Visualisation and Uncertainty

Jana Kleineberg | Polina Levontin | Jo Lindsay Walton

The current state of knowledge

"Even if the effects of visualizing uncertainty and its influence on reasoning are not fully understood, it has been shown that the **graphical display of uncertainty has positive effects on performance**."

Riveiro et al. (2014)

"**We do not have a comprehensive understanding** of the parameters that influence successful uncertainty visualization, nor is it easy to determine how close we are to achieving such an understanding."

MacEachren et al. (2005)

MacEachren et al. (2005), Riveiro et al. (2014),

The current state of knowledge

- ➔ Because we lack a comprehensive framework of how visualisation works, it is important to approach visualisation on a case-by-case basis
- ➔ There are many fancy and interesting visualisations out there whose effectiveness has not been studied at all
- ➔ Even the case-by-case approach doesn't guarantee success: it's possible for a designer and a client to fall in love with a visualisation and fail to see that it is not doing its job effectively

In this clinic we hope to ...

- ➔ Review some **current research** around visualising uncertainty
- ➔ Discuss **sample visualisations** and visualisation techniques
- ➔ **Share** one another's experience from our various fields
- ➔ Raise awareness of some **pitfalls** -- how can visualisation itself be a source of uncertainty?
- ➔ Try some open-ended **exercises** as an additional way of sharing experience and perspectives
- ➔ Briefly look at some more **experimental** approaches

Visualisation may impact ...

1. Decision outcomes
2. Correctness of decisions
3. Kinds of errors made
4. Decision time
5. Confidence in a decision
6. Willingness to make a decision
7. How much workload decision-making causes
8. How a decision is made

Effects are **case-specific**, and are impossible to predict in advance without empirical evidence and testing with intended audiences

Kinkeldey et al. (2015)

Biased self-perception

1. Decision outcomes
2. Correctness of decisions
3. Kinds of errors made
4. Decision time
5. Confidence in a decision
6. Willingness to make a decision
7. How much workload decision-making causes
8. How a decision is made

These impacts are not correlated as we might expect

Several studies indicate that **confidence** is not correlated with the **correctness.** In general, self-perception is subject to a variety of biases.

Kinkeldey et al. (2015)

Where do we encounter visualisation?

- ➔ **Where do you encounter visualisation in your professional work?**
- ➔ Do these visualisations have **drawbacks** and downsides?
- ➔ How can visualisations *themselves* be a source of uncertainty?
- ➔ How can visualisations effectively communicate their own limitations?
- ➔ What would you ideally *like* visualisation to be able to do that it usually doesn't?
- ➔ Feel free to define "visualisation" as broadly as you like … any graphical presentation of information could count

Methodological problems with obtaining knowledge on visualisations

- ➔ Small sample sizes for audience testing
- ➔ Inappropriate subjects (students rather than decision-makers)
- ➔ Reproducibility
- ➔ Small effect sizes
- ➔ Transferability issues
- ➔ Individual, cultural, and other differences that are difficult to control for
- ➔ Aesthetic appreciation of data may be misleading
  (“Yes, I love this visualization! Now here is my terrible decision based on it.”)
- ➔ Biased self-perception (“Like the visualization, my decision was amazing.”)

Where do we visualise uncertainty?

- **Where do you encounter uncertainty in your professional lives?**
- Of course, there are many kinds of uncertainty, so you may want to focus on the kinds which you could potentially do something about
- How is visualisation *currently* used to identify, understand, manage, and communicate this uncertainty?
- How *could* visualisation be used to deal with uncertainty? Feel free to be imaginative / speculative …
- And if you have time: how does uncertainty in your professional lives relate to working with others? How might visualisations help to create shared frames of reference? Where do the dangers lie?

Types of uncertainty: some prompts ...

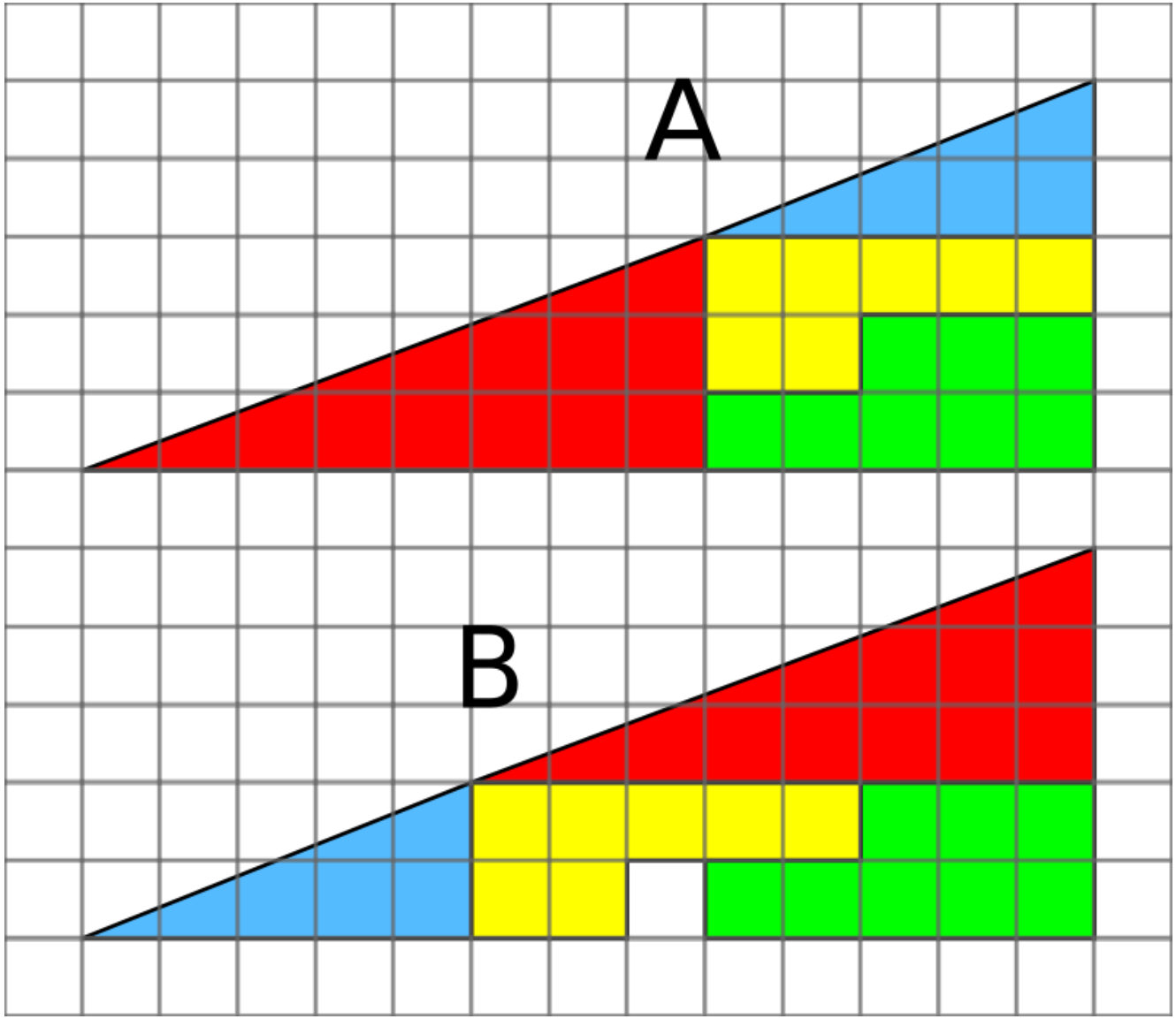
- ➔ **Accuracy** – the difference between observation and reality
- ➔ **Precision** – the quality of the estimate or measurement
- ➔ **Completeness** – the extent to which information is comprehensive
- ➔ **Consistency** – the extent to which information elements agree
- ➔ **Lineage** – the pathway through which information has been passed
- ➔ **Currency** – the time span from occurrence to information presentation
- ➔ **Credibility** – the reliability of the information source
- ➔ **Subjectivity** – the extent to which the observer influences the observation
- ➔ **Interrelatedness** – the dependence on other information
- ➔ **Experimental** – the width of a random distribution of observations
- ➔ **Geometric** – the region within which a spatial observation lies

Chung and Wark (2016)

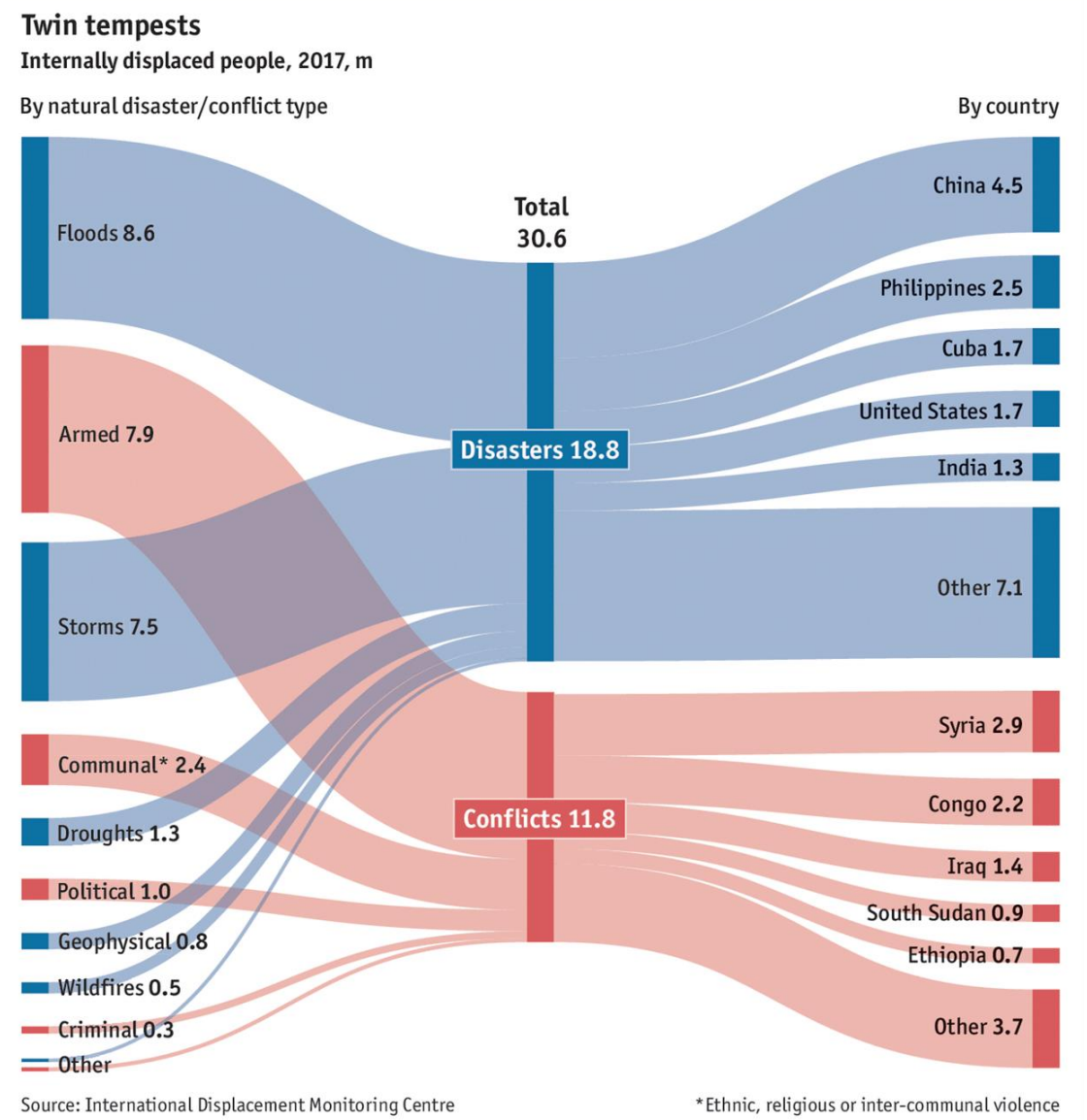
Uncertainties can be introduced at any stage during information processing

- ➔ **Acquisition**: introduced by the measurement or sampling processes
- ➔ **Transformation:** during processing
- ➔ **Visualisation**: introduced during visualisation process

Pang et al. (1996)



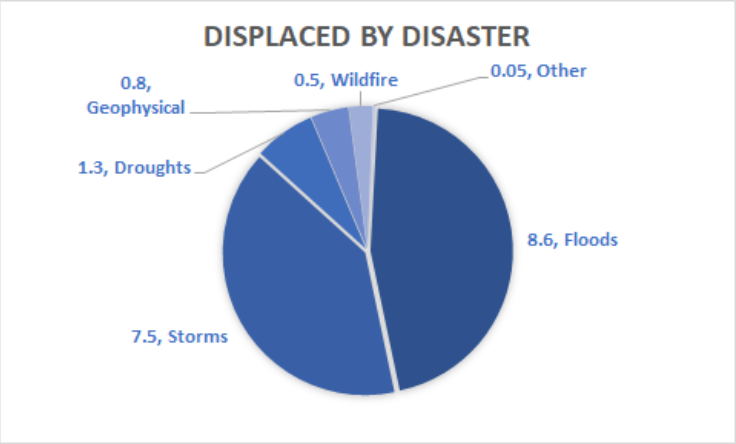
Missing square puzzle

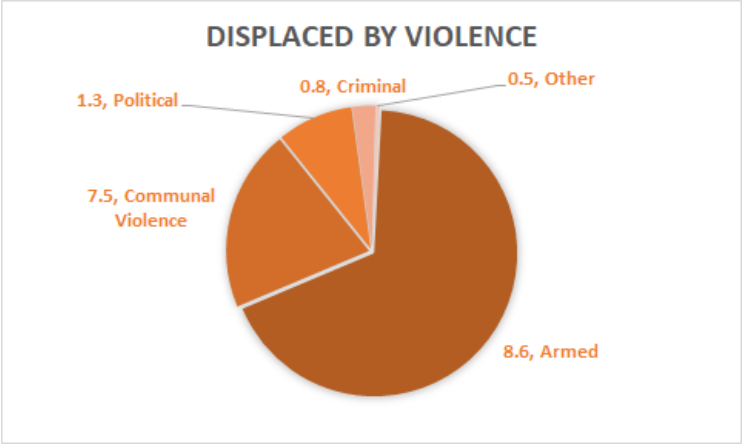

Twin tempests
Internally displaced people, 2017, m
By natural disaster/conflict type
By country
Floods 8.6
Armed 7.9
Storms 7.5
Communal* 2.4
Droughts 1.3
Political 1.0
Geophysical 0.8
Wildfires 0.5
Criminal 0.3
Other
Total
30.6
Disasters 18.8
Conflicts 11.8
China 4.5
Philippines 2.5
Cuba 1.7
United States 1.7
India 1.3
Other 7.1
Syria 2.9
Congo 2.2
Iraq 1.4
South Sudan 0.9
Ethiopia 0.7
Other 3.7
Source: International Displacement Monitoring Centre
*Ethnic, religious or inter-communal violence

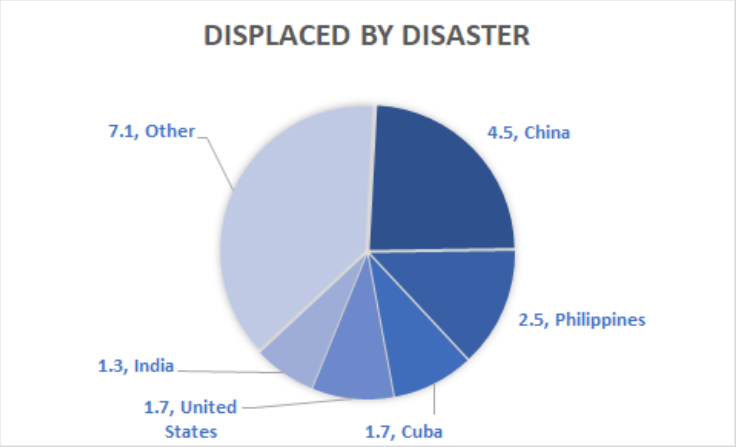
Twin tempests: analysis

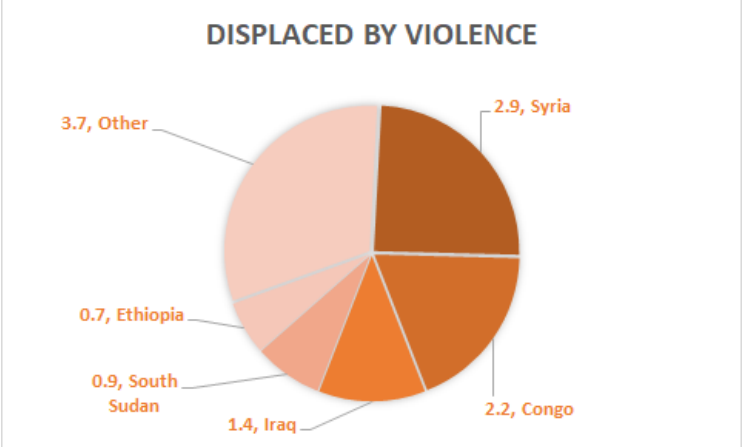
- ➔ Is this diagram visualising the number of displaced in Syria due to storms?
- ➔ Why is “other” split? Was this whole chart a big mistake?
- ➔ Also may give the impression e.g. “Floods” impacts China more than US
- ➔ The sense of motion may make the user assume that this is about migration across borders
- ➔ Tendency to read left-right (in the West), whereas perhaps this is best read from the centre
- ➔ If the underlying data is available, could a more detailed breakdown be represented? I.e. start with “Disasters” and “Conflicts” on the left?
- ➔ Association of Africa and Middle East with conflict
- ➔ Are these separate studies woven appropriately? Why this order? Are overlapping areas meaningful? Are there interactions among these causes?

