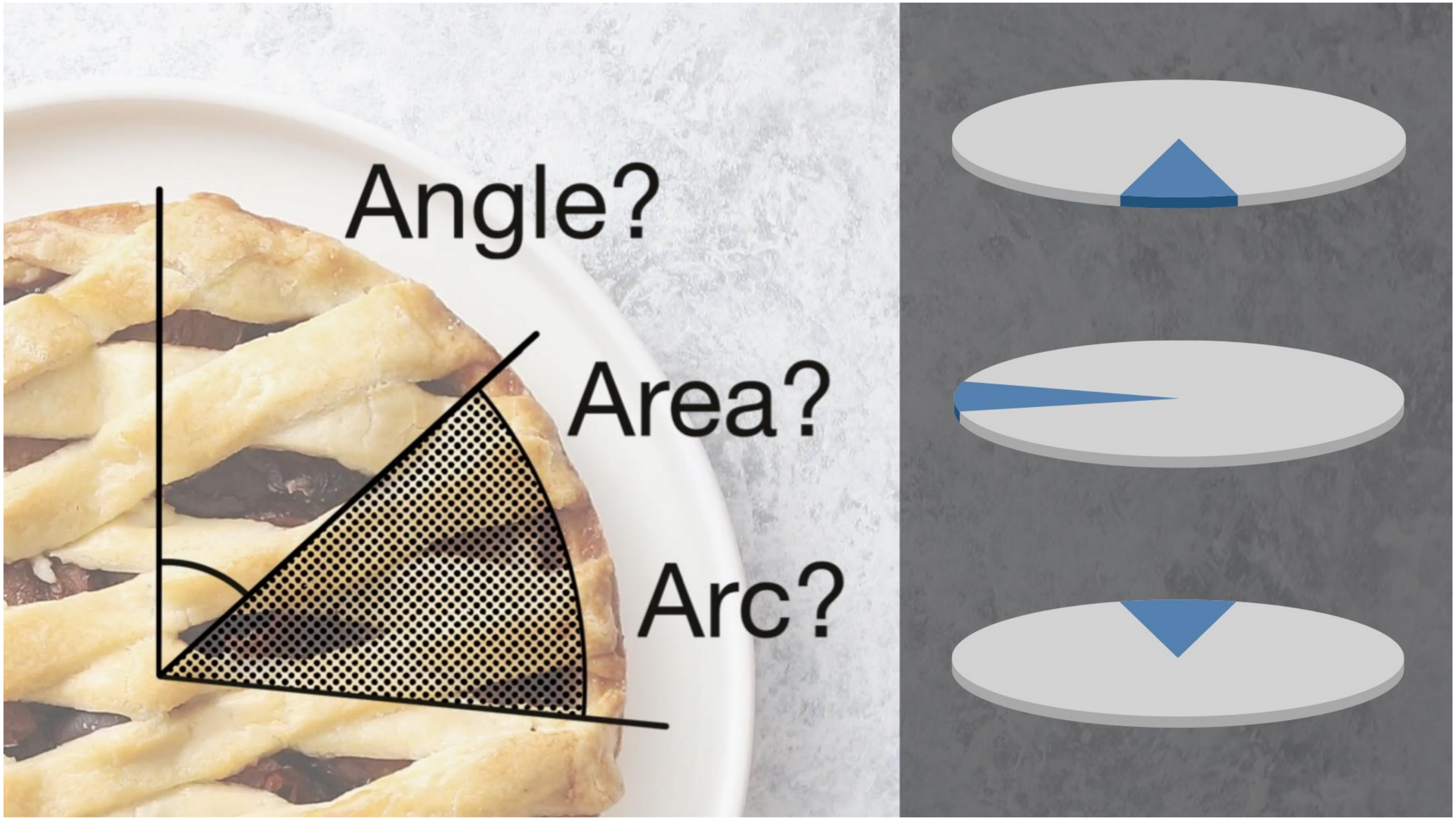
	Müller-Lyer illusion
	Waite-Massaro illusion

Size constancy illusion



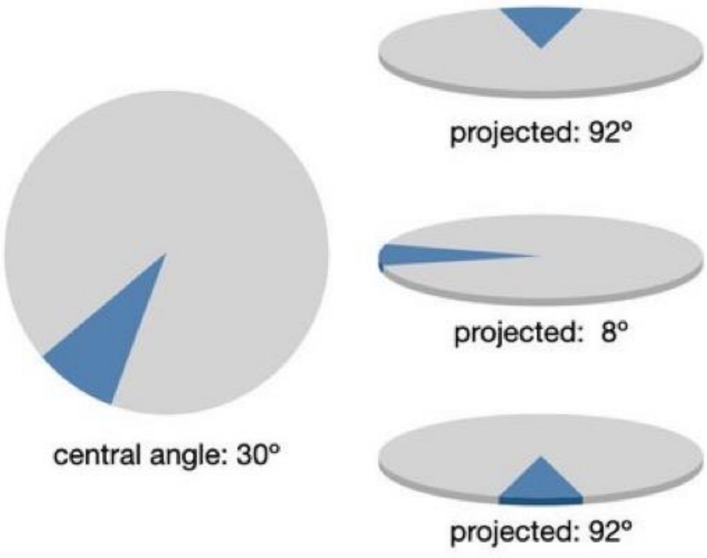
(Ponzo illusion)

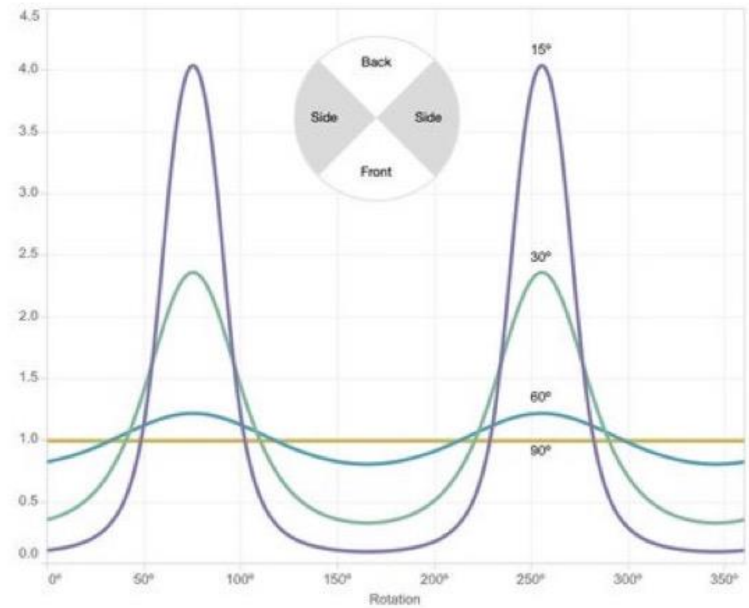
Implications for 3D



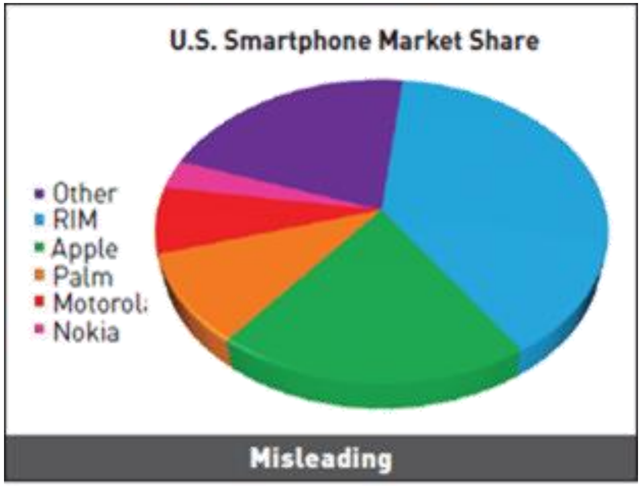
https://eagereyes.org/blog/2019/eagereyestv-episode-3-3d-pie-charts-for-science

