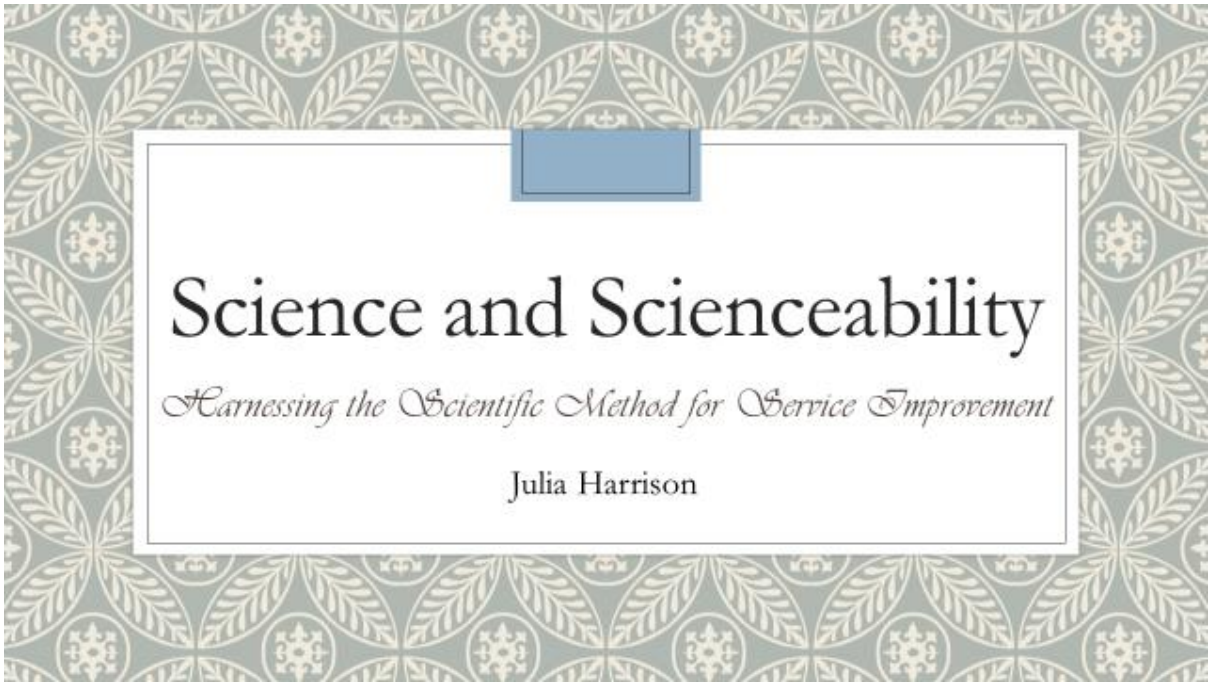
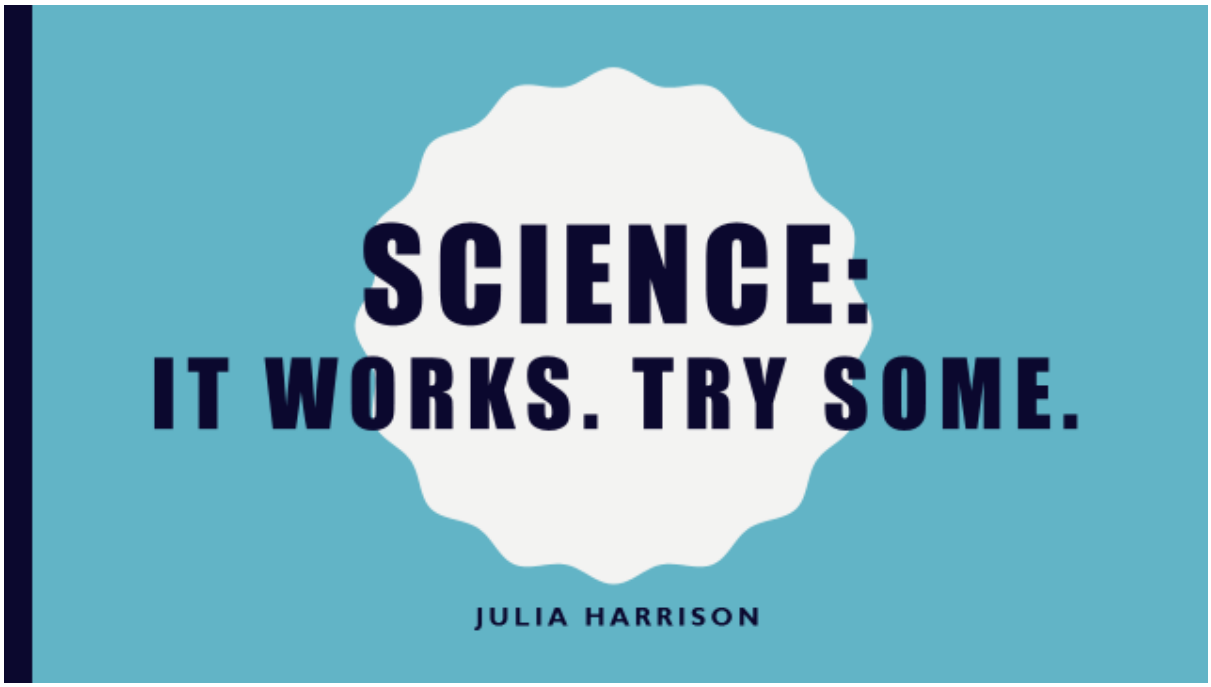


"It is a truth universally acknowledged, that a technology organisation in possession of a good budget, must be in want of a service improvement project."

