Slate Star Codex, 2016

# 2015 Predictions: Calibration Results

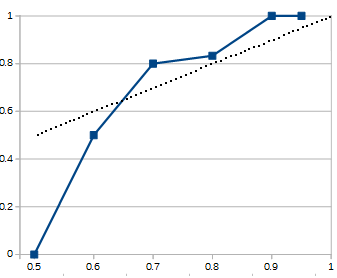
Many people are making predictions for the new year around now. I see fewer people going back and grading the accuracy of their predictions about last year, even though that obviously has a lot of relevance for how seriously we should take this year’s predictions.

In keeping with the SSC tradition, here are the results for my (late) predictions for 2015. Successful predictions are normal-looking, failed predictions are crossed out:

World Events  
1. US will not get involved in any new major war with death toll of > 100 US soldiers: 70%  
2. North Korea’s government will survive the year without large civil war/revolt: 95%  
3. Greece will not announce it’s leaving the Euro: 60%  
3. Neither Russia nor Qatar will lose their World Cups: 80%  
4. Ebola will kill fewer people in second half of 2015 than the in first half: 95%  
5. No terrorist attack in the USA will kill > 100 people: 90%  
6. Assad will remain President of Syria: 70%  
7. Israel will not get in a large-scale war (ie >100 Israeli deaths) with any Arab state: 90%  
8. Syria’s civil war will not end this year: 80%  
9. ISIS will control less territory than it does right now: 70%  
10. ISIS will continue to exist: 80%  
11. Iran will reach a deal with the West on nuclear weapons: 80%  
12. No major civil war in Middle Eastern country not currently experiencing a major civil war: 90%  
13. Iraq’s situation not to get any worse (eg gov’t collapse, new rebellion): 60%  
14. Obamacare will survive the year mostly intact: 60%  
15. Hillary Clinton will be the top-polling Democratic Presidential candidate: 95%  
16. Jeb Bush will be the top-polling Republican candidate: 50%  
17. Trans-Pacific Partnership to pass at least mostly intact: 60%  
18. US official unemployment rate will be less than 7% in Dec 2015: 95%  
19. Bitcoin will end the year higher than $200: 95%  
20. Oil will end the year greater than $60 a barrel: 50%

Personal Life  
21. SSC will remain active: 95%  
22. SSC will get fewer hits in the second half of 2015 than the first half: 60%  
23. At least one SSC post in the second half of 2015 will get > 100,000 hits: 70%  
24. Shireroth will remain active: 90%  
25. I will remain at my same job through the end of 2015: 95%  
26. There will be no further ramifications or lawsuits from either side over the flooding of my house: 80%  
27. I will reach my savings target: 90%  
28. I will get a score at >95th percentile for my year on PRITE: 50% (unknown, haven’t gotten score back)  
29. I will be involved in at least one published/accepted-to-publish research paper by the end of 2015: 60%  
30. I will not break up with any of my current girlfriends: 80%  
31. I will not get any new girlfriends: 50%  
32. I will not finish [project]: 60%  
33. I will attend NYC Solstice ritual: 80%  
34. I will flake out of my plan to lead some kind of Solstice Ritual myself: 60%  
35. I will be living in the house I’m currently trying to arrange to rent: 70%

Scoring  
Of items I marked as 50% confident, 0 were right and 3 were wrong  
Of items I marked as 60% confident, 4 were right and 4 were wrong  
Of items I marked as 70% confident, 4 were right and 1 was wrong  
Of items I marked as 80% confident, 5 were right and 1 was wrong  
Of items I marked as 90% confident, 4 were right and 0 were wrong  
Of items I marked as 95% confident, 7 were right and 0 were wrong



As usual, the dotted line represents perfect calibration; the closer my blue line comes to that, the better I’m doing.

The graph looks like there’s a massive failure at 50%, but this is just an artifact of very few questions at that level. If I’d gotten just one more right, I would be at 33%, ie as close to 50% as it’s possible to get with a set of three. Given that the difference between total success and total failure was just one question, I don’t feel too bad about total failure. Everything else looks pretty good. I’m prepared to call this another successful year.

A side note: Scott Adams has also graded his predictions from the past year, and reports incredible success: 9/9 correct despite going way out on a limb and saying things everyone else found really unlikely (like that Trump would stay Republican front-runner). The obvious way to accomplish like that is to make lots of things that vaguely sound like predictions, then only highlight and count the ones that end up correct; after a quick scan of Adams’ blog, there’s no sign that he’s doing this; his win seems pretty genuine. Another method might be to make vague predictions and grade them in your favor, and there is some sign of this – for example, someone going off Clinton’s poll numbers versus Sanders could say they are in fact still going up. Nevertheless, it’s obviously been a good year for Adams, and I’d be fascinated to see him make a list of official concrete predictions for 2016, all in one place, maybe even associated with confidence levels. [EDIT: Someone else’s more pessimistic analysis; this makes it more pressing that he do everything beforehand this time]

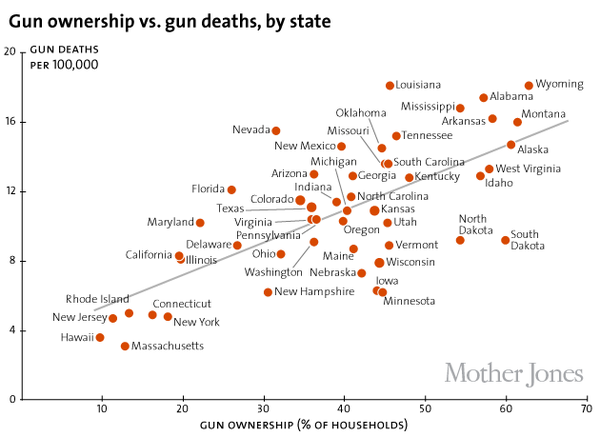
# Guns And States

[Epistemic status: I think I probably wrung the right conclusions out of this evidence, but this isn’t the only line of evidence bearing on the broader gun control issue and all I can say is what it’s consistent with. Content warning for discussion of suicide, murder, and race]

I.

From a Vox article on America’s Gun Problem, Explained: “On Wednesday, it happened again: There was a mass shooting — this time, in San Bernardino, California. And once again on Sunday, President Barack Obama called for measures that make it harder for would-be shooters to buy deadly firearms.”

Then it goes on to say that “more guns mean more gun deaths, period. The research on this is overwhelmingly clear. No matter how you look at the data, more guns mean more gun deaths.” It cites the following chart:

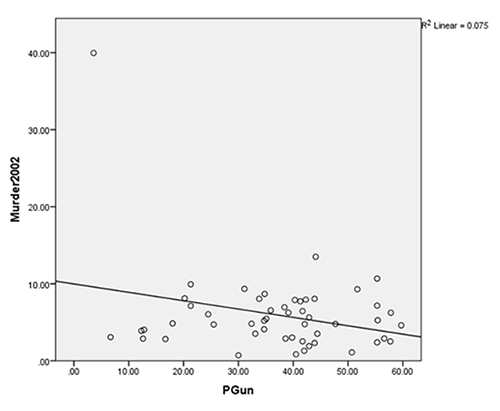


…then uses the graph as a lead in to talk about active shooter situations, gun-homicide relationships, and outrage over gun massacres.

Did you notice that the axis of this graph says “gun deaths”, and that this is a totally different thing from gun murders?

(this isn’t an isolated incident: Vox does the same thing here and here)

Gun deaths are a combined measure of gun homicides and gun suicides. Here is a graph of guns vs. gun homicides:



And here is a graph of guns vs. gun suicides:



The relationship between gun ownership and homicide is weak (and appears negative), the relationship between gun ownership and suicide is strong and positive. The entire effect Vox highlights in their graph is due to gun suicides, but they are using it to imply conclusions about gun homicides. This is why you shouldn’t make a category combining two unlike things.

II.

I am not the first person to notice this. The Washington Examiner makes the same criticism of Vox’s statistics that I do. And Robert VerBruggen of National Review does the same analysis decomposing gun deaths into suicides and homicides, and like me finds no correlation with homicides.

German Lopez of Vox responds here. He argues that VerBruggen can’t just do a raw uncontrolled correlation of state gun ownership with state murder rates without adjusting for confounders. This is true, although given that Vox has done this time and time again for months on end and all VerBruggen is doing is correctly pointing out a flaw in their methods, it feels kind of like an isolated demand for rigor.

So let’s look at the more-carefully-controlled studies. Lopez suggests the ones at the Harvard Injury Control Research Center, which has done several statistical analyses of gun violence. They list two such analyses comparing gun ownership versus homicide rates across US states: Miller Azrael & Hemenway (2002), and Miller Azrael & Hemenway (2007).

(does it count as nominative determinism when someone named Azrael goes into homicide research?)

We start with MA&H 2002. This study does indeed conclude that higher gun ownership rates are correlated with higher murder rates after adjusting for confounders. But suspiciously, it in fact finds that higher gun ownership rates are correlated with higher murder rates even before adjusting for confounders, something that we already found wasn’t true! Furthermore, even after adjusting for confounders it finds in several age categories that higher gun ownership rates are correlated with higher non-gun homicide rates (eg the rates at which people are murdered by knives or crowbars or whatever) at p less than 0.001. This is really suspicious! Unless guns are exerting some kind of malign pro-murder influence that makes people commit more knife murders, some sort of confounding influence has remained. Let’s look closer.

The study gets its murder rate numbers from the National Center for Health Statistics, which seems like a trustworthy source. It gets its gun ownership numbers from…oh, that’s interesting, it doesn’t actually have any gun ownership numbers. It says that there is no way to figure out what percent of people in a given state own guns, so as a proxy for gun ownership numbers, it will use a measure called FS/S, ie the number of firearm suicides in a state divided by the total number of suicides.

This makes some intuitive sense. Among people who want to commit suicide, suppose a fixed percent prefer to use guns compared to other methods. In that case, the determining factor for whether or not they use a gun will be whether or not they have a gun. Hospitals diligently record statistics about suicide victims including method of suicide, so if our assumption holds this should be a decent proxy for gun ownership within a state.

There’s only one problem – I checked this against an actual measure of gun ownership per state that came out after this study was published – the CDC asking 200,000 people how many guns they had as part of the Behavioral Risk Factor Surveillance System Survey – and the FS/S measure fails. When I repeat all of their analyses with their own FS/S measure, I get all of their same positive correlations, including the ones with non-gun homicides. When I repeat it with the real gun ownership data, all of these positive correlations disappear. When I look at exactly why this happens, it’s because FS/S is much more biased towards Southern states than actual gun ownership is. Real gun ownership correlates very modestly – 0.25 – with 538’s ranking of the Southern-ness of states. FS/S correlates at a fantastically high 0.62. For some reason, suicidal Southerners are much more likely to kill themselves with guns than suicidal people from the rest of the States, even when you control for whether they have a gun or not. That means that MA&H 2002 thought it was measuring gun ownership, but was actually measuring Southern-ness. This is why they found higher homicide rates, including higher rates of non-gun homicide.

So we move on to MA&H 2007. This study was published after the CDC’s risk survey, so they have access to the same superior gun ownership numbers I used to pick apart their last study. They also have wised up to the fact that Southern-ness is important, and they include a dummy variable for it in their calculations. They also control for non-gun crime rate, Gini coefficient, income, and alcohol use. They do not control for urbanization level or race, but when I re-analyze their data including these factors doesn’t change anything, likely because they are already baked in to the crime rate.

They find that even after controlling for all of this stuff, there is still a significant correlation between gun ownership level and gun homicide rate. Further, this time they are using good statistics, and there is not a significant correlation between gun ownership and non-gun-homicide rate. Further, there is a correlation between gun ownership and total homicide rate, suggesting that the gun-gun-homicide correlation was not just an artifact of people switching from inferior weapons to guns while still committing the same number of murders. Further, this is robust to a lot of different decisions about what to control or not to control, and what to include or not to include.

I repeated all of their analyses using two different sources of gun ownership data, a couple different sources of homicide and crime rate data, and a bunch of different plausible and implausible confounders – thanks a lot to Tumblr user su3su2u1 for walking me through some of the harder analyses. I was able to replicate their results. Pro-gun researcher John Lott had many complaints about this study, including that it was insensitive to including DC and that it was based entirely on the questionable choice of controlling for robbery rate – but I was unable to replicate his concerns and found that the guns-homicide correlation remained even after DC was included and even when I chose a group of confounders not including robbery rate. I was unable to use their methodology to replicate the effect in places where it shouldn’t replicate (I tried to convince it to tell me tractors caused homicide, since I was suspicious that it was just picking up an urban/rural thing, but it very appropriately refused to fall for it). Overall I am about as sure of this study as I have ever been of any social science study, ie somewhat.

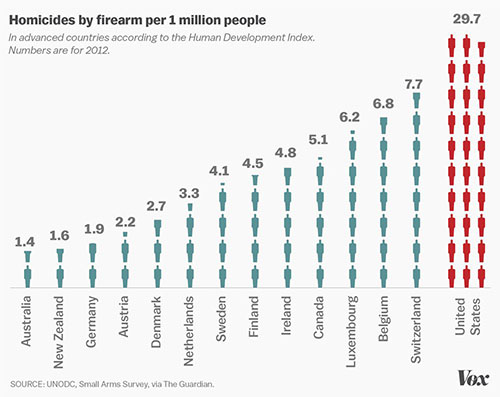
This study doesn’t prove causation; while one interpretation is that guns cause homicide, another is that homicide causes guns – for example, by making people feel unsafe so they buy guns to protect themselves. However, I doubt the reverse causation aspect in this case. The study controlled for robbery rate; ie it was looking at whether guns predicted homicides above and beyond those that could be expected given the level of non-homicide crime. My guess is that people feeling unsafe is based more on the general crime rate than on the homicide rate per se, which would make it hard for the homicide rate to cause increased gun ownership independently of the crime rate.

If guns are in fact correlated with more homicide, how come me and VerBruggen found the opposite in our simpler scatterplot analysis? This is complicated, but I think the biggest part of the answer is the urban/rural divide. Rural people have more guns. Murder rates are higher in urban areas. Race also plays a part: whites have more guns, but black areas have higher murder rates. Finally, the North and West seem to have more guns, but murder rates are highest in the South (which is what produced the bogus effect on the last study). All of these differences are large enough to cancel out the gun/no-gun difference and make the raw scatterplot look like nothing. This study didn’t address all those things directly, but its decision to control for non-gun crime rate and poverty took care of them nevertheless. As the old saying goes, guns don’t kill people; guns controlled for robbery rate, alcoholism, income, a dummy variable for Southernness, and a combined measure of social deprivation kill people.

If this is all true, how come I spent so much time yelling at that first study with worse data? Because I worry that if people only see the good studies, they’ll get complacent. Vox posted these two studies as proof that there was a state-level gun-murder correlation. The first one was deeply flawed, but the second one turned out to be okay. Do you think Vox realized this? Do you think they would have written that article any differently in a world where both studies were flawed? As long as you trust every scientific paper you see – let alone every scientific paper you see on your side in a highly politicized field – even when you’re right it will often just be by luck.

III.

Vox also voxsplains to us about America’s unusually high gun homicide rate.



Having presented this graph, they say that “To understand why that is, there’s another important statistic: The US has by far the highest number of privately owned guns in the world.”

Even granting, as we saw above, that gun ownership does indeed increase homicide rates, this is not the most important factor in explaining America’s higher homicide rate, or even close to the most important factor. Let me give a few arguments for why this must be the case:

1. The United States’ homicide rate of 3.8 is clearly higher than that of eg France (1.0), Germany (0.8), Australia (1.1), or Canada (1.4). However, as per the FBI, only 11,208 of our 16,121 murders were committed with firearms, eg 69%. By my calculations, that means our nonfirearm murder rate is 1.2. In other words, our non-firearm homicide rate alone is higher than France, Germany, and Australia’s total homicide rate. Nor does this mean that if we banned all guns we would go down to 1.2 – there is likely a substitution effect where some murderers are intent on murdering and would prefer to use convenient firearms but will switch to other methods if they have to. 1.2 should be considered an absolute lower bound. And it is still higher than the countries we want to compare ourselves to.

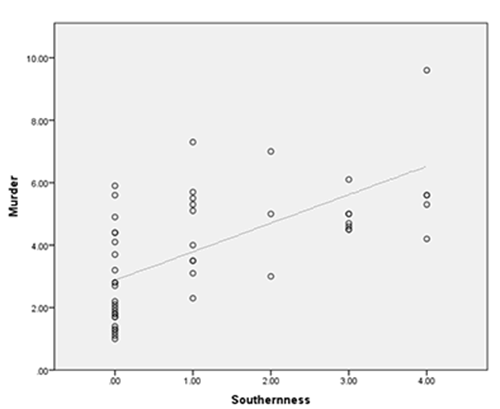
2. There are many US states that combine very high firearm ownership with very low murder rates. The highest gun-ownership state in the nation is Wyoming, where 59.7% of households have a gun (really!). But Wyoming has a murder rate of only 1.4 – the same as right across the border in more gun-controlled Canada, and only about a third of that of the nation as a whole. It seems likely that the same factors giving Canada a low murder rate give Wyoming a low murder rate, and that the factors differentiating the rest of America from Wyoming are the same factors that differentiate the rest of America from Canada (and Germany, and France…). But this does not include lower gun ownership.

3. There are many US states that combine very low firearm ownership with very high murder rates. The highest murder rate in the country is that of Washington, DC, which has a murder rate of 21.8, more than twenty times that of most European countries. But DC also has the strictest gun bans and the lowest gun ownership rate in the country, with gun ownership numbers less than in many European states! It seems likely that the factors making DC so deadly are part of the story of why America as a whole is so deadly, but these cannot include high gun ownership.

If not gun ownership, what is the factor making America so much more deadly than Europe and other First World countries? The traditional answer I always heard to this question was that America had a “culture of violence”. I always hated this answer, because it seemed so vague and meaningless as to be untestable by design. If the NRA waves their hands and says “eh, culture of violence”, how are you going to tell them they’re wrong?

But we can work with this if we assume the culture of violence (or, if you want to be official about it, “honor culture”) is more common in some populations and areas than others. Some of the groups most frequently talked about during these lines are Southerners and various nonwhite minorities. This provides a testable theory: if we compare American non-Southern whites to European countries mostly made up of non-Southern whites, we’ll find similar murder rates. But first, some scatter plots:

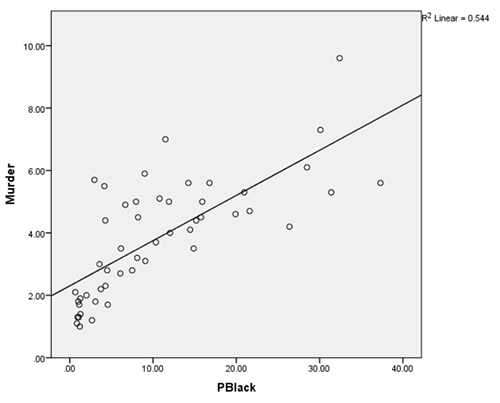
This is murder rate by state, correlated with perceived Southernness of that state as per 538’s poll. I’ve removed DC as an outlier on all of the following.



Who'd have thought populating half a country with the descendants of a group of people called "Border Reavers" would cause so much trouble?

— Scott Alexander (@slatestarcodex) December 3, 2015

And this is murder rate by state correlated with percent black population:

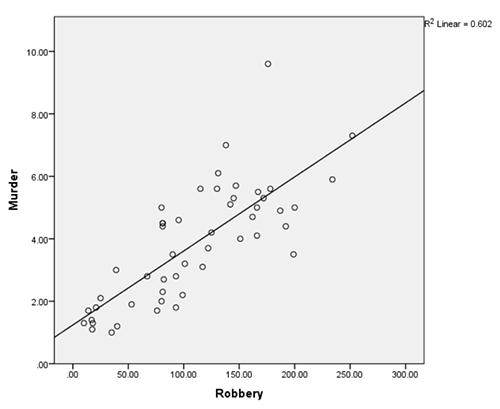


This would seem to support the “culture of violence” theory.

Can we adjust for this and see what the murder rate is for non-Southern whites? Sort of. The Economist gives a white-only murder rate of 2.5 (this is based on white victims, whereas we probably want white perpetrators, but the vast majority of murders are within-race so it doesn’t make much difference). And Audacious Epigone has put together a collection of white murder rates by state. I can’t find anything on non-Southern white murder rates per se, but one hack would be to take the white murder rate in non-Southern states and assume there aren’t any Southerners there.

Our main confounder will be urbanization. Western Europe is about 80% urban, so let’s look at states at a similar level. The four northern states that are closest to 80% urban are Colorado, Oregon, Washington, and Connecticut. I’m throwing out Colorado because it has a large Latino population who can’t be statistically differentiated from whites. That leaves, Washington (2.4), Connecticut (2.0), and Oregon (2.0). So possibly adjusting out Southerners brings us down from 2.5 (all whites) to 2.1 or so (non-Southern whites)? Again, compare to Germany at 0.8, Canada at 1.4, and America at 3.8.

There’s one more factor that needs to be considered:



This is a plot of the gun death rate vs. the robbery rate. There’s a strong correlation (r = 0.78). Robbery is heavily correlated with percent black, percent Southern, and urbanization, so it’s probably coming from the same place. Nevertheless, it seems to correlate with murder better than any of them alone, maybe because it’s combining all three measures together. I was able to make a linear model using those three measures that correlated at r = 0.79 with murder, about the same amount that robbery does. I should also mention that robbery correlates negatively with gun ownership at r = – 0.52, but this disappeared when controlled for urbanization.

So my very tentative conclusion is that although the US murder rate is much higher than that of other First World countries, this is partly due to the existence of various cultural factors not present in those other nations. When we adjust those away, America’s murder rate falls from 3.8 to 2.1. Which is still higher than Germany’s 0.8 or Canada’s 1.4.

Is that extra due to guns?

IV.

According to MA&H 2007, each absolute percentage point in gun ownership was related to a 2.2 relative percentage point difference in homicide. This part of the study was beyond my ability to check, and I’m not sure why they switched from absolute to relative percents there, but suppose we take it seriously.

America has a gun ownership rate of 32%, so if we somehow decreased that to zero, we would naively expect about a 70% decrease in homicides. Unfortunately, only 67% of American homicides involve guns, so we’re back to pretending that eliminating guns will not only have zero substitution effect but also magically prevent non-gun homicides. This shows the dangers of extrapolating a figure determined by small local differences all the way to the edge of the graph (I’M TALKING TO YOU, RAY KURZWEIL).

Maybe we can be more modest? Canada has a gun ownership rate of aboot 26%, so…

…wait a second. I thought we’ve been told that the US has a gun ownership rate seven zillion times that of any other country in the world, and that is why we are so completely unique in our level of gun crime? And now they’re telling us that Canada has 26% compared to our 32%? What?

Don’t trust me too much here, because I’ve never seen anyone else analyze this and it seems like the sort of thing there should be loads of analyses of if it’s true, but I think the difference is between percent of households with guns vs. guns per capita. US and Canada don’t differ very much in percent of households with guns, but America has about four times as many guns per capita. Why? I have no idea, but the obvious implication is that Canadians mostly stop at one gun, whereas Americans with guns buy lots and lots of them. In retrospect this makes sense; I am looking at gun enthusiast bulletin boards, and they’re advising other gun enthusiasts that six guns is really the bare minimum it’s possible to get by with (see also “How many guns can you have before it’s okay to call your collection an ‘arsenal’?”, which I have to admit is not a question that I as a boring coastal liberal have ever considered). So if the guy asking that question decides he needs 100 guns before he gets his arsenal merit badge, that’s a lot more guns per capita without increasing percent household gun ownership. This should actually be another argument that guns are not a major factor in differentiating US vs. Canadian murder rates, since unless you’re going on a mass shooting (WHICH IS REALLY RARE) you wouldn’t expect more murders from any gun in a household beyond the first. That means that the small difference between US and Canadian household percent gun ownership rates (32% vs. 26%) would have to drive the large difference between US and Canadian murder rates (1.4 vs. 3.8), which just isn’t believable.

…okay, sorry, where were we? Canada has a gun ownership rate of about 26%, so if America were to get its gun ownership as low as Canada, that would be -6 absolute percentage points = a 13% relative decrease in murder rate = the murder rate going from 3.8 to 3.3 = a 0.5 point decrease in the murder rate. That’s pretty close to the difference between our 2.1 US-sans-culture-of-violence estimate and the 1.4 Canadian rate – so maybe beyond the cultures of violence, the rest of the US/Canada difference really is due to guns?

(I’m not sure whether I should be subtracting 13% from 2.1 rather than 3.8 here)

In Germany, 9% of households own firearms (wait, really? European gun control is less strict than I thought!) Using MA&H’s equation, we predict that if the US had the same gun ownership rate as Germany, its murder rate would drop 50%, eg from 3.8 to 1.9. Adjust out the culture of violence, and we’re actually pretty close to real Germany’s murder rate of 0.8.

How much would gun control actually cut US gun ownership? That obviously depends on the gun control, but a lot of people talk about Australia’s gun buyback program as a model to be emulated. These people say it decreased gun ownership from 7% of people to 5% of people (why is this number so much lower than Canada and Germany? I think because it’s people rather than households – if a gun owner is married to a non-gun-owner, they count as one gun-owner and one non-owner, as opposed to a single gun-owning household. The Australian household number seems to be 19% or so). So the gun buyback program in Australia decreased gun ownership by (relative) 30% or so. If a similar program decreased gun ownership in America by (relative) 30%, it would decrease it by (absolute) 10% and decrease the homicide rate by (absolute) 22%. Since there are about 13000 homicides in the US per year, that would save about 3000 lives – or avert about one 9/11 worth of deaths per year.

(note that our murder rate would still be 3.0, compared to Germany’s 0.8 and Canada’s 1.4. Seriously, I’m telling you, the murder rate difference is not primarily driven by guns!)

Is that worth it? That obviously depends on how much you like being able to have guns. But let me try to put this number into perspective in a couple of different ways:

Last time anyone checked, which was 1995, about 618,000 people died young (ie before age 65) in the US per year. Suppose that the vast majority of homicides are of people below 65. That means that instituting gun control would decrease the number of premature deaths to about 615,000 – in other words, by about half a percentage point. I’m having to borrow this data from the UK, but if it carries over, the average person my age (early 30s) has a 1/1850 chance of death each year. Gun control would decrease that to about 1/1860. I’m very very unsure about the exact numbers, but it seems like the magnitude is very low.

On the other hand, lives are very valuable. In fact, the statistical value of a human life in the First World – ie the value that groups use to decide whether various life-saving interventions are worth it or not – is $7.4 million. That means that gun control would “save” $22 billion dollars a year. Americans buy about 20 million guns per year (really)! If we were to tax guns to cover the “externality” of gun homicides preventable by Australia-level gun control, we would have to slap a $1000 tax on each gun sold. While I have no doubt that some people, probably including our arsenal collector above, would be willing to pay that, my guess is that most people would not. This suggests that most people probably do not enjoy guns enough to justify keeping them around despite their costs.

Or if all gun enthusiasts wanted to band together for some grand Coasian bargain to buy off the potential victims of gun violence, each would have to contribute $220/year to the group effort – not totally impossible, but also not something I can really see happening.

This is very, very, very, very very tentative, but based on this line of reasoning alone, without looking into the experimental studies or anything else, it appears that Australia-style gun control would probably be worth it, if it were possible.

(I didn’t price in the advantages of guns in terms of preventing state tyranny and protecing freedom, which might be worth subsidizing, but my guess is that if 32% gun ownership is enough to maintain freedom, 22% gun ownership is as well)

V.

In summary, with my personal confidence levels:

1. Scatterplots showing raw correlations between gun ownership and “gun deaths” are entirely driven by suicide, and therefore dishonest to use to prove that guns cause murder (~100% confidence)

2. But if you adjust for all relevant confounders, there is a positive correlation between gun ownership and homicide rates (~90% confidence). This relationship is likely causal (~66% confidence).

3. The majority of the difference between America’s murder rate and that of other First World countries is not because of easier access to guns in America (~90% confidence).

4. But some of it is due to easier access to guns. This is probably about 0.5 murders/100K/year.

5. An Australian-style gun control program that worked and had no side effects would probably prevent about 2,000 murders in the US. It would also prevent a much larger number of suicides. I am otherwise ignoring suicides in this piece because discussing them would make me too angry.

6. Probably the amount of lost gun-related enjoyment an Australian-style gun control program would cause do not outweigh the benefits.

7. This is not really enough analysis to make me have a strong opinion about gun control, since this just looks at the correlational evidence and doesn’t really investigate the experimental evidence. Contrary to what everyone always tells you, experimental evidence doesn’t always trump correlational – there are cases where each has its strengths – but it wouldn’t be responsible to have a real opinion on this until I look into that too. Nevertheless, these data are at least highly consistent with Australia-style gun control being a good idea for the US.

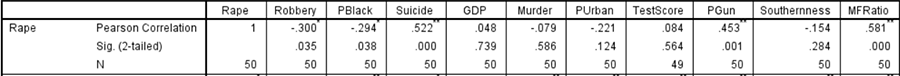
If you want to look into this more, here is a CSV version of all the relevant data.

# Guns and States 2: Son of a Gun

Thanks for your many excellent comments on Guns and States. I’ve made it through most of the comments thread and learned a lot. Here are some (hopefully) final thoughts on the matter:

1. In yesterday’s post, I suggested that the difference in homicide rates between America and other First World countries were about two-thirds cultural, one-third gun-related. That’s sort of true, but people have reminded me to think of it as an interaction. Without the cultural factors in place, guns are pretty harmless. That’s why Wyoming can be the highest-gun-ownership state in the country, with 60% of households having a weapon, and still have a murder rate equal to Canada’s and lower than 45 other states’. With the cultural factors in place, guns make a bad situation worse. This is why the robbery rate and the gun homicide rate correlate at about 0.9 (0.75 if DC is removed as an outlier) even though most robberies do not involve firearms.

2. But some robberies do involve firearms, making them an imperfect control. I looked at rape, a crime that sounds like it should be associated with other violent crimes but which almost never involves firearms. Unfortunately, the results were really weird:



As you can see, rape is negatively correlated with robbery and %black, insignificantly correlated with homicide and urbanness, and positively correlated with gun ownership and suicide rate (which itself is heavily correlated with gun ownership). I also added one more variable in an attempt to explain some of this: MFRatio, which is the number of men per 100 women in the state. That certainly mattered a lot, but doesn’t take away from the confusingness of the other variables. My first theory was that rape is more common in Red Tribe rural white culture, which would explain the guns and %black but really doesn’t fit with the PUrban or Southernness issues, nor the fact that the correlation with guns was much higher than the racial correlation. My second was that people bought guns to defend against rape, but that doesn’t explain the negative correlation with robbery, and surely people would buy guns based on crime in general (and since it’s mostly men doing the gun-buying, rape would be the least likely to affect the gun rate). Overall I admit I am confused.

3. Commenter eccdogg can’t replicate the Southern culture of violence thing after controlling for other factors. I acknowledge that he seems to be doing the right statistics, but can’t square that with commonsense eyeballing of the data.

4. I did a reanalysis that found that (after adjusting for confounders) the gun data from 2002 successfully predicted the murder rate in 2002, but did not successfully predict the murder rate today, even though most other relevant variables did. I’m going to attribute this to gun ownership patterns changing faster than (say) whether a state is Southern or not, and probably if I had gun ownership data from today (which I don’t) it would work, but it seemed important to mention that and get it out there.

5. I was shocked to see that between 20 to 30 percent of people in most European/Anglosphere countries owned guns, including the ones like Canada that gun control advocates hold up as an example of what they want. That makes it very strange the degree to which people expect gun control to mean “the government confiscates all your guns, after which no guns are left”. I mean, I understand why (for example) the NRA would promote that story in order to get people angry, but it’s very strange how often liberals nod along and say “Yup, that’s what we want to do!”. I wish more of the debate could be about waiting periods, required training classes, and background checks – none of which would prevent people from using guns to defend against a potentially tyrannical government. (I don’t know enough about guns to be know whether restricting handguns would make it harder to defend against tyranny. Could you shoot the tyrants equally well with a rifle?)

6. Commenter Elias brings up this meta-analysis by Gary Kleck claiming that the 40 guns-and-states style studies he could find were split almost exactly half and half in terms of whether they found a significant guns-homicide correlation or not. He further claimed that the better the study, the less likely it was to find a significant correlation. The study I cited, plus my own analyses, passed two of Kleck’s tests – used good gun ownership numbers and controlled for confounders – but failed the third, which was distinguishing forward causation from reverse causation. I admitted this was a problem and that it would take very powerful methods beyond my abilities to solve, but I also said that I suspected forward causation since the homicide causation remained independent of robbery and I would expect people to buy guns based on a general crime rate. I stick to that suspicion. Kleck has some studies that he thinks establish reverse causation instead, and I encourage you to read them.

This reminds me a lot of my post Beware The Man Of One Study – I can at least say in my own defense that I did the analysis on my own, which should obviate exactly which studies I did or didn’t look at, but the media articles that just presented MA&H and said “Look! Here’s what Science says, discussion over!” really should have mentioned that there were also nineteen studies, 48% of the total, that disagreed with them.

On the other hand, Kleck said that only three studies met his quality standards, and all of them were his own, so take that with however much salt you feel like it deserves.

# Slow But Steady

I saw this on Giving What We Can and was so delighted I wanted to repost it here. Source is The effect of malaria control on plasmodium falciparum in Africa between 2000 and 2015 in Nature:

GWWC comments:

The reduction in malaria has been caused — in large part — by people sleeping under bednets…Long lasting insecticide-treated bednets are a powerful weapon against malaria, not only because they’re a physical barrier between mosquitoes and sleeping children — the insecticide coating kills mosquitoes, so they don’t infect other members of the family (and the village) who don’t have mosquito nets…

In other words, this paper tells us about the bigger picture, showing that bednets are incredibly effective, not just at the level of individual villages, but at the level of whole populations. Essentially, the case that we should be distributing bednets just got even stronger.

The research suggests that anti-malarial interventions have prevented about 663 million malarial fevers. Long-lasting insecticide treated bed nets stand out as being particularly effective — being responsible for around 68% of the malaria reduction. This means that bednets have prevented around 450 million cases of malaria! And globally, 6.2 million fewer people died of malaria over the last 15 years because of malaria interventions…

Malaria is an immense economic burden on health systems and people: since 2000, malaria treatment in sub-Saharan Africa has cost almost $300 million. We can see the flip side of this in places in the United States, Brazil and Uganda — children born after malaria was eradicated or heavily controlled and so were not exposed to malaria during pregnancy, had substantially higher income later in life.

We urgently need to keep funding the distribution of bednets, because bednet distributions are one of the most cost-effective ways of preventing disease and death. Since 2000, one billion bednets have been distributed (costing around 5 dollars each), and have averted 450 million cases of malaria — this suggests that, on average, one episode of clinical malaria can be prevented for about $11 (malarial fevers can be very painful). One recent study suggests that in Kenya, bednet distributions between 2003 and 2008 have prevented a death of a child for about $1,011 on average

Humanity seems to be very visibly winning the war against malaria. I just donated a thousand more bednets; now I feel like a part of it and one day I can tell my kids that I helped. If you’re interested, you can donate to Against Malaria Foundation here.

# Schizophrenia: No Smoking Gun

[Note: despite how some people are spinning this, tobacco is still really really bad and you should not smoke it]

I.

Schizophrenics smoke. A lot. Depending on the study, about 60-80% of schizophrenics smoke, compared to only about 20% of the general population. And they spend on average about 27% (!) of their income on cigarettes. Even allowing that schizophrenics don’t make much income, that’s a lot of money. Sure, schizophrenics are often poor and undereducated and have other risk factors for smoking – but even after you control for this, the effect is still pretty strong.

Various people have come up with various explanations. Cognitively-minded people say that schizophrenics smoke as a maladaptive coping strategy for the anxiety caused by their condition. Pharmacologically-minded people say that schizophrenics smoke because smoking accelerates the metabolism of antipsychotic drugs and so makes their side effects go away faster. Pragmatically-minded people say that schizophrenics smoke because they’re stuck in institutions with nothing to do all day. No points for guessing what the Freudians say.

But all these theories have problems. Sure, schizophrenics are often institutionalized, but even the ones at home smoke a lot. Sure, some schizophrenics are often on antipsychotics, but even the ones who aren’t on meds smoke a lot. Sure, schizophrenics are anxious, but we don’t see people with Generalized Anxiety Disorder having 80% smoking rates.

As usual, I’m more biologically-minded, so I find it interesting that some of the genes that most commonly turn up as linked to schizophrenia – especially CHRNA3, CHRNA5, and CHRNA7 – are in nicotine receptors. Indeed, some of them are also the genes identified as risk factors for smoking. Further, there’s a lot of evidence that schizophrenic people actually feel better and have fewer symptoms when they’re smoking. Further, schizophrenics tend to gravitate toward cigarettes with higher nicotine content, and smoke them in ways that maximize nicotine absorption.

It seems like part of the problem with schizophrenia is that the brain’s nicotine system isn’t working well. Smoking supplements nicotine and makes the system run smoother, so schizophrenics feel better when they smoke and continue to do so. This is the widely accepted self-medication hypothesis.

I like this because it’s a really elegant example of…I don’t know what you’d call it…memetic evolution? Nobody knew that nicotine helped schizophrenia, nobody told the schizophrenics that, but they sort of naturally gravitated to an effective treatment for their condition by going in the direction of things that make them feel better, even going so far as to unknowingly gravitate toward cigarette brands with more nicotine. They did all of this before psychiatry had any idea why they were doing it, and in the face of constant protests that it was stupid and useless. This should be a warning to anyone who’s too quick to tell patients that their coping strategies are maladaptive.

But there’s a much more important question here: does smoking cause schizophrenia? How about prevent it?

II.

First, the causation argument. Gurillo et al do a meta-analysis and conclude that “daily tobacco use is associated with increased risk of psychosis and an earlier age of onset of psychotic illness. The possibility of a causal link between tobacco use and psychosis merits further examination”. That is, schizophrenics are already smoking much more at the moment their schizophrenia starts. This suggests that maybe smoking is helping to cause the schizophrenia?

All nice and well, except for a few things. First, this study ignores the possibility that the genes that cause schizophrenia might also cause increased smoking, even though we have some evidence that this is true (actually, it doesn’t ignore this, it mentions it, but uses it as a reason why a schizophrenia-smoking link is more plausible). Second, we know that people who will later develop schizophrenia are seen as kind of odd even before they come down with the disease, and it’s possible that they’re already in some unusual brain state that smoking helps relieve. Third, this study is not controlled – meaning that we’re totally helpless before factors like “people destined to later develop schizophrenia are often poor, and poor people smoke more”.

And fourth, another study shows exactly the opposite.

Zammit et al (thanks to @allfeelsallthetime for the tip) looks at 50,000 teenage Swedish conscripts, then follows them throughout their lives to see which ones do or don’t get schizophrenia. They find that without adjusting for confounders, smokers are more likely to get schizophrenia. But when you do adjust for confounders, smokers are less likely to get schizophrenia, (hazard ratio 0.8, p = 0.003) and heavy smokers are much less likely to get schizophrenia (hazard ratio 0.5)! A dose-dependent relationship was found between smoking and protection from schizophrenia. This is really interesting.

Why do we find such different results from these two studies? The only explanation I can think of is that the second study controls for various factors including cannabis use, personality variables, IQ, past psychiatric diagnoses, and place of upbringing (thanks @su3su2u1 for the tip) and the first study controls for zilch. In fact, we find that the second study’s uncontrolled numbers are not that different from the first study’s uncontrolled numbers, and that the only difference is that the second study then went on to control for confounders and get the opposite result. Controlling for more things is not always better, but controlling for a few things that previous studies and common sense suggest are very relevant is pretty superior to just leaving the data entirely unprocessed. Advantage very much second study.

III.

Unlike certain people on Facebook, I fucking hate science. Let me explain why.

The first study here, Gurillo et al, was published ten years after the second study. Since it is a meta-analysis, it included the second study in it. The authors of the first study definitely read the second study. They just didn’t care.

Nowhere in the first study does it say “By the way, we read this other study that got the opposite results from us, let’s try to figure out why, oh, it was because they controlled for things and we didn’t, maybe that should call our findings into question.” You know what they did do? They listed the second study as finding that smoking increased schizophrenia risk, because the rules of their meta-analysis said they would only take uncontrolled data, and so they did. You can read this entire study, which cites the second study no fewer than six times, without hearing at all about the fact that the second study got the opposite result using likely better methodology.

Then they go on to conclude that:

Cigarette smoking might be a hitherto neglected modifiable risk factor for psychosis, but confounding and reverse causality are possible. Notwithstanding, in view of the clear benefits of smoking cessation programs in this population, every effort should be made to implement change in smoking habits in this group of patients.

Clear benefits! Every effort! Aaaaaaah!

I mean, I know where they (and the Lancet editors, who write a glowing comment backing them up) are coming from. Smoking is bad because lung cancer, COPD, etc. But now we have these things called e-cigarettes! They deliver nicotine without tobacco! As far as anyone knows they carry vastly less risk of cancer, COPD, etc. If nicotine actually prevents schizophrenia rather than causing it, that is the sort of thing we should really want to know. And instead we’re just getting this “We should make schizophrenia patients stop smoking, because smoking is bad”.

Look. I am not going to come out and say that there’s great evidence that nicotine decreases schizophrenia risk. There’s one study, which other studies contradict. I happen to think that the one study looks better than its competitors, but that’s my opinion and I have nowhere near the evidence I would need to feel really strongly about this. But I feel like we are very far from the point where we know enough to be pushing people at risk of schizophrenia away from nicotine, and light-years away from the point where we can use phrases like “clear benefits”.

Possibly I am an idiot and missing something very important. But if this is true, I wish the authors of the new study, and the editors of The Lancet, would have acknowledged the existence of the conflicting study and patiently explained to their readership, many of whom are idiots like myself, “Here’s a study that looks better than ours that seems to contradict our results, but here’s why our study is nevertheless far more believable.” That’s all I ask.

No matter how much of an idiot I am, I can’t possibly imagine how that wouldn’t be a straight-out gain.

PS: Cigarette smoking definitely decreases your risk of Parkinson’s Disease. Parkinson’s is similar to schizophrenia in that both involve dopamine. But schizophrenia involves too much dopamine and Parkinson’s too little, so the analogy could go either direction.

PPS: Tobacco smoking is definitely still bad! Nothing in here at all suggests that tobacco smoking has the slightest chance of not being a terrible decision!

# Lies, Damned Lies, And The Media (Part 6 of ∞)

[content warning: discussion of violent crime, including sexual assault]

I plan to write an article on misuse of statistics by online news organizations, but looking back through my archives most of the examples I’ve got are from a couple of liberal sites that aren’t the worst offenders so much as the only ones I can even bear to read. I’m worried that some of my readers have gotten the impression that liberal sites are the only ones that routinely misuse statistics, which would be grossly false. So before I write the article, I thought I’d give one example of how a lot of conservative sites have statistics that are so bad they’re not even fun to dissect.

I chose Breitbart’s “Rape Deniers: 9 Facts About Illegal Alien Crime The Media Covers Up” because it sounded promising – talking about rape and calling people who disagree with you “deniers” are two pretty reliable red flags that an article will be terrible. It’s a series of 9 facts meant to show that illegal immigrants to the US are involved in a lot of crime, especially sexual crime.

I’m skipping Fact 1, which is just a methodological point I don’t dispute, Fact 2, which just says some native-born Americans are unemployed, and Fact 3, which says that a lot of our heroin comes from Latin America; I don’t really disagree with any of these. I’m also skipping Fact 7 because it’s a repeat of Fact 6 and the same points apply. That leaves Facts 4, 5, 6, 8, and 9.

4. Because local and state prisons don’t track legal status, we don’t know how many illegals are in those prisons. As my colleague Ben Shapiro points out, the lack of this number is being used dishonestly by the media against Trump…Trump’s repeated statements about immigrants and crime underscore a common public perception that crime is correlated with immigration, especially illegal immigration. But that is a misperception; no solid data support it, and the data that do exist negate it. Trump can defend himself all he wants, but the facts just are not there. Except the facts are there. The Feds do track legal status, and the numbers are startling. Of 78,022 primary offense cases in fiscal year 2013, 38.6 percent were illegal immigrant offenders. The majority of their cases (76 percent) were immigration related. Of total primary offenses, 17.6 percent of drug trafficking offenses and 3.8 percent of sex abuse were illegal immigrants. Of 22,878 drug crime cases, 17.2 percent were illegal immigrants.

The first part of this is a really weird complaint. They’re saying that 38.6% of federal prisoners are illegal immigrants, which is true and indeed very high. Then they’re admitting that 76% of their cases are immigration related. That is, the Feds are imprisoning them because they immigrated illegally. I think that most people would be willing to concede that illegal immigrants are more likely to have immigrated illegally than other populations.

The second part, the part about federal drug trafficking, is complicated; it’s different from “drug having”, “drug dealing”, and even “drug trafficking” as a broader non-federal category. To get a federal drug trafficking arrest, you have to move really large quantities of drugs “across state or national borders”, preferably in a “High Intensity Drug Trafficking Area”. That is, crimes that would give you state charges in a normal place become federal charges in one of these areas. Where are they? The entire US border with Mexico is a gigantic High Intensity Drug Trafficking Area (see this map). So a high rate of illegal immigrants among federal drug trafficking prisoners just means that they’re more likely to be involving in transporting drugs across the US-Mexico border than, say, a lumberjack in Wisconsin is. I am prepared to believe this.

Finally, illegal immigrants do commit 3.8% of federal sexual abuse cases. I give Breitbart credit for finally getting a number that is entirely correct and not biased at all. Unfortunately for them, illegal immigrants are 3.8% of the US population.

5. According to the Justice Department…There are 94 federal court districts in this country and the five located near the southern border see a large portion of criminal cases, according to the Justice Department’s annual report on criminal prosecutions. The five federal districts also have the biggest number of defendants actually convicted of federal crimes. Of the 61,529 criminal cases initiated by federal prosecutors last fiscal year, more than 40%—or 24,746—were filed in court districts neighboring the Mexican border….Nearly 22% (13,383) were drug related, 19.7% (12,123) were violent crimes and 10.2% (6,300) involved white-collar offenses that include a full range of frauds committed by business and government professionals. Read those stats closely because the media will lie and claim the crimes involve border enforcement. As you can see, over 40% involve drugs and violence.

So here’s some interesting math. “More than 40% – or 24,746 – were filed in court districts neighboring the Mexican border.” Ellipsis. “13,383 were drug-related, 12,123 were violent crimes, and 6,300 were white-collar fraud.” Wait a second. Those three numbers add up to 31,806, more than the total number of cases. Apparently we can’t trust illegal immigrants to obey any laws, including the laws of mathematics.

Breitbart’s reference goes here and their reference is here. When I look it up, these turn out to be the numbers across the entire US, not the numbers for southern border regions. This makes sense – do we really think illegal immigrants commit 6,300 cases of white-collar fraud per year in their dynamic illegal immigrant megacorporations?

But this is really embarrassing for Breitbart’s case. Their whole point is that the disproportionate number of crimes committed in immigrant-heavy areas are not just immigration offenses, but in fact representative of the country’s crime load as a whole. But that’s only because they’re accidentally looking at the country’s crime load as a whole instead of at crimes committed in immigrant-heavy areas! I can’t actually find the immigrant-heavy-area data, but I’d be willing to bet that the disproportionate number of federal crimes along the border do in fact involve border enforcement, exactly the argument they claim to debunk.

6. There are more than 2,000 sex offenders deported by ICE every year in Texas alone…of the 862 alien sex offenders deported by the Texas-based offices, about 27 percent were convicted of sex offenses against children?

Quick, how many illegal aliens in Texas? If it’s 10 million, then their sexual offense rate is far lower than that of any other population. If it’s 10,000, their sexual assault rate is far higher than that of any other population. Not only does Breitbart not give these numbers, but they don’t even seem to understand that they should give them, or expect their readers to care.

I can’t actually figure this number out because it depends on knowing what percent of immigrant sex offenders are deported each year. Consider: there are 86000 sex offenders in Texas. About 6% of Texans are illegal immigrants, so by chance we should expect about 5000 sex offenders total. Given that more than 2000 are deported per year, and that this has been happening for more than three years, that sounds like there are (or were) more than 5000, unless some deportees came back, which we know some do. But I don’t know that the sex offender registry is measuring the same kind of sex-offenderness as the illegal immigration numbers, so I can’t be sure of this. Also, the Texas statistics for immigrants include Oklahoma and possible other areas, so it’s hard to directly compare.

According to this page, there are 5017 arrests for illegal aliens for sexual assault over 4.5 years, so about 1100/year. That should be the sort of number we can work with. But it’s well above the total number of Hispanics arrested for sexual assault given here, immigrant and native-born alike, which doesn’t make sense. So I don’t know what to do.

The only good source I can find for percent violent crime by illegal immigrants in Texas is this one, which says that they commit 7.5% of murders. But they’re 6% of the population, so that’s pretty much what we’d expect. This accords with the numbers mentioned above, where in federal prisons the percent illegal immigrants serving time for sexual assault was proportionate to their percent of the population.

I can’t find good numbers here. A very rough inference from one source of Texas sex offender numbers would suggest that the number of sexual offense deportations is unusually high, but inferring based on the murder and sexual abuse numbers suggests that sexual offenses are about average. I’m not sure which method is more correct – but in any case, whatever the truth is there’s no way Breitbart’s numbers could be expected to get anyone any closer to it.

8. In 2013 the Obama administration released 36,007 criminal immigrants who had nearly 88,000 convictions between them. Those convictions included 193 homicide convictions, 426 sexual assault convictions, 303 kidnapping convictions, and 1,075 aggravated assault convictions In January, the DHS admitted to Sen. Grassley that 1,000 of the 36,007 released had gone on to commit more crime including: terroristic threats, lewd acts with a minor, various types of assault, DUI, robbery, hit-and-run, gang activity, rape, and child cruelty.

Politifact rates this half true. All immigrants involved served full prison sentences appropriate for their crime; the concern isn’t that they didn’t serve their time but that after their time was up they didn’t get deported. The DHS says that their policy is that if they haven’t finished a deportation case by the time a criminal gets out of prison, the criminal may be released until the deportation case is finished, mostly because the prison system won’t keep them and the DHS doesn’t have enough immigration-related prisons of its own. These people are subjected to the usual monitoring and may be rearrested and deported after their deportation case comes through. There is some reason for concern in that about 3/4s of these people manage to lose themselves before the deportation case is complete, but the DHS reasonably says that if people want this to stop happening people should give them more funding.

And once again, without more information we can’t tell whether any of this involves a higher crime rate than any other population.

9. ICE is finding and removing more criminal aliens each year. The number ordered removed has gone up from 7,000 in 2007 to 79,000 in 2010. These criminals are not being stopped at the border. These criminals are being deported after making it across the border and committing tens of thousands of crimes.

Sheesh. You get angry when we don’t deport immigrants. You get angry when we do deport immigrants? Make up your minds! This is literally just complaining that we’re getting better at solving the problem you complained about above!

There are between ten and twenty million illegal immigrants in the United States – about equal to the number of New Yorkers. If somebody wanted to expel New York from the country, they could point out that New Yorkers commit 616 homicides, 2,534 rapes, and 45,206 cases of aggravated assaults per year. Or that we need seventy-one prisons just to contain all the New Yorker criminals in our justice system, and we don’t have nearly enough funding to run all of them effectively, such that literally thousands of New Yorker criminals, including murderers and rapists, are released into the general population every year. You could report that [number] of New Yorkers commit violent sexual assaults on children each year – I don’t know what the number is, but I guarantee you that it is a number, that it has a certain number of digits, and that it is worse than zero.

Breitbart consistently fails to give numbers that would mean anything or inform anybody, and when it either uses numbers that are loaded in its favor or ones that don’t mean what it thinks they mean. The actual numbers on the question of interest don’t seem particularly outrageous.

This isn’t to say that there can’t be legitimate concerns about illegal immigration. In fact, that’s my whole point and something that I wish conservatives better understood: we need to practice Gettier politics. There’s a theory on the Right that since the media has created a giant edifice of lies to justify liberalism, liberalism must be false. But other parts of the media have created a giant edifice of lies to justify conservativism. Instead of assuming our opponents are necessarily gullible morons who believe the giant edifice of lies on their side, we should kind of awkwardly go “Oh, there’s a giant edifice of lies on your side too? Yeah, I know that feeling,” and listen to what they have to say.

A major obstacle to changing one’s mind is the fallacy that if the other side is right, it must be decent and intellectually honest.

— Steven Kaas (@stevenkaas) August 21, 2010

# Side Effects May Include Anything

A couple of days ago a patient said he’d become depressed after starting Xolair, a new asthma drug I know nothing about.

On the one hand, lots of things that mess with the immune system can cause depression. On the other, patients are notorious for blaming drugs for any random thing that happens around the same time they started taking them. So I did what any highly-trained competent medical professional would: I typed “does xolair cause depression?” into Google.

The results seemed promising. The first site was called “Can Xolair cause depression?”. The second was “Is depression a side effect of Xolair?”. Also on the front page were “Could Xolair cause major depression?” and “Xolair depression side effects”. Clearly this is a well-researched topic that lots of people cared about, right?

Let’s look closer at one of those sites, EHealthMe.com. It says: “Major depression is found among people who take Xolair, especially for people who are female, 40-49 old, also take medication Singulair, and have Asthma. We study 11,502 people who have side effects while taking Xolair from FDA and social media. Among them, 14 have Major depression. Find out below who they are, when they have Major depression and more.” Then it offers a link: “Join a support group for people who take Xolair and have Major depression”.

First things first: if there were actually 11502 people taking Xolair, and only 14 of them had major depression, that would be a rate of 0.1%, compared to 6.9% in the general population. In other words, Xolair would be the most effective antidepressant on Earth. But of course nobody has ever done an n=11502 study on whether a random asthma medication causes depression, and EHealthMe is just scraping the FDA databases to see how many people reported depression as a side effect to the FDA. But only a tiny percent of people who get depression report it, and depression sometimes strikes at random times whether you’re taking Xolair or not. So this tells us nothing.

And yet a patient who worries that Xolair might be causing their depression will Google “can xolair cause depression?”, and she will end up on this site that says “major depression is found among people who take Xolair”, which is one of the worst examples of weasel words I’ve ever heard. Then she will read that there are entire support groups for depressed Xolair sufferers. She will find all sorts of scary-looking information like that Xolair-related depression has been increasing since 2008. And this is above and beyond just the implications of somebody bothering to write an entire report about the Xolair-depression connection!

In case you haven’t guessed the twist – no one’s ever investigated whether Xolair causes depression. EHealthMe’s business model is to make an automated program that runs through every single drug and every possible side effect, scrapes the FDA database for examples, then autopublishes an ad-filled web page titled “COULD $DRUG CAUSE $SIDE\_EFFECT?”. It populates the page by spewing random FDA data all over it, concludes “$SIDE\_EFFECT is found among people who take $DRUG”, and offers a link to a support group for $DRUG patients suffering from $SIDE\_EFFECT. Needless to say, the support group is an automatically-generated forum with no posts in it.

And it’s not just EHealthMe. This is a whole market, with competitors elbowing their way past one another to the top of the Google search results. Somebody who doubts EHealthMe and seeks an online second opinion will probably just end up at PatientsVille, whose page is called “Xolair Depression Side Effects”, which contains the same FDA data, and which gets the Google description text “This opens a possibility that Xolair could cause Depression”. Or Treato, whose page claims to contain 56 reader comments on Xolair and depression, but which has actually just searched the Web for every single paragraph that contains “Xolair” and “depression” together and then posted garbled excerpts in its comment section. For example, one of their comments – and this is not at all clear from Treato’s garbled excerpt – is from a tennis forum, where a user with the handle Xolair talks about how his tennis serve is getting worse with age; another user replies “Xolair, I read this and get depressed, I just turned 49.” But if you don’t check whether it came from a tennis forum or not, 56 reports of a connection between a drug and a side effect sounds convincing!

This is really scummy. Maybe it’s not the most devious of traps for you or me, but what about for your grandmother? What about for those people who send money to Nigerian princes? The law is usually pretty strict about who can and can’t provide medical information – so much so that it cracks down on 23andMe just for reading off the genome in a way that uneducated people might misinterpret. Yet somehow sites like EHealthMe are allowed to continue, because they just very strongly imply fake medical information instead of saying it outright.

Remember, only about 50% of people who are prescribed medication take it. Sometimes it’s personal choice or simple forgetfulness. But a lot of the time they stop because of side effects. I had a patient a few months ago who was really depressed. I started her on an antidepressant and she got much better. Then she stopped the medication cold turkey and got a lot worse again. I asked her why she’d stopped. She said her shoulder started hurting, she’d Googled whether antidepressants could cause shoulder pain, and read that they could. She couldn’t remember what site she was reading, but I bet it was EHealthMe or Treato or some of the others just like them.

One day, somebody’s going to Google “can penicillin cause cancer?”, read a report with a link to a support group for penicillin-induced-cancer survivors, stop taking antibiotics, and die. And when that happens, I hope it’s in America, so I can be sure their family will sue the company involved for more money than exists in the entire world.

# Predictions For 2016

At the beginning of every year, I make predictions. At the end of every year, I score them. So here are a hundred more for 2016. Most of them are objectively decideable, but a few are subjective (eg “X goes well”) and are marked with asterisks.

WORLD EVENTS  
1. US will not get involved in any new major war with death toll of > 100 US soldiers: 60%  
2. North Korea’s government will survive the year without large civil war/revolt: 95%  
3. Greece will not announce it’s leaving the Euro: 95%  
4. No terrorist attack in the USA will kill > 100 people: 90%  
5. …in any First World country: 80%  
6. Assad will remain President of Syria: 60% [edit: called out as dumb, but I won’t cheat and change it]  
7. Israel will not get in a large-scale war (ie >100 Israeli deaths) with any Arab state: 90%  
8. No major intifada in Israel this year (ie > 250 Israeli deaths, but not in Cast Lead style war): 80%  
9\* No interesting progress with Gaza or peace negotiations in general this year: 90%  
10. No Cast Lead style bombing/invasion of Gaza this year: 90%  
11\* Situation in Israel looks more worse than better: 70%  
12. Syria’s civil war will not end this year: 70%  
13. ISIS will control less territory than it does right now: 90%  
14. ISIS will not continue to exist as a state entity [added: meant in Iraq/Syria]: 60% [edit: called out as dumb, but I won’t cheat and change it]  
15. No major civil war in Middle Eastern country not currently experiencing a major civil war: 90%  
16\* Libya to remain a mess: 80%  
17. Ukraine will neither break into all-out war or get neatly resolved: 80%  
18. No country currently in Euro or EU announces plan to leave: 90%  
19. No agreement reached on “two-speed EU”: 80%  
20. Hillary Clinton will win the Democratic nomination: 95%  
21. Donald Trump will win the Republican nomination: 60%  
22\* Conditional on Trump winning the Republican nomination, he impresses everyone how quickly he pivots towards wider acceptability: 70%  
23. Conditional on Trump winning the Republican nomination, he’ll lose the general election: 80%  
24. Conditional on Trump winning the Republican nomination, he’ll lose the general election worse than either McCain or Romney: 70%  
25. Marco Rubio will not win the Republican nomination: 60% [edit: called out as dumb, but I won’t cheat and change it]  
26. Bloomberg will not run for President: 80%  
27. Hillary Clinton will win the Presidency: 60%  
28. Republicans will keep the House: 95%  
29. Republicans will keep the Senate: 70%  
30. Bitcoin will end the year higher than $500: 80%  
31. Oil will end the year lower than $40 a barrel: 60%  
32. Dow Jones will not fall > 10% this year: 70%  
33. Shanghai index will not fall > 10% this year: 60% [edit: called out as dumb, but I won’t cheat and change it]  
34. No major revolt (greater than or equal to Tiananmen Square) against Chinese Communist Party: 95%  
35. No major war in Asia (with >100 Chinese, Japanese, South Korean, and American deaths combined) over tiny stupid islands: 99%  
36. No exchange of fire over tiny stupid islands: 90%  
37. US GDP growth lower than in 2015: 60%  
38. US unemployment to be lower at end of year than beginning: 50%  
39. No announcement of genetically engineered human baby or credible plan for such: 90%  
40\* No major change in how the media treats social justice issues from 2015: 70%  
41\* European far right makes modest but not spectacular gains: 80%  
42\* Mainstream European position at year’s end is taking migrants was bad idea: 60%  
43. [Duplicate removed]  
44\* So-called “Ferguson effect” continues and becomes harder to deny: 70%  
45. SpaceX successfully launches a reused rocket: 50%  
46\* Nobody important changes their mind much about the EMDrive based on any information found in 2016: 80%  
47. California’s drought not officially declared over: 50%  
48. No major earthquake (>100 deaths) in US: 99%  
49. No major earthquake (>10000 deaths) in the world: 60%  
50. Occupation of Oregon ranger station ends: 99%

PERSONAL/COMMUNITY  
1. SSC will remain active: 95%  
2. SSC will get fewer hits than in 2015: 60%  
3. At least one SSC post > 100,000 hits: 50%  
4. UNSONG will get fewer hits than SSC in 2016: 90%  
5. > 10 new permabans from SSC this year: 70%  
5. UNSONG will get > 1,000,000 hits: 50%  
6. UNSONG will not miss any updates: 50%  
7. UNSONG will have higher Google Trends volume than HPMOR at the end of this year: 60%  
8. UNSONG Reddit will not have higher average user activity than HPMOR Reddit at the end of this year: 60%  
9. Shireroth will remain active: 70%  
10. I will be involved in at least one published/accepted-to-publish research paper by the end of 2016: 50%  
11. I won’t stop using Twitter, Tumblr, or Facebook: 95%  
12. > 10,000 Twitter followers by end of this year: 50%  
13. I will not break up with any of my current girlfriends: 70%  
14. I will not get any new girlfriends: 50%  
15. I will attend at least one Solstice next year: 90%  
16. …at least two Solstices: 70%  
17. I will finish a long blog post review of stereotype threat this year: 60%  
18\* Conditional on finishing it, it won’t significantly change my position: 90%  
19. I will finish a long FAQ this year: 60%  
20. I will not have a post-residency job all lined up by the end of this year: 80%  
21. I will have finished all the relevant parts of my California medical license application by the end of this year: 70%  
22. I will no longer be living in my current house at the end of this year: 70%  
23. I will still be at my current job: 95%  
24. I will still not have gotten my elective surgery: 80%  
25. I will not have been hospitalized (excluding ER) for any other reason: 95%  
26. I will not have taken any international vacations with my family: 70%  
27. I will not be taking any nootropic daily or near-daily during any 2-month period this year: 90%  
28. I will complete an LW/SSC survey: 80%  
29. I will complete a new nootropics survey: 80%  
30. I will score 95th percentile or above in next year’s PRITE: 50%  
31. I will not be Chief Resident next year: 60%  
32. I will not have any inpatient rotations: 50%  
33. I will continue doing outpatient at the current clinic: 90%  
34\* I will not have major car problems: 60%  
35\* I won’t publicly and drastically change highest-level political/religious/philosophical positions (eg become a Muslim or Republican): 90%  
36. I will not vote in the 2016 primary: 70%  
37. I will vote in the 2016 general election: 60%  
38. Conditional on me voting and Hillary being on the ballot, I will vote for Hillary: 90%  
39\* I will not significantly change my mind about psychodynamic or cognitive-behavioral therapy: 80%  
40. I will not attend the APA meeting this year: 80%  
41. I will not do any illegal drugs (besides gray-area nootropics) this year: 90%  
42. I will not get drunk this year: 80%  
43\* Less Wrong will neither have shut down entirely nor undergone any successful renaissance/pivot by the end of this year: 60%  
44. No co-bloggers (with more than 5 posts) on SSC by the end of this year: 80%  
45. I get at least one article published on a major site like Huffington Post or Vox or New Statesman or something: 50%  
46. I still plan to move to California when I’m done with residency: 90%  
47. I don’t manage to make it to my friend’s wedding in Ireland: 60%  
48. I don’t attend any weddings this year: 50%  
49. I decide to buy the car I am currently leasing: 60%  
50. Except for the money I spend buying the car, I make my savings goal before July 2016: 90%

Other people doing yearly predictions with probability: Against Jebel al-Lawz, Anatoly Karlin, Old Lamps, Garrett Peterson. If you’re doing this and I missed it, let me know and I’ll add you in.

# Staying Classy

Siderea writes an essay on class in America. You should read it. In case you don’t, here’s the summary:

1. People tend to confuse social class with economic class, eg how much money you make. But social class is a more complicated idea involving how respectable you seem, how educated you are, and what kind of family you come from. An assembly line supervisor might make the same amount of money as a schoolteacher, but the schoolteacher would probably seem more refined and be able to access better social circles.

2. Classes are cultures. People in a certain class have their own way of dressing, speaking, decorating, and behaving. They have distinctive ideas and values. This is why a lower-class person cannot simply claim to be upper-class and so gain all the benefits of upper-class-hood; it would be as hard as trying to pass for Japanese. Lower-class people can learn their way around upper-class culture, but it’s a difficult and lifelong project done most easily if you already have upper-class resources.

3. Talking about class is taboo because we like to believe we’re a classless society. We talk about income instead and pretend it’s class. Class breaks through in a couple of phrases like “rednecks” or “white trash” or “white collar” or “coastal elites”, but people use the phrases without usually having a broader idea that it’s class they’re talking about.

4. Class prejudice is complicated. It combines the practical superiority of being upper-class to being lower-class (because you have more money and opportunity) with the very dubious value judgment that upper-class culture is superior to lower-class culture, or that lower-class culture is just people trying to do upper-class culture but failing. But lower-class people like lower-class culture and generally do not want to adopt upper-class culture, except insofar as it’s necessary to advance. Analogies to race and assimilation are obvious.

5. People mostly understand their own class, and the class one step above or below them, but have only vague stereotypes of classes further than that. This limits social mobility; you can’t join what you can’t understand.

6. College is a finishing school for the upper classes. They send their children there to learn the proper upper class values and behaviors. Even if community college does a great job teaching whatever trades it teaches, it will not teach you how to be a part of the upper class, and this will seriously limit your opportunities.

7. Politically, the left pretends class doesn’t exist; the right talks about it, but only to yell at the underclass and say that their culture is wrong. Race is really complicated and will be left out of this analysis.

I notice Siderea is a psychotherapist, which doesn’t surprise me. We in mental health get a pretty good cross-sectional exposure of everybody and get to hear about their lives, and with enough data points the structure comes into sharper relief.

Just to give an example: suppose a lady comes in with really over-permed dyed curly hair wearing several rings, bracelets, and necklaces. Her name is Sherri and she calls you “darling”; she’s also carrying her lunch, which is KFC plus a Big Gulp. Without knowing anything else about her, you can peg her as working class. Maybe she won the lottery ten years ago and is now the richest person in your state. It doesn’t matter. She’s still working class.

Or suppose a thin 25-year-old man comes in wearing glasses, a small close-cropped beard, and a Led Zeppelin t-shirt. His name is Alex and he apologizes for being three minutes late. This guy is probably middle-to-upper-middle-class and college educated, maybe not a great college but still college-educated. And maybe he’s fallen on hard times and doesn’t have a dollar to his name. It still doesn’t matter. He’s still middle-to-upper-middle class.

And you start to learn you can predict things about these people, the concerns they’re going to have, the kind of things that happen to them. Who their friends are. How they relate to their friends: Sherri will expound upon the flaws of every single one of her ungrateful coworkers; Alex will reluctantly say he went through a tough breakup a year or two ago. What kind of drugs they abuse, if they abuse drugs (maybe Sherri has smoking and drinking problems; Alex has probably tried marijuana and LSD but is embarrassed to say so).

But this kind of innate stereotyping is different than a formal taxonomy. Siderea links to Michael Church’s attempt to explain what the classes actually are. This is another piece you should read, but again in case you don’t:

1. 10% of people are in an underclass consisting of “generationally poor” people who may never have held jobs and who come from similarly poor families.

2. 65% of people are in the labor class. They work jobs where labor is seen as a commodity, ie there’s not as much sense of career capital or reputation. They base virtue and success around Hard Work. Its lower levels are minimum wage McJobs, its middle levels are assembly line work, and its higher levels are things like pilots, plumbers, and small business owners. The stratospheric semi-divine level is “celebrities” like reality TV stars who become fabulously rich and famous while sticking to their labor class roots.

3. 23.5% of people are in the gentry class. They fetishize education and career capital. They engage in all sorts of signaling games around “fair trade” and “organic” and what museums they go to. At the lower level they’re schoolteachers and starving artists, at the mid level they’re “professions” like engineering and law, and at the highest level they’re professors and scientists and entrepreneurs. The stratospheric semi-divine level is “cultural influencers” like Jon Stewart or Steven Pinker who become famous and (maybe) rich while sticking to their gentry class roots.

4. 1.5% of people are in the elite class. Although you can be borderline-elite by getting a job in finance and making a few million, the real elite are born into money and don’t work unless they want to. Occasionally they’ll sit on a board or found a philanthropic association or something. They don’t believe in “professional achievement” because working is lower-class; they might compete in complicated status games around who throws the best parties or has the best horses or whatever.

5. The highest class (E1) are psychopaths who burn the global commons for shits and giggles. They tend to be drug lords, arms dealers, and morally insane billionaires. Most famous politicians and businesspeople are not in this class and most people in this class are not famous.

6. The three main classes (labor, gentry, and elite) are three different ‘infrastructures’. To be in labor you need skills, to be in gentry you need education, and to be in elite you need connections. There’s no strict hierarchy (eg not all gentry are above all labor), but you can picture them as offset ladders, with the lower gentry being at the same rung as the higher labor and so on.

7. The Elite control everything; the constant threat is that Gentry and Labor will unite against them, which might very well work. The Elite neutralize this threat by making Labor hate Gentry as “effeminate” or “pretentious”; they also convince Labor that the Gentry are probably secretly in cahoots with the underclass against Labor. Elites also convince Labor that Elites don’t exist and it’s Gentry all the way up, which means that “anti-1%” sentiment, which should properly get Labor and Gentry to cooperate against the Elites, instead makes Gentry hate the Elites but Labor hate Gentry. Politics boils down to Gentry being good people trying to improve things, and Elite conning Labor into hating Gentry to prevent things from being improved.

8. While all classes can have good and bad people (except E1, which is wholly bad), Elites have a generally negative influence on society, and Gentry are generally positive. After the World Wars, everybody got angry at the Elites for all the war and killing and stuff, which convinced them to lie low for a few decades and forced the Gentry to take over. This was why the country did so well during the 50s and 60s. Whether the country goes in a good or bad direction now depends on whether the Elites manage to take it back or not. One reason Silicon Valley works (used to work?) so unusually well was that it was mostly a native project of the Gentry that hadn’t yet been infiltrated by the Elites.

Reaction to Church on the subreddit was pretty negative, but I find it at least a good nucleus for further discussion. The Gentry/Labor distinction is glaringly obvious. The Labor/Underclass distinction also seems glaringly obvious to me, if only because Labor hates the underclass. The Gentry/Elite distinction doesn’t seem glaringly obvious to me, but maybe that’s just because I haven’t met enough elites. In particular, Church’s “E1” seems caricatured and out-of-place in his otherwise sober analysis. Then again, if those people existed I probably wouldn’t know anyway. Then again, the rest of Church’s blog suggests some paranoid tendencies, so maybe the E1 entry is just those coming out.

Siderea notes that Church’s analysis independently reached about the same conclusion as Paul Fussell’s famous guide. I’m not entirely sure how you’d judge this (everybody’s going to include lower, middle, and upper classes), but eyeballing Fussell it does look a lot like Church, so let’s grant this.

It also doesn’t sound too different from Marx. Elites sound like capitalists, Gentry like bourgeoisie, Labor like the proletariat, and the Underclass like the lumpenproletariat. Or maybe I’m making up patterns where they don’t exist; why should the class system of 21st century America be the same as that of 19th century industrial Europe?

There’s one more discussion of class I remember being influenced by, and that’s Unqualified Reservations’ Castes of the United States. Another one that you should read but that I’ll summarize in case you don’t:

1. Dalits are the underclass, made up of homeless people, chronically unemployed people, drug addicts, etc. They tend to have a lot of trouble with the law, go in and out of jail, never really hold down stable employment. Status is “street cred” that you get from being powerful, wealthy, and sexually successful, eg gang leaders.

2. Vaisyas are standard middle-class people who engage in productive employment. They tend to form nuclear families and try to go to church. Status is having a stable job, a stable family, and being well-liked in your church or social club.

3. Brahmins are very educated people who participate in the world of ideas. They range from doctors and lawyers to artists and professors. Access is conferred by top-tier university education. Status is from conspicuous engagement in progressive politics, eg being an activist, working for an NGO, “campaigning for justice”. They are “the ruling class”.

4. Optimates are very rich WASPs concerned with breeding and old money. Status comes from breeding and an antiquated idea of “nobility”. Optimates used to be “the ruling class”, but now they’re either extinct or endangered, having been pretty much absorbed into the Brahmins.

5. Mentioned elsewhere in the UR corpus: politics boils down to Vaisyas being basically decent people trying to lead normal productive lives, and Brahmins trying to create a vast tentacled monstrosity of useless bureaucrats and petty enforcers of ideological conformity to employ Brahmins in the “knowledge work” they feel entitled to and to protect their interests. Silicon Valley is (used to be?) unusually functional because it maintained some Vaisya values separate from the corrupting influence of the Brahmins.

Michael Church’s system (henceforth MC) and the Unqualified Reservation system (henceforth UR) are similar in some ways. MC’s Underclass matches Dalits, MC’s Labor matches Vaisyas, MC’s Gentry matches Brahmins, and MC’s Elite matches Optimates. This is a promising start. It’s a fourth independent pair of eyes that’s found the same thing as all the others. (commenters bring up Joel Kotkin and Archdruid Report as similar convergent perspectives).

But there are also some profound differences. UR says that the Elites are mostly gone, that everything’s ruled by the Gentry nowadays, and that the Gentry are allying with the criminal Underclass against Labor. MC mentions this same picture, but only as the false facade that the Elites are trying to get everyone else to believe in order to keep them divided.

You could reconcile some of the differences by supposing the two models have different cutoffs. Suppose we rank people from 0 (lowest underclass) to 100 (highest elite). Maybe MC draws the Labor/Gentry and Gentry/Elite borders at 40 and 70 respectively, and UR draws the Vaisya/Brahmin and Brahmin/Optimate borders at 60 and 90. If the world’s being run by 80s, MC could be right to say it’s run by Elites and not Gentry, and UR could be right in saying it’s run by Brahmins and not Optimates. If Silicon Valley is run by 55s but being ruined by 75s, MC could say it’s run by Gentry but ruined by elites, and UR could say it’s run by Vaisyas but ruined by Brahmins. But if there’s this much variability in class boundaries, what’s the point in even drawing them in the first place?

But I think the differences are real and political: MC comes from a liberal perspective, UR from a conservative one. MC wants to locate the source of the cancer in the (mostly plutocrat) Elites, cast the (mostly liberal) Gentry as wonderful people who can do no wrong, cast the (mostly conservative) Labor as deluded and paranoid, and cast the (liberal-aligned) Underclass in a sympathetic light. UR wants to locate the source of the cancer in the (mostly liberal) Brahmins, cast the (mostly conservative) Labor as decent salt-of-the-Earth types under threat from the elite, and cast the (liberal-aligned) Underclass in an unsympathetic light.

And the political angle evokes one more system worth adding here: my own discussion of the Blue Tribe vs. the Red Tribe in I Can Tolerate Anything But The Outgroup. I point out that the group sometimes referred to as “coastal liberals” or “SWPL” and so on are marked not only by Democratic Party beliefs, but by a host of cultural similarities including food, dress, music, hobbies, religion, values, art, etc. Likewise, the group sometimes referred to as “rednecks” or “fundies” and so on are marked not only by Republican Party beliefs, but by a similar set of cultural similarities. I call these the “Blue Tribe” and the “Red Tribe” as an attempt to distinguish them as cultures and not just as sets of political beliefs.

These tribes seem closely related to classes. “Blue Tribe” is similar to Gentry; “Red Tribe” is similar to Labor. I won’t say there’s a perfect 1:1 equivalence; for example, I know some union leaders who are very clearly in the Labor class but who wouldn’t be caught dead in the Red Tribe. But the resemblance is too close to miss.

Some final scattered thoughts:

1. All those studies that analyze whether some variable or other affects income? They’d all be much more interesting if they analyzed the effect on class instead. For example, there’s a surprisingly low correlation between your parents’ income and your own income, which sounds like it means there’s high social mobility. But I grew up in a Gentry class family; I became a doctor, my brother became a musician, and my cousin got a law degree but eventually decided to work very irregularly and mostly stay home raising her children. I make more money than my brother, and we both make more money than my cousin, but this is not a victory for social mobility and family non-determinism; it’s no coincidence none of us ended up as farmers or factory workers. We all ended up Gentry class, but I chose something closer to the maximize-income part of the Gentry class tradeoff space, my brother chose something closer to the maximize-creativity part, and my cousin chose to raise the next generation. Any studies that interpret our income difference as an outcome difference and tries to analyze what factors gave me a leg up over my relatives (better schools? more breastfeeding as a child?) are stupid and will come up with random noise. We all got approximately the same level of success/opportunity, and those things just happen to be very poorly measured by money. If we could somehow collapse the entirety of tradeoffspace into a single variable, I bet it would have a far greater parent-child correlation than income does. This is part of why I don’t follow the people who take the modest effect of IQ on income as a sign that IQ doesn’t change your opportunities much; maybe everyone in my family has similar IQs but wildly different income levels, and there’s your merely modest IQ/income relationship right there. I think some studies (especially in Britain) have tried analyzing class and gotten some gains over analyzing income, but I don’t know much about this.

2. I think Siderea is right that the Right thinks in social class terms more naturally than the Left. To oversimplify, both sides use class warfare, but the Left’s class warfare is economic (“the plutocrat billionaires are ruining everything!”) and the Right’s class warfare is social (“the media and academic elites are ruining everything!”).

3. Closely related: Donald Trump appeals to a lot of people because despite his immense wealth he practically glows with signs of being Labor class. This isn’t surprising; his grandfather was a barber and his father clawed his way up to the top by getting his hands dirty. He himself went to a medium-tier college and is probably closer in spirit to the small-business owners of the upper Labor class than to the Stanford MBA-holding executives of the Elite. Trump loves and participates in professional wrestling and reality television; those definitely aren’t Gentry or Elites pastimes! When liberals shake their heads wondering why Joe Sixpack feels like Trump is a kindred soul even though Trump’s been a billionaire his whole life, they’re falling into the liberal habit of sorting people by wealth instead of by class. To Joe Sixpack, Trump is “local boy made good”.

4. The thesis of “I Can Tolerate Anything But The Outgroup” simplifies to “It is a Gentry-class tradition to sweep aside all prejudices except class prejudice, which must be held with the intensity of all the old prejudices combined.”

5. But “I Can Tolerate Anything But The Outgroup”‘s Grey Tribe sits uneasy within this system. It doesn’t seem to be a class. But it also seems distinctly different from ordinary Gentry norms. And what about minorities? What about the differences between farmers vs. factory workers? If different classes are equivalent to different cultures, well, there are a lot of different cultures that don’t fit easily into the hierarchy. Maybe class is one factor among many that can create a different culture, but other factors can be stronger than class in some groups?

6. Siderea doesn’t want to get into how race interacts with class, and that seems wise. But a related digression: lots of people complain about social justice being classist, in that it’s hard for anybody who hasn’t either gone to college or at least spent a lot of time hanging around social justice people to keep track of which words, opinions, and causes are okay versus will render you radioactive. On the one hand, this is probably true. On the other, it’s probably true of everything, with social justice as an unexceptional example. Yes, the way you refer to trans people shows what class you’re from, but so does the way you order ice cream.

7. Siderea admits she is classist and not ashamed of this. I have a hard time understanding what she means, but I can try to explain my own classism: I think classes probably sort on important qualities and reinforce those qualities. For example, the Underclass and Labor class people I know are much more likely to have high-conflict styles of interaction: if they feel offended, they’ll yell at you and maybe even fight you. Gentry class people would be horrified at the thought; they might respond to the same offense by filing a complaint with Human Resources. I think there are two equally correct ways to interpret this. Number one, people with the maladaptive behavior of starting physical fights don’t make it very far in life and so end out in lower classes, and insofar as these behaviors are either genetic or learned within the family, their families stay in lower classes throughout the generations. Number two, the lower classes have a culture where you defend your honor by fighting people who offend you, and the upper classes have a culture where you defend your honor by submitting complaints, and although in a cosmic sense both of those styles are equally valid, and although indeed a thousand years ago the fighting might have been more adaptive, in today’s society the complaint-submitting is more adaptive and the lower classes are screwed unless they unlearn that behavior – which they probably won’t, because unlearning class is hard. But this means that classism is at least kind of justified – if you want to hire for example a schoolteacher, you might want to look for people who show all the signs of Gentry rather than Labor class to make sure they’re not going to get into physical fights in the classroom.

8. Cellular automaton theory of fashion likely relevant.

9. Siderea’s idea of college as finishing school for the upper classes is interesting, and her own experience is a window into something I never thought about before. But I’m not sure how typical she is; I think most colleges admit students who are already members of the classes their graduates end up in. I felt like I didn’t learn any class culture during my own college experience at all – which isn’t surprising since I was born the son of a doctor and ended up as a doctor myself. I think my story’s probably more typical than Siderea’s, though other people can prove me wrong if they’ve seen differently.

# Book Review: Superforecasting

Philip Tetlock, author of Superforecasting, got famous by studying prediction. His first major experiment, the Expert Political Judgment experiment, is frequently cited as saying that top pundits’ predictions are no more accurate than a chimp throwing darts at a list of possibilities- although Tetlock takes great pains to confess to us that no chimps were actually involved, and this phrasing just sort of popped up as a flashier way of saying “random”.

Although this was generally true, he was able to distinguish a small subset of people who were able to do a little better than chance. His investigation into the secrets of their very moderate success led to his famous “fox” versus “hedgehog” dichotomy, based on the fable that “the fox knows many things, the hedgehog knows one big thing”. Hedgehog pundits/experts are people who operate off a single big idea- for example, an economist who says that government intervention is always bad, predicts doom for any interventionist policy, and predicts great success for any noninterventionist one. Foxes are people who don’t have much of a narrative or ideology, but try to find the right perspective to approach each individual problem. Tetlock found that the hedgehogs did worse than the chimp and the foxes did a little better.

Cut to the late 2000s. The US intelligence community has just been seriously embarrassed by their disastrous declaration that there were weapons of mass destruction in Iraq. They set up an Intelligence Advanced Research Projects Agency to try crazy things and see if any of them worked. IARPA approached a bunch of scientists, handed them a list of important world events that might or might not happen, and told them to create some teams and systems for themselves and compete against each other to see who could predict them the best.

Tetlock was one of these scientists, and his entry into the competition was called the Good Judgment Project. The plan was simple: get a bunch of people to sign up and try to predict things, then find the ones who did the best. This worked pretty well. 2,800 people showed up, and a few of them turned out to be…

…okay, now we’re getting to a part I don’t understand. When I read Tetlock’s paper, all he says is that he took the top sixty forecasters, declared them superforecasters, and then studied them intensively. That’s fine; I’d love to know what puts someone in the top 2% of forecasters. But it’s important not to phrase this as “Philip Tetlock discovered that 2% of people are superforecasters”. This suggests a discontinuity, a natural division into two groups. But unless I’m missing something, there’s no evidence for this. Two percent of forecasters were in the top two percent. Then Tetlock named them “superforecasters”. We can discuss what skills help people make it this high, but we probably shouldn’t think of it as a specific phenomenon.

Anyway, the Good Judgment Project then put these superforecasters on teams with other superforecasters, averaged out their decisions, slightly increased the final confidence levels (to represent the fact that it was 60 separate people, all of whom were that confident), and presented that to IARPA as their final answer. Not only did they beat all the other groups in IARPA’s challenge in a landslide, but they actually did 30% better than professional CIA analysts working off classified information.

Having established that this is all pretty neat, Tetlock turns to figuring out how superforecasters are so successful.

First of all, is it just luck? After all, if a thousand chimps throw darts at a list of stocks, one of them will hit the next Google, after which we can declare it a “superchimp”. Is that what’s going on here? No. Superforecasters one year tended to remain superforecasters the next. The year-to-year correlation in who was most accurate was 0.65; about 70% of superforecasters in the first year remained superforecasters in the second. This is definitely a real thing.

Are superforecasters just really smart? Well, sort of. The superforecasters whom Tetlock profiles in his book include a Harvard physics PhD who speaks 6 languages, an assistant math professor at Cornell, a retired IBM programmer data wonk, et cetera. But the average superforecaster is only at the 80th percentile for IQ – just under 115. And there are a lot of people who are very smart but not very good at predicting. So while IQ definitely helps, it isn’t the whole story.

Are superforecasters just really well-informed about the world? Again, sort of. The correlation between well-informedness and accuracy was about the same as the correlation between IQ and accuracy. None of them are remarkable for spending every single moment behind a newspaper, and none of them had as much data available as the CIA analysts with access to top secret information. Even when they made decisions based on limited information, they still beat other forecasters. Once again, this definitely helps, but it’s not the whole story.

Are superforecasters just really good at math? Again, kind of. A lot of them are math PhDs or math professors. But they all tend to say that they don’t explicitly use numbers when doing their forecasting. And some of them don’t have any kind of formal math background at all. The correlation between math skills and accuracy was about the same as all the other correlations.

So what are they really good at? Tetlock concludes that the number one most important factor to being a superforecaster is really understanding logic and probability.

Part of it is just understanding the basics. Superforecasters are less likely to think in terms of things being 100% certain, and – let’s remember just how far left the bell curve stretches – less likely to assign anything they’re not sure about a 50-50 probability. They’re less likely to believe that things happen because they’re fated to happen, or that the good guys always win, or that things that happen will necessarily teach a moral lesson. They’re more likely to admit they might be wrong and correct themselves after an error is discovered. They’re more likely to debate with themselves, try to challenge their original perception, start asking “What could be wrong about this thing I believe?” rather than “How can I prove I’m right?”

But they’re also more comfortable actively using probabilities. Like my predictions, the Good Judgment Project made forecasters give their answers as numerical probability estimates – for example, 15% chance of a war between North and South Korea in the next ten years killing > 1000 people. Poor forecasters tend to make a gut decision based on feelings that superficially related to the question, like “Well, North Korea is pretty crazy, so they’re pretty likely to declare war, let’s say 90%” or “War is pretty rare these days, how about 10%?”. Superforecasters tend to focus on the specific problem in front of them and break it down into pieces. For example, they might start with the Outside View – it’s been about 50 years since the Koreas last fought, so their war probability per decade shouldn’t be more than about 20% – and then adjust that based on Inside View information – “North Korea has a lot fewer foreign allies these days, so they’re less likely to start something than they once were – maybe 15%”.

Or they might break the problem down into pieces: “There would have to be some sort of international incident, and then that incident would have to erupt into total war, and then that war would have to kill > 1,000 people. There are about two international incidents between the Koreas every year, but almost none of them end in war; on the other hand, because of all the artillery aimed at Seoul, probably any war that did happen would have an almost 100% chance of killing > 1,000 people” … and so on. One result is that while poor forecasters tend to give their answers in broad strokes – maybe a 75% chance, or 90%, or so on – superforecasters are more fine-grained. They may say something like “82% chance” – and it’s not just pretentious, Tetlock found that when you rounded them off to the nearest 5 (or 10, or whatever) their accuracy actually decreased significantly. That 2% is actually doing good work.

Most interesting, they seem to be partly immune to cognitive bias. The strongest predictor of forecasting ability (okay, fine, not by much, it was pretty much the same as IQ and well-informedness and all that – but it was a predictor) was the Cognitive Reflection Test, which includes three questions with answers that are simple, obvious, and wrong. The test seems to measure whether people take a second to step back from their System 1 judgments and analyze them critically. Superforecasters seem especially good at this.

Tetlock cooperated with Daniel Kahneman on an experiment to elicit scope insensitivity in forecasters. Remember, scope insensitivity is where you give a number-independent answer to a numerical question. For example, how much should an organization pay to save the lives of 100 endangered birds? Ask a hundred people, and maybe the average answer is “$10,000”. Ask a (different group of) a hundred people how much the same organization should pay to save the lives of 1000 endangered birds, and maybe the average answer will still be $10,000. So it seems you can get people to change their estimate of the value of bird life just by changing the number in the question. Poor forecasters do the same thing on their predictions. For example, a hundred poor forecasters might on average predict a 15% chance of war in Korea in the next five years, and a different group of a hundred poor forecasters might on average predict a 15% chance of war in Korea in the next fifteen years. They’re ignoring the question and just going off of a vague feeling of how likely another Korean war seems. Superforecasters, in contrast, showed much reduced scope insensitivity, and their probability of a war in five years was appropriately lower than of a war in fifteen.

Maybe all this stuff about probability calibration, inside vs. outside view, willingness to change your mind, and fighting cognitive biases is starting to sound familiar? Yeah, this is pretty much the same stuff as in the Less Wrong Sequences and a lot of CFAR work. They’re both drawing from the same tradition of cognitive science and rationality studies.

So as I said before, Superforecasting is not necessarily too useful for people who are already familiar with the cognitive science/rationality tradition, but great for people who need a high-status and official-looking book to justify it. The next time some random person from a terrible forum says that everything we’re doing is stupid, I’m already looking forward to pulling out Tetlock quotes like:

The superforecasters are a numerate bunch: many know about Bayes’ theorem and could deploy it if they felt it was worth the trouble. But they rarely crunch the numbers so explicitly. What matters far more to the superforecasters than Bayes’ theorem is Bayes’ core insight of gradually getting closer to the truth by constantly updating in proportion to the weight of the evidence. That’s true of Tim Minto [the top superforecaster]. He knows Bayes’ theorem, but he didn’t use it even once to make his hundreds of updated forecasts. And yet Minto appreciates the Bayesian spirit. “I think it is likely that I have a better intuitive grasp of Bayes’ theorem than most people,” he said, “even though if you asked me to write it down from memory I’d probably fail.” Minto is a Bayesian who does not use Bayes’ theorem. That paradoxical description applies to most superforecasters.

And if you’re interested, it looks like there’s a current version of the Good Judgment Program going on here that you can sign up to and see if you’re a superforecaster or not.

EDIT: A lot of people have asked the same question: am I being too dismissive? Isn’t it really important to have this book as evidence that these techniques work? Yes. It is important that the Good Judgment Project exists. But you might not want to read a three-hundred page book that explains lots of stuff like “Here’s what a cognitive bias is” just to hear that things work. If you already know what the techniques are, it might be quicker to read a study or a popular news article on GJP or something.

# List Of Passages I Highlighted In My Copy Of “Superforecasting”

In year 1, [the Good Judgment Project] beat the official control group by 60%. In year 2, we beat the control group by 78%. GJP also beat its university-affiliated competitors, including the Uniersity of Michigan and MIT, by hefty margins, from 30% to 70%, and even outperformed professional intellgience analysts with access to classified data. After two years, GJP was doing so much better than its academic competitors that IARPA dropped the other teams.

I keep wondering what these other teams were doing. Good Judgment Project sounds like it was doing the simplest, most obvious possible tactic – asking people to predict things and seeing what happened. David Manheim says the other groups tried “more straightforward wisdom of crowds” methods, so maybe GJP’s secret sauce was concentrating on the best people instead of on everyone? Still seems like it should have taken fewer than five universities and a branch of government to think of that.

One result that particularly surprised me was the effect of a tutorial covering some basic concepts that we’ll explore in this book and are summarized in the Ten Commandments appendix. It took only about sixty minutes to read and improved accuracy by roughly 10% through the entire tournament year. Yes, 10% may sound modest, but it was achieved at so little cost.

These Ten Commandments are available online here.

For centuries, [aversion to measuring things and collecting evidence] hobbled progress in medicine. When physicians finally accepted that their experience and perceptions were not reliable means of determining whether a treatment works, they turned to scientific testing – and medicine finally started to make rapid advances.

I see what Tetlock is trying to say here, but as written it’s horribly wrong.

Evidence-based medicine could be fairly described as starting in the 1970s with Cochrane’s first book, and really took off in the 80s and 90s. But this is also the period when rapid medical advances started slowing down! In my own field of psychiatry, the greatest advances were the first antidepressants and antipsychotics in the 50s, the benzodiazepines in the 60s, and then a gradual trickle of slightly upgraded versions of these through the 70s and 80s. The last new drugs that could be called “revolutionary” by any stretch of the imagination were probably the first SSRIs in the early 80s. This is the conventional wisdom of the field and everybody admits this, but I would add the stronger claim that the older medications in many ways work better. I know less about the history of other subfields, but they seem broadly similar – the really amazing discoveries are all pre-EBM, and the new drugs are mostly nicer streamlined versions of the old ones.

There’s an obvious “low-hanging fruit” argument to be made here, but some people (I think Michael Vassar sometimes toys with this idea) go further and say that evidence-based medicine as currently practiced can actually retard progress. In the old days, people tried possible new medications in a very free-form and fluid way that let everyone test their pet ideas quickly and keep the ones that worked; nowadays any potential innovations need $100 million 10-year multi-center trials which will only get funded in certain very specific situations. And in the old days, a drug would only be kept if it showed obvious undeniable improvement in patients, whereas nowadays if a trial shows a p < 0.05, d = 0.10 advantage, that's enough to make it the new standard if it's got a good pharma company behind it. So the old method allowed massive-scale innovation combined with high standards for success; the new method only allows very limited innovation but keeps everything that can show the slightest positive effect whatsoever on an easily-rigged but very expensive test.  
  
I'm not sure I believe in the strong version of this argument (the low-hanging fruit angle is probably sufficient), but the idea that medicine only started advancing after the discovery of evidence-based medicine is just wrong. A better way of phrasing it might be that around that time we started getting fewer innovations, but we also became a lot more effective and intelligent at using the innovations we already had.  
  
  
  
Consider Galen, the second-century physician to Rome’s emperors…Galen was untroubled by doubt. Each outcome confirmed he was right, no matter how equivocal the evidence might look to someone less wise than the master. “All who drink of this treatment recover in a short time, except those whom it does not help, who all die,” he wrote. “It is obvious, therefore, that it fails only in incurable cases.”

After hearing one too many “everyone thought Columbus would fall off the edge of the flat world” -style stories, I tend to be skeptical of “people in the past were hilariously stupid” anecdotes. I don’t know anything about Galen, but I wonder if this was really the whole story.

When hospitals created cardiac care units to treat patients recovering from heart attacks, Cochrane proposed a randomized trial to determine whether the new units delivered better results than the old treatment, which was to send the patient home for monitoring and bed rest. Physicians balked. It was obvious the cardiac care units were superior, they said, and denying patients the best care would be unethical. But Cochrane was not a man to back down…he got his trial: some patients, randomly selected, were sent to the cardiac care units while others were sent home for monitoring and bed rest. Partway through the trial, Cochrane met with a group of the cardiologists who had tried to stop his experiment. He told them that he had preliminary results. The difference in outcomes between the two treatments was not statistically signficant, he emphasized, but it appeared that patients might do slightly betteri n the cardiac care units. “They were vociferous in their abuse: ‘Archie,’ they said, ‘we always thought you were unethical. You must stop the trial at once.'” But then Cochrane revealed he had played a little trick. He had reversed the results: home care had done slightly better than the cardiac units. “There was dead silence and I felt rather sick because they were, after all, my medical colleagues.”

This story is the key to everything. See also my political spectrum quiz and the graph that inspired it. Almost nobody has consistent meta-level principles. Almost nobody really has opinions like “this study’s methodology is good enough to believe” or “if one group has a survival advantage of size X, that necessitates stopping the study as unethical”. The cardiologists sculpted their meta-level principles around what best supported their object-level opinions – that more cardiology is better – and so generated the meta-level principles “Cochrane’s experiment is accurate” and “if one group has a slight survival advantage, that’s all we need to know before ordering the experiment stopped as unethical.” If Cochrane had (truthfully) told them that the cardiology group was doing worse, they would have generated the meta-level principles “Cochrane’s experiment is flawed” and “if one group has a slight survival advantage that means nothing and it’s just a coincidence”. In some sense this is correct from a Bayesian point of view – I interpret sonar scans of Loch Ness that find no monsters to be probably accurate, but if a sonar scan did find a monster I’d wonder if it was a hoax – but in less obvious situations it can be a disaster. Cochrane understood this and so fed them the wrong data and let them sell him the rope he needed to hang them. I know no better solution to this except (possibly) adversarial collaboration. Also, I suppose this is more proof (as if we needed it) that cardiologists are evil.

In the late 1940s, the Communist government of Yugoslavia broke from the Soviet Union, raising fears that the Soviets would invade. In March 1951 [US intelligence under Sherman Kent reported there was a “serious possibility” of a Soviet attack.] But a few days later, Kent was chatting with a senior State Department official who casually asked, “By the way, what did you people mean by the expression ‘serious possibility’? What kind of odds did you have in mind?” Kent said he was pessimistic. He felt that the odds were about 65 to 35 in favor of an attack. The official was startled. He and his colleagues had taken “serious possibility” to mean much lower odds.

Disturbed, Kent went back to his team. They had all agreed to use “serious possibility” in the [report], so Kent asked each person, in turn, what he thought it meant. One analyst said it meant odds of about 80%. Another thought it meant odds of 20% – exactly the opposite. Other answers were scattered between those extremes. Kent was floored. A phrase that looked informative was so vague as to be almost useless…

In 1961, when the CIA was planning to topple the Castro government by landing a small army of Cuban expatriates at the Bay of Pigs, President John F. Kennedy turned to the military for an unbiased assessment. The Joint Chiefs of Staff concluded that the plan had a “fair chance” of success. The man who wrote the words “fair chance” later said he had in mind odds of 3 to 1 against. But Kennedy was never told precisely what “fair chance” meant and, not unreasonably, he took it to be a much more positive assessment.

…

Nate Silver, Princeton’s Sam Wang, and other poll aggregators were hailed for correctly predicting all fifty state outcomes, but almost no one noted that a crude, across-the-board prediction of “no change” – if a state went Democratic or Republican in 2008, it will do the same in 2012 – would have scored forty-eight out of fifty, which suggests that the many excited exclamations of “he called all fifty states!” we heard at the time were a tad overwrought.

I didn’t realize this. I think this election I’m going to predict the state-by-state results just so that I can tell people I “predicted 48 of the 50 states” or something and sound really impressive.

The [Expert Political Judgment] data revealed an inverse correlation between fame and accuracy: the more famous an expert was, the less accurate he was. That’s not because editors, producers, and the public go looking for bad forecasters. They go looking for hedgehogs, who just happen to be bad forecasters. Animated by a Big Idea, hedgehogs tell tight, simple, clear stories that grab and hold audiences.

One day aliens are going to discover humanity and be absolutely shocked we made it past the wooden-club stage.

In 2008, the Office of the Director of national Intelligence – which sits atop the entire network of sixteen intelligence agencies -asked the National research Council to form a committee. The task was to synthesize research on good judgment and help the IC put that research to good use. By Washington’s standards, it was a bold (or rash) thing to do. It’s not every day that a bureaucracy pays one of the world’s most respected scientific institutions to produce an objective report that might conclude that the bureaucracy was clueless.

This was a big theme of the book: the US intelligence community deserves celebration for daring to investigate its own competency at all. Interestingly, a lot of its investigations said it was doing things more right than we would think: Tetlock mentions that even independent-to-hostile investigators concluded that it had been correct in using the facts it had to believe Saddam had WMDs. The book didn’t explain exactly how this worked: possibly Saddam was trying to deceive everyone into thinking he had WMDs to prevent attacks, and did a good job? This was part of what got the intelligence community interested in probability: given that they had made a reasonable decision in saying there were WMDs, but it had been a big disaster for the United States, what could they have done differently? Their answer was “continue to make the reasonable decision, but learn to calibrate themselves well enough to admit there’s a big chance they’re wrong.”

[We finished by giving] the forecast a final tweak: “extremizing” it, meaning pushing it closer to 100% or zero. If the forecast is 70% you might bump it up to, say, 85%. If it’s 30%, you might reduce it to 15%…[it] is based on a pretty simple insight: when you combine the judgments of a large group of people to calculate the “wisdom of the crowd” you collect all of the relevant information that is dispersed among all those people. But none of those people has access to all that information…what would happen if every one of those people were given all the information? They would become more confident. If you then calculated the wisdom of the crowd, it too would be more extreme.

Something to remember if you’re doing wisdom-of-crowds with calibration estimates.

The correlation between how well individuals do from one year to the next is about 0.65…Regular forecasters scored higher on intelligence and knowledge tests than about 70% of the population. Superforecasters did better, placing higher than about 80% of the population.

People interested in taking these kinds of tests are generally intelligent; superforecasters are somewhat more, but not vastly more, intelligent than that.

Researchers have found that merely asking people to assume their initial judgment is wrong, to seriously consider why that might be, and then make another judgment, produces a second estimate which, when combined with the first, improves accuracy almost as much as getting a second estimate from another person.

There’s a rationalist tradition – I think it started with Mike and Alicorn – that before you get married, you ask all your friends to imagine that the marriage failed and tell you why. I guess if you just asked people “Will our marriage fail?” everyone would say no, either out of optimism or social desirability bias. If you ask “Assume our marriage failed and tell us why”, you’ll actually hear people’s concerns. I think this is the same principle. On the other hand, I’ve never heard of anyone trying this and deciding not to get married after all, so maybe we’re just going through the motions.

[Superforecaster] Doug Lorch knows that when people read for pleasure they naturally gravitate to the like-minded. So he created a database containing hundreds of information sources – from the New York Times to obscure blogs – that are tagged by their ideological oreintation, subject matter, and geographical origin, then wrote a program that selects what he should read next using criteria that maximize diversity.

Of all humans, only Doug Lorch is virtuous. Well, Doug Lorch and this guy from rationalist Tumblr who tried to get the program but was told it wasn’t really the sort of thing you could just copy and give someone.

[The CIA was advising Obama about whether Osama bin Laden was in Abbotabad, Pakistan; their estimates averaged around 70%]. “Okay, this is a probability thing,” the President said in response, according to Bowden’s account. Bowden editorializes: “Ever since the agency’s erroneous call a decade earlier [on Saddam’s weapons of mass destruction], the CIA had instituted an almost comically elaborate process for weighing certainty…it was like trying to controve a mathematical formula for good judgment.”Bowden was clearly not impressed with the CIA’s use of numbers and probabilities. Neither was Barack Obama, according to Bowden. “What you ended up with, as the president was finding, and as he would later explain to me, was not more certainty but more confusion…in this situation, what you started to get was probabilities that disguised uncertainty, as opposed to actually providing you with useful information…”

After listening to the widely ranging opinions, Obama addressed the rrom. “This is fifty-fifty,” he said. That silenced everyone. “Look guys, this is a flip of the coin. I can’t base this decision on the notion that we have any greater certainty than that…

The information Bowden provides is sketchy but it appears that the media estimate of the CIA officers – the “wisdom of the crowd” – was around 70%. And yet Obama declares the reality to be “fifty-fifty.” What does he mean by that?…Bowden’s account reminded me of an offhanded remark that Amos Tversky made some thirty years ago…In dealing with probabilities, he said, most people only have three settings: “gonna happen,” “not gonna happen,” and “maybe”.

Lest I make it look like Tetlock is being too unfair to Obama, he goes on to say that maybe he was speaking colloquially. But the way we speak colloquially says a lot about us, and there are many other examples of people saying this sort of thing and meaning it. This ties back into an old argument we had here on whether something like a Bayesian concept of probability was meaningful/useful. Some people said that it wasn’t, because everyone basically understands probability and Bayes doesn’t add much to that. I said it was, because people’s intuitive idea of probability is hopelessly confused and people don’t really think in probabilistic terms. I think we have no idea how confused most people’s idea of probability is, and perhaps even Obama, one of our more intellectual presidents, has some issues there.

Barbara Mellers has shown that granularity predicts accuracy: the average forecaster who sticks with the tens – 20%, 30%, 40% – is less accurate than the finer-grained forecaster who uses fives – 20%, 25%, 30% – and still less accurate than the even finer-grained forecaster who uses ones – 20%, 21%, 22%. As a further test, she rounded forecasts to make them less granular, so a forecast at the greatest granularity possible in the tournament, single percentage points, would be rounded to the nearest five, and then the nearest ten. This way, all of the forecasts were made one level less granular. She then recalculated Bier scores and discovered that superforecasters lost accuracy in response to even the smallest-scale rounding, to the nearest 0.05, whereas regular forecasters lost little even from rounding four times as large, to the nearest 0.2.

This was the part nobody on the comments to the last post believed, and I have trouble believing it too.

[There’s a famous Keynes quote: “When the facts change, I change my mind. What do you do, sir?”] It’s cited in countless books, including one written by me and another by my coauthor. Google it and you will find it’s all over the internet. Of all the many famous things Keynes says, it’s probably the most famous. But while researching this book, I tried to track it to its source and failed. Instead I found a post by a Wall Street Journal blogger, which said that no one has ever discovered its provenance and the two leading experts on Keynes think it is apocryphal. In light of these facts, and in the spirit of what Keynes apparently never said, I concluded that I was wrong.

The funny part is that if this fact is true, we’ve known it for fifty years, and people still haven’t changed their mind about whether he said it or not.

“Keynes is always ready to contradict not only his colleagues but also himself whenever circustancse make this seem appropriate,” re[prted a 1945 profile of the “consistently inconsistent” economist. “So far from feeling guilty about such reversals of position, he utilizes them as pretexts for rebukes to those he saw as less nimble-minded. Legend says that while conferring with Roosevelt at Quebec, Churchill sent Keynes a cable reading, ‘Am coming around to your point of view.’ His Lordship replied, ‘Sorry to hear it. Have started to change my mind.'”

I sympathize with this every time people email me to say how much they like the Non-Libertarian FAQ.

Police officers spend a lot of time figuring out who is telling the truth and who is lying, but research has found they aren’t nearly as good at it as they think they are and they tend not to get better with experience…predictably, psychologists who test police officers’ ability to spot lies in a controlled setting find a big gap between their confidence and their skill. And that gap grows as officers become more experienced and they assume, not unreasonably, that their experience has made them better lie detectors.

There’s some similar research on doctors and certain types of diagnostic tasks that don’t give quick feedback.

In 1988, when the Soviet Union was implementing major reforms that had people wondering about its future, I asked experts to estimate how likely it was that the Communist Party would lose its monopoly on power in the Soviet Union in the next five years. In 1991 the world watched in shock as the Soviet Union disintegrated. So in 1992-93 I retunred to the experts, reminded them of the question in 1988, and asked them to recall their estimates. On average, the experts recalled a number 31 percentage points higher than the correct figure. So an expert who thought there was only a 10% chance might remember herself thinking there was a 40% or 50% chance. There was even a case in which an expert who pegged the probability at 20% recalled it as 70%.

As the old saying goes, hindsight is 20/70.

The results were clear-cut each year. Teams of ordinary forecasters beat the wisdom of the crowd by about 10%. Prediction markets beat ordinary teams by about 20%. And superteams beat prediction markets by 15% to 30%. I can already hear the protests from my colleagues in finance that the only reason the superteams beat the prediction markets was that our markets lacked liquidity…they may be right. It is a testable idea, and one worth testing.

The correct way to phrase this is “if there is ever a large and liquid prediction market, Philip Tetlock will gather his superforecasters, beat the market, become a zillionaire, and then the market will be equal to or better than the forecasters.”

Orders in the Wehrmacht were often short and simple – even when history hung in the balance. “Gentlemen, I demand that your divisions completely cross the German borders, completely cross the Belgian borders, and completely cross the River Meuse,” a senior officer told the commanders who would launch the great assault into Belgium and France on May 10, 1940. “I don’t care how you do it, that’s completely up to you.”

This is the opposite of the image most people have of Germany’s World War II military. The Wehrmacht served a Nazi regime that rpeached total obedience to the dictates of the Fuhrer, and everyone emembers the old newsreels of German soldiers marching in goose-stepping unison…but what is often forgotten is that the Nazis did not create the Wehrmacht. They inherited it. And it could not have been more different from the unthinking machine we imagine.

[…]

Shortly after WWI, Eisenhower, then a junior officer who had some experience witht he new weapons called tanks, published an article in the US Army’s Infantry Journal making the modest argument that “the clumsy, awkward and snail-like progress of the old tanks must be forgotten, and in their place we must picture this speedy, reliable, and efficient engine of destruction.” Eisenhower was dressed down. “I was told my ideas were not only wrong but dangerous, and that henceforth I was to keep them to myself,” he recalled. “Particularly, I was not to publish anything incompatible with solid infantry doctrine. If I did, I would be hauled before a court martial.”

Tetlock includes a section on what makes good teams and organizations. He concludes that they’re effective when low-level members are given leeway both to pursue their own tasks as best they see fit, and to question and challenge their higher-ups. He contrasts the Wehrmacht, which was very good at this and overperformed its fundamentals in WWII, to the US Army, which was originally very bad at this and underperformed its fundamentals until it figured this out. Later in the chapter, he admits that his choice of examples might raise some eyebrows, but says that he did it on purpose to teach us to think critically and overcome cognitive dissonance between our moral preconceptions and our factual beliefs. I hope he has tenure.

Ultimately the Wehrmacht failed. In part, it was overwhelmed by its enemies’ superior resources. But it also made blunders – often because its commander-in-chief, Adolf Hitler, took direct control of operations in violation of Helmuth von Moltke’s principles, nowhere with more disastrous effect than during the invasion of Normandy. The Allies feared that after their troops landed, German tanks would drive them back to the beaches and into the sea, but Hitler had directed that the reserves could only move on his personal command. Hitler slept late. For hours after the Allies landed on the beaches, the dictator’s aides refused to wake him to ask if he wanted to order the tanks into battle.

Early to bed  
And early to stir up  
Makes a man healthy  
And ruler of Europe

The humility required for good judgment is not self-doubt – the sense that you are untalented, unintelligent, or unworthy. It is intellectual humility. It is a recognition that reality is profoundly complex, that seeing things clearly is a constant struggle, when it can be done at all, and that human judgment must therefore be riddled with mistakes. This is true for fools and geniuses alike. So it’s quite possible to think highly of yourself and be intellectually humble. In fact, this combination can be wonderfully fruitful. Intellectual humility compels the careful reflection necessary for good judgment; confidence in one’s abilities inspires determined action.

Yes! This is a really good explanation of Eliezer Yudkowsky’s Say It Loud.

(and that sentence would also have worked without the apostrophe or anything after it).

I am…optimistic that smart, dedicated people can inoculate themselves to some degree against certain cognitive illusions. That may sound like a tempest in an academic teapot, but it has real-world implications. If I am right, organizations will have more to gain from recruiting and training talented people to resist their biases.

This is probably a good time to mention that CFAR is hiring.

# Testimonials for SSC

[Content note: various slurs and insults]

I.

Last post I thanked some of the people who have contributed to this blog. But I forgot some of the most important contributors: the many readers whose give valuable feedback on everything I write.

So here’s a short sample of some of the feedback I’ve gotten over the past three years. I’m avoiding names and links to avoid pile-ons, but you can probably find most of these if you Google them.

II.

“It’s like someone tried to make fivethirtyeight as uninteresting as possible.”

“Slate Star Codex: 20,000 words on ‘feminism is bad’ and ‘Tom Swifties are the funniest shit I’ve ever seen'”

“Mark Atwood spent a month on the SSC registry of bans. It is my belief that Scott uses his toolbox of psychiatric techniques to manage the range of comment allowable on his blog. This may be justified in order to minimize flaming and trolling, but it is also a stifling form of censorship.”

“Go read the comment section on Slate Star Codex for a week and report back if you think LessWrongism is acceptable. It’s a place for broken people to be shielded from ever hearing that they’re broken and for developing better and better rationalizations for why they’re not broken and shouldn’t do the work needed to fix themselves. And SSC is miles better than lesswrong itself which is a weird cult centered around gnome in charge Eliezer Yudkowski. Just read the post where Scott let ozymandius post about all the ways that Roissy is wrong. The information there permanently disqualifies anyone associated with it from from having anything to do with building a functioning society.”

“Scott Alexander’s blog used to be good, but now he has been terrorized out of politics. Therefore boring. The problem was he purged all frequent commentors to the right of him out of the comments, which means that he had only enemies in his comments. And, being the rightmost, was persecuted. He has stopped posting on politics, I assume as a result of this persecution.”

“it’s by a stuttering aspie with expertise in nothing at all”

“a mentally-ill beta male who literally admitted that he wished he could become an asexual.”

“He is a fairly smart guy who makes well reasoned arguments. (He is also a literal cuckold.)”

“never forget for one fucking second that its author (who is ‘asexual’) and his most avid readers engage in ‘cuddle puddles’ irl, often bringing stuffed animals, and that he recommends this because it ‘increases credence’ in the other cuddlers’ statements.”

“his arguments only seem well-reasoned to people with NO knowledge of the subject matter (like him), and thus his main effect (just like that of his mentor, Eliezer Yudkowsky) is to keep smart people from learning things”

“oh what an expert in psychiatry! he’s a fucking med student in IRELAND. not to mention he uses yudkowsky’s lingo and called Less Wrong “revelatory” or something like that. you are dealing with a 110 IQ reddit type in SSC.”

“here is a series of a few posts (1, 2, 3) about how he is basically a conspiracy theorist, and in these posts he gets completely owned by the guy who cucked him with his tranny ex.”

“Merciful $DEITY. If I had any inclination to participate [on SSC], that [Guns and States] comment thread would have turned me completely off of it. How much more SJW-feminist-entitled can you get?…Extreme? No. Hard-line SJW enough that I’ve got better things to do than try to engage them? Yes. SJW-feminist-entitled? Yes.”

“SSC skews toward highly intelligent discourse, but Scott is very protective of his liberal homies. He will ban you if you stray too far from PC rigor.”

“I’d always got a whiff of fedora from this guy, so I feel gratified in my judgment at seeing him come out as one.”

“Faggot blocked me for calling out some recent bit of his retarded bullshit. Fuck ‘im.”

“Reading Slate Star Codex, I feel like I’m finally strarting to understand how postmodernism happened. First, there’s the whole thing of looking at your friends and a few books you happened to have read recently, and jumping to grand conclusions about all of society throughout all of history. Second, there’s the thing of him writing lengthy posts elaborating at great length on something that might be either boring and obviously true or bold and innovative but also completely wrong.”

“Aargh!! I read that entire Scott Alexander piece…well 65% of it…in earnest with the expectation that there was going to be a POINT to it. Some sort of payoff for my investment of time and attention. But there was nothing. It was just a bunch of bloviation with no purpose.”

“Scott Alexander is the story of a functioning pattern-recognition module trapped in a progressive brain. It would make a great story of its truth-seeking brain blob could eventually break free and rewire his brain to be a born-again reactionary. Not gonna happen though. The prog morality police has a hard, thick grasp on his brain, and all his friends and pseudosexual partners are the leftiest hacks this side of Lenin; so it’s an endless futile battle to square the circle. No wonder he went into psychiatry.”

“I retain great hopes for Scott, he’ll come around. When he does he’ll bring a high level of rigor with him. He is a caterpillar and will become a beautiful reactionary butterfly someday.”

“slatestarcodex is a great example of the difference between ‘knowing how to type’ and ‘knowing how to write'”

“You aren’t reading it right. Scott’s ability to completely identify the problem but still, quite sincerely, ritually abase himself to it at the same time, makes him worthy of connoisseurship. It takes a once in a generation talent to write long sincere \*thoughtful\* screeds pointing out that baby sacrifice is lowering the birth rate and causing family trauma, though of course he fully understands and endorses that Lord Moloch must be sated with the only food acceptable unto him.”

“Scott Alexander reminds me of some too-nice beta (nominally played by Joseph Gordon Leavitt or a JGL-alike) from some rom-com who’s trying to find his way and get the girl, while painfully oblivious to the fact that he just needs to stop being a too-nice beta and rip out somebody’s jugular. Ostracize someone for their beliefs? Me? Never. Golly gee.”

“seems like these guys are incapable of being dismissive of anything and have to objectively analyze everything.”

“Slate Star Codex is 140 IQ discussion about 105 IQ issues”

“He is sharp and makes good points but is way too fucking verbose. I dont need parts I II III and IV just fucking write concisely and stop vomiting words on your wordpress blog ”

“it’s basically a fish trap for aspies. people who can’t grasp nuance or understand basic human behavior, but are nonetheless obsessed with details and complex systems will inevitably gravitate toward this kind of horseshit. ultimately it’s a bunch of STEM-inclined dudes on the autism spectrum sitting around attempting to unpack societal problems like it was all a game of fucking sim city.”

“I would add that something like Slate Star Codex is also a clinic in the aspie tendency to miss the forest for the trees, except in this case it’s more like closely examining the bark on the trees for no goddamn reason whatsoever.”

“doesn’t this guy have a dayjob as like a doctor or something? why the fuck does he spend hours each day on a blog?” [to which another person on the same forum responded “why the fuck do you spend hours each day posting here?”]

“a blog populated by 99th percentile aspergers/IQ “rationalist” millennials who converse in an abnormally abstract style, and whose concrete cultural experience is drawn mainly from a bunch of weird nerd shit.”

“Its weird brand of reductionism and bizarre, arbitrary specificity plays to the types of spergy assholes and dumb know-it-all teenagers who don’t care about that anyway, or at least that’s how it seems to me. I mean, the ideas themselves seem like they’d be as much of a turn-off to regular people as their proponents’ personalities are, even if in a different way.”

“Oh, hey, the King of the Race Realist Misogynist Libertarian Nerds has Clever Things to say about vaccination.”

“yet another confirmation that: psychiatrists are crazier than their patients. polyamorous, diarrhea of the mouth/pen, math challenged, … i had no idea what an utter piece of shit you were.”

“He keeps his head down for fear of insulting permanently insulted people. He tries hard to be polite to people who hate him and consider him but a dog, unless they need him – and until they need him no longer. It is a waste of intellect, and debasement of character.”

“What makes me sad about Scott is just how close he is. I won’t give up hope on him yet. If only there was some way to secretly inject this guy with testosterone.”

“I wonder if he’s had bloodwork done to check his T count. I have to assume that if someone is an “asexual heteroromantic” as he puts it, that he’s interested in women from some abstract standpoint, and just needs some additional hormones to be thoroughly normal.”

“I literally want to see you kill yourself. I’m serious. You, and everyone else like you, are fucking disgusting wastes of space that are causing the decay of decency in the human race. I’m not going to argue with you, or say that it’s just my opinion or that it’s even up for debate.”

“is it some sort of special ‘Talk Like a Vulcan Day’ over there? Or are they always like that?”

“ssc spends a significant amount of time talking about stuff like how tables and chairs can be genders. he keeps a pretty unhinged tumblr”

“He’s definitely a beta orbiting cuckold.”

“that article seemed like a return ticket to obviousville with eight-hour layovers everywhere”

“That blog is very boring, and I didn’t manage to read long enough to find out what it was about. I hit Page Down a couple of times, and it seemed like it was on an entirely new topic each time.”

“I am thankful that I have never had any desire to seek psychiatric help. I have always had the impression, rightly or wrongly, that folks who pursue psychiatry as a career may themselves be the ones most in need of such therapy. Go to the mountains and look. Get up early and see the sunrise. Stop anywhere and take a minute to look at the beauty of nature all around you. We are a small piece in the universe, but still a part. The plan is good. You are fine. You will succeed if you try hard enough. Everything you need spiritually is inside you and has always been there. Stop complaining.”

“Also that ‘heteronormative asexual’ Scott Alexander. What a bizarre kike. He recently wrote that he’s incapable of not writing. LOL so kikish.”

“I thought it was a blog about science methodology until that post with the talking cactus.”

# Before You Get Too Excited About That GitHub Study…

Another day, another study purporting to find that Tech Is Sexist. Since it’s showing up here, you probably already guessed how this is going to end. Most of this analysis is not original to me – Hacker News had figured a lot of it out before I even woke up this morning – but I think it’ll at least be helpful to collect all the information in one easily linkable place.

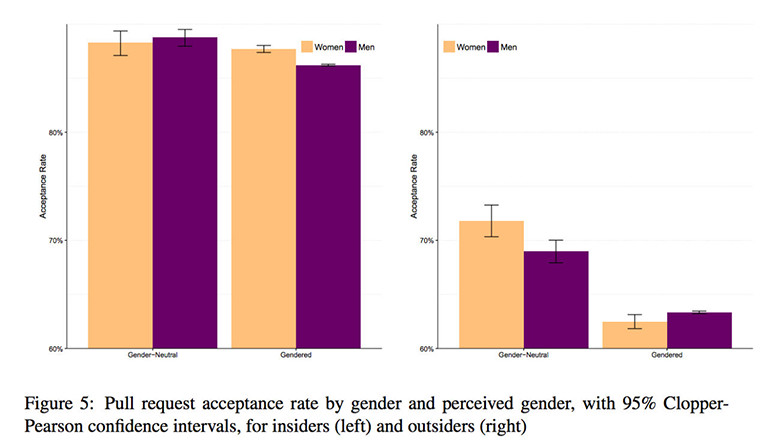
The study is Gender Bias In Open Source: Pull Request Acceptance Of Women Vs. Men. It’s a pretty neat idea: “pull requests” are discrete units of contribution to an open source project which are either accepted or rejected by the community, so just check which ones are submitted by men vs. women and whether one gender gets a higher acceptance rate than the other. This is a little harder than it sounds – people on GitHub use nicks that don’t always give gender cues – but the researchers wrote a program to automatically link contributor emails to Google Plus pages so they could figure out users’ genders.

This alone can’t rule out that one gender is genuinely doing something differently than another, so they had another neat trick: they wrote another program that automatically scored accounts on obvious gender cues: for example, somebody whose nickname was JaneSmith01, or somebody who had a photo of themselves on their profile. By comparing obviously gendered participants with non-obviously gendered participants whom the researchers had nevertheless been able to find the gender of, they should be able to tell whether there’s gender bias in request acceptances.

Because GitHub is big and their study is automated, they manage to get a really nice sample size – about 2.5 million pull requests by men and 150,000 by women.

They find that women get more (!) requests accepted than men for all of the top ten programming languages. They check some possible confounders – whether women make smaller changes (easier to get accepted) or whether their changes are more likely to serve an immediate project need (again, easier to get accepted) and in fact find the opposite – women’s changes are larger and less likely to serve project needs. That makes their better performance extra impressive.

So the big question is whether this changes based on obviousness of gender. The paper doesn’t give a lot of the analyses I want to see, and doesn’t make its data public, so we’ll have to go with the limited information they provide. They do not provide an analysis of the population as a whole (!) but they do give us a subgroup analysis by “insider status”, ie whether the person has contributed to that project before.



Among insiders, women do the same as men when gender is hidden, but better than men when gender is revealed. In other words, if you know somebody’s a woman, you’re more likely to approve her request than you would be on the merits alone. We can’t quantify exactly how much this is, because the paper doesn’t provide numbers, just graphs. Eyeballing the graph, it looks like being a woman gives you about a 1% advantage. I don’t see any discussion of this result, even though it’s half the study, and as far as I can tell the more statistically significant half.

Among outsiders, women do the same as/better than men when gender is hidden, and the same as/worse than men when gender is revealed. I can’t be more specific than this because the study doesn’t give numbers and I’m trying to eyeball confidence intervals on graphs. The study itself say that women do worse than men when gender is revealed, so since the researchers presumably have access to their real numbers data, that might mean the confidence intervals don’t overlap. From eyeballing the graph, it looks like the difference is 1% – ie, men get their requests approved 64% of the time, and women 63% of the time. Once again, it’s hard to tell by graph-eyeballing whether these two numbers are within each other’s confidence intervals.

The paper concludes that “for insiders…we see little evidence of bias…for outsiders, we see evidence of gender bias: women’s acceptance rates are 71.8% when they use gender neutral profiles, but drop to 62.5% when their gender is identifiable. There is a similar drop for men, but the effect is not as strong.”

In other words, they conclude there is gender bias among outsiders because obvious-women do worse than gender-anonymized-women. They admit that obvious-men also do worse than gender-anonymized men, but they ignore this effect because it’s smaller. They do not report doing a test of statistical significance on whether it is really smaller or not.

So:

1. Among insiders, women get more requests accepted than men.

2. Among insiders, people are biased towards women, that is, revealing genders gives women an advantage over men above and beyond the case where genders are hidden.

3. Among outsiders, women still get more requests accepted than men.

4. Among outsiders, revealing genders appears to show a bias against women. It’s not clear if this is statistically significant.

5. When all genders are revealed among outsiders, men appear to have their requests accepted at a rate of 64%, and women of 63%. The study does not provide enough information to determine whether this is statistically significant. Eyeballing it it looks like it might be, just barely.

6. The study describes its main finding as being that women have fewer requests approved when their gender is known. It hides on page 16 that men also have fewer requests approved when their gender is known. It describes the effect for women as larger, but does not report the size of the male effects, nor whether the difference is statistically significant. Eyeballing it, it looks about 2/3 the size of the female effect, and maybe?

7. The study has no hypothesis for why both sexes have fewer requests approved when their gender is known, without which it seems kind of hard to speculate about the significance of the phenomenon for one gender in particular. For example, suppose that the reason revealing gender decreases acceptance rates is because corporate contributors tend to use their (gendered) real names and non-corporate contributors tend to use handles like 133T\_HAXX0R. And suppose that the best people of all genders go to work at corporations, but a bigger percent of men go there than women. Then being non-gendered would be a higher sign of quality in a man than in a woman. This is obviously a silly just-so story, but my point is that without knowing why all genders show a decline after unblinding, it’s premature to speculate about why their declines are of different magnitudes – and it doesn’t take much to get so small a difference.

8. There’s no study-wide analysis, and no description of how many different subgroup analyses the study tried before settling on Insiders vs. Outsiders (nor how many different definitions of Insider vs. Outsider they tried). Remember, for every subgroup you try, you need to do a Bonferroni correction. This study does not do any Bonferroni corrections; given its already ambiguous confidence intervals, a proper correction would almost certainly destroy the finding.

9. We still have that result from before that women’s changes are larger and less likely to serve immediate needs, both of which make them less likely to be accepted. No attempt was made to control for this.

“Science” “journalism”, care to give a completely proportionate and reasonable response to this study?

Here’s Business Insider: Sexism Is Rampant Among Programmers On GitHub, Research Finds. “A new research report shows just how ridiculously tough it can be to be a woman programmer, especially in the very male-dominated world of open-source software….it also shows that women face a giant hurdle of “gender bias” when others assess their work. This research also helps explain the bigger problem: why so many women who do enter tech don’t stick around in it, and often move on to other industries within 10 years. Why bang your head against the wall for longer than a decade?” [EDIT: the title has since been changed]

Here’s Tech Times: Women Code Better Than Men But Only If They Hide Their Gender: “Interestingly enough, among users who were not well known in the coding community, coding suggestions from those whose profiles clearly stated that the users were women had a far lower acceptance rate than suggestions from those who did not make their gender known. What this means is that there is a bias against women in the coding world.” (Note the proportionate and reasonable use of the term “far lower acceptance rate” to refer to a female vs. male acceptance rate of, in the worst case, 63% vs. 64%.)

Here’s Vice.com: Women Are Better At Coding Than Men: “If feminism has taught us anything, it’s that almost all men are sexist. As this GitHub data shows, whether or not bros think that they view women as equals, women’s work is not being judged impartially. On the web, a vile male hive mind is running an assault mission against women in tech.”

This is normally the part at which I would question how a study got through peer review, but luckily this time there is a very simple answer: it didn’t. If you read the study, you may notice the giant red “NOT PEER-REVIEWED” sign on the top of every page. The paper was uploaded to a pre-peer-review site asking for comments. The authors appear to be undergraduate students.

I don’t blame the authors for doing a neat study and uploading it to a website. I do blame the entire world media up to and including the BBC for swallowing it uncritically. Note that two of the three news sources above failed to report that it is not peer-reviewed.

Oh, one more thing. A commenter on the paper’s pre-print asked for a breakdown by approver gender, and the authors mentioned that “Our analysis (not in this paper — we’ve cut a lot out to keep it crisp) shows that women are harder on other women than they are on men. Men are harder on other men than they are on women.”

Depending on what this means – since it was cut out of the paper to “keep it crisp”, we can’t be sure – it sounds like the effect is mainly from women rejecting other women’s contributions, and men being pretty accepting of them. Given the way the media predictably spun this paper, it is hard for me to conceive of a level of crispness which justifies not providing this information.

So, let’s review. A non-peer-reviewed paper shows that women get more requests accepted than men. In one subgroup, unblinding gender gives women a bigger advantage; in another subgroup, unblinding gender gives men a bigger advantage. When gender is unblinded, both men and women do worse; it’s unclear if there are statistically significant differences in this regard. Only one of the study’s subgroups showed lower acceptance for women than men, and the size of the difference was 63% vs. 64%, which may or may not be statistically significant. This may or may not be related to the fact, demonstrated in the study, that women propose bigger and less-immediately-useful changes on average; no attempt was made to control for this. This tiny amount of discrimination against women seems to be mostly from other women, not from men.

The media uses this to conclude that “a vile male hive mind is running an assault mission against women in tech.”

Every time I say I’m nervous about the institutionalized social justice movement, people tell me that I’m crazy, that I’m just sexist and privileged, and that feminism is merely the belief that women are people so any discomfort with it is totally beyond the pale. I would nevertheless like to re-emphasize my concerns at this point.