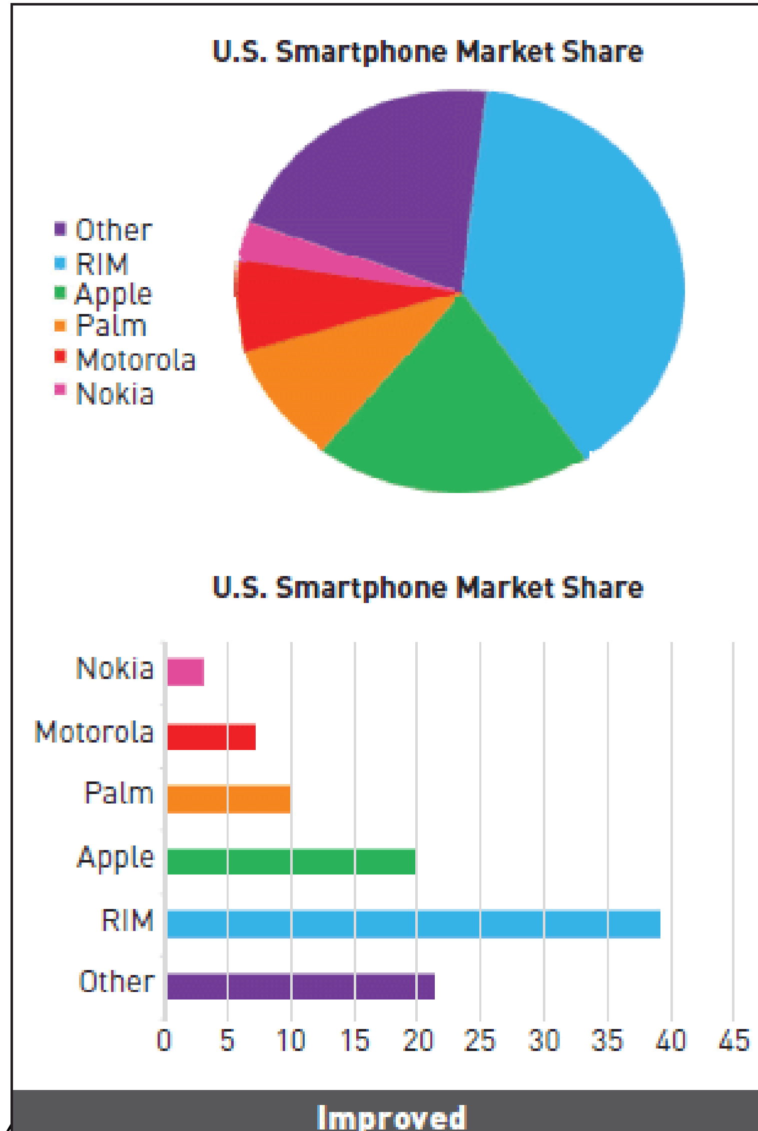
Implications for 3D



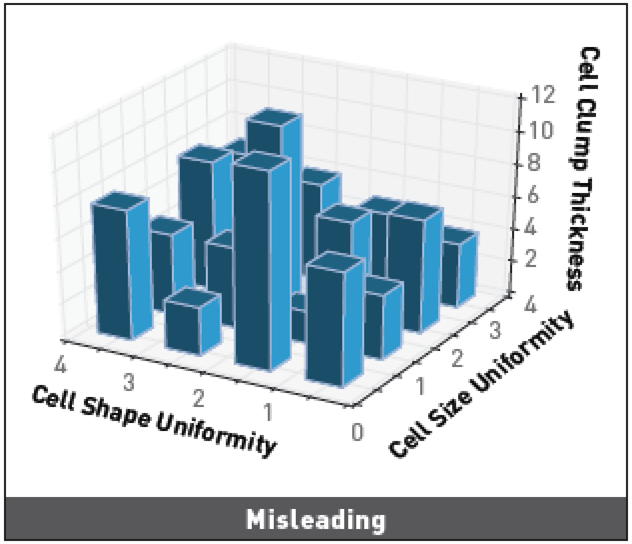


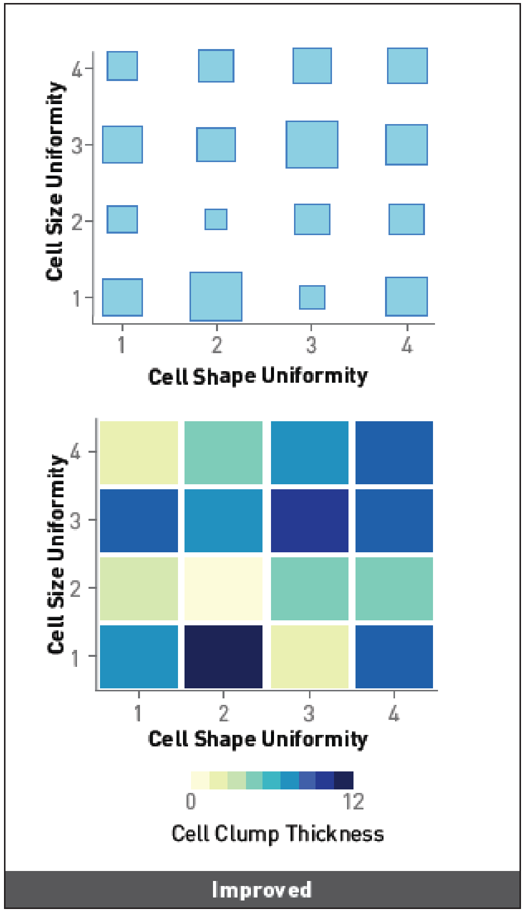
Implications for 3D



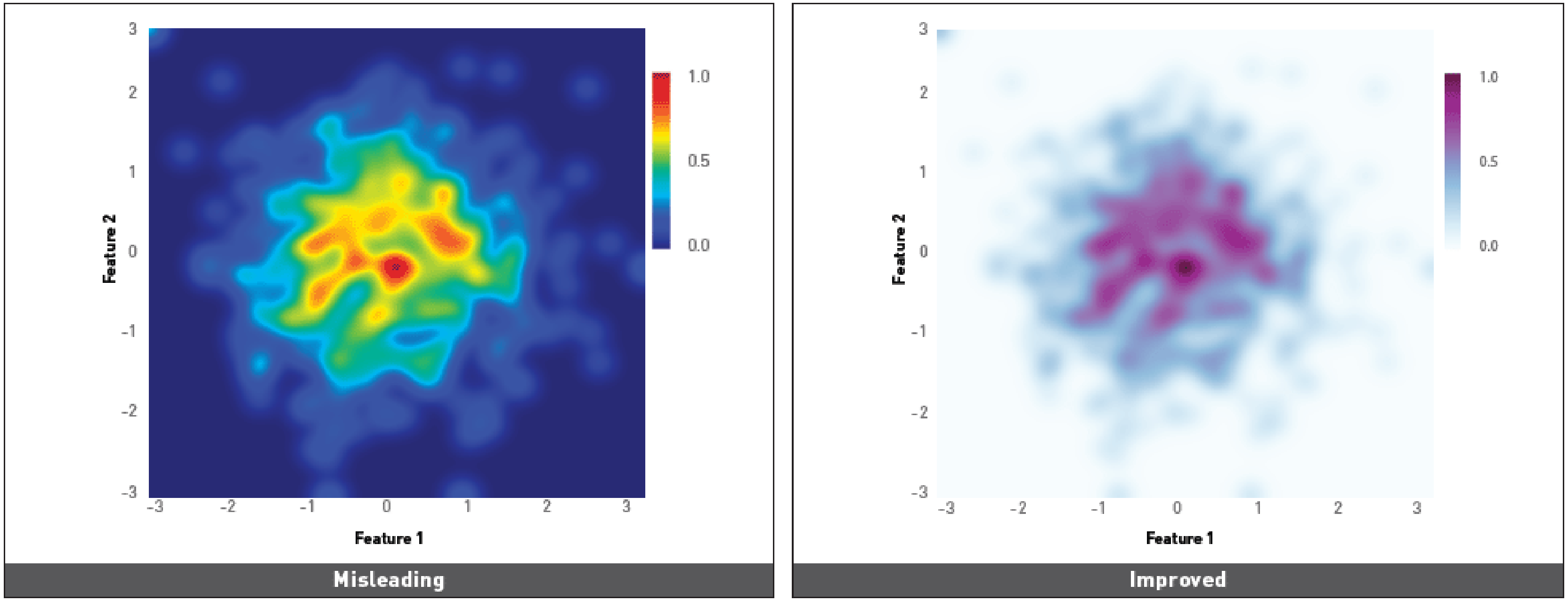


Implications for 3D

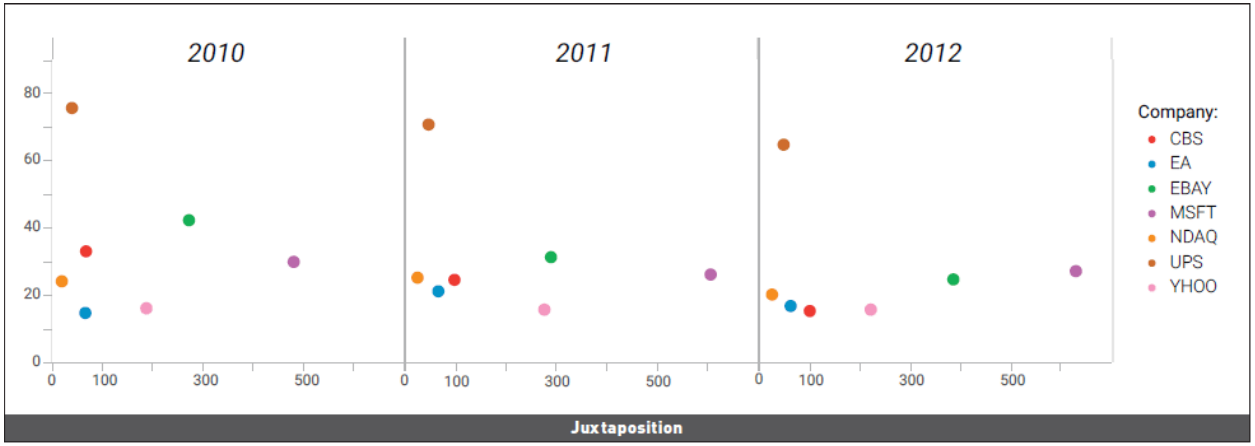


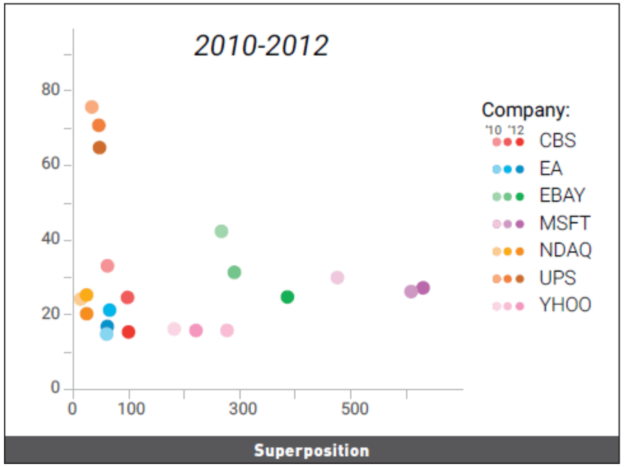


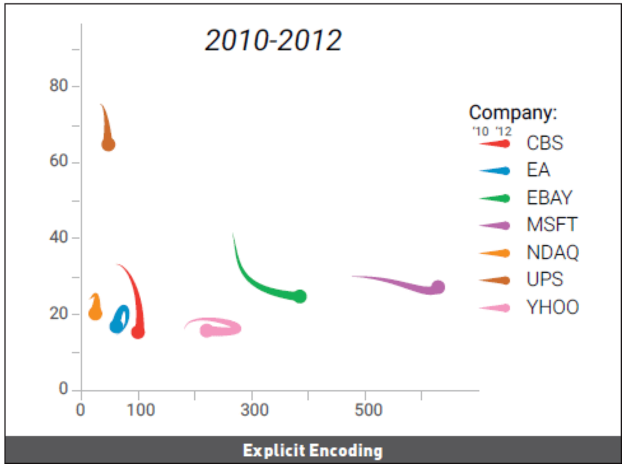
Rainbow palettes



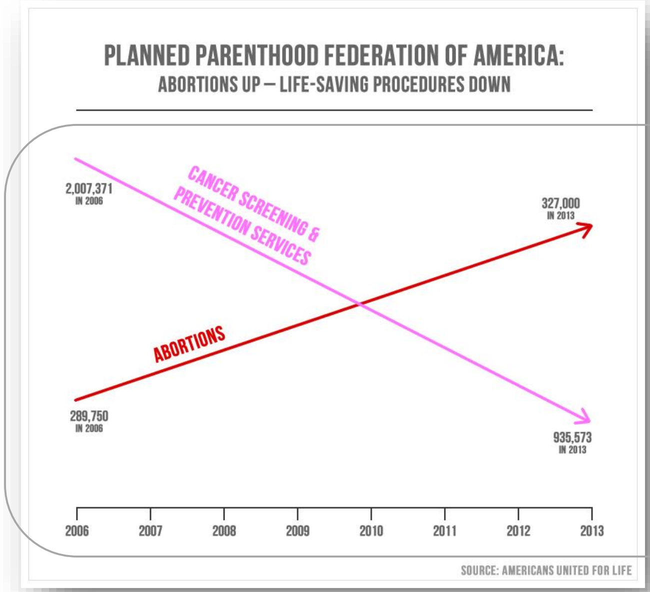
