


< PREDICTABLY WRONG >

Preface

291

by **Eliezer Yudkowsky**  3 min read  11th Mar 2015 8 comments

You hold in your hands a compilation of two years of daily blog posts. In retrospect, I look back on that project and see a large number of things I did completely wrong. I'm fine with that. Looking back and *not* seeing a huge number of things I did wrong would mean that neither my writing nor my understanding had improved since 2009. *Oops* is the sound we make when we improve our beliefs and strategies; so to look back at a time and not see anything you did wrong means that you haven't learned anything or changed your mind since then.

It was a mistake that I didn't write my two years of blog posts with the intention of helping people do better in their everyday lives. I wrote it with the intention of helping people solve big, difficult, important problems, and I chose impressive-sounding, abstract problems as my examples.

In retrospect, this was the second-largest mistake in my approach. It ties in to the *first* -largest mistake in my writing, which was that I didn't realize that the big problem in learning this valuable way of thinking was figuring out how to practice it, not knowing the theory. I didn't realize that part was the priority; and regarding this I can only say "Oops" and "Duh."

Yes, sometimes those big issues really are big and really are important; but that doesn't change the basic truth that to master skills you need to practice them and it's harder to practice on things that are further away. (Today the Center for Applied Rationality is working on repairing this huge mistake of mine in a more systematic fashion.)

A third huge mistake I made was to focus too much on rational belief, too little on rational action.

The fourth-largest mistake I made was that I should have better organized the content I

was presenting in the sequences. In particular, I should have created a wiki much earlier, and made it easier to read the posts in sequence.

That mistake at least is correctable. In the present work Rob Bensinger has reordered the posts and reorganized them as much as he can without trying to rewrite all the actual material (though he's rewritten a bit of it).

My fifth huge mistake was that I—as I saw it—tried to speak plainly about the stupidity of what appeared to me to be stupid ideas. I did try to avoid the fallacy known as Bulverism, which is where you *open* your discussion by talking about how stupid people are for believing something; I would always discuss the issue first, and only afterwards say, “And so this is stupid.” But in 2009 it was an open question in my mind whether it might be important to have some people around who expressed contempt for homeopathy. I thought, and still do think, that there is an unfortunate problem wherein treating ideas courteously is processed by many people on some level as “Nothing bad will happen to me if I say I believe this; I won’t lose status if I say I believe in homeopathy,” and that derisive laughter by comedians can help people wake up from the dream.

Today I would write more courteously, I think. The discourtesy did serve a function, and I think there were people who were helped by reading it; but I now take more seriously the risk of building communities where the normal and expected reaction to low-status outsider views is open mockery and contempt.

Despite my mistake, I am happy to say that my readership has so far been amazingly good about *not* using my rhetoric as an excuse to bully or belittle others. (I want to single out Scott Alexander in particular here, who is a nicer person than I am and an increasingly amazing writer on these topics, and may deserve part of the credit for making the culture of *Less Wrong* a healthy one.)

To be able to look backwards and say that you’ve “failed” implies that you had goals. So what was it that I was trying to do?

There is a certain valuable way of thinking, which is not yet taught in schools, in this present day. This certain way of thinking is not taught systematically at all. It is just absorbed by people who grow up reading books like *Surely You’re Joking, Mr. Feynman* or who have an unusually great teacher in high school.

Most famously, this certain way of thinking has to do with science, and with the experimental method. The part of science where you go out and look at the universe instead of just making things up. The part where you say “Oops” and give up on a bad theory when the experiments don’t support it.

But this certain way of thinking extends beyond that. It is deeper and more universal than a pair of goggles you put on when you enter a laboratory and take off when you leave. It applies to daily life, though this part is subtler and more difficult. But if you can’t say “Oops” and give up when it looks like something isn’t working, you have no choice but to keep shooting yourself in the foot. You have to keep reloading the shotgun and you have to keep pulling the trigger. You know people like this. And somewhere, someplace in your life you’d rather not think about, you *are* people like this. It would be nice if there was a certain way of thinking that could help us stop doing that.

In spite of how large my mistakes were, those two years of blog posting appeared to help a surprising number of people a surprising amount. It didn’t work reliably, but it worked sometimes.

In modern society so little is taught of the skills of rational belief and decision-making, so little of the mathematics and sciences underlying them . . . that it turns out that just reading through a massive brain-dump full of problems in philosophy and science can, yes, be surprisingly good for you. Walking through all of that, from a dozen different angles, can sometimes convey a glimpse of the central rhythm.

Because it is all, in the end, one thing. I talked about big important distant problems and neglected immediate life, but the laws governing them aren’t actually different. There are huge gaps in which parts I focused on, and I picked all the wrong examples; but it is all in the end one thing. I am proud to look back and say that, even after all the mistakes I made, and all the other times I said “Oops” . . .

Even five years later, it still appears to me that this is better than nothing.

—Eliezer Yudkowsky, February 2015

291

Next:

Biases: An Introduction

6 comments 90 points

Log in to save where you left off

4 comments, sorted by top scoring

Highlighting new comments since Today at 6:53 AM

[-] **Brian Mwangi** 8mo  < 6 >

Great intro.

Can't wait to say "Oops" more often.

"Anyone that doesn't change their mind should consider themselves a fossil"

[-] **DragonGod** 1y  < 5 >

This was very refreshing to read. I'm glad EY has realised that mocking silly ideas doesn't actually help (it makes adherents of the idea double down and be much less likely to listen to you and may also alienate some neutrals. This is particularly true for ideas which have gained currency like the Abrahamic religions). I wasn't able to recommend The Sequences to Christian friends previously because of it's antireligiosity — here's hoping this version would be better.

↑ [-] **Elo** 1y  < 2 > Response to previous version

Welcome! You might like to post on a recent welcome thread. Do look around and do stick around.

[-] **Crell** 2mo  < 3 >