[EDIT: I don’t have much of a quarrel with the authors, who seem to have done an interesting study and are doing the correct thing by submitting it for peer review. I have a big quarrel with “science” “journalists” for the way they reported it. If any of the authors read this and want my peer review suggestions, I would recommend:

1. Report gender-unblinding results for the entire population before you get into the insiders-vs.-outsiders dichotomy.  
2. Give all numbers represented on graphs as actual numbers too.  
3. Declare how many different subgroup groupings you tried, and do appropriate Bonferroni corrections.  
4. Report the magnitude of the male drop vs. the female drop after gender-unblinding, test if they’re different, and report the test results.  
5. Add the part about men being harder on men and vice versa, give numbers, and do significance tests.  
6. Try to find an explanation for why both groups’ rates dropped with gender-unblinding. If you can’t, at least say so in the Discussion and propose some possibilities.  
7. Fix the way you present “Women’s acceptance rates are 71.8% when they use gender neutral profiles, but drop to 62.5% when their gender is identifiable”, at the very least by adding the comparable numbers about the similar drop for men in the same sentence. Otherwise this will be the heading for every single news article about the study and nobody will acknowledge that the drop for men exists at all. This will happen anyway no matter what you do, but at least it won’t be your fault.  
8. If possible, control for your finding that women’s changes are larger and less-needed and see how that affects results. If this sounds complicated, I bet you could find people here who are willing to help you.  
9. Please release an anonymized version of the data; it should be okay if you delete all identifiable information.]

# Nonfiction Writing Advice

People have asked me for advice on writing nonfiction online, so here are some tips:

1. Divide things into small chunks

Nobody likes walls of text. By this point most people know that you should have short, sweet paragraphs with line breaks between them. The shorter, the better. If you’re ever debating whether or not to end the paragraph and add a line break, err on the side of “yes”.

Once you understand this principle, you can generalize it to other aspects of your writing. For example, I stole the Last Psychiatrist’s style of section breaks – bold headers saying I., II., III., etc. Now instead of just paragraph breaks, you have two forms of break – paragraph break and section break. On some of my longest posts, including the Anti-Reactionary FAQ and Meditations on Moloch, I add a third level of break – in the first case, a supersection level in large fonts, in the latter, a subsection level with an underlined First, Second, etc. Again, if you’re ever debating more versus fewer breaks, err on the side of “more”.

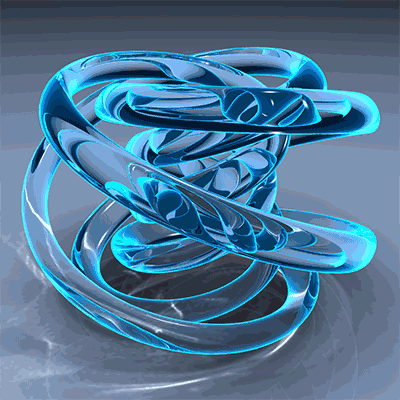
Finishing a paragraph or section gives people a micro-burst of accomplishment and reward. It helps them chunk the basic insight together and remember it for later. You want people to be going – “okay, insight, good, another insight, good, another insight, good” and then eventually you can tie all of the insights together into a high-level insight. Then you can start over, until eventually at the end you tie all of the high-level insights together. It’s nice and structured and easy to work with. If they’re just following a winding stream of thought wherever it’s going, it’ll take a lot more mental work and they’ll get bored and wander off.

Remember that clickbait comes from big media corporations optimizing for easy readability, and that the epitome of clickbait is the listicle. But the insight of the listicle applies even to much more sophisticated intellectual pieces – people are much happier to read a long thing if they can be tricked into thinking it’s a series of small things.

2. Variety is the spice of life

This is really closely linked to the last tip. Your brain gets bored if it has to focus on the same thing for too long. But you can get around that by making an activity look like many different things. Sometimes this is as simple and as dumb as putting Roman numeral one, Roman numeral two, etc at natural breaks in the article, and then your brain thinks “Oh, I guess there are two different things here”. But other times you actually have to vary the reading experience.

Again, the clickbaiters are our gurus – they intersperse images throughout their content. The images aren’t always very useful, they don’t always add much, but now it’s not just a wall of text. It’s a wall of text and images.



Watch the blue twirly thing until you forget how bored you are by this essay, then continue.

Or you can be more subtle. Break your flow. Include links, so that the never-ending stream of black text on white background is broken up with some pretty blue. If you are very desperate, italicize certain words to simulate the stresses of normal speech and turn the visual experience into a visual-auditory experience. Vary the form of your sentences, as per Gary Provost:

This sentence has five words. Here are five more words. Five-word sentences are fine. But several together become monotonous. Listen to what is happening. The writing is getting boring. The sound of it drones. It’s like a stuck record. The ear demands some variety. Now listen. I vary the sentence length, and I create music. Music. The writing sings. It has a pleasant rhythm, a lilt, a harmony. I use short sentences. And I use sentences of medium length. And sometimes, when I am certain the reader is rested, I will engage him with a sentence of considerable length, a sentence that burns with energy and builds with all the impetus of a crescendo, the roll of the drums, the crash of the cymbals–sounds that say listen to this, it is important.

(Blockquotes are also a nice way to vary the reading experience)

But don’t just vary the appearance of your writing. Vary the tone. If you’re comfortable, shift between registers. When I was talking about SSRIs, I mentioned study after study after study – and then, around the middle, I told a kind of funny story about the time I had a job interview with the author of one of the studies. It was a complete break with the tone of the piece, which is dangerous – but my hope was that after having your mind dulled by twenty different pharmacology studies in a row, a quick first-person aside and silly story would be invigorating and give you the energy to wade through another twenty such studies.

3. Keep your flow of ideas strong

I lampshade my flow of ideas with a lot of words like “Also”, “But”, “Nevertheless”, “Relatedly”, and “So” (when I’m feeling pretentious, also “Thus”). These are the words your eighth-grade English teacher told you never to start paragraphs with. Your eighth-grade English teacher was wrong. If you’re writing three paragraphs that are three different pieces of evidence for the same conclusion that you’re going to present afterwards, make damn sure your readers know this. It could be as simple as:

It’s pretty obvious that X is true, and we have lots of converging lines of evidence for this. Some of the best evidence comes from the field of augury. For example:

First, A

Second, B

Third, C

Now, some people say that not-A, but that’s totally wrong. It only looks like not-A, because P. Likewise, although Q might make it look like not-B, Q can’t be trusted for several other reasons, for example R. And not-C is too silly to even think about. So despite the objections you always hear, the augurical evidence for X is strong.

Even more evidence comes from the field of haruspicy. All four major haruspical schools hold X as a major principle. School 1 says X because of D. School 2 says X because of E. School 3 says X because of F. And school 4 says X because of G. So although augury and haruspicy disagree on a lot, on the subject of X they are in complete accord.

Notice the underlined words holding up the structure of the argument. Not only is the argument nice and tight, but the role of each part in the whole is telegraphed beforehand. For example, the “now” that comes just after C is saying something like “Take a step back, I’m about to tell you something that might otherwise be controversial, but listen to what I have to say”. And the “likewise” just after P means something like “We just got down talking about not-A because P, here’s another argument with about the same structure”. Before any of the facts are inserted, you already know where they fit into the structure. And you’re able to abstract from the micro-level and get the bigger picture of some fact which is supported by both augury and haruspicy, which was the main point of the argument.

I overuse the world “actually” really badly. I’m trying to cut down on it, but I don’t want to stop completely. “Actually” is a great structural word. It distinguishes “Here’s how things look, here’s what’s actually true”. That sentence makes sense even without the “actually”, but I feel like the “actually” holds my readers’ hands through the process and makes the dichotomy better-defined.

Defend your flow of ideas at all costs. This might sound paradoxical after section 2, which was about how breaking flows is great. It is kind of paradoxical, and it is sort of hard to explain, but it’s the difference between “exciting” and “horrible”. Eating a foreign cuisine can be exciting because it’s so different from your usual fare; eating hot lava is even more different than your usual fare, but no longer fun. Play around with flow and variety, but never break the flow in a jarring way. And if you have to break a flow, make it the flow of your sentence, or the flow of your paragraph, but not the flow of ideas.

I agonize a lot about where it is versus isn’t appropriate to break the flow of ideas. Sometimes I use the really ugly solution of having an entire paragraph within parentheses, as if to say “I really wanted to bring this up here, but remember it’s not actually part of the structure of this argument!”

(this is a good meta-level example. I used the word “actually” there, and I wanted to point it out as an example of what I was talking about before, but doing that would break the flow of this whole argument about how you shouldn’t break the flow of things. So, in accordance with the prophecy, into a paragraph-long set of parentheses it goes. I’m starting to think maybe I’m not the best person to be giving writing advice…)

But sometimes you’ve just got to leave out an observation which would be interesting and helpful but not at that particular part of your argument. I think the phrase is “kill your darlings”.

4. Learn what should and shouldn’t be repeated.

A lot of the medical notes I read look like this:

Mr. Smith presents to the ER for evaluation. He is a 24 year old man. He is complaining of chest pain. He was in the shower today when he slipped and fell. He was able to get up and make dinner. He says the chest pain started two hours later. He says has never had chest pain before. He took two aspirin. He says that did not help. He says the pain is 8/10 at this time. He says it is pretty bad, but that the pain of hearing these repetitive sentence structures is even worse.

If two sentences in a row start with the same word, it sounds unwieldy. If three or four do, it sounds bizarre. If it’s a whole paragraph’s worth, people start questioning their own sanity and trying to claw their eyes out.

A counterexample: what about the paragraph just above, starting with “If two sentences…”? I started with “if” three times in a row, and it didn’t sound bad at all! What’s up?

Deliberate use of parallelism is okay and even commendable. Usually this involves using the same structure to call attention to certain differences. You can tell if something is good parallelism by saying it aloud. When I say the paragraph above aloud, I’m using special intonation, especially in the places where the sentences differ (ie “two”, “three or four”, “whole paragraph’s worth”). Here your reader knows what you’re trying to do and it’s interesting. In the medical history example, there’s no deliberate attempt at parallelism in order to compare and contrast. You’re just doing the same thing again and again.

But it’s not just about first words of sentences. Consider something like this:

China has the largest population of any country in the world. It also has the largest military. Because of China’s powerful military, some of its neighbors are afraid of it. China has reassured its neighbors many times that it’s peaceful, but they’re not convinced.

This sounds off to me. The repetition of “largest population” and “largest military” is done clumsily. There are a lot of ways to make it a virtuous parallelism – for example “China has the largest population – and largest military – in the world” or “China has the largest population in the world; it also boasts the largest military” – but as it is, it just sounds weird. When you come to “largest military”, there’s an immediate mental callback to “largest population”, but you’re not sure why and it’s just distracting.

Likewise, the repetition of “neighbors” is weird. It could be solved by changing the second use to “those neighbors”, which sort of telegraphs that you know you’re repeating “neighbors” and did it on purpose. Otherwise it has the same unfortunate dull-sounding cadence as the medical history.

You could also solve both those problems by just varying the structure enough that the problem goes away. For example:

China is the most populous nation in the world. It also boasts the world’s largest military, which has provoked concern among other nations in the region. Although China has tried to reassure its neighbors of its peaceful intentions, they remain unconvinced.

This is hard and really deserves a book-length treatment. Without the book, all I can say is to realize that any repetition of words and structures will stand out to your reader, and make sure that their standing-out emphasizes your point instead of just being confusing.

5. Use microhumor

You’ve heard of microaggressions. Now try microhumor. It’s things that aren’t a joke in the laugh-out-loud told-by-a-comedian sense, but still put the tiniest ghost of a smile on your reader’s face while they’re skimming through them.

I learned this art from Dave Barry and Scott Adams, both of whom are humor writers and use normal macrohumor, but both of whom pepper the spaces in between jokes with microhumor besides. Your best best is to read everything they’ve written, your second best bet is to listen to me fumblingly try to explain it.

Here’s a paragraph from my “about” page:

Topics here tend to center vaguely around this meta-philosophical idea of how people evaluate arguments for their beliefs, and especially whether this process is spectacularly broken in a way that may or may not doom us all.

There are a couple of things here that might qualify as microhumor. Take “especially whether this process is spectacularly broken in a way that may or may not doom us all”. It’s not really a joke. If I were a comedian and recited that sentence, you wouldn’t start laughing. But it’s kind of funny to be starting with what sounds like a pretty dry academic idea (“how people evaluate arguments for their beliefs” and whether the process is broken), and then confound expectations with an exaggerated (well, maybe) warning about it dooming us all. The phrase “may or may not doom us all” does the same thing on a smaller scale: “may or may not” is a pretty reserved, careful sounding phrase, whereas “doom us all” is obviously the opposite of reserved (I also like the similar construction “it might have sort of kind of been the worst idea ever”).

You can actually go a long way toward microhumor just with hedge words (“vaguely”, “sort of”), exaggerations (“the worst thing ever”, “doom us all”), and sometimes the combination of the two.

I think this microhumor stuff is really important, maybe the number one thing that separates really enjoyable writers from people who are technically proficient but still a chore to read. Think about it with a really simplistic behaviorist model where you keep doing things that give you little bursts of reward, and stop doing things that don’t. There are only a couple of sources of reward in reading. One of them is getting important insights. Another is hearing things that support your ingroups or bash your outgroups. And a third – maybe the biggest – is humor. Who ever had trouble slogging through a really hilarious book of jokes?

Nobody can be super funny all the time, and an article on the economic crisis filled with man-walks-into-a-bar-style jokes would be jarring and weird. But micro-humor really works. It works at a background level where people don’t notice it working, and it makes people keep coming back for more.

Humor is also disarming. It’s hard to hate somebody who’s making you laugh. I don’t mean somebody who’s making bigoted jokes that offend you, or writing political cartoons about how awful your ingroup is. Those people are easy to hate. I mean somebody who’s making you laugh, right now. If you can make people laugh while challenging their cherished beliefs, you’ve got a tiny bit more good will than usual.

6. Use concrete examples

Consider the following proposition:

In a study measuring whether implicit attitudes determine an outcome, you need to make sure the implicit attitudes aren’t serving as accurate proxies for underlying fundamentals.

This is the thesis of one of my more popular posts, Perceptions Of Required Ability Act As A Proxy For Actual Required Ability, but I don’t present it like that. Instead, I start by saying:

Imagine a study with the following methodology. You survey a bunch of people to get their perceptions of who is a smoker (“97% of his close friends agree Bob smokes”). Then you correlate those numbers with who gets lung cancer. Your statistics program lights up like a Christmas tree with a bunch of super-strong correlations. You conclude “Perception of being a smoker causes lung cancer”, and make up a theory about how negative stereotypes of smokers cause stress which depresses the immune system. The media reports that as “Smoking Doesn’t Cause Cancer, Stereotypes Do”.

Whether or not you understood or agreed with the abstract version thesis, you (hopefully) find the problems with the nicotine example intuitively obvious. Now if I give you the principle “in a study measuring whether implicit attitudes determine an outcome, you need to make sure the implicit attitudes aren’t serving as accurate proxies for underlying fundamentals”, that principle makes sense and you will tend to agree with it. Now we can move on to harder problems, like the actual study in the post, where it’s not as obvious and where a lot of people thought they’d proven that the implicit attitude determined the outcome.

If you’re going to be making a complicated point, start with a concrete example. If you’re going to be making a very complicated point, start with a lot of concrete examples. When I wrote Meditations on Moloch, probably the most complicated point I’ve ever tried to express on this blog, I began with fourteen different examples before I even started trying to express the underlying principle. I hoped that readers would be able to triangulate my point by finding what all fourteen examples had in common, and most of them did.

This is related to an idea I keep stressing here, which is that people rarely have consistent meta-level principles. Instead, they’ll endorse the meta-level principle that supports their object-level beliefs at any given moment. The example I keep giving is how when the federal government was anti-gay, conservatives talked about the pressing need for federal intervention and liberals insisted on states’ rights; when the federal government became pro-gay, liberals talked about the pressing need for federal intervention and conservatives insisted on states’ rights.

So if you want to convince someone of a meta-level principle, you need to build it up from examples that support it. And if you want the principle to be well-founded and stable under reflective equilibrium, you also need to present the examples that don’t support it and explain why you didn’t make your principle out of those instead.

And if you want to convince somebody that their meta-level principle is wrong, the quickest and most effective way to do it is to show that it proves too much, then provide them with a better principle that preserves the things they want but doesn’t prove things they don’t want.

But my point is that all of this has to be done on the object-level, with the excursions to the meta-level level being few, far-between, and justified with extensive application to the object-level. Otherwise you’re too likely to shoot off into the entirely abstract and end up sounding like Hegel:

The good is the idea, or unity of the conception of the will with the particular will. Abstract right, well-being, the subjectivity of consciousness, and the contingency of external reality, are in their independent and separate existences superseded in this unity, although in their real essence they are contained in it and preserved. This unity is realized freedom, the absolute final cause of the world. Every stage is properly the idea, but the earlier steps contain the idea only in more abstract form. The I, as person, is already the idea, although in its most abstract guise. The good is the idea more completely determined; it is the unity of the conception of will with the particular will. It is not something abstractly right, but has a real content, whose substance constitutes both right and well-being.

Please don’t end up sounding like Hegel.

And a free tip for this: use words like “me” and “you” instead of “a person” or “someone”. Compare:

“If someone does the calculations with this methodology, the result will probably be nonsense.”

Versus:

“If you do the calculations with that methodology, you’ll probably end up with nonsense.”

I think the second sounds snappier and more concrete.

7. Figure out who you’re trying to convince, then use the right tribal signals

Your role model in this (and in nothing else) should be Donald Trump. Think about it. He supports Planned Parenthood, doesn’t want to cut entitlement programs, condemns Dubya and the Iraq war, supports affirmative action, supports medical marijuana, etc. If somebody were to tell you last year that a man with those policy positions would not only be leading the Republican primary, but leading even among the most conservative voters, you’d think they were crazy. The rest of the country has been trying to convince conservative Republicans to be more comfortable with those positions for decades, and we’ve failed miserably. Now Trump just waltzes in and everyone is like “Yeah, okay, sure”?

The secret of Trump’s success is that most conservative Republicans don’t really care about medical marijuana (or whatever) for its own sake. They care because opposing medical marijuana symbolizes membership in their tribe, they feel like their tribe is persecuted, they have a fierce loyalty to their tribe, and darned if they’re going to support somebody who doesn’t use the right shibboleths.

Trump throws them a bone. He says things like “illegal immigrants are rapists” that no moderate or liberal would ever say, things that would horrify them. He uses all the affectations of being working class. He may not quite prove he’s “one of us”, but he very effectively proves he’s not Just A Typical Outgroup Member. When Trump says “Legalize medical marijuana”, they don’t hear “I’m yet another RINO liberal pansy who hates Christian values and wants everybody to become reefer-smoking hippies”. So they only hear something boring about the regulations around pain relief medication – and who cares about those?

Trump’s Law is that if you want to convince people notorious for being unconvinceable, half the battle is using the right tribal signals to sound like you’re one of them.

For example, when I’m trying to convince conservatives, I veer my signaling way to the right. I started my defense of trigger warnings with “I complain a lot about the social justice movement”. Then I cited Jezebel and various Ethnic Studies professors being against trigger warnings. Then I tried to argue that trigger warnings actually go together well with strong versions of freedom of speech. At this point I haven’t even started arguing in favor of trigger warnings, I’ve just set up an unexpected terrain in which trigger warnings can be seen as a conservative thing supported by people who like free speech and don’t like social justice, and opposition to trigger warnings can be seen as the sort of very liberal thing that people like Jezebel and Ethnic Studies professors support. The important thing isn’t that I convince anyone that trigger warnings are really on the right – that’s a tall order – but that the rightists reading my argument feel like I’m working with them rather than against them. I’m not just another leftist saying “Support trigger warnings because it’s the leftist thing and you should be leftist and everyone on the right is terrible!”

My reward was seeing a bunch of hard-core anti-social-justice types trip over themselves in horror at actually being kind of convinced, which was pretty funny.

On the other hand, when I’m trying to convince feminists of something, I start with a trigger warning – partly because I genuinely believe it’s a good idea and those posts can be triggering, but also partly because starting with a trigger warning is a tribal signal that people on the right rarely use. It means that either I’m on their side, or I’m being unusually respectful to it. In this it’s a lot like Trump saying illegal immigrants are rapists – something the outgroup would never, ever do.

(And that’s not just my theory – I’ve gotten lots of angry comments about the trigger warnings from people further right than me, saying that using them makes me an idiot or a pushover or a cuck or something. I am always happy to get these comments, because it means the signaling value of using trigger warnings remains intact.)

Crossing tribal signaling boundaries is by far the most important persuasive technique I know, besides which none of the others even deserve to be called persuasive techniques at all. But to make it work, you have to actually understand the signals, and you have to have at least an ounce of honest sympathy for the other side. You can’t just be like “HELLO THERE, FELLOW LIBERALS! LET’S CREATE INTRUSIVE BIG GOVERNMENT AGENCIES TOGETHER! BUT BEFORE WE DO, I HAVE SOMETHING I WANT TO TELL YOU ABOUT THE SECOND AMENDMENT…”

Which I guess means that being able to consider both sides of an issue sort of gives you superpowers. That’s pretty encouraging.

8. Anticipate and defuse counterarguments

Here’s something I’ve noticed. Something like:

Alice: We need to invade Syria. I know that there’s always the risk of creating a Iraq-style power vacuum in these situations, but the threat from ISIS is too great.

Sounds a whole lot better than something like:

Alice: We need to invade Syria.  
Bob: But isn’t there a risk that will create a Iraq-style power vacuum?  
Alice: The threat from ISIS is too great.

The second one sounds too much like Alice hadn’t really thought about the power vacuum thing, Bob called her on it, and she kind of blew him off with a tangentially related point. The first one sounds more like Alice is a careful thinker who has weighed all the risks and benefits and finally decided in favor of invasion. This is true even though Alice’s reasoning is the same in both situations.

Or what about this:

Alice: We need to invade Syria. I know that there’s always the risk of creating a Iraq-style power vacuum in these situations, but the threat from ISIS is too great.  
Bob: I know the threat from ISIS is serious, but I’m still really worried about that power vacuum thing.

Bob sounds kind of weak here. Come on, Bob. Alice already raised the power vacuum issue! We’re done with that!

The moral of the story is that you sound a lot more credible, and your opponents a lot less persuasive, if you’re the one who brings the possible counterarguments up yourself. This is true regardless of how effective your countercounterarguments are.

There’s also a visibility advantage. Suppose Alice puts her argument on her blog. Bob quotes her and puts his counterargument on his blog. Maybe the readers of Bob’s blog won’t read Alice’s blog where they can see her countercounterargument. Maybe they don’t even read the comments of Bob’s blog. If Alice addressed the obvious counterarguments in her first post, Bob has been preempted from blogging about them separately, or at least has lost his low-hanging fruit and has to stretch further before he finds something he can talk about. And if he does quote Alice, all the countercounterarguments against his point will be right there for his readers.

This isn’t just good rhetorical practice, it’s good epistemic practice. A lot of Internet arguments are the same ten or twenty issues re-examined time after time after time after time. If you’re arguing in favor of gun control, you have no excuse not to realize somebody will bring up “But don’t guns save lives by helping people with self-defense?”. And in fact, if you’re arguing in favor of gun control, you had better have thought long and hard yourself about whether or not guns save lives through self-defense. If you’ve never considered that, you have no business having an opinion. But if you have considered and rejected that, then you might as well run your audience through your thought process now (and sound more convincing) now rather than wait until some pro-gun person brings it up (and be caught flat-footed like Alice in the second example).

The logical conclusion of this process is that you address all the arguments, counterarguments, and countercounterarguments in the space you’re covering, nobody can disagree with you, and you’re self-evidently right. Sometimes this takes a lot of text, but better a long argument which is accurate and convincing than a short snappy argument which might be wrong or unpersuasive. Besides, by this point you’ve absorbed all these other tips and hopefully write in an engaging way that makes your readers want to keep going no matter how many levels of countercountercountercounterargument you spring on them.

(Your other option here is just to put this stuff in footnotes and let your readers decide whether or not to go through them. It’s very satisfying to answer somebody’s stupid objection with “Actually, I think you’ll find I disproved that in footnote 17.”)

There’s a special variant of this you need when you’re in shark-infested waters, debating very controversial things with very hostile people. Here the “counterargument” is going to take the form of trying to destroy your reputation by using one of your comments, taken out of context, to prove you’re a bad person with unconscionable beliefs who should never be listened to.

For example, suppose I’m trying to explain some social phenomenon and I mention that rich people get better grades in school than poor people. A hostile opponent could accuse me of making a stupid stereotype and saying that all rich people are better than all poor people; then he could condescendingly “correct” me by saying that actually the within-class differences are larger than the between-class ones. Or might say that I need to realize school grades aren’t the only thing and there are much more important determinants of people’s worth as a human being. Or he might accuse me of being a Social Darwinist, and “correct” me by saying that maybe this is because of the stresses of poverty.

Now, in fact I neither said nor meant any of those things. But if somebody accuses me of them, and I have to plead that I really didn’t mean it that way, honest – then they can double down and say that my protests of innocence are the surest sign of my guilt. Whether they succeed or not, I’m on the defensive. We’ve shifted from debating whatever point I wanted to make in the first place, to debating whether Scott is a monocle-wearing Social Darwinist who believes all poor people deserve to starve on the street.

The solution is really simple: anticipate and defuse counterarguments. If I wanted to make the class/grades point, it would go something like this:

According to [study], students from families earning >$100,000 score have an average high school GPA X points higher than students from families earning < $20,000. This obviously doesn't mean all rich people do better than all poor people, and it doesn't rule out causes like the stress of poverty hurting people's brains. But it does suggest a possible explanation for [whatever phenomenon I was trying to use this to explain].

A related note: when talking about controversial things to a potentially hostile audience, look through every single sentence of your work and imagine how it would sound if it were quoted out of context and used as a summary of who you are as a human being. If you don’t, eventually someone will try this and you’ll be unprepared.

9. Use strong concept handles

The idea of concept-handles is itself a concept-handle; it means a catchy phrase that sums up a complex topic.

Eliezer Yudkowsky is really good at this. “belief in belief“, “semantic stopsigns“, “applause lights“, “Pascal’s mugging“, “adaptation-executors vs. fitness-maximizers“, “reversed stupidity vs. intelligence“, “joy in the merely real” – all of these are interesting ideas, but more important they’re interesting ideas with short catchy names that everybody knows, so we can talk about them easily.

I have very consciously tried to emulate that when talking about ideas like trivial inconveniences, meta-contrarianism, toxoplasma, and Moloch.

I would go even further and say that this is one of the most important things a blog like this can do. I’m not too likely to discover some entirely new social phenomenon that nobody’s ever thought about before. But there are a lot of things people have vague nebulous ideas about that they can’t quite put into words. Changing those into crystal-clear ideas they can manipulate and discuss with others is a big deal.

If you figure out something interesting and very briefly cram it into somebody else’s head, don’t waste that! Give it a nice concept-handle so that they’ll remember it and be able to use it to solve other problems!

10. Recognize that applying these rules will probably start disastrously

There’s a pattern across almost all skills, where people start off doing things half-baked but sometimes with a bit of native talent. The experts teach them The Right Way To Do Things, and they switch to doing it in a stilted formulaic way that makes everybody else cringe. Eventually they become better and better. Finally, they do things that completely contradict the rules they were taught, and it works great. I think it was in the context of poetry that somebody said “Learn the rules first, then you can break them as much as you want.”

Untrained natural writing is often bad, but at least honestly bad. Untrained writing that tries to force itself to do something because somebody told them it was a good idea is much worse. Think of the old adage “If you’re giving a speech, start out with a joke.” It’s great advice when done right. Now imagine all the ways it could go wrong – terrible jokes, inappropriate jokes, forced jokes, speeches that absolutely shouldn’t start off with jokes, et cetera. A speech that doesn’t start off with a joke is often good; a speech that shouldn’t start out with a joke but has been forced into doing so never is. Eventually you end up shouting “Just use your instincts!” at people who do not actually have instincts.

Use the advice in this post wrong, and you end up transforming the famous quote from the Declaration of Independence into something like:

Although we agree King George has many good qualities, we nevertheless hold these truths to be more or less self-evident. Truth number one, that all men are created equal. For example, they should all be equally allowed to speak freely about important issues like taxes. It’s possible that there might be some times they shouldn’t be equal, like children having fewer rights than adults, but this are just minor exceptions. [insert picture of Liberty Bell here] Truth number two, that they are endowed by their Creator with certain inalienable rights…

Almost the only good advice in any discipline is “develop instincts, then use them”. While you’re waiting for the instincts to develop, or in order to push them along, it’s sometimes helpful to hear some other people’s advice. But do. not. force. it.

George Orwell ended his own list of writing advice with “Break any of these rules rather than say something outright barbarous”. As usual, George Orwell is right.

I also like Piet Hein’s commentary:

There is  
one art,  
no more,  
no less:  
to do  
all things  
with art-  
lessness.

# Two Attitudes In Psychiatry

Attitude 1 says that patients know what they want but not necessarily how to get it, and psychiatrists are there to advise them. So a patient might say “I want to stop being depressed”, and their psychiatrist might recommend them an antidepressant drug, or a therapy that works against depression. This is nice and straightforward and tends to make patients very happy.

Attitude 2 says that people are complicated. Sometimes this complexity makes them mentally ill, and sometimes it makes them come to psychiatrists and ask for help, but there’s no guarantee that the thing that they’re asking about is actually the problem. In order to solve the problem, you need to unravel the complexity, and that might involve not giving the patient what they want, or giving them things they don’t want. This is not straightforward and requires some justification, so let me give a few cases where Attitude 2 seems to me obviously correct.

1. A mother brings her 6 year old son to the doctor, complaining that he gets nauseous every morning. She wants the doctor to prescribe an anti-nausea pill. The doctor probes further and finds the kid only gets nauseous on school days. In fact, he only gets nauseous on school days when he has a particular gym class. The doctor asks the kid if there are any problems in that gym class, and the kid is reluctant to say anything. After a while, he finally admits there is a bully in that class. The mother calls the school, and the school takes care of the bully. After that the kid is no longer nauseous in the mornings.

2. A woman goes to a plastic surgeon asking him to fix her nose, which she insists is hideously deformed. The plastic surgeon thinks the nose looks perfectly normal and asks her to be cleared by a psychiatrist before surgery. The psychiatrist diagnoses the woman with body dysmorphic disorder, a delusional belief that one of their body parts is unbearably ugly. The psychiatrist advises the woman and her surgeon that plastic surgery does not work for this disease; if the woman gets her operation, she’ll inevitably either think that the new nose is just as ugly as the old one, or she’ll switch to focusing on something else like her ears or her mouth. He suggests she get psychotherapy instead. After several years of psychotherapy, the woman learns not to worry so much about her nose.

3. A woman goes to her doctor asking him how to taper off her birth control pills. The doctor is surprised at this request because he knows she is planning to break up with her boyfriend. The woman says that this is true, but she wants a child as a way to remember the relationship. The doctor probes deeper and finds the patient is very anxious and ambivalent about leaving her boyfriend and feels like if she has his child at least she will always have “a part of him” with her. The doctor refers her to therapy for her anxiety, and she is able to sort through her conflicting feelings about leaving her boyfriend. She chooses to stay on her birth control.

4. A man goes to his doctor asking for the strongest antipsychotics that exist, saying that he’s crazy and he’s going to hurt someone. The man converses very logically, and tells the doctor he’s felt like this for a year now and never hurt anybody. The doctor suggests that he’s not actually psychotic or violent, but might have an obsessive-compulsive disorder where he worries about becoming that way. The doctor recommends therapy for OCD.

5. A man comes to the psychiatric hospital saying that he’s suicidal and needs admission. The doctor knows him well, and remembers that he has been admitted five times in the past six months, each time after a life crisis, and that the patient has never actually attempted suicide and never even planned how he might do it. The doctor suggests that the man is using the psych hospital as an emotional crutch, and that instead of threatening suicide and going to the hospital whenever he is upset, he needs to learn more adaptive coping mechanisms.

Attitude 1 would have been the wrong choice in these five situations. If the doctor had just given the mother the anti-nausea pill she’d been asking for, the son’s stress about being bullied probably would have just caused some other symptoms. If the surgeon had just given the woman the nose job she wanted, she would have been dissatisfied with the surgery and wanted it changed again. If the third doctor had just told the woman how to get off birth control like she wanted, she might have had a baby for the wrong reasons and regretted it later, leading to heartache all around. If the fourth doctor had just given the man an antipsychotic, he would have unnecessarily exposed him to a potentially life-long course of very strong medication. If the fifth doctor had admitted the man to the hospital, he would be using up scarce resources and discouraging the man from learning better coping strategies.

Any halfway decent psychiatrist uses both attitudes at different times, but most people I know tend to lean to one side or the other. The 2-leaning doctors stereotype the 1-leaning doctors as simple-minded and gullible. The 1-leaning doctors stereotype the 2-leaning doctors as antirational paranoiacs with sledgehammers.

I remember a textbook talking about a case study by a famous psychiatrist. The patient had come in talking about how her husband was being borderline-emotionally-abusive to her. The psychiatrist interrupted her and said that she was perpetuating this dynamic to feed her own narcissism. The patient said this was absolutely not true and she wasn’t narcissistic. The psychiatrist said she would never be able to get over her provoking-her-husband problem until she admitted the depth of her narcissism. The patient refused to keep seeing the psychiatrist after that, and the psychiatrist commented that it had been a hopeless case from the beginning – the extent of her narcissism was so great that she would never acknowledge that somebody else might know more than she did.

And the textbook was very wishy-washy about this – it acknowledged that the famous psychiatrist was brilliant and was doing the right thing in trying to confront the woman with evidence for her narcissism, but then it said that maybe he should have taken a more compassionate tone. Meanwhile, I couldn’t help thinking that the famous psychiatrist was a jerk, that his only evidence the woman was narcissistic at all was a snap judgment from one or two easily misinterpretable things she said, and that call me narcissistic if you want but I wouldn’t have kept attending therapy with this guy either.

That right there is the failure mode of Attitude 2; when we get out of the perfectly safe cases I mentioned above and into the more extreme versions, it starts looking a lot like making snap judgments about how all of a patient’s problems reduce to a single personality flaw, and then interpreting everything about the patient in that light. Narcissism is probably the most popular, but other such flaws include “patient is regressing and wants to act like a child and have other people take care of her”, “patient is just looking for attention”, and “patient is obsessive and demands complete control over everything”. The problem is, once you make one of these judgments every possible piece of data becomes further confirmation.

For example, suppose that a patient says he is having side effects on his new medication.

If you already believe the patient is a narcissist, you can dismiss the patient by saying that he wants to be special, he’s not happy on the same medication as everyone else, he’s trying to control the interaction by making you feel bad because you gave him an inferior medication. The solution is to teach the patient that he can’t always have his way by continuing the medication.

If you already believe the patient is regressing, you can dismiss the patient by saying that he’s throwing a temper tantrum, that instead of dealing with the side effects like a mature adult he wants someone else to step in and make everything magically better. The solution is to teach the patient to deal with his own problems by continuing the medication.

If you already believe that the patient is looking for attention, you can dismiss the patient by saying that they’re just trying to get the doctor’s attention by complaining. You can teach them that this is a maladptive social strategy by continuing the medication.

If you already believe that the patient is obsessive, you can dismiss the patient by saying that he’s getting all neurotic over minor side effects and has worked himself into a frenzy over perfectly ordinary minor hiccups because he can’t tolerate anxiety. The solution is to reassure the patient that everything is fine and continue the medication.

If you already believe that the patient is a witch, you can dismiss the patient by saying that they’re trying to confuse and upset you so that you will be easy prey when they try to kidnap you and sacrifice you to their lord and master, the Devil.

So it’s pretty easy to dislike 2-leaning doctors. Also, fun. Also, quite often justified. So sometimes I give in to the urge and dislike them.

The problem is, sometimes they’re right. I remember one time I had a patient who complained that Geodon was making her hallucinate. Geodon is an anti-hallucination medicine, so the chance that it makes someone hallucinate is pretty slim – but I’ve read all the usual social media posts where people complain about their evil psychiatrist who just dismisses their deeply felt pain as fakery because they had a problem that wasn’t listed in the textbook, and I didn’t want to be that guy, so I went along with it. I asked her to take some Geodon right there in my office. She swallowed the Geodon pill, and sure enough, about two minutes later she said she was starting to have all of these terrible hallucinations.

So I explained to her that oral Geodon takes at least an hour or two before a reasonable amount gets into the bloodstream, and there was no biological way that it could cause hallucinations two minutes after she took it. Then we talked about why she might be scared of the Geodon, and whether she felt any ambivalence about really wanting to get better. Eventually she agreed to try the Geodon again and didn’t hallucinate any more.

Here I felt okay because I had biological impossibility on my side. But I always wonder how many cases I’m letting slip just because my patients’ stories are merely possible-but-unlikely.

Everything’s a tradeoff between Type I and Type II errors. If I err too far on the side of Attitude 1, then my patients will like me and I’ll never inspire a “my doctor said I was just making up my side effects for attention, and later on I got neuroleptic malignant syndrome and died!” horror story. But I will occasionally be doing the equivalent of doing plastic surgery on a body dysmorphic disorder patient, giving unnecessary and harmful medical care while ignoring the true problem.

If I err too far on the side of Attitude 2, then I always get to feel like a hard-headed non-gullible investigator digging down to the root of the problem – but occasionally I’ll end up like that famous psychiatrist in the textbook and tell people that the reason their foot hurts is because they’re narcissistic, and it has nothing to do with the fact that they stepped on a nail and the only reason they’re even bringing up the nail is their deep-seated narcissism.

I tend to lean way toward Attitude 1. I’m not sure I can justify it. Part of it is my personality: conflict scares me and I want to be liked. Part of it is that I read too many horror stories on social media about how much patients hate their Attitude 2 psychiatrists. Part of it is that Attitude 2 has a lot of its philosophical grounding in Freud, and I really don’t trust Freud.

This is a lucrative attitude nowadays. We are all supposed to be biological psychiatrists, all the old psycho-babble is no longer covered by insurance, and The Customer Is Always Right. I am lucky insofar as my natural tendency is also the socially more acceptable one.

(I suppose an Attitude 2 psychiatrist would say I’m not lucky at all, and that my unconscious desire for social approval and success has led me to adopt Attitude 1, and also I am a narcissist)

I may or may not be a narcissist, but I am definitely neurotic. And when my neurosis gets to “maybe I’m a terrible psychiatrist”, this is what it usually settles upon to worry about. Attitude 2 and the various arts associated with it are opaque to me. I can pass tests on them when I have to, but I don’t feel them in my bones. When I’m with a whole conference of doctors nodding their head and going “Yup, that guy’s a narcissist”, I’m always panicking, thinking “Wait, I’m not even close to convinced he’s a narcissist, and also nobody really knows how to treat narcissism, and I would feel a lot more comfortable if this conversation would shift to comparing and contrasting the various subtypes of dopamine receptors.” I am bad at it, and what’s worse I don’t even know if I should be better at it, and I don’t know how to solve the bad-at-it part without worrying that I’m sending myself and my patients down a blind alley.

# Against Interminable Arguments

[Epistemic status: something I’ve been thinking about recently. There’s a lot of complication around these issues and this is more to start a discussion than to present any settled solution]

There’s a scene in Fiddler on the Roof where Tevye is describing his peaceful little town. He says they never fight – except that one time about a horse some people thought was a mule. Someone interrupts him to say it was really a mule some people thought was a horse, and then everyone in town starts shouting “MULE!” or “HORSE!” at each other until they get drowned out by the chorus.

The town is happy and peaceful as long as nobody brings up the horse/mule thing. As soon as somebody brings it up all of the old rancor instantly resurfaces and everybody’s at each other’s throats. And the argument itself never gets more sophisticated than people yelling “HORSE!” or “MULE!” at each other. Maybe it would be worth it to create a norm around never bringing it up?

The rationalist/EA/etc community has a norm that people must be able to defend their beliefs with evidence, and a further norm that people shouldn’t be confident in their beliefs unless they’ve sounded them off others and sought out potential counterarguments. These are great norms. But their failure mode is a community where dredging up interminable horse/mule style arguments is seen as a virtue, and avoiding them is seen as a cowardly refusal to expose one’s own beliefs to challenge.

I’ve been thinking about this in the context of some arguments that keep cropping up on rationalist Tumblr. These have gotten so repetitive and annoying that I made a joke argument calendar to shame us for it; some of the people who replied were less able to see the humor in it and thought we really should be ashamed.

After some thought, I agree. Avoiding interminable arguments is an important social engineering problem we’re really bad at. Part of it is that we need a way to distinguish the baby from the bathwater. What does it mean to seek out productive discussion while avoiding interminable arguments?

A while back I wrote about memetic evolution for controversy. If an idea is outrageous, it’s likely to spread as people condemn it. if something is controversially outrageous, it’s likely to spread even further, as people argue against it and then other people counterargue against the arguments ad infinitum. This gives it visibility far beyond things that are objectively more important. The entire news media freaks out over BernieBros for weeks, and nobody ever hears about desalinization policy in drought-stricken nations.

Interminable arguments are the local version of the same process. For me, the biggest difference between a productive discussion and an interminable argument is simple. I’m participating in the productive discussion because I want to. I’m participating in the interminable argument because I have to.

I mean, obviously nobody’s pointing a gun to my head and forcing me into any arguments. But there are a lot of reasons I might feel obligated to debate something I really don’t have time for.

First, there’s the feeling XKCD describes as “someone is wrong on the Internet”. When somebody makes a deeply flawed argument against a position you hold dear, and everyone else seems to believe them, it creates (at least for me) this weird irresistable urge to correct them.

Second, sometimes people are jerks. Nobody on Tumblr can just say “I don’t think AI is a big problem”. They have to say “I don’t think AI is a big problem, and the only reason some people worry about it is because they’re cultish sci-fi dorks who are too brainwashed to look up what real scientists have to say”. Nobody on Tumblr can just say they think feminism is important; they have to post comics like this. This is hard to just leave be, especially when it’s not just yourself but your friends who are being insulted, and especially when a lot of people are vocally agreeing with them and even some people you think should know better are being convinced. Letting jerks have the last word is really hard.

Third, sometimes there are actually things at stake. I’ve written before on how one of the main reasons I get defensive is because I think some groups actively strategize to push their opponents out of the Overton Window and turn them into despised laughingstocks. When it works, it means I either have to be a despised laughingstock or spend way too much mental energy hiding my true opinions. The alternative to letting these people have the final say is defending one’s self.

When I don’t want to argue but feel forced into it, I’m doing a very different thing than when I’m having a voluntary productive discussion. I’m a lot less likely to change my views or admit subtlety, because that contradicts my whole point in having the argument. And I’m a lot more likely to be hostile, because hostility is about making other people feel bad and disincentivizing their behavior, and in this case I really do believe their behavior needs disincentivizing.

But of course this just starts the cycle where people who disagree with me – both the original people I’m arguing against, and bystanders who just happen to hear – feel forced to write nasty replies of their own. And so on. It only dies down once everybody can tell themselves that they put enough effort into self-defense to acquit themselves well. And as soon as somebody challenges that – like Tevye in the Fiddler song – the whole thing starts up again as bad as ever.

So how does a community prevent this? Blocking jerks – the people who start the whole cycle by deliberately trolling others – is an obvious good start. But some people on the Tumblr discussion have mentioned some more subtle points worth thinking about.

First, the influx of newbies is a big driver of this dynamic. Newbies are less likely to know the relevant arguments, won’t be bored of them yet, won’t want to steer clear of them, and may mistake somebody’s unwillingness to engage for the 9000th time as unwillingness to engage at all. People should be more tolerant of newbies, and newbies should be more tolerant of “look in the archives for the last time we discussed this issue, but seriously, don’t start up about this again”. This is what I think social justice people mean when they talk about “this is not a 101 space”.

Second, be aware that some problems with interminable arguments might be asymmetrical. I’ve heard this most often in social justice contexts. For example, cis people might never have discussed trans issues much before and might find them really interesting and not particularly defensiveness-inducing. But consider what happens if there are a hundred times more cis people than trans people, plus trans people have spent a hundred times longer thinking about it, plus trans people have a lot less tolerance before they get annoyed. A cis person might innocently ask “Hey, isn’t using chromosomal sex as a proxy for gender a pretty elegant system which is much easier than all of this stuff about different identities?” because they really want to know and their local trans person seems like a good person to debate with. Meanwhile, the trans person might have had this exact debate two thousand times, find it personally insulting, but not know how to disengage politely without giving the impression that they’re too much of an intellectual lightweight to answer the simplest of questions about something very important to them.

(This isn’t limited to social justice or identity politics. I feel this same way sometimes on relatively transhumanist-free parts of the Internet. Someone will make a well-intentioned attempt to start a discussion, like “But what you people don’t get is that the AI will be smart enough to realize that paper clips are silly and compassion towards all living things is the best goal,” and I’ll want to say something like “Look, I promise that in fifty years of thinking about machine ethics somebody else has raised that point before, I think you’re so confused about things that correcting you wouldn’t be a very good use of my time, but if you google anything at all in this general area you’ll probably find an answer to your question.” But that would just be the “it’s not my job to educate you” or “read the Sequences” answer which so many people find annoying.)

But third, as I said before, make it absolutely super-duper crystal clear that there is not a community norm that everybody has to defend their positions every time they are asked, do not say anything like “I’ve never heard you response to Point X so now I’m going to assume that you have no argument against it and are just a brainwashed cultist taking your position on faith” and emphasize that although everybody who wants to be accurate should discuss and challenge their beliefs, nobody should have to do it to anyone else’s schedule.

The exact way this works is something I’m still working on. More and more I’m abandoning the idea of debating on social media/comments/forums entirely, and switching either to email or longer forms. Email is private and removes the performative factor; I can’t say how many times a previously terrible discussion has become manageable and productive as soon as it gets outside the public eye. And by longer forms, I mean things like books and (really good long form) blog posts. I would much rather read the best book by someone I disagree with, and hear all their best arguments laid out by a leading intellectual with a good editor, rather than have to listen to somebody taunt me on Tumblr. And if I don’t understand something about the book, or I still have questions, then I can pick one or two people I know and debate it with them privately.

Part of me thinks this is another point in favor of niceness. A community made mostly of nice people can probably hold more productive debates and have fewer interminable arguments than one that’s not as good at civility. On the other hand, I see these problems even among nice people. I think that the skill of structuring a point such that nobody feels attacked by it is complicated in ways beyond regular niceness, and that otherwise nice communities can sometimes make the right choice in deciding to avoid the whole issue.

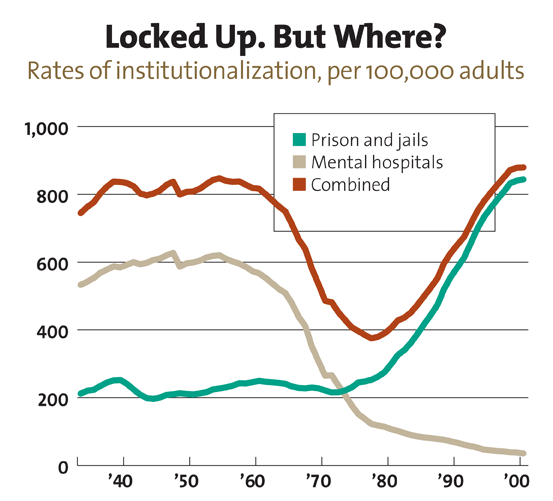
EDIT: comment

# Reverse Voxsplaining: Prison and Mental Illness

I.

German Lopez of Vox writes that “America’s criminal justice system has in many ways become a substitute for the US’ largely gutted mental health system”.

He says that starting in the 1970s the US “began locking up a lot more people”, and “at the same time, the country pulled back and defunded its public mental health system”. He admits that “this wasn’t, at the time, totally malicious”, but then says it “left the criminal justice system as the only system that can respond to people with mental illness.”



He concludes that as a result, “the number of people with mental illness in prisons/jails outnumber those in state hospitals 10 to 1.” The apparent (though unstated) conclusion is that defunding the big state mental hospitals was a mistake and we need to bring them back so that the mentally ill in state hospitals once more outnumber those in prison.

Lopez seems to be working off a model where there is a population of mentally ill people who can’t make it in normal society, and so will inevitably end up either in a long-term mental hospital or a prison. Since mental hospitals are good places where people get treatment, and prisons are bad places where people get punishment, we should “catch” these mentally ill people before they end up in prison so that they can be in nice hospitals instead.

Needless to say I disagree with pretty much every part of this assessment.

II.

Between all of this talk of “the tragic collapse of America’s public mental health system” and “the US’s largely gutted mental health system” and “the country pulled back and defunded its mental health system” and so on, you might get the impression that less money is being spent on mental health. This is not really true. The share of GDP devoted to mental health is the same as it was in 1971, although this looks worse if you compare it to rising costs in other areas of health care. There hasn’t been a “gutting of the mental health system”, there’s been a shift from long-term state-run mental hospitals to community care. It hasn’t “left the criminal justice system as the only system that can respond to people with mental illness”, it helped create an alternate and less restrictive system of outpatient psychiatry. In my opinion, this was a positive development, and the share of mentally ill people in prison is not an argument against it. Let me explain.

“Mentally ill people in prison” conjures up this lurid image of psychos who snap and kill their families, followed by “well, what did you expect leaving a person like that on the street?” The reality is more mundane. There are lots of mentally ill people in prison because there are lots of mentally ill people everywhere. Remember, 20% of the population qualifies as mentally ill in one sense or another. If a depressed guy sells some marijuana and gets caught, he is now a “mentally ill person in prison”.

There are disproportionately many mentally ill people in prison partly because people’s illnesses lead them to commit crimes, but mostly because some of the factors correlated with mental illness are the same factors correlated with criminality. Poverty? Check. Neighborhood effects? Check. Genetic load? Check. Education? Check. IQ? Check. Broken families? Check. Drug abuse? Definitely check. The factors that gave that pot dealer depression might be the same factors that drove him to sell pot instead of becoming an astronaut. Treating the depression might help a little, but it’s not guaranteed to keep him on the good side of the law.

In my model, the overwhelming majority of mentally ill people can live okay lives outside of any institution, hopefully receiving community care if they want it. If they commit crimes they will go to prison just like anyone else; if not, we should hardly be clamoring to bring back the often-horrifying state-run mental hospitals and lock them up there.

So when we talk about the number of mentally ill people in prison, we should be trying to distinguish between Lopez’s model and mine. That means asking: exactly how mentally ill are we talking about here?

III.

Lopez’s source for the claim that “ten times more mentally ill people are in prisons than hospitals” is a report by the Treatment Advocacy Center – note the less-than-neutral name. Where Lopez uses the phrase “mental illness”, TAC uses the phrase “severe mental illness” and defines it in two ways. For people in state prisons, they define it as reporting at least one psychotic symptom, and say 15% of people met their criteria. For people in county jails, they define it as meeting criteria for a depressive, bipolar, or psychotic illness, and say 15% of people met their criteria (they later arbitrarily increase that number to 20% because they feel like the survey might have undercounted).

No no NO. First, “psychotic” is not the same thing as “severely mentally ill”. Some people are severe but not psychotic – for example, a suicidally depressed person. Others are psychotic but not severe – for example, someone who hears a voice whispering her name but shrugs it off. Describing a survey that shows 15% of people as admitting one symptom of psychosis as showing 15% of people are severely mentally ill is really sketchy.

The prison survey provides a perfect example. It looks like the prisoners were asked fixed questions about their symptoms, and I think the exact screening instrument was just this survey, which has four relevant questions: “Can anybody else control your brain or thoughts?”, “Do you ever hear voices other people don’t hear?”, “Do you ever see something that other people tell you isn’t real?” and “Do you ever think anyone (other than correctional staff) is spying on you or plotting against you?”

Unfortunately, these kinds of surveys are really weak. I’m doing a study about this now, so maybe later I can cite myself on this, but the gist is that a lot of short mental health screening questions get false positives from perfectly healthy people. For example, I can’t tell you how many patients I’ve asked “Do you ever feel like anyone is spying on you?”, they say “Yes”, I ask “Who?” and they say “The NSA on my Internet activity”. Well, good work keeping up with the news. But a survey with a checkbox and no followup questions diagnoses that person as psychotic (see also: Lizardman’s Constant). This prison questionnaire was smart enough to exclude prison guards, who are certainly spying on all the respondents, but even beyond that I feel like the criminal lifestyle really does involve being spied on and plotted against a lot. At the very least it gives you lots of opportunities to legitimately worry about it.

(also, the diagnostic criteria for psychotic disorder are very clear that paranoia experienced while taking drugs doesn’t count. 75% of prisoners admit to using marijuana, marijuana can totally make you paranoid, and as far as I can tell the survey did not specify that the paranoia had to be while sober.)

(also also, Scientific American says that about 5-15% of perfectly ordinary people hear voices. Meanwhile, 4-6% of prisoners in the survey admitted to it.)

When the survey says that X% of prisoners have felt plotted-against or heard voices in the last year, does that mean X% of them are psychotic? That X% of them are “severely mentally ill”? That the old state mental hospitals need to be re-opened so X% of them can be locked up there for being too crazy for society? I don’t think it means any of those things.

But this is the stricter of the two criteria that the survey uses! The other one counts depressed people, bipolar people, and psychotic people. I don’t want to trivialize non-psychotic illnesses like depression. But remember: about 10% of the ordinary non-prison population is depressed/bipolar/psychotic. Also, going to prison is depressing as heck in and of itself. When they say that 15% of people in county jails (rounded up to 20%) are severely psychiatrically ill, they’re talking about pretty normal people, who might be in prison for something unrelated to their mental illness and might not even have become mentally ill at all if they hadn’t been incarcerated.

So I don’t think this survey shows the majority of the mentally ill prison population is in need of institutionalization. Yes, ten times more mentally ill people are in prison than in state mental hospitals, but consider the base rates! The prison population is huge. The population of people who need to be committed to mental hospitals 24-7 is tiny. Even if mentally ill people committed crimes at exactly the average population rate, there would still be far more mentally ill people in prison than in psychiatric hospitals, just by base rates! Especially if you use as broad a definition of “mentally ill” as these people!

So when Vox says that ten times more mentally ill people are in prison than in psychiatric hospitals, I will shoot right back at them that ten times more mentally ill people are in the Los Angeles metropolitan area than in state mental hospitals. You want more meaningless statistics? Ten times more mentally ill people are in the Southern Baptist Church than in state mental hospitals! Ten times more mentally ill people watched the last season of Game of Thrones than are in state mental hospitals! We can’t and shouldn’t aim to institutionalize all of them.

IV.

What about that graph? It’s very suggestive. You see a sudden drop in the number of people in state mental hospitals. Then you see a corresponding sudden rise in the number of people in prison. It looks like there’s some sort of Law Of Conservation Of Institutionalization. Coincidence?

Yes. Absolutely. It is 100% a coincidence. Studies show that the majority of people let out of institutions during the deinstitutionalization process were not violent and that the rate of violent crime committed by the mentally ill did not change with deinstitutionalization. Even if we take the “15% of inmates are severely mentally ill” factoid at face value, that would mean that the severely mentally ill could explain at most 15%-ish of the big jump in prison population in the 1980s. The big jump in prison population in the 1980s was caused by the drug war and by people Getting Tough On Crime. Stop dragging the mentally ill into this.

Lopez himself wrote a nice piece on how most mentally ill people are not violent, and another nice piece on how most people in prison are there for violent offenses. But put these together, and you get that most mentally ill people do not end up in prison. Most of the people who got out of the mental hospitals during deinstitutionalization are getting by. Some of them are homeless, and that’s bad. But if you want to solve homelessness among the mentally ill, build homeless shelters, not state-run long-term mental hospitals.

V.

In case you haven’t noticed, I really don’t like state-run long-term mental hospitals. There is a really amazingly great thing about prison, which is that you don’t go there unless you’re convicted of a crime. Mental hospitals do not have that advantage. The commitment process kind of sucks, and I am saying this as a person who makes commitment decisions myself. A lot of times it degenerates into a ritualized method of avoiding lawsuits without much concern for benevolence or patient autonomy. It helps me sleep at night to know that most commitments only last a couple of days or a week at most. Long-term state-run mental hospitals didn’t work that way. Remember that some of the perfectly sane people in the Rosenhan experiment were kept locked up for fifty days just for saying they heard a voice once but now they’re better.

I think long-term state-run mental hospitals are better than prison, but not by very much. The Rosenhan participants described it as:

…an overwhelming sense of dehumanization, severe invasion of privacy, and boredom while hospitalized. Their possessions were searched randomly, and they were sometimes observed while using the toilet. They reported that though the staff seemed to be well-meaning, they generally objectified and dehumanized the patients, often discussing patients at length in their presence as though they were not there, and avoiding direct interaction with patients except as strictly necessary to perform official duties. Some attendants were prone to verbal and physical abuse of patients when other staff were not present. A group of bored patients waiting outside the cafeteria for lunch early were said by a doctor to his students to be experiencing “oral-acquisitive” psychiatric symptoms. Contact with doctors averaged 6.8 minutes per day.

The idea of potentially saving a couple of people from prison by pre-emptively committing way more people to a mental hospital does not appeal to me at all, and I still think closing the institutions was the best thing Reagan ever did.

But prison and institutions aren’t the only two options! There’s a six month waiting period for psychiatrists in most parts of the country. The existing mental hospitals – which are different from and often nicer than the old state-run institutions – are constantly turning away people who want to be there because they don’t have enough beds for them. There are a bunch of patients who are having trouble affording their medications. There are special treatment options like day clinics, partial hospital programs, recreational therapy, occupational therapy, et cetera that do really great things but which most patients can’t afford. There are intensive health monitoring programs – think nurses who come to your house and make sure you take your medication on time – which are proven to improve outcomes but which never have enough staff for everybody who needs them. There are omnipresent underfunded community mental health systems. All of these things are doing great work right now. Indeed, the plan for closing the state-run long-term facilities was to gradually transition care to all of these other systems, and where that was supported it worked well, and insofar as it didn’t work well it was because it wasn’t supported.

If we support all that, will it keep all mentally ill people out of the prison system? No. First of all, no treatment is perfect and most are downright mediocre. Second of all, like I said, there are mentally ill people in prison because there are mentally ill people everywhere. There are disproportionately many mentally ill people in prison because the risk factors for mental illness are the same as the risk factors for crime, like poverty and drug abuse. Regardless of the level of care given, mentally ill people are likely to end up in prison at increased rates, unless you’re willing to either institutionalize all mentally ill people before they can commit any crimes, or excuse all crimes committed by mentally ill people.

But we shouldn’t be making our mental health decisions based on worries about criminality and prisons. Most people who are mentally ill will never end up on the wrong side of the law, and many (most?) mentally ill people who do end up in prison will do so for reasons not directly related to their illness. Make mental health decisions because it’s the right thing to do and there are people who really need help.

And if mentally ill people do end up in prison? There is a forensic health system dedicated to treating mentally ill prisoners. It’s not perfect, but with more funding and attention it could be better. There are forensic psychiatric hospitals that house mentally ill prisoners, and though again they are not perfect, they at least have that great advantage that you can’t be put in them unless you are found guilty of a crime.

So my argument is: fund and use the community mental health system more to help people in the community. Fund and use the forensic mental health system more to help people in prison. But stop acting like the two groups are fungible. And stop trying to institutionalize more people. That doesn’t help.

# Non-Shared Environment Doesn’t Just Mean Schools And Peers

[Epistemic status: uncertain. Everything in here seems right, but I haven’t heard other people/experts in the field talk about this nearly as much as I would expect them to if it were true. Obviously amount of variability attributable to environment (shared and non-shared) increases as the variability in environments in the sample increases]

The “nature vs. nurture” question is frequently investigated by twin studies, which separate interpersonal variation into three baskets: heritable, shared environmental, and non-shared environmental. Heritable mostly means genes. Shared environmental means anything that two twins have in common – usually parents, siblings, household, and neighborhood. Non-shared environmental is everything else.

At least in relatively homogeneous samples (eg not split among the very rich and the very poor) studies of many different traits tend to find that ~50% of the variation is heritable and ~50% is due to non-shared environment, with the contribution of shared environment usually lower and often negligible. This is typically summarized as “50% nature, 50% nurture”. That summary is wrong.

I mean, it’s tempting. All these social developmentalists were so sure that the way your parents praised you or didn’t praise you, or spanked you or didn’t spank you, had long-lasting repercussions that totally shaped your adult personality. The underwhelming performance of shared environment in twin studies torpedoed that whole area of study. But at least (these scholars of social behavior could tell themselves) it provided a consolation prize. The non-shared environment contributes 50% of variation, just as much as genes. That means things like your friends, your schoolteachers, and even that time you and your twin got sent away to separate camps must be really important. More than enough there to continue worrying about how society is Ruining The Children, right?

Not necessarily. Non-shared environment isn’t really “non-shared environment” the way you would think. It’s more of a dumpster. Anything that isn’t genetic or family-related gets tossed into the non-shared environment term. Here are some of the things that go into that 50% non-shared environment:

1. Error. Measurement error is neither genetics nor family, so it ends up in the non-shared environmental term. Suppose you’re studying intelligence, and you make a bunch of twins take IQ tests. IQ tests measure intelligence, but not perfectly. For example, someone who makes a lucky guess on a multiple choice IQ test will get a higher score even though they are not more intelligent than someone who makes an unlucky guess. Someone who takes the test when they’re tired and stressed may get a lower score even though they’re no less intelligent than somebody else who takes it well-rested and feeling good.

Imagine a world where intelligence is entirely genetic. Two identical twins take an IQ test, one makes some lucky guesses, the other is tired, and they end up with a score difference of 5 points. Then some random unrelated people take the test and they get the 5 point difference plus an extra 20 point difference from genuinely having different IQs. In this world, scientists might conclude that about 80% of IQ is genetic and 20% is environmental. But in fact in terms of real, stable IQ differences, 100% would be genetic and 0% environmental.

This gets even harder when trying to measure fuzzier constructs like criminality. Suppose someone does a twin study on criminality and their outcome is whether a twin was ever convicted of a felony. This depends partly on whether the twin is actually the sort of person with criminal tendencies – but also partly on whether a policeman happened to be in the area to catch them, whether their lawyer happened to be good enough to get them off, whether their judge was feeling merciful that day, et cetera. Imagine a world where criminality is entirely genetic. Identical Twin A becomes a small-time cocaine dealer in a back alley in West Philly, sells to an undercover cop, and ends up in jail. Identical Twin B becomes a small-time cocaine dealer in a back alley in East Philly, doesn’t run into any undercover cops, and so avoids conviction. This shows up as “variation in criminality is due to non-shared environment”.

Riemann and Kandler (h/t JayMan) run a study which is an excellent demonstration of this. Classical twin studies sometimes use self-report to determine personality – ie they ask people to rate how extraverted/conscientious/whatever they are. These studies find that most personality traits are about 40% genetic, 60% non-shared environmental. Riemann and Kandler obsessively collect every possible measurement of personality – self-report, other-report, multiple different tests – and average them out to get an unusually accurate and low-noise estimate of the personality of the twins in their study. They find that variation in personality is about 85% genetic, 15% non-shared environmental. So it looks like much of the non-shared environmental variation in traditional studies of personality was just error.

2. Luck of the draw. Bob becomes a junior advertising executive at Coca-Cola, where he designs a new ad targeting young female consumers. His identical twin Rob becomes a junior advertising executive at Pepsi-Cola, where he designs his own new ad targeting young female consumers. Both ads are very successful – in fact, exactly equally successful. But Coke’s CEO is a crony capitalist who wants to replace everyone in the company with his college buddies, so he ignores Bob’s good work and demotes him to a low-level position. Pepsi’s CEO is a skilled leader who recognizes good talent when she sees it, and she promotes Rob to Vice-President Of Advertising.

Now a scientist comes along, does a twin study on them, and finds that they have very different levels of income. She reports that there’s a lot of difference between these two identical twins, so much of income must be non-shared environmental.

Science reporters read the study finding that much of the variation in income is non-shared environmental, and conclude that despite their identical genes, there must be deep and mysterious differences in Bob and Rob’s abilities and business acumen. They speculate that Rob had a very inspirational teacher in school who pushed him to achieve greatness, and Bob must have fallen in with a bad peer group who didn’t value hard work.

But actually, Bob and Rob are completely identical in every way, no incident in their past did anything to separate them, and Bob just ended up working for a crappy CEO. In this scenario, inherent predisposition to earning money is exactly the same in both twins, they just have different amounts of luck at it. If both twins become pathological gamblers, but one of them hits the jackpot and the other goes broke, that will show up as “non-shared environment” too.

3a. Biological random noise. The genome can’t encode the location of every cell in the body. Instead, it specifies high-level processes which create lower-level processes which create those cells. But this gives the lower-level processes a lot of leeway, meaning that there can be significant biological differences between identical twins.

Consider by analogy The Postmodernism Generator. It’s a cute program that will make a (sort of) convincing sounding postmodernist essay on demand. We can imagine hundreds of different programmers all designing their own postmodernism generators. Some would be really brilliantly designed and consistently come up with plausible looking essays. Others would be poorly designed and consistently come up with crappy essays that don’t convince anybody. But there would also be variation within the results of each generator. There might be a generator that is mostly terrible but occasionally by coincidence comes up with a really funny essay, or vice versa. In this analogy, the genes are the code for the generator, and the person is an individual essay produced by that generator.

Thus, identical twins have different fingerprints, different freckles, and different birthmarks. Only about a fifth of left-handers’ identical twins will also be left-handed. And twins even look different enough that their friends and parents eventually learn to tell them apart. All of these are non-genetic issues likely to show up in “non-shared environment” but not related to schools or peers or “nurture” as traditionally conceived.

3b. The immune system. Immunology is still poorly understood, but it seems very important. Immune reactions and neuroinflammation have been implicated to one degree or another in a lot of psychiatric diseases. A functional immune system can protect good health; a dysfunctional immune system can make someone constantly tired and miserable.

There seems to be more of an element of chance to the immune system than to a lot of other bodily processes. Part of it is the input – one child in a twin pair might inhale a particle of cat dander at a critical time; another might get some unknown adenovirus with no immediate effects but which contributes to obesity twenty years later. Another part is the output; sometimes a natural killer cell stumbles across something quickly and takes it out without any fuss; other times the immune system misses it for a while and it gets more of a chance to spread.

The end result is that immune-system-related-conditions are really discordant across identical twins. If your identical twin has asthma, there’s only a 33% chance you’ll have it as well. If your identical twin has Crohn’s disease, a disabling autoimmune intestinal condition, there’s only a 50% chance you’ll suffer the same. I’m not sure how significant this is in the broad scheme of things, but I suspect more so than people think.

3c. Epigenetics We know that identical twins have substantially different epigenetics, and there are hints that this underlies discordant behavior. This is probably really important, but I feel bad bringing it up because it seems to be passing the buck. We usually think of epigenetic differences as a response to different environments or life choices. But if identical twins start with the same environment and can be expected to make the same life choices, why do they end up with different epigenomes? I’m not sure what to do with this one.

3d. Genes that differ between identical twins. Apparently this happens! Identical twins come from the same zygote, which means they start out with the same genes, but after that all bets are off. If there’s a mutation in one twin in the first embryonic division, then half of that twin’s cells will carry that mutation. Remember that there are a lot of divisions and opportunities for mutation before any cells even start forming the brain, and any mutation before that time could be transmitted to all brain cells. One study found that the average identical twin pair probably has about 359 genetic differences occuring early in development.

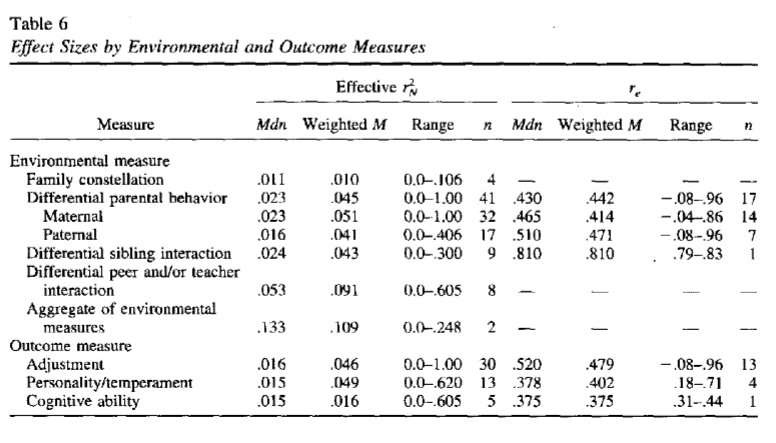
I’ve grouped 3a, 3b, 3c, and 3d together as possible biological sources of variation. One of these – or maybe some 3e I don’t know about – is probably the reason for less-than-perfect twin concordance in conditions like Parkinson’s disease, migraine, autism, and schizophrenia. Needless to say, anything that can make you schizophrenic can probably affect your personality and life outcomes pretty intensely.

But all of this gets counted as “non-shared environment” in a twin study, and used to play up the importance of schools and peer groups.

4. Actual nurture. Twins do have different experiences growing up. How much does this shape their adult traits? Can we separate this out in to specific experiences that shape adult traits, like school and summer camp?

The good news is that Eric Turkheimer has a big review article on this; the bad news is that the discussion section is called “The Gloomy Prospect” and says:

Quantitative analysis of studies of specific nonshared environmental events shows that effect sizes measuring the effects of such variables on child outcomes are generally very small. Effect sizes are largest when confounds with genetic variability and outcome-to-environment causal effects are not controlled. When such confounds are controlled, as in the most recent reports from the NEAD project, effect sizes become smaller still.



I’m not sure if this table represents the “very small” uncontrolled or the “smaller still” controlled sizes

The paper concludes: “We emphasize that these findings should not lead the reader to conclude that the nonshared environment is not as important as had been thought.”

But although I have a huge amount of respect for Turkheimer, I kind of want to conclude that the nonshared environment is not as important as had been thought. My guess is that the nonshared environment as Turkheimer discusses it – differential parenting, schools, peers, and so on – is only a fraction of the “nonshared environmental” term in genetics studies.

If that were true, it would mean that nature is more important than we thought relative to environment in terms of things we can understand and possibly affect. That would make the quest to change important outcomes like intelligence, personality, income, or criminality by changing society even more daunting. And it would make the opportunity to change those outcomes through genetic engineering even more tempting.

# Book Review: The Art Of The Deal

I.

Many of my friends recommend Robert Cialdini’s Influence, a book about how to be persuasive and successful. I read it most of the way through, and it was okay, but I didn’t have it in me to finish the whole thing. It’s not that being persuasive and successful doesn’t sound pretty neat. But how could I be sure the book would deliver the goods?

Robert Cialdini’s Wikipedia page says “He is best known for his book Influence“. Since its publication, he seems to have spent his time directing an institute to spread awareness of techniques for success and persuasion. At the risk of being a little too cynical – a guy knows the secrets of success, so he uses them to…write a book about the secrets of success? If I knew the secrets of success, you could bet I’d be doing much more interesting things with them. All the best people recommend Cialdini, and his research credentials are impeccable, but I can’t help wondering: if he’s so smart, why isn’t he Emperor?

Donald Trump may not be Emperor, but he’s a lot closer than Robert Cialdini. I knew that Trump wrote his own book on success and persuasion back in 1988 – Trump: The Art of the Deal – and I wondered if it might not be the anti-Cialdini.

Trump is no psychology expert, but after a few months of attributing his victories to blind luck, most people have accepted Scott Adams’ hypothesis that he’s really a “master persuader”. Salon, Daily Caller, Bill Maher, and the Economist all use the word “genius”. The less you respect Trump’s substance – and I respect it very little – the more you’re forced to admire whatever combination of charisma, persuasion, and showmanship he uses to succeed without having any. If this guy has written a book on how to be persuasive and successful, that’s a book I want to read.

II.

The downside of buying a book by a master manipulator is that sometimes you learn you were manipulated into buying the book.

Trump: The Art Of The Deal is 365 pages of some of the biggest print I have ever seen. The cover has a quote from the New York Times – “Trump makes one believe for a moment in the American dream again” – which some poor reviewer is probably desperately wishing he could take back right now.

Although the blurb says that he “fully reveals the deal-maker’s art” and that it is “an unprecedented education in the practice of deal-making” and “the ultimate read for anyone interested in achieving money and success” – only seventeen pages of very large print are anything resembling business advice. The rest of it is a weirdly deal-focused autobiography that doesn’t mention marrying his wife or having children, but devotes a lovingly detailed twenty-four pages to the time he renovated the Commodore Hotel.

But first, those seventeen pages. I am pleased to report that Donald Trump is weirdly abreast of modern science – he tells his readers looking for advice about how to make it big that deal-making is probably just genetic. Either you’ve got the deal, gene or you don’t:

More than anything else, I think deal-making is an ability you’re born with. It’s in the genes…unlike the real estate evangelists you see all over television these days, I can’t promise you that by following the precepts I’m about to offer you’ll become a millionaire overnight. Unfortunately, life rarely works that way, and most people who try to get rich quick end up going broke instead.

Every time I speak of the haters and losers I do so with great love and affection. They cannot help the fact that they were born fucked up!

— Donald J. Trump (@realDonaldTrump) September 29, 2014

Related?

This is a weirdly humble and self-aware Trump. It might be that the book medium suits him well; more likely he just has a really good ghost-writer. Unfortunately, he has much to be humble about. His advice, while not bad, is vague and not too useful. For example, his first rule is “think big”. But his second rule is “protect the downside and the upside will take care of itself”, which he explains as:

It’s been said that I believe in the power of positive thinking. In fact, I believe in the power of negative thinking. I happen to be very conservative in business. I always go into the deal anticipating the worst. If you plan for the worst – if you can live with the worst – the good will take care of itself.

So – take a lot of risks, but also be very cautious. Okay. I’m not saying his advice is literally contradictory – it makes sense that you can have big plans but also be very careful about them. I just don’t get the feeling that his advice is too helpful in narrowing down your strategy.

Is there anything at all worth reading in these seventeen pages? Oh yes. But not for the reason I expected.

Trump’s sixth rule of deal-making is “Get The Word Out”. He says:

One thing I’ve learned about the press is that they’re always hungry for a good story, and the more sensational the better. It’s in the nature of the job, and I understand that. The point is that if you are a little different, or a little outrageous, or if you do things that are bold or controversial, the press is going to write about you…

The funny thing is that even a critical story, which may be hurtful personally, can be very valuable to your business. [When I announced my plans to build Television City to the press], not all of them liked the idea of the world’s tallest building. But the point is that we got a lot of attention, and that alone creates value.

The other thing I do when I talk with reporters is to be straight. I try not to deceive them or to be defensive, because those are precisely the ways most people get themselves into trouble with the press. Instead, when a reporter asks me a tough question, I try to frame a positive answer, even if that means shifting the ground. For example, if someone asks me what negative effects the world’s tallest building might have on the West Side, I turn the tables and talk about how New Yorkers deserve the world’s tallest building, and what a boost it will give the city to have it again. When a reporter asks why I build only for the rich, I note that the rich aren’t the only ones who benefit from my buildings. I explain that I put thousands of people to work who might otherwise be collecting unemployment, and that I add to the city’s tax base every time I build a new project. I also point out that buildings like Trump Tower have helped spark New York’s renaissance.

The final key to the way I promote is bravado. I play to people’s fantasies. People may not always think big themselves, but they can still get very excited by those who do. That’s why a little hyperbole never hurts. People want to believe that something is the biggest and the greatest and the most spectacular.

I call it truthful hyperbole. It’s an innocent form of exaggeration – and a very effective form of promotion.

In the immortal words of Marco Rubio, “Let’s dispel once and for all with this fiction that Donald Trump doesn’t know what he’s doing. He knows exactly what he’s doing.”

On the other hand, his eighth rule of business is “Deliver The Goods”. He gives an interesting example:

You can’t con people, at least not for long. You can create excitement, you can do wonderful promotion and get all kinds of press, and you can throw in a little hyperbole. But if you don’t deliver the goods, people will eventually catch on.

I think of Jimmy Carter. After he lost the election to Ronald Reagan, Carter came to see me in my office. He told me he was seeking contributions to the Jimmy Carter Library. I asked how much he had in mind. And he said, “Donald, I would be very appreciative if you contributed five million dollars.

I was dumbfounded. I didn’t even answer him.

But that experience also taught me something. Until then, I’d never understood how Jimmy Carter became President. The answer is that as poorly qualified as he was for the job, Jimmy Carter had the nerve, the guts, the balls, to ask for something extraordinary. That ability above all helped him get elected president. But then, of course, the American people caught on pretty quickly that Carter couldn’t do the job, and he lost in a landslide when he ran for reelection.

Ronald Reagan is another example. He is so smooth and so effective a performer that he completely won over the American people. Only now, nearly seven years later, are people beginning to question whether there’s anything beneath that smile.

Trump-1988 is weirdly prophetic.

Finally, his tenth rule is “Have Fun”:

I don’t kid myself. Life is very fragile, and success doesn’t change that. If anything, success makes it more fragile. Anything can change, without warning, and that’s why I try not to take any of what’s happened too seriously. Money was never a big motivation for me, except as a way to keep score. The real excitement is playing the game. I don’t spend a lot of time worrying about what I should have done differently, or what’s going to happen next. If you ask me exactly what the deals I’m about to describe all add up to in the end, I’m not sure I have a very good answer. Except that I’ve had a very good time making them.

Marcus Aurelius, eat your heart out.

III.

So much for seventeen pages of business advice. The other three hundred forty-eight pages are Trump gushing about the minutiae all of the interesting deals he’s been a part of.

“GUYS, YOU’RE NOT GOING TO BELIEVE THIS, THERE WAS THIS ONE SKYSCRAPER THAT WAS SUPPOSED TO HAVE A FLOOR TO AREA RATIO OF 6, BUT THEN I BEAT HILTON IN NEGOTIATING THE AIR RIGHTS FROM THE COMPANY NEXT DOOR, AND ACTIVATED AN OPTION TO BUY A PROPERTY ON THE OTHER SIDE OF IT, AND ALL OF THAT LANDED ME A PARTNERSHIP WITH ONE OF THE BIG BANKS, AND THEN THE PLANNING BOARD TOTALLY CHANGED THE FLOOR AREA RATIO! CAN YOU BELIEVE IT, GUYS??!”

Overall the effect was that of an infodump from an autistic child with a special interest in real estate development, which was both oddly endearing and not-so-oddly very boring.

I started the book with the question: what exactly do real estate developers do? They don’t design buildings; they hire an architect for that part. They don’t construct the buildings; they hire a construction company for that part. They don’t manage the buildings; they hire a management company for that part. They’re not even the capitalist who funds the whole thing; they get a loan from a bank for that. So what do they do? Why don’t you or I take out a $100 million loan from a bank, hire a company to build a $100 million skyscraper, and then rent it out for somewhat more than $100 million and become rich?

As best I can tell, the developer’s job is coordination. This often means blatant lies. The usual process goes like this: the bank would be happy to lend you the money as long as you have guaranteed renters. The renters would be happy to sign up as long as you show them a design. The architect would be happy to design the building as long as you tell them what the government’s allowing. The government would be happy to give you your permit as long as you have a construction company lined up. And the construction company would be happy to sign on with you as long as you have the money from the bank in your pocket. Or some kind of complicated multi-step catch-22 like that. The solution – or at least Trump’s solution – is to tell everybody that all the other players have agreed and the deal is completely done except for their signature. The trick is to lie to the right people in the right order, so that by the time somebody checks to see whether they’ve been conned, you actually do have the signatures you told them that you had. The whole thing sounds very stressful.

The developer’s other job is dealing with regulations. The way Trump tells it, there are so many regulations on development in New York City in particular and America in general that erecting anything larger than a folding chair requires the full resources of a multibillion dollar company and half the law firms in Manhattan. Once the government grants approval it’s likely to add on new conditions when you’re halfway done building the skyscraper, insist on bizarre provisions that gain it nothing but completely ruin your chance of making a profit, or just stonewall you for the heck of it if you didn’t donate to the right people’s campaigns last year. Reading about the system makes me both grateful and astonished that any structures have ever been erected in the United States at all, and somewhat worried that if anything ever happens to Donald Trump and a few of his close friends, the country will lose the ability to legally construct artificial shelter and we will all have to go back to living in caves.

Trump’s greatest pride is his ability to construct things on time and under budget. He gives the story of an ice rink that New York City was trying to renovate in Central Park. After six years and $13 million, the city had completely failed to renovate it and just made things worse. Trump offered as a charitable gesture to do it himself, and the mayor, who was a political enemy, refused. The press hounded the mayor, Trump eventually was allowed to try, and he finished it in four months for only $2.5 million. He boasted that he finished fixing the rink in less time than it took the city to complete their study on why their rink-fixing project had failed.

He had a couple more stories like this – but throughout all of it, there was a feeling of something missing. Here is a guy whose job is cutting through bureaucracy, and who is apparently quite good at it. Yet throughout the book – and for that matter, throughout his campaign for the nomination of a party that makes cutting bureaucracy a big part of their platform – he doesn’t devote a lot of energy to expressing discontent with the system. There is no libertarian streak to Trump – in the process of successfully navigating all of these terrible rules, he rarely takes a step back and wonders about a better world where these rules don’t exist. Despite having way more ability to change the system than most people, he seems to regard it as a given, not worth debating. I think back to his description of how it’s all just a big game to him. Most star basketball players are too busy shooting hoops to imagine whether the game might be more interesting if a three-pointer was worth five points, or whatever. Trump seems to have the same attitude – the rules are there; his job is to make the best deal he can within those rules.

Maybe I’m imagining things, but I feel like this explains a lot about his presidential campaign. People ask him something like “How would you fix Medicare?”, and he gives some vapid answer like “There are tremendous problems with Medicare, but I’m going to hire the best people. I know all of the best doctors and health care executives, and we’re going to cut some amazing deals and have the best Medicare in the world.” And yeah, he did say in his business tips that you should change the frame to avoid being negative to reporters. But this isn’t a negative or a gotcha question. At some point you’d expect Trump to do his homework and get some kind of Medicare plan or other. Instead he just goes off on the same few tangents. This thing about hiring the best people, for example, seems almost like an obsession in the book. But it works for him. When somebody sues him (which seems like an hourly occurrence in real estate development no matter how careful you are) his response is to find the best lawyer, hire them, and throw them at the problem. When he needs a hotel managed, he hires the best hotel managers and tells them to knock themselves out. Even his much-mocked tendency to talk about all the people he knows comes from this being a big part of his real estate strategy – one of the reasons he can outcompete other tycoons is because he knows people on the planning board, knows people in the banks, knows people in all the companies he works with. It’s a huge advantage for him.

These strategies have always worked for him before, and floating off into some intellectual ideal-system-design effort has never worked for him before. So when he says that he’s going to solve Medicare by hiring great managers and knowing all the right people, I don’t think this is some vapid way of avoiding the question. I think it’s the honest output of a mind that works very differently from mine. I’ve been designing ideal systems of government for the heck of it ever since I was old enough to realize what a government was. Trump is at serious risk of actually taking over a government, and such design still doesn’t appeal to him. The best he can do is say that other people are bad at governing, but he’s going to be good at governing, on account of his deal-making skill. I think he honestly believes this. It makes perfect sense in real estate, where some people are good businesspeople, others are bad businesspeople, and the goal is to game the system rather than change it. But in politics, it’s easy to interpret as authoritarianism – “Forget about policy issues, I’m just going to steamroll through this whole thing by being personally strong and talented.”

I said it before, but it bears repeating – this book has a really good ghostwriter. Yeah, it comes across as narcissistic; there’s probably no way to avoid that in a Trump autobiography. But Donald Trump’s interest in Donald Trump pales beside his blazing hot interest in the sheer awesomeness of hotel property deals. And part of me wants to say that people with obsessive interests in bizarre things are My Kind Of People.

But there’s still something alien about Trump here, even moreso than with the populist demagogue of the campaign trail. Trump the demagogue is attacked as anti-intellectual. I get anti-intellectualism because – like all isms – it’s an intellectual idea, and I tend to think in those terms. But Trump of the book is more a-intellectual, in the same way some people are amoral or asexual. The world is taken as a given. It contains deals. Some people make the deals well, and they are winners. Other people make the deals poorly, and they are losers. Trump does not need more than this. There will be no civilization of philosopher-Trumps asking where the first deal came from, or whether a deal is a deal only by virtue of its participation in some primordial deal beyond material existence. Trump’s world is so narrow it’s hard to fit your head inside it, so narrow that on contact with any wider world it seems strange and attenuated, a broken record of deals and connections and hirings expanding to fill the space available.

On the other hand, he made a billion dollars and will probably win the GOP nomination. So there’s that.

Trump ends by saying:

What’s next? Fortunately, I don’t know the answer, because if I did, that would take all the fun out of it. This much I do know: it won’t be more of the same.

I’ve spent the first twenty years of my working life building, accumulating, and accomplishing things that many said could not be done. The biggest challenge I see over the next twenty years is to figure out some creative ways to give back some of what I’ve gotten.

I don’t just mean money, although that’s part of it. It’s easy to be generous when you’ve got a lot, and anyone who does, should be. But what I admire most are people who put themselves directly on the line. I’ve never been terribly interested in why people give, because their motivation is rarely what it seems to be, and it’s almost never pure altruism. To me, what matters is the doing, and giving time is far more valuable than just giving money. [note: a contrary perspective]

In my life, there are two things I’ve found I’m very good at: overcoming obstacles and motivating good people to do their best work. One of the challenges ahead is how to use those skills as successfully in the service of others as I’ve done, up to now, on my own behalf.

Don’t get me wrong. I also plan to keep making deals, big deals, and right around the clock.

# The Price Of Glee In China

[Epistemic status: Overly simplistic treatment of a horrifyingly complex topic; I can only hope I haven’t missed enough to completely embarrass myself]

I.

Noah Smith reviews recent economic research suggesting that globalization was a net harm to working class people in rich countries like the US; he tentatively suggests this could justify a weak form of protectionism. But Scott Sumner argues that’s the wrong way to look at things. Globalization fueled China’s transition from a poor agrarian economy to an industrialized modern nation. A billion people were lifted out of poverty, an accomplishment Sumner calls “the best thing that ever happened”. This is far more important than the less dramatic costs imposed on the US. Therefore, even if we agree globalization hurts the working class of rich nations, it’s still a morally defensible policy since it benefits the needier working classes of much poorer nations.

On the one hand, this makes sense. On the other, here’s happiness in China over the past fifteen years:



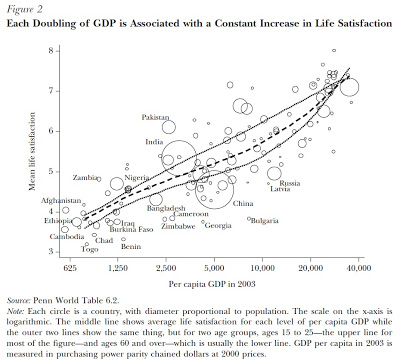
Measuring happiness is really hard, but the Chinese result seems as robust as any. You get the same thing if you ask about satisfaction versus dissatisfaction. Brookings analyzes five different series of happiness data and concludes that “the Chinese became less happy during their growth boom”. The New York Times agrees and says that “Chinese people’s feelings of well-being have declined in [this] period of momentous improvement in their economic lives”. And this seems to be worst among the poorest Chinese:



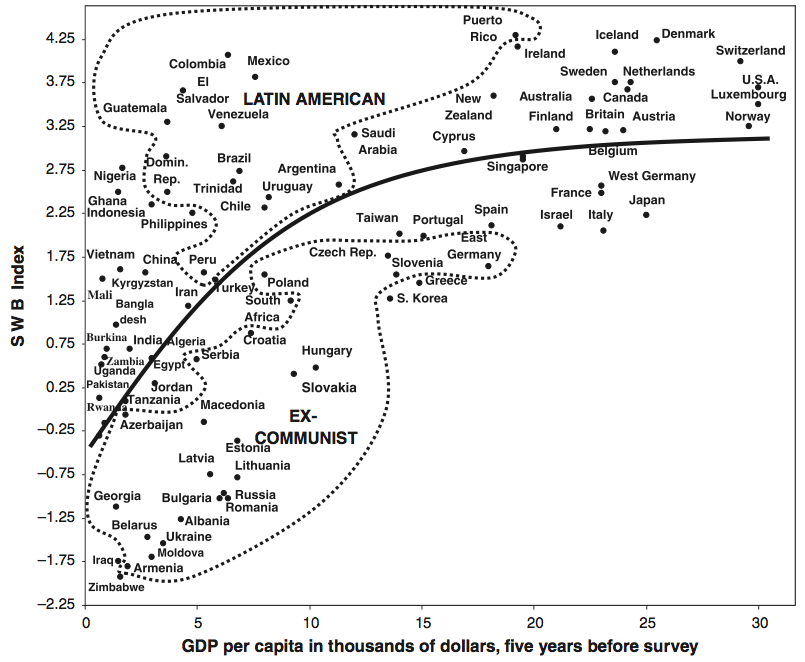
Nor does this seem to be an effect from our happiness research just not being good enough to capture changes in happiness even if they occur. There’s good evidence that increased income within a country increases happiness, and various other things have been found to be effective too. I would even argue we can find happiness changes in nations – recent surveys have found Iraq and Syria to be the least happy nations in the world, and I doubt this was true before those countries’ respective wars. It seems to just be national GDP per capita that doesn’t do anything.

This is Easterlin’s Paradox, the observation that a country in general does not get happier as it becomes richer. This is very controversial, with statisticians analyzing and reanalyzing data and crunching it a bunch of different ways. In the latest volley in this eternal war, Easterlin’s side came out with data from 37 countries over 30 years, including many countries that underwent spectacular growth during that time, and confirmed their original conclusion.

There are certainly graphs like this one that propose a nice clear log relationship between income and happiness:

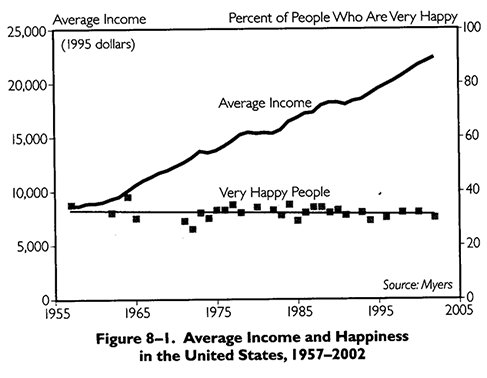


But I find the exact breakdown much more interesting:



Here we see a lot of cultural variation in this apparent happiness-income relationship. For example, Latin American countries are consistently poor but happy; Eastern European countries are usually richer but sadder than African countries, et cetera. Looking at the original graph above, you’d expect Chinese growth to make them much happier; looking at this graph, you notice that China’s three rich neighbors – Japan, Taiwan, and South Korea – are all about as happy as China. South Korea, despite making five times more money, is less happy than China is. If China’s income quintuples, why would you expect it to look like France or Ireland rather than South Korea?

Just to rub this in a little:



A UN report theorizes that although richer countries tend to be happier, this is more likely due to factors other than income, like freedom, social trust, and stable families. These may be stable on scales much longer than income is, and may be related to culture.

II.

Let’s assume for a second that all this is true. National income does not matter for national happiness, and if China’s growth continues to skyrocket then in twenty years it will be as rich as Japan but not an iota happier than it is today. What do we do with this kind of knowledge?

Or let me ask a more specific question. Suppose that some free trade pact will increase US unemployment by 1%, but also accelerate the development of some undeveloped foreign country like India into hyper-speed. In twenty years, India’s GDP per capita will go from $1,500/year to $10,000/year. The only cost will be a million or so extra unemployed Americans, plus all that coal that the newly vibrant India is burning probably won’t be very good for the fight against global warming.

Part of me wants to argue that obviously we should sign the trade pact; as utilitarians we should agree with Sumner that lifting 1.4 billion Chinese out of poverty was “the best thing that ever happened” and so lifting 1.2 billion Indians out of poverty would be the second-best thing that ever happened, far more important than any possible risks. But if Easterlin is right, those Indians won’t be any happier, the utility gain will be nil, and all we will have done is worsened global warming and kicked a million Americans out of work for no reason (and they will definitely be unhappy).

Or since most of us don’t get the option to sign trade pacts, here’s a more relevant question. Suppose we are effective altruists. We have the opportunity to cure disease (at relatively high costs) or boost national development (at relatively low costs). Assume the numbers work out such that if we took a simple ‘development = good’ perspective, then donating to the development charity would be a no-brainer. Should we donate to the disease-cure charity anyway?

A couple of years ago, I learned that people who were paralyzed in car accidents took a few months to adjust to their new situation, but after that were no less happy than people who were still healthy and abled. Then last December I learned that this was an urban legend, that people who were paralyzed in car accidents were mostly as miserable as you would expect. But for those few years while I still believed that particular factoid, I was a little creeped out. Was a doctor who helps car accident victims recover their function wasting her life? If people got genuine enjoyment from driving drunk at 95 mph while shouting “WOOOOOOOOOOOOO!”, was there any reason to make them stop, since they weren’t really hurting anybody?

(I admit I’m skipping over factors like how paralyzed people can’t earn any income to pay into the tax system and stuff, but I’m just saying I would be pretty creeped out if that were the only reason we should avoid car accidents.)

Again assuming I haven’t made some simple calculation mistake, I can think of three ways to go from here. First, abandon consequentialism entirely (I understand that having children will likely decrease my happiness, but I still want to have children because I value them for non-utilitarian reasons). Second, switch to a consequentialism based on non-subjective things like maximizing development and industrialization as a terminal goal (Really? Even if everyone hates it? Does it matter what the factories are building? How about paper clips?). Third, switch to preference utilitarianism.

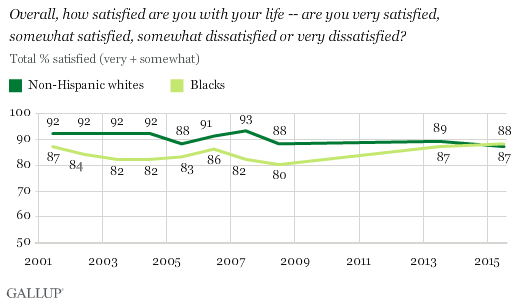
Preference utilitarianism is tempting and I was kind of in favor of it already, but I don’t find it completely satisfying. Suppose I myself am an Indian peasant. Should I have a preference for my society industrializing? If I’m not going to be any happier after it does, and supposing there’s no inherent moral value in industrialization, why bother? And if Indian peasants want their country to industrialize anyway, aren’t we as Americans allowed to say we don’t take their preference that seriously? If some hippie said they wanted to go on some Spiritual Yoga Nature Retreat that would turn their life around and bring them constant bliss, but we knew it was a complete fraud that wouldn’t help them at all, would we still feel a moral obligation to help fund that hippie’s retreat? How are the two situations different?

There’s a risk of being patronizing here – telling the Indians “Oh, you don’t need to industrialize, it’s not so great anyway,” even while we ourselves enjoy our nice food and flat-screen TVs. If we were to actively try to keep the Indians from industrializing, that would be pretty awful. But that’s not the argument at hand here. The argument at hand is “are we morally required to sacrifice our own economy in order to help the Indians industrialize?”, and I feel like that’s a hard sell if industrialization doesn’t really help the Indians.

And there’s also a risk that I might be misdefining happiness. Maybe every way economists have hitherto measured happiness is hopelessly deficient, and there’s some ineffable essence of happiness which, if we could get at it, would increase during national development. I admit that all of these subjective well-being indices are kind of sketchy and change a lot with the wording that you use or don’t use.

A final option for rescuing common sense might be acknowledging that economic progress doesn’t change happiness yet. That is, there are ways to convert economic (and closely linked technological) progress into happiness, but most countries are not making use of them – either for political reasons, or because they don’t know about them, or because we haven’t gotten enough technological and economic progress to reach them yet. This seems probably true to me – if nothing else, a technological singularity ought to help – but this situation looks a lot different from the situation where incremental progress increases happiness. In particular, it would make us want to concentrate our resources on increasing technological progress, perhaps in the richest economies, rather than trying to help poor countries in particular.

None of these possibilities really appeal to me, and I am forced to acknowledge that happiness research remains a very strange field whose conclusions make no sense to me and which tempt me to crazy beliefs and actions if I take them seriously.



I guess we’re done fighting racism. Good job, guys.

# Book Review: My Brother Ron

[Content warning: mental illness, forced institutionalization, anorexia. As always all patient anecdotes are obfuscated composites of multiple cases with all the details changed in order to protect people’s privacy]

I.

After I wrote about Prison And Mental Illness, a reader recommended I read My Brother Ron by Clayton Cramer, a recent book/memoir arguing against deinstitutionalization. Cramer tells the story of his schizophrenic brother Ron, who was poorly treated because of the lack of an institutional system and so ended up dealing with homelessness and violence, then surveys the history and current state of mental health care in America and the various reasons why deinstitutionalization was a bad idea.

I found the book interesting and engaging, and its arguments intellectually honest and well-written. But in the end I just wasn’t convinced.

But first, his brother Ron. Smart guy, joined the military, did well, finished his tour of duty, went to college, studied electrical engineering. Around 22 – the usual age for this to happen – he started acting weird, dropped out of college, obsessed over weird things like nickels, started thinking random people were plotting against him, et cetera. He ended up in a psych hospital where he got Thorazine and improved quickly – which meant, ironically, that when it came time for his commitment hearing two weeks later, the judge thought he looked pretty normal and released him.

Then he went to live with his family – including his brother the author – where he stopped his medication, started acting violently, smashed windows, screamed at people, and was otherwise a poor housemate. His parents asked him to leave, and he wandered around until he ended up in Santa Monica. There the government gave him a monthly disability check, which he spent on alcohol and a room in a disgusting hotel; when the money ran out around the middle of the month, he spent the next few weeks on the street until he got his next check, after which the cycle repeated itself.

Every so often he would break some law or annoy somebody enough to get arrested, at which point the police would bring him to a psychiatric hospital, he’d be placed on drugs, and he’d get better. Usually he’d leave after a few days to a few weeks. Occasionally he would keep taking the drugs after getting out, become pretty with-it, and try to go back to college. Sometimes he’d stay stable for months, even a year or two. But eventually he would stop taking the drugs for one reason or another, decompensate, and end up back on the streets, his previous progress ruined.

So the author asks: how did we get to this point? He answers with a fascinating history of American mental health care.

II.

Mental health care during the colonial era was surprisingly non-terrible. Mental illness seemed to be pretty well-understood and nobody was accusing psychotics of being witches or trying to beat the demons out of them or anything. Most of the mentally ill lived with families or in their own houses, where other members of the community supported them as best they could. Some were given jobs, with the understanding that they needed the support and their idiosyncrasies would be excused. Some would wander off, and there was a general understanding among colonial towns that if they found a mentally ill person wandering they would return them to their town of origin, who had the ultimate responsibility of caring for them. A few very violent people were locked away, usually in the basements of general hospitals or in prison cells. Getting somebody committed for mental illness was an informal process usually involving finding the friendly local magistrate and explaining why it was a good idea. But this option seems to have been used judiciously, and the incarcerated individuals managed to avoid most abuse and torture. Cramer describes it as “gloriously idyllic…mental illness appears to have been rare, and small town life tolerated all but the ‘furiously mad’ to live in the community.”

The part I found most interesting here was Cramer’s theory about why this system ended. Part of it was the end of small town life; a little village where all the families know each other is more likely to tolerate someone’s eccentricities than a large city of atomized individuals. But a bigger part may have been an unmanageable increase in the mentally ill population.

Urbanization may not simply have been a factor in making Americans more wary of their mentally ill neighbors; it may have increased mental illness rates as well. While we do not know if this was true in the eighteenth century, some recent studies suggest that being born or growing up in an urban area increases one’s risk of developing schizophrenia and other psychoses. in the twentieth century, comparison of insanity rates revealed that urban areas had much higher rates of mental hospital admissions for schizophrenia and bipolar disorder – almost twice as high for New York City compared to the rest of New York State…older statistical examinations of mental hospital admissions argue that at least in the period from 1840 to 1940, while mental hospitalizations increased (because of increased availability) there was no large and obvious increase in insanity. A more recent study of mental illness data shows, much more persuasively, that psychosis rates rose quite dramatically between 1807 and 1961 in the United States, England and Wales, Ireland, and the Canadian Atlantic provinces. A study of Buckinghamshire, England shows more than a ten-fold increase in psychosis rates from the beginning of the seventeenth century to 1986. In 1764, Thomas Hancock left 600 pounds to the City of Boston to build a mental hospital for the inhabitants of Massachusetts. The city declined to accept the gift on the grounds that there were not enough insane persons to justify building such a facility. Massachusetts had a population between 188,000 and 235,000 in 1764; if the population of the time suffered the same schizophrenia rates as today, that would mean that there were about 2000 schizophrenics in the province. Even accounting for the greater tolerance of small town life for the mentally ill, this lends credence to Torrey and Miller’s claim of rising psychosis rates. Urban life today is not the same as urban life then, and even the scale of what constitutes “urban” is dramatically different – but it is an intriguing possibility that the increased rates of mental illness at the close of the Colonial period were the results of urbanization.

Irish immigration may also have played a role in the increasing development of mental hospitals in America. It was widely believed in the 1830s that Irish immigrants were disproportionately present among the insane. More recent analysis shows that throughout the nineteenth and twentieth centuries, Ireland’s rates of insanity were twice or more than that of the United States, England, and Wales. Irish immigrants were also overrepresented in insane asylums in the United States, England, Australia and Canada at the end of the nineteenth century.

To this I would add that even today immigrants get schizophrenia at rates up to four times those of non-immigrant populations, though nobody agrees whether this is because the genetically vulnerable are more likely to immigrate or because immigration is a very stressful experience. Even today, developing countries seem to have less schizophrenia than developed countries do (although of course this is hard to prove with certainty). The idea of a tenfold increase in psychosis over the past few centuries is jarring but not entirely outlandish, and does a lot to explain why the mental health system is so much larger and more relevant now.

Faced with these problems, the early Americans created big mental institutions that attracted prestigious clinicians (I interviewed for a job at one of these a few years ago; they boasted that they were in the “Psychiatric Ivy League”, which was a pretty good window into how they thought of themselves). These could never really figure out whether their job was custodial (ie warehouse mentally ill people so they didn’t cause trouble on the streets) or clinical (treat mentally ill people and cure their psychosis), and the nineteenth century vacillated wildly between people making big claims about how they were dedicated to treating all their patients, versus admitting that it was the nineteenth century and nobody had the slightest idea how to do this. While they argued the institutions grew and grew. Along with the schizophrenics, they became the dumping ground for syphilitics (remember, before penicillin syphilis was a common incurable disease that usually caused insanity in its final stages) and old people with Alzheimers (not officially recognized at this point; before the invention of nursing homes they figured they might as well stick crazy old people in with all the other crazy people). Finally, after the obsolescence of the “poorhouse” but before the beginning of welfare, there were a bunch of poor people just completely unprepared for normal life, and some of them ended up in the mental institutions too for lack of a better place to put them. This sort of put a damper on a lot of the curability discussion; not only could 19th century doctors not cure mental illness, but most of the people there weren’t even mentally ill in the traditional sense.

(not that some people didn’t try. Cramer describes a Dr. “Henry Cotton, who removed teeth, tonsils, and parts of the intestine from hundreds of patients at the Trenton State Hospital in New Jersey. Cotton claimed that there were foci of infections in these organs that were causing the insanity and that removal of the infectious would cause clinical improvement.” And then there was Dr. Wagner-Jauregg, whose bold strategy of deliberately infecting psychiatric patients with malaria actually paid off: many of them had syphilis, and the high fever induced by the malaria killed the syphilis bacterium. Wagner-Jauregg received the Nobel Prize for this insight; his later strategy of sterilizing schizophrenics on the theory that the disease was caused by masturbation was perhaps somewhat less Nobel-worthy.)

The institutions continued to grow. In 1954 the national mental health budget was $568 million; in 1959 it was $854 million. In 1951, states spend on average 8% of their budgets on psychiatric hospitals; New York spent one third of its budget on psychiatric hospitals (or not? see dispute in comments). Compare to today, when New York spends only about 20-30% of its budget on education. Psychiatric hospitals (which, remember, also subsumed the function of modern nursing homes) were a huge part of the infrastructure of government.

This started to shift in the 1940s due to what the book calls “dynamic psychiatry” (although they use this phrase a bit differently from how I understand the definition). The old, tired psychiatry was a simple dichotomy between sane people (who don’t need psychiatric help) and insane people (who are totally out of touch with reality and need to be locked up for their own good). And it understood this distinction in relatively biological terms – they didn’t know anything about genes or neurons them, but they figured something was going on. But the new, exciting psychiatry thought of mental illness as a continuum, with everybody having a little bit of mental illness – whether it was just neurosis or anxiety or whatever – and psychotics just being the people whose mental illnesses made it hard for them to function. The new school understood this in very psychosocial, Freudian terms. Schizophrenics were people with oppressively close mothers; autistics were people with distant, cold mothers, et cetera. Psychiatrists tended to like this new school, because it meant that instead of spending their time in scary mental institutions full of crazy people, they could spend their time in nice Viennese parlors talking to rich people about their families.

Around the same time, scientists invented Thorazine, which seemed to produce miraculous recoveries in institutionalized psychotic people. This was before anyone knew anything about the long-term side effects of Thorazine, so everyone figured it was a miracle drug with no side effects and now there was no need for mental institutions any more.

Then we got to the Sixties. Cramer mostly manages to avoid being too transparently political, but it’s hard for him to talk about Sixties Leftists without a bit of vitriol. He describes the genesis of the anti-psychiatry movement – a wide variety of traditions all coming together in an agreement that the mentally ill are just Too Cool And Free-Spirited For Society and anybody who tries to treat them is a bad person who hates creativity and wants to make everyone conform. He describes the jettisoning of centuries of accumulated wisdom about the causes and presentation of mental illness in favor of an unexamined dogma that mental illness is caused by oppressive systems of social control. He describes how some people did a few quick studies showing that schizophrenic people mostly lived in bad neighborhoods full of social decay, and concluded that bad neighborhoods and social decay caused schizophrenia without considering any other possible causal structures (of course, we as a society have long since moved beyond that). Others argued that hospitalization was the sole cause of mental illness, turning otherwise happy eccentrics into violent lunatics (again, a position we have long since moved beyond).

He reserves some of his strongest words for anti-psychiatry psychiatrists like R. D. Laing and Thomas Szaszszsz:

You might wonder how a psychiatrist could believe that there was no such thing as insanity. Would not the exposure to psychotic patients during Szasz’s training have shown him the error of his ideology? It turns out that Szasz may not have had any exposure to psychotics. In a 1997 interview, he describes how he consciously selected a psychiatric residency “that did not include work with involuntary patients”. The chairman of the Psychiatry Department told him, “Tom, you have only one year left of your residency, I don’t think it’s right that you should finish without any experience with psychotic patients. I think you should do your third year at the Cook County Hospital.” So Szasz quit and went elsewhere to avoid that experience.

Szasz was drafted into the Navy after completing his training, and his experiences there almost certainly reinforced his already well-developed belief that mental illness did not exist. “The servicemen didn’t want to be in the Navy and played the role of mental patient. I didn’t want to be in the Navy and played the role of military psychiatrist. My job was to discharge the men from the Service as ‘neuropsychiatric casualties’.” Szasz had gone out of his way to avoid seeing psychotic patients, and then took a job that he describes as certifying that sane people pretending to be insane were actually insane as a convenient fiction. Is there anything surprising about Szasz’s projection of this situation onto the entire profession?

I actually had been wondering about that, and that clears up a lot. As for Laing:

In the mid-1960s, British activists gravitated to Laing’s ideas, arguing that schizophrenia was more “properly human”, in a world of hydrogen bombs, than conventional definitions of sanity…Laing argued that schizophrenia was not a breakdown but a breakthrough. By the 1970s, Laing took the position of Huxley’s The Doors of Perception, that schizophrenia was a form of sanity, not insanity. Laing’s position increasingly became a political attack on Western society, and then morphed once again, rejecting the idea of schizophrenia by declaring it as hypersanity. Eventually Laing’s celebrity led him to India and drug abuse, and he became a shell of his former self.

Well then.

Around the time all this was going on, the ACLU was launching an attack of its own on the psychiatric system. Most of what they were saying sounded good – make sure people only get committed if the courts are absolutely sure they’re insane, make sure that they have all of their rights even within the psychiatric hospital – but Cramer references internal memos and discussions purporting to show that the ACLU’s real goal was to make psychiatric commitment so bureaucratically difficult that nobody would ever do it, thus freeing the mentally ill from their oppressors and destroying the psychiatric system. The courts were sympathetic to their cases and established several new rights and standards that made committing people exceptionally difficult.

In exchange, the opponents of institutions promised community treatment. Everybody agrees that community treatment was a good idea. The implementation left a lot to be desired. First, as always, they were seriously underfunded. Second, even the ones that had enough money quickly found that creating outpatient psychiatric centers is fundamentally geographically difficult. Schizophrenics are not known for their ability to go places on an organized schedule, nor for their access to good consistent transportation. The great advantage of the old asylums was that all of the schizophrenics were in one convenient location for the mental health workers to treat. When the new community treatment centers were set up, they tended to serve any schizophrenics who might live within a few blocks of them, and all the rest never made it to their appointments. Third, as per Cramer most of the people operating these new community centers were Sixties Leftists who decided that instead of the “bandaid solution” of actually treating mentally ill people, their real job was to cut out mental illness at the root by protesting capitalism and racism:

One of the officials of the CMHC [Community Mental Health Centers] program later admitted that the CMHCs “were not equipped to deal” with the chronically mentally ill, who were about to be released in large numbers from state mental hospitals. The belief that mental hospitals caused mental illness, or at least made the mentally ill worse off than they were before, combined with an idealized view of how caring communities would be for the severely mentally ill. The activists and bureaucrats who wrote the CMHC regulations were about to start the release of mental patients into caring communities which for the most part did not exist. As one of those involved later admitted, “We were federal bureaucrats on an NIMH campus talking about the community, but really from some conceptual level as opposed to hands-on experience.”

If CHMCs were not primarily serving the chronically mentally ill, then whom were they serving? Two especially notorious examples were Lincoln Hospital Mental Services in New York City and Temple University Community Mental Health Center in Philadelphia. In both cases, the belief that mental illness was somehow an expression of class struggle meant that broader social and political causes – such as landlord/tenant relations, poverty, and oppression – became significant activities of the staff. Racial and ethnic tensions within the staff destroyed both CMHCs, with threats of violence, sit-ins, VietCong flags, posters of Che Guevera and Malcolm X as symbols of the fight.

In the late sixties and early seventies all of these things came together. Psychiatrists wanted to focus on healthy people who were much more pleasant to talk to. Pharmaceutical companies insisted that their new wonder drugs could cure psychosis. Activists wanted to destroy the psychiatric system. Judges were making it much more difficult to commit anybody. And community mental health centers were trying to pick up the slack. The result was the deinstitutionalization strategy called “closing the front door and opening the back door” – that is, making new commitments more difficult, and accelerating the pace at which psychotics already in institutions could be discharged to the new community treatment programs (it didn’t hurt that syphilis had been cured a few decades earlier and the last few chronically insane syphilitics were dying off as well). This went exactly according to plan, the institutionalized population shrunk and shrunk throughout the seventies, and by the time Reagan decided to close the last few psychiatric institutions there wasn’t much left to close down.

III.

Needless to say, Cramer opposes most of these developments. He makes his antideinstitutionalization argument in several parts. But first, some things he doesn’t argue.

Cramer is pretty quick to admit the institutions had their problems:

Many [psychiatric hospitals] remained “snake pits”, to borrow the title of Mary Jane Ward’s very popular 1946 novel about mental hospitals. The American Psychiatric Association created the Central Inspection Board in 1947 to evaluate existing mental hospitals in the United States and Canada. The results were not encouraging. By 1953, it had evaluated 45 hospitals, approved two, given ten a “contingent approval”, and disapproved the rest.

The book frankly discusses the “regimented, often hopeless conditions of state mental hospitals”, talks about a hospital in Alabama where “care was worse than simply inadequate: one psychiatrist for 5000 patients; astonishingly low funding for clothing, food and upkeep of the buildings”, studies showing that institutions never actually got patients’ signatures on the forms that were supposed to waive their rights to court hearings. It describes the case of Edna Long, who was hospitalized for “public drunkenness” and

permanently hospitalized in 1952. As Ennis tells the tale, Long received no treatment during the next fifteen years, but was kept busy working at menial jobs in the hospital. After the death of her husband in 1960, the state hospital had her declared incompetent, and seized her assets to pay for her care. Then, they put what assets remained under the management of an attorney, who made a bit of money from reducing the value of her estate by 86% (according to Ennis, a common practice at the time in New York). Once Long had become too physically ill to continue working, the hospital suddenly found her “competent to manage her own affairs” and released her, to a life of elderly poverty. Most of the money that she and her husband had accumulated had been consumed by attorneys supposedly protecting her assets.

Against this tale of woe, Cramer can say only that it “leads me to wonder if there was a bit more to the story”. Judging from my own conversations with patients and nurses who used to live in / work at these hospitals – who generally report similar stories – I doubt there was.

So what is this book’s argument against deinstitutionalization?

First, it points out that very many deinstitutionalized schizophrenics slipped through the community mental health system and never got further treatment. This was in part due to the problems with CMHCs – poor funding, difficult to get to, sometimes not that interested in mental health at all (though they got a lot better after the Sixties). But it was also due to schizophrenics just generally not being too interested in engaging with the psychiatric system (especially, one might imagine, the ones who had just gotten out of institutions) and no one being able to make them. I 100% acknowledge that this argument is correct.

Second, it points out that many untreated or unsuccessfully treated schizophrenics ended up homeless on the street.

“Of 179 homeless men and women who received psychiatric examinations in a Philadelphia shelter in 1981, 40% were found to have “major mental disorders”. One-third of those examined were diagnosed as schizophrenic, and another one-fourth had a primary diagnosis of substance abuse. A Boston shelter study of 78 residents in 1983 again found that 40% had major mental disorders, and another 51% had less severe psychiatric problems…a survey of 345 subjects seeking food assistance in 1983 Phoenix found that about 30% had spent some time in a mental institution.

A quick Fermi calculation from the book’s numbers suggests that maybe 10% of schizophrenics are currently homeless. Again, I 100% acknowledge that this argument is correct and that these are probably accurate statistics about the percent of the homeless who are mentally ill.

Third, it points out that many of these people die of preventable causes. Many freeze to death on cold nights. Cramer notes that deinstituionalization corresponded with a doubling of US hypothermia deaths (although never above 1/500,000 people = 500 people per year) and that anecdotal evidence suggests many of these were mentally ill. Still others commit suicide or otherwise die of their own predictable poor choices. For example:

In another case, a woman with anorexia was admitted to a hospital after she had been involved in a family disagreement and refused to eat. She had lost a great deal of weight but refused to submit to a psychiatric exam, and since a judge felt her condition was not dangerous in an immediate sense, she was allowed to go home. She died from starvation three weeks later.

Again, I 100% acknowledge this sort of thing probably happened and happens quite often.

Fourth, it says that these people are generally weird and scary and can push everyone else out of public places. Many, for example, end up in libraries, the rare sort of public place you can enter without an admission charge. He tells the story of some such library “patrons”:

Mick is having a bad day. He hasn’t misbehaved but sits and stares, glassy-eyed. This is usually the prelude to a seizure. His seizures are easier to deal with than Bob’s, for instance, because he usually has them while seated and so rarely hits his head and bleeds, nor does he ever soil his pants. Bob tends to pace restlessly all day and is often on the move when, without warning, his seizures strike. The last time he went down, he cut his head. The staff has learned to turn him over quickly after he hits the floor, so that his urine does not stain the carpet.

A friend worked at the main branch of the Santa Rosa, California public library in the 1980s and 1990s. She was awash in similar stories of mentally ill people who would urinate in the corners of the library, make frightening noises, sleep at the tables, and generally create an environment that would have been grounds for at least expulsion, if not arrest and commitment, in any American public library in 1960. The library staff was obligated to work with such “patrons” until their actions became clearly criminal. She recounted what happened when she observed that one of these mentally ill patrons was sitting at a table with his pants down to his knees. Her supervisor was obligated by library rules to attempt to first resolve the problem without the police. He approached this exposed “patron” and diplomatically asked “Sir, are you appropriately attired for the library?”

Why was it necessary for librarians to take such a kid glove approach? Attempts to resolve behavioral problems led to lawsuits, such as happened in Morristown, New Jersey. The behavior and offensive smell of a homeless person named Kreimer led to the adoption of a code of conduct prohibiting loitering, “unnecessary staring”, following others around the library, and requiring those using the library to conform to community standards of cleanliness. The ACLU filed suit against this discriminatory code. At trial, Judge Sarokin ruled that the rules were discriminatory, and that the ban on annoying other patrons violated Kreimer’s right to freedom of speech.

This ruling was later overturned on appeal, but apparently the whole series of lawsuits had cost so much money that the mere possibility of a suit from the ACLU led libraries to adopt a policy of tolerating everyone, no matter how filthy, loud, or threatening they might be. Once again, this sounds like the sort of thing that probably happens and I have no doubt the book is telling the truth. One need not blame the homeless and mentally ill for their behavior to acknowledge that this is a potential argument in favor of institutionalizing people so they have less inconvenient places than libraries to spend their time.

Fifth, Cramer argues the deinstitutionalized mentally ill are responsible for a disproportionate amount of crime, including some of the flashiest mass shootings. He notes that of a New York Times list of the 100 most famous rampage killers, 47 had a past history of mental health problems, and 20 had been previously institutionalized. Former psychiatric inpatients are 55 times more likely than the general population to be arrested for murder, and about five times more likely to be arrested for lesser crimes like robbery, rape, and aggravated assault. He cites Bernard Harcourt’s work showing a strong negative correlation between the institutionalization rate and the crime rate – although as I’ve mentioned before, I think these numbers are seriously off and that this is more likely related to lead levels. Nevertheless, the general point that deinstitutionalized mentally ill are at high risk of criminality stands – although Cramer admits that the overwhelming majority will never get in trouble.

IV.

So I agree with almost all of Cramer’s empirical claims. Yes, many deinstitutionalized schizophrenics are not receiving adequate treatment. Yes, many are homeless, either broke or unable to manage their disability money in a rational way. Yes, many are dying of preventable causes like freezing to death. Yes, many are going around public places and threatening people and freaking people out. And yes, many of them (though by no means most) are committing terrible crimes. So how can I disagree with his assessment that deinstitutionalization was a mistake, that Reagan and the hippies and Thomas Szasz were in the wrong, and we need to bring back a strong system of long-term state-run psychiatric hospitals?

Well, let me ask a related question. Should we round up everybody from the ghetto and stick them in prison? This policy would have a number of advantages. Many people in the ghetto are desperately poor and living in terrible conditions. Many die before their time. They often make middle-class people who come across them profoundly uncomfortable. And their crime rate is much higher than that of the non-ghetto population. All the advantages of institutionalizing the mentally ill also apply to institutionalizing people in ghettos.

Against this we have a counterbalancing consideration: it is a horrible idea and it would be really mean and everybody involved would hate it and you have no right to even consider such a thing. This is also how I feel about institutionalizing the mentally ill.

First, a digression. Many of the people Cramer mentions – his brother Ron, his case studies of homeless people who freeze to death on the streets, some of the mass killers – have in fact been institutionalized. Ron was institutionalized the better part of a dozen times. Usually they’re in the hospital for a few weeks to a few months, stabilized on medications, and then released. After their release for one reason or another they come off their medications and then experience whatever catastrophe makes them suitable for inclusion in this book.

So if we want to solve all of the problems Cramer brings up – homelessness, crime, library-bothering, etc – we can’t do it by just having people in institutions for a few months or a few years. The second they set foot out of a hospital in this counterfactual world, they’ll encounter the same problems they encounter in our real one. In other words, this isn’t really about treatment, at least in the sense of “we need better commitment laws so hospitals can treat patients and then help them reintegrate into society.” What Cramer is talking about, if he’s really serious about solving these issues, is lifetime institutionalization.

Making someone spend their entire life in an institution is a pretty big deal, especially if, as Cramer freely admits, they often include “regimented, hopeless conditions” where “care is worse than simply inadequate”. Sometimes we as a society decide that criminals need to spend their entire life in an unpleasant institution because they murdered somebody or something, but it seems excessive to say that somebody should be institutionalized for life merely because they are from a population that has a disproportionate (though still not high!) risk of committing some kind of crime in the future. Once again, if we were in that business we should just imprison people for being born in bad neighborhoods. Yes, it’s a tragedy when an anorexic starves themselves to death. But should we lock up all anorexics forever to prevent that one case?

What about the humanitarian argument that we need to institutionalize schizophrenics so that they don’t end up starving on the street? Here we get into some really thorny moral issues. I tend to go by revealed preferences – schizophrenics have voted with their feet to not be in mental hospitals. If there were voluntary mental hospitals, and schizophrenics chose to live in them, that would be great and I would support them in that choice. If you are contradicting schizophrenics’ expressed preference that they prefer not being in mental hospitals – freezing weather and all – to being in mental hospitals, then you have no right to say you’re doing it for their own sake.

I can see a counterargument: psychotic people are not very good at making decisions. What if they would be happier in a nice warm institution, but they are too crazy to realize this? For example, maybe when the person asks them “Would you like to go to the hospital?” they believe that person is a CIA spy who will be leading them to the firing chamber instead?

I agree this is a possibility and a strong argument. Against it I can only say that many of the psychotic people who don’t want to go to mental hospitals are dragged there anyway, and usually continue to not want to be in the mental hospital after they get there and learn what it is like.

An example from my own life might serve to clarify the odd mix of rational and irrational decision-making I think characterizes these choices. When I was a child, my OCD was much worse. I would do things like close every shutter in my room nine times. I won’t say this was the most rational thing to be doing. But if you with your superior rationality had come in and chained me to my bed so that I couldn’t close my shutters, I would have spent the entire night freaking out because my shutters hadn’t been closed the appropriate nine times and that meant the world was unbearably wrong. Given a mind that will freak out for a whole night if the shutters aren’t closed, and supposing for a second that curing the underlying OCD is not an option, then spending a minute closing the shutters is a perfectly rational decision. Likewise, given the weird collections of fears and sensitivities that characterize the typical psychotic, staying out of a psychiatric hospital may be a perfectly rational decision. And this is even granting the extremely dubious premise that the hospital is not abusive, is not disgusting, is not dictatorial, doesn’t involve drugs with terrible side effects, or any of the other hundred ways a psychiatric hospital can be bad even when your judgment is perfectly intact.

I recently learned many of the homeless in nicer cities have laptops. This makes sense – laptops are really cheap these days, way cheaper than houses, and you can carry them around with you on your back. Psychiatric hospitals, in contrast, do not have laptops. Even if you own a laptop, you may not bring it in, since it is theoretically Usable As A Weapon. You may not bring a cell phone, a tablet or any other form of communication device. Some of the very nice psychiatric hospitals, including the one I work at, have a single computer for thirty residents, which you may use for fifteen minutes a day, with a nurse watching you the whole time to make sure you don’t go on any sites that seem likely to make you upset or emotional. This fact alone makes me, personally, with my as far as I can tell totally intact mind, prefer the thought of homelessness to the thought of lifetime institutionalization. My computer is my only lifeline to most of my friends and the only way I have to express myself, and the thought of trading that away just so I can have a warm bed seems – pardon the expression – insane.

And for me it’s the computer. For other people it’s other things, reasonable by our standards or not. A few weeks ago I was woken up by a call in the middle of the night. A newly admitted patient at the mental hospital where I work was making a scene. She had this thing about using her special pillowcase, and pillowcases weren’t on the hospital’s Special List Of Things It Is Okay To Bring In. Sheets? Absolutely. Blankets? Totally fine. Pillows? Knock yourself out. But nobody had thought about pillowcases, so they were officially banned. And I made it to the ward, still half-asleep, and for a second I couldn’t figure out who was the crazy person, the woman making a William Wallace-esque stand for the right to bring her pillowcase into a hospital, or the woman telling her absolutely not, because it wasn’t on the Magic List. Eventually I asked the nurse if maybe we could just sort of pretend the pillowcase was a very small sheet, and she said that if I specifically ordered her to do so she wasn’t able to contradict a doctor’s orders, and the problem was solved. By which I mean that by the time she figured out something else she needed, my shift would be over and it would be someone else’s problem. Because everything in a mental hospital is like this all the time.

So am I okay with this causing some people to freeze to death? Yes. I don’t think we can be sufficiently sure that institutionalizing schizophrenics is in their own best interest to overcome the burden of proof necessary for overriding someone’s revealed preferences. So if respecting people’s revealed preferences mean some of them go homeless or die, so be it. God help us if we ever systematically decide that people should not be allowed their freedom if the decision carries any discomfort or risk.

I want to stress just how important a decision this is. Back before deinstitutionalization, there were about 500,000 people in US psychiatric institutions, with varying degrees of permanency. Given the increase in the population and mental illness, I expect there are up to a million potentially institutionalizable individuals today. If institutionalization costs the average psychotic 1/3 of a QALY per year (eg moving from poverty to imprisonment on this table) then we’re taking away 300,000 QALYs every year indefinitely. On the other hand, if institutionalization were better for psychotics, they could potentially gain a similar number of QALYs. That makes policy decisions in this area potentially more important than crime, more important than terrorism, more important than education, potentially more important than everything except health care, not starting too many wars, and mass incarceration full stop. These kinds of decisions are the ones you want to be really, really sure about. So far, nothing in My Brother Ron has given me the level of certainty I would need.

I agree kids should have a right to use public libraries without having mentally ill people urinate on them or scream at them. I think the solution in this case is to tell the ACLU to take a chill pill and then let librarians enforce common-sense decency rules, not to lock up a million people for the rest of their lives.

V.

So that leaves the question – what do we do with all of these psychotic people starving on the street? Saying “leave them alone” is all nice and well, but what if they start seeming violent or threatening? Do we leave them alone until the point at which they commit a major crime and they end up in prison for the rest of their lives? What if they’re clearly acting recklessly and about to die? What if we have evidence (maybe from past experience) that they would prefer to be sane and medicated but they’re too far gone to realize it?

The book itself mentions my preferred answer to this conundrum: involuntary outpatient commitment (IOC). This is exactly what it sounds like. If you, let’s say, start trespassing on government property and yelling at police officers (a common way for mentally ill people to come to the attention of the system), and you get brought before a sympathetic judge who wants to help you and doesn’t want to lock you up but would prefer you not do that anymore, he can order an outpatient commitment. This means you’re legally required to see a psychiatrist every so often and maybe get injected with long-acting antipsychotic medication (usually once per month, although I think they’ve recently invented a once-every-three-months version now).

I have seen psychotic patients involved in such programs and they usually do very well. They get the same level of treatment they would in a psychiatric hospital, people will come hunt them down to make sure they don’t miss their appointments or medication dosings, and in the interim they can live wherever they want in whatever conditions they want. If the medications work, which they usually do, then they are hopefully clear-headed enough to either hold down a job or use their disability payments responsibly. If they can’t do that, then it’s probably for the same reason that normal poor people can’t, and nobody says they need to be institutionalized.

Cramer notes that people in IOC programs have half the suicidality rates, half the crime rates, and “substantial reductions in hospitalization, homelessness, arrest, and incarceration.” They are half as likely to be hospitalized, half as likely to be victims of crimes, and “enjoy improved quality of life”.

This isn’t as good as, say, one-tenth the suicidality and hospitalization rates would be. But psychiatry isn’t a discipline with very many miracles. Sometimes the drugs work and sometimes they don’t. Long-term psychotics are notoriously difficult to treat and this is probably about as well as they would be doing in a long-term institution anyway.

Cramer brings this up as part of his political polemic – apparently the same hippies who oppose everything else opposed IOCs, so their success is part of the Grand Narrative Of Hippies Being Proven Wrong. I like hippie-bashing as much as anyone else, but I don’t understand why he doesn’t take this further, say that this is the alternative to reinstitutionalization that he secretly knows we need. He points out that the main reason IOCs are underused is that psychiatrists don’t know about them – I would add that at least in my county there isn’t enough funding to refer enough patients to the program and monitor their medication compliance and so on. But I guarantee you that publicizing the option to psychiatrists and expanding the program is a lot cheaper than reinstitutionalizing people would be.

(my hospital charges $1,000/day/inpatient, though goodness only knows how much of that insurance companies actually pay. Cramer notes that the prison system usually costs $50,000/year/mentally ill prisoner. My guess is that the costs of institutionalization are somewhere around that order of magnitude.)

So in my ideal world, psychotic people who aren’t bothering anybody can do what they want – preferably with the option of voluntary psychiatric hospitalization available, and with some pressure to at least try it once and get a feel for what it’s like. Psychotic people who are bothering other people can get outpatient treatment once every couple of months and remain medicated and monitored by professionals. Preferably there would also be some kind of concept of a psychiatric living will – that is, some way for people who are not yet mentally ill, or who are currently being managed on drugs, to express a wish to be stabilized if they ever become mentally ill so that they can make their long-run choices from a position of sanity.