Animation

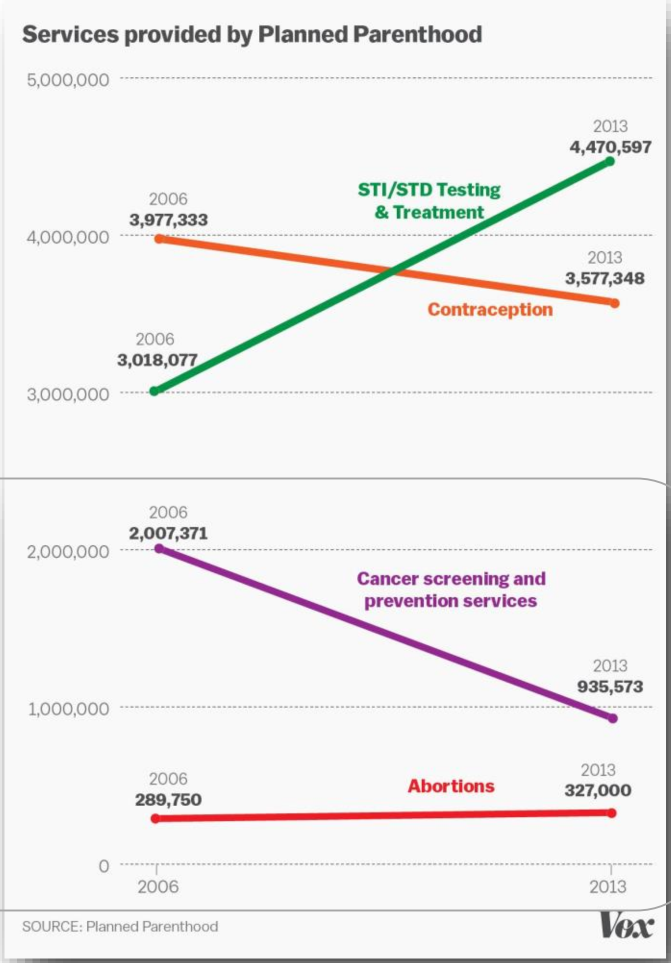

2010
2011
2012
Company:
CBS
EA
EBAY
MSFT
NDAQ
UPS
YHOO
0
20
40
60
80
100
300
500
Juxtaposition


2010-2012
Company:
'10 '12
CBS
EA
EBAY
MSFT
NDAQ
UPS
YHOO
0
20
40
60
80
100
300
500
Superposition

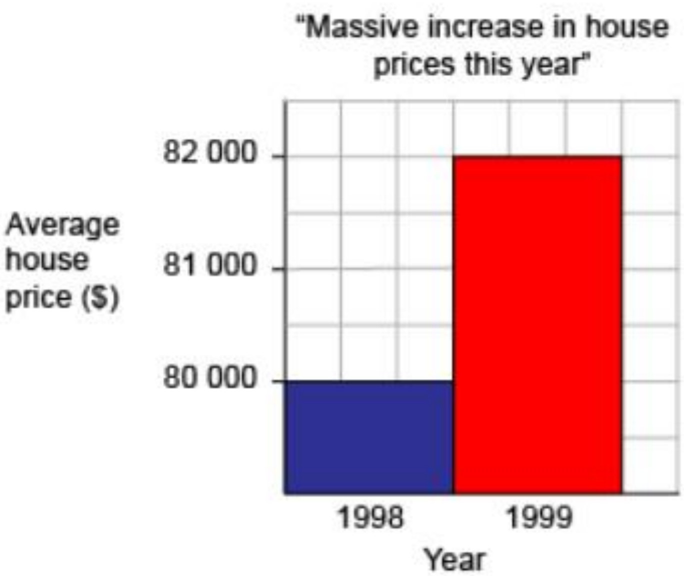

2010-2012
Company:
'10 '12
CBS
EA
EBAY
MSFT
NDAQ
UPS
YHOO
0
20
40
60
80
100
300
500
Explicit Encoding