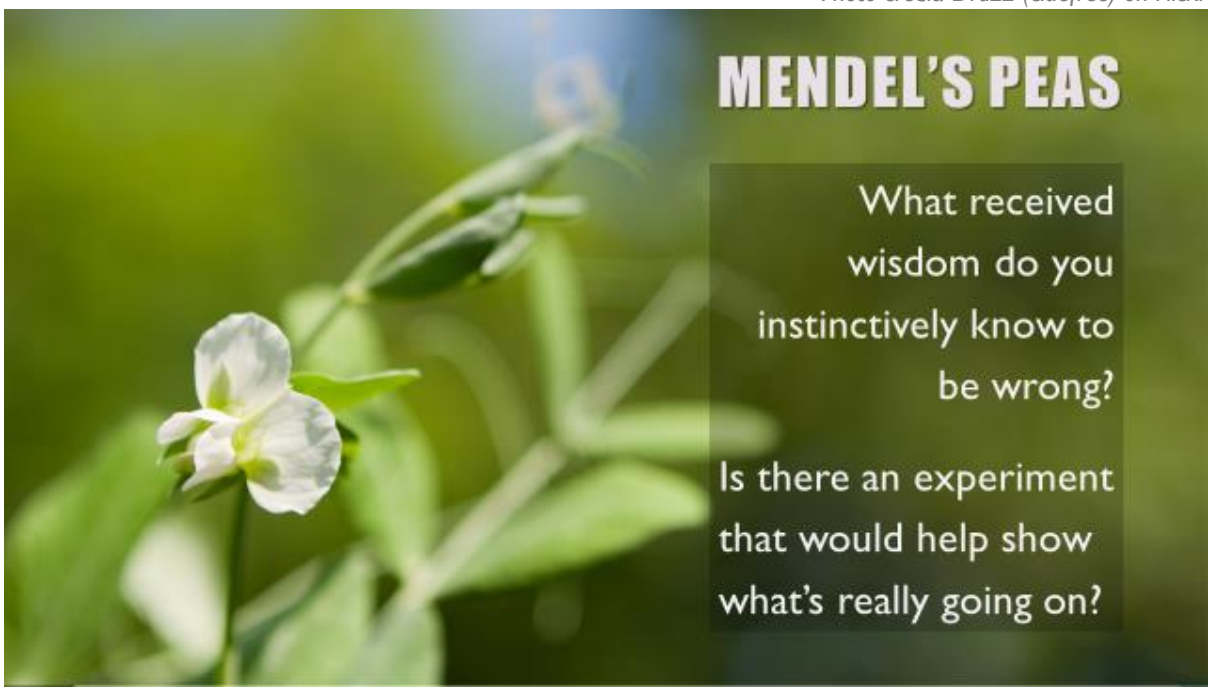
Apologies to Jane Austen, but I can't keep going in this style.



Is this better?

Caveat: I am not a scientist. And no amount of reading popular science books or enthusiasm will make me one. But I can try to learn from science and apply the scientific method to ITSM.