Twin tempests: analysis









Assumptions for visualising uncertainty, and the challenge of deep uncertainty

"First, it is assumed that uncertainty, or at least uncertainty of interest, is both **knowable** and **identifiable**. Similarly, to be visualized, uncertainty must be **quantifiable**, such as through statistical estimates, quantitative ranges, or qualitative statements (e.g., less or more uncertain). Moreover, evaluations define effectiveness as an ability to identify specific uncertainty values, which assumes that **identifying specific** uncertainty values is useful to decision-makers and that the values of interest can be quantified. Lastly, there is an assumption that the quantification of uncertainty is **beneficial**, **applicable** to the decision task, and **usable** by the decision-maker, even if users do not currently work with uncertainty in that way. **These assumptions pose a challenge for visualizing uncertainty to support decision making under deep uncertainty**, where quantification of uncertainty is not possible or necessarily desirable. In this way, current approaches to uncertainty visualization are more normative in nature, reflecting what researchers think decision-makers need to know about uncertainty."

Deitrick and Wentz (2015), emphasis added

11-STEP STRATEGY *for* UNCERTAINTY VISUALIZATION DESIGN

1.IDENTIFY
the **UNCERTAINTY** visualisation task

2.UNDERSTAND
the **DATA** that is to be visualised

3.UNDERSTAND WHY
WHY uncertainty needs to be visualised and **HOW** it needs to help the user

4.DECIDE
which uncertainty should be visualized

5.DECIDE
on a definition of uncertainty

6.DETERMINE
the specific causes of the uncertainty

7.DETERMINE
the causal categories of the uncertainty

8.DETERMINE
the visualisation requirements

9.PREPARE
uncertainty for visualisation

10.TRY
different techniques

11.FEEDBACK
Get audience opinions and criticisms

Lapinski A.L. S. (2009), graphic reformatted

Graphic design principles: intrinsic vs. extrinsic techniques

Two general categories:

Intrinsic representation techniques
integrate uncertainty by varying the appearance of the data
(e.g. shape, texture, brightness, opacity, hue)

Extrinsic representation techniques
addition of geometry to describe uncertainty
(e.g. arrows, error bars, charts)

Pang et al. (1996)

Graphic design principles: five intuitive approaches

- ➔ Modification of **graphical attribute**
  (colour, texture, blurring, opacity)
- ➔ Addition of **artefacts**
  (icons, glyphs, contours, iso-surfaces)
- ➔ **Acoustic** and **haptic** feedback
  (coupled with a visual display)
- ➔ **Animation** of graphical attributes
  (e.g. speed, duration, motion blur, range, or extent of motion)
- ➔ **Dynamic** displays
  (interactivity, e.g. additional information on mouse-over)

Chung and Wark (2016)

Graphic design principles: areas to consider

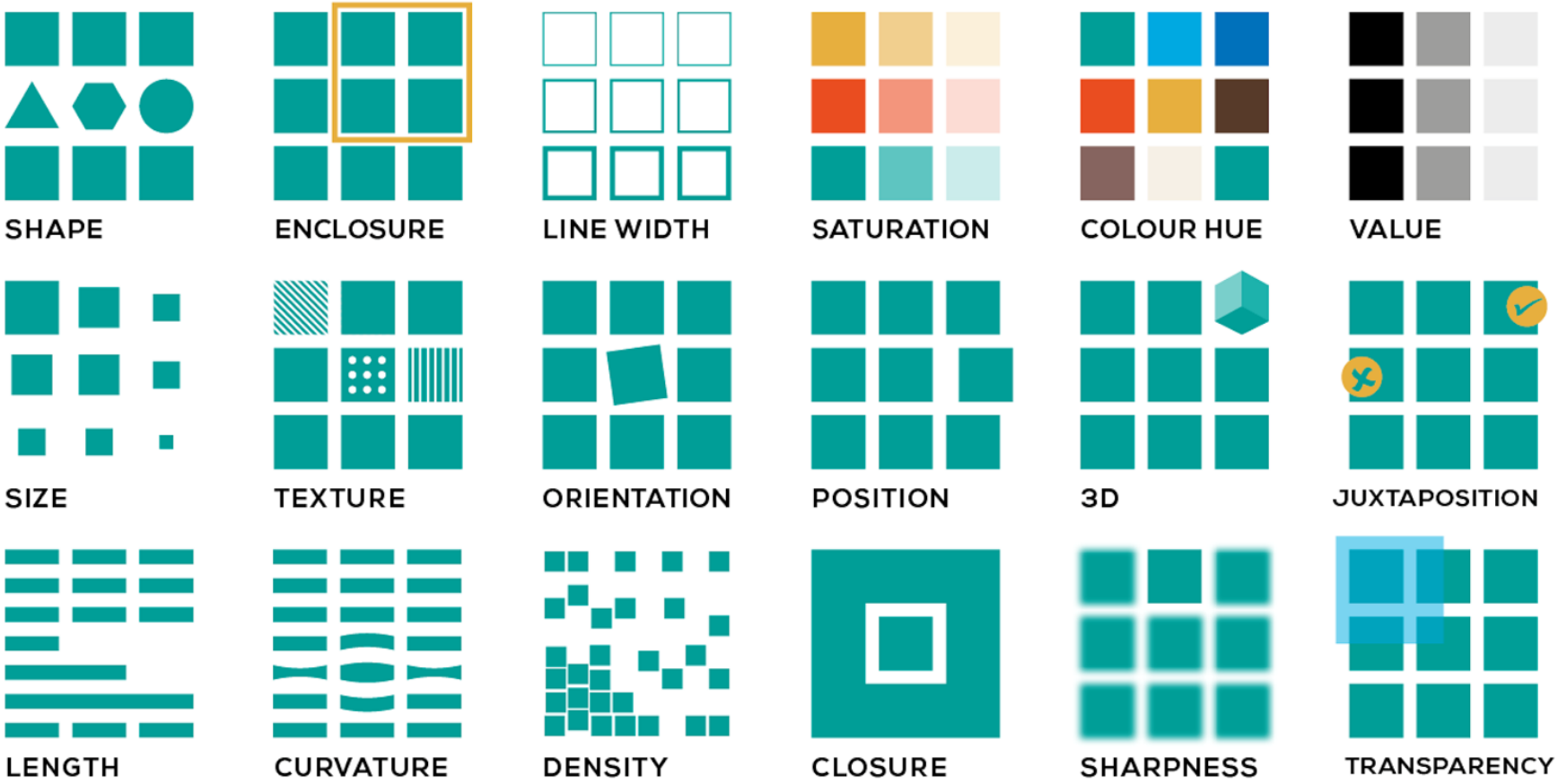
Effective graphic design takes account of:

- ➔ the **specific task** at hand (Hegarty, 2011)
- ➔ **expressiveness** of the display (Kosslyn, 2006)
- ➔ data-ink **ratio** (Tufte, 2001)
- ➔ issues of **perception** (Kosslyn, 2006; Tversky, Morrison, & Betrancourt, 2002; Wickens & Hollands, 2000)
- ➔ **pragmatics** of the display, including making the most relevant information salient (Bertin, 1983; Dent, 1999; Kosslyn, 2006)

It also takes account of semantics:

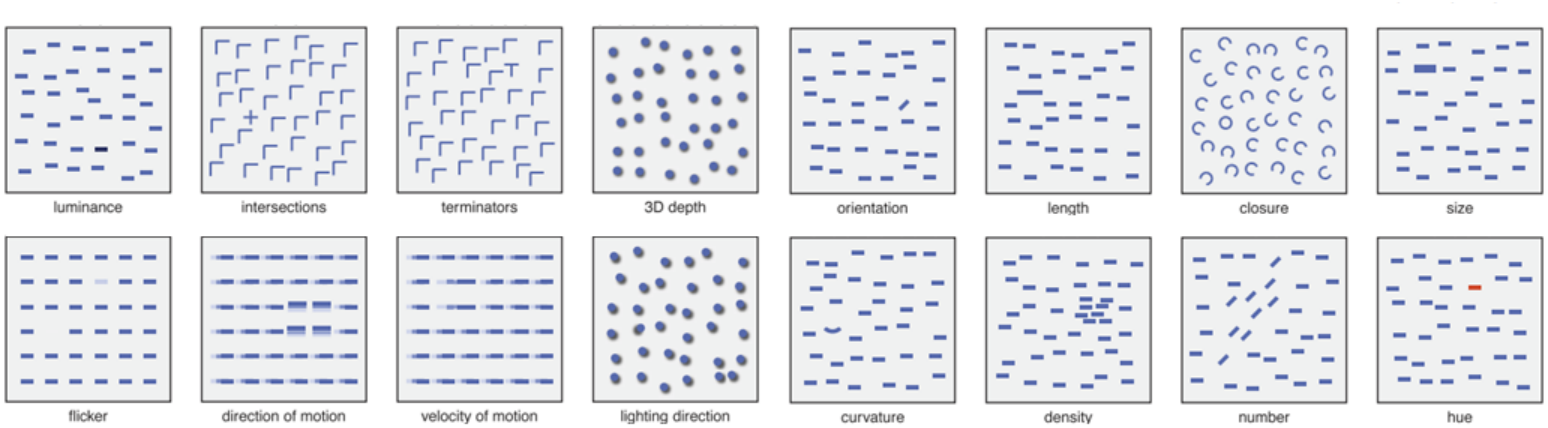
- ➔ **compatibility** between the form of the graphic and its meaning (Bertin, 1983; Kosslyn, 2006; Zhang, 1996)
- ➔ **usability** of the display, such as including appropriate knowledge (Kosslyn, 2006)

ATTRIBUTES *of* GRAPHICS



Graphic design principles: visual salience

- ➔ **Visual salience** refers to the conspicuousness of a visual element relative to its surroundings (Itti & Koch, 2001). Visual salience relies on so-called 'pre-attentive' visual processing. **Colour** (hue) is one well-known example.
- ➔ Other visual features that have been identified as pre-attentive: length, width, size, curvature, number, terminators, intersection, closure, intensity, flicker, direction of motion, binocular lustre, stereoscopic depth, 3D depth cues, and lighting direction.



itti & Koch (2001)

PRINCIPLES *of* DESIGN AESTHETICS

PATTERN

CONTRAST

EMPHASIS

BALANCE

RYTHM

PROPORTION

VARIETY

UNITY

MOVEMENT

HARMONY

Graphicacy or the principle of appropriate knowledge

- ➔ Perhaps attractively designed graphics and interfaces are perceived as having more value, and therefore capable of encoding more values?
- ➔ The effectiveness of a graphic is certainly influenced by the knowledge the viewer has about **the conventions of the graphic type**.
- ➔ Kosslyn (2006) called this the **principle of appropriate knowledge**.

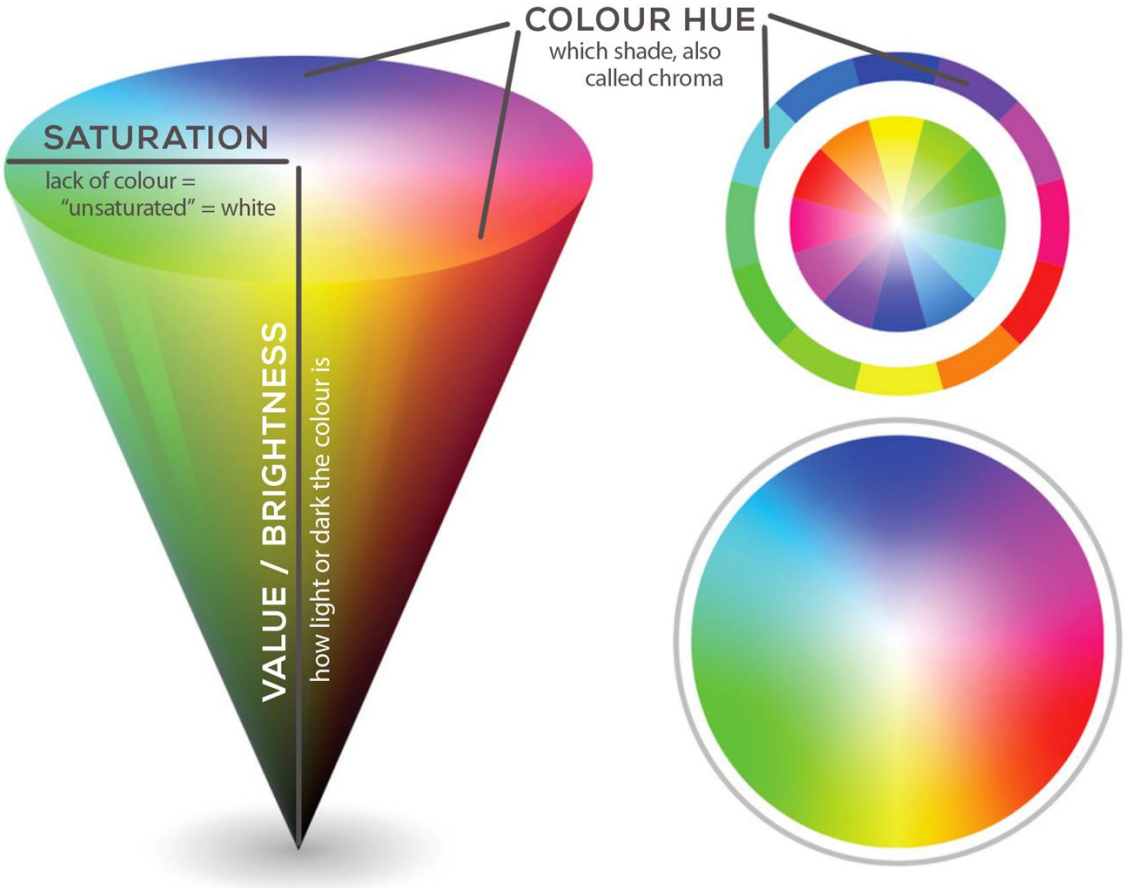
Kosslyn (2006), Boone et al. (2018)