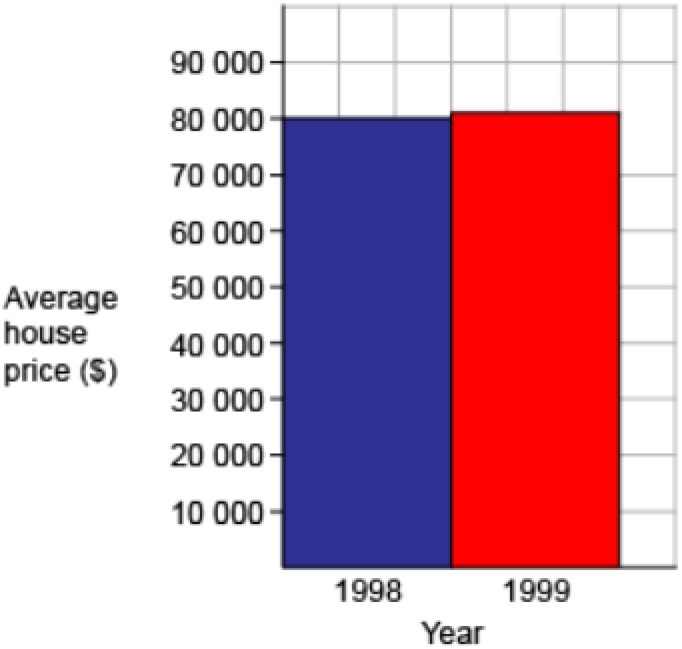
Biased labeling





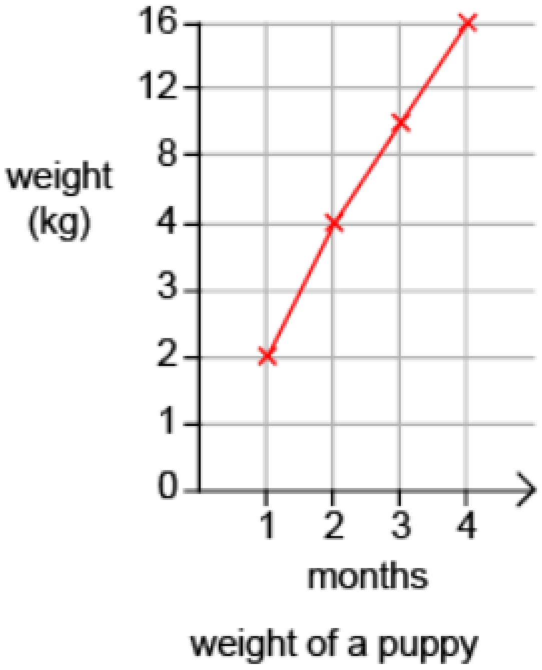
Truncated axis

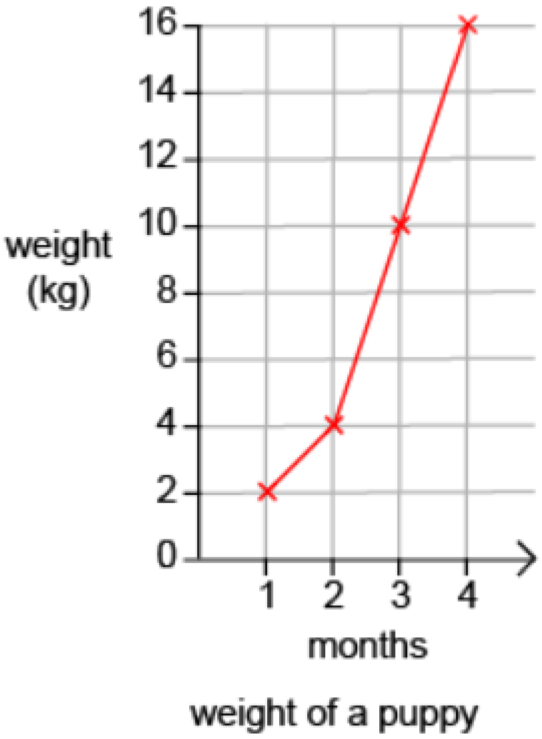




https://topdrawer.aamt.edu.au/Statistics/Misunderstandings/Misleading-graphs/Misleading-scales

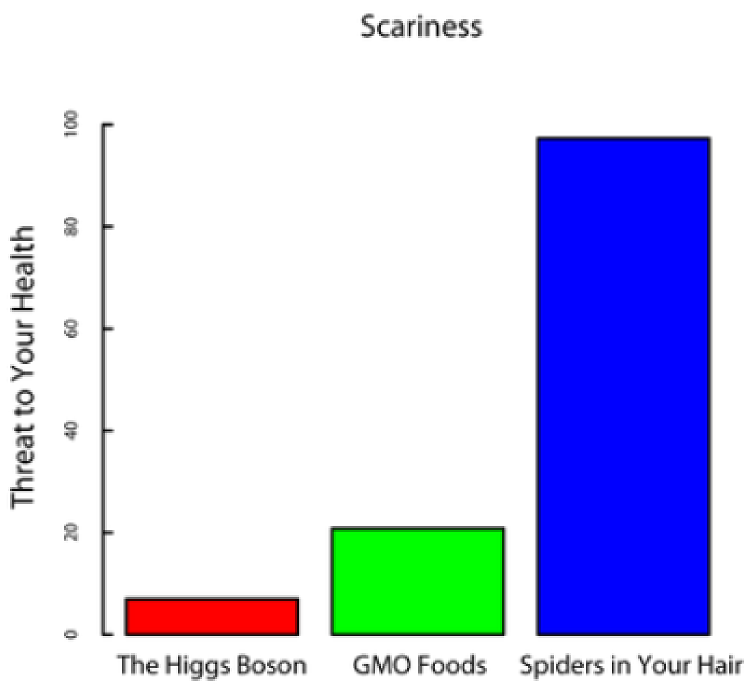
Improper intervals/scales



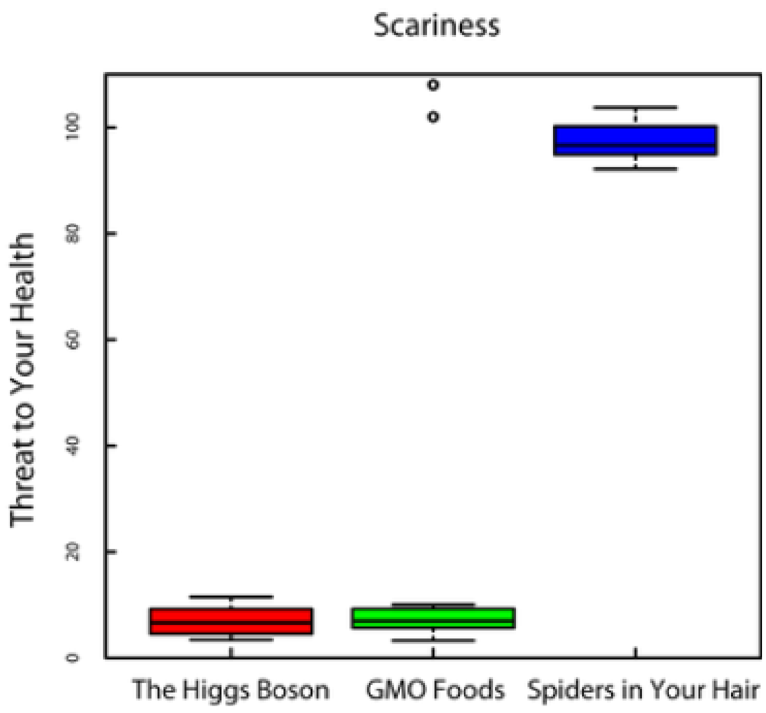


https://topdrawer.aamt.edu.au/Statistics/Misunderstandings/Misleading-graphs/Misleading-scales

