

## What could help fix belief problems?

[cross-post from Medium]

Last part (of 4) from a talk I gave about hacking belief systems.

### Who can help?

My favourite quote is “ignore everything they say, watch everything they do”. It was originally advice to girls about men, but it works equally well with both humanity and machines.

If you want to understand what is happening behind a belief-based system, you’re going to need a contact point with reality. Be aware of what people are saying, but also watch their actions; follow the money, and follow the data, because everything leaves a trace somewhere if you know how to look for it (looking may be best done as a group, not just to split the load, but also because diversity of viewpoint and opinion will be helpful).

Verification and validation (V&V) become important. Verification means going there. For most of us, verification is something we might do up front, but rarely do as a continuing practice. Which, apart from making people easy to phish, also makes us vulnerable to deliberate misinformation. Want to believe stuff? You need to do the leg-work of cross-checking that the source is real, finding alternate sources, or getting someone to physically go look at something and send photos (groups like findyr still do this). Validation is asking whether this is a true representation of what’s happening. Both are important; both can be hard.

One way to start V&V is to find people who are already doing this, and learn from them. Crisismappers used a bunch of V&V techniques, including checking how someone’s social media profile had grown, looking at activity patterns, tracking and contacting sources, and not releasing a datapoint until at least 3 messages had come in about it. We also used data proxies and crowdsourcing from things like satellite images (spending one painful Christmas counting and marking all the buildings in Somalia’s Afgooye region), automating that crowdsourcing where we could.

Another group that’s used to V&V is the intelligence community, e.g. intelligence analysts (IAs). As Heuer puts it in the *Psychology of Intelligence Analysis* (see the CIA’s online library), “*Conflicting information of uncertain reliability is endemic to intelligence analysis, as is the need to make rapid judgments on current events even before all the evidence is in*”. Both of these groups (IAs and crismappers) have been making rapid judgements about what to believe, how to handle deliberate misinformation, what to share and how to present uncertainty in it for years now. And they both have toolsets.

### Is there a toolset?

Graphs from a session on statistical significance, AB testing and friends. We can ‘clearly’ see when the two distributions are separate or nearly the same, but need tools to work out where the boundary between these two states are.

The granddaddy of belief tools is statistics. Statistics helps you deal with and code for uncertainty; it also makes me uncomfortable, and makes a lot of other people uncomfortable. My personal theory on that is it’s because statistics tries to describe uncertainty, and there’s always that niggling feeling that there’s something else we haven’t quite considered when we use do this.

Statistics is nice once you’ve pinned down your problem and got it into a nicely quantified form. But most data on belief isn’t like that. And that’s where some of the more interesting qualitative tools come in, including structured analytics (many of which are described in Heuer’s book “structured analytic techniques”). I’ll talk more about these in a later post, but for now I’ll share my favourite technique: Analysis of Competing Hypotheses (ACH).

ACH is about comparing multiple hypotheses; these might be problem hypotheses or solution hypotheses at any level of the data science ladder (e.g. from explanation to prediction and learning). The genius of ACH isn’t that it encourages people to come up with multiple competing explanations for something; it’s in the way that a ‘winning’ explanation is selected, i.e. by collecting evidence \*against\* each hypothesis and choosing the one with the least disproof, rather than trying to find evidence to support the “strongest” one. The basic steps (Heuer) are:

- Hypothesis. Create a set of potential hypotheses. This is similar to the hypothesis generation used in hypothesis-driven development (HDD), but generally has graver potential consequences than how a system’s users are likely to respond to a new website, and is generally encouraged to be wilder and more creative.
- Evidence. List evidence and arguments for each hypothesis.
- Diagnostics. List evidence against each hypothesis; use that evidence to compare the relative likelihood of different hypotheses.
- Refinement. Review findings so far, find gaps in knowledge, collect more evidence

(especially evidence that can remove hypotheses).

- Inconsistency. Estimate the relative likelihood of each hypothesis; remove the weakest hypotheses.
- Sensitivity. Run sensitivity analysis on evidence and arguments.
- Conclusions and evidence. Present most likely explanation, and reasons why other explanations were rejected.

This is a beautiful thing, and one that I suspect several of my uncertainty in AI colleagues were busy working on back in the 80s. It could be extended (and has been); for example, DHCP creativity theory might come in useful in hypothesis generation.

### **What can help us change beliefs?**

Boyd's original OODA loop diagram

Which us brings us back to one of the original questions. What could help with changing beliefs? My short answer is "other people", and that the beliefs of both humans and human networks are adjustable. The long answer is that we're all individuals (it helps if you say this in a Monty Python

voice), but if we view humans as systems, then we have a number of pressure points, places that we can be disrupted, some of which are nicely illustrated in Boyd's original OODA loop diagram (but do this gently: go too quickly with belief change, and you get shut down by cognitive dissonance). We can also be disrupted as a network, borrowing ideas from biostatistics, and creating idea 'infections' across groups. Some of the meta-level things that we could do are:

- Teaching people about disinformation and how to question (this works best on the young; we old 'uns can get a bit set in our ways)
- Making belief differences visible
- Breaking down belief 'silos'
- Building credibility standards
- Creating new belief carriers (e.g. meme wars)