Appropriate knowledge

Maps are the most obvious example of encoded visual variables

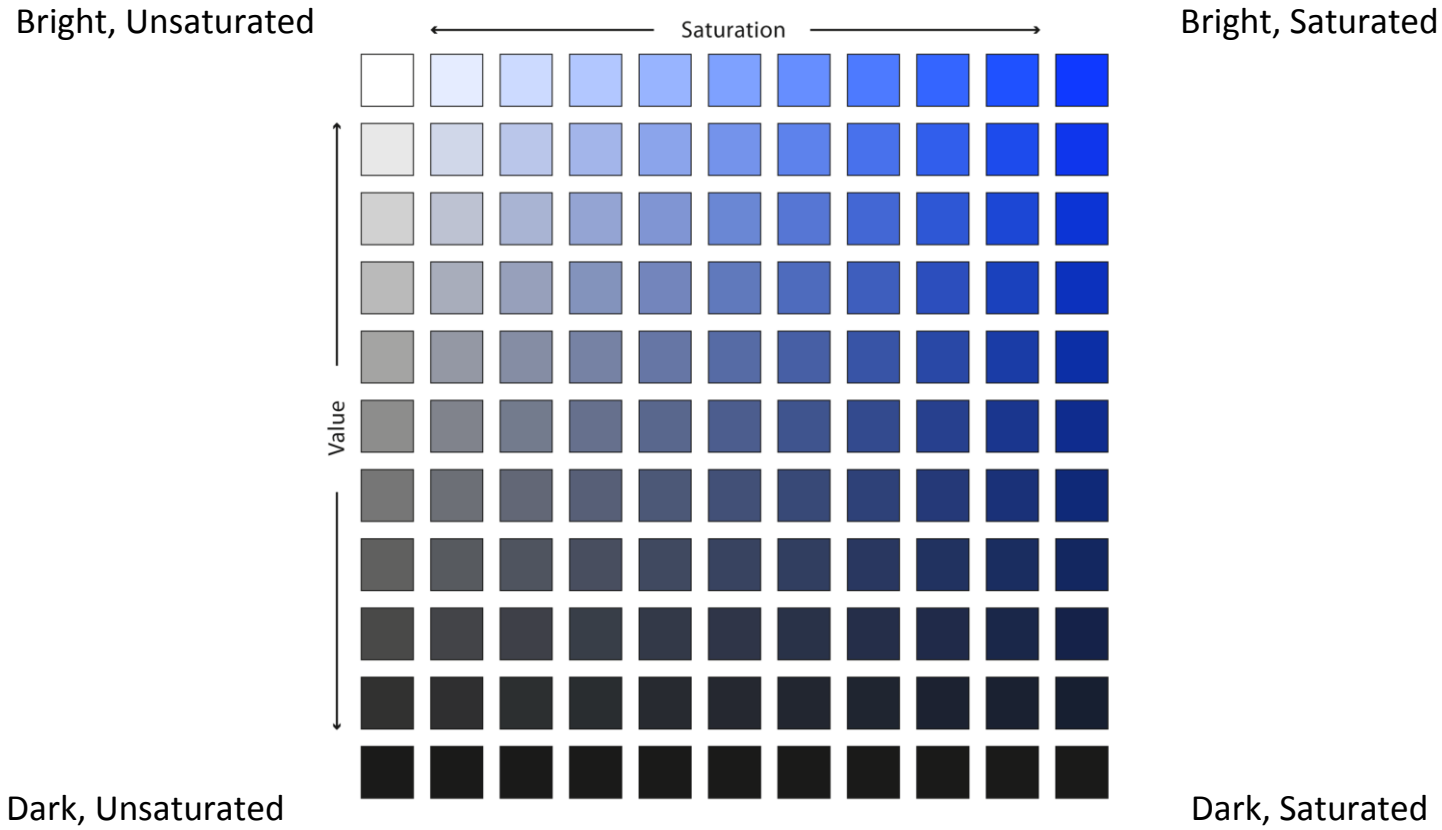
Graphicacy or the principle of appropriate knowledge

- ➔ The conventions of the display are typically encoded in a legend expressing the correspondence between visual variables and their meaning.
- ➔ Who is the audience for the visualisation? What training and experience will they have? How will the visualisation be curated?
- ➔ **Could the user believe they have appropriate knowledge when they don't?** How does the visualisation signal that it is being interpreted correctly or incorrectly?



Graphical attributes: colour

Hue | value | saturation | tint | chroma | lightness



Graphical attributes: colour

Value (AKA brightness, luminosity) vs. saturation

Colour hue was explored by MacEachren (1992) for representing uncertainty, suggesting that this technique is best used for novice users.

The evidence for the usability of **colour value** (also called lightness, brightness, or luminosity) is conflicting. Some experiments indicate value is *not effective* (Schweizer and Goodchild 1992) and others indicating *the converse* (Leitner and Buttenfield 2000, Aerts et al. 2003).

Where it is effective, *darker values are typically associated with more certainty, and lighter values with more uncertainty* (MacEachren 1992, Buttenfield 1993, McGranghan 1993, Van Der Wel et al. 1994).

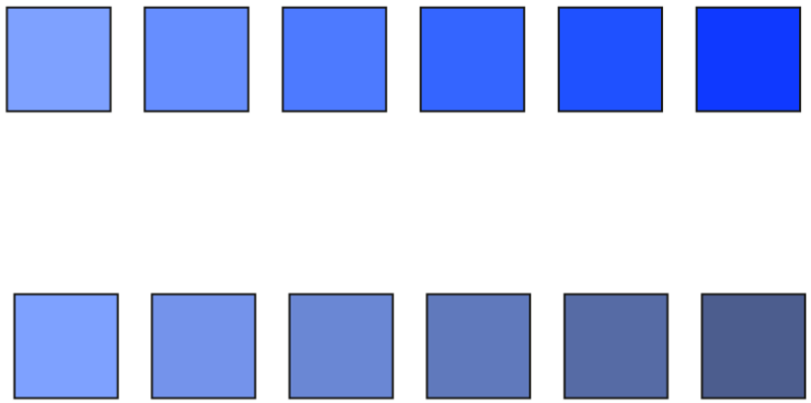
Saturation has been found not to be particularly effective or suitable for representing uncertainty (Buttenfield 1993, MacEachren et al. 2012).

Graphical attributes: colour

Hue | value | saturation

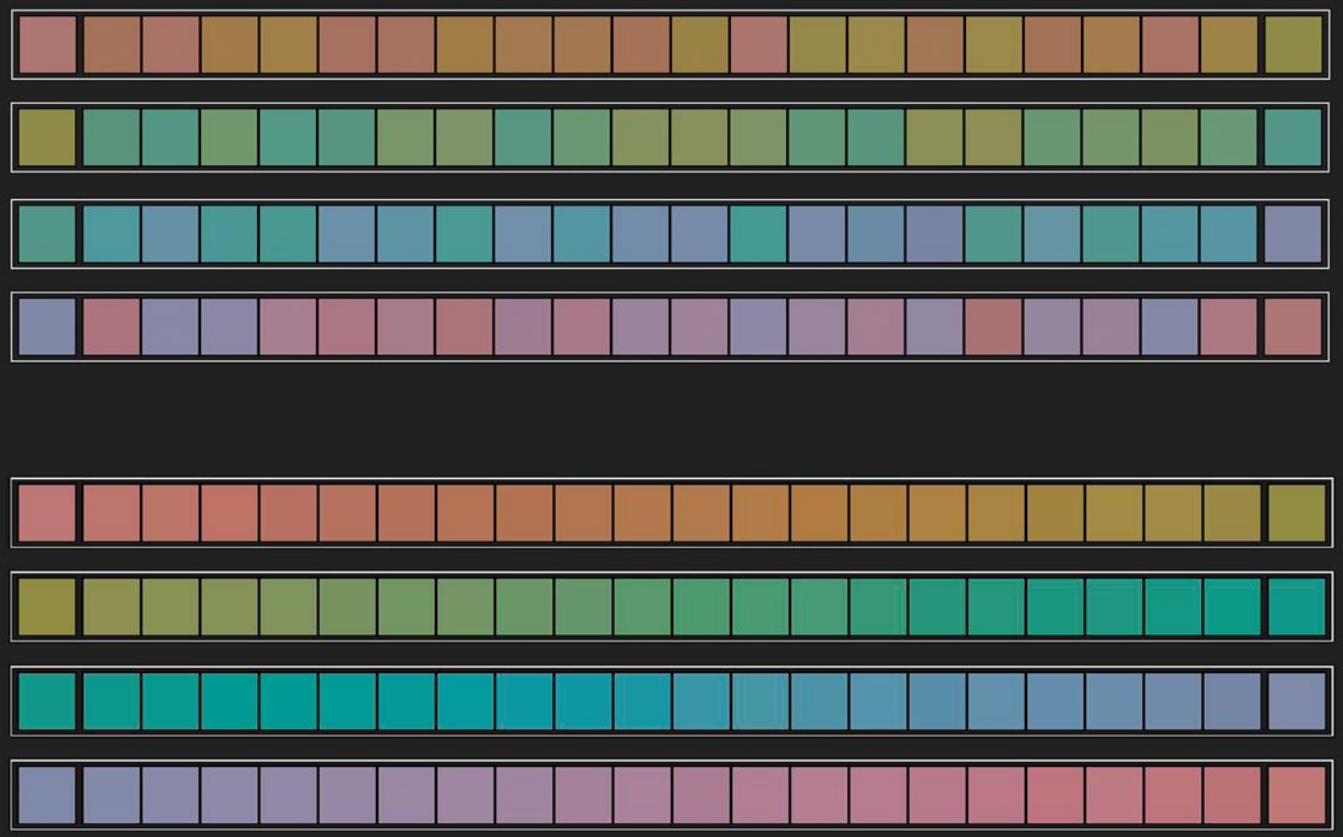
The evidence for the usability of **colour value** is conflicting. Where it is effective, *darker values are typically associated with more certainty*.

Saturation has been found not to be particularly effective or suitable for representing uncertainty.



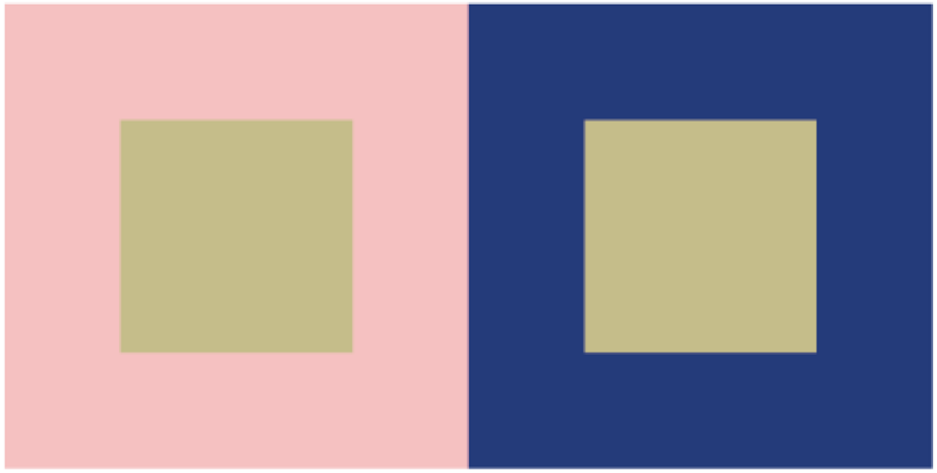
Graphical attributes: colour

Value (AKA brightness, luminosity) vs. saturation



Graphical attributes: colour value

Can be difficult to differentiate

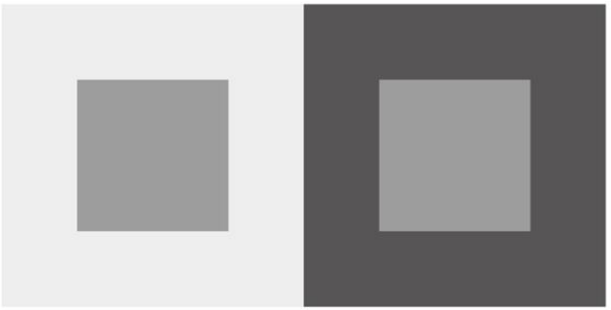


Colour phenomena

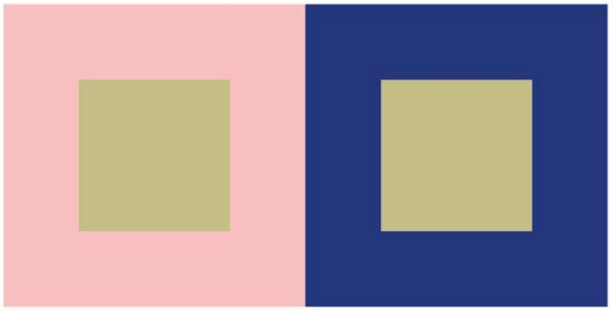
Optical illusions | simultaneous contrast

Colour phenomena

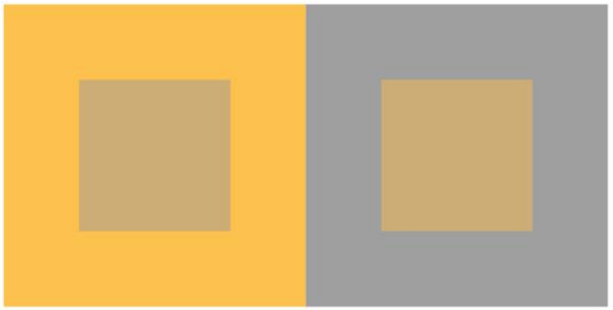
Optical illusions | simultaneous contrast



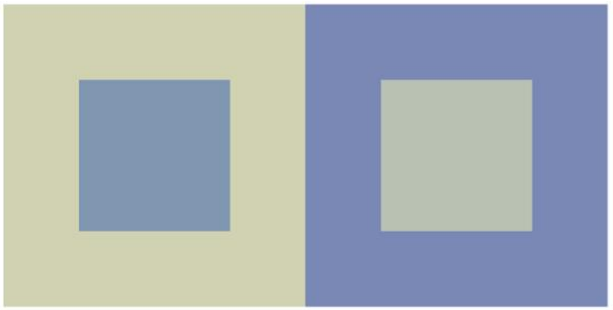
Surround a colour with a lighter color and it will appear darker; surround a color with a darker color and it will appear lighter.



Surround a colour with different hues and it will shift in appearance towards the complementary hue of the surrounding color.



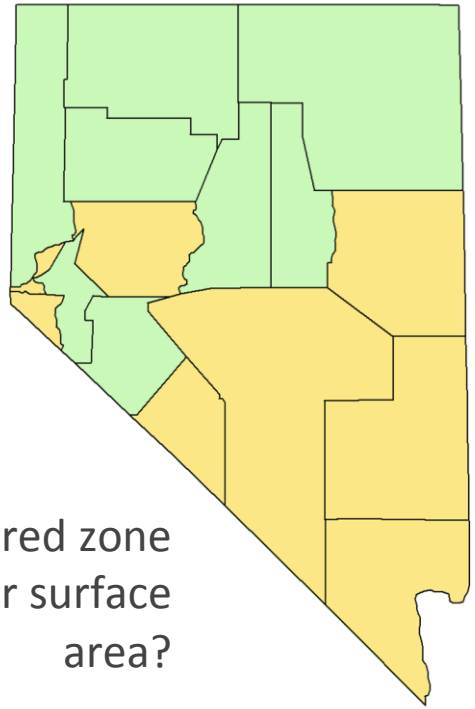
Surround a colour with a less saturated color and it will appear more saturated; with a more saturated color it will appear less saturated.