Photo credit: Drazz (cluefree) on Flickr

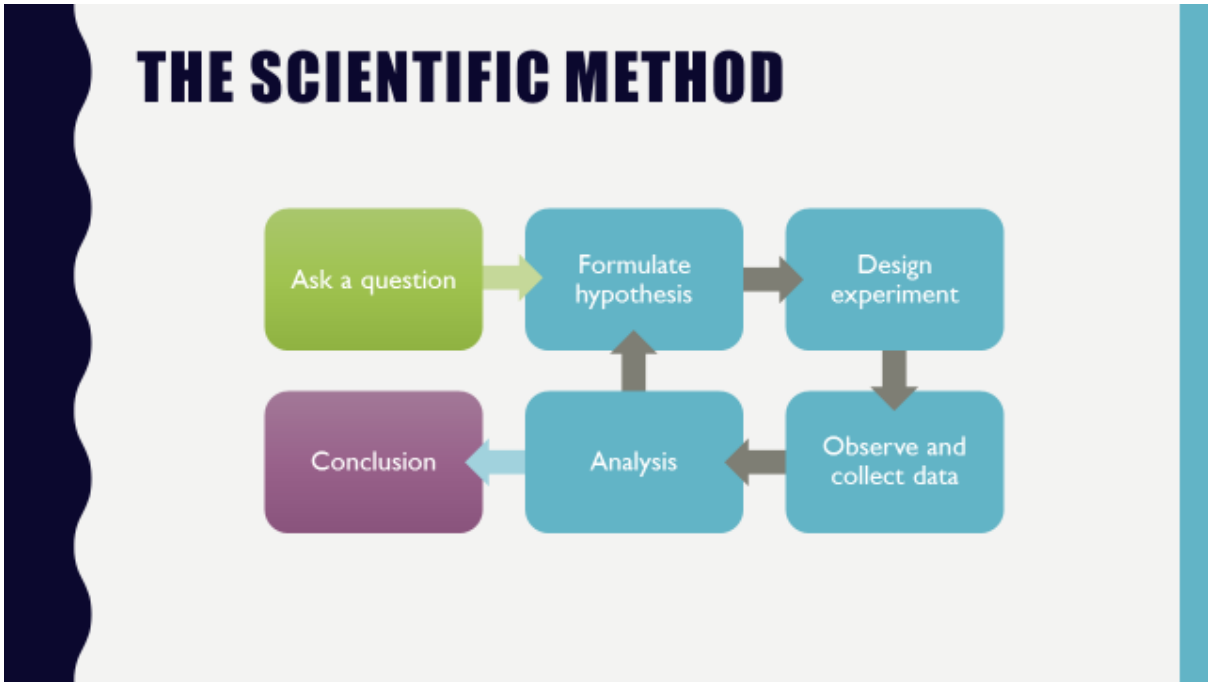


Conventional wisdom in the mid-nineteenth century was that offspring were a smooth blend of the characteristics of their parents. But Gregor Mendel noticed that this didn't explain how, for example, two brown-eyed people could have a blue-eyed child. So he set out to understand how traits were passed on from one generation to the next.

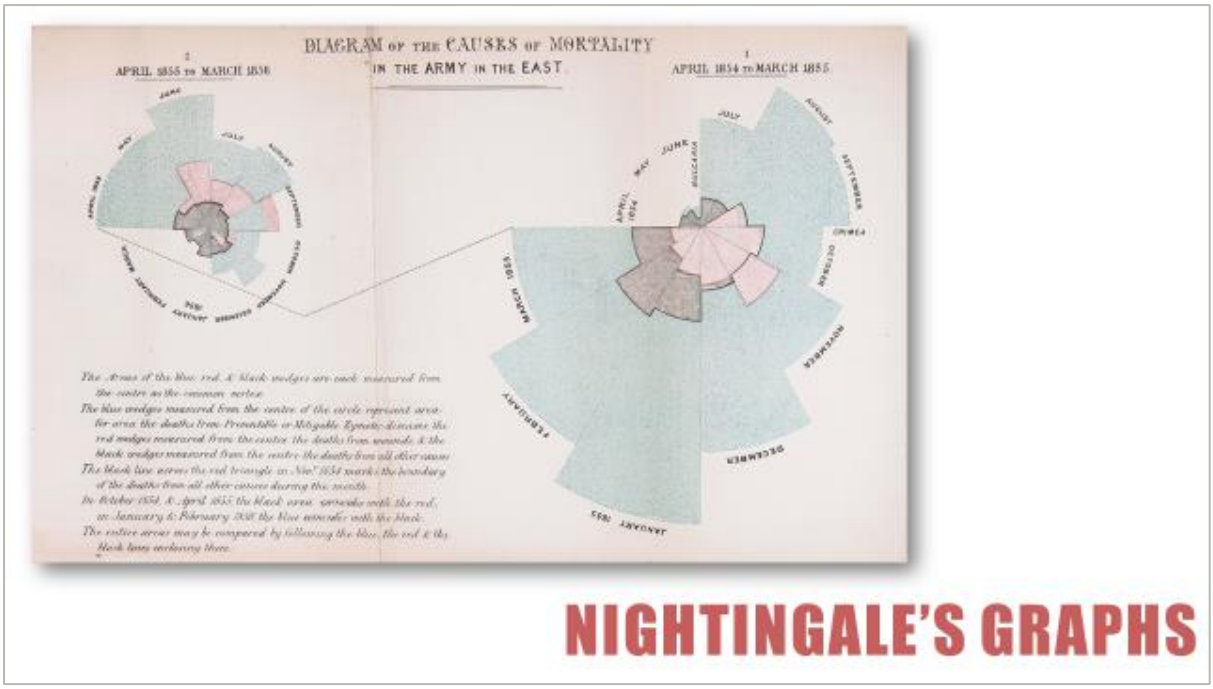
He designed an experiment to test this by breeding peas together in a systematic way and making detailed notes on characteristics of the resulting plants. His conclusions laid the foundations for what we know now about dominant and recessive traits in heredity.

Challenge: what received wisdom within your organisation do you instinctively know to be wrong? Is there an experiment that would help show what's really going on?

Example: tiered change approval processes may be helpful in reducing risk. But do they add value relative to the amount of effort they consume? How could we prove or disprove something like that? What kind of evidence could we gather?



The scientific method is a systematic process of **learning** by collecting and analysing evidence.



Another person who challenged received wisdom was Florence Nightingale.

During her work as a nurse in the Crimean war she noticed that more men were dying from disease than from their wounds, and started making changes to nursing care and ward hygiene to prevent and combat infection. And those changes worked.