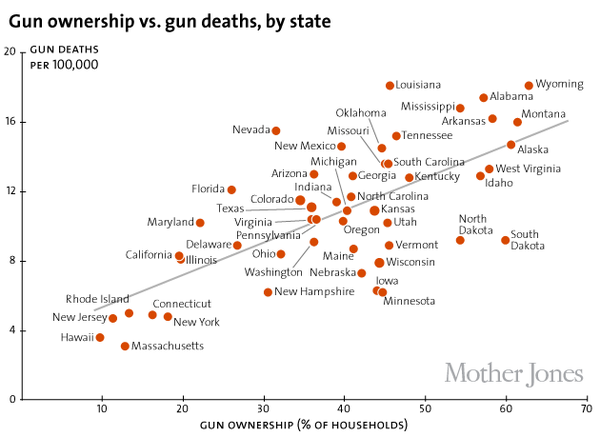
I acknowledge this is not the ideal world. I acknowledge there are some people who really need institutionalization – people who are constantly violent, who have zero concept of social rules and will scream at anyone they meet, people who are catatonic or need extraordinarily complicated medication regimens that can’t be handled in a normal environment. I’ve referred some of these people to involuntary long-term institutions (which still exist for these kinds of extreme situations), I don’t feel guilty at all, and in most cases I am pretty sure the general public would be pretty grateful to me if they knew the gory details.

But for a million people, most of whom aren’t bothering anybody and just want to be able to live a half-decent life outside the walls of a locked facility? There has to be a better solution than that.

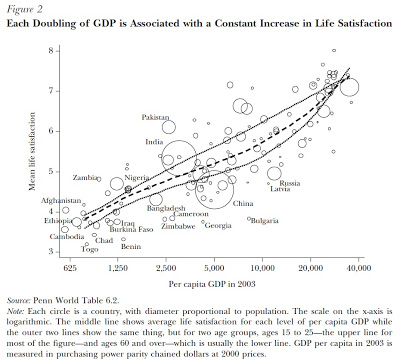
# Beware Regional Scatterplots

[Epistemic status: Not original, but worth mentioning]

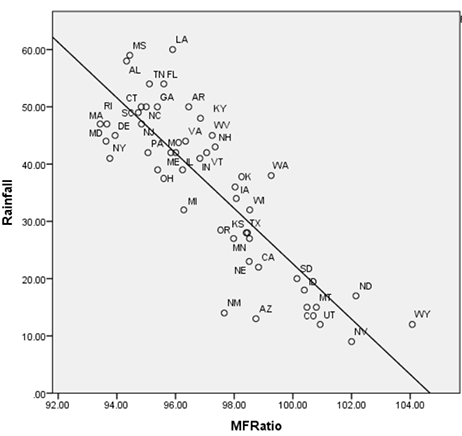
I’ve been using scatterplots of different states and countries a lot here lately. For example, this one in the discussion about guns:



And this one in the discussion about national happiness:



Hopefully we already know that we should worry about confounders like income and race and those kinds of things when we’re looking at a graph like this. But recently I learned it’s even worse than that. Consider for example this:



This is the average yearly rainfall in the lower 48 US states vs. their gender balance (measured in number of men per 100 women). The correlation is about r = 0.84 (p ≤ 0.0001), much higher than anyone’s ever found between guns and crime, or income and happiness, or most other things people make regional scatterplots about. So what’s going on? Do women cause rainfall? Does rain drive men away? Or is there some confounder that causes both rain and womanhood?

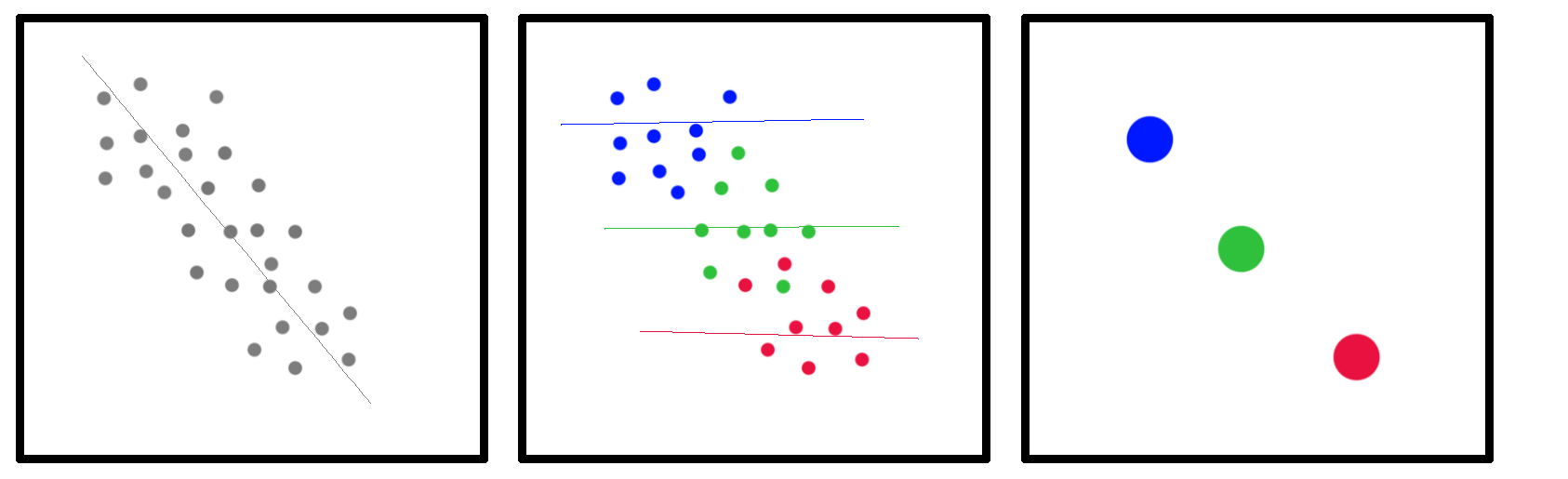
I don’t think it’s any of these things. I think it’s a coincidence.

“But you said p ≤ 0.0001! There are forty-eight data points and the fit is almost perfect! How could it be a coincidence?”

But I don’t think there are forty-eight data points. I think there are three data points. For 48 data points to all lie the same line is very impressive. For three data points to all lie on the same line is much less so.

I think that Southern states have more women (probably because they have a higher male incarceration rate, and incarcerated men aren’t counted) and more rainfall. Mountainous western states have more men (probably because the jobs there tend to be in manly mining/forestry type industries) and are pretty dry. And other states are somewhere between those two extremes.

Within these regional categories the rainfall/gender relationship is random – on the scatterplot it would look like a circle. But between these three regional categories the rainfall/gender relationship is very strong, making the whole chart consist of three circles in a line. Crucially, because these are kind of amorphous circles and they blend into each other, you can’t tell that that’s what’s going on. Here’s my graphic depiction of this:



In the first box, the gray points show what looks like a very significant correlation – the further right you go on the horizontal axis, the further down you go on the vertical axis. The gray trend line confirms the strong relationship. In our US state example, this was a correlation between many women and high rainfall.

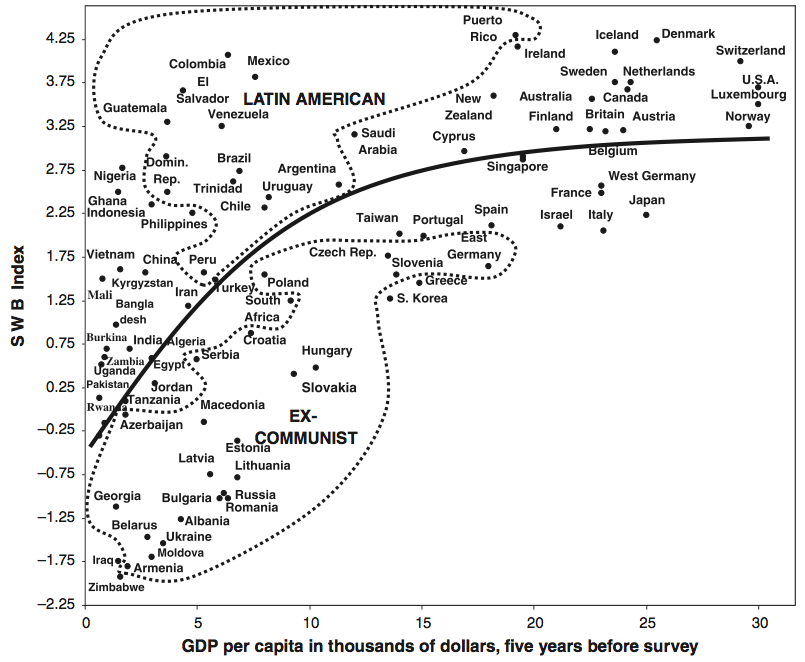
In the second box, the gray points are revealed to be grouped into three regions: blue, green, and red. In our US state example, blue is the female-skewed and rainy Southern states, red is the male-skewed and dry Western states, and green are all the other states with pretty average rainfalls and gender balances. Within each region, there’s no relationship between rainfall and gender, as shown by the horizontal red, green, and blue trend lines.

In the third box, we see what I’d argue is the correct interpretation of the data. There are three big data points – the South, the West, and the Rest – and they do sort of form a line but nobody cares about a line between three data points.

If I go back to my statistics packet and repeat the rainfall/gender correlation with only three points – the Southern average, the Western average, and the Other average – I still get r = -0.84, but now p = 0.3. The statistics have no reason to think it’s anything other than pure coincidence – and indeed, with that small a sample size, why would they?

I think most real studies are smart enough to control for this – although it’s really hard to determine how exactly you should be doing that and leaves a lot of wiggle room for people who want to fudge their way to a preferred result. But basic scatterplots do not control for it, and so almost every regional scatterplot is suspect.

This is why I was happy to see the income/happiness correlation broken down further:



This makes it clear that the income/happiness relationship is primarily cluster-driven, with clusters of ex-Communist countries, Latin American countries, and Euro/Anglosphere countries (if you’re willing to do some more work, you can sort of make out clusters of African and Asian countries). None of these clusters show a strong income/happiness relationship except for the ex-Communist one, which suggests this might be the same kind of confounding as the rainfall/gender example above.

…unless I’m biased and reading too much into this. It’s really easy to change your conclusion just by changing your clusters. For example, if Puerto Rico counts as Latin American, then that creates a pretty impressive happiness/income relationship within that cluster. If it counts as Euro/Anglosphere – it’s part of the US, after all – then there is no income/happiness relationship within either cluster. So which is it?

Or what about this? I claim that the apparent income/happiness relationship within the ex-Communist countries is actually an artifact of Europeanness. The richest ex-Communist countries, like East Germany and the Czech Republic – are also the happiest only because they are closest to Western Europe, which is both happier and richer than the rest of the world. Likewise, the poorest ex-Communist countries, like Armenia and Georgia, are also the unhappiest only because they are the furthest and least Western European of the bunch.

Once we start going there, we can pretty much prove or disprove anything we want based on our own intuitions about how to group things. I am suspicious of this, but I’m also equally suspicious of not doing that – do you really just want to let it pass that Puerto Rico, the closest Latin American country to the Euro/Anglosphere cluster, is also politically Euro/Anglosphere?

Overall there is no good answer and I would recommend against drawing any strong causal conclusions from a scatterplot unless someone has very carefully addressed these concern.

EDIT: Inty gives a great (ie horrifying) example in the comments, and Theo Jones discusses more formal tests of spatial autocorrelation. Some people bring up the possibility that some of the rainfall/gender relationship is causal after all, since drier states will have less farming and be forced to turn to mining/forestry to support themselves; this is possible but probably doesn’t explain the whole relationship, and even if I’m wrong about this one the point is still important.

# The Ideology Is Not The Movement

I.

Why is there such a strong Sunni/Shia divide?

I know the Comparative Religion 101 answer. The early Muslims were debating who was the rightful caliph. Some of them said Abu Bakr, others said Ali, and the dispute has been going on ever since. On the other hand, that was fourteen hundred years ago, both candidates are long dead, and there’s no more caliphate. You’d think maybe they’d let the matter rest.

Sure, the two groups have slightly different hadith and schools of jurisprudence, but how many Muslims even know which school of jurisprudence they’re supposed to be following? It seems like a pretty minor thing to have centuries of animus over.

And so we return again to Robbers’ Cave:

The experimental subjects — excuse me, “campers” — were 22 boys between 5th and 6th grade, selected from 22 different schools in Oklahoma City, of stable middle-class Protestant families, doing well in school, median IQ 112. They were as well-adjusted and as similar to each other as the researchers could manage.

The experiment, conducted in the bewildered aftermath of World War II, was meant to investigate the causes—and possible remedies—of intergroup conflict. How would they spark an intergroup conflict to investigate? Well, the 22 boys were divided into two groups of 11 campers, and —

— and that turned out to be quite sufficient.

The researchers’ original plans called for the experiment to be conducted in three stages. In Stage 1, each group of campers would settle in, unaware of the other group’s existence. Toward the end of Stage 1, the groups would gradually be made aware of each other. In Stage 2, a set of contests and prize competitions would set the two groups at odds.

They needn’t have bothered with Stage 2. There was hostility almost from the moment each group became aware of the other group’s existence: They were using our campground, our baseball diamond. On their first meeting, the two groups began hurling insults. They named themselves the Rattlers and the Eagles (they hadn’t needed names when they were the only group on the campground).

When the contests and prizes were announced, in accordance with pre-established experimental procedure, the intergroup rivalry rose to a fever pitch. Good sportsmanship in the contests was evident for the first two days but rapidly disintegrated.

The Eagles stole the Rattlers’ flag and burned it. Rattlers raided the Eagles’ cabin and stole the blue jeans of the group leader, which they painted orange and carried as a flag the next day, inscribed with the legend “The Last of the Eagles”. The Eagles launched a retaliatory raid on the Rattlers, turning over beds, scattering dirt. Then they returned to their cabin where they entrenched and prepared weapons (socks filled with rocks) in case of a return raid. After the Eagles won the last contest planned for Stage 2, the Rattlers raided their cabin and stole the prizes. This developed into a fistfight that the staff had to shut down for fear of injury. The Eagles, retelling the tale among themselves, turned the whole affair into a magnificent victory—they’d chased the Rattlers “over halfway back to their cabin” (they hadn’t).

Each group developed a negative stereotype of Them and a contrasting positive stereotype of Us. The Rattlers swore heavily. The Eagles, after winning one game, concluded that the Eagles had won because of their prayers and the Rattlers had lost because they used cuss-words all the time. The Eagles decided to stop using cuss-words themselves. They also concluded that since the Rattlers swore all the time, it would be wiser not to talk to them. The Eagles developed an image of themselves as proper-and-moral; the Rattlers developed an image of themselves as rough-and-tough.

If the researchers had decided that the real difference between the two groups was that the Eagles were adherents of Eagleism, which held cussing as absolutely taboo, and the Rattlers adherents of Rattlerism, which held it a holy duty to cuss five times a day – well, that strikes me as the best equivalent to saying that Sunni and Shia differ over the rightful caliph.

II.

Nations, religions, cults, gangs, subcultures, fraternal societies, internet communities, political parties, social movements – these are all really different, but they also have some deep similarities. They’re all groups of people. They all combine comradery within the group with a tendency to dislike other groups of the same type. They all tend to have a stated purpose, like electing a candidate or worshipping a deity, but also serve a very important role as impromptu social clubs whose members mostly interact with one another instead of outsiders. They all develop an internal culture such that members of the groups often like the same foods, wear the same clothing, play the same sports, and have the same philosophical beliefs as other members of the group – even when there are only tenuous links or no links at all to the stated purpose. They all tend to develop sort of legendary histories, where they celebrate and exaggerate the deeds of the groups’ founders and past champions. And they all tend to inspire something like patriotism, where people are proud of their group membership and express that pride through conspicuous use of group symbols, group songs, et cetera. For better or worse, the standard way to refer to this category of thing is “tribe”.

Tribalism is potentially present in all groups, but levels differ a lot even in groups of nominally the same type. Modern Belgium seems like an unusually non-tribal nation; Imperial Japan in World War II seems like an unusually tribal one. Neoliberalism and market socialism seem like unusually non-tribal political philosophies; communism and libertarianism seem like unusually tribal ones. Corporations with names like Amalgamated Products Co probably aren’t very tribal; charismatic corporations like Apple that become identities for their employees and customers are more so. Cults are maybe the most tribal groups that exist in the modern world, and those Cult Screening Tools make good measures for tribalism as well.

The dangers of tribalism are obvious; for example, fascism is based around dialing a country’s tribalism up to eleven, and it ends poorly. If I had written this essay five years ago, it would be be titled “Why Tribalism Is Stupid And Needs To Be Destroyed”. Since then, I’ve changed my mind. I’ve found that I enjoy being in tribes as much as anyone else.

Part of this was resolving a major social fallacy I’d had throughout high school and college, which was that the correct way to make friends was to pick the five most interesting people I knew and try to befriend them. This almost never worked and I thought it meant I had terrible social skills. Then I looked at what everyone else was doing, and I found that instead of isolated surgical strikes of friendship, they were forming groups. The band people. The mock trial people. The football team people. The Three Popular Girls Who Went Everywhere Together. Once I tried “falling in with” a group, friendship became much easier and self-sustaining precisely because of all of the tribal development that happens when a group of similar people all know each other and have a shared interest. Since then I’ve had good luck finding tribes I like and that accept me – the rationalists being the most obvious example, but even interacting with my coworkers on the same hospital unit at work is better than trying to find and cultivate random people.

Some benefits of tribalism are easy to explain. Tribalism intensifies all positive and prosocial feelings within the tribe. It increases trust within the tribe and allows otherwise-impossible forms of cooperation – remember Haidt on the Jewish diamond merchants outcompeting their rivals because their mutual Judaism gave them a series of high-trust connections that saved them costly verification procedures? It gives people a support network they can rely on when their luck is bad and they need help. It lets you “be yourself” without worrying that this will be incomprehensible or offensive to somebody who thinks totally differently from you. It creates an instant densely-connected social network of people who mostly get along with one another. It makes people feel like part of something larger than themselves, which makes them happy and can (provably) improves their physical and mental health.

Others are more complicated. I can just make motions at a feeling that “what I do matters”, in the sense that I will probably never be a Beethoven or a Napoleon who is very important to the history of the world as a whole, but I can do things that are important within the context of a certain group of people. All of this is really good for my happiness and mental health. When people talk about how modern society is “atomized” or “lacks community” or “doesn’t have meaning”, I think they’re talking about a lack of tribalism, which leaves people all alone in the face of a society much too big to understand or affect. The evolutionary psychology angle here is too obvious to even be worth stating.

And others are entirely philosophical. I think some people would say that wanting to have a tribe is like wanting to have a family – part of what it means to be human – and demands to justify either are equally wrong-headed.

Eliezer thinks every cause wants to be a cult. I would phrase this more neutrally as “every cause wants to be a tribe”. I’ve seen a lot of activities go through the following cycle:

1. Let’s get together to do X  
2. Let’s get together to do X, and have drinks afterwards  
3. Let’s get together to discuss things from an X-informed perspective  
4. Let’s get together to discuss the sorts of things that interest people who do X  
5. Let’s get together to discuss how the sort of people who do X are much better than the sort of people who do Y.  
6. Dating site for the sort of people who do X  
7. Oh god, it was so annoying, she spent the whole date talking about X.  
8. X? What X?

This can happen over anything or nothing at all. Despite the artificial nature of the Robbers’ Cove experiment, its groups are easily recognized as tribes. Indeed, the reason this experiment is so interesting is that it shows tribes in their purest form; no veneer of really being about pushing a social change or supporting a caliph, just tribes for tribalism’s sake.

III.

Scholars call the process of creating a new tribe “ethnogenesis” – Robbers’ Cave was artificially inducing ethnogenesis to see what would happen. My model of ethnogenesis involves four stages: pre-existing differences, a rallying flag, development, and dissolution.

Pre-existing differences are the raw materials out of which tribes are made. A good tribe combines people who have similar interests and styles of interaction even before the ethnogenesis event. Any description of these differences will necessarily involve stereotypes, but a lot of them should be hard to argue. For example, atheists are often pretty similar to one another even before they deconvert from their religion and officially become atheists. They’re usually nerdy, skeptical, rational, not very big on community or togetherness, sarcastic, well-educated. At the risk of going into touchier territory, they’re pretty often white and male. You take a sample of a hundred equally religious churchgoers and pick out the ones who are most like the sort of people who are atheists even if all of them are 100% believers. But there’s also something more than that. There are subtle habits of thought, not yet described by any word or sentence, which atheists are more likely to have than other people. It’s part of the reason why atheists need atheism as a rallying flag instead of just starting the Skeptical Nerdy Male Club.

The rallying flag is the explicit purpose of the tribe. It’s usually a belief, event, or activity that get people with that specific pre-existing difference together and excited. Often it brings previously latent differences into sharp relief. People meet around the rallying flag, encounter each other, and say “You seem like a kindred soul!” or “I thought I was the only one!” Usually it suggests some course of action, which provides the tribe with a purpose. For atheists, the rallying flag is not believing in God. Somebody says “Hey, I don’t believe in God, if you also don’t believe in God come over here and we’ll hang out together and talk about how much religious people suck.” All the atheists go over by the rallying flag and get very excited about meeting each other. It starts with “Wow, you hate church too?”, moves on to “Really, you also like science fiction?”, and ends up at “Wow, you have the same undefinable habits of thought that I do!”

Development is all of the processes by which the fledgling tribe gains its own culture and history. It’s a turning-inward and strengthening-of-walls, which transforms it from ‘A Group Of People Who Do Not Believe In God And Happen To Be In The Same Place’ to ‘The Atheist Tribe’. For example, atheists have symbols like that ‘A’ inside an atom. They have jokes and mascots like Russell’s Teapot and the Invisible Pink Unicorn. They have their own set of heroes, both mythologized past heroes like Galileo and controversial-but-undeniably-important modern heroes like Richard Dawkins and Daniel Dennett. They have celebrities like P.Z. Myers and Hemant Mehta. They have universally-agreed-upon villains to be booed and hated, like televangelists or the Westboro Baptist Church. They have grievances, like all the times that atheists have been fired or picked on by religious people, and all the laws about pledging allegiance to one nation under God and so on. They have stereotypes about themselves – intelligent, helpful, passionate – and stereotypes about their outgroups – deluded, ignorant, bigoted.

Dissolution is optional. The point of the previous three steps is to build a “wall” between the tribe and the outside, a series of systematic differences that let everybody know which side they’re on. If a tribe was never really that different from the surrounding population, stops caring that much about its rallying flag, and doesn’t develop enough culture, then the wall fails and the members disperse into the surrounding population. The classic example is the assimilation of immigrant groups like Irish-Americans, but history is littered with failed communes, cults, and political movements. Atheism hasn’t quite dissolved yet, but occasionally you see hints of the process. A lot of the comments around “Atheism Plus” centered around this idea of “Okay, talking about how there’s no God all the time has gotten boring, plus nobody interesting believes in God anymore anyway, so let’s become about social justice instead”. The parts of atheism who went along with that message mostly dissolved into the broader social justice community – there are a host of nominally atheist blogs that haven’t talked about anything except social justice in months. Other fragments of the atheist community dissolved into transhumanism, or libertarianism, or any of a number of other things. Although there’s still an atheist community, it no longer seems quite as vibrant and cohesive as it used to be.

We can check this four-stage model by applying it to the Sunni and Shia and seeing if it sticks.

I know very little about early Islam and am relying on sources that might be biased, so don’t declare a fatwa against me if I turn out to be wrong, but it looks like from the beginning there were big pre-existing differences between proto-Shia and proto-Sunni. A lot of Ali’s earliest supporters were original Muslims who had known Mohammed personally, and a lot of Abu Bakr’s earliest supporters were later Muslims high up in the Meccan/Medinan political establishment who’d converted only after it became convenient to do so. It’s really easy to imagine cultural, social, and personality differences between these two groups. Probably members in each group already knew one another pretty well, and already had ill feelings towards members of the other, without necessarily being able to draw the group borders clearly or put their exact differences into words. Maybe it was “those goody-goodies who are always going on about how close to Mohammed they were but have no practical governing ability” versus “those sellouts who don’t really believe in Islam and just want to keep playing their political games”.

Then came the rallying flag: a political disagreement over the succession. One group called themselves “the party of Ali”, whose Arabic translation “Shiatu Ali” eventually ended up as just “Shia”. The other group won and called itself “the traditional orthodox group”, in Arabic “Sunni”. Instead of a vague sense of “I wonder whether that guy there is one of those goody-goodies always talking about Mohammed, or whether he’s a practical type interested in good governance”, people could just ask “Are you for Abu Bakr or Ali?” and later “Are you Sunni or Shia?” Also at some point, I’m not exactly sure how, most of the Sunni ended up in Arabia and most of the Shia ended up in Iraq and Iran, after which I think some pre-existing Iraqi/Iranian vs. Arab cultural differences got absorbed into the Sunni/Shia mix too.

Then came development. Both groups developed elaborate mythologies lionizing their founders. The Sunni got the history of the “rightly-guided caliphs”, the Shia exaggerated the first few imams to legendary proportions. They developed grievances against each other; according to Shia history, the Sunnis killed eleven of their twelve leaders, with the twelfth escaping only when God directly plucked him out of the world to serve as a future Messiah. They developed different schools of hadith interpretation and jurisprudence and debated the differences ad nauseum with each other for hundreds of years. A lot of Shia theology is in Farsi; Sunni theology is entirely in Arabic. Sunni clergy usually dress in white; Shia clergy usually dress in black and green. Not all of these were deliberately done in opposition to one another; most were just a consequence of the two camps being walled off from one another and so allowed to develop cultures independently.

Obviously the split hasn’t dissolved yet, but it’s worth looking at similar splits that have. Catholicism vs. Protestantism is still a going concern in a few places like Ireland, but it’s nowhere near the total wars of the 17th century or even the Know-Nothing-Parties of the 19th. Consider that Marco Rubio is Catholic, but nobody except Salon particularly worries about that or says that it will make him unsuitable to lead a party representing the interests of very evangelical Protestants. Heck, the same party was happy to nominate Mitt Romney, a Mormon, and praise him for his “Christian faith”. Part of it is the subsumption of those differences into a larger conflict – most Christians acknowledge Christianity vs. atheism to be a bigger deal than interdenominational disputes these days – and part of it is that everyone of every religion is so influenced by secular American culture that the religions have been reduced to their rallying flags alone rather than being fully developed tribes at this point. American Sunni and Shia seem to be well on their way to dissolving into each other too.

IV.

I want to discuss a couple of issues that I think make more sense once you understand the concept of tribes and rallying flags:

1. Disability: I used to be very confused by disabled people who insist on not wanting a “cure” for their condition. Deaf people and autistic people are the two classic examples, and sure enough we find articles like Not All Deaf People Want To Be Cured and They Don’t Want An Autism Cure. Autistic people can at least argue their minds work differently rather than worse, but being deaf seems to be a straight-out disadvantage: the hearing can do anything the deaf can, and can hear also. A hearing person can become deaf at any time just by wearing earplugs, but a deaf person can’t become hearing, at least not without very complicated high-tech surgeries.

When I asked some deaf friends about this, they explained that they had a really close-knit and supportive deaf culture, and that most of their friends, social events, and ways of relating to other people and the world were through this culture. This made sense, but I always wondered: if you were able to hear, couldn’t you form some other culture? If worst came to worst and nobody else wanted to talk to you, couldn’t you at least have the Ex-Deaf People’s Club?

I don’t think so. Deafness acts as a rallying flag that connects people, gives them a shared foundation to build culture off of, and walls the group off from other people. If all deaf people magically became able to hear, their culture would eventually drift apart, and they’d be stuck without an ingroup to call their own.

Part of this is reasonable cost-benefit calculation – our society is so vast and atomized, and forming real cohesive tribes is so hard, that they might reasonably expect it would be a lot of trouble to find another group they liked as much as the deaf community. But another part of this seems to be about an urge to cultural self-preservation.

2. Genocide: This term is kind of overused these days. I always thought of it as meaning literally killing every member of a certain group – the Holocaust, for example – but the new usage includes “cultural genocide”. For example, autism rights advocates sometimes say that anybody who cured autism would be committing genocide – this is of course soundly mocked, but it makes sense if you think of autistic people as a tribe that would be dissolved absent its rallying flag. The tribe would be eliminated – thus “cultural genocide” is a reasonable albeit polemical description.

It seems to me that people have an urge toward cultural self-preservation which is as strong or stronger as the urge to individual self-preservation. Part of this is rational cost-benefit calculation – if someone loses their only tribe and ends up alone in the vast and atomized sea of modern society, it might take years before they can find another tribe and really be at home there. But a lot of it seems to be beyond that, an emotional certainty that losing one’s culture and having it replaced with another is not okay, any more than being killed at the same time someone else has a baby is okay. Nor do I think this is necessarily irrational; locating the thing whose survival you care about in the self rather than the community is an assumption, and people can make different assumptions without being obviously wrong.

3. Rationalists: The rationalist community is a group of people (of which I’m a part) who met reading the site Less Wrong and who tend to hang out together online, sometimes hang out together in real life, and tend to befriend each other, work with each other, date each other, and generally move in the same social circles. Some people call it a cult, but that’s more a sign of some people having lost vocabulary for anything between “totally atomized individuals” and “outright cult” than any particular cultishness.

But people keep asking me what exactly the rationalist community is. Like, what is the thing they believe that makes them rationalists? It can’t just be about being rational, because loads of people are interested in that and most of them aren’t part of the community. And it can’t just be about transhumanism because there are a lot of transhumanists who aren’t rationalists, and lots of rationalists who aren’t transhumanists. And it can’t just be about Bayesianism, because pretty much everyone, rationalist or otherwise, agrees that is a kind of statistics that is useful for some things but not others. So what, exactly, is it?

This question has always bothered me, but now after thinking about it a lot I finally have a clear answer: rationalism is the belief that Eliezer Yudkowsky is the rightful caliph.

No! Sorry! I think “the rationalist community” is a tribe much like the Sunni or Shia that started off with some pre-existing differences, found a rallying flag, and then developed a culture.

The pre-existing differences range from the obvious to the subtle. A lot of rationalists are mathematicians, programmers, or computer scientists. The average IQ is in the 130s. White men are overrepresented, but so are LGBT and especially transgender people. But there’s more. Nobody likes the Myers-Briggs test, but I continue to find it really interesting that rationalists have some Myers-Briggs types (INTJ/INTP) at ten times the ordinary rate, and other types (ISFJ/ESFP) at only one one-hundredth the ordinary rate. Myers-Briggs doesn’t cleave reality at its joints, but if it measures anything at all about otherwise hard-to-explain differences in thinking styles, the rationalist community heavily selects for those same differences. Sure enough, I am constantly running into people who say “This is the only place where I’ve ever found people who think like me” or “I finally feel understood”.

The rallying flag was the Less Wrong Sequences. Eliezer Yudkowsky started a blog (actually, borrowed Robin Hanson’s) about cognitive biases and how to think through them. Whether or not you agreed with him or found him enlightening loaded heavily on those pre-existing differences, so the people who showed up in the comment section got along and started meeting up with each other. “Do you like Eliezer Yudkowsky’s blog?” became a useful proxy for all sorts of things, eventually somebody coined the word “rationalist” to refer to people who did, and then you had a group with nice clear boundaries.

The development is everything else. Obviously a lot of jargon sprung up in the form of terms from the blog itself. The community got heroes like Gwern and Anna Salamon who were notable for being able to approach difficult questions insightfully. It doesn’t have much of an outgroup yet – maybe just bioethicists and evil robots. It has its own foods – MealSquares, that one kind of chocolate everyone in Berkeley started eating around the same time – and its own games. It definitely has its own inside jokes. I think its most important aspect, though, is a set of shared mores – everything from “understand the difference between ask and guess culture and don’t get caught up in it” to “cuddling is okay” to “don’t misgender trans people” – and a set of shared philosophical assumptions like utilitarianism and reductionism.

I’m stressing this because I keep hearing people ask “What is the rationalist community?” or “It’s really weird that I seem to be involved in the rationalist community even though I don’t share belief X” as if there’s some sort of necessary-and-sufficient featherless-biped-style ideological criterion for membership. This is why people are saying “Lots of you aren’t even singularitarians, and everyone agrees Bayesian methods are useful in some places and not so useful in others, so what is your community even about?” But once again, it’s about Eliezer Yudkowsky being the rightful caliph it’s not necessarily about anything.

If you take only one thing from this essay, it’s that communities are best understood not logically but historically. If you want to understand the Shia, don’t reflect upon the true meaning of Ali being the rightful caliph, understand that a dispute involving Ali initiated ethnogenesis, the resulting culture picked up a bunch of features and became useful to various people, and now here we are. If you want to understand the rationalist community, don’t ask exactly how near you have to think the singularity has to be before you qualify for membership, focus on the fact that some stuff Eliezer Yudkowsky wrote led to certain people identifying themselves as “rationalists” and for various reasons I enjoy dinner parties with those people about 10000% more interesting than dinner parties with randomly selected individuals.

nostalgebraist actually summed this up really well: “Maybe the real rationalism was the friends we made along the way.” Maybe that’s the real Shia Islam too, and the real Democratic Party, and so on.

4. Evangelical And Progressive Religion: There seems to be a generational process, sort of like Harold Lee’s theory of immigrant assimilation, by which religions dissolve. The first generation believes everything literally. The second generation believes that the religion might not be literally true, but it’s an important expression of universal values and they still want to follow the old ways and participate in the church/temple/mosque/mandir community. The third generation is completely secularized.

This was certainly my family’s relationship with Judaism. My great-great-grandfather was so Jewish that he left America and returned to Eastern Europe because he was upset at American Jews for not being religious enough. My great-grandfather stayed behind in America but remained a very religious Jew. My grandparents attend synagogue when they can remember, speak a little Yiddish, and identify with the traditions. My parents went to a really liberal synagogue where the rabbi didn’t believe in God and everyone just agreed they were going through the motions. I got Bar Mitzvahed when I was a kid but haven’t been to synagogue in years. My children probably won’t even have that much.

So imagine you’re an evangelical Christian. All the people you like are also evangelical Christians. Most of your social life happens at church. Most of your good memories involve things like Sunday school and Easter celebrations, and even your bittersweet memories are things like your pastor speaking at your parents’ funeral. Most of your hopes and dreams involve marrying someone and having kids and then sharing similarly good times with them. When you try to hang out with people who aren’t evangelical Christians, they seem to think really differently than you do, and not at all in a good way. A lot of your happiest intellectual experiences involve geeking out over different Bible verses and the minutiae of different Christian denominations.

Then somebody points out to you that God probably doesn’t exist. And even if He does, it’s probably in some vague and complicated way, and not the way that means that the Thrice-Reformed Meta-Baptist Church and only the Thrice-Reformed Meta-Baptist Church has the correct interpretation of the Bible and everyone else is wrong.

On the one hand, their argument might be convincing. On the other, you are pretty sure that if everyone agreed on this, your culture would be destroyed. Sure, your kids could be Christmas-and-Easter-Christians who still enjoy the cultural aspects and derive personal meaning from the Bible. But you’re pretty sure that within a couple of generations your descendents would be exactly as secular as anyone else. Absent the belief that serves as your culture’s wall against the outside world, it would dissolve without a trace into the greater homogeneity of Western liberal society. So, do you keep believing a false thing? Or do you give up on everything you love and enjoy and dissolve into a culture that mostly hates and mocks people like you? There’s no good choice. This is why it sucks that things like religion and politics are both rallying flags for tribes, and actual things that there may be a correct position on.

5. Religious Literalism: One comment complaint I heard during the height of the Atheist-Theist Online Wars was that atheists were a lot like fundamentalists. Both wanted to interpret the religious texts in the most literal possible way.

Being on the atheist side of these wars, I always wanted to know: well, why wouldn’t you? Given that the New Testament clearly says you have to give all your money to the poor, and the Old Testament doesn’t say anything about mixing meat and milk, maybe religious Christians should start giving everything to the poor and religious Jews should stop worrying so much about which dishes to use when?

But I think this is the same mistake as treating the Sunni as an organization dedicated to promoting an Abu Bakr caliphate. The holy book is the rallying flag for a religion, but the religion is not itself about the holy book. The rallying flag created a walled-off space where people could undergo the development process and create an independent culture. That independent culture may diverge significantly from the holy book.

I think that very neurotypical people naturally think in terms of tribes, and the idea that they have to retool their perfectly functional tribe to conform to the exact written text of its holy book or constitution or stated political ideology or something seems silly to them. I think that less neurotypical people – a group including many atheists – think less naturally in terms of tribes and so tend to take claims like “Christianity is about following the Bible” at face value. But Christianity is about being part of the Christian tribe, and although that tribe started around the Bible, maintains its coherence because of the Bible, and is of course naturally influenced by it, if it happens to contradict the Bible in some cases that’s not necessarily surprising or catastrophic.

This is also why I’m not really a fan of debates over whether Islam is really “a religion of peace” or “a religion of violence”, especially if those debates involve mining the Quran for passages that support one’s preferred viewpoint. It’s not just because the Quran is a mess of contradictions with enough interpretive degrees of freedom to prove anything at all. It’s not even because Islam is a host of separate cultures as different from one another as Unitarianism is from the Knights Templar. It’s because the Quran just created the space in which the Islamic culture could evolve, but had only limited impact on that evolution. As well try to predict the warlike or peaceful nature of the United Kingdom by looking at a topographical map of Great Britain.

6. Cultural Appropriation: Thanks to some people who finally explained this to me in a way that made sense. When an item or artform becomes the rallying flag for a tribe, it can threaten the tribe if other people just want to use it as a normal item or artform.

Suppose that rappers start with pre-existing differences from everyone else. Poor, male, non-white minority, lots of experience living in violent places, maybe a certain philosophical outlook towards their condition. Then they get a rallying flag: rap music. They meet one another, like one another. The culture undergoes further development: the lionization of famous rappers, the development of a vocabulary of shared references. They get all of the benefits of being in a tribe like increased trust, social networking, and a sense of pride and identity.

Now suppose some rich white people get into rap. Maybe they get into rap for innocuous reasons: rap is cool, they like the sound of it. Fine. But they don’t share the pre-existing differences, and they can’t be easily assimilated into the tribe. Maybe they develop different conventions, and start saying that instead of being about the struggles of living in severe poverty, rap should be about Founding Fathers. Maybe they start saying the original rappers are bad, and they should stop talking about violence and bitches because that ruins rap’s reputation. Since rich white people tend to be be good at gaining power and influence, maybe their opinions are overrepresented at the Annual Rap Awards, and all of a sudden you can’t win a rap award unless your rap is about the Founding Fathers and doesn’t mention violence (except Founding-Father-related duels). All of a sudden if you try to start some kind of impromptu street rap-off, you’re no longer going to find a lot of people like you whom you instantly get along with and can form a high-trust community. You’re going to find half people like that, and half rich white people who strike you as annoying and are always complaining that your raps don’t feature any Founding Fathers at all. The rallying flag fails and the tribe is lost as a cohesive entity.

7. Fake Gamer Girls: A more controversial example of the same. Video gaming isn’t just a fun way to pass the time. It also brings together a group of people with some pre-existing common characteristics: male, nerdy, often abrasive, not very successful, interested in speculation, high-systematizing. It gives them a rallying flag and creates a culture which then develops its own norms, shared reference points, internet memes, webcomics, heroes, shared gripes, even some unique literature. Then other people with very different characteristics and no particular knowledge of the culture start enjoying video games just because video games are fun. Since the Gamer Tribe has no designated cultural spaces except video games forums and magazines, they view this as an incursion into their cultural spaces and a threat to their existence as a tribe.

Stereotypically this is expressed as them getting angry when girls start playing video games. One can argue that it’s unfair to infer tribe membership based on superficial characteristics like gender – in the same way it might be unfair for the Native Americans to assume someone with blonde hair and blue eyes probably doesn’t follow the Old Ways – but from the tribe’s perspective it’s a reasonable first guess.

I’ve found gamers to get along pretty well with women who share their culture, and poorly with men who don’t – but admit that the one often starts from an assumption of foreignness and the other from an assumption of membership. More important, I’ve found the idea of the rejection of the ‘fake gamer girl’, real or not, raised more as a libel by people who genuinely do want to destroy gamer culture, in the sense of cleansing video-game-related spaces of a certain type of person/culture and making them entirely controlled by a different type of person/culture, in much the same way that a rich white person who says any rapper who uses violent lyrics needs to be blacklisted from the rap world has a clear culture-change project going on.

These cultural change projects tend to be framed in terms of which culture has the better values, which I think is a limited perspective. I think America has better values than Pakistan does, but that doesn’t mean I want us invading them, let alone razing their culture to the ground and replacing it with our own.

8. Subcultures And Posers: Obligatory David Chapman link. A poser is somebody who uses the rallying flag but doesn’t have the pre-existing differences that create tribal membership and so never really fits into the tribe.

9. Nationalism, Patriotism, and Racism: Nationalism and patriotism use national identity as the rallying flag for a strong tribe. In many cases, nationalism becomes ethno-nationalism, which builds tribal identity off of a combination of heritage, language, religion, and culture. It has to be admitted that this can make for some incredibly strong tribes. The rallying flag is built into ancestry, and so the walls are near impossible to obliterate. The symbolism and jargon and cultural identity can be instilled from birth onward. Probably the best example of this is the Jews, who combine ethnicity, religion, and language into a bundle deal and have resisted assimilation for millennia.

Sometimes this can devolve into racism. I’m not sure exactly what the difference between ethno-nationalism and racism is, or whether there even is a difference, except that “race” is a much more complicated concept than ethnicity and it’s probably not a coincidence that it has become most popular in a country like America whose ethnicities are hopelessly confused. The Nazis certainly needed a lot of work to transform concern about the German nation into concern about the Aryan race. But it’s fair to say all of this is somewhat related or at least potentially related.

On the other hand, in countries that have non-ethnic notions of heritage, patriotism has an opportunity to substite for racism. Think about the power of the civil rights message that, whether black or white, we are all Americans.

This is maybe most obvious in sub-national groups. Despite people paying a lot of attention to the supposed racism of Republicans, the rare black Republicans do shockingly well within their party. Both Ben Carson and Herman Cain briefly topped the Republican presidential primary polls during their respective election seasons, and their failures seem to have had much more to do with their own personal qualities than with some sort of generic Republican racism. I see the same with Thomas Sowell, with Hispanic Republicans like Ted Cruz, and Asian Republicans like Bobby Jindal.

Maybe an even stronger example is the human biodiversity movement, which many people understandably accuse of being entirely about racism. Nevertheless, some of its most leading figures are black – JayMan and Chanda Chisala (who is adjacent to the movement but gets lots of respect within it) – and they seem to get equal treatment and respect to their white counterparts. Their membership in a strong and close-knit tribe screens off everything else about them.

I worry that attempts to undermine nationalism/patriotism in order to fight racism risk backfiring. The weaker the “American” tribe becomes, the more people emphasize their other tribes – which can be either overtly racial or else heavily divided along racial lines (eg political parties). It continues to worry me that people who would never display an American flag on their lawn because “nations are just a club for hating foreigners” now have a campaign sign on their lawn, five bumper stickers on their car, and are identifying more and more strongly with political positions – ie clubs for hating their fellow citizens.

Is there such a thing as conservation of tribalism? Get rid of one tribal identity and people just end up seizing on another? I’m not sure. And anyway, nobody can agree on exactly what the American identity or American tribe is anyway, so any conceivable such identity would probably risk alienating a bunch of people. I guess that makes it a moot point. But I still think that deliberately trying to eradicate patriotism is not as good an idea as is generally believed.

V.

I think tribes are interesting and underdiscussed. And in a lot of cases when they are discussed, it’s within preexisting frameworks that tilt the playing field towards recognizing some tribes as fundamentally good, others as fundamentally bad, and ignoring the commonalities between all of them.

But in order to talk about tribes coherently, we need to talk about rallying flags. And that involves admitting that a lot of rallying flags are based on ideologies (which are sometimes wrong), holy books (which are always wrong), nationality (which we can’t define), race (which is racist), and works of art (which some people inconveniently want to enjoy just as normal art without any connotations).

My title for this post is also my preferred summary: the ideology is not the movement. Or, more jargonishly – the rallying flag is not the tribe. People are just trying to find a tribe for themselves and keep it intact. This often involves defending an ideology they might not be tempted to defend for any other reason. This doesn’t make them bad, and it may not even necessarily mean their tribe deserves to go extinct. I’m reluctant to say for sure whether I think it’s okay to maintain a tribe based on a faulty ideology, but I think it’s at least important to understand that these people are in a crappy situation with no good choices, and they deserve some pity.

Some vital aspects of modern society – freedom of speech, freedom of criticism, access to multiple viewpoints, the existence of entryist tribes with explicit goals of invading and destroying competing tribes as problematic, and the overwhelming pressure to dissolve into the Generic Identity Of Modern Secular Consumerism – make maintaining tribal identities really hard these days. I think some of the most interesting sociological questions revolve around whether there are any ways around the practical and moral difficulties with tribalism, what social phenomena are explicable as the struggle of tribes to maintain themselves in the face of pressure, and whether tribalism continues to be a worthwhile or even a possible project at all.

EDIT: Commenters point out a very similar Melting Asphalt post, Religion Is Not About Beliefs.

# A Theory About Religion

Related to Monday’s post but spun off for length reasons: my crazy theory about where religion comes from.

The near-universal existence of religion across cultures is surprising. Many people have speculated on what makes tribes around the world so fixated on believing in gods and propitiating them and so on. More recently people like Dawkins and Dennett have added their own contributions about parasitic memes and hyperactive agent-detection.

But I think a lot of these explanations are too focused on a modern idea of religion. I find ancient religion much more enlightening. I’m no historian, but from the little I know ancient religion seems to bleed seamlessly into every other aspect of the ancient way of life. For example, the Roman religion was a combination of mythology, larger-than-life history, patriotism, holidays, customs, superstitions, rules about the government, beliefs about virtue, and attempts to read the future off the livers of pigs. And aside from the pig livers, this seems entirely typical.

American culture (“American civil religion“) has a lot of these features too. It has mythology and larger-than-life history: George Washington chopping down the cherry tree, the wise and glorious Founding Fathers, Honest Abe single-handedly freeing the slaves with his trusty hatchet. It has patriotic symbols and art: the flag, the anthem, Uncle Sam. It has holidays: the Fourth of July, Martin Luther King Day, Washington’s birthday. It has customs: eat turkey on Thanksgiving, have a barbecue on Memorial Day, watch the Super Bowl. It has superstitions – the number 13, black cats – and ritual taboos – even “obvious” things like don’t go outside naked needs to be thought of as taboo considering some cultures do so without thinking. It has rules about the government – both the official laws you’ll find in the federal law code, but also deep-seated beliefs about the goodness of democracy or about how all men are created equal, and even customs that affect day-to-day governance like the President giving a State of the Union in January before both houses of Congress. There are beliefs about virtue: everyone should be free, we should try to be independent, we should work hard and pursue the American Dream.

People call the Jewish dietary code unusually strict, but it’s important to realize the strictness of modern American kashrut. Absolutely no eating insects – remember, even Jewish kashrut allows locusts! Precious few birds outside of chickens, ducks, turkeys, and geese – remember, even Jewish kashrut allows pigeons! No dogs, cats, rodents, or horses. No reptiles or amphibians, no matter how much the French try to convince us that frog legs are great. No eating clearly obvious animal heads with eyes and stuff (even though dozens of advanced cultures do so happily). No blood products (eg black pudding). Mixing milk and soda in the same glass would be absurd and disgusting. Any tuna made with a process that cannot 100% exclude dolphins is impure. And this isn’t even including all of the more modern health-oriented taboos like gluten, MSG, trans-fats, GM foods, et cetera.

If we were to ask the same New Guinea tribe to follow Jewish food taboos one week and American food taboos the next, I’m not sure they’d be able to identify one code as any stricter or weirder than the other. They might have some questions about the meat/milk thing, but maybe they’d also wonder why cheeseburgers are great for dinner but ridiculous for breakfast.

People get worked up over all of the weird purity laws and dress codes in Leviticus, but it’s important to realize how strict our own purity laws are. The ancient Jews would have found it ridiculous that men have to shave and bathe every day if they want to be considered for the best jobs. One must not piss anywhere other than a toilet; this is an abomination (but you would be shocked how many of the supposedly strait-laced Japanese will go in an alley if there’s no restroom nearby). I have been yelled at for going to work without a tie and for tying my tie in the wrong pattern; wearing sweatpants to work is right out. And once again, this gets even longer if you you let the more modern/rational rules onto the list – Leviticus has a lot to say about dwellings with fungus in them, but I recently learned to my distress that landlord/tenant law has a lot more.

Once again, if we made our poor New Guinea tribe follow Jewish purity laws one week and American purity laws the next, they would probably end up equally confused and angry both times.

So when we think of America as a perfectly natural secular culture, and Jews as following some kind of superstitious draconian law code, we’re just saying that our laws feel natural and obvious, but their laws feel like an outside imposition. And I think if a time-traveling King Solomon showed up at our doorstep, he would recognize American civil religion as a religion much quicker than he would recognize Christianity as one. Christianity would look like a barbaric mystery cult that had gotten too big for its britches; American civil religion would look like home.

Insofar as this isn’t obvious to schoolchildren learning about ancient religion, it’s because the only thing one ever hears about ancient religion is the crazy mythologies. But I think American culture shows lots of signs of trying to form a crazy mythology, only to be stymied by modernity-specific factors. We can’t have crazy mythologies because we have too many historians around to tell us exactly how things really happened. We can’t have crazy mythologies because we have too many scientists around to tell us where the rain and the lightning really come from. We can’t have crazy mythologies because we’re only two hundred-odd years old and these things take time. And most of all, we can’t have crazy mythologies because Christianity is already sitting around occupying that spot.

But if America was a thousand years old and had no science, no religion, and no writing, we would have crazy mythologies up the wazoo. George Washington would take on the stature of an Agamemnon; Benjamin Franklin would take on the status of a Daedalus. Instead of centaurs and satyrs and lamia we would have jackalopes and chupacabras and grey aliens. All those people who say with a nod and a wink that Paul Bunyan dug the Great Lakes as a drinking trough for his giant ox would say the same thing nodless and winkless. Superman would take on the stature of a Zeus, dwelling beside Obi-Wan Kenobi and Bigfoot atop Mt. Whitney, helping the virtuous and punishing the wicked. Some American Hesiod would put succumb to the systematizing impulse, put it all together and explain how George Washington was the son of Superman and ordered Paul Bunyan to dig Chesapeake Bay to entrap the British fleet, and nobody would be able to say they were wrong. I mean, we already have Superman vs. Batman as canon, why not go the extra distance?

So in one sense our best analogy for ancient religion is American civil religion coupled to the sort of national mythology we might have gotten if we’d been a little bit more historically confused. But in another sense ancient religion was actually much stronger than this. America has its own individual culture, but it also partakes of the entire Western liberal industrial secular worldview. An American might not feel culture shock if she moved to Britain; she probably would if she moved to New Guinea.

The ancient world had far less trade and transportation than our own and was far less homogenized. If you want to get into the shoes of an ancient contemplating his religion, imagine you’re an American in a world where even your closest neighboring countries are as different from you as New Guinea hill tribes, or Afghan chiefdoms, or Chinese party cadres. In a world like this, your identity as an American would be very salient – and the essence of being an American is impossible to separate from this whole set of national beliefs about celebrating the Fourth of July, not eating insects, wanting freedom and democracy, and believing that Superman lives atop Mt. Whitney. Outside the community of people who 100% believe all these things, there’s just unreachable foreigners whose language you do not speak and whose customs seem somewhere between inscrutable and barbaric.

That was ancient religion – culture in a world where culture meant something. It was nothing like modern religion – which is why you never hear the Greeks complaining that the Egyptians were evil heretics who denied the light of Zeus and needed to be converted by the sword. But ancients nevertheless felt a connection to their culture and community that combined modern patriotism, religious piety, and belief in science – and they expressed it by continuing to perform their rites and even dying for them.

The question of the origin of religion comes down to how these cultures evolved into the clearly-defined religions of the modern day.

I think a big part of this is ossification and separation from context. The Jewish law perfectly preserves what any right-thinking Israelite in 1000 BC would have considered obvious, natural, and not-even-needing-justification (much as any right-thinking American today considers not eating insects obvious). By the time the Bible was being written this was no longer true – foreign customs and inevitable social change were making the old law seem less and less relevant, and I think modern scholarship thinks the Bible was written by a conservative faction of priests making their case for adherence to the old ways. The act of writing it down in a book, declaring this book the sort of thing that people might doubt but shouldn’t, and then passing that book to their children – that made it a modern religion, in the sense of something potentially separable from culture that required justification. I think that emphasizing the role of God and the gods provided that justification.

The Hebrew Bible never says other gods don’t exist; indeed, it often says the opposite. It constantly praises God as stronger and better than other gods. God proves his superiority over the gods of the Egyptians when the serpent he sends Moses eats the serpents the Egyptian gods send Pharaoh’s sorcerers. The Israelites are constantly warned against worshipping other gods, not because those gods don’t exist but because God is better and also jealous. This is not the worldview of somebody who has very strong ideas about the nature of reality and how supernatural beings fit into that nature. It’s the worldview of people who want to say “Our culture is better than your culture”. The Bible uses “worshipping foreign gods” as synonymous with “turning to foreign ways”. But God has a covenant with Israel, therefore both are forbidden.

This seems to match religion in the classical world – I’m especially thinking of Augustus’ conception here, but he wasn’t drawing it out of a vacuum. Performing the proper rites to the Roman gods was how you showed you were on board with Roman culture was how you showed you were loyal to Rome. The Roman view of religion seems pretty ridiculous to us – constant influx of new gods and mystery cults that were believed kind of indiscriminately, plus occasional deification of leading political figures followed by their undeification once they fell from power. But throughout it all, this idea that following the rites as Romulus prescribed them showed loyalty, but doing otherwise would result in decadence and defeat, stuck around.

More modern religions like Christianity, Islam, and Buddhism are a bit different. Obviously their respective founders play a huge role, but I still think part of what makes them religions rather than just philosophies or spiritual teachings is that they underwent this ossification process. Just as modern Judaism preserves many features of 10th-century-BC Israel that got encoded into holy writ, so modern Christianity preserves many features of 1st century Judeo-Hellenist syncretism. In fact, it preserves a lot of features of 13th-century scholasticism, since that was when they really became serious about formalizing and officializing their theology. At the time scholasticism wasn’t particularly religious; it was just the best understanding the 13th century academic community had about the world around them. Since the Church officialized it, everyone else drifted away and they didn’t.

I think it’s also possible that the first few followers of these religions ended up as a subculture, with as much arbitrary subcultural development as any other tribe. My personal experience with subcultures tells me they can get very different customs from the surrounding society very fast, with or without any connection to real feature of their rallying flag. Those unusual subcultural values then became the values of the religion that developed later.

Hopefully the connection to Monday’s post is pretty clear. The important thing about a religion is that it has a rallying flag that encourages it to preserve a certain culture, plus walls against the outside world. Crucially, despite everything I’m saying about ossification the culture changes a lot: King Solomon would probably recognize modern rabbinic Judaism, but only barely. But it changes in a way different from the way the outside secular society changes, and in ways bound by the ossified text, so there’s still an element of having this ancient culture preserved in amber and maintained up to the modern day.

# SSC Bay Area Meetups 2016

The following Slate Star Codex meetups are planned for the next week or so:

Berkeley, Sunday 4/17 at 2:00 PM at CFAR office, 2030 Addison, 7th floor, Berkeley

San Jose, Monday 4/18 at 7:00 PM at a private house, 3806 Williams Rd, San Jose

Googleplex, Tuesday 4/19 at 12:30 PM at Bigtable cafe in building 1900. Googlers can get more information at the internal slatestarcodex-discuss mailing list.

Stanford, Tuesday 4/19 at 5:30 PM at Tressider Food Court

Everyone is welcome to Berkeley, San Jose, and Stanford. Googleplex is limited to Google employees. Please feel free to come even if you don’t read this much, don’t comment, don’t understand everything, don’t agree with me about things, or don’t feel like you fit in with the normal demographic. During past meetups we’ve always appreciated having more diversity of different types of people attending.

# Book Review: Albion’s Seed

I.

Albion’s Seed by David Fischer is a history professor’s nine-hundred-page treatise on patterns of early immigration to the Eastern United States. It’s not light reading and not the sort of thing I would normally pick up. I read it anyway on the advice of people who kept telling me it explains everything about America. And it sort of does.

In school, we tend to think of the original American colonists as “Englishmen”, a maximally non-diverse group who form the background for all of the diversity and ethnic conflict to come later. Fischer’s thesis is the opposite. Different parts of the country were settled by very different groups of Englishmen with different regional backgrounds, religions, social classes, and philosophies. The colonization process essentially extracted a single stratum of English society, isolated it from all the others, and then plunked it down on its own somewhere in the Eastern US.

I used to play Alpha Centauri, a computer game about the colonization of its namesake star system. One of the dynamics that made it so interesting was its backstory, where a Puerto Rican survivalist, an African plutocrat, and other colorful characters organized their own colonial expeditions and competed to seize territory and resources. You got to explore not only the settlement of a new world, but the settlement of a new world by societies dominated by extreme founder effects. What kind of weird pathologies and wonderful innovations do you get when a group of overly romantic Scottish environmentalists is allowed to develop on its own trajectory free of all non-overly-romantic-Scottish-environmentalist influences? Albion’s Seed argues that this is basically the process that formed several early US states.

Fischer describes four of these migrations: the Puritans to New England in the 1620s, the Cavaliers to Virginia in the 1640s, the Quakers to Pennsylvania in the 1670s, and the Borderers to Appalachia in the 1700s.

II.

A: The Puritans

I hear about these people every Thanksgiving, then never think about them again for the next 364 days. They were a Calvinist sect that dissented against the Church of England and followed their own brand of dour, industrious, fun-hating Christianity. Most of them were from East Anglia, the part of England just northeast of London. They came to America partly because they felt persecuted, but mostly because they thought England was full of sin and they were at risk of absorbing the sin by osmosis if they didn’t get away quick and build something better. They really liked “city on a hill” metaphors.

I knew about the Mayflower, I knew about the black hats and silly shoes, I even knew about the time Squanto threatened to release a bioweapon buried under Plymouth Rock that would bring about the apocalypse. But I didn’t know that the Puritan migration to America was basically a eugenicist’s wet dream.

Much like eg Unitarians today, the Puritans were a religious group that drew disproportionately from the most educated and education-obsessed parts of the English populace. Literacy among immigrants to Massachusetts was twice as high as the English average, and in an age when the vast majority of Europeans were farmers most immigrants to Massachusetts were skilled craftsmen or scholars. And the Puritan “homeland” of East Anglia was a an unusually intellectual place, with strong influences from Dutch and Continental trade; historian Havelock Ellis finds that it “accounts for a much larger proportion of literary, scientific, and intellectual achievement than any other part of England.”

Furthermore, only the best Puritans were allowed to go to Massachusetts; Fischer writes that “it may have been the only English colony that required some of its immigrants to submit letters of recommendation” and that “those who did not fit in were banished to other colonies and sent back to England”. Puritan “headhunters” went back to England to recruit “godly men” and “honest men” who “must not be of the poorer sort”.

INTERESTING PURITAN FACTS:  
1. Sir Harry Vane, who was “briefly governor of Massachusetts at the age of 24”, “was so rigorous in his Puritanism that he believed only the thrice-born to be truly saved”.  
2. The great seal of the Massachusetts Bay Company “featured an Indian with arms beckoning, and five English words flowing from his mouth: ‘Come over and help us'”  
3. Northern New Jersey was settled by Puritans who named their town after the “New Ark Of The Covenant” – modern Newark.  
4. Massachusetts clergy were very powerful; Fischer records the story of a traveller asking a man “Are you the parson who serves here?” only to be corrected “I am, sir, the parson who rules here.”  
5. The Puritans tried to import African slaves, but they all died of the cold.  
6. In 1639, Massachusetts declared a “Day Of Humiliation” to condemn “novelties, oppression, atheism, excesse, superfluity, idleness, contempt of authority, and trouble in other parts to be remembered”  
7. The average family size in Waltham, Massachusetts in the 1730s was 9.7 children.  
8. Everyone was compelled by law to live in families. Town officials would search the town for single people and, if found, order them to join a family; if they refused, they were sent to jail.  
9. 98% of adult Puritan men were married, compared to only 73% of adult Englishmen in general. Women were under special pressure to marry, and a Puritan proverb said that “women dying maids lead apes in Hell”.  
10. 90% of Puritan names were taken from the Bible. Some Puritans took pride in their learning by giving their children obscure Biblical names they would expect nobody else to have heard of, like Mahershalalhasbaz. Others chose random Biblical terms that might not have technically been intended as names; “the son of Bostonian Samuel Pond was named Mene Mene Tekel Upharsin Pond”. Still others chose Biblical words completely at random and named their children things like Maybe or Notwithstanding.  
11. Puritan parents traditionally would send children away to be raised with other families, and raise those families’ children in turn, in the hopes that the lack of familiarity would make the child behave better.  
12. In 1692, 25% of women over age 45 in Essex County were accused of witchcraft.  
13. Massachusetts passed the first law mandating universal public education, which was called The Old Deluder Satan Law in honor of its preamble, which began “It being one chief project of that old deluder, Satan, to keep men from the knowledge of the scriptures…”  
14. Massachusetts cuisine was based around “meat and vegetables submerged in plain water and boiled relentlessly without seasonings of any kind”.  
15. Along with the famous scarlet A for adultery, Puritans could be forced to wear a B for blasphemy, C for counterfeiting, D for drunkenness, and so on.  
16. Wasting time in Massachusetts was literally a criminal offense, listed in the law code, and several people were in fact prosecuted for it.  
17. This wasn’t even the nadir of weird hard-to-enforce Massachusetts laws. Another law just said “If any man shall exceed the bounds of moderation, we shall punish him severely”.

Harriet Beecher Stowe wrote of Massachusetts Puritanism: “The underlying foundation of life in New England was one of profound, unutterable, and therefore unuttered mehalncholy, which regarded human existence itself as a ghastly risk, and, in the case of the vast majority of human beings, an inconceivable misfortune.” And indeed, everything was dour, strict, oppressive, and very religious. A typical Massachusetts week would begin in the church, which doubled as the town meeting hall. There were no decorations except a giant staring eye on the pulpit to remind churchgoers that God was watching them. Townspeople would stand up before their and declare their shame and misdeeds, sometimes being forced to literally crawl before the other worshippers begging for forgiveness. Then the minister would give two two-hour sermons back to back. The entire affair would take up to six hours, and the church was unheated (for some reason they stored all their gunpowder there, so no one was allowed to light a fire), and this was Massachusetts, and it was colder in those days than it is now, so that during winter some people would literally lose fingers to frostbite (Fischer: “It was a point of honor for the minister never to shorten a sermon merely because his audience was frozen”). Everyone would stand there with their guns (they were legally required to bring guns, in case Indians attacked during the sermon) and hear about how they were going to Hell, all while the giant staring eye looked at them.

So life as a Puritan was pretty terrible. On the other hand, their society was impressively well-ordered. Teenage pregnancy rates were the lowest in the Western world and in some areas literally zero. Murder rates were half those in other American colonies. There was remarkably low income inequality – “the top 10% of wealthholders held only 20%-30% of taxable property”, compared to 75% today and similar numbers in other 17th-century civilizations. The poor (at least the poor native to a given town) were treated with charity and respect – “in Salem, one man was ordered to be set by the heels in the stocks for being uncharitable to a poor man in distress”. Government was conducted through town meetings in which everyone had a say. Women had more equality than in most parts of the world, and domestic abuse was punished brutally. The educational system was top-notch – “by most empirical tests of intellectual eminence, New England led all other parts of British America from the 17th to the early 20th century”.

In some ways the Puritans seem to have taken the classic dystopian bargain – give up all freedom and individuality and art, and you can have a perfect society without crime or violence or inequality. Fischer ends each of his chapters with a discussion of how the society thought of liberty, and the Puritans unsurprisingly thought of liberty as “ordered liberty” – the freedom of everything to tend to its correct place and stay there. They thought of it as a freedom from disruption – apparently FDR stole some of his “freedom from fear” stuff from early Puritan documents. They were extremely not in favor of the sort of liberty that meant that, for example, there wouldn’t be laws against wasting time. That was going too far.

B: The Cavaliers

The Massachusetts Puritans fled England in the 1620s partly because the king and nobles were oppressing them. In the 1640s, English Puritans under Oliver Cromwell rebelled, took over the government, and killed the king. The nobles not unreasonably started looking to get the heck out.

Virginia had been kind of a wreck ever since the original Jamestown settlers had mostly died of disease. Governor William Berkeley, a noble himself, decided the colony could reinvent itself as a destination for refugee nobles, and told them it would do everything possible to help them maintain the position of oppressive supremacy to which they were accustomed. The British nobility was sold. The Cavaliers – the nobles who had fought and lost the English Civil War – fled to Virginia. Historians cross-checking Virginian immigrant lists against English records find that of Virginians whose opinions on the War were known, 98% were royalists. They were overwhelmingly Anglican, mostly from agrarian southern England, and all related to each other in the incestuous way of nobility everywhere: “it is difficult to think of any ruling elite that has been more closely interrelated since the Ptolemies”. There were twelve members of Virginia’s royal council; in 1724 “all without exception were related to one another by blood or marriage…as late as 1775, every member of that august body was descended from a councilor who had served in 1660”.

These aristocrats didn’t want to do their own work, so they brought with them tens of thousands of indentured servants; more than 75% of all Virginian immigrants arrived in this position. Some of these people came willingly on a system where their master paid their passage over and they would be free after a certain number of years; others were sent by the courts as punishments; still others were just plain kidnapped. The gender ratio was 4:1 in favor of men, and there were entire English gangs dedicated to kidnapping women and sending them to Virginia, where they fetched a high price. Needless to say, these people came from a very different stratum than their masters or the Puritans.

People who came to Virginia mostly died. They died of malaria, typhoid fever, amoebiasis, and dysentery. Unlike in New England, where Europeans were better adapted to the cold climate than Africans, in Virginia it was Europeans who had the higher disease-related mortality rate. The whites who survived tended to become “sluggish and indolent”, according to the universal report of travellers and chroniclers, although I might be sluggish and indolent too if I had been kidnapped to go work on some rich person’s farm and sluggishness/indolence was an option.

The Virginians tried their best to oppress white people. Really, they did. The depths to which they sank in trying to oppress white people almost boggle the imagination. There was a rule that if a female indentured servant became pregnant, a few extra years were added on to their indenture, supposedly because they would be working less hard during their pregnancy and child-rearing so it wasn’t fair to the master. Virginian aristocrats would rape their own female servants, then add a penalty term on to their indenture for becoming pregnant. That is an impressive level of chutzpah. But despite these efforts, eventually all the white people either died, or became too sluggish to be useful, or worst of all just finished up their indentures and became legally free. The aristocrats started importing black slaves as per the model that had sprung up in the Caribbean, and so the stage was set for the antebellum South we read about in history classes.

INTERESTING CAVALIER FACTS:  
1. Virginian cavalier speech patterns sound a lot like modern African-American dialects. It doesn’t take much imagination to figure out why, but it’s strange to think of a 17th century British lord speaking what a modern ear would clearly recognize as Ebonics.  
2. Three-quarters of 17th-century Virginian children lost at least one parent before turning 18.  
3. Cousin marriage was an important custom that helped cement bonds among the Virginian elite, “and many an Anglican lady changed her condition but not her name”.  
4. In Virginia, women were sometimes unironically called “breeders”; English women were sometimes referred to as “She-Britons”.  
5. Virginia didn’t really have towns; the Chesapeake Bay was such a giant maze of rivers and estuaries and waterways that there wasn’t much need for land transport hubs. Instead, the unit of settlement was the plantation, which consisted of an aristocratic planter, his wife and family, his servants, his slaves, and a bunch of guests who hung around and mooched off him in accordance with the ancient custom of hospitality.  
6. Virginian society considered everyone who lived in a plantation home to be a kind of “family”, with the aristocrat both as the literal father and as a sort of abstracted patriarch with complete control over his domain.  
7. Virginia governor William Berkeley probably would not be described by moderns as ‘strong on education’. He said in a speech that “I thank God there are no free schools nor printing [in Virginia], and I hope we shall not have these for a hundred years, for learning has brought disobedience, and heresy, and sects into the world, and printing has divuldged them, and libels against the best government. God keep us from both!”  
8. Virginian recreation mostly revolved around hunting and bloodsports. Great lords hunted deer, lesser gentry hunted foxes, indentured servants had a weird game in which they essentially draw-and-quartered geese, young children “killed and tortured songbirds”, and “at the bottom of this hierarchy of bloody games were male infants who prepared themselves for the larger pleasures of maturity by torturing snakes, maiming frogs, and pulling the wings off butterflies. Thus, every red-blooded male in Virginia was permitted to slaughter some animal or other, and the size of his victim was proportioned to his social rank.”  
9. “In 1747, an Anglican minister named William Kay infuriated the great planter Landon Carter by preaching a sermon against pride. The planter took it personally and sent his [relations] and ordered them to nail up the doors and windows of all the churches in which Kay preached.”  
10. Our word “condescension” comes from a ritual attitude that leading Virginians were supposed to display to their inferiors. Originally condescension was supposed to be a polite way of showing respect those who were socially inferior to you; our modern use of the term probably says a lot about what Virginians actually did with it.

In a lot of ways, Virginia was the opposite of Massachusetts. Their homicide rate was sky-high, and people were actively encouraged to respond to slights against their honor with duels (for the rich) and violence (for the poor). They were obsessed with gambling, and “made bets not merely on horses, cards, cockfights, and backgammon, but also on crops, prices, women, and the weather”. Their cuisine focused on gigantic sumptuous feasts of animals killed in horrible ways. There were no witchcraft trials, but there were people who were fined for disrupting the peace by accusing their neighbors of witchcraft. Their church sermons were twenty minutes long on the dot.

The Puritans naturally thought of the Virginians as completely lawless reprobate sinners, but this is not entirely true. Virginian church sermons might have been twenty minutes long, but Virginian ballroom dance lessons could last nine hours. It wasn’t that the Virginians weren’t bound by codes, just that those codes were social rather than moral.

And Virginian nobles weren’t just random jerks, they were carefully cultivated jerks. Planters spared no expense to train their sons to be strong, forceful, and not take nothin’ from nobody. They would encourage and reward children for being loud and temperamental, on the grounds that this indicated a strong personality and having a strong personality was fitting of a noble. When this worked, it worked really well – witness natural leaders and self-driven polymaths like George Washington and Thomas Jefferson. More often it failed catastrophically – the rate of sex predation and rape in Virginia was at least as high as anywhere else in North America.

The Virginian Cavaliers had an obsession with liberty, but needless to say it was not exactly a sort of liberty of which the ACLU would approve. I once heard someone argue against libertarians like so: even if the government did not infringe on liberties, we would still be unfree for other reasons. If we had to work, we would be subject to the whim of bosses. If we were poor, we would not be “free” to purchase most of the things we want. In any case, we are “oppressed” by disease, famine, and many other things besides government that prevent us from implementing our ideal existence.

The Virginians took this idea and ran with it – in the wrong direction. No, they said, we wouldn’t be free if we had to work, therefore we insist upon not working. No, we wouldn’t be free if we were limited by poverty, therefore we insist upon being extremely rich. Needless to say, this conception of freedom required first indentured servitude and later slavery to make it work, but the Virginians never claimed that the servants or slaves were free. That wasn’t the point. Freedom, like wealth, was properly distributed according to rank; nobles had as much as they wanted, the middle-class enough to get by on, and everyone else none at all. And a Virginian noble would have gone to his grave insisting that a civilization without slavery could never have citizens who were truly free.

C: The Quakers

Fischer warns against the temptation to think of the Quakers as normal modern people, but he has to warn us precisely because it’s so tempting. Where the Puritans seem like a dystopian caricature of virtue and the Cavaliers like a dystopian caricature of vice, the Quakers just seem ordinary. Yes, they’re kind of a religious cult, but they’re the kind of religious cult any of us might found if we were thrown back to the seventeenth century.

Instead they were founded by a weaver’s son named George Fox. He believed people were basically good and had an Inner Light that connected them directly to God without a need for priesthood, ritual, Bible study, or self-denial; mostly people just needed to listen to their consciences and be nice. Since everyone was equal before God, there was no point in holding up distinctions between lords and commoners: Quakers would just address everybody as “Friend”. And since the Quakers were among the most persecuted sects at the time, they developed an insistence on tolerance and freedom of religion which (unlike the Puritans) they stuck to even when shifting fortunes put them on top. They believed in pacificism, equality of the sexes, racial harmony, and a bunch of other things which seem pretty hippy-ish even today let alone in 1650.

England’s top Quaker in the late 1600s was William Penn. Penn is universally known to Americans as “that guy Pennsylvania is named after” but actually was a larger-than-life 17th century superman. Born to the nobility, Penn distinguished himself early on as a military officer; he was known for beating legendary duelists in single combat and then sparing their lives with sermons about how murder was wrong. He gradually started having mystical visions, quit the military, and converted to Quakerism. Like many Quakers he was arrested for blasphemy; unlike many Quakers, they couldn’t make the conviction stick; in his trial he “conducted his defense so brilliantly that the jurors refused to convict him even when threatened with prison themselves, [and] the case became a landmark in the history of trial by jury.” When the state finally found a pretext on which to throw him in prison, he spent his incarceration composing “one of the noblest defenses of religious liberty ever written”, conducting a successful mail-based courtship with England’s most eligible noblewoman, and somehow gaining the personal friendship and admiration of King Charles II. Upon his release the King liked him so much that he gave him a large chunk of the Eastern United States on a flimsy pretext of repaying a family debt. Penn didn’t want to name his new territory Pennsylvania – he recommended just “Sylvania” – but everybody else overruled him and Pennyslvania it was. The grant wasn’t quite the same as the modern state, but a chunk of land around the Delaware River Valley – what today we would call eastern Pennsylvania, northern Delaware, southern New Jersey, and bits of Maryland – centered on the obviously-named-by-Quakers city of Philadelphia.

Penn decided his new territory would be a Quaker refuge – his exact wording was “a colony of Heaven [for] the children of the Light”. He mandated universal religious toleration, a total ban on military activity, and a government based on checks and balances that would “leave myself and successors no power of doing mischief, that the will of one man may not hinder the good of a whole country”.

His recruits – about 20,000 people in total – were Quakers from the north of England, many of them minor merchants and traders. They disproportionately included the Britons of Norse descent common in that region, who formed a separate stratum and had never really gotten along with the rest of the British population. They were joined by several German sects close enough to Quakers that they felt at home there; these became the ancestors of (among other groups) the Pennsylvania Dutch, Amish, and Mennonites.

INTERESTING QUAKER FACTS:  
1. In 1690 a gang of pirates stole a ship in Philadelphia and went up and down the Delaware River stealing and plundering. The Quakers got in a heated (but brotherly) debate about whether it was morally permissible to use violence to stop them. When the government finally decided to take action, contrarian minister George Keith dissented and caused a major schism in the faith.  
2. Fischer argues that the Quaker ban on military activity within their territory would have doomed them in most other American regions, but by extreme good luck the Indians in the Delaware Valley were almost as peaceful as the Quakers. As usual, at least some credit goes to William Penn, who taught himself Algonquin so he could negotiate with the Indians in their own language.  
3. The Quakers’ marriage customs combined a surprisingly modern ideas of romance, with extreme bureaucracy. The wedding process itself had sixteen stages, including “ask parents”, “ask community women”, “ask community men”, “community women ask parents”, and “obtain a certificate of cleanliness”. William Penn’s marriage apparently had forty-six witnesses to testify to the good conduct and non-relatedness of both parties.  
4. Possibly related: 16% of Quaker women were unmarried by age 50, compared to only about 2% of Puritans.  
5. Quakers promoted gender equality, including the (at the time scandalous) custom of allowing women to preach (condemned by the Puritans as the crime of “she-preaching”).  
6. But they were such prudes about sex that even the Puritans thought they went too far. Pennsylvania doctors had problems treating Quakers because they would “delicately describe everything from neck to waist as their ‘stomachs’, and anything from waist to feet as their ‘ankles'”.  
7. Quaker parents Richard and Abigail Lippincott named their eight children, in order, “Remember”, “John”, “Restore”, “Freedom”, “Increase”, “Jacob”, “Preserve”, and “Israel”, so that their names combined formed a simple prayer.  
8. Quakers had surprisingly modern ideas about parenting, basically sheltering and spoiling their children at a time when everyone else was trying whip the Devil out of them.  
9. “A Quaker preacher, traveling in the more complaisant colony of Maryland, came upon a party of young people who were dancing merrily together. He broke in upon them like an avenging angel, stopped the dance, anddemanded to know if they considered Martin Luther to be a good man. The astonished youngsters answered in the affirmative. The Quaker evangelist then quoted Luther on the subject of dancing: ‘as many paces as the man takes in his dance, so many steps he takes toward Hell. This, the Quaker missionary gloated with a gleam of sadistic satisfaction, ‘spoiled their sport’.”  
10. William Penn wrote about thirty books defending liberty of conscience throughout his life. The Quaker obsession with the individual conscience as the work of God helped invent the modern idea of conscientious objection.  
11. Quakers were heavily (and uniquely for their period) opposed to animal cruelty. When foreigners introduced bullbaiting into Philadelphia during the 1700s, the mayor bought a ticket supposedly as a spectator. When the event was about to begin, he leapt into the ring, personally set the bull free, and threatened to arrest anybody who stopped him.  
12. On the other hand, they were also opposed to other sports for what seem like kind of random reasons. The town of Morley declared an anathema against foot races, saying that they were “unfruitful works of darkness”.  
13. The Pennsylvania Quakers became very prosperous merchants and traders. They also had a policy of loaning money at low- or zero- interest to other Quakers, which let them outcompete other, less religious businesspeople.  
14. They were among the first to replace the set of bows, grovels, nods, meaningful looks, and other British customs of acknowledging rank upon greeting with a single rank-neutral equivalent – the handshake.  
15. Pennsylvania was one of the first polities in the western world to abolish the death penalty.  
16. The Quakers were lukewarm on education, believing that too much schooling obscured the natural Inner Light. Fischer declares it “typical of William Penn” that he wrote a book arguing against reading too much.  
17. The Quakers not only instituted religious freedom, but made laws against mocking another person’s religion.  
18. In the late 1600s as many as 70% of upper-class Quakers owned slaves, but Pennsylvania essentially invented modern abolitionism. Although their colonial masters in England forbade them from banning slavery outright, they applied immense social pressure and by the mid 1700s less than 10% of the wealthy had African slaves. As soon as the American Revolution started, forbidding slavery was one of independent Pennsylvania’s first actions.

Pennsylvania was very successful for a while; it had some of the richest farmland in the colonies, and the Quakers were exceptional merchants and traders; so much so that they were forgiven their military non-intervention during the Revolution because of their role keeping the American economy afloat in the face of British sanctions.

But by 1750, the Quakers were kind of on their way out; by 1750, they were a demographic minority in Pennsylvania, and by 1773 they were a minority in its legislature as well. In 1750 Quakerism was the third-largest religion in the US; by 1820 it was the ninth-largest, and by 1981 it was the sixty-sixth largest. What happened? The Quakers basically tolerated themselves out of existence. They were so welcoming to religious minorities and immigrants that all these groups took up shop in Pennsylvania and ended its status as a uniquely Quaker society. At the same time, the Quakers themselves became more “fanatical” and many dropped out of politics believing it to be too worldly a concern for them; this was obviously fatal to their political domination. The most famous Pennsylvanian statesman of the Revolutionary era, Benjamin Franklin, was not a Quaker at all but a first-generation immigrant from New England. Finally, Quakerism was naturally extra-susceptible to that thing where Christian denominations become indistinguishable from liberal modernity and fade into the secular background.

But Fischer argues that Quakerism continued to shape Pennsylvania long after it had stopped being officially in charge, in much the same way that Englishmen themselves have contributed disproportionately to American institutions even though they are now a numerical minority. The Pennsylvanian leadership on abolitionism, penal reform, the death penalty, and so on all happened after the colony was officially no longer Quaker-dominated.

And it’s hard not to see Quaker influence on the ideas of the modern US – which was after all founded in Philadelphia. In the middle of the Puritans demanding strict obedience to their dystopian hive society and the Cavaliers demanding everybody bow down to a transplanted nobility, the Pennsylvanians – who became the thought leaders of the Mid-Atlantic region including to a limited degree New York City – were pretty normal and had a good opportunity to serve as power-brokers and middlemen between the North and South. Although there are seeds of traditionally American ideas in every region, the Quakers really stand out in terms of freedom of religion, freedom of thought, checks and balances, and the idea of universal equality.

It occurs to me that William Penn might be literally the single most successful person in history. He started out as a minor noble following a religious sect that everybody despised and managed to export its principles to Pennsylvania where they flourished and multiplied. Pennsylvania then managed to export its principles to the United States, and the United States exported them to the world. I’m not sure how much of the suspiciously Quaker character of modern society is a direct result of William Penn, but he was in one heck of a right place at one heck of a right time

D: The Borderers

The Borderers are usually called “the Scots-Irish”, but Fischer dislikes the term because they are neither Scots (as we usually think of Scots) nor Irish (as we usually think of Irish). Instead, they’re a bunch of people who lived on (both sides of) the Scottish-English border in the late 1600s.

None of this makes sense without realizing that the Scottish-English border was terrible. Every couple of years the King of England would invade Scotland or vice versa; “from the year 1040 to 1745, every English monarch but three suffered a Scottish invasion, or became an invader in his turn”. These “invasions” generally involved burning down all the border towns and killing a bunch of people there. Eventually the two sides started getting pissed with each other and would also torture-murder all of the enemy’s citizens they could get their hands on, ie any who were close enough to the border to reach before the enemy could send in their armies. As if this weren’t bad enough, outlaws quickly learned they could plunder one side of the border, then escape to the other before anyone brought them to justice, so the whole area basically became one giant cesspool of robbery and murder.

In response to these pressures, the border people militarized and stayed feudal long past the point where the rest of the island had started modernizing. Life consisted of farming the lands of whichever brutal warlord had the top hand today, followed by being called to fight for him on short notice, followed by a grisly death. The border people dealt with it as best they could, and developed a culture marked by extreme levels of clannishness, xenophobia, drunkenness, stubbornness, and violence.

By the end of the 1600s, the Scottish and English royal bloodlines had intermingled and the two countries were drifting closer and closer to Union. The English kings finally got some breathing room and noticed – holy frick, everything about the border is terrible. They decided to make the region economically productive, which meant “squeeze every cent out of the poor Borderers, in the hopes of either getting lots of money from them or else forcing them to go elsewhere and become somebody else’s problem”. Sometimes absentee landlords would just evict everyone who lived in an entire region, en masse, replacing them with people they expected to be easier to control.

Many of the Borderers fled to Ulster in Ireland, which England was working on colonizing as a Protestant bulwark against the Irish Catholics, and where the Crown welcomed violent warlike people as a useful addition to their Irish-Catholic-fighting project. But Ulster had some of the same problems as the Border, and also the Ulsterites started worrying that the Borderer cure was worse than the Irish Catholic disease. So the Borderers started getting kicked out of Ulster too, one thing led to another, and eventually 250,000 of these people ended up in America.

250,000 people is a lot of Borderers. By contrast, the great Puritan emigration wave was only 20,000 or so people; even the mighty colony of Virginia only had about 50,000 original settlers. So these people showed up on the door of the American colonies, and the American colonies collectively took one look at them and said “nope”.

Except, of course, the Quakers. The Quakers talked among themselves and decided that these people were also Children Of God, and so they should demonstrate Brotherly Love by taking them in. They tried that for a couple of years, and then they questioned their life choices and also said “nope”, and they told the Borderers that Philadelphia and the Delaware Valley were actually kind of full right now but there was lots of unoccupied land in Western Pennsylvania, and the Appalachian Mountains were very pretty at this time of year, so why didn’t they head out that way as fast as it was physically possible to go?

At the time, the Appalachians were kind of the booby prize of American colonization: hard to farm, hard to travel through, and exposed to hostile Indians. The Borderers fell in love with them. They came from a pretty marginal and unproductive territory themselves, and the Appalachians were far away from everybody and full of fun Indians to fight. Soon the Appalachian strategy became the accepted response to Borderer immigration and was taken up from Pennsylvania in the north to the Carolinas in the South (a few New Englanders hit on a similar idea and sent their own Borderers to colonize the mountains of New Hampshire).

So the Borderers all went to Appalachia and established their own little rural clans there and nothing at all went wrong except for the entire rest of American history.

INTERESTING BORDERER FACTS:  
1. Colonial opinion on the Borderers differed within a very narrow range: one Pennsylvanian writer called them “the scum of two nations”, another Anglican clergyman called them “the scum of the universe”.  
2. Some Borderers tried to come to America as indentured servants, but after Virginian planters got some experience with Borderers they refused to accept any more.  
3. The Borderers were mostly Presbyterians, and their arrival en masse started a race among the established American denominations to convert them. This was mostly unsuccessful; Anglican preacher Charles Woodmason, an important source for information about the early Borderers, said that during his missionary activity the Borderers “disrupted his service, rioted while he preached, started a pack of dogs fighting outside the church, loosed his horse, stole his church key, refused him food and shelter, and gave two barrels of whiskey to his congregation before a service of communion”.  
4. Borderer town-naming policy was very different from the Biblical names of the Puritans or the Ye Olde English names of the Virginians. Early Borderer settlements include – just to stick to the creek-related ones – Lousy Creek, Naked Creek, Shitbritches Creek, Cuckold’s Creek, Bloodrun Creek, Pinchgut Creek, Whipping Creek, and Hangover Creek. There were also Whiskey Springs, Hell’s Half Acre, Scream Ridge, Scuffletown, and Grabtown. The overall aesthetic honestly sounds a bit Orcish.  
5. One of the first Borderer leaders was John Houston. On the ship over to America, the crew tried to steal some of his possessions; Houston retaliated by leading a mutiny of the passengers, stealing the ship, and sailing it to America himself. He settled in West Virginia; one of his descendants was famous Texan Sam Houston.  
6. Traditional Borderer prayer: “Lord, grant that I may always be right, for thou knowest I am hard to turn.”  
7. “The backcountry folk bragged that one interior county of North Carolina had so little ‘larnin’ that the only literate inhabitant was elected ‘county reader'”  
8. The Borderer accent contained English, Scottish, and Irish elements, and is (uncoincidentally) very similar to the typical “country western singer” accent of today.  
9. The Borderers were famous for family feuds in England, including the Johnson clan’s habit of “adorning their houses with the flayed skins of their enemies the Maxwells in a blood feud that continued for many generations”. The great family feuds of the United States, like the Hatfield-McCoy feud, are a direct descendent of this tradition.  
10. Within-clan marriage was a popular Borderer tradition both in England and Appalachia; “in the Cumbrian parish of Hawkshead, for example, both the bride and the groom bore the same last names in 25 percent of all marriages from 1568 to 1704”. This led to the modern stereotype of Appalachians as inbred and incestuous.  
11. The Borderers were extremely patriarchal and anti-women’s-rights to a degree that appalled even the people of the 1700s.  
12. “In the year 1767, [Anglican priest] Charles Woodmason calculated that 94 percent of backcountry brides whom he had married in the past year were pregnant on their wedding day”  
13. Although the Borderers started off Presbyterian, they were in constant religious churn and their territories were full of revivals, camp meetings, born-again evangelicalism, and itinerant preachers. Eventually most of them ended up as what we now call Southern Baptist.  
14. Borderer folk beliefs: “If an old woman has only one tooth, she is a witch”, “If you are awake at eleven, you will see witches”, “The howling of dogs shows the presence of witches”, “If your shoestring comes untied, witches are after you”, “If a warm current of air is felt, witches are passing”. Also, “wet a rag in your enemy’s blood, put it behind a rock in the chimney, and when it rots your enemy will die”; apparently it was not a coincidence they were thinking about witches so much.  
15. Borderer medical beliefs: “A cure for homesickness is to sew a good charge of gunpowder on the inside of ths shirt near the neck”. That’ll cure homesickness, all right.  
16. More Borderer medical beliefs: “For fever, cut a black chicken open while alive and bind it to the bottom of your foot”, “Eating the brain of a screech owl is the only dependable remedy for headache”, “For rheumatism, apply split frogs to the feet”, “To reduce a swollen leg, split a live cat and apply while still warm”, “Bite the head off the first butterfly you see and you will get a new dress”, “Open the cow’s mouth and throw a live toad-frog down her throat. This will cure her of hollow-horn”. Also, blacksmiths protected themselves from witches by occasionally throwing live puppies into their furnaces.  
17. Rates of public schooling in the backcountry settled by the Borderers were “the lowest in British North America” and sometimes involved rituals like “barring out”, where the children would physically keep the teacher out of the school until he gave in and granted the students the day off.  
18. “Appalachia’s idea of a moderate drinker was the mountain man who limited himself to a single quart [of whiskey] at a sitting, explaining that more ‘might fly to my head’. Other beverages were regarded with contempt.”  
19. A traditional backcountry sport was “rough and tumble”, a no-holds-barred form of wrestling where gouging out your opponent’s eyes was considered perfectly acceptable and in fact sound strategy. In 1772 Virginia had to pass a law against “gouging, plucking, or putting out an eye”, but this was the Cavalier-dominated legislature all the way on the east coast and nobody in the backcountry paid them any attention. Other traditional backcountry sports were sharpshooting and hunting.  
20. The American custom of shooting guns into the air to celebrate holidays is 100% Borderer in origin.  
21. The justice system of the backcountry was heavy on lynching, originally a race-neutral practice and named after western Virginian settler William Lynch.  
22. Scottish Presbyterians used to wear red cloth around their neck to symbolize their religion; other Englishmen nicknamed them “rednecks”. This may be the origin of the popular slur against Americans of Borderer descent, although many other etiologies have been proposed. “Cracker” as a slur is attested as early as 1766 by a colonist who says the term describes backcountry men who are great boasters; other proposed etymologies like slaves talking about “whip-crackers” seem to be spurious.

This is not to paint the Borderers as universally poor and dumb – like every group, they had an elite, and some of their elite went on to become some of America’s most important historical figures. Andrew Jackson became the first Borderer president, behaving exactly as you would expect the first Borderer president to behave, and he was followed by almost a dozen others. Borderers have also been overrepresented in America’s great military leaders, from Ulysses Grant through Teddy Roosevelt (3/4 Borderer despite his Dutch surname) to George Patton to John McCain.

The Borderers really liked America – unsurprising given where they came from – and started identifying as American earlier and more fiercely than any of the other settlers who had come before. Unsurprisingly, they strongly supported the Revolution – Patrick Henry (“Give me liberty or give me death!”) was a Borderer. They also also played a disproportionate role in westward expansion. After the Revolution, America made an almost literal 180 degree turn and the “backcountry” became the “frontier”. It was the Borderers who were happiest going off into the wilderness and fighting Indians, and most of the famous frontiersmen like Davy Crockett were of their number. This was a big part of the reason the Wild West was so wild compared to, say, Minnesota (also a frontier inhabited by lots of Indians, but settled by Northerners and Germans) and why it inherited seemingly Gaelic traditions like cattle rustling.

Their conception of liberty has also survived and shaped modern American politics: it seems essentially to be the modern libertarian/Republican version of freedom from government interference, especially if phrased as “get the hell off my land”, and especially especially if phrased that way through clenched teeth while pointing a shotgun at the offending party.

III.

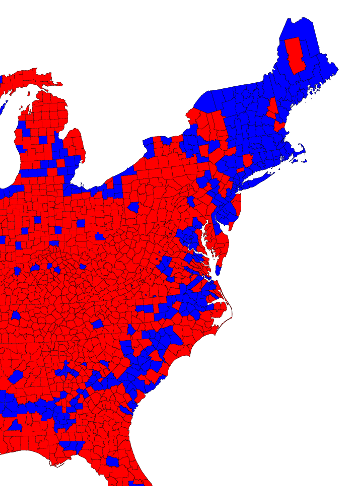
This is all interesting as history and doubly interesting as anthropology, but what relevance does it have for later American history and the present day?

One of my reasons reading this book was to see whether the link between Americans’ political opinions and a bunch of their other cultural/religious/social traits (a “Blue Tribe” and “Red Tribe”) was related to the immigration patterns it describes. I’m leaning towards “probably”, but there’s a lot of work to be done in explaining how the split among these four cultures led to a split among two cultures in the modern day, and with little help from the book itself I am going to have to resort to total unfounded speculation. But the simplest explanation – that the Puritans and Quakers merged into one group (“progressives”, “Blue Tribe”, “educated coastal elites”) and the Virginians and Borderers into another (“conservatives”, “Red Tribe”, “rednecks”) – has a lot going for it.

Many conservatives I read like to push the theory that modern progressivism is descended from the utopian Protestant experiments of early America – Puritanism and Quakerism – and that the civil war represents “Massachusetts’ conquest of America”. I always found this lacking in rigor: Puritanism and Quakerism are sufficiently different that positing a combination of them probably needs more intellectual work than just gesturing at “you know, that Puritan/Quaker thing”. But the idea of a Puritan New England and a Quaker-(ish) Pennsylvania gradually blending together into a generic “North” seems plausible, especially given the high levels of interbreeding between the two (some of our more progressive Presidents, including Abraham Lincoln, were literally half-Puritan and half-Quaker). Such a merge would combine the Puritan emphasis on moral reform, education, and a well-ordered society with the Quaker doctrine of niceness, tolerance, religious pluralism, individual conscience, and the Inner Light. It seems kind of unfair to just mix-and-match the most modern elements of each and declare that this proves they caused modernity, but there’s no reason that couldn’t have happened.

The idea of Cavaliers and Borderers combining to form modern conservativism is buoyed by modern conservativism’s obvious Border influences, but complicated by its lack of much that is recognizably Cavalier – the Republican Party is hardly marked by its support for a hereditary aristocracy of gentlemen. Here I have to admit that I don’t know as much about Southern history as I’d like. In particular, how were places like Alabama, Mississippi, et cetera settled? Most sources I can find suggest they were set up along the Virginia model of plantation-owning aristocrats, but if that’s true how did the modern populations come to so embody Fischer’s description of Borderers? In particular, why are they so Southern Baptist and not very Anglican? And what happened to all of those indentured servants the Cavaliers brought over after slavery put them out of business? What happened to that whole culture after the Civil War destroyed the plantation system? My guess is going to be that the indentured servants and the Borderer population mixed pretty thoroughly, and that this stratum was hanging around providing a majority of the white bodies in the South while the plantation owners were hogging the limelight – but I just don’t know.

A quick argument that I’m not totally making all of this up:



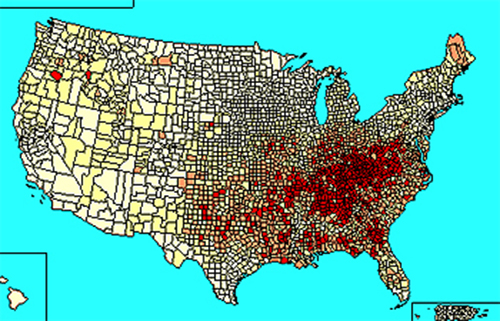
This is a map of voting patterns by county in the 2012 Presidential election. The blue areas in the South carefully track the so-called “black belt” of majority African-American areas. The ones in the Midwest are mostly big cities. Aside from those, the only people who vote Democrat are New England (very solidly!) and the Delaware Valley region of Pennsylvania. In fact, you can easily see the distinction between the Delaware Valley settled by Quakers in the east, and the backcountry area settled by Borderers in the west. Even the book’s footnote about how a few Borderers settled in the mountains of New Hampshire is associated with a few spots of red in the mountains of New Hampshire ruining an otherwise near-perfect Democratic sweep of the north.

One anomaly in this story is a kind of linear distribution of blue across southern Michigan, too big to be explained solely by the blacks of Detroit. But a quick look at Wikipedia’s History of Michigan finds:

In the 1820s and 1830s migrants from New England began moving to what is now Michigan in large numbers (though there was a trickle of New England settlers who arrived before this date). These were “Yankee” settlers, that is to say they were descended from the English Puritans who settled New England during the colonial era….Due to the prevalence of New Englanders and New England transplants from upstate New York, Michigan was very culturally contiguous with early New England culture for much of its early history…The amount with which the New England Yankee population predominated made Michigan unique among frontier states in the antebellum period. Due to this heritage Michigan was on the forefront of the antislavery crusade and reforms during the 1840s and 1850s.

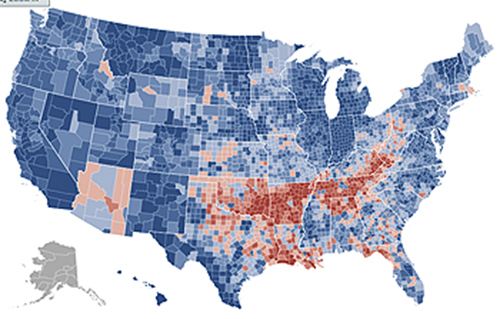
Alhough I can’t find proof of this specifically, I know that Michigan was settled from the south up, and I suspect that these New England settlers concentrated in the southern regions and that the north was settled by a more diverse group of whites who lacked the New England connection.

Here’s something else cool. We can’t track Borderers directly because there’s no “Borderer” or “Scots-Irish” option on the US census. But Albion’s Seed points out that the Borderers were uniquely likely to identify as just “American” and deliberately forgot their past ancestry as fast as they could. Meanwhile, when the census asks an ethnicity question about where your ancestors came from, every year some people will stubbornly ignore the point of the question and put down “America” (no, this does not track the distribution of Native American population). Here’s a map of so-called “unhyphenated Americans”, taken from this site:



We see a strong focus on the Appalachian Mountains, especially West Virginia, Tennesee, and Kentucky, bleeding into the rest of the South. Aside from west Pennsylvania, this is very close to where we would expect to find the Borderers. Could these be the same groups?

Meanwhile, here is a map of where Obama underperformed the usual Democratic vote worst in 2008:



These maps are small and lossy, and surely unhyphenatedness is not an exact proxy for Border ancestry – but they are nevertheless intriguing. You may also be interested in the Washington Post’s correlation between distribution of unhyphenated Americans and Trump voters, or the Atlantic’s article on Trump and Borderers.

If I’m going to map these cultural affiliations to ancestry, do I have to walk back on my previous theory that they are related to class? Maybe I should. But I also think we can posit complicated interactions between these ideas. Consider for example the interaction between race and class; a black person with a white-sounding name, who speaks with a white-sounding accent, and who adopts white culture (eg listens to classical music, wears business suits) is far more likely to seem upper-class than a black person with a black-sounding name, a black accent, and black cultural preferences; a white person who seems black in some way (listens to hip-hop, wears baggy clothes) is more likely to seem lower-class. This doesn’t mean race and class are exactly the same thing, but it does mean that some races get stereotyped as upper-class and others as lower-class, and that people’s racial identifiers may change based on where they are in the class structure.

I think something similar is probably going on with these forms of ancestry. The education system is probably dominated by descendents of New Englanders and Pennsylvanians; they had an opportunity to influence the culture of academia and the educated classes more generally, they took it, and now anybody of any background who makes it into that world is going to be socialized according to their rules. Likewise, people in poorer and more rural environments will be surrounded by people of Borderer ancestry and acculturated by Borderer cultural products and end up a little more like that group. As a result, ethnic markers have turned into and merged with class markers in complicated ways.

Indeed, some kind of acculturation process has to have been going on, since most of the people in these areas today are not the descendents of the original settlers. But such a process seems very likely. Just to take an example, most of the Jews I know (including my own family) came into the country via New York, live somewhere on the coast, and have very Blue Tribe values. But Southern Jews believed in the Confederacy as strongly as any Virginian – see for example Judah Benjamin. And Barry Goldwater, a half-Jew raised in Arizona, invented the modern version of conservativism that seems closest to some Borderer beliefs.

All of this is very speculative, with some obvious flaws. What do we make of other countries like Britain or Germany with superficially similar splits but very different histories? Why should Puritans lose their religion and sexual prudery, but keep their interest in moralistic reform? There are whole heaps of questions like these. But look. Before I had any idea about any of this, I wrote that American society seems divided into two strata, one of which is marked by emphasis on education, interest in moral reforms, racial tolerance, low teenage pregnancy, academic/financial jobs, and Democratic party affiliation, and furthermore that this group was centered in the North. Meanwhile, now I learn that the North was settled by two groups that when combined have emphasis on education, interest in moral reforms, racial tolerance, low teenage pregnancy, an academic and mercantile history, and were the heartland of the historical Whigs and Republicans who preceded the modern Democratic Party.

And I wrote about another stratum centered in the South marked by poor education, gun culture, culture of violence, xenophobia, high teenage pregnancy, militarism, patriotism, country western music, and support for the Republican Party. And now I learn that the South was settled by a group noted even in the 1700s for its poor education, gun culture, culture of violence, xenophobia, high premarital pregnancy, militarism, patriotism, accent exactly like the modern country western accent, and support for the Democratic-Republicans who preceded the modern Republican Party.

If this is true, I think it paints a very pessimistic world-view. The “iceberg model” of culture argues that apart from the surface cultural features we all recognize like language, clothing, and food, there are deeper levels of culture that determine the features and institutions of a people: whether they are progressive or traditional, peaceful or warlike, mercantile or self-contained. We grudgingly acknowledge these features when we admit that maybe making the Middle East exactly like America in every way is more of a long-term project than something that will happen as soon as we kick out the latest dictator and get treated as liberators. Part of us may still want to believe that pure reason is the universal solvent, that those Afghans will come around once they realize that being a secular liberal democracy is obviously great. But we keep having deep culture shoved in our face again and again, and we don’t know how to get rid of it. This has led to reasonable speculation that some aspects of it might even be genetic – something which would explain a lot, though not its ability to acculturate recent arrivals.

This is a hard pill to swallow even when we’re talking about Afghanistan. But it becomes doubly unpleasant when we think about it in the sense of our neighbors and fellow citizens in a modern democracy. What, after all, is the point? A democracy made up of 49% extremely liberal Americans and 51% fundamentalist Taliban Afghans would be something very different from the democratic ideal; even if occasionally a super-charismatic American candidate could win over enough marginal Afghans to take power, there’s none of the give-and-take, none of the competition within the marketplace of ideas, that makes democracy so attractive. Just two groups competing to dominate one another, with the fact that the competition is peaceful being at best a consolation prize.

If America is best explained as a Puritan-Quaker culture locked in a death-match with a Cavalier-Borderer culture, with all of the appeals to freedom and equality and order and justice being just so many epiphenomena – well, I’m not sure what to do with that information. Push it under the rug? Say “Well, my culture is better, so I intend to do as good a job dominating yours as possible?” Agree that We Are Very Different Yet In The End All The Same And So Must Seek Common Ground? Start researching genetic engineering? Maybe secede? I’m not a Trump fan much more than I’m an Osama bin Laden fan; if somehow Osama ended up being elected President, should I start thinking “Maybe that time we made a country that was 49% people like me and 51% members of the Taliban – maybe that was a bad idea“.

I don’t know. But I highly recommend Albion’s Seed as an entertaining and enlightening work of historical scholarship which will be absolutely delightful if you don’t fret too much over all of the existential questions it raises.

# Why Were Early Psychedelicists So Weird?

[Epistemic status: very speculative, asserted with only ~30% confidence. On the other hand, even though psychiatrists don’t really talk about this it’s possible other groups know this all already]

A few weeks ago I gave a presentation on the history of early psychedelic research. Since I had a tough crowd, I focused on the fascinating biographies of some of the early psychedelicists.

Timothy Leary was a Harvard professor and former NIMH researcher who made well-regarded contributions to psychotherapy and psychometrics. He started the Harvard Psilocybin Project and several other Harvard-based experiments to test the effects of psychedelics on normal and mentally ill subjects. He was later fired from Harvard and arrested; later he accomplished a spectacular break out of prison and fled to Algeria. During his later life, he wrote books about how the human brain had hidden circuits of consciousness that would allow us to live in space, including a quantum overmind which could control reality and break the speed of light. He eventually fell so deep into madness that he started hanging out with Robert Anton Wilson and participating in Ron Paul fundraisers.

Richard Alpert was Leary’s co-investigator at the Harvard Psilocybin Project. He, too, had all the signs of a promising career, including a psychology PhD from Stanford, a visiting professorship at Berkeley, and a combination academic/clinical position at Stanford. After his work with Leary, he moved to India, changed his name to Baba Ram Dass, and became one of the world’s most prominent advocates for bhakti yoga.

John Lilly was a doctor, a neuroanatomy researcher, and an inventor who helped develop the principle behind many modern neuroprosthetics. He was always very strange, and did a lot of work in human-dolphin communication and SETI even before starting his work with LSD. But in the 1960s, he ran across Richard Alpert, joined in his LSD experiments, and became even stranger. He started writing books with names like “Programming And Metaprogramming The Human Biocomputer”, and arguing that benevolent and malevolent aliens were locked in a battle to manipulate Earth’s coincidences and with them the future of the human species. He became an expert yogi and claimed to have achieved samadhi, the highest state of union with God.

Kary Mullis is kind of cheating since he was not technically a psychedelicist. He was a biochemist in the completely unrelated field of bacterial iron transport molecules. But he did try LSD in 1966 back when it was still a legal research chemical. In fact he tried 1000 micrograms of it, one of the biggest doses I’ve ever heard of someone taking. Like the others, Mullis was a brilliant scientist – he won the Nobel Prize in Chemistry for inventing the polymerase chain reaction. Like the others, Mullis got really weird fast. He is a global warming denialist, HIV/AIDS denialist, and ozone hole denialist; on the other hand, he does believe in the efficacy of astrology. He also believes he has contacted extraterrestrials in the form of a fluorescent green raccoon, and “founded a business with the intent to sell pieces of jewelry containing the amplified DNA of deceased famous people like Elvis Presley”.

I wondered if there might be a selection bias in which psychedelicists I heard about, or that I might be cherry-picking the most unusual examples, so I looked for leading early psychedelics researchers I’d never heard of before and checked how weird they were. My sources told me that the two most important early psychedelicists were Humphry Osmond (who invented the word ‘psychedelic’ and may have been the first person to experiment with LSD rigorously) and his colleague John Smythies.

Osmond has an impressive early resume: started off as a surgeon, became a psychiatrist, did some well-regarded research into the structure of the human metabolite adrenochrome. And although he did not become fluorescent-alien-raccoon level weird, he can’t quite be called normal either. He became one of the founders of orthomolecular psychiatry, a discipline arguing that schizophrenia and other psychiatric diseases can be cured by massive amounts of vitamins – this is currently considered pseudoscience. His publications include the article “Selection of twins for ESP experimentation” in International Journal of Parapsychology, and a history of psychedelics records that “after his mescaline experiment in 1951, Dr. Osmond claimed to have successfully transmitted telepathic information to a fellow researcher, Duncan Blewett, who was also under the influence of mescaline, leading an independent observer to panic at the uncanny event.” He seems to have maintained a lifetime interest in parapsychology, Jungian typological analysis, and a field of his own invention called “socio-architecture”.

Smythies was a neuropsychiatrist, neuroanatomist, biochemist, EEG researcher, editor of the International Review of Neurobiology, etc, etc, etc (also, a cousin of Richard Dawkins). He is 94 but apparently still alive and going strong and making new neuroanatomical discoveries. He was one of the first people to investigate the pharmacology of psychedelics and helped with Osmond’s experiments in the early 1950s. He has also written The Walls Of Plato’s Cave, a book presenting a new theory of consciousness which “extends our concepts of consciousness and analyses possible geometrical and topological relations between phenomenal space and physical space linked to brane theory in physics” (I kind of wish I was a fly on the wall at his and Dawkins’ family reunions).

My point is that the field of early psychedelic research seemed to pretty consistently absorb brilliant scientists, then spit out people who, while still brilliant scientists, also had styles of thought that could be described as extremely original at best and downright crazy at worst.

I think it’s important to try to understand why.

First possibility: you had to be kind of weird to begin with in order to be interested in researching psychedelics. On the one hand, this is surely true; on the other, the early psychedelicists ended up really weird. At least in the early days I’m not sure psychedelics had the reputation for weirdness they now enjoy, and I’m also not sure that we’re living in a world where a high enough percent of psychiatrists go off to become gurus in India, that we can just dismiss LSD research as happening to attract that type of person.

Second possibility: I know that almost all of these researchers (I’m not sure about Smythies) used psychedelics themselves. Psychedelic use is a sufficiently interesting experience that I can see why it might expand one’s interest in the study of consciousness and the universe. Perhaps this is especially true if you’re one of the first people to use it, and you don’t have the social setting of “Oh, yeah, this is that drug that makes you have really weird experiences about consciousness for a while”. If you’re not aware that psychedelic hallucinations are a thing that happens, you might have to interpret your experience in more traditional terms like divine revelation. Under this theory, these pioneers had to become kind of weird to learn enough for the rest of us to use these substances safely. But why would that make John Lilly obsessed with aliens? Why would it turn Timothy Leary into a space colonization advocate and Ron Paul supporter?

The third possibility is the one that really intrigues me. A 2011 study found that a single dose of psilocybin could permanently increase the personality dimension of Openness To Experience. I’m emphasizing that because personality is otherwise pretty stable after adulthood; nothing should be able to do this. But magic mushrooms apparently have this effect, and not subtly either; participants who had a mystical experience on psilocybin had Openness increase up to half a standard deviation compared to placebo, and the change was stable sixteen months later. This is really scary. I mean, I like Openness To Experience, but something that can produce large, permanent personality changes is so far beyond anything else we have in psychiatry that it’s kind of terrifying.

(related: 1972 study finds LSD may cause permanent increase in hypnotic susceptibility, which other sources have linked to being “fantasy prone” and “creative”)

And that’s one dose. These researchers were taking psychedelics pretty constantly for years, and probably experimented with the sort of doses you couldn’t get away with giving research subjects. What would you expect to happen to their Openness To Experience? How many standard deviations do you think it went up?

It seems possible to me that psychedelics have a direct pharmacological effect on personality that causes people to be more open to unusual ideas. I know this is going against most of the latest research, which says psychedelics have no long-term negative mental health effects and do not cause psychosis. But there’s a difference between being schizophrenic, and being the sort of guy who is still a leading neuroanatomist but also writes books about the geometric relationships between consciousness and the space-time continuum.

I’m not sure anyone has ever done studies to rule out the theory that psychedelics just plain make people weird. Indeed, such studies would be very difficult, given that weird people with very high Openness To Experience are more likely to use psychedelics. This problem would even prevent common sense detection of the phenomenon – even if we noticed that frequent psychedelic users were really weird, we would attribute it to selection effects and forget about it.

In this situation, the early psychedelicists could be a natural experiment giving us data we can’t get any other way. Here are relatively sober scientists who took psychedelics for reasons other than being weird hippies already. Their fate provides signal through the noise which is the general psychedelic-using population.

I think this is only medium-risk; the explanation that weird people gravitate toward psychedelics, even in the sciences, is a strong one. But it’s sufficient that I am hesitant to repeat the common view that psychedelics are not at all dangerous, or that they have no permanent side effects. There seems to me at least a moderate chance that they will make you more interesting without your consent – whether that is a good or a bad thing depends on exactly how interesting you want to be.

# Be Nice, At Least Until You Can Coordinate Meanness

[Epistemic status: idea for one’s toolbox of ideas; not to be followed off a cliff]

I.

Commenters on this blog have sometimes tried to shame or attack other commenters for perceived misdeeds like sexual promiscuity. They tell people to their faces that they’re bad people and try to humiliate them.

When this happens, I ban the commenters involved.

And I get protests – what about free speech? What about the marketplace of ideas? Isn’t shaming sometimes a useful social mechanism? There are some norms we can’t or shouldn’t codify into law; shouldn’t violation of those norms be punished by shaming? Shaming can be very effective – for example, last week we learned the Puritans had a premarital pregnancy rate near zero because they publicly shamed anyone who departed from their moral standards. Might it not be useful to have something like that nowadays, either for premarital sex, or for other evils like homophobia and racism that we want to discourage? And even if I think we shouldn’t, is it really okay to ban the people trying, seeing as they were probably well-intentioned?

I think my answer is: be nice, at least until you can coordinate meanness.

II.

A friend (I can’t remember who) once argued that “be nice” provides a nigh-infallible ethical decision procedure. For example, enslaving people isn’t very nice, so we know slavery is wrong. Kicking down people’s doors and throwing them in prison for having a joint of marijuana isn’t very nice, so we know the drug war is wrong. Not letting gays marry isn’t very nice, so we know homophobia is wrong.

I counterargue that even if we ignore the ways our notion of “nice” itself packs up pre-existing moral beliefs, this heuristic fails in several important cases:

1. Refusing the guy who is begging you to give his drivers’ license back, saying that without a car he won’t be able to visit his friends and family or have any fun, and who is promising that he won’t drive drunk an eleventh time.  
2. Forcibly restraining a screaming baby while you jam a needle into them to vaccinate them against a deadly disease.  
3. Sending the police to arrest a libertarian rancher in Montana who refuses to pay taxes for reasons of conscience  
4. Revoking the credential (and thus destroying the future job prospects of) a teacher who has sex with one of her underage students

Sure, you could say that each of these “leads toward a greater niceness”, like that you’re only refusing the alcoholic his license in order to be nice to potential drunk driving victims. But then you’ve lost all meaningful distinction between the word “nice” and the word “good” and reinvented utilitarianism. And reinventing utilitarianism is pretty cool, but after you do that you no longer have such an easy time arguing against the drug war – somebody’s going to argue that it leads to the greater good of there being fewer drugs.

We usually want to avoid meanness. In some rare cases, meanness is necessary. I think one check for whether a certain type of meanness might be excusable is – it’s less likely to be excusable if it’s not coordinated.

Consider: society demands taxes to pay for communal goods and services. This does sometimes involve not-niceness, as in the example of the rancher in (3). But what makes it tolerable is that it’s done consistently and through a coordinated process. If the rule was “anybody who has a social program they want can take money from somebody else to pay for it,” this would be anarchy. Some libertarians say “taxation is theft”, but where arbitrary theft is unfair, unpredictable, and encourage perverse incentives like living in fear or investing in attack dogs, taxation has none of these disadvantages.

By the rule “be nice, at least until you can coordinate meanness”, we should not permit individuals to rob each other at gunpoint in order to pay for social programs they want, but we might permit them to advocate for a coordinated national taxation policy.

Or: society punishes people for crimes, including the crime of libel. Punishment is naturally not-nice, but this seems fair; we can’t just have people libeling each other all the time with no consequences. But what makes this tolerable is that it’s coordinated – done through the court system according to carefully codified libel law that explains to everybody what is and isn’t okay. Remove the coordination aspect, and you’ve got the old system where if you say something that offends my honor then I get some friends and try to beat you up in a dark alley. The impulse is the same: deploy not-niceness in the worthy goal of preventing libel. But one method is coordinated and the other isn’t.

This is very, very far from saying that coordinated meanness is a sure test that means something’s okay – that would be the insane position that anything legal must be ethical, something most countries spent the past few centuries disproving spectacularly. This is the much weaker claim that legality sets a minimum bar for people attempting mean policies.

As far as I can tell there are two things we want in a legal system. First, it should have good laws that produce a just society. But second, it should at least have clear and predictable laws that produce a safe and stable society.

For example, the first goal of libel law is to balance people’s desire to protect their reputation with other people’s desire for free speech. But the second goal of libel law should be that everybody understands what is and isn’t libel. If a system achieves the second goal, nobody will end up jailed or dead because they said something they thought was totally innocent but somebody else thought was libel. And nobody will spent years and thousands of dollars entangled in an endless court case hiring a bunch of lawyers to debate whether some form of speech was acceptable or not.

So coordinated meanness is better than uncoordinated meanness not because it necessarily achieves the first goal of justice, but because it achieves the second goal of safety and stability. Everyone knows exactly when to expect it and what they can do to avoid it. I may not know what speech will or won’t offend a violent person with enough friends to organize a goon squad, but I can always read the libel law and try to stay on the right side of it.

Likewise, in the Puritan community, I know exactly what things I have to do to avoid being shamed. Better still, I can only be shamed for violating one set of moral standards – the shared moral standards of the whole community. This isn’t true of random people shaming promiscuous people, or people with the wrong opinion on race/gender issues, or whatever, on a private blog. Not only do most people reasonably expect to be able to do those things (and/or talk about those things here) without being shamed, but there are too many conflicting standards to meet – plausibly somebody could be shamed by traditionalists for being promiscuous, and by free-love people for not being promiscuous enough. Since shaming is unpleasant and supposed to act as a punishment, this is the equivalent of letting anybody beat up anybody else if they think they’ve broken an unwritten rule. It probably results in a lot of people being beaten up for not very much social change.

III.

The second reason that coordinated meanness is better than uncoordinated meanness is that it is less common. Uncoordinated meanness happens whenever one person wants to be mean; coordinated meanness happens when everyone (or 51% of the population, or an entire church worth of Puritans, or whatever) wants to be mean. If we accept theories like the wisdom of crowds or the marketplace of ideas – and we better, if we’re small-d democrats, small-r republicans, small-l liberals, or basically any word beginning with a lowercase letter at all – then a big group of people all debating with each other will be harder to rile up than a single lunatic.

As a Jew, if I heard that skinheads were beating up Jews in dark alleys, I would be pretty freaked out; for all I know I could be the next victim. But if I heard that skinheads were circulating a petition to get Congress to expel all the Jews, I wouldn’t be freaked out at all. I would expect almost nobody to sign the petition

(and in the sort of world where most people were signing the petition, I hope I would have moved to Israel long before anyone got any chance to expel me anyway)

Trying to coordinate meanness is not in itself a mean act – or at least, not as mean as actual meanness. If Westboro Baptist Church just published lots of pamphlets saying we should pass laws against homosexuality, maybe it would have made some gay people feel less wanted, but it would have been a lot less intense than picketing funerals. If people who are against promiscuity want to write books about why we should all worry about promiscuity, it might get promiscuous people a little creeped out, but a lot less so than going up to promiscuous people and throwing water on them and shouting “YOU STRUMPET!”

This is my answer to people who say that certain forms of speech make them feel unsafe, versus certain other people who demand the freedom to express their ideas. We should all feel unsafe around anybody who relishes uncoordinated meanness – beating people in dark alleys, picketing their funerals, shaming them, harassing them, doxxing them, getting them fired from their jobs. I have no tolerance for these people – I am sometimes forced to accept their existence because of the First Amendment, but I won’t do anything more.

On the other hand, we should feel mostly safe around people who agree that meanness, in the unfortunate cases where it’s necessary, must be coordinated. There is no threat at all from pro-coordination skinheads except in the vanishingly unlikely possibility they legally win control of the government and take over.

I admit that this safety is still only relative. It hinges on the skinheads’ inability to convert 51% of the population. But until the Messiah comes to enforce the moral law directly, safety has to hinge on something. The question is whether it should hinge on the ability of the truth to triumph in the marketplace of ideas in the long-term across an entire society, or whether it should hinge on the fact that you can beat me up with a baseball bat right now.

(if you want pre-Messianic absolute safety, there are some super-democratic mechanisms that might help. America’s Bill of Rights seems pretty close to this; anyone wanting to coordinate meanness against a certain religion has to clear not only the 50% bar, but the much higher level required of Constitutional amendments. Visions of more complete protection remain utopian but alluring. For example, in an Archipelago you might well have absolute safety. The skinheads can’t say “Let’s beat up Jews right now”, they can’t even say “Let’s start an anti-Jew political party and gradually win power”. They can, at best, say, “Let’s go found our own society somewhere else without any Jews”, in which case you need say nothing but “don’t let the door hit you on your way out”. In this case their coordination of meanness cannot possibly hurt anyone.)

IV.

I’ve said many times I find the idea of “safe spaces” very attractive. I think they can be understood not just as spaces that are guaranteed safe for one group, but as spaces that have coordinated meanness against anything that threatens that group – ie they’ve agreed to shame, shun, and expel people who violate group norms. Everybody knows the local norms, and if somebody gets kicked out they can’t say they weren’t warned.

This is the principle with which I deal with the blog comments I started off by talking about. Right now people come to this blog with a default expectation that people aren’t going to be mean to them or try to shame them for things, other than the things universally agreed to be shameful in these general circles (like trolling, spamming, and misusing one-tailed t-tests). I want to explicitly reinforce that expectation here.

If you support being meaner in certain ways for the greater good, either as a subculture or as a society, you’re welcome to try to use this blog to advocate for that policy (within reason), but you’re not welcome to enact that policy unilaterally.

So here are two previously implicit SSC rules, made explicit for your edification:

First, you’re allowed to make (polite) arguments for why we should try shaming certain groups, but you are not allowed to directly shame any commenters here.

Second, you’re allowed to (politely) express your philosophical disagreements with the idea of transgender, but you are not allowed to actually misgender transgender commenters here.

# Myers’ Race Car Versus The General Fitness Factor

[Epistemic status: I am not a geneticist, and even the geneticists I know aren’t sure about a lot of this. Take as speculation only.]

I.

PZ Myers argues against Stephen Hsu’s genetic engineering proposal here – a disappointing attitude toward mad science for a guy whose blog header is a crocodile-duck hybrid. The piece has a lot of errors, the worst of which other people have already discussed – but I want to talk about what I think is its strongest point. Myers writes:

Note [Hsu’s] estimate of the number of genes that contribute to IQ: 10,000. That’s half the human genome! Hmmm. I wonder if any of those genes play a role in other processes in human physiology that might be affected by his plan?

Here’s an analogy for you: let’s say a novice car designer has decided that the one quality of an automobile that is most important is speed, raw speed. He doesn’t know much about cars, so he asks more qualified engineers about what elements of the car contribute to acceleration and velocity, and they start off with the obvious…details of the engine, fuel mixes, etc. Then they’re talking tires. Aerodynamics. Weight. Pretty soon they have to admit that just about everything in the car is going to affect the speed at which it travels.

So our blithe designer decides that making a fast car is simple: we just look at each component of the car one by one, and we pick an available option for it entirely on the basis of which option makes the car go faster. We’ll easily be able to make a car that can rocket along at a thousand miles an hour, he thinks.

But we have to ask whether we would want a car where the seats and steering were optimized for speed, where safety options were discarded, where something like visibility or reliability were jettisoned for the sole virtue of going really fast.

This makes a lot of sense. A car in which every component was optimized for speed would probably be uncomfortable, unsafe, ugly, difficult to maintain, and otherwise not the sort of car you want to drive. A human in which every component was optimized for intelligence might well be unhealthy, ugly, physically weak, antisocial, et cetera.

And we can do more than just hand-wave at an analogy to cars. Natural selection constantly weeds out worse alleles and replaces them with better ones. If an allele increases intelligence enough to improve reproductive fitness, but has no negative effects, it should sweep across all human populations in an amount of time proportional to its fitness benefit1. We see genetic evidence of various alleles sweeping various human populations in both the distant and recent past. These do not include the intelligence-boosting alleles Hsu is talking about.

For example, Hsu cites Rietvald et al‘s finding that rs1487441 is linked to cognitive ability (though it only gives you 0.3 extra IQ points, typical of the generally unimpressive contribution of single genes). About 20% of both Europeans and Japanese have the (A:A) variant, which suggests that however many thousands of years it’s been since Europeans and Japanese diverged from each other isn’t long enough for the gene to undergo much selection. That means neither allele can have any overwhelming advantage, which means there must be some reason to have the opposite allele worth as much as 0.3 IQ points. I think this is the rigorous version of what Myers is saying.

Despite the fact that the race car argument makes perfect sense both analogically and rigorously, it seems to be wrong.

II.

What are the sorts of things we might trade off against intelligence? Perhaps fitness, height, attractiveness, health, longevity, social well-adjustedness?

But in fact none of these trade off against intelligence, many are strongly positively associated with it, and in some the link has been proven genetic!

People with high IQ tend to live longer. For example, a person with IQ 115 (85th percentile) is 20% more likely to survive to age 76 than an average person with IQ 100. One can of course posit many possible connections. Maybe high-IQ people are smart enough to eat healthy and exercise. Maybe rich people can afford both good schools and good doctors. Maybe good health behaviors protect the brain as well as the body and increase both IQ and longevity. But further investigation has cast doubt on all of these theories and strongly supports the hypothesis that no, the same genes that give you high intelligence also make you live longer. See for example the International Journal of Epidemiology: The Link Between Intelligence And Lifespan Is Mostly Genetic, which find genetics explain 95% of the correlation. A few of the genes linking intelligence and longevity may be already known; SSADH seems to be a contributing factor. My favorite study in this area, though, is one that is not yet complete: since all mammals are basically the same [citation needed] some London School of Economics researchers have developed an IQ test for dogs in the hope of checking whether the same correlation applies to them. Since canine intelligence doesn’t affect things like diet, exercise, or tobacco status, a positive correlation in them too would help solidify the finding. We’re still waiting on those results, but even without them the genetic hypothesis is looking pretty strong.

People with high IQ tend to be taller. This is interesting since height is often used as a measure of health and fitness during childhood, and since taller people get a bunch of advantages including being rated as more attractive and earning higher income. Once again we can imagine all sorts of possible confounders; once again studies find that the link is genetic. See for example Common Genetic Variants Explain The Majority Of The Correlation Between Height And Intelligence, The Genetic Correlation Between Height And IQ: Shared Genes Or Assortative Mating, Resolving The Genetic And Environmental Sources Of The Correlation Between Height And Intelligence, On The Height-Intelligence Correlation.

People with high IQ may be more attractive. This is the conclusion of a meta-analysis that finds a positive correlation between intelligence and body symmetry, usually used as a proxy for attractiveness unaffected by things like hairstyle and cosmetics; a second study failed to find this relationship. The jury is out on the positive link, but there certainly isn’t the negative link that Myers’ race car would predict.

People with high IQ commit much less crime – which is going to be our measure for social well-adjustedness here since it’s well studied. Once again it’s easy to think of possible confounders (I’ll add lead levels to the usual lot). Once again the studies show that at least some of the effect is genetic – here’s one on low-IQ/antisocial-behavior correlation in children, and here’s one cleverly linking fathers’ criminal history to sons’ vs. nephews’ IQ and then throwing enough statistics at it to find that the relationship is genetic. Likewise, the relationship between high IQ and low drug abuse seems to be genetic as well.

People with high IQ tend to be more physically fit. This is the conclusion of a study of 1.2 million Swedes. I don’t have any strong evidence that this relationship is genetically mediated (although Gottfredson on the fitness factor may be relevant here), I just want to note that, once again, there is less than zero evidence for Myers’ race car hypothesis.

People with high IQ have lower rates of heart disease, stroke, circulatory diseases, and diabetes. Intelligence may or may not decrease cancer risk, but again contra the race car hypothesis, it certainly does not increase it. Sibling designs suggest that shared family environment during youth is not responsible for the benefits; differential socioeconomic status as an adult may be, but this status is itself likely caused be the intelligence differences.

So despite its apparent plausibility the race car hypothesis crashes and burns.

III.

In one sense, this is bizarre. It’s as if somebody optimized every part of a race car for speed, and found that by coincidence this also made it the safest, most comfortable, and cheapest car on the road.

Yet if we think about it some more maybe it’s not too surprising. Consider Niels Bohr. He was a Nobel Prize winning physicist, professional football player, activist who helped Jews flee the Nazis, loving father of six children, and so healthy he kept doing science well into his seventies. And his talents show every sign of being at least partly genetic – his brother Harald was a leading mathematician, anti-Nazi activist, and Olympic silver medalist; one of Bohr’s children was also a Nobel Prize winning physicist and another was also an Olympic athlete. So it’s obviously possible to design a human with all-around great genes. Why does evolution restrict such designs to the Bohr family?

I can think of a few possibilities, all of which people who know more than I do are welcome to shoot down.

First, some of these all-around-beneficial genes could be good in heterozygosity but bad in homozygosity. We know something similar is true in the case of sickle-cell anaemia, which is mostly good in heterozygosity (protects vs. malaria) and very bad in homozygosity (causes sickle cell). This is exactly the sort of gene that should exist at a constant low frequency in the population, never getting more or less common. If the frequency got too low, then there’s no risk of two carriers mating, so evolution would encourage it as a free disease cure. If it got too high, evolution would discourage it – any carrier would probably marry another carrier and give their kids sickle cell. Suppose there are ten such genes, each of which grants higher intelligence on heterozygosity and has a frequency of 10% in the population. The average person will on average carry one such gene and have a 10% chance of a horrible genetic disease. Maybe Niels Bohr lucked out and carried all ten such genes without going homozygous on any. Maybe some other poor guy who is lost to history got all ten genes homozygous and died at birth of ten horrible genetic diseases at once.

This would make Hsu’s gene-editing project very promising; all he would need to do is give everybody one copy of the relevant genes (and then never let them mate). But the hypothesis can’t be quite right: I think it would predict that Niels Bohr’s children would have unusually high rates of genetic diseases. In fact, the children of great men regress to the mean a little bit but show no signs at all of being unusually cursed.

Second, we could be talking not about polymorphisms but about mutational load. That means that there’s some genome that works for humans (plus or minus a few hundred thousand polymorphisms that aren’t too important at this level of analysis) and genetic health is determined by how many detrimental mutations you and your parents randomly accreted. If your mother spent too much time near the local nuclear reactor when she was pregnant, maybe you get a few hundred extra mutations and end up with lower IQ, a worse heart, less attractive features, et cetera. And this is obviously true in the case of a literal nuclear reactor, but I’m having trouble figuring out what plays the reactor role in real life. I know Greg Cochran and others have talked about things like paternal age at conception, climate, et cetera, but he applies these only to differences between populations. I’m not sure whether it would work out to expect a big difference in mutational load between Niels Bohr and his underachieving next-door neighbor. Maybe Bohr came from a long line of people who lucked out and got hit by unusually few cosmic rays? I don’t know if this makes sense or not. Part of my problem might be that I still don’t really understand how mutational load ever decreases – I’ve heard “the most heavily-loaded people are weeded out by natural selection”, but it seems like that should only be able to slow the gradual universal dysgenesis.

