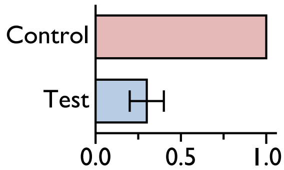
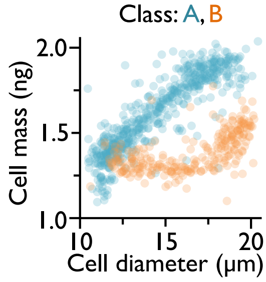
Figuring out Good Figures

If a picture is worth a thousand words, how many words is a scientific figure, chart or graph worth?

The content of figures is actually best thought of in terms of number of numbers...

This is a very simple figure showing just two numbers: that the test is 0.30±0.10 of the control.

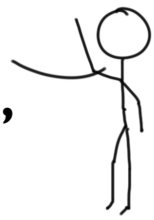




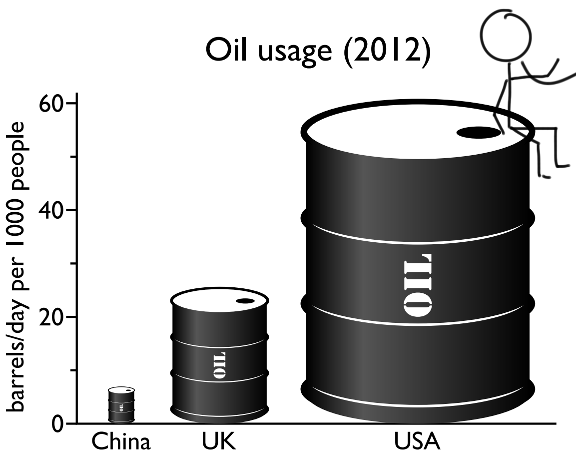
This figure is much more complex, and presents over two thousand numbers simultaneously.

Thinking about figures as a collection of numbers might seem weirdly abstract. Of course, figures aren't all **just** numbers. They often also include pieces of data better presented as images, 3D structures, etc. So why think about them as numbers? It lets you test one important thing: how good is the figure?

> “ Graphical excellence begins with telling the truth about the data. ”
>
> Edward Tufte, 1983



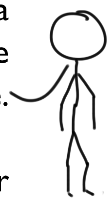
What is telling the truth about data? In short it is making sure the numbers (the actual data) that were used to make the figure are clear. Lying with a graph or figure is just as easy as accidentally misphrasing or deliberately obfuscating text. So what does a lying figure look like?



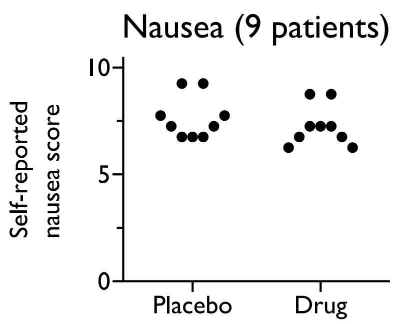
But this figure can't be lying; it's got a clear scale and labels!

Your brain intuitively interprets the **volume** of the barrels. It looks like the UK uses about an eighth as much oil per day as the USA. Actually the height of the barrel shows the true usage; that the UK uses about half that of the USA. The figure **tricks** you into a 400% error in extracting the data.

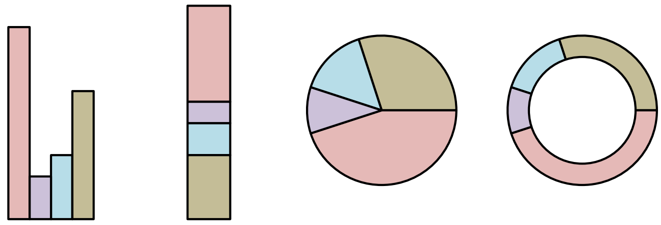
It might seem silly, but this figure is lying too. By adding artificial structure to the data (arranging the points as smiley faces) the true values of the data points are hard to see.



Which did you notice first: the sad face, or the lower mean, in the drug-treated group?