You can surround two different colours with two other colours to make them appear more similar.

Colour phenomena

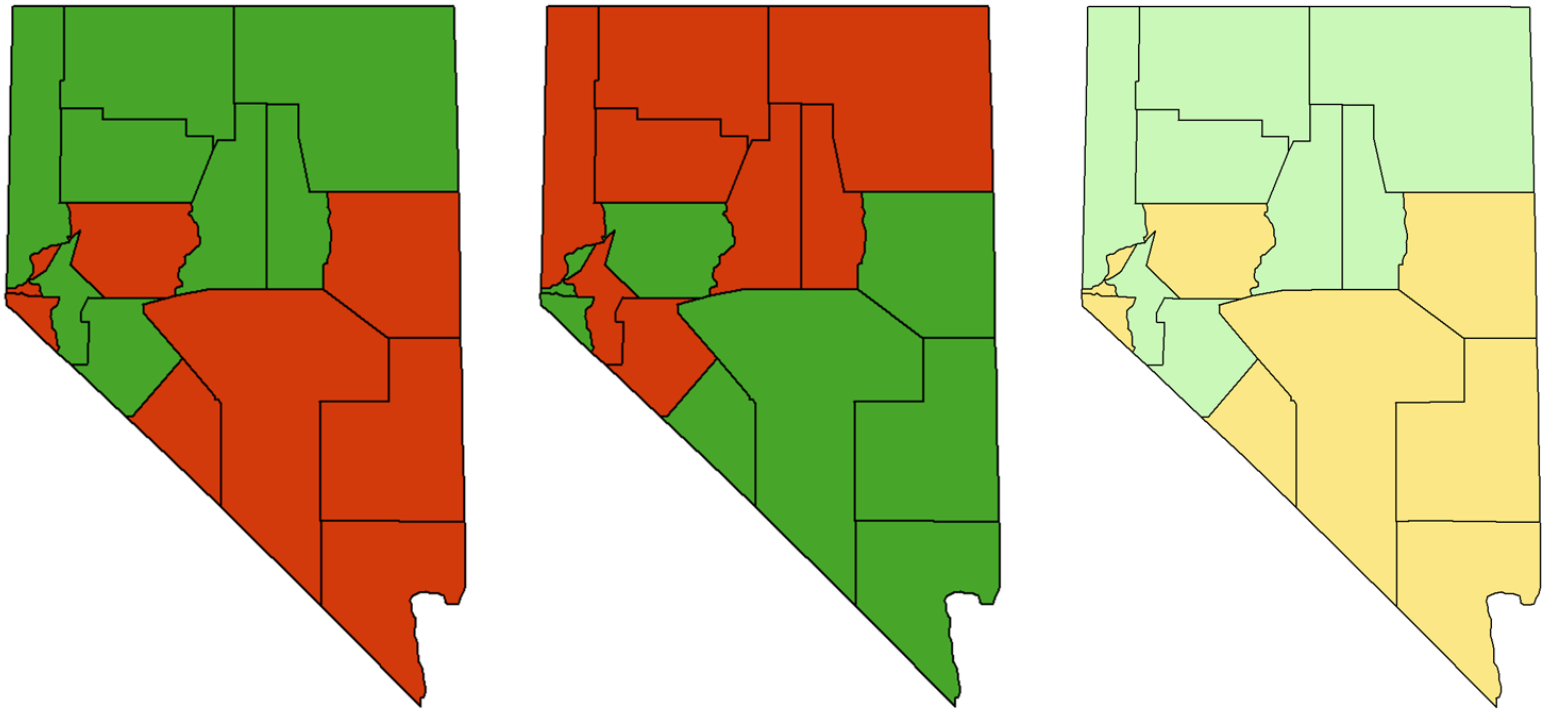
Optical illusions | simultaneous contrast



Which coloured zone has the larger surface area?

Size colour illusion

Optical illusions | size-colour illusion



Size colour illusion

Optical illusions | size-colour illusion

Visualisation of variability

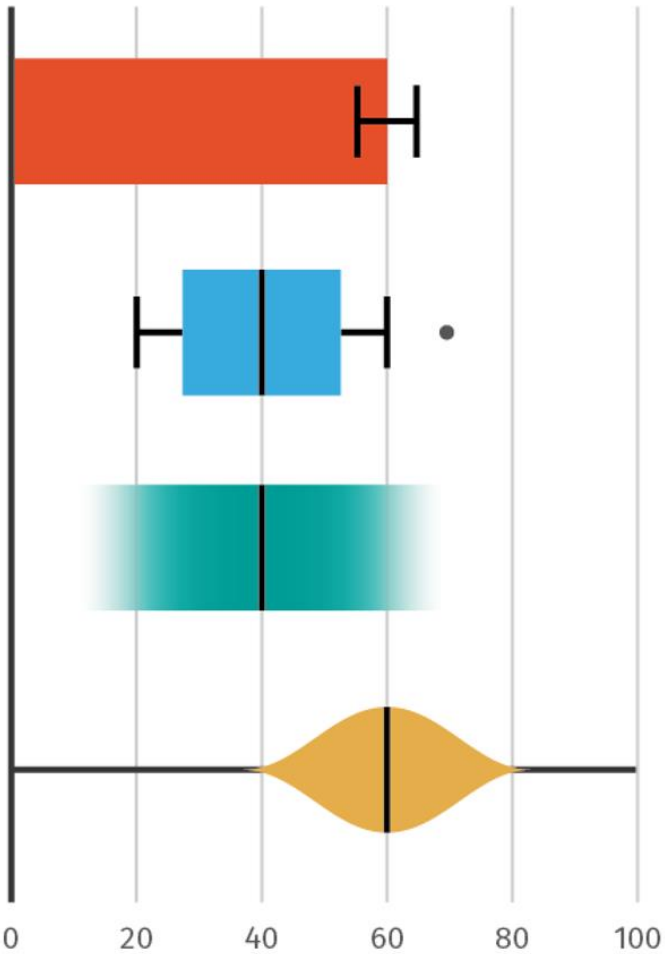
Cognitive bias: people assume uniform distributions regardless of what the underlying distribution is (e.g., normal)

"In psychological science, one example that has received much attention is the interpretation of confidence intervals representing uncertainty.

Several studies have shown that both students and advanced researchers **misunderstand the distribution underlying the confidence interval to be uniform** (e.g., Belia, Fidler, Williams, & Cumming, 2005; Zwick, Zapata-Rivera, & Hegarty, 2014).

Recent approaches have modified confidence intervals with different graphical encodings such as gradients or violin-like shapes to give more information about the distribution, finding some improvements compared to bar charts with error bars."

Correll and Gleicher (2014), emphasis added



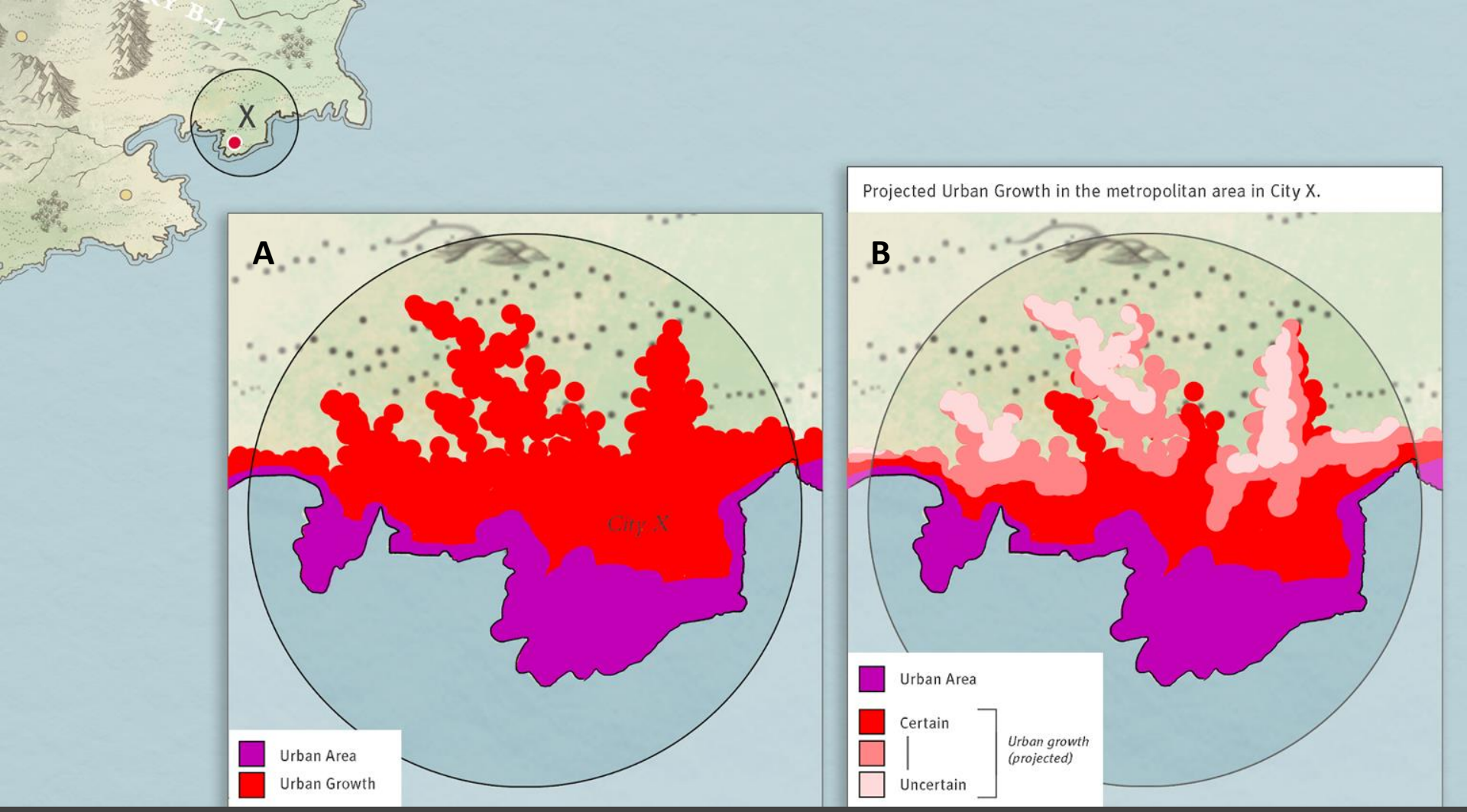
Bar chart with error bars
length of bar = proportional to the values they represent.
whiskers represent error margin (here ±5)

Modified box plot
whiskers represent error margin
center line = median (50% percentile)
dots = outliers (extreme datapoints)

Gradient plot
Point estimate with probability density function shown as a gradient

Violin plot
symmetrical probability density

Correll and Gleicher (2014)



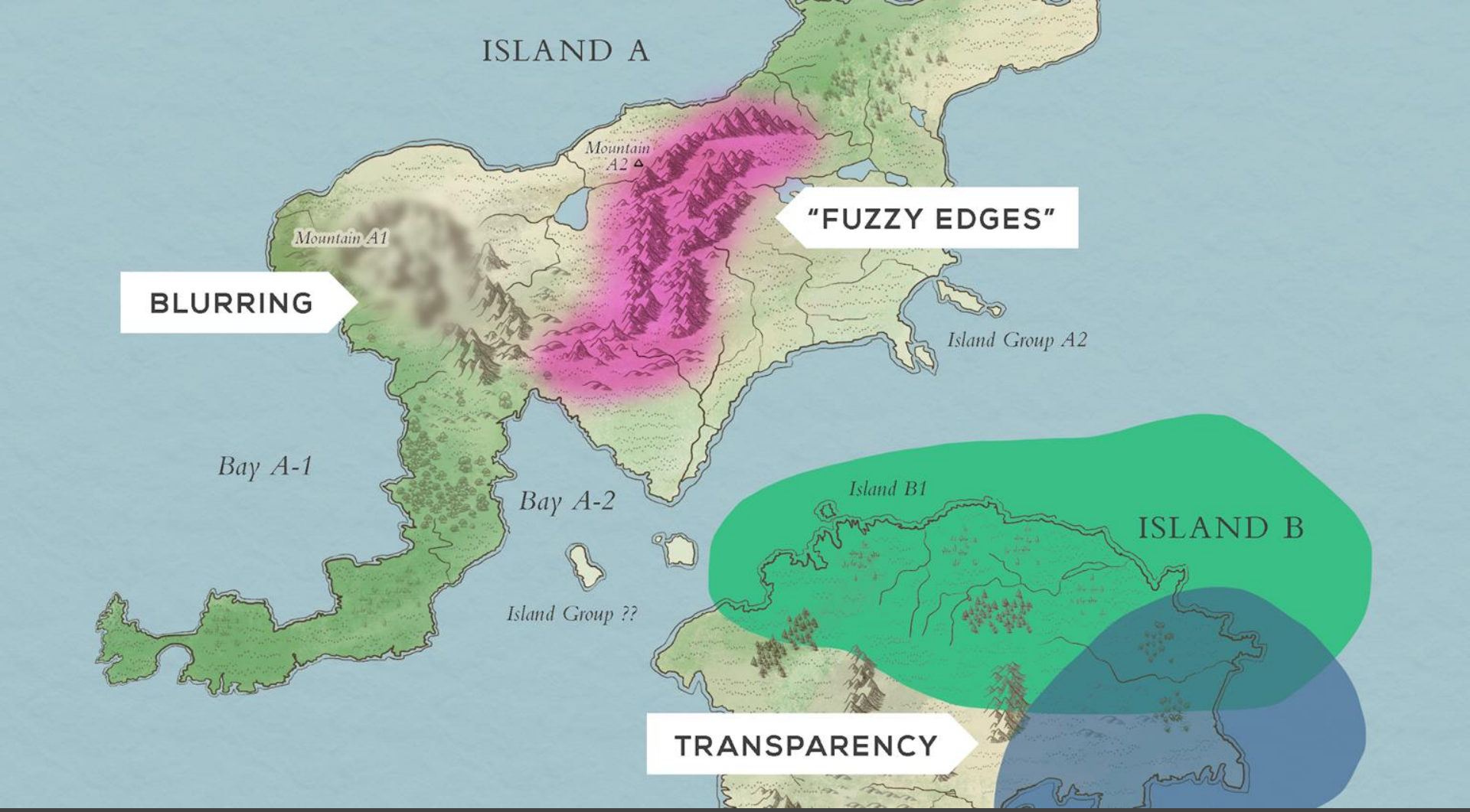
Modification of graphical attributes

Static comparison with a model result (A) and its uncertainty (B).
[single-hue, colour-lightness, side-by-side]

Aerts et al. (2003)