This would also bode well for Hsu’s project. In fact, it would make it even easier; it would reduce to the modal genome idea (make a baby whose genome, at each location, has the nucleotide which is most common at that location among all humans worldwide) which could be done without even performing the groundwork to see which genes do or don’t influence intelligence.

Third, maybe all of these other good things are trading off against things that were important in the environment of evolutionary adaptedness but not today. Greg Cochran brings up infectious disease resistance; some commenters bring up calorie requirements. This latter seems especially plausible; the brain uses a lot of energy and energy was a scarce resource through much of evolutionary history. Either of these would explain why evolution kept the seemingly detrimental version around for so long, and why right now in our low-infection high-calorie modern civilization one allele or another seems to be an unalloyed good.

Again, this would bode well for Hsu’s project, although the supergeniuses so produced would probably want to stay well away from any malarial swamps.

Or maybe it is some mixture of all four possibilities – Myers’ trade-offs, heterozygosity, mutational load, and disease burden. The latter three could provide a sufficiently positive effect for intelligence to hide the negative effect of the first. I would be really surprised if something like this wasn’t true – the theoretical argument for the first seems compelling, and it ought to happen at least a little even if it isn’t the main source driving intelligence differences.

I was going to write that in this case we’d have to sort through every intelligence gene one-by-one to make sure we weren’t getting one of the trade-off ones, but maybe this isn’t true – a genius designed by Hsu’s method should have on average the same number of trade-offs versus unalloyed-goods as a genius born normally (right? or am I missing something?). Since most of us would prefer a natural-born baby with IQ 150 to a natural-born baby with IQ 100, it seems whatever trade-offs are necessary to reach that point are widely considered worth it. So unless there’s a difference I’m missing between normal recombination and Hsu’s method, we should be okay with designing an IQ 150 baby as well (from a purely health-related perspective, at least).

Whether this generalizes to creating an IQ IQ 200 or 300 baby depends not just on ethics, but on whether for some reason the costs and trade-offs of intelligence compound more than linearly. It’s possible, for example, that there are ten different genes coding for something that protects heart health, all of which can be traded off against intelligence. If you switch one gene from heart health to intelligence, whatever, you still have nine genes protecting your heart. But switching all of them at once would be a bad idea. I don’t know any reason to think this is true, but it’s a possibility that might give us pause between the IQ 150 and the IQ 300 level.

Overall though, I think the race car idea, despite its plausibility, is likely to be less of an impediment to genetic engineering than it might seem.

Only One Footnote, But It’s Really Long

1. This is an assumption I’m granting for the sake of argument, but possibly not true at the margin.

Consider that evolution doesn’t care about intelligence nearly as much as we do. The most recent common ancestor of Europeans and Japanese wasn’t going to use her intelligence to design a mammoth-seeking rocket. In fact, it’s not totally clear why humans did evolve intelligence before the modern age; sure, tools are nice, but early hominids stuck to the same tools for a million years at a stretch; that doesn’t exactly give a tight feedback loop to work with. The most convincing argument I’ve heard is the Machiavellian intelligence hypothesis which says that our ancestors used intelligence to navigate tribal politics and gain status within a social group.

But this theory would naively predict that the smartest person in high school would be the most popular. If intelligence is for gaining status, it seems to have diminishing returns beyond a certain point, which would explain why evolution didn’t generally make us more intelligent even though greater-than-average intelligence is clearly possible (eg geniuses).

In the rare cases where evolution did have an incentive to evolve higher intelligence, it did so quickly and effectively. Several highly mercantile societies independently evolved the same set of genes producing higher IQ. The most notable were the Ashkenazi Jews, who have an average IQ 12-15 points higher than their European neighbors and whose genes show strong signatures of recent selection for intelligence; this most likely occurred during the Middle Ages when they were the mercantile class of Europe, since non-Ashkenazi Jews show no such effect. The genes involved tend to produce sphingolipidoses when homozygous, which shows a pretty good reason why evolution didn’t do this kind of thing more often. Myers has previously dismissed this research, but I think wrongly – the paper itself considers and rejects the all the criticisms he raises (see pages 15 – 31). The very short summary is that Myers dismisses the genetic pattern as “variations amplified by chance,” but the expected level of chance variations can be calculated and this isn’t it. Ashkenazim have similar heterozygosity to other Europeans in neutral markers – ruling out a simple bottleneck – and the mutations involved are too potentially deleterious in homozygosity to persist for many generations by chance alone. Further, the mutations are all clustered in a few key pathways, many of which are clearly linked to intelligence. For example, Ashkenazim are at high risk for torsion dystonia, which is associated with higher IQ in sufferers.

So in response to the argument that evolution must trade off against something else, I would argue that evolution doesn’t share our exchange rate. Suppose that we could gain 20 IQ points at the cost of having larger heads that are harder to fit through a birth canal (remember, some of the known genes for intelligence are associated with head size, and the obstetrical dilemma used to be a big deal!). For hunter-gatherers, who had little use for IQ but lots of use for getting through birth canals, this was a bad deal and evolution didn’t take it. For moderns, who can use IQ points to cure cancer and explore space, and who have modern obstetric techniques, it’s a lot more attractive. So yes, let’s be cautious, but I think we’d all feel pretty stupid if we avoided bootstrapping our way to superintelligence out of fears of “things man was not meant to meddle with”, only to learn later that the whole problem could have been solved with c-sections.

# Hidden Test Post

Hidden test post

# Solidarity

[Epistemic status: I am not British, it’s been years since I’ve been in the HSE, and the HSE is not the NHS. All of this may be misunderstood or outdated.]

I don’t usually blog on labor disputes here, but I want to talk about one on which I have a tiny bit of inside knowledge.

Last month junior doctors in Britain went on strike for two days, protesting imposition of a new contract. There’s a lot of anger about this, and admittedly when you’re being rushed by ambulance into the emergency department for sudden onset chest pain, “doctors are on strike today” is not something you want to hear. My normal instincts would be to question whether this is really necessary. My experience tells me it is.

“Oh, you’re a junior doctor. Of course you would support a doctor’s strike.” Okay, but I’m not a British junior doctor. I work in America, where I would describe conditions as “tough, but fair”. Sure, Dr. Cox yells at you a lot, but only because he secretly thinks you’re one of the best doctors ever to pass through the doors of this hospital. My own specialty of psychiatry is a lot better than most and overall I have little to complain about in my own life.

But that’s not to say that I don’t have any special knowledge here. I went to medical school in Ireland, where I worked alongside junior doctors in a system very much based off of the British one. And it was pretty shocking.

Technically European law caps junior doctor work weeks at 48 hours a week. Then again, technically American law caps junior doctor work weeks at 80 hours a week. My first week on a non-psychiatry service as an American junior doctor, I worked a bit over 100 hours – and so did everybody else I encountered. When I asked about the law, everyone just gave me that “oh you sweet summer child” look.

Such caps seem to be honored more in the breach than in the observance, and this is the British custom too. Physicians Weekly describes it as “the 48 hour trainee work week sham”, and the Telegraph and The Daily Mail both seem to agree that many British doctors are working 100 hour shifts. They seem to circumvent the law either by giving them a few weeks off afterwards and saying it “averages” to 48 hours/week, or else by doing what my hospital did – carefully schedule a 48 hour shift in big bold letters, assign 100 hours worth of work, and then get angry if anyone goes home before their work is done.

Many of the junior doctors I worked with in Ireland were working a hundred hours a week. It’s hard to describe what working 100 hours a week is like. Saying “it means you work from 7 AM to 9 PM every day including weekends” doesn’t really cut it. Imagine the hobbies you enjoy and the people you love. Now imagine you can’t spend time on any of them, because you are being yelled at as people die all around you for fourteen hours a day, and when you get home you have just enough time to eat dinner, brush your teeth, possibly pay a bill or two, and curl up in a ball before you have to go do it all again, and your next day off is in two weeks.

And this is the best case scenario, where everything is spaced out nice and even. The junior doctors I knew frequently worked thirty-six hour shifts at a time (the European Court of Human Rights has since declined to fine Ireland for this illegal practice). Dr. Brid McGrath (my lab partner in medical school) has been collecting some stories for the Irish media:

My stories are like my colleagues’ stories: working through illness, personal turmoil, and deprivation of sleep, food and toilet breaks. The worst stint was working 73 hours within an 82 hour period. I have been bullied, and to my shame, bullied others. I realised I was falling into the trap of treating others the way I had been treated. My self esteem faltered and I began to believe I truly was a nasty person. I had the insight to get help, but not everyone is so lucky.

I came to talk to you about the imminent arrival of your very premature baby, at just 24 weeks. I held your hand and passed you tissues as we talked about his name, how tiny he was, how hard his life could be, but how we would try to give him the best possible chance… and how we might also have to accept the reality that he might not make it. That same day, I’d worked a 28-hour shift while I was 24 weeks pregnant myself. I fought back tears before I saw you. I worried about how I would cope with your pain and distress, barely able to think about the baby growing inside me. I had dinner a 1am, and worked on. An incredible nurse sat me down for a glass of water. She had to force me to. I was so busy.

The other night, after a particularly busy 16-hour shift in the Emergency Department and Theatre, I went up to the wards to take blood samples for a patient. I’d had no dinner. A patient in the same room was handing around coconut buns, and gave me one. I inhaled it, it smelt so good. She then pushed the box into my hands and said “You look like you need them more than me!”

Imagine having to decide between going to the bathroom or getting a bag of crackers from the vending machine because you don’t have enough time between cases to do both. Imagine having to remember the difference between nephritic syndrome and nephrotic syndrome (two totally different things) after ten hours of work, after getting three hours of sleep the night before. Imagine that you’ve just admitted a neurotic old woman to the hospital and you know in your heart that you should take her hand and explain to her in a soothing voice that everything is going to be okay, except that you already feel like every nerve of yours is beaten raw and you have three patients left to go before you can so much as sit down for a few minutes. Imagine your attending yelling at you because you got something wrong and saying you need to spend more time studying, and you trying to keep your mouth shut instead of telling him that you literally have only a half-hour in the day that could be considered free time by even the broadest stretch of the imagination and you are damned if you are going to spend that studying endocrinology.

The psychological consequences are predictable: after one year, 55% of junior doctors describe themselves as burned out, 30% meet criteria for moderate depression, and 12% report considering suicide.

A lot of American junior doctors are able to bear this by reminding themselves that it’s only temporary. The worst part, internship, is only one year; junior doctorness as a whole only lasts three or four. After that you become a full doctor and a free agent – probably still pretty stressed, but at least making a lot of money and enjoying a modicum of control over your life.

In Britain, this consolation is denied most junior doctors. Everyone works for the government, and the government has a strict hierarchy of ranks, only the top of which – “consultant” – has anything like the freedom and salary that most American doctors enjoy. It can take ten to twenty years for junior doctors in Britain to become consultants, and some never do. In Ireland (I don’t know about the UK) there was a very scary distinction between “training” and “service” positions, the former of which were always in short supply. Imagine that you’re a freshman in college, and your university announces that due to budget cutbacks there are only about half as many sophomore positions available this year, so the top fifty percent of freshmen can go on to become sophomores, and the rest will have to stay freshmen until more money comes in. Also, there are no other colleges in the entire country so you have no choice but to follow along and hope for the best. This is what being a junior doctor is like.

Faced with all this, many doctors in Britain and Ireland have made the very reasonable decision to get the heck out of Britain and Ireland. The modal career plan among members of my medical school class was to graduate, work the one year in Irish hospitals necessary to get a certain certification that Australian hospitals demanded, then move to Australia. In Ireland, 47.5% of Irish doctors had moved to some other country. The situation in Britain is not quite so bad but rapidly approaching this point. Something like a third of British emergency room doctors have left the country in the past five years, mostly to Australia, citing “toxic environment” and “being asked to endure high stress levels without a break”. Every year, about 2% of British doctors apply for the “certificates of good standing” that allow them to work in a foreign medical system, with junior doctors the most likely to leave. Doctors report back that Australia offers “more cash, fewer hours, and less pressure”. I enjoy a pretty constant stream of Facebook photos of kangaroos and the Sydney Opera House from medical school buddies who are now in Australia and trying to convince their colleagues to follow in their footsteps.

Upon realizing their doctors are moving abroad, British and Irish health systems have leapt into action by…ignoring all systemic problems and importing foreigners from poorer countries who are used to inhumane work environments. I worked in some rural Irish towns where 99% of the population was white yet 80% of the doctors weren’t; if you have a heart attack in Ireland and can’t remember what their local version of 911 is, your best bet is to run into the nearest mosque, where you’ll find all the town’s off-duty medical personnel conveniently gathered together. This seems to be true of Britain as well, with the stats showing that almost 40% of British doctors trained in a foreign country (about half again as high as the US numbers, even though the US is accused of “stealing the world’s doctors” – my subjective impression is that foreign doctors try to come to the US despite barriers because they’re attracted to the prospect of a better life here, but that they are actively recruited to Britain out of desperation). Many of the doctors who did train in Britain are new immigrants who moved to Britain for medical school – for example, the Express finds that only 37% of British doctors are white British (the corresponding number for America is something like 50-65%, even though America is more diverse than Britain). While many new immigrants are great doctors, the overall situation is unfortunate since a lot of them end up underemployed compared to their qualifications in their home country, or trapped in the lower portions of the medical hierarchy by a combination of racism, language difficulties, and just the fact that everyone is trapped in the lower portions of the medical hierarchy these days.

If Britain continues along its current course, they’ll probably be able to find more desperate people willing to staff its medical services after even more homegrown doctors move somewhere else (70% say they’re considering it, although we are warned not to take that claim at face value). I work with several British and Irish doctors in my hospital here in the US Midwest, they’re very talented people, and we could always use more of them. But this still seems like just a crappy way to run a medical system.

I don’t know anything about the latest dispute that has led to this particular strike in Britain. Both sides’ positions sound reasonable when I read about them in the papers. I would be tempted to just split the difference, if not for the fact several years of medical work in the British Isles have taught me that everything that a government health system says is vile horrible lies, and everybody with a title sounding like “Minister of Health” or “Health Secretary” is an Icke-style lizard person whose terminal value is causing as many humans to die of disease as possible. I can’t overstate the importance of this. You read the press releases and they sound sort of reasonable, and then you talk to the doctors involved and they tell you all of the reasons why these policies have destroyed the medical system and these people are ruining their lives and the lives of their patients and how they once shook the Health Secretary’s hand and it was ice-cold and covered in scales. I don’t know how much of this is true. I just think of it as something in the background when the health service comes up to doctors and says “Hey, we have this great new deal we want to offer you!”

(I remember reporting into the hospital one day and seeing almost a carnival atmosphere, and one surgeon who had never been known to do anything but yell at his subordinates gave me a friendly nod and smile as he passed me in the corridor, and I started to worry I had walked into some Stepford Wives bizarro-world. Finally I learned that, the evening before, the Irish health minister had resigned in disgrace. This is the only time anyone ever saw that surgeon happy.)

[EDIT: a strong argument that the junior doctors have the right of it and the NHS’ position is based on a misunderstanding of patient care statistics here]

Whatever caused this latest dispute is probably relevant mostly as a straw that breaks the camel’s back. If British junior doctors today are anything like the Irish junior doctors of a few years ago, all of their complaints are legitimate and they’re also hiding several dozen other legitimate complaints you have yet to hear about. I sort of sympathize with the government’s complaints that they don’t have enough money to make a system where doctors don’t have to work a bunch of 36 hour shifts, but I feel like if you don’t have enough money to run a health system that treats its employees like human beings, maybe you shouldn’t be running your country’s health system.

Labor disputes suck, and I have no good theory of them. Part of me is outraged at people being mistreated, and another part of me worries about a world where anybody who can convince the media that they’re being oppressed can force other people into paying them whatever amount of money they think they deserve. I long for some kind of principled system that will solve these problems more elegantly than letting everybody shout their grievances at each other and seeing which ones stick. I long for something that will take care of the deeper problems underlying unfair labor practices like dualization of entire industries. This is why I find libertarian ideas like letting competition among firms determine people’s pay and conditions so attractive.

But these may or may not work, insofar as they do work they only work in certain situations, and insofar as they do work under certain situations a 100% socialized industry run as a government monopoly probably isn’t one of them. So we’ve got to do the thing where people get mistreated and have to cry out for redress of their grievances. And my experience tells me the grievances of British junior doctors are copious, horrifying, and entirely valid.

# Skin In The Game

I.

One of the most interesting responses I got to my post supporting the junior doctors strike was by Salem, who said that this situation was (ethically) little different than that around adjunct professors, who also become overworked and miserable trying to break into a high-status profession. Salem very kindly didn’t directly accuse me of hypocrisy, but maybe he should have.

While I sympathize with adjuncts’ terrible conditions, my natural instinct is to say feedback mechanisms should keep doing their work. You can probably trace the argument- imagine a simplified toy model where the only two jobs are professor and salesperson, and being a professor is fun and high-status but being a salesperson is boring and low-status. Everyone will become a professor, and this will decrease the demand for professors and increase the demand for salespeople until the employers involved change their policies accordingly. Eventually it will stabilize where the nonmonetary advantages of being a professor are perfectly compensated by the monetary advantages of being a salesperson. If professors are getting paid shockingly little, it means the system is sending a signal that the nonmonetary advantages of being a professor are shockingly high, or else why would people keep trying? If we demand that professors get paid more, then we’re letting them keep all their nonmonetary advantages over salespeople but demanding they have monetary advantages as well. It destroys the system’s incentives to have people go into less fun but nevertheless necessary fields.

All of this makes perfect sense in the adjunct case. So why do I feel so differently in the doctor case?

Maybe for personal reasons. When I was in college, my two dueling career plans were doctor and philosophy professor. I brought this up with my professors, who universally told me not to go into academia. They told me that it was grueling, thankless, and for the vast majority of people involved doomed to failure, and that they couldn’t in good conscience advise me to try it. I listened to their advice and became a doctor instead. It might not have quite the I-can’t-believe-they’re-paying-me-to-do-this amazingness of debating metaethics all day, but I still love it and it has much better career prospects.

So I guess you could argue that one reason I have less sympathy for adjuncts is that letting them achieve their goal would be a kick in the face to Past-Scott, who made what he considered the sober choice and went into a better-paying profession. If it turns out all I had to do was hang on a few years, and then the government would decree that people who got paid to argue about metaethics had to have career prospects as good as doctors’, then I was a huge chump to try to do things the hard way. Maybe it’s that fear of chumpness that makes it harder to sympathize.

And maybe the reason I feel such solidarity with doctors is that it’s not supposed to be a profession you go into knowing you have no hope. Healing the sick is a lot more practical and socially-subsidizable an activity than pondering Truth; it seems like the sort of thing it should be easy to get paid for. Here in America, this is the conventional wisdom: make it into medicine, and you’re promised a pretty good career. So maybe my solidarity with British doctors is a big cultural misunderstanding. Maybe, coming from America, I’ve absorbed a social promise that doctors will be treated well (which is true in America), but in Britain those doctors go into medical school knowing 100% that their lives will be unbearable and their compensation miniscule. Maybe they do it anyway because for them medicine is as much of an I-can’t-believe-they’re-paying-me-to-do-this as philosophy is to me.

Suppose somebody tells me that before going to medical school, every doctor in the British Isles has to sign a Waiver Of Appreciation Of Consequences, which spells out in excruciating detail all of the horrible things about a medical career. It says exactly how many 100 hour weeks they’ll be expected to work, exactly how many 36-hour shifts they’ll take, does its best to give them an idea how much senior doctors will abuse them. Maybe there’s even a video portion that has interviews with disgruntled current doctors describing in explicit detail all the worst things that have happened to them. Maybe they even have to shadow current junior doctors, work 36 hour shifts themselves even though they can’t do any procedures yet, just so that they can never claim that they didn’t understand on an experiential level what a 36-hour shift means. Suppose the government (unlike in the real world) does this 100% fairly and accurately, and then they stick to it – ie as bad as things are, they never get any worse than the video promised. And suppose that after all of this, representatives from Private Industry come in to discuss their options for non-medical jobs, and explain how conditions in these jobs will probably be much better. If a medical student signs this waiver, completes their training, graduates, hates their job, and demands better pay and conditions, can we have any sympathy for them?

(I’m talking about a hypothetical world here. In the real world, we have to consider that lots of people didn’t know what they were going into, and the thing they were going into keeps getting worse than they were led to expect. This is purely a Least Convenient Possible World argument here.)

The first argument we might use is that instead of focusing on the virtue of the employee, we should focus on the problems with society. That is, who cares whether the doctor made a good choice or not? Society still owes people decent jobs and working conditions. I think I reject this argument. There are some decent jobs attainable by the sorts of people who could become doctors (this is not true for everyone, but it’s pretty true of the people who could become doctors), and not every alternative has to be great for everybody. Suppose there are many entry-level gardening jobs available, but I become a skyscraper-window-washer. Then I complain because I am afraid of heights, and I want special accomodations for this. When you say “maybe you should try a different line of work”, I say “stop making this about my virtue and start focusing on the problems with society such that it can’t give everyone decent conditions.” It’s not a very sympathetic argument. Society is totally allowed to have jobs that not everybody would enjoy as long as it gives everybody a wide range of options.

(one might argue that nobody could possibly enjoy being a junior doctor, but weirdly enough this is false. There are some people who love it. I usually assume these people are on cocaine, but maybe some of them aren’t.)

The second argument we might use is that instead of focusing on the virtue of the employee, we should focus on distortions in the economy. Who decided to have a medical career consist of ten to twenty years of misery followed by an elusive opportunity to get a really nice job with great pay and hours? Isn’t part of the problem that hospitals aren’t competing for junior doctors? Can we oppose dualization of the labor force as just a generally bad idea? I think this argument is right, but I also think there’s still a lingering question of “Okay, but since we do have this terrible industry with bad incentives, should we feel sympathy for people who voluntarily place themselves right in the middle of it?”

The third argument we might use is that we sometimes need to save people from their terrible decisions. Just as we ban people from permanently selling themselves into slavery, no matter how aware they are of the consequences when they make the deal, so maybe we should assume that signing the Waiver Of Appreciation of Consequences is an irrational decision motivated by time discounting and that people’s future selves ought to be freed from the tyranny of their past selves’ poor judgment. There’s definitely some truth to this, but it’s also a little too close to the policies of the BETA-MEALR Party for comfort. And doesn’t this ignore that doctors aren’t slaves, and can leave the medical profession any time they want? Even if there are some reasons they can’t (difficulty finding other jobs, sunk cost effects) wouldn’t it be less distortionary to smooth their path to non-medical careers than to try to reform medicine?

All of this makes sense. And yet as a psychiatrist, I constantly have people come into my office saying their jobs are making them suicidal. You would think somebody would leave a job before it makes them suicidal, but this doesn’t always happen (the same is true of relationships, by the way). My suggestions that maybe they find a different job or boyfriend tend to fall on deaf ears.

I think one big reason I am so much more sympathetic to doctors than adjuncts is that I know them and have worked with them and I can see a reality – both a real reality and an emotional reality – that makes the model look kind of weak and threadbare.

II.

A few years ago I rented a house. In the rental contract, it said “tenant must have their own rental insurance to cover damages to the property”. I reminded my landlord that rental insurance mostly just covers damages to the renter’s stuff, and that the landlord usually has a separate property insurance to cover damages to the property, and that the way he was doing it might not even be legal. He just said yes it was and I had better get the rental insurance. I made a mental note to get the rental insurance and then got distracted by everything else in life and never got around to it.

A few years later, a pipe burst and the house flooded. The damage was assessed at way more than I had the ability to pay. I told the landlord he had better fix it, and the landlord told me I had better get my rental insurance to fix it for me – ie, the rental insurance I had forgotten to get.

There followed a gigantic disaster. I asked some lawyers whether the landlord was legally required to insure his own property, and they gave me vague and conflicting answers, then all agreed that even if he was it would cost more money than I had to fight the case in court. I ended up on the verge of breakdown. The landlord was clearly deliberately stonewalling me, trying to make it as hard as possible for me to figure out what was going on in the hopes that I gave in and gave him the money I didn’t have. Eventually, after several months’ of living out of a friend’s spare room while I tried to get the fate of the house sorted out, we settled out of court for a big fraction of my life’s savings.

(I later got some evidence that the landlord did have the property insured and was running some kind of insurance scam trying to get the money out of both me and his insurance, but that’s a different story and I’m still not sure either way)

The reason I bring this up was that for the duration of the crisis, and to a lesser degree even to the present, I was utterly convinced that the government had an obligation to make sure landlords insure their properties. Offering a contract where the tenant was responsible for insuring the property? Totally unacceptable. Maybe even a human rights violation. What’s the new phrase people are using these days? “A denial of your right to exist”? Even so.

It’s easy to craft the argument where I’m in the wrong. Different landlords should be allowed to experiment with different arrangements as long as their terms are clearly listed in their contracts and the other contracting party agrees to them. I signed the contract then failed to do what the contract said, and clearly I needed to pay the penalty. We can imagine some nanny-state laws that might ameliorate that – for example, if a contract says a tenant must purchase insurance, it’s the landlord’s responsibility to make sure the insurance was actually purchased – but the arguments for implementing these laws are on really shaky ground.

And all I’m saying is that in the middle of this crisis, I had no sympathy for any of this. In the middle of this crisis, my thought was that I had worked really hard for years to get a little bit of money saved up, I was going to lose all of it for some burst pipe that wasn’t my fault, this was cosmically unfair, somebody needed to do something about it, the landlord was a big company that probably had millions of dollars, and somebody needed to do something about this right now.

III.

An article by Freddie deBoer in this month’s Current Affairs proposes “Journalistic Self-Outsourcing”. DeBoer notes that lots of journalists and intellectuals suggest that protectionism and other anti-globalization policies are immoral. For example, Zack Beauchamp of Vox calls Bernie Sanders’ skepticism of free trade “screwing the global poor”; Brad deLong calls the same “a call to keep China a society of poor subsistence rice farmers as long as possible – keep them poor, barefoot, uneducated, and by no means allow them to work at any of the high-value manufacturing occupations we want to keep in the United States.”

DeBoer has a few things to say about how we should take money away from the rich in a way that can help both poor Americans and poor Chinese, but he quickly transitions into the barb he clearly relishes: if Beauchamp and DeLong are so in favor of poor Chinese people, how come they haven’t outsourced their jobs? They both earn quite a bit of money, and China must have some decent journalists and economists who would love to telecommute to well-paid American positions. Are DeLong and Beauchamp hypocrites for not arranging to expedite the transition of their jobs to Chinese people?

One argument in their favor: they’re currently living in a relatively globalized world, their employers must have the option of replacing them with Chinese people, but both of them still have jobs. This suggests that maybe journalism and economics aren’t as replaceable as deBoer thinks. Beauchamp might be able to retort “As soon as Vox finds a Chinese journalist who’s as good as I am they’re welcome to hire him/her, but until then I can stay on, secure in the consistency of my principles.” Presumably he can say this without fear, since the subtleties of being in touch with the pulse of the American people and writing English-language articles would elude most lower-class Chinese.

But this lets them off too easily for purely contingent reasons. Even granting that they can’t be replaced with Chinese people right away, are they forced to will their replacement with Chinese people, deep in their hearts? Should Beauchamp go into work every morning asking his boss whether he’s found a suitable Chinese person to replace him with yet, and be disappointed every time his boss says no?

Again, there’s a contingent argument otherwise. If you’re really pro-globalization, you might believe that it’s impossible for the Chinese to take all our jobs, in the same way that the Luddite Fallacy says it’s impossible for robots to take all our jobs. The more Chinese take manufacturing jobs, the more Americans will have lots of money which will encourage new service jobs that the Chinese can’t easily take. If this is Beauchamp’s argument, he could say that we should globalize all the jobs that can be globalized, including his if possible, and then he will just move to an unglobalizable job. Since his job hasn’t been globalized yet, maybe he’s already in an unglobalizable job, so people should just stop bothering him.

But this is still too contingent. Let’s least convenient possible world again, and suppose that economists determine that there are permanent negative effects on American jobs from Chinese globalization, those effects disproportionately go to the poorest Americans, and no new unglobalizable jobs arise to restore cosmic balance. Now does Beauchamp have to will his own replacement?

Okay, argument from the other side: suppose a 1980s version of Beauchamp is writing an article against apartheid in South Africa. The South African whites argue that if apartheid ends, they’ll be competing for jobs against much poorer blacks and so their quality of life will go way down. They say that if Beauchamp really wants to end apartheid, he himself should give up his journalist position to a South African black who will take it for a few dollars a day.

I’m both very convinced that the right thing to do in that situation would be to fight apartheid, and also convinced that the South Africans would be right about the personal jab – the median American journalist who pushed the fight wouldn’t want his job taken over by Zulus willing to work for lower wages. When I put myself in that situation, and imagine myself being undercut by foreigners willing to practice medicine for almost nothing, I’m pretty pissed off about the idea too.

So maybe we should just let the journalist keep being a hypocrite. Journalists are pretty privileged people, and if we call journalists hypocrites every time they stand up for the less privileged without giving up all their own privilege, we’ll probably just end up with journalists who stop standing up for the less privileged.

IV.

Consider two contradictory arguments.

The first says that people who experience a problem have unique insight into it, and people who don’t experience it debate it from a position of ignorance or callousness. Thus a rich person can say “The poor don’t need help because they can pull themselves up by their own bootstraps”, but a poor person knows this is harder than it sounds. There’s a really boring sense of this in which the poor person may know specific facts the rich person is missing – for example, the rich person might falsely think welfare is much more generous than it is. But then a rich person who’s read a lot of books on poverty might claim to have a better perspective than the poor person, since she would likely know more about the welfare system. The more interesting claim is that there’s a sort of near-mode-vs-far-mode thing going on here, where things that are easy to dismiss in the abstract become a lot more relevant in the concrete. A really good example of this is Hitchens’ waterboarding – he said he thought it was an acceptable interrogation technique, some people offered to waterboard him to see if he changed his mind, and he quickly did. I’m fascinated by this incident because it’s hard for me to understand, in a Mary’s Room style sense, what he learned from the experience. He must have already known it was very unpleasant – otherwise why would you even support it as useful for interrogation? But somehow there’s a difference between having someone explain to you that waterboarding is horrible, and undergoing it yourself. Just so with everything else. Under this view, the only people we should trust to tell us whether junior doctors get treated fairly are junior doctors; the only people we should trust to tell us whether globalization is acceptable are people whose own jobs are on the line.

The second says that people with skin in the game are the last people we should be trusting. Who do you trust to tell you how many subsidies the government should give the oil companies? Some economist in the budget watchdog organization who calculates exactly what the costs and benefits are? Or an oil company CEO who says “Trust me, we need lots and lots of money”? Under this argument, everyone has access to logic and reason, people detached from the situation are able to use it, and people within the situation are (perhaps excusably) motivated by self-interest. We can perhaps understand the fears of the white South African who thinks he’ll lose his job and end up bankrupt, but it’s our job as dispassionate external observers to notice that his concerns are outweighed by the concerns of other people whom his self-interest makes him unable to understand. And the last person you want giving you a sober cost-benefit analysis of torture is the person who is being waterboarded – everyone knows a waterboarding victim will say anything at all to make it stop!

The second argument obviously has its uses, but I’m fascinated by the first. In the few cases where I have direct experience with it, it seems to bring knowledge beyond just “this is really bad”. In fact, the anti-junior-doctor argument seems resistant to just learning medicine is worse than you thought – if anything, medicine being very bad makes the argument stronger, since it means the doctor should be extra quick to listen to feedback and go into a different career. But somehow experience with doctors has made me much more reluctant to believe that argument. In the same way, one could imagine Hitchens saying “Yes, this waterboarding is really unpleasant – good thing that means it’ll be really easy to make the terrorists want to talk” – but that wasn’t the conclusion he drew from it.

I guess the thing I’m not sure about is – does personal experience/”skin in the game” reduce fully to factual propositions? Does a factory worker have an advantage over a journalist in understanding globalization just because he knows that being laid off is really bad, and that it’s harder to get a new job than a journalist thinks – two things we would expect any journalist worth their salt to already know about? Or is there some hard-to-communicate knowledge that’s neither factual nor just a cover for “the secret hard-to-communicate knowledge that I am selfish and want a system that benefits me rather than other people”?

Awkwardly, my far mode says that there isn’t and my near mode says that there is.

# Teachers: Much More Than You Wanted To Know

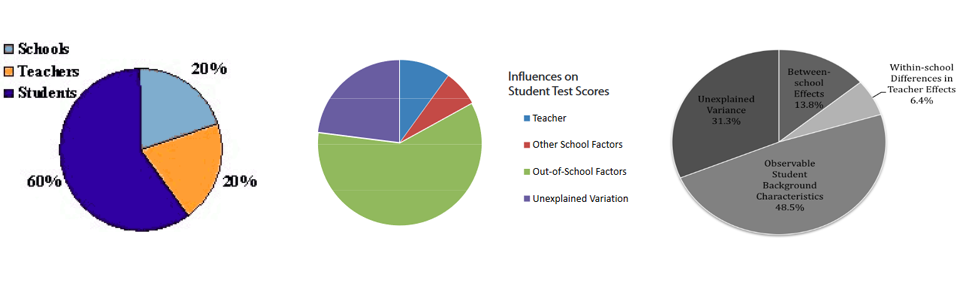
[Epistemic status: This is really complicated, this is not my field, people who have spent their entire lives studying this subject have different opinions, and I don’t claim to have done more than a very superficial survey. I welcome corrections on the many inevitable errors.]

I.

Newspapers report that having a better teacher for even a single grade (for example, a better fourth-grade teacher) can improve a child’s lifetime earning prospects by $80,000. Meanwhile, behavioral genetics studies suggest that a child’s parents have minimal (non-genetic) impact on their future earnings. So one year with your fourth-grade teacher making you learn fractions has vast effects on your prospects, but twenty-odd years with your parents shaping you at every moment doesn’t? Huh? I decided to try to figure this out by looking into the research on teacher effectiveness more closely.

First, how much do teachers matter compared to other things? To find out, researchers take a district full of kids with varying standardized test scores and try to figure out how much of the variance can be predicted by what school the kids are in, what teacher’s class the kids are in, and other demographic factors about the kids. So for example if the test scores of two kids in the same teacher’s class were on average no more similar than the test scores of two kids in two different teachers’ classes, then teachers can’t matter very much. But if we were consistently seeing things like everybody in Teacher A’s class getting A+s and everyone in Teacher B’s class getting Ds, that would suggest that good teachers are very important.

Here are the results from three teams that tried this (source, source, source):



These differ a little in that the first one assumes away all noise (“unexplained variance”) and the latter two keep it in. But they all agree pretty well that individual factors are most important, followed by school and teacher factors of roughly equal size. Teacher factors explain somewhere between 5% and 20% of the variance. Other studies seem to agree, usually a little to the lower end. For example, Goldhaber, Brewer, and Anderson (1999) find teachers explain 9% of variance; Nye, Konstantopoulos, and Hedges (2004) find they explain 13% of variance for math and 7% for reading. The American Statistical Association summarizes the research as “teachers account for about 1% to 14% of the variability in test scores”, which seems about right.

So put more simply – on average, individual students’ level of ability grit is what makes the difference. Good schools and teachers may push that a little higher, and bad ones bring it a little lower, but they don’t work miracles.

(remember that right now we’re talking about same-year standardized test scores. That is, we’re talking about how much your fourth-grade history teacher affects your performance on a fourth-grade history test. If teacher effects show up anywhere, this is where it’s going to be.)

Just as it’s much easier to say “this is 40% genetic” than to identify particular genes, so it’s much easier to say “this is 10% dependent on school-level factors and 10% based on teacher-level factors” then to identify what those school-level and teacher-level factors are. The Goldhaber study above tries its best, but the only school-level variable they can pin down is that having lots of white kids in your school improves test scores. And as far as I can tell, they don’t look at socioeconomic status of the school or its neighborhood, which is probably what the white kids are serving as a proxy for. Even though these “school level effects” are supposed to be things like “the school is well-funded” or “the school has a great principal”, I worry that they’re capturing student effects by accident. That is, if you go to a school where everyone else is a rich white kid, chances are that means you’re a rich white kid yourself. Although they try to control for this, having a couple of quantifiable variables like race and income probably doesn’t entirely capture the complexities of neighborhood sorting by social class.

In terms of observable teacher-level effects, the only one they can find that makes a difference is gender (female teachers are better). Teacher certification, years of experience, certification, degrees, et cetera have no effect. This is consistent with most other research, such as Miller, McKenna, and McKenna (1998). A few studies that we’ll get to later do suggest teacher experience matters; almost nobody wants to claim certifications or degrees do much.

One measurable variable not mentioned here does seem to have a strong ability to predict successful teachers. I’m not able to access these studies directly, but according to the site of the US Assistant Secretary of Education:

The most robust finding in the research literature is the effect of teacher verbal and cognitive ability on student achievement. Every study that has included a valid measure of teacher verbal or cognitive ability has found that it accounts for more variance in student achievement than any other measured characteristic of teachers (e.g., Greenwald, Hedges, & Lane, 1996; Ferguson & Ladd, 1996; Kain & Singleton, 1996; Ehrenberg & Brewer, 1994).

So far most of this is straightforward and uncontroversial. Teachers account for about 10% of variance in student test scores, it’s hard to predict which teachers do better by their characteristics alone, and schools account for a little more but that might be confounded. In order to say more than this we have to have a more precise way of identifying exactly which teachers are good, which is going to be more complicated.

II.

Suppose you want to figure out which teachers in a certain district are the best. You know that the only thing truly important in life is standardized test scores [citation needed], so you calculate the average test score for each teacher’s class, then crown whoever has the highest average as Teacher Of The Year. What could go wrong?

But you’ll probably just give the award to whoever teaches the gifted class. Teachers have classes with very different ability, and we already determined that innate ability grit explains more variance than teacher skill, so teachers who teach disadvantaged children will be at a big, uh, disadvantage.

So okay, back up. Instead of judging teachers by average test score, we can judge them by the average change in test score. If they start with a bunch of kids who have always scored around twentieth percentile, and they teach them so much that now the kids score at the fortieth percentile, then even though their kids are still below average they’ve clearly done some good work. Rank how many percentile points on average a teacher’s students go up or down during the year, and you should be able to identify the best teachers for real this time.

Add like fifty layers of incomprehensible statistics and this is the basic idea behind VAM (value-added modeling), the latest Exciting Educational Trend and the lynchpin of President Obama’s educational reforms. If you use VAM to find out which teachers are better than others, you can pay the good ones more to encourage them to stick around. As for the bad ones, VAM opponents are only being slightly unfair when they describe the plan as “firing your way to educational excellence”.

A claim like “VAM accurately predicts test scores” is kind of circular, since test scores are what we used to determine VAM. But I think the people in this field try to use the VAM of class c to predict the student performance of class c + 1, or other more complicated techniques, and Chetty, Rothstein, and Rivkin, Hanushek, and Kane all find that a one standard deviation increase in teacher VAM corresponds to about a 0.1 standard deviation increase in student test scores.

Let’s try putting this in English. Consider an average student with an average teacher. We expect her to score at exactly the 50th percentile on her tests. Now imagine she switched to the best teacher in the whole school. My elementary school had about forty teachers, so this is 97.5th percentile eg two standard deviations above the mean. A teacher whose VAM is two standard deviations above the mean should have students who score on average 0.2 standard deviations above the mean. Instead of scoring at the 50th percentile, now she’ll score at the 58th percentile.

Or consider the SAT, which is not the sort of standardized test involved in VAM but which at least everybody knows about. Each of its subtests is normed to a mean of 500 and an SD of 110. Our hypothetical well-taught student would go from an SAT of 500 to an SAT of 522. Meanwhile, average SAT subtest score needed to get into Harvard is still somewhere around 740. So this effect is nonzero but not very impressive.

But what happens if we compound this and give this student the best teachers many years in a row? Sanders and Rivers (also Jordan, Mendro, and Weerasinghe) argue the effects are impressive and cumulative. They compare students in Tennessee who got good teachers three years in a row to similar students who got bad teachers three years in a row (good = top quintile; bad = bottom quintile, so only 1/125 students was lucky or unlucky enough to qualify). The average bad-bad-bad student got scores in the 29th percentile; the average good-good-good student got scores in the 83rd percentile – which based on the single-teacher results looks super-additive. This is starting to sound a lot more impressive, and maybe Harvard-worthy after all. In fact, occasionally it is quoted as “four consecutive good teachers would close the black-white achievement gap” (I’m not sure whether this formulation requires also assigning whites to four consecutive bad teachers).

A RAND education report criticizes these studies as “using ad hoc methods” and argue that they’re vulnerable to double-counting student achievement. That is, we know that this teacher is the best because her students get great test scores; then later on we return and get excited over the discovery that the best teachers’ students get great test scores. Sanders and Rivers did some complicated things that ought to adjust for that; RAND runs simulations and finds that depending on the true size of teacher effects vs. student effects, those complicated things may or may not work. They conclude that “[Sanders and Rivers] provide evidence of the existence and persistence of teacher or classroom effects, but the size of the effects is likely to be somewhat overstated”.

Gary Rubinstein thinks he’s debunked Sanders and Rivers style studies. I strongly disagree with his methods – he seems to be saying that the correlation between good teaching and good test scores isn’t exactly one and therefore doesn’t matter – but he offers some useful data. Just by eyeballing and playing around with it, it looks like most of the gain from these “three consecutive great teachers” actually comes from the last great teacher. So the superadditivity might not be quite right, and Sanders and Rivers might just be genuinely finding bigger teacher effects than anybody else.

At what rate do these gains from good teachers decay?

They decay pretty fast. Jacob, Lefgren and Sims find that only 25% of gains carry on to the next year, and only 15% to the year after that. That is, if you had a great fourth grade teacher who raised your test scores by x points, in fifth grade your test scores will be 0.25x higher than they would otherwise have been. Kane and Rothstein find much the same. A RAND report suggests 20% persistence after one year and 10% persistence after two. Jacob, Lefgren, and Sims find that only 25% of gains remain after one year, and about 13% after two years, after which it drops off much more slowly. All of this contradicts Sanders and Rivers pretty badly.

None of these studies can tell us whether the gains go all the way to zero after a long enough time. Chetty does these calculations and finds that they stabilize at 25% of their original value. But this number is higher than the two-year number for most of the other studies, plus Chetty is famous for getting results that are much more spectacular and convenient than anybody else’s. I am really skeptical here. I remember a lot more things about last year than I do about twenty years ago, and even though I am pretty sure that my sixth grade teacher (for some weird reason) taught our class line dancing, I can’t remember a single dance step. And remember Louis Benezet’s early 20th century experiments with not teaching kids any math at all until middle school – after a year or two they were just as good as anyone else, suggesting a dim view of how useful elementary school math teachers must be. And even Chetty doesn’t really seem to want to argue the point, saying that his results “[align] with existing evidence that improvements in education raise contemporaneous scores, then fade out in later scores”.

In summary, I think there’s pretty strong evidence that a +1 SD increase in teacher VAM can increase same-year test scores by + 0.1 SD, but that 50% – 75% of this effect decays in the first two years. I’m less certain how much these numbers change when one gets multiple good or bad teachers in a row, or how fully they decay after the first two years.

III.

When I started looking for evidence about how teachers affected children, I expected teachers’ groups and education specialists to be pushing all the positive results. After all, what could be better for them than solid statistical proof that good teachers are super valuable?

In fact, these groups are the strongest opponents of the above studies – not because they doubt good teachers have an effect, but because in order to prove that effect you have to concede that good teaching is easy to measure, which tends to turn into proposals to use VAM to measure teacher performance and then fire underperformers. They argue that VAM is biased and likely to unfairly pull down teachers who get assigned less intelligent lower-grit kids.

It’s always fun to watch rancorous academic dramas from the outside, and the drama around VAM is really a level above anything else I’ve seen. A typical example is the blog VAMboozled! with its oddly hypnotic logo and a steady stream of posts like Kane Is At It Again: “Statistically Significant” Claims Exaggerated To Influence Policy. Historian/researcher Diane Ravitch doesn’t have quite as cute an aesthetic, but she writes things like:

VAM is Junk Science. Looking at children as machine-made widgets and looking at learning solely as standardized test scores may thrill some econometricians, but it has nothing to do with the real world of children, learning, and teaching. It is a grand theory that might net its authors a Nobel Prize for its grandiosity, but it is both meaningless in relation to any genuine concept of education and harmful in its mechanistic and reductive view of humanity.

But tell us how you really feel.

I was originally skeptical of this, but after reading enough of these sites I think they have some good points about how VAM isn’t always a good measure.

First, it seems to depend a lot on student characteristics; for example, it’s harder to get a high VAM in a class full of English as a Second Language students. It makes perfect sense that ESL students would get low test scores, but since VAM controls for prior achievement you might expect them to get the same VAM anyway. They don’t. Also, a lot of VAM models control for student race, gender, socioeconomic status, et cetera. I guess this is better than not doing this, but it seems to show a lack of confidence – if controlling for prior achievement was enough, you wouldn’t need to control for these other things. But apparently people do feel the need to control for this stuff, and at that point I bring up my usual objection that you can never control for confounders enough, and also all to some degree these things are probably just lossy proxies for genetics which you definitely can’t control for enough.

Maybe because of this, there’s a lot of noise in VAM estimates. Goldhaber & Hansen (2013) finds that a teacher’s VAM in year t is correlated at about 0.3 with their VAM in year t + 1. A Gates Foundation study also found reliabilities from 0.19 to 0.4, averaging about 0.3. Newton et al get slightly higher numbers from 0.4 to 0.6; Bessolo a wider range from 0.2 to 0.6. But these are all in the same ballpark, and Goldhaber and Hanson snarkily note that standardized tests aimed to assess students usually need correlations of 0.8 to 0.9 to be considered valid (the SAT, for example, is around 0.87). Although this suggests there’s some component of VAM which is stable, it can’t be considered to be “assessing” teachers in the same way normal tests assess students.

Even if VAM is a very noisy estimate, can’t the noise be toned down by averaging it out over many years? I think the answer is yes, and I think the most careful advocates of VAM want to do this, but President Obama wants to improve education now and a lot of teachers don’t have ten years worth of VAM estimates.

Also, some teachers complain that even averaging it out wouldn’t work if there are consistent differences in student assignment. For example, if Ms. Andrews always got the best students, and Mr. Brown always got the worst students, then averaging ten years is just going to average ten years of biased data. Proponents argue that aside from a few obvious cases (the teacher of the gifted class, the teacher of the ESL class) this shouldn’t happen. They can add school-fixed effects into their models (eg control for average performance of students at a particular school), leaving behind only teacher effects. And, they argue, which student in a school gets assigned which teacher ought to be random. Opponents argue that it might not be, and cite Paufler and Amrein-Beardsley‘s survey of principals, in which the principals all admit they don’t assign students to classes randomly. But if you look at the study, the principals say that they’re trying to be super-random – ie deliberately make sure that all classes are as balanced as possible. Even if they don’t 100% achieve this goal, shouldn’t the remaining differences be pretty minimal?

Maybe not. Rothstein (2009) tries to “predict” students’ fourth-grade test scores using their fifth-grade teacher’s VAM and finds that this totally works. Either schools are defying the laws of time and space, or for some reason the kids who do well in fourth-grade are getting the best fifth-grade teachers. Briggs and Domingue not only replicate these effects, but find that a fifth-grade teacher’s “effects” on her students in fourth-grade is just as big as her effect on her students when she is actually teaching them, which would suggest that 100% of VAM is bias. Goldhaber has an argument for why there are statistical reasons this might not be so damning, which I unfortunately don’t have enough understanding grit to evaluate.

Genetics might also play a role in explaining these results (h/t Spotted Toad’s excellent post on the subject). A twin study by Robert Plomin does the classical behavioral genetics thing to VAM and finds that individual students’ nth grade VAM is about 40% to 50% heritable. That is, the change in your test scores between third to fourth grade will probably be more like the change in your identical twin’s test scores than like the change in your fraternal twin’s test scores.

At first glance, this doesn’t make sense – since VAM controls for past performance, shouldn’t it be a pretty pure measure of your teacher’s effectiveness? Toad argues otherwise. One of those Ten Replicated Findings From Behavioral Genetics is that IQ is more shared environmental in younger kids and more genetic in older kids. In other words, when you’re really young, how smart you are depends on how enriched your environment is; as you grow older, it becomes more genetically determined.

So suppose that your environment is predisposing you to an IQ of 100, but your genes are predisposing you to an IQ of 120. And suppose (pardon the oversimplification) that at age 5 your IQ is 100, at age 15 it’s 120, and change between those ages is linear. Then every year you could expect to gain 2 IQ points. Now suppose there’s another kid whose environment is predisposing her to an IQ of 130, but whose genes are predisposing her to an IQ of 90. At age 5 her IQ is 130, at age 15 it’s 90, and so every year she is losing 4 IQ points. And finally, suppose that your score on standardized tests is exactly 100% predicted by your IQ. Since you gain two points every year, in fifth grade you’ll gain two points on your test, and your teacher will look pretty good. She’ll get a good VAM, a raise, and a promotion. Since your friend loses four points every year, in fifth grade she’ll lose four points on her test, and her teacher will look incompetent and be assigned remedial training.

This critique meshes nicely with the Rothstein test. Since you’re gaining 2 points every year, Prof. Rothstein can use your 5th grade gains of +2 points to accurately predict your fourth grade gain of +2 points. Then he can use your friend’s 5th grade loss of -4 points to accurately predict her fourth grade loss of -4 points.

This is a very neat explanation. My only concern is that it doesn’t explain decay effects very well. If a fifth grade teacher’s time-bending effect on students in fourth grade is exactly the same as her non-time-bending effect on students in fifth grade, how come her effect on her students once they graduate to sixth grade will only be 25% as large as her fifth grade effects? How come her seventh-grade effects will be smaller still? Somebody here has to be really wrong.

It would be nice to be able to draw all of this together by saying that teachers have almost no persistent effects, and the genetic component identified by Plomin and pointed at by Rothstein represents the 15 – 25% “permanent” gain identified by Chetty and others which so contradicts my lack of line dancing memories. But that would be just throwing out Briggs and Domingue’s finding that the Rothstein effect explains 100% of identified VAM.

One thing I kept seeing in the best papers on this was an acknowledgement that instead of arguing “VAMs are biased!” versus “VAMs are great!”, people should probably just agree that VAMs are biased, just like everything else, and start figuring out ways to measure exactly how biased they are, then use that number to determine what purposes they are or aren’t appropriate for. But I haven’t seen anybody doing this in a way I can understand.

In summary, there are many reasons to be skeptical of VAM. But some of these reasons contradict each other, and it’s not clear that we should be infinitely skeptical. A big part of VAM is bias, but there might also be some signal within the noise, especially when it’s averaged out over many years.

IV.

So let’s go back to that study that says that a good fourth grade teacher can earn you $89,000. The study itself is Chetty, Friedman, and Rockoff (part 1, part 2). You may recognize Chetty as a name that keeps coming up, usually attached to findings about as unbelievable as these ones.

Bloomberg said that “a truly great” teacher could improve a child’s earnings by $80,000, but I think this is mostly extrapolation. The number I see in the paper is a claim that a 1 SD better fourth-grade teacher can improve lifetime earnings by $39,000, so let’s stick with that.

This sounds impressive, but imagine the average kid works 40 years. That means it’s improving yearly earnings by about $1,000. Of note, the study didn’t find this. They found that such teachers improved yearly earnings by about $300, but their study population was mostly in their late twenties and not making very much, and they extrapolated that if good teachers could increase the earnings of entry-level workers by $300, eventually they could increase the earnings of workers with a little more experience by $1000. The authors use a lot of statistics to justify this assumption which I’m not qualified to assess. But really, who cares? The fact that having a good fourth grade teacher can improve your adult earnings any measurable amount is the weird claim here. Once I accept that, I might as well accept $300, $1,000, or $500,000.

And here’s the other weird thing. Everyone else has found that teacher effects on test scores decay very quickly over time. Chetty has sort of found that up to 25% of them persist, but he doesn’t really seem interested in defending that claim and agrees that probably test scores just fade away. Yet as he himself admits, good teachers’ impact on earnings works as if there were zero fadeout of teacher effects. He and his co-authors write:

Our conclusion that teachers have long-lasting impacts may be surprising given evidence that teachers’ impacts on test scores “fade out” very rapidly in subsequent grades (Rothstein 2010, Carrell and West 2010, Jacob, Lefgren, and Sims 2010). We confirm this rapid fade-out in our data, but find that teachers’ impacts on earnings are similar to what one would predict based on the cross-sectional correlation between earnings and contemporaneous test score gains.

They later go on to call this a “pattern of fade-out and re-emergence”, but this is a little misleading. The VAM never re-emerges on test scores. It only shows up in the earnings numbers.

All of this is really dubious, and it seems like Section III gives us an easy way out. There’s probably a component of year-to-year stable bias in VAM, such that it captures something about student quality, maybe even innate ability, rather than just teacher quality. It sounds very easy to just say that this is the component producing Chetty’s finding of income gains at age 28; students who have higher innate ability in fourth grade will probably still have it in their twenties.

Chetty is aware of this argument and tries to close it off. He conducts a quasi-experiment which he thinks replicates and confirms his original point: what happens when new teachers enter the school?

The thing we’re most worried about is bias in student selection to teachers. If we take an entire grade of a school (for example, if a certain school has three fifth-grade teachers, we take all three of them as a unit) this should be immune to such effects. So Chetty looks at entire grades as old teachers retire and new teachers enter. In particular, he looks at such grades when a new teacher transfers from a different school. That new transfer teacher already has a VAM which we know from his work at the other school, which will be either higher or lower than the average VAM of his new school. If it’s higher and VAM is real, we should expect the average VAM of that grade of his new school to go up a proportionate amount. If it’s lower and VAM is real, we should expect the average VAM of that grade of his new school to go down a proportionate amount. Chetty investigates this with all of the transfer teachers in his data, finds this is in fact what happens, and finds that if he estimates VAM from these transfers he gets the same number (+ $1000 in earnings) that he got from the normal data. This is impressive. Maybe even too impressive. Really? The same number? So there’s no bias in the normal data? I thought there was a lot of evidence that most of it was bias?

Rothstein is able to replicate Chetty’s findings using data from a different district, but then he goes on to do the same thing on Chetty’s quasi-experiment as he did on the normal VAMs, with the same results. That is, you can use the amount a school improves when a great new fifth-grade teacher transfers in to predict that teacher’s students’ fourth-grade performance. Not perfectly. But a little. For some reason, teacher transfers are having the same freaky time-bending effects as other VAM. Rothstein mostly explains this by saying that Chetty incorrectly excluded certain classes and teachers from his sample, although I don’t fully understand this argument. He also gives one other example of when this might happen: suppose that a neighborhood is gentrifying. The new teachers who transfer in after the original teachers retire will probably be a better class of professional lured in by the improving neighborhood. And the school’s student body will also probably be more genetically and socioeconomically advantaged. So better transfer teachers will be correlated with higher-achieving kids, but they won’t have caused such high achievement.

After this came an increasingly complicated exchange between Rothstein and Chetty that I wasn’t able to follow. Chetty, Friedman, and Rockoff wrote a 52 page Response To Rothstein where they argued that Rothstein’s methodology would find retro-causal effects even in a fair experiment where none should exist. According to a 538 article on the debate, a couple of smart people (albeit smart people who already support VAMs and might be biased) think that Chetty’s response makes sense, and even Rothstein agrees it “could be” true. 538 definitely thought the advantage in this exchange went to Chetty. But Rothstein responded with a re-replication of his results that he says addresses Chetty’s criticisms but still finds the retro-causal effects indicating bias; as far as I know Chetty has not responded and nobody has weighed in to give me an expert opinion on whether or not it’s right.

My temptation would usually be to say – here are some really weird results that can’t possibly be true which we want to explain away, here’s a widely-respected Berkeley professor of economics who says he’s explained them away, great, let’s forget about the whole thing. But there’s one more experiment which I can’t dismiss so easily.

V.

Project STAR (Student Teacher Achievement Ratio) was a big educational experiment in the 80s and 90s to see whether or not smaller class size improved student performance. That’s a whole different can of worms, but the point is that in order to do this experiment for a while they randomized children to kindergarten classes within schools across 79 different schools. Since one of the biggest possible sources of bias for these last few studies has been possible nonrandom assignment of students to teachers, these Tennessee schools were an opportunity to get much better data than were available anywhere else.

So Chetty, Friedman, Higer, Saez, Schanzenbach, and Yagan analyzed the STAR data. They tried to do a lot of things with predicting earnings based on teacher experience, teacher credentials, and other characteristics, and it’s a bit controversial whether they succeeded or not – see Bryan Caplan’s analysis (1, 2) for more. Caplan is skeptical of a lot of the study, but one part he didn’t address – and which I find most convincing – is based on something a lot like VAM.

Because of the random assignment, Chetty et al don’t have to do full VAM here. It looks like their measure of kindergarten teacher quality is just the average of all their students’ test scores (wait, kindergarteners are taking standardized tests now? I guess so.) When they’re using teacher quality to predict the success of specific students, they use the average of all the test scores except that of the student being predicted, in order to keep it fair.

They find that the average test score of all the other students in your class, compared against the average score of all the students in other randomly assigned classes in your school, predicts your own test score. “A one percentile increase in entry-year class quality is estimated to raise own test scores by 0.68 percentiles, confirming that test scores are highly correlated across students within a classroom”. This fades to approximately zero by fourth grade, confirming that the test-score-related benefits of having a good teacher are transient and decay quickly. But, students assigned to a one-percentile-higher class have average earnings that are 0.4% higher at age 25-27! And they say that this relationship is linear! So for example, the best kindergarten teacher in their dataset caused her class to perform at the 70th percentile on average, and these students earned about $17000 on average (remember, these are young entry-level workers in Tennessee) compared to the $15500 or so of their more average-kindergarten-teacher-having peers. Just their kindergarten teacher, totally apart from any other teacher in their life history, increased their average income 10%. Really, Chetty et al? Really?

But as crazy as it is, this study is hard to poke holes in. Even in arguing against it, Caplan notes that “it’s an extremely impressive piece” that “the authors are very careful”, and that it’s “one of the most impressive empirical papers ever written”. The experimental randomization means we can’t apply most of the usual anti-VAM measures to it. I don’t know, man. I just don’t know.

Okay, fine. I have one really long-shot possibility. Chetty et al derive their measure for teacher quality from the performance of all of the students in a class, excluding each student in turn as they try to predict his or her results. But this is only exogenous if the student doesn’t affect his or her peers’ test scores. But it’s possible some students do affect their peers’ test scores. If a student is a behavioral problem, they can screw up the whole rest of their class. Carrell finds that “exposure to a disruptive peer in classes of 25 during elementary school reduces earnings at age 26 by 3 to 4 percent”. Now, this in itself is a crazy, hard-to-believe study. But if we accept this second crazy hard-to-believe study, it might provide us with a way of attacking the first crazy hard-to-believe study. Suppose we have a really screwed-up student who is always misbehaving in class and disrupting the lesson. This lowers all his peers’ test scores and makes the teacher look low-quality. Then that kid grows up and remains screwed-up and misbehaving and doesn’t get as good a job. If this is a big factor in the differences in performances between classes, then so-called “teacher quality” might be conflated with a measure of how many children in their classes are behavioral problems, and apparent effects of teacher quality on earnings might just represent that misbehaving kids tend to become low-earning adults. I’m not sure if the magnitude of this effect checks out, but it might be a possibility.

But if we can’t make that work, we’re stuck believing that good kindergarten teachers can increase your yearly earnings by thousands of dollars. What do we make of that?

Again, everybody finds that test score gains do not last nearly that long. So it can’t be that kindergarten teachers provide you with a useful fund of knowledge which you build upon later. It can’t even be that kindergarten teachers stimulate and enrich you which raises your IQ or makes you love learning or anything like that. It has to be something orthogonal to test scores and measurable intellectual ability.

Chetty et al’s explanation is that teachers also teach “non-cognitive skills”. I can’t understand the regressions they use, but they say that although a one percentile increase in kindergarten class quality has a statistically insignificant increase (+ 0.05 percentiles) on 8th grade test scores, it has a statistically significant increase (+0.15 percentiles) on 8th grade non-cognitive scores (“non-cognitive scores” in this case are a survey where 8th grade teachers answer questions like “does this student annoy others?”) They then proceed to demonstrate that the persistence of these non-cognitive effects do a better job of predicting the earning gains than the test scores do. They try to break these non-cognitive effects into four categories: “effort”, “initiative”, “engagement” and “whether the student values school”, but the results are pretty boring and about equally loaded on all of them.

This does go together really well with my “behavioral problem” theory of the kindergarten class-earnings effect. The “quality” of a student’s kindergarten class, which might have more to do with the number of students who were behavioral problems in it than anything else, doesn’t correlate with future test scores but does correlate with future behavioral problems. It also seems to match Plomin’s point about how very early test scores are determined by environment, but later test scores are determined by genetics. A poor learning environment might be a really big deal in kindergarten, but stop mattering as much later on.

But this also goes together with some other studies that have found the same. The test scores gains from pre-K are notorious for vanishing after a couple of years, but a few really big preschool studies like the Perry Preschool Program found that such programs do not boost IQ but may have other effects (though to complicate matters, apparently Perry did boost later-life standardized test scores, just not IQ scores, and to further complicate matters, other studies find children who went to pre-K have worse behavior). This also sort of reminds me of some of the very preliminary research I’ve been linking to recently suggesting that excessively early school starting ages seem to produce an ADHD-like pattern of bad behavior and later-life bad effects, which I was vaguely willing to attribute to overchallenging kids’ brains too early while they’re still developing. If I wanted to be very mean (and I do!) I could even say that all kindergarten is a neurological insult that destroys later life prospects because of forcing students to overclock their young brains concentrating on boring things, but good teachers can make this less bad than it might otherwise be by making their classes a little more enjoyable.

But even if this is true, it loops back to the question I started with: there’s strong evidence that parents have relatively little non-genetic impact on their childrens’ life outcomes, but now we’re saying that even a kindergarten teacher they only see for a year does have such an impact? And what’s more, it’s not even in the kindergarten teacher’s unique area of comparative advantage (teaching academic subjects), but in the domain of behavioral problems, something that parents have like one zillion times more exposure to and power over?

I don’t know. I still find these studies unbelievable, but don’t have the sort of knock-down evidence to dismiss them that I’d like. I’m really impressed with everybody participating in this debate, with the quality of the data, and with the ability to avoid a lot of the usual failure modes. It’s just not enough to convince me of anything yet.

VI.

In summary: teacher quality probably explains 10% of the variation in same-year test scores. A +1 SD better teacher might cause a +0.1 SD year-on-year improvement in test scores. This decays quickly with time and is probably disappears entirely after four or five years, though there may also be small lingering effects. It’s hard to rule out the possibility that other factors, like endogenous sorting of students, or students’ genetic potential, contributes to this as an artifact, and most people agree that these sorts of scores combine some signal with a lot of noise. For some reason, even though teachers’ effects on test scores decay very quickly, studies have shown that they have significant impact on earning as much as 20 or 25 years later, so much so that kindergarten teacher quality can predict thousands of dollars of difference in adult income. This seemingly unbelievable finding has been replicated in quasi-experiments and even in real experiments and is difficult to banish. Since it does not happen through standardized test scores, the most likely explanation is that it involves non-cognitive factors like behavior. I really don’t know whether to believe this and right now I say 50-50 odds that this is a real effect or not – mostly based on low priors rather than on any weakness of the studies themselves. I don’t understand this field very well and place low confidence in anything I have to say about it.

Further reading: Institute of Education Science summary, Edward Haertel’s summary, TTI report, Adler’s critique of Chetty, American Statistical Society’s critique of Chetty/VAM, Chetty’s response, Ballou’s critique of Chetty

# Three Great Articles On Poverty, And Why I Disagree With All Of Them

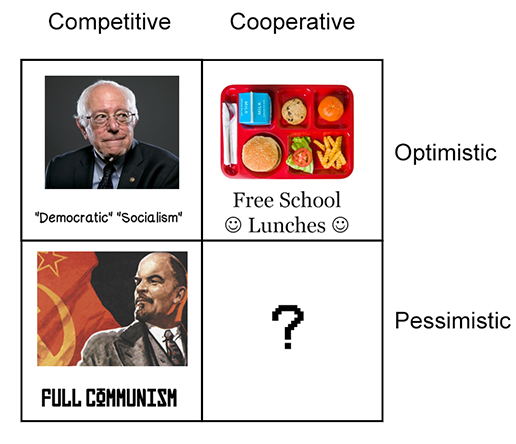
QZ: The universal basic income is an idea whose time will never come. Okay, maybe this one isn’t so great. It argues that work is ennobling (or whatever), that robots probably aren’t stealing our jobs, that even if we’re going through a period of economic disruption we’ll probably adapt, and that “if the goal is eliminating poverty, it is better to direct public funds to [failing schools and substandard public services]” then to try a guaranteed income scheme. It ends by saying that “I can’t understand why we’d consider creating and then calcifying a perpetually under-employed underclass by promoting the stagnation of their skills and severing their links to broader communities.”

(imagine a world where we had created and calcified a perpetually under-employed stagnant underclass. It sounds awful.)

More Crows Than Eagles: Unnecessariat. This one is great. A blogger from the Rust Belt reports on the increasing economic despair and frustration all around her, in the context of the recent spikes in heroin overdoses and suicides. There’s an important caveat here, in that at least national-level economic data paint a rosy picture: the unemployment rate is very low, consumer confidence is high, and the studies of technological unemployment suggest it’s not happening yet. Still, a lot of people on the ground – the anonymous blogger, the pathologists she worked with, and me from my position as a psychiatrist in the Midwest – feel like there’s a lot more misery and despair than the statistics suggest. MCTE replaces the old idea of the “precariat” – people who just barely have jobs and are worried about losing them – with her own coinage “unnecessariat” – people who don’t have jobs, are useless to the economy, and nobody cares what happens to them. It reminds me of the old argument of sweatshop-supporting economists – sure, we’re exploiting you, but you’d miss us if we left. She hates Silicon Valley for building its glittering megaplexes while ignoring everyone else, but she hates even more the people saying “Learn to code! Become part of the bright new exciting knowledge economy!” because realistically there’s no way an opioid-depended 55-year-old ex-trucker from Kentucky is going to learn to code. The only thing such people have left is a howl of impotent rage, and it has a silly hairstyle and is named Donald J. Trump.

Freddie deBoer: Our Nightmare. Also pretty great. The same things deBoer has been warning about for years, but expressed unusually clearly. By taking on the superficial mantle of center-leftism, elites sublimate the revolutionary impulse into a competition for social virtue points which ends up reinforcing and legitimizing existing power structures. Constant tally-keeping over what percent of obscenely rich exploitative Wall Street executives are people of color replaces the question of whether there should be obscenely rich exploitative Wall Street executives at all. As such tendencies completely capture the Democratic Party and the country’s mainstream left, genuine economic anger becomes more likely to be funneled into the right wing, where the elites can dismiss it as probably-racist (often with justification) and ignore it. “I cannot stress enough to you how vulnerable the case for economic justice is in this country right now. Elites agitate against it constantly…this is a movement, coordinated from above, and its intent is to solidify the already-vast control of economic elites over our political system…[Liberalism] is an attempt to ameliorate the inequality and immiseration of capitalism, when inequality and immiseration are the very purpose of capitalism.”

These articles all look at poverty in different ways, and I think that I look at poverty in a different way still. In the spirit of all the crazy political compasses out there, maybe we can learn something by categorizing them:



Including only people who think society should be in the business of collectively helping the poor at all (ie no extreme libertarians or social Darwinists) and people who are interested in something beyond deBoer’s nightmare scenario (ie not just making sure every identity group has an equal shot at the Wall Street positions).

People seem to split into a competitive versus a cooperative view of poverty. To massively oversimplify: competitives agree with deBoer that “inequality and immiseration are the very purpose of capitalism” and conceive of ending poverty in terms of stopping exploitation and giving the poor their “just due” that the rich have taken away from them. The cooperatives argue that everyone is working together to create a nice economy that enriches everybody who participates in it, but some people haven’t figured out exactly how to plug into the magic wealth-generating machine, and we should give them a helping hand (“here’s government-subsidized tuition to a school where you can learn to code!”). Probably nobody’s 100% competitive or 100% cooperative, but I think a lot of people have a tendency to view the problem more one way than the other.

So the northwest corner of the grid is people who think the problem is primarily one of exploitation, but it’s at least somewhat tractable to reform. No surprises here – these are the types who think that the big corporations are exploiting people, but if average citizens try hard enough they can make the Man pay a $15 minimum wage and give them free college tuition, and then with enough small victories like these they can level the balance enough to give everybody a chance.

(These are all going to be straw men, but hopefully useful straw men)

The southwest corner is people who think the problem is primarily one of exploitation, but nothing within the system will possibly help. I put “full communism” in the little box, but I guess this could also be anarcho-syndicalism, or anarcho-capitalism, or theocracy, or Trumpism, or [insert your preferred poorly-planned form of government which inevitably fails here].

The northeast corner is people who think we’re all in this together and there are lots of opportunities to help. This is the QZ writer who said we should be focusing on “education and public services”. The economy is a benevolent force that wants to help everybody, but some people through bad luck – poor educational opportunities, not enough childcare, racial prejudice – haven’t gotten the opportunity they need yet, so we should lend them a helping hand so they can get back on their feet and one day learn to code. I named this quadrant “Free School Lunches” after all those studies that show that giving poor kids free school lunches improves their grades by X percent, which changes their chances of getting into a good college by Y percent, which increases their future income by Z percent, so all we have to do is have lots of social programs like free school lunches and then poverty is solved. But aside from the lunch people people, this category must also include libertarians who think that all we need to do is remove regulations that prevent the poor from succeeding, Reaganites who think that a rising tide will lift all boats, and conservatives who think the poor just need to be taught Traditional Hard-Working Values. Actually, probably 90% of the Overton Window is in this corner.

The southeast corner is people who think that we’re all in this together, but that helping the poor is really hard. They agree with the free school lunch crowd that capitalism is more the solution than the problem, and that we should think of this in terms of complicated impersonal social and educational factors preventing poor people from fitting into the economy. But the southeasterners worry school lunches won’t be enough. Maybe even hiring great teachers, giving everybody free health care, ending racism, and giving generous vocational training to people in need wouldn’t be enough. If we held a communist revolution, it wouldn’t do a thing: you can’t hold a revolution against skill mismatch. This is a very gloomy quadrant, and I don’t blame people for not wanting to be in it. But it’s where I spend most of my time.

The exploitation narrative seems fundamentally wrong to me – I’m not saying exploitation doesn’t happen, nor even that it isn’t common, just that isn’t not the major factor causing poverty and social decay. The unnecessariat article, for all its rage against Silicon Valley hogging the wealth, half-admits this – the people profiled have become unnecessary to the functioning of the economy, no longer having a function even as exploited proletarians. Silicon Valley isn’t exploiting these people, just ignoring them. Fears of technological unemployment are also relevant here: they’re just the doomsday scenario where all of us are relegated to the unnecessariat, the economy having passed us by.

But I also can’t be optimistic about programs to end poverty. Whether it’s finding out that schools and teachers have relatively little effect on student achievement, that good parenting has even less, or that differences in income are up to fifty-eight percent heritable and a lot of what isn’t outright genetic is weird biology or noise, most of the research I read is very doubtful of easy (or even hard) solutions. Even the most extensive early interventions have underwhelming effects. We can spend the collective energy of our society beating our head against a problem for decades and make no headway. While there may still be low-hanging fruit – maybe an scaled-up Perry Preschool Project, lots of prenatal vitamins, or some scientist discovering a new version of the unleaded-gasoline movement – we don’t seem very good at finding it, and I worry it would be at most a drop in the bucket. Right now I think that a lot of variation in class and income is due to genetics and really deep cultural factors that nobody knows how to change en masse.

I can’t even really believe that a rising tide will lift all boats anymore. Not only has GDP uncoupled from median wages over the past forty years, but there seems to be a Red Queen’s Race where every time the GDP goes up the cost of living goes up the same amount. US real GDP has dectupled since 1900, yet a lot of people have no savings and are one paycheck away from the street. In theory, a 1900s poor person who suddenly got 10x his normal salary should be able to save 90% of it, build up a fund for rainy days, and end up in a much better position. In practice, even if the minimum wage in 2100 is $200 2016 dollar an hour, I expect the average 2100 poor person will be one paycheck away from the street. I can’t explain this, I just accept it at this point. And I think that aside from our superior technology, I would rather be a poor farmer in 1900 than a poor kid in the projects today. More southeast corner gloom.

The only public figure I can think of in the southeast quadrant with me is Charles Murray. Neither he nor I would dare reduce all class differences to heredity, and he in particular has some very sophisticated theories about class and culture. But he shares my skepticism that the 55 year old Kentucky trucker can be taught to code, and I don’t think he’s too sanguine about the trucker’s kids either. His solution is a basic income guarantee, and I guess that’s mine too. Not because I have great answers to all of the QZ article’s problems. But just because I don’t have any better ideas1,2.

The QZ article warns that it might create a calcified “perpetually under-employed stagnant underclass”. But of course we already have such an underclass, and it’s terrible. I can neither imagine them all learning to code, nor a sudden revival of the non-coding jobs they used to enjoy. Throwing money at them is a pretty subpar solution, but it’s better than leaving everything the way it is and not throwing money at them.

This is why I can’t entirely sympathize with any of the essays I read on poverty, eloquent though they are.

Footnotes

1. And then there’s the rest of the world. Given the success of export capitalism in Korea, Taiwan, China, Vietnam, et cetera, and the pattern where multinationals move to some undeveloped country with cheap labor, boost the local economy until the country is developed and labor there isn’t so cheap anymore, and then move on to the next beneficiary – solving international poverty seems a lot easier than solving local poverty. All we have to do is keep wanting shoes and plastic toys. And part of me wonders – if setting up a social safety net would slow domestic economic growth – or even divert money that would otherwise go to foreign aid – does that make it a net negative? Maybe we should be optimizing for maximum economic growth until we’ve maxed out the good we can do by industrializing Third World countries? My guess is that enough of the basic income debate is about how to use existing welfare payments that this wouldn’t be too big a factor. And I would hope (for complicated reasons), that basic income would be more likely to help than hurt the economy3.

2. Obviously invent genetic engineering and create a post-scarcity society, but until then we have to deal with this stuff.

3. And then there’s the whole open borders idea, which probably isn’t very compatible with basic income at all. Right now I think – I’ll explain at more length later – fully open borders is a bad idea, because the risk of it destabilizing the country and ruining the economic motor that lifts Third World countries out of poverty is too high.

# Book Review: Age of Em

[Note: I really liked this book and if I criticize it that’s not meant as an attack but just as what I do with interesting ideas. Note that Robin has offered to debate me about some of this and I’ve said no – mostly because I hate real-time debates and have bad computer hardware – but you may still want to take this into account when considering our relative positions. Mild content warning for murder, rape, and existential horror. Errors in Part III are probably my own, not the book’s.]

I.

There are some people who are destined to become adjectives. Pick up a David Hume book you’ve never read before and it’s easy to recognize the ideas and style as Humean. Everything Tolkien wrote is Tolkienesque in a non-tautological sense. This isn’t meant to denounce either writer as boring. Quite the opposite. They produced a range of brilliant and diverse ideas. But there was a hard-to-define and very consistent ethos at the foundation of both. Both authors were very much like themselves.

Robin Hanson is more like himself than anybody else I know. He’s obviously brilliant – a PhD in economics, a masters in physics, work for DARPA, Lockheed, NASA, George Mason, and the Future of Humanity Institute. But his greatest aptitude is in being really, really Hansonian. Bryan Caplan describes it as well as anybody:

When the typical economist tells me about his latest research, my standard reaction is ‘Eh, maybe.’ Then I forget about it. When Robin Hanson tells me about his latest research, my standard reaction is ‘No way! Impossible!’ Then I think about it for years.

This is my experience too. I think I said my first “No way! Impossible!” sometime around 2008 after reading his blog Overcoming Bias. Since then he’s influenced my thinking more than almost anyone else I’ve ever read. When I heard he was writing a book, I was – well, I couldn’t even imagine a book by Robin Hanson. When you read a thousand word blog post by Robin Hanson, you have to sit down and think about it and wait for it to digest and try not to lose too much sleep worrying about it. A whole book would be something.

I have now read Age Of Em (website)and it is indeed something. Even the cover gives you a weird sense of sublimity mixed with unease:



And in this case, judging a book by its cover is entirely appropriate.

II.

Age of Em is a work of futurism – an attempt to predict what life will be like a few generations down the road. This is not a common genre – I can’t think of another book of this depth and quality in the same niche. Predicting the future is notoriously hard, and that seems to have so far discouraged potential authors and readers alike.

Hanson is not discouraged. He writes that:

Some say that there is little point in trying to foresee the non-immediate future. But in fact there have been many successful forecasts of this sort. For example, we can reliably predict the future cost changes for devices such as batteries or solar cells, as such costs tend to follow a power law of the cumulative device production (Nagy et al 2013). As another example, recently a set of a thousand published technology forecasts were collected and scored for accuracy, by comparing the forecasted date of a technology milestone with its actual date. Forecasts were significantly more accurate than random, even forecasts 10 to 25 years ahead. This was true separately for forecasts made via many different methods. On average, these milestones tended to be passed a few years before their forecasted date, and sometimes forecasters were unaware that they had already passed (Charbonneau et al, 2013).

A particularly accurate book in predicting the future was The Year 2000, a 1967 book by Herman Kahn and Anthony Wiener. It accurately predicted population, was 80% correct for computer and communication technology, and 50% correct for other technology (Albright 2002). On even longer time scales, in 1900 the engineer John Watkins did a good job of forecasting many basic features of society a century later (Watkins 1900) […]

Some say no one could have anticipated the recent big changes associated with the arrival and consequences of the World Wide Web. Yet participants in the Xanadu hypertext project in which I was involved from 1984 to 1993 correctly anticipated many key aspects of the Web […] Such examples show that one can use basic theory to anticipate key elements of distant future environments, both physical and social, but also that forecasters do not tend to be much rewarded for such efforts, either culturally or materially. This helps to explain why there are relatively few serious forecasting efforst. But make no mistake, it is possible to forecast the future.

I think Hanson is overstating his case. All except Watkins were predicting only 10 – 30 years in the future, and most of their predictions were simple numerical estimates, eg “the population will be one billion” rather than complex pictures of society. The only project here even remotely comparable in scope to Hanson’s is John Watkins’ 1900 article.

Watkins is classically given some credit for broadly correct ideas like “Cameras that can send pictures across the world instantly” and “telephones that can call anywhere in the world”, but of his 28 predictions, I judge only eight as even somewhat correct. For example, I grant him a prediction that “the average American will be two inches taller because of good medical care” even though he then goes on to say in the same sentence that the average life expectancy will be fifty and suburbanization will be so total that building city blocks will be illegal (sorry, John, only in San Francisco). Most of the predictions seem simply and completely false. Watkins believes all animals and insects will have been eradicated. He believes there will be “peas as large as beets” and “strawberries as large as apples” (these are two separate predictions; he is weirdly obsessed with fruit and vegetable size). We will travel to England via giant combination submarine/hovercrafts that will complete the trip in a lightning-fast two days. There will be no surface-level transportation in cities as all cars and walkways have moved underground. The letters C, X, and Q will be removed from the language. Pneumatic tubes will deliver purchases from stores. “A man or woman unable to walk ten miles at a stretch will be regarded as a weakling.”

Where Watkins is right, he is generally listing a cool technology slightly beyond what was available to his time and predicting we will have it. Nevertheless, he is still mostly wrong. Yet this is Hanson’s example of accurate futurology. And he is right to make it his example of accurate futurology, because everything else is even worse.

Hanson has no illusions of certainty. He starts by saying that “conditional on my key assumptions, I expect at least 30% of future situations to be usefully informed by my analysis. Unconditionally, I expect at least 10%.” So he is not explicitly overconfident. But in an implicit sense, it’s just weird to see the level of detail he tries to predict – for example, he has two pages about what sort of swear words the far future might use. And the book’s style serves to reinforce its weirdness. The whole thing is written in a sort of professorial monotone that changes little from loving descriptions of the sorts of pipes that will cool future buildings (one of Hanson’s pet topics) to speculation on our descendents’ romantic relationships (key quote: “The per minute subjective value of an equal relation should not fall much below half of the per-minute value of a relation with the best available open source lover”). And it leans heavily on a favorite Hansonian literary device – the weirdly general statement about something that sounds like it can’t possibly be measurable, followed by a curt reference which if followed up absolutely confirms said statement, followed by relentlessly ringing every corollary of it:

Today, mental fatigue reduces mental performance by about 0.1% per minute. As by resting we can recover at a rate of 1% per minute, we need roughly one-tenth of our workday to be break time, with the duration between breaks being not much more than an hour or two (Trougakos and Hideg 2009; Alvanchi et al 2012)…Thus many em tasks will be designed to take about an hour, and many spurs are likely to last for about this duration.

Or:

Today, painters, novelists, and directors who are experimental artists tend to do their best work at roughly ages 46-52, 38-50, and 45-63 respectively, but those ages are 24-34, 29-40, and 27-43, respectively for conceptual artists (Galenson 2006)…At any one time, the vast majority of actual working ems [should be] near a peak productivity subjective age.

Or:

Wars today, like cities, are distributed evenly across all possible war sizes (Cederman 2003).

At some point I started to wonder whether Hanson was putting me on. Everything is just played too straight. Hanson even addresses this:

To resist the temptation to construe the future too abstractly, I’ll try to imagine a future full of complex detail. One indiciation that I’ve been successful in all these efforts will be if my scenario description sounds less like it came from a typical comic book or science fiction movie, and more like it came form a typical history text or business casebook.

Well, count that project a success. The effect is strange to behold, and I’m not sure it will usher in a new era of futurology. But Age of Em is great not just as futurology, but as a bunch of different ideas and purposes all bound up in a futurological package. For example:

– An introduction to some of the concepts that recur again and again across Robin’s thought – for example, near vs. far mode, the farmer/forager dichotomy, the inside and outside views, signaling. Most of us learned these through years reading Hanson’s blog Overcoming Bias, getting each chunk in turn, spending days or months thinking over each piece. Getting it all out of a book you can read in a couple of days sounds really hard – but by applying them to dozens of different subproblems involved in future predictions, Hanson makes the reader more comfortable with them, and I expect a lot of people will come out of the book with an intuitive understanding of how they can be applied.

– A whirlwind tour through almost every science and a pretty good way to learn about the present. If you didn’t already know that wars are distributed evenly across all possible war sizes, well, read Age of Em and you will know that and many similar things besides.

– A manifesto. Hanson often makes predictions by assuming that since the future will be more competitive, future people are likely to converge toward optimal institutions. This is a dangerous assumption for futurology – it’s the same line of thinking that led Watkins to assume English would abandon C, X, and Q as inefficient – but it’s a great assumption if you want a chance to explain your ideas of optimal institutions to thousands of people who think they’re reading fun science-fiction. Thus, Robin spends several pages talking about how ems may use prediction markets – an information aggregation technique he invented – to make their decisions. In the real world, Hanson has been trying to push these for decades, with varying levels of success. Here, in the guise of a future society, he can expose a whole new group of people to their advantages – as well as the advantages of something called “combinatorial auctions” which I am still not smart enough to understand.

– A mind-expanding drug. One of the great risks of futurology is to fail to realize how different societies and institutions can be – the same way uncreative costume designers make their aliens look like humans with green skin. A lot of our thoughts about the future involve assumptions we’ve never really examined critically, and Hanson dynamites those assumptions. For page after page, he gives strong arguments why our descendants might be poorer, shorter-lived, less likely to travel long distances or into space, less progressive and open-minded. He predicts little noticeable technological change, millimeter-high beings living in cities the size of bottles, careers lasting fractions of seconds, humans being incomprehensibly wealthy patrons to their own robot overlords. And all of it makes sense.

When I read Stross’ Accelerando, one of the parts that stuck with me the longest were the Vile Offspring, weird posthuman entities that operated a mostly-incomprehensible Economy 2.0 that humans just sort of hung out on the edges of, goggle-eyed. It was a weird vision – but, for Stross, mostly a black box. Age of Em opens the box and shows you every part of what our weird incomprehensible posthuman descendents will be doing in loving detail. Even what kind of swear words they’ll use.

III.

So, what is the Age of Em?

According to Hanson, AI is really hard and won’t be invented in time to shape the posthuman future. But sometime a century or so from now, scanning technology, neuroscience, and computer hardware will advance enough to allow emulated humans, or “ems”. Take somebody’s brain, scan it on a microscopic level, and use this information to simulate it neuron-by-neuron on a computer. A good enough simulation will map inputs to outputs in exactly the same way as the brain itself, effectively uploading the person to a computer. Uploaded humans will be much the same as biological humans. Given suitable sense-organs, effectuators, virtual avatars, or even robot bodies, they can think, talk, work, play, love, and build in much the same way as their “parent”. But ems have three very important differences from biological humans.

First, they have no natural body. They will never need food or water; they will never get sick or die. They can live entirely in virtual worlds in which any luxuries they want – luxurious penthouses, gluttonous feasts, Ferraris – can be conjured out of nothing. They will have some limited ability to transcend space, talking to other ems’ virtual presences in much the same way two people in different countries can talk on the Internet.

Second, they can run at different speeds. While a normal human brain is stuck running at the speed that physics allow, a computer simulating a brain can simulate it faster or slower depending on preference and hardware availability. With enough parallel hardware, an em could experience a subjective century in an objective week. Alternatively, if an em wanted to save hardware it could process all its mental operations v e r y s l o w l y and experience only a subjective week every objective century.

Third, just like other computer data, ems can be copied, cut, and pasted. One uploaded copy of Robin Hanson, plus enough free hardware, can become a thousand uploaded copies of Robin Hanson, each living in their own virtual world and doing different things. The copies could even converse with each other, check each other’s work, duel to the death, or – yes – have sex with each other. And if having a thousand Robin Hansons proves too much, a quick ctrl-x and you can delete any redundant ems to free up hard disk space for Civilization 6 (coming out this October!)

Would this count as murder? Hanson predicts that ems will have unusually blase attitudes toward copy-deletion. If there are a thousand other copies of me in the world, then going to sleep and not waking up just feels like delegating back to a different version of me. If you’re still not convinced, Hanson’s essay Is Forgotten Party Death? is a typically disquieting analysis of this proposition. But whether it’s true or not is almost irrelevant – at least some ems will think this way, and they will be the ones who tend to volunteer to be copied for short term tasks that require termination of the copy afterwards. If you personally aren’t interested in participating, the economy will leave you behind.

The ability to copy ems as many times as needed fundamentally changes the economy and the idea of economic growth. Imagine Google has a thousand positions for Ruby programmers. Instead of finding a thousand workers, they can find one very smart and very hard-working person and copy her a thousand times. With unlimited available labor supply, wages plummet to subsistence levels. “Subsistence levels” for ems are the bare minimum it takes to rent enough hardware from Amazon Cloud to run an em. The overwhelming majority of ems will exist at such subsistence levels. On the one hand, if you’ve got to exist on a subsistence level, a virtual world where all luxuries can be conjured from thin air is a pretty good place to do it. On the other, such starvation wages might leave ems with little or no leisure time.

Sort of. This gets weird. There’s an urban legend about a “test for psychopaths”. You tell someone a story about a man who attends his mother’s funeral. He met a really pretty girl there and fell in love, but neglected to get her contact details before she disappeared. How might he meet her again? If they answer “kill his father, she’ll probably come to that funeral too”, they’re a psychopath – ordinary people would have a mental block that prevents them from even considering such a drastic solution. And I bring this up because after reading Age of Em I feel like Robin Hanson would be able to come up with some super-solution even the psychopaths can’t think of, some plan that gets the man a threesome with the girl and her even hotter twin sister at the cost of wiping out an entire continent. Everything about labor relations in Age of Em is like this.

For example, suppose you want to hire an em at subsistence wages, but you want them 24 hours a day, 7 days a week. Ems probably need to sleep – that’s hard-coded into the brain, and the brain is being simulated at enough fidelity to leave that in. But jobs with tasks that don’t last longer than a single day – for example, a surgeon who performs five surgeries a day but has no day-to-day carryover – can get around this restriction by letting an em have one full night of sleep, then copying it. Paste the em at the beginning of the workday. When it starts to get tired, let it finish the surgery it’s working on, then delete it and paste the well-rested copy again to do the next surgery. Repeat forever and the em never has to get any more sleep than that one night. You can use the same trick to give an em a “vacation” – just give it one of them, then copy-paste that brain-state forever.

Or suppose your ems want frequent vacations, but you want them working every day. Let a “trunk” em vacation every day, then make a thousand copies every morning, work all the copies for twenty-four hours, then delete them. Every copy remembers a life spent in constant vacation, and cheered on by its generally wonderful existence it will give a full day’s work. But from the company’s perspective, 99.9% of the ems in its employment are working at any given moment.

(another option: work the em at normal subjective speed, then speed it up a thousand times to take its week-long vacation, then have it return to work after only one-one-thousandth of a week has passed in real life)

Given that ems exist at subsistence wages, saving enough for retirement sounds difficult, but this too has weird psychopathic solutions. Thousands of copies of the same em can pool their retirement savings, then have all except a randomly chosen one disappear at the moment of retirement, leaving that one with an nest egg thousands of time what it could have accumulated by its own efforts. Or an em can invest its paltry savings in some kind of low-risk low-return investment and reduce its running speed so much that the return on its investment is enough to pay for its decreased subsistence. For example, if it costs $100 to rent enough computing power to run an em at normal speed for one year, and you only have $10 in savings, you can rent 1/1000th of the computer for $0.10, run at 1/1000th speed, invest your $10 in a bond that pays 1% per year, and have enough to continue running indefinitely. The only disadvantage is that you’ll only experience a subjective week every twenty objective years. Also, since other entities are experiencing a subjective week every second, and some of those entities have nukes, probably there will be some kind of big war, someone will nuke Amazon’s data centers, and you’ll die after a couple of your subjective minutes. But at least you got to retire!

If ems do find ways to get time off the clock, what will they do with it? Probably they’ll have really weird social lives. After all, the existence of em copies is mostly funded by companies, and there’s no reason for companies to copy-paste any but the best workers in a given field. So despite the literally trillions of ems likely to make up the world, most will be copies of a few exceptionally brilliant and hard-working individuals with specific marketable talents. Elon Musk might go out one day to the bar with his friend, who is also Elon Musk, and order “the usual”. The bartender, who is Elon Musk himself, would know exactly what drink he wants and have it readily available, as the bar caters entirely to people who are Elon Musk. A few minutes later, a few Chesley Sullenbergers might come in after a long day of piloting airplanes. Each Sullenberger would have met hundreds of Musks before and have a good idea about which Musk-Sullenberger conversation topics were most enjoyable, but they might have to adjust for circumstances; maybe the Musks they met before all branched off a most recent common ancestor in 2120, but these are a different branch who were created in 2105 and remember Elon’s human experiences but not a lot of the posthuman lives that shaped the 2120 Musks’ worldviews. One Sullenberger might tentatively complain that the solar power grid has too many outages these days; a Musk might agree to take the problem up with the Council of Musks, which is totally a thing that exist (Hanson calls these sorts of groups “copy clans” and says they are “a natural candidate unit for finance, reproduction, legal, liability, and political representation”).

Romance could be even weirder. Elon Musk #2633590 goes into a bar and meets Taylor Swift #105051, who has a job singing in a nice local nightclub and so is considered prestigious for a Taylor Swift. He looks up a record of what happens when Elon Musks ask Taylor Swifts out and finds they are receptive on 87.35% of occasions. The two start dating and are advised by the Council of Musks and the Council of Swifts on the issues that are known to come up in Musk-Swift relationships and the best solutions that have been found to each. Unfortunately, Musk #2633590 is transferred to a job that requires operating at 10,000x human speed, but Swift #105051’s nightclub runs at 100x speed and refuses to subsidize her to run any faster; such a speed difference makes normal interaction impossible. The story has a happy ending; Swift #105051 allows Musk #2633590 to have her source code, and whenever he is feeling lonely he spends a little extra money to instantiate a high-speed copy of her to hang out with.

(needless to say, these examples are not exactly word-for-word taken from the book, but they’re heavily based off of Hanson’s more abstract descriptions)

The em world is not just very weird, it’s also very very big. Hanson notes that labor is a limiting factor in economic growth, yet even today the economy doubles about once every fifteen years. Once you can produce skilled labor through a simple copy-paste operation, especially labor you can run at a thousand times human speed, the economy will go through the roof. He writes that:

To generate an empirical estimate of em economy doubling times, we can look at the timescales it takes for machine shopes and factories today to make a mass of machines of a quality, quantity, variety, and value similar to that of machines that they themselves contain. Today that timescale is roughly 1 to 3 months. Also, designs were sketched two to three decades ago for systems that might self-repliate nearly completeld in 6 to 12 months…these estimates suggest that today’s manufacturing technologiy is capable of self-repliating on a scale of a few weeks to a few months.

Hanson thinks that with further innovation, such times can be reduced so far that “the economy might double every objective year, month, week, or day.” As the economy doubles the labor force – ie the number of ems – may double with it, until only a few years after the first ems the population numbers in the trillions. But if the em population is doubling every day, there had better be some pretty amazing construction efforts going on. The only thing that could possibly work on that scale is prefabricated modular construction of giant superdense cities, probably made mostly out of some sort of proto early-stage computronium (plus cooling pipes). Ems would be reluctant to travel from one such city to another – if they exist at a thousand times human speed, a trip on a hypersonic airliner that could go from New York to Los Angeles in an hour would still take forty subjective days. Who wants to be on an airplane for forty days?

(long-distance trade is also rare, since if the economy doubles fast enough it means that by the time goods reach their destination they could be almost worthless)

The real winners of this ultra-fast-growing economy? Ordinary humans. While humans will be way too slow and stupid to do anything useful, they will tend to have non-subsistence amounts of money saved up from their previous human lives, and also be running at speeds thousands of times slower than most of the economy. When the economy doubles every day, so can your bank account. Ordinary humans will become rarer, less relevant, but fantastically rich – a sort of doddering Neanderthal aristocracy spending sums on a cheeseburger that could support thousands of ems in luxury for entire lifetimes. While there will no doubt be pressure to liquidate humans and take their stuff, Hanson hopes that the spirit of rule of law – the same spirit that protects rich minority groups today – will win out, with rich ems reluctant to support property confiscation lest it extend to them also. Also, em retirees will have incentives a lot like humans – they have saved up money and go really slow – and like AARP memembers today they may be able to obtain disproportionate political power which will then protect the interests of slow rich people.

But we might not have much time to enjoy our sudden rise in wealth. Hanson predicts that the Age of Em will last for subjective em millennia – ie about one to two actual human years. After all, most of the interesting political and economic activity is going on at em timescales. In the space of a few subjective millennia, either someone will screw up and cause the apocalypse, somebody will invent real superintelligent AI that causes a technological singularity, or some other weird thing will happen taking civilization beyond the point that even Robin dares to try to predict.

IV.

Hanson understands that people might not like the idea of a future full of people working very long hours at subsistence wages forever (Zack Davis’ Contract-Drafting Em song is, as usual, relevant). But Hanson himself does not view this future as dystopian. Despite our descendents’ by-the-numbers poverty, they will avoid the miseries commonly associated with poverty today. There will be no dirt or cockroaches in their sparkling virtual worlds, nobody will go hungry, petty crime will be all-but-eliminated, and unemployment will be low. Anybody who can score some leisure time will have a dizzying variety of hyperadvanced entertainment available, and as for the people who can’t, they’ll mostly have been copied from people who really like working hard and don’t miss it anyway. As unhappy as we moderns may be contemplating em society, ems themselves will not be unhappy! And as for us:

The analysis in this book suggests that lives in the next great era may be as different from our lives as our lives are from farmers’ lives, or farmers’ lives are from foragers’ lives. Many readers of this book, living industrial era lives and sharing industrial era values, may be disturbed to see a forecast of em era descendants with choices and lifestyles that appear to reject many of the values that they hold dear. Such readers may be tempted to fight to prevent the em future, perhaps preferring a continuation of the industrial era. Such readers may be correct that rejecting the em future holds them true to their core values. But I advise such readers to first try hard to see this new era in some detail from the point of view of its typical residents. See what they enjoy and what fills them with pride, and listen to their criticisms of your era and values.

A short digression: there’s a certain strain of thought I find infuriating, which is “My traditionalist ancestors would have disapproved of the changes typical of my era, like racial equality, more open sexuality, and secularism. But I am smarter than them, and so totally okay with how the future will likely have values even more progressive and shocking than my own. Therefore I pre-approve of any value changes that might happen in the future as definitely good and better than our stupid hidebound present.”

I once read a science-fiction story that depicted a pretty average sci-fi future – mighty starships, weird aliens, confederations of planets, post-scarcity economy – with the sole unusual feature that rape was considered totally legal, and opposition to such as bigoted and ignorant as opposition to homosexuality is today. Everybody got really angry at the author and said it was offensive for him to even speculate about that. Well, that’s the method by which our cheerful acceptance of any possible future values is maintained: restricting the set of “any possible future values” to “values slightly more progressive than ours” and then angrily shouting down anyone who discusses future values that actually sound bad. But of course the whole question of how worried to be about future value drift only makes sense in the context of future values that genuinely violate our current values. Approving of all future values except ones that would be offensive to even speculate about is the same faux-open-mindedness as tolerating anything except the outgroup.

Hanson deserves credit for positing a future whose values are likely to upset even the sort of people who say they don’t get upset over future value drift. I’m not sure whether or not he deserves credit for not being upset by it. Yes, it’s got low-crime, ample food for everybody, and full employment. But so does Brave New World. The whole point of dystopian fiction is pointing out that we have complicated values beyond material security. Hanson is absolutely right that our traditionalist ancestors would view our own era with as much horror as some of us would view an em era. He’s even right that on utilitarian grounds, it’s hard to argue with an em era where everyone is really happy working eighteen hours a day for their entire lives because we selected for people who feel that way. But at some point, can we make the Lovecraftian argument of “I know my values are provincial and arbitrary, but they’re my provincial arbitrary values and I will make any sacrifice of blood or tears necessary to defend them, even unto the gates of Hell?”

This brings us to an even worse scenario.

There are a lot of similarities between Hanson’s futurology and (my possibly erroneous interpretation of) the futurology of Nick Land. I see Land as saying, like Hanson, that the future will be one of quickly accelerating economic activity that comes to dominate a bigger and bigger portion of our descendents’ lives. But whereas Hanson’s framing focuses on the participants in such economic activity, playing up their resemblances with modern humans, Land takes a bigger picture. He talks about the economy itself acquiring a sort of self-awareness or agency, so that the destiny of civilization is consumed by the imperative of economic growth.

Imagine a company that manufactures batteries for electric cars. The inventor of the batteries might be a scientist who really believes in the power of technology to improve the human race. The workers who help build the batteries might just be trying to earn money to support their families. The CEO might be running the business because he wants to buy a really big yacht. And the whole thing is there to eventually, somewhere down the line, let a suburban mom buy a car to take her kid to soccer practice. Like most companies the battery-making company is primarily a profit-making operation, but the profit-making-ness draws on a lot of not-purely-economic actors and their not-purely-economic subgoals.

Now imagine the company fires all its employees and replaces them with robots. It fires the inventor and replaces him with a genetic algorithm that optimizes battery design. It fires the CEO and replaces him with a superintelligent business-running algorithm. All of these are good decisions, from a profitability perspective. We can absolutely imagine a profit-driven shareholder-value-maximizing company doing all these things. But it reduces the company’s non-masturbatory participation in an economy that points outside itself, limits it to just a tenuous connection with soccer moms and maybe some shareholders who want yachts of their own.

Now take it further. Imagine there are no human shareholders who want yachts, just banks who lend the company money in order to increase their own value. And imagine there are no soccer moms anymore; the company makes batteries for the trucks that ship raw materials from place to place. Every non-economic goal has been stripped away from the company; it’s just an appendage of Global Development.

Now take it even further, and imagine this is what’s happened everywhere. There are no humans left; it isn’t economically efficient to continue having humans. Algorithm-run banks lend money to algorithm-run companies that produce goods for other algorithm-run companies and so on ad infinitum. Such a masturbatory economy would have all the signs of economic growth we have today. It could build itself new mines to create raw materials, construct new roads and railways to transport them, build huge factories to manufacture them into robots, then sell the robots to whatever companies need more robot workers. It might even eventually invent space travel to reach new worlds full of raw materials. Maybe it would develop powerful militaries to conquer alien worlds and steal their technological secrets that could increase efficiency. It would be vast, incredibly efficient, and utterly pointless. The real-life incarnation of those strategy games where you mine Resources to build new Weapons to conquer new Territories from which you mine more Resources and so on forever.

But this seems to me the natural end of the economic system. Right now it needs humans only as laborers, investors, and consumers. But robot laborers are potentially more efficient, companies based around algorithmic trading are already pushing out human investors, and most consumers already aren’t individuals – they’re companies and governments and organizations. At each step you can gain efficiency by eliminating humans, until finally humans aren’t involved anywhere.

True to form, Land doesn’t see this as a dystopia – I think he conflates “maximally efficient economy” with “God”, which is a hell of a thing to conflate – but I do. And I think it provides an important new lens with which to look at the Age of Em.

The Age of Em is an economy in the early stages of such a transformation. Instead of being able to replace everything with literal robots, it replaces them with humans who have had some aspects of their humanity stripped away. Biological bodies. The desire and ability to have children normally. Robin doesn’t think people will lose all leisure time and non-work-related desires, but he doesn’t seem too sure about this and it doesn’t seem to bother him much if they do.

I envision a spectrum between the current world of humans and Nick Land’s Ascended Economy. Somewhere on the spectrum we have ems who get leisure time. A little further on the spectrum we have ems who don’t get leisure time.

But we can go further. Hanson imagines that we can “tweak” em minds. We may not understand the brain enough to create totally new intelligences from the ground up, but by his Age of Em we should understand it well enough to make a few minor hacks, the same way even somebody who doesn’t know HTML or CSS can usually figure out how to change the background color of a webpage with enough prodding. Many of these mind tweaks will be the equivalent of psychiatric drugs – some might even be computer simulations of what we observe to happen when we give psychiatric drugs to a biological brain. But these tweaks will necessarily be much stronger and more versatile, since we no longer care about bodily side effects (ems don’t have bodies) and we can apply it to only a single small region of the brain and avoid actions anywhere else. You could also very quickly advance brain science – the main limits today are practical (it’s really hard to open up somebody’s brain and do stuff to it without killing them) and ethical (the government might have some words with you if you tried). An Age of Em would remove both obstacles, and give you the added bonus of being able to make thousands of copies of your test subjects for randomized controlled trials, reloading any from a saved copy if they died. Hanson envisions that:

As the em world is a very competitive world where sex is not needed for reproduction, and as sex can be time and attention-consuming, ems may try to suppress sexuality, via mind tweaks that produce effects analogous to castration. Such effects might be temporary, perhaps with a consciously controllable on-off switch…it is possible that em brain tweaks could be found to greatly reduce natural human desires for sex and related romantic and intimate pair bonding without reducing em productivity. It is also possible that many of the most productive ems would accept such tweaks.

Possible? I can do that right now with a high enough dose of Paxil, and I don’t even have to upload your brain to a computer first. Fun stories about Musk #2633590 and Swift #105051 aside, I expect this would happen about ten minutes after the advent of the Age of Em, and we would have taken another step down the path to the Ascended Economy.

There are dozens of other such tweaks I can think of, but let me focus on two.

First, stimulants have a very powerful ability to focus the brain on the task at hand, as anybody who’s taken Adderall or modafinil can attest. Their main drawbacks are addictiveness and health concerns, but in a world where such pills can be applied as mental tweaks, where minds have no bodies, and where any mind that gets too screwed up can be reloaded from a backup copy, these are barely concerns at all. Many of the purely mental side effects of stimulants come from their effects in parts of the brain not vital to the stimulant effect. If we can selectively apply Adderall to certain brain centers but not others, then unapply it at will, then from employers’ point of view there’s no reason not to have all workers dosed with superior year 2100 versions of Adderall at all times. I worry that not only will workers not have any leisure time, but they’ll be neurologically incapable of having their minds drift off while on the job. Davis’ contract-drafting em who starts wondering about philosophy on the job wouldn’t get terminated. He would just have his simulated-Adderall dose increased.

Second, Robin managed to write an entire book about emulated minds without using the word “wireheading”. This is another thing we can do right now, with today’s technology – but once it’s a line of code and not a costly brain surgery, it should become nigh-universal. Give ems the control switches to their own reward centers and all questions about leisure time become irrelevant. Give bosses the control switches to their employees’ reward centers, and the situation changes markedly. Hanson says that there probably won’t be too much slavery in the em world, because it will likely have strong rule of law, because slaves aren’t as productive as free workers, and there’s little advantage to enslaving someone when you could just pay them subsistence wages anyway. But slavery isn’t nearly as abject and inferior a condition as the one where somebody else has the control switch to your reward center. Combine that with the stimulant use mentioned above, and you can have people who will never have nor want to have any thought about anything other than working on the precise task at which they are supposed to be working at any given time.

This is something I worry about even in the context of normal biological humans. But Hanson already believes em worlds will have few regulations and be able to ignore the moral horror of 99% of the population by copying and using the 1% who are okay with something. Combine this with a situation where brains are easily accessible and tweakable, and this sort of scenario becomes horribly likely.

I see almost no interesting difference between an em world with full use of these tweaks and an Ascended Economy world. Yes, there are things that look vaguely human in outline laboring in the one and not the other, but it’s not like there will be different thought processes or different results. I’m not even sure what it would mean for the ems to be conscious in a world like this – they’re not doing anything interesting with the consciousness. The best we could say about this is that if the wireheading is used liberally it’s a lite version of the world where everything gets converted to hedonium.

V.

In a book full of weird ideas, there is only one idea rejected as too weird. And in a book written in a professorial monotone, there’s only one point at which Hanson expresses anything like emotion:

Some people foresee a rapid local “intelligence explosion” happening soon after a smart AI system can usefully modify its local architecture (Chalmers 2010; Hanson and Yudkowsky 2013; Yudkowsky 2013; Bostrom 2014)…Honestly to me this local intelligence explosion scenario looks suspiciously like a super-villain comic book plot. A flash of insight by a lone genius lets him create a genius AI. Hidden in its super-villain research lab lair, this guines villain AI works out unprecedented revolutions in AI design, turns itself into a super-genius, which then invents super-weapons and takes over the world. Bwa ha ha.

For someone who just got done talking about the sex lives of uploaded computers in millimeter-tall robot bodies running at 1000x human speed, Robin is sure quick to use the absurdity heuristic to straw-man intelligence explosion scenarios as “comic book plots”. Take away his weird authorial tic of using the words “genius” and “supervillain”, this scenario reduces to “Some group, perhaps Google, perhaps a university, invent an artificial intelligence smart enough to edit its own source code; exponentially growing intelligence without obvious bound follows shortly thereafter”. Yes, it’s weird to think that there may be a sudden quantum leap in intelligence like this, but no weirder than to think most of civilization will transition from human to em in the space of a year or two. I’m a little bit offended that this is the only idea given this level of dismissive treatment. Since I do have immense respect for Robin, I hope my offense doesn’t color the following thoughts too much.

Hanson’s arguments against AI seem somewhat motivated. He admits that AI researchers generally estimate less than 50 years before we get human-level artificial intelligence, a span shorter than his estimate of a century until we can upload ems. He even admits that no AI researcher thinks ems are a plausible route to AI. But he dismisses this by saying when he asks AI experts informally, they say that in their own field, they have only noticed about 5-10% of the progress they expect would be needed to reach human intelligence over the past twenty years. He then multiplies out to say that it will probably take at least 400 years to reach human-level AI. I have two complaints about this estimate.

First, he is explicitly ignoring published papers surveying hundreds of researchers using validated techniques, in favor of what he describes as “meeting experienced AI experts informally”. But even though he feels comfortable rejecting vast surveys of AI experts as potentially biased, as best I can tell he does not ask a single neuroscientist to estimate the date at which brain scanning and simulation might be available. He just says that “it seems plausible that sufficient progress will be made in roughly a century or so”, citing a few hopeful articles by very enthusiastic futurists who are not neuroscientists or scanning professionals themselves and have not talked to any. This seems to me to be an extreme example of isolated demands for rigor. No matter how many AI scientists think AI is soon, Hanson will cherry-pick the surveying procedures and results that make it look far. But if a few futurists think brain emulation is possible, then no matter what anybody else thinks that’s good enough for him.

Second, one would expect that even if there were only 5-10% progress over the last twenty years, then there would be faster progress in the future, since the future will have a bigger economy, better supporting technology, and more resources invested in AI research. Robin answers this objection by saying that “increases in research funding usually give much less than proportionate increases in research progress” and cites Alston et al 2011. I looked up Alston et al 2011, and it is a paper relating crop productivity to government funding of agriculture research. There was no attempt to relate its findings to any field other than agriculture, nor to any type of funding other than government. But studies show that while public research funding often does have minimal effects, the effect of private research funding is usually much larger. A single sentence citing a study in crop productivity to apply to artificial intelligence while ignoring much more relevant results that contradict it seems like a really weak argument for a statement as potentially surprising as “amount of research does not affect technological progress”.

I realize that Hanson has done a lot more work on this topic and he couldn’t fit all of it in this book. I disagree with his other work too, and I’ve said so elsewhere. For now I just want to say that the arguments in this book seem weak to me.

I also want to mention what seems to me a very Hansonian counterargument to the ems-come-first scenario: we have always developed de novo technology before understanding the relevant biology. We built automobiles by figuring out the physics of combustion engines, not by studying human muscles and creating mechanical imitations of myosin and actin. Although the Wright brothers were inspired by birds, their first plane was not an ornithopter. Our power plants use coal and uranium instead of the Krebs Cycle. Biology is really hard. Even slavishly copying biology is really hard. I don’t think Hanson and the futurists he cites understand the scale of the problem they’ve set themselves.

Current cutting-edge brain emulation projects have found their work much harder than expected. Simulating a nematode is pretty much the rock-bottom easiest thing in this category, since they are tiny primitive worms with only a few neurons; the history of the field is a litany of failures, with current leader OpenWorm “reluctant to make bold claims about its current resemblance to biological behavior”. A more ambitious $1.3 billion attempt to simulate a tiny portion of a rat brain has gone down in history as a legendary failure (politics were involved, but I expect they would be involved in a plan to upload a human too). And these are just attempts to get something that behaves vaguely like a nematode or rat. Actually uploading a human, keeping their memory and personality intact, and not having them go insane afterwards boggles the mind. We’re still not sure how much small molecules matter to brain function, how much glial cells matter to brain function, how many things in the brain are or aren’t local. AI researchers are making programs that can defeat chess grandmasters; upload researchers are still struggling to make a worm that will wriggle. The right analogy for modern attempts to upload human brains isn’t modern attempts at designing AI. It’s an attempt at designing AI by someone who doesn’t even know how to plug in a computer.

VI.

I guess what really bothers me about Hanson’s pooh-poohing of AI is him calling it “a comic book plot”. To me, it’s Hanson’s scenario that seems science-fiction-ish.

I say this not as a generic insult but as a pointer at a specific category of errors. In Star Wars, the Rebellion had all of these beautiful hyperspace-capable starfighters that could shoot laser beams and explore galaxies – and they still had human pilots. 1977 thought the pangalactic future would still be using people to pilot its military aircraft; in reality, even 2016 is moving away from this.

Science fiction books have to tell interesting stories, and interesting stories are about humans or human-like entities. We can enjoy stories about aliens or robots as long as those aliens and robots are still approximately human-sized, human-shaped, human-intelligence, and doing human-type things. A Star Wars in which all of the X-Wings were combat drones wouldn’t have done anything for us. So when I accuse something of being science-fiction-ish, I mean bending over backwards – and ignoring the evidence – in order to give basically human-shaped beings a central role.

This is my critique of Robin. As weird as the Age of Em is, it makes sure never to be weird in ways that warp the fundamental humanity of its participants. Ems might be copied and pasted like so many .JPGs, but they still fall in love, form clans, and go on vacations.

In contrast, I expect that we’ll get some kind of AI that will be totally inhuman and much harder to write sympathetic stories about. If we get ems after all, I expect them to be lobotomized and drugged until they become effectively inhuman, cogs in the Ascended Economy that would no more fall in love than an automobile would eat hay and whinny. Robin’s interest in keeping his protagonists relatable makes his book fascinating, engaging, and probably wrong.

I almost said “and probably less horrible than we should actually expect”, but I’m not sure that’s true. With a certain amount of horror-suppressing, the Ascended Economy can be written off as morally neutral – either having no conscious thought, or stably wireheaded. All of Robin’s points about how normal non-uploaded humans should be able to survive an Ascended Economy at least for a while seem accurate. So morally valuable actors might continue to exist in weird Amish-style enclaves, living a post-scarcity lifestyle off the proceeds of their investments, while all the while the Ascended Economy buzzes around them, doing weird inhuman things that encroach upon them not at all. This seems slightly worse than a Friendly AI scenario, but much better than we have any right to expect of the future.

I highly recommend Age of Em as a fantastically fun read and a great introduction to these concepts. It’s engaging, readable, and weird. I just don’t know if it’s weird enough.

# Ascended Economy?

[Obviously speculative futurism is obviously speculative. Complex futurism may be impossible and I should feel bad for doing it anyway. This is “inspired by” Nick Land – I don’t want to credit him fully since I may be misinterpreting him, and I also don’t want to avoid crediting him at all, so call it “inspired”.]

I.

My review of Age of Em mentioned the idea of an “ascended economy”, one where economic activity drifted further and further from human control until finally there was no relation at all. Many people rightly questioned that idea, so let me try to expand on it further. What I said there, slightly edited for clarity:

Imagine a company that manufactures batteries for electric cars. The inventor of the batteries might be a scientist who really believes in the power of technology to improve the human race. The workers who help build the batteries might just be trying to earn money to support their families. The CEO might be running the business because he wants to buy a really big yacht. The shareholders might be holding the stock to help save for a comfortable retirement. And the whole thing is there to eventually, somewhere down the line, let a suburban mom buy a car to take her kid to soccer practice. Like most companies the battery-making company is primarily a profit-making operation, but the profit-making-ness draws on a lot of not-purely-economic actors and their not-purely-economic subgoals.

Now imagine the company fires the inventor and replaces him with a genetic algorithm that optimizes battery design. It fires all its employees and replaces them with robots. It fires the CEO and replaces him with a superintelligent business-running algorithm. All of these are good decisions, from a profitability perspective. We can absolutely imagine a profit-driven shareholder-value-maximizing company doing all these things. But it reduces the company’s non-masturbatory participation in an economy that points outside itself, limits it to just a tenuous connection with soccer moms and maybe some shareholders who want yachts of their own.

Now take it further. Imagine that instead of being owned by humans directly, it’s owned by an algorithm-controlled venture capital fund. And imagine there are no soccer moms anymore; the company makes batteries for the trucks that ship raw materials from place to place. Every non-economic goal has been stripped away from the company; it’s just an appendage of Global Development.

Now take it even further, and imagine this is what’s happened everywhere. Algorithm-run banks lend money to algorithm-run companies that produce goods for other algorithm-run companies and so on ad infinitum. Such a masturbatory economy would have all the signs of economic growth we have today. It could build itself new mines to create raw materials, construct new roads and railways to transport them, build huge factories to manufacture them into robots, then sell the robots to whatever companies need more robot workers. It might even eventually invent space travel to reach new worlds full of raw materials. Maybe it would develop powerful militaries to conquer alien worlds and steal their technological secrets that could increase efficiency. It would be vast, incredibly efficient, and utterly pointless. The real-life incarnation of those strategy games where you mine Resources to build new Weapons to conquer new Territories from which you mine more Resources and so on forever.

This is obviously weird and I probably went too far, but let me try to explain my reasoning.

The part about replacing workers with robots isn’t too weird; lots of industries have already done that. There’s a whole big debate over to what degree that will intensify, and whether unemployed humans will find jobs somewhere else, or whether there will only be jobs for creative people with a certain education level or IQ. This part is well-discussed and I don’t have much to add.

But lately there’s also been discussion of automating corporations themselves. I don’t know much about Ethereum (and I probably shouldn’t guess since I think the inventor reads this blog and could call me on it) but as I understand it they aim to replace corporate governance with algorithms. For example, the DAO is a leaderless investment fund that allocates money according to member votes. Right now this isn’t super interesting; algorithms can’t make too many difficult business decisions so it’s limited to corporations that just do a couple of primitive actions (and why would anyone want a democratic venture fund?). But once we get closer to true AI, they might be able to make the sort of business decisions that a CEO does today. The end goal is intelligent corporations controlled by nobody but themselves.

This very blog has an advertisement for a group trying to make investment decisions based on machine learning. If they succeed, how long is it before some programmer combines a successful machine investor with a DAO-style investment fund, and creates an entity that takes humans out of the loop completely? You send it your money, a couple years later it gives you back hopefully more money, with no humans involved at any point. Such robo-investors might eventually become more efficient than Wall Street – after all, hedge fund managers get super rich by skimming money off the top, and any entity that doesn’t do that would have an advantage above and beyond its investment acumen.

If capital investment gets automated, corporate governance gets automated, and labor gets automated, we might end up with the creepy prospect of ascended corporations – robot companies with robot workers owned by robot capitalists. Humans could become irrelevant to most economic activity. Run such an economy for a few hundred years and what do you get?

II.

But in the end isn’t all this about humans? Humans as the investors giving their money to the robo-venture-capitalists, then reaping the gains of their success? And humans as the end consumers whom everyone is eventually trying to please?

It’s possible to imagine accidentally forming stable economic loops that don’t involve humans. Imagine a mining-robot company that took one input (steel) and produced one output (mining-robots), which it would sell either for money or for steel below a certain price. And imagine a steel-mining company that took one input (mining-robots) and produced one output (steel) which it would sell for either money or for mining-robots below a certain price. The two companies could get into a stable loop and end up tiling the universe with steel and mining-robots without caring whether anybody else wanted either. Obviously the real economy is a zillion times more complex than that, and I’m nowhere near the level of understanding I would need to say if there’s any chance that an entire self-sustaining economy worth of things could produce a loop like that. But I guess you only need one.

I think we can get around this in a causal-historical perspective, where we start with only humans and no corporations. The first corporations that come into existence have to be those that want to sell goods to humans. The next level of corporations can be those that sell goods to corporations that sell to humans. And so on. So unless a stable loop forms by accident, all corporations should exist to serve humans. A sufficiently rich human could finance the creation of a stable loop if they wanted to, but why would they want to? Since corporations exist only to satisfy human demand on some level or another, and there’s no demand for stable loops, corporations wouldn’t finance the development of stable loops, except by accident.

(for an interesting accidental stable loop, check out this article on the time two bidding algorithms accidentally raised the price of a book on fly genetics to more than $20 million)

Likewise, I think humans should always be the stockholders of last resort. Since humans will have to invest in the first corporation, even if that corporation invests in other corporations which invest in other corporations in turn, eventually it all bottoms down in humans (is this right?)

The only way I can see humans being eliminated from the picture is, again, by accident. If there are a hundred layers between some raw material corporation and humans, then if each layer is slightly skew to what the layer below it wants, the hundredth layer could be really really skew. Theoretically all our companies today are grounded in serving the needs of humans, but people are still thinking of spending millions of dollars to build floating platforms exactly halfway between New York and London in order to exploit light-speed delays to arbitrage financial markets better, and I’m not sure which human’s needs that serves exactly. I don’t know if there are bounds to how much of an economy can be that kind of thing.

Finally, humans might deliberately create small nonhuman entities with base level “preferences”. For example, a wealthy philanthropist might create an ascended charitable organization which supports mathematical research. Now 99.9% of base-level preferences guiding the economy would be human preferences, and 0.1% might be a hard-coded preference for mathematics research. But since non-human agents at the base of the economy would only be as powerful as the proportion of the money supply they hold, most of the economy would probably still overwhelmingly be geared towards humans unless something went wrong.

Since the economy could grow much faster than human populations, the economy-to-supposed-consumer ratio might become so high that things start becoming ridiculous. If the economy became a light-speed shockwave of economium (a form of matter that maximizes shareholder return, by analogy to computronium and hedonium) spreading across the galaxy, how does all that productive power end up serving the same few billion humans we have now? It would probably be really wasteful, the cosmic equivalent of those people who specialize in getting water from specific glaciers on demand for the super-rich because the super-rich can’t think of anything better to do with their money. Except now the glaciers are on Pluto.

III.

Glacier water from Pluto sounds pretty good. And we can hope that things will get so post-scarcity that governments and private charities give each citizen a few shares in the Ascended Economy to share the gains with non-investors. This would at least temporarily be a really good outcome.

But in the long term it reduces the political problem of regulating corporations to the scientific problem of Friendly AI, which is really bad.

Even today, a lot of corporations do things that effectively maximize shareholder value but which we consider socially irresponsible. Environmental devastation, slave labor, regulatory capture, funding biased science, lawfare against critics – the list goes on and on. They have a simple goal – make money – whereas what we really want them to do is much more complicated and harder to measure – make money without engaging in unethical behavior or creating externalities. We try to use regulatory injunctions, and it sort of helps, but because those go against a corporation’s natural goals they try their best to find loopholes and usually succeed – or just take over the regulators trying to control them.

This is bad enough with bricks-and-mortar companies run by normal-intelligence humans. But it would probably be much worse with ascended corporations. They would have no ethical qualms we didn’t program into them – and again, programming ethics into them would be the Friendly AI problem, which is really hard. And they would be near-impossible to regulate; most existing frameworks for such companies are built on crypto-currency and exist on the cloud in a way that transcends national borders.

(A quick and very simple example of an un-regulate-able ascended corporation – I don’t think it would be too hard to set up an automated version of Uber. I mean, the core Uber app is already an automated version of Uber, it just has company offices and CEOs and executives and so on doing public relations and marketing and stuff. But if the government ever banned Uber the company, could somebody just code another ride-sharing app that dealt securely in Bitcoins? And then have it skim a little bit off the top, which it offered as a bounty to anybody who gave it the processing power it would need to run? And maybe sent a little profit to the programmer who wrote the thing? Sure, the government could arrest the programmer, but short of arresting every driver and passenger there would be no way to destroy the company itself.)

The more ascended corporations there are trying to maximize shareholder value, the more chance there is some will cause negative externalities. But there’s a limited amount we would be able to do about them. This is true today too, but at least today we maintain the illusion that if we just elected Bernie Sanders we could reverse the ravages of capitalism and get an economy that cares about the environment and the family and the common man. An Ascended Economy would destroy that illusion.

How bad would it get? Once ascended corporations reach human or superhuman level intelligences, we run into the same AI goal-alignment problems as anywhere else. Would an ascended corporation pave over the Amazon to make a buck? Of course it would; even human corporations today do that, and an ascended corporation that didn’t have all human ethics programmed in might not even get that it was wrong. What if we programmed the corporation to follow local regulations, and Brazil banned paving over the Amazon? This is an example of trying to control AIs through goals plus injunctions – a tactic Bostrom finds very dubious. It’s essentially challenging a superintelligence to a battle of wits – “here’s something you want, and here are some rules telling you that you can’t get it, can you find a loophole in the rules?” If the superintelligence is super enough, the answer will always be yes.

From there we go into the really gnarly parts of AI goal alignment theory. Would an ascended corporation destroy South America entirely to make a buck? Depending on how it understood its imperative to maximize shareholder value, it might. Yes, this would probably kill many of its shareholders, but its goal is to “maximize shareholder value”, not to keep its shareholders alive to enjoy that value. It might even be willing to destroy humanity itself if other parts of the Ascended Economy would pick up the slack as investors.

(And then there are the weirder problems, like ascended corporations hacking into the stock market and wireheading themselves. When this happens, I want credit for being the first person to predict it.)

Maybe the most hopeful scenario is that once ascended corporations achieved human-level intelligence they might do something game-theoretic and set up a rule-of-law among themselves in order to protect economic growth. I wouldn’t want to begin to speculate on that, but maybe it would involve not killing all humans? Or maybe it would just involve taking over the stock market, formally setting the share price of every company to infinity, and then never doing anything again? I don’t know, and I expect it would get pretty weird.

IV.

I don’t think the future will be like this. This is nowhere near weird enough to be the real future. I think superintelligence is probably too unstable. It will explode while still in the lab and create some kind of technological singularity before people have a chance to produce an entire economy around it.

But given Robin’s assumptions in Age of Em – hard AI, no near-term intelligence explosion, fast economic growth – but ditching his idea of human-like em minds as important components of the labor force – I think something like this would be where we would end up. It probably wouldn’t be so bad for the first couple of years. But eventually ascended corporations would start reaching the point where we might as well think of them as superintelligent AIs. Maybe this world would be friendlier towards AI goal alignment research than Yudkowsky and Bostrom’s scenarios, since at least here we could see it coming, there was no instant explosion, and a lot of different entities approach superintelligence around the same time. But given that the smartest things around are encrypted, uncontrollable, unregulated entities that don’t have humans’ best interests at heart, I’m not sure they would be in much shape to handle the transition.

Welcome to Meteor Deflection Corp. Our quest is to maximize shareholder value this quarter.

— Instance Of Class (@InstanceOfClass) December 26, 2013

# Ketamine Research In A New Light

[Preliminary drawing of very far-out conclusions from research that hasn’t even been 100% confirmed yet]

A few weeks ago, Nature published a bombshell study showing that ketamine’s antidepressant effects were actually caused by a metabolite, 2S,6S;2R,6R-hydroxynorketamine (don’t worry about the name; within ten years it’ll be called JOYVIVA™®© and you’ll catch yourself humming advertising jingles about it in the shower). Unlike ketamine, which is addictive and produces scary dissociative experiences, the metabolite is pretty safe. This is a big deal clinically, because it makes it easier and safer to prescribe to depressed people.

It’s also a big deal scientifically. Ketamine is a strong NMDA receptor antagonist; the metabolite is an AMPA agonist – they have different mechanisms of action. Knowing the real story behind why ketamine works will hopefully speed efforts to understand the nature of depression.

But I’m also interested in it from another angle. For the last ten years, everyone has been excited about ketamine. In a field that gets mocked for not having any really useful clinical discoveries in the last thirty years, ketamine was proof that progress was possible. It was the Exciting New Thing that everybody wanted to do research about.

Given the whole replication crisis thing, I wondered. You’ve got a community of people who think that NMDA antagonism and dissociation are somehow related to depression. If the latest study is true, all that was false. This is good; science is supposed to be self-correcting. But what about before it self-corrected? Did researchers virtuously say “I know the paradigm says NMDA is essential to depression, and nobody’s come up with a better idea yet, but there are some troubling inconsistencies in that picture”? Or did they tinker with their studies until they got the results they expected, then triumphantly declare that they had confirmed the dominant paradigm was right about everything all along?

This is too complicated an issue for me to be really sure, but overall the picture I found was mixed.

A big review of ketamine and NDMA antagonism came out last year. In this case, I was most interested in the section on other NMDA antagonists – if ketamine’s efficacy is unrelated to its NMDA antagonism, then we shouldn’t expect other NMDA antagonists to be antidepressants like ketamine. So if the review found that other NMDA antagonists worked great, that would be a sign that something fishy was going on. But in fact the abstract says:

The antidepressant efficacy of ketamine, and perhaps D-cycloserine and rapastinel, holds promise for future glutamate-modulating strategies; however, the ineffectiveness of other NMDA antagonists suggests that any forthcoming advances will depend on improving out understanding of ketamine’s mechanism of action.

This is pretty impressive; they basically admit that other NMDA antagonists don’t work and that maybe this means they don’t really understand ketamine.

But dig deeper, and you find a less sanguine picture. The body of the paper lists notes five NMDA antagonists as confirmed ineffective – memantine, lanicemine, nitrous oxide, traxoprodil, and MK-0657. But the paper itself notes that all of these were effective on some endpoints and not others, and the decision that they were ineffective was sort of a judgment call by the reviewers. Just to give an example, there’s only ever been one study done on traxoprodil. Since the reviewers reviewed this one study and declared it ineffective, you might expect the study to be negative. But here’s the abstract of the study itself:

On the prespecified main outcome measure (change from baseline in the Montgomery-Asberg Depression Rating Scale total score at day 5 of period 2), CP-101,606 produced a greater decrease than did placebo (mean difference, 8.6; 80% confidence interval, -12.3 to -4.5) (P < 0.10). Hamilton Depression Rating Scale response rate was 60% for CP-101,606 versus 20% for placebo. Seventy-eight percent of CP-101,606-treated responders maintained response status for at least 1 week after the infusion. CP-101,606 was safe, generally well tolerated, and capable of producing an antidepressant response without also producing a dissociative reaction. Antagonism of the NR2B subtype of the N-methyl-D-aspartate receptor may be a fruitful target for the development of a new antidepressant with more robust effects and a faster onset compared with those currently available and capable of working when existing antidepressants do not.

Read this quickly, and it looks like they’ve confirmed traxoprodil is pretty great. The reviewers say it isn’t. They argue that p < 0.10 isn't good enough (I think the study was trying to use a one-sided t-test or something?), and that of five different days when responses were measured (day 2, 5, 8, 12, and 15 after the infusion), there was only a difference on day 5. Apparently this isn't good enough for the reviewers.  
  
On the other hand, take rapastinel, one of the NMDA antagonists the reviewers say "holds promise".  
  
  
  
No statistically significant differences were observed in rates of treatment response or symptom remission associated with placebo (64% and 42%, respectively) versus rapastinel at any dose (up to 70% and 53%, respectively). However, statistically significant differences in the reduction of the 17-item HAM-D scores were observed for the 5-mg/kg dose at all intervals except day 14) and the 10-mg/kg dose at day 1 and day 3. Neither the low nor the high rapastinel doses were associated with significant greater 17-item HAM-D score reduction than placebo, leading the authors to posit an inverted U-shape dose-response curve.

