Imperial College London

Visual Communication

John Pinney

March 2023

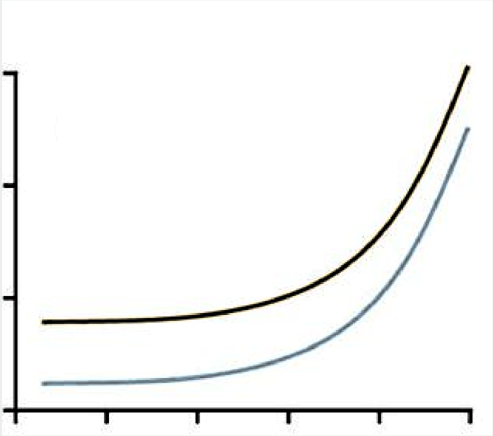
Visual estimations

Human perception treats different types of visual stimuli in different ways.

Some types of representation are easier to decode and compare than others.

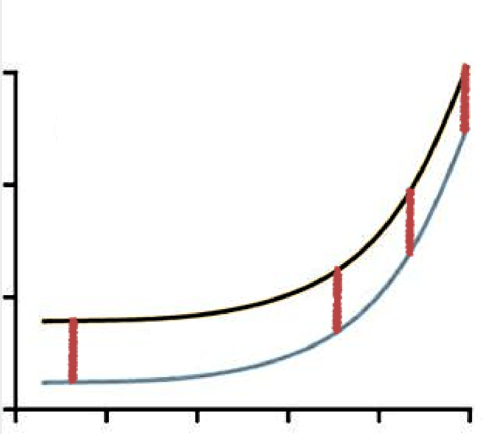
Our tendency to visually assemble elements that are close together or similar to each other can also cause biases in perception that can distort the message communicated. [1]

Visual estimations



This plot shows two variables that increase over time.
What proportion of the initial difference (y_1 - y_2) is the final difference?

Visual estimations

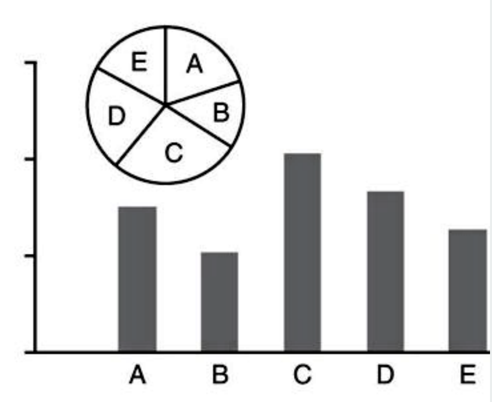


In fact, there is no change in the difference between the curves.

Visual estimations

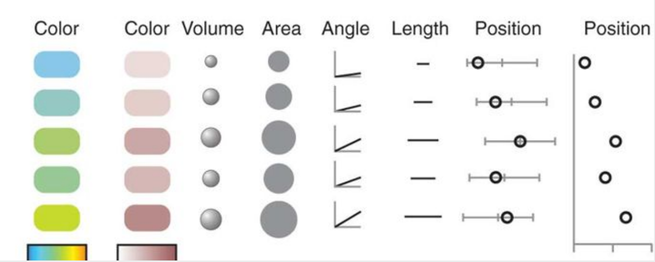
Order the letters by decreasing value.

Visual estimations



Comparing pie chart areas is notoriously difficult.
A bar chart is much easier to decode.

Visual estimations



The data in each column are the same, but the perceived values may be quite different.

Principles of data graphics

Edward Tufte presents the following principles for visual communication. [2]

- Tell the truth

  Show the data and avoid distorting what they have to say.

- Show data variation

  Lead the viewer to focus on the substance of the findings rather than the methodology or graphical design.