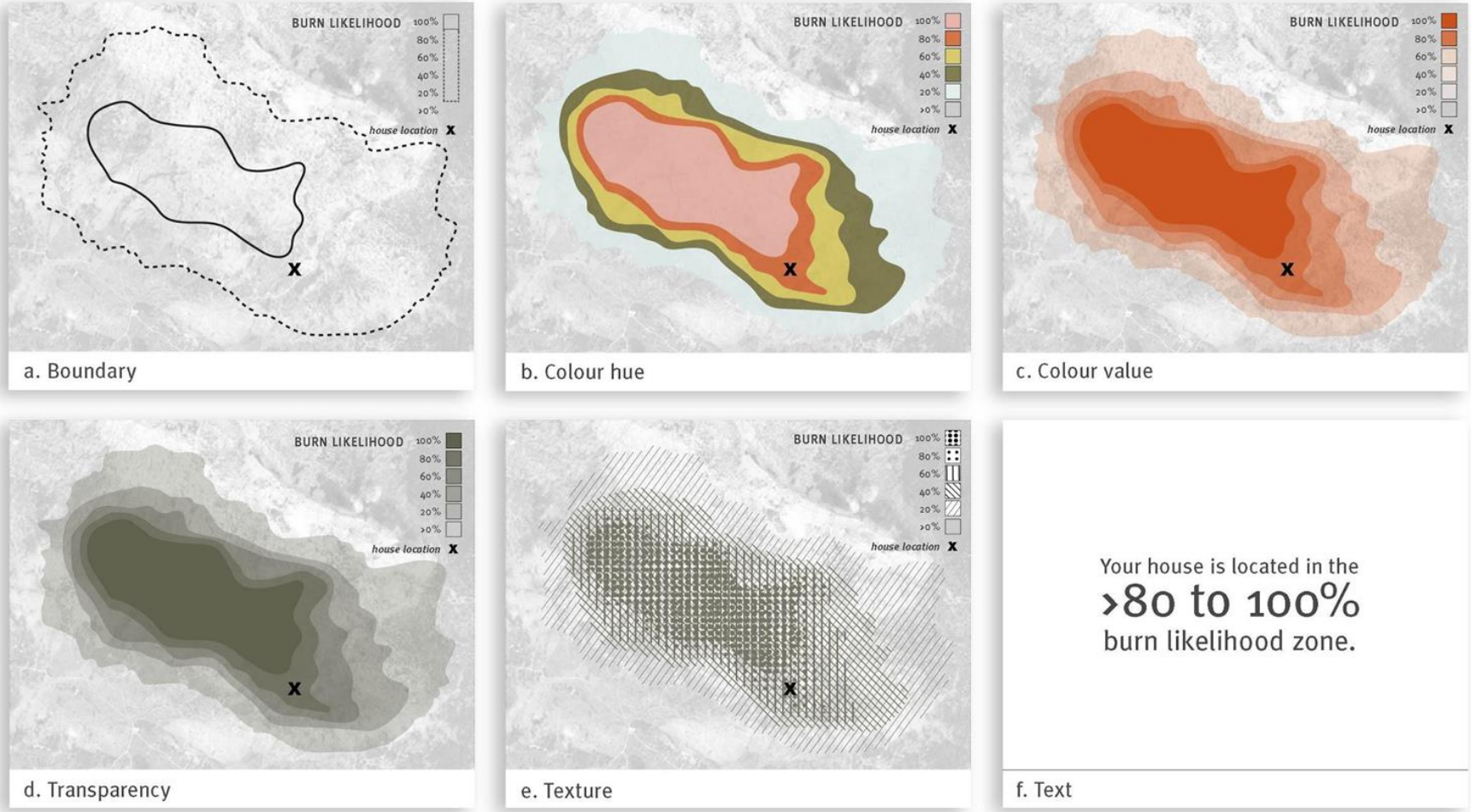
Modification of graphical attributes

Fictitious map; using hue to indicate uncertainty.
[bi-colour, colour-hue]



Modification of graphical attributes

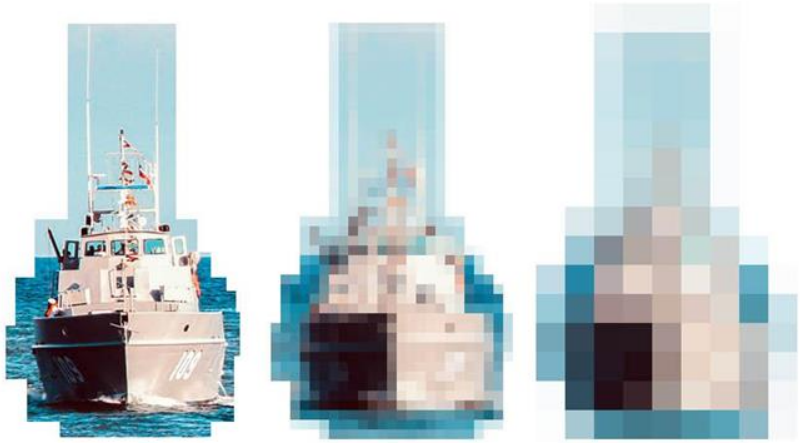
Using "fuzziness", transparency, blurring to denote uncertainty.



Modification of graphical attributes

Experiment: choose to stay or leave a home based upon the interpretation of the potential and uncertain impact of a wildfire. Study of different of methods.

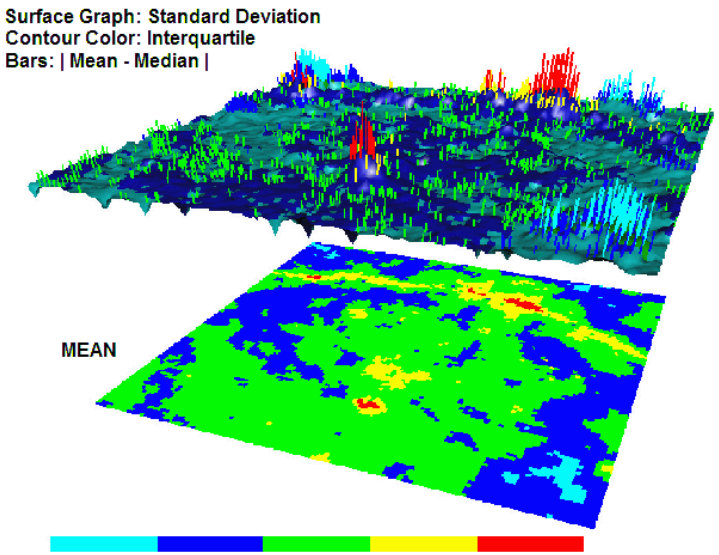
Cheong et al. (2016)



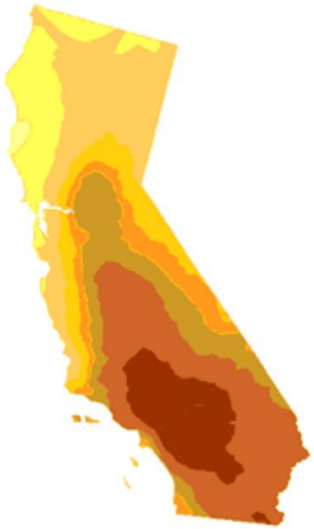
Modification of graphical attributes

Pixelation or adding of noise to show uncertainty.

Interpolation predicts values for cells in a raster from a limited number of sample data points. It can be used to predict unknown values for any geographic point data, such as elevation, rainfall, chemical concentrations, and noise levels.



Point locations of ozone monitoring stations

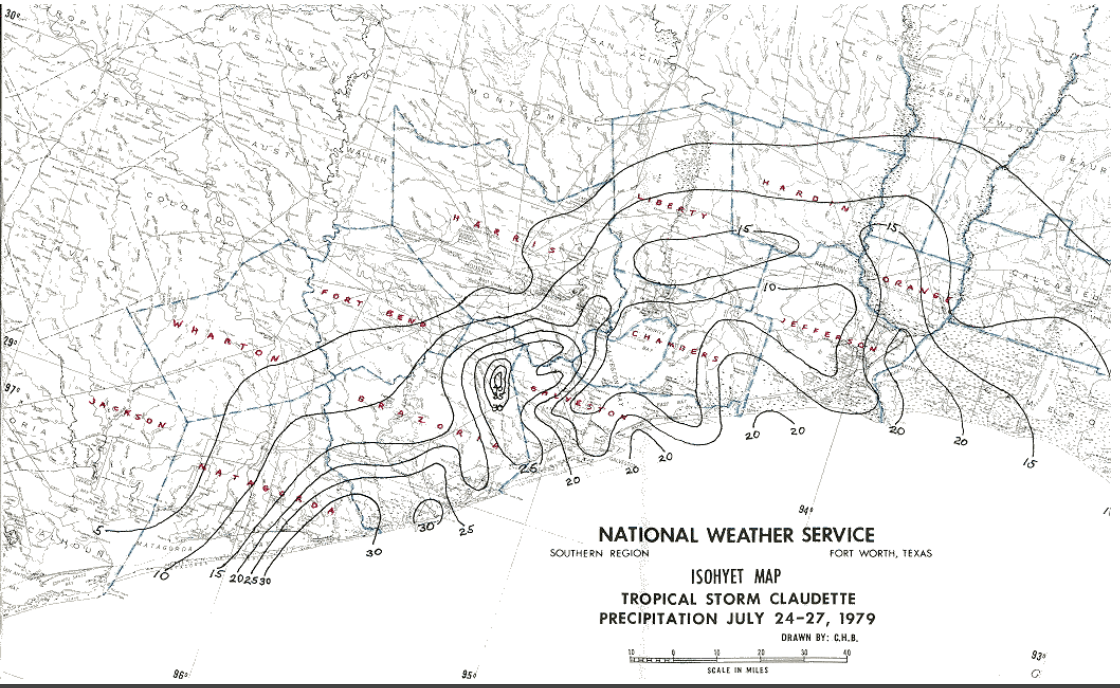


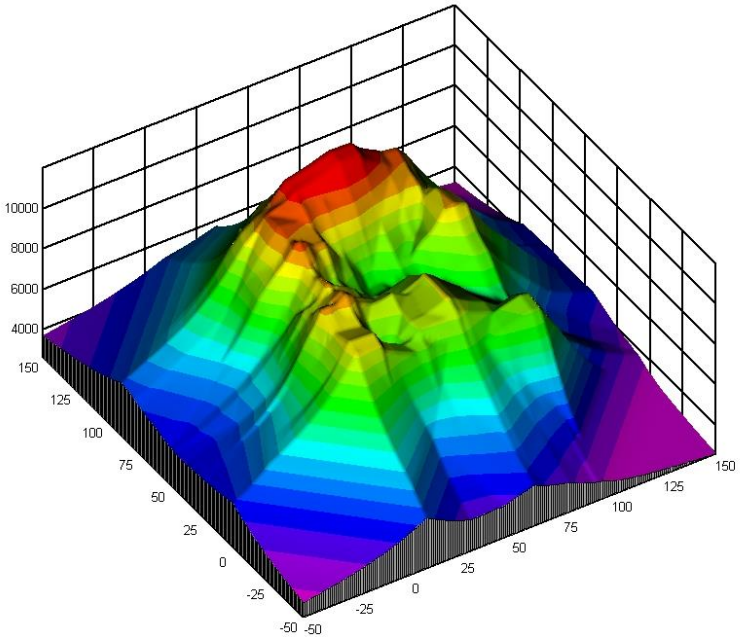
Interpolated prediction surface

Surface interpolation

Predictions of similar characteristics / spatial correlation.
https://slvg.soe.ucsc.edu/areas/unvis.html (many other examples)

An **isosurface** is a three-dimensional analog of an isoline. It is a surface that represents points of a constant value (e.g. pressure, temperature, velocity, density) within a volume of space; in other words, it is a level set of a continuous function whose domain is 3D-space.





Graphical attributes

Isosurfaces

The **shape, size, and orientation of symbols** have been used to represent information about uncertainty in maps, as well as the arrangement (pattern) of groups of symbols.

Howard and MacEachren (1996) found linear patterns overlaid on top of standard maps to be an effective way of communicating levels of uncertainty.

The use of glyphs to represent uncertainty has also been studied (Pang et al. 1997, Cliburn et al. 2002). The findings indicate that glyphs can be successful in the context, although more suited to use by experts as they can be visually overwhelming (Pang et al. 1997).

Addition of artefacts

Symbols & glyphs.

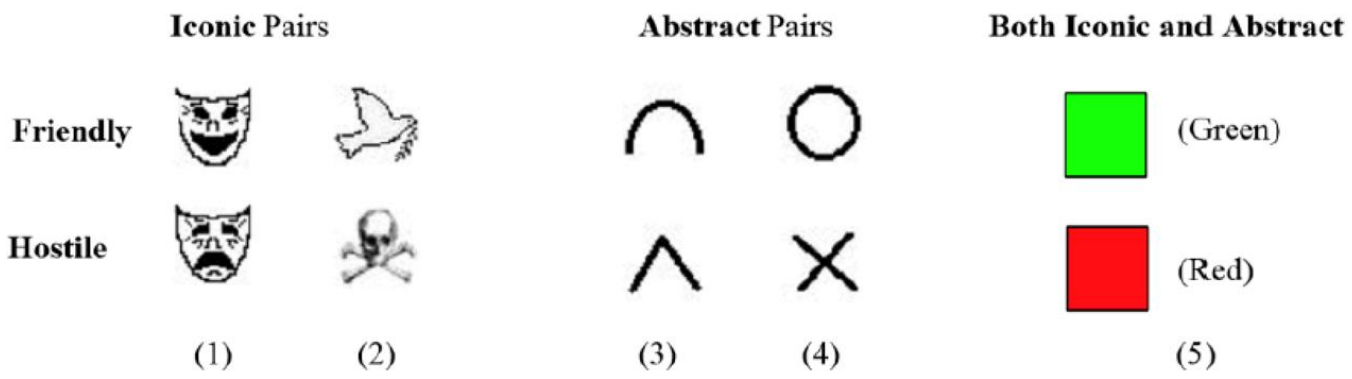
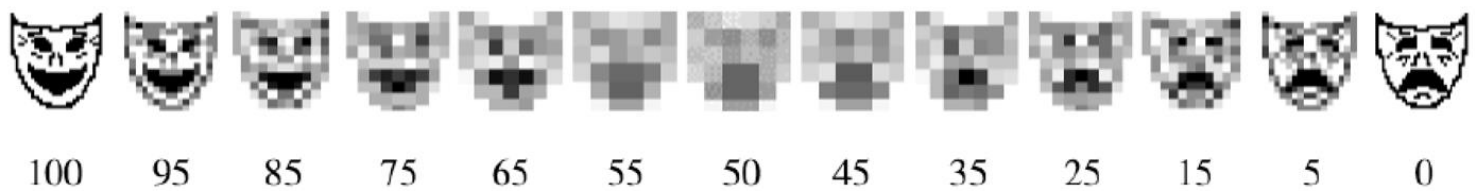


Figure 20: Icons representing object identity (hostile or friendly)(Bisantz et al., 1999)

They then showed that subjects were able to understand different levels of uncertainty associated with the identity of radar contacts as hostile or friendly using a series of 13 degraded and blended icons, as illustrated in Figure 21.



Addition of artefacts: icons

Icons representing a range of probabilities (hostile or friendly).
The numbers indicate the probability of friendly.