Misleading graphs



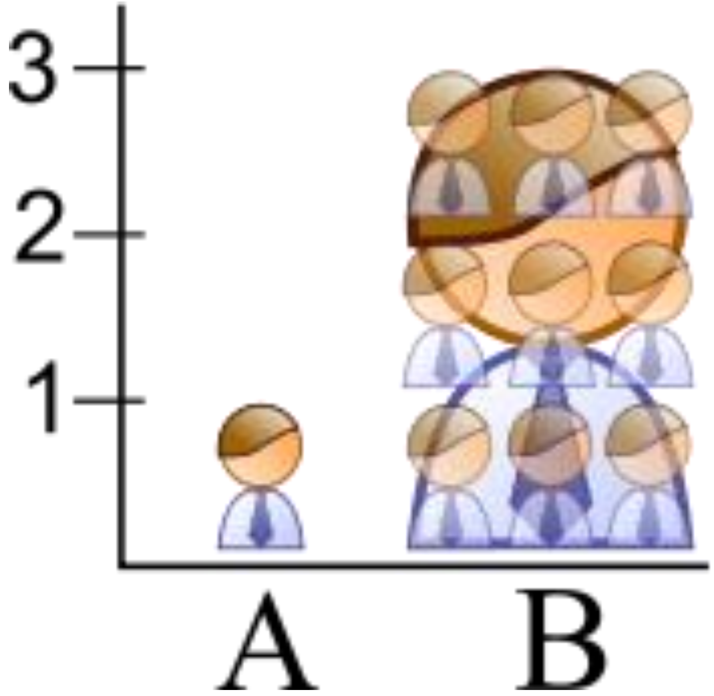
Bar Graph



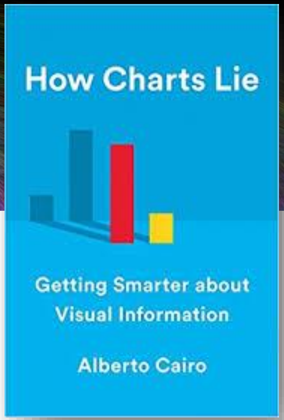
Box Plot

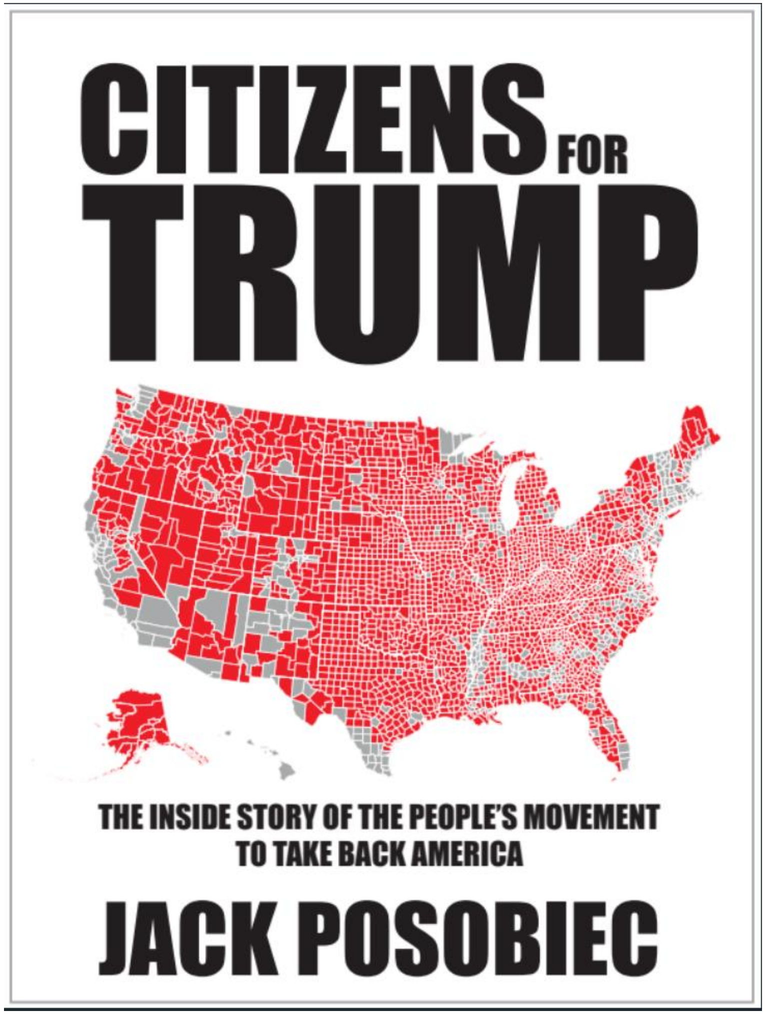
https://figureoneblog.wordpress.com/2014/03/12/misleading-with-pictures-the-pitfalls-of-data-visualization/