At the time, data was usually represented in tables of numbers, and experts in medicine would be used to reading information that way. But for Nightingale's changes in the field hospitals to become common practice, she needed to convince army generals and politicians, not doctors, who were not used to consuming information in that way. So she came up with the visualisations you can see on the slide.

The graphic is a radial, time-based diagram. Each segment is a month, going clockwise. The chart on the right is before Nightingale's interventions, the chart on the left is from after. Even without being able to read the labels, if you learn that the blue segments represent the number of soldiers dying from illness and infection, and that the two charts are to the same scale, you can see the size of the problem, and the change between the two charts. Now imagine the effect that must have had on the audience who had never seen data represented like this before.

Challenge: What do you already have the data to demonstrate, but you're not able to convince others? Can you present it differently to make a better case?

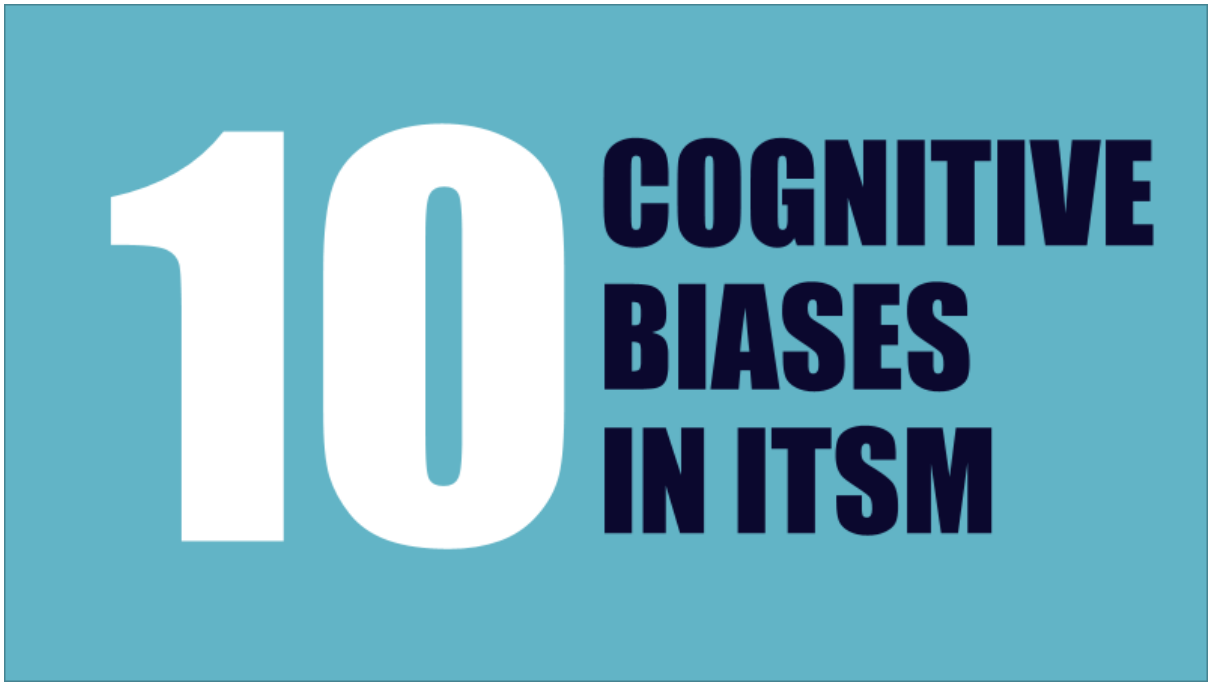
Think about Kanban and the advantages of being able to visualise work as cards on a board, like being able to see bottlenecks or relative volume of different work types (e.g. by card colours). A team whose stakeholders walk past their board regularly might benefit from the extra visibility this gives them.

Both Mendel and Nightingale started out with an observation or a hunch, but needed data to back it up. So does this mean we should trust our intuition more?

Intuition is lay-speak for *heuristic cognition* – mental shortcuts. As we experience the world, there is just too much information coming at us all the time for us to possibly process it, so our brain has to use patterns and shortcuts, and make assumptions.

But sometimes, the shortcuts lead us to the wrong conclusion. That's **bias**.

But it seems like our examples came from people with great intuition. And when you read books about successful business people, often they do too. So should we trust our intuition more? Or could it be that the people who trusted their intuition and failed horribly didn't go on to make history? That's **survivor bias**.



Here's a list of my ten favourite cognitive biases as they apply to ITSM:

1. Risk taking works, all our senior leaders took big risks in their careers!

Or is it just that all the other risk takers got fired and these guys got lucky? We may never know, thanks to **SURVIVOR BIAS**. This is the bias that takes success stories and assumes the steps they took led to success, even though many failures may have resulted from following the exact same steps.

2. I don't have any biases.

Oh yes you do. And the first one you'll need to control for is **BLIND-SPOT BIAS**, where we can see other people's biases, but not our own.

3. We hire the smartest developers, why would we invest in the QA platform?

OVERCONFIDENCE in our own abilities or those of our organisation is just another type of bias.

4. We can rate that change as Low Risk.

You can, and maybe you should. Or maybe it's **OPTIMISM BIAS**. Related to overconfidence, optimism bias also applies to outcomes we're not directly involved in.