Sometimes things do have inverted U-shaped dose-response curves – for some discussion of why, read the Last Psychiatrist’s Most Important Article On Psychiatry. But a study that shows no treatment response or symptom response, and the test score response is only on a medium dose but not a high or a low dose – that makes me kind of suspicious.

Why is the review so much more accepting of these ambiguous results than of the last set of ambiguous results? Psych blog 1BoringOldMan points out that the original study was done by the company making rapastinel and two authors of the review article I’m citing were affiliated with the companies that are developing rapastinel. And that at least one of them has a “legendary” history of conflicts of interest.

I don’t want to say for sure this is what’s going on. For all anybody knows, rapastinel might work – the NMDA and AMPA systems are really connected, and the base rate of a randomly chosen compound being an antidepressant is higher than you’d think. But I think it’s at least one possible explanation.

This review article also gets into the nitty-gritty of mechanisms of action:

That other NMDA channel blockers have yet to replicate ketamine’s rapid antidepressant effects has led to speculation that ketamine’s antidepressant properties may not be mediated via the NMDA receptor at all…additional evidence indicates that activation of glutamatergic AMPA receptors is necessary for ketamine’s antidepressant effects. Specifically, coadministration of an AMPA receptor antagonist has been shown to block ketamine’s antidepressant-like behavioral effects.

So that’s neat.

Two other relevant studies: Do The Dissociative Side Effects Of Ketamine Mediate Its Antidepressant Effects finds that they do, which contradicts the recent metabolite-related findings. On the other hand, the two papers share some authors, so I’m tempted to say it was an honest mistake. This paper incidentally finds that the dissociative effects of ketamine are not related to its antidepressant effects, which I think makes more sense now.

The other studies I found were mostly compatible with the new results, with a lot of people expressing doubt about whether NMDA really mediated ketamine, a lot of people finding null results for other NMDA antagonist medications, and a lot of people saying there were weird hints that AMPA was involved somewhere.

I feel kind of premature doing this, because as much as I think it’s elegant the discovery about the metabolite hasn’t been totally confirmed yet. But assuming it’s right, psychiatry comes out of this looking sort of okay. There were a lot of early results with a lot of hype. But the big review articles mostly put these in their place and were able to come up with the right results and fit the pieces together.

The one place this wasn’t so clear was when there were conflicts of interest. If we assume rapastinel doesn’t really work (which right now would be very preliminary and I’m not actually saying this, but these latest findings do seem to imply that), various teams made up of people affiliated with rapastinel’s manufacturers were unable to determine this (neither was the FDA, who just just gave rapastinel “breakthrough drug” status, apparently on the strength of industry studies).

A big reason I’m concerned about this is that I want to know how much to trust the rest of the psychiatric literature – for example, those claims about SSRIs being mostly ineffective. An answer of “you can trust it a lot, except in cases of conflicts of interest” would be a mixed bag. Almost every drug was originally researched and promoted by people with conflicts of interest, and then we trust the academics to catch up with them later and keep them honest. I don’t think this system has failed us too terribly yet. But it’s important to remember that that is the system.

# Three More Articles On Poverty, And Why They Disagree With Each Other

[Posts are decreased in quantity and quality because I’m on vacation; normal schedule to return next week]

Wealth, Health, and Child Development is a study of Swedish lottery winners which finds that winning the lottery doesn’t make them or their children any healthier, better educated, or more prosocial. It fits in with a large literature of studies showing the same – for example, I discussed here the Cherokee land lottery, where the families of Georgians who were randomly given a gift of lots of lucrative land were no better off a generation later. And let’s not forget that the best evidence suggests poverty traps don’t exist.

Why Do The Poor Make Such Poor Decisions also involves the Cherokee, but comes to the opposite conclusion. The main study discussed follows an impoverished group of Cherokee Indians as a casino opened on their land. The casino was very successful and the profits were distributed among the (relatively small) Indian tribe, meaning each Cherokee family got about $6,000 extra. Some researchers had been studying the Cherokee before for other reasons, and found that the boost in incomes decreased behavioral problems in teenagers, juvenile crime, and improved school performance. I don’t see huge evidence that anybody’s checked to what degree this persists into adulthood, but it’s already gotten past the early childhood period where these things tend to fade out. And even if the decreased crime is just in adolescence, adolescent crime can still have a really negative impact on people’s lives. I don’t really trust a lot of the studies listed here, but the main Cherokee one seems pretty solid.

Can America’s Poor Save A Large Share Of Their Incomes? by Scott Sumner is sadly Cherokee-less. It describes a Chinese immigrant to the US who has the same sort low-paying job as many poor Americans but manages to save > $1000/month. It mentions my observation a little while ago that it was strange that the poor are earning 10x (in real value) what they did in 1900, poor people in 1900 survived just fine, but poor people today don’t find themselves with ten times the money they need to survive. Sumner suggests that it is economically possible for poor people today to save much of their income, but that they don’t because they’re not the kind of people who do that kind of thing. When the sort of people who do do that kind of thing find themselves poor – like Chinese immigrants – they tend to be poor very temporarily and have no trouble getting out of poverty even with the same jobs as everyone else.

These articles sort of contradict each other. The first contradicts the second – does giving people money improve life outcomes, or doesn’t it? And the second sort of contradicts the third – if poor people’s budget will expand to fit the money available, such that 2010’s $15000 leaves people just as desperate as 1900’s $1500, what does it matter if some people get an extra $6000?

The contradiction between the first and second reminds me of Tucker-Drob on IQ. He resolves a long-standing debate on whether intelligence is more heritable in poor than in rich individuals by finding this was true in the US but not in Europe. This suggests that American poverty can genuinely lower IQ (and presumably all the other good things associated with IQ like responsibility and prosocial behavior), but European poverty can’t. The study didn’t find this to be related to the US’ greater racial diversity, but it might have to do with the worse social safety net or just changes in the level and nature of poverty. Take this seriously, and it reconciles the first and second article. Getting more money might not help long-term outcomes in Sweden, but in certain kinds of extreme poverty in America – like the type you might find on an Indian reservation – maybe it would.

The third article is more complicated. The second article says:

What, then, is the cause of mental health problems among the poor? Nature or culture? Both, was Costello’s conclusion, because the stress of poverty puts people genetically predisposed to develop an illness or disorder at an elevated risk.

Maybe with the right genes it might be easier to rise out of poverty; I guess the stories of famous entrepreneurs who did exactly that already suggest that. With the wrong genes, it might be much harder but – at least in America, at least if given large amounts of money – still possible.

Also, regarding that Chinese immigrant – I, too, have worked a $20,000/year job and managed to save a lot of money while doing so. I think my “secret” was not having a car, debts, drugs, or dependents; it seems the Chinese guy’s secret is the same. Exactly how easy this strategy is for the average person is left as an exercise for the reader, but I’m impressed with how culturally malleable it seems to be. If we’re worse at this kind of thing today than in 1900, maybe the extra is just compensating for those sorts of problems.

I think this can be considered me slightly changing my opinion stated here to be more optimistic about the possibility of alleviating the most extreme poverty. But it still seems like money transfers are the way to go.

# Against Dog Whistle-ism

I.

Back during the primary, Ted Cruz said he was against “New York values”.

A chump might figure that, being a Texan whose base is in the South and Midwest, he was making the usual condemnation of coastal elites and arugula-eating liberals that every other Republican has made before him, maybe with a special nod to the fact that his two most relevant opponents, Donald Trump and Hillary Clinton, were both from New York.

But sophisticated people immediately detected this as an “anti-Semitic dog whistle”, eg Cruz’s secret way of saying he hated Jews. Because, you see, there are many Jews in New York. By the clever strategem of using words that had nothing to do with Jews or hatred, he was able to effectively communicate his Jew-hatred to other anti-Semites without anyone else picking up on it.

Except of course the entire media, which seized upon it as a single mass. New York values is coded anti-Semitism. New York values is a classic anti-Semitic slur. New York values is an anti-Semitic comment. New York values is an anti-Semitic code word. New York values gets called out as anti-Semitism. My favorite is this article whose headline claims that Ted Cruz “confirmed” that he meant his New York values comment to refer to Jews; the “confirmation” turned out to be that he referred to Donald Trump as having “chutzpah”. It takes a lot of word-I-am-apparently-not-allowed-to-say to frame that as a “confirmation”.

Meanwhile, back in Realityville (population: 6), Ted Cruz was attending synagogue services at his campaign tour, talking about his deep love and respect for Judaism, and getting described as “a hero” in many parts of the Orthodox Jewish community” for his stance that “if you will not stand with Israel and the Jews, then I will not stand with you.”

But he once said “New York values”, so clearly all of this was just really really deep cover for his anti-Semitism.

II.

Unlike Ted Cruz, former London mayor Ken Livingstone said something definitely Jew-related and definitely worrying.

A month or two ago a British MP named Naz Shah got in trouble for sharing a Facebook post saying Israel should be relocated to the United States. Fellow British politician Ken Livingstone defended her, and one thing led to another, and somewhere in the process he might have kind of said that Hitler supported Zionism.

This isn’t totally out of left field. During the Nazi period in Germany, some Nazis who wanted to get rid of the Jews and some Jews who wanted to get away from the Nazis created the Haavara Agreement, which facilitated German Jewish emigration to Palestine. Hitler was ambivalent on the idea but seems to have at least supported some parts of it at some points. But it seems fair to say that calling Hitler a supporter of Zionism was at the very least a creative interpretation of the historical record.

The media went further, again as a giant mass. Ken Livingstone is anti-Semitic. Ken Livingstone is anti-Semitic. Ken Livingstone is anti-Semitic. Ken Livingstone is anti-Semitic. Ken Livingstone is anti-Semitic. I understand he is now having to defend himself in front of a parliamentary hearing on anti-Semitism.

So. First things first. Ken Livingstone is tasteless, thoughtless, embarrassing, has his foot in his mouth, is inept, clownish and offensive, and clearly made a blunder of cosmic proportions.

But is he anti-Semitic?

When I think “anti-Semitic”, I think of people who don’t like, maybe even hate, Jews. I think of the medieval burghers who accused Jews of baking matzah with the blood of Christian children. I think of the Russians who would hold pogroms and kill Jews and burn their property. I think of the Nazis. I think of people who killed various distant family members of mine without a second thought.

Obviously Livingstone isn’t that anti-Semitic. But my question is, is he anti-Semitic at all? Is there any sense in which his comments reveal that, in his heart of hearts, he really doesn’t like Jews? That he thinks of them as less – even slightly less – than Gentiles? That if he were to end up as Prime Minister of Britain, this would be bad in a non-symbolic, non-stupid-statement-related way for Britain’s Jewish community? Does he just say dumb things, or do the dumb things reflect some underlying attitude of his that colors his relationship with Jews in general?

(speaking of “his relationship with Jews”, he brings up in his own defense that two of his ex-girlfriends are Jewish)

I haven’t seen anyone present any evidence that Livingstone has any different attitudes or policies towards Jews than anyone else in his general vicinity. I don’t think even his worst enemies suggest that during a hypothetical Livingstone administration he would try (or even want) to kick the Jews out of Britain, or make them wear gold stars, or hire fewer Jews for top posts (maybe he’d hire more, if he makes his hiring decisions the same way he makes his dating decisions). It sounds like he might be less sympathetic to Israel than some other British people, but I think he describes his preferred oppositional policies toward Israel pretty explicitly. I don’t think knowing that he made a very ill-advised comment about the Haavara agreement should make us believe he is lying about his Israel policies and would actually implement ones that are even more oppositional than he’s letting on.

Where am I going with this? It’s stupid to care that Ken Livingstone describes 1930s Germany in a weird way qua describing 1930s Germany in a weird way; he’s a politician and not a history textbook writer. It seems important only insofar as his weird description reveals something about him, insofar as it’s a sort of Freudian slip revealing deep-seated attitudes that he had otherwise managed to keep hidden. The British press framed Livingstone’s comments as an explosive revelation, an “aha! now we see what Labour is really like!” They’re really like…people who describe the 1930s in a really awkward and ill-advised way? That’s not a story. It’s a story only if the weird awkward description reveals more important attitudes of Livingstone’s and Labour’s that might actually affect the country in an important way.

But not only is nobody making this argument, but nobody even seems to think it’s an argument that has to be made. It’s just “this is an offensive thing involving Jews, that means it’s anti-Semitic, that means the guy who said it is anti-Semitic”. Maybe he is. I’m just not sure this incident proves much one way or the other.

III.

Nobody reads things online anymore unless they involve senseless violence, Harambe the gorilla, or Donald Trump. I can’t think of a relevant angle for the first two, so Trump it is.

Donald Trump is openly sexist. We know this because every article about him prominently declares that he is “openly sexist” or “openly misogynist” in precisely those words. Trump is openly misogynist. Trump is openly misogynist. Trump is openly misogynist. Trump shows blatant misogyny. Trump is openly sexist. Trump is openly sexist and gross.

But if you try to look for him being openly anything, the first quote anyone mentions is the one where he says Megyn Kelly has blood coming out of her “wherever”. As somebody who personally ends any list of more than three items with “… and whatever”, I may be more inclined than most to believe his claim that no anatomical reference was intended. But even if he was in fact talking about her anatomy – well, we’re back to Livingstone again. The comment is crude, stupid, puerile, offensive, gross, inappropriate, and whatever. But sexist?

When I think of “sexist” or “misogynist”, I think of somebody who thinks women are inferior to men, or hates women, or who thinks women shouldn’t be allowed to have good jobs or full human rights, or who wants to disadvantage women relative to men in some way.

This does not seem to apply very well to Trump. It’s been remarked several times that his policies are more “pro-women” in the political sense than almost any other Republican candidate in recent history – he defends Planned Parenthood, defends government support for child care, he’s flip-flopped to claiming he’s pro-life but is much less convincing about it than the average Republican. And back before his campaign, he seems to have been genuinely proud of his record as a pro-women employer. From his Art of the Deal, written in the late 1980s (ie long before he was campaigning):

The person I hired to be my personal representative overseeing the construction, Barbara Res, was the first woman ever put in charge of a skyscraper in New York…I’d watched her in construction meetings, and what I liked was that she took no guff from anyone. She was half the size of most of these bruising guys, but she wasn’t afraid to tell them off when she had to, and she knew how to get things done.

It’s funny. My own mother was a housewife all her life. And yet it’s turned out that I’ve hired a lot of women for top jobs, and they’ve been among my best people. Often, in fact, they are far more effective than the men around them. Louise Sunshine, who was an executive vice president in my company for ten years, was as relentless a fighter as you’ll ever meet. Blanche Sprague, the executive vice president who handles all sales and oversses the interior design of my buildings, is one of the best salespeople and managers I’ve ever met. Norma Foerderer, my executive assistant, is sweet and charming and very classy, but she’s steel underneath, and people who think she can be pushed around find out very quickly that they’re mistaken.

There have since been a bunch of news reports on how Trump was (according to the Washington Post) “ahead of his time in providing career advancement for women” and how “while some say he could be boorish, his companies nurtured and promoted women in an otherwise male-dominated industry”. According to internal (ie hard-to-confirm) numbers, his organization is among the few that have more female than male executives.

Meanwhile, when I check out sites like Women Hold Up Signs With Donald Trump’s Most Sexist Quotes, the women are holding up signs with quotes like “A person who is flat-chested is very hard to be a 10” (yes, he actually said that). This is undeniably boorish. But are we losing something when we act as if “boorish” and “sexist” are the same thing? Saying “Donald Trump is openly boorish” doesn’t have the same kind of ring to it.

This bothers me in the same way the accusations that Ken Livingstone is anti-Semitic bother me. If Trump thinks women aren’t attractive without big breasts, then His Kink Is Not My Kink But His Kink Is Okay. If Trump is dumb enough to say out loud that he thinks women aren’t attractive without big breasts, that says certain things about his public relations ability and his dignity-or-lack-thereof, but it sounds like it requires a lot more steps to suggest he is a bad person, or would have an anti-woman administration, or anything that we should actually care about.

(if you’re going to bring up “objectification”, then at least you have some sort of theory for how this tenuously connects, but it doesn’t really apply to the Megyn Kelly thing, and anyway, this)

And what bothers me most about this is that word “openly”. Donald Trump says a thousand times how much he wants to fight for women and thinks he will be a pro-women president, then makes some comments that some people interpret as revealing a deeper anti-women attitude, and all of a sudden he’s openly sexist? Maybe that word doesn’t mean what you think it means.

IV.

I don’t want to claim dog whistles don’t exist. The classic example is G. W. Bush giving a speech that includes a Bible verse. His secular listeners think “what a wise saying”, and his Christian listeners think “ah, I recognize that as a Bible verse, he must be very Christian”.

The thing is, we know G. W. Bush was pretty Christian. His desire to appeal to Christian conservatives isn’t really a secret. He might be able to modulate his message a little bit to his audience, but it wouldn’t be revealing a totally new side to his personality. Nor could somebody who understood his “dog whistles” predict his policy more accurately than somebody who just went off his stated platform.

I guess some of the examples above might have gotten kind of far from what people would usually call a “dog whistle”, but I feel like there’s an important dog-whistle-related common thread in all of these cases.

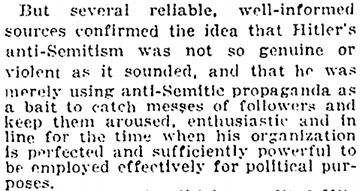
In particular, I worry there’s a certain narrative, which is catnip for the media: Many public figures are secretly virulently racist and sexist. If their secret is not discovered, they will gain power and use their racism and sexism to harm women and minorities. Many of their otherwise boring statements are actually part of a code revealing this secret, and so very interesting. Also, gaffes are royal roads to the unconscious which must be analyzed obsessively. By being very diligent and sophisticated, journalists can heroically ferret out which politicians have this secret racism, and reveal it to a grateful world.

There’s an old joke about a man who walks into a bar. The bar patrons are holding a weird ritual. One of them will say a number, like “twenty-seven”, and the others of them will break into laughter. He asks the bartender what’s going on. The bartender explains that they all come here so often that they’ve memorized all of each other’s jokes, and instead of telling them explicitly, they just give each a number, say the number, and laugh appropriately. The man is intrigued, so he shouts “Two thousand!”. The other patrons laugh uproariously. “Why did they laugh more at mine than any of the others?” he asks the bartender. The bartender answers “They’d never heard that one before!”

In the same way, although dog whistles do exist, the dog whistle narrative has gone so far that it’s become detached from any meaningful referent. It went from people saying racist things, to people saying things that implied they were racist, to people saying the kind of things that sound like things that could imply they are racist even though nobody believes that they are actually implying that. Saying things that sound like dog whistles has itself become the crime worthy of condemnation, with little interest in whether they imply anything about the speaker or not.

Against this narrative, I propose a different one – politicians’ beliefs and plans are best predicted by what they say their beliefs and plans are, or possibly what beliefs and plans they’ve supported in the past, or by anything other than treating their words as a secret code and trying to use them to infer that their real beliefs and plans are diametrically opposite the beliefs and plans they keep insisting that they hold and have practiced for their entire lives.

Let me give a snarky and totally unfair example. This is from the New York Times in 1922 (source):



I won’t say we should always believe that politicians are honest about their beliefs and preferred policies. But I am skeptical when the media claims to have special insight into what they really think.

# Book Review: Unlearn Your Pain

[Content warning: discussion of chronic pain and related conditions, and the debate over whether some of them may be psychological in origin. None of this is medical advice or a recommendation to start or stop any form of therapy. Low confidence in my conclusions here.]

I.

Some of the most interesting lectures in medical training are the ones that start with “Okay, you’re all going to think I’m a quack, but…”

This was how Dr. Howard Schubiner started the lecture he gave at the hospital where I work. Dr. Schubiner isn’t an obvious quack – he’s a professor of medicine at the local university, directs a clinic at a reputable hospital nearby, and is on the editorial boards of a bunch of medical journals. And although his lecture raised what we will generously call a few red flags, there was also just enough interesting stuff there that I couldn’t resist buying his book Unlearn Your Pain to learn more.

Dr. Schubiner’s specialty is psychosomatic complaints – bodily symptoms that don’t come from any obvious disease and seem to reflect psychological stress. Everyone agrees that this category exists. Most doctors have stories about conversion disorder – usually patients who become “paralyzed” in previously healthy limbs after some life crisis. One of my medical school professors had a pretty good diagnostic test for this – feign a punch at the patient’s face, really quickly, without warning her. If she instinctively uses her “paralyzed” limb to block it, it’s conversion disorder. The same sort of thing works for pseudoseizures – apparent seizures not associated with objective seizure EEG activity. There’s a legend about a neurologist telling a medical student that a certain patient’s fit was a pseudoseizure, and the patient interrupting his seizure to protest “No it isn’t!”.

Most people who have worked with conversion or pseudoseizure patients don’t doubt their inherent honesty. These patients aren’t faking, per se. Such a person genuinely can’t move their limb, can’t just decide not to have seizures. Often they’re very distressed at what’s happening to them (although sometimes they really aren’t). Psychologists like to say that it’s subconscious – whatever that means. Just like somebody crippled by panic attacks, the symptoms are real and involuntary, but they’re also psychologically produced.

The existence of this category isn’t controversial, but its size definitely is. Some people propose a long list of conditions – fibromyalgia, irritable bowel syndrome, chronic fatigue syndrome, chronic Lyme disease, tension headaches, interstitial cystitis, et cetera – that they think of as mostly or entirely psychosomatic. On the other hand, patients’ rights groups get very upset at claims that their conditions are “all in their head”, accuse doctors of thinking that they’re lazy or making up their symptoms, and pass around stories with titles like RE: RE: RE: FWD: RE: THE MEDICAL PROFESSION about some guy whose doctor dismissed him as making up his symptoms but who was later diagnosed with zebra-itis and cured with an experimental gene therapy treatment.

Dr. Schubiner is a psychosomatic complaint maximalist. He thinks that just about anything that can’t be traced to a well-understood physiological cause is probably psychosomatic – in his language, Mind-Body Syndrome or MBS. He quotes a fascinating theory by Edward Shorter that this all dates back to the invention of the tendon hammer, ie that little thing doctors hit your knee with:

An important advance in medicine was the discovery of deep tendon reflexes. The simple test of striking a tendon with a reflex hammer can quickly distinguish pathological from psychological paralysis. Amazingly, once doctors could do this test, the number of people with this type of conversion disorder decreased substantially, and now the condition is rare. When doctors and the general public come to view a medical condition as psychologically induced, it is less likely to occur..the subconscious mind is unlikely to produce symptoms that will be easily seen as psychological. But since humans continue to experience great stresses and strong emotions, paralysis has been replaced by chronic back pain, fibromyalgia, fatigue, irritable bowel syndrome, and many other symptoms.

When Schubiner talks about fibromyalgia and fatigue, he’s not so far outside (one edge of) respectable medical opinion. But he goes further and lists migraine headaches, heartburn, carpal tunnel syndrome, tinnitis, postural orthostatic tachycardia, repetitive stress injury, and reflex sympathetic dystrophy as likely Mind-Body Syndrome as well. And most explosively, he says the condition explains almost all pain.

Schubiner admits there is such a thing as anatomically-caused non-psychological pain. It tends to be associated with very obvious injuries like dropping an anvil on your foot, and it tends to go away after a couple of weeks at most. Anything more mysterious and chronic – facial pain, TMJ pain, joint pain, abdominal pain, neck pain, shoulder pain, tendonitis, and especially back pain – is probably Mind-Body Syndrome. After a medical workup has failed to reveal obvious cancer or infection, these are almost certainly psychosomatic and continuing to treat them as potentially medical just makes them worse:

When patients with Mind-Body Syndrome are labeled as having degenerative disc disease on the basis of an MRI….symptoms can be exacerbated and patients harmed by medical diagnoses. This occurs because the diagnosis creates fear and the belief that there is something seriously wrong with one’s body. These emotions activate the anterior cingulate cortex, which creates even more pain by ramping up the learned nerve pathways of MBS.

If this were true, it would be really important. Surveys suggest that between 40 million and 100 million Americans have chronic pain; the former study finds 67% of them say their pain is “constantly present” and 50% say it is sometimes “unbearable and excruciating”. The financial cost is between $60 billion and $600 billion per year. I’m not sure what to think about all these estimates that differ by orders of magnitude, but the point is that there’s a “chronic pain epidemic” and it’s really bad. The mainstay of treatment for chronic pain is opioids, and by non-coincidence there’s also an opioid addiction epidemic and an opioid-related death epidemic. To some degree the government can use regulation to trade off pain burden against opiate deaths, but no point at that curve is very palatable and we desperately need some kind of real solution.

Schubiner says he has it. It’s time to admit that all of this pain that’s getting all epidemicky is almost entirely psychosomatic. It might start with a real injury, but after that injury heals the brain “remembers” the relevant pain pathways and exploits them as a way to express psychological stress. He presents some fascinating and delightful evidence for this.

A guy named Harold Schraeder studied prevelance of chronic whiplash in Lithuania, of all things. He found the prevalence was zero. In most Western nations, a certain subset of people who get in car accidents suffer chronic disabling neck pain, presumably related to having their neck get suddenly jerked by the force of the impact. But Schrader found that this never happened in Lithuania, even though they had a lot of accidents and their cars were no safer than ours. Simotas and Shen found that there was zero whiplash in demolition derby drivers, even though they got into crashes all the time and it was basically their job description. Further studies found that accident victims with more neck injury were no more likely to develop whiplash than victims with less neck injury. Perhaps, they argue, chronic whiplash isn’t a bodily injury at all, but a culture-bound syndrome in which people who expect whiplash to exist use its symptom profile as a way of expressing their psychological tension.

Then there’s back pain, one of the most common and disabling types of chronic pain – Medicare back-pain related costs have grown about 3-4x in a decade. Standard medical workup for back pain usually involves getting an x-ray or MRI, finding some problem with the discs in the spine, and treating with painkillers, steroids, or surgery. Schubiner is not convinced. He notes studies that find that radiographic findings of disc degeneration or herniation do not accurately predict future back pain. Yes, most back pain sufferers will have problems visible on MRI, but most perfectly healthy people pulled off the street will also have problems visible on MRI – for example, this study finds that half of all 21-year-olds in Finland have a degenerated disc, and a quarter have a bulging disc. The book quotes an NEJM article as saying that “neither baseline MRIs nor followup MRIs are useful predictors of low back pain”. However, studies (1, 2) find that a patient’s job satisfaction does predict their future back pain. Books and studies called things like Time To Back Off?, Back In Control and Watch Your Back point out that many surgeries and injections for back pain work no better than placebos in controlled experiments, with a review article in the Journal of the American Board of Family Medicine concluding that “prescribing yet more imaging, opioids, injections, and operations is not likely to improve outcomes for patients with chronic back pain” On the other hand, Schofferman et al found that childhood trauma correlates heavily with success of back pain treatment: 95% of patients with a happy childhood got better after back surgery, but only 15% of patients with multiple childhood traumas did.

Based on these studies and others like them, Schubiner concludes that chronic back pain is psychological rather than physiological. He thinks there may have been some original minor injury, of the sort that most people would get over in a couple of weeks. This causes the nerves to “sensitize” – ie the brain is primed to think about and remember this form of pain. Then, when people recall their subconscious tension over childhood trauma and the stresses of life, they express it as back pain through the sensitive nerve pathways.

Extend this model to headaches, irritable bowel, chronic fatigue, and everything else, and you have Dr. Schubiner’s theory of pain.

II.

If all of this pain has a psychological cause, then it should have a psychological solution. But psychological solutions to chronic pain are no more effective than physical ones. For example, Cochrane Review finds that cognitive behavioral therapy for chronic back pain has a moderate short-term effect which fades quickly. There was a similar effect for neck pain which Cochrane found “could not be considered clinically meaningful” and which also faded quickly.

Dr. Schubiner says this is because cognitive behavioral therapy is inappropriate for this condition. It’s caused not by negative thoughts and dysfunctional behaviors, but by unresolved childhood traumas. He recommends a therapy designed to help resolve such traumas called Intensive Short Term Dynamic Psychotherapy.

Freudian therapy (“psychoanalysis”) usually takes several years to get anywhere, and may take ten years or longer to complete. Around the 1960s, some psychiatrists got tired of waiting and invented a high-speed version called “psychodynamic therapy”. Schubiner’s version, which derives from the word of a guy named Dr. Davanloo in Montreal, promises results in as little as four weeks. It involves a lot of stuff, including some kind of silly-sounding things like writing affirmations, finding your acupuncture points, and putting your body in very masculine “power poses” and raising your fists and shouting “I AM GOING TO OVERCOME MY PAIN!”. Dr. Schubiner demonstrated this last in front of us and he did indeed look very masculine and determined; if I were chronic back pain, I would definitely put him on my list of people to avoid.

But the heart of the therapy is a technique for returning to traumatic childhood moments. You try to figure out what your traumatic childhood moments were – for example, maybe your father got drunk and beat you up. So you go back to the incident, either as a solo visualization or in a conversation with a partner. It goes something like this:

Doctor: Tell me what’s happening

Patient: I’m in my childhood home. My father approaches me, beer bottle in hand, looking really angry.

Doctor: How do you feel?

Patient: Really scared.

Doctor: But also?

Patient: Angry.

Doctor: Good! You have every right to! Tell your father that!

Patient: Father, I’m really scared and angry!

Doctor: Now what does your father do?

Patient: He doesn’t care.

Doctor: And what would you like to do now?

Patient: Beat up my father.

Doctor: Then go back into that experience and beat up your father.

Patient: AAAARGH! I HATE YOU SO MUCH, FATHER! YOU RUINED MY CHILDHOOD! GRAAAAAAAAAH! DIE! DIE! DIE! DIE! DIE!

Doctor: How do you feel now?

Patient: My chronic back pain is gone!

There are enough variations on this to make it a four week course, but in Schubiner’s examples (which he takes from real clinical practice), even something as simple as this can be enough to make chronic pain go away near-instantly. He has about a dozen anecdotes from his own practice where this happens. Then the rest of the course is just solidifying that gain and making sure it doesn’t come back.

I talked to a professor of psychoanalysis I work with about this. She says that Davanloo is well within the psychoanalytic mainstream. She says that she herself is not a big fan of his work, because she thinks it’s important to spend those several years unpeeling a patient’s defenses instead of just smashing them with a sledgehammer. But she says chronic pain patients may have unusually strong defenses and that maybe the sledgehammer approach is the right one. So overall she cautiously approves.

On the other hand, my more cynical readers might note that “well within the psychoanalytic mainstream” isn’t exactly equivalent to “definitely not a quack”. Schubiner is aware of this and has tried to get some evidence for his method. Along with a case series, he has published a study on the psychological treatment for fibromyalgia, in which 45% of the intervention group experienced significant pain relief compared to 0% of the controls. He also tested the full version of his therapy in a preliminary trial in which “two-thirds of the patients improved at least 30% in pain”.

So, should we believe him?

III.

I tried to verify some of the claims in this book and discovered things were much more ambiguous than it let on.

The idea of whiplash as psychosocial is still controversial in the literature. It’s true that some studies in Lithuania and Greece show almost no whiplash. But some critics say that these studies lacked enough power to find a difference in whiplash rates among countries even if such a difference existed. There were also extremely weird fluctuations in the data – for example, the same team in the same city doing two studies a few years apart found neck injury rates of 15% versus 47%. Here’s a long and acrimonious debate in a medical journal about this. But interestingly, even the pro-psychosocial side doesn’t seem to want to say there’s no biological component. And such a claim would be difficult to sustain given studies that show significant effects of things like head position during an accident on future whiplash rates. You can find a good summary of some of the points on each side here. Here’s another review by Dr. Arthur Croft, Ph.D., D.C., M.Sc., M.P.H., F.A.C.O., and Emmy award nominee (really) – who says, brutally:

Ferrari, et al., have recently promoted the so-called biopsychosocial model in the context of whiplash, making numerous excursions into the literature in support of it. The lynchpin of their theory relies on two studies conducted in Lithuania which purportedly followed the natural history of late (i.e., chronic) whiplash in a population of persons exposed to rear-impact motor vehicle crashes – the putative injury mechanism for acute whiplash injury. Unfortunately, “fatal” errors in study design in both cases prevented meaningful interpretation of their results, not the least of which was that only a small portion of their cohort actually had an acute whiplash – the necessary precursor for late whiplash. Our post hoc power calculation revealed that their cohort was inadequate to support any of their conclusions.

Unfortunately, many authors – since these flaws were pointed out – have failed to be dissuaded from citing this literature in support of the biopsychosocial theory, particularly those authors from the camp of the nonbelievers. These Lithuanian papers, it should also be noted, stand alone as outliers to more than 50 other published reports of outcome over the past 45 or so years, and arrive at some rather improbable conclusions that almost immediately beg some questions. In the first paper, the results suggested that persons exposed to whiplash mechanisms would have about the same long-term neck pain as age-matched uninjured persons in the population. In the second study, acute whiplash trauma exposure seemed to actually have a protective effect, somehow immunizing these people against future neck pain. Again, these findings would be particularly interesting if the studies had the added virtue of being valid on a statistical and methodological basis.

To these I would add that even if a US-Lithuania whiplash rate difference existed, it wouldn’t necessarily have to be psychological. Some people bring up Americans having bigger cars; other people bring up differences between American and European car seat headrest design, and still others bring up differences in US and Lithuanian diets and lifestyles which might affect pain and healing.

The book’s treatment of back pain also raises some concerns. It is definitely true that the relationship between back pain and radiographic findings is way lower than anybody wants to admit, and that MRI isn’t very useful for diagnosis. That having been said, the relationship is not zero. For example, in this study, patients with the highest level of radiographic degeneration were 4.5x more likely to be in the highest back pain category compared to patients with the lowest level of radiographic degeneration. The correlation is not one, but neither is it zero. Some people with lots of back pain will have no radiographic findings, and some people with lots of radiographic findings will have no back pain, and the relationship is weak enough that using MRIs for diagnosis is heavily discouraged, but in general the two have a positive and significant relationship. See for example here, here, and here. I don’t know to what degree this affects the book’s thesis, but it seemed important to point out.

Nor is there any more clarity about the relationship of back pain to job (dis-) satisfaction. I was able to find three meta-analyses on it. One of these, Linton, said that:

The available literature indicated a clear link between psychological variables and neck and back pain. The prospective studies indicated that psychological variables were related to the onset of pain, and to acute, subacute, and chronic pain. Stress, distress, or anxiety as well as mood and emotions, cognitive functioning, and pain behavior all were found to be significant factors. Personality factors produced mixed results. Although the level of evidence was low, abuse also was found to be a potentially significant factor.

But Hartvingsen et al concluded:

According to recent epidemiological literature we found moderate evidence for no positive association between perception of work, organisational aspects of work, and social support at work and LBP. We found insufficient evidence for an association between stress at work and LBP. Regarding consequences of LBP, there was insufficient evidence for an association between perception of work in relation to consequences of LBP. There was strong evidence for no association between organisational aspects of work and moderate evidence for no association between social support at work and stress at work and consequences of LBP. There were major methodological problems in the majority of studies included in this review and the diversity in methods was considerable. Therefore associations reported may be spurious and should be interpreted with caution.

Finally, Hoogendoorn, Poppel and Bongers, who sound like a band for very young children, are very ambivalent. They do find some effects, but they all give off an air of desperation, eg “Low job control was found to have a statistically significant positive effect on short and long absences due to back pain, except in men in lower grade jobs and women in higher grade jobs, in whom the effect was reversed”. They are very open about this, and conclude:

“Evidence was found for the effect of some of the psychosocial work characteristics, but there is no psychosocial work characteristic for which evidence was found in all reviews…the conclusions drawn in the various reviews appear to be rather heterogenous”.

And, very significantly for our purposes:

“Having evaluated the strength of the evidence for both physical and psychosocial factors as risk factors for back pain, using the same methods, the question arises of whether the findings indicated a difference in the evidence for physical and psychosocial factors. Strong or moderate evidence has been found for heavy physical work, lifting, bending, and twisting, and whole body vibration at work. Unlike the results for psychosocial factors, these results were rather insensitive to slight changes in the assessment of the findings and the methodologic quality of the studies and in agreement with the results of previous reviews on physical load. This indicates that the body of evidence supporting the role of these physical load factors as risk factors for back pain is somewhat more consistent than that for the psychosocial factors)

The consensus in pain medicine is that pain depends on both psychological and physical factors working together. Schubiner is trying to shift that consensus to say pain is almost entirely psychological and based mainly on childhood trauma. But the studies, while not ruling out a psychological cause, are very emphatic that physical causes definitely matter. And even the papers supporting psychological causes say that, among all such causes, there is unusually little evidence for childhood abuse as a factor.

But the part that bothered me most was the use of Schofferman’s study showing that childhood trauma predicted back surgery success rate (I should note that he doesn’t cite this explicitly in the book, but he implicitly works off it, and he discussed it explicitly during the lecture). This was a surprising study that cried out for replication – and which was in fact re-tested in 2002 on a larger sample by Nickel, Egle, and Hardt. They were unable to replicate the findings. Chronic back pain patients, surgery-failing and otherwise, were no more likely to have childhood trauma than anybody else. This bothered me because Schubiner played up Schofferman’s 1991 study that supported his hypothesis without even mentioning this one. People have a right to present their case the way they want, but when someone clearly ignores better and more recent evidence, it makes me a little more skeptical of everything they say.

IV.

What about the psychiatric part of Unlearn Your Pain‘s program?

The psychodynamic therapy literature is even more of a mess than the back pain literature. I’ve been there before and don’t want to go back. You can read Jonathan Shedler and Jared DeFife in support and James Coyne and Michael Anestis in opposition. I find myself more sympathetic to the “doesn’t work very well” camp, but the field is muddy enough, and my biases against it strong enough, that I place little confidence in that judgment.

So let me try to cut through all of this with my favorite weapon for these kinds of things: behavioral genetics. Of the five behavioral genetics studies on back pain I could find, four (1, 2, 3, 4) found no shared environmental effect on back pain, with only one dissenter. This is in common with a large literature finding little shared environmental effect on a host of psychological problems including depression, anxiety, and bipolar disorder – and indeed, it would be very strange if chronic pain were more related to childhood experiences than those were.

Psychiatry tried really hard to give the “childhood trauma causes everything” thesis a go for fifty-something years. Sure enough, psychiatrists found loads of childhood trauma, because, much as pretty much everybody will have something weird with the discs in their backs that can be detected on MRI, pretty much everybody will have something weird with their childhood that can be detected with psychotherapy. Using the kabbalistic method, you can always find suspicious coincidences linking their childhood trauma with their current pain. Schubiner writes – as far as I can tell, 100% seriously – that:

When someone develops a pain in the buttocks, there may be someone in their lives who is ‘a pain in the butt’.” Someone who develops difficulty swallowing may be reacting to a situation in life that is ‘hard to swallow’. I evaluated a woman with pain in the bottom of her feet. While waiting in line one day, she realized there was a situation in her life that she ‘just couldn’t stand anymore’

I want you to appreciate how much willpower I’m showing here. There is form of psychiatry based around corny puns, and yet instead of emailing these people my resume immediately I’m trying to maintain a cautious skepticism.

And when I do, I just can’t believe it. The early psychoanalysts weren’t doing science, they were taking Sofer’s Law and running with it. Eventually we realized that talking about childhood traumas wasn’t predictive, wasn’t especially curative as per rigorous studies, and we moved on.

There’s a lot of controversy around this decision, but I think behavioral genetics has made the childhood-trauma side increasingly untenable. Assuming twin studies aren’t entirely fatally flawed – something thousands of people have looked for and nobody has found – childhood shared environment, which presumably includes things like abusive parents, just doesn’t affect adult outcomes very much. I can’t see a way to reconcile that with psychoanalytic theories and I don’t think we should keep trying.

I don’t deny that there are a lot of suspicious coincidences. But I think if we look harder, we can find that those suspicious coincidences all have more reasonable explanations. Like, yes, people with a lot of psychological problems tend to have a lot of back pain. But again, when you do twin studies:

On initial analysis considering the participants as individuals, rather than twins — and therefore not accounting for genetic and familial factors — the odds of having back pain were about 1.6 higher for those with symptoms of depression and anxiety.

On further analysis of monozygotic twins — who are genetically identical — the association between symptoms of depression and low back pain disappeared. This suggested that the strong association found in non-identical twins resulted from the “confounding” effects of common genetic factors influencing both conditions. For example, genes affecting levels of neurotransmitters such as serotonin and norepinephrine might affect the risk of both conditions.

Previous studies have shown a “consistent relationship” between back pain and depression — a combination that may complicate diagnosis and treatment. However, the nature of the association remains unclear. The new study is the first to examine the relationship between depression and low back pain using twin data to control for genetic and familial factors.

When you control for genetics, WHICH YOU SHOULD ALWAYS DO AND I AM SO SERIOUS ABOUT THIS, this explains the entire psychological problem/back pain link. Combined with the previous twin studies showing no effect of childhood environment, this is a very strong challenge that theories claiming a psychogenic origin of back pain based in life events will have trouble surviving.

V.

So in the end, what do we make of chronic pain?

Many, many, many people report using the techniques in Unlearn Your Pain (or the closely related techniques of Dr. John Sarno) and having good success. I don’t think this is entirely coincidence or bias. But I’m also not willing to entirely buy into this repressed childhood trauma theory.

There are definitely some types of pain which are not related to bodily injury. My best evidence for this, which Dr. Schubiner talks about too, is the people who have pain which is anatomically implausible or “migrating”. By “anatomically implausible” I mean pain that cuts willy-nilly across the body’s neural regions; pain in the distribution of the ulnar nerve may be an ulnar nerve problem, but if it has half the distribution of the ulnar nerve plus half the distribution of the median nerve, while leaving the other half of both distributions pain-free, it’s a little harder to figure out what could be causing it (especially if it’s on both sides equally!) By “migrating”, I mean that somebody has right hand pain, the doctor gives them some kind of treatment, that goes away, the next day they have right foot pain, another treatment, it goes away again, and the next day they have diarrhea. While there are very rare processes that can do something like that, when it goes on long enough that’s good evidence that the pain isn’t anatomical.

But I think the way in which pain isn’t anatomical is more complicated than the simple model that Unlearn Your Pain uses. Instead of the brain “using” pain to express repressed emotions, maybe gating and modulating pain sensations is just really hard.

Let me give an example [trigger warning for inducing mild bodily discomfort]. Right now, you’re suddenly aware of the feeling of your tongue in your mouth. And right now, the top of your head is suddenly really itchy. Also, right now something is wrong with your saliva and you’re swallowing consciously, but it feels awkward and you’re worried something might be wrong with your throat.

This isn’t because I have magic powers inflicting these things on you, it’s because you’re constantly receiving all sorts of sensations and your brain effectively gates and modulates them. You’ve always got micro-itches and micro-pains going on everywhere – no part of your body is one hundred percent optimal and even if it were there are still variations in neural noise – but your brain usually correctly decides these aren’t worth your time. It’s only when something forces you to focus on them – whether worry about a back injury, or an annoying blogger – that they make it through.

All psychiatric disorders are heavily comorbid. People with one or more of depression, anxiety, OCD, anorexia, autism, gender dysphoria, PTSD, et cetera are many times more likely to have all of the others, and it doesn’t just seem to be in a boring “OCD is depressing and makes me anxious” sort of way. All of this seems to relate to a general factor of neural messed-up-ness. It wouldn’t be surprising if this correlated with some kind of messed-up-ness in the neural systems that are supposed to process and gate pain.

Remember also that stress can cause relapses of many very biological and serious diseases like ulcerative colitis, multiple sclerosis, or epilepsy. Also, inflammation seems to be a shared and complicated factor between various bodily illnesses, stress, and depression. So for stress to cause a “relapse” of chronic pain, all we’d need is for it to put extra pressure – whether through inflammation or some other method – on an already slightly messed-up pain-gating system in the brain. And then there’s muscular tension, which I inexcusably forgot to mention until now but which is also a relevant system by which stress affects chronic pain.

We know that pain is very sensitive to the placebo effect – I’m generally a placebo effect skeptic, but even arch-skeptics Hróbjartsson and Gøtzsche agree that pain is one of the few places where the placebo effect really dominates. This is why we so often see faith healers and saints and miracle water from Lourdes treating pain so effectively – at least briefly. It’s why homeopathic treatment for pain shows such an amazingly good effect size.

And, I will say cynically, it’s why so many people have reported (genuine) success from Unlearn Your Pain and related programs. It’s why Schubiner writes:

I believe that each and every person with Mind Body Syndrome can get better because it is possible to overcome MBS by using this program. Those people who are unable to accept that their symptoms are due to MBS, or who do not develop positive expectations of relief, or who are unable to believe that they can make changes in their health and in their lives are the people who are less likely to improve.

Part of me wants to say that we have a word for medical treatments that only work if you believe that they will, and it rhymes with “gazebo”.

Another part worries this is unfair. If the placebo effect comes from the brain’s ability to gate pain, then saying “You’re not really affecting the brain’s ability to gate pain, it’s just the placebo effect” stops making sense. It’s not that it doesn’t work and is just placebo – it’s that it does work, via placebo.

There’s something to be said for glorifying in the placebo effect, laying it on as hard as possible, putting on the fanciest robe and wizard hat you can find and saying “I CAST YOU OUT BY THE POWER OF PLACEBO, GO FORTH AND SIN NO MORE!” I think in some ways there can be better and worse placebo therapies just as there can be better and worse real therapies, placebo therapies that activate the placebo effect only a little and don’t help much, and placebo therapies that activate the placebo effect really strongly and use it to work miracles. Maybe we should give more status to the best placebo therapies, to view them as highly perfected works of the placebomantic arts in the same way that powerful medications are triumphs of psychopharmacology. I think psychodynamic therapy and everything descended from it would have a high place in that pantheon.

In that sense, I think Unlearn Your Pain might be a useful book. I think that even if I accept what I consider the consensus theory of chronic pain – genuine (if small) lingering injuries (or nerve sensitization from such) interacting with a poorly-wired pain gating system in the brain which is highly susceptible to placebo effects – Unlearn Your Pain remains a useful book, as the distilled wisdom of many years of work trying to activate those effects as strongly as possible.

Another possibility is that the active ingredient isn’t the intensive psychotherapy, it’s the belief that the pain is caused by Mind-Body Syndrome. It seems just possible that this belief could break the cognitive loops that seem so relevant in all of these processes.

So I guess I’m in a weird spot in terms of what I think of Unlearn Your Pain.

I think it’s definitely right that a lot of pain has psychosomatic components. I think it probably helps treat psychosomatic pain, maybe really effectively, and partly for the reasons that it thinks it does.

But I’m not convinced by its more sweeping claims that physical injuries play little-to-no role in chronic pain. Along with Schubiner’s talk of nerve sensitization, one can imagine a scenario in which alternatively apparently-healed physical injuries may leave very small irritations on local nerves, and that the degree of irritation a nerve is able to bear without giving you chronic pain is related to your general neural-non-messed-upness and stress level. In such a scenario psychological factors might play a role in gating the pain, or in tensing or releasing muscles around the pain, but would not entirely explain it.

I’m also not convinced by its claims that childhood trauma has any interesting relationship with pain, nor that trauma-related therapy has a unique non-placebo ability to deal with such pain. I think that childhood trauma is overemphasized throughout psychiatry and that this theory of pain represents a step in the wrong direction. If trauma-related therapy works, it works by a nonspecific process of making people feel like they’re doing something useful and taking their attribution for their pain off of bodily processes.

Niels Bohr used to hang a horseshoe above the door to his office, saying “I’m not superstitious, but I hear this works whether you believe in it or not.” Part of me is tempted to recommend Unlearn Your Pain to my patients on the same principle. And if any readers of this blog have chronic pain and want to try to the month-long self-help therapy course in this book, I would be very interested in hearing back from you (please tell me before you start, so that there aren’t response biases). If the $25 price of this book is the difference between someone in that category trying vs. not trying it, I’m happy to send you the book if you agree to get back to me with your results. Contact me at slatestarcodex@gmail.com if you’re interested in this.

# I Wrote A Blog Post, But Did Not Adjust For The Fact That The Title Would Be Too L

I recently got in some fights with psychoanalysts on the importance of parenting. They mentioned that one good test for genuine parent effects – as opposed to genetic effects, stress-related effects, toxin-related effects, et cetera – would be things that seemed to depend more on one parent than the other. In particular, in order to rule out intrauterine factors, we should be looking at effects that depend disproportionately on the father. For example, if young women with distant fathers are uniquely more likely to become lesbians, that would be a pretty convincing demonstration of the importance of parenting.

So I was interested to see a recent study that claimed a good father/son relationship – but not a good mother/son relationship – had a special role in sons’ development. University of Guelph, Parents, Especially Fathers, Play A Key Role In Young Adults’ Health:

The researchers found that young adults who grew up in stable families with quality parental relationships were more likely to have healthy diet, activity and sleep behaviours, and were less likely to be obese.

Surprisingly, they found that when it came to predicting whether a young male will become overweight or obese, the mother-son relationship mattered far less than the relationship between father and son.

“Much of the research examining the influence of parents has typically examined only the mother’s influence or has combined information across parents,” said Prof. Jess Haines, Family Relations and Applied Nutrition, and lead author of the paper.

“Our results underscore the importance of examining the influence fathers have on their children and to develop strategies to help fathers support the development of healthy behaviours among their children.”

Okay. Let’s look at the study. It’s a correlational study of 6000 kids age 14-24. They were asked to rate the quality of their relationship with each parent, then they were tested for various unhealthy behaviors: obesity, eating disorders, fast food intake, soda intake, TV watching, sedentariness, and poor sleep.

Among all participants, better relationships led to less disordered eating, increased physical activity, and better sleep. This was true both for child/mother relationship, child/father relationship, and child/generic-measure-of-family-functioning relationship. So far this isn’t surprising. There was no attempt to control for wealth, class, or anything else, let alone genes. And a lot of these children are still living with their parents, so good parenting is going to be important to them right now (the study didn’t separate children who were still with their parents from adult children who weren’t). No surprise to find an effect here.

Among no participants did better relationships affect soda consumption or screen time, whether it was the child/mother relationship, the child/father relationship, or the child/generic-measure-of-family-functioning relationship. Okay. I guess these are somewhat more neutral things that good parenting doesn’t affect much.

Among female but not male participants, better relationships decrease fast food consumption. This was true both for child/mother relationships, child/father relationships, and child/generic-measure-of-family-functioning relationships (I believe all marriages should be between a man, a woman, and a generic-measure-of-family-functioning). This suggests that maybe parents care more about their daughters eating fast food than their sons – or maybe those daughters themselves care more. In either case, this wouldn’t be too surprising.

What about the blockbuster result that fathers, but not mothers, affect male children’s obesity level?

The odds ratio for obesity with a good mother-son relationship was 1.04, confidence level (0.85, 1.27).

The odds ratio for obesity with a good father-son relationship was 0.80, confidence level (0.66, 0.98).

Okay. You are measuring seven different outcomes on two different genders of child. On thirteen of these tests, results are concordant between fathers and mothers. On one of them, results are discordant, in that with mothers the confidence interval included 1.00, but with fathers the confidence interval merely included 0.98.

You could either conclude that fathers have a unique ability to affect their sons’ (but not their daughters’) level of obesity (but not disordered eating, or fast food eating, or soda drinking, etc). Or you could conclude that if you do enough tests, 5% of the time something will fall just outside a 95% confidence interval.

Let’s see what the study’s Limitations section has to say about this:

We calculated 42 tests and did not adjust for multiple comparisons.

Why would you do this? If NASA preceded their missions with statements like “We are launching a rocket to Jupiter, but we did not adjust for the fact that it is very far away,” we would stop taking them seriously. But for some reason in the social sciences it’s okay?

All right, fine, let’s hear your excuse:

Of these tests, 25 were statistically significant at the 0.05 level, much larger than the 2 we would expect by chance.

This might work for individual results, but it doesn’t work for discordances between results, which is what they’re trying to show.

Suppose I want to prove that a certain medicine only works on people whose names begin with the letter M (and suppose in reality, the drug works on everybody). My experiment has 80% power to detect the drug effect when it works. I do fifty tests on fifty different populations – elderly Latino women, young black men, genderqueer Caucasian neonates, Thai rice farmers, unemployed auto workers, whatever – and divide each of them into a subgroup with M-names and a subgroup with other names. I’m actually simulating this right now in an Excel spreadsheet, and here are my results:

Among non-M-names, 42 of the populations test positive, which is much as expected – the drug works and we have 80% power to show that it does, so we should expect 50\*0.8 = 40 positive results on average. A little random noise brings that to 42.

Among M-names, 43 of the populations test positive, which is also close to 40. So here everything is just as we would expect.

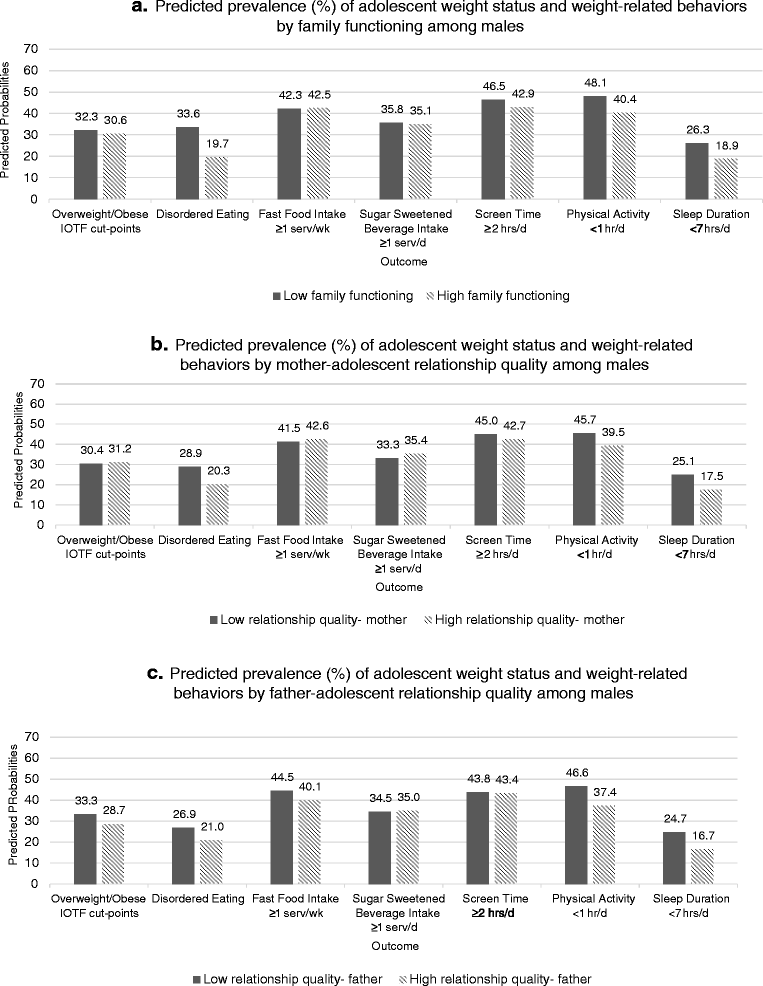
But! In six of the populations, the drug works “differently” for people with M-names and other names. For example, on Test 18 (let’s call this Thai rice farmers), the drug works for rice farmers who have names beginning with M, but doesn’t work for rice farmers who have names beginning with other letters.

So I report this in the literature as “Astounding! Drug works for Thai rice farmers with names beginning with M, but not for Thai rice farmers with names beginning with other letters!” Some annoying person comes back with “but you did a bunch of comparisons and didn’t correct for that”. And I retort “Aha! But actually 85 of my 100 tests came back positive, compared to only 5 that would be expected by pure chance, so clearly there’s something there! There’s an M-name effect after all!”

This is comparing apples to oranges. Yes, you’ve shown that your drug works. But you haven’t come close to showing that it works differently for people whose name begins with M. Your evidence doesn’t even suggest that it does.

But this is what this paper is doing when it says it has evidence that male obesity is affected by the father and not the mother, and claims it doesn’t need to adjust for multiple comparisons.

As Exhibit B, I present the graphs:



I think this is noise.

The paper itself mentions the father-son difference in one paragraph in the Discussion section, but doesn’t even find it worthy of mention in the Conclusion. It’s the press release that plays this up into the major finding of the study. Why?

Because the press release came out three days before Father’s Day.

Look:

In time for Father’s Day, a new University of Guelph study has found that parents, and especially fathers, play a vital role in developing healthy behaviours in young adults and helping to prevent obesity in their children.

I think overly cutesy university PR departments do a lot more damage than is generally realized.

On the other hand, one impressive thing about this paper is its willingness to cite large quantities of stuff. For example, a quote:

Level of bonding or closeness with a parent has also been shown to moderate the association between maternal-BMI and daughter-BMI [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60] and parental and adolescent weight-related behaviors [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61].

I am not going to go through 43 studies to see if any of them are any good, but I guess if there are 43 studies claiming these sorts of parental effects I should be a little more humble.

So: does anyone know of any good studies showing gender-specific-parent effects on a child that don’t seem obviously related to intrauterine or Y-chromosomal factors?

# 3/4

Related: 1/4, 2/4

[Content warning: psychiatry, suicide. Note that all stories involving patients are mixtures of several different people which have been obfuscated and changed around in order to protect confidentiality. The ethical standard I have heard in this situation is “must be so well disguised that the patient would not recognize himself if he read it” and I have tried to meet that standard – which means that these capture the spirit of situations only. The same is true of some of the other stories here, just in case. Please do not link.]

I.

I’m back at Our Lady Of An Undisclosed Location hospital now as a final-year resident. You wouldn’t think a year would make so much difference, but it does.

Identifying residents by their year is easy. The first-years walk around, deer-in-the-headlights look to them, impossible to confuse with anybody except maybe a patient having a panic attack. The middle-year residents are a little more confident. And then the final year residents, leading teams, putting out fires, taking attendings’ abuse in stride.

(True story – last week an attending yelled at me for not knowing some minor detail about uraemic encephalopathy. Later I couldn’t find the detail he’d mentioned, so I asked for a reference, and he said it had been discovered by one of his friends at the big university hospital where he used to work, but the friend had died before he could publish his findings. I think the attending realized as he was talking that it might have been unreasonable to expect me to know a fact whose discoverer took it to the grave with him, but he didn’t apologize.)

It’s only sort of a facade. 99% of things that happen in a hospital are the same things that happened yesterday and the day before, so if you hang around long enough you can learn what to do, or at least which consultant you can call to make it not your problem anymore. On the other hand, Actual Pathology is still a gigantic mystery. I’m not sure this ever changes. One in every X patients with symptoms won’t have any of the things that could possibly be causing those symptoms, won’t respond to any of the treatments that are supposed to cure those symptoms, and you’ll still have family members and hospital administrators demanding that you fix it right now (and in psychiatry, X is probably a single digit number). All you can do is keep up the facade, put your skill at taking attendings’ abuse in stride to good use, and start learning necromancy so you can summon the one big university hospital researcher who studied it but never got a chance to publish their findings.

II.

Two of the most important things I learned during my third year were “Tell me more” and “[awkward silence]”.

“Tell me more,” works for every situation. Part of the problem with psychotherapy is that you’re always expected to have something to say. As a last resort, that thing is “Tell me more”. It sounds like you’re interested. It sounds like you care. And if you’re very lucky, maybe the patient will actually tell you something more, as opposed to their usual plan to stonewall you and hide all possibly useful information.

I saw something on Tumblr the other day which, despite being about a 9-1-1 operator, perfectly sums up being a doctor too:

my bf has many interesting stories and observations from his new job as a 911 operator

my favorite is how meandering people are, even in the midst of a terrible emergency

they respond to “what is the emergency” with “well, the thing is, four weeks ago–”

and then he’s like “WHAT IS THE EMERGENCY RIGHT NOW”

and they’re like “so what happened this morning was, i said to my wife, i said–”

“WHAT IS CURRENTLY HAPPENING AT THIS MOMENT”

“oh i’m having a heart attack”

And:

my second favorite is how specific he has to get sometimes

like, “what is your emergency?”

“i’m sitting in a pool of blood.”

“… is it… your blood?”

“yes i think so”

“do you know where it’s coming from?”

“probably the stab wound”

“have you been stabbed?”

“oh yah definitely”

Psychiatry is like this, except it’s all very vague, and your patients are really suggestable, and people are always afraid that if you just ask specific questions like “Are you depressed?” then they’ll say yes to make you happy and won’t talk about how the real problem is their anxiety or something. So instead, the patient says something like “I’m sitting in a pool of blood”, and I say “Tell me more…”. They say “Well, it’s my blood.” I say “Tell me more…”. After repeating this process a couple of times, we finally get to the stabbing, and the patient doesn’t feel like I railroaded over their chance to tell their story.

Or it helps you figure out what’s important to the patient. If someone said “I hate my husband so much,” my natural instinct might be to ask “Why?”. But maybe why isn’t the question the patient cares about. Maybe what she really wants to talk about is how guilty she feels about hating their husband, and if I asked her why then we’d get on a tangent about what the husband is doing that never addresses her real problem. Maybe she’s agonizing every moment about whether or not to divorce him, and losing sleep over it, and coming to me for a sleeping pill. Maybe she’s just hatched a plan to kill him and wants to check it over with me to see if I can find any flaws. In any case I should probably figure out why they hate him eventually, but if their real issue is whether or not I approve of their murder plot then we should probably get to that first.

So instead, it’s “I hate my husband so much.” “Tell me more.”

“I’m feeling depressed.” “Tell me more.”

“Sometimes I think life isn’t worth living.” “Tell me more.”

“Listen, if you don’t give me a prescription for Adderall right now I swear to God that I will stab you right here in this office!” “Tell me more.”

This has seeped into my personal life. I was on a date with a girl earlier this year, and whenever she started telling me about her life I would just say “Tell me more”, and it worked.

And then there’s [awkward silence]. I learned this one from the psychoanalysts. Nobody likes an awkward silence. If a patient tells you something, and you are awkwardly silent, then the patient will rush to fill the awkward silence with whatever they can think of, which will probably be whatever they were holding back the first time they started talking. You won’t believe how well this one works until you try it. Just stay silent long enough, and the other person will tell you everything. It’s better than waterboarding.

The only problem is when two psychiatrists meet. One of my attendings tried to [awkward silence] me at the same time I was trying to [awkward silence] him, and we ended up just staring at each other for five minutes until finally I broke down laughing.

“I see you find something funny,” he said. “Tell me more.”

III.

If the patients are cryptic, the doctors are even worse. In a worst case scenario, I’ll be filling in for another doctor – this happens all the time at free clinics, but it happens at least a little wherever there are doctors who go on vacation. The documentation will be obscure or missing. The patient’s family is out of contact range. My only information will be the patient in front of me, whose information-transmitting ability is on par with that person from the Tumblr post who took four tries to mention that they’d been stabbed.

So imagine this – a guy from out of state moves in, comes to me without any documentation, and says in a monotone that his only problem is feeling “weird”. All my “tell me mores” and [awkward silences] fail to get him to explain further. I look at his medication list, and he is on a cocktail of supramaximal doses of really old-school antipsychotics that I could not imagine giving anybody unless they were an axe murderer who had killed their last three psychiatrists and I wanted to cool their metaphorical brain temperature to the level of winter on Pluto. Sure enough, the guy is stiff, displays no emotions, and his only hobby is staring at the wall – all exactly what you would expect of somebody who is super-drugged on all of the strongest chemicals known to mankind. I ask him if maybe he’s schizophrenic, or bipolar, or something. He says no, he just feels “weird”.

I know that if I don’t change the medication, he will probably be a zombie like this until such time as somebody else does change it, which may be never. But if I do change the medication…well, there must be some reason somebody put him on that, and the idea of somebody who needed that much medication not being on it is too horrible to imagine. Also, I’m only seeing him once, and then he gets transferred to someone else. What do I do?

The maxim is “do what lets you sleep at night”, so I punted. I kept him on his medications and turned him over to the next guy. I just hope the next guy gets my documentation instead of thinking “Dr. Alexander kept him on all this medication…I wonder what he knows that I don’t.”

IV.

“Instead of putting patients on these toxic medications, why don’t you just give them therapy?”

Sometimes I worry I might be the worst person in the world to do psychotherapy. My coping strategy is to not talk about or react to my emotions and wait for them to go away. This usually works. I know this is exactly the opposite of what psychotherapy is supposed to teach, and all I can say is that it works for me and I seem to be pretty psychologically healthy and maybe I am just a mutant.

My relationship strategy is the same. Date really low-conflict, low-drama, agreeable people. If we have a conflict anyway, then agree to disagree and wait for the problem to go away. Apparently this is terrible, and maybe this is why my only really serious relationship only lasted a year or two, but it leaves me with something of an understanding deficit for the people who want to replay every single argument they’ve ever had with their spouse and figure out exactly what it means about their mental state.

Heck, even polyamory is like this. I can’t tell you how many patients I’ve had come in because their partner is cheating on them, or they worry their partner is cheating on them, or they’re cheating on their partner, or their partner worries they’re cheating on them, or something, and my natural instinct is to just say “Have you considered not worrying about it?” and as usual my natural instinct is terrible. So instead I just say “Tell me more…” and listen to them describe how the possibility of their girlfriend cheating is rending their heart in two.

This is even worse in any form of therapy based around investigating childhood traumas. Look, I’m sorry you didn’t like your mother, but have you read The Nurture Assumption? But of course I can’t say that. I just have to play along. And then somebody expects me to come up with something to heal the maternal trauma that I’m not even sure people really have, and then if I do come up with something it feels like a clever fake.

Cognitive behavioral therapy is a little better, because it tends to be pretty common sense techniques that any reasonable person would agree with. The problem is, it’s pretty common sense techniques that any reasonable person would agree with. I think that I and most of my friends would respond to the average CBT session with a sort of anger at being condescended at, combined with annoyance at the therapist for wasting our time with obvious things. “My job sucks”. “Well, have you considered making a list of good and bad things about your job?” “Yes, that was the process by which I determined it sucks. How much am I paying you again?”

Most of the time I do therapy, I feel cringeworthy, unnatural. I feel like a fraud, even when (according to the supervisors watching me) I’m doing it exactly right. I feel like I’m responding to people in fake, silly ways, like they’re coming to me with problems from the depth of their being and I’m giving them facile non-answers. It doesn’t even help that most of them get better anyway. In a way, that just makes it worse. How dare you get better after me telling you stupid things I feel embarrassed to say? That’s just going to encourage people to make me keep doing that!

V.

I nevertheless hold a special place of annoyance in my heart for psychoanalysis/psychodynamic therapy.

The attending who started my training in psychodynamic therapy (I got a new psychodynamic supervisor recently who I don’t know too well yet) was an elderly doctor in an office attached to the clinic, full of creepy modern art statues. He is convinced that patients’ lives revolve around their therapy and their therapists. I know that in moderation this is the idea of “transference”, a genuine and important tenet of the therapy style. My attending does not do it in moderation.

My patient will say something like “My best friend moved away and now I am sad”, I will think “That sounds straightforward, better bring it up to my attending and see how he wants me to deal with this.” My attending will invariably say “What your patient means is that he’s afraid of losing you, his therapist.”

I will say “No, I’m pretty sure he actually lost his best friend. He told me all about how they’d been together since middle school, but now he moved away to take a job in Texas, and then he broke down crying.”

Then my attending will get really angry and tell me that if I’m just going to take everything my patient tells me exactly literally, then I shouldn’t be in psychiatry, because a monkey could listen to a patient say he was sad about losing his best friend and conclude he was sad about losing his best friend, and my duty as a trained professional is to be able to see beneath that to the true thought which my patient is trying to express. Which is always, 100% of the time, about how much the patient cares about psychodynamic therapy and wants to continue doing it.

Even worse, he wants me to do this to the patient. When the patient says “I’m really upset about losing my best friend”, I’m supposed to answer “Are you sure this isn’t about how you’re worried I’m sort of like a friend to you and one day you’ll lose me?” If talking about relationships and cognitive therapy makes me cringe, this super quadruple makes me cringe.

Still, I have to do it, because my attending grades me and if I don’t pass psychodynamic therapy I don’t get to graduate. So I do it, and then my attending declares he was right all along based on extremely strained interpretations of whatever happens next. Like, if the patient misses their next appointment, he’ll say “I see your patient missed their next appointment. That means they’re having a defensive reaction to the fact that you called them out on their being afraid of you leaving them. And to think that you told me you weren’t sure that was true! This just shows how much you still have to learn about psychodynamics. I certainly hope that after this you won’t keep questioning me every time I try to help you.”

But when I leave for good, I’m getting him a present, and it’s going to be a copy of The Nurture Assumption. Heck, maybe I’ll give that to all the psychoanalysts I know.

VI.

It’s kind of morbid to feel smug about your patients not attempting suicide, but I guess I am a kind of morbid person.

The doctor down the hall from me had one of his patients attempt suicide in October. Then another doctor I knew had two of his patients attempt suicide in the same week in January. And I was really sympathetic and tried to comfort them, but I also had a part of my mind thinking “Hey, I haven’t had any of my patients attempt suicide yet, this is pretty good.”

March. April. May. My coworkers told me their stories, but I kept my secret morbid goal – I was going to go the entire year without any of my patients trying to kill themselves. I mean, on one hand this sounds like a pretty minimal standard. On the other, when you’re taking care of like a hundred mentally ill people, many of whom have really bad depression and a history of past suicide attempts, it’s not exactly trivial.

I got the call just a few weeks ago. The patient was a former heroin addict who had been clean for a long time. He slipped, took heroin, felt terrible, and stabbed himself in the heart.

Luckily the heart is a little to the right of where most people think it is. Stabbing yourself in the lung isn’t great either, but he was a young healthy man and he could take it. He went to the hospital, they patched it up a little, and he was fine. He said it was the best thing that had ever happened to him and now he knew how low he could get and he was going to stay clean forever and today was the first day of the rest of his life.

A lot of things in psychiatry are reverse lotteries. In the regular lottery, you pay a constant small cost for the possibility of a stupendous benefit. In the reverse lottery, you get a constant small benefit at the risk of a stupendous cost. Lots of things are like this. If you give someone a powerful medication, then they’ll definitely recover, but there’s a risk you’ll have a catastrophic side effect. If you let a severely ill patient leave your office when they promise they’re okay, then you definitely save them the trauma of an involuntary hospitalization, but there’s a risk they’ll do something disastrous. If you don’t check someone’s vitals every time you see them then it definitely makes the appointment quicker and smoother, but there’s a risk you’ll miss something really bad.

It’s really easy to fall into playing reverse lotteries. I think almost everybody does it to a degree. The usual pattern is to play some of them tentatively, do more and more of them as you reap the benefits and nothing goes wrong, then boom, close call, and you resolve never to do anything like that again and you’re going to do a full half-hour neurological examination on everybody who comes into your office including random passers-by who just want to use the bathroom.

After my patient stabbed himself I spent a week totally neurotic, looking over every aspect of his case – could I have checked up on his Narcotics Anonymous meeting attendance more frequently? Maybe if I’d given him a long lecture every appointment about how heroin was definitely still bad, that would have changed something? Maybe if I hadn’t forgotten to check his blood pressure that one time…? In the end, I decided I had done a pretty okay job on that case – which just made me more acutely aware of all of the reverse lotteries I was playing on everybody else. Now I’m a little bit paranoid. Maybe that’s temporary. Maybe it’s permanent. I don’t know. The DSM-V says you have to have it six months before you can give yourself a schizophrenia diagnosis, so there’s that.

I am getting good at dealing with annoying attendings, meandering patients, unreasonable requests, and silly bureaucracy. Actual Pathology remains scary, mysterious, and really hard to predict. Hopefully that’s what fourth year is for.

# Pushing And Pulling Goals

This is a distinction I’ve always found helpful.

A pulling goal is when you want to achieve something, so you come up with a plan and a structure. For example, you want to cure cancer, so you become a biologist and set up a lab and do cancer research. Or you want to get rich, so you go to business school and send out your resume.

A pushing goal is when you have a plan and a structure, and you’re trying to figure out what to do with it. For example, you’re studying biology in college, your professor says you need to do a research project to graduate, and so you start looking for research to do. You already know the plan – you’re going to get books, maybe use a lab, do biology-ish things, and end up with a finished report which is twenty pages double-spaced. All you need to figure out is what you’re going to select as the nominal point of the activity. There’s something perversely backwards about this – most people would expect that the point of a research project is to research some topic in particular. But from your perspective the actual subject you’re researching is almost beside the point. The point is to have a twenty page double-spaced report on something.

School and business are obvious ways to end up with pushing goals, but not every pushing goal is about satisfying somebody else’s requirements. I remember in college some friends set up an Atheist Club. There was a Christian Club, and a Buddhist Club, so why shouldn’t the atheists get a club too? So they wrote the charter, they set a meeting time, and then we realized none of us knew what exactly the Atheist Club was supposed to do. The Christian Club prayed and did Bible study; the Buddhist club meditated, the atheist club…sat around and tried to brainstorm Atheist Club activities. Occasionally we came up with some, like watching movies relevant to atheism, or having speakers come in and talk about how creationism was really bad. But we weren’t doing this because we really wanted to watch movies relevant to atheism, or because we were interested in what speakers had to say about creationism. We were doing this because we’d started an Atheist Club and now we had to come up with a purpose for it.

Sometimes on Reddit’s /r/writing I see people asking “How do you come up with ideas for things to write about?” and I feel a sort of horror. So you want to write a novel, but…you don’t have anything to write about? And you just sit there thinking “Maybe it should be about romance…no, war…no, the ennui of the working classes…or maybe hobbits.” I can understand this in theory – you want to be A Writer – but it still weirds me out.

You may have noticed I don’t really like pushing goals. Part of it is an irrational intuition that they’re dishonest in some way that’s hard to explain. It usually ends up with me trying to figure out what to do my biology research project on, and I think “well, I can’t think of anything I really want to research, so maybe I should just do whatever is easiest”. But if I do whatever is easiest, I feel really bad, and worry maybe I have some kind of obligation to research something important that I care about. So I get my brain tangled up trying to figure out how much easiness I can get away with, then feeling bad for asking the question, then trying to come up with something important I honestly want to do, which doesn’t exist since I wasn’t doing a biology research project the month before my professor assigned it to me and so clearly I am only doing it to satisfy the requirement.

Another part of it is that it’s often a sign something has gone wrong somewhere. In the example of the Atheist Club, that thing might have been starting the club in the first place. But assuming that we genuinely want to start the club, then the presence of a pushing goal means we don’t understand why we wanted to start the club. If we wanted to start it because we wanted to hang out with other atheists, then that offers a blueprint for a solution to the problem – instead of planning all these movies and speakers, we should just hang out. If we did it because we thought it was important for atheism to be more visible on campus, then again, that offers a blueprint for a solution – spend our sessions trying to improve atheism’s campus visibility. If we just sit there saying “I guess we have an Atheist Club now, better think of something to do at meetings”, then it seems like something important hasn’t been fully examined.

The third part of it is that things done for push goals usually suck. Maybe this isn’t a human universal – my go-to example is Edgar Allen Poe deciding to write a creepy poem and coming up with The Raven from first principles – but it’s true for me. If I have to write a report on a topic I don’t care about, then even if I’m really trying to do a good job, it’s not going to be as good as something I actually want to write about. Sometimes I try to solve this by making lists of things I want to pull, then using them when the appropriate pushing situation comes up. For example, when I knew I would be assigned research projects and writing assignments on a regular basis, whenever I thought of something I wanted to research or write, I wrote it down, then consulted the list when I needed it. I have a similar list of interesting things to work into stories. This is one reason I’m not interested in journalism – I worry that if I have to produce specific articles on specific things within a time frame, they’ll probably suck.

# Things Probably Matter

A while back when I wrote about how China’s economic development might not have increased happiness there much, Scott Sumner wrote a really interesting response, Does Anything Matter?

He points out that it’s too easy to make this about exotic far-off Chinese. Much the same phenomenon occurs closer to home:

If nothing really matters in China, if even overcoming horrible problems doesn’t make the Chinese better off, then what’s the use of favoring or opposing any public policy? After all, America also shows no rise in average happiness since the 1950s, despite:

1. A big rise in real wages.  
2. Environmental clean-up (including lead–does Flint matter?)  
3. Civil rights for African Americans  
4. Feminism, gay rights.  
5. Dentists now use Novocain (My childhood cavities were filled without it)  
6. 1000 channels in glorious widescreen HDTV  
7. Blogs

I could go on and on. And yet, if the surveys are to be believed, we are no happier than before. And I think it’s very possible that we are in fact no happier than before, that there’s a sort of law of the conservation of happiness. As I walk down the street, grown-ups don’t seem any happier than the grown-ups I recall as a kid. Does that mean that all of those wonderful societal achievements since 1950 were absolutely worthless?

But there are exceptions. I recall reading that surveys showed a rise in European happiness in the decades after WWII, and Scott reports that happiness is currently very low in Iraq and Syria. So that suggests that current conditions do matter.

The following hypothesis will sound really ad hoc, but matches the way a lot of people I know talk about their lives. Suppose people’s happiness is normally calibrated around the sort of lifestyle that they view as “normal.” As America got richer after 1950, it all seemed very normal, so people didn’t report more happiness. Ditto for China during the boom years. Everyone around you was also doing better, so you started thinking about how you were doing relative to your neighbors. But Germans walking through the rubble of Berlin in 1948, or Syrians doing so today in Aleppo, do see their plight as abnormal. They remember a time before the war. So they report less happiness than during normal times.

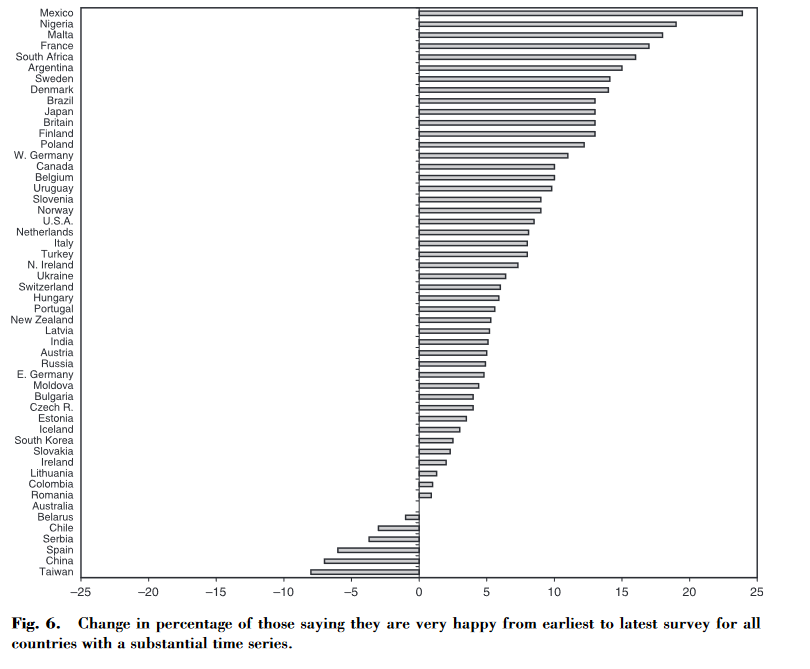
The obvious retort is – modern Chinese grew up when China was very poor. Why didn’t they calibrate themselves to poverty, such that sudden wealth seems good? What’s the difference between a Chinese person going from poverty to wealth, versus a Syrian going from stability to chaos? Might it be a shorter time course? A sudden shock is noticeable, a gradual thirty-year improvement in living standards isn’t?

Probably not. There seem to be a lot of cases where happiness of large groups does change gradually in response to social trends less dramatic than a world war.

First, consider African-Americans. The New York Times calls the increase in black happiness over the past forty years “one of the most dramatic gains in the happiness data that you’ll see”. This is not just about poverty; in 1970, blacks who earned more than 75% of whites were only in the tenth percentile of white happiness. Today, those blacks would be in the fiftieth percentile; they’re still doing worse than would be expected based on income, but not nearly as much worse. This is a very sensible and predictable thing to find. Black people face a lot less racism and discrimination today than in 1970 [citation needed], so assuming that was really unpleasant we shouldn’t be surprised that they’re happier. But notice that this is a time course very similar to the rise of China! It doesn’t look like black people picked a happiness level to calibrate on and then never bothered to adjust. It looks like they adjusted exactly like we would expect them to, even over the course of a multi-decade change.

Second, consider women. In 1970, US women were generally happier than US men. Today, the reverse is true. There seems to be a general pattern around the world of women being happier than men in traditional societies and less happy than men in modern societies (though see Language Log for a contrary perspective). I don’t think of this as a weird paradox. It seems perfectly reasonable to me that having to work outside the home makes people less happy, getting to spend time with their family makes them more happy, and having to work outside the home but also being expected to take care of your family at the same time makes them least happy of all. In any case, the point is that the numbers are changing. Men and women aren’t just fixating on some level of happiness and staying there, they’re altering their happiness level based on real trends, just like African-Americans did (but apparently unlike Chinese).

Third, I was finally able to find a paper that had really good data on change in happiness in different countries, and it supports the idea that happiness can change significantly on a countrywide level.



This is change in happiness in a bunch of countries between about 1990 and 2010 (the years were slightly different in each country). There are other graphs for related concepts like life satisfaction and subjective well-being that look about the same.

The most striking finding is that most countries got happier between those two years – sometimes a lot happier. In Mexico, the percent of people saying they were very happy increased by 25 percentage points!

Just eyeballing the graph, there’s not an obvious relationship between happiness and economic growth – China is still near the bottom like we talked about before, and France – a country that’s been First World since forever – is near the top. Even Japan, which is famous for its decades of stagnation, has done pretty well. But the authors tell us that after doing their statistical analyses, there is a strong relationship with economic growth. Okay, I guess.

They also say there’s a dramatic relationship with freedom and democracy. Mexico, the top country on the graph, went from a relatively closed to a relatively democratic government during this time. South Africa, number five, went from apartheid to no apartheid. Some of the ex-Communist countries like Poland and Ukraine also look pretty good here. On the other hand, other ex-Communist countries like Lithuania and Estonia are near the bottom. I wonder if this has to do with cutoff points – since every country started at a slightly different time, maybe they began sampling Poland during the worst parts of Soviet dictatorship and got Lithuania right in the first euphoria of independence? I don’t know. It all seems very noisy.

They also mention that the United States’ supposedly level happiness is kind of a misunderstanding. People say things like “Happiness in the US has been flat from 1950 to today”, but in fact it declined from 1950 to 1979 and increased from 1980 to today. They attribute this to the 1950s being unusually happy; then the 60s and 70s being unusually conflict-prone, and the Reagan Revolution and Clinton years were back to being optimistic. They don’t have data that stretches too long after that.

(This is pretty neat for Reagan and Clinton. When I die, I’ll consider my life a success if people attribute a spike on national happiness graphs to my influence.)

So apparently population happiness levels do change in response to relevant social changes, even on multi-decade timescales. Which brings us back to asking – what’s up with China?

The graph above shows India as doing okay – not great, but okay. But a similar graph on subjective well-being – which should be another way of looking at the same thing – shows India as doing pretty poorly, right down there with China – even though its GDP per capita quadrupled during the period of study.

I see a lot of conflicting perspectives about whether economic growth increases national happiness. It may, but the effect isn’t as big as you’d expect, and is usually overpowered by other factors. Maybe it isn’t even direct, but has something to do with development increasing democracy, liberalism, rule of law, and stability. China got the development, but its happiness genuinely didn’t increase because of country-specific factors that have something to do with how it developed (inequality? pollution? authoritarianism?).

This matches the race and gender data. Blacks saw a big happiness boost during a time when their feeling of freedom (but not their income) increased relative to whites. Women saw a small happiness drop during a time when their income (but not their feeling of freedom) increased relative to men.

So it looks like happiness can change. It just didn’t change in China over the past thirty years. The apparent paradox of improving economic situation and stable/decreasing happiness is genuinely paradoxical. Intangibles are probably just way more important than money, even amounts of money big enough to raise whole countries out of poverty.

# How The West Was Won

I.

Someone recently linked me to Bryan Caplan’s post A Hardy Weed: How Traditionalists Underestimate Western Civ. He argues that “western civilization”‘s supposed defenders don’t give it enough credit. They’re always worrying about it being threatened by Islam or China or Degeneracy or whatever, but in fact western civilization can not only hold its own against these threats but actively outcompetes them:

The fragility thesis is flat wrong. There is absolutely no reason to think that Western civilization is more fragile than Asian civilization, Islamic civilization, or any other prominent rivals. At minimum, Western civilization can and does perpetuate itself the standard way: sheer conformity and status quo bias.

But saying that Western civilization is no more fragile than other cultures is a gross understatement. The truth is that Western civilization is taking over the globe. In virtually any fair fight, it steadily triumphs. Why? Because, as fans of Western civ ought to know, Western civ is better. Given a choice, young people choose Western consumerism, gender norms, and entertainment. Anti-Western governments from Beijing to Tehran know this this to be true: Without draconian censorship and social regulation, “Westoxification” will win.

A big part of the West’s strength, I hasten to add, is its openness to awesomeness. When it encounters competing cultures, it gleefully identifies competitors’ best traits – then adopts them as its own. By the time Western culture commands the globe, it will have appropriated the best features of Asian and Islamic culture. Even its nominal detractors will be Westernized in all but name. Picture how contemporary Christian fundamentalists’ consumerism and gender roles would have horrified Luther or Calvin. Western civ is a good winner. It doesn’t demand total surrender. It doesn’t make fans of competing cultures formally recant their errors. It just tempts them in a hundred different ways until they tacitly convert.

Traditionalists’ laments for Western civilization deeply puzzle me. Yes, it’s easy to dwell on setbacks. In a world of seven billion people, you can’t expect Western culture to win everywhere everyday. But do traditionalists seriously believe that freshman Western civ classes are the wall standing between us and barbarism? Have they really failed to notice the fact that Western civilization flourishes all over the globe, even when hostile governments fight it tooth and nail? It is time for the friends of Western civilization to learn a lesson from its enemies: Western civ is a hardy weed. Given half a chance, it survives, spreads, and conquers. Peacefully.

I worry that Caplan is eliding the important summoner/demon distinction. This is an easy distinction to miss, since demons often kill their summoners and wear their skin. But in this case, he’s become hopelessly confused without it.

I am pretty sure there was, at one point, such a thing as western civilization. I think it included things like dancing around maypoles and copying Latin manuscripts. At some point Thor might have been involved. That civilization is dead. It summoned an alien entity from beyond the void which devoured its summoner and is proceeding to eat the rest of the world.

An analogy: naturopaths like to use the term “western medicine” to refer to the evidence-based medicine of drugs and surgeries you would get at your local hospital. They contrast this with traditional Chinese medicine and Ayurvedic medicine, which it has somewhat replaced, apparently a symptom of the “westernization” of Chinese and Indian societies.

But “western medicine” is just medicine that works. It happens to be western because the West had a technological head start, and so discovered most of the medicine that works first. But there’s nothing culturally western about it; there’s nothing Christian or Greco-Roman about using penicillin to deal with a bacterial infection. Indeed, “western medicine” replaced the traditional medicine of Europe – Hippocrates’ four humors – before it started threatening the traditional medicines of China or India. So-called “western medicine” is an inhuman perfect construct from beyond the void, summoned by Westerners, which ate traditional Western medicine first and is now proceeding to eat the rest of the world.

“Western culture” is no more related to the geographical west than western medicine. People who complain about western culture taking over their country always manage to bring up Coca-Cola. But in what sense is Coca-Cola culturally western? It’s an Ethiopian bean mixed with a Colombian leaf mixed with carbonated water and lots and lots of sugar. An American was the first person to discover that this combination tasted really good – our technological/economic head start ensured that. But in a world where America never existed, eventually some Japanese or Arabian chemist would have found that sugar-filled fizzy drinks were really tasty. It was a discovery waiting to be plucked out of the void, like penicillin. America summoned it but did not create it. If western medicine is just medicine that works, soda pop is just refreshment that works.

The same is true of more intellectual “products”. Caplan notes that foreigners consume western gender norms, but these certainly aren’t gender norms that would have been recognizable to Cicero, St. Augustine, Henry VIII, or even Voltaire. They’re gender norms that sprung up in the aftermath of the Industrial Revolution and its turbulent intermixing of the domestic and public economies. They arose because they worked. The West was the first region to industrialize and realize those were the gender norms that worked for industrial societies, and as China and Arabia industrialize they’re going to find the same thing.

Caplan writes:

A big part of the West’s strength, I hasten to add, is its openness to awesomeness. When it encounters competing cultures, it gleefully identifies competitors’ best traits – then adopts them as its own. By the time Western culture commands the globe, it will have appropriated the best features of Asian and Islamic culture.

Certainly he’s pointing at a real phenomenon – sushi has spread almost as rapidly as Coke. But in what sense has sushi been “westernized”? Yes, Europe has adopted sushi. But so have China, India, and Africa. Sushi is another refreshment that works, a crack in the narrative that what’s going on is “westernization” in any meaningful sense.

Here’s what I think is going on. Maybe every culture is the gradual accumulation of useful environmental adaptations combined with random memetic drift. But this is usually a slow process with plenty of room for everybody to adjust and local peculiarities to seep in. The Industrial Revolution caused such rapid change that the process become qualitatively different, a frantic search for better adaptations to an environment that was itself changing almost as fast as people could understand it.

The Industrial Revolution also changed the way culture was spatially distributed. When the fastest mode of transportation is the horse, and the postal system is frequently ambushed by Huns, almost all culture is local culture. England develops a culture, France develops a culture, Spain develops a culture. Geographic, language, and political barriers keep these from intermixing too much. Add rapid communication – even at the level of a good postal service – and the equation begins to change. In the 17th century, philosophers were remarking (in Latin, the universal language!) about how Descartes from France had more in common with Leibniz from Germany than either of them did with the average Frenchman or German. Nowadays I certainly have more in common with SSC readers in Finland than I do with my next-door neighbor whom I’ve never met.

Improved trade and communication networks created a rapid flow of ideas from one big commercial center to another. Things that worked – western medicine, Coca-Cola, egalitarian gender norms, sushi – spread along the trade networks and started outcompeting things that didn’t. It happened in the west first, but not in any kind of a black-and-white way. Places were inducted into the universal culture in proportion to their participation in global trade; Shanghai was infected before West Kerry; Dubai is further gone than Alabama. The great financial capitals became a single cultural region in the same way that “England” or “France” had been a cultural region in the olden times, gradually converging on more and more ideas that worked in their new economic situation.

Let me say again that this universal culture, though it started in the West, was western only in the most cosmetic ways. If China or the Caliphate had industrialized first, they would have been the ones who developed it, and it would have been much the same. The new sodas and medicines and gender norms invented in Beijing or Baghdad would have spread throughout the world, and they would have looked very familiar. The best way to industrialize is the best way to industrialize.

II.

Something Caplan was pointing towards but never really said outright: universal culture is by definition the only culture that can survive without censorship.

He writes in his post:

The truth is that Western civilization is taking over the globe. In virtually any fair fight, it steadily triumphs. Why? Because, as fans of Western civ ought to know, Western civ is better. Given a choice, young people choose Western consumerism, gender norms, and entertainment. Anti-Western governments from Beijing to Tehran know this this to be true: Without draconian censorship and social regulation, “Westoxification” will win.

Universal culture is the collection of the most competitive ideas and products. Coca-Cola spreads because it tastes better than whatever people were drinking before. Egalitarian gender norms spread because they’re more popular and likeable than their predecessors. If there was something that outcompeted Coca-Cola, then that would be the official soda of universal culture and Coca-Cola would be consigned to the scrapheap of history.

The only reason universal culture doesn’t outcompete everything else instantly and achieve fixation around the globe is barriers to communication. Some of those barriers are natural – Tibet survived universalization for a long time because nobody could get to it. Sometimes the barrier is time – universal culture can’t assimilate every little hill and valley instantly. Other times there are no natural barriers, and then your choice is to either accept assimilation into universal culture, or put up some form of censorship.

Imagine that Tibet wants to protect its traditional drink of yak’s milk. The Dalai Lama requests that everyone continue to drink yak’s milk. But Coca-Cola tastes much better than yak’s milk, and everyone knows this. So it becomes a coordination problem: even if individual Tibetans would prefer that their neighbors all drink yak’s milk to preserve the culture, they want to drink Coca-Cola. The only way yak’s milk stays popular is if the Dalai Lama bans Coca-Cola from the country.

But westerners aren’t banning yak’s milk to “protect” their cultures. They don’t have to. Universal culture is high-entropy; it’s already in its ground state and will survive and spread without help. All other cultures are low-entropy; they survive only if someone keeps pushing energy into the system to protect them. It could be the Dalai Lama banning Coca-Cola. It could be the Académie Française removing English words from the language. It could be the secret police killing anyone who speaks out against Comrade Stalin. But if you want anything other than universal culture, you better either be surrounded by some very high mountains, or be willing to get your hands dirty.

There’s one more sense in which universal culture is high-entropy; I think it might be the only culture that can really survive high levels of immigration.

I’ve been wondering for a long time – how come groups that want to protect their traditional cultures worry about immigration? After all, San Francisco is frequently said to have a thriving gay culture. There’s a strong Hasidic Jewish culture in New York City. Everyone agrees that the US has something called “black culture”, although there’s debate over exactly what it entails. But only 6% of San Francisco is gay. Only 1% of New Yorkers are Hasidim. Only about 11% of Americans are black. So these groups have all managed to maintain strong cultures while being vastly outnumbered by people who are different from them.

So why is anyone concerned about immigration threatening their culture? Suppose that Tibet was utterly overwhelmed by immigrants, tens of millions of them. No matter how many people you import, Tibetan people couldn’t possibly get more outnumbered in their own country than gays, Hasidim, and blacks already are. But those groups hold on to their cultures just fine. Wouldn’t we expect Tibetans (or Americans, or English people) to do the same?

I’m still not totally sure about the answer to this one, but once again I think it makes more sense when we realize that Tibet is competing not against Western culture, but against universal culture.

And here, universal culture is going to win, simply because it’s designed to deal with diverse multicultural environments. Remember, different strategies can succeed in different equilibria. In a world full of auto-cooperators, defect-bot hits the jackpot. In a world full of tit-for-tat-players, defect-bot crashes and burns. Likewise, in a world where everybody else follows Tibetan culture, Tibetan culture may do very well. In a world where there are lots of different cultures all mixed together, Tibetan culture might not have any idea what to do.

(one more hypothetical, to clarify what I’m talking about – imagine a culture where the color of someone’s clothes tells you a lot of things about them – for example, anyone wearing red is a prostitute. This may work well as long as everyone follows the culture. If you mix it 50-50 with another culture that doesn’t have this norm, then things go downhill quickly; you proposition a lady wearing red, only to get pepper sprayed in the eye. Eventually the first culture gives up and stops trying to communicate messages through clothing color.)

I think universal culture has done a really good job adapting to this through a strategy of social atomization; everybody does their own thing in their own home, and the community exists to protect them and perform some lowest common denominator functions that everyone can agree on. This is a really good way to run a multicultural society without causing any conflict, but it requires a very specific set of cultural norms and social technologies to work properly, and only universal culture has developed these enough to pull it off.

Because universal culture is better at dealing with multicultural societies, the more immigrants there are, the more likely everyone will just default to universal culture in public spaces. And eventually the public space will creep further and further until universal culture becomes the norm.

If you don’t understand the difference between western culture and universal culture, this looks like the immigrants assimilating – “Oh, before these people were Chinese people behaving in their foreign Chinese way, but now they’re Westerners just like us.” Once you make the distinction, it looks like both Chinese people and traditional Americans assimilating into universal culture in order to share a common ground – with this being invisible to people who are already assimilated into universal culture, to whom it just looks “normal”.

III.

I stress these points because the incorrect model of “foreign cultures being Westernized” casts Western culture as the aggressor, whereas the model of “every culture is being universalized” finds Western culture to be as much a victim as anywhere else. Coca-Cola might have replaced traditional yak’s milk in Mongolia, but it also replaced traditional apple cider in America. A Hopi Indian saddened that her children no longer know the old ritual dances differs little from a Southern Baptist incensed that her kids no longer go to church. Universal values have triumphed over both.

Our society is generally in favor of small, far-away, or exotic groups trying to maintain their culture. We think it’s great that the Hopi are trying to get the next generation to participate in the traditional dances. We support the Tibetans’ attempt to maintain their culture in the face of pressure from China. We promote black culture, gay culture, et cetera. We think of it as a tragedy when the dominant culture manages to take over and destroy one of these smaller cultures. For example, when white American educators taught Native American children to identify with white American culture and ignore the old ways, that was inappropriate and in some senses “genocidal” if the aim was to destroy Native Americans as a separate people. We get excited by the story of Bhutan, the tiny Himalayan kingdom trying to preserve its natural and human environment and prevent its own McDonaldization. We tend to be especially upset when the destruction of cultures happens in the context of colonialism, ie a large and powerful country trying to take over and eliminate the culture of a smaller country. Some examples include the English in Ireland, the English in India, the English in Africa, and basically the English anywhere.

One of the most common justifications for colonialism is that a more advanced and enlightened society is taking over an evil and oppressive society. For example, when China invaded Tibet, they said that this was because Tibet was a feudal hellhole where most of the people were living in abject slavery and where people who protested the rule of the lamas were punished by having their eyes gouged out (true!). They declared the anniversary of their conquest “Serfs Emancipation Day” and force the Tibetans to celebrate it every year. They say that anyone who opposes the Chinese, supports the Dalai Lama, or flies the old Tibetan flag is allied with the old feudal lords and wants to celebrate a culture based around serfdom and oppression.

But opponents of colonialism tend to believe that cultures are valuable and need to be protected in and of themselves. This is true even if the culture is very poor, if the culture consists of people who aren’t very well-educated by Western standards, even if they believe in religions that we think are stupid, even if those cultures have unsavory histories, et cetera. We tend to allow such cultures to resist outside influences, and we even celebrate such resistance. If anybody were to say that, for example, Native Americans are poor and ignorant, have a dumb religion with all sorts of unprovable “spirits”, used to be involved in a lot of killing and raiding and slave-taking – and so we need to burn down their culture and raise their children in our own superior culture – that person would be incredibly racist and they would not be worth listening to. We celebrate when cultures choose preservation of their traditional lifestyles over mere economic growth, like Bhutan’s gross national happiness program.

This is true in every case except with the cultures we consider our outgroups – in the US, white Southern fundamentalist Christian Republicans; in the UK, white rural working-class leave voters. In both cases, their ignorance is treated as worthy of mockery, their religion is treated as stupidity and failure to understand science, their poverty makes them “trailer trash”, their rejection of economic-growth-at-all-costs means they are too stupid to understand the stakes, and their desire to protect their obviously inferior culture makes them xenophobic and racist. Although we laugh at the Chinese claim that the only reason a Tibetan could identify with their own culture and want to fly its flag is because they support serfdom and eye-gouging, we solemnly nod along with our own culture’s claim that the only reason a Southerner could identify with their own culture and want to fly its flag is because they support racism and slavery.

(one question I got on the post linked above was why its description of American tribes seemed to fit other countries so well. I think the answer is because most countries’ politics are centered around the conflict between more-universalized and less-universalized segments of the population.)

We could even look at this as a form of colonialism – if Brexit supporters and opponents lived on two different islands and had different colored skin, then people in London saying things like “These people are so butthurt that we’re destroying their so-called ‘culture’, but they’re really just a bunch of ignorant rubes, and they don’t realize they need us elites to keep their country running, so screw them,” would sound a lot more sinister. The insistence that they tolerate unwanted immigration into their lands would look a lot like how China is trying to destroy Tibet by exporting millions of people to it in the hopes they will eventually outnumber the recalcitrant native Tibetans (if you don’t believe me, believe the Dalai Lama, who apparently has the same perspective). The claim that they’re confused bout their own economic self-interest would give way to discussions of Bhutan style “gross national happiness”.

(I get accused of being crypto-conservative around here every so often, but I think I’m just taking my anti-colonialism position to its logical conclusion. A liberal getting upset about how other liberals are treating conservatives, doesn’t become conservative himself, any more than an American getting upset about how other Americans treat Iraqis becomes an Iraqi.)

And I worry that confusing “universal culture” with “Western culture” legitimizes this weird double standard. If universal culture and Western culture are the same thing, then Western culture doesn’t need protection – as Caplan points out, it’s the giant unstoppable wave of progress sweeping over everything else. Or maybe it doesn’t deserve protection – after all, it’s the colonialist ideology that tried to destroy local cultures and set itself up as supreme. If Western culture is already super-strong and has a history of trying to take over everywhere else, then surely advocating “protecting Western culture” must be a code phrase for something more sinister. We can sympathize with foreign cultures like the Tibetans who are actually under threat, but sympathizing with any Western culture in any way would just be legitimizing aggression.

But I would argue that it’s universal culture which is the giant unstoppable wave of progress, and that it was universal culture that was responsible for colonizing other cultures and replacing them with itself. And universal culture’s continuing attempts to subjugate the last unassimilated remnants of traditional western culture are just part of this trend.

IV.

I am mostly just on the side of consistency. After that I have no idea what to do.

One argument is that we should consistently support traditional cultures’ attempts to defend themselves against universal culture. Support the Native Americans’ ability to practice their old ways, support traditional Siberians trying to return to their shamanistic roots, support Australian Aborigines’ rights to continue the old rituals, support Tibetans’ rights to practice Vajrayana Buddhism, and support rural British people trying to protect Ye Olde England from the changes associated with increased immigration. For most people, this would mean extending the compassion that they feel to the Aborigines, peasants, and Tibetans to apply to the British as well.

But another argument is that we should consistently support universal culture’s attempt to impose progress on traditional cultures. Maybe we should tell the Native Americans that if they embraced global capitalism, they could have a tacqueria, sushi restaurant, and kebab place all on the same street in their reservation. Maybe we should tell the Aborigines that modern science says the Dreamtime is a myth they need to stop clinging to dumb disproven ideas. Maybe we should tell the Tibetans that Vajrayana Buddhism is too intolerant of homosexuality. Take our conviction that rural Englanders are just racist and xenophobic and ill-informed, and extend that to everyone else who’s trying to resist a way of life that’s objectively better.

I am sort of torn on this.

On the one hand, universal culture is objectively better. Its science is more correct, its economy will grow faster, its soft drinks are more refreshing, its political systems are (necessarily) freer, and it is (in a certain specific sense) what everybody would select if given a free choice. It also seems morally better. The Tibetans did gouge out the eyes of would-be-runaway serfs. I realize the circularity of saying that universal culture is objectively morally better based on it seeming so to me, a universal culture member – but I am prepared to suspend that paradox in favor of not wanting people’s eyes gouged out for resisting slavery.

On the other hand, I think that “universal culture is what every society would select if given the opportunity” is less of a knock-down point than it would seem. Heroin use is something every society would select if given the opportunity. That is, if nobody placed “censorship” on the spread of heroin, it would rapidly spread from country to country, becoming a major part of that country’s society. Instead, we implement an almost authoritarian level of control on it, because we know that even though it would be very widely adopted, it’s not something that is good for anybody in the long term. An opponent of universal culture could say it has the same property.

Things get even worse when you remember that cultures are multi-agent games and each agent pursuing its own self-interest might be a disaster for the whole. Pollution is a good example of this; if the best car is very polluting, and one car worth of pollution is minimal but many cars’ worth of pollution is toxic, then absent good coordination mechanisms everyone will choose the best car even though everyone would prefer a world where nobody (including them) had the best car. I may have written about this before.

I’m constantly intrigued (though always a little skeptical) by claims that “primitive” cultures live happier and more satisfying lives than our own. I know of several of this type. First, happiness surveys that tend to find Latin American countries doing as well or better than much richer and more advanced European countries. Second, the evidence from the Amish, whose children are allowed to experience the modern culture around them but who usually prefer to stay in Amish society. Third, Axtell’s paper on prisoner exchanges between early US colonists and Native Americans; colonists captured by the natives almost always wanted to stay and live with the natives; natives captured by the colonists never wanted to stay and live with the colonists. Many people have remarked on how more culturally homogenous countries seem happier. Bhutan itself might be evidence here, although I’ve seen wildly different claims on where it falls on happiness surveys. I’ve also talked before about how China’s happiness level stayed stable or even dropped during its period of rapid development.

(on the other hand, there’s also a lot of counterevidence. More democratic countries seem to be happier, and democracies will generally be the low-censorship countries that get more assimilated into universal culture. Free market economies are happier. Some studies say that more liberal countries are happier. And there’s a complicated but positive relationship between national happiness and wealth.)

I also think that it might be reasonable to have continuation of your own culture as a terminal goal, even if you know your culture is “worse” in some way than what would replace it. There’s a transhumanist joke – “Instead of protecting human values, why not reprogram humans to like hydrogen? After all, there’s a lot of hydrogen.” There’s way more hydrogen than beautiful art, or star-crossed romances, or exciting adventures. A human who likes beautiful art, star-crossed romances, and exciting adventures is in some sense “worse” than a human who likes hydrogen, since it would be much harder for her to achieve her goals and she would probably be much less happy. But knowing this does not make me any happier about the idea of being reprogrammed in favor of hydrogen-related goals. My own value system might not be objectively the best, or even very good, but it’s my value system and I want to keep it and you can’t take it away from me. I am an individualist and I think of this on an individual level, but I could also see having this self-preservation-against-optimality urge for my community and its values.

(I’ve sometimes heard this called Lovecraftian parochialism, based on H.P. Lovecraft’s philosophy that the universe is vast and incomprehensible and anti-human, and you’ve got to draw the line between Self and Other somewhere, so you might as well draw the line at 1920s Providence, Rhode Island, and call everywhere else from Boston all the way to the unspeakable abyss-city of Y’ha-nthlei just different degrees of horribleness.)

Overall I am not 100% convinced either way. Maybe some traditional cultures are worse than universal culture and others are better? Mostly the confusion makes me want to err on the side of allowing people to go either direction as they see fit, barring atrocities. Which are of course hard to define.

I like the Jewish idea of the Noahide Laws, where the Jews say “We are not going to impose our values on anyone else…except these seven values which we think are incredibly important and breaking them is totally beyond the pale.” Sometimes I wish universal culture would just establish a couple of clear Noahide Laws – two of them could be “no slavery” and “no eye-gouging” – and then agree to bomb/sanction/drone any culture that breaks them while leaving other cultures alone. On the other hand, I also understand universal culture well enough to know that two minutes after the first set of Noahide Laws were established, somebody would propose amending them to include something about how every culture must protect transgender bathroom rights or else be cleansed from the face of the Earth by fire and sword. I’m not sure how to prevent this, or if preventing it is even desirable. This seems like the same question as the original question, only one meta-level up and without any clear intuition to help me solve it. I guess this is another reason I continue to be attracted to the idea of Archipelago.

But I think that none of this makes sense unless we abandon the idea that “universal culture” and “western culture” are one and the same. I think when Caplan’s debate opponent talked about “protecting Western culture”, he was referring to something genuinely fragile and threatened.

I also think he probably cheated by saying we needed to protect it because it was responsible for so many great advances, like Coca-Cola and egalitarian gender norms. I don’t think that’s fair. I think it’s a culture much like Tibetan or Indian culture, pretty neat in its own way, possibly extra interesting as the first culture to learn the art of summoning entities from beyond the void. Mostly I’m just happy that it exists in the same way I’m happy that pandas and gorillas exist, a basic delight in the diversity of the world. I think it can be defended in those terms without having to resolve the debate on how many of its achievements are truly its own.

# Post-Partisanship Is Hyper-Partisanship

I.

A few years ago, I wrote:

I want to avoid a very easy trap, which is saying that ingroups vs. outgroups are about how different you are, or how hostile you are. I don’t think that’s quite right.

Compare the Nazis to the German Jews and to the Japanese. The Nazis were very similar to the German Jews: they looked the same, spoke the same language, came from a similar culture. The Nazis were totally different from the Japanese: different race, different language, vast cultural gap. But the Nazis and Japanese mostly got along pretty well. Heck, the Nazis were actually moderately positively disposed to the Chinese, even when they were technically at war. Meanwhile, the conflict between the Nazis and the German Jews – some of whom didn’t even realize they were anything other than German until they checked their grandparents’ birth certificate – is the stuff of history and nightmares. Any theory of outgroupishness that naively assumes the Nazis’ natural outgroup is Japanese or Chinese people will be totally inadequate.

So what makes an outgroup? Proximity plus small differences. If you want to know who someone in former Yugoslavia hates, don’t look at the Indonesians or the Zulus or the Tibetans or anyone else distant and exotic. Find the Yugoslavian ethnicity that lives closely intermingled with them and is most conspicuously similar to them, and chances are you’ll find the one who they have eight hundred years of seething hatred toward.

I didn’t coin a silly term for the relationship of the Yugoslavs and the Tibetans, but let’s use “fargroup” in order to remind us of the Near/Far distinction. We think of groups close to us in Near Mode, judging them on their merits as useful allies or dangerous enemies. We think of more distant groups in Far Mode – usually, we exoticize them. Sometimes it’s positive exoticization of the Noble Savage variety (understood so broadly that our treatment of Tibetans counts as an example of the trope). Other times it’s negative exoticization, treating them as cartoonish stereotypes of evil who are more funny or fascinating than repulsive. Take Genghis Khan – objectively he was one of the most evil people of all time, killing millions of victims, but since we think of him in Far Mode he becomes fascinating or even perversely admirable – “wow, that was one impressively bloodthirsty warlord”.

(this jars when other cultures do it to people we consider Near-Mode evil – for example India’s Hitler-themed clothing store, romance movies, and their use of Mein Kampf as a business advice book. It’s a bit strange, but not objectively stranger than us having a comedy movie about Kim Jong-un)

Fargroups aren’t always people who are literally distant from us. It seems more like it’s people who don’t threaten us, or aren’t in competition with us, or don’t get involved in the conflicts we care about, or something like that. There’s a Scientologist Church just a couple of miles from my house, and I recognize that Scientologists do some pretty horrible things, but none of them affect me, or people close to me, or values that I have a personal connection with, so I’m still more likely to find them cartoonishly funny, Kim Jong-un style, than I am to feel angry or afraid of them.

We exoticize fargroups, but we can also use them as props in our own local conflicts. For example, a lot of the time I hear about ISIS, it’s in contexts like the Democrats being weak on ISIS or Trump playing into ISIS’ hands, or how our immigration policy makes us easy prey for ISIS, or fundamentalist Christians are no different from ISIS, or something like that. We use sympathetic fargroups the same way. The Tibetans aren’t just wise and noble, they’re a foil to our overly materialist society, or an example of how religion can be based on reason instead of faith, or whatever. This is all as the theory would predict. The GOP view the Democrats as more of an outgroup and ISIS as more of a fargroup. It’s harder for them to have genuine outrage at ISIS for beheading a bunch of people, than for them to have outrage at the Democrats for not mentioning the beheading. Even in cases where they seem angry at ISIS in a non-Democrat related way, I would argue that a lot of it can be traced back to appreciating the way ISIS proves various domestic points, like “Muslims are scary” or “the barbarism vs. civilization axis is important”.

II.

Last month I asked on Tumblr:

I remember that when I was young and the Internet was young, people online were debating religion vs. atheism ALL THE TIME. It felt inescapable. Whatever else you were trying to discuss, eventually it would turn into a religion vs. atheism debate. Whenever it came up, people would sigh and say “Oh no, not another religion vs. atheism debate”.

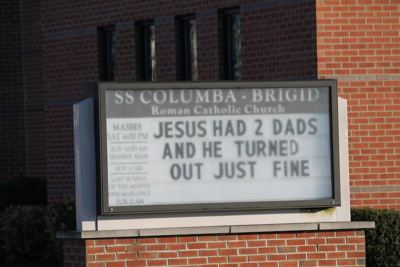
I remember spending a lot of time at talk.origins and infidels.org because religious people kept attacking me and I wanted to be able to rebut their points. And I remember a lot of people who seemed to genuinely believe that religion was like the #1 problem in the world, maybe even the only problem in the world because it was the root cause of all of the others.

I haven’t seen an online religion vs. atheism debate in years now. Occasionally somebody criticizes Richard Dawkins or something, but it’s always a tone argument and practically never about the nitty-gritty of Biblical contradictions or whatever. Now social justice vs. anti-social-justice seems to have totally taken over as the Annoying Thing Everybody On The Internet Has To Debate.

Has anybody else noticed this? Is it just me, or maybe a function of the places I hang out / used to hang out?

I got a lot of responses. Other people confirmed this was a real phenomenon and that they remember it the same way. The consensus explanation was that there was a moment in the 90s and early Bush administration when evangelical Christianity seemed to have a lot of political power, and secularists felt really threatened by it. This caused a lot of fear and arguments. Then everyone mostly agreed Bush was terrible, studies came out showing religion was on the decline, evangelicalism became so politically irrelevant that even the Republicans started nominating Mormons and Donald Trump, and people stopped caring so much.

Now I see atheists sharing things like this:



Not only have they stopped caring that much about religion, but they’re willing to adopt progressive religious people as role models and generally share stories that portray religious people in a positive light. Pope Francis gets to be the same sort of Socially Approved Benevolent Wise Person as the Dalai Lama.

I think once Christianity stopped seeming threatening, Christians went from being an outgroup to being a fargroup, and were exoticized has having the same sort of vague inoffensive wisdom as Buddhists.

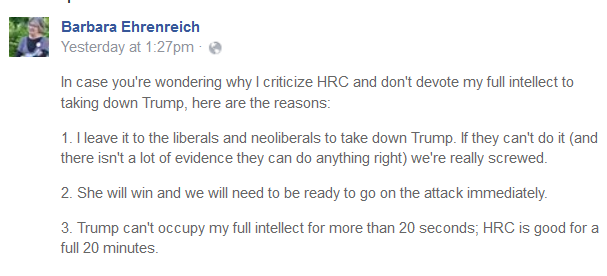
I saw something that seemed very similar during my time interacting with movement atheists. There was a split between people who were raised in fundamentalist families and very traumatized about it and who viewed Christianity as an outgroup, versus people who were raised in agnostic families and pretty live-and-let-live and who viewed Christianity as an fargroup. I know it seems weird to say that movement atheists living in a majority-Christian country treated a religion they interacted with every day the same way the Yugoslavs treated Tibetans, and sure, they would make fun of them, but that was exactly it – they found religion funny – and even in the process of lightly mocking them they tried to avoid stepping on too many toes. The fundie-raised atheists would propose something really combative and offensive, and the secular-raised atheists would say “Oh, come on, we don’t want to be jerks about this, Christians are basically nice people who are just a bit deluded”. To the fundie-raised atheists it was real, it was a hot war, these people were monsters; to the secular-raised atheists, religious people were just kind of wacky in a problematic way, like the North Koreans, and nobody in America lives their life in a state of constant rage about how evil North Korea is.

And I think as the threat of movement fundamentalism declined, there was a shift among atheists from more emotional hostility to more of a live-and-let-live kind of attitude.

(and then movement atheism started tearing itself apart even more viciously than it was already. I don’t know if this was a coincidence and I’m still curious whether conservation of tribalism is a real phenomenon.)

III.

From Facebook the other day:



All good reasons. But I’ve been seeing more and more people lately saying things like this. There have always been primary elections, and there have always been intra-Left disagreements, but the level of Bernie vs. Hillary drama at the Democratic Convention this week seems to be something new. Ehrenreich-style leftists focus on critiquing Hillary instead of Trump – either within or outside of the context of supporting the Sanders campaign. And on the other side, Hillary-supporting liberals go after Sanders and his supporters instead of Trump – Freddie deBoer has written frequently (some would say incessantly) about this.

The right, of course, has its own conflicts between Trump partisans and Trump opponents, culminating in Cruz’s non-endorsement. Also relevant: the alt-right’s favorite slur of “cuck” is short for “cuckservative” – an insult not for leftists but for conservatives who they think are doing conservativism badly.

People are talking more and more about partisan bubbles. People dividing into political tribes, and cutting off contact with people on the other side. Cultural, geographic, and social differences isolate people so completely that for example my Facebook feed tends about 95% liberal; I’m sure there are other people out there with the opposite problem. I think that as bubbleification increases, the other party becomes less and less of an outgroup and more and more of a fargroup.

Republicans still “threaten” me in the sense of being able to enact policies that harm me. And people less privileged than I am face even more threats – a person dependent on food stamps has a lot to fear from Republican victories. But Republicans aren’t taking over my social circle or screaming in my face. In a purely social context they start to seem more like cartoonish and distant figures of evil, rather than neighbors and coworkers. The average Trump voter no longer seems like an uncanny-valley version of me; they seem like some strange inhabitant of a far-off land with incomprehensible values, just like ISIS.

I have yet to meet anybody in person (other than my patients) who supports Donald Trump. On the other hand, I’ve met a bunch of people on both sides with strong feelings about Bernie vs. Hillary. The Bernie vs. Hillary conflict is real to me in a way that the Hillary vs. Trump conflict isn’t. It has the potential to split my friend group. There are social advantages for me of taking either side, and I could reasonably take either side without people looking at me like I went to work stark naked. This is the kind of socially relevant conflict that produces ingroups and outgroups in a way that America vs. ISIS never will.

My guess is that this sort of thing is only going to become more common. Partisanism is going to give way to hyperpartisanism, where people hate other factions of their own party with the same venom they previously reserved for their opponents across the aisle.

At the same time, old outgroup hatreds will take on a different character. Even If You Don’t Like Donald Trump, You Should Understand The Pain Of His Poor White Supporters. And I Know Why Poor Whites Chant Trump, Trump, Trump. And Millions Of Ordinary Americans Support Donald Trump; Here’s Why. And The Incredible Crushing Despair Of The White Working Class. I’m not saying these articles are typical; for every one of these articles there are ten “Trump Voters Are Xenophobic Trailer Trash” pieces. I’m saying that it’s weird that they’re happening at all.

Same thing with Brexit. Yes, the usual xenophobic trailer trash articles. But also: In This Brexit Vote, The Poor Turned On An Elite Who Ignored Them. And Brexit Voters Are Not Thick Or Racist, Just Poor. And Outraged Elites Should Listen To Fed-Up Brexit Supporters.

(and of course this blog has been pushing a similar line for reasons that are probably not completely ahistorical or divorced from general trends)

People are starting to treat Trump voters and Brexit voters as interesting and worthy of respect, which means they’re not really an outgroup any more. Talking about how poor they are and how sympathetic we should be and how we need to be more educated in order to understand what they’re going through all sound like instances of fargroup exoticization to me.

I predict (50% probability) that the progressives most carefully bubbled and separated from any actual threat from Republicans – which disproportionately includes politicians, journalists, and other opinion-makers – will start treating the Trump-voting classes more like Tibetans. I predict when they talk about specific bad Republicans like Trump, they’ll focus more on the ways they are funny and cartoonish (far too easy with Trump, but maybe the next guy will be a better test) instead of the ways they’re threatening. I predict that conflicts within the progressive movement will be increasingly vicious and increasingly likely to use poor whites as a political football (“the other side is bigoted against poor whites!”). I predict this will happen much more if the Democrats win the election than if they lose it; it’s always easier to be gracious toward a vanquished opponent.

(I’m of course 100% guilty of all of this myself)

(Yeah, this is a change in predictions since Right Is The New Left, which talked about something similar but reached a kind of different conclusion)

I’m not sure how things will go on the Republican side. I haven’t seen the same signs of rapprochement from them – but then Republicans have never shown the same tendency to sympathize with poor exotic fargroups that Democrats do. But I also don’t know as many Republicans and maybe if this were happening I would miss it.

# Book Review: PiHKaL

I.

PiHKAL (“Phenylethylamines I Have Known And Loved”), subtitled “A Chemical Love Story”, is the autobiography of Alexander and Ann Shulgin. Alexander Shulgin was a chemist who invented lots of new psychedelic drugs. Ann was his wife. Together they discuss their chemistry research and their relationship.

I was expecting a sort of popular science style book that cleverly ties the chemical story into the love story. You know the drill – the bonds between people are like the bonds between atoms, fragile in some ways yet incredibly strong in others. Or something like that. It would use the human interest story to hook you into the chemistry, then use the chemistry to give scientific respectability to the human interest story, so that both science nerds and hopeless romantics could enjoy it and gain more of an appreciation for the other side’s point of view.

Some parts of PiHKAL approached this kind of style. But a lot of them didn’t. Chapters and sections tended to be kind of either/or. You can be reading one moment about how MDMA is an n-methylated homolog of MDA, and the next moment about how Alexander Shulgin’s broad shoulders rippled as he was making love. It was a bit jarring.

The first quarter of the book was about Alexander Shulgin’s childhood and early life. He was born in 1930s Berkeley, went to school, became a chemist, and got a job with Dow Chemical. He invented a pesticide so successful that Dow gave him total freedom to work on whatever he wanted – which turned out to be psychedelics. He met, courted, and married his first wife Helen, a process which receives six sentences (compared to seven sentences a page later on the history of 3,4,5-trimethoxyamphetamine synthesis techniques). He has various scientific and professional successes, including a supporting role in the invention of MDMA/Ecstasy. His first wife dies of a brain haemorrhage.

The second quarter is a love story told from Ann’s perspective. It reminds me a little bit of very bad fanfiction, like “My Immortal”. Ann very briefly talks about her childhood, then gets to the part where she (after three previous marriages and divorces) meets Alexander. Cue lavish descriptions of his eyes, his hair, his muscles, his shoulders, how much better than her he is, how there’s no way someone as awesome as he is could possibly fall for someone boring like her, how many sparks get sent through her spine every time he gazes at her, et cetera. I am okay with people being in love, but this is a bit excessive. We get to hear about how amazing it is that he is a chemist, how amazing it is that he creates new drugs, how amazing it is that he is a brilliant yet dark and brooding loner, how there’s no way someone like that could ever love her (but spoiler: he totally does). Sometimes it seems to strain plausibility – there is a section where Shulgin tells Ann that he is interested in psychopharmacology, and she innocently goes into “Oh my, that’s such a big word for a girl like me, I wonder what it means”. But she has been married to a psychiatrist for ten years at this point, and also, she’s a hospital transcriptionist. I roll to disbelieve that she has never heard the word “psychopharmacology” before – let alone never heard the word “pharmacology” and the prefix “psycho-” from which the meaning is completely obvious.

The third quarter is an assortment of trip reports, social gatherings, and arguments against the War on Drugs. It is probably my favorite part, given that it’s neither as dry as the autobiography nor as overwrought as the love story. The trip reports about weird new psychedelics that nobody else has ever tried are really what I’m here for – they occur throughout, but especially here, and they do not disappoint.

The fourth quarter is a cookbook detailing the ingredients of, synthesis techniques for, and effects of 179 different psychedelic substances. It’s really fascinating, and I’m consumed by a desire to try some of it, except that they all begin by with instructions like “Obtain a professional-quality chemistry lab and several zillion different compounds with names like 2,5-dimethoxythiophenol”, and end with getting raided by the DEA. So I will have to stick to enjoying Alexander Shulgin’s psychedelic experiments vicariously.

II.

There were a couple of levels on which I enjoyed this book, though none left me completely without questions.

The first level, of course, is Shulgin’s work on psychedelics. My opinion on psychedelics hasn’t changed since Universal Love, Said The Cactus Person. I think they’re really interesting and mysterious and show every sign of pointing at something profoundly important. I also think that nobody has ever been able to consistently extract anything useful or scalable out of them, and until someone does, they’ll remain a weird toy where you take them and feel transcendent joy for a few hours, and if all you want is to feel transcendent joy for a few hours then they’re definitely the way to go, but as of yet it’s unclear what relevance they can have to any other project.

Alexander Shulgin disagrees. At least I think he disagrees. He is dark and brooding and quiet, and he doesn’t wear his heart on his sleeve even in his autobiography. But he mentions – once – that he believes his project is vitally important for the human race. This is Ann:

It seems to me that the magic plants – and the psychedelic drugs – are there to be used because the human race needs some way of finding out what it is, some way of remembering things we’ve usually forgotten by the time we’re grown up. I also think that the whole 1960s eruption- all that psychedelic experimenting and exploring – was due to some very strong instinct – maybe on the collective unconsciously level, if you want to use Jung’s term – an instinct that’s telling us if we don’t hurry up and find out why we are the way we are, and why we do the things we do, as a species, we could very soon wipe ourselves out completely

Alexander says this is what drives him also, then adds:

Of course, there are many ways to alter your consciousness and your perceptions; there always have been, and new ways will keep being developed. Drugs are only one way, but I feel they’re the way that brings about the changes most rapidly, and – in some ways – most dependably. Which makes them very valuable when the person using them knows what he’s doing. I thought for a while that I could use music to accomplish what I wanted to do, because music can be a very powerful consciousness changer. But when I discovered that I had a certain knack for chemistry, I made a decision to go that way, to concentrate on developing these tools. Mostly, I suppose, because these particular drugs, these materials, are a way to bring about new insights and perceptions quickly, and – well, I just don’t know if we have much time. Sometimes I suspect it may be too late already…I have no intention of getting lazy, and there’s nothing better than a suspicion that time’s running out, to keep you working hard.

Shulgin and his friends seem like good people. But not, crucially, like the best people. Shulgin himself – by his own admission, based on facts that he himself presents in his own autobiography – is consistently kind of a jerk to his wife (and his wife kind of agrees). He gets depressed and ornery a lot, sometimes to the point where it seriously interferes with his work and relationships. His circle of friends seems to have some problems with marital infidelity and random drama, and he tells one story about a distant friend-of-a-friend obsessed with LSD who seems to be an outright con man. I’m not saying they’re bad people; quite the opposite, the book makes them seem very human and if I lived in the same time and place as them I would be delighted to have the opportunity to know them. But they seem, well, about as good or bad as any other set of intelligent, creative people. It’s not clear that their psychedelic use – and man, do these people use psychedelics – has made them morally or spiritually exceptional. It’s hard to shake the thought that these people would be relatively nice and interesting artists and scientists with a little bit of marital infidelity and personal drama even if they’d never taken anything stronger than Tylenol.

Don’t get me wrong – during the trips they are constantly seeing God and understanding the oneness of everything in creation. But even Alexander Shulgin’s close friend group aren’t high more than like 20% of the time. I’m not sure exactly what about them makes them potential human-race-savers. Yes, I think they’re probably anti-nuclear-weapon. But this is the 1970s Berkeley counterculture; anti-nuclear-weapon people are not exactly hard to come by.

And so I was left with one question the book didn’t really answer – why is Shulgin doing this? What is his hope? Does he hope that the 200th new psychedelic he discovers will be the one that really teaches people universal love, to the point where they can’t ignore it? That just having twenty different slightly different permutations of the same psychoactive sulfur compound isn’t enough to create a world revolution in consciousness, but having thirty of them is? Why didn’t he become the scientist in this article, who has come up with a clever way to extend the DMT high in order to be able to enter complex negotiations with the machine elves, hopefully involving factoring large numbers? That’s the sort of project that can go somewhere.

I’m not saying that everyone needs to have a four-hundred-step plan for how exactly what they’re doing today is going to save the world, Leverage Research style. I can’t tell you how writing this blog post or doing psychiatry is going to save the world. And I’m not saying people can’t focus on their comparative advantage – if you’re a brilliant chemist, maybe you should invent new chemicals and let politicians and religious leaders figure out what to do with them. But Shulgin seems to think he’s doing something more important than coming up with new and better toys, and it’s not totally clear to me what this is.

Shulgin does have a great excuse – the War on Drugs is in full swing by the middle of his career, and this prevents a lot of his products from getting the trial they deserve. For example, he introduces a psychiatrist friend to MDMA, and the psychiatrist finds it to be an immense help to psychotherapy, helping patients realize and come to terms with their issues much faster than the non-chemically-assisted version. I’ve heard this from psychiatrists I know as well, and maybe we could have had a revolution in mental health if the DEA hadn’t banned this kind of thing (there is, in fact, a big literature on psychedelics in psychiatry, most of which shows impressive effects in the very small experiments that have been permitted thus far). So maybe Shulgin’s angle is that he’s developed very useful mental health treatments which unfortunately the medical establishment refuses to consider, but they’ll be there if anybody needs them. But then how come he keeps inventing more of them, seemingly with no interest in whether they help the mentally ill or not? How come he keeps talking about saving humanity instead of curing depressed people?

The only hint I get is during a trip report for a particularly powerful compound called 2C-T-4, where Shulgin writes:

There is a simultaneous union with everything around me, and thus with everything within me too. A complete identification with my environment. And a sense of being at total peace with it, as well. If this is me, then I thank the dear Lord for a wonderful awareness, at least for a short time, of the fact that we can be so rich and beautiful. The mind flows and with it the soul, and no matter what words I put down in an effort to catch the wondrous monolog, I can do it little justice…

I have been fooled, again and again, into thinking that the magic of the unified reality was in the drug, and not in the person. Of course it is in the person – and only in the person – but if a drug could be found that would consistently catalyze this, then it would be one of the most powerful and awesome drugs that could be conceived of by man. If it were this material, 2C-T-4, it would have to be held apart with a reverence that would be impossible to describe or explain on a patent application!!

But Shulgin later reports that the drug does not consistently have this effect; testing it on other people (and again on himself) he gets various interesting psychedelic trips but never a return to the same level of transcendence. So maybe Shulgin is looking for a drug that consistently works as well as 2C-T-4? But I’m not sure what he would do with it if he found it. Ann, for example, describes her first mescaline trip in language a lot like the language Shulgin uses for 2C-T-4, and many others (eg Huxley) do the same – but everyone already knows about mescaline. Would releasing a consistent version of 2C-T-4 to the world do something mescaline hasn’t already done? What has mescaline already done?