Misleading pictographs

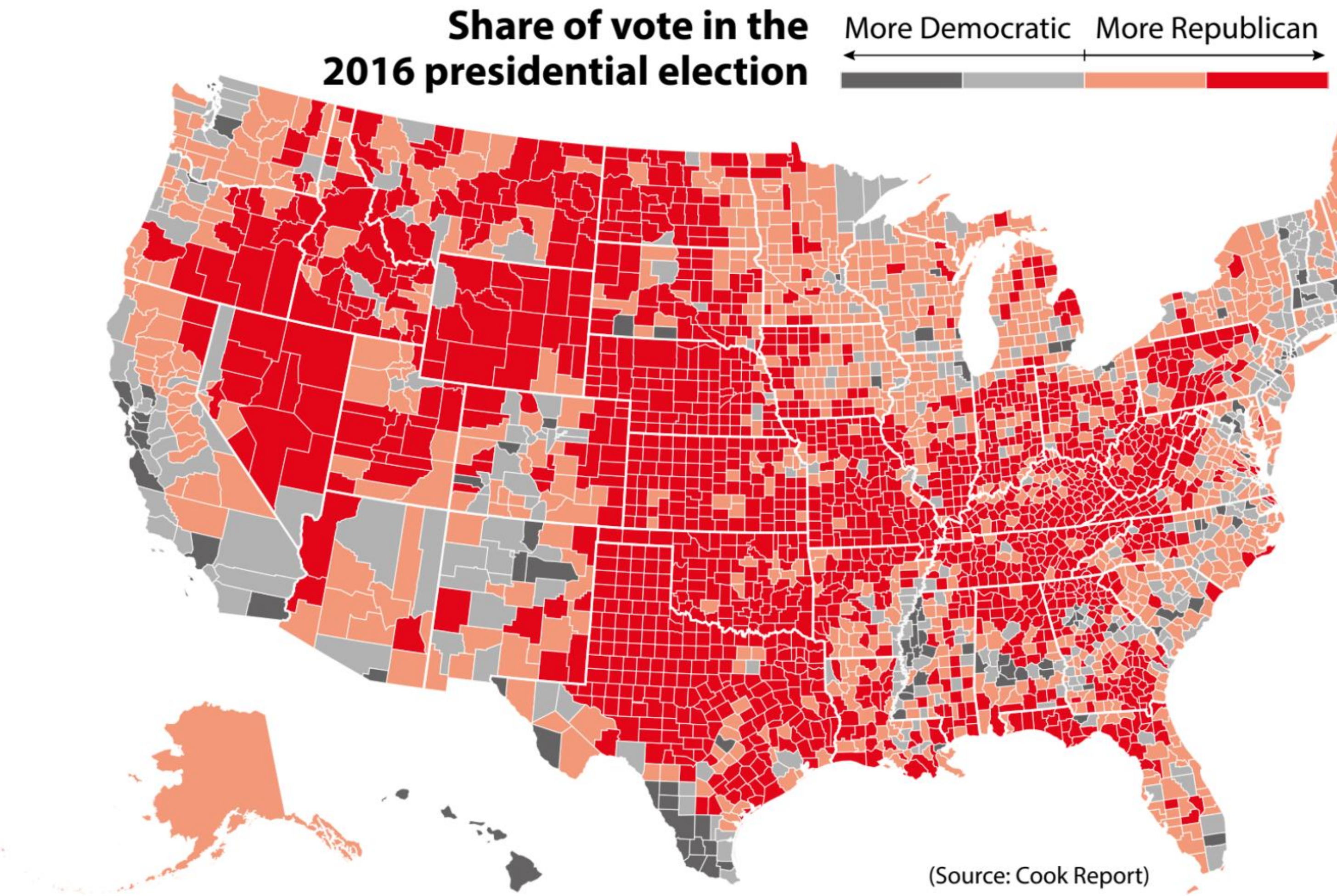


Misleading maps



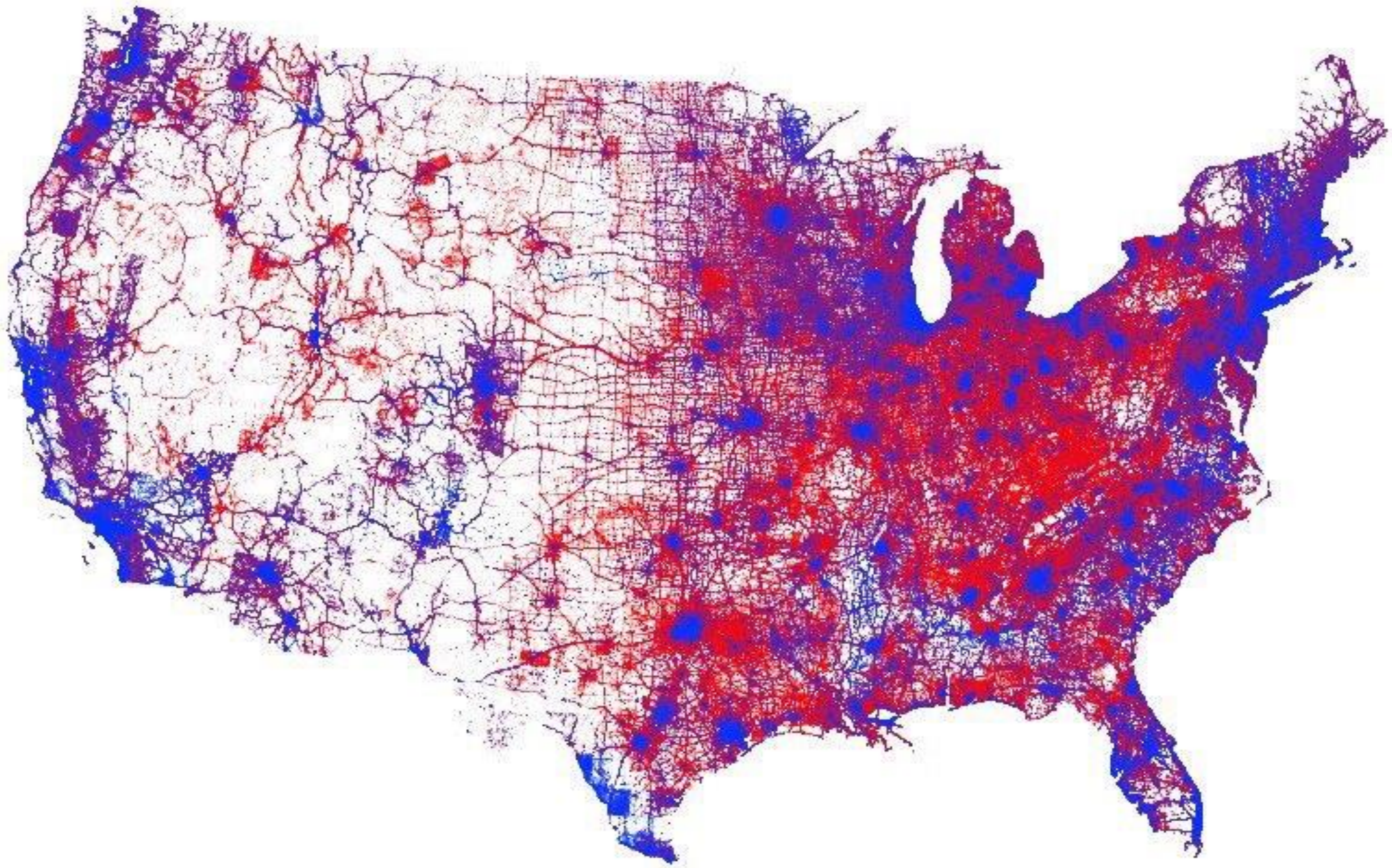


Misleading maps



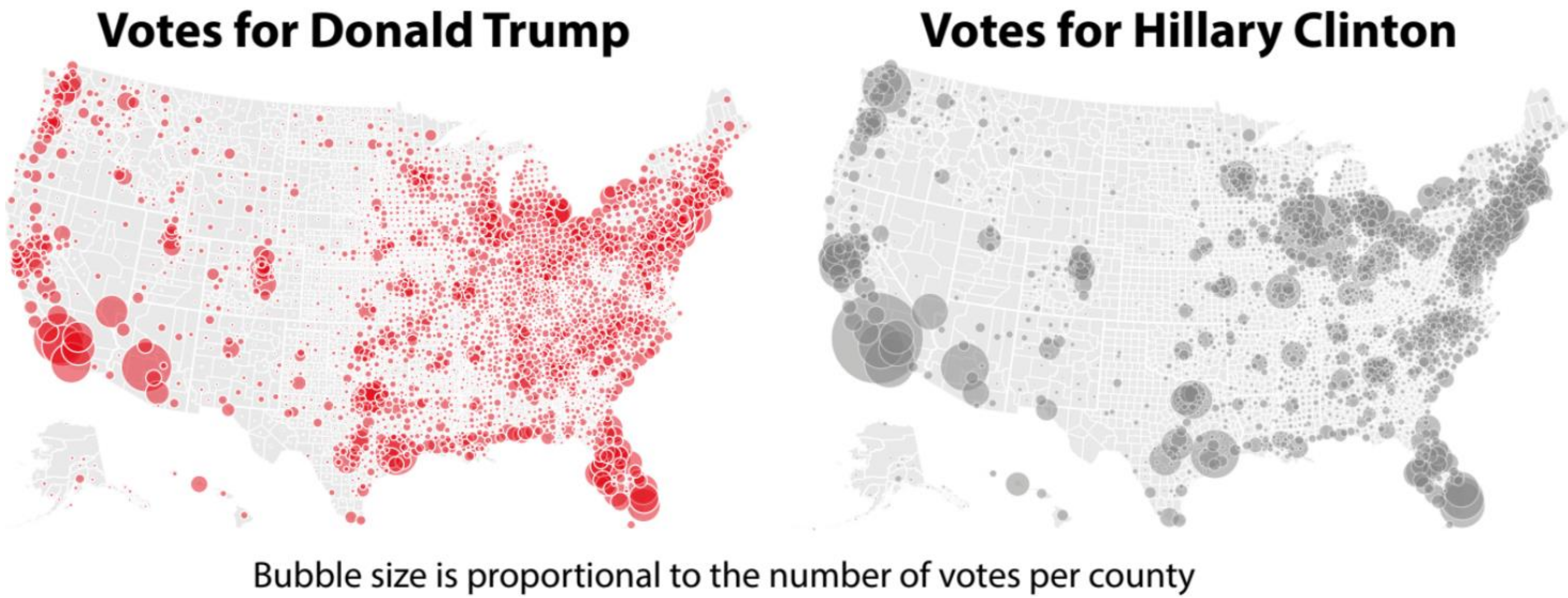
Misleading maps

Dasymetric Dot Density



http://cartonerd.blogspot.com/2018/03/dotty-election-map.html

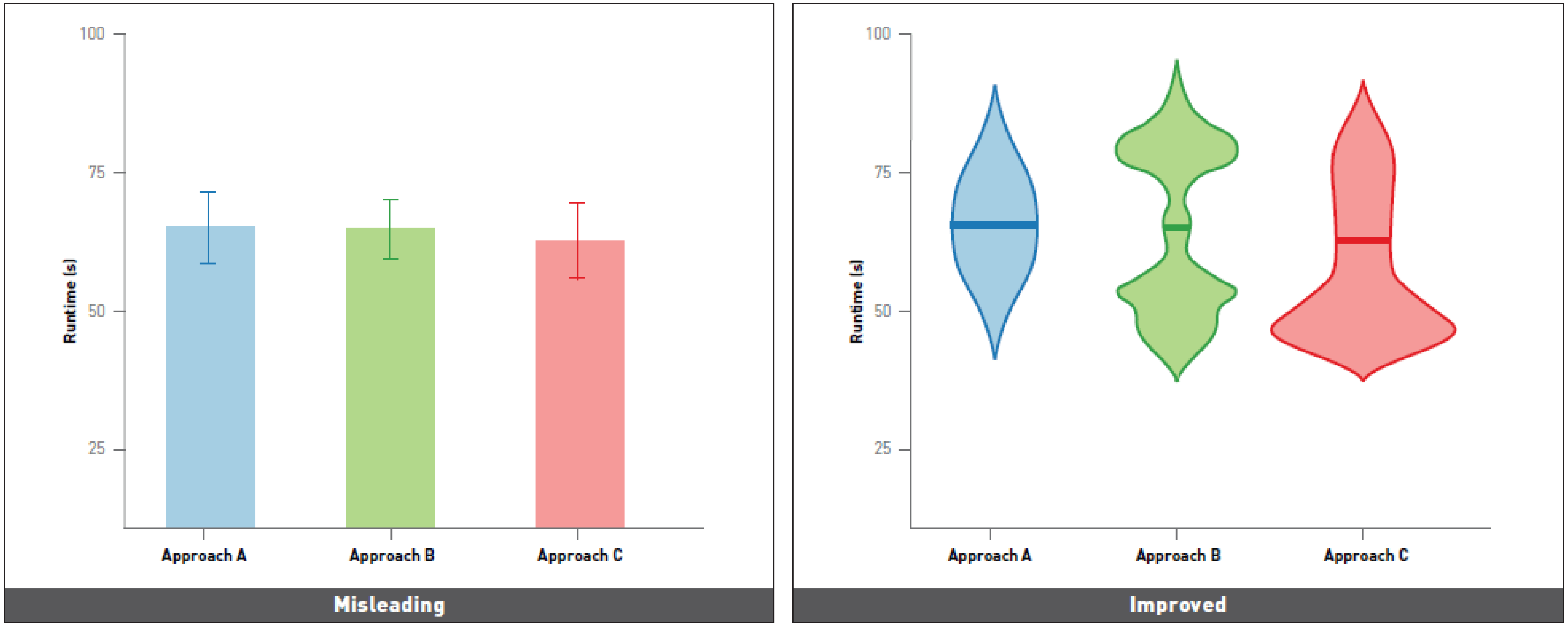
Misleading maps



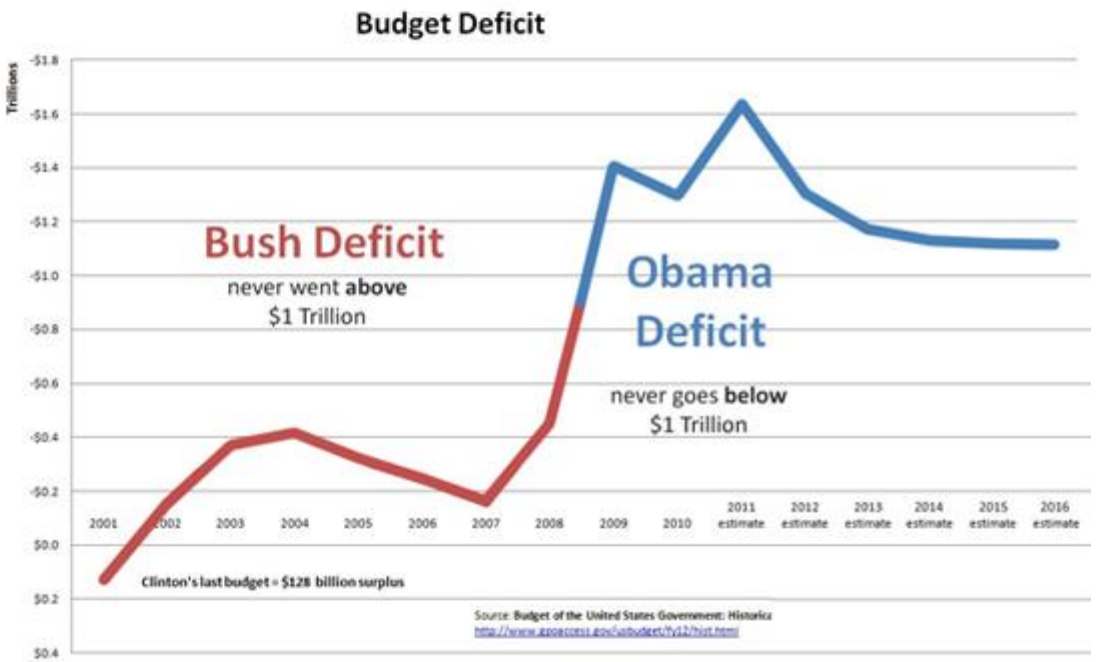