5. Everyone has a two-hour in-person CAB meeting, so it must be a good idea.

You'll never get anything done on a Thursday afternoon, thanks to **THE BANDWAGON EFFECT** - thinking something is a good idea because everyone else is doing it.

6. Thanks for that report that took a month to prepare. I'm nearly ready to make a decision, but first can you break down the data by postcode?

Some decision makers just don't seem to like making decisions. It could be because of **INFORMATION BIAS**. That's the bias that tells us if we have just one more piece of information, we can remove the risk of a decision. And then another piece. And another piece.

7. The dashboard is green, everything is fine.

You may have heard of the watermelon effect (dashboards that are green on the outside, but red on the inside). We may be tempted to believe them because of the **AVAILABILITY HEURISTIC**, where we over-value the information that's in front of us, and don't go looking for the information that's really important.

8. Password resets are down over the summer, which must be as a result of my great leadership of the Remember Your Password campaign!

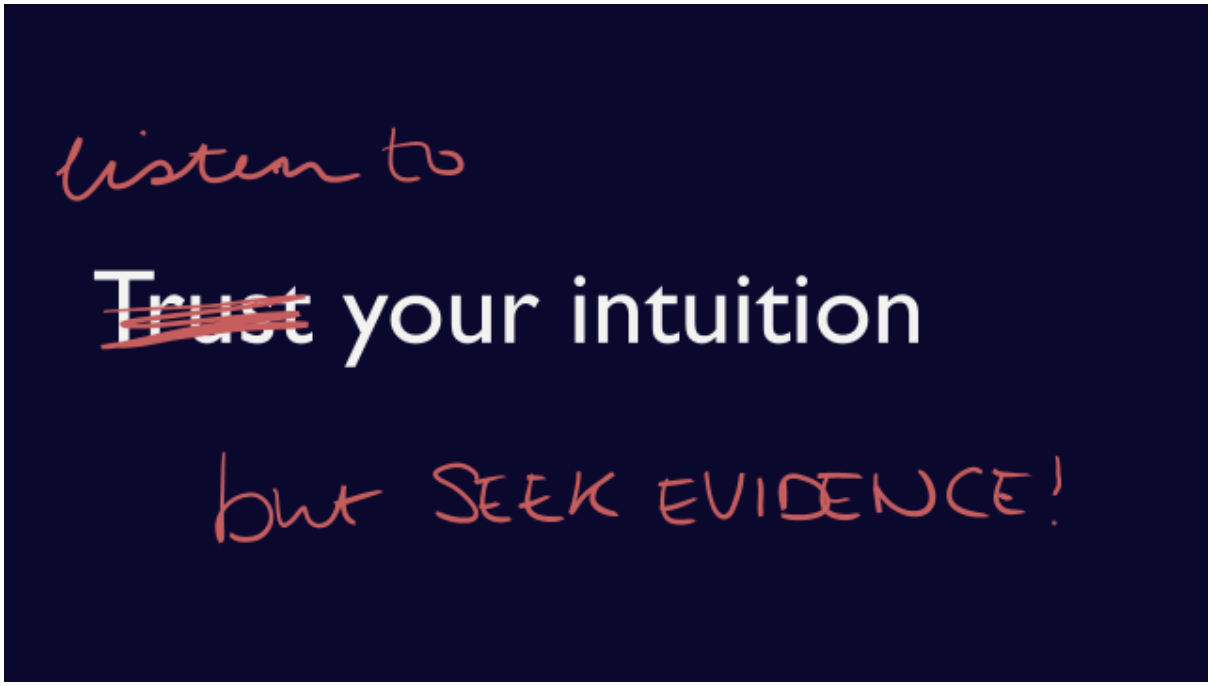
Or maybe it's just that a quarter of the users are on holiday. We'll find out in September whether it's a **FAULTY CAUSAL ASSUMPTION**. Thinking that one thing caused another, when they may be coincidental – also known as *post hoc ergo procter hoc* (after therefore because of) is more of a logical fallacy than a bias, but it's one of the most common and most dangerous so I can't leave it out on a technicality.

9. People love puns about science and Jane Austen.

And if I've succumbed to this one, obviously you all would as well, it's **PROJECTION BIAS** – believing that other people think and feel the same as we do.

10. Since I started learning about cognitive biases, I see them everywhere!

OBSERVATIONAL SELECTION BIAS happens when, once you become aware of a phenomenon, you start to see it everywhere. Which you probably will tomorrow.



Intuition is helpful, but it's where our biases live. To minimise the effects of our biases we should take a more systematic approach and seek evidence.

But how do we do that?

Photo credit: redwoodbusdriver on Instagram

In Office 2007, Microsoft introduced the Ribbon interface. That came as a response to a specific challenge: Microsoft had a wish list on their website, where anyone could submit a suggestion for a new product feature. But about 80% of the requests were for things that already existed. The team knew they had a problem with discoverability – users weren't able to find the functionality they wanted, so assumed it didn't exist.