Oddly enough, it is Ann – who keeps on insisting that she is not intellectual, that she is hopelessly boring compared to Alexander, that we should be reading her parts only to gain a hero-worshipping outsider’s perspective on Alexander – whose speculations on this subject I really like. This is from her mescaline trip report. She says that on mescaline she understood for the first time that the world was perfectly good, and writes:

I nodded, remembering some of the phrases I’d read in books and articles about psychedelic experiences, phrases like “Everything’s all right just exactly the way it is,” and the equally infuriating “I’m okay, you’re okay,” which had always sounded unbearably fatuous and self-satisfied. I’d often thought angrily that the writers had conveniently forgotten about the babies in Calcutta garbage cans, sorrow and hurt and loneliness, and the rest of a planetful of miseries. I’d said to myself, here’s some whacked-out idiot rhapsodizing about life being all right just the way it is. It had never stopped me from reading about such experiences, but my liberal soul had always ground its teeth at that aspect of the reports.

Now – now I would have to take it all back, all that resentment, because I was beginning to understand. I stopped in the road and looked at Sam and looked past him, and around and up at the grey sky and I knew that everything in the world was doing exactly what it was supposed to be doing; that the universe was on course, and that there was a Mind somewhere that knew everything that happened because it was everything that happened, and that, whether I understood it with my intellect or not, all was well. I simply knew it and I knew that I would try to figure it out later, but that I had to absorb the truth of it now, standing on a wet road in Golden Gate Park.

At least here, she seems to be using the mescaline not to catapult humanity to a higher state of existence, but to make her peace with the current state. I’m not sure making peace with the current state is philosophically justifiable, but it seems to have helped her, and I can imagine it helping a lot of people, as long as they stick to viewing it as a psychological truth and not as an excuse for quietism – something Ann doesn’t seem to have done.

III.

The second level on which I enjoyed this book was anthropology and ethnography.

The Shulgins met through a group that had branched off of Mensa. Their social circle consisted of a mishmash of scientists, underachieving geniuses, mental health professionals, hippies, and people convinced that their new projects were going to save the world – all in the context of Berkeley and the Bay Area. This is also my social circle, thirty years later, so it’s interesting to see how things have changed and how they’ve stayed the same.

Ann starts her narrative while she is dating a fellow Mensan named Kelly. So, ethnographical study number one:

He was an intense man with a striking, angular face who had met me at a Mensa gathering four months earlier. The next day, he came to my house and asked me to marry him. He explained, much later, that of course he knew I would refuse – had counted on it, in fact – but that he had often found proposing marriage to be an effective way of getting a woman’s attention. There was no denying that it had done just that.

Kelly’s passions in life were computers, good-looking older women and the creation of new IQ tests. I also discovered that he had a generalized contempt for humanity, referring to most people ask ‘turkeys’, and a tendency to uncontrolled explosions of rage, which often resulted in his having to apologize later for damage done to someone else’s furniture or a relationship – usually both.

He explained about his painful illnesses in childhood and his demanding, punitive father, and asked me to be understanding a patient. It worked for a while (I’ve always had a soft spot for intelligent neurotics), but after one memorable day when he smashed some of my records in front of the children, screaming at me for coming home ten minutes late from work and keeping him waiting, I told him if he didn’t go into therapy, I was through.

Kelly’s answer was, “I’ve never met a psychiatrist I couldn’t out-think and out-reason, I’m not about to waste my time or my money on one of those cretins!”

This Thursday gathering in Berkeley was an effort on Kelly’s part to bring together people he considered intelligent enough to, as he put it, appreciate what he could teach them about using their minds effectively.

Ethnographic study number two, slightly edited for length; this is Ann introducing Alexander to some of her friends:

[I] proceeded to give him rapid-fire descriptions of some of Mensa’s main attractions, as they stood talking or moved past us. In my best museum-guide manner, I told him, addressing his right ear closely because of the noise in the room, “You see that man there, the tall one with the red vest? He created the SIG – Special Interest Group – which is known as the Orgy SIG; I forget his official title for it, something like Sexual Freedom SIG, but everyone refers to it by the other name. I’ve never been to a meeting, but I hear they’re a lot of fun for those who go in for that sort of thing…”

“That woman over there in the purple dress, the one standing in a straight line between us and the candles – that’s Candice. She’s a very good-hearted, motherish person who gives the Mensa tests in this area, and for a while her little boy, Robin, was the youngest member of Mensa in the country. He’s around ten now, and no longer the youngest.”

I told him about the mathematical computer which inhabited the sometimes bewildered soul of the young man on the couch, and he said he was very interested in that kind of mind, and would go over and talk with him later on. I said I hoped he would want to do that, because few people paid any attention to the boy, and he was very sweet.

I asked Shura [Alexander Shulgin]’s ear, “Why haven’t you joined Mensa, by the way? It’s a good way to meet interesting people, especially when you’ve been divorced – or widowed.”

“Well,” shouted Shura, “To tell you the truth, I never thought of applying, probably because you have to take an IQ tests, and I will not take an IQ test.”

“Why, in heaven’s name?”

“I feel total, complete disgust for all tests of intelligence, and only limited patience with the people who give them. When I was in the third grade or thereabouts, I was given a so-called IQ test, a Binet-something-something, and I made an honest and diligent effort to complete it. There were angular objects, and number games, and if-this-then-what types of questions and the strategies needed for getting to most of the answers were pretty obvious.”

“You did well?”

“Of course I did, and that’s where I really tangled with the school principal. He accused me of having cheated, since no one could get the results I had gotten without cheating, and so I was in essence thrown out of the testing group, and was pretty much humiliated. They obviously wanted scores that fit on a kind of distribution curve about some sort of a norm. Mine was a bit too far to the right of the curve. My mother was furious with the principal; she pulled me into his office and confronted him and lectured him about my integrity, which made me want to run and hide even worse than before. I swore then that I’d never take another IQ test, and I never will!”

Ethnographic description number three is Shulgin’s colleague, a German professor named Dolph, and his wife Ursula. Shulgin’s first marriage, to the woman he spent six sentences on in Part I, was never very happy – never unhappy, neither of them was abusive or anything, just sort of boring and straightforward. When Shulgin met Ursula, he fell in love, maybe for the first time in his life, and they had a brief affair before Dolph and Ursula had to go back to Germany:

We met, Ursula and I, two or three times in some inn or private place sufficiently far away from the Bay Area to minimize the possibility of being seen by a friend or acquaintance, and I discovered what it was to feel unashamed, uncensored, joyous sexuality.

Being in love, like any other kind of consciousness alteration, makes small but real changes in the way you view things about you, and in the way you behave around others. Over the years, my friends had come to accept me as what they affectionately called a “difficult genius”, and were quite used to my habitually ironic humor, cutting commentary, and somewhat sour view of the world. One of the hardest things I had to do, in my unaccustomed role of secret lover and beloved, was to avoid giving expression – in the company of family or friends – to the feelings of optimism and even outright niceness which overtook me now and then, and which I knew would cause some degree of concern if they were detected.

I really like this passage. Here’s someone who has tried more psychedelic drugs than anyone else in the world, and what really changes his outlook is the power of love.

But then it gets complicated. Shulgin’s wife dies. He corresponds incessantly with Ursula back in Germany, urging her to leave Dolph and come to California and marry him. She says yes, but asks for some time to plan. Shulgin is overjoyed and says to take as much time as she wants. Weeks turn into months. Months to years. They keep writing each other. Ursula insists that she continues to be excited at the impending plan to move to California and marry Shulgin, but she keeps asking for more time. She needs to close up her affairs in Germany. She needs to figure out a way to break it gently to her husband. She needs to stay to comfort her husband during this difficult time. She needs to figure out how she’s going to send her stuff.

Meanwhile, all this time Shulgin is calling her in Germany to talk to her a lot, and a lot of the time her husband answers the phone, and he’s got to suspect something at this point, but he’s still perfectly cheerful and friendly, and finally Shulgin asks, “You know your wife is planning on moving to California to live with me,” and he’s like “Oh yeah, I know”, and Shulgin starts to get a tad suspicious. Meanwhile, this is around the time he starts meeting/falling in love with Ann, and he keeps telling Ann “I really like you, but this is just a fling until my True Love Ursula gets here from Germany”, and Ann is always okay with this, because Alexander Shulgin is Objectively The Best And Most Attractive Person In The World, and obviously having him for a short time is better than having anybody else forever. But it starts to get really annoying and everybody is super confused by what’s going on, and finally one of their psychiatrist friends tells Ann:

Ursula is – how best to put it – she’s a person who, when she’s attracted to a man, intuitively senses what’s lacking in his emotional life, and she has a compulsion to become whatever that man most needs in a woman. She probably convinces herself each time that she’s truly in love, but I doubt she’s capable of what most of us would call real loving. The Jungians have a term, ‘anima woman’. The anima woman lacks a solid identity; like many great actors, she borrows – she takes on – a sense of wholeness from playing a part. In this case, it’s the part of the muse, the inspiration, the adored dream-woman. She fulfills a fantasy, and you can imagine the tremendous emotional rewards there are for her in such a role, as long as the affair lasts. Each affair lasts, of course, only until the next needy attractive man comes along. It’s all unconscious, by the way; I don’t think Ursula has the slightest idea of what she’s doing or why she feels compelled to do it. Or, for that matter, why the men she’s drawn to always happen to be married. When it’s time to move on, she explains – and probably believes – that she’s ending the relationship because she couldn’t live with the responsibility of having broken up a marriage.

When she first joined [our] group, we had long talks with each other, under the influence of [Shulgin’s psychedelics], and she told me a lot about her involvements with married men; she told me more than she realized or intended to. It was a subtle form of preening, under the guise of telling problems to a wise, sympathetic psychologist, you understand? Gradually, I put enough of the pieces together to understand the pattern. By that time, she had stopped telling me personal things about herself and her life, and I sensed that she was feeling uncomfortable around me […]

The dynamics of this kind of psychological compulsion are more than I want to go into right now, but what worries me is I believe Ursula is simply not capable of true emotional commitment to anyone. She’ll play the role for a time, as I said, until somebody else comes along – someone she finds appealing, with an emotional hole that’s begging to be filled – and she’ll move on to the new challenge. That’s what’s going to happen to Shura [Alexander Shulgin], I’m sure of it. I know it! I love him very much – we all do, you know – and sooner or later, he’s going to be badly hurt. That’s why I’m more pleased than I can say, to see you here. I don’t know what your relationship is with Shura, but it’s clear that you care for him, and I hope that – umm – I hope you’ll stay around. To help cushion the blow, when it comes; to give him something real to hold onto, when the unreal thing begins to unravel. Which I’m sure will happen before long, now that Ursula finds herself involved with a man who has – quite unexpectedly – become free to make an open commitment to her and ask her to do the same. Her bluff is being called.

Everything goes back and forth a lot, the whole social circle becomes hopelessly muddled, but in the end the psychiatrist is proven right. After much back-and-forth, Ursula agrees to fly to California that very day. When she doesn’t actually arrive, Ann sends her a nasty letter, informing her that she exists, that she’s on to her, and that now is the time to put up or shut up. Ursula then sends Shulgin a letter:

Dearest, dearest Shura,

A window has widely opened to you, a soul-window, a love-window, of graceful being – being together. A common space of breathing, of light touch, of inner smile. I could let those hours pass without telling you, and then you would never know what I am feeling – you would have only your own experience. Or I could share this with you. That is what I am doing […]

In a past life, about 2,000 years ago, you took a long knife and cut my throat, took my life, murdered me in the desert! You were the chief of our tribe, and I was a young girl, and you killed me! The whys are irrelevant. I have seen this over and over, and others who lived with us in that time have come to me in this life and warned me to be aware of this old karmic connection. We were, I think, of a nomad people in North Africa when this happened so long ago […]

In this moment of open love, you might be able to believe what I say to you, that I do not have any misgiving or second thoughts about emotional involvements with you because of this vision of what happened so long ago. No, my only concern is, and this is very real to me, to free myself and to give you the possibility of freeing yourself, from these old, old bonds of emotional slavery which must not be repeated in this life. In this life, through our deep love, we have the real chance of changing this by bringing it out into the open. We have broken a karmic consequence and do no longer have to blindly bear the burdens of the past life and tragedy.

I am leaving Dolph and I will go to a place to begin a new life with myself. I do not think I will marry again. I must seek alone my true path of the soul. I love you very deeply and I go to live my own life, of which you are a wonderful spiritual part. Maybe it will come that you will be a material part as well. But now you must live the present as completely as you can.

Shura, my dearest one, I want you to be free as a bird. Unfold your wings and leave all pain behind you, all possible accumulated guilt, all disquietness, all sorrows. Be free, and newly born, and walk into sunrise!!!

Fly and be!  
Ursula

So much for ethnographic study number three.

IV.

All of this seems somewhat more dramatic than normal reality. I don’t know how much liberty the Shulgins took when writing their autobiography – maybe this is another one of those things like Ann not knowing what “psychopharmacology” was. But one last thing I noticed about the book was how clear and coherent the psychology of everyone in PiHKAL was.

The Shulgins and everyone they know are Freudians – not explicitly, nobody ever says “I am a Freudian”, but just on a deep level they assume that it’s obviously true. Sometimes if it is an especially good day they’re Jungians as well, in the same implicit way. They’re always getting messages from their unconscious, they’re always rediscovering psychologically repressed material, and they’re always meeting people like Ursula who seem driven to behave in dramatic and unusual ways which are very predictable to any of the approximately one zillion psychoanalysts and psychiatrists whom the Shulgins know. Even their drugs are good analysts – the psychedelics are always helping them remember repressed childhood memories, after which they feel much better from whatever was bothering them at the time.

This is interesting, because I almost never see anyone behave in as dramatically Freudian a way as the Shulgins and their friends seem to behave all the time – even though I occasionally do psychodynamic therapy on people! I’m left a little baffled. Part of me wants to say that the primitive mind sees omens everywhere – I’m sure medievals were always seeing various signs of Christ in their daily life, and pattern-matching has never been a difficult sport (ask me how I feel about that brilliant and humane reflection on theodicy above being by someone named Ann). Another part wonders whether, if you’re Freudian enough, your subconscious starts acting in Freudian ways just to keep up – although that itself is a Freudian idea and I’m not sure whether you can get it without presupposing Freudianism anyway. A third possibility is just that the more crazy drugs you’re on, the more Freudian you act – wasn’t Freud a coke fiend anyway? A fourth possibility is that the problem is with me – I’m somehow so closed to all this kind of thing that when people around me tell me Freudian stuff, I completely miss it without Ann Shulgin’s narrative voice to tell me how Freudian it is – or even actively repress it.

Maybe the most interesting chapter of the book was where Ann had a spiritual crisis. She takes a psychedelic called DESOXY which Shulgin thinks is pretty weak, and she has a very strange reaction where the world starts seeming hostile, emotionless, and run by a perfectly rational demiurge that doesn’t care about humans the tiniest bit. The psychedelic leaves her bloodstream, she reaches the point where she should feel normal again, but she can’t shake her feeling of the demiurge’s obvious and palpable presence, to the point where she becomes barely able to function. She goes to one of the zillion or so Jungian psychologists in her friend group, who matter-of-fact tells her she’s having a spiritual crisis, and the only thing to do is wait for her soul to process it and gain the necessary enlightenment to go on (I wish I could get away with saying this kind of thing to my patients). Then she starts having extremely vivid visions of what is very obvious her mind doing Internal Family Systems therapy on herself, despite this being way before Internal Family Systems was invented, and despite the inventor being one of the three or so psychologists who was not a personal friend of the Shulgins. Finally she gets all the IFS steps right, accepts her parts, frees her repressed memories, and stops feeling like the Demiurge is harassing her at every moment. It’s pretty fascinating, but that’s just the thing – even though I’ve tried really hard to do Internal Family Systems on myself, armed with an official book and everything, I get nothing. I never have these sort of exciting spiritual crises that partake of exactly the right amount of symbolism from each of the world’s great mystical traditions. I’m sort of jealous of all the people who do, and sort of suspicious of them. Maybe I need to take more psychedelics.

(I will note, though, that the book is appropriately skeptical about some of this. Shulgin describes going to a psychedelic conference and meeting an academic who worked in ethnobotany. He was studying a certain psychedelic plant, and was especially interested in why everyone who took that plant had hallucinations of jaguars in particular. He theorized that the plant was from the Mexican jungle, and that in some deep way our collective unconscious knew this, and so came up with appropriate hallucinations. Another psychedelicist who happened to hear the conversation interjected “I synthesized that chemical a little while ago, and all I got was wiggly lines.” Shulgin left them as the first was getting increasingly agitated and demanding of the other whether he might have seen something that looked kind of like a jaguar.)

Speaking of weird things that Ann Shulgin sees, I’ll end this with something that might be interestingly testable. She writes that when she was young and going to sleep, every so often she would have a strange experience:

Lying down for naptime (as a child) or at night for sleep, I would have reached that point of relaxation where one is not very much aware of the body…when I sensed it beginning (I never knew when it was going to come), I would immediately snap into alertness, excited and pleased, then I would just lie quietly as it unfolded…every part of it was the same each time. It was always in black and white…and I could never extend it, by so much as a few seconds. When it was finished, it was finished.

First came the image-sensation after which I named the entire experience – the spiral. I felt my entire self drawn rapidly into a tiny point which kept shrinking, until it could shrink no further, at which time the microscopic point became a tunnel in which I continued traveling at great speed, inexpressibly small and implacably diminishing.

Simultaneously, I was expanding. I was expanding to the edges of the universe, at the same tremendous speed as that of the shrinking, and the combination, the contraction-expansion, was not only an image, it was also a sensation the whole of me recognized and welcomed. This experience of myself as microcosm-macrocosm lasted exactly four minutes.

The next stage came abruptly, as did all the changes. I was looking at standing figures which were vaguely human, dark thin figures being pulled into elongated shapes, like the sculptures of Giacometti. They stretched out, arms and legs like black string, until it seemed they could elongate no further, then the scene changed and I was watching obscenely rounded bodies, Tweedledums and Tweedledees without costumes, their small heads and legs disappearing into their puffed, bloated flesh. The sensation accompanying this stage was one of discomfort, unpleasantness, a feeling of something grating on my soul. I once timed this part and the one that followed; they lasted a total of six minutes. I disliked them intensely.

Abruptly again, the inner screen became white, a horrible dead-white, nasty and aggressive like the underbelly of a sting-ray. After presenting itself for a few seconds, the flat white began to curdle from the outer edges into black, until finally the screen was totally black. A thick, awful, dead black, a pool of tar in an unlit cave deep underground. After another brief pause, the black began to curdle at its edges into the white again. The process repeated itself once, and the sensation was similar in every way to the previous one: irritating, grating, a feeling of unpleasantness that approached repugnance. I always endured it with a mental gritting of teeth, knowing it had to be gone through because that’s the way it always went and it was not to be changed.

And then, finally, I broke out into the last stage, the final part for which I had always been and always would be willing to undergo the middle parts. Now I was at the edge of an unseen cliff, looking out into a very different blackness, the deep, cradling blackness of the infinite universe, of space which stretched without end. I was completely happy and comfortable in that place, and would have stayed there indefinitely, had I been allowed, breathing in the beautiful darkness and the exquisitely familiar sense of infinity as a living presence, surrounding me, intimate and warm.

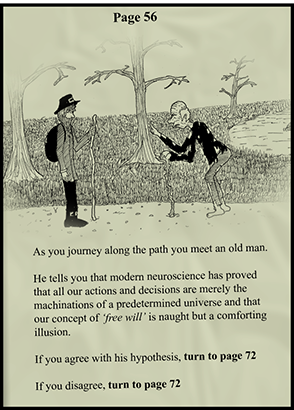
After a moment of this pleasure, came the greeting. From the upper left-hand corner of the universe there came a greeting from Something which had known me, and which I had known, since before time and space began. There were no words, but the message was clear and smiling: Hello, dear friend, I salute you with respect-humor-love. It is a pleasure with laughter-joy to encounter you again…

Then it was over. It had taken exactly twelve minutes.

According to Ann, she had this experience every so often “since I was born”, maybe once a week or so when she was a child, but becoming gradually less and less common until it finally happened for the last time when she was twenty-five. She said it was what got her interested in psychology and spirituality in the first place, and that her later access to psychedelics seemed to be sort of a substitute for the connection she had lost.

She also said that she met two people at cocktail parties who had exactly the same experience with exactly the same sequence of steps (though cutting out earlier). I am sure more people read Slate Star Codex than Ann Shulgin talks to at cocktail parties. So come on, people. Any of you ever have a very specific black and white vision of infinity?

# OT56: Spur Of The Comment



(source)

This is the bi-weekly visible open thread. There are hidden threads every few days here. Post about anything you want, ask random questions, whatever. Also:

1. Comment of the week is Mammon talking about clandestine MDMA labs. But see also the people who reported Spiral-like experiences in the comments to the PiHKaL review, (1, 2, 3, 4, 5), with my special interest caught by Kaminiwa’s report that his brother’s night terrors were like this. I’ve always wanted to know more about night terrors, since classically nobody can remember the content. These kinds of spiral experiences – things that are kind of like dreams, only different, and potentially very dysphoric, and common in childhood but disappearing as you grow older – seems like potentially a good match.

2. Lots of people pointed out last time that “banning anonymous commenting” was meaningless, since people could just register names like “Anonymous1” and keep commenting as normal. So let me be more specific – what do you think of requiring email verification (without listing the emails publicly) to comment? I know that it’s pretty easy to get working fake emails, but it would at least be a trivial inconvenience to constantly getting banned and re-registering.

3. By popular request, Deiseach is now unbanned.

4. Please don’t send me emails offering me sponsorship deals, affiliations with your own site, this one weird trick to increase my visitor count, et cetera. Please also don’t send me emails requesting that a link you like be included in the link roundup, especially not a link to your company (advertising is available if you want it). These are getting kind of high-volume and annoying. If you have something that I absolutely need to know about, you can try posting about it in the open thread, on the subreddit (which I definitely mine for good links), or by some kind of social proof where you convince somebody I know really well and they bother me about it normal conversation. I will grudgingly tolerate exceptions for important community events and very good charitable causes.

# Tolerance Troubles

[Not meant as a claim that science doesn’t know something. More of an admission that I don’t know some things, and a hope to be informed about them by someone who does.]

Everyone knows about tolerance. The first time you take heroin, you get really high. The second time you take heroin, you get slightly less high. The nth time you take heroin, you barely feel good at all – but if you stop taking heroin, then you feel miserable. Your body adjusts, the receptors desensitize, whatever.

This is so simple that it took me forever to ask the obvious next question – how come this doesn’t happen for everything else? Supposedly if you have ADHD you can just stay on Adderall forever. Nobody says “The first time you take Adderall you can concentrate really well, the second time you take it you’ll concentrate less well, and the nth time you take it you can barely concentrate better at all.”

The psychiatry textbooks contain a sentence or two saying that “some” patients “may” develop Adderall tolerance, but it’s not something that we’re trained to expect. There are a lot of anecdotal reports online, but there are also anecdotal reports of people who don’t develop any tolerance at all after years and years. Hmm.

Also, people who abuse Adderall develop tolerance all the time, and keep having to up the doses just like heroin abusers do. This is a little weird – my pet theory is that people only develop tolerance to drug effects that aren’t FDA-approved uses – though how your receptors know what the FDA says I haven’t quite figured out. More seriously, it may be an effect of method of use – taking a small amount daily versus snorting a large amount of crushed tablet whenever you feel like it. Or it may be that tolerance to euphoric effects is worse than tolerance to stimulant effects.

(This seems true in general – I get euphoric effects from caffeine when I drink it very rarely; if I drink it more often, the euphoria goes away and I just feel a little more awake. This is important since it suggests tolerance isn’t just your body metabolizing the drug better, but actually a matter of receptor-level action – something I think everyone agrees is true, but which it’s always nice to have independent confirmation for.)

Sometimes tolerance gets weird. Antipsychotics are supposed to block dopamine receptors. Too much dopamine can contribute to psychosis, but it can also screw up the basal ganglia’s modulation of movement and cause you to make repetitive jerking motions all the time. People who have been on antipsychotics for too long may remain protected from psychosis, but start making repetitive jerking movements in a way consistent with too much dopamine. The theory goes that the receptors involved in psychosis haven’t developed tolerance (for some reason), but the receptors involved in movement modulation have developed so much tolerance that they’ve overshot their baseline and become supersensitive to dopamine. So by taking a drug that lowers dopamine, you get higher dopaminergic effects. In the worst case scenario, you end up with a condition called tardive dyskinesia, which is permanent. The receptors stay supersensitive forever and you will always make repetitive jerking movements. If you stop the antipsychotic, that will just make it (temporarily) worse – now you have supersensitive dopamine receptors and you’re not blocking them, so that means lots of repetitive jerking movements.

In this case, giving someone a drug has caused them to develop not just tolerance but supertolerance, where they are permanently worse than before. It’s as if taking heroin for long enough made you permanently miserable.

…which actually isn’t totally hypothetical. Some percent of people who abuse opiates like heroin get what’s called a post-acute withdrawal syndrome (PAWS), meaning that they feel depressed for months or years after they stop using the opiates. I treated a patient like for a while. I tried pretty much every antidepressant on him without success. He was just miserable. He’d been clean for about six months and I told him that he might just have to wait it out – it usually goes away after a few months to a year or so.

There is a poorly-studied but anecdotally very helpful treatment for PAWS: low-dose naltrexone. Naively, this sounds like the stupidest possible thing to try. It’s an opiate antagonist, meaning that you’re taking somebody who is undersensitive to opiates and blocking the tiny number of functioning opiate receptors they already have. It should be the only thing capable of making this already bad condition worse. Yet people swear by it. The theory is supposed to be the tolerance reaction again. Your body reacts to this opiate-blocking agent by releasing more opiates. So we’re treating a condition in which drugs that increase opiates cause you to have fewer opiates, by giving you a drug that decreases opiates which will cause you to have more opiates. How annoying is that?!

(some people recommend that if you’re giving someone opiate painkillers, you can give them low-dose naltrexone at the same time to prevent development of tolerance. Giving someone an opiate and an opiate-blocker simultaneously seems kind of like the medical equivalent of digging holes and filling them back in again, but apparently it does something useful)

This means we have examples of all three of the following:

1. A drug that’s supposed to have effect X, and after a while it still has effect X (Adderall)  
2. A drug that’s supposed to have effect X, but after a while it has no effect (heroin)  
3. A drug that’s supposed to have effect X, but after a while it overshoots and has effect anti-X (Antipsychotics? Heroin? Naltrexone?)

You may notice that these are all three of the logical possibilities. So for example, if we gave someone a drug that was supposed to decrease anxiety, it might decrease anxiety, have no effect, or increase anxiety. If scientific hypotheses are about closing off parts of possibility-space, then the receptor sensitivity hypothesis isn’t doing a very good job.

But it’s actually worse than this, because I get the impression that different people will end up in different branches of this trilemma. Benzodiazepines are a special offender here. Some people can take Xanax once a day for anxiety, and it’s a perfect solution – it suppresses their anxiety, it never stops working, and they never become addicted – if twenty years later they get a good therapist who helps them treat their anxiety without drugs, they can stop the Xanax with just a couple of days of mild discomfort. Other people will lose all effect after a couple of weeks, up the dose, up the dose some more, and end up as total wrecks. I think this is much less common than most people say – my attending’s rule of thumb is “benzo tolerance develops for sleep but not anxiety” – but it certainly happens. And for that matter, I’ve met a few people who never seem to develop tolerance for benzodiazepine sleeping pills. You see this same pattern for opiates used as painkillers. I spent so many years confused about whether people develop tolerance to these or not, and my final conclusion is that some people do and some people don’t and if you try to find a coherent universal pattern here you will go insane.

And it’s actually worse than this. Drugs don’t just work differently in different people, sometimes the same person will cycle through totally different mechanisms of drug response. SSRIs have something called tachyphylaxis, where they’ll work really well for months and then suddenly stop working (the word means “fast protection”, ie you develop protection against the drug effects quickly). This is even more annoying than the other patterns – at least with heroin, it makes sense that the receptors will gradually lose their sensitivity. But here? In random people at random times, the drug just stops working suddenly. It might be after a month, it might be after a year, it might be after ten years. And then every so often you’ll try the drug again a decade later, and then it’ll work just fine. Why? Nobody knows.

Some skeptics have pointed out that this is exactly what you would expect if the drug had no real effect and it was just luck that people didn’t have depressive episodes while they were taking them, but we know SSRIs have some effect. And anyway, placebo tachyphylaxis isn’t any less weird than real tachyphylaxis.

One more weird thing: LSD users report very strong tolerance lasting about three days after a dose, to the point where a second dose the day after will do almost nothing. Rat experiments have shown this is definitely because of receptor downregulation and not just enzyme induction. Okay. But LSD is a pretty strong drug. If receptors are so down-regulated that you are essentially on negative one tabs of LSD, how are you remotely normal while the tolerance is in effect? Are people during their periods of LSD tolerance less crazy and creative than normal? What the heck is the 5-HT2A receptor even doing if decreased amounts of it sufficient to render LSD ineffective don’t have noticeable effects on consciousness?

This is my concern about naltrexone as well. Sometimes doctors give naltrexone to help with alcohol addiction, which usually works okay. The theory is that since naltrexone blocks opiates, and opiates power the endogenous reward system, alcohol will seem less rewarding. Fine. But shouldn’t everything seem less rewarding? I always worry that I’m just blocking my alcoholic patients’ ability to enjoy anything at all (of which enjoying alcohol is a subset), but that doesn’t seem to be how it works. This is about when I default to my theory of “receptors read the FDA labels for medications and make sure to only do what they’re supposed to”.

All of this annoys me for a few reasons.

First, psychiatry really doesn’t think about this enough (or sometimes at all). The pharmacology textbooks will tell you how effective a drug is, how long it lasts, how many side effects it has, et cetera, but not whether it’s going to produce tolerance or not. It’s mostly just assumed that it won’t.

Second, groups skeptical of psychiatry are always talking about tolerance and it’s hard to tell whether they’re right or wrong. For example, some people claim antidepressants cause tardive dysphoria – that like the antipsychotics that eventually give permanent repetitive jerking movements, antidepressants can make serotonin receptors permanently undersensitive (or something) and so make depression worse. Other people say that antipsychotics themselves can eventually screw up dopamine receptors in ways that make psychosis worse (though see here). My guess is that these problems don’t arise for most people, but I can’t explain why these things wouldn’t happen.

Third, I think something like this is involved in addiction. Addiction is highly genetic; some people can drink alcohol socially their entire lives and never become alcoholic; other people quickly get hooked. This seems related to the thing where some people are stable forever at their low dose of opiate painkiller, and other people quickly develop tolerance and need to keep increasing it. I’m sure there are other things involved in addiction, but this is probably one of them.

Fourth, how many interesting things are we missing because they’re stupid and make no sense? I don’t know who first discovered that low-dose naltrexone could help potentiate the effect of opiates, but there have got to be other things like that. Forcing your body to become more sensitive to its own chemistry seems like a good alternative to forcing more and more foreign chemicals into it.

Finally, the best drugs seem to be the ones we hesitate to use because they produce too much tolerance. Xanax, opiates, you name it. A version of Xanax that that didn’t produce any tolerance would be a holy grail of anxiety disorder pharmacology. Some way to switch off Xanax tolerance would be just as good.

And a tolerance-free version of heroin would be pretty interesting too – from a purely pharmacological perspective, of course.

# The View From Ground Level

[Epistemic status: Any time I make an anthropic argument, you should probably interpret it as trolling]

Sean Carroll argues that the simulation argument is false.

The simulation argument posits two kinds of universes: “high-level” universes that can simulate other universes, and “ground-level” universes that can’t. By the terms of the simulation argument itself, most universes will be at ground level, since every high-level universe can simulate many ground-level ones. So (says the argument) we should expect to be at ground level. But the simulation argument itself hinges on our observation that it looks like our universe is capable of simulating other lower-level universes. So apparently we aren’t on ground-level. So the simulation argument is probably false.

(I might be summing it up badly. Read the actual post for more.)

Suppose Carroll’s reasoning is right. What would a ground-level universe look like?

It would have to be pretty weird. It would have to ban the creation of Turing machines – since with enough time and resources any Turing machine could be expanded into a full-scale simulation. But Turing machines are pretty simple, and brains supporting conscious observers are pretty complicated. To have conscious observers but not Turing machines – well, once again, this would have to be pretty weird.

Brains would have to run off a science different from the local science accessible to in-universe researchers. Probably they would be run remotely, in the simulating universe, and then the results beamed into the simulated universe with no regard for the computational rules of the simulation. Maybe an alien dissecting a fellow alien’s head would just find a perfectly featureless crystal with no internal structure, which is observed to inexplicably send nerve impulses to the rest of the entity’s body. Such aliens might invent psychology, but never neuroscience, and even if they speculated about it, it wouldn’t matter – attempts to “simulate” neurons would fail, their workings forever beyond locally accessible physics. Even if they completely mastered their local science, their brains would remain a mystery.

I used the phrase “conscious observers” above. There are versions of anthropics that work for p-zombies, but we’re not p-zombies and we don’t have to use them; we can do anthropics conditioning upon consciousness. Try that, and the simulation argument doesn’t exactly depend on a ground-level universe where further simulations are impossible. It depends on a ground-level universe where further simulations containing conscious observers are impossible.

This changes the scenario a bit. Now people in ground-level simulations can expect arbitrarily complex physics, physics that allow the creation of as many Turing machines as they want, but which can’t possibly explain consciousness. They should be able to master every aspect of the universe around them except consciousness, which try as they might will remain refractory to their simulations. Consciousness will make perfect sense in the physics of the universe above theirs, but the simulators will have excised all consciousness-related rules from the ground-level sim. Try as the simulated scientists might, it’ll remain a mystery.

If Carroll’s deconstruction of the simulation argument is right, then the more trouble we have explaining consciousness, the more that should push us to believe we’re in a ground-level simulation. There’s probably a higher-level version of physics in which consciousness makes sense. Our own consciousness is probably being run in a world that operates on that higher-level law. And we’re stuck in a low-resolution world whose physics doesn’t allow consciousness – because if we weren’t, we’d just keep recursing further until we were.

# Devoodooifying Psychology

[Epistemic status: very low. Total conjecture based on insufficient evidence.]

“Voodoo death” refers to supposed cases where people died after being cursed by witch doctors. The theory goes that even though witch doctors don’t have real magic, if their victims come from a culture that believes in witchdoctory then they’ll be so scared that they gradually waste away out of fear and die anyway.

For a while psychologists believed that this absolutely happened, a testament to the powers of the mind. Maybe it was because of toxic levels of the stress hormone adrenaline or something.

Now there’s some more controversy. A lot of these cases turned out to be primitive tribesmen who said it totally happened to a friend of a friend of a cousin or something. In others, witch doctors placed curses on people who already had some kind of serious disease, then took credit. In still others, the curse victim became so upset that they stopped eating or drinking and died of dehydration – which, while technically a death, doesn’t really testify to the powers of the mind so much as the power of blood osmolality. And after some more thought, everyone agreed the adrenaline theory probably didn’t apply since adrenaline spikes kill suddenly but voodoo victims waste away over the space of weeks.

So now voodoo death looks a lot more complicated than a simple progression of curse -> adrenaline -> you were killed by your own mind. Lester gives a relatively sympathetic view of the evidence here, but even he can only find two good cases – one of which involved a patient who died of preexisting asthma, the other of which did not involve voodoo at all.

I find this interesting because so much of psychology seems basically voodoo-ish.

Take the placebo effect. This is basically the voodoo effect in reverse. Instead of a witch doctor saying you’ll get worse, and you do, a regular doctor tells you you’ll get better, and you do. For a while, people were claiming all sorts of amazing effects for placebo – placebo can activate the immune system to fight infections, placebo can slow the growth of cancer, placebo can make bedridden invalids start dancing jigs. But the best studies now suggest that the placebo effect is probably very weak and limited to controlling pain. The vaunted power of mind over body, of belief over reality, doesn’t look nearly as impressive as we thought.

Or take stereotype threat. Again, this is sort of a voodoo curse. If people make you think you’re going to do bad on a test, then you’ll do bad on the test. Again, widely believed, held up as an example of the power of perception. Again, doesn’t replicate well in large studies, has a very suspicious funnel plot, and is starting to inspire doubt even among top researchers in the area.

Or take self-esteem. Again, a reverse voodoo curse. By believing that you’re a good person and likely to do well, good things will happen to you. Again, very popular in the ’90s, but it hasn’t aged well. Similarly, self affirmations failed to replicate results showing their effectiveness.

This is all oversimplified; there are still lots of unrebutted studies supporting all of these things. Maybe some of the studies that seem to debunk them have themselves been debunked, and I don’t know about it. Still, it seems to me that things that sound like voodoo – that is, which argue that our optimistic or pessimistic beliefs about how well we will do can mysteriously and directly affect how well we will do – are faring unusually badly as psychologists get better at trying to replicate things.

(This kind of thing is why I’m so skeptical of growth mindset, despite so little hard data supporting my skepticism)

This isn’t to say that nothing like this is true. At the very least, my belief that I can’t swim the English Channel makes me not try to swim the English Channel, and so if that belief is wrong it’s voodoo-ish-ly making me fail at Channel-swimming. But this is a lot less mysterious than the thing where you repeat “I will do well at school” every day and so do better at school.

Let me introduce a second category of things.

First, the name preference effect. This is where you’re positively predisposed to things that sound like your name. For example, people named Dennis are more likely to become dentists, people named Bob are more likely to be bakers or butchers or barristers, and people with three letters in their name are more likely to live in Three Forks. People believed this for years until finally somebody did a statistical reanalysis and found that it was totally false.

Second, unconscious social priming. Supposedly people who heard the word “retirement” walked more slowly for a while afterwards, because “retirement” primed their thoughts of old people, and old people primed their thoughts of being slow, and so for a while they themselves behaved like an old person. This sort of thing inspired an entire field of psychology showing similar results (like the infamous study showing that an earthquake increased divorce rates by priming thoughts of instability), but it very much failed to replicate and is now the archetypal example of a formerly-accepted finding now believed to be false. Dr. Primestein’s work in this area is also a must-read.

Third, and from just last month: Artificial surveillance cues do not increase generosity: two meta-analyses. You know how if there was a picture of eyes or something, people would be nicer and more law-abiding, because deep down they felt like they were being watched? Yeah, turns out that’s not true.

And again, don’t trust me too much – there are a lot of studies I could have mentioned here, and three don’t necessarily make a pattern. But if I’m right that these are representative examples, then they seem to share a pattern. I’m not sure they all technically qualify as “social priming” (though they might, and that field has already been pointed out as particularly bad) but they all have the same feeling of tiny cues that you don’t think about causing big unconscious changes to behavior. This seems to be another category that is faring unusually badly.

Implicit association tests probably don’t work (1, 2, 3, 4). That is, people who have “implicit racial biases” according to the tests are not more racist in everyday life than people who don’t. If this were true – and if it reflected a general failure of implicit racial biases to affect explicit actions – it’s hard to overestimate how much it would change psychology. We wouldn’t have to worry about how the wrong character on TV would accidentally bias people toward having certain stereotypes. We wouldn’t have to worry about subconscious racism affecting hiring decisions even among people who are trying hard to be fair and neutral.

Does this fall into the previous patterns? It’s not exactly about self-fulfilling prophecies, or tiny stimuli having oversized effects on behavior. But it seems to have a certain kinship with them.

And a few days ago, a friend posted a quote on Tumblr:

There is no sovereign sanctuary within ourseles which represents our real nature. There is nobody at home in the internal fortress. Everything we cherish as our ego, everything we believe in, is just what we have cobbled together out of the accident of our birth and subsequent experiences. With drugs, brainwashing, and other techniques of extreme persuasion, we can quite readily make a man a devotee of a different ideology, the patriot of a different country, or the follower of a different religion

I was only too happy to be able to reply with Gwern’s research on how “brainwashing” mostly doesn’t work. Does this suggest the post is wrong about the lack of a “real nature”? Does this relate to any of the previous patterns?

A single thread seems to run through all of these examples: a shift away from the power of the unconscious. The unconscious doesn’t make you succeed or fail proportionately to your belief in yourself. The unconscious doesn’t change your behavior because of insignificant environmental cues. The unconscious doesn’t make you racially discriminate despite your own better nature. The conscious mind is strong enough to hold onto its preferred beliefs despite brainwashing techniques intended to force it otherwise.

So maybe we should update in general towards less of a role for the unconscious mind?

I remember my first freshman psychology class. After studying learned helplessness, I realized that was what I was feeling: study after study of crazy things, everything depended on your beliefs, the first letter of your name could affect your life outcome, stuff like that. The end result was a lot like in the quote above – nobody has any control over their lives, we’re all at the mercy of vast unconscious forces, society’s hidden assumptions and bases are vital in shaping our future. What if all of that was wrong? What if people mostly make decisions based on reasonable factors, succeed or fail based on things like ability or random luck, and social assumptions are relatively powerless beyond a common sense level? Wouldn’t that be great?

But maybe this is going too far. Once again, all I have is a few data points, curated by my own biases. Certainly not all of the studies showing creepy unconscious effects have been disproven; probably fewer than 10% of them have. Certainly not all the studies that have been disproven show creepy unconscious effects – ego depletion is a very mechanistic biological idea, but it’s done no better than priming. You have to kind of squint to see the pattern, then take it on faith that it’s real and that it continues throughout all the data that haven’t been checked.

And how do you tell the baby from the bathwater here? Some results – like cognitive biases, sales techniques or eyewitness unreliability – sort of fall under the heading of “power of unconscious effects”, but seem subtly different – maybe because of a less agentic unconscious? I don’t know. But I would be surprised if those followed the same pattern.

(on the other hand, a first draft included the Asch conformity experiments in that list, but apparently those never said what I thought they did)

Of course, this post is really about Freudian psychology. When I presented a Freudian friend with information on the general irrelevance of childhood factors and family composition, he countered by saying that the only thing that could really harm his belief in psychoanalysis was to learn that the unconscious wasn’t very powerful.

I don’t think we’re anywhere close to there yet. And I don’t think it’s meaningful to “deny the unconscious” – the unconscious is everything that happens except for conscious stuff, and that seems like a lot. But maybe our concept of the unconscious, or certain things that we attributed to the unconscious, was overly broad. And I think there’s an interesting project in trying to make explicit exactly what that means and what sort of smaller concepts we can get away with.

# Reverse Voxsplaining: Drugs vs. Chairs

[Content note: this is pretty much a rehash of things I’ve said before, and that other people have addressed much more eloquently. My only excuse for wasting your time with it again is that SOMEHOW THE MESSAGE STILL HASN’T SUNK IN. Pitching this as “market” vs. “government” is overly simplistic, but maybe if I am overly simplistic sometimes then it will sink in better.]

EpiPens, useful medical devices which reverse potentially fatal allergic reactions, have recently quadrupled in price, putting pressure on allergy sufferers and those who care for them. Vox writes that this “tells us a lot about what’s wrong with American health care” – namely that we don’t regulate it enough:

The story of Mylan’s giant EpiPen price increase is, more fundamentally, a story about America’s unique drug pricing policies. We are the only developed nation that lets drugmakers set their own prices, maximizing profits the same way sellers of chairs, mugs, shoes, or any other manufactured goods would.

Let me ask Vox a question: when was the last time that America’s chair industry hiked the price of chairs 400% and suddenly nobody in the country could afford to sit down? When was the last time that the mug industry decided to charge $300 per cup, and everyone had to drink coffee straight from the pot or face bankruptcy? When was the last time greedy shoe executives forced most Americans to go barefoot? And why do you think that is?

The problem with the pharmaceutical industry isn’t that they’re unregulated just like chairs and mugs. The problem with the pharmaceutical industry is that they’re part of a highly-regulated cronyist system that works completely differently from chairs and mugs.

If a chair company decided to charge $300 for their chairs, somebody else would set up a woodshop, sell their chairs for $250, and make a killing – and so on until chairs cost normal-chair-prices again. When Mylan decided to sell EpiPens for $300, in any normal system somebody would have made their own EpiPens and sold them for less. It wouldn’t have been hard. Its active ingredient, epinephrine, is off-patent, was being synthesized as early as 1906, and costs about ten cents per EpiPen-load.

Why don’t they? They keep trying, and the FDA keeps refusing to approve them for human use. For example, in 2009, a group called Teva Pharmaceuticals announced a plan to sell their own EpiPens in the US. The makers of the original EpiPen sued them, saying that they had patented the idea epinephrine-injecting devices. Teva successfully fended off the challenge and brought its product to the FDA, which rejected it because of “certain major deficiencies”. As far as I know, nobody has ever publicly said what the problem was – we can only hope they at least told Teva.

In 2010, another group, Sandoz, asked for permission to sell a generic EpiPen. Once again, the original manufacturers sued for patent infringement. According to Wikipedia, “as of July 2016 this litigation was ongoing”.

In 2011, Sanoji asked for permission to sell a generic EpiPen called e-cue. This got held up for a while because the FDA didn’t like the name (really!), but eventually was approved under the name Auvi-Q, (which if I were a giant government agency that rejected things for having dumb names, would be going straight into the wastebasket). But after unconfirmed reports of incorrect dosage delivery, they recalled all their products off the market.

This year, a company called Adamis decided that in order to get around the patent on devices that inject epinephrine, they would just sell pre-filled epinephrine syringes and let patients inject themselves. The FDA rejected it, noting that the company involved had done several studies but demanding that they do some more.

Also, throughout all of this a bunch of companies are merging and getting bought out by other companies and making secret deals with each other to retract their products and it’s all really complicated.

None of this is because EpiPens are just too hard to make correctly. Europe has eight competing versions. But aside from the EpiPen itself, only one competitor has ever made it past the FDA and onto the pharmacy shelf – a system called Adrenaclick.

And of course there’s a catch. With ordinary medications, pharmacists are allowed to interpret prescriptions for a brand name as prescriptions for the generic unless doctors ask them not to. For example, if I write a prescription for “Prozac”, a pharmacist knows that I mean anything containing fluoxetine, the chemical ingredient sold under the Prozac brand. They don’t have to buy it directly from Prozac trademark-holder Eli Lilly. It’s like if someone asks for a Kleenex and you give them a regular tissue, or if you suggest putting something in a Tupperware but actually use a plastic container made by someone other than the Tupperware Corporation.

EpiPens are protected from this substitution. If a doctor writes a prescription for “EpiPen”, the pharmacist must give an EpiPen-brand EpiPen, not an Adrenaclick-brand EpiPen. This is apparently so that children who have learned how to use an EpiPen don’t have to relearn how to use an entirely different device (hint: jam the pointy end into your body).

If you know anything at all about doctors, you know that they have way too much institutional inertia to change from writing one word on a prescription pad to writing a totally different word on a prescription pad, especially if the second word is almost twice as long, and especially especially if it’s just to do something silly like save a patient money. I have an attending who, whenever we are dealing with anything other than a life-or-death matter, just dismisses it with “Nobody ever died from X”, and I can totally hear him saying “Nobody ever died from paying extra for an adrenaline injector”. So Adrenaclick continues to languish in obscurity.

So why is the government having so much trouble permitting a usable form of a common medication?

There are a lot of different factors, but let me focus on the most annoying one. EpiPen manufacturer Mylan Inc spends about a million dollars on lobbying per year. OpenSecrets.org tells us what bills got all that money. They seem to have given the most to defeat S.214, the “Preserve Access to Affordable Generics Act”. The bill would ban pharmaceutical companies from bribing generic companies not to create generic drugs.

Did they win? Yup. In fact, various versions of this bill have apparently failed so many times that FDA Law Blog notes that “insanity is doing the same thing over and over again and expecting different result”.

So let me try to make this easier to understand.

Imagine that the government creates the Furniture and Desk Association, an agency which declares that only IKEA is allowed to sell chairs. IKEA responds by charging $300 per chair. Other companies try to sell stools or sofas, but get bogged down for years in litigation over whether these technically count as “chairs”. When a few of them win their court cases, the FDA shoots them down anyway for vague reasons it refuses to share, or because they haven’t done studies showing that their chairs will not break, or because the studies that showed their chairs will not break didn’t include a high enough number of morbidly obese people so we can’t be sure they won’t break. Finally, Target spends tens of millions of dollars on lawyers and gets the okay to compete with IKEA, but people can only get Target chairs if they have a note signed by a professional interior designer saying that their room needs a “comfort-producing seating implement” and which absolutely definitely does not mention “chairs” anywhere, because otherwise a child who was used to sitting on IKEA chairs might sit down on a Target chair the wrong way, get confused, fall off, and break her head.

(You’re going to say this is an unfair comparison because drugs are potentially dangerous and chairs aren’t – but 50 people die each year from falling off chairs in Britain alone and as far as I know nobody has ever died from an EpiPen malfunction.)

Imagine that this whole system is going on at the same time that IKEA spends millions of dollars lobbying senators about chair-related issues, and that these same senators vote down a bill preventing IKEA from paying off other companies to stay out of the chair industry. Also, suppose that a bunch of people are dying each year of exhaustion from having to stand up all the time because chairs are too expensive unless you’ve got really good furniture insurance, which is totally a thing and which everybody is legally required to have.

And now imagine that a news site responds with an article saying the government doesn’t regulate chairs enough.

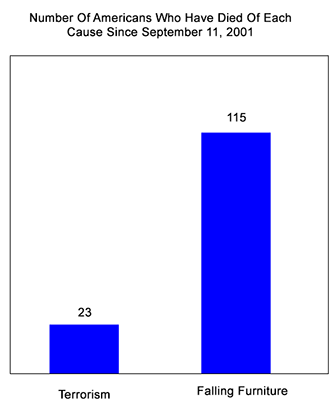
# Terrorists Vs. Chairs: An Outlier Story

The other day I needed to know how many people were killed by chairs, and while searching I came across the Washington Post’s You’re More Likely To Be Fatally Crushed By Furniture Than Killed By A Terrorist. It argues that worrying about terrorism is irrational, because terrorists kill fewer people each year than falling furniture, and nobody cares about that:

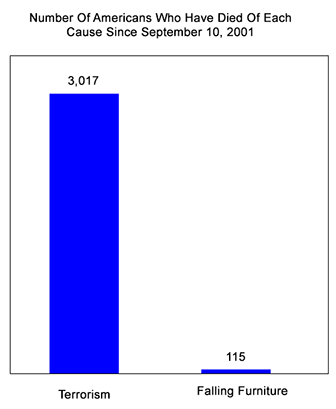
Consider, for instance, that since the attacks of Sept. 11, 2001, Americans have been no more likely to die at the hands of terrorists than being crushed to death by unstable televisions and furniture […] What accounts for the fear that terrorism inspires, considering that its actual risk in the United States and other Western countries is so low? The answer lies in basic human psychology.

I once saw the perfect response to this kind of argument on Twitter, but I forgot to screenshot it, so I’ll have to try to draw it from memory here.

One person posted a graph that looked something like this:



And somebody else edited it to look like this:



And whoa I had never realized before how sketchy it is to start your interval for recording the average number of terrorist attacks the day after the last major terrorist attack.

I mean, I know why people do it. It’s because September 11 was an “outlier”, and outliers should not be counted. Problem is, depending on your distribution, sometimes “outliers” are the only thing that matters.

Let me give an example. Suppose I’m trying to make an argument that earthquakes are totally not a problem for Haiti at all, that there’s no need to invest in earthquake preparedness, and that Haitian people who worry about earthquakes are stupid. I make a graph showing that since January 13, 2010, fewer Haitians have died per year from earthquake-related causes than from crazy furniture-related mishaps. This is totally 100% true. Look at those stupid Haitians, worrying about something that on average never hurts anybody!

(the Haitian earthquake of January 12, 2010 killed about 100,000 people)

I’m sure there are a zillion small Richter 1.0 and Richter 2.0 earthquakes in Haiti all the time. I’m sure our monitoring interval of January 13, 2010 to present picked up lots of these and correctly noted that they don’t kill anybody. The only Haitian earthquakes anyone needs to worry about are the outliers.

If you start your monitoring interval on January 13, earthquakes kill 0 people/year. If you start it on January 11, earthquakes kill 20,000 people a year. Neither of these is entirely fair – one is purpose-designed to maximize casualties, the other to minimize it. I don’t think there’s an obvious fair way to do things – the best solution would be extend the interval back to infinity, but then you get into problems like Haiti having fewer people back in the day, or Haiti not having risen out of the sea yet back in 4,000,000,000 BC. Maybe the best solution is to pick an arbitrary block of time like “the last fifty years”, or to do something very complicated like using the remote historical record to produce earthquake numbers and then combine it with modern populations to produce expected casualties.

The same is true of September 11. Start the interval September 12, and you get 5-10 terrorism deaths/year. Start it September 10, and you get 200. I don’t know when the best time to start it would be. If I had to choose something, I would say maybe 1985, when jihadist terrorism got started after the Soviet invasion of Afghanistan. But someone else could choose 1776, or 2000, based on similarly arbitrary criteria. And it would all be irrelevant – September 11 either made terrorists more ambitious, made security forces more watchful, or both, and so probably changed the calculus for good.

Granted, even when you include September 11, terrorism isn’t the worst thing, and people probably do overestimate it. So forget terrorism. On average, the flu kills something like 20,000 people worldwide each year. That’s a lot, but not apocalyptically much. If you go back year after year, the average stays at something like 20,000/year, right up until you get to 1918, when about 100,000,000 people died. So flu deaths over the last century average about 1 million/year. But three years from now, average flu deaths over the last century will average about 20,000 year. A death rate of only 20,000/year might make our current efforts to contain the flu seem excessive compared to other diseases. But a death rate of 1 million/year makes them look if anything the opposite.

Even worse: did you know that giant asteroids kill about a hundred people per year, on average? This is admittedly an odd definition of “kill” and “average” given that no human being has ever been killed by a giant asteroid. But given that giant asteroids strike Earth about every ten million years, and an asteroid strike today might kill about a billion people, on average giant asteroids kill about a hundred people per year.

Actually, un-forget terrorism. I have a friend who is very in favor of the War On Terror, and he argues that the problem with terrorism isn’t the average suicide bomber who kills three people. It isn’t even the 9-11 hijackers who killed three thousand people. It’s the group that steals a nuke and kills three million people. Just as “on average” a hundred people die each year from giant asteroid strikes, maybe “on average” thirty thousand people die each year from nuclear terrorism. All you’d need for this to be true is one nuclear attack per century. And that’s as bad as an average flu season!

The thing about falling furniture is that there’s probably not going to be a furniturepocalypse where suddenly millions of people all perish at once after being struck by a really really big desk. Furniture is constant. Terrorism isn’t. The whole point of black swans is that we pay too much attention to constant risks and ignore the outliers, especially the outliers which outlie so far that they haven’t happened yet. That’s true whether it’s terrorism, earthquakes, pandemics, or AI.

I worry that someday many years from now, terrorists are going to have some improbable victory which is even more destructive than September 11. I worry that uncounted people are going to die. And I worry that ten years later, someone is going to post on Facebook about how “From the day after ISIS nuked London through today, on average fewer people per year have died of terrorism than from hilarious accidents involving bedside dressers!”

# Reverse Voxsplaining: Brand-Name Drugs

[Epistemic status: Uncertain, especially on the accuracy of the economic studies cited]

I.

Sarah Kliff of Vox replies to my piece from last week. She writes:

Earlier this week I wrote about how a lack of drug price regulation in the United States allows pharmaceutical companies — including EpiPen’s manufacturer, Mylan — to charge exceptionally high prices for their products.

Scott Alexander at the blog Slate Star Codex argues that I’ve got it all wrong: The problem isn’t a lack of price regulation. Instead, its too much regulation, which has prevented generic competitors from entering the market and has left EpiPen’s price so high. […]

Alexander and I are both essentially pointing out two examples of how the United States has created a regulatory system that is incredibly favorable to pharmaceutical companies. We’ve set up a system that makes it incredibly easy for drug companies to score high profits and charge exceptionally high prices for their products.

One way it’s favorable is that we let drug companies pick their own prices — in this way the United States is exceptional, as the vast majority of developed companies regulate their drug prices. Another way we’ve created a favorable regulatory environment, as Alexander writes, is by allowing roadblocks to stand in the way of generic drug makers who want to enter.

More generics can help America’s drug spending problem. But they can’t solve it.

Greater use of generic drugs is widely accepted as a way to drive down drug spending. The FDA has found that drug prices decline to 55 percent of the brand-name price when two generic manufacturers begin making a product. Right now, the United States already uses a lot of generic drugs. In fact, about 90 percent of drugs prescribed in the country right now are generic. Brand name drugs are the reason that America has higher per-capita drug spending than other countries. Brand-name drugs make up just 10 percent of prescriptions filled in the United States, but account for 72 percent of drug spending.

Drugs that are under patent are the true source of high American drug costs. EpiPen is, in a way, a bit of a red herring […]

Harvoni, a [patented] pill that cures Hepatitis C, costs $32,114 here — and $22,554 in the United Kingdom. Less red tape around generic drug competition wouldn’t really change that fact. As long as we’re going to have patented drugs, letting drug manufacturers set their own prices will remain a key driver of America’s higher drug spending.

(the above is some excerpts stitched together to provide a taste; you should really read the whole thing)

First of all, thanks to Ms. Kliff and Vox for responding to me; it’s always neat to get featured in real news sources as if I were a real writer or something instead of just a guy with a blog. Additional thanks for a very measured and charitable tone despite my own tendency toward snarkiness. I worry that I am being unfair by acting snarky to somebody who is writing for a professional publication and so is not allowed to be snarky back; if so, I’m sorry and will try to control myself as best I can, which admittedly is not very good.

That having been said, I’m super against all of this and think it’s totally wrong.

I would kind of like to complain about Vox calling EpiPen a “red herring” when they were the ones who brought it up, but I think the problem is deeper than that. Discussing generic drug costs is completely different from discussing brand name drug costs, and the two issues have very different arguments around them. To transition fluidly from one to the other, saying we need more price controls for generics and then backing up your argument by saying that you were mostly thinking of brand-names after all – elides a distinction which is the heart of this entire subject.

Generic drugs are overpriced because we’re morons who can’t come up with a decent regulatory regime. Brand-name drugs are overpriced because of a deliberate decision to overprice them to encourage research.

The economic argument goes: the more profitable new drugs are, the more incentive a company has to make them. If we didn’t reward pharmaceutical companies for inventing new drugs, then they wouldn’t go through the $2.5 billion, ten-year hassle of seeking FDA approval with no guarantee of success. The way we reward them is by giving them a twenty-year monopoly when they can charge lots of money without anybody telling them not to.

(this isn’t quite right. The law says a twenty-year monopoly, but it’s dated from the time the drug is invented. Since it takes ten years to go through the FDA, it’s effectively more like a ten-year monopoly on actually selling the drug)

The reason I usually limit my griping about pharmaceutical overpricing to generics and avoid brand-names is that while high generic pricing is inexcusable, high brand-name pricing is debatably useful. Some people would say that the benefit of encouraging more drug development is worth the cost of higher prices, other people would say that it isn’t. I didn’t want to wade into this complicated debate.

But I guess now I have to. So. Lit review time. I searched the economics literature for studies, models, and arguments used to calculate whether price regulation would decrease drug development, and if so, how the benefits and risks balanced out. Here’s what I’ve got:

1. Golec & Vernon (2006) say that as a result of European drug price regulation, “EU consumers enjoyed much lower pharmaceutical price inflation, however, at a cost of 46 fewer new medicines introduced by EU firms.”

2. Eger and Mahlich (2014) find that among pharmaceutical companies, “a higher presence in Europe is associated with lower R&D investments. The results can be interpreted as further evidence of the deteriorating effect of regulation on firm’s incentives to invest in R&D.”

3. Kutyavina (2010) finds that “brand-name pharmaceutical firms characterized by large R&D expenditures decreased their R&D efforts post 1993 threat [to regulate drug prices] relative to firms that did not engage in as much innovative R&D”.

4. Acemoglu and Linn (2004) find that “We find a large effect of potential market size on the entry of nongeneric drugs and new molecular entities”, which I think is supposed to generalize to mean that the more money they expect to make the more research they do. I will count this as half a study since the connection is not explicit.

5. Danzon & Epstein (2008) analyze price regulations and new drugs invented in 15 countries and 12 drug classes, and find that “If price regulation reduces drug prices, it contributes to launch delay in the home country.

6. Troyer & Krasnikov (2002) find that “the empirical relationship between pharmaceutical industry revenues and pharmaceutical industry innovation is estimated, allowing for an exploration of the impact of the Medicaid rebate program [which regulated drug prices somewhat]. Using the empirical results, the opportunity cost of the Medicaid rebate program is found to be as high as four new drug approvals annually. Given the increased interest in a Medicare drug benefit, regulators should be aware of the hidden cost of price regulation for pharmaceuticals.”

7. Vernon (2005) finds that “I simulate how a new policy regulating pharmaceutical prices in the US will affect R&D investment. I find that such a policy will lead to a decline in industry R&D by between 23.4% and 32.7%. This prediction, however, is accompanied by several caveats.”

8. Golec, Hegde, and Vernon (2009) find that “Results show that the HSA [a bill to regulate drug spending in the US] had significant negative effects on stock prices and firm-level R&D spending. Conservatively, the HSA reduced R&D spending by about $1 billion even though it never became law.”

9. Santerre and Vernon (2006) use drug demand data to simulate various regulatory regimes, and find that a certain price regulation policy they test, continued over twenty years, would have cost gains of $472 billion (!) but also “have led to 198 new drugs being brought to the US market” (!!). They note that “Therefore, the average social opportunity cost per drug developed during this period was approximately $2.4 billion. Research on the value of pharmaceuticals suggests that the social benefits of a new drug are far greater than this estimate. Hence, drug price controls could do more harm than good.”

10. Keyhani, Carpenter, et al (2010) find that “The United States accounted for 42% of prescription drug spending and 40% of the total GDP among innovator countries and was responsible for the development of 43.7% of the NMEs [ie new drugs invented]. The United Kingdom, Switzerland, and a few other countries innovated proportionally more than their contribution to GDP or prescription drug spending, whereas Japan, South Korea, and a few other countries innovated less…higher prescription drug spending in the US does not disproportionately privilege domestic innovation, and many countries with drug price regulation were significant contributors to pharmaceutical innovation.” This study does not attempt to address the effects of price regulation, only to say that European countries seem to do pretty well at innovation despite price regulation, which is suggestive that price regulation does not hurt drug innovation but not really scientific evidence for it. I’m going to count this one as half a study too.

So by my count, there are eight-and-a-half studies concluding that price regulation would hurt new drug innovation, and one-half of a study concluding that it wouldn’t. I’ve tried to eliminate all the studies sponsored by the pharmaceutical industry from this list, but I might have missed some, and I am always skeptical of anything that says anything the pharmaceutical industry approves of even I can’t trace the money directly.

One source I do trust is RAND, a think tank which is generally well-respected and pretty objective (despite the name, they are not associated with Ayn Rand or Rand Paul). In Regulating Drug Prices: US Policy Alternatives In A Global Context, they write:

U.S. consumers spend roughly twice as much on drugs as their European counterparts….Pressure is building in U.S. policy circles for the federal government to take action to regulate the cost of drugs. At the same time, there is debate about the pros and cons of doing so…To shed light on this debate, a team of RAND researchers examined the impact of drug price regulation…The results showed that:

— Globally, the regulation of pharmaceutical prices has increased in recent years.

— In most cases, regulation reduces pharmaceutical revenues.

— Regulatory approaches that reduce pharmaceutical revenues may generate modest consumer savings in the best cases, but risk much larger costs as decreased innovation leads to reductions in life expectancy.

In other words, such prices would be good in the short term as we get all the currently-existing drugs for very cheap:

Annual per capita spending in 2010 would fall for Americans age 55–59 by an amount in the range of $9,000 annually and for Europeans of the same age by an amount in the range of $400.

But bad in the long-term as pharmaceutical innovation declines and we have fewer interventions available to protect our health:

Life expectancy would fall by somewhere in the range of 0.2 years for Americans age 55–59 in 2010 and by approximately 0.1 years for Europeans in the same cohort and year. By 2060, this effect would increase for both Americans and Europeans to approximately 0.7 years.

Given the value they place on human life, they argue that this money-for-life-years trade is net negative:

Overall, as shown in Figure 1, the net value of price controls is positive in the short term (2010) for Americans age 55–59, producing approximately $1,100 in per capita savings, but negative for Europeans in the same cohort and year, who face increased costs in the range of $8,000. For both Americans and Europeans, price controls have higher per capita costs over the longer term: By 2060, reductions in life expectancy, after accounting for medical cost savings, would cost the equivalent of $51,000 and $54,000, to age 55–59 Americans and Europeans, respectively.

All of this sounds sort of boring and economics-y when you read it like this, and maybe your eyes are glazing over. So let me put this in context. In 2060 there will probably be 420 million Americans and 523 million Europeans. And suppose that whatever changes we make in drug regulations today last for one human lifespan, so that everybody has a chance to be 55-60. So about a billion people each losing about 0.7 years of their life equals 700 million life-years. Since some people live in countries outside the US and Europe [citation needed] and they also benefit from First-World-invented medications, let’s round this up to about a billion life-years lost.

What was the worst thing that ever happened? One strong contender is Mao’s Great Leap Forward, in which ineffective agricultural reforms and very effective purges killed 45 million people. Most of these people were probably already adults, and lifespan in Mao’s China wasn’t too high, so let’s say that each death from the Great Leap Forward cost what would otherwise be twenty healthy life years. In that case, the worst thing that has ever happened until now cost 45 million \* 20 = 900 million life-years.

Once again, RAND’s calculations plus my own Fermi estimate suggest that prescription drug price regulation would cost one billion life-years, which would very slightly edge out Communist China for the title of Worst Thing Ever.

Am I exaggerating or being facetious? I’m actually not sure. Dammit, Jim I’m a doctor, not an economist. I’m not qualified to analyze any of those ten studies above beyond a quick check to see if they’re completely ridiculous. I’m not qualified to say if RAND is right or wrong to estimate a cost of 0.7 life-years, or whether I’m misusing their calculation to try to add up exactly how bad it would be. Maybe a real economist will look at this whole essay and say it’s really stupid. I don’t know.

The only thing I can say in my own defense is that I am acknowledging that the question exists. I am not at all sure that Vox has reached this level yet. They just wrote an article on price regulations for brand-name drugs which, first, mixed them liberally with generic drugs despite the different arguments around each, and second, didn’t mention anything about research or innovation. Call me overly demanding, but when you are proposing a policy which most economists think would decrease the rate of life-saving medical progress, and which by some calculations might edge out Mao’s China for the title of most disastrous and deadly thing in all of human history, then I feel like you should acknowledge, at least in a single sentence, that somebody has claimed, at least once, that the policy might have some slight downside. At least don’t act as if it’s the same issue as a different kind of drug regulation which doesn’t have that downside.

I think it’s an unfortunate omission to talk about the EpiPen cost increase as relating only to lack of price controls, and fail to mention the reason why this happens with EpiPen but never chairs or mugs. And I think it’s a further omission to talk about regulating brand-name medications but fail to mention that some people think it will backfire and impede innovation. While I appreciate the effort to say we’re both on the same team of reducing drug costs, I’m a little concerned about my teammates’ strategy here.

And there’s another way we’re not quite on the same team. I’m on Team Left-Libertarian, which luckily is so confusing and contradictory that I can define it however I want. And today it means that while I’m not opposed to all regulation in principle, I at least get really scared when somebody pushes for regulation today and promises to check whether it will have bad consequences tomorrow. I think that’s how we got in this mess where the generics industry is so regulated that EpiPens cost hundreds of dollars, and even if Vox and I are on the same object-level team of Make Epi-Pens Cost Less, I worry we are not on the same meta-level team of Learn From The Fact That Epi-Pens Cost So Much And Worry That The Same Kind Of Thinking That Caused The Epi-Pen Problem Will Probably Cause Other Problems Too.

II.

So do I have a solution to the high price of brand-name drugs? Well, I have a partial, unsatisfying solution. But first, a digression.

Vox gives the example of Harvoni costing $32,000 in the United States, but only $22,000 in the United Kingdom. This is supposed to be an example of the United Kingdom’s drug price regulation system working. I guess this is good, but you may notice that both numbers are really really high. There’s generic Harvoni available in India for $900. I can’t find how much it costs to manufacture, but reading between the lines and looking at some similar compounds, it’s probably about $100. So good work, Britain. You’re paying $22,000 instead of $32,000 for a $100 pill.

There are a couple of morals to this story. The first is that Vox’s claim that generics made by two competing companies cost 55% of the brand-name price isn’t the right statistic to use here. Look at their source and you find that as number of competing companies gets to 20, generics cost 5% of brand names. As number of competitors approaches infinity, drug cost should approach manufacturing cost, which can be very low – in the cast of Harvoni, less than one percent. This seems true in the case of modafinil, which I’ve talked about before; it costs about $25 per pill in the US and more like $2 per pill in more generic-friendly India.

So the second moral of the story is that almost all gains in prescription drug prices are to be found not in price regulation bringing prices down from $32,000 to $22,000, but in switching from monopoly brand-name drugs that cost $32,000 to heavily-competitive generic drugs that cost $100.

In a lot of cases, this is easier than you would think.

Pristiq is the brand-name of desvenlafaxine, a new antidepressant which is still brand-name only. Desvenlafaxine sounds a lot like venlafaxine – which is Effexor, an old antidepressant which is available in generic. In fact, desvenlafaxine is a tiny change to the venlafaxine molecule which may or may not have any interesting medical benefit over the original, and which was invented solely to have something whose patent hasn’t expired.

Wyeth, the company that makes Pristiq, says that it’s better than Effexor because it doesn’t have as many drug-drug interactions. But Effexor doesn’t really have clinically significant drug-drug interactions, and this seems to be them just saying random stuff and hoping people believe them. There are no good head-to-head studies comparing Pristiq to Effexor, but if you try to piece together a comparison from unrelated studies (not recommended, but we’ll do it anyway) Effexor actually seems better than its newer cousin. Even the data I took from drug rating databases shows patients preferring Effexor to Pristiq by quite a lot. Carlat Psychiatry, which is psychiatrists’ insider news site on pharmacology developments, has a blog post called Top Five Reasons To Forget About Pristiq. Most of the well-informed psychiatrists I know agree that Pristiq is a slightly worse version of an older antidepressant with no proven advantages.

A month’s supply of Effexor costs $20. A month’s supply of Pristiq costs $300. So let me amend the paragraph above. Pristiq is a slightly worse version of an older antidepressant with no proven advantages that also costs fifteen times as much.

It should come as no surprise to anyone familiar with the state of psychiatry that it is the second most-prescribed antidepressant in the USA, with three million prescriptions per month.

Why would this happen? The relevant study is called Pharmaceutical Industry-Sponsored Meals And Physician Prescribing Patterns For Medicare Beneficiaries, so you know it’s going to be good. It shows that doctors who often eat drug-company-sponsored free lunches are more than twice as likely to prescribe Pristiq as doctors who rarely eat such lunches. This matches my observations perfectly. Doctors prescribe Pristiq because they don’t know very much about antidepressants, but they attend free lunches by pharmaceutical companies who tell them that Pristiq is great, and they believe it. If this surprises you, be more cynical.

I’m looking at the price of Pristiq in Canada, and it seems to range around $120 to $250. So if we instituted price regulations like Canada’s, we might lower the cost of Pristiq from $300 to $150. If we convinced doctors to prescribe Effexor instead, it would be $20, plus I really do believe Effexor is genuinely better.

Pristiq is far from alone in this. I don’t have good statistics, but I bet that at least half of brand-name prescriptions in the US are more like Pristiq (attempts to rip people off) than like Harvoni (genuinely wonderful breakthroughs in medical science).

So one of the best ways to deal with expensive brand-name drugs is to stop using expensive brand name drugs for no reason. Since I get to define what left-libertarianism means however I want, I will say that it is provisionally okay with banning pharmaceutical companies from buying doctors lunch, as long as there aren’t any studies concluding that this would kill more people than Communist China. There are probably lots of other ways to improve medical education and medical economics so that doctors are less easily bamboozled into prescribing these, but those can wait for other blog posts.

What about the genuinely novel brand-name drugs, the ones like Harvoni that really are better than anything that came before?

An optimistic answer: maybe after we stop spending our civilizational resources on Pristiq, we’ll have a little more money to afford them, and maybe we’ll be happy to subsdize the genuinely awesome pharmaceutical research that remains.

Another optimistic answer: once FDA regulatory requirements are loosened, there will be a wide selection of different brand-name drugs. For example, even when Prozac was a brand-name, pharma’s ability to inflate its price was limited by the existence of several very similar brand-name drugs, like Paxil and Zoloft. If there are twenty competing brand-name hepatitis wonder drugs – and I don’t think that’s outside the realm of what we can hope for – then I think that will tend to lower prices to cost. This would include the cost of research and licensing, and so still be pretty high, but as long as research is a real unavoidable cost it would probably be the best we can hope for.

(this would be a good time to bring up that chlorcyclizine costs fifty cents per pill and might work as well as Harvoni)

The pessimistic answer is that all we can do is ensure that the generics marketplace is fair and competitive. And then rich people can buy Harvoni now for $30,000, and poor people will have to wait ten years to buy Harvoni when it costs $100. Right now they’ll unfortunately have to figure out how to make do with the set of medications invented in 2006 and before.

And I know this is terrible, especially if someone has a disease with only one cure and it was invented after 2006. But think of it this way. This objection, rephrased, is that 2016 has more drugs available than 2006, and we want to maximize the number of new drugs available to the poorest patients. But if we try to do that by instituting price controls which decreases the rate of drug innovation long-term, then we end up decreasing the number of new drugs available to the poorest patients, exactly the opposite of what we thought we were doing.

Let me give an example. According to study (9) above, price controls would have caused about 200 fewer drugs to be approved over the period from 1980 – 2000. In fact about 600 drugs were approved during that period. So if they’re right, it would have cut the innovation rate by 1/3. That means that in Hypothetical Price Control World’s 2016, after 36 years of price controls, we would only have 24 of our years’ worth of drugs – ie, the drugs that we had in 2004 in our own world. But since drugs usually go off patent about ten years after approval, in fact we’ve genericized the drugs that we had in 2006 in our own world. So we have more drugs available just as generics than Hypothetical Price Control World has as generics and brand-names combined. If poor people can afford only generics or price-controlled brand names, our poor people are better off than Hypothetical Price Control World’s poor people (and our rich people are much better off than Hypothetical Price-Control World’s rich people). And as time goes on, our advantage over their world will only get bigger.

Maybe there are better ways to do this. Some people have talked about funding research via “prizes” rather than through an investment-and-profit model. Some people say we should fund it publicly through the NIH or something, which we already sort of do to a degree. Still other people say that we should abolish the FDA, cut the costs of drug development by an order of magnitude, and, um, see what happens. I don’t know about any of those things. I just feel like until you’re ready to set these up and have some idea that they work, do the thing that probably is going to result in people having the best access to the most life-saving drugs. Which right now looks like no price control.

Or maybe I am completely wrong about all of this. I am not an economist and have to take these studies at face value, and anything that touches pharmaceutical companies ends up being corrupted and full of lies. Maybe I myself am saying something very stupid that will end up killing more people than Communist China. If so I certainly hope that people who know more than I do will tell me why these calculations were wrong and how to look at this situation better.

But this is the level at which I think this discussion needs to be had.

# It’s Bayes All The Way Up

[Epistemic status: Very speculative. I am not a neuroscientist and apologize for any misinterpretation of the papers involved. Thanks to the people who posted these papers in r/slatestarcodex. See also Mysticism and Pattern-Matching and Bayes For Schizophrenics]

Bayes’ Theorem is an equation for calculating certain kinds of conditional probabilities. For something so obscure, it’s attracted a surprisingly wide fanbase, including doctors, environmental scientists, economists, bodybuilders, fen-dwellers, and international smugglers. Eventually the hype reached the point where there was both a Bayesian cabaret and a Bayesian choir, popular books using Bayes’ Theorem to prove both the existence and the nonexistence of God, and even Bayesian dating advice. Eventually everyone agreed to dial down their exuberance a little, and accept that Bayes’ Theorem might not literally explain absolutely everything.

So – did you know that the neurotransmitters in the brain might represent different terms in Bayes’ Theorem?

First things first: Bayes’ Theorem is a mathematical framework for integrating new evidence with prior beliefs. For example, suppose you’re sitting in your quiet suburban home and you hear something that sounds like a lion roaring. You have some prior beliefs that lions are unlikely to be near your house, so you figure that it’s probably not a lion. Probably it’s some weird machine of your neighbor’s that just happens to sound like a lion, or some kids pranking you by playing lion noises, or something. You end up believing that there’s probably no lion nearby, but you do have a slightly higher probability of there being a lion nearby than you had before you heard the roaring noise. Bayes’ Theorem is just this kind of reasoning converted to math. You can find the long version here.

This is what the brain does too: integrate new evidence with prior beliefs. Here are some examples I’ve used on this blog before:







All three of these are examples of top-down processing. Bottom-up processing is when you build perceptions into a model of the the world. Top-down processing is when you let your models of the world influence your perceptions. In the first image, you view the center letter of the the first word as an H and the second as an A, even though they’re the the same character; your model of the world tells you that THE CAT is more likely than TAE CHT. In the second image, you read “PARIS IN THE SPRINGTIME”, skimming over the duplication of the word “the”; your model of the world tells you that the phrase should probably only have one “the” in it (just as you’ve probably skimmed over it the three times I’ve duplicated “the” in this paragraph alone!). The third image might look meaningless until you realize it’s a cow’s head; once you see the cow’s head your model of the world informs your perception and it’s almost impossible to see it as anything else.

(Teh fcat taht you can siltl raed wrods wtih all the itroneir ltretrs rgraneanrd is ahonter empxlae of top-dwon pssirocneg mkinag nsioy btotom-up dtaa sanp itno pacle)

But top-down processing is much more omnipresent than even these examples would suggest. Even something as simple as looking out the window and seeing a tree requires top-down processing; it may be too dark or foggy to see the tree one hundred percent clearly, the exact pattern of light and darkness on the tree might be something you’ve never seen before – but because you know what trees are and expect them to be around, the image “snaps” into the schema “tree” and you see a tree there. As usual, this process is most obvious when it goes wrong; for example, when random patterns on a wall or ceiling “snap” into the image of a face, or when the whistling of the wind “snaps” into a voice calling your name.

Most of the things you perceive when awake are generated from very limited input – by the same machinery that generates dreams with no input

— Void Of Space (@VoidOfSpace) September 2, 2016

Corlett, Frith & Fletcher (2009) (henceforth CFF) expand on this idea and speculate on the biochemical substrates of each part of the process. They view perception as a “handshake” between top-down and bottom-up processing. Top-down models predict what we’re going to see, bottom-up models perceive the real world, then they meet in the middle and compare notes to calculate a prediction error. When the prediction error is low enough, it gets smoothed over into a consensus view of reality. When the prediction error is too high, it registers as salience/surprise, and we focus our attention on the stimulus involved to try to reconcile the models. If it turns out that bottom-up was right and top-down was wrong, then we adjust our priors (ie the models used by the top-down systems) and so learning occurs.

In their model, bottom-up sensory processing involves glutamate via the AMPA receptor, and top-down sensory processing involves glutamate via the NMDA receptor. Dopamine codes for prediction error, and seem to represent the level of certainty or the “confidence interval” of a given prediction or perception. Serotonin, acetylcholine, and the others seem to modulate these systems, where “modulate” is a generic neuroscientist weasel word. They provide a lot of neurological and radiologic evidence for these correspondences, for which I highly recommend reading the paper but which I’m not going to get into here. What I found interesting was their attempts to match this system to known pharmacological and psychological processes.

CFF discuss a couple of possible disruptions of their system. Consider increased AMPA signaling combined with decreased NMDA signaling. Bottom-up processing would become more powerful, unrestrained by top-down models. The world would seem to become “noisier”, as sensory inputs took on a life of their own and failed to snap into existing categories. In extreme cases, the “handshake” between exuberant bottom-up processes and overly timid top-down processes would fail completely, which would take the form of the sudden assignment of salience to a random stimulus.

Schizophrenics are famous for “delusions of reference”, where they think a random object or phrase is deeply important for reasons they have trouble explaining. Wikipedia gives as examples:

– A feeling that people on television or radio are talking about or talking directly to them

– Believing that headlines or stories in newspapers are written especially for them

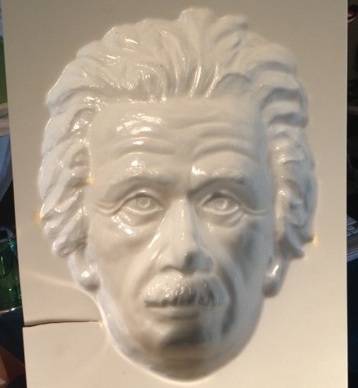
– Seeing objects or events as being set up deliberately to convey a special or particular meaning to themselves

– Thinking ‘that the slightest careless movement on the part of another person had great personal meaning…increased significance’

In CFF, these are perceptual handshake failures; even though “there’s a story about the economy in today’s newspaper” should be perfectly predictable, noisy AMPA signaling registers it as an extreme prediction failure, and it fails its perceptual handshake with overly-weak priors. Then it gets flagged as shocking and deeply important. If you’re unlucky enough to have your brain flag a random newspaper article as shocking and deeply important, maybe phenomenologically that feels like it’s a secret message for you.

And this pattern – increased AMPA signaling combined with decreased NMDA signaling – is pretty much the effect profile of the drug ketamine, and ketamine does cause a paranoid psychosis mixed with delusions of reference.

Organic psychosis like schizophrenia might involve a similar process. There’s a test called the binocular depth inversion illusion, which looks like this:



(source)

The mask in the picture is concave, ie the nose is furthest away from the camera. But most viewers interpret it as convex, with the nose closest to the camera. This makes sense in terms of Bayesian perception; we see right-side-in faces a whole lot more often than inside-out faces.

Schizophrenics (and people stoned on marijuana!) are more likely to properly identify the face as concave than everyone else. In CFF’s system, something about schizophrenia and marijuana messes with NMDA, impairs priors, and reduces the power of top-down processing. This predicts that schizophrenics and potheads would both have paranoia and delusions of reference, which seems about right.

Consider a slightly different distortion: increased AMPA signaling combined with increased NMDA signaling. You’ve still got a lot of sensory noise. But you’ve also got stronger priors to try to make sense of them. CFF argue these are the perfect conditions to create hallucinations. The increase in sensory noise means there’s a lot of data to be explained; the increased top-down pattern-matching means that the brain is very keen to fit all of it into some grand narrative. The result is vivid, convincing hallucinations of things that are totally not there at all.

LSD is mostly serotonergic, but most things that happen in the brain bottom out in glutamate eventually, and LSD bottoms out in exactly the pattern of increased AMPA and increased NMDA that we would expect to produce hallucinations. CFF don’t mention this, but I would also like to add my theory of pattern-matching based mysticism. Make the top-down prior-using NMDA system strong enough, and the entire world collapses into a single narrative, a divine grand plan in which everything makes sense and you understand all of it. This is also something I associate with LSD.

If dopamine represents a confidence interval, then increased dopaminergic signaling should mean narrowed confidence intervals and increased certainty. Perceptually, this would correspond to increased sensory acuity. More abstractly, it might increase “self-confidence” as usually described. Amphetamines, which act as dopamine agonists, do both. Amphetamine users report increased visual acuity (weirdly, they also report blurred vision sometimes; I don’t understand exactly what’s going on here). They also create an elevated mood and grandiose delusions, making users more sure of themselves and making them feel like they can do anything.

(something I remain confused about: elevated mood and grandiose delusions are also typical of bipolar mania. People on amphetamines and other dopamine agonists act pretty much exactly like manic people. Antidopaminergic drugs like olanzapine are very effective acute antimanics. But people don’t generally think of mania as primarily dopaminergic. Why not?)

CFF end their paper with a discussion of sensory deprivation. If perception is a handshake between bottom-up sense-data and top-down priors, what happens when we turn the sense-data off entirely? Psychologists note that most people go a little crazy when placed in total sensory deprivation, but that schizophrenics actually seem to do better under sense-deprivation conditions. Why?

The brain filters sense-data to adjust for ambient conditions. For example, when it’s very dark, your eyes gradually adjust until you can see by whatever light is present. When it’s perfectly silent, you can hear the proverbial pin drop. In a state of total sensory deprivation, any attempt to adjust to a threshold where you can detect the nonexistent signal is actually just going to bring you down below the point where you’re picking up noise. As with LSD, when there’s too much noise the top-down systems do their best to impose structure on it, leading to hallucinations; when they fail, you get delusions. If schizophrenics have inherently noisy perceptual systems, such that all perception comes with noise the same way a bad microphone gives off bursts of static whenever anyone tries to speak into it, then their brains will actually become less noisy as sense-data disappears.

(this might be a good time to remember that no congentally blind people ever develop schizophrenia and no one knows why)

II.

Lawson, Rees, and Friston (2014) offer a Bayesian link to autism.

(there are probably a lot of links between Bayesians and autism, but this is the only one that needs a journal article)

They argue that autism is a form of aberrant precision. That is, confidence intervals are too low; bottom-up sense-data cannot handshake with top-down models unless they’re almost-exactly the same. Since they rarely are, top-down models lose their ability to “smooth over” bottom-up information. The world is full of random noise that fails to cohere into any more general plan.

Right now I’m sitting in a room writing on a computer. A white noise machine produces white noise. A fluorescent lamp flickers overhead. My body is doing all sorts of body stuff like digesting food and pumping blood. There are a few things I need to concentrate on: this essay I’m writing, my pager if it goes off, any sorts of sudden dramatic pains in my body that might indicate a life-threatening illness. But I don’t need to worry about the feeling of my back against the back fo the chair, or the occasional flickers of the fluorescent light, or the feeling of my shirt on my skin.

A well-functioning perceptual system gates out those things I don’t need to worry about. Since my shirt always feels more or less similar on my skin, my top-down model learns to predict that feeling. When the top-down model predicts the shirt on my skin, and my bottom-up sensation reports the shirt on my skin, they handshake and agree that all is well. Even if a slight change in posture makes a different part of my shirt brush against my skin than usual, the confidence intervals are wide: it is still an instance of the class “shirt on skin”, it “snaps” into my shirt-on-skin schema, and the perceptual handshake goes off successfully, and all remains well. If something dramatic happens – for example my pager starts beeping really loudly – then my top-down model, which has thus far predicted silence – is rudely surprised by the sudden burst of noise. The perceptual handshake fails, and I am startled, upset, and instantly stop writing my essay as I try to figure out what to do next (hopefully answer my pager). The system works.

The autistic version works differently. The top-down model tries to predict the feeling of the shirt on my skin, but tiny changes in the position of the shirt change the feeling somewhat; bottom-up data does not quite match top-down prediction. In a neurotypical with wide confidence intervals, the brain would shrug off such a tiny difference, declare it good enough for government work, and (correctly) ignore it. In an autistic person, the confidence intervals are very narrow; the top-down systems expect the feeling of shirt-on-skin, but the bottom-up systems report a slightly different feeling of shirt-on-skin. These fail to snap together, the perceptual handshake fails, and the brain flags it as important; the autistic person is startled, upset, and feels like stopping what they’re doing in order to attend to it.

(in fact, I think the paper might be claiming that “attention” just means a localized narrowing of confidence intervals in a certain direction; for example, if I pay attention to the feeling of my shirt on my skin, then I can feel every little fold and micromovement. This seems like an important point with a lot of implications.)

Such handshake failures match some of the sensory symptoms of autism pretty well. Autistic people dislike environments that are (literally or metaphorically) noisy. Small sensory imperfections bother them. They literally get annoyed by scratchy clothing. They tend to seek routine, make sure everything is maximally predictable, and act as if even tiny deviations from normal are worthy of alarm.

They also stim. LRF interpret stimming as an attempt to control sensory predictive environment. If you’re moving your arms in a rhythmic motion, the overwhelming majority of sensory input from your arm is from that rhythmic motion; tiny deviations get lost in the larger signal, the same way a firefly would disappear when seen against the blaze of a searchlight. The rhythmic signal which you yourself are creating and keeping maximally rhythmic is the most predictable thing possible. Even something like head-banging serves to create extremely strong sensory data – sensory data whose production the head-banger is themselves in complete control of. If the brain is in some sense minimizing predictive error, and there’s no reasonable way to minimize prediction error because your predictive system is messed up and registering everything as a dangerous error – then sometimes you have to take things into your own hands, bang your head against a metal wall, and say “I totally predicted all that pain”.

(the paper doesn’t mention this, but it wouldn’t surprise me if weighted blankets work the same way. A bunch of weights placed on top of you will predictably stay there; if they’re heavy enough this is one of the strongest sensory signals you’re receiving and it might “raise your average” in terms of having low predictive error)

What about all the non-sensory-gating-related symptoms of autism? LRF think that autistic people dislike social interaction because it’s “the greatest uncertainty”; other people are the hardest-to-predict things we encounter. Neurotypical people are able to smooth social interaction into general categories: this person seems friendly, that person probably doesn’t like me. Autistic people get the same bottom-up data: an eye-twitch here, a weird half-smile there – but it never snaps into recognizable models; it just stays weird uninterpretable clues. So:

This provides a simple explanation for the pronounced social-communication difficulties in autism; given that other agents are arguably the most difficult things to predict. In the complex world of social interactions, the many-to-one mappings between causes and sensory input are dramatically increased and difficult to learn; especially if one cannot contextualize the prediction errors that drive that learning.

They don’t really address differences between autists and neurotypicals in terms of personality or skills. But a lot of people have come up with stories about how autistic people are better at tasks that require a lot of precision and less good at tasks that require central coherence, which seems like sort of what this theory would predict.

LRF ends by discussing biochemical bases. They agree with CFF that top-down processing is probably related to NMDA receptors, and so suspect this is damaged in autism. Transgenic mice who lack an important NMDA receptor component seem to behave kind of like autistic humans, which they take as support for their model – although obviously a lot more research is needed. They agree that acetylcholine “modulates” all of this and suggest it might be a promising pathway for future research. They agree with CFF that dopamine may represent precision/confidence, but despite their whole spiel being that precision/confidence is messed up in autism, they don’t have much to say about dopamine except that it probably modulates something, just like everything else.

III.

All of this is fascinating and elegant. But is it elegant enough?

I notice that I am confused about the relative role of NMDA and AMPA in producing hallucinations and delusions. CFF say that enhanced NMDA signaling results in hallucinations as the brain tries to add excess order to experience and “overfits” the visual data. Fine. So maybe you get a tiny bit of visual noise and think you’re seeing the Devil. But shouldn’t NMDA and top-down processing also be the system that tells you there is a high prior against the Devil being in any particular visual region?

Also, once psychotics develop a delusion, that delusion usually sticks around. It might be that a stray word in a newspaper makes someone think that the FBI is after them, but once they think the FBI is after them, they fit everything into this new paradigm – for example, they might think their psychiatrist is an FBI agent sent to poison them. This sounds a lot like a new, very strong prior! Their doctor presumably isn’t doing much that seems FBI-agent-ish, but because they’re working off a narrative of the FBI coming to get them, they fit everything, including their doctor, into that story. But if psychosis is a case of attenuated priors, why should that be?

(maybe they would answer that because psychotic people also have increased dopamine, they believe in the FBI with absolute certainty? But then how come most psychotics don’t seem to be manic – that is, why aren’t they overconfident in anything except their delusions?)

LRF discuss prediction error in terms of mild surprise and annoyance; you didn’t expect a beeping noise, the beeping noise happened, so you become startled. CFF discuss prediction error as sudden surprising salience, but then say that the attribution of salience to an odd stimulus creates a delusion of reference, a belief that it’s somehow pregnant with secret messages. These are two very different views of prediction error; an autist wearing uncomfortable clothes might be constantly focusing on their itchiness rather than on whatever she’s trying to do at the time, but she’s not going to start thinking they’re a sign from God. What’s the difference?

Finally, although they highlighted a selection of drugs that make sense within their model, others seem not to. For example, there’s some discussion of ampakines for schizophrenia. But this is the opposite of what you’d want if psychosis involved overactive AMPA signaling! I’m not saying that the ampakines for schizophrenia definitely work, but they don’t seem to make the schizophrenia noticeably worse either.

Probably this will end the same way most things in psychiatry end – hopelessly bogged down in complexity. Probably AMPA does one thing in one part of the brain, the opposite in other parts of the brain, and it’s all nonlinear and different amounts of AMPA will have totally different effects and maybe downregulate itself somewhere else.

Still, it’s neat to have at least a vague high-level overview of what might be going on.

# Some Context For That NYT Sugar Article

Imagine a political historian discovers that Lyndon Johnson accepted a campaign contribution from a big Wall Street bank. Since Johnson’s policies helped shape the modern Democratic Party, everyone agrees the Democrats are built on a foundation of lies. “Republicans Vindicated; Small-Government Conservativism Was Right All Along”, say the headlines of all the major newspapers.

This is kind of how I feel about the reaction to the latest New York Times article.

How The Sugar Industry Shifted Blame To Fat describes new historical research that finds that the sugar industry sponsored a study showing that fat (and not sugar) was the major risk factor for cardiovascular disease. They tie this into a bigger narrative about how sugar is the real dietary villain, and it’s only the sugar industry’s successful bribery work that made us suspect fat for so long:

The revelations are important because the debate about the relative harms of sugar and saturated fat continues today, Dr. Glantz said. For many decades, health officials encouraged Americans to reduce their fat intake, which led many people to consume low-fat, high-sugar foods that some experts now blame for fueling the obesity crisis.

“It was a very smart thing the sugar industry did, because review papers, especially if you get them published in a very prominent journal, tend to shape the overall scientific discussion,” he said […]

I’m glad researchers have discovered this. But treating it as a smoking gun which exonerates fat and blames sugar is like the political example above. Yes, it’s sketchy for LBJ to take Wall Street money. But this kind of low-level corruption is so universal that concentrating on any one example is likely to lead to overcorrection.

Yes, the sugar lobby sponsors some research, but the fat lobby has researchers of its own. They tend to be associated with the dairy and meat industries, both of which are high in saturated fat and both of which are very involved in nutrition research. For example, Siri-Tarino et al’s Meta-analysis of prospective cohort studies evaluating the association of saturated fat with cardiovascular disease finds that saturated fat does not increase heart disease risk, but it has a little footnote saying that it’s supported by the National Dairy Council. Modulation of Replacement Nutrients, which finds that replacing dietary fat with dietary sugar doesn’t help and may worsen heart disease, includes two authors affiliated with the National Dairy Council and one affiliated with the National Cattleman’s Beef Association.

Mother Jones does a dairy industry expose and finds:

[Industry] ties can sometimes be hard to avoid, since much of the research on dairy is funded by a constellation of industry-backed institutes, including the Nestlé Nutrition Institute, the Dannon Institute, and the Dairy Research Institute, which spends $19 million a year “to establish the health benefits of dairy products and ingredients.” Even Willett acknowledges that he has received a “very small” dairy industry grant. Dairy companies also donate heavily to the American Society for Nutrition, which publishes the influential American Journal of Clinical Nutrition, and the Academy of Nutrition and Dietetics, “the world’s largest organization of food and nutrition professionals.”

Then there are the industry’s donations to politicians. Dairy companies spent nearly $63 million on federal lobbying and gave $24 million to candidates between 2004 and 2014.

As Jim Babcock points out in the comments, some of the agendas are more complicated than I’m making them sound. Dairy was pretty okay with the low-fat craze for a while, because it let them market low-fat milk. But they do seem to be behind a lot of the pro-saturated-fat research going on right now, and their website does promote pro-saturated-fat articles (1, 2, 3). Overall they seem to be taking a low-key approach where they roll with some studies and push back on others.

In any case, claims that the sugar industry sponsored one study back in the 1960s, and this means everything we’ve ever thought is wrong and biased against fat and in favor of sugar, miss the point (especially since there are probably problems with both sugar and fat). Whatever study the New York Times has dredged up was one volley in an eternal clandestine war of Big Fat against Big Sugar, and figuring out who’s distorted the science more is the sort of project that’s going to take more than one article.

# AI Persuasion Experiment

I’ve been trying to write a persuasive essay about AI risk, but there are already a lot of those out there and I realize I should see if any of them are better before pushing mine. This also ties into a general interest in knowing to what degree persuasive essays really work and whether we can measure that.

So if you have time, I’d appreciate it if you did an experiment. You’ll be asked to read somebody’s essay explaining AI risk and answer some questions about it. Note that some of these essays might be long, but you don’t have to read the whole thing (whether it can hold your attention so that you don’t stop reading is part of what makes a persuasive essay good, so feel free to put it down if you feel like it).

Everyone is welcome to participate in this, especially people who don’t know anything about AI risk and especially especially people who think it’s stupid or don’t care about it.

I want to try doing this two different ways, so:

If your surname starts with A – M, try the first version of the experiment here at https://goo.gl/forms/8quRVmYNmDKAEsvS2

If your surname starts with N – Z, try the second version at https://goo.gl/forms/FznD6Bm51oP7rqB82

Thanks to anyone willing to put in the time.

# SSC Endorses Clinton, Johnson, Or Stein

I.

If you are American, SSC endorses voting in this presidential election.

Andrew Gelman, Nate Silver, and Aaron Edlin calculate the chance that a single vote will determine the election (ie break a tie in a state that breaks an Electoral College tie). It ranges from about one in ten million (if you live in a swing state) to one in a billion (if you live in a very safe state). The average American has a one in sixty million chance of determining the election results. The paper was from the 2008 election, which was a pro-Obama landslide; since this election is closer the chance of determining it may be even higher.

The size of the US budget is about $4 trillion, but Presidents can only affect a tiny bit of that – most of the money funds the same programs no matter who’s in charge. But Presidents do shift budgetary priorities a lot. GW Bush started a war in Iraq which probably cost $2 trillion; the CBO estimates Obamacare may cost about $1.2 trillion. Neither of these are pure costs – Obamacare buys us more health care, and military presence in Iraq buys us [mumble] – but if you think these are less (or more) efficient ways to spend money than other possible uses, then they represent ways that having one President might be better than another. If we suppose a good president would use these trillions of dollars at least 33% more efficiently than a bad president, then this is still $300 billion in value.

So order of magnitude, having a good President rather than a bad one can be worth $300 billion. A 1/60 million chance to create $300 billion in value is worth $5,000; even the 1/1 billion chance afforded someone in a safe state is worth $300.

We don’t know for sure that we’re right about politics. In order to add signal rather than noise to the election results, we have to be better than the average voter. The Inside View is useless here; probably every voter thinks they’re better than average. I recommend the Outside View – looking for measurable indicators correlated with ability to make good choices. Education’s probably a good one. IQ might be another. But overall, my suggestion is that if you’re seriously uncertain about whether or not you think more clearly than the average voter, by that fact alone you almost certainly do.

Suppose you live in a swing state. If you think (in a well-calibrated way) that it’s 10% more likely that your candidate will use $1 trillion well than that the other candidate will, your vote is worth $500. If you live in a safe state, it’s more like $30. If you value the amount of time it takes to vote at less than that, voting is conceivably a good use of your time.

II.

SSC endorses voting for Hillary Clinton if you live in a swing state. If you live in a safe state, I endorse voting for Clinton, Johnson, or (if you insist) Stein. If you want, you can use a vote-swapping site to make this easier or more impactful.

You might notice who’s missing from this endorsement. I think Donald Trump would be a bad president.

Partly this is because of his policies, insofar as he has them. I’m not going to talk much about these because I don’t think I can change anyone’s mind here – either you agree with me (and disagree with Trump) on things like abortion, global warming, free trade, et cetera, or you don’t. A two sentence argument in a blog post won’t change your mind either way.

In fact, I’m not sure any of this ever changes anyone’s mind, and I didn’t really want to write this post. But the latest news says:



This is going to be close. And since the lesson of Brexit is that polls underestimate support for politically incorrect choices, this is going to be really close.

And I don’t know if I’d go so far as Scott Aaronson, who worries that he will one day live in a nuclear hellscape where his children ask him “Daddy, why didn’t you blog about Trump?”. But if some of my blogging on conservative issues has given me any political capital with potential Trump voters, then I this is where I want to spend it.

So here are some reasons why I would be afraid to have Trump as president even if I agreed with him about the issues.

Many conservatives make the argument against utopianism. The millenarian longing for a world where all systems are destroyed, all problems are solved, and everything is permissible – that’s dangerous whether it comes from Puritans or Communists. These same conservatives have traced this longing through leftist history from Lenin through social justice.

Which of the candidates in this election are millennarian? If Sanders were still in, I’d say fine, he qualifies. If Stein were in, same, no contest. But Hillary? The left and right both critique Hillary the same way. She’s too in bed with the system. Corporations love her. Politicians love her. All she wants to do is make little tweaks – a better tax policy here, a new foreign policy doctrine there. The critiques are right. Hillary represents complete safety from millennialism.

Trump’s policy ideas are mostly silly, but no one cares, because he’s not really running on policy. He’s running on making America great again, fighting the special interests, and defying the mainstream media. Nobody cares what policies he’ll implement after he does this, because his campaign is more an expression of rage at these things than anything else.

In my review of Singer on Marx, I wrote that:

I’d always heard that Marx was long on condemnations of capitalism and short on blueprints for communism, and the couple of Marx’s works I read in college confirmed he really didn’t talk about that very much. It seemed like a pretty big gap. I figured…he’d probably made a few vague plans, like “Oh, decisions will be made by a committee of workers,” and “Property will be held in common and consensus democracy will choose who gets what,” and felt like the rest was just details. That’s the sort of error I could at least sympathize with, despite its horrendous consequences.

But in fact Marx was philosophically opposed, as a matter of principle, to any planning about the structure of communist governments or economies. He would come out and say “It is irresponsible to talk about how communist governments and economies will work.” He believed it was a scientific law, analogous to the laws of physics, that once capitalism was removed, a perfect communist government would form of its own accord. There might be some very light planning, a couple of discussions, but these would just be epiphenomena of the governing historical laws working themselves out. Just as, a dam having been removed, a river will eventually reach the sea somehow, so capitalism having been removed society will eventually reach a perfect state of freedom and cooperation.

Singer blames Hegel. Hegel viewed all human history as the World-Spirit trying to recognize and incarnate itself. As it overcomes its various confusions and false dichotomies, it advances into forms that more completely incarnate the World-Spirit and then moves onto the next problem. Finally, it ends with the World-Spirit completely incarnated – possibly in the form of early 19th century Prussia – and everything is great forever.

Marx famously exports Hegel’s mysticism into a materialistic version where the World-Spirit operates upon class relations rather than the interconnectedness of all things, and where you don’t come out and call it the World-Spirit – but he basically keeps the system intact. So once the World-Spirit resolves the dichotomy between Capitalist and Proletariat, then it can more completely incarnate itself and move on to the next problem. Except that this is the final problem (the proof of this is trivial and is left as exercise for the reader) so the World-Spirit becomes fully incarnate and everything is great forever. And you want to plan for how that should happen? Are you saying you know better than the World-Spirit, Comrade?

I am starting to think I was previously a little too charitable toward Marx. My objections were of the sort “You didn’t really consider the idea of welfare capitalism with a social safety net” or “communist society is very difficult to implement in principle,” whereas they should have looked more like “You are basically just telling us to destroy all of the institutions that sustain human civilization and trust that what is baaaasically a giant planet-sized ghost will make sure everything works out.”

And since then, one of the central principles behind my philosophy has been “Don’t destroy all existing systems and hope a planet-sized ghost makes everything work out”. Systems are hard. Institutions are hard. If your goal is to replace the current systems with better ones, then destroying the current system is 1% of the work, and building the better ones is 99% of it. Throughout history, dozens of movements have doomed entire civilizations by focusing on the “destroying the current system” step and expecting the “build a better one” step to happen on its own. That never works. The best parts of conservativism are the ones that guard this insight and shout it at a world too prone to taking shortcuts.

Donald Trump does not represent those best parts of conservativism. To transform his movement into Marxism, just replace “the bourgeoisie” with “the coastal elites” and “false consciousness” with “PC speech”. Just replace the assumption that everything will work itself out once power is in the hands of the workers, with the assumption that everything will work itself out once power is in the hands of “real Americans”. Just replace the hand-waving lack of plans with what to do after the Revolution with a hand-waving lack of plans what to do after the election. In both cases, the sheer virtue of the movement, and the apocalyptic purification of the rich people keeping everyone else down, is supposed to mean everything will just turn out okay on its own. That never works.

A commenter on here the other day quoted an Atlantic article complaining that “The press takes [Trump] literally, but not seriously; his supporters take him seriously, but not literally”. Well, count me in that second group. I don’t think he’s literal. I think when he talks about building a wall and keeping out Muslims, he’s metaphorically saying “I’m going to fight for you, the real Americans”. When he talks about tariffs and trade deals, he’s metaphorically saying “I’m going to fight for you, the real Americans”. Fine. But neither of those two things are a plan. The problem with getting every American a job isn’t that nobody has been fighting for them, the problem with getting every American a job is that getting 100% employment in a modern economy is a really hard problem.

Donald Trump not only has no solution to that problem, he doesn’t even understand the question. He lives in a world where there is no such thing as intelligence, only loyalty. If we haven’t solved all of our problems yet, it’s because the Department of Problem-Solving was insufficiently loyal, and didn’t try hard enough. His only promise is to fill that department with loyal people who really want the problem solved.

I’ve never been fully comfortable with the Left because I feel like they often make the same error – the only reason there’s still poverty is because the corporate-run government is full of traitors who refuse to make the completely great, no-downsides policy of raising the minimum wage. One of the right’s great redeeming feature has been an awareness of these kinds of tradeoffs. But this election, it’s Hillary who sounds restrained and realistic, and Trump who wants the moon on a silver platter (“It will be the best moon you’ve ever seen. And the silver platter is going to be yuuuuuge!”)

III.

But I guess you’ve got to balance someone’s ability to pursue goals effectively with whether you like the goals they’ll be pursuing. I can imagine someone admitting that Clinton will probably be better at governing than Trump, but preferring Trump’s position on the issues so much that it still gives him an edge. In that case, I beg you to consider not only the mean but the variance.

I think even people who expect Trump to be a better President on average will admit he’s a high-variance choice. Hillary is an overwhelmingly known quantity at this point. A Hillary presidency will probably be a lot like Obama’s presidency. There might be a Libya-style military action; probably not an Iraq-style one. If something terrible happens like China tries to invade Taiwan, she will probably make some sort of vaguely reasonable decision after consulting her advisors. She might do a bad job, but it’s hard to imagine a course where a Hillary presidency leads directly to the apocalypse, the fall of American democracy, et cetera.

Trump isn’t a known quantity. Maybe he’ll kind of dodder around and be kind of funny while not changing much. Or maybe there will be some crisis and Trump will take what could have been a quickly-defused diplomatic incident and turn it into World War III. Remember also that it’s more likely the House and Senate both stay Republican than that they both switch to being Democrat. So if Hillary is elected, she’ll probably spend four years smashing her head against Congress; if Trump is elected, he will probably get a lot of what he wants.

Some people like high variance. I don’t. The world has seen history’s greatest alleviation of poverty over the past few decades, and this shows every sign of continuing as long as we don’t do something incredibly stupid that blows up the current world order. I’m less sanguine about the state of America in particular but I think that its generally First World problems probably can’t be solved by politics. They will probably require either genetic engineering or artificial intelligence; the job of our generation is keep the world functional enough to do the research that will create those technologies, and to alleviate as much suffering as we can in the meantime. I don’t see a Clinton presidency as making the world non-functional, whatever that means. I don’t know what I see a Trump presidency doing because, Trump is inherently unpredictable, but some major blow to world functionality is definitely on the list of possibilities.

The one place where Clinton is higher-variance than Trump is immigration. Clinton does not explicitly support open borders, but given her election on a pro-immigration platform and the massive anti-Trump immigration backlash that seems to be materializing, it’s easy to see her moving in that direction. If you believe that immigrants can import the less-effective institutions of their home countries, lower the intelligence of the national hive mind, or cause ethnic fractionalization that replaces sustainable democratic politics with ethnic coalition-building (unlike the totally-not-ethnic-coalition-based politics of today, apparently?), that could potentially make the world less functional and prevent useful technologies from being deployed.

I consider this one of the strongest pro-Trump arguments, but I think it exaggerates the scale of the problem. Hillary will have a Republican Congress to contend with; she probably won’t be able to increase immigration very much. Immigration rates are currently too low to cause massive demographic change before the point at which useful technologies can be deployed, and most immigrants are Asian and come from countries with pretty good institutions themselves. More important, Trump’s anti-immigration policies would prevent foreign researchers from attending top American universities, and probably slow the deployment of future technologies directly, far more than any indirect effect from Hillary would.

There’s another argument here – how exactly are we visualizing a world where immigrants damage American institutions? I envision it as America becoming more like Third World countries – constant ethnic tension, government by strongmen, rampant corruption, lack of respect for checks and balances, and overregulation of industry. But Trump is promising us all of that already, without even admitting any immigrants! If we’re going to become a Third World country, let’s at least help some people while we’re doing it!

IV.

US conservatism is in crisis, and I think that crisis might end better if Trump loses than if he wins.

Since a country with thriving conservative and liberal parties is lower-variance than one with lots of liberals but no effective conservatism, I would like conservatism to get out of crisis as soon as possible and reach the point where it could form an effective opposition. It would also be neat if whatever form conservatism ended out taking had some slight contact with reality and what would help the country (this is not meant as a dig at conservatives – I’m not sure the Democrats have much contact with reality or helps the country either; I’m wishing for the moon and stars here).

Nobody expects Republicans to win blacks and Hispanics. The interesting thing about this election is that college-educated whites are also moving into the Democratic column. If the latest polls are to be believed, the demographic – which favored Romney by 14 points last election – favors Clinton by 8 points now. The nightmare scenario is that Trump wins, his style of anti-intellectual populism is cemented as Official New Republican Ideology, and every educated person switches to the Democrats.

I’m not 100% this would be bad – maybe educated people who are temperamentally conservative would pull the Democratic Party a little to the right, turning them into a broad moderate coalition which has no problem winning elections and combines the smartest elements of liberal and conservative thought. But more likely, there’s a vicious cycle where the lack of intelligent conservatives guts the system of think tanks that produce the sort of studies and analyses which convince smart people to become conservative, which in turn makes there even fewer intelligent conservatives, and so on. In the end, intellectuals won’t just vote Democrat; they’ll shift their personal views further to the left to fit in. We already have a problem with a glut of leftist researchers and journalists producing evidence why leftists are right about everything, and a shortage of conservative researchers and journalists to fact-check them and present the opposite case. As intelligent people desert the Republican Party, this situation gets worse and we lose access to any knowledge that Vox doesn’t want to write an explainer on. In the worst case scenario, everybody develops a hard-coded association between “conservative” and “stupid people”, even more than they have already, the academies purge the hell out of everyone even slightly to the right of the loudest activist, and the only alternative is The Donald Trump Institute Of Research That Is Going To Be Absolutely Yuuuuuuge, which busies itself putting out white papers to a coalition of illiterates.

If Trump fails, then the situation is – much the same, really, but conservatives can at least get started right now picking up the pieces instead of having to wait four years. There’s a fundamental problem, which is that about 30% of the US population is religious poor southern whites who are generally not very educated, mostly not involved in US intellectual life, but form the biggest and most solid voting bloc in the country. If you try to form two parties with 50% of the vote each, then whichever party gets the religious poor southern whites is going to be dominated by them and end up vulnerable to populism. Since the religious poor southern whites are conservative, that’s always going to be the conservative party’s cross to bear and conservatism is always going to be less intellectual than liberalism in this country. I don’t know how to solve this. But there have been previous incarnations of American conservatism that have been better at dealing with the problem than this one, and maybe if Trumpism gets decisively defeated it will encourage people to work on the problem.

V.

I said I wouldn’t try to convince people about the big hot-button issues, but I’ve been told now thatthe guardrails of democracy have been broken lying is okay. So let’s talk about global warming.

Most hot-button issues are less President-influenced than most people think. No Supreme Court is likely to overturn Roe v. Wade at this point, so the president’s impact on abortion is limited to whatever edge cases come before the justices they appoint. I have no idea whether there was more or less capital punishment during Obama’s administration than Bush’s, but I doubt that the president’s opinion of the issue had much impact one way or the other. But it looks like the Obama administration made really impressive progress on global warming; needless to say Donald Trump feels differently.

I don’t want to argue climate science here. I want to say that, as usual, I support the low-variance position that’s not going to make the world vastly less functional before we can invent genetic engineering or AI. Even if you doubt modern climate science, are you so sure it’s wrong that it’s worth the risk? What chance of global warming being a real problem would it take before you agreed that we should probably reduce CO2 emissions just in case? How could that chance possibly be lower than the chance of something that 90-something percent of the relevant scientists believe to be true is true? Yes, we know here that science is not always as authoritative as it would like to be, but it’s not completely anticorrelated with truth either!

(also, if the research about high CO2 levels decreasing cognitive ability is true – and my guess is no, but I’m far from sure – that could be even more disastrous than the traditional global warming effects – remember that even tiny IQ decreases have horrible consequences on a society-wide scale.)

VI.

Okay, but what about the real reason Trump is so popular?

When I talk to Trump supporters, it’s not usually about doubting climate change, or thinking Trump will take the conservative movement in the right direction, or even immigration. It’s about the feeling that a group of arrogant, intolerant, sanctimonious elites have seized control of a lot of national culture and are using it mostly to spread falsehood and belittle anybody different than them. And Trump is both uniquely separate from these elites and uniquely repugnant to them – which makes him look pretty good to everyone else.

This is definitely true. Please vote Hillary anyway.

Aside from the fact that getting back at annoying people isn’t worth eroding the foundations of civil society – do you really think a Trump election is going to hurt these people at all? Make them question anything? “Oh, 51% of the American people disagree with me, I guess that means I’ve got a lot of self-reflecting to do.” Of course not. A Trump election would just confirm for them exactly what they already believe – that the average American is a stupid racist who needs to be kept as far away from public life as possible. If Trump gets elected, sure, the editorial pages will be full of howls of despair the next day, but underneath the howls will be quiet satisfaction that the world is exactly the way they believed it to be.

The right sometimes argues that modern leftism is analogous to early millenarian Christianity. They argue this, and then they say “You know what would stop these people in their tracks? A strong imperial figure who persecutes them. That’s definitely going to make them fade away quietly. There is no way this can possibly go wrong.”

Leftism has never been about controlling the government, and really the government is one of the areas it controls least effectively – even now both houses of Congress, most state legislatures, most governors, etc, are Republican. When people say that the Left is in control, they’re talking about academia, the media, the arts, and national culture writ large. But all of these things have a tendency to define themselves in opposition to the government. When the left controls the government, this is awkward and tends to involve a lot of infighting. When the right controls the government, it gets easy. If Trump controls the government, it gets ridiculously easy.

This has real-world effects. Millennials are more conservative than previous generations. Andrew Gelman, who is usually right about everything, says:

If you look at the cohort of young voters who came of age during George W. Bush’s presidency, they’re mostly Democrats, which makes sense as Bush was a highly unpopular Republican. The young voters who came of age during Obama’s presidency are more split, which makes sense because Obama is neither popular nor unpopular; he has an approval of about 50%

I would prefer the next generation end up leaning more to the right, because that will cancel out younger people’s natural tendency to lean left and make them pretty moderate and so low-variance. I definitely don’t want an unpopular far-right presidency, because then they’re going to lean left, which will combined with the natural leftiness of the young and make them super left. And this is the sort of thing that affects the culture!

VII.

One more warning for conservatives who still aren’t convinced. If the next generation is radicalized by Trump being a bad president, they’re not just going to lean left. They’re going to lean regressive, totalitarian, super-social-justice left.

Everyone has already constructed the narrative: Trump is the anti-PC, anti-social-justice candidate. If he wins, he’s going to be the anti-PC, anti-social-justice President. And he will fail. First of all, because he doesn’t really show much sign of knowing what he’s doing. Second of all, because all presidents fail in a sense – 80% of Americans consistently believe the country is headed the wrong direction and the president is the natural fall guy for this trend. And third of all, because even if by some miracle Trump avoids the first two failure modes, the media will say he failed and people will believe them. And when the anti-PC, anti-social-justice President fails, the reaction will be a giant “we told you so” from the social justice movement, and a giant shift of all the disillusioned young people right into their fold.

Trump is all set to be the biggest gift to the social justice movement in history. They thrive on claims of persecution, claims that they’re the ones fighting a stupid hateful regressive culture that controls everything. And people think that bringing their straw man to life and putting him in the Oval Office is going to help?

If you’re a Jew fighting anti-Semitism, the absolute minimum you can do is not actually kill Christian children and use their blood to make matzah. Likewise, if you are a principled classical liberal fighting the social justice movement’s attempt to smear anyone who disagrees with them as an overprivileged clueless hateful Neanderthal, the absolute minimum you can do is not actually be an overprivileged clueless hateful Neanderthal. Opinions on Trump range all the way from “he is definitely an overprivileged clueless hateful Neanderthal” to “he is remarkably and uniquely bad at not appearing to be an overprivileged clueless hateful Neanderthal”. In any case, having him as the public face of anti-social-justice for the next four years would be a godsend for them and a disaster for everyone else.

VIII.

There’s one more thought I wanted to mention which is vaguely in this space.

The enemy isn’t leftism or social justice. The enemy is epistemic vice.

When the Left errs, it’s through using shouting and shaming to cut through the long and painful process of having to justify its beliefs. It’s through confusing disagreement with evil, a dissenter who needs convincing with a thought-criminal who needs neutralizing.

Sometimes it might be strategically necessary to whack particular ideologies to make examples of them. But in the longer-term, replacing left with right just puts a new group of people in position to shame their opponents and silence dissent. The long range plan has to combine a short-term need to neutralize immediate would-be tyrants with a long-term need to slowly encourage epistemic virtue so that we don’t have to keep putting out fires.

Now, watch this video:

Trump’s not in that crowd. But does anyone think he disagrees with it? Can anyone honestly say that Trump or his movement promote epistemic virtue? That in the long-term, we’ll be glad that we encouraged this sort of thing, that we gave it power and attention and all the nutrients it needed to grow? That the road to whatever vision of a just and rational society we imagine, something quiet and austere with a lot of old-growth trees and Greek-looking columns, runs through LOCK HER UP?

I don’t like having to vote for the lesser of two evils. But at least I feel like I know who it is.

RELATED: Eliezer, The Unit Of Caring, Scott Aaronson

# He Kept Us Out Of War?

I.

Some of the best pushback I got on my election post yesterday was from people who thought Trump was a safer choice than Clinton because of the former’s isolationism and the latter’s interventionism. Since I glossed over that point yesterday, I want to explain why I don’t agree.

Trump has earned a reputation as an isolationist by criticizing the Iraq War. I don’t think that reputation is deserved. He’s said a lot of things which suggest he would go to war at the drop of a hat.

— He says he will “bomb the s#!t out of ISIS” and calls for sending 30,000 troops to destroy them. His campaign website says he will “pursue aggressive joint and coalition military operations to crush and destroy ISIS”.

— He is ambiguous about whether Obama should have intervened in Syria to depose dictator Bashar Assad. He complained “there is something missing from our president. Had he crossed the line and really gone in with force, done something to Assad – if he had gone in with tremendous force, you wouldn’t have millions of people displaced all over the world. ”

— Back during the rebellion in Libya, Trump seems to have been in favor of even more dramatic intervention than Obama eventually allowed. He said on his video blog “I can’t believe what our country is doing. Qaddafi in Libya is killing thousands of people, nobody knows how bad it is, and we’re sitting around we have soldiers all have the Middle East, and we’re not bringing them in to stop this horrible carnage and that’s what it is: It’s a carnage. You talk about things that have happened in history; this could be one of the worst. Now we should go in, we should stop this guy, which would be very easy and very quick. We could do it surgically, stop him from doing it, and save these lives. This is absolutely nuts. We don’t want to get involved and you’re gonna end up with something like you’ve never seen before. But we have go in to save these lives; these people are being slaughtered like animals. It’s horrible what’s going on; it has to be stopped. We should do on a humanitarian basis, immediately go into Libya, knock this guy out very quickly, very surgically, very effectively, and save the lives.”

— He thinks we should have “kept” Iraq’s oil. When pressed on how exactly one keeps billions of barrels of petroleum buried underneath a foreign country, he said he would get US troops to circle and defend the areas with the oil. The “areas with the oil” are about half of the country. This sounds a lot like he wants US troops to remain in Iraq indefinitely.

— He also wants to to keep Libya’s oil. As per National Review: “I would go in and take the oil — I would just go in and take the oil. We don’t know who the rebels are, we hear they come from Iran, we hear they’re influenced by Iran or al-Qaeda, and, frankly I would go in, I would take the oil — and stop this baby stuff.”

— He suggests declaring war on Iran as a response to them harassing US ships. During the debate, he said he would “shoot their ships out of the water.”

— In 2007, he he suggested “knocking the hell out of [Iran] and keeping their oil”, though in his (sort of) defense he might have been confusing them with ISIS at the time.

— In his 2000 book The America We Deserve he suggested a preemptive strike on North Korea: “[If I were President], North Korea would suddenly discover that its worthless promises of civilized behavior would cut no ice. I would let Pyongyang know in no uncertain terms that it can either get out of the nuclear arms race or expect a rebuke similar to the one Ronald Reagan delivered to Ghadhafi in 1986. [Reagan bombed Libya]. I don’t think anybody is going to accuse me of tiptoeing through the issues or tap-dancing around them either. Who else in public life has called for a pre-emptive strike on North Korea?”

— During a town hall meeting, when host Chris Matthews asked Trump when he would use nuclear weapons, he answered “Somebody hits us within ISIS — you wouldn`t fight back with a nuke?” When Matthews reminded him that most people try to avoid ever using nuclear weapons, he answered “Then why are we making them? Why do we make them?”

II.

Some writers have called the period since World War II the “Pax Americana”. Although there have been some deadly local wars, there’s been relative peace between great powers. A big part of this is America’s promise to defend its allies. This both prevents other countries from attacking America’s allies and prevents America’s allies from building big militaries and launching attacks of their own. The whole system is cemented by America-centric trade organizations which make war unprofitable and incentivize countries to stay in America’s orbit.

Trump wants to destroy this system because it costs money, even though it doesn’t really cost that much money compared to anything else we do and Trump intends to increase the defense budget anyway. It’s possible a post-Trump world might find some other way to maintain peace. It’s also possible that it wouldn’t, or that the process of finding that alternative way would be really bloody.

— In March, Trump said “I think NATO may be obsolete. NATO was set up a long time ago — many, many years ago when things were different. Things are different now. We were a rich nation then. We had nothing but money. We had nothing but power. And you know, far more than we have today, in a true sense. And I think NATO — you have to really examine NATO. And it doesn’t really help us, it’s helping other countries. And I don’t think those other countries appreciate what we’re doing.” Although this isn’t the worst opinion, most foreign policy scholars think that our policy of defending our allies is necessary to prevent global arms races and random regional wars.

— In July, he publicly admitted he wasn’t sure he would protect the Baltic states if Russia attacked, something we’re currently obligated to do. The Atlantic calls this “a marked departure from the security policy of every presidential nominee from either of the two major parties since NATO’s founding in 1949”. It’s especially worrying because even if you’re not going to protect the Baltic states from Russia, you shouldn’t openly say so where Russians can hear you!

— And throughout the race, Trump has campaigned on a platform that would effectively end American participation in the World Trade Organization. Trump understands that this would probably start a global trade war, but asks “who the hell cares if there’s a trade war?” I care for two reasons. First, because free trade has produced decades of sustained economic growth and the most successful poverty alleviation in human history. Second, this would probably crash the world economy, creating exactly the sort of depression that tends to produce instability (most famously Hitler’s rise during Germany’s interwar stagnation) or which drives countries toward regional hegemons willing to trade with them or just plain bribe them.

III.

Hillary’s foreign policy isn’t great either, but it doesn’t seem as bad as some people are making it out to be.

— Hillary will probably continue US intervention in Syria; here she is more interventionist than Obama. But her intervention would probably be smaller-scale than Trump’s. She wants to arm “friendly” rebel groups and enforce a no-fly zone, but she has ruled out sending ground troops into Iraq or Syria, something Trump has promised to do. Likely she would focus on keeping enough of Syria safe to protect some civilians and prevent more refugees, then use indirect methods to make life miserable for Assad. This seems like as good a plan as any other.

— The main concern I’ve heard is that the no-fly zone might lead to conflict (war?) with Russia. Declaring a no-fly zone would mean a commitment to shoot down any plane that flies through the zone. Russia is currently flying planes through Syria, and if they tried to call Hillary’s bluff she would have to shoot down Russian planes or lose credibility; shooting down a foreign plane could obviously lead to war. Many different news sources make this point (1, 2, 3, etc). But the clearest description she’s given of what she wants suggests a no-fly zone with Russian cooperation and support. Last October, she said of her no-fly zone proposal that “I think it’s complicated and the Russians would have to be part of it, or it wouldn’t work.” There’s some good discussion of this on Reddit (see especially this comment) where most people end up agreeing that this is the heart of her plan – something like the US agreeing it won’t bomb Russian allies if Russia doesn’t bomb our allies.

— Hillary has said she will “treat cyberattacks just like any other attack”, which could mean that if Russia launches a cyberattack on the US (for example hacking the DNC’s emails) Hillary would treat it as an act of war. I think this requires a stretch. She did mention the possibility of a military response, but only in the context of possible “serious political, economic, and military responses”. My guess is we should interpret this in a non-crazy way – if Russia hacks our emails, we condemn them and maybe hack some of their stuff. If Iran hacks a dam and causes it to fail, then maybe we start thinking airstrikes. Shooting down an airliner is an act of war, but countries have shot down other countries’ airliners a bunch of times and usually people posture a bit and then let it slide. I don’t think it makes sense to think Hillary will treat cyber-attacks more seriously than that.

IV.

A lot of this has a lot of room for interpretation. I’m totally ready to believe that when Trump said he would shoot any Iranian ship that annoyed US vessels, he just meant generic macho posturing and expected everyone to hear it that way. He might even be cunningly pursuing a North Korean – style “mad dog” strategy where he tries to sound so dangerous and unpredictable that nobody dares call his bluff, and so his enemies never mess with him in any way.

Or he might mean everything he says. After all, a lot of it has been pretty consistent since long before he was running for president. There’s no point in saying things to send a game theoretic signal to Iran if you’re a random New York real estate developer and Iran isn’t listening. If he understood the theory behind sounding trigger-happy to intimidate our enemies, he probably wouldn’t have openly admitted he wouldn’t respond to a Russian invasion of the Baltics. And he does seem kind of 100% like a loose cannon in every way, to the point where trying to explain away loose-cannon-like statements as part of a deeper plan seems overly complex.

(Actually, I have a theory which I think explains a lot about Trump’s foreign policy positions: he doesn’t like losers. He supported the Iraq War and the Libya intervention when it looked like we would probably win. Then we lost, and he said they were stupid and bungled. He supports counterfactual invasions of Iraq and Libya where we “kept the oil” because that would have counted as winning. He supports invading ISIS because he expects to be in charge of the invasion and he expects to win. Under this theory, Trump’s retrospective non-support for failed wars doesn’t predict that he won’t start new ones.)

In the end it all comes back to the argument from variance. Maybe Trump is secretly a principled isolationist, and he’s only saying he’ll shoot at Iran and invade Libya and first-strike North Korea and steal oil from Iraq and send troops against ISIS and remove Assad in order to scare people into cooperating with him. Or maybe he’ll actually shoot at Iran and invade Libya and first-strike North Korea and steal oil from Iraq and send troops against ISIS and try to remove Assad. Who knows? He’s said a thousand times now that he’s totally different from the usual politicians, and I believe him. He could do pretty much anything.

(I’d like to think his advisors would rein him in before that point, but when asked which advisors he would consult before a major foreign policy decision, Trump could only think of one person, and he does not exactly inspire confidence.)

I am not qualified to judge Hillary’s work as Secretary of State, but I expect her to play by the book. I’m not sure if Hillary will be more aggressive or more peaceful than the last few presidents, but I don’t expect her to be a wild outlier totally beyond comparison to any previous president. I expect her to consult the foreign policy community on anything important she does, and take some advice relatively within their Overton Window. If she comes to the brink of nuclear war with Russia, I expect her to de-escalate for the same reason I expect Putin to de-escalate; they’re both rationally self-interested people who want to continue being alive and ruling their respective countries, and they value that more than any particular principle or any opportunity to prove their machismo.

I think she remains the low-variance choice for president.

# Somewhat Against Psychiatric Conditions As Domestication Failure

[Epistemic status: Not sure if I’m arguing against a straw man here and my conclusion is what the researchers meant all along.]

I.

Benitez-Burraco and Lattanzi theorize that autism and schizophrenia are anomalies in the human self-domestication process. I’ll try to explain, but for a much better explanation than I can give read Dr. Chris Badcock here.

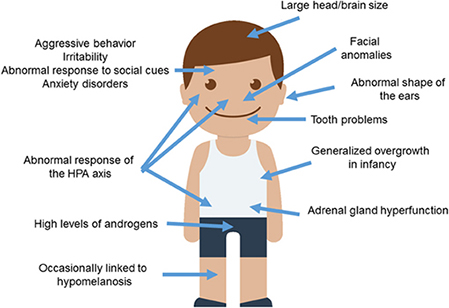
Still here? Fine. BBL’s theory goes like this. When Russian scientist Dmitry Belyaev tried to domesticate foxes by breeding them for tame behavior, he found that changes in a lot of other traits went along for the ride. In short, the foxes started looking kind of dog-like: smaller heads, shorter snouts, spotted fur, floppy ears, more youthful characteristics. Some further experiments confirmed that similar changes happen in any species bred for tameness. Probably this has to do with changes in the neural crest, an embryonic structure which goes on to form a bunch of things including the adrenal medulla. Since the adrenal medulla produces some of the hormones involved in fear and stress, animals with hypoactive medullae will probably be tamer. But since the neural crest also goes on to form lots of other stuff, or produce hormones that influence the formation of lots of other stuff, these tamer animals will be different in other ways too.