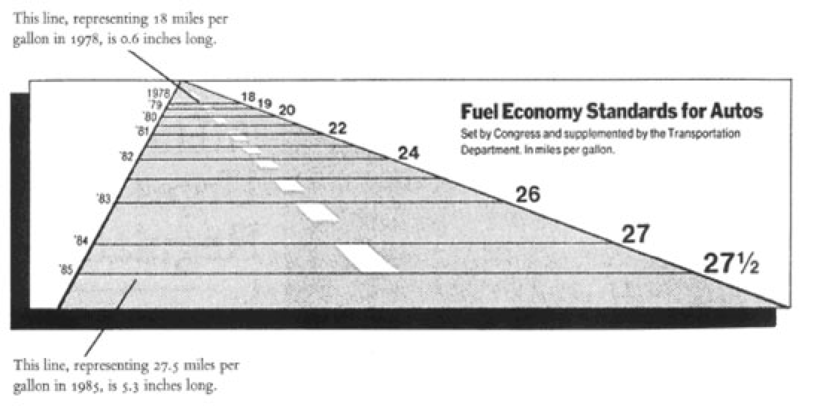
- Make large data sets coherent

  Present data efficiently and encourage the eye to compare different pieces of data.

- Reveal the data at several levels of detail

  From a broad overview to the fine structure.

Visual Integrity



Visual Integrity

It is easy to design graphics that deliberately mislead the viewer.
We can quantify the amount of distortion using Tufte's lie factor:

$$\text{lie factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

where

$$\text{size of effect} = \frac{|\text{second value} - \text{first value}|}{\text{first value}}$$

The lie factor of the fuel economy graph is 14.8.

(numerical change = 53% but graphical change = 783%)

Data-Ink

Good graphical representations should maximize data-ink and erase as much non-data-ink as possible.

The data-ink ratio is defined as

$$1 - \text{proportion of ink that can be erased without loss of information}$$

Data-Ink

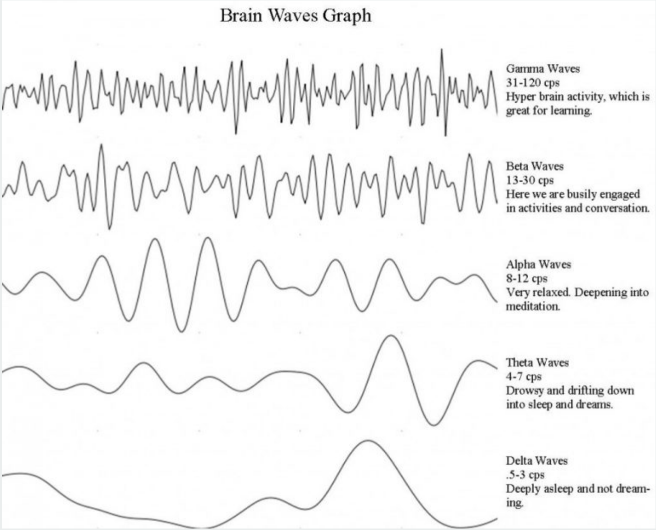
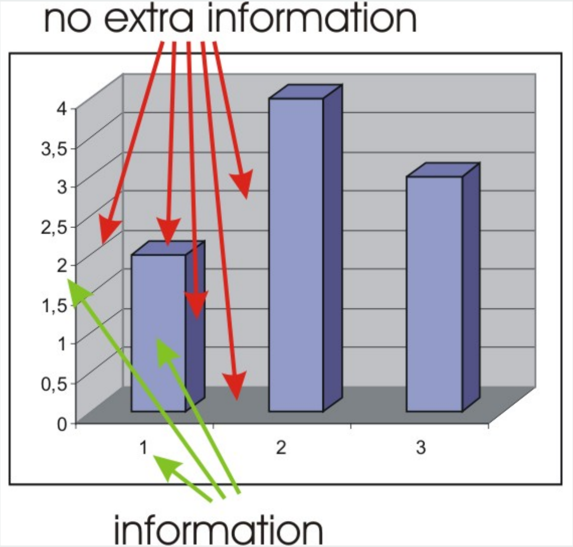

Brain Waves Graph
Gamma Waves
31-120 cps
Hyper brain activity, which is great for learning.
Beta Waves
13-30 cps
Here we are busily engaged in activities and conversation.
Alpha Waves
8-12 cps
Very relaxed. Deepening into meditation.
Theta Waves
4-7 cps
Drowsy and drifting down into sleep and dreams.
Delta Waves
.5-3 cps
Deeply asleep and not dreaming.