Instrumentation (that is, code within an application to gather information about how it is used) was used to log which features were used and where people clicked to find them. In addition, groups of users were invited into labs where eye-tracking cameras were used to see where they were looking on the screen, and mock-ups of new designs were tested.

Challenge: How much do you really know about your users' experience? What data are you already collecting but not using? What can you only learn by observation?

Are you using performance monitoring for event and capacity management? Could it also be used to detect differences between production and pilot software, to understand the impact of potential changes before they become widespread?

Photo credit: tombullock on Flickr

Some things we can instrument or observe directly, some things we need to ask. But when we ask humans we introduce bias.

One type of bias associated with intervention (change) is the placebo effect. Usually this refers to medical trials, but might it occur more widely?

A placebo is a dummy pill, injection or other medical intervention. It looks to the patient like 'the real thing' but it does nothing. So the placebo effect is where, if someone believes that an intervention will help, they may think it has, even if it did nothing. (In medicine the placebo effect may *actually* help the patient get better, not just think they have.)

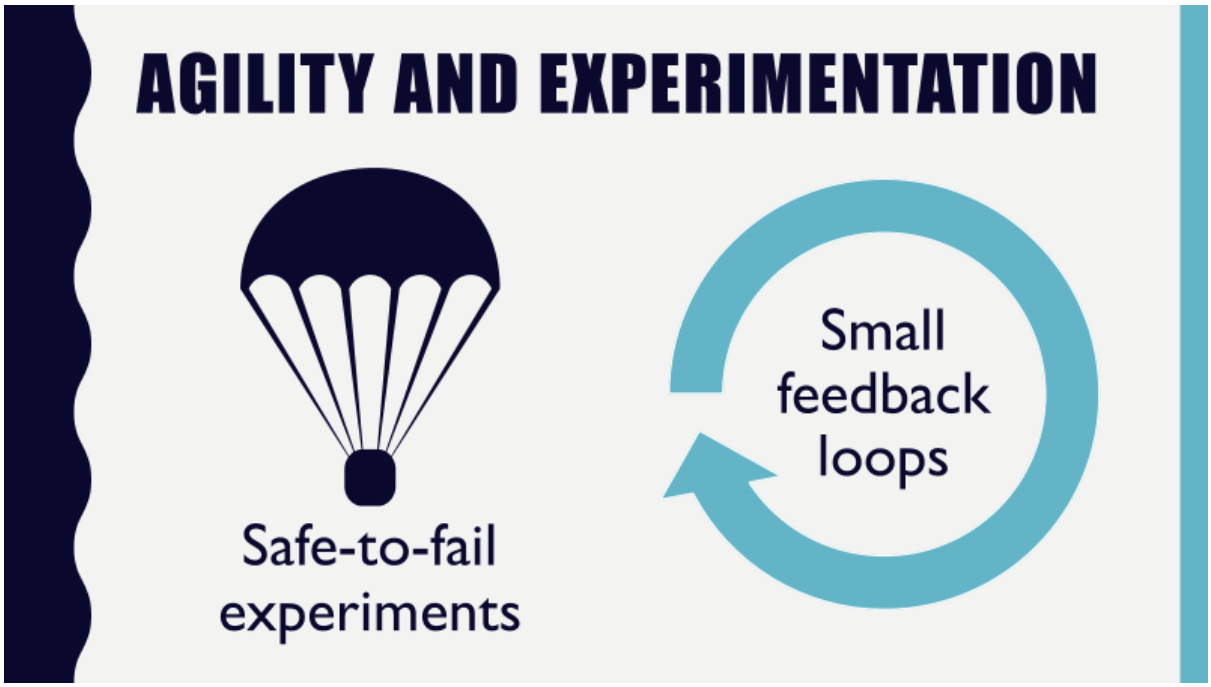
In ITSM this could manifest as users finding a new software version to be faster than the previous one, when there were only cosmetic changes. Once I overheard a user commenting on how well a network change was performing when, in fact, it had been cancelled the previous night.

The opposite can also be true (nocebo effect) – if someone is pessimistic about a change and expects it to go badly, they may perceive a service has degraded even if it's performing no worse than before.

Challenge: Which of your measurements are subjective? Might there be placebo and nocebo effects? How could you control for them?

I had this puzzle to solve with a software update some years ago: we hoped it would fix an intermittent issue, but the users had a workaround and had long since stopped logging incidents for it, so it would be difficult to tell if their experience had changed. We'd need to ask them to report any new occurrences, but that made the current baseline unreliable as a comparison.

So we doubled the size of our pilot group and told the users we were testing an update, and asked them to log all issues through the service desk. Then we rolled out the change to only half the group at random, and compared the level of reporting between the two halves. (Notice that we were careful not to lie to anyone in our communications!)



The key for ITSM is keeping things agile.

Safe-to-fail can mean easy to roll back. If you don't already test that changes can be quickly rolled back, this is an opportunity to reduce risk and enable more experimentation.

Small feedback loops result in learning quickly – what's the cheapest, shortest experiment to find out what we need to know right now?

Pilot rollouts are an example of both these things that we've been doing in ITSM for years.

Photo credit: Andrés Nieto Porras (anieto2k) on Flickr

Advertising has been running these kinds of experiments for decades.