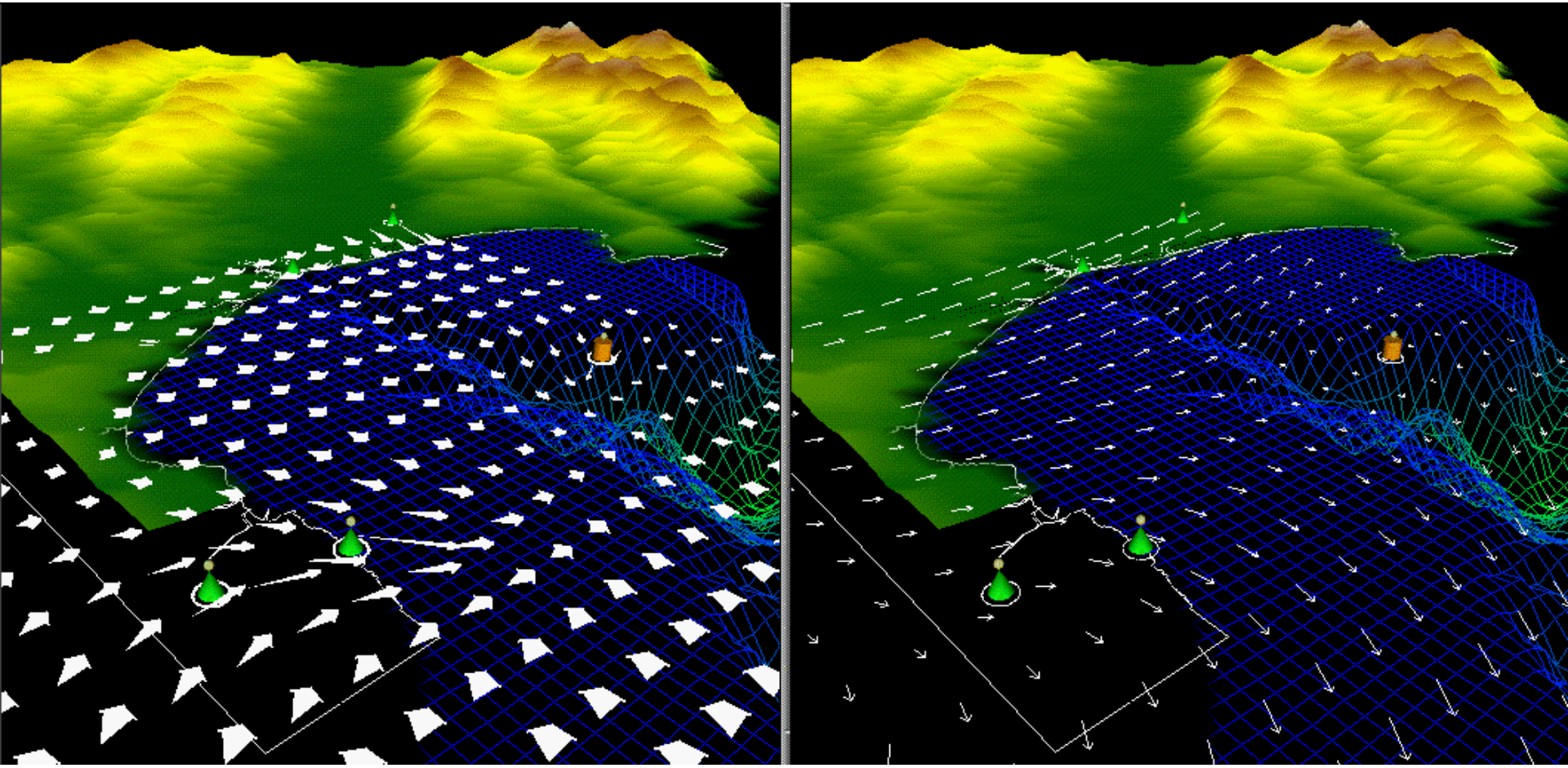
Finger and Bisantz, 2002

Addition of artefacts: icon arrays

Isotype

Emoji

We are using graphics more in everyday communication, but how well are we understanding each other … ?



Addition of artefacts: glyphs

Uncertainty glyphs and vector glyphs to visualize uncertain winds and ocean currents.
Link to source (and many other examples)

Wittenbrink et al. (1996)

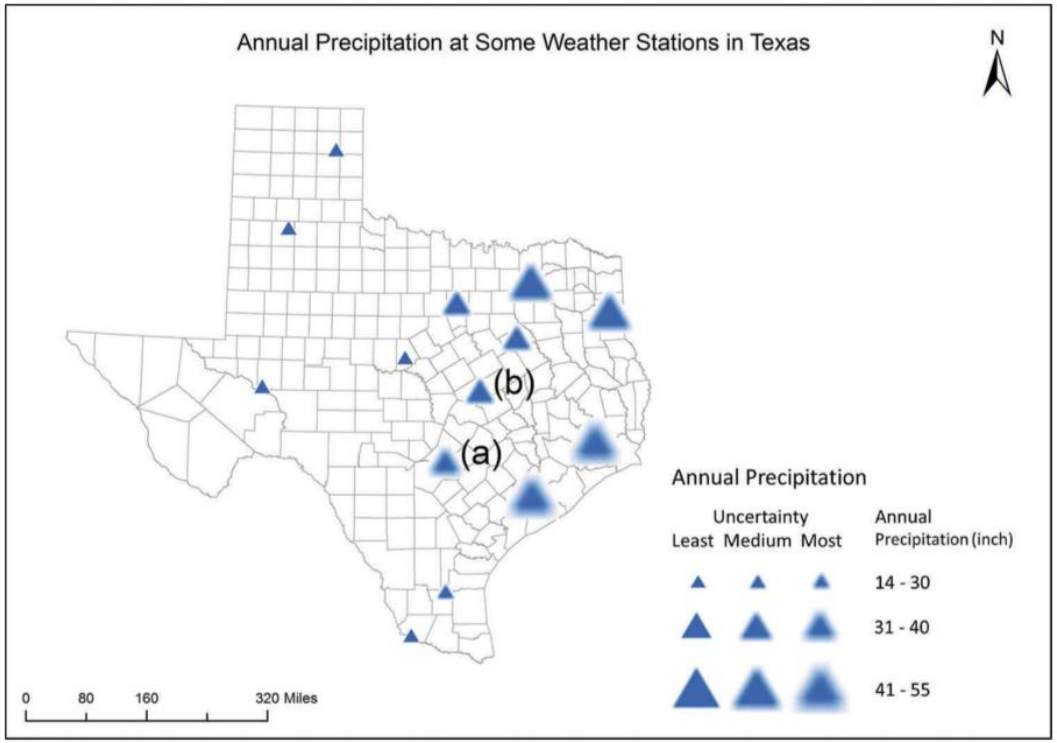
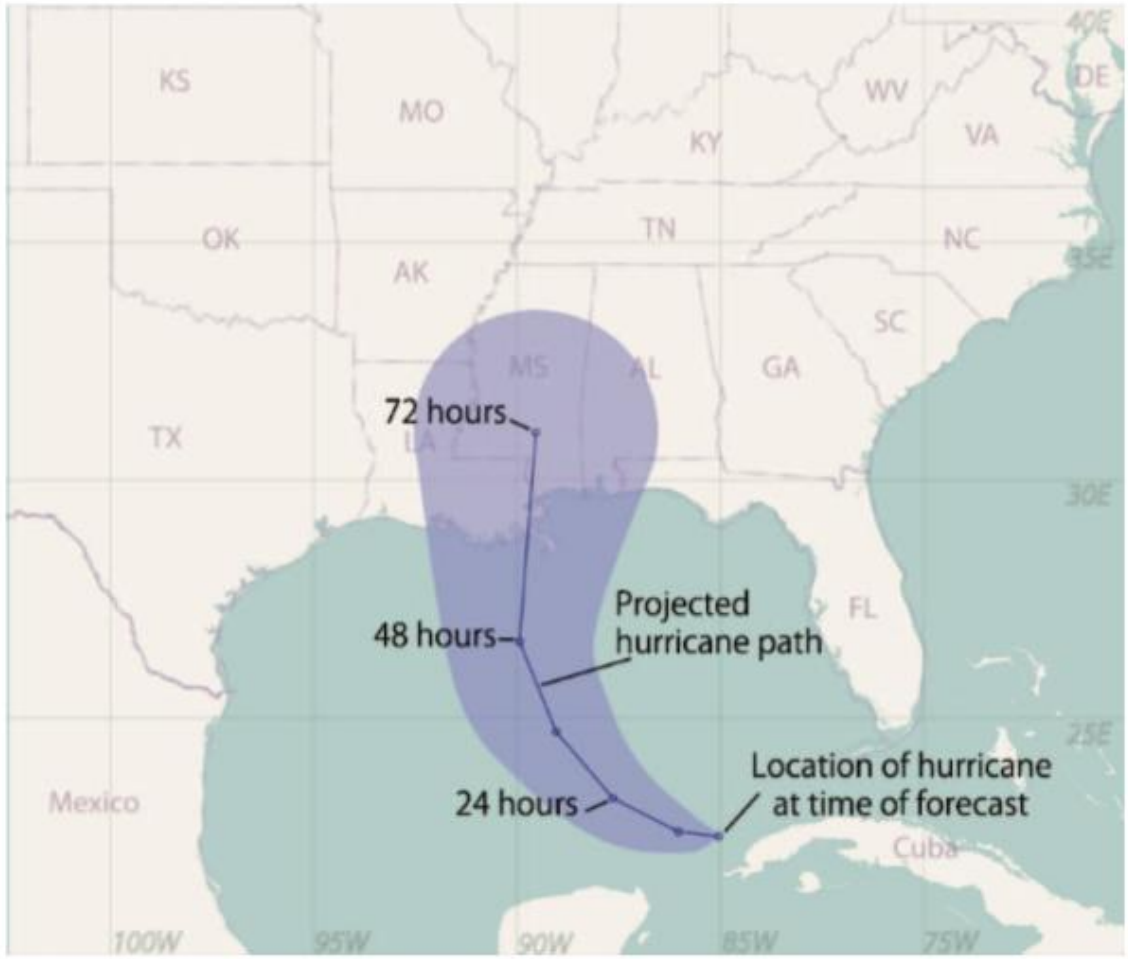


Figure 4. Selection task between two locations (a) and (b) by precipitation and related uncertainty [reprinted with permission from Scholz and Lu (2014)].

Using fuzziness and glyphs/icons

Fuzziness, glyphs

Kinkeldey et al. (2015)



Visualizations of uncertainty in data are often presented to the public without explanations of their graphical conventions and are often misunderstood by nonexperts.

The Cone of Uncertainty.

Graphical Conventions

Boone, A.P. et al. (2018)

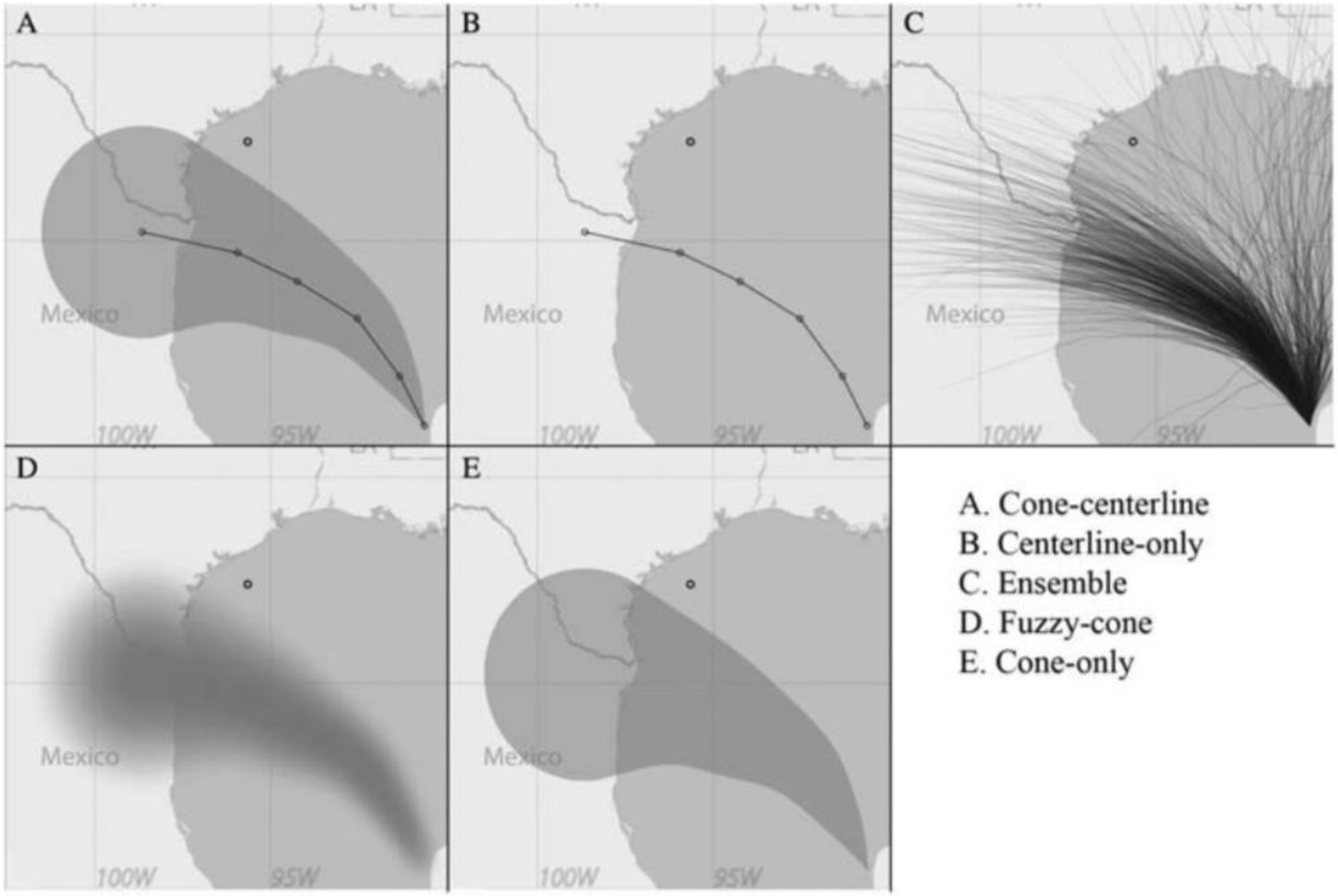
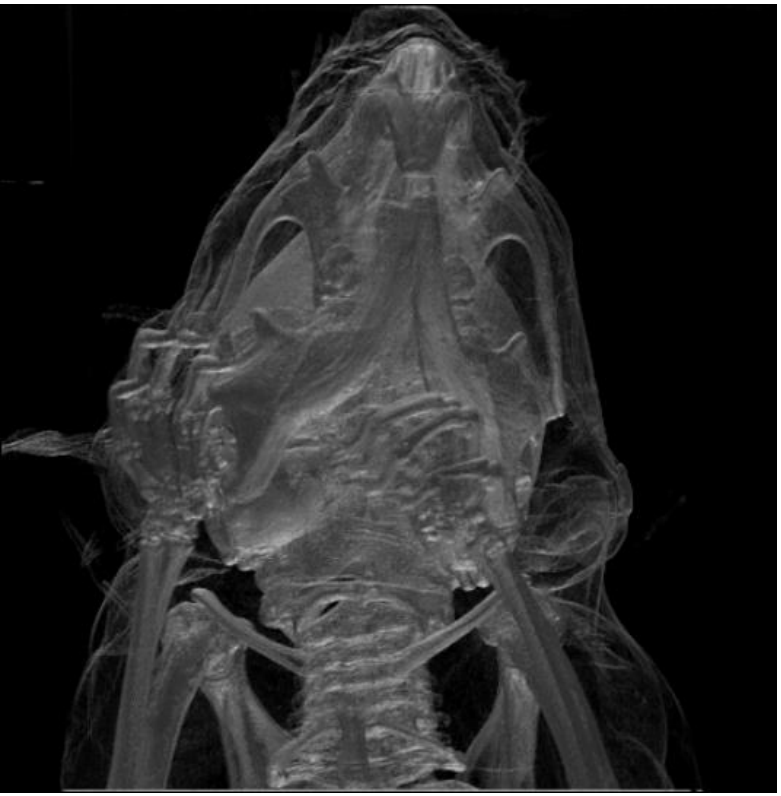


Figure 2. One example of each of the visualizations (presented in color in the actual study with dark blue track lines, light blue cones, and red oil rig locations).

Ruginski, I. T., al. (2016)

Volumetric rendering



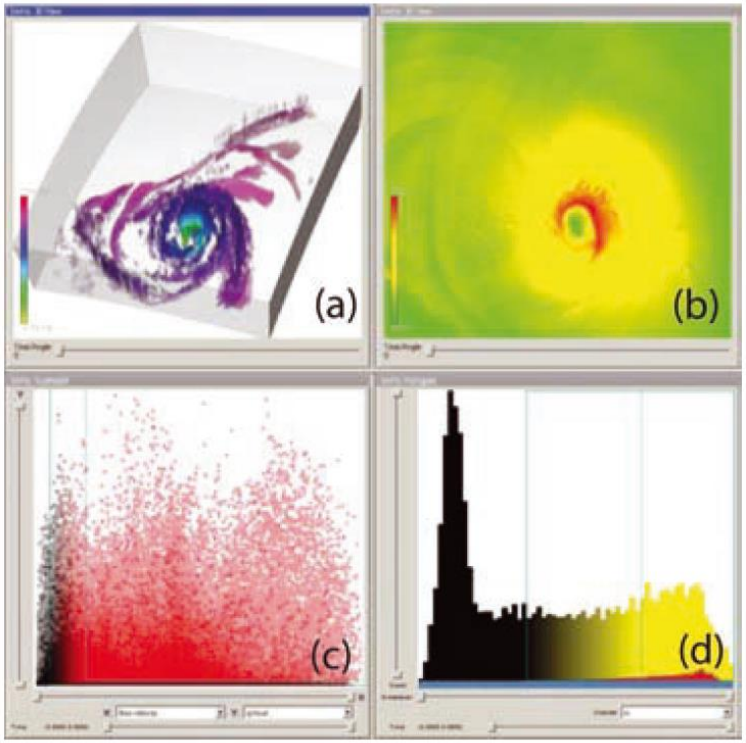
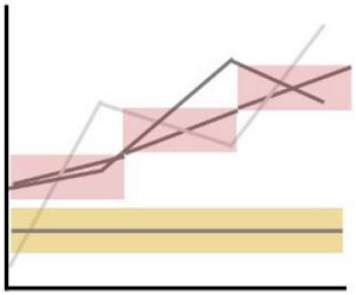


Figure 16: *An example of combined attribute and volumetric views. The 3D view (a) shows the location of data points in space with pressure mapped to colour. A 2D slice (b) shows the velocity close to the eye of the storm. Two attribute views [scatterplot of velocity vs. cloud density (c) and a histogram of temperature (d)] are used to select which cells are shown.*

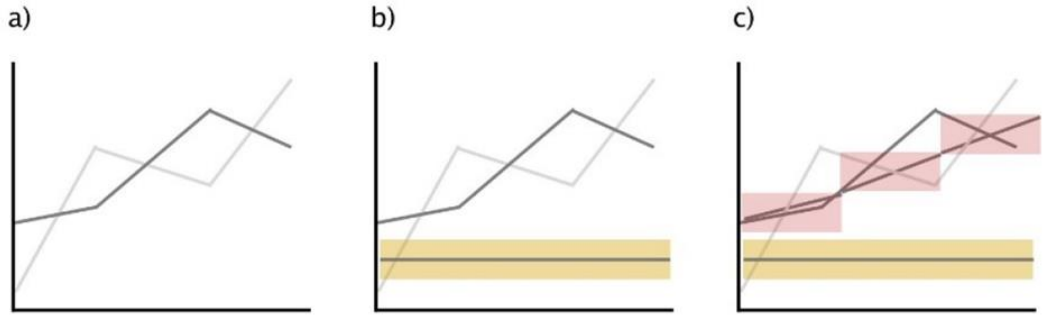
rendering sample, wikipedia

Fuchs and Hauser (2009)

Before:



After (information sequentially built-up):

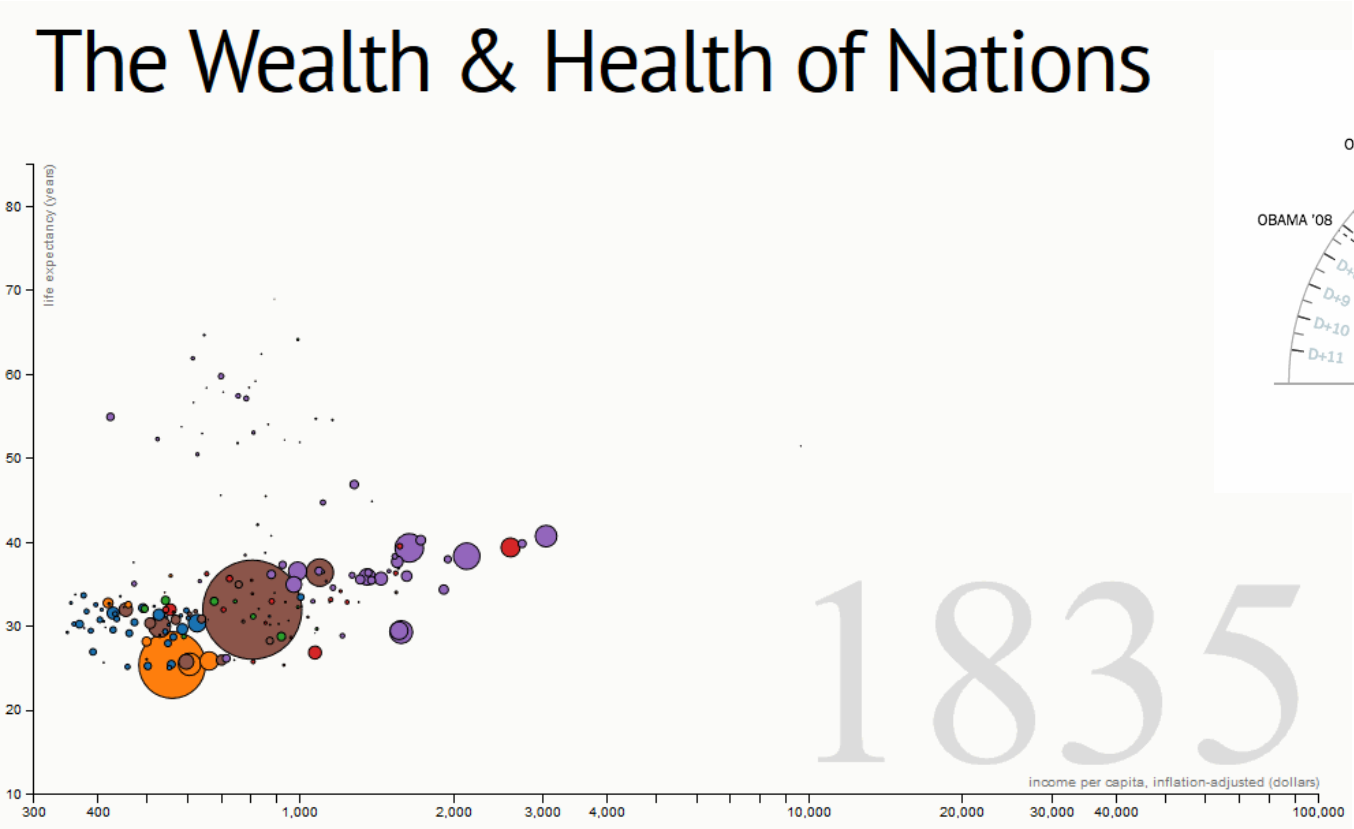


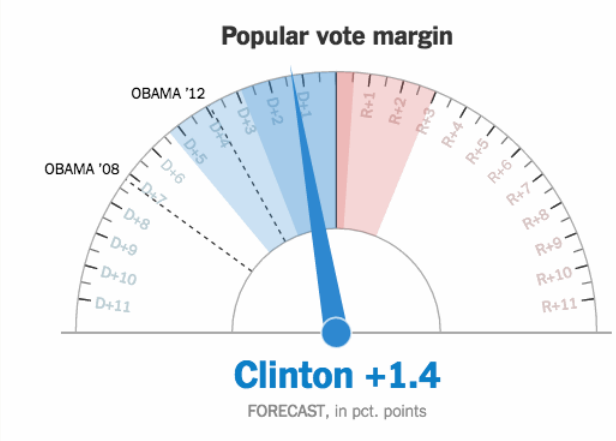
- ➔ Technology enables us to present information dynamically, for instance, layering one image on top of another on top of another.
- ➔ This sequential approach can be used in print or in live presentations.

Animation

Montages

Harold et al. (2007)





Animation

On the left Gapminder's Wealth & Health of Nations, made famous by Hans Rosling's 2006 TED talk. It shows the dynamic fluctuation in per-capita income (*x*), life expectancy (*y*) and population (radius) of 180 nations over the last 209 years. Nations are colored by geographic region; mouseover to read their names. On the right the "Live Presidential Forecast" New York Times 2016 election jitter gauge. The jittering was random, but the jitter range was not.

Sonification is the use of non-speech audio to convey information or perceptualize data. Auditory perceptions open possibilities as an alternative or complement to visualization techniques.

Example: the rate of clicking of a **Geiger counter**

Sonification faces many challenges to widespread use for presenting and analyzing data; studies show it is difficult, to provide adequate context for interpreting sonifications of data.

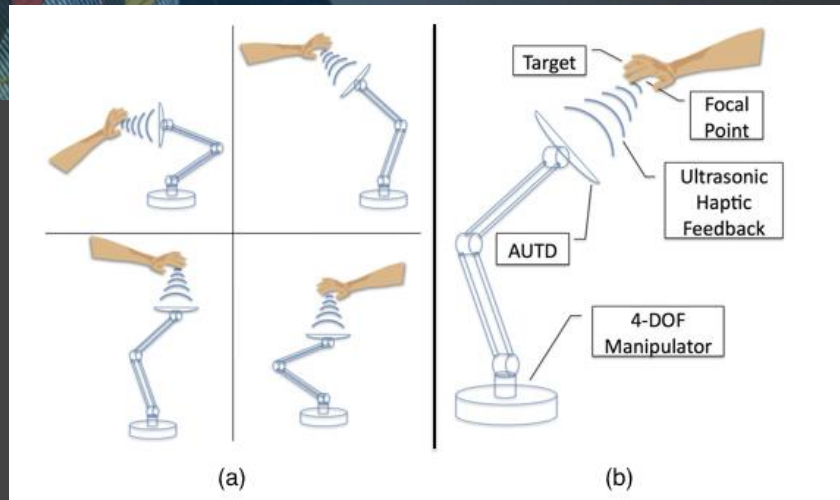
Acoustic feedback

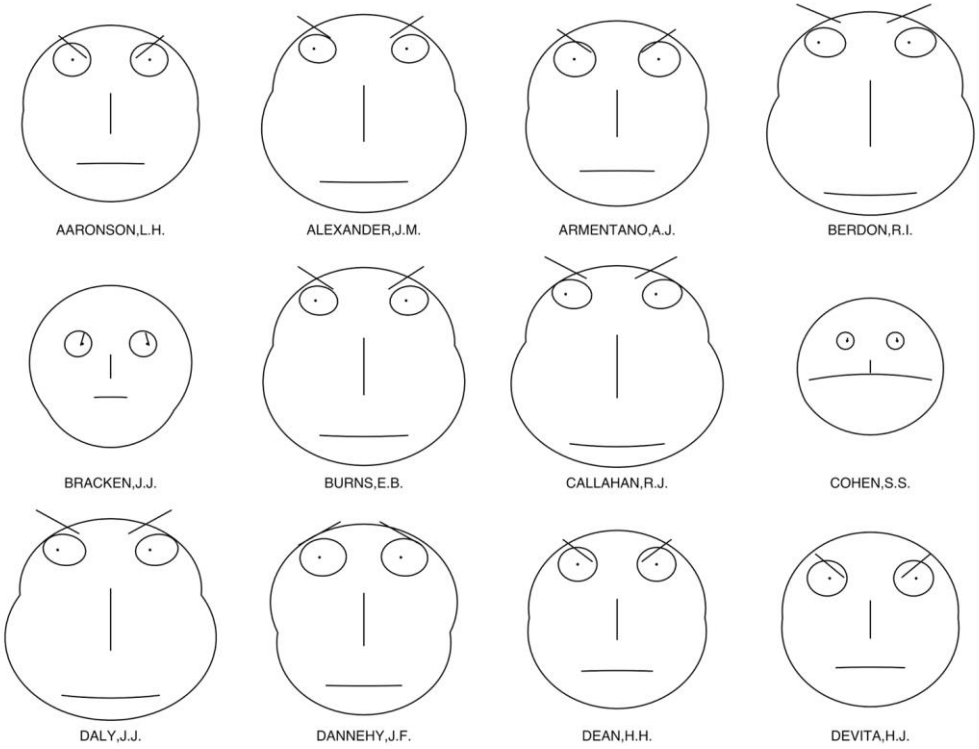
Sonification | alerts | warnings
sample: https://www.nytimes.com/interactive/2017/10/02/us/vegas-guns.html

Kramer, Gregory, ed. (1994), Smith et al. (2005)

Haptic feedback

Vibration, game controllers, virtual reality.





Individual parts represent values by their shape, size, placement and orientation.

The character as a more efficient method of representing data as **a large portion of the human brain is devoted to facial recognition**.

Chernoff faces

Herman Chernoff proposed the use of the shape of a human face (and its individual parts) to display multivariate data in 1973.

[picture source wikipedia]

Figure 24: Examples of stimulus faces with trustworthy scores (lower score means higher trustworthy and vice versa)(Smith et al., 2005)

Avatars

Uncertainty described by descriptors (unlikely, possible, certain, etc.) or through perceived “trustworthiness” of a virtual avatar.

Chung and Wark (2016)

GAN faces

NVIDIA uses a generative adversarial network (GAN) to generate some extremely realistic faces.
No research on using more realistic faces to influence “trustworthiness”.

picture source NVIDIA, copyrighted, do not distribute.

Interactive or playable visualizations have become more popular and end up standing out more

Purposes:

- ➔ Making the data more engaging or playful and showing the data in manageable portions
- ➔ Newer interaction techniques include customization and gamification

But:

- ➔ Interactivity not always necessary to have a successful visualization
- ➔ Can sometimes negatively affect the understanding of the data
- ➔ The distinction between “interactive” and “static” can be blurred (e.g. reorienting a map, folding it, writing on it, may all be forms of interaction)

Dynamic displays

Interactive

A. Figueras (2015)

Filtering	only show me the data in which I am interested
Selecting	mark or track items in which I am interested
Abstract/Elaborate	adjust the level of abstraction of the data
Overview and Explore	overview first, zoom and filter, then details-on-demand
Connect/Relate	show me how this data is related
Reconfigure	give me a different arrangement of the data
Encode	give me a different representation of the data
History	allow me to retrace the steps I take in the exploration of the data
Extraction of features	allow me to extract data in which I am interested
Participation/Collaboration	allow me to contribute to the data
Gamification	show me the data in a more playful way

Table 2: Proposed taxonomy

Dynamic displays

Interactive

A. Figueras (2015)

The use of interactivity and animation has been extensively discussed in information visualization research, but there has been some controversy in relation to its benefits. Additionally, there is still little empirical evidence about its efficacy in terms of improving understanding of the data and there is little research that points out guidelines of how to incorporate it successfully, or that proves that playable visualizations are indeed more enjoyable and popular among users.

Dynamic displays

Interactive

A. Figueras (2015)

Workshop activity:
Exploring new signifiers of uncertainty

“Uncertainty and graphicacy: How should statisticians, journalists, and designers reveal uncertainty in graphics for public consumption?”

‘The first time we see a novel graph, chart, or map, it’s unlikely that we’ll know how to read it at a glance. Before we can decode a graphic, we need to understand its logic, grammar, and conventions.’

