

Earning credibility in post-factual science?

Björn Brembs 

Published February 2, 2016

Citation

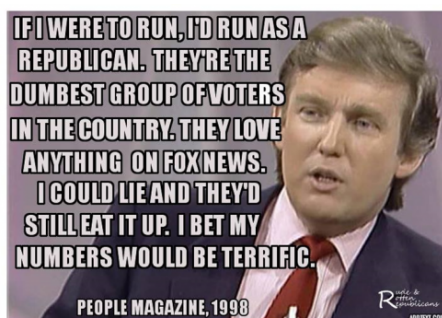
Brembs, B. (2016, February 2). Earning credibility in post-factual science?. *Bjoern.brembs.blog*.
<https://doi.org/10.59350/cc9hb-21877>

Keywords

Science Politics, Politics, Post-factual Science, Society, Truthiness

Abstract

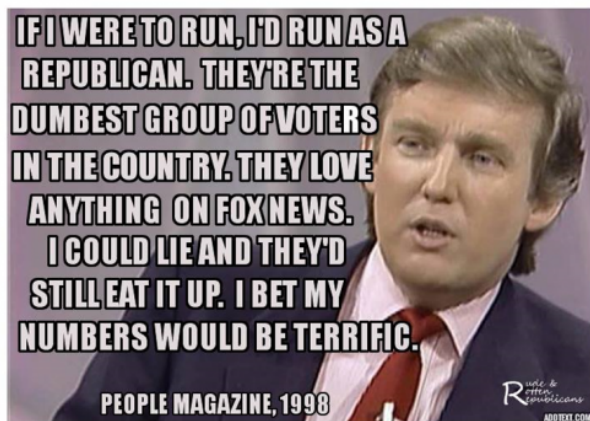
What do these two memes have in common? While they may have more than one thing in common, the point important for now is that despite both having an air of plausibility or ‘truthiness’ around them, they’re both false: neither [...] [↓ Read the rest of this entry...](#)



Copyright

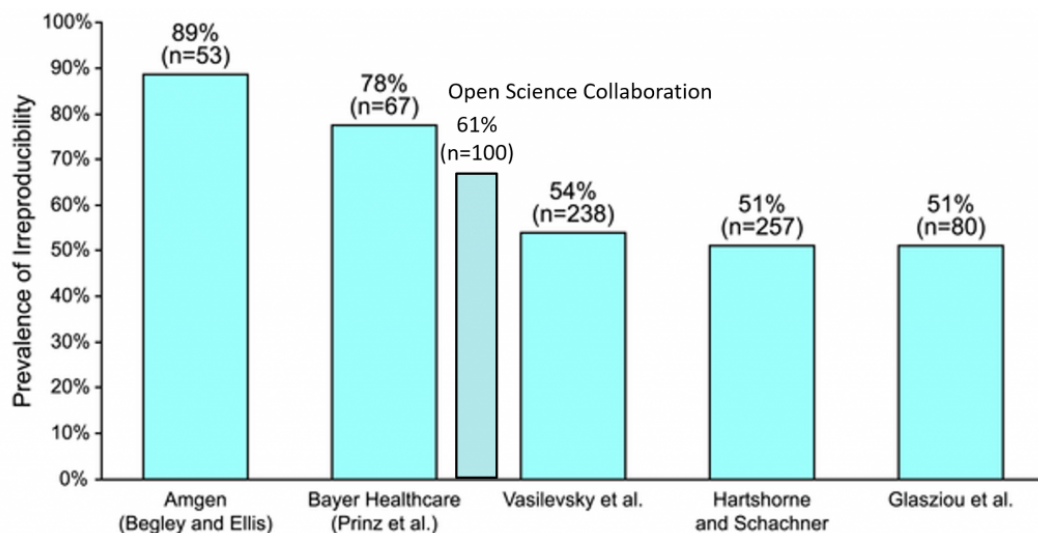
Copyright © Björn Brembs 2016. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

What do these two memes have in common?



While they may have more than one thing in common, the point important for now is that despite both having an air of plausibility or 'truthiness' around them, they're both false: neither has Donald Trump [ever said](#) these words to *People* magazine, nor do gay canvassers [have such an effect](#) on people's attitudes (even though the quoted statement about this research was indeed published in *Science* magazine).