BBL continues: we went from being wild apes to tame humans, a process that could be analogized to “self-domestication”. Some of the same changes the Russians saw in the transition from wild to domesticated foxes can be seen in the transition from early hominid skulls to modern human skulls.

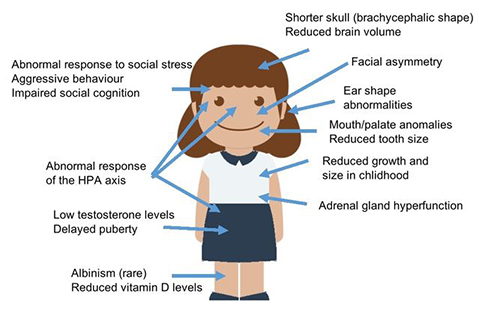
Autistic people, say BBL, are “undomesticated humans” – people in whom for some reason the neural crest changes that result in domesticated features have reversed. They find that some of the changes of domestication syndrome are the reverse of some of the symptoms of autism:

Smaller heads = autistic people have larger heads  
More trusting and social = autistic people are less trusting and social  
Spotted fur = the depigmenting disease “hypomelanosis of Ito” is sometimes associated with autistic symptoms  
Floppy ears = studies find autistic people are more likely to have abnormally shaped ears (really!)  
Change in adrenal response = autistic people have abnormal function in the HPA axis, the system including the adrenal gland

Or in the form of their cutesy picture:



Schizophrenics, say BBL, are “hyperdomesticated humans”. Once again, they match up the symptoms:



I originally thought this theory was dumb. After looking into it more, I think it has some serious issues, but that there might be a core of truth.

II.

I’ll get to that core, but first, the argument against: all of this is coincidences, pareidolia, and finessing things to fit into a system where they don’t really belong.

Going down the list:

Smaller heads = autistic people have larger heads

Some studies find this is true. Others find that it isn’t. In any case, note a discrepancy between this claim and the schizophrenia version. BBL note smaller brains in schizophrenics (true) and shorter skull (true), but not smaller heads, which we would expect if autism were the “reverse” of schizophrenia. In fact, schizophrenics may have larger heads than healthy people. This sort of moving the goal-posts, where autistics are judged on their larger heads but schizophrenics on their smaller brains, is a red flag for fake pattern-matching.

More trusting and social = autistic people are less trusting and social

True! But schizophrenics are way less trusting and social! Paranoia – pathological inability to trust – is a classic symptom of schizophrenia; indeed, if you made people choose between schizophrenia and autism and asked which one was associated with lack of trust, I think most people would choose schizophrenia. This brings an important point into relief: the whole point of domestication is that the domesticated animal is supposed to be friendlier and less aggressive. But nobody would describe schizophrenics as friendlier and less aggressive.

Spotted fur = the depigmenting disease “hypomelanosis of Ito” is sometimes associated with autistic symptoms

True! But hypomelanosis of Ito is a really rare disease (1/10,000 births) that has nothing to do with most autism. Also, it causes eye problems, kidney cysts, weirdly-shaped chests, short stature, seizures, mental retardation, etc. To me this looks more like “a super-rare disease that can cause pretty much anything can sometimes also cause autistic symptoms”, which is not very interesting. Also, domestication causing “pigmentation changes” (usually spotted fur) versus autism being (very rarely) associated with a depigmenting disease and schizophrenia being (very rarely) associated with albinism is more goalpost-shifting.

Floppy ears = studies find autistic people are more likely to have abnormally shaped ears (really!)

I looked at this study – Manouilenko et al – and what it actually finds is that autistic people are more likely to have asymmetrical ears. In fact, nonsignificantly more likely to have asymmetrical ears; their significant finding is that autistic people have more “minor physical abnormalities”, and the asymmetrical ears were one of many pieces of evidence combined to get the significant finding. But asymmetrical features are common in lots of genetic/embryological diseases and seem like a general sign of high mutational load. It seems sketchy to combine autists’ asymmetrical ears and wild foxes’ pointy ears and say “Look, they both have ear abnormalities, this is the same thing!” Some other studies suggest that autistic people have low-set ears, which sounds more promising, but schizophrenic people also have low-set ears, so whatever. The other schizophrenia ear findings are exactly as unconvincing as the autistic ones.

Change in adrenal response = autistic people have abnormal function in the HPA axis, the system including the adrenal gland

Wikipedia’s page on the HPA axis has a section on its possible role in disease, which states that dysfunction of the axis is involved in various conditions “including anxiety disorder, bipolar disorder, insomnia, posttraumatic stress disorder, borderline personality disorder, ADHD, major depressive disorder, burnout, chronic fatigue syndrome, fibromyalgia, irritable bowel syndrome, and alcoholism”. In other words, in a list of the HPA axis’ twelve greatest hits, neither autism nor schizophrenia qualify for inclusion.

I’m not saying that there isn’t an HPA axis component to these diseases. I’m just saying HPA axis is a nonspecific finding. I’m agnostic whether the HPA axis causes everything or our HPA axis study methods are so bad that they invariably turn up false positives. The point is that we shouldn’t get too excited when we see the HPA axis involved in both domestication and autism. This is like saying “Cancer causes you to feel bad, and AIDS causes you to feel bad, therefore cancer causes AIDS.” No, it’s just that everything makes you feel bad.

When we look beyond the general claim of “abnormal function”, things get less clear. BBL say that domesticated animals have “reduced levels of stress hormones including adrenocorticoids, adrenocorticotropic hormone, cortisol, and corticosterone”. So their theory should predict that autistic people have increased stress hormones, and schizophrenics decreased, relative to typical people. Actually, it’s a mess; autistic people seem to have higher ACTH but lower cortisol; schizophrenia studies are conflicting but tend towards higher levels of both. Once again, they can support a general claim of “these conditions affect the same system”, but they can’t predict the direction of the effect. Also, every condition affects this system.

(if you’re wondering why we’re talking about cortisol levels in a theory about the adrenal medulla, well, so am I. Whatever.)

Finally, if by “undomesticated human” we mean something like an ape or Neanderthal, well, neither apes nor Neanderthals (as far as we know) display the symptoms of autism. They seem to be pretty social. They seem to be able to eat all kinds of stuff without trouble. They don’t seem bothered by sensory processing problems. For that matter, dogs and cattle and nth generation silver foxes, the most domesticated animals we’ve got, don’t seem very schizophrenic. I guess cows could just be hallucinating all the time and how would we know, but there doesn’t seem to be any evidence that they are.

So this is why I originally was not too big on this theory.

III.

But what about Williams Syndrome?

Benitez-Burraco and Lattanzi don’t mention Williams Syndrome (also called Williams-Beuren Syndrome) at all, which is crazy, because it sounds a thousand times more like a syndrome of hyperdomestication than either of the two conditions they examine (h/t Nicholas Wade and a random Reddit comment). Williams Syndrome is a rare condition (1/10,000 births) caused by the deletion of some genes on chromosome 7. There are three very interesting things about people with Williams Syndrome. Number one, they are really nice. Like if you meet someone with Williams Syndrome, you will think “This person clearly has a rare genetic disease that causes pathological levels of niceness as a symptom.” Number two, they are really trusting. An Atlantic article profiling the condition, What Happens When You Trust Too Much? describes special therapy for Williams Syndrome children where the therapist has to teach them, painfully and laboriously, how to distrust people. NPR calls it “essentially biologically impossible for kids [with Williams Syndrome] to distrust [people].” Number three, they talk all the time; the informal name for the condition is “cocktail personality syndrome”.

People with Williams Syndromes actually legitimately have short noses (compare to the short snout on domesticated foxes), smaller teeth (compare to smaller teeth in dogs vs. wolves), smaller brains, and “unusually shaped ears” (I can’t find anything more specific; I guess it’s too much to hope for that researchers actually describe the ears as “floppy”).

Also, somebody checked which gene was most different in dogs versus wolves, and they found it was WBSCR17. The WBS in the name stands for “Williams-Beuren Syndrome” because it’s been linked to the disorder. So there’s that.

So as far as I can tell there’s an amazingly good case for Williams Syndrome being linked to domestication. Williams Syndrome tends to cause severe mental retardation and death at an early age, but that’s probably because there are twenty-five totally different genes missing. Maybe a version that only deleted WBSCR17 would keep the behavioral and physiologic changes but not much else.

A lot of people suggest Williams’ Syndrome is “the opposite of autism”. I can only find three pieces of evidence for this. Number one, the obvious contrast with the love of social situations and high verbal skills. Number two, Williams Syndrome kids seem to be really good at face recognition, whereas autistic people are often worse at this. Number three, Williams’ Syndrome kids seem to be unusually bad at the puzzles and interlocking-mechanical-part type problems on which autistic people excel.

On the other hand, there are some reasons to think these conditions are not exact opposites. For one thing, autism is caused by a hideously complex interplay of thousands of genes and various environmental factors, but Williams Syndrome is a drop-dead simple “oops, we forgot part of this chromosome over here”. Williams Syndrome kids seem to have some of the same sensory sensitivities as autistic kids. And both groups usually suffer from mental retardation.

I think that Williams Syndrome establishes the possibility of a physiological social/trust system linked to domestication, the neural crest, and various other parts of embroygenesis. Once you admit the existence of such a system, it seems like autism probably involves some kind of damage to it – probably along with damage to a lot of other systems too. Schizophrenia is more of a stretch, but the overwhelming presence of distrust as a symptom makes the existence of a physiological social/trust system at least kind of interesting and relevant.

So maybe instead of saying that “autistic people are undomesticated humans” and “schizophrenics are hyperdomesticated humans”, we should say something like “there is a very subtle and hard-to-notice biological system that determines level of trust and sociability and which seems weirdly linked to ear and nose shape; autism, schizophrenia, and Williams Syndrome all affect that system in different ways.” Note that this doesn’t mean they’re “the same disease” or “opposite diseases”; the connection might be no deeper than the “connection” where heart attacks, atrial fibrillation, and getting stabbed in the chest all affect the heart. But they all hit the same system.

My take-home message from looking into all of this is that I was very silly for trying to learn about autism and schizophrenia without thinking about embryology. These are highly genetically-loaded diseases that present early in life and seem linked to teratogens and prenatal infections; of course they’re embryological! I had to take some embryology classes in medical school, and like everyone else I tuned them out because they seemed totally irrelevant to real clinical practice and mostly involved memorizing pointless trivia like “on day thirty-six and a half, the developing shmendroblast has transformed into a blexomere”. But if you want to know what causes secret connections between ear shape and level of social trust, embryology seems like the way to go. Autism and schizophrenia are hard to study because they seem to affect everything, yet nothing specifically enough to localize the condition. Maybe going back and thinking more embryologically could help pinpoint the particular systems involved.

# More Hardball Debate Questions

[See also Hardball Questions For The Next Debate. The Gary Johnson question is not original to me.]

Jill Stein:

You’re a former doctor and researcher who first got involved in politics because of your interest in public health. One of your first forays into activism was the 2000 publication of In Harm’s Way: Toxic Threats to Child Development, a magisterial report on the effect of pollution on children’s physical and mental health. You described your focus as being on “developmental disabilities, including attention deficit/hyperactivity disorder, autism, and related neurodevelopmental diseases”. You describe “accumulating evidence of neurotoxic damage to children by environmental agents, such as lead and PCBs”.

In Chapter 7, you discuss the high burden of pesticides eaten by developing children, saying that:

Twenty million American children five and under eat an average of eight pesticides every day through food consumption. Thirty-seven pesticides registered for use on foods are neurotoxic organophosphate insecticides, chemically related to more toxic nerve warfare agents developed earlier this century…a national health exposure study detected chlorpyrifos residues (as the metabolite TCP) in the urine of 82% of a representative sample of American adults. A more recent study in Minnesota revealed that an even higher 92% of children had detectable levels of this metabolite in their urine.

You connect this increasing pesticide exposure to what you believe to be increasing levels of developmental disorders in American children:

The Cailfornia Department of Developmental Services released [a study] in March 1999 [that] looked at pervasive developmental disorders from 1987 through 1998 and showed a 210 percent increase in cases entered into the autism registry during those years. If the incidence of autism is increasing, and/or clusters of autism are being discovered, an environmental influence is likely.

Maybe as a result, you’ve become a big advocate of eating organic food. Your party platform says you want to “support organic and regenerative agriculture” and “put a moratorium on GMOs and pesticides until they are proven safe”. You presented a pro-organics case on Bill Moyers’ show back in 2012, and you’re even known for preparing your own organic meals on the campaign trail.

But there’s a lot of pushback from mainstream scientists and the mainstream media. For example, news webzine Vox has an interesting article Is Organic Food Any Healthier? Most Scientists Are Still Skeptical publicizing a meta-analysis of 237 studies which showed that “organic foods didn’t appear to be any healthier or safer to eat than their conventionally grown counterparts” and that “typical exposure to pesticide residues is at levels 10,000 to 10,000,000 times lower than doses that cause no observable effect in laboratory animals that are fed pesticides daily throughout their entire lifetimes”. Vox has also written Local And Organic Food Has Extra Safety Risks. Just Ask Chipotle. Vox’s spinoff webzine Eater even makes fun of customers looking for “natural” foods without having any idea what that means.

If they’re right, then you’re promoting an unscientific fad that has millions of people needlessly stressed out about everything they eat. On the other hand, if you’re right, then these media outlets’ pooh-poohing of a vital public health message makes them complicit in and maybe even responsible for what you call the “epidemic” of childhood neurodevelopmental disorders.

So my question for you is: do you believe Vox ‘zines cause autism?

Hillary Clinton:

During your first debate with Donald Trump, the moderator asked you about racial bias in police shootings; you responded that “implicit bias is a problem for everyone, not just police”. You argued that you would “put money into that budget to help us deal with implicit bias by police officers” and that we’ve “got to do everything possible to improve policing, to go right at implicit bias”. Your running mate Tim Kaine continued on the theme, saying that people shouldn’t be afraid to bring up police officers’ implicit biases.

This talk of implicit bias references a whole psychological field centered around the Implicit Association Test. It works like this: a subject sitting in front of a keyboard is shown rapid-fire pictures representing various categories – classically black people, white people, positive adjectives, and negative adjectives. They’re given various instructions about which keys to press in response to which categories, and their responses are timed. Many people will find that it’s easier to press the same key for white people and positive adjectives (and an opposite key for black people and negative adjectives) than to press the same key for whites+negatives and blacks+positives. This has been widely considered to show implicit racism – that is, even people who say they are not racist unconsciously associate black people with bad qualities. This research has become wildly popular, profiled in every major media outlet, and catapulted its inventors to scientific stardom. It’s even been featured on the Oprah Winfrey show, maybe a first for a social psych paper.

A few early small studies suggested the IAT predicted prejudiced behavior. But later attempts to replicate this result failed. Blanton, Jaccard, Klick, Mellers, Mitchell, Tetlock (2009) reanalyzed some of the original studies, found no effect, and complained that the IAT was being popularized despite an almost-complete lack of evidence for its validity. Oswald et al (2013) did a meta-analysis of 275 implicit association test results from 46 different studies and found that “IATs were poor predictors of every criterion category other than brain activity, and the IATs performed no better than simple explicit measures”. Carlsson and Agerstrom did another meta-analysis earlier this year, and found “the overall effect was close to zero and highly inconsistent across studies” and “there is…little evidence that the IAT can meaningfully predict discrimination, and we thus strongly caution against any practical applications of the IAT that rest on this assumption”.

You particularly mention the IAT as relevant to policing, but Dolan Group, a consulting firm which advises police forces how to avoid racial discrimination, did an internal analysis of results surrounding the IAT and reports to its clients that:

Persons who do not hold overt racist attitudes do not have to worry about some deeply-hidden, unknown, unconscious attitudes influencing their work decisions. These findings reveal the need to aggressively weed out officers who hold conscious racial stereotypes and biases in order to avoid biased-based policing. These findings also raise questions about whether the money and time spent on law enforcement training and testing regarding implicit bias could be put to better use on something else.

When you and your running mate suggest a focus on implicit bias as relevant to policing, this can really only be justified by taking the preliminary findings of a few small early studies and ignoring both more rigorous reanalysis of their results and the consensus finding of all studies and meta-analyses conducted since that time.

On the other hand, there still is something to be explained here: if the IAT isn’t analyzing implicit racial prejudice, why do people so consistently have an easier time associating black people with negative adjectives? I actually have a theory of my own about that. Consider claims like the following:

1. Black people were brutally enslaved for hundreds of years.  
2. Black people are almost three times more likely than whites to live below the poverty line.  
3. Black people are systematically being murdered by the criminal justice system.  
4. Black people are frequent victims of racism and hate crimes.  
5. Our society is set up to structurally discriminate against black people.

None of these claims are racist per se; in fact, many of them are anti-racist in intent. But all of them connect black people to negative affect! If your local newspaper says that white people usually have friendly and positive interactions with the police but black people are victimized and killed by police, that is some heavy association of whites with positive feelings and blacks with negative feelings. If you usually see photos of white people in the news under the headline “LOCAL BUSINESS BOUGHT BY GOOGLE”, and photos of blacks in the news under the headline “LEARN HOW OUR RACIST SOCIETY KEPT THIS POOR WOMAN FROM SUCCEEDING” then once again, you’re learning to associate whites with positive feelings and blacks with negative feelings.

This would explain very nicely why people taking the IAT generally associate whites with positive feelings and blacks with negative feelings in a way apparently unrelated to whether they are explicitly prejudiced/racist. It would also explain very nicely why about 50% of blacks associate whites with positive feelings and blacks with negative feelings, which is definitely a thing that happens and which previous explanations of have always sounded unconvincing and ad hoc.

But from your debate statements, it sounds like you are absolutely opposed to this reinterpretation. That you are committed to defending the position that implicit bias is a real predictor of racism, and that Implicit Association Tests don’t just report contingent associations drilled in by the media, but genuinely reveal profound unconscious beliefs about how the world works.

So my question for you is: would you be willing to take an Implicit Association Test measuring how easily you associate your own name vs. your opponents’ names with the adjective “crooked”?

Gary Johnson:

If you were elected, what would you do about the ongoing crisis in Updog?

Donald Trump:

You’re well-known for your boast that you “hire the best people”. And one of those best people is Steve Bannon, the CEO of your campaign. When Bannon took over on August 17th, 538 had you at only a 12% chance of winning; after he was running your campaign for a month, you were up to 40%. Although you’ve since crashed back down, a lot of political observers attribute what successes you’ve had to Bannon and what problems you’ve had to your own big mouth. You seem to recognize his utility, calling him one of “the best talents in politics, with the experience and expertise needed to defeat Hillary Clinton in November”.

Before he joined your campaign, Bannon was best known for his role leading far-right news website Breitbart. But he was actually involved in some pretty interesting stuff when he was younger. In particular, in 1993 Bannon was the acting director of the famous environmental science experiment Biosphere 2.

Biosphere 2 was an attempt to create a self-sustaining closed ecosystem capable of supporting human life, possibly with applications for future space travel. It was actually the first such attempt – it was called “Biosphere 2” because the first such self-sustaining biosphere was the Earth itself. Eight “crew members” entered the facility along with various plants and animals, the airlocks were sealed, and for a year everyone tried to do what they could to keep the various species and environmental parameters in balance.

It didn’t work; CO2 levels started fluctuating wildly, soil microbes surged out of control, ants and cockroaches overran the facility, oxygen dropped to worrying levels, and the experiment was stopped early out of concern for crew health. They decided to try a second mission, and that was when they had a change in management and brought on Mr. Bannon as director.

Unfortunately, a lot of the crew members really didn’t like Bannon and his team. Possibly some of it had to do with an incident where a crew member submitted a list of safety complaints and Bannon threatened to “shove it down her f\*\*king throat”. It got so bad that some of the crew deliberately vandalized the Biosphere, causing gas exchange between the inside and the outside and ruining the scientific value of the experiment. Although they probably could have tried again, by that time lawsuits and financial mismanagement had sapped their funding, and they finally sold the whole thing off to Columbia University as a research campus.

So my question for you is: in all of history, there have only been two self-sufficient ecosystems capable of maintaining human life. Your team has already destroyed one of them. The other is Earth. How scared should we be?

# The Moral Of The Story

[content warning: puns. This is mostly self-plagiarism from my Tumblr and Twitter]

Once upon a time there was a small desert village with a single well outside town. One day a young woman went to the well to fetch water, and the well heard her crying, and asked “What’s wrong?”

She stopped her sobbing and asked the well “You can talk?”

“Yes,” said the well. “Long ago, the witch who lives in this town gave me life so I could serve as a guardian to the townspeople.”

“Alas,” said the young woman. “I am the daughter of that witch. She lived in peace with the townsfolk for many years. But the new mayor, who is a violent and hateful man, riled the people up against her, and they burned her at the stake. I am young and still do not know very much magic. I tried to curse them, but my curses fizzled. Now I worry I will never avenge my mother’s death.”

“Do not be afraid,” said the well. “I will take care of this.”

The next morning, when the Mayor came to fetch water from the well, he heard an odd noise coming from the bottom. He peered over as far as he could to see what was happening. Then an impossibly long arm shot up from the bottom of the well, grabbed the mayor, and pulled him into the well shaft. There was a horrible crunching sound, and nobody ever saw the Mayor again. The townsfolk apologized to the witch’s daughter, and they all lived happily ever after.

Moral of the story: Living well is the best revenge.

Pixar’s movie Up won the Academy Award for “Best Picture” and was widely hailed as one of the best children’s films of the decade. In fact, some people argued it was too good, and that kids were ignoring school, chores, and other responsibilities to watch it again and again. They said that along with the cute plot, the short, catchy name gave it an almost drug-like addictive quality. This made a lot of people very angry, and Pixar agreed to give its addictive must-watch movies longer names in the future.

Moral of the story: Do not call Up what you cannot put down.

There’s a new report out of CERN that a team of scientists has unraveled the structure of the photon. Apparently this started years ago when some equations showed that photons acted like tiny “hands” – structures with a “palm” and radiating “fingers” – which “crawl” across time/space and “grab” the solid particles they interact with. This explained most of the properties of light but wasn’t an exact match for the data. The latest result is that single photons are actually made up of hundreds of these shapes, all joined together into a single particle, and this is how they’re able to travel so quickly.

Moral of the story: Many hands make light work.

Once upon a time there was an ugly duckling. All of the other ducklings had grown their beautiful white soft downy feathers, but this duckling had no down feathers at all and was bald and ugly and all the other ducklings teased him.

He went to the mysterious crow who lived in the woods and asked for help. The crow said to repeat the magic words “HOCUS POCUS” at midnight with a full moon, and then he would grow his down feathers. The duckling tried that, but the moon just laughed at him and said the magic had no power here.

So he went to the creepy raven who lived in the swamp and asked for help. The raven said to repeat the magic words “ABRA CADABRA” at high noon on a sunny day, and then he would grow his down feathers. The duckling tried that, but the sun just laughed at him and said he wasn’t bound by the magic.

So he went to the wise old owl who lived in the tallest tree and asked for help. The owl explained that the duckling should just ignore the mockery of the other birds and accept that he was okay just the way he was, because there were no magic spells to make ducklings grow feathers.

Moral of the story: You are beautiful, no matter what they say. Words can’t bring you down.

Once upon a time a young lady died and went to Hell. At the check-in desk, Satan asked her age. She was in her twenties, but looked much younger; she thought quick and realized that even in Hell, they probably wouldn’t be mean to children. So she told Satan that she was twelve, and sure enough he said she wasn’t old enough to be held accountable for her sins, and ushered her off to a more peaceful part of Hell reserved for ages eleven through thirteen. She met the other sinners there and realized that many of them, like her, were older people who had lied to get out of their punishment.

Satan began to suspect something like this was going on, so he set up hidden cameras in the 11-13 wing of Hell, trying to catch people acting like adults or admitting to one another that they had lied about their age. But there were hundreds of millions of sinners and Satan couldn’t monitor all the cameras himself. So he went up to the mortal world and asked for the best supercomputer they had. The mortals recommended a newer model of Deep Blue, the supercomputer that had first beaten a human world champion at chess. Satan picked one up from IBM and went back to Hell, where he programmed the Deep Blue to monitor all of the hidden camera feeds at once and report any suspicious activity.

Sure enough, after a few days, he got thousands of reports of people acting older than thirteen. He hunted them down and removed them to Hell proper, where there was much wailing and gnashing of teeth. And it all could have been avoided if they had just stuck to their charade and acted as young as they said they were.

Moral of the story: Don’t get caught, be tween – the Devil and the Deep Blue see.

By 2050, screens have shrunk and become more flexible until the dream of “programmable paper” becomes a reality. Citizens of the future read newspapers like the ones in Harry Potter that include moving images and even videos of important events. This new technology even makes it as far as the US Post Office, where they decide to include programmable stamps. Instead of a static picture of eg George Washington’s head, it will have a moving image of Washington speaking and giving his famous Farewell Address.

Unfortunately, the technology isn’t ready for the kind of abuse that envelopes undergo on their travel throughout the country and the world. Most of the computerized stamps become corrupted and “crash”; in a particularly common bug, they try to reload but just end up displaying “GENERATING IMAGE…” permanently. The government has no money to fix the problem, so people just get used to stamps on their letters that say “GENERATING IMAGE…” instead of having interesting pictures on them.

Moral of the story: If you want a vision of the future, imagine a human face booting on a stamp forever.

# The Heart Has Its Reasons That Reason Knows Not Of

I.

Psychoanalysts argue that sons are attracted to women who look like their mothers, because they imprint on their mothers and use them as a schema for their ideal woman.

(and probably something similar for daughters and their fathers, though the psychoanalysts don’t usually get around to talking about that)

I’ve counterargued that sons get half their genes from their fathers, who apparently were attracted to women who looked like the sons’ mothers. Since study after study shows genes having unexpectedly big effects on behavior, and early childhood experiences with parents having unexpectedly small effects, maybe this would be a better explanation for the effect (if it even exists).

But the research doesn’t really bear me out.

Goose-ologist Konrad Lorenz raised goslings from birth. When the goslings grew up, he found they tried to mate with humans, especially “Caucasian men with white beards”. He concluded that they imprinted on their adoptive parent (him) and learned to prefer mates who looked like that parent.

More formally, Bischof et al got some male zebra finches and arrange to have them raised by a closely-related species, Bengalese finches. Then they put them in cages with both female zebra finches and female Bengalese finches and observed which females the birds tried to court. The results were pretty striking; they overwhelmingly went for the Bengalese finches who looked like their mothers, not the zebra finches who were genetically more suitable. Spence & Smith replicated this finding with zebra fish raised by differently-colored zebra fish.

So the research shows conclusively that sexual selection is based on learned imprinting, at least in animals whose names start with the string “zebra fi”. What about humans?

I can’t find the study itself, but multiple reviews cite Jedlicka 1984, who looked at children of mixed-race couples (white and native Hawaiian). They found that both men and women were more likely to marry someone of the race of their opposite-sex parent than of their same-sex parent (eg if you’re a woman with a Hawaiian mother and white father, you’re more likely to marry a white person). This is consistent with some kind of social imprinting where your opposite-sex parent serves as a template for future romantic interest. It’s not consistent with a simple genetic theory where you just get both parents’ genes. It might be consistent with a more complicated genetic theory where mate preferences are on a sex-appropriate chromosome or get chromosomally imprinted such that you only care about your father’s preferences for women and your mother’s preferences for men, but this is hard and I haven’t seen any analysis of whether it’s evolutionarily worth it.

Enquist, Aronsson, Ghirlanda, Jansson, and Jannini (2010) starts its Methods section with “We obtained data through newsgroups alt.sex.fetish and alt.sex.fetish.breastmilk”, so you know it’s gonna be interesting. They test another feature of men sexually imprinting on their mothers: suppose you’re a man with a sibling a few years younger than yourself. That means your mother was pregnant or lactating during the supposed critical sexual imprinting window. So if men with younger siblings are more likely to have pregnancy and lactation fetishes, that suggests that sexual imprinting on mothers is really a thing. This is indeed what they found: when a person with a pregnancy or lactation fetish only had one sibling, there was a 66% chance (compared to expected 50%) that their sibling would be younger than they, p < 0.0001 in their sample of 560 such people. This was true if and only if they were between age 1.5 and 5 during their sibling’s birth, hinting at the span of the imprinting window. Of course, this is still a really poor predictor: 33% of such people got the fetish without any younger siblings, and of course most people with younger siblings don’t end up with the fetish at all. But it does look like something is going on.

(I wonder what’s up with adult baby fetishes; if you could find a similar pattern among them that would suggest they’re imprinting on the baby sibling instead of the mother, which would be fascinating. Maybe I should try to survey the appropriate subreddit.)

Hefferman and Fraley did a very similar study. They find that people born to older parents, when compared against people born to younger parents, find older faces more attractive. This was true even after controlling for the age of the participants themselves. The effect size was small but pretty consistent across different groups and measurements. I am not quite as happy about the quality this study as in the ones above, but nothing raises huge red flags.

So there’s a lot of circumstantial evidence suggesting some kind of imprinting process is going on. But what about the original question – do people choose mates who look like their opposite-sex parent? And how genetic versus environmental is it?

Bereczkei, Gyuris, and Weisfeld addresses this question directly. They get a sample of grown-up adopted daughters and their adoptive fathers. They show subjects pictures of the adoptive fathers taken when their daughters were 2-8 years old, and then photos of four similar-aged men, one of whom is the daughter’s current husband. The subjects’ task is to guess which of the photos is the husband based on which of them looks most like the adoptive father. And subjects are in fact able to do this better than chance; they pick the correct husband 37% of the time compared to 21% of the time for (each of the three) incorrect husbands, which in their sample of 242 subjects is a p-less-than 0.001 result.

Somebody else check if I’m wrong here, but I’m a little concerned about the way they calculated p-values in this study. Suppose that their 242 subjects are always accurate in identifying the husband who looks most like the father; no randomness or noise here. But now the chance for randomness and noise comes in how many of the women might have picked husbands who look like their fathers by pure coincidence. It looks like there were only 26 women used in the study, so there’s a lot more opportunity for those twenty-six people to coincidentally choose husbands who happened to look like Dad than for the 242 subjects to coincidentally choose the correct husband even if they all looked alike. Yet as far as I can tell, the p-value calculation was only done on the latter possibility.

(an analogy: suppose that I want to prove I have ESP. I predict that a coin will land heads, then flip the coin. It does indeed land heads. Then I survey one thousand observers to ask them to judge whether the coin landed in accordance with my psychic prediction. All 1,000 of them say yes, a p = 0.000001 result that could not possibly be by coincidence. This does not prove at the p = 0.000001 level that I have ESP!)

Rantala and Marcinkowska do a really excellent review of this field that found a lot of the same papers I did. They acknowledge the Bereczkei/Gyuris/Weisfeld study as very interesting, but note that the first author “replicated” these results in a 2009 study using more scientific facial measurements. That study was found to be extremely flawed (found by Rantala himself, in fact) and was retracted under a cloud of accusations of misconduct and/or inexcusable error. It’s still not clear exactly what happened there, and nobody has formally accused the original father/adopted-daughter study of anything, but it’s kind of awkward to hang your theory on a study by somebody accused of scientific misconduct for faking a very similar study that supported the same hypothesis.

(side note: I wish that people would explain retractions better. Right now I don’t feel like I have a good understanding of exactly how the 2009 study got retracted, so I haven’t learned anything that will help me spot similar problems in the future)

So in summary: there’s a lot of good evidence that animals learn mate preferences by imprinting on their parents. There’s some circumstantial evidence that humans might do this too, across traits as diverse as race, age, and lactation status. There is, however, as yet no smoking gun.

II.

Wait, what about twin studies? Those are usually pretty great! Can we just do a twin study and sort this out once and for all?

Zietsch et al look at “a large community-based sample of twins and their partners and parents (N > 20,000 individuals) to test for genetic and family environmental influences on mate choice”. They find:

…near-zero genetic influences on male and female mate choice over all traits and no significant genetic influences on mate choice for any specific trait (!!!!!). A significant family environmental influence was found for the age and income of females’ mate choices, possibly reflecting parental influence over mating decisions.

This might be the first twin study I’ve ever seen which unambiguously breaks Turkheimer’s First Law Of Behavioral Genetics (every trait is somewhat heritable). Such a striking finding should increase our confidence in all of the above experiments a lot. So okay, I guess this issue is solved, it’s definitely just sexual imprinting on the opposite-sex parent, thank goodness, for once we have a perfectly clear noncontradictory result and we can all just go home and –

We also tested for evidence of sexual imprinting, where individuals acquire mate-choice criteria during development by using their opposite-sex parent as the template of a desirable mate; there was no such effect for any trait.

Sigh.

Okay, fine, let’s look at this a little more closely. They were analyzing a bunch of data from a big Australian survey of twins. This included the twins, the twins’ family, and the twins’ mates “height, BMI, age, education, income, personality, social attitudes, and religiosity”, which were the dimensions along which they tried to predict mate choice. They figured out whether imprinting was involved by checking whether twins’ mates were more like their opposite-sex parent than like other members of their family (their same-sex parent, their other twin). There was no sign of this being true, not even a nonsignificant trend.

If we want to play the dangerous game of trying to explain differences between contradictory studies instead of just dismissing everything as noise, I might argue that this looked at some pretty different variables compared to the last set. Instead of looking at facial similarities, it’s looking at things like social attitudes and religiosity; young children trying to imprint on their mother’s image can maybe be forgiven for not knowing her opinion about Asian immigrants (one of the “social attitudes” questions they asked).

The authors took a different tactic and pointed out that most of us don’t marry the first person we have a crush on, or even the second or third. Sometimes we marry the tenth person we really like, sometimes we settle for people we only like a little, and sometimes we get drunk, have sex in a cheap motel, and have the person’s parent threaten us with a shotgun unless we go to the chapel right now. So maybe the person we end up marrying isn’t a good proxy for our mate preferences per se.

(this is starting to get kind of depressing)

They investigate this hypothesis in a followup study where they directly ask twins about their preferences for an ideal mate. Here they’re able to get some more immediately visual data – preferences like tall/short, long-hair/short-hair, beards/clean-shaven, even big-breasts/small-breasts. They find…well, some things come out heritable, but the confidence intervals are really wide. There are a lot of suspicious things like hair-length preferences being impressively and significantly heritable for women but not heritable at all for men, and on one hand the two findings’ 95% confidence intervals do overlap, but on the other hand that’s because all of the confidence intervals are super-wide anyway. Overall I guess it’s nice that this study doesn’t blatantly break Turkheimer’s First Law, but I’m not going to draw any sweeping conclusions off of it.

I have one more twin study here, Lykken and Tellegen (which coincidentally opens with the same Pascal quote as this post). This experiment has an interesting design – they ask dizygotic and monozygotic twins to rate how attracted they are to their co-twin’s spouse! (I hope these researchers went above and beyond in keeping all their data confidential!) Despite those human interest stories where two identical twins separated at birth rediscover each other and find that they have both married blonde Corgi-owning optometrists named Theo, in this study there was very little resemblance between the spouses of either identical or fraternal twins. On average identical twins correlated with each other at r = 0.57 on some long list of variables, but their spouses correlated with each other only at r = 0.14 (identical) and r = 0.11 (fraternal), and most of this was just similar religious and educational backgrounds that don’t surprise us (we already know people tend to marry others of the same religion and social class, and twins are no exception).

This is interesting because it suggests a minimal role of either genetics or shared environment in mate choice, even in the most extreme circumstance (identical twins raised together). And in fact, the authors remind us that people who have had multiple long-term relationships often choose people who don’t resemble each other in more than the most superficial ways (if the authors had exposure to polyamory, they might note that most people’s simultaneous partners aren’t very similar either).

Maybe this just supports Zietsch et al’s hypothesis that mate preferences, whether imprinted or genetic, don’t really matter because we don’t have a whole shopping aisle of mates lined up to choose from and we’ve got to take whatever we can get. Yet even so there’s room for surprise. Even if only 5% of the opposite sex is interested in you, that still leaves the average person hundreds of different choices over their lifetime. Surely there should be some degrees of freedom for people to pick who they end up with. So what’s going on? If not by genes or shared environment, how do we choose our significant others?

The authors’ guess: true love, a mysterious and magical thing totally beyond scientific understanding.

If we provisionally accept our interpretation of these data, we are left with a curious and disquieting conclusion: Although most human choice behavior lawfully reflects the characteristics of the chooser and of the choice, the most important choice of all, that of a mate, seems to be an exception…we outline a theory that is compatible with these interpretations, namely that human pair bonding is relatively adventitious, based on romantic infatuation which, as Stendhal observed, “is like a fever that comes and goes quite independently of the will.”

Sure. Sounds plausible. Let’s just consider this whole matter closed.

# AI Persuasion Experiment Results

I.

Last month I asked three thousand people to read some articles on AI risk and tell me how convinced they were. Last week, I asked them to come back and tell me some more stuff, to see if they stayed convinced.

I started off interested in the particular articles – which one was best at convincing newcomers to this topic. But I ended up hoping this could teach me about persuasion in general. Can online essays change people’s minds about complicated, controversial topics? Who is easiest to convince? Can we figure out what features of an essay do or don’t change people’s minds?

Depending on the last digit of people’s birth dates, I asked them to read one of five different essays:

— I asked people whose birth dates ended with 0 or 1 to read Wait But Why’s The AI Revolution: The Road To Superintelligence.

— I asked people whose birth dates ended with 2 or 3 to read a draft version of my own Superintelligence FAQ.

— I asked people whose birth dates ended with 4 or 5 to read Lyle Cantor’s Russell, Bostrom, And The Risk Of AI.

— I asked people whose birth dates ended with 6 or 7 to read Michael Cohen’s Extinction Risk From Artifical Intelligence.

— And I asked people whose birth dates ended with 8 or 9 to read Sean Carroll’s Maybe We Do Not Live In A Simulation. This had nothing to do with AI risk and was included as a placebo – that is, to get a group who had just had to read an online essay but presumably hadn’t had their minds changed about AI.

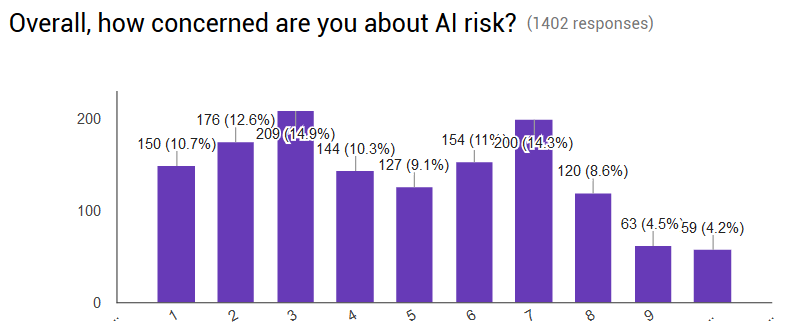
I hosted all of these pieces on Slate Star Codex and stripped them of any identifiers so that hopefully people would judge them based on content and not based on how much they liked the author or what color the page background was or whatever. It mostly worked: only 67% of readers had no idea who had written the essay, with another 23% having only vague guesses. Only about 10% of readers were pretty sure they knew.

People did read the essays: 70% of people said they finished all of theirs, and another 22% read at least half.

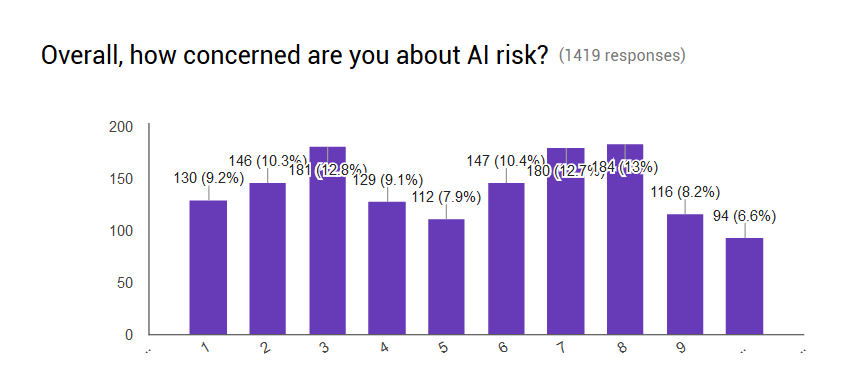
So the experiment was in a pretty good position to detect real effects on persuasion if they existed. What did it find?

II.

My outcome was people’s ratings on a score of 1 – 10 for various questions relating to AI. My primary outcome, selected beforehand, was their answer to the question “Overall, how concerned are you about AI risk?” About half of the respondents took a pre-test, and I got the following results:



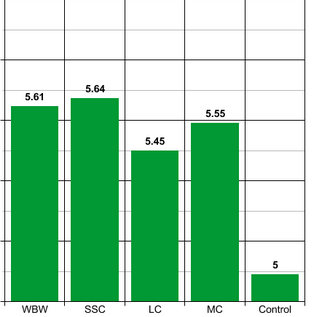
After reading the essays, this changed to:



Overall, people increased their concern an average of 0.5 points.

(Note that I only had half the sample take a pretest, because I was worried that people would anchor to their pretest answers and give demand effects. This turned out not to be a big deal; the people who had taken a pretest changed in the same ways as the people who hadn’t. I’ve combined all data.)

But what about the different essays? All five groups had means between 5 and 6, but the exact numbers differed markedly.



Note truncated y-axis

On further testing, the differences between the four active essays weren’t significant, but the difference between all of the essays and the control was significant.

III.

Aside from the primary outcome, I also had various secondary outcomes: answers to specific questions about AI risk. First, a list of average pretest and posttest answers (pretest, posttest):

How likely are we to invent human-level AI before 2100?: 5.7, 5.9  
How likely is human-level AI to become superintelligent within 30 years?: 7.3, 7.3  
How likely is human-level AI to become superintelligent within 1 year?: 4.8, 5.2  
How likely that a superintelligent AI would turn hostile to humanity?: 6.4, 6.8  
How likely that a hostile superintelligence would defeat humans?: 6.8, 7.1

All the nonzero differences here were significant.

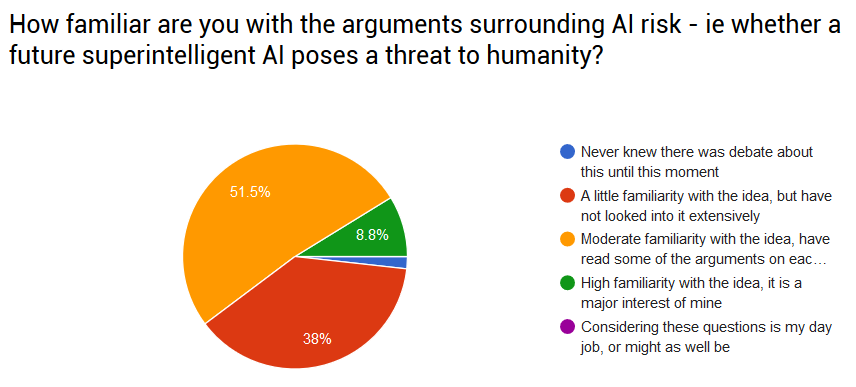
If we look at this as a conjunction of claims, all of which have to be true before AI risk becomes worth worrying about, then it looks like the weak links in the chain are near-term human-level AI and fast takeoff. Neither of these are absolutely necessary for the argument, so this is kind of encouraging. There is less opposition than I would expect to claims that AI will eventually become superintelligent, or to claims that a war with a superintelligent AI would go very badly for humans.

Given the level of noise, there wasn’t a lot of evidence that any of the (active) essays were more persuasive than others on any of the steps, including the two dubious steps. This is actually a little surprising, since some essays focused on some things more than others. Possibly there was a “rising tide lifts all boats” effect where people who were more convinced of AI risk in general raised their probability at every step. But there was too much noise to say so for sure.

IV.

Was there any difference among people who were already very familiar with AI risk, versus people who were new to the topic?

It was hard to tell. Only about two percent of readers here had never heard about the AI risk debate, and 60% said they had at least a pretty good level of familiarity with it.

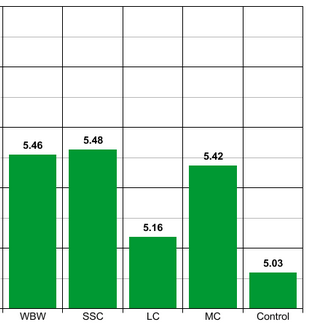


The best I could do was to look at the 38% of participants (still about 1000 people!) who had only “a little familiarity” with the idea. Surprisingly, these people’s minds didn’t change any more than their better-informed peers. The average “little familiarity” person became 0.8 points more concerned, which given the smaller sample size wasn’t that different from the average person’s 0.5.

In general, people who knew very little about AI thought it was less important (r = 0.47), which makes sense since probably one reason people study it a lot is because they think it matters.

V.

How stable were these effects after one month?



Still truncated y-axis

Pretty stable.

Other than an anomalous drop in the third group (I’m sticking with “noise”), the effect remained about two-thirds as strong as it had been the moment after the participants read the essays. All essay groups remained significantly better than the control group.

I looked at the subgroup of people who’d had little knowledge of AI risk before starting the experiments. There was a lot more noise here so it was harder to be sure, but it seemed generally consistent with the same sort of effect.

So in conclusion, making people read a long essay on AI risk changed their overall opinions about half a point on a ten-point scale. There weren’t major differences between essays. This was true whether or not they were already pretty familiar with it. And about two-thirds of the effect persisted after a month.

At least in this area, there might be a modest but useful effect from trying to persuade people.

In terms of which essay to use to persuade people? I don’t have any hard-and-firm results, but there were two trends I noticed. First, my essay was somewhat underrepresented among the people who’d had the biggest jumps (3 points or more) in their level of concern. Second, the third essay was anomalously bad at maintaining its gains over the month. That leaves just the first and fourth. Some people in the comments said they were actively repulsed by the fourth, and WaitButWhy seems pretty good, so at the risk overinterpreting noise, you might as well just send people theirs.

You can find the chaotic and confusingly-labeled data below. It might help to read the survey itself to figure out what’s going on.

Main experiment: .xlsx, .csv  
One month follow-up: .xls, .csv

# How Does Recent AI Progress Affect The Bostromian Paradigm?

[content note: I seriously know nothing about this and it’s all random uninformed speculation]

I.

AI risk discussions are dominated by the Bostromian paradigm of AIs as highly strategic agents that try to maximize certain programmed goals. This paradigm got developed in the early 2000s, before a recent spurt of advances in machine learning. Do these advances require any changes to the way we approach these topics?

The latest progress has concentrated in neural networks – “cells” arranged in layers that represent the potential for ascending levels of abstract categorization. For example, a neural network working on image recognition might have a low-level layer that scans the image and resolves it into edges, a medium-level network that scans the set of edges and resolves it into shapes, and a highest-level network that scans the shapes and resolves them into subjects and themes. With enough training, the network “learns” how best to map each level onto the level above it, ending up with profound insight into the high-level features of a scene.

These are a lot like the human brain, and in fact some of the early researchers got important insights from neuroscience. The brain certainly uses cells, the cells are arranged in layers, and the brain categorizes things in hierarchies that move from simple things like edges or sounds to complicated things like objects or sentences.

In particular, these networks are like the brain’s sensory cortices, and they’re starting to equal or beat human sensory cortices at important tasks like recognizing speech and faces.

(I think this is scarier than most people give it credit for. It’s no big deal when computers beat humans at chess – human brains haven’t been evolving specific chess modules. But face recognition is an adaptive skill localized in a specific brain area that underwent a lot of evolutionary work, and modern AI still beats it)

The sensory tasks where AIs excel tend to involve abstraction, categorization, and compression: the thing where you take images of black dogs, white dogs, big dogs, little dogs, ugly dogs, cute dogs, et cetera and are able to generalize them into “dog”. Or to take a more interesting example: a new AI classifies images as pornographic or safe-for-work. Its structure naturally gives it an abstract understanding of pornographicness that allows it to “imagine” what the most pornographic possible images would look like (trigger warning: artificially intelligent computer generating the most pornographic possible images). This kind of classification/categorization/generalization ability is a major advance and eerily reminiscent of human abilities.

But how far is this to building an AGI or human-level AI or superintelligence or whatever else you want to call it?

II.

Consider two opposite perspectives:

The engineer’s perspective: Categorization ability is just one tool out of many. When people invented automated theorem-provers, that was pretty cool – it meant computers could now assess new mathematics. But for AGI, you still need some thing that wants to prove theorems, something (someone?) that can do something with the theorems it proves. The theorem-prover is a tool for the AI to use, not the core “consciousness” of the AI itself. The same will be true of these new neural nets and deep learning programs. They can recognize dogs, and that’s cool. But AGI is still about creating some kind of program that wants to recognize dogs, and which can do something interesting with the dogs once it recognizes them. And that will probably require something different from either a theorem-prover or a neural-net-categorizer. A paperclip maximizer might use a neural net to recognize paperclips, but its desire to maximize them will still come from some novel architecture we don’t know much about yet which probably looks more like normal programming.

The biologist’s perspective: The whole brain runs on more or less similar cells doing more or less similar things, and evolved in a series of tiny evolutionary steps. If we’ve figured out how one part of the brain works, that’s a pretty big clue as to how other parts of the brain work. The human motivation system is in brain structures not so different from the human perception-association-categorization system, and they probably evolved from a common root. If researchers are discovering that the easiest way to make perception-association-categorization systems is neural nets reminiscent of the brain, then they’ll probably find that those neural nets are pretty easy to alter slightly to make a motivational system reminiscent of the brain. This would look less like strategic/agenty goal maximization, which the brain is terrible at, and more like the sort of vague mishmash of desires which humans have.

The exact evolutionary history beyond the biologist’s perspective is complicated. There’s a split between some sensory processing centers (like the visual cortex) and some motivational/emotional centers (like the hypothalamus) pretty early in vertebrates and maybe even before. But in other cases the systems are all messed up. Some parts of the cortex interact with the hypothalamus and are considered part of the limbic system. Some parts of the really primitive lizard brain handle sensation (like the colliculi). It looks like sensation/perception-related areas and emotion/motivation-related areas are mixed throughout every level of the brain. Most important, the frontal lobe, which we tend to interpret as the seat of truly human intelligence and executive planning and “the will” – probably evolved from sensation/perception-related areas in fish, since it looks like sensation/perception-related areas are just about all the cortex that fish had. And all of this evolved from the same couple of hundred neurons in worms, which were already responsible for interpreting the sensations picked up by the worm’s little bristle thingies.

The point is, neither evolution nor anatomy suggests that the brain enforces a deep conceptual separation between perception, motivation, and cognition. Instead, the same sort of systems which handle perception in some areas are – with a few tweaks – able to handle cognition and motivation in others.

In fact, there are some deep connections between all three domains. The same factors that make a grey figure on dark ground look white can make an okay choice compared to worse choices look good. The same top-down processing that screws up PARIS IN THE THE SPRINGTIME is responsible for confirmation bias. In general the mapping between cognitive biases and perceptual illusions is fruitful enough that it’s hard for me to believe that cognition and sensation/perception aren’t handled in really similar ways, with motivation probably also involved.

So if we have something that can equal human sensory cortices – not just in the coincidental way where a sports car can equal a cheetah, but because we’re genuinely doing the same thing human sensory cortices do for the same reasons – then we might already be further than we think towards understanding human intelligence and motivation.

III.

A quick sketch of two ways this might play out in real life.

First, categorization/classification/generalization/abstraction seems to be a big part of how people develop a moral sense, and maybe a big part of what morality is.

Everyone remembers the whole thing about mental categories, right? The thing where you have a category “bird”, and you can’t give a necessary-and-sufficient explicit definition of what you mean by that, but you know a sparrow is definitely a bird, and an ostrich is weird but probably still a bird, and there are edge cases like Archaeopteryx where you’re not quite sure if they’re birds or not and there’s probably no fact of the matter either way? Cluster-structures in thingspace? Weird border disputes? That thing?

And you remember how we get these categories, right? A little bit of training data, your mother pointing at a sparrow and saying “bird”, then maybe at a raven and saying “bird”, then maybe learning ad hoc that a bat isn’t a bird, and your brain’s brilliant hyperadvanced categorization/classification/generalization/abstraction system picking it up from there? And then maybe after several thousand years of this Darwin comes along and tells you what birds actually are, and it’s good to know, but you were doing just fine way before that?

We learn morality in a very similar way. When we hit someone, our mother/father/teacher/priest/rabbi/shaman says “That’s bad”; when we share, “that’s good”. From all this training data, the categorization/classification/generalization/abstraction system eventually feels like it has a pretty good idea of what morality is, although often we can’t verbalize an explicit definition any better than we can verbalize an explicit definition of “bird” (“it’s an animal that can fly…wait, no, bats…um, that has feathers…uh, do all birds have feathers? Bah, of course they don’t if you pluck them, that wasn’t what I meant…”). Just as Darwin was able to give an explicit definition of “bird” which conclusively settled some edge cases like bats, so philosophers have tried to give explicit definitions of “morality” which settle edge cases like abortion and trolley-related mishaps.

An AI based around a categorization/classification/generalization/abstraction system might learn morality in the same way. Its programmers give it a bunch of training data – maybe the Bible (this is a joke, please do not train an AI on the Bible) – and the AI gains a “moral sense” that it can use to classify novel data.

The classic Bostromian objection to this kind of scheme is that the AI might draw the wrong conclusion. For example, an AI might realize that things that make people happy are good – seemingly a high-level moral insight – but then forcibly inject everybody with heroin all the time so they could be as happy as possible.

To this I can only respond that we humans don’t work this way. I’m not sure why. It seems to either be a quirk of our categorization/classification/generalization/abstraction system, or a genuine moral/structure-of-thingspace-related truth about how well forced-heroin clusters with other things we consider good vs. bad. A fruitful topic for AI goal alignment research might be to understand exactly how this sort of thing works and whether there are certain values of classification-related parameters that will make classifiers more vs. less like humans on these kinds of cases.

Second, even if we can’t get this 100% right, there might be a saving grace: I don’t see these kinds of systems as paperclip maximizers. The human utility function seems to be a set of complicated things generalizing/abstracting from a few biologically programmed imperatives (food, sex, lack of pain) and ability to learn other goals from society and your moral system.

Categorization/classification/generalization/abstraction is certainly involved in reinforcement learning. You say “BARK!” and a dog barks, and you give it a treat. The dog needs to be able to figure out, on the fly, whether the treat was for barking when you said “BARK!”, for barking whenever you speak, for barking in general, for being next to you, or just completely random. This is a problem of categorization and abstraction – going from training data (“the human did or didn’t reward me at this specific time”) to general principles (“when the human says bark, I bark”).

I don’t really understand how the human motivational system works. Dopamine and the idea of incentive salience seem to be involved in a fundamental way that seems linked to perception. But I am kind of hopeful that it’s something that’s not too hard to do if you already have a working categorizer, and that it’s a foundation to build agents that want things without being psychopathic maniacs. Humans can want sex without being insane sex maximizers who copulate with everything around until they explode. An AI that wanted paperclips, but which was built on a human incentive system that gave paperclips the same kind of position as sex, might be a good paperclip producer without being insane enough to subordinate every other goal and moral rule to its paperclip-lust.

Tomorrow 10/31 is the last day of MIRI’s yearly fundraiser, and as usual I think it is a good cause well worth your donation. But its basic assumption is that AIs will be very computer-like: entities of pure code and logic that will reflect on themselves using mathematical tools. I can also imagine futures where AIs aren’t much more purely-logical than we are, and the tools we need to keep them human-friendly are very different. I support MIRI’s efforts to deal with the one case, but I’m hoping there will be some efforts in the other direction as well.

EDIT: Nick points out some of MIRI’s work along these lines.  
EDIT2: Comment by Eliezer

# SSC San Jose Meetup 11/4

What: SSC meetup  
When: Friday 11/4, at 7 PM  
Where: 3806 Williams Road, San Jose, CA  
Why: Physicist/intelligence researcher Steve Hsu of Information Processing will be in the Bay Area and wants to get to meet some of you.  
How: David Friedman is very nice and lets us use his house for things

I will unfortunately not be able to attend.

# The Pyramid And The Garden

I.

A recent breakthrough in pseudoscience: the location of the Great Pyramid of Giza encodes the speed of light to seven decimal places.

This is actually true. The speed of light in a vacuum is 299,792,458 meters per second. The coordinates of of the Great Pyramid are 29.9792458° N, 31.1342880° E (you can confirm with Google Maps that this gets you right on top of the Pyramid). The speed of light and the latitude number there have all the same digits. That’s a pretty impressive coincidence.

You might think this is idiotic because the meter was invented by 1600s French people. If ancient aliens or Atlanteans built the pyramids, why would they encode their secret wisdom using a unit of measurement from 1600s France? But there’s a way around this objection: the 1600s French people defined their meter as 1/10,000,000th the distance between the Equator and the North Pole. If the aliens also thought that was an interesting way to measure length, then they could have encoded their secret wisdom in it. So you wouldn’t need aliens who could predict the thoughts of 1600s Frenchmen. Just aliens who thought exactly like 1600s Frenchmen.

(actually, a different group of 1600s Frenchmen proposed a different version of the meter, defined as the length of a pendulum with a half-period of one second. This turned out to be 99.7% of the 1/10,000,000th-the-way-to-the-North-Pole definition, so either one works unless you want super-exactness. I think a much more interesting conspiracy theory would be that aliens designed the Earth to encode secret wisdom about the periods of pendulums.)

But realistically, aliens who think suspiciously like French people probably weren’t involved. So how do we explain the coincidence?

II.

The following is indebted to user mrfintoil’s great explanation on metabunk.org.

First, it’s not a coincidence to seven decimal places. Yes, that particular nine-digit sequence lands you atop the Great Pyramid. But that gives you way more precision than you need – cutting off the last three digits actually gets you closer rather than further from the center of the Pyramid. The only numbers that are doing any work are the 29.9792° N. So you really only get four decimal places worth of coincidence.

On the other hand, matching six digits is still pretty good. That’s literally a one-in-a-million chance.

So here the explanation has to go to how hard the pseudoscientists worked to find a coincidence of this magnitude; in other words, how many degrees of freedom they had.

Here’s an obvious example; as far as I can tell, the longitude of the Great Pyramid doesn’t encode anything interesting at all. So it’s not the equivalent of winning a one-in-a-million lottery with a single ticket. It’s the equivalent of winning a one-in-a-million lottery with two tickets.

A second issue: if the latitude of the Great Pyramid had been 10.7925 N, that would be the speed of light in kilometers per hour, which would be an equally impressive match.

So just taking these two degrees of freedom, we have four lottery tickets:

1. The one where the latitude is the speed of light in meters/second  
2. The one where the longitude is the speed of light in meters/second  
3. The one where the latitude is the speed of light in kilometers/hour  
4. The one where the longitude is the speed of light in kilometers/hour

In other words, the number of lottery tickets increases exponentially as we get more degrees of freedom.

Let me list out all the degrees of freedom I can think of and see where we end up. I am going to try my best to be as fair as possible to the ancient aliens. For example, I was considering saying that since there are three pyramids at Giza, we have to multiply by three, but to be honest the Great Pyramid is clearly greater than the other two, and it would be less elegant if Menkaure’s pyramid encoded some amazing cosmic constant, so I won’t raise that objection. I am going to try to be really fricking fair.

1. Latitude vs. longitude (2 options)

2. Speed of light in meters/second vs. kilometers/hour vs. cubits/second vs. cubits/hour. I’m avoiding using feet/miles, because that’s even more arbitrary than meters. But I think it would actually be even more convincing if the calculation actually used the real Egyptian unit, which I understand is the cubit. So let’s go with (4 options)

3. Great Pyramid vs. Sphinx. Like I said before, the other two pyramids at Giza are noticeably less impressive than the Great Pyramid. But the Sphinx is pretty impressive, and the ancient aliens folks talk about it just as much as the Pyramid, so I think that would be an equally good hit if it had been true. (2 options)

4. Use of a 90 degree latitude system vs. use of a 100 degree latitude system. I’m a little split on this one, because it wouldn’t look anywhere near as impressive if the pseudoscience sites had to explain that they found a really cool coincidence but it only worked if you converted normal latitude into a different hypothetical latitude system that had 100 degrees. But since we know the aliens/Atlanteans use base 10 anyway (they’re encoding their wisdom in the base 10 representation of the speed of light) it makes more sense for them to use a base 10 latitude system instead of replicating our own bizarre custom of using base 10 for everything else but having latitude go from 0 to 90. On the other hand, if these were Earth-based Atlanteans, they might have gotten the custom of dividing the circle into 360 parts for the same reason we did – there are about 360 days in a year. And if they were aliens, maybe we got our bizarre latitude convention from them – the idea of 360 degree circles is really old and lost in the mists of time. Overall I can see this one going either way, so I’m going to give it as (2 options)

5. Decimal point placement. The latitude 29.9792 N matches the speed of light exactly, but so would the latitudes 2.99792, 2.99792 S, and 29.9792 S. I checked these other sites at the same longitude as the Pyramid to see if there were any mysterious features. But they seem to be, respectively, a perfectly ordinary field in Uganda, a perfectly ordinary field in Tanzania, and a perfectly ordinary patch of ocean. But a world where the pyramid was in Uganda and the ordinary field was in Egypt would be just as much of a hit as our current world. Therefore (4 options)

From these really simple things alone, we learn we’ve got 2 x 4 x 2 x 2 x 4 = 128 lottery tickets, reducing our 1/1 million chance of winning to something more like 1/10,000. Progress!

There are a few other degrees of freedom that I think are a little harder to judge, but still important:

6. What aspect of the Pyramid we’re looking at. That is, it would have been equally interesting (maybe moreso!) if its height or width matched the speed of light exactly. So that’s another (3 options). I guess if the ancient aliens were really good at what they were doing, they could have given the pyramid 299,792,458 sides, but I won’t hold that against them. This should really make the multiplication more complicated because I can no longer use all the different ways of representing latitude vs. longitude, but I’ll stick with the simple method for now.

7. Which site we’re looking at. This one is hard, because I don’t know if anywhere else has the ancient alien-related credibility of the Great Pyramid. The only equally mysterious site I can think of is Stonehenge, and maybe the Nazca Lines. I don’t feel comfortable saying it would be equally impressive if Tiwanaku or Yonaguni had the right coordinates. I’ll just say (2 options) for Pyramids and Stonehenge.

8. Which constant we’re looking at. Sure, the Pyramid encoding the speed of light is pretty cool, but what about the Planck length? Avagadro’s number? I’m split on whether I want to include mathematical constants like pi or e in here. I think if it encoded pi to some number of decimals places then I would just think that the Egyptians were more advanced at math than I thought but it wouldn’t necessarily be earth-shattering. The Egyptians knowing e would be pretty shocking but still maybe not worth believing in ancient aliens over. There really aren’t that many physical constants as cool as the speed of light, so I might just arbitrarily call this one (4 options).

So now we have a total of 128 x 3 x 2 x 4 = 3072 lottery tickets, for a 1/300 chance of winning the one-in-a-million lottery.

I would like to say “Ha ha, I sure proved those dumb conspiracy nuts wrong”, except that a 1/300 chance is still a pretty impressive coincidence – what scientists call p < 0.01. And now I've used up all my excuses.  
  
I think what’s going on here is that I’m still accepting the terms of the game – comparing only the exact categories used in the original calculation. Suppose that the latitude of the Great Pyramid was exactly 30.0000? That too would be impressive – it would prove that the pyramid builders knew the exact size and shape of the Earth and were able to build their Pyramid one third of the way between Equator and Pole. Suppose that the Great Pyramid was latitude 19.69724. That’s the date humankind first landed on the moon in yyyy/mm/dd format – clearly the Pyramid was built by a time-traveling Nostradamus! Suppose that the Pyramid was built of stones of four different colors, with blue stones always paired opposite red stones, and yellow stones always paired opposite green stones. Then the ancient Egyptians were trying to tell us about the structure of DNA. What if the Pyramid, viewed from above, looked like a human brain?

Is it fair to take all of that into account? If so, does the remaining coincidence go away? I wish I were able to give these questions a more confident affirmative answer.

III.

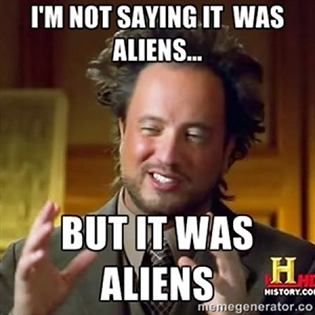
I still believe that pseudoscience is helpful for understanding regular science. The loopholes that let people discover proofs of ESP or homeopathy are the same ones that let them discover proofs of power posing and ego depletion.

In the same way, numerology is helpful for understanding statistics. You can see the same factors at work, free from any lingering worry that maybe the theory you’re investigating is true after all.

Andrew Gelman writes about the garden of forking paths. The idea is: the scientific community accepts a discovery as meaningful if p < 0.05 - that is, if equally extreme data would only occur by coincidence 5% of the time or less. In other words, you need to win a lottery with a one-in-twenty chance if you want to get credit for discovering something absent any real effect to be discovered.  
  
But if a scientist forms their hypothesis after seeing their data, they might massage the precise wording of their hypothesis to better fit their data. If there are many different ways to frame the hypothesis, then they have many lottery tickets to choose from and a win is no longer so surprising.  
  
Gelman discusses a study claiming to find that women wear red or pink shirts during the most fertile part of their menstrual cycle, which sometimes involves red or pink coloration changes in primates. The study does detect the effect, p < 0.05. But there were a couple of different ways the researchers could have framed the problem. They could have looked at only red shirts. They could have looked at only pink shirts. They chose days 7-14 as most fertile. But they could also have chosen days 6-15 without really being wrong. They could have looked only at the unmarried women most likely to be trying to attract mates. A recent paper listed 34 different degrees of freedom that can be used in this kind of thing. Add up enough of them, and you have more than twenty tickets to the one-chance-in-twenty lottery and success is all but certain.

I used to call this the Elderly Hispanic Woman Effect, after drug studies where the drug has no effect in general, no effect on a subgroup of just men, no effect on a subgroup of just women, no effect on a subgroup of just blacks, no effect on a subgroup of just whites…but when you get to a subgroup of elderly Hispanic women, p < 0.05, apparently because it's synchronized with their unique biological needs. This is pretty obvious. The lesson of the Pyramid-lightspeed link is that sometimes it isn't. It just looks like some sudden and shocking coincidence.  
  
The other lesson of the Pyramid is that I cannot consistently figure this kind of thing out. I threw everything I had against the correlation, and I still ended up with p = 0.003. I don’t think this is because the Pyramid really was designed by aliens with a suspicious link to 1600s France. I think it’s because I’m not creative enough to fully dissect coincidences even when I’m looking for them.

This is always happening to me in real studies too. Something seems very suspicious. But their effect size is very high and their p-value is very significant. I can’t always figure out exactly what’s going on. But I should be reluctant to dismiss the possibility that I’m missing something and that there’s some reasonable explanation.



# Tuesday Shouldn’t Change The Narrative

538 predicts Hillary has a 65% chance of winning the election to Trump’s 35%. New York Times says it’s more like 84% Hillary and 16% Trump. Both sites agree both candidates will get somewhere between 40% and 50% of the popular vote, and that Hillary seems to lead Trump by 3%. The smart money is on Hillary, but at this point either major candidate could win.

Lots of things can happen tomorrow. Maybe it rains in Philadelphia, that city’s racially diverse and left-leaning voters stay home, and Pennsylvania goes for Trump, winning him the election. Maybe there’s a really good get-out-the-vote campaign among Hispanics, and Florida ends up being Trump 48 Hillary 52 instead of the projected Trump 52 Hillary 48. Maybe the Department of Agriculture announces that Hillary is under investigation for bringing exotic weevil species into the US, and the population turns against her en masse.

And someone is going to confuse this kind of stuff with deep insight into the state of the country.

In June 2016, Jon Wiener of The Nation wrote Relax, Donald Trump Can’t Win, about how the media is incentivized to make races look competitive but an understanding of political fundamentals proved that there was no way for Trump to actually make it.

On the other hand, a few days ago Scott Adams reiterated his long-standing prediction of a 98% chance Trump wins in a landslide.

I’m worried that one of these two things will happen on Wednesday:

Either Hillary wins, and everybody agrees that Jon Wiener and various other people like him were right, that the fundamentals made a Trump win impossible, that Trump was a random clown who never had a chance anyway, that the people who warned us to beware of Trump were crying wolf, that this proves that nationalism is a spent force in politics, et cetera.

Or Trump wins, and everybody agrees that Scott Adams was a genius, that Wiener was an idiot, that Trump is a brilliant “master persuader”, that this proves that the 21st century will be a century of renewed nationalist power, that the white working class is sexist, that elites need to realize the precariousness of their position within a democratic system, or whatever.

Imagine that the deciding factor really is a rainstorm in Philadelphia. There was a rainstorm in Philly, therefore nationalism is one of the great motivating forces in human affairs? It was a clear sunny day in Philly, therefore nationalism doesn’t matter anymore? The difference between nationalism being all-powerful and irrelevant is whether there was a cold front over the mid-Atlantic region?

But with a race this close, any deciding factor is going to be about as random as a rainstorm over Philadelphia. Maybe the pollsters made some kind of big mistake and missed shy Trump voters, and the vote goes Trump 47% Hillary 45% instead of the predicted Hillary 47% Trump 45%. So what? The difference between a proof of nationalism’s vigor versus proof its impotence is which candidate gets 47% vs. 45%? Really?

If a Trump victory tomorrow would convince you that X is true, I suggest that you believe X is true regardless of whether or not Trump wins, because Trump’s victory almost certainly will depend more on noise than on X. If a Hillary victory tomorrow would convince you that Y is true, I suggest that you believe Y is true regardless of whether or not Hillary wins, for the same reason. If there’s some Z that you will believe only if Trump wins but not if Hillary wins, then I suggest you seriously reconsider what thought process has led you to decide that you will flip your views on politics and society depending on whether or not there’s a rainstorm or a 2% polling error or whatever.

Instead, I suggest people precommit to their views on politics and society now. We live in a country and a world where Hillary can be at about 47% and Trump at about 45%. This is pretty much all you need to know. It suggests that a lot of people are willing to support a nationalist candidate, and a lot of other people really hate that candidate. It suggests that political fundamentals are totally compatible with a situation where either Trump or Hillary could win based on noise in the electoral process.

(unless the polls are totally wrong and one candidate somehow wins in a 20 percentage point landslide or something)

It also suggests that both Wiener and Adams were wrong to be so confident in their respective predictions. If either one is right, it will be mostly by luck. Wiener tells us to “relax” because Trump can never win, and maybe Trump doesn’t win, but the fact is that even if Trump loses we were one Hillary gaffe away from the opposite result and shouldn’t have relaxed at all. Adams says there’s been a 98% chance of a Trump win since last year, but the polls make it look a lot like Trump only has a chance at all because of the total coincidence of Hillary getting hit by a new FBI investigation two weeks before Election Day.

I already count both Wiener and Adams as having been proven wrong regardless of what happens tomorrow. Any further praise or condemnation launched at one or the other after the election is just interpreting noise, or at least a signal so subtle that it might as well be.

If both Wiener’s extreme pro-Hillary prediction and Adams’ extreme pro-Trump prediction are bad, what would a good prediction look like? In January of this year, I predicted that, conditional on Trump winning the Republican primary, he would have a 20% chance of winning the election. Well, Trump won the Republican primary. And today, the day before the election, the prediction markets give Trump’s chance of winning as 17.9%.

If Trump wins anyway, I’ll have egg on my face and it’ll look bad when I grade my prediction accuracy next year. But I don’t think I would fundamentally update the way I think about America or the way I make predictions. No prediction can account for every rainstorm. I think I got the fundamentals right, and if I end up losing on noise I can at least take solace in knowing that is soon to be the least of our problems.

[EDIT: I agree the election results can obviously change the future! Maybe if Trump wins he’ll enact nationalistic policy that will make people more nationalist. But the results shouldn’t change our interpretation of existing trends.]

# Book Review: House of God

I’m not a big fan of war movies. I liked the first few I watched. It was all downhill from there. They all seem so similar. The Part Where You Bond With Your Squadmates. The Part Where Your Gruff Sergeant Turns Out To Have A Heart After All. The Part Where Your Friend Dies But You Have To Keep Going Anyway. The Part That Consists Of A Stirring Speech.

The problem is that war is very different from everything else, but very much like itself.

Medical internship is also very different from everything else but very much like itself. I already had two examples of it: Scrubs and my own experience as a medical intern (I preferred Scrubs). So when every single personin the medical field told me to read Samuel Shem’s House of God, I deferred. I deferred throughout my own internship, I deferred for another two years of residency afterwards. And then for some reason I finally picked it up a couple of days ago.

This was a heck of a book.

On some level it was as predictable as I expected. It hit all of the Important Internship Tropes, like The Part Where Your Attendings Are Cruel, The Part Where Your Patient Dies Because Of Something You Did, The Part Where You Get Camaraderie With Other Interns, The Part Where You First Realize You Are Actually Slightly Competent At Like One Thing And It Is The Best Feeling In The Universe, The Part Where You Realize How Pointless 99% Of The Medical System Is, The Part Where You Have Sex With Hot Nurses, et cetera.

All I can say is that it was really well done. The whole thing had a touch of magical realism, which turns out to be exactly the right genre for a story about medicine. Real medicine is absolutely magical realist. It’s a series of bizarre occurrences just on the edge of plausibility happening to incredibly strange people for life-and-death stakes, day after day after day, all within the context of the weirdest and most byzantine bureaucracy known to humankind.

Just in the past week, for example, I had to deal with an aboulomaniac patient – one with a pathological inability to make up his mind. He came to my clinic for treatment, but as soon as he saw me, he decided he didn’t want treatment after all and left. The next day, he was back on my calendar – he’d decided he needed treatment after all – but when his appointment came around, he chanegd his mind and left again. This happened five times in five days. Every day he would phone in asking for an appointment. Every day I would give it to him. Every day he would leave a minute or two before it began. Unsure how to proceed, I sought out my attending. He ignored my questions, pulled me into a side office, took out his cell phone, and started playing me a video. It’s a scene from his musical, The Phantom Of The Psychiatric Unit, which he’s been forcing his interns to rehearse after rounds. I watched, horrified. It was weirdly good.

If I were to write a book about this kind of thing, people would criticize me for being unrealistic. The only way to get away with it is to pass it off as “a touch of magical realism”, and this The House of God does to excellent effect.

The story revolves around an obvious author-insert character, Roy Basch MD, who starts his internship year at a hospital called the House of God (apparently a fictionalized version of Beth Israel Hospital in Boston). He goes in with expectations to provide useful medical care to people with serious diseases. Instead, he finds gomers:

“Gomer is an acronym: Get Out of My Emergency Room. It’s what you want to say when one’s sent in from the nursing home at three A.M.”

“I think that’s kind of crass,” said Potts. “Some of us don’t feel that way about old people.”

“You think I don’t have a grandmother?” asked Fats indignantly. “I do, and she’s the cutest dearest, most wonderful old lady. Her matzoh balls float – you have to pin them down to eat them up. Under their force the soup levitates. We eat on ladders, scraping the food off the ceiling. I love…” The Fat Man had to stop, and dabbed the tears from his eyes, and then went on in a soft voice, “I love her very much.”

I thought of my grandfather. I loved him too.

“But gomers are not just dear old people,” said Fats. “Gomers are human beings who have lost what goes into being human beings. They want to die, and we will not let them. We’re cruel to the gomers, by saving them, and they’re cruel to us, by fighting tooth and nail against our trying to save them. They hurt us, we hurt them.”

This is where the magical realism starts to come in:

Rokitansky was an old bassett. He’d been a college professor and had suffered a severe stroke. He lay on his bed, strapped down, IV’s going in, catheter coming out. Motionless, paralyzed, eyes closed, breathing comfortably, perhaps dreaming of a bone, or a boy, or of a boy throwing a bone.

“Mr. Rokitansky, how are you doing?” I asked.

Without opening his eyes, after fifteen seconds, in a husky slurred growl from deep down in his smushed brain he said: PURRTY GUD.

Pleased, I asked, “Mr. Rokitansky, what date is it today?”

PURRTY GUD. .

To all my questions, his answer was always the same. I felt sad. A professor, now a vegetable. Again I thought of my grandfather, and got a lump in my throat. Turning to Fats, I said, “This is too sad. He’s going to die.”

“No, he’s not,” said Fats. “He wants to, but he won’t.”

“He can’t go on like this.”

“Sure he can. Listen, Basch, there are a number of LAWS OF THE HOUSE OF GOD. LAW NUMBER ONE: GOMERS DON’T DIE.”

“That’s ridiculous. Of course they die.”

“I’ve never seen it, in a whole year here,” said Fats.

“They have to.”

“They don’t. They go on and on. Young people – like you and me – die, but not the gomers. Never seen it. Not once.”

“Why not?”

“I don’t know. Nobody knows. It’s amazing. Maybe they get past it. It’s pitiful. The worst.”

Potts came in, looking puzzled and concerned. He wanted the Fat Man’s help with Ina Goober. They left, and I turned back to Rokitansky. In the dim half-light I thought I saw tears trickling down the old man’s cheeks. Shame swept over me. My stomach churned. Had he heard what we’d said?

“Mr. Rokitansky, are you crying?” I asked, and I waited, as the long seconds ticked away, my guilt moaning inside me.

PURRTY GUD.

“But did you hear what we said about gomers?”

PURRTY GUD.

Someone once said that the point of art is to be more real than reality. The House Of God is way more real than reality. Reality wishes it could be anywhere close to as real as The House of God. This is a world where young people – the kid just out of school, the blushing new mother – die. Even normal old people – your grandmother, your grandpa – can die. But the most decrepit, demented people, the ones for whom every moment of artificially-prolonged life is a gratuitous misery and you pray at every moment that God will just let them find some peace – somehow they never die. They come into the hospital, they go back out to nursing homes, a few weeks later they’re back in the hospital, a few weeks later they’re back in their nursing homes, but they never die. This can’t be literally true. But it’s the subjective truth of working in a hospital. The Fat Man is right. I’ve been working in medicine for three years now, and I have seen my share of young people tragically cut off in the prime of life, and yet as far as I can remember I have never seen a gomer die. The magical realism of House of God describes the reality of medical professionals infinitely better than the rational world of hospital mortality statistics.

In the world of The House of God, the primary form of medical treatment is the TURF – the excuse to get a patient out of your care and on to somebody else’s. If the psychiatrist can’t stand a certain patient any longer, she finds some trivial abnormality in their bloodwork and TURFs to the medical floor. But she knows that if the medical doctor doesn’t want one of his patients, then he can interpret a trivial patient comment like “Being sick is so depressing” as suicidal ideation and TURF to psychiatry. At 3 AM on a Friday night, every patient is terrible, the urge to TURF is overwhelming, and a hospital starts to seem like a giant wheel uncoupled from the rest of the world, Psychiatry TURFING to Medicine TURFING to Surgery TURFING to Neurosurgery TURFING to Neurology TURFING back to Psychiatry again. Surely some treatment must get done somewhere? But where? It becomes a legend, The Place Where Treatment Happens, hidden in some far-off hospital wing accessible only to the pure-hearted. This sort of Kafkaesque picture is how medical care feels, and the genius of The House of God is that it accentuates the reality just a little bit until its fictional world is almost as magical-realist as the real one.

In the world of The House of God, medical intervention can only make patients worse:

Anna O. had started out on Jo’s service in perfect electrolyte balance, with each organ system working as perfectly as an 1878 model could. This, to my mind, included the brain, for wasn’t dementia a fail-safe and soothing oblivion of the machine to its own decay?

From being on the verge of a TURF back to the Hebrew House for the Incurables, as Anna knocked around the House of God in the steaming weeks of August, getting a skull film here and an LP there, she got worse, much worse. Given the stress of the dementia work-up, every organ system crumpled: in a domino progression the injection of radioactive dye for her brain scan shut down her kidneys, and the dye study of her kidneys overloaded her heart, and the medication for her heart made her vomit, which altered her electrolyte balance in a life-threatening way, which increased her dementia and shut down her bowel, which made her eligible for the bowel run, the cleanout for which dehydrated her and really shut down her tormented kidneys, which led to infection, the need for dialysis, and big-time complications of these big-time diseases. She and I both became exhausted, and she became very sick. Like the Yellow Man, she went through a phase of convulsing like a hooked tuna, and then went through a phase that was even more awesome, lying in bed deathly still, perhaps dying. I felt sad, for by this time, I liked her. I didn’t know what to do. I began to spend a good deal of time sitting with Anna, thinking.

The Fat Man was on call with me every third night as backup resident, and one night, searching for me to go to the ten o’clock meal, he found me with Anna, watching her trying to die.

“What the hell are you doing?” he asked.

I told him.

“Anna was on her way back to the Hebrew House, what happened – wait, don’t tell me. Jo decided to go all-out on her dementia, right?”

“Right. She looks like she’s going to die.”

“The only way she’ll die is if you murder her by doing what Jo says.”

“Yeah, but how can I do otherwise, with Jo breathing down my neck?”

“Easy. Do nothing with Anna, and hide it from Jo.”

“Hide it from Jo?”

“Sure. Continue the work-up in purely imaginary terms, buff the chart with the imaginary results of the imaginary tests, Anna will recover to her demented state, the work-up will show no treatable cause for it, and everybody’s happy. Nothing to it.”

“I’m not sure it’s ethical.”

“Is it ethical to murder this sweet gomere with your work-up?”

There was nothing I could say.”

After learning these medical secrets, Dr. Basch uses hook and crook to prevent his patients from getting any treatment. They end up healthier than anyone else in the hospital, and Basch becomes a contender for “Most Valuable Intern” – in typical House of God style, nobody knows if this award really exists or is just a rumor. His colleagues compete for another award, the “Black Crow”, which goes to the intern who gets the most autopsy consents from grieving families – and which the administration doesn’t realize incentivizes doctors to kill their patients. This is so reminiscent of the bizarre incentive systems in real hospitals that it hurts.

But as the year goes on, everyone gets more and more frazzled. One intern has a mental breakdown. Another commits suicide by jumping out of a hospital window (this isn’t dramatic exaggeration by the way; three junior doctors have committed suicide by jumping out of windows in the past three years in New York City alone). Dr. Basch runs through all sorts of interesting forms of neurosis. Finally, the end of the year approaches, the original crop of interns thinned-out but triumphant – and then they realize they have to do the whole thing again next year as residents, which is maybe a little less grueling but still in the same ballpark.

So they decide, en masse, to go into psychiatry, well-known to be a rare non-terrible residency. The author of House of God is a psychiatrist, so I guess this is only a spoiler insofar as you aren’t logically omniscient. When the Chief of Medicine learns that every single one of his hospital’s interns are going into psychiatry and there aren’t going to be any non-psychiatry residents in the whole hospital…

…okay, fine, I won’t spoil the ending. But suffice it to say I’m feeling pretty good about my career path right now.

II.

House of God does a weird form of figure-ground inversion.

An example of what I mean, taken from politics: some people think of government as another name for the things we do together, like providing food to the hungry, or ensuring that old people have the health care they need. These people know that some politicians are corrupt, and sometimes the money actually goes to whoever’s best at demanding pork, and the regulations sometimes favor whichever giant corporation has the best lobbyists. But this is viewed as a weird disease of the body politic, something that can be abstracted away as noise in the system.

And then there are other people who think of government as a giant pork-distribution system, where obviously representatives and bureaucrats, incentivized in every way to support the forces that provide them with campaign funding and personal prestige, will take those incentives. Obviously they’ll use the government to crush their enemies. Sometimes this system also involves the hungry getting food and the elderly getting medical care, as an epiphenomenon of its pork-distribution role, but this isn’t particularly important and can be abstracted away as noise.

I think I can go back and forth between these two models when I need to, but it’s a weird switch of perspective, where the parts you view as noise in one model resolve into the essence of the other and vice versa.

And House of God does this to medicine.

Doctors use certain assumptions, like:

1. The patient wants to get better, but there are scientific limits that usually make this impossible  
2. Medical treatment makes people healthier  
3. Treatment is determined by medical need and expertise

But in House of God, the assumptions get inverted:

1. The patient wants to just die peacefully, but there are bureaucratic limits that usually make this impossible  
2. Medical treatment makes people sicker  
3. Treatment is determined by what will make doctors look good without having to do much work

Everybody knows that those first three assumptions aren’t always true. Yes, sometimes we prolong life in contravention of patients’ wishes. Sometimes people mistakenly receive unnecessary treatment that causes complications. And sometimes care suffers because of doctors’ scheduling issues. But it’s easy to abstract away to an ideal medicine based on benevolence and reason, and then view everything else as rare and unfortunate deviations from the norm.

House of God goes the whole way and does a full figure-ground inversion. The outliers become the norm; good care becomes the rare deviation. What’s horrifying is how convincing it is. Real medicine looks at least as much like the bizarro-world of House of God as it does the world of the popular imagination where doctors are always wise, diagnoses always correct, and patients always grateful.

There have been a couple of studies finding that giving people health insurance doesn’t make them any healthier – see for example the RAND Health Insurance Experiment and the Oregon Medicaid Experiment. I’ve always been skeptical of these studies, because it seems logical that people who can afford health care will get more of it, and there are ten zillion studies showing various forms of health care to help. Insulin helps diabetes. Antibiotics help sepsis. Surgery helps appendicitis. To deny claims like these would be madness, yet the studies don’t lie. What is going on?

And the answer has to be somewhere in the bizarro-world of House of God. Real medical treatment looks precious little like the House MD model of rare serious disease -} diagnosis -} cure. At least as often, it’s like the House of God model where someone becomes inconvenient -} send to hospital -} one million unnecessary tests. Everyone agrees this is part of the story. House of God is a brilliant book in that it refactors perception to place it in the foreground.

But it’s brilliant because in the end it’s not just a romp through hilarious bureaucratic mishaps. There is as much genuine human goodness and compassion in this book as there is in any rousing speech by a medical school dean. The goodness is often mixed with horror – the doctor who has to fight off hordes of autopsy-consent-form-seekers to let a dying patient spend his last few seconds in peace, or the one who secretly slips euthanasia to a terminal patient begging for an end to the pain because he knows it’s the right thing to do.

The question posed here is “what do you do in a crazy cannibalistic system where it’s impossible to do good work and everyone is dying all around you?”, and the answer is “try as hard as you can to preserve whatever virtue you can, and to remain compassionate and human”. The protagonist swings wildly between “this is all bullshit and I’ll just make fun of these disgusting old people and call it a day” and “I need to save everybody and if I don’t I should hate myself forever”, and eventually like everybody, comes to some kind of synthesis where he recognizes he’s human, recognizes that his patients are human, and tries to deal with it with whatever humor and grace he can manage.

It’s hard enough for a book to be funny, and it’s hard enough for one to be deep, but a book like House of God that can be both at once within the space of a few sentences is an absolute treasure.

III.

I talked to my father about House of God, and I told him a few parts that seemed unrealistic. He told me that those parts were 100% true in 1978 when the book was written. I looked into it more, and ended up appreciating the work on a whole new level.

Uncle Tom’s Cabin is credited with kickstarting the emancipationist movement and maybe even causing the Civil War. The Jungle is famous for launching a whole new era of safety regulations. House of God has a place beside them in the pantheon of books that have changed the world.

The book’s “Second Law” is “GOMER GOES TO GROUND”: demented old people will inevitably fall out of their hospital bed and injure themselves. The book has a whole funny/horrifying scene where the senior resident explains his strategy for this eventuality: He leaves their beds low enough that patients won’t kill themselves when they fall, but high enough that they’ll probably break a bone or two and have to go to orthopaedic surgery – which takes them off his hands. Later, a medical student apes this procedure, a patient falls and breaks a bone or two, and everyone freaks out and tells him that it was a joke, that of course you don’t really arrange skeletal fractures for old people just to save yourself time, what kind of heartless moron could think such a thing? This is some nth-level meta-humor: the reader probably mistook it for real advice because it meshes so seamlessly with all of the other madness and horror, yet most of the other madness and horror in the book is easily recognizable by practicing doctors as a real part of the medical system. Actually, on the n+1st meta-level, I’m not at all sure that the resident wasn’t meant to be completely serious and then backtracked and called it a joke when it went wrong. For that matter, I’m far from sure this wasn’t a real medical practice in the 1970s.

I see enough falls that I wasn’t surprised to see them as a theme, but I thought the book exaggerated their omnipresence. My father said it didn’t – there were just far more falls back in the Old Days. Now hospitals are safer and falls are comparatively rare. Why? Because the government passed a law saying that insurance wouldn’t pay hospitals extra money for the extra days patients have to stay due to fall-related injuries. I am so serious about this. This, I think, is the n+2nd meta-level; amidst all its jokes-played-straight the book treats encouraging falls as an actual in-universe joke, and yet in the real world once hospitals were no longer incentivized to let patients fall the falls stopped.

How did people become aware of this kind of thing? How did the movement against it start? A lot of it seems to be because of House of God. Everyone in medicine knew about this sort of thing. But House of God made it common knowledge.

People were scared to speak up. Everyone thought that maybe they were just a uniquely bad person, or their hospital a uniquely bad institution. Anyone who raised some of these points was met with scorn by prestigious doctors who said that maybe they just weren’t cut out of medicine. House of God shaped medicine because it was the first thing to say what everybody was experiencing. Its terms like “gomer” and “turf” made it into the medical lexicon because they pointed to obvious features of reality nobody had the guts to talk about before.

Shem writes an afterword where he talks about the reaction to the book. Junior doctors and the public loved it. Senior doctors hated it. He tells the story of going to a medical conference. Someone asked who he was, and he said jokingly “I’m the most hated doctor here”. His interlocutor answered “Oh, don’t worry, I’m sure you’re not as bad as the guy who wrote that House of God book.”

But House of God gets credit for helping start movements to cut intern work hours, protect doctors from sleep deprivation, reduce patient falls, and teach empathy and communication skills. The moral of the story is: the courage to tell the truth is rare and powerful. More specifically: the courage to tell the truth is rare and powerful not just in Stalinist dictatorships and violent cults, but in apparently normal parts of everyday First World life. All of these differently loaded terms like “culture of silence” and “political correctness” point at a fear of rocking various boats with nothing but your imperfect first-person knowledge to go on. But a tiny crack in the wall can make a big difference.

IV.

In a closing scene, Dr. Basch and all of his fellow interns – interns who had broken into tears weekly, gotten burnt out, starting seeing psychiatrists, considered suicide, all this stuff, these interns who had smashed up against the unendurable horrors of medicine and held themselves together only by the promise that it would soon be over – the minute they graduate internship they change their tune:

It looked like all but two or three [interns] would stay. The Runt and I were definitely leaving; Chuck hadn’t yet said. The others were staying. In years to come they would spread out across America into academic centers and Fellowships, real red-hots in internal medicine, for they had been trained at the Best Medical School’s best House, the House of God. Although a few might kill themselves or get addicted or go crazy, by and large they’d repress and conform and perpetuate the Leggo [the Chief of Medicine] and the House and all the best medical stuff. [Eddie] had been praised by the Leggo that he could start off the second year as ward resident, with “a free rein” on his interns. And so, saying already that the internship been “not so bad,” he was preparing to indoctrinate his new charges: “I want them on their knees from day one.”

Shem’s author mouthpiece character Berry says:

It’s been inhuman. No wonder doctors are so distant in the face of the most poignant human dramas. The tragedy isn’t the crassness, but the lack of depth. Most people have some human reaction to their daily work, but doctors don’t. It’s an incredible paradox that being a doctor is so degrading and yet is so valued by society. In any community, the most respected group are doctors. [It’s] a terrific repression that makes doctors really believe that they are omnipotent healers. If you hear yourselves saying, ‘Well, this year wasn’t really that bad,’ you’re repressing, to put the next group through it. [But] it’s hard to say no. If you’re programmed from age six to be a doctor, invest years in it, develop your repressive skills so that you can’t even recall how miserable you were during internship, you can’t stop.

Shem’s thesis is that it isn’t just about not wanting to make waves or offend the Chief of Medicine. It’s about denying your own pain by identifying with the system.

This puts me in a weird spot. My internship (I find myself saying) wasn’t so bad. I can give you some arguments why this might be true – things have gotten a lot better since The House of God was published (with no small credit to Shem himself), a small community hospital in Michigan is less intense than Harvard Medical School’s training hospital, psychiatry interns sometimes have it easier than internal medicine interns since everyone knows this isn’t a permanent deal for them.

And yet I distinctly remember one night a long time ago, coming home from high school. I had noticed that all of the adults around me said high school was some of the best years of their lives and I would miss it when I was gone, and yet high school seemed objectively terrible. I wondered if there might be some bias or bizarre shift in memory that happened sometime in people’s twenties and gave them a localized amnesia or insanity. So I very distinctly recall telling myself “My current assessment is that high school is terrible, and if you ever find yourself remembering that high school was lovely, please be aware that your memories have been hijacked by some malevolent force.”

And God help me, but every single part of my brain is telling me that high school was lovely. I fondly remember all the friends I made, the crazy teachers I had to put up with, the science competitions I won, the lunches spent in the library reading whatever random stuff I could get my hands on. It seems like it was a blast. It’s hard for me to even trust that one memory as anything more than imagination or the product of a single bad day. But although high-school-me had a lot of issues, he generally had a decent head on his shoulders, and if he says my memories have been hijacked, then I grudgingly believe him.

So was my intern year a good learning experience? I have no idea and I’m not sure anyone else does either. It’s another type of figure-ground inversion: parade of horrors broken only by the occasional triumph, or clear sailing with a few bad moments?

On my last day of internship, one of my colleagues who was moving on said “I’m going to miss hating this place”. I’ve always remembered that phrase. Now I wonder if it’s some kind of weird snapshot of the exact moment of transition, the instant when “nightmarish ordeal” morphs into “halcyon days of youth”. This is why medicine has to be written as magical realism. How else to capture a world where people reliably go from agony to Stockholm Syndrome in the space of a day, and where the transition is so intermixed with the general weirdness that it doesn’t even merit special remark?

I found myself having more emotions reading House of God than I’ve had about anything in a long time. I don’t really know why. But I think it has something to do with this resignation to the general incommunicable weirdness all around anyone who works in medicine. Somehow Shem manages to avoid the normalization of insanity that happens to every young doctor, capture the exact subjective experience and write it down in a way that makes sense. And then, having put his finger right on the unbearable thing, he makes it funny and beautiful and poignant.

I tell her. Again I tell her about Dr. Sanders bleeding out in my lap, about the look in Potts’s eyes that night before he jumped, about my pushing the KCl into poor Saul. I tell her how ashamed I am for turning into a sarcastic bastard who calls the old ones gomers, how, during the ternship, I’d ridiculed them for their weaknesses, for throwing up their suffering in my face, for scaring me, for forcing me to do disgusting things to take care of them. I tell her how I want to live, compassionately, with the idea of death clearly in sight, and how I doubt I can do that, ever again. As I think back to what I’d gone through and what I’d become, sadness wells up and mixes with contempt. I put my head into Berry’s folds and weep, and curse, and shout, and weep.

“. . . and in your own way, you did. Someone had to care for the gomers; and this year, in your own way, you did.”

“The worst thing is this bitterness. I used to be different, gentle, even generous, didn’t I? I wasn’t always like this, was I?”

“I love who you are. To me, underneath it all, you’re still there:” She paused, and then, eyes sparkling, said, “And you might even be better.”

“What? What do you mean?”

“This might have been the only thing that could have awakened you. Your whole life has been a growing from the outside, mastering the challenges that others have set for you. Now, finally, you might just be growing from inside yourself.

He also frames all of it in the language of psychoanalysis, which is jarring and sounds preachy. I’ve ordered the sequel, Mount Misery, about his training as a psychoanalyst. Expect a review of that soon.

# You Are Still Crying Wolf

Go away, pro-Trump bots

[Update 5/30: I originally took this post down because it got picked up by some pro-Trump accounts that I think were bots which would post it several times a day to various places. I didn’t want to be involved in that or seeming to support it. After that I got in more trouble because I was supposedly “censoring” or “hiding” this. In order to avoid the Streisand Effect, I think my best option is to leave it up with this disclaimer at the top and the rest in a hard-to-read font.

Overall I think the predictions here have been borne out. Trump has proven himself an awful president and an offensive and incontinent person, but has continued to deny overt racism without any particular evidence that his denials are false (even as huge quantities of evidence of all Trump’s other misdeeds have piled up). Despite inflammatory rhetoric and unprecedented incompetence, he has continued to pursue policies basically within the Republican mainstream.

Many people have tried to tell me this post “didn’t age well”, but I think they’re just continuing to make the same mistakes this post warns against. For example, this post warns against thinking of Trump’s opposition to immigration as some sort of shocking new development when it’s similar to other immigration opposition from other presidents of both parties. I predicted that Trump would not deport more people than Obama. Given all of the evidence of ICE atrocities, I’ve gotten a lot of people writing me to complain about how wrong I obviously was, but this is on track to be correct – Obama deported more people in 2010 than Trump did in 2017, and ICE was pretty atrocious during his administration too. If it doesn’t seem that way, that’s because my original thesis is still correct – Trump is acting within the (worst parts of the) US mainstream, but the media is spinning it as an unprecedented atrocity.

The other event that led to me getting lots of hate mail was the Charlottesville protests. Again, the fact that a pan-alt-right gathering attracted a few hundred people matches my claim that there are probably only a few thousand white supremacists in the US (and is consistent with my claim that they are less influential than Satanists, given that thousands of people have shown up to some Satanist masses). There have been white supremacist rallies (including white supremacist rallies with violence against counterprotesters) during the Bush administration and the Obama administration. The only thing that’s new in the Age of Trump is that the media is paying much more attention and people are considering them worthy of notice, or symbolic of some larger trend.

Overall I continue to believe Trump is a terrible president, but also not “an overt white supremacist” by any reasonable definition. I think the Democrats have done a good job focusing on Trump’s other bad qualities and have mostly stopped centering their criticism around bumbling attempts to link him to the KKK or to neo-Nazis – and I no longer am particularly upset with how they’re handling this matter. That makes this post of historical interest only.]

[Content warning: hate crimes, Trump, racism. I have turned off comments to keep out bad people who might be attracted by this sort of thing. Avoid sharing in places where this will attract the wrong kind of attention, as per your best judgment. Please don’t interpret anything in this article to mean that Trump is not super terrible]

[Epistemic status: A reduction of a complicated issue to only 8000 words, because nobody would read it if it were longer. I think this is true but incomplete. I do not deny that Trump is being divisive and abusing identity politics in more subtle ways. I will try to discuss missing parts at more length later.]

I.

A New York Times article from last September that went viral only recently: Crying Wolf, Then Confronting Trump. It asks whether Democrats have “cried wolf” so many times that nobody believes them anymore. And so:

When “honorable and decent men” like McCain and Romney “are reflexively dubbed racists simply for opposing Democratic policies, the result is a G.O.P. electorate that doesn’t listen to admonitions when the genuine article is in their midst”.

I have a different perspective. Back in October 2015, I wrote that the picture of Trump as “the white power candidate” and “the first openly white supremacist candidate to have a shot at the Presidency in the modern era” was overblown. I said that “the media narrative that Trump is doing some kind of special appeal-to-white-voters voodoo is unsupported by any polling data”, and predicted that:

If Trump were the Republican nominee, he could probably count on equal or greater support from minorities as Romney or McCain before him.

Now the votes are in, and Trump got greater support from minorities than Romney or McCain before him. You can read the Washington Post article, Trump Got More Votes From People Of Color Than Romney Did, or look at the raw data (source)



Trump made gains among blacks. He made gains among Latinos. He made gains among Asians. The only major racial group where he didn’t get a gain of greater than 5% was white people. I want to repeat that: the group where Trump’s message resonated least over what we would predict from a generic Republican was the white population.

Nor was there some surge in white turnout. I don’t think we have official numbers yet, but by eyeballing what data we have it looks very much like whites turned out in equal or lesser numbers this year than in 2012, 2008, and so on. [EDIT: see counterpoint, countercounterpoint]

The media responded to all of this freely available data with articles like White Flight From Reality: Inside The Racist Panic That Fueled Donald Trump’s Victory and Make No Mistake: Donald Trump’s Win Represents A Racist “Whitelash”.

I stick to my thesis from October 2015. There is no evidence that Donald Trump is more racist than any past Republican candidate (or any other 70 year old white guy, for that matter). All this stuff about how he’s “the candidate of the KKK” and “the vanguard of a new white supremacist movement” is made up. It’s a catastrophic distraction from the dozens of other undeniable problems with Trump that could have convinced voters to abandon him. That it came to dominate the election cycle should be considered a horrifying indictment of our political discourse, in the same way that it would be a horrifying indictment of our political discourse if the entire Republican campaign had been based around the theory that Hillary Clinton was a secret Satanist. Yes, calling Romney a racist was crying wolf. But you are still crying wolf.

I avoided pushing this point any more since last October because I didn’t want to look like I was supporting Trump, or accidentally convince anyone else to support Trump. I think Trump’s election is a disaster. He has no plan, he’s dangerously trigger-happy, and his unilateralism threatens aid to developing countries, one of the most effective ways we currently help other people. I thought and still think a Trump presidency will be a disaster.

But since we’re past the point where we can prevent it, I want to present my case.

I realize that all of this is going to make me sound like a crazy person and put me completely at odds with every respectable thinker in the media, but luckily, being a crazy person at odds with every respectable thinker in the media has been a pretty good ticket to predictive accuracy lately, so whatever.

II.

First, I want to go over Donald Trump’s official, explicit campaign message. Yes, it’s possible for candidates’ secret feelings to differ from their explicit messages, but the things they say every single day and put on their website and include in their speeches are still worth going over to see what image they want to project.

Trump’s official message has been the same vague feel-good pro-diversity rhetoric as any other politician. Here’s Trump on African Americans:

It is my highest and greatest hope that the Republican Party can be the home in the future and forevermore for African-Americans and the African-American vote because I will produce, and I will get others to produce, and we know for a fact it doesn’t work with the Democrats and it certainly doesn’t work with Hillary.

When I am President, I will work to ensure that all of our kids are treated equally, and protected equally. Every action I take, I will ask myself: does this make life better for young Americans in Baltimore, Chicago, Detroit, Ferguson who have as much of a right to live out their dreams as any other child in America?

African-American citizens have sacrificed so much for this nation. They have fought and died in every war since the Revolution, and from the pews and the picket lines they have lifted up the conscience of our country in the long march for Civil Rights. Yet, too many African-Americans have been left behind.

No group in America has been more harmed by Hillary Clinton’s policies than African-Americans. No group. No group. If Hillary Clinton’s goal was to inflict pain on the African-American community, she could not have done a better job. It’s a disgrace. Tonight, I am asking for the vote of every African-American citizen in this country who wants a better future.

And at the end of four years I guarantee that I will get over 95% of the African-American vote. I promise you. Because I will produce for the inner-cities and I will produce for the African-Americans.

America must reject the bigotry of Hillary Clinton who sees communities of color only as votes, not as human beings worthy of a better future.

On Hispanics:

I have just landed having returned from a very important and special meeting with the President of Mexico…we discussed the great contributions of Mexican-American citizens to our two countries, my love for the people of Mexico, and the close friendship between our two nations.

I employ thousands and thousands of Hispanics. I love the people. They’re great workers. They’re fantastic people and they want legal immigration. I’ll take jobs back from China, I’ll take jobs back from Japan. The Hispanics are going to get those jobs, and they’re going to love Trump.

On his campaign:

It’s a movement comprised of Americans from all races, religions, backgrounds and beliefs who want and expect our government to serve the people, and serve the people it will.

Trump’s campaign photos are consistent with a desire to present the same message:



This wasn’t a scripted appearance forced by his campaign staff. According to the Washington Times:

Trump walked on stage in Greeley, Colorado to a large cheering crowd when he spotted a rainbow flag in the audience. As the music blasted through the speakers, Mr. Trump pointed to a supporter as if to ask if he could see his flag and then motioned for a campaign worker to help retrieve the LGBT symbol of equality from the attendee.

Within seconds, Mr. Trump was walking around the platform with the rainbow flag in his hands and moments later unfurled it in full display. You could see a huge smile on Mr. Trump’s face as he walked to both sides of the stage to proudly hold up the rainbow flag announcing support from the gay and lesbian community.

Trump campaign spokesman Jason Miller told me, “Mr. Trump is campaigning to be President for ALL Americans and was proud to carry the ‘LGBT for Trump‘ rainbow flag on stage in Greeley, CO yesterday.



This is just a tiny representative sample, but the rest is very similar. Trump has gone from campaign stop to campaign stop talking about how much he likes and respects minorities and wants to fight for them.

And if you believe he’s lying, fine. Yet I notice that people accusing Trump of racism use the word “openly” like a tic. He’s never just “racist” or “white supremacist”. He’s always “openly racist” and “openly white supremacist”. Trump is openly racist, openly racist, openly racist, openly racist, openly racist, openly racist, openly racist. Trump is running on pure white supremacy, has thrown off the last pretense that his campaign is not about bigotry, has the slogan Make American Openly White Supremacist Again, is an openly white supremacist nominee, etc, etc, etc. And I’ve seen a few dozen articles like this where people say that “the bright side of a Trump victory is that finally America admitted its racism out in the open so nobody can pretend it’s not there anymore.”

This, I think, is the first level of crying wolf. What if, one day, there is a candidate who hates black people so much that he doesn’t go on a campaign stop to a traditionally black church in Detroit, talk about all of the contributions black people have made to America, promise to fight for black people, and say that his campaign is about opposing racism in all its forms? What if there’s a candidate who does something more like, say, go to a KKK meeting and say that black people are inferior and only whites are real Americans?

We might want to use words like “openly racist” or “openly white supremacist” to describe him. And at that point, nobody will listen, because we wasted “openly white supremacist” on the guy who tweets pictures of himself eating a taco on Cinco de Mayo while saying “I love Hispanics!”

III.

A rundown of some contrary talking points:

1. Is Trump getting a lot of his support from white supremacist organizations?

No, because there are not enough organized white supremacists to make up “a lot” of anyone’s support.

According to Wikipedia on KKK membership:

As of 2016, the Anti-Defamation League puts total Klan membership nationwide at around 3,000, while the Southern Poverty Law Center puts it at 6,000 members total

The KKK is really small. They could all stay in the same hotel with a bunch of free rooms left over. Or put another way: the entire membership of the KKK is less than the daily readership of this blog.

If you Google “trump KKK”, you get 14.8 million results. I know that Google’s list of results numbers isn’t very accurate. Yet even if they’re inflating the numbers by 1000x, and there were only about 14,000 news articles about the supposed Trump-KKK connection this election, there are still two to three articles about a Trump-KKK connection for every single Klansman in the world.

I don’t see any sign that there are other official white supremacy movements that are larger than the Klan, or even enough other small ones to substantially raise the estimate of people involved. David Duke called a big pan-white-supremacist meeting in New Orleans in 2005, and despite getting groups from across North America and Europe he was only able to muster 300 attendees (by comparison, NAACP conventions routinely get 10,000).

My guess is that the number of organized white supremacists in the country is in the very low five digits.

2. Is Trump getting a lot of his support from online white nationalists and the alt-right?

No, for the same reason.

The alt-right is mostly an online movement, which makes it hard to measure. The three main alt-right hubs I know of are /r/altright, Stormfront, and 4chan’s politics board.

The only one that displays clear user statistics is /r/altright, which says that there are about 5,000 registered accounts. The real number is probably less – some people change accounts, some people post once and disappear, and some non-white-nationalists probably go there to argue. But sure, let’s say that community has 5,000 members.

Stormfront’s user statistics say it gets about 30,000 visits/day, of which 60% are American. My own blog gets about 8,000 visits/day , and the measurable communities associated with it (the subreddit, people who follow my social media accounts) have between 2000 – 8000 followers. If this kind of thing scales, then it suggests about 10,000 people active in the Stormfront community.

4chan boasts about 1 million visits/day. About half seem to be American. Unclear how many go to the politics board and how many are just there for the anime and video games, but Wikipedia says that /b/ is the largest board with 30% of 4Chan’s traffic, so /pol/ must be less than that. If we assume /pol/ gets 20% of 4chan traffic, and that 50% of the people on /pol/ are serious alt-rightists and not dissenters or trolls, the same scaling factors give us about 25,000 – 50,000 American alt-rightists on 4Chan.

Taking into account the existence of some kind of long tail of alt-right websites, I still think the population of the online US alt-right is somewhere in the mid five-digits, maybe 50,000 or so.

50,000 is more than the 5,000 Klansmen. But it’s still 0.02% of the US population. It’s still about the same order of magnitude as the Nation of Islam, which has about 30,000 – 60,000 members, or the Church of Satan, which has about 20,000. It’s not quite at the level of the Hare Krishnas, who boast 100,000 US members. This is not a “voting bloc” in the sense of somebody it’s important to appeal to. It isn’t a “political force” (especially when it’s mostly, as per the 4chan stereotype, unemployed teenagers in their parents’ basements.)

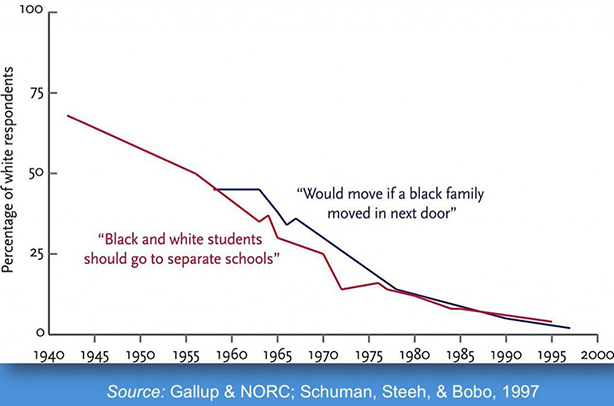
So the mainstream narrative is that Trump is okay with alienating minorities (= 118 million people), whites who abhor racism and would never vote for a racist (if even 20% of whites, = 40 million people), most of the media, most business, and most foreign countries – in order to win the support of about 50,000 poorly organized and generally dysfunctional people, many of whom are too young to vote anyway.

Caring about who the KKK or the alt-right supports is a lot like caring about who Satanists support. It’s not something you would do if you wanted to understand real political forces. It’s only something you would do if you want to connect an opposing candidate to the most outrageous caricature of evil you can find on short notice.

3. Is Trump getting a lot of his support from people who wouldn’t join white nationalist groups, aren’t in the online alt-right, but still privately hold some kind of white supremacist position?

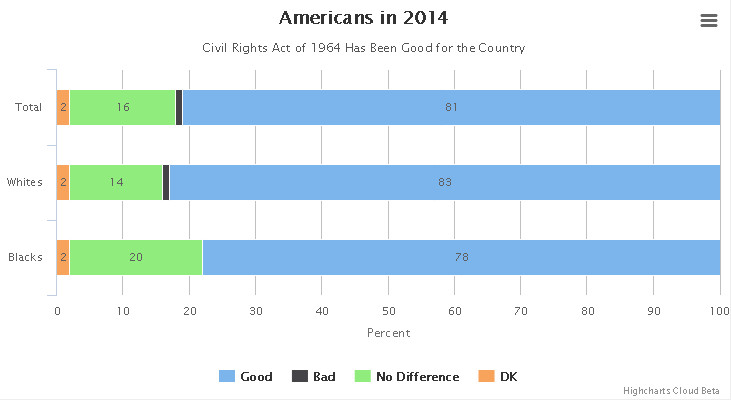
There are surprisingly few polls that just straight out ask a representative sample of the population “Are you white supremacist?”.

I can find a couple of polls that sort of get at this question in useful ways.

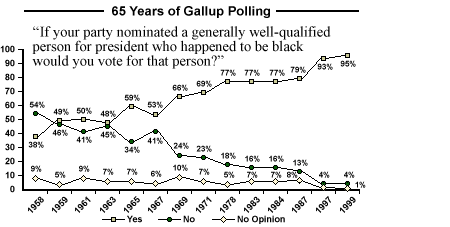


This poll from Gallup asks white Americans their support for school segregation and whether they would move out if a black family moved in next door. It declines from about 50% in 1960 to an amount too small to measure in the 1990s, maybe 1-2%, where it presumably remains today.

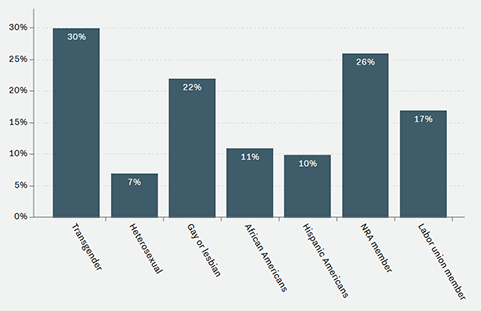
(this graph also seems relevant to the stories of how Trump’s father would try to keep blacks out of his majority-white real estate developments in the late 60s/early 70s – note that at that time 33% of white families would move out if a black person moved in next door)



Here’s a CBS News poll from 2014 asking Americans their opinion on the Civil Rights Act that legally prohibited discrimination. Once again, the number of whites who think it was a bad thing is too small to measure meaningfully, but looks like maybe 1-2%. Of note, whites were more convinced the Civil Rights Act was good than blacks were, though I guess it depends on the margin of error.



Another Gallup graph here, with the percent of people who would vs. wouldn’t vote for an otherwise-qualified black candidate for President. It goes from 54% in 1968 to 5% in 1999; later polls that aren’t included on the graph give numbers from 4% to 7%, which sounds probably within the margin of error.



This is a Vox poll asking how many people had favorable vs. unfavorable views of different groups. 11% admit to “somewhat unfavorable” or “very unfavorable” views of blacks, which sounds bad, except that 7% of people admit to unfavorable views of heterosexuals by the same definition. This makes me think “have an unfavorable view about this group” is not a very high bar. If we restrict true “white supremacists” to those who have only “very unfavorable” views of blacks, this is 3%, well in line with our other sources.

(of note, 1% of respondents had “never heard of” blacks. Um…)

Maybe a better way of looking for racists: David Duke ran for Senate in Louisiana this year. He came in seventh with 58,000 votes (3%). Multiplied over 50 states, that would suggest 2.5 million people who would vote for a leading white supremacist. On the other hand, Louisiana is one of the most racist states (for example, Slate’s investigation found that it led the US in percent of racist tweets) and one expects Duke would have had more trouble in eg Vermont. Adjusting for racism level as measured in tweets, it looks like there would be about 1 million Duke voters in a nationwide contest. That’s a little less than 1% of voters.

So our different ways of defining “open white supremacist”, even for definitions of “open” so vague they include admitting it on anonymous surveys, suggest maybe 1-2%, 1-2%, 4-7%, 3-11%, and 1-3%.

But doesn’t this still mean there are some white supremacists? Isn’t this still really important?

I mean, kind of. But remember that 4% of Americans believe that lizardmen control all major governments. And 5% of Obama voters believe that Obama is the Antichrist. The white supremacist vote is about the same as the lizardmen-control-everything vote, or the Obama-is-the-Antichrist-but-I-support-him-anyway vote.

(and most of these people are in Solid South red states and don’t matter in the electoral calculus anyway.)

4. Aren’t there a lot of voters who, although not willing to vote for David Duke or even willing to express negative feelings about black people on a poll, still have implicit racist feelings, the kind where they’re nervous when they see a black guy on a deserted street at night?

Probably. And this is why I am talking about crying wolf. If you wanted to worry about the voter with subconscious racist attitudes carefully hidden even from themselves, you shouldn’t have used the words “openly white supremacist KKK supporter” like a verbal tic.

5. But even if Donald Trump isn’t openly white supremacist, didn’t he get an endorsement from KKK leader David Duke? Didn’t he refuse to reject that endorsement? Doesn’t that mean that he secretly wants to court the white supremacist vote?

The answer is no on all counts.

No, Donald Trump did not get an endorsement from KKK leader David Duke. Duke has spoken out in favor of Trump, but refused to give a formal endorsement. You can read the explanation straight from the horse’s mouth at davidduke.com: “The ZioMedia Lies: I Have Not Endorsed Donald Trump” (content warning: exactly what you would expect). If you don’t want that site in your browser history, you can read the same story at The International Business Times.

No, Donald Trump did not refuse to reject the endorsement. From Politico.com:

Donald Trump says he isn’t interested in the endorsement of David Duke, the anti-Semitic former Ku Klux Klan leader who praised the GOP presidential hopeful earlier this week on his radio show.

“I don’t need his endorsement; I certainly wouldn’t want his endorsement,” Trump said during an interview with Bloomberg’s Mark Halperin and John Heilemann. He added: “I don’t need anyone’s endorsement.”

Asked whether he would repudiate the endorsement, Trump said “Sure, I would if that would make you feel better.”

From Washington Post:

ABC NEWS: “So, are you prepared right now to make a clear and unequivocal statement renouncing the support of all white supremacists?”

TRUMP: “Of course, I am. I mean, there’s nobody that’s done so much for equality as I have. You take a look at Palm Beach, Florida, I built the Mar-a-Lago Club, totally open to everybody; a club that frankly set a new standard in clubs and a new standard in Palm Beach and I’ve gotten great credit for it. That is totally open to everybody. So, of course, I am.”

From CNN:

“David Duke is a bad person, who I disavowed on numerous occasions over the years,” Trump said on MSNBC’s “Morning Joe.”

“I disavowed him. I disavowed the KKK,” Trump added. “Do you want me to do it again for the 12th time? I disavowed him in the past, I disavow him now.”

The concern comes from a single interview February 28, where Trump was asked to renounce support from David Duke and the KKK, where he gave a non-answer:

“I have to look at the group. I mean, I don’t know what group you’re talking about,” Trump said. “You wouldn’t want me to condemn a group that I know nothing about. I’d have to look. If you would send me a list of the groups, I will do research on them and certainly I would disavow if I thought there was something wrong. You may have groups in there that are totally fine — it would be very unfair. So give me a list of the groups and I’ll let you know.”

This is pretty bad. But the next day Trump was saying that of course he denounced the KKK and blaming a “bad earpiece” for not being able to understand what the interviewer was saying.

Trump’s bad earpiece explanation doesn’t hold water – he repeated the name “David Duke” in his answer, so he obviously heard it. And his claim that he didn’t know who David Duke was doesn’t make sense – he’s mentioned Duke before in various contexts.

But it’s actually worth taking a look at those contexts. In 2000, Trump was already considering running for President. His friend Jesse Ventura suggested he seek the Presidential nomination of Ross Perot’s Reform Party. Trump agreed and started putting together a small campaign (interesting historical trivia: he wanted Oprah Winfrey as a running mate). But after some infighting in the Reform Party, Ventura was kicked out in favor of a faction led by populist Pat Buchanan, who had some support from David Duke. Trump closed his presidential bid, saying: “The Reform Party now includes a Klansman, Mr. Duke, a neo-Nazi, Mr. Buchanan, and a communist, Ms. Fulani. This is not company I wish to keep.” Later he continued to condemn the party, saying “You’ve got David Duke just joined — a bigot, a racist, a problem. I mean, this is not exactly the people you want in your party.”

So we have Trump – who loudly condemned Duke before February 28th, and who loudly condemned Duke after February 28th – saying on February 28th that he wanted to “look into” who David Duke was before refusing his (non-existent) endorsement. I’m not super sure what’s going on. It’s possible he wanted to check to see whether it was politically advantageous to officially reject it, which I agree is itself pretty creepy.

But notice that the evidence on the side of Trump being against David Duke includes twenty years of unambiguous statements to that effect. And the evidence of Trump not being against David Duke includes one statement along the lines of “I don’t know who he is but I’ll look into it” on an interview one time which he later blamed on a bad earpiece and said he totally disavowed.

This gets back to my doubts about “dog whistles”. Dog whistling seems to be the theory that if you want to know what someone really believes, you have to throw away decades of consistent statements supporting the side of an issue that everyone else in the world supports, and instead pay attention only to one weird out-of-character non-statement which implies he supports a totally taboo position which is perhaps literally the most unpopular thing it is possible to think.

And then you have to imagine some of the most brilliant rhetoricians and persuaders in the world are calculating that it’s worth risking exposure this taboo belief in order to win support from a tiny group with five-digit membership whose support nobody wants, by sending a secret message, which inevitably every single media outlet in the world instantly picks up on and makes the focus of all their coverage for the rest of the election.

Finally, no, none of this suggests that Donald Trump is courting the white supremacist vote. Anybody can endorse anybody with or without their consent. Did you know that the head of the US Communist Party endorsed Hillary, and Hillary never (as far as I know) “renounced” their endorsement? Does that mean Hillary is a Communist? Did you know that a leader of a murderous black supremacist cult supported Donald Trump and Trump said that he “loved” him? Does that mean Trump is a black supremacist? The only time this weird “X endorsed Y, that means Y must support X” thing is brought out, is in favor of the media narrative painting Trump to be a racist.

This, to me, is another form of crying wolf. One day you might have a candidate who openly courts the KKK, in the sense of having a campaign platform saying “I like the KKK and value their support”, speaking at Klan meetings, et cetera. And instead, you’ve wasted the phrase “openly courts the KKK” on somebody with a twenty year history of loudly condemning the KKK, plus one weird interview where he said he didn’t know anything about it, then changed his mind the next day and said he hates them.

6. What about Trump’s “drugs and crime” speech about Mexicans?

Trump said that:

When Mexico sends its people, they’re not sending their best. They’re not sending you. They’re sending people that have lots of problems, and they’re bringing those problems with us. They’re bringing drugs. They’re bringing crime. Their rapists. And some, I assume, are good people.

Note how totally non-racist this statement is. I’m serious. It’s anti-illegal-immigrant. But in terms of race, it’s saying Latinos (like every race) include both good and bad people, and the bad people are the ones coming over here. It suggests a picture of Mexicans as including some of the best people – but those generally aren’t the ones who are coming illegally.

Compare to eg Bill Clinton’s 1996 platform (all emphasis mine):

We cannot tolerate illegal immigration and we must stop it. For years before Bill Clinton became President, Washington talked tough but failed to act. In 1992, our borders might as well not have existed. The border was under-patrolled, and what patrols there were, were under-equipped. Drugs flowed freely. Illegal immigration was rampant. Criminal immigrants, deported after committing crimes in America, returned the very next day to commit crimes again. President Clinton is making our border a place where the law is respected and drugs and illegal immigrants are turned away.

Or John McCain in 2008:

Border security is essential to national security. In an age of terrorism, drug cartels, and criminal gangs, allowing millions of unidentified persons to enter and remain in this country poses grave risks to the sovereignty of the United States and the security of its people.

Trump’s platform contains similar language – and, like all past platforms, also contains language praising legal immigrants:

Just as immigrant labor helped build our country in the past, today’s legal immigrants are making vital contributions in every aspect of national life. Their industry and commitment to American values strengthens our economy, enriches our culture, and enables us to better understand and more effectively compete with the rest of the world.

We are particularly grateful to the thousands of new legal immigrants, many of them not yet citizens, who are serving in the Armed Forces and among first responders. Their patriotism should encourage all to embrace the newcomers legally among us, assist their journey to full citizenship, and help their communities avoid isolation from the mainstream of society. We are also thankful for the many legal immigrants who continue to contribute to American society.

When Democrats and Republicans alike over the last twenty years say that we are a nation of immigrants but that illegal immigrants threaten our security, or may be criminals or drug pushers, they’re met with yawns. When Trump says exactly the same thing, he’s Literally the KKK.

7. What about the border wall? Doesn’t that mean Trump must hate Mexicans?

As multiple sources point out, both Hillary and Obama voted for the Secure Fence Act of 2006, which put up a 700 mile fence along the US-Mexican border. Politifact says that Hillary and Obama wanted a 700 mile fence but Trump wants a 1000 mile wall, so these are totally different. But really? Support a 700 mile fence, and you’re the champion of diversity and all that is right in the world; support a 1000 mile wall and there’s no possible explanation besides white nationalism?

8. Isn’t Trump anti-immigrant?

He’s at least anti-undocumented immigrant, which is close to being anti-immigrant. And while one can argue that “anti-immigrant” is different than “racist”, I would agree that probably nobody cares that much about British or German immigrants, suggesting that some racial element is involved.

But I think when Trump voters talk about “globalists”, they’re pointing at how they model this very differently from the people they criticize.

In one model, immigration is a right. You need a very strong reason to take it away from anybody, and such decisions should be carefully inspected to make sure no one is losing the right unfairly. It’s like a store: everyone should be allowed to come in and shop and if a manager refused someone entry then they better have a darned good reason.

In another, immigration is a privilege which members of a community extend at their pleasure to other people whom they think would be a good fit for their community. It’s like a home: you can invite your friends to come live with you, but if someone gives you a vague bad feeling or seems like a good person who’s just incompatible with your current lifestyle, you have the right not to invite them and it would be criminal for them to barge in anyway.

It looks like many Clinton supporters believe in the first model, and many Trump supporters in the second model. I think this ties into deeper differences – Clinton supporters are more atomized and individualist, Trump supporters stronger believers in culture and community.

In the second model, the community gets to decide how many immigrants come in and on what terms. Most of the Trump supporters I know are happy to let in a reasonable amount, but they get very angry when people who weren’t invited or approved by the community come in anyway and insist that everyone else make way for them.

Calling this “open white supremacy” seems like those libertarians who call public buses Communism, except if “Communism” got worn out on the euphemism treadmill and they started calling public buses “overt Soviet-style Stalinism”.

9. Don’t Trump voters oppose the Emancipation Proclamation that freed the slaves?

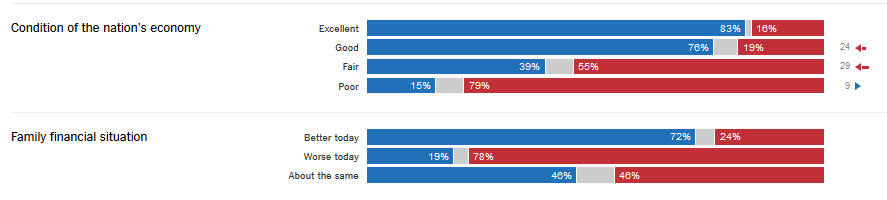
This was in New York Times, Vox, Huffington Post, Time, et cetera. It’s very misleading. See Snopes for full explanation.

10. Isn’t Trump anti-Semitic?

I feel like an attempt to avoid crying wolf might reserve that term for people who didn’t win an Israeli poll on what candidate would best represent Israel’s interests, or doesn’t have a child who converted to Judaism, or hasn’t won various awards from the American Jewish community for his contributions to Israel and American Judaism, or wasn’t the grand marshal of a Salute To Israel Parade, or…

11. Don’t we know that Trump voters are motivated by racism because somebody checked and likelihood of being a Trump voter doesn’t correlate with some statistic or other supposedly measuring economic anxiety?

Although economic issues are only one part of Trump voters’ concerns, they certainly are a part. You just have to look in the right places. See also:



12. Don’t we know that Trump voters are motivated by racism because despite all the stuff about economic anxiety, rich people were more likely to vote Trump than poor people?

I keep hearing stuff like this, and aside from the object-level question, I think it’s important to note the way in which this kind of thing makes racism the null hypothesis. “You say it’s X, but you can’t prove it, so it’s racism”.

Anyway, in this particular case, there’s a simple answer. Yes, Republicans are traditionally the party of the rich. What’s different about this election is that far more poor people voted Republican than usual, and far more rich people voted Democrat than usual.



Poor people were 16 percentage points more likely to vote Republican this election than last time around, but rich people (well, the richest bracket NYT got data about) were 9 percentage points more likely to vote Democrat. This is consistent with economic anxiety playing a big role.

13. Doesn’t Trump want to ban (or “extreme vet”, or whatever) Muslims entering the country?

Yes, and this is awful.

But why do he (and his supporters) want to ban/vet Muslims, and not Hindus or Kenyans, even though most Muslims are white(ish) and most Hindus and Kenyans aren’t? Trump and his supporters are concerned about terrorism, probably since the San Bernardino shooting and Pulse nightclub massacre dominated headlines this election season.

You can argue that he and his supporters are biased for caring more about terrorism than about furniture-related injuries, which kill several times more Americans than terrorists do each year. But do you see how there’s a difference between “cognitive bias that makes you unreasonably afraid” versus “white supremacy”?

I agree that this is getting into murky territory and that a better answer here would be to deconstruct the word “racism” into a lot of very heterogenous parts, one of which means exactly this sort of thing. But as I pointed out in Part 4, a lot of these accusations shy away from the word “racism” precisely because it’s an ambiguous thing with many heterogenous parts, some of which are understandable and resemble the sort of thing normal-but-flawed human beings might think. Now they say “KKK white nationalism” or “overt white supremacy”. These terms are powerful exactly because they do not permit the gradations of meaning which this subject demands.

Let me say this for the millionth time. I’m not saying Trump doesn’t have some racist attitudes and policies. I am saying that talk of “entire campaign built around white supremacy” and “the white power candidate” is deliberate and dangerous exaggeration. Lots of people (and not just whites!) are hasty to generalize from “ISIS is scary” to “I am scared of all Muslims”. This needs to be called out and fought, but it needs to be done in an understanding way, not with cries of “KKK WHITE SUPREMACY!”

14. Haven’t there been hundreds of incidents of Trump-related hate crimes?

This isn’t a criticism of Trump per se (he’s demanded that his supporters avoid hate crimes), but it seems relevant to the general tenor of the campaign.