… and much more

https://www.nytimes.com/interactive/2020/10/30/opinion/election-results-maps.html

Misleading graphs

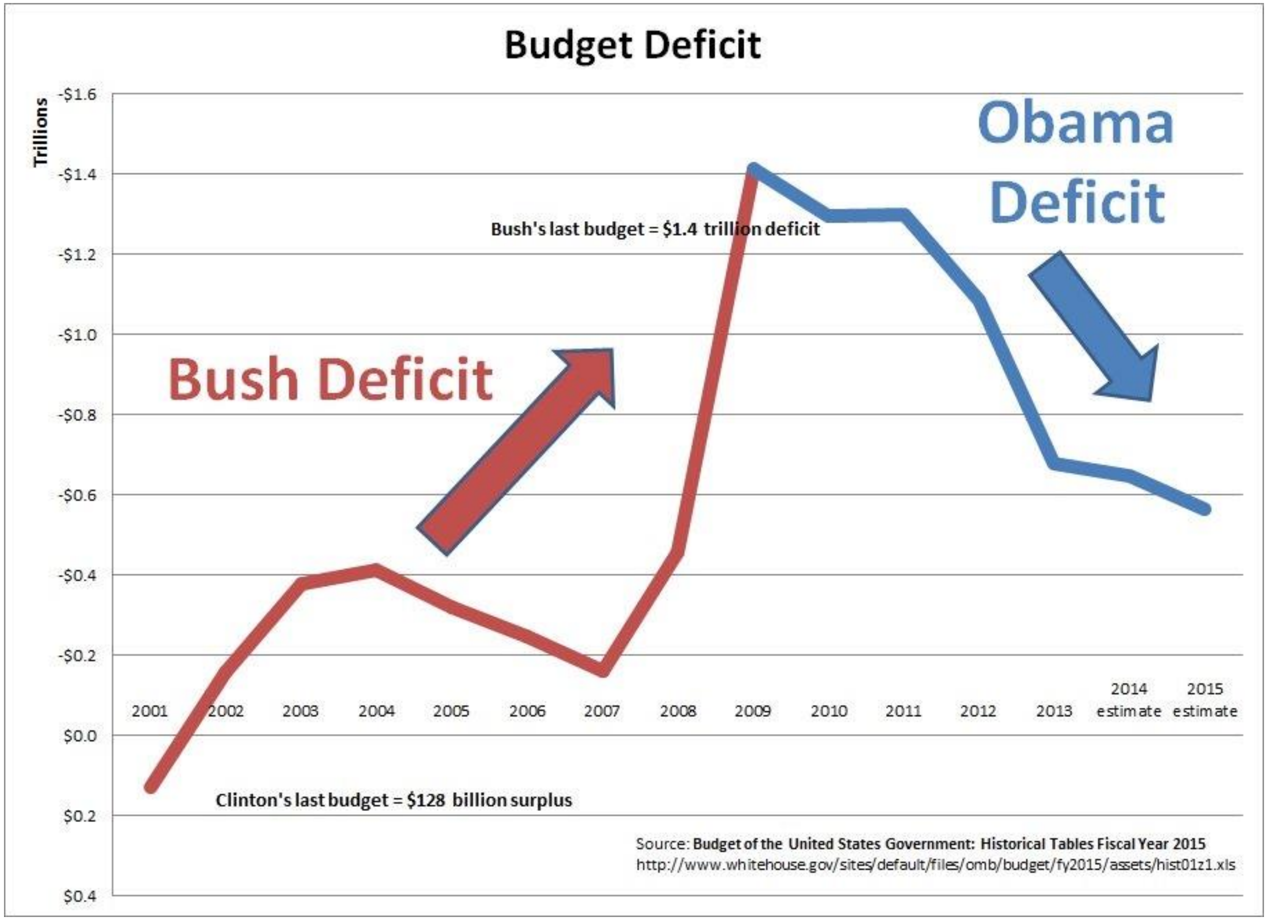


Cultural Bias



Signal value issues. The red line feels more negative. From http://www.politicalmathblog.com

Cultural Bias



– Benjamin Haydon

“FORTUNATELY FOR SERIOUS MINDS, A BIAS RECOGNIZED IS A BIAS STERILIZED.”