Since the early 20th century, advertisers have tested the effectiveness of their ads by, for example, placing newspaper ads in just some regions, and comparing sales with regions where the ads weren't run, or against different versions of an ad.

Now services like Google Adwords allow advertisers to test different versions of an online ad and receive instant feedback about how each is performing.

Challenge: Where could you use A/B testing to improve interactions with your users, colleagues and partners?

For example, most organisations acknowledge that, while surveying users for feedback is easy, and response rates tend to be low and polarised. Could you randomly test different versions of survey invitation to see which perform best and make improvements?

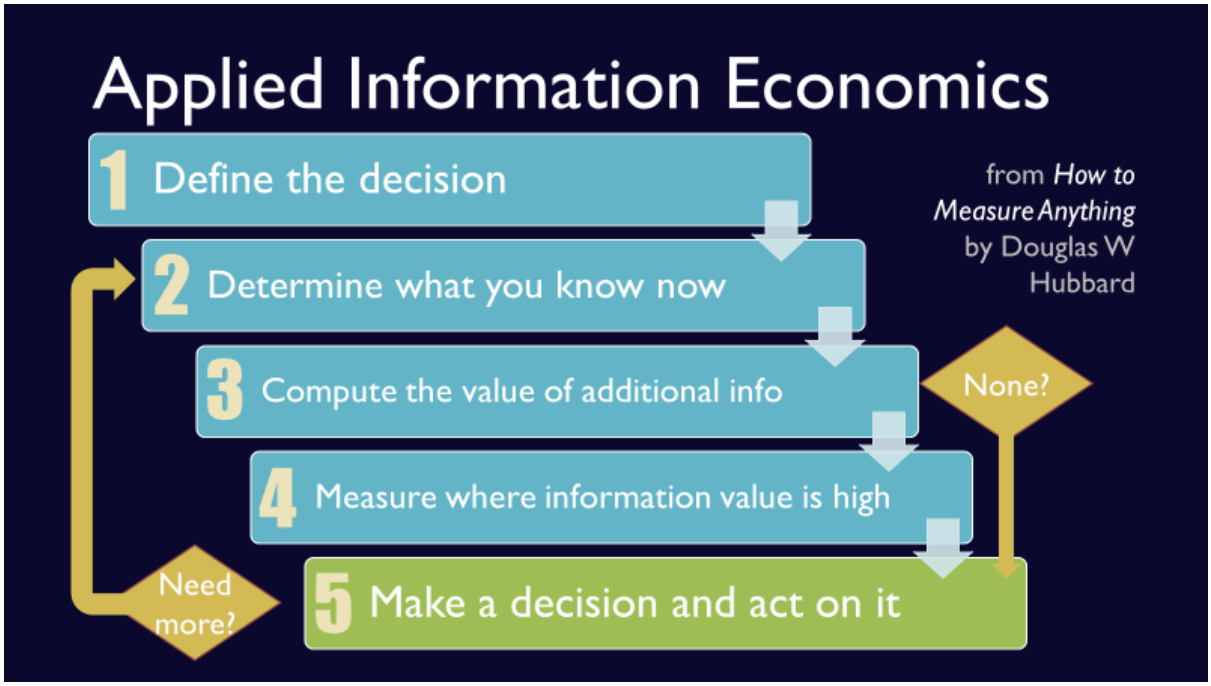
In 2009, Google's Gmail product was using a slightly different shade of blue for advertising hyperlinks to the one used on Google's search result pages. Rather than choose one or the other, it was decided to test forty (or forty-one by some accounts) different shades. People using Google search over the duration of the test would see links in one of the colours being tested, chosen at random.

The 'winning' shade of blue was associated with a $200M increase in advertising revenue from the extra clicks it generated.

Challenge: is there anything in ITSM where you'd benefit from testing so many different options?