The issue of false facts has become so rampant in current politics, that some have [dubbed](#) our common era "post-factual". While on the surface, "post-factual science" appears to be an oxymoron, recent evidence raises the tantalizing possibility that at least the literature of the life-sciences, broadly speaking, may be on track to leaving the reality-based community (although more research is required):



Irreproducibility loosely defined as in "not replicated", or "difficult to replicate with imprecise methods". Note that not all of these are replication studies and not all of the replication studies are properly published themselves, with data available etc. To my knowledge, only the Open Science study is reliable in that respect. Sources: <https://journals.plos.org/plosbiology/article?id=info:doi/10.1371/journal.pbio.1002165> and <https://science.sciencemag.org/content/349/6251/aac4716>

Obviously, this data can only hold for experimental sciences and even there I would expect huge differences between the different sub-fields. Nevertheless, the frequency of retractions in biomedicine is rising exponentially, the 'majority of research findings' was [estimated](#) and more recently supported (at least tentatively for the findings analyzed/replicated in the above six studies) to be false and the public trust in science is [eroding](#). Maybe it is not too early to start asking the question: if we can't even trust the scientific literature, what information is trustworthy? Or, put differently: if traditional (some would say legacy) strategies for assigning credibility are failing us, are there more modern strategies which can replace or at least support them?

Personally, I think this is one of the most urgent and important challenges of the post-internet globalized society. It may well be that science, which brought us the internet in the first place, may also be the right place to start looking for a solution.

Truth has never been an easy concept. Some may even argue that a large portion of humanity is probably quite happy with it being rather malleable and pliable. It wasn't really until the 20th century that [epistemologists](#) have explicitly formulated a convention of what constitutes a scientific fact and how they can be used to derive and test scientific statements. Obviously, in life outside of science we are still far from an explicit convention, which complicates matters even further.

Be that as it may, neither within nor outside of science can we expect every individual to always fact-check, question and investigate every single statement or bit of information ever encountered, no matter how much that may be desired. There will always, inevitably have to be shortcuts and we have been taking them, legitimately, I would argue, since the beginning of humanity.

Initially, these shortcuts involved authority. With enlightenment and the shedding of both religious, societal and political shackles, a (competing?) more democratic and informal shortcut was conceived (without ever completely replacing authority): popularity. If a sufficiently large number of information outlets confirmed a statement, it was considered sufficient for believing it. Both shortcuts are still being used today both inside and outside of science, for very legitimate reasons. However, in part because the choice of authority is often arbitrary and may even be subject to partisan dispute, there are few, if any, universal authorities left. Even in science, the outlets with the most authority have been shown to publish the [least reliable](#) science. This erosion of the 'authority' shortcut accelerated with the dawn of the internet age: never before was it so easy to get so many people to repeat so many falsehoods with such conviction.

The 'wisdom of the crowd' seems to suggest that a sufficiently large crowd can be at least as accurate as a small number of expert authorities, if not more so. Social media have the uncanny ability of always aggregating what subjectively feels like a "sufficiently large crowd" to solidly refute any and all authority, whether that would be on the moon landings, the 9/11 attacks, vaccination effectiveness/risk, climate change, crime/gun control, or spherical earth. Obviously, this not only constitutes a fatal misunderstanding of how the crowd becomes wise, it

also contributes to an unjustified, exaggerated distrust in entities which do have significant expertise and hence credibility.

There are several reasons as to why social media, as currently implemented are notoriously incapable of getting anywhere near an ideal ‘wisdom of the crowd’ effect. For one, social feedback loops tend to aggregate people who think alike, i.e., reduce heterogeneity, when [diversity](#) is one of the most decisive factors in achieving crowd wisdom. Second, with our stone-age concept of a crowd, we may intellectually understand, but fail to intuitively grasp that any group of less than a few tens of millions is more of an intimate gathering rather than a crowd, on internet scales. With today’s social media implementation, it is comparatively easy to quickly gather a few million people who all share some fringe belief.

For instance, if we round the incidence of schizophrenia to about 1% of the population and assume, for simplicity’s sake, that all of them harbor either auditory hallucinations or some other forms of delusions. Hence, if 1% of all [internet users](#) were delusional in some way or form and only half of them aggregated in an online patient forum, we’d be talking about more than 15 million individuals. I’m pretty sure that such a patient site would deliver some quite astounding news feed with an amazing commentariat. And this is just one of a growing list of psychiatric disorders. How many nominally healthy users would feel compelled to believe the news from this site, re-share items and comment approvingly? Most far-out-there communities have orders of magnitudes less users but disproportionately large visibility. Indeed, some of those communities may appear just like a subforum of such a patient site, but without a single user actually being diagnosed with any psychiatric disorder whatsoever.