Ask yourself

- Why am I seeing what I'm seeing?
- Why am I interpreting it the way I am interpreting it?
- Have I looked at all related data?

Recent Research

COGNITIVE BIAS IN INTERACTIVE VISUAL ANALYTICS

Cognitive Bias in Interactive Visual Analytics

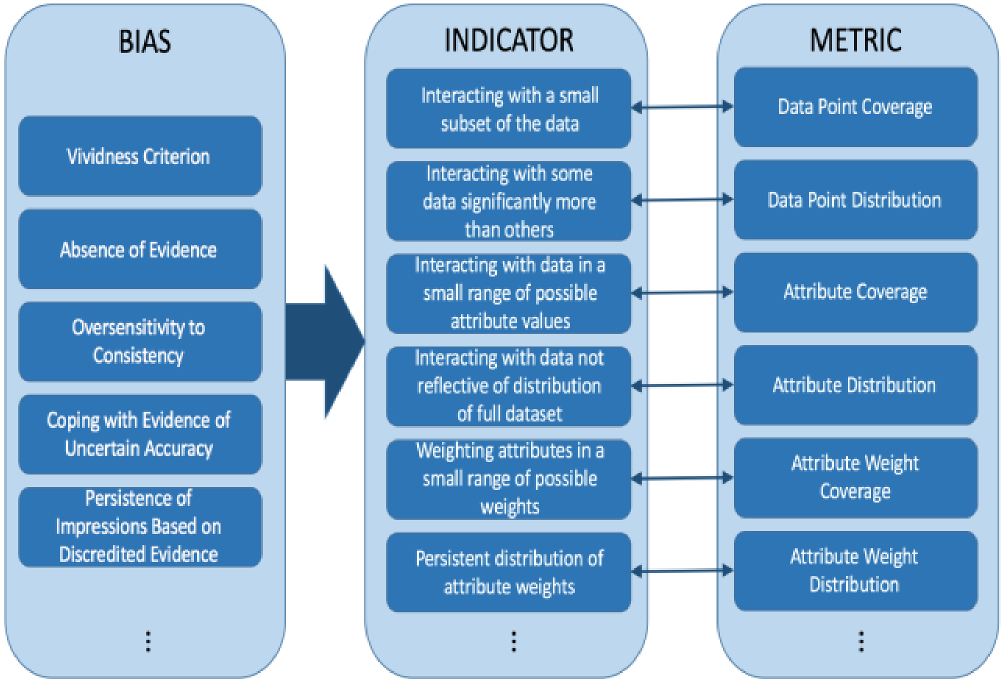


Figure 1: Cognitive biases result in behavioral indicators that are measurable by the proposed metrics. We scope this paper to those indicators and metrics depicted above, but there are numerous other biases, behavioral indicators, and ways to measure those indicators.

Wall et al. (2017), *Warning , Bias May Occur : A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics, VAST symposium, URL: https://www.cc.gatech.edu/~ewall9/media/papers/BiasVAST17.pdf*

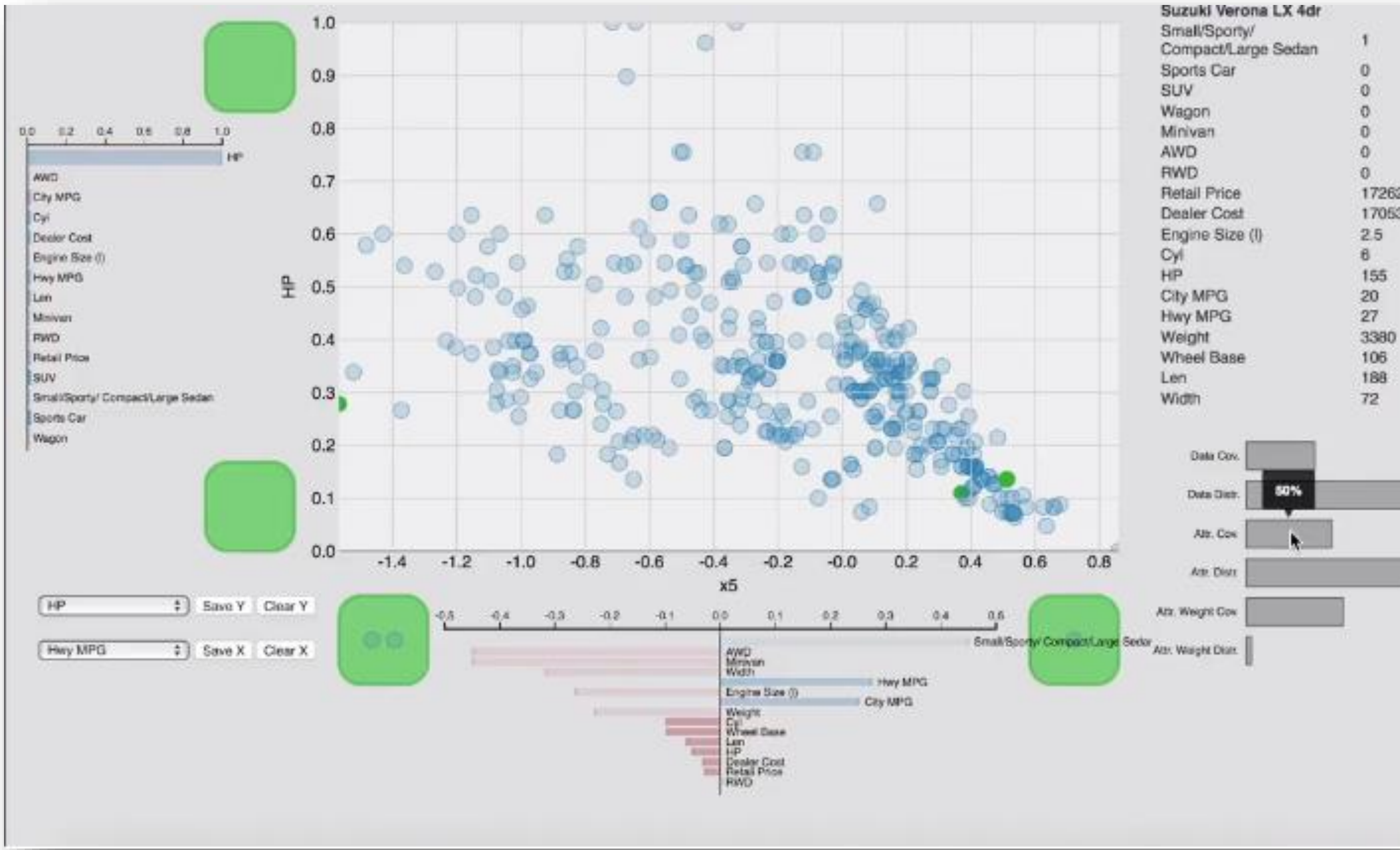
Cognitive Bias in Interactive Visual Analytics

Bias	Description	Interaction Manifestation
Vividness Criterion	humans rely more heavily on information that is specific or personal than information that is abstract or lacking in detail	e.g., analyst frequently returns to / interacts with data points that are rich in detail
Absence of Evidence	humans tend to focus their attention on the information that is present, ignoring other significant pieces of evidence that may be missing	e.g., analyst filters out a subset of data, forgets about it, and makes future decisions without accounting for the missing data
Oversensitivity to Consistency	humans tend to choose hypotheses that encompass the largest subset of evidence	e.g., analyst interacts almost exclusively with data that supports the largest encompassing hypothesis, dismissing other data
Coping with Evidence of Uncertain Accuracy	humans tend to choose to accept or reject a piece of evidence wholly and seldom account for the probability of its accuracy	e.g., analyst filters out data that supports a seemingly unlikely hypothesis, thus fully rejecting it
Persistence of Impressions Based on Discredited Evidence	humans tend to continue to believe information even after it has been discredited (also known as the *continued influence effect*)	e.g., analyst continues to interact with data supporting a hypothesis that has been disproved

Table 1: Cognitive biases relevant to intelligence analysis [38] that produce the measurable behavioral indicators we focus on in this paper.

In R. J. Heuer Jr. *Psychology of Intelligence Analysis*. Washington, DC, 1999

Detecting Cognitive Bias



https://www.youtube.com/watch?v=W9LWi3oXjM0

Wall et al. (2017), *Warning , Bias May Occur : A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics, VAST symposium, URL: https://www.cc.gatech.edu/~ewall9/media/papers/BiasVAST17.pdf*