If a figure lies to you that certainly doesn't mean the author is trying to mislead you. It is often very hard to present data in the clearest and hardest to misinterpret way.

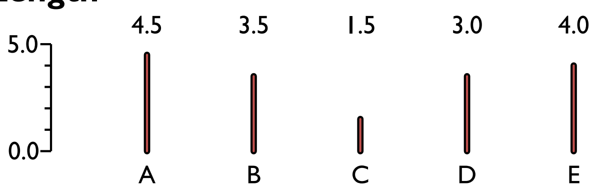


These four charts show exactly the same data, but which is clearest? Is there actually one that is better and easier to read than the rest?

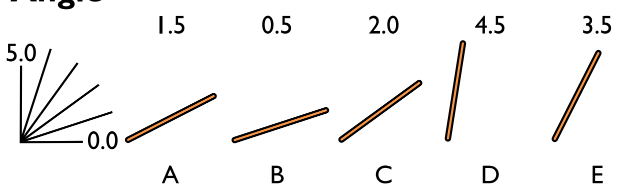
Why not test yourself?

Below are six different data presentation methods, each showing five measurements (labelled A, B, C, D and E) between 0.0 and 5.0. In each set of five measurements there is one deliberate mistake, which differs from the stated value by ±0.5. Can you identify which one it is? The answers are at the bottom of the page.

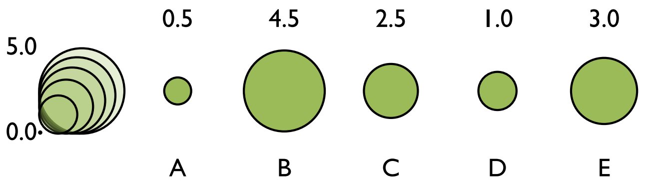
Length



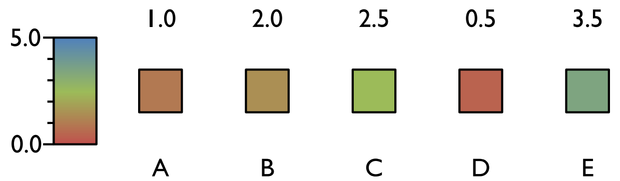
Angle



Area

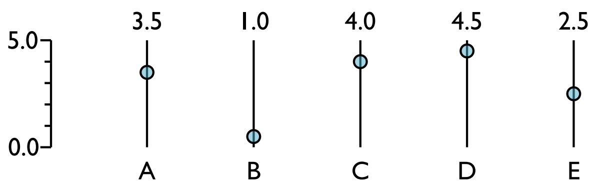


Colour

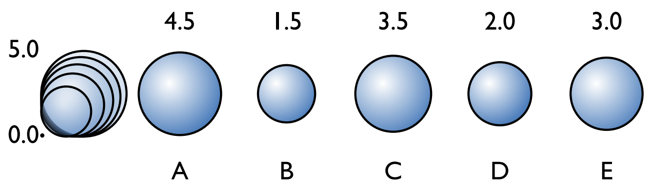


Apologies if you are colour blind, but that's part of the point!

Position



Volume



Most people have a similar order of ability in accurately interpreting data presented in these different ways:

Position is easier than Length is easier than Angle is easier than Area is easier than Volume is easier than Colour

This has important consequences: It means that people typically find pie charts (areas) are harder to read accurately than doughnut charts (angle), and both are harder than bar charts (length). It also means that images (colour brightness), and especially fluorescence colocalisation (colour hue or shade) are very hard to interpret quantitatively by eye.

It is almost impossible to make simple rules for designing the perfect, easy to understand figure for sharing precious scientific data, but with a little thought and energy a lot can be done to help your readers!

Answers: Length: D (actual value 3.5), Angle: B (actual value 1.0), Area: C (actual value 2.0), Colour: B (actual value 1.5), Position: B (actual value 0.5), Volume: D (actual value 2.5).

By Richard Wheeler 2013

Tufte ER (1983) The visual display of quantitative information. *Graphics Press*, Cheshire, CT, USA.
Wong B (2010) Points of view: Design of data figures. *Nat Methods* 7(9):665.