So if we want to take advantage of the [micro-expertise](#) of the individuals in a crowd, that crowd needs to be not only ‘sufficiently’ large for the task at hand, it more importantly needs to be sufficiently diverse, or size quickly becomes almost completely irrelevant. From numerous examples in science and elsewhere, it seems straightforward to argue that we need to harness the individual micro-expertise that *anyone* can have, without making the mistake to attribute this expertise to *everyone*. Authority alone cannot serve as a reliable shortcut for credibility, but neither can popularity alone. Here is an idea of how one might combine them.

I may be wrong, but at least for now I would argue that we probably cannot start from scratch, assigning every institution, organization and individual the same kind of credibility. We cannot and should not undo history and track records based in evidence: there are information sources that have a higher credibility than others.

Further, we probably need a score or at least ranks that get computed iteratively and recursively. For any individual or piece of information to gather points or climb ranks, there need to be arbiters that already have some credibility – another reason why we likely won’t be able to start with a level playing field. What is less clear is how such a scoring/ranking system ought to be designed: I somehow have the impression that it ought to be difficult to earn credibility, shouldn’t it? Of course, it’s usually always “innocent until proven guilty”, but is this a practical approach when doling out credibility? Should we all start with a credibility of 100 and then lose it? Or should we start with 0 and then gain? Does such a score have to go negative?

So far, these ideas have been very vague and general. Here are my first thoughts on how one may go about implementing such a system in science. A prerequisite for such a system is, of course, a [modern scholarly information infrastructure](#). This won't work with the 350 year-old technology we call 'journals'.

Because we need diversity and inclusiveness, one would never prevent anybody from posting whatever they have discovered. However, if someone described a discovery from a known research institution, that discovery would receive more credibility than if it were posted by a layman without a track record (even though both scores would still be relatively low at the point of publication). Similarly, if the author list contained professors, the article would receive more credibility than if there were only graduate student authors. Yet more credibility would be assigned if the data and code relevant for the discovery were openly available as well. Once this initial stage had been completed, the document and its affiliated authors and institutions can earn even more credibility, for instance if the code gets verified, or the data checked for completeness and documentation. Those doing the verification and reviewing also need to be diverse, so also here there should not be a principle limit. However, the weight given to each verification (or lack thereof) will be different according to the scores of the person doing the verification. More credit for reproducible data analysis (e.g. via [docker](#) et al.) and if the narrative accompanying the data/code is supported by them. This whole process would be similar to current peer-review, albeit more fine-grained, likely by more people (each contributing a smaller fraction of the reviewing work) and not necessarily on each and every article.

This process continues to accrue (or lose) credibility inasmuch as the article is receiving attention with consequences, e.g., how many views for each citation (in accordance with a citation typology, e.g. [CITO](#)), how many views for each comment, endorsement or recommendation, etc. This is one possible way of normalizing for field size, another could be by analyzing citation networks (as in, e.g. [RCR](#)). Clearly, most credibility ought to be associated with the experiments in each article being actually reproduced by independent laboratories (i.e., a special kind of citation).

In this iterative process, each participant receives credit in various ways for their various activities. Credibility would be just one of several factors being constantly monitored. Points can be awarded both for receiving (and passing!) scrutiny from others as well as scrutinizing other people's work. The resulting system is thought to allow everyone to check the track record of everyone else for some data on how reliable the work of this person (or institution, or community) has been so far, along with more details on how the track record was generated.

Obviously, there are plenty of feedback loops involved in this system, so care has to be taken to make these loops balance each other. The feedback loops found in many biological processes would serve as excellent examples of how to accomplish this. Complex systems like this are also known to be notoriously difficult to game.

Those are still very rough ideas, without a clear picture of the most suitable or effective implementation, yet, or whether the desired outcomes can actually be achieved in this way. I

also have no good idea how one would take such a system to leverage it outside of science. I would like to hope, however, that by starting on the easier case of science, we may be able to approach a related system for the society at large.