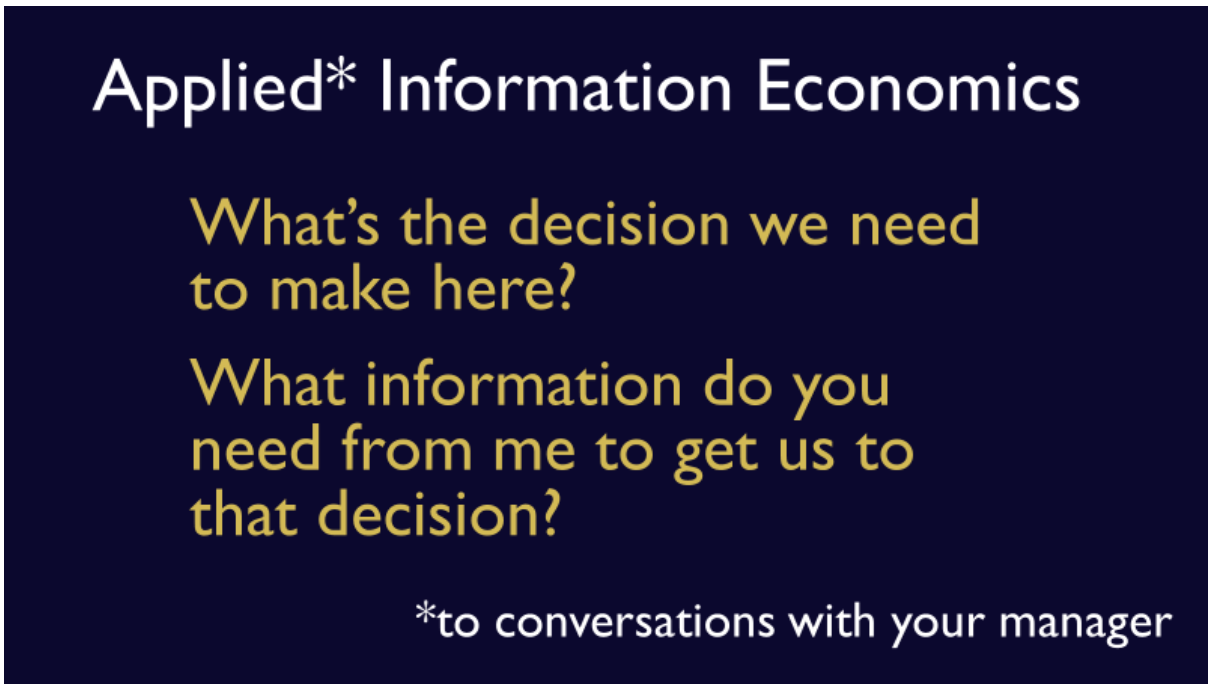
Maybe this is just a cautionary tale. Focus on value for your organisation, not satisfying idle curiosity.

And if you test multiple options, one *will win*. That doesn't necessarily mean it's better, it may just have won by chance. Testing for **statistical significance** will help quantify how likely it is that the result is meaningful and not just random.

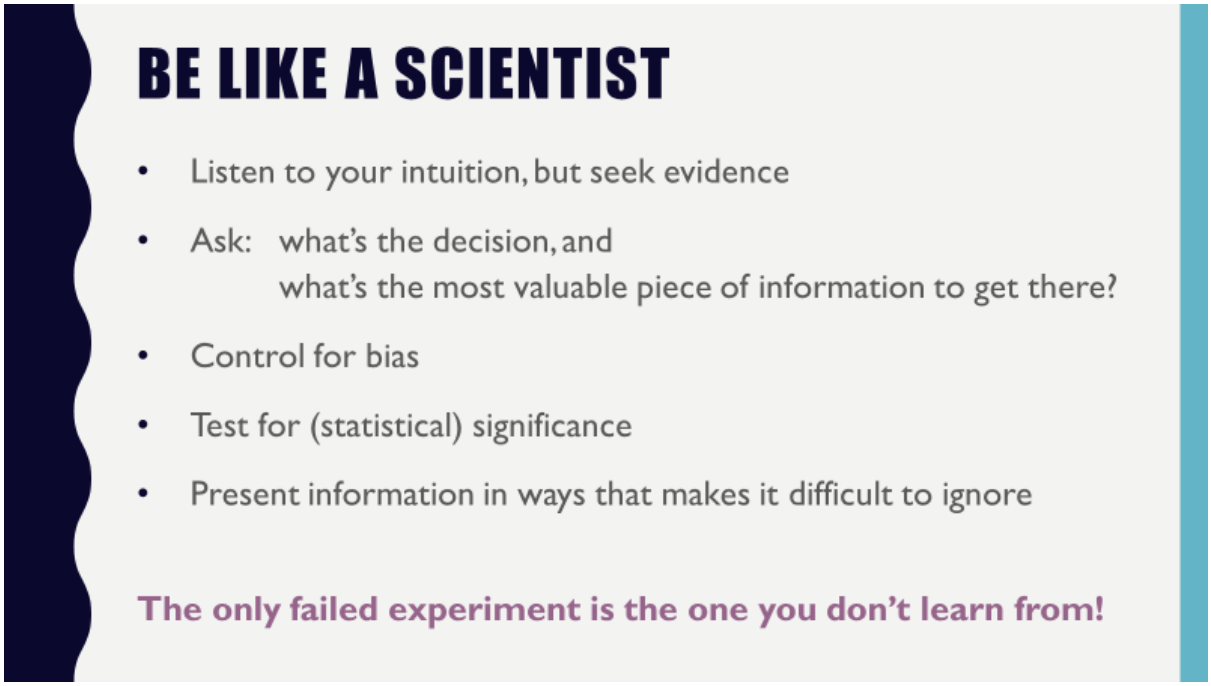


Instrumentation makes things cheap (free in some cases) to collect data, but that doesn't mean we should report on everything. Even if the data is free, your attention isn't. Looking at a hundred metrics usually comes at the cost of giving your full attention to the few that matter most.

The Applied Information Economics method comes from *How to Measure Anything* by Douglas W Hubbard.



If you ever find yourself sitting in front of a decision-maker in your organisation, presenting report after report, business case after business case, and seemingly getting no closer to a decision, it could be that **information bias** is a factor. Try using these two questions to break the deadlock.



You can find me on twitter at twitter.com/JuliafromIT