

[REPOST] Epistemic Learned Helplessness

Posted on June 3, 2019 by Scott Alexander



This is a slightly edited repost of an essay from my old LiveJournal.

*
**

A friend recently complained about how many people lack the basic skill of believing arguments. That is, if you have a valid argument for something, then you should accept the conclusion. Even if the conclusion is unpopular, or inconvenient, or you don't like it. He envisioned an art of rationality that would make people believe something *after it had been proven to them*.

And I nodded my head, because it sounded reasonable enough, and it wasn't until a few hours later that I thought about it again and went "Wait, no, that would be a terrible idea."

I don't think I'm overselling myself too much to expect that I could argue circles around the average uneducated person. Like I mean that on most topics, I could demolish their position and make them look like an idiot. Reduce them to some form of "Look, everything you say fits together and I can't explain why you're wrong, I

just know you are!” Or, more plausibly, “Shut up I don’t want to talk about this!”

And there are people who can argue circles around me. Maybe not on every topic, but on topics where they are experts and have spent their whole lives honing their arguments. When I was young I used to read pseudohistory books; Immanuel Velikovsky’s [Ages in Chaos](#) is a good example of the best this genre has to offer. I read it and it seemed so obviously correct, so perfect, that I could barely bring myself to bother to search out rebuttals.

And then I read the rebuttals, and they were so obviously correct, so devastating, that I couldn’t believe I had ever been so dumb as to believe Velikovsky.

And then I read the rebuttals to the rebuttals, and they were so obviously correct that I felt silly for ever doubting.

And so on for several more iterations, until the labyrinth of doubt seemed inescapable. What finally broke me out wasn’t so much the lucidity of the consensus view so much as starting to sample different crackpots. Some were almost as bright and rhetorically gifted as Velikovsky, all presented insurmountable evidence for their theories, and all had mutually exclusive ideas. After all, Noah’s Flood couldn’t have been a cultural memory both of the fall of Atlantis *and* of a change in the Earth’s orbit, let alone of a lost Ice Age civilization or of megatsunamis from a meteor strike. So given that at least some of those arguments are wrong and all seemed practically proven, I am obviously just gullible in the field

of ancient history. Given a total lack of independent intellectual steering power and no desire to spend thirty years building an independent knowledge base of Near Eastern history, I choose to just accept the ideas of the prestigious people with professorships in Archaeology, rather than those of the universally reviled crackpots who write books about [Venus being a comet](#).

You could consider this a form of epistemic learned helplessness, where I know any attempt to evaluate the arguments is just going to be a bad idea so I don't even try. If you have a good argument that the Early Bronze Age worked completely differently from the way mainstream historians believe, I just don't want to hear about it. If you insist on telling me anyway, I will nod, say that your argument makes complete sense, and then totally refuse to change my mind or admit even the slightest possibility that you might be right.

(This is the correct Bayesian action: if I know that a false argument sounds just as convincing as a true argument, argument convincingness provides no evidence either way. I should ignore it and stick with my prior.)

I consider myself lucky in that my epistemic learned helplessness is circumscribed; there are still cases where I'll trust the evidence of my own reason. In fact, I trust it in most cases other than infamously deceptive arguments in fields I know little about. But I think the average uneducated person doesn't and shouldn't. Anyone anywhere – politicians, scammy businessmen, smooth-talking romantic partners – would be able to argue them into anything. And so

they take the obvious and correct defensive maneuver – they will never let anyone convince them of any belief that sounds “weird”.

(and remember that, if you grow up in the right circles, beliefs along the lines of “astrology doesn’t work” sound “weird”.)

This is starting to resemble ideas like compartmentalization and [taking ideas seriously](#). The only difference between their presentation and mine is that I’m saying that for 99% of people, 99% of the time, taking ideas seriously is the wrong strategy. Or, at the very least, it should be the last skill you learn, after you’ve learned every other skill that allows you to know which ideas are or are not correct.

The people I know who are best at taking ideas seriously are those who are smartest and most rational. I think people are working off a model where these co-occur because you need to be very clever to resist your natural and detrimental tendency not to take ideas seriously. But I think they might instead co-occur because you have to be really smart in order for taking ideas seriously not to be immediately disastrous. You have to be really smart not to have been talked into enough terrible arguments to develop epistemic learned helplessness.

Even the smartest people I know have a commendable tendency not to take certain ideas seriously. Bostrom’s [simulation argument](#), the [anthropic doomsday argument](#), [Pascal’s Mugging](#) – I’ve never heard anyone give a coherent argument against any of these, but

I've also never met anyone who fully accepts them and lives life according to their implications.

A friend tells me of a guy who once accepted fundamentalist religion because of Pascal's Wager. I will provisionally admit that this person "takes ideas seriously". Everyone else gets partial credit, at best.

Which isn't to say that some people don't do better than others. Terrorists seem pretty good in this respect. People used to talk about how terrorists must be very poor and uneducated to fall for militant Islam, and then someone did a study and found that they were disproportionately well-off, college educated people (many were engineers). I've heard a few good arguments in this direction before, things like how engineering trains you to have a very black-and-white right-or-wrong view of the world based on a few simple formulae, and this meshes with fundamentalism better than it meshes with subtle liberal religious messages.

But to these I'd add that a sufficiently smart engineer has never been burned by arguments above his skill level before, has never had any reason to develop epistemic learned helplessness. If Osama comes up to him with a really good argument for terrorism, he thinks "Oh, there's a good argument for terrorism. I guess I should become a terrorist," as opposed to "Arguments? You can prove anything with arguments. I'll just stay right here and not blow myself up."

Responsible doctors are at the other end of the spectrum from terrorists here. I once heard someone rail against how doctors totally ignored all the latest and most exciting medical studies. The same person, practically in the same breath, then railed against how [50% to 90% of medical studies are wrong](#). These two observations are not unrelated. Not only are there so many terrible studies, but pseudomedicine (not the stupid homeopathy type, but [the type that links everything to some obscure chemical on an out-of-the-way metabolic pathway](#)) has, for me, proven much like pseudohistory – unless I am an expert in that particular subsubfield of medicine, it can sound very convincing even when it's very wrong.

The medical establishment offers a shiny tempting solution. First, a total unwillingness to trust anything, no matter how plausible it sounds, until it's gone through an endless cycle of studies and meta-analyses. Second, a bunch of Institutes and Collaborations dedicated to filtering through all these studies and analyses and telling you what lessons you should draw from them.

I'm glad that some people never develop epistemic learned helplessness, or develop only a limited amount of it, or only in certain domains. It seems to me that although these people are more likely to become terrorists or Velikovskians or homeopaths, they're also the only people who can figure out if something basic and unquestionable is wrong, and make this possibility well-known enough that normal people start becoming willing to consider it.

But I'm also glad epistemic learned helplessness exists. It seems like a pretty useful social safety valve most of the time.