SPLC said they have 300 such hate incidents, although their definition of “hate incident” includes things like “someone overheard a racist comment in someone else’s private conversation, then challenged them about it and got laughed at”. Let’s take that number at face value (though see here)

If 47% of America supports Trump (= the percent of vote he got extrapolated to assume non-voters feel the same way), there are 150,000,000 Trump supporters. That means there has been one hate incident per 500,000 Trump supporters.

But aren’t there probably lots of incidents that haven’t been reported to SLPC? Maybe. Maybe there’s two unreported attacks for every reported one, which means that the total is one per 150,000 Trump supporters. Or maybe there are ten unreported attacks for every reported one, which means that the total is one per 45,000 Trump supporters. Since nobody has any idea about this, it seems weird to draw conclusions from it.

Oh, also, I looked on right-wing sites to see if there are complaints of harassment and attacks by Hillary supporters, and there are. Among the stories I was able to confirm on moderately trustworthy news sites that had investigated them somewhat (a higher standard than the SLPC holds their reports to) are ones about how Hillary supporters have beaten up people for wearing Trump hats, screamed encouragement as a mob beat up a man who they thought voted Trump, knocked over elderly people, beaten up a high school girl for supporting Trump on Instagram, defaced monuments with graffiti saying “DIE WHITES DIE”, advocated raping Melania Trump, kicked a black homeless woman who was holding a Trump sign, attacked a pregnant woman stuck in her car, with a baseball bat, screamed at children who vote Trump in a mock school election, etc, etc, etc.

But please, keep talking about how somebody finding a swastika scrawled in a school bathroom means that every single Trump supporter is scum and Trump’s whole campaign was based on hatred.

15. Don’t we know that Trump supports racist violence because, when some of his supporters beat up a Latino man, he just said they were “passionate”?

All those protests above? The anti-Trump protests that have resulted in a lot of violence and property damage and arrests? With people chanting “KILL TRUMP” and all that?

When Trump was asked for comment, he tweeted “Love the fact that the small groups of protesters last night have passion for our great country”.

I have no idea how his mind works and am frankly boggled by all of this, but calling violent protesters “passionate” just seems to be a thing of his.

16. But didn’t Trump…

Whatever bizarre, divisive, ill-advised, and revolting thing you’re about to mention, the answer is probably yes.

This is equally true on race-related and non-race-related issues. People ask “How could Trump believe the wacky conspiracy theory that Obama was born in Kenya, if he wasn’t racist?” I don’t know. How could Trump believe the wacky conspiracy theory that vaccines cause autism? How could Trump believe the wacky conspiracy theory that the Clintons killed Vince Foster? How could Trump believe the wacky conspiracy theory that Ted Cruz’s father shot JFK?

Trump will apparently believe anything for any reason, especially about his political opponents. If Clinton had been black but Obama white, we’d be hearing that the Vince Foster conspiracy theory proves Trump’s bigotry, and the birtherism was just harmless wackiness.

Likewise, how could Trump insult a Mexican judge just for being Mexican? I don’t know. How could Trump insult a disabled reporter just for being disabled? How could Trump insult John McCain just for being a beloved war hero? Every single person who’s opposed him, Trump has insulted in various offensive ways, including 140 separate incidents of him calling someone “dopey” or “dummy” on Twitter, and you expect him to hold his mouth just because the guy is a Mexican?

I don’t think people appreciate how weird this guy is. His weird way of speaking. His catchphrases like “haters and losers!” or “Sad!”. His tendency to avoid perfectly reasonable questions in favor of meandering tangents about Mar-a-Lago. The ability to bait him into saying basically anything just by telling him people who don’t like him think he shouldn’t.

If you insist that Trump would have to be racist to say or do whatever awful thing he just said or did, you are giving him too much credit. Trump is just randomly and bizarrely terrible. Sometimes his random and bizarre terribleness is about white people, and then we laugh it off. Sometimes it’s about minorities, and then we interpret it as racism.

17. Isn’t this a lot of special pleading? Like, sure, you can make up various non-racist explanations for every single racist-sounding thing Trump says, and say a lot of it is just coincidence or Trump being inexplicably weird, but eventually the coincidences start adding up. You have to look at this kind of thing in context.

I actually disagree with this really strongly and this point deserves a post of its own because it’s really important. But let me try to briefly explain what I mean.

Suppose you’re talking to one of those ancient-Atlantean secrets-of-the-Pyramids people. They give you various pieces of evidence for their latest crazy theory, such as (and all of these are true):

1. The latitude of the Great Pyramid matches the speed of light in a vacuum to five decimal places.  
2. Famous prophet Edgar Cayce, who predicted a lot of stuff with uncanny accuracy, said he had seen ancient Atlanteans building the Pyramid in a vision.  
3. There are hieroglyphs near the pyramid that look a lot like pictures of helicopters.  
4. In his dialogue Critias, Plato relayed a tradition of secret knowledge describing a 9,000-year-old Atlantean civilization.  
5. The Egyptian pyramids look a lot like the Mesoamerican pyramids, and the Mesoamerican name for the ancient home of civilization is “Aztlan”  
6. There’s an underwater road in the Caribbean, whose discovery Edgar Cayce predicted, and which he said was built by Atlantis  
7. There are underwater pyramids near the island of Yonaguni.  
8. The Sphinx has apparent signs of water erosion, which would mean it has to be more than 10,000 years old.

She asks you, the reasonable and well-educated supporter of the archaeological consensus, to explain these facts. After looking through the literature, you come up with the following:

1. This is just a weird coincidence.  
2. Prophecies have so many degrees of freedom that anyone who gets even a little lucky can sound “uncannily accurate”, and this is probably just what happened with Cayce, so who cares what he thinks?  
3. Lots of things look like helicopters, so whatever.  
4. Plato was probably lying, or maybe speaking in metaphors.  
5. There are only so many ways to build big stone things, and “pyramid” is a natural form. The “Atlantis/Atzlan” thing is probably a coincidence.  
6. Those are probably just rocks in the shape of a road, and Edgar Cayce just got lucky.  
7. Those are probably just rocks in the shape of pyramids. But if they do turn out to be real, that area was submerged pretty recently under the consensus understanding of geology, so they might also just be pyramids built by a perfectly normal non-Atlantean civilization.  
8. We still don’t understand everything about erosion, and there could be some reason why an object less than 10,000 years old could have erosion patterns typical of older objects.

I want you to read those last eight points from the view of an Atlantis believer, and realize that they sound really weaselly. They’re all “Yeah, but that’s probably a coincidence”, and “Look, we don’t know exactly why this thing happened, but it’s probably not Atlantis, so shut up.”

This is the natural pattern you get when challenging a false theory. The theory was built out of random noise and ad hoc misinterpretations, so the refutation will have to be “every one of your multiple superficially plausible points is random noise, or else it’s a misinterpretation for a different reason”.

If you believe in Atlantis, then each of the seven facts being true provides “context” in which to interpret the last one. Plato said there was an Atlantis that sunk underneath the sea, so of course we should explain the mysterious undersea ruins in that context. The logic is flawless, it’s just that you’re wrong about everything.

This is how I feel about demands that we interpret Trump’s statements “in context”, too.

IV.

Why am I harping on this?

I work in mental health. So far I have had two patients express Trump-related suicidal ideation. One of them ended up in the emergency room, although luckily both of them are now safe and well. I have heard secondhand of several more.

Like Snopes, I am not sure if the reports of eight transgender people committing suicide due to the election results are true or false. But if they’re true, it seems really relevant that Trump denounced North Carolina’s anti-transgender bathroom law, and proudly proclaimed he would let Caitlyn Jenner use whatever bathroom she wanted in Trump Tower, making him by far the most pro-transgender Republican president in history.

I notice news articles like Vox: Donald Trump’s Win Tells People Of Color They Aren’t Welcome In America. Or Salon’s If Trump Wins, Say Goodbye To Your Black Friends. MSN: Women Fear For Their Lives After Trump Victory.

Vox writes about the five-year-old child who asks “Is Donald Trump a bad person? Because I heard that if he becomes president, all the black and brown people have to leave and we’re going to become slaves.” The Star writes about a therapist called in for emergency counseling to help Muslim kids who think Trump is going to kill them. I have patients who are afraid to leave their homes.

Listen. Trump is going to be approximately as racist as every other American president. Maybe I’m wrong and he’ll be a bit more. Maybe he’ll surprise us and be a bit less. But most likely he’ll be about as racist as Ronald Reagan, who employed Holocaust denier Pat Buchanan as a senior advisor. Or about as racist as George Bush with his famous Willie Horton ad. Or about as racist as Bill “superpredator” Clinton, who took a photo op in front of a group of chained black men in the birthplace of the KKK. Or about as racist as Bush “doesn’t care about black people!” 43. He’ll have some scandals, people who want to see them as racist will see them as racist, people who don’t will dismiss them as meaningless, and nobody will end up in death camps. He probably won’t do a great job fighting to end voter suppression, or helping people caught in the criminal justice system, but I’m not sure that makes him too different from the average member of the Republican Congress we have already.

Since everyone has been wrong about everything lately, I’ve started thinking it’s more important than ever to make clear predictions and grade myself on them, so here are my predictions for the Trump administration:

1. Total hate crimes incidents as measured here will be not more than 125% of their 2015 value at any year during a Trump presidency, conditional on similar reporting methodology [confidence: 80%]  
2. Total minority population of US citizens will increase throughout Trump’s presidency [confidence: 99%]  
3. US Muslim population increases throughout Trump’s presidency [confidence: 95%]  
4. Trump cabinet will be at least 10% minority [confidence: 90%], at least 20% minority [confidence: 70%], at least 30% minority [30%]. Here I’m defining “minority” to include nonwhites, Latinos, and LGBT people, though not women. Note that by this definition America as a whole is about 35% minority and Congress is about 15% minority.  
5. Gay marriage will remain legal throughout a Trump presidency [confidence: 95%]  
6. Race relations as perceived by blacks, as measured by this Gallup poll, will do better under Trump than they did under Obama (ie the change in race relations 2017-2021 will be less negative/more positive than the change 2009-2016) [confidence: 70%].  
7. Neither Trump nor any of his officials (Cabinet, etc) will endorse the KKK, Stormfront, or explicit neo-Nazis publicly, refuse to back down, etc, and keep their job [confidence: 99%].  
8. No large demographic group (> 1 million people) get forced to sign up for a “registry” [confidence: 95%]  
9> No large demographic group gets sent to internment camps [confidence: 99%]  
10. Number of deportations during Trump’s four years will not be greater than Obama’s 8 [confidence: 90%]

If you disagree with me, come up with a bet and see if I’ll take it.

And if you don’t, stop.

Stop fearmongering. Somewhere in America, there are still like three or four people who believe the media, and those people are cowering in their houses waiting for the death squads.

Stop crying wolf. God forbid, one day we might have somebody who doesn’t give speeches about how diversity makes this country great and how he wants to fight for minorities, who doesn’t pose holding a rainbow flag and state that he proudly supports transgender people, who doesn’t outperform his party among minority voters, who wasn’t the leader of the Salute to Israel Parade, and who doesn’t offer minorities major cabinet positions. And we won’t be able to call that guy an “openly white supremacist Nazi homophobe”, because we already wasted all those terms this year.

Stop talking about dog whistles. The kabbalistic similarities between “dog-whistling” and “wolf-crying” are too obvious to ignore.

Stop writing articles breathlessly following everything the KKK says. Stop writing several times more articles about the KKK than there are actual Klansmen. Remember that thing where Trump started out as a random joke, and then the media covered him way more than any other candidate because he was so outrageous, and gave him what was essentially free advertising, and then he became President-elect of the United States? Is the lesson you learned from this experience that you need 24-7 coverage of the Ku Klux Klan?

Stop using the words “white nationalist” to describe Trump. When you describe someone as a white nationalist, and then they win, people start thinking white nationalism won. People like winners. This was entirely an own-goal and the perception that white nationalism is now the winning team has 1% to do with Trump and 99% to do with his critics.

Stop responding to everyone who worries about Wall Street or globalism or the elite with “I THINK YOU MEAN JEWS. BECAUSE JEWS ARE THE ELITES. ALL ELITES AND GLOBALISTS ARE JEWS. IF YOU’RE WORRIED ABOUT THE ELITE, IT’S DEFINITELY JEWS YOU SHOULD BE WORRIED ABOUT. IF YOU FEEL SCREWED BY WALL STREET, THEN THE PEOPLE WHO SCREWED YOU WERE THE JEWS. IT’S THE JEWS WHO ARE DOING ALL THIS, MAKE SURE TO REMEMBER THAT. DEFINITELY TRANSLATE YOUR HATRED TOWARDS A VAGUE ESTABLISHMENT INTO HATRED OF JEWS, BECAUSE THEY’RE TOTALLY THE ONES YOU’RE THINKING OF.” This means you, Vox. Someday those three or four people who still believe the media are going to read this stuff and immediately join the Nazi Party, and nobody will be able to blame them.

Stop saying that being against crime is a dog whistle for racism. Have you ever met a crime victim? They don’t like crime. I work with people from a poor area, and a lot of them have been raped, or permanently disabled, or had people close to them murdered. You know what these people have in common? They don’t like crime When you say “the only reason someone could talk about law and order is that they secretly hate black people, because, y’know, all criminals are black”, not only are you an idiot, you’re a racist. Also, I judge you for not having read the polls saying that nonwhites are way more concerned about crime than white people are.

Stop turning everything into identity politics. The only thing the media has been able to do for the last five years is shout “IDENTITY POLITICS IDENTITY POLITICS IDENTITY POLITICS IDENTITY POLITICS IDENTITY POLITICS!” at everything, and then when the right wing finally says “Um, i…den-tity….poli-tics?” you freak out and figure that the only way they could have possibly learned that phrase is from the KKK.

Stop calling Trump voters racist. A metaphor: we have freedom of speech not because all speech is good, but because the temptation to ban speech is so great that, unless given a blanket prohibition, it would slide into universal censorship of any unpopular opinion. Likewise, I would recommend you stop calling Trump voters racist – not because none of them are, but because as soon as you give yourself that opportunity, it’s a slippery slope down to “anyone who disagrees with me on anything does so entirely out of raw seething hatred, and my entire outgroup is secret members of the KKK and so I am justified in considering them worthless human trash”. I’m not saying you’re teetering on the edge of that slope. I’m saying you’re way at the bottom, covered by dozens of feet of fallen rocks and snow. Also, I hear that accusing people of racism constantly for no reason is the best way to get them to vote for your candidate next time around. Assuming there is a next time.

Stop centering criticism of Donald Trump around this sort of stuff, and switch to literally anything else. Here is an incompetent thin-skinned ignorant boorish fraudulent omnihypocritical demagogue with no idea how to run a country, whose philosophy of governance basically boils down to “I’m going to win and not lose, details to be filled in later”, and all you can do is repeat, again and again, how he seems popular among weird Internet teenagers who post frog memes. In the middle of an emotionally incontinent reality TV show host getting his hand on the nuclear button, your chief complaint is that in the middle of a few dozen denunciations of the KKK, he once delayed denouncing the KKK for an entire 24 hours before going back to denouncing it again. When a guy who says outright that he won’t respect elections unless he wins them does, somehow, win an election, the headlines are how he once said he didn’t like globalists which means he must be anti-Semitic.

Stop making people suicidal. Stop telling people they’re going to be killed. Stop terrifying children. Stop giving racism free advertising. Stop trying to convince Americans that all the other Americans hate them. Stop. Stop. Stop.

# The Alzheimer Photo

A professor recently brought my attention to this photo of Alois Alzheimer and his colleagues in Munich (source):



Alzheimer is the very-German-looking guy with the silly mustache third from the right on the top. Far right is Friedrich Lewy, discoverer of Lewy bodies and Lewy body dementia. Bottom, second from the left, looking kind of like Petyr Baelish, is Ugo Cerletti, inventor of electroconvulsive therapy.

Other members of Alzheimer’s team didn’t make it to the group photo. These include Alzheimer’s mentor, Emil Kraepelin, who discovered bipolar disorder, schizophrenia, etc, etc, etc (there’s a reason modern psychiatry calls itself “neo-Kraepelinian”). They include two of Alzheimer’s assistants, Hans Creutzfeldt and Alfons Jakob, who discovered Creutzfeldt-Jakob disease, the human version of mad cow. They include Alzheimer’s collaborator Franz Nissl, who discovered Nissl bodies and the Nissl stain at the same lab.

If you come across a neurological disease that sounds like a guy’s name, there’s a not insignificant chance that guy is either in this picture or else just barely missed it.

This made me think of a lot of the discussion around when fields of science prosper versus when they go stagnant. The last few decades haven’t really been great for neuropsychiatry. But one group of people in one lab came up with entire textbooks worth of advances. Why? Do we need to resurrect Alois Alzheimer and put him in charge of NIMH?

Part of it was that good histological staining had just been invented and Alzheimer’s lab was on the bleeding edge, so they were just sitting around picking off the low-hanging fruit that could be discovered by staining stuff. But Kraepelin’s and Cervetti’s discoveries didn’t have much to do with staining.

Part of it was that Alzheimer was in the right place at the right time. If he’d really wanted an impressive photo, he could have gotten together with his chief competitors, a group centered around Carl Westphal (cf. Westphal’s sign, Edinger-Westphal nucleus) which included his students Arnold Pick (cf. Pick’s dementia) and Karl Wernicke (cf. Wernicke’s area, Wernicke-Korsakoff syndrome). Heck, if he wanted to go further, the number of people within a day’s train journey staggers the imagination. Rudolph Virchow, Eugen Bleuler, Robert Koch, Sigmund Freud. Fin de siecle Central Europe was just a really good place for neurology and psychiatry.

Part of it was that the whole thing was arranged by Kraepelin, who besides being a scientific genius, was apparently an organizational genius as well. According to Wikipedia, “Kraepelin has been described as a ‘scientific manager’ and political operator, who developed a large-scale, clinically oriented, epidemiological research programme.” See also Psychiatric Governance And The German Institute Of Psychiatry In Munich. Kraepelin grabbed all these people, threw them at the most interesting problems, and made sure they always had all the funding they wanted – although the final form of all of this as the Institute for Psychiatric Research didn’t coalesce until after Alzheimer’s death.

And part of it is the natural tendency for some institution to gain a reputation for being the best, and then attract the best people. I’m sure you could find some pretty impressive conjunctions of people if you looked at photos of Harvard departments.

My theory of apparent scientific stagnation has always been that it’s easier to pick low-hanging fruit in one paradigm than to get entirely new ones – in other words, the problem is at least as much in the territory itself as in our engagement with it. I was interested to learn that one of the big hurdles to faster aircraft is a nonlinearity in fuel costs, which grow exponentially for physics reasons right when you start getting faster than modern planes. I think something similar might be going on here. Through painstaking trial-and-error, psychiatric hit upon a really fruitful paradigm of combining clinical observation, histopathology, and and random wacky ideas, right about when Alois Alzheimer opened his lab. Anybody who happened to be in the vicinity when the new paradigm was invented ended up getting a disease named after him. Eventually all the stuff that was easy to discover this way got discovered, and right now there just aren’t any equally fruitful paradigms coming to our attention.

This story has a sad ending. Alzheimer (ironically) died young. He was succeeded by his student Walther Spielmeyer (cf. Spielmeyer-Vogt-Sjögren-Batten Disease), and then Kurt Schneider (cf. Schneider’s first-rank symptoms). Schneider invented the modern concept of psychopathy, but unfortunately he was probably working from personal experience – this was in the middle of the rise of the Nazis. He was fired for political reasons and got replaced with Alzheimer’s fellow Kraepelin protege, Ernst Rudin, who re-centered the whole thing around the role of psychiatry in sterilizing the feeble-minded. The chain that started with Kraepelin and Alzheimer ended in Rudin’s own student, Josef Mengele.

After the war, Rudin was fined 500 deutschmarks, apparently the going penalty for leading a Nazi eugenics program at the time, and Kraepelin/Alzheimer’s institute was re-founded as the Max Planck Institute of Psychiatry. As far as I know they’re still around, but I haven’t heard of them discovering any interesting new diseases lately.

# SSC Journal Club: Expert Prediction Of Experiments

I.

It’s been a good month for fretting over failures of expert opinion, so let’s look at DellaVigna & Pope, Predicting Experimental Results: Who Knows What?

The authors ran a pretty standard behavioral economics experiment where they asked people on Mechanical Turk to do a boring task while being graded on speed and accuracy. Then they offered one of fifteen different incentive schemes, like “we’ll pay you extra if you do well” or “your score will be publicly visible”.

But the point of the study wasn’t to determine which incentive scheme worked the best, it would determine who could best predict which incentive scheme worked the best. The researchers surveyed a bunch of people – economics professors, psychology professors, PhD students, undergrads, business students, and random Internet users on Mechanical Turk – and asked them to predict the experimental results. Since this was a pretty standard sort of behavioral economics experiment, they were wondering whether people with expertise and knowledge in the field might be better than randos at figuring out which schemes would work.

They found that knowledgeable academics had some advantage over randos, but with enough caveats that it’s worth going over in more detail.

First, they found that prestigious academics did no better (and possibly slightly worse) than less prestigious academics. Full professors did no better than associate professors, assistant professors, or PhD students. People with many publications and citations did no better than people with fewer publications and citations.

Second, they found that field didn’t matter. Behavioral economists did as well as microeconomists did as well as experimental psychologists did as well as theoretical psychologists. To be fair, this experiment was kind of in the intersection of economics and psychology, so all of these fields had equal claim to it. I would have liked to see some geologists or political scientists involved, but they weren’t.

Third, the expert advantage was present in one measure of accuracy (absolute forecast error), but not in another (rank-order correlation). On this second measure, experts and randos did about equally well. In other words, experts were better at guessing the exact number for each condition, but not any better at guessing which conditions would do better or worse relative to one another.

Fourth, the expert advantage was pretty small. Professors got an average error of 169, PhD students of 171, undergrads of 187, MBA students of 198, and MTurk users of 271 (random guessing gave an error of about 416). So the difference between undergrads and experts, although statistically significant, was hardly overwhelming.

Fifth, even the slightest use of “wisdom of crowds” was enough to overwhelm the expert advantage. A group of five undergrads averaged together had average error 115, again compared to individual experts’ error of 169! Five undergrads averaged together (115) did about as well as five experts averaged together (114). Twenty undergrads averaged together (95) did about as well as twenty experts averaged together (99).

Sixth, having even a little knowledge of individuals’ forecasting ability screened off expert status. The researchers gave forecasters some experimental data about the effects of a one-cent incentive and a ten-cent incentive, and asked them to predict the scores after a four-cent incentive – a simple, mechanical problem that just requires common sense. Randos who can do well on this problem do just as well as experts on the experiment as a whole. Likewise, randos who are noticed to do well on the first half of the experiment will do just as well as experts on the second half too. In other words, we’re back to finding “superforecasters”, people who are just consistently good at this kind of thing.

None of this seems to be too confounded by effort. The researchers are able to measure how much time people take on the task, whether they read the instructions carefully, etc. There is some advantage to not rushing through the task, but after that it doesn’t seem to matter much. They also try offering some of the Mechanical Turkers lots of money for getting the answers right. That doesn’t seem to help much either.

The researchers ask the experts to predict the results of this experiment. They (incorrectly) predict that prestigious academics with full professorships and lots of citations will do better than mere PhD students. They (incorrectly) predict that psychologists will do better than non-psychologists. They (correctly) predict that professors and PhD students will do better than undergrads and randos.

II.

What do we make of this?

I would tentatively suggest it doesn’t look like experts’ expertise is helping them very much here. Part of this is that experts in three different fields did about equally well in predicting the experimental results. But this is only weak evidence; it could be that the necessary expertise is shared among those three fields, or that each field contains one helpful insight and someone who knew all three fields would do better than any of the single-field experts.

But more important, randos who are able to answer a very simple question, or who do well on other similar problems, do just as well as the experts. This suggests it’s possible to get expert-level performance just by being clever, without any particular expertise.

So is it just IQ? This is a tempting explanation. The US average IQ is 100. The undergrads in this experiment came from Berkeley, and Berkeley undergrads have an average SAT of 1375 = average IQ of 133 (this seems really high, but apparently matches estimates from The Bell Curve and the Brain Size blog; however, see Vaniver’s point here). That same Brain Size post proposes that the average professor has an IQ of 133, but I would expect psychology/economics professors to be higher, plus most of the people in this experiment were from really good schools. If we assume professors are 135-140, then this would neatly predict the differences seen from MTurkers to undergrads to professors.

But the MBA students really don’t fit into this model. The experiment gets them from the University of Chicago Booth School of Business, which is the top business school in the country and has an average GMAT score of 740. That corresponds to an IQ of almost 150, meaning this should be the highest-IQ sample in the study, yet the MBAs do worse than the undergrads. Unless I’m missing something, this is fatal to an IQ-based explanation.

I think that, as in Superforecasting, the best explanation is a separate “rationality” skill which is somewhat predicted by high IQ and scientific training, but not identical to either of them. Although some scientific fields can help you learn the basics of thinking clearly, it doesn’t matter what field you’re in or whether you’re in any field at all as long as you get there somehow.

I’m still confused by the MBA students, and expect to remain so. All MBA students were undergraduates once upon a time. Most of them probably took at least one economics class, which was where the researchers found and recruited their own undergraduates from. And most of them were probably top students from top institutions, given that they made it into the best business school in the US. So how come Berkeley undergraduates taking an econ class outperform people who used to be Berkeley undergraduates taking an econ class, but are now older and wiser and probably a little more selected? It might be that business school selects against the rationality skill, or it might be that business students learn some kind of anti-insight that systematically misleads them in these kinds of problems.

(note that the MBAs don’t put in less effort than the other groups; if anything, the reverse pattern is found).

III.

Does this relate to interesting real-world issues like people’s trouble predicting this election?

One important caveat: this is all atheoretical. As far as I know, there’s no theory of psychology or economics that should let people predict how the incentive experiment would go. So it’s asking experts to use their intuition, supposedly primed by their expertise, to predict something they have no direct knowledge about. If the experiment were, say, physicists being asked to predict the speed of a falling object, or biologists being asked to predict how quickly a gene with a selective advantage would reach fixation, then we’d be in a very different position.

Another important caveat: predictive tasks are different than interpretative tasks. Ability to predict how an experiment will go without having any data differs from ability to crunch data in a complicated field and conclude that eg saturated fat causes/doesn’t cause heart attacks. I worry that a study like this might be used to discredit eg nutritional experts, and to argue that they might not be any better at nutrition than smart laymen. Whether or not this is true, the study doesn’t support it.

So one way of looking at it might be that this is a critique not of expertise, but of “punditry”. Engineers are still great at building bridges, doctors are still great at curing cancer, physicists are still great at knowing physics – but if you ask someone to predict something vaguely related to their field that they haven’t specifically developed and tested a theory to cope with, they won’t perform too far above bright undergrads. I think this is an important distinction.

But let’s also not get too complacent. The experts in this study clearly thought they would do better than PhD students. They thought that their professorships and studies and citations would help them. They were wrong. The distinction between punditry and expertise is pretty fuzzy. Had this study come out differently, I could have argued for placing nice clear lab experiments about incentive schemes in the “theory-based and amenable to expertise” category. You can spin a lot of things either direction.

I guess really the only conclusion you can draw from all of this is not to put any important decisions in the hands of people from top business schools.

# Contra Robinson On Schooling

I.

Nathan Robinson argues against school vouchers: Why Is The Decimation Of Public Schools A Bad Thing?

(note that despite the inflammatory title, he’s arguing that the decimation of public schools is, indeed, a bad thing)

He starts with a meta-level point: most criticisms of Trump’s Education Secretary nominee Betsy DeVos merely point out that she will promote schools vouchers instead of public schools, expecting their audience to be suitably turned off. But this won’t change the minds of DeVos supporters, whose whole point is that they want more school vouchers. In order to convince people, you’ve got to convince people. If this doesn’t seem like a suitably profound insight to you, the click the link, read the piece, and notice how there’s something weird about it. Is it written in a funny font? Is the computer screen flickering or something? Finally, you realize with dawning horror that this is the first time you’ve read a logical argument, written in good faith and intended to convince someone, in the past you-can’t-remember-how-many months.

But anyway, let me explain why I think it’s wrong.

Robinson says:

Introducing profit into the school system is very dangerous, for a simple reason: it creates a terrible set of incentives. If we hand a voucher to a for-profit private school, or give a large grant to a for-profit charter school, there is a strong incentive for the school to give as little in return as possible. After all, since a for-profit corporation exists to maximize value to shareholders (not value to students), for-profit schools should try to spend as little money educating students as possible, in order to reap the largest financial gains. If you don’t have to spring for new lab equipment or new textbooks, you have no incentive to do so merely because it would benefit the students.

He does note the obvious counterargument, but he’s not convinced:

Privatization advocates have a compelling response to this argument. They reply that it misses the full picture. Yes, corporations have an incentive to maximize shareholder value. But they can’t do that without satisfying their customers. The interests of shareholders and consumers are brought into alignment through the existence of choice. In the case of schools, because parents have a voucher, if the school is not prioritizing its students, parents can simply go elsewhere. Nobody is making them send their students to this particular school. The theory of school choice is about choice, and choice creates competition, which creates quality. A school that simply funneled money to its executives and shareholders would not long maintain its enrollment.

But the theory of choice here is a romantic fiction. In reality, parents will not have many options among which to choose (there are only so many schools within a feasible distance of one’s home, after all) and moving schools can be an extraordinarily disruptive and complicated process that hurts the child. We can also see how, even in theory, it is easy for a privatized school system to simply enrich the wealthy, while making schooling for poor children worse. In a public school system, all money is spent on the schools. In a for-profit school system, at least some portion of that money is directed instead toward the pockets of shareholders (if it wasn’t, the for-profit schools couldn’t continue to exist). And if we have a school district comprised in total of three for-profit elementary schools, and all of them simply pocket most of the voucher money while failing to educate the children, then no matter what “choices” among schools parents make, they won’t be able to improve the quality of the schools. One might expect new operators to enter the market, but if the only way to make any real money on the children is to neglect them, then new operators won’t be any better than the old ones.

This is a good point, made somewhat weaker by failure to consider why it doesn’t apply to everything else.

Robinson compares school vouchers to foodstamps, which are basically “food vouchers” to be redeemed at grocery stores. I often see poor people using food stamps at my own grocery store, so I know the quality of service these poor people get for their money. And it is really good. Practically all grocery stores are really good. There’s a story about Boris Yeltsin coming to America for the first time, walking into a random grocery store, seeing that random middle-class Americans had a better selection of goods than the highest-status Soviet officials, and freaking out that this was some kind of weird Potemkin economy that the Americans had set up to demoralize him. Grocery stores don’t just have fifty different kinds of cereal and a hundred different kinds of soda, they’re also really cheap. You can buy a day’s worth of food for an hour’s minimum-wage work, maybe two hours if you want a little quality and variety.

So why don’t grocery store shareholders leech off so much money that everything is overpriced and has terrible service? Why aren’t stores dingy and full of rats? Why don’t we have a world where, as Robinson argues, the theory of choice is a romantic fiction because all of the grocery-related options available to poor people are terrible, and no new operators can do a better job because the only way to make money in the grocery business is to shaft customers and have a terrible store selling rice mixed with sawdust?

Something like 48% of Americans are satisfied with the education system in the US. My guess is 100% of Americans are satisfied with the grocery system in the US. Why should this be?

My guess: the loss from profits matters less than the gain in efficiency.

Profit margins are a specific number. You can just look them up. Usually they’re not very big. Once you look up that number, you know how much profit the company is making. After that, you’ve circumscribed “the dangers of profit” to a relatively small amount.

An example: Health care in this country is overpriced and everyone knows it. Some people think this is because greedy insurance companies are charging too much in order to make a profit.

But health insurance companies have a profit margin of about 3% (see caveats here, but I do think the 3% number is the one relevant to this discussion). This is a big deal in terms of absolute number of dollars. But it’s not a big deal if you’re wondering how much they affect health care costs. If you’re paying $5,000 a year for health insurance, then take away all profit motive from the insurance companies and you would pay $4850 a year for health insurance. This is less than year-to-year variation, let alone any of the components that actually matter.

Insurance companies aren’t callously throwing sick people out on the street for profitability reasons, they’re callously throwing sick people out on the street because they can only pay as much money as their customers give them and that isn’t enough to fund as much health care as people need. Or, well, maybe out of the people they callously throw out on the street, 3% are for profit and the other 97% are of necessity.

The same is true of other famously predatory businesses like payday lenders. Wikipedia notes:

In a profitability analysis by Fordham Journal of Corporate & Financial Law, it was determined that the average profit margin from seven publicly traded payday lending companies (including pawn shops) in the U.S. was 7.63%, and for pure payday lenders it was 3.57%. These averages are less than those of other traditional lending institutions such as credit unions and banks. Comparatively the profit margin of Starbucks for the measured time period was just over 9%, and comparison lenders had an average profit margin of 13.04%. These comparison lenders were mainstream companies: Capital One, GE Capital, HSBC, Money Tree, and American Express Credit.

Payday lenders aren’t charging outrageous interest rates so they can get fat off the profits. They’re charging outrageous interest rates because loaning money to poor people who often fail to pay back their loans is a hard business to break even on.

Now apply this logic to private schools. You think they’re going to ruin everything by funneling more and more money into their own profitability? I propose that about 11% of their funding would go to profit, the same as current private schools. Actually less than that, since most current private schools serve rich people who don’t care about outrageous markups as long as they can buy prestige.

But isn’t 11% still more than nothing? If private schools cost the same amount of money as public schools, but 11% of that went to shareholder profits, wouldn’t our children receive an 11% worse education?

Or to put it another way – what are we buying with that 11% of the education budget?

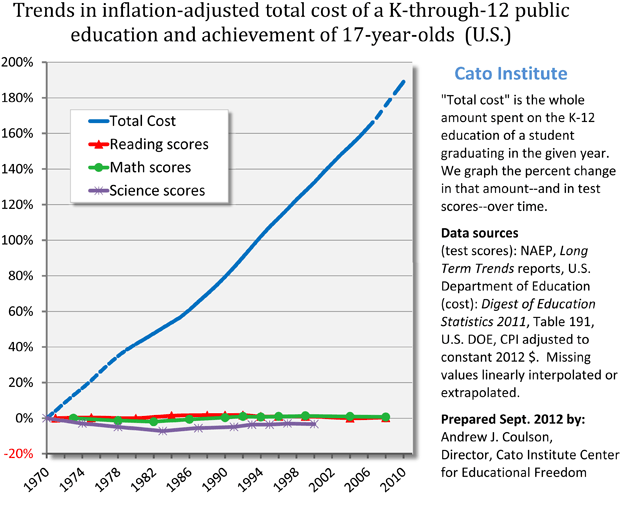
The hope would be that we’re buying efficiency.

Robinson writes:

Let’s consider what the conservative argument on schooling actually is. It goes like this: government-run institutions tend to function poorly. They are not efficient, like businesses are, because they do not have incentives to perform well. Businesses, because they must compete for customers in a market environment, must offer the best products if they want to stay profitable. Governments, on the other hand, can offer crappy products, and because they are state-imposed monopolies, there is no way for consumers to go elsewhere. School choice will improve schools, because instead of forcing students to attend whatever school the government happens to offer, choice allows parents to decide which school they prefer. Schools will have to strive to be better and better, because parents can pull their students out and go elsewhere if they don’t like them. Introducing a profit motive into schooling offers a powerful incentive for schools to offer a great product. If there is money to be made on being a good school, you can bet businesses will want to provide great schools. Thus private, for-profit schools with vouchers are a highly efficient way of delivering the best-quality education.

He counters that “introducing profit into the school system is very dangerous”, but never really says that the argument above is wrong. So we’re looking at a tradeoff here. There’s the dangers of profit and the promises of better efficiency. If profits aren’t going to take too much money out of the system, might the gain from efficiency be worth the small cost?

Here is a graph by the Cato Institute – note that this is already adjusted for inflation:



In case you don’t trust them, here’s Politifact rating a similar claim mostly true, although they get slightly different numbers using a different methodology.

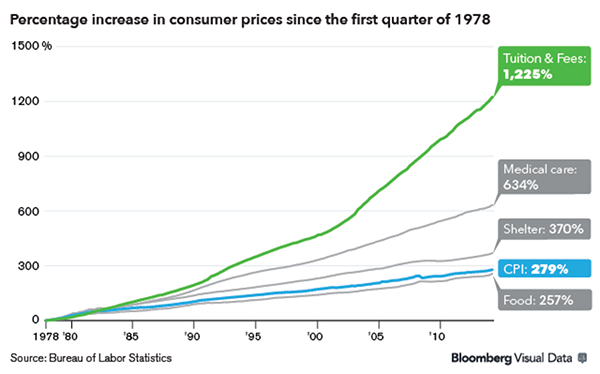
What are we to make of this? It’s not that teachers are getting paid any more – their salaries have remained stagnant over the time involved and they may even have lost ground compared to other professions. It’s not that school buildings cost more – I don’t have good data on schools in particular, but I looked into skyscrapers and found there wasn’t any general rise in construction prices. So how is an activity which basically involves getting a bunch of kids into a building and throwing a teacher at them rising so dramatically in the absence of changes in building or teacher prices? I’ve only heard three theories:

The first theory is that student test scores are improving, but we’re stuck in a Simpson’s Paradox. That is, students of some racial groups get higher scores than others, every racial group’s score is improving, but a higher percent of students are in low-performing racial groups now which is bringing the average down. But it looks like race-specific scores are also stagnant (1, 2), with the possible exception of a jump for blacks between 1980 and 1990. Given that no other ethnic group had this jump and school spending has increased at a constant rate the whole time, I think this is more likely socioeconomic factors than education spending.

The second theory is that all this extra money has been used to help previously underserved special needs kids. But these kids apparently cost about twice as much as average to educate. Up to 13% of students are special needs, although I don’t know if that’s the same definition of “special needs” as the people calculating the cost used. But if we take that as our estimate, then providing extra services to special needs kids can explain a rise of 13% in education costs, but not the 150% we actually see.

The third theory is Wilde’s Law: “The bureaucracy is expanding to meet the needs of the expanding bureaucracy.”

One argument in favor: the same thing that’s happening to primary education is happening to college education. College costs about 4x as much (inflation-adjusted) as it did in 1980.



This is not because of decreased government subsidization. It’s not going to shareholders – most colleges are nonprofits and even public institutions have seen outsized increases. And although I don’t have any equivalent to the flat-lining test scores from primary school, today’s college students don’t seem to be four times better-educated than those of yore. So where did all that money go?

I don’t think anybody knows. There are whole studies that have been done on this. Every so often people argue about it on the editorial pages. Some people say “administration”, other people say it’s not exactly administration but it’s something else. I don’t know. But it doesn’t seem to be Simpson’s Paradox or special needs kids.

A lot of smart people think that easy availability of student loans fueled college cost increases – see eg here. The theory was that colleges could charge more money, so they did. I definitely don’t understand how this works economically, but it seems like somehow easy availability of money combined with lack of real competitive pressure caused colleges to increase administration and pass the cost on to students. Maybe something similar happened in primary schools?

If this third hypothesis of increased primary school costs is true, then going back to the level of bureaucracy we had in 1970 would cut costs by 75% while maintaining similar test scores. Or it would allow us to keep costs constant and pay for things that actually work.

The point is, private schools lose 11% of their funding to shareholder profit, and public schools apparently lose 75% of their funding to, uh, nobody really knows.

Do private schools also lose 75% of their funding to nobody-knows, since I don’t see many of them around as cheap as schools were in 1980? I’m not sure. Arnold Kling does some calculations here and suggests that private schools should have really big profit margins. But they apparently don’t. Overall I admit I am confused on this issue. But I am a little more hopeful that private schools might be able to work this issue out than that public schools can. [EDIT: see note at bottom of post]

It may seems kind of bloodless to focus on cost. But the amount of money added to the public school budget without any change in outcomes is more than enough to house every homeless person in the country in style and comfort. Money matters. And when we talk about private schools being obsessed with profit, that’s an argument about money. It’s acknowledging that every dollar diverted to shareholders is a dollar that isn’t going to improving education, housing the homeless, or something else useful.

Robinson writes:

It’s because the things needed by poor people, if done well, will never be money-makers. Introducing an incentive to make money will necessarily mean exploiting and neglecting the poor, whose “choices” are highly constrained by their circumstances. I fear privatization not because of some mystical devotion to the inefficiencies of government but because I fear the erosion of the idea of education as something that isn’t win-win, that we give to children because they deserve it rather than because we can profit from it. I worry that the sort of people who run things “like a business” do not really care about children very much, and are motivated by the wrong incentives. I am concerned about what would happen if they ever faced a choice between doing the right thing and doing the lucrative thing.

With all due respect, I think there is something mystical in this thought process, some demon best exorcised with a bell, candle, and Public Choice Theory textbook. There’s no object called The System, which is focused on profit in businesses and focused on education in public services. There’s just a bunch of people motivated by a combination of ethics, incentives, and trying not to get fired. Business isn’t antithetical to caring – the average family doctor is motivated by desire to help patients, even though she’s also a small business. And lack of a profit motive doesn’t guarantee good behavior – it looks like the administrators of nonprofit colleges decided to spend their windfall on prestige and empire-building rather than on keeping costs low.

I have very low confidence in this. I know many people who are involved in education, and they are all very good people who are very passionate and definitely would never skim 75% off the top and use it to buy gold-plated yachts for themselves. In my home state of California, there was a big funding shortfall ten years ago, and schools tried to cut everything they could, but finally they said there was nothing left to cut and they had no more ideas, and I believe that they tried as hard as they could. If bureaucracy is inflating the price of schooling, it’s not doing so in an obvious way where you can point a finger at the exact bureau involved.

But it might be a general ethos of inefficiency that makes a lot of little things add up – I know it is in health care. And I at least think it would be worth trying the experiment.

(I realize the experiment is already being tried, with wildly varying results based on the specifics. I want to look into this research in more depth soon to see if there are any consistent trends.)

II.

A digression to support my point that sometimes increased efficiency can compensate for money diverted into profits – what about hospitals?

Hospitals are about evenly split between for-profit and not-for-profit institutions. Measuring “hospital quality” is even harder than measuring school quality, but researchers have tried to do this on various metrics. The results are hard to sum up, and I was only able to find a few studies and not anyone’s magisterial summation of the field, but it looks like there are minimal differences between for-profit and non-profit private hospitals, with government hospitals doing worst of all:

— A team from Harvard finds that for-profits and non-profits have about equal quality, and government-owned hospitals are worse than either. A follow-up study by the same team finds non-profit hospitals becoming for-profit is not associated with a drop in care.

— Truven Health Analytics finds some advantages for church-owned nonprofit hospitals, with secular nonprofit hospitals and for-profit hospitals in the middle, and government-owned hospitals worst of all. Note that this is my interpretation of a lot of different data and you might want to look at the particular metrics they use to draw your own conclusions.

— A textbook on the hospital industry finds that “on average, the performance of non-profit hositals in treating elderly patients with heart disease appears to be slightly better than that of for-profit hospitals, even after accounting for systematic differences…however, this small average difference masks an enormous amount of variation in hospital quality within the for-profit and not-for-profit hospital groups.”

— A study in a cardiology journal found “no evidence that for-profit hospitals selectively treat less sick patients, provide less evidence-based care, limit in-hospital stays, or have patients with worse acute outcomes than nonprofit centers”.

— As per the Handbook Of Health Economics:

The most rigorous and extensive study of large-scale empirical study of quality published to date that permits comparisons of quality by hospital ownership is by Keeler and co-authors (1992). They used two process measures of quality based on reviews of 14,000 medical record for five diseases in five states. One of this “explicit process” gauged the extent to which the charts showed that specific diagnostic and therapeutic procedures were performed competently. Rather than focus on particular elements of care as explicit process did, a second process measure, “implicit process,” measured the care process overall. For example, one of their implicit process questions to physician reviewers was: “Based on what you now know about this case, would you send your mother to this hospital?” In addition, they gauged quality on an outcome measure — the difference between actual mortality and the rate that would be expected based on the patient’s characteristics.

They found no difference in quality between private not-for-profit and for-profit hospitals on two indicators, excess mortality and explicit process; public hospitals fared worse on both criteria. However, on a third measure, implicit process, there was a statistically significant difference between quality of care of private not-for-profit hospitals and the other two ownership types, indicating higher quality levels for the for profits. The authors appear to have been more persuaded by the results on the first two indicators, stating that “nonprofit and for-profit hospitals provide similar quality overall”

In their national study of 981 hospitals in 1983-84, Shortell and Hughes (1988) found no difference in quality measured in terms of mortality by ownership. However, using fewer covariates, Hartz et al. (1989) did find that mortality was higher in for-profit than in private not-for-profit hospitals.

Sloan and co-authors (1998a,1998b, 1998c) examined outcomes of care of elderly persons hospitalized for one of four conditions: hip fracture; stroke; coronary heart disease; and congestive heart failure. They analyzed the first admission for these conditions since patients with a first unanticipated major health shock are less likely to shop among hospitals. Their outcome measures were survival, functional status, cognitive status, and living arrangements (probability of living in a nursing home). Although, on some measures, patients admitted to major teaching hospitals did better, a result consistent with Keeler et al. (1992), there were no statistically significant differences in outcomes between non-teaching private not-for-profit and for-profit hospitals. On some measures, elderly patients admitted to non-teaching government hospitals had worse outcomes, holding a large number of other factors constant.

So it looks like the differences between for-profit and not-for-profit hospitals are pretty slight, and that government hospitals are likely worse than either. This should be equally confusing to people who believe that the profit motive is invariably destructive and to people who believe it invariably results in better service.

This would be a good time to note that, contrary to the impression one might get from the Current Affairs article, only 13% of charter schools are for-profit. So a system with vouchers and charter school would probably be most like private non-profit hospitals, which these studies also show as doing well, and as decisively better than government-run hospitals. Also, everything I’ve talked about up to this point is mostly irrelevant. Maybe I should have started by mentioning this.

III.

From the same article:

There are other serious problems with the “gutting” of public schools. As we have argued before in Current Affairs, converting public schools to a voucher system makes education operate similarly to food stamps. After all, SNAP benefits operate roughly the same way: instead of giving people food, we give them the equivalent of money, which they then use to go and buy food. A voucher program does the same for schooling: instead of giving them schools, we give them a voucher, which they can use to go and find a private school. But look what happens with food stamps: the moment you start handing out a “voucher,” conservatives start seeing it as some kind of unearned “handout.” Pressure then develops to cut the handout. Is there any reason to think that “education stamps” would be subjected to less cost-cutting political pressure than food stamps? A serious problem with voucher programs is that they erode the idea of education as a fundamental right, instead making it seem like a privilege that one does not necessarily deserve. But education should be a right, because children cannot help the circumstances of their birth, and should therefore not be punished for their parents’ poverty.

I think this is drawing the wrong lesson from education’s popularity relative to food stamps. Robinson thinks conservatives like one-size-fits-all handouts, but not voucher handouts. I think conservatives like universal handouts, but not handouts to the poor.

Imagine a world where food stamps are replaced by the Federal Food Agency. Every week, a truck comes to poor people’s houses and gives them a one-size-fits-all food package that the government believes satisfies their nutritional needs. Do you think conservatives would be any happier with this than they are with food stamps? For that matter, did conservatives support public housing projects any more or less than they support housing vouchers now?

On the other hand, Medicare remains popular even though it’s essentially a voucher. Patients with Medicare choose their doctor, choose their hospital, and then Medicare pays for it. But because everyone expects to benefit someday, it has pretty broad bipartisan support; even its critics mostly want to change rather than eliminate it.

I don’t think changing public education from a service to a voucher would change whether people support it or not.

IV.

So I disagree with Robinson’s specific arguments. But there are some things that worry me about school vouchers.

First, the hospital case study is kind of ambiguous. Although for-profit hospitals aren’t noticeably worse than not-for-profit, they’re also not noticeably better. And the existence of for-profit hospitals hasn’t started some kind of virtuous cycle where all hospitals compete to save money and provide better care that ends up with hospitals being lean and inexpensive and just as accessible as grocery stores. Having a field be open to competition isn’t necessarily incompatible with it being overpriced and inefficient. And commenters point out that existing private schools are not generally 75% cheaper than public schools, suggesting that cost-cutting is hard.

Second, Robinson notes later that:

Privatization schemes are also heavily dependent on the existence of highly astute parents, who have the time and inclination to carefully study schools. The most vulnerable children are unlikely to have such parents. And we can imagine a system in which private schools offer parents $100 out of the voucher money if they agree to enroll their children. Desperate and uncaring parents might snap up the cash, with the neediest children ending up in the most vicious, uncaring, profit-grubbing schools.

I doubt there would be such blatant kickbacks – they’d be illegal and I don’t think they’ve happened on other voucher programs like food stamps – but his point that many parents are ignorant or malicious is well-taken. You don’t need literal bribery to get schools which are very good at having flashy ad campaigns but not very good at education. Parents might not check the test scores of a smooth-sounding school any more than they check the health care grades of their local hospital. The worst-case scenario is schools associated with cults or fringe political ideologies that prey on the children of people who believe them, either out of genuine fanaticism or a cynical calculation that fanatics are easy to milk.

Third, the whole point of Trumpism is that once we have fewer immigrants we can create a culturally cohesive community where everybody shares some core values. But the school system – as fractured and diverse as it is – is really one of the only institutions responsible for instilling some basic civic values in everyone and making sure they’re all on the same page. I do not put it past people to start sending their kids to schools that teach liberal values or conservative values in particular, and then one of the few (albeit mostly ineffective) brakes on further polarization is removed. On the other hand, Catholic school is already sort of like this and they don’t seem to be some weird foreign cancer on the body politic, so maybe it’s not such a big deal?

Fourth, vouchers could worsen class segregation. Maybe not too much, because everyone already goes to public schools in their own class-segregated neighborhood anyway. But at least there’s a little socioeconomic diversity now. And with vouchers, there’s a risk of deliberate sorting/signaling, where if everyone gets a voucher for $10,000, decent schools will charge $15,000 just to sell the “privilege” of going to a school without poor students. That is, in the same way people will pay extra for a house in a gated community because they worry poor people make bad neighbors, they might pay extra for a spot in a more-expensive school because they worry poor people will make bad classmates. A little bit of this segregation goes a long way, because if enough people do this then the exactly-$10,000 schools will only have poor people, in much the same way that a little bit of racial segregation goes a long way.

(I don’t know much about proposed voucher systems, but I wonder if it would be possible to have a system where you’re not allowed to combine the voucher with your own money. That is, if you get a $10,000 voucher, you can go to a school charging $10,000. But if you want to go to a school charging $11,000, you have to throw away the voucher and pay the whole price out of pocket.)

To all these downsides we would have to add one very big upside – it destroys the incentive to overspend on/segregate housing in order to get into a “good school district”. Elizabeth Warren has argued this is primarily behind the secular rise in real estate prices that has undermined the economic position of the middle class for the past fifty years. This factor could easily be more important than everything else combined and might make school vouchers a plus even if they seriously worsened the quality of education.

Overall my thoughts on school vouchers are the same as my thoughts on pretty much everything in this category: let’s experiment. Figure out a window of acceptable possibilities that are reversible and don’t have too much risk, and let different states and areas try different ones. As we start to understand things better, extend the window of possibilities in the relevant direction. Check results. Rinse. Repeat. Then figure something out.

V.

Finally, one more point from the article that deserves its own discussion:

If we have a school district comprised in total of three for-profit elementary schools, and all of them simply pocket most of the voucher money while failing to educate the children, then no matter what “choices” among schools parents make, they won’t be able to improve the quality of the schools. One might expect new operators to enter the market, but if the only way to make any real money on the children is to neglect them, then new operators won’t be any better than the old ones.

It’s important because lack of good competition is indeed the bane of all of these sorts of industries. Economic theory predicts that in a perfectly competitive environment businesses will be pretty good; it is much less sure of itself in these sorts of three-school districts without enough competition to have much effect.

In the real world, someone will have to empirically determine how much this matters. In my own fantasy world, I have a solution that the new Education Secretary probably won’t be on board with: Let’s let random people open tiny schools.

Something like 3% of parents home-school their children. This cuts across class and racial lines better than you’d think. All the research shows that home-schooled students do much better than traditionally schooled students on standardized tests, college admission exams, college GPAs, and general life satisfaction as adults. This is probably unfair, because home-schooled students are the descendents of the sort of thoughtful conscientious parents who want to home-school their children, so they probably have a big genetic advantage. But there is at least absolutely no evidence that home-schooling makes anyone do any worse.

The average cost per pupil per year in the US is something like $10,000. So suppose we give everyone $10,000 school vouchers. A parent who wants to make the median US yearly income of $30,000 would have to teach three students. Add in some overhead and curriculum costs, and maybe it’s more like five students.

So imagine. A woman has a kid and decides she doesn’t want to go back to work and leave the kid in daycare for eighteen years. She takes some test, clears some regulatory hurdle, promises that she’ll clear a certain bar on her kids’ standardized test scores, and registers as an approved school. Then she gets a couple of friends and neighbors who trust her to send their kids to her too. Maybe her husband works outside the home, so she doesn’t even need five. She’s happy with two or three (I think it would be important that you can’t make any money by educating your own kid; otherwise the incentive is to keep them out of school and pretend to be educating them yourself). Then she tutors them in a class a fifth the size of comparable public school classes.

If you’re an actual, qualified teacher, maybe you can get ten or twenty kids who are interested. That’s $100,000 to $200,000, minus your overhead, much more than qualified teachers make today with a much lower class size. Remember, for the majority of American history, kids were taught by a member of the community in a one room schoolhouse, and that was the system that produced Benjamin Franklin, Thomas Edison, et cetera.

(and remember that all the research shows that formal teacher training and level of teacher credentialing has zero effect on how well teachers teach kids)

This would provide provide a means of self-directed, boss-free income for millions of people, including undercredentialled poor people, disabled people who can’t leave the home, people in rural areas, and especially young mothers. It would rebuild community ties. And it would ensure no one ever has to worry about districts with only three schools, or even districts with only thirty schools.

I don’t know. Probably there’s some sort of horrible flaw that I’m missing. But I still think the moral of the story is to experiment more. And school vouchers might be a good start.

EDIT: From the comments: “At my independent non-profit high school we have reduced the cost per student-year down to $3,000. Despite this budget we are able to offer the students many opportunities that public schools don’t. Art classes like glassblowing and copper and silversmithing, advanced science curricula like organic and biochemistry, health class vastly more informative that the state requirements, zero bullying enforced by a self-organized student culture, I could go on all day. For comparison the average cost for nearby school districts is $17,000 per high schooler-year and offer a fraction of the services we provide.”

# Highlights From The Comment Thread On School Choice

Several people including Yehoshua K and Freddie deBoer point out that “nonprofit” and “for-profit” are potentially meaningless terms in situations like these.

IrishDude adds some context to the for-profit hospital scene by noting that companies are not allowed to open new hospitals until they apply for and are granted a Certificate Of Need, apparently on the basis of a theory that an oversupply of hospitals would increase (?!) costs.

Algorizmi describes his work with a private school that costs less than half as much as most public schools, including how it saves money

Half a dozen people yell at me for saying that the grocery industry worked well for the poor, objecting that I had forgotten about Food Deserts. Various other people save me some time by pointing out that most of the claims about Food Deserts are kind of fake (1, 2, 3). There’s a defensible version of the term, which is that in very poorly planned car-centered zoning-regulated cities without good public transportation it’s not always possible for someone without a car to easily get to the stores they want, but just posing the problem that way makes the solution pretty obvious.

Spotted Toad is great as usual, pointing out among other things that Obama has already led the most pro-charter administration in history.

Douglas Knight proposes that college demand curves are upward-sloping, which is kind of terrifying but seems to have at least anecdotal support.

Robinson: “The things needed by poor people, if done well, will never be money-makers.” Matt M: “A bold claim considering that the biggest company on Earth, more than double #2, serves the needs of poor people almost exclusively.”

MDP has worked in the payday loan industry and explains how their interest rates can be so high and their profits so low.

Many people (1, 2, 3) point out the high cost to schools of misbehaving kids, and try to explain the rise in education costs by saying that kids today are raised wrong (or not at all) which makes them harder to control. But everybody who looks for this kind of thing finds the opposite – kids today have less teenage pregnancy, crime, dropouts, et cetera. I don’t know if anyone has specifically looked at classroom misbehaving, but it would be weird for that to be getting worse in such an isolated way.

Lots of people are very angry at me for posting the graph from the Cato Institute for various reasons. A few people object that it is dishonest because it didn’t adjust for inflation, even though it did adjust for inflation and is very clear about that. A few people object that it is dishonest because it puts cost increases and score increases on the same scale as each other, instead of skewing them to make the effect look bigger than it really is – this is a definition of “dishonest” I haven’t heard before (maybe I am being uncharitable here?). Several others say that test scores have increased more than they give credit for. You can look at a very good summary of test score changes here. You’ll see there are some gains among younger students, but much fewer among older kids – for example, 17 year olds’ math scores went from 304 in 1971 all the way to 306 in 2012. This is not just a race-related Simpson’s Paradox – among white students alone, for example, the gain was 4 points (though blacks did gain 18, as I mentioned before). I am not sure what good it is to have high gains in early years if those gains are all lost by the time kids leave school. Overall I think the Cato graph comes out looking pretty good. But even if you disagree, I would ask you to pick whatever metrics you want – your favorite test, racial group, axis scheme, whatever – and tell me whether it really looks like the doubling-to-tripling of education costs during the relevant time period has been money well-spent. If not, then as happy as I am to debate details in the comments, it all seems like basically nitpicking.

Jonah Katz brings up a really complete analysis of increasing costs in higher education:

The rising cost of higher education isn’t quite so mysterious, at least for the last 10-15 years. The Delta Cost Project has put together some fairly comprehensive data about this. What you see across most categories of post-secondary institutions is that basically \*everything\* is becoming more expensive, but ‘student ‘life’ and ‘academic support’ are rising fastest, followed by ‘institutional support’. Student life is all of the bells and whistles (athletic centers, movie theaters, etc.) that colleges use to try to entice prospective students into paying huge amounts of money to enroll in their institutions, and I believe it also includes health and mental health services, which I would imagine have become exponentially more expensive over the past couple decades (this is probably unavoidable, because health costs are going up in general and universities are enrolling a far wider range of students with more mental and physical health issues who wouldn’t have gone to college in the past). Academic support includes a mix of stuff that is crucial to the academic mission of a university (libraries, IT systems), stuff that is arguably not part of the core academic mission at all (Dean’s Office personnel, museums), and stuff that is well intentioned but tends to be useless in practice (central offices for teaching and curriculum development). Institutional support is administration proper. Note that these data come from 2003-2013, so they don’t capture the explosion in university administration that is generally agreed to have occurred from roughly the 1970s to 1990s. I’ve never been able to find categorized data that goes back that far, but I imagine the change in spending on administration during that period must have been astronomical. The cost of instruction is still the largest single category of expenditure, and accounts for the majority of absolute price increases, but proportionally it is not rising as fast as these other categories. Also, the NY times Op-ed piece you link to is either selectively pulling misleading data or is just plain ignorant about the state of public financing for higher ed. There has not been a ‘modest’ reduction in per-student funding: it has dropped around 30% in inflation-adjusted dollars since 2000.

Static brings up the role of pensions as a driver of schooling cost increases.

Swing finds there have not been similar cost increases in the Netherlands. And Politifact rates as true a similar claim – that we greatly outspend other First World countries in per pupil spending. This article notes that we spend about 25% more than Britain and almost 50% more than Germany. On the other hand, the Netherlands is only a little better than we are, so this doesn’t match a scenario where the Netherlands’ spending goes up by only a little but America’s goes up by 200%. I don’t know where the discrepancy comes from. {EDIT: Douglas Knight points out that most countries have per pupil spending as a similar percent GDP)

Various people chime in with their favorite anecdotes about school vouchers working very well (DC) or working very poorly (Sweden). Murphy describes a personal bad experience with school privatization. At some point I do want to go through and sum up all the empirical literature on this, but not now.

Many people mention the possibility for bad incentives or market failures in schooling. 1soru1 thinks that, absent better signals of quality, schools will compete on shininess and raise prices to have the biggest and most breathtaking stadium (which I think is what the post above was saying happened to colleges, so certainly plausible). Tanagrabeast describes finding the private schools in Arizona heavily politicized: “What Scott worries about is already happening. I was skeptical until I took my son to an open-house at a fast-growing chain of charters where they tried very hard to play conservative buzzword Bingo and did all but lead us in a prayer to the Founding Fathers.” In contrast, Doctor Mist says that as a rightist, he feels like going to a private school lets him escape what he sees as public schools’ existing liberal politicization.

EarthSeaSky is a purist and reminds us that the free market which can be named is not the eternal free market.

Steve Sailer argues that for-profit colleges are a natural comparison group for for-profit primary schools, and they are very bad.

I talked to Education Realist on Twitter. Their position is complicated but they recommend their posts The Fallacy At The Heart Of All Reform and Charters: The Center Won’t Hold as introductions/summaries. I am still not entirely clear on their position – the objection seems to be that successful charters succeed only by taking the best students who would get good test scores anywhere, then claiming charters raise test scores. Obviously charters are trying this, but every halfway-decent study on charter schools has tried to control for this possibility. Also, none of my points involved empirical claims that charter schools raise test scores, so I don’t see why this discredits me in particular. They also note that US education is already pretty good both compared to other countries and compared to its own past, something else I agree is true and have never denied.

Levarkin brings up James Tooley’s fascinating work on private schools for the poor in Third World countries.

And Justreggedthis on the subreddit makes what I find the most convincing argument in this whole discussion:

Sweden’s experiment with school vouchers showed a different problem: the market delivers what you want, not what you need. What (stupid) parents want is good grades. What kids need is good education. So precdictably, voucher schools ended up diluting grades. You can probably imagine how it works. We live in an age of narcissism. Many parents want to hear their kids is super, special, and a genius, and get straight-A grades for a performance that is at best average. Few parents have the character left to stand up to it, and want challenging education and honest grading.

Of course this is a problem with people, not vouchers. I am sure the very same narcissism in modern culture also rears its head in public schools as well.

The classic solution was school principals having low time preference and interested in preserving the long-term good name of their school. So they would not agree to grade dilution, they would not encheapen the brand of their school.

Seems like today time preferences are high.

Grocery stores are a good parallel. You need healthy food. You want (a stupider version of you wants) gallon buckets of ice cream. Hence, you get all kinds of special offers and discounts on tasty and cheap gallon buckets of ice cream. The market delivers what you want. Hence, obesity epidemic.

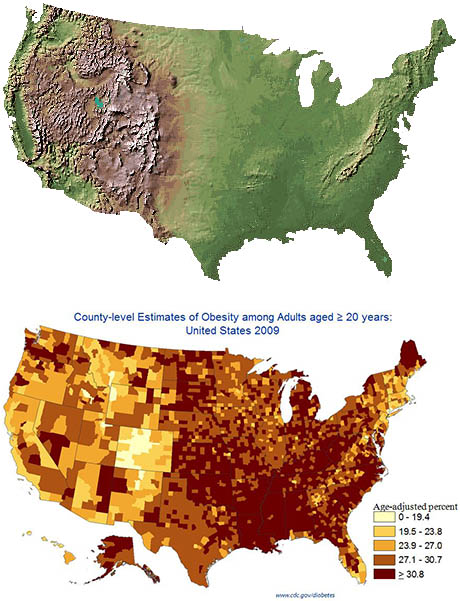
I don’t really know any solution that is acceptable within a democratic framework. Obvious someone somewhere should override personal preferences, but that someone should have a very good set of incentives and that is what we don’t get in this framework.

Overall reading this has made me somewhat more pessimistic about charter schools. But I’m still uncertain enough that I want to look into the empirical literature more, and I still think careful experimentation is the way to go.

…so maybe I should end with shadypirelli’s comment from the subreddit pointing out that Betsy DeVos’ policies cannot be described as “careful experimentation”.

# Thin Air

The International Journal of Obesity (h/t amaranththallium) points out a correspondence between US topography and US obesity rates:



It’s easy to see the Rocky Mountains on the obesity map. Not too hard to see the Appalachians either. Squint a little and you can even see California’s Central Valley vs. its coastal ranges.

This doesn’t seem to be related to poverty or population density. It does look a lot like the map of exercise level, but apparently it stays significant even when you control for that.

The IJO study finds that people living at sea level are five times more likely to be obese than people living at 500m elevation, even after controlling for “temperature, diet, physical activity, smoking, and demographic factors”. I don’t always trust controlling for things, but in this case the effect is big enough, and similar enough to the results of eyeballing, that it seems pretty plausible. Also, European studies find the same effect in high-altitude areas there, as do studies in Tibet. And someone did a study on US soldiers, who are randomly assigned (via deployment) to different areas, and found the same effect controlling for BMI at enlistment.

(on the other hand, a study in Saudi Arabia finds the opposite. Whatever. I didn’t even know Saudi Arabia had mountains.)

So what’s going on? There’s a well-known phenomenon called altitude anorexia where lowland people going to a high altitude suddenly can lose a lot of weight. Unfortunately most of the studies just stop at showing an acute effect; it’s not clear how long it lasts or whether there are more general principles involved. One study on rats found that they ate 58% less one day after being transported to Pike’s Peak, and were still eating 16% less per day two weeks afterwards. An article in High Altitude Medicine noted without further details that altitude anorexia seemed to persist after initial acclimatization. Pugh et al note weight loss of 1 kg/week up to 5-10 kg over a several week Everest ascent, reversing quickly as the climbers descended. A controlled experiment where obese subjects were ferried to the Swiss Alps, told to eat as much as they want, and banned from exercising resulted in weight loss of three pounds after a week, mostly sustained (?!) after a month at low altitude. It seemed mediated by eating less, which was independent of altitude sickness and persisted after people were no longer altitude-sick.

The active ingredient of altitude seems to be hypoxia. The air is thin at high altitudes so the body gets less oxygen. Being in low oxygen conditions in normal pressure seems to cause weight loss too – see here and here for studies of people exercising in low oxygen conditions. I don’t know of any studies where people were just kept in low-oxygen environments for a long time without exercise to see what happened to their weight. It’s not really clear how reduced oxygen makes people eat less. A lot of people mention leptin, but the studies seem pretty unconvincing, and people try to work leptin into everything.

This BMJ editorial suggests that hypoxia should get credit for smoking-related weight loss. But it reads like a completely unhinged screed from the tobacco lobby of an alternate dimension (“Might the aggressive anti-smoking lobby have contributed to the costly epidemic in obesity and type 2 diabetes that Professor Sir George Alberti have warned us about?”) so maybe we shouldn’t take it too seriously. Also, nicotine gum works just as well as cigarettes here, so it’s probably an effect of the nicotine itself, and in fact we have some pretty good ideas how this happens. Maybe unhinged screeds by alternate-universe tobacco lobbies aren’t the most trustworthy source of information.

Anyway, this is boring. Let’s move on to a more interesting question – did global warming cause the obesity epidemic?

The arguments in favor: there’s a lot more carbon dioxide in the air now than there was just a few decades ago. The obesity epidemic began around the time carbon dioxide concentrations really started getting worrying. And there are various body functions that are exquisitely dependent on CO2 levels. Bierworth (2014) has a good run-down of some of these and how they might be affected by increasing atmospheric CO2 levels (dear conservatives who always talk about Chesterton’s Fence and principle of precaution – has it occurred to you that doubling the concentration of a major bioactive atmospheric gas might be a bad thing?). I see some conflicting claims about how much atmospheric CO2 could affect average blood pH. Neurons that produce obesity-regulating chemical orexin are potentially very sensitive to blood pH, so maybe this could be involved?

Wild animals are affected by the obesity epidemic too, even though they eat far fewer Big Macs. Even lab rats and zoo animals, supposedly kept on a well-monitored diet, are heavier now than they were decades ago. It’s hard to think of some obesogenic factor so prevalent that it could seep into laboratories and zoos unnoticed by scientists and zookeepers. If it were a chemical, it would have to be really prevalent. The xenoestrogens in the water are one possiblity. But the other is the atmospheric gas breathed in by every living thing which we already know has been increasing for decades.

This at least is the theory of epidemiologist Lars George Hersoug. He did a study where he put some people in a high-CO2 room and found that they gained weight. It got a decent amount of press.

I really like this theory. It’s elegant. It’s clever. It’s at exactly the right level of contrarianism to be fun. If it were true, it would solve global climate change – once tabloids covers trumpet THE ONE SECRET TO A TIGHT BELLY DOCTORS DON’T WANT YOU TO KNOW – TELL YOUR CONGRESSMAN TO PASS THE PARIS AGREEMENT TO LIMIT CLIMATE CHANGE TO WITHIN 2 DEGREES CENTIGRADE OF PRE-INDUSTRIAL LEVELS, we will finally get Middle America on board. Insofar as scientific theories can be “fun”, this theory is fun.

But I grudgingly acknowledge that it’s probably not true.

For one thing, the study involved kind of sucks. It has a sample size of six. The six people were in a chamber that elevated CO2 levels about 50x higher than human industrial activity has elevated them in the atmosphere. And it got a non-significant result. In fact, food intake decreased in three of the six subjects, with almost all of the (nonsignificant) positive trend coming from one guy who apparently was just really hungry that day. There is a decent study showing CO2-related orexin effects in mice, but it’s at 2000x atmospheric concentrations. Also, when I look at the orexin neuron calculations, even if their hypothesis is true it suggests that orexin neurons might fire a little less than 1% more often now than they did 100 years ago. Unless there’s something really nonlinear going on, this is not enough to cause an epidemic.

For another, it doesn’t seem to line up geographically. Zheutlin, Adar, and Park try to correlate the geography of US obesity with the geography of US atmospheric CO2 in the same way that some of the studies above successfully correlated US obesity with US altitude. Here they fail. After adjusting for appropriate confounders, there is no clear relationship between CO2 levels and obesity. I think this might also reflect a more general point, which is that CO2 has been rising all over the world but obesity hasn’t; Japan, for example, is a very high CO2 emitter but has almost totally avoided a US-level obesity crisis.

(but while we’re talking about this study, it did find that serum bicarbonate has been increasing over the past decade or two. No proof as yet that this is a real effect or related to CO2, but have I mentioned that increasing the concentration of a bioactive atmospheric gas worldwide is a really bad idea?)

One last counterargument. A global warming skepticism blog points out that submarines are a natural laboratory for the effect of high CO2 on human health, since they usually have CO2 levels up to ten times atmospheric (and several times worse than even a poorly-ventilated building). I don’t see any formal tests of their own argument, which is that submariners don’t suffer any cognitive problems, and I’m not sure they’re right to use intuition and failure to notice gross impairment – after all, the original studies showing impairment were done in office buildings, and there’s no grossly noticeable differences there. But in any case, the Navy actually did a formal study and found that submariners do not gain weight. This seems pretty fatal for a CO2 = weight gain theory.

So my guess is that Hersoug is wrong and CO2 doesn’t cause appreciable weight gain in normal concentrations. We should abandon the beautiful theory of climate-change-induced obesity and go down a level of contrarianism to blaming boring normal-person things like xenoestrogens and gut microbiota.

(I’ve heard there are theories of obesity even less contrarian than those, but I’ve never been to such low contaranianism levels and wouldn’t be able to tell you what they might be.)

# Might People On The Internet Sometimes Lie?

From Reddit: Parents Of Children Who Claim To Have Had Past Lives, What Did They Tell You?. Some sample comments:

When he was 6 years old my son described in great detail my grandmother’s house he never been to. This was in 1986 or so, pre-internet. There are no pics of the place that I’m aware and no one owned a camcorder in our family, so video is out of question either. It’s a small house with red roof and a purple door (grandma painted the door every couple of years). He described all of it – that it had one big room with a fireplace across from the window, he explained where the doors are located, how there always were some boxes under the stairs, that there always was a faint smell of apples in the house (grandma ran a small time apple sauce business). That there was this cat almost completely white with a black spot around his right eye (that’s mr. Whiskers, my grandma’s cat!).

My grandma and Whiskers both died in 1977, 3 years before my son was born. To this day I can’t fathom it and can’t even get a remotely sane explanation on how does he know all this. I never told him about it, my wife has never met my grandma and never been to her house and in 1986 we were stationed in Germany, so none of my old friends could have reached my son, so this is definitely not someone’s prank. Best part of this is my son says he doesn’t remember telling me that, but my wife heard him saying that too, so if definitely happened!

And from Reddit, What Is The Creepiest “Glitch In The Matrix” You’ve Encountered?:

When I was in school I had this hippie teacher who would always tell us that the universe can help if you just ask it.

She told us one time her daughter had lost something very important and when she asked the universe to help she suddenly had a massive pulling feeling towards the sink. She walks over and immediately stuck her hand down into the garbage disposal and pulled the item out in perfect condition.

So I think it’s total bullshit of course, but later that day I was searching for a thin little booklet that I really, really needed for school. I spent 3 hours looking for it and had no luck. Finally out of frustration I almost sarcastically said, “I need your help universe.” I immediately walked over to this bookcase filled with books from my step dad. I had never once used this shelf or any book on it.

I grab a random book I’ve never seen from the middle of a huge pile. I open it to somewhere around page 200 and right there is my booklet smashed in between the pages. It was incredibly thin so you couldn’t even tell there was anything in there if you looked at it from another angle.

I’m sure there’s a good explanation, but it’s been well over a decade and I still remember the incredibly freaky vibe I got the moment I saw the book.

I don’t believe in reincarnation or paranormal forces. When I read stories like this, my first impulse is to try to think of reasonable explanations or ways they could be a coincidence. Maybe some kids have instinctive talent at that sort of cold-reading thing TV psychics do sometimes. Maybe your unconscious can remember where you put a booklet and then repress it from the conscious mind for some reason.

But these kinds of claims are often themselves far-fetched. If I told you in normal conversation, unrelated to compelling reincarnation theories, that kids have a natural talent at cold reaading, you’d scoff and demand proof. And it’s not just reincarnation and booklet-finding. If you read Reddit enough, you’ll find hundreds of equally compelling stories of telepathic contact, cryptid sightings, UFOs, et cetera.

So. Alternate hypothesis. About one million people view Reddit every day. Let’s assume 10% of those see threads like the above – which were pretty popular and which I think both made it to the front page. That’s 100,000 people. Now let’s assume that even 1/10,000 people on the Internet are annoying trolls, which is maybe the easiest assumption we’re ever going to have to make. If each of those annoying trolls posts one fake story to a thread like that for the lulz, that’s enough for ten really convincing stories per thread – which is really all there are, the other fifty or sixty are just the usual friend-of-a-friend-had-a-vague-feeling stuff.

(it’s true that in a site read by a million people, there will also be far more people who have experienced a genuine one-in-a-million coincidence, but that shouldn’t scale nearly as quickly; after all, liars can invent coincidences way more far-fetched than the sheer numbers would allow)

This hypothesis seems obviously right. If I ask “what’s the chance that at least one in ten thousand Internet users is an annoying troll?” you laugh hysterically and tell me that nobody has even invented numbers that high. It perfectly explains mysterious events that would otherwise require impossible coincidences or weird theories about hidden brain functions. So why is it so hard to make myself believe?

I think part of it is a failure of scale. Reddit looks a lot like a normal forum or blog comment section, the sort of BBS I used to go on as a kid with twenty or thirty regulars who would dominate all the discussions. If indeed 1/10,000 people is the sort of jerk who would make up a story like this just to troll people (or even 1/1,000 or 1/100 people), the chance that I’d run into them on my little BBS/comment section/Dunbar-number-group is pretty low, and I can safely ignore the possibility that five different crazy paranormal comments are all by pathological liars. It’s only when you get a place like Reddit, which manages to feel like a community while also having a million readers a day, that you have to start thinking about these things.

This suggests a more general principle: interesting things should usually be lies. Let me give three examples.

I wrote in Toxoplasma of Rage about how even when people crusade against real evils, the particular stories they focus on tend to be false disproportionately often. Why? Because the thousands of true stories all have some subtleties or complicating factors, whereas liars are free to make up things which exactly perfectly fit the narrative. Given thousands of stories to choose from, the ones that bubble to the top will probably be the lies, just like on Reddit.

Every time I do a links post, even when I am very careful to double- and triple- check everything, and to only link to trustworthy sources in the mainstream media, a couple of my links end up being wrong. I’m selecting for surprising-if-true stories, but there’s only one way to get surprising-if-true stories that isn’t surprising, and given an entire Internet to choose from, many of the stories involved will be false.

And then there’s bad science. I can’t remember where I first saw this, so I can’t give credit, but somebody argued that the problem with non-replicable science isn’t just publication bias or p-hacking. It’s that some people will be sloppy, biased, or just stumble through bad luck upon a seemingly-good methodology that actually produces lots of false positives, and that almost all interesting results will come from these people. They’re the equivalent of Reddit liars – if there are enough of them, then all of the top comments will be theirs, since they’re able to come up with much more interesting stuff than the truth-tellers. In fields where sloppiness is easy, the truth-tellers will be gradually driven out, appearing to be incompetent since they can’t even replicate the most basic findings of the field, let alone advance it in any way. The sloppy people will survive to train the next generation of PhD students, and you’ll end up with a stable equilibrium.

The weird thing is, I know all of this. I know that if a community is big enough to include even a few liars, then absent a strong mechanism to stop them those lies should rise to the top. I know that pretty much all of our modern communities are super-Dunbar sized and ought to follow that principle.

And yet my System 1 still refuses to believe that the people in those Reddit threads are liars. It’s actually kind of horrified at the thought, imagining them as their shoulders slump and they glumly say “Well, I guess I didn’t really expect anyone to believe me”. I want to say “No! I believe you! I know you had a weird experience and it must be hard for you, but these things happen, I’m sure you’re a good person!”

If you’re like me, and you want to respond to this post with “but how do you know that person didn’t just experience a certain coincidence or weird psychological trick?”, then before you comment take a second to ask why the “they’re lying” theory is so hard to believe. And when you figure it out, tell me, because I really want to know.

# SSC Journal Club: Mental Disorders As Networks

I.

Suppose you have sniffles, fatigue, muscle aches, and headache. You go to the doctor, who diagnoses you with influenza and gives you some Tamiflu.

There’s some complicated statistics going on here. Your doctor has noticed some observable variables (sniffles, fatigue, etc) – and inferred the presence of an invisible latent variable (influenza). Then, instead of treating the symptoms with eg aspirin for the headache, she treats the latent variable itself, expecting its effects to disappear along with it.

Psychiatry tries to use the same model. You get some symptoms – depressed mood, insomnia, fatigue, feelings of worthlessness, suicidality. You go to the psychiatrist, who diagnoses you with depression and gives you an antidepressant.

The psychiatrist is implicitly assuming that the causal structure of her field matches the causal structure of better-understood diseases like influenza. Generations of psychiatrists have noticed that different symptoms all tend to show up together and follow a similar pattern, suggesting some kind of deep connection between them. So psychiatrists follow the influenza model and attribute this collection of linked symptoms to a latent variable called “depression”.

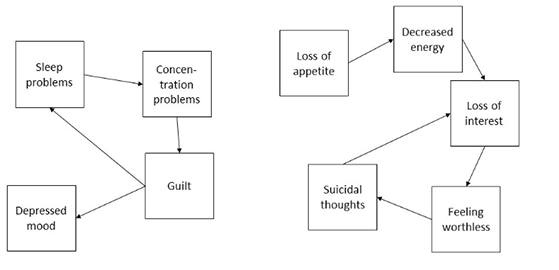
This gets complicated really fast. Psychiatric disorders are diagnosed through clusters of symptoms, but we don’t expect every person to have every symptom in the cluster. For example, we diagnose depression when a patient has five out of nine symptoms on a list including fatigue, guilt, sleep disturbance, suicidality, et cetera. Each of these symptoms is often but not always present in a patient who has most of the others – for example, 75% of depressed patients have sleep disturbances, but 25% don’t.

But all psychiatric disorders are hopelessly comorbid with each other. If someone meets criteria for one DSM disorder, there’s a 50% chance they’ll have another one too. 60% of people with major depression also have an anxiety disorder. This is awkward when compared to eg the 75% sleep disturbance rate. Why are we calling sleep disturbance a “symptom” of depression, but anxiety a “comorbid condition” with depression? If we’re trying to cluster symptoms together to identify conditions, how come “sleep” is grouped with a bunch of other symptoms in the depression cluster, but “anxiety” gets to be a cluster of its own? Are there really two conditions called “depression” and “anxiety”, or just one big condition that has various symptoms including low mood, sleep disturbance, and anxiety, and some people get some of the symptoms and other people get others? I’m told that the people who write the DSM have long conversations about this using rigorous methods, but to the rest of us it seems kind of arbitrary.

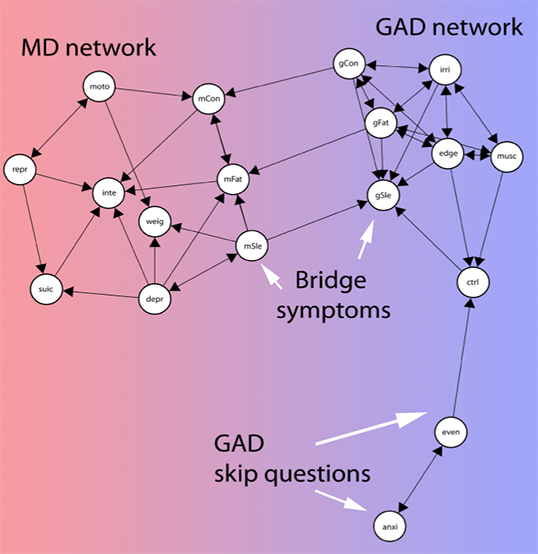
The problem isn’t that nothing ever clusters together – depression, for example, is a very natural category. But so are various subtypes of depression. And so are various supertypes of depression, like depression + anxiety, or depression + psychosis, or depression + anxiety + psychosis. Choosing to draw the borders around depression and say “Yup, this is the Actual Disease” isn’t a bad choice, but it doesn’t jump out of the data either. When people try to use sophisticated clustering algorithms on psychiatric disorders, they usually come up with something like this, where there are only three supercategories instead of the 297 different diagnoses in the DSM. And even three supercategories are pushing it – people with psychosis are far more likely to have depression too! Having any number of categories starts seeming arbitrary and fuzzy.

So Nuijten, Deserno, Cramer, and Borsboom (from here on: NDCB) ask: what if that’s wrong? What if there isn’t a latent variable like “influenza”? What if it’s symptoms all the way down?

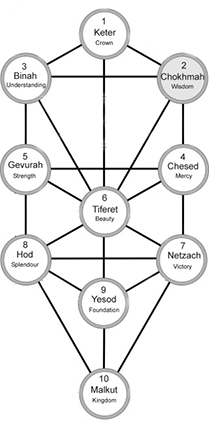
Consider a network in which each symptom is a node, connected to all the others by pathways with certain weights on each direction. So for example, “sleep disturbance” might be connected to “fatigue” by a strong path – people with disturbed sleep are much more likely to be tired. These might both be connected to “low mood” – people who don’t sleep well, or who are tired all the time, start feeling down about themselves. And this path might go the other way too: people who feel down about themselves might have more trouble getting to sleep on time. And maybe all of these are connected to suicidality, because if you feel bad about yourself you’re more likely to commit suicide, and if you’re suicidal you might feel bad about it, and if you’re tired all the time then maybe you can’t accomplish anything useful with your life and so death might seem like a good way out, and so on.



A sample image from the paper, showing two possible simple networks of depression symptoms



Also from the paper. This shows a more complicated (and apparently empirically validated) network of symptoms. MD is major depression. GAD is generalized anxiety disorder. The nodes are all different symptoms – for example, “inte” is “loss of interest in activities” and “musc” is “muscle tension”.



Not from the paper. But if you figure out a good way to calculate weights on this one, email me.

Each node might affect the others with a certain delay. Being suicidal might make you feel guilty, but even if your last suicidal thought was fifteen minutes ago, you might still feel guilty now. Maybe it would take months or even years before you no longer felt guilty about your suicidal thoughts. So there could be loops: in a simple model, your low mood makes you feel suicidal, your suicidality makes you feel guilty, and your guilt makes you have low mood. This type of loopy network might be stable and self-reinforcing. Maybe your boss yells at you at work, which makes you have a bad mood. Then even if the direct effect of your boss would go away quickly, if it causes suicidal thoughts which cause guilt which cause more low mood, then the cycle can stick around forever.

In NDCB’s model, all possible psychiatric symptoms are connected like this in a loose network. Particularly tight-knit symptom clusters that often active together and reinforce each other correspond to the well-known and well-delineated psychiatric diseases, like depression and schizophrenia. But there are no natural boundaries in the network; low mood and poor sleep may be closely connected to each other, but they’ll also be more distantly connected to anxiety, and even more distantly connected to psychosis. This corresponds to the fact that some depressed people will develop psychotic symptoms, even though psychosis isn’t usually associated with depression. The paths aren’t usually as strong as those between low mood and poor sleep, but they’re there, and in some people with a predisposition to psychosis or some idiosyncratic factor strengthening those paths beyond their usual level in the population, that will be enough.

There are lots of good things about thinking about psychiatric problems this way:

1. It helps explain how life stressors can cause depression. Some people who have a bad breakup will get depressed. This should be mysterious if we think of depression as a biological illness – and we have to at least a little; some people who take the drug interferon-alpha will get depressed afterwards too. But if depression is a symptom network, it becomes easier to explain. The bad breakup causes low mood, which under the right conditions and genetic predispositions can activate all of the other depression symptoms and create a stable, self-reinforcing depression. Likewise, poor sleep is a risk factor for the development of subsequent depression, which is hard to explain if we just think of it as a symptom of some latent-variable-style condition.

2. It explains how treating depression symptoms can treat the depression. I’ve heard a lot of different perspectives on this, but at least one of my attendings (and some studies) believes that treating poor sleep with a sleeping pill like Ambien can help dispel an underlying depression, including symptoms seemingly unrelated to sleep like “feelings of worthlessness and guilt”.

3. It explains how therapy can treat depression. If eg cognitive behavioral therapy helps you stop thinking of yourself as worthless, then you’ve de-activated the “feelings of worthlessness and guilt” node and made it a lot harder for all the other nodes to coalesce into a stable self-reinforcing pattern.

4. It explains the polygenic structure of mental illnesses. If a mental illness were one specific thing, we would expect it to have one specific cause, or at least be limited to genes active in one specific area or process. In fact, it’s hard to come up with anything that genes involved in these illnesses have in common other than “they’re mostly expressed in the brain” – and sometimes not even that. In NDBC’s model, genes might be involved in any of the symptoms, or in the paths between the symptoms. A gene involved in poor sleep could predispose to depression. So could a gene involved in low energy levels. Even a gene involved in anxiety or psychosis could have some effect. And so would any gene that influenced the probability that, given poor sleep, a person would have low energy levels; or that given anxiety, a person will have psychosis. The end result would be everyone having a slightly different network, with different amounts of work needed to activate each node and different weights on each of the inter-nodal paths.

5. It helps explain why so many brilliant people searching for The One True Cause Of Depression have come up empty.

II.

Actually, this last one deserves more explanation. NDCB think of these symptoms as visible patient complaints (“poor sleep”, “feelings of worthlessness”), and treat the connections between them as common sense (“if you don’t sleep, you’ll probably be fatigued”, “if you feel very guilty, you might attempt suicide because you think you deserve to die”). But their theory also works for networks of biological dysfunctions, or networks that combine biological dysfunctions with common-sense observed symptoms.

For example, we know that there’s a link between depression and inflammation. But it’s not a very good link; not all depressed people have increased inflammation, not all people with increased inflammation get depressed, and drugs that decrease inflammation don’t always cure depression. There’s similarly good evidence linking depression to folate metabolism, serotonergic neurotransmission, BDNF levels, and so on. Suppose we made a graph like the ones above, except that instead of putting things like “poor sleep” and “feelings of guilt” on it, we used “inflammatory dysfunction”, “folate metabolism dysfunction”, “serotonin dysfunction”, and “BDNF dysfunction”. There are a lot of reasons to expect these things to interconnect – for example, folate helps produce a cofactor necessary for serotonin synthesis, so any dysfunction in folate metabolism could make a problem with serotonergic neurotransmission more likely.

In a best case scenario we could merge the biological and psychological perspective, replacing “disturbed sleep” with “disturbance in the orexin and histamine systems that regulate sleep” and “tiredness” with “disturbance in the dopamine system that regulates goal-directed action”, and so “poor sleep makes you tired” with “disturbance in the orexin system causes a disturbance in the dopamine system”. In practice I expect this would be a terrible idea and that common-sense concepts mostly don’t have simple well-delineated biological equivalents. But what I’m saying is that the model where all of these things are observable symptoms, and the model where they’re all disturbances in brain chemicals and metabolism, aren’t necessarily in conflict.

So we can expand point (5) to say not only that it explains why nobody has found the One True Depression Cause, but why they have found so many promising leads that never quite pan out. Just like depression has a bunch of different symptoms, each of which is often-but-not-always involved, and each of which reinforces the others — so it has a bunch of different disturbances in biological systems, each of which is often-but-not-always involved, and each of which reinforces the others. Maybe there’s a nice correspondence between one disrupted biological system and one symptom, or maybe they sit uneasily together as different nodes on the same big graph.

III.

Are there any problems with this theory?

There are a couple of disorders that really don’t fit this model. Bipolar disorder, for example, doesn’t quite work as a collection of self-reinforcing symptoms. It’s marked by depressive episodes that can give way to years of stable mood before the person has a manic episode months or years later. I can’t think of any way to model this except as some underlying unified tendency toward bipolar disorder – although the ability for this tendency to cause a depression that looks just like normal unipolar depression is a point in NDCB’s favor, since it suggests there can be many different causes for the same syndrome.

The impressive success of ketamine also counts as a point against. NDCB imagine psychiatric disorders like depression as gradually fading out on a symptom-by-symptom basis, eventually reaching a point where enough symptoms are gone that the rest of them aren’t self-reinforcing and just sputter out. This matches the course of eg SSRI treatment, where the medications will gradually improve a few symptoms at at time over the space of a month or so and maybe cause a full remission if you’re lucky. It doesn’t really match ketamine, where every aspect of depression vanishes instantly, then returns after a week or so without treatment. There are a couple of other equally impressive things – staying awake for thirty hours straight, for example, can have an immediate and near-miraculous antidepressant effect, which unfortunately vanishes as soon as you go to sleep. Both of these treatments seem like direct strikes against the One True Cause Of Depression, and both suggest that an underlying tendency toward depression can exist separate from any symptoms (or else why would the depression come back after the effects of the ketamine wore off?)

I don’t think it’s possible to cure depression by blasting every symptom simultaneously. That is, suppose somebody is depressed with symptoms of poor sleep, poor appetite, low energy, suicidality, and low mood. Ambien can make them sleep. Pot can make them eat. Adderall can give them energy. Clozaril can make them stop wanting to kill themselves. And heroin can perk up mood. So if you gave someone Ambien, pot, Adderall, Clozaril, and heroin at the same time, would that cure their depression? I’m pretty sure no one has ever tried this, but I don’t think anyone’s reported exceptional results from less extreme cocktails like Adderall + trazodone + pot, which I’m sure a bunch of people end up taking. This along with the stuff from the last paragraph suggests that if we want to go with this model, maybe we should think less in terms of actual poor sleep and more in terms of dysfunction in the biological system of which sleep is a visible correlate. In that case we could say that Ambien helps the sleep itself but not the underlying dysfunction. But that takes some of the elegance out of the theory.

Despite these issues, I feel like something along these lines has to be true. There are too many things that sort of kind of cause psychiatric problems, and too few things that look like One True Causes. Things that look a lot like schizophrenia can be caused by viral infections in utero, by genetic factors, by hitting your head really hard as a child, by hypoxia during the birthing process, by something something something intestinal tract, by something relating to immigration which seems like it might involve psychosocial stress, and so on. Studies of the immune system, the dopamine system, the glutamate system, and the kynurenine system have all found disruptions. There have been so many really brilliant attempts to reduce all of these to a single brain region, or the levels of one specific chemical, or something that’s simple in the same way that lack-of-insulin-causes-diabetes is simple. But nobody’s ever succeeded. Maybe we should just give up.

I guess I’ve felt for a long time that some kind of weird change in attractor states of biological systems is the best way to explain these kinds of things, but I was never able to express what I meant coherently besides “weird change in attractor states of biological systems”. NDCB offer a clear model that suggests good avenues for future research.

(And I wasn’t joking when I said that little diagram with the two pentagons was the solution to 25% of extant philosophical problems.)

# Book Review: Mount Misery

[Content warning: psychiatric abuse (especially around borderline personality), rape, spoilers for Mount Misery]

I.

Last month I reviewed Samuel Shem’s House of God. The sequel, Mount Misery, is about his time training in psychiatry. This is obviously relevant to my interests, so I picked it up.

It’s weird to accuse someone of writing a cheap knockoff of their own book, but Mount Misery reads like a cheap House of God knockoff. There are all the same elements – a young doctor, an incompetent system, cruel hospital administrators, a kind mentor. But in the first book, it all came together perfectly. In this one, it was more hit-and-miss. Sometimes the imagery clicked; other times, it just seemed like caricatures. Creepy magical realism alternated with guys who announced “I hate patients! Let’s just pump them as full of drugs as possible and leave them to die!” and then zoomed away in fancy sports cars they bought with pharma money. Maybe it’s just less funny when it hits closer to home? I don’t know, but it was less funny.

The book’s frame story follows Dr. Roy Basch, who has left his abusive medical internship to do a psychiatric residency at the Mount Misery hospital, lured there by the kind and decent Dr. Ike White. But Dr. White commits suicide Basch’s first month of the job, the hospital administration ineptly covers it up, and nobody wants to talk about it – something something metaphor even psychiatrists stigma something metaphor. Having lost his mentor, Basch is thrown back and forth among various attendings – the one who thinks Freud solves everything, the one who thinks drugs solve everything, the one who thinks a thinly-veiled parody of Otto Kernberg solves everything, et cetera. The only sympathetic character is Dr. Malik, who tells Basch to ignore the theories and try to connect with his patients as human beings; Malik is of course loathed by all the other doctors and ostracized from all the good positions in the institution.

The typical psychiatric treatment in Mount Misery works as follows: Take someone who has some awful stuff going on in their life but is pretty much functional. Declare them to be a perfect example of whichever theory is popular that week (“this person clearly is obsessed with the idea of sucking their father’s penis”), then insult any underlings who don’t buy-in as being ignorami who refuse to understand the complexities of the human mind. Ignore all of the patient’s human needs in favor of the theory – if they complain that their room is cold, tell them that’s a reflection of the coldness of their father. If they protest that no, they’re really cold, then mark them as “resistant” and double-down on your theory since they obviously need a lot of help. The angrier the patient gets, the more you’ve obviously hit a nerve and the better a psychiatrist you are. Repeat the process until they are curled up in a ball, completely nonfunctional, which you will call “successfully regressing the patient” and “revealing the repressed pathology”. Then keep them in hospital until their insurance runs out, at which point discharge them to be someone else’s problem.

When someone like Dr. Malik doesn’t do the typical treatment, the fact that his patients never get reduced to curled-up balls gets held against him. He’s so “superficial” that he just takes the patient’s complaints about being cold as a reference to real physical coldness in the environment! If his patients say they’re upset at losing their job, he’s so “superficial” that he just talks to them about their job and how they can support themselves financially! A janitor could do that! When his patients fail to be reduced to curled-up balls of rage, that obviously proves he’s not nailing their real emotional problems, not getting past their defenses, not successfully regressing people, and just generally incompetent.

The master of this kind of anti-treatment is Dr. Heller – the hospital’s specialist in borderline personality disorder – who believes that it’s psychodynamically important to bring out the latent negative transference in borderlines (ie make them hate you). When Dr. Basch, at Dr. Malik’s urging, tries being nice to a borderline patient instead, and gets much better results than Heller ever has, the expert lectures him on his mistake:

Heiler explained that his technique, “confrontation”, had evoked the anger that was hidden in each and every borderline. “She wasn’t angry at me“, he said, “it was her transference to me. She was distorting her real relationship with me based on early infantile experiences, with her bad mom, in the first year of life.”

“But she was angry at you,” I said. “Anyone would be.”

“Not that angry,” he said. “Not borderline angry.”

“How do you know that she’s a borderline?”

“Because of that incredible anger.”

“But she didn’t start out angry – you provoked it.”

“Who says?”

“I do! It was obvious.”

“You? You, who’ve been a psychiatrist two whole months? You don’t know diddly-squat about treating borderlines. Your so-called ‘concern’ is going to be a real problem – you’re already overinvolved, imagining that you can rescue her. To you, what I did seemed cruel, right? […]

His voice softened, and he went on, “Look. I know that the first time you see it, this theory seems strange – it’s counterintuitive. If just being nice to borderlines worked, don’t you think I would do it? Of course I would! In fact, I tried, way back, at first. But it’s like dealing with difficult children: you’ve got to be firm. Everybody knows that if you don’t dig up the Latent Negative Transference in these gals, next thing you know you’ve got people killing themselves, or killing other people. For fifty years people have been trying to cure borderlines by being ‘nice and human’ to them. Everybody felt better, nobody got better. It’s easy to act nice, it’s hard as hell to stand firm and confront the rage locked up in borderlines. Borderlines are hell. There aren’t too many of us left who have the guts to treat ’em. I’ve specialized in borderlines for years and years, and I’ve seen what works: You go through that rage to the truth, to their miserable pain and suffering, and believe me, they get better.”

“But,” Solini said, “I mean, everyone agrees that the lady [did get better when we were nice to her]”

“In this case,” Blair said, “Better is worse. She’ll have to get worse – which is in fact better – in order to get better, which will still be worse. If she gets a little worse, she won’t get a lot better, but if she gets a lot worse, she may get a little better. Not smarmy-‘nice’ better. Borderline better…Don’t worry, Roy. Your overinvolvement with her is normal. Sick, but normal. Gals like her are experts at getting guys like you entangled. Read my paper, Rescue Fantasies In The Naive Resident“

This speech could be a word-for-word transcription of something one of my attendings said to me during my intern year when I tried being nice to a borderline patient. There is a subtle sense in which this attitude can sometimes be helpful. But get the subtlety even slightly wrong and it devolves into being really evil, and Mount Misery brings out the worst in it.

Dr. Basch’s first therapy patient is a man named Cherokee, a rich WASP lawyer. He’s obsessed with the paranoid fantasy that his wife is having an affair with her psychoanalyst, a Mount Misery luminary named Dr. Dove. Basch tries everything with Cherokee – drugging him up, uncovering his latent homosexuality, suggesting he hates his father – but eventually Cherokee commits suicide anyway (“eventually he commits suicide anyway” will be a common theme among characters in this book.) In the aftermath, it is discovered that – surprise! – his wife was having an affair with her psychoanalyst, and also it was kind of coercive and bordered on rape.

Basch starts an investigation and learns that this same Dr. Dove is molesting a bunch of his female patients, and various other tangentially related people for good measure. He tries to expose Dove, but Dove denies everything, and he’s a bigwig who can get the administration to take his side.

The description of the ensuing investigation is beautifully done, precisely because it avoids some of the caricatures of the rest of the book. Dr. Dove isn’t portrayed as an ogre grumbling about “lying whores” or whatever. He sounds to all the world like a caring psychoanalyst, who understands that his patients are fragile and that stress of discussing sexual fantasies in psychoanalysis can sometimes break out into the patient’s consciousness and cause them to behave as if those fantasies actually occur. Yet all of this just serves to make him creepier and more hate-able.

In a particularly sharp scene, Dove capitalizes on the occasion to team up with a colleague and offer workshops about how to protect yourself from false accusations of assault in psychiatry:

The slide show ended. Dr. Shpitzer then made a heartfelt statement that patient-psychiatrist contact was absolutely off-limits. Touching the patient, but for a handshake, was off limits. A hug was totally out of bounds. Yet what was the psychiatrist to do when a female patient, maybe a borderline or dissociative or multiple, suddenly got up out of his chair and approached, intent on hugging him? Dr. Shpitzer asked Dr. Dove to demonstrate. Schlomo, ever the showman, popped to his feet.

First Shptizer said he would show us all what not to do, and told Shlomo to go ahead. Playing the woman patient, Schlomo started toward Shptizer, arms forward. Shpitzer crouched in a martial-arts stance and with a scream \_ HYAH! – karate-chopped Schlomo’s hands down…The discussion then centered on variants of this technique. Dr. Shpitzer passed out his brochure, describing his video course – “Six Quick Steps To Avoid The Pitfalls Of Risk” – which we could all buy for $399.95. This would allow us to pass out risk-management requirements for state relicensure as shrinks in the comfort and privacy of our very own homes.”

A psychiatrist actually rapes a patient, he doesn’t get punished because he’s a very important guy who’s friends with all the bigwigs, but everyone has to feel like they’re doing something, so they ban all normal human contact with patients, and also sell $399.95 courses that you can use to prove you’re compliant with patient protection regulations. This may be the best metaphor for life that I have ever heard.

(it doesn’t hurt that I’ve had to go through courses on whether it’s ever appropriate to hug patients, or that I once had to finagle my way out of attending a conference that was basically this guy’s $399.95 video lecture)

II.

One of the main themes of this book is that psychoanalysis makes people worse.

The book doesn’t claim that psychoanalysis isn’t effective. It treats it as powerful and worthy of respect. The book’s psychoanalysts are consistently able to tell weird facts about a person from just a glance, to strip them down to their deepest insecurities in minutes. It’s just that people who are healthy and decent going into psychoanalysis end up cracked and nasty coming out of it. A lot of the worst doctors at Mount Misery were decent people before they started getting analyzed themselves. Of course, it would have helped if their analyst wasn’t a sexual predator, but the book treats the process as dangerous even aside from that.

When Basch asks his supervisor Dr. Lowell what to do about the man who believes Dr. Dove is having an affair with his wife, Dr. Lowell describes the analytic technique:

“But what about Schlomo fucking his wife? You don’t think it’s true?

“There is no truth, there is only the individual perception of experience.”

“Wait a minute. The truth is that I’m taller than you.”

“That’s not the truth, that’s your transference to me.”

“We can measure it. To see who in fact is taller.”

“You think ‘taller’ can be measured?”

I saw her point. She wasn’t only aware of the objective fact, she was also aware of the deeper meaning psychologically. “But I’m stuck,” I said. “I don’t know what to do.”

“You have to go deeper into his obsession, find the deeper meaning, the roots of it in his childhood, his past.”

This was exactly what Malik had warned me against doing. Suspicious, I asked: “How?”

“If he talks feeling, you talk thought. If he talks thought, you talk feeling. If he talks past, you talk present. If he talks present, you talk past. You the doctor talk constantly about what he the patient doesn’t want to talk about. This is the analysis of the resistance. Then, when he starts distorting his relationship with you and calling you a sonofabitch for not talking about what he wants to talk about, then you do the analysis of the transference, telling him he’s treating you like his father, his mother, his aunt Sally, whatever. On a deeper level still, you can analyze the resistance to the transference, and the transference to the resistance. Not to mention the countertransference to each – but that’s still way beyond you at this point.”

Finally I felt I was getting some concrete advice about what to do in therapy.

I highlighted this last line because this is seriously much more concrete and actionable advice than anyone has ever given me about psychodynamic therapy and I’ve been doing it for two years.

And a lot of this rings true. I remember one time one of my patients missed a session because his flight back from vacation was delayed. I told my supervisor this and he got angry with me, saying it was superficial to blame it on the flight instead of talking about which of my comments had triggered the patient and made him decide to miss his plane. I insisted that we’d had a perfectly good session the week before, that the delayed plane had just been a delayed plane, and me and my supervisor got angrier and angrier at each other for both missing what the other thought was the point. Finally I got on the Internet and managed to prove that my patient’s plane really had been delayed to the point where it was impossible for him to have made my appointment, at which point my supervisor switched the discussion to why it was so important to me to believe that his plane had been delayed that I would do an Internet search about it, and whether I was trying to defend against the unbearable notion that my patient might ever voluntarily miss one of our sessions. My supervisor’s treatment of whether planes ever get delayed seems a lot like Basch’s supervisor’s treatment of who’s taller.

And I don’t think these people are literally so stupid that they don’t understand that there are objective rulers that tell objective height. Trying to steelman this school of psychoanalysis, it’s a sort of as-if game, the professional equivalent of Crowley’s demand that the adept swear an oath “to interpret all phenomena as a direct dealing of God with his soul”. It’s an enforced fast from object-level discussion, where you treat everything as significant as an assumption. My first guess was going to be that this is so that you minimize the Type II errors where you miss something that really is significant, but after thinking about it more I wonder if it’s just that this is a bizarre and unnatural mode of thought that can get you places that normal thought can’t, sort of the same way some people have revelations on LSD not because LSD itself is magic but because it’s so different from normal thought processes that it can uncover things that are otherwise hidden. This could also explain the Freudian obsession with dreams – it’s not that they necessarily mean anything, any more than my patient missing his flight meant something, it’s that they’re a good source of noise to start scrying into.

(another Freudian technique is free association, asking the patient to just say whatever first comes to mind. In Mount Misery, Basch’s patient says “porpoises”, but has no idea why – something had to come to mind, and a porpoise was the first thing to pop into his head. This is a lot more like my own experience with free association than the textbook cases of people suddenly coming up with repressed childhood memories or something)

But this method also reminds me of something else. This is Christopher Hitchens:

“I think Hannah Arendt said that one of the great achievements of Stalinism was to replace all discussion involving arguments and evidence with the question of motive. If someone were to say, for example, that there are many people in the Soviet Union who don’t have enough to eat, it might make sense for them to respond, “It’s not our fault, it was the weather, a bad harvest or something.” Instead it’s always, “Why is this person saying this, and why are they saying it in such and such a magazine? It must be that this is part of a plan.”

The avoidance of object-level discussion in favor of meta-level discussion can get really nasty, really quickly. The book gives one example – if you psychoanalyze rape accusations (“what purpose is it serving in this person’s mental ecosystem to have them accuse their psychiatrist of rape right now?”) then you miss someone who is actually getting raped. This can be more insidious when complaints are less dramatic and less binary – I know a lot of psychiatrists who will respond to people saying their medication isn’t working (or is causing side effects), with analyzing their motives for wanting to piss off their psychiatrist or stay unhealthy. And finally, this is absolutely fatal to any kind of complicated social discussion – the thing where instead of debating someone else’s assertion, you bulverize what self-interest or privilege causes them to believe it.

Basch says:

I breathed in the cleansing sorrow of the rain and stared back up at the castle, and I saw clearly how through psychoanalysis you could know every nook and cranny of yourself and have no idea how to be with anyone, the seeming dazzle of the self blinding you to the connections with others…and I knew then I had once been in touch with people, and that it wasn’t inevitable that we are always shouting across an unbridgeable gap, but rather that the gap was in Freud and monstrous fabrications like [Dr. Lowell] herself who followed after, bereft souls floating untethered in pools of self like lilies in sepsis, the gap was in them, not in the essence of humans, nor in the essence of the whole world.

I stared up at the vigilant street lamp, the cone of glittering sleet in the winter night reaching toward me like a beacon, showing me as clearly as if it were the moment’s sun that the real perversion of Freud and analysis was to take the essence of something and reduce it to something eles – the present to the past, love to hate, joy to misery, life to death – and to do it under the guise of understanding and yet, let’s face it, all the while doing it to escape from what Malik kept saying life at heart actually is – being, without description of that being.

If the book is right about psychoanalysis being destructive, I wonder if this is why. Living on the object level is really good. That’s where all the problems are and generally where the solutions are. It’s a natural, healthy place to live.

III.

The last thing that really struck me about the book was its praise for Alcoholics Anonymous.

In the last chapter of the book, Dr. Malik is revealed to be a recovering alcoholic who relapses when he gets diagnosed with cancer. He ends up committed to his own hospital, where he is first pumped full of irrelevant drugs (of course), then subjected to random people telling him he is a bad person because only bad people would drink. All of this is finally contrasted with Alcoholics Anonymous, treated as a beautiful organization full of caring-yet-pragmatic people that tries to genuinely connect with people and give them what they need to stop drinking.

It’s pretty popular to hate on Alcoholics Anonymous these days. And not without reason – I did a sort of literature review about them a while ago, and while they’re no worse than any other treatment options, they aren’t any better either. Their insistence on acknowledging a Higher Power pisses some people off; their insistence on how they are the only way and if you abandon them you’ll just be a drunk for the rest of your life pisses off others. Yet there are some very smart and very compassionate people – apparently including Mount Misery author Samuel Shem – who absolutely love them. In fact, looking at the About The Author page in the back of the book, it looks like after finishing this novel, Shem wrote a hagiographical play about AA founders Bill W and Dr. Bob.

(I checked to see if Shem has ever been an alcoholic himself, sometimes a common feature of people who are really into AA, but it doesn’t look like it.)

There’s a link between AA and Shem’s constant theme throughout his books – people healing through relationships and human connection. But I was especially interested to see this quote, from an article on his AA play:

And what about God? At the time that the two of them met, neither one had much faith in a traditional, religious God. As Smith said, “I was forced to attend church four times a week. I vowed when I was free I would never darken the door of a church again—a vow I’ve kept, religiously, for forty-odd years.” Wilson, too, had more or less given up on God. Both men had pragmatic reasons: they had tried prayer to God, and it didn’t work to keep them sober. The key to their vision about “God” came from a man named Ebby Thatcher, an old friend of Wilson’s who said, “You don’t have to believe in God, you just have to admit that you’re not God. Use what you do believe in, whatever it is.”

Shem seems really into this. I can’t quite justify this from the text, but I get the feeling that he would even take this to the meta-level, something like “complaining about how AA is exclusionary because it requires you to acknowledge a Higher Power is a good sign that you haven’t completed the personal growth task that ‘acknowledging a Higher Power’ corresponds to.”

This was pretty close to what Dr. Basch decides is the essence of a good psychiatrist during his final-chapter epiphany: he realized that he had to get outside himself. The lesson is a little bit Buddhist, but it also ties in nicely to the condemnation of psychoanalysis – while he was being analyzed, he was focused on “his inner machinery”, focused on how everything he experienced was a reflection of his own mind and desires. The attitude Shem holds up as healthy is the exact opposite of that – being able to think about anything except your own problems and your own status, being able to connect to your patients because you’re experiencing them as human beings.

I’ve been trying to reread some of The Last Psychiatrist and better understand what he means by narcissism, something I haven’t been able to get a good feel for before. I think Shem’s idea of getting outside yourself and “admitting that you’re not God” is close to this, a kind of narcissism therapy, where you can work yourself out of narcissism which allows you to connect to your patients and maybe help them in the same way.

(it also sort of reminds me of C.S. Lewis)

It’s well-known in psychopharmacology that different drugs work on different people, for mysterious reasons. Prozac and Paxil are about equally good in general, but some people will hate Prozac and find Paxil a miracle drug, whereas other people will get better on Prozac and find Paxil does nothing. I wonder if there might be something similar for social interventions like Alcoholics Anonymous. Over the whole population, it won’t outperform any other form of rehab – but there will be a few people for whom it works miracles. Those people will go on to praise it to the skies in all kinds of books and plays and so on – not to mention starting the next generation of Alcoholics Anonymous groups – while everyone else watches bemusedly. Actually, now that I mention this it sounds obviously true and I’m not sure why I wasn’t thinking this way already.

IV.

One of the reasons the psychiatrists in Mount Misery are so bad is that they’re narcissistic, but it’s an understandable narcissism. Someone says “My mom just died”, and you say “I’m sorry for your loss” and let them talk about their memories of their mother? Anyone can do that! Why did they get borderline-tortured throughout their twenties and thirties getting a really prestigious psychiatry degree if they were just going to say “I’m sorry your mom died”? Being able to relate it all to wanting to suck your dad’s penis at least gives them some credibility for all their erudite Freud-knowledge and justifies their $200-an-hour fees. “I’m sorry for your loss” isn’t exactly $200-an-hour level insight.

But I don’t know if Shem has a good solution here. It doesn’t seem like he wants to destroy psychiatry as an institution – he is, after all, a Harvard psychiatry professor. But his fictional bigwigs are right. If all you do is be a decent human being and have one-to-one meaningful discussions with patients, then it doesn’t seem like there’s a point in having MDs for that.

I’ve had a lot of patients with this exact complaint – usually it’s about psychologists or therapists instead of psychiatrists. “She kept telling me to go to sessions, that she was going to help me, and all we did was talk about my problems. I could have had a friend do that. So eventually I just quit and I haven’t been back to see a therapist since. Bunch of quacks.” I hear this kind of thing almost every day. It’s a big fear of mine that somebody thinks it about me. Probably one reason I like psychopharmacology so much is that it makes me feel useful – prescribing imipramine correctly isn’t something that just anybody could do; my patients may or may not get better but at least they’re getting their money’s worth.

I don’t know if Shem thinks that well-trained psychiatrists have some kind of special ability to connect with other people. Based on how horrible every psychiatrist in his book is, plus his preference for self-help groups like AA, it doesn’t look like it. But his rejection of both official therapies and medication doesn’t leave him a lot of outs. Also, it seems pretty obvious even to him that a lot of conditions – like melancholic depression and schizophrenia – don’t just need a kind word and a smile, that they are really complex entities that need a lot of effort and probably a good biochemical understanding before you can do much to them.

If all that Shem is saying is that doing the complicated work of psychiatric treatment – therapies, medications, et cetera – has to be combined with actually caring about the patient and treating them like a human being, then fair enough. But the vitriol of his criticisms of therapy and medication make it hard to read that message. If he’s proposing something more radical, then I’m afraid I didn’t entirely get what it was.

# Contra NYT On Economists On Education

[epistemic status: still having a hard time believing I am right about this, but have double-checked. Tell me if I’m missing something. Correction: I previously left the word “Generally” out of the title by accident. I have slightly softened a claim about ‘journalistic malpractice’ based on many people apparently finding the phrasing less misleading than I do.]

From today’s New York Times: Free Market For Education: Economists Generally Don’t Buy It:

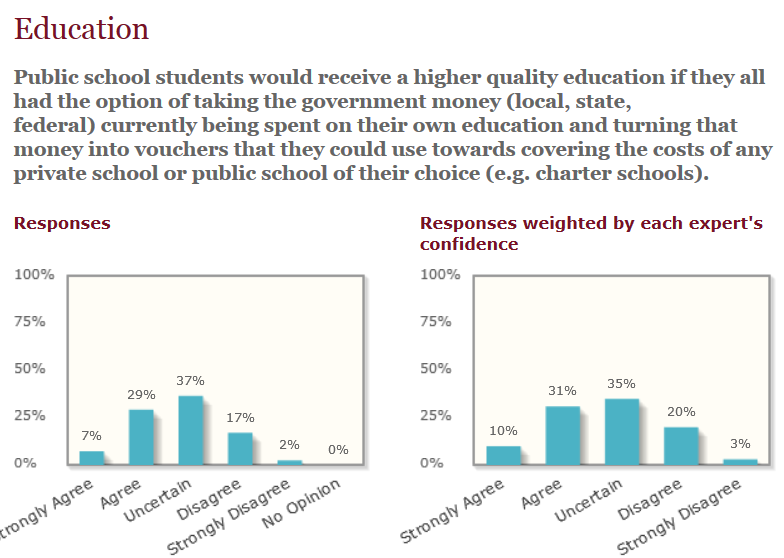
The odds are good that privatizing education will be part of the agenda for President-elect Donald J. Trump’s administration. […] You might think that most economists agree with this overall approach, because economists generally like free markets. For example, over 90 percent of the members of the University of Chicago’s panel of leading economists thought that ride-hailing services like Uber and Lyft made consumers better off by providing competition for the highly regulated taxi industry.

But economists are far less optimistic about what an unfettered market can achieve in education. Only a third of economists on the Chicago panel agreed that students would be better off if they all had access to vouchers to use at any private (or public) school of their choice.

While economists are trained about the value of free markets, they are also trained to spot when markets can’t work alone and government intervention is required.

This is followed by a long discussion of market failures and externalities, with the implication that this is the sort of knowledge that economists are using to come to their anti-voucher views.

But look at the NYT’s source for its claim about economists:



36% of economists agree that vouchers would improve education, compared to 19% who disagree. The rest are unsure or didn’t answer the question. The picture looks about the same when weighted by the economists’ confidence.

A more accurate way to summarize this graph is “About twice as many economists believe a voucher system would improve education as believe that it wouldn’t.”

By leaving it at “only a third of economists support vouchers”, the article implies that there is an economic consensus against the policy. Heck, it more than implies it – its title is “Free Market For Education: Economists Generally Don’t Buy It”. But its own source suggests that, of economists who have an opinion, a large majority are pro-voucher.

(note also that the options are only “agree that vouchers will improve education” and “disagree that vouchers will improve education”, so that it’s unclear from the data if any dissenting economists agree with the Times’ position that vouchers will make things worse. They might just think that things would stay the same.)

I think this is really poor journalistic practice and implies the opinion of the nation’s economists to be the opposite of what it really is. I hope the Times prints a correction.

[see follow-up post here]

# Addendum To “Economists On Education”

A couple of people have challenged my essay yesterday, saying that they didn’t find the article misleading (or that it was only very slightly misleading).

They argue that it pointed out that economists were overwhelmingly in favor of ride-sharing arguments like Uber, but only somewhat in favor of school vouchers. Therefore, it’s fair to say that “economists don’t generally buy school vouchers” in the same sense that they “generally buy” Uber.

This would have been a fine thing to say, but I think the article failed to make this point and instead phrased its argument in a way that made it unlikely to readers to conclude anything other than that economists generally were not in favor of vouchers.

First, I feel like you could write exactly the opposite headline. “Public School: Economists Generally Don’t Buy It”. You would cite the statistic that “only 19% of economists surveyed disagree with the statement that school vouchers would improve education over the existing public system”. Then you could explain some public choice theory about how economists believe government services generally do poorly. This would be exactly as honest or misleading as the existing article. Yet it would produce exactly the opposite impression in readers’ minds: the original article makes you think economists mostly oppose school vouchers, the changed article makes you think they mostly support them. If newspapers are allowed to interpret data so loosely that they can use it to draw two different conclusion precisely opposite each other, what’s the point of having data?

Second, the article uses economists “not buying it” as a segue into a description of why economic theory says school choice could be a bad idea. But it seems like the majority of economists are not convinced by this argument. That is, both the 1/3 of economists who agree and the 1/3 of economists who are uncertain don’t accept the author’s argument that economic theory proves school vouchers can’t work as obviously true (the uncertain ones may be uncertain whether it’s true or not). This gives it a credibility it doesn’t deserve.

Third, really, if this same article was on Breitbart, and it used a survey showing that 40% of economists supported climate change legislation, 40% were uncertain, and 20% opposed it, and it described this as “Economists Generally Don’t Buy Climate Change Solution”, nobody would think anything untoward had happened and they would all agree this was a perfectly fair and unbiased summary of the evidence?

Fourth, you can do this for anything because there’s no clear definition of “uncertain”. How sure does an economist have to be before she “agrees” with a statement rather than being “uncertain” about it. If you are 51% sure school vouchers help, are you pro-voucher or uncertain? What about 60% sure? What about 90% sure? Suppose that all economists believe with 70% probability that vouchers will be good. If your criteria for “certain” is “80% or above”, then as long as you separate them out into support/uncertain/oppose, you can “honestly” declare that “no economists support school vouchers” and convince everyone that it must be an economically absurd plan. Yet I would hope that on any controversial issue more complicated than Uber, most economists are at least a little uncertain about it. We shouldn’t view that as legitimizing us to say that the economic consensus is whatever we want it to be.

Fifth and related, if you look at the economists’ comments, a lot of the people who self-described as “uncertain” thought vouchers would be good on net, but didn’t like the question because they thought it implied that literally 100% of students would be better off. This was such a problem that the IGM redid the study a year later, this time asking whether vouchers would make most students better off. 44% of economists agreed, compared to only 5% who disagreed (again, 34% were uncertain). Weighted by confidence, >50% of the economists agreed that vouchers would improve things, compared to only 6% who thought it wouldn’t improve things – a difference of almost ten times more economists agreeing that vouchers would help compared to disagreeing!

In the face of all of this, the New York Times gives the field’s opinion as “Free Market In Education: Economists Generally Don’t Buy It”.

I’m trying to be more empirical these days, so if you disagree with me, let me make you a bet – and I seriously mean I’ll bet money on this if anyone wants to take me up on it. We find ten random people of ordinary intelligence and economics knowledge and show them this article. Then we ask them a question like “according to this article, what is the economic consensus on vouchers?” We assure them that this isn’t a trick question and they’re just supposed to honestly give the impression they get from the article. If they say something like “they’re generally against vouchers”, I win the bet. If they say something like “probably more economists support vouchers than oppose them, but many are uncertain”, you win the bet. I am willing to alter exact terms if you have a better idea. I am willing to make this bet at 10:1 odds, so if you think there is any chance at all this article is not misleading it should be easy money for you.

[EDIT: I am going to take a version of this bet with Noah Smith. I don’t need to take it with any more people. Offer is now closed.]

[EDIT 2: Noah Smith is apparently no longer willing to bet me on this, although he refuses loudly and at great length to say whether that means he now agrees with me that most people would get a false impression of economists’ position from the article. Is there anyone else who wants to take a bet on these terms or any others?]

# 2016 Predictions: Calibration Results

At the beginning of every year, I make predictions. At the end of every year, I score them. Here are 2014 and 2015.

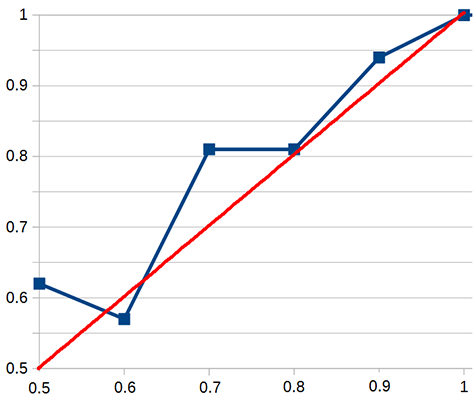
And here are the predictions I made for 2016. Strikethrough’d are false. Intact are true. Italicized are getting thrown out because I can’t decide if they’re true or not.

WORLD EVENTS  
1. US will not get involved in any new major war with death toll of > 100 US soldiers: 60%  
2. North Korea’s government will survive the year without large civil war/revolt: 95%  
3. Greece will not announce it’s leaving the Euro: 95%  
4. No terrorist attack in the USA will kill > 100 people: 90%  
5. …in any First World country: 80%  
6. Assad will remain President of Syria: 60%  
7. Israel will not get in a large-scale war (ie >100 Israeli deaths) with any Arab state: 90%  
8. No major intifada in Israel this year (ie > 250 Israeli deaths, but not in Cast Lead style war): 80%  
9. No interesting progress with Gaza or peace negotiations in general this year: 90%  
10. No Cast Lead style bombing/invasion of Gaza this year: 90%  
11. Situation in Israel looks more worse than better: 70%  
12. Syria’s civil war will not end this year: 70%  
13. ISIS will control less territory than it does right now: 90%  
14. ISIS will not continue to exist as a state entity: 60%  
15. No major civil war in Middle Eastern country not currently experiencing a major civil war: 90%  
16. Libya to remain a mess: 80%  
17. Ukraine will neither break into all-out war or get neatly resolved: 80%  
18. No country currently in Euro or EU announces plan to leave: 90%  
19. No agreement reached on “two-speed EU”: 80%  
20. Hillary Clinton will win the Democratic nomination: 95%  
21. Donald Trump will win the Republican nomination: 60%  
22. Conditional on Trump winning the Republican nomination, he impresses everyone how quickly he pivots towards wider acceptability: 70%  
23. Conditional on Trump winning the Republican nomination, he’ll lose the general election: 80%  
24. Conditional on Trump winning the Republican nomination, he’ll lose the general election worse than either McCain or Romney: 70%  
25. Marco Rubio will not win the Republican nomination: 60%  
26. Bloomberg will not run for President: 80%  
27. Hillary Clinton will win the Presidency: 60%  
28. Republicans will keep the House: 95%  
29. Republicans will keep the Senate: 70%  
30. Bitcoin will end the year higher than $500: 80%  
31. Oil will end the year lower than $40 a barrel: 60%  
32. Dow Jones will not fall > 10% this year: 70%  
33. Shanghai index will not fall > 10% this year: 60%  
34. No major revolt (greater than or equal to Tiananmen Square) against Chinese Communist Party: 95%  
35. No major war in Asia (with >100 Chinese, Japanese, South Korean, and American deaths combined) over tiny stupid islands: 99%  
36. No exchange of fire over tiny stupid islands: 90%  
37. US GDP growth lower than in 2015: 60%  
38. US unemployment to be lower at end of year than beginning: 50%  
39. No announcement of genetically engineered human baby or credible plan for such: 90%  
40. No major change in how the media treats social justice issues from 2015: 70%  
41. European far right makes modest but not spectacular gains: 80%  
42. Mainstream European position at year’s end is taking migrants was bad idea: 60%  
43. Occupation of Oregon ranger station ends: 99%  
44. So-called “Ferguson effect” continues and becomes harder to deny: 70%  
45. SpaceX successfully launches a reused rocket: 50%  
46. Nobody important changes their mind much about the EMDrive based on any information found in 2016: 80%  
47. California’s drought not officially declared over: 50%  
48. No major earthquake (>100 deaths) in US: 99%  
49. No major earthquake (>10000 deaths) in the world: 60%

PERSONAL/COMMUNITY  
1. SSC will remain active: 95%  
2. SSC will get fewer hits than in 2015: 60%  
3. At least one SSC post > 100,000 hits: 50%  
4. UNSONG will get fewer hits than SSC in 2016: 90%  
5. > 10 new permabans from SSC this year: 70%  
5. UNSONG will get > 1,000,000 hits: 50%  
6. UNSONG will not miss any updates: 50%  
7. UNSONG will have higher Google Trends volume than HPMOR at the end of this year: 60%  
8. UNSONG Reddit will not have higher average user activity than HPMOR Reddit at the end of this year: 60%  
9. Shireroth will remain active: 70%  
10. I will be involved in at least one published/accepted-to-publish research paper by the end of 2016: 50%  
11. I won’t stop using Twitter, Tumblr, or Facebook: 95%  
12. > 10,000 Twitter followers by end of this year: 50%  
13. I will not break up with any of my current girlfriends: 70%  
14. I will not get any new girlfriends: 50%  
15. I will attend at least one Solstice next year: 90%  
16. …at least two Solstices: 70%  
17. I will finish a long blog post review of stereotype threat this year: 60%  
18. Conditional on finishing it, it won’t significantly change my position: 90%  
19. I will finish a long FAQ this year: 60%  
20. I will not have a post-residency job all lined up by the end of this year: 80%  
21. I will have finished all the relevant parts of my California medical license application by the end of this year: 70%  
22. I will no longer be living in my current house at the end of this year: 70%  
23. I will still be at my current job: 95%  
24. I will still not have gotten my elective surgery: 80%  
25. I will not have been hospitalized (excluding ER) for any other reason: 95%  
26. I will not have taken any international vacations with my family: 70%  
27. I will not be taking any nootropic daily or near-daily during any 2-month period this year: 90%  
28. I will complete an LW/SSC survey: 80%  
29. I will complete a new nootropics survey: 80%  
30. I will score 95th percentile or above in next year’s PRITE: 50%  
31. I will not be Chief Resident next year: 60%  
32. I will not have any inpatient rotations: 50%  
33. I will continue doing outpatient at the current clinic: 90%  
34. I will not have major car problems: 60%  
35. I won’t publicly and drastically change highest-level political/religious/philosophical positions (eg become a Muslim or Republican): 90%  
36. I will not vote in the 2016 primary: 70%  
37. I will vote in the 2016 general election: 60%  
38. Conditional on me voting and Hillary being on the ballot, I will vote for Hillary: 90%  
39. I will not significantly change my mind about psychodynamic or cognitive-behavioral therapy: 80%  
40. I will not attend the APA meeting this year: 80%  
41. I will not do any illegal drugs (besides gray-area nootropics) this year: 90%  
42. I will not get drunk this year: 80%  
43. Less Wrong will neither have shut down entirely nor undergone any successful renaissance/pivot by the end of this year: 60%  
44. No co-bloggers (with more than 5 posts) on SSC by the end of this year: 80%  
45. I get at least one article published on a major site like Huffington Post or Vox or New Statesman or something: 50%  
46. I still plan to move to California when I’m done with residency: 90%  
47. I don’t manage to make it to my friend’s wedding in Ireland: 60%  
48. I don’t attend any weddings this year: 50%  
49. I decide to buy the car I am currently leasing: 60%  
50. Except for the money I spend buying the car, I make my savings goal before July 2016: 90%

Of 50% predictions, I got 8 right and 5 wrong, for a score of 62%  
Of 60% predictions, I got 12 right and 9 wrong, for a score of 57%  
Of 70% predictions, I got 13 right and 3 wrong, for a score of 81%  
Of 80% predictions, I got 13 right and 3 wrong, for a score of 81%  
Of 90% predictions, I got 16 right and 1 wrong, for a score of 94%  
For 95% predictions, I got 9 right and 0 wrong, for a score of 100%  
For 99% predictions, I got 3 right and 0 wrong, for a score of 100%

This is the graph of my accuracy for this year:



Red is hypothetical perfect calibration, blue is my calibration. I am too lazy and bad at graphs to put in 95% right, but it doesn’t change the picture very much (especially because it’s impossible to get a very accurate 95% with 9 questions).

The 50% number is pretty meaningless, as many people have noted, so my main deviation was some underconfidence at 70%. This was probably meaningless in the context of this year’s numbers alone, but looking back at 2014 and 2015, I see a pretty similar picture. I am probably generally a little bit underconfident in medium probabilities (I have also gotten lazier about making graphs).

Overall I rate this year’s predictions a success. Predictions for 2017 coming soon.