Chart Junk

"The interior decoration of graphics generates a lot of ink which does not tell the viewer anything new. The purpose of the decoration varies - to make the graphic appear more scientific, to enliven the display, to give the designer an opportunity to exercise artistic skill. Regardless of the cause, it is all non-data-ink or redundant data-ink, and it is often chartjunk."[2]

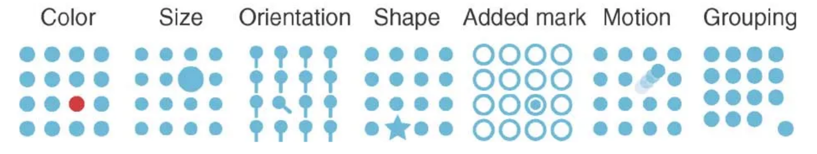
Chart Junk



Salience

The visual salience of a graphical element is the extent to which it stands out from its surroundings.[3]

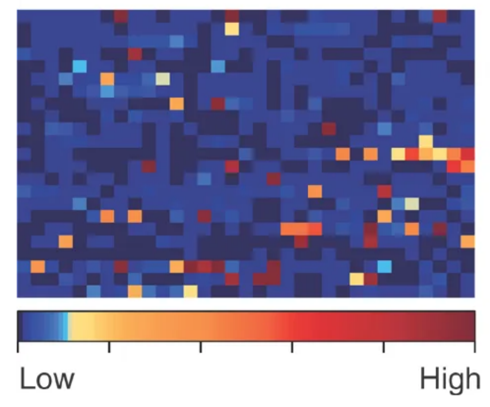
Salience



Examples of visual features that make objects distinct.

Salience vs relevance

It is important that the features that are most noticeable (high salience) are those that are most important to the message communicated (high relevance).[4]

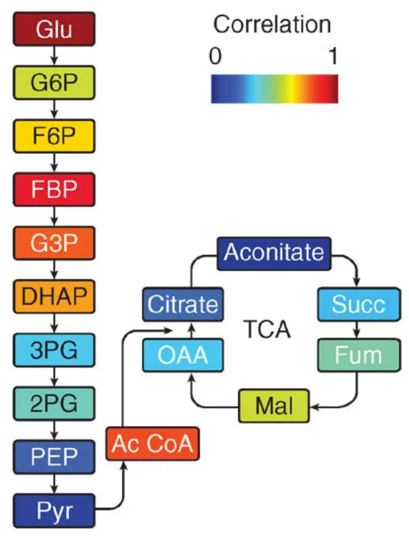


Use of colour

Colour can be very helpful for distinguishing between classes (categorical data).

It can be more difficult to construct an effective visualisation for quantitative data using a colour scale, because perceived change in hue / saturation / brightness do not map evenly to values [5].

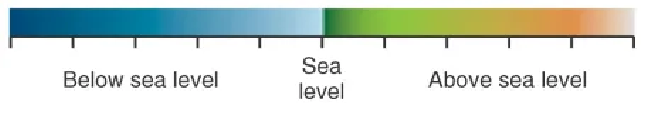
Use of colour



Rainbow scales have particularly uneven transitions and can be difficult to interpret.

Use of colour

Changes in hue have very high salience, so can be useful to emphasise zero-crossings where appropriate (e.g. a topographical map)



In general it will help communication to avoid relying on colour to communicate information [6].