A. Cairo

- This is an **open-ended experiment**, designed to stretch us and spark conversation … we can define success however we like
- We have some **dummy data** that is fairly complicated, confusing, and contradictory
- In groups, let’s try to interpret this data, then think about **how this uncertainty could be visualised**
- You can actually create visualisations, and/or describe them
- If you want, you can use this as an opportunity to think more about dynamic displays and interactivity

Activity Scenario

Getting creative with uncertainty visualisation

- ➔ Technology has lowered the barriers to creating bioweapons (such as engineered pandemics)
- ➔ Small groups and even individuals with relatively ordinary knowledge and financial resources may pose global catastrophic risks
- ➔ Imagine you are monitoring nine AI models, drawing on big data sources
- ➔ Each model is designed to seek patterns associated with one particular thing: three models look for **bioterror-related patterns**, three models look for **other criminal or regulatory issues**, three look for normal, everyday **research and development**

Activity Scenario

Getting creative with uncertainty visualisation

- Each model is designed to seek patterns associated with one particular thing: three models look for **bioterror-related patterns**, three models look for **other criminal or regulatory issues**, three are kind of a control, looking for normal, everyday **research and development**

Source Name	Task of Source is to Match ...	CLASSIFICATION
Model 1	R&D	R&D
Model 2	R&D	Not R&D
Model 3	R&D	Unknown
Model 4	Terror	Crime
Model 5	Terror	Crime
Model 6	Terror	Crime
Model 7	Crime	Terror
Model 8	Crime	Terror
Model 9	Crime	Not Terror

- Each classification comes with loads of caveats (as you'll see in a moment)
- Your models also interact to output an overall estimated classification
- Let's say this classification has just flipped to the possible detection of bioterror-related activity

Activity Scenario

Getting creative with uncertainty visualisation

Source Name	Task of Source is to Match ...	Dependencies (Do some models usually covary?)	Final Weighting (Based on Priority, Past Performance and Situational Relevance)	Model Priority	Model Past Performance	Model Situational Relevance	Final Data Quality Score (Based on Data Timeliness, Precision and Accuracy, and Completeness)	Data Timeliness	Data Precision and Accuracy	Data Completeness	CLASSIFICATION
Model 1	R&D	Model 4	56	100	1.00	0.56	0.88	00 hr 52 min	56%	0.19	R&D
Model 2	R&D	None	50	50	0.89	0.99	0.94	00 hr 26 min	N/A	0.39	Not R&D
Model 3	R&D	None	32	35	0.87	0.90	0.98	00 hr 07 min	N/A	0.44	Not R&D
Model 4	Terror	Model 1	35	100	0.81	0.35	1.00	00 hr 14 min	85%	0.73	Crime
Model 5	Terror	Model 9	1	50	0.97	0.02	0.96	00 hr 16 min	N/A	0.08	Not Crime
Model 6	Terror	None	15	25	0.98	0.59	0.01	06 hr 00 min	N/A	0.64	Unknown
Model 7	Crime	None	83	100	0.96	0.83	0.99	00 hr 24 min	N/A	0.38	Terror
Model 8	Crime	None	20	50	0.94	0.41	0.79	01 hr 00 min	N/A	0.17	Not Terror
Model 9	Crime	Model 5	28	35	0.92	0.79	1.00	00 hr 21 min	N/A	0.70	Terror

Classified: Terror

Model Name	Model Analytic Focus	Composite Model Weighting	Priority (Influences Model Weighting)	Past Performance (Influences Model Weighting)	Situational Relevance (Influences Model Weighting)	Label Estimate	Composite Data Quality Score	Data Latency (Influences DQ Score)	Data Precision and Accuracy (Influences DQ Score)	Data Completeness (Influences DQ Score)	Potential Covariance
Model 4	Crime	60	100	0.97	0.6	Crime?	0.8	01 hr 12 min	97%	0.13	Model 1
Model 5	Crime	8	50	0.99	0.15	Crime?	1	00 hr 02 min	65%	0.24	Model 9
Model 6	Crime	5	25	0.98	0.22	Crime?	1	00 hr 02 min	N/A	0.67	None
Model 7	Terror	72	100	0.91	0.72	Terror	0.7	01 hr 12 min	N/A	0.12	None
Model 8	Terror	30	50	0.95	0.61	Terror	0.88	00 hr 43 min	66%	0.15	None
Model 9	Terror	15	35	0.84	0.44	Not Terror?	1	00 hr 19 min	67%	0.88	Model 5
Model 1	R&D	46	100	0.88	0.47	R&D?	0.85	00 hr 50 min	60%	0.52	Model 4
Model 2	R&D	14	50	0.84	0.27	Not R&D?	0.91	00 hr 02 min	N/A	0.19	None
Model 3	R&D	27	35	0.86	0.78	Unknown?	0.94	00 hr 38 min	N/A	0.99	None

Overall Classification: **Potential Terror Threat!**

Activity Scenario

Feedback

- How might the analyst want to visualize this data for themselves?
- What should the analyst and/or decision support tool show to the decision-maker?
- What might we tell the decision-maker?
- What if we're not there to explain our visualization? What if the visualization is likely to escape "into the wild" -- moving among multiple stakeholders, or given to the media?
- How did you imagine the analyst's role here? How much influence does the analyst have? What counts as undue influence, and how do they avoid it?
- What other promising ideas did we come up with?
- Any constructive feedback on each other's ideas?
- How did it feel? What felt most difficult? How could this difficulty have been mitigated?
- Any questions for each other?
- Anything emerging here we'd like to explore in the next session?
- Any other comments?

Activity Scenario

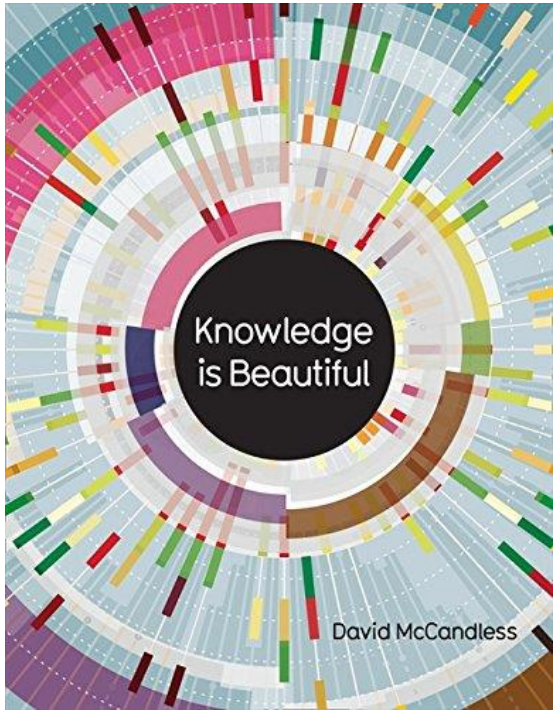
Debrief

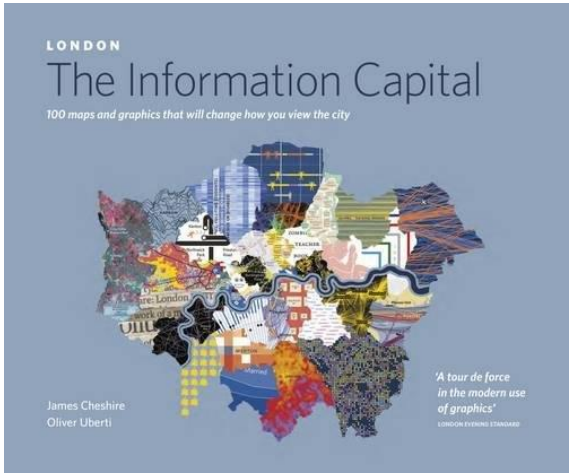
- ➔ The text data itself is already a visualization -- how might this have influenced your activity in the groups?
- ➔ **Boundaries** “help partition an information space into zones of relative semantic homogeneity” (Fabrikant and Skupin 2005).
- ➔ Is there a risk of treating data within a containment as more similar than data across boundaries?
- ➔ Observers who read left-to-right and top-to-bottom **may pay less attention to the bottom or the far-right cells** of a row or column (Sütterlin, Brunner, & Opwis, 2008)
- ➔ In two-dimensional arrays, **observers may pay less attention to whatever is tucked in the corners** (Meißner, Musalem, & Huber 2016)
- ➔ Red-Green-Amber was used in a fairly idiosyncratic way in one of the datasets
- ➔ The questions marks are probably a terrible idea too

Activity Scenario

Debrief

Recommended books





Resources: general background

How do you create interactive data visualisations?
Nesta Sparks lecture by Cath Sleeman
https://www.youtube.com/watch?v=ctSl8tYEEDY

Visualising the uncertainty in data by Nathan Yau (UCLA)
https://flowingdata.com/2018/01/08/visualizing-the-uncertainty-in-data/

Visualising conflict data:
https://www.acleddata.com/

Jean Golding Institute 'Beauty of Data' competition
https://www.flickr.com/photos/bristoluniversity/sets/72157699232459344/with/41800387840/

Resources: design

- Free images:
  https://pixabay.com/
  https://unsplash.com/
  https://thenounproject.com/

- Catalogue of examples
  http://www.rethinkingvis.com/#all
  https://datavizcatalogue.com/

- Tutorials
  https://flowingdata.com/category/tutorials/

Resources: platforms for data visualisations

Microsoft

- https://powerbi.microsoft.com/en-us/

R shiny

- https://shiny.rstudio.com/gallery/

Tableau

- https://www.tableau.com/

D3

- https://github.com/d3/d3/wiki/Gallery

Resources: social media on visualising data

- Financial Times @ftdata
- NYT Graphics @nytgraphics
- https://twitter.com/GuardianVisuals
- The Pudding @puddingviz or https://pudding.cool/
- Andy Kirk @visualisingdata or http://www.visualisingdata.com/

Future collaborations

Jana Kleineberg

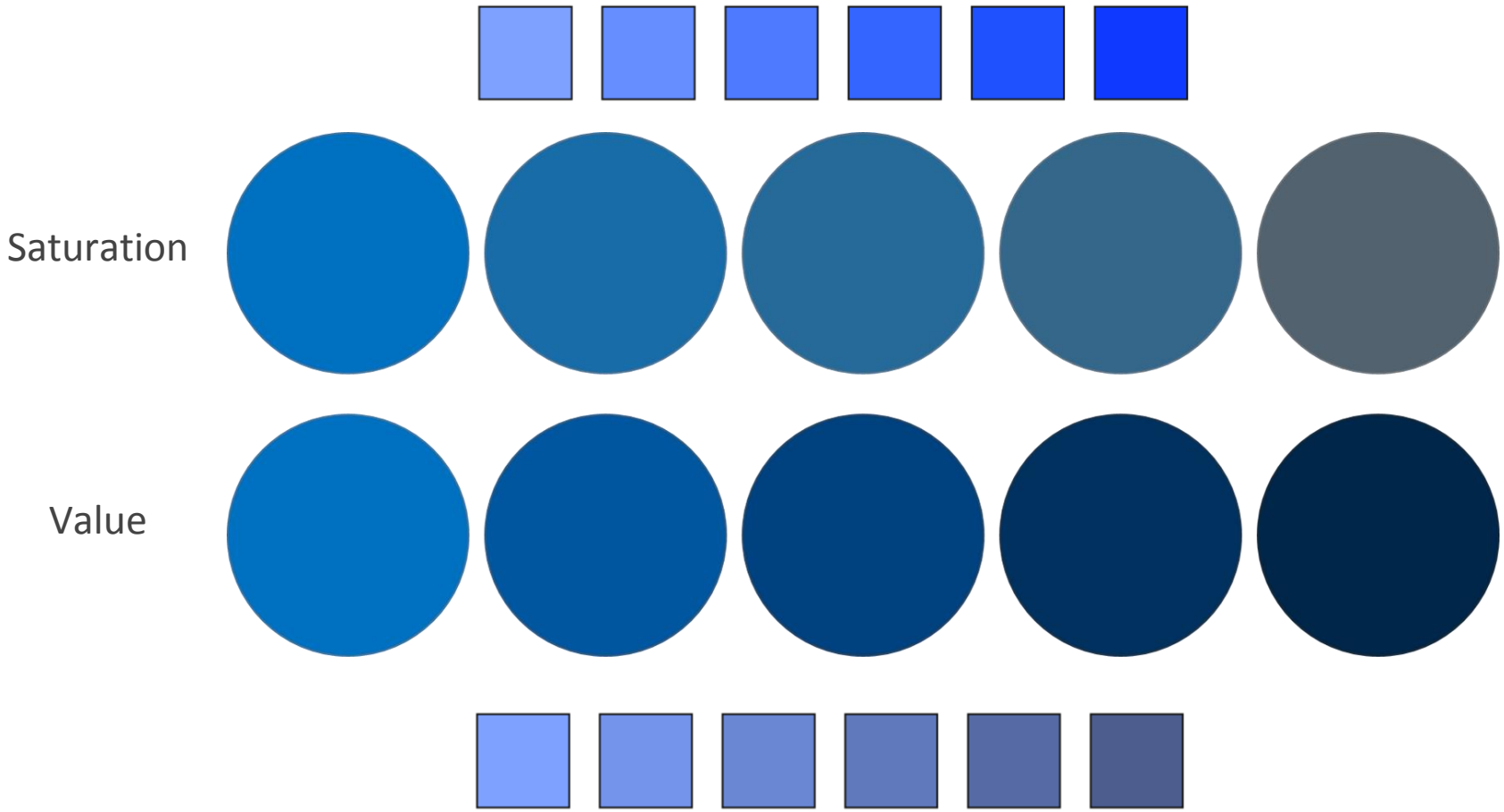
http://www.kleineberg.co.uk/

Or contact us:

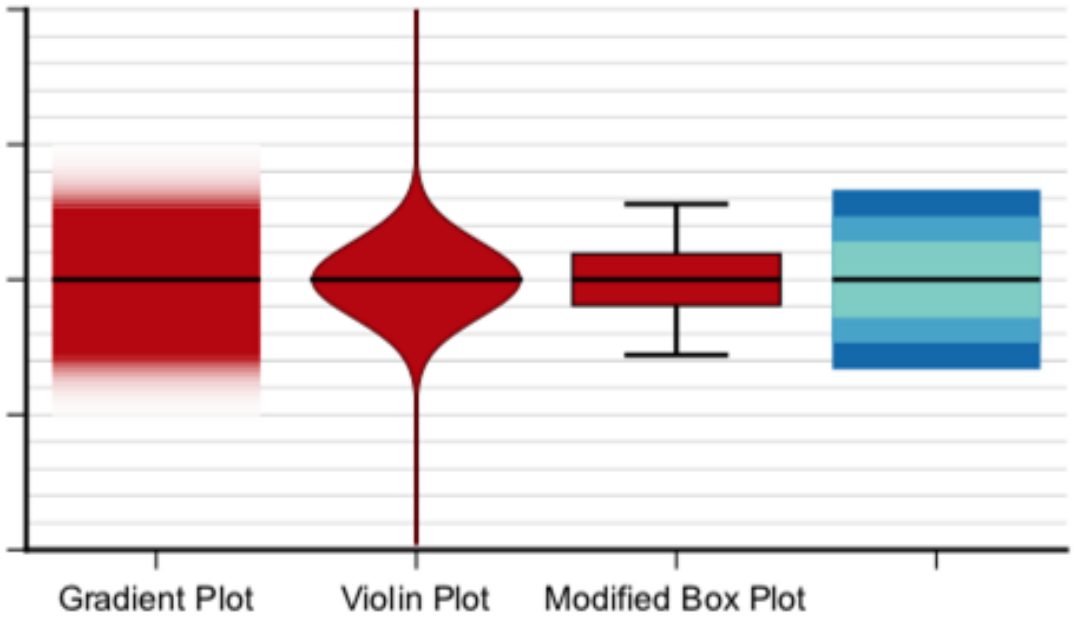
http://www.seaplusplus.co.uk/

http://au4dmnetworks.co.uk/

Unused Slides



Powerpoint’s idea of saturation?



The colored bars represent standard error, 95% and 99% confidence interval respectively