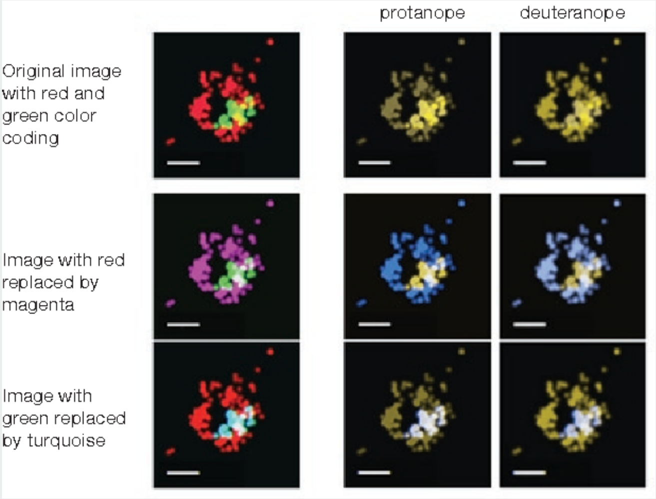
Colour blindness

Problems with colour vision are very common - around 4.5% of people in the UK are affected.

Any audience bigger than 15 is *more likely than not* to include a colour-blind person.

It is easy to adjust colours to avoid coding information as red/green (the most common form of colour blindness). [7]

Colour blindness



Colour blindness

If several colours are needed, use a palette that maximises the distinction between different colours.

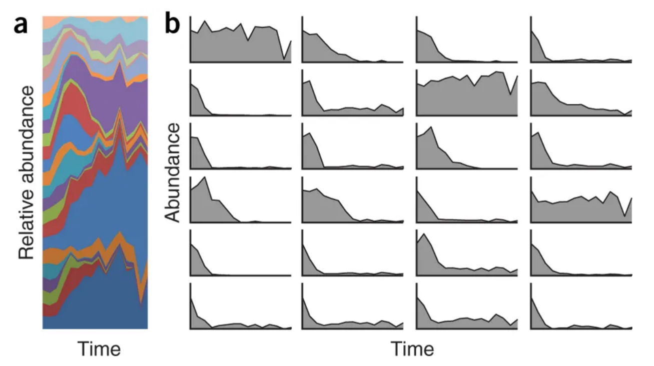
Color	Color name	RGB (1–255)	CMYK (%)	P	D
	Black	0, 0, 0	0, 0, 0, 100		
	Orange	230, 159, 0	0, 50, 100, 0		
	Sky blue	86, 180, 233	80, 0, 0, 0		
	Bluish green	0, 158, 115	97, 0, 75, 0		
	Yellow	240, 228, 66	10, 5, 90, 0		
	Blue	0, 114, 178	100, 50, 0, 0		
	Vermillion	213, 94, 0	0, 80, 100, 0		
	Reddish purple	204, 121, 167	10, 70, 0, 0		

Small multiples

Presenting *classes of behaviour* in high-dimensional data is a common visualisation goal.

By preparing lower-dimensional *slices* of the data, we can draw attention to the similarities and differences in behaviours [8].

Small multiples



Further reading

Since 2013, Nature Methods has run a regular feature looking at data visualisation best practice, called *Points of View*. [9].

These articles cover the most common data types and can be a great source of inspiration for difficult visualisation problems.

References i

[1] Wong, B. Design of data figures. Nat Methods 7, 665 (2010)

[2] E. R. Tufte. The Visual Display of Quantitative Information, 2nd Edition. Graphics Press, Cheshire, Connecticut, 2001.

[3] Wong, B. Salience. Nat Methods 7, 773 (2010)

[4] Wong, B. Salience to relevance. Nat Methods 8, 889 (2011)

[5] Gehlenborg, N., Wong, B. Mapping quantitative data to color. Nat Methods 9, 769 (2012)

[6] Wong, B. Avoiding color. Nat Methods 8, 525 (2011)

[7] Wong B. Points of view: Color blindness. Nat Methods 8:441 (2011)

References ii

[8] Shoresh, N., Wong, B. Data exploration. Nat Methods 9, 5 (2012)

[9] `http://blogs.nature.com/methagora/2013/07/data-visualization-points